

ОСНОВЫ
АВТОМАТИ-
ЧЕСКОГО
УПРАВЛЕНИЯ

ОСНОВЫ автоматического управления



ОСНОВЫ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ

Под редакцией В. С. ПУГАЧЕВА

Издание третье,
исправленное и дополненное



ИЗДАТЕЛЬСТВО «НАУКА»
ГЛАВНАЯ РЕДАКЦИЯ
ФИЗИКО-МАТЕМАТИЧЕСКОЙ ЛИТЕРАТУРЫ
МОСКВА 1974

Основы автоматического управления. Под редакцией В. С. Пугачева. Издательство «Наука», Главная редакция физико-математической литературы, 1974, 720 стр.

В книге дано изложение основ современной теории автоматического управления. Особенностью книги является новый подход к изложению теории линейных систем. За основу принята общая теория линейных систем, а теория стационарных линейных систем выводится из общей теории как частный случай. В связи с этим за основную характеристику линейной системы принята весовая функция. При этом, в отличие от традиционного подхода, операционное исчисление не применяется в качестве инструмента для развития теории, а играет лишь вспомогательную роль как удобное средство для определения передаточных функций стационарных линейных систем по данным весовым функциям, и наоборот. В качестве основного метода определения весовых функций линейных систем дается метод моделирования сопряженных систем. Подчеркивается, что аппарат теории дифференциальных уравнений принципиально недостаточен для теории автоматических систем вследствие того, что многие применяемые в них элементы не могут быть описаны обыкновенными дифференциальными уравнениями или даже уравнениями в частных производных.

Второй особенностью книги является статистический подход к исследованию процессов управления. Большую часть книги занимают методы изучения поведения автоматических систем при случайных возмущениях. Особое внимание уделяется статистической теории оптимальных систем.

Книга содержит общую теорию линейных систем, включающую методы математического описания непрерывных и дискретных линейных систем, теорию их соединений и структурных преобразований, методы исследования устойчивости линейных систем, теорию их статистического анализа и методы определения их оптимальных характеристик; методы математического описания нелинейных систем, теорию их устойчивости, методы исследования автоколебаний в нелинейных системах, приближенные методы статистического анализа, основанные на обычной и статистической линеаризации, методы определения оптимальных характеристик нелинейных систем по различным статистическим критериям и элементы теории оптимального управления. В частности, излагается решение задачи оптимального обнаружения сигналов в шумах и дается постановка задачи оптимальной селекции сигналов из всей совокупности принимаемых системой сигналов. Даются краткие сведения об основных элементах автоматических систем и их характеристиках и краткое изложение основ теории информации. В заключительной главе книги рассматривается общая проблема управления при отсутствии полной информации и дается краткое изложение основных принципов построения самонастраивающихся систем.

Табл. 8. Илл. 385. Библ. 92 назв.

© Издательство «Наука», 1974.

О 30501 — 026
053(01) — 74 169-74

ОГЛАВЛЕНИЕ

Из предисловия к первому изданию	9
Из предисловия ко второму изданию	12
Предисловие к третьему изданию	13
Глава 1. Основные понятия теории автоматического управления . .	17
§ 1.1. Понятие об управлении. Основные принципы управления	17
§ 1.2. Структура автоматической системы и ее составные элементы	23
§ 1.3. Детерминированные и стохастические системы	29
§ 1.4. Оператор системы	31
§ 1.5. Линейные и нелинейные системы. Принципы суперпозиции	33
§ 1.6. Стационарные и нестационарные системы	37
§ 1.7. Непрерывные и дискретные (импульсные) системы	38
Глава 2. Характеристики линейных систем	41
§ 2.1. Единичная импульсная функция. Разложение произвольной функции на элементарные импульсы	41
§ 2.2. Весовые функции линейных систем	48
§ 2.3. Характеристика реакции линейной системы на показателное возмущение. Частотная характеристика	56
§ 2.4. Стационарные линейные системы. Передаточная функция и частотная характеристика стационарной линейной системы	61
§ 2.5. Определение характеристик стационарных линейных систем, описываемых дифференциальными уравнениями	72
§ 2.6. Элементарные стационарные линейные звенья	79
Глава 3. Линейные элементы автоматических систем	86
§ 3.1. Состав и назначение элементов автоматических систем	86
§ 3.2. Потенциометрические датчики	89
§ 3.3. Индуктивные датчики	92
§ 3.4. Тахометры	96
§ 3.5. Гироскопические измерители угловой скорости	97
§ 3.6. Акселерометры	99
§ 3.7. Магнитоэлектрические чувствительные элементы	101
§ 3.8. Сельсины	104
§ 3.9. Электронные усилители	110
	1*

§ 3.10. Электромашиные усилители	114
§ 3.11. Линейные функциональные преобразователи	118
§ 3.12. Импульсные линейные элементы	124
§ 3.13. Двигатели постоянного тока	126
§ 3.14. Двигатели переменного тока	129
§ 3.15. Длинные линии	131
§ 3.16. Объект управления как элемент автоматической системы	138
Глава 4. Структурные схемы и методы определения характеристик линейных систем	144
§ 4.1. Соединения систем и их элементов. Структурные схемы	144
§ 4.2. Весовые функции соединений	151
§ 4.3. Определение весовых функций методом сопряженных систем	155
§ 4.4. Линейная система, описываемая одним дифференциаль- ным уравнением	164
§ 4.5. Линейная система, описываемая системой дифференци- альных уравнений	176
§ 4.6. Соединения стационарных линейных систем	183
§ 4.7. Структурные преобразования линейных систем	188
§ 4.8. Метод логарифмических частотных характеристик . . .	195
Глава 5. Дискретные линейные системы	201
§ 5.1. Дискретное (импульсное) управление. Способы модуля- ции импульсов	201
§ 5.2. Характеристики дискретных линейных систем	203
§ 5.3. Стационарные дискретные линейные системы	211
§ 5.4. Стационарные дискретные линейные системы, описывае- мые разностными уравнениями	218
Глава 6. Устойчивость и качество линейных систем	226
§ 6.1. Определение устойчивости. Общее условие устойчивости линейной системы	226
§ 6.2. Устойчивость стационарных линейных систем. Критерии Рауса и Гурвица	230
§ 6.3. Устойчивость стационарных линейных систем. Критерий Найквиста. Запасы устойчивости	235
§ 6.4. Устойчивость дискретных линейных систем	247
§ 6.5. Переходные процессы в линейных системах	249
§ 6.6. Синтез линейных систем методом частотных характе- ристик	253
Глава 7. Методы исследования точности линейных систем	259
§ 7.1. Случайные возмущения, действующие в автоматических системах	259
§ 7.2. Общие методы исследования точности линейных систем	261
§ 7.3. Определение установившихся систематических ошибок стационарных линейных систем	278

§ 7.4.	Определение установившейся дисперсии выходной переменной стационарной линейной системы	284
§ 7.5.	Определение установившейся дисперсии выходной переменной стационарной дискретной линейной системы	291
§ 7.6.	Применение теории линейных систем для нахождения интегральных канонических представлений случайных функций	294
§ 7.7.	Определение моментов выходных переменных интегрированием дифференциальных уравнений	306
Глава 8.	Характеристики нелинейных систем	313
§ 8.1.	Основные свойства нелинейных систем	313
§ 8.2.	Типовые нелинейности автоматических систем	314
§ 8.3.	Линеаризация нелинейных характеристик путем разложения в ряд	318
§ 8.4.	Гармоническая линеаризация нелинейных характеристик	319
§ 8.5.	Статистическая линеаризация нелинейных характеристик	325
§ 8.6.	Совместная гармоническая и статистическая линеаризация нелинейных характеристик	333
Глава 9.	Нелинейные элементы автоматических систем	338
§ 9.1.	Электролитические датчики вертикали	338
§ 9.2.	Термосопротивления	341
§ 9.3.	Магнитные усилители	343
§ 9.4.	Нелинейные функциональные преобразователи на диодах	357
§ 9.5.	Вращающиеся трансформаторы	360
§ 9.6.	Гальваномагнитные элементы	362
§ 9.7.	Реле	365
§ 9.8.	Модуляторы и демодуляторы	381
§ 9.9.	Электродвигатели с фрикционными муфтами	388
§ 9.10.	Электродвигатели с редукторами	390
§ 9.11.	Гидравлические двигатели	392
§ 9.12.	Пневматические двигатели	397
§ 9.13.	Характерные нелинейности объектов управления	403
Глава 10.	Устойчивость и автоколебания нелинейных систем	405
§ 10.1.	Общее определение устойчивости	405
§ 10.2.	Исследование устойчивости нелинейных систем методом линеаризации (первый метод Ляпунова)	410
§ 10.3.	Прямой метод исследования устойчивости нелинейных систем (второй метод Ляпунова)	413
§ 10.4.	Фазовые траектории и особые точки нелинейных систем	416
§ 10.5.	Исследование нелинейных систем методом точечных отображений	432

§ 10.6. Приближенное исследование нелинейных систем методом гармонической линеаризации	437
§ 10.7. Влияние случайных возмущений на параметры автоколебаний стационарных нелинейных систем	444
Глава 11. Методы исследования точности нелинейных систем . . .	451
§ 11.1. Приближенное исследование точности нелинейных систем методами линеаризации	451
§ 11.2. Применение метода статистической линеаризации для исследования точности стационарных систем	456
§ 11.3. Применение метода статистической линеаризации для исследования точности систем, описываемых дифференциальными уравнениями	463
§ 11.4. Исследование точности автоматических систем методом статистических испытаний	466
Глава 12. Передача информации по каналам связи	471
§ 12.1. Основные задачи теории информации. Каналы связи	471
§ 12.2. Сигналы, используемые для передачи информации. Измерение количества информации, содержащейся в сигналах	473
§ 12.3. Статистические свойства сигналов	479
§ 12.4. Передача информации по каналу с шумами. Пропускная способность канала	483
Глава 13. Задачи статистической теории оптимальных систем . . .	489
§ 13.1. Постановка задач теории оптимальных систем	489
§ 13.2. Статистические критерии оптимальности автоматических систем	495
§ 13.3. Определение оптимальных параметров системы, имеющей заданную структуру	502
§ 13.4. Общее условие минимума средней квадратической ошибки	509
§ 13.5. Уравнения, определяющие оптимальную линейную систему	515
§ 13.6. Общий анализ уравнений, определяющих оптимальную линейную систему	530
§ 13.7. Уравнения, определяющие оптимальную дискретную линейную систему	537
Глава 14. Методы определения оптимальных линейных систем . . .	539
§ 14.1. Определение оптимальной линейной системы в случае белого шума на входе	539
§ 14.2. Определение оптимальной линейной системы в случае стационарных сигнала и помехи и бесконечного интервала наблюдения	542
§ 14.3. Общая формула для определения оптимальной линейной системы	547

§ 14.4. Определение оптимальной линейной системы в случае стационарной помехи и конечного интервала наблюдения	553
§ 14.5. Синтез оптимальной стационарной линейной системы методом логарифмических частотных характеристик	574
§ 14.6. Определение оптимальной линейной системы методом канонических разложений	583
§ 14.7. Определение оптимальной дискретной линейной системы	587
§ 14.8. Оптимальные системы, описываемые дифференциальными уравнениями	590
Глава 15. Определение оптимальных нелинейных систем	608
§ 15.1. Общий метод определения оптимальной системы	608
§ 15.2. Определение оптимальной системы в случае линейной зависимости входной функции от параметров сигнала	622
§ 15.3. Определение оптимальных систем по критерию минимума средней квадратической ошибки	627
§ 15.4. Оптимальные системы для обнаружения сигналов в шумах	630
§ 15.5. Определение оптимальных систем для селекции сигналов	639
Глава 16. Методы определения оптимального управления	642
§ 16.1. Принцип максимума	642
§ 16.2. Применение принципа максимума для решения типовых задач	649
§ 16.3. Динамическое программирование	655
§ 16.4. Применение динамического программирования к непрерывным системам	660
Глава 17. Самонастраивающиеся системы	665
§ 17.1. Управление при неполной информации	665
§ 17.2. Принципы построения самонастраивающихся систем	669
§ 17.3. Стабилизация динамических характеристик объекта управления с помощью корректирующих цепей	671
§ 17.4. Самонастраивающиеся системы с настройкой параметров по эталону	675
§ 17.5. Самонастраивающиеся системы с экстремальной настройкой параметров	679
§ 17.6. О некоторых перспективах	688
Приложения	691
1. Основные характеристики элементарных стационарных звеньев	692
2. Номограммы для определения логарифмических частотных характеристик системы, замкнутой отрицательной жесткой обратной связью, по логарифмическим частотным характеристикам соответствующей разомкнутой системы, и наоборот	694
3. Номограммы для определения основных характеристик систем, имеющих типовые логарифмические амплитудные характеристики	696

4. Типовые характеристики нелинейных звеньев	700
5. Формулы для статистических и гармонических коэффициентов усиления нелинейных звеньев	702
6. Значения функции $\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_0^u e^{-\frac{t^2}{2}} dt$	708
7. Значения функции $\Phi'(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$	709
Литература	710
Предметный указатель	714

ИЗ ПРЕДИСЛОВИЯ К ПЕРВОМУ ИЗДАНИЮ

Современный этап развития автоматики характеризуется значительным усложнением задач теории автоматического управления в связи с переходом от автоматизации отдельных объектов к комплексной автоматизации производственных процессов, от управления отдельными объектами к одновременному управлению большим числом взаимосвязанных объектов. Вследствие этого традиционное содержание теории автоматического регулирования и управления стало слишком узким. Возникла необходимость включения в теорию автоматического управления ряда новых вопросов, связанных с передачей и преобразованием информации в автоматических системах. Так, например, для решения задач управления многими объектами существенно уметь оценивать количество информации, передаваемой по линиям связи между объектами и от объектов к элементам системы управления, а также пропускную способность каналов связи. А так как передача информации по каналам связи неизбежно сопровождается действием шумов и помех, то возникает необходимость использовать в конструкциях систем управления и для их расчета современные методы обнаружения сигналов в шумах и выделения их из шумов, т. е. методы оптимальной статистической обработки информации. Эти методы необходимы также для автоматического контроля процессов управления и своевременного обнаружения и устранения неисправностей. Наконец, вследствие того, что для автоматизации сложных процессов решающую роль играет точность работы систем управления, одним из важнейших разделов современной теории автоматического управления являются методы исследования точности автоматических систем.

В современных сложных автоматических системах для обработки информации приходится применять вычислительные машины как непрерывного действия, так и цифровые. Поэтому процессы передачи и преобразования информации в современных системах могут протекать как непрерывно, так и дискретно. Отсюда следует, что в современной теории автоматического управления необходимо изучать не только непрерывные, но и дискретные процессы в автоматических системах.

На основании приведенных соображений излагаемые в книге основы теории автоматического управления включают методы

определения динамических характеристик непрерывных и дискретных автоматических систем, исследования их устойчивости и качества, методы структурного анализа и синтеза, методы расчета точности автоматических систем, вопросы передачи информации и методы оптимальной обработки информации.

Теоретической основой методов исследования точности автоматических систем, изучения вопросов передачи информации и методов оптимальной обработки информации являются теория вероятностей и математическая статистика. Кроме того, изучение общих динамических свойств автоматической системы с учетом всего многообразия возможных входных сигналов также невозможно без привлечения статистических методов. Поэтому изложение общей теории автоматического управления ведется в книге в значительной мере на основе статистических методов. Статистический подход к исследованию точности и оценке качества систем управления рассматривается в книге как фундамент для решения основных проблем современной теории автоматического управления и ее дальнейшего развития.

В современных задачах автоматического управления часто приходится иметь дело с существенно нестационарными системами и процессами. Поэтому за основу в книге принята общая теория динамических систем, применяемая как к стационарным, так и к нестационарным системам. Теория стационарных систем выводится из общей теории как частный случай. В связи с этим пришлось отказаться от привычного для специалистов в области автоматики и радиотехники подхода к построению теории стационарных линейных систем, основанного на применении операционного исчисления. Вся теория автоматического управления развивается в книге без применения операционного исчисления. По нашему мнению, изложение от этого только выигрывает, становится более наглядным, теснее связанным с физическими представлениями и менее формальным.

Изложение теории автоматического управления, как прикладной инженерной науки, было бы недостаточно конкретным и читатель не мог бы самостоятельно освоить практическое применение излагаемых в книге методов без практических примеров. Поэтому в книге уделено большое внимание иллюстрации излагаемых методов примерами. Некоторые примеры имеют общий характер и применимы к различным конкретным системам, составленным из различных элементов, основанных на использовании разных физических явлений. Наличие таких примеров позволяет применять излагаемые в книге методы инженерам самых различных специальностей. Кроме таких общих примеров, в книге имеется еще некоторое количество конкретных технических примеров, в которых рассматриваются вполне определенные автоматические системы, состоящие из определенных элементов. Эти примеры

позволят читателю глубже вникнуть в техническую сущность задач автоматизации.

Всякая автоматическая система состоит из ряда элементов. Динамические свойства системы в целом определяются динамическими свойствами входящих в ее состав элементов. Поэтому мы сочли полезным включить в книгу некоторые сведения об основных элементах автоматических систем. Однако при этом мы ограничились только основными принципами устройства элементов, знание которых необходимо для определения их динамических характеристик, имея в виду, что для расчета автоматических систем имеют значение только динамические свойства элементов, а не их конкретная физическая реализация. Поэтому вопросы расчета и проектирования элементов автоматизации в книге не излагаются.

Изложение теории автоматического управления строится в книге следующим образом: сначала излагаются общие способы математического описания динамических систем и их динамических характеристик, затем даются некоторые сведения об элементах автоматических систем и их динамических характеристиках и после этого рассматриваются структуры автоматических систем, методы определения их динамических характеристик по динамическим характеристикам входящих в их состав элементов и более простых систем, и изучаются методы анализа и синтеза автоматических систем.

В настоящее время существует точная и достаточно полная теория линейных систем. Нелинейные системы изучены пока еще значительно меньше. В частности, не существует общей точной теории нелинейных систем. Между тем практически в любой системе всегда имеются нелинейности, более или менее существенно влияющие на ее поведение. Поэтому для практики необходимо уметь исследовать и рассчитывать не только линейные, но и нелинейные системы. Для исследования нелинейных систем приходится, как правило, прибегать к приближенным методам, которые практически все основаны на применении в том или ином виде линейной теории, в частности различных методов линеаризации. Вследствие этого мы сочли целесообразным изложить сначала теорию линейных систем по вышеприведенной схеме, затем по той же схеме методы исследования нелинейных систем и после этого основы теории передачи информации и методы определения оптимальных алгоритмов обработки информации и нахождения оптимальных систем для обнаружения и выделения сигналов в присутствии шумов и помех.

При разработке содержания книги мы тщательно следили за тем, чтобы в нее попало только собственно то, что может быть отнесено непосредственно к современной теории автоматического управления, и избегали включения общих математических вопро-

сов, которые можно рассматривать как математический аппарат, лежащий в основе соответствующих разделов теории автоматического управления. В частности, мы нарушили традицию и не включили в теорию передачи информации общие сведения о вероятностном содержании понятий энтропии, информации, количества информации и их основных свойствах, считая, что эти вопросы относятся к общей теории вероятностей и применимы не только к задачам теории управления и связи. Вследствие этого удалось добиться весьма компактного изложения элементов теории передачи информации по каналам связи в рамках одной небольшой главы. Без такого отбора материала книги было бы невозможно изложить основные вопросы современной теории автоматического управления в рамках одной сравнительно небольшой книги.

В окончательной подготовке книги к печати деятельную помощь авторам оказал Л. А. Чульский, которому авторы приносят свою глубокую благодарность.

Авторы заранее благодарят всех читателей, которые сочтут возможным прислать в издательство свои замечания относительно содержания книги и замеченных недостатков и опечаток.

Октябрь 1962 г.

Авторы

ИЗ ПРЕДИСЛОВИЯ КО ВТОРОМУ ИЗДАНИЮ

Во втором издании книга дополнена несколькими новыми параграфами, разъяснениями и примерами. Кроме того, в нее внесены необходимые исправления замеченных неточностей и опечаток. Глава о самонастраивающихся системах (глава 12 первого издания книги) перенесена в конец книги, что дает возможность читателю глубже понять принципы применения оптимальных статистических методов обработки информации в самонастраивающихся системах.

Октябрь 1966 г.

Авторы

ПРЕДИСЛОВИЕ К ТРЕТЬЕМУ ИЗДАНИЮ

В третьем издании книга дополнена несколькими новыми параграфами и одной главой. В частности, в главу 14 добавлен § 14.8, содержащий изложение теории оптимальных систем, описываемых обыкновенными дифференциальными уравнениями (фильтров Калмана). Новая глава 16 содержит изложение элементов теории оптимального управления детерминированными системами. В связи с этим глава 16 второго издания в третьем издании стала главой 17. Кроме этих дополнений, в книгу внесены исправления замеченных неточностей и опечаток, а также сделаны некоторые редакционные улучшения и необходимые добавления, связанные с современным развитием теории управления.

Книга содержит 17 глав и приложения. В главе 1 даются общие представления о процессе управления, о составе автоматических систем и об их математическом описании. В главе 2 рассматриваются общие характеристики линейных систем. В главе 3 даются некоторые сведения об элементах автоматических систем, которые с достаточной для практики точностью можно считать линейными. Глава 4 посвящена методам определения характеристик линейных систем. В частности, здесь излагаются методы определения характеристик линейной системы по характеристикам входящих в ее состав более простых систем (в частности, элементов), методы нахождения и моделирования сопряженных систем, рассматриваются структурные схемы автоматических систем и правила их преобразования. В заключение главы 4 рассматриваются стационарные линейные системы и их исследование методом частотных характеристик. В главе 5 из общей теории линейных систем, развитой в главе 2, выводятся основы теории дискретных линейных систем. Глава 6 посвящена основным методам исследования устойчивости линейных систем. В главе 7 излагаются методы исследования точности линейных систем, имеющих заданные характеристики, работающих под действием как управляющих сигналов, так и случайных возмущений. В главе 8 рассматриваются основные особенности нелинейных систем, вводится понятие элементарного безынерционного нелинейного звена и рассматриваются характеристики нелинейных звеньев. В частности, излагаются методы гармонической и статистической линеаризации, а также метод совместной статистической и гармонической линеаризации. В главе 9

излагаются некоторые сведения об основных нелинейных элементах автоматических систем. Глава 10 содержит методы исследования устойчивости нелинейных систем, методы исследования автономных систем второго порядка с помощью фазовой плоскости, в частности метод точечных отображений, и приближенные методы исследования автоколебаний нелинейных систем, основанные на гармонической и совместной гармонической и статистической линеаризации. В главе 11 излагаются приближенные методы исследования точности нелинейных систем, основанные на обычной и статистической линеаризации нелинейностей. Кроме того, рассматриваются некоторые вопросы приближенного исследования точности автоматических систем методом статистических испытаний. В главе 12 излагаются элементы теории передачи информации по каналам связи. Здесь вводится понятие канала связи, рассматриваются статистические свойства каналов связи и передаваемых по ним сигналов и выводится основная формула, определяющая пропускную способность канала связи. В главе 13 дается постановка общей задачи оптимальной обработки информации, рассматриваются различные статистические критерии оптимальности и выводится основное уравнение, определяющее оптимальную систему по критерию минимума средней квадратической ошибки. Из этого уравнения выводятся уравнения, определяющие оптимальную линейную систему. В последнем параграфе как частный случай выводятся уравнения, определяющие оптимальную дискретную линейную систему по критерию минимума средней квадратической ошибки. В главе 14 излагаются методы решения основного интегрального уравнения, к которому приводится задача определения оптимальных линейных систем по критерию минимума средней квадратической ошибки. В последнем параграфе главы 14 изложены методы определения оптимальных систем, описываемых обыкновенными дифференциальными уравнениями (теория фильтров Калмана). Глава 15 посвящена общему методу нахождения оптимальных алгоритмов обработки информации по любым статистическим критериям и применению этого метода для решения задач выделения полезных сигналов из шумов, обнаружения полезных сигналов, искаженных помехами, и селекции полезных сигналов, сопровождаемых ложными сигналами. В главе 16 изложены элементы теории оптимального управления. Здесь даны основные сведения о принципе максимума Л. С. Понтрягина и его учеников и его применении к задачам оптимального управления, а также о методе динамического программирования Р. Беллмана в применении к задачам оптимизации управления. Последняя глава 17 содержит основные сведения о так называемых самонастраивающихся системах. В приложениях даны таблицы и номограммы, необходимые для практического применения излагаемых в книге методов.

Все главы книги взаимно связаны и вместе составляют единое целое. Однако некоторые главы все же имеют известную автономность, что дает возможность пользоваться книгой не только в целом, но и по частям.

Так, например, читатель, желающий изучить общие основы теории управления и не интересующийся конкретными элементами автоматических систем, может без ущерба пропустить главы 3 и 9.

Читатель, желающий ознакомиться с основными понятиями теории управления и методами исследования линейных систем при определенных (не случайных) входных возмущениях, может ограничиться чтением глав 1—6 и § 7.3.

Читатель, желающий ознакомиться еще и с приближенными методами исследования нелинейных систем при определенных входных возмущениях, может ограничиться чтением глав 1—6, § 7.3 и глав 8, 10, за исключением §§ 8.5, 8.6 и 10.7.

Прочитав главы 1, 2, 4, 7, 8, 11, 13 и 14, за исключением §§ 7.5, 13.7 и 14.7, читатель может изучить основы статистической теории непрерывных линейных систем. При этом для понимания изложенного в главе 7 достаточно предварительно прочесть главы 1 и 2. Для понимания изложенного в главе 11 достаточно прочесть главы 1, 2, 4, 7 и 8. Глава 13 требует знания содержания глав 1, 2 и 4. Для чтения главы 14 достаточно знать содержание глав 1, 2, 4 и §§ 7.4, 7.6, 13.5, 13.6 и 13.7.

Чтобы получить основные представления об автоматических системах и принципах устройства самонастраивающихся систем, достаточно прочесть главы 1, 2, § 4.1 и главу 17. Главы 12 и 16 можно читать отдельно и независимо от остальных глав книги. Наконец, для понимания изложенного в главе 15 достаточно прочесть §§ 13.1, 13.2 и 14.6.

Книга рассчитана в основном на инженеров и студентов, знакомых с основами высшей математики, физики, электротехники и радиотехники в объеме программ высших технических учебных заведений, а также с элементами теории вероятностей в объеме книги В. С. Пугачева [54].

Для читателей, желающих перед чтением книги вспомнить основы электротехники и радиотехники, можно рекомендовать книги А. С. Касаткина, М. А. Перекалина и П. С. Сергеева [30] и В. Ф. Власова [13].

Ясно, что чрезвычайно обширный материал, затронутый в книге, невозможно изложить с исчерпывающей полнотой в малом объеме. Поэтому в книге даны лишь основы автоматического управления. Для детального изучения различных разделов теории автоматического управления отсылаем читателя к специальной литературе. В частности, для более детального изучения методов исследования непрерывных стационарных линейных систем

можно рекомендовать книгу Г. В. Бode [8], книгу Д. К. Ньютона, Л. А. Гулда и Д. Ф. Кайзера [46] или коллективный труд «Основы автоматического регулирования» [47], том 1, «Теория», главы I—XXI. Для изучения стационарных дискретных линейных систем можно рекомендовать книгу Я. З. Цыпкина [72]. Приближенные методы исследования нелинейных систем читатель может изучить по книге Е. П. Попова и И. П. Пальтова [49] или по книге «Основы автоматического регулирования» [47], том 1, «Теория», главы XXVI—XXXIX. Специальные методы исследования релейных систем читатель может изучить по книге Я. З. Цыпкина [73]. Для более глубокого изучения статистических методов теории автоматического управления читателю рекомендуется обратиться к книгам В. С. Пугачева [53] и И. Е. Казакова и Б. Г. Доступова [28]. Для изучения теории оптимального управления можно рекомендовать книгу Л. С. Понтрягина, В. Г. Болтянского, Р. В. Гамкрелидзе и Е. Ф. Мищенко [48], а для изучения динамического программирования — книги Р. Беллмана [77] и Р. Беллмана и Р. Калабы [78]. Наконец, элементы автоматических систем можно изучить по книге Б. С. Сотскова [67]. Обстоятельное изложение основ современной теории автоматического управления читатель может найти в трехтомном труде А. А. Воропова [82].

Книга представляет собой коллективный труд. Ее авторы:

В. С. Пугачев — главы 1, 2, 4 (исключая § 4.5 и пример 4.1.4), 6, 7 (исключая § 7.7), 11 (исключая §§ 11.3 и 11.4), 12, 13 (исключая пример 13.3.2), 14 (исключая § 14.5), 15 и §§ 3.16, 9.13, 17.1 и 17.6;

И. Е. Казаков — главы 8, 10, 16 и §§ 17.2, 17.3 и 17.4;

Д. Н. Гладков — глава 3 (исключая §§ 3.7, 3.15 и 3.16), § 9.3 и часть § 9.8;

Л. Г. Евланов — глава 9 (исключая §§ 9.3, 9.13, часть § 9.7 и часть § 9.8), §§ 3.7 и 17.5;

С. В. Мальчиков — §§ 4.5, 7.7, 11.3, 11.4 и 14.5;

А. Ф. Мишаков — § 3.15, часть § 9.7, примеры 4.1.4 и 13.3.2;

В. Д. Седов — глава 5.

В подготовке первого издания книги принимал также участие покойный В. И. Соколов.

Общий замысел книги и разработка ее плана принадлежат В. С. Пугачеву, написавшему большую часть книги на основе читанных им на протяжении ряда лет курсов лекций.

Март 1973 г.

Авторы

ОСНОВНЫЕ ПОНЯТИЯ ТЕОРИИ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ

§ 1.1. Понятие об управлении. Основные принципы управления

В окружающем нас мире повсюду протекают различные процессы управления. В частности, человек в своей практической деятельности постоянно использует различные физические явления, законы природы для достижения определенных целей. При этом он в известной мере управляет силами природы, заставляя их работать в необходимом направлении. Особенно ярко это выражается с появлением техники. Человек управляет различными средствами транспорта, машинами в различных отраслях промышленности, химическими реакциями и разными другими процессами. Процессы управления протекают также в других областях деятельности человека. Так, полководец во время войны управляет боевыми действиями состоящих в его распоряжении войск и боевой техники. Руководитель предприятия, учреждения или отрасли промышленности управляет деятельностью подчиненной ему организации. Процессы управления непрерывно протекают и в живой природе. В результате этих процессов в живом организме поддерживается практически постоянным ритм сердца и дыхания, температура тела и вообще обеспечивается функционирование всех органов и жизнедеятельность всего организма.

Управление — это такая организация того или иного процесса, которая обеспечивает достижение определенных целей.

Бурное развитие автоматики, электроники и вычислительной техники в послевоенный период привело к внедрению автоматики буквально во все области деятельности человека. Автоматика и автоматизация становятся главным направлением развития всей техники. Совершается переход от автоматизации отдельных простейших производственных операций к комплексной автоматизации средств производства и производственных процессов. Появляются различные автоматические производственные линии и даже целые автоматические заводы. Решающую роль в развитии автоматики сыграло появление математических машин. Если в прошлом машины предназначались для освобождения человека от тяжелого физического труда, то сейчас наступила эпоха создания машин принципиально нового типа, выполняющих некото-

рые функции человеческого мозга, в частности осуществляющих управление различными сложными процессами. Роль человека при этом сводится к организации работы автоматических систем и средств вычислительной техники, такому их проектированию, соединению и использованию, которое обеспечивает получение необходимых результатов с наименьшими затратами. Для решения этих задач необходима наука о процессах управления и их общих закономерностях — *теория управления*.

В связи с повышением требований к эффективности управления всеми отраслями народного хозяйства в последнее десятилетие возникла необходимость научной организации всех процессов управления, в том числе и осуществляемых с участием людей. Это выдвинуло перед теорией управления новые задачи. Если раньше было достаточно уметь исследовать процессы автоматического управления техническими устройствами, то теперь необходимо исследовать и рассчитывать также процессы управления сложными системами, содержащими коллективы людей, с помощью систем управления, в которых главную роль играют люди — соответствующие руководители. Поэтому теория управления перестала быть только теорией автоматического управления. В ней появилось новое научное направление — теория автоматизированных систем управления (АСУ). Теория автоматического управления стала лишь частью общей теории управления.

Чтобы понять основные принципы управления, рассмотрим, например, процесс управления автомобилем. Сидя за рулем, водитель видит перед собой дорогу и находящиеся на ней предметы, наблюдает, куда идет машина, и на основании этого принимает решение, надо ли изменить направление движения машины и если надо, то куда и насколько следует повернуть руль. Анализируя этот процесс, мы видим в нем следующие основные элементы. Во-первых, получение информации о направлении, в котором должна идти машина, т. е. информации о задаче управления. Эту информацию водитель получает при помощи зрения. Во-вторых, получение информации о результатах управления. Водителю недостаточно видеть перед собой дорогу, он должен видеть, куда идет машина. Эту информацию он также получает с помощью зрения. В-третьих, анализ полученной информации и принятие на основе этого анализа решения о необходимых управляющих действиях. Наконец, в-четвертых, исполнение принятого решения. Эти четыре элемента составляют основу всякого управления. Если исключить хотя бы один из них, то управление автомобилем станет невозможным. Например, если завязать водителю глаза или вывести из строя его мозг или руки, то управление станет невозможным.

Рассмотренный пример показывает, что в общем случае процесс управления состоит из следующих четырех элементов: получение информации о задачах управления, получение информации

о результатах управления, т. е. о поведении объекта управления, анализ полученной информации и выработка решения и исполнение решения, т. е. осуществление управляющих действий. В соответствии с этим для организации процесса управления необходимо иметь источники информации о задачах управления и результатах управления, устройства для анализа получаемой информации и выработки решения и исполнительные устройства, осуществляющие управление объектом.

Из приведенного анализа процесса управления следует, что в организации процесса управления большую, а во многих случаях и решающую роль играет получение информации о результатах управления. При наличии этой информации решение об управляющих действиях существенно зависит от результатов управления. Иными словами, первичный элемент — управляющее действие — зависит от вторичного элемента — вызываемого управляющим действием поведения объекта управления. Получается как бы замкнутый круг: причина, вызывающая изменение состояния объекта управления, ставится в зависимость от того, какой результат она вызывает. Такая связь причины и следствия называется *обратной связью*. Принцип управления с использованием информации о результатах управления называется *принципом обратной связи*.

Принцип обратной связи лежит в основе подавляющего большинства процессов управления. В частности, принцип обратной связи лежит в основе почти всей деятельности человека. Так, например, если человек протягивает руку, чтобы взять какой-либо предмет, то он смотрит на предмет и подсознательно непрерывно управляет движением руки, исправляя ошибки направления движения руки к предмету. Чтобы убедиться в этом, достаточно исключить получение информации о результатах управления, завязав человеку глаза, и наблюдать движение его руки.

Однако в некоторых отдельных случаях принцип обратной связи использовать не удается из-за практической невозможности получить информацию о результатах управления. В таких случаях управление сложными процессами становится невозможным. Например, управление автомобилем при отсутствии информации о том, куда он движется, совершенно невозможно. Однако в некоторых случаях закон изменения состояния объекта управления во времени заранее известен и практически не зависит от результатов управления, а зависит только от управляющих действий. Так, например, при автоматическом управлении токарным станком, изготавливающим детали одного определенного образца, чрезвычайно трудно организовать измерение деталей в процессе обточки, т. е. получение информации о результатах управления. Вследствие этого практически невозможно управлять движением резца с учетом результатов этого управления, т. е. с использова-

нием принципа обратной связи. Однако закон изменения положения резца во времени в процессе обработки различных деталей весьма мало зависит от разброса параметров отдельных заготовок и является практически одним и тем же для всех деталей. Поэтому в данном случае можно задать положение резца как определенную функцию времени и осуществить автоматическое перемещение его по этому закону.

Управление, обеспечивающее заданный закон изменения состояния объекта управления во времени, независимо от результатов управления, называется *программным управлением* или управлением *по разомкнутому циклу*. Закон изменения состояния объекта управления во времени при этом называется *программой управления*. В отличие от управления по разомкнутому циклу, управление с использованием принципа обратной связи называется часто управлением *по замкнутому циклу*. Область применения программного управления в чистом виде ограничена в основном промышленными автоматами (автоматическими станками). Значительно более широкое распространение нашло такое программное управление, при котором сама программа изменения состояния объекта управления осуществляется с помощью обратной связи. В таких случаях программа управления служит информацией о задачах управления, а состояние объекта управления определяется с помощью измерителей, которые дают информацию о результатах управления и тем самым осуществляют обратную связь. В подобных случаях управление по существу не является программным, а представляет собой управление с обратной связью. Так, например, при управлении ракетой сразу после пуска часто задается определенная программа изменения угла наклона ее продольной оси к горизонту с течением времени. Истинный угол наклона оси ракеты к горизонту измеряется и сравнивается с программным, и на основе этого сравнения вырабатывается сигнал управления. Таким образом, принцип обратной связи является основным принципом управления.

Совокупность всех устройств, обеспечивающих управление каким-либо объектом, называется *системой управления*. Изложенное показывает, что всякая система управления должна содержать источники информации о задачах управления и результатах управления, устройства, анализирующие информацию и вырабатывающие решение об управляющих действиях, и исполнительные устройства. Некоторыми из этих элементов или всеми сразу может, в частности, служить человек. Например, водитель автомобиля выполняет сразу функции всех перечисленных элементов. В таких случаях система управления включает человека и не является автоматической.

Если функции всех элементов системы управления выполняются различными устройствами без непосредственного уча-

ствия человека, то система управления называется *автоматической*. Примерами автоматических систем управления могут служить автопилот, управляющий полетом самолета, система самонаведения ракеты, следящая система, обеспечивающая автоматическое сопровождение самолета лучом радиолокатора, устройства управления процессом вычислений в электронной цифровой машине и т. д.

Система управления, в которой решения об управляющих действиях принимаются людьми, а автоматические устройства используются только для сбора, обработки и представления информации о задачах и результатах управления и для сравнительного анализа различных возможных вариантов решений, называется *автоматизированной*. Примером автоматизированной системы управления может служить система управления отраслью промышленности, содержащая устройства и вычислительные машины, автоматически осуществляющие сбор и обработку необходимой информации, анализирующие различные варианты решений и определяющие их сравнительные характеристики.

При изучении процесса управления приходится рассматривать совместную работу объекта управления и системы управления и их взаимодействие. Только при этом условии можно изучить процессы управления, основанные на принципе обратной связи. Объект управления вместе с системой управления представляет собой сложную динамическую систему, которую мы будем называть в дальнейшем для краткости просто *системой*. Совокупность технического объекта и автоматической системы управления им будем называть *автоматической системой*.

В некоторых случаях задачей управления является обеспечение постоянства некоторой физической величины. Такой частный вид управления обычно называется *регулированием*. Автоматическая система управления, обеспечивающая регулирование значения какой-либо физической величины, называется *регулятором*. Автоматическая система, состоящая из регулируемого объекта и регулятора, называется *системой автоматического регулирования*. Процессы автоматического регулирования широко распространены в живой природе. Примерами могут служить регулирование ритма сердца, регулирование ритма дыхания, регулирование температуры тела и т. д.

Простейшие автоматические системы управления, а именно системы автоматического регулирования, появились очень давно, по существу вместе с первыми машинами. Первые же попытки создания машин показали, что никакая машина не может нормально функционировать без соответствующих регуляторов. Так например, для нормального функционирования паровой машины оказалось необходимым поддерживать постоянное давление пара в паровом котле и постоянную скорость вращения махового колеса.

В качестве примера простейшего регулятора на рис. 1.1.1 приведен схематический чертеж центробежного регулятора скорости вращения махового колеса паровой машины, который был изобретен в 1784 г. Джемсом Уаттом. Этот регулятор состоит из двух грузов 2, подвешенных на шарнирах вдоль оси вертикального вала 1. Рычаги грузов связаны с муфтой 3, которая может перемещаться вдоль вала 1. Муфта 3 в свою очередь связана рычагом с заслонкой (дресселем) 4, положение которой определяет

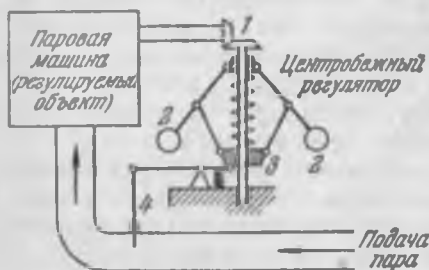


Рис. 1.1.1.

сечение отверстия, через которое пар поступает из котла в цилиндр паровой машины. При определенной угловой скорости вала 1 грузы 2, муфта 3 и заслонка 4 находятся в определенном положении. Конструкция рычагов и заслонки выбирается таким образом, чтобы при нужной угловой скорости сечение отверстия для прохода пара было таким, чтобы машина вращала вал именно с этой

скоростью. Тогда, если скорость вращения вала 1 превысит заданную величину, то центробежная сила грузов увеличится, вследствие чего грузы поднимутся и переместят муфту 3 и заслонку 4. При этом подача пара в машину уменьшится, что вызовет уменьшение угловой скорости вала 1, т. е. приближение ее к заданному значению. Если угловая скорость вала станет меньше заданного значения, то грузы опустятся и переместят заслонку вверх, вследствие чего подача пара в машину увеличится. В этом простейшем регуляторе грузы являются и измерителем отклонения угловой скорости от заданного значения, и исполнительными устройствами, осуществляющими управление заслонкой, а устройства для анализа информации полностью отсутствуют.

Простейшие регуляторы обеспечивают функционирование машин, управляемых человеком, и на протяжении многих десятилетий были достаточны. И лишь с развитием физики и особенно электроники, с развертыванием производства электронных приборов, с появлением быстродействующих математических машин, могущих в течение коротких интервалов времени выполнять сложные математические и логические операции, стало возможным создание сложных систем управления, осуществляющих управление сложными процессами и одновременное управление большим числом объектов.

Соответственно развитию техники и ее возможностей развивалась и теория автоматического управления. Вначале, когда суще-

ствовали лишь простейшие регуляторы, теория автоматического регулирования была по существу лишь собранием примеров применения методов теории дифференциальных уравнений. Затем стали создаваться специальные методы исследования систем автоматического регулирования, поведение которых описывается дифференциальными уравнениями, и теория автоматического регулирования стала специальной отраслью прикладной теории дифференциальных уравнений. В послевоенный период теория автоматического управления стала бурно развиваться. Новые технические средства, новые задачи управления потребовали новых научных методов, привлечения различных существующих математических методов и разработки новых. В результате сейчас теория автоматического управления развилась в самостоятельную прикладную техническую науку, располагающую своими специфическими эффективными методами исследования автоматических систем, их расчета и проектирования.

Мы ограничимся в этой книге изложением основ теории автоматического управления, изучим общие методы расчета и исследования автоматических систем и их составных частей, а также общие принципы решения задач управления.

§ 1.2. Структура автоматической системы и ее составные элементы

Изложенное в предыдущем параграфе показывает, что автоматическая система, как правило, состоит из *объекта управления* и *системы управления*. Система управления имеет в своем составе *источники информации* о задачах управления и результатах управления, *устройства для анализа* вводимой в систему управления информации и выработки решения об управляющих действиях и *исполнительные устройства*, осуществляющие управление.

Источниками информации служат различные измерители, определяющие различные физические величины, которые характеризуют задачи управления и поведение объекта управления. Так, например, в системе самонаведения ракеты основным источником информации о задачах управления и результатах управления является локатор, определяющий положение цели относительно ракеты. В системе телеуправления источником информации о задачах управления служит локатор, определяющий координаты цели, а источником информации о результатах управления служит локатор, определяющий координаты телеуправляемой ракеты. В простейшей системе управления полетом самолета — автопилоте — источником информации о задачах управления является программное устройство, определяющее заданное положение осей самолета в пространстве, а источником информации о результатах

управления служит гироскопический измеритель углов отклонения осей самолета от заданных направлений.

Анализ информации в простейших системах управления состоит в сравнении требуемых и фактических значений величин, определяющих состояние объекта. Разность между требуемым и фактическим значениями каждой величины, принятой за характеристику состояния объекта, называется *параметром управления*, или *рассогласованием*, или *сигналом ошибки*. Однако вырабатывать управляющее действие только в зависимости от текущего значения параметра управления без учета закона его изменения не всегда целесообразно. Так, например, если рулевой ведет лодку в заданном направлении таким образом, что при отклонении ее оси от этого направления влево держит руль направления повернутым вправо до тех пор, пока отклонение не станет равным нулю, и лишь тогда ставит руль в нейтральное положение, то лодка будет по инерции продолжать поворот направо и ее ось окажется отклоненной от заданного направления вправо, вследствие чего рулевой будет вынужден повернуть руль влево. В результате такого управления ось лодки будет совершать непрерывные колебания около заданного направления, а рулевой будет вынужден все время поворачивать руль то вправо, то влево. Очевидно, что для того, чтобы не допустить этих колебаний, рулевой должен ставить руль в нейтральное положение несколько раньше, до того, как отклонение оси лодки от заданного направления станет равным нулю. Иными словами, рулевой должен отклонять руль не только в зависимости от текущего значения отклонения оси лодки от заданного направления, но и с некоторым предвидением закона его изменения. Для автоматического осуществления такого управления необходимо определять некоторые параметры закона изменения параметров управления во времени, в простейшем случае — их производные по времени, которые характеризуют тенденцию их изменения. Для этого в систему управления необходимо ввести устройства, осуществляющие дифференцирование параметров управления или более общее их преобразование, т. е. выработку решения об управляющих действиях с учетом закона изменения параметров управления. Подобные устройства обычно называются *функциональными преобразователями*. Совокупность функциональных устройств, предназначенных для простейших преобразований параметров управления, например для их приближенного дифференцирования, в простых системах управления отдельными объектами называется *корректирующей цепью*. В сложных системах управления параметры управления подвергаются сложным преобразованиям, и совокупность устройств, осуществляющих эти преобразования, уже не сводится к простой корректирующей цепи, а представляет собой сложный комплекс преобразующих, а иногда и вычислительных устройств. Это

особенно характерно для систем управления, в которых для анализа информации и выработки решений используются вычислительные машины.

Решение об управляющих действиях формируется в автоматической системе управления в виде сигналов управления, приводящих в действие исполнительные устройства. Исполнительные устройства обычно представляют собой двигатели различных типов или источники энергии.

На основании изложенного можно построить общую схему автоматической системы, изображенную на рис. 1.2.1. Из этой

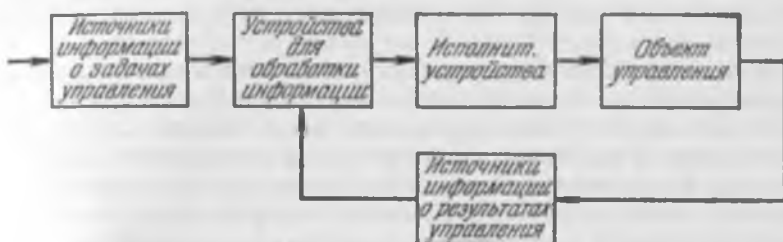


Рис. 1.2.1.

схемы видно, что автоматическая система состоит из целого ряда элементов, имеющих различное назначение и выполняющих различные функции. Типичными элементами автоматической системы управления являются: 1) измерители различных физических величин, называемые часто *чувствительными элементами* и *датчиками*, вводящие в систему управления информацию о задачах управления и результатах управления; 2) функциональные преобразователи или вычислительные устройства, осуществляющие определение параметров управления и заданное их преобразование или более сложный анализ информации и вырабатывающие сигналы управления; 3) исполнительные устройства, осуществляющие управление.

С усложнением задач автоматического управления становятся все более сложными и системы управления. Но, чем сложнее система, чем большее число элементов она содержит, тем более вероятны ее отказы из-за выхода из строя отдельных элементов. Отыскание неисправных элементов в сложных системах представляет собой весьма сложную задачу и требует много времени. И чем сложнее система, тем труднее найти в ней причину отказа и выявить отказавший элемент. Поэтому для нормального функционирования сложной системы управления, для нормальной ее эксплуатации необходимо автоматизировать контроль состояния всех ее частей. Эту задачу наиболее целесообразно решать путем ввода в саму конструкцию системы, наряду с перечисленными основ-

ными элементами, соответствующих приборов и устройств для автоматического контроля состояния системы и режима работы всех ее элементов, а также для сигнализации о неисправностях и указания неисправных элементов. Такие приборы и устройства контроля должны органически входить в конструкцию системы и составлять ее неотъемлемую часть. Поэтому такие входящие в конструкцию системы устройства контроля обычно называются *элементами встроенного контроля*. Встроенный контроль улучшает эксплуатационные качества автоматических систем, сокращает время регламентных работ по их обслуживанию и повышает эффективность их применения. Поэтому задача проектирования автоматических систем с элементами встроенного контроля является важнейшей задачей современной техники.

Состояние объекта управления, особенно сложного, обычно характеризуется большим числом величин. Однако обычно лишь небольшое число этих переменных существенно для процесса управления. Так, например, при управлении полетом самолета с помощью простейшего автопилота существенными переменными являются лишь углы, определяющие положение его осей в пространстве, а все остальные переменные — координаты центра массы самолета, векторы его скорости и ускорения — не входят в выражения параметров управления и не оказывают непосредственного влияния на работу системы управления. Это обстоятельство, характерное для задач управления, определяет коренное отличие теории управления, например, от теоретической механики и многих других областей науки, в которых необходимо рассматривать все переменные, полностью определяющие состояние изучаемого объекта в каждый данный момент времени. Из очень большого числа переменных, определяющих состояние объекта управления, лишь немногие существенны для организации процесса управления. А именно важны лишь те переменные, которые измеряются в системе управления и используются для определения параметров управления.

Интересующие нас величины, характеризующие процесс управления, называются *выходными переменными*, или *выходными функциями*, или *выходными сигналами* системы. Точки системы, в которых выходные сигналы могут наблюдаться в виде определенных физических величин, называются *выходами* системы.

На любой объект управления и на любой элемент системы управления всегда действуют различные внешние возмущения. Точки системы, в которых приложены внешние возмущения, называются *входами* системы, а сами внешние возмущения называются *входными переменными*, или *входными функциями*, или *входными сигналами* системы. В частности, входными переменными автоматической системы являются сигналы измерителей, дающих информацию о задачах управления.

Сигналы измерителей, вводящих в систему управления информацию о задачах управления, которые действовали бы в случае абсолютно точных измерений, представляют собой *полезные входные сигналы* автоматической системы. Действительные сигналы этих измерителей всегда содержат ошибки измерений и помехи и поэтому представляют собой полезные входные сигналы с наложенными на них помехами, т. е. как бы смесь полезных сигналов с помехами. Полезные сигналы играют решающую роль в организации процессов управления. Все прочие входные сигналы, не связанные с источниками информации о задачах и результатах управления, так же как и ошибки источников информации, представляют собой *помехи* и не играют положительной роли в процессе управления.

Следует заметить, что одни и те же сигналы в автоматической системе могут быть внешними возмущениями, т. е. входными сигналами одних элементов системы и одновременно выходными сигналами других элементов. Так, например, величина, определяющая состояние исполнительного устройства, является входным сигналом для объекта управления и в то же время выходным сигналом системы управления. Для всей же автоматической системы в целом этот сигнал является внутренним и не может быть отнесен ни к входным, ни к выходным переменным. Поэтому понятие входов и выходов является условным и относится всегда только к определенной системе. Если рассматривается объект управления, то управляющий сигнал системы управления будет входной переменной, а величины, определяющие состояние объекта управления, будут выходными переменными. Если рассматривается система управления, то для нее входными переменными будут величины, определяющие состояние объекта управления, и величины, определяющие задачи управления, а выходной переменной будет сигнал управления. Если же рассматривается вся автоматическая система в целом, то входными переменными будут величины, определяющие задачи управления, а выходными — величины, определяющие состояние объекта управления.

Системы с одним входом и одним выходом обычно называются *одномерными*. Системы с несколькими входами и выходами называются *многомерными*. В частном случае многомерная система может иметь несколько входов и один выход или один вход и несколько выходов.

Встречаются и такие системы, у которых действующие входные возмущения распределены непрерывно вдоль некоторой линии, на поверхности или в пространстве. Такие системы можно рассматривать как системы с непрерывным множеством входов. Точно так же встречаются системы (или части систем), для описания которых необходимо определить некоторую величину во всех точках некоторой области. Такие системы можно рассматривать как

системы с непрерывным множеством выходов, распределенных в соответствующей области. Системы обоих типов, так же как и системы, у которых и входы и выходы распределены непрерывно в некоторых областях, называются *распределенными* системами.

Следует заметить, что полезные управляющие сигналы вводятся в распределенную систему почти всегда только в конечном числе определенных точек и выходные сигналы также снимаются в конечном числе точек. Поэтому распределенную систему практически всегда можно рассматривать как систему с конечным числом входов и выходов. Если при этом, кроме конечного числа полезных сигналов, на систему действуют непрерывно распределенные возмущения, то их можно рассматривать как внутренние шумы системы.

В дальнейшем, говоря о системе, мы будем подразумевать как произвольную автоматическую систему, так и любую часть или

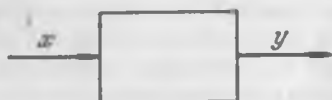


Рис. 1.2.2.



Рис. 1.2.3.

элемент автоматической или автоматизированной системы. При этом, интересуясь только функционированием системы и не интересуясь ее устройством, мы будем изображать систему схематически в виде прямоугольника со стрелками, указывающими входы и выходы (рис. 1.2.2 и 1.2.3). Очевидно, что, рассматривая многомерную систему, можно одной буквой обозначить совокупность всех ее входных переменных и соответственно одной стрелкой показать все ее входы на схеме. Точно так же можно обозначить одной буквой совокупность всех выходных переменных системы и соответственно на схеме показать одной стрелкой все выходы системы.

Общие принципы исследования и проектирования автоматических систем применимы к любым автоматическим системам, как одномерным, так и многомерным. Однако изложение этих принципов проще для одномерных систем. Поэтому в дальнейшем мы будем, как правило, рассматривать одномерные системы, имея в виду, что вся излагаемая теория применима и к многомерным системам. И лишь в некоторых местах будем показывать, как рассматриваемые понятия распространяются на многомерные системы.

Мы видим, какую большую роль для современной автоматики играет вычислительная техника. Сложные системы управления, особенно системы, предназначенные для управления большим

количеством различных объектов или для управления сложными и разнообразно протекающими процессами, совершенно немислимы без быстродействующих цифровых машин. Примером такой сложной системы, как мы уже отмечали, может служить система управления боевыми действиями войск или система управления сложным производственным процессом. Поэтому все развитие автоматике в настоящее время тесно связано с развитием вычислительной техники. Вычислительная техника нужна для автоматики как в качестве одного из главнейших элементов сложных автоматических систем, так и для выполнения научных исследований по разработке принципов создания автоматических систем и изучению протекающих в них процессов.

§ 1.3. Детерминированные и стохастические системы

В современной теории управления приходится встречаться с двумя видами систем. Одни обладают полной определенностью поведения, т. е. ведут себя всегда одинаково в одинаковых условиях. Для других характерна некоторая неопределенность поведения; работая много раз в совершенно одинаковых условиях, они ведут себя в разных случаях различно. Поэтому необходимо различать системы по степени определенности их поведения, которая может быть охарактеризована разбросом выходных сигналов при одном и том же входном сигнале.

Система, которая отвечает на один и тот же входной сигнал всегда одним и тем же вполне определенным выходным сигналом, называется *детерминированной*. Система, отвечающая на один и тот же входной сигнал в разных случаях различными выходными сигналами, называется *недетерминированной*.

Примером детерминированной системы может служить идеальный дифференциатор — система, осуществляющая дифференцирование входного сигнала. Выходным сигналом этой системы всегда служит производная входного сигнала. Примером недетерминированной системы может служить человек, следящий осью оптического прибора за движущимся объектом, скажем самолетом. Входным сигналом в данном случае служит закон движения прямой, соединяющей глаз человека с движущимся объектом, а выходным — закон движения оси прибора. Легко понять, что, повторяя слежение при одном и том же законе движения объекта много раз, человек никогда не осуществит двух совершенно одинаковых законов движения оси прибора. Другим примером недетерминированной системы может служить завод. Продукция завода при одном и том же дневном плановом задании в разные дни оказывается различной.

Во многих случаях разброс выходных сигналов недетерминированной системы при одном и том же входном сигнале подчи-

и только тогда, когда линейной комбинации любых входных возмущений соответствует та же линейная комбинация соответствующих выходных функций. Это свойство линейных систем, выраженное формулой (1.5.1), обычно называется *принципом суперпозиции*. Поэтому линейные системы можно определить как такие системы, для которых справедлив принцип суперпозиции.

Для того чтобы система была линейной, необходимо и достаточно выполнение следующих двух условий:

1) сумме любых двух входных возмущений соответствует сумма соответствующих двух выходных переменных;

2) при любом усилении входного возмущения без изменения его формы выходная переменная претерпевает точно такое же усиление, также не изменяя своей формы.

Необходимость этих условий очевидна. Так как формула (1.5.1) справедлива для любого n и любых чисел $c_1, \dots, c_n$, то, полагая $n = 2$, $c_1 = c_2 = 1$, получаем

$$A \{x_1(t) + x_2(t)\} = Ax_1(t) + Ax_2(t). \quad (1.5.2)$$

Полагая $n = 1$, получим при произвольных c и $x(t)$

$$A \{cx(t)\} = cAx(t). \quad (1.5.3)$$

Для доказательства достаточности условий (1.5.2) и (1.5.3) заметим, что из этих условий вытекают формулы

$$\begin{aligned} A \{c_1x_1(t) + c_2x_2(t)\} &= A \{c_1x_1(t)\} + A \{c_2x_2(t)\} = \\ &= c_1Ax_1(t) + c_2Ax_2(t), \end{aligned} \quad (1.5.4)$$

$$\begin{aligned} A \left\{ \sum_{v=1}^n c_v x_v(t) \right\} &= A \left\{ \sum_{v=1}^{n-1} c_v x_v(t) + c_n x_n(t) \right\} = \\ &= A \left\{ \sum_{v=1}^{n-1} c_v x_v(t) \right\} + A \{c_n x_n(t)\} = A \left\{ \sum_{v=1}^{n-1} c_v x_v(t) \right\} + c_n Ax_n(t), \end{aligned} \quad (1.5.5)$$

справедливые для любых чисел $n, c_1, \dots, c_n$ и для любых функций $x_1(t), \dots, x_n(t)$. Формула (1.5.4) показывает, что из условий (1.5.2) и (1.5.3) следует справедливость принципа суперпозиции для случая двух слагаемых. Формула (1.5.5) показывает, что принцип суперпозиции выполняется для n слагаемых, если он выполняется для $n - 1$ слагаемых. Из этой формулы по индукции следует справедливость принципа суперпозиции при любом числе n слагаемых, поскольку он справедлив для случая двух слагаемых. Таким образом, принцип суперпозиции является следствием условий (1.5.2) и (1.5.3), что и доказывает достаточность этих условий.

Подчеркнем, что для линейности системы необходимо, чтобы принцип суперпозиции соблюдался при любом числе слагаемых, при любом выборе постоянных c_v и функций $x_v(t)$.

Примерами линейных операторов могут служить оператор дифференцирования

$$y(t) = Dx(t) = \frac{d}{dt} x(t), \quad (1.5.6)$$

линейный интегральный оператор

$$y(t) = \int_{t_0}^t g(t, \tau) x(\tau) d\tau \quad (1.5.7)$$

и более общий линейный интегро-дифференциальный оператор

$$y(t) = \sum_{p=0}^N \int_{t_0}^t g_p(t, \tau) x^{(p)}(\tau) d\tau. \quad (1.5.8)$$

К линейному интегральному оператору или к линейному интегро-дифференциальному оператору приводится оператор решения произвольного обыкновенного линейного дифференциального уравнения

$$\begin{aligned} a_n(t) y^{(n)}(t) + a_{n-1}(t) y^{(n-1)}(t) + \dots + a_1(t) y'(t) + \\ + a_0(t) y(t) = b_n(t) x^{(n)}(t) + b_{n-1}(t) x^{(n-1)}(t) + \dots + \\ + b_1(t) x'(t) + b_0(t) x(t). \end{aligned} \quad (1.5.9)$$

Нелинейным называется любой оператор, для которого принцип суперпозиции не имеет места или справедлив только при некоторых вполне определенных функциях $x_1(t), \dots, x_n(t)$ и числах $c_1, \dots, c_n$.

Система с нелинейным оператором называется *нелинейной*.

В качестве примеров нелинейных операторов можно привести нелинейный интегральный оператор

$$y(t) = \int_{t_0}^t \varphi(x(\tau), \tau, t) d\tau, \quad (1.5.10)$$

где $\varphi(x, \tau, t)$ — данная функция, нелинейная относительно переменной x , а оператор решения нелинейного дифференциального уравнения

$$y''(t) + k \sin y(t) = x(t). \quad (1.5.11)$$

Принцип суперпозиции значительно облегчает исследование линейных систем по сравнению с нелинейными. Благодаря принципу суперпозиции теория линейных дифференциальных уравнений разработана в самом общем виде для уравнений любого порядка, в то время как теория нелинейных дифференциальных уравнений по существу отсутствует, и мы можем решать в анали-

тической форме только нелинейные дифференциальные уравнения частных видов невысокого порядка. Вот почему для решения всех математических вопросов, возникающих в приложениях, обращаются в первую очередь к линейным методам. При этом даже нелинейные системы стараются приближенно рассматривать как линейные. В результате появились различные методы *линеаризации* нелинейных систем, т. е. приближенной замены нелинейных систем практически равноценными линейными.

Из справедливости принципа суперпозиции для линейных систем при любом числе слагаемых и любом выборе функций $x_\nu(t)$ и чисел c_ν следует, что он применим не только к суммам, но и к интегралам. Другими словами, если входное возмущение системы представляет собой сумму бесконечно большого числа бесконечно малых элементарных возмущений, то выходная переменная линейной системы представляет собой сумму соответствующих бесконечно малых реакций на эти элементарные возмущения. Математически это выражается формулой

$$A_t \left\{ \int_{\lambda_1}^{\lambda_2} c(\lambda) x(t, \lambda) d\lambda \right\} = \int_{\lambda_1}^{\lambda_2} c(\lambda) A_t x(t, \lambda) d\lambda, \quad (1.5.12)$$

где индекс t у оператора A показывает, что этот оператор действует над функцией аргумента t , а λ при этом рассматривается как фиксированный параметр. Эта формула выражает *принцип суперпозиции в интегральной форме*. Для доказательства достаточно представить интеграл в виде предела последовательности сумм. Для каждой суммы принцип суперпозиции справедлив. Таким образом, для любого члена этой последовательности справедлива формула (1.5.1). Следовательно, при переходе к пределу получится формула (1.5.12), если интеграл в правой части существует.

Принцип суперпозиции дает возможность выразить реакцию линейной системы на любое возмущение через ее реакцию на определенный вид элементарных возмущений. Для этого достаточно разложить произвольное возмущение $x(t)$ на элементарные возмущения выбранного типа. Тогда, зная реакцию линейной системы на элементарные возмущения этого типа, мы можем при помощи принципа суперпозиции определить ее реакцию на произвольное возмущение $x(t)$. Таким образом, для определения реакции линейной системы на произвольное возмущение достаточно знать ее реакцию на выбранный стандартный тип элементарных возмущений. Иными словами, любая линейная система полностью характеризуется ее реакцией на какой-нибудь стандартный тип возмущений. В зависимости от выбора стандартного типа возмущений мы получим разные характеристики линейной системы. Каждая такая характеристика будет исчерпывающей, так как знания ее

достаточно для нахождения реакции линейной системы на любое возмущение.

Уравнения, описывающие поведение линейной системы, всегда линейны. И наоборот, если все уравнения, описывающие поведение системы, линейны, то данная система линейна. Если среди уравнений, описывающих поведение объекта управления и процессы в элементах системы управления, имеется хотя бы одно нелинейное, то система нелинейна.

§ 1.6. Стационарные и нестационарные системы

Система называется *стационарной*, если ее реакция на любое данное возмущение зависит только от интервала времени между данным моментом и моментом начала действия возмущения. Пусть $x(t)$ — произвольная функция, равная нулю при $t < t_0$. Тогда, согласно данному определению, реакция $y(t)$ стационарной системы на возмущение $x(t)$ зависит только от интервала времени $t - t_0$, $y(t) = f(t - t_0)$. Если то же самое возмущение будет действовать на стационарную систему начиная с момента $t_1 = t_0 + a$, то оно будет описываться функцией $x(t - a)$, а реакция системы будет представлять собой функцию $f(t - t_1) = f(t - t_0 - a) = y(t - a)$. Таким образом, стационарную систему можно определить как такую систему, у которой при любом сдвиге во времени входного возмущения без изменения его формы выходная переменная претерпевает такой же сдвиг во времени без изменения своей формы (рис. 1.6.1, а и б). Нестационарные системы характерны тем, что при сдвиге входного возмущения во времени без изменения формы их выходные переменные не только сдвигаются во времени, но и изменяют форму (1.6.1, в).

Стационарные системы могут быть как линейными, так и нелинейными. Нестационарные системы также могут быть как линейными, так и нелинейными. Линейные системы могут быть как стационарными, так и нестационарными.

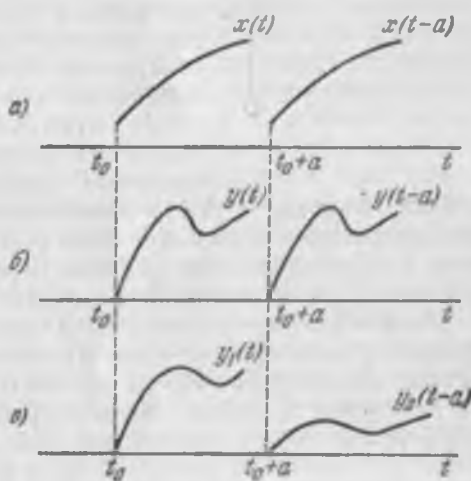


Рис. 1.6.1.

Примером стационарной системы может служить обычный неизменяемый маятник (математический или физический). Независимо от того, в какой момент на маятник, находящийся в покое, начнет действовать возмущение (сила) заданной формы, например синусоидальное возмущение заданной амплитуды и частоты, маятник будет совершать одни и те же колебания.

Примером нестационарной системы может служить математический маятник, длина которого изменяется в зависимости от времени по заданному закону. Такой маятник можно представить как материальную точку, подвешенную на нерастяжимой невесомой нити, перекинутой через ось и наматываемой на барабан часовым механизмом (рис. 1.6.2). Длина такого маятника в момент начала действия на него возмущения будет зависеть от момента начала действия возмущения. Следовательно, и характер колебаний (в частности, частота) такого маятника под действием возмущения заданной формы будет зависеть от момента начала действия возмущения.

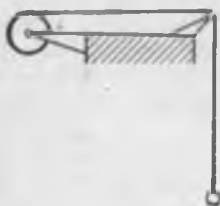


Рис. 1.6.2.

Из теоретической механики известно, что колебания маятника точно описываются нелинейным дифференциальным уравнением второго порядка (1.5.11). Следовательно, математический маятник постоянной длины является стационарной нелинейной системой, а математический маятник переменной длины является нестационарной нелинейной системой.

Если ограничиться изучением только малых колебаний математического маятника, т. е. таких его колебаний, при которых линейные перемещения его конца малы по сравнению с его длиной, то колебания маятника можно приближенно описать линейным дифференциальным уравнением второго порядка

$$y''(t) + ky(t) = x(t). \quad (1.6.1)$$

Следовательно, при малых колебаниях математический маятник постоянной длины можно приближенно рассматривать как стационарную линейную систему, а маятник переменной длины — как нестационарную линейную систему.

§ 1.7. Непрерывные и дискретные (импульсные) системы

Входные сигналы в автоматических системах могут действовать непрерывно в течение всего времени работы системы или только в определенные моменты времени (точнее, в течение коротких интервалов времени), разделенные промежутками времени, в течение которых они не действуют на систему. Системы первого типа называются *непрерывными*, а системы второго типа — *дискретными*.

или *импульсными*. Примерами непрерывных систем могут служить регулятор Уатта, схематически изображенный на рис. 1.1.1, простейший автопилот, предназначенный для управления положением осей самолета в пространстве, или система управления пушечной установкой на самолете. Примерами дискретных систем являются импульсная система телеуправления полетом ракеты, передающая сигналы управления на ракету дискретно, в течение коротких промежутков времени, а также цифровая вычислительная машина. В последнем случае не только входные сигналы вводятся в машину дискретно, но и выходные переменные (результаты решения задачи) изменяются дискретно в определенные моменты времени. Ракета с дискретной системой управления может служить примером дискретной системы с дискретным вводом входных сигналов и с непрерывно изменяющимися выходными переменными.

Дискретное управление приходится неизбежно применять в сложных системах, содержащих вычислительные машины со сложной программой математических и логических операций. В таких системах высокую точность вычислений, связанных с обработкой поступающей в систему информации и выработкой решения об управляющих сигналах, можно получить только с помощью цифровых машин. Поэтому система неизбежно получается дискретной. В частности, без применения цифровых машин и дискретного управления невозможно обойтись в таких системах, как системы управления большим количеством объектов.

Дискретное управление иногда применяется также с целью повысить помехозащищенность автоматических систем, что имеет особенно важное значение для управления военной техникой. Так как в любой автоматической или автоматизированной системе управления объектами военной техники целью управления является поражение объектов противника, то информация о задачах управления в таких системах представляет собой информацию о состоянии и действиях противника. Используя для получения этой информации различные физические явления (например, электромагнитные волны в радиолокации или инфракрасное излучение военных объектов противника), мы всегда даем противнику возможность использовать те же самые физические явления для передачи ложных сигналов и помех, нарушающих работу наших систем управления. Дискретное управление дает возможность добывать и вводить в систему управления информацию о противнике лишь в течение коротких промежутков времени, разделенных относительно длинными интервалами времени, в течение которых система управления остается «закрытой» для внешних сигналов. Этим во многих случаях можно несколько уменьшить возможность для противника мешать работе системы управления, вводить в нее помехи.

Наконец, некоторые источники информации, как, например, радиолокаторы с импульсным излучением, по существу являются дискретными источниками входных сигналов систем управления. Системы с дискретными источниками информации неизбежно оказываются дискретными. Однако если импульсы дискретного источника информации следуют друг за другом через короткие промежутки времени, т. е. с большой частотой, как это имеет место в импульсных радиолокаторах, то такие источники информации с достаточной для практики точностью можно считать непрерывными.

ХАРАКТЕРИСТИКИ ЛИНЕЙНЫХ СИСТЕМ

§ 2.1. Единичная импульсная функция. Разложение произвольной функции на элементарные импульсы

В общей теории линейных систем удобно пользоваться в качестве стандартных возмущений, на которые можно раскладывать любые возмущения, единичными мгновенными импульсами. Чтобы сформировать понятие единичного импульса и дать его математическое описание, обратимся к известным понятиям теоретической механики. Как известно, импульсом постоянной силы называется произведение величины этой силы на время ее действия. Согласно второму закону Ньютона изменение количества движения точки равно импульсу действующей на нее силы. Следовательно, данной величине действующего на точку постоянной массы импульса соответствует вполне определенное конечное изменение скорости точки. Если мы будем неограниченно уменьшать время действия силы и пропорционально увеличивать величину силы так, чтобы импульс оставался неизменным, то изменение скорости точки под действием этой силы будет оставаться постоянным, а время, в течение которого происходит это изменение скорости, будет неограниченно уменьшаться. В пределе, при нулевой длительности импульса, скорость точки будет мгновенно изменяться на величину импульса, деленную на массу точки. Ускорение точки при этом будет представлять собой такую функцию времени, которая имеет бесконечно большое значение в течение определенного бесконечно малого промежутка времени, равна нулю вне этого промежутка времени, а интеграл от которой конечен (равен мгновенному изменению скорости тела). Такими свойствами обладает импульсная δ -функция, впервые введенная в науку знаменитым английским физиком Дираком.

Импульсной δ -функцией называется функция, равная нулю всюду, кроме начала координат, принимающая бесконечное значение в начале координат, и притом так, что интеграл от нее по любому интервалу, содержащему начало координат, равен единице:

$$\delta(t) = 0 \quad \text{при } t \neq 0, \quad \delta(0) = \infty, \quad (2.1.1)$$

$$\int_{-\varepsilon}^{\varepsilon} \delta(t) dt = 1 \quad \text{при любом } \varepsilon > 0. \quad (2.1.2)$$

Функцию, обладающую такими свойствами, можно получить, например, как предел положительного прямоугольного импульса, имеющего единичную площадь, когда длительность этого импульса стремится к нулю (рис. 2.1.1). Еще удобнее определить δ -функцию как предел при $h \rightarrow \infty$ функции

$$\delta_h(t) = \frac{h}{\sqrt{\pi}} e^{-h^2 t^2}. \quad (2.1.3)$$

График этой функции представлен на рис. 2.1.2. Очевидно, что при любом $t \neq 0$ функция $\delta_h(t)$ стремится к нулю при $h \rightarrow \infty$ *).

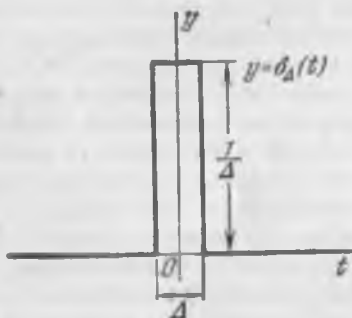


Рис. 2.1.1.

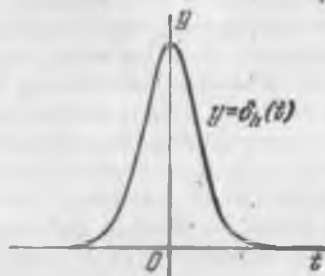


Рис. 2.1.2.

При $t = 0$ эта функция неограниченно возрастает при $h \rightarrow \infty$. Наконец, при любом $\varepsilon > 0$

$$\int_{-\varepsilon}^{\varepsilon} \delta_h(t) dt = \frac{h}{\sqrt{\pi}} \int_{-\varepsilon}^{\varepsilon} e^{-h^2 t^2} dt = \frac{1}{\sqrt{\pi}} \int_{-h\varepsilon}^{h\varepsilon} e^{-u^2} du. \quad (2.1.4)$$

При $h \rightarrow \infty$ это выражение стремится к единице (см. [63], т. II, § 8 или [54]).

Практически удобно еще потребовать, чтобы δ -функция была четной. В этом случае, кроме (2.1.2), имеют место равенства

$$\int_{-\varepsilon}^0 \delta(t) dt = \int_0^{\varepsilon} \delta(t) dt = \frac{1}{2} \quad \text{при любом } \varepsilon > 0. \quad (2.1.5)$$

*) Действительно, раскрывая неопределенность по правилу Лопиталья, получаем

$$\lim_{h \rightarrow \infty} h e^{-h^2 t^2} = \lim_{h \rightarrow \infty} \frac{h}{e^{h^2 t^2}} = \lim_{h \rightarrow \infty} \frac{1}{2h t^2 e^{h^2 t^2}} = 0.$$

Докажем, что для любой непрерывной функции $x(t)$ и для любого $\varepsilon > 0$ справедлива формула

$$\int_{t-\varepsilon}^{t+\varepsilon} x(\tau) \delta(t-\tau) d\tau = \int_{t-\varepsilon}^{t+\varepsilon} x(\tau) \delta(\tau-t) d\tau = x(t). \quad (2.1.6)$$

Для доказательства заметим, что подынтегральная функция в этой формуле на основании (2.1.1) равна нулю всюду, кроме одной точки $\tau = t$, причем функция $x(\tau)$ равна в этой точке $x(t)$. Поэтому интегралы в (2.1.6) не изменятся, если заменить в них функцию $x(\tau)$ ее значением $x(t)$ в фиксированной точке $\tau = t$. Тогда можно будет вынести величину $x(t)$ за знак интеграла, а интеграл будет равен единице вследствие (2.1.2):

$$\begin{aligned} \int_{t-\varepsilon}^{t+\varepsilon} x(\tau) \delta(t-\tau) d\tau &= \int_{t-\varepsilon}^{t+\varepsilon} x(t) \delta(t-\tau) d\tau = \\ &= x(t) \int_{t-\varepsilon}^{t+\varepsilon} \delta(t-\tau) d\tau = x(t) \int_{-\varepsilon}^{\varepsilon} \delta(\sigma) d\sigma = x(t). \end{aligned}$$

Таким образом, интеграл от произведения δ -функции на любую непрерывную функцию равен значению этой функции при том значении переменной интегрирования, при котором аргумент δ -функции обращается в нуль.

Так как подынтегральная функция в формуле (2.1.6) равна нулю при всех значениях τ , кроме $\tau = t$, то пределы интегрирования в (2.1.6) можно произвольно расширить. В результате для любого интервала (a, b) , содержащего точку t , $a < t < b$, получим

$$\int_a^b x(\tau) \delta(t-\tau) d\tau = \int_a^b x(\tau) \delta(\tau-t) d\tau = x(t). \quad (2.1.7)$$

Заметим, что формулы (2.1.6) и (2.1.7) справедливы и в случае, когда точка t является точкой разрыва первого рода функции $x(\tau)$, если ее значение в точке разрыва определить как среднее арифметическое ее значений справа и слева в этой точке:

$$x(t) = \frac{x(t-0) + x(t+0)}{2} = \lim_{\varepsilon \rightarrow 0} \frac{x(t-\varepsilon) + x(t+\varepsilon)}{2}.$$

Действительно, рассуждая так же, как и раньше, получим в силу (2.1.5)

$$\begin{aligned} \int_{t-\varepsilon}^{t+\varepsilon} x(\tau) \delta(t-\tau) d\tau &= \int_{t-\varepsilon}^t x(\tau) \delta(t-\tau) d\tau + \int_t^{t+\varepsilon} x(\tau) \delta(t-\tau) d\tau = \\ &= x(t-0) \int_{-\varepsilon}^0 \delta(\sigma) d\sigma + x(t+0) \int_0^{\varepsilon} \delta(\sigma) d\sigma = \frac{1}{2} x(t-0) + \frac{1}{2} x(t+0), \end{aligned}$$

что и доказывает высказанное утверждение. Таким образом, формула (2.1.7) может быть переписана в более общем виде:

$$\int_a^b x(\tau) \delta(t-\tau) d\tau = \int_a^b x(\tau) \delta(\tau-t) d\tau = \frac{x(t-0) + x(t+0)}{2} \quad (2.1.8)$$

при $a < t < b$.

В практических приложениях удобно бывает пользоваться также производными импульсной δ -функции. Их можно определить, например, как пределы при $h \rightarrow \infty$ соответствующих производных функции $\delta_h(t)$, которая выражается формулой (2.1.3). Пользуясь производными δ -функции, мы можем формально дифференцировать равенства (2.1.7), являющиеся тождествами относительно t при изменении t в интервале (a, b) . В результате, предполагая, что функция $x(t)$ имеет непрерывные производные до n -го порядка, получим

$$\int_a^b x(\tau) \delta'(t-\tau) d\tau = - \int_a^b x(\tau) \delta'(\tau-t) d\tau = x'(t) \quad (2.1.9)$$

и вообще

$$\int_a^b x(\tau) \delta^{(k)}(t-\tau) d\tau = (-1)^k \int_a^b x(\tau) \delta^{(k)}(\tau-t) d\tau = x^{(k)}(t) \quad (2.1.10)$$

$$(k=1, 2, \dots, n).$$

Эти равенства имеют место при любом значении t , заключенном в интервале интегрирования $a < t < b$.

Формулы (2.1.7) и (2.1.10) можно вывести совершенно строго путем предельного перехода при $h \rightarrow \infty$ в соответствующих формулах, содержащих функцию $\delta_h(t)$, определяемую формулой (2.1.3), вместо δ -функции.

Формулы (2.1.7) и (2.1.10) дают возможность компактно записывать в форме интегралов сложные выражения, содержащие интегралы и суммы. Пусть, например, мы имеем выражение

$$I = \int_a^b g(t) x(t) dt + \sum_{p=0}^N \sum_{k=1}^n c_{pk} x^{(p)}(t_k), \quad (2.1.11)$$

где $t_1, \dots, t_n$ — произвольные значения t в интервале интегрирования (a, b) . Вводя функцию

$$g_1(t) = g(t) + \sum_{p=0}^N \sum_{k=1}^n c_{pk} \delta^{(p)}(t_k - t) \quad (2.1.12)$$

и пользуясь формулами (2.1.7) и (2.1.10), можно написать формулу (2.1.11) в виде

$$I = \int_a^b g_1(t) x(t) dt. \quad (2.1.13)$$

Заметим еще, что δ -функцию можно рассматривать как производную единичной ступенчатой функции $1(t)$, определяемой равенствами

$$1(t) = \begin{cases} 0 & \text{при } t < 0, \\ 1 & \text{при } t > 0. \end{cases} \quad (2.1.14)$$

Действительно, на основании (2.1.1) и (2.1.2)

$$\int_{-\infty}^t \delta(\tau) d\tau = \begin{cases} 0 & \text{при } t < 0, \\ 1 & \text{при } t > 0, \end{cases}$$

или

$$\int_{-\infty}^t \delta(\tau) d\tau = 1(t). \quad (2.1.15)$$

Дифференцируя эту формулу, получим

$$\delta(t) = 1'(t). \quad (2.1.16)$$

Таким образом, используя δ -функцию, мы получаем возможность дифференцировать разрывные функции.

Формула (2.1.7) показывает, что любая функция в любом интервале (a, b) может быть разложена на элементарные импульсы. Действительно, формула (2.1.7) представляет функцию $x(t)$ в виде суммы бесконечно большого числа бесконечно малых слагаемых вида $x(\tau) \delta(t - \tau) d\tau$. Каждое такое слагаемое представляет собой бесконечно малый мгновенный импульс величины $x(\tau) d\tau$, действующий в момент времени $t = \tau$, так как $\delta(t - \tau)$ по определению является единичным мгновенным импульсом, действующим в момент $t = \tau$.

Если функция $x(t)$ задана при всех значениях t и мы хотим представить ее разложением на элементарные импульсы (2.1.7) на всей числовой оси, то мы должны положить в (2.1.7) $a = -\infty$, $b = \infty$. Тогда получим, ограничиваясь только первым интегралом,

$$x(t) = \int_{-\infty}^{\infty} x(\tau) \delta(t - \tau) d\tau \quad (-\infty < t < \infty). \quad (2.1.17)$$

Ввиду большой важности этой формулы для всей теории автоматического управления, мы дадим еще один ее вывод, более

наглядный и связанный с физическими представлениями. Для этого введем единичный прямоугольный импульс конечной длительности Δ (рис. 2.1.1):

$$\delta_{\Delta}(t) = \begin{cases} \frac{1}{\Delta} & \text{при } |t| < \frac{\Delta}{2}, \\ 0 & \text{при } |t| > \frac{\Delta}{2}. \end{cases} \quad (2.1.18)$$

Разобьем ось времени на равные интервалы длительности Δ и обозначим середину k -го интервала через τ_k ($k = 0, \pm 1, \pm 2, \dots$).

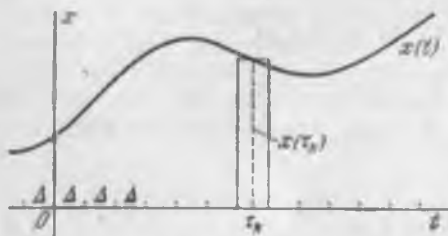


Рис. 2.1.3.

Значение функции $x(t)$ в точке τ_k равно $x(\tau_k)$ (рис. 2.1.3). Построим прямоугольный импульс длительности Δ , имеющий высоту $x(\tau_k)$, действующий в промежутке времени $(\tau_k - \frac{\Delta}{2}, \tau_k + \frac{\Delta}{2})$. Очевидно, что единичный прямоугольный импульс, действующий в этом промежутке времени, выражается функцией

$\delta_{\Delta}(t - \tau_k)$. Высота этого импульса равна $\frac{1}{\Delta}$. Чтобы получить импульс высоты $x(\tau_k)$, необходимо функцию $\delta_{\Delta}(t - \tau_k)$ разделить на $\frac{1}{\Delta}$ и умножить на $x(\tau_k)$. Таким образом, один прямоугольный импульс, имеющий ту же ординату $x(\tau_k)$, что и данная функция $x(t)$ в точке $t = \tau_k$, действующий в интервале времени $(\tau_k - \frac{\Delta}{2}, \tau_k + \frac{\Delta}{2})$, математически

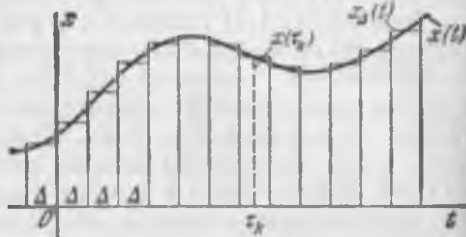


Рис. 2.1.4.

выражается функцией $x(\tau_k) \delta_{\Delta}(t - \tau_k) \Delta$. Сумма таких импульсов, соответствующих интервалам, на которые мы разбили ось t , представляет ступенчатую функцию $x_{\Delta}(t)$, имеющую в каждом интервале постоянное значение, равное значению функции $x(t)$ в середине этого интервала (рис. 2.1.4):

$$x_{\Delta}(t) = \sum_{k=-\infty}^{\infty} x(\tau_k) \delta_{\Delta}(t - \tau_k) \Delta. \quad (2.1.19)$$

При малой длине интервалов Δ эта функция может быть принята приближенно совпадающей с данной функцией $x(t)$. Чем меньше Δ , тем точнее будет совпадение. В пределе при $\Delta \rightarrow 0$ ступенчатая функция $x_\Delta(t)$ стремится к $x(t)$. При этом функция $\delta_\Delta(t - \tau)$ стремится к δ -функции $\delta(t - \tau)$, а сумма стремится к соответствующему интегралу. В результате переход к пределу при $\Delta \rightarrow 0$ в (2.1.19) дает формулу (2.1.17).

Формула (2.1.19) при любом Δ дает разложение соответствующей ступенчатой функции $x_\Delta(t)$ на прямоугольные импульсы длительности Δ . Следовательно, формула (2.1.17), являющаяся пределом формулы (2.1.19) при $\Delta \rightarrow 0$, дает разложение функции $x(t)$ на мгновенные импульсы.

Для дальнейшего нам понадобится еще выразить импульсную δ -функцию в форме интеграла Фурье. Для этого мы сначала представим интегралом Фурье единичный прямоугольный импульс $\delta_\Delta(t)$, определяемый формулой (2.1.18), а затем переходом к пределу при $\Delta \rightarrow 0$ получим искомое представление интегралом Фурье δ -функции.

Из анализа известно, что любая абсолютно интегрируемая функция $f(t)$ (т. е. такая функция, интеграл от абсолютной величины которой в бесконечных пределах имеет конечное значение) может быть представлена интегралом Фурье [50]

$$f(t) = \int_{-\infty}^{\infty} \varphi(\omega) e^{i\omega t} d\omega, \quad (2.1.20)$$

где

$$\varphi(\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(\tau) e^{-i\omega\tau} d\tau. \quad (2.1.21)$$

Подставляя в (2.1.21) выражение (2.1.18) прямоугольного импульса, будем иметь

$$\begin{aligned} \varphi(\omega) &= \frac{1}{2\pi\Delta} \int_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} e^{-i\omega\tau} d\tau = \frac{1}{2\pi\Delta} \left. \frac{e^{-i\omega\tau}}{(-i\omega)} \right|_{-\frac{\Delta}{2}}^{\frac{\Delta}{2}} = \\ &= \frac{1}{\pi\omega\Delta} \frac{e^{i\omega\frac{\Delta}{2}} - e^{-i\omega\frac{\Delta}{2}}}{2i} = \frac{1}{\pi\omega\Delta} \sin \frac{\omega\Delta}{2}. \end{aligned} \quad (2.1.22)$$

Подставляя в (2.1.20) выражение (2.1.22) изображения Фурье прямоугольного импульса, получим

$$\delta_\Delta(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} \frac{2}{\omega\Delta} \sin \frac{\omega\Delta}{2} d\omega. \quad (2.1.23)$$

Переходя к пределу при $\Delta \rightarrow 0$ и принимая во внимание, что при этом $\frac{2}{\omega\Delta} \sin \frac{\omega\Delta}{2} \rightarrow 1$, получим

$$\delta(t) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega t} d\omega. \quad (2.1.24)$$

Эта формула дает разложение мгновенного единичного импульса на гармонические колебания всех возможных частот.

Заметим, что интеграл в (2.1.24) представляет собой расходящийся несобственный интеграл. Однако это не мешает формально пользоваться формулой (2.1.24) и получать правильные результаты. В частности, для любой абсолютно интегрируемой функции $x(t)$ выражение δ -функции (2.1.24) можно формально подставить в формулу (2.1.17). Изменив после этого порядок интегрирования, получим на основании известной формулы Фурье в правой части $x(t)$.

§ 2.2. Весовые функции линейных систем

Рассмотрим линейную систему с одним входом и одним выходом. Разложив действующее на линейную систему возмущение $x(t)$ на элементарные импульсы, т. е. представив функцию $x(t)$ в виде интеграла (2.1.17), воспользуемся принципом суперпозиции в интегральной форме (1.5.12). В данном случае параметром λ является момент действия очередного элементарного импульса τ ,

$$x(t, \lambda) = \delta(t - \tau), \quad c(\lambda) d\lambda = x(\tau) d\tau.$$

Следовательно, на основании принципа суперпозиции (1.5.12) реакция $y(t)$ линейной системы на произвольное возмущение $x(t)$ имеет вид

$$y(t) = Ax(t) = \int_{-\infty}^{\infty} x(\tau) A_t \delta(t - \tau) d\tau, \quad (2.2.1)$$

где индекс t у оператора A показывает, что этот оператор действует над функцией $\delta(t - \tau)$, рассматриваемой как функция t при фиксированном значении τ . Формула (2.2.1) показывает, что для нахождения реакции линейной системы на произвольное возмущение $x(t)$ достаточно знать ее реакцию на единичный импульс $\delta(t - \tau)$, действующий на нее в произвольный момент τ . Эта реакция зависит от переменных t и τ , т. е. от момента действия импульса τ и текущего момента t :

$$g(t, \tau) = A_t \delta(t - \tau). \quad (2.2.2)$$

Функция $g(t, \tau)$, определяемая формулой (2.2.2), является исчерпывающей характеристикой линейной системы и называется ее *весовой* или *импульсной переходной функцией*. Таким образом,

весовая или импульсная переходная функция линейной системы представляет собой реакцию этой системы в момент t на единичный импульс, действующий на систему в момент τ .

Пользуясь понятием весовой функции, мы можем записать зависимость (2.2.1) между входной и выходной переменными произвольной линейной системы в виде

$$y(t) = \int_{-\infty}^{\infty} g(t, \tau) x(\tau) d\tau. \quad (2.2.3)$$

Таким образом, оператор любой линейной системы может быть представлен в форме линейного интегрального оператора.

Выясним смысл термина «весовая функция». Для этого рассмотрим рис. 2.2.1, где единичный импульс, действующий на систему в момент τ , изображен вертикальной стрелкой, а кривая $g(t, \tau)$ изображает реакцию системы на этот импульс, т. е. ее весовую функцию, рассматриваемую как функция t при данном фиксированном значении τ .

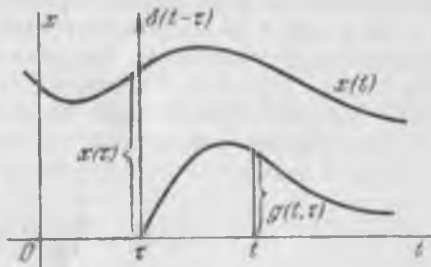


Рис. 2.2.1.

Согласно формуле (2.2.3) для определения значения выходной переменной линейной системы в момент времени t необходимо значение $x(\tau)$, которое имело входное возмущение $x(t)$ в момент τ , умножить на ординату весовой функции $g(t, \tau)$ при данном значении t (рис. 2.2.1) и все подобные произведения просуммировать. Таким образом, весовая функция $g(t, \tau)$ характеризует «удельный вес» возмущения, которое действовало в момент времени τ , в формировании выходной переменной системы в данный момент времени t . Этим и объясняется смысл термина «весовая функция».

Весовая функция любой реально существующей системы равна нулю при значении второго аргумента, большем значения первого:

$$g(t, \tau) = 0 \quad \text{при } \tau > t. \quad (2.2.4)$$

Это свойство отражает тот факт, что никакая физическая система не может реагировать в данный момент на возмущение, которое будет действовать на нее позже. Это свойство мы будем называть *условием физической возможности системы* *). Никакая система,

*) Чаще это условие называют условием физической осуществимости. Однако этот термин неточен, так как осуществимость предполагает умение эту систему сделать, а между тем далеко не всякую систему, удовлетворяющую условию (2.2.4), можно осуществить техническими средствами.

весовая функция которой не удовлетворяет этому условию, физически немыслима.

Теоретически мы можем рассматривать как физически возможные, так и физически невозможные системы. Более того, в общей теории линейных систем, как мы увидим ниже, большую роль играет один класс физически невозможных систем. Кроме того, часто приходится проектировать системы для решения таких задач, которые не могут быть идеально точно решены физическими возможными системами. В таких случаях приходится рассматривать идеальные физически невозможные системы для оценки точности проектируемых физически возможных систем.

Рассмотрим теперь физически возможную линейную систему, находящуюся в покое до момента t_0 , предполагая, что возмущение $x(t)$ действует на нее начиная с момента t_0 . Для такой системы в формуле (2.2.3) подынтегральная функция равна нулю при $\tau < t_0$ и при $\tau > t$. Действительно, возмущение $x(t)$ не действует на систему до момента t_0 , что равноценно предположению $x(\tau) = 0$ при $\tau < t_0$. При $\tau > t$ весовая функция $g(t, \tau)$ равна нулю вследствие условия физической возможности (2.2.4). Следовательно, для физически возможной системы соотношение (2.2.3) между входной и выходной переменными переписывается таким образом:

$$y(t) = \int_{t_0}^t g(t, \tau) x(\tau) d\tau. \quad (2.2.5)$$

Имея в виду, что в возмущение $x(t)$ могут входить импульсы типа δ -функции, условимся считать, что пределы интегрирования всегда включаются в интервал интегрирования, т. е. интеграл в (2.2.5) понимать как предел интеграла от $t_0 - \varepsilon$ до $t + \varepsilon$ при $\varepsilon \rightarrow 0$.

Рассмотрим типовые идеальные системы, осуществляющие элементарные операции анализа.

1. Система тождественного преобразования, или, что одно и то же, идеальная следящая система, представляет собой такую систему, у которой выходная и входная переменные тождественно равны друг другу: $y(t) \equiv x(t)$. Сравнивая формулы (2.1.17) и (2.2.3), видим, что весовая функция идеальной следящей системы есть δ -функция:

$$g_s(t, \tau) = \delta(t - \tau). \quad (2.2.6)$$

2. Идеальный экстраполятор представляет собой такую систему, у которой выходная переменная опережает входную переменную на заданный интервал времени $a > 0$: $y(t) = x(t + a)$. Для того чтобы найти весовую функцию идеального экстраполятора, заменим в формуле (2.1.17) t величиной $t + a$ и сравним полученную формулу с (2.2.3). Тогда получим

$$g_e(t, \tau) = \delta(t - \tau + a) = \delta(t - (\tau - a)). \quad (2.2.7)$$

Эта формула показывает, что идеальный экстраполятор реагирует на единичный импульс, действующий в момент времени τ , единичным импульсом в момент времени $\tau - a$, предшествующий моменту τ . Следовательно, идеальный экстраполятор является физически невозможной линейной системой.

3. *Идеальное запаздывающее звено* получается из экстраполятора при $a < 0$. Поэтому, полагая в (2.2.7) $a = -l$, получим весовую функцию запаздывающего звена:

$$g_3(t, \tau) = \delta(t - \tau - l) = \delta(t - (\tau + l)). \quad (2.2.8)$$

4. *Безынерционный усилитель* производит усиление входной величины. При этом коэффициент усиления в общем случае может быть функцией времени. Таким образом, для безынерционного усилителя $y(t) = k(t)x(t)$. Умножая формулу (2.1.17) на k и сравнивая результат с (2.2.3), видим, что весовая функция безынерционного усилителя имеет вид

$$g_4(t, \tau) = k(t)\delta(t - \tau). \quad (2.2.9)$$

5. *Идеальное дифференцирующее звено*. Его выходная переменная в каждый момент времени равна значению производной входной переменной по времени в тот же момент: $y(t) = x'(t)$. Дифференцируя формулу (2.1.17) по t и сравнивая результат с (2.2.3), убеждаемся в том, что весовая функция идеального дифференцирующего звена определяется формулой

$$g_5(t, \tau) = \delta'(t - \tau). \quad (2.2.10)$$

Таким образом, весовая функция идеального дифференцирующего звена представляет собой производную δ -функции.

6. *Идеальное интегрирующее звено* дает на выходе функцию

$$y(t) = \int_{t_0}^t x(\tau) d\tau. \quad (2.2.11)$$

Сравнивая эту формулу с (2.2.3), мы видим, что весовая функция идеального интегрирующего звена равна единице при $t > \tau$ и нулю при $t < \tau$, т. е. представляет собой единичную ступенчатую функцию:

$$g_6(t, \tau) = 1(t - \tau). \quad (2.2.12)$$

Пример 2.2.1. Найдем в качестве примера весовую функцию объекта управления, представляющего собой двигатель, у которого управляющим входным сигналом является вращающий момент M , выходным сигналом — угловая скорость вала ω , а сопротивление вращению вала пропорционально угловой скорости ω .

Обозначив момент инерции вала двигателя вместе с присоединенными к нему деталями через J , можно написать дифференциальное уравнение вращательного движения вала:

$$J\dot{\omega} = M - h\omega, \quad (2.2.13)$$

где $h\omega$ — момент сил сопротивления. Для определения весовой функции двигателя следует задать входной сигнал M в виде мгновенного единичного импульса $\delta(t - \tau)$. Тогда выходной сигнал ω по определению и будет представлять собой искомую весовую функцию $g(t, \tau)$. В результате для определения весовой функции $g(t, \tau)$ получим уравнение

$$J\ddot{g} + h\dot{g} = \delta(t - \tau), \quad (2.2.14)$$

где точка означает дифференцирование по t , а τ рассматривается как фиксированный момент действия единичного импульса. Интегрируя уравнение (2.2.14) и учитывая, что двигатель является физической возможной системой, вследствие чего $g(t, \tau) = 0$ при $t < \tau$, получим при $t > \tau$

$$g(t, \tau) = e^{-\frac{h}{J}t} \int_{-\infty}^t e^{\frac{h}{J}\sigma} \frac{\delta(\sigma - \tau)}{J} d\sigma = \frac{1}{J} e^{-\frac{h}{J}(t-\tau)}, \quad (2.2.15)$$

Пример 2.2.2. Если в условиях предыдущего примера за выходной сигнал двигателя принять угол поворота вала φ , то будем иметь $\omega = \dot{\varphi}$, и дифференциальное уравнение вращательного движения вала примет вид

$$J\ddot{\varphi} + h\dot{\varphi} = M. \quad (2.2.16)$$

Для определения весовой функции зададим входной сигнал в виде единичного импульса $M = \delta(t - \tau)$. Тогда получим для весовой функции $g_1(t, \tau)$ дифференциальное уравнение

$$J\ddot{g}_1 + h\dot{g}_1 = \delta(t - \tau). \quad (2.2.17)$$

Интегрируя это уравнение, получим

$$\frac{d}{dt} g_1(t, \tau) = g(t, \tau) = \begin{cases} 0 & \text{при } t < \tau, \\ \frac{1}{J} e^{-\frac{h}{J}(t-\tau)} & \text{при } t > \tau \end{cases}$$

и

$$g_1(t, \tau) = \int_{t_0}^t g(\sigma, \tau) d\sigma = \int_{t_0}^{\tau} 0 \cdot d\sigma + \frac{1}{J} \int_{\tau}^t e^{-\frac{h}{J}(\sigma-\tau)} d\sigma = \frac{1}{h} [1 - e^{-\frac{h}{J}(t-\tau)}]. \quad (2.2.18)$$

Пример 2.2.3. Объектом управления является цилиндрический нагреватель (рис. 2.2.2). Нагреваемое вещество находится внутри цилиндра. Управляющим входным сигналом является количество тепла $q(t)$, подводимое к единице площади внешней поверхности цилиндра в единицу времени. Целью управления является обеспечение заданного закона изменения температуры внутренней поверхности цилиндрической стенки нагревателя. Эта температура и является в данном случае выходным сигналом нагревателя.

Если толщина стенки нагревателя мала по сравнению с его диаметром, то можно пренебречь разностью площадей внешней и внутренней поверхностей стенки. Тогда, если выбрать начало координат O на внешней поверхности стенки нагревателя и направить ось Ox по нормали к ней внутрь стенки (рис. 2.2.3), то закон изменения температуры в стенке $\theta(t, x)$ будет определяться одномерным уравнением теплопроводности (см., например, [63],

т. II, § 24)

$$\frac{\partial \theta}{\partial t} = \frac{k}{c\rho} \frac{\partial^2 \theta}{\partial x^2}, \quad (2.2.19)$$

где k — удельная теплопроводность, c — удельная теплоемкость, а ρ — плотность материала стенки нагревателя.

Для полного описания процесса изменения температуры в стенке нагревателя необходимо добавить к уравнению (2.2.19) граничные условия — уравнение подвода тепла к внешней поверхности стенки и уравнение отвода тепла от внутренней поверхности стенки:

$$\left(\frac{\partial \theta}{\partial x} \right)_{x=0} = -\frac{q(t)}{k}, \quad \left(\frac{\partial \theta}{\partial x} \right)_{x=l} = -\frac{h}{k} \theta(t, l), \quad (2.2.20)$$

где в дополнение к предыдущим обозначениям h — коэффициент теплопередачи от стенки нагревателя к нагреваемому веществу.

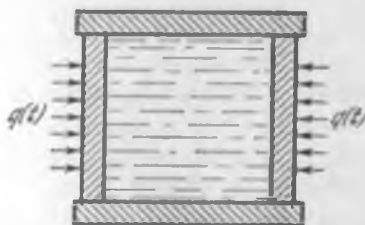


Рис. 2.2.2.

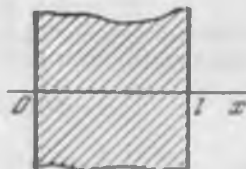


Рис. 2.2.3.

Для определения весовой функции нагревателя следует принять входной сигнал $q(t)$ равным $\delta(t - \tau)$. Проинтегрировав уравнение (2.2.19) при граничных условиях (2.2.20) при $q(t) = \delta(t - \tau)$, найдем весовую функцию нагревателя как значение температуры $\theta(t, x)$ на внутренней поверхности стенки: $g(t, \tau) = \theta(t, l)$.

Легко проверить непосредственной подстановкой, что уравнению (2.2.19) и второму граничному условию (2.2.20) удовлетворяет функция

$$\theta(t, x) = \int_{-\infty}^{\infty} \alpha(\omega) f(x, \omega) e^{i\omega t} d\omega, \quad (2.2.21)$$

где

$$f(x, \omega) = (k \sqrt{i\omega - ah}) e^{\sqrt{i\omega - ah} x} + (k \sqrt{i\omega + ah}) e^{-\sqrt{i\omega + ah} x}, \quad (2.2.22)$$

$a^2 = k/c\rho$, а $\alpha(\omega)$ — произвольная функция. Этой произвольной функцией можно распорядиться так, чтобы удовлетворить и первому граничному условию (2.2.20). Полагая $q(t) = \delta(t - \tau)$ и пользуясь представлением δ -функции в виде интеграла Фурье (2.1.24), перепишем первое граничное условие (2.2.20) в виде

$$\left(\frac{\partial \theta}{\partial x} \right)_{x=0} = -\frac{1}{2\pi k} \int_{-\infty}^{\infty} e^{i\omega(t-\tau)} d\omega.$$

Подставляя сюда выражение $\partial\theta/\partial x$, полученное дифференцированием формулы (2.2.21), будем иметь

$$\int_{-\infty}^{\infty} \alpha(\omega) f'_x(0, \omega) e^{i\omega t} d\omega = -\frac{1}{2\pi k} \int_{-\infty}^{\infty} e^{i\omega(t-\tau)} d\omega.$$

Это условие обращается в тождество, если положить

$$\alpha(\omega) f'_x(0, \omega) = -\frac{1}{2\pi k} e^{-i\omega\tau}. \quad (2.2.23)$$

Для определения $f'_x(0, \omega)$ продифференцируем формулу (2.2.22) по x и положим после этого $x = 0$. Получим

$$f'_x(0, \omega) = -\varphi(i\omega) \sqrt{i\omega}; \quad (2.2.24)$$

где для краткости положено

$$\varphi(i\omega) = \left(\frac{k}{a} \sqrt{i\omega} + h\right) e^{\sqrt{i\omega} \frac{l}{a}} - \left(\frac{k}{a} \sqrt{i\omega} - h\right) e^{-\sqrt{i\omega} \frac{l}{a}}. \quad (2.2.25)$$

Подставляя выражение (2.2.24) в (2.2.23) и решая полученное уравнение относительно $\alpha(\omega)$, будем иметь

$$\alpha(\omega) = \frac{1}{2\pi k \varphi(i\omega) \sqrt{i\omega}} e^{-i\omega\tau}. \quad (2.2.26)$$

Подставляя это выражение $\alpha(\omega)$ в (2.2.21), получим

$$\theta(t, x) = \frac{1}{2\pi k} \int_{-\infty}^{\infty} \frac{f(x, \omega)}{\varphi(i\omega) \sqrt{i\omega}} e^{i\omega(t-\tau)} d\omega. \quad (2.2.27)$$

Наконец, для определения весовой функции положим в (2.2.27) $x = l$. Тогда, принимая во внимание, что на основании (2.2.22) $f(l, \omega) = 2k \sqrt{i\omega}$, получим для искомой весовой функции нагревателя формулу

$$g(t, \tau) = \theta(t, l) = \frac{1}{\pi} \int_{-\infty}^{\infty} e^{i\omega(t-\tau)} \frac{d\omega}{\varphi(i\omega)}. \quad (2.2.28)$$

Перейдем теперь к многомерным линейным системам, имеющим любое число входов и выходов.

На основании принципа суперпозиции действие каждого входного возмущения на многомерную линейную систему можно рассматривать отдельно, а затем результаты действия отдельных возмущений на каждом выходе просуммировать.

Для изучения реакции на k -м выходе линейной системы на возмущение, действующее на одном только h -м входе, можно рассматривать эту систему как систему с одним входом и одним выходом. Тогда для вычисления реакции на k -м выходе линейной системы на любое возмущение, действующее на h -м входе, достаточно будет знать соответствующую весовую функцию $g_{kh}(t, \tau)$. Эта весовая функция представляет собой реакцию на k -м выходе

линейной системы на единичный импульс, действующий на h -м входе в момент τ при отсутствии возмущений на остальных входах. Совокупность весовых функций $g_{kh}(t, \tau)$ ($k = 1, \dots, m$; $h = 1, \dots, n$), соответствующих всем входам и всем выходам, является исчерпывающей характеристикой многомерной линейной системы.

Зная весовые функции многомерной линейной системы, соответствующие всем входам и выходам, можно для вычисления ее реакции на каждом выходе на возмущение, действующее только на каком-нибудь одном входе, применить формулу (2.2.3) или (2.2.5) (для физически возможных систем). Суммируя полученные результаты для каждого отдельного выхода по всем входам, найдем все выходные переменные рассматриваемой линейной системы, соответствующие одновременному действию любых возмущений на всех ее входах.

Таким образом, изложенная теория применима как к одномерным, так и к многомерным линейным системам.

Согласно изложенному выходные переменные линейной системы, имеющей n входов и m выходов, выражаются через ее весовые функции $g_{kh}(t, \tau)$ ($k = 1, \dots, m$; $h = 1, \dots, n$) формулой

$$y_k(t) = \sum_{h=1}^n \int_{-\infty}^{\infty} g_{kh}(t, \tau) x_h(\tau) d\tau \quad (k = 1, \dots, m). \quad (2.2.29)$$

Для физически возможных линейных систем все весовые функции равны нулю при $\tau > t$. Поэтому для физически возможной многомерной линейной системы, находящейся в покое до момента t_0 , подынтегральные функции в (2.2.29) равны нулю при $\tau < t_0$ и при $\tau > t$, вследствие чего формула (2.2.29) может быть переписана в виде

$$y_k(t) = \sum_{h=1}^n \int_{t_0}^t g_{kh}(t, \tau) x_h(\tau) d\tau \quad (k = 1, \dots, m). \quad (2.2.30)$$

Заметим, что если ввести матрицу весовых функций

$$g(t, \tau) = \begin{bmatrix} g_{11}(t, \tau) & g_{12}(t, \tau) & \dots & g_{1n}(t, \tau) \\ g_{21}(t, \tau) & g_{22}(t, \tau) & \dots & g_{2n}(t, \tau) \\ \dots & \dots & \dots & \dots \\ g_{m1}(t, \tau) & g_{m2}(t, \tau) & \dots & g_{mn}(t, \tau) \end{bmatrix} \quad (2.2.31)$$

и векторные входной и выходной сигналы $x(t)$ и $y(t)$ с составляющими $x_1(t), \dots, x_n(t)$ и $y_1(t), \dots, y_m(t)$ соответственно, то формулы (2.2.29) и (2.2.30) можно коротко записать в виде (2.2.3) и (2.2.5). Таким образом, формулы (2.2.3) и (2.2.5) справедливы для систем с любым числом входов и выходов.

§ 2.3. Характеристика реакции линейной системы на показательное возмущение. Частотная характеристика

Взяв в качестве элементарных возмущений, на которые можно разложить произвольное возмущение, единичные импульсы, т. е. δ -функции, мы получили характеристику линейной системы — весовую функцию и выразили через нее реакцию линейной системы на любое возмущение.

Тип элементарных возмущений, на которые можно разложить произвольное возмущение, можно выбрать различными способами. В зависимости от выбора типа элементарных возмущений получаются различные характеристики линейных систем. Во многих задачах практики удобно взять в качестве элементарных возмущений гармонические колебания всех возможных частот. Известно, что при весьма общих условиях любую функцию можно разложить в ряд Фурье или представить интегралом Фурье. Поэтому, зная реакцию линейной системы на гармонические колебания всех возможных частот и пользуясь принципом суперпозиции, мы можем определить реакцию системы на произвольное возмущение.

Рассмотрим некоторую функцию $x(t)$. Если она абсолютно интегрируема, то ее можно представить интегралом Фурье:

$$x(t) = \int_{-\infty}^{\infty} c(i\omega) e^{i\omega t} d\omega, \quad (2.3.1)$$

где

$$c(i\omega) = \frac{1}{2\pi} \int_{-\infty}^{\infty} x(t) e^{-i\omega t} dt. \quad (2.3.2)$$

Если возмущение $x(t)$ действует на входе линейной системы, то на основании принципа суперпозиции выходная функция системы будет равна

$$y(t) = Ax(t) = \int_{-\infty}^{\infty} c(i\omega) A_t e^{i\omega t} d\omega, \quad (2.3.3)$$

где A — оператор системы. В этой формуле $A_t e^{i\omega t}$ представляет собой реакцию линейной системы на гармонические колебания частоты ω . Эта реакция может быть принята за характеристику линейной системы. Зная эту характеристику, мы можем вычислить по формуле (2.3.3) реакцию линейной системы на произвольное возмущение, которое можно представить рядом или интегралом Фурье. Обычно эту характеристику несколько обобщают и рассматривают в качестве элементарного возмущения показательную функцию e^{st} , где s — произвольный комплексный параметр.

Если параметр s имеет чисто мнимое значение, то e^{st} описывает гармонические колебания; если этот параметр имеет комплексное значение с отрицательной действительной частью, то e^{st} представляет собой затухающие гармонические колебания; если параметр s имеет комплексное значение с положительной действительной частью, то e^{st} — расходящиеся гармонические колебания; наконец, если s — действительное число, то e^{st} можно рассматривать как затухающие или расходящиеся гармонические колебания нулевой частоты. Таким образом, семейство показательных функций e^{st} с комплексным параметром s охватывает гармонические колебания всех возможных частот, как с постоянными амплитудами, так и с амплитудами, изменяющимися со временем по экспоненциальному закону.

Из физических соображений ясно, что любая система, которую мы будем возбуждать гармоническими колебаниями, будет реагировать на них тоже каким-то колебательным движением. Естественно определить такую характеристику линейной системы, которая сама не будет являться колебательным процессом, а будет характеризовать преобразование амплитуды и сдвиг фазы выходных колебаний по отношению к входным. Чтобы получить такую характеристику, можно разделить реакцию системы на само входное показательное возмущение e^{st} , т. е. принять за характеристику линейной системы функцию

$$z(t, s) = \frac{A_1 e^{st}}{e^{st}}. \quad (2.3.4)$$

Функцию $z(t, s)$ мы будем называть *характеристикой реакции линейной системы на показательное возмущение*. Таким образом, характеристикой реакции линейной системы на показательное возмущение называется отношение ее реакции на возмущение, представляющее собой показательную функцию времени, к этой показательной функции. В общем случае характеристика реакции линейной системы на показательное возмущение зависит не только от параметра s , но и от времени t .

В частном случае при чисто мнимых значениях параметра s мы получаем так называемую *частотную характеристику* линейной системы:

$$z(t, i\omega) = \frac{A_1 e^{i\omega t}}{e^{i\omega t}}. \quad (2.3.5)$$

Частотная характеристика линейной системы, в общем случае также зависит не только от частоты, но и от времени.

На основании формулы (2.3.5) реакция линейной системы на гармонические колебания данной частоты ω и единичной амплитуды $e^{i\omega t}$ равна произведению значения частотной характеристики

системы при данном значении частоты на входную функцию $e^{i\omega t}$:

$$A_1 e^{i\omega t} = z(t, i\omega) e^{i\omega t}. \quad (2.3.6)$$

Найдем теперь реакцию линейной системы на входные гармонические колебания произвольной амплитуды и фазы

$$x(t) = a e^{i(\omega t + \varphi)}. \quad (2.3.7)$$

На основании принципа суперпозиции и формулы (2.3.6) выходная переменная системы при входном возмущении (2.3.7) определяется формулой

$$\begin{aligned} y(t) &= A_1 \{a e^{i(\omega t + \varphi)}\} = a e^{i\varphi} A_1 e^{i\omega t} = \\ &= a e^{i\varphi} \cdot z(t, i\omega) e^{i\omega t} = a z(t, i\omega) e^{i(\omega t + \varphi)}. \end{aligned} \quad (2.3.8)$$

Частотная характеристика линейной системы в общем случае является комплексной функцией. Как и всякую комплексную величину, ее можно представить в показательной форме:

$$z(t, i\omega) = |z(t, i\omega)| e^{i \arg z(t, i\omega)}. \quad (2.3.9)$$

Подставляя это выражение в (2.3.8), получим

$$y(t) = a |z(t, i\omega)| \exp \{i[\omega t + \varphi + \arg z(t, i\omega)]\}. \quad (2.3.10)$$

Эта формула показывает, что линейная система реагирует на гармонические колебания входной величины в общем случае колебаниями переменной амплитуды

$$b(t) = a |z(t, i\omega)| \quad (2.3.11)$$

и переменного сдвига фазы

$$\psi(t) = \varphi + \arg z(t, i\omega). \quad (2.3.12)$$

Таким образом, модуль частотной характеристики линейной системы определяет изменение амплитуды колебаний данной частоты при прохождении их через эту систему, а аргумент частотной характеристики определяет изменение фазы колебаний. А именно модуль частотной характеристики представляет собой коэффициент усиления амплитуды гармонических колебаний (в общем случае переменный) при прохождении их через линейную систему. Аргумент частотной характеристики представляет собой сдвиг фазы выходных колебаний по отношению к входным колебаниям. Вследствие этого модуль частотной характеристики линейной системы $|z(t, i\omega)|$ обычно называется *амплитудной* частотной характеристикой системы, а аргумент частотной характеристики $\arg z(t, i\omega)$ — *фазовой* частотной характеристикой системы.

Пользуясь частотной характеристикой линейной системы, мы можем записать соотношение (2.3.3) между входной и выходной

переменными в следующем виде:

$$y(t) = \int_{-\infty}^{\infty} c(i\omega) z(t, i\omega) e^{i\omega t} d\omega. \quad (2.3.13)$$

Сравним эту формулу с (2.2.3). Как в первом, так и во втором случаях за характеристику системы принимается некоторая функция двух переменных: текущего времени t и какого-то параметра. В качестве такого параметра в формуле (2.2.3) принимается момент действия единичного импульса. В формуле (2.3.13) параметром является круговая частота гармонических колебаний ω . Для того чтобы воспользоваться формулой (2.2.3), необходимо знать только весовую функцию $g(t, \tau)$ и само входное возмущение. Для того чтобы воспользоваться формулой (2.3.13), нужно сначала найти преобразование Фурье $c(i\omega)$ входного сигнала $x(t)$. Таким образом, для произвольных линейных систем применение частотных характеристик обязательно включает лишнюю операцию перехода к преобразованию Фурье. Очевидно, что это будет общей закономерностью. Какими бы элементарными возмущениями мы ни пользовались, кроме δ -функции, мы обязательно должны будем находить какое-то преобразование входного возмущения, а потом уже определять выходную переменную системы. И только использование в качестве элементарного возмущения δ -функции дает существенное упрощение, так как при этом исключается необходимость дополнительного преобразования входного возмущения. Таким образом, для произвольных линейных систем применение частотных характеристик нерационально. Только для систем с медленно изменяющимися параметрами, когда можно считать на значительных отрезках времени параметры системы постоянными, использование частотных характеристик может представить некоторые выгоды.

Сделаем еще одно замечание о частотных характеристиках и характеристиках реакции линейных систем на показательное возмущение. Если рассматривать эти характеристики для произвольных моментов включения системы, то они будут зависеть еще от момента начала действия показательного возмущения t_0 и, таким образом, будут фактически функциями трех аргументов: t, s, t_0 , а не двух. Чтобы избежать этого усложнения, частотные характеристики и характеристики реакции линейных систем на показательное возмущение обычно определяют для случая установившейся работы системы, т. е. такого режима работы, когда возмущение действует на систему неограниченно долго, т. е. для $t_0 = -\infty$. Это не мешает с помощью частотной характеристики находить реакцию системы на возмущения, действующие начиная с произвольного момента времени t_0 . Для этого при определении преобразования Фурье $c(i\omega)$ входного сигнала $x(t)$ следует положить функцию $x(t)$ равной нулю при $t < t_0$.

Весовая функция, как мы видели, является полной, исчерпывающей характеристикой линейной системы. Частотная характеристика и характеристика реакции системы на показательное возмущение также являются исчерпывающими характеристиками линейной системы. Естественно возникает мысль, что между весовой функцией и частотной характеристикой или характеристикой реакции системы на показательное возмущение должна существовать связь. Чтобы найти эту связь, достаточно вычислить реакцию системы на показательное возмущение с помощью весовой функции по формуле (2.2.3):

$$A_1 e^{st} = \int_{-\infty}^{\infty} g(t, \tau) e^{s\tau} d\tau. \quad (2.3.14)$$

Чтобы получить характеристику реакции системы на показательное возмущение, нужно полученный интеграл разделить на входное возмущение e^{st} . Тогда получим

$$z(t, s) = \frac{1}{e^{st}} \int_{-\infty}^{\infty} g(t, \tau) e^{s\tau} d\tau = \int_{-\infty}^{\infty} g(t, \tau) e^{-s(t-\tau)} d\tau. \quad (2.3.15)$$

Для физически возможной системы формула (2.3.15) приобретает вид

$$z(t, s) = \int_{-\infty}^t g(t, \tau) e^{-s(t-\tau)} d\tau. \quad (2.3.16)$$

Полагая в формулах (2.3.15) и (2.3.16) $s = i\omega$, получим выражение частотной характеристики через весовую функцию. В результате частотная характеристика физически возможной линейной системы выразится формулой

$$z(t, i\omega) = \int_{-\infty}^t g(t, \tau) e^{-i\omega(t-\tau)} d\tau. \quad (2.3.17)$$

Выразим теперь весовую функцию линейной системы через частотную характеристику. Для этого нужно разложить δ -функцию на элементарные гармонические колебания. Согласно формуле (2.1.24)

$$\delta(t-\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{i\omega(t-\tau)} d\omega. \quad (2.3.18)$$

Сравнивая эту формулу с общей формулой (2.3.1) интеграла Фурье, приходим к заключению, что преобразование Фурье δ -функции определяется формулой

$$c(i\omega) = \frac{1}{2\pi} e^{-i\omega\tau}. \quad (2.3.19)$$

Подставляя это выражение в формулу (2.3.13), выразим реакцию системы на δ -функцию, т. е. ее весовую функцию, через частотную характеристику:

$$g(t, \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} z(t, i\omega) e^{i\omega(t-\tau)} d\omega. \quad (2.3.20)$$

Таким образом, зная весовую функцию системы, можно определить ее характеристику реакции на показательное возмущение и частотную характеристику, и, наоборот, зная частотную характеристику системы, можно определить ее весовую функцию.

Заметим, что установившийся режим колебаний под действием гармонических входных колебаний существует не для всех линейных систем, а только для устойчивых систем (определение устойчивости см. в гл. 6). Вследствие этого формулы (2.3.13) и (2.3.20) применимы только к устойчивым линейным системам.

Совершенно так же определяются характеристики реакции многомерной линейной системы на показательные возмущения, действующие на различных входах, и ее частотные характеристики. При этом характеристики реакции системы на показательные возмущения и частотные характеристики связаны с соответствующими весовыми функциями тем же формулами (2.3.16), (2.3.17) и (2.3.20).

§ 2.4. Стационарные линейные системы.

Передаточная функция и частотная характеристика стационарной линейной системы

Изучим подробнее стационарные линейные системы. На основании определения стационарной системы реакция стационарной линейной системы на единичный импульс зависит только от интервала времени между моментом действия импульса τ и данным моментом t . Иными словами, весовая функция стационарной линейной системы зависит только от разности аргументов $t - \tau$:

$$g(t, \tau) = w(t - \tau). \quad (2.4.1)$$

Вследствие этого формула (2.2.5) дает следующее выражение для реакции физически возможной стационарной линейной системы на любое возмущение $x(t)$:

$$y(t) = \int_{-\infty}^t w(t - \tau) x(\tau) d\tau. \quad (2.4.2)$$

Для общности мы полагаем нижний предел интегрирования равным $-\infty$, чтобы охватить случай неограниченно долго действующих возмущений. В частном случае, когда возмущение начинается

действовать на систему в момент t_0 , в формуле (2.4.2) следует положить $x(\tau) = 0$ при $\tau < t_0$.

Производя в формуле (2.4.2) замену переменных $\xi = t - \tau$, приведем ее к виду

$$y(t) = \int_0^{\infty} w(\xi) x(t - \xi) d\xi. \quad (2.4.3)$$

Найдем характеристику реакции физически возможной стационарной линейной системы на показательное возмущение. Подставляя выражение (2.4.1) в общую формулу (2.3.16), получим

$$z(t, s) = \int_{-\infty}^t w(t - \tau) e^{-s(t - \tau)} d\tau. \quad (2.4.4)$$

Замена $\xi = t - \tau$ приводит эту формулу к виду

$$z(t, s) = \int_0^{\infty} w(\xi) e^{-s\xi} d\xi. \quad (2.4.5)$$

Правая часть этой формулы не зависит от времени t . Следовательно, для любой стационарной линейной системы характеристика реакции на показательное возмущение, и в частности, частотная характеристика, не зависит от времени и представляет собой некоторую функцию параметра показательной функции s . Мы обозначим эту функцию через $\Phi(s)$. Тогда формула (2.4.5) примет вид

$$\Phi(s) = \int_0^{\infty} w(\xi) e^{-s\xi} d\xi. \quad (2.4.6)$$

В частном случае при $s = i\omega$ получаем частотную характеристику стационарной линейной системы

$$\Phi(i\omega) = \int_0^{\infty} w(\xi) e^{-i\omega\xi} d\xi. \quad (2.4.7)$$

Согласно определению характеристики реакции линейной системы на показательное возмущение установившаяся реакция стационарной линейной системы на показательное возмущение e^{st} представляет собой произведение

$$y(t) = \Phi(s) e^{st}, \quad (2.4.8)$$

Таким образом, показательные функции времени проходят через стационарную линейную систему, не изменяя своей формы,

а только умножаясь на постоянный коэффициент $\Phi(s)$, зависящий от параметра показательной функции. Функции, обладающие таким свойством, называются *инвариантными функциями* линейной системы. Формула (2.4.8) показывает, что показательные функции являются инвариантными функциями для любой стационарной линейной системы. Множитель, на который умножается показательная функция e^{st} , проходя через стационарную линейную систему, естественно назвать *передаточной функцией* системы. Вследствие того что семейство показательных функций является общим семейством инвариантных функций для всех стационарных линейных систем, передаточные функции особенно удобны для исследования стационарных линейных систем.

Заметим, что для некоторых классов нестационарных линейных систем тоже можно определить инвариантные функции, которые при прохождении через систему не изменяют своей формы, а лишь умножаются на соответствующие постоянные множители. Рассмотрим в качестве примера систему, поведение которой описывается линейным дифференциальным уравнением с переменными коэффициентами

$$\sum_{k=0}^n a_k(t) y^{(k)} = \sum_{k=0}^n b_k(t) x^{(k)}.$$

Положив $y = \lambda x$, где λ — постоянный коэффициент, получим следующее уравнение для определения инвариантной функции $x(t)$ рассматриваемой системы:

$$\sum_{k=0}^n [\lambda a_k(t) - b_k(t)] x^{(k)}(t) = 0.$$

По теореме существования решений дифференциальных уравнений это уравнение имеет решение (и не одно), если функция $\lambda a_n(t) - b_n(t)$ нигде не обращается в нуль на действительной оси. Изменяя параметр λ , мы получим семейство инвариантных функций рассматриваемой системы.

Однако, как легко понять, разные нестационарные линейные системы в общем случае имеют различные инвариантные функции, и пока еще не удалось найти достаточно широкого класса нестационарных систем, которые имели бы общее семейство инвариантных функций. Вследствие этого попытки распространения на нестационарные линейные системы простых алгебраических методов, основанных на использовании передаточных функций, которые оказались весьма плодотворными в теории стационарных линейных систем, пока еще не увенчались успехом.

Так как частотная характеристика стационарной линейной системы $\Phi(i\omega)$ не зависит от времени, то и ее амплитудная и фазовая частотные характеристики $|\Phi(i\omega)|$ и $\arg \Phi(i\omega)$ не зависят

от времени. Поэтому, как показывает формула (2.3.10), стационарная линейная система реагирует на бесконечно долго действующие гармонические колебания входной переменной с амплитудой a и фазой φ гармоническими колебаниями выходной переменной с той же частотой, амплитудой

$$b = a | \Phi(i\omega) | \quad (2.4.9)$$

и фазой

$$\psi = \varphi + \arg \Phi(i\omega). \quad (2.4.10)$$

Таким образом, амплитудная частотная характеристика стационарной линейной системы представляет собой коэффициент изменения амплитуды гармонических колебаний при прохождении их через данную систему, т. е. отношение амплитуды гармонических колебаний выходной переменной к амплитуде гармонических колебаний входной переменной. Фазовая частотная характеристика стационарной линейной системы представляет собой сдвиг фазы выходных гармонических колебаний по отношению к входным гармоническим колебаниям. При этом, если фазовая частотная характеристика $\arg \Phi(i\omega)$ положительна, то колебания на

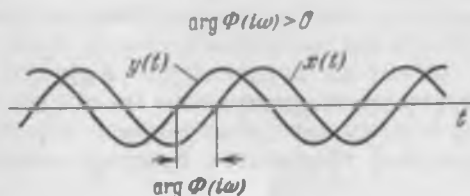


Рис. 2.4.1.

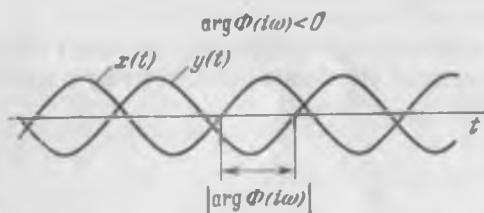


Рис. 2.4.2.

выходе системы опережают по фазе входные колебания на величину $\arg \Phi(i\omega)$ (рис. 2.4.1); если фазовая частотная характеристика отрицательна, то колебания на выходе системы отстают по фазе от входных колебаний на величину $|\arg \Phi(i\omega)|$ (рис. 2.4.2).

Найдем передаточные функции идеальных систем, осуществляющих основные операции анализа, рассмотренных в § 2.2.

1. *Идеальная следящая система* при входной функции e^{st} дает на выходе ту же функцию e^{st} . Следовательно, передаточная функция идеальной следящей системы тождественно равна единице:

$$\Phi_c(s) \equiv 1. \quad (2.4.11)$$

2. *Идеальный экстраполятор* при входной функции e^{st} дает на выходе $e^{s(t+a)}$. Следовательно, его передаточная функция равна

$$\Phi_e(s) = e^{sa} \quad (a > 0). \quad (2.4.12)$$

3. *Идеальное запаздывающее звено* при входной функции e^{st} дает на выходе $e^{s(t-l)}$. Следовательно, его передаточная функция равна

$$\Phi_z(s) = e^{-ls} \quad (l > 0). \quad (2.4.13)$$

Полагая в (2.4.12) и (2.4.13) $s = i\omega$, видим, что идеальные экстраполятор и запаздывающее звено не изменяют амплитуду гармонических колебаний. При этом экстраполятор дает опережение входных колебаний по фазе, пропорциональное частоте, а запаздывающее звено дает отставание от входных колебаний по фазе, также пропорциональное частоте.

4. *Безынерционный усилитель* с постоянным коэффициентом усиления k при входной функции e^{st} дает на выходе ke^{st} . Следовательно, его передаточная функция равна коэффициенту усиления k :

$$\Phi_y(s) = k. \quad (2.4.14)$$

Усилитель с переменным коэффициентом усиления, очевидно, не является стационарной системой. Предоставляем читателю самостоятельно найти для него характеристику реакции на показательную функцию $z(t, s)$.

5. *Идеальный дифференциатор* при входной функции e^{st} дает на выходе se^{st} . Следовательно, его передаточная функция равна

$$\Phi_d(s) = s. \quad (2.4.15)$$

Отсюда находим частотную характеристику:

$$\Phi_d(i\omega) = i\omega = \omega e^{i\frac{\pi}{2}}. \quad (2.4.16)$$

Таким образом, идеальный дифференциатор дает усиление гармонических колебаний, пропорциональное частоте, и опережение выходных колебаний по фазе $\pi/2$, независимое от частоты.

6. *Идеальный интегратор* при входной функции e^{st} при $s > 0$ дает на выходе $\frac{1}{s}e^{st}$. Следовательно, его передаточная функция равна

$$\Phi_{\pi}(s) = \frac{1}{s}. \quad (2.4.17)$$

Полагая здесь $s = i\omega$, находим частотную характеристику интегратора:

$$\Phi_{\pi}(i\omega) = \frac{1}{i\omega} = \frac{1}{\omega} e^{-i\frac{\pi}{2}}. \quad (2.4.18)$$

Отсюда видно, что идеальный интегратор дает усиление гармонических колебаний, обратно пропорциональное частоте, и отставание выходных колебаний по фазе $\pi/2$, независимое от частоты.

Пример 2.4.1. Формулы (2.2.15) и (2.2.18) показывают, что весовые функции двигателя, рассмотренного в примерах 2.2.1 и 2.2.2, зависят только от $t - \tau$. Следовательно, этот двигатель является стационарной линейной системой. Для определения его передаточной функции $\Phi(s)$ возьмем показательный входной сигнал $M = e^{st}$. Тогда на выходе будем иметь по определению $\omega = \Phi(s) e^{st}$. Подставляя эти выражения в (2.2.13) и решая полученное уравнение, найдем передаточную функцию двигателя:

$$\Phi(s) = \frac{1}{Js + h}. \quad (2.4.19)$$

Аналогично находим передаточную функцию двигателя $\Phi_1(s)$ для случая, когда за его выходной сигнал принимается угол поворота вала:

$$\Phi_1(s) = \frac{1}{s(Js + h)}. \quad (2.4.20)$$

Те же результаты получаются, если воспользоваться формулой (2.4.6), подставив в нее найденные в примерах 2.2.1 и 2.2.2 выражения весовых функций $\omega(t - \tau) = g(t, \tau)$ и $\omega_1(t - \tau) = g_1(t, \tau)$.

Пример 2.4.2. Формула (2.2.28) показывает, что рассмотренный в примере 2.2.3 нагреватель является стационарной линейной системой. Для определения его передаточной функции зададим входной сигнал $q(t) = e^{st}$ и будем искать решение уравнения (2.2.19) в виде $\theta(t, x) = \theta(x, s) e^{st}$. Подставив это выражение в (2.2.19), получим после сокращения на e^{st}

$$\theta(x, s)s = a^2 \theta_x''(x, s) \quad (a^2 = k/c\rho)$$

или

$$\theta_x''(x, s) - \frac{s}{a^2} \theta(x, s) = 0.$$

Общий интеграл этого уравнения выражается формулой

$$\theta(x, s) = c_1 e^{\sqrt{s} \frac{x}{a} + c_2 e^{-\sqrt{s} \frac{x}{a}}. \quad (2.4.21)$$

Граничные условия (2.2.20) при $q(t) = e^{st}$, $\theta(t, x) = \theta(x, s) e^{st}$ принимают вид

$$\theta_x'(0, s) = -\frac{1}{k}, \quad \theta_x'(l, s) = -\frac{h}{k} \theta(l, s).$$

Подставляя сюда выражение (2.4.21), получим уравнения для определения постоянных c_1 и c_2 :

$$\begin{aligned} \sqrt{s}(c_1 - c_2) &= -\frac{a}{k}, \\ \sqrt{s}(c_1 e^{\sqrt{s} \frac{l}{a}} - c_2 e^{-\sqrt{s} \frac{l}{a}}) &= -\frac{ha}{k}(c_1 e^{\sqrt{s} \frac{l}{a}} + c_2 e^{-\sqrt{s} \frac{l}{a}}). \end{aligned}$$

Решая эти уравнения, находим

$$c_1 = \frac{k \sqrt{s} - ah}{k\varphi(s) \sqrt{s}} e^{-\sqrt{s} \frac{l}{a}}, \quad c_2 = \frac{k \sqrt{s} + ah}{k\varphi(s) \sqrt{s}} e^{\sqrt{s} \frac{l}{a}},$$

где $\varphi(s)$ — функция, определяемая формулой (2.2.25) с заменой $i\omega$ на s . Подставляя полученные выражения c_1 и c_2 в (2.4.21), получим

$$\theta(x, s) = \frac{(k\sqrt{s} - ah)e^{\sqrt{s}\frac{x-l}{a}} + (k\sqrt{s} + ah)e^{-\sqrt{s}\frac{x-l}{a}}}{k\varphi(s)\sqrt{s}}.$$

Наконец, полагая $x = l$, находим передаточную функцию нагревателя:

$$\Phi(s) = \theta(l, s) = \frac{2}{\varphi(s)}. \quad (2.4.22)$$

Формулы (2.4.9) и (2.4.10) показывают, что коэффициент изменения амплитуды и сдвиг фазы колебаний при прохождении через систему в общем случае зависят от частоты колебаний ω .

Обычные механические и электромеханические системы (вообще любые системы, содержащие механические элементы) обладают инерцией, вследствие чего под действием колебаний очень высокой частоты такие системы практически остаются в покое. Грубо говоря, при большой частоте колебаний такая система не успевает еще начать движение в одном направлении, как входное возмущение уже начинает «толкать» ее в противоположную сторону. Это свойство реальных физических систем дает основание ввести чисто практическое понятие полосы пропускания системы. *Полосой пропускания* системы называется диапазон частот гармонических колебаний, «пропускаемых» системой, т. е. проходящих через систему с практически заметными колебаниями выходной переменной. Гармонические колебания частот, лежащих вне полосы пропускания системы, практически не проходят через систему, т. е. колебания выходной переменной для этих частот имеют столь малую амплитуду по сравнению с амплитудой входных колебаний, что их можно не принимать во внимание и с достаточной для практики точностью считать, что они отсутствуют.

Подчеркнем, что понятие полосы пропускания системы является чисто инженерным понятием и не имеет строгого математического определения. Физически любая система пропускает колебания всех частот. Однако для одних частот выходная переменная системы совершает заметные колебания, а для других частот выходная переменная колеблется с едва заметной амплитудой. Поэтому в зависимости от того, какой величиной амплитуды выходных колебаний по сравнению с амплитудой входных колебаний мы условимся пренебрегать, полоса пропускания одной и той же системы будет иметь различные значения. Обычно считают возможным пренебрегать амплитудой выходных колебаний, меньшей чем 5% амплитуды входных колебаний. В соответствии с этим полосой пропускания системы считают диапазон частот, за пределами которого амплитудная частотная характеристика системы меньше чем 0,05.

Частота ω_p , для которой амплитудная частотная характеристика имеет максимум, называется *резонансной частотой* системы,

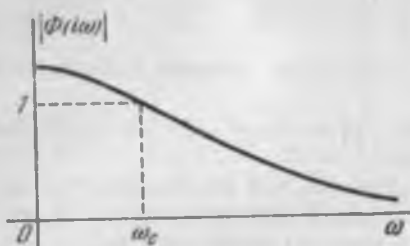


Рис. 2.4.3.

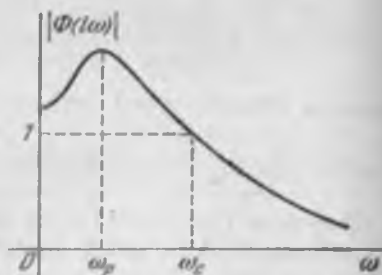


Рис. 2.4.4.

так как при этой частоте гармонические колебания, проходящие через систему, получают наибольшее усиление. В зависимости от

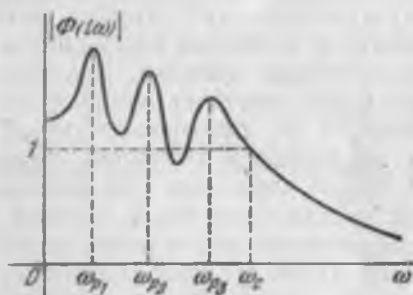


Рис. 2.4.5.

характера частотной характеристики стационарная линейная система может совсем не иметь резонансной частоты (рис. 2.4.3), иметь одну резонансную частоту (рис. 2.4.4) или несколько резонансных частот (рис. 2.4.5).

Частота ω_c , при которой амплитудная частотная характеристика, уменьшаясь, переходит от значений, больших единицы, к значениям, меньшим единицы, и при дальнейшем увеличении частоты остается

меньше единицы, называется *частотой среза* системы (рис. 2.4.3, 2.4.4 и 2.4.5).

Формула (2.3.20) дает следующее выражение весовой функции стационарной линейной системы через ее частотную характеристику:

$$w(t - \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(j\omega) e^{j\omega(t - \tau)} d\omega. \quad (2.4.23)$$

Из этой формулы следует, что не существует нестационарной линейной системы, для которой показательные функции были бы инвариантными, так как правая часть этой формулы зависит только от разности $t - \tau$, если первый множитель не зависит от t .

Формулу (2.4.23) можно переписать также в виде

$$w(t - \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \overline{\Phi(i\omega)} e^{-i\omega(t-\tau)} d\omega, \quad (2.4.24)$$

так как весовая функция $w(t - \tau)$ действительна и, следовательно, совпадает со своей сопряженной величиной.

Производя замену переменных $\xi = t - \tau$, можно переписать формулу (2.4.23) в виде

$$w(\xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(i\omega) e^{i\omega\xi} d\omega. \quad (2.4.25)$$

Пример 2.4.3. Полагая в формуле (2.2.28) примера 2.2.3 $g(t, \tau) = w(t - \tau)$, $\xi = t - \tau$ и сравнивая полученное равенство с (2.4.25), снова получим для передаточной функции нагревателя формулу (2.4.22).

Весовая функция $w(\xi)$ физически возможной стационарной линейной системы равна нулю при отрицательных значениях аргумента ξ . Формула (2.4.6) показывает, что передаточная функция стационарной линейной системы $\Phi(s)$ представляет собой преобразование Лапласа ее весовой функции. Пользуясь терминами операционного исчисления, можно сказать, что передаточная функция стационарной линейной системы представляет собой изображение ее весовой функции. Это обстоятельство дает возможность пользоваться для определения передаточных функций стационарных линейных систем по их весовым функциям и наоборот готовыми таблицами формул операционного исчисления и таким образом избежать вычисления интегралов (2.4.6) и (2.4.25)*).

При практическом применении частотных характеристик их часто изображают в логарифмическом масштабе, откладывая по оси абсцисс величину $\lg \omega$, а по оси ординат величину $\lg |\Phi(i\omega)|$ для амплитудной характеристики и величину $\arg \Phi(i\omega)$ для фазовой характеристики. Построенные таким образом частотные характеристики обычно называются *логарифмическими частотными характеристиками* системы.

Величина $\lg |\Phi(i\omega)|$ характеризует усиление амплитуды колебаний данной системой. Для измерения усиления пользуются особыми единицами. За единицу усиления принимают такое усиление, при котором мощность сигнала увеличивается в 10 раз. Эти единица называется *белом*. Таким образом, усиление системы равно 2 белам, если она усиливает мощность сигнала в 100 раз, 3 белам, если она усиливает мощность сигнала в 1000 раз и т. д. В соответствии с основным правилом десятичной системы для

* Читатель может найти таблицу преобразований Лапласа, например, в [17].

измерения усиления применяются также более крупные единицы: *декабелы*, *гектобелы*, *килобелы* и т. д. — и более мелкие единицы: *децибелы*, *сантибелы*, *миллибелы* и т. д. В автоматике и радиотехнике принято измерять усиление в *децибелах* (дБ).

Так как мощность сигнала пропорциональна квадрату его амплитуды (например, мощность электрического тока пропорциональна квадрату силы тока), то при

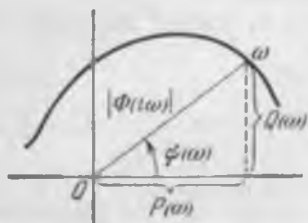


Рис. 2.4.6.

усилении амплитуды колебаний данной системой, равном 1 белу, величина $\lg |\Phi(i\omega)|^2 = 2 \lg |\Phi(i\omega)|$ равна единице, при усилении, равном 2 белам, эта величина равна 2 и т. д. Следовательно, усиление амплитуды колебаний данной системой, выраженное в белах, численно равно $2 \lg |\Phi(i\omega)|$, а усиление, выраженное в децибелах, численно равно $20 \lg |\Phi(i\omega)|$. Поэтому при практическом построении логариф-

мических частотных характеристик обычно откладывают по оси ординат прямо величину $20 \lg |\Phi(i\omega)|$, которая выражает усиление в децибелах.

Изменение частоты в 2 раза обычно называется, согласно принятой в акустике терминологии, изменением частоты на *октаву*. Изменение частоты в 10 раз называют изменением на *декаду*.

Величина

$$L(\omega) = 20 \lg |\Phi(i\omega)|, \quad (2.4.26)$$

рассматриваемая как функция величины $\lg \omega$, называется *логарифмической амплитудной частотной характеристикой* системы.

Величина

$$\psi(\omega) = \arg \Phi(i\omega), \quad (2.4.27)$$

рассматриваемая как функция величины $\lg \omega$, называется *логарифмической фазовой частотной характеристикой* системы.

В §§ 4.5 и 4.7 читателю станет ясно, почему логарифмические частотные характеристики особенно удобны для исследования стационарных линейных систем.

Комплексное число $\Phi(i\omega)$ при каждом данном значении частоты ω можно изобразить вектором на комплексной числовой плоскости (рис. 2.4.6). При изменении частоты ω конец этого вектора опишет кривую, которая называется *годографом* частотной характеристики или *амплитудно-фазовой* характеристикой системы.

Амплитудно-фазовая характеристика системы представляет собой графическое изображение зависимости между амплитудной частотной и фазовой частотной характеристиками системы $|\Phi(i\omega)|$ и $\psi(\omega)$ в полярной системе координат.

При использовании логарифмических частотных характеристик амплитудно-фазовую характеристику системы удобнее строить в декартовых координатах, откладывая по осям координат величины $L(\omega)$ в децибелах и $\varphi(\omega)$ в градусах (рис. 2.4.7).

Комплексное число $\Phi(i\omega)$ можно выразить также в обычной алгебраической форме:

$$\Phi(i\omega) = P(\omega) + iQ(\omega), \quad (2.4.28)$$

где $P(\omega)$ и $Q(\omega)$ — действительные функции частоты ω . Функция $P(\omega)$ называется *действительной частотной характеристикой*, а $Q(\omega)$ — *мнимой частотной характеристикой* системы.

Подставляя выражение (2.4.28) в (2.4.25) и пользуясь известной формулой Эйлера, получим

$$\begin{aligned} w(\xi) &= \frac{1}{2\pi} \int_{-\infty}^{\infty} [P(\omega) + iQ(\omega)] (\cos \omega\xi + i \sin \omega\xi) d\omega = \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} [P(\omega) \cos \omega\xi - Q(\omega) \sin \omega\xi] d\omega + \\ &+ \frac{i}{2\pi} \int_{-\infty}^{\infty} [P(\omega) \sin \omega\xi + Q(\omega) \cos \omega\xi] d\omega. \end{aligned} \quad (2.4.29)$$

Но весовая функция $w(\xi)$ действительна. Следовательно, ее мнимая часть равна нулю, и мы можем переписать полученную формулу в виде

$$w(\xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} [P(\omega) \cos \omega\xi - Q(\omega) \sin \omega\xi] d\omega \quad (2.4.30)$$

Заметим теперь, что для физически возможной системы $w(\xi) = 0$ при $\xi < 0$. Следовательно, полагая $\xi > 0$ и заменяя в (2.4.30) ξ на $-\xi$, можем написать

$$0 = \frac{1}{2\pi} \int_{-\infty}^{\infty} [P(\omega) \cos \omega\xi + Q(\omega) \sin \omega\xi] d\omega. \quad (2.4.31)$$

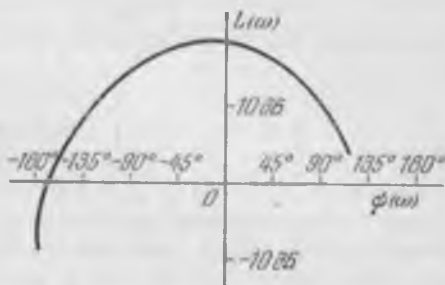


Рис. 2.4.7.

Складывая эту формулу с (2.4.30), получим

$$w(\xi) = \frac{1}{\pi} \int_{-\infty}^{\infty} P(\omega) \cos \omega \xi d\omega \quad (\xi > 0). \quad (2.4.32)$$

Совершенно таким же образом, вычитая (2.4.31) из (2.4.30), получим

$$w(\xi) = -\frac{1}{\pi} \int_{-\infty}^{\infty} Q(\omega) \sin \omega \xi d\omega \quad (\xi > 0). \quad (2.4.33)$$

Формулы (2.4.32) и (2.4.33) показывают, что физически возможная стационарная линейная система полностью характеризуется одной своей действительной частотной характеристикой или одной мнимой частотной характеристикой. Зная действительную или мнимую частотную характеристику такой системы, можно найти по формуле (2.4.32) или (2.4.33) ее весовую функцию, а следовательно, и любые другие характеристики.

Действительная и мнимая частотные характеристики системы связаны с ее амплитудной и фазовой частотными характеристиками очевидными соотношениями (рис. 2.4.6):

$$P(\omega) = |\Phi(i\omega)| \cos \psi(\omega), \quad Q(\omega) = |\Phi(i\omega)| \sin \psi(\omega), \quad (2.4.34)$$

$$|\Phi(i\omega)|^2 = P^2(\omega) + Q^2(\omega), \quad \psi(\omega) = \operatorname{arctg} \frac{Q(\omega)}{P(\omega)}. \quad (2.4.35)$$

При этом следует иметь в виду, что во второй формуле (2.4.35) следует взять значение арктангенса в первой четверти, если $P(\omega) > 0$, $Q(\omega) > 0$; во второй четверти, если $P(\omega) < 0$, $Q(\omega) > 0$; в третьей четверти, если $P(\omega) < 0$, $Q(\omega) < 0$, и в четвертой четверти, если $P(\omega) > 0$, $Q(\omega) < 0$. Это легко видно непосредственно из рис. 2.4.6.

Многомерная стационарная линейная система полностью характеризуется частотными характеристиками, определяющими прохождение гармонических колебаний от различных входов к различным выходам. При этом частотная характеристика многомерной системы, соответствующая h -му входу и k -му выходу, определяет изменение амплитуды (усиление) гармонических колебаний и сдвиг фазы при прохождении их от h -го входа к k -му выходу при отсутствии возмущений на всех остальных входах.

§ 2.5. Определение характеристик стационарных линейных систем, описываемых дифференциальными уравнениями

Рассмотрим подробнее стационарные линейные системы, поведение которых описывается обыкновенными дифференциальными уравнениями. Сначала докажем, что любая система, поведение которой описывается линейным дифференциальным уравнением

с постоянными коэффициентами, является стационарной линейной системой.

Пусть x — входная переменная системы, а y — ее выходная переменная, связанная с входной переменной x линейным дифференциальным уравнением с постоянными коэффициентами

$$\begin{aligned} a_n \frac{d^ny}{dt^n} + a_{n-1} \frac{d^{n-1}y}{dt^{n-1}} + \dots + a_1 \frac{dy}{dt} + a_0 y = \\ = b_m \frac{d^mx}{dt^m} + b_{m-1} \frac{d^{m-1}x}{dt^{m-1}} + \dots + b_1 \frac{dx}{dt} + b_0 x. \end{aligned} \quad (2.5.1)$$

Из теории дифференциальных уравнений известно, что в случае, когда правая часть линейного уравнения, равная в данном случае

$$f(t) = b_m \frac{d^mx}{dt^m} + b_{m-1} \frac{d^{m-1}x}{dt^{m-1}} + \dots + b_1 \frac{dx}{dt} + b_0 x, \quad (2.5.2)$$

представляет собой сумму нескольких слагаемых, интеграл этого уравнения равен сумме интегралов, соответствующих отдельным слагаемым в правой части. При умножении правой части на постоянную интеграл линейного дифференциального уравнения умножается на ту же постоянную. Это означает, что для системы, описываемой линейным дифференциальным уравнением, выполнены условия (1.5.2) и (1.5.3), т. е. справедлив принцип суперпозиции. Следовательно, любая система, поведение которой описывается линейными дифференциальными уравнениями (безразлично, с постоянными или переменными коэффициентами), линейна.

Далее, рассматриваемая система стационарна, так как вследствие постоянства коэффициентов уравнения (2.5.1) реакция этой системы на любое возмущение, изменяющееся по заданному закону, не зависит от момента начала действия этого возмущения.

Для определения передаточной функции системы, описываемой уравнением (2.5.1), следует, согласно определению, найти реакцию этой системы на возмущение, представляющее собой показательную функцию

$$x(t) = e^{st}. \quad (2.5.3)$$

Для этого необходимо проинтегрировать уравнение (2.5.1) для случая $x(t) = e^{st}$. Дифференцируя формулу (2.5.3), находим

$$\frac{dx}{dt} = se^{st}, \quad \frac{d^2x}{dt^2} = s^2e^{st}$$

и вообще

$$\frac{d^kx}{dt^k} = s^k e^{st} \quad (k = 1, \dots, m). \quad (2.5.4)$$

Подставляя полученные выражения в правую часть уравнения (2.5.1) и вынося за скобки показательную функцию, получим

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_1 \frac{dy}{dt} + a_0 y = (b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0) e^{st}. \quad (2.5.5)$$

Как известно из теории дифференциальных уравнений (см., например, [68]), если правая часть линейного дифференциального уравнения с постоянными коэффициентами представляет собой показательную функцию, умноженную на постоянную величину, то частный интеграл этого уравнения следует искать в виде произведения той же показательной функции на некоторую другую постоянную, если только параметр показательной функции s не является корнем характеристического уравнения

$$a_n v^n + a_{n-1} v^{n-1} + \dots + a_1 v + a_0 = 0. \quad (2.5.6)$$

Следовательно, предполагая это условие выполненным, мы можем искать частный интеграл уравнения (2.5.5) в виде

$$y = c e^{st}, \quad (2.5.7)$$

где c — неизвестная постоянная. Подставляя выражение (2.5.7) в уравнение (2.5.5) и имея в виду, что

$$\frac{d^k}{dt^k} e^{st} = s^k e^{st} \quad (k = 1, \dots, n), \quad (2.5.8)$$

получим

$$(a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0) c e^{st} = (b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0) e^{st}. \quad (2.5.9)$$

Сокращая это уравнение на показательную функцию и решая относительно c , получим

$$c = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0}. \quad (2.5.10)$$

Подставляя это выражение в (2.5.7), найдем частный интеграл уравнения (2.5.5):

$$y = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0} e^{st}. \quad (2.5.11)$$

Эта формула определяет реакцию рассматриваемой системы на показательное возмущение e^{st} . Для определения передаточной функции системы остается разделить выражение (2.5.11) на e^{st} . Тогда получим

$$\Phi(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0}. \quad (2.5.12)$$

Формула (2.5.11) показывает, что характеристика реакции системы, описываемой линейным дифференциальным уравнением с постоянными коэффициентами, на показательное возмущение e^{st} не зависит от времени t . Это еще раз доказывает, что система стационарна.

Изложенное показывает, что для определения передаточной функции стационарной линейной системы, описываемой обыкновенным линейным дифференциальным уравнением с постоянными коэффициентами, следует заменить в этом уравнении оператор дифференцирования d/dt параметром s , входное возмущение x единицей, а выходную переменную y неизвестной передаточной функцией $\Phi(s)$ и решить полученное уравнение относительно $\Phi(s)$.

Формула (2.5.12) показывает, что передаточная функция стационарной линейной системы, описываемой обыкновенным дифференциальным уравнением, представляет собой дробно-рациональную функцию параметра s . Корни характеристического уравнения (2.5.6) являются значениями параметра s , при которых знаменатель передаточной функции системы (2.5.12) обращается в нуль. Такие значения аргумента, при которых знаменатель дробно-рациональной функции обращается в нуль, называются *полюсами* этой функции. Таким образом, корни характеристического уравнения (2.5.6) являются полюсами передаточной функции системы, описываемой уравнением (2.5.1). При значениях s , являющихся полюсами передаточной функции, уравнение (2.5.5) не имеет решения вида (2.5.7).

Итак, мы доказали, что любая система, поведение которой описывается линейным дифференциальным уравнением с постоянными коэффициентами, линейна и стационарна, и нашли ее передаточную функцию.

Можно доказать и обратное положение: если поведение стационарной линейной системы описывается линейным дифференциальным уравнением, то все коэффициенты этого уравнения постоянны. Для доказательства достаточно заметить, что независимо от того, постоянны или переменны коэффициенты $a_0, a_1, \dots, a_n, b_0, b_1, \dots, b_m$ уравнения (2.5.1), уравнение (2.5.5) вследствие стационарности системы имеет решение вида $y = \Phi(s) e^{st}$, где $\Phi(s)$ — передаточная функция системы, не зависящая от времени t . Поэтому уравнение (2.5.5) будет удовлетворено, если положить в нем $y = \Phi(s) e^{st}$. Тогда, принимая во внимание (2.5.8), получим после сокращения на e^{st}

$$(a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0) \Phi(s) = b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0. \quad (2.5.13)$$

Решая это уравнение относительно $\Phi(s)$, мы снова получим формулу (2.5.12). Так как левая часть формулы (2.5.12) не зависит от времени t , то и правая ее часть не может зависеть от времени.

Следовательно, все коэффициенты $a_0, a_1, \dots, a_n, b_0, b_1, \dots, b_m$ либо постоянны, либо представляют собой произведения постоянных величин на одну и ту же функцию времени. В последнем случае исходное уравнение (2.5.1) можно сократить на эту функцию, после чего это уравнение будет уравнением с постоянными коэффициентами. Таким образом, мы доказали, что если поведение стационарной линейной системы описывается дифференциальным уравнением, то это уравнение представляет собой линейное дифференциальное уравнение с постоянными коэффициентами.

В дальнейшем нам будет удобно записывать дифференциальные уравнения с постоянными коэффициентами сокращенно в операторной форме. Для этого введем полиномы

$$\left. \begin{aligned} F(s) &= a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0 = \sum_{k=0}^n a_k s^k, \\ H(s) &= b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0 = \sum_{k=0}^m b_k s^k. \end{aligned} \right\} \quad (2.5.14)$$

Тогда дифференциальное уравнение (2.5.1) запишется коротко в операторной форме:

$$F(D) y = H(D) x, \quad (2.5.15)$$

где $D = d/dt$ — оператор дифференцирования по времени. Формула (2.5.12) для передаточной функции системы примет вид

$$\Phi(s) = \frac{H(s)}{F(s)}. \quad (2.5.16)$$

Для определения весовой функции системы, описываемой уравнением (2.5.1), достаточно выделить в выражении (2.5.12) передаточной функции целую часть (если $m > n$), разложить дробную часть на элементарные дроби и после этого воспользоваться таблицей преобразований Лапласа. В § 4.4 мы выведем общую формулу для весовой функции стационарной линейной системы другим методом.

Решая формально уравнение (2.5.15) относительно y , можем написать

$$y = \frac{H(D)}{F(D)} x \quad (2.5.17)$$

или, приписывая во внимание (2.5.16),

$$y = \Phi(D) x. \quad (2.5.18)$$

Эта формула показывает, что оператор стационарной линейной системы, поведение которой описывается линейным дифференциальным уравнением, можно символически выразить через ее передаточную функцию. Для этого достаточно заменить в выражении передаточной функции системы аргумент s оператором

дифференцирования по времени $D = d/dt$. Обобщая этот результат, можно записать в виде (2.5.18) уравнение любой стационарной линейной системы. Таким образом, оператор любой стационарной линейной системы можно выразить как функцию оператора дифференцирования по времени, заменив в выражении передаточной функции этой системы аргумент s оператором D .

Формула (2.5.12) формально определяет передаточную функцию стационарной линейной системы, поведение которой описывается дифференциальным уравнением (2.5.1), при всех значениях s , кроме совпадающих с корнями характеристического уравнения (2.5.6). Однако физически эта передаточная функция существует не при всех значениях s . Действительно, формула (2.5.11) определяет реакцию системы на показательное возмущение не при всех условиях. Эта реакция в общем случае представляет собой общий интеграл уравнения (2.5.5), а не частный. Для получения общего интеграла уравнения (2.5.5) следует к найденному частному интегралу (2.5.11) добавить общий интеграл соответствующего однородного уравнения

$$a_n \frac{d^n y}{dt^n} + a_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \dots + a_1 \frac{dy}{dt} + a_0 y = 0. \quad (2.5.19)$$

Из теории дифференциальных уравнений известно, что если корни характеристического уравнения (2.5.6) $v_1, \dots, v_n$ все различны, то общий интеграл уравнения (2.5.19) представляет собой линейную комбинацию показательных функций $e^{v_1 t}, \dots, e^{v_n t}$ с произвольными коэффициентами. Если некоторые корни характеристического уравнения совпадают, например $v_k = v_{k+1} = \dots = v_{k+l}$, то функции $e^{v_{k+1} t}, \dots, e^{v_{k+l} t}$ заменяются функциями $t e^{v_k t}, \dots, t^l e^{v_k t}$. Таким образом, в случае, когда характеристическое уравнение не имеет кратных корней, общий интеграл уравнения (2.5.5) определяется формулой

$$y = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0} e^{st} + c_1 e^{v_1 t} + \dots + c_n e^{v_n t}, \quad (2.5.20)$$

где $c_1, \dots, c_n$ — произвольные постоянные. Отсюда следует, что установившаяся реакция рассматриваемой системы на показательное возмущение e^{st} , независимая от начальных условий, существует только в том случае, когда все корни характеристического уравнения $v_1, \dots, v_n$ имеют отрицательные действительные части, а действительная часть параметра s отрицательна или равна нулю. Если это условие не выполнено, то реакция системы на возмущение e^{st} неограниченно возрастает. Однако и в этом случае можно говорить об установившейся реакции системы на возмущение e^{st} , если первое слагаемое в правой части (2.5.20) растет при $t \rightarrow \infty$ быстрее, чем все остальные слагаемые. В этом случае при достаточно длительном времени работы системы t реакция ее на

возмущение e^{st} будет практически выражаться одним первым слагаемым в правой части формулы (2.5.20).

Для определения передаточной функции системы разделим формулу (2.5.20) на e^{st} :

$$\frac{y}{e^{st}} = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0} + c_1 e^{(v_1-s)t} + \dots + c_n e^{(v_n-s)t}. \quad (2.5.21)$$

Если действительные части всех разностей $v_1 - s, \dots, v_n - s$ отрицательны, то все показательные функции в (2.5.21) стремятся к нулю при $t \rightarrow \infty$. В этом случае передаточная функция системы, представляющая собой отношение реакции этой системы на бесконечно долго действующее на нее возмущение e^{st} к самому возмущению e^{st} , определяется формулой (2.5.12). Если хотя бы одно из чисел $v_1 - s, \dots, v_n - s$ имеет положительную действительную часть, то соответствующая показательная функция в (2.5.21) неограниченно возрастает при $t \rightarrow \infty$, и, следовательно, передаточная функция системы не существует.

Таким образом, передаточная функция стационарной линейной системы, поведение которой описывается дифференциальным уравнением (2.5.1), существует только в области значений s , действительные части которых больше действительных частей всех корней характеристического уравнения $v_1, \dots, v_n$. Иными словами, передаточная функция этой системы существует только в полуплоскости комплексного параметра s , расположенной справа от вертикальной прямой, проходящей через корень характеристического уравнения с наибольшей действительной частью (эта полуплоскость заштрихована на рис. 2.5.1). Левее этой прямой и на самой прямой передаточная функция не существует, несмотря на то что формула (2.5.12) формально определяет ее при всех значениях s , кроме точек $v_1, \dots, v_n$. Этот вывод справедлив и в случае кратных корней характеристического уравнения, так как при любом $r > 0$ произведение $t^r e^{(v_k-s)t}$ стремится к нулю при $t \rightarrow \infty$, если действительная часть параметра s больше действительной части корня v_k характеристического уравнения, и неограниченно возрастает, если действительная часть параметра s меньше или равна действительной части корня v_k .

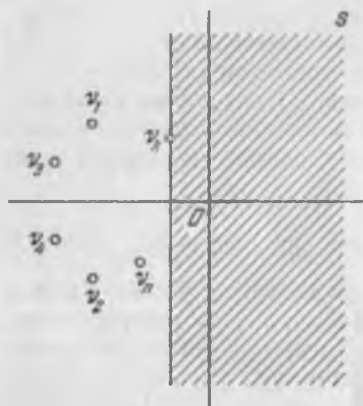


Рис. 2.5.1.

Таким образом, передаточная функция стационарной линейной системы, поведение которой описывается дифференциальным уравнением (2.5.1), существует только в области значений s , действительные части которых больше действительных частей всех корней характеристического уравнения $v_1, \dots, v_n$. Иными словами, передаточная функция этой системы существует только в полуплоскости комплексного параметра s , расположенной справа от вертикальной прямой, проходящей через корень характеристического уравнения с наибольшей действительной частью (эта полуплоскость заштрихована на рис. 2.5.1). Левее этой прямой и на самой прямой передаточная функция не существует, несмотря на то что формула (2.5.12) формально определяет ее при всех значениях s , кроме точек $v_1, \dots, v_n$. Этот вывод справедлив и в случае кратных корней характеристического уравнения, так как при любом $r > 0$ произведение $t^r e^{(v_k-s)t}$ стремится к нулю при $t \rightarrow \infty$, если действительная часть параметра s больше действительной части корня v_k характеристического уравнения, и неограниченно возрастает, если действительная часть параметра s меньше или равна действительной части корня v_k .

Из сказанного ясно, что частотная характеристика, представляющая собой значение передаточной функции на мнимой оси $s = i\omega$, существует для рассматриваемой системы только в том случае, когда действительные части всех корней характеристического уравнения отрицательны. В § 6.2 мы увидим, что это условие необходимо и достаточно для устойчивости системы. Таким образом, в соответствии с общим утверждением в конце § 2.3, частотная характеристика физически существует только для устойчивой системы, описываемой уравнением (2.5.1). При этом формула (2.5.12) при $s = i\omega$ формально определяет частотную характеристику для любой системы, поведение которой описывается дифференциальным уравнением (2.5.1).

§ 2.6. Элементарные стационарные линейные звенья

В современной теории линейных систем большую роль играют восемь типов простейших систем, обычно называемых *элементарными звеньями*. В §§ 2.2 и 2.4 мы уже познакомились с четырьмя типами элементарных звеньев, а именно с усилителем, дифференциатором, интегратором и запаздывающим звеном *). В § 2.4 мы видели, что усилитель с постоянным коэффициентом усиления k , дифференциатор, интегратор и запаздывающее звено являются стационарными линейными системами, передаточные функции которых определяются формулами (2.4.14), (2.4.15), (2.4.17) и (2.4.13). Изменив обозначения в (2.4.13), мы перепишем эти формулы в виде

$$\Phi_y(s) = k, \quad \Phi_d(s) = s, \quad \Phi_{\text{и}}(s) = \frac{1}{s}, \quad \Phi_z(s) = e^{-\tau s}, \quad (2.6.1)$$

где τ — время запаздывания.

Кроме этих уже знакомых нам типов элементарных звеньев, рассматривают еще четыре типа элементарных стационарных звеньев, передаточные функции которых определяются формулами

$$\left. \begin{aligned} \Phi_{\phi_1}(s) &= k(Ts + 1), & \Phi_{\phi_2}(s) &= k(T^2s^2 + 2\zeta Ts + 1), \\ \Phi_a(s) &= \frac{k}{Ts + 1}, & \Phi_k(s) &= \frac{k}{T^2s^2 + 2\zeta Ts + 1}. \end{aligned} \right\} \quad (2.6.2)$$

Звено, передаточные функции которых определяются первыми двумя формулами (2.6.2), называются соответственно *форсирующим звеном первого порядка* и *форсирующим звеном второго порядка*. Звено, передаточная функция которого определяется третьей формулой (2.6.2), называется *апериодическим звеном* или

*) Для краткости мы отбрасываем прилагательные «идеальный» и «безынерционный», которыми пользовались в §§ 2.2 и 2.4, чтобы подчеркнуть, что эти элементарные звенья выполняют соответствующие операции идеально точно.

инерционным звеном первого порядка. Звено, передаточная функция которого определяется последней формулой (2.6.2) при $|\zeta| < 1$, называется *колебательным звеном*. Условие $|\zeta| < 1$ необходимо для того, чтобы квадратный трехчлен $T^2s^2 + 2\zeta Ts + 1$ имел комплексные корни.

Из (2.6.1) и (2.6.2) следует, что передаточные функции дифференциатора и форсирующих звеньев первого и второго порядка при $k = 1$ являются обратными величинами по отношению соответственно к передаточным функциям интегратора, аperiodического

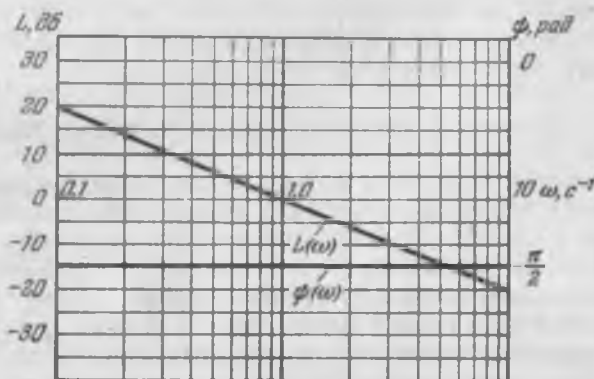


Рис. 2.6.1.

и запаздывающего звеньев. Вследствие этого логарифмические амплитудные и фазовые частотные характеристики первых трех звеньев только знаками отличаются от соответствующих характеристик остальных трех звеньев. Поэтому достаточно изучить только три последних звена и запаздывающее звено.

Имея в виду, что наиболее удобным практическим методом исследования стационарных линейных систем является метод логарифмических частотных характеристик, найдем амплитудно-фазовые и логарифмические частотные характеристики интегратора, аperiodического, колебательного и запаздывающего звеньев.

Согласно формуле (2.4.18) амплитудно-фазовой характеристикой идеального интегратора является мнимая ось. Логарифмические амплитудная и фазовая частотные характеристики определяются формулами

$$L_{\text{и}}(\omega) = -20 \lg \omega, \quad \Psi_{\text{и}}(\omega) = -\frac{\pi}{2}. \quad (2.6.3)$$

Таким образом, логарифмическая амплитудная характеристика интегратора представляет собой прямую с отрицательным наклоном 20 дБ/дек, пересекающую ось абсцисс при частоте $\omega = 1$, а фазовая характеристика — прямую, параллельную оси абсцисс (рис. 2.6.1).

Перейдем к аperiodическому звену. Величины k и T в третьей формуле (2.6.2) называются соответственно *коэффициентом усиления* и *постоянной времени* аperiodического звена.

Для определения частотной характеристики аperiodического звена положим в третьей формуле (2.6.2) $s = i\omega$. Тогда получим

$$\begin{aligned} \Phi_a(i\omega) &= \frac{k}{Ti\omega + 1} = k \frac{1 - Ti\omega}{1 + T^2\omega^2} = \\ &= \frac{k}{\sqrt{1 + T^2\omega^2}} \exp\{-i \operatorname{arctg} T\omega\}. \end{aligned} \quad (2.6.4)$$

Отсюда находим амплитудную и фазовую частотные характеристики аperiodического звена

$$A_a(\omega) = \frac{k}{\sqrt{1 + T^2\omega^2}}, \quad \psi_a(\omega) = -\operatorname{arctg} T\omega \quad (2.6.5)$$

и его логарифмическую амплитудную частотную характеристику:

$$L_a(\omega) = 20 \lg k - 20 \lg \sqrt{1 + T^2\omega^2}. \quad (2.6.6)$$

На рис. 2.6.2 представлена половина амплитудно-фазовой характеристики аperiodического звена при $T > 0$, $k > 0$, соответствующая области положительных частот. Она представляет собой полуокружность радиуса $k/2$ с центром на действительной оси, оканчивающуюся при $\omega = \infty$ в начале координат.

Из (2.6.6) видно, что при очень малых частотах

$$L_a(\omega) \approx 20 \lg k, \quad (2.6.7)$$

а при очень больших частотах

$$L_a(\omega) \approx 20 \lg k - 20 \lg |T| \omega. \quad (2.6.8)$$

Это значит, что логарифмическая амплитудная характеристика имеет асимптотами прямую, параллельную оси абсцисс, проходящую выше нее на $20 \lg k$, и прямую (2.6.8), имеющую наклон -20 дБ/дек и пересекающую первую асимптоту при частоте $\omega = 1/|T|$, которая называется *сопрягающей частотой* звена. Прямая (2.6.7) является асимптотой логарифмической амплитудной характеристики в области низких частот, а прямая (2.6.8) — асимптотой в области высоких частот.

На рис. 2.6.3 представлены логарифмические частотные характеристики аperiodического звена при $k = 1$, $T = 1$.

Легко видеть, что логарифмическая амплитудная частотная характеристика аperiodического звена может быть приближенно представлена в виде ломаной, состоящей из отрезков двух

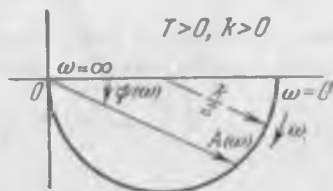


Рис. 2.6.2.

пересекающихся при $\omega = 1/|T|$ асимптот. Максимальная ошибка, равная $20 \lg \sqrt{2} \approx 3$ дБ, получится при $\omega = 1/|T|$.

Из (2.6.5) видно, что амплитудная характеристика аperiodического звена не зависит от знака постоянной времени T , а фазовая характеристика имеет знак, противоположный знаку T .

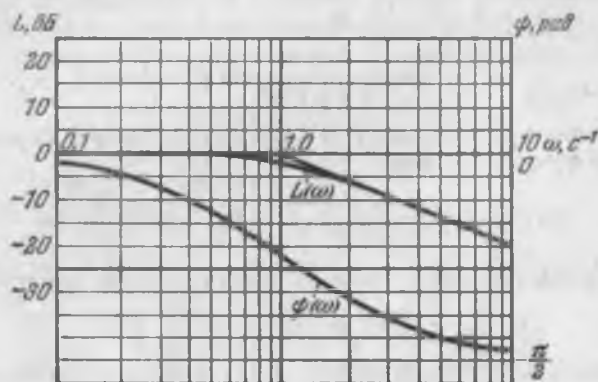


Рис. 2.6.3.

Перейдем к колебательному звену. Величины k , T и ζ в последней формуле (2.6.2), определяющей передаточную функцию колебательного звена, называются соответственно *коэффициентом усиления*, *постоянной времени* и *коэффициентом колебательности* (или *декрементом затухания*).

Для определения частотной характеристики колебательного звена положим в последней формуле (2.6.2) $s = i\omega$:

$$\Phi_K(i\omega) = \frac{k}{1 + 2T\zeta i\omega - T^2\omega^2} = k \frac{1 - T^2\omega^2 - 2T\zeta i\omega}{(1 - T^2\omega^2)^2 + 4\zeta^2 T^2\omega^2}. \quad (2.6.9)$$

Отсюда находим амплитудную и фазовую частотные характеристики колебательного звена

$$\left. \begin{aligned} A_K(\omega) &= \frac{k}{\sqrt{(1 - T^2\omega^2)^2 + 4\zeta^2 T^2\omega^2}}, \\ \psi_K(\omega) &= -\arctg \frac{2\zeta T\omega}{1 - T^2\omega^2} \end{aligned} \right\} \quad (2.6.10)$$

и его логарифмическую амплитудную частотную характеристику

$$L_K(\omega) = 20 \lg k - 20 \lg \sqrt{(1 - T^2\omega^2)^2 + 4\zeta^2 T^2\omega^2}. \quad (2.6.11)$$

На рис. 2.6.4 представлена амплитудно-фазовая характеристика колебательного звена при $\zeta > 0$, $k > 0$ (величина T для колебательного звена всегда считается положительной).

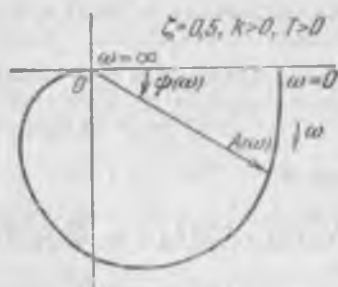


Рис. 2.6.4.

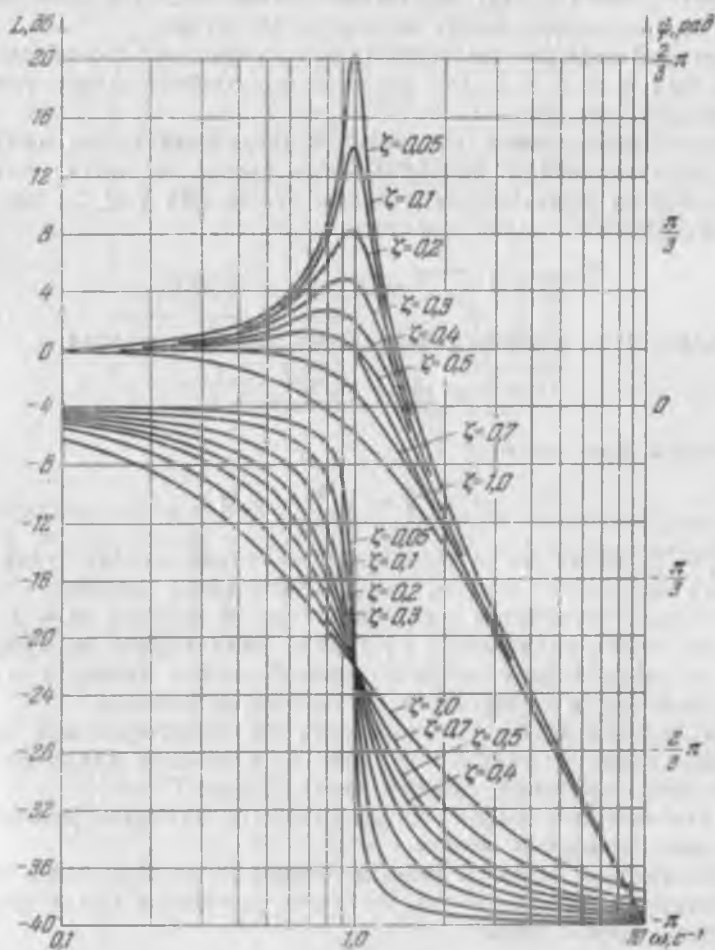


Рис. 2.6.5.

Из формулы (2.6.11) видно, что при очень малых ω

$$L_k(\omega) \approx 20 \lg k, \quad (2.6.12)$$

а при очень больших ω

$$L_k(\omega) \approx 20 \lg k - 40 \lg T\omega. \quad (2.6.13)$$

Отсюда следует, что логарифмическая амплитудная характеристика колебательного звена имеет две асимптоты, пересекающиеся при частоте $\omega = 1/T$, которая называется, как и в случае апериодического звена, *сопрягающей частотой*. При этом высокочастотная асимптота имеет наклон -40 дБ/дек.

Логарифмические частотные характеристики колебательного звена при $k = T = 1$ для различных положительных значений ζ приведены на рис. 2.6.5.

Если приблизительно представить логарифмическую амплитудную характеристику колебательного звена в виде ломаной, состоящей из отрезков двух асимптот, то при $\zeta < \zeta_0$, где ζ_0 — положительный корень уравнения

$$4\zeta^2 \sqrt{1 - \zeta^2} = 1 \quad (\zeta_0 \approx 0,547),$$

наибольшая по абсолютной величине ошибка, равная

$$20 \lg 2|\zeta| \sqrt{1 - \zeta^2},$$

получится при частотах

$$\omega = (1 - 2\zeta^2)^{\pm 1/2} / T;$$

при $\zeta = \zeta_0$ такая же по абсолютной величине ошибка будет и при частоте $\omega = 1/T$; при $\zeta > \zeta_0$ наибольшая ошибка, равная $20 \lg 2|\zeta|$, получится при сопрягающей частоте $\omega = 1/T$.

При малых значениях $|\zeta|$ замена амплитудной логарифмической характеристики асимптотической может привести к большим ошибкам в окрестности сопрягающей частоты.

Из (2.6.10) видно, что амплитудная характеристика колебательного звена не зависит от знака ζ , а фазовая характеристика имеет знак, противоположный знаку ζ (при $T > 0$).

Колебательное звено, у которого $\zeta = 0$, называют *резонансным* или *консервативным звеном*.

Рассмотрим, наконец, запаздывающее звено. Полагая в последней формуле (2.6.1) $s = i\omega$, получим частотную характеристику запаздывающего звена:

$$\Phi_s(i\omega) = e^{-i\omega\tau}. \quad (2.6.14)$$

Отсюда находим амплитудную и фазовую частотные характеристики запаздывающего звена:

$$A_s(\omega) = 1, \quad \psi_s(\omega) = -\tau\omega = -\tau e^{j\ln \omega} = -\tau e^{2,303 \lg \omega}. \quad (2.6.15)$$

Таким образом, амплитудно-фазовая характеристика запаздывающего звена представляет собой окружность единичного радиуса, центром которой служит начало координат.

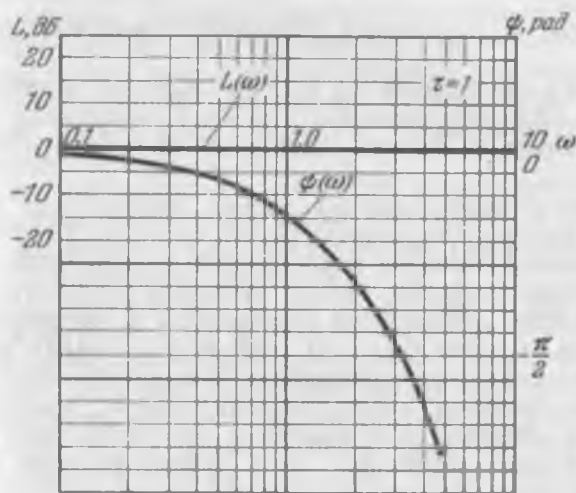


Рис. 2.6.6.

Логарифмические амплитудная и фазовая частотные характеристики запаздывающего звена при $\tau = 1$ представлены на рис. 2.6.6.

Глава 3

ЛИНЕЙНЫЕ ЭЛЕМЕНТЫ АВТОМАТИЧЕСКИХ СИСТЕМ

§ 3.1. Состав и назначение элементов автоматических систем

Как уже было сказано в § 1.2, информация, необходимая для управления объектом, вводится в систему управления с помощью соответствующих *измерителей*, которые часто называются также *датчиками*.

В автоматических системах удобно использовать в качестве носителей информации электрические сигналы. Поэтому выходными сигналами датчиков должны быть электрические сигналы. *Электрическим датчиком* принято называть устройство, преобразующее линейное или угловое перемещение в электрический сигнал. Однако входным сигналом системы управления может быть не только механическое перемещение. Для преобразования любого входного воздействия в электрический сигнал применяют чувствительные элементы или сочетания чувствительных элементов с электрическими датчиками. Под *чувствительным элементом* понимается такой элемент, который тем или иным образом реагирует на данное возмущение. Электрический датчик является чувствительным элементом, но не всякий чувствительный элемент является электрическим датчиком. Если данная физическая величина, например давление воздуха, не может быть достаточно просто непосредственно преобразована в электрическую величину, то выбирается такой чувствительный элемент, который преобразует эту величину в перемещение, чтобы с помощью электрического датчика преобразовать это перемещение в электрический сигнал. Например, давление воздуха может быть преобразовано в перемещение с помощью упругой спиралеобразной трубки с запаянным концом. Трубка в этом случае является чувствительным элементом. Обычно чувствительный элемент конструктивно выполняется совместно с электрическим датчиком и все устройство называется датчиком величины, которая действует на вход. Например, комбинация спиральной трубки с электрическим датчиком называется датчиком давления.

В автоматических системах наиболее широкое применение нашли потенциометрические, индуктивные, тахометрические и гироскопические датчики, фотоэлектрические, термоэлектрические, пьезоэлектрические, инерционные и пружинные чувствительные элементы.

Исполнительным устройством автоматической системы управления служит двигатель, перемещающий орган управления (например, руль самолета или ракеты) или сам объект управления (например, подвижную артиллерийскую установку на самолете) в соответствии с сигналом управления. Исполнительное устройство часто называют также *сервомотором* или *серводвигателем*. Исполнительное устройство преобразует управляющий сигнал в механическое перемещение органа управления или самого управляемого объекта. При этом используется либо энергия самого управляющего сигнала, либо энергия дополнительного источника. В последнем случае работа исполнительного устройства определяется положением так называемого *управляющего элемента*. Перемещение управляющего элемента осуществляется энергией управляющего сигнала.

В качестве исполнительных устройств в автоматических системах используются электродвигатели постоянного и переменного тока, электродвигатели с электромагнитными муфтами, гидравлические и пневматические двигатели.

Мощность сигналов, необходимых для приведения в действие исполнительных устройств, обычно значительно превышает мощность, получаемую на выходе датчиков и функциональных преобразователей. Поэтому для обеспечения работы автоматических систем необходимы промежуточные элементы — усилители.

Усилителем будем называть устройство, преобразующее энергию постороннего источника в выходной сигнал, который превосходит по мощности входной сигнал, изменяется в соответствии с входным сигналом и имеет одинаковую с ним физическую природу. В принципе усиливаться могут электрические (ток, напряжение) и механические (перемещение, давление) сигналы. В зависимости от сигнала различают усилители тока, напряжения, перемещения, давления и т. д. Выбор типа и схемы усилителя определяется типом датчика входного сигнала, а также типом и мощностью исполнительного органа. Если датчики сигналов электрические, то, естественно, применяются усилители тока или усилители напряжения. И в том и в другом случае входной сигнал будет усиливаться по мощности. Но в первом случае выходным сигналом будет ток, а во втором — напряжение. Усилители напряжения используются при большом входном сопротивлении следующего за усилителем элемента системы, а усилители тока — при малом.

В автоматических системах используются магнитные, электромашинные, релейные, ламповые, полупроводниковые и другие усилители.

Если датчик формирует сигнал в виде постоянного тока и исполнительным устройством является двигатель постоянного тока, то усиление может производиться усилителем мощности без специ-

ального преобразования сигнала. Если же датчик формирует сигнал в виде постоянного тока, а исполнительным устройством является двигатель переменного тока, то наряду с усилением необходимо преобразование постоянного тока в переменный. Амплитуда полученного в результате преобразования переменного тока должна быть пропорциональна величине постоянного тока, а фаза переменного тока должна определяться знаком (полярностью) постоянного тока. Устройство, выполняющее это преобразование, называется *модулятором*. В случае, когда датчик формирует сигнал в виде переменного тока, а исполнительное устройство работает на постоянном токе, необходимо преобразование переменного тока в постоянный. При этом величина постоянного тока должна быть пропорциональна амплитуде переменного тока, а знак постоянного тока должен определяться фазой переменного тока. Устройства, выполняющие эти функции, называются *демодуляторами* или *фазовыми дискриминаторами*. Обычно применяются ламповые, полупроводниковые и электромеханические модуляторы и демодуляторы.

Для работы некоторых схем фазовых дискриминаторов и в ряде других случаев необходимы одинаковые по амплитуде, но противоположные по фазе сигналы переменного тока. Для получения таких сигналов в автоматических системах используются *фазоинверторы*. Фазоинверторы выполняются на электронных лампах и на полупроводниках.

В автоматических системах часто по конструктивным или другим соображениям элементы отстоят на значительных расстояниях друг от друга. В связи с этим возникает необходимость дистанционной передачи сигналов. Передача электрических сигналов может быть осуществлена по проводам или по радио. Для передачи сигналов, представляющих собой механические перемещения, применяют специальные электрические дистанционные передачи, в которых механическое перемещение преобразуется в электрический сигнал, подвергающийся после передачи по проводам или по радио обратному преобразованию в механическое перемещение.

Для формирования управляющих сигналов по сигналам датчиков применяются различные функциональные преобразователи, входящие в состав корректирующих цепей и других устройств автоматических систем. В качестве функциональных преобразователей часто применяются интегрирующие, дифференцирующие и суммирующие операционные усилители, стационарные и нестационарные электрические цепи. В сложных системах роль функциональных преобразователей могут выполнять вычислительные машины как непрерывного действия, так и цифровые.

На основании изложенного в § 1.1 элементами автоматических систем являются и объекты управления.

§ 3.2. Потенциометрические датчики

Потенциометрический датчик (потенциометр) предназначен для преобразования перемещения в напряжение постоянного или переменного тока.

Простейший потенциометр (рис. 3.2.1) представляет собой активное сопротивление, выполненное в виде непрерывно намотанной на каркас высокоомной проволоки, к которому прикладывается переменное или постоянное напряжение u_0 . Входной сигнал (перемещение) задается с помощью контактной щетки (движок), которая может перемещаться по потенциометру. Выходной сигнал (напряжение) снимается с движка потенциометра.

При непрерывной намотке проволоки напряжение u_0 , подводимое к датчику, распределяется равномерно по длине потенциометра. Поэтому при разомкнутой выходной цепи (сопротивление нагрузки $R_n = \infty$), т. е. в режиме холостого хода, потенциал щетки пропорционален перемещению x :

$$u_x = \frac{u_0}{l} x, \quad (3.2.1)$$

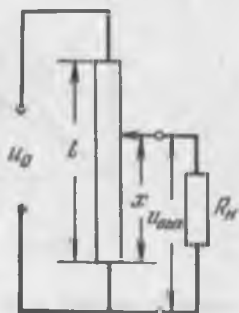


Рис. 3.2.1.

где l — длина потенциометра, x — перемещение движка. При включенной нагрузке ($R_n \neq \infty$) напряжение, снимаемое щеткой потенциометра, не будет равно u_x . В самом деле, при включенной нагрузке полное сопротивление цепи равно

$$R = R_0 - R_x + \frac{R_x R_n}{R_x + R_n} \quad (3.2.2)$$

где R_0 — сопротивление всего потенциометра, $R_x = R_0 x/l$ — сопротивление потенциометра, соответствующее перемещению x . Ток i в той части потенциометра, которая не шунтирована нагрузкой, равен u_0/R , а падение напряжения на этой части потенциометра равно $i(R_0 - R_x)$. Следовательно, напряжение, снимаемое с движка потенциометра, определяется формулой

$$u_{\text{вых}} = u_0 - i(R_0 - R_x) = \frac{u_0 R_x}{R_0 + \frac{R_x R_0}{R_n} - \frac{R_x^2}{R_n}}$$

или

$$u_{\text{вых}} = \frac{u_0 x}{l + \frac{R_0}{R_n} x - \frac{R_0 x^2}{l R_n}} \quad (3.2.3)$$

Таким образом, зависимость выходного сигнала потенциометра $u_{\text{вых}}$ от входного сигнала x в общем случае нелинейна. На

рис. 3.2.2 графически изображены законы изменения напряжений u_x и $u_{вых}$ в зависимости от $\xi = R_x/R_0 = x/l$. При одном и том же значении ξ разность напряжений $u_x - u_{вых}$ будет зависеть от отношения $R_0/R_n = \alpha$. Чем меньше α , тем ближе кривая $u_{вых}$ к прямой u_x . Практически при $\alpha = 0,1 \div 0,01$, т. е. когда сопротивление нагрузки превышает сопротивление потенциометра в $10 \div 100$ раз, отклонение закона изменения $u_{вых}$ от линейного

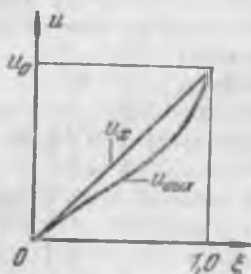


Рис. 3.2.2.

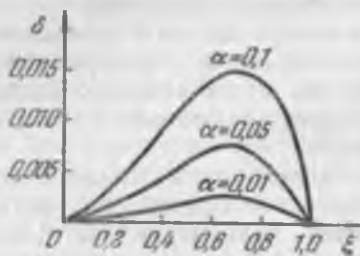


Рис. 3.2.3.

пренебрежимо мало, и потенциометр можно считать линейным элементом. На рис. 3.2.3 представлены графики, показывающие зависимость относительного отклонения

$$\frac{u_x - u_{вых}}{u_0} = \delta \quad (3.2.4)$$

от ξ . Из этих графиков видно, что при $\alpha = 0,1$ максимальное отклонение составляет примерно 1,5% от u_0 . Относительное отклонение δ имеет максимальное значение при $x \approx 2l/3$. Таким образом, для того чтобы характеристика потенциометра была близка к линейной, сопротивление нагрузки должно быть достаточно велико по сравнению с сопротивлением потенциометра. В автоматических системах это условие нетрудно выполнить.

Учитывая сделанные замечания, заменим точную зависимость (3.2.3) выходного напряжения потенциометра $u_{вых}$ от перемещения движка x приближенной линейной зависимостью

$$u_{вых} = kx. \quad (3.2.5)$$

Следовательно, потенциометрический датчик, как звено автоматической системы, можно приближенно считать идеальным линейным усилителем с коэффициентом усиления $k = u_0/l$. Однако по функциям, выполняемым в автоматической системе, потенциометр является датчиком, но не усилителем, так как физическая природа его входного и выходного сигналов неодинакова.

Следует иметь в виду, что, принимая формулу (3.2.5) в качестве динамической характеристики потенциометрического датчика, не-

обходимо нагрузку рассматривать как самостоятельный динамический элемент, который может быть как безынерционным, так и инерционным.

Рассмотренный простейший потенциометр является однополярным. Перемещение x и напряжению $u_{\text{вых}}$ могут быть только одного знака. Для преобразования в напряжение перемещения, которое может быть как положительным, так и отрицательным,

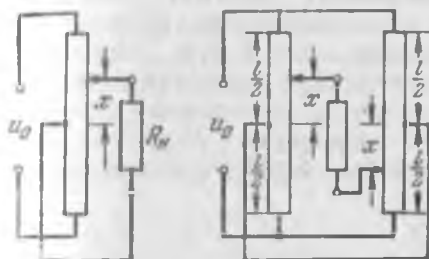


Рис. 3.2.4.

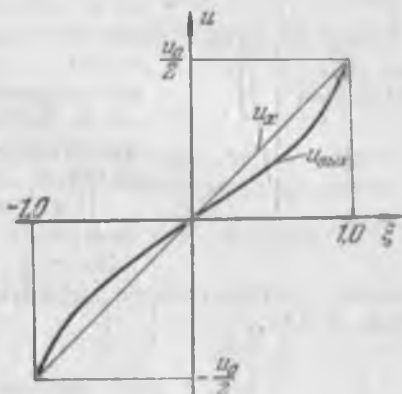


Рис. 3.2.5.

необходимы потенциометрические датчики, у которых знак выходного напряжения изменяется при изменении знака входного сигнала. К таким датчикам относятся двухтактные потенциометры, показанные на рис. 3.2.4.

В первом датчике на рис. 3.2.4 выходное напряжение снимается с движка и средней точки потенциометра, которая принимается за начало отсчета перемещения движка. Если потенциометр питается постоянным током, то при прохождении движком средней точки потенциометра знак выходного напряжения изменяется. Если потенциометр питается переменным током, то при прохождении средней точки фаза выходного напряжения изменяется на 180° . Амплитуда же выходного напряжения пропорциональна перемещению x .

Во втором датчике на рис. 3.2.4 выходное напряжение снимается с двух движков, перемещающихся симметрично относительно средних точек двух потенциометров. Очевидно, что при данном перемещении x разность потенциалов между движками у второго датчика на рис. 3.2.4 в два раза больше разности потенциалов между движком и средней точкой у первого датчика. Следовательно, коэффициент усиления второго датчика в два раза больше коэффициента усиления первого датчика.

На рис. 3.2.5 представлены характеристики двухтактного потенциометрического датчика. При $\alpha \neq 0$ действительные характеристики датчиков не будут совпадать с характеристикой холостого

хода, соответствующей $\alpha = 0$. При данном значении α отклонение характеристики от линейной у второго датчика на рис. 3.2.4 меньше, чем у первого.

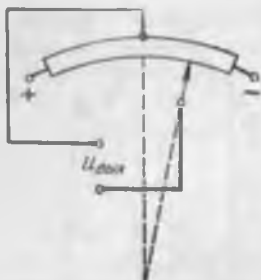


Рис. 3.2.6.

К числу достоинств потенциометрических датчиков следует отнести простоту конструкции, малые габариты и вес, возможность питания как постоянным, так и переменным током. Благодаря этим достоинствам потенциометрические датчики широко применяются в автоматических системах. Основным недостатком потенциометрических датчиков является наличие скользящих контактов.

В потенциометрах, предназначенных для измерения углового перемещения, движок движется по некоторой окружности и каркас потенциометра должен иметь соответствующую форму (рис. 3.2.6).

§ 3.3. Индуктивные датчики

Индуктивные датчики также предназначены для преобразования механического перемещения в электрический сигнал. Работа индуктивного датчика основана на изменении индуктивного сопротивления катушки со стальным сердечником при перемещении подвижного якоря.

На рис. 3.3.1 показана принципиальная схема простейшего индуктивного датчика. Обмотка 1 через сопротивление подключена к сети переменного тока. Магнитный поток обмотки 1 проходит через стальной сердечник 2, воздушный зазор и замыкается через подвижный якорь 3. Индуктивность L катушки 1, измеряемая в генри, определяется формулой

$$L = \frac{w\Phi}{i}, \quad (3.3.1)$$

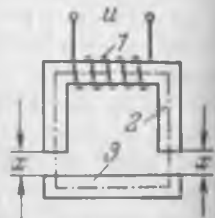


Рис. 3.3.1.

где w — число витков катушки, Φ — магнитный поток в веберах, i — ток катушки в амперах. Если x — величина воздушного зазора в см, а S_m — площадь сечения магнитопровода в см², то магнитный поток равен

$$\Phi = \frac{0,4\pi i w}{R_m} = \frac{0,4\pi i w}{R_{ст} + \frac{2x}{S_m}} \quad (3.3.2)$$

где R_m — магнитное сопротивление цепи, складывающееся из сопротивления стального магнитопровода $R_{ст}$ и сопротивления

двух воздушных зазоров, равного $2x/S_M$ (считаем площадь воздушного зазора $S_x = S_M$). Подставляя выражение (3.3.2) в (3.3.1), получим

$$L = \frac{0,4\pi\omega^2}{R_{ст} + \frac{2x}{S_M}}, \quad (3.3.3)$$

Эффективное значение тока в катушке I определяется формулой

$$I = \frac{U}{\sqrt{R^2 + \omega_0^2 L^2}}, \quad (3.3.4)$$

где U — эффективное значение напряжения сети, ω_0 — частота питающего катушку I переменного тока, R — активное сопротивление катушки I . Подставляя выражение (3.3.3) индуктивности L в (3.3.4) и учитывая, что сопротивление двух воз-



Рис. 3.3.2.

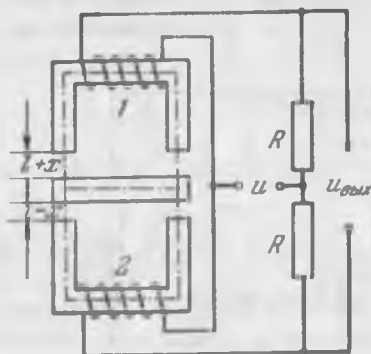


Рис. 3.3.3.

душных зазоров значительно больше сопротивления стального сердечника $R_{ст}$, а индуктивное сопротивление обмотки $\omega_0 L$ значительно больше ее активного сопротивления R , получим

$$I \approx \frac{x}{0,2\pi\omega_0\omega^2 S_M} U = kx. \quad (3.3.5)$$

На рис. 3.3.2 сплошной линией представлена характеристика индуктивного датчика, определяемая формулой (3.3.5). Пунктирной линией показана точная характеристика датчика, построенная по формулам (3.3.4) и (3.3.3).

Простейший индуктивный датчик является однотактным, т. е. знаки его входного и выходного (начальная фаза тока) сигналов не могут изменяться. Кроме того, для изменения величины x к якору простейшего датчика необходимо прикладывать большие и зависящие от x усилия.

Широкое практическое применение находят двухтактные индуктивные датчики. Двухтактные датчики включаются по

дифференциальной или мостовой схеме. На рис. 3.3.3 представлена схема дифференциального индуктивного датчика, состоящего из двух одинаковых простейших индуктивных датчиков с общим якорем. Входным сигналом дифференциального индуктивного датчика является x — перемещение якоря относительно среднего положения. На основании (3.3.5) токи в катушках 1 и 2 равны соответственно $I_1 = k(l + x)$, $I_2 = k(l - x)$. Выходным сигналом датчика является напряжение, эффективное значение которого равно

$$U_{\text{вых}} = I_1 R - I_2 R = 2kRx. \quad (3.3.6)$$

Изменение знака входного сигнала x приводит к изменению фазы выходного переменного напряжения на 180° . Характеристика датчика линейна. Электромеханические усилия, действующие на

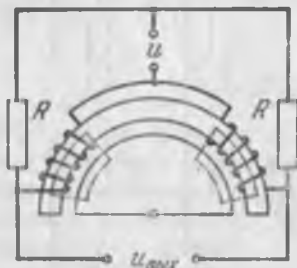


Рис. 3.3.4.

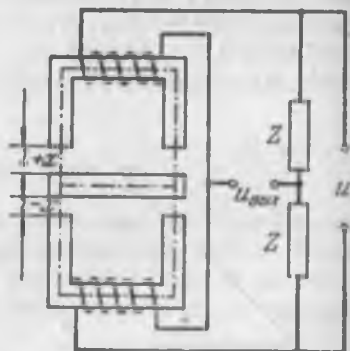


Рис. 3.3.5.

якорь от двух катушек, в значительной степени взаимно компенсируются практически на всем рабочем диапазоне измеряемых перемещений.

На рис. 3.3.4 представлена схема дифференциального индуктивного датчика поворотного типа, который служит для преобразования угловых перемещений в электрические сигналы. При повороте якоря этого датчика изменяется индуктивность обмоток за счет изменения площади сечения воздушного зазора S_x . Изменение индуктивности приводит к связи угла поворота якоря с выходным сигналом.

На рис. 3.3.5 представлена мостовая схема включения простейших индуктивных датчиков. В качестве сопротивлений Z чаще всего используются дроссели с постоянной индуктивностью. Сигнал на выходе появляется при разбалансе моста, к которому приводит отклонение якоря от начального положения.

Работа всех рассмотренных датчиков основана на изменении индуктивности L . Существуют датчики, работа которых основана

на изменении коэффициента взаимной индукции двух катушек M . Такие датчики называются *трансформаторными* или *индукционными*. На рис. 3.3.6 представлена схема трансформаторного индуктивного датчика. Датчик имеет две обмотки питания 3 и одну выходную обмотку 2. Обмотки 3 включены так, что их потоки в сердечнике выходной обмотки 2 направлены в противоположные стороны. Если магнитные сопротивления обеих обмоток питания одинаковы, то потоки и наводимые ими э. д. с. в обмотке 2 взаимно компенсируются, вследствие чего выходной сигнал датчика равен нулю. Если магнитные сопротивления обмоток питания неодинаковы, т. е. неодинаковы коэффициенты взаимной индукции этих обмоток с обмоткой 2, то выходной сигнал датчика будет равен

$$u_{\text{вых}} = e_1 - e_2, \quad (3.3.7)$$

где e_1 и e_2 — э. д. с., наведенные в обмотке 2 магнитными потоками обмоток питания. Учитывая, что

$$e_v = -M_v \frac{di}{dt}, \quad M_v = \frac{w_2 w_3}{R_{Mv}} \quad (v = 1, 2),$$

где M_1 и M_2 — коэффициенты взаимной индукции обмотки 2 с первой и второй обмотками 3, получим

$$u_{\text{вых}} = w_2 w_3 \left[\frac{1}{R_{M1}} - \frac{1}{R_{M2}} \right] \frac{di}{dt}, \quad (3.3.8)$$

где R_{M1} и R_{M2} — магнитные сопротивления обмоток 3, w_2 и w_3 — числа витков обмотки 2 и одной обмотки 3. При постоянном значении магнитной проницаемости сердечников μ и постоянной длине магнитопровода l_M магнитные сопротивления определяются только величинами площадей сечений магнитопроводов обмоток 3:

$$R_{M1} = \frac{l_M}{\mu S_1}, \quad R_{M2} = \frac{l_M}{\mu S_2}. \quad (3.3.9)$$

Площади S_1 и S_2 могут изменяться путем поворота якоря 1. Если изменение площадей происходит пропорционально углу поворота якоря x , отсчитываемого от нейтрального (среднего) положения, то, как это следует из (3.3.8) и (3.3.9), амплитуда выходного сигнала будет пропорциональна входному сигналу x .

Выбором ферромагнитного материала и величины амплитуды переменного напряжения u можно обеспечить синусоидальный закон изменения выходного сигнала в зависимости от x . Особенностью трансформаторного датчика является отсутствие электри-

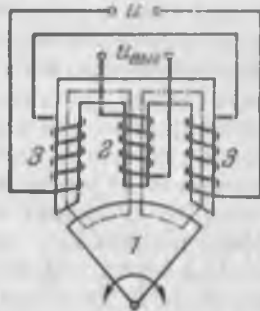


Рис. 3.3.6.

ческой связи между выходной цепью и цепью питания. Между ними существует только магнитная связь.

Основным недостатком индуктивных датчиков является сильная зависимость их характеристик от частоты источника питания. Недостатками являются также работа датчиков только на переменном токе и трудность их регулировки. Основное достоинство индуктивных датчиков — их высокая надежность, обусловленная отсутствием скользящих контактов. Положительными качествами индуктивных датчиков является также большой коэффициент усиления, высокая разрешающая способность (т. е. минимальное перемещение, при котором выходное напряжение датчика заметно отличается от нуля) и больший, чем у потенциометрических датчиков, коэффициент полезного действия. Коэффициент передачи дифференциального датчика, например, может составлять несколько сотен вольт на миллиметр перемещения. Разрешающая способность некоторых индуктивных датчиков при тщательной экранировке и регулировке их цепей может измеряться сотыми долями микрона. Благодаря этому индуктивные датчики особенно пригодны для измерения очень малых перемещений.

Изложенное показывает, что с точки зрения динамических свойств индуктивные датчики, так же как и потенциометрические, являются безынерционными линейными усилителями. Однако это справедливо только при условии, что входной сигнал (перемещение или поворот якоря) изменяется достаточно медленно по сравнению с изменением питающего датчик переменного напряжения, что обычно всегда имеет место.

§ 3.4. Тахометры

Тахометром или *тахогенератором* (ТГ) называется малогабаритный генератор постоянного тока с независимым возбуждением, э. д. с. которого линейно зависит от числа оборотов якоря. Таким образом, тахогенератор является электрическим датчиком. Входным сигналом датчика служит угловая скорость вала. Выходным сигналом является напряжение.

Если ротор генератора вращается с угловой скоростью ω , а Φ_v есть поток возбуждения, то в обмотке якоря будет наводиться э. д. с., равная $e = c\Phi_v\omega$, где c — конструктивная постоянная тахогенератора. При постоянном потоке возбуждения Φ_v э. д. с. в обмотке якоря, а следовательно, и напряжение на выходных клеммах ненагруженного тахогенератора будут пропорциональны угловой скорости ω :

$$u_{\text{вых}} = e = k\omega. \quad (3.4.1)$$

Однако нагрузка ТГ искажает линейную зависимость напряжения $u_{\text{вых}}$ от скорости вращения ротора. Причиной искажения яв-

дается реакция якоря, оказывающая влияние на магнитный поток возбуждения. Уменьшить реакцию якоря можно уменьшением тока якоря, т. е. увеличением сопротивления нагрузки тахогенератора. Чем меньше ток якоря, тем меньше искажается линейная характеристика датчика.

В целях компенсации реакции якоря иногда используют дополнительную обмотку, расположенную на статоре, включенную последовательно с обмоткой якоря и с сопротивлением нагрузки R_n (рис. 3.4.1).

При наличии нагрузки и компенсационной обмотки

$$u_{\text{вых}} = e - i_n R_n = c (i_n w_n + i_n w_n - i_n w_n) \omega - i_n R_n, \quad (3.4.2)$$

где i_n , i_k , i_n — токи в цепях возбуждения, компенсации и якоря, w_n , w_k — числа витков обмоток возбуждения и компенсации, w_n — зависящее от ω число витков фиктивной обмотки, эквивалентной якорю, R_n — сопротивление якоря. Из рис. 3.4.1 находим

$$i_n = i_n = \frac{u_{\text{вых}}}{R_0}, \quad i_k = i_n \frac{R_n}{R_k + R_n} = \frac{u_{\text{вых}}}{R_1}, \quad i_n = \frac{u_n}{R_n}, \quad (3.4.3)$$

$$R_0 = R_n + \frac{R_k R_n}{R_k + R_n}, \quad R_1 = R_0 \frac{R_k + R_n}{R_n}, \quad (3.4.4)$$

где R_n , R_k — сопротивления обмоток возбуждения и компенсации, R_n , R_n — сопротивления нагрузки и шунта. Полагая

$$a_1 = w_n/R_n, \quad a_2 = w_k/R_1, \quad a_3 = w_n/R_0, \quad a_4 = 1 + (R_n/R_0)$$

и исключая из (3.4.2), (3.4.3) и (3.4.4) i_n , i_n , i_k , i_n , приведем формулу (3.4.2) к виду

$$u_{\text{вых}} = \frac{c u_n a_1 \omega}{a_4 - c (a_2 - a_3) \omega}. \quad (3.4.5)$$

При заданной нагрузке путем подбора сопротивления R_n можно в достаточно большом диапазоне изменения ω обеспечить приближенное равенство $a_2 \approx a_3$ и получить практически линейную зависимость входного напряжения от угловой скорости.

§ 3.5. Гироскопические измерители угловой скорости

Гироскопический измеритель угловой скорости состоит из двухстепенного гироскопа, являющегося чувствительным элементом, и потенциометрического (или какого-нибудь другого) датчика (рис. 3.5.1). При вращении корпуса гироскопа вокруг

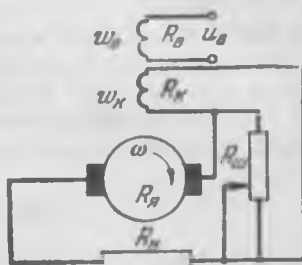


Рис. 3.4.1.

оси YY со скоростью Ω возникает гироскопический момент, приближенно равный

$$M_r = K\Omega, \quad (3.5.1)$$

где K — кинетический момент ротора гироскопа. Для уравновешивания этого момента используют пружину. Как известно, развиваемая пружиной сила пропорциональна деформации, которая в свою очередь пропорциональна углу поворота рамки гироскопа θ вокруг оси ZZ (рис. 3.5.1):

$$M_p = -cl^2\theta, \quad (3.5.2)$$

где c — жесткость пружины, а l — расстояние между точкой крепления пружины к рамке гироскопа и осью вращения рамки ZZ .

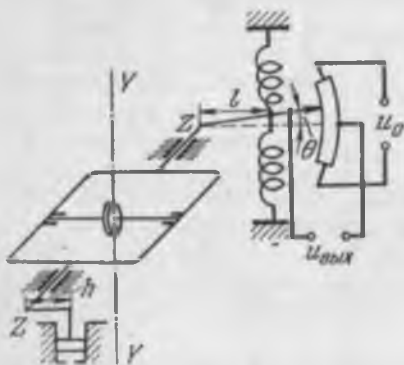


Рис. 3.5.1.

При длительном вращении корпуса гироскопа с постоянной скоростью Ω колебания рамки затухнут, и рамка будет находиться в равновесии под действием гироскопического момента M_r и момента пружины M_p . Приравнявая нулю сумму этих моментов (3.5.1) и (3.5.2), приходим к заключению, что угол отклонения рамки гироскопа от нейтрального положения θ будет пропорционален угловой скорости корпуса гироскопа Ω относительно оси YY . На этом и основано применение двух-

степенного гироскопа в качестве чувствительного элемента измерителя угловой скорости.

Для определения динамических характеристик гироскопического измерителя угловой скорости составим уравнение движения рамки гироскопа. Кроме гироскопического момента и момента пружины на рамку гироскопа действует момент усилия демпфера, обеспечивающего затухание колебаний рамки гироскопа. Усилие демпфера пропорционально скорости движения поршня, равной угловой скорости рамки гироскопа $\dot{\theta}$, умноженной на расстояние h точки приложения усилия демпфера от оси вращения рамки. Следовательно, демпфирующий момент выражается формулой

$$M_d = -vh^2\dot{\theta}, \quad (3.5.3)$$

где v — коэффициент демпфирования. Используя выражения (3.5.1), (3.5.2) и (3.5.3) моментов, действующих на рамку гироскопа, получим следующее уравнение движения рамки гироскопа,

скопа относительно его корпуса, справедливое для достаточно малых значений угла θ :

$$J\ddot{\theta} + \nu h^2\dot{\theta} + cl^2\theta = K\Omega, \quad (3.5.4)$$

где J — экваториальный момент инерции гироскопа.

Таким образом, двухстепенной гироскоп, предназначенный для измерения угловой скорости, является колебательным звеном (см. приложение 1).

Для преобразования угла θ , являющегося выходной величиной чувствительного элемента, в электрический сигнал чаще всего используется потенциометрический датчик. Датчик устанавливается на корпус гироскопа. Движок потенциометрического датчика механически связывается с рамкой гироскопа (рис. 3.5.1).

§ 3.6. Акселерометры

Акселерометром называется измеритель ускорения. Чувствительным элементом акселерометра является груз 1 (рис. 3.6.1), подвешенный на двух пружинах 2. Датчиком, преобразующим механическое перемещение груза в электрический сигнал, может быть потенциометр 5. В некоторых акселерометрах применяются индуктивные и другие датчики. Обычно акселерометр имеет демпфер 4, успокаивающий колебания груза 1. Чувствительный элемент, датчик и демпфер помещаются в корпус 3 акселерометра, который закрепляется вблизи центра массы объекта, ускорение которого измеряется. Осью чувствительности измерителя является направление XX перемещения груза относительно корпуса акселерометра.

При отсутствии ускорения, направленного по оси чувствительности, груз располагается в нейтральном положении, при котором $u_{\text{вых}} = 0$. При наличии ускорения вдоль оси чувствительности груз под действием силы инерции сместится от нейтрального положения в сторону, противоположную направлению ускорения, и займет такое положение, при котором сила инерции уравнивается силой упругости пружин. Вместе с грузом сместится от начального положения и движок потенциометра, вследствие чего на выходе измерителя появится напряжение $u_{\text{вых}} \neq 0$. Так как сила упругости пружин пропорциональна смещению груза x , а сила инерции пропорциональна ускорению, то

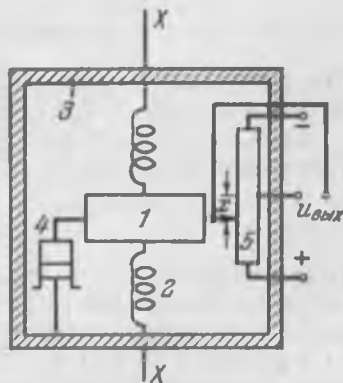


Рис. 3.6.1.

выходной сигнал акселерометра при движении с постоянным ускорением будет пропорционален ускорению.

Для нахождения точной зависимости смещения груза акселерометра от ускорения составим уравнение движения груза относительно корпуса акселерометра. На груз действуют сила упругости пружины, усилие демпфера, направленные в сторону, противоположную перемещению x , и сила тяжести. Составляющая

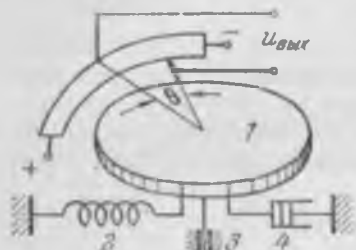


Рис. 3.6.2.

полного ускорения груза вдоль оси X , очевидно, равна $\ddot{x} + j_x$, где j_x — составляющая ускорения объекта вдоль оси x . Следовательно, уравнение движения груза имеет вид

$$m(\ddot{x} + j_x) = -cx - v\dot{x} + mg_x, \quad (3.6.1)$$

где m — масса груза, g_x — составляющая ускорения силы тяжести по оси XX , c — жесткость пружины, а v — коэффициент демпфирования. После элементарных преобразований уравнение (3.6.1) принимает вид

$$\ddot{x} + \frac{v}{m}\dot{x} + \frac{c}{m}x = -(j_x - g_x). \quad (3.6.2)$$

Это уравнение показывает, что акселерометр является колебательным звеном (см. приложение 1).

Если объект движется так, что величина $j_x - g_x$ постоянна, то уравнение (3.6.2) имеет интеграл

$$x = -\frac{m}{c}(j_x - g_x). \quad (3.6.3)$$

Таким образом, при постоянной величине $j_x - g_x$ груз акселерометра после затухания колебаний установится в положении равновесия, определяемом формулой (3.6.3).

Уравнение (3.6.2) и формула (3.6.3) показывают, что акселерометр измеряет всегда только разность между соответствующими составляющими ускорения объекта и ускорения силы тяжести. Поэтому любой акселерометр всегда работает с систематической ошибкой, равной соответствующей составляющей ускорения силы тяжести. Эта ошибка может быть компенсирована соответствующим смещением корпуса потенциометра относительно корпуса акселерометра. Для этого необходимо иметь на движущемся объекте вычислитель составляющей ускорения силы тяжести по оси чувствительности акселерометра и соответствующие исполнительные устройства, смещающие корпус потенциометра. Однако обычно в системах управления полетом такая компенсация систематических ошибок акселерометров не применяется. Для исклю-

чения ошибок, связанных с ускорением силы тяжести, акселерометры часто устанавливают на платформу, стабилизированную в горизонтальном положении.

Аналогично устроены акселерометры, предназначенные для измерения угловых ускорений (рис. 3.6.2). В данном случае под действием момента сил инерции груз поворачивается относительно

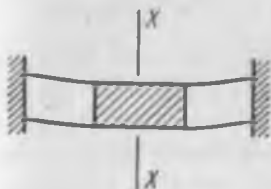


Рис. 3.6.3.

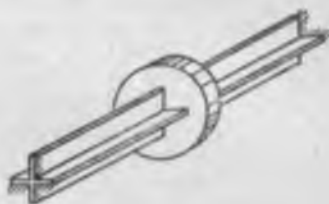


Рис. 3.6.4.

корпуса акселерометра. Уравнение вращения груза относительно корпуса акселерометра имеет вид

$$\ddot{\theta} + 2a\dot{\theta} + b^2\theta = -w_{\theta}, \quad (3.6.4)$$

где w_{θ} — угловое ускорение объекта относительно оси вращения груза, $2a$ и b^2 — коэффициенты, характеризующие соответственно демпфирование и жесткость пружин.

Точность и чувствительность акселерометра сильно зависят от конструкции подвески его груза. Подвеска должна обеспечить движение груза только в одном направлении и с наименьшим трением. Должна быть обеспечена также линейная зависимость усилий, действующих на груз со стороны элементов подвески, от перемещения и скорости груза относительно корпуса акселерометра. В акселерометрах, измеряющих линейное ускорение, наиболее целесообразной является подвеска груза на плоских пружинах, закрепленных с обоих концов (рис. 3.6.3). В акселерометрах, служащих для измерения углового ускорения, может быть использована крестообразная пружина, работающая на кручение (рис. 3.6.4).

§ 3.7. Магнитоэлектрические чувствительные элементы

Магнитоэлектрическим чувствительным элементом называется устройство, служащее для преобразования электрического напряжения в механическое перемещение (обычно угловое перемещение). Магнитоэлектрический элемент состоит из подвижной катушки 1 и неподвижного постоянного магнита 2 (рис. 3.7.1). Входной переменной устройства является напряжение $u_{\text{вх}}$, подаваемое на катушку через спиральные пружины 3, которые

устанавливают подвижную систему в нейтральное положение при отсутствии входного сигнала. Выходной переменной является угол поворота α катушки.

Катушка вращается под действием электромагнитного момента M_e и момента пружин M_n . Поэтому уравнение движения катушки имеет вид

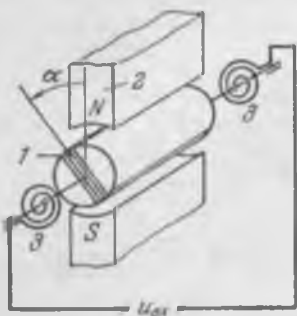


Рис. 3.7.1.

$$J\ddot{\alpha} = M_e + M_n, \quad (3.7.1)$$

где J — момент инерции катушки. Момент пружин пропорционален углу поворота катушки:

$$M_n = -c\alpha, \quad (3.7.2)$$

где c — жесткость пружин. Электромагнитный момент пропорционален току i в катушке и магнитному потоку Φ . Поскольку у постоянного магнита $\Phi = \text{const}$, то

$$M_e = k_1 i, \quad (3.7.3)$$

где k_1 — постоянный коэффициент пропорциональности. Ток в катушке i определяется уравнением

$$L \frac{di}{dt} + Ri + k_2 \frac{d\alpha}{dt} = u_{\text{вх}}, \quad (3.7.4)$$

где $k_2 \dot{\alpha}$ — э. д. с., наводимая в катушке при вращении ее со скоростью $\dot{\alpha}$, R и L — соответственно омическое сопротивление и индуктивность катушки.

Если величина индуктивности катушки мала, то приближенно можно считать, что $L = 0$, и уравнение для тока принимает вид

$$i = \frac{1}{R} u_{\text{вх}} - \frac{k_2}{R} \dot{\alpha}. \quad (3.7.5)$$

Подставляя это выражение тока в формулу для электромагнитного момента (3.7.3), получим

$$M_e = \frac{k_1}{R} u_{\text{вх}} - \frac{k_1 k_2}{R} \dot{\alpha}. \quad (3.7.6)$$

На основании формул (3.7.2) и (3.7.6) уравнение движения катушки (3.7.1) принимает вид

$$T^2 \ddot{\alpha} + 2\zeta T \dot{\alpha} + \alpha = k u_{\text{вх}}, \quad (3.7.7)$$

где $T = \sqrt{J/c}$ — постоянная времени, $\zeta = k_1 k_2 / 2R \sqrt{Jc}$ — коэффициент затухания, k — коэффициент усиления. Таким образом, по своим динамическим свойствам магнитоэлектрический чувствительный элемент является колебательным звеном (см. приложение 1).

Магнитоэлектрические чувствительные элементы нашли широкое распространение в электроизмерительных приборах (вольтметрах, амперметрах и т. п.), а также в автоматических системах управления как элементы согласования различных типов электрических усилителей с гидравлическими и пневматическими исполнительными устройствами. В последнем случае, как правило, вместо постоянного магнита используется электромагнит. Типичная схема устройства показана на рис. 3.7.2. Якорь 1 может вращаться относительно оси 2. На ярме 3 намотаны четыре обмотки подмагничивания w_n , питаемые постоянным током. За счет тока в катушках w_n ярмо приобретает свойства электромагнита с определенной полярностью. На якорь 1 намотаны обмотки управления w_y , к которым приложено входное напряжение $u_{вх}$.

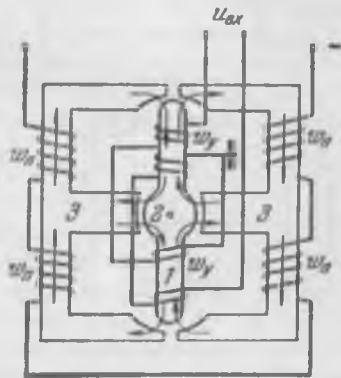


Рис. 3.7.2.

При отсутствии входного напряжения устойчивым положением якоря является нейтральное положение, когда все воздушные зазоры между ярмом и якорем одинаковы и, следовательно, сопротивление магнитных цепей минимально.

Таким образом, роль пружины в данной схеме играет восстанавливающий момент, возникающий за счет изменения сопротивлений воздушных зазоров при смещении якоря. При $u_{вх} \neq 0$ в якорь и ярмо возникает магнитный поток управления. Взаимодействие полюсов электромагнитов ярма и якоря создает момент, поворачивающий якорь на угол, пропорциональный входному напряжению.

Рассмотренная дифференциальная схема устройства обеспечивает хорошую сбалансированность подвижных частей, что необходимо в условиях действия ускорений, широкий диапазон линейности и высокую чувствительность. По своим динамическим свойствам такой элемент также является колебательным звеном, если пренебречь индуктивностью катушек w_y . Максимальный момент, развиваемый магнитоэлектрическим элементом, имеет порядок $100 \div 500$ гс·см.

§ 3.8. Сельсины

Сельсинами называются специальные машины переменного тока, которые в зависимости от того, как включены их статорные и роторные обмотки, могут работать в качестве датчиков или в качестве элементов дистанционной передачи. Датчиками являются сельсины, работающие в *трансформаторном* режиме. В дистанционных передачах используются сельсины, работающие в *индикаторном* режиме. Кроме того, существуют так называемые *дифференциальные сельсины*, которые применяются как в сельсинных датчиках, так и в дистанционных передачах.

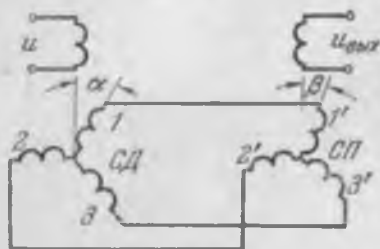


Рис. 3.8.1.

Сельсины, работающие в трансформаторном или индикаторном режиме, имеют по четыре обмотки. Три обмотки соединяются в звезду и имеют магнитные оси, сдвинутые друг относительно друга на 120° . Совокупность трех расположенных таким образом обмоток обычно называют *трехфазной* обмоткой. Если трехфазная обмотка располагается на статоре, то четвертая *однофазная* обмотка располагается на роторе, и наоборот. Первый вариант применяется в малоомощных сельсинах, второй — в сельсинах большой мощности.

Обычно сельсины применяются попарно. Один называется *сельсином-датчиком (СД)*, а другой — *сельсином-приемником (СП)*.

На рис. 3.8.1 показаны сельсины, работающие в трансформаторном режиме. Статорными являются однофазные обмотки. Роторные обмотки *СД* соединены с соответствующими обмотками ротора *СП*. Статорная обмотка *СД* питается переменным током. Концы статорной обмотки *СП* выведены к выходным клеммам.

Пара включенных таким образом сельсинов является датчиком, преобразующим угловое перемещение ротора *СД* в электрический сигнал, снимаемый со статорной обмотки *СП*. Входным сигналом датчика является угол поворота ротора *СД* относительно ротора *СП*, т. е. разность в угловых положениях роторов. Выходным сигналом является напряжение, наводимое в статорной обмотке *СП*.

Для вывода приближенной формулы для выходного напряжения *СП* пренебрежем активными сопротивлениями обмоток, учитывая, что они малы по сравнению с их реактивными сопротивлениями. Кроме того, пренебрежем э.д.с. самонадукции и э.д.с. взаимонадукции обмоток, так как сумма этих э.д.с. мала по сравнению с э.д.с., наведенной в обмотке потоком возбуждения. Тогда эффективные значения э.д.с. в обмотках ротора

определяются формулами

$$\left. \begin{aligned} E_1 &= E_m \cos \alpha, \\ E_2 &= E_m \cos (\alpha - 120^\circ), \\ E_3 &= E_m \cos (\alpha + 120^\circ), \end{aligned} \right\} \quad (3.8.1)$$

где E_m — эффективное значение э. д. с., наводимой в обмотке I потоком возбуждения при $\alpha = 0$. Обозначим через Z сопротивление каждой фазовой обмотки ротора, считая, что они одинаковы. Тогда, если пренебречь сопротивлением соединительных проводов и учесть, что взаимное влияние обмоток роторов, расположенных на большом удалении друг от друга, отсутствует, а магнитопроводы сельсиннов работают в ненасыщенном режиме, то эффективные значения токов в фазовых обмотках выразятся формулами

$$I_1 = \frac{E_1}{2Z}, \quad I_2 = \frac{E_2}{2Z}, \quad I_3 = \frac{E_3}{2Z}. \quad (3.8.2)$$

Эти токи создают переменные магнитные потоки, которые в однофазной статорной обмотке *СП* будут индуцировать электродвижущие силы, эффективные значения которых соответственно равны

$$\left. \begin{aligned} E'_1 &= k_1 I_1 \cos \beta, \\ E'_2 &= k_1 I_2 \cos (\beta - 120^\circ), \\ E'_3 &= k_1 I_3 \cos (\beta + 120^\circ), \end{aligned} \right\} \quad (3.8.3)$$

где k_1 — коэффициент пропорциональности. Эффективное значение выходного напряжения сельсина приемника будет равно

$$U_{\text{вых}} = E'_1 + E'_2 + E'_3. \quad (3.8.4)$$

Подставляя сюда выражения (3.8.3.) и учитывая (3.8.2) и (3.8.1), получим

$$U_{\text{вых}} = \frac{k_1 E_m}{2Z} [\cos \beta \cos \alpha + \cos (\beta - 120^\circ) \cos (\alpha - 120^\circ) + \\ + \cos (\beta + 120^\circ) \cos (\alpha + 120^\circ)]$$

или, после тригонометрических преобразований,

$$U_{\text{вых}} = \frac{3k_1 E_m}{4Z} \cos (\beta - \alpha) = U_m \cos (\beta - \alpha), \quad (3.8.5)$$

где U_m — амплитудное значение напряжения при нулевом рассогласовании $\beta - \alpha$. Эта формула показывает, что напряжение на выходных клеммах сельсинного датчика не зависит от абсолютного углового положения роторов, а зависит только от их относительного углового положения. Если положение ротора сельсина-приемника зафиксировать, то по величине $U_{\text{вых}}$ можно судить об угле поворота ротора сельсина-датчика. При фиксированном значении

$\beta = \frac{\pi}{2}$ выходное напряжение сельсинного датчика будет пропорционально $\cos\left(\frac{\pi}{2} - \alpha\right) = \sin \alpha$. При малых углах α можно считать $U_{\text{вых}}$ пропорциональным α .

На рис. 3.8.2 показаны сельсины, работающие в индикаторном режиме. В отличие от сельсинов, работающих в трансформаторном режиме, здесь статорная однофазная обмотка $СП$, так же как и статорная обмотка $СД$, подключена к источнику переменного тока.

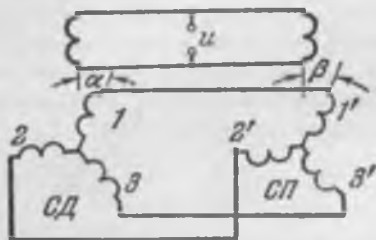


Рис. 3.8.2.

Положение ротора $СП$ не фиксируется, он может свободно вращаться на своей оси.

Переменное напряжение питания создает в статорных обмотках обоих сельсинов магнитные потоки, направления которых совпадают с направлениями осей статорных обмоток. Эти магнитные потоки индуцируют в фазовых обмотках роторов электродвижущие силы. При одинаковом положении роторов $СД$ и $СП$ по отношению к статорным обмоткам ($\alpha = \beta$) эти э. д. с. в соответствующих фазовых обмотках равны между собой и противоположны по направлению. Следовательно, результирующие э. д. с. в каждой паре соединенных между собой фазовых обмоток равны нулю и ток в цепях роторов отсутствует ($I = I_2 = I_3 = 0$). Если же ротор $СД$ повернут на некоторый угол относительно ротора $СП$, то э. д. с. в соответствующих фазовых обмотках роторов окажутся различными по величине. Результирующие э. д. с. в фазовых обмотках роторов не будут равны нулю, и в цепи роторов возникнут токи I_1, I_2, I_3 . Взаимодействие результирующих магнитных потоков роторов с магнитными потоками обмоток статоров создает вращающие моменты. Поэтому, если ротор $СД$ после разворота на заданный угол α остановить, то ротор $СП$ под действием вращающего момента будет вращаться до тех пор, пока не станет в положение $\beta = \alpha$. При непрерывном вращении ротора $СД$ с определенной скоростью ротор $СП$ будет вращаться с той же скоростью и следить за ротором $СД$. Вращающий момент обеспечивает синхронность движения роторов, поэтому его называют *синхронизирующим моментом*.

По аналогии с (3.8.1) для фазовых обмоток ротора сельсин-приемника, работающего в индикаторном режиме, можно записать

$$\left. \begin{aligned} E'_1 &= E_m \cos \beta, \\ E'_2 &= E_m \cos (\beta - 120^\circ), \\ E'_3 &= E_m \cos (\beta + 120^\circ). \end{aligned} \right\} \quad (3.8.6)$$

Эффективные значения результирующих э. д. с. в обмотках роторов будут равны

$$\Delta E_1 = E'_1 - E_1, \quad \Delta E_2 = E'_2 - E_2, \quad \Delta E_3 = E'_3 - E_3. \quad (3.8.7)$$

После подстановки сюда выражений (3.8.1) и (3.8.6) и тригонометрических преобразований получим

$$\left. \begin{aligned} \Delta E_1 &= 2E_m \sin \frac{\alpha + \beta}{2} \sin \frac{\alpha - \beta}{2}, \\ \Delta E_2 &= 2E_m \sin \left(\frac{\alpha + \beta}{2} - 120^\circ \right) \sin \frac{\alpha - \beta}{2}, \\ \Delta E_3 &= 2E_m \sin \left(\frac{\alpha + \beta}{2} + 120^\circ \right) \sin \frac{\alpha - \beta}{2}. \end{aligned} \right\} \quad (3.8.8)$$

Отсюда следует, что результирующие э. д. с. одновременно обращаются в нуль лишь в случае $\alpha = \beta$, т. е. при одинаковом положении роторов *СД* и *СП*.

Результирующий магнитный поток имеет составляющую, направленную по оси статорной обмотки (продольная составляющая), и составляющую, перпендикулярную оси статорной обмотки (поперечная составляющая). Продольная составляющая ослабляет магнитный поток обмотки статора, но не создает вращающего момента. При малых углах рассогласования $\beta - \alpha$ влияние продольной составляющей незначительно. Поперечная составляющая создает синхронизирующий момент. Величина синхронизирующего момента определяется формулой

$$M_{\text{синх}} = M_m \sin \theta, \quad (3.8.9)$$

где $\theta = \beta - \alpha$, M_m — момент, действующий на ротор *СП* при рассогласовании θ , равном $\pi/2$.

На ротор *СД* будет действовать момент, также равный $M_{\text{синх}}$, так как обмотки роторов одинаковы и величины токов, текущих по соответственным обмоткам роторов, одинаковы; но знак этого момента будет противоположным, так как направления токов в соответственных обмотках противоположны.

Формула (3.8.9) показывает, что синхронизирующий момент равен нулю при $\theta = 0^\circ$ и $\theta = 180^\circ$, т. е. в пределах изменения θ от 0° до 360° в системе возможны два состояния равновесия. Однако при $\theta = 180^\circ$ положение ротора *СП* неустойчиво, так как при любом $\theta \neq \pm \pi$ момент, действующий на ротор приемника, увеличивает β и уменьшает α при $\alpha > \beta$ и наоборот при $\alpha < \beta$.

Практически вследствие наличия моментов трения в подшипниках и контактных кольцах, момента нагрузки, неточной балансировки роторов, неодинаковости конструктивных и электрических параметров *СД* и *СП* и других факторов равновесное состояние сельсинной системы наступает при $\alpha \neq \beta$, т. е. слежение ротора

СП за ротором СД происходит с некоторой ошибкой. Ошибка слежения обычно определяется экспериментально и для сельсинов первого и второго классов точности не превосходит $1,5^\circ$. Ошибка может быть значительно уменьшена, если входной сигнал подается в сельсинную систему через редуктор, а выходной сигнал снимается с помощью такого же редуктора.

Заметим, что выводы, полученные для случая, когда однофазные обмотки расположены на статорах СП и СД, остаются в силе и для случая, когда однофазные обмотки расположены на роторах.

Дифференциальный сельсин имеет шесть обмоток. Три статорные обмотки, так же как и три роторные, соединены в звезду с магнитными осями, сдвинутыми друг относительно друга на 120° .

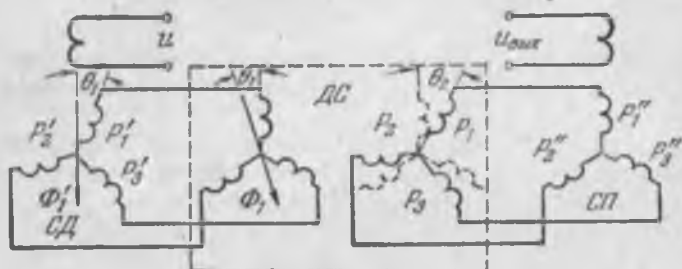


Рис. 3.8.3.

Дифференциальный сельсин включается между двумя обычными сельсинами и, по существу, обеспечивает возможность введения дополнительного входного сигнала.

На рис. 3.8.3. представлена схема, в которой дифференциальный сельсин (ДС) работает в трансформаторном режиме и используется в качестве датчика. Поворот ротора СД на угол θ_1 вызывает поворот результирующего магнитного потока статора ДС на тот же угол, но в противоположную сторону. Если ротор ДС занимает такое положение (пунктир), когда магнитные оси его обмоток параллельны соответствующим магнитным осям обмоток статора, то э.д.с., наведенные в обмотках P_1 , P_2 и P_3 ДС, будут такими же, как э.д.с. в обмотках P_1' , P_2' , P_3' , так как поток статора ДС Φ_1 занимает относительно обмоток ротора такое же положение, как поток статора Φ_1' СД относительно обмоток его ротора. Точно такие же э.д.с. будут наводиться в обмотках ротора СП. Выходное напряжение $U_{\text{вых}}$ будет изменяться в соответствии с формулой (3.8.5) при $\beta - \alpha = \theta_1$. В этом случае ДС реализует индуктивную (а не проводную) связь обмоток роторов СД и СП.

Если ротор ДС повернуть на угол θ_2 относительно статора, то положение магнитного потока Φ_1' по отношению к обмоткам P_1 , P_2 , P_3 изменится на угол θ_2 . Магнитный поток роторных обмоток

СП также изменит свое положение на угол θ_2 . В соответствии с этим изменится выходное напряжение. Аналогичный результат можно было бы получить, не поворачивая ротор ДС, а дополнительно повернув ротор СД на угол θ_2 . Следовательно, с помощью ДС организуется второй вход в сельсинный датчик. Функция дифференциального сельсина отличается от функции обычного сельсина-датчика тем, что ДС является не только сельсином-датчиком, но и осуществляет передачу сигнала к сельсину-приемнику от другого сельсина-датчика.

На рис. 3.8.4 представлена схема управления положением ротора дифференциального сельсина при помощи двух обычных сельсинов-датчиков. На этой схеме дифференциальный сельсин

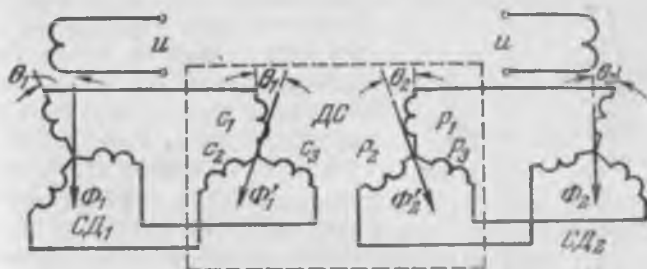


Рис. 3.8.4.

выделен пунктиром. Статорные обмотки ДС обозначены буквами c , роторные — буквами p . Если ротор СД₁ повернут против часовой стрелки на угол θ_1 , то результирующий магнитный поток статора дифференциального сельсина ДС, соединенного с ротором СД₁, повернется на тот же угол θ_1 , но по часовой стрелке и займет положение Φ'_1 . При повороте ротора СД на угол θ_2 по часовой стрелке результирующий магнитный поток Φ_2 ротора ДС, соединенного с ротором СД, повернется против часовой стрелки на тот же угол θ_2 . Взаимодействие потоков Φ'_1 и Φ'_2 создает вращающий момент, который развернет ротор ДС на угол $\theta_d = \theta_1 + \theta_2$, т. е. до совпадения направлений потоков Φ'_1 и Φ'_2 . Таким образом, угол поворота ротора θ_d дифференциального сельсина является алгебраической суммой углов поворота роторов двух сельсинов-датчиков θ_1 и θ_2 , которые являются входными сигналами. В рассмотренной схеме дифференциальный сельсин работает в индикаторном режиме и используется в качестве сумматора углов поворотов.

Использование дифференциальных сельсинов позволяет производить управление объектами из нескольких пунктов.

§ 3.9. Электронные усилители

Большое применение в автоматических системах управления получили *электронные усилители* переменного и постоянного тока.

На рис. 3.9.1 изображен каскад лампового усилителя переменного тока на сопротивлениях. Для этого усилителя справедливы следующие соотношения:

$$\left. \begin{aligned} i_a &= i + i_1, & i &= s u_{вх} + \frac{u_a}{R_i}, & u_a &= -R_0 i_a, \\ u_{вх} &= R_1 i_1, & u_a &= u_{вх} + u_c, & i_1 &= C \frac{du_c}{dt}, \end{aligned} \right\} \quad (3.9.1)$$

где i_a , i , i_1 — переменные токи, проходящие через анодную нагруз-

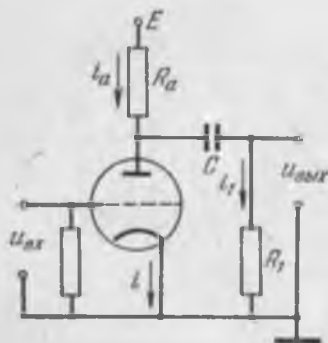


Рис. 3.9.1.

ку, лампу и сопротивление R_1 соответственно, $u_{вх}$, u_a — переменные напряжения на входе усилителя и аноде лампы, u_c — напряжение на конденсаторе, s — крутизна сеточной характеристики, при работе лампы в линейном режиме (т. е. на линейном участке характеристики) — величина постоянная, R_i — внутреннее сопротивление лампы. Остальные обозначения ясны из рис. 3.9.1.

Исключая из уравнений (3.9.1) переменные i_a , i , i_1 , u_a , u_c после алгебраических преобразований, получим дифференциальное уравнение работы усилителя:

$$T \dot{u}_{вх} + u_{вх} = -k \dot{u}_{вх}, \quad (3.9.2)$$

где

$$k = \frac{s R_1 R_i R_a C}{R_i + R_a}, \quad T = \left(R_1 + \frac{R_i R_a}{R_i + R_a} \right) C.$$

Постоянная времени T в усилителях имеет порядок $10^{-4} \div 10^{-6}$ с. Поэтому усилитель на сопротивлениях является практически безынерционным элементом.

Представляет интерес определение динамических свойств усилителя по отношению к огибающей *) несущей частоты. Предположим, что на входе усилителя действует сигнал вида

$$u_{вх} = u \cos \omega_0 t, \quad (3.9.3)$$

*) Огибающей высокочастотных колебаний (не обязательно гармонических) называется их амплитуда, которая обычно представляет собой медленно изменяющуюся функцию времени, т. е. такую функцию, изменение которой в пределах одного периода колебаний мало.

где u — огибающая. Решение дифференциального уравнения (3.9.2) при входном сигнале вида (3.9.3) выражается формулой

$$u_{\text{вых}} = u_1 \cos \omega_0 t + u_2 \sin \omega_0 t, \quad (3.9.4)$$

где u_1 и u_2 — также медленно изменяющиеся функции времени. Подставляя выражения (3.9.3) и (3.9.4) в уравнение (3.9.2) и сравнивая коэффициенты при синусах и косинусах в левой и правой частях уравнения, получим уравнения для определения u_1 и u_2 :

$$\left. \begin{aligned} (TD + 1) u_1 + T\omega_0 u_2 &= -kDu, \\ -T\omega_0 u_1 + (TD + 1) u_2 &= k\omega_0 u. \end{aligned} \right\} \quad (3.9.5)$$

Пренебрегая здесь производными медленно изменяющихся функций и решая полученные алгебраические уравнения относительно u_1 и u_2 , находим

$$\left. \begin{aligned} u_1 &= -\frac{kT\omega_0^2}{1 + T^2\omega_0^2} u, \\ u_2 &= \frac{k\omega_0}{1 + T^2\omega_0^2} u. \end{aligned} \right\} \quad (3.9.6)$$

Отсюда видно, что по отношению к медленно изменяющейся огибающей входного сигнала усилитель является безынерционным элементом. В автоматических системах при использовании сигнала на переменном токе (в модуляторах, демодуляторах, двигателях и др.) обычно ограничиваются только одной составляющей u_2 выходного сигнала лампового усилителя.

Усилители постоянного тока предназначены для усиления медленно изменяющихся сигналов постоянного тока, так называемых *инфранизкочастотных* сигналов. Управляющая сетка каждого каскада усилителя постоянного тока связывается непосредственно с анодом предшествующего каскада. Так как анод находится под высоким положительным потенциалом, а связанная с ним управляющая сетка должна иметь отрицательный потенциал относительно катода, то возникает трудность создания необходимого режима работы электронных ламп в усилителях постоянного тока.

Вторая особенность электронных усилителей постоянного тока состоит в случайных изменениях выходного сигнала при нулевом значении входного сигнала. Это явление называется *дрейфом нуля*. Дрейф нуля вызывается нестабильностью источников питания, медленными случайными изменениями электронной эмиссии ламп и другими причинами. Проблема борьбы с дрейфом нуля в усилителях постоянного тока является весьма острой. В усилителях переменного тока связь каскадов осуществляется через разделительные емкости. Для таких усилителей помехи, приводящие к дрейфу нуля, легко отфильтровываются, так как

разделительные емкости являются для них сопротивлениями, близкими к бесконечности.

Третьей особенностью усилителей постоянного тока является необходимость применения мостовых или двухтактных схем включения нагрузки, которые способны обеспечить изменение знака выходного сигнала при изменении полярности сигнала на входе.

На рис. 3.9.2 показана принципиальная схема трехкаскадного усилителя постоянного тока. Необходимый режим работы ламп обеспечивается путем потенциометрической межкаскадной связи,

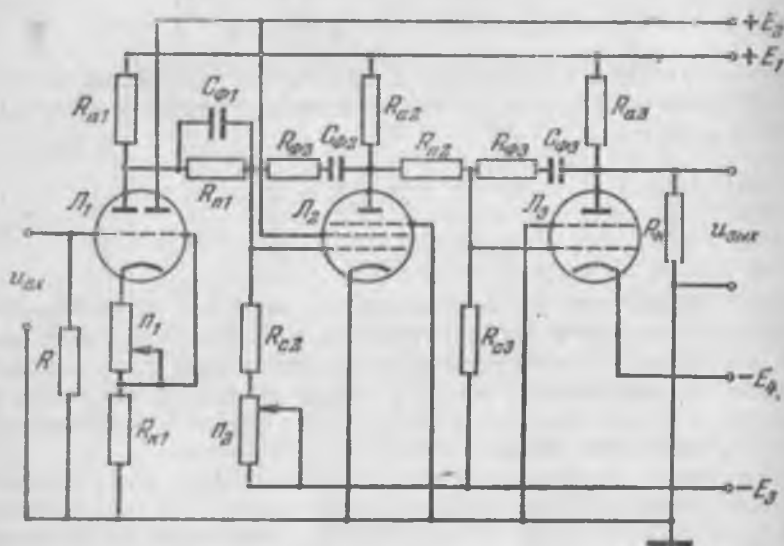


Рис. 3.9.2.

осуществляемой с помощью делителей напряжения R_{a1} , $R_{п1}$, R_{c2} , R_{a2} , $R_{п2}$, R_{c3} . С помощью регулировочного потенциометра $П_2$ производится настройка нуля усилителя.

На левой половине двойного триода $Л_1$ собран первый каскад усилителя, на пентоде $Л_2$ — второй. Третий каскад усилителя является выходным. Он собран по мостовой схеме на лучевом тетраде $Л_3$. На правой половине лампы $Л_1$ собрана схема компенсации дрейфа нуля. Сущность компенсации состоит в автоматическом изменении смещения на левой половине лампы $Л_1$ путем изменения тока в катодной нагрузке лампы. Изменение катодного тока производится с помощью правой половины лампы $Л_1$ в соответствии со случайным изменением ее катодного напряжения. Конденсаторы $C_{Ф1}$, $C_{Ф2}$, $C_{Ф3}$ совместно с сопротивлениями $R_{Ф2}$ и $R_{Ф3}$ являются фильтрами высоких частот.

Сопrotивление нагрузки включено в диагональ моста, плечами которого являются сопротивление R_{a3} , лампа L_3 и источники питания с напряжениями E_1 и E_4 . Благодаря такому включению нагрузки и при соответствующем выборе рабочей точки на характеристике усилителя достигается изменение знака выходного напряжения при изменении знака входного сигнала. Коэффициент усиления такого усилителя постоянного тока может иметь значение в пределах 40 000—50 000. Усилители постоянного тока находят широкое применение в моделирующих устройствах.

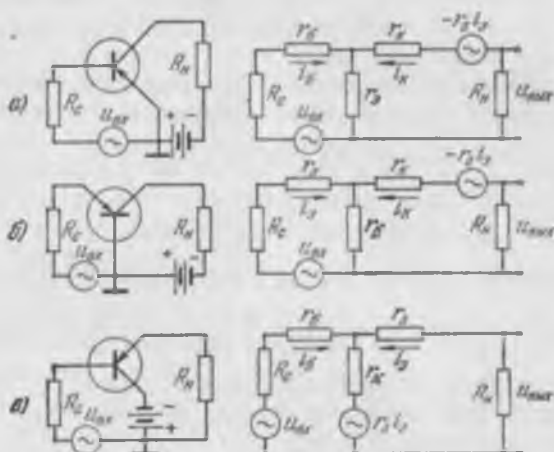


Рис. 3.9.3.

Иногда для усиления сигналов постоянного тока применяют усилители с предварительным преобразованием постоянного тока в переменный. После усиления сигнала на переменном токе производится (если это необходимо) обратное преобразование переменного тока в постоянный.

В последнее время в автоматических системах получили распространение также усилители на кристаллических триодах. На рис. 3.9.3 показаны схемы включения полупроводниковых триодов, а также эквивалентные схемы, которые используются для расчета основных характеристик усилителей, работающих в линейном режиме. Принцип действия полупроводниковых триодов, их характеристики и эквивалентные схемы описаны в [13].

Основными параметрами кристаллических триодов являются: сопротивление базы r_6 , сопротивление эмиттера $r_в$, сопротивление коллектора r_k и генераторное сопротивление $r_г$. Сопротивление базы r_6 обычно имеет порядок сотен Ом как для точечных, так и для плоскостных триодов. Сопротивление эмиттера $r_в$ имеет

порядок сотен для точечных и десятков Ом для плоскостных триодов. Сопротивление коллектора r_k имеет порядок десятков кОм для точечных и единиц МОм для плоскостных триодов. Генераторное сопротивление r_g равно нескольким десяткам кОм для точечных триодов и около МОм для плоскостных [13].

Внутренние сопротивления кристаллических триодов могут считаться чисто активными до частот порядка десятков кГц. При более высоких частотах переменного тока, идущего через полупроводник, начинает сказываться реактивное сопротивление. Кроме основных параметров, на схемах показаны сопротивление входной цепи R_c , сопротивление нагрузки R_n , входное напряжение $u_{вх}$, выходное напряжение $u_{вых}$.

Для усилителя с заземленным эмиттером, в соответствии с эквивалентной схемой, справедливы следующие соотношения:

$$\left. \begin{aligned} u_{вх} &= (R_c + r_g) i_g + r_g i_b, \\ -r_g i_g + i_k r_k + r_g i_b + u_{вых} &= 0, \\ i_g &= i_b + i_k, \quad u_{вых} = R_n i_k. \end{aligned} \right\} \quad (3.9.7)$$

Исключая переменные i_g , i_k , i_b , после преобразований получим

$$u_{вых} = \frac{(r_g - r_g) R_n}{r_g (r_g - r_g) + (R_c + r_g + r_g) (r_k - r_g + r_g + R_n)} u_{вх}. \quad (3.9.8)$$

Отсюда видно, что полупроводниковый триод с заземленным эмиттером является безынерционным усилителем. Наибольшее усиление по напряжению будет при $R_n = \infty$, $R_c = 0$:

$$k_{max} = \frac{r_g - r_g}{r_g + r_g}, \quad (3.9.9)$$

Величина k_{max} достигает 10^3 для точечных триодов и $2 \cdot 10^3$ для плоскостных триодов, что следует из приведенных выше порядков величин сопротивлений триодов.

Наибольший коэффициент усиления по напряжению имеют усилители с заземленным эмиттером. Усилители с заземленной базой имеют существенно меньший коэффициент усиления по напряжению, а коэффициент усиления по напряжению усилителей с заземленным коллектором меньше единицы. Последние аналогичны катодным повторителям. Наибольшим входным сопротивлением обладают усилители с заземленным коллектором порядка 1 МОм).

§ 3.10. Электромашинные усилители

Электромашинный усилитель (ЭМУ) представляет собой генератор постоянного тока, ротор которого вращается двигателем постоянного или переменного тока. ЭМУ предназначен для усиления маломощных электрических сигналов. Усиление входного

сигнала производится за счет энергии двигателя, вращающего ротор ЭМУ.

Различают *однокаскадные* и *двухкаскадные* ЭМУ. Однокаскадный ЭМУ — это обычный генератор постоянного тока с независимым возбуждением, конструктивно выполненный совместно с двигателем, вращающим его ротор. Такой усилитель обеспечивает усиление по мощности в несколько десятков раз. Наиболее широкое применение в автоматических системах нашли двухкаскадные ЭМУ. Рассмотрим их подробнее.

Двухкаскадный ЭМУ (рис. 3.10.1) — это генератор постоянного тока с двумя парами щеток на коллекторе. Одна пара щеток располагается на нейтральной линии, а другая — по линии, перпендикулярной нейтральной. Щетки на нейтральной линии называются *поперечными*. Две другие щетки называются *рабочими*. Поперечные щетки замкнуты накоротко. К рабочим щеткам подключается цепь нагрузки. Обмотка возбуждения с числом витков w_y в ЭМУ служит обмоткой управления, к которой подводится входной сигнал u_y .

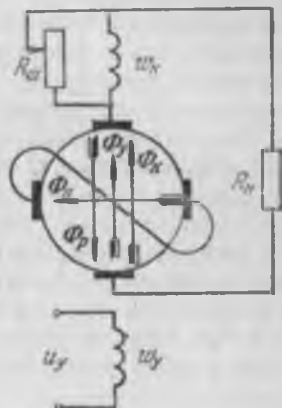


Рис. 3.10.1.

Принцип действия ЭМУ состоит в следующем. Управляющий сигнал u_y возбуждает магнитный поток Φ_y . При вращении якоря в этом потоке с постоянной скоростью Ω в якорной обмотке наводится э. д. с., которая снимается поперечными щетками. Вследствие того что поперечные щетки замкнуты накоротко, в их цепи возникает достаточно большой ток. Этот ток образует поперечный поток Φ_n , который сдвинут в пространстве на 90° относительно управляющего потока Φ_y . Неподвижный поперечный поток Φ_n наводит э. д. с. во вращающейся якорной обмотке. Эта э. д. с. снимается рабочими щетками. Рабочие щетки не снимают э. д. с., наводимой потоком Φ_y , так как установлены по оси, перпендикулярной нейтральной линии, что приводит к равной нулю сумме э. д. с., наведенных во всех проводниках якорной обмотки потоком Φ_y . Таким образом, в двухкаскадном электромашинном усилителе мощность усиливается в две ступени. Первая ступень — от обмотки управления до поперечной цепи. Вторая ступень — от поперечной цепи до рабочей. Каждая ступень усиливает мощность примерно в 100 раз, и следовательно, полное усиление мощности, которое дает ЭМУ, достигает 10 000.

При подключении нагрузки по рабочей цепи пойдет ток. В соответствии с ним возникает поток реакции якоря Φ_p , действующий навстречу потоку управления Φ_y . Поток Φ_p будет искажать поток

Подставляя сюда выражения (3.15.13) постоянных c_1 и c_2 , получаем окончательное выражение для передаточной функции $\Phi(s)$:

$$\Phi(s) = \frac{2Z_n(s) \gamma(s)}{[1 + Z_n(s) \gamma(s)] e^{\beta(s)l} - [1 - Z_n(s) \gamma(s)] e^{-\beta(s)l}}. \quad (3.15.15)$$

Выражение (3.15.15) передаточной функции $\Phi(s)$ показывает, что в общем случае однородная длинная линия представляет собой сложный динамический элемент, обладающий, в зависимости от соотношения величин ее основных параметров R , L , C , G и вида нагрузки $Z_n(s)$, в той или иной мере свойствами запаздывающего, колебательного и апериодического звеньев. Проанализируем, как изменяется передаточная функция $\Phi(s)$ в зависимости от изменения вида нагрузки и выбора определенных сочетаний основных параметров линии.

Нагрузка на конце длинной линии возникает только при соединении ее с каким-либо другим элементом автоматической системы. Поэтому влияние нагрузки на динамические свойства линии можно изучить методами главы 4, в которой показано, как по динамическим характеристикам отдельных звеньев определяются динамические характеристики их соединений. Для изучения динамических свойств самой длинной линии, ни с чем не соединенной, следует предположить, что нагрузка на конце линии отсутствует. Это равноценно случаю, когда сопротивление нагрузки R_n бесконечно велико, т. е. $R_n = \infty$. Полагая в (3.15.15) $Z_n(s) = \infty$, находим передаточную функцию длинной линии, рассматриваемой как изолированное звено автоматической системы:

$$\Phi(s) = \frac{2}{e^{\beta(s)l} + e^{-\beta(s)l}} = \frac{1}{\operatorname{ch} \beta(s)l} = \operatorname{sch} \beta(s)l. \quad (3.15.16)$$

Практически влиянием нагрузки на передаточную функцию длинной линии можно пренебречь в случае, когда сопротивление нагрузки очень велико, например, когда коаксиальный или двухпроводный кабель применяется для подачи напряжения на сетку электронной лампы, обладающей большим омическим нагрузочным сопротивлением (обычно порядка нескольких Мом).

Выбирая параметры длинной линии и нагрузку на ее конце соответствующим образом, можно получить линию с желательными динамическими свойствами.

В системах автоматического управления обычно применяются только так называемые *согласованные* линии, т. е. такие линии, которые не дают отражения прямого сигнала, приходящего к концу линии, от нагрузки. Линия становится согласованной, когда нагрузка $Z_n(s)$ выбирается равной волновому сопротивлению линии:

$$Z_n(s) = R_n + sL_n = \sqrt{\frac{R+sL}{G+sC}} = \frac{1}{\gamma(s)}. \quad (3.15.17)$$

Из этого равенства видно, что согласованности линий в общем случае можно добиться только при каком-нибудь одном значении s , так как это равенство не является тождеством относительно s . Поэтому для обеспечения согласованности линии обычно используются переменным током определенной частоты ω_0 и сопротивление нагрузки выбирают из условия (3.15.17) при $s = i\omega_0$. В этом случае при передаче сигналов переменным током частоты ω_0 линия будет работать в режиме бегущей волны. Полагая $\beta(i\omega_0) = \mu + iv$, получим

$$\Phi(i\omega_0) = e^{-\mu l - ivl}. \quad (3.15.18)$$

Длинная линия в этом случае работает как запаздывающее звено с временем запаздывания $\tau = vl/\omega_0$ и коэффициентом усиления $k = e^{-\mu l}$. Сигнал в такой линии распространяется со скоростью ω_0/v и ослабляется в e^μ раз на единицу длины линии.

Полной согласованности линии при любом значении s можно добиться, если выбрать ее параметры так, чтобы удовлетворялось условие $RC = LG$. Действительно, в этом случае $R/L = G/C$, $\gamma(s) = \sqrt{C/L}$ и условие (3.15.17) принимает вид

$$R_n + sL_n = \sqrt{L/C}. \quad (3.15.19)$$

Этому условию можно удовлетворить, приняв $R_n = \sqrt{L/C}$, $L_n = 0$. При этом выражение (3.15.15) для передаточной функции длинной линии примет вид

$$\Phi(s) = e^{-i\sqrt{RC}} \cdot e^{-s\sqrt{LC}}. \quad (3.15.20)$$

Эта формула показывает, что в этом случае длинная линия представляет собой запаздывающее звено с временем запаздывания $\tau = l\sqrt{LC}$ и коэффициентом ослабления сигнала $k = \exp\{-l\sqrt{RG}\}$. Сигнал распространяется по такой линии без искажения со скоростью, равной $1/\sqrt{LC}$, и ослабляется в $\exp\{\sqrt{RG}\}$ раз на каждую единицу длины линии. Такие линии обычно применяются в виде реальных линий задержки.

Чтобы получить значительное время задержки сигнала, необходимо выбрать большую длину линии (десятки — сотни метров), что может привести к большим трудностям в выборе основных параметров линии, позволяющих выполнить условия $RC = LG$. Кроме того, линия получается громоздкой. Поэтому для получения значительных времен задержки выбирают линию специальной конструкции, состоящей из последовательного соединения отдельных ячеек, имеющих значительные собственные индуктивности L_n и емкости C_n . Тогда $\tau_s = n\sqrt{L_n C_n}$, где n — число ячеек, входящих в линию задержки. Часто применяют непрерывную спиральную линию задержки, у которой изолированный

управления. Для устранения вредного влияния продольного потока в ЭМУ предусматривается компенсационная обмотка w_k , которая включается последовательно с обмоткой якоря и нагрузкой в цепь рабочих щеток. Компенсационную обмотку включают так, чтобы протекающий по ней ток нагрузки создавал компенсационный поток Φ_k , направленный против реакции якоря Φ_r . Когда поток Φ_k равен потоку Φ_r , усилитель работает в режиме полной компенсации. Если $\Phi_k < \Phi_r$ или $\Phi_k > \Phi_r$, то усилитель работает соответственно в режиме недокомпенсации или перекомпенсации. Для регулировки ЭМУ и установления необходимого значения потока Φ_k компенсационная обмотка шунтируется переменным сопротивлением $R_{ш}$.

Магнитный поток в обмотке управления пропорционален намагничивающей силе $\Phi_y = c_1 w_y i_y$. При вращении якоря с угловой скоростью Ω на поперечных щетках создается э. д. с., равная $e_p = k\Phi_y\Omega$. Если сопротивление короткозамкнутой цепи равно $R_{п}$, то в этой цепи потечет ток $i_{п} = e_p/R_{п}$. Этот ток создает магнитный поток $\Phi_{п} = c_2 w_{я} i_{п}$. Вращение якоря в потоке $\Phi_{п}$ приводит к появлению на рабочих щетках электродвижущей силы $e_p = k\Phi_{п}\Omega$. Таким образом, зависимость напряжения на рабочих щетках e_p от управляющего тока i_y выражается формулой

$$e_p = \frac{k^2 c_1 c_2 w_y w_{я} \Omega^2 i_y}{R_{п}}. \quad (3.10.1)$$

Зависимость (3.10.1) является *характеристикой холостого хода* ЭМУ. Так как все величины в правой части формулы (3.10.1), кроме i_y , постоянны для данного ЭМУ, то характеристика холостого хода ЭМУ линейна. Линейность характеристики сохраняется до значений i_y , при которых наступает насыщение магнитной системы усилителя.

Выходное напряжение ЭМУ u_p определяется формулой

$$u_p = e_p - i_p R_{вн}, \quad (3.10.2)$$

где i_p — ток в нагрузке, $R_{вн}$ — сопротивление участка рабочей цепи, в который входят: обмотка якоря, щетки и компенсационная обмотка, зашунтированная сопротивлением $R_{ш}$. Величина e_p в (3.10.2) в этом случае зависит не только от управляющего тока i_y , но и от тока в нагрузке i_p . Величина $R_{вн}$ зависит от доли сопротивления $R_{ш}$, подключенного к компенсационной обмотке. Зависимость u_p от тока i_p называется *внешней характеристикой* ЭМУ. При постоянном входном сигнале и соответственно подобранном положении движка переменного сопротивления $R_{ш}$ внешняя характеристика ЭМУ является линейной. Наклон прямой определяется степенью компенсации реакции якоря, которая в свою очередь зависит от сопротивления $R_{ш}$, подключенного к компенсационной обмотке. Обычно ЭМУ настраивают на недо-

компенсированный режим. Этому режиму соответствует отрицательный наклон внешней характеристики.

Усилительные свойства ЭМУ обычно характеризуются коэффициентом усиления по мощности

$$k_p = \frac{e_p^2}{RR_y i_p^2}, \quad (3.10.3)$$

где R — сопротивление рабочей цепи, R_y — сопротивление обмотки управления. Равенства (3.10.1) и (3.10.3) показывают, что коэффициент усиления по мощности k_p пропорционален четвертой степени угловой скорости ротора. Поэтому угловая скорость ротора ЭМУ сильно влияет на его коэффициент усиления. Это накладывает жесткие требования на стабилизацию угловой скорости Ω в ЭМУ, применяемых в автоматических системах управления. В современных ЭМУ $\Omega = 3000 \div 8000$ об/мин.

Определим динамические характеристики ЭМУ, предполагая для простоты, что он работает в режиме полной компенсации. Так как магнитная система усилителя не насыщена, то индуктивность цепей можно считать постоянной, а магнитные потоки — пропорциональными токам. Тогда э. д. с. самоиндукции e_c будет равна $e_c = -L di/dt$. На основании этих замечаний можно написать уравнения напряжений для рабочей, поперечной и управляющей цепей соответственно в виде

$$k_1 \Omega L_{\Pi} i_{\Pi} - L_p \frac{di_p}{dt} = i_p (R_{\Pi\Pi} + R_{\Pi}), \quad (3.10.4)$$

$$k_2 \Omega L_y i_y - L_{\Pi} \frac{di_{\Pi}}{dt} = i_{\Pi} R_{\Pi}, \quad (3.10.5)$$

$$u_y - L_y \frac{di_y}{dt} = i_y R_y, \quad (3.10.6)$$

где L_{Π} , L_p , L_y — индуктивности поперечной, рабочей (включая нагрузку) и управляющей цепей, R_{Π} — активное сопротивление нагрузки, $R_{\Pi\Pi}$, R_y — активные сопротивления поперечной и управляющей цепей, k_1 , k_2 — конструктивные постоянные ЭМУ, а i_{Π} , i_p , i_y — токи в поперечной, рабочей и управляющей цепях.

В качестве нагрузки электромашиного усилителя обычно используется исполнительный двигатель постоянного тока с независимым возбуждением. Скорость вращения вала такого двигателя регулируется путем изменения напряжения на якоре. Следовательно, выходным сигналом ЭМУ следует считать выходное напряжение, которое прикладывается к нагрузке, т. е.

$$u_p = i_p R_{\Pi}. \quad (3.10.7)$$

Э. д. с. самоиндукции и падение напряжения на внутреннем сопротивлении в рабочей цепи пренебрежимо малы по сравнению

внутренний кабель оплетается спиральным экраном. Такая линия имеет значительные собственные индуктивность и емкость при сравнительно короткой общей длине. Поэтому ослаблением амплитуды сигнала в ней можно пренебречь (так как величины l , R и G будут малыми, а L и C — большими).

§ 3.16. Объект управления как элемент автоматической системы

Как уже было отмечено в §§ 1.1 и 1.2, объект управления является одним из элементов автоматической системы, от которого существенно зависят динамические свойства и поведение всей системы в целом. Поэтому, изучая элементы автоматических систем и их динамические свойства, нельзя не коснуться и объектов управления.

Объектом управления может быть любая физическая система, и вследствие этого переменными, определяющими состояние объекта управления, могут быть любые физические величины: координаты, определяющие положения механических частей объекта управления, давления, плотности, температуры и скорости течения жидкостей и газов, токи, напряжения и другие электрические и электромагнитные величины, световые потоки, химические величины и т. д. Поэтому для описания поведения объекта управления в общем случае необходимо составить уравнения всех физических и химических процессов, связывающие интересующие нас величины, определяющие мгновенное состояние объекта управления, с управляющими воздействиями. Ввиду большого многообразия возможных объектов управления практически невозможно указать все возможные типы уравнений объекта и определяемых ими его динамических характеристик. Поэтому мы ограничимся здесь лишь рассмотрением некоторых объектов управления.

В некоторых случаях положение объекта управления однозначно определяется положением выходного вала или штока исполнительного устройства (т. е. объект управления практически жестко связан с выходной деталью исполнительного устройства). Так, например, положение подвижной пушечной установки на самолете полностью определяется углами поворота выходных валов исполнительных электродвигателей или гидродвигателей, с которыми установка связана зубчатыми или червячными передачами. В таких случаях за уравнение, описывающее поведение объекта управления, можно принять уравнение движения выходного элемента исполнительного устройства, добавив к его массе (моменту инерции) приведенную массу (соответственно момент инерции) объекта управления и включив в число действующих на выходной элемент исполнительного устройства сил приведенные силы, действующие на объект управления. Приведенные

массы, моменты инерции и силы определяются по известным формулам теории механизмов [7].

Для примера рассмотрим один канал управления подвижной пушечной установкой в случае, когда исполнительное устройство представляет собой электродвигатель постоянного тока с управлением напряжением в цепи якоря. Согласно изложенному мы должны в этом случае добавить к моменту инерции ротора двигателя в уравнении (3.13.5) приведенный момент инерции установки $k^2 J_1 / \eta$, где k — передаточное число от выходного вала двигателя к установке, η — коэффициент полезного действия передачи, а к действующим силам добавить приведенный к валу двигателя момент сил сопротивления, действующих на установку. Предполагая, что момент сил сопротивления пропорционален угловой скорости установки $k_Q \Omega$, получим в результате следующее уравнение движения ротора двигателя с установкой:

$$T \dot{\Omega} + \Omega = k_1 u_{вх}, \quad (3.16.1)$$

где

$$T = \frac{R_n \left(J + \frac{k^2 J_1}{\eta} \right)}{R_n c_T + c_e c_m \Phi_n^2 + k_Q R_n}, \quad k_1 = \frac{c_m \Phi_n}{R_n c_T + c_e c_m \Phi_n^2 + k_Q R_n}.$$

Во многих случаях исполнительные устройства приводят в движение только управляющие органы объекта управления, например рули самолета или ракеты. В таких случаях необходимо составить уравнения, определяющие динамику объекта управления, а в уравнениях исполнительного устройства учесть приведенную массу управляющих органов объекта управления и действующие на них приведенные силы.

Одним из важнейших объектов управления в авиации является летательный аппарат — самолет или крылатая ракета. Движение самолета или крылатой ракеты описывается известными уравнениями механики: уравнениями движения центра массы, уравнениями вращательного движения вокруг центра массы и кинематическими уравнениями. В полете на самолет или ракету действуют сила тяжести, сила тяги двигательной установки, аэродинамические силы и моменты и в некоторых случаях момент силы тяги двигательной установки. Вывод полных уравнений пространственного движения самолета или ракеты сложен и не входит в задачу этой книги. Поэтому мы ограничимся здесь выводом уравнений движения самолета в вертикальной плоскости.

Обозначим через m массу самолета (или ракеты), v — скорость его центра массы, θ — угол наклона вектора скорости к горизонту, ψ — угол наклона оси самолета к горизонту, называемый обычно *углом тангажа*, J_z — центральный момент инерции самолета относительно оси, перпендикулярной плоскости его движе-

с э. д. с., наведенной внешним магнитным полем. Поэтому уравнения (3.10.4), (3.10.5) и (3.10.6), описывающие работу ЭМУ, можно заменить приближенными уравнениями

$$u_p = k_1 \Omega L_n i_n, \quad L_n \frac{di_n}{dt} + R_n i_n = k_2 \Omega L_y i_y, \quad L_y \frac{di_y}{dt} + R_y i_y = u_y.$$

Исключая из этих уравнений i_n и i_y , получим дифференциальное уравнение в операторной форме, связывающее входное напряжение u_y и выходное напряжение u_p ЭМУ:

$$(T_n D + 1)(T_y D + 1) u_p = k_u u_y, \quad (3.10.8)$$

где

$$T_n = \frac{L_n}{R_n}, \quad T_y = \frac{L_y}{R_y}, \quad k_u = k_1 k_2 \Omega^2 T_n T_y. \quad (3.10.9)$$

Величина k_u представляет собой статический коэффициент усиления ЭМУ по напряжению, T_n и T_y являются соответственно постоянными времени поперечной цепи и цепи управления.

§ 3.11. Линейные функциональные преобразователи

В качестве функциональных преобразователей в автоматических системах можно применять любые автоматические системы и вычислительные устройства. Функциональный преобразователь с данным оператором можно спроектировать из различных элементов, как непрерывных, так и дискретных. В частности, в автоматических системах широко применяются в качестве функциональных преобразователей пассивные электрические цепи, составленные из омических сопротивлений, конденсаторов и индуктивностей.

Для определения динамических характеристик пассивной электрической цепи достаточно применить к каждому элементу цепи законы Кирхгофа и учесть, что ток i и напряжение u в активном сопротивлении R , емкости C и индуктивности L связаны соответственно уравнениями

$$u = Ri, \quad C \frac{du}{dt} = i, \quad u = L \frac{di}{dt}. \quad (3.11.1)$$

В результате получим дифференциальные уравнения цепи. Зная дифференциальные уравнения цепи, можно методом, изложенным в § 2.5, найти передаточную функцию цепи.

Можно также непосредственно найти передаточную функцию цепи, вычислив ее эквивалентное сопротивление и выходное напряжение с учетом того, что эквивалентные сопротивления конденсатора C и индуктивности L равны соответственно $1/Cs$ и Ls .

Получив выражение передаточной функции электрической цепи и задав необходимые значения ее параметров (т. е. коэффи-

коэффициентов при различных степенях s в числителе и знаменателе), можно подобрать сопротивления, емкости и индуктивности цепи таким образом, чтобы она имела заданную (или близкую к заданной) передаточную функцию.

Применим эти общие правила к цепочке, составленной из одного омического сопротивления R и одного конденсатора емкости C (рис. 3.11.1). Для этой цепочки имеем уравнения

$$u_{\text{вх}} = Ri, \quad u_{\text{вх}} = u_c + Ri, \quad i = C\dot{u}_c. \quad (3.11.2)$$

Исключая из этих уравнений u_c и i , получим дифференциальное уравнение цепи

$$T\dot{u}_{\text{вх}} + u_{\text{вх}} = Tu_{\text{вх}}, \quad T = RC. \quad (3.11.3)$$

Величина $T = RC$ называется *постоянной времени* цепочки RC .

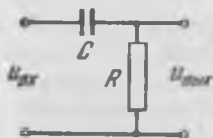


Рис. 3.11.1.

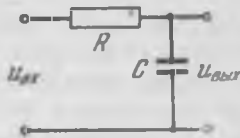


Рис. 3.11.2.

Для нахождения передаточной функции цепочки RC достаточно вычислить ее сопротивление $R + 1/Cs$ и ток $i = u_{\text{вх}}/(R + 1/Cs)$. Выходное напряжение $u_{\text{вх}}$ представляет собой падение напряжения на омическом сопротивлении $u_{\text{вх}} = Ri$. Подставляя сюда найденное значение тока и положив $u_{\text{вх}} = e^{st}$, $u_{\text{вх}} = \Phi(s)e^{st}$, найдем передаточную функцию цепочки RC :

$$\Phi(s) = \frac{Ts}{1 + Ts}. \quad (3.11.4)$$

Тот же результат получится, если определить передаточную функцию по дифференциальному уравнению (3.11.3) способом, изложенным в § 2.5.

При достаточно малой постоянной времени T второй член в знаменателе формулы (3.11.4) мал по сравнению с единицей и, следовательно, передаточная функция цепочки RC близка к передаточной функции идеального дифференциатора s , умноженной на T . Вследствие этого рассмотренная цепочка RC обычно применяется для приближенного определения производной медленно изменяющегося сигнала и называется *дифференцирующей*.

Если за выход цепочки RC принять напряжение на конденсаторе (рис. 3.11.2), то аналогично получим

$$C\dot{u}_{\text{вх}} = i, \quad u_{\text{вх}} = u_{\text{вх}} + Ri. \quad (3.11.5)$$

Подставляя выражение i из первого уравнения во второе, получим дифференциальное уравнение цепочки:

$$T\dot{u}_{\text{вых}} + u_{\text{вых}} = u_{\text{вх}}, \quad T = RC, \quad (3.11.6)$$

и ее передаточную функцию:

$$\Phi(s) = \frac{1}{1 + Ts}. \quad (3.11.7)$$

Таким образом, цепочка RC , изображенная на рис. 3.11.2, является апериодическим звеном (см. приложение 1). Написав формулу (3.11.7) в виде

$$\Phi(s) = \frac{\frac{1}{T}}{\frac{1}{T} + s},$$

убеждаемся в том, что при достаточно большой постоянной времени T передаточная функция цепочки RC близка к передаточной функции идеального интегратора $1/s$, умноженной на $1/T$. Вследствие этого цепочку RC , собранную по схеме рис. 3.11.2, обычно называют *интегрирующей*.

В случае, когда требуется получить линейную комбинацию сигнала и его производной, можно применить цепь, состоящую из конденсатора и двух сопротивлений, изображенную на рис. 3.11.3. Для этой цепи имеем

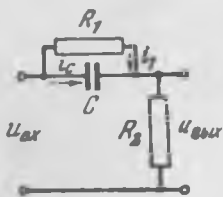


Рис. 3.11.3.

$$u_{\text{вых}} = R_2 i, \quad u_{\text{вх}} = u_c + u_{\text{вых}}, \quad i = i_c + i_1,$$

$$R_1 i_1 = u_c, \quad i_c = C \dot{u}_c.$$

Исключая из этих уравнений i , u_c , i_c , i_1 , получим дифференциальное уравнение цепи

$$k\tau \dot{u}_{\text{вых}} + u_{\text{вых}} = k(\tau \dot{u}_{\text{вх}} + u_{\text{вх}}), \quad k = \frac{R_2}{R_1 + R_2}, \quad \tau = R_1 C \quad (3.11.8)$$

и ее передаточную функцию

$$\Phi(s) = k \frac{1 + \tau s}{1 + k\tau s}. \quad (3.11.9)$$

Если сопротивление R_2 достаточно мало по сравнению с R_1 , то величина k будет малой и выходной сигнал цепи будет близок к $k(u_{\text{вх}} + \tau \dot{u}_{\text{вх}})$, т. е. к линейной комбинации входного сигнала и его производной. Обычно берут k от 0,1 до 0,3. Подбором величин R_1 и C можно обеспечить заданное значение отношения $1/\tau$ коэффициентов при входном сигнале и его производной. Что касается

абсолютных величин этих коэффициентов, то их можно сделать любыми, соединив выход цепи, изображенной на рис. 3.11.3, с электронным усилителем с соответствующим коэффициентом усиления.

Предоставляем читателю самостоятельно убедиться в том, что цепочки, составленные из одного омического сопротивления

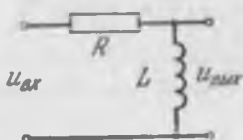


Рис. 3.11.4.

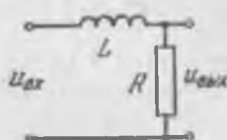


Рис. 3.11.5.

R и одной индуктивности L , изображенные на рис. 3.11.4 и 3.11.5, при $L/R = T$ равноценны цепочкам RC , представленным на рис. 3.11.1 и 3.11.2 соответственно.

Кроме пассивных электрических цепей, в качестве функциональных преобразователей в системах управления часто используются активные цепи с электронными усилителями. Эти цепи

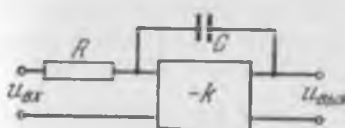


Рис. 3.11.6.

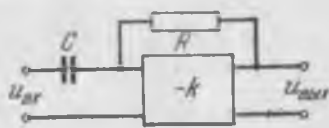


Рис. 3.11.7.

называются активными потому, что они используют не только энергию входного сигнала, но и энергию источников питания усилителей. Примеры активных цепей с усилителями даны на рис. 3.11.6 и 3.11.7.

При определении динамических характеристик активных цепей можно считать входные сопротивления усилителей достаточно большими, чтобы токи, ответвляемые в усилители, были пренебрежимо малыми по сравнению с токами во внешних цепях. Тогда, имея в виду, что, согласно изложенному в § 3.9 (рис. 3.9.1), нижний входной провод и нижний выходной провод усилителя на рис. 3.11.6 и 3.11.7 имеют одинаковый потенциал, для определения динамических характеристик активного функционального преобразователя достаточно будет, как и в случае пассивных цепей, написать вытекающие из основных законов электротехники соотношения для всех его внешних электрических цепей.

Применим изложенные общие правила к функциональному преобразователю с электронным усилителем, схема которого

представлена на рис. 3.11.6. Для этого преобразователя

$$u_{вх} + u_c + Ri = u_{вых}, \quad u_{вых} = -k(u_{вх} + Ri), \quad C\dot{u}_c = i.$$

Исключая из этих уравнений u_c и i , получим дифференциальное уравнение преобразователя:

$$\left(1 + \frac{1}{k}\right) RC\dot{u}_{вых} + \frac{1}{k} u_{вых} = -u_{вх}. \quad (3.11.10)$$

В случае усилителя с очень большим коэффициентом усиления k членами с коэффициентом $1/k$ можно пренебречь. Тогда уравнение (3.11.10) примет вид

$$\dot{u}_{вых} = -\frac{1}{RC} u_{вх}, \quad (3.11.11)$$

т. е. выходной сигнал функционального преобразователя будет пропорционален интегралу от входного сигнала.

Аналогично для активной цепи с электронным усилителем, собранной по схеме рис. 3.11.7, имеем

$$u_{вх} + u_c + Ri = u_{вых}, \quad u_{вых} = -k(u_{вх} + u_c), \quad C\dot{u}_c = i.$$

Исключая из этих уравнений i и u_c , получим дифференциальное уравнение цепи:

$$\frac{1}{k+1} RC\dot{u}_{вых} + u_{вых} = -\frac{k}{k+1} RC\dot{u}_{вх}. \quad (3.11.12)$$

При очень большом коэффициенте усиления k в уравнении (3.11.12) можно пренебречь первым членом и единицей в знаменателе правой части. Тогда уравнение (3.11.12) примет вид

$$u_{вых} = -RC\dot{u}_{вх}, \quad (3.11.13)$$

т. е. выходной сигнал цепи будет пропорционален производной входного сигнала.

Таким образом, усилитель постоянного тока с очень большим коэффициентом усиления, включенный по схеме рис. 3.11.6, является интегрирующим, а такой же усилитель, включенный по схеме рис. 3.11.7, является дифференцирующим. Такие усилители постоянного тока, называемые обычно *операционными*, широко применяются в моделирующих устройствах.

Функциональные преобразователи на переменном токе применяются редко. Чаще всего производится предварительное преобразование переменного тока в постоянный, функциональное преобразование сигнала на постоянном токе, а затем, если необходимо, обратное преобразование постоянного тока в переменный.

Аналогично можно построить нестационарные функциональные преобразователи, используя переменные сопротивления, емкости и индуктивности. Для составления дифференциальных

уравнений электрической цепи с переменными сопротивлениями, емкостями и индуктивностями достаточно применить ко всем элементам цепи основные законы электротехники, учитывая, что ток i и напряжение u в переменной емкости C и индуктивности L связаны соответственно уравнениями

$$i = \frac{d(Cu)}{dt}, \quad u = \frac{d(Li)}{dt}, \quad (3.11.14)$$

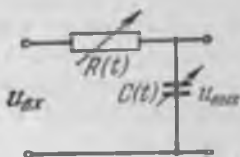


Рис. 3.11.8.

которыми следует в этом случае пользоваться вместо второго и третьего уравнения (3.11.1).

На рис. 3.11.8 показана пассивная цепь с переменными сопротивлением и емкостью. Для этой цепи справедливы уравнения

$$u_{\text{вх}} = iR + u_c, \quad u_{\text{вых}} = u_c, \quad i = i_c, \quad i_c = \dot{q}, \quad q = Cu_c.$$

Исключая из этих уравнений i , i_c , u_c , q , получим дифференциальное уравнение нестационарной цепи:

$$RC\dot{u}_{\text{вых}} + (RC + 1)u_{\text{вых}} = u_{\text{вх}}. \quad (3.11.15)$$

Для определения весовой функции этой нестационарной цепи следует на вход ее подать единичный импульс, т. е. заменить входное напряжение $u_{\text{вх}}$ единичной импульсной функцией $\delta(t - \tau)$. Тогда получим для весовой функции $g(t, \tau)$ уравнение

$$RC\dot{g}(t, \tau) + (RC + 1)g(t, \tau) = \delta(t - \tau). \quad (3.11.16)$$

Искомая весовая функция $g(t, \tau)$ представляет собой интеграл дифференциального уравнения (3.11.16), равный нулю при всех $\tau < t$. Легко проверить непосредственной подстановкой, что этот интеграл выражается формулой *)

$$g(t, \tau) = \begin{cases} 0 & \text{при } t < \tau, \\ f(t)\varphi(\tau) & \text{при } t > \tau, \end{cases} \quad (3.11.17)$$

где

$$\left. \begin{aligned} f(t) &= \exp \left\{ - \int_{t_0}^t \frac{RC + 1}{RC} dt \right\}, \\ \varphi(\tau) &= \frac{1}{R(\tau)C(\tau)} \exp \left\{ \int_{t_0}^{\tau} \frac{RC + 1}{RC} dt \right\}. \end{aligned} \right\} \quad (3.11.18)$$

Формулы (3.11.17) и (3.11.18) определяют весовую функцию нестационарной цепи при известных законах изменения величин

*) Формула (3.11.17) является частным случаем формулы (4.4.30), которая будет выведена в примере 4.4.1.

C и R со временем. Однако на практике часто возникает обратная задача. Весовая функция нестационарного линейного функционального преобразователя задается в форме (3.11.17) и требуется определить законы изменения сопротивления R и емкости C . Для решения этой задачи положим во второй формуле (3.11.18) $t = \tau$, после чего прологарифмируем формулы (3.11.18) и продифференцируем полученные равенства. Тогда получим дифференциальные уравнения первого порядка для R и C , которые легко интегрируются и в результате дают

$$\left. \begin{aligned} R &= \frac{a}{\varphi(t)} \exp \left\{ \int_{t_0}^t f(\eta) \varphi(\eta) d\eta \right\}, \\ C &= \frac{1}{af(t)} \exp \left\{ - \int_{t_0}^t f(\eta) \varphi(\eta) d\eta \right\}, \end{aligned} \right\} \quad (3.11.19)$$

где a — произвольная постоянная.

Формулы (3.11.19) определяют закон изменения (программу) сопротивления и емкости нестационарной цепи (рис. 3.11.8) по данным функциям $f(t)$ и $\varphi(t)$.

§ 3.12. Импульсные линейные элементы

В § 1.6 было отмечено, что в технике находят широкое применение дискретные (импульсные) системы управления. Характерной

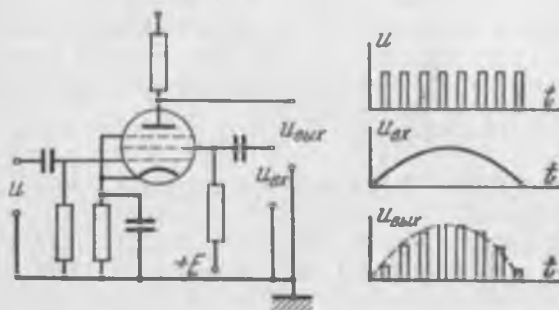


Рис. 3.12.1.

особенностью таких систем является наличие импульсных элементов. *Импульсным* называется такой элемент, который преобразует непрерывный входной сигнал в последовательность кратковременных импульсов.

Линейный импульсный элемент осуществляет *амплитудную модуляцию*, т. е. преобразует непрерывный входной сигнал в после-

довательность импульсов, амплитуды которых пропорциональны значениям входного сигнала в дискретные моменты времени. При этом форма импульсов может быть различной, например прямоугольной или треугольной.

На рис. 3.12.1 показан линейный импульсный элемент с электронной лампой. На управляющую сетку лампы поступает периодическая последовательность прямоугольных импульсов одинаковой амплитуды u (nT_n). На экранирующую сетку лампы наряду

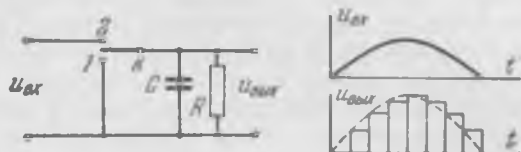


Рис. 3.12.2.

с постоянным напряжением, обеспечивающим необходимый режим работы лампы, подается напряжение $u_{вх}(t)$, являющееся входным сигналом. Импульсы, действующие на управляющую сетку, периодически отпирают и запирают лампу. Величина анодного тока открытой лампы определяется величиной входного сигнала $u_{вх}(t)$, приложенного к экранирующей сетке. Таким образом, на сопротивлении нагрузки будет получаться выходное напряжение $u_{вых}(nT_n)$, представляющее собой последовательность импульсов, амплитуды которых пропорциональны значениям входного сигнала в дискретные моменты времени $t = nT_n$.

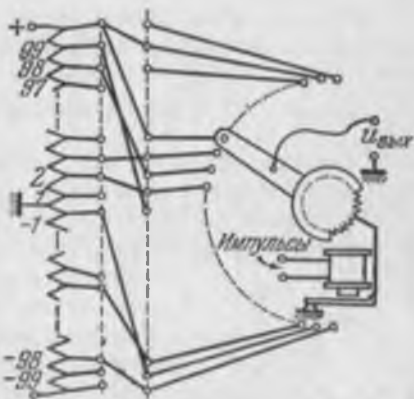


Рис. 3.12.3.

На рис. 3.12.2 приведена схема другого линейного импульсного элемента, называемого *фиксатором*. Фиксатор предназначен для преобразования непрерывного входного сигнала $u_{вх}(t)$ в последовательность прямоугольных импульсов или в ступенчатую функцию. Этот элемент фиксирует значения входного сигнала в дискретные моменты $t = nT_n$ и «помнит» каждое значение в течение периода T_n . Ключ K замыкается на короткий промежуток времени с клеммой 1, а затем с клеммой 2. В момент замыкания ключа с клеммой 1 с конденсатора сбрасывается заряд, при замыкании ключа с клеммой 2 конденсатор заряжается до текущего значения

$u_{bx}(t)$. В течение периода T_n ключ находится в нейтральном состоянии, конденсатор сохраняет полученный заряд.

В моделирующих устройствах широкое применение нашли шаговые искатели, применяющиеся для реализации переменных коэффициентов уравнений или заданных возмущений. На рис. 3.12.3 представлен элемент, основой которого является 200-ламельный шаговый искатель. К обмотке электромагнита (шагового реле) подводятся импульсы тока, под действием которых электромагнит дискретно с помощью храпового колеса переводит токосъемную щетку с одной ламели на другую. Значения напряжений, подводимых к ламелям от делителя напряжений, могут быть набраны ступенчато по любому закону. На делителе имеется сто интервалов деления положительного напряжения и сто интервалов отрицательного. Коммутация клемм делителя и ламелей осуществляется на специальном наборном поле. На выходе устройства формируется ступенчатая функция, дискретные значения которой определяются номером импульса, поданного на реле, и величиной потенциала, подведенного к клемме делителя напряжения, с которой в данный момент соединена токосъемная щетка искателя. Рассмотренное устройство с шаговым искателем является линейным импульсным элементом. Оно преобразует непрерывную функцию в последовательность импульсов с амплитудами, пропорциональными дискретным значениям функции.

§ 3.13. Двигатели постоянного тока

В системах управления в качестве исполнительных устройств часто применяются двигатели постоянного тока.

Управление двигателем постоянного тока производится путем изменения напряжения $u_{я}$, приложенного к цепи якоря, при неизменном потоке возбуждения (рис. 3.13.1) или путем изменения потока возбуждения при неизменном напряжении $u_{я}$, приложенном к цепи якоря.

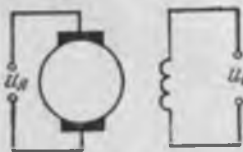


Рис. 3.13.1.

Из теории электрических машин известно, что ток в цепи якоря $i_{я}$, поток возбуждения $\Phi_{в}$, вращающий момент $M_{вр}$, скорость вращения вала двигателя Ω и противозлектродвижущая сила $e_{ц}$, наводимая в обмотке

якоря, в установившемся движении связаны следующими зависимостями:

$$e_{ц} = c_e \Phi_{в} \Omega, \quad M_{вр} = c_m \Phi_{в} i_{я}, \quad u_{я} = i_{я} R_{я} + e_{ц}, \quad (3.13.1)$$

где

$$c_e = \frac{PN \cdot 10^{-8}}{2\pi a}, \quad c_m = \frac{c_e}{9,81},$$

а N — число проводников якоря, a — число пар параллельных ветвей, P — число пар полюсов, $R_{\text{я}}$ — активное сопротивление цепи якоря. Исключая из уравнений (3.13.1) величины $e_{\text{я}}$ и $i_{\text{я}}$, получим

$$M_{\text{вр}} = \frac{c_m}{R_{\text{я}}} \Phi_{\text{в}} u_{\text{я}} - \frac{c_e c_m}{R_{\text{я}}} \Phi_{\text{в}}^2 \Omega. \quad (3.13.2)$$

Эта формула определяет *механическую характеристику* двигателя постоянного тока.

При управлении путем изменения напряжения, приложенного к цепи якоря, поток возбуждения $\Phi_{\text{в}}$ постоянен, а управляющим входным сигналом является $u_{\text{я}}$. В этом случае, как видно из (3.13.2), вращающий момент двигателя линейно связан с управляющим сигналом. Это обеспечивает возможность плавного изменения скорости двигателя в широком диапазоне.

При управлении путем изменения потока возбуждения $u_{\text{я}} = \text{const}$, а управляющим сигналом является $\Phi_{\text{в}}$. В этом случае вращающий момент нелинейно зависит от управляющего сигнала. Однако практически при управлении двигателем с помощью потока возбуждения предусматриваются меры (в цепь якоря включается добавочное сопротивление), обеспечивающие независимость тока в якорной цепи от угловой скорости вала Ω . Поэтому в этом случае можно пренебречь вторым слагаемым в правой части формулы (3.13.2) и принять линейную зависимость вращающего момента от управляющего потока:

$$M_{\text{вр}} \approx \frac{c_m}{R_{\text{я}}} \Phi_{\text{в}} u_{\text{я}}. \quad (3.13.3)$$

Для определения динамических характеристик двигателя составим уравнение движения ротора двигателя:

$$J \dot{\Omega} = M_{\text{вр}} - c_{\tau} \Omega - M_{\text{с}}, \quad (3.13.4)$$

где J — момент инерции всех вращающихся масс, приведенный к ротору двигателя, c_{τ} — коэффициент вязкого трения, $M_{\text{с}}$ — момент сопротивления *). Подставляя в уравнение (3.13.4) выражение (3.13.2) вращающего момента, получим

$$J \dot{\Omega} + \left(c_{\tau} + \frac{c_e c_m}{R_{\text{я}}} \Phi_{\text{в}}^2 \right) \Omega = \frac{c_m}{R_{\text{я}}} \Phi_{\text{в}} u_{\text{я}} - M_{\text{с}}. \quad (3.13.5)$$

Выходной величиной может быть или угловая скорость вала Ω , или угол поворота ротора двигателя α , определяемый кинематическим уравнением

$$\dot{\alpha} = \Omega. \quad (3.13.6)$$

*) Относительно приведения масс см. любой курс теории механизмов, например [7].

При управлении первым способом входным сигналом двигателя является напряжение $u_{я}$. В этом случае уравнение (3.13.5) может быть представлено в виде

$$(T_1 D + 1) \Omega = k_1 u_{я} - k_c M_c, \quad (3.13.7)$$

где

$$T_1 = \frac{J R_{я}}{R_{я} c_T + c_e c_m \Phi_{я}^2}, \quad k_1 = \frac{c_m \Phi_{я}}{R_{я} c_T + c_e c_m \Phi_{я}^2}, \quad k_c = \frac{R_{я}}{R_{я} c_T + c_e c_m \Phi_{я}^2}. \quad (3.13.8)$$

При управлении вторым способом входным сигналом двигателя является напряжение $u_{в}$, приложенное к обмотке возбуждения. Управляющий поток $\Phi_{в}$ и управляющее напряжение $u_{в}$ связаны уравнениями

$$\Phi_{в} = \frac{L_{в} i_{в}}{w_{в}}, \quad u_{в} = i_{в} R_{в} + L_{в} \frac{di_{в}}{dt}. \quad (3.13.9)$$

Исключая из уравнений (3.13.3), (3.13.4) и (3.13.9) величины $\Phi_{я}$, $i_{я}$, $M_{вр}$ и считая момент сопротивления M_c постоянным, получим

$$(T_2 D + 1)(T_{в} D + 1) \Omega = k_2 u_{в} - k_{м} M_c, \quad (3.13.10)$$

где

$$T_2 = \frac{J}{c_T}, \quad T_{в} = \frac{L_{в}}{R_{в}}, \quad k_2 = \frac{c_m u_{я} L_{в}}{w_{в} c_T R_{я} R_{в}}, \quad k_{м} = \frac{1}{c_T}. \quad (3.13.11)$$

Уравнения (3.13.7) и (3.13.10) написаны для случая, когда выходным сигналом двигателя является угловая скорость Ω . Подставляя в эти уравнения выражение Ω из уравнения (3.13.6), можно получить уравнения для случая, когда выходным сигналом считается угол поворота вала α .

Из уравнений (3.13.7) и (3.13.10) следует, что двигатель постоянного тока с независимым возбуждением является линейной динамической системой. Сравнение динамических характеристик двигателей при первом и втором способах управления позволяет сделать вывод, что быстродействие двигателя при управлении изменением потока возбуждения меньше, чем при управлении изменением напряжения на якоре. Это объясняется, во-первых, тем, что при управлении вторым способом на быстродействие двигателя оказывает влияние не только электрохимическая постоянная времени T_2 , но и постоянная времени цепи возбуждения $T_{в}$. Во-вторых, электрохимическая постоянная времени двигателя при втором способе управления, как показывает сравнение формул (3.13.8) и (3.13.11), больше, чем при первом способе.

Кроме большего быстродействия управление путем изменения напряжения, приложенного к цепи якоря, дает более плавное изменение скорости вращения двигателя и больший диапазон

изменения скорости. Эти обстоятельства позволяют отдать предпочтение способу управления двигателем путем изменения напряжения на якоре. При управлении этим способом требуется большое усиление управляющего сигнала по мощности, так как ток в якровой цепи должен иметь большую величину. Обычно для этих целей используется электромашиный усилитель. При управлении с помощью обмотки возбуждения в качестве усилителей могут использоваться усилители на электронных лампах, так как токи управления в этом случае малы.

§ 3.14. Двигатели переменного тока

Весьма широкое применение в качестве исполнительных устройств в системах управления находят двухфазные асинхронные двигатели. Особенно широкое распространение получили двухфазные асинхронные двигатели с тонкостенным ротором. Этому способствует целый ряд достоинств таких двигателей: отсутствие щеток и коллектора, незначительный момент инерции ротора, простота и стабильность усиления управляющего сигнала при помощи усилителей переменного тока и т. д.

Ротор двухфазного асинхронного двигателя выполняется в виде «беличьего колеса», образованного несколькими параллельными проводниками, замкнутыми накоротко торцевыми кольцами, или в виде тонкостенного стакана, сделанного из немагнитного материала. Статор двигателя имеет две обмотки, которые укладываются в пазы пакета из железных пластин так, чтобы магнитные оси их были взаимно перпендикулярными (рис. 3.14.1). Одна из обмоток статора является возбуждающей и подключается к источнику питания переменного тока с фиксированным напряжением. Вторая обмотка является управляющей и обычно питается напряжением переменного тока той же частоты, что и напряжение возбуждения, но через управляющий усилитель. Управляющее напряжение сдвинуто по фазе на 90° относительно напряжения возбуждения. Это приводит к вращению магнитного поля. Вращающееся магнитное поле индуцирует в стенках ротора вихревые токи, которые, взаимодействуя с магнитным потоком, обуславливают появление вращающего момента, увлекающего ротор в сторону вращения магнитного поля. Для получения фазового сдвига в 90° последовательно с управляющей обмоткой включают емкость, как показано на рис. 3.14.1.

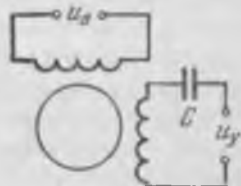


Рис. 3.14.1.

Если нагрузка на оси якоря отсутствует, то скорость вращения якоря Ω будет равна скорости вращения магнитного поля Ω_m .

При наличии момента нагрузки скорость вращения якоря замедляется до такой величины, которая необходима для создания вращающего момента, равного моменту нагрузки. Разность $\Omega_m - \Omega$ называется *скольжением* асинхронного двигателя. При постоянном моменте нагрузки скольжение постоянно и скорость вращения якоря также постоянна. Для изменения направления вращения ротора достаточно изменить фазу управляющего напряжения на 180° .

Намагничивающая сила вращающегося ротора взаимодействует с намагничивающей силой статора так же, как намагни-

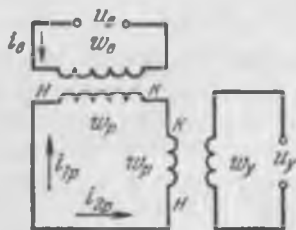


Рис. 3.14.2.

чивающая сила неподвижного ротора, если токи в неподвижном роторе остаются по величине и фазе (относительно э. д. с.) теми же, что и при вращении. Это позволяет представить ротор в виде неподвижной двухфазной обмотки. Магнитная ось одной фазы такой эквивалентной обмотки ротора параллельна магнитной оси обмотки возбуждения статора, ось другой фазы параллельна магнитной оси управляющей обмотки

статора (рис. 3.14.2). Пользуясь, как и везде, стандартными обозначениями для токов и чисел витков обмоток с соответствующими индексами, можем написать следующее выражение магнитного потока, создаваемого обмоткой возбуждения и эквивалентной обмоткой ротора:

$$\Phi_{py} = \frac{i_p w_p + i_{1p} w_p}{R_m}, \quad (3.14.1)$$

где, в дополнение к стандартным обозначениям, i_{1p} — ток в первой фазе эквивалентной обмотки ротора, R_m — магнитное сопротивление в цепи магнитного потока. Аналогичной формулой определяется магнитный поток, создаваемый обмоткой управления и эквивалентной обмоткой ротора:

$$\Phi_{py} = \frac{i_y w_y + i_{2p} w_p}{R_m}, \quad (3.14.2)$$

где i_{2p} — ток во второй фазе эквивалентной обмотки ротора. Обычно $w_y = w_v = w_c$, т. е. обмотки статора одинаковы.

Вращающий момент $M_{вр}$ является результатом взаимодействия тока i_{1p} с потоком Φ_{py} и тока i_{2p} с потоком Φ_{py} :

$$M_{вр} = w_p (i_{1p} \Phi_{py} - i_{2p} \Phi_{py}).$$

Подставляя сюда выражения потоков (3.14.1) и (3.14.2), получим

$$M_{вр} = M (i_{1p} i_y - i_{2p} i_v), \quad (3.14.3)$$

где $M = w_p w_c / R_m$ — коэффициент взаимной индукции.

Пренебрегая потоками рассеяния, реактивным сопротивлением обмоток статора и ротора, влиянием обмоток ротора на обмотки статора и э. д. с., наведенной обмоткой управления в роторе при его вращении, и учитывая (3.14.1) и (3.14.2), можем записать уравнения, описывающие процессы в обмотках статора и ротора, в виде

$$\left. \begin{aligned} u_{\text{вх}} \cos \omega t &= i_y R_c, & U_0 \sin \omega t &= i_b R_c, \\ M \frac{di_b}{dt} &= i_{1p} R_p, & M \frac{di_y}{dt} &= i_{2p} R_p - M \Omega i_b, \end{aligned} \right\} \quad (3.14.4)$$

где $u_{\text{вх}} \cos \omega t$ и $U_0 \sin \omega t$ — соответственно управляющее напряжение, модулированное входным сигналом $u_{\text{вх}}$, и напряжение возбуждения. Решая уравнения (3.14.4) относительно i_b , i_y , i_{1p} , i_{2p} и подставляя результат в (3.14.3), получим следующее выражение для вращающего момента:

$$M_{\text{вр}} = \frac{M^2 \omega U_0}{R_p R_c^2} \left(u_{\text{вх}} - \frac{U_0}{2\omega} \Omega \right) + \frac{M^2 U_0^2}{2 R_p R_c^2} \Omega \cos 2\omega t. \quad (3.14.5)$$

Учтем теперь, что ротор двигателя, вследствие своей инерционности, практически не будет реагировать на колебания вращающего момента удвоенной несущей частоты 2ω . Вследствие этого в (3.14.5) можно пренебречь колеблющейся составляющей вращающего момента. Тогда для вращающего момента получим выражение, аналогичное выражению (3.13.2). Следовательно, работа двигателя переменного тока описывается уравнением (3.13.7), в котором

$$\left. \begin{aligned} u_y &= u_{\text{вх}}, & k_1 &= \frac{2M^2 \omega U_0}{2R_p R_c^2 c_T + M^2 U_0^2}, \\ k_c &= \frac{2R_p R_c^2}{2R_p R_c^2 c_T + M^2 U_0^2}, & T_1 &= \frac{2J R_p R_c^2}{2R_p R_c^2 c_T + M^2 U_0^2}. \end{aligned} \right\} \quad (3.14.6)$$

§ 3.15. Длинные линии

В современных системах автоматического управления отдельные блоки и агрегаты часто располагаются на значительных расстояниях друг от друга. Для соединения элементов в единую систему широко применяются специальные линии передачи сигналов. Для этого используются двухпроводные или однопроводные линии, экранированные кабели, фидеры и волноводы. У подобных элементов основные электрические характеристики — сопротивление, емкость, индуктивность — не сосредоточены в каких-то отдельных частях, а распределены непрерывно вдоль линии. Такие элементы обычно называются *элементами с распределенными*

параметрами. К элементам с распределенными параметрами можно также отнести различные линии задержки.

Строго говоря, почти все реальные элементы, входящие в состав систем автоматического управления, также представляют собою элементы с распределенными параметрами. Но в подавляющем большинстве случаев реальный элемент можно с достаточной степенью точности заменить упрощенной моделью — системой с сосредоточенными параметрами и идеальными линиями связи, в которых отсутствуют потери энергии и запаздывание при прохождении сигналов. Такая замена дает возможность определить динамические характеристики элемента с достаточной точностью лишь в том случае, когда эквивалентное сопротивление всех линий связи достаточно мало по сравнению с эквивалентным сопротивлением соединяемых ими деталей. Все перечисленные выше элементы характерны тем, что они представляют собой чистые линии связи, у которых распределение параметров вдоль линии существенно влияет на их динамические свойства. Подобные элементы, так же как и элементы, у которых параметры распределены непрерывно вдоль поверхности или по объему, необходимо рассматривать как системы с распределенными параметрами.

Проводные линии связи, кабели, фидеры, волноводы и линии задержки представляют собой различные типы так называемых *длинных линий* и описываются одними и теми же уравнениями. Поэтому мы будем определять динамические характеристики длинной линии независимо от ее конкретной технической реализации.

Любая длинная линия характеризуется следующими распределенными по ее длине параметрами: активным сопротивлением, индуктивностью, емкостью и активной проводимостью между проводами линии, обусловленной наличием определенного конечного омического сопротивления их изоляции и промежутка между ними. В общем случае эти параметры могут быть распределены по длине линии неравномерно. Такие линии называются *неоднородными*. Если перечисленные параметры распределены вдоль линии равномерно, то линия называется *однородной*. На практике в подавляющем большинстве случаев применяются только однородные длинные линии.

Обозначим через R , L , C и G соответственно омическое сопротивление, индуктивность, емкость и межпроводную проводимость единицы длины однородной линии. Напряжение и ток в каждом сечении линии будут функциями времени t и координаты сечения x : $u = u(t, x)$, $i = i(t, x)$. Условимся считать, что началу линии соответствует $x = 0$, а концу — $x = l$, где l — полная длина линии.

Для составления системы дифференциальных уравнений, описывающих длинную линию, представим ее бесконечно малый отрезок длины dx упрощенной моделью, изображенной на

рис. 3.15.1. Изменение напряжения между проводами линии от сечения x до сечения $x + dx$ произойдет из-за падения напряжения на активном сопротивлении $R dx$ и на индуктивности $L dx$ от резка dx :

$$u(t, x + dx) - u(t, x) = -iR dx - L \frac{\partial i}{\partial t} dx.$$

С другой стороны, с точностью до бесконечно малых второго порядка

$$u(t, x + dx) - u(t, x) = \frac{\partial u}{\partial x} dx.$$

Таким образом, имеем

$$\frac{\partial u}{\partial x} dx = -iR dx - L \frac{\partial i}{\partial t} dx. \quad (3.15.1)$$

Изменение тока, протекающего в линии между теми же сечениями,

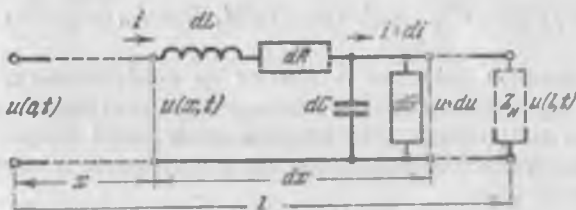


Рис. 3.15.1.

вызывается утечкой тока через проводимость $G dx$ и током заряда емкости $C dx$:

$$i(t, x + dx) - i(t, x) = -Gu dx - C \frac{\partial u}{\partial t} dx.$$

С другой стороны, с точностью до бесконечно малых второго порядка

$$i(t, x + dx) - i(t, x) = \frac{\partial i}{\partial x} dx.$$

Таким образом,

$$\frac{\partial i}{\partial x} dx = -Gu dx - C \frac{\partial u}{\partial t} dx. \quad (3.15.2)$$

Сократив уравнения (3.15.1) и (3.15.2) на dx , получим систему двух линейных дифференциальных уравнений в частных производных, описывающую процесс в длинной линии при действии на ее входе любого напряжения и при произвольной нагрузке на ее конце (выходе):

$$u'_x = -Ri - Li', \quad i'_x = -Gu - Ci'. \quad (3.15.3)$$

пия, x — горизонтальное перемещение центра массы самолета, а y — высоту его полета (рис. 3.16.1). Уравнения движения центра массы самолета будем рассматривать в проекциях на касательную и нормаль к его траектории. Как известно из теоретической механики, касательная и нормальная составляющие вектора ускорения точки в плоском движении равны соответственно $\dot{v}$ и $v\dot{\theta}$. Поэтому уравнения движения центра массы самолета в проекциях на касательную и нормаль к его траектории имеют вид

$$\left. \begin{aligned} m\dot{v} &= T \cos \alpha - X - mg \sin \theta, \\ mv\dot{\theta} &= T \sin \alpha + Y - mg \cos \theta, \end{aligned} \right\} \quad (3.16.2)$$

где X — касательная составляющая главного вектора аэродинамических сил, называемая обычно *силой лобового сопротивления*,

Y — нормальная составляющая главного вектора аэродинамических сил, называемая *подъемной силой*, T — сила тяги двигательной установки, а $\alpha = \vartheta - \theta$ — угол между вектором скорости центра массы самолета и его продольной осью, называемый *углом атаки*. При малых значениях угла атаки α можно принять в уравнениях (3.16.2) $\sin \alpha = \alpha$, $\cos \alpha = 1$.

Для полного определения движения центра массы самолета необходимо добавить к динамическим уравнениям (3.16.2) кинематические уравнения движения центра массы:

$$\dot{x} = v \cos \theta, \quad \dot{y} = v \sin \theta. \quad (3.16.3)$$

Наконец, уравнение движения самолета вокруг центра массы в вертикальной плоскости имеет вид

$$J_z \ddot{\vartheta} = M_z + M_T \quad (3.16.4)$$

где M_z — аэродинамический момент, а M_T — момент силы тяги двигательной установки.

Первое уравнение (3.16.2) определяет скорость полета самолета и поэтому должно рассматриваться отдельно как уравнение объекта управления в системе управления скоростью. Эта система обеспечивает такое управление тягой двигателя, при котором скорость v изменяется по заданному закону или, в частном случае, сохраняет определенное постоянное значение. Поэтому для системы управления полетом, обеспечивающей необходимое изменение направления полета, первое уравнение (3.16.3) роли не играет,

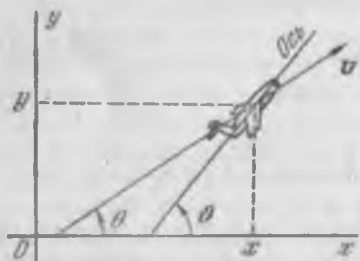


Рис. 3.16.1.

и при исследовании процесса управления скорость полета v в остальных уравнениях можно считать известной функцией времени или постоянной.

Подъемная сила Y и аэродинамический момент M_z обычно выражаются формулами

$$Y = c_y S \frac{\rho v^2}{2}, \quad M_z = m_z l S \frac{\rho v^2}{2}, \quad (3.16.5)$$

где ρ — плотность воздуха, S — некоторая характерная площадь (для крылатых летательных аппаратов обычно площадь крыльев), l — характерный линейный размер (для крылатых летательных аппаратов — средняя хорда крыла), а c_y и m_z — безразмерные аэродинамические коэффициенты, определяемые экспериментально.

Аэродинамический коэффициент подъемной силы c_y при малых углах атаки можно приближенно считать пропорциональным углу атаки α :

$$c_y = c_y^\alpha \alpha. \quad (3.16.6)$$

Аэродинамический момент M_z представляет собой сумму трех моментов: момента крыльев и фюзеляжа, который при малом угле атаки приближенно можно считать линейной функцией угла атаки, добавочного момента, создаваемого рулем высоты, приближенно пропорционального углу отклонения руля высоты δ , и момента сопротивления, приближенно пропорционального угловой скорости самолета $\omega_z = \dot{\theta}$. В соответствии с этим аэродинамический коэффициент m_z обычно выражают приближенной формулой

$$m_z = m_z^\alpha + m_z^\alpha \alpha + m_z^{\omega_z} \frac{l \omega_z}{v} + m_z^\delta \delta. \quad (3.16.7)$$

На основании формул (3.16.5), (3.16.6) и (3.16.7) второе уравнение (3.16.2) и уравнение (3.16.4) при малых углах атаки α можно переписать в виде

$$\left. \begin{aligned} \ddot{\theta} &= A_\alpha \alpha - \frac{g \cos \theta}{v}, \\ \ddot{\theta} + a_\theta \dot{\theta} + a_\alpha \alpha &= a_0 - a_\delta \delta, \end{aligned} \right\} \quad (3.16.8)$$

где

$$\left. \begin{aligned} A_\alpha &= \frac{c_y^\alpha S \rho v}{2m} + \frac{T}{mv}, & a_\theta &= -\frac{m_z^{\omega_z} l^2 S \rho v}{2J_z}, \\ a_\alpha &= -\frac{m_z^\alpha l S \rho v^2}{2J_z}, & a_\delta &= -\frac{m_z^\delta l S \rho v^2}{2J_z}, \\ a_0 &= \frac{M_T}{J_z} + \frac{m_z^\alpha l S \rho v^2}{2J_z}. \end{aligned} \right\} \quad (3.16.9)$$

Время t не входит явно в уравнения однородной линии (3.15.3). Поэтому процесс в ней при одном и том же входном напряжении протекает всегда одинаково, независимо от момента его начала. Это означает, что длинная линия представляет собой стационарную линейную систему. Следовательно, ее динамические свойства полностью определяются ее передаточной функцией $\Phi(s)$. Для определения передаточной функции $\Phi(s)$ приложим ко входу линии напряжение, изменяющееся по показательному закону $u(t, 0) = e^{st}$, и будем искать решение системы уравнений (3.15.3) в виде

$$u(t, x) = \varphi(s, x) e^{st}, \quad i(t, x) = \psi(s, x) e^{st}, \quad (3.15.4)$$

где $\varphi(s, x)$ и $\psi(s, x)$ — некоторые неизвестные функции, не зависящие от времени. При этом учтем граничные условия в начале и конце линии:

$$u(t, 0) = e^{st}, \quad u(t, l) = R_{\text{н}} i(t, l) + L_{\text{н}} \frac{\partial i(t, l)}{\partial t}. \quad (3.15.5)$$

Второе граничное условие выбрано из соображения, что длинные линии, применяющиеся в системах автоматического управления, обычно нагружены или чистым омическим сопротивлением, или последовательным соединением омического и индуктивного сопротивлений.

Подставляя выражения $u(t, x)$, $i(t, x)$ из (3.15.4) в уравнения (3.15.3), после сокращения на e^{st} получаем для определения функций $\varphi(s, x)$ и $\psi(s, x)$ уравнения

$$\left. \begin{aligned} \varphi'_x(s, x) &= -(R + Ls) \psi(s, x), \\ \psi'_x(s, x) &= -(G + Cs) \varphi(s, x). \end{aligned} \right\} \quad (3.15.6)$$

Эти уравнения образуют систему двух обыкновенных линейных дифференциальных уравнений с постоянными коэффициентами, зависящими от параметра s . Применяя обычный метод интегрирования систем линейных уравнений с постоянными коэффициентами, составляем характеристическое уравнение:

$$\begin{vmatrix} -k & -(R + Ls) \\ -(G + Cs) & -k \end{vmatrix} = 0$$

или

$$k^2 - (R + Ls)(G + Cs) = 0. \quad (3.15.7)$$

Отсюда, вводя обозначение

$$\beta(s) = \sqrt{(R + Ls)(G + Cs)}, \quad (3.15.8)$$

находим $k = \pm \beta(s)$. Этим двум корням характеристического

уравнения (3.15.7) соответствуют два частных интеграла системы уравнений (3.15.6):

$$\begin{aligned}\varphi_1(s, x) &= e^{\beta(s)x}, \quad \psi_1(s, x) = a_1 e^{\beta(s)x}, \\ \varphi_2(s, x) &= e^{-\beta(s)x}, \quad \psi_2(s, x) = a_2 e^{-\beta(s)x}.\end{aligned}$$

Для определения постоянных a_1 и a_2 подставим эти частные интегралы по очереди в любое из уравнений (3.15.6) и потребуем, чтобы оно обратилось в тождество относительно x . Тогда, полагая для краткости

$$\gamma(s) = \sqrt{\frac{G + Cs}{R + Ls}}, \quad (3.15.9)$$

получим $a_1 = -\gamma(s)$, $a_2 = \gamma(s)$. Таким образом, мы нашли два линейно независимых частных интеграла системы уравнений (3.15.6). Общий интеграл этой системы уравнений определяется формулами

$$\left. \begin{aligned}\varphi(s, x) &= c_1 e^{\beta(s)x} + c_2 e^{-\beta(s)x}, \\ \psi(s, x) &= -[c_1 e^{\beta(s)x} - c_2 e^{-\beta(s)x}] \gamma(s).\end{aligned} \right\} \quad (3.15.10)$$

Для определения постоянных интегрирования c_1 и c_2 воспользуемся граничными условиями (3.15.5). Подставляя в них выражения (3.15.4) и сокращая на e^{st} , приведем эти условия к виду

$$\varphi(s, 0) = 1, \quad \varphi(s, l) = Z_n(s) \psi(s, l), \quad (3.15.11)$$

где $Z_n(s) = R_n + sL_n$. Теперь остается подставить сюда выражения $\varphi(s, 0)$, $\varphi(s, l)$ и $\psi(s, l)$, вытекающие из (3.15.10). Тогда получим следующую систему линейных алгебраических уравнений для определения c_1 и c_2 :

$$\left. \begin{aligned}c_1 + c_2 &= 1, \\ c_1 e^{\beta(s)l} + c_2 e^{-\beta(s)l} &= -Z_n(s) [c_1 e^{\beta(s)l} - c_2 e^{-\beta(s)l}] \gamma(s).\end{aligned} \right\} \quad (3.15.12)$$

Решая эту систему уравнений, находим c_1 и c_2 :

$$\left. \begin{aligned}c_1 &= -\frac{[1 - Z_n(s) \gamma(s)] e^{-\beta(s)l}}{[1 + Z_n(s) \gamma(s)] e^{\beta(s)l} - [1 - Z_n(s) \gamma(s)] e^{-\beta(s)l}}, \\ c_2 &= \frac{[1 + Z_n(s) \gamma(s)] e^{\beta(s)l}}{[1 + Z_n(s) \gamma(s)] e^{\beta(s)l} - [1 - Z_n(s) \gamma(s)] e^{-\beta(s)l}}.\end{aligned} \right\} \quad (3.15.13)$$

Передаточная функция однородной длинной линии $\Phi(s)$ в случае, когда выходной переменной является напряжение на конце линии $u(t, l)$, на основании (3.15.4) будет равна

$$\Phi(s) = \frac{u(t, l)}{e^{st}} = \varphi(s, l) = c_1 e^{\beta(s)l} + c_2 e^{-\beta(s)l}. \quad (3.15.14)$$

Подставляя сюда выражения (3.15.13) постоянных c_1 и c_2 , получаем окончательное выражение для передаточной функции $\Phi(s)$:

$$\Phi(s) = \frac{2Z_n(s) \gamma(s)}{[1 + Z_n(s) \gamma(s)] e^{\beta(s)l} - [1 - Z_n(s) \gamma(s)] e^{-\beta(s)l}}. \quad (3.15.15)$$

Выражение (3.15.15) передаточной функции $\Phi(s)$ показывает, что в общем случае однородная длинная линия представляет собой сложный динамический элемент, обладающий, в зависимости от соотношения величин ее основных параметров R , L , C , G и вида нагрузки $Z_n(s)$, в той или иной мере свойствами запаздывающего, колебательного и апериодического звеньев. Проанализируем, как изменяется передаточная функция $\Phi(s)$ в зависимости от изменения вида нагрузки и выбора определенных сочетаний основных параметров линии.

Нагрузка на конце длинной линии возникает только при соединении ее с каким-либо другим элементом автоматической системы. Поэтому влияние нагрузки на динамические свойства линии можно изучить методами главы 4, в которой показано, как по динамическим характеристикам отдельных звеньев определяются динамические характеристики их соединений. Для изучения динамических свойств самой длинной линии, ни с чем не соединенной, следует предположить, что нагрузка на конце линии отсутствует. Это равноценно случаю, когда сопротивление нагрузки R_n бесконечно велико, т. е. $R_n = \infty$. Полагая в (3.15.15) $Z_n(s) = \infty$, находим передаточную функцию длинной линии, рассматриваемой как изолированное звено автоматической системы:

$$\Phi(s) = \frac{2}{e^{\beta(s)l} + e^{-\beta(s)l}} = \frac{1}{\operatorname{ch} \beta(s)l} = \operatorname{sch} \beta(s)l. \quad (3.15.16)$$

Практически влиянием нагрузки на передаточную функцию длинной линии можно пренебречь в случае, когда сопротивление нагрузки очень велико, например, когда коаксиальный или двухпроводный кабель применяется для подачи напряжения на сетку электронной лампы, обладающей большим омическим нагрузочным сопротивлением (обычно порядка нескольких Мом).

Выбирая параметры длинной линии и нагрузку на ее конце соответствующим образом, можно получить линию с желательными динамическими свойствами.

В системах автоматического управления обычно применяются только так называемые *согласованные* линии, т. е. такие линии, которые не дают отражения прямого сигнала, приходящего к концу линии, от нагрузки. Линия становится согласованной, когда нагрузка $Z_n(s)$ выбирается равной волновому сопротивлению линии:

$$Z_n(s) = R_n + sL_n = \sqrt{\frac{R + sL}{G + sC}} = \frac{1}{Y(s)}. \quad (3.15.17)$$

Из этого равенства видно, что согласованности линий в общем случае можно добиться только при каком-нибудь одном значении s , так как это равенство не является тождеством относительно s . Поэтому для обеспечения согласованности линии обычно пользуются переменным током определенной частоты ω_0 и сопротивление нагрузки выбирают из условия (3.15.17) при $s = i\omega_0$. В этом случае при передаче сигналов переменным током частоты ω_0 линия будет работать в режиме бегущей волны. Полагая $\beta(i\omega_0) = \mu + iv$, получим

$$\Phi(i\omega_0) = e^{-\mu l - ivl}. \quad (3.15.18)$$

Длинная линия в этом случае работает как запаздывающее звено с временем запаздывания $\tau = vl/\omega_0$ и коэффициентом усиления $k = e^{-\mu l}$. Сигнал в такой линии распространяется со скоростью ω_0/v и ослабляется в e^μ раз на единицу длины линии.

Полной согласованности линии при любом значении s можно добиться, если выбрать ее параметры так, чтобы удовлетворялось условие $RC = LG$. Действительно, в этом случае $R/L = G/C$, $\gamma(s) = \sqrt{C/L}$ и условие (3.15.17) принимает вид

$$R_n + sL_n = \sqrt{LC}. \quad (3.15.19)$$

Этому условию можно удовлетворить, приняв $R_n = \sqrt{LC}$, $L_n = 0$. При этом выражение (3.15.15) для передаточной функции длинной линии примет вид

$$\Phi(s) = e^{-i\sqrt{RG}} \cdot e^{-s\sqrt{LC}}. \quad (3.15.20)$$

Эта формула показывает, что в этом случае длинная линия представляет собой запаздывающее звено с временем запаздывания $\tau = l\sqrt{LC}$ и коэффициентом ослабления сигнала $k = \exp\{-l\sqrt{RG}\}$. Сигнал распространяется по такой линии без искажения со скоростью, равной $1/\sqrt{LC}$, и ослабляется в $\exp\{\sqrt{RG}\}$ раз на каждую единицу длины линии. Такие линии обычно применяются в виде реальных линий задержки.

Чтобы получить значительное время задержки сигнала, необходимо выбрать большую длину линии (десятки — сотни метров), что может привести к большим трудностям в выборе основных параметров линии, позволяющих выполнить условия $RC = LG$. Кроме того, линия получается громоздкой. Поэтому для получения значительных времен задержки выбирают линию специальной конструкции, состоящей из последовательного соединения отдельных ячеек, имеющих значительные собственные индуктивности L_n и емкости C_n . Тогда $\tau_0 = n\sqrt{L_n C_n}$, где n — число ячеек, входящих в линию задержки. Часто применяют непрерывную спиральную линию задержки, у которой изолированный

В случае полета, близкого к горизонтальному, угол наклона вектора скорости к горизонту θ мал и можно принять $\cos \theta \approx 1$, $\sin \theta \approx \theta$. Тогда, подставляя во второе уравнение (3.16.8) $\dot{\theta} = \theta + \alpha$ и заменяя $\dot{\theta}$, $\ddot{\theta}$ их выражениями из первого уравнения (3.16.8), приведем уравнения (3.16.8) и второе уравнение (3.16.3) к виду

$$\ddot{\theta} = A_{\alpha}\alpha - \frac{g}{v}, \quad \ddot{\alpha} + c_{\alpha}\dot{\alpha} + c_{\alpha}\alpha = c_0 - c_{\delta}\delta, \quad \dot{y} = v\theta, \quad (3.16.10)$$

где

$$\left. \begin{aligned} c_{\alpha} &= a_{\dot{\theta}} + A_{\alpha}, & c_{\alpha} &= a_{\alpha} + a_{\dot{\theta}}A_{\alpha} + \dot{A}_{\alpha}, \\ c_0 &= a_0 + a_{\dot{\theta}}\frac{g}{v} - \frac{g\dot{v}}{v^3}, & c_{\delta} &= a_{\delta}. \end{aligned} \right\} \quad (3.16.11)$$

Таким образом, при полете с малым углом атаки α и малым углом наклона вектора скорости к горизонту θ движение самолета в вертикальной плоскости приближенно описывается линейными уравнениями (3.16.10). При полете с переменной скоростью v , представляющей собой известную функцию времени, все коэффициенты уравнений (3.16.10) при неизвестных α , $\dot{\alpha}$, δ и θ и свободные члены g/v и c_0 являются известными функциями времени. При полете с постоянной скоростью все эти величины постоянны.

В уравнениях (3.16.10) угол отклонения руля δ является входной переменной объекта управления, а выходной переменной может служить, в зависимости от задачи управления, любая из величин θ , α , y или любая их линейная комбинация. Так, например, если задачей системы управления является стабилизация оси самолета в заданном направлении, то за выходную переменную объекта управления следует принять угол тангажа $\vartheta = \theta + \alpha$. При этом за параметр управления следует принять отклонение угла тангажа от заданного значения: $\Delta = \vartheta - \vartheta_0$. Если задачей системы управления является обеспечение полета самолета на заданной высоте H , то за выходную переменную объекта управления следует принять высоту полета y , а за параметр управления — отклонение высоты полета от заданной:

$$\Delta = y - H.$$

В некоторых случаях ракету приходится стабилизировать относительно продольной оси. Для проектирования системы стабилизации необходимо иметь уравнение движения ракеты вокруг продольной оси. Обозначая угловую скорость ракеты относительно продольной оси через ω_x и принимая во внимание, что при вращении ракеты вокруг продольной оси на нее действует аэродинамический момент, приближенно пропорциональный угловой скорости ω_x , и, кроме того, действует аэродинамический момент,

пропорциональный углу отклонения элеронов δ , можем написать уравнение движения ракеты вокруг продольной оси в виде

$$J_x \dot{\omega}_x = \left(m_x^{\delta} \delta + m_x^{\omega_x} \frac{l \omega_x}{v} \right) S l \frac{\rho v^2}{2}, \quad (3.16.12)$$

где l — размах ракеты. Уравнение (3.16.12) можно представить в виде

$$T \dot{\omega}_x + \omega_x = k \delta, \quad (3.16.13)$$

где

$$T = -\frac{2J_x}{m_x^{\omega_x} S l^2 \rho v}, \quad k = -\frac{m_x^{\delta} v}{m_x^{\omega_x} l}. \quad (3.16.14)$$

Входной переменной объекта управления в уравнениях (3.16.12) и (3.16.13) является угол отклонения элеронов δ , а выходной переменной — угловая скорость ракеты относительно продольной оси ω_x . Уравнение (3.16.13) показывает, что ракета в движении вокруг продольной оси может рассматриваться как апериодическое звено. Постоянная времени T и коэффициент усиления этого звена k в общем случае при полете с переменными скоростью и высотой являются функциями времени. В частном случае при горизонтальном равномерном полете постоянная времени T и коэффициент усиления k постоянны и, следовательно, ракета является стационарным апериодическим звеном.

СТРУКТУРНЫЕ СХЕМЫ И МЕТОДЫ ОПРЕДЕЛЕНИЯ
ХАРАКТЕРИСТИК ЛИНЕЙНЫХ СИСТЕМ

§ 4.1. Соединения систем и их элементов.

Структурные схемы

Каждая сложная система состоит из ряда более простых систем, взаимодействующих между собой определенным образом. В зависимости от характера взаимодействия этих систем они могут быть связаны между собой различными способами. Основными типами соединений систем в сложных системах являются последовательное соединение, параллельное соединение и обратная связь.

Последовательным соединением систем называется такое соединение, когда выход каждой системы связывается с входом следующей системы, т. е. когда выходная переменная каждой системы служит входной переменной для следующей системы (рис. 4.1.1).

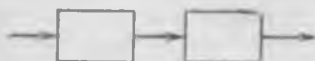


Рис. 4.1.1.



Рис. 4.1.2.

При этом предполагается, что соединяемые системы обладают *направленным действием*. Это значит, что при соединении выхода одной системы с входом другой системы характеристики первой системы не изменяются, т. е. что при данном входном сигнале $x(t)$ системы ее выходная переменная представляет собой одну и ту же функцию $y(t)$ независимо от того, соединен или не соединен выход этой системы с входом другой системы. Это условие далеко не всегда соблюдается на практике. Поэтому физическое последовательное соединение двух систем, как мы увидим дальше, является часто более сложным соединением, чем последовательное соединение в смысле данного определения.

Параллельным соединением систем называется такое соединение, при котором входная переменная подается одновременно на

несколько систем, а их выходные переменные суммируются (рис. 4.1.2*).

Обратной связью называется соединение выхода системы с ее входом (рис. 4.1.3). Если выходная переменная системы непосредственно подается на ее вход без какого-либо преобразования, то обратная связь называется *жесткой*. Жесткая обратная связь может быть *положительной* или *отрицательной* в зависимости от того, суммируется выходная переменная системы с ее входной переменной или вычитается из нее. Если в цепь обратной связи



Рис. 4.1.3.



Рис. 4.1.4.

включена некоторая система, преобразующая выходную переменную основной системы, то такая обратная связь называется *гибкой* (рис. 4.1.4).

Для удобства мы всегда будем рассматривать отрицательную гибкую обратную связь, что, очевидно, не нарушает общности, так как перемену знака всегда можно включить в преобразование, выполняемое системой, включенной в цепь обратной связи. Очевидно, что жесткую обратную связь можно рассматривать как частный случай гибкой обратной связи, когда система, включенная в цепь обратной связи, представляет собой идеальную следящую систему (отрицательная жесткая обратная связь) или безынерционное усиительное звено с коэффициентом усиления, равным — 1 (положительная жесткая обратная связь).

Очевидно, что при любых видах соединений линейных систем система, полученная в результате соединения, будет линейной.

В предыдущей главе мы видели, что практически во всех случаях, когда выходная переменная элемента представляет собой электрический ток или напряжение, подключение к выходу этого элемента другого элемента изменяет ток или напряжение на выходе первого элемента. Это значит, что соединение элементов не является последовательным соединением в смысле данного в начале параграфа определения. Происходит это потому, что при соединении выходных проводов одной электрической цепи с входными проводами другой получается параллельное соединение электри-

*) Сумматоры изображаются на схемах кружком, разделенным на квадранты двумя наклонными диаметрами. Если сумматор производит вычитание, то квадрант, к которому подводится сигнал с обратным знаком, делается черным, как на рис. 4.1.3 и 4.1.4.

Глава 4

СТРУКТУРНЫЕ СХЕМЫ И МЕТОДЫ ОПРЕДЕЛЕНИЯ ХАРАКТЕРИСТИК ЛИНЕЙНЫХ СИСТЕМ

§ 4.1. Соединения систем и их элементов.

Структурные схемы

Каждая сложная система состоит из ряда более простых систем, взаимодействующих между собой определенным образом. В зависимости от характера взаимодействия этих систем они могут быть связаны между собой различными способами. Основными типами соединений систем в сложных системах являются последовательное соединение, параллельное соединение и обратная связь.

Последовательным соединением систем называется такое соединение, когда выход каждой системы связывается с входом следующей системы, т. е. когда выходная переменная каждой системы служит входной переменной для следующей системы (рис. 4.1.1).



Рис. 4.1.1.



Рис. 4.1.2.

При этом предполагается, что соединяемые системы обладают *направленным действием*. Это значит, что при соединении выхода одной системы с входом другой системы характеристики первой системы не изменяются, т. е. что при данном входном сигнале $x(t)$ системы ее выходная переменная представляет собой одну и ту же функцию $y(t)$ независимо от того, соединен или не соединен выход этой системы с входом другой системы. Это условие далеко не всегда соблюдается на практике. Поэтому физическое последовательное соединение двух систем, как мы увидим дальше, является часто более сложным соединением, чем последовательное соединение в смысле данного определения.

Параллельным соединением систем называется такое соединение, при котором входная переменная подается одновременно на

несколько систем, а их выходные переменные суммируются (рис. 4.1.2*).

Обратной связью называется соединение выхода системы с ее входом (рис. 4.1.3). Если выходная переменная системы непосредственно подается на ее вход без какого-либо преобразования, то обратная связь называется *жесткой*. Жесткая обратная связь может быть *положительной* или *отрицательной* в зависимости от того, суммируется выходная переменная системы с ее входной переменной или вычитается из нее. Если в цепь обратной связи

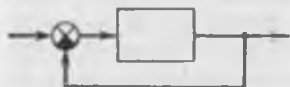


Рис. 4.1.3.



Рис. 4.1.4.

включена некоторая система, преобразующая выходную переменную основной системы, то такая обратная связь называется *гибкой* (рис. 4.1.4).

Для удобства мы всегда будем рассматривать отрицательную гибкую обратную связь, что, очевидно, не нарушает общности, так как перемену знака всегда можно включить в преобразование, выполняемое системой, включенной в цепь обратной связи. Очевидно, что жесткую обратную связь можно рассматривать как частный случай гибкой обратной связи, когда система, включенная в цепь обратной связи, представляет собой идеальную следящую систему (отрицательная жесткая обратная связь) или безынерционное усилительное звено с коэффициентом усиления, равным -1 (положительная жесткая обратная связь).

Очевидно, что при любых видах соединений линейных систем система, полученная в результате соединения, будет линейной.

В предыдущей главе мы видели, что практически во всех случаях, когда выходная переменная элемента представляет собой электрический ток или напряжение, подключение к выходу этого элемента другого элемента изменяет ток или напряжение на выходе первого элемента. Это значит, что соединение элементов не является последовательным соединением в смысле данного в начале параграфа определения. Происходит это потому, что при соединении выходных проводов одной электрической цепи с входными проводами другой получается параллельное соединение электри-

*) Сумматоры изображаются на схемах кружком, разделенным на квадранты двумя наклонными диаметрами. Если сумматор производит вычитание, то квадрант, к которому подводится сигнал с обратным знаком, делается черным, как на рис. 4.1.3 и 4.1.4.

ческих цепей, вследствие чего изменяется общая проводимость полученной сложной цепи, а следовательно, и токи, и падения напряжения в ее элементах. Мы видели также, что для уменьшения влияния присоединяемого элемента на выходной сигнал данного элемента необходимо, чтобы входное сопротивление присоединяемого элемента было велико по сравнению с выходным сопротивлением данного элемента, так как только в этом случае проводимость цепи данного элемента будет незначительно изменяться при присоединении к нему другого элемента. Точно так же и присоединение масс к выходному валу исполнительного устройства, например электродвигателя, изменяет его динамические свойства. И только при малом моменте нагрузки по сравнению с вращающим моментом двигателя присоединение к валу двигателя нагрузки не оказывает существенного влияния на его динамические свойства. При большом моменте нагрузки, как, например, в рассмотренном в § 3.16 случае, когда исполнительное устройство приводит в движение непосредственно объект управления, необходимо определять динамические характеристики исполнительного устройства совместно с присоединенными к нему массами.

Пример 4.1.1. Рассмотрим последовательное соединение двух цепочек RC (рис. 4.1.5), считая входной переменной каждой цепочки напряжение, питающее эту цепочку, а выходной переменной — падение напряжения на конденсаторе.

Из чисто физических соображений ясно, что при данном входном напряжении x падение напряжения y на первом конденсаторе, т. е. на выходе первой цепочки, будет раз личным в случае одной первой цепочки и в слу-

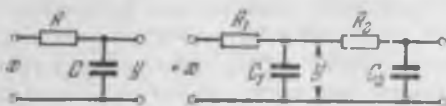


Рис. 4.1.5.

чае последовательного соединения двух цепочек. Действительно, подключение входа второй цепочки к выходу первой равноценно включению в цепь параллельно с первым конденсатором дополнительной линии, представляющей собой последовательно соединенные второе омическое сопротивление и

второй конденсатор. При таком включении изменяется эквивалентное сопротивление всей цепи, а следовательно, и ток в первом сопротивлении, и падение напряжения на этом сопротивлении. Вследствие этого изменится и падение напряжения на первом конденсаторе, т. е. на входе второй цепочки. Таким образом, последовательное присоединение второй цепочки к первой изменяет выходную переменную первой цепочки при данном входном напряжении. Поэтому последовательное соединение двух цепочек RC не является последовательным соединением систем в смысле данного выше определения.

Однако в случае, когда сопротивление R_2 достаточно велико, ток, ответвляемый во вторую цепочку при ее подключении к первой, будет пренебрежимо малым и изменением падения напряжения на первом конденсаторе можно будет пренебречь. Следовательно, при достаточно большом входном сопротивлении второй цепочки по сравнению с выходным сопротивлением первой цепочки их последовательное соединение можно приближенно считать последовательным соединением систем в смысле данного в начале параграфа определения.

Изложенное показывает, что для «развязки» соединяемых элементов, чтобы динамические характеристики каждого из них оставались при соединении неизменными, необходимо согласовать входные и выходные сопротивления соединяемых элементов (усилия в случае механических входов и выходов). Если такое согласование оказывается практически невозможным и в нем нет практической необходимости, то следует рассматривать соединяемые элементы совместно и определять непосредственно динамические характеристики соединения, как это мы делали, изучая динамические характеристики элементов в предыдущей главе и составляя уравнения работы элемента в целом. Однако во многих случаях развязка соединяемых элементов необходима. Если это невозможно сделать согласованием сопротивлений, то применяют специальные развязывающие элементы. В качестве таких элементов обычно применяются усилители с большим входным сопротивлением, катодные повторители, согласующие трансформаторы и другие элементы, имеющие большое входное сопротивление.

Схема, показывающая, какие элементы входят в состав данной системы и как они соединены между собой, называется *структурной схемой* данной системы. Для каждой данной системы, кроме естественной структурной схемы, показывающей ее фактическую структуру, можно составить много других структурных схем. Это дает возможность преобразовывать структурные схемы исследуемых систем и приводить их к виду, удобному для исследования. В § 4.7 мы выведем основные правила преобразования структурных схем.

Пример 4.1.2. Колебательное звено с передаточной функцией

$$\Phi_k(s) = \frac{k}{T^2 s^2 + 2\zeta Ts + 1}$$

Можно представить различными структурными схемами. Так, например, взяв дифференциальное уравнение колебательного звена, можно представить

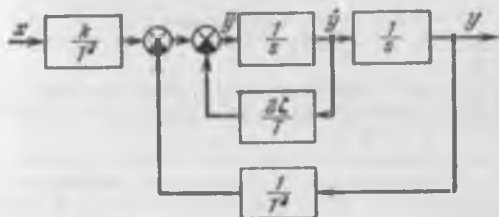


Рис. 4.1.6.

это звено как систему, составленную из двух интеграторов и трех усилителей с соответствующими обратными связями (рис. 4.1.6). С другой стороны, его можно представить как последовательное соединение усилителя с коэффициентом усиления k и идеальной следящей системы с двумя обратными

ческих цепей, вследствие чего изменяется общая проводимость полученной сложной цепи, а следовательно, и токи, и падения напряжения в ее элементах. Мы видели также, что для уменьшения влияния присоединяемого элемента на выходной сигнал данного элемента необходимо, чтобы входное сопротивление присоединяемого элемента было велико по сравнению с выходным сопротивлением данного элемента, так как только в этом случае проводимость цепи данного элемента будет незначительно изменяться при присоединении к нему другого элемента. Точно так же и присоединение масс к выходному валу исполнительного устройства, например электродвигателя, изменяет его динамические свойства. И только при малом моменте нагрузки по сравнению с вращающим моментом двигателя присоединение к валу двигателя нагрузки не оказывает существенного влияния на его динамические свойства. При большом моменте нагрузки, как, например, в рассмотренном в § 3.16 случае, когда исполнительное устройство приводит в движение непосредственно объект управления, необходимо определять динамические характеристики исполнительного устройства совместно с присоединенными к нему массами.

Пример 4.1.1. Рассмотрим последовательное соединение двух цепочек RC (рис. 4.1.5), считая входной переменной каждой цепочки напряжение, питающее эту цепочку, а выходной переменной — падение напряжения на конденсаторе.

Из чисто физических соображений ясно, что при данном входном напряжении x падение напряжения y на первом конденсаторе, т. е. на выходе первой цепочки, будет раз личным в случае одной первой цепочки и в случае

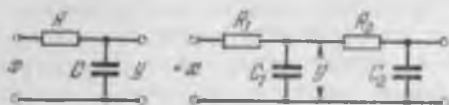


Рис. 4.1.5.

последовательного соединения двух цепочек. Действительно, подключение входа второй цепочки к выходу первой равноценно включению в цепь параллельно с первым конденсатором дополнительной линии, представляющей собой последовательно соединенные второе омическое сопротивление и

второй конденсатор. При таком включении изменяется эквивалентное сопротивление всей цепи, а следовательно, и ток в первом сопротивлении, и падение напряжения на этом сопротивлении. Вследствие этого изменится и падение напряжения на первом конденсаторе, т. е. на входе второй цепочки. Таким образом, последовательное присоединение второй цепочки к первой изменяет выходную переменную первой цепочки при данном входном напряжении. Поэтому последовательное соединение двух цепочек RC не является последовательным соединением систем в смысле данного выше определения.

Однако в случае, когда сопротивление R_2 достаточно велико, ток, ответвляемый во вторую цепочку при ее подключении к первой, будет пренебрежимо малым и изменением падения напряжения на первом конденсаторе можно будет пренебречь. Следовательно, при достаточно большом входном сопротивлении второй цепочки по сравнению с выходным сопротивлением первой цепочки их последовательное соединение можно приближенно считать последовательным соединением систем в смысле данного в начале параграфа определения.

Изложенное показывает, что для «развязки» соединяемых элементов, чтобы динамические характеристики каждого из них оставались при соединении неизменными, необходимо согласовать входные и выходные сопротивления соединяемых элементов (усилия в случае механических входов и выходов). Если такое согласование оказывается практически невозможным и в нем нет практической необходимости, то следует рассматривать соединяемые элементы совместно и определять непосредственно динамические характеристики соединения, как это мы делали, изучая динамические характеристики элементов в предыдущей главе и составляя уравнения работы элемента в целом. Однако во многих случаях развязка соединяемых элементов необходима. Если это невозможно сделать согласованием сопротивлений, то применяют специальные развязывающие элементы. В качестве таких элементов обычно применяются усилители с большим входным сопротивлением, катодные повторители, согласующие трансформаторы и другие элементы, имеющие большое входное сопротивление.

Схема, показывающая, какие элементы входят в состав данной системы и как они соединены между собой, называется *структурной схемой* данной системы. Для каждой данной системы, кроме естественной структурной схемы, показывающей ее фактическую структуру, можно составить много других структурных схем. Это дает возможность преобразовывать структурные схемы исследуемых систем и приводить их к виду, удобному для исследования. В § 4.7 мы выведем основные правила преобразования структурных схем.

Пример 4.1.2. Колебательное звено с передаточной функцией

$$\Phi_k(s) = \frac{k}{T^2 s^2 + 2\zeta Ts + 1}$$

можно представить различными структурными схемами. Так, например, взяв дифференциальное уравнение колебательного звена, можно представить

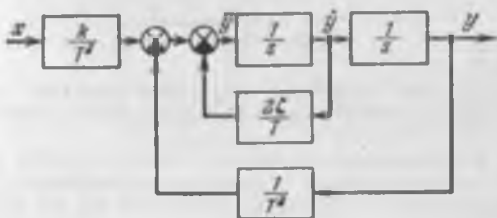


Рис. 4.1.6.

это звено как систему, составленную из двух интеграторов и трех усилителей с соответствующими обратными связями (рис. 4.1.6). С другой стороны, его можно представить как последовательное соединение усилителя с коэффициентом усиления k и идеальной следящей системы с двумя обратными

связями, содержащими идеальные дифференциаторы и усилители (рис. 4.1.7). Наконец, изучив правила нахождения передаточных функций сложных стационарных линейных систем по данным передаточным функциям входящих в их состав более простых систем, изложенные в § 4.6, читатель может убедиться в том, что колебательное звено можно представить в виде последовательного соединения интегратора, аperiodического звена с постоянной

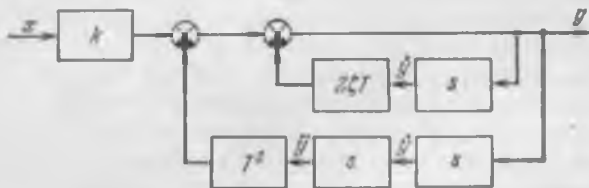


Рис. 4.1.7.

времени $T_1 = T/2\zeta$ и коэффициентом усиления $k_1 = 1/2\zeta T$ и усилителя с коэффициентом усиления k с жесткой отрицательной обратной связью, охватывающей первые два звена (рис. 4.1.8).

Пример 4.1.3. На основании уравнений (3.16.10) самолет или крылатую ракету при полете с постоянной скоростью на заданной высоте можно рассматривать как стационарную линейную систему, входным сигналом

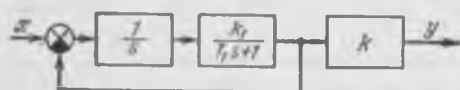


Рис. 4.1.8.

которой является угол отклонения руля высоты δ (управляющее воздействие), а выходной переменной — высота полета y . Эту систему можно представить в виде последовательного соединения колебательного звена, соответствующего второму уравнению (3.16.10), безынерционного усилителя с ко-

эффициентом усиления A_α , идеального интегрирующего звена, усилителя с коэффициентом усиления v и второго интегрирующего звена (рис. 4.1.9).

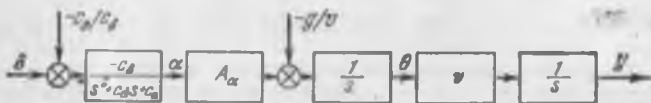


Рис. 4.1.9.

Величина $-g/v$ в первом уравнении (3.16.10) и величина $-c_\delta/c_\alpha$ во втором уравнении (3.16.10) представляют собой постоянно действующие известные возмущения.

При полете с переменной скоростью колебательное звено на схеме рис. 4.1.9 будет нестационарным вследствие переменности коэффициентов c_α^* , c_α и c_δ , а коэффициенты усиления усилителей A_α и v будут известными функциями времени.

Пример 4.1.4. Рассмотрим теперь более сложный пример построения структурной схемы. На рис. 4.1.10 представлена электрокинематическая схема одного канала управления подвижной пушечной установкой на самолете. На основании этой схемы составим сначала вспомогательную схему, показывающую, из каких элементов состоит система и как передаются сигналы между ее элементами (рис. 4.1.11). После этого определяем динамически

а из уравнения (3.13.7) при $M_c = 0$ и уравнения (3.13.6) — передаточную функцию двигателя постоянного тока:

$$\Phi_d(s) = \frac{k_d}{s(T_d s + 1)}.$$

Наконец, составляем уравнения, связывающие токи и напряжения в элементах цепи гибкой обратной связи (рис. 4.1.10):

$$u_{\text{гос}} = R_3 i_3, \quad \left(R_3 + R_4 + \frac{1}{Cs} \right) i_3 = R_1 i_1, \quad i_1 + i_3 = i_2, \quad i_1 R_1 + i_2 R_2 = u_{\text{вых}},$$

где $u_{\text{вых}}$ — выходное напряжение ЭМУ*). Полагая $u_{\text{вых}} = e^{st}$, $u_{\text{гос}} = \Phi_{\text{гос}}(s) e^{st}$ и исключая токи i_1 , i_2 и i_3 , находим передаточную функцию цепи гибкой обратной связи:

$$\Phi_{\text{гос}}(s) = \frac{k_1 T_1 s}{T_2 s + 1}, \quad k_1 T_1 = \frac{R_1 R_3 C}{R_1 + R_2}, \quad T_2 = \left(\frac{R_1 R_2}{R_1 + R_2} + R_3 + R_4 \right) C.$$

После этого строим структурную схему системы управления с указанием

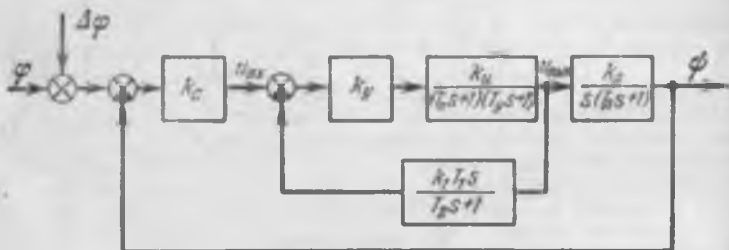


Рис. 4.1.12.

передаточных функций ее звеньев (рис. 4.1.12, где через k_c и k_p обозначены коэффициенты усиления сельсина и электронного усилителя).

Для построения теории линейных систем полезно ввести еще понятие взаимно обратных систем. Две системы называются *взаимно обратными*, если в результате их последовательного соединения получается идеальная следящая система, т. е. если при любом входном сигнале первой системы выходной сигнал второй системы в каждый момент времени равен входному сигналу первой системы. Легко понять, что две взаимно обратные системы остаются взаимно обратными при изменении порядка их соединения (рис. 4.1.13). Действительно, если входному сигналу x первой системы соответствует ее выходной сигнал y , то

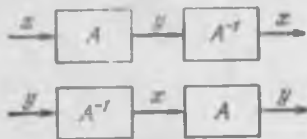


Рис. 4.1.13.

*) Теоретически сопротивления R_2 и R_4 должны быть одинаковыми, чтобы напряжения обратной связи на лампах L_1 и L_2 были равны. Однако практически для компенсации неодинаковости характеристик ламп сопротивления R_2 и R_4 делают регулируемыми. Регулировкой добиваются того, чтобы выходные напряжения обеих ламп усилителя были одинаковыми.

по определению взаимно обратных систем входному сигналу y второй системы соответствует сигнал x на ее выходе. Следовательно, при изменении порядка соединения систем входному сигналу y полученного соединения будет соответствовать его выходной сигнал, также равный y , что и доказывает высказанное утверждение.

§ 4.2. Весовые функции соединений

Для исследования линейных систем важно уметь определять весовые функции сложных систем по данным весовым функциям входящих в их состав систем. Для этого достаточно научиться находить весовые функции различных соединений систем по данным весовым функциям соединяемых систем.

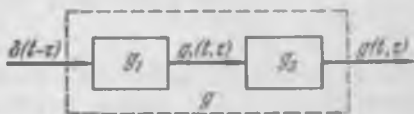


Рис. 4.2.1.

Рассмотрим последовательное соединение двух линейных систем, весовые функции которых $g_1(t, \tau)$ и $g_2(t, \tau)$ заданы (рис. 4.2.1). Найдем весовую функцию $g(t, \tau)$ сложной системы, полученной в результате соединения. Для этого найдем ее реакцию на единичный импульс. В результате действия единичного импульса на входе первой системы ее выходная переменная будет представлять собой ее весовую функцию $g_1(t, \tau)$. Следовательно, входной сигнал второй системы выражается формулой

$$x_2(t) = g_1(t, \tau). \quad (4.2.1)$$

При этом выходная переменная второй системы согласно общей формуле (2.2.3) выражается формулой

$$y_2(t) = \int_{-\infty}^{\infty} g_2(t, \sigma) x_2(\sigma) d\sigma = \int_{-\infty}^{\infty} g_2(t, \sigma) g_1(\sigma, \tau) d\sigma. \quad (4.2.2)$$

Здесь переменную интегрирования мы обозначили буквой σ , так как буквой τ в данном случае обозначен момент действия на систему единичного импульса $\delta(t - \tau)$. Но выходная переменная второй системы в данном случае представляет собой и выходную переменную всего последовательного соединения рассматриваемых систем, т. е. весовую функцию полученной в результате соединения системы:

$$y_2(t) = g(t, \tau). \quad (4.2.3)$$

Сравнивая эту формулу с (4.2.2), получаем следующую формулу

а из уравнения (3.13.7) при $M_c = 0$ и уравнения (3.13.6) — передаточную функцию двигателя постоянного тока:

$$\Phi_d(s) = \frac{k_d}{s(T_d s + 1)}$$

Наконец, составляем уравнения, связывающие токи и напряжения в элементах цепи гибкой обратной связи (рис. 4.1.10):

$$u_{\text{гос}} = R_3 i_3, \quad \left(R_3 + R_4 + \frac{1}{C_k}\right) i_3 = R_1 i_1, \quad i_1 + i_3 = i_2, \quad i_1 R_1 + i_2 R_2 = u_{\text{вых}},$$

где $u_{\text{вых}}$ — выходное напряжение ЭМУ *). Полагая $u_{\text{вых}} = e^{st}$, $u_{\text{гос}} = \Phi_{\text{гос}}(s) e^{st}$ и исключая токи i_1 , i_2 и i_3 , находим передаточную функцию цепи гибкой обратной связи:

$$\Phi_{\text{гос}}(s) = \frac{k_1 T_1 s}{T_2 s + 1}, \quad k_1 T_1 = \frac{R_1 R_3 C}{R_1 + R_2}, \quad T_2 = \left(\frac{R_1 R_2}{R_1 + R_2} + R_3 + R_4\right) C.$$

После этого строим структурную схему системы управления с указанным

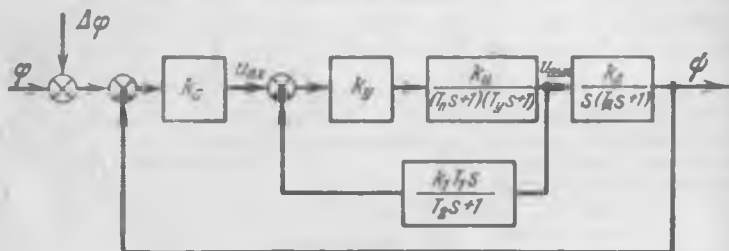


Рис. 4.1.12.

передаточных функций ее звеньев (рис. 4.1.12, где через k_c и k_v обозначены коэффициенты усиления сельсина и электронного усилителя).

Для построения теории линейных систем полезно ввести еще понятие взаимно обратных систем. Две системы называются *взаимно обратными*, если в результате их последовательного соединения получается идеальная следящая система, т. е. если при любом входном сигнале первой системы выходной сигнал второй системы в каждый момент времени равен входному сигналу первой системы.

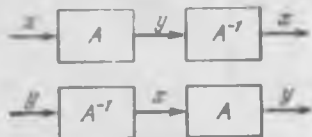


Рис. 4.1.13.

Легко понять, что две взаимно обратные системы остаются взаимно обратными при изменении порядка их соединения (рис. 4.1.13). Действительно, если входному сигналу x первой системы соответствует ее выходной сигнал y , то

*) Теоретически сопротивления R_2 и R_4 должны быть одинаковыми, чтобы напряжения обратной связи на лампах L_1 и L_2 были равны. Однако практически для компенсации неодинаковости характеристик ламп сопротивления R_2 и R_4 делают регулируемыми. Регулировкой добиваются того, чтобы выходные напряжения обеих ламп усилителя были одинаковыми.

по определению взаимно обратных систем входному сигналу y второй системы соответствует сигнал x на ее выходе. Следовательно, при изменении порядка соединения систем входному сигналу y полученного соединения будет соответствовать его выходной сигнал, также равный y , что и доказывает высказанное утверждение.

§ 4.2. Весовые функции соединений

Для исследования линейных систем важно уметь определять весовые функции сложных систем по данным весовым функциям входящих в их состав систем. Для этого достаточно научиться находить весовые функции различных соединений систем по данным весовым функциям соединяемых систем.

Рассмотрим последовательное соединение двух линейных систем, весовые функции которых $g_1(t, \tau)$ и $g_2(t, \tau)$ заданы (рис. 4.2.1). Найдем весовую функцию $g(t, \tau)$ сложной системы, полученной в результате соединения. Для этого найдем ее реакцию на единичный импульс. В результате действия единичного импульса на входе первой системы ее выходная переменная будет представлять собой ее весовую функцию $g_1(t, \tau)$. Следовательно, входной сигнал второй системы выражается формулой

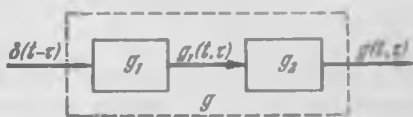


Рис. 4.2.1.

$$x_2(t) = g_1(t, \tau). \quad (4.2.1)$$

При этом выходная переменная второй системы согласно общей формуле (2.2.3) выражается формулой

$$y_2(t) = \int_{-\infty}^{\infty} g_2(t, \sigma) x_2(\sigma) d\sigma = \int_{-\infty}^{\infty} g_2(t, \sigma) g_1(\sigma, \tau) d\sigma. \quad (4.2.2)$$

Здесь переменную интегрирования мы обозначили буквой σ , так как буквой τ в данном случае обозначен момент действия на систему единичного импульса $\delta(t - \tau)$. Но выходная переменная второй системы в данном случае представляет собой и выходную переменную всего последовательного соединения рассматриваемых систем, т. е. весовую функцию полученной в результате соединения системы:

$$y_2(t) = g(t, \tau). \quad (4.2.3)$$

Сравнивая эту формулу с (4.2.2), получаем следующую формулу

для весовой функции последовательного соединения двух линейных систем:

$$g(t, \tau) = \int_{-\infty}^{\infty} g_2(t, \sigma) g_1(\sigma, \tau) d\sigma. \quad (4.2.4)$$

В частном случае физически возможных систем $g_1(\sigma, \tau) = 0$ при $\sigma < \tau$, $g_2(t, \sigma) = 0$ при $\sigma > t$. Следовательно, для физически возможных систем подынтегральная функция в (4.2.4) отлична от нуля только в пределах $\tau \leq \sigma \leq t$, и формула (4.2.4) принимает вид

$$g(t, \tau) = \int_{\tau}^t g_2(t, \sigma) g_1(\sigma, \tau) d\sigma. \quad (4.2.5)$$

Эта формула справедлива при $t > \tau$. При $t < \tau$ подынтегральная функция в (4.2.4) равна нулю при всех σ . Таким образом, мы приходим к выводу, что последовательное соединение физически возможных систем всегда дает физически возможную систему. Согласно принятому в § 2.2 условию пределы интегрирования в формуле (4.2.5) всегда включаются в интервал интегрирования.

Применяя формулы (4.2.4) и (4.2.5) последовательно, можно найти весовую функцию системы, полученной в результате последовательного соединения любого числа линейных систем.

Заметим, что результат последовательного соединения линейных систем в общем случае зависит от порядка их соединения. Действительно, меняя местами соединяемые системы на рис. 4.2.1, мы должны будем поменять местами и весовые функции g_1 и g_2 в формуле (4.2.4). В результате весовая функция последовательного соединения двух систем будет

$$g'(t, \tau) = \int_{-\infty}^{\infty} g_1(t, \sigma) g_2(\sigma, \tau) d\sigma. \quad (4.2.6)$$

Это выражение в общем случае не совпадает с (4.2.4). Читатель легко может в этом убедиться на примере последовательного соединения усилителя с переменным коэффициентом усиления и интегратора. Только в отдельных частных случаях выражение (4.2.6) тождественно совпадает с (4.2.4). В § 4.6 мы увидим, что таким важным частным случаем является случай стационарных линейных систем.

Рассмотрим теперь параллельное соединение линейных систем, имеющих известные весовые функции g_1 и g_2 (рис. 4.2.2). Для определения весовой функции $g(t, \tau)$ сложной системы, полученной в результате соединения, найдем ее реакцию на единичный импульс $\delta(t - \tau)$. Имея в виду, что при подаче на входы соединяе-

мых систем единичного импульса $\delta(t - \tau)$ их выходные переменные равны соответствующим весовым функциям $g_1(t, \tau)$ и $g_2(t, \tau)$, получаем

$$g(t, \tau) = g_1(t, \tau) + g_2(t, \tau). \quad (4.2.7)$$

Таким образом, при параллельном соединении линейных систем их весовые функции суммируются.

Для изучения систем с обратной связью воспользуемся понятием взаимно обратных систем. Рассмотрим последовательное соединение двух взаимно обратных линейных систем (рис. 4.2.3). Весовую функцию линейной системы, обратной по отношению к системе с весовой функцией $g(t, \tau)$, будем обозначать символом $g^-(t, \tau)$. Весовая функция последовательного соединения взаимно обратных систем равна $\delta(t - \tau)$. С другой стороны, весовую функцию последовательного соединения можно вычислить, применяя

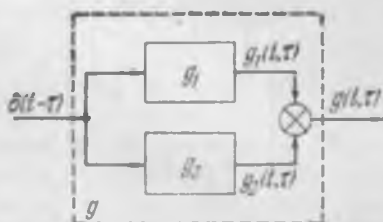


Рис. 4.2.2.

общую формулу (4.2.4). В результате получим

$$\int_{-\infty}^{\infty} g^-(t, \sigma) g(\sigma, \tau) d\sigma = \delta(t - \tau). \quad (4.2.8)$$

Это соотношение представляет собой основное свойство весовых функций взаимно обратных линейных систем. Изменяя порядок соединения систем, получим другое соотношение:



Рис. 4.2.3.

$$\int_{-\infty}^{\infty} g(t, \sigma) g^-(\sigma, \tau) d\sigma = \delta(t - \tau). \quad (4.2.9)$$

Таким образом, весовые функции двух взаимно обратных линейных систем удовлетворяют двум интегральным соотношениям, которые в общем случае друг с другом не совпадают.

Для физически возможных взаимно обратных линейных систем формулы (4.2.8) и (4.2.9) могут быть переписаны в виде

$$\int_{\tau}^t g^-(t, \sigma) g(\sigma, \tau) d\sigma = \delta(t - \tau), \quad (4.2.10)$$

$$\int_{\tau}^t g(t, \sigma) g^-(\sigma, \tau) d\sigma = \delta(t - \tau). \quad (4.2.11)$$

Весовую функцию $g(t, \tau)$ системы с обратной связью, изображенной на рис. 4.2.4, можно определить тем же стандартным приемом, что и в предыдущих случаях, проследив прохождение через систему единичного импульса. Но такой путь сравнительно

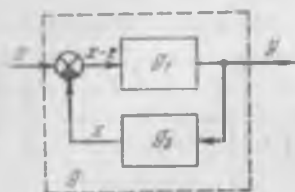


Рис. 4.2.4

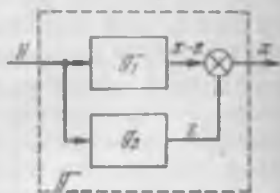


Рис. 4.2.5.

сложен и приводит к интегральному уравнению относительно искомой весовой функции $g(t, \tau)$. Значительно проще находится весовая функция обратной системы. Легко видеть, что системой, обратной по отношению к системе с обратной связью, является параллельное соединение системы, обратной по отношению к системе, стоящей в прямой цепи, и системы, стоящей в цепи обратной связи (рис. 4.2.5). Действительно, обозначив входное возмущение рассматриваемой системы через x , выходную переменную

через y , а выходную переменную системы в цепи обратной связи через z , видим, что входной переменной системы в прямой цепи будет $x-z$, а ее выходной переменной будет y . При подаче на вход параллельного соединения систем с весовыми функциями g_1 и g_2 возмущения y мы получим на выходе первой системы величину $x-z$, а на выходе второй системы, как и в

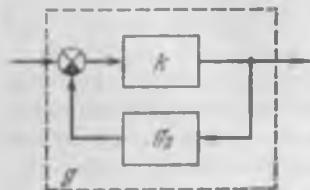


Рис. 4.2.6.

основной системе, величину z . Суммируя, эти величины дадут на выходе параллельного соединения величину x (рис. 4.2.5). Таким образом, последовательное соединение систем, изображенных на рис. 4.2.4 и 4.2.5, дает идеальную следящую систему. Следовательно, эти системы являются взаимно обратными.

Применяя формулу (4.2.7) для весовой функции параллельного соединения линейных систем, находим

$$g^-(t, \tau) = g_1^-(t, \tau) + g_2^-(t, \tau). \quad (4.2.12)$$

Эта формула определяет весовую функцию линейной системы, обратной по отношению к системе с обратной связью, изображенной на рис. 4.2.4.

Рассмотрим, в частности, безынерционный усилитель с обратной связью (рис. 4.2.6). Так как системой, обратной по отношению

к усилителю с коэффициентом усиления k , очевидно, является усилитель с коэффициентом усиления $1/k$, то формула (4.2.12) дает для весовой функции обратной системы

$$g^-(t, \tau) = \frac{1}{k} \delta(t - \tau) + g_2(t, \tau). \quad (4.2.13)$$

При очень большом коэффициенте усиления k первым членом в правой части формулы (4.2.13) можно пренебречь. Тогда получим

$$g^-(t, \tau) \approx g_2(t, \tau). \quad (4.2.14)$$

Отсюда следует, что

$$g(t, \tau) \approx g_2^-(t, \tau). \quad (4.2.15)$$

Таким образом, систему, обратную по отношению к данной линейной системе с весовой функцией $g_2(t, \tau)$, можно приближенно реализовать, включив данную систему в цепь обратной связи усилителя с большим коэффициентом усиления.

§ 4.3. Определение весовых функций методом сопряженных систем

Наиболее естественным способом определения весовой функции данной линейной системы, особенно если эта система реально существует, является способ непосредственного экспериментального определения ее реакции на единичный импульс. Подавая на вход системы единичный импульс $\delta(t - \tau)$ или моделируя действие единичного импульса на входе системы, мы получим на выходе весовую функцию системы в зависимости от первого аргумента t при данном фиксированном значении момента подачи единичного импульса τ . Для полного определения таким способом весовой функции как функции двух переменных необходимо определить реакции системы на единичные импульсы, действующие в различные моменты τ . При этом мы определим поверхность, изображающую весовую функцию системы, ее сечениями, параллельными оси t (рис. 4.3.1).

Обратим теперь внимание на следующее обстоятельство. Для вычисления по формуле (2.2.3) реакции линейной системы на

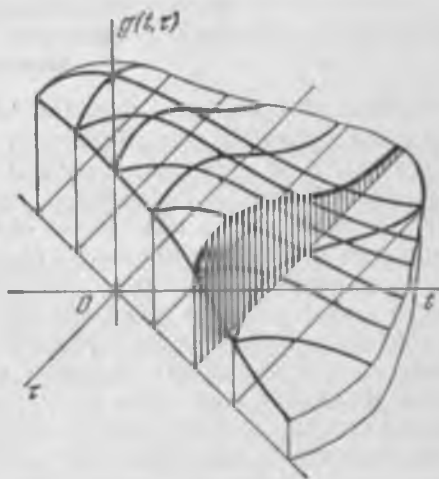


Рис. 4.3.1.

произвольное возмущение в один определенный момент времени t необходимо знать весовую функцию системы как функцию второго аргумента τ при фиксированном значении первого аргумента, т. е. как функцию момента действия на систему единичного импульса τ при фиксированном рассматриваемом моменте времени t . Иными словами, для вычисления реакции линейной системы на произвольное возмущение в момент времени t необходимо знать одно сечение поверхности, изображающей весовую функцию системы, параллельное оси τ (рис. 4.3.1). Это сечение можно, конечно, приближенно определить, если найти достаточно большое число

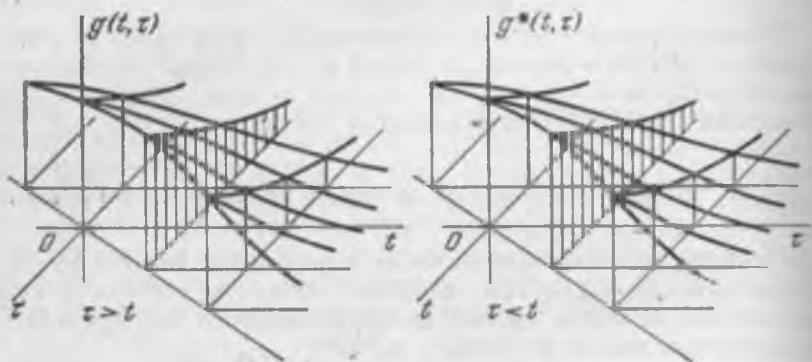


Рис. 4.3.2.

сечений поверхности, параллельных оси t . Однако во многих задачах практики нас интересует лишь конечный момент работы системы, т. е. результат работы системы в один определенный момент времени. В подобных случаях для вычисления по формуле (2.2.3) достаточно знать только одно сечение поверхности, изображающей весовую функцию, параллельное оси τ . Для приближенного определения этого сечения по точкам путем непосредственного моделирования действия на систему единичного импульса необходимо многократно производить моделирование для различных значений момента действия импульса τ , и при этом каждое моделирование дает лишь одну точку интересующего нас сечения поверхности, изображающей весовую функцию. Чтобы избежать многократного моделирования данной системы при определении ее весовой функции, естественно попытаться заменить данную систему другой системой, имеющей ту же весовую функцию, но с измененными ролями аргументов. Иными словами, заменить данную систему такой системой, для которой весовая функция изображается той же поверхностью, что и весовая функция данной системы, но оси координат меняются местами: ось t становится осью τ , и наоборот (рис. 4.3.2). Если такую систему можно найти,

то, давая на ее вход единичный импульс, мы получим на выходе сечение поверхности весовой функции основной системы, параллельное оси t . Таким образом, основная идея заключается в том, чтобы в весовой функции поменять роли аргументов, заменить данную систему такой системой, которая имеет ту же весовую функцию, но с измененными ролями аргументов. Тогда первый аргумент стал бы играть роль момента подачи импульса, а второй — роль текущего момента времени и мы могли бы определить необходимое сечение поверхности, изображающей весовую функцию данной системы, параллельное оси t , одним моделированием действия единичного импульса на некоторую другую линейную систему.

Линейная система, весовая функция которой совпадает с весовой функцией данной линейной системы при изменении ролей аргументов t и τ , называется системой, *сопряженной* с данной линейной системой. Иначе это можно сформулировать так: линейная система, весовая функция которой получается из весовой функции данной линейной системы перестановкой аргументов, называется системой, сопряженной с данной линейной системой.

Весовые функции сопряженных систем мы будем отмечать звездочкой в виде верхнего индекса. Таким образом, если данная линейная система имеет весовую функцию $g(t, \tau)$, то весовая функция сопряженной системы будет $g^*(t, \tau)$. В этом обозначении первый аргумент, как всегда, представляет собой текущее время, а второй — момент действия единичного импульса. Согласно определению сопряженной системы

$$g^*(t, \tau) = g(\tau, t). \quad (4.3.1)$$

Это равенство выражает тот факт, что весовые функции двух взаимно сопряженных систем изображаются одной и той же поверхностью в координатах, различающихся обозначением осей t и τ (рис. 4.3.2).

Очевидно, что система, сопряженная с физически возможной системой, является физически невозможной системой. На первый взгляд это кажется серьезным затруднением. Однако это затруднение легко преодолевается. Действительно, вследствие условия (2.2.4) физической возможности данной системы сопряженная система в каждый данный момент t должна реагировать только на возмущения, которые следуют за данным моментом t , и не может реагировать на предшествующие возмущения. Если изменить направление течения времени на обратное (реверсировать время), то понятия «предшествующий» и «последующий» поменяются местами и сопряженная система станет физически возможной. Следовательно, сопряженную систему можно моделировать, изменив направление отсчета времени, т. е. моделируя отрицательное время натуральным физическим временем. Это обстоятельство делает метод сопряженных систем исключительно

ценным для практики. Моделируя сопряженную систему в отрицательном времени, можно получить как раз то сечение весовой функции данной системы, которое необходимо для расчетов по формуле (2.2.3).

Для практического применения метода сопряженных систем необходимо уметь находить системы, сопряженные с данными. Для этого достаточно, во-первых, научиться определять системы, сопряженные с некоторыми элементарными системами, во-вторых, научиться находить системы, сопряженные с системами, полученными из элементарных систем путем различных соединений.

Найдем сначала системы, сопряженные с элементарными звеньями. Будем по-прежнему обозначать входной сигнал системы буквой x , а выходной сигнал — буквой y . Входную переменную сопряженной системы будем обозначать буквой ξ , а выходную — буквой η .

1. *Усилитель* с коэффициентом усиления $k(t)$ имеет весовую функцию $g_y(t, \tau) = k(t)\delta(t - \tau)$ (формула (2.2.9)). Согласно определению (4.3.1) весовая функция сопряженной системы определится формулой

$$g_y^*(t, \tau) = g_y(\tau, t) = k(\tau)\delta(\tau - t).$$

Поскольку δ -функция равна нулю при $t \neq \tau$, можно заменить $k(\tau)$ на $k(t)$, ничего не изменяя. Тогда, принимая во внимание четность δ -функции, получим

$$g_y^*(t, \tau) = k(t)\delta(t - \tau). \quad (4.3.2)$$

Таким образом, системой, сопряженной с усилителем, является тот же усилитель. Иными словами, усилитель представляет собой *самосопряженную* систему. В частности, идеальная следящая система, представляющая собой усилитель с $k = 1$, является самосопряженной системой.

2. *Дифференциатор* имеет весовую функцию $\delta'(t - \tau)$. Весовая функция сопряженной системы определяется согласно (4.3.1) формулой

$$g_d^*(t, \tau) = g_d(\tau, t) = \delta'(\tau - t). \quad (4.3.3)$$

Рассмотрим соотношение между входной и выходной переменными этой системы. На основании общей формулы (2.2.3) имеем

$$\eta(t) = \int_{-\infty}^{\infty} g_d^*(t, \tau) \xi(\tau) d\tau = \int_{-\infty}^{\infty} \delta'(\tau - t) \xi(\tau) d\tau \quad (4.3.4)$$

или, пользуясь формулой (2.1.9),

$$\eta(t) = -\xi'(t) = -\frac{d\xi(t)}{dt}. \quad (4.3.5)$$

Таким образом, операцией, сопряженной с операцией дифференцирования по времени, является дифференцирование по отрицательному времени. Или иначе: системой, сопряженной с идеальным дифференциатором, является система, производящая дифференцирование по отрицательному времени. При моделировании сопряженной системы в отрицательном времени систему, сопряженную с дифференциатором, следует моделировать тем же дифференциатором.

3. *Интегратор* имеет в качестве весовой функции единичную ступенчатую функцию, $1(t - \tau)$. Следовательно, сопряженная система имеет весовую функцию

$$g_{\Pi}^*(t, \tau) = g_{\Pi}(\tau, t) = 1(\tau - t). \quad (4.3.6)$$

Рассмотрим связь между входной и выходной переменными этой системы. На основании общей формулы (2.2.3) имеем

$$\eta(t) = \int_{-\infty}^{\infty} g_{\Pi}^*(t, \tau) \xi(\tau) d\tau = \int_{-\infty}^{\infty} 1(\tau - t) \xi(\tau) d\tau = \int_t^{\infty} \xi(\tau) d\tau, \quad (4.3.7)$$

так как единичная ступенчатая функция равна нулю при отрицательном аргументе и единице при положительном аргументе. После замены переменной интегрирования $\tau = -\sigma$ (4.3.7) примет вид

$$\eta(t) = \int_{-\infty}^{-t} \xi(-\sigma) d\sigma. \quad (4.3.8)$$

Эта формула показывает, что операцией, сопряженной с интегрированием по времени, является интегрирование по отрицательному времени. Таким образом, при моделировании в отрицательном времени интегратор в сопряженной системе сохраняет свою роль.

4. *Запаздывающее звено* с временем запаздывания l имеет весовую функцию $g_z(t, \tau) = \delta(t - \tau - l)$ (формула (2.2.8)). Весовая функция сопряженной системы на основании (4.3.1) определяется формулой

$$g_z^*(t, \tau) = g_z(\tau, t) = \delta(\tau - t - l) = \delta(t - \tau + l). \quad (4.3.9)$$

Сравнение этой формулы с (2.2.7) показывает, что запаздывающее звено и экстраполятор с тем же параметром l являются взаимно сопряженными системами. При моделировании сопряженной системы в отрицательном времени запаздывающие звенья и экстраполяторы сохраняют свою роль.

Таким образом, при моделировании сопряженных систем в отрицательном времени усилители, дифференциаторы, интеграторы и запаздывающие звенья сохраняют свои роли, т. е. усилители моделируются усилителями, дифференциаторы — диф-

ферециаторами, интеграторы — интеграторами, запаздывающие звенья — запаздывающими звеньями.

Легко доказать, что и вообще любые стационарные системы при моделировании сопряженной системы при обратном течении времени остаются неизменными. Действительно, пусть $g(t, \tau) = w(t - \tau)$ — весовая функция стационарной линейной системы. Весовой функцией соответствующей сопряженной системы будет, согласно определению, $g^*(t, \tau) = w(\tau - t)$. Для моделирования сопряженной системы в отрицательном времени необходимо сде-

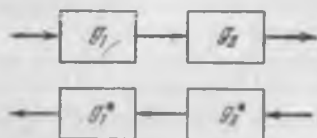


Рис. 4.3.3.

лать замену переменных $s = t_1 - t$ и соответственно $\sigma = t_1 - \tau$. Тогда будем иметь $\tau - t = s - \sigma$. При этом весовая функция сопряженной системы примет вид $w(s - \sigma)$, что и доказывает наше утверждение. В частности, формирующие звенья первого и второго порядка, апериодические и колебатель-

ные звенья при моделировании сопряженных систем в отрицательном времени сохраняют свою роль.

Таким образом, мы доказали, что при моделировании сопряженных систем в отрицательном времени все элементарные звенья, рассмотренные в § 2.6, и усилители с переменными коэффициентами усиления сохраняют свою роль, т. е. моделируются теми же самыми звеньями. При этом, само собой разумеется, все переменные коэффициенты усиления следует программировать в отрицательном времени.

Найдем теперь систему, сопряженную с последовательным соединением двух систем, имеющих весовые функции g_1 и g_2 (рис. 4.3.3). Весовая функция последовательного соединения выражается через весовые функции соединяемых систем формулой (4.2.4). Весовая функция сопряженной системы, по определению, получается перестановкой аргументов t и τ :

$$g^*(t, \tau) = \int_{-\infty}^{\infty} g_2(\tau, \sigma) g_1(\sigma, t) d\sigma, \quad (4.3.10)$$

или

$$g^*(t, \tau) = \int_{-\infty}^{\infty} g_1^*(t, \sigma) g_2^*(\sigma, \tau) d\sigma. \quad (4.3.11)$$

Сравнивая эту формулу с (4.2.4), приходим к заключению, что системой, сопряженной с последовательным соединением линейных систем, является последовательное соединение соответствующих сопряженных систем, взятых в обратном порядке (рис. 4.3.3).

При параллельном соединении линейных систем, как мы видели, их весовые функции суммируются. Поэтому, переставляя

аргументы t и τ в формуле (4.2.7), приходим к выводу, что системой, сопряженной с параллельным соединением линейных систем, является параллельное соединение соответствующих сопряженных систем (рис. 4.3.4).

Рассмотрим последовательное соединение двух взаимно обратных систем (рис. 4.3.5). Согласно выведенному правилу сопряженная система представляет собой последовательное соединение соответствующих сопряженных систем, взятых в обратном порядке. Но при последова-

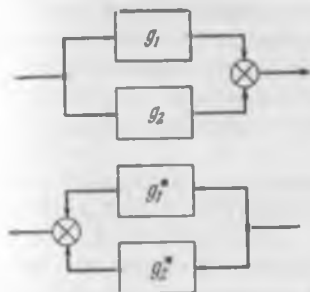


Рис. 4.3.4.

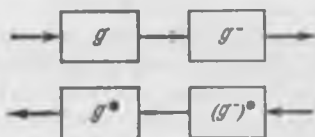


Рис. 4.3.5.

тельном соединении двух взаимно обратных систем получается идеальная следящая система, которая является самосопряженной.

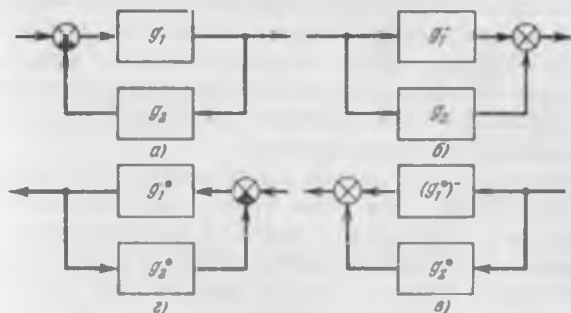


Рис. 4.3.6.

Следовательно, системы, сопряженные с двумя взаимно обратными системами, также являются взаимно обратными. Математически это можно записать в виде формулы:

$$(g^-)^* = (g^*)^- \quad (4.3.12)$$

Рассмотрим теперь систему с весовой функцией g_1 , охваченную отрицательной обратной связью, в цепь которой включена система с весовой функцией g_2 (рис. 4.3.6, а). Для нахождения сопряженной системы применим следующий прием. Построим сначала обратную систему, которая по доказанному в предыдущем парагра-

рафе представляет собой параллельное соединение систем с весовыми функциями g_1 и g_2 (рис. 4.3.6, б). Системой, сопряженной с этой обратной системой, будет на основании полученных результатов параллельное соединение соответствующих сопряженных систем (рис. 4.3.6, в). Система, обратная по отношению к этому параллельному соединению, будет сопряженной с исходной системой вследствие формулы (4.3.12). Но система, изображенная на рис. 4.3.6, в, отличается от системы, изображенной на рис. 4.3.6, б, только тем, что все элементарные системы заменены соответствующими сопряженными. Следовательно, и система, сопряженная с данной системой, получается простой заменой в данной системе всех входящих в нее систем соответствующими сопряженными системами (рис. 4.3.6, г).

Таким образом, система, сопряженная с системой с обратной связью, получается из основной системы путем замены составляющих ее систем соответствующими сопряженными системами. При этом, изображая сопряженные системы с обратным направлением прохождения сигналов, как это сделано на рис. 4.3.3, 4.3.4, 4.3.5 и 4.3.6, мы должны будем заменить точку разветвления сумматором и наоборот.

На основании изложенного можно дать следующую сводку правил построения сопряженных систем:

1. При моделировании сопряженной системы в отрицательном времени все элементарные звенья (усилители, дифференциаторы, интеграторы, форсирующие, апериодические, колебательные и запаздывающие звенья), а также все стационарные системы сохраняют свои функции. Все прочие линейные звенья заменяются соответствующими сопряженными.

2. Направление всех связей на структурной схеме изменяется на противоположное. При этом точки разветвления заменяются сумматорами, а сумматоры — точками разветвления.

Эти правила дают возможность весьма просто и быстро находить сопряженные системы для любых систем, составленных из идеальных систем, осуществляющих основные операции анализа, и стационарных систем. А именно необходимо наложить на структурную схему данной системы лист прозрачной бумаги, перерисовать на него все системы, входящие в состав данной системы, заменить сумматоры точками разветвления и наоборот, а направление всех связей заменить противоположным. В результате получим структурную схему сопряженной системы *).

*) Эти правила нахождения сопряженных систем были впервые выведены для частного случая систем, состоящих из конечного числа идеальных интеграторов, усилителей и сумматоров, Лэнингом и Бэттином. Для произвольных линейных систем эти правила были установлены А. В. Солодовым [66] и Н. М. Сотским, получившими их независимо друг от друга различными методами.

Итак, для определения весовой функции данной линейной системы в зависимости от второго аргумента τ при фиксированном значении первого аргумента t следует построить структурную схему этой системы и, пользуясь выведенными правилами, построить структурную схему соответствующей сопряженной системы. Найденную таким образом сопряженную систему следует моделировать в обратном времени, начиная с интересующего нас момента t , моделируя при этом действие на систему в момент t единичного импульса. При этом все стационарные звенья исследуемой системы останутся в сопряженной системе неизменными. Это дает возможность включать в схему моделирования сопряженной системы реальные стационарные звенья, входящие в состав исследуемой системы.

Пример 4.3.1. Найти систему, сопряженную с системой, поведение которой описывается линейным дифференциальным уравнением

$$a_2 \ddot{y} + a_1 \dot{y} + a_0 y = x. \quad (4.3.13)$$

Составим для этой системы схему моделирования, после чего найдем структурную схему сопряженной системы, пользуясь выведенными правилами.

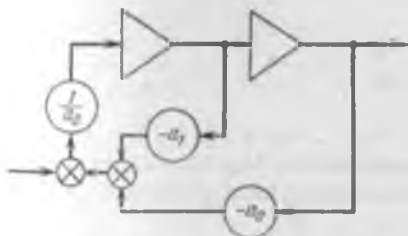


Рис. 4.3.7.

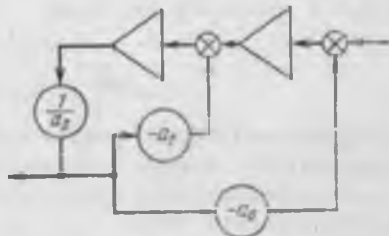


Рис. 4.3.8.

Легко проверить, что уравнению (4.3.13) соответствует структурная схема системы, изображенная на рис. 4.3.7, где треугольниками обозначены интеграторы, а кружочками — безынерционные усилители (блоки коэффициентов) и сумматоры. Согласно сформулированным выше общим правилам чертим те же элементарные звенья, которые входят в состав рассматриваемой системы: два интегратора и три блока коэффициентов, сумматоры заменяем точками разветвления, а точки разветвления — сумматорами; направление всех связей изменяем на обратное. В результате получим схему моделирования сопряженной системы, изображенную на рис. 4.3.8.

Для составления дифференциального уравнения сопряженной системы обозначим через u отрицательное время: $u = -t$, через ξ — выходную переменную первого интегратора и через ω — выходную переменную второго интегратора. Тогда из схемы рис. 4.3.8 будут следовать уравнения

$$\frac{d\xi}{du} = \xi - a_0 \eta, \quad \frac{d\omega}{du} = \xi - a_1 \eta, \quad \eta = \frac{\omega}{a_2}.$$

Дифференцируя второе уравнение по u и заменяя в полученном равенстве $\frac{d\xi}{du}$ и ω их выражениями из остальных двух уравнений, исключим

переменные ξ и ω и в результате получим уравнение

$$\frac{d^2}{du^2} (a_2 \eta) + \frac{d}{du} (a_1 \eta) + a_0 \eta = \xi. \quad (4.3.14)$$

Это и есть уравнение, связывающее входную и выходную переменные сопряженной системы при моделировании ее в отрицательном времени. Переходя к натуральному времени, получим уравнение сопряженной системы в виде

$$\frac{d^2}{dt^2} (a_2 \eta) + \frac{d}{dt} (a_1 \eta) + a_0 \eta = \xi. \quad (4.3.15)$$

§ 4.4. Линейная система, описываемая одним дифференциальным уравнением

Применим изложенную в предыдущих параграфах общую теорию линейных систем к часто встречающемуся на практике случаю, когда поведение физически возможной системы описывается обыкновенным линейным дифференциальным уравнением *)

$$\sum_{k=0}^n a_k y^{(k)} = \sum_{k=0}^m b_k x^{(k)} \quad (4.4.1)$$

или, в развернутом виде,

$$\begin{aligned} a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = \\ = b_m x^{(m)} + b_{m-1} x^{(m-1)} + \dots + b_1 x' + b_0 x. \end{aligned} \quad (4.4.2)$$

Коэффициенты этого уравнения $a_0, a_1, \dots, a_n$ и $b_0, b_1, \dots, b_m$ в общем случае могут быть произвольными функциями времени.

В дальнейшем нам будет удобно записывать уравнение (4.4.1) в сокращенном виде:

$$Fy = Hx, \quad (4.4.3)$$

где через F и H обозначены линейные дифференциальные операторы:

$$F = \sum_{k=0}^n a_k D^k, \quad H = \sum_{k=0}^m b_k D^k. \quad (4.4.4)$$

Эти операторы представляют собой полиномы относительно оператора дифференцирования по времени $D = d/dt$, в общем случае с переменными коэффициентами.

Очевидно, что система, описываемая дифференциальным уравнением (4.4.3), представляет собой последовательное соединение двух линейных систем: системы, осуществляющей линейную дифференциальную операцию H над входным возмущением x и

*) Естественно, что поведение линейной системы может описываться только линейными уравнениями, так как принцип суперпозиции применим только к линейным уравнениям.

дающей на выходе функцию

$$x_1 = Hx, \quad (4.4.5)$$

и системы, описываемой линейным дифференциальным уравнением

$$Fy = x_1. \quad (4.4.6)$$

Определив весовые функции этих двух систем, мы можем найти весовую функцию системы, описываемой уравнением (4.4.3), по формуле (4.2.4).

Для определения весовой функции системы, осуществляющей линейную дифференциальную операцию H над входным возмущением, достаточно заменить в уравнении этой системы (4.4.5)

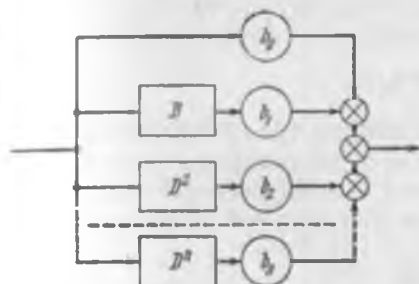


Рис. 4.4.1.

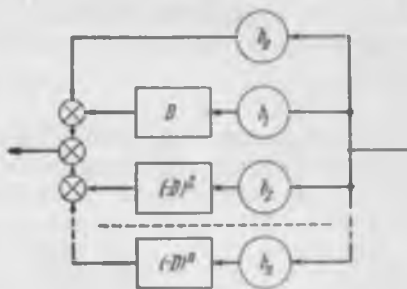


Рис. 4.4.2.

входное возмущение x единичным импульсом $\delta(t - \tau)$. Тогда получим для весовой функции $h(t, \tau)$ этой системы формулу

$$h(t, \tau) = H_t \delta(t - \tau), \quad (4.4.7)$$

где индекс t у оператора H указывает, что все коэффициенты этого оператора рассматриваются как функции t и дифференцирование осуществляется по переменной t при фиксированном τ . Подставляя в формулу (4.4.7) выражение (4.4.4) оператора H , получим развернутую формулу для весовой функции $h(t, \tau)$:

$$h(t, \tau) = \sum_{k=0}^m b_k(t) \delta^{(k)}(t - \tau). \quad (4.4.8)$$

Для определения сопряженной системы воспользуемся выведенными в предыдущем параграфе общими правилами нахождения сопряженных систем. Для этого начертим структурную схему системы, осуществляющей над входным возмущением линейную дифференциальную операцию H (рис. 4.4.1). Сопряженная система получится, если мы заменим дифференцирующие звенья звеньями, выполняющими дифференцирование по отрицательному времени, сумматор — точкой разветвления и наоборот и изменим направление всех связей на обратное. Полученная в результате схема сопряженной системы представлена на рис. 4.4.2. Связь между

переменные ξ и ω и в результате получим уравнение

$$\frac{d^2}{du^2} (a_2 \eta) + \frac{d}{du} (a_1 \eta) + a_0 \eta = \xi. \quad (4.3.14)$$

Это и есть уравнение, связывающее входную и выходную переменные сопряженной системы при моделировании ее в отрицательном времени. Переходя к натуральному времени, получим уравнение сопряженной системы в виде

$$\frac{d^2}{dt^2} (a_2 \eta) - \frac{d}{dt} (a_1 \eta) + a_0 \eta = \xi. \quad (4.3.15)$$

§ 4.4. Линейная система, описываемая одним дифференциальным уравнением

Применим изложенную в предыдущих параграфах общую теорию линейных систем к часто встречающемуся на практике случаю, когда поведение физически возможной системы описывается обыкновенным линейным дифференциальным уравнением *)

$$\sum_{k=0}^n a_k y^{(k)} = \sum_{k=0}^m b_k x^{(k)} \quad (4.4.1)$$

или, в развернутом виде,

$$\begin{aligned} a_n y^{(n)} + a_{n-1} y^{(n-1)} + \dots + a_1 y' + a_0 y = \\ = b_m x^{(m)} + b_{m-1} x^{(m-1)} + \dots + b_1 x' + b_0 x. \end{aligned} \quad (4.4.2)$$

Коэффициенты этого уравнения $a_0, a_1, \dots, a_n$ и $b_0, b_1, \dots, b_m$ в общем случае могут быть произвольными функциями времени.

В дальнейшем нам будет удобно записывать уравнение (4.4.1) в сокращенном виде:

$$Fy = Hx, \quad (4.4.3)$$

где через F и H обозначены линейные дифференциальные операторы:

$$F = \sum_{k=0}^n a_k D^k, \quad H = \sum_{k=0}^m b_k D^k. \quad (4.4.4)$$

Эти операторы представляют собой полиномы относительно оператора дифференцирования по времени $D = d/dt$, в общем случае с переменными коэффициентами.

Очевидно, что система, описываемая дифференциальным уравнением (4.4.3), представляет собой последовательное соединение двух линейных систем: системы, осуществляющей линейную дифференциальную операцию H над входным возмущением x и

*) Естественно, что поведение линейной системы может описываться только линейными уравнениями, так как принцип суперпозиции применим только к линейным уравнениям.

дающей на выходе функцию

$$x_1 = Hx, \quad (4.4.5)$$

и системы, описываемой линейным дифференциальным уравнением

$$Fy = x_1. \quad (4.4.6)$$

Определив весовые функции этих двух систем, мы можем найти весовую функцию системы, описываемой уравнением (4.4.3), по формуле (4.2.4).

Для определения весовой функции системы, осуществляющей линейную дифференциальную операцию H над входным возмущением, достаточно заменить в уравнении этой системы (4.4.5)

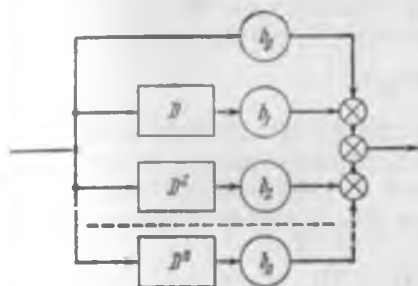


Рис. 4.4.1.

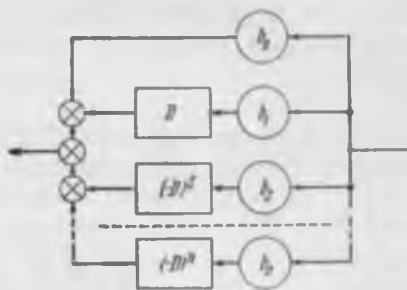


Рис. 4.4.2.

входное возмущение x единичным импульсом $\delta(t - \tau)$. Тогда получим для весовой функции $h(t, \tau)$ этой системы формулу

$$h(t, \tau) = H_t \delta(t - \tau), \quad (4.4.7)$$

где индекс t у оператора H указывает, что все коэффициенты этого оператора рассматриваются как функции t и дифференцирование осуществляется по переменной t при фиксированном τ . Подставляя в формулу (4.4.7) выражение (4.4.4) оператора H , получим развернутую формулу для весовой функции $h(t, \tau)$:

$$h(t, \tau) = \sum_{k=0}^m b_k(t) \delta^{(k)}(t - \tau). \quad (4.4.8)$$

Для определения сопряженной системы воспользуемся выведенными в предыдущем параграфе общими правилами нахождения сопряженных систем. Для этого начертим структурную схему системы, осуществляющей над входным возмущением линейную дифференциальную операцию H (рис. 4.4.1). Сопряженная система получится, если мы заменим дифференцирующие звенья звеньями, выполняющими дифференцирование по отрицательному времени, сумматор — точкой разветвления и наоборот и изменим направление всех связей на обратное. Полученная в результате схема сопряженной системы представлена на рис. 4.4.2. Связь между

выходной переменной η этой системы и ее входной переменной ξ , как следует из рис. 4.4.2, имеет вид

$$\eta = \sum_{k=0}^m (-1)^k \frac{d^k}{dt^k} (b_k \xi). \quad (4.4.9)$$

Линейный дифференциальный оператор

$$H^* = \sum_{k=0}^m (-1)^k \frac{d^k}{dt^k} (b_k \cdot), \quad (4.4.10)$$

где точкой заменена входная функция, над которой этот оператор действует, называется *сопряженным* с оператором H . При помощи сопряженного оператора уравнение (4.4.9) переписется в виде

$$\eta = H^* \xi. \quad (4.4.11)$$

Для определения весовой функции $g(t, \tau)$ системы, поведение которой описывается линейным дифференциальным уравнением (4.4.6), заменим входное возмущение x_1 этой системы единичным импульсом $\delta(t - \tau)$. Тогда получим для весовой функции $g(t, \tau)$ уравнение

$$F_t g(t, \tau) = \delta(t - \tau), \quad (4.4.12)$$

где, как и раньше, индекс t у оператора F означает, что все коэффициенты этого оператора рассматриваются как функции t и дифференцирование производится по переменной t при фиксированном τ . Подставляя в (4.4.12) выражение (4.4.4) оператора F , можно представить дифференциальное уравнение (4.4.12) в развернутом виде:

$$\sum_{k=0}^n a_k(t) g_t^{(k)}(t, \tau) = \delta(t - \tau). \quad (4.4.13)$$

Это уравнение представляет собой обыкновенное линейное дифференциальное уравнение, определяющее весовую функцию рассматриваемой системы как функцию первого аргумента t при фиксированном значении второго аргумента τ . Весовая функция $g(t, \tau)$ определяется как интеграл уравнения (4.4.12), равный нулю при всех $t < \tau$.

Покажем теперь, что δ -функцию в правой части уравнения (4.4.12) или (4.4.13) можно заменить соответствующими начальными условиями. Для этого заметим, что единичный импульс в правой части уравнения (4.4.12) вводит δ -функцию $\delta(t - \tau)$ в старший член уравнения $a_n(t) g_t^{(n)}(t, \tau)$, что равноценно слагаемому

$$\frac{1}{a_n(\tau)} \delta(t - \tau)$$

в выражении n -й производной весовой функции $g_i^{(n)}(t, \tau)$. Этому слагаемому соответствует скачкообразное изменение $n-1$ -й производной $g_i^{(n-1)}(t, \tau)$ при $t = \tau$ на величину $1/a_n(\tau)$. При этом производные $g_i^{(n-2)}(t, \tau), \dots, g_i'(t, \tau)$ и сама функция $g(t, \tau)$ остаются непрерывными в точке $t = \tau$. А так как при $t < \tau$ функция $g(t, \tau)$ тождественно равна нулю в силу условия физической возможности, то функции $g(t, \tau), g_i'(t, \tau), \dots, g_i^{(n-2)}(t, \tau)$ равны нулю при $t = \tau$. Функция $g_i^{(n-1)}(t, \tau)$ также равна нулю при всех $t < \tau$. Следовательно, при $t = \tau$ она скачком изменяется от нуля до $1/a_n(\tau)$ вследствие действия на рассматриваемую систему единичного импульса $\delta(t - \tau)$.

Изложенное показывает, что весовую функцию рассматриваемой системы можно определить как интеграл однородного линейного дифференциального уравнения

$$F_t g(t, \tau) = 0, \quad (4.4.14)$$

удовлетворяющий начальным условиям

$$\left. \begin{aligned} g(\tau, \tau) = g_i'(\tau, \tau) = \dots = g_i^{(n-2)}(\tau, \tau) = 0, \\ g_i^{(n-1)}(\tau, \tau) = \frac{1}{a_n(\tau)}, \end{aligned} \right\} \quad (4.4.15)$$

где для краткости положено

$$g_i^{(k)}(\tau, \tau) = [g_i^{(k)}(t, \tau)]_{t=\tau} = \left[\frac{\partial^k g(t, \tau)}{\partial t^k} \right]_{t=\tau}. \quad (4.4.16)$$

Таким образом, действие единичного импульса на линейную систему, описываемую дифференциальным уравнением (4.4.6), можно моделировать начальными условиями (4.4.15) в точке $t = \tau$, и все трудности, связанные с непосредственным вводом в систему единичного импульса, отпадают.

Для нахождения системы, сопряженной с системой, поведение которой описывается уравнением (4.4.6), достаточно заметить, что эта система является обратной по отношению к системе, осуществляющей над входным возмущением линейную дифференциальную операцию F . Системой, сопряженной с этой обратной системой, по доказанному выше является система, осуществляющая над входной переменной сопряженную линейную дифференциальную операцию F^* :

$$F^* = \sum_{k=0}^n (-1)^k \frac{d^k}{dt^k} (a_k \cdot). \quad (4.4.17)$$

Переходя к обратной системе, приходим к заключению, что системой, сопряженной с описываемой уравнением (4.4.6), является система, поведение которой описывается линейным

дифференциальным уравнением

$$F^* \eta = \xi. \quad (4.4.18)$$

Это уравнение на основании формулы (4.4.17) может быть переписано в развернутом виде:

$$\sum_{h=0}^n (-1)^h \frac{d^h}{dt^h} (a_h \eta) = \xi, \quad (4.4.19)$$

Уравнение (4.4.18) или (4.4.19) называется в теории дифференциальных уравнений *сопряженным* с уравнением (4.4.6).

Для определения весовой функции сопряженной системы достаточно заменить в уравнении (4.4.18) входную переменную ξ единичным импульсом $\delta(t - \tau)$. Тогда получим для весовой функции сопряженной системы уравнение

$$F_1^* g^*(t, \tau) = \delta(t - \tau). \quad (4.4.20)$$

Поменяв здесь местами аргументы t и τ , можно после этого заменить весовую функцию $g^*(\tau, t)$ сопряженной системы весовой функцией $g(t, \tau)$ основной системы. Тогда весовая функция $g(t, \tau)$ основной системы, как функция второго аргумента τ при фиксированном значении t , определится дифференциальным уравнением

$$F_t^* g(t, \tau) = \delta(t - \tau), \quad (4.4.21)$$

где индекс t указывает, что все коэффициенты оператора F рассматриваются как функции τ и дифференцирование производится по переменной τ при фиксированном значении t . На основании формулы (4.4.17) уравнение (4.4.21) может быть переписано в развернутом виде следующим образом:

$$\sum_{h=0}^n (-1)^h \frac{d^h}{d\tau^h} [a_h(\tau) g(t, \tau)] = \delta(t - \tau). \quad (4.4.22)$$

Это уравнение представляет собой обыкновенное линейное дифференциальное уравнение относительно весовой функции $g(t, \tau)$, рассматриваемой как функция τ при фиксированном значении t . Поэтому мы и пишем производные по τ как обыкновенные производные, а t рассматриваем как параметр, имеющий данное фиксированное значение.

Импульсную δ -функцию в правой части уравнения (4.4.22) также можно заменить соответствующими условиями в точке $\tau = t$. Для этого заметим, что если принять за независимую переменную $-\tau$, то старшим членом в уравнении (4.4.22) будет $a_n(\tau) \frac{d^n}{d(-\tau)^n} g(t, \tau)$. Следовательно, действие единичного импульса $\delta(t - \tau)$ согласно изложенному выше равноценно скачкообразному изменению $n - 1$ -й производной $\frac{d^{n-1}}{d(-\tau)^{n-1}} g(t, \tau)$ в точке

$\tau = t$ на $\frac{1}{a_n(t)}$. Поэтому, учитывая, что при $\tau > t$ функция $g(t, \tau)$ и все ее производные равны нулю, приходим к заключению, что действие единичного импульса на сопряженную систему равноценно скачкообразному изменению $n-1$ -й производной $g_\tau^{(n-1)}(t, \tau)$ от нуля до $(-1)^{n-1}/a_n(t)$, когда τ переходит через точку t , убывая, а производные $g_\tau^{(n-2)}(t, \tau), \dots, g'_\tau(t, \tau)$ и сама весовая функция $g(t, \tau)$ непрерывны и равны нулю при $\tau = t$. Таким образом, весовая функция $g(t, \tau)$, рассматриваемая как функция τ при фиксированном t , представляет собой интеграл однородного линейного дифференциального уравнения

$$F_\tau^* g(t, \tau) = 0, \quad (4.4.23)$$

удовлетворяющий условиям

$$\left. \begin{aligned} g(t, t) = g'_\tau(t, t) = \dots = g_\tau^{(n-2)}(t, t) = 0, \\ g_\tau^{(n-1)}(t, t) = \frac{(-1)^{n-1}}{a_n(t)}, \end{aligned} \right\} \quad (4.4.24)$$

где, как и раньше, положено

$$g_\tau^{(h)}(t, t) = [g_\tau^{(h)}(t, \tau)]_{\tau=t} = \left[\frac{\partial^h g(t, \tau)}{\partial \tau^h} \right]_{\tau=t}. \quad (4.4.25)$$

Перейдем к определению весовой функции $w(t, \tau)$ системы, поведение которой описывается уравнением (4.4.3). Как уже было сказано, эта система представляет собой последовательное соедине-

ние системы, осуществляющей над входной переменной линейную



Рис. 4.4.3.

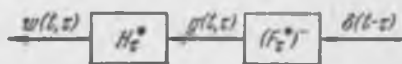


Рис. 4.4.4.

дифференциальную операцию H , и системы, описываемой дифференциальным уравнением (4.4.6). На рис. 4.4.3 эта вторая система обозначена F^- . Согласно результатам предыдущего параграфа сопряженная система представляет собой последовательное соединение системы, описываемой дифференциальным уравнением (4.4.18), и системы, осуществляющей линейную дифференциальную операцию H^* (рис. 4.4.3). Изменяя роли аргументов t и τ и подавая на вход сопряженной системы единичный импульс, мы получим на выходе первой системы весовую функцию $g(t, \tau)$, рассматриваемую как функция τ при фиксированном t , а на выходе всего последовательного соединения — весовую функцию $w(t, \tau)$ системы, описываемой уравнением (4.4.3), рассматриваемую как функция τ при фиксированном t (рис. 4.4.4). Следовательно,

$$w(t, \tau) = H_\tau^* g(t, \tau). \quad (4.4.26)$$

Подставляя сюда выражение (4.4.10) оператора H^* , получим

$$w(t, \tau) = \sum_{h=0}^m (-1)^h \frac{d^h}{d\tau^h} [b_h(\tau) g(t, \tau)]. \quad (4.4.27)$$

Заметим, что формулу (4.4.27) можно получить и непосредственно применением формулы (4.2.4) для весовой функции последовательного соединения двух линейных систем. В результате, принимая во внимание (4.4.8), получим

$$w(t, \tau) = \int_{-\infty}^{\infty} g(t, \sigma) h(\sigma, \tau) d\sigma = \sum_{h=0}^m \int_{-\infty}^{\infty} g(t, \sigma) b_h(\sigma) \delta^{(h)}(\sigma - \tau) d\sigma. \quad (4.4.28)$$

Отсюда, применяя формулу (2.1.10), получаем формулу (4.4.27).

Пример 4.4.1. Найти весовую функцию линейной системы, поведение которой описывается уравнением

$$a_1(t) \dot{y} + a_0(t) y = x. \quad (4.4.29)$$

Согласно изложенному весовая функция рассматриваемой системы определяется как интеграл однородного линейного дифференциального уравнения

$$a_1(t) g'_t(t, \tau) + a_0(t) g(t, \tau) = 0,$$

удовлетворяющий условию

$$g(\tau, \tau) = \frac{1}{a_1(\tau)}.$$

Легко проверить, что этот интеграл выражается формулой

$$g(t, \tau) = \frac{1}{a_1(\tau)} \exp \left\{ - \int_{\tau}^t \frac{a_0(\sigma)}{a_1(\sigma)} d\sigma \right\}. \quad (4.4.30)$$

Эту весовую функцию можно также рассматривать как функцию второго аргумента τ при фиксированном t и определить как интеграл сопряженного линейного дифференциального уравнения

$$-\frac{d}{d\tau} [a_1(\tau) g(t, \tau)] + a_0(\tau) g(t, \tau) = 0,$$

удовлетворяющий условию

$$g(t, t) = \frac{1}{a_1(t)}.$$

Весовая функция обратной системы получится, если в уравнении (4.4.29) рассматривать y как входную переменную, а x как выходную переменную и заменить входную переменную y единичным импульсом $\delta(t - \tau)$:

$$g^-(t, \tau) = a_1(t) \delta'(t - \tau) + a_0(t) \delta(t - \tau). \quad (4.4.31)$$

В частном случае постоянных коэффициентов a_0 и a_1 , полагая $a_0 = 1/k$, $a_1 = T/k$, приведем уравнение (4.4.29) к обычному виду уравнения аperiodического звена:

$$T\dot{y} + y = kx. \quad (4.4.32)$$

Формула (4.4.30) примет в этом случае вид

$$g(t, \tau) = \frac{k}{T} e^{-\frac{t-\tau}{T}}. \quad (4.4.33)$$

Формула (4.4.31) для весовой функции обратной системы, которая представляет собой форсирующее звено первого порядка, примет вид

$$g^-(t, \tau) = \frac{1}{k} [T\delta'(t-\tau) + \delta(t-\tau)]. \quad (4.4.34)$$

Пример 4.4.2. Найти весовую функцию системы, поведение которой описывается линейным дифференциальным уравнением

$$\ddot{y} + 2a\dot{y} + b^2y = ke^{\gamma t}(x + bx), \quad (4.4.35)$$

где a , b , k и γ — некоторые постоянные, $0 < a < b$. Найти также весовую функцию обратной системы.

Согласно изложенной теории находим сначала весовую функцию $g(t, \tau)$ системы, уравнение которой получается заменой всей правой части уравнения (4.4.35) возмущением x_1 :

$$\ddot{y} + 2a\dot{y} + b^2y = x_1.$$

Эта весовая функция может быть определена как интеграл однородного уравнения

$$g_1''(t, \tau) + 2ag_1'(t, \tau) + b^2g_1(t, \tau) = 0,$$

удовлетворяющий условиям $g(\tau, \tau) = 0$, $g_1'(\tau, \tau) = 1$.

Этот интеграл находится обычным способом интегрирования линейных дифференциальных уравнений с постоянными коэффициентами. В результате, полагая $\omega_0 = \sqrt{b^2 - a^2}$, получим

$$g(t, \tau) = \frac{e^{-a(t-\tau)}}{2i\omega_0} [e^{i\omega_0(t-\tau)} - e^{-i\omega_0(t-\tau)}]$$

или, в действительной форме,

$$g(t, \tau) = \frac{1}{\omega_0} e^{-a(t-\tau)} \sin \omega_0(t-\tau). \quad (4.4.36)$$

Эти формулы определяют весовую функцию колебательного звена.

Для определения весовой функции $w(t, \tau)$ системы, поведение которой описывается уравнением (4.4.35), применим формулу (4.4.27). В результате получим

$$\begin{aligned} w(t, \tau) &= -k \frac{d}{d\tau} [e^{\gamma\tau} g(t, \tau)] + kbe^{\gamma\tau} g(t, \tau) = \\ &= \frac{k}{2i\omega_0} e^{-a_1 t + a_1 \tau} [(b - a_1 + i\omega_0) e^{i\omega_0(t-\tau)} - (b - a_1 - i\omega_0) e^{-i\omega_0(t-\tau)}], \end{aligned} \quad (4.4.37)$$

где $a_1 = a + \gamma$.

Для определения весовой функции обратной системы будем рассматривать y в уравнении (4.4.35), как входное возмущение, а x — как выходную переменную и применим тот же метод. Сначала находим весовую функцию $p(t, \tau)$ упрощенной системы, уравнение которой получится, если в уравнении (4.4.35) заменить всю левую часть одним возмущением y_1 :

$$ke^{\gamma t}x + kbe^{\gamma t}x = y_1.$$

Это уравнение представляет собой частный случай уравнения (4.4.29), когда

$a_0(t) = kbe^{\gamma t}$, $a_1(t) = ke^{\gamma t}$. Поэтому для определения весовой функции $p(t, \tau)$ можно воспользоваться формулой (4.4.30). В результате получим

$$p(t, \tau) = \frac{1}{k} e^{-bt + (b-\gamma)\tau}. \quad (4.4.38)$$

Для определения весовой функции $w^-(t, \tau)$ системы, обратной по отношению к описываемой уравнением (4.4.35), воспользуемся формулой (4.4.27). Тогда получим

$$w^-(t, \tau) = p_{\tau}^{\sim}(t, \tau) - 2ap_{\tau}'(t, \tau) + b^2 p(t, \tau). \quad (4.4.39)$$

При вычислении производных функции $p(t, \tau)$ необходимо учесть, что формула (4.4.38) определяет ее только при $t \geq \tau$, а при $t < \tau$ она тождественно равна нулю. Тогда получим

$$\begin{aligned} p_{\tau}'(t, \tau) &= \frac{1}{k} e^{-bt} \frac{d}{d\tau} [e^{(b-\gamma)\tau} \mathbf{1}(t-\tau)] = \\ &= \frac{b-\gamma}{k} e^{-bt + (b-\gamma)\tau} \mathbf{1}(t-\tau) - \frac{1}{k} e^{-bt + (b-\gamma)\tau} \delta(t-\tau). \end{aligned}$$

Учитывая, что $\delta(t-\tau) = 0$ при всех $\tau \neq t$, можно заменить в последнем слагаемом множитель $e^{(b-\gamma)\tau}$ величиной $e^{(b-\gamma)t}$, так как это ничего не изменит. В результате получим

$$p_{\tau}'(t, \tau) = \frac{b-\gamma}{k} e^{-bt + (b-\gamma)\tau} \mathbf{1}(t-\tau) - \frac{1}{k} e^{-\gamma t} \delta(t-\tau).$$

Дифференцируя эту формулу по τ , получим

$$p_{\tau}^{\sim}(t, \tau) = \frac{(b-\gamma)^2}{k} e^{-bt + (b-\gamma)\tau} \mathbf{1}(t-\tau) - \frac{b-\gamma}{k} e^{-\gamma t} \delta(t-\tau) + \frac{1}{k} e^{-\gamma t} \delta'(t-\tau).$$

Подставляя полученные выражения $p_{\tau}'(t, \tau)$ и $p_{\tau}^{\sim}(t, \tau)$ в (4.4.39), после элементарных преобразований найдем следующее окончательное выражение весовой функции $w^-(t, \tau)$ при $t \geq \tau$:

$$\begin{aligned} w^-(t, \tau) &= \frac{2b(b-a) - \gamma(2b-2a-\gamma)}{k} e^{-bt + (b-\gamma)\tau} + \\ &+ \frac{1}{k} e^{-\gamma t} [(2a-b+\gamma)\delta(t-\tau) + \delta'(t-\tau)]. \end{aligned} \quad (4.4.40)$$

Применим теперь изложенную теорию к стационарным системам. В этом случае по доказанному в § 2.5 все коэффициенты $a_0, a_1, \dots, a_n, b_0, b_1, \dots, b_m$ операторов F и H постоянны. При этом удобно в явной форме показывать, что операторы F и H являются полиномами относительно оператора дифференцирования по времени $D = d/dt$, и писать, как в § 2.5, $F(D)$ и $H(D)$ вместо F и H . Тогда уравнение (4.4.3) примет вид

$$F(D)y = H(D)x. \quad (4.4.41)$$

Вследствие постоянства коэффициентов $b_0, b_1, \dots, b_m$ формула (4.4.10) дает следующее выражение линейного дифференциального оператора, сопряженного с оператором H :

$$H^*(D)\xi = \sum_{\lambda=0}^m (-1)^{\lambda} D^{\lambda} (b_{\lambda}\xi) = \sum_{\lambda=0}^m b_{\lambda} (-D)^{\lambda} \xi. \quad (4.4.42)$$

уравнений:

$$c_1(v_1 - v_n)e^{v_1\tau} + c_2(v_2 - v_n)e^{v_2\tau} + \dots + c_{n-1}(v_{n-1} - v_n)e^{v_{n-1}\tau} = 0,$$

$$c_1(v_1 - v_n)v_1e^{v_1\tau} + c_2(v_2 - v_n)v_2e^{v_2\tau} + \dots$$

$$\dots + c_{n-1}(v_{n-1} - v_n)v_{n-1}e^{v_{n-1}\tau} = 0,$$

$$\dots$$

$$c_1(v_1 - v_n)v_1^{n-2}e^{v_1\tau} + c_2(v_2 - v_n)v_2^{n-2}e^{v_2\tau} + \dots$$

$$\dots + c_{n-1}(v_{n-1} - v_n)v_{n-1}^{n-2}e^{v_{n-1}\tau} = \frac{1}{a_n}.$$

Вычитая из каждого последующего уравнения, начиная с последнего, предыдущее уравнение, умноженное на v_{n-1} , мы исключим из полученных уравнений величину c_{n-1} . При этом в коэффициентах при неизвестных $c_1, c_2, \dots, c_{n-2}$ появятся соответственно дополнительные множители $v_1 - v_{n-1}, v_2 - v_{n-1}, \dots, v_{n-2} - v_{n-1}$. Продолжая этот процесс, мы по очереди исключим все неизвестные $c_n, c_{n-1}, c_{n-2}, \dots, c_2$ и получим в результате одно уравнение с одной неизвестной c_1 :

$$c_1(v_1 - v_n)(v_1 - v_{n-1}) \dots (v_1 - v_3)(v_1 - v_2)e^{v_1\tau} = \frac{1}{a_n}.$$

Решая это уравнение, находим

$$c_1 = \frac{e^{-v_1\tau}}{a_n(v_1 - v_n)(v_1 - v_{n-1}) \dots (v_1 - v_2)}.$$

Но

$$a_n(v_1 - v_n)(v_1 - v_{n-1}) \dots (v_1 - v_2) = F'(v_1). \quad (4.4.47)$$

Для того чтобы убедиться в этом, достаточно разложить полином $F(v)$ на множители, продифференцировать полученное выражение по v и после этого положить $v = v_1$. Пользуясь формулой (4.4.47), можем переписать полученную для c_1 формулу в виде

$$c_1 = \frac{e^{-v_1\tau}}{F'(v_1)}.$$

А так как уравнения (4.4.46) совершенно симметричны относительно неизвестных $c_1, \dots, c_n$, то из полученной формулы вытекает следующая общая формула для коэффициентов $c_1, \dots, c_n$:

$$c_r = \frac{e^{-v_r\tau}}{F'(v_r)} \quad (r = 1, \dots, n). \quad (4.4.48)$$

Подставляя выражение (4.4.48) в (4.4.45), получим

$$g(t - \tau) = \sum_{r=1}^n \frac{1}{F'(v_r)} e^{v_r(t-\tau)}. \quad (4.4.49)$$

Для нахождения весовой функции $w(t - \tau)$ системы, описываемой уравнением (4.4.41), воспользуемся формулой (4.4.26). Тогда, принимая во внимание (4.4.43), получим

$$w(t - \tau) = H \left(-\frac{d}{d\tau} \right) g(t - \tau) = \sum_{r=1}^n \frac{H(v_r)}{F'(v_r)} e^{v_r(t-\tau)}. \quad (4.4.50)$$

При пользовании формулами (4.4.49) и (4.4.50) следует иметь в виду, что при $m < n$ они определяют весовые функции $g(t - \tau)$ и $w(t - \tau)$ только при $t > \tau$, а при $t < \tau$ эти весовые функции равны нулю. Если $m > n$, то формула (4.4.50) выражает $w(t - \tau)$ только при $t > \tau$. Чтобы определить $w(t - \tau)$ и при $t = \tau$, в этом случае необходимо добавить линейную комбинацию функций $\delta(t - \tau)$, $\delta'(t - \tau)$, ..., $\delta^{(m-n)}(t - \tau)$ с соответственно определенными коэффициентами. Для нахождения этих коэффициентов достаточно применить общую формулу (4.4.27), выполнив в ней дифференцирования с учетом разрывов дифференцируемых функций, как это сделано в примере 4.4.2.

Пример 4.4.3. Найти по формуле (4.4.49) весовую функцию колебательного звена. Подставляя в формулу (4.4.49) выражение

$$F(s) = \frac{1}{k} (s^2 + 2as + b^2)$$

и имея в виду, что корни этого полинома равны

$$v_1 = -a + i\omega_0, \quad v_2 = -a - i\omega_0, \quad \text{где } \omega_0 = \sqrt{b^2 - a^2},$$

и что

$$F'(v_1) = \frac{2i\omega_0}{k}, \quad F'(v_2) = -\frac{2i\omega_0}{k},$$

получим

$$\begin{aligned} g(t - \tau) &= \frac{k}{2i\omega_0} e^{-a(t-\tau)} [e^{i\omega_0(t-\tau)} - e^{-i\omega_0(t-\tau)}] = \\ &= \frac{k}{\omega_0} e^{-a(t-\tau)} \sin \omega_0(t - \tau). \end{aligned} \quad (4.4.51)$$

Это выражение при $k = 1$ совпадает с найденным в примере 4.4.2. Полагая в соответствии с § 2.6 $b = 1/T$, $a = \zeta/T$ и включая множитель T^2 в коэффициент k , приведем формулу (4.4.51) к виду

$$g(t - \tau) = \frac{k}{T \sqrt{1 - \zeta^2}} e^{-\frac{\zeta}{T}(t-\tau)} \sin \frac{\sqrt{1 - \zeta^2}}{T} (t - \tau). \quad (4.4.52)$$

Пример 4.4.4. Найти весовую функцию стационарной линейной системы, поведение которой описывается уравнением

$$\ddot{y} + 2a\dot{y} + b^2y = k(\dot{x} + bx), \quad (4.4.53)$$

где $b > a > 0$. Замечая, что в данном случае

$$H(s) = k(s + b), \quad F(s) = s^2 + 2as + b^2,$$

находим корни полинома $F(s)$:

$$v_1 = -a + i\omega_0, \quad v_2 = -a - i\omega_0, \quad \text{где } \omega_0 = \sqrt{b^2 - a^2}.$$

Так как в данном случае

$$F'(v_1) = 2i\omega_0, \quad F'(v_2) = -2i\omega_0, \\ H(v_1) = k(b - a + i\omega_0), \quad H(v_2) = k(b - a - i\omega_0),$$

то формула (4.4.50) дает

$$w(t - \tau) = \frac{k}{2i\omega_0} e^{-a(t-\tau)} [(b - a + i\omega_0) e^{i\omega_0(t-\tau)} - (b - a - i\omega_0) e^{-i\omega_0(t-\tau)}] \quad (4.4.54)$$

или, в действительной форме,

$$w(t - \tau) = \frac{k}{\omega_0} e^{-a(t-\tau)} [(b - a) \sin \omega_0(t - \tau) + \omega_0 \cos \omega_0(t - \tau)]. \quad (4.4.55)$$

Формула (4.4.54) является, конечно, частным случаем формулы (4.4.37), так как уравнение (4.4.53) получается как частный случай уравнения (4.4.35) при $\gamma = 0$.

§ 4.5. Линейная система, описываемая системой дифференциальных уравнений

На практике поведение линейной автоматической системы обычно описывается не одним обыкновенным дифференциальным уравнением, а системой таких уравнений. При этом каждое из уравнений r -го порядка, входящее в систему, можно заменить системой r дифференциальных уравнений первого порядка. Такая замена удобна, например, при решении дифференциальных уравнений на аналоговых или цифровых вычислительных машинах. Таким образом, мы должны рассмотреть методы определения весовых функций автоматических систем, описываемых системой дифференциальных уравнений вида

$$\dot{y}_k = \sum_{l=1}^n a_{kl} y_l + x_k \quad (k = 1, \dots, n). \quad (4.5.1)$$

Если автоматическая система описывается уравнениями (4.5.1), то ее можно рассматривать как систему с n входами и n выходами. При этом, конечно, не обязательно на все входы должны быть поданы входные сигналы и не обязательно все выходы могут интересоваться исследователя. Функция x_v представляет собой входной сигнал на v -м входе, а y_μ — выходной сигнал на μ -м выходе.

Покажем прежде всего, каким образом можно заменить одно дифференциальное уравнение r -го порядка системой r дифференциальных уравнений первого порядка, не содержащих производных от входного сигнала. Рассмотрим дифференциальное уравнение

$$y^{(r)} + a_{r-1} y^{(r-1)} + \dots + a_1 \dot{y} + a_0 y = \\ = b_{r-1} x^{(r-1)} + \dots + b_1 \dot{x} + b_0 x \quad (4.5.2)$$

В этом уравнении коэффициент при старшей производной выходной переменной положен равным единице, что, очевидно, не уменьшает общности, так как если этот коэффициент не равен нулю, то на него можно разделить обе части уравнения. Положим $y_1 = y$ и запишем следующую систему r уравнений первого порядка:

$$\begin{aligned} \dot{y}_1 &= y_2 + q_1 x, \\ \dot{y}_2 &= y_3 + q_2 x, \\ &\dots \dots \dots \\ \dot{y}_{r-2} &= y_{r-1} + q_{r-2} x, \\ \dot{y}_{r-1} &= y_r + q_{r-1} x, \\ \dot{y}_r &= -a_0 y_1 - a_1 y_2 - \dots - a_{r-1} y_r + q_r x, \end{aligned} \quad (4.5.3)$$

где $q_v (v = 1, \dots, n)$ — неопределенные пока переменные или постоянные коэффициенты. Определим коэффициенты q_v таким образом, чтобы система уравнений (4.5.3) была эквивалентна одному дифференциальному уравнению (4.5.2). Для этого исключим из уравнений (4.5.3) все переменные, кроме $y_1 = y$. Сначала исключим из последнего уравнения y_r . Для этого подставим выражение y_r из предпоследнего уравнения (4.5.3) в последнее уравнение:

$$\begin{aligned} \dot{y}_{r-1} + a_{r-1} y_{r-1} + a_{r-2} y_{r-1} = \\ = -a_0 y_1 - a_1 y_2 - \dots - a_{r-3} y_{r-2} + \frac{d}{dt} (q_{r-1} x) + a_{r-1} q_{r-1} x + q_r x. \end{aligned} \quad (4.5.4)$$

Чтобы исключить переменную y_{r-1} , подставим в (4.5.4) выражение y_{r-1} из третьего снизу уравнения (4.5.3). После этого из полученного уравнения аналогичным способом можно будет исключить y_{r-2} . Продолжая таким образом исключать переменные, мы придем к следующему уравнению, содержащему лишь одну переменную $y_1 = y$:

$$\begin{aligned} y^{(r)} + a_{r-1} y^{(r-1)} + \dots + a_1 \dot{y} + a_0 y = \\ = \frac{d^{r-1}}{dt^{r-1}} (q_1 x) + \\ + a_{r-1} \frac{d^{r-2}}{dt^{r-2}} (q_1 x) + \frac{d^{r-2}}{dt^{r-2}} (q_2 x) + \\ + a_{r-2} \frac{d^{r-3}}{dt^{r-3}} (q_1 x) + a_{r-1} \frac{d^{r-3}}{dt^{r-3}} (q_2 x) + \frac{d^{r-3}}{dt^{r-3}} (q_3 x) + \\ \dots \dots \dots \end{aligned}$$

$$\begin{aligned}
& + a_2 \frac{d}{dt} (q_1 x) + a_3 \frac{d}{dt} (q_2 x) + a_4 \frac{d}{dt} (q_3 x) + \dots + \\
& + a_{r-1} \frac{d}{dt} (q_{r-2} x) + \frac{d}{dt} (q_{r-1} x) + a_1 q_1 x + \\
& + a_2 q_2 x + a_3 q_3 x + \dots + a_{r-2} q_{r-2} x + a_{r-1} q_{r-1} x + q_r x. \quad (4.5.5)
\end{aligned}$$

Левая часть этого уравнения тождественна левой части уравнения (4.5.2). Теперь надо выбрать коэффициенты $q_v (v = 1, \dots, n)$ так, чтобы правые части уравнений (4.5.5) и (4.5.2) также совпадали. Для этого следует приравнять коэффициенты при производных входной переменной x одинакового порядка в правых частях (4.5.5) и (4.5.2) и решить полученные уравнения относительно неизвестных коэффициентов $q_v (v = 1, \dots, n)$.

Заметим, что $x^{(r-1)}$ входит только в первую строчку правой части уравнения (4.5.5), $x^{(r-2)}$ — только в первые две строчки, $x^{(r-3)}$ — в первые три строчки и т. д. Поэтому коэффициент q_1 определяется сразу: $q_1 = b_{r-1}$. После этого коэффициент q_2 выражается через коэффициенты q_1 , a_{r-1} и b_{r-2} . Затем определяется коэффициент q_3 и т. д. Таким образом, все коэффициенты $q_v (v = 1, \dots, n)$ определяются последовательно. При этом каждый последующий коэффициент определяется через предыдущие.

Полагая теперь в (4.5.3)

$$q_v x = x_v \quad (v = 1, \dots, n), \quad (4.5.6)$$

приведем уравнения (4.5.3) к виду (4.5.1).

Пример 4.5.1. Заменить дифференциальное уравнение

$$\ddot{y} + a_1 \dot{y} + a_0 y = b_1 \dot{x} + b_2 x \quad (4.5.7)$$

эквивалентной системой дифференциальных уравнений первого порядка. Согласно изложенному полагаем $y_1 = y$ и пишем уравнения

$$\left. \begin{aligned} \dot{y}_1 &= y_2 + q_1 x, \\ \dot{y}_2 &= -a_0 y_1 - a_1 y_2 + q_2 x. \end{aligned} \right\} \quad (4.5.8)$$

Исключая y_2 , получаем

$$\ddot{y} + a_1 \dot{y} + a_0 y = \frac{d}{dt} (q_1 x) + a_1 q_1 x + q_2 x = q_1 \dot{x} + (q_1 + a_1 q_1 + q_2) x. \quad (4.5.9)$$

Приравнявая коэффициенты при x и $\dot{x}$ в правых частях уравнений (4.5.7) и (4.5.9), получаем

$$\left. \begin{aligned} q_1 &= b_1 \\ q_2 &= b_2 - a_1 b_1 - b_1. \end{aligned} \right\} \quad (4.5.10)$$

Покажем теперь, как определяются весовые функции автоматической системы, описываемой системой дифференциальных уравнений (4.5.1). Так как такая система имеет n входов и n выходов,

то, согласно изложенному в § 2.2, для ее полного описания необходимо определить совокупность n^2 весовых функций $g_{kh}(t, \tau)$ ($k, h = 1, \dots, n$), соответствующих всем входам и всем выходам. Для определения весовых функций $g_{kh}(t, \tau)$ ($k = 1, \dots, n$) необходимо проинтегрировать систему (4.5.1) при

$$x_h = \delta(t - \tau), \quad x_v = 0 \quad (v = 1, \dots, h-1, h+1, \dots, n).$$

При этом весовой функцией $g_{kh}(t, \tau)$ будет являться переменная на k -м выходе системы. Импульсную δ -функцию, входящую в h -ю строку системы, можно заменить единичным начальным значением h -й переменной y_h , так как при интегрировании δ -функции возникает единичный скачок. Иными словами, для определения весовой функции $g_{kh}(t, \tau)$ как функции первого аргумента при фиксированном втором надо интегрировать однородную систему дифференциальных уравнений

$$\frac{d}{dt} g_{kh}(t, \tau) = \sum_{l=1}^n a_{kl}(t) g_{lh}(t, \tau) \quad (k = 1, \dots, n) \quad (4.5.11)$$

при начальных условиях

$$g_{hh}(\tau, \tau) = 1, \quad g_{kh}(\tau, \tau) = 0 \text{ при } k \neq h \quad (4.5.12)$$

и наблюдать k -ю переменную.

Для определения весовой функции $g_{kh}(t, \tau)$ как функции второго аргумента воспользуемся методом сопряженных систем, изложенным в § 4.3. Для простоты рассмотрим сначала систему двух дифференциальных уравнений первого порядка

$$\left. \begin{aligned} \dot{y}_1 &= a_{11}y_1 + a_{12}y_2 + x_1, \\ \dot{y}_2 &= a_{21}y_1 + a_{22}y_2 + x_2. \end{aligned} \right\} \quad (4.5.13)$$

Структурная схема такой системы изображена на рис. 4.5.1. Пользуясь правилами, сформулированными в § 4.3, построим сопряженную систему (рис. 4.5.2). Как показано в § 4.3, чтобы определить весовую функцию системы как функцию второго аргумента τ , необходимо моделировать сопряженную систему в обратном времени $\sigma = t - \tau$, подав на соответствующий вход единичный импульс $\delta(\sigma)$. Так, если нас интересует первый выход системы, то надо подать δ -функцию $\delta(\sigma)$ на первый вход сопряженной системы. Тогда переменная $z_1(\sigma)$ на первом выходе сопряженной системы будет представлять собой весовую функцию $g_{11}(t, t - \sigma)$, связывающую первый вход с первым выходом системы, а $z_2(\sigma)$ — весовую функцию $g_{12}(t, t - \sigma)$, связывающую второй вход с первым выходом системы. Если δ -функцию $\delta(\sigma)$ подать на второй вход сопряженной системы, то переменная $z_1(\sigma)$ будет представ-

лать собой весовую функцию $g_{21}(t, t - \sigma)$, а $z_2(\sigma)$ — весовую функцию $g_{22}(t, t - \sigma)$. Естественно, что практически δ -функции на входах сопряженной системы заменяются единичными начальными значениями переменных z_1 или z_2 в зависимости от того, для какого выхода требуется получить весовые функции.

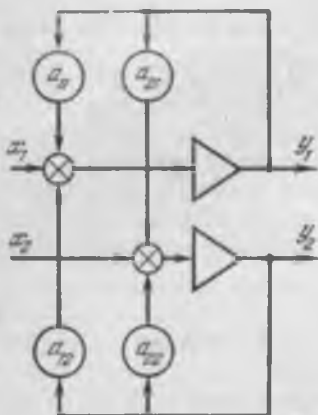


Рис. 4.5.1.

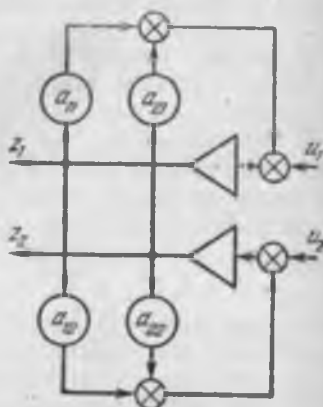


Рис. 4.5.2.

Переходя от структурной схемы сопряженной системы к системе дифференциальных уравнений в соответствии с рис. 4.5.2, получаем

$$\left. \begin{aligned} \frac{dz_1}{d\sigma} &= a_{11}z_1 + a_{21}z_2, \\ \frac{dz_2}{d\sigma} &= a_{12}z_1 + a_{22}z_2, \end{aligned} \right\} \quad (4.5.14)$$

где коэффициенты a_{ij} рассматриваются как функции переменной $\sigma = t - \tau$. На основании изложенного заключаем, что для определения весовой функции $g_{rh}(t, \tau)$ ($r, h = 1, 2$), $\tau = t - \sigma$, необходимо интегрировать систему дифференциальных уравнений (4.5.14) при начальных условиях

$$z_r(0) = 1, \quad z_k(0) = 0 \quad (k \neq r). \quad (4.5.15)$$

При этом переменная интегрирования σ в системе (4.5.14) пробегает значения от 0 до $t - t_0$, в то время как реальное время изменяется в обратном направлении от t до t_0 . Если перейти от обратного времени σ к реальному $\tau = t_0 - \sigma$, то уравнения (4.5.14) запишутся в виде

$$\left. \begin{aligned} \dot{z}_1 &= -a_{11}z_1 - a_{21}z_2, \\ \dot{z}_2 &= -a_{12}z_1 - a_{22}z_2. \end{aligned} \right\} \quad (4.5.16)$$

начальные условия (4.5.15) заменяются граничными условиями

$$z_r(t) = 1, \quad z_k(t) = 0 \quad (k \neq r). \quad (4.5.17)$$

Система дифференциальных уравнений (4.5.16) называется сопряженной с системой (4.5.14). Сравнивая (4.5.14) и (4.5.16), видим, что сопряженная система получается из исходной по следующим правилам. В исходной системе дифференциальных уравнений отбрасываются входные сигналы x_k , знаки перед коэффициентами a_{ki} заменяются на обратные, а матрица коэффициентов транспонируется, т. е. коэффициенты, стоящие в k -й строке, заменяются коэффициентами, стоящими в k -м столбце.

Покажем, что сопряженная система уравнений, построенная по этим правилам, определяет весовые функции исходной системы как функции второго аргумента и в общем случае, т. е. в случае системы n дифференциальных уравнений первого порядка. Система, сопряженная с системой (4.5.1), будет иметь вид

$$\dot{z}_k = - \sum_{l=1}^n a_{lk} z_l \quad (k=1, \dots, n). \quad (4.5.18)$$

Умножим k -е уравнение (4.5.18) на $g_{kh}(t, \tau)$, а k -е уравнение (4.5.11) на $z_k(t)$, заменив для удобства дальнейших выкладок аргумент t на σ , и сложим все полученные уравнения:

$$\begin{aligned} \sum_{k=1}^n \left[g_{kh}(\sigma, \tau) \frac{d}{d\sigma} z_k(\sigma) + z_k(\sigma) \frac{d}{d\sigma} g_{kh}(\sigma, \tau) \right] = \\ = - \sum_{k=1}^n \sum_{l=1}^n a_{lk}(\sigma) z_l(\sigma) g_{kh}(\sigma, \tau) + \sum_{k=1}^n \sum_{l=1}^n a_{kl}(\sigma) z_k(\sigma) g_{lh}(\sigma, \tau). \end{aligned} \quad (4.5.19)$$

Так как значения сумм не зависят от того, какой буквой обозначить индексы суммирования, то в одной из двойных сумм в правой части полученной формулы можно поменять обозначения l на k , а k на l . Тогда сразу будет видно, что обе двойные суммы тождественны и вся правая часть формулы (4.5.19) равна нулю, и мы получим

$$\sum_{k=1}^n \frac{d}{d\sigma} [g_{kh}(\sigma, \tau) z_k(\sigma)] = 0. \quad (4.5.20)$$

Интегрируем полученное равенство в пределах от τ до t :

$$\sum_{k=1}^n [g_{kh}(t, \tau) z_k(t) - g_{kh}(\tau, \tau) z_k(\tau)] = 0. \quad (4.5.21)$$

Имея в виду, что система (4.5.11) интегрируется при начальных условиях (4.5.12), получаем из (4.5.21):

$$z_h(\tau) = \sum_{k=1}^n g_{kh}(t, \tau) z_k(t). \quad (4.5.22)$$

Теперь видно, что если сопряженную систему уравнений интегрировать при граничных условиях при $\tau = t$:

$$z_r(t) = 1, \quad z_k(t) = 0 \text{ при } k \neq r, \quad (4.5.23)$$

то будем иметь

$$z_h(\tau) = g_{rh}(t, \tau) \quad (h = 1, \dots, n). \quad (4.5.24)$$

Итак, мы доказали, что и в общем случае системы n дифференциальных уравнений первого порядка сопряженная система определяет весовые функции как функции второго аргумента τ при фиксированном значении первого аргумента t . Таким образом, для определения весовой функции системы $g_{rh}(t, \tau)$ как функции τ необходимо интегрировать систему дифференциальных уравнений

$$\frac{dg_{rh}(t, \tau)}{d\tau} = - \sum_{l=1}^n a_{lh}(\tau) g_{rl}(t, \tau) \quad (h = 1, \dots, n) \quad (4.5.25)$$

при граничных условиях при $\tau = t$:

$$g_{rr}(t, t) = 1, \quad g_{rh}(t, t) = 0, \text{ когда } h \neq r. \quad (4.5.26)$$

Интегрирование сопряженной системы с заданными граничными условиями не представляет никаких практических трудностей, так как путем замены переменной интегрирования по формуле $\tau = t - \sigma$ заданные конечные значения (4.5.23) переменных $z_1, \dots, z_n$ при $\tau = t$ заменяются начальными значениями (4.5.15) при $\sigma = 0$.

Пример 4.5.2. Найти весовые функции $g_{11}(t, \tau)$ и $g_{12}(t, \tau)$ как функции τ при фиксированном t для системы, описываемой системой дифференциальных уравнений

$$\left. \begin{aligned} \dot{y}_1 &= a_{11}(t) y_1 + x_1, \\ \dot{y}_2 &= a_{21}(t) y_1 + a_{22}(t) y_2 + x_2. \end{aligned} \right\} \quad (4.5.27)$$

На основании изложенного искомые весовые функции определяются системой дифференциальных уравнений

$$\left. \begin{aligned} \frac{dg_{11}(t, \tau)}{d\tau} &= -a_{11}(\tau) g_{11}(t, \tau), \\ \frac{dg_{12}(t, \tau)}{d\tau} &= -a_{12}(\tau) g_{11}(t, \tau) - a_{22}(\tau) g_{12}(t, \tau) \end{aligned} \right\} \quad (4.5.28)$$

при граничных условиях при $\tau = t$:

$$g_{11}(t, t) = 1, \quad g_{12}(t, t) = 0. \quad (4.5.29)$$

Если сделать замену переменной интегрирования по формуле $\tau = t - \sigma$ и рассматривать переменные как функции σ , то будем иметь

$$\left. \begin{aligned} \frac{dg_{11}(t, t-\sigma)}{d\sigma} &= -a_{21}(t-\sigma)g_{12}(t, t-\sigma), \\ \frac{dg_{12}(t, t-\sigma)}{d\sigma} &= a_{12}(t-\sigma)g_{11}(t, t-\sigma) + a_{22}(t-\sigma)g_{12}(t, t-\sigma). \end{aligned} \right\} \quad (4.5.30)$$

В этом случае условия (4.5.29) выполняются при начальном значении $\sigma = 0$.

§ 4.6. Соединения стационарных линейных систем

Выведем теперь зависимости между передаточными функциями соединений стационарных линейных систем и передаточными функциями соединяемых систем.

Рассмотрим последовательное соединение двух стационарных линейных систем с передаточными функциями $\Phi_1(s)$ и $\Phi_2(s)$ (рис. 4.6.1). Предположим, что на входе этого последовательного соединения неограниченно долго действует возмущение, представляющее собой показательную функцию времени e^{st} . Это возмущение при прохождении через первую систему умножается на передаточную функцию $\Phi_1(s)$. Таким образом, на выходе первой системы и на входе второй будет функция $\Phi_1(s)e^{st}$. В силу линейности второй системы ее выходная переменная будет равна произведению ее реакции на показательное возмущение e^{st} на постоянный множитель $\Phi_1(s)$, т. е. $\Phi_1(s)\Phi_2(s)e^{st}$. Таким образом, последовательное соединение стационарных линейных систем дает стационарную линейную систему, передаточная функция которой $\Phi(s)$ равна произведению передаточных функций соединяемых систем:

$$\Phi(s) = \Phi_1(s)\Phi_2(s). \quad (4.6.1)$$

Формула (4.6.1) легко распространяется на последовательное соединение любого числа стационарных линейных систем. А именно, если передаточные функции n соединяемых последовательно систем равны $\Phi_1(s), \dots, \Phi_n(s)$, то передаточная функция соединения $\Phi(s)$ определяется формулой

$$\Phi(s) = \Phi_1(s)\Phi_2(s)\dots\Phi_n(s). \quad (4.6.2)$$

Так как модуль произведения комплексных чисел равен произведению модулей сомножителей, а аргумент равен сумме аргу-

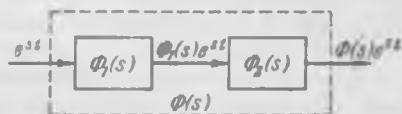


Рис. 4.6.1.

ментов сомножителей, то из (4.6.2) вытекают следующие формулы для амплитудной и фазовой частотных характеристик последовательного соединения:

$$|\Phi(i\omega)| = \prod_{k=1}^n |\Phi_k(i\omega)|, \quad (4.6.3)$$

$$\arg \Phi(i\omega) = \sum_{k=1}^n \arg \Phi_k(i\omega). \quad (4.6.4)$$

Логарифмируя формулу (4.6.3), получим

$$\lg |\Phi(i\omega)| = \sum_{k=1}^n \lg |\Phi_k(i\omega)|. \quad (4.6.5)$$

Формулы (4.6.4) и (4.6.5) показывают, что при последовательном соединении стационарных линейных систем их фазовые частотные характеристики и логарифмические амплитудные частотные характеристики суммируются.

Формула (4.6.2) показывает, что результат последовательного соединения стационарных линейных систем не зависит от порядка их соединения.

Рассмотрим теперь две взаимно обратные стационарные линейные системы. В результате их последовательного соединения получается идеальная следящая система, передаточная функция которой по доказанному в § 2.4 равна единице. Поэтому из формулы (4.6.1) следует, что передаточные функции двух взаимно обратных стационарных линейных систем являются взаимно обратными величинами. Если данная система имеет передаточную функцию $\Phi(s)$, то передаточная функция обратной системы будет $1/\Phi(s)$.

Для определения передаточной функции параллельного соединения

стационарных линейных систем с весовыми функциями $\Phi_1(s)$ и $\Phi_2(s)$ (рис. 4.6.2) предположим, что на входе этого соединения неограниченно долго действует возмущение, представляющее собой показательную функцию времени e^{st} . В результате выходные переменные соединяемых систем будут равны соответственно $\Phi_1(s)e^{st}$ и $\Phi_2(s)e^{st}$. Выходная переменная соединения будет равна сумме этих выражений. Следовательно, передаточная функция параллельного соединения стационарных линейных систем равна сумме передаточных функций соединяемых систем:

$$\Phi(s) = \Phi_1(s) + \Phi_2(s). \quad (4.6.6)$$

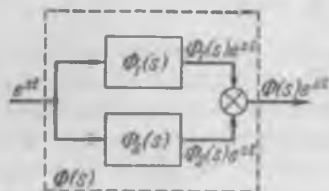


Рис. 4.6.2.

Эта формула, очевидно, также распространяется на любое число соединенных параллельно систем.

Переходим к определению передаточных функций стационарных линейных систем, замкнутых обратными связями. Рассмотрим

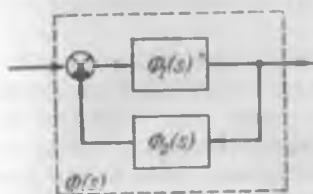


Рис. 4.6.3.

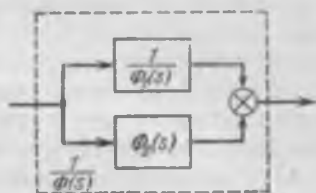


Рис. 4.6.4.

систему, состоящую из стационарной линейной системы с передаточной функцией $\Phi_1(s)$, замкнутой отрицательной обратной связью, содержащей стационарную линейную систему с передаточной функцией $\Phi_2(s)$ (рис. 4.6.3).

Для определения передаточной функции $\Phi(s)$ этой системы рассмотрим обратную систему, которая по доказанному имеет передаточную функцию $1/\Phi(s)$. На основании изложенного в § 4.2 эта обратная система представляет собой параллельное соединение системы с передаточной функцией $1/\Phi_1(s)$ и системы с передаточной функцией $\Phi_2(s)$ (рис. 4.6.4). Поэтому, пользуясь формулой (4.6.6), получаем

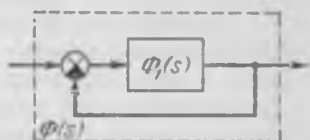


Рис. 4.6.5.

$$\frac{1}{\Phi(s)} = \frac{1}{\Phi_1(s)} + \Phi_2(s) = \frac{1 + \Phi_1(s) \Phi_2(s)}{\Phi_1(s)}.$$

Отсюда находим передаточную функцию интересующей нас системы с обратной связью (рис. 4.6.3):

$$\Phi(s) = \frac{\Phi_1(s)}{1 + \Phi_1(s) \Phi_2(s)}. \quad (4.6.7)$$

В частном случае жесткой отрицательной обратной связи (рис. 4.6.5) $\Phi_2(s) = 1$ и формула (4.6.7) принимает вид

$$\Phi(s) = \frac{\Phi_1(s)}{1 + \Phi_1(s)}. \quad (4.6.8)$$

Мы видим, что при любых соединениях стационарных линейных систем всегда получаются стационарные линейные системы, передаточные функции и частотные характеристики которых определяются при помощи элементарных алгебраических действий по данным передаточным функциям (соответственно частотным характеристикам) соединяемых систем. Вследствие этого метод

частотных характеристик получил широкое практическое распространение.

Вычисления по формулам (4.6.4) и (4.6.5) практически выполняются графически путем суммирования ординат логарифмических частотных характеристик соединяемых последовательно систем.

Для вычисления частотной характеристики системы с отрицательной жесткой обратной связью по формуле (4.6.8) обычно пользуются специальными номограммами, которые позволяют находить логарифмические частотные характеристики системы, замкнутой отрицательной жесткой обратной связью, по данным логарифмическим частотным характеристикам разомкнутой системы. Такая номограмма дана в приложении 2.

Для нахождения логарифмических частотных характеристик системы с гибкой обратной связью можно пользоваться той же номограммой. Для этого представим формулу (4.6.7) в виде

$$\Phi(s) = \frac{\Phi_1(s) \Phi_2(s)}{1 + \Phi_1(s) \Phi_2(s)} \frac{1}{\Phi_2(s)}. \quad (4.6.9)$$

Таким образом, систему с гибкой обратной связью можно представить в виде последовательного соединения системы с передаточной функцией $\Phi_1(s) \Phi_2(s)$, охваченной отрицательной жесткой обратной связью, и системы с передаточной функцией $1/\Phi_2(s)$. К этому же выводу можно прийти чисто структурными преобразованиями, которые будут рассмотрены в следующем параграфе.

Логарифмические частотные характеристики параллельного соединения двух систем также можно найти при помощи номограммы, определяющей логарифмические частотные характеристики замкнутой системы по данным логарифмическим частотным характеристикам разомкнутой системы. Для этого достаточно перейти к обратной системе. Мы знаем, что системой, обратной по отношению к параллельному соединению систем с передаточными функциями $\Phi_1(s)$ и $\Phi_2(s)$, является система, полученная в результате замыкания системы с передаточной функцией $1/\Phi_1(s)$ отрицательной обратной связью, содержащей систему с передаточной функцией $\Phi_2(s)$. Определив при помощи номограммы логарифмические частотные характеристики этой обратной системы и изменив у них знаки, мы и получим логарифмические частотные характеристики параллельного соединения систем с передаточными функциями $\Phi_1(s)$ и $\Phi_2(s)$. Применяя этот прием последовательно, можно найти логарифмические частотные характеристики параллельного соединения любого числа стационарных линейных систем.

Практические приемы построения логарифмических частотных характеристик стационарных линейных систем будут изложены в § 4.8.

Легко видеть, что при всех возможных соединениях стационарных линейных систем, описываемых обыкновенными дифференциальными уравнениями, в результате всегда получаются стационарные линейные системы, поведение которых также описывается обыкновенными дифференциальными уравнениями. Действительно, в этом случае передаточные функции всех соединяемых систем представляют собой дробно-рациональные функции s . Все выведенные формулы для передаточных функций соединений в этом случае определяют передаточную функцию соединения $\Phi(s)$ также в виде дробно-рациональной функции s . Представив эту функцию в виде отношения двух полиномов и заменив s оператором дифференцирования по времени D , мы найдем операторы $F(D)$ и $H(D)$ и получим дифференциальное уравнение вида (2.5.15), связывающее входную и выходную переменные x, y рассматриваемого соединения стационарных линейных систем.

Заметим, что на основании доказанного в конце § 2.5 определяемая любой из выведенных выше формул передаточная функция $\Phi(s)$ соединения стационарных линейных систем, описываемых обыкновенными дифференциальными уравнениями, существует при всех s , действительные части которых больше действительных частей всех полюсов функции $\Phi(s)$. При этом в случае, когда рассматриваемое соединение содержит цепи обратных связей, передаточные функции систем, входящих в соединение, могут и не существовать при всех значениях s , при которых существует $\Phi(s)$. Это, очевидно, не мешает пользоваться при вычислении $\Phi(s)$ формальными выражениями вида (2.5.12) для передаточных функций всех соединяемых систем. Этим объясняется, в частности, возможность пользоваться формальными выражениями частотных характеристик дифференциатора, интегратора, формирующих звеньев и неустойчивых апериодических ($T < 0$) и колебательных ($\zeta < 0$) звеньев для получения частотной характеристики любого соединения стационарных линейных систем. Следует лишь помнить, что полученная в результате частотная характеристика соединения физически существует только в том случае, когда все полюсы передаточной функции этого соединения имеют отрицательные действительные части.

Пример 4.6.1. Физическое последовательное соединение двух цепочек RC , рассмотренное в примере 4.1.1, с точки зрения теории автоматического управления представляет собой последовательное соединение этих двух элементов, замкнутое обратной связью, содержащей дифференцирующее звено с передаточной функцией $R_1 C_2 s$ (рис. 4.6.6). Действительно, вычислив изложенным в § 3.11 методом передаточную функцию физического последовательного соединения двух цепочек RC (рис. 4.1.5), получим

$$\Phi(s) = \frac{1}{(T_1 s + 1)(T_2 s + 1) + R_1 C_2 s} = \frac{1}{1 + \frac{R_1 C_2 s}{(T_1 s + 1)(T_2 s + 1)}},$$

где $T_1 = R_1 C_1$, $T_2 = R_2 C_2$. Сравнивая это выражение с (4.6.7) и принимая во внимание (4.6.1), убеждаемся в справедливости высказанного утверждения. И только при очень малой емкости C_2 (и пропорционально большом сопротивлении R_2 , чтобы произведение $T_2 = R_2 C_2$ оставалось неизменным) это соединение можно рассматривать как чистое последовательное соединение без обратной связи.

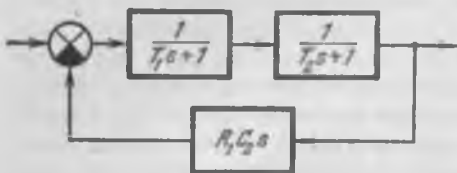


Рис. 4.6.6.

Интересно отметить, что не только в рассмотренном примере, но и вообще физическое присоединение любого элемента к выходу

любого другого элемента практически всегда добавляет к последовательному соединению ту или иную обратную связь. Иными словами, присоединение к выходу любого элемента другого элемента, представляющего собой нагрузку на выходе первого элемента, дает не только эффект последовательного соединения, но и дополнительный эффект обратной связи. Это легко проследить на примерах исполнительных устройств и длинной линии, рассмотренных в предыдущей главе. Учет сил, действующих на исполнительное устройство со стороны приводимого им в движение механизма, в уравнениях работы исполнительного устройства равноценен подаче на вход исполнительного устройства дополнительного сигнала, зависящего от его выходного сигнала, т. е. обратной связи. Точно так же учет нагрузки на конце длинной линии в уравнениях, описывающих процесс в длинной линии, равноценен охвату ее обратной связью, содержащей звено с передаточной функцией $(e^{\beta(s)l} - e^{-\beta(s)l})/2Z_n(s) \gamma(s)$. Это непосредственно следует из выражений (3.15.15) и (3.15.16) для передаточных функций длинной линии с учетом и без учета нагрузки и формулы (4.6.7).

§ 4.7. Структурные преобразования линейных систем

Мы видели, что структурная схема автоматической системы определена неоднозначно, одной и той же системе соответствует бесчисленное множество структурных схем. Это дает возможность выбирать наиболее простые и удобные для исследования структурные схемы и тем самым значительно сократить объем необходимых вычислений. Для нахождения наиболее простых и удобных структурных схем линейных систем необходимо научиться преобразовывать структурные схемы линейных систем.

Так как всякая структурная схема представляет линейную систему в виде соответствующего соединения более простых линейных систем, сумматоров и точек разветвления, то любое пре-

образование структурной схемы сводится к поочередной попарной перестановке ее соседних звеньев. Поэтому достаточно научиться переставлять различные соседние звенья структурной схемы — точки разветвления, которые мы будем называть для краткости *узлами*, сумматоры и линейные системы.

Ясно, что все преобразования в структурной схеме и, в частности, перестановки соседних звеньев должны производиться так,

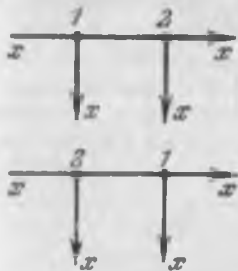


Рис. 4.7.1.

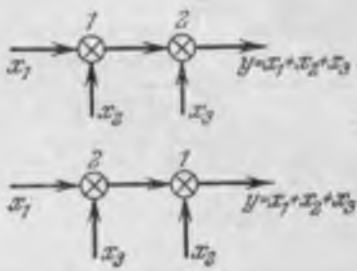


Рис. 4.7.2.

чтобы все входные и выходные сигналы каждого преобразуемого участка схемы оставались неизменными. В этом состоит основной принцип структурных преобразований, который обеспечивает равноценность исходной и преобразованной схем в том смысле, что обе они соответствуют одной и той же линейной системе.

Из сформулированного общего принципа вытекают следующие правила структурных преобразований:

1) Два узла всегда можно менять местами (рис. 4.7.1).

2) Два сумматора всегда можно менять местами (рис. 4.7.2).

3) При переносе узла через сумматор по ходу сигнала (или сумматора через узел против хода сигнала) следует добавить связь между линией второго входа сумматора и ответвлением, направленную по ходу сигнала в прямой цепи и содержащую усилитель с коэффициентом усиления -1 (рис. 4.7.3).

4) При переносе узла через сумматор против хода сигнала (или сумматора через узел по ходу сигнала) следует добавить связь между линией второго входа сумматора и ответвлением, направленную против хода сигнала в прямой цепи (рис. 4.7.4).

5) При переносе узла через линейную систему по ходу сигнала необходимо включить в ответвление обратную линейную систему (рис. 4.7.5).

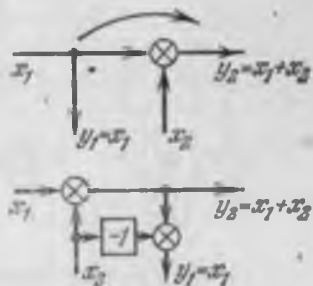


Рис. 4.7.3.

6) При переносе узла через линейную систему против хода сигнала необходимо включить в ответвление такую же линейную систему (рис. 4.7.6).

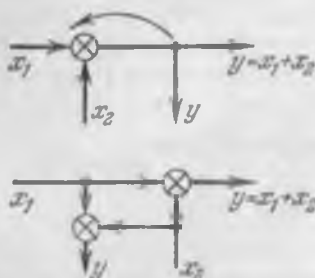


Рис. 4.7.4.

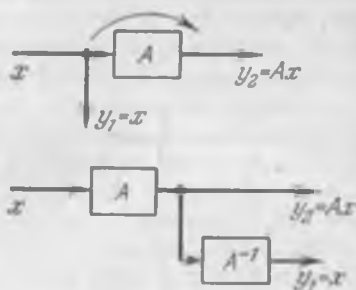


Рис. 4.7.5.

7) При переносе сумматора через линейную систему по ходу сигнала необходимо включить в линию второго входа сумматора такую же линейную систему (рис. 4.7.7).

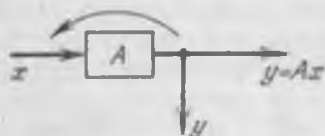


Рис. 4.7.6.

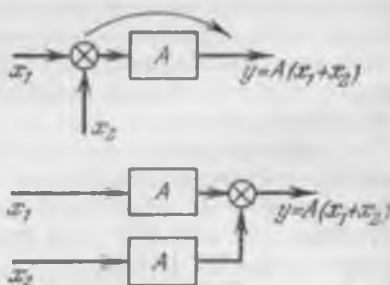


Рис. 4.7.7.

8) При переносе сумматора через линейную систему против хода сигнала необходимо включить в линию второго входа сумматора обратную систему (рис. 4.7.8).

9) Ветви параллельного соединения можно менять местами (рис. 4.7.9).

10) В системе с обратной связью можно менять местами систему, находящуюся в прямой цепи, и систему, включенную в цепь обратной связи, заменив при этом обе системы соответствующими обратными системами (рис. 4.7.10). Это правило легко выводится из предыдущего переходом к обратной системе.

Все выведенные правила структурных преобразований линейных систем являются абсолютно общими и применимы к любым линейным системам.

Для стационарных линейных систем на основании результатов предыдущего параграфа добавляется еще одно очень важное правило:

11) Последовательно соединенные стационарные линейные системы можно менять местами.

Это же правило применимо к частному виду нестационарных линейных систем, а именно к усилителям:

12) Последовательно соединенные усилители, как стационарные, так и нестационарные, можно менять местами.

Для произвольных нестационарных систем нельзя изменять порядок последовательного соединения без ввода дополнительных изменений в структурную схему.

Чтобы вывести правила перестановки нестационарных систем, заметим, что любая нестационарная система, поведение которой описывается дифференциальными уравнениями, может быть составлена из одних только дифференциаторов и усилителей с переменными коэффициентами усиления или из одних только интеграторов и усилителей с переменными коэффициентами усиления (см. §§ 4.4 и 4.5). Поэтому достаточно научиться переставлять нестационарный усилитель с дифференциатором и интегратором.

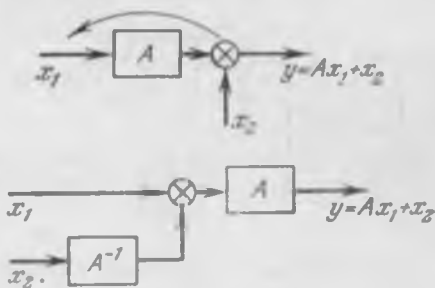


Рис. 4.7.8.

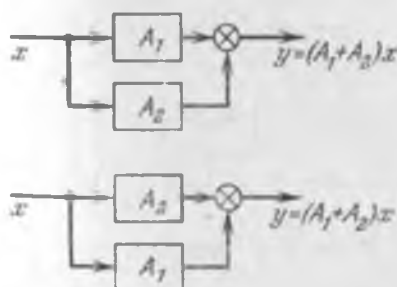


Рис. 4.7.9.

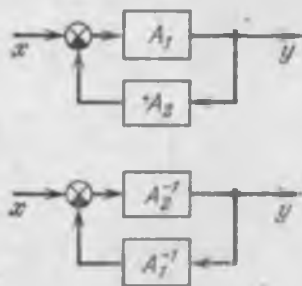


Рис. 4.7.10.

Применяя сформулированный в начале параграфа общий принцип преобразования структурных схем, получаем следующие правила:

13) При переносе усилителя с коэффициентом усиления k (ι) через дифференциатор по ходу сигнала необходимо параллельно

с дифференциатором включить усилитель с коэффициентом усиления $k'(t)/k(t)$ (рис. 4.7.11).

14) При переносе усилителя с коэффициентом усиления $k(t)$ через дифференциатор против хода сигнала необходимо параллельно с дифференциатором включить усилитель с коэффициентом усиления $-k'(t)/k(t)$ (рис. 4.7.12).

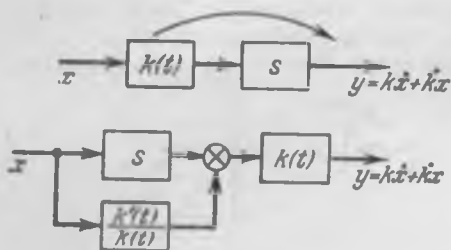


Рис. 4.7.11.

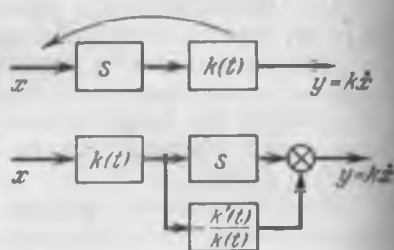


Рис. 4.7.12.

Переходом к обратным системам получаем из этих двух правил еще следующие правила:

15) При переносе усилителя с коэффициентом усиления $k(t)$ через интегратор по ходу сигнала необходимо замкнуть интегратор отрицательной обратной связью, содержащей усилитель с коэффициентом усиления $k'(t)/k(t)$ (рис. 4.7.13).

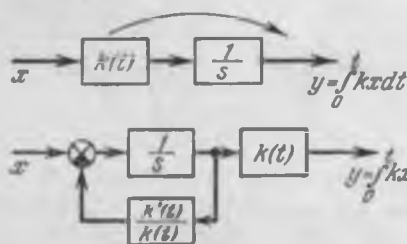


Рис. 4.7.13.

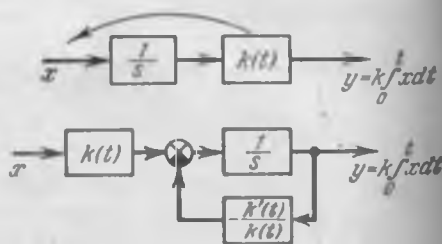


Рис. 4.7.14.

16) При переносе усилителя с коэффициентом усиления $k(t)$ через интегратор против хода сигнала необходимо замкнуть интегратор отрицательной обратной связью, содержащей усилитель с коэффициентом усиления $-k'(t)/k(t)$ (рис. 4.7.14).

Эти правила структурных преобразований нестационарных линейных систем выведены В. И. Соколовым [64] и А. В. Солодовым [66].

Пример 4.7.1. Для замены гибкой обратной связи жесткой (рис. 4.7.15, а) меняем местами узел и систему в цепи обратной связи A_2 (рис. 4.7.15, б).

Пример 4.7.2. Рассмотрим перенос параллельно включенной цепи с одного звена на другое по линии последовательно соединенных стационарных линейных систем. Так как последовательно соединенные стационарные линейные системы можно как угодно переставлять, то взаимно обратные системы, вводимые при переносе узла и сумматора через одну и ту же систему, компенсируются. Вследствие этого при переносе параллельной цепи с n -го звена на первое или наоборот необходимо учесть только перенос узла через первое звено и сумматора через n -е звено (рис. 4.7.16).

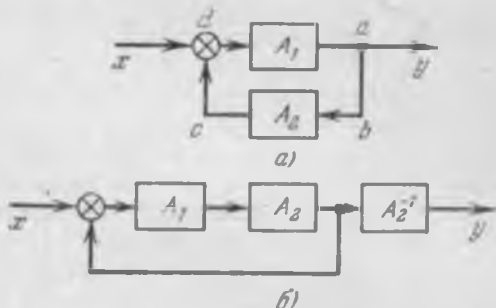


Рис. 4.7.15.

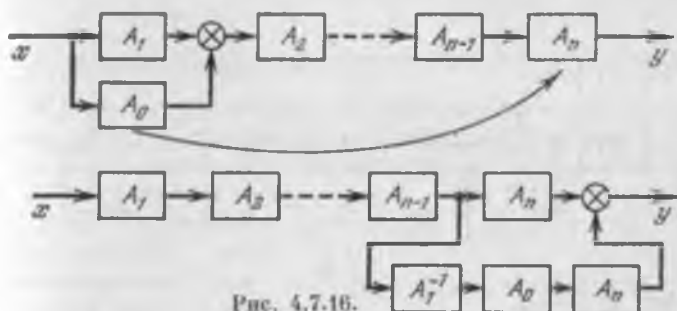


Рис. 4.7.16.

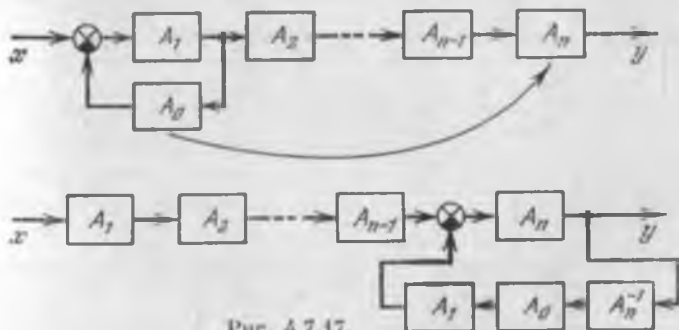


Рис. 4.7.17.

Пример 4.7.3. Совершенно так же убеждаемся в том, что при переносе цепи обратной связи с первого звена на n -е или наоборот по линии последовательно соединенных стационарных линейных систем необходимо учитывать только перенос сумматора через первое звено и узла через n -е звено (рис. 4.7.17).

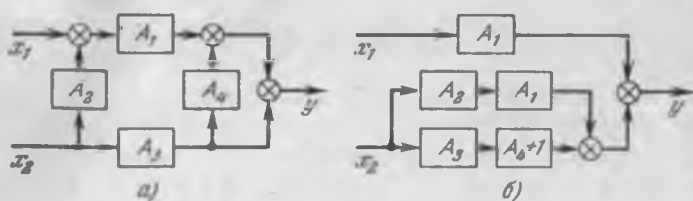


Рис. 4.7.18.

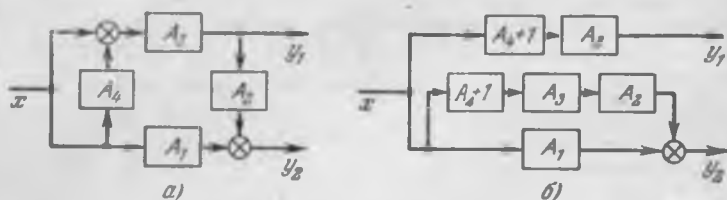


Рис. 4.7.19.

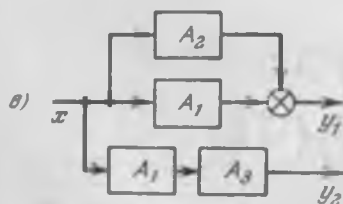
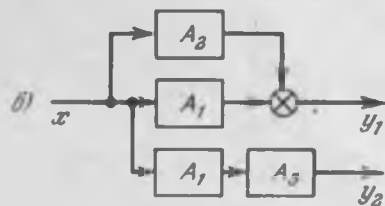
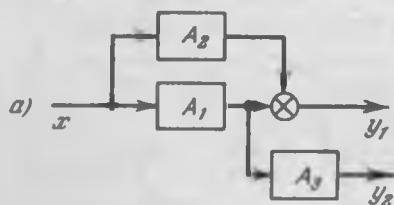


Рис. 4.7.20.

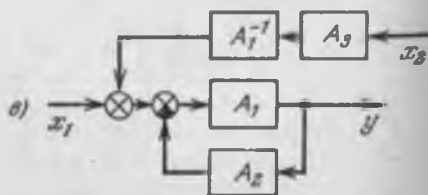
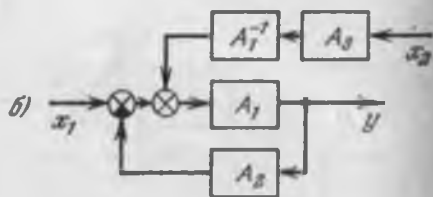
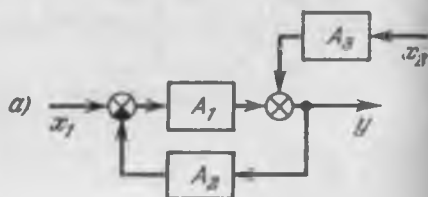


Рис. 4.7.21.

Пример 4.7.4. На рис. 4.7.18, *а* представлена схема системы с двумя входами и одним выходом. На рис. 4.7.18, *б* показана соответствующая преобразованная схема.

Пример 4.7.5. На рис. 4.7.19, *а* представлена схема с одним входом и двумя выходами. На рис. 4.7.19, *б* показана соответствующая преобразованная схема.

Выведенные правила позволяют, в частности, выносить узлы и сумматоры из петель параллельных соединений и обратных связей. Для этого достаточно выполнить такие перестановки в петле, чтобы посторонние (подлежащие выносу) узлы оказались рядом с основным узлом, а посторонние сумматоры — рядом с основным сумматором. Тогда после перестановки узлов и сумматоров все посторонние узлы и сумматоры выйдут из петли.

Пример 4.7.6. На рис. 4.7.20, *а* представлена схема с петлей параллельного соединения и добавочным узлом на одной из ветвей. На рис. 4.7.20, *б* и 4.7.20, *в* показаны последовательные этапы выноса узла из петли.

Пример 4.7.7. На рис. 4.7.21, *а* представлена схема с петлей обратной связи, внутри которой расположен сумматор (посторонний). На рис. 4.7.21, *б* и 4.7.21, *в* показаны последовательные этапы выноса сумматора из петли.

§ 4.8. Метод логарифмических частотных характеристик

В § 2.6 было показано, что логарифмические амплитудные частотные характеристики усилителей, дифференциаторов, интеграторов и запаздывающих звеньев представляют собой прямые линии, а логарифмические амплитудные характеристики форсирующих, апериодических и колебательных звеньев могут быть приближенно представлены ломаной, образованной двумя прямолинейными лучами — низкочастотной и высокочастотной асимптотами. Вследствие этого логарифмические амплитудные характеристики любых последовательных соединений элементарных звеньев приближенно изображаются ломаными. Эти ломаные обычно называются *асимптотическими* логарифмическими амплитудными частотными характеристиками.

Построение асимптотической логарифмической амплитудной характеристики последовательного соединения элементарных звеньев сводится к следующим простым операциям (рис. 4.8.1):

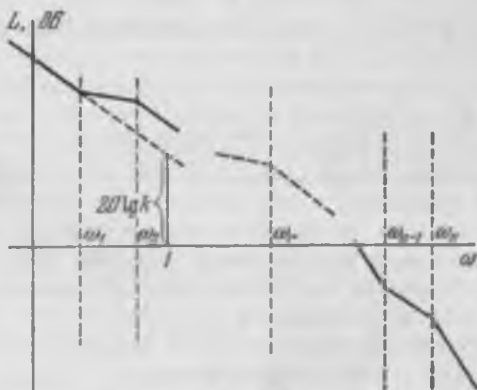


Рис. 4.8.1.

1) На оси абсцисс отмечаются (в логарифмическом масштабе) сопрягающие частоты всех форсирующих, апериодических и колебательных звеньев, входящих в соединение, и через полученные точки проводятся вертикальные прямые.

2) В точке $\omega = 1$ откладывается ордината $L(1) = 20 \lg k$, где k — общий коэффициент усиления соединения, равный произведению коэффициентов усиления всех входящих в соединение звеньев. Через полученную точку проводится прямая с наклоном $20(m - n)$ дБ/дек, где m — число дифференциаторов, а n — число интеграторов в системе. Отрезок этой прямой, расположенный левее первой вертикальной прямой (первой сопрягающей частоты), представляет собой низкочастотную часть асимптотической логарифмической амплитудной характеристики системы.

3) При пересечении с каждой вертикальной прямой наклон асимптотической характеристики изменяется на 20 дБ/дек, если соответствующая сопрягающая частота относится к форсирующему звену первого порядка, на —20 дБ/дек, если она относится к апериодическому звену, на 40 дБ/дек, если она относится к форсирующему звену второго порядка, и на —40 дБ/дек, если она относится к колебательному звену.

Чтобы избежать систематических ошибок при построении асимптотической характеристики, целесообразно вычислять ее ординаты $L(\omega_r)$ в точках излома по очевидной формуле

$$L(\omega_r) = L(\omega_{r-1}) + 20m_r [\lg \omega_r - \lg \omega_{r-1}], \quad (4.8.1)$$

где $20m_r$ — наклон асимптотической характеристики в интервале частот (ω_{r-1}, ω_r) .

Если желательно построить точную логарифмическую амплитудную частотную характеристику соединения, то следует добавить к полученной асимптотической характеристике соответствующие поправки. Эти поправки могут быть определены или по графикам логарифмических частотных характеристик апериодического и колебательного звеньев (рис. 2.6.3 и 2.6.5), или с помощью специальной линейки [65].

Фазовые частотные характеристики последовательного соединения на основании формулы (4.6.4) получаются суммированием соответствующих ординат фазовых частотных характеристик входящих в соединение звеньев.

Для построения логарифмических частотных характеристик систем, полученных в результате замыкания систем с найденными характеристиками отрицательными жесткими обратными связями, и параллельных соединений систем используются специальные номограммы (приложение 2). На этих номограммах по оси абсцисс откладываются значения фазовой характеристики разомкнутой системы ψ в градусах, а по оси ординат — значения ее логарифмической амплитудной характеристики L в децибелах. В этой

системе координат построены два семейства кривых: семейство кривых постоянного усиления амплитуды замкнутой системы и семейство кривых постоянного сдвига фаз замкнутой системы. Для вывода уравнений этих семейств кривых достаточно положить в формуле (4.6.8) $\Phi_1(i\omega) = Ae^{i\psi}$ и приравнять постоянным величинам модуль и аргумент полученного выражения. В результате получим уравнение первого семейства кривых (т. е. кривых постоянного усиления амплуды замкнутой системы):

$$\frac{A}{\sqrt{1 + 2Ak_1 \cos \psi + A^2}} = \text{const}, \quad A = 10^{\frac{L}{20}}, \quad (4.8.2)$$

и уравнение второго семейства кривых (т. е. кривых постоянного сдвига фаз замкнутой системы):

$$\frac{\sin \psi}{A + \cos \psi} = \text{const}. \quad (4.8.3)$$

Из номограммы видно, что при значениях амплитудной частотной характеристики разомкнутой системы, меньших -20 дБ, амплитудная и фазовая частотные характеристики замкнутой системы практически совпадают с соответствующими характеристиками разомкнутой системы, а при ее значениях, больших 20 дБ, амплитудная частотная характеристика замкнутой системы не превышает по модулю 1 дБ, фазовая же характеристика не превышает $\pm 5^\circ$, т. е. практически они совпадают с соответствующими осями абсцисс. Поэтому номограмму замыкания практически необходимо использовать только в интервалах частот, в которых $|L(\omega)| < 20$ дБ.

Пример 4.8.1. В качестве примера рассмотрим построение логарифмических частотных характеристик системы, схема которой представлена на рис. 4.8.2. При помощи структурного преобразования выписим звено из

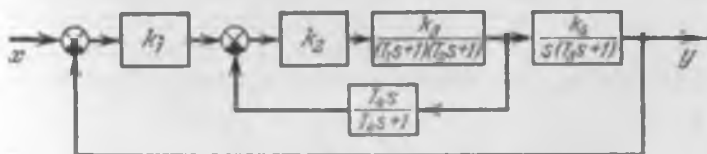


Рис. 4.8.2.

цепи обратной связи и таким образом заменяем гибкую обратную связь жесткой (пример 4.7.1). Преобразованная схема представлена на рис. 4.8.3.

Построим частотные характеристики системы, обведенной на рис. 4.8.3 пунктирной линией. Общий коэффициент усиления этой системы $k = k_2 k_3 T_4 = 4$. Определяем сопрягающие частоты в порядке их возрастания: $\omega_1 = 1/T_2 = 20$, $\omega_2 = 1/T_4 = 50$, $\omega_3 = 1/T_1 = 125$. При частоте $\omega = 1$ откладываем

ординату, равную $20 \lg k = 12,04$ дБ, и через полученную точку A (рис. 4.8.4) проводим асимптоту низких частот с наклоном 20 дБ/дек. В точке B наклон становится равным 0, в точке C — 20 дБ/дек и, наконец, в точке D — 40 дБ/дек. Фазовая частотная характеристика этой системы получается путем суммирования фазовых характеристик элементарных звеньев.

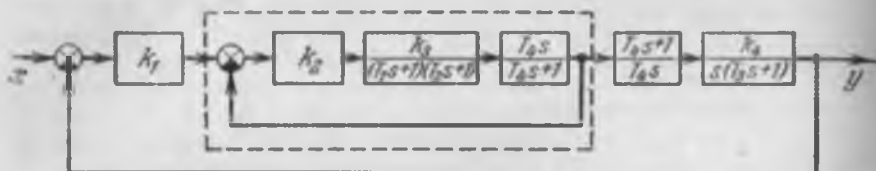


Рис. 4.8.3.

С помощью номограммы (приложение 2) находим логарифмические частотные характеристики системы, полученной путем замыкания жесткой отрицательной обратной связью системы, обведенной на рис. 4.8.3 пунктирным прямоугольником. Логарифмические частотные характеристики этой замкнутой системы отмечены на рис. 4.8.4 индексом «з».

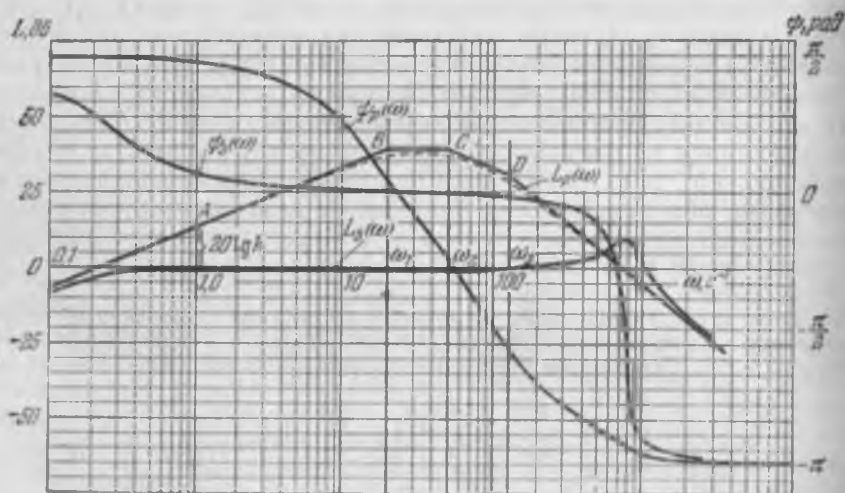


Рис. 4.8.4.

Совершенно так же строится логарифмические частотные характеристики второй последовательной цепи элементарных звеньев, входящих в состав исследуемой системы. Эти логарифмические частотные характеристики показаны на рис. 4.8.5.

Наконец, логарифмические частотные характеристики всей исследуемой разомкнутой системы, отмеченные на рис. 4.8.5 индексом 0, получаются суммированием найденных характеристик второй последовательной цепи и соответствующих характеристик замкнутой системы, изображенных на рис. 4.8.4.

Изложенный метод позволяет быстро строить логарифмические частотные характеристики любых стационарных линейных систем, в том числе и очень сложных. При этом, как мы увидим в главе 6, поправки к асимптотическим характеристикам практически необходимо вносить только в тех интервалах частот, в которых фазовая характеристика разомкнутой системы близка к $\pm\pi$.

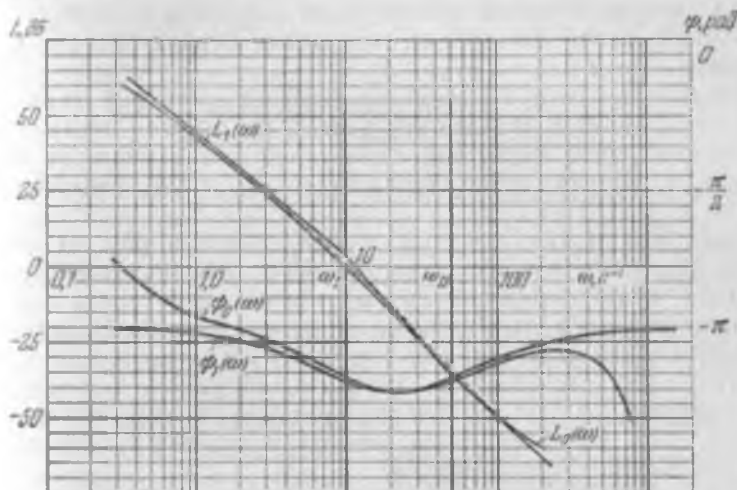


Рис. 4.8.5.

Впрочем, существует большой класс стационарных линейных систем, для которых можно обойтись без точного построения фазовых характеристик. А именно для систем, передаточные функции которых не имеют ни нулей, ни полюсов в правой полуплоскости комплексной переменной s , называемых обычно *минимально-фазовыми* *), фазовая характеристика полностью определяется амплитудной характеристикой. Эта зависимость фазовой характеристики минимально-фазовой системы от амплитудной может быть выражена различными уравнениями. Мы приведем здесь без вывода следующую формулу, справедливую для любой минимально-фазовой системы:

$$\Psi(\omega) = \frac{\pi}{40} \frac{dL(\omega)}{d \lg \omega} + \frac{1}{20\pi} \int_0^{\infty} \left[\frac{dL(\mu)}{d \lg \mu} - \frac{dL(\omega)}{d \lg \omega} \right] \ln \frac{\mu + \omega}{|\mu - \omega|} \frac{d\mu}{\mu}, \quad (4.8.4)$$

где фазовая характеристика $\Psi(\omega)$ выражена в радианах, а амплитудная характеристика $L(\omega)$ в децибелах.

*) Это название объясняется тем, что такие системы имеют минимальное отставание по фазе при любом значении частоты при данном значении амплитудной характеристики. Подробнее об этом см. [8].

Формула (4.8.4) показывает, что фазовая характеристика минимально-фазовой системы при любом значении частоты ω выражается через наклон логарифмической амплитудной частотной характеристики $dL(\omega)/d \lg \omega$. Из этого можно сделать некоторые интересные выводы.

1) Если в некотором достаточно большом интервале частот наклон логарифмической амплитудной частотной характеристики постоянен и равен $20n$ дБ/дек, то фазовая характеристика в этом интервале частот близка к $n\pi/2$.

2) Наиболее резкое изменение фазовой частотной характеристики происходит в области тех частот, где наклон логарифмической амплитудной частотной характеристики изменяется интенсивно, т. е. вблизи сопрягающих частот $\omega_i = 1/T_i$ элементарных звеньев.

Класс минимально-фазовых систем очень широк. В частности, он включает, как нетрудно убедиться, апериодическое, колебательное и форсирующие звенья первого и второго порядка с положительными постоянными времени T и коэффициентами колебательности ζ . Поэтому во многих случаях можно выполнить большую часть расчетной работы по одной только логарифмической амплитудной частотной характеристике разомкнутой системы, которая, как было показано, строится исключительно просто даже для очень сложных систем.

ДИСКРЕТНЫЕ ЛИНЕЙНЫЕ СИСТЕМЫ

§ 5.1. Дискретное (импульсное) управление.
Способы модуляции импульсов

В § 1.6 было показано, что входные возмущения, в частности, сигналы, несущие информацию о задачах управления, могут действовать в автоматических системах непрерывно или дискретно. Там же были приведены примеры непрерывных и дискретных систем.

Изложенная в предыдущих главах, за исключением §§ 2.5 и 4.4, общая теория линейных систем применима к любым линейным системам, как непрерывным, так и дискретным. Хотя мы и рассматривали входное возмущение $x(t)$ как функцию непрерывно изменяющегося времени, однако нигде не делали предположения, что это входное возмущение воспринимается системой непрерывно, в течение всего времени работы системы. Поэтому все изученные в предыдущих главах характеристики линейных систем и методы их исследования, кроме методов, изложенных в §§ 2.5 и 4.4, относятся в равной мере как к непрерывным, так и к дискретным линейным системам. И лишь теория, изложенная в §§ 2.5 и 4.4, в которых изучались непрерывные линейные системы, описываемые дифференциальными уравнениями, неприменима к дискретным системам. Однако дискретные линейные системы, как частный вид линейных систем, обладают некоторыми специфическими свойствами, которые целесообразно изучить и учитывать при расчете и проектировании дискретных систем.

Для дискретного ввода входных сигналов в систему необходимо изменять некоторые параметры входных импульсов в зависимости от значения входных сигналов. Изменение какого-либо параметра импульсов в зависимости от входного сигнала называется *модуляцией* импульсов входным сигналом. Устройство, формирующее последовательность импульсов, зависящую от входного сигнала, называется *импульсным элементом*.

Обычно форма импульсов сохраняется при модуляции неизменной. При этом в принципе возможно изменять в зависимости от значений входного сигнала один из трех параметров импульса: амплитуду, длительность или момент начала действия импульса. В соответствии с этим различают три вида модуляции импульсов:

Формула (4.8.4) показывает, что фазовая характеристика минимально-фазовой системы при любом значении частоты ω выражается через наклон логарифмической амплитудной частотной характеристики $dL(\omega)/d \lg \omega$. Из этого можно сделать некоторые интересные выводы.

1) Если в некотором достаточно большом интервале частот наклон логарифмической амплитудной частотной характеристики постоянен и равен $20n$ дБ/дек, то фазовая характеристика в этом интервале частот близка к $n\pi/2$.

2) Наиболее резкое изменение фазовой частотной характеристики происходит в области тех частот, где наклон логарифмической амплитудной частотной характеристики изменяется интенсивно, т. е. вблизи сопрягающих частот $\omega_i = 1/T_i$ элементарных звеньев.

Класс минимально-фазовых систем очень широк. В частности, он включает, как нетрудно убедиться, аperiodическое, колебательное и форсирующие звенья первого и второго порядка с положительными постоянными времени T и коэффициентами колебательности ζ . Поэтому во многих случаях можно выполнить большую часть расчетной работы по одной только логарифмической амплитудной частотной характеристике разомкнутой системы, которая, как было показано, строится исключительно просто даже для очень сложных систем.

ДИСКРЕТНЫЕ ЛИНЕЙНЫЕ СИСТЕМЫ

§ 5.1. Дискретное (импульсное) управление.
Способы модуляции импульсов

В § 1.6 было показано, что входные возмущения, в частности, сигналы, несущие информацию о задачах управления, могут действовать в автоматических системах непрерывно или дискретно. Там же были приведены примеры непрерывных и дискретных систем.

Изложенная в предыдущих главах, за исключением §§ 2.5 и 4.4, общая теория линейных систем применима к любым линейным системам, как непрерывным, так и дискретным. Хотя мы и рассматривали входное возмущение $x(t)$ как функцию непрерывно изменяющегося времени, однако нигде не делали предположения, что это входное возмущение воспринимается системой непрерывно, в течение всего времени работы системы. Поэтому все изученные в предыдущих главах характеристики линейных систем и методы их исследования, кроме методов, изложенных в §§ 2.5 и 4.4, относятся в равной мере как к непрерывным, так и к дискретным линейным системам. И лишь теория, изложенная в §§ 2.5 и 4.4, в которых изучались непрерывные линейные системы, описываемые дифференциальными уравнениями, неприменима к дискретным системам. Однако дискретные линейные системы, как частный вид линейных систем, обладают некоторыми специфическими свойствами, которые целесообразно изучить и учитывать при расчете и проектировании дискретных систем.

Для дискретного ввода входных сигналов в систему необходимо изменять некоторые параметры входных импульсов в зависимости от значения входных сигналов. Изменение какого-либо параметра импульсов в зависимости от входного сигнала называется *модуляцией* импульсов входным сигналом. Устройство, формирующее последовательность импульсов, зависящую от входного сигнала, называется *импульсным элементом*.

Обычно форма импульсов сохраняется при модуляции неизменной. При этом в принципе возможно изменять в зависимости от значений входного сигнала один из трех параметров импульса: амплитуду, длительность или момент начала действия импульса. В соответствии с этим различают три вида модуляции импульсов:

1) амплитудно-импульсная модуляция (АИМ), при которой амплитуда импульсов a зависит от значения входного сигнала в момент начала действия импульса t_k (рис. 5.1.1);

2) широтно-импульсная модуляция (ШИМ), при которой длительность импульса T_k зависит от значения входного сигнала в момент начала действия импульса (рис. 5.1.2);

3) временная импульсная модуляция (ВИМ), при которой временной сдвиг T_c (запаздывание) импульса зависит от значения входного сигнала в определенный момент времени (рис. 5.1.3).

При амплитудно-импульсной и широтно-импульсной модуляции модулирующий сигнал изменяет площадь (т. е. интенсивность) импульсов. При временной импульсной модуляции площадь импульса остается постоянной.

Все рассмотренные типы модуляции могут быть осуществле-

ны так, чтобы модулируемый параметр импульсов изменялся в зависимости от значений модулирующего сигнала в дискретные моменты времени. Амплитудная модуляция может быть осуществлена



Рис. 5.1.1

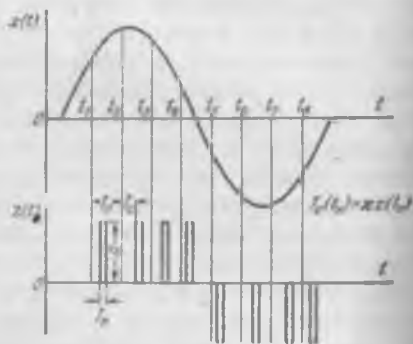


Рис. 5.1.2

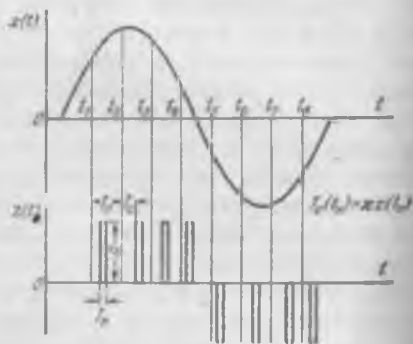


Рис. 5.1.3

и таким образом, чтобы ордината каждого импульса изменялась непрерывно в зависимости от входного сигнала в течение времени действия импульса. Однако дискретные системы с такой формой амплитудной модуляции импульсов не обладают существенными

специфическими свойствами, облегчающими их исследование, вследствие чего их следует изучать общими методами предыдущей главы. Излагаемая в этой главе теория к ним неприменима.

Зависимость модулируемого параметра вырабатываемых импульсным элементом импульсов от соответствующих дискретных значений входной переменной называется *характеристикой* импульсного элемента. Характеристика импульсного элемента может быть линейной или нелинейной. Импульсный элемент с линейной характеристикой является линейным элементом, а импульсный элемент с нелинейной характеристикой — нелинейным.

Импульсные элементы различаются также по форме и характеру модуляции импульсов, по частоте импульсов и их относительной длительности.

Обычно импульсные элементы работают периодически, вырабатывая по одному импульсу за каждый период. Период следования импульсов T_n называется *периодом повторения* импульсов или *тактом* дискретной системы. Величина $\omega_n = 2\pi/T_n$ представляет собой *частоту* повторения импульсов. Отношение длительности одного импульса (средней в случае широтной модуляции) T_n к периоду повторения импульсов $\gamma = T_n/T_n$ представляет собой *относительную длительность* импульсов. Величина $1 - \gamma$ называется *скважностью* импульсного элемента.

§ 5.2. Характеристики дискретных линейных систем

Так как входная переменная $x(t)$ действует на дискретную систему только в определенные моменты времени t_k ($k = 0, \pm 1, \pm 2, \dots$), то ее выходная переменная является функцией времени и значений $x(t_k)$ ($k = 0, \pm 1, \pm 2, \dots$) входной переменной.

Обозначим реакцию дискретной линейной системы на кратковременное входное возмущение, равное единице и действующее только в течение времени действия k -го импульса, через $g_k(t)$. Тогда ее реакция на кратковременное возмущение, равное $x(t_k)$ и действующее только в течение времени действия k -го импульса, будет на основании принципа суперпозиции равна $g_k(t) x(t_k)$. Реакция дискретной линейной системы на всю последовательность импульсов, модулированных входным возмущением $x(t)$, в силу принципа суперпозиции определится формулой

$$y(t) = \sum_{k=-\infty}^{\infty} g_k(t) x(t_k). \quad (5.2.1)$$

Это и есть основная формула, выражающая зависимость выходной переменной дискретной линейной системы от входного сигнала.

Функции $g_k(t)$ полностью характеризуют дискретную линейную систему, так как, зная эти функции, можно вычислить реак-

цию дискретной линейной системы на любое входное возмущение $x(t)$. Функции $g_k(t)$ определяют долю, или удельный вес, значений входной переменной, действующих в различные моменты времени t_k , в формировании выходной переменной системы в любой момент времени t . Вследствие этого функции $g_k(t)$ называются *весовыми коэффициентами* дискретной линейной системы.

Любая физически возможная система может реагировать в данный момент времени t только на возмущения, действующие на нее в предшествующие моменты времени. Поэтому для любой физически возможной дискретной системы

$$g_k(t) = 0 \text{ при } t < t_k \quad (k = 0, \pm 1, \pm 2, \dots). \quad (5.2.2)$$

Для физически возможной дискретной линейной системы, находящейся в покое до момента t_0 , формула (5.2.1) может быть переписана в виде

$$y(t) = \sum_{t_0 \leq t_k \leq t} g_k(t) x(t_k). \quad (5.2.3)$$

где неравенство под знаком суммы показывает, что суммирование распространяется только на моменты действия импульсов t_k , заключенные в интервале $[t_0, t]$.

Для определения весовой функции $g(t, \tau)$ дискретной линейной системы достаточно найти ее реакцию на единичный импульс, действующий в момент τ . Полагая в формуле (5.2.1) $x(t) = \delta(t - \tau)$, получим следующую формулу для весовой функции $g(t, \tau)$ дискретной линейной системы:

$$g(t, \tau) = \sum_{k=-\infty}^{\infty} g_k(t) \delta(t_k - \tau). \quad (5.2.4)$$

Таким образом, весовая функция любой дискретной линейной системы представляет собой линейную комбинацию δ -функций. Наоборот, всякая линейная система, весовая функция которой является линейной комбинацией δ -функций, дискретна. Действительно, подставляя выражение (5.2.4) в формулу (2.2.3) и принимая во внимание правило интегрирования δ -функций, получим формулу (5.2.1). Из этой формулы следует, что входное возмущение $x(t)$ действует на систему только в моменты времени t_k , т. е. что система дискретна.

Определив весовую функцию дискретной линейной системы формулой (5.2.4), можно найти по формулам § 2.3 ее характеристику реакции на показательное возмущение и частотную характеристику. Полагая в формулах § 4.2 весовые функции равными линейным комбинациям δ -функций, получим формулы для весовых коэффициентов различных соединений дискретных линейных систем. Наконец, формулы § 4.3 определяют системы, сопряжен-

ные с дискретными линейными системами. При этом выведенные в § 4.3 правила моделирования сопряженных систем дадут возможность определять значения всех весовых коэффициентов системы в определенный момент времени t однократным моделированием сопряженной системы.

Рассмотрим дискретную линейную систему, представляющую собой последовательное соединение импульсного элемента и непрерывной линейной системы с весовой функцией $g_1(t, \tau)$ (рис. 5.2.1). Обозначим функцию, описывающую форму импульсов, вырабатываемых импульсным элементом, через $\eta(t)$ (рис. 5.2.2). Тогда

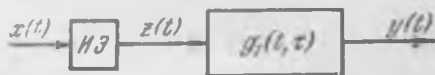


Рис. 5.2.1

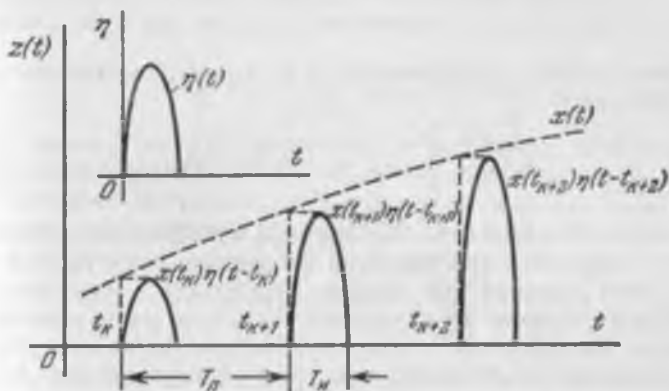


Рис. 5.2.2

в случае амплитудной модуляции выходной сигнал линейного импульсного элемента $z(t)$ выразится формулой *)

$$z(t) = \sum_{k=-\infty}^{\infty} x(t_k) \eta(t - t_k). \quad (5.2.5)$$

Эта функция является входной переменной непрерывной линейной системы с весовой функцией $g_1(t, \tau)$. Следовательно, на основании общей формулы (2.2.3) выходная переменная рассматриваемой дискретной системы выражается формулой

$$y(t) = \int_{-\infty}^{\infty} g_1(t, \tau) z(\tau) d\tau = \sum_{k=-\infty}^{\infty} x(t_k) \int_{-\infty}^{\infty} g_1(t, \tau) \eta(\tau - t_k) d\tau. \quad (5.2.6)$$

*) Очевидно, что без потери общности функцию $\eta(t)$ можно выбрать так, чтобы при входном возмущении $x(t)$, тождественно равном единице, импульсы, вырабатываемые импульсным элементом, были равны $\eta(t - t_k)$ ($k = 0, \pm 1, \pm 2, \dots$).

Так как функция $\eta(\tau - t_k)$ отлична от нуля только в интервале $t_k \leq \tau \leq t_k + T_n$, где T_n — длительность импульса, то формулу (5.2.6) можно переписать в виде

$$y(t) = \sum_{k=-\infty}^{\infty} x(t_k) \int_{t_k}^{t_k+T_n} g_1(t, \tau) \eta(\tau - t_k) d\tau. \quad (5.2.7)$$

Сравнивая эту формулу с (5.2.1), получаем следующую формулу для весовых коэффициентов рассматриваемой дискретной линейной системы:

$$g_k(t) = \int_{t_k}^{t_k+T_n} g_1(t, \tau) \eta(\tau - t_k) d\tau. \quad (5.2.8)$$

Производя замену переменных $\sigma = \tau - t_k$, приведем формулу (5.2.8) к виду

$$g_k(t) = \int_0^{T_n} g_1(t, t_k + \sigma) \eta(\sigma) d\sigma. \quad (5.2.9)$$

Эта формула показывает, что весовые коэффициенты последовательного соединения импульсного элемента и непрерывной линейной системы зависят от формы импульсов, вырабатываемых импульсным элементом, и весовой функции непрерывной части системы.

Из формулы (5.2.9) видно, что практически для любой физической возможной дискретной линейной системы начальные значения весовых коэффициентов $g_k(t)$ равны нулю: $g_k(t_k) = 0$ ($k = 0, 1, 2, \dots$), так как при $t = t_k$ подынтегральная функция тождественно равна нулю в интервале интегрирования. Начальные значения весовых коэффициентов $g_k(t_k)$ могут быть отличными от нуля только в случае, когда весовая функция непрерывной части системы $g_1(t, \tau)$ содержит линейную комбинацию δ -функции и ее производных: $\delta(t - \tau)$, $\delta'(t - \tau)$, $\delta''(t - \tau)$, $\dots$, или в теоретическом случае идеализированного импульсного элемента, вырабатывающего без запаздывания мгновенные импульсы: $\eta(\sigma) = \delta(\sigma)$, если при этом $g_1(t, t) \neq 0$.

В случае широтной модуляции мы будем для простоты считать, что импульсный элемент вырабатывает прямоугольные импульсы постоянной величины a , длительность которых пропорциональна значениям входного сигнала в соответствующие моменты времени. Тогда входной сигнал $z(t)$ непрерывной части системы будет равен a в интервалах времени $[t_k, t_k + \alpha x(t_k)]$ и нулю вне этих интервалов. Следовательно, применяя формулу (2.2.3), мы получим следующее выражение для выходной переменной рас-

смаатриваемой дискретной линейной системы:

$$y(t) = a \sum_{k=-\infty}^{\infty} \int_{t_k}^{t_k + \alpha x(t_k)} g_1(t, \tau) d\tau. \quad (5.2.10)$$

Очевидно, что система, у которой зависимость между входной и выходной переменными определяется формулой (5.2.10), нелинейна. Однако при малой длительности импульсов ее можно приближенно рассматривать как линейную. Действительно, при достаточно малой длительности каждого импульса весовую функцию $g_1(t, \tau)$ можно считать в каждом интеграле в (5.2.10) практически постоянной и равной ее значению $g_1(t, t_k)$ в начале интервала интегрирования. Тогда, вынося ее за знак интеграла, можем выполнить интегрирование и переписать формулу (5.2.10) в виде

$$y(t) \approx \alpha x \sum_{k=-\infty}^{\infty} g_1(t, t_k) x(t_k). \quad (5.2.11)$$

Отсюда видно, что при широтной модуляции кратковременных прямоугольных импульсов последовательное соединение импульсного элемента и непрерывной системы с весовой функцией $g_1(t, \tau)$ можно считать линейной дискретной системой, весовые коэффициенты которой определяются формулой

$$g_k(t) = \alpha x g_1(t, t_k). \quad (5.2.12)$$

Величина αx , представляющая собой произведение амплитуды импульса на его длительность при единичном входном сигнале, определяет интенсивность импульсов.

Аналогично можно прийти к заключению, что при временной модуляции импульсов произвольной формы последовательное соединение импульсного элемента и непрерывной линейной системы лишь приближенно можно рассматривать как линейную дискретную систему, если максимальный возможный временной сдвиг импульса достаточно мал, чтобы весовую функцию непрерывной части системы $g_1(t, \tau)$ можно было считать приблизительно линейной функцией τ в диапазоне возможных значений временного сдвига.

В некоторых случаях выходная переменная дискретной линейной системы интересует нас лишь в определенные моменты времени t_l ($l = 0, \pm 1, \pm 2, \dots$) или система дает выходную переменную только в определенные моменты времени t_l . Полагая в формуле (5.2.1) $t = t_l$, найдем значения выходной переменной системы в интересующие нас моменты времени t_l ($l = 0, \pm 1, \pm 2, \dots$):

$$y(t_l) = \sum_{k=-\infty}^{\infty} g_k(t_l) x(t_k). \quad (5.2.13)$$

Полагая для краткости

$$x_k = x(t_k), \quad y_l = y(t_l), \quad g_k(t_l) = g_{lk}, \quad (5.2.14)$$

можем переписать формулу (5.2.13) окончательно в виде

$$y_l = \sum_{k=-\infty}^{\infty} g_{lk} x_k \quad (l=0, \pm 1, \pm 2, \dots). \quad (5.2.15)$$

Аналогично в случае, когда последовательность моментов фиксации выходной переменной системы совпадает с последовательностью моментов действия импульсов $t_l' = t_l$ ($l = 0, \pm 1, \pm 2, \dots$), формула (5.2.3) при $t = t_l$ даст следующее выражение значений выходной переменной физически возможной дискретной линейной системы:

$$y_l = \sum_{k=0}^l g_{lk} x_k \quad (l=0, 1, 2, \dots) \quad (5.2.16)$$

Формула (5.2.16) и аналогичные формулы для дискретных систем с несколькими входами и выходами описывают, в частности, работу цифровой математической машины в случае, когда результаты вычислений линейно зависят от исходных данных, т. е. цифровой машины с линейной программой. В этом случае исходные данные для вычислений, вводимые в каждом шаге вычислений, являются входными сигналами, а результаты вычислений — выходными переменными. Весовые коэффициенты g_{lk} определяют программу вычислений.

Формула (5.2.16) на простейшем примере показывает, что динамические характеристики цифровой (так же, как и непрерывной) математической машины полностью определяются программой вычислений, т. е. решаемой задачей.

Легко сообразить, что при достаточно большой частоте повторения импульсов всякую дискретную систему можно приближенно рассматривать как непрерывную. Действительно, реакция любой системы на кратковременное единичное возмущение, действующее в течение одного периода повторения импульсов T_n , при бесконечно малом T_n является бесконечно малой порядка T_n . Поэтому функцию $g_k(t)$ при малом $T_n = \Delta t$ можно выразить формулой

$$g_k(t) \approx g(t, t_k) T_n = g(t, t_k) \Delta t, \quad (5.2.17)$$

где $g(t, t_k)$ — некоторая функция, не зависящая от T_n . Так, например, в случае последовательного соединения импульсного элемента и непрерывной линейной системы с весовой функцией $g_1(t, \tau)$ мы можем выразить зависимость импульсов данной формы от периода повторения импульсов, приняв функцию $\eta(t)$ в виде

$$\eta(t) = h\left(\frac{t}{T_n}\right), \quad (5.2.18)$$

где $h(\xi)$ — функция, отличная от нуля только в интервале $0 \leq \xi \leq \gamma = T_n/T_n$. Тогда формула (5.2.9) даст

$$g_k(t) = \int_0^{T_n} g_1(t, t_k + \sigma) h\left(\frac{\sigma}{T_n}\right) d\sigma = T_n \int_0^{\gamma} g_1(t, t_k + \xi T_n) h(\xi) d\xi.$$

При достаточно малом периоде повторения импульсов T_n весовая функция g_1 в полученном интеграле будет практически постоянной в интервале интегрирования, близкой к своему значению $g_1(t, t_k + 0)$. Вынося ее за знак интеграла и полагая

$$g(t, t_k) = g_1(t, t_k + 0) \int_0^{\gamma} h(\xi) d\xi, \quad (5.2.19)$$

мы и получим формулу (5.2.17).

Подставляя выражение (5.2.17) в (5.2.1), получим

$$y(t) \approx \sum_{k=-\infty}^{\infty} g(t, t_k) x(t_k) \Delta t, \quad (5.2.20)$$

причем относительная ошибка этого равенства стремится к нулю при $\Delta t = T_n \rightarrow 0$. Но при $\Delta t \rightarrow 0$ сумма в (5.2.20) стремится к интегралу (2.2.3). Следовательно, при достаточно малом T_n дискретную линейную систему можно приближенно рассматривать как непрерывную линейную систему с весовой функцией $g(t, t_k) = g_k(t)/T_n$.

В частном случае последовательного соединения импульсного элемента и стационарной линейной системы, описываемой дифференциальным уравнением, условие приближительного постоянства весовой функции непрерывной части системы на протяжении периода повторения импульсов T_n , как показывает формула (4.4.50), равноценно условию близости к единице всех величин $e^{v_1 T_n}, \dots, e^{v_n T_n}$, где $v_1, \dots, v_n$ — корни характеристического уравнения непрерывной части системы. Но, разложив числитель и знаменатель выражения (2.5.12) на множители, мы представим передаточную функцию непрерывной части системы в виде (2.6.1). Таким образом, непрерывная часть системы может быть представлена в данном случае в виде последовательного соединения интегрирующих или дифференцирующих, апериодических, колебательных и форсирующих звеньев. При этом корни характеристического уравнения $v_1, \dots, v_n$ будут обратно пропорциональны постоянным времени апериодических и колебательных звеньев, входящих в состав непрерывной части системы. Следовательно, условие малости величин $v_1 T_n, \dots, v_n T_n$ с физической точки зрения является условием малости периода повторения импульсов T_n по сравнению с постоянными времени непрерывной части

системы. Таким образом, последовательное соединение импульсного элемента и непрерывной стационарной линейной системы можно считать непрерывной стационарной линейной системой, если период повторения импульсов достаточно мал по сравнению с постоянными времени непрерывной части системы.

Пример 5.2.1. Найти весовые коэффициенты импульсного запаздывающего звена, для которого $y(t_l) = x(t_l - T)$.

Импульсное запаздывающее звено можно представить в виде последовательного соединения импульсного элемента и непрерывного запаздывающего звена. Подставляя выражение $g_1(t, \tau) = \delta(t - \tau - T)$ весовой функции запаздывающего звена в формулу (5.2.9), получим

$$g_k(t) = \int_0^T \delta(t - t_k - \sigma - T) \eta(\sigma) d\sigma = \eta(t - t_k - T). \quad (5.2.21)$$

Таким образом, импульсное запаздывающее звено повторяет входные импульсы с запаздыванием T .

В частном случае единичных прямоугольных импульсов $\eta(\sigma) = 1$ при $0 \leq \sigma \leq T_n$. Если при этом запаздывание T равно целому числу периодов повторения $T = nT_n$, то, полагая в (5.2.21) $t = t_l = lT_n$, $t_k = kT_n$, получим

$$g_{lk} = g_k(t_l) = \begin{cases} 1 & \text{при } l = k + n, \\ 0 & \text{при } l \neq k + n. \end{cases} \quad (5.2.22)$$

Пример 5.2.2. Найти весовые коэффициенты дискретной линейной системы, представляющей собой последовательное соединение импульсного элемента, вырабатывающего единичные прямоугольные импульсы длительности T_n с периодом повторения T_n , и апериодического звена.

Подставляя в формулу (5.2.9) выражение (4.4.33) весовой функции апериодического звена и принимая во внимание, что $\eta(\sigma) = 1$ при $0 \leq \sigma \leq T_n$, получим:

$$g_k(t) = \left. \begin{aligned} & \frac{k}{T} \int_0^{t-t_k} e^{-\frac{t-t_k-\sigma}{T}} d\sigma = k(1 - e^{-\frac{t-t_k}{T}}) \\ & \text{при } t_k < t < t_k + T_n, \\ & \frac{k}{T} \int_0^{T_n} e^{-\frac{t-t_k-\sigma}{T}} d\sigma = ke^{-\frac{t-t_k}{T}} (e^{\frac{T_n}{T}} - 1) \\ & \text{при } t > t_k + T_n. \end{aligned} \right\} \quad (5.2.23)$$

Пример 5.2.3. Найти весовые коэффициенты последовательного соединения импульсного элемента, вырабатывающего единичные прямоугольные импульсы, и непрерывной системы, описываемой дифференциальным уравнением

$$\ddot{y} + 2a\dot{y} + b^2y = ke^{\gamma t}(x + bx),$$

где a, b, k, γ — некоторые постоянные, $0 < a < b$. Весовая функция этой системы была найдена в примере 4.4.2 и выражается формулой (4.4.37).

Подставляя в (5.2.9) выражение (4.4.37) и учитывая, что $\gamma(\sigma) = 1$ при $0 < \sigma \leq T_n$, получим

$$\begin{aligned} g_h(t) &= \frac{k}{2i\omega_0} \int_0^{T_n} e^{-at + a_1(t_h + \sigma)} [(b - a_1 + i\omega_0) e^{i\omega_0(t - t_h - \sigma)} - \\ &\quad - (b - a_1 - i\omega_0) e^{-i\omega_0(t - t_h - \sigma)}] d\sigma = \\ &= \frac{k}{2i\omega_0} e^{-at + a_1 t_h} \left\{ e^{a_1 T_n} \left[\frac{b - a_1 + i\omega_0}{a_1 - i\omega_0} e^{i\omega_0(t - t_h - T_n)} - \right. \right. \\ &\quad \left. \left. - \frac{b - a_1 - i\omega_0}{a_1 + i\omega_0} e^{-i\omega_0(t - t_h - T_n)} \right] - \frac{b - a_1 + i\omega_0}{a_1 - i\omega_0} e^{i\omega_0(t - t_h)} + \right. \\ &\quad \left. + \frac{b - a_1 - i\omega_0}{a_1 + i\omega_0} e^{-i\omega_0(t - t_h)} \right\} \quad \text{при } t > t_h + T_n. \quad (5.2.2) \end{aligned}$$

где $\omega_0 = \sqrt{b^2 - a^2}$, $a_1 = a + \gamma$. Предоставляем читателю самостоятельно привести найденные выражения весовых коэффициентов к действительной форме, а также найти их выражения при $t_h < t < t_h + T_n$.

§ 5.3. Стационарные дискретные линейные системы

Строго говоря, никакая дискретная система не может быть стационарной в смысле данного в § 1.6 определения, так как при любом сдвиге входного возмущения во времени на величину, не равную целому числу периодов повторения импульсов (кривая 2 на рис. 5.3.1), значения входного возмущения $x(t_h)$, действующие на систему, изменяются (точки на кривой 2) и вследствие этого выходной сигнал системы не только сдвигается во времени, но и изменяет свою форму. Однако если пренебречь интервалами времени, меньшими периода повторения импульсов T_n , и рассматривать только сдвиги во времени, кратные периоду повторения импульсов, то можно говорить о стационарных или нестационарных дискретных системах. В соответствии с этим мы дадим следующее определение стационарной дискретной системы: дискретная система называется *стационарной*, если при сдвиге во времени входного возмущения без изменения его формы на интервал времени, кратный периоду повторения импульсов, выходная переменная сдвигается во времени на такой же интервал без изменения своей формы. Из этого определения следует, что дискретная система может быть стационарной только в том

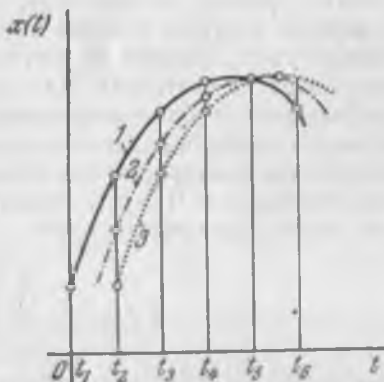


Рис. 5.3.1

случае, когда действующие на систему значения входного возмущения следуют друг за другом через равные промежутки времени. Если значения входной переменной вводятся в систему через неравные интервалы времени, то дискретная система нестационарна.

Весовой коэффициент $g_h(t)$ линейной дискретной системы по определению есть ее реакция на кратковременное возмущение (т. е. возмущение длительности, меньшей периода повторения импульсов T_n), равное единице в момент $t_h = hT_n$. Если линейная дискретная система стационарна, то при сдвиге кратковременного единичного возмущения во времени на интервал kT_n ее реакция сдвинется во времени на тот же интервал kT_n , не изменяя формы, т. е. будет равна $g_h(t - kT_n)$. С другой стороны, реакция системы на единичное возмущение, действующее в момент $t_{h+k} = (h+k)T_n$, будет равна ее весовому коэффициенту $g_{h+k}(t)$. Следовательно, для стационарной дискретной линейной системы при любых k и h

$$g_{h+k}(t) = g_h(t - kT_n). \quad (5.3.1)$$

Полагая здесь $h = 0$, получим при любом k

$$g_k(t) = g_0(t - kT_n) \quad (k = 0, \pm 1, \pm 2, \dots). \quad (5.3.2)$$

Таким образом, весовые коэффициенты стационарной дискретной линейной системы представляют собой одну и ту же функцию, сдвинутую во времени на интервалы, кратные периоду повторения импульсов. В частности, для дискретной линейной системы, представляющей собой последовательное соединение импульсного элемента, вырабатывающего последовательность равноотстоящих импульсов, и непрерывной стационарной линейной системы с весовой функцией $w(t - \tau)$, формула (5.2.9) дает следующее выражение весовых коэффициентов:

$$g_k(t) = g_0(t - kT_n) = \int_0^{T_n} w(t - kT_n - \sigma) \eta(\sigma) d\sigma. \quad (5.3.3)$$

Полагая в (5.3.2), в частности, $t = t_l = lT_n$, получим

$$g_{lk} = g_k(t_l) = g_0((l - k)T_n) \quad (k, l = 0, \pm 1, \pm 2, \dots) \quad (5.3.4)$$

Таким образом, весовые коэффициенты g_{lk} для стационарной дискретной линейной системы зависят только от разности индексов. Мы в дальнейшем будем обозначать их w_m ($m = 0, \pm 1, \pm 2, \dots$). Для физически возможных стационарных дискретных линейных систем $g_0(t) = 0$ при $t < 0$ и, следовательно,

$$w_m = 0 \text{ при } m < 0. \quad (5.3.5)$$

При этом на основании изложенного в предыдущем параграфе практически для всех реальных дискретных линейных систем $w_0 = 0$.

На основании формул (5.3.3) и (5.3.4) весовые коэффициенты w_m последовательного соединения импульсного элемента и непрерывной стационарной линейной системы определяются формулой

$$w_m = \int_0^{T_n} w(mT_n - \sigma) \eta(\sigma) d\sigma \quad (m = 0, \pm 1, \pm 2, \dots). \quad (5.3.6)$$

Формула (5.2.15), выражающая последовательность значений выходной переменной через последовательность значений входной переменной, для стационарной дискретной линейной системы принимает вид

$$y_l = \sum_{k=-\infty}^{\infty} w_{l-k} x_k = \sum_{m=-\infty}^{\infty} m_m x_{l-m}. \quad (5.3.7)$$

Рассмотрим теперь действие на стационарную дискретную линейную систему показательного возмущения $x(t) = e^{st}$. В этом случае

$$x_{l-m} = e^{s(l-m)T_n}$$

и формула (5.3.7) дает

$$y_l = \sum_{m=-\infty}^{\infty} w_m e^{s(l-m)T_n} = e^{slT_n} \sum_{m=-\infty}^{\infty} w_m e^{-smT_n}. \quad (5.3.8)$$

Таким образом, реакция стационарной дискретной линейной системы на показательное возмущение e^{st} в момент $t = lT_n$ равна значению этого возмущения в тот же момент, умноженному на функцию

$$\Phi(s) = \sum_{m=-\infty}^{\infty} w_m e^{-smT_n}, \quad (5.3.9)$$

зависящую только от параметра s . Функция $\Phi(s)$ является передаточной функцией стационарной дискретной линейной системы.

Полагая в (5.3.9) $s = i\omega$, находим частотную характеристику стационарной дискретной линейной системы:

$$\Phi(i\omega) = \sum_{m=-\infty}^{\infty} w_m e^{-imT_n\omega}. \quad (5.3.10)$$

Эта формула показывает, что частотная характеристика стационарной дискретной линейной системы является периодической функцией частоты ω с периодом $\Omega = 2\pi/T_n$. Формула (5.3.10) определяет частотную характеристику стационарной дискретной линейной системы ее рядом Фурье. Коэффициентами этого ряда

являются весовые коэффициенты системы. Следовательно, пользуясь известной формулой для коэффициентов ряда Фурье, мы выразим весовые коэффициенты стационарной дискретной линейной системы через ее частотную характеристику:

$$w_m = \frac{1}{\Omega} \int_0^{\Omega} \Phi_1(i\omega) e^{imT_n} d\omega = \frac{T_n}{2\pi} \int_0^{\frac{2\pi}{T_n}} \Phi(i\omega) e^{imT_n} d\omega. \quad (5.3.11)$$

В случае физически возможной стационарной дискретной линейной системы интегралы в (5.3.11) равны нулю при $m < 0$ и формула (5.3.9) принимает вид

$$\Phi(s) = \sum_{m=0}^{\infty} w_m e^{-smT_n}. \quad (5.3.12)$$

Формулы (5.3.9) и (5.3.12) показывают, что передаточные функции стационарных дискретных линейных систем являются функциями величины

$$z = e^{sT_n}. \quad (5.3.13)$$

Обозначая передаточную функцию физически возможной стационарной дискретной линейной системы, рассматриваемую как функцию параметра z , через $\Psi(z)$, можем переписать формулу (5.3.12) в виде

$$\Psi(z) = \sum_{m=0}^{\infty} w_m z^{-m}. \quad (5.3.14)$$

При этом, конечно, имеют место тождества

$$\Phi(s) \equiv \Psi(e^{sT_n}), \quad \Psi(z) \equiv \Phi\left(\frac{1}{T_n} \ln z\right). \quad (5.3.15)$$

Легко понять, что все формулы § 4.6 справедливы и для передаточных функций соединений стационарных дискретных линейных систем.

В некоторых случаях передаточную функцию стационарной дискретной линейной системы удобно рассматривать как функцию параметра

$$v = \frac{z-1}{z+1} = \frac{e^{sT_n}-1}{e^{sT_n}+1}. \quad (5.3.16)$$

Подставив выражение z из (5.3.16):

$$z = \frac{1+v}{1-v}, \quad (5.3.17)$$

в (5.3.15) и обозначив передаточную функцию, рассматриваемую как функция v , через $\Omega(v)$, получим следующие соотношения

между передаточными функциями Ω , Ψ и Φ :

$$\Omega(v) = \Psi\left(\frac{1+v}{1-v}\right) = \Phi\left(\frac{1}{T_n} \ln \frac{1+v}{1-v}\right). \quad (5.3.18)$$

Используя (5.3.16), получим еще соотношения

$$\Psi(z) = \Omega\left(\frac{z-1}{z+1}\right), \quad \Phi(s) = \Omega\left(\frac{e^{sT_n}-1}{e^{sT_n}+1}\right). \quad (5.3.19)$$

Пример 5.3.1. Найти передаточные функции $\Psi(z)$ и $\Omega(v)$ запаздывающего импульсного звена.

Весовые коэффициенты запаздывающего импульсного звена на основании (5.2.22) определяются формулой

$$w_m = g_k(t) = \begin{cases} 1 & \text{при } m = l - k = n, \\ 0 & \text{при } m = l - k \neq n. \end{cases}$$

Подставляя это выражение в формулу (5.3.14), получим

$$\Psi(z) = z^{-n}. \quad (5.3.20)$$

Подставляя сюда выражение (5.3.17) переменной z через v , найдем

$$\Omega(v) = \left(\frac{1-v}{1+v}\right)^n. \quad (5.3.21)$$

Пример 5.3.2. Найти передаточную функцию замкнутой дискретной системы (рис. 5.3.2), импульсный элемент которой вырабатывает прямоугольные импульсы длительности T_n с периодом повторения T_n .

Найдем передаточную функцию $\Psi_1(z)$ разомкнутой системы. Подставляя в формулу (5.3.14) выражение (5.2.23) для весовых коэффициентов разомкнутой импульсной системы и учитывая, что $l - k = m$, получим формулу для передаточной функции $\Psi_1(z)$:

$$\Psi_1(z) = k \left(e^{\frac{\gamma T_n}{T}} - 1 \right) \sum_{m=1}^{\infty} e^{-\frac{m T_n}{T}} z^{-m}$$

или, суммируя геометрическую прогрессию,

$$\Psi_1(z) = k z_1 \frac{z_1^{-\gamma} - 1}{z - z_1}, \quad z_1 = e^{-\frac{T_n}{T}}. \quad (5.3.22)$$

Подставляя выражение (5.3.22) в формулу (4.5.8), найдем передаточную функцию замкнутой системы:

$$\Psi(z) = \frac{\Psi_1(z)}{1 + \Psi_1(z)} = \frac{k z_1 (z_1^{-\gamma} - 1)}{z - z_1 (k + 1 - k z_1^{-\gamma})}. \quad (5.3.23)$$

Замена переменной z переменной v по формуле (5.3.17) в (5.3.22) дает для разомкнутой системы

$$\Omega_1(v) = k z_1 (z^{-\gamma} - 1) \frac{1 - v}{1 - z_1 + (1 + z_1)v}. \quad (5.3.24)$$

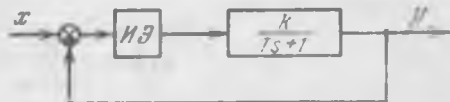


Рис. 5.3.2

Передаточная функция $\Omega(v)$ замкнутой системы определяется формулой

$$\Omega(v) = \frac{\Omega_1(v)}{1 + \Omega_1(v)} = \frac{kz_1(z_1^{-\gamma} - 1)(1 - v)}{1 - z_1 + k_1z_1(z_1^{-\gamma} - 1) + [1 + z_1 - kz_1(z_1^{-\gamma} - 1)]v} \quad (5.3.25)$$

Для определения значений выходных переменных стационарных линейных систем с непрерывно изменяющимися выходными переменными в промежутке между импульсами можно пользоваться общей формулой (5.2.1), заменив в ней весовые коэффициенты $g_k(t)$ их выражением (5.3.2). При этом практически удобно определять момент времени t между двумя соседними импульсами относительно приращением времени после каждого импульса ε . Иными словами, любой момент времени в интервале между импульсами $lT_{\Pi} < t < (l+1)T_{\Pi}$ удобно определить формулой

$$t = lT_{\Pi} + \varepsilon T_{\Pi} \quad (0 < \varepsilon < 1). \quad (5.3.26)$$

Подставляя это выражение в формулу (5.2.1) и полагая для краткости

$$\left. \begin{aligned} y(lT_{\Pi} + \varepsilon T_{\Pi}) &= y_l(\varepsilon), \\ g_k(lT_{\Pi} + \varepsilon T_{\Pi}) &= g_0((l-k)T_{\Pi} + \varepsilon T_{\Pi}) = w_{l-k}(\varepsilon), \end{aligned} \right\} \quad (5.3.27)$$

получим для физически возможной стационарной дискретной линейной системы

$$y_l(\varepsilon) = \sum_{k=-\infty}^l w_{l-k}(\varepsilon) x_k. \quad (5.3.28)$$

Заменой индекса $m = l - k$ приведем эту формулу к виду

$$y_l(\varepsilon) = \sum_{m=0}^{\infty} w_m(\varepsilon) x_{l-m} \quad (0 \leq \varepsilon < 1; l = 0, \pm 1, \pm 2, \dots). \quad (5.3.29)$$

В случае показательного входного возмущения $x(t) = e^{st}$, $x_{l-m} = e^{s(l-m)T_{\Pi}}$, и формула (5.3.29) дает

$$y_l(\varepsilon) = e^{slT_{\Pi}} \sum_{m=0}^{\infty} w_m(\varepsilon) e^{-smT_{\Pi}}. \quad (5.3.30)$$

Функция

$$\Phi(s, \varepsilon) = \sum_{m=0}^{\infty} w_m(\varepsilon) e^{-smT_{\Pi}} \quad (0 \leq \varepsilon < 1) \quad (5.3.31)$$

называется *передаточной функцией* стационарной дискретной линейной системы с непрерывным выходом. При $s = i\omega$ формула (5.3.31) определяет частотную характеристику стационарной дискретной линейной системы с непрерывным выходом. Зависимость передаточной функции и частотной характеристики от ε отражает тот факт, что дискретная система может быть стацио-

нарной лишь по отношению к сдвигам во времени, кратным периоду повторения импульсов, и не может быть стационарной в полном смысле, т. е. по отношению к любым сдвигам во времени.

Так как на основании второй формулы (5.3.27) $w_m(0) = g_0(mT_n) = w_m$, то при $\varepsilon = 0$ передаточная функция стационарной дискретной линейной системы совпадает с ранее введенной передаточной функцией, определяемой формулой (5.3.12):

$$\Phi(z, 0) = \Phi(z). \quad (5.3.32)$$

Рассматривая передаточную функцию стационарной дискретной линейной системы с непрерывным выходом как функцию параметра $z = e^{sT_n}$, можно переписать формулу (5.3.31) в виде

$$\Psi(z, \varepsilon) = \sum_{m=0}^{\infty} w_m(\varepsilon) z^{-m} \quad (0 \leq \varepsilon < 1). \quad (5.3.33)$$

Формула (5.3.12) может быть получена из общей формулы (2.4.6), если заменить в ней весовую функцию $w(\xi)$ линейной комбинацией δ -функций:

$$w(\xi) = \sum_{m=0}^{\infty} w_m \delta(\xi - mT_n) \quad (5.3.34)$$

Иными словами, и для дискретной стационарной линейной системы передаточная функция является преобразованием Лапласа ее весовой функции. Преобразование Лапласа линейной комбинации (5.3.34) δ -функций, которое, само собой разумеется, является частным случаем преобразования Лапласа, называется *дискретным преобразованием Лапласа* числовой последовательности w_m . По аналогии функцию $\Psi(z)$, определяемую формулой (5.3.14), обычно называют *z-преобразованием* последовательности весовых коэффициентов w_m . Для нахождения передаточных функций стационарных дискретных линейных систем по их весовым коэффициентам и наоборот можно пользоваться имеющимися таблицами дискретного преобразования Лапласа и z -преобразования [72].

Пример 5.3.3. Найти весовые коэффициенты $w_m(\varepsilon)$ и передаточную функцию $\Psi(z, \varepsilon)$ дискретной системы, рассмотренной в примере 5.2.2.

Полагая в формуле (5.2.23) $t = lT_n + \varepsilon T_n$, $t_k = kT_n$, $l - k = m$, получим выражение для весовых коэффициентов $w_m(\varepsilon)$:

$$w_0(\varepsilon) = g_h(kT_n + \varepsilon T_n) = \begin{cases} k(1 - z_1^\varepsilon) & \text{при } 0 \leq \varepsilon < \gamma, \\ kz_1^\varepsilon(z_1^{-\gamma} - 1) & \text{при } \gamma \leq \varepsilon < 1, \end{cases} \quad (5.3.35)$$

$$w_m(\varepsilon) = g_h(lT_n + \varepsilon T_n) = kz_1^{m+\varepsilon}(z_1^{-\gamma} - 1) \quad (m=1, 2, \dots).$$

Подставляя эти выражения в формулу (5.3.33), находим передаточную функцию системы:

$$\begin{aligned} \Psi(z, \varepsilon) &= k(1 - z_1^\varepsilon) + kz_1^\varepsilon(z_1^{-\gamma} - 1) \sum_{m=1}^{\infty} z_1^m z^{-m} = \\ &= k(1 - z_1^\varepsilon) + kz_1^{1+\varepsilon} \frac{z_1^{-\gamma} - 1}{z - z_1} \quad \text{при } 0 \leq \varepsilon < \gamma, \\ \Psi(z, \varepsilon) &= kz_1^\varepsilon(z_1^{-\gamma} - 1) \sum_{m=0}^{\infty} z_1^m z^{-m} = kz_1^\varepsilon \frac{(z_1^{-\gamma} - 1)z}{z - z_1} \\ &\quad \text{при } \gamma \leq \varepsilon < 1. \end{aligned} \quad (5.3.36)$$

Полагая в первой формуле (5.3.36) $\varepsilon = 0$, получим ранее выведенную формулу (5.3.22).

§ 5.4. Стационарные дискретные линейные системы, описываемые разностными уравнениями

Рассматривая входную и выходную переменные дискретной системы только в определенные моменты времени $t_k = kT_n$ ($k = 0, \pm 1, \pm 2, \dots$), мы лишаемся возможности давать времени бесконечно близкие значения и совершать предельные переходы, неограниченно сближая эти значения. Вследствие этого становится неудобным характеризовать скорость изменения функции ее производной. Поэтому, изучая поведение дискретных систем лишь в определенные равноотстоящие моменты времени, целесообразно вместо производных ввести конечные разности функций.

Разностью первого порядка или первой разностью функции $f(t)$ называется ее приращение при изменении аргумента на данную величину T_n :

$$\Delta f(t) = f(t + T_n) - f(t). \quad (5.4.1)$$

Полагая здесь $t = kT_n$ и вводя для краткости для любой функции $f(t)$ обозначение

$$f_k = f(kT_n) \quad (k = 0, \pm 1, \pm 2, \dots), \quad (5.4.2)$$

получим

$$\Delta f_k = f_{k+1} - f_k \quad (k = 0, \pm 1, \pm 2, \dots). \quad (5.4.3)$$

Разностью второго порядка или второй разностью функции $f(t)$ называется разность первого порядка ее первой разности. Применяя формулу (5.4.3) к первой разности функции $f(t)$, получим для ее второй разности формулу

$$\Delta^2 f_k = \Delta f_{k+1} - \Delta f_k. \quad (5.4.4)$$

Таким образом, пользуясь понятием конечной разности функции, можно определить ее разности любых порядков рекуррентным

соотношением

$$\Delta^l f_k = \Delta^{l-1} f_{k+1} - \Delta^{l-1} f_k \quad (l = 1, 2, 3, \dots). \quad (5.4.5)$$

Очевидно, что все разности функции могут быть выражены через ее значения $f_k, f_{k+1}, f_{k+2}, \dots$. Действительно, из (5.4.4) и (5.4.3) следует, что

$$\Delta^2 f_k = f_{k+2} - 2f_{k+1} + f_k.$$

Подставляя это выражение в формулу (5.4.5) при $l = 3$, находим

$$\Delta^3 f_k = f_{k+3} - 3f_{k+2} + 3f_{k+1} - f_k.$$

Пользуясь методом полной математической индукции, легко получаем следующую общую формулу:

$$\Delta^l f_k = \sum_{h=0}^l (-1)^{l-h} C_l^h f_{k+h}, \quad (5.4.6)$$

где C_l^h — биномиальные коэффициенты.

Соотношение, связывающее значения неизвестной функции f_k и ее разностей различных порядков $\Delta f_k, \Delta^2 f_k, \dots$, называется *уравнением в конечных разностях* или *разностным уравнением*. Если это соотношение линейно, то оно представляет собой *линейное разностное уравнение*.

Разностное уравнение может содержать или значения неизвестной функции и ее разностей различных порядков при одном и том же значении аргумента $t = kT_n$, или значения неизвестной функции при различных равноотстоящих значениях аргумента. При помощи формул (5.4.3), (5.4.5) и (5.4.6) можно преобразовать разностное уравнение из одной формы в другую.

Разностное уравнение, содержащее разности неизвестной функции до порядка n или значения неизвестной функции от f_k до f_{k+n} , называется *разностным уравнением n -го порядка*.

Если поведение дискретной линейной системы описывается разностным уравнением, то на основании изложенного это уравнение всегда может быть написано в виде

$$\begin{aligned} a_n y_{k+n} + a_{n-1} y_{k+n-1} + \dots + a_1 y_{k+1} + a_0 y_k = \\ = b_m x_{k+m} + b_{m-1} x_{k+m-1} + \dots + b_1 x_{k+1} + b_0 x_k. \end{aligned} \quad (5.4.7)$$

Совершенно так же, как в § 2.5, легко доказывается, что для того, чтобы дискретная линейная система, описываемая разностным уравнением (5.4.7), была стационарной, необходимо и достаточно, чтобы все коэффициенты уравнения (5.4.7) были постоянными.

Разностное уравнение (5.4.7) может быть записано в операторной форме. Для этого введем оператор сдвига ∇ , определяющий сдвиг функции на период повторения T_n :

$$\nabla f(t) = f(t + T_n). \quad (5.4.8)$$

Таким образом, передаточная функция стационарной дискретной линейной системы, поведение которой описывается разностным уравнением, всегда является дробно-рациональной функцией переменной $z = e^{sT_n}$.

Мы видим, что передаточные функции стационарных дискретных линейных систем, описываемых разностными уравнениями, определяются совершенно так же, как и передаточные функции непрерывных стационарных линейных систем, описываемых дифференциальными уравнениями. В случае непрерывной системы оператор дифференцирования в дифференциальном уравнении заменяется параметром показательной функции s , а в случае дискретной системы оператор сдвига в разностном уравнении заменяется величиной $z = e^{sT_n}$. В обоих случаях в результате получается алгебраическое уравнение для передаточной функции системы. Заметим, что операторы дифференцирования и сдвига связаны тем же самым соотношением, что и величины s и z . Для доказательства разложим функцию f в ряд Тейлора с центром в точке t_k :

$$\begin{aligned} f_{k+1} &= f(t_k + T_n) = f(t_k) + T_n f'(t_k) + \dots + \frac{T_n^n}{n!} f^{(n)}(t_k) + \dots = \\ &= \left[1 + T_n \frac{d}{dt} + \dots + \frac{1}{n!} \left(T_n \frac{d}{dt} \right)^n + \dots \right] f(t_k). \end{aligned} \quad (5.4.16)$$

Ряд в квадратных скобках можно формально просуммировать. В результате получим

$$1 + T_n \frac{d}{dt} + \dots + \frac{1}{n!} \left(T_n \frac{d}{dt} \right)^n + \dots = e^{T_n \frac{d}{dt}} = e^{T_n D},$$

и формула (5.4.16) примет вид

$$f_{k+1} = e^{T_n D} f_k. \quad (5.4.17)$$

В данном случае $e^{T_n D}$ является сокращенной формой записи тех операций, которые необходимо произвести над значением функции f в точке t_k , чтобы получить ее значение в точке $t_{k+1} = t_k + T_n$. Сравнивая формулу (5.4.17) с формулой (5.4.9) при $l = 1$, получим формальное соотношение между операторами сдвига и дифференцирования:

$$\nabla = e^{T_n D}. \quad (5.4.18)$$

Сравнивая эту формулу с (5.3.13), убеждаемся в том, что операторы сдвига и дифференцирования связаны тем же соотношением, что и величины z и s .

Определим теперь передаточную функцию $\Psi(z)$ стационарной дискретной линейной системы, представляющей собой последовательное соединение импульсного элемента и непрерывной стационарной линейной системы, описываемой дифференциальным

уравнением

$$F(D)y = H(D)x, \quad (5.4.19)$$

предполагая, что степень n полинома F больше степени m полинома H . Сначала находим весовые коэффициенты системы. При этом для простоты будем предполагать, что полином $F(s)$ имеет только простые корни $\nu_1, \dots, \nu_n$. В этом случае весовая функция системы, описываемой уравнением (5.4.19), определяется формулой (4.4.50):

$$w(\xi) = \sum_{r=1}^n \frac{H(\nu_r)}{F'(\nu_r)} e^{\nu_r \xi} \quad (\xi \geq 0).$$

Подставляя это выражение в формулу (5.3.6), получим

$$w_0 = 0, \quad w_m = \sum_{r=1}^n \frac{H(\nu_r)}{F'(\nu_r)} e^{\nu_r m T_0} \int_0^{T_0} e^{-\nu_r \sigma} \eta(\sigma) d\sigma \quad (m = 1, 2, \dots).$$

Полагая для краткости

$$z_r = e^{\nu_r T_0}, \quad \kappa_r = \int_0^{T_0} e^{-\nu_r \sigma} \eta(\sigma) d\sigma \quad (r = 1, \dots, n), \quad (5.4.20)$$

можем переписать полученную формулу в виде

$$w_0 = 0, \quad w_m = \sum_{r=1}^n \kappa_r \frac{H(\nu_r)}{F'(\nu_r)} z_r^m \quad (m = 1, 2, \dots). \quad (5.4.21)$$

Зная весовые коэффициенты системы, можно воспользоваться для определения передаточной функции общей формулой (5.3.14). Подставив в нее выражение (5.4.21) весовых коэффициентов и изменив порядок суммирования, получим

$$\Psi(z) = \sum_{r=1}^n \kappa_r \frac{H(\nu_r)}{F'(\nu_r)} \sum_{m=1}^{\infty} z_r^m z^{-m}. \quad (5.4.22)$$

Бесконечный ряд в этой формуле является геометрической прогрессией со знаменателем $z_r z^{-1}$ и с первым членом $z_r z^{-1}$. Суммируя эту прогрессию, получим

$$\Psi(z) = \sum_{r=1}^n \kappa_r \frac{H(\nu_r)}{F'(\nu_r)} \frac{z_r z^{-1}}{1 - z_r z^{-1}}$$

или

$$\Psi(z) = \sum_{r=1}^n \frac{H(\nu_r)}{F'(\nu_r)} \frac{\kappa_r z_r}{z - z_r}. \quad (5.4.23)$$

Эта формула показывает, что передаточная функция последовательного соединения импульсного элемента и стационарной линейной системы, описываемой дифференциальным уравнением, является дробнорациональной функцией переменной z .

Если полином $F(s)$ имеет кратные корни, то весовая функция непрерывной части системы представляет собой линейную комбинацию показательных функций, умноженных на полиномы относительно времени. В этом случае в выражении (5.4.21) весовых коэффициентов w_m появятся еще слагаемые, пропорциональные величинам $(mT_n)^h z_r^m$, где h — целое положительное число, не превосходящее степени l полинома $F(s)$. Эти слагаемые дадут в выражении передаточной функции системы (5.4.22) дополнительные члены, пропорциональные рядам вида

$$\sum_{m=1}^{\infty} (mT_n)^h z_r^m z^{-m}.$$

Эти ряды легко суммируются. В самом деле, замечая, что

$$(mT_n)^h z_r^m = \frac{\partial^h}{\partial v_r^h} e^{v_r m T_n} = \frac{\partial^h z_r^m}{\partial v_r^h},$$

находим

$$\sum_{m=1}^{\infty} (mT_n)^h z_r^m z^{-m} = \frac{\partial^h}{\partial v_r^h} \sum_{m=1}^{\infty} z_r^m z^{-m} = \frac{\partial^h}{\partial v_r^h} \frac{z_r}{z - z_r}. \quad (5.4.24)$$

В результате и в случае кратных корней полинома $F(s)$ мы получим дробно-рациональную передаточную функцию рассматриваемой системы.

Приводя дроби в (5.4.23) или в аналогичной формуле в случае кратных корней полинома $F(s)$ к общему знаменателю, можем выразить передаточную функцию в виде отношения двух полиномов (5.4.15) и написать соответствующее разностное уравнение. Следовательно, поведение стационарной дискретной линейной системы, представляющей собой последовательное соединение импульсного элемента и стационарной линейной системы, описываемой дифференциальным уравнением, описывается некоторым разностным уравнением, которое легко находится по данному дифференциальному уравнению непрерывной части системы и данной функции $\eta(t)$, определяющей форму импульсов, вырабатываемых импульсным элементом.

Передаточная функция $\Omega(v)$ последовательного соединения импульсного элемента и стационарной линейной системы, описываемой обыкновенным дифференциальным уравнением, также представляет собой дробно-рациональную функцию параметра v , так как формула (5.3.17) выражает z как рациональную функцию v .

Пример 5.4.1. Найти передаточную функцию $\Psi(z)$ и разностное уравнение дискретной системы, представляющей собой последовательное соединение импульсного элемента и апериодического звена с постоянной времени T и коэффициентом усиления k , предполагая импульсы прямоугольными: $\eta(\sigma) = 1$.

Для определения передаточной функции $\Psi(z)$ воспользуемся формулой (5.4.23). Единственный в данном случае корень полинома

$$F(s) = Ts + 1$$

в уравнении (5.4.19) для апериодического звена равен $v_1 = -1/T$. Так как в данном случае $H(s) = k$, $F'(s) = T$, то формула (5.4.23) дает

$$\Psi(z) = \frac{k}{T} \frac{x_1 z_1}{z - z_1}, \quad z_1 = e^{-\frac{T}{T}}. \quad (5.4.25)$$

Остается определить коэффициент x_1 . Вторая формула (5.4.20) в данном случае дает

$$x_1 = \int_0^{T_H} e^{\frac{\sigma}{T}} d\sigma = T(e^{\frac{T_H}{T}} - 1).$$

Подставляя это выражение в (5.4.25), получим

$$\Psi(z) = k z_1 \frac{z_1^{-\gamma} - 1}{z - z_1}, \quad \gamma = \frac{T_H}{T}. \quad (5.4.26)$$

Разностное уравнение рассматриваемой дискретной системы имеет вид

$$(\nabla - z_1) y_k = k z_1 (z_1^{-\gamma} - 1) x_k$$

или

$$y_{k+1} - z_1 y_k = k z_1 (z_1^{-\gamma} - 1) x_k. \quad (5.4.27)$$

Пример 5.4.2. Найти передаточную функцию $\Psi(z)$ и разностное уравнение последовательного соединения импульсного элемента, вырабатывающего прямоугольные импульсы $\eta(\sigma) = 1$, и колебательного звена.

Используя результаты примера 4.4.3, находим по формуле (5.4.23):

$$\left. \begin{aligned} \Psi(z) &= \frac{k}{2i\omega_0} \left(\frac{x_1 z_1}{z - z_1} - \frac{x_2 z_2}{z - z_2} \right), \\ z_1 &= e^{-(a-i\omega_0)T_H}, \quad z_2 = e^{-(a+i\omega_0)T_H}, \end{aligned} \right\} \quad (5.4.28)$$

Формула (5.4.20) дает в данном случае следующие выражения для чисел x_1 и x_2 :

$$\begin{aligned} x_1 &= \int_0^{T_H} e^{(a-i\omega_0)\sigma} d\sigma = \frac{e^{(a-i\omega_0)T_H} - 1}{a - i\omega_0} = \frac{z_1^{-\gamma} - 1}{a - i\omega_0}, \\ x_2 &= \int_0^{T_H} e^{(a+i\omega_0)\sigma} d\sigma = \frac{e^{(a+i\omega_0)T_H} - 1}{a + i\omega_0} = \frac{z_2^{-\gamma} - 1}{a + i\omega_0}. \end{aligned}$$

Приводя дроби в (5.4.28) к общему знаменателю, приведем полученное выражение передаточной функции к виду

$$\Psi(z) = k \frac{(\kappa_1 z_1 - \kappa_2 z_2) z + (\kappa_2 - \kappa_1) z_1 z_2}{2i\omega_0 [z^2 - (z_1 + z_2) z + z_1 z_2]}.$$

Отсюда видно, что поведение последовательного соединения импульсного элемента и колебательного звена описывается линейным разностным уравнением второго порядка с постоянными коэффициентами:

$$\begin{aligned} y_{k+2} - (z_1 + z_2) y_{k+1} + z_1 z_2 y_k &= \\ &= \frac{k}{2i\omega_0} [(\kappa_1 z_1 - \kappa_2 z_2) x_{k+1} + (\kappa_2 - \kappa_1) z_1 z_2 x_k]. \end{aligned} \quad (5.4.29)$$

Легко видеть, что все коэффициенты этого уравнения действительны вследствие того, что z_1 , z_2 и κ_1 , κ_2 являются парами сопряженных комплексных чисел.

УСТОЙЧИВОСТЬ И КАЧЕСТВО ЛИНЕЙНЫХ СИСТЕМ

§ 6.1. Определение устойчивости. Общее условие устойчивости линейной системы

Всякая автоматическая система должна быть, прежде всего, работоспособной. Это значит, что она должна нормально функционировать и быть нечувствительной к неизбежным посторонним возмущениям различного рода. В частности, она должна нормально функционировать при действии на нее случайных помех и шумов. Иными словами, автоматическая система должна работать устойчиво, несмотря на действие на нее различных посторонних возмущений. Чтобы научиться проектировать устойчивые автоматические системы, необходимо сначала дать определение устойчивости.

Линейная система называется *устойчивой*, если ее выходная переменная остается ограниченной при любых ограниченных по абсолютной величине входных возмущениях. На основании принципа суперпозиции, справедливого для любых линейных систем, это определение можно также сформулировать следующим образом: линейная система называется устойчивой, если ее выходная переменная остается сколь угодно малой при любых достаточно малых по абсолютной величине входных возмущениях.

Для того чтобы физически возможная линейная система была устойчивой, необходимо и достаточно, чтобы ее весовая функция $g(t, \tau)$ удовлетворяла при всех t_0 условию

$$\lim_{t \rightarrow \infty} \int_{t_0}^t |g(t, \tau)| d\tau < c, \quad (6.1.1)$$

где c — некоторая конечная величина *).

Для доказательства достаточности условия (6.1.1) предположим, что входное возмущение представляет собой функцию $x(t)$,

*) На практике встречаются только такие системы, весовые функции которых ограничены при всех t и τ , кроме, может быть, дискретного ряда значений τ , которым соответствуют линейные комбинации δ -функций. Для таких систем интеграл в (6.1.1) конечен при любых конечных t_0 и t и условие (6.1.1) будет выполнено (или не выполнено) при всех t_0 , если оно выполнено (соответственно не выполнено) при каком-нибудь одном значении t_0 .

ограниченную по абсолютной величине: $|x(t)| \leq \eta$, где η — некоторая положительная величина. Тогда вследствие (6.1.1) будем иметь

$$|y(t)| = \left| \int_{t_0}^t g(t, \tau) x(\tau) d\tau \right| \leq \int_{t_0}^t |g(t, \tau)| |x(\tau)| d\tau \leq \eta \int_{t_0}^t |g(t, \tau)| d\tau < \eta c. \quad (6.1.2)$$

Таким образом, при выполнении условия (6.1.1) выходная переменная системы остается ограниченной при ограниченном входном возмущении. Из (6.1.2) следует также, что выходная переменная системы будет оставаться по абсолютной величине меньше любого заданного положительного числа ε при любых входных возмущениях, не превосходящих по абсолютной величине $\eta = \varepsilon/c$.

Докажем теперь, что условие (6.1.1) необходимо. Для этого предположим, что система устойчива и что для нее условие (6.1.1) не выполнено, т. е.

$$\lim_{t \rightarrow \infty} \int_{t_0}^t |g(t, \tau)| d\tau = \infty. \quad (6.1.3)$$

Вследствие предполагаемой устойчивости системы при любых t_0 , t' , t , $t_0 < t' \leq t$, справедливо неравенство

$$\left| \int_{t_0}^{t'} g(t, \tau) x(\tau) d\tau \right| < k(t), \quad (6.1.4)$$

где $k(t)$ — некоторая ограниченная функция, а $x(t)$ — любая функция, не превосходящая по абсолютной величине единицу: $|x(t)| \leq 1$. Возьмем теперь произвольную неограниченно возрастающую последовательность чисел $a_1 < a_2 < \dots < a_n < \dots$, $\lim_{n \rightarrow \infty} a_n = \infty$. Вследствие (6.1.3) для любого $\eta > 0$ существует

такая неограниченная последовательность моментов времени $t_0 < t_1 < \dots < t_n < \dots$, $t_n \rightarrow \infty$ при $n \rightarrow \infty$, для которой

$$\int_{t_{n-1}}^{t_n} |g(t_n, \tau)| d\tau > k(t_n) + \frac{a_n}{\eta} \quad (n = 1, 2, \dots). \quad (6.1.5)$$

Рассмотрим теперь входное возмущение вида *)

$$x(\tau) = \eta \operatorname{sgn} g(t_n, \tau) \text{ при } t_{n-1} < \tau < t_n \quad (n = 1, 2, \dots). \quad (6.1.6)$$

*) Функция $\operatorname{sgn} x$ (читается «сигнум икс») представляет собой знак аргумента x , т. е. равна $+1$ при $x > 0$ и -1 при $x < 0$. Вследствие этого при любых x имеет место тождество $x \operatorname{sgn} x = |x|$.

Это входное возмущение не превосходит по абсолютной величине произвольную величину $\eta > 0$. Выходная переменная системы в момент t_n при входном возмущении (6.1.6) будет равна

$$\begin{aligned} y(t_n) &= \int_{t_0}^{t_{n-1}} g(t_n, \tau) x(\tau) d\tau + \int_{t_{n-1}}^{t_n} g(t_n, \tau) x(\tau) d\tau = \\ &= \int_{t_0}^{t_{n-1}} g(t_n, \tau) x(\tau) d\tau + \eta \int_{t_{n-1}}^{t_n} |g(t_n, \tau)| d\tau. \end{aligned} \quad (6.1.7)$$

Но из (6.1.4) следует, что

$$\int_{t_0}^{t_{n-1}} g(t_n, \tau) x(\tau) d\tau > -\eta k(t_n), \quad (6.1.8)$$

так как в данном случае $|x(\tau)| = \eta$. На основании (6.1.5) и (6.1.8) из (6.1.7) вытекает неравенство

$$y(t_n) > -\eta k(t_n) + \eta \left[k(t_n) + \frac{a_n}{\eta} \right] = a_n \rightarrow \infty. \quad (6.1.9)$$

Таким образом, если условие (6.1.1) для линейной системы не выполнено, то существует сколь угодно малое входное возмущение, при котором выходная переменная системы неограниченно возрастает при $t \rightarrow \infty$, т. е. система не может быть устойчивой. Полученное противоречие доказывает необходимость условия (6.1.1).

Из (6.1.1) следует, что для непрерывной устойчивой системы

$$\lim_{t \rightarrow \infty} g(t, \tau) = 0 \quad (6.1.10)$$

при любом τ . Однако это условие не достаточно для устойчивости непрерывной системы.

В теоретических исследованиях часто приходится рассматривать и такие системы, у которых выходной сигнал содержит линейную комбинацию производных входного сигнала. Мы будем называть такие системы *дифференцирующими*. Весовая функция линейной дифференцирующей системы содержит линейную комбинацию δ -функции и ее производных. Дифференцирующие системы всегда неустойчивы в смысле данного выше определения. Поэтому, говоря об устойчивости дифференцирующей линейной системы, мы всегда будем подразумевать устойчивость системы, весовая функция которой получается из весовой функции данной дифференцирующей системы путем отбрасывания δ -функции и ее производных.

Можно также определить устойчивость дифференцирующей системы как свойство иметь ограниченный выходной сигнал при любых входных сигналах, ограниченных вместе со своими производными до соответствующего порядка.

Пример 6.1.1. Определить, устойчиво или неустойчиво аperiodическое звено, весовая функция которого определяется формулой

$$g(t, \tau) = \frac{k}{T} e^{-\frac{t-\tau}{T}}.$$

В этом случае

$$\int_{t_0}^t |g(t, \tau)| d\tau = \frac{|k|}{|T|} \int_{t_0}^t e^{-\frac{t-\tau}{T}} d\tau = |k| \left| 1 - e^{-\frac{t-t_0}{T}} \right|.$$

Таким образом, при положительной постоянной времени T интеграл (6.1.1) при любых t_0 и t не превосходит абсолютной величины коэффициента усиления аperiodического звена k . При отрицательной постоянной времени T интеграл в (6.1.1) неограниченно возрастает. Таким образом, аperiodическое звено устойчиво при положительной постоянной времени и неустойчиво при отрицательной постоянной времени.

Пример 6.1.2. Определить, устойчиво или неустойчиво колебательное звено, поведение которого описывается дифференциальным уравнением

$$y'' + 2ay' + b^2y = x, \quad |a| < b.$$

Весовая функция колебательного звена определяется формулой (пример 4.4.2)

$$g(t, \tau) = \frac{1}{\omega_0} e^{-a(t-\tau)} \sin \omega_0(t-\tau),$$

где $\omega_0 = \sqrt{b^2 - a^2}$. Следовательно, в этом случае

$$\begin{aligned} \int_{t_0}^t |g(t, \tau)| d\tau &= \int_0^{t-t_0} |g(t, t-\xi)| d\xi = \frac{1}{\omega_0} \int_0^{\omega_0(t-t_0)} \left| g\left(t, t - \frac{\eta}{\omega_0}\right) \right| d\eta = \\ &= \frac{1}{\omega_0^2} \sum_{k=1}^N \int_{(k-1)\pi}^{k\pi} e^{-\frac{a\eta}{\omega_0}} |\sin \eta| d\eta + \frac{1}{\omega_0^2} \int_{N\pi}^{\omega_0(t-t_0)} e^{-\frac{a\eta}{\omega_0}} |\sin \eta| d\eta = \\ &= \frac{1}{\omega_0^2} \sum_{k=1}^N e^{-\frac{a(k-1)\pi}{\omega_0}} \int_0^\pi e^{-\frac{a\sigma}{\omega_0}} \sin \sigma d\sigma + \frac{1}{\omega_0^2} \int_{N\pi}^{\omega_0(t-t_0)} e^{-\frac{a\eta}{\omega_0}} |\sin \eta| d\eta, \end{aligned}$$

где $N = [\omega_0(t-t_0)/\pi]$. Выполняя интегрирование и суммируя геометрическую прогрессию, получим

$$\begin{aligned} \int_{t_0}^t |g(t, \tau)| d\tau &= \frac{1 + e^{-\frac{a\pi}{\omega_0}} - 2e^{-\frac{a(N+1)\pi}{\omega_0}}}{b^2(1 - e^{-\frac{a\pi}{\omega_0}})} + \\ &+ \frac{(-1)^{N-1}}{\omega_0 b^2} e^{-a(t-t_0)} [a \sin \omega_0(t-t_0) + \omega_0 \cos \omega_0(t-t_0)]. \quad (6.1.11) \end{aligned}$$

При $a > 0$ правая часть этого равенства остается конечной при $t \rightarrow \infty$ (при этом, очевидно, и $N \rightarrow \infty$). При $a < 0$ правая часть равенства (6.1.11) неогра-

* Через $[x]$ обозначена целая часть числа x , т. е. наибольшее целое число, содержащееся в числе x .

Это входное возмущение не превосходит по абсолютной величине произвольную величину $\eta > 0$. Выходная переменная системы в момент t_n при входном возмущении (6.1.6) будет равна

$$\begin{aligned} y(t_n) &= \int_{t_0}^{t_{n-1}} g(t_n, \tau) x(\tau) d\tau + \int_{t_{n-1}}^{t_n} g(t_n, \tau) x(\tau) d\tau = \\ &= \int_{t_0}^{t_{n-1}} g(t_n, \tau) x(\tau) d\tau + \eta \int_{t_{n-1}}^{t_n} |g(t_n, \tau)| d\tau. \end{aligned} \quad (6.1.7)$$

Но из (6.1.4) следует, что

$$\int_{t_0}^{t_{n-1}} g(t_n, \tau) x(\tau) d\tau > -\eta k(t_n), \quad (6.1.8)$$

так как в данном случае $|x(\tau)| = \eta$. На основании (6.1.5) и (6.1.8) из (6.1.7) вытекает неравенство

$$y(t_n) > -\eta k(t_n) + \eta \left[k(t_n) + \frac{a_n}{\eta} \right] = a_n \rightarrow \infty. \quad (6.1.9)$$

Таким образом, если условие (6.1.1) для линейной системы не выполнено, то существует сколь угодно малое входное возмущение, при котором выходная переменная системы неограниченно возрастает при $t \rightarrow \infty$, т. е. система не может быть устойчивой. Полученное противоречие доказывает необходимость условия (6.1.1).

Из (6.1.1) следует, что для непрерывной устойчивой системы

$$\lim_{t \rightarrow \infty} g(t, \tau) = 0 \quad (6.1.10)$$

при любом τ . Однако это условие не достаточно для устойчивости непрерывной системы.

В теоретических исследованиях часто приходится рассматривать и такие системы, у которых выходной сигнал содержит линейную комбинацию производных входного сигнала. Мы будем называть такие системы *дифференцирующими*. Весовая функция линейной дифференцирующей системы содержит линейную комбинацию δ -функции и ее производных. Дифференцирующие системы всегда неустойчивы в смысле данного выше определения. Поэтому, говоря об устойчивости дифференцирующей линейной системы, мы всегда будем подразумевать устойчивость системы, весовая функция которой получается из весовой функции данной дифференцирующей системы путем отбрасывания δ -функции и ее производных.

Можно также определить устойчивость дифференцирующей системы как свойство иметь ограниченный выходной сигнал при любых входных сигналах, ограниченных вместе со своими производными до соответствующего порядка.

Пример 6.1.1. Определить, устойчиво или неустойчиво апериодическое звено, весовая функция которого определяется формулой

$$g(t, \tau) = \frac{k}{T} e^{-\frac{t-\tau}{T}}.$$

В этом случае

$$\int_{t_0}^t |g(t, \tau)| d\tau = \frac{|k|}{|T|} \int_{t_0}^t e^{-\frac{t-\tau}{T}} d\tau = |k| \left| 1 - e^{-\frac{t-t_0}{T}} \right|.$$

Таким образом, при положительной постоянной времени T интеграл (6.1.1) при любых t_0 и t не превосходит абсолютной величины коэффициента усиления апериодического звена k . При отрицательной постоянной времени T интеграл в (6.1.1) неограниченно возрастает. Таким образом, апериодическое звено устойчиво при положительной постоянной времени и неустойчиво при отрицательной постоянной времени.

Пример 6.1.2. Определить, устойчиво или неустойчиво колебательное звено, поведение которого описывается дифференциальным уравнением

$$y'' + 2ay' + b^2y = x, \quad |a| < b.$$

Весовая функция колебательного звена определяется формулой (пример 4.4.2)

$$g(t, \tau) = \frac{1}{\omega_0} e^{-a(t-\tau)} \sin \omega_0(t-\tau),$$

где $\omega_0 = \sqrt{b^2 - a^2}$. Следовательно, в этом случае

$$\begin{aligned} \int_{t_0}^t |g(t, \tau)| d\tau &= \int_0^{t-t_0} |g(t, t-\xi)| d\xi = \frac{1}{\omega_0} \int_0^{t-t_0} \left| g\left(t, t - \frac{\eta}{\omega_0}\right) \right| d\eta = \\ &= \frac{1}{\omega_0^2} \sum_{k=1}^N \int_{(k-1)\pi}^{k\pi} e^{-\frac{a\eta}{\omega_0}} |\sin \eta| d\eta + \frac{1}{\omega_0^2} \int_{N\pi}^{\omega_0(t-t_0)} e^{-\frac{a\eta}{\omega_0}} |\sin \eta| d\eta = \\ &= \frac{1}{\omega_0^2} \sum_{k=1}^N e^{-\frac{a(k-1)\pi}{\omega_0}} \int_0^{\pi} e^{-\frac{a\sigma}{\omega_0}} \sin \sigma d\sigma + \frac{1}{\omega_0^2} \int_{N\pi}^{\omega_0(t-t_0)} e^{-\frac{a\eta}{\omega_0}} |\sin \eta| d\eta, \end{aligned}$$

где $N = [\omega_0(t-t_0)/\pi]$. Выполняя интегрирование и суммируя геометрическую прогрессию, получим

$$\begin{aligned} \int_{t_0}^t |g(t, \tau)| d\tau &= \frac{1 + e^{-\frac{a\pi}{\omega_0}} - 2e^{-\frac{a(N+1)\pi}{\omega_0}}}{b^2(1 - e^{-\frac{a\pi}{\omega_0}})} + \\ &+ \frac{(-1)^{N-1}}{\omega_0 b^2} e^{-a(t-t_0)} [a \sin \omega_0(t-t_0) + \omega_0 \cos \omega_0(t-t_0)]. \quad (6.1.11) \end{aligned}$$

При $a > 0$ правая часть этого равенства остается конечной при $t \rightarrow \infty$ (при этом, очевидно, и $N \rightarrow \infty$). При $a < 0$ правая часть равенства (6.1.11) неогра-

*) Через $[x]$ обозначена целая часть числа x , т. е. наибольшее целое число, содержащееся в числе x .

неченно возрастает при $t \rightarrow \infty$. Если $a = 0$, то правая часть равенства (6.1.11) неопределенна. Раскрывая эту неопределенность по правилу Лопиталя, убеждаемся в том, что при $a = 0$ интеграл (6.1.11) также неограниченно возрастает при $t \rightarrow \infty$. Таким образом, колебательное звено устойчиво, если $a > 0$, и неустойчиво, если $a \leq 0$.

§ 6.2. Устойчивость стационарных линейных систем.

Критерии Рауса и Гурвица

Для того чтобы условие (6.1.1) было выполнено для физически возможной стационарной линейной системы, необходимо и достаточно, чтобы все особые точки передаточной функции $\Phi(s)$ этой системы лежали в левой полуплоскости комплексной переменной s . В этом случае весовая функция стационарной линейной системы $w(t - \tau)$ убывает при $t \rightarrow \infty$ быстрее, чем $e^{-\varepsilon(t-\tau)}$, где ε — некоторое положительное число, и интеграл (6.1.1) ограничен при $t \rightarrow \infty$. Мы не будем доказывать здесь это утверждение в общем виде, а ограничимся частным случаем физически возможных стационарных линейных систем, поведение которых описывается обыкновенными дифференциальными уравнениями.

Рассмотрим стационарную линейную систему, поведение которой описывается обыкновенным линейным дифференциальным уравнением с постоянными коэффициентами

$$F(D)y = H(D)x. \quad (6.2.1)$$

Если все корни характеристического уравнения $v_1, \dots, v_n$ различны, то весовая функция стационарной линейной системы определяется формулой (4.5.50):

$$g(t, \tau) = w(t - \tau) = \sum_{r=1}^n \frac{H(v_r)}{F'(v_r)} e^{v_r(t-\tau)}. \quad (6.2.2)$$

Обозначим через α_r, β_r соответственно действительную и мнимую части корня v_r характеристического уравнения. Так как нумерация корней характеристического уравнения безразлична, то мы предположим, что α_1 — наибольшее из чисел $\alpha_1, \dots, \alpha_r$. Тогда будет $\alpha_r - \alpha_1 \leq 0$ ($r = 1, \dots, n$). Представим весовую функцию системы формулой

$$w(t - \tau) = e^{(\alpha_1 + \varepsilon)(t-\tau)} \varphi(t - \tau), \quad (6.2.3)$$

где согласно (6.2.2)

$$\varphi(t - \tau) = \sum_{r=1}^n \frac{H(v_r)}{F'(v_r)} e^{(\alpha_r - \alpha_1 - \varepsilon + i\beta_r)(t-\tau)},$$

а ε — произвольно малое положительное число. Так как показатели функции в выражении функции φ все по модулю меньше

единицы, то

$$|\varphi(t-\tau)| < \sum_{r=1}^n \left| \frac{H(v_r)}{F'(v_r)} \right| < c, \quad (6.2.4)$$

где c — некоторая постоянная. Если характеристическое уравнение имеет кратные корни, то весовая функция системы представляет собой линейную комбинацию произведений показательных функций на полиномы. Обозначая и в этом случае через α_1 наибольшую из действительных частей корней характеристического уравнения, выразим весовую функцию системы формулой (6.2.3). При этом функция φ будет представлять собой сумму произведений полиномов на показательные функции с отрицательными действительными частями показателей и, следовательно, будет непрерывной ограниченной функцией. Таким образом, весовая функция стационарной линейной системы, поведение которой описывается дифференциальным уравнением, всегда может быть выражена формулой (6.2.3), где ε — произвольно малое положительное число, а функция φ непрерывна и ограничена:

$$|\varphi(t-\tau)| < c. \quad (6.2.5)$$

На основании (6.2.3) имеем

$$\int_{t_0}^t |w(t-\tau)| d\tau = \int_0^{t-t_0} |w(\xi)| d\xi = \int_0^{t-t_0} e^{(\alpha_1+\varepsilon)\xi} |\varphi(\xi)| d\xi. \quad (6.2.6)$$

Пользуясь теоремой о среднем и выполняя интегрирование, получим

$$\int_{t_0}^t |w(t-\tau)| d\tau = |\varphi(\xi)|_{\text{ср}} \frac{e^{(\alpha_1+\varepsilon)(t-t_0)} - 1}{\alpha_1 + \varepsilon}. \quad (6.2.7)$$

При любом $\alpha_1 < 0$ положительную величину ε можно выбрать меньшей, чем $|\alpha_1|$, чтобы было $\alpha_1 + \varepsilon < 0$. Тогда интеграл (6.2.7) не будет превосходить величину $c/|\alpha_1 + \varepsilon|$ при любых значениях t_0 и t .

Для исследования случая $\alpha_1 \geq 0$ положим $\varepsilon = 0$. В этом случае функция $\varphi(\xi)$ стремится к постоянной при $\xi \rightarrow \infty$, если характеристическое уравнение не имеет ни одного кратного корня с действительной частью α_1 , и неограниченно возрастает при $\xi \rightarrow \infty$, если характеристическое уравнение имеет кратные корни с действительной частью α_1 . Следовательно, при $\varepsilon = 0$ величина $|\varphi(\xi)|_{\text{ср}}$ в формуле (6.2.7) не может неограниченно убывать при $t \rightarrow \infty$. Но в таком случае из формулы (6.2.7) при $\varepsilon = 0$ следует, что при любом $\alpha_1 \geq 0$ интеграл от абсолютной величины весовой функции системы неограниченно возрастает при $t \rightarrow \infty$ *).

*) При $\alpha_1 = \varepsilon = 0$ дробь в правой части равенства (6.2.7) равна $t - t_0$. В этом можно убедиться, выполнив интегрирование в (6.2.6) при $\alpha_1 = \varepsilon = 0$, после выноса функции $\varphi(\xi)$ за знак интеграла средним значением.

Таким образом, мы доказали, что стационарная линейная система, поведение которой описывается обыкновенным дифференциальным уравнением, устойчива тогда и только тогда, когда все корни ее характеристического уравнения имеют отрицательные действительные части $\alpha_r \leq \alpha_1 < 0$ ($r = 1, \dots, n$). Иными словами, стационарная линейная система, описываемая уравнением (6.2.1), устойчива тогда и только тогда, когда все полюсы ее передаточной функции

$$\Phi(s) = \frac{H(s)}{F(s)} \quad (6.2.8)$$

лежат в левой полуплоскости комплексной переменной s .

Доказанное утверждение позволяет свести задачу исследования устойчивости стационарных линейных систем, описываемых дифференциальными уравнениями, к чисто алгебраической задаче нахождения условий, которым должны удовлетворять коэффициенты полинома для того, чтобы все его корни имели отрицательные действительные части. Такие условия были впервые найдены в 1877 г. Раусом [59]. Другая форма условий, которым должны удовлетворять коэффициенты полинома для того, чтобы все действительные части его корней были отрицательными, была найдена Гурвицем, работа которого была опубликована в 1895 г. [16]. Мы приведем здесь условия Рауса и Гурвица без доказательств.

Для того чтобы дать условия Рауса и Гурвица в общепринятой форме, мы несколько изменим обозначения коэффициентов, которые мы применяли в §§ 2.5 и 4.4, и представим полином $F(s)$ в виде

$$F(s) = a_0 s^n + a_1 s^{n-1} + \dots + a_{n-1} s + a_n. \quad (6.2.9)$$

Докажем прежде всего, что для того, чтобы полином $F(s)$ имел только корни с отрицательными действительными частями, необходимо, чтобы все его коэффициенты имели один и тот же знак. Так как от перемены знаков всех коэффициентов корни полинома не изменяются, то достаточно рассмотреть случай, когда $a_0 > 0$. На основании известной теоремы алгебры полином $F(s)$ может быть выражен в форме

$$F(s) = a_0 (s - v_1) (s - v_2) \dots (s - v_n). \quad (6.2.10)$$

Из алгебры известно, что корни полинома с действительными коэффициентами или действительны, или являются попарно сопряженными комплексными величинами. Для отрицательного действительного корня $v_r = -\alpha$

$$s - v_r = s + \alpha.$$

Для двух комплексных сопряженных корней с отрицательными действительными частями $v_r = -\alpha + i\beta$, $v_{r+1} = -\alpha - i\beta$

$$\begin{aligned} (s - v_r) (s - v_{r+1}) &= [s - (-\alpha + i\beta)] [s - (-\alpha - i\beta)] = \\ &= s^2 + 2\alpha s + \alpha^2 + \beta^2. \end{aligned}$$

Таким образом, полином $F(s)$ является произведением линейных двучленов и квадратных трехчленов с положительными коэффициентами. Отсюда следует, что все коэффициенты полинома $F(s)$ положительны, что и требовалось доказать. Образует теперь матрицу

$$\begin{vmatrix} \cdot & \cdot & \cdot & \cdot & \cdot \\ c_1 & c_3 & c_5 & \dots \\ c_0 & c_2 & c_4 & \dots \\ b_1 & b_3 & b_5 & \dots \\ b_0 & b_2 & b_4 & \dots \\ a_1 & a_3 & a_5 & \dots \\ a_0 & a_2 & a_4 & \dots \end{vmatrix} \quad (6.2.11)$$

по следующему закону. Элементами нижней строки являются коэффициенты полинома $F(s)$ с четными номерами. Элементами второй снизу строки являются коэффициенты полинома $F(s)$ с нечетными номерами. Элементы b_i следующих двух строк определяются формулами

$$\left. \begin{aligned} b_0 &= \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix}, & b_2 &= \begin{vmatrix} a_1 & a_5 \\ a_0 & a_4 \end{vmatrix} + \dots + \begin{vmatrix} a_1 & a_{2k+3} \\ a_0 & a_{2k+2} \end{vmatrix}, \\ b_1 &= \begin{vmatrix} b_0 & b_2 \\ a_1 & a_3 \end{vmatrix}, & b_3 &= \begin{vmatrix} b_0 & b_4 \\ a_1 & a_5 \end{vmatrix} + \dots + \begin{vmatrix} b_0 & b_{2k+2} \\ a_1 & a_{2k+3} \end{vmatrix}. \end{aligned} \right\} \quad (6.2.12)$$

Элементы c_i определяются теми же формулами (6.2.12), в которых величины a и b заменены соответственно величинами b и c . И так, далее, элементы каждых двух последующих строк выражаются через элементы двух предыдущих строк формулами (6.2.12).

Раус доказал, что для того, чтобы действительные части всех корней полинома $F(s)$ были отрицательными, необходимо и достаточно, чтобы все элементы первого столбца матрицы (6.2.11) были положительными:

$$a_0 > 0, a_1 > 0, b_0 > 0, b_1 > 0, c_0 > 0, c_1 > 0, \dots \quad (6.2.13)$$

Пример 6.2.1. В случае уравнения (6.2.1) первого порядка

$$F(s) = a_0 s + a_1,$$

все величины $b_i, c_i, \dots$, определяемые формулами (6.2.12), равны нулю и условия устойчивости (6.2.13) сводятся к

$$a_0 > 0, a_1 > 0. \quad (6.2.14)$$

Таким образом, для системы, поведение которой описывается дифференциальным уравнением первого порядка, условие одинаковости знаков коэффициентов уравнения не только необходимо, но и достаточно.

Пример 6.2.2. В случае уравнения (6.2.1) второго порядка

$$F(s) = a_0 s^2 + a_1 s + a_2,$$

$b_0 = a_1 a_2$, а все остальные величины $b_r, c_r, \dots$ равны нулю. Условия устойчивости (6.2.13) в этом случае принимают вид

$$a_0 > 0, \quad a_1 > 0, \quad a_2 > 0. \quad (6.2.15)$$

Таким образом, для устойчивости системы, описываемой дифференциальным уравнением второго порядка, необходимо и достаточно, чтобы все коэффициенты этого уравнения имели одинаковые знаки.

Пример 6.2.3. В случае уравнения (6.2.1) третьего порядка

$$F(s) = a_0 s^3 + a_1 s^2 + a_2 s + a_3$$

и формулы (6.2.12) дают

$$b_0 = a_1 a_2 - a_0 a_3, \quad b_1 = b_0 a_3.$$

Все остальные величины $b_r, c_r, \dots$ равны нулю. Условия устойчивости (6.2.13) в этом случае принимают вид

$$a_0 > 0, \quad a_1 > 0, \quad a_1 a_2 - a_0 a_3 > 0, \quad a_3 > 0. \quad (6.2.16)$$

Отсюда видно, что для устойчивости стационарной системы, описываемой линейным дифференциальным уравнением третьего порядка, условие положительности коэффициентов необходимо, но не достаточно.

Пример 6.2.4. В случае уравнения (6.2.1) четвертого порядка

$$F(s) = a_0 s^4 + a_1 s^3 + a_2 s^2 + a_3 s + a_4$$

и формулы (6.2.12) дают

$$b_0 = a_1 a_2 - a_0 a_3, \quad b_2 = a_1 a_4, \quad b_1 = b_0 a_3 - b_2 a_1, \quad c_0 = b_1 b_2.$$

Все остальные величины $b_r, c_r, \dots$ равны нулю. Условия устойчивости (6.2.13) в этом случае принимают вид

$$a_0 > 0, \quad a_1 > 0, \quad a_1 a_2 - a_0 a_3 > 0, \quad a_1 a_2 a_3 - a_0 a_3^2 - a_1^2 a_4 > 0, \quad a_4 > 0. \quad (6.2.17)$$

Гурвиц доказал, что для того, чтобы действительные части всех корней полинома $F(s)$ были отрицательными, необходимо и достаточно, чтобы определитель

$$\Delta_n = \begin{vmatrix} a_1 & a_3 & a_5 & \dots & 0 & 0 \\ a_0 & a_2 & a_4 & \dots & 0 & 0 \\ 0 & a_1 & a_3 & \dots & 0 & 0 \\ \dots & \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_{n-1} & 0 \\ 0 & 0 & 0 & \dots & a_{n-2} & a_n \end{vmatrix} \quad (6.2.18)$$

и все его диагональные миноры, отделенные в (6.2.18) пунктирными линиями, были положительными:

$$\left. \begin{aligned} \Delta_1 &= a_1 > 0, \quad \Delta_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix} > 0, \\ \Delta_3 &= \begin{vmatrix} a_1 & a_3 & a_5 \\ a_0 & a_2 & a_4 \\ 0 & a_1 & a_3 \end{vmatrix} > 0, \quad \dots, \quad \Delta_{n-1} > 0, \quad \Delta_n > 0. \end{aligned} \right\} \quad (6.2.19)$$

Раскладывая определитель (6.2.18) по элементам последнего столбца, получим

$$\Delta_n = a_n \Delta_{n-1}. \quad (6.2.20)$$

Таким образом, условие положительности определителя Δ_n может быть заменено условием положительности последнего коэффициента a_n . Тогда получим следующие условия устойчивости:

$$a_1 > 0, \quad \Delta_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix} > 0, \quad \dots, \quad \Delta_{n-1} > 0, \quad a_n > 0. \quad (6.2.21)$$

Пример 6.2.5. В условиях примера 6.2.3 определитель Гурвица и его диагональные миноры выражаются формулами

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

Условия устойчивости (6.2.21) принимают в этом случае найденную ранее форму (6.2.16).

Пример 6.2.6. В условиях примера 6.2.4 определитель Гурвица и его диагональные миноры выражаются формулами

$$\Delta_4 = \begin{vmatrix} a_1 & a_3 & 0 & 0 \\ a_0 & a_2 & a_4 & 0 \\ 0 & a_1 & a_3 & 0 \\ 0 & a_0 & a_2 & a_4 \end{vmatrix}, \quad \Delta_1 = a_1, \quad \Delta_2 = \begin{vmatrix} a_1 & a_3 \\ a_0 & a_2 \end{vmatrix}, \quad \Delta_3 = \begin{vmatrix} a_1 & a_3 & 0 \\ a_0 & a_2 & a_4 \\ 0 & a_1 & a_3 \end{vmatrix}.$$

Таким образом, условия устойчивости (6.2.21) в данном случае, как нетрудно видеть, сводятся к найденным ранее неравенствам (6.2.17).

Приведенные примеры показывают, что критерии устойчивости Рауса и Гурвица по существу равноценны. Однако вычисление величин $b_r, c_r, \dots$ по формулам (6.2.12) обычно бывает значительно проще, чем вычисление определителя Гурвица и его миноров.

Легко видеть, что первые четыре диагональных минора определителя Гурвица $\Delta_1, \Delta_2, \Delta_3, \Delta_4$ равны соответственно величинам $a_1, b_0, b_1, c_0/a_1$. Остальные диагональные миноры определителя Гурвица также выражаются через элементы матрицы (6.2.11).

§ 6.3. Устойчивость стационарных линейных систем.

Критерий Найквиста. Запасы устойчивости

Изложенные в предыдущем параграфе аналитические методы исследования устойчивости Рауса и Гурвица удобны в случае, когда поведение стационарной линейной системы описывается дифференциальным уравнением сравнительно невысокого порядка. Для систем, описываемых дифференциальными уравнениями

высоких порядков, исследование устойчивости по критериям Рауса и Гурвица требует громоздких и трудоемких вычислений. Кроме того, само нахождение дифференциального уравнения сложной системы связано с громоздкими выкладками и трудоемкими вычислениями. Между тем, как мы видели в § 4.6, частотные характеристики легко находятся для любых сколь угодно сложных систем простыми графическими и алгебраическими операциями. Поэтому, естественно, возникает вопрос, нельзя ли непосредственно по частотным характеристикам системы определить, устойчива она или нет. Частотные критерии устойчивости стационарных линейных систем были найдены Найквистом и Михайловым. Особенно удобным с практической точки зрения является критерий Найквиста [45]. Его мы здесь и изложим.

Очевидно, что при любых последовательных и параллельных соединениях устойчивых систем всегда будет получаться устойчивая система. Если же среди соединяемых последовательно или параллельно систем имеется хотя бы одна неустойчивая, то и вся система, полученная в результате соединения, будет неустойчивой. Поэтому исследование устойчивости любой линейной системы, полученной путем последовательного и параллельного соединения любого количества элементарных систем, сводится к исследованию устойчивости отдельных элементарных систем, входящих в состав этой системы. Для элементарных звеньев эта задача решается совершенно просто изложенными в предыдущем параграфе методами. При этом легко определяются все нули и полюсы передаточных функций элементарных звеньев. Зная полюсы передаточных функций элементарных звеньев, легко определить, какие полюсы в правой полуплоскости будет иметь передаточная функция системы, полученной путем последовательных и параллельных соединений этих звеньев, в случае, если она неустойчива. Таким образом, для полного исследования устойчивости сложных стационарных линейных систем остается исследовать влияние на устойчивость обратных связей. При этом, очевидно, достаточно рассмотреть только случай жесткой отрицательной обратной связи, так как гибкая обратная связь по доказанному в § 4.7 простыми структурными преобразованиями сводится к жесткой. Эту задачу решает критерий Найквиста, который позволяет, зная, какие полюсы имеет передаточная функция системы в правой полуплоскости и на мнимой оси, решить вопрос о том, устойчивой или неустойчивой будет система при замыкании ее жесткой отрицательной обратной связью.

Передаточная функция $\Phi(s)$ стационарной линейной системы приводит в соответствие каждому значению ее аргумента s комплексное число $\varphi = \Phi(s)$, которое может быть изображено на комплексной числовой плоскости φ . При изменении s изображающая это число точка будет описывать на плоскости комплексной пере-

менной s некоторую кривую C (рис. 6.3.1, а). При этом соответствующая точка $\varphi = \Phi(s)$ будет описывать некоторую другую кривую Γ в плоскости переменной φ (рис. 6.3.1, б). Таким образом, передаточная функция $\Phi(s)$ устанавливает отображение комплексной плоскости переменной s на плоскость переменной φ . Кривая, описываемая точкой $\varphi = \Phi(s)$ на плоскости переменной φ , является отображением кривой, описываемой точкой s на плоскости переменной s . В качестве примера такого отображения можно указать амплитудно-фазовую характеристику системы, которая, как следует из изложенного в § 2.4, является отображением мнимой оси плоскости переменной s .

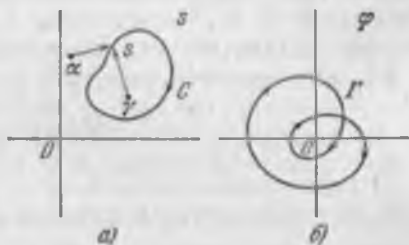


Рис. 6.3.1.

Предположим теперь, что точка s описывает некоторую замкнутую кривую C по ходу часовой стрелки. На рис. 6.3.1 видно, что для любой точки α , лежащей вне области, ограниченной кривой C , изменение аргумента комплексного числа $s - \alpha$ в результате обхода точкой s кривой C равно нулю. Для любой точки γ , лежащей внутри области, ограниченной кривой C , вектор, изображающий комплексное число $s - \gamma$, поворачивается по часовой стрелке на угол 2π при обходе точкой s кривой C , и, следовательно, аргумент комплексного числа $s - \gamma$ изменяется при этом на -2π .

Предположим теперь, что функция $\Phi(s)$ имеет внутри области, ограниченной кривой C , h нулей $\mu_1, \dots, \mu_h$ и l полюсов $\nu_1, \dots, \nu_l$, а на кривой C не имеет ни нулей, ни полюсов. В этом случае

$$\varphi = \Phi(s) = \frac{(s - \mu_1) \dots (s - \mu_h)}{(s - \nu_1) \dots (s - \nu_l)} \Psi(s), \quad (6.3.1)$$

где $\Psi(s)$ не имеет ни нулей, ни полюсов в области, ограниченной кривой C , и на самой кривой C . При обходе точкой s кривой C по ходу часовой стрелки аргумент числителя в (6.3.1) получает приращение $-2\pi h$, а аргумент знаменателя получает приращение $-2\pi l$. Следовательно, аргумент дроби в (6.3.1) изменится на $2\pi(l - h)$. Аргумент функции $\Psi(s)$ при этом не изменится. Таким образом, при обходе точкой s кривой C по часовой стрелке аргумент величины $\varphi = \Phi(s)$ изменяется на $2\pi(l - h)$. Это значит, что если $h > l$, то точка $\varphi = \Phi(s)$ обходит начало координат на плоскости φ по часовой стрелке $h - l$ раз, а если $h < l$, то точка $\varphi = \Phi(s)$ обходит начало координат против часовой стрелки $l - h$ раз. Таким образом, мы доказали, что если функция $\Phi(s)$ имеет внутри области, ограниченной замкнутой кривой C , h

нулей и l полюсов и не имеет ни нулей, ни полюсов на кривой C , то при обходе точкой z кривой C по часовой стрелке соответствующая точка $\varphi = \Phi(s)$ обходит начало координат по часовой стрелке $h - l$ раз.

Применим теперь доказанное предложение к передаточным функциям стационарных линейных систем с обратными связями. Рассмотрим стационарную линейную систему с передаточной функцией $\Phi(s)$, охваченную отрицательной жесткой обратной связью. Передаточная функция замкнутой системы, согласно § 4.6, определяется формулой

$$\Psi(s) = \frac{\Phi(s)}{1 + \Phi(s)}. \quad (6.3.2)$$

Предположим сначала, что разомкнутая система устойчива, т. е. что передаточная функция $\Phi(s)$ не имеет полюсов в правой

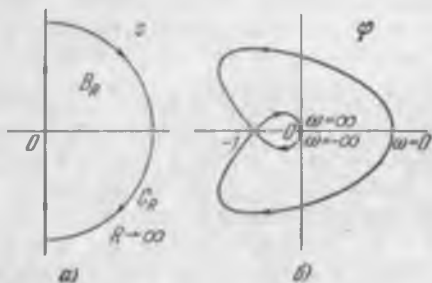


Рис. 6.3.2.

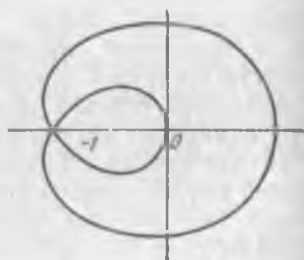


Рис. 6.3.3.

полуплоскости и на мнимой оси. Для устойчивости замкнутой системы необходимо и достаточно, чтобы знаменатель $1 + \Phi(s)$ дроби (6.3.2) не имел нулей в правой полуплоскости и на мнимой оси. Очевидно, что началу координат плоскости переменной $1 + \Phi(s)$ соответствует точка -1 плоскости переменной $\varphi = \Phi(s)$. Поэтому для определения числа нулей функции $1 + \Phi(s)$ можно пользоваться плоскостью переменной $\Phi(s)$, помня при этом, что роль начала координат в этом случае играет точка -1 . Возьмем на плоскости переменной s контур C_R , состоящий из отрезка мнимой оси $(-iR, iR)$ и полуокружности большого радиуса R (рис. 6.3.2). Так как при любом R функция $\Phi(s)$, а следовательно, и функция $1 + \Phi(s)$ не имеют полюсов на контуре C_R и в ограниченном им полукруге B_R , то при обходе точкой z этого контура точка $\Phi(s)$ обходит точку -1 по ходу часовой стрелки столько раз, сколько нулей имеет функция $1 + \Phi(s)$ в полукруге B_R . А так как это справедливо при любом R , то можно перейти к пределу при $R \rightarrow \infty$. Таким образом, число нулей функции $1 + \Phi(s)$ в правой полуплоскости равно числу обходов по часовой стрелке

амплитудно-фазовой характеристикой разомкнутой системы точки — 1.

Из сказанного следует, что в случае устойчивой разомкнутой системы для устойчивости замкнутой системы необходимо и достаточно, чтобы амплитудно-фазовая характеристика разомкнутой системы не охватывала точку —1 и не проходила через нее (рис. 6.3.2). В этом и состоит критерий Найквиста для случая устойчивой разомкнутой системы. На рис. 6.3.3 изображена амплитудно-фазовая характеристика разомкнутой системы в случае неустойчивой замкнутой системы. Амплитудно-фазовая характеристика разомкнутой системы в этом случае обходит точку —1 по часовой стрелке два раза. Следовательно, передаточная функция замкнутой системы имеет два полюса в правой полуплоскости.

Перейдем теперь к случаю неустойчивой разомкнутой системы. Предположим, что передаточная функция разомкнутой системы $\Phi(s)$ имеет k полюсов в правой полуплоскости и не имеет полюсов на мнимой осп. В этом случае функция $1 + \Phi(s)$ также имеет k полюсов в правой полуплоскости. Следовательно, чтобы функция $1 + \Phi(s)$ не имела нулей в правой полуплоскости, необходимо и достаточно, чтобы кривая $\Phi(i\omega)$ обходила точку —1 против часовой стрелки k раз.

Таким образом, если передаточная функция разомкнутой системы $\Phi(s)$ имеет k полюсов в правой полуплоскости, то для устойчивости замкнутой системы необходимо и достаточно, чтобы амплитудно-фазовая характеристика разомкнутой системы обходила точку —1 против часовой стрелки k раз.

Заметим теперь, что амплитудно-фазовые характеристики всех систем с действительными параметрами, а только такие системы и встречаются в задачах практики, симметричны относительно действительной осп. Поэтому достаточно рассматривать только одну половину амплитудно-фазовой характеристики, соответствующую изменению частоты ω от нуля до ∞ и считать полуобход амплитудно-фазовой характеристикой точки —1.

Чтобы дать общую формулировку критерия устойчивости Найквиста, условимся считать каждое пересечение амплитудно-фазовой характеристикой разомкнутой системы отрезка действительной осп ($-\infty, -1$) сверху вниз за $+1$ пересечение, а каждое пересечение снизу вверх — за -1 пересечение. Тогда общее число пересечений амплитудно-фазовой характеристикой разомкнутой системы отрезка ($-\infty, -1$) будет равно разности между числом пересечений сверху вниз и числом пересечений снизу вверх. При этом условии каждому обходу амплитудно-фазовой характеристикой разомкнутой системы точки —1 против часовой стрелки соответствует одно пересечение отрезка действительной осп ($-\infty, -1$). Если амплитудно-фазовая характеристика разомкну-

той системы при $\omega = 0$ начинается на отрезке $(-\infty, -1)$, то этому будет соответствовать $+1/2$ или $-1/2$ пересечения отрезка $(-\infty, -1)$ в зависимости от того, вниз или вверх от этого отрезка идет амплитудно-фазовая характеристика при возрастании ω . Мы можем теперь сформулировать найденное общее условие

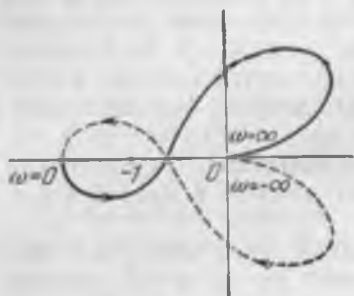


Рис. 6.3.4.

устойчивости замкнутой системы следующим образом. Если передаточная функция $\Phi(s)$ разомкнутой системы имеет k полюсов в правой полуплоскости, то для устойчивости замкнутой системы необходимо и достаточно, чтобы амплитудно-фазовая характеристика разомкнутой системы в диапазоне положительных частот пересекала отрезок действительной оси $(-\infty, -1)$ $k/2$ раз. Для иллюстрации на рис. 6.3.4 показана амплитудно-фазовая характеристика разомкнутой системы для случая, когда

передаточная функция разомкнутой системы $\Phi(s)$ имеет один полюс в правой полуплоскости и замкнутая система устойчива. На рис. 6.3.5 и 6.3.6 показаны амплитудно-фазовые характеристики разомкнутых систем для случая, когда передаточные функции разомкнутых

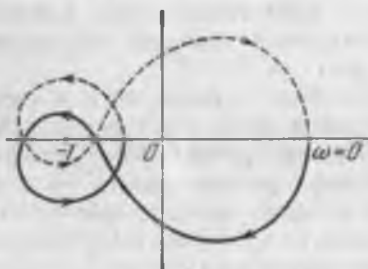


Рис. 6.3.5.

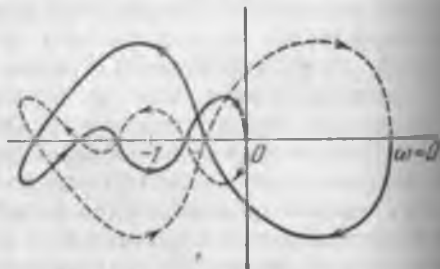


Рис. 6.3.6.

систем имеют по два полюса в правой полуплоскости и замкнутые системы устойчивы.

Если передаточная функция разомкнутой системы имеет полюсы на мнимой оси, то контур в плоскости s следует деформировать таким образом, чтобы обойти полюсы, лежащие на мнимой оси, по полуокружностям малого радиуса. Так как разомкнутая система в этом случае неустойчива, то мы условимся обходить эти полюсы слева, т. е. относить их к правой полуплоскости (рис. 6.3.7, а). При обходе слева полюса iv , расположенного на мнимой оси, аргумент комплексного числа $s - iv$ уменьшается

на π и, следовательно, аргумент дроби $1/(s - iv)$ увеличивается на π . Вследствие этого аргумент передаточной функции разомкнутой системы $\Phi(s)$ также увеличивается на π . Иными словами, обходу слева полюса передаточной функции разомкнутой системы, расположенного на мнимой оси, по полуокружности малого радиуса соответствует обход начала координат в плоскости $\Phi(s)$ по кривой, близкой к полуокружности большого радиуса, против

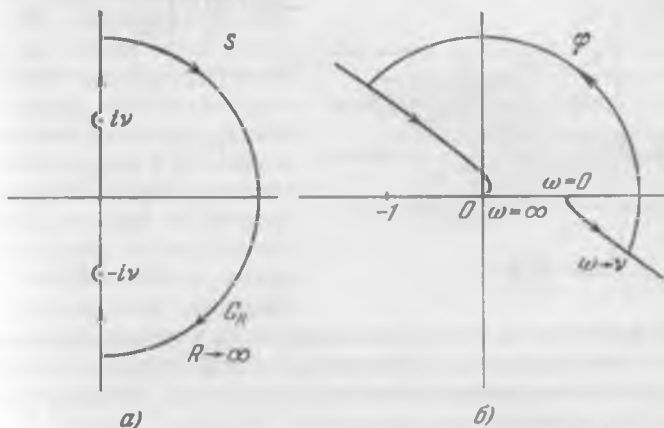


Рис. 6.3.7.

часовой стрелки (рис. 6.3.7, б). После добавления такого обхода к амплитудно-фазовой характеристике разомкнутой системы для каждого полюса, расположенного на мнимой оси, сформулированный общий критерий устойчивости Найквиста будет применим и к этому случаю *).

Посмотрим теперь, как можно подсчитать число пересечений отрезка $(-\infty, -1)$ амплитудно-фазовой характеристикой разомкнутой системы по логарифмическим частотным характеристикам. Так как фазовая характеристика системы определена только с точностью до периода 2π , то, чтобы сделать ее однозначной, условимся считать ее значения заключенными в интервале $(-2\pi, 0]$. Тогда пересечению амплитудно-фазовой характеристикой разомкнутой системы отрезка $(-\infty, -1)$ сверху вниз будет соответствовать пересечение фазовой характеристикой уровня $-\pi$ снизу вверх при положительной логарифмической амплитудной характеристике. Следовательно, пересечение фазовой характеристикой уровня $-\pi$ снизу вверх при положительной логарифмической амплитудной характеристике следует считать за $+1$ пересечение,

*) При этом каждый кратный полюс следует считать за соответствующее число полюсов.

а пересечение уровня $-\pi$ сверху вниз при положительной логарифмической амплитудной характеристике следует считать за -1 пересечение.

Если передаточная функция разомкнутой системы имеет полюс на мнимой оси iv , то, устремив к нулю радиус полуокружности,

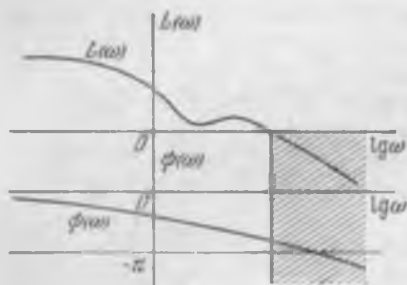


Рис. 6.3.8.

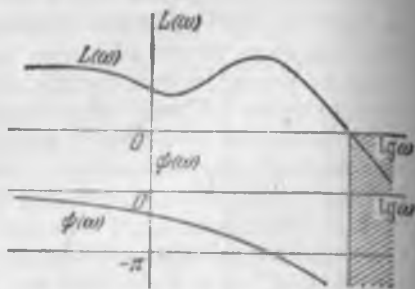


Рис. 6.3.9.

по которой обходится этот полюс, приходим к выводу, что логарифмическая амплитудная характеристика стремится к бесконечности, а фазовая характеристика увеличивается скачком на π при

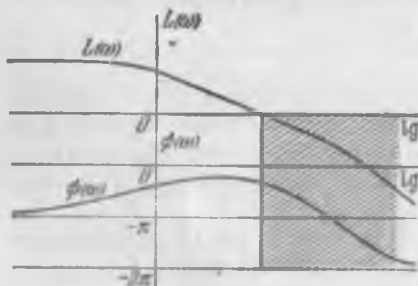


Рис. 6.3.10.

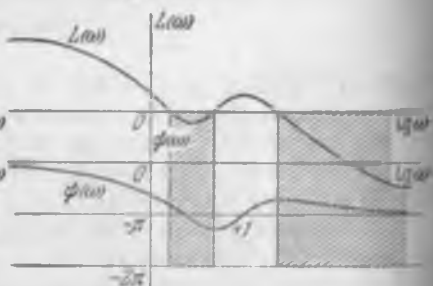


Рис. 6.3.11.

$\omega = v$. При этом перескок через уровень $-\pi$ также следует считать за пересечение уровня $-\pi$. Считая такие перескоки по логарифмическим частотным характеристикам, следует не забывать и полюсы, расположенные в начале координат, которым соответствует значение $\lg \omega = -\infty$. Если фазовая характеристика асимптотически стремится к уровню $-\pi/2$ при $\lg \omega \rightarrow -\infty$, где l — остаток от деления кратности полюса в начале координат на 4, то этот уровень следует считать достигнутым после скачкообразного увеличения фазовой характеристики на величину $\pi/2$, умножаемую на кратность полюса, и, исходя из этого, определять

число перескоков фазовой характеристики через уровень $-\pi$ при $\lg \omega = -\infty$.

Таким образом, если передаточная функция разомкнутой системы $\Phi(s)$ имеет k полюсов в правой полуплоскости и на мнимой оси, то для устойчивости замкнутой системы необходимо и достаточно, чтобы число пересечений логарифмической фазовой характеристикой разомкнутой системы уровня $-\pi$ снизу вверх при положительной логарифмической амплитудной характеристике превышало число пересечений сверху вниз на $k/2$. В частности, если разомкнутая система устойчива, то для устойчивости замкнутой системы необходимо и достаточно, чтобы число пересечений логарифмической фазовой характеристикой уровня $-\pi$ снизу вверх при положительной логарифмической амплитудной характеристике было равно числу пересечений сверху вниз. На рис. 6.3.8, 6.3.9, 6.3.10, 6.3.11 и 6.3.12 показаны логарифмические частотные характеристики для случаев, для которых амплитудно-фазовые

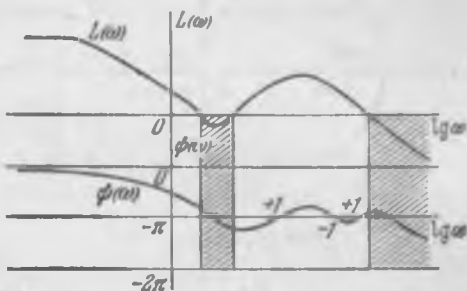


Рис. 6.3.12.

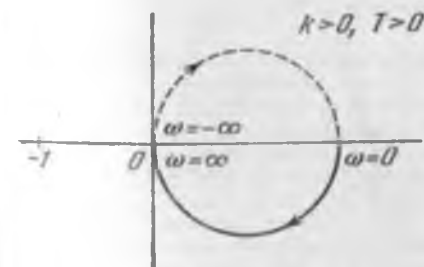


Рис. 6.3.13.

характеристики соответственно на рис. 6.3.2, 6.3.3, 6.3.4, 6.3.5 и 6.3.6.

Пример 6.3.1. Рассмотрим случай передаточной функции разомкнутой системы

$$\Phi(s) = \frac{k}{Ts + 1}.$$

При любых положительных k , T разомкнутая система устойчива. Ее амплитудно-фазовая характеристика представляет собой окружность, половина которой показана на рис. 6.3.13.

Рис. 6.3.13 показывает, что замкнутая система в этом случае всегда устойчива. При отрицательном T разомкнутая система неустойчива. Ее передаточная функция имеет один полюс $-1/T$ в правой полуплоскости. Следовательно, для устойчивости замкнутой системы в этом случае необходимо, чтобы число пересечений амплитудно-фазовой характеристикой разомкнутой системы отрезка $(-\infty, -1)$ действительной оси было равно $1/2$. На рис. 6.3.14 показана амплитудно-фазовая характеристика разомкнутой системы в случае, когда $T < 0$, $k < -1$. Так как число пересечений амплитудно-фазовой характеристикой отрезка $(-\infty, -1)$ действительной оси в этом случае равно $1/2$, то замкнутая система устойчива. На рис. 6.3.15 показана амплитудно-фазовая характеристика разомкнутой

системы в случае, когда $T < 0$, $-1 < k < 0$. В этом случае амплитудно-фазовая характеристика не пересекает отрезок $(-\infty, -1)$. Следовательно, замкнутая система в этом случае неустойчива.

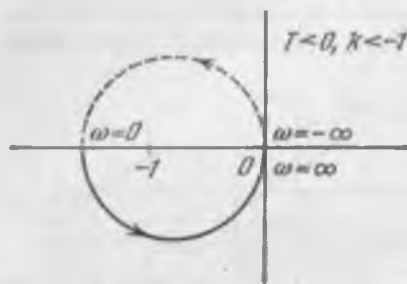


Рис. 6.3.14.

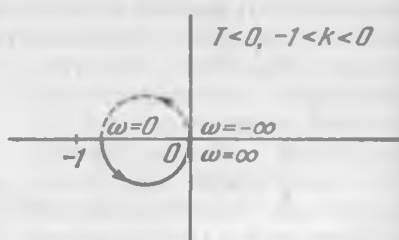


Рис. 6.3.15.

Пример 6.3.2. Передаточная функция разомкнутой системы

$$\Phi(s) = \frac{k}{s(s^2 + 2as + b^2)}, \quad k > 0,$$

при $a > 0$ имеет один полюс на мнимой оси. Следовательно, для устойчивости замкнутой системы в этом случае необходимо, чтобы число пересечений

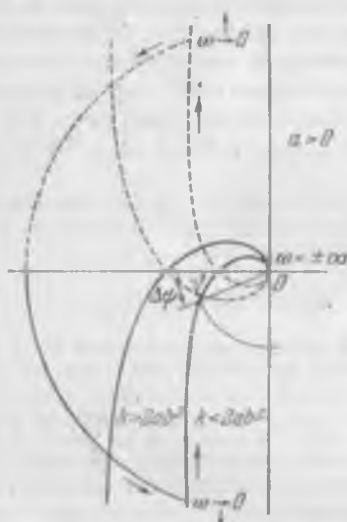


Рис. 6.3.16.

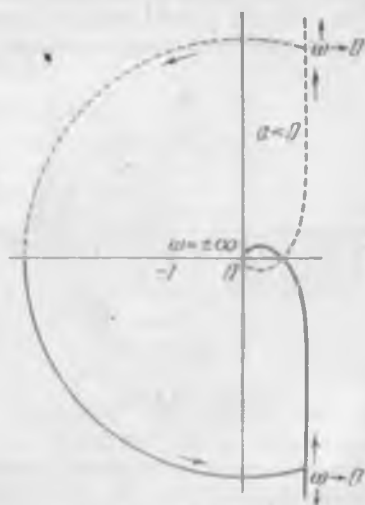


Рис. 6.3.17.

амплитудно-фазовой характеристикой разомкнутой системы отрезка $(-\infty, -1)$ действительной оси было равно $1/2$. При $k < 2ab^2$ амплитудно-фазовая характеристика разомкнутой системы не пересекает отрезок $(-\infty, -1)$. Дополняя ее согласно изложенному полуокружностью большего радиуса, обходящей

начало координат против часовой стрелки, видим, что часть этой полуокружности, соответствующая положительным частотам, начинается на отрезке $(-\infty, -1)$ и идет вниз от него (сплошная линия на рис. 6.3.16). Следовательно, число пересечений амплитудно-фазовой характеристикой разомкнутой системы, дополненной полуокружностью большого радиуса, отрезка $(-\infty, -1)$ сверху равно $1/2$, и замкнутая система в этом случае устойчива. При $k > 2ab^2$ число пересечений отрезка $(-\infty, -1)$ равно $-1/2$, и следовательно, замкнутая система неустойчива. При $a < 0$ передаточная функция

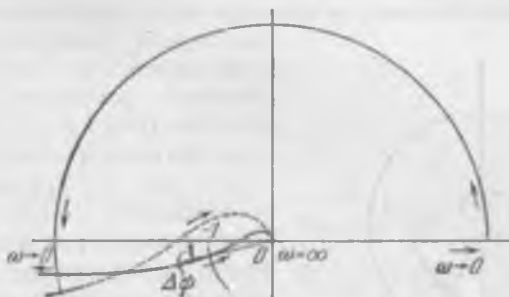


Рис. 6.3.18.

разомкнутой системы имеет один полюс на мнимой оси и два полюса в правой полуплоскости. Рис. 6.3.17 показывает, что в этом случае замкнутая система всегда неустойчива.

Пример 6.3.3. Передаточная функция разомкнутой системы

$$G(s) = \frac{k(1 + \tau s)}{s^2(1 + 2\zeta Ts + T^2 s^2)}$$

при положительных ζ , T имеет один двойной полюс на мнимой оси. Следовательно, для устойчивости замкнутой системы в этом случае необходимо, чтобы амплитудно-фазовая характеристика разомкнутой системы пересекала отрезок $(-\infty, -1)$ в положительном диапазоне частот один раз. На рис. 6.3.18 показаны амплитудно-фазовые характеристики разомкнутой системы, дополненные окружностью большого радиуса, обходящей начало координат против часовой стрелки, для двух комбинаций значений параметров системы. Сплошная кривая пересекает отрезок $(-\infty, -1)$ один раз и, следовательно, соответствует устойчивой замкнутой системе. Пунктирная кривая пересекает отрезок $(-\infty, -1)$ нуль раз и, следовательно, соответствует неустойчивой замкнутой системе.

Мы видим, что метод частотных характеристик дает возможность сравнительно просто исследовать устойчивость любых стационарных линейных систем. Приведенные примеры показывают, что по частотным характеристикам можно не только судить об устойчивости стационарных линейных систем, но и подбирать некоторые их параметры так, чтобы обеспечить устойчивость. Предоставляем читателю самостоятельно проверить, что результаты рассмотренных примеров совпадают с результатами исследования устойчивости тех же систем по критериям Рауса и Гурвица. Для рассмотренных простейших примеров это нетрудно

сделать. Для более сложных систем такая проверка потребовала бы большого труда.

Совокупность значений параметров системы, при которых система устойчива, называется *областью устойчивости* системы. Совокупность значений параметров системы, при которых система неустойчива, называется *областью неустойчивости*. Совокупность значений параметров, при которых совершается переход из области устойчивости в область неустойчивости и наоборот,



Рис. 6.3.19.

называется *границей области устойчивости*. В примере 6.3.2 областью устойчивости является отрезок $0 < k < 2ab^2$, областью неустойчивости — отрезок $2ab^2 < k < \infty$, а границей области устойчивости — точка $k = 2ab^2$. При проектировании автоматических систем необходимо не только обеспечить устойчивость, но и выбрать параметры системы в области устойчивости достаточно далеко от границы области устойчивости. Если параметры системы близки к границе области устойчивости, то достаточно небольших измене-

ний этих параметров, чтобы система стала неустойчивой. При изготовлении элементов автоматических систем невозможно обеспечить заданные параметры элементов абсолютно точно. Параметры элементов всегда имеют некоторый разброс в пределах допусков производства. Вследствие этого, если выбрать параметры системы близко к границе устойчивости, то может случиться, что одни экземпляры спроектированной системы будут устойчивыми, а другие неустойчивыми. Кроме того, в этом случае качество системы будет невысоким. Чтобы избежать этого, и следует выбирать параметры системы подальше от границы области устойчивости.

Для суждения о степени близости системы к границе области устойчивости обычно пользуются запасами устойчивости. *Запасом устойчивости по фазе* или *избытком фазы* устойчивой замкнутой системы называется увеличение запаздывания по фазе на частоте среза (т. е. при значении частоты ω , при котором амплитудно-фазовая характеристика разомкнутой системы входит внутрь круга единичного радиуса и в дальнейшем не выходит из него), при котором система доходит до границы области устойчивости. Из рис. 6.3.19 видно, что запас устойчивости по фазе определяется формулой

$$\Delta\varphi = \pi - |\varphi(\omega_c)| \quad (-\pi < \varphi(\omega_c) < 0). \quad (6.3.3)$$

Запасом устойчивости по амплитуде называется относительное увеличение коэффициента усиления разомкнутой системы, при котором устойчивая замкнутая система доходит до границы области устойчивости (рис. 6.3.19). Запас устойчивости по амплитуде обычно измеряется в децибелах. В этом случае запас устойчивости по амплитуде равен абсолютной величине отрицательной ординаты логарифмической амплитудной характеристики $L(\omega)$ при той частоте $\omega = \omega_\pi$, при которой запаздывание по фазе достигает $-\pi$.

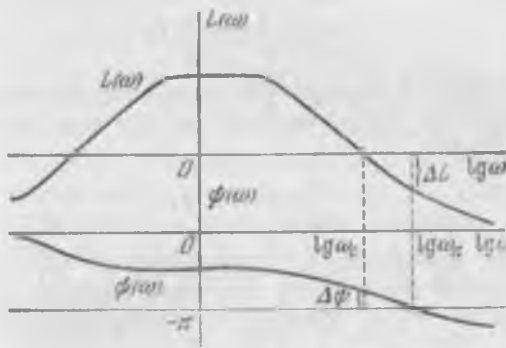


Рис. 6.3.20.

На рис. 6.3.20 показано, как определяются запасы устойчивости по амплитуде и фазе по логарифмическим частотным характеристикам.

При проектировании автоматических систем рекомендуется выбирать запас устойчивости по фазе не меньше 30° , а в случае возможности — не меньше 45° . Запас устойчивости по амплитуде рекомендуется брать не меньше 6 дБ, чему соответствует допустимое увеличение коэффициента усиления разомкнутой системы без выхода из области устойчивости приблизительно вдвое.

Пример 6.3.4. В условиях примера 6.3.2 запасы устойчивости по амплитуде и фазе равны соответственно приблизительно 3,3 дБ и $23,6^\circ$ (рис. 6.3.16). В условиях примера 6.3.3 запасы устойчивости по амплитуде и фазе приблизительно равны соответственно 6 дБ и 14° (рис. 6.3.18).

§ 6.4. Устойчивость дискретных линейных систем

Общий критерий устойчивости (6.1.1) применим к любым линейным системам, в частности к дискретным. Очевидно, что при исследовании устойчивости дискретной системы достаточно рассматривать только дискретную последовательность моментов времени (моментов съема информации). Согласно изложенному в § 5.2 весовая функция дискретной линейной системы

представляет собой линейную комбинацию δ -функций:

$$g(t_l, \tau) = \sum_{k=-\infty}^l g_{lk} \delta(\tau - t_k), \quad (6.4.1)$$

где g_{lk} — весовые коэффициенты системы. Подставляя выражение (6.4.1) в (6.1.1), получим общее необходимое и достаточное условие устойчивости дискретной линейной системы:

$$\lim_{l \rightarrow \infty} \sum_{k=k_0}^l |g_{lk}| < c. \quad (6.4.2)$$

Для стационарной дискретной линейной системы $g_{lk} = w_{l-k}$ и заменой индекса суммирования $m = l - k$ условие (6.4.2) приводится к виду

$$\sum_{m=0}^{\infty} |w_m| < c. \quad (6.4.3)$$

Таким образом, необходимым и достаточным условием устойчивости стационарной дискретной линейной системы является абсолютная сходимость ряда, членами которого являются ее весовые коэффициенты.

Передаточная функция стационарной линейной дискретной системы, согласно (5.3.14), определяется формулой

$$\Psi(z) = \sum_{m=0}^{\infty} w_m z^{-m}. \quad (6.4.4)$$

Из сходимости ряда (6.4.3) следует, что ряд (6.4.4) сходится при $z = 1$. Но в таком случае он сходится и сумма его конечна и при любом z^{-1} , по модулю меньшем единицы, т. е. при любом z , по модулю превышающем единицу. Следовательно, передаточная функция устойчивой стационарной дискретной линейной системы конечна всюду вне единичного круга плоскости комплексной переменной z с центром в начале координат.

Таким образом, для устойчивости стационарной дискретной линейной системы необходимо и достаточно, чтобы все полюсы (и другие особые точки) ее передаточной функции $\Psi(z)$ лежали внутри круга единичного радиуса с центром в начале координат.

Формула (5.3.16) показывает, что единичному кругу плоскости переменной z соответствует левая полуплоскость переменной $v = (z - 1)/(z + 1)$, а окружности единичного радиуса плоскости z соответствует мнимая ось переменной v . Следовательно, для устойчивости стационарной дискретной линейной системы необходимо и достаточно, чтобы все полюсы ее передаточной функции $\Omega(v) = \Psi((1+v)/(1-v))$, рассматриваемой как функция комплексной переменной v , лежали в левой полуплоскости. Отсюда сле-

дует, что к стационарным дискретным линейным системам полностью применимы критерий Найквиста и изложенные в предыдущем параграфе частотные методы исследования устойчивости. При этом переменная v имеет чисто мнимые значения $v = i\lambda$, а роль частоты играет величина λ , связанная с частотой ω очевидным соотношением

$$\lambda = \frac{e^{i\omega T_n} - 1}{i(e^{i\omega T_n} + 1)}, \quad (6.4.5)$$

вытекающим из (5.3.16) при $s = i\omega$, $v = i\lambda$.

Заметим, что критерий Найквиста и частотные методы исследования устойчивости применимы также и в том случае, если пользоваться непосредственно частотными характеристиками стационарных дискретных линейных систем и плоскостью параметра s , так как единичному кругу плоскости $z = e^{sT_n}$ соответствует часть левой полуплоскости переменной s , лежащая между двумя горизонтальными прямыми $\text{Im}\{s\} = \omega_0$ и $\text{Im}\{s\} = \omega_0 + 2\pi/T_n$ при любом ω_0 . При этом одному обходу единичной окружности точкой z соответствует перемещение точки s из точки $i\omega_0$ в точку $i\omega_0 + 2\pi i/T_n$. Следовательно, для устойчивости стационарной дискретной линейной системы необходимо и достаточно, чтобы все полюсы ее передаточной функции $\Phi(s) = \Psi(e^{sT_n})$, рассматриваемой как функция комплексной переменной s , лежали в левой полуплоскости. При этом вследствие периодичности передаточной функции достаточно брать значения частоты ω в любом интервале длины $2\pi/T_n$.

§ 6.5. Переходные процессы в линейных системах

Рассмотрим стационарную линейную систему, поведение которой описывается дифференциальным уравнением (6.2.1). Общий интеграл уравнения (6.2.1) выражается формулой

$$y(t) = c_1 e^{v_1 t} + c_2 e^{v_2 t} + \dots + c_n e^{v_n t} + \int_{t_0}^t w(t-\tau) x(\tau) d\tau, \quad (6.5.1)$$

где $v_1, \dots, v_n$ — корни характеристического уравнения, предполагаемые для простоты различными. Формула (6.5.1) показывает, что общий интеграл уравнения (6.2.1) представляет собой сумму его частного интеграла, соответствующего действию входного сигнала $x(t)$:

$$y_n(t) = \int_{t_0}^t w(t-\tau) x(\tau) d\tau, \quad (6.5.2)$$

и общего интеграла соответствующего однородного уравнения:

$$y_{св}(t) = c_1 e^{\nu_1 t} + c_2 e^{\nu_2 t} + \dots + c_n e^{\nu_n t}, \quad (6.5.3)$$

где постоянные $c_1, \dots, c_n$ определяются начальными значениями $y(t_0)$, $y'(t_0)$, $\dots$, $y^{(n-1)}(t_0)$ выходной переменной системы y и ее производных $y', \dots, y^{(n-1)}$.

Формула (6.5.2) определяет движение системы, вызываемое действием только одной функции $x(t)$. Оно отсутствует, если входная функция системы $x(t)$ тождественно равна нулю *). Мы будем называть это движение *невозмущенным* движением системы.

Формула (6.5.3) определяет движение системы при отсутствии входной переменной, вызываемое только начальными отклонениями от состояния равновесия. В теории устойчивости начальные отклонения системы от состояния равновесия обычно называются *начальными возмущениями*. Движение системы, определяемое формулой (6.5.3), называется *переходным процессом* или *возмущением*. Формула (6.5.1) определяет *возмущенное* движение системы, состоящее из невозмущенного движения и наложенного на него переходного процесса.

Из доказанного в § 6.2 следует, что стационарная линейная система устойчива тогда и только тогда, когда переходный процесс затухает при $t \rightarrow \infty$. Иными словами, стационарная линейная система устойчива, если при любых начальных возмущениях отклонение возмущенного движения от невозмущенного стремится к нулю при $t \rightarrow \infty$. При этом отклонение возмущенного движения от невозмущенного будет сколь угодно малым, если начальные возмущения достаточно малы. Эти свойства устойчивой системы были положены в основу определения устойчивости А. М. Ляпуновым, который в 1892 г. впервые создал точную и строгую теорию устойчивости любых систем, поведение которых описывается обыкновенными дифференциальными уравнениями [42]. До Ляпунова существовали только методы исследования устойчивости стационарных линейных систем, описываемых обыкновенными дифференциальными уравнениями, разработанные Раусом и Гурвицем. Кроме того, классической работе Ляпунова по теории устойчивости предшествовала интересная работа Н. Е. Жуковского [24].

А. М. Ляпунов дал следующее определение устойчивости любой системы, поведение которой описывается обыкновенными дифференциальными уравнениями: система называется устойчивой, если при всех $t > t_0$ отклонение возмущенного движения от невозмущенного сколь угодно мало при достаточно малых начальных

*) Мы говорим здесь не о механическом движении, а о движении в обобщенном смысле, понимая под словом *движение* любой процесс изменения состояния системы.

возмущениях в момент t_0 ; система называется асимптотически устойчивой, если отклонение возмущенного движения от невозмущенного стремится к нулю при $t \rightarrow \infty$.

Таким образом, стационарная линейная система устойчива в смысле определения § 6.1 тогда и только тогда, когда она асимптотически устойчива по Ляпунову. Это распространяется на любые линейные системы и, как мы увидим в § 10.1, и на нелинейные системы, описываемые обыкновенными дифференциальными уравнениями. Таким образом, определение устойчивости, данное в § 6.1, равноценно определению асимптотической устойчивости по Ляпунову.

Так как практически устойчивыми можно считать только асимптотически устойчивые системы, то в дальнейшем, говоря об устойчивости, мы всегда будем иметь в виду асимптотическую устойчивость по Ляпунову или устойчивость в смысле определения, данного в § 6.1.

За невозмущенное движение системы можно взять любое ее движение, т. е. интеграл (6.5.1) уравнения (6.2.1), соответствующий любым фиксированным значениям постоянных $c_1, \dots, c_n$. Тогда возмущенным будет любое движение системы, соответствующее другим значениям постоянных $c_1, \dots, c_n$.

Очевидно, что для устойчивой линейной системы при любом выборе невозмущенного движения возмущенное движение всегда стремится к невозмущенному при $t \rightarrow \infty$.

С практической точки зрения важна не только устойчивость системы, но и характер затухания ее переходных процессов. Если переходный процесс длится долго и система совершает большие колебания, то такая система хотя и устойчива, но не может считаться хорошей. Поэтому кроме устойчивости необходимо исследовать также и качество системы.

Качество линейной системы принято оценивать по виду переходного процесса. При этом за стандартный переходный процесс для стационарной линейной системы принимают ее реакцию на единичную ступенчатую функцию $1(t - \tau)$, называемую обычно *переходной функцией*. Так как $1(t - \tau)$ есть интеграл от $\delta(t - \tau)$, то переходную функцию данной линейной системы можно определить как весовую функцию последовательного соединения интегрирующего звена и данной системы. Тогда, пользуясь формулой (4.2.5) и имея в виду, что весовой функцией интегрирующего звена является $1(t - \tau)$, получим следующее выражение переходной функции $h(t - \tau)$ стационарной линейной системы через ее весовую функцию $w(t - \tau)$:

$$h(t - \tau) = \int_{\tau}^t w(t - \sigma) 1(\sigma - \tau) d\sigma = \int_{\tau}^t w(t - \sigma) d\sigma = \int_0^{t-\tau} w(\xi) d\xi. \quad (6.5.4)$$

Полагая здесь $\tau = 0$ и дифференцируя по t , получим выражение весовой функции стационарной линейной системы через ее переходную функцию:

$$\dot{h}(t) = h'(t). \quad (6.5.5)$$

Подставляя выражение (6.2.2) весовой функции стационарной линейной системы в (6.5.4), приходим к заключению, что разность между переходной функцией устойчивой стационарной линейной системы и ее установившимся значением при $t \rightarrow \infty$ выражается формулой (6.5.3) при некоторых значениях постоянных $c_1, \dots, c_n$,

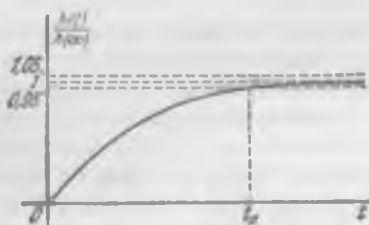


Рис. 6.5.1.

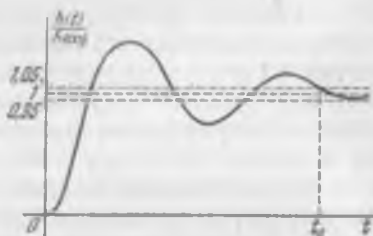


Рис. 6.5.2.

т. е. является одним из возможных переходных процессов системы. Поэтому она и может служить стандартом переходного процесса.

Подставляя в (6.5.4) выражение (2.4.32) весовой функции стационарной линейной системы через ее действительную частотную характеристику $P(\omega)$, выполняя интегрирование по ξ и принимая во внимание, что для всех систем с действительными весовыми функциями $P(\omega)$ является четной функцией, получим

$$h(t) = \frac{1}{\pi} \int_{-\infty}^{\infty} \frac{P(\omega) \sin \omega t}{\omega} d\omega = \frac{2}{\pi} \int_0^{\infty} \frac{P(\omega) \sin \omega t}{\omega} d\omega. \quad (6.5.6)$$

Этой формулой обычно пользуются для приближенного определения переходной функции системы по ее частотной характеристике. При этом для вычисления интеграла удобно аппроксимировать кривую $P(\omega)/\omega$ ломаной по возможности с небольшим числом отрезков, начиная от некоторого достаточно малого значения ω_0 . В интервале же $0 \leq \omega < \omega_0$ следует вынести $P(\omega)$ за знак интеграла средним значением и положить $\sin \omega t \approx \omega t$.

Типичный характер переходных функций стационарных линейных систем показан на рис. 6.5.1 и 6.5.2. Значение $h(\infty)$ переходной функции представляет собой ее установившееся значение. Время t_n от начала переходного процесса до момента, когда разность $h(t) - h(\infty)$ становится по абсолютной величине меньше, чем $0,05h(\infty)$, обычно называется *временем переходного процесса*.

Если в течение времени переходного процесса t_n функция $h(t) - h(\infty)$ изменяет знак, то каждый максимум абсолютной величины этой разности называется *перерегулированием*. Число экстремумов разности $h(t) - h(\infty)$ в течение времени переходного процесса t_n называется *числом перерегулирований*. Обычно принято считать переходный процесс удовлетворительным, если число перерегулирований за время переходного процесса t_n не превышает двух. Требования к величине перерегулирования могут быть различными в зависимости от конкретного характера системы. В некоторых случаях перерегулирование вообще не допускается.

Понятие переходной функции можно обобщить и на нестационарные линейные системы. Пользуясь формулой (4.2.5), получим следующее выражение переходной функции линейной системы $h(t, \tau)$ через ее весовую функцию $g(t, \tau)$:

$$h(t, \tau) = \int_{\tau}^t g(t, \sigma) 1(\sigma - \tau) d\sigma = \int_{\tau}^t g(t, \sigma) d\sigma. \quad (6.5.7)$$

Дифференцируя эту формулу по τ , получим выражение весовой функции линейной системы через ее переходную функцию:

$$g(t, \tau) = -\frac{\partial h(t, \tau)}{\partial \tau} \quad (6.5.8)$$

Формулы (6.5.4) и (6.5.5) являются частными случаями формул (6.5.7) и (6.5.8) соответственно.

§ 6.6. Синтез линейных систем методом частотных характеристик

Мы видели в §§ 4.6, 4.8 и 6.3, что метод частотных характеристик очень удобен для исследования стационарных линейных систем, особенно если пользоваться логарифмическими частотными характеристиками. Благодаря своей простоте метод логарифмических частотных характеристик оказывается также простейшим методом синтеза стационарных линейных систем. В § 4.8 было отмечено, что для широкого класса минимально-фазовых систем фазовая характеристика полностью определяется амплитудной характеристикой. А так как все устойчивые стационарные линейные звенья являются минимально-фазовыми, то практически и все проектируемые стационарные линейные системы оказываются минимально-фазовыми и, следовательно, полностью характеризуются своими амплитудными частотными характеристиками. Это сильно упрощает задачу синтеза стационарных линейных систем.

В конце § 4.8 было отмечено, что если логарифмическая амплитудная характеристика минимально-фазовой системы имеет достаточно длинный участок с наклоном 20 дБ/дек, то ее фазовая

характеристика в середине этого участка близка к $1\pi/2$. Отсюда следует, что при проектировании стационарной линейной системы следует стремиться к тому, чтобы частота среза разомкнутой системы ω_c приходилась примерно на середину достаточно длинного участка ее логарифмической амплитудной характеристики с наклоном -20 дБ/дек. Это обеспечит запас устойчивости по фазе, близкий к $\pi/2$, и достаточный запас устойчивости по амплитуде замкнутой системы. Достичь этого можно соответствующим выбором корректирующих устройств и их параметров.

Чтобы облегчить определение основных характеристик проектируемых систем, составлены специальные номограммы для систем

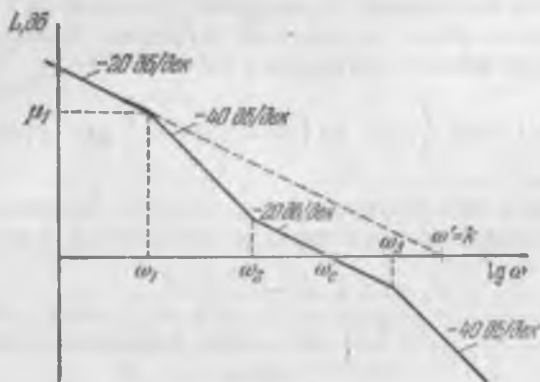


Рис. 6.6.1.

с типовыми асимптотическими характеристиками соответствующих разомкнутых систем. В приложении 3 даны такие номограммы для разомкнутых систем с одним интегратором, имеющих типовую асимптотическую амплитудную характеристику, показанную на рис. 6.6.1. Передаточная функция такой разомкнутой системы определяется формулой

$$\Phi(s) = \frac{k(T_2 s + 1)}{s(T_1 s + 1)(T_3 s + 1)} = \frac{kT_2}{T_1 T_3} \frac{s + \omega_2}{(s + \omega_1)(s + \omega_3)}. \quad (6.6.1)$$

Номограммы дают возможность определять следующие динамические характеристики замкнутых систем: время переходного процесса t_n ; максимум переходной функции h_m (рис. 6.6.2); время t_1 достижения этого максимума; частоту ω_1 колебаний переходной функции; максимум амплитудной характеристики A_m ; частоту ω_m , при которой достигается этот максимум. Для этого на номограммах представлены зависимости величин h_m , A_m , $\omega_c t_n/10$, $\omega_c t_1/10$, ω_1/ω_c и ω_m/ω_c от отношений ω_1/ω_c , ω_3/ω_c и ординаты асимптотической характеристики μ_1 в первой точке излома (при $\omega = \omega_1$). В приложении 3 эти графики даны для четырех значений

ω_3/ω_c : 1, 2, 4, ∞ . Для каждого значения ω_3/ω_c даны две серии графиков. На верхней части номограммы даны графики зависимости h_m и A_m от ω_1/ω_c для пяти значений μ_1 : 20, 30, 40, 60 и 80 дБ. На нижней части номограммы даны графики зависимости величин $\omega_c t_p/10$, $\omega_c t_i/10$, ω_1/ω_c и ω_m/ω_c от ω_1/ω_c для тех же значений μ_1 . Подобные номограммы имеются также для типовой разомкнутой системы с двумя интеграторами [31, 47].

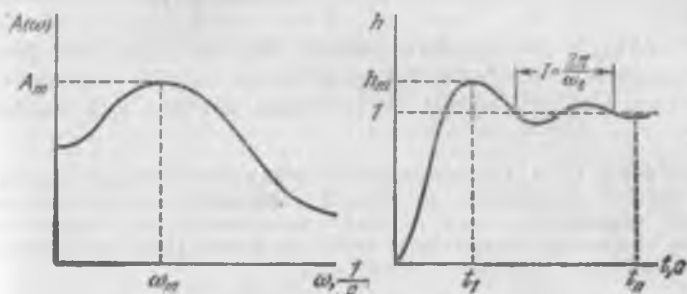


Рис. 6.6.2.

Очевидно, что ордината μ_1 асимптотической характеристики в первой точке ее излома связана с сопрягающими частотами ω_1 , ω_2 и частотой среза ω_c соотношением

$$\lg \frac{\omega_2}{\omega_c} = \frac{\mu_1}{20} + 2 \lg \frac{\omega_1}{\omega_c}. \quad (6.6.2)$$

Поэтому задание величин μ_1 и ω_1/ω_c определяет также величину ω_2/ω_c .

Применение номограмм основных динамических характеристик стационарных линейных систем с соответствующими типовыми асимптотическими характеристиками разомкнутых систем дает возможность свести проектирование системы управления для стационарного линейного объекта к следующим простым операциям.

1. Задаются приемлемыми динамическими характеристиками проектируемой системы (t_p , h_m , t_i и др.).

2. По этим характеристикам по номограммам определяют допустимые значения параметров типовой амплитудной характеристики μ_1 , ω_1/ω_c и ω_3/ω_c . При этом, чтобы обеспечить достаточное удаление частоты среза ω_c от концов интервала частот с наклоном амплитудной характеристики — 20 дБ/дек, рекомендуется выбирать величину ω_3/ω_c от 2 до 4, а μ_1 выбирать так, чтобы отношение ω_3/ω_2 было порядка 10.

3. Руководствуясь соображениями технического порядка, выбирают необходимые элементы системы управления: измерители, усилители, исполнительные устройства и другие необходимые элементы и задают допустимые параметры этих элементов.

Для полученной таким путем разомкнутой системы строится амплитудная характеристика.

4. Если полученную характеристику не удастся приблизить к типовой путем выбора параметров элементов системы управления (что, как правило, и бывает на практике), то стремятся достичь этого путем ввода соответствующих корректирующих звеньев. При этом, естественно, стремятся достичь наилучшего приближения с помощью возможно более простых корректирующих устройств.

5. После окончательного выбора корректирующих устройств и их параметров определяют динамические характеристики спроектированной системы путем поверочного расчета или моделирования.

Пример 6.6.1. Применим изложенный способ к задаче синтеза одного канала системы управления подвижной пушечной установкой на самолете. Объектом управления здесь служит электродвигатель, осуществляющий слежение установкой за выходным сигналом прицельного устройства. Передаточная функция двигателя имеет вид

$$\Phi_d(s) = \frac{k_d}{s(T_d s + 1)} \quad (6.6.3)$$

Примем для определенности $k_d = 1 \text{ с}^{-1}$, $T_d = 0,5 \text{ с}$.

Задаемся приемлемыми характеристиками системы: $t_{\Pi} \approx 0,6 \text{ с}$, $h_m \leq 1,2$ — и выбираем согласно рекомендации $\omega_3/\omega_c = 4$, $\omega_3/\omega_2 = 10$. Кроме того, примем для упрощения $\omega_1 = 1/T_d = 2 \text{ с}^{-1}$. По верхней номограмме (приложение 3) для $\omega_3/\omega_c = 4$ находим для разных значений μ_1 приемлемые значения отношения ω_1/ω_c , при которых $h_m \leq 1,2$, а по нижней номограмме находим соответствующие значения $\omega_c t_{\Pi}$:

$$\begin{aligned} \mu_1 = 20; \quad \omega_1/\omega_c &\leq 0,2; \quad \omega_c t_{\Pi} \leq 5,5; \\ \mu_1 = 30; \quad \omega_1/\omega_c &\leq 0,09; \quad \omega_c t_{\Pi} \leq 7; \\ \mu_1 = 40; \quad \omega_1/\omega_c &\leq 0,04; \quad \omega_c t_{\Pi} \leq 7,2. \end{aligned}$$

Отсюда видно, что, приняв $\mu_1 = 20 \text{ дБ}$, $\omega_1/\omega_c = 0,2$ и, следовательно, $\omega_c = \omega_1/0,2 = 5\omega_1 = 10 \text{ с}^{-1}$, получим $t_{\Pi} = 0,55 \text{ с}$, $\omega_2/\omega_c \approx 0,4$, $\omega_3/\omega_2 \approx 10$, т. е. типовая характеристика будет удовлетворять поставленным требованиям (ломаная 1 на рис. 6.6.3).

Выбираем в качестве датчика параметра управления сельсинны, работающие в трансформаторном режиме. Как было показано в § 3.8, пару сельсинов можно считать усилителем. Коэффициент усиления их обозначим k_c . В качестве усилителя напряжения выбираем электронный усилитель, коэффициент усиления которого обозначим k_y . В качестве усилителя мощности выбираем электромашиный усилитель. Его передаточная функция, как мы видели в § 3.10, определяется формулой

$$\Phi_o(s) = \frac{k_y}{(T_y s + 1)(T_{\Pi} s + 1)}, \quad (6.6.4)$$

где T_y — постоянная времени цепи входной обмотки, T_{Π} — постоянная времени короткозамкнутой цепи. Исходя из характеристик электромашиных усилителей, применяемых в авиационной технике, принимаем $T_y = 0,008 \text{ с}$, $T_{\Pi} = T_3 = 1/\omega_3 = 0,025 \text{ с}$.

На рис. 6.6.4, а представлена принципиальная, а на рис. 6.6.4, б — структурная схема полученной системы. Ее асимптотическая амплитудная

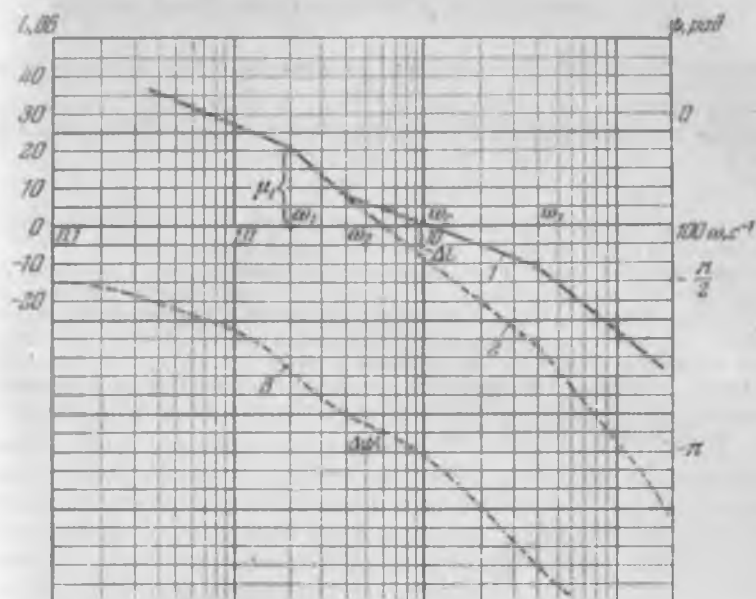


Рис. 6.6.3.

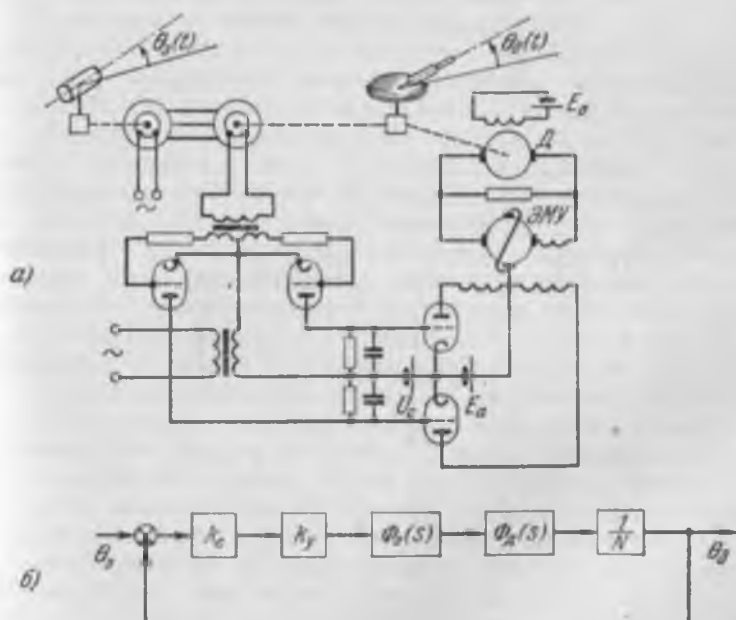


Рис. 6.6.4.

и фазовая характеристики показаны на рис. 6.6.3 пунктиром (ломаная 2 и кривая 3).

Из рис. 6.6.3 видно, что можно добиться совпадения характеристик системы с типовой с помощью корректирующего звена с передаточной функцией

$$\Phi_k(s) = T_2 s + 1.$$

Такое корректирующее звено можно реализовать с помощью цепочки RC , включенной в цепь обратной связи электронного усилителя.

Действительно, звено с передаточной функцией $T_2 s + 1$ является системой, обратной по отношению к апериодическому звену, которое легко реализуется цепочкой RC . А систему, обратную по отношению к данной системе, как мы видели в конце § 4.2, можно приближенно реализовать, включив данную систему в цепь обратной связи усилителя с большим коэффициентом усиления. На основании сказанного в § 3.11 такое корректирующее звено можно также реализовать с помощью цепочки RC во входной цепи электронного усилителя.

Глава 7

МЕТОДЫ ИССЛЕДОВАНИЯ ТОЧНОСТИ ЛИНЕЙНЫХ СИСТЕМ

§ 7.1. Случайные возмущения, действующие в автоматических системах

Всякая автоматическая система работает в реальных условиях под действием случайных возмущений. И если в простейших системах регулирования случайные возмущения малы и не оказывают существенного влияния на работу системы, то в сложных системах случайные возмущения в значительной мере, а иногда и полностью определяют весь процесс работы системы. Так, например, самолеты летают в реальной атмосфере, которая находится в состоянии непрерывного движения и изменения (турбулентность атмосферы). Вследствие этого самолеты всегда находятся под непрерывным действием случайных сил и моментов, вызываемых движением атмосферы. Это знакомо каждому, кому случалось попадать во время полетов на самолетах в «болтанку», которая представляет собой не что иное, как движение самолета, вызываемое случайными аэродинамическими силами и моментами. Случайные аэродинамические силы и моменты существенно влияют на движение самолета, а следовательно, и на работу его системы управления. Эти случайные возмущения заложены в самой природе физического явления, используемого для управления полетом самолета, — аэродинамических сил и моментов, возникающих вследствие относительного движения воздуха и самолета. Если спроектировать систему управления самолетом (автопилот) без учета случайных возмущений, то может оказаться, что при высоком качестве этой системы с точки зрения теории устойчивости, изложенной в предыдущей главе, она будет очень плохо управлять самолетом во время «болтанки». Проектируя систему управления полетом, конструктор должен заботиться об удобствах пассажиров самолета, добиваясь наименьшего влияния на самолет «болтанки». А это можно сделать только тогда, когда мы научимся оценивать влияние случайных сил и моментов на движение самолета и работу его системы управления.

Вторым примером системы, на которую существенно влияют случайные возмущения, может служить следящая система радиолокатора. Вследствие колебаний самолета, координаты которого измеряет радиолокатор, и электромагнитных процессов в атмосфере сигнал, отраженный от самолета, совершает беспорядочные

случайные колебания. Вследствие этого антенна радиолокатора принимает наряду с отраженным от самолета полезным сигналом, несущим информацию о его координатах, наложенные на полезный сигнал случайные электромагнитные колебания. Полностью отделить полезный сигнал от случайных колебаний принципиально невозможно, так как случайные колебания возникают вследствие внутренних свойств того самого физического явления — отражения и распространения электромагнитных волн, — которое используется для получения информации о координатах самолета. Вследствие случайных колебаний отраженного от самолета сигнала исполнительное устройство следящей системы всегда получает управляющий сигнал, искаженный помехой, что вызывает ошибку сопровождения самолета радиолокатором. Кроме того, на антенну радиолокатора действуют случайные аэродинамические силы и моменты, которые также оказывают влияние на работу исполнительного устройства и на ошибку сопровождения.

В качестве третьего примера можно указать любую промышленную автоматическую систему. Любая такая система работает не изолированно, а в соседстве с другими системами. Поэтому в элементах любой системы, кроме связанных с ее работой электромагнитных полей и токов, работой других систем наводятся дополнительные электромагнитные поля и токи, которые являются случайными возмущениями для данной системы. Такие внешние помехи существенно влияют на работу промышленной автоматической системы, особенно если по соседству с ней имеются источники мощных электромагнитных полей. Наконец, в самих элементах любой автоматической системы всегда возникают постоянные «шумы», т. е. случайные колебания (флуктуации) токов, вызываемые тепловым движением молекул в сопротивлениях и бомбардировкой электронами электродов электронных ламп.

Приведенные примеры показывают, что случайные возмущения, действующие на автоматические системы, могут быть внешними и внутренними. Внешние случайные возмущения связаны со средой, в которой работает система, и могут или действовать на объект управления, или поступать в систему управления через датчики информации (измерители) вместе с полезной информацией о задачах управления и о состоянии объекта управления (результатах управления).

В системах управления военного назначения, наряду с естественными внешними помехами, могут действовать и искусственные помехи, специально организованные противником с целью помешать работе систем управления и дезорганизовать управление военной техникой и войсками. Особенно уязвимы для искусственных помех датчики информации, через которые и поступают в систему управления организованные противником помехи. При этом в качестве источника помех противник использует те самые

физические явления, которые заложены в основу устройства систем получения информации.

Из сказанного можно сделать вывод, что устранить влияние случайных возмущений — шумов и помех — на работу автоматических систем принципиально невозможно, так как они порождаются теми самыми физическими явлениями, которые используются для устройства автоматических систем. Можно говорить лишь о большей или меньшей степени влияния случайных возмущений. Но в таком случае при проектировании автоматических систем необходимо стремиться к тому, чтобы свести влияние случайных возмущений — шумов и помех — к минимуму. А для этого необходимо научиться оценивать влияние случайных возмущений на работу автоматических систем, оценивать точность их работы под действием случайных возмущений. Это особенно важно для автоматических систем военного назначения, которые должны получать и обрабатывать информацию о действиях противника на предельных дальностях действия источников информации и управлять при этом действиями своих войск и военной техникой, когда сигналы от датчиков информации весьма слабы и вследствие этого сильно искажены шумами и помехами или когда противник применяет искусственные помехи.

Изучение работы автоматических систем под действием случайных возмущений является предметом статистической теории (статистической динамики) систем управления. После первых работ советских ученых [4, 52] (см. также библиографию в [53]) и американских ученых [12, 18], посвященных статистической теории автоматических систем, статистические (вероятностные) методы теории автоматического управления развивались быстрыми темпами и в настоящее время являются одним из важнейших и наиболее перспективных разделов общей теории процессов управления.

§ 7.2. Общие методы исследования точности линейных систем

Действующие в автоматических системах случайные возмущения обычно накладываются на полезные сигналы. В результате этого входные переменные любой системы практически всегда представляют собой случайные функции времени.

Всякая автоматическая система предназначена для работы с различными полезными сигналами. Множество всех полезных входных сигналов автоматической системы можно рассматривать как множество возможных реализаций некоторой случайной функции $S(t)$. Таким образом, и полезные входные сигналы автоматических систем можно рассматривать как случайные функции. Однако при исследовании точности автоматических систем часто важно определить зависимость точности от условий их работы,

т. е. от того, какой полезный сигнал действует на входе. Поэтому точность автоматической системы с известными характеристиками обычно исследуется для каждой данной реализации входного полезного сигнала. Иными словами, при анализе точности автоматических систем обычно учитывается только случайность ошибок измерений, шумов и всех других видов помех, а полезные сигналы считаются определенными функциями времени. При таких условиях математические ожидания входных случайных функций обычно представляют собой полезные входные сигналы, если только измерительные элементы системы свободны от систематических ошибок, которые, как правило, легко устраняются соответствующей регулировкой. Случайные колебания входных переменных около их математических ожиданий, которые обычно называются *флуктуациями* входных переменных, представляют собой помехи, нарушающие нормальную работу системы и приводящие к случайным ошибкам в выходных переменных. Вследствие случайных колебаний входных переменных системы ее выходные переменные тоже совершают случайные колебания, т. е. представляют собой случайные функции времени. При этом математические ожидания выходных переменных представляют собой выходные полезные сигналы системы.

Разности между фактическими полезными выходными сигналами системы и требуемыми выходными сигналами представляют собой *систематические ошибки* системы. Случайные колебания (флуктуации) выходных переменных представляют собой *случайные ошибки* системы.

Задача исследования точности автоматической системы состоит в определении ее систематических ошибок и вероятностных характеристик случайных ошибок, т. е. элементов рассеивания выходных переменных. При этом обычно бывает достаточно определить дисперсии выходных переменных, а в случае системы с несколькими выходами иногда и корреляционные моменты выходных переменных, характеризующие степень вероятностной связи между ними. Изложенное показывает, что для нахождения систематических ошибок системы необходимо определить математические ожидания ее выходных переменных. Следовательно, задача исследования точности автоматической системы приводится к нахождению математических ожиданий, дисперсий и корреляционных функций ее выходных переменных. Для определения этих величин необходимо знать, как данная система преобразует входные переменные, т. е. необходимо знать оператор системы.

На основании изложенного общая задача исследования точности одномерной системы формулируется следующим образом: на систему с известным оператором A действует возмущение $X(t)$, представляющее собой случайную функцию времени; зная вероятностные характеристики случайной функции $X(t)$, требуется

найти математическое ожидание и дисперсию выходной переменной системы $Y(t)$. Эта задача принципиально легко решается для линейных систем. При этом для определения математического ожидания и дисперсии выходной переменной линейной системы необходимо знать лишь математическое ожидание и корреляционную функцию входного случайного возмущения. Зная эти характеристики, можно также определить корреляционную функцию выходной переменной линейной системы. Таким образом, мы приходим к следующей задаче: зная математическое ожидание $m_x(t)$ и корреляционную функцию $K_x(t, t')$ действующего на линейную систему входного случайного возмущения и оператор этой системы, найти математическое ожидание $m_y(t)$, дисперсию $D_y(t)$ и корреляционную функцию $K_y(t, t')$ выходной переменной системы $Y(t)$.

Поставленная задача решается различными способами в зависимости от того, какие характеристики линейной системы заданы (т. е. как задан ее оператор). Если известна весовая функция физически возможной линейной системы $g(t, \tau)$, то ее выходная переменная на основании общей формулы (2.2.5) выражается формулой

$$Y(t) = \int_{t_0}^t g(t, \tau) X(\tau) d\tau. \quad (7.2.1)$$

Для определения математического ожидания выходной переменной системы заметим, что интеграл в (7.2.1) является пределом последовательности соответствующих сумм, которые представляют собой линейные комбинации значений случайной функции $X(t)$ в дискретном ряде точек. Как известно из теории вероятностей, математическое ожидание линейной функции случайных величин равно той же линейной функции математических ожиданий этих величин. Поэтому для вычисления математических ожиданий сумм, аппроксимирующих интеграл (7.2.1), достаточно заменить в них случайную функцию $X(t)$ ее математическим ожиданием $m_x(t)$. Следовательно, и при переходе к пределу случайная функция $X(t)$ заменится в формуле (7.2.1) ее математическим ожиданием $m_x(t)$, и мы получим

$$m_y(t) = \int_{t_0}^t g(t, \tau) m_x(\tau) d\tau. \quad (7.2.2)$$

Эта формула показывает, что операции математического ожидания и интегрирования практически всегда можно переставлять.

Для вычисления корреляционной функции выходной переменной $Y(t)$ найдем соответствующую центрированную случайную функцию $Y^0(t)$. Вычитая формулу (7.2.2) почленно из (7.2.1),

получим

$$Y(t) - m_y(t) = \int_{t_0}^t g(t, \tau) [X(\tau) - m_x(\tau)] d\tau,$$

или

$$Y^0(t) = \int_{t_0}^t g(t, \tau) X^0(\tau) d\tau. \quad (7.2.3)$$

Давая здесь переменной t значение t' и изменяя обозначение переменной интегрирования, от которого определенный интеграл не зависит, можем написать

$$Y^0(t') = \int_{t_0}^{t'} g(t', \tau') X^0(\tau') d\tau', \quad (7.2.4)$$

Перемножая формулы (7.2.3) и (7.2.4) почленно, будем иметь

$$\begin{aligned} Y^0(t) Y^0(t') &= \int_{t_0}^t g(t, \tau) X^0(\tau) d\tau \cdot \int_{t_0}^{t'} g(t', \tau') X^0(\tau') d\tau' = \\ &= \int_{t_0}^t \left\{ \int_{t_0}^{t'} g(t', \tau') X^0(\tau') d\tau' \right\} g(t, \tau) X^0(\tau) d\tau = \\ &= \int_{t_0}^t \left\{ \int_{t_0}^{t'} g(t, \tau) g(t', \tau') X^0(\tau) X^0(\tau') d\tau' \right\} d\tau = \\ &= \int_{t_0}^t \int_{t_0}^{t'} g(t, \tau) g(t', \tau') X^0(\tau) X^0(\tau') d\tau d\tau'. \end{aligned}$$

Отсюда, пользуясь возможностью изменять порядок операций математического ожидания и интегрирования, находим

$$\begin{aligned} M[Y^0(t) Y^0(t')] &= M \left[\int_{t_0}^t \int_{t_0}^{t'} g(t, \tau) g(t', \tau') X^0(\tau) X^0(\tau') d\tau d\tau' \right] = \\ &= \int_{t_0}^t \int_{t_0}^{t'} M[g(t, \tau) g(t', \tau') X^0(\tau) X^0(\tau')] d\tau d\tau' = \\ &= \int_{t_0}^t \int_{t_0}^{t'} g(t, \tau) g(t', \tau') M[X^0(\tau) X^0(\tau')] d\tau d\tau'. \end{aligned} \quad (7.2.5)$$

Согласно определению корреляционной функции случайной функции

$$M[X^0(\tau)X^0(\tau')] = K_x(\tau, \tau'), \quad M[Y^0(t)Y^0(t')] = K_y(t, t').$$

Подставляя эти выражения в формулу (7.2.5), получим следующую зависимость между корреляционными функциями входной и выходной переменных физически возможной линейной системы:

$$K_y(t, t') = \int_{t_0}^t \int_{t_0}^{t'} g(t, \tau) g(t', \tau') K_x(\tau, \tau') d\tau d\tau'. \quad (7.2.6)$$

Для определения дисперсии выходной переменной линейной системы достаточно положить в формуле (7.2.6) $t' = t$. Тогда получим

$$D_y(t) = K_y(t, t) = \int_{t_0}^t \int_{t_0}^t g(t, \tau) g(t, \tau') K_x(\tau, \tau') d\tau d\tau'. \quad (7.2.7)$$

Эта формула показывает, что для вычисления дисперсии выходной переменной линейной системы необходимо знать корреляционную функцию входного случайного возмущения. Знания дисперсии входной переменной недостаточно для определения дисперсии выходной переменной.

Общие формулы для математических ожиданий, корреляционных функций и дисперсий выходных переменных любых линейных систем, в том числе и физически невозможных, отличаются от (7.2.2), (7.2.6) и (7.2.7) только тем, что вместо интегралов в конечных пределах в них фигурируют интегралы в пределах от $-\infty$ до ∞ .

Таким образом, задача исследования точности линейных систем решается принципиально очень просто. Достаточно знать математическое ожидание и корреляционную функцию входного случайного возмущения и весовую функцию линейной системы. Тогда определение математического ожидания и дисперсии выходной переменной системы сведется к вычислению интегралов, что может быть сделано любым известным способом, точным или приближенным.

Формулы (7.2.2), (7.2.6) и (7.2.7) справедливы для любых линейных систем. В частности, они справедливы и для дискретных линейных систем. Согласно изложенному в § 5.2 весовая функция физически возможной дискретной линейной системы представляет собой линейную комбинацию δ -функций:

$$g(t, \tau) = \sum_{t_k \leq t} g_k(t) \delta(\tau - t_k).$$

Подставляя это выражение в формулы (7.2.2), (7.2.6) и (7.2.7), получим соответственно следующие формулы для математического ожидания, корреляционной функции и дисперсии выходной переменной дискретной линейной системы:

$$m_y(t) = \sum_{t_0 \leq t_h \leq t} g_h(t) m_x(t_h), \quad (7.2.8)$$

$$K_y(t, t') = \sum_{t_0 \leq t_h \leq t} \sum_{t_0 \leq t'_h \leq t'} g_h(t) g_h(t') K_x(t_h, t'_h), \quad (7.2.9)$$

$$D_y(t) = \sum_{t_0 \leq t_h, t'_h \leq t} g_h(t) g_h(t') K_x(t_h, t'_h). \quad (7.2.10)$$

Пример 7.2.1. Найти математическое ожидание и дисперсию выходной переменной аperiodического звена (например, цепочки RC), предполагая, что на входе этого звена, начиная с момента t_0 , действует возмущение, представляющее собой стационарную случайную функцию с математическим ожиданием m_x и корреляционной функцией

$$K_x(t, t') = k_x(t - t') = D e^{-\alpha|t - t'|}. \quad (7.2.11)$$

Подставляя в формулу (7.2.2) выражение (4.4.33) весовой функции аperiodического звена, находим математическое ожидание выходной переменной:

$$m_y(t) = \frac{k m_x}{T} \int_{t_0}^t e^{-\frac{t-\tau}{T}} d\tau = k m_x \left(1 - e^{-\frac{t-t_0}{T}}\right). \quad (7.2.12)$$

Подставляя выражение (4.4.33) весовой функции и выражение (7.2.11) корреляционной функции в формулу (7.2.7), находим дисперсию выходной переменной:

$$D_y(t) = \frac{k^2 D}{T^2} \int_{t_0}^t \int_{t_0}^t e^{-\frac{t-\tau}{T}} e^{-\frac{t-\tau'}{T}} e^{-\alpha|\tau-\tau'|} d\tau d\tau'. \quad (7.2.13)$$

Для вычисления этого интеграла разобьем интервал интегрирования по переменной τ' на две части: (t_0, τ) и (τ, t) . Тогда получим

$$\begin{aligned} D_y(t) &= \frac{k^2 D}{T^2} e^{-\frac{2t}{T}} \int_{t_0}^t \left\{ \int_{t_0}^{\tau} e^{\frac{\tau+\tau'}{T}} e^{-\alpha(\tau-\tau')} d\tau' + \int_{\tau}^t e^{\frac{\tau+\tau'}{T}} e^{-\alpha(\tau'-\tau)} d\tau' \right\} d\tau = \\ &= \frac{k^2 D}{1 - \alpha^2 T^2} \left\{ 1 - \alpha T - 2e^{-\left(\frac{1}{T} + \alpha\right)(t-t_0)} + (1 + \alpha T) e^{-\frac{2}{T}(t-t_0)} \right\}. \end{aligned} \quad (7.2.14)$$

Для исследования точности линейных систем можно также применить метод канонических представлений случайных функций. Если выразить входное случайное возмущение $X(t)$

каким-либо каноническим разложением *):

$$X(t) = m_x(t) + \sum_v V_v x_v(t), \quad (7.2.15)$$

то на основании принципа суперпозиции выходная переменная линейной системы выразится таким же каноническим разложением:

$$Y(t) = m_y(t) + \sum_v V_v y_v(t), \quad (7.2.16)$$

где функции $m_y(t)$, $y_v(t)$ представляют собой соответственно результаты преобразования рассматриваемой системой функций $m_x(t)$, $x_v(t)$. Таким образом, обозначая оператор линейной системы буквой A , можем написать

$$m_y(t) = A m_x(t), \quad (7.2.17)$$

$$y_v(t) = A x_v(t). \quad (7.2.18)$$

После нахождения координатных функций y_v корреляционная функция и дисперсия выходной переменной Y определяются по известным общим формулам теории случайных функций:

$$K_y(t, t') = \sum_v D_v y_v(t) \overline{y_v(t')}, \quad (7.2.19)$$

$$D_y(t) = \sum_v D_v |y_v(t)|^2. \quad (7.2.20)$$

Для нахождения функций $m_y(t)$ и $y_v(t)$ можно пользоваться различными методами в зависимости от того, какие характеристики линейной системы заданы. Если известна весовая функция системы, то математическое ожидание и координатные функции выходной переменной системы можно определить, пользуясь общей формулой (2.2.5), связывающей входную и выходную переменные линейной системы. В результате получим для математического ожидания выходной переменной $m_y(t)$ формулу (7.2.2), а для координатных функций $y_v(t)$ получим аналогичную формулу:

$$y_v(t) = \int_{t_0}^t g(t, \tau) x_v(\tau) d\tau. \quad (7.2.21)$$

Если весовая функция системы неизвестна и оператор системы задан уравнениями, описывающими поведение системы (дифференциальными, разностными или любыми другими), то теми же самыми уравнениями на основании принципа суперпозиции определяются математические ожидания и соответствующие координаты

*) См. учебник по теории вероятностей В. С. Пугачева [54], § 6.2 или [53], гл. 9.

натные функции входящих в уравнения переменных. Иными словами, для определения математических ожиданий выходных переменных линейной системы следует заменить в уравнениях этой системы все переменные, как известные, так и неизвестные, их математическими ожиданиями и решить полученные уравнения. Для определения координатных функций выходных переменных линейной системы следует заменить в уравнениях этой системы все переменные соответствующими координатными функциями и решить полученные уравнения.

При практическом применении этого метода необходимо иметь в виду, что если в уравнения, описывающие поведение системы, входят в виде слагаемых известные функции времени (т. е. действующие на систему неслучайные возмущения), то их следует рассматривать как случайные функции с нулевыми дисперсиями и нулевыми координатными функциями. Поэтому при составлении уравнений для определения координатных функций выходных переменных необходимо отбрасывать все неслучайные возмущения, т. е. известные функции времени, входящие в уравнения системы в виде слагаемых. В уравнениях для математических ожиданий переменных все такие неслучайные функции времени необходимо сохранить.

Пример 7.2.2. Движение самолета или крылатой ракеты в вертикальной плоскости в режиме прямолинейного горизонтального полета с учетом ветра описывается приближенными дифференциальными уравнениями

$$\left. \begin{aligned} \dot{y} &= v\theta, \quad \dot{\theta} = A_\alpha \alpha + A_\alpha \frac{W}{v} - \frac{g}{v}, \\ \ddot{\alpha} + c_\alpha \dot{\alpha} + c_\alpha \alpha &= \left(\frac{A_\alpha}{v} \dot{v} - c_\alpha \right) \frac{W}{v} - A_\alpha \frac{\dot{W}}{v} + c_0 - c_0 \delta, \end{aligned} \right\} \quad (7.2.22)$$

где W — вертикальная составляющая вектора скорости ветра, а остальные обозначения — те же, что в § 3.16. При полете в неспокойной атмосфере (а она практически всегда неспокойна в той или иной степени) действующая на самолет или ракету вертикальная скорость ветра, вызывающая «болтанку», является случайной функцией времени *). Вследствие этого и все остальные переменные, входящие в уравнения движения (7.2.22), являются случайными функциями времени. Если скорость полета v считать известной функцией времени, что обычно всегда возможно, то все коэффициенты при переменных θ , α , $\dot{\alpha}$, δ , W , $\dot{W}$ в уравнениях (7.2.22) будут известными функциями времени (в частном случае при $v = \text{const}$ — постоянными). Слагаемое $-g/v$ во втором уравнении и слагаемое $c_0 - c_0 \delta$ в третьем уравнении при полете с закрепленным рулем ($\delta = \text{const}$) будут также известными функциями времени, играющими роль неслучайных возмущений. Согласно изложенному, для матема-

*) Строго говоря, скорость ветра является случайной функцией времени и координат точки пространства. Однако при полете летательного аппарата его координаты являются функциями времени. Вследствие этого действующий на летательный аппарат ветер можно рассматривать как случайную функцию времени.

тических ожиданий элементов движения самолета или ракеты получаем уравнения

$$\left. \begin{aligned} \dot{m}_y &= v m_0, \quad \dot{m}_\theta = A_\alpha m_\alpha + A_\alpha \frac{m_w}{v} - \frac{g}{v}, \\ \ddot{m}_\alpha + c_\alpha \dot{m}_\alpha + c_\alpha m_\alpha &= \left(\frac{A_\alpha}{v} \dot{v} - c_\alpha \right) \frac{m_w}{v} - A_\alpha \frac{\dot{m}_w}{v} + c_0 - c_\delta \delta. \end{aligned} \right\} \quad (7.2.23)$$

В уравнениях для координатных функций переменных $y_v(t)$, $\theta_v(t)$, $\alpha_v(t)$ следует отбросить неслучайные возмущения $-g/v$ и $c_0 - c_\delta \delta$, так как их координатные функции равны нулю. В результате получим уравнения

$$\left. \begin{aligned} \dot{y}_v &= v \theta_v, \quad \dot{\theta}_v = A_\alpha \alpha_v + A_\alpha \frac{w_v}{v}, \\ \ddot{\alpha}_v + c_\alpha \dot{\alpha}_v + c_\alpha \alpha_v &= \left(\frac{A_\alpha}{v} \dot{v} - c_\alpha \right) \frac{w_v}{v} - A_\alpha \frac{\dot{w}_v}{v} \quad (v=1, 2, \dots). \end{aligned} \right\} \quad (7.2.24)$$

В случае, когда оператор системы задан уравнениями, для полного определения поведения системы необходимо задать еще начальные условия, т. е. начальные значения переменных и их производных или конечных разностей определенных порядков.

Если начальные условия заданы и не являются случайными, то их следует учитывать только при решении уравнений, определяющих математические ожидания переменных. А именно уравнения для математических ожиданий переменных всегда следует решать при тех же неслучайных начальных условиях, при которых рассматривается поведение системы. Уравнения для координатных функций всегда должны решаться при пулевых начальных условиях. Это видно непосредственно из формулы (7.2.16). Начальные значения функции $Y(t)$ и ее производных или конечных разностей до определенного порядка в момент t_0 могут быть неслучайными только в том случае, когда начальные значения всех координатных функций $y_v(t)$ и их производных или конечных разностей соответствующих порядков равны нулю.

Если поведение системы рассматривается при случайных начальных условиях, то движение системы в силу принципа суперпозиции можно рассматривать как сумму двух движений: движения, вызываемого постоянным действием входных случайных возмущений, при начальных значениях переменных и их производных или конечных разностей, равных математическим ожиданиям случайных начальных значений соответствующих переменных, и свободного движения, вызываемого случайными отклонениями начальных значений переменных от их математических ожиданий.

Первое движение исследуется так же, как движение системы при неслучайных начальных условиях: уравнения для математических ожиданий переменных решаются при начальных значе-

ниях переменных, равных математическим ожиданиям случайных начальных значений соответствующих переменных, а уравнения для координатных функций решаются при нулевых начальных условиях. После нахождения координатных функций корреляционные функции и дисперсии интересующих нас переменных вычисляются по формулам (7.2.19) и (7.2.20) соответственно.

Второе движение, вызываемое случайными отклонениями начальных условий от их математических ожиданий, исследуется путем решения соответствующих однородных уравнений, т. е. уравнений, описывающих поведение исследуемой системы при отсутствии входных возмущений, с учетом случайных начальных условий. В результате все переменные, входящие в уравнения системы, определяются как линейные функции начальных значений переменных и их производных или конечных разностей, которые являются обычными случайными величинами. Поэтому дисперсии и корреляционные функции интересующих нас переменных легко вычисляются по известным формулам теории вероятностей для линейных функций случайных величин.

Для определения дисперсий и корреляционных функций переменных, соответствующих суммарному движению системы, следует воспользоваться формулами для дисперсии и корреляционной функции суммы двух случайных функций. Так как случайные начальные условия в задачах практики обычно всегда независимы от входных случайных возмущений, то для вычисления дисперсий и корреляционных функций интересующих нас переменных в суммарном движении системы под действием входных случайных возмущений и случайных начальных условий достаточно сложить соответственно дисперсии и корреляционные функции переменных в движении под действием входных случайных возмущений и тех же переменных в движении, вызываемом случайными отклонениями начальных условий от их математических ожиданий.

Пример 7.2.3. В условиях примера 7.2.2, рассматривая движение самолета или крылатой ракеты при заданных неслучайных начальных условиях

$$t = t_0, \quad y = y_0, \quad \theta = \theta_0, \quad \alpha = \alpha_0, \quad \dot{\alpha} = \dot{\alpha}_0, \quad (7.2.25)$$

мы должны, согласно изложенному, интегрировать уравнения (7.2.23) для математических ожиданий переменных при тех же начальных условиях:

$$t = t_0, \quad m_y = y_0, \quad m_\theta = \theta_0, \quad m_\alpha = \alpha_0, \quad m_{\dot{\alpha}} = \dot{\alpha}_0. \quad (7.2.26)$$

Уравнения (7.2.24) для координатных функций всегда необходимо интегрировать при нулевых начальных условиях:

$$t = t_0, \quad y_v = \theta_v = \alpha_v = \dot{\alpha}_v = 0 \quad (v = 1, 2, \dots). \quad (7.2.27)$$

Если начальные значения переменных $y, \theta, \alpha, \dot{\alpha}$ случайны, то движение самолета или ракеты следует рассматривать как сумму двух движений: дви-

жения, определяемого дифференциальными уравнениями (7.2.22) с заменой в них переменных y, θ, α переменными $y^{(1)}, \theta^{(1)}, \alpha^{(1)}$ при начальных условиях

$$t = t_0, \quad y^{(1)} = m_{y_0}, \quad \theta^{(1)} = m_{\theta_0}, \quad \alpha^{(1)} = m_{\alpha_0}, \quad \dot{\alpha}^{(1)} = m_{\dot{\alpha}_0}, \quad (7.2.28)$$

и свободного движения, определяемого соответствующими однородными уравнениями

$$\ddot{y}^{(2)} = \nu \theta^{(2)}, \quad \ddot{\theta}^{(2)} = A_{\alpha} \alpha^{(2)}, \quad \ddot{\alpha}^{(2)} + c_{\alpha} \dot{\alpha}^{(2)} + c_{\alpha} \alpha^{(2)} = 0 \quad (7.2.29)$$

при центрированных случайных начальных условиях

$$t = t_0, \quad y^{(2)} = y_0^0, \quad \theta^{(2)} = \theta_0^0, \quad \alpha^{(2)} = \alpha_0^0, \quad \dot{\alpha}^{(2)} = \dot{\alpha}_0^0. \quad (7.2.30)$$

Для исследования первого движения следует интегрировать уравнения (7.2.23) для математических ожиданий при начальных условиях (7.2.28). В результате будут определены математические ожидания переменных y, θ, α в суммарном движении, так как математические ожидания переменных $y^{(2)}, \theta^{(2)}, \alpha^{(2)}$ в свободном движении при начальных условиях (7.2.30) равны нулю. После интегрирования уравнений (7.2.24) для координатных функций при нулевых начальных условиях (7.2.27) дисперсии интересующих нас переменных, например, отклонения по высоте $y^{(1)}(t)$, определяются по формуле (7.2.20):

$$D_{y^{(1)}}(t) = \sum_v D_v |y_v(t)|^2. \quad (7.2.31)$$

Для исследования второго движения необходимо проинтегрировать уравнения (7.2.29) при начальных условиях (7.2.30). В результате получим интересующие нас переменные, например отклонение по высоте $y^{(2)}(t)$, как линейные комбинации начальных значений переменных:

$$y^{(2)}(t) = y_0^0 y_y(t) + \theta_0^0 y_{\theta}(t) + \alpha_0^0 y_{\alpha}(t) + \dot{\alpha}_0^0 y_{\dot{\alpha}}(t), \quad (7.2.32)$$

где $y_y(t), y_{\theta}(t), y_{\alpha}(t), y_{\dot{\alpha}}(t)$ — некоторые вполне определенные функции времени t . Пользуясь известными формулами теории вероятностей для линейных функций случайных величин, находим дисперсию величины $y^{(2)}(t)$, считая для простоты начальные значения $y_0^0, \theta_0^0, \alpha_0^0, \dot{\alpha}_0^0$ некоррелированными:

$$D_{y^{(2)}}(t) = D_{y_0^0} y_y^2(t) + D_{\theta_0^0} y_{\theta}^2(t) + D_{\alpha_0^0} y_{\alpha}^2(t) + D_{\dot{\alpha}_0^0} y_{\dot{\alpha}}^2(t). \quad (7.2.33)$$

Считая случайные начальные значения переменных $y_0, \theta_0, \alpha_0, \dot{\alpha}_0$ независимыми от случайной скорости ветра W , найдем дисперсию отклонения по высоте в суммарном движении с учетом как действующего на самолет или ракету ветра, так и случайных начальных условий по формуле

$$D_y(t) = D_{y^{(1)}}(t) + D_{y^{(2)}}(t). \quad (7.2.34)$$

Совершенно аналогично, выразив входное случайное возмущение интегральным каноническим представлением (см. [54], § 6.5 или [53], § 67)

$$X(t) = m_x(t) + \int_{\lambda_1}^{\lambda_2} V(\lambda) x(t, \lambda) d\lambda, \quad (7.2.35)$$

получим в силу принципа суперпозиции в интегральной форме аналогичное интегральное каноническое представление выходной переменной линейной системы:

$$Y(t) = m_y(t) + \int_{\lambda_1}^{\lambda_2} V(\lambda) y(t, \lambda) d\lambda, \quad (7.2.36)$$

где $y(t, \lambda)$ при каждом данном значении параметра λ есть результат прохождения через данную систему соответствующей координатной функции $x(t, \lambda)$:

$$y(t, \lambda) = A_i x(t, \lambda). \quad (7.3.37)$$

В зависимости от того, как задан оператор линейной системы, функции $y(t, \lambda)$ определяются различными изложенными выше способами. В частности, если задана весовая функция системы $g(t, \tau)$, координатные функции $y(t, \lambda)$ определяются формулой

$$y(t, \lambda) = \int_{t_0}^t g(t, \tau) x(\tau, \lambda) d\tau, \quad (7.2.38)$$

аналогичной формуле (7.2.21).

Если оператор системы задан уравнениями, то координатные функции $y(t, \lambda)$ определяются теми же уравнениями, в которых все переменные заменены соответствующими координатными функциями при данном значении λ , при нулевых начальных условиях.

После нахождения координатных функций $y(t, \lambda)$ корреляционная функция и дисперсия выходной переменной исследуемой линейной системы определяются на основании известных общих формул теории случайных функций (см., например, [54], § 6.5) следующими формулами:

$$K_y(t, t') = \int_{\lambda_1}^{\lambda_2} G(\lambda) y(t, \lambda) \overline{y(t', \lambda)} d\lambda, \quad (7.2.39)$$

$$D_y(t) = \int_{\lambda_1}^{\lambda_2} G(\lambda) |y(t, \lambda)|^2 d\lambda, \quad (7.2.40)$$

где $G(\lambda)$ — интенсивность белого шума $V(\lambda)$ в формуле (7.2.35).

При практическом применении метода интегральных канонических представлений для исследования точности линейных систем часто оказывается невозможным определить функции $y(t, \lambda)$ аналитическими методами для всех значений параметра λ . В таких случаях интегралы (7.2.39) и (7.2.40) приходится находить численными или графическими методами. При этом необходимо предварительно определить функции $y(t, \lambda)$ для конечного числа

значений параметра λ , что можно сделать, пользуясь теми же методами, с помощью которых находятся координатные функции канонического разложения $y_v(t)$.

Метод канонических разложений дает возможность использовать для исследования точности линейной системы саму систему, или ее действующий макет, или модель. Подавая на вход системы по очереди в достаточно большом масштабе функции $m_x(t)$ и $x_v(t)$ [или $x(t, \lambda)$], получим на выходе соответствующие функции $m_y(t)$ и $y_v(t)$ [или $y(t, \lambda)$].

Аналогично определяются функции $m_y(t)$, $y_v(t)$ в случае задания других характеристик линейной системы. Так, например, если задана частотная характеристика линейной системы, то функции $m_y(t)$, $y_v(t)$ могут быть определены по формулам (2.3.2) и (2.3.13), в которых функция $x(t)$ должна быть заменена соответствующей функцией $m_x(t)$, $x_v(t)$, а $y(t)$ — функцией $m_y(t)$, $y_v(t)$.

Мы видим, таким образом, что функции $m_y(t)$ и $y_v(t)$ определяются различными способами в зависимости от того, как задан оператор исследуемой линейной системы.

Пример 7.2.4. Решить задачу, рассмотренную в примере 7.2.1, методом интегральных канонических представлений.

Входное случайное возмущение $X(t)$ в данном случае представляет собой стационарную случайную функцию времени t . Следовательно, оно может быть выражено интегральным каноническим представлением (см. [54], §§ 5.3 и 6.5)

$$X(t) = m_x + \int_{-\infty}^{\infty} V(\omega) e^{i\omega t} dt, \quad (7.2.41)$$

координатными функциями которого являются показательные функции $e^{i\omega t}$, выражающие гармонические колебания всех возможных частот. При этом интенсивность белого шума $V(\omega)$, представляющая собой спектральную плотность случайной функции $X(t)$, определяется формулой

$$G(\omega) = s_x(\omega) = \frac{D}{\pi} \frac{\alpha}{\alpha^2 + \omega^2}. \quad (7.2.42)$$

Подставляя это выражение в (7.2.40) и принимая во внимание, что частота ω , играющая в данном случае роль параметра λ , изменяется в пределах $(-\infty, \infty)$, получим следующую формулу для дисперсии выходной переменной системы:

$$D_y(t) = \frac{D\alpha}{\pi} \int_{-\infty}^{\infty} \frac{|y(t, \omega)|^2 d\omega}{\alpha^2 + \omega^2}, \quad (7.2.43)$$

где функция $y(t, \omega)$ представляет собой реакцию рассматриваемой системы на гармонические колебания $e^{i\omega t}$, действующие на нее начиная с момента t_0 .

Для определения функций $y(t, \omega)$, соответствующих всем возможным значениям частоты ω , можно, согласно изложенному, применить различные методы. Подставляя в формулу (7.2.38) выражение (4.4.33) весовой функции

$g(t, \tau)$ и принимая во внимание, что $x(\tau, \omega) = e^{i\omega\tau}$, получим

$$y(t, \omega) = \frac{k}{T} \int_0^t e^{-\frac{t-\tau}{T} + i\omega\tau} d\tau = \frac{k [e^{i\omega t} - e^{-\frac{t-t_0}{T} + i\omega t_0}]}{1 + iT\omega}. \quad (7.2.44)$$

Подставляя это выражение в (7.2.43), найдем дисперсию выходной переменной $Y(t)$:

$$D_y(t) = \frac{k^2 D\alpha}{\pi} \int_{-\infty}^{\infty} \frac{|e^{i\omega t} - e^{-\frac{t-t_0}{T} + i\omega t_0}|^2}{(\alpha^2 + \omega^2)(1 + T^2\omega^2)} d\omega. \quad (7.2.45)$$

Наконец, выразив показательные функции по формуле Эйлера через тригонометрические функции, приведем формулу (7.2.45) к действительной форме:

$$\begin{aligned} D_y(t) &= \frac{k^2 D\alpha}{\pi} \int_{-\infty}^{\infty} \frac{(\cos \omega t - e^{-\frac{t-t_0}{T}} \cos \omega t_0)^2 + (\sin \omega t - e^{-\frac{t-t_0}{T}} \sin \omega t_0)^2}{(\alpha^2 + \omega^2)(1 + T^2\omega^2)} d\omega = \\ &= \frac{k^2 D\alpha}{\pi} \int_{-\infty}^{\infty} \frac{1 - 2e^{-\frac{t-t_0}{T}} \cos \omega(t-t_0) + e^{-\frac{2(t-t_0)}{T}}}{(\alpha^2 + \omega^2)(1 + T^2\omega^2)} d\omega. \end{aligned} \quad (7.2.46)$$

В результате вычисления интеграла в этой формуле снова получается формула (7.2.14).

Координатные функции $y(t, \omega)$ на основании изложенного можно также определить как интеграл уравнения рассматриваемой системы, в котором входная и выходная переменные заменены соответствующими координатными функциями $x(t, \omega)$ и $y(t, \omega)$. В результате получаем дифференциальное уравнение

$$Ty'_i(t, \omega) + y(t, \omega) = ke^{i\omega t}. \quad (7.2.47)$$

Функция $y(t, \omega)$ представляет собой интеграл уравнения (7.2.47) при нулевом начальном условии $y(t_0, \omega) = 0$. Этот интеграл выражается формулой (7.2.44), выведенной ранее другим способом.

При $m \neq 0$ 7.2.5. Для изучения поведения системы, рассмотренной в предыдущем примере, при случайном начальном условии $Y(t) = Y_0$, $m_{i/0} = 0$ необходимо к движению системы, исследованному в предыдущем примере, добавить свободное движение, определяемое соответствующим однородным уравнением и случайным начальным условием

$$T\dot{Y}^{(2)} + Y^{(2)} = 0, \quad Y^{(2)}(t_0) = Y_0. \quad (7.2.48)$$

Интеграл этого уравнения, удовлетворяющий начальному условию, определяется формулой

$$Y^{(2)}(t) = Y_0 e^{-\frac{t-t_0}{T}}. \quad (7.2.49)$$

Дисперсия этого интеграла определяется формулой

$$D_{y^{(2)}}(t) = D_{y_0} e^{-2\frac{t-t_0}{T}}. \quad (7.2.50)$$

Дисперсия $D_{y(1)}(t)$ выходной переменной системы под действием входного случайного возмущения при нулевом начальном условии определяется формулой (7.2.45), выведенной в предыдущем примере. Суммируя дисперсии $D_{y(1)}$ и $D_{y(2)}$, получим следующее выражение дисперсии выходной переменной рассматриваемой системы при случайном начальном условии:

$$D_y(t) = D_{y0} e^{-2\frac{t-t_0}{T}} + \frac{k^2 D\alpha}{\pi} \int_{-\infty}^{\infty} \frac{|e^{i\omega t} - e^{-\frac{t-t_0}{T} + i\omega t_0}|^2}{(\alpha^2 + \omega^2)(1 + T^2 \omega^2)} d\omega. \quad (7.2.51)$$

Пример 7.2.6. Найти математическое ожидание и дисперсию выходной переменной аperiodического звена, на которое действует нестационарное случайное возмущение $X(t)$, представляющее собой произведение данной функции $f(t)$ на стационарную случайную функцию времени $X_1(t)$ с показательной корреляционной функцией, которая выражается формулой (7.2.11), и математическим ожиданием m_{X_1} . В данном случае легко находится интегральное каноническое представление входного случайного возмущения. Так как любая стационарная случайная функция выражается интегральным каноническим представлением в виде интеграла Фурье (7.2.41), то случайная функция $X(t)$ выражается интегральным каноническим представлением вида

$$X(t) = m_{X_1} f(t) + \int_{-\infty}^{\infty} V(\omega) f(t) e^{i\omega t} d\omega, \quad (7.2.52)$$

Координатные функции этого интегрального канонического представления определяются формулой

$$x(t, \omega) = f(t) e^{i\omega t}, \quad (7.2.53)$$

т. е. представляют собой произведения показательных функций на одну и ту же функцию $f(t)$.

Подставляя выражение (4.4.33) весовой функции аperiodического звена и выражение $m_{X_1} f(t)$ математического ожидания входной переменной в формулу (7.2.2), находим математическое ожидание выходной переменной Y :

$$m_y(t) = \frac{k m_{X_1}}{T} \int_{t_0}^t e^{-\frac{t-\tau}{T}} f(\tau) d\tau. \quad (7.2.54)$$

Подставляя выражение (4.4.33) весовой функции и выражение (7.2.53) координатных функций $x(t, \omega)$ в формулу (7.2.38), находим координатные функции выходной переменной Y :

$$y(t, \omega) = \frac{k}{T} \int_{t_0}^t e^{-\frac{t-\tau}{T} + i\omega \tau} f(\tau) d\tau. \quad (7.2.55)$$

Эти координатные функции можно также определить как интеграл дифференциального уравнения

$$T y'_t(t, \omega) + y(t, \omega) = k f(t) e^{i\omega t}, \quad (7.2.56)$$

удовлетворяющий нулевому начальному условию $y(t_0, \omega) = 0$.

Чтобы дать пример аналитического решения задачи до конца, рассмотрим теперь частный случай, когда

$$f(t) = e^{i\omega t}. \quad (7.2.57)$$

Подставляя это выражение в формулы (7.2.54) и (7.2.55), получим

$$m_y(t) = \frac{km_{x1}}{T} \int_{t_0}^t e^{-\frac{t-\tau}{T} + \mu\tau} d\tau = \frac{km_{x1} \left[e^{\mu t} - e^{-\frac{t-t_0}{T} + \mu t_0} \right]}{1 + T\mu}. \quad (7.2.58)$$

$$y(t, \omega) = \frac{k}{T} \int_{t_0}^t e^{-\frac{t-\tau}{T} + (\mu + i\omega)\tau} d\tau = \frac{k \left[e^{(\mu + i\omega)t} - e^{-\frac{t-t_0}{T} + (\mu + i\omega)t_0} \right]}{1 + T(\mu + i\omega)}. \quad (7.2.59)$$

Подставляя последнее выражение в формулу (7.2.43), получим следующую формулу для дисперсии выходной переменной Y :

$$\begin{aligned} D_y(t) &= \frac{k^2 D\alpha}{\pi} \int_{-\infty}^{\infty} \frac{\left| e^{(\mu + i\omega)t} - e^{-\frac{t-t_0}{T} + (\mu + i\omega)t_0} \right|^2}{(\alpha^2 + \omega^2) [(1 + T\mu)^2 + T^2\omega^2]} d\omega = \\ &= \frac{k^2 D\alpha}{\pi} \int_{-\infty}^{\infty} \frac{e^{2\mu t} - 2e^{-\frac{t-t_0}{T} + \mu(t+t_0)} \cos \omega(t-t_0) + e^{-2\frac{t-t_0}{T} + 2\mu t_0}}{(\alpha^2 + \omega^2) [(1 + T\mu)^2 + T^2\omega^2]} d\omega. \end{aligned} \quad (7.2.60)$$

Дисперсия выходной переменной системы характеризует случайный разброс выходной переменной этой системы и может служить мерой случайных ошибок системы. Для оценки систематической ошибки системы необходимо вычесть из математического ожидания ее выходной переменной требуемый выходной полезный сигнал. Предположим, что система должна выполнять заданную линейную операцию L над входным полезным сигналом $m_x(t)$. Тогда требуемым выходным сигналом системы будет $Lm_x(t)$ и систематическая ошибка системы определится формулой

$$\varepsilon(t) = m_y(t) - Lm_x(t). \quad (7.2.61)$$

В частном случае следящей системы требуемый выходной сигнал системы тождественно равен полезному входному сигналу $m_x(t)$ и формула (7.2.61) принимает вид

$$\varepsilon(t) = m_y(t) - m_x(t). \quad (7.2.62)$$

Изложенные методы исследования точности линейных систем применимы и к многомерным системам. Если на входах многомерной линейной системы действуют независимые случайные возмущения, то следует определить одним из изложенных методов математические ожидания и дисперсии выходных переменных для каждого входного возмущения в отдельности. Тогда математическое ожидание и дисперсия каждой выходной переменной

при одновременном действии всех входных случайных возмущений будут равны соответственно сумме математических ожиданий и сумме дисперсий составляющих выходной переменной, вызванных действием различных входных возмущений.

Если случайные возмущения, действующие на различных входах, коррелированы, то, учитывая сказанное в конце § 2.2, векторный входной сигнал $Y(t)$ следует выразить через векторный входной сигнал $X(t)$ и матрицу весовых функций системы $g(t, \tau)$ той же формулой (7.2.1). Отсюда следует справедливость формулы (7.2.2) для математического ожидания выходного сигнала для многомерных линейных систем. Повторив почти буквально выкладки, приведшие нас к формуле (7.2.6) (с учетом особенностей матричной алгебры), легко убеждаемся в том, что матрица корреляционных и взаимных корреляционных функций составляющих выходного сигнала многомерной линейной системы $K_y(t, t')$ определяется формулой

$$K_y(t, t') = \int_{t_0}^t \int_{t_0}^{t'} g(t, \tau) K_x(\tau, \tau') g(t', \tau')^T d\tau d\tau', \quad (7.2.63)$$

где $K_x(t, t')$ — матрица корреляционных и взаимных корреляционных функций составляющих входного сигнала, а верхний индекс «т» означает операцию транспонирования матрицы [83]. Как и следовало ожидать, в частном случае одномерной системы формула (7.2.63) совпадает с (7.2.6).

Предоставляем читателю самостоятельно убедиться в том, что формулы (7.2.6) и (7.2.63) остаются справедливыми, если заменить в них центральные моменты второго порядка $K_x(\tau, \tau')$ и $K_y(t, t')$ начальными моментами второго порядка $\Gamma_x(\tau, \tau')$ и $\Gamma_y(t, t')$ соответственно.

Аналогично обобщается формула (7.2.9) на многомерные дискретные линейные системы.

Совершенно так же убеждаемся в том, что формулы (7.2.17) — (7.2.19), (7.2.21), (7.2.37) — (7.2.39) справедливы для многомерных линейных систем, если понимать в них векторные координатные функции $x_v(t)$, $y_v(t)$, $x(t, \lambda)$, $y(t, \lambda)$ как матрицы-столбцы, а черту как операцию транспонирования матрицы с одновременной заменой всех ее комплексных элементов соответствующими сопряженными величинами.

Заметим в заключение, что все выведенные формулы легко распространяются на физически невозможные линейные системы. Для этого достаточно заменить интегрирование по переменным τ, τ' по конечному интервалу интегрированием по всей числовой оси.

Более полное изложение методов исследования точности многомерных линейных систем читатель может найти в книге В. С. Пугачева [53].

§ 7.3. Определение установившихся систематических ошибок стационарных линейных систем

Применим изложенный в предыдущем параграфе метод определения математических ожиданий выходных переменных линейных систем к устойчивым стационарным линейным системам, работающим в установившемся режиме (т. е. когда переходный процесс закончился).

Для определения установившегося математического ожидания стационарной линейной системы следует положить в формуле (7.2.2) $t_0 = -\infty$. Тогда, принимая во внимание, что весовая функция стационарной линейной системы зависит только от разности ее аргументов, получим

$$m_y(t) = \int_{-\infty}^t w(t-\tau) m_x(\tau) d\tau = \int_0^{\infty} w(\xi) m_x(t-\xi) d\xi. \quad (7.3.1)$$

Разложим функцию $m_x(t-\xi)$ в ряд Тейлора с центром в точке t :

$$m_x(t-\xi) = \sum_{r=0}^{\infty} (-1)^r \frac{m_x^{(r)}(t)}{r!} \xi^r. \quad (7.3.2)$$

Подставляя это выражение в (7.3.1), получим

$$m_y(t) = \sum_{r=0}^{\infty} (-1)^r \frac{m_x^{(r)}(t)}{r!} \int_0^{\infty} \xi^r w(\xi) d\xi. \quad (7.3.3)$$

Величины

$$\mu_r = \int_0^{\infty} \xi^r w(\xi) d\xi \quad (r = 0, 1, 2, \dots) \quad (7.3.4)$$

обычно называются *моментами* весовой функции стационарной линейной системы.

Покажем, что моменты весовой функции очень просто выражаются через значения передаточной функции системы и ее производных в начале координат. Для этого воспользуемся формулой (2.4.6), на основании которой

$$\int_0^{\infty} w(\xi) e^{-s\xi} d\xi = \Phi(s). \quad (7.3.5)$$

Полагая здесь $s = 0$, находим момент нулевого порядка весовой функции стационарной линейной системы:

$$\mu_0 = \int_0^{\infty} w(\xi) d\xi = \Phi(0). \quad (7.3.6)$$

Дифференцируя формулу (7.3.5) r раз по s , получим

$$(-1)^r \int_0^{\infty} \xi^r w(\xi) e^{-s\xi} d\xi = \Phi^{(r)}(s).$$

Полагая здесь $s = 0$, находим момент r -го порядка весовой функции стационарной линейной системы:

$$\mu_r = \int_0^{\infty} \xi^r w(\xi) d\xi = (-1)^r \Phi^{(r)}(0) \quad (r = 1, 2, \dots). \quad (7.3.7)$$

Подставляя найденные выражения моментов весовой функции в формулу (7.3.3), получим

$$m_y(t) = \sum_{r=0}^{\infty} \frac{\Phi^{(r)}(0)}{r!} m_x^{(r)}(t). \quad (7.3.8)$$

Обозначим теперь через $\Phi_T(s)$ передаточную функцию идеальной стационарной линейной системы, точно осуществляющей требуемое преобразование L входного полезного сигнала. Применяя к этой системе формулу (7.3.8), получим следующее выражение требуемого выходного сигнала:

$$Lm_x(t) = \sum_{r=0}^{\infty} \frac{\Phi_T^{(r)}(0)}{r!} m_x^{(r)}(t). \quad (7.3.9)$$

Систематическая ошибка системы определяется согласно (7.2.61) формулой

$$e(t) = m_y(t) - Lm_x(t) = \sum_{r=0}^{\infty} \frac{1}{r!} [\Phi^{(r)}(0) - \Phi_T^{(r)}(0)] m_x^{(r)}(t). \quad (7.3.10)$$

Величины

$$c_r = \frac{1}{r!} [\Phi^{(r)}(0) - \Phi_T^{(r)}(0)] \quad (r = 0, 1, 2, \dots) \quad (7.3.11)$$

обычно называются *коэффициентами ошибок*. Формула (7.3.10) для случая полезных сигналов, представляющих собой полиномы не выше n -й степени относительно времени, принимает на основании (7.3.11) вид

$$e(t) = \sum_{r=0}^n c_r m_x^{(r)}(t). \quad (7.3.12)$$

Первые три коэффициента ошибок в соответствии с их физическим смыслом имеют еще специальные названия.

Величина

$$c_0 = \Phi(0) - \Phi_T(0) \quad (7.3.13)$$

называется *коэффициентом статической ошибки* или *коэффициентом ошибки по положению*. Системы, для которых $c_0 \neq 0$, отрабатывают постоянный сигнал с установившейся постоянной ошибкой, отличной от нуля. В самом деле, при $m_x(t) = x_0 = \text{const}$ формула (7.3.12) дает $e = c_0 x_0$.

Системы, для которых $c_0 = 0$, называются *астатическими*. Астатические системы отрабатывают любой постоянный сигнал в установившемся режиме без систематической ошибки.

Величина

$$c_1 = \Phi'(0) - \Phi'_T(0) \quad (7.3.14)$$

называется *коэффициентом ошибки по скорости*. Если $c_0 = 0$, $c_1 \neq 0$, то система отрабатывает сигнал, изменяющийся с постоянной скоростью, с постоянной установившейся ошибкой. Действительно, при $m_x(t) = x_0 + vt$, $c_0 = 0$, $c_1 \neq 0$ установившаяся систематическая ошибка системы согласно (7.3.12) равна $e = c_1 v$.

Наконец, величина

$$c_2 = \frac{1}{2} [\Phi''(0) - \Phi''_T(0)] \quad (7.3.15)$$

называется *коэффициентом ошибки по ускорению*. Если $c_0 = c_1 = 0$, $c_2 \neq 0$, то система отрабатывает сигнал, изменяющийся с постоянным ускорением, с постоянной установившейся ошибкой, так как при $m_x(t) = x_0 + vt + \frac{1}{2} at^2$, $c_0 = c_1 = 0$, $c_2 \neq 0$ установившаяся систематическая ошибка системы согласно (7.3.12) равна $e = c_2 a$.

Для следящей системы $\Phi_T(s) = 1$ и формулы (7.3.13), (7.3.14), (7.3.15) и (7.3.11) для коэффициентов ошибок принимают вид

$$c_0 = \Phi(0) - 1, \quad c_r = \frac{1}{r!} \Phi^{(r)}(0) \quad (r = 1, 2, \dots). \quad (7.3.16)$$

Отсюда видно, что для астатической следящей системы всегда должно быть выполнено условие $\Phi(0) = 1$, т. е. значение передаточной функции (а следовательно, и частотной характеристики) в начале координат должно быть равно единице.

Формула (7.3.12) показывает, что если первые $k+1$ коэффициентов ошибок равны нулю: $c_0 = c_1 = \dots = c_k = 0$, то система отрабатывает без установившейся ошибки любые входные сигналы, представляющие собой полиномы не выше чем k -й степени. Чем большее количество первых коэффициентов ошибок равно нулю, тем более разнообразные полезные сигналы система отрабатывает без установившейся ошибки, т. е. тем совершеннее система с точки зрения передачи полезных сигналов при отсутствии помех. Поэтому среди астатических систем различают астатические системы различных порядков.

Астатической системой первого порядка называется система, для которой $c_0 = 0$, $c_1 \neq 1$. *Астатической системой второго порядка* называется система, для которой $c_0 = c_1 = 0$, $c_2 \neq 0$. И вообще *астатической системой k -го порядка* называется система, для которой $c_0 = c_1 = \dots = c_{k-1} = 0$, $c_k \neq 0$. Астатические системы первого порядка обрабатывают без установившейся ошибки любые постоянные сигналы. Астатические системы второго порядка обрабатывают без установившейся ошибки любые постоянные сигналы и любые линейные функции времени. Вообще астатические системы k -го порядка обрабатывают без установившейся ошибки любые сигналы, представляющие собой постоянные или полиномы относительно времени не выше $k - 1$ -й степени.

Пример 7.3.1. Фильтр RC с постоянной времени $RC = T$ предназначен для отфильтровывания от помехи различных полезных сигналов, представляющих собой постоянные величины, или линейные функции времени, или квадратные трехчлены относительно времени. Найти коэффициенты ошибок фильтра и установившиеся систематические ошибки. Передаточная функция фильтра определяется формулой

$$\Phi(s) = \frac{1}{1 + Ts}. \quad (7.3.17)$$

Дифференцируя эту формулу и полагая после этого $s = 0$, получим

$$\Phi(0) = 1, \quad \Phi'(0) = -T, \quad \Phi''(0) = 2T^2.$$

По своему назначению рассматриваемый фильтр является следящей системой. Поэтому коэффициенты ошибок определяются для него по формулам (7.3.16). В результате находим

$$c_0 = \Phi(0) - 1 = 0, \quad c_1 = \Phi'(0) = -T, \quad c_2 = \frac{1}{2} \Phi''(0) = T^2. \quad (7.3.18)$$

Таким образом, рассматриваемый фильтр является астатической системой первого порядка. Для случая полезных сигналов, представляющих собой квадратные полиномы относительно времени:

$$m_x(t) = x_0 + vt + \frac{1}{2} at^2, \quad (7.3.19)$$

где величины x_0 , v и a представляют собой соответственно начальное значение, начальную скорость (т. е. начальное значение первой производной) и начальное ускорение (т. е. вторую производную) полезного сигнала, формула (7.3.12) дает следующее выражение установившейся систематической ошибки фильтра:

$$e(t) = -T(v + at) + aT^2. \quad (7.3.20)$$

Пример 7.3.2. Найти коэффициенты ошибок по положению, скорости и ускорению для следящей системы, исполнительное устройство которой представляет собой двигатель постоянного тока с электромеханической постоянной времени T_2 , а корректирующее дифференцирующее устройство включено в прямую цепь. В этом случае система состоит из последовательно соединенных дифференцирующего звена с передаточной функцией

$$\Phi_1(s) = k_1 \frac{1 + \tau_1 s}{1 + T_1 s},$$

аперiodического звена с передаточной функцией

$$\Phi_2(s) = \frac{k_2}{1 + T_2 s}$$

и интегрирующего звена, охваченных жесткой отрицательной обратной связью (рис. 7.3.1). Первое из этих звеньев преобразует параметр управления следящей системы, представляющий собой разность между входной и выходной переменными системы. Второе вырабатывает скорость изменения выходной переменной по данному преобразованному (скорректированному) параметру управления. Третье звено вырабатывает выходной сигнал следящей

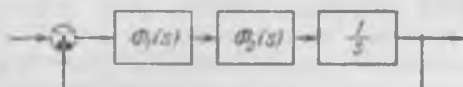


Рис. 7.3.1.

системы путем интегрирования скорости его изменения. Само собой разумеется, второе и третье звенья физически представляют собой одно целое — исполнительное устройство следящей системы.

Передаточная функция рассматриваемой следящей системы, согласно формулам § 4.6, равна

$$\Phi(s) = \frac{k_1 k_2 (1 + \tau_1 s)}{k_1 k_2 (1 + \tau_1 s) + s (1 + T_1 s) (1 + T_2 s)}. \quad (7.3.21)$$

Дифференцируя эту формулу дважды, полагая $s = 0$ и пользуясь формулами (7.3.16), находим коэффициенты ошибок рассматриваемой следящей системы:

$$e_0 = 0, \quad e_1 = -\frac{1}{k_1 k_2}, \quad e_2 = \frac{1}{k_1 k_2} \left(\frac{1}{k_1 k_2} + \tau_1 - T_1 - T_2 \right).$$

Наконец, пользуясь формулой (7.3.12), находим установившуюся систематическую ошибку рассматриваемой следящей системы для случая входного полезного сигнала, представляющего собой квадратный трехчлен (7.3.19) относительно времени t :

$$e(t) = -\frac{v + at}{k_1 k_2} + \frac{a}{k_1 k_2} \left(\frac{1}{k_1 k_2} + \tau_1 - T_1 - T_2 \right). \quad (7.3.22)$$

Легко видеть, что для того, чтобы замкнутая следящая система была астатической k -го порядка, достаточно, чтобы передаточная функция разомкнутой системы $\Phi_1(s)$ имела в начале координат полюс k -го порядка:

$$\Phi_1(s) = \frac{1}{s^k} \Psi_1(s), \quad \Psi_1(0) \neq 0. \quad (7.3.23)$$

Действительно, из (4.6.8) следует, что передаточная функция замкнутой следящей системы в данном случае выражается формулой

$$\Phi(s) = \frac{\Psi_1(s)}{s^k + \Psi_1(s)} = \frac{1}{1 + \frac{s^k}{\Psi_1(s)}}. \quad (7.3.24)$$

Так как $\Psi_1(s)$ — аналитическая функция в окрестности начала координат и $\Psi_1(0) \neq 0$, то при достаточно малых s правая часть формулы (7.3.24) может быть представлена в виде сходящейся геометрической прогрессии

$$\Phi(s) = 1 - \frac{s^k}{\Psi_1(s)} + \frac{s^{2k}}{\Psi_1^2(s)} - \dots$$

Дифференцируя эту формулу k раз и полагая в полученных формулах и в (7.3.24) $s = 0$, убеждаемся в том, что

$$\Phi(0) = 1, \quad \Phi'(0) = \dots = \Phi^{(k-1)}(0) = 0, \quad \Phi^{(k)}(0) = \frac{k!}{\Psi_1(0)}.$$

Подставляя эти выражения в (7.3.16), убеждаемся в том, что в данном случае $c_0 = c_1 = \dots = c_{k-1} = 0$, $c_k \neq 0$, что и доказывает высказанное утверждение.

Заметим, что для справедливости формул (7.3.8) и (7.3.10) необходимо, чтобы входной полезный сигнал $m_x(t)$ мог быть представлен рядом Тейлора (7.3.2), сходящимся при всех значениях ξ . Поэтому практически формулы (7.3.8) и (7.3.10) применимы только для сигналов, представляющих собой полиномы относительно времени. При этом ввиду того, что для сложных систем частотные характеристики обычно определяются приближенными методами, чаще всего графически, точное определение производных передаточной функции по существу невозможно. Вследствие этого практическое применение формул (7.3.8) и (7.3.10) обычно ограничивается случаями постоянных полезных сигналов или сигналов, представляющих собой полиномы первой или второй степени относительно времени.

Установившиеся систематические ошибки стационарной линейной системы можно вычислить также другим методом, если известна частотная характеристика системы.

Предположим, что входной полезный сигнал $m_x(t)$ можно приближенно представить тригонометрическим полиномом

$$m_x(t) = a_0 + \sum_{r=1}^n (a_r \cos \omega_r t + b_r \sin \omega_r t). \quad (7.3.25)$$

Тогда на основании принципа суперпозиции установившееся значение математического ожидания выходной переменной системы определится формулой

$$m_y(t) = \Phi(0) a_0 + \sum_{r=1}^n |\Phi(i\omega_r)| \{a_r \cos[\omega_r t + \arg \Phi(i\omega_r)] + b_r \sin[\omega_r t + \arg \Phi(i\omega_r)]\}. \quad (7.3.26)$$

Таким образом, зная амплитудную и фазовую частотные характеристики стационарной линейной системы, можно весьма просто

находить ее установившиеся систематические ошибки для любых входных полезных сигналов, которые могут быть приближенно представлены тригонометрическими полиномами.

Формула (7.3.26) легко обобщается на случай сигналов, которые можно с достаточной точностью представить суммой синусов с амплитудами, изменяющимися по показательному закону. В этом случае

$$m_x(t) = a_0 e^{\mu_0 t} + \sum_{r=1}^n e^{\mu_r t} (a_r \cos \omega_r t + b_r \sin \omega_r t) \quad (7.3.27)$$

и на основании принципа суперпозиции установившееся значение математического ожидания выходной переменной системы определится формулой

$$\begin{aligned} m_y(t) = \\ = \Phi(\mu_0) a_0 e^{\mu_0 t} + \sum_{r=1}^n |\Phi(\mu_r + i\omega_r)| e^{\mu_r t} \{a_r \cos[\omega_r t + \arg \Phi(\mu_r + i\omega_r)] + \\ + b_r \sin[\omega_r t + \arg \Phi(\mu_r + i\omega_r)]\}. \end{aligned} \quad (7.3.28)$$

Для практического вычисления установившихся систематических ошибок стационарных линейных систем формулы (7.3.26) и (7.3.28) во многих случаях оказываются удобнее, чем формула (7.3.8).

§ 7.4. Определение установившейся дисперсии выходной переменной стационарной линейной системы

Из теории случайных функций известно, что любая стационарная случайная функция $X(t)$ выражается интегральным каноническим представлением, координатными функциями которого являются функции $e^{i\omega t}$, описывающие гармонические колебания всех возможных частот:

$$X(t) = m_x + \int_{-\infty}^{\infty} V(\omega) e^{i\omega t} d\omega, \quad (7.4.1)$$

где $V(\omega)$ — белый шум (т. е. случайная функция ω с некоррелированными значениями при различных значениях ω), интенсивность которого равна спектральной плотности $s_x(\omega)$ случайной функции $X(t)$ (см. [54], §§ 5.2, 5.3 и 6.5 или [53], § 76). Вследствие этого для исследования точности стационарной линейной системы, работающей в установившемся режиме под действием входного возмущения, представляющего собой стационарную случайную функцию времени, естественно воспользоваться методом интегральных канонических представлений и методом частотных характеристик.

Координатная функция $y(t, \omega)$ выходной переменной, как установившаяся реакция стационарной линейной системы на гармонические колебания $e^{i\omega t}$, определяется формулой

$$y(t, \omega) = \Phi(i\omega) e^{i\omega t}, \quad (7.4.2)$$

где $\Phi(i\omega)$ — частотная характеристика системы. Подставляя выражение (7.4.2) в формулы (7.2.39) и (7.2.40) и принимая во внимание, что интенсивность $G(\omega)$ белого шума $V(\omega)$ в данном случае равна спектральной плотности $s_x(\omega)$ стационарной случайной функции $X(t)$, получим следующие формулы для корреляционной функции и дисперсии выходной переменной стационарной линейной системы:

$$\begin{aligned} r_y(t, t') &= \int_{-\infty}^{\infty} s_x(\omega) \Phi(i\omega) e^{i\omega t} \overline{\Phi(i\omega)} e^{-i\omega t'} d\omega = \\ &= \int_{-\infty}^{\infty} s_x(\omega) |\Phi(i\omega)|^2 e^{i\omega(t-t')} d\omega, \end{aligned} \quad (7.4.3)$$

$$D_y(t) = \int_{-\infty}^{\infty} s_x(\omega) |\Phi(i\omega)|^2 d\omega. \quad (7.4.4)$$

Эти формулы показывают, что при неограниченно долгом действии стационарного случайного возмущения на стационарную линейную систему дисперсия ее выходной переменной постоянна, а корреляционная функция зависит только от разности аргументов. Следовательно, выходная переменная стационарной линейной системы, работающей неограниченно долго под действием стационарного случайного возмущения, является стационарной случайной функцией времени. Из формул (7.4.3) и (7.4.4) следует, что спектральная плотность выходной переменной стационарной линейной системы равна произведению квадрата модуля ее частотной характеристики и спектральной плотности входного стационарного случайного возмущения:

$$s_y(\omega) = s_x(\omega) |\Phi(i\omega)|^2. \quad (7.4.5)$$

Для вычисления установившихся дисперсий выходных сигналов стационарных линейных систем по формуле (7.4.4) в случае рациональных функций $s_x(\omega)$ и $\Phi(s)$ удобно пользоваться формулой

$$\int_{-\infty}^{\infty} \frac{b_0(i\omega)^{2n-2} + b_1(i\omega)^{2n-4} + \dots + b_{n-2}(i\omega)^2 + b_{n-1}}{|a_0(i\omega)^n + a_1(i\omega)^{n-1} + \dots + a_{n-1}(i\omega) + a_n|^2} d\omega = (-1)^{n+1} \frac{\pi}{a_0} \frac{D_n}{\Delta_n}, \quad (7.4.6)$$

где

$$\Delta_n = \begin{vmatrix} c_{11} & c_{12} & \dots & c_{1n} \\ c_{21} & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ c_{n1} & c_{n2} & \dots & c_{nn} \end{vmatrix}, \quad D_n = \begin{vmatrix} b_0 & c_{12} & \dots & c_{1n} \\ b_1 & c_{22} & \dots & c_{2n} \\ \dots & \dots & \dots & \dots \\ b_{n-1} & c_{n2} & \dots & c_{nn} \end{vmatrix}, \quad (7.4.7)$$

$$\left. \begin{aligned} c_{pq} &= a_{2p-q} \quad \text{при} \quad 0 \leq 2p-q \leq n, \\ c_{pq} &= 0 \quad \text{при} \quad 2p-q < 0 \quad \text{и} \quad \text{при} \quad 2p-q > n. \end{aligned} \right\} \quad (7.4.8)$$

Для применения формулы (7.4.6) следует представить знаменатель дроби $s_x(\omega) / |\Phi(i\omega)|^2$ в виде квадрата модуля полинома относи-

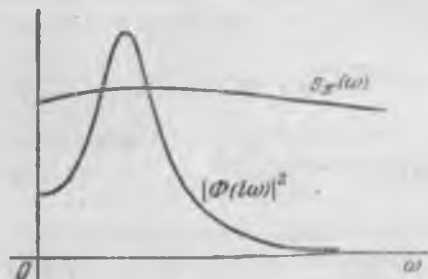


Рис. 7.4.1.

тельно $i\omega$ с неотрицательными коэффициентами (см. § 7.6), а числитель в виде полинома относительно $i\omega$ (который всегда содержит только четные степени $i\omega$). При этом определяются коэффициенты $a_0, a_1, \dots, a_n, b_0, b_1, \dots, b_{n-1}$, после чего можно будет применить формулы (7.4.6)–(7.4.8).

В задачах практики стационарная линейная система часто имеет полосу пропускания, узкую по сравнению с полосой

частот входного возмущения. В таких случаях часто оказывается возможным считать спектральную плотность входного возмущения $s_x(\omega)$ постоянной в пределах полосы пропускания системы и равной, например, ее значению $s_x(\omega_p)$ при резонансной частоте системы ω_p . Для иллюстрации на рис. 7.4.1 приведены графики функций $|\Phi(i\omega)|^2$ и $s_x(\omega)$ для этого случая. В этом случае формула (7.4.4) может быть заменена приближенной формулой

$$D_v \approx s_x(\omega_p) \Delta\omega, \quad (7.4.9)$$

где

$$\Delta\omega = \int_{-\infty}^{\infty} |\Phi(i\omega)|^2 d\omega. \quad (7.4.10)$$

Эта величина обычно называется *эффективной полосой пропускания* стационарной линейной системы. Таким образом, во многих задачах исследование точности стационарной линейной системы практически сводится к определению ее эффективной полосы пропускания $\Delta\omega$.

Заметим, что применение формулы (7.4.9) по существу равноценно замене действующего на систему стационарного случайного возмущения белым шумом с постоянной спектральной плот-

ностью $s_x(\omega_p)$ и интенсивностью $2\pi s_x(\omega_p)$ *). Таким образом, при исследовании действия широкополосного шума на узкополосную стационарную линейную систему можно с достаточной для практики точностью считать широкополосный шум белым.

Заметим, что формула (7.4.2) справедлива только для случая бесконечно долгого действия гармонических колебаний $e^{i\omega t}$ на устойчивую стационарную линейную систему. Поэтому и формулы (7.4.3), (7.4.4) и (7.4.9) применимы только к случаю бесконечно долгого действия стационарного случайного возмущения на устойчивую стационарную линейную систему. Практически это означает, что формулы (7.4.2), (7.4.3), (7.4.4) и (7.4.9) можно применять только к устойчивым стационарным линейным системам, проработавшим под действием стационарных случайных возмущений достаточно долго (в течение времени, превосходящего время переходного процесса). К неустойчивым системам, а также к устойчивым стационарным линейным системам, работающим в переходных режимах, формулы (7.4.2), (7.4.3), (7.4.4) и (7.4.9) неприменимы. Во всех подобных случаях следует пользоваться общими методами, изложенными и проиллюстрированными примерами в § 7.2. При этом, если поведение стационарной линейной системы описывается дифференциальными уравнениями, то и координатные функции выходной переменной $y(t, \omega)$ будут представлять собой интегралы дифференциальных уравнений, полученных путем замены входного возмущения в уравнениях системы гармоническими колебаниями $e^{i\omega t}$, удовлетворяющие нулевым начальным условиям. Эти интегралы будут представлять собой не чистые гармонические колебания, а результат наложения свободных колебаний на соответствующие гармонические колебания. При этом выходная переменная рассматриваемой системы, как мы видели на примерах § 7.2, уже не будет стационарной случайной функцией. Поэтому теория стационарных случайных функций принципиально недостаточна для исследования точности стационарных линейных систем, работающих под действием стационарных случайных возмущений. Это принципиальное ограничение теории стационарных случайных функций является ее существенным недостатком и вызывает необходимость изучения и практического применения теории нестационарных случайных функций.

Пример 7.4.1. Решить задачу, рассмотренную в примерах 7.2.1 и 7.2.4, для случая неограниченно долгого действия возмущения ($t_0 = -\infty$). Подставляя в формулу (7.4.4) выражение (2.6.4) частотной характеристики аperiodического звена и выражение (7.2.42) спектральной плотности $s_x(\omega)$,

*) Из теории случайных функций известно, что стационарная случайная функция с постоянной спектральной плотностью, равной s_0 , является стационарным белым шумом времени t , интенсивность которого равна $2\pi s_0$ (см. [54], § 5.4 или [53], § 76, пример 1).

получим

$$D_y = \frac{k^2 D \alpha}{\pi} \int_{-\infty}^{\infty} \frac{d\omega}{(\alpha^2 + \omega^2)(1 + T^2 \omega^2)}. \quad (7.4.11)$$

Применим для вычисления интеграла формулу (7.4.6). Для этого представим (7.4.11) в виде

$$D_y = \frac{k^2 D \alpha}{\pi} \int_{-\infty}^{\infty} \frac{d\omega}{|T(i\omega)^2 + (1 + \alpha T)i\omega + \alpha|^2}.$$

Отсюда видно, что в данном случае $n = 2$, $a_0 = T$, $a_1 = 1 + \alpha T$, $a_2 = \alpha$, $b_0 = 0$, $b_1 = 1$. Формулы (7.4.8) дают

$$c_{11} = a_1 = 1 + \alpha T, \quad c_{12} = a_0 = T, \quad c_{21} = 0, \quad c_{22} = a_2 = \alpha.$$

После этого по формулам (7.4.7) находим

$$\Delta_2 = \begin{vmatrix} 1 + \alpha T & T \\ 0 & \alpha \end{vmatrix} = \alpha(1 + \alpha T), \quad D_2 = \begin{vmatrix} 0 & T \\ 1 & \alpha \end{vmatrix} = -T.$$

Подставив эти выражения в (7.4.6), получим окончательно

$$D_y = \frac{k^2 D \alpha}{\pi} (-1) \frac{\pi}{T} \frac{-T}{\alpha(1 + \alpha T)} = \frac{k^2 D}{1 + \alpha T}. \quad (7.4.12)$$

Формулу (7.4.12), конечно, можно получить при $T > 0$ из более общей формулы (7.2.14), полагая в ней $t_0 = -\infty$. Однако в случае неустойчивой системы, когда $T < 0$, формула (7.2.14) дает при $t_0 = -\infty$ бесконечную дисперсию выходной переменной. Этот факт иллюстрирует высказанное выше утверждение о неприменимости формул (7.4.3) и (7.4.4) к неустойчивым системам.

Формулы (7.2.14) и (7.2.51) могут также служить иллюстрацией нашего утверждения, что выходная переменная стационарной линейной системы, работающей под действием стационарного случайного возмущения в переходном режиме, является нестационарной случайной функцией, так как ее дисперсия зависит от времени t . Дисперсия выходной переменной практически становится постоянной только при достаточно больших значениях $t - t_0$ и при $T > 0$, когда показательная функция $\exp \left\{ -\frac{t - t_0}{T} \right\}$ мала по сравнению с единицей. В этом случае дисперсия выходной переменной может быть приближенно вычислена по формуле (7.4.11) или (7.4.12). Практически формулы (7.4.11) и (7.4.12) дают достаточную точность при $t - t_0 \geq 3T$.

Приведенный пример показывает, что изложенная в § 7.2 теория дает общие методы исследования точности линейных систем, в результате применения которых получаются точные формулы, применимые к любым линейным системам, работающим как в установившихся, так и в переходных режимах. В частном случае для устойчивых стационарных линейных систем, работающих под действием стационарных случайных возмущений в установившемся режиме, изложенная теория дает простые формулы (7.4.3) и (7.4.4). Эти формулы определяют точное значение дисперсии выходной переменной системы для случая бесконечно долгого действия возмущения и могут служить для приближенного определения

дисперсии выходной переменной системы для моментов времени, удаленных от момента начала действия возмущения t_0 больше, чем на время переходного процесса.

Пример 7.4.2. Найти установившуюся дисперсию выходной переменной следящей системы, рассмотренной в примере 7.3.2, для случая, когда корреляционная функция помехи, накладывающейся на входной полезный сигнал системы, определяется формулой (7.2.11), предполагая, что величина $1/\alpha$ мала по сравнению с постоянными времени τ_1 , T_1 , T_2 (т. е. интервал корреляции входной помехи мал по сравнению с τ_1 , T_1 , T_2).

Полоса пропускания рассматриваемой следящей системы имеет тот же порядок, что и величины τ_1^{-1} , T_1^{-1} , T_2^{-1} . Диапазон изменения частоты ω , в котором спектральная плотность входного случайного возмущения $s_x(\omega)$, определяемая формулой (7.2.42), изменяется не больше чем на 10%, имеет порядок $0,3\alpha$. Так как по условию величины τ_1^{-1} , T_1^{-1} , T_2^{-1} малы по сравнению с α , то спектральную плотность входного возмущения можно считать в данном случае практически постоянной и равной D_x/α в пределах полосы пропускания следящей системы. Следовательно, дисперсию выходной переменной можно вычислить по формуле (7.4.9). Для этого находим по формуле (7.4.10) эффективную полосу пропускания следящей системы. Полагая в (7.3.21) $s = i\omega$ и подставляя полученное выражение в (7.4.10), после несложных преобразований получим

$$\Delta\omega = \int_{-\infty}^{\infty} \frac{k_1^2 k_2^2 [1 - \tau_1^2 (i\omega)^2] d\omega}{|k_1 k_2 + (1 + k_1 k_2 \tau_1) i\omega + (T_1 + T_2) (i\omega)^2 + T_1 T_2 (i\omega)^3|^2}.$$

Воспользовавшись для вычисления интеграла формулой (7.4.6), получим

$$\Delta\omega = \pi k_1 k_2 \frac{T_1 + T_2 + k_1 k_2 \tau_1^2}{(1 + k_1 k_2 \tau_1) (T_1 + T_2) - k_1 k_2 T_1 T_2}. \quad (7.4.13)$$

Подставляя это выражение и выражение D_x/α спектральной плотности в (7.4.9), получим следующую приближенную формулу для установившейся дисперсии выходной переменной рассматриваемой следящей системы:

$$D_y \approx \frac{D_x k_1 k_2}{\alpha} \cdot \frac{T_1 + T_2 + k_1 k_2 \tau_1^2}{(1 + k_1 k_2 \tau_1) (T_1 + T_2) - k_1 k_2 T_1 T_2}. \quad (7.4.14)$$

Для дальнейшего нам понадобится еще вывести формулу для взаимной корреляционной функции выходной переменной Y стационарной линейной системы с другой случайной функцией Z . Предположим, что случайная функция Z стационарна и стационарно связана со входной стационарной случайной функцией X стационарной линейной системы. Выразим случайные функции Y и Z интегральными каноническими представлениями. Принимая во внимание (7.4.1) и (7.4.2), можем написать интегральное каноническое представление выходной переменной системы Y в виде

$$Y(t) = m_y + \int_{-\infty}^{\infty} V(\omega) \Phi(i\omega) e^{i\omega t} d\omega. \quad (7.4.15)$$

Интегральное каноническое представление стационарной случай-
ной функции Z имеет вид

$$Z(t) = m_z + \int_{-\infty}^{\infty} U(\omega) e^{i\omega t} d\omega, \quad (7.4.16)$$

где $U(\omega)$ — некоторый белый шум аргумента ω . Так как действительная случайная функция Z совпадает со своей комплексной сопряженной величиной, то на основании (7.4.16) можем написать

$$Z(t) = m_z + \int_{-\infty}^{\infty} \overline{U(\omega)} e^{-i\omega t} d\omega. \quad (7.4.17)$$

Из (7.4.15) и (7.4.17) следует, что

$$\begin{aligned} K_{yz}(t, t') &= M[Y^0(t) Z^0(t')] = \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} M[V(\omega) \overline{U(\omega')}] \Phi(i\omega) e^{i(\omega t - \omega' t')} d\omega d\omega'. \end{aligned} \quad (7.4.18)$$

Но взаимная корреляционная функция белых шумов V и U выражается через взаимную спектральную плотность $s_{xz}(\omega)$ случайных функций X и Z формулой (см. [54], § 5.6 или [53], § 78)

$$K_{vu}(\omega, \omega') = M[V(\omega) \overline{U(\omega')}] = s_{xz}(\omega) \delta(\omega - \omega'). \quad (7.4.19)$$

Подставляя это выражение в (7.4.18), получим

$$K_{yz}(t, t') = \int_{-\infty}^{\infty} s_{xz}(\omega) \Phi(i\omega) e^{i\omega t} d\omega \int_{-\infty}^{\infty} e^{-i\omega' t'} \delta(\omega - \omega') d\omega'$$

или, выполняя интегрирование по ω' ,

$$K_{yz}(t, t') = \int_{-\infty}^{\infty} s_{xz}(\omega) \Phi(i\omega) e^{i\omega(t-t')} d\omega. \quad (7.4.20)$$

Формула (7.4.20) показывает, что взаимная корреляционная функция случайных функций Y и Z зависит только от разности аргументов $t - t'$ и, следовательно, случайные функции Y и Z стационарно связаны.

Сравнивая (7.4.20) с известным выражением взаимной корреляционной функции стационарных и стационарно связанных случайных функций Y и Z через их взаимную спектральную плотность ([54], § 5.6 или [53], § 78):

$$k_{yz}(t-t') = \int_{-\infty}^{\infty} s_{yz}(\omega) e^{i\omega(t-t')} d\omega,$$

получим

$$s_{yz}(\omega) = s_{xz}(\omega) \Phi(i\omega). \quad (7.4.21)$$

Таким образом, взаимная спектральная плотность выходной переменной Y стационарной линейной системы с любой стационарной (и стационарно связанной со входной переменной X) случайной функцией Z равна взаимной спектральной плотности входной переменной X со случайной функцией Z , умноженной на частотную характеристику системы. Аналогично находим

$$s_{zy}(\omega) = s_{zx}(\omega) \overline{\Phi(i\omega)} = s_{zx}(\omega) \Phi(-i\omega). \quad (7.4.22)$$

Формулы (7.4.20), (7.4.21) и (7.4.22) также применимы только к устойчивым системам и к моментам времени, удаленным от начального момента больше чем на время переходного процесса.

Из формул (7.4.21) и (7.4.22) можно, в частности, получить выражения взаимных спектральных плотностей входной и выходной переменных стационарной линейной системы. Для этого достаточно принять случайную функцию Z тождественно совпадающей со входной случайной функцией X . Тогда получим

$$\left. \begin{aligned} s_{yx}(\omega) &= s_x(\omega) \Phi(i\omega), \\ s_{xy}(\omega) &= s_x(\omega) \overline{\Phi(i\omega)} = s_x(\omega) \Phi(-i\omega). \end{aligned} \right\} \quad (7.4.23)$$

§ 7.5. Определение установившейся дисперсии выходной переменной стационарной дискретной линейной системы

Все формулы предыдущего параграфа справедливы как для непрерывных, так и для дискретных линейных систем для значений t и t' , совпадающих с моментами действия импульсов $t_k = kT_n$ ($k = 0, \pm 1, \pm 2, \dots$). Однако для определения установившейся дисперсии выходной переменной стационарной дискретной линейной системы удобнее воспользоваться интегральным каноническим представлением стационарной случайной функции $X(t)$ в дискретном ряде равноотстоящих точек:

$$\begin{aligned} X(t_k) = X(kT_n) &= m_x + \int_0^{2\pi/T_n} V_d(\omega) e^{i\omega k T_n} d\omega \quad (7.5.1) \\ (k &= 0, \pm 1, \pm 2, \dots), \end{aligned}$$

где $V_d(\omega)$ — белый шум, интенсивность которого равна спектральной плотности $s_x^*(\omega)$ стационарной случайной последовательности $X(t_k)$, а T_n — период повторения импульсов (такт) исследуемой дискретной линейной системы (см. [54], § 5.7). Дискретные координатные функции входного случайного возмущения выражаются в данном случае формулой

$$x_k(\omega) = x(t_k, \omega) = e^{i\omega k T_n} \quad (k = 0, \pm 1, \pm 2, \dots). \quad (7.5.2)$$

Координатные функции выходной переменной системы в установленном режиме выразятся аналогичной формулой

$$y_k(\omega) = y(t_k, \omega) = \Phi(i\omega) e^{i\omega k T_n} = \Psi(e^{i\omega T_n}) e^{i\omega k T_n} \quad (7.5.3)$$

$$(k=0, \pm 1, \pm 2, \dots).$$

Подставляя это выражение в формулы (7.2.39) и (7.2.40) и принимая во внимание, что в данном случае

$$\lambda = \omega, \quad \lambda_1 = 0, \quad \lambda_2 = \frac{2\pi}{T_n}, \quad G(\omega) = s_x^d(\omega),$$

получим следующие формулы для корреляционной функции и дисперсии выходной переменной дискретной линейной системы:

$$k_y(t_k - t_l) = k_y((h-l)T_n) = \int_0^{2\pi/T_n} s_x^d(\omega) |\Phi(i\omega)|^2 e^{i\omega(h-l)T_n} d\omega, \quad (7.5.4)$$

$$D_y = \int_0^{2\pi/T_n} s_x^d(\omega) |\Phi(i\omega)|^2 d\omega. \quad (7.5.5)$$

Интегрирование в формулах (7.5.4) и (7.5.5) производится вдоль отрезка мнимой оси на плоскости комплексной переменной s от 0 до точки $2\pi i/T_n$. При переходе к переменной $z = e^{sT_n}$ этот отрезок перейдет в единичную окружность C_1 плоскости z . Поэтому при замене переменных $z = e^{i\omega T_n}$ формулы (7.5.4) и (7.5.5) примут вид

$$k_y(mT_n) = \frac{1}{i} \int_{C_1} \sigma_x^d(z) |\Psi(z)|^2 z^{m-1} dz, \quad (7.5.6)$$

$$D_y = \int_{C_1} \sigma_x^d(z) |\Psi(z)|^2 \frac{dz}{iz}, \quad (7.5.7)$$

где

$$\sigma_x^d(z) = \frac{1}{T_n} s_x^d \left(\frac{1}{iT_n} \ln z \right), \quad (7.5.8)$$

а интегрирование производится вдоль единичной окружности.

Если перейти в формулах (7.5.4)–(7.5.7) к переменной $\lambda = v/i = (z-1)/i(z+1) = (e^{i\omega T_n} - 1)/i(e^{i\omega T_n} + 1)$, то, учитывая, что единичной окружности плоскости комплексной переменной z соответствует вся мнимая ось плоскости переменной v ,

получим

$$k_y(mT_n) = \int_{-\infty}^{\infty} \bar{s}_x(\lambda) |\Omega(i\lambda)|^2 \left(\frac{1+i\lambda}{1-i\lambda} \right)^m \frac{d\lambda}{1+\lambda^2} =$$

$$= \int_{-\infty}^{\infty} \bar{s}_x(\lambda) |\Omega(i\lambda)|^2 \sum_{h=0}^m C_{2m}^{2h} (i\lambda)^{2h} \frac{d\lambda}{(1+\lambda^2)^{m+1}}, \quad (7.5.9)$$

$$D_y = \int_{-\infty}^{\infty} \bar{s}_x(\lambda) |\Omega(i\lambda)|^2 \frac{d\lambda}{1+\lambda^2}, \quad (7.5.10)$$

где

$$s_x(\lambda) = 2\sigma_x^d \left(\frac{1+i\lambda}{1-i\lambda} \right) = \frac{2}{T_n} s_x^d \left(\frac{2}{T_n} \operatorname{arctg} \lambda \right). \quad (7.5.11)$$

Формулы (7.5.9) и (7.5.10) в случае рациональных функций $\bar{s}_x(\lambda)$ и $\Omega(v)$ удобны тем, что интегралы в них могут быть вычислены по формуле (7.4.6).

Формулы (7.5.4)–(7.5.7), (7.5.9) и (7.5.10), как и формулы предыдущего параграфа, применимы только к устойчивым дискретным стационарным линейным системам для моментов времени, достаточно удаленных от момента начала работы системы.

Пример 7.5.1. Найти установившуюся дисперсию выходной переменной стационарной дискретной линейной системы, состоящей из импульсного устройства, вырабатывающего прямоугольные импульсы с периодом повторения T_n , и аperiodического звена, если входное возмущение представляет собой стационарную случайную функцию, интервал корреляции которой меньше периода повторения импульсов T_n . В этом случае значения входного возмущения, действующие на систему, не коррелированы (т. е. входное случайное возмущение представляет собой импульсный белый шум с периодом повторения T_n). В примере 5.4.1 была найдена передаточная функция рассматриваемой системы:

$$\Psi(z) = kz_1 \frac{z_1^{-\gamma} - 1}{z - z_1}, \quad z_1 = e^{-T_n/T}, \quad \gamma = \frac{T_n}{T}. \quad (7.5.12)$$

Спектральная плотность стационарной последовательности некоррелированных импульсов $s_x^d(\omega)$ постоянна и равна $T_n D/2\pi$, где D — дисперсия каждого из этих импульсов. Следовательно, как показывает формула (7.5.8), функция $\sigma_x^d(z)$ в данном случае постоянна и равна $D/2\pi$. Подставляя это выражение спектральной плотности $\sigma_x^d(z)$ и выражение (7.5.12) передаточной функции системы в формулу (7.5.7), находим дисперсию выходной переменной системы:

$$D_y = \frac{Dk^2 z_1^2}{2\pi i} (z_1^{-\gamma} - 1)^2 \int_{\mathcal{C}_1} \frac{dz}{z |z - z_1|^2}. \quad (7.5.13)$$

Формула (7.5.5) в этом случае дает

$$D_y = \frac{T_n Dk^2 z_1^2}{2\pi} (z_1^{-\gamma} - 1)^2 \int_0^{2\pi/T_n} \frac{d\omega}{|e^{i\omega T_n} - z_1|^2} =$$

$$= \frac{Dk^2 z_1^2}{2\pi} (z_1^{-\gamma} - 1)^2 \int_0^{2\pi} \frac{d\varphi}{1 - 2z_1 \cos \varphi + z_1^2}. \quad (7.5.14)$$

К этому же виду приводится формула (7.5.10) заменой переменных $z = e^{i\omega T_1} = e^{i\psi}$. Выполняя интегрирование в формуле (7.5.14), получим

$$D_y = Dk^2 z_1^2 \frac{(z_1^{-\gamma} - 1)^2}{1 - z_1^2}. \quad (7.5.15)$$

Если воспользоваться формулой (7.5.10), то, принимая во внимание, что в данном случае

$$\Omega(v) = \Psi \left(\frac{1+v}{1-v} \right) = kz_1 (z_1^{-\gamma} - 1) \frac{1-v}{1-z_1+(1+z_1)v}, \quad \tilde{z}_x(\lambda) = \frac{D}{\pi}, \quad (7.5.16)$$

получим

$$D_y = \frac{Dk^2 z_1^2}{\pi} (z_1^{-\gamma} - 1)^2 \int_{-\infty}^{\infty} \frac{d\lambda}{|1-z_1+(1+z_1)i\lambda|^2}. \quad (7.5.17)$$

Применив для вычисления интеграла формулу (7.4.6), снова получим формулу (7.5.15).

§ 7.6. Применение теории линейных систем для нахождения интегральных канонических представлений случайных функций

Из теории случайных функций известно, что для нахождения интегрального канонического представления случайной функции $X(t)$ в данном интервале (t_0, t_1) изменения аргумента t достаточно найти три функции $x(t, \lambda)$, $a(t, \lambda)$, $G(\lambda)$, удовлетворяющие условиям (см. [54], § 6.5 или [53], § 67)

$$\int_{t_0}^{t_1} x(t, \lambda) \overline{a(t, \lambda')} dt = \delta(\lambda - \lambda'), \quad (7.6.1)$$

$$\int_{\lambda_1}^{\lambda_2} x(t, \lambda) \overline{a(t', \lambda)} d\lambda = \delta(t - t'), \quad (7.6.2)$$

$$x(t, \lambda) = \frac{1}{G(\lambda)} \int_{t_0}^{t_1} K_x(t, t') a(t', \lambda) dt'. \quad (7.6.3)$$

Если такие функции найдены, то случайная функция

$$V(\lambda) = \int_{t_0}^{t_1} \overline{a(t, \lambda)} X^0(t) dt \quad (7.6.4)$$

будет представлять собой белый шум, интенсивностью которого является функция $G(\lambda)$, а случайная функция $X(t)$ выразится через белый шум $V(\lambda)$ интегральным каноническим представлением

$$X(t) = m_x(t) + \int_{\lambda_1}^{\lambda_2} V(\lambda) x(t, \lambda) d\lambda \quad (t_0 \leq t \leq t_1). \quad (7.6.5)$$

Корреляционная функция белого шума $V(\lambda)$ определяется формулой

$$K_v(\lambda, \lambda') = G(\lambda) \delta(\lambda - \lambda'). \quad (7.6.6)$$

В общей теории интегральных канонических представлений параметр λ может быть произвольной величиной и в соответствии с этим $V(\lambda)$ является белым шумом не в физическом, а в обобщенном, математическом смысле. Однако в приложениях часто оказывается возможным рассматривать случайную функцию времени как результат прохождения некоторого физического белого шума, представляющего собой случайную функцию времени, через некоторую линейную систему. В этом случае параметр λ представляет собой время, а функцию $x(t, \lambda)$ можно рассматривать как весовую функцию системы, формирующей данную случайную функцию времени из белого шума. Функция $a(t, \lambda)$ в этом случае может рассматриваться как весовая функция обратной системы, преобразующей данную случайную функцию в белый шум. Эти соображения наводят на мысль попытаться применить для нахождения интегральных канонических представлений случайных функций теорию линейных систем. Как будет видно из дальнейшего, это во многих случаях приводит к положительным результатам.

В § 4.2 было показано, что весовые функции двух взаимно обратных физически возможных линейных систем в любом интервале (t_0, t_1) связаны соотношениями (4.2.10) и (4.2.11), которые можно переписать в других обозначениях в виде

$$\int_{t_0}^{t_1} w(t, \lambda) w^-(\lambda', t) dt = \delta(\lambda - \lambda') \quad (t_0 \leq \lambda, \lambda' \leq t_1), \quad (7.6.7)$$

$$\int_{t_0}^{t_1} w(t, \lambda) w^-(\lambda, t') d\lambda = \delta(t - t') \quad (t_0 \leq t, t' \leq t_1). \quad (7.6.8)$$

Здесь мы произвольно расширили пределы интегрирования, имея в виду, что вследствие условия физической возможности подынтегральная функция в первой формуле отлична от нуля только при $\lambda < \lambda'$ в интервале (λ, λ') , а подынтегральная функция во второй формуле отлична от нуля только при $t > t'$ в интервале (t', t) . Соотношения (7.6.7) и (7.6.8) по существу не отличаются от (7.6.1) и (7.6.2). Поэтому, если известны весовые функции w и w^- двух взаимно обратных линейных систем, то, принимая во внимание, что эти весовые функции действительны, можно удовлетворить условиям (7.6.1) и (7.6.2), приняв

$$x(t, \lambda) = w(t, \lambda), \quad a(t, \lambda) = w^-(\lambda, t). \quad (7.6.9)$$

Таким образом, весовые функции двух взаимно обратных линейных систем всегда удовлетворяют двум условиям (7.6.1) и (7.6.2)

в любом интервале (t_0, t_1) , если у одной из них первый аргумент рассматривать как независимую переменную, а второй — как параметр, а у другой наоборот. Следовательно, для нахождения интегрального канонического представления случайной функции времени $X(t)$ достаточно среди всех известных пар весовых функций взаимно обратных физически возможных линейных систем найти такую пару, для которой удовлетворяется и условие (7.6.3) в некотором интервале (t_0, t_1) . Критерием выполнения условия (7.6.3), как это видно из формул (7.6.3) и (7.6.9), может служить независимость от времени t выражения

$$\frac{1}{w(t, \lambda)} \int_{t_0}^{t_1} K_x(t, t') w^-(\lambda, t') dt'.$$

Если это выражение при некоторых значениях t_0 и t_1 не зависит от времени t , то, полагая при этих t_0 и t_1

$$G(\lambda) = \frac{1}{w(t, \lambda)} \int_{t_0}^{t_1} K_x(t, t') w^-(\lambda, t') dt', \quad (7.6.10)$$

мы определим функцию $G(\lambda)$, удовлетворяющую совместно с функциями (7.6.9) условию (7.6.3).

На основании формул (7.6.4) и (7.6.9) случайная функция

$$V(\lambda) = \int_{t_0}^{t_1} w^-(\lambda, t) X^0(t) dt \quad (7.6.11)$$

будет при этом белым шумом параметра λ , интенсивность которого равна $G(\lambda)$. Случайная функция $X(t)$ выразится через этот белый шум интегральным каноническим представлением

$$X(t) = m_x(t) + \int_{t_0}^{t_1} w(t, \lambda) V(\lambda) d\lambda \quad (t_0 \leq t \leq t_1). \quad (7.6.12)$$

Корреляционная функция случайной функции $X(t)$ выразится соответствующим интегральным каноническим представлением *)

$$K_x(t, t') = \int_{t_0}^{t_1} G(\lambda) w(t, \lambda) w(t', \lambda) d\lambda \quad (t_0 \leq t, t' \leq t_1). \quad (7.6.13)$$

*) Для вывода формулы (7.6.13) достаточно воспользоваться формулой (7.2.6), расширив в ней пределы интегрирования и заменив верхние пределы обоих интегралов величиной t_1 , предполагая, что $t, t' < t_1$. Тогда, принимая во внимание (7.6.6), получим

$$K_x(t, t') = \int_{t_0}^{t_1} \int_{t_0}^{t_1} w(t, \lambda) w(t', \lambda') G(\lambda) \delta(\lambda - \lambda') d\lambda d\lambda'.$$

Выполняя здесь интегрирование по λ' , мы и получим формулу (7.6.13).

Параметр λ в данном случае имеет смысл времени. Поэтому белый шум $V(\lambda)$ в данном случае является случайной функцией времени, т. е. представляет собой белый шум в обычном, а не в обобщенном смысле. Это дает возможность перейти в формулах (7.6.11), (7.6.12) и (7.6.13) к обычным обозначениям теории линейных систем. Тогда, учитывая, что весовые функции w и w^- удовлетворяют условию физической возможности, перепишем формулы (7.6.11), (7.6.12) и (7.6.13) в виде

$$V(t) = \int_{t_0}^t w^-(t, \tau) X^0(\tau) d\tau, \quad (7.6.14)$$

$$X(t) = m_x(t) + \int_{t_0}^t w(t, \tau) V(\tau) d\tau, \quad (7.6.15)$$

$$K_x(t, t') = \int_{t_0}^{\min[t, t']} G(\tau) w(t, \tau) w(t', \tau) d\tau, \quad (7.6.16)$$

где через $\min[t, t']$ обозначено наименьшее из чисел t и t' .

Таким образом, мы видим, что проблема отыскания интегрального канонического представления случайной функции времени сводится к нахождению двух взаимно обратных физически возможных линейных систем, одна из которых преобразует данную случайную функцию в белый шум, а другая формирует эту случайную функцию из белого шума. В соответствии с этим иногда линейную систему с весовой функцией $w(t, \tau)$, формирующую данную случайную функцию из белого шума, называют *формирующим фильтром*. При этом задачу нахождения интегрального канонического представления случайной функции времени ставят более конкретно как задачу нахождения формирующего фильтра.

Пример 7.6.1. В примере 4.4.1 мы видели, что функции

$$w(t, \tau) = \begin{cases} \frac{1}{a_1(\tau)} \exp \left\{ - \int_{\tau}^t \frac{a_0(\sigma)}{a_1(\sigma)} d\sigma \right\} & \text{при } t > \tau, \\ 0 & \text{при } t < \tau \end{cases}$$

и

$$w^-(t, \tau) = a_1(t) \delta'(t - \tau) + a_0(t) \delta(t - \tau)$$

являются весовыми функциями двух взаимно обратных линейных систем. Предположим теперь, что функции $a_1(t)$ и $a_0(t)$ при всех t положительны и ограничены, и рассмотрим случайную функцию $X(t)$, корреляционная функция которой определяется формулой

$$K_x(t, t') = \begin{cases} q_1(t) q_2(t') & \text{при } t > t', \\ q_1(t') q_2(t) & \text{при } t < t', \end{cases} \quad (7.6.17)$$

где

$$\left. \begin{aligned} q_1(t) &= \exp \left\{ - \int_0^t \frac{a_0(\sigma)}{a_1(\sigma)} d\sigma \right\}, \\ q_2(t) &= q_1(t) \int_{-\infty}^t \frac{d\tau}{a_1^2(\tau) q_1^2(\tau)}. \end{aligned} \right\} \quad (7.6.18)$$

В данном случае при любых $\lambda < t_1$ и $t < t_1$

$$\begin{aligned} \int_{-\infty}^{t_1} K_x(t, t') w^-(\lambda, t') dt' &= \int_{-\infty}^{\lambda} K_x(t, t') w^-(\lambda, t') dt' = \\ &= q_2(t) \int_t^{\lambda} q_1(t') [a_1(\lambda) \delta'(\lambda - t') + a_0(\lambda) \delta(\lambda - t')] dt' = \\ &= q_2(t) [a_1(\lambda) q_1'(\lambda) + a_0(\lambda) q_1(\lambda)] \quad \text{при } t < \lambda, \\ \int_{-\infty}^{t_1} K_x(t, t') w^-(\lambda, t') dt' &= \int_{-\infty}^{\lambda} K_x(t, t') w^-(\lambda, t') dt' = \\ &= q_1(t) \int_{-\infty}^{\lambda} q_2(t') [a_1(\lambda) \delta'(\lambda - t') + a_0(\lambda) \delta(\lambda - t')] dt' = \\ &= q_1(t) [a_1(\lambda) q_2'(\lambda) + a_0(\lambda) q_2(\lambda)] \quad \text{при } t > \lambda. \end{aligned}$$

Но из формул (7.6.18) следует, что

$$\left. \begin{aligned} \frac{q_1'(\lambda)}{q_1(\lambda)} &= \frac{d \ln q_1(\lambda)}{d\lambda} = - \frac{d}{d\lambda} \int_0^{\lambda} \frac{a_0(\sigma)}{a_1(\sigma)} d\sigma = - \frac{a_0(\lambda)}{a_1(\lambda)}, \\ \frac{q_2'(\lambda)}{q_2(\lambda)} &= \frac{d \ln q_2(\lambda)}{d\lambda} = \frac{d \ln q_1(\lambda)}{d\lambda} + \frac{d}{d\lambda} \ln \int_{-\infty}^{\lambda} \frac{d\tau}{a_1^2(\tau) q_1^2(\tau)} = \\ &= - \frac{a_0(\lambda)}{a_1(\lambda)} + \frac{1}{a_1^2(\lambda) q_1^2(\lambda) \int_{-\infty}^{\lambda} \frac{d\tau}{a_1^2(\tau) q_1^2(\tau)}}, \\ &= - \frac{a_0(\lambda)}{a_1(\lambda)} + \frac{1}{a_1^2(\lambda) q_1(\lambda) q_2(\lambda)}. \end{aligned} \right\} \quad (7.6.19)$$

откуда

$$a_1(\lambda) q_1'(\lambda) + a_0(\lambda) q_1(\lambda) = 0, \quad a_1(\lambda) q_2'(\lambda) + a_0(\lambda) q_2(\lambda) = \frac{1}{a_1(\lambda) q_1(\lambda)}.$$

Следовательно,

$$\begin{aligned} \int_{-\infty}^{t_1} K_x(t, t') w^-(\lambda, t') dt' &= 0 = w(t, \lambda) \quad \text{при } t < \lambda, \\ \int_{-\infty}^{t_1} K_x(t, t') w^-(\lambda, t') dt' &= \frac{q_1(t)}{a_1(\lambda) q_1(\lambda)} = w(t, \lambda) \quad \text{при } t > \lambda. \end{aligned}$$

Таким образом, функции $x(t, \lambda) = w(t, \lambda)$, $a(t, \lambda) = w^-(\lambda, t)$ и $G(\lambda) = 1$ удовлетворяют всем трем условиям (7.6.1), (7.6.2) и (7.6.3), если принять $t_0 = \lambda_1 = -\infty$, $\lambda_2 = t_1$, где t_1 — произвольный момент времени. Следовательно, рассматриваемая случайная функция $X(t)$ выражается интегральным каноническим представлением (7.6.15) при $t_0 = -\infty$ в любом полубесконечном интервале времени $(-\infty, t_1)$.

К сожалению, в настоящее время не существует общих методов нахождения такой линейной системы, которая преобразует данную случайную функцию в белый шум, и соответствующей обратной системы — формирующего фильтра. Однако в некоторых частных случаях эта задача решается сравнительно просто. В частности, эта задача легко решается для любой стационарной случайной функции $X(t)$, спектральная плотность которой является дробно-рациональной функцией.

Предположим, что спектральная плотность стационарной случайной функции $X(t)$ выражается формулой

$$s_x(\omega) = \frac{P(\omega)}{Q(\omega)}, \quad (7.6.20)$$

где $P(\omega)$ и $Q(\omega)$ — некоторые полиномы относительно ω . Так как $s_x(\omega)$ — четная функция, то полиномы $P(\omega)$ и $Q(\omega)$ содержат только четные степени частоты ω . При этом в задачах практики степень $2m$ числителя $P(\omega)$ всегда бывает меньше степени $2n$ знаменателя $Q(\omega)$, так как только в этом случае интеграл от спектральной плотности, определяющий дисперсию случайной функции X , может сходиться и дисперсия случайной функции X может быть конечной. Кроме того, для сходимости интеграла от спектральной плотности необходимо, чтобы знаменатель $Q(\omega)$ в выражении спектральной плотности (7.6.20) не обращался в нуль ни при каком действительном значении частоты ω . Наконец, вследствие того, что спектральная плотность существенно положительна при действительных значениях частоты ω , полином $P(\omega)$ обычно также не имеет действительных корней и все коэффициенты полиномов $P(\omega)$ и $Q(\omega)$ действительны. Из этих свойств полиномов $P(\omega)$ и $Q(\omega)$ следует, что корни каждого из них являются попарно сопряженными комплексными величинами и каждому корню соответствует корень того же полинома, противоположный по знаку. Иными словами, корни каждого из полиномов $P(\omega)$ и $Q(\omega)$ расположены в плоскости комплексной переменной ω симметрично относительно действительной и мнимой осей (рис. 7.6.1).

Разложим теперь полиномы $P(\omega)$ и $Q(\omega)$ на множители и отберем в полученных разложениях множители, соответствующие корням, расположенным в верхней полуплоскости комплексной переменной ω . Добавив к таким множителям в разложении полинома $P(\omega)$ корень квадратный из коэффициента A_{2m} при ω^{2m} в этом полиноме и множитель i^m , мы получим полином степени m

относительно ω , все корни которого расположены в верхней полуплоскости. Обозначим этот полином $H(i\omega)$. Легко видеть, что все коэффициенты полинома $H(i\omega)$, рассматриваемого как полином относительно $i\omega$, положительны. Действительно, возьмем какой-нибудь один корень полинома $P(\omega)$, расположенный в верхней полуплоскости (таких корней полином $P(\omega)$ имеет m). Пусть этот корень будет $\alpha + i\beta$, где $\beta > 0$. Если $\alpha \neq 0$, то полином $P(\omega)$

по доказанному имеет корень $-\alpha + i\beta$, также расположенный в верхней полуплоскости. Следовательно, полином $H(i\omega)$ представляет собой произведение положительного числа на множители вида

$$i(\omega - \alpha - i\beta) i(\omega + \alpha - i\beta)$$

и

$$i(\omega - i\beta),$$

где $\beta > 0$. Но

$$i(\omega - \alpha - i\beta) i(\omega + \alpha - i\beta) =$$

$$= (i\omega)^2 + 2\beta(i\omega) + \alpha^2 + \beta^2,$$

$$i(\omega - i\beta) = i\omega + \beta,$$

вследствие чего произведение любого числа таких множителей представляет собой полином относительно $i\omega$, все коэффициенты которого положительны.

Оставшиеся множители в разложении полинома $P(\omega)$, соответствующие корням, расположенным в нижней полуплоскости, умноженные на $(-i)^m$ и на $\sqrt{A_{2m}}$, образуют полином, который получается из $H(i\omega)$ изменением знака у ω , т. е. полином $H(-i\omega)$. Действительно, каждому корню $\alpha + i\beta$ полинома $P(\omega)$, расположенному в верхней полуплоскости, соответствует корень $\alpha - i\beta$, расположенный в нижней полуплоскости. Поэтому каждому множителю вида

$$i(\omega - \alpha - i\beta) i(\omega + \alpha - i\beta) = (i\omega)^2 + 2\beta(i\omega) + \alpha^2 + \beta^2$$

или

$$i(\omega - i\beta) = i\omega + \beta$$

полинома $H(i\omega)$ соответствует оставшийся множитель вида

$$-i(\omega - \alpha + i\beta) (-i)(\omega + \alpha + i\beta) =$$

$$= (-i\omega)^2 + 2\beta(-i\omega) + \alpha^2 + \beta^2$$

или соответственно

$$-i(\omega + i\beta) = -i\omega + \beta,$$

что доказывает высказанное утверждение.

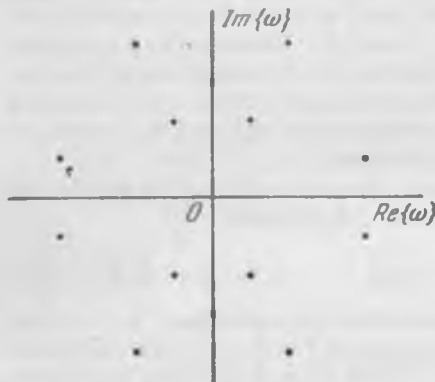


Рис. 7.6.1.

Таким образом, отобрав в разложении полинома $P(\omega)$ множители, соответствующие корням, расположенным в верхней полуплоскости, и добавив множитель $i^m \sqrt{A_{2m}}$, мы можем представить полином $P(\omega)$ в виде

$$P(\omega) = H(i\omega) H(-i\omega), \quad (7.6.21)$$

где $H(i\omega)$ — полином относительно $i\omega$ с положительными коэффициентами.

Совершенно так же, отобрав в разложении полинома $Q(\omega)$ множители, соответствующие корням, расположенным в верхней полуплоскости, добавив к ним множитель $i^n \sqrt{B_{2n}}$, где B_{2n} — коэффициент при старшей степени ω в полиноме $Q(\omega)$, и обозначив образованный этими множителями полином через $F(i\omega)$, мы представим полином $Q(\omega)$ в виде

$$Q(\omega) = F(i\omega) F(-i\omega), \quad (7.6.22)$$

где $F(i\omega)$ — полином относительно $i\omega$ с положительными коэффициентами.

Пользуясь формулами (7.6.21) и (7.6.22), мы можем представить выражение (7.6.20) спектральной плоскости $s_x(\omega)$ для действительных значений ω в виде

$$s_x(\omega) = \left| \frac{H(i\omega)}{F(i\omega)} \right|^2. \quad (7.6.23)$$

Заметим теперь, что функцию

$$\Phi(s) = \frac{H(s)}{F(s)} \quad (7.6.24)$$

можно рассматривать как передаточную функцию некоторой стационарной линейной системы. Тогда функция

$$\frac{1}{\Phi(s)} = \frac{F(s)}{H(s)} \quad (7.6.25)$$

будет представлять собой передаточную функцию обратной системы. Так как умножение комплексного числа на мнимую единицу i представляет собой поворот изображающего это число вектора на угол $\pi/2$ против часовой стрелки и все корни полиномов $H(i\omega)$ и $F(i\omega)$, рассматриваемых как функции ω , лежат в верхней полуплоскости переменной ω , то все корни полиномов $H(s)$ и $F(s)$, рассматриваемых как функции переменной $s = i\omega$, лежат в левой полуплоскости переменной s (рис. 7.6.2). Следовательно, $\Phi(s)$ и $1/\Phi(s)$ являются передаточными функциями двух устойчивых взаимно обратных стационарных линейных систем.

Сравним теперь формулу (7.6.23) с (7.4.5). Вспоминая, что стационарная случайная функция с постоянной спектральной плотностью s_0 представляет собой белый шум с интенсивностью $2\pi s_0$

(см. примечание на стр. 287), приходим к заключению, что стационарную случайную функцию с дробно-рациональной спектральной плотностью можно рассматривать как результат прохождения белого шума с единичной спектральной плотностью через устойчи-

вую стационарную линейную систему с передаточной функцией $\Phi(s)$, определяемой формулой (7.6.24). Эта система и представляет собой формирующий фильтр, соответствующий данной стационарной случайной функции $X(t)$.

Зная передаточные функции двух взаимно обратных стационарных линейных

систем, одна из которых преобразует стационарную случайную функцию $X(t)$ в белый шум, а другая формирует ее из этого белого шума, можно по формуле (2.4.23) найти соответствующие весовые функции:

$$w(t, \tau) = w(t - \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(i\omega) e^{i\omega(t-\tau)} d\omega, \quad (7.6.26)$$

$$w^-(t, \tau) = w^-(t - \tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega(t-\tau)}}{\Phi(i\omega)} d\omega. \quad (7.6.27)$$

Как мы видели в § 7.4, формула (7.4.5) применима только к устойчивым стационарным линейным системам, работающим под действием стационарных случайных возмущений бесконечно долго. Из этого условия следует, во-первых, что предыдущие заключения справедливы только в том случае, когда все нули и полюсы передаточной функции $\Phi(s)$ лежат в левой полуплоскости, что обеспечено изложенным способом построения этой функции. Кстати, это условие обеспечивает существование единственной такой передаточной функции (конечно, с точностью до произвольного постоянного множителя). Во-вторых, из условия применимости формулы (7.4.5) следует, что случайную функцию $X(t)$ можно рассматривать только как результат бесконечно долгого действия стационарного белого шума $V(t)$ с единичной спектральной плотностью на стационарную линейную систему с передаточной функцией $\Phi(s)$, а белый шум $V(t)$ представляет собой результат бесконечно долгого прохождения стационарной случайной функции $X(t)$ через стационарную линейную систему с передаточной функцией $1/\Phi(s)$. Поэтому для стационарной случайной функции $X(t)$ в формулах (7.6.14), (7.6.15) и (7.6.16)

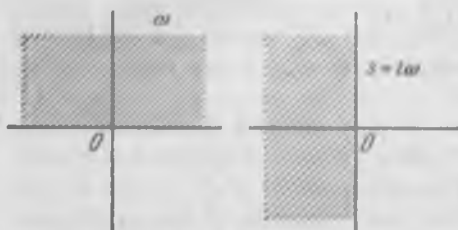


Рис. 7.6.2.

необходимо взять $t_0 = -\infty$. Только при этом условии они будут справедливы. Таким образом, изложенный метод дает интегральное каноническое представление любой стационарной случайной функции $X(t)$, имеющей дробно-рациональную спектральную плотность, в любом полубесконечном интервале $(-\infty, t_1)$.

Заметим еще, что поведение системы с рациональной передаточной функцией $\Phi(s)$, определяемой формулой (7.6.24), на основании § 2.5 описывается линейным дифференциальным уравнением с постоянными положительными коэффициентами

$$F\left(\frac{d}{dt}\right) X^0 = H\left(\frac{d}{dt}\right) V. \quad (7.6.28)$$

Из того факта, что все коэффициенты уравнения (7.6.28) действительны, следует на основании сказанного в § 4.4, что весовые функции w и w^* этой системы и обратной системы действительными. Но в таком случае интегралы в формулах (7.6.26) и (7.6.27) действительны и, следовательно, совпадают со своими комплексными сопряженными величинами. Поэтому в формулах (7.6.26) и (7.6.27) можно изменить знак у мнимой единицы i . Тогда, изменяя обозначение переменной интегрирования, получим

$$w(t-\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \Phi(-i\mu) e^{-i\mu(t-\tau)} d\mu, \quad (7.6.29)$$

$$w^*(t-\tau) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-i\mu(t-\tau)}}{\Phi(-i\mu)} d\mu. \quad (7.6.30)$$

Практически для вычисления интегралов в этих формулах можно пользоваться таблицами преобразования Лапласа (см., например, [17]).

В некоторых случаях бывает удобно выразить случайную функцию $X(t)$ через белый шум, интенсивность которого тождественно равна единице. Согласно примечанию на стр. 287 такой белый шум является стационарной случайной функцией с постоянной спектральной плотностью, равной $1/2\pi$. Следовательно, для того чтобы выразить стационарную случайную функцию $X(t)$ с дробно-рациональной спектральной плотностью через белый шум с единичной интенсивностью, достаточно в предыдущих рассуждениях включить в полином $H(i\omega)$ дополнительный множитель $\sqrt{2\pi}$ или в полином $F(i\omega)$ дополнительный множитель $1/\sqrt{2\pi}$. Тогда получим вместо формулы (7.6.23)

$$s_x(\omega) = \frac{1}{2\pi} \left| \frac{H(i\omega)}{F(i\omega)} \right|^2. \quad (7.6.31)$$

Вследствие того, что корреляционные функции и спектральные плотности случайных функций обычно определяются по результатам

экспериментов, они всегда бывают известны лишь приближенно. Поэтому практически любую спектральную плотность можно аппроксимировать дробно-рациональной функцией. Отсюда следует, что изложенный метод нахождения интегрального канонического представления применим практически к любым стационарным случайным функциям.

Пример 7.6.2. Найти интегральное каноническое представление стационарной случайной функции $X(t)$, спектральная плотность и корреляционная функция которой определяются формулами

$$s_x(\omega) = \frac{D}{\pi} \frac{\alpha}{\alpha^2 + \omega^2}, \quad k_x(\tau) = D e^{-\alpha |\tau|}. \quad (7.6.32)$$

В данном случае числитель $P(\omega)$ является постоянной величиной ($m=0$), а знаменатель имеет два чисто мнимых корня $\omega = \pm i\alpha$. Следовательно, в данном случае можно принять

$$H(i\omega) = \sqrt{\frac{D\alpha}{\pi}}, \quad F(i\omega) = i(\omega - i\alpha) = i\omega + \alpha.$$

Тогда получим

$$s_x(\omega) = | \Phi(i\omega) |^2, \quad \Phi(s) = \sqrt{\frac{D\alpha}{\pi}} \frac{1}{s + \alpha}. \quad (7.6.33)$$

Таким образом, формирующим фильтром в данном случае является апериодическое звено с постоянной времени $T = 1/\alpha$ и коэффициентом усиления $k = \sqrt{D/\pi\alpha}$ [см. формулу (2.6.4)]. При этом рассматриваемая случайная функция X будет представлять собой результат бесконечно долгого действия на это звено белого шума с интенсивностью, равной 2π . Полагая

$$H(i\omega) = \sqrt{2D\alpha}, \quad F(i\omega) = i\omega + \alpha, \quad (7.6.34)$$

получим

$$s_x(\omega) = \frac{1}{2\pi} | \Phi(i\omega) |^2, \quad \Phi(s) = \frac{\sqrt{2D\alpha}}{s + \alpha}. \quad (7.6.35)$$

Таким образом, случайную функцию X в данном случае можно рассматривать как результат бесконечно долгого действия белого шума единичной интенсивности на апериодическое звено с постоянной времени $T = 1/\alpha$ и коэффициентом усиления $k = \sqrt{2D/\alpha}$.

Определив по формулам (7.6.26) и (7.6.27) или (7.6.29) и (7.6.30) или по таблицам преобразований Лапласа весовые функции w и w^- , выразим рассматриваемую стационарную случайную функцию $X(t)$ интегральным каноническим представлением (7.6.15) при $t_0 = -\infty$ в интервале $(-\infty, t_1)$, где t_1 — любой момент времени. Впрочем, как мы увидим дальше, для решения практических задач в данном случае нет необходимости определять весовые функции w и w^- , а вполне достаточно знать передаточную функцию $\Phi(s)$.

Пример 7.6.3. Найти интегральное каноническое представление стационарной случайной функции $X(t)$, спектральная плотность которой

определяется формулой

$$s_x(\omega) = \frac{Da}{\pi} \frac{b^2 + \omega^2}{b^4 + 2(a^2 - \omega_0^2)\omega^2 + \omega^4}, \quad \omega_0^2 = b^2 - a^2,$$

где $b > a > 0$ *).

В данном случае числитель $P(\omega)$ имеет два корня $\omega = \pm ib$, а знаменатель $Q(\omega)$ — четыре корня $\pm(\omega_0 \pm ia)$. Отобрав из этих корней те, которые лежат в верхней полуплоскости, согласно изложенному методу, полагаем

$$\left. \begin{aligned} H(i\omega) &= \sqrt{2Da} i(\omega - ib) = \sqrt{2Da} (i\omega + b), \\ F(i\omega) &= i(\omega - \omega_0 - ia) i(\omega + \omega_0 - ia) = (i\omega)^2 + 2a(i\omega) + b^2. \end{aligned} \right\} \quad (7.6.36)$$

Тогда будем иметь

$$s_x(\omega) = \frac{1}{2\pi} \left| \frac{H(i\omega)}{F(i\omega)} \right|^2 = \frac{1}{2\pi} \left| \frac{\sqrt{2Da} (i\omega + b)}{(i\omega)^2 + 2a(i\omega) + b^2} \right|^2.$$

Следовательно, стационарную случайную функцию X в данном случае можно рассматривать как результат бесконечно долгого действия белого шума с единичной интенсивностью на устойчивую стационарную линейную систему (формирующий фильтр) с передаточной функцией

$$\Phi(s) = \frac{H(s)}{F(s)} = \frac{\sqrt{2Da}(s+b)}{s^2 + 2as + b^2}.$$

Весовые функции этой системы и обратной системы можно в случае необходимости вычислить по формулам (7.6.26) и (7.6.27) или (7.6.29) и (7.6.30) или по таблицам преобразования Лапласа.

Как уже было отмечено, изложенный способ нахождения интегрального канонического представления применим практически к любым стационарным случайным функциям. Представляет интерес обобщить его на нестационарные случайные функции времени. Мы видели, что поведение системы, связывающей стационарную случайную функцию, имеющую дробно-рациональную спектральную плотность, с белым шумом, описывается линейным дифференциальным уравнением с постоянными коэффициентами (7.6.28). Поэтому естественным обобщением класса стационарных случайных функций с дробно-рациональными спектральными плотностями является класс нестационарных случайных функций времени, связанных с белым шумом линейными дифференциальными уравнениями с переменными коэффициентами вида

$$FX^0 = HV, \quad (7.6.37)$$

где F и H — линейные дифференциальные операторы вида (4.4.4) с произвольными достаточно гладкими переменными коэффициентами. Однако в настоящее время мы еще не умеем находить линейное дифференциальное уравнение, определяющее приближенную зависимость между любой данной случайной функцией

*) Предоставляем читателю самостоятельно проверить, что корреляционная функция этой случайной функции выражается формулой

$$k_x(\tau) = De^{-a|\tau|} \cos \omega_0 \tau.$$

и белым шумом. Для стационарной случайной функции, как мы видели, для нахождения такого дифференциального уравнения достаточно аппроксимировать ее спектральную плотность дробно-рациональной функцией и определить изложенным методом полиномы $H(s)$ и $F(s)$. Поэтому применение изложенного метода к нестационарным случайным функциям времени пока еще ограничивается случаями, когда удастся найти приближенную зависимость между интересующей нас случайной функцией и белым шумом в форме линейного дифференциального уравнения. В таких случаях весовые функции w и w^* могут быть определены методами, изложенными в §§ 4.3 и 4.4.

Пример 7.6.4. Весовые функции w и w^* примера 7.6.1 соответствуют связи между случайной функцией X и белым шумом V в форме линейного дифференциального уравнения первого порядка (см. пример 4.4.1):

$$a_1(t) \dot{X} + a_0(t) X = V. \quad (7.6.38)$$

Следовательно, случайная функция X , имеющая корреляционную функцию вида (7.6.17), всегда связана с некоторым белым шумом уравнением первого порядка вида (7.6.38). Для определения коэффициентов $a_0(t)$ и $a_1(t)$ этого уравнения по данным функциям $q_1(t)$, $q_2(t)$ достаточно воспользоваться уравнениями (7.6.19), из которых следует формула

$$\frac{d}{dt} \ln \frac{q_2(t)}{q_1(t)} = \frac{1}{a_1^2(t) q_1(t) q_2(t)}. \quad (7.6.39)$$

Определив отсюда функцию $a_1(t)$, мы найдем после этого функцию $a_0(t)$ из первого уравнения (7.6.19). Заметим, что эти операции имеют практический смысл только в том случае, когда отношение $q_2(t)/q_1(t)$ является возрастающей функцией, так как только в этом случае уравнение (7.6.39) определяет действительную функцию $a_1(t)$ *).

§ 7.7. Определение моментов выходных переменных интегрированием дифференциальных уравнений

Рассмотрим автоматическую систему, описываемую системой обыкновенных линейных дифференциальных уравнений конечного порядка, на которую действуют случайные возмущения, представляющие собой белые шумы. В § 4.5 было показано, что уравнения такой системы всегда можно привести к системе дифференциальных уравнений первого порядка

$$\dot{Y}_k = \sum_{l=1}^n a_{kl} Y_l + V_k \quad (k=1, \dots, n), \quad (7.7.1)$$

где a_{kl} — переменные или постоянные коэффициенты, а V_k — белые шумы.

*) Можно показать, что это условие необходимо также для того, чтобы функция, определяемая формулой (7.6.17), могла быть корреляционной функцией некоторой случайной функции.

В § 7.2 показано, что математические ожидания выходных переменных рассматриваемой системы можно получить, если в системе дифференциальных уравнений (7.7.1) входные сигналы и начальные условия заменить их математическими ожиданиями:

$$\dot{m}_k = \sum_{l=1}^n a_{kl} m_l + m_{v_k} \quad (k = 1, \dots, n). \quad (7.7.1)$$

Интегрирование системы (7.7.2) не представляет никаких практических трудностей. Поэтому мы рассмотрим только центрированные случайные составляющие выходных сигналов, определяемые системой дифференциальных уравнений

$$\dot{Y}_k^0 = \sum_{l=1}^n a_{kl} Y_l^0 + V_k^0 \quad (k = 1, \dots, n), \quad (7.7.2)$$

которая получается вычитанием из уравнений (7.7.1) соответствующих уравнений (7.7.2).

Продифференцируем попарные произведения всех выходных переменных системы:

$$\frac{d}{dt} (Y_i^0 Y_j^0) = \dot{Y}_i^0 Y_j^0 + Y_j^0 \dot{Y}_i^0 \quad (i, j = 1, \dots, n). \quad (7.7.3)$$

Подставляя в правую часть этого равенства выражения производных выходных переменных системы из (7.7.3), получим

$$\frac{d}{dt} (Y_i^0 Y_j^0) = \sum_{l=1}^n a_{il} Y_j^0 Y_l^0 + \sum_{l=1}^n a_{jl} Y_i^0 Y_l^0 + Y_j^0 V_i^0 + Y_i^0 V_j^0 \quad (7.7.4)$$

$$(i, j = 1, \dots, n).$$

Применяя к обеим частям равенства операцию математического ожидания и пользуясь возможностью изменять порядок операции дифференцирования и математического ожидания, получим

$$\frac{d}{dt} M [Y_i^0 Y_j^0] = \sum_{l=1}^n a_{il} M [Y_j^0 Y_l^0] + \sum_{l=1}^n a_{jl} M [Y_i^0 Y_l^0] +$$

$$+ M [Y_j^0 V_i^0] + M [Y_i^0 V_j^0] \quad (i, j = 1, \dots, n). \quad (7.7.5)$$

Входящие в эти уравнения математические ожидания $M [Y_i^0 V_j^0]$ представляют собой дисперсии (при $i = j$) и корреляционные моменты (при $i \neq j$) выходных сигналов системы. Вводя обозначения

$$\theta_{ij} = M [Y_i^0 Y_j^0] \quad (i, j = 1, \dots, n), \quad (7.7.6)$$

перепишем уравнения (7.7.5) в виде

$$\dot{\theta}_{ij} = \sum_{l=1}^n (a_{il} \theta_{jl} + a_{jl} \theta_{il}) + M [Y_j^0 V_i^0] + M [Y_i^0 V_j^0] \quad (7.7.7)$$

$$(i, j = 1, \dots, n).$$

и белым шумом. Для стационарной случайной функции, как мы видели, для нахождения такого дифференциального уравнения достаточно аппроксимировать ее спектральную плотность дробно-рациональной функцией и определить изложенным методом полиномы $H(s)$ и $F(s)$. Поэтому применение изложенного метода к нестационарным случайным функциям времени пока еще ограничивается случаями, когда удастся найти приближенную зависимость между интересующей нас случайной функцией и белым шумом в форме линейного дифференциального уравнения. В таких случаях весовые функции w и w^- могут быть определены методами, изложенными в §§ 4.3 и 4.4.

Пример 7.6.4. Весовые функции w и w^- примера 7.6.1 соответствуют связи между случайной функцией X и белым шумом V в форме линейного дифференциального уравнения первого порядка (см. пример 4.4.1):

$$a_1(t) \dot{X} + a_0(t) X = V. \quad (7.6.38)$$

Следовательно, случайная функция X , имеющая корреляционную функцию вида (7.6.17), всегда связана с некоторым белым шумом уравнением первого порядка вида (7.6.38). Для определения коэффициентов $a_0(t)$ и $a_1(t)$ этого уравнения по данным функциям $q_1(t)$, $q_2(t)$ достаточно воспользоваться уравнениями (7.6.19), из которых следует формула

$$\frac{d}{dt} \ln \frac{q_2(t)}{q_1(t)} = \frac{1}{a_1^2(t) q_1(t) q_2(t)}. \quad (7.6.39)$$

Определив отсюда функцию $a_1(t)$, мы найдем после этого функцию $a_0(t)$ из первого уравнения (7.6.19). Заметим, что эти операции имеют практический смысл только в том случае, когда отношение $q_2(t)/q_1(t)$ является возрастающей функцией, так как только в этом случае уравнение (7.6.39) определяет действительную функцию $a_1(t)$ *).

§ 7.7. Определение моментов выходных переменных интегрированием дифференциальных уравнений

Рассмотрим автоматическую систему, описываемую системой обыкновенных линейных дифференциальных уравнений конечного порядка, на которую действуют случайные возмущения, представляющие собой белые шумы. В § 4.5 было показано, что уравнения такой системы всегда можно привести к системе дифференциальных уравнений первого порядка

$$\dot{Y}_k = \sum_{l=1}^n a_{kl} Y_l + V_k \quad (k = 1, \dots, n), \quad (7.7.1)$$

где a_{kl} — переменные или постоянные коэффициенты, а V_k — белые шумы.

*) Можно показать, что это условие необходимо также для того, чтобы функция, определяемая формулой (7.6.17), могла быть корреляционной функцией некоторой случайной функции.

В § 7.2 показано, что математические ожидания выходных переменных рассматриваемой системы можно получить, если в системе дифференциальных уравнений (7.7.1) входные сигналы и начальные условия заменить их математическими ожиданиями:

$$\dot{m}_k = \sum_{l=1}^n a_{kl} m_l + m_{v_k} \quad (k = 1, \dots, n). \quad (7.7.2)$$

Интегрирование системы (7.7.2) не представляет никаких практических трудностей. Поэтому мы рассмотрим только центрированные случайные составляющие выходных сигналов, определяемые системой дифференциальных уравнений

$$\dot{Y}_k^0 = \sum_{l=1}^n a_{kl} Y_l^0 + V_k^0 \quad (k = 1, \dots, n), \quad (7.7.3)$$

которая получается вычитанием из уравнений (7.7.1) соответствующих уравнений (7.7.2).

Продифференцируем попарные произведения всех выходных переменных системы:

$$\frac{d}{dt} (Y_i^0 Y_j^0) = \dot{Y}_i^0 Y_j^0 + \dot{Y}_j^0 Y_i^0 \quad (i, j = 1, \dots, n). \quad (7.7.4)$$

Подставляя в правую часть этого равенства выражения производных выходных переменных системы из (7.7.3), получим

$$\frac{d}{dt} (Y_i^0 Y_j^0) = \sum_{l=1}^n a_{il} Y_l^0 Y_j^0 + \sum_{l=1}^n a_{jl} Y_i^0 Y_l^0 + Y_j^0 V_i^0 + Y_i^0 V_j^0 \quad (7.7.5)$$

$$(i, j = 1, \dots, n).$$

Применяя к обеим частям равенства операцию математического ожидания и пользуясь возможностью изменять порядок операций дифференцирования и математического ожидания, получим

$$\frac{d}{dt} M [Y_i^0 Y_j^0] = \sum_{l=1}^n a_{il} M [Y_l^0 Y_j^0] + \sum_{l=1}^n a_{jl} M [Y_i^0 Y_l^0] +$$

$$+ M [Y_j^0 V_i^0] + M [Y_i^0 V_j^0] \quad (i, j = 1, \dots, n). \quad (7.7.6)$$

Входящие в эти уравнения математические ожидания $M [V_i^0 V_j^0]$ представляют собой дисперсии (при $i = j$) и корреляционные моменты (при $i \neq j$) выходных сигналов системы. Вводя обозначение

$$\theta_{ij} = M [Y_i^0 Y_j^0] \quad (i, j = 1, \dots, n), \quad (7.7.7)$$

перепишем уравнения (7.7.6) в виде

$$\dot{\theta}_{ij} = \sum_{l=1}^n (a_{il} \theta_{lj} + a_{jl} \theta_{il}) + M [Y_j^0 V_i^0] + M [Y_i^0 V_j^0] \quad (7.7.8)$$

$$(i, j = 1, \dots, n).$$

Найдем входящие в (7.7.8) математические ожидания $M[Y_v^0 V_\mu^0]$. На основании формулы (2.2.30) можно написать

$$Y_v^0(t) = \sum_{h=1}^n \int_{t_0}^t g_{vh}(t, \tau) V_h^0(\tau) d\tau. \quad (7.7.9)$$

Отсюда находим

$$M[Y_v^0(t) V_\mu^0(t)] = \sum_{h=1}^n \int_{t_0}^t g_{vh}(t, \tau) M[V_\mu^0(t) V_h^0(\tau)] d\tau. \quad (7.7.10)$$

Так как согласно принятым допущениям $V_1, \dots, V_n$ являются белыми шумами, то

$$M[V_\mu^0(t) V_h^0(\tau)] = G_{\mu h}(t) \delta(t - \tau), \quad (7.7.11)$$

где $G_{\mu h}(t)$ ($\mu, h = 1, \dots, n$) — интенсивности и взаимные интенсивности белых шумов $V_1, \dots, V_n$. Подставляя выражение (7.7.11) в (7.7.10), получаем

$$M[Y_v^0(t) V_\mu^0(t)] = \sum_{h=1}^n G_{\mu h}(t) \int_{t_0}^t g_{vh}(t, \tau) \delta(t - \tau) d\tau. \quad (7.7.12)$$

Заметим теперь, что весовая функция $g_{vh}(t, \tau)$ терпит разрыв в точке $\tau = t$. Поэтому для вычисления интегралов в (7.7.12) следует применить формулу (2.1.8):

$$\int_{t_0}^t g_{vh}(t, \tau) \delta(t - \tau) d\tau = \frac{g_{vh}(t, t-0) + g_{vh}(t, t+0)}{2}. \quad (7.7.13)$$

Для физически возможных систем $g_{vh}(t, t + \varepsilon) = 0$ при любом $\varepsilon > 0$ и, следовательно, $g_{vh}(t, t + 0) = 0$. Что касается предела слева $g_{vh}(t, t - 0)$, то его мы обозначали в § 4.5 просто через $g_{vh}(t, t)$. Поэтому можно написать

$$\int_{t_0}^t g_{vh}(t, \tau) \delta(t - \tau) d\tau = \frac{1}{2} g_{vh}(t, t). \quad (7.7.14)$$

Но, как показано в § 4.5, $g_{vv}(t, t) = 1$, $g_{vh}(t, t) = 0$ при $v \neq h$. Поэтому из (7.7.12) и (7.7.14) получаем

$$M[Y_v^0(t) V_\mu^0(t)] = \frac{1}{2} G_{\mu v}(t) \quad (v, \mu = 1, \dots, n). \quad (7.7.15)$$

Подставляя выражение (7.7.15) в (7.7.8) и учитывая, что $G_{ij}(t) = G_{ji}(t)$, получаем окончательно:

$$\dot{\theta}_{ij} = \sum_{l=1}^n (a_{il}\theta_{jl} + a_{jl}\theta_{il}) + G_{ij}(t) \quad (i, j = 1, \dots, n). \quad (7.7.16)$$

Систему уравнений (7.7.16) можно компактно записать в виде одного матричного уравнения. Для этого выходные переменные $Y_1, \dots, Y_n$ будем считать составляющими вектора Y , а белые шумы $V_1, \dots, V_n$ — составляющими вектора V . Тогда, представив векторы матрицами-столбцами, можем записать систему уравнений (7.7.1) в виде одного матричного уравнения:

$$\dot{Y} = AY + V, \quad (7.7.17)$$

где A — матрица коэффициентов уравнений (7.7.1):

$$A = \begin{bmatrix} a_{11} & a_{12} & \dots & a_{1n} \\ a_{21} & a_{22} & \dots & a_{2n} \\ \dots & \dots & \dots & \dots \\ a_{n1} & a_{n2} & \dots & a_{nn} \end{bmatrix}.$$

Вводя матрицы

$$\Theta = \begin{bmatrix} \theta_{11} & \theta_{12} & \dots & \theta_{1n} \\ \theta_{21} & \theta_{22} & \dots & \theta_{2n} \\ \dots & \dots & \dots & \dots \\ \theta_{n1} & \theta_{n2} & \dots & \theta_{nn} \end{bmatrix}, \quad G = \begin{bmatrix} G_{11} & G_{12} & \dots & G_{1n} \\ G_{21} & G_{22} & \dots & G_{2n} \\ \dots & \dots & \dots & \dots \\ G_{n1} & G_{n2} & \dots & G_{nn} \end{bmatrix},$$

замечаем, что в силу симметричности матрицы Θ первая сумма в (7.7.16) представляет собой элемент i -й строки и j -го столбца произведения матриц $A \Theta$, а вторая сумма — элемент i -й строки и j -го столбца произведения матрицы Θ на транспонированную матрицу A^T . Поэтому систему уравнений (7.7.16) можно записать в виде одного матричного уравнения:

$$\dot{\Theta} = A \Theta + \Theta A^T + G. \quad (7.7.18)$$

Предоставляем читателю самостоятельно убедиться в том, что если математические ожидания белых шумов $V_k(t)$ равны нулю, $m_{v_k} = 0$ ($k = 1, \dots, n$), то уравнения (7.7.16) и (7.7.18) остаются в силе, если заменить в них корреляционную матрицу $\Theta = \|\theta_{ij}\|$ случайного вектора Y матрицей начальных моментов второго порядка его составляющих $\Gamma = \|\gamma_{ij}\|$, $\gamma_{ij} = M[Y_i Y_j]$ ($i, j = 1, \dots, n$).

Таким образом, мы получили систему обыкновенных линейных дифференциальных уравнений для определения дисперсий и корреляционных моментов выходных сигналов системы. Интегрируя эту систему аналитически, если это возможно, или с помощью вычислительных машин, можно получить статистические характеристики сразу всех выходных сигналов как функции времени. Правда, за эту возможность приходится расплачиваться сильным ростом порядка системы дифференциальных уравнений по сравнению с порядком исходной системы уравнений. Действительно, если

поведение исследуемой системы описывается системой дифференциальных уравнений порядка n , то в систему уравнений (7.7.16) будут входить n уравнений для дисперсий и $n(n-1)/2$ уравнений для корреляционных моментов выходных сигналов ($\theta_{ij} = \theta_{ji}$). Следовательно, порядок системы уравнений (7.7.16) будет равен $N = n + n(n-1)/2 = n(n+1)/2$. В таблице 7.7.1 приведена зависимость порядка системы дифференциальных уравнений для дисперсий и корреляционных моментов выходных сигналов от порядка исходной системы.

ТАБЛИЦА 7.7.1

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
N	1	3	6	10	15	21	28	36	45	55	66	78	91	105	120

Следовательно, изложенным методом целесообразно пользоваться, только если порядок исходной системы невысок. В противном случае следует пользоваться одним из общих методов, описанных в § 7.2.

Если исследуемая система и белые шумы на ее входах стационарны, то для определения установившихся значений дисперсий и корреляционных моментов выходных сигналов следует в системе дифференциальных уравнений (7.7.16) положить $\dot{\theta}_{ij} = 0$. Тогда искомые характеристики точности работы системы определяются из системы линейных алгебраических уравнений

$$\sum_{i=1}^n (a_{ii}\theta_{ji} + a_{ji}\theta_{ii}) = -G_{ij} \quad (i, j = 1, \dots, n). \quad (7.7.19)$$

Заметим, что изложенный метод пригоден и в случае, когда входные сигналы не являются белыми шумами, но могут быть представлены в виде преобразования белых шумов некоторой линейной системой — формирующим фильтром, — описываемой системой обыкновенных дифференциальных уравнений конечного порядка. В этом случае к дифференциальным уравнениям, описывающим автоматическую систему, следует добавить дифференциальные уравнения, описывающие формирующий фильтр. При этом надо помнить, что заданный случайный процесс на выходе формирующего фильтра практически всегда устанавливается лишь после затухания переходного процесса в нем. Поэтому дифференциальные уравнения для моментов выходных переменных формирующего фильтра следует начинать интегрировать раньше, чем уравнения для моментов выходных сигналов исследуемой системы и их корреляционных моментов с выходными сигналами формирующего фильтра. И лишь спустя некоторое время, превосходящее время

переходных процессов в формирующем фильтре, можно начинать совместное интегрирование всех уравнений.

Если входной случайный сигнал системы стационарен, то формирующий фильтр и действующий на него белый шум стационарны. В этом случае дифференциальные уравнения для моментов выходных сигналов формирующего фильтра в установившемся режиме превращаются в алгебраические. Определив из этих уравнений моменты выходных сигналов формирующего фильтра, можно подставить их в дифференциальные уравнения для остальных моментов.

Пример 7.7.1. Определить дисперсию на выходе системы, описываемой дифференциальным уравнением первого порядка

$$\dot{Y} = -a(t) Y + X(t), \quad (7.7.20)$$

где $X(t)$ — стационарная случайная функция, математическое ожидание которой равно нулю, а спектральная плотность и корреляционная функция определяются формулами (7.6.32). Начальное значение Y при $t = 0$ является случайной величиной с нулевым математическим ожиданием и известной дисперсией $D_Y(0)$.

В примере 7.6.2 показано, что случайную функцию $X(t)$ можно рассматривать как результат бесконечно долгого действия белого шума единичной интенсивности на апериодическое звено с постоянной времени $T = 1/\alpha$ и коэффициентом усиления $K = \sqrt{2D/\alpha}$. Следовательно, случайная функция $X(t)$ связана с белым шумом V единичной интенсивности дифференциальным уравнением

$$\dot{X} = -\alpha X + \sqrt{2D\alpha} V. \quad (7.7.21)$$

Положим

$$Y = Y_1, \quad X = Y_2, \quad \sqrt{2D\alpha} V = V_2.$$

Тогда уравнения (7.7.20) и (7.7.21) запишутся в виде

$$\left. \begin{aligned} \dot{Y}_1 &= -a(t) Y_1 + Y_2, \\ \dot{Y}_2 &= -\alpha Y_2 + V_2, \end{aligned} \right\} \quad (7.7.22)$$

где V_2 — белый шум интенсивности $2D\alpha$.

Чтобы получить уравнения для моментов второго порядка, положим в (7.7.16) $n = 2$, $a_{11} = -a(t)$, $a_{12} = 1$, $a_{21} = 0$, $a_{22} = -\alpha$, $G_{11} = G_{12} = 0$, $G_{22} = 2D\alpha$:

$$\left. \begin{aligned} \dot{\theta}_{11} &= -2a(t) \theta_{11} + 2\theta_{12}, \\ \dot{\theta}_{12} &= -[a(t) + \alpha] \theta_{12} + \theta_{22}, \\ \dot{\theta}_{22} &= -2\alpha \theta_{22} + 2D\alpha. \end{aligned} \right\} \quad (7.7.23)$$

Так как вследствие стационарности $X(t)$ нас интересует только установившееся значение дисперсии случайной функции на выходе формирующего фильтра, то в последнем уравнении нужно положить $\dot{\theta}_{22} = 0$, после чего получим, как и следовало ожидать, $\theta_{22} = D$. Первые два уравнения (7.7.23)

при этом примут вид

$$\left. \begin{aligned} \dot{\theta}_{11} &= -2a(t) + 2\theta_{12}, \\ \dot{\theta}_{12} &= -[a(t) + \alpha]\theta_{12} + D. \end{aligned} \right\} \quad (7.7.24)$$

Эту систему уравнений следует проинтегрировать при начальных условиях $\theta_{11}(0) = D_H(0)$, $\theta_{12}(0) = 0$. Дисперсией на выходе рассматриваемой системы является величина θ_{11} . Если коэффициент $a(t)$ не зависит от времени, $a(t) = a$, и требуется определить установившееся значение θ_{11} , то, полагая $\dot{\theta}_{11} = \dot{\theta}_{12} = 0$, получаем два алгебраических уравнения

$$\left. \begin{aligned} -2a\theta_{11} + 2\theta_{12} &= 0, \\ -(a + \alpha)\theta_{12} + D &= 0. \end{aligned} \right\} \quad (7.7.25)$$

Решая эти уравнения, получаем установившееся значение дисперсии выходной переменной системы:

$$D_H = \theta_{11} = \frac{D}{a(a + \alpha)}. \quad (7.7.26)$$

Глава 8

ХАРАКТЕРИСТИКИ НЕЛИНЕЙНЫХ СИСТЕМ

§ 8.1. Основные свойства нелинейных систем

Согласно данному в § 1.4 определению нелинейной системой называется динамическая система, для которой не выполняется принцип суперпозиции. Система является нелинейной при наличии в ней хотя бы одного элемента, для которого не выполняется принцип суперпозиции. Из этого определяющего свойства нелинейных систем вытекают другие присущие им характерные черты, связанные с тем, что процессы в нелинейных системах и собственные их движения гораздо разнообразнее и сложнее процессов и движений в линейных системах. Так, в нелинейных системах возможны без внешних возмущений устойчивые колебания определенной амплитуды и частоты, называемые обычно *автоколебаниями*. Такие режимы невозможны в линейных системах. Система может иметь несколько различных видов устойчивых периодических процессов с разными частотами и соответствующими вполне определенными амплитудами. Такие режимы нелинейной системы являются результатом внутренних нелинейных свойств элементов системы. Появление и установление того или иного вида периодического режима зависит от начальных условий. Таким образом, в отличие от линейных систем, в нелинейной системе сам характер собственных движений (колебания, затухающие колебания, расходящиеся колебания, аperiodический процесс) зависит от начальных условий. Одна и та же система при различных начальных условиях может совершать различные по своему характеру движения. Так, например, если маятнику в состоянии покоя сообщить в некоторый начальный момент толчок, т. е. начальную скорость, то при достаточно малой начальной скорости он будет совершать колебания, а при достаточно большой начальной скорости — аperiodическое движение — вращение в одном направлении.

При периодическом внешнем входном возмущении в нелинейной системе могут наблюдаться особые случаи резонанса, несвойственные линейным системам и характеризующиеся неоднозначной зависимостью амплитуды вынужденных колебаний от частоты. Резонанс в нелинейной системе возможен и на дробной частоте, т. е. на частоте, равной некоторой доле частоты входного возмущения. Наконец, некоторым нелинейным системам свойственны

явления параметрического возбуждения и резонанса, заключающиеся в том, что в системе возникают колебания при периодическом изменении какого-либо ее параметра.

Особые свойства нелинейных систем широко используются в технике. На этих свойствах основано генерирование электромагнитных колебаний, выпрямление переменного тока, умножение и деление частот и другие процессы. Существует большое число нелинейных автоматических систем, в которых рационально используются нелинейные характеристики отдельных элементов и на этой основе получаются хорошие практические результаты. По динамическим качествам нелинейные автоматические системы во многих случаях превосходят линейные системы.

Однако в некоторых случаях нелинейности характеристик являются вредными факторами. Их надо либо устранять, либо выбирать режимы работы таким образом, чтобы нелинейности не оказывали существенного влияния на процессы в системе.

Особенности поведения нелинейных систем и многообразие процессов в них создают трудности точного их математического описания и теоретического изучения. Несмотря на это, задачи исследования нелинейных систем, несравненно более трудные, чем задачи исследования линейных систем, постепенно приобретают в современной технике все более актуальное значение. Можно выделить две основные проблемы теории нелинейных систем.

Первая проблема — исследование влияния нелинейностей характеристик некоторых реальных элементов на процессы в системах, которые в основном являются линейными и проектируются как линейные. В этих случаях наличие нелинейных зависимостей изменяет только количественные характеристики процессов в системах. Необходимо учесть их влияние или спроектировать систему так, чтобы процессы в ней не выходили из границ линейности характеристик элементов.

Вторая проблема — исследование принципиально нелинейных систем, имеющих в своем составе нелинейные элементы, которые определяют характерные особенности динамических процессов в таких системах. Эта проблема включает задачи анализа процессов в нелинейных системах и синтеза систем, обладающих заданными динамическими характеристиками.

§ 8.2. Типовые нелинейности автоматических систем

Реальные элементы автоматических систем, строго говоря, в большинстве случаев нелинейны, и лишь в известных пределах можно считать их линейными. Так, например, закон Ома дает линейную зависимость падения напряжения на сопротивлении от тока. Однако в действительности сопротивление проводника зависит от его температуры, которая в свою очередь зависит от протека-

ющего по проводнику тока. Вследствие этого зависимость падения напряжения от тока получается значительно более сложной и совсем не линейной. И лишь при допущении, что ток изменяется в достаточно малом диапазоне, чтобы можно было пренебречь изменением сопротивления, можно принять зависимость падения напряжения от тока в проводнике линейной. Точно так же, как мы видели в § 3.2, зависимость напряжения на выходе потенциометра от перемещения движка лишь приближенно можно считать линейной. И даже более точная нелинейная зависимость, полученная в § 3.2, является идеализированной и вследствие этого приближенной. В действительности напряжение на выходе потенциометра мало изменяется при перемещении движка до тех пор, пока имеется контакт движка с данным витком, и быстро изменяется при разрыве контакта с данным витком и входе движка в контакт с новым витком. Вследствие этого действительная зависимость напряжения на выходе потенциометра от перемещения движка близка к ступенчатой кривой, и лишь при достаточно малой толщине проволоки можно пренебречь зубцами и считать эту кривую плавной, как это было сделано в § 3.2.

Таким образом, практически любая автоматическая система является нелинейной и лишь приближенно может считаться линейной. Однако нелинейности такого характера оказывают пренебрежимо малое влияние на свойства системы, вследствие чего теория линейных систем, изложенная в предыдущих главах, дает хорошие результаты в применении к практическим задачам.

Как было отмечено в предыдущем параграфе, в состав автоматических систем часто входят существенно нелинейные элементы, которые существенно изменяют характер системы и придают ей такие свойства, которые никак не могут быть исследованы в рамках линейной теории. Среди нелинейных элементов автоматических систем особую роль играют так называемые *безынерционные* нелинейности, не обладающие заметным запаздыванием. Мы будем называть *элементарным безынерционным звеном* любую систему, выходная переменная которой в каждый данный момент времени зависит только от значения входной переменной в тот же момент времени и не зависит от того, как изменяется входная переменная до данного момента. Таким образом, оператором элементарного безынерционного звена является обычная функциональная зависимость между входной и выходной переменными. Эта функциональная зависимость называется *характеристикой* элементарного безынерционного звена.

Очевидно, что безынерционные усилители, входящие в состав линейных систем, являются элементарными безынерционными звеньями с линейными характеристиками.

В задачах практики обычно всегда удается все нелинейные свойства системы отнести к безынерционным звеньям, т. е. считать

с достаточной точностью любую систему состоящей из ряда линейных систем и элементарных нелинейных безынерционных звеньев, соединенных различными связями.

Характеристики элементарных нелинейных звеньев можно разделить на слабые нелинейности и существенные нелинейности.

К первой группе относятся такие нелинейные характеристики, которые при малом диапазоне изменения входного сигнала или при малом его отклонении от изменяющегося среднего значения могут быть заменены линейными. Отклонение от линейной зависимости между входным и выходным сигналами у таких элементов объясняется влиянием особенностей их устройства, естественным ограничением мощности источников энергии или насыщением. Такие характеристики можно аппроксимировать полиномами или выразить через однозначные аналитические функции. На рис. 8.2.1 приведена типовая характеристика усилителя

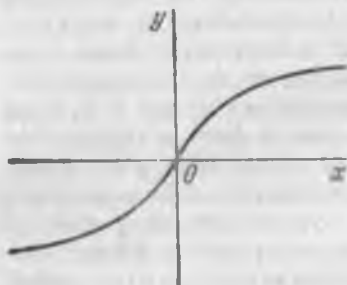


Рис. 8.2.1.

с насыщением. Насыщение наступает постепенно при больших значениях входного сигнала x . Такая зависимость может быть приближенно аппроксимирована четным полиномом невысокой степени или какой-нибудь другой дифференцируемой функцией, например линейной комбинацией синусов. Наличие подобных нелинейностей в системе часто мало сказывается на ее динамике, а именно при малых значениях входных сигналов таких нелинейностей система ведет себя как линейная и только при больших входных сигналах проявляются ее нелинейные свойства, которые необходимо учитывать.

Ко второй группе относятся нелинейные характеристики, которые являются существенно нелинейными функциями, например разрывными или близкими к разрывным. Зависимость между входной и выходной переменными для таких элементов чаще всего может быть приближенно представлена в виде кусочно-линейных функций. Пример типовой существенно нелинейной характеристики приведен на рис. 8.2.2.

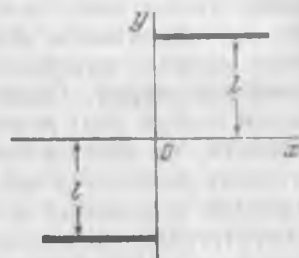


Рис. 8.2.2.

Нелинейности второй группы чрезвычайно разнообразны, а элементы с такими характеристиками имеют широкое применение в автоматических системах. Преобразование любого входного

сигнала этими нелинейными элементами всегда нелинейное. Типовые характеристики элементарных нелинейных звеньев приведены в приложении 5. Там же приведены формулы, показывающие зависимость выходной переменной y от входной переменной x . Для некоторых нелинейных звеньев выходная переменная y зависит не только от значения входной переменной x , но и от направления ее изменения, т. е. от знака скорости $\dot{x}$. Характеристики таких звеньев оказываются неоднозначными функциями входной переменной x .

Мы будем обозначать характеристику любого элементарного безынерционного нелинейного звена буквой φ :

$$y = \varphi(x). \quad (8.2.1)$$

Однако каждый раз будем оговаривать, однозначна или неоднозначна эта нелинейность или будем указывать в числе аргументов функции φ величину $\operatorname{sgn} \dot{x}$.

Нелинейности 1, 2, 3, 4 в приложении 5 характерны для различных усилителей. Нелинейность 5 характерна для устройств специального типа, применяемых для форсирования управляющего сигнала при больших рассогласованиях. Нелинейности 5 и 6 являются идеализированными характеристиками многих электромагнитных приборов и других устройств. Характеристики 7, 8 и 13 являются неоднозначными. Характеристика 7 типична для трехпозиционных, а характеристика 8 типична для двухпозиционных переключающих элементов. Нелинейность 13 характеризует обычно зависимость между входным и выходным сигналами в механических передачах. Нелинейности 9, 10, 11 характеризуют работу различных радиотехнических элементов, выпрямителей или других устройств, обеспечивающих выпрямление или функциональное нелинейное преобразование сигнала. Характеристика 12 типична для обыкновенного электромагнита или двигателя перегулируемой тяги.

Следует отметить, что приведенные в приложении 5 типовые нелинейности характеризуют не только названные элементы. Одна и та же нелинейность может выражать зависимость между входным и выходным сигналами для ряда типовых нелинейных элементов, основанных на различных физических принципах. Иными словами, одна и та же нелинейность может характеризовать функциональные свойства различных реальных элементов. Пользуясь характеристиками нелинейных звеньев, можно определить выходную переменную звена, если входная переменная задана как функция времени.

Однако при теоретических, а иногда и при экспериментальных исследованиях нелинейных систем не всегда удобно пользоваться реальными характеристиками нелинейных звеньев. Во многих случаях возможно и целесообразно заменить реальные нелиней-

ности некоторыми приближенными (в некотором смысле эквивалентными) зависимостями между входной и выходной переменными. Особенно широкое распространение получили различные методы линеаризации нелинейностей, основанные на замене действительных зависимостей между входными и выходными переменными приближенными линейными зависимостями. При этом, само собой разумеется, линеаризацию необходимо производить так, чтобы учесть хотя бы приближенно, в среднем нелинейные свойства элементов (т. е. чтобы для линеаризованных элементов не выполнялся принцип суперпозиции).

§ 8.3. Линеаризация нелинейных характеристик путем разложения в ряд

Во многих практических задачах изменение входного сигнала нелинейного звена бывает настолько малым, что нелинейная характеристика в пределах изменения входного сигнала может быть приближенно заменена линейной.

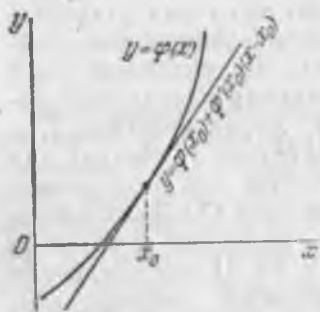


Рис. 8.3.1.

Линеаризация нелинейности состоит в замене характеристики нелинейного звена приближенной линейной зависимостью, определяемой первыми членами разложения характеристики в ряд Тейлора. Это можно сделать только для однозначных дифференцируемых функций $\varphi(x)$.

Пусть нелинейная характеристика $\varphi(x)$ является дифференцируемой функцией. Если входной сигнал x нелинейного звена мало отклоняется от некоторого среднего значения x_0 , то, применяя формулу Тейлора и отбрасывая остаточный член выше первого порядка относительно $\Delta x = x - x_0$, получим приближенную зависимость

$$y \approx \varphi(x_0) + \varphi'(x_0)(x - x_0). \quad (8.3.1)$$

Замена нелинейной зависимости (8.2.1) приближенной линейной зависимостью (8.3.1) с геометрической точки зрения представляет собой замену кривой $y = \varphi(x)$ касательной к ней в точке x_0 (рис. 8.3.1).

Реально действующие в автоматических системах сигналы обычно содержат помехи, представляющие собой случайные функции времени. Поэтому входной сигнал нелинейного звена в автоматической системе, как правило, является случайной функцией времени. Его можно представить в виде

$$X(t) = m_x(t) + X^0(t), \quad (8.3.2)$$

где $m_x(t)$ — математическое ожидание входного сигнала, включающее и его регулярную составляющую (полезный сигнал), а $X^0(t)$ — центрированная случайная составляющая входного сигнала, имеющая равное нулю математическое ожидание. В этом случае практически удобно производить линеаризацию нелинейной характеристики относительно центрированного входного случайного сигнала, т. е. взять за центр разложения математическое ожидание входного сигнала. При этом нелинейную характеристику надо разложить в ряд Тейлора относительно центрированной случайной функции X^0 и отбросить члены ряда выше первой степени. В результате получим приближенное равенство

$$Y \approx \varphi(m_x) + \varphi'(m_x) X^0. \quad (8.3.3)$$

Замена нелинейной зависимости выходного сигнала от входного приближенной зависимостью (8.3.3) при переменном математическом ожидании входного сигнала равносильна замене кривой $y = \varphi(x)$ подвижной касательной к ней в точке $x = m_x(t)$, которая «обкатывает» кривую $y = \varphi(x)$ с течением времени. Таким образом, приближенная зависимость (8.3.3) линейна только относительно случайных колебаний (флуктуаций) X^0 входного сигнала и остается нелинейной относительно полезного сигнала m_x . Вследствие этого принцип суперпозиции для линеаризованной характеристики несправедлив.

Степень точности линейной зависимости (8.3.3) может быть оценена по максимально возможной величине отброшенных членов в области практически возможных реализаций случайных сигналов [53].

§ 8.4. Гармоническая линеаризация нелинейных характеристик

В целом ряде практических задач, связанных с исследованием нелинейных систем, приходится рассматривать действие на нелинейное звено сигнала в виде синусоидальных колебаний постоянной амплитуды a :

$$x = a \sin \omega t. \quad (8.4.1)$$

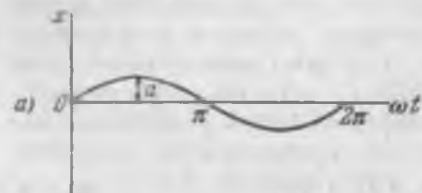
Выходной сигнал нелинейного звена в этом случае также будет периодическим. Этот периодический выходной сигнал для некоторых типовых нелинейных звеньев с нечетными характеристиками показан на рис. 8.4.1. Однако при гармонических колебаниях входной переменной колебания выходной переменной нелинейного звена уже не будут гармоническими.

Любая периодическая функция может быть разложена в ряд Фурье. Поэтому выходной сигнал нелинейного звена при синусоидальном входном возмущении всегда может быть представлен

формулой

$$y(t) = \varphi^* + \sum_{v=1}^{\infty} (a_v \sin v\omega t + b_v \cos v\omega t), \quad (8.4.2)$$

где φ^* , a_v , b_v определяются известными формулами теории рядов Фурье [50]:



$$\varphi^* = \frac{1}{2\pi} \int_0^{2\pi} \varphi(a \sin \psi) d\psi, \quad (8.4.3)$$

$$a_v = \frac{1}{\pi} \int_0^{2\pi} \varphi(a \sin \psi) \sin v\psi d\psi,$$

$$b_v = \frac{1}{\pi} \int_0^{2\pi} \varphi(a \sin \psi) \cos v\psi d\psi$$

$$(v = 1, 2, \dots).$$

(8.4.4)

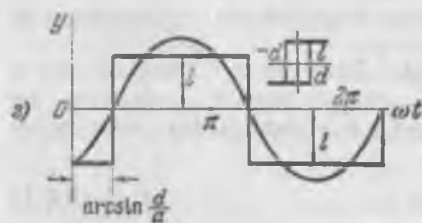
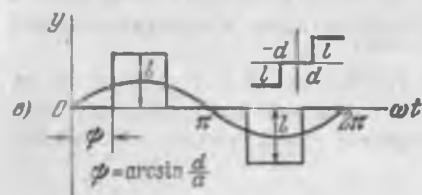
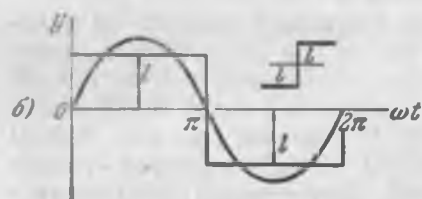


Рис. 8.4.1.

Величина φ^* представляет собой постоянную составляющую, или среднее значение выходного сигнала. Амплитуды a_v и b_v гармонических составляющих выходного сигнала быстро уменьшаются по мере увеличения номера составляющей.

Мы видим, что в отличие от линейных систем, нелинейные звенья реагируют на гармонические колебания одной определенной частоты гармоническими колебаниями различных частот (теоретически — бесконечной последовательности частот). В сумме эти гармонические колебания дают негармонические колебания выходного сигнала. Эта основная особенность

нелинейных систем присуща, как мы видим, даже простейшим нелинейным системам — элементарным безынерционным нелинейным звеньям. Именно вследствие этой особенности резонанс в нелинейной системе оказывается возможным, когда частота входного возмущения не совпадает с резонансной частотой системы, а равна резонансной частоте, деленной на целое положительное число.

Нелинейное звено в автоматической системе включается практически всегда последовательно с линейными частями системы. Линейная часть является, как правило, фильтром низких частот и подавляет высокочастотные колебания. Вследствие этого часто оказывается возможным приближенно пренебречь гармониками выше первой на выходе нелинейного звена. Тогда, принимая во внимание (8.4.1), получим следующую приближенную зависимость выходной переменной элементарного безынерционного нелинейного звена от его входной переменной:

$$y \approx \varphi^* + qx + \frac{q'}{\omega} \dot{x}, \quad (8.4.5)$$

где

$$q = \frac{a_1}{a} = \frac{1}{\pi a} \int_0^{2\pi} \psi(a \sin \psi) \sin \psi d\psi, \quad (8.4.6)$$

$$q' = \frac{b_1}{a} = \frac{1}{\pi a} \int_0^{2\pi} \psi(a \sin \psi) \cos \psi d\psi. \quad (8.4.7)$$

При неоднозначной характеристике нелинейного звена возрастанию и убыванию аргумента соответствуют разные ветви функции ψ . В этом случае интегралы в формулах (8.4.3), (8.4.4), (8.4.6) и (8.4.7) должны быть разбиты на две части, соответствующие возрастанию $\sin \psi$ в интервале от $-\frac{\pi}{2}$ до $\frac{\pi}{2}$ и убыванию $\sin \psi$ в интервале от $\frac{\pi}{2}$ до $\frac{3\pi}{2}$.

Таким образом, пренебрегая высшими гармониками (обертонами), в разложении в ряд Фурье выходной переменной, мы получили приближенную линейную зависимость между входной и выходной переменными нелинейного звена. Иными словами, под действием гармонических колебаний нелинейное звено как бы линеаризуется и может рассматриваться приближенно как линейное звено.

Коэффициенты q и q' называются *гармоническими коэффициентами усиления* нелинейного звена. Они зависят от характеристики нелинейного звена и от амплитуды входного синусоидального сигнала.

Зависимость φ^* и коэффициентов q и q' от амплитуды входного сигнала a отражает нелинейные свойства звена. Вследствие этой зависимости принцип суперпозиции неприменим для линеаризованного звена.

Приближенное представление нелинейной зависимости (8.2.1) при синусоидальном входном сигнале линейной зависимостью (8.4.5) обычно называется *гармонической линеаризацией*. Основанный на этом приеме метод исследования нелинейных систем называется *методом гармонической линеаризации* или *методом гармо-*

нического баланса. Основы этого метода заложены в работах Н. М. Крылова и Н. Н. Боголюбова [35—37]. Дальнейшее развитие метод гармонической линеаризации получил в применении к автоматическим системам в работах Л. С. Гольдфарба [15] и Е. П. Попова и И. П. Пальтова [49].

Заметим, что для любой нечетной характеристики Φ функция Φ^* равна нулю: $\Phi^* = 0$. Отметим также, что для любой однозначной характеристики Φ второй коэффициент q' равен нулю: $q' = 0$. Это соответствует отсутствию сдвига фаз между первой гармоникой на выходе нелинейного звена и входным синусоидальным сигналом. На рис. 8.4.1, б, в изображена первая гармоника для таких нелинейностей. Если нелинейная характеристика неоднозначна, то первая гармоника выходного сигнала сдвинута по фазе по отношению к входному синусоидальному колебанию. Этот случай изображен на рис. 8.4.1, г. В этом случае оба гармонических коэффициента усиления отличны от нуля.

Выражение (8.4.5) можно представить также в форме

$$y \approx \Phi^* + ca \sin(\omega t + \psi), \quad (8.4.8)$$

где

$$\left. \begin{aligned} c &= \sqrt{q^2 + q'^2}, \quad \psi = \arctg \frac{q'}{q}, \\ c \cos \psi &= q, \quad c \sin \psi = q'. \end{aligned} \right\} \quad (8.4.9)$$

Величина c характеризует усиление амплитуды, а ψ — сдвиг фазы гармонических колебаний нелинейным звеном. В отличие от линейных систем, усиление амплитуды и сдвиг фазы гармонически линеаризованным нелинейным звеном зависят от амплитуды входного сигнала.

При теоретических исследованиях автоматических систем часто удобно рассматривать гармонический входной сигнал в комплексной форме

$$x = ae^{i\omega t}. \quad (8.4.10)$$

Тогда первая гармоника выходного сигнала принимает вид

$$y - \Phi^* = cae^{i(\omega t + \psi)}. \quad (8.4.11)$$

Если ввести комплексный гармонический коэффициент усиления или амплитудно-фазовую характеристику нелинейного звена для первой гармоники

$$\Phi_n(a) = ce^{i\psi}, \quad (8.4.12)$$

то зависимость между входным и выходным сигналами нелинейного звена можно написать в форме

$$y - \Phi^* = \Phi_n(a) x. \quad (8.4.13)$$

Пользуясь формулами (8.4.9) и (8.4.12), получим для комплексного гармонического коэффициента усиления другое выражение:

$$\Phi_n(a) = q(a) + iq'(a). \quad (8.4.14)$$

Комплексный гармонический коэффициент усиления безынерционного нелинейного звена или его амплитудно-фазовая характеристика не зависит от частоты, но зависит от амплитуды гармонических колебаний. В этом состоит основное отличие нелинейного безынерционного звена от линейного инерционного. Для линейного звена, как мы видели, амплитудно-фазовая характеристика не зависит от амплитуды, но зависит от частоты входных гармонических колебаний.

Гармонические коэффициенты усиления типовых нелинейностей даны в приложении 6.

Если входной сигнал нелинейного звена представляет собой несимметричные колебания

$$x = x_0 + a \sin \omega t, \quad (8.4.15)$$

где x_0 — постоянный сигнал, то, учитывая только первую гармонику в соответствии с принципом гармонической линеаризации, получим следующее приближенное выражение для выходной переменной нелинейного звена:

$$y \approx \varphi^*(a, x_0) + q(a, x_0) a \sin \omega t + q'(a, x_0) a \cos \omega t, \quad (8.4.16)$$

где

$$\varphi^* = \frac{1}{2\pi} \int_0^{2\pi} \varphi(x_0 + a \sin \psi) d\psi, \quad (8.4.17)$$

$$q = \frac{1}{\pi a} \int_0^{2\pi} \varphi(x_0 + a \sin \psi) \sin \psi d\psi, \quad (8.4.18)$$

$$q' = \frac{1}{\pi a} \int_0^{2\pi} \varphi(x_0 + a \sin \psi) \cos \psi d\psi. \quad (8.4.19)$$

Для нечетных характеристик

$$\varphi(-x) = -\varphi(x)$$

и постоянную составляющую выходного сигнала φ^* можно представить в виде

$$\varphi^* = q_0(a, x_0) x_0, \quad (8.4.20)$$

где q_0 — коэффициент усиления по постоянной составляющей, равный

$$q_0 = \frac{1}{2\pi x_0} \int_0^{2\pi} \varphi(x_0 + a \sin \psi) d\psi. \quad (8.4.21)$$

Формулу (8.4.16) можно переписать также в виде, аналогичном (8.4.5):

$$y \approx \varphi^*(a, x_0) + q(a, x_0) (x - x_0) + \frac{q'(a, x_0)}{\omega} \dot{x}. \quad (8.4.22)$$

Пример 8.4.1. На входе нелинейного звена с характеристикой, изображенной на рис. 8.4.2, а, действует сигнал вида (8.4.15). Вычислим коэффициент усиления q_0 по постоянной составляющей и гармоническим

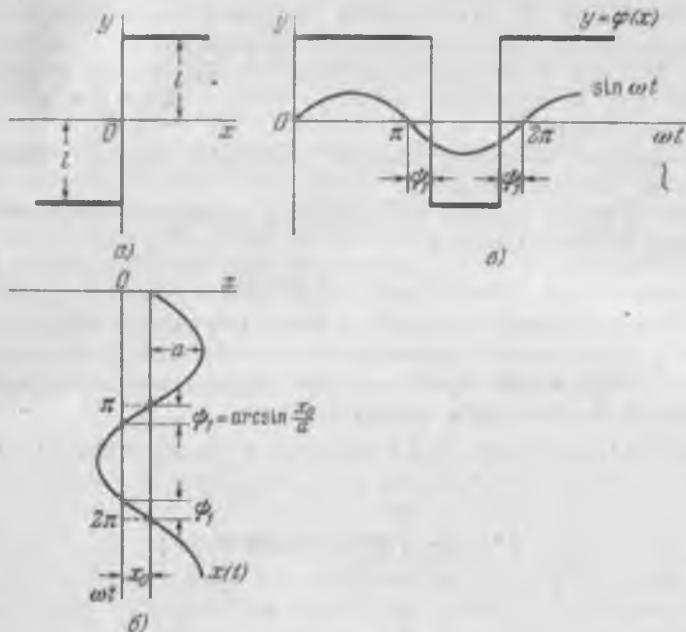


Рис. 8.4.2.

коэффициенты усиления q и q' . На рис. 8.4.2, б показан входной сигнал $x(t)$, а на рис. 8.4.2, в — выходной сигнал $y = \varphi(x)$. Применяя для вычисления коэффициента q_0 формулу (8.4.21), получим

$$q_0 = \begin{cases} \frac{2l}{\pi x_0} \arcsin \frac{x_0}{a} & \text{при } a > x_0, \\ \frac{l}{x_0} & \text{при } a < x_0. \end{cases} \quad (8.4.23)$$

Вычисляя по формуле (8.4.18) коэффициент q , получим

$$q = \begin{cases} \frac{4l}{\pi a} \sqrt{1 - \frac{x_0^2}{a^2}} & \text{при } a > x_0, \\ 0 & \text{при } a < x_0. \end{cases} \quad (8.4.24)$$

Так как характеристика нелинейного звена однозначна, то в данном случае $q' = 0$. В частном случае при $x_0 = 0$ из формул (8.4.23), (8.4.20) и (8.4.24) получаем

$$q_0 = \frac{2l}{\pi a}, \quad \varphi^* = 0, \quad q = \frac{4l}{\pi a}. \quad (8.4.25)$$

Пример 8.4.2. Вычислим гармонические коэффициенты усиления для однозначной характеристики, представленной на рис. 8.4.3, предполагая.

что на входе звена действует сигнал вида (8.4.1). В этом случае $\varphi^* = 0$, а для вычисления коэффициента q интеграл в формуле (8.4.6) должен быть записан в виде

$$q = \frac{1}{\pi a} \left\{ - \int_{-\frac{\pi}{2}}^{\psi_1} l \sin \psi d\psi + \int_{\psi_1}^{\frac{\pi}{2}} l \sin \psi d\psi \right\} + \\ + \frac{1}{\pi a} \left\{ \int_{\frac{\pi}{2}}^{\pi + \psi_1} l \sin \psi d\psi - \int_{\pi + \psi_1}^{\frac{3\pi}{2}} l \sin \psi d\psi \right\}, \quad (8.4.26)$$

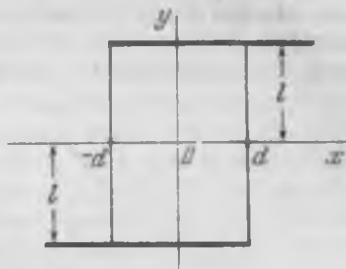


Рис. 8.4.3.

где первые два интеграла соответствуют ветви функции φ для случая возрастания $\sin \psi$, а вторые два интеграла соответствуют ветви кривой φ для случая убывания $\sin \psi$, а $\psi_1 = \arcsin \frac{d}{a}$. Вычислив интегралы, получим

$$q = \begin{cases} \frac{4l}{\pi a} \sqrt{1 - \frac{d^2}{a^2}} & \text{при } a > d, \\ 0 & \text{при } a < d. \end{cases} \quad (8.4.27)$$

Для коэффициента q' имеем

$$q' = \frac{1}{\pi a} \left\{ - \int_{-\frac{\pi}{2}}^{\psi_1} l \cos \psi d\psi + \int_{\psi_1}^{\frac{\pi}{2}} l \cos \psi d\psi \right\} + \\ + \frac{1}{\pi a} \left\{ \int_{\frac{\pi}{2}}^{\pi + \psi_1} l \cos \psi d\psi - \int_{\pi + \psi_1}^{\frac{3\pi}{2}} l \cos \psi d\psi \right\}. \quad (8.4.28)$$

Отсюда получим

$$q' = \begin{cases} -\frac{4ld}{\pi a^2} & \text{при } a > d, \\ 0 & \text{при } a < d. \end{cases} \quad (8.4.29)$$

§ 8.5. Статистическая линеаризация нелинейных характеристик

Изложенный в § 8.3 способ линеаризации нелинейностей неприемлем к нелинейным звеньям, имеющим характеристики с угловыми точками и разрывами (т. е. существенно нелинейные характеристики). Такие функции не имеют в каждой точке производной и не могут быть разложены в ряд Тейлора. Однако в предыдущем параграфе мы видели, что при синусоидальном входном сигнале

существенно нелинейные характеристики можно линеаризовать методом гармонической линеаризации, который учитывает основные законы прохождения колебаний через нелинейные звенья. Поэтому, естественно, возникает мысль и при случайном входном сигнале попытаться линеаризовать существенно нелинейные

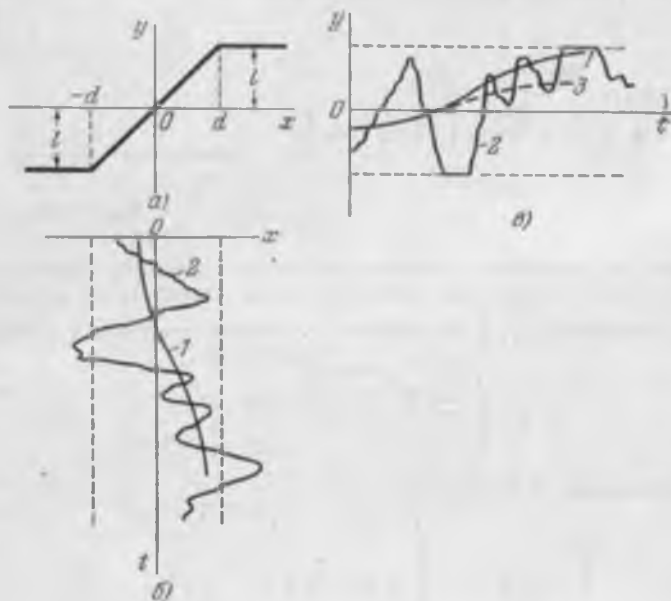


Рис. 8.5.1.

характеристики таким образом, чтобы учесть основные законы совместного прохождения полезных сигналов и случайных возмущений через нелинейные звенья.

Метод приближенной замены нелинейных характеристик эквивалентными в вероятностном смысле линейными зависимостями называется *методом статистической линеаризации* *).

Чтобы уяснить сущность метода статистической линеаризации, рассмотрим совместное прохождение плавно изменяющегося полезного сигнала и помехи через нелинейное звено. На рис. 8.5.1, а изображена характеристика нелинейного звена. На рис. 8.5.1, б кривая 1 показывает изменение полезного сигнала, который можно считать математическим ожиданием входного случайного сигнала $m_x(t)$, а кривая 2 — изменение полного входного сигнала $m_x(t) + X^0(t)$ с учетом помехи (флуктуаций) входного сигнала. На

*) Этот метод был разработан и опубликован практически одновременно И. Е. Казаковым в СССР [26] и Бутоном в США [10]. (Прим. ред.)

рис. 8.5.1, в кривая 1 показывает изменение выходного сигнала нелинейного звена при отсутствии на входе помехи, а кривая 2 — полный выходной сигнал при наличии на входе полезного сигнала и помехи. Пунктирной кривой 3 на рис. 8.5.1, в показано статистическое среднее значение (математическое ожидание) выходного сигнала. Из этих рисунков видно, что при малом уровне флуктуаций входного сигнала кривая 3 будет мало отклоняться от кривой 1. При большом уровне флуктуаций входного сигнала среднее значение выходного сигнала будет близко к нулю, так как выходной сигнал будет беспорядочно колебаться между крайними значениями. Следовательно, полезный сигнал практически не будет проходить через нелинейное звено. Иными словами, флуктуации подавляют полезный сигнал в нелинейном звене, уменьшают эффективный коэффициент усиления нелинейного звена. Чем выше уровень флуктуаций, тем сильнее подавляется полезный сигнал. При этом величина флуктуаций на выходе рассмотренного нелинейного звена также ограничивается.

Идея статистической линеаризации состоит в том, чтобы заменить истинную зависимость между входным и выходным сигналами нелинейного звена такой приближенной зависимостью, линейной относительно централизованного случайного входного сигнала, которая в основном учитывает рассмотренную физическую картину совместного прохождения полезного сигнала и случайной помехи через нелинейное звено. А именно статистически линеаризованная зависимость должна с достаточной точностью определять полезный сигнал и уровень флуктуаций на выходе нелинейного звена.

В результате статистической линеаризации зависимость между входной и выходной переменными нелинейного звена

$$Y = \varphi(X) \quad (8.5.1)$$

заменяется приближенной зависимостью вида

$$U = \varphi_0 + k_1 X^0. \quad (8.5.2)$$

Величина φ_0 представляет собой выходной полезный сигнал нелинейного звена, зависимость которого от входного полезного сигнала является *статистической характеристикой* нелинейного звена. Коэффициент k_1 представляет собой *статистический коэффициент усиления нелинейного звена по случайной составляющей*.

Для нелинейных звеньев, имеющих нечетные характеристики, функция φ_0 может быть представлена в форме

$$\varphi_0 = k_0 m_x, \quad (8.5.3)$$

где k_0 — *статистический коэффициент усиления нелинейного звена по математическому ожиданию*.

Величина φ_0 и статистические коэффициенты усиления нелинейного звена k_0 и k_1 определяются из основного условия сохранения величины полезного сигнала и уровня флуктуаций на выходе нелинейного звена. Так как полезный сигнал представляет собой математическое ожидание, а уровень флуктуаций характеризуется дисперсией выходного сигнала, то величины φ_0 , k_0 , k_1 можно определить из условия равенства математического ожидания m_u и дисперсии D_u случайной величины U , определяемой формулой (8.5.2), соответственно математическому ожиданию и дисперсии выходного сигнала Y нелинейного звена:

$$m_u = m_y, \quad D_u = D_y. \quad (8.5.4)$$

Так как случайная величина U представляет собой линейную функцию случайной величины X , то на основании известных свойств математических ожиданий и дисперсий математическое ожидание и дисперсия случайной величины U определяются формулами

$$m_u = \varphi_0, \quad D_u = k_1^2 D_x. \quad (8.5.5)$$

На основании первой из этих формул первое условие (8.5.4) дает

$$\varphi_0 = m_y. \quad (8.5.6)$$

Подставляя это значение φ_0 в (8.5.3), находим коэффициент k_0 для нечетной характеристики:

$$k_0 = \frac{m_y}{m_x}. \quad (8.5.7)$$

Подставляя выражение (8.5.5) дисперсии величины U во второе условие (8.5.4), находим коэффициент k_1 :

$$k_1 = \pm \sqrt{\frac{D_y}{D_x}} = \pm \frac{\sigma_y}{\sigma_x}, \quad (8.5.8)$$

где $\sigma_y = \sqrt{D_y}$, $\sigma_x = \sqrt{D_x}$ — средние квадратические отклонения выходного и входного сигналов нелинейного звена Y и X соответственно. Знак коэффициента k_1 определяется характером функции φ . Если функция $\varphi(x)$ возрастает около точки $x = m_x$, то следует взять $k_1 > 0$, а если функция $\varphi(x)$ убывает, то следует взять $k_1 < 0$.

Условия (8.5.4) обеспечивают точный учет изменения нелинейным звеном полезного сигнала и уровня флуктуаций. При этом, конечно, закон распределения выходного сигнала нелинейного звена при статистической линеаризации искажается. В этом отношении метод статистической линеаризации вполне равноценен методу гармонической линеаризации, который обеспечивает правильный учет первой гармоники на выходе нелинейного звена, но искажает характер колебаний выходного сигнала.

Учитывая, что нелинейные звенья в автоматических системах работают совместно с линейными, для которых законы преобразования случайных сигналов определяются не столько их дисперсиями, сколько корреляционными функциями, приходим к заключению, что можно допустить при статистической линеаризации некоторую ошибку в дисперсии выходного сигнала с целью лучше приблизить его корреляционную функцию.

На основании изложенных соображений возникает идея положить в основу метода статистической линеаризации другой принцип, а именно принцип минимума средней квадратической ошибки от замены нелинейной зависимости (8.5.1) приближенной линейной зависимостью (8.5.2):

$$\eta = M[(Y - U)^2] = \min. \quad (8.5.9)$$

Подставляя сюда выражение для U из (8.5.2) и используя известные свойства математических ожиданий, получим следующее условие:

$$\eta = M[Y^2] + \varphi_0^2 + k_1^2 D_x - 2\varphi_0 m_y - 2k_1 M[YX^0] = \min. \quad (8.5.10)$$

При известных m_y , $M[Y^2]$, D_x и $M[YX^0]$ величина η является функцией параметров φ_0 и k_1 . Приравнявая нулю частные производные этой функции по φ_0 и k_1 , получим уравнения

$$\varphi_0 - m_y = 0, \quad (8.5.11)$$

$$k_1 D_x - M[YX^0] = 0. \quad (8.5.12)$$

Уравнение (8.5.11) дает для φ_0 ранее полученную формулу (8.5.6). Решая уравнение (8.5.12), получаем следующую формулу для коэффициента k_1 :

$$k_1 = \frac{1}{D_x} M[YX^0]. \quad (8.5.13)$$

Легко показать, что найденные выражения φ_0 и k_1 действительно обеспечивают минимум величины η . Для этого дадим величинам φ_0 и k_1 в выражении (8.5.10) приращения, равные соответственно u_0 и u_1 . Тогда получим

$$\begin{aligned} \eta &= M[(Y - \varphi_0 - u_0 - (k_1 + u_1)X^0)^2] = \\ &= M[(Y - \varphi_0 - k_1 X^0)^2] + u_0^2 + u_1^2 D_x + \\ &\quad + 2u_0(\varphi_0 - m_y) + 2u_1(k_1 D_x - M[YX^0]). \end{aligned} \quad (8.5.14)$$

Если φ_0 и k_1 определяются уравнениями (8.5.11) и (8.5.12), то формула (8.5.14) принимает вид

$$\eta = M[(Y - \varphi_0 - k_1 X^0)^2] + u_0^2 + u_1^2 D_x. \quad (8.5.15)$$

Отсюда видно, что η принимает минимальное значение только при $u_0 = u_1 = 0$, т. е. при φ_0 и k_1 , определяемых уравнениями (8.5.11) и (8.5.12).

Для вычисления величин k_0 , k_1 и φ_0 по полученным формулам необходимо знать одномерную плотность вероятности $f_1(x; t)$ входной случайной функции $X(t)$ нелинейного звена, которую мы будем обозначать сокращенно $f(x)$. Тогда формулы (8.5.6), (8.5.7), (8.5.8) и (8.5.13) примут вид

$$\varphi_0 = \int_{-\infty}^{\infty} \varphi(x) f(x) dx, \quad (8.5.16)$$

$$k_0 = \frac{1}{m_x} \int_{-\infty}^{\infty} \varphi(x) f(x) dx, \quad (8.5.17)$$

$$k_1^{(1)} = \pm \frac{1}{\sigma_x} \left\{ \int_{-\infty}^{\infty} \varphi^2(x) f(x) dx - \varphi_0^2 \right\}^{1/2} \quad (8.5.18)$$

$$k_1^{(2)} = \frac{1}{\sigma_x^2} \int_{-\infty}^{\infty} (x - m_x) \varphi(x) f(x) dx, \quad (8.5.19)$$

где индексы сверху у коэффициентов k_1 указывают способ линеаризации.

Пользуясь формулами (8.5.16)–(8.5.19), можно определить функцию φ_0 и статистические коэффициенты усиления k_0 , k_1 для типовых нелинейных звеньев при заданной одномерной плотности вероятности $f(x)$ входной случайной функции. Однако при работе нелинейного звена в замкнутой системе плотность вероятности его входного сигнала заранее неизвестна. Даже и после проведенного исследования системы часто удается определить только некоторые числовые характеристики входных случайных сигналов нелинейных звеньев. Выход из этого затруднения достаточно прост. Дело в том, что изменение формы закона распределения $f(x)$ в широких пределах не оказывает существенного влияния на коэффициенты k_0 и k_1 и функцию φ_0 . Кроме того, пелтиевые звенья, как правило, соединяются в системах с инерционными линейными звеньями, а закон распределения выходной переменной инерционной линейной системы оказывается обычно близким к нормальному при любом законе распределения входной переменной. Закон распределения выходной переменной тем ближе к нормальному, чем инерционнее линейная система (т. е. чем больше ее постоянные времени). Учитывая эти соображения, можно принять в формулах (8.5.16)–(8.5.19) закон распределения входной переменной нелинейного звена нормальным:

$$f(x) = \frac{1}{\sigma_x \sqrt{2\pi}} \exp \left\{ -\frac{(x - m_x)^2}{2\sigma_x^2} \right\}. \quad (8.5.20)$$

При этом статистические коэффициенты усиления нелинейного звена k_0 , $k_1^{(1)}$, $k_1^{(2)}$ и его статистическая характеристика Φ_0 будут функциями математического ожидания m_x и среднего квадратического отклонения σ_x входного сигнала.

При нормальном законе распределения входного сигнала функция Φ_0 выражается формулой

$$\Phi_0 = \int_{-\infty}^{\infty} \frac{\varphi(x)}{\sqrt{2\pi}\sigma_x} e^{-\frac{(x-m_x)^2}{2\sigma_x^2}} dx. \quad (8.5.21)$$

Если продифференцировать это выражение по m_x и сравнить результат с (8.5.19), то получим

$$k_1^{(2)} = \frac{\partial \Phi_0}{\partial m_x}. \quad (8.5.22)$$

Следовательно, при нормальном законе распределения входного сигнала нелинейного звена коэффициент $k_1^{(2)}$ можно вычислять по формуле (8.5.22).

Таким образом, в отличие от обычной линеаризации при помощи ряда Тейлора, статистическая линеаризация дает выходной полезный сигнал и коэффициент передачи случайной составляющей как функции не только полезного сигнала, но и уровня флуктуаций входного сигнала. При этом, так же как и обычная линеаризация, статистическая линеаризация заменяет характеристику нелинейного звена функцией, линейной только относительно центрированной случайной составляющей входного сигнала, но нелинейной относительно полезного сигнала. Вследствие этого принцип суперпозиции не выполняется для статистически линеаризованных звеньев, и таким образом их основные нелинейные свойства сохраняются.

В случае неоднозначных характеристик нелинейных звеньев функция $\varphi(x)$ в формулах (8.5.16)–(8.5.19) заменяется функцией $\varphi(x, \operatorname{sgn} \dot{x})$, одномерная плотность вероятности $f(x)$ входной случайной функции $X(t)$ заменяется совместной плотностью вероятности $f(x, \dot{x})$ значений случайной функции $X(t)$ и ее производной $X'(t)$ при любом данном t , а интегралы заменяются двойными интегралами по переменным $x, \dot{x}$.

Как показывают расчеты, при выборе k_1 по первому и второму способам в корреляционной функции выходного сигнала получаются односторонние ошибки разных знаков. Поэтому в качестве коэффициента k_1 целесообразно брать среднее арифметическое значений этого коэффициента, полученных по первому и второму способам:

$$k_1 = \frac{k_1^{(1)} + k_1^{(2)}}{2}. \quad (8.5.23)$$

В приложении 6 даны расчетные формулы для статистических коэффициентов усиления k_0 , k_1 и статистической характеристики φ_0 типовых нелинейностей при нормальном законе распределения входного сигнала *).

Пример 8.5.1. Вычислим статистические коэффициенты усиления k_0 , $k_1^{(1)}$, $k_1^{(2)}$ для нелинейности 5 (приложение 5), представленной на рис. 8.4.2, а.

Так как нелинейность в данном случае нечетная, то, применяя формулу (8.5.17), находим статистический коэффициент усиления по полезному сигналу:

$$k_0 = -\frac{l}{m_x \sqrt{2\pi} \sigma_x} \int_{-\infty}^0 \exp \left\{ -\frac{(x-m_x)^2}{2\sigma_x^2} \right\} dx + \\ + \frac{l}{m_x \sqrt{2\pi} \sigma_x} \int_0^{\infty} \exp \left\{ -\frac{(x-m_x)^2}{2\sigma_x^2} \right\} dx.$$

После замены переменных $x - m_x = t\sigma_x$ получим

$$k_0 = \frac{l}{m_x} \left\{ \frac{1}{\sqrt{2\pi}} \int_{-\frac{m_x}{\sigma_x}}^{\infty} e^{-\frac{t^2}{2}} dt - \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{\frac{m_x}{\sigma_x}} e^{-\frac{t^2}{2}} dt \right\}$$

или

$$k_0 = \frac{2l}{m_x} \Phi \left(\frac{m_x}{\sigma_x} \right), \quad (8.5.24)$$

где $\Phi(x)$ — известная в теории вероятностей функция. Таблица этой функции дана в приложении 7.

Аналогично, применяя формулы (8.5.18) и (8.5.19) и производя ту же замену переменных, находим

$$k_1^{(1)} = \frac{l}{\sigma_x} \sqrt{1 - 4\Phi^2 \left(\frac{m_x}{\sigma_x} \right)}, \quad (8.5.25)$$

$$k_1^{(2)} = \frac{l}{\sigma_x} \frac{2}{\sqrt{2\pi}} \exp \left\{ -\frac{1}{2} \frac{m_x^2}{\sigma_x^2} \right\}. \quad (8.5.26)$$

Формула (8.5.26) может быть получена также путем дифференцирования по m_x равенства

$$\varphi_0 = k_0 m_x = 2l\Phi \left(\frac{m_x}{\sigma_x} \right). \quad (8.5.27)$$

*) Здесь изложен метод статистической линейаризации И. Е. Казакова. Метод статистической линейаризации американского ученого Бутона основан на замене характеристики нелинейного звена линейной зависимостью $U = kX$ с одним общим коэффициентом усиления для полезного сигнала и помехи, который определяется из условия (8.5.9). Статистическая линейаризация И. Е. Казакова совершеннее статистической линейаризации Бутона, так как она точнее учитывает основные законы совместного прохождения полезного сигнала и помехи через нелинейности. Относительно дальнейшего развития метода статистической линейаризации см. [27]. (Прим. ред.)

§ 8.6. Совместная гармоническая и статистическая линеаризация нелинейных характеристик

Входной сигнал нелинейного звена автоматической системы часто представляет собой сумму синусоидального сигнала определенной частоты и амплитуды и случайной функции $X(t)$:

$$Z(t) = a \sin \omega t + X(t) = m_x + a \sin \omega t + X^0(t). \quad (8.6.1)$$

Если применить в этом случае метод статистической линеаризации, то статистическая характеристика нелинейного звена и статистические коэффициенты усиления по математическому ожиданию и случайной составляющей будут периодическими функциями времени в силу периодической зависимости от времени математического ожидания входного сигнала. При применении метода гармонической линеаризации гармонические коэффициенты усиления будут случайными величинами в силу наличия случайной составляющей во входном сигнале. Поэтому в подобном случае целесообразно применить совместную статистическую и гармоническую линеаризацию, т. е. заменить характеристику нелинейного звена

$$Y(t) = \varphi(Z(t)) \quad (8.6.2)$$

приближенной зависимостью

$$Y(t) \approx \varphi_0^* + \kappa a \sin \omega t + \kappa' a \cos \omega t + \kappa_1 X^0(t), \quad (8.6.3)$$

линейной относительно синусоидальной и центрированной случайной составляющих входного сигнала. При этом, так же как в случаях одной гармонической и одной статистической линеаризации, полезную составляющую φ_0^* выходного сигнала нелинейного звена можно в случае нечетной характеристики φ принять пропорциональной систематической составляющей входного сигнала (т. е. полезному входному сигналу) m_x :

$$\varphi_0^* = \kappa_0 m_x. \quad (8.6.4)$$

Величины φ_0^* , κ , κ' , κ_0 и κ_1 можно определить различными способами из условия правильного учета передачи нелинейным звеном полезного сигнала, первой гармоники и уровня флуктуаций выходного сигнала.

Предположим, что математическое ожидание m_x и дисперсия $D_x = \sigma_x^2$ случайной части $X(t)$ входного сигнала $Z(t)$ нелинейного звена изменяются достаточно медленно для того, чтобы можно было считать их постоянными в пределах одного периода синусоидальной части входного сигнала. Производя статистическую линеаризацию характеристики (8.5.1) нелинейного звена изложенным в предыдущем параграфе методом, получим следующую приближенную зависимость между входным и выходным сигналами

нелинейного звена:

$$Y(t) \approx \varphi_0(m_x + a \sin \omega t, \sigma_x) + k_1(m_x + a \sin \omega t, \sigma_x) X^n(t). \quad (8.6.5)$$

Здесь мы показываем в явной форме зависимость φ_0 и k_1 от математического ожидания $m_z = m_x + a \sin \omega t$ и среднего квадратического отклонения $\sigma_z = \sigma_x$ входного сигнала нелинейного звена.

Вследствие наличия периодической составляющей в математическом ожидании входного сигнала функции φ_0 и k_1 оказываются периодическими функциями времени. Следовательно, к ним можно применить метод гармонической линеаризации. Согласно этому методу представляем функцию φ_0 при данном фиксированном значении σ_x рядом Фурье и оставляем в этом ряде только постоянную (систематическую) составляющую и первую гармонику. В результате получаем

$$\varphi_0(m_x + a \sin \omega t, \sigma_x) \approx \varphi_0^*(a, m_x, \sigma_x) + \kappa(a, m_x, \sigma_x) a \sin \omega t + \kappa'(a, m_x, \sigma_x) a \cos \omega t, \quad (8.6.6)$$

где согласно формулам (8.4.17), (8.4.18) и (8.4.19)

$$\varphi_0^*(a, m_x, \sigma_x) = \frac{1}{2\pi} \int_0^{2\pi} \varphi_0(m_x + a \sin \psi, \sigma_x) d\psi, \quad (8.6.7)$$

$$\kappa(a, m_x, \sigma_x) = \frac{1}{\pi a} \int_0^{2\pi} \varphi_0(m_x + a \sin \psi, \sigma_x) \sin \psi d\psi, \quad (8.6.8)$$

$$\kappa'(a, m_x, \sigma_x) = \frac{1}{\pi a} \int_0^{2\pi} \varphi_0(m_x + a \sin \psi, \sigma_x) \cos \psi d\psi. \quad (8.6.9)$$

Разложив в ряд Фурье статистический коэффициент усиления нелинейного звена по случайной составляющей k_1 в формуле (8.6.5), мы должны оставить только постоянную составляющую, чтобы избежать появления нелинейного члена — произведения синусоидальной и центрированной случайной составляющих. В результате получим приближенную формулу

$$k_1(m_x + a \sin \omega t, \sigma_x) \approx k_1(a, m_x, \sigma_x) = \frac{1}{2\pi} \int_0^{2\pi} k_1(m_x + a \sin \psi, \sigma_x) d\psi. \quad (8.6.10)$$

Подставляя выражения (8.6.6) и (8.6.10) в формулу (8.6.5), мы и получим приближенную линеаризованную зависимость (8.6.3) между составляющими входного сигнала и выходным сигналом нелинейного звена.

Функции φ_0 и k_1 в (8.6.7)–(8.6.10) вычисляются по формулам (8.5.16), (8.5.18), (8.5.19) и (8.5.21):

$$\varphi_0(m_x + a \sin \psi, \sigma_x) = \int_{-\infty}^{\infty} \varphi(x + a \sin \psi) f(x) dx, \quad (8.6.11)$$

$$k_1(m_x + a \sin \psi, \sigma_x) = \frac{1}{2} [k_1^{(1)}(m_x + a \sin \psi, \sigma_x) + k_1^{(2)}(m_x + a \sin \psi, \sigma_x)], \quad (8.6.12)$$

$$k_1^{(1)}(m_x + a \sin \psi, \sigma_x) = \pm \frac{1}{\sigma_x} \left\{ \int_{-\infty}^{\infty} \varphi^2(m_x + a \sin \psi) f(x) dx - \varphi_0^2(m_x + a \sin \psi, \sigma_x) \right\}^{1/2}, \quad (8.6.13)$$

$$k_1^{(2)}(m_x + a \sin \psi, \sigma_x) = \frac{1}{\sigma_x^2} \int_{-\infty}^{\infty} (x - m_x) \varphi(x + a \sin \psi) f(x) dx \quad (8.6.14)$$

для нормальной плотности вероятности $f(x)$, определяемой формулой (8.5.20).

В случае нечетной характеристики нелинейного звена формулы (8.6.4) и (8.6.7) дают

$$x_0(a, m_x, \sigma_x) = \frac{q_0^*(a, m_x, \sigma_x)}{m_x} = \frac{1}{2\pi m_x} \int_0^{2\pi} \varphi_0(m_x + a \sin \psi, \sigma_x) d\psi. \quad (8.6.15)$$

Полученные формулы показывают, что статистические коэффициенты усиления x_0 и x_1 представляют собой усредненные за период изменения гармонической составляющей статистические коэффициенты усиления k_0 и k_1 . Коэффициенты x и x' являются гармоническими коэффициентами усиления для статистической характеристики φ_0 , получаемой из функции φ статистическим усреднением, как это следует из (8.6.11). Наконец, функция φ_0^* есть постоянная составляющая функции φ_0 , вычисленная путем усреднения φ_0 за период изменения гармонической составляющей.

Другим способом определения φ_0^* , x_0 , x , x' , x_1 в (8.6.3) и (8.6.4), является применение сначала метода гармонической линеаризации, а потом метода статистической линеаризации. При этом полученные в результате гармонической линеаризации постоянная составляющая φ^* и гармонические коэффициенты усиления q и q' будут вследствие формул (8.4.17) — (8.4.19) функциями амплитуды синусоидальной составляющей a и случайной составляющей X .

Применяя к постоянной составляющей φ^* метод статистической линейаризации и производя статистическое усреднение коэффициентов q и q' , мы снова получим формулы (8.6.7) — (8.6.15). Центрированными случайными составляющими в коэффициентах q и q' следует пренебречь, чтобы избежать появления произведения синусоидальной и центрированной случайной составляющих. Самой собой разумеется, порядок интегрирования в формулах (8.6.7) — (8.6.15) можно изменить и произвести интегрирование сначала по φ , а затем по x .

Величины φ_0^* , x_0 , x , x' , x_1 можно также определить из условия совпадения средних значений математических ожиданий и дисперсий случайных функций $Y(t)$ и $U(t)$ за период синусоидальной составляющей (первый способ) или из условия минимума среднего значения за период среднего квадрата ошибки (8.5.9) (второй способ). При этом снова получатся формулы (8.6.7) — (8.6.15).

При теоретических исследованиях входной сигнал иногда удобно представить в комплексной форме, как и в § 8.4:

$$Z(t) = m_x + ae^{i\omega t} + X^0(t). \quad (8.6.16)$$

Тогда, применяя совместную гармоническую и статистическую линейаризацию, представим функцию Y в виде

$$Y(t) = \varphi_0^* + a\Psi_{\Pi}e^{i\omega t} + x_1X^0, \quad (8.6.17)$$

где

$$\Psi_{\Pi} = x + ix'. \quad (8.6.18)$$

Выведенные формулы показывают, что совместная статистическая и гармоническая линейаризация нелинейных характеристик дает приближенную линейную зависимость выходного сигнала нелинейного звена от синусоидальной и центрированной случайной составляющих входного сигнала и в то же время учитывает взаимодействие систематической, синусоидальной и случайной составляющих при их совместном прохождении через нелинейное звено. Это выражается в том, что коэффициенты линейаризованной зависимости определяются как функции амплитуды синусоидальной составляющей и математического ожидания и дисперсии случайной составляющей входного сигнала. Иными словами, совместная статистическая и гармоническая линейаризация заменяет нелинейную зависимость между самими сигналами приближенной нелинейной зависимостью между их основными параметрами — постоянными (или медленно изменяющимися) составляющими, амплитудами и фазами синусоидальных составляющих и дисперсиями случайных составляющих, — а зависимость между быстро изменяющимися (синусоидальными и случайными) составляющими заменяет приближенной линейной зависимостью. В этом и заключается сила методов статистической и гармонической линейаризации, причина их широкого распространения и практической эффектив-

ности. Заменяя нелинейные зависимости между быстро изменяющимися составляющими сигналов линейными, методы статистической и гармонической линеаризации дают возможность использовать для приближенного исследования нелинейных систем хорошо разработанные методы теории линейных систем. Учитывая приближенно зависимость между основными параметрами постоянных, синусоидальных и случайных составляющих сигналов, методы статистической и гармонической линеаризации с достаточной в большинстве случаев для практики точностью отражают нелинейные свойства нелинейных звеньев и вследствие этого дают возможность приближенно исследовать различные процессы в нелинейных системах.

В некоторых случаях при изучении колебаний в нелинейных системах можно не интересоваться случайными составляющими выходных сигналов нелинейных звеньев, однако полностью пренебречь случайными составляющими входных сигналов не удастся вследствие их значительного влияния на амплитуды и частоты колебаний. В таких случаях можно пренебречь последним, случайным слагаемым в формуле (8.6.3). При этом зависимость параметров колебаний от случайных возмущений будет учтена в форме зависимости величин φ_0^* , χ_0 , χ и χ' от дисперсии случайной составляющей σ_x^2 .

В приложении 6 приведены формулы для величин χ , χ_1 при отсутствии во входном сигнале постоянной составляющей для типовых нелинейных характеристик.

В главах 10 и 11 будет показано применение методов гармонической и статистической линеаризации для приближенного исследования нелинейных систем.

НЕЛИНЕЙНЫЕ ЭЛЕМЕНТЫ АВТОМАТИЧЕСКИХ СИСТЕМ

§ 9.1. Электролитические датчики вертикали

Электролитический датчик является измерителем отклонения данного направления от вертикали. Конструктивно датчик выполнен в виде ограниченной частью сферы полости, заполненной токопроводящей жидкостью — электролитом (рис. 9.1.1). В корпусе расположены четыре электрода, изолированных друг от друга. Электролит не заполняет всего объема полости, и образующийся воздушный пузырек при вертикальном положении датчика накрывает около половины площади каждого из четырех электродов. При отклонении датчика от вертикали изменяется соотношение площадей перекрытия электродов электролитом и воздушным пузырьком. В результате изменяется проводимость электрической цепи и ток в ней. На этом основан принцип действия датчика.

Электрическая схема электролитического датчика показана на рис. 9.1.2. К электродам через обмотки трансформаторов $Tr1$, $Tr2$ подведено переменное напряжение u . Через электроды датчика протекает переменный ток, величина которого определяется проводимостью G участка электрод — электролит — корпус датчика:

$$G = G_0 + \nu S, \quad (9.1.1)$$

где G_0 — проводимость при полном накрытии электрода воздушным пузырьком, S — площадь накрытия электрода электролитом, ν — коэффициент пропорциональности.

Напряжения u_1 , u_2 на вторичных обмотках трансформаторов $Tr1$ и $Tr2$ являются выходными сигналами и характеризуют углы отклонения α , β оси датчика от вертикали в двух взаимно перпендикулярных плоскостях. Амплитуды напряжений пропорциональны величинам углов α и β , а фазы определяют направления отклонений (знаки углов α и β). Определим зависимость напряжения u_1 от угла α . Напряжение u_1 определяется разностью токов i_1 и i_3 в первичной обмотке трансформатора $Tr1$:

$$u_1 = Z (i_1 - i_3), \quad (9.1.2)$$

где $Z = \sqrt{R^2 + \omega^2 L^2}$ — полное сопротивление первичной обмотки $Tr1$. Токи i_1 и i_3 определяются проводимостями цепей 1 и 3:

$$i_1 = uG_1 = u (G_0 + \nu S_1), \quad i_3 = uG_3 = u (G_0 + \nu S_3). \quad (9.1.3)$$

Подставляя эти выражения в (9.1.2), получим

$$u_1 = Zv(S_1 - S_3) a. \quad (9.1.4)$$

Таким образом, амплитуда напряжения u_1 , равная $Zv(S_1 - S_3) a$, где a — амплитуда питающего напряжения u , пропорциональна

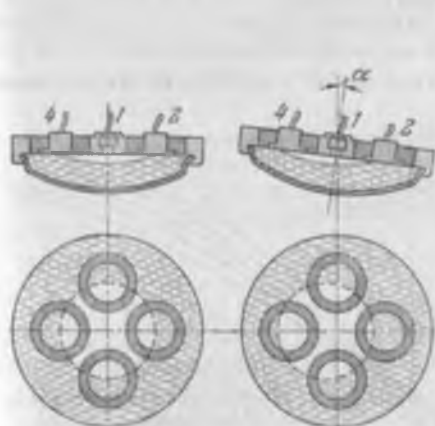


Рис. 9.1.1.

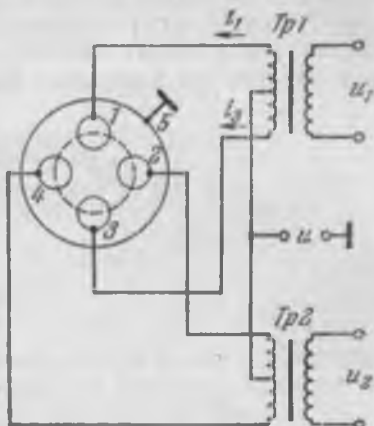


Рис. 9.1.2.

разности площадей покрытия электродов электролитом. При вертикальном положении датчика $S_1 = S_3$ и $u_1 = 0$.

Зависимость разности площадей $\Delta S = S_1 - S_3$ от угла наклона датчика, вычисленная для одного из промышленных образцов датчика, представлена на рис. 9.1.3. Эта зависимость нелинейна, причем нелинейность выражается в наличии ограничения. Участки ограничения соответствуют полному перекрытию площадей электрода воздушным пузырьком или электролитом. Такой же характер в силу формулы (9.1.4) имеет зависимость амплитуды a_1 выходного напряжения датчика u_1 от угла отклонения оси датчика от вертикали α :

$$a_1 = Zva(S_1 - S_3). \quad (9.1.5)$$

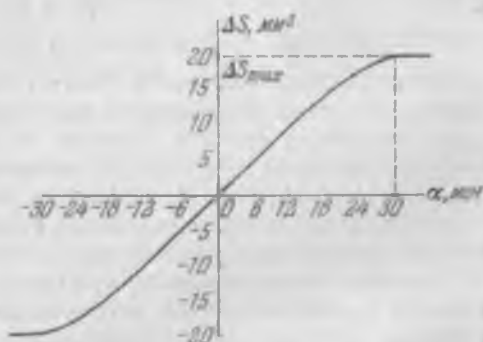


Рис. 9.1.3.

Реальная характеристика датчика несколько отличается от вычисленной вследствие изменения формы воздушного пузырька при отклонении датчика от вертикали.

Нелинейную характеристику электролитического датчика можно приближенно представить в виде кусочно-линейной характеристики типа чистого ограничения (приложение 5, нелинейность 1). Кроме рассмотренной нелинейности, в электролитическом датчике имеется еще одна существенная нелинейность, обусловленная взаимным влиянием каналов. При наличии угла наклона датчика в другой плоскости β изменяется крутизна характеристики, определяющей зависимость амплитуды выходного напряжения u_1 от угла α . На рис. 9.1.4 представлено семейство характеристик датчика,

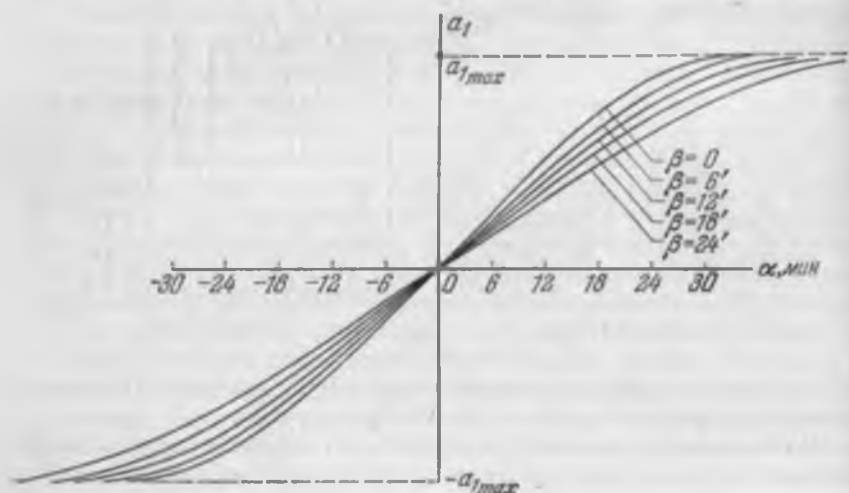


Рис 9.1.1.

соответствующих плоскости угла α , в зависимости от параметра β . Таким образом, параметры идеализированной кусочно-линейной характеристики датчика зависят от угла β . Точно так же параметры характеристики датчика, соответствующей наклону оси датчика в плоскости угла β , зависят от угла α . Эти зависимости необходимо учитывать при вычислении гармонических и статистических коэффициентов передачи электролитического датчика.

Изложенное показывает, что электролитический датчик можно считать безынерционным нелинейным звеном типа ограничителя с двумя взаимно связанными каналами.

Электролитические датчики нашли широкое распространение в гироскопических приборах как первичные датчики вертикали места. Простота устройства, небольшие габариты и высокая чувствительность (пороговая чувствительность порядка $1'-2'$) являются их положительными качествами.

§ 9.2. Термосопротивления

Термосопротивление является датчиком температуры. Чувствительным элементом датчика служит платиновая или медная проволока 1 (рис. 9.2.1), намотанная на каркас из изоляционного материала 2. Каркас с проволокой помещен в защитную трубку 3 из алюминия или латуни. При изменении температуры среды,

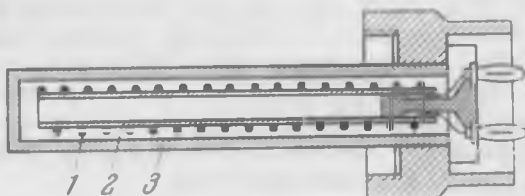


Рис. 9.2.1.

в которой находится датчик, изменяется сопротивление проводника (проволоки). На этом основан принцип действия датчика.

Зависимость электрического сопротивления проводника от его температуры выражается формулой

$$R = R_0 \varphi(\theta - \theta_0), \quad (9.2.1)$$

где θ — температура проводника, R_0 — сопротивление проводника при температуре θ_0 , $\varphi(\theta - \theta_0)$ — определенная функция разности температур. В широком диапазоне температур функция $\varphi(\theta - \theta_0)$ хорошо аппроксимируется экспоненциальной зависимостью

$$\varphi(\theta - \theta_0) = e^{-\alpha(\theta - \theta_0)}. \quad (9.2.2)$$

В узком диапазоне температур функция $\varphi(\theta - \theta_0)$ с достаточной точностью аппроксимируется полиномом

$$\varphi(\theta - \theta_0) = 1 + \alpha(\theta - \theta_0) + \beta(\theta - \theta_0)^2 + \dots \quad (9.2.3)$$

Коэффициенты α , β , ... в выражениях (9.2.2) и (9.2.3) определяются экспериментально.

Выведем зависимость температуры проводника θ от температуры среды τ , в которой он находится. Для этого вычислим изменение количества тепла dQ , содержащегося в проводнике, при изменении его температуры на $d\theta$:

$$dQ = mc d\theta, \quad (9.2.4)$$

где c — теплоемкость материала проводника, m — его масса. С другой стороны, согласно закону Ньютона, количество тепла, получаемое датчиком от внешней среды, пропорционально разности температур среды и датчика, площади поверхности датчика

и времени:

$$dQ = kS (\tau - \theta) dt, \quad (9.2.5)$$

где k — коэффициент теплопередачи. На основании закона сохранения энергии левые части в уравнениях (9.2.4) и (9.2.5) равны. Поэтому из (9.2.4) и (9.2.5) вытекает дифференциальное уравнение, связывающее температуру датчика θ и температуру среды τ :

$$T\dot{\theta} + \theta = \tau, \quad (9.2.6)$$

где $T = cm/kS$ — постоянная времени термосопротивления. Таким образом, динамические свойства термосопротивления описываются уравнениями (9.2.1) и (9.2.6), которые определяют зависимость сопротивления датчика, а следовательно и тока в нем, от температуры среды. На основании этих уравнений термосопротивление можно представить в виде последовательного соединения апериодического звена и нелинейности с экспоненциальной или квадратичной характеристикой (рис. 9.2.2).

Термосопротивления обычно включаются в мостовые схемы. На рис. 9.2.3 показана возможная схема включения датчика. Термосопротивление R_θ включено в одно из плеч электрического моста.

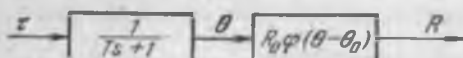


Рис. 9.2.2.

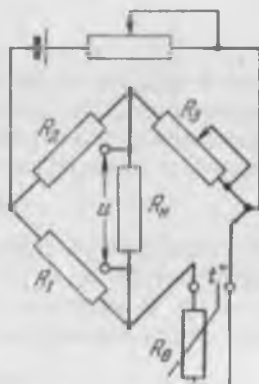


Рис. 9.2.3.

Переменным сопротивлением R_θ мост балансируется на определенную температуру θ_0 . При изменении температуры среды τ и, следовательно, температуры датчика θ мост разбалансируется и через сопротивление нагрузки R_n , включенное в диагональ моста, течет ток. Входной величиной датчика является температура окружающей среды τ , выходной — напряжение u на сопротивлении нагрузки R_n .

Термосопротивления нашли широкое распространение для измерения температур различных сред в пределах $200-1000^\circ\text{C}$. В авиационной технике они применяются для измерения температуры в прецизионных гироскопических приборах, температуры водяной и масляной систем двигателя, температуры газов газотурбинных двигателей и т. д. Параметры некоторых электрических термосопротивлений для случая, когда функция φ выражается

формулой (9.2.3), имеют порядок $R_0 = 50 \div 100 \text{ Ом}$; $T = 20 \text{ с} \div 10 \text{ мин}$; $\alpha = 0,00428 \text{ град}^{-1}$ для проводниковой меди в диапазоне $0 \div 200^\circ \text{C}$; $\alpha = 0,00394 \text{ град}^{-1}$, $\beta = -5,8 \cdot 10^7 \text{ град}^2$ для платиновых термосопротивлений в диапазоне $200 \div 800^\circ \text{C}$; $\alpha = 0,0121 \text{ град}^{-1}$ для теллуросеребряного термосопротивления.

Особенно широкое распространение в последнее время получили так называемые *термисторы* — полупроводники с большим отрицательным температурным коэффициентом. Температурный коэффициент у термисторов больше, чем у металлов, в 5—10 раз. Поэтому схемы тепловых датчиков с термисторами обладают высокой чувствительностью. Термисторы имеют более высокие динамические свойства, чем металлические проводники. Постоянная времени термисторов в зависимости от размеров и формы датчиков имеет значение $T = 0,01 \div 50 \text{ с}$ [47].

§ 9.3. Магнитные усилители

Магнитные усилители нашли широкое применение в системах автоматического управления как усилители мощности и как суммирующие устройства (см., например, [14] и [60]). На вход магнитного усилителя действуют сигналы постоянного напряжения или

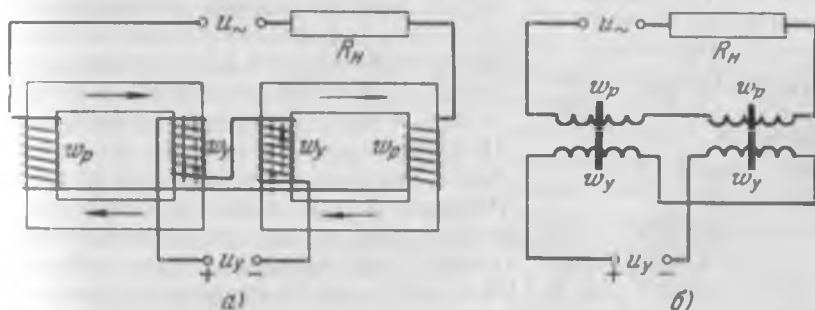


Рис. 9.3.1.

тока. Выходным сигналом усилителя является переменный ток. Закон изменения этого тока зависит от режима работы усилителя. Часто в схему усилителя включаются выпрямительные элементы, с помощью которых переменный ток преобразуется в постоянный.

Простейший магнитный усилитель с одним входом представлен на рис. 9.3.1, а. Он состоит из двух одинаковых *дросселей насыщения*. На внутренние стержни сердечников дросселей намотана обмотка управления с числом витков $2w_y$, на внешние стержни намотана рабочая обмотка с числом витков $2w_p$. Рабочая обмотка питается переменным напряжением u_m . В цепь рабочей обмотки включена нагрузка с сопротивлением R_n . К управляющей обмотке

подается входной сигнал u_{γ} . Управляющая обмотка наматывается так, чтобы направления магнитных потоков Φ_{-} , создаваемых ею во внутренних стержнях обоих дросселей, совпадали. Рабочая обмотка наматывается так, чтобы направления потоков Φ_{+} , создаваемых ею во внутренних стержнях, были противоположны. На рис. 9.3.1, а потоки Φ_{-} показаны пунктирными стрелками, а потоки Φ_{+} — сплошными. На рис. 9.3.1, б показана условная схема простейшего магнитного усилителя.

Работа магнитного усилителя основана на связи и взаимном влиянии магнитных потоков, создаваемых входным и выходным



Рис. 9.3.2.

создаваемых обмотками, намотанными на его стержни. Состояние дросселя определяется токами, проходящими через обмотки. В соответствии с рис. 9.3.2 каждый дроссель магнитного усилителя может находиться только в одном из двух состояний: насыщенном или ненасыщенном.

Современные магнитные усилители работают в режиме, при котором за период изменения напряжения u_{γ} каждый дроссель один раз переходит из ненасыщенного состояния в насыщенное и наоборот. Для двух дросселей простейшего магнитного усилителя могут быть следующие сочетания состояний: оба дросселя не насыщены; один дроссель насыщен, а другой не насыщен; оба дросселя насыщены.

Ненасыщенное состояние дросселя возможно только при равной нулю намагничивающей силе. Следовательно, ненасыщенному состоянию одного дросселя соответствует равенство

$$\omega_p i_p + \omega_y i_y = 0, \quad (9.3.1)$$

а ненасыщенному состоянию другого — равенство

$$w_p i_p - w_y i_y = 0. \quad (9.3.2)$$

Разные знаки в равенствах (9.3.1) и (9.3.2) объясняются тем, что направление управляющего тока в одном дросселе совпадает с положительным направлением тока в рабочей обмотке, а в другом — с отрицательным. В случае, когда оба дросселя не насыщены, приведенные равенства должны удовлетворяться одновременно. Это возможно только при $i_p = i_y = 0$. Таким образом, в случае, когда оба дросселя не насыщены, ток в нагрузке равен нулю. При ненасыщенном состоянии дросселей напряжение u_- компенсируется противоэлектродвижущей силой ϵ , наведенной в рабочей обмотке, т. е. $u_- + \epsilon = 0$. В соответствии с законом электромагнитной индукции

$$u_- = 2w_p \frac{d\Phi_-}{dt}. \quad (9.3.3)$$

Если $u_- = U_0 \sin \omega t$, то из (9.3.3) следует, что

$$\Phi_- = -\frac{U_0}{2w_p \omega} \cos \omega t. \quad (9.3.4)$$

Равенство (9.3.4) определяет закон изменения магнитного потока, действующего в магнитном усилителе. Переменный магнитный поток наводит э. д. с. в обеих частях обмотки управления. Однако благодаря показанной на рис. 9.3.1 намотке обмоток э. д. с., наводимые в обмотке управления, взаимно компенсируются.

Рассмотрим теперь случай, когда один дроссель насыщен, а другой не насыщен. В насыщенном дросселе $B = B_{кр} = \text{const}$ и $\Phi = \Phi_{кр} = \text{const}$. Неизменность магнитной индукции B при любом изменении напряженности поля H свидетельствует о том, что магнитная проницаемость материала сердечника в таком состоянии равна нулю. Вследствие этого и индуктивность обмоток сердечника равна нулю. Следовательно, равно нулю сопротивление обмоток переменному току. Таким образом, при насыщенном состоянии одного дросселя можно считать, что часть рабочей обмотки, приходящаяся на этот дроссель, замкнута накоротко. Ненасыщенному состоянию другого дросселя соответствует равенство (9.3.1) или (9.3.2). В этом случае при $i_y \neq 0$ рабочий ток $i_p \neq 0$ и

$$u_- = u_p + i_p R_n, \quad (9.3.5)$$

где u_p — падение напряжения на рабочей обмотке ненасыщенного дросселя. Величина u_p зависит от сопротивления цепи управляющей обмотки. Дело в том, что в случае, когда один дроссель насыщен, а другой нет, переменная э. д. с., наведенная изменяющимся магнитным потоком ненасыщенной части рабочей обмотки в обмотке управления, не будет компенсироваться, так как магнитный поток в насыщенной части рабочей обмотки постоянен и не

наводит в обмотке управления противоэлектродвижущей силы. В результате в обмотке управления будет наводиться переменный ток. Так как для ненасыщенного дросселя сумма намагничивающих сил должна быть равна нулю, то ток обмотки управления повторяет форму тока рабочей цепи. Иначе говоря, ненасыщенный дроссель работает в режиме трансформатора тока. Сопротивление

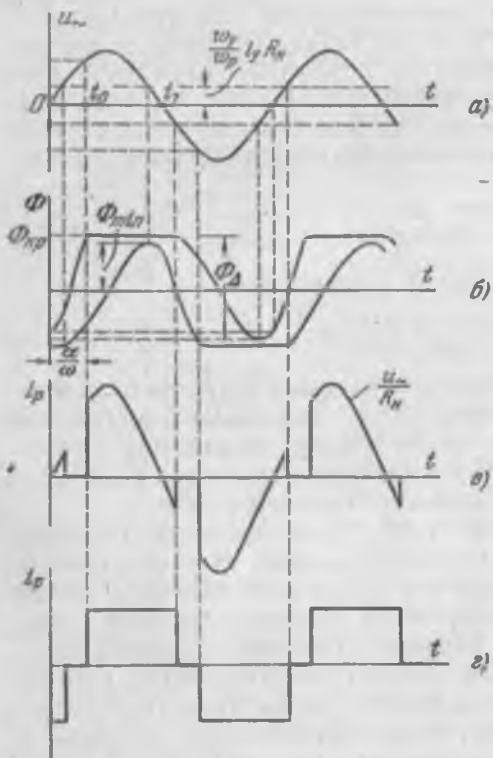


Рис. 9.3.3.

обмотки управления является нагрузкой трансформатора. Следовательно, u_p зависит от сопротивления цепи обмотки управления. При небольшом сопротивлении цепи управления величина u_p будет пренебрежимо малой по сравнению с $i_p R_n$. При этом в соответствии с равенством (9.3.5) ток в рабочей цепи определяется сопротивлением нагрузки R_n и имеет форму отрезков синусоиды. Если сопротивление управляющей цепи переменному току велико, то влияние наведенной в цепи управления э. д. с. на входной сигнал незначительно, и им можно пренебречь. В этом случае ток в рабочей обмотке не зависит от напряжения $u_{\sim}$ и в соответствии с равенством (9.3.1) или (9.3.2) целиком определяется током в цепи управления.

В состоянии, когда оба дросселя насыщены, рабочая и управляющая цепи не зависят друг от друга. Такой режим является нежелательным и практически не допускается.

Из произведенного анализа следует, что рабочий цикл магнитного усилителя состоит в переходе из состояния, когда один дроссель насыщен, в состояние, когда оба дросселя не насыщены, затем в состояние, когда другой дроссель насыщен, и т. д. Для всего рабочего цикла справедливо равенство

$$w_y i_y = \pm w_p i_p, \quad (9.3.6)$$

в котором i_y и i_p — мгновенные значения токов в управляющей и рабочей цепях.

На рис. 9.3.3 приведены диаграммы изменения напряжения, токов и потоков в цепи рабочей обмотки для случая, когда $u_y = \text{const} > 0$. На рис. 9.3.3, а показано изменение тока в нагрузке при малом сопротивлении цепи управляющей обмотки, на рис. 9.3.3, б — при большом. При уменьшении u_y время пребывания дросселей в насыщенном состоянии уменьшается, уменьшается ширина (по основанию) импульсов рабочего тока. Если изменить полярность управляющего сигнала, но не изменять его величину, то изменится только порядок насыщения дросселей, характер изменения тока i_p останется прежним, не изменится ни величина, ни фаза рабочего тока. Из рис. 9.3.3 видно, что ток в нагрузке является переменным, но не синусоидальным.

На основании изложенного можно записать

$$\omega_y |I_y| = \omega_p I_p, \quad (9.3.7)$$

где I_y — среднее значение тока в управляющей обмотке за период напряжения питания, I_p — среднее значение тока i_p рабочей обмотки. Равенство (9.3.7) является

основным соотношением магнитного усилителя. На рис. 9.3.4 показана зависимость I_p от I_y , называемая *статической характеристикой* усилителя. Сплошной линией показана характеристика, соответствующая равенству (9.3.7), полученному при допущении, что кривая намагничивания дросселей прямоугольна. Реальная характеристика магнитного усилителя показана пунктирной линией.

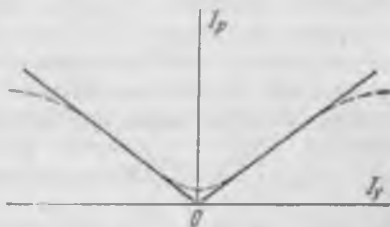


Рис. 9.3.4.

Отличие реальной характеристики от теоретической обусловлено прежде всего отличием реальной кривой намагничивания от прямоугольной. В действительности при равной нулю намагнивающей силе насыщение дросселя произойти не может. Поэтому не могут, в частности, одновременно выполняться равенства (9.3.1) и (9.3.2). Однако ввиду большой крутизны кривых намагничивания современных материалов, из которых выполняются сердечники усилителей, достаточно весьма малого тока i_y , чтобы возник постоянный управляющий поток Φ и дроссели попеременно насыщались. Допустимость пользования идеальной прямоугольной кривой намагничивания оправдывается близостью реальной и теоретической характеристик усилителя в широком диапазоне изменения входного сигнала (рис. 9.3.4).

Динамические свойства усилителя в сильной степени зависят от сопротивления цепи управляющей обмотки переменному току.

Если это сопротивление велико, то рабочая обмотка практически не оказывает влияния на входной сигнал. В этом случае статическая характеристика усилителя описывает его динамические свойства, т. е. магнитный усилитель можно считать безынерционным нелинейным звеном. Если же сопротивление цепи управляющей обмотки мало, то рабочая обмотка оказывает сильное влияние на процессы, происходящие в цепи управления. Наводимые в обмотках э. д. с. приводят к запаздыванию рабочего тока относительно входного сигнала u_y .

Выведем дифференциальное уравнение, характеризующее динамические свойства простейшего магнитного усилителя в случае, когда сопротивление цепи управляющей обмотки мало. Рассмотрим наиболее типичный случай, когда длительность переходного процесса в усилителе значительно больше периода $T = 2\pi/\omega$ питающего рабочую обмотку переменного напряжения. В этом случае переходный процесс можно характеризовать изменением средних значений тока и напряжения в рабочей обмотке и магнитных потоков.

Под средними значениями переменного тока и напряжения принято понимать средние значения их абсолютных величин за половину периода $T/2$. Знак средних значений тока в обмотке управления и управляющего напряжения за время T совпадает со знаком входного сигнала. За среднее значение магнитного потока в сердечнике одного дросселя приближенно принимается

$$\Phi_{1c} = \frac{\Phi_{\max} + \Phi_{\min}}{2} \operatorname{sgn} U_y = \left(\Phi_{\max} - \frac{\Phi_{\max} - \Phi_{\min}}{2} \right) \operatorname{sgn} U_y,$$

где $\Phi_{\max}$ и $\Phi_{\min}$ — соответственно максимальное и минимальное значения потока в течение периода. Множитель $\operatorname{sgn} U_y$ появился вследствие того, что при положительном управляющем напряжении $\Phi_{1c} > 0$, $\Phi_{2c} < 0$, а при отрицательном, наоборот, $\Phi_{1c} < 0$, $\Phi_{2c} > 0$. Из рис. 9.3.3, б видно, что

$$\Phi_{\max} - \Phi_{\min} = \Phi_{\Delta}, \quad \Phi_{\max} = \Phi_{кр}, \quad \Phi_{2c} = -\Phi_{1c}.$$

Следовательно,

$$\Phi_{1c} = \left(\Phi_{кр} - \frac{\Phi_{\Delta}}{2} \right) \operatorname{sgn} U_y, \quad \Phi_{2c} = \left(-\Phi_{кр} + \frac{\Phi_{\Delta}}{2} \right) \operatorname{sgn} U_y. \quad (9.3.8)$$

Уравнение напряжений в цепи управляющей обмотки имеет вид

$$u_y = e_y + i_y R_y. \quad (9.3.9)$$

Электродвижущая сила e_y является результирующей э. д. с., наведенной в управляющих обмотках магнитными потоками двух дросселей:

$$e_y = e_2 - e_1 = -w_y \frac{d}{dt} (\Phi_2 - \Phi_1).$$

Подставляя это выражение в уравнение (9.3.9), получим

$$u_y = -w_y \frac{d}{dt} (\Phi_2 - \Phi_1) + i_y R_y. \quad (9.3.10)$$

Проинтегрируем левую и правую части равенства (9.3.10) в пределах от 0 до $\Delta t = 2\pi/\omega = T$:

$$\int_0^{\Delta t} u_y dt = -w_y \int_{\Delta\Phi_0}^{\Delta\Phi_T} d(\Phi_2 - \Phi_1) + R_y \int_0^{\Delta t} i_y dt.$$

Поделив левую и правую части полученного равенства на Δt , получим

$$U_y = -w_y \frac{\Delta\Phi_T - \Delta\Phi_0}{\Delta t} + R_y I_y$$

или

$$|U_y| \operatorname{sgn} U_y = -w_y \frac{\Delta\Phi_T - \Delta\Phi_0}{\Delta t} + R_y |I_y| \operatorname{sgn} U_y, \quad (9.3.11)$$

где U_y и I_y — средние значения мало изменяющихся за период T напряжения и тока в управляющей обмотке, $\Delta\Phi_0$ — значение разности магнитных потоков при $t = 0$, $\Delta\Phi_T$ — значение разности потоков при $t = \Delta t$. Разность $\Delta\Phi_T - \Delta\Phi_0$ есть приращение разности потоков за время Δt , т. е. за период. Если длительность переходного процесса значительно больше времени Δt , то приращения разности потоков, получающиеся в течение периода, можно заменить приращениями разностей их средних значений, изменяющихся во время переходного процесса, т. е.

$$\Delta\Phi_T - \Delta\Phi_0 = [\Phi_{2c} - \Phi_{1c}]_n - [\Phi_{2c} - \Phi_{1c}]_{n-1}, \quad (9.3.12)$$

где $[\Phi_{2c} - \Phi_{1c}]_n$ и $[\Phi_{2c} - \Phi_{1c}]_{n-1}$ — разности средних значений магнитных потоков соответственно за n -й и $(n-1)$ -й периоды. Подставляя выражения (9.3.8) средних значений магнитных потоков в (9.3.12), получим

$$\begin{aligned} [\Phi_{2c} - \Phi_{1c}]_n - [\Phi_{2c} - \Phi_{1c}]_{n-1} = \\ = ([\Phi_\Delta]_n - [\Phi_\Delta]_{n-1}) \operatorname{sgn} U_y = \Delta\Phi_\Delta \operatorname{sgn} U_y, \end{aligned} \quad (9.3.13)$$

где $\Delta\Phi_\Delta$ — приращение разности $\Phi_\Delta = \Phi_{\max} - \Phi_{\min}$ за период. Подставив полученное из (9.3.12) и (9.3.13) выражение приращения разности потоков в (9.3.11) и заменив отношение конечных приращений $\Delta\Phi_\Delta/\Delta t$ производной $d\Phi_\Delta/dt$, после сокращения на $\operatorname{sgn} U_y$, получим

$$|U_y| = w_y \frac{d\Phi_\Delta}{dt} + R_y |I_y|. \quad (9.3.14)$$

Магнитный поток при ненасыщенном состоянии сердечника изменяется в соответствии с формулой (9.3.4). Из рис. 9.3.3, б видно,

что при небольших значениях управляющего сигнала $\Phi_{\min} = U_0 \cos \alpha / 2\omega_p \omega$, а максимальное значение потока равно $\Phi_{\text{кр}} = U_0 / 2\omega_p \omega$ и достигается в момент времени $t = t_0$. Угол $\alpha = \omega t_0$ является углом насыщения. По формуле (9.3.4) определяем

$$\Phi_{\Delta} = \Phi_{\text{кр}} - \Phi_{\min} = \frac{U_0}{2\omega_p \omega} (1 - \cos \alpha).$$

Подставляя это выражение Φ_{Δ} в (9.3.14) и выполняя дифференцирование, получим

$$|U_Y| = -\frac{\omega_Y U_0 \sin \alpha}{2\omega_p \omega} \frac{d\alpha}{dt} + R_Y |I_Y|$$

или

$$|U_Y| = -\frac{\omega_Y U_0 \sin \alpha}{2\omega_p \omega} \frac{d\alpha}{dI_p} \frac{dI_p}{dt} + R_Y |I_Y|. \quad (9.3.15)$$

Среднее значение рабочего тока I_p при малых значениях управляющего сигнала U_Y определяется формулой

$$I_p = \frac{1}{t_1} \int_{t_0}^{t_1} \frac{U_0}{R_H} \sin \omega t dt = \frac{U_0}{\pi R_H} (1 + \cos \alpha).$$

Отсюда видно, что

$$\frac{dI_p}{d\alpha} = -\frac{U_0}{\pi R_H} \sin \alpha, \quad \frac{d\alpha}{dI_p} = -\frac{\pi R_H}{U_0 \sin \alpha}. \quad (9.3.16)$$

Подставляя полученное выражение производной $d\alpha/dI_p$ в (9.3.15), учитывая (9.3.7) и полагая $\omega = 2\pi f$, получим дифференциальное уравнение, описывающее работу магнитного усилителя:

$$|U_Y| = \frac{\omega_Y R_H}{4\omega_p f} \frac{dI_p}{dt} + \frac{\omega_p}{\omega_Y} R_Y I_p$$

или

$$(TD + 1) I_p = k |U_Y|, \quad (9.3.17)$$

где

$$T = \frac{\omega_Y^2 R_H}{4f\omega_p^2 R_Y}, \quad k = \frac{\omega_Y}{\omega_p R_Y}. \quad (9.3.18)$$

Таким образом, при малом сопротивлении управляющей цепи переменному току магнитный усилитель можно считать аperiодическим звеном, постоянная времени и коэффициент усиления которого определяются формулами (9.3.18).

Магнитный усилитель обычно характеризуется коэффициентом усиления мощности:

$$k_p = \frac{R_H I_p^2}{R_Y I_Y^2} = \frac{R_H \omega_Y^2}{R_Y \omega_p^2} = R_Y R_H k^2. \quad (9.3.19)$$

Сравнивая выражения для T и k_p , замечаем, что

$$k_p = 4fT. \quad (9.3.20)$$

Для повышения коэффициента усиления в магнитном усилителе часто используется положительная обратная связь. На рис. 9.3.5 приведена схема магнитного усилителя с внешней положительной обратной связью. Положительная обратная связь осуществляется путем введения дополнительных обмоток обратной связи w_{oc} , которые включены в схему так же, как и соответствующие части управляющей обмотки. Выпрямитель B служит для преобразования выходного переменного тока в постоянный. Постоянный ток, пропорциональный среднему значению тока нагрузки, протекает по обмоткам обратной связи. При положительной полярности входного сигнала u_y магнитные потоки обмоток управления и обмоток обратной связи складываются. Это приводит к увеличению коэффициента усиления усилителя. Изменение полярности входного сигнала вызывает изменение направления потоков в управляющих обмотках, но направления потоков в обмотках обратной связи остаются прежними. Магнитные потоки обмоток управления и обмоток обратной связи будут вычитаться. Это приводит к уменьшению коэффициента усиления. На рис. 9.3.6

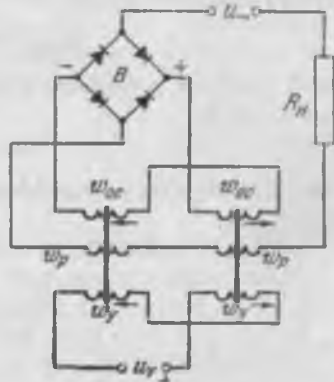


Рис. 9.3.5.



Рис. 9.3.6.

показана статическая характеристика магнитного усилителя с положительной обратной связью. Как видно, магнитный усилитель с обратной связью неодинаково реагирует на входные сигналы разной полярности. В реальных автоматических системах входной сигнал может колебаться около нулевого значения. Для обеспечения одинакового коэффициента усиления при изменении полярности входного сигнала характеристику усилителя смещают относительно начального положения вдоль оси абсцисс. Смещение осуществляется с помощью специальных обмоток, называемых обмотками смещения, которые питаются постоянным током, полярность которого соответствует большей крутизне характери-

менения полярности входного сигнала характеристику усилителя смещают относительно начального положения вдоль оси абсцисс. Смещение осуществляется с помощью специальных обмоток, называемых обмотками смещения, которые питаются постоянным током, полярность которого соответствует большей крутизне характери-

стикп. Следует, однако, заметить, что введение положительной обратной связи приводит к некоторому смещению минимума статической характеристики в сторону отрицательных токов и без специальных обмоток смещения. Это объясняется тем, что в простейших усилителях при нулевом входном сигнале рабочий ток не равен нулю, следовательно, не равен нулю ток в обмотке обратной связи, не равно нулю подмагничивающее поле и ток в нагрузке не является минимальным.

При наличии положительной обратной связи вместо равенства (9.3.7) следует записать

$$\omega_p I_p = | \omega_y I_y + \omega_{oc} I_{oc} |,$$

где I_{oc} — ток в обмотках обратной связи. Обычно $I_{oc} \approx I_p$. Тогда при $I_y > 0$

$$\omega_p I_p (1 - \beta) = \omega_y I_y, \quad (9.3.21)$$

где $\beta = \omega_{oc}/\omega_p$ — коэффициент обратной связи (обычно $0 < \beta < 1$).

Дифференциальное уравнение, описывающее работу магнитного усилителя с обратной связью, получится, если в уравнение (9.3.15) подставить производную da/dI_p из (9.3.16) и выражение I_y из (9.3.21). При этом постоянная времени магнитного усилителя с обратной связью определится формулой

$$T = \frac{R_n \omega_y^3}{4f R_y \omega_p^2 (1 - \beta)}. \quad (9.3.22)$$

Подставляя в (9.3.19) отношение I_p/I_y из (9.3.21), получим формулу для коэффициента усиления магнитного усилителя с обратной связью:

$$k_p = \frac{R_n \omega_y^3}{R_y \omega_p^2 (1 - \beta)^2} \quad (9.3.23)$$

Сравнивая формулы (9.3.22) и (9.3.23), находим

$$T = \frac{k_p}{4f} (1 - \beta) \quad (9.3.24)$$

Из формул (9.3.22) и (9.3.23) следует, что положительная обратная связь увеличивает коэффициент усиления и одновременно приводит к увеличению инерционности усилителя. Однако при одинаковых коэффициентах усиления постоянная времени усилителя с положительной обратной связью меньше постоянной времени усилителя без обратной связи.

Из (9.3.18) и (9.3.22) следует, что для уменьшения инерционности магнитного усилителя необходимо повышать частоту f источника переменного напряжения $u_{\sim}$. Инерционность магнитного усилителя снижается также, если при заданном коэффициенте усиления

по мощности использовать несколько каскадов усиления. Пусть, например, двухкаскадный магнитный усилитель имеет по первому каскаду коэффициент усиления k_1 , а по второму k_2 при общем коэффициенте усиления по мощности $k_p = k_1 k_2$. Постоянную времени двухкаскадного усилителя можно приближенно определить как сумму

$$T_1 + T_2 = \frac{1-\beta}{4f} (k_1 + k_2), \quad (9.3.25)$$

где T_1 и T_2 — постоянные времени первого и второго каскадов соответственно, определяемые по формуле (9.3.24). Постоянная времени однокаскадного усилителя T , имеющего коэффициент усиления k_p , также определяется равенством (9.3.24). Разделив почленно (9.3.24) на (9.3.25), получим

$$\frac{T}{T_1 + T_2} = \frac{k_p}{k_1 + k_2}.$$

Обычно $k_1 \gg 1$ и $k_2 \gg 1$. При этом $k_p = k_1 k_2 > k_1 + k_2$ и, следовательно, $T > T_1 + T_2$.

Простейшие магнитные усилители без обратной связи и с обратной связью обладают существенными недостатками: они нечувствительны к полярности входного сигнала и не обеспечивают нулевого тока в нагрузке при нулевом сигнале на входе усилителя. Эти недостатки устраняются применением мостовой и дифференциальной схем включения простейших магнитных усилителей. Включение двух простейших магнитных усилителей по дифференциальной схеме образует двухтактный дифференциальный магнитный усилитель. Два простейших магнитных усилителя, включенных по мостовой схеме, образуют двухтактный мостовой магнитный усилитель.

На рис. 9.3.7 представлена принципиальная схема дифференциального магнитного усилителя. Каждый из простейших усилителей имеет статическую характеристику, смещенную относительно оси ординат приблизительно до середины рабочего участка (рис. 9.3.8). Подобное смещение достигается, как уже отмечалось, включением специальных обмоток смещения. Обмотки управления включены так, что в одном усилителе намагничивающие поля, создаваемые обмотками управления и смещения, совпадают по направлению, а в другом усилителе имеют противоположные направления. Благодаря этому ток управления увеличивает выходной сигнал в одном усилителе и уменьшает в другом. Результирующий ток в нагрузке равен разности $I_{II} = I_1 - I_2$. Сопровождение $R_{сж}$ служит для балансировки, в результате которой обеспечивается прохождение через нуль характеристики дифференциального магнитного усилителя, показанной на рис. 9.3.8. Практически изменение знака среднего значения тока нагрузки

дифференциального усилителя при изменении полярности тока управления означает изменение на 180° фазы выходного сигнала.

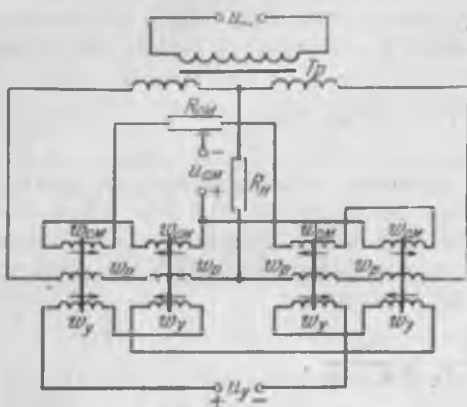


Рис. 9.3.7.

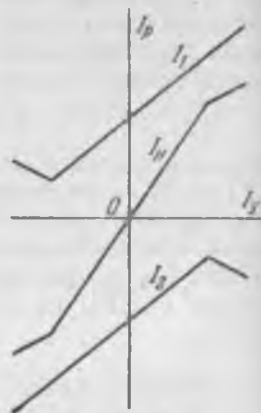


Рис. 9.3.8.

На рис. 9.3.9 представлена схема мостового магнитного усилителя. Каждое плечо моста образовано одной из обмоток переменного тока, причем в противоположных плечах находятся обмотки,

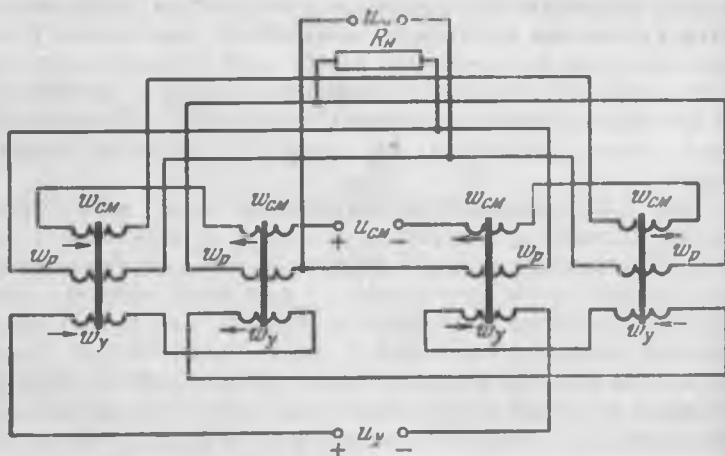


Рис. 9.3.9.

расположенные на одном сердечнике. При отсутствии входного сигнала мост уравновешен и выходное напряжение на нагрузке R_H равно нулю. При подаче входного сигнала равновесие моста нарушается, так как в одной паре противоположных плеч ток

уменьшается, а в другой — увеличивается. На нагрузке появляется напряжение, фаза которого будет определяться полярностью входного сигнала u_y . Так же как и в схеме рис. 9.3.7, в мостовой схеме в одном из сердечников магнитные поля обмотки смещения и обмотки управления действуют в противоположных направлениях, а во втором — в одном направлении.

Если от магнитного усилителя нужно получить выходное напряжение, превышающее напряжение источника питания или

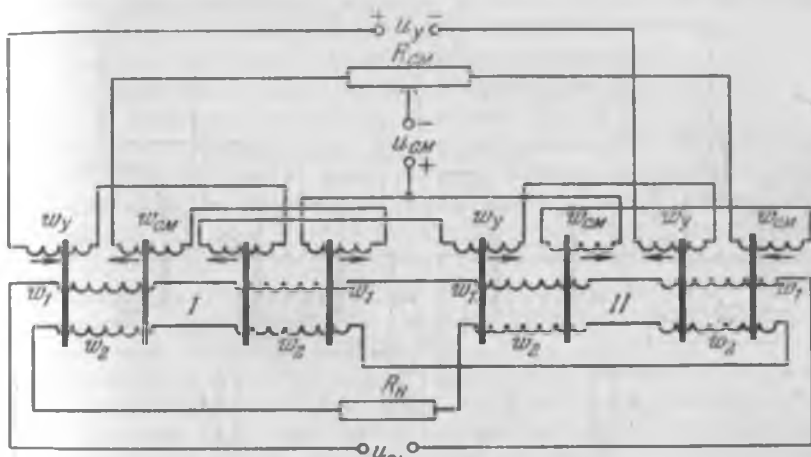


Рис. 9.3.10.

значительно меньше его, то применяется трансформаторная схема магнитного усилителя. На рис. 9.3.10 представлена схема трансформаторного магнитного усилителя. Первичные обмотки (четыре обмотки с числом витков w_1 каждая) соединены последовательно и создают магнитные потоки одного направления. Вторичные обмотки, в цепь которых включена нагрузка R_n , соединены тоже последовательно, но так, что потоки обмоток дросселей I и II взаимно противоположны. При входном сигнале, равном нулю, потоки в сердечниках I и II одинаковы. Во вторичных обмотках сердечников I и II падают равные э.д.с., которые взаимно компенсируются. Тока в цепи нагрузки не будет. Если входной сигнал u_y отличен от нуля, то в сердечниках I и II создаются разные подмагничивающие поля, так как в сердечнике I поток обмотки смещения совпадает по направлению с потоком обмотки управления, а в сердечнике II эти потоки направлены друг против друга. Магнитные потоки в сердечниках I и II не будут одинаковыми. В цепи вторичных обмоток наведется результирующая э.д.с. В нагрузке появится ток, величина и фаза которого

будут зависеть от величины входного сигнала u_y и его полярности соответственно. Величина выходного напряжения определяется соотношением чисел витков w_1 и w_2 .

Если элементы автоматической системы, включенные после магнитного усилителя, работают на постоянном токе, то магнитные усилители делаются с выходом на постоянном токе. На рис. 9.3.11 показан магнитный усилитель с выходом на постоянном токе. Он

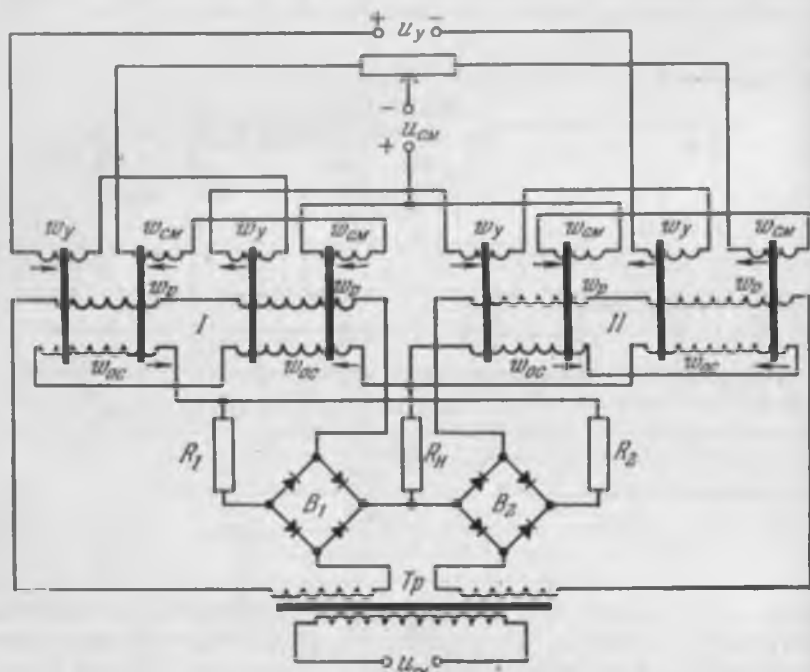


Рис. 9.3.11.

имеет положительную обратную связь и дополнительное подмагничивание, осуществляемое при помощи обмоток смещения. Когда входной сигнал u_y равен нулю, в обмотках w_p проходит переменный ток одинаковой амплитуды, вследствие чего по нагрузке R_n протекают равные, но противоположно направленные постоянные токи i_1 и i_2 . При $u_y \neq 0$ один из сердечников подмагничивается сильнее, а другой в меньшей степени, так как в одном из них поле смещения и поле управления действуют в одном направлении, а в другом — в противоположных направлениях. Это приводит к тому, что амплитуды переменных токов в обмотках сердечников I и II, а также токи i_1 и i_2 в нагрузке R_n не будут одинаковыми. Изменение полярности входного сигнала вызовет изменение направления постоянного тока нагрузки.

Схема дифференциального усилителя, приведенная на рис. 9.3.7, применяется тогда, когда последующий элемент, например асинхронный двигатель, имеет лишь одну управляющую обмотку переменного тока. Мостовая схема усилителя имеет то же значение. Коэффициент полезного действия мостового магнитного усилителя, вследствие меньших потерь в обмотках, несколько больше коэффициента полезного действия дифференциального усилителя. Преимуществом трансформаторного усилителя является простота согласования с нагрузкой.

Часто исполнительный элемент автоматической системы имеет две дифференциальные обмотки управления. В этом случае обмотки в качестве нагрузки включаются непосредственно в цепи простейших магнитных усилителей, составляющих двухтактный усилитель, а питание усилителей остается общим.

Все рассмотренные схемы усилителей имеют одну управляющую обмотку и рассчитаны на один входной сигнал. Магнитный усилитель может иметь несколько управляющих обмоток, являющихся его входами. В этом случае выходной сигнал будет результатом суммарного действия всех входных сигналов усилителя. Выбором чисел витков управляющих обмоток обеспечивается необходимое соотношение коэффициентов усиления слагаемых, входящих в суммарный выходной сигнал.

Магнитные усилители нашли широкое применение в автоматических системах управления. Они надежны в эксплуатации, обладают большим сроком службы, нечувствительны к вибрациям, вследствие того что не имеют подвижных частей. Они допускают большие перегрузки, не нуждаются в предварительном разогреве, как, например, электронные усилители, и обладают более высоким коэффициентом полезного действия, чем электронные усилители.

Двухтактные магнитные усилители обладают большой зоной линейности, которая практически позволяет считать их линейными элементами. Коэффициент усиления магнитных усилителей достигает значения $10^3 \div 10^6$ на один каскад. Это позволяет усиливать слабые сигналы постоянного тока мощностью $10^{-10} \div 10^{-17}$ В·А [19].

Основным недостатком магнитного усилителя является его инерционность. Постоянная времени магнитного усилителя с большим коэффициентом усиления может исчисляться секундами.

§ 9.4. Нелинейные функциональные преобразователи на диодах

Различные нелинейные операции над сигналами можно осуществить электрическими цепями с диодами. Основной характеристикой диода является его *вольтамперная характеристика* $i = \varphi(u)$, показывающая зависимость тока i , протекающего через диод, от приложенного к нему напряжения u . На рис. 9.4.1

изображены вольтамперные характеристики для вакуумного (рис. 9.4.1, а) и полупроводникового (рис. 9.4.1, б) диодов. Точка O на характеристиках является точкой покоя, т. е. точкой, характеризующей состояние системы при отсутствии входного сигнала ($u = 0$). При положительном напряжении на диоде его сопротивление

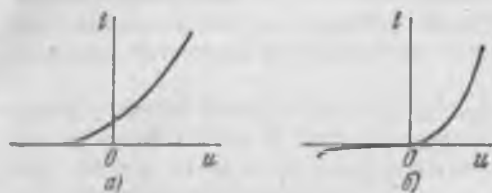


Рис. 9.4.1.

невелико и поэтому через диод протекает ток i , зависящий от величины напряжения u . При отрицательном значении напряжения u сопротивление диода очень велико и через диод ток не протекает (в случае вакуумного диода) или он очень незначите-

лен (в случае полупроводникового диода). Зависимость сопротивления диода от знака приложенного напряжения и используется для создания различных нелинейных устройств и схем. Нелинейная вольтамперная характеристика диода обычно аппрок-

симируется экспоненциальной или кусочно-линейной функцией. На рис. 9.4.2 приведены кусочно-линейные характеристики для вакуумного (рис. 9.4.2, а) и полупроводникового (рис. 9.4.2, б) диодов. Диод можно считать безынерционным нелинейным элементом, если спектр ча-

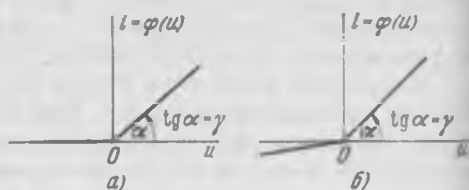


Рис. 9.4.2.

стот входного сигнала (напряжение на диоде u) не содержит частот выше 100 кГц . При больших частотах следует принимать во внимание межэлектродную емкость диода. Нелинейная характеристика диода позволяет реализовать большое количество различных нелинейностей путем соединения нескольких диодов в простые схемы. Рассмотрим две схемы.

На рис. 9.4.3 изображена схема на диодах, дающая на выходе модуль входного напряжения. Напряжение $u_{вх}$ приложено к диоду D_1 и, после изменения знака на инверторе I (усилитель с коэффициентом усиления -1), к диоду D_2 . При $u_{вх} > 0$ диод D_1 открыт и через сопротивление нагрузки R течет ток i_1 . При $u_{вх} < 0$ диод D_1 закрыт, а диод D_2 открыт. Через сопротивление нагрузки R течет тот же ток i_1 в том же направлении, что и при $u_{вх} > 0$. Считая за выходной сигнал напряжение $u_{вых}$ на сопротивлении R , получим следующую зависимость выходного напряжения от входного:

$$u_{вых} = k |u_{вх}|, \quad (9.4.1)$$

где $k = R\gamma / (1 + R\gamma)$ — коэффициент передачи, а γ — крутизна вольтамперной характеристики диода. При $R\gamma \geq 50$, что обычно имеет место, коэффициент передачи $k \approx 1$. Характеристика (9.4.1) представлена на рис. 9.4.4. Нелинейность такого типа называется *модулем* или *двухполупериодным линейным детектором*. Если

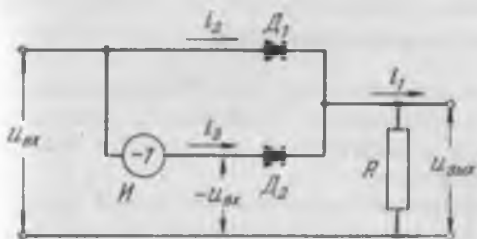


Рис. 9.4.3.

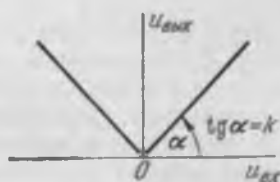


Рис. 9.4.4.

поменять входы и выходы диодов местами, т. е. перевернуть диоды, то схема будет давать на выходе модуль входного напряжения с отрицательным знаком:

$$u_{\text{вых}} = -k |u_{\text{вх}}|. \quad (9.4.2)$$

В системах управления иногда приходится решать задачу выбора наибольшей (наименьшей) из нескольких величин. На рис. 9.4.5 изображена схема на диодах, осуществляющая нелинейную операцию выделения наибольшей величины. Входные напряжения $u_k (k=1, 2, \dots, n)$ приложены к диодам D_k . Если все входные напряжения u_k положительны, то открытым будет только тот диод, анод которого имеет более высокий потенциал. Поэтому выходное напряжение $u_{\text{вых}}$ будет равно наибольшему из входных напряжений $u_1, u_2, \dots, u_n$. Чтобы схема работала и при отрицательных значениях входных напряжений, потенциал точки a (рис. 9.4.5) понижается за счет напряжения смещения $-u_c$, которое выбирается из условия

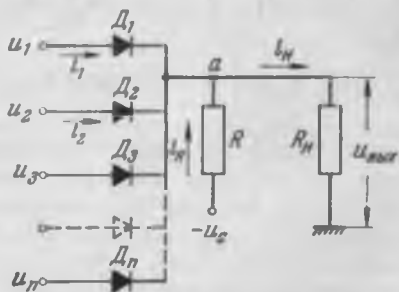


Рис. 9.4.5.

$$\left| u_c \frac{R_n}{R + R_n} \right| > |u_{k_{\text{max}}}| \quad (k=1, 2, \dots, n). \quad (9.4.3)$$

Зависимость выходного напряжения от входных в данной схеме

выражается формулой

$$u_{\text{вых}} = \max \{u_1, \dots, u_n\} \quad (9.4.4)$$

Схема выделения наименьшей величины получается из предыдущей перевертыванием диодов и изменением знака напряжения смещения.

§ 9.5. Вращающиеся трансформаторы

Задачи, связанные с геометрическими преобразованиями (решение треугольников, преобразование координат, разложение и построение векторов и пр.), эффективно решаются с помощью вращающихся трансформаторов.

Вращающийся трансформатор представляет собой электрическую машину с неявно выраженными полюсами ротора и статора, на которых размещаются по две обмотки со смещенными одна относительно другой на 90° электрическими осями.

Для разложения вектора на составляющие в прямоугольной системе координат, преобразования прямоугольных координат вектора в полярные и наоборот служит *синусно-косинусный* вращающийся трансформатор.

Для разложения вектора на составляющие необходимо подать на одну из статорных обмоток переменное напряжение u (рис. 9.5.1).

Переменный ток будет создавать магнитный поток, направленный вдоль оси обмотки возбуждения и пропорциональный амплитуде переменного напряжения. Вторая статорная обмотка замыкается при этом накоротко для уничтожения поперечной составляющей магнитного потока, которая может появиться за счет реакции роторных обмоток, работающих на нагрузку, или за счет электрической и магнитной асимметрии конструкции вращающегося трансформатора. Нулевым положением ротора считается положение, при котором ось одной из его обмоток совпадает по направлению с осью обмотки возбуждения. При повороте ротора на угол α вектор магнитного потока индуцирует в обмотках ротора э. д. с., величины которых пропорциональны проекциям вектора магнитного потока на оси обмоток ротора:

$$u_1 = ku \sin \alpha, \quad u_2 = ku \cos \alpha. \quad (9.5.1)$$

На плоскости, вращающейся с угловой скоростью ω , соответст-

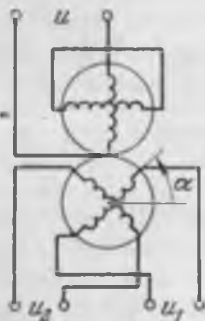


Рис. 9.5.1.

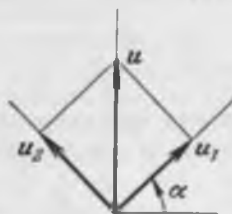


Рис. 9.5.2.

вующей частоте переменного напряжения u , величины u'_1 , u'_2 , u представляются векторами. Очевидно, что u'_1 , u'_2 пропорциональны составляющим вектора u (рис. 9.5.2).

Для решения задачи поворота координатных осей на обе статорные обмотки подаются переменные напряжения, совпадающие друг с другом по фазе и имеющие амплитуды, пропорциональные составляющим вектора в исходной системе координат (рис. 9.5.3). Продольная и поперечная составляющие магнитного потока

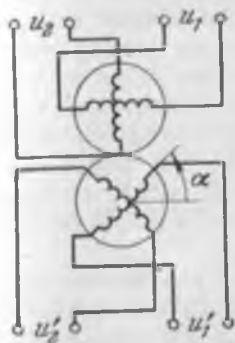


Рис. 9.5.3.

статорной обмотки, суммируясь, образуют результирующий магнитный поток, величина которого пропорциональна модулю вектора (рис. 9.5.4). Вектор магнитного поля наводит в

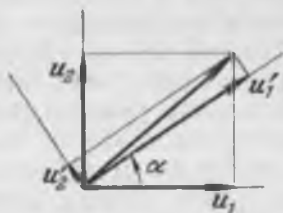


Рис. 9.5.4.

обмотках ротора, повернутого относительно нулевого положения на угол α , напряжения, пропорциональные проекциям вектора на оси обмоток ротора:

$$u'_1 = ku_1 \cos \alpha + ku_2 \sin \alpha, \quad u'_2 = -ku_1 \sin \alpha + ku_2 \cos \alpha. \quad (9.5.2)$$

В данном случае синусно-косинусный трансформатор можно рассматривать как электромеханическую модель, в которой оси исходной координатной системы совпадают с осями статорных обмоток, а оси новой системы координат совпадают с осями роторных обмоток. Уравнения (9.5.2) являются общими для синусно-косинусного вращающегося трансформатора.

Наиболее важным показателем качества работы вращающегося трансформатора является его точность. Если вращающийся трансформатор питается от сети, а обмотки ротора не нагружены (режим холостого хода), то погрешность его работы определяется эксцентриситетом воздушного зазора, влиянием зубьев на статоре и роторе и другими причинами технологического и конструктивного характера. При выполнении вращающихся трансформаторов по первому классу точности ошибки от этих причин не превышают 0,1%. При работе вращающихся трансформаторов в каскадных системах, т. е. при последовательном соединении их, погрешности работы значительно больше. В схеме последовательного соеди-

нения нескольких вращающихся трансформаторов питание статорных обмоток последующих трансформаторов осуществляется от роторных обмоток предыдущих. Поэтому амплитуда напряжения питания непостоянна и зависит от положения ротора. Вследствие нелинейности характеристики намагничивания магнитопровода вращающегося трансформатора изменяется его входное сопротивление, что приводит к появлению ошибки. Изменение входных и выходных сопротивлений вращающегося трансформатора часто приводит к нарушению симметрии сопротивлений в обмотках. Чем выше коэффициент трансформации, тем больше сказывается ошибка из-за асимметрии и тем труднее она устраняется. Для уменьшения перечисленных ошибок применяют развязку каскадов вращающихся трансформаторов с помощью усилителей и используют пермаллой в качестве материала для магнитопровода [47]. Рассматривая вращающийся трансформатор как элемент автоматических систем, можно представить его как звено, имеющее три входа и два выхода. Связь между входными и выходными сигналами описывается уравнениями (9.5.2).

Динамические свойства вращающихся трансформаторов аналогичны свойствам индуктивных датчиков. Вращающийся трансформатор можно считать безынерционным элементом, так как величина α обычно изменяется достаточно медленно, чтобы ее можно было считать практически постоянной в пределах нескольких периодов изменения входных напряжений.

§ 9.6. Гальваномагнитные элементы

Среди новых физических принципов, используемых для реализации функциональных преобразователей, заслуживает внимания эффект Холла, который состоит в следующем. Рассмотрим полупроводник в форме прямоугольного параллелепипеда с размерами

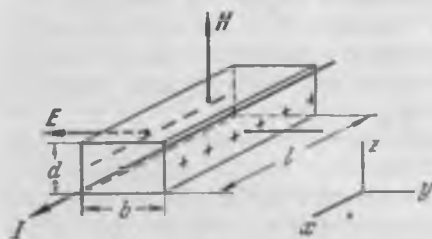


Рис. 9.6.1.

l, b, d (рис. 9.6.1), по которому течет ток I в направлении оси x . При наложении магнитного поля с напряженностью H вдоль оси z на каждый электрон в полупроводнике будет действовать сила Лоренца

$$F = e (v \times H), \quad (9.6.1)$$

где e — заряд электрона, v — скорость электрона. Под действием этой силы электроны будут отклоняться вдоль оси y . Поскольку размеры полупроводника ограничены, то поверхность, к которой отклоняются электроны, получает отрицательный заряд, а противоположная ей поверхность получает положительный заряд.

Эти объемные заряды и создают в полупроводнике поперечное электрическое поле E , называемое обычно *полем Холла*. Очевидно, что напряженность электрического поля Холла можно определить из условия равенства лоренцевой силы и силы отталкивания, возникающей за счет напряженности поля Холла:

$$eE = e(v \times H), \quad (9.6.2)$$

откуда

$$E = v \times H. \quad (9.6.3)$$

Скорость движения электронов можно выразить через ток:

$$v = R_H \frac{I}{bd} \cdot 10^{-8}, \quad (9.6.4)$$

где $R_H = 3\pi\mu/8\sigma$ — постоянная Холла, μ — подвижность электронов, а σ — электропроводность пластинки полупроводника. Подставляя выражение (9.6.4) в формулу (9.6.3) и проектируя полученное равенство на ось y , получим

$$E_y = -R_H \frac{IH}{bd} \cdot 10^{-8}. \quad (9.6.5)$$

Выражая ток I через напряжение u_1 , приложенное к пластинке полупроводника: $I = obdu_1/l$, и учитывая, что разность потенциалов, возникающая на боковых гранях полупроводника вдоль оси y (э. д. с. Холла), равна произведению напряженности поля E_y на b , получим

$$u_y = k_1 u_1 H, \quad (9.6.6)$$

где $k_1 = 3\pi\mu b \cdot 10^{-8}/8l$, величины u_y , u_1 выражаются в вольтах, H — в эрстедах, b и l — в сантиметрах, μ — в $\text{см}^2/\text{В} \cdot \text{с}$. Таким образом, напряжение на боковых гранях пластины полупроводника пропорционально произведению напряжения u_1 и напряженности магнитного поля H .

Эффект Холла может быть использован для создания различных элементов автоматических систем, осуществляющих бесконтактное измерение или функциональное преобразование переменных величин. С помощью этих элементов можно измерять напряженность переменных и постоянных магнитных полей, силу тока в электрической цепи, осуществлять модуляцию или детектирование сигналов, производить гармонический анализ сигнала, усиливать и генерировать сигналы различной формы.

Для измерения напряженности переменных и постоянных магнитных полей достаточно поместить пластину полупроводника в исследуемое магнитное поле. В соответствии с формулой (9.6.6) э. д. с. Холла при фиксированном напряжении u_1 будет пропорциональна напряженности магнитного поля. Небольшие размеры датчика э. д. с. Холла (изготавливаются пленочные датчики размером $0,01 \div 0,02$ мм) позволяют точно измерять топографию маг-

нитного поля в малых объемах и зазорах. Измерение напряженности магнитного поля можно производить в очень широком диапазоне за счет изменения коэффициента усиления датчика с помощью постоянного напряжения u_1 . Малая инерционность датчика по отношению к напряженности магнитного поля позволяет измерять переменные магнитные поля с частотами до 10^{11} Гц.

Для измерения больших токов, в десятки тысяч ампер, используется магнитное поле измеряемого тока. При постоянном напряжении u_1 напряженность магнитного поля, создаваемого током,

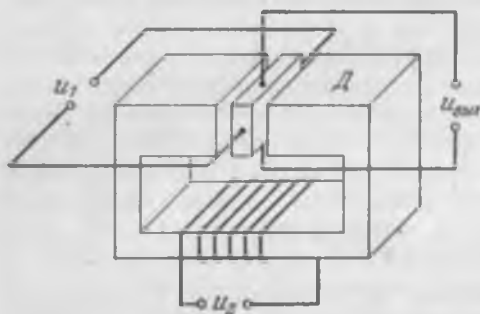


Рис. 9.6.2.

пропорциональна току в проводнике. Бесконтактный способ имеет большие преимущества перед способом непосредственного измерения тока, поскольку изготовление шунтов для токов в десятки тысяч ампер является трудной задачей.

На рис. 9.6.2 представлена схема множительного устройства. Напряжение u_1 , являющееся входным, приложено к пластинке

полупроводника. Магнитное поле создается дросселем D , в зазоре которого помещена пластина полупроводника. Напряженность поля можно изменять, увеличивая или уменьшая ток в катушке дросселя, поскольку $H = k_2 i$, где k_2 — коэффициент пропорциональности. Ток в катушке связан со вторым входным напряжением u_2 уравнением

$$T \frac{di}{dt} + i = \frac{1}{R} u_2, \quad T = \frac{L}{R}, \quad (9.6.7)$$

где L — индуктивность катушки дросселя, а R — активное сопротивление катушки. Постоянная времени T имеет порядок $0,01 \div 0,05$ с. Выходное напряжение (з. д. с. Холла) снимается с боковых граней пластины полупроводника. Если u_2 — медленно изменяющаяся функция времени, то можно не учитывать инерционность связи между входным напряжением и током в катушке дросселя, т. е. положить $T = 0$. Инерционность эффекта Холла по отношению к напряжению u_1 и напряженности H магнитного поля очень мала (постоянная времени имеет порядок $10^{-11} \div 10^{-13}$ с.). Поэтому ее можно не учитывать. При сделанных допущениях связь выходного напряжения с входными выражается соотношением

$$u_y = k u_1 u_2. \quad (9.6.8)$$

Электрическая схема включения полупроводниковой пластинки датчика э. д. с. Холла показана на рис. 9.6.3. Входами являются напряжения u_1 и u_2 . Выходное напряжение снимается с боковых граней пластинки датчика. Таким образом, схема, приведенная на рис. 9.6.3, может быть использована как множительное устройство. Сомножители подаются на входы элемента в виде напряжений u_1 и u_2 . Точность вычисления в данной схеме имеет порядок $0,1 \div 0,3\%$.

Для модуляции переменного напряжения необходимо несущую подавать в виде напряжения u_1 , а модулирующий сигнал в виде напряжения u_2 . В соответствии с формулой (9.6.8) на выходе элемента получается модулированный сигнал.

Для построения квадратичного детектора нужно включить катушку дросселя последовательно с пластиной полупроводника и подать на вход полученной цепи напряжение переменного тока, модулированного сигналом. При этом в силу формулы (9.6.8) выходное напряжение будет пропорционально квадрату тока, протекающего через катушку дросселя и одновременно через пластину полупроводника: $u_y = cki^2$. Учитывая, что ток является переменным, имеем

$$u_y = \frac{1}{2} cka^2 (1 + \cos 2\omega t), \quad (9.6.9)$$

где a — амплитуда тока, c — коэффициент пропорциональности. Составляющая двойной частоты может быть отделена фильтром. В результате получим выходное напряжение, пропорциональное квадрату амплитуды тока.

В настоящее время полупроводниковые пластины изготавливаются из Ge, HgSe, HgTe, InSb и других материалов, характеризующихся высокой электронной проводимостью. Для полупроводниковой пластины из селенида ртути HgSe основные параметры в формулах (9.6.6) и (9.6.8) имеют следующий порядок: $b/l = 0,4$, $k = 0,03 \div 0,05$, $\mu = 10\,000$, $u_{y\max} = 0,05$ В.

Основными преимуществами гальваномагнитных элементов, основанных на эффекте Холла, являются высокая механическая прочность, малая инерционность и высокая точность [23].

§ 9.7. Реле

Реле представляет собой устройство, предназначенное для замыкания и размыкания электрических цепей. При этом часто входной сигнал, управляющий работой реле, имеет малую мощность, а замыкаемая реле электрическая цепь питается источни-

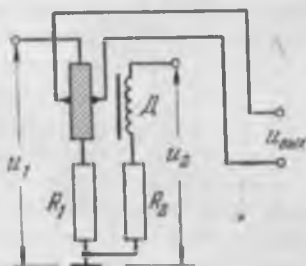


Рис. 9.6.3.

ком значительно большей мощности. В таких случаях реле играет роль усилителя мощности.

На рис. 9.7.1 представлена схема простейшего электромагнитного реле, состоящего из якоря 1 с пружиной 4, сердечника 2, катушки 3, ярма 5 и контактов 6. Входной величиной является напряжение, подаваемое на катушку 3, а выходной — напряжение в цепи, замыкаемой контактами 6. При приложении к катушке 3 входного напряжения через нее протекает ток, создающий в сердечнике 2, ярме 5, якоре 1 и воздушном зазоре между якорем и сердечником магнитный поток. При этом возникает сила, притягивающая якорь к сердечнику. Якорь

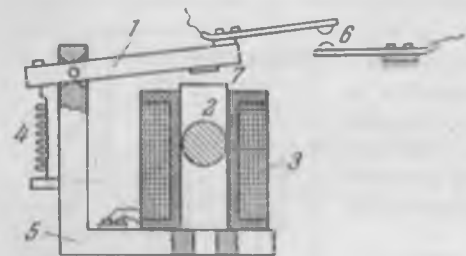


Рис. 9.7.1.

преодолевая усилие возвратной пружины 4, и замыкает контакты 6. При полном замыкании контактов дальнейшее движение якоря ограничивается упором 7.

В зависимости от назначения реле контакты реле могут быть разомкнуты при отсутствии входного напряжения или замкнуты.

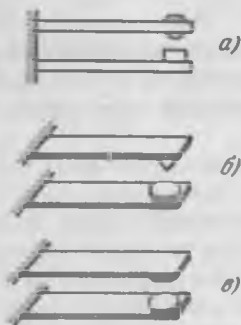


Рис. 9.7.2.

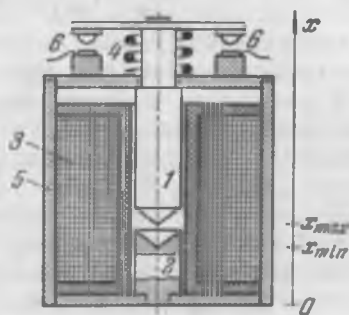


Рис. 9.7.3.

Контакты первого типа обычно называются *нормально разомкнутыми*, а контакты второго типа — *нормально замкнутыми*. По конструктивному оформлению контакты реле (рис. 9.7.2) бывают точечные (а), линейные (б) и плоскостные (в). При замкнутых контактах поверхности контактных тел прижимаются друг к другу с некоторым усилием F_A , называемым *контактным давлением*:

Для высокочувствительных реле применяются контактные давления порядка $1 \div 2$ гс, для реле с меньшей чувствительностью $5 \div 10$ гс и для обычных малочувствительных реле (типа телефонных) $20 \div 50$ гс.

Кроме рассмотренного поворотного электромагнитного реле, называемого так вследствие поворотной конструкции якоря, широко распространены электромагнитные реле втяжного (плунжерного) типа (рис. 9.7.3) [67]. Реле втяжного типа имеет те же составные части, что и поворотное реле. Поэтому номера частей реле на рис. 9.7.3 соответствуют номерам частей реле на рис. 9.7.1.

Реле можно создать не только с помощью электромагнита, но и на электронных лампах, кристаллических триодах, магнитных усилителях и других устройствах. В соответствии с этим различают *электромеханические, электронные, ионные и магнитные* реле. К электромеханическим относятся электромагнитные, магнитозлектрические и электродинамические реле. Электронные реле можно построить на вакуумных лампах и на кристаллических триодах. Ионные реле реализуются на тиратронах.

Рассмотрим подробнее физические процессы, протекающие в электромагнитном реле на примере реле втяжного типа (рис. 9.7.3). При подаче напряжения $u_{вх}$ на управляющую обмотку реле в ней возникает ток i , в результате чего в сердечнике создается магнитный поток и на якорь начинает действовать сила притяжения. В момент, когда эта сила станет равной силе пружины, якорь начинает двигаться и через некоторое время замыкает контакты. Дойдя до упора, якорь останавливается и остается в этом положении до тех пор, пока напряжение $u_{вх}$ на управляющей обмотке не будет снято. После того как напряжение $u_{вх}$ станет равным нулю, ток в обмотке реле начинает уменьшаться, вследствие чего начинает уменьшаться и сила притяжения якоря к сердечнику. В момент, когда сила притяжения сравняется с действующей в противоположном направлении силой пружины, якорь начнет движение и через некоторое время разомкнет контакты. После этого якорь будет продолжать движение до упора.

Время с момента подачи напряжения на обмотку реле до момента начала движения якоря $T_{тр}$ называется *временем трогания*. Ток в катушке $i_{ср}$, при котором якорь начинает движение, называется *током срабатывания*. Время от момента подачи напряжения $u_{вх}$ до момента замыкания контактов $T_{ср}$ называется *временем срабатывания* реле. Время $T_{тpo}$ с момента снятия напряжения $u_{вх}$ до начала обратного движения якоря называется *временем трогания при отпуске*. Время $T_{от}$ с момента снятия напряжения $u_{вх}$ до момента размыкания контактов реле называется *временем отпущения* реле. Ток в катушке $i_{от}$, при котором начинается движение якоря под действием пружины, называется *током отпущения*. Величины $T_{тр}$, $T_{ср}$, $T_{тpo}$, $T_{от}$, а также время

$T_{\text{дв}}$ движения якоря до упора при включении и время $T_{\text{дв.о}}$ движения якоря до упора при отпускании являются основными *временными характеристиками* реле.

Для вычисления временных характеристик реле необходимо составить уравнение движения якоря и уравнение изменения тока в катушке и интегрировать их совместно. Для составления уравнения движения якоря необходимо сначала найти силу притяжения якоря к сердечнику. Для этого напомним уравнение, определяющее ток в катушке:

$$u_{\text{вх}} = Ri + L \frac{di}{dt} + i \frac{dL}{dt}, \quad (9.7.1)$$

где R и L — омическое сопротивление и индуктивность катушки соответственно. Изменение индуктивности L с течением времени вызывается изменением величины воздушного зазора x при движении якоря реле, как это следует из формулы

$$L = \frac{0,4\pi w^2 \cdot 10^{-8}}{R_{\mu} + \frac{x}{S}}, \quad (9.7.2)$$

где R_{μ} — сумма магнитных сопротивлений магнитопровода и якоря, S — площадь сечения магнитного потока в воздушном зазоре, w — число витков обмотки. Для определения силы притяжения, действующей на якорь, умножим уравнение (9.7.1) на ток i и проинтегрируем по времени, принимая, что $i = 0$ при $t = 0$:

$$\int_0^t u_{\text{вх}} i dt = R \int_0^t i^2 dt + \int_0^t L \frac{di}{dt} i dt + \int_0^t i^2 \frac{dL}{dt} dt. \quad (9.7.3)$$

Выполняя во втором члене правой части интегрирование по частям, получаем соотношение, выражающее закон сохранения энергии:

$$\int_0^t u_{\text{вх}} i dt = R \int_0^t i^2 dt + \frac{1}{2} Li^2 + \frac{1}{2} \int_0^t i^2 \frac{dL}{dt} dt. \quad (9.7.4)$$

Левая часть этого равенства есть энергия, полученная от входного сигнала. Первый член правой части представляет собой количество тепла, выделенное на активном сопротивлении. Величина $Li^2/2$ представляет собой энергию магнитного поля катушки. Последний член в правой части есть энергия, затраченная на изменение индуктивности. Но изменение индуктивности происходит только за счет перемещения якоря. Следовательно, последний член в правой части формулы (9.7.4) равен механической энергии:

$$\frac{1}{2} \int_0^t i^2 \frac{dL}{dt} dt = \int_0^t F \frac{dx}{dt} dt, \quad (9.7.5)$$

где F — сила притяжения якоря. Дифференцируя равенство (9.7.5) по времени, получаем формулу для силы притяжения:

$$F = F(x, i) = \frac{1}{2} i^2 \frac{dL}{dx}. \quad (9.7.6)$$

Зависимость $F(x, i)$ силы притяжения якоря от положения якоря, определяемого величиной x , и тока в обмотке реле i и зависимость силы пружины от x являются *тяговыми* характеристиками реле. График функции $F(x, i)$ при данном фиксированном значении тока i приведен на рис. 9.7.4.

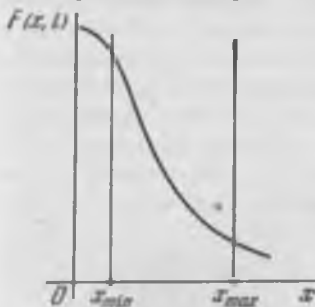


Рис. 9.7.4.



Рис. 9.7.5.

Движение якоря происходит под действием трех сил: силы притяжения, силы сопротивления возвратной пружины и силы реакции контактов при их замыкании $F_k(x)$. Следовательно, уравнение движения якоря имеет вид

$$m\ddot{x} = F(x, i) - c(x - x_0) - F_k(x), \quad (9.7.7)$$

где c — жесткость пружины. Зависимость контактного давления $F_k(x)$ от перемещения x показана на рис. 9.7.5.

Уравнения (9.7.1) и (9.7.7) описывают электромеханические процессы в реле.

Время трогания якоря $T_{тр}$, время движения якоря $T_{дв}$ и время срабатывания реле $T_{ср}$ можно определить путем интегрирования уравнений (9.7.1) и (9.7.7) при начальных условиях $t=0$, $i=0$, $x=x_{max}$. Время трогания якоря $T_{тр.0}$, время его движения при отпускании $T_{дв.0}$ и время отпускания $T_{от}$ можно определить путем интегрирования уравнений (9.7.1) и (9.7.7) при $u_{вх}=0$ и начальных условиях $t=0$, $i=i_0$, $x=x_{min}$, где i_0 — установившееся значение тока в обмотке после срабатывания реле. Интегрирование уравнений (9.7.1) и (9.7.7) можно выполнить численно или с помощью моделирования. Временные характеристики реле, так же как и его тяговые характеристики, могут быть определены и экспериментально.

На рис. 9.7.6 приведена кривая изменения тока при включении и выключении реле, полученная экспериментально. Постоянные времени соответствуют максимальному и минимальному значениям индуктивности $T_{\max} = L_{\max}/R$, $T_{\min} = L_{\min}/R$.

Перейдем к изучению динамических свойств реле как элемента автоматической системы. Время срабатывания и время отпущения реле обычно бывают достаточно малы, чтобы ими можно было пренебречь по сравнению с временами переходных процессов в других элементах и цепях автоматических систем. В таких случаях можно пренебречь индуктивностью в уравнении (9.7.1), т. е.

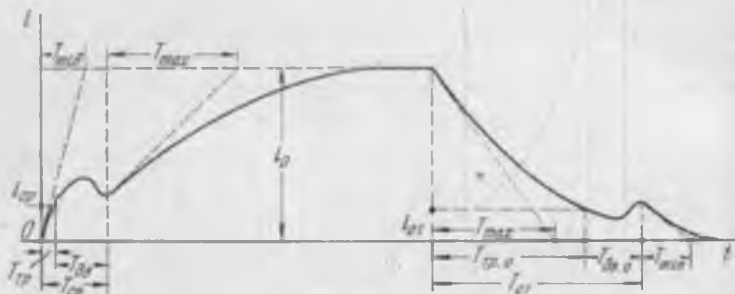


Рис. 9.7.6.

электрической постоянной времени катушки реле $T = L/R$, и временем движения якоря и считать, что контакты реле мгновенно замыкаются, когда входное напряжение $u_{вх}$ достигает значения $u_{ср} = Ri_{ср}$, и мгновенно размыкаются, когда $u_{вх}$, уменьшаясь, достигает значения $u_{от} = Ri_{от}$. Таким образом, реле рассмотренного типа можно приближенно считать безынерционным нелинейным звеном с характеристикой, изображенной на рис. 9.7.7. Если временем срабатывания и временем отпущения пренебречь нельзя, то их обычно учитывают как запаздывание включения и выключения выходного напряжения $u_{вых}$. В таких случаях реле можно рассматривать как последовательное соединение идеального запаздывающего звена с данным запаздыванием τ и безынерционного нелинейного звена с характеристикой, изображенной на рис. 9.7.7.

Рассмотренное электромагнитное реле не реагирует на знак сигнала, вследствие чего оно называется *нейтральным*. В системах управления, где необходимо различать знак сигнала, применяют *поляризованные* реле, которые замыкают одни или другие контакты в зависимости от полярности управляющего напряжения, т. е. реагируют не только на величину, но и на знак входного сигнала.

На рис. 9.7.8 приведена схема современного поляризованного реле. Якорь 1 может поворачиваться вокруг оси 2. С якорем связан подвижный контакт 3. Хвостовик якоря реле расположен в воздушном зазоре магнитопровода. Входной величиной является напряжение, подаваемое на катушку 6. Магнитные потоки постоянного магнита и тока в катушке проходят по различным магнитопроводам, общей частью которых является воздушный зазор.

При отсутствии тока в управляющей обмотке и среднем положении якоря постоянный магнит 5 создает равные магнитные потоки по обеим сторонам якоря. Поток постоянного магнита от северного полюса раздваивается и идет через якорь реле к южному полюсу постоянного магнита.

При протекании по обмоткам реле управляющего тока возникает

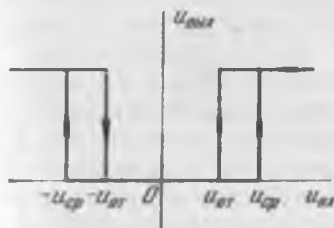


Рис. 9.7.7.

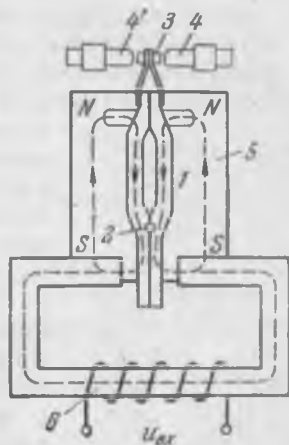


Рис. 9.7.8.

дополнительный магнитный поток. В одной половине зазора магнитный поток от управляющего тока направлен так же, как от постоянного магнита, а в другой половине — противоположно. Вследствие этого на якорь реле действует результирующая сила, которая вызывает его перемещение и замыкание контакта якоря 3 с неподвижным контактом 4 или 4'.

Временные процессы в поляризованных реле аналогичны процессам в рассмотренных выше простейших реле. Однако поляризованные реле обладают большей чувствительностью и быстродействием. Это объясняется тем, что, во-первых, у поляризованных реле нет возвратных пружин, а во-вторых, при малейшем отклонении якоря от нейтрального положения сила, создаваемая постоянным магнитом, помогает силе электромагнита ускорять движение якоря.

Поскольку тяговые характеристики поляризованного реле несколько специфичны, рассмотрим их подробнее. Сила, действующая на якорь поляризованного реле рассматриваемого типа,

определяется формулой

$$F = \frac{10^{-7}}{8\pi S} [(\Phi_y + \Phi_1)^2 - (\Phi_y - \Phi_2)^2], \quad (9.7.8)$$

где магнитные потоки управляющей обмотки Φ_y и постоянного магнита Φ_1, Φ_2 соответственно равны

$$\Phi_y = \frac{0,4\pi (iw)}{R_\mu + \frac{2x_0}{S}}, \quad \Phi_1 = \frac{0,4\pi (iw)_0}{R_\mu + \frac{x_0 - x}{S}}, \quad \Phi_2 = \frac{0,4\pi (iw)_0}{R_\mu + \frac{x_0 + x}{S}}. \quad (9.7.9)$$

В этих формулах (iw) — ампер-витки управляющей обмотки, $(iw)_0$ — магнитодвижущая сила постоянного магнита, $2x_0$ — ширина воздушного зазора, R_μ — сопротивление магнитопровода (без учета воздушного зазора), по которому проходит поток управляющей обмотки, R_μ — сопротивление магнитопровода (без учета воздушного зазора), по которому проходит поток постоянного магнита, S — площадь сечения магнитного потока в воздушном зазоре.

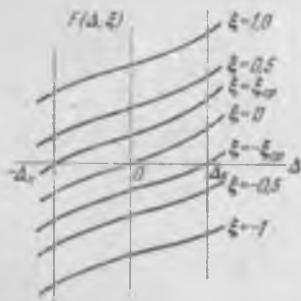


Рис. 9.7.9.

Формулы (9.7.8) и (9.7.9) определяют зависимость силы, действующей на якорь поляризованного реле, от его перемещения x и тока i в управляющей обмотке, т. е. тяговую характеристику поляризованного реле. Обычно силу F выражают как функцию относительного перемещения якоря $\Delta = x/x_0$ и относительных

ампер-витков управляющей обмотки $\xi = (iw)/(iw)_0$. Типовые тяговые характеристики поляризованного реле приведены на рис. 9.7.9. Из рис. 9.7.9 видно, что нейтральное положение якоря неустойчиво даже при $\xi = 0$. Чем больше якорь отклоняется от нейтрального положения, тем больше отклоняющая его сила. Поэтому при снятии управляющего напряжения, т. е. когда ξ становится равным нулю, якорь остается в том из двух крайних положений, в котором он находился до снятия управляющего напряжения. Вследствие этого такое поляризованное реле является *двухпозиционным*.

Для вычисления контактного давления в поляризованном реле достаточно положить в формулах (9.7.8) и (9.7.9) $iw = 0$ ($\xi = 0$) и принять перемещение якоря x равным его значению x_k , при котором происходит замыкание одного из контактов. Для вычисления тока срабатывания $i_{ср}$, т. е. такого значения тока, при котором сила, прижимающая якорь к контакту, переходит через нуль и начинается движение якоря, достаточно положить в формулах (9.7.8) и (9.7.9) $x = x_k$, $F = 0$, $i = i_{ср}$ и решить полу-

ченное уравнение относительно $i_{\text{ср}}$. После определения $i_{\text{ср}}$ можно найти и минимальное значение u_d абсолютной величины входного напряжения $u_{\text{вх}}$, при котором начинается переброска якоря из одного положения в другое.

Так же как в случае простейшего реле, временем срабатывания реле и переброски якоря из одного положения в другое часто

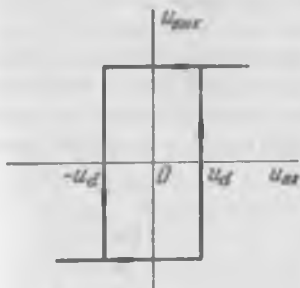


Рис. 9.7.10.

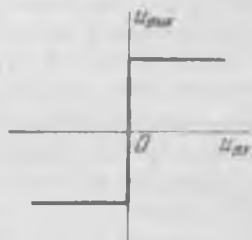


Рис. 9.7.11.

можно пренебрегать. В этом случае двухпозиционное поляризованное реле можно рассматривать как безынерционное звено с характеристикой гистерезисного типа, изображенной на рис. 9.7.10. Если пренебречь и величиной u_d и допустить, что перебрасывание якоря происходит мгновенно при изменении знака $u_{\text{вх}}$, то можно считать, что реле имеет идеализированную однозначную характеристику, изображенную на рис. 9.7.11.

Если временем срабатывания реле пренебрегать нельзя, то двухпозиционное поляризованное реле можно рассматривать как последовательное соединение идеального запаздывающего звена с данным временем запаздывания τ и безынерционного нелинейного звена с характеристикой, изображенной на рис. 9.7.10, или идеализированной характеристикой, изображенной на рис. 9.7.11.

Для создания трехпозиционного реле с устойчивым средним положением якоря при отсутствии управляющего тока можно укрепить якорь на пластинчатой пружине. Выбирая необходимую жесткость пружины, можно получить суммарную тяговую характеристику реле с отрицательной крутизной и тем самым обеспечить устойчивость нейтрального положения при $\xi = 0$. На рис. 9.7.12 приведены типовые тяговые характеристики трехпо-

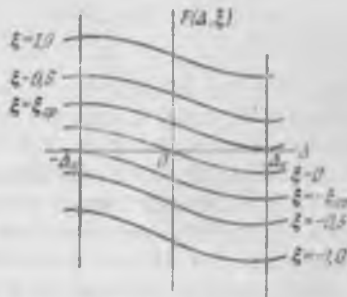


Рис. 9.7.12.

виционного поляризованного реле. Эти тяговые характеристики получаются добавлением в правой части формулы (9.7.8) усилия пружины, равного — cx , где c — жесткость пружины. Вычисление тока срабатывания и контактного давления для трехпозиционного реле производится совершенно так же, как для двухпозиционного.

Если пренебречь временами срабатывания реле и переброски якоря, то трехпозиционное реле можно рассматривать как безынерционное нелинейное звено, характеристика которого приведена на рис. 9.7.13. Если не учитывать разницу в токах срабатывания и отпускания реле, то эта характеристика заменяется идеализированной однозначной характеристикой, изображенной на

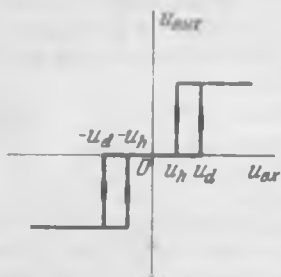


Рис. 9.7.13.



Рис. 9.7.14.

рис. 9.7.14. Наконец, неучет мертвой зоны (пороговой чувствительности реле) приводит к идеальной релейной характеристике, изображенной на рис. 9.7.11. Для приближенного учета времени срабатывания трехпозиционного реле можно считать его последовательным соединением идеального запаздывающего звена с данным временем запаздывания τ и звена с релейной характеристикой типа рис. 9.7.13 или рис. 9.7.14, или рис. 9.7.11.

Электромагнитные реле получили очень широкое распространение в системах управления. Они применяются в системах связи (телефон, телеграф), сигнализации и блокировки, в вычислительной технике и во многих других устройствах. Широкое распространение реле обусловлено рядом их положительных качеств: простотой устройства, небольшими габаритами, большим усилением по мощности и быстрым действием. Некоторые основные характеристики современных реле приведены в таблице 9.7.1.

Недостатком электромагнитных реле является наличие подвижного якоря и контактов, которые могут служить причиной отказов. Поэтому с точки зрения надежности целесообразнее применять бесконтактные реле. В настоящее время широко применяются бесконтактные магнитные реле.

ТАБЛИЦА 9.7.1

Тип реле	Мощность срабатыва- ния, Вт	Мощность управления, Вт	Время сраба- тывания, мс	Коэффициент усиления по мощности
Электромагнитное нейтральное	$10^{-3} \div 10^3$	$0,1 \div 10^4$	$1 \div 200$	$10 \div 10^4$
Электромагнитное поляризованное	$0,005 \div 0,5$	$10 \div 10^2$	$1 \div 15$	$10 \div 10^6$
Магнитоэлектрическое	$10^{-10} \div 10^{-11}$	$0,1 \div 2$	$0,01 \div 2 \cdot 10^3$	$10^9 \div 10^{10}$

Для выяснения принципа устройства бесконтактного магнитного реле рассмотрим магнитный усилитель с положительной обратной связью. Характеристика магнитного усилителя, собранного по дифференциальной схеме (рис. 9.3.7), имеет линейный участок и участки насыщения, обусловленные конечной мощностью выходного сигнала. Эта характеристика может быть приближенно представлена как нелинейность типа ограничения. Обозначим через k крутизну среднего линейного участка характеристики усилителя. Если замкнуть усилитель положительной обратной связью с коэффициентом усиления β , то входным сигналом усилителя будет $I_{вх} + \beta I_{вых}$. Если этот сигнал по абсолютной величине меньше $I_{\max}/k$, то входной и выходной сигналы усилителя связаны соотношением

$$I_{вых} = k (I_{вх} + \beta I_{вых}). \quad (9.7.10)$$

Отсюда находим

$$I_{вых} = \frac{k}{1 - k\beta} I_{вх}. \quad (9.7.11)$$

Эта формула показывает, что введение обратной связи позволяет изменять наклон средней линейной части характеристики усилителя. Части характеристики, соответствующие ограничению выходного сигнала, при этом не изменяются, так как выходной сигнал усилителя не может быть по абсолютной величине больше $I_{\max}$. При $\beta = 1/k$ линейный участок (9.7.11) становится вертикальным и характеристика усилителя принимает вид идеальной релейной характеристики, изображенной на рис. 9.7.11. Однако для реальных усилителей обычно приходится устанавливать значение $\beta = 1,05/k \div 1,10/k$. Из формулы (9.7.11) следует, что при $\beta > 1/k$ наклон характеристики усилителя становится отрицательным и она становится петлеобразной, т. е. гистерезисной (рис. 9.7.15). Действительная характеристика магнитного усилителя, работающего в релейном режиме, вследствие плавности перехода из насыщенного состояния в непасыщенное и обратно, отклоняется от рассмотренной идеальной характеристики и имеет вид сглаженной релейной характеристики (пунктирная кривая на рис. 9.7.15).

Мы показали, как можно получить бесконтактное поляризованное реле путем ввода в дифференциальный магнитный усилитель положительной обратной связи. Совершенно так же можно получить нейтральное бесконтактное реле путем замыкания положительной обратной связью простейшего магнитного усилителя, рассмотренного в § 9.3. Из рис. 9.3.6, показывающего изменение характеристики магнитного усилителя под действием обратной связи, ясно, что подбором коэффициента усиления сигнала обратной связи можно добиться того, что средний участок характеристики усилителя будет вертикальным или даже иметь отрицательный наклон. В последнем случае характеристика усилителя будет петлеобразной, близкой по характеру к одной половине характеристики нейтрального реле, изображенной на рис. 9.7.7.

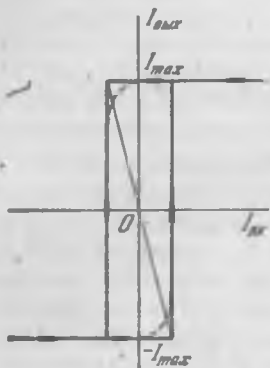


Рис. 9.7.15.

Заметим, что характеристика любого усилителя (электронного, полупроводникового) всегда нелинейна вследствие ограничения мощности выходного сигнала. Поэтому принципиально каждый усилитель можно превратить в бесконтактное реле путем введения положительной обратной связи с соответствующим коэффициентом усиления. Однако практически в таком режиме можно использовать лишь магнитный усилитель, ввиду большой стабильности его характеристик.

Магнитные материалы, обладающие хорошо выраженными гистерезисными свойствами, также могут быть использованы для создания бесконтактных реле. Для некоторых материалов петля гистерезиса имеет прямоугольную форму. Характеристикой близости петли гистерезиса к прямоугольной служит коэффициент прямоугольности, т. е. отношение остаточной индукции B_r к индукции насыщения B_s . Для железоникелевых сплавов типа «перминвар» марок 50 НП, 65 НП, 34 НКМП

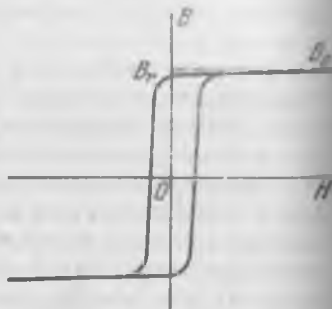


Рис. 9.7.16.

коэффициент прямоугольности достигает значений 0,93; 0,97; 0,98 соответственно. Высокий коэффициент прямоугольности характерен и для некоторых марок ферритов — неметаллических соединений окислов железа, никеля, цинка, марганца и других металлов. На рис. 9.7.16

приведена характеристика кривой намагничивания для ферритового сердечника. Наибольшее распространение магнитные гистерезисные элементы получили в вычислительной технике и в импульсных системах автоматики [60]. На рис. 9.7.17 приведена схема бесконтактного реле, применяемого в качестве элемента памяти в вычислительных машинах. На ферритовый сердечник, имеющий прямоугольную петлю гистерезиса, намотаны три обмотки: w_y , w_1 , w_2 . Схема приводится в исходное положение, соответствующее отрицательной остаточной индукции $-B_r$, импульсом тока i_1 , поступающим в обмотку w_1 . Если в обмотке w_y импульса тока нет, то система сохраняет свое состояние. При подаче импульса тока i_y в обмотку w_y система переводится в состояние с остаточной индукцией $+B_r$. Два устойчивых положения бесконтактного реле соответствуют числам 0 и 1 двоичной системы счисления, широко применяемой в вычислительных машинах. Таким образом, знак остаточной индукции определяет цифру 0 или 1. Для считывания из ячейки памяти соответствующей цифры в обмотку w_1 снова подается импульс тока i_1 , переводящий систему в исходное состояние $-B_r$. При этом индукция в сердечнике не изменяется, если был записан нуль, и изменяется от $+B_r$ до $-B_r$, если была записана единица, что сопровождается появлением импульса э. д. с. в выходной обмотке сердечника. Эта э. д. с. и позволяет определить, что в ячейке памяти была записана единица.

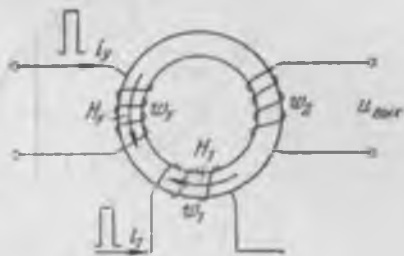


Рис. 9.7.17.

Бесконтактные реле характерны высокой надежностью, малыми габаритами и высоким быстродействием. Поэтому они являются перспективными элементами систем управления.

В автоматических системах управления нашли широкое применение *тиратронные* реле благодаря их высокой чувствительности и большой надежности. У тиратронных реле мощность выходного сигнала может достигать 10^3 Вт при мощности входного сигнала 10^{-3} Вт.

Основным элементом тиратронного реле является тиратрон — трех- или четырехэлектродная лампа, наполненная неон, аргон или парами ртути. В тиратроне за счет термоэлектронной эмиссии катода происходит ионизация молекул газа. Для управления анодным током служит сетка. Однако действие сеточного напряжения на изменение анодного тока в тиратроне существенно отличается от аналогичного действия в вакуумных лампах. Управляющее действие сетки тиратрона заключается в создании

дополнительного (обратного анодного) электрического поля между сеткой и катодом, благодаря которому становится невозможной полная ионизация молекул газа. Поэтому при напряжении на сетке, меньшем напряжения зажигания $u_{сз}$, сетка играет роль экрана, препятствующего распространению ионизации газа на все пространство между анодом и катодом. В этом случае существует динамическое равновесие между числами образующихся в единицу времени ионизированных и рекомбинированных молекул (нейтральных молекул, образовавшихся при соединении свободных электронов с ионами). При достижении u_c напряжения

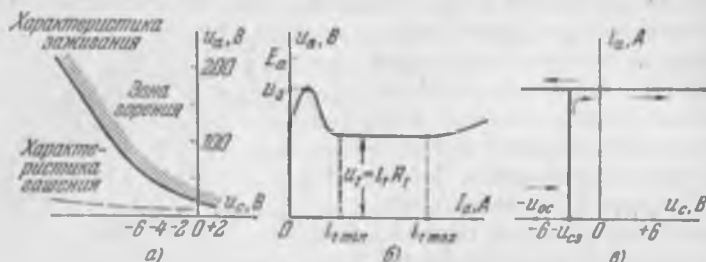


Рис. 9.7.18.

$u_{сз}$ это динамическое равновесие нарушается, начинается лавинообразный процесс ионизации газа между катодом и анодом, и тиратрон зажигается. Время лавинообразного процесса (зажигания) длится $10^{-6} \div 10^{-4}$ с. Ток в цепи анода, протекающий после зажигания тиратрона, определяется потоком электронов и ионов. Поэтому при одинаковом токе накала анодный ток тиратронов в несколько раз больше анодного тока вакуумных ламп. В момент зажигания тиратрона его внутреннее сопротивление скачком падает и соответственно возрастает анодный ток. Поэтому зависимость, связывающая сеточное напряжение с анодным током, представляет собой специфичную релейную характеристику. Падение напряжения на горящем тиратроне не зависит от величины анодного тока и является постоянным для данного типа тиратрона при постоянных напряжении источника питания и температуре окружающей среды. Параметры зажигания и горения тиратрона определяются характеристикой зажигания (рис. 9.7.18, а) и вольт-амперной характеристикой (рис. 9.7.18, б). Характеристика тиратрона, связывающая входную и выходную переменные, показана на рис. 9.7.18, в.

После зажигания тиратрона сетка теряет свое управляющее свойство. Поэтому изменением напряжения на сетке тиратрон погасить нельзя, даже при подаче отрицательного потенциала, превышающего в несколько раз первоначальный потенциал запира-ния. Это происходит потому, что положительные ионы газа, запол-

няющие после зажигания тиратрона все внутреннее пространство, окружают сетку со всех сторон и, подходя к ней, присоединяют недостающий электрон, превращаясь в нейтральные молекулы газа, которые благодаря тепловому движению будут уходить от сетки, уступая место другим ионам. Такой процесс рекомбинации нейтрализует влияние электрического поля сетки тиратрона. Для его гашения необходимо или разорвать анодную цепь, или уменьшить анодное напряжение до напряжения гашения. Практически в простых схемах применяемых тиратронных реле для гашения чаще всего используют питание анодной цепи переменным

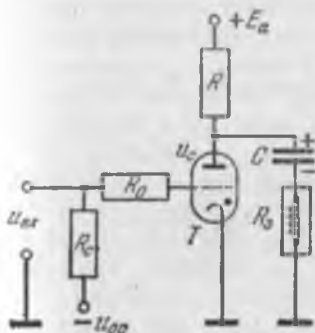


Рис. 9.7.19.

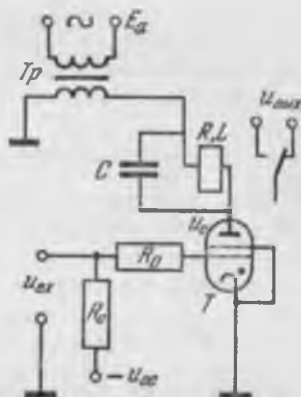


Рис. 9.7.20.

током или применяют конденсатор, создающий при разряде в анодной цепи протпвоэлектродвижущую силу, уменьшающую u_a до напряжения гашения.

Рассмотрим две типовые схемы тиратронных реле: а) импульсного реле с питанием анодной цепи постоянным током (рис. 9.7.19) и б) реле переключения с питанием анодной цепи переменным током (рис. 9.7.20).

В первой схеме параллельно тиратрону включена цепь, состоящая из конденсатора C и запала R_a . Для запираания тиратрона подается напряжение отрицательного смещения $-u_{oc}$ ($u_{oc} > u_{ca}$) через сопротивление R_c . Сопротивление R_0 служит для ограничения тока сетки. При $u_{вх} < u_{oc} - u_{ca}$ тиратрон не горит, конденсатор C заряжается через достаточно большое сопротивление R и сопротивление R_a . Ток зарядки, протекающий через R_a , мал и не вызывает срабатывания запала. При подаче на вход положительного напряжения $u_{вх} \geq u_{oc} - u_{ca}$ напряжение на сетке u_c станет равным u_{ca} , тиратрон зажжется и замкнет цепь конденсатора и запала. Горящий тиратрон имеет малое сопротивление R_t ,

поэтому ток разрядки конденсатора через R_a и R_T будет больше тока, необходимого для срабатывания запала, и последний после разогрева сопротивления до температуры зажигания капсюля сработает.

Во второй схеме в анодную цепь, питаемую через трансформатор Tr , включено электромагнитное реле (или, в бесконтактном варианте, управляющая обмотка магнитного реле). Цени сетки

устроены так же, как и в первой схеме. Конденсатор C вместе с обмоткой реле (R, L) служит фильтром для сглаживания импульсов анодного тока, протекающего через обмотку реле при зажигании тиратрона. Когда u_c станет равным u_{ca} (при $u_{ax} \geq u_{oc} - u_{ca}$), тиратрон будет зажигаться и гаснуть в каждый положительный полупериод изменения E_a , и в его анодной цепи будет протекать импульсный постоянный ток. Когда значение анодного тока в первый положительный полупериод достигнет величины

тока срабатывания электромагнитного реле, последнее сработает и замкнет контакты выходной цепи. Реле останется замкнутым на все время импульсного зажигания тиратрона, так как за время

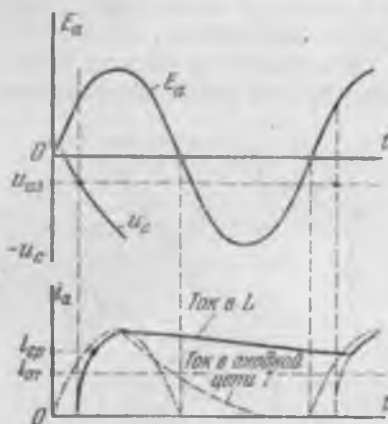


Рис. 9.7.21.

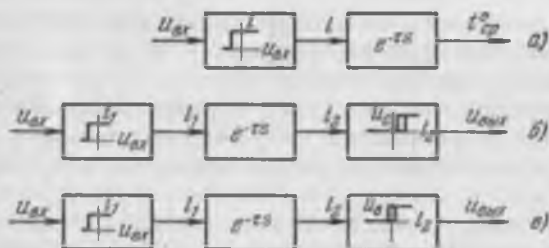


Рис. 9.7.22.

его горения в каждом положительном полупериоде конденсатор заряжается, а за время отрицательного полупериода, когда тиратрон не горит, разряжается через обмотку реле, причем минимальное значение тока разрядки превышает ток отпускания реле (рис. 9.7.21). При уменьшении значения сигнала на входе до $u_{ax} < u_{oc} - u_{ca}$ тиратрон перестанет зажигаться в каждый

положительный полупериод, реле отпустит и разомкнет контакты выходной цепи. Тиратронное реле как элемент автоматической системы управления можно приближенно представить последовательным соединением безынерционного нелинейного и запаздывающего звеньев с нелинейностью, характеристика которой показана на рис. 9.7.18, в.

Структурная схема первого из рассмотренных реле изображена на рис. 9.7.22, а, а второго на рис. 9.7.22, б для случая электромагнитного реле и на рис. 9.7.22, в для случая бесконтактного реле. В первой схеме время запаздывания τ зависит от постоянных C , R и R_T , а также от времени разогрева капсюля до температуры зажигания $t_{ср}$. Во второй схеме τ зависит от частоты питающего напряжения, времени срабатывания реле и постоянной времени сглаживающего фильтра.

§ 9.8. Модуляторы и демодуляторы

Передачу сигналов в автоматических системах от одного элемента к другому часто целесообразно осуществлять при помощи переменного тока. Переменный ток (не обязательно синусоидальный), с помощью которого осуществляется передача сигнала, называется *несущей*. Для передачи сигнала необходимо изменять в зависимости от этого сигнала какой-нибудь параметр переменного тока. Этот процесс называется *модуляцией* несущей сигналом. В соответствии с тем, какой параметр несущей изменяется в зависимости от модулирующего сигнала, различают амплитудную модуляцию, частотную модуляцию и фазовую модуляцию.

При амплитудной модуляции амплитуда переменного напряжения является определенной функцией, обычно линейной, передаваемого сигнала $x(t)$. В соответствии с этим модулированный сигнал $u(t)$ в случае синусоидального переменного тока выражается формулой

$$u(t) = U(t) \sin \omega t, \quad (9.8.1)$$

где ω — частота несущей, а

$$U(t) = U_0 [1 + kx(t)]. \quad (9.8.2)$$

Переменная амплитуда $U(t)$ сигнала называется *огibaющей* сигнала. Величина U_0 представляет собой амплитуду сигнала при $x(t) = 0$, т. е. амплитуду напряжения несущей. Величина

$$\Delta U = U(t) - U_0 = U_0 kx(t), \quad (9.8.3)$$

пропорциональная модулирующему сигналу $x(t)$, называется *отклонением* огибающей. Максимальное относительное изменение амплитуды модулированного сигнала

$$m = k |x(t)|_{\max} \quad (9.8.4)$$

называется *коэффициентом модуляции*.

Устройства, осуществляющие преобразование медленно изменяющегося сигнала постоянного тока $u_{вх}$ в сигнал переменного тока (не обязательно синусоидального), имеющего амплитуду, пропорциональную $u_{вх}$, называются *модуляторами*. Модуляторами называются также устройства, осуществляющие частотную или фазовую модуляцию.

Для выделения передаваемого сигнала $x(t)$ из модулированного переменного напряжения $u(t)$ необходимо преобразовать это напряжение в напряжение постоянного тока. Эта операция

называется *демодуляцией*. Демодуляция сигналов переменного тока осуществляется *демодуляторами* или *фазовыми дискриминаторами*. В радиотехнике операция выделения модулирующего сигнала обычно называется *детектированием* и в соответствии с этим устройства, выполняющие эту операцию, называются *детекторами*.

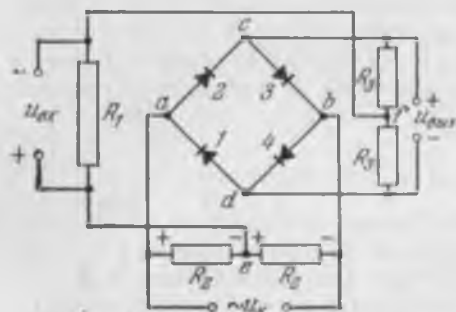


Рис. 9.8.1.

На рис. 9.8.1 показана схема кольцевого модулятора,

выполненного на полупроводниковых диодах. Принцип работы кольцевого модулятора основан на свойстве диодов пропускать ток в одном направлении и не пропускать его в противоположном направлении (§ 9.4).

Напряжение несущей, называемое также *коммутирующим* напряжением u_n , приложено между точками a и b моста, в плечи которого включены диоды 1, 2, 3, 4 с одинаковыми вольтамперными характеристиками. Медленно изменяющееся модулирующее напряжение $u_{вх}$ приложено к сопротивлению R_1 . Для нормальной работы модулятора необходимо, чтобы амплитуда коммутирующего напряжения была значительно больше максимального значения медленно изменяющегося входного напряжения.

Режим работы модулятора определяется переменным коммутирующим напряжением u_n . В полупериоде коммутирующего напряжения, в котором его плюс приложен к точке a , как показано на рис. 9.8.1, коммутирующие токи идут в направлении от точки a через диоды 2 и 3 к точке b , от точки a через диод 2, сопротивления R_2 и R_1 к точке e ; от точки e через сопротивления R_1 , R_2 и диод 3 к точке b . В результате через сопротивление R_3 один ток идет от точки c к точке f , другой, — наоборот, от точки f к точке c . Таким образом, коммутирующее напряжение при $u_{вх} = 0$ создает в сопротивлении R_3 два тока, равных по величине и противоположных по направлению. Это означает, что потенциалы

точек c и f одинаковы и фактически в сопротивлении R_3 тока не будет. В другом полупериоде коммутирующего напряжения к точке a будет приложен минус, к точке b — плюс и ток пойдет через диоды 4 и 1 . При этом потенциалы точек d и f при $u_{вх} = 0$ будут одинаковы и тока в сопротивлении R_3 также не будет. Таким образом, напряжение u_k не оказывает влияния на выходной сигнал.

Если $u_{вх} \neq 0$ и имеет полярность, как показано на рис. 9.8.1, то потенциалы точек c и d больше потенциала точки f , причем разность потенциалов между точками c и d и точкой f пропорциональна $u_{вх}$.

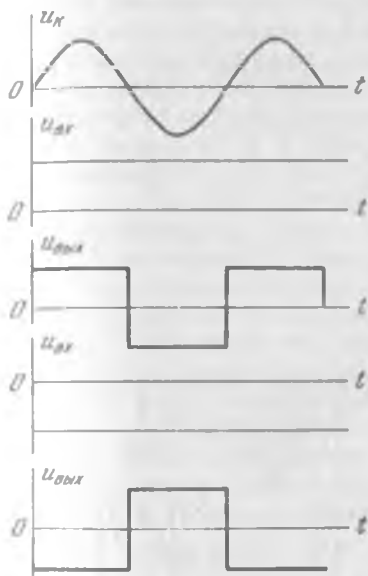


Рис 9.8.2.

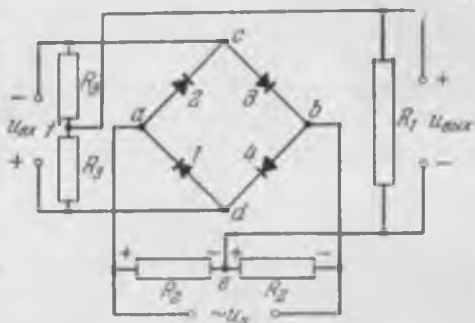


Рис. 9.8.3.

Поэтому в полупериоде коммутирующего напряжения, когда его плюс приложен к точке a , через верхнее (на схеме) сопротивление R_3 будет протекать ток от точки c к точке f , а в другом полупериоде коммутирующего напряжения ток будет протекать через нижнее сопротивление от точки d к точке f . В результате на выходе модулятора формируется переменное напряжение с амплитудой, пропорциональной входному напряжению. На рис. 9.8.2 показаны графики изменения сигналов $u_{вх}$, $u_{кв}$ и $u_{вых}$. При изменении полярности входного напряжения потенциал точки f становится больше потенциала точек c и d . Поэтому ток через сопротивление R_3 в соответствующих полупериодах коммутирующего напряжения будет протекать в обратных направлениях. Таким образом, при изменении знака входного сигнала фаза выходного напряжения изменяется на 180° .

В качестве демодулятора можно применить рассмотренный кольцевой модулятор, если поменять местами роли входа и выхода (рис. 9.8.3).

Рассмотрим сначала случай, когда фаза переменного входного напряжения $u_{вх}$ совпадает с фазой коммутирующего напряжения. Пусть в полупериод, когда к точке a приложен положительный потенциал коммутирующего напряжения, к точке d будет приложен плюс входного напряжения. В этом случае открыты диоды 2 и 3 и ток от точки f течет через сопротивления R_1 , R_2 и диод 2 к точке c . С изменением фазы коммутирующего напряжения одновременно изменяется фаза входного напряжения,

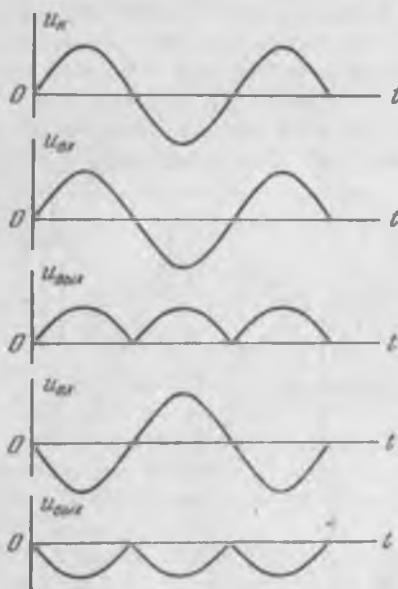


Рис. 9.8.4.

так как их частоты одинаковы. При этом открываются диоды 4 и 1. Ток от точки f течет через сопротивления R_1 , R_2 и диод 4 к точке d . В обоих полупериодах ток через сопротивление R_1 течет в одном направлении независимо от полярности вход-

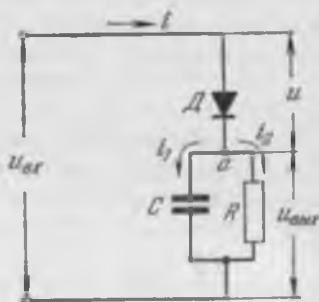


Рис. 9.8.5.

ного напряжения. Полярность выходного напряжения изменится только в том случае, если фаза входного напряжения изменится на 180° по отношению к фазе коммутирующего напряжения. На рис. 9.8.4 показаны графики изменения сигналов.

Рассмотрим теперь принцип устройства амплитудного детектора, принципиальная схема которого изображена на рис. 9.8.5. Детектор состоит из диода D и фильтра RC . Фильтр служит для уменьшения высокочастотных пульсаций напряжения на нагрузке и для увеличения коэффициента передачи детектора. Из схемы устройства вытекают следующие уравнения:

$$\left. \begin{aligned} i &= i_1 + i_2, \quad i = \varphi(u), \\ u &= u_{вх} - u_{вых}, \\ i_{вых} &= \frac{i_1}{C}, \quad u_{вых} = i_2 R, \end{aligned} \right\} \quad (9.8.5)$$

где $\varphi(u)$ — вольтамперная характеристика диода (см. § 9.4). Исключая переменные i , i_1 , i_2 , u , получаем нелинейное дифференциальное уравнение, связывающее входное и выходное напряжения:

$$T\dot{u}_{\text{вых}} + u_{\text{вых}} = R\varphi(u_{\text{вх}} - u_{\text{вых}}), \quad T = RC. \quad (9.8.6)$$

Структурная схема детектора, соответствующая уравнению (9.8.6), изображена на рис. 9.8.6. Из схемы следует, что детектор

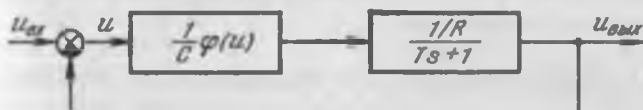


Рис. 9.8.6.

можно рассматривать как систему, состоящую из последовательно соединенных безынерционного нелинейного звена с характеристикой $\varphi(u)$ и апериодического звена, охваченных жесткой отрицательной обратной связью.

Рассмотрим процесс детектирования амплитудно-модулированного входного сигнала

$$u_{\text{вх}} = U(t) \sin \omega t, \quad (9.8.7)$$

т. е. процесс выделения амплитуды $U(t)$.

На рис. 9.8.7 изображены модулированный входной сигнал $u_{\text{вх}}$ (сплошная тонкая линия) и выходное напряжение $u_{\text{вых}}$, снимаемое с конденсатора C . При положительных напряжениях на аноде диода, т. е. при $u_{\text{вх}} > u_{\text{вых}}$, диод открыт, вследствие чего конденсатор быстро заряжается током, протекающим через диод. При отрицательном значении входного напряжения диод закрыт и конденсатор C разряжается. При этом напряжение на конденсаторе в среднем повторяет закон изменения амплитуды $U(t)$ входного сигнала (9.8.7). Если амплитуда входного напряжения $U(t)$ в некотором интервале времени убывает быстрее, чем напряжение на конденсаторе, то выходной сигнал детектора не успевает следить за изменением амплитуды $U(t)$ (участок AB на рис. 9.8.7), вследствие чего возникает ошибка детектора, обусловленная инерционностью схемы. Для устранения подобных ошибок необходимо выбирать сопротивление R и емкость C таким образом, чтобы скорость падения напряжения на конденсаторе превышала



Рис. 9.8.7.

максимально возможную скорость изменения модулирующего сигнала. Если модулирующий полезный сигнал $x(t)$ имеет ограниченную полосу частот Ω , то для выполнения сформулированного условия достаточно, чтобы постоянная времени $T = RC$ удовлетворяла неравенству

$$T \leq \frac{\sqrt{1-m^2}}{m\Omega}, \quad m = \max \{|x(t)|\}, \quad (9.8.8)$$

которое мы приводим здесь без доказательства. При выполнении сформулированного условия детектор можно считать практически безынерционным элементом.

§ 9.9. Электродвигатели с фрикционными муфтами

Одним из недостатков электродвигателей как исполнительных устройств в автоматических системах является их большая инерционность. Для устранения влияния инерционности двигателя на работу автоматической системы применяют фрикционные муфты.

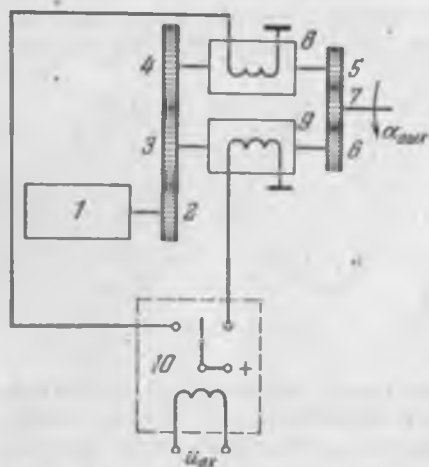


Рис. 9.9.1.

Схема устройства двигателя с фрикционной муфтой показана на рис. 9.9.1. Двигатель 1 через шестерню 2 вращает с постоянной скоростью в различных направлениях шестерни 3 и 4. Движение шестерен 3 и 4 может передаваться шестерням 5 и 6 через фрикционные муфты 8 и 9. Шестерни 5 и 6 соединены с шестерней 7, угол поворота которой является выходной переменной исполнительного устройства. Входной переменной является напряжение $u_{\text{вх}}$, подаваемое на управляющую обмотку трехпозиционного поляризованного реле 10. В зависимости от полярности входного напряжения якорь реле замыкает те или другие контакты, и напряжение источника, подключенное к якорю, прикладывается к обмотке соответствующей фрикционной муфты. Фрикционная муфта включается, и вращение от двигателя передается на выходной вал. Полярность входного напряжения определяет направление вращения выходного вала.

В системах управления обычно применяются электромагнитные и порошковые фрикционные муфты. Схема устройства элект-

трогательной муфты показана на рис. 9.9.2. При подаче тока в обмотку 1 электромагнита возникает магнитный поток, который создает усилие, перемещающее якорь 2 и прижимающее фрикционное кольцо 3 к вращающемуся кольцу 6. В результате трения между кольцами 3 и 6 вращение передается от шестерни 4 к шестерне 5.

Рассмотрим динамику разгона выходного вала муфты при включении электромагнита. От момента приложения напряжения на обмотку электромагнита до момента соприкосновения фрикционных дисков проходит время $T_{тр} + T_{дв}$, определяемое процессами нарастания тока в обмотке и движения якоря электромагнита. При соприкосновении фрикционных дисков возникает момент трения, равный

$$M_{тр} = f l F_v, \quad (9.9.1)$$

где F_v — сила притяжения электромагнита, f — коэффициент трения, l — средний радиус фрикционного кольца. Когда момент трения превысит момент нагрузки M_n , приложенный к выходному валу, последний начинает вращаться. Движение выходного вала описывается уравнением

$$J \dot{\Omega} = M_{тр} - M_n, \quad (9.9.2)$$

где Ω — угловая скорость выходного вала, J — приведенный момент инерции выходного вала.

При постоянных значениях момента трения и момента нагрузки скорость вращения выходного вала нарастает по линейному закону до тех пор, пока не достигнет значения, при котором проскальзывание фрикционных колец 3 и 6 прекратится. Постоянное значение скорости Ω_0 определяется оборотами двигателя и передаточным отношением от вала двигателя к выходному валу. В существующих конструкциях электромагнитных муфт время разгона при отсутствии нагрузки с учетом времени срабатывания реле составляет $10^{-3} \div 5 \cdot 10^{-3}$ с.

При отклонении муфты и малом моменте нагрузки возможно длительное вращение выходного вала по инерции (выбег). Для предотвращения выбега применяются специальные тормозные муфты, включающиеся при отключении фрикционных дисков.

Вследствие малости времени разгона выходного вала часто можно считать двигатель с фрикционной муфтой безынерционным звеном с релейной характеристикой (рис. 9.9.3). Если желательно

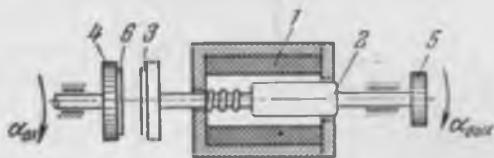


Рис. 9.9.2.

учесть время разгона выходного вала, то можно считать двигатель с фрикционной муфтой последовательным соединением идеального запаздывающего звена и звена с релейной характеристикой, изображенной на рис. 9.9.3.

Порошковые фрикционные муфты бывают двух типов: дисковые и цилиндрические. Конструкция дисковой муфты изображена

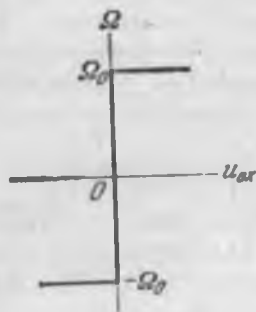


Рис. 9.9.3.

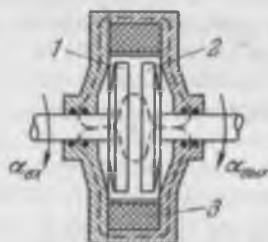


Рис. 9.9.4.

на рис. 9.9.4. В пространстве между ведущим 1 и ведомым 2 дисками находится ферромагнитная масса, которая может иметь жидкую или твердую основу.

Ферромагнитная масса на жидкой основе представляет собой смесь порошка железа с маслом. Ферромагнитная масса на твердой основе представляет собой смесь порошков железа и графита или талька.

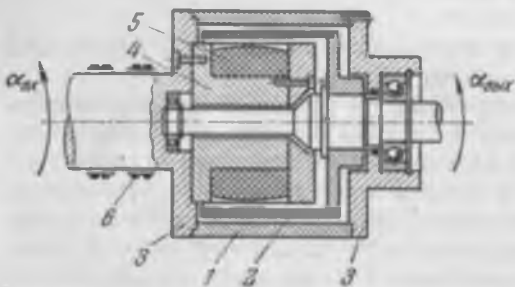


Рис. 9.9.5.

При отсутствии тока в обмотке 3 на ведомый диск передается только малый момент, обусловленный трением в ферромагнитной массе. Если в обмотке муфты течет ток, то между дисками возникает магнитное поле и свойства ферромагнитной

массы резко изменяются. Частицы железа, располагаясь по силовым линиям поля (пунктирная линия на рис. 9.9.4), образуют цепочки. Смесь как бы загустевает, связываясь в единое целое массу. В результате возникает значительный момент трения, определяемый напряженностью магнитного поля.

Конструкция цилиндрической порошковой муфты изображена на рис. 9.9.5. Стальной цилиндр 1 связан с ведущим валом, а топи-

костенный стакан 2 — с ведомым. С ведущей частью жестко связаны торцевые крышки 3 из немагнитного материала и магнитопровод 4, в котором уложена управляющая обмотка 5. Питание на обмотку подается через контактные кольца 6. Ферромагнитная масса заполняет объем между цилиндром 1, стаканом 2 и магнитопроводом 4. За счет такой конструкции удается увеличить рабочую поверхность фрикционной муфты вдвое.

Схема включения порошковых муфт может быть релейной (рис. 9.9.1) или линейной (рис. 9.9.6). Возможность использования линейной схемы основана на пропорциональной зависимости момента трения муфты от управляющего тока. Если токи в обмотках электромагнитов муфт i_1 и i_2 равны, то моменты трения муфт одинаковы, противоположны по знаку и выходной вал не вращается. Когда токи i_1 и i_2 не равны, возникает момент пропорциональный разности токов $M = k(i_1 - i_2)$.

Порошковые муфты более экономичны и малогабаритны по сравнению с электромагнитными. Так, например, при одинаковом токе управления порошковые муфты обеспечивают момент трения в полтора раза больший, чем электромагнитные.

Основными преимуществами электродвигателей с фрикционными муфтами являются малая инерционность и малая потребляемая мощность сигнала управления. Повышение быстродействия достигается, во-первых, тем, что управление осуществляется дисками муфт, моменты инерции которых значительно меньше, чем моменты инерции роторов двигателей. Во-вторых, двигатель непрерывно вращается, имея большую кинетическую энергию, вследствие чего присоединение к валу двигателя неподвижного выходного вала муфты не оказывает существенного влияния на режим его работы. Небольшие массы подвижных частей муфт не требуют большой мощности управляющего сигнала. Поэтому быстродействие передачи с фрикционными муфтами в десятки и сотни раз больше, чем передачи с реверсивным двигателем.

Для сравнения приведем следующие цифры. Электродвигатель ДИД-0,5, широко применяемый в маломощных следящих системах, имеет электромеханическую постоянную времени 0,1 с. Если использовать этот двигатель с электромагнитными фрикционными муфтами, то запаздывание исполнительного устройства может быть доведено до 0,005 с.

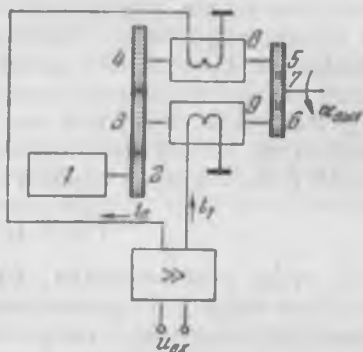


Рис. 9.9.6.

§ 9.10. Электродвигатели с редукторами

В следящих системах электродвигатели обычно имеют редуктор с зубчатой передачей. В системе электродвигатель — редуктор наблюдаются нелинейные явления. Из них наиболее характерны: насыщение в магнитной системе двигателя, нелинейное трение и люфт в механической передаче.

В двигателях постоянного тока зависимость между током управления и магнитным потоком линейна только в небольшом диапазоне, за пределами которого, вследствие явления насыщения стальной части магнитопровода, характеристика носит ярко выраженный нелинейный характер. С увеличением тока управления, или, что то же, управляющего напряжения u , в случае двигателя с управлением через обмотку якоря магнитный поток, а следовательно, и вращающий момент линейно возрастает. Затем наступает насыщение магнитопровода и характеристика принимает вид ограничителя. Учитывая это, уравнение движения выходного вала двигателя постоянного тока с управлением через обмотку якоря (3.13.7) необходимо записать в виде

$$T\dot{\Omega} + \Omega = \varphi(u) - k_c M_c, \quad (9.10.1)$$

где $\varphi(u)$ — нелинейная функция управляющего напряжения (ограничитель). В установившемся режиме при больших значениях управляющего напряжения угловая скорость Ω постоянна и не зависит от управляющего напряжения. Таким образом, насыщение магнитопровода приводит к ограничению угловой скорости вращения выходного вала двигателя. Трение в механической передаче создает момент сопротивления M_c :

$$M_c = M_1 + M_2, \quad (9.10.2)$$

где $M_1 = M_0 \operatorname{sgn} \Omega$ — момент сухого трения, $M_2 = k_1 \Omega^r$ — момент вязкого трения, который при малых скоростях имеет линейный характер ($r = 1$), а при больших — квадратичный ($r = 2$).

Сухое трение не зависит от величины скорости и определяется лишь ее знаком. Важной особенностью сухого трения является то, что при равенстве нулю скорости момент трения может принимать любое значение в пределах $-M_0 \leq M_1 \leq M_0$, причем в каждый данный момент времени момент трения равен сумме всех других моментов, включая момент сил инерции. Поэтому, если скорость стала равной нулю и сумма других моментов окажется по абсолютной величине меньше M_0 , то система остановится. Застой будет продолжаться до тех пор, пока сумма моментов не превысит величину M_0 , после чего снова начнется движение системы. Если в процессе движения системы при $\Omega = 0$ всегда оказывается, что сумма всех моментов больше момента трения, то застой не будет.

Исследование работы электродвигателя с редуктором в общем случае с учетом силы сухого трения чрезвычайно затруднено. Значительного упрощения можно достигнуть, если считать двигатель безынерционным. Полагая в (9.10.1) $T = 0$, учитывая только сухое трение $M_c = M_1 = M_0 \operatorname{sgn} \Omega$ и рассматривая линейную часть характеристики момента управления $\varphi(u) = ku$, получим

$$\left. \begin{aligned} \Omega + M_0^* \operatorname{sgn} \Omega &= ku \quad \text{при} \quad \Omega \neq 0, \\ -M_0^* &\leq ku \leq M_0^* \quad \text{при} \quad \Omega = 0, \end{aligned} \right\} \quad (9.10.3)$$

где $M_0^* = k_c M_0$. Отсюда определяется угловая скорость Ω как нелинейная функция управляющего напряжения u :

$$\Omega = \begin{cases} ku - M_0^* & \text{при} \quad u > M_0^*/k, \\ 0 & \text{при} \quad |u| \leq M_0^*/k, \\ ku + M_0^* & \text{при} \quad u \leq -M_0^*/k. \end{cases} \quad (9.10.4)$$

Из этой формулы следует, что сухое трение приводит к возникновению зоны нечувствительности. Совместное действие двух факторов: насыщения в магнитной системе двигателя и сухого трения

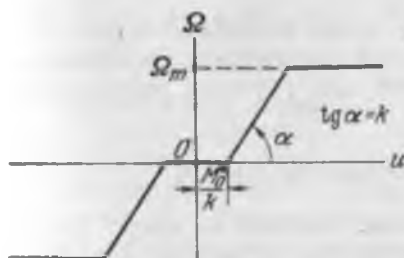


Рис. 9.10.1.

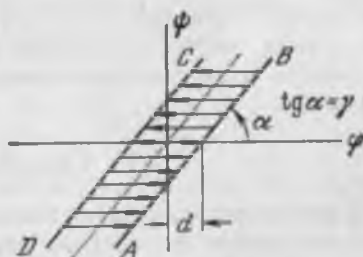


Рис. 9.10.2.

в механической передаче (без учета инерционности двигателя) — приводит к нелинейной зависимости $\Omega = \Omega(u)$, изображенной на рис. 9.10.1.

Рассмотрим теперь влияние люфта. При наличии люфта, т. е. зазора в механической передаче, перемещение ведущего звена не вызывает перемещения ведомого звена до тех пор, пока не будет выбран зазор. Связь между входной и выходной переменными соответствует прямой AB на рис. 9.10.2, где через $2d$ обозначена величина зазора, отнесенная к ведущему звену, а через $\gamma = \operatorname{tg} \alpha$ — передаточное отношение. При перемене знака производной входной величины ведомое звено как бы отключается от ведущего, пока последнее не выберет зазор; при этом зависимость между входной и выходной переменными теперь описывается прямой CD на рис. 9.10.2.

Причиной люфта являются неточность изготовления и одностороннее соприкосновение зубьев шестерен. Так, например, для двигателя ДИД-0,5 с редуктором, имеющим передаточное отношение $\gamma = \psi/\varphi = 0,002$, величина люфта на выходе редуктора составляет 0,032 рад.

Нелинейность типа люфта оказывает существенное влияние на характер работы динамической системы, вызывая статические ошибки и автоколебания. При изготовлении редуктора принимаются меры к тому, чтобы люфт был меньше, чем половина ошибки, допускаемой для автоматической системы. В прецизионных системах обычно применяются различные способы устранения люфта.

Структурная схема системы электродвигатель — редуктор с учетом инерционности двигателя, насыщения магнитной системы



Рис. 9.10.6.

двигателя, нелинейного трения и люфта приведена на рис. 9.10.3. Эта схема является типичной для различных исполнительных устройств и широко используется при расчетах автоматических систем.

§ 9.11. Гидравлические двигатели

Гидравлические исполнительные устройства имеют ряд важных преимуществ перед электрическими, а именно они обладают большей мощностью и большим быстродействием, чем электродвигатели, при тех же габаритах и весах. Это объясняется высоким коэффициентом использования веса конструкции (напряженность силы в гидравлических двигателях достигает $200 \div 300$ кгс/см², а в электрических двигателях магнитное поле создает напряженность силы порядка $4 \div 6$ кгс/см²).

Существует много конструкций гидродвигателей [40]. Рассмотрим широко распространенную схему с золотниковым распределителем (рис. 9.11.1). Гидравлический шестеренчатый насос 1 постоянной производительности обеспечивает подачу жидкости с постоянной скоростью в аккумулятор давления 3. Давление рабочей жидкости составляет $150 \div 280$ ат. После того как давление в аккумуляторе достигает определенной величины, клапан 2 переключает насос на резервуар 4. Из аккумулятора 3 рабочая жидкость поступает в золотниковый распределитель 6. При отсутствии входного сигнала $u_{вх}$, подаваемого на обмотку электромагнита 7, якорь 8 находится в нейтральном положении и золотник 9

закрывает доступ жидкости в рабочий цилиндр 10. При наличии входного сигнала электромагнит поворачивает якорь 8, перемещая золотник 9. Смещение золотника вправо открывает доступ для рабочей жидкости через канал 11 в рабочий цилиндр. Одновременно открывается доступ для рабочей жидкости, находящейся справа от поршня, в канал 12 и далее на слив в резервуар 4. Под действием разности давлений рабочей жидкости поршень 5 перемещается вправо. Выходной величиной гидродвигателя является перемещение y штока рабочего цилиндра.

Для исследования динамических свойств гидродвигателя составим уравнения движения поршня в рабочем цилиндре:

$$m\ddot{y} = (p_1 - p_2)S - R - G \quad (|y| \leq y_{\max}), \quad (9.11.1)$$

где m — приведенная масса поршня с учетом присоединенных к нему масс, y — перемещение поршня, $(p_1 - p_2)S$ — сила давления жидкости, R — сила сопротивления, G — сопротивление нагрузки, $y_{\max}$ — максимальное перемещение поршня от среднего (нулевого) положения. Сила сопротивления R создается за счет трения. В реальных системах сила трения может быть представлена в виде суммы жидкого и сухого трения:

$$R = c\dot{y} + N \operatorname{sgn} \dot{y}, \quad (9.11.2)$$

где $c\dot{y}$ — сила жидкого (вязкого) трения, $N \operatorname{sgn} \dot{y}$ — сила сухого трения. Сопротивление нагрузки зависит от перемещения поршня y . Эту зависимость часто можно считать линейной.

Определим разность давлений $p_1 - p_2$. Очевидно, что при смещении золотника вправо от нейтрального положения $x > 0$ и давление p_1 равно разности давления в аккумуляторе p_0 и перепадов давлений за счет потерь в трубопроводах и отверстиях золотника:

$$p_1 = p_0 - \Delta p_1 - \Delta p_2. \quad (9.11.3)$$

Давление p_2 , при условии равенства нулю давления в резервуаре 4, равно перепаду давления за счет потерь в трубопроводах

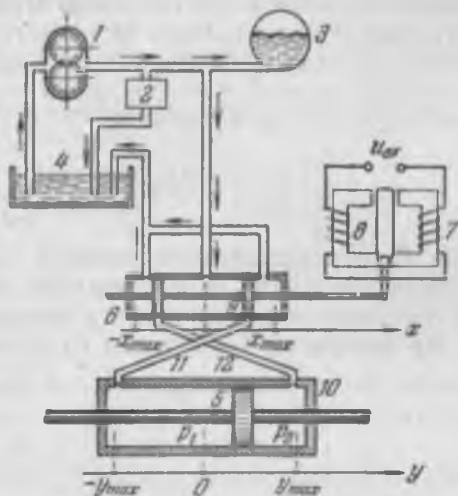


Рис. 9.11.1.

Δp_3 и в отверстиях золотника Δp_4 :

$$p_2 = \Delta p_3 + \Delta p_4. \quad (9.11.4)$$

Как известно из гидродинамики, потери давления для несжимаемой жидкости происходят за счет трения в жидкости и местных гидравлических сопротивлений. При малой длине трубопроводов, что обычно имеет место в гидродвигателях рассматриваемого типа, потери на трение незначительны и их можно не учитывать. Потери давления за счет местных сопротивлений пропорциональны скоростному напору жидкости:

$$\Delta p_2 + \Delta p_3 = \sum_{i=1}^n \zeta_i \frac{\gamma}{2g} w_i^2, \quad (9.11.5)$$

где ζ_i — коэффициенты местных сопротивлений, w_i — скорость потока жидкости, γ — удельный вес жидкости. В соответствии с условием неразрывности жидкости скорость ее истечения w_i в произвольном i -м сечении гидросистемы можно выразить через скорость движения поршня $\dot{y}$ и отношение площадей поршня S и отверстия S_i в i -м сечении:

$$w_i = \frac{S}{S_i} \dot{y}. \quad (9.11.6)$$

Исключая величины w_i из уравнений (9.11.5) и (9.11.6), получим

$$\Delta p_2 + \Delta p_3 = \sum_{i=1}^n \zeta_i \frac{\gamma S^2}{2g S_i^3} \dot{y}^2 = \frac{\gamma S^2 r}{2g} \dot{y}^2, \quad (9.11.7)$$

где $r = \sum_{i=1}^n \zeta_i / S_i$. Потери давления в отверстиях золотника пропорциональны скоростному напору жидкости:

$$\Delta p_1 + \Delta p_4 = 2 \frac{\gamma}{2g} w_s^2, \quad (9.11.8)$$

где w_s — скорость истечения в отверстии золотника, равная

$$w_s = \frac{S}{f \varphi(x)} \dot{y}. \quad (9.11.9)$$

В формуле (9.11.9) f — полная площадь отверстия золотника, $\varphi(x)$ — нелинейная функция, определяющая зависимость относительной площади открытой части отверстия от перемещения золотника. Эта функция, конечно, зависит от формы отверстия. На рис. 9.11.2 приведены графики функции $\varphi(x)$ для случая прямоугольного отверстия и ширины буртика золотника, равной ширине отверстия (рис. 9.11.2, а), и большей ширины буртика (рис. 9.11.2, б), а также для отверстия круглой формы

(рис. 9.11.2, в). Исключая из уравнений (9.11.8) и (9.11.9) скорость истечения w , получим

$$\Delta p_1 + \Delta p_4 = \frac{\gamma S^2}{g f^2 \varphi^2(x)} \dot{y}^2. \quad (9.11.10)$$

Таким образом, разность давлений $p_1 - p_2$ равна

$$p_1 - p_2 = p_0 - \frac{\gamma S^2}{g f^2 \varphi^2(x)} \dot{y}^2 - \frac{\gamma S^2 r}{2g} \dot{y}^2. \quad (9.11.11)$$

При $x < 0$ давления p_1 и p_2 поменяются местами, вследствие чего

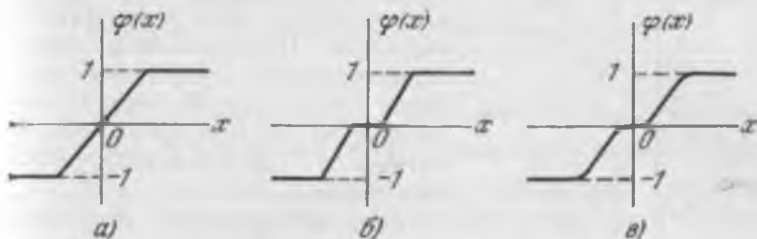


Рис. 9.11.2.

изменится знак разности $p_1 - p_2$. Поэтому можно написать следующую формулу:

$$p_1 - p_2 = \left\{ p_0 - \frac{\gamma S^2}{g} \left[\frac{1}{f^2 \varphi^2(x)} + \frac{r}{2} \right] \dot{y}^2 \right\} \operatorname{sgn} x. \quad (9.11.12)$$

Подставляя выражения (9.11.2) и (9.11.12) в уравнение (9.11.1), получим дифференциальное уравнение, связывающее перемещение золотника x и перемещение поршня гидродвигателя y :

$$m \ddot{y} + c \dot{y} + N \operatorname{sgn} \dot{y} + \frac{\gamma S^2}{g} \left[\frac{1}{f^2 \varphi^2(x)} + \frac{r}{2} \right] \dot{y}^2 \operatorname{sgn} x = S p_0 \operatorname{sgn} x - G. \quad (9.11.13)$$

К уравнению (9.11.13) следует добавить дифференциальное уравнение движения поршня золотника

$$m_1 \ddot{x} = F_m - R_1 \quad (|x| < x_{\max}), \quad (9.11.14)$$

где m_1 — приведенная масса поршня золотника с присоединенными к нему деталями, F_m — сила, развиваемая электромагнитом, R_1 — сила трения, $x_{\max}$ — максимальное по абсолютной величине перемещение поршня золотника, при котором рычаг δ (рис. 9.11.1) доходит до упора. Сила F_m , создаваемая электромагнитом, зависит от тока i_y в его обмотке и перемещения золотника x :

$$F_m = F_m(x, i_y). \quad (9.11.15)$$

В свою очередь ток i_y в обмотке электромагнита связан с приложенным к нему входным напряжением $u_{вх}$ дифференциальным уравнением

$$T \frac{di_y}{dt} + i_y = \frac{1}{R} u_{вх}, \quad T = \frac{L}{R}, \quad (9.11.16)$$

где R и L — сопротивление и индуктивность обмотки электромагнита соответственно.

Система трех уравнений (9.11.13), (9.11.14) и (9.11.16) полностью определяет динамические свойства гидродвигателя. Эти урав-

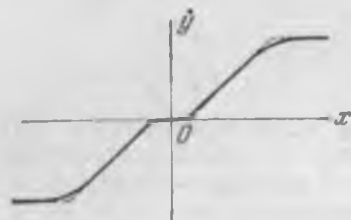


Рис. 9.11.3.

нения показывают, что гидродвигатель является сложной нелинейной системой. Предоставляем читателю самостоятельно представить эту систему структурной схемой, составленной из элементарных линейных звеньев и безынерционных нелинейных звеньев.

Для упрощенного исследования динамических свойств гидродвигателя заметим, что обычно силы инер-

ции, силы вязкого трения и сила нагрузки G малы по сравнению с силой давления $(p_1 - p_2) S$. Поэтому, пренебрегая силами инерции $m\ddot{y}$, силами вязкого трения $c\dot{y}$ и нагрузки G , получим из уравнения (9.11.13) приближенное уравнение

$$N \operatorname{sgn} \dot{y} + \frac{\gamma S^3}{g} \left[\frac{1}{f^2 \varphi^2(x)} + \frac{r}{2} \right] \dot{y}^3 \operatorname{sgn} x = S p_0 \operatorname{sgn} x. \quad (9.11.17)$$

Принимая во внимание, что знак $\dot{y}$ всегда совпадает со знаком x : $\operatorname{sgn} \dot{y} = \operatorname{sgn} x$, и решая уравнение (9.11.17) относительно $\dot{y}$, получим

$$\dot{y} = \frac{f}{S} \varphi(x) \sqrt{\frac{2g(Sp_0 - N)}{\gamma S |2 + rf^2 \varphi^2(x)|}}. \quad (9.11.18)$$

Определяемая этой формулой зависимость скорости движения поршня $\dot{y}$ от перемещения золотника x (сплошная линия на рис. 9.11.3) близка к нелинейности типа ограничения с зоной нечувствительности (пунктирная линия на рис. 9.11.3). Зона нечувствительности возникает вследствие того, что ширина буртика поршня золотника больше ширины отверстия золотника, что обычно всегда делается для обеспечения надежного перекрытия отверстия.

Что касается зависимости перемещения золотника x от входного напряжения $u_{вх}$, то ее можно приближенно представить в форме последовательного соединения идеального запаздывающего звена, безынерционного линейного усилителя и ограничителя.

Таким образом, если объединить ограничение величины x с ограничением величины y и не учитывать нагрузку ($G = 0$), то гидравлический двигатель можно приближенно представить в виде последовательного соединения идеального запаздывающего звена, нелинейного звена с характеристикой типа ограничения с зоной нечувствительности, интегрирующего звена и нелинейного звена типа ограничения. Такая приближенная структурная схема гидравлического двигателя изображена на рис. 9.11.4.



Рис. 9.11.4,

Если учесть в уравнении (9.11.17) еще сопротивление нагрузки G , зависящее от перемещения поршня y , то величина y будет функцией не только x , но и y . Иными словами, характеристика первого нелинейного звена на структурной схеме на рис. 9.11.4 $y = \varphi(x)$ будет зависеть еще от y , как от параметра. В соответствии с этим на структурной схеме гидродвигателя (рис. 9.11.4) добавится обратная связь выхода с первым нелинейным звеном. Эта обратная связь имеет сложный нелинейный характер, так как сигнал обратной связи y не суммируется со входным сигналом x , а изменяет саму форму характеристики звена. Нелинейное звено в этом случае имеет два входа, на один из которых подается входной сигнал x , а на другой — сигнал обратной связи y .

Заметим, что сопротивление нагрузки (т. е. силы сопротивления, действующие на приводимое исполнительным устройством в движение звено объекта управления) всегда можно рассматривать как сигнал обратной связи.

§ 9.12. Пневматические двигатели

В автоматических системах в качестве исполнительных устройств часто используются пневматические двигатели. Принципиальные схемы пневматических двигателей мало отличаются от схем гидравлических двигателей. В качестве распределительных устройств применяются золотниковые и струйные схемы. На рис. 9.12.1 приведена схема пневматического двигателя, имеющего распределительное устройство в виде струйной трубки. Это устройство, отличающееся простотой, малыми габаритами, экономичностью изготовления и высокой надежностью, имеет хорошие динамические свойства. Струйная трубка 1 может перемещаться относительно нейтрального положения, например,

с помощью магнитоэлектрического устройства, рассмотренного в § 3.7, на ограниченный угол. Трубка шарниром 2 соединена с магистралью высокого давления. Воздух или какой-нибудь другой газ, используемый в качестве рабочего тела, проходя через внутреннюю полость трубки и сопло, ударяется о пластинку 3, имеющую два приемных отверстия, которые соединены с поло-

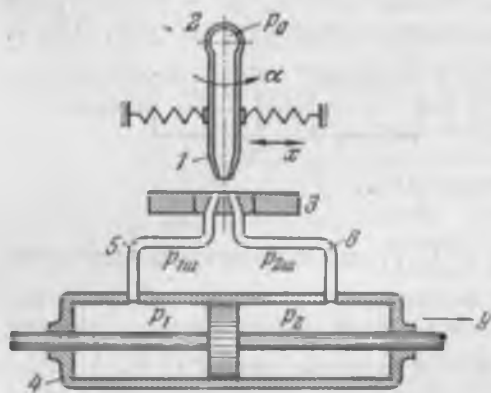


Рис. 9.12.1.

стями силового цилиндра 4. Если сопло расположено симметрично относительно приемных отверстий, то давление в обеих полостях цилиндра одинаково и поршень находится в покое. При смещении трубки относительно отверстий давление в одной полости становится больше, чем в другой, и поршень приходит в движение.

Динамические свойства пневматического двигателя определяются уравнением движения струйной трубки, уравнением дви-

жения поршня и уравнением, определяющим зависимость перепада давлений в полостях цилиндра от угла поворота трубки. Входной переменной пневматического двигателя является напряжение $u_{вх}$ на катушке магнитоэлектрического устройства. За выходную переменную пневматического двигателя примем величину y — перемещение поршня в цилиндре.

Движение струйной трубки описывается уравнением колебательного звена (см. § 3.7):

$$T_1^2 \ddot{\alpha} + 2\zeta T_1 \dot{\alpha} + a = k_1 u_{вх}, \quad (9.12.1)$$

где α — угол поворота трубки, T_1 — постоянная времени, ζ — коэффициент затухания, k_1 — коэффициент усиления.

Движение поршня в цилиндре описывается уравнением

$$m \ddot{y} = (p_1 - p_2) S - R - G \quad (|y| < y_{max}), \quad (9.12.2)$$

аналогичным уравнению движения поршня в гидродвигателе. Основными особенностями работы пневматических двигателей по сравнению с гидравлическими являются сжимаемость рабочего тела и нелинейная зависимость силы трения от скорости движения. Перепад давлений $p = p_1 - p_2$ в цилиндре пневматического двигателя вследствие сжимаемости рабочего тела определяется не

только потерями давления вследствие внутреннего трения, местных сопротивлений и утечек, но и сжимаемостью рабочего тела. Что касается силы трения R , то для пневматических двигателей ее можно представить в виде суммы силы вязкого трения, приблизительно характеризующейся кубической зависимостью от скорости движения поршня, и силы сухого трения:

$$R(\dot{y}) = c_1 \dot{y} + c_2 \dot{y}^3 + N \operatorname{sgn} \dot{y}, \quad (9.12.3)$$

где через N обозначена величина силы сухого трения. Сопротивление нагрузки G зависит от перемещения поршня y и может также зависеть от его скорости $\dot{y}$.

Определим зависимость перепада давлений p от угла поворота α струйной трубки. При отклонении струйной трубки в отводящих штуцерах 5 и 6 (рис. 9.12.1) устанавливаются давления $p_{1ш}$, $p_{2ш}$, линейно зависящие от угла поворота трубки α . Зависимости $p_{1ш}(\alpha)$, $p_{2ш}(\alpha)$ представлены графически на рис. 9.12.2. При отклонении струйной трубки на максимальный угол $\alpha_{\max}$ весь воздух, вытекающий из сопла (без учета потерь), попадает в какой-нибудь один отводящий штуцер, и давление в нем соответствует максимальному давлению p_0 , подводимому из магистрали высокого давления. Минимальное давление равно атмосферному p_a .

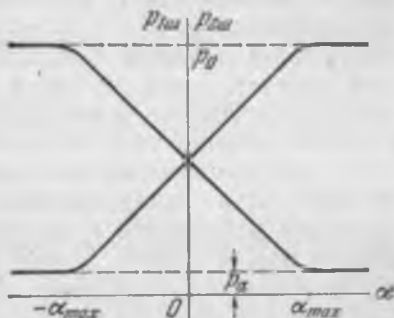


Рис. 9.12.2.

Для упрощения дальнейших выкладок будем считать, что атмосферное давление равно нулю ($p_a = 0$). Это допущение равноценно тому, что за начало отсчета давлений принимается атмосферное давление. В соответствии с зависимостями $p_{1ш}(\alpha)$ и $p_{2ш}(\alpha)$ и сделанным допущением ($p_a = 0$) запишем соотношения (рис. 9.12.2)

$$p_{1ш} + p_{2ш} = \zeta_1 p_0, \quad (9.12.4)$$

$$p_{1ш} - p_{2ш} = \zeta_1 p_0 \Phi(\varepsilon), \quad (9.12.5)$$

где

$$\Phi(\varepsilon) = \begin{cases} -1 & \text{при } \alpha \leq -\alpha_{\max}, \\ \varepsilon & \text{при } |\alpha| < \alpha_{\max}, \\ 1 & \text{при } \alpha > \alpha_{\max}, \end{cases} \quad (9.12.6)$$

$\varepsilon = \alpha/\alpha_{\max}$ — относительный угол поворота струйной трубки, ζ_1 — коэффициент, учитывающий потери давления в сопле и от-

водящих штуцерах. Эти потери в зависимости от температуры окружающей среды составляют от 5 до 25%. Следовательно, коэффициент ζ_1 имеет значение $\zeta_1 = 0,95 \div 0,75$. Давления $p_{1ш}$ и $p_{2ш}$ в отводящих штуцерах и давления p_1, p_2 в полостях цилиндра различаются незначительно. Поэтому весовые расходы рабочего тела в отводящих штуцерах с достаточной точностью можно определить линейными соотношениями

$$G_{1ш} = k_2 (p_{1ш} - p_1), \quad G_{2ш} = k_2 (p_{2ш} - p_2), \quad (9.12.7)$$

где k_2 — коэффициент пропорциональности. Весовые расходы в полостях цилиндра выражаются формулами

$$G_1 = \frac{d(\gamma_1 V_1)}{dt}, \quad G_2 = \frac{d(\gamma_2 V_2)}{dt}, \quad (9.12.8)$$

где γ_1, γ_2 — удельные веса рабочего тела в полостях цилиндра, $V_1 = S(y_{\max} + y)$, $V_2 = S(y_{\max} - y)$ — объемы полостей, $y_{\max}$ — максимальное перемещение поршня цилиндра от среднего положения, S — площадь поршня. Выразим удельные веса γ_1, γ_2 через давления, принимая, что процессы сжатия и разрежения рабочего тела являются изотермическими. Фактически эти процессы являются политропическими, больше приближаясь при быстром наполнении и истечении рабочего тела к адиабатическим, чем к изотермическим (которые характерны для медленных процессов). Для быстродействующих пневматических двигателей процессы наполнения и истечения рабочего тела с высокой точностью можно считать адиабатическими. Однако описание работы пневматического двигателя при адиабатическом процессе значительно сложнее, чем при изотермическом. Как показывают эксперименты и расчеты, вполне допустимо считать, что процессы в пневматическом двигателе являются изотермическими. Так, расчеты показывают, что при сделанном допущении расчетное время нарастания давления в полостях цилиндра увеличивается на $10 \div 20\%$ по сравнению с истинным [6]. Это можно учесть в дальнейшем введением поправочного коэффициента. Как известно, для изотермического процесса имеет место соотношение

$$\frac{p_1}{\gamma_1} = \frac{p_2}{\gamma_2} = k_3.$$

Подставляя выражения удельных весов γ_1, γ_2 и объемов V_1, V_2 в уравнения (9.12.7), получим

$$\left. \begin{aligned} G_1 &= \frac{S}{k_3} [\dot{p}_1 (y_{\max} + y) + p_1 \dot{y}], \\ G_2 &= \frac{S}{k_3} [\dot{p}_2 (y_{\max} - y) - p_2 \dot{y}]. \end{aligned} \right\} \quad (9.12.9)$$

Условие сохранения весовых расходов в отводящих штуцерах и полостях пневмоцилиндра позволяет записать следующие

равенства:

$$G_{1\text{ш}} = G_1, \quad G_{2\text{ш}} = G_2. \quad (9.12.10)$$

Подставим сюда выражения весовых расходов из (9.12.7) и (9.12.9). Тогда получим

$$\left. \begin{aligned} k_2 k_3 (p_{1\text{ш}} - p_1) &= S [\dot{p}_1 (y_{\text{max}} + y) + p_1 \dot{y}], \\ k_2 k_3 (p_{2\text{ш}} - p_2) &= S [\dot{p}_2 (y_{\text{max}} - y) - p_2 \dot{y}]. \end{aligned} \right\} \quad (9.12.11)$$

Вычитая из первого уравнения (9.12.11) второе, получим

$$\begin{aligned} k(p_{1\text{ш}} - p_{2\text{ш}}) - k(p_1 - p_2) &= \\ &= S y_{\text{max}} (\dot{p}_1 - \dot{p}_2) + S y (\dot{p}_1 + \dot{p}_2) + S \dot{y} (p_1 + p_2), \end{aligned} \quad (9.12.12)$$

где $k = k_2 k_3$. Подставляя в (9.12.12) выражение разности давлений в отводящих штуцерах и полагая $p = p_1 - p_2$, получим

$$S y_{\text{max}} \dot{p} + k p = k \zeta_1 p_0 \Phi(\varepsilon) - S \frac{d}{dt} [y (p_1 + p_2)]. \quad (9.12.13)$$

Экспериментальные данные показывают, что сумма давлений в полостях цилиндра является величиной постоянной и пропорциональной давлению p_0 :

$$p_1 + p_2 = \zeta_2 p_0, \quad (9.12.14)$$

где ζ_2 — коэффициент, учитывающий потери давления. Обычно величина коэффициента ζ_2 имеет порядок $\zeta_2 = 0,90 \div 0,75$. Поэтому с достаточной точностью можно считать $\zeta_1 \approx \zeta_2$. С учетом соотношения (9.12.14) уравнение (9.12.13) принимает вид

$$T \dot{p} + p = \zeta_1 p_0 \left[\Phi(\varepsilon) - \frac{S \dot{y}}{k} \right], \quad (9.12.15)$$

где $T = S y_{\text{max}} / k$ — постоянная времени процесса изменения перепада давления.

Уравнения (9.12.1), (9.12.2) и (9.12.15), которые можно переписать в виде

$$\left. \begin{aligned} T_1 \ddot{\alpha} + 2 \zeta T_1 \dot{\alpha} + \alpha &= k_1 u_{\text{вх}}, \\ T \dot{p} + p &= \zeta_1 p_0 \left[\Phi(\varepsilon) - \frac{S \dot{y}}{k} \right], \quad \varepsilon = \frac{\alpha}{\alpha_{\text{max}}}, \\ m \ddot{y} &= S p - R - G \quad (|y| < y_{\text{max}}), \end{aligned} \right\} \quad (9.12.16)$$

описывают работу пневматического двигателя. Из этих уравнений видно, что пневматический двигатель, так же как и гидравлический, является сложной нелинейной системой. Рекомендуем читателю самостоятельно составить структурную схему этой системы.

Пневматические двигатели обладают большим быстродействием. Их инерционность очень мала по сравнению с инерционностью объектов управления. Вследствие этого при теоретическом исследовании систем управления в большинстве случаев можно ограничиться грубо приближенным описанием динамических свойств пневматических двигателей или даже считать их безынерционными, мгновенно отрабатывающими задаваемое системой управления положение управляющего элемента объекта управления (например, руля самолета или ракеты).

Для приближенного описания динамических свойств пневматического двигателя можно, так же как и в случае гидравлического двигателя, пренебречь силами инерции $T_1 \ddot{\alpha}$ и $m \ddot{y}$ и силами вязкого трения $2\zeta T_1 \dot{\alpha}$ и $c_1 \dot{y} + c_2 \dot{y}^3$ в уравнениях движения магнитоэлектрического элемента и поршня двигателя. Кроме того, пренебрежем инерционностью образования перепада давления p , т. е. постоянной времени T во втором уравнении (9.12.16), и сопротивлением нагрузки G . Тогда в выражении (9.12.3) силы трения R останется только один последний член и уравнения (9.12.16) заменятся упрощенными уравнениями

$$\alpha = k_1 u_{\text{вх}}, \quad p = \zeta_1 p_0 \left[\varphi(\varepsilon) - \frac{S \dot{y}}{k} \right], \quad \varepsilon = \frac{\alpha}{\alpha_{\text{max}}}, \quad Sp - N \operatorname{sgn} \dot{y} = 0.$$

Исключая из этих уравнений α , ε и p и полагая $\kappa = k_1/\alpha_{\text{max}}$, получим

$$\frac{N}{\zeta_1 S p_0} \operatorname{sgn} \dot{y} = \varphi(\kappa u_{\text{вх}}) - \frac{S \dot{y}}{k}. \quad (9.12.17)$$

Отсюда видно, что поршень двигателя находится в покое при всех значениях входного сигнала $u_{\text{вх}}$, при которых $|\varphi(\kappa u_{\text{вх}})| < N/\zeta_1 S p_0$, т. е. при $|u_{\text{вх}}| < N/\kappa \zeta_1 S p_0$. При $u_{\text{вх}} > N/\kappa \zeta_1 S p_0$ скорость поршня $\dot{y}$ положительна ($\operatorname{sgn} \dot{y} = 1$), а при $u_{\text{вх}} < -N/\kappa \zeta_1 S p_0$ скорость поршня $\dot{y}$ отрицательна ($\operatorname{sgn} \dot{y} = -1$). Таким образом, уравнение (9.12.17) дает следующую зависимость скорости поршня двигателя от входного сигнала:

$$\dot{y} = \begin{cases} \frac{k}{S} \left[\varphi(\kappa u_{\text{вх}}) + \frac{N}{\zeta_1 S p_0} \right] & \text{при } u_{\text{вх}} < -\frac{N}{\kappa \zeta_1 S p_0}, \\ 0 & \text{при } -\frac{N}{\kappa \zeta_1 S p_0} < u_{\text{вх}} < \frac{N}{\kappa \zeta_1 S p_0}, \\ \frac{k}{S} \left[\varphi(\kappa u_{\text{вх}}) - \frac{N}{\zeta_1 S p_0} \right] & \text{при } u_{\text{вх}} > \frac{N}{\kappa \zeta_1 S p_0}. \end{cases} \quad (9.12.18)$$

Эта зависимость представляет собой нелинейность типа ограничителя с зоной нечувствительности (приложение 5, нелинейность 3, рис. 9.11.3).

Так как максимальное значение функции φ равно 1, то из (9.12.18) вытекает следующее выражение для максимальной скорости поршня:

$$\dot{y}_{\max} = \frac{k}{S} \left(1 - \frac{N}{\zeta_1 S p_0} \right). \quad (9.12.19)$$

Отсюда можно выразить величину k/S через $\dot{y}_{\max}$. Тогда формула (9.12.18) примет вид

$$\dot{y} = \begin{cases} \dot{y}_{\max} \frac{\zeta_1 S p_0 \varphi(\kappa u_{\text{вх}}) - N}{\zeta_1 S p_0 - N} & \text{при } u_{\text{вх}} < -\frac{N}{\kappa \zeta_1 S p_0}, \\ 0 & \text{при } -\frac{N}{\kappa \zeta_1 S p_0} < u_{\text{вх}} < \frac{N}{\kappa \zeta_1 S p_0}, \\ \dot{y}_{\max} \frac{\zeta_1 S p_0 \varphi(\kappa u_{\text{вх}}) - N}{\zeta_1 S p_0 - N} & \text{при } u_{\text{вх}} > \frac{N}{\kappa \zeta_1 S p_0}. \end{cases} \quad (9.12.20)$$

Чтобы приближенно учесть запаздывание магнитоэлектрического устройства и образования перепада давления, обычно в формулах (9.12.18) и (9.12.20) берут запаздывающее значение входного сигнала $u_{\text{вх}}$, т. е. считают, что эти формулы определяют значение скорости поршня $y'(t)$ по значению входного сигнала $u_{\text{вх}}(t - \tau)$. Запаздывание τ для современных пневматических двигателей имеет порядок $0,01 \div 0,02$ с.

На основании изложенного пневматический двигатель можно приближенно представить той же структурной схемой (рис. 9.11.4), что и гидравлический двигатель.

Учет сопротивления нагрузки G приводит, так же как и в случае гидравлического двигателя, к изменению скоростной характеристики двигателя, т. е. характеристики первого нелинейного звена на рис. 9.11.4. Это равноценно добавлению на структурной схеме на рис. 9.11.4 обратной связи выхода с первым нелинейным звеном, а в случае зависимости G и от $\dot{y}$ — еще дополнительному замыканию скоростной характеристики соответствующей нелинейной обратной связью.

§ 9.13. Характерные нелинейности объектов управления

Большинство объектов управления лишь приближенно могут рассматриваться как линейные. Во многих случаях при исследовании автоматических систем приходится учитывать нелинейности как в элементах системы управления, так и в объекте управления. В предыдущих параграфах были выяснены основные нелинейности исполнительных устройств. Эти нелинейности приводят к нелинейности объекта управления в случае, когда объект управления связан жесткой механической передачей с исполнительным устройством, как, например, в случае подвижной пупечной установки на самолете.

В § 3.16 было показано, что самолет или крылатую ракету можно приближенно считать линейным объектом управления. Однако в некоторых задачах приходится учитывать нелинейности самолета или ракеты. Основной существенной нелинейностью самолета или ракеты является ограничение отклонения рулей. Эта нелинейность, имеющая характер идеального ограничения, определяется зависимостью истинного отклонения руля δ от теоретического отклонения, задаваемого системой управления δ_a (рис. 9.13.1). Учет этой нелинейности особенно необходим в случае высокого

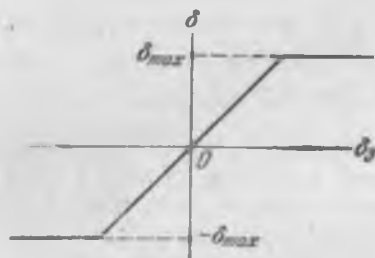


Рис. 9.13.1.

уровня помех в сигнале управления. В этом случае задаваемое системой управления отклонение руля δ_a содержит случайные колебания большой амплитуды, которые вследствие изложенного в § 8.5 будут подавлять полезный сигнал на нелинейности и тем самым снижать эффективность управления, т. е. как бы уменьшать фактическое эффективное значение коэффициентов a_δ и c_δ в уравнениях (3.16.8) и (3.16.10). Кроме этой существенной нелинейности,

в уравнениях движения самолета (ракеты) имеются и другие нелинейности. Так, например, угол наклона вектора скорости самолета к горизонту θ входит в уравнения движения (3.16.2), (3.16.8) и в кинематические уравнения (3.16.3) нелинейно. Кроме того, аэродинамические коэффициенты c_y и m_z лишь приближенно могут быть выражены линейными функциями (3.16.6), (3.16.7). В действительности, как показывают результаты экспериментальных исследований в аэродинамических трубах, c_y и m_z являются сложными нелинейными функциями угла атаки α и угла отклонения руля высоты δ , а в случае пространственного движения — и переменных, определяющих поперечное движение самолета, перпендикулярное к вертикальной плоскости. Конечно, и в этом случае можно выразить аэродинамические коэффициенты c_y и m_z формулами вида (3.16.6) и (3.16.7), но только коэффициенты c_y^α , m_z^α , m_z^δ и $m_z^{\delta\delta}$ следует при этом считать функциями угла атаки α и угла отклонения руля высоты δ . Таким образом, движение самолета или ракеты с учетом нелинейности также описывается уравнениями (3.16.3) и (3.16.8), только коэффициенты A_α , a_δ , a_α и a_δ в них являются функциями α и δ . Кроме того, к этим уравнениям добавляется условие $|\delta| \leq \delta_{\max}$.

УСТОЙЧИВОСТЬ И АВТОКОЛЕБАНИЯ НЕЛИНЕЙНЫХ СИСТЕМ

§ 10.1. Общее определение устойчивости

Данное в § 6.5 определение устойчивости по Ляпунову применимо как к линейным, так и к нелинейным системам, поведение которых описывается дифференциальными уравнениями. Чтобы формулировать это определение математически, положим, что система дифференциальных уравнений, описывающая поведение исследуемой системы, приведена к нормальной форме Коши, т. е. к системе уравнений первого порядка, решенных относительно производных:

$$\dot{y}_k = \varphi_k(t, y_1, \dots, y_n) \quad (k = 1, \dots, n), \quad (10.1.1)$$

где $y_1, \dots, y_n$ — величины, определяющие состояние системы в каждый данный момент времени. Одна или несколько из этих величин могут быть выходными переменными системы, а остальные представляют собой вспомогательные переменные, которые необходимо ввести для приведения уравнений к виду (10.1.1). Правые части уравнений (10.1.1) представляют собой заданные функции времени t и величин $y_1, \dots, y_n$.

Если правые части уравнений (10.1.1) не зависят явно от времени t , то это означает, что на систему не действуют никакие внешние возмущения, представляющие собой функции времени, а все действующие силы (конечно, обобщенные) зависят только от состояния системы, т. е. являются внутренними. Поэтому системы, описываемые уравнениями вида (10.1.1), с правыми частями, зависящими только от величин $y_1, \dots, y_n$ и не зависящими явно от времени t , называются *автономными*.

Если в правых частях уравнений (10.1.1) явно зависят от времени лишь внешние возмущения, то система, описываемая уравнениями (10.1.1), стационарна. В общем случае, когда функциями времени в уравнениях (10.1.1) являются не только внешние возмущения, но и некоторые параметры системы, эта система нестационарна. Очевидно, что любая автономная система стационарна, но не всякая стационарная система автономна. Стационарная система автономна лишь в том случае, когда на нее не действуют никакие внешние (входные) возмущения.

Каждой совокупности начальных условий $t_0, y_{10}, \dots, y_{n0}$, не являющейся особой точкой правых частей уравнений (10.1.1),

соответствует вполне определенное движение системы, определяемое решением уравнений (10.1.1).

Выберем из всех возможных движений системы какое-либо одно и условимся называть его *невозмущенным* по терминологии А. М. Ляпунова [42]. Этому движению соответствует вполне определенная совокупность начальных условий $t_0, y_1^{(0)}, \dots, y_n^{(0)}$. Роль невозмущенного движения может играть всякое возможное движение системы, и выбор движения, которое считается невозмущенным, произволен. Решение уравнений (10.1.1), соответствующее невозмущенному движению, будем обозначать $y_1^{(0)}, \dots, y_n^{(0)}$.

Всякое другое движение, отличное от невозмущенного, условимся называть *возмущенным*. Возмущенное и невозмущенное движения являются возможными движениями одной и той же системы и определяются одними и теми же дифференциальными уравнениями (10.1.1), только при различных начальных условиях.

Данное в § 6.5 определение устойчивости по Ляпунову может быть математически сформулировано следующим образом. Невозмущенное движение называется *устойчивым* по отношению к величинам $y_1, \dots, y_n$, если при любом сколь угодно малом $\varepsilon > 0$ существует такое положительное число $\delta = \delta(\varepsilon)$, что все возмущенные движения, для которых в любой начальный момент t_0

$$|y_k - y_k^{(0)}| < \delta \quad (k = 1, \dots, n), \quad (10.1.2)$$

при любом $t > t_0$ удовлетворяют условию

$$|y_k - y_k^{(0)}| < \varepsilon \quad (k = 1, \dots, n). \quad (10.1.3)$$

Невозмущенное движение называется *асимптотически устойчивым*, если существует такое положительное число δ , что все возмущенные движения, для которых в любой начальный момент t_0 выполнены неравенства (10.1.2), асимптотически стремятся к невозмущенному движению при $t \rightarrow \infty$:

$$y_k \rightarrow y_k^{(0)} \quad \text{при} \quad t \rightarrow \infty \quad (k = 1, \dots, n). \quad (10.1.4)$$

Если существует хотя бы одно такое число $\varepsilon > 0$, что при любом сколь угодно малом $\delta > 0$ существует возмущенное движение, удовлетворяющее условию (10.1.2), для которого при некоторых значениях $t > t_0$ неравенства (10.1.3) не выполняются, то невозмущенное движение называется *неустойчивым*.

Этим определениям можно дать геометрическую интерпретацию. Для этого будем изображать отклонение возмущенного движения системы от невозмущенного в каждый данный момент времени t точкой n -мерного пространства с прямоугольными

декартовыми координатами *)

$$\eta_1 = y_1 - y_1^{(0)}, \quad \eta_2 = y_2 - y_2^{(0)}, \quad \dots, \quad \eta_n = y_n - y_n^{(0)}.$$

Всякое пространство, в котором декартовыми координатами точки являются величины, определяющие мгновенное состояние некоторой системы, называется *фазовым пространством* этой системы. Точка фазового пространства, соответствующая состоянию системы в данный момент t , называется *изображающей точкой*.



Рис. 10.1.1.

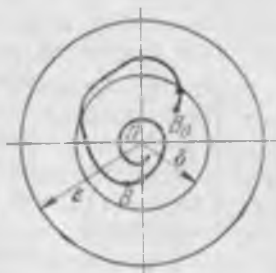


Рис. 10.1.2.

Изменению состояния системы со временем соответствует движение изображающей точки в фазовом пространстве. Любая траектория, по которой может двигаться изображающая точка при изменении состояния системы, называется *фазовой траекторией*. Изучение фазовых траекторий, как мы увидим в §§ 10.4 и 10.5, во многих случаях значительно облегчает исследование свойств автономных нелинейных систем.

Очевидно, что в рассматриваемом фазовом пространстве невозмущенному движению соответствует состояние покоя изображающей точки в начале координат. Невозмущенное движение устойчиво, если все возможные фазовые траектории, начинающиеся внутри сферы C_δ достаточно малого радиуса δ , не выходят из сферы C_ϵ сколь угодно малого наперед заданного радиуса ϵ (рис. 10.1.1). Невозмущенное движение асимптотически устойчиво, если все возможные фазовые траектории, начинающиеся внутри сферы C_δ достаточно малого радиуса δ , стремятся к началу координат при $t \rightarrow \infty$ (рис. 10.1.2). Невозмущенное движение неустой-

*) В случае $n = 2$ это пространство является плоскостью, а в случае $n = 3$ — обычным трехмерным пространством. Поэтому при $n = 2$ и $n = 3$ всем вводимым ниже понятиям можно поставить в соответствие наглядные геометрические образы. При $n > 3$ таких наглядных образов не существует. Однако и в этом случае геометрическая терминология удобна тем, что она дает возможность по аналогии с двумерным и трехмерным пространствами легче осваивать вводимые понятия и иллюстрировать их в случае необходимости наглядными образами на плоскости или в трехмерном пространстве.

чно, если существуют фазовые траектории, начинающиеся как угодно близко к началу координат, выходящие при некоторых t за пределы некоторой фиксированной сферы C_ε (рис. 10.1.3).

Иными словами, невозмущенное движение устойчиво, если при любом как угодно малом радиусе ε сферы C_ε существует такая сфера C_δ радиуса δ , что изображающая точка B не может выйти за пределы сферы C_ε , начав движение из любой точки B_0 сферы C_δ . Невозмущенное движение асимптотически устойчиво, если существует такая сфера C_δ , что изображающая точка стремится к началу координат, начав движение из любой точки B_0 сферы C_δ . Невозмущенное движение неустойчиво, если существует такая сфера C_ε и как угодно близко к началу координат можно найти такую точку B_0 , что изображающая точка выйдет за пределы сферы C_ε , начав движение в точке B_0 .



Рис. 10.1.3.

При исследовании нелинейных систем обычно принимают за координаты изображающей точки в фазовом пространстве сами величины $y_1, \dots, y_n$, определяющие состояние системы в любой момент времени t . Приведенная геометрическая интерпретация определений устойчивости и неустойчивости остается

справедливой и в этом случае, если вместо сфер C_ε и C_δ с центром в начале координат рассматривать сферы с центром в подвижной точке, изображающей состояние системы в невозмущенном движении.

В отличие от линейных систем, одна и та же нелинейная система может быть устойчивой в одних режимах и неустойчивой в других режимах. Иными словами, одни движения или состояния равновесия нелинейной системы могут быть устойчивыми, а другие движения или состояния равновесия той же самой системы могут быть неустойчивыми. Поэтому нельзя говорить об устойчивости или неустойчивости нелинейной системы вообще, а можно говорить только об устойчивости или неустойчивости различных режимов движения (работы) нелинейной системы. Именно поэтому, давая определение устойчивости, мы отнесли это понятие не к системе, а к некоторому движению системы, которое мы назвали невозмущенным. В качестве такого невозмущенного движения естественно выбирать некоторое требуемое движение или состояние равновесия системы, в котором она должна находиться при нормальном функционировании и от которого ее отклоняют различные неизбежные в любой системе возмущения, в том числе и случайные.

В определении устойчивости по Ляпунову невозмущенное и возмущенное движения рассматриваются всегда при одних и тех же правых частях уравнений (10.1.1), а следовательно, при

одних и тех же входных сигналах системы, представляющих собой вполне определенные заданные функции времени. Поэтому одна и та же система при одних и тех же начальных условиях невозмущенного движения $t_0, y_1^{(0)}, \dots, y_n^{(0)}$ может быть устойчивой по Ляпунову при одних входных сигналах и неустойчивой при других входных сигналах. Между тем в задачах практики приходится оценивать устойчивость работы системы при различных входных сигналах, заранее неизвестных. В таких случаях можно взять невозмущенное движение системы при каких-нибудь вполне определенных входных сигналах, а сравнивать с ним придется все возможные возмущенные движения, соответствующие всем возможным достаточно малым изменениям входных сигналов и начальным отклонениям. При этом функции F_k в уравнениях (10.1.1) будут различными для всех возмущенных движений и будут отличаться от функций F_k , соответствующих невозмущенному движению, т. е. каждое возмущенное движение будет определяться своими дифференциальными уравнениями, отличными от дифференциальных уравнений невозмущенного движения. Это обстоятельство является одной из причин того, что определение устойчивости по Ляпунову недостаточно для практики. Именно поэтому в главе 6 в основу теории устойчивости линейных систем было положено другое определение устойчивости, данное в § 6.1. Это же определение устойчивости можно с соответствующими видоизменениями перенести на нелинейные системы.

Предположим, что при изменении формы входных сигналов правые части дифференциальных уравнений (10.1.1) изменяются непрерывно, т. е. что при замене входных сигналов достаточно близкими к ним другими входными сигналами функции F_k в уравнениях (10.1.1) изменяются сколь угодно мало. В этом случае дифференциальные уравнения возмущенных движений можно написать в виде

$$\dot{y}_k = \Phi_k(t, y_1, \dots, y_n) + F_k(t, y_1, \dots, y_n) \quad (k=1, \dots, n), \quad (10.1.5)$$

где F_k — функции, сколь угодно малые по абсолютной величине при любых достаточно малых отклонениях входных сигналов системы от входных сигналов, соответствующих невозмущенному движению. Различным входным сигналам системы соответствуют различные функции F_k в уравнениях (10.1.5). В частности, для входных сигналов, при которых рассматривается невозмущенное движение системы, все функции F_k тождественно равны нулю. Рассматривая возмущенные движения, соответствующие всем возможным входным сигналам, мы должны оставить функции F_k неопределенными и дать определение устойчивости, независимое от вида функций F_k .

Видоизменяя определение § 6.1, мы назовем невозмущенное движение системы *устойчивым*, если отклонение ее возмущенного

движения от невозмущенного остается сколь угодно малым при любых достаточно малых начальных отклонениях и при любых входных сигналах, достаточно близких к входным сигналам, соответствующим невозмущенному движению. Иными словами, невозмущенное движение системы называется устойчивым, если при любом сколь угодно малом $\varepsilon > 0$ существуют такие числа $\delta = \delta(\varepsilon)$ и $\delta' = \delta'(\varepsilon)$, что для всех возмущенных движений, удовлетворяющих условию (10.1.2), при любом $t > t_0$ выполняются неравенства (10.1.3), если

$$|F_k(t, y_1^{(0)}, \dots, y_n^{(0)})| < \delta' \quad (k = 1, \dots, n). \quad (10.1.6)$$

Можно доказать, что при довольно общих условиях невозмущенное движение системы устойчиво в смысле данного определения, если оно асимптотически устойчиво по Ляпунову. Это утверждение справедливо, в частности, для любых установившихся и периодических невозмущенных движений стационарных систем *) [74]. В главе 6 мы видели, что это утверждение справедливо для любой стационарной линейной системы, поведение которой описывается дифференциальным уравнением.

Таким образом, исследование устойчивости нелинейной системы практически сводится к исследованию ее асимптотической устойчивости по Ляпунову. Этим объясняется большое практическое значение теории устойчивости движения, созданной А. М. Ляпуновым.

§ 10.2. Исследование устойчивости нелинейных систем методом линеаризации (первый метод Ляпунова)

А. М. Ляпунов создал два общих метода исследования устойчивости нелинейных систем. Первый метод Ляпунова основан на линеаризации уравнений, описывающих поведение системы. Обычно уравнения системы линеаризуются относительно отклонений элементов возмущенного движения системы от соответствующих элементов невозмущенного движения, т. е. относительно разностей $\eta_l = y_l - y_l^{(0)}$ ($l = 1, \dots, n$). Линеаризация уравнений осуществляется методом § 8.3, т. е. разложением всех входящих в уравнения функций в ряд Тейлора и отбрасыванием всех членов выше первой степени относительно величин $\eta_1, \dots, \eta_n$:

$$\varphi_k(t, y_1, \dots, y_n) \approx \varphi_k(t, y_1^{(0)}, \dots, y_n^{(0)}) + \sum_{l=1}^n \frac{\partial \varphi_k(t, y_1^{(0)}, \dots, y_n^{(0)})}{\partial y_l^{(0)}} \eta_l \quad (k = 1, \dots, n). \quad (10.2.1)$$

*) Определение установившегося движения будет дано в следующем параграфе.

Подставляя эти выражения в уравнения (10.1.1) и принимая во внимание, что невозмущенное движение удовлетворяет системе уравнений (10.1.1), вследствие чего

$$\dot{y}_k^{(0)} = \varphi_k(t, y_1^{(0)}, \dots, y_n^{(0)}) \quad (k = 1, \dots, n), \quad (10.2.2)$$

получим следующую систему линейных дифференциальных уравнений для $\eta_1, \dots, \eta_n$:

$$\dot{\eta}_k = \sum_{l=1}^n a_{kl} \eta_l \quad (k = 1, \dots, n), \quad (10.2.3)$$

где коэффициенты a_{kl} определяются формулой

$$a_{kl} = \frac{\partial \varphi_k(t, y_1^{(0)}, \dots, y_n^{(0)})}{\partial y_l^{(0)}} \quad (k, l = 1, \dots, n). \quad (10.2.4)$$

Величины $\eta_1, \dots, \eta_n$, определяющие отклонение возмущенного движения от невозмущенного, обычно называются *вариациями* переменных $y_1, \dots, y_n$, определяющих состояние системы. Приближенные линейные уравнения (10.2.3) называются *уравнениями в вариациях* или *линейными уравнениями первого приближения*.

При весьма общих условиях А. М. Ляпунов доказал следующие теоремы, которые мы сформулируем без доказательства:

1. Невозмущенное движение нелинейной системы асимптотически устойчиво, если определяемая соответствующими уравнениями в вариациях линейная система асимптотически устойчива.

2. Невозмущенное движение нелинейной системы неустойчиво, если определяемая соответствующими уравнениями в вариациях линейная система неустойчива.

Эти теоремы Ляпунова дают возможность исследовать устойчивость различных режимов работы нелинейных систем методами теории устойчивости линейных систем.

Если все коэффициенты a_{kl} системы уравнений в вариациях постоянны, то невозмущенное движение называется *установившимся*. Для исследования устойчивости установившихся движений нелинейных систем можно применить изложенные в §§ 6.2 и 6.3 методы исследования устойчивости стационарных линейных систем.

В случае периодического невозмущенного движения (колебаний) автономной нелинейной системы коэффициенты a_{kl} уравнений в вариациях являются периодическими функциями времени. Методы исследования устойчивости систем в случае, когда уравнения в вариациях являются уравнениями с периодическими коэффициентами, также разработаны Ляпуновым [42] (см. также [74]).

В общем случае, когда коэффициенты уравнений в вариациях являются любыми функциями времени, для исследования устойчивости системы можно применить общий критерий § 6.1. Для

этого следует определить весовые функции линейной системы, определяемой уравнениями в вариациях, соответствующие всем входам системы и всем переменным $y_1, \dots, y_n$, и вычислить для них интегралы (6.1.1). Все эти вычисления можно выполнить с помощью моделирования соответствующей сопряженной системы. Выполнив эти вычисления для различных моментов времени t , можно судить об устойчивости системы. Если интегралы (6.1.1) для весовых функций, соответствующих всем входам системы и всем переменным $y_1, \dots, y_n$, перестают заметно изменяться с увеличением t , начиная с некоторого момента, то исследуемая нелинейная система устойчива по отношению к переменным $y_1, \dots, y_n$. Если же хотя бы один из этих интегралов неограниченно возрастает с увеличением t , то система неустойчива. Заметим, что этот способ исследования устойчивости является чисто практическим и не дает возможности делать строгие выводы об устойчивости и неустойчивости вследствие невозможности осуществить с помощью моделирующих устройств бесконечные процессы, связанные с переходом к пределу в формуле (6.1.1). Однако для практики такой метод исследования устойчивости вполне достаточен и дает возможность безошибочно судить об устойчивости движения в течение любых интересующих нас конечных отрезков времени.

Для исследования устойчивости можно также применить непосредственное моделирование системы или линейной системы, определяемой уравнениями в вариациях, при различных начальных условиях.

Для приближенного исследования устойчивости нелинейных систем методом линеаризации в случае произвольных переменных коэффициентов a_k , уравнений в вариациях (10.2.3) можно также применить прием «замораживания» коэффициентов. Этот прием состоит в том, что переменные коэффициенты уравнений в вариациях заменяются их постоянными значениями, соответствующими какому-нибудь выбранному моменту времени t_1 , и исследуется устойчивость стационарной линейной системы, определяемой полученными уравнениями с постоянными коэффициентами. Это исследование производится для ряда характерных моментов времени $t_1, t_2, \dots$ работы системы. Если для всех выбранных моментов времени стационарные линейные системы, полученные «замораживанием» коэффициентов уравнений в вариациях, устойчивы, то исследуемую нестационарную систему можно практически считать устойчивой. При этом в случае определенного конечного времени работы системы, например в случае работы системы управления полетом снаряда, когда в момент попадания снаряда в цель или его разрыва работа системы управления, естественно, прекращается, систему можно считать практически устойчивой и в том случае, если система, полученная «замораживанием» коэффициентов

тов уравнений в вариациях, становится неустойчивой, когда до момента конца работы системы остается время, значительно меньшее времени переходного процесса системы.

Заметим, что с практической точки зрения совершенно не обязательно выбирать в качестве центра разложения входящих в уравнения функций исследуемое невозмущенное движение системы, т. е. интеграл уравнений (10.1.1), описывающих поведение системы. Можно в качестве центра разложения взять любые функции $y_1^{(0)}, \dots, y_n^{(0)}$, не удовлетворяющие уравнениям (10.1.1), но близкие к интегралу уравнений (10.1.1), соответствующему невозмущенному движению системы. Так, например, если некоторые из переменных $y_1, \dots, y_n$, скажем $y_1, \dots, y_m$, являются по физическому смыслу малыми величинами при устойчивой работе системы, то в качестве функций $y_1^{(0)}, \dots, y_m^{(0)}$ можно взять нули, а остальные функции $y_{m+1}^{(0)}, \dots, y_n^{(0)}$ определить путем интегрирования соответствующей укороченной системы уравнений, полученной из (10.1.1) путем замены нулями всех малых величин.

В заключение заметим, что изложенный метод исследования устойчивости нелинейных систем, основанный на линеаризации уравнений, применим лишь к системам, содержащим только элементарные нелинейные звенья с непрерывными гладкими характеристиками. К системам с существенно нелинейными звеньями этот метод неприменим. В таких случаях часто оказывается полезным второй метод Ляпунова, который мы изложим в следующем параграфе. Для более подробного изучения теории устойчивости рекомендуем читателю обратиться к специальной литературе (см., например, [42, 43], а также [39, 41, 74]).

§ 10.3. Прямой метод исследования устойчивости нелинейных систем (второй метод Ляпунова)

Второй метод Ляпунова основан на обобщении и развитии элементарных представлений, связанных с положениями равновесия материальной точки в консервативном силовом поле. Из теоретической механики известно, что точки минимума потенциальной энергии являются положениями устойчивого равновесия материальной точки, а точки максимума потенциальной энергии — положениями неустойчивого равновесия (теорема Лежен-Дирихле). Основная идея второго метода Ляпунова состоит в нахождении такой функции координат точки фазового пространства данной системы $v(\eta_1, \dots, \eta_n)$, которая была бы до некоторой степени аналогична потенциальной энергии покоящейся материальной точки в обычном пространстве.

Для изложения второго метода исследования устойчивости, разработанного А. М. Ляпуновым, введем следующие определения. Функция $v(\eta_1, \dots, \eta_n)$ называется *знакопостоянной*, если она

имеет один и тот же знак всюду в некоторой области, содержащей начало координат, за исключением некоторых точек, в которых она равна нулю. Знакопостоянная функция, равная нулю лишь в начале координат, называется *законоопределенной*: *определенно-положительной* или *определенно-отрицательной*, в зависимости от знака.

Для любой определено-положительной функции $v(\eta_1, \dots, \eta_n)$ уравнение $v(\eta_1, \dots, \eta_n) = c$ изображает однопараметрическое семейство замкнутых (по крайней мере для малых значений c) поверхностей в фазовом пространстве данной системы. При этом вблизи начала координат поверхность, соответствующая любому достаточно малому значению c , находится целиком внутри

области, ограниченной любой поверхностью, которой соответствует большее значение c . При $c \rightarrow 0$ поверхность стягивается в точку — начало координат.

Из геометрических соображений непосредственно ясно, что если при любом движении системы изображающая точка в фазовом пространстве может двигаться вблизи начала координат только

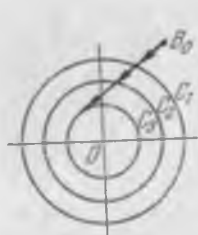


Рис. 10.3.1.

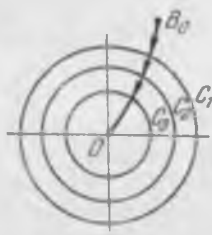


Рис. 10.3.2.

внутри любой поверхности $v(\eta_1, \dots, \eta_n) = c$ или по ней и не может выходить из ограниченной этой поверхностью области наружу (рис. 10.3.1), то система устойчива. Если же изображающая точка может двигаться вблизи начала координат только внутри любой поверхности $v(\eta_1, \dots, \eta_n) = c$ (рис. 10.3.2), то система асимптотически устойчива. Но если изображающая точка движется внутрь поверхности $v(\eta_1, \dots, \eta_n) = c$, то определено-положительная функция $v(\eta_1, \dots, \eta_n)$ убывает при движении системы и, следовательно, ее полная производная по времени отрицательна. Если изображающая точка движется по поверхности $v = c$, то полная производная функции v по времени равна нулю. Иными словами, полная производная функции $v(\eta_1, \dots, \eta_n)$ знакопостоянна отрицательна, если изображающая точка может двигаться только внутрь любой поверхности $v = c$ или по ней, и определено-отрицательна, если изображающая точка может двигаться только внутрь поверхности $v = c$.

Для любой определено-отрицательной функции $v(\eta_1, \dots, \eta_n)$ ее полная производная в рассмотренных случаях будет положительной.

Таким образом, мы приходим к следующим двум теоремам Ляпунова.

1. Если дифференциальные уравнения возмущенного движения таковы, что можно найти знакоопределенную функцию v , полная производная которой по времени

$$\frac{dv}{dt} = \sum_{k=1}^n \frac{\partial v}{\partial \eta_k} \eta_k \quad (10.3.1)$$

в силу этих уравнений знакопостоянна и имеет знак, противоположный знаку функции v , или тождественно равна нулю, то невозмущенное движение устойчиво.

2. Если дифференциальные уравнения возмущенного движения таковы, что можно найти знакоопределенную функцию v , полная производная которой по времени в силу этих уравнений знакоопределенна и имеет знак, противоположный знаку функции v , то невозмущенное движение асимптотически устойчиво.

Изложенный метод оценки устойчивости дает достаточные условия устойчивости. Это означает, что невыполнение этих условий еще не означает, что движение системы неустойчиво. Движение системы может быть устойчивым и при невыполнении условий сформулированных теорем Ляпунова.

Метод исследования устойчивости с помощью функции Ляпунова $v(\eta_1, \dots, \eta_n)$ весьма эффективен и применим к любым нелинейным системам, так как он не накладывает никаких ограничений на правые части уравнений движения (10.1.1).

К сожалению, практическое применение этого метода часто осложняется трудностью нахождения функции Ляпунова $v(\eta_1, \dots, \eta_n)$. Общих рекомендаций по нахождению этой функции для произвольных систем не существует. Однако Ляпунов дал один способ построения функции $v(\eta_1, \dots, \eta_n)$ для широкого круга нелинейных задач, связанных с исследованием устойчивости установившихся движений. Этот способ заключается в том, что уравнения, описывающие поведение системы, сначала линеаризуются. После этого функция v ищется в виде квадратичной формы с неопределенными коэффициентами:

$$v(\eta_1, \dots, \eta_n) = \sum_{k,l=1}^n \alpha_{kl} \eta_k \eta_l. \quad (10.3.2)$$

Коэффициенты α_{kl} определяются из условия

$$\frac{dv}{dt} = \sum_{j=1}^n \frac{\partial v}{\partial \eta_j} \dot{\eta}_j = -A \sum_{r=1}^n \eta_r^2, \quad (10.3.3)$$

где A — любая постоянная.

Подставляя в формулу (10.3.3) выражения производных $\frac{\partial v}{\partial \eta_j}$, полученные дифференцированием формулы (10.3.2), и выражения

производных $\dot{\eta}_i$ из линеаризованных уравнений движения и сравнивая коэффициенты при одинаковых членах $\eta_k \eta_l$ в левой и правой частях равенства, получаем алгебраические уравнения для определения коэффициентов α_{kl} . После нахождения коэффициентов α_{kl} формула (10.3.2) определит функцию Ляпунова v для линеаризованных уравнений. Эта функция может быть использована для оценки устойчивости исходной нелинейной системы.

Другой метод нахождения функции Ляпунова для широкого класса систем, содержащих одно произвольное элементарное нелинейное звено, указал А. И. Лурье [41].

§ 10.4. Фазовые траектории и особые точки нелинейных систем

Для получения достаточно полного представления о характере возможных движений системы без непосредственного интегрирования ее дифференциальных уравнений весьма плодотворным является изучение многообразия фазовых траекторий и особых точек равновесия в фазовом пространстве системы [5, 51, 68]. Особенно наглядны фазовые траектории и особые точки для автономных систем второго порядка, т. е. на фазовой плоскости *). На примерах исследования фазовых траекторий автономных систем второго порядка выявляются основные особенности нелинейных систем вообще.

Уравнение автономной системы второго порядка может быть представлено в виде

$$\ddot{y} = \varphi(y, \dot{y}), \quad (10.4.1)$$

где φ — известная функция выходной переменной и ее производной, в общем случае нелинейная. Полагая $\dot{y} = z$, приведем уравнение (10.4.1) к системе двух уравнений первого порядка:

$$\dot{y} = z, \quad \dot{z} = \varphi(y, z). \quad (10.4.2)$$

Фазовыми координатами системы являются ее выходная переменная y и скорость ее изменения $z = \dot{y}$. Разделив уравнения (10.4.2) одно на другое, получим дифференциальное уравнение фазовых траекторий:

$$\frac{dz}{dy} = \frac{\varphi(y, z)}{z}. \quad (10.4.3)$$

Уравнение (10.4.3) однозначно определяет касательную к фазовой траектории во всех точках, кроме тех, в которых одновременно

*) Системой второго порядка мы называем для краткости систему, поведение которой описывается одним дифференциальным уравнением второго порядка или системой двух дифференциальных уравнений первого порядка.

выполняются равенства

$$\varphi(y, z) = 0, \quad z = 0. \quad (10.4.4)$$

В этих точках не существует определенного направления касательной к траектории. Точки такого типа называются *особыми*. Из этих точек могут исходить многие траектории. Через каждую точку фазовой плоскости, для которой не удовлетворяются одновременно оба уравнения (10.4.4), проходит только одна фазовая траектория.

Из (10.4.2) и (10.4.4) следует, что в особых точках производные фазовых координат системы равны нулю. Следовательно, особые точки являются *точками равновесия* системы. Для отыскания на фазовой плоскости точек равновесия (особых точек) необходимо совместно решить уравнения (10.4.4).

Прежде всего выясним характер возможных особых точек и оценим устойчивость равновесия системы в этих точках. При этом, в соответствии с определением устойчивости по Ляпунову, будем интересоваться поведением системы только в близкой окрестности положения равновесия. Для выяснения поведения системы при малых отклонениях от состояния равновесия воспользуемся методом линеаризации уравнений (10.4.2), предполагая, что функция φ имеет непрерывные первые производные по y и z . Пусть $y_0, z_0 = 0$ — координаты особой точки, т. е. какое-нибудь решение уравнений (10.4.4). Полагая $y = y_0 + \eta, z = \zeta$ и учитывая малость величин η, ζ , можем написать

$$\varphi(y, z) = \varphi(y_0, 0) + a\eta + b\zeta + \varphi_1(\eta, \zeta), \quad (10.4.5)$$

где a, b — постоянные коэффициенты, определяемые формулами

$$a = \varphi'_y(y_0, 0), \quad b = \varphi'_z(y_0, 0), \quad (10.4.6)$$

а $\varphi_1(\eta, \zeta)$ — малая величина высшего порядка по сравнению с η и ζ . Подставляя выражение (10.4.5) функции $\varphi(y, z)$ во второе уравнение (10.4.2), отбрасывая малые высшего порядка, а также принимая во внимание, что координаты особой точки $(y_0, 0)$ удовлетворяют уравнениям (10.4.4), получим систему линейных уравнений первого приближения:

$$\dot{\eta} = \zeta, \quad \dot{\zeta} = a\eta + b\zeta. \quad (10.4.7)$$

Решая эти уравнения первого приближения, можно определить движение соответствующей линейной системы и ее фазовые траектории вблизи особой точки и оценить устойчивость соответствующего положения равновесия. На основании теорем А. М. Ляпунова, сформулированных в § 10.2, полученные результаты будут справедливы и для исходной нелинейной системы.

Системе уравнений (10.4.7) соответствует характеристическое уравнение

$$\lambda^2 - b\lambda - a = 0. \quad (10.4.8)$$

Корни этого уравнения полностью определяют интеграл уравнений (10.4.7), т. е. поведение системы около положения равновесия, а следовательно, и характер особой точки $(y_0, 0)$. Поэтому рассмотрим возможные типы корней характеристического уравнения (10.4.8).

1. *Случай комплексных сопряженных корней $\lambda = h \pm i\omega$.*

Решение уравнений (10.4.7) имеет вид колебаний:

$$\left. \begin{aligned} \eta &= me^{ht} \sin(\omega t + \psi), \\ \zeta &= hme^{ht} \sin(\omega t + \psi) + \omega me^{ht} \cos(\omega t + \psi), \end{aligned} \right\} \quad (10.4.9)$$

где m, ψ — постоянные интегрирования, определяемые начальными условиями. При увеличении t на период $2\pi/\omega$ обе координаты

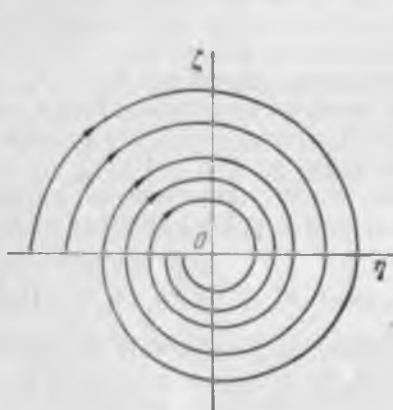


Рис. 10.4.1.

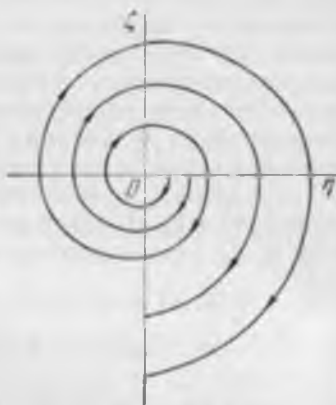


Рис. 10.4.2.

пропорционально убывают при $h < 0$ и возрастают при $h > 0$. Изображающая точка стремится к началу координат при $t \rightarrow \infty$, если $h < 0$. Изображающая точка удаляется от начала координат при $h > 0$. Так как

$$\frac{\zeta}{\eta} = h + \omega \operatorname{ctg}(\omega t + \psi)$$

и так как при увеличении t на период $2\pi/\omega$ котангенс, убывая, дважды пробегает все значения от $-\infty$ до ∞ , то радиус-вектор изображающей точки поворачивается за время $2\pi/\omega$ по часовой стрелке на угол 2π . Следовательно, в случае отрицательной действительной части h корней характеристического уравнения изображающая точка B приближается к особой точке $(y_0, 0)$ по спирали (рис. 10.4.1). В случае положительной действительной части h

корней характеристического уравнения изображающая точка удаляется по спирали от особой точки $(y_0, 0)$ (рис. 10.4.2). При различных начальных условиях движение будет происходить по различным спиралям.

Особая точка такого типа называется *фокусом*. При $h < 0$ это будет устойчивый фокус, соответствующий устойчивому положению равновесия системы, а при $h > 0$ — неустойчивый фокус, соответствующий неустойчивому положению равновесия системы.

2. Случай действительных корней λ_1, λ_2 одного знака.

Решение уравнений (10.4.7) имеет вид

$$\left. \begin{aligned} \eta &= m_1 e^{\lambda_1 t} + m_2 e^{\lambda_2 t}, \\ \zeta &= m_1 \lambda_1 e^{\lambda_1 t} + m_2 \lambda_2 e^{\lambda_2 t}, \end{aligned} \right\} \quad (10.4.10)$$

где m_1 и m_2 — постоянные интегрирования, определяемые начальными условиями. Движение системы является аperiodическим. Изображающая точка асимптотически приближается к особой точке, если λ_1 и λ_2 отрицательны, и удаляется от нее, если λ_1 и λ_2 положительны.

Для выяснения характера фазовых траекторий вблизи особой точки рассмотрим вначале частный случай начальных условий, когда $m_2 = 0$. В этом случае уравнение фазовых траекторий получается делением второй формулы (10.4.10) на первую:

$$\zeta = \lambda_1 \eta. \quad (10.4.11)$$

Таким образом, при $m_2 = 0$ изображающая точка движется по прямой. В другом частном случае, при $m_1 = 0$, получаем

$$\zeta = \lambda_2 \eta. \quad (10.4.12)$$

Таким образом, при $m_1 = 0$ изображающая точка движется по другой прямой. Чтобы получить уравнение фазовых траекторий в общем случае, исключим из уравнений (10.4.10) время t . Для этого умножим первое равенство сначала на λ_1 и вычтем из второго, а затем умножим первое равенство на λ_2 и вычтем из второго. В результате получим

$$\zeta - \lambda_1 \eta = m_2 e^{\lambda_2 t} (\lambda_2 - \lambda_1),$$

$$\zeta - \lambda_2 \eta = m_1 e^{\lambda_1 t} (\lambda_1 - \lambda_2),$$

откуда, исключая время t , находим уравнение фазовых траекторий в виде

$$\zeta - \lambda_1 \eta = c (\zeta - \lambda_2 \eta)^{\frac{\lambda_2}{\lambda_1}}, \quad (10.4.13)$$

где

$$\frac{1}{c} = \frac{[m_1 (\lambda_1 - \lambda_2)]^{\frac{\lambda_2}{\lambda_1}}}{m_2 (\lambda_2 - \lambda_1)}.$$

Определив из уравнения (10.4.13) по правилу дифференцирования неявных функций производную $d\zeta/d\eta$ в особой точке, убеждаемся в том, что все фазовые траектории касаются в особой точке той из прямых (10.4.11) и (10.4.12), которая соответствует меньшему по абсолютной величине из корней λ_1, λ_2 характеристического уравнения.

Семейство фазовых траекторий, определяемое уравнением (10.4.13), для случаев $\lambda_1, \lambda_2 < 0$ и $\lambda_1, \lambda_2 > 0$ представлено соответственно на рис. 10.4.3 и 10.4.4.

Таким образом, через особую точку в этом случае проходят все фазовые траектории, из которых четыре представляют собой полу-прямые. Изображающая точка по любой фазовой траектории

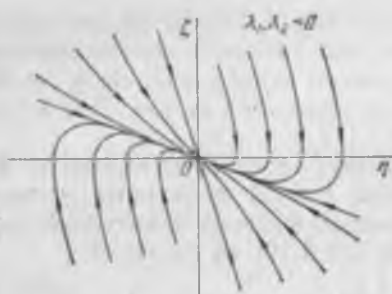


Рис. 10.4.3.

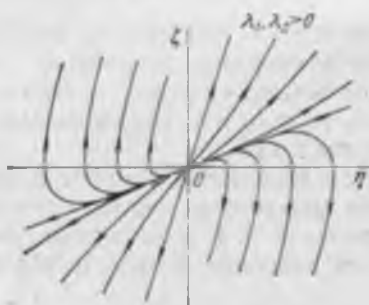


Рис. 10.4.4.

с течением времени асимптотически приближается к особой точке в случае отрицательных корней λ_1, λ_2 характеристического уравнения и удаляется от нее в случае положительных корней λ_1, λ_2 характеристического уравнения. Особая точка такого типа называется *узлом*. При $\lambda_1, \lambda_2 < 0$ узел устойчив, при $\lambda_1, \lambda_2 > 0$ — неустойчив.

Все эти качественные выводы остаются в силе и в том случае, когда $\lambda_1 = \lambda_2$, только прямые (10.4.11) и (10.4.12) в этом случае совпадают.

3. Случай чисто мнимых корней $\lambda = \pm i\omega$.

Решение уравнений (10.4.7) в этом случае принимает форму незатухающих колебаний:

$$\begin{aligned}\eta &= m \sin(\omega t + \psi), \\ \zeta &= m\omega \cos(\omega t + \psi),\end{aligned}\tag{10.4.14}$$

где m, ψ — постоянные интегрирования, определяемые начальными условиями. Возводя в квадрат первое уравнение (10.4.14), деленное на m , и второе, деленное на $m\omega$, и складывая полученные таким образом равенства, находим уравнение семейства фазовых

траекторий:

$$\frac{\eta^2}{m^2} + \frac{\xi^2}{m^2\omega^2} = 1. \quad (10.4.15)$$

Таким образом, семейство фазовых траекторий представляет собой семейство подобных эллипсов с полуосями m и $m\omega$ (рис. 10.4.5). Через каждую точку плоскости проходит один эллипс. Вся фазовая плоскость оказывается заполненной вложенными друг в друга эллипсами. Изображающая точка движется по любому эллипсу по часовой стрелке вокруг особой точки равновесия. Такая особая точка называется *центром*. Подобные точки равновесия характерны для консервативных систем, в которых нет притока и расхода энергии.

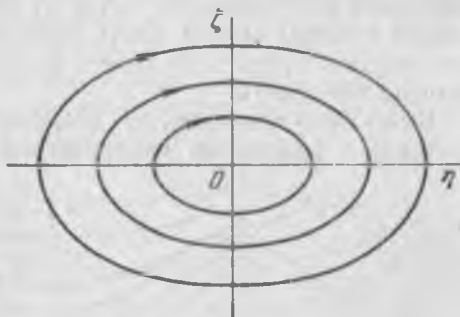


Рис. 10.4.5.

Линейная система в этом случае устойчива по Ляпунову. Однако относительно устойчивости исходной нелинейной системы ничего сказать нельзя, так как линеаризованная система устойчива не асимптотически и теоремы § 10.2 неприменимы. Необходимы дополнительные исследования. Эти исследования в каждом частном случае показывают, является ли рассматриваемая особая точка фактически центром, как, например, в случае маятника без трения, или нет.

4. *Случай двух действительных корней λ_1, λ_2 различных знаков ($\lambda_1 < 0, \lambda_2 > 0$).*

В этом случае показатель степени в уравнении (10.4.13) отрицателен, вследствие чего при обращении одной из величин $\xi - \lambda_1\eta$ и $\xi - \lambda_2\eta$ в нуль другая становится бесконечной. Это свидетельствует о том, что фазовые траектории имеют в данном случае прямые (10.4.11) и (10.4.12) асимптотами (рис. 10.4.6). Таким образом, в данном случае только две прямолинейные фазовые траектории (10.4.11) и (10.4.12) проходят через особую точку. Где бы ни находилась изображающая точка в начальный момент, она всегда в конечном итоге станет удаляться от положения равновесия. Только в случае движения изображающей точки точно по асимптоте (10.4.11) она приближается к положению равновесия. Но стоит изображающей точке отклониться от этой прямой, как она уйдет от особой точки. Такая особая точка является *седлом*. Седло всегда соответствует положению неустойчивого равновесия системы.

Для линейной системы уравнения (10.4.4) линейны и вследствие этого могут иметь только одно решение. Таким образом, линейная

система может иметь только одно положение равновесия. Все полученные выше выводы о характере фазовых траекторий справедливы для всей фазовой плоскости. Следовательно, характер особой точки для линейной системы полностью определяет ее поведение при любых отклонениях от особой точки, т. е. на всей фазовой плоскости.

Для нелинейной системы характер особой точки не определяет поведения изображающей точки на всей фазовой плоскости. Нелинейная система может иметь несколько положений равновесия, а ее фазовые траектории имеют различный вид в разных областях фазовой плоскости.

Выше было показано, что при чисто мнимых корнях характеристического уравнения консервативная линейная система может

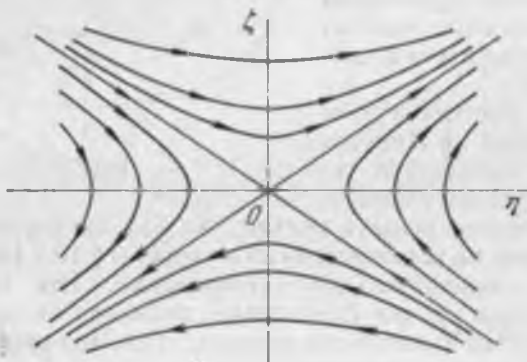


Рис. 10.4.6.

совершать периодические движения любой амплитуды и одной строго определенной частоты. Вся фазовая плоскость такой системы заполнена вложенными друг в друга замкнутыми траекториями. Нелинейная система также может иметь в этом случае особую точку — центр, около которого все фазовые траектории являются замкнутыми кривыми. Однако, кроме таких фазовых траекторий, для нелинейной системы в этом случае обязательно будут существовать в других областях фазовой плоскости и незамкнутые фазовые траектории. Примером такой системы может служить обычный плоский маятник (математический или физический). Фазовые траектории плоского маятника, не обладающего затуханием, изображены на рис. 10.4.7.

Маятник имеет два положения равновесия: нижнюю точку устойчивого равновесия и верхнюю точку неустойчивого равновесия. При отклонении от положения устойчивого равновесия маятник совершает колебания с амплитудой, равной начальному отклонению. Фазовая траектория его является замкнутой кривой,

охватывающей точку устойчивого равновесия — центр. При увеличении начального отклонения до π маятник при нулевой скорости ($z = 0$) оказывается в верхней точке неустойчивого равновесия — седле. Движение маятника может происходить из этой точки в любом направлении. Этому движению соответствуют фазовые траектории, отмеченные на рис. 10.4.7 жирными линиями. Если маятнику сообщить достаточно большую начальную скорость, то он станет вращаться относительно точки подвеса в соответствующем направлении, а его скорость нигде не будет обращаться в нуль. Фазовые траектории, соответствующие этому случаю движения, расположены вне траекторий, отмеченных на

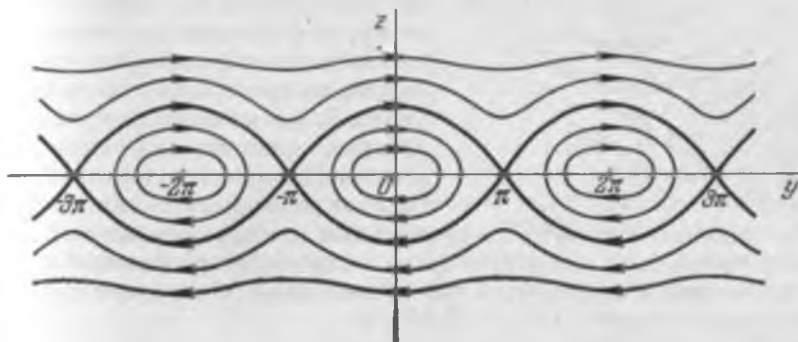


Рис. 10.4.7.

рис. 10.4.7 жирными линиями. Последние являются особыми фазовыми траекториями — *сепаратрисами*, проходящими через особые точки типа седел, и служат разграничительными кривыми, разделяющими области с траекториями разных типов. Следовательно, фазовая плоскость маятника содержит особые точки типа центра и седла и заполнена интегральными кривыми двух видов: эллипсами внутри сепаратрис и волнообразными кривыми вне сепаратрис.

Однако в задачах практики часто встречаются и такие нелинейные системы, у которых существуют только отдельные изолированные замкнутые фазовые траектории, называемые *предельными циклами*. Всякая траектория, соседняя с замкнутой траекторией, уже не является замкнутой, она либо наматывается на предельный цикл, либо сходит с него. Если фазовые траектории, близкие к предельному циклу, наматываются на него, то предельный цикл *устойчив* и ему соответствует устойчивое периодическое движение системы. Если при движении по фазовой траектории, близкой к предельному циклу, изображающая точка удаляется от него, то предельный цикл *неустойчив* и соответствующее периодическое движение системы неустойчиво.

На рис. 10.4.8 показана фазовая диаграмма системы, имеющей особую точку неустойчивого равновесия типа неустойчивого фокуса и один устойчивый предельный цикл. Изображающие точки, находящиеся внутри предельного цикла, движутся по развертывающейся спирали, приближаясь к предельному циклу. Все изображающие точки, находящиеся вне предельного цикла, с течением времени по свертывающимся спиральям приближаются к предельному циклу. Таким образом, если система не находится в состоянии равновесия, то с течением времени в ней устанавливается периодический режим, не зависящий от начальных условий. Такие периодические колебания называются *автоколебаниями*. На рис. 10.4.9 показана фазовая диаграмма нелинейной системы с одним неустойчивым предельным циклом. Предельный цикл может также быть *полуустойчивым* в случае, когда фазовые траектории с одной стороны приближаются к нему, а с другой стороны удаляются от него (рис. 10.4.10).

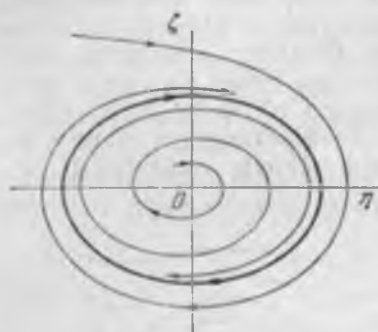


Рис. 10.4.8.

Неустойчивый и полуустойчивый предельные циклы соответствуют неустойчивым автоколебаниям в системе. Неустойчивые автоколебания практически невозможны, так как при малейшем возмущении изображающая точка сходит с соответствующего предельного цикла и удаляется от него.

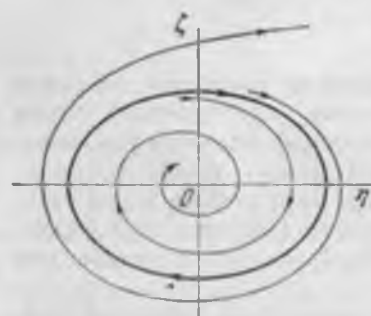


Рис. 10.4.9.

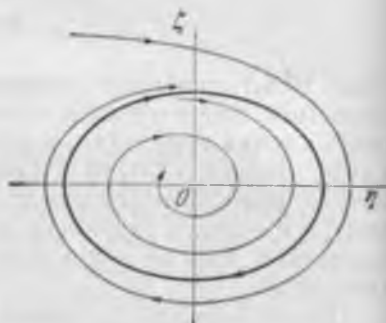


Рис. 10.4.10.

Если система имеет несколько предельных циклов, соответствующих одной и той же особой точке, то неустойчивые и устойчивые предельные циклы всегда чередуются.

В этих случаях конечное положение системы зависит от структуры ее фазового портрета (от свойств системы) и от исходного толчка — возмущения, т. е. от положения изображающей точки в начальный момент времени. Если, например, система имеет неустойчивый фокус, то ближайший к нему предельный цикл является устойчивым (сплошная линия), затем следует неустойчивый (пунктир) и устойчивый циклы, чередуясь, как показано на рис. 10.4.11. В этом случае при любом возмущении в системе всегда устанавливается один из устойчивых режимов автоколебаний. При этом устанавливается тот режим, предельный цикл которого на фазовом портрете ближе всего к начальному положению изображающей точки системы. Такие системы носят название систем с *мягким режимом* возбуждения колебаний. Если система имеет устойчивый фокус, то ближайший к нему предельный цикл является неустойчивым. Затем могут следовать, чередуясь, устойчивые и неустойчивые предельные циклы, как показано

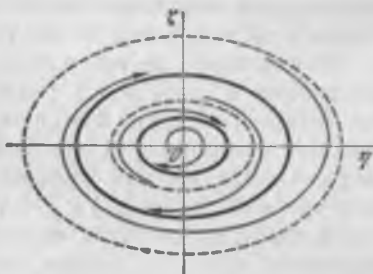


Рис. 10.4.11.

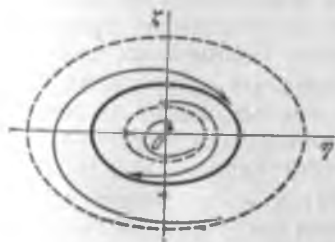


Рис. 10.4.12.

на рис. 10.4.12. В этом случае колебания в системе возбуждаются только при достаточно больших начальных отклонениях или при достаточно большом возмущении, при которых изображающая точка окажется вне первого неустойчивого цикла. При этом система раскачивается до возникновения автоколебательного режима, соответствующего ближайшему устойчивому циклу. Если в начальный момент изображающая точка окажется

внутри первого неустойчивого цикла, то колебания в системе затухают. Такие системы называются системами с *жестким режимом* возбуждения колебаний.

Таким образом, фазовая плоскость нелинейной системы может иметь весьма сложную структуру. В окрестности особых точек фазовые траектории могут иметь различный характер. Границами между различными частями фазовой плоскости являются особые фазовые траектории — сепаратрисы.

В § 8.2 мы видели, что характеристики элементарных нелинейных звеньев обычно считаются кусочно-линейными функциями, которые изображаются ломаными линиями, иногда разрывными. Угловым точкам и точкам разрыва этих ломаных на фазовой

плоскости нелинейной системы соответствуют некоторые линии, которые называются *линиями переключения*. В зависимости от характера входных сигналов нелинейных звеньев линии переключения могут быть прямыми или кривыми. Однако чаще всего входные сигналы нелинейных звеньев являются линейными функциями переменных, определяющих состояние системы. В этом случае линии переключения являются прямыми.

Изложенный метод исследования особых точек на фазовой плоскости и поведения системы вблизи положения равновесия неприменим в случае ломаных характеристик нелинейных звеньев, так как правая часть уравнения (10.4.1) не имеет в этом случае непрерывных первых производных и, следовательно, линеаризация уравнений (10.4.1) и (10.4.2) невозможна. Однако особые точки систем с существенными нелинейностями имеют тот же общий характер, что и в случае систем с гладкими характеристиками нелинейных звеньев. Только в некоторых случаях вместо особой точки может появиться особый отрезок оси y , каждая точка которого соответствует возможному состоянию покоя системы, как, например, в случае системы с зоной нечувствительности.

В случае ломаных характеристик нелинейных звеньев линии переключения делят фазовую плоскость на ряд областей. Поэтому построение фазовых траекторий системы приходится вести по участкам. При этом начальные значения переменных на каждом участке должны быть равны их конечным значениям на предыдущем участке. Если входные сигналы всех нелинейных звеньев являются линейными функциями переменных, определяющих состояние системы, и все нелинейные звенья имеют кусочно-линейные характеристики, то поведение системы в каждой области фазовой плоскости описывается линейными дифференциальными уравнениями с постоянными коэффициентами. При пересечении изображающей точкой линии переключения некоторые из коэффициентов скачками изменяют свои значения и вновь остаются постоянными до пересечения следующей линии переключения. Так как любые линейные дифференциальные уравнения с постоянными коэффициентами легко интегрируются стандартными методами, то интегрирование уравнений нелинейной системы и построение ее фазовых траекторий по участкам для таких систем не представляет принципиальных затруднений.

Пример 10.4.1. Рассмотрим простейшую релейную следящую систему (рис. 10.4.13), содержащую линейное исполнительное устройство — двигатель постоянного тока с передаточной функцией $k/s (Ts + 1)$, поляризованное реле и отрицательную обратную связь с тахогенератором *). Структурную схему такой системы изобразим в виде рис. 10.4.14. Для простоты будем сначала считать тахогенератор $TГ$ отключенным. В этом случае обратная связь будет жесткой и $T_1 = 0$. Из рис. 10.4.14* вытекает следующее

*) Примеры 10.4.1—10.4.5 и 10.5.1 заимствованы в основном из [34].

уравнение системы в этом случае:

$$T\ddot{y} + \dot{y} = k\varphi(x - y), \quad (10.4.16)$$

где x — входной сигнал системы. Так как в дальнейшем будем рассматривать только собственные движения системы, то положим $x = 0$. Кроме того, положим $\varphi(y) = l\psi(y)$. Тогда, вводя переменную $z = \dot{y}$, заменим уравнение (10.4.16) равноценной системой уравнений первого порядка:

$$\left. \begin{aligned} \dot{y} &= z, \\ T\dot{z} &= -z - kl\psi(y). \end{aligned} \right\} \quad (10.4.17)$$

Разделив второе уравнение системы (10.4.17) на первое, получим дифференциальное уравнение фазовых траекторий:

$$T \frac{dz}{dy} = -\frac{kl\psi(y)}{z} - 1. \quad (10.4.18)$$

Принимая во внимание, что функция $-\psi$ для реле имеет постоянное значение (равное -1 ; 0 или 1), обозначим это постоянное значение через $v = -\psi(y)$. Тогда, интегрируя уравнение (10.4.18) при начальных условиях $t = t_0$, $y = y_0$, $z = z_0$, получим

$$y = y_0 + T \left(z_0 - z + klv \ln \frac{z_0 - klv}{z - klv} \right). \quad (10.4.19)$$

Пользуясь этой формулой, можно построить фазовые траектории изображающей точки для различных типов реле в следящей системе. Рассмотрим сначала двухпозиционное поляризованное реле с идеальной характеристикой 5 (приложения 5 и 6). Переключение реле происходит при $y = 0$, т. е. линией переключения является ось z (отмечена штриховкой на рис. 10.4.15). В этом случае $v = 1$ при $y < 0$ и $v = -1$ при $y > 0$. Следовательно, уравнение фазовой траектории имеет вид

$$y = y_0 + T \left(z_0 - z + kl \ln \frac{z_0 - kl}{z - kl} \right) \quad \text{при } y < 0, \quad (10.4.20)$$

$$y = y_0 + T \left(z_0 - z - kl \ln \frac{z_0 + kl}{z + kl} \right) \quad \text{при } y > 0. \quad (10.4.21)$$

Предположим, что в начальный момент $t = 0$ в следящей системе имеется рассогласование $y_0 =$

a_0 , а начальная скорость выходного вала $z_0 = 0$, т. е. изображающая точка в начальный момент имеет координаты $(a_0, 0)$ в левой полуплоскости (рис. 10.4.15). Для построения фазовой траектории надо использовать уравнение (10.4.20). Это уравнение справедливо до тех пор, пока фазовая траектория не пересечет ось z в точке $z = b_0$. В этой точке рассогласование уменьшилось до нуля, но скорость выходного вала возросла

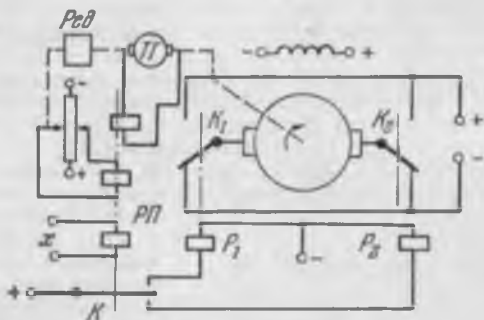


Рис. 10.4.13.

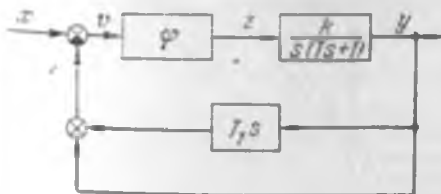


Рис. 10.4.14.

плоскости нелинейной системы соответствуют некоторые линии, которые называются *линиями переключения*. В зависимости от характера входных сигналов нелинейных звеньев линии переключения могут быть прямыми или кривыми. Однако чаще всего входные сигналы нелинейных звеньев являются линейными функциями переменных, определяющих состояние системы. В этом случае линии переключения являются прямыми.

Изложенный метод исследования особых точек на фазовой плоскости и поведения системы вблизи положения равновесия неприменим в случае ломаных характеристик нелинейных звеньев, так как правая часть уравнения (10.4.1) не имеет в этом случае непрерывных первых производных и, следовательно, линеаризация уравнений (10.4.1) и (10.4.2) невозможна. Однако особые точки систем с существенными нелинейностями имеют тот же общий характер, что и в случае систем с гладкими характеристиками нелинейных звеньев. Только в некоторых случаях вместо особой точки может появиться особый отрезок оси y , каждая точка которого соответствует возможному состоянию покоя системы, как, например, в случае системы с зоной нечувствительности.

В случае ломаных характеристик нелинейных звеньев линии переключения делят фазовую плоскость на ряд областей. Поэтому построение фазовых траекторий системы приходится вести по участкам. При этом начальные значения переменных на каждом участке должны быть равны их конечным значениям на предыдущем участке. Если входные сигналы всех нелинейных звеньев являются линейными функциями переменных, определяющих состояние системы, и все нелинейные звенья имеют кусочно-линейные характеристики, то поведение системы в каждой области фазовой плоскости описывается линейными дифференциальными уравнениями с постоянными коэффициентами. При пересечении изображающей точкой линии переключения некоторые из коэффициентов скачками изменяют свои значения и вновь остаются постоянными до пересечения следующей линии переключения. Так как любые линейные дифференциальные уравнения с постоянными коэффициентами легко интегрируются стандартными методами, то интегрирование уравнений нелинейной системы и построение ее фазовых траекторий по участкам для таких систем не представляет принципиальных затруднений.

Пример 10.4.1. Рассмотрим простейшую релейную следящую систему (рис. 10.4.13), содержащую линейное исполнительное устройство — двигатель постоянного тока с передаточной функцией $k/s (Ts + 1)$, поляризованное реле и отрицательную обратную связь с тахогенератором *). Структурную схему такой системы изобразим в виде рис. 10.4.14. Для простоты будем сначала считать тахогенератор $TГ$ отключенным. В этом случае обратная связь будет жесткой и $T_1 = 0$. Из рис. 10.4.14* вытекает следующее

*) Примеры 10.4.1—10.4.5 и 10.5.1 заимствованы в основном из [34].

уравнение системы в этом случае:

$$T\ddot{y} + \dot{y} = k\varphi(x - y), \quad (10.4.16)$$

где x — входной сигнал системы. Так как в дальнейшем будем рассматривать только собственные движения системы, то положим $x = 0$. Кроме того, положим $\varphi(y) = l\psi(y)$. Тогда, вводя переменную $z = \dot{y}$, заменим уравнение (10.4.16) равноценной системой уравнений первого порядка:

$$\left. \begin{aligned} \dot{y} &= z, \\ T\dot{z} &= -z - kl\psi(y). \end{aligned} \right\} \quad (10.4.17)$$

Разделив второе уравнение системы (10.4.17) на первое, получим дифференциальное уравнение фазовых траекторий:

$$T \frac{dz}{dy} = -\frac{kl\psi(y)}{z} - 1. \quad (10.4.18)$$

Принимая во внимание, что функция $-\psi$ для реле имеет постоянное значение (равное -1 ; 0 или 1), обозначим это постоянное значение через $v = -\psi(y)$. Тогда, интегрируя уравнение (10.4.18) при начальных условиях $t = t_0$, $y = y_0$, $z = z_0$, получим

$$y = y_0 + T \left(z_0 - z + klv \ln \frac{z_0 - klv}{z - klv} \right). \quad (10.4.19)$$

Пользуясь этой формулой, можно построить фазовые траектории изображающей точки для различных типов реле в следящей системе. Рассмотрим сначала двухпозиционное поляризованное реле с идеальной характеристикой 5 (приложения 5 и 6). Переключение реле происходит при $y = 0$, т. е. линией переключения является ось z (отмечена штриховкой на рис. 10.4.15). В этом случае $v = 1$ при $y < 0$ и $v = -1$ при $y > 0$. Следовательно, уравнение фазовой траектории имеет вид

$$y = y_0 + T \left(z_0 - z + kl \ln \frac{z_0 - kl}{z - kl} \right) \quad \text{при } y < 0, \quad (10.4.20)$$

$$y = y_0 + T \left(z_0 - z - kl \ln \frac{z_0 + kl}{z + kl} \right) \quad \text{при } y > 0. \quad (10.4.21)$$

Рис. 10.4.14.

$= a_0$, а начальная скорость выходного вала $z_0 = 0$, т. е. изображающая точка в начальный момент имеет координаты $(a_0, 0)$ в левой полуплоскости (рис. 10.4.15). Для построения фазовой траектории надо использовать уравнение (10.4.20). Это уравнение справедливо до тех пор, пока фазовая траектория не пересечет ось z в точке $z = b_0$. В этой точке рассогласование уменьшилось до нуля, но скорость выходного вала возросла

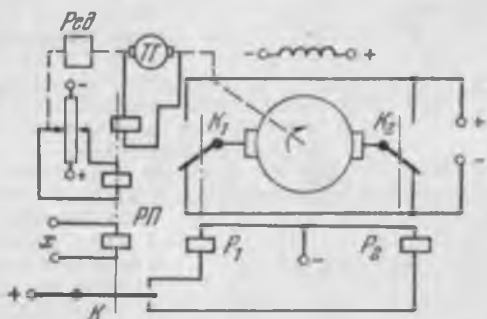
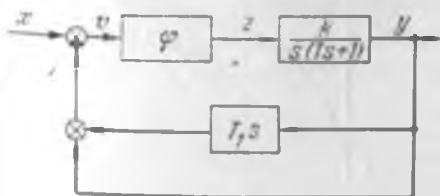


Рис. 10.4.13.



до значения b_0 . Реле переключается при переходе через эту точку. С этого момента фазовая траектория определяется уравнением (10.4.21). При построении фазовой траектории по уравнению (10.4.21) надо положить начальные значения y_0 и z_0 равными конечным значениям y и z , удовлетворяющим

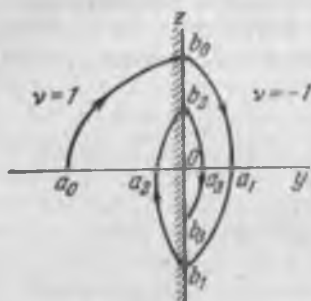


Рис. 10.4.15.

уравнению (10.4.20), т. е. $y_0 = 0$, $z_0 = b_0$. Уравнение (10.4.21) справедливо, пока фазовая траектория не пересечет ось z в точке b_1 . В этой точке опять рассогласование равно нулю: $y = 0$, но выходной вал движется в противоположном направлении со скоростью $z = b_1$. В этот момент реле вновь переключается, и для построения следующего участка фазовой траектории до точки $z = b_2$ необходимо использовать уравнение (10.4.20) и т. д. Как показывают расчеты и построение, изображающая точка приближается к началу координат по траектории типа спирали (рис. 10.4.15). Сама же система в силу наличия в ней сил сопротивления совершает затухающие колебания, асимптотически приближаясь к началу координат — точке устойчивого равновесия. Частота колебаний увеличивается при уменьшении амплитуды и в пределе, при $t \rightarrow \infty$, стремится к бесконечности. В этом можно убедиться, если рассмотреть интеграл уравнений (10.4.17) при начальных условиях $t = 0$, $y = y_0$, $z = z_0$:

$$y = y_0 + Tz_0(1 - e^{-\frac{t}{T}}) + vktT \left[\frac{t}{T} - (1 - e^{-\frac{t}{T}}) \right], \quad (10.4.22)$$

$$z = z_0 e^{-\frac{t}{T}} + vkt(1 - e^{-\frac{t}{T}}), \quad v = -\operatorname{sgn} y. \quad (10.4.23)$$

Пользуясь этими уравнениями, можно построить графики изменения отклонения y и скорости $z = \dot{y}$ по участкам (переходный процесс). Так как пере-

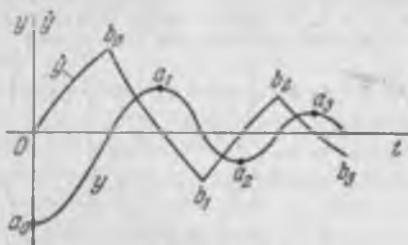


Рис. 10.4.16.

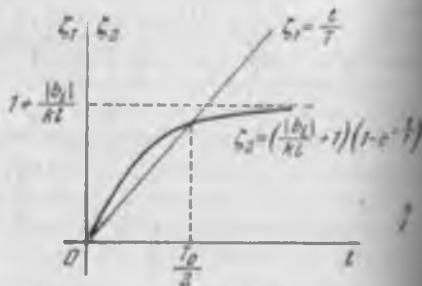


Рис. 10.4.17.

ключение (перемена знака v) происходит при $y = 0$, то начальные условия для каждого участка, кроме первого, имеют вид $y_0 = 0$, $z_0 = b_i$ ($i = 0, 1, 2, \dots$). Для первого участка начальные условия определяются начальным состоянием системы. На рис. 10.4.16 даны кривые переходного процесса для рассматриваемого примера.

Для определения времени, в течение которого система совершает одно полуколебание и отклонение y сохраняет постоянный знак, следует в урав-

нения (10.4.22) положить $y_0 = y = 0$, $z_0 = b_1$, $v = -\operatorname{sgn} b_1$. Тогда получим следующее уравнение:

$$\left(\frac{|b_1|}{kl} + 1\right) \left(1 - e^{-\frac{t}{T}}\right) = \frac{t}{T}. \quad (10.4.24)$$

Это уравнение легко решается графически. Для этого достаточно построить графики его левой и правой частей и определить точку пересечения полученных кривых (рис. 10.4.17). Так как при $b_1 \rightarrow 0$ угловой коэффициент касательной к кривой, соответствующей левой части уравнения (10.4.24), в начале координат стремится к $1/T$, то время полукоселебания стремится к нулю по мере уменьшения амплитуды колебаний. Это и доказывает, что частота колебаний неограниченно возрастает при приближении системы к положению равновесия.

Пример 10.4.2. В условиях предыдущего примера рассмотрим слу-

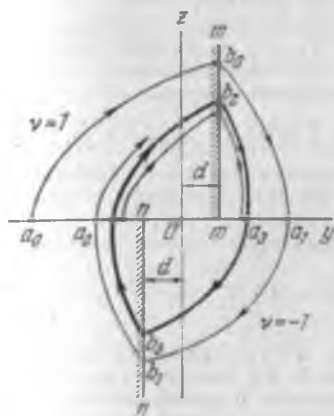


Рис. 10.4.18.

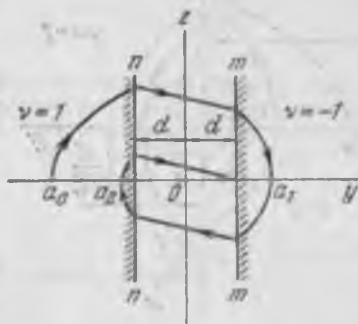


Рис. 10.4.19.

чай двухпозиционного реле с зоной неоднозначности (приложения 5 и 6, нелинейность 8). Переключение реле происходит, когда $y = d$ при $z > 0$ и $y = -d$ при $z < 0$, т. е. линиями переключения являются прямые $y = \pm d$. На рис. 10.4.18 они обозначены буквами л, т и штриховкой. В этом случае

$$v = 1 \quad \text{при } y < d, z > 0 \text{ и при } y < -d, z < 0,$$

$$v = -1 \quad \text{при } y > d, z > 0 \text{ и при } y > -d, z < 0.$$

Фазовая траектория в этом случае строится по участкам совершенно так же, как в предыдущем примере, по уравнениям (10.4.20) и (10.4.21). Вычисления и построение показывают, что фазовые траектории в этом случае не стремятся к началу координат, а сходятся к некоторому предельному циклу (рис. 10.4.18). Это значит, что в системе устанавливаются автоколебания с определенными амплитудой и частотой, а положение равновесия $y = z = 0$ является неустойчивым фокусом. Амплитуда и частота автоколебаний не зависят от начальных условий, а зависят только от параметров системы: величины координатного запаздывания d реле, постоянной времени T и коэффициента усиления k исполнительного устройства.

Пример 10.4.3. В условиях примера 10.4.1 рассмотрим случай идеального реле с зоной нечувствительности (нелинейность 6, приложения 5 и 6). Переключение такого реле происходит, когда $y = \pm d$. Линии переключения изображены на рис. 10.4.19 штриховкой. В этом случае имеются три области:

левее линии переключения $y = -d$ ($v = 1$), область между линиями переключения $y = \pm d$ ($v = 0$) и область правее линии переключения $y = d$ ($v = -1$). Фазовая траектория в этом случае, так же как и в предыдущем случае, строится по участкам, по уравнению (10.4.20) в первой области ($v = 1$), по уравнению (10.4.19) при $v = 0$ во второй области и по уравнению (10.4.21) в третьей области ($v = -1$). Фазовые траектории в этом случае представляют собой спиральные кривые (рис. 10.4.19). Система успокаивается в положении, когда координата y попадает в зону нечувствительности $(-d, +d)$ при значении $z = 0$. Таким образом, в системе происходят затухающие колебания. В частном случае, при начальных отклонениях, меньших зоны нечувствительности $|y_0| \leq d$, $z_0 = 0$ и система остается в покое.

Пример 10.4.4. Рассмотрим теперь влияние отрицательной скоростной обратной связи, осуществляемой с помощью тахогенератора, на дина-

мические характеристики простейшей релейной следящей системы. В этом случае $T_1 \neq 0$ и из рис. 10.4.14 вытекает следующее уравнение, описывающее поведение системы:

$$T\ddot{y} + \dot{y} = k\varphi(x - y - T_1\dot{y}). \quad (10.4.25)$$

Рассматривая только собственные движения системы, положим $x = 0$ и приведем уравнение (10.4.25) к системе двух уравнений первого порядка:

$$\begin{aligned} \dot{y} &= z, & T\dot{z} &= -z - k\varphi(y + T\xi z), \end{aligned} \quad (10.4.26)$$

где $\xi = T_1/T$. Предположим, как и в примере 10.4.2, что реле является двухпозиционным с зоной неоднозначности (нелинейность 8, приложения 5 и 6). В этом случае переключение реле происходит, когда $y + T\xi z = d$ при $z > 0$

и $y + T\xi z = -d$ при $z < 0$, т. е. линиями переключения являются прямые mm , nn на рис. 10.4.20. Следовательно, $v = 1$ слева от прямой mm при $z > 0$ и слева от прямой $n - n$ при $z < 0$; $v = -1$ справа от прямой $m - m$ при $z > 0$ и справа от прямой $n - n$ при $z < 0$. Фазовые траектории системы определяются уравнением (10.4.20) в части фазовой плоскости, расположенной слева от линий переключения, и уравнением (10.4.21) в части плоскости, расположенной справа от линий переключения.

Мы видим, что введение скоростной отрицательной обратной связи приводит к наклону линий переключения против движения изображающей точки. Вследствие этого переключения происходят с упреждением по отношению к моменту достижения отклонением величины d (границы зоны запаздывания).

Фазовые траектории в этом случае также строятся по участкам. Вычисления и построения показывают, что фазовые траектории в этом случае сходятся так же, как и в примере 10.4.2, к предельному циклу (рис. 10.4.20). Однако при тех же параметрах системы амплитуда автоколебаний снижается, а частота повышается, что ведет к улучшению качества следящей системы.

Пример 10.4.5. С помощью отрицательной скоростной обратной связи можно существенно улучшить переходный процесс в нелинейной системе. Например, можно получить оптимальный переходный процесс без перерегулирования с одним переключением реле, если ввести соответствующую нелинейную гибкую обратную связь.

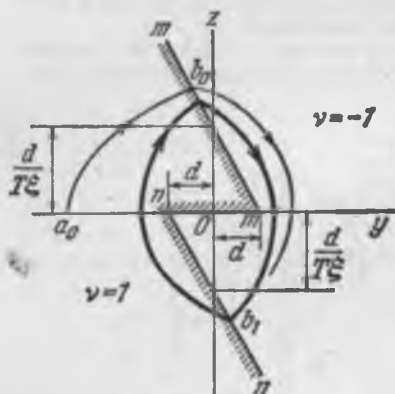


Рис. 10.4.20.

Рассмотрим, например, простейшую нелинейную следящую систему второго порядка, состоящую из двойного интегрирующего звена (безынерционный двигатель) и релейного управляющего элемента с идеальной характеристикой, охваченных отрицательной обратной связью с нелинейным преобразованием производной выходного сигнала системы. Уравнение этой следящей системы, как легко видеть на рис. 10.4.21, имеет вид

$$\ddot{y} = k_1 \Phi(x - y - f(\dot{y})), \quad (10.4.27)$$

где $\Phi(u) = l \operatorname{sgn} u$, а x — входной сигнал. Для исследования переходного процесса в системе, т. е. ее собственных движений, положим $x = 0$.

Тогда, полагая $k = k_1 l$, $z = \dot{y}$, приведем уравнение (10.4.27) к системе двух уравнений первого порядка:

$$\dot{y} = z, \quad \dot{z} = kv, \quad (10.4.28)$$

где $v = -1$ при $y + f(z) > 0$ и $v = 1$ при $y + f(z) < 0$. Таким образом, кривая $y + f(z) = 0$ является в данном случае линией переключения. Разделив второе уравнение (10.4.28) на первое, получим дифференциальное уравнение

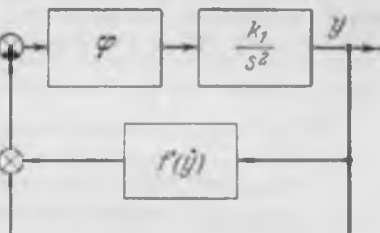


Рис. 10.4.21.

фазовых траекторий, интеграл которого выражается формулой

$$z^2 - z_0^2 = 2kv(y - y_0). \quad (10.4.29)$$

На рис. 10.4.22 изображены два семейства фазовых траекторий, соответствующих $v=1$ и $v=-1$.

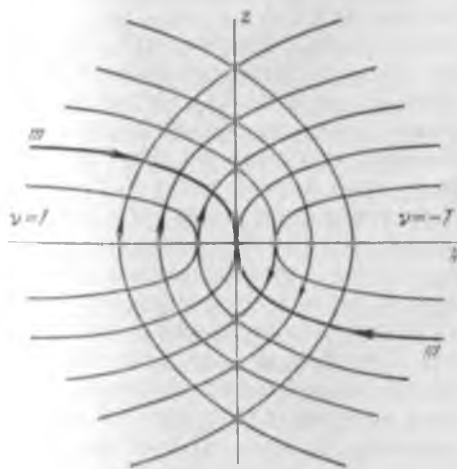


Рис. 10.4.22.

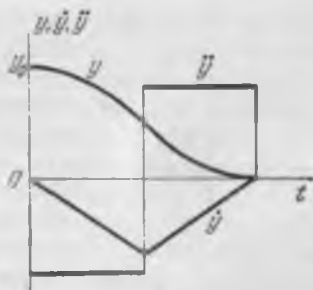


Рис. 10.4.23.

Поставим теперь задачу выбрать линию переключения реле так, чтобы переходный процесс при любых начальных отклонениях заканчивался за одно подколебание. Легко сообразить, что такой линией переключения является кривая, образованная половинами проходящих через начало координат фазовых траекторий двух семейств (кривая mn на рис. 10.4.22). В этом случае изображающая точка, попав на линию переключения, пойдет по ней к началу координат. Такой переходный процесс называется *оптимальным по быстротечности*.

Для определения соответствующего преобразования производной в цепи обратной связи напомним уравнение найденной оптимальной линии переключения. Для этого положим в уравнении (10.4.29) $y_0 = z_0 = 0$, $v = -1$

при $z > 0$, $v = 1$ при $z < 0$. Сравнивая полученное таким образом уравнение линии переключения с уравнением $y + f(z) = 0$, находим функцию $f(z)$:

$$f(z) = \frac{1}{2k} z^2 \operatorname{sgn} z. \quad (10.4.30)$$

На рис. 10.4.23 изображены графики переходного процесса в рассматриваемой нелинейной следящей системе. В этом случае переходный процесс в системе при любом начальном отклонении протекает на участке разгона с максимальным ускорением, а на участке торможения, когда переключается реле, с максимальным замедлением.

Общая теория оптимальных процессов управления разработана в трудах Л. С. Понтрягина, Р. В. Гамкрелидзе, В. Г. Болтянского, Е. Ф. Мищенко [48]. Задача оптимизации системы по быстродействию впервые рассматривалась А. А. Фельдбаумом [71].

§ 10.5. Исследование нелинейных систем методом точечных отображений

Для количественного исследования фазовых траекторий и автоколебаний нелинейных систем можно применить метод точечных отображений, развитый А. А. Андроновым [5].

Для пояснения идеи этого метода предположим, что изображающая точка в какой-то момент времени занимает положение 1 на полупрямой Oz (рис. 10.5.1). После обхода вокруг начала координат изображающая точка пересечет полупрямую Oz , вообще говоря, в некоторой другой точке 2. Так как через каждую точку полупрямой Oz можно провести одну и только одну фазовую траекторию, то обходу изображающей точки вокруг начала координат соответствует переход каждой точки полупрямой Oz в некоторую другую точку той же полупрямой. Иными словами, обходу изображающей точки вокруг начала координат (или другой особой точки на фазовой плоскости) соответствует

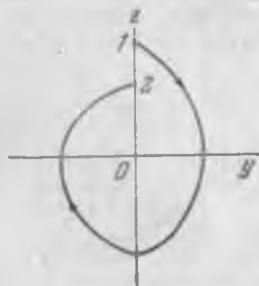


Рис. 10.5.1.

точечное преобразование любой полупрямой, выходящей из начала координат, в саму себя. Если при этом преобразовании какая-нибудь точка полупрямой переходит в саму себя (т. е. остается неподвижной), то через эту точку проходит замкнутая фазовая траектория — предельный цикл. Следовательно, для нахождения предельных циклов и параметров соответствующих автоколебаний достаточно найти неподвижные точки точечного преобразования какой-нибудь полупрямой, выходящей из соответствующей особой точки фазовой плоскости.

Каждой точке выбранной полупрямой соответствует некоторое положительное число — расстояние этой точки от начала координат.

нат (или от соответствующей особой точки). Следовательно, точечное отображение этой полупрямой при обходе изображающей точки вокруг начала координат определяет однозначную возрастающую функцию $v = f(z)$, которая может быть изображена кривой в координатах (z, v) (рис. 10.5.2). Каждая точка этой кривой с одинаковыми абсциссой и ординатой, т. е. точка пересечения этой кривой с биссектрисой координатного угла $v = z$, является неподвижной точкой отображения и определяет предельный цикл. Участкам кривой $v = f(z)$, лежащим ниже биссектрисы координатного угла $v = z$, соответствуют спиральные фазовые траектории, по которым изображающая точка приближается к началу координат. Участкам кривой $v = f(z)$, лежащим выше биссектрисы координатного угла $v = z$, соответствуют спиральные фазовые траектории, по которым изображающая точка удаляется от начала координат. Таким образом, взаимное расположение кривой точечного отображения $v = f(z)$ и биссектрисы координатного угла $v = z$ полностью определяет поведение системы около положения равновесия. Чтобы определить характер колебаний системы при любом начальном положении z_0 изображающей точки на полуоси $z > 0$, достаточно провести из этой точки прямую, параллельную оси ординат, до пересечения с кривой $v = f(z)$. Затем из полученной точки пересечения провести горизонтальную прямую до пересечения с биссектрисой координатного угла $v = z$. Из полученной точки пересечения снова провести вертикальную прямую до пересечения с кривой $v = f(z)$, и т. д. В результате получим ступенчатую линию («лестницу»), которая определяет на оси абсцисс последовательность точек пересечения изображающей точкой полуоси $z > 0$ (рис. 10.5.2). Рассматривая это построение, убеждаемся в том, что устойчивым предельным циклом и положению равновесия соответствуют такие точки пересечения кривой точечного преобразования $v = f(z)$ с прямой $v = z$, в которых кривая $v = f(z)$ имеет меньший наклон к оси абсцисс, чем прямая $v = z$. Неустойчивым предельным циклом и положению равновесия соответствуют такие точки пересечения кривой $v = f(z)$ с прямой $v = z$, в которых кривая $v = f(z)$ имеет больший наклон к оси абсцисс, чем прямая $v = z$. На рис. 10.5.2 кривая $v = f(z)$ имеет две точки пересечения с прямой $v = z$, соответ-

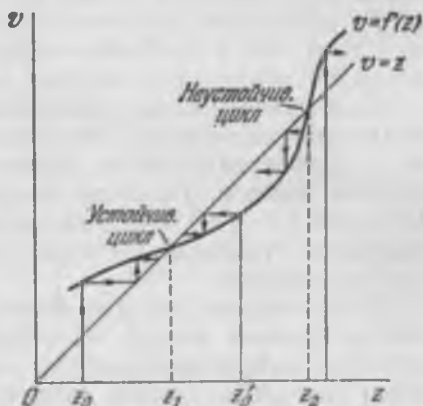


Рис. 10.5.2.

ствующие двум предельным циклам с амплитудами колебаний скорости z_1 и z_2 , причем первый предельный цикл соответствует устойчивым автоколебаниям, а второй — неустойчивым.

Если кривая $v = f(z)$ и прямая $v = z$ не пересекаются, то предельного цикла не существует и автоколебания в системе невозможны.

Во многих задачах фазовая плоскость оказывается симметричной относительно оси абсцисс и оси ординат. В таких случаях удобно рассматривать преобразование полуоси $z > 0$ в полуось $z < 0$ при полубоюде изображающей точки вокруг начала координат, которому также соответствует некоторая функция $v = f(z)$.

При практическом применении метода точечных отображений исходными данными для исследования нелинейной системы являются ее дифференциальные уравнения и характеристики ее нелинейных звеньев, заданные аналитически или экспериментально. Функция $v = f(z)$ определяется путем интегрирования дифференциальных уравнений системы, так же как и при построении фазовых траекторий.

В тех случаях, когда на фазовой плоскости есть линии переключения, удобнее вместо преобразования полуоси $z > 0$ изучать преобразование какой-нибудь линии переключения саму в себя. Если фазовая плоскость симметрична относительно осей координат, то удобно изучать отображение точек линии переключения в положительной полуплоскости в соответствующие точки линии переключения в отрицательной полуплоскости.

Пример 10.5.1. Применим метод точечных отображений к нелинейной системе примера 10.4.2, имеющей двухпозиционное реле с зоной неоднозначности. Фазовая плоскость такой системы симметрична в том смысле, что любая фазовая траектория, начинающаяся на одной линии переключения, при повороте вокруг начала координат на 180° совпадает с фазовой траекторией, начинающейся на таком же расстоянии от начала координат на другой линии переключения (рис. 10.4.18). Следовательно, в данном случае удобно рассматривать точечное отображение линии переключения $m - m$ в линию переключения $n - n$. Для определения функции $v = f(z)$ достаточно рассмотреть фазовую траекторию, выходящую из точки (d, z) линии переключения $m - m$ и пересекающую линию переключения $n - n$ в некоторой точке $(-d, -v)$. Для этого следует заменить в уравнении (10.4.21) величины y_0, z_0, y, z соответственно величинами $d, z, -d, -v$. Тогда после несложных преобразований получаем уравнение кривой $v = f(z)$ в неявном виде:

$$(1 - \theta) e^{\theta + \Delta} = (1 + \zeta) e^{-\zeta}, \quad \zeta = \frac{z}{kl}, \quad \theta = \frac{v}{kl}, \quad \Delta = \frac{2d}{klT}. \quad (10.5.1)$$

Для построения кривой $v = f(z)$ по уравнению (10.5.1) можно построить графики функций

$$F_0(\zeta) = (1 + \zeta) e^{-\zeta}, \quad F_1(\theta) = (1 - \theta) e^{\theta + \Delta}, \quad (10.5.2)$$

откладывая аргументы ζ и θ в одном и том же масштабе по оси абсцисс (рис. 10.5.3). По этим графикам точки кривой $v = f(z)$ находятся следующим приемом: задаваясь значением z_1 величины z , проводим из точки $\zeta_1 =$

z_1/kl оси абсцисс вертикальную прямую до пересечения с кривой $F_0(\zeta)$; из полученной точки пересечения проводим горизонтальную прямую до пересечения с кривой $F_1(\theta)$; абсцисса θ_1 этой точки пересечения, умноженная на kl , будет значением v_1 функции $f(z)$ при $z = z_1$.

Построив кривую $v = f(z)$ (рис. 10.5.4), видим, что рассматриваемая система имеет один устойчивый предельный цикл с амплитудой колебаний скорости, равной $z^* = kl\zeta^*$.

Заметим, что в случае неявного уравнения кривой точечного отображения типа $F_1(\theta) = F_0(\zeta)$ нет необходимости строить эту кривую, а достаточно построить графики левой и правой частей определяющего ее уравнения (рис. 10.5.3). Точки пересечения кривых $F_0(\zeta)$ и $F_1(\theta)$ определяют предельные циклы, поскольку в этих точках $v = z$. Предельный цикл устойчив,

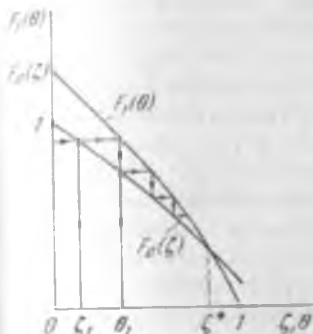


Рис. 10.5.3.

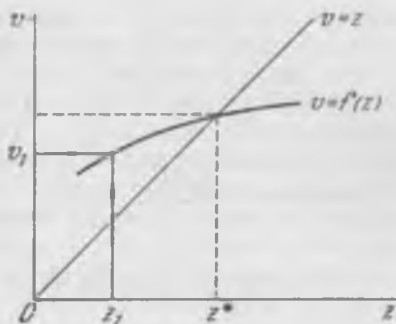


Рис. 10.5.4.

если в точке пересечения кривых $F_0(\zeta)$ и $F_1(\theta)$ кривая $F_0(\zeta)$ имеет меньший по абсолютной величине угол наклона к оси абсцисс, и неустойчив, если кривая $F_0(\zeta)$ имеет больший по абсолютной величине угол наклона к оси абсцисс.

Определив амплитуду z^* колебаний скорости выходной переменной системы, можно найти и все другие параметры автоколебаний — амплитуду колебаний выходной переменной y^* и период T_0 (а следовательно, и частоту). Для этого положим в формуле (10.4.23) $t = T_0/2$, $z_0 = -z^*$, $z = z^*$, $v = 1$. Тогда получим

$$z^* = -z^*e^{-\frac{T_0}{2T}} + kl\left(1 - e^{-\frac{T_0}{2T}}\right). \quad (10.5.3)$$

Отсюда находим период автоколебаний:

$$T_0 = 2T \ln \frac{kl + z^*}{kl - z^*}. \quad (10.5.4)$$

Частота автоколебаний равна $\omega_0 = 2\pi/T_0$. Амплитуду y^* колебаний выходной переменной можно определить по формуле (10.4.19), положив в ней $v = 1$, $y_0 = -d$, $z_0 = -z^*$, $y = -y^*$, $z = 0$. Тогда получим

$$y^* = d + T \left[z^* - kl \ln \left(1 + \frac{z^*}{kl} \right) \right]. \quad (10.5.5)$$

Полагая, например, $d = 0.1$; $T = 0.1$; $k = 1$; $l = 1$, вычисляем величину Δ по последней формуле (10.5.1) и строим графики функций $F_0(\zeta)$ и $F_1(\theta)$

(рис. 10.5.5) в соответствии с формулами (10.5.2). Точка их пересечения дает амплитуду автоколебаний скорости $z^* = k\zeta^* \approx 0,96$. После этого по формулам (10.5.4) определяем период и частоту автоколебаний: $T_0 \approx 0,36$ с, $\omega_0 \approx 17,5$ с⁻¹. Наконец, по формуле (10.5.5) определяем амплитуду автоколебаний: $y^* \approx 0,167$.

Метод исследования нелинейных систем с помощью фазовой плоскости и метод точечных отображений очень наглядны и сравнительно просты для систем второго порядка. Для систем третьего порядка эти методы становятся уже значительно более сложными. Для систем выше третьего порядка эти точные методы исследования

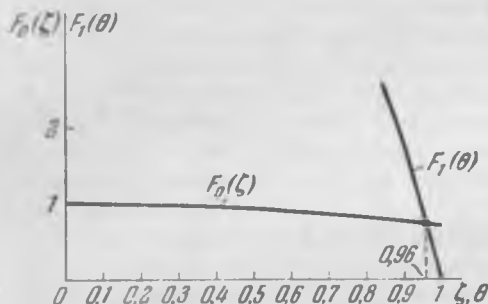


Рис. 10.5.5.

нелинейных систем практически неприменимы. Между тем на практике в подавляющем большинстве случаев приходится иметь дело с системами, поведение которых описывается дифференциальными уравнениями высоких порядков. Поэтому необходимы другие методы исследования нелинейных систем.

Для исследования периодических режимов в ре-

лейных системах, содержащих один релейный элемент и произвольную стационарную линейную часть, разработан специальный точный метод, основанный на представлении передаточной функции линейной части системы в виде суммы элементарных дробей и вычислении реакции линейной части системы на последовательность прямоугольных знакопеременных импульсов. В результате получаются трансцендентные уравнения относительно периода колебаний, зависящие от типа реле и вида передаточной функции линейной части. Этот метод дает возможность вычислить точное значение периода автоколебаний, после чего могут быть найдены амплитудные значения периодически изменяющихся переменных [73].

В связи с трудностью точного исследования нелинейных систем большое значение имеют приближенные методы. Пожалуй, наиболее эффективным и вследствие этого наиболее распространенным приближенным методом исследования нелинейных систем является метод гармонической линеаризации. Для систем, близких к линейным, часто применяется так называемый метод малого параметра, основанный на разложении неизвестных элементов движения системы в степенные ряды относительно малого параметра [35]. Этот метод дает хорошие результаты в случае, когда совокупность нелинейных членов в дифференциальных уравнениях системы

мала по сравнению с каждым из линейных членов. Малым параметром в этом случае является искусственно вводимый или естественно присутствующий в уравнениях системы общий постоянный множитель всех нелинейных членов.

Мы ограничимся здесь изложением метода гармонической линеаризации.

§ 10.6. Приближенное исследование нелинейных систем методом гармонической линеаризации

Метод гармонической линеаризации в большинстве случаев позволяет достаточно просто и с достаточной точностью исследовать стационарные нелинейные системы. Этот метод дает возможность оценивать устойчивость нелинейных систем, определять амплитуду и частоту автоколебаний, а также выбирать корректирующие цепи, обеспечивающие приемлемые или заданные характеристики нелинейных систем. При применении метода гармонической линеаризации предполагается существование в системе сигнала, близкого к синусоидальному, и на основании этого производится гармоническая линеаризация характеристик нелинейных звеньев методом, изложенным в § 8.4. Возможность применения этого метода к стационарным нелинейным системам в основном определяется близостью периодического движения системы к синусоидальному. Это условие обычно удовлетворяется в случае, когда линейные части системы являются фильтрами низких частот, т. е. хорошо отфильтровывают обертоны.

Как и раньше, будем считать, что стационарная система состоит из стационарных линейных систем и элементарных безынерционных нелинейных звеньев. Передаточные функции линейных частей системы и характеристики нелинейных звеньев будем считать известными. Для нахождения параметров автоколебаний системы методом гармонической линеаризации выразим входные и выходные переменные всех линейных частей и нелинейных звеньев системы в виде сумм неизвестных постоянных составляющих и синусоидальных составляющих с неизвестными амплитудами и начальными фазами и неизвестной общей для всех переменных частотой. При этом начальную фазу одной из переменных (обычно интересующей нас выходной переменной системы) можно считать равной нулю. Тогда каждая из входящих в уравнения системы переменных будет зависеть от трех неизвестных параметров: одна — от постоянной составляющей, амплитуды и частоты, а каждая из остальных — от постоянной составляющей, амплитуды и начальной фазы. Если произвести гармоническую линеаризацию характеристик всех нелинейных звеньев и сравнить в левой и правой частях уравнений системы постоянные составляющие, амплитуды и фазы синусоидальных составляющих, то каждое

уравнение системы даст три уравнения. В результате получим столько уравнений, сколько имеется неизвестных постоянных. Решив эти уравнения, найдем частоту автоколебаний, постоянные составляющие, амплитуды и начальные фазы всех переменных, входящих в уравнения системы. Если полученные уравнения не имеют решения, то автоколебания в системе невозможны.

Покажем детально, как применяется изложенный общий метод в случае системы с одним нелинейным звеном с характеристикой Φ и линейной частью с передаточной функцией $\Phi(s)$ (рис. 10.6.1). При этом совершенно безразлично, где включено нелинейное звено. В частности, оно может быть включено в прямой цепи, до или после линейной части.

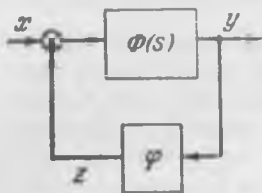


Рис. 10.6.1.

Предполагая в нелинейной системе смещенные автоколебания при $x = 0$, представим все переменные в виде суммы постоянных составляющих и синусоидальных слагаемых с неизвестными амплитудой, частотой и начальной фазой одного из них:

$$y = y^* + a_y \sin \omega_0 t, \quad z = z^* + a_z \sin (\omega_0 t + \psi_z). \quad (10.6.1)$$

Осуществляя гармоническую линейризацию нелинейного звена и пользуясь формулами (8.4.8), (8.4.9), (8.4.12), представим переменную z в виде

$$z = \Phi^* + |\Phi_n(a_y, y^*)| a_y \sin (\omega_0 t + \psi_n), \quad (10.6.2)$$

где $\Phi_n(a_y, y^*)$ — комплексный гармонический коэффициент усиления, а $\psi_n = \arg \Phi_n(a_y, y^*)$. Выходную переменную y линейной части системы в силу принципа суперпозиции выразим через переменную z в следующем виде:

$$y = -\Phi(0) z^* - |\Phi(i\omega_0)| a_z \sin [\omega_0 t + \psi_z + \psi(\omega_0)]. \quad (10.6.3)$$

Приравнивая правые части в формулах (10.6.1) и (10.6.2) для z и в формулах (10.6.1) и (10.6.3) для y , получим

$$z^* + a_z \sin (\omega_0 t + \psi_z) = \Phi^* + |\Phi_n(a_y, y^*)| a_y \sin (\omega_0 t + \psi_n), \quad (10.6.4)$$

$$y^* + a_y \sin \omega_0 t = -\Phi(0) z^* - |\Phi(i\omega_0)| a_z \sin [\omega_0 t + \psi_z + \psi(\omega_0)]. \quad (10.6.5)$$

Сравнивая в правых и левых частях уравнений (10.6.4) и (10.6.5) постоянные составляющие, а также амплитуды и фазы синусоидальных составляющих, получим уравнения

$$\left. \begin{aligned} \Phi^*(a_y, y^*) &= z^*, & \Phi(0) z^* &= -y^*, \\ |\Phi_n(a_y, y^*)| a_y &= a_z, & |\Phi(i\omega_0)| a_z &= a_y, \\ \psi_n(a_y, y^*) &= \psi_z, & \psi_z + \psi(\omega_0) &= \pi. \end{aligned} \right\} \quad (10.6.6)$$

Исключая из уравнений (10.6.6) переменные a_z, z^*, ψ_z , получим три уравнения для определения переменных y^*, a_y, ω_0 :

$$y^* = -\Phi(0) \Phi^*(a_y, y^*), \quad (10.6.7)$$

$$|\Phi(i\omega_0)| \cdot |\Phi_n(a_y, y^*)| = 1, \quad (10.6.8)$$

$$\arg \Phi(i\omega_0) + \arg \Phi_n(a_y, y^*) = \pi. \quad (10.6.9)$$

Решить систему уравнений (10.6.7), (10.6.8) и (10.6.9) в общем виде невозможно. Для решения этих уравнений можно применить различные приближенные методы и, в частности, метод последовательных приближений. Удобно также воспользоваться графическими способами. Рассмотрим возможные способы графического решения этой системы уравнений.

Первый способ состоит в следующем. Задаваясь значениями y^* , находим из уравнения (10.6.7) соответствующие значения a_y . По результатам этих вычислений строим на плоскости параметров y^*, a_y кривую 1 (рис. 10.6.2). Затем при определенном значении ω_0 задаемся значениями y^* и определяем из уравнений (10.6.8)

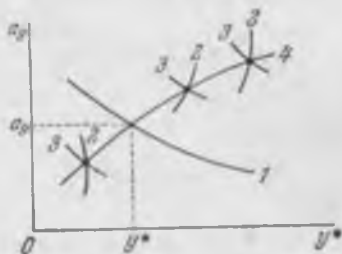


Рис. 10.6.2.

и (10.6.9) соответствующие значения a_y . По результатам этих вычислений строим кривую 2 в координатах y^*, a_y , соответствующую уравнению (10.6.8), и кривую 3, соответствующую уравнению (10.6.9). Точка пересечения кривых 2 и 3 является решением уравнений (10.6.8) и (10.6.9) для заданного значения ω_0 . Построив кривые 2 и 3 для ряда значений ω_0 , соединим точки их пересечения кривой 4. Точка пересечения кривой 4 с кривой 1 определяет значения y^* и a_y , удовлетворяющие всем трем уравнениям (10.6.7), (10.6.8) и (10.6.9). Для определения частоты автоколебаний ω_0 следует подставить найденные значения y^* и a_y в уравнение (10.6.8) или в уравнение (10.6.9). В результате получим уравнение для определения ω_0 . Частоту автоколебаний можно также определить по шкале значений ω_0 , получаемой на кривой 4 при ее построении по достаточно большому числу точек. Однако этот способ менее точен, чем предыдущий.

Второй способ аналогичен первому, но отличается только приемом отыскания кривой 4. Объединим уравнения (10.6.8) и (10.6.9) в одно уравнение

$$\Phi(i\omega_0) \Phi_n(a_y, y^*) = -1 \quad (10.6.10)$$

и перепишем его в виде

$$\Phi(i\omega_0) = -\frac{1}{\Phi_n(a_y, y^*)}. \quad (10.6.11)$$

уравнение системы даст три уравнения. В результате получим столько уравнений, сколько имеется неизвестных постоянных. Решив эти уравнения, найдем частоту автоколебаний, постоянные составляющие, амплитуды и начальные фазы всех переменных, входящих в уравнения системы. Если полученные уравнения не имеют решения, то автоколебания в системе невозможны.

Покажем детально, как применяется изложенный общий метод в случае системы с одним нелинейным звеном с характеристикой Φ

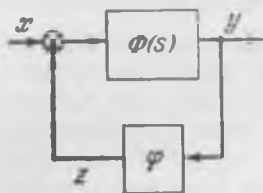


Рис. 10.6.1.

и линейной частью с передаточной функцией $\Phi(s)$ (рис. 10.6.1). При этом совершенно безразлично, где включено нелинейное звено. В частности, оно может быть включено в прямой цепи, до или после линейной части.

Предполагая в нелинейной системе смещенные автоколебания при $x=0$, представим все переменные в виде суммы постоянных составляющих и синусоидальных слагаемых с неизвестными амплитудой, частотой и начальной фазой одного из них:

$$y = y^* + a_y \sin \omega_0 t, \quad z = z^* + a_z \sin (\omega_0 t + \psi_z). \quad (10.6.1)$$

Осуществляя гармоническую линеаризацию нелинейного звена и пользуясь формулами (8.4.8), (8.4.9), (8.4.12), представим переменную z в виде

$$z = \varphi^* + |\Phi_n(a_y, y^*)| a_y \sin (\omega_0 t + \psi_n), \quad (10.6.2)$$

где $\Phi_n(a_y, y^*)$ — комплексный гармонический коэффициент усиления, а $\psi_n = \arg \Phi_n(a_y, y^*)$. Выходную переменную y линейной части системы в силу принципа суперпозиции выразим через переменную z в следующем виде:

$$y = -\Phi(0) z^* - |\Phi(i\omega_0)| a_z \sin [\omega_0 t + \psi_z + \psi(\omega_0)]. \quad (10.6.3)$$

Приравнивая правые части в формулах (10.6.1) и (10.6.2) для z и в формулах (10.6.1) и (10.6.3) для y , получим

$$z^* + a_z \sin (\omega_0 t + \psi_z) = \varphi^* + |\Phi_n(a_y, y^*)| a_y \sin (\omega_0 t + \psi_n), \quad (10.6.4)$$

$$y^* + a_y \sin \omega_0 t = -\Phi(0) z^* - |\Phi(i\omega_0)| a_z \sin [\omega_0 t + \psi_z + \psi(\omega_0)]. \quad (10.6.5)$$

Сравнивая в правых и левых частях уравнений (10.6.4) и (10.6.5) постоянные составляющие, а также амплитуды и фазы синусоидальных составляющих, получим уравнения

$$\left. \begin{aligned} \varphi^*(a_y, y^*) &= z^*, & \Phi(0) z^* &= -y^*, \\ |\Phi_n(a_y, y^*)| a_y &= a_z, & |\Phi(i\omega_0)| a_z &= a_y, \\ \psi_n(a_y, y^*) &= \psi_z, & \psi_z + \psi(\omega_0) &= \pi. \end{aligned} \right\} \quad (10.6.6)$$

Исключая из уравнений (10.6.6) переменные a_z, z^*, ψ_z , получим три уравнения для определения переменных y^*, a_y, ω_0 :

$$y^* = -\Phi(0) \varphi^*(a_y, y^*), \quad (10.6.7)$$

$$|\Phi(i\omega_0)| \cdot |\Phi_n(a_y, y^*)| = 1, \quad (10.6.8)$$

$$\arg \Phi(i\omega_0) + \arg \Phi_n(a_y, y^*) = \pi. \quad (10.6.9)$$

Решить систему уравнений (10.6.7), (10.6.8) и (10.6.9) в общем виде невозможно. Для решения этих уравнений можно применить различные приближенные методы и, в частности, метод последовательных приближений. Удобно также воспользоваться графическими способами. Рассмотрим возможные способы графического решения этой системы уравнений.

Первый способ состоит в следующем. Задаваясь значениями y^* , находим из уравнения (10.6.7) соответствующие значения a_y . По результатам этих вычислений строим на плоскости параметров y^*, a_y кривую 1 (рис. 10.6.2). Затем при определенном значении ω_0 задаемся значениями y^* и определяем из уравнений (10.6.8)

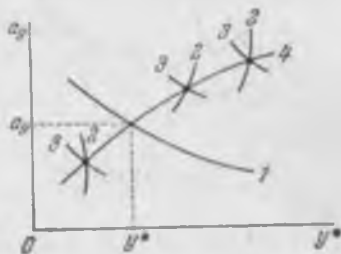


Рис. 10.6.2.

и (10.6.9) соответствующие значения a_y . По результатам этих вычислений строим кривую 2 в координатах y^*, a_y , соответствующую уравнению (10.6.8), и кривую 3, соответствующую уравнению (10.6.9). Точка пересечения кривых 2 и 3 является решением уравнений (10.6.8) и (10.6.9) для заданного значения ω_0 . Построив кривые 2 и 3 для ряда значений ω_0 , соединим точки их пересечения кривой 4. Точка пересечения кривой 4 с кривой 1 определяет значения y^* и a_y , удовлетворяющие всем трем уравнениям (10.6.7), (10.6.8) и (10.6.9). Для определения частоты автоколебаний ω_0 следует подставить найденные значения y^* и a_y в уравнение (10.6.8) или в уравнение (10.6.9). В результате получим уравнение для определения ω_0 . Частоту автоколебаний можно также определить по шкале значений ω_0 , получаемой на кривой 4 при ее построении по достаточно большому числу точек. Однако этот способ менее точен, чем предыдущий.

Второй способ аналогичен первому, но отличается только приемом отыскания кривой 4. Объединим уравнения (10.6.8) и (10.6.9) в одно уравнение

$$\Phi(i\omega_0) \Phi_n(a_y, y^*) = -1 \quad (10.6.10)$$

и перепишем его в виде

$$\Phi(i\omega_0) = -\frac{1}{\Phi_n(a_y, y^*)}. \quad (10.6.11)$$

Построим на комплексной плоскости (рис. 10.6.3) амплитудно-фазовую характеристику линейной части системы $\Phi(i\omega)$ и обратную амплитудно-фазовую характеристику нелинейного звена с обратным знаком $-1/\Phi_n(a_y, y^*)$ для разных значений y^* при

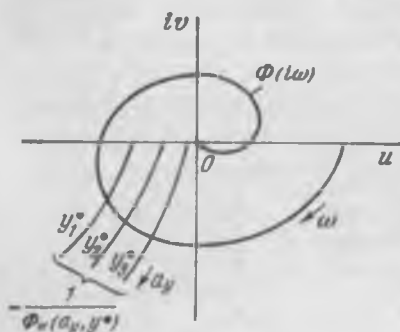


Рис. 10.6.3.

изменении ω и a_y от 0 до ∞ . Точки пересечения этих кривых определяют зависимость $a_y = f(y^*)$, соответствующую уравнению (10.6.11), т. е. кривую 4 (рис. 10.6.2). Далее определяем точку пересечения ее с кривой 1, построенной по уравнению (10.6.7), и находим ω_0 , как и в первом способе.

Наконец, для решения системы уравнений (10.6.7), (10.6.8), (10.6.9) можно воспользоваться методом логарифмических частотных характеристик. Для этого следует построить логарифмические частот-

ные характеристики системы с разомкнутой у сумматора цепью обратной связи для различных значений a_y при различных фиксированных значениях y^* :

$$L(\omega, a_y, y^*) = 20 \lg |\Phi(i\omega)| + 20 \lg |\Phi_n(a_y, y^*)|, \quad (10.6.12)$$

$$\psi(\omega, a_y, y^*) = \arg \Phi(i\omega) + \arg \Phi_n(a_y, y^*). \quad (10.6.13)$$

Значения a_y , y^* , ω , при которых $L(\omega, a_y, y^*) = 0$, а $\psi(\omega, a_y, y^*) = -\pi$ удовлетворяют уравнению (10.6.10), полученному в результате объединения уравнений (10.6.8) и (10.6.9), т. е. определяют кривую 4. Точка пересечения этой кривой с кривой 1 определяет решение уравнений (10.6.7), (10.6.8) и (10.6.9).

В частном случае, когда $\Phi^*(a_y, 0) = 0$, уравнение (10.6.7) удовлетворяется при $y^* = 0$, т. е. автоколебания в системе симметричны. Уравнения (10.6.8) и (10.6.9) в этом случае принимают вид

$$|\Phi(i\omega_0)| \cdot |\Phi_n(a_y)| = 1, \quad (10.6.14)$$

$$\psi(\omega_0) + \arg \Phi_n(a_y) - \pi \quad (10.6.15)$$

Из этих уравнений вытекает одно уравнение

$$\Phi(i\omega_0) \Phi_n(a_y) = -1. \quad (10.6.16)$$

Для его решения можно воспользоваться описанным уже способом построения на комплексной числовой плоскости амплитудно-фазовой характеристики линейной части $\Phi(i\omega)$ и обратной амплитудно-фазовой характеристики нелинейного звена с обратным знаком $-1/\Phi_n(a_y)$ (второй способ). Точка пересечения этих

двух кривых определяет частоту ω_0 и амплитуду a_y автоколебаний (рис. 10.6.4).

Можно также воспользоваться построением логарифмических частотных характеристик разомкнутой системы

$$L(\omega, a_y) = 20 \lg |\Phi(i\omega)| + 20 \lg |\Phi_n(a_y)|, \quad (10.6.17)$$

$$\psi(\omega, a_y) = \arg \Phi(i\omega) + \arg \Phi_n(a_y) \quad (10.6.18)$$

для различных значений a_y (рис. 10.6.5). Как следует из (10.6.16), значения a_y и ω , при которых $L(\omega, a_y) = 0$, $\psi(\omega, a_y) = -\pi$, являются искомыми амплитудой и частотой автоколебаний. Иными

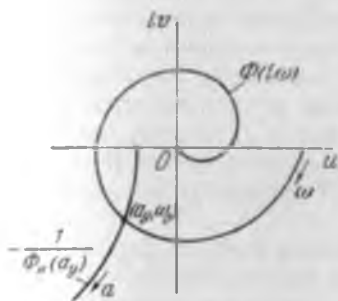


Рис. 10.6.4.

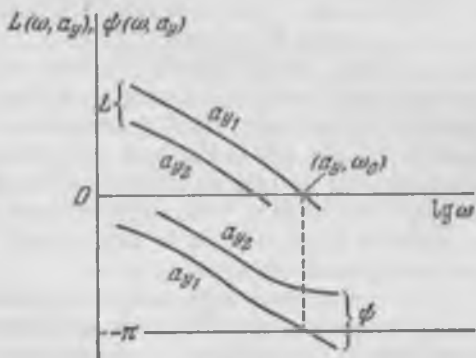


Рис. 10.6.5.

словами, если существует такое значение a_y , при котором фазовая характеристика (10.6.18) пересекает уровень $-\pi$ на частоте среза, т. е. при значении ω , при котором логарифмическая амплитудная характеристика (10.6.17) пересекает ось абсцисс, то в системе возможны автоколебания. Соответствующие значения амплитуды a_y и частоты среза ω_0 приблизительно равны амплитуде и частоте автоколебаний.

В частном случае, когда $q'(a_y) \equiv 0$, т. е. нелинейное звено имеет однозначную характеристику, $\Phi_n(a_y) = q(a_y)$ и из (10.6.16) следуют уравнения

$$q(a_y) P(\omega_0) = -1, \quad q(a_y) Q(\omega_0) = 0, \quad (10.6.19)$$

где $P(\omega)$ и $Q(\omega)$ — действительная и мнимая частотные характеристики линейной части системы. Сократив второе уравнение (10.6.19) на $q(a_y)$, получим уравнение для определения частоты автоколебаний:

$$Q(\omega_0) = 0. \quad (10.6.20)$$

Таким образом, частота автоколебаний совершенно не зависит от характеристики нелинейного звена и определяется только

свойствами линейной части системы. После решения уравнения (10.6.20), определяющего частоту автоколебаний, из первого уравнения (10.6.19) найдем

$$q(a_y) = -\frac{1}{P(\omega_0)}. \quad (10.6.21)$$

По этому значению $q(a_y)$, пользуясь графиком $q = q(a_y)$, определяем амплитуду автоколебаний a_y .

Если не существует действительных значений y^* и положительных значений a_y и ω_0 , удовлетворяющих уравнениям (10.6.7), (10.6.8) и (10.6.9), то автоколебания в системе невозможны.

Определив приближенно периодическое движение системы путем решения уравнений (10.6.7), (10.6.8) и (10.6.9), необходимо оценить его устойчивость, чтобы определить, соответствует ли оно реально существующим в системе автоколебаниям. Особенно важно это сделать, когда существует несколько периодических решений. Для исследования устойчивости периодического движения системы можно применить первый метод Ляпунова, изложенный в § 10.2. Для этого необходимо составить соответствующие уравнения в вариациях и исследовать устойчивость определяемой ими линейной системы.

Для приближенной оценки устойчивости автоколебаний можно также воспользоваться приближенными способами [35, 49], основанными на следующих соображениях. Движение системы, близкое к найденному периодическому, можно приближенно считать синусоидальным с медленно изменяющейся амплитудой u , приближающейся к амплитуде автоколебаний a_y в случае устойчивых автоколебаний и удаляющейся от a_y в случае неустойчивых автоколебаний. Если автоколебания в системе устойчивы и имеют амплитуду a_y , то при замене этой амплитуды в уравнении (10.6.10) величиной $u > a_y$ пара чисто мнимых корней $\pm i\omega_0$ должна замениться парой комплексных сопряженных корней с отрицательной действительной частью, а при замене a_y величиной $u < a_y$ корни $\pm i\omega_0$ должны замениться парой комплексных сопряженных корней с положительной действительной частью. Для проверки характера изменения корней $\pm i\omega_0$ при замене в уравнении (10.6.10) величины a_y величиной u удобнее всего пользоваться критерием Найквиста. Так как уравнение (10.6.10) имеет чисто мнимые корни $\pm i\omega_0$ при $u = a_y$, то амплитудно-фазовая характеристика разомкнутой системы

$$\Phi_p(i\omega, u, y^*) = \Phi(i\omega) \Phi_n(u, y^*) \quad (10.6.22)$$

при $u = a_y$ проходит через точку -1 плоскости комплексной переменной (кривая 1 на рис. 10.6.6). Если автоколебания системы устойчивы, то при замене величины a_y близкой к ней величиной $u > a_y$ кривая $\Phi_p(i\omega, u, y^*)$ должна сместиться в сторону умень-

шения запаздывания по фазе на частоте среза (кривая 2 на рисунке 10.6.6), а при замене амплитуды a_y близкой к ней величиной $u < a_y$ кривая $\Phi_p(i\omega, u, y^*)$ должна сместиться в сторону увеличения запаздывания по фазе на частоте среза (кривая 3 на рис. 10.6.6). В случае неустойчивых автоколебаний должна наблюдаться обратная картина. В случае полуустойчивого предельного цикла кривая $\Phi_p(i\omega, u, y^*)$ смещается в одну и ту же сторону при $u > a_y$ и при $u < a_y$.

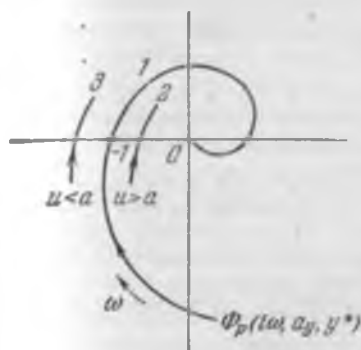


Рис. 10.6.6.

Пример 10.6.1. Определить автоколебания в следящей системе (рис. 10.6.7), имеющей линейную часть с передаточной функцией

$$\Phi(s) = \frac{k}{s(T_1 s + 1)(T_2 s + 1)}$$

и идеальное реле с характеристикой $\varphi(v) = l \operatorname{sgn} v$. Так как релейное звено

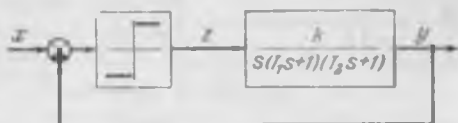


Рис. 10.6.7.

обладает нечетной характеристикой, то согласно изложенному $y^* = 0$ и для определения частоты ω_0 и амплитуды a_y автоколебаний имеем уравнение (10.6.16), которое в данном случае принимает вид

$$\frac{k}{\omega_0 (T_1 i \omega_0 + 1)(T_2 i \omega_0 + 1)} \Phi_n(a_y) = -1. \quad (10.6.23)$$

Но для идеального реле $q'(a_y) = 0$, $q(a_y) = 4l/\pi a_y$ (приложение 6, нелинейность 5). Следовательно, уравнение (10.6.23) может быть переписано в виде

$$\frac{4kl}{\pi a_y} + i\omega_0 (T_1 i \omega_0 + 1)(T_2 i \omega_0 + 1) = 0. \quad (10.6.24)$$

Разделяя в этом уравнении действительную и мнимую части, получим два уравнения для определения a_y и ω_0 :

$$\frac{4kl}{\pi a_y} - \omega_0^2 (T_1 + T_2) = 0, \quad \omega_0 - T_1 T_2 \omega_0^3 = 0, \quad (10.6.25)$$

которые, очевидно, равноценны уравнениям (10.6.19). Из второго уравнения (10.6.25) находим частоту автоколебаний:

$$\omega_0 = \frac{1}{\sqrt{T_1 T_2}}. \quad (10.6.26)$$

Подставляя эту величину в первое уравнение (10.6.25) и решая его относительно a_y , находим амплитуду автоколебаний:

$$a_y = \frac{4}{\pi} l k \frac{T_1 T_2}{T_1 + T_2}. \quad (10.6.27)$$

Оценим устойчивость найденного периодического режима. Для этого воспользуемся приближенным способом оценки изменения частотной характеристики разомкнутой системы при изменении параметра a_y . Частотная характеристика разомкнутой цепи при замене амплитуды a_y величиной u определяется формулой (10.6.22), которая в данном случае имеет вид

$$\Phi_p(i\omega, u) = \frac{4ki}{\pi i [i(1 - T_1 T_2 \omega^2) \omega - (T_1 + T_2) \omega^2]} \quad (10.6.28)$$

Так как $\Phi_p(i\omega_0, a_y) = -1$, то, полагая в (10.6.28) $\omega = \omega_0$, получим

$$\Phi_p(i\omega_0, u) = -\frac{a_y}{u}. \quad (10.6.29)$$

Отсюда видно, что если $u > a_y$, то $\Phi_p(i\omega_0, u) > -1$, а если $u < a_y$, то $\Phi_p(i\omega_0, u) < -1$. Следовательно, найденное периодическое движение системы устойчиво.

Пример 10.6.2. Определить автоколебания в следующей системе примеров 10.4.2 и 10.5.1 (рис. 10.4.14) при $T_1 = 0$.

В этом случае уравнение (10.6.16) для определения частоты ω_0 и амплитуды автоколебаний a_y принимает вид

$$\frac{k}{i\omega_0(Ti\omega_0 + 1)} \Phi_n(a_y) = -1. \quad (10.6.30)$$

Комплексный гармонический коэффициент усиления $\Phi_n(a_y)$ в данном случае равен (приложение 6, нелинейность 8)

$$\Phi_n(a_y) = \frac{4l}{\pi a_y} \sqrt{1 - \frac{a^2}{a_y^2}} - i \frac{4ld}{\pi a_y^3}. \quad (10.6.31)$$

Для определения амплитуды a_y и частоты ω_0 автоколебаний построим амплитудно-фазовую характеристику линейной части $\Phi(i\omega) = \frac{k}{i\omega(Ti\omega + 1)}$

и обратную амплитудно-фазовую характеристику нелинейного звена с обратным знаком. На рис. 10.6.8 представлены эти кривые для числовых данных примера 10.5.1. Точка пересечения их определяет приближенные значения амплитуды и частоты автоколебаний в системе: $a_y \approx 0,165$, $\omega_0 \approx 17,0$. Сравнение этих результатов с данными вычислений точным способом в примере 10.5.1 показывает хорошее их совпадение: значение амплитуды совпадает с точностью до 1%, а значение частоты — с точностью до 3%.

Устойчивость автоколебаний может быть проверена тем же способом, что и в предыдущем примере.

§ 10.7. Влияние случайных возмущений на параметры автоколебаний стационарных нелинейных систем

Изложенный в предыдущем параграфе метод определения параметров автоколебаний нелинейных систем дает достаточно высокую точность, если уровень случайных возмущений, действующих на систему, достаточно мал по сравнению с амплитудами колебаний

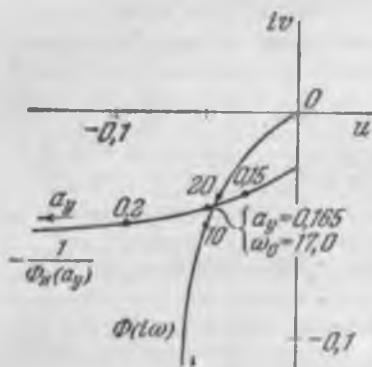


Рис. 10.6.8.

всех переменных, входящих в уравнения, описывающие поведение системы. Если это условие не соблюдено, то изложенный метод может привести к значительным ошибкам как при определении параметров автоколебаний, так и при решении вопроса о возможности автоколебаний в системе. Случайные возмущения в этом случае существенно влияют на поведение системы. Поэтому при значительном уровне случайных возмущений (шумов и помех) необходимо учитывать эти случайные возмущения при исследовании нелинейной системы. В частности, параметры автоколебаний системы необходимо определять с учетом случайных возмущений. Это можно сделать, применяя метод совместной гармонической и статистической линеаризации, изложенный в § 8.6. Для применения этого метода необходимо каждую из переменных, входящих в уравнения стационарной системы, представить в виде суммы неизвестной постоянной составляющей, синусоидальной составляющей с неизвестными амплитудой, начальной фазой и частотой и центрированной случайной составляющей с неизвестной дисперсией. Выполняя совместную гармоническую и статистическую линеаризацию всех нелинейных звеньев системы и сравнивая постоянные составляющие, а также амплитуды и фазы синусоидальных составляющих в левых и правых частях всех уравнений системы, получим, как и в предыдущем параграфе, уравнения для определения неизвестных параметров регулярных составляющих всех входящих в уравнения системы переменных. Однако на основании изложенного в § 8.6 эти уравнения будут содержать еще неизвестные дисперсии входных сигналов нелинейных звеньев. Для составления необходимых дополнительных уравнений предположим, что действующие на систему случайные возмущения являются стационарными случайными функциями времени. Тогда достаточно сравнить дисперсии левых и правых частей всех уравнений системы, пользуясь при этом известными формулами для дисперсий линейных функций случайных величин и формулой (7.4.4) для дисперсий выходных переменных стационарных линейных систем. Решив полученные в результате уравнения, мы найдем частоту автоколебаний, постоянные составляющие, амплитуды и начальные фазы синусоидальных составляющих и дисперсии всех переменных, входящих в уравнения системы.

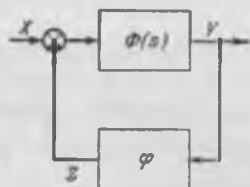


Рис. 10.7.1.

Покажем подробно применение изложенного общего метода на примере стационарной системы с одним нелинейным звеном с характеристикой φ и стационарной линейной частью с передаточной функцией $\Phi(s)$ (рис. 10.7.1). При этом так же, как и в § 10.6, нелинейное звено может быть включено и в прямую цепь, до или после линейной части. На систему действует помеха,

представляющая собой стационарную случайную функцию времени $X(t)$.

Предполагая в системе наличие смещенных автоколебаний, представим переменные Y и Z в виде сумм постоянных составляющих, синусоидальных составляющих с неизвестными амплитудами, частотой и начальной фазой для одной из них, а также случайных центрированных составляющих:

$$\left. \begin{aligned} Y &= y^* + a_y \sin \omega_0 t + Y^0, \\ Z &= z^* + a_z \sin (\omega_0 t + \psi_z) + Z^0. \end{aligned} \right\} \quad (10.7.1)$$

Применяя совместную статистическую и гармоническую линеаризацию к нелинейному звену, как изложено в § 8.6, получим приближенное равенство

$$Z \approx \varphi_0^* + |\Psi_n(a_y, y^*, \sigma_y)| a_y \sin (\omega_0 t + \psi_n) + \kappa_1 Y^0. \quad (10.7.2)$$

Так как математическое ожидание входного сигнала линейной части системы $X - Z$ в данном случае представляет собой разность постоянной и синусоидального сигнала, то для определения математического ожидания выходной переменной Y линейной части системы можно применить формулу (7.3.26). В результате получим

$$m_y = \Phi(0)(m_x - z^*) - |\Phi(i\omega_0)| a_z \sin [\omega_0 t + \psi_z + \psi(\omega_0)]. \quad (10.7.3)$$

Приравнявая математические ожидания правых частей формул (10.7.1) соответственно правой части формулы (10.7.3) и математическому ожиданию правой части формулы (10.7.2), получим

$$y^* + a_y \sin \omega_0 t = \Phi(0)(m_x - z^*) - |\Phi(i\omega_0)| a_z \sin [\omega_0 t + \psi_z + \psi(\omega_0)], \quad (10.7.4)$$

$$z^* + a_z \sin (\omega_0 t + \psi_z) = \varphi_0^* + |\Psi_n(a_y, y^*, \sigma_y)| a_y \sin (\omega_0 t + \psi_n). \quad (10.7.5)$$

Сравнивая постоянные составляющие, амплитуды и фазы синусоидальных составляющих в формулах (10.7.4), (10.7.5), получим систему уравнений

$$y^* = \Phi(0)[m_x - \varphi_0^*(a_y, y^*, \sigma_y)], \quad (10.7.6)$$

$$|\Phi(i\omega_0)| \cdot |\Psi_n(a_y, y^*, \sigma_y)| = 1, \quad (10.7.7)$$

$$\arg \Psi_n(a_y, y^*, \sigma_y) + \arg \Phi(i\omega_0) = \pi. \quad (10.7.8)$$

Для составления четвертого уравнения заметим, что формула (10.7.2) заменяет нелинейное звено по отношению к случайной составляющей Y^0 линейным усилителем с коэффициентом усиления κ_1 . Поэтому для нахождения установившейся дисперсии σ_y^2 выходного сигнала системы Y можно применить формулу (7.4.4).

Для этого находим сначала по формуле (4.5.7) передаточную функцию полученной линейной системы (рис. 10.7.2):

$$\Phi_1(s) = \frac{\Phi(s)}{1 + \chi_1(a_y, y^*, \sigma_y) \Phi(s)}. \quad (10.7.9)$$

Подставляя это выражение в (7.4.4), получим следующую формулу для дисперсии выходной переменной системы:

$$\sigma_y^2 = \int_{-\infty}^{\infty} \left| \frac{\Phi(i\omega)}{1 + \chi_1(a_y, y^*, \sigma_y) \Phi(i\omega)} \right|^2 s_x(\omega) d\omega, \quad (10.7.10)$$

где $s_x(\omega)$ — спектральная плотность случайной функции $X(t)$.

Таким образом, мы получили четыре уравнения (10.7.6), (10.7.7), (10.7.8) и (10.7.10) для определения четырех неизвестных y^* , a_y , ω_0 , σ_y . Если спектральная плотность $s_x(\omega)$ представляет собой дробно-рациональную функцию, то интеграл в формуле (10.7.10) может быть выражен в конечном виде (приложение 4). Заметим еще, что уравнения (10.7.7) и (10.7.8) можно записать в виде одного уравнения

$$\Phi(i\omega_0) \Psi_n(a_y, y^*, \sigma_y) = -1. \quad (10.7.11)$$

Для решения уравнений (10.7.6), (10.7.7), (10.7.8) и (10.7.10) можно применить следующий прием. Зададимся рядом значений σ_y и для каждого из этих значений σ_y решим уравнения (10.7.6), (10.7.7) и (10.7.8) любым из методов, изложенных в предыдущем параграфе. В результате получим y^* , a_y и ω_0 как функции σ_y (т. е. уровня шумов на входе нелинейного звена). После этого можно будет определить для каждого значения σ_y коэффициент $\chi_1(a_y, y^*, \sigma_y)$ и вычислить правую часть уравнения (10.7.10). По этим данным можно построить кривую зависимости правой части уравнения (10.7.10) от σ_y . Точка пересечения этой кривой с параболой, изображающей зависимость левой части уравнения (10.7.10) от σ_y , определяет значение σ_y , при котором удовлетворяются все уравнения (10.7.6), (10.7.7), (10.7.8) и (10.7.10). Соответствующие этому значению σ_y значения y^* , a_y и ω_0 представляют собой искомые постоянную составляющую, амплитуду и частоту автоколебаний.

Если не существует решения уравнений (10.7.6), (10.7.7), (10.7.8) и (10.7.10) с положительными значениями a_y и ω_0 , то автоколебания в системе невозможны.

Определив изложенным методом параметры автоколебаний с учетом помех, необходимо оценить их устойчивость, что может быть сделано, например, так, как показано в предыдущем параграфе.

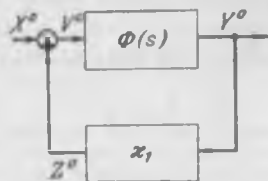


Рис. 10.7.2.

В частном случае однозначной характеристики Φ нелинейного звена $\kappa'(a_y, y^*, \sigma_y) \equiv 0$, $\Psi_n(a_y, y^*, \sigma_y) = \kappa(a_y, y^*, \sigma_y)$ и из уравнения (10.7.11) вытекает уравнение (10.6.20) для определения частоты автоколебаний. Таким образом, в случае однозначной характеристики нелинейного звена частота автоколебаний ω_0 не зависит ни от параметров нелинейного звена, ни от уровня шумов в системе и определяется только свойствами линейной части системы. После нахождения ω_0 величины y^* , a_y и σ_y определяются совместным решением уравнений (10.7.6), (10.7.10) и уравнения

$$\kappa(a_y, y^*, \sigma_y) P(\omega_0) = -1, \quad (10.7.12)$$

вытекающего в этом случае из (10.7.11). Если, кроме того, $m_x = 0$ и $\Phi_0^*(a_y, 0, \sigma_y) \equiv 0$, то уравнение (10.7.6) удовлетворяется при $y^* = 0$ и в системе возможны симметричные автоколебания. В этом случае после нахождения ω_0 из уравнения (10.6.20) величины a_y и σ_y определяются совместным решением уравнений (10.7.10) и (10.7.12).

Пример 10.7.1. Определить автоколебания и условия их существования в релейной системе, рассмотренной в примере 10.6.1 (рис. 10.6.7), при стационарном случайном возмущении $X(t)$, имеющем равное нулю математическое ожидание и корреляционную функцию

$$k_x(\tau) = D e^{-\alpha|\tau|}.$$

В этом случае постоянная составляющая сигнала равна нулю и уравнение (10.7.11) принимает вид

$$\frac{k}{i\omega_0(T_1 i\omega_0 + 1)(T_2 i\omega_0 + 1)} \kappa(a_y, \sigma_y) = -1, \quad (10.7.13)$$

где a_y — амплитуда колебаний входного сигнала реле, а σ_y^2 — его дисперсия. Коэффициент $\kappa(a_y, \sigma_y)$ для идеального реле равен

$$\kappa(a_y, \sigma_y) = \frac{l}{a_y} B_0(\lambda), \quad (10.7.14)$$

$$\lambda = \frac{a_y}{\sigma_y \sqrt{2}}.$$

В приложении 6 приведена формула, определяющая функцию $B_0(\lambda)$:

$$B_0(\lambda) = \frac{2\lambda}{\sqrt{\pi}} \left(1 - \frac{1}{4} \lambda^2 + \frac{1}{16} \lambda^4 - \dots \right), \quad (10.7.15)$$

и график для определения ее значений. Отделяя в уравнении (10.7.13) действительную и мнимую части, получим

$$k\kappa(a_y, \sigma_y) - (T_1 + T_2) \omega_0^2 = 0, \quad (10.7.16)$$

$$\omega_0 - T_1 T_2 \omega_0^3 = 0. \quad (10.7.17)$$

Из уравнения (10.7.17) определяем частоту автоколебаний, которая определяется той же формулой (10.6.26), что и при отсутствии случайных

возмущений. Из уравнения (10.7.16), учитывая (10.6.26), получаем

$$\kappa(a_v, \sigma_v) = \frac{1}{k} \cdot \frac{T_1 + T_2}{T_1 T_2} \quad (10.7.18)$$

Это уравнение содержит две неизвестные величины a_v и σ_v . Чтобы получить дополнительное уравнение, запишем выражение для дисперсии величины V на входе в нелинейный элемент:

$$\sigma_v^2 = \frac{D\alpha}{\pi} \int_{-\infty}^{\infty} \left| \frac{i\omega (T_1 i\omega + 1) (T_2 i\omega + 1)}{i\omega (T_1 i\omega + 1) (T_2 i\omega + 1) + k\kappa_1(a_v, \sigma_v)} \right|^2 \frac{d\omega}{\omega^2 + \alpha^2} \quad (10.7.19)$$

где коэффициент $\kappa_1(a_v, \sigma_v)$ для идеального реле равен

$$\kappa_1(a_v, \sigma_v) = \frac{l}{\sigma_v} C_0(\lambda),$$

$$\lambda = \frac{a_v}{\sigma_v \sqrt{2}} \quad (10.7.20)$$

Функция $C_0(\lambda)$ приведена в приложении 6. Вычислив интеграл в формуле (10.7.19) с помощью таблицы (приложение 4), получим

$$\sigma_v^2 = D\alpha \frac{b_0(-a_1 a_4 + a_2 a_3) - a_0(a_3 b_1 + a_1)}{a_0(a_0 a_3^2 + a_1^2 a_4 - a_1 a_2 a_3)} \quad (10.7.21)$$

где $a_0 = T_1 T_2$, $a_1 = T_1 + T_2 + T_1 T_2 \alpha$, $a_2 = (T_1 + T_2) \alpha + 1$, $a_3 = k\kappa_1 + \alpha$, $a_4 = k\alpha\kappa_1$, $b_0 = -T_1^2 T_2^2$, $b_1 = T_1^2 + T_2^2$. В силу зависимости коэффициента κ_1 от a_v и σ_v равенство (10.7.21) является уравнением, связывающим величины a_v и σ_v . Для определения амплитуды автоколебаний a_v и среднего квадратического отклонения случайного возмущения на входе нелинейного элемента надо совместно решить два уравнения (10.7.18), (10.7.21). Практически удобно воспользоваться графическим способом решения этих уравнений. Для этого, задаваясь значениями σ_v , по уравнению (10.7.18) определяем соответствующие значения a_v . В результате на плоскости параметров a_v , σ_v можно построить кривую 1 (рис. 10.7.3). Далее для значений a_v и соответствующих значений σ_v , снятых с кривой 1 на рис. 10.7.3, вычисляем правую часть уравнения (10.7.21):

$$\zeta = D\alpha \frac{b_0(-a_1 a_4 + a_2 a_3) - a_0(a_3 b_1 + a_1)}{a_0(a_0 a_3^2 + a_1^2 a_4 - a_1 a_2 a_3)} \quad (10.7.22)$$

и строим кривую в координатах a_v , σ_v , откладывая значение ζ по оси σ_v (кривая 2 на рис. 10.7.3). Точка пересечения полученных двух кривых определяет искомые значения a_v , σ_v , так как эти значения a_v , σ_v удовлетворяют обоим уравнениям (10.7.18) и (10.7.21). Амплитуда колебаний выходной переменной системы в данном случае, очевидно, равна a_v ; $a_v = a_v$. Результаты расчетов показывают, что амплитуда автоколебаний получается меньше, чем при отсутствии случайных возмущений.

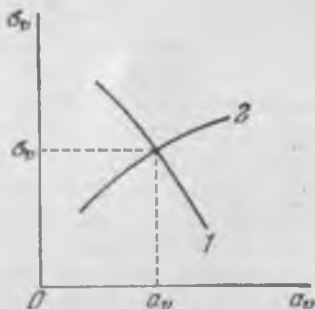


Рис. 10.7.3.

Формула (10.7.18) позволяет также определить значения величины σ_0 , при которых автоколебания в системе становятся невозможными. Можно также определить значения параметров системы, в частности коэффициента k , при которых автоколебания в системе в присутствии случайных возмущений определенного уровня становятся невозможными. Для решения этих вопросов заметим, что $B_0^{(1)}(\lambda) = B_0(\lambda)/\lambda$ является убывающей функцией, т. е. $B_0^{(1)}(\lambda) < B_0^{(1)}(0)$ при любом $\lambda > 0$. Формулы (10.7.14) показывают, что вследствие этого имеет место неравенство

$$\kappa(\sigma_0, \sigma_0) < \kappa(0, \sigma_0) = \frac{1}{\sigma_0} \sqrt{\frac{2}{\pi}}, \quad (10.7.23)$$

справедливое при любых значениях $\sigma_0 > 0$.

Из (10.7.18) и (10.7.23) вытекает следующее необходимое условие возможности автоколебаний в системе:

$$\frac{T_1 + T_2}{kT_1T_2} < \frac{1}{\sigma_0} \sqrt{\frac{2}{\pi}}. \quad (10.7.24)$$

МЕТОДЫ ИССЛЕДОВАНИЯ ТОЧНОСТИ НЕЛИНЕЙНЫХ СИСТЕМ

§ 11.1. Приближенное исследование точности нелинейных систем методами линеаризации

В настоящее время не существует общей точной статистической теории нелинейных систем, позволяющей находить вероятностные характеристики выходных переменных любых нелинейных систем. Точная статистическая теория существует лишь для некоторых частных видов нелинейных систем (см. [53], §§ 106, 107, а также [52], [55] и [56]). Поэтому для исследования точности нелинейных систем, имеющих заданные характеристики, обычно применяют различные приближенные методы. Наибольшее распространение получали методы, основанные на линеаризации уравнений нелинейных систем, как обычной, так и статистической. Другие приближенные методы исследования точности нелинейных систем основаны на разложении выходных переменных систем в степенные ряды относительно случайных коэффициентов канонического разложения входных случайных функций (см. [53], § 108). Одним из наиболее удобных таких методов является метод эквивалентных возмущений Б. Г. Доступова [20, 21] (см. также [53], § 108).

Ввиду того что нелинейные уравнения, не содержащие случайных функций, можно решать с помощью современных математических машин, как моделирующих, так и цифровых, для исследования точности нелинейных систем достаточно линеаризовать их уравнения лишь относительно центрированных случайных функций, т. е. относительно случайных отклонений, входящих в уравнения переменных от их математических ожиданий. При этом уравнения системы останутся нелинейными относительно неслучайных полезных сигналов, представляющих собой математические ожидания входящих в уравнения системы переменных.

Как и в предыдущей главе, мы предположим, что нелинейная система состоит из линейных систем и безынерционных элементарных нелинейных звеньев. Тогда для линеаризации уравнений системы достаточно будет линеаризовать зависимости между случайными составляющими входных и выходных сигналов всех элементарных нелинейных звеньев.

Рассмотрим элементарное нелинейное звено с характеристикой

$$Y = \varphi(X). \quad (11.1.1)$$

Если характеристика этого звена φ непрерывна и изображается кривой без угловых точек, то для линейризации зависимости (11.1.1) относительно центрированных случайных функций достаточно заменить кривую $y = \varphi(x)$ касательной к ней в точке $x = m_x$. Тогда получим следующую приближенную зависимость между входным и выходным сигналами нелинейного звена (см. § 8.3):

$$Y \approx \varphi(m_x) + \varphi'(m_x) X^0. \quad (11.1.2)$$

Эта зависимость линейна только относительно центрированной случайной функции X^0 и нелинейна относительно полезного входного сигнала m_x .

Если характеристика нелинейного звена φ существенно нелинейна, например разрывна или имеет угловые точки, то для линейризации зависимости (11.1.1) относительно центрированных случайных функций придется применить метод статистической линейризации (см. § 8.5). В результате получим следующую приближенную зависимость между входным и выходным сигналами нелинейного звена:

$$Y \approx \varphi_0(m_x, \sigma_x) + k_1(m_x, \sigma_x) X^0. \quad (11.1.3)$$

Эта зависимость также линейна только относительно центрированной случайной функции X^0 и нелинейна относительно m_x . Ее принципиальное отличие от зависимости (11.1.2) заключается в том, что коэффициенты φ_0 и k_1 зависят не только от математического ожидания m_x входного сигнала нелинейного звена, но и от его дисперсии $D_x = \sigma_x^2$, т. е. от уровня шумов на входе нелинейного звена.

Заменяя в уравнениях нелинейной системы для всех ее нелинейных звеньев нелинейные зависимости типа (11.1.1) приближенными зависимостями вида (11.1.2) или (11.1.3), получим приближенные уравнения системы, линейные относительно случайных функций. При этом, если система содержит только нелинейные звенья с непрерывными гладкими характеристиками, то коэффициенты линейризованных уравнений системы будут зависеть только от математических ожиданий входных сигналов нелинейных звеньев. Если же система содержит существенно нелинейные элементарные звенья, то коэффициенты линейризованных уравнений системы будут зависеть еще от дисперсий входных сигналов нелинейных звеньев. Во всех случаях линейризация уравнений системы равноценна замене данной системы другой системой, линейной по отношению к передаче случайных возмущений (флуктуаций) и нелинейной по отношению к передаче полезных сигналов. Линейризация уравнений системы дает возможность применить для приближенного исследования ее точности статистическую теорию линейных систем, изложенную в главе 7.

Согласно изложенному в § 7.2 общему методу, уравнения для математических ожиданий выходных переменных системы, поведение которой описывается линейными относительно случайных функций уравнениями, получаются из уравнений данной системы путем замены всех случайных функций их математическими ожиданиями. Производя такую замену в (11.1.2), получим

$$m_y \approx \varphi(m_x). \quad (11.1.4)$$

Сравнивая эту формулу с (11.1.1), видим, что математические ожидания входных и выходных переменных элементарных нелинейных звеньев с непрерывными гладкими характеристиками связаны теми же уравнениями, что и сами переменные. На основании § 7.2 математические ожидания входных и выходных переменных линейных систем, входящих в состав нелинейной системы, также связаны теми же уравнениями, что и сами переменные. Отсюда следует, что если система содержит только элементарные нелинейные звенья с непрерывными гладкими характеристиками, то математические ожидания ее выходных переменных определяются теми же уравнениями, которыми описывается поведение системы, с заменой в них всех случайных величин их математическими ожиданиями. В отношении начальных условий для математических ожиданий переменных остается в силе все сказанное по этому поводу в § 7.2.

Вычитая (11.1.4) из (11.1.2), получим линеаризованную зависимость между центрированными случайными функциями на входе и выходе нелинейного звена:

$$Y^0 \approx \varphi'(m_x) X^0. \quad (11.1.5)$$

После интегрирования уравнений для математических ожиданий переменных все коэффициенты $\varphi'(m_x)$ в зависимостях вида (11.1.5) будут известными функциями времени. Таким образом, для центрированных случайных функций получатся линейные уравнения с известными коэффициентами. Поэтому для определения дисперсий выходных переменных системы можно будет применить изложенные в главе 7 методы исследования точности линейных систем.

Итак, исследование точности системы, содержащей только нелинейные звенья с непрерывными гладкими характеристиками, сводится к решению уравнений, описывающих поведение системы, для математических ожиданий выходных переменных, определению коэффициентов уравнений системы, линеаризованных относительно центрированных случайных функций, и применению теории точности линейных систем для нахождения дисперсий выходных переменных.

Пример 11.1.1. Рассмотрим систему, представляющую собой аperiodическое звено, охваченное отрицательной обратной связью, в цепь которой включено нелинейное звено с кубической характеристикой. Из рис. 11.1.1

следует, что поведение этой системы описывается уравнениями

$$T\dot{Y} + Y = k(X - U), \quad U = aY^3.$$

Линеаризуя второе уравнение путем разложения нелинейной функции в ряд Тейлора в окрестности точки m_Y , получим следующие уравнения для математического ожидания выходной переменной системы и центрированных случайных функций:

$$T\dot{m}_Y + m_Y = k(m_X - am_Y^3), \quad T\dot{Y}^0 + (1 + 3kam_Y^2)Y^0 = kX^0.$$

Принтегрировав первое уравнение при заданных начальных условиях, найдем m_Y как функцию t . После этого коэффициент во втором уравнении станет известной функцией времени и дисперсию выходной переменной системы Y можно будет определить любым из методов, изложенных в § 7.2.

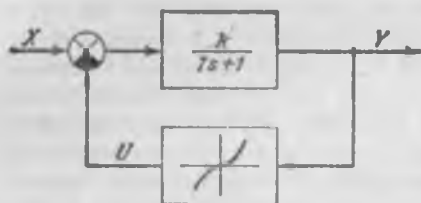


Рис. 11.1.1.

Заменяя случайные функции их математическими ожиданиями в (11.1.3), получим

$$m_Y \approx \varphi_0(m_X, \sigma_X). \quad (11.1.6)$$

Таким образом, при применении метода статистической линеаризации

сама форма зависимости между математическими ожиданиями входного и выходного сигналов нелинейного звена зависит от дисперсии входного сигнала и не совпадает с характеристикой нелинейного звена. Вследствие этого уравнения для математических ожиданий переменных не совпадают с уравнениями, описывающими поведение системы, и не могут быть решены отдельно от уравнений, определяющих случайные отклонения переменных от их математических ожиданий, так как все функции вида (11.1.6) остаются неопределенными, пока не найдены дисперсии входных сигналов всех нелинейных звеньев системы.

Для получения уравнений для центрированных случайных функций следует вычесть равенство (11.1.6) из (11.1.3). Тогда получим

$$Y^0 \approx k_1(m_X, \sigma_X) X^0. \quad (11.1.7)$$

Добавляя к таким зависимостям для всех нелинейных звеньев системы уравнения входящих в данную систему линейных систем, получим полную систему уравнений, линейных относительно центрированных случайных функций. Коэффициенты уравнений вида (11.1.7) остаются неизвестными до тех пор, пока не определены математические ожидания и дисперсии входных сигналов всех нелинейных звеньев. Поэтому методы главы 7 не могут быть непосредственно применены для определения дисперсий переменных.

Таким образом, для систем, содержащих существенно нелинейные звенья, необходимо совместно решать уравнения, определяющие математические ожидания, и уравнения, определяющие дисперсии переменных.

Для решения этой задачи в общем случае можно применить метод канонических разложений (§ 7.2). Применив этот метод к линеаризованным уравнениям системы, получим для координатных функций переменных линейные уравнения с коэффициентами, зависящими от неизвестных математических ожиданий и дисперсий входных сигналов нелинейных звеньев. Добавив к уравнениям, определяющим математические ожидания и координатные функции, формулы вида (7.2.20) для дисперсий входных сигналов нелинейных звеньев, получим полную бесконечную систему уравнений для математических ожиданий, координатных функций и дисперсий всех переменных, описывающих поведение исследуемой системы (координат системы). Ограничиваясь конечными отрезками канонических разложений, получим конечную систему уравнений для приближенного определения математических ожиданий, координатных функций и дисперсий координат системы. Эта система уравнений может быть решена любым численным методом на цифровой вычислительной машине или с помощью моделирующих устройств.

В некоторых частных случаях задачу совместного определения математических ожиданий и дисперсий входных сигналов нелинейных звеньев удается свести к применению графических методов решения уравнений или даже решить ее аналитически. Один такой частный случай будет рассмотрен в следующем параграфе.

Пример 11.1.2. Рассмотрим систему, состоящую из последовательно соединенных релейного элемента и апериодического звена, охваченных отрицательной обратной связью (рис. 11.1.2). Уравнения этой системы имеют вид

$$\left. \begin{aligned} T\dot{Y} + Y &= l \operatorname{sgn} V, \\ V &= X - Y. \end{aligned} \right\} \quad (11.1.8)$$



Рис. 11.1.2.

Произведя статистическую линеаризацию характеристики релейного элемента, получим следующие уравнения для математических ожиданий переменных и для централизованных случайных функций:

$$T\dot{m}_Y + m_Y = \varphi_0(m_Y, \sigma_Y), \quad m_Y = m_X - m_Y, \quad (11.1.9)$$

$$T\dot{Y}^0 + [1 + k_1(m_Y, \sigma_Y)] Y^0 = k_1(m_Y, \sigma_Y) X^0, \quad V^0 = Y^0 - X^0. \quad (11.1.10)$$

Коэффициенты $\varphi_0 = k_0 m_Y$ и k_1 для релейного элемента определяются формулами, приведенными в приложении 5 (нелинейность 5). Первое из этих уравнений не может быть проинтегрировано, пока не известна величина σ_Y . С другой стороны, σ_Y нельзя определить из второго уравнения методами главы 7, пока не известен коэффициент k_1 , зависящий от неизвестных m_Y и σ_Y .

Применив метод канонических разложений, получим для координатных функций переменных Y и V уравнения

$$T \dot{y}_v + [1 + k_1(m_v, \sigma_v)] y_v = k_1(m_v, \sigma_v) x_v, \quad v_v = y_v - x_v \quad (v=1, 2, \dots). \quad (11.1.11)$$

К уравнениям (11.1.9) и (11.1.11) следует добавить формулу

$$\sigma_v^2 = \sum_{v=1}^{\infty} D_v |v_v|^2, \quad (11.1.12)$$

определяющую дисперсию входного сигнала V нелинейного звена. Ограничиваясь в канонических разложениях n первыми членами, получим систему $2n + 3$ совместных уравнений (11.1.9), (11.1.11) и (11.1.12), приближенно определяющую $2n + 3$ неизвестных функций времени $m_v, m_v, y_1, v_1, \dots, y_n, v_n, \sigma_v$. После нахождения этих функций дисперсия выходной переменной Y определится по формуле (7.2.20).

При применении изложенного метода к автоколебательным системам, т. е. к таким системам, в которых возможны устойчивые автоколебания, математические ожидания переменных получаются в виде колеблющихся функций времени, стремящихся к некоторым периодическим функциям. Если частота колебаний математических ожиданий сравнима со средней частотой колебаний случайных составляющих входных сигналов нелинейных звеньев, то изложенный метод может оказаться неточным. В подобных случаях можно получить более точные результаты, применяя совместную статистическую и гармоническую линеаризацию, как было изложено в § 10.7.

§ 11.2. Применение метода статистической линеаризации для исследования точности стационарных систем

Рассмотрим сначала стационарную систему с одним нелинейным звеном, работающую в установившемся режиме под действием стационарного входного сигнала $X(t)$. Если нелинейное звено не находится в цепи обратной связи и не охватывается обратной связью (рис. 11.2.1), то дисперсия



Рис. 11.2.1.

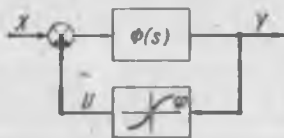


Рис. 11.2.2.

выходной переменной системы может быть вычислена точно. Для этого достаточно определить корреляционную функцию выходного сигнала нелинейного звена по данной двумерной плотности вероятности входного сигнала X , пользуясь обычными методами теории функциональных зависимостей между случайными вели-

чинами (см. [54], § 3.10 или [53], § 30). После этого можно будет найти спектральную плотность выходного сигнала нелинейного звена и определить дисперсию выходной переменной системы Y по формуле (7.4.4).

Если нелинейное звено находится в цепи обратной связи (рис. 11.2.2) или охватывается обратной связью (рис. 11.2.3 и рис. 11.2.4), то точное определение дисперсии выходной переменной системы в общем случае невозможно и для ее нахождения

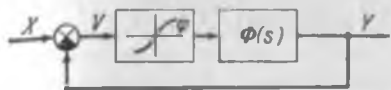


Рис. 11.2.3.

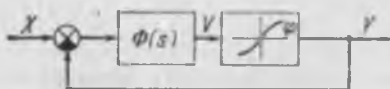


Рис. 11.2.4.

можно применить метод линеаризации, изложенный в предыдущем параграфе, статистической или обычной, в зависимости от характеристики звена.

Обозначая выходной сигнал нелинейного звена через U , применяя статистическую линеаризацию нелинейного звена и заменяя в (11.1.3) функцию φ_0 ее выражением (8.5.3): $\varphi_0 = k_0 m_y$, напомним для случая нелинейного звена в цепи обратной связи уравнения (11.1.6) и (11.1.7) в виде

$$m_u = k_0 m_y, \quad U^0 = k_1 Y^0, \quad (11.2.1)$$

где k_0 и k_1 являются известными функциями величин m_y и σ_y . Формулы (11.2.1) показывают, что если считать k_0 и k_1 неизвестными постоянными, то в результате статистической линеаризации нелинейное звено заменяется безынерционным усилителем с разными коэффициентами усиления полезного сигнала и помехи. Таким образом, при постоянных k_0 и k_1 статистическая линеаризация заменяет данную систему двумя разными линейными системами для полезного сигнала и для флуктуаций. Пользуясь формулой (4.6.7), находим передаточные функции линеаризованной системы для полезного сигнала и для флуктуаций:

$$\Phi_y^{(0)}(s) = \frac{\Phi(s)}{1 + k_0 \Phi(s)}, \quad \Phi_y^{(1)}(s) = \frac{\Phi(s)}{1 + k_1 \Phi(s)}, \quad (11.2.2)$$

где $\Phi(s)$ — передаточная функция линейной части системы.

Считая передаточную функцию $\Phi_y^{(0)}(s)$ известной, применим формулу (7.3.8) для математического ожидания выходной переменной системы. Тогда, принимая во внимание, что математическое ожидание стационарной входной переменной системы $X(t)$ постоянно, получим

$$m_y = \frac{\Phi(0) m_x}{1 + k_0(m_y, \sigma_y) \Phi(0)}, \quad (11.2.3)$$

где мы теперь показываем в явной форме, что коэффициент k_0 является известной функцией величин m_y и σ_y . Уравнение (11.2.3) содержит две неизвестные величины m_y и σ_y и поэтому может быть решено только совместно с уравнением, определяющим дисперсию выходной переменной системы σ_y^2 .

Для определения дисперсии выходной переменной системы Y применим формулу (7.4.4). В результате получим

$$\sigma_y^2 = \int_{-\infty}^{\infty} s_x(\omega) \left| \frac{\Phi(i\omega)}{1 + k_1(m_y, \sigma_y)\Phi(i\omega)} \right|^2 d\omega, \quad (11.2.4)$$

где также показана явно зависимость k_1 от m_y и σ_y .

Решив уравнения (11.2.3) и (11.2.4) относительно неизвестных величин m_y и σ_y , мы и найдем математическое ожидание и дисперсию выходной переменной рассматриваемой системы. Эти уравнения могут быть решены любым приближенным методом. В частности, можно применить метод последовательных приближений или графический метод.

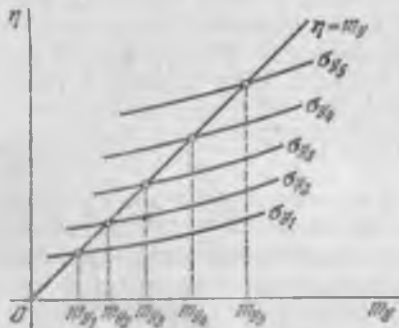


Рис. 11.2.5.

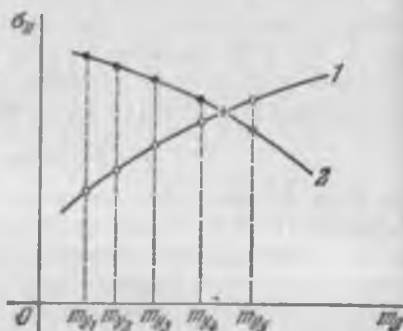


Рис. 11.2.6.

Для решения уравнений (11.2.3) и (11.2.4) методом последовательных приближений необходимо задать исходные грубо приближенные значения k_0 и k_1 и вычислить в первом приближении m_y и σ_y по формулам (11.2.3) и (11.2.4). После этого можно уточнить значения k_0 и k_1 и вычислить m_y и σ_y во втором приближении. И так далее. Процесс заканчивается, когда два последовательных приближения совпадут в пределах принятой точности вычислений. Обычно практически бывает достаточно двух или трех приближений.

Для решения уравнений (11.2.3) и (11.2.4) графическим методом заменим уравнение (11.2.3) равноценной системой уравнений

$$\eta = m_y, \quad \eta = \frac{\Phi(0) m_x}{1 + k_0(m_y, \sigma_y) \Phi(0)}. \quad (11.2.5)$$

Первому из этих уравнений соответствует биссектриса координатного угла в декартовых прямоугольных координатах (m_y, η). Второму уравнению (11.2.5) соответствует семейство кривых в координатах (m_y, η) с параметром σ_y . Построив кривые, соответствующие второму уравнению (11.2.5), для ряда значений параметра σ_y и определив точки пересечения их с биссектрисой координатного угла (рис. 11.2.5), мы найдем значения абсцисс m_y кривой (11.2.3), соответствующие выбранным значениям ординаты σ_y , после чего можем по найденным точкам построить кривую (11.2.3) в координатах (m_y, σ_y) (кривая 1 на рис. 11.2.6).

После этого строим в тех же координатах (m_y, σ_y) кривую 2, соответствующую уравнению

$$\xi^2 = \int_{-\infty}^{\infty} s_x(\omega) \left| \frac{\Phi(i\omega)}{1 + k_1(m_y, \sigma_y)\Phi(i\omega)} \right|^2 d\omega, \quad (11.2.6)$$

откладывая величину ξ по оси σ_y и рассматривая величину σ_y как функцию величины m_y , определяемую уравнением (11.2.3), или, что то же, кривой 1 на рис. 11.2.6. Кривая (11.2.6), очевидно, пересекает кривую (11.2.3) в той же точке, что и кривая (11.2.4). Поэтому точка пересечения кривых (11.2.3) и (11.2.6) дает решение уравнений (11.2.3) и (11.2.4).

Заметим, что кривая 1, соответствующая уравнению (11.2.3), может быть построена и другим способом. А именно, задаваясь различными значениями m_y , можно определить из уравнения (11.2.3) соответствующие значения σ_y и по полученным таким образом точкам построить кривую 1.

Совершенно аналогично решается задача определения математического ожидания и дисперсии выходной переменной стационарной системы при других способах включения нелинейного звена. Для определения математического ожидания и дисперсии входного сигнала нелинейного звена достаточно принять за выходную переменную системы входной сигнал нелинейного звена и в соответствии с этим преобразовать структурные схемы систем.

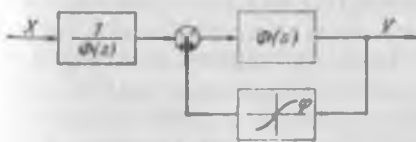


Рис. 11.2.7.

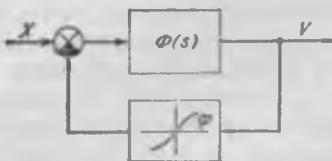


Рис. 11.2.8.

Изображенные на рис. 11.2.3 и 11.2.4, в схемы, представленные на рис. 11.2.7 и 11.2.8 соответственно. Тогда, применив изложенный метод, можно найти математическое ожидание и дисперсию

входного сигнала нелинейного звена и его статистические коэффициенты усиления k_0 и k_1 . После этого, вернувшись к первоначальной структурной схеме, можно найти ее передаточные функции по полезному сигналу и по флуктуациям и по формулам (7.3.8) и (7.4.4) найти математическое ожидание и дисперсию выходной переменной исследуемой системы.

Совершенно аналогично метод статистической линеаризации применяется к одномерным стационарным системам с несколькими нелинейными звеньями и к многомерным стационарным системам. Заменив нелинейные звенья линейными усилителями с соответствующими статистическими коэффициентами усиления и выразив при помощи формул (7.3.8) и (7.4.4) математические ожидания и дисперсии входных сигналов всех нелинейных звеньев, получим соответствующее количество уравнений для определения неизвестных математических ожиданий и дисперсий входных сигналов нелинейных звеньев. В результате решения этих уравнений будут определены математические ожидания и дисперсии входных сигналов нелинейных звеньев и соответствующие статистические коэффициенты усиления этих звеньев. После этого можно будет определить по тем же формулам (7.3.8) и (7.4.4) математическое ожидание и дисперсию выходной переменной системы.

Однако в случае системы с двумя или большим количеством нелинейных звеньев графическое решение уравнений, определяющих математические ожидания и дисперсии входных сигналов нелинейных звеньев, может стать очень громоздким. Поэтому в общем случае для системы с несколькими нелинейными звеньями системы уравнений, определяющие математические ожидания и дисперсии входных сигналов нелинейных звеньев, можно решить практически только методом последовательных приближений. И лишь в частных случаях, когда эти уравнения распадаются на независимые пары уравнений, их можно решить изложенным выше графическим методом.

Заметим, что в случае астатической следящей системы порядка k с нелинейным звеном, включенным по схеме рис. 11.2.4, установившееся значение математического ожидания входного сигнала нелинейного звена будет постоянным, если математическое ожидание входного сигнала системы представляет собой полином степени k относительно времени t . Поэтому изложенный метод полностью применим к астатическим следящим системам в случае, когда входной сигнал представляет собой сумму стационарной случайной функции и полинома относительно времени, степень которого не выше порядка астатизма системы.

Изложенный метод применим только к таким системам, в которых невозможны автоколебания, так как только в случае отсутствия автоколебаний m_y может быть постоянным при постоянном m_x . Поэтому перед тем, как применять этот метод, необходимо

исследовать систему на автоколебания, что можно сделать с помощью метода точечных отображений (§ 10.5) или метода гармонической линеаризации (§ 10.6). Если исследование покажет, что в системе возможны устойчивые автоколебания, то для исследования ее точности необходимо применить метод совместной статистической и гармонической линеаризации, как показано в § 10.7.

Пример 11.2.1. Рассмотрим следящую систему примера 10.4.1, состоящую из исполнительного устройства с передаточной функцией

$$\Phi(s) = \frac{k}{s(Ts+1)} \quad (11.2.7)$$

и управляющего им реле с идеальной релейной характеристикой в прямой цепи (рис. 11.2.9), предполагая, что входной полезный сигнал представляет

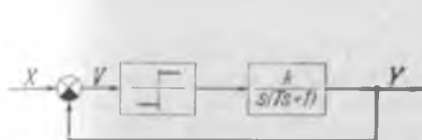


Рис. 11.2.9.

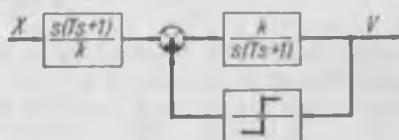


Рис. 11.2.10.

собой линейную функцию времени $a + bt$, а помеха является стационарной случайной функцией с нулевым математическим ожиданием, корреляционная функция и спектральная плотность которой определяются формулами

$$k_x(\tau) = De^{-\alpha|\tau|}, \quad S_x(\omega) = \frac{D}{\pi} \cdot \frac{\alpha}{\alpha^2 + \omega^2}. \quad (11.2.8)$$

В примере 10.4.1 мы видели, что автоколебания в рассматриваемой системе невозможны. Поэтому для исследования ее точности можно применить изложенный метод.

Согласно изложенному перестраиваем структурную схему данной системы так, чтобы выходной переменной системы служил входной сигнал нелинейного звена V (рис. 11.2.10). Тогда, применяя формулы (4.6.1) и (4.6.7), находим передаточные функции линеаризованной системы по полезному сигналу $\Phi_v^{(0)}(s)$ и по помехе $\Phi_v^{(1)}(s)$:

$$\begin{aligned} \Phi_v^{(0)}(s) &= \frac{s(Ts+1)}{s(Ts+1) + kk_0}, \\ \Phi_v^{(1)}(s) &= \frac{s(Ts+1)}{s(Ts+1) + kk_1}. \end{aligned} \quad (11.2.9)$$

Принимая во внимание, что в данном случае $m_x(t) = a + bt$, находим по формуле (7.3.8) математическое ожидание входного сигнала нелинейного звена:

$$m_0 = \frac{b}{kk_0(m_0, \sigma_0)}. \quad (11.2.10)$$

Формула (7.4.4) дает следующее выражение для дисперсии входного сигнала нелинейного звена:

$$\sigma_0^2 = \frac{D\alpha}{\pi} \int_{-\infty}^{\infty} \frac{T^2(i\omega)^4 - (i\omega)^2}{|(\alpha + i\omega)\{i\omega(Ti\omega + 1) + kk_1(m_0, \sigma_0)\}|^2} d\omega.$$

Применив для вычисления интеграла формулу (7.4.6), получим

$$\sigma_v^2 = D\alpha \frac{1 + [\alpha + kk_1(m_v, \sigma_v)]T}{\alpha + kk_1(m_v, \sigma_v) + \alpha^2 T} \quad (11.2.11)$$

В уравнениях (11.2.10) и (11.2.11) статистические коэффициенты усиления звена с идеальной релейной характеристикой k_0 и k_1 определяются формулами (приложение 6, нелинейность 5)

$$k_0(m_v, \sigma_v) = \frac{2l}{m_v} \Phi\left(\frac{m_v}{\sigma_v}\right), \quad k_1(m_v, \sigma_v) = \frac{l}{\sigma_v} \psi\left(\frac{m_v}{\sigma_v}\right), \quad (11.2.12)$$

где

$$\psi(z) = \frac{1}{2} \sqrt{1 - 4\Phi^2(z)} + \Phi'(z). \quad (11.2.13)$$

Подставляя выражение k_0 из (11.2.12) в уравнение (11.2.10), приведем его к виду

$$\Phi\left(\frac{m_v}{\sigma_v}\right) = \frac{b}{2kl}. \quad (11.2.14)$$

Таким образом, в данном случае кривая 1 представляет собой прямую $m_v/\sigma_v = z$, где z — значение аргумента функции $\Phi(z)$, при котором она равна $b/2kl$. Определив z по таблице функции $\Phi(u)$ (приложение 6), строим прямую 1 (рис. 11.2.11).

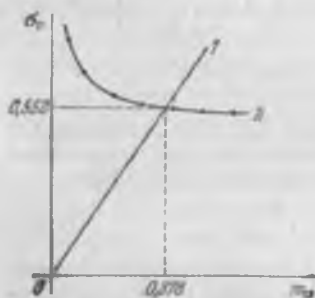


Рис. 11.2.11.

Для построения кривой 2 подставим в уравнение (11.2.11) выражение k_1 из (11.2.12), заменив в нем величину σ_v ее выражением m_v/z , соответствующим уравнению (11.2.14). В результате получим уравнение

$$\zeta^2 = D\alpha \frac{(1 + \alpha T)m_v + klT\psi(z)}{\alpha(1 + \alpha T)m_v + kl\psi(z)}. \quad (11.2.15)$$

Задаваясь рядом значений m_v , определяем по уравнению (11.2.15) соответствующие значения ζ и, откладывая их по оси σ_v , строим кривую 2 (рис. 11.2.11). Точка пересечения кривых 1 и 2 определяет значения m_v и σ_v , удовлетворяющие уравнениям (11.2.10) и (11.2.11). Определив m_v и σ_v , находим по

формулам (11.2.12) и (11.2.13) статистические коэффициенты усиления нелинейного звена k_0 и k_1 .

Для определения математического ожидания и дисперсии выходной переменной рассматриваемой системы возвращаемся к первоначальной структурной схеме системы (рис. 11.2.9) и, пользуясь формулой (4.6.8), находим передаточные функции линеаризованной системы по полезному сигналу $\Phi_Y^{(0)}(s)$ и по помехе $\Phi_Y^{(1)}(s)$:

$$\Phi_Y^{(0)}(s) = \frac{kk_0}{s(Ts + 1) + kk_0}, \quad \Phi_Y^{(1)}(s) = \frac{kk_1}{s(Ts + 1) + kk_1}. \quad (11.2.16)$$

После этого по формулам (7.3.8) и (7.4.4) находим математическое ожидание и дисперсию выходной переменной Y системы:

$$m_Y(t) = a + bt - \frac{b}{kk_0}, \quad (11.2.17)$$

$$\sigma_Y^2 = \frac{D\alpha k^2 k_1^2}{\pi} \int_{-\infty}^{\infty} \frac{d\omega}{|(\alpha + i\omega)\{i\omega(Ti\omega + 1) + kk_1\}|^2}.$$

Применив для вычисления интеграла формулу (7.4.6), получим

$$\sigma_y^2 = Dkk_1 \frac{1 + \alpha T}{\alpha + kk_1 + \alpha^2 T} \quad (11.2.18)$$

На рис. 11.2.11 кривые 1 и 2 построены для следующих числовых значений параметров системы и входного сигнала: $k = 2$, $l = T = a = b = 1$, $\alpha = 10$, $\sigma_x = \sqrt{D} = 0,5$. В результате получается $m_r = 0,378$, $\sigma_r = 0,552$, $k_0 = 1,324$, $k_1 = 1,362$ и формулы (11.2.17) и (11.2.18) дают

$$m_y = 0,622 + t, \quad \sigma_y = 0,258.$$

§ 11.3. Применение метода статистической линеаризации для исследования точности систем, описываемых дифференциальными уравнениями

При исследовании точности нестационарных систем методом статистической линеаризации основная трудность заключается в том, что математические ожидания и дисперсии входных сигналов нелинейных звеньев являются переменными, вследствие чего и статистические характеристики нелинейных звеньев φ_0 , k_0 и k_1 представляют собой функции времени. Как было отмечено в § 11.1, эту трудность в общем случае можно преодолеть, применив к статистически линеаризованной системе метод канонических разложений. Однако это связано с большим объемом вычислений. В некоторых случаях, когда поведение исследуемой нестационарной системы описывается системой дифференциальных уравнений не очень высокого порядка, целесообразно применить к статистически линеаризованной системе метод § 7.7 [1]. Применение этого метода совместно с методом статистической линеаризации удобно тем, что дает математические ожидания, дисперсии и корреляционные моменты всех выходных переменных одновременно. Поэтому наличие в системе многих нелинейностей практически не очень сильно усложняет исследование. Кроме того, при этом вычисление статистических характеристик нелинейных звеньев производится одновременно с интегрированием уравнений для моментов, так же как и при применении метода канонических разложений.

Рассмотрим автоматическую систему, описываемую системой обыкновенных нелинейных дифференциальных уравнений первого порядка

$$\dot{Y}_r = \sum_{l=1}^n \varphi_{rl}(t, Y_l) + V_r \quad (r = 1, \dots, n), \quad (11.3.1)$$

где V_r — белые шумы, а $\varphi_{rl}(t, Y_l)$ — любые заданные функции. Некоторые из этих функций могут быть, в частности, линейными, т. е. иметь вид $a_{rl}(t) Y_l$. В соответствии с общим методом § 11.1 производим статистическую линеаризацию всех нелинейных

функций, входящих в (11.3.1):

$$\varphi_{rl}(t, Y_l) \approx \varphi_{0rl}(t, m_l, \sigma_l) + k_{1rl}(t, m_l, \sigma_l) Y_l^1, \quad (11.3.2)$$

где $m_l = M[Y_l]$, $\sigma_l = \sqrt{D[Y_l]}$. Если некоторые из функций $\varphi_{rl}(t, Y_l)$ являются нечетными, то для них вместо формулы (11.3.2) можно написать

$$\varphi_{rl}(t, Y_l) \approx k_{0rl}(t, m_l, \sigma_l) m_l + k_{1rl}(t, m_l, \sigma_l) Y_l^1. \quad (11.3.3)$$

В дальнейшем будем опускать аргумент t в выражениях для коэффициентов статистической линеаризации, помня об их явной зависимости от времени.

Подставляя выражение (11.3.2) в (11.3.1), получаем

$$\dot{Y}_r = \sum_{l=1}^n \varphi_{0rl}(m_l, \sigma_l) + \sum_{l=1}^n k_{1rl}(m_l, \sigma_l) Y_l^1 + V_r \quad (r = 1, \dots, n). \quad (11.3.4)$$

Для получения уравнений для математических ожиданий заменим в уравнениях (11.3.4) все случайные функции их математическими ожиданиями. Получим

$$\dot{m}_r = \sum_{l=1}^n \varphi_{0rl}(m_l, \sigma_l) + m_{vr} \quad (r = 1, \dots, n), \quad (11.3.5)$$

где $m_{vr} = M[V_r]$. В соответствии с общим замечанием в § 11.1 эту систему уравнений нельзя проинтегрировать, так как в нее входят, кроме неизвестных математических ожиданий $m_1, \dots, m_n$, еще неизвестные средние квадратические отклонения $\sigma_1, \dots, \sigma_n$. Поэтому к уравнениям (11.3.5) необходимо добавить уравнения, определяющие дисперсии и корреляционные моменты переменных. Для составления этих уравнений воспользуемся методом § 7.7. Тогда, принимая во внимание, что $\sigma_v = \sqrt{\theta_{vv}}$ ($v = 1, \dots, n$), получим

$$\dot{\theta}_{ij} = \sum_{l=1}^n [k_{1il}(m_l, \sqrt{\theta_{ll}}) \theta_{jl} + k_{1jl}(m_l, \sqrt{\theta_{ll}}) \theta_{il}] + G_{ij}(t) \quad (11.3.6) \\ (i, j = 1, \dots, n).$$

Уравнения (11.3.5) и (11.3.6) образуют замкнутую систему совместных уравнений, определяющую неизвестные функции времени $m_v, \theta_{v\mu}$ ($v, \mu = 1, \dots, n$). Эта система содержит n уравнений первого порядка (11.3.5) и $n(n+1)/2$ уравнений второго порядка (11.3.6). Следовательно, порядок полной системы уравнений, подлежащей интегрированию, равен $N = n + n(n+1)/2 = n(n+3)/2$. В таблице 11.3.1 показана зависимость порядка системы уравнений, определяющей моменты первого и второго порядков, от порядка исходной системы уравнений.

ТАБЛИЦА 11.3.1

n	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
N	2	5	9	14	20	27	35	44	54	65	77	90	104	119	135

Мы видим, что порядок системы уравнений, определяющей моменты выходных сигналов, быстро растет с увеличением порядка системы уравнений, описывающей поведение исследуемой системы. Вследствие этого изложенный метод, так же как и метод § 7.7, можно практически применять только в случаях, когда поведение исследуемой системы описывается системой дифференциальных уравнений невысокого порядка.

Интересно отметить, что при применении более общего метода канонических разложений (см. конец § 11.1) порядок системы уравнений, определяющей моменты первого и второго порядков выходных сигналов, растет не так быстро с увеличением порядка n исходной системы уравнений. Действительно, если ограничиться в канонических разложениях p первыми членами, то система уравнений, определяющих математические ожидания и координатные функции всех переменных, будет иметь порядок $(p + 1)n$. Таким образом, при применении метода канонических разложений порядок системы уравнений, определяющей моменты первого и второго порядков выходных сигналов, растет линейно с увеличением порядка исходной системы уравнений n , в то время как при применении только что изложенного метода порядок системы уравнений для моментов растет как n^2 . Вследствие этого для систем, описываемых системами дифференциальных уравнений высокого порядка, метод канонических разложений оказывается более выгодным, чем метод этого параграфа.

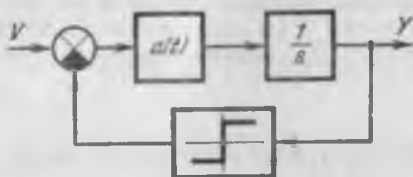


Рис. 11.3.1.

Пример 11.3.1. Рассмотрим систему, состоящую из последовательного соединения усилителя с переменным коэффициентом усиления и интегратора, охваченного отрицательной обратной связью, в цепь которой включено нелинейное звено с релейной характеристикой. Из рис. 11.3.1 следует, что поведение этой системы описывается дифференциальным уравнением

$$\dot{Y} = -a(t) \operatorname{sgn} Y + a(t) V. \quad (11.3.7)$$

Пусть V представляет собой белый шум интенсивности $G(t)$. Так как нелинейность, входящая в уравнение (11.3.7), является нечетной, то уравнения для математического ожидания и случайной составляющей можно взять

в виде

$$\dot{m}_y = -a(t) k_0(m_y, \sigma_y) m_y + a(t) m_0, \quad (11.3.8)$$

$$\dot{Y}^0 = -a(t) k_1(m_y, \sigma_y) Y^0 + a(t) V^0. \quad (11.3.9)$$

Напишем дифференциальное уравнение для дисперсии. Полагая в (11.3.8) $n = 1$, $G_{11} = a(t) G(t)$, $0_{11} = D_y$ и имея в виду, что $k_{111}(m_1, \sigma_1) = k_1(m_y, \sigma_y)$, получаем

$$\dot{D}_y = -2a(t) k_1(m_y, \sigma_y) D_y + a^2(t) G(t). \quad (11.3.10)$$

Подставляя в уравнения (11.3.8) и (11.3.10) выражения $k_0(m_y, \sigma_y)$ и $k_1(m_y, \sigma_y)$ для релейной характеристики нелинейного звена в случае аппроксимации нелинейности по второму способу (приложения 4 и 5, нелинейность 4), получаем

$$\left. \begin{aligned} \dot{m}_y &= -2la(t) \Phi\left(\frac{m_y}{\sqrt{D_y}}\right) + a(t) m_0, \\ \dot{D}_y &= -\frac{4la(t)}{\sqrt{2\pi}} \sqrt{D_y} e^{-\frac{m_y^2}{2D_y}} + a^2(t) G(t). \end{aligned} \right\} \quad (11.3.11)$$

Эту систему уравнений следует интегрировать при начальных значениях m_y и D_y , равных соответственно математическому ожиданию и дисперсии случайного начального значения выходного сигнала системы.

Если $a(t) = a = \text{const}$, белый шум V стационарен: $G(t) = G = \text{const}$, и требуется найти установившиеся значения m_y и D_y , то в уравнениях (11.3.11) следует положить $\dot{m}_y = \dot{D}_y = 0$:

$$2l\Phi\left(\frac{m_y}{\sqrt{D_y}}\right) = m_0, \quad \frac{4l}{\sqrt{2\pi}} \sqrt{D_y} e^{-\frac{m_y^2}{2D_y}} = aG. \quad (11.3.12)$$

Эта система уравнений решается очень легко. Из первого уравнения определяется значение функции $\Phi(m_y/\sqrt{D_y}) = m_0/2l$. Затем, пользуясь таблицей функции $\Phi(u)$ (приложение 6), можно определить значение аргумента $z = m_y/\sqrt{D_y}$, после чего дисперсия D_y находится из второго уравнения (11.3.12):

$$D_y = \frac{\pi a^2 G^2}{8l^2} e^{z^2}. \quad (11.3.13)$$

Наконец, математическое ожидание m_y находится по формуле

$$m_y = z \sqrt{D_y} = \frac{aG}{2l} \sqrt{\frac{\pi}{2}} e^{\frac{1}{2}z^2}. \quad (11.3.14)$$

§ 11.4. Исследование точности автоматических систем методом статистических испытаний

Изложенные способы исследования нелинейных систем, основанные на линеаризации уравнений, оказываются неприменимыми к некоторым системам. Кроме того, применяя эти методы, мы не имеем возможности оценивать степень точности получаемых ре-

результатов, оценивать, насколько найденные значения характеристик точности близки к их истинным значениям. Вследствие этого возникает необходимость в непосредственном моделировании исследуемых систем и действия на них случайных возмущений. Это необходимо как для непосредственного исследования некоторых сложных систем, к которым трудно применить изложенные в предыдущих параграфах методы, так и для контрольных исследований с целью оценить точность и надежность результатов, получаемых с помощью этих методов.

Метод непосредственного моделирования систем под действием случайных возмущений обычно называется *методом статистических испытаний* *). Этот метод применяется как для непосредственного исследования различных систем, так и для исследования вероятностных моделей различных математических задач, например задач вычисления определенных интегралов, особенно многократных [9]. При исследовании системы методом статистических испытаний ее работа моделируется в условиях, близких к реальным. При этом на входы системы или ее модели подаются реализации случайных возмущений. В качестве модели системы могут быть использованы ее дифференциальные уравнения, решаемые на математической машине. Точность работы системы определяется путем статистической обработки результатов испытаний.

Предположим, что на вход системы в каждом опыте подается реализация $x_r(t)$ случайной функции $X(t)$ и в результате с выхода снимается соответствующая реализация $y_r(t)$ случайной функции $Y(t)$. Проведя n опытов, можно получить оценки $m_y^*(t)$ и $D_y^*(t)$, математического ожидания $m_y(t)$ и дисперсии $D_y(t)$ выходной переменной системы $Y(t)$, пользуясь известными формулами математической статистики:

$$m_y^*(t) = \frac{1}{n} \sum_{r=1}^n y_r(t), \quad D_y^*(t) = \frac{1}{n-1} \sum_{r=1}^n [y_r(t) - m_y^*(t)]^2. \quad (11.4.1)$$

Точность получаемых результатов можно оценивать средними квадратическими отклонениями оценок $m_y^*(t)$ и $D_y^*(t)$. Из теории вероятностей известно, что дисперсии оценок (11.4.1) определяются формулами (см. [53], § 111)

$$D[m_y^*(t)] = \frac{D_y(t)}{n}, \quad D[D_y^*(t)] = \frac{2D_y^2(t)}{n-1}. \quad (11.4.2)$$

Соответственно средние квадратические отклонения оценок определяются формулами

$$\sigma_{m_y^*} = \sqrt{\frac{D_y(t)}{n}} = \frac{\sigma_y(t)}{\sqrt{n}}, \quad \sigma_{D_y^*} = \sqrt{\frac{2}{n-1}} D_y(t). \quad (11.4.3)$$

*) Этот метод называется также *методом Монте-Карло*.

В качестве характеристик точности оценок (11.4.1) можно принять их относительные средние квадратические отклонения:

$$\frac{\sigma_{m_y^*}}{\sigma_y(t)} = \frac{1}{\sqrt{n}}, \quad \frac{\sigma_{D_y^*}}{D_y(t)} = \sqrt{\frac{2}{n-1}}. \quad (11.4.4)$$

При необходимости оценить абсолютные средние квадратические отклонения оценок $m_y^*(t)$ и $D_y^*(t)$ следует заменить в формулах (11.4.3) неизвестную величину $D_y(t)$ ее оценкой $D_y^*(t)$.

Для более полной оценки точности полученных результатов можно вычислить доверительные вероятности для заданных границ оценок $m_y^*(t)$ и $D_y^*(t)$, т. е. вероятности различных отклонений оценок от соответствующих вероятностных характеристик. Имея в виду, что при достаточно большом числе опытов n законы распределения оценок $m_y^*(t)$ и $D_y^*(t)$ близки к нормальному, на основании известных формул теории вероятностей получаем (см. [54], § 2.5)

$$\left. \begin{aligned} \alpha_1 &= P(|m_y^*(t) - m_y(t)| < \varepsilon_1) = 2\Phi\left(\frac{\varepsilon_1}{\sigma_{m_y^*}}\right), \\ \alpha_2 &= P(|D_y^*(t) - D_y(t)| < \varepsilon_2) = 2\Phi\left(\frac{\varepsilon_2}{\sigma_{D_y^*}}\right). \end{aligned} \right\} \quad (11.4.5)$$

Здесь ε_1 и ε_2 — отклонения оценок $m_y^*(t)$ и $D_y^*(t)$ от истинных значений, α_1 и α_2 — соответствующие доверительные вероятности.

Подставляя в (11.4.5) выражения (11.4.3) средних квадратических отклонений оценок, получим

$$\alpha_1 = 2\Phi\left(\frac{\varepsilon_1}{\sigma_y(t)} \sqrt{n}\right), \quad \alpha_2 = 2\Phi\left(\frac{\varepsilon_2}{D_y(t)} \sqrt{\frac{n-1}{2}}\right). \quad (11.4.6)$$

Вводя относительные отклонения оценок

$$v_1 = \frac{\varepsilon_1}{\sigma_y(t)}, \quad v_2 = \frac{\varepsilon_2}{D_y(t)}, \quad (11.4.7)$$

получим окончательно

$$\alpha_1 = 2\Phi(v_1 \sqrt{n}), \quad \alpha_2 = 2\Phi\left(v_2 \sqrt{\frac{n-1}{2}}\right). \quad (11.4.8)$$

В таблицах 11.4.1 и 11.4.2 приведены зависимости требуемого числа испытаний n от заданных доверительных вероятностей α_1 и α_2 и относительных отклонений v_1 и v_2 .

Относительные отклонения v_1 и v_2 определяют точность полученных результатов, а доверительные вероятности α_1 и α_2 — их надежность.

Из таблиц видно, что с ростом требований к точности и надежности результатов необходимое число испытаний резко возрастает. При этом повышение требований к надежности ведет к менее

быстрому росту необходимого числа испытаний, чем повышение требований к точности. Поэтому метод статистических испытаний может быть практически использован только в тех случаях, когда требуется получить точность решения порядка $15 \div 20\%$. И лишь в единичных случаях, когда требуется получить высокую точность, можно применять этот метод при очень большом числе n .

ТАБЛИЦА 11.4.1

Требуемое число испытаний для определения математического ожидания

$\alpha_1 \backslash v_1$	0,2	0,15	0,10	0,05	0,01
0,6	18	31	70	281	7 000
0,7	27	47	108	431	10 800
0,8	41	73	164	651	16 400
0,9	68	121	272	1090	27 200

ТАБЛИЦА 11.4.2

Требуемое число испытаний для определения дисперсии

$\alpha_2 \backslash v_2$	0,2	0,15	0,10	0,05	0,01
0,6	37	63	141	563	14 000
0,7	55	95	217	863	21 600
0,8	83	147	239	1300	32 800
0,9	137	243	545	2180	54 400

Следует отметить, что приведенные формулы и таблицы получены без учета ошибок самой математической машины и, следовательно, справедливы только в том случае, когда ошибки машины малы по сравнению с искомыми ошибками системы. В случае,

если ошибки машины соизмеримы с ошибками исследуемой системы, для получения надежного результата требуется очень большое число испытаний и применение метода становится нецелесообразным.

Низкая точность результатов является первым недостатком метода статистических испытаний. Вторым его недостатком является необходимость генератора случайных функций, имеющих заданные вероятностные характеристики. Создание такого генератора возможно далеко не для всех случайных функций. Практически просто создать генератор стационарной случайной функции, спектральная плотность которой постоянна в некотором диапазоне частот, т. е. случайной функции, которую можно приближенно считать белым шумом. Тогда для формирования случайной функции с заданной корреляционной функцией необходимо иметь устройство, преобразующее белый шум в требуемую случайную функцию, — формирующий фильтр. В § 7.6 мы видели, что для некоторых часто встречающихся случайных функций формирующие фильтры просты и могут быть легко реализованы. В более сложных случаях для формирования случайной функции с заданной корреляционной функцией можно использовать моделирующие устройства в соединении с генератором белого шума. Таким способом можно получить нестационарные случайные функции с разнообразными характеристиками. Для этого достаточно подать вырабатываемый генератором белый шум на вход модели, описываемой линейным дифференциальным уравнением с переменными коэффициентами. Еще большее многообразие случайных функций можно получить, если использовать модели, описываемые нелинейными дифференциальными уравнениями.

Для образования случайных функций в цифровых машинах необходимо генерировать последовательности случайных чисел. Преобразуя такую последовательность в машине по соответствующей программе, можно получить случайную последовательность с произвольно заданными вероятностными характеристиками. В некоторых задачах оказывается возможным пользоваться вместо случайных чисел последовательностями так называемых псевдослучайных чисел (подробнее об этом см. [9]).

Трудность формирования случайных функций в математических машинах и трудоемкость исследований ограничивают применимость метода статистических испытаний.

ПЕРЕДАЧА ИНФОРМАЦИИ ПО КАНАЛАМ СВЯЗИ

§ 12.1. Основные задачи теории информации. Каналы связи

Рассматривая в § 1.1 общие задачи управления, мы видели, что для организации процесса управления необходима информация о целях управления, о результатах управления и о всей обстановке, в которой происходит процесс управления. Эта информация передается от одного элемента системы управления к другому. Естественно, что вопросы, связанные с передачей информации, например вопросы объема передаваемой информации, скорости передачи информации, времени, необходимого для передачи заданного объема информации, имеют первостепенное значение для теории управления, особенно для сложных систем, предназначенных для управления большим количеством объектов и процессов.

Из теории вероятностей известно, что информация есть одна из форм проявления зависимости между разными явлениями. Образно выражаясь, можно сказать, что информация есть след, оставляемый одним явлением на другом. Благодаря этому следу можно по результатам наблюдения одного явления определить некоторые черты другого явления. Эти обстоятельства и лежат в основе всех устройств для передачи информации *). Используя различные физические явления, человек оставляет на них определенные следы в виде сигналов, несущих информацию. Наблюдая эти явления, другие люди или автоматические устройства изучают содержащиеся в данном явлении сигналы и таким образом выделяют передаваемую информацию.

Совокупность устройств, с помощью которых некоторое физическое явление используется для передачи информации, называется *каналом связи (передачи информации)* или, короче, просто *каналом*. Таким образом, общее понятие канала включает любые физические явления, которые могут быть носителями информации, например: ток, текущий по проводам, радиоволны, световые сигналы, звуковые сигналы и т. п.

*) Если явления независимы, то течение каждого из них не зависит от того, как протекает другое явление; одно из них не оставляет следа на другом. В этом случае ни одно из этих явлений не содержит никакой информации о другом. Поэтому для передачи информации можно пользоваться только зависимыми явлениями.

Процесс передачи информации можно представить схемой, изображенной на рис. 12.1.1. Информация X преобразуется в передатчике 1 в сигнал U , носителем которого является канал 2. В результате в приемнике 3 попадает сигнал V , из которого выделяется информация Y .

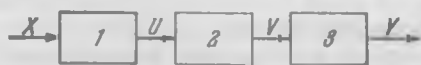


Рис. 12.1.1.

Любое физическое явление, используемое для передачи информации, всегда связано с бесчисленным множеством других явлений. Это приводит к тому, что на закономерности этого явления всегда накладываются случайные факторы, которые искажают передаваемые сигналы и затрудняют их обнаружение и выделение на приемном конце. Поэтому информация Y , выделяемая в приемнике (рис. 12.1.1), практически всегда несколько отличается от передаваемой информации X . Задача проектирования систем передачи информации заключается в том, чтобы сделать эти искажения по возможности минимальными. Таким образом, никакой канал не может передать информацию абсолютно точно, без искажений. В любом канале к передаваемым сигналам примешиваются случайные помехи или, что одно и то же, шумы. Поэтому передачу информации по различным каналам необходимо изучать с учетом помех. Принимая во внимание, что и само понятие информации является существенно вероятностным, нетрудно прийти к заключению, что в основе того раздела общей теории управления, в котором изучаются вопросы передачи информации — теории информации, — лежат вероятностные методы.

Для передачи информации необходимо прежде всего преобразовать подлежащую передаче информацию в определенного вида сигналы. На приемном конце эти сигналы надо обнаружить, отделить от шумов и выделить из них передаваемую информацию. Операция преобразования информации в сигналы называется *кодированием*. Обратная операция — преобразование передаваемых сигналов в информацию — называется *декодированием*. Кодирование и декодирование являются необходимыми элементами любого процесса передачи информации. Максимальная скорость передачи информации, т. е. максимальное количество информации, которое может быть передано по данному каналу в единицу времени, называется *пропускной способностью* канала.

Основными задачами теории информации являются кодирование информации, изучение свойств сигналов, несущих информацию, в частности определение количества содержащейся в них информации, расчет пропускной способности каналов, обнаружение сигналов в шумах, отделение их от шумов и декодирование. На основании изучения перечисленных задач можно поставить

общую задачу комплексного проектирования системы передачи информации, исходя из условия получения максимального эффекта на приемном конце при данных условиях. Однако такая комплексная задача слишком сложна, и в настоящее время наука не располагает средствами для ее решения. Поэтому в современной теории информации стараются найти оптимальное решение перечисленных задач по отдельности.

В этой главе мы изучим свойства сигналов, используемых для передачи информации, рассмотрим некоторые вопросы кодирования, изучим основные свойства каналов связи и научимся рассчитывать их пропускную способность. Задачи обнаружения сигналов в шумах и наилучшего отделения их от шумов, которые будут изучены в последующих главах, возникают не только в теории информации, но и в других разделах теории автоматического управления, в частности в теории проектирования следящих систем и в теории самонастраивающихся и самоорганизующихся систем.

§ 12.2. Сигналы, используемые для передачи информации. Измерение количества информации, содержащейся в сигналах

Во многих случаях информация передается при помощи языка как, например, в телеграфных передачах и телефонных разговорах. В таких случаях любое сообщение представляет собой конечную последовательность знаков — букв и промежутков между словами. Для передачи информации в подобных случаях используются сигналы, представляющие собой конечные последовательности знаков, например точек, тире и промежутков при применении азбуки Морзе в обычном телеграфе. Задача кодирования при этом заключается в установлении соответствия между передаваемыми знаками и буквами или словами передаваемого текста. Таким образом, одним из типичных сигналов, содержащих информацию и используемых для передачи информации, являются дискретные последовательности конечного числа знаков. Так как разные сообщения кодируются различными последовательностями знаков, то каждый знак последовательности следует рассматривать как случайную величину, которая может принимать различные возможные значения. Следовательно, в общем случае последовательность знаков, несущую информацию, необходимо рассматривать как случайную последовательность. При этом каждое конкретное сообщение будет представлять собой реализацию этой случайной последовательности.

Общее количество информации, содержащейся в данной конечной последовательности знаков, равно энтропии этой последовательности, так как после того, как все знаки последовательности становятся известными, в этой последовательности не остается никакой неопределенности (средняя условная энтропия системы конеч-

ного числа прерывных случайных величин относительно самой себя равна пулю) *). Таким образом, если последовательность содержит N знаков, представляющих собой случайные величины $X_1, \dots, X_N$, то полное количество информации в данной последовательности знаков равно энтропии совокупности случайных величин $X_1, \dots, X_N$:

$$H[X_1, \dots, X_N] = -M[\log_2 f(X_1, \dots, X_N)], \quad (12.2.1)$$

где $f(x_1, \dots, x_N)$ — совместная плотность вероятности случайных величин $X_1, \dots, X_N$. Количество информации, содержащееся в среднем в одном знаке, равно энтропии последовательности, приходящейся в среднем на один знак:

$$H_1 = -\frac{1}{N} M[\log_2 f(X_1, \dots, X_N)]. \quad (12.2.2)$$

Однако непосредственное вычисление количества информации по формулам (12.2.1) и (12.2.2) практически невозможно, так как число знаков в сообщении N обычно очень велико. Поэтому при определении количества информации, содержащейся в данной последовательности знаков, обычно вычисляют среднюю условную энтропию каждого знака, учитывая не полную зависимость между всеми знаками сообщения $X_1, \dots, X_N$, а лишь зависимость каждого знака X_k от небольшого числа непосредственно предшествующих знаков $X_{k-1}, \dots, X_{k-l}$. Эта зависимость определяется на основании изучения статистических свойств языка. Подобные исследования показывают, что для практических целей вполне достаточно учитывать зависимость каждого знака не больше чем от восьми непосредственно предшествующих.

На основании изложенного будем рассматривать последовательность знаков сигнала как стационарную цепь Маркова (см. [54], § 4.8), у которой условный закон распределения каждой случайной величины X_k относительно всех предшествующих величин зависит только от значений l непосредственно предшествующих величин $X_{k-1}, \dots, X_{k-l}$. Пусть $x_1, \dots, x_n$ — совокупность возможных значений каждой случайной величины X_k ($k = 1, \dots, N$), $p_{i_1, \dots, i_l}$ — вероятность того, что $X_{k-1}, \dots, X_{k-l}$ примут значения $x_{i_1}, \dots, x_{i_l}$ соответственно, $q_{i_1, \dots, i_l, j}$ — условная вероятность того, что величина X_k примет значение x_j при условии, что $X_{k-1}, \dots, X_{k-l}$ принимают значения $x_{i_1}, \dots, x_{i_l}$. Тогда условная энтропия случайной величины X_k относительно события $X_{k-1} = x_{i_1}, \dots, X_{k-l} = x_{i_l}$ определится формулой

*) Относительно понятий энтропии, средней условной энтропии и количества информации см. [54], гл. 7.

(см. [54], § 7.2)

$$H[X_k | X_{k-1} = x_{i_1}, \dots, X_{k-l} = x_{i_l}] = - \sum_{j=1}^n q_{i_1, \dots, i_l, j} \log_2 q_{i_1, \dots, i_l, j}. \quad (12.2.3)$$

Для вычисления средней условной энтропии случайной величины X_k относительно величин $X_{k-1}, \dots, X_{k-l}$ следует умножить выражение (12.2.3) на вероятность события $X_{k-1} = x_{i_1}, \dots, X_{k-l} = x_{i_l}$ и просуммировать по всем значениям $i_1, \dots, i_l$. В результате получим

$$\begin{aligned} H_1 &= H_{x_{k-l}, \dots, x_{k-1}}[X_k] = \\ &= - \sum_{i_1, \dots, i_l, j=1}^n p_{i_1, \dots, i_l} q_{i_1, \dots, i_l, j} \log_2 q_{i_1, \dots, i_l, j}. \end{aligned} \quad (12.2.4)$$

Эта формула и определяет количество информации, содержащееся в среднем в одном знаке последовательности *). Для сокращения записи мы будем обозначать совокупность индексов $i_1, \dots, i_l$ одной буквой l . Тогда формула (12.2.4) переписится в виде

$$H_1 = - \sum_{l, j=1}^n p_l q_{lj} \log_2 q_{lj}. \quad (12.2.5)$$

Пользуясь этой формулой, необходимо помнить, что сумма в этой формуле в общем случае является $l + 1$ -кратной.

Дискретные последовательности знаков, каждый из которых может принимать конечное число возможных значений, используются в качестве сигналов, несущих информацию, в телеграфе и в математических машинах дискретного действия, в частности в цифровых машинах.

*) Покажем, что в рассматриваемом случае среднее количество информации, содержащееся в одном знаке, можно считать равным $H_{x_{k-l}, \dots, x_{k-1}}[X_k]$. Для этого воспользуемся известной формулой (см. [54], § 7.2)

$$H[X_1, \dots, X_N] = H[X_1] + H_{x_1}[X_2] + \dots + H_{x_1, \dots, x_{N-1}}[X_N]. \quad (A)$$

Если условное распределение каждой случайной величины X_k зависит только от l непосредственно предшествующих, то каждый член суммы в (A), кроме l первых, равен $H_{x_{k-l}, \dots, x_{k-1}}[X_k]$ и формулы (12.2.1) и (12.2.2) дают

$$H_1 = \frac{H[X_1] + H_{x_1}[X_2] + \dots + H_{x_1, \dots, x_{l-1}}[X_l]}{N} + \frac{N-l}{N} H_{x_{k-l}, \dots, x_{k-1}}[X_k].$$

Отсюда видно, что если N велико по сравнению с l , то можно принять

$$H_1 = H_{x_{k-l}, \dots, x_{k-1}}[X_k].$$

Кроме дискретных последовательностей знаков, для передачи информации часто используются непрерывные функции времени. Примерами таких непрерывных сообщений могут служить устная речь, представляющая собой звуковые колебания воздуха, музыка, непрерывные колебания электромагнитного поля, используемые для передачи звука по радио, непрерывные колебания электромагнитного поля, вызываемые разверткой изображения при телевизионных передачах, сигналы, передаваемые по линиям обратной связи в непрерывных автоматических системах управления, и т. д. Так же как и дискретные последовательности знаков, непрерывные сигналы, несущие информацию, представляют собой случайные

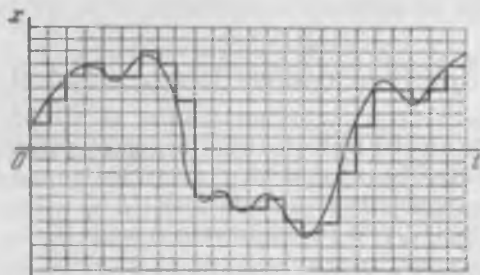


Рис. 12.2.1.

функции времени, а каждое конкретное непрерывное сообщение является возможной реализацией случайной функции.

Непрерывные сигналы всегда можно преобразовать в дискретные последовательности знаков. Это можно сделать различными способами. Так, например, разложив случайную функцию, представляющую со-

бой несущий информацию сигнал, в какой-нибудь ряд, например представив ее каноническим разложением, и ограничиваясь конечным отрезком ряда, можно вместо самого сигнала передать последовательность значений коэффициентов ряда.

Можно также аппроксимировать функцию ступенчатой функцией. Тогда достаточно будет передать последовательность ее значений в дискретном ряде точек (рис. 12.2.1). Эта операция обычно называется *квантованием сигнала по времени*. Кроме того, практически имеет смысл ограничиться дискретной шкалой передаваемых значений функции, т. е. заменить значение функции при любом фиксированном значении времени, представляющее собой в общем случае непрерывную случайную величину, прерывной случайной величиной, имеющей конечное число возможных значений. Эту операцию обычно называют *квантованием сигнала по уровню*.

Для обеспечения возможности однозначной расшифровки принимаемых сигналов обычно стараются выбирать шаг между соседними передаваемыми уровнями сигнала, т. е. «квант» изменения сигнала, так, чтобы он превышал уровень шума не меньшей мере в два раза. Так как уровень шумов обычно увеличивается, а сигнал ослабляется по мере увеличения дальности передачи информации, то для обеспечения надежной передачи информации на большие расстояния обычно вводят в линии связи промежуточные усилите-

ли, расположенные на таких расстояниях друг от друга, на которых уровень шумов не превышает половины кванта сигнала. Эти устройства однозначно и без ошибки восстанавливают принятый сигнал, воспроизводят его в очищенном от шумов виде и передают дальше по линии связи. На этом принципе основано, в частности, устройство радиорелейной линии.

Так как обычно каналы связи имеют ограниченную полосу пропускания, то для передачи информации практически целесообразно пользоваться сигналами с ограниченным спектром. Установим связь между диапазоном частот, содержащихся в спектре сигнала, и необходимой частотой передачи его значений. Для этого рассмотрим любую функцию $f(t)$, представимую интегралом Фурье (см., например, [50], ч. I, § 21)

$$f(t) = \int_{-\infty}^{\infty} F(\nu) e^{2\pi i \nu t} d\nu, \quad (12.2.6)$$

где $F(\nu)$ — изображение Фурье (спектр) функции $f(t)$:

$$F(\nu) = \int_{-\infty}^{\infty} f(t) e^{-2\pi i \nu t} dt. \quad (12.2.7)$$

Если функция $f(t)$ имеет ограниченный спектр, то $F(\nu)$ отлична от нуля только в некоторых пределах $|\nu| < \Omega$, и формула (12.2.6) примет вид

$$f(t) = \int_{-\Omega}^{\Omega} F(\nu) e^{2\pi i \nu t} d\nu \quad (12.2.8)$$

Всякая функция, заданная на конечном интервале, может быть разложена в ряд Фурье. Следовательно, функция $F(\nu)$, рассматриваемая как функция ν в интервале $-\Omega < \nu < \Omega$, может быть разложена в ряд Фурье с периодом по ν , равным 2Ω :

$$F(\nu) = \sum_{n=-\infty}^{\infty} a_n e^{\frac{i n \pi \nu}{\Omega}}. \quad (12.2.9)$$

Коэффициенты этого ряда определяются известной формулой теории рядов Фурье

$$a_n = \frac{1}{2\Omega} \int_{-\Omega}^{\Omega} F(\nu) e^{-\frac{i n \pi \nu}{\Omega}} d\nu. \quad (12.2.10)$$

На основании (12.2.8) эта формула может быть переписана в виде

$$a_n = \frac{1}{2\Omega} f\left(-\frac{n}{2\Omega}\right) \quad (n = 0, \pm 1, \pm 2, \dots). \quad (12.2.11)$$

Так как суммирование по n в (12.2.9) производится по всем положительным и отрицательным значениям n , то индекс n можно заменить на $-n$. Тогда, подставляя выражение (12.2.11) в формулу (12.2.9), получим

$$F(v) = \sum_{n=-\infty}^{\infty} a_n e^{-\frac{in\pi v}{\Omega}} = \frac{1}{2\Omega} \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2\Omega}\right) e^{-\frac{in\pi v}{\Omega}}. \quad (12.2.12)$$

Подставим теперь полученное выражение $F(v)$ в формулу (12.2.8):

$$f(t) = \frac{1}{2\Omega} \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2\Omega}\right) \int_{-\Omega}^{\Omega} e^{\frac{i\pi v}{\Omega}(2\Omega t - n)} dv. \quad (12.2.13)$$

Но

$$\begin{aligned} \int_{-\Omega}^{\Omega} e^{\frac{i\pi v}{\Omega}(2\Omega t - n)} dv &= \frac{\Omega}{\pi i(2\Omega t - n)} e^{\frac{i\pi v}{\Omega}(2\Omega t - n)} \Big|_{-\Omega}^{\Omega} = \\ &= \frac{\Omega}{\pi i(2\Omega t - n)} [e^{i\pi(2\Omega t - n)} - e^{-i\pi(2\Omega t - n)}] = \frac{2\Omega}{\pi} \cdot \frac{\sin \pi(2\Omega t - n)}{2\Omega t - n}. \end{aligned}$$

Подставляя это выражение в (12.2.13), получим окончательно

$$f(t) = \sum_{n=-\infty}^{\infty} f\left(\frac{n}{2\Omega}\right) \frac{\sin \pi(2\Omega t - n)}{\pi(2\Omega t - n)}. \quad (12.2.14)$$

Эта формула показывает, что любая функция с ограниченным спектром ширины 2Ω полностью определяется своими значениями в дискретном ряде точек, отделенных одна от другой интервалом времени $1/2\Omega$. Это положение было впервые доказано В. А. Котельниковым в 1933 г. [32, 33] и поэтому называется в нашей литературе теоремой Котельникова. В иностранной литературе это положение обычно называется теоремой отсчетов и связывается с именем основоположника теории информации американского ученого К. Шеннона *).

Формула (12.2.14) показывает, что если сигнал имеет спектр, ограниченный полосой частот 2Ω , то его значения необходимо передавать через интервалы времени $\Delta t = 1/2\Omega$. Иными словами, необходимо передавать $m = 1/\Delta t = 2\Omega$ значений сигнала в секунду, т. е. передавать сигнал с частотой $m = 2\Omega$.

*) Многие частные вопросы, связанные с передачей информации по каналам связи, в частности статистическая природа информации, были ясны некоторым ученым и до Шеннона. Однако лишь Шеннону в 1948 г. удалось отчетливо определить основные понятия теории информации и вывести основные ее законы [75].

После замены непрерывного сигнала дискретным количество информации, содержащееся в непрерывном сигнале, может быть вычислено так же, как и для дискретного сигнала.

Пример 12.2.1. Замкнутая система наведения самонаводящегося снаряда имеет полосу пропускания $\Omega = 2$ Гц. Для улучшения помехозащищенности системы решено применить дискретную систему измерения ошибки наведения, дающую ошибку в восьмых долях максимального значения ошибки (т. е. с семнадцатью уровнями ошибки). С каким периодом повторения T_n должна работать система измерения ошибки наведения и какое количество информации в секунду она должна передавать в блок формирования сигнала управления?

Спектр процесса изменения координат центра массы снаряда, а следовательно и процесса изменения ошибки наведения, ограничен полосой пропускания замкнутой системы наведения Ω . Следовательно, согласно формуле (12.2.14), система должна измерять ошибку наведения снаряда с периодом повторения $T_n = 1/2\Omega = 1/4$ с. Считая значения ошибки наведения, отделенные интервалом времени $1/4$ с, практически независимыми и все семнадцать значений ошибки $ke_{\max}$ ($k = 0, \pm \frac{1}{8}, \dots, \pm 1$) равновероятными,

находим количество информации, которое должно содержаться в каждом импульсе:

$$H_1 = - \sum_{i=1}^{17} p_i \log_2 p_i = \log_2 17 \approx 4,09 \frac{\text{бит}}{\text{имп}}.$$

Скорость передачи информации в блок формирования команд должна быть равна

$$H = \frac{H_1}{T_n} = \frac{4,09}{0,25} \approx 16,36 \frac{\text{бит}}{\text{с}}.$$

§ 12.3. Статистические свойства сигналов

Изучим подробнее статистические свойства сигналов, связанные с количеством информации, передаваемым по каналам связи. Предположим, что передаваемый сигнал является стационарной эргодической случайной последовательностью знаков, причем каждый знак представляет собой случайную величину с возможными значениями $x_1, \dots, x_n$. Кроме того, будем считать последовательность знаков сигнала цепью Маркова, в которой условный закон распределения каждого знака зависит только от значений l непосредственно предшествующих знаков. В этом случае количество информации, приходящееся в среднем на один знак последовательности, определяется формулой (12.2.5).

Оценим вероятность Q типичной реализации сигнала, в которой первые l знаков принимают значения $x_{1_1}, \dots, x_{1_l}$, а переход от значений $x_{1_1}, \dots, x_{1_l}$ предшествующих знаков к значению $x_{1_{l+1}}$ происходит $M_{1_1, \dots, 1_{l+1}} = M_{1j}$ раз ($i_1, \dots, i_l, j = 1, \dots, n$). Пользуясь обозначениями предыдущего параграфа, можем

написать

$$Q = p_{I_1, \dots, I_l} \prod_{i_1, \dots, i_l, j=1}^n q_{i_1, \dots, i_l, j}^{M_{i_1, \dots, i_l, j}} = p_I \prod_{i, j=1}^n q_{ij}^{M_{ij}}, \quad (12.3.1)$$

где $p_I = p_{I_1, \dots, I_l}$ — вероятность того, что первые l знаков примут значения $x_{I_1}, \dots, x_{I_l}$, а q_{ij} — вероятность перехода от значений $x_{i_1}, \dots, x_{i_l}$ предшествующих знаков к значению x_j . Так как заранее неизвестно, какие значения будут иметь знаки последовательности сигнала, то величины p_I , M_{ij} , а следовательно и Q , являются случайными. Так как по предположению последовательность знаков является эргодической стационарной, то при достаточно большом числе знаков последовательности N частота перехода от $x_{i_1}, \dots, x_{i_l}$ к x_j будет как угодно близка к его вероятности, т. е. при любых $\varepsilon > 0$, $\delta > 0$ и $i_1, \dots, i_l, j = 1, \dots, n$

$$P \left(\left| \frac{M_{ij}}{N} - p_i q_{ij} \right| < \varepsilon \right) > 1 - \delta. \quad (12.3.2)$$

Полагая

$$M_{ij} = N (p_i q_{ij} + S_{ij}), \quad (12.3.3)$$

можем переписать формулу (12.3.2) в виде

$$P (|S_{ij}| < \varepsilon) > 1 - \delta \quad (i_1, \dots, i_l, j = 1, \dots, n). \quad (12.3.4)$$

Логарифмируя формулу (12.3.1) и принимая во внимание (12.3.3), можем написать

$$\begin{aligned} \log_2 Q &= \log_2 p_I + \sum_{i, j=1}^n M_{ij} \log_2 q_{ij} = \\ &= \log_2 p_I + N \sum_{i, j=1}^n p_i q_{ij} \log_2 q_{ij} + N \sum_{i, j=1}^n S_{ij} \log_2 q_{ij}. \end{aligned} \quad (12.3.5)$$

Отсюда на основании (12.2.5) получаем

$$\frac{1}{N} \log_2 Q = \frac{1}{N} \log_2 p_I - H_1 + \sum_{i, j=1}^n S_{ij} \log_2 q_{ij}. \quad (12.3.6)$$

При достаточно большом N последнее слагаемое в (12.3.6) на основании (12.3.4) будет сколь угодно мало с вероятностью, сколь угодно близкой к единице *). Поэтому для любых $\eta > 0$ и $\delta > 0$ имеем

$$P \left(\left| \frac{1}{N} \log_2 \frac{1}{Q} - H_1 \right| < \eta \right) > 1 - \delta. \quad (12.3.7)$$

*) Ни одно из q_{ij} в (12.3.1) и (12.3.6) не может быть равно нулю, так как в этом случае $Q = 0$ и соответствующая реализация последовательности знаков не может быть типичной.

Но из неравенства

$$\left| \frac{1}{N} \log_2 \frac{1}{Q} - H_1 \right| < \eta$$

следуют неравенства

$$2^{-N(H_1+\eta)} < Q < 2^{-N(H_1-\eta)}. \quad (12.3.8)$$

Таким образом, мы доказали, что из всех возможных реализаций последовательности знаков с вероятностью, как угодно близкой к единице, появляется одна из реализаций, вероятности которых заключены в пределах (12.3.8) при любом $\eta > 0$, если только последовательность достаточно длинна. Грубо говоря, можно сказать, что практически всегда появляется одна из возможных реализаций последовательности, имеющих вероятность, равную $Q = 2^{-NH_1}$. Число таких реализаций не может быть больше 2^{NH_1} . А так как суммарная вероятность этих реализаций как угодно близка к единице, то можно считать, что число практически возможных реализаций равно 2^{NH_1} и вероятность каждой из них равна 2^{-NH_1} .

Пример 12.3.1. Сообщение передается двоичным кодом. Энтропия последовательности на один знак равна 0,75 бит. Найти общее число возможных реализаций и число практически возможных реализаций последовательности, содержащей 100 знаков. В данном случае $N = 100$, $n = 2$, $H_1 = 0,75$. Общее число возможных реализаций равно $n^N = 2^{100}$. Число практически возможных реализаций на основании изложенного равно $2^{NH_1} = 2^{100 \cdot 0,75} = 2^{75}$. Отношение числа всех возможных реализаций к числу практически возможных равно

$$\frac{n^N}{2^{NH_1}} = \frac{2^{100}}{2^{75}} = 2^{25} = 1024^2,5 > 30 \cdot 10^4.$$

Таким образом, в данном случае число практически возможных реализаций сигнала составляет меньше чем одну тридцатимиллионную долю числа всех возможных реализаций. Для более длинных последовательностей эта доля будет еще значительно меньше. Так, например, для $N = 1000$ отношение числа всех возможных реализаций к числу практически возможных реализаций больше 10^{75} .

Пример 12.3.2. Передача сообщения ведется на русском языке буквами русского алфавита. Энтропия русской речи, приходящаяся в среднем на одну букву, приблизительно равна двум двоичным знакам. Определить число практически возможных сообщений, составленных из 100 букв, и сравнить его с числом всех возможных последовательностей по 100 букв русского алфавита. Так как в русском алфавите 32 буквы, то $n = 32$, $N = 100$, $H_1 = 2$. Отсюда $2^{NH_1} = 2^{100 \cdot 2} = 2^{200}$ и

$$\frac{n^N}{2^{NH_1}} = \frac{32^{100}}{2^{200}} = \frac{32^{100}}{32^{200}} = 32^{-100} = 1024^{-30} > 10^{90}.$$

Полученный результат имеет чрезвычайно большое значение для практики, так как позволяет резко ограничить число практически возможных реализаций передаваемых сигналов. А именно доказанное предложение и рассмотренные примеры показывают, что из полного числа $n^N = 2^{N \log_2 n}$ всех возможных реализаций

сигнала практически возможно появление лишь 2^{NH_1} реализаций. При большом N и $H_1 < H_{1\max} = \log_2 n$ число 2^{NH_1} составляет лишь ничтожно малую долю числа n^N . Таким образом, лишь ничтожно малая доля всех возможных реализаций оказывается практически возможной, а подавляющее число остальных реализаций имеет ничтожно малую суммарную вероятность появления.

Изложенное показывает, какую большую роль играет энтропия случайной последовательности для практики. Зная энтропию последовательности, можно определить число практически возможных ее реализаций и учитывать в практических расчетах только это относительное небольшое число реализаций, а с подавляющим числом остальных реализаций совершенно не считаться.

Второй из рассмотренных примеров показывает, что из всех возможных последовательностей, составленных из 100 русских букв, лишь ничтожно малая доля представляет собой имеющие смысл последовательности русских букв, а остальные являются случайными наборами букв. Это говорит о том, что язык обладает большой избыточностью. Действительно, так как

$$2^{NH_1} = n^{\frac{NH_1}{\log_2 n}},$$

то каждой имеющей смысл последовательности из N букв можно поставить в соответствие другую последовательность, составленную из $s = NH_1 / \log_2 n$ букв. Иными словами, все имеющие смысл сообщения, составленные из N букв, можно закодировать последовательностями, составленными из s букв каждая. При этом все возможные последовательности из s букв будут иметь смысл. Так как для любого европейского языка величина $H_1 / \log_2 n$ не превышает 0,5, то все имеющие смысл сообщения можно закодировать приблизительно вдвое более короткими последовательностями букв. Это и доказывает, что лишь около 50% букв в нашей речи независимы, а остальные практически полностью определяются законами грамматики. Иными словами, наш язык обладает большой избыточностью, и лишь приблизительно 50% букв в нем несут информацию.

Избыточность языка позволяет практически достоверно расшифровывать сообщения с ошибками, заменяя ошибочные буквы правильными. Именно поэтому мы легко обнаруживаем опечатки в тексте и исправляем их. Если бы язык не обладал избыточностью, то каждая ошибочная буква меняла бы смысл написанного и опечатки было бы невозможно обнаруживать и исправлять.

Применим изложенные соображения к задаче кодирования сообщений. Для того чтобы закодировать сообщение, необходимо последовательности его знаков поставить в соответствие другую последовательность знаков. Простейшим примером кодирования может служить замена последовательности букв алфавита последователь-

ностями точек и тире азбуки Морзе. Другим примером кодирования может служить сопоставление звукам устной речи последовательностей букв. Кодированием является также сопоставление определенных групп импульсов различным командам, передаваемым по линии телеуправления.

Из теории вероятностей известно, что при любом преобразовании случайных величин содержащееся в них количество информации о некоторых других случайных величинах не может увеличиться, а может только уменьшиться. И только в случае обратимого функционального преобразования количество информации остается неизменным. Отсюда следует, что кодирование может сопровождаться потерей информации и только в случае обратимого кода, т. е. кода, допускающего однозначную расшифровку (декодирование), количество информации остается неизменным. Поэтому необходимо всегда стремиться применять обратимые коды.

Очевидно, что при кодировании надо стараться по возможности сокращать текст, так как при этом сокращается время передачи сообщений и увеличивается скорость передачи информации. Выше мы видели, что вследствие избыточности речи можно при кодировании сжать текст в $\log_2 n/H_1$ раз. При этом обеспечивается взаимно однозначное соответствие между имеющими смысл последовательностями, содержащими по N букв, и всеми возможными последовательностями по $s = NH_1/\log_2 n$ букв, т. е. обеспечивается обратимость кода и возможность его однозначной расшифровки. Легко понять, что при сжатии текста больше чем в $\log_2 n/H_1$ раз код станет необратимым, так как число возможных последовательностей букв будет меньше числа практически возможных сообщений.

Однако практическое осуществление кодирования с максимальным сжатием текста очень затруднительно вследствие практической невозможности перебрать все возможные варианты сообщений различной длительности. Поэтому практически осуществляется кодирование по целых сообщений, а отдельных групп знаков. При этом, очевидно, чем чаще встречается данная группа знаков, тем более короткой последовательностью знаков ее следует кодировать. Иными словами, при кодировании следует стремиться к тому, чтобы кодировать более короткими последовательностями знаков те группы знаков, которые имеют большую вероятность появления.

§ 12.4. Передача информации по каналу с шумами.

Пропускная способность канала

Вследствие искажения передаваемых сигналов шумами, неизбежно действующими в канале, каждому принятому сигналу Y , вообще говоря, не будет соответствовать один возможный переданный сигнал X . Иными словами, после приема сигнала Y переданный

сигнал X не будет полностью определен, а остается случайным. Неопределенность переданного сигнала X будет характеризоваться средней условной энтропией X относительно Y . Поэтому количество информации о переданном сигнале X , содержащееся в принятом сигнале Y , определяется формулой ([54], § 7.4)

$$I_Y[X] = H[X] - H_Y[X]. \quad (12.4.1)$$

Количество информации в каждом знаке передаваемого сигнала в зависимости от способа кодирования будет различным. Поэтому количество информации, передаваемое по каналу в единицу времени, будет зависеть от способа кодирования и от числа передаваемых в единицу времени знаков сигнала. Максимальное количество информации, которое может быть передано по данному каналу в единицу времени, называется *пропускной способностью* этого канала. Считая, что через $H[X]$ и $H_Y[X]$ в (12.4.1) обозначены соответственно энтропия и средняя условная энтропия сигнала, приходящиеся на единицу времени, и обозначая пропускную способность канала через C , можем записать сформулированное определение пропускной способности канала в виде

$$C = \max \{I_Y[X]\} = \max \{H[X] - H_Y[X]\}. \quad (12.4.2)$$

Предположим, что информация передается по каналу со скоростью H , меньшей пропускной способности канала C , и оценим вероятность ошибки при расшифровке принятого сигнала. Для этого обозначим через $H^{(m)}$ и $H_Y^{(m)}$ соответственно значения $H[X]$ и $H_Y[X]$, при которых достигается максимум в (12.4.2). Тогда (12.4.2) даст

$$C = H^{(m)} - H_Y^{(m)}. \quad (12.4.3)$$

На основании доказанного в предыдущем параграфе максимальное число практически возможных реализаций переданного сигнала данной длительности T при том способе кодирования, при котором $H[X]$ достигает максимального значения $H^{(m)}$, равно $2^{TH^{(m)}}$, а число практически возможных реализаций сигнала X , соответствующих одной реализации принятого сигнала Y , при этом равно $2^{TH_Y^{(m)}}$. Иными словами, максимальное число практически возможных различных последовательностей знаков, каждая из которых может быть передана в течение времени T по данному каналу, равно $2^{TH^{(m)}}$. При этом каждой принятой последовательности знаков соответствует $2^{TH_Y^{(m)}}$ практически возможных переданных последовательностей. Это положение схематически изображено на рис. 12.4.1.

Если информация, содержащаяся в сигнале X , передается со скоростью H бит/с, то число практически возможных реализаций этого сигнала длительности T равно $2^{TH} < 2^{TH^{(m)}}$. Рассмотрим такой

способ кодирования, при котором каждой из практически возможных реализаций переданного сигнала сопоставляется одна из $2^{TH^{(m)}}$ последовательностей знаков, т. е. из общего числа $2^{TH^{(m)}}$ точек на рис. 12.4.1 выбирается 2^{TH} точек. При этом вероятность того, что данная точка будет занята одной из возможных реализаций передаваемого сигнала, равна $2^{TH}/2^{TH^{(m)}} = 2^{-T(H^{(m)}-H)}$. Из теории вероятностей известно, что вероятность появления хотя бы одного из нескольких событий не может быть больше суммы вероятностей этих событий. Поэтому вероятность того, что при кодировании будет занята хотя бы одна из $2^{TH_Y^{(m)}}$ точек, соответствующих данной реализации принятого сигнала Y , кроме той, которая соответствует действительно переданной реализации сигнала X , будет меньше чем $2^{TH_Y^{(m)}} \cdot 2^{-T(H^{(m)}-H)}$, т. е. для вероятности ошибки $P_{ош}$ при передаче сигнала X имеем неравенство

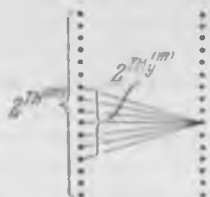


Рис. 12.4.1.

$$P_{ош} \leq 2^{TH_Y^{(m)}} \cdot 2^{-T(H^{(m)}-H)} = 2^{-T(H^{(m)}-H_Y^{(m)}-H)}$$

или, принимая во внимание (12.4.3),

$$P_{ош} \leq 2^{-T(C-H)}. \quad (12.4.4)$$

При $H < C$ это выражение стремится к нулю при $T \rightarrow \infty$. Таким образом, мы доказали, что при передаче информации по каналу с пропускной способностью C со скоростью $H < C$ можно закодировать передаваемый сигнал так, что вероятность ошибки будет как угодно мала при достаточно большой длительности передаваемого сигнала T . Это предположение представляет собой основную теорему Шеннона [75]. Более тонкие математические методы дают возможность доказать, что эта теорема верна и при $H = C$ [69].

Очевидно, однако, что возможность передачи информации по каналу с шумами с как угодно малой вероятностью ошибки может быть достигнута лишь ценой запаздывания, необходимого для расшифровки принятого сигнала, так как декодирование можно осуществить лишь после того, как будет передана вся последовательность знаков сигнала.

Рассмотрим теперь случай передачи информации по каналу, имеющему пропускную способность C , со скоростью $H > C$. Из (12.4.2) следует, что при любом кодировании

$$H - H_Y \leq C, \quad (12.4.5)$$

откуда

$$H_Y \geq H - C. \quad (12.4.6)$$

Таким образом, при передаче информации по каналу с пропускной

способностью C со скоростью $H > C$ неопределенность в принятом сигнале не может быть меньше чем $H - C$ бит/с. В этом состоит вторая основная теорема Шеннона.

Число практически возможных реализаций переданного сигнала при данном принятом сигнале в этом случае на основании результатов предыдущего параграфа равно 2^{TH_v} , причем все эти реализации практически одинаково вероятны и вероятность каждой из них равна 2^{-TH_v} . Следовательно, вероятность выбора действительно переданной реализации из 2^{TH_v} практически возможных при данном принятом сигнале равна 2^{-TH_v} , а вероятность ошибки равна $1 - 2^{-TH_v}$. Отсюда, принимая во внимание (12.4.6), получаем для вероятности ошибки неравенство

$$P_{\text{ош}} = 1 - 2^{-TH_v} \geq 1 - 2^{-T(H-C)}. \quad (12.4.7)$$

Из этого неравенства следует, что вероятность ошибки стремится к единице при $T \rightarrow \infty$, если $H > C$. Этот результат показывает, что передача информации со скоростью, превышающей пропускную способность канала, практически невозможна.

Перейдем к выводу формулы для расчета пропускной способности канала. Предположим, что канал имеет полосу пропускания Ω , т. е. что он передает только гармонические колебания всех частот от 0 до Ω Гц. Обозначим через P среднюю мощность сигнала, несущего информацию, а через N — среднюю мощность шумов в канале. Несущий информацию сигнал $X(t)$ и шум $V(t)$ будем считать независимыми эргодическими стационарными случайными функциями времени с тождественно равными нулю математическими ожиданиями и дисперсиями D_x и D_v соответственно. Величины D_x и D_v вследствие эргодичности сигнала и шума пропорциональны их средним мощностям P и N .

Из теории вероятностей известно, что количество информации о случайной величине X , содержащееся в случайной величине Y , равно количеству информации о величине Y , содержащемуся в величине X :

$$I_y[X] = I_x[Y] = H[Y] - H_x[Y]. \quad (12.4.8)$$

Если шум V суммируется в канале с сигналом X , то

$$Y(t) = X(t) + V(t) \quad (12.4.9)$$

$$H_x[Y] = H[V]. \quad (12.4.10)$$

Таким образом, средняя условная энтропия принимаемого сигнала Y равна энтропии шума и, следовательно, не зависит от способа кодирования информации.

На основании (12.4.8), (12.4.9) и (12.4.10) формулу (12.4.2) для пропускной способности канала можно переписать в виде

$$C = \max \{H[X + V]\} - H[V]. \quad (12.4.11)$$

Так как шум обычно можно считать нормально распределенной случайной функцией времени, то энтропию его значения при данном значении t можно вычислить по формуле (см. [54], § 7.3)

$$H_1[V] = \log_2 \sqrt{2\pi e D_v}. \quad (12.4.12)$$

При этом значения шума, разделенные промежутком времени между двумя последовательно передаваемыми значениями сигнала, можно считать независимыми. Тогда, если в секунду по каналу передается s значений сигнала, то энтропия шума в единицу времени будет равна

$$H[V] = s \log_2 \sqrt{2\pi e D_v}. \quad (12.4.13)$$

Для вычисления $\max \{H[X + V]\}$ вспомним, что из всех случайных величин, имеющих одну и ту же дисперсию $D_x + D_v$, максимальной энтропией обладают нормально распределенные случайные величины. Поэтому наибольшее количество информации в каждом значении сигнала можно получить, если пользоваться для передачи информации нормально распределенным случайным сигналом. В этом сигнале наибольшая возможная энтропия на один знак сигнала равна

$$\max \{H_1[X + V]\} = \log_2 \sqrt{2\pi e (D_x + D_v)}. \quad (12.4.14)$$

А так как совместная энтропия нескольких случайных величин не может быть больше суммы их энтропий и равна сумме их энтропий только когда они независимы, то для получения максимальной энтропии сигнала $X + V$ на единицу времени необходимо пользоваться сигналом X с независимыми значениями. В этом случае при передаче в единицу времени s значений сигнала будем иметь

$$\max \{H[X + V]\} = s \log_2 \sqrt{2\pi e (D_x + D_v)}. \quad (12.4.15)$$

Из (12.4.11), (12.4.13) и (12.4.15), следует, что при передаче s значений сигнала в единицу времени максимальное количество информации, которое может быть передано в единицу времени по каналу, равно

$$C_s = \max \{H[X + V]\} - H[V] = \frac{s}{2} \log_2 \left(1 + \frac{D_x}{D_v} \right). \quad (12.4.16)$$

Остается теперь определить наибольшее возможное число значений сигнала s , которое может быть передано по каналу в единицу времени. Так как канал имеет полосу пропускания Ω , то по нему может передаваться только сигнал, имеющий спектр, ограниченный полосой частот Ω . На основании теоремы Котельникова такой сигнал полностью определяется своими значениями, отделенными промежутком времени $T_n = 1/2\Omega$. Поэтому максимальное число значений сигнала, которое может быть передано в единицу времени, равно $s = 1/T_n = 2\Omega$. Следовательно, пропускная способность

канала, представляющая собой максимальное значение выражения (12.4.16), равна

$$C = \Omega \log_2 \left(1 + \frac{D_x}{D_s} \right). \quad (12.4.17)$$

Заменяя здесь отношение дисперсий сигнала и шума отношением их средних мощностей, получим окончательно следующую формулу для пропускной способности канала:

$$C = \Omega \log_2 \left(1 + \frac{P}{N} \right). \quad (12.4.18)$$

Формула (12.4.18) показывает, что пропускная способность канала определяется его полосой пропускания Ω и отношением сигнала к шуму в канале $\gamma = \sqrt{P/N}$. Зная эти параметры, можно определить пропускную способность, и, наоборот, задаваясь пропускной способностью C и одним из параметров Ω , γ , можно определить значение другого параметра, при котором обеспечивается заданная пропускная способность канала.

Пример 12.4.1. Определить необходимую пропускную способность канала телеуправления, если рассеивание телеуправляемых снарядов можно считать круговым со средним квадратическим отклонением σ , рассеивание неуправляемых снарядов того же типа является круговым со средним квадратическим отклонением σ_1 , а время полета снаряда до цели равно t . Определить необходимую полосу пропускания канала при отношении сигнала к шуму в канале γ .

Считая закон рассеивания снарядов нормальным, находим энтропию рассеивания снарядов в случае телеуправления и в случае отсутствия управления:

$$H_{\text{туп}} = 2 \log_2 \sigma \sqrt{2\pi e}, \quad H_{\text{нуп}} = 2 \log_2 \sigma_1 \sqrt{2\pi e}.$$

Следовательно, количество информации, которое канал передачи команд телеуправления должен передавать на снаряд за время его полета, равно

$$I = 2 \log_2 \frac{\sigma_1}{\sigma}.$$

Необходимая пропускная способность канала равна

$$C = \frac{I}{t} = \frac{2}{t} \log_2 \frac{\sigma_1}{\sigma}.$$

Сравнивая это выражение с (12.4.18), находим

$$\frac{2}{t} \log_2 \frac{\sigma_1}{\sigma} = \Omega \log_2 (1 + \gamma^2),$$

откуда

$$\Omega = \frac{2}{t} \frac{\log_2 \sigma_1 - \log_2 \sigma}{\log_2 (1 + \gamma^2)} = \frac{2}{t} \frac{\lg \sigma_1 - \lg \sigma}{\lg (1 + \gamma^2)}.$$

Пример 12.4.2. В условиях примера 12.2.1 найти необходимое отношение сигнала к шуму в канале. Необходимая полоса пропускания канала, как было показано в примере 12.2.1, равна $\Omega = 2$ Гц, а необходимая скорость передачи информации $C = 16,36$ бит/с. Следовательно, пользуясь формулой (12.4.18), находим

$$\gamma^2 = \frac{P}{N} = 2^{\frac{C}{\Omega}} - 1 = 2^{8,18} - 1 = 289, \quad \gamma = 17.$$

ЗАДАЧИ СТАТИСТИЧЕСКОЙ ТЕОРИИ
ОПТИМАЛЬНЫХ СИСТЕМ

§ 13.1. Постановка задач теории оптимальных систем

В главах 7 и 11 были изложены методы исследования точности автоматических систем, имеющих заданные характеристики. Однако на практике часто возникает другая задача: спроектировать систему, т. е. определить ее характеристики, таким образом, чтобы она обладала наибольшей возможной точностью при данных условиях. Систему, обладающую наибольшей возможной точностью с какой-нибудь определенной точки зрения среди всех систем данного класса, обычно называют *оптимальной*. Раздел теории автоматического управления, развивающий методы непосредственного определения оптимальных систем, называется *теорией оптимальных систем*. Величина, характеризующая качество системы, максимальное или минимальное значение которой достигается для оптимальной системы, называется *критерием оптимальности*.

Оптимальную систему можно найти подбором с помощью методов, изложенных в главах 7 и 11. Для этого надо рассчитать различные варианты системы одним из изложенных в главах 7 и 11 методов, сравнить их и выбрать вариант, для которого принятый критерий имеет наименьшее или наибольшее, в зависимости от характера задачи, значение. Однако этот путь сложен и требует большого количества вычислений, которые часто бывают очень трудоемкими.

Задачу непосредственного определения оптимальной системы можно ставить двумя способами. В первой постановке задается структура системы и требуется найти оптимальные значения ее числовых параметров, при которых ее точность будет наилучшей с точки зрения выбранного критерия. В такой постановке задача определения оптимальной системы сводится к обычной задаче отыскания экстремума функции одной или нескольких переменных. Выразив с помощью методов главы 7 или главы 11 принятый критерий качества системы через ее неизвестные параметры, мы можем найти такие значения этих параметров, при которых критерий качества имеет наибольшее или наименьшее значение в зависимости от его смысла. При этом в зависимости от того, как выражается критерий качества через параметры системы, можно применить для нахождения экстремума известные аналитические или числен-

ные методы. В § 13.3 будут даны примеры нахождения оптимальных значений параметров системы, имеющей заданную структуру.

Во второй постановке задачи система считается полностью неизвестной и требуется определить ее характеристики (в общем случае — оператор) так, чтобы она была оптимальной с точки зрения принятого критерия качества. Идеальным решением задачи в такой постановке, на первый взгляд, является определение структуры оптимальной системы и всех ее параметров. Однако пока еще не существует теории, способной дать решение в такой форме. Да и вряд ли такое решение удовлетворило бы конструкторов. При проектировании системы необходимо учитывать не только точность, но и многие другие факторы. Проектируемая система должна удовлетворять многим, подчас противоречивым требованиям. Требование близости к оптимуму по точности является, конечно, одним из важнейших, но далеко не единственным. Не менее важными являются требования надежности и простоты обслуживания системы. Поэтому проектирование любой системы обычно представляет собой ряд компромиссных решений с целью наилучшим образом удовлетворить всем предъявляемым к системе требованиям. Поэтому определение структуры и параметров системы, оптимальных с точки зрения какого-нибудь одного критерия, например критерия точности, обычно не дает существенной помощи конструктору. Гораздо важнее дать конструктору метод, с помощью которого он может, оставаясь в пределах достаточной близости к оптимальному решению с точки зрения точности, иметь достаточно свободы для того, чтобы разумным образом удовлетворить и всем остальным требованиям. Теория оптимальных систем дает методы определения операторов оптимальных систем и соответствующих экстремальных значений принятого критерия качества. Иными словами, она позволяет находить предельные потенциальные качества систем данного назначения для данных условий работы. Сравнивая рассматриваемые при проектировании варианты системы с теоретической оптимальной системой, конструктор может судить о том, насколько эти варианты близки к оптимуму с точки зрения точности.

В большинстве практических задач возможны значительные отступления от оптимальных характеристик без существенного ухудшения системы с точки зрения принятого критерия качества (точности). Это характерное свойство задач теории оптимальных систем является весьма положительным фактором, так как позволяет конструктору после определения предельной, потенциальной точности проектируемой системы в широких пределах варьировать ее структуру и параметры без существенного отклонения от оптимальной точности и тем самым удовлетворять многим другим требованиям, предъявляемым к проектируемой системе, в частности требованиям простоты и надежности.

Для уяснения характера задач теории оптимальных систем рассмотрим в качестве примера задачу проектирования оптимальной следящей системы, предназначенной для воспроизведения некоторой переменной величины. Всякая следящая система предназначена для работы в различных условиях, для воспроизведения различных по характеру сигналов. Так, например, следящая система радиолокатора, предназначенная для измерения координат летящих самолетов, должна следить за самолетами, летящими с различными скоростями, различными курсами, на разных высотах. Эта следящая система должна воспроизводить все возможные различные законы изменения входной переменной. Вследствие этого естественно рассматривать полезный сигнал, который должна вос-



Рис. 13.1.1.

производить следящая система, как случайную функцию времени, а различные конкретные виды этого сигнала — как возможные реализации этой случайной функции. Кроме полезного сигнала, на входе следящей системы всегда действует помеха. И даже при отсутствии внешней помехи в аппаратуре следящей системы возникают внутренние шумы, которые являются случайными функциями времени. Таким образом, следящая система, которую необходимо спроектировать, получает на входе сумму полезного сигнала $S(t)$ и помехи или шума $N(t)$ (рис. 13.1.1). На выходе следящая система должна с максимальной точностью воспроизвести сигнал $S(t)$. Решить эту задачу абсолютно точно принципиально невозможно. Объясняется это тем, что действующие на входе системы шумы и помехи, накладываются на полезный сигнал, при любой конструкции системы вызывают непрерывные случайные колебания — флуктуации — выходного сигнала. Вследствие этого выходной сигнал системы ни при каких условиях нельзя сделать точно равным требуемому сигналу. Действительный выходной сигнал системы будет некоторой случайной функцией $W^*(t)$, не совпадающей с подлежащим воспроизведению полезным сигналом $S(t)$. Разность

$$E(t) = W^*(t) - S(t) \quad (13.1.1)$$

представляет собой ошибку следящей системы. При проектировании следящей системы возникает задача обеспечения в известном смысле минимальной ошибки при всех возможных условиях работы системы.

Задача нахождения оптимальной следящей системы является далеко не единственной задачей такого рода. Так, например, часто требуется спроектировать экстраполирующую систему, у которой требуемым значением выходной переменной в каждый данный момент времени t является значение полезного сигнала $S(t + \Delta)$

в некоторый будущий момент времени $t + \Delta$ ($\Delta > 0$). Для такой системы ошибка выражается формулой

$$E(t) = W^*(t) - S(t + \Delta). \quad (13.1.2)$$

В общей теории линейных систем (глава 2) мы видели, что идеальный экстраполятор построить принципиально невозможно, даже в условиях полного отсутствия шумов и помех, так как он является физически невозможной системой. Поэтому возникает задача нахождения такой системы, для которой ошибка экстраполяции имеет минимальное значение.

В качестве третьего примера рассмотрим задачу проектирования дифференцирующего устройства, представляющего собой систему, требуемым выходным сигналом которой в каждый данный момент t является значение производной $S'(t)$ полезного сигнала. Ошибка дифференциатора определяется формулой

$$E(t) = W^*(t) - S'(t). \quad (13.1.3)$$

Как и в предыдущих рассмотренных примерах, построить идеальную систему оказывается невозможным, и возникает задача нахождения такой системы, для которой ошибка имеет минимальное возможное значение.

Все рассмотренные задачи являются частными случаями задачи проектирования системы, предназначенной для выполнения какой-то заданной операции L над полезным сигналом $S(t)$, содержащимся во входном сигнале системы. Во втором из рассмотренных примеров эта операция представляет собой сдвиг во времени (т. е. экстраполяцию) на величину Δ . В третьем примере операция L представляет собой дифференцирование. Таким образом, в общем случае требуемым выходным сигналом проектируемой системы является результат выполнения некоторой операции L над полезным сигналом:

$$W(t) = LS(t). \quad (13.1.4)$$

Ошибка системы при этом будет равна

$$E(t) = W^*(t) - W(t). \quad (13.1.5)$$

Оптимальной системой, предназначенной для выполнения операции L над входным сигналом, будет такая система, ошибка которой имеет минимальное (в известном смысле) значение.

Перейдем теперь к общей постановке основной задачи теории оптимальных систем. Обозначим через $Z(t)$ входную функцию системы. В общем случае безразлично, каким образом она составлена из полезного сигнала и помехи. Обычно входная функция представляет собой сумму полезного сигнала $S(t)$ и помехи $N(t)$:

$$Z(t) = S(t) + N(t). \quad (13.1.6)$$

Требуемый выходной сигнал системы обозначим через $W(t)$. В общем случае безразлично, как он связан с полезным сигналом. Обычно он представляет собой результат заданного преобразования полезного сигнала и выражается формулой (13.1.4), где L — заданный оператор. Действительным выходным сигналом системы является некоторая случайная функция $W^*(t)$, которая представляет собой результат преобразования данной системой входной функции $Z(t)$:

$$W^*(t) = AZ(t), \quad (13.1.7)$$

где A — оператор системы.

Принципиальное отличие действительного выходного сигнала системы $W^*(t)$ от требуемого $W(t)$ состоит в том, что действительный выходной сигнал $W^*(t)$ всегда представляет собой результат преобразования полной входной функции $Z(t)$, состоящей из полезного сигнала и помехи, в то время как требуемый выходной сигнал $W(t)$ является результатом преобразования одного полезного сигнала. Иными словами, в отличие от требуемой выходной переменной системы, действительная выходная переменная всегда содержит помеху, от которой принципиально невозможно полностью избавиться.

На основании изложенного основная задача теории оптимальных систем заключается в нахождении такой системы, которая, получая на входе случайную функцию $Z(t)$, воспроизводит на выходе с максимальной возможной точностью другую случайную функцию $W(t)$. Математически эта задача сводится к нахождению оптимального оператора системы A , обеспечивающего наилучшую точность системы, т. е. наилучшее воспроизведение требуемого выходного сигнала $W(t)$.

Сформулированная основная задача теории оптимальных систем относится к классу задач, рассматриваемых в математической статистике. Вследствие этого вполне естественной тенденцией является максимальное использование достижений современной математической статистики для решения задач теории оптимальных систем.

Действительный выходной сигнал системы $W^*(t)$ в соответствии с принятой в статистике терминологией представляет собой *оценку* требуемого выходного сигнала $W(t)$. Задачей оптимальной системы является автоматическая выработка оптимальной оценки требуемого выходного сигнала по данной входной функции. Основная задача теории оптимальных систем, поставленная выше, может рассматриваться как задача нахождения оптимального преобразования входной функции, дающего в результате оптимальную оценку требуемого выходного сигнала. Таким образом, по самому своему существу теория оптимальных систем не дает возможности находить оптимальные системы в том смысле, в каком это интересно конструктора, т. е. находить структуру и параметры оптималь-

ных систем, а определяет лишь *математические операции*, которые должна производить оптимальная система над входной функцией. Вследствие этого теория оптимальных систем по существу не является теорией статистического синтеза автоматических систем, хотя ее часто и называют так.

Поставленная основная задача теории оптимальных систем остается неопределенной до тех пор, пока не будет установлено, что значит требование минимальности ошибки системы, т. е. пока не будет выбран критерий качества, экстремальное значение которого должно достигаться для оптимальной системы. Вопрос этот не может быть решен в рамках самой теории оптимальных систем. При решении его нужно руководствоваться прежде всего характером задач, которые должна решать искомая оптимальная система, и соображениями о том, какие величины могут служить мерой успешности решения этих задач. При этом легко убедиться в том, что не существует единственного универсального критерия качества для всех систем и что для оценки качества (точности) систем различного назначения целесообразно пользоваться различными критериями.

Пример 13.1.1. Рассмотрим задачу проектирования оптимального корректирующего устройства для стационарной линейной следящей системы.

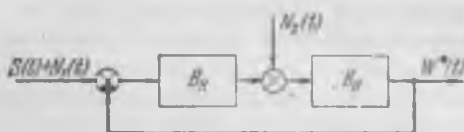


Рис. 13.1.2.

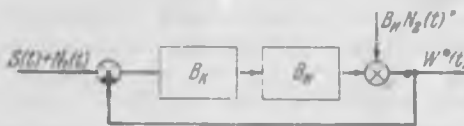


Рис. 13.1.3.

исполнительное устройство которой задано (рис. 13.1.2), предполагая, что на следящую систему действуют две помехи: $N_1(t)$, складывающаяся с полезным сигналом $S(t)$, и $N_2(t)$, складывающаяся с выходным сигналом корректирующего устройства и действующая на входе исполнительного устройства.

Обозначим через B_k и B_n соответственно операторы (линейные) корректирующего устройства и исполнительного устройства. Тогда элементарными структурными преобразованиями, показанными на рис. 13.1.3 и 13.1.4, мы приведем поставленную задачу к определению оптимальной замкнутой системы

мы, обведенной на рис. 13.1.4 пунктирным прямоугольником, которая должна вырабатывать наилучшую оценку полезного сигнала $S(t) - B_n N_2(t)$, получая на входе случайную функцию $S(t) + N_1(t) - B_n N_2(t)$. Определив передаточную функцию оптимальной замкнутой системы и зная передаточную функцию исполнительного устройства, мы можем по формулам § 4.6 найти оптимальную передаточную функцию корректирующего устройства. Таким образом, задача проектирования оптимального корректирующего устройства сводится к сформулированной основной задаче теории оптимальных систем в случае, когда входной функцией системы является

$$Z(t) = S(t) - B_n N_2(t) + N_1(t),$$

в требуемым выходным сигналом —

$$W(t) = S(t) - B_N N_2(t).$$

Пример 13.1.2. Предположим, что требуется определить динамические характеристики некоторой системы в процессе ее работы. В этом случае

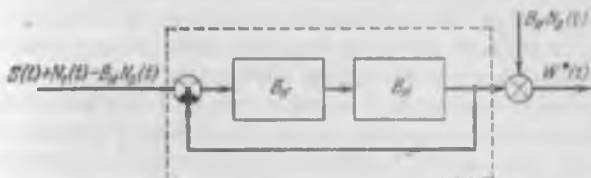


Рис. 13.1.4.

обычно записывают входную и выходную переменные системы как функции времени и по этим записям находят их вероятностные характеристики. Динамические характеристики системы определяются ее оператором, т. е. законом, по которому каждой данной входной функции приводится в соответствие выходная функция. Обозначим через $Z(t)$ входную функцию системы, а через $W(t)$ ее выходную функцию. Вследствие ошибок в измерении и в записи входной и выходной функций точное определение оператора системы принципиально невозможно. Поэтому задачу определения характеристик системы приходится ставить следующим образом: найти такую модель изучаемой системы, выходная переменная которой $W^*(t)$ насколько возможно более близка к выходной переменной системы $W(t)$ (рис. 13.1.5). Таким образом, задача экспериментального определения динамических характеристик какой-либо системы в процессе ее нормальной работы также сводится к поставленной выше основной задаче теории оптимальных систем. При этом требуемой выходной переменной искомой модели системы является выходная переменная системы $W(t)$.



Рис. 13.1.5.

§ 13.2. Статистические критерии оптимальности автоматических систем

Ошибка системы, определяемая формулой (13.1.5), является случайной функцией времени. Поэтому она не может непосредственно служить оценкой точности системы в силу своей неопределенности. Вследствие этого естественно взять какое-то ее вероятностное среднее значение. Учитывая, что знак ошибки в большинстве случаев нас не интересует, естественно принять за характеристику точности системы математическое ожидание квадрата ошибки:

$$\eta = M [E^2(t)]. \quad (13.2.1)$$

Эта величина представляет собой начальный момент второго порядка ошибки системы. Положительный корень квадратный из этой

величины обычно называется *средней квадратической ошибкой* системы. Таким образом, оптимальной системой можно считать такую систему, которая обладает минимальной средней квадратической ошибкой.

Критерий минимума средней квадратической ошибки является простейшим с математической точки зрения и обычно приводит к наиболее простым методам определения оптимальных систем. Однако далеко не во всех задачах он может служить подходящей мерой качества системы. Поэтому теория оптимальных систем не может ограничиться методами нахождения оптимальных систем по критерию минимума средней квадратической ошибки.

Величина η , как момент второго порядка ошибки, может быть выражена через математическое ожидание и дисперсию ошибки, т. е. является функцией математического ожидания и дисперсии ошибки. Поэтому естественным обобщением критерия минимума средней квадратической ошибки является критерий экстремума заданной функции математического ожидания и дисперсии ошибки:

$$f(M[E], D[E]) = \text{extremum}. \quad (13.2.2)$$

Это обобщение критерия минимума средней квадратической ошибки предложено Н. И. Андреевым, который вывел общие необходимые и достаточные условия для определения оптимальной системы по этому критерию в произвольном классе систем и дал общее решение задачи определения оптимальной линейной системы [2, 3] (см. также [53], гл. 16 и 17*). Если закон распределения ошибки полностью определяется ее математическим ожиданием и дисперсией, например, если ошибка распределена нормально, то частным случаем критерия (13.2.2) является критерий максимума вероятности того, что ошибка не выйдет из заданных пределов:

$$P(|E(t)| < \varphi(t)) = \max, \quad (13.2.3)$$

где $\varphi(t)$ — заданная функция.

Перечисленные критерии основаны на количественной оценке ошибки приближения, т. е. разности между фактическим и требуемым выходными сигналами. Вследствие этого они могут применяться только для решения некоторых задач воспроизведения сигналов. В задачах обнаружения сигналов, для которых имеет значение не величина ошибки, а лишь качественный факт наличия или отсутствия ошибки, приходится пользоваться другими критериями, например критерием минимума вероятности ошибочного решения. Предполагая, что система решает, что сигнал есть, если ее выходной сигнал W^* превышает некоторый пороговый уровень s , и ре-

*) Независимо от Н. И. Андреева решение задачи определения оптимальной линейной системы по критерию вида (13.2.2) было дано В. С. Рыбаковым.

имеет, что сигнала нет, если $W^* \leq c$, можем записать этот критерий в виде

$$P_{\text{ош}} = P \left(\frac{W \neq 0}{W^* \leq c} \right) + P \left(\frac{W = 0}{W^* > c} \right) = \min. \quad (13.2.4)$$

Вторым критерием для решения задачи обнаружения сигналов может служить критерий условного минимума вероятности ошибочного решения при заданной условной вероятности ложного обнаружения сигнала, когда он отсутствует во входной случайной функции, т. е. критерий (13.2.4) при дополнительном условии

$$P(W^* > c \mid W = 0) = \alpha. \quad (13.2.5)$$

Этот критерий обычно называется *критерием Неймана — Пирсона*.

Заметим теперь, что все рассмотренные выше критерии оптимальности, кроме (13.2.2), а также большое количество различных других статистических критериев можно представить в общей форме

$$M[l(W, W^*)] = \min, \quad (13.2.6)$$

где $l(W, W^*)$ — некоторая функция требуемого и действительного выходных сигналов. В зависимости от выбора функции $l(W, W^*)$ получается тот или иной критерий оптимальности.

В частности, при

$$l(W, W^*) = (W^* - W)^2 = E^2 \quad (13.2.7)$$

формула (13.2.6) дает критерий минимума средней квадратической ошибки (13.2.1).

При

$$l(W, W^*) = \begin{cases} 1 & \text{при } |W^* - W| \geq \varphi(t), \\ 0 & \text{при } |W^* - W| < \varphi(t) \end{cases} \quad (13.2.8)$$

получается критерий минимума вероятности выхода ошибки из заданного интервала $(-\varphi(t), \varphi(t))$. Действительно, при таком выборе функции $l(W, W^*)$

$$\begin{aligned} M[l(W, W^*)] &= 1 \cdot P(|E(t)| \geq \varphi(t)) + 0 \cdot P(|E(t)| < \varphi(t)) = \\ &= P(|E(t)| \geq \varphi(t)) = 1 - P(|E(t)| < \varphi(t)). \end{aligned}$$

Чтобы представить в виде (13.2.6) критерий минимума вероятности ошибки при решении задачи обнаружения сигнала (13.2.4), возьмем такую функцию $l(W, W^*)$, которая равна нулю при правильном решении задачи обнаружения, т. е. когда $W = 0, W^* \leq c$ или $W^* \neq 0, W^* > c$, и равна единице при ошибочном решении, т. е. когда $W \neq 0, W^* \leq c$ или $W = 0, W^* > c$:

$$l(W, W^*) = \begin{cases} 0 & \text{при } W = 0, W^* \leq c \text{ и при } W \neq 0, W^* > c, \\ 1 & \text{при } W \neq 0, W^* \leq c \text{ и при } W = 0, W^* > c. \end{cases} \quad (13.2.9)$$

Математическое ожидание этой функции равно произведению значения при правильном решении на вероятность правильного решения плюс произведение ее значения при ошибочном решении на вероятность ошибочного решения:

$$M[l(W, W^*)] = 0 \cdot (1 - p_{\text{ош}}) + 1 \cdot p_{\text{ош}} = p_{\text{ош}}. \quad (13.2.10)$$

Для отыскания условного минимума вероятности $p_{\text{ош}}$ при дополнительном условии (13.2.5), согласно методу неопределенных множителей Лагранжа, необходимо найти минимум величины

$$p_1 = p_{\text{ош}} + \lambda_1 P(W^* > c | W = 0), \quad (13.2.11)$$

где λ_1 — неопределенный множитель, и после этого определить λ_1 так, чтобы было выполнено условие (13.2.5). Обозначая через p и q соответственно вероятность наличия и вероятность отсутствия сигнала во входной случайной функции ($q = 1 - p$), можем переписать равенство (13.2.11) в виде

$$\begin{aligned} p_1 &= pP(W^* \leq c | W \neq 0) + q\left(1 + \frac{\lambda_1}{q}\right)P(W^* > c | W = 0) = \\ &= P\left(\frac{W \neq 0}{W^* \leq c}\right) + \lambda P\left(\frac{W = 0}{W^* > c}\right), \end{aligned} \quad (13.2.12)$$

где

$$\lambda = 1 + \frac{\lambda_1}{q}. \quad (13.2.13)$$

Из формулы (13.2.12) видно, что критерий минимума величины p_1 также можно представить в форме (13.2.6), если определить функцию $l(W, W^*)$ формулой

$$l(W, W^*) = \begin{cases} 1 & \text{при } W \neq 0, W^* \leq c, \\ \lambda & \text{при } W = 0, W^* > c, \\ 0 & \text{при } W \neq 0, W^* > c \text{ и при } W = 0, W^* \leq c, \end{cases} \quad (13.2.14)$$

где λ — неопределенный параметр.

Таким образом, критерий минимума вероятности ошибки и критерий условного минимума вероятности ошибки при заданной условной вероятности ложного обнаружения сигнала при его отсутствии во входной случайной функции являются частными случаями критерия (13.2.6), причем для первого критерия функция l определена полностью формулой (13.2.9), а для второго критерия она содержит неопределенный параметр λ , который должен быть определен после нахождения минимума из дополнительного условия (13.2.5).

В задачах практики часто оказывается полезным также критерий (13.2.6) при выборе функции l в виде

$$l(W, W^*) = 1 - e^{-\lambda^2(W^* - W)^2}. \quad (13.2.15)$$

Таким образом, мы получаем следующий общий принцип оценки качества системы и выбора критерия оптимальности. Качество решения задачи в каждом конкретном случае оценивается некоторой функцией $l(W, W^*)$, значение которой определяется конкретными реализациями требуемого выходного сигнала W и действительного выходного сигнала системы W^* . Эту функцию целесообразно назвать *функцией потерь*. Качество решения задачи в среднем для данной реализации требуемого выходного сигнала W при всех возможных реализациях действительного выходного сигнала W^* , соответствующих данной реализации сигнала W , оценивается условным математическим ожиданием функции потерь при данной реализации требуемого выходного сигнала:

$$\rho(A | W) = M[l(W, W^*) | W]. \quad (13.2.16)$$

Эту величину обычно называют *условным риском*. Условный риск зависит от оператора A , определяющего W^* , и от реализации требуемого выходного сигнала W , что и отражено в обозначениях в формуле (13.2.16). Наконец, среднее качество решения задачи при всех возможных реализациях сигналов W и W^* характеризуется математическим ожиданием условного риска, равным безусловному математическому ожиданию функции потерь:

$$R(A) = M[\rho(A | W)] = M[l(W, W^*)]. \quad (13.2.17)$$

Эту величину естественно назвать *средним риском*. Установив таким образом меру качества системы, мы можем принять за критерий оптимальности системы *критерий минимума среднего риска*, который имеет форму (13.2.6).

Если требуется обеспечить наилучшее решение задачи при некоторых дополнительных условиях, то, представив эти дополнительные условия в виде

$$\sigma_k(A) = M[\varphi_k(W, W^*)] = \alpha_k \quad (k = 1, \dots, N) \quad (13.2.18)$$

и применяя метод неопределенных множителей Лагранжа, мы сведем задачу к отысканию минимума величины

$$R_1(A) = R(A) + \sum_{k=1}^N \lambda_k \sigma_k(A) = M[l(W, W^*) + \sum_{k=1}^N \lambda_k \varphi_k(W, W^*)]. \quad (13.2.19)$$

Эту величину, очевидно, можно рассматривать как средний риск соответствующий функции потерь

$$l_1(W, W^*) = l(W, W^*) + \sum_{k=1}^N \lambda_k \varphi_k(W, W^*), \quad (13.2.20)$$

зависящей от неопределенных параметров $\lambda_1, \dots, \lambda_N$. В рассмотренном выше случае критерия Неймана — Пирсона для

обнаружения сигналов дополнительное условие (13.2.5) может быть представлено в форме (13.2.18), если принять

$$\varphi_1(W, W^*) = \begin{cases} 0 & \text{при } W \neq 0 \text{ и при } W = 0, W^* \leq c, \\ \frac{1}{q} & \text{при } W = 0, W^* > c \end{cases} \quad (13.2.21)$$

Таким образом, практически все статистические критерии являются частными случаями критерия минимума среднего риска, соответствующими различным способам выбора функции потерь. При этом функция потерь полностью определена и известна заранее, если требуется обеспечить безусловный минимум среднего риска, и зависит от неопределенных параметров, если требуется обеспечить условный минимум среднего риска при некоторых дополнительных условиях.

Критерии минимума среднего риска, соответствующие всем возможным определенным функциям потерь, которые могут содержать неопределенные параметры, обычно называются *бейсовыми критериями* по имени английского ученого Бейеса (Bayes). Таким образом, все рассмотренные выше критерии являются бейсовыми. Решение задачи определения оптимальной оценки сигнала по бейсову критерию называется *бейсовым решением*.

К сожалению, вопрос о выборе критерия оптимума, который, очевидно, сводится к выбору функции потерь l , не всегда может быть практически решен на основе анализа назначения и условий работы проектируемой системы. Решение этого вопроса в значительной мере определяется имеющимися в нашем распоряжении данными о вероятностных характеристиках входного и требуемого выходного сигналов. Очевидно, что для определения оптимальной системы по критерию (13.2.6) при произвольной функции потерь необходимо полное знание закона распределения входного и требуемого выходного сигналов. Только для некоторых частных видов функции потерь задача определения оптимальной системы может быть решена на основе неполного знания закона распределения случайных функций Z и W . Так, например, как мы увидим в §§ 13.5—13.7, для определения оптимальной линейной системы по критерию минимума средней квадратической ошибки достаточно знать математические ожидания случайных функций Z , W , корреляционную функцию случайной функции Z и взаимную корреляционную функцию случайных функций W , Z . С другой стороны, если известны только математические ожидания и корреляционные функции случайных функций Z , W и нет никаких других данных об их законе распределения, то нет никакого смысла искать оптимальную систему среди нелинейных систем. Действительно, при данных математических ожиданиях и корреляционных функциях случайные функции Z и W могут иметь самые разнообразные законы распределения и, в частности, могут быть распре-

делены нормально. А при нормальном распределении случайных функций Z и W , как мы увидим в § 15.3, оптимальной системой среди всех возможных систем по отношению к любому критерию вида (13.2.6), соответствующему функции потерь l , представляющей собой функцию модуля ошибки, является некоторая линейная система. А так как нормальный закон распределения вследствие его широкого распространения во многих случаях может считаться наиболее вероятным, то оптимальная линейная система будет наиболее вероятной оптимальной системой в классе всех возможных систем. Вследствие этого часто приходится ограничиваться определением оптимальной линейной системы по критерию минимума средней квадратической ошибки. При нормальном законе распределения сигнала и помех эта система оказывается оптимальной и с точки зрения многих других критериев.

Если известен только условный закон распределения входного сигнала Z относительно требуемого выходного сигнала W , а закон распределения сигнала W неизвестен, то оптимальную систему можно искать по критерию минимума максимального возможного значения условного риска $\rho(A | W)$:

$$\max_W \rho(A | W) = \min, \quad (13.2.22)$$

т. е. наибольшего из значений условного риска, соответствующих всем возможным реализациям сигнала W . Этот критерий, называемый обычно *минимаксным*, обеспечивает наилучшую работу системы в наихудших возможных условиях. Вследствие этого он особенно подходит для расчета систем, работающих в условиях организованного противодействия, которое осуществляется путем выбора реализации сигнала, обеспечивающего наихудшую точность решения задачи. В случае, когда закон распределения требуемого выходного сигнала неизвестен или даже вообще не существует, минимаксный критерий часто оказывается целесообразным потому, что в этом случае неизвестно, насколько часто могут встречаться реализации требуемого выходного сигнала, при которых точность системы близка к наихудшей, а минимаксный критерий обеспечивает наименьшую возможную потерю точности именно для таких сигналов.

В некоторых случаях, когда система управления должна работать в условиях организованного противодействия, т. е. когда имеется противник, стремящийся дезорганизовать, нарушить работу системы управления, требуемый выходной сигнал $W(t)$ зависит от некоторых параметров, которыми противник может располагаться. Его задача состоит в таком выборе значений этих параметров, чтобы нанести по возможности больший ущерб работе системы управления, воспрепятствовать решению поставленных перед системой управления задач. Задачей системы управления в таких случаях является организовать управление так, чтобы

свести к возможному минимуму эффект противодействия, т. е. наносимый противником ущерб работе системы управления. Такая обстановка типична, например, для управления боевыми действиями войск. Другим примером может служить стрельба управляемым снарядом по самолету, совершающему оборонительный маневр. В этом случае самолет стремится маневрировать так, чтобы по возможности уменьшить вероятность его поражения снарядом. Система управления снарядом должна вырабатывать при этом такие сигналы управления, чтобы свести к минимуму снижение вероятности поражения самолета, т. е. эффект его маневрирования. Третьим примером может служить наведение управляемого снаряда с радиолокационной системой управления на цель при наличии организованных помех. В этом случае противник стремится выбрать помехи так, чтобы по возможности больше помешать работе системы управления полетом снаряда, свести к минимуму вероятность поражения цели. Система управления должна при этом вырабатывать такие сигналы управления, чтобы вероятность поражения цели в условиях противодействия была возможно больше.

Задачи определения оптимальных решений, когда имеются две стороны, интересы которых противоположны и каждая из которых стремится получить максимальную выгоду для себя и нанести максимальный ущерб противоположной стороне, являются предметом сравнительно нового раздела теории вероятностей, так называемой теории игр. Для создания оптимальных систем управления в условиях организованного противодействия необходимо привлекать теорию игр. Впрочем, в настоящее время работы по организации процессов управления с помощью методов теории игр еще только начинаются и теория систем управления, вырабатывающих оптимальные решения в конфликтных ситуациях, еще пока недостаточно развита [90, 91].

§ 13.3. Определение оптимальных параметров системы, имеющей заданную структуру

Если структура системы задана, то для определения оптимальных значений ее параметров необходимо найти зависимость принятого критерия качества от параметров системы. В простейших случаях эту зависимость можно найти в аналитической форме. В таких случаях оптимальные значения параметров системы определяются системой уравнений, полученных приравливанием нулю частных производных критерия качества по неизвестным параметрам системы.

Пример 13.3.1. Найти оптимальное значение постоянной времени фильтра RC , предназначенного для отфильтровывания полезного сигнала от помехи, если полезный сигнал представляет собой линейную функцию

времени со случайными коэффициентами, имеющими нулевые математические ожидания, а помеха представляет собой белый шум постоянной интенсивности k , независимый от полезного сигнала. За критерий качества системы принять установившееся значение средней квадратической ошибки или, что одно и то же, математическое ожидание квадрата разности между установившимся значением выходной переменной фильтра и ее требуемому значением (равным по условию задачи входному полезному сигналу).

В данном случае входным сигналом фильтра является случайная функция

$$Z(t) = U_1 + U_2 t + X(t),$$

где U_1 и U_2 — случайные величины, математические ожидания которых равны нулю, а $X(t)$ — независимый от величин U_1, U_2 белый шум. Требуемым выходным сигналом фильтра является случайная функция

$$W(t) = U_1 + U_2 t.$$

Ошибка фильтра представляет собой сумму ошибки воспроизведения полезного сигнала и результата прохождения через фильтр помехи. Так как по своему назначению фильтр является следящей системой, то для вычисления ошибки воспроизведения полезного сигнала можно воспользоваться формулой (7.3.12) и найденными в примере 7.3.1 коэффициентами ошибок. В результате получим

$$E(t) = -U_2 T + Y_{\text{ш}}(t),$$

где $Y_{\text{ш}}(t)$ — выходной шум фильтра, который на основании изложенного в § 7.4 является стационарной случайной функцией. Дисперсия выходного шума находится по данной спектральной плотности $k/2\pi$ входного белого шума и данной передаточной функции фильтра по формуле (7.4.4):

$$D_{Y_{\text{ш}}} = \frac{k}{2\pi} \int \frac{d\omega}{|1 + T i\omega|^2} = \frac{k}{2T}.$$

Так как случайная величина U_2 и помеха $X(t)$ по условию независимы и их математические ожидания равны нулю, то квадрат средней квадратической ошибки фильтра, равный в данном случае дисперсии ошибки, определяется формулой

$$\eta = M[E^2(t)] = T^2 D_{u_2} + D_{Y_{\text{ш}}} = T^2 D_{u_2} + \frac{k}{2T}.$$

Для нахождения оптимальной постоянной времени T' приравняем нулю значение производной величины η по T при $T = T'$ и решаем полученное уравнение. В результате получим

$$T' = \sqrt[3]{\frac{k}{4D_{u_2}}}.$$

Подставляя это выражение в формулу для η , находим минимальную среднюю квадратическую ошибку, соответствующую оптимальной постоянной времени T' :

$$\sqrt{\eta_{\text{min}}} = T' \sqrt{3D_{u_2}} = \frac{1}{2} \sqrt{\frac{3k}{T'}} = \sqrt{\frac{3}{2}} \cdot \sqrt[3]{\frac{k^2 D_{u_2}}{2}}.$$

Проанализируем теперь, насколько ухудшается точность фильтра, если выбирать различные значения постоянной времени T , отличные от оптимального значения T' . Для этого построим график зависимости величины η/η_{min}

от отношения T/T' (рис. 13.3.1). Из этого графика видно, что при отклонении постоянной времени от оптимального значения на $30 \div 35\%$, средняя квадратическая ошибка фильтра увеличивается всего на $5 \div 6\%$. Отсюда можно сделать вывод, что значительные отклонения от оптимальной постоянной времени в данном случае не приводят к существенному ухудшению

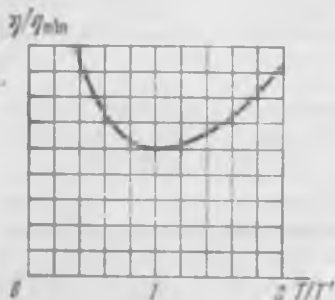


Рис. 13.3.1.

точности фильтра. Этот пример может служить иллюстрацией высказанного в § 13.1 общего положения, что во многих задачах практики можно допускать значительные отклонения от оптимальных характеристик системы без существенного ухудшения ее точности.

В сложных случаях, когда уравнения, полученные приравнованием нулю производных критерия качества по параметрам системы, нельзя решить аналитически или когда не удастся выразить критерий качества аналитически через параметры системы, для нахождения оптимальных

значений параметров системы приходится прибегать к приближенным численным методам. Наиболее эффективным численным методом нахождения минимума функции является метод наискорейшего спуска и различные его разновидности. Случай, когда требуется найти максимум функции, очевидно, приводится к задаче нахождения минимума изменением знака функции.

Для выяснения идеи метода наискорейшего спуска рассмотрим сначала случай, когда требуется найти оптимальные значения двух параметров системы α_1, α_2 . Критерий качества является функцией $f(\alpha_1, \alpha_2)$, минимум которой соответствует оптимальной системе. Равноценным с точки зрения принятого критерия качества системам соответствует на плоскости параметров α_1, α_2 кривая

$$f(\alpha_1, \alpha_2) = c, \quad (13.3.1)$$

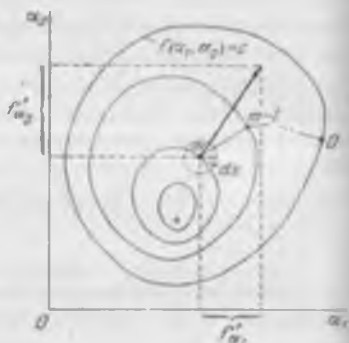


Рис. 13.3.2.

которая является линией уровня поверхности $z = f(\alpha_1, \alpha_2)$. Различным значениям параметра c соответствуют различные линии уровня (рис. 13.3.2).

Пусть $(\alpha_1^{(m)}, \alpha_2^{(m)})$ — точка, для которой вычислено значение функции f . Проведем из этой точки как из центра окружность бесконечно малого радиуса ds и найдем на этой окружности точку, в которой функция f имеет минимальное значение. Предполагая, что существуют непрерывные частные производные функции f по α_1 ,

α_2 . находим дифференциал функции f в точке $(\alpha_1^{(m)}, \alpha_2^{(m)})$:

$$df = f'_{\alpha_1} d\alpha_1 + f'_{\alpha_2} d\alpha_2 = (f'_{\alpha_1} \cos \varphi + f'_{\alpha_2} \sin \varphi) ds, \quad (13.3.2)$$

где аргументы функций f'_{α_1} , f'_{α_2} для краткости опущены. Для нахождения минимума выражения (13.3.2) приравниваем нулю его производную по φ :

$$-f'_{\alpha_1} \sin \varphi + f'_{\alpha_2} \cos \varphi = 0.$$

Отсюда находим

$$\cos \varphi = \pm k f'_{\alpha_1}, \quad \sin \varphi = \pm k f'_{\alpha_2}, \quad \frac{1}{k} = \sqrt{f'^2_{\alpha_1} + f'^2_{\alpha_2}}, \quad (13.3.3)$$

где должны быть взяты либо оба верхних знака, либо оба нижних. Очевидно, что df будет положительным, если взять верхние знаки, и отрицательным, если взять нижние знаки. Таким образом, нижние знаки в (13.3.3) соответствуют минимуму df и мы доказали, что при смещении от точки $(\alpha_1^{(m)}, \alpha_2^{(m)})$ на данную бесконечно малую величину ds наилучшее приближение к минимуму функции f достигается, если выбрать смещение в направлении вектора с составляющими $-f'_{\alpha_1}$, $-f'_{\alpha_2}$. Но вектор с составляющими f'_{α_1} , f'_{α_2} представляет собой вектор градиента функции f на плоскости параметров α_1 , α_2 , нормальный к линии уровня (13.3.1) функции f (рис. 13.3.2). Следовательно, для быстрейшего приближения к минимуму функции f , т. е. для скорейшего спуска по поверхности, изображающей функцию f , необходимо из каждой точки двигаться в направлении, противоположном направлению вектора градиента функции f . Иными словами, в каждой расчетной точке необходимо выбирать приращения параметров системы пропорциональными соответствующим составляющим вектора градиента функции f с отрицательным коэффициентом пропорциональности — s_m :

$$\Delta \alpha_1^{(m)} = -s_m f'_{\alpha_1}(\alpha_1^{(m)}, \alpha_2^{(m)}), \quad \Delta \alpha_2^{(m)} = -s_m f'_{\alpha_2}(\alpha_1^{(m)}, \alpha_2^{(m)}).$$

В результате получим на плоскости α_1 , α_2 ломаную, вершинами которой являются расчетные точки (рис. 13.3.2). Если увеличивать неограниченно число расчетных точек, неограниченно уменьшая отрезки ломаной, то в пределе получим кривую, по которой будет течь по поверхности функции f вода, вылитая в какой-либо точке этой кривой.

Совершенно так же доказывается, что для скорейшего приближения к минимуму функции n параметров $f(\alpha_1, \dots, \alpha_n)$ необходимо для каждой расчетной комбинации значений параметров $\alpha_1^{(m)}, \dots, \alpha_n^{(m)}$ выбирать приращения параметров $\alpha_1, \dots, \alpha_n$ пропорциональными соответствующим частным производным функции f с отрицательным коэффициентом пропорциональности — s_m , т. е. двигаться в n -мерном пространстве параметров $\alpha_1, \dots, \alpha_n$ в направлении, противоположном вектору градиента функции f .

Для доказательства положим

$$d\alpha_l = \xi_l ds \quad (l = 1, \dots, n), \quad \xi_1^2 + \dots + \xi_n^2 = 1. \quad (13.3.4)$$

Тогда дифференциал функции f выразится формулой

$$df = \sum_{l=1}^n f'_{\alpha_l} d\alpha_l = \sum_{l=1}^n f'_{\alpha_l} \xi_l ds. \quad (13.3.5)$$

Так как $\xi_1, \dots, \xi_n$ связаны вторым уравнением (13.3.4), то для нахождения минимума df в данном случае удобнее всего воспользоваться методом неопределенных множителей Лагранжа. Согласно этому методу для нахождения минимума df следует приравнять нулю частные производные по $\xi_1, \dots, \xi_n$ выражения

$$\sum_{l=1}^n f'_{\alpha_l} \xi_l + \lambda \sum_{l=1}^n \xi_l^2.$$

В результате получим

$$f'_{\alpha_l} + 2\lambda \xi_l = 0 \quad (l = 1, \dots, n). \quad (13.3.6)$$

После решения этих уравнений относительно $\xi_1, \dots, \xi_n$ величина λ определится из второго уравнения (13.3.4). В результате получим

$$\xi_l = -k f'_{\alpha_l} \quad (l = 1, \dots, n), \quad \frac{1}{k} = \sqrt{\sum_{j=1}^n f'^2_{\alpha_j}}, \quad (13.3.7)$$

что и требовалось доказать.

Таким образом, метод наискорейшего спуска дает следующую последовательность расчетных комбинаций значений параметров $\alpha_1, \dots, \alpha_n$, начинающуюся в произвольно выбранной исходной точке $(\alpha_1^{(0)}, \dots, \alpha_n^{(0)})$:

$$\alpha_l^{(m+1)} = \alpha_l^{(m)} - s_m f'_{\alpha_l}(\alpha_1^{(m)}, \dots, \alpha_n^{(m)}) \quad (13.3.8)$$

$$(l = 1, \dots, n; m = 0, 1, 2, \dots).$$

Вопрос о выборе длин шагов, т. е. чисел s_m , решается в каждом конкретном случае опытом. При этом можно рекомендовать пользоваться следующими общими соображениями. При слишком малых числах s_m приближение к минимуму будет медленным и объем вычислений будет большим. При слишком больших s_m может случиться так, что функция f при переходе от m -й точки к $m + 1$ -й возрастет (т. е. произойдет «перескок» через минимум). Поэтому числа s_m желательно выбирать возможно большими, но все же достаточно малыми для того, чтобы функция f убывала при переходе из каждой расчетной точки в следующую. Наиболее рациональным является такой выбор чисел s_m , при котором вектор градиента

поворачивается приблизительно на 90° при переходе из каждой расчетной точки в следующую, т. е. при котором скалярное произведение векторов градиента функции f в соседних расчетных точках близко к нулю. В тех случаях, когда вычисление вектора градиента функции f значительно сложнее, чем вычисление самой функции f , П. М. Сотский рекомендует вычислять значения функции f для ряда значений s_m и выбирать каждый раз такое значение s_m , при котором функция f имеет наименьшее значение. Легко сообразить, что при таком способе подбора значений s_m вектор градиента функции f будет поворачиваться приблизительно на 90° при переходе от каждой расчетной точки к следующей. При этом число расчетных точек, в которых придется вычислить градиент функции f , будет близким к минимальному, возможному при данном выборе исходной точки $(\alpha_1^{(0)}, \dots, \alpha_n^{(0)})$.

Пример 13.3.2. Найти оптимальные значения постоянной времени T и коэффициента усиления k аperiodического звена, предназначенного для отфильтровывания от помехи $X(t)$ полезного сигнала, представляющего собой линейную функцию времени $U_1 + U_2 t$ со случайными коэффициентами U_1, U_2 имеющими равные нулю математические ожидания и дисперсии D_1, D_2 соответственно. Независимая от полезного сигнала помеха $X(t)$ является стационарной случайной функцией с равным нулю математическим ожиданием и корреляционной функцией $k_x(\tau) = D e^{-\alpha|\tau|}$. За критерий качества принять среднее по времени значение математического ожидания квадрата установившейся ошибки системы в интервале времени $0 < t < T_p$.

Применяя для вычисления установившейся ошибки воспроизведения полезного сигнала формулу (7.3.12), получим следующее выражение полной установившейся ошибки системы:

$E(t) = c_0(U_1 + U_2 t) + c_1 U_2 + Y_{ш}(t) = (k - 1)(U_1 + U_2 t) - k U_2 T + Y_{ш}(t)$, где $Y_{ш}(t)$ — выходной шум системы, дисперсия которого была найдена в примере 7.4.4:

$$D[Y_{ш}(t)] = \frac{k^2 D}{1 + \alpha T}.$$

Так как математические ожидания случайных величин U_1, U_2 и помехи $X(t)$ равны нулю, то математическое ожидание квадрата ошибки системы равно ее дисперсии и определяется формулой

$$D[E(t)] = (k - 1)^2 D_1 + (t - kt + kT)^2 D_2 + \frac{k^2 D}{1 + \alpha T}. \quad (13.3.9)$$

Среднее значение математического ожидания квадрата ошибки системы в интервале времени $0 < t < T_p$ определяется формулой

$$D_{ср} = f(k, T) = \frac{1}{T_p} \int_0^{T_p} D[E(t)] dt = (k - 1)^2 D_1 + \left[\frac{1}{3} (k - 1)^2 T_p^2 - k(k - 1) T_p T + k^2 T^2 \right] D_2 + \frac{k^2 D}{1 + \alpha T}. \quad (13.3.10)$$

Таким образом, задача сводится к нахождению минимума функции двух переменных $f(k, T)$, определяемой формулой (13.3.10). Дифференцируя

формулу (13.3.10) по k и T , получим

$$\left. \begin{aligned} \frac{\partial f}{\partial k} &= 2(k-1)D_1 + \\ &+ \left[\frac{2}{3}(k-1)T_p^3 - (2k-1)T_pT + 2kT^2 \right] D_2 + \frac{2kD}{1+\alpha T}, \\ \frac{\partial f}{\partial T} &= k[2kT - (k-1)T_p] - \frac{k^2 D \alpha}{(1+\alpha T)^2} \end{aligned} \right\} \quad (13.3.11)$$

Для нахождения оптимальных значений k и T обычным способом следует приравнять нулю производную $\partial f/\partial k$ и из полученного уравнения выразить величину k через T . Подставив это выражение в уравнение $\partial f/\partial T = 0$, получим кубическое уравнение для определения оптимальной постоянной времени T . Это уравнение можно решить обычными способами решения алгебраических уравнений, например графически.

Найдем теперь оптимальные значения k и T методом наискорейшего спуска для случая $D_1 = 10D$, $D_2 = D$, $T_p = 12$ с, $\alpha = 100$ с⁻¹. Задаваясь начальными значениями $k^{(0)} = 0,5$, $T^{(0)} = 0,5$ и применяя изложенный метод, получаем результаты, приведенные в таблице 13.3.1.

ТАБЛИЦА 13.3.1

m	k	T	$f(k, T)$	$\frac{\partial f}{\partial k}$	$\frac{\partial f}{\partial T}$	θ_m
1	0,5	0,5	16,067	-57,73	+3,25	0,008
2	0,96	0,474	0,536	-9,41	+0,085	0,009
3	1,045	0,462	0,114	-0,34	+0,40	0,04
4	1,059	0,446	0,113	+1,32	+0,21	0,01
5	1,046	0,445	0,107	-0,0007	+0,34	0,65
6	1,047	0,225	0,097	+2,60	-0,285	0,0088
7	1,024	0,227	0,065	+0,123	-0,004	0,016
8	1,022	0,227	0,065	-0,003	+0,018	0,165
9	1,0224	0,224	0,064	0	0	—

Полученные после восьми шагов результаты практически совпадают с точным решением в пределах принятой точности вычислений, как читатель сам может убедиться, решив кубическое уравнение, определяющее в данном случае оптимальную постоянную времени.

Метод наискорейшего спуска и его различные разновидности позволяют находить минимумы сложных функций, зависящих от большого числа аргументов. Однако при очень большом числе аргументов минимизация функций требует большого объема вычислений. В этих случаях выгоднее пользоваться методом *случайного поиска*, который требует меньшего объема вычислений при минимизации функций очень большого числа аргументов [58]. Для облегчения и ускорения вычислений при отыскании минимумов функций целесообразно пользоваться специальными вычислительными машинами — так называемыми *оптимизаторами*.

§ 13.4. Общее условие минимума средней квадратической ошибки

Среди всех рассмотренных в § 13.2 критериев качества автоматических систем простейшим с математической точки зрения является критерий минимума средней квадратической ошибки. Кроме того, как было сказано в § 13.2, при нормальном распределении полезных сигналов и помех оптимальная линейная система, найденная по критерию минимума средней квадратической ошибки, оказывается оптимальной и с точки зрения многих других критериев. Наконец, как мы увидим в § 15.1, решение задачи определения оптимальной системы по критерию минимума среднего риска (13.2.6) при любом выборе функции потерь при весьма общих условиях приводится к уравнениям такого же типа, как и уравнения, определяющие оптимальную линейную систему по критерию минимума средней квадратической ошибки. Все эти причины приводят к тому, что методы определения оптимальных линейных систем по критерию минимума средней квадратической ошибки являются основой всей статистической теории оптимальных систем.

На основании изложенного мы в первую очередь будем изучать методы определения оптимальных линейных систем по критерию минимума средней квадратической ошибки, ограничиваясь для простоты случаем одномерных систем, т. е. систем с одним входом и одним выходом. В главе 15 будет изложен общий метод определения оптимальных систем среди всех возможных систем (как линейных, так и нелинейных) по любым критериям, которые можно представить в форме критерия минимума среднего риска (13.2.6) при различных выборах функции потерь.

Выведем сначала общее условие минимума средней квадратической ошибки системы.

Предположим, что в некотором классе систем R требуется найти оптимальную систему, для которой средняя квадратическая ошибка имеет наименьшее возможное значение. Класс систем R может быть любым множеством систем, обладающих определенными свойствами, например множеством всех линейных систем с одним входом и одним выходом или множеством всех возможных систем, для которых интересующая нас случайная функция $Z(t)$ может быть входным сигналом.

Обозначим через A оператор оптимальной системы в классе R , а через B оператор любой системы класса R . Тогда выходной сигнал оптимальной системы будет равен

$$W^*(t) = AZ(t), \quad (13.4.1)$$

а выходной сигнал любой системы класса R будет равен

$$W_i(t) = BZ(t). \quad (13.4.2)$$

Для вывода общего условия, которому должен удовлетворять оператор оптимальной системы, вычислим среднюю квадратическую

ошибку произвольной системы класса R . Выполняя элементарные преобразования и опуская для краткости аргумент t , получим

$$M[(W_1^* - W)^2] = M[(W^* - W + W_1^* - W^*)^2] = M[(W^* - W)^2] + \\ + M[(W_1^* - W^*)^2] + 2M[(W^* - W)(W_1^* - W^*)]. \quad (13.4.3)$$

Предположим, что нам удалось найти в классе R такую систему с оператором A , выходная переменная W^* которой удовлетворяет условию

$$M[(W^* - W)(W_1^* - W^*)] = 0 \quad (13.4.4)$$

для всех систем класса R , т. е. для выходной переменной W_1^* любой системы класса R при входной функции $Z(t)$. Тогда равенство (13.4.3) примет вид

$$M[(W_1^* - W)^2] = M[(W^* - W)^2] + M[(W_1^* - W^*)^2]. \quad (13.4.5)$$

Правая часть этого равенства представляет собой сумму двух неотрицательных слагаемых. Поэтому, отбрасывая второе слагаемое, получим неравенство

$$M[(W_1^* - W)^2] \geq M[(W^* - W)^2]. \quad (13.4.6)$$

Это неравенство показывает, что средняя квадратическая ошибка системы, выходная переменная которой W^* удовлетворяет равенству (13.4.4) для выходных переменных W_1^* всех систем класса R ,

не может быть больше средней квадратической ошибки какой-нибудь другой системы класса R . Иными словами, система, удовлетворяющая условию (13.4.4), всегда является оптимальной. Условие (13.4.4) является, таким образом, достаточным условием минимума средней квадратической ошибки.

Предположим теперь, что класс систем R , среди которых

требуется найти оптимальную систему, обладает таким свойством, что он содержит все системы, полученные путем параллельного соединения любых входящих в него систем с подключенными последовательно к их выходам безынерционными усилителями с постоянными коэффициентами усиления. Это означает, что если мы выберем из класса R любые системы и обозначим через $B_1, B_2, \dots, B_n$ их операторы, то и система, структурная схема которой представлена на рис. 13.4.1, при любых коэффициентах усиления $c_1, c_2, \dots, c_n$ безынерционных усилителей тоже является системой класса R . Иными словами, составляя из любых систем класса R системы типа изображенной на рис. 13.4.1, мы не можем получить систе-

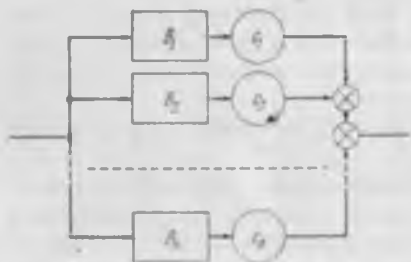


Рис. 13.4.1.

му, которая не входила бы в класс R . Найдем математическое выражение этого свойства. Обозначим выходные переменные систем, из которых составлена система, изображенная на рис. 13.4.1, соответственно через W_1^* , W_2^* , ..., W_n^* . Тогда выходная переменная системы, изображенной на рис. 13.4.1, будет равна

$$W_0^* = c_1 W_1^* + c_2 W_2^* + \dots + c_n W_n^*. \quad (13.4.7)$$

Следовательно, сформулированное выше свойство рассматриваемого класса систем R означает, что любая линейная комбинация выходных переменных систем класса R всегда является выходной переменной некоторой системы того же класса R . Иными словами, множество выходных переменных всех систем класса R таково, что любая линейная комбинация элементов этого множества принадлежит этому множеству. Множество, обладающее таким свойством называется *линейным пространством*. Таким образом, мы рассматриваем сейчас класс систем R , обладающий таким свойством, что множество выходных переменных всех входящих в него систем, при подаче на их входы одной и той же входной функции, представляет собой линейное пространство. Коротко мы будем говорить, что при этом и сам класс систем R является *линейным пространством*.

Пример 13.4.1. Множество всех возможных систем с одним входом и одним выходом является, очевидно, линейным пространством, так как, подключая к выходам любых систем с одним входом и одним выходом безынерционные усилители и соединяя параллельно полученные таким образом пепочки, мы всегда получим систему с одним входом и одним выходом, т. е. систему, принадлежащую рассматриваемому классу.

Пример 13.4.2. Множество всех линейных систем с одним входом и одним выходом также является линейным пространством, так как на основании изложенного в главах 2 и 4 безынерционные усилители являются линейными системами и любые последовательные и параллельные соединения линейных систем с одним входом и одним выходом всегда дают линейную систему с одним входом и одним выходом.

Пример 13.4.3. Чтобы убедиться в том, что существуют классы систем, не являющиеся линейными пространствами, рассмотрим множество всех аperiodических звеньев с разными постоянными времени и с одинаковыми коэффициентами усиления, равными единице. В этом классе систем мы искали оптимальный фильтр в примере 13.3.1. Очевидно, что параллельное соединение двух аperiodических звеньев с разными постоянными времени T_1 и T_2 не является аperiodическим звеном, так как передаточная функция такого параллельного соединения на основании (4.6.6) выражается формулой

$$\Phi(s) = \frac{1}{1+T_1 s} + \frac{1}{1+T_2 s} = 2 \frac{1 + \frac{T_1 + T_2}{2}s}{(1+T_1 s)(1+T_2 s)}$$

и при $T_1 \neq T_2$ не является передаточной функцией аperiodического звена. Таким образом, класс аperiodических звеньев не является линейным пространством.

Итак, предположим, что класс систем R , среди которых ищется оптимальная система, представляет собой линейное пространство. В этом случае разность выходных переменных $W_i^* - W^*$ оптимальной системы и любой другой системы класса R также является выходной переменной некоторой системы класса R . Наоборот, если взять выходную переменную любой системы класса R , то, сложив ее с выходной переменной оптимальной системы W^* , мы снова получим выходную переменную W_i^* некоторой системы класса R . Отсюда следует, что выходная переменная любой системы класса R может быть представлена в виде разности выходной переменной некоторой системы класса R и выходной переменной оптимальной системы. Таким образом, разность $W_i^* - W^*$ в равенстве (13.4.4) можно рассматривать как выходную переменную произвольной системы класса R , которую можно также обозначить через W_i^* . Поэтому для случая, когда класс систем R , в котором ищется оптимальная система, представляет собой линейное пространство,

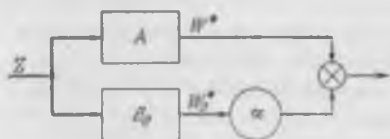


Рис. 13.4.2.

достаточное условие минимума средней квадратической ошибки (13.4.4) может быть переписано в виде

$$M[(W^* - W)W_i^*] = 0, \quad (13.4.8)$$

где W_i^* — выходная переменная любой системы класса R .

Докажем теперь, что условие (13.4.8) является не только доста-

точным, но и необходимым условием минимума средней квадратической ошибки, если класс систем, в котором требуется найти оптимальную систему, представляет собой линейное пространство. Для доказательства достаточно показать, что при невыполнении условия (13.4.8) W^* не может быть выходной переменной оптимальной системы. Предположим, что в классе R существует такая система с оператором B_0 , для выходной переменной W_0^* которой условие (13.4.8) не выполнено. Для определенности предположим, например, что

$$M[(W^* - W)W_0^*] > 0. \quad (13.4.9)$$

Рассмотрим систему, представляющую собой параллельное соединение предполагаемой оптимальной системы, выходная переменная которой равна W^* , и системы с выходной переменной W_0^* с подключенным последовательно к ее выходу безынерционным усилителем с коэффициентом усиления α (рис. 13.4.2). Эта система принадлежит классу систем R , так как класс R по предположению является линейным пространством. Покажем, что коэффициент усиления α всегда можно выбрать так, чтобы рассматриваемая система была лучше предполагаемой оптимальной в случае, когда

имеет место неравенство (13.4.9) или противоположное неравенство. Этим и будет доказано, что оптимальная система не может не удовлетворять условию (13.4.8), т. е. что условие (13.4.8) необходимо. Выходная переменная системы, изображенной на рис. 13.4.2, выражается формулой

$$W_i^* = W^* + \alpha W_0^*,$$

откуда

$$W_i^* - W^* = \alpha W_0^*$$

и, следовательно,

$$M[(W_i^* - W^*)^2] = \alpha^2 M[(W_0^*)^2],$$

$$M[(W^* - W)(W_i^* - W^*)] = \alpha M[(W^* - W)W_0^*].$$

Подставляя эти выражения в формулу (13.4.3), будем иметь

$$M[(W_i^* - W)^2] = M[(W^* - W)^2] + \alpha \{2M[(W^* - W)W_0^*] + \alpha M[(W_0^*)^2]\}. \quad (13.4.10)$$

Если взять достаточно малое по абсолютной величине отрицательное значение α , то при выполнении неравенства (13.4.9) выражение в фигурных скобках будет положительным и, следовательно, второе слагаемое в правой части равенства (13.4.10) будет отрицательным. Отбрасывая это слагаемое, мы получим

$$M[(W_i^* - W)^2] < M[(W^* - W)^2]. \quad (13.4.11)$$

Это неравенство показывает, что при любом достаточно малом по абсолютной величине отрицательном значении коэффициента усиления α система, изображенная на рис. 13.4.2, будет лучше предполагаемой оптимальной, если имеет место неравенство (13.4.9).

Если

$$M[(W^* - W)W_0^*] < 0,$$

то, выбирая любое достаточно малое положительное значение α , мы снова получим неравенство (13.4.11), и система, изображенная на рис. 13.4.2, будет лучше предполагаемой оптимальной. Следовательно, оптимальная система не может не удовлетворять условию (13.4.8), т. е. это условие необходимо.

Итак, мы доказали, что если класс систем R , в котором ищется оптимальная система, является линейным пространством, то необходимым и достаточным условием оптимальности системы с выходной переменной W^* является выполнение условия (13.4.8) для выходных переменных W_i^* всех систем класса R . В этом случае оптимальная система обязательно удовлетворяет условию (13.4.8), если она существует. Если же класс систем R не является линей-

ным пространством, то условие (13.4.4) может не быть необходимым и оптимальная система может не удовлетворять этому условию. В последнем случае в классе R не существует такой системы, которая удовлетворяла бы условию (13.4.4), так как вследствие достаточности этого условия система, удовлетворяющая ему, обязательно является оптимальной.

Чтобы получить окончательную форму уравнения, определяющего оператор оптимальной системы по критерию минимума средней квадратической ошибки, подставим в (13.4.8) выражения (13.4.1) и (13.4.2) выходных переменных оптимальной системы и произвольной системы того класса, в котором ищется оптимальная система. Тогда получим

$$M \{ [AZ(t) - W(t)] BZ(t) \} = 0 \quad (B \in R). \quad (13.4.12)$$

Здесь запись $B \in R$ означает, что равенство должно быть выполнено для операторов B всех систем класса R .

Уравнение (13.4.12) является общим необходимым и достаточным условием минимума средней квадратической ошибки, определяющим оператор оптимальной системы A в любом классе систем R , представляющем собой линейное пространство. Это уравнение может служить не только для определения оптимальных линейных систем, но и для определения оптимальных систем в различных классах нелинейных систем.

Уравнение (13.4.12) определяет оптимальную систему, для которой мгновенное значение средней квадратической ошибки для каждого момента времени t имеет наименьшее возможное значение. Таким образом, уравнение (13.4.12) минимизирует среднюю квадратическую ошибку для любого момента времени t . В частном случае, когда выходные переменные всех систем класса R и требуемый выходной сигнал $W(t)$ являются эргодическими стационарными случайными функциями времени, средняя квадратическая ошибка любой системы класса R постоянна и с вероятностью единица совпадает с соответствующим средним по времени, вычисленным для любой реализации входного сигнала $Z(t)$ и требуемого выходного сигнала $W(t)$:

$$\eta = \lim_{T \rightarrow \infty} \frac{1}{2T} \int_{-T}^T E^2(t) dt. \quad (13.4.13)$$

В этом случае уравнение (13.4.12) минимизирует среднее значение квадрата ошибки по времени. Однако уравнение (13.4.12), как мы видели, имеет значительно более общий характер и определяет оптимальные системы и в тех случаях, когда выходные переменные систем рассматриваемого класса не являются эргодическими стационарными случайными функциями.

§ 13.5. Уравнения, определяющие оптимальную линейную систему

Выведенное в предыдущем параграфе общее уравнение (13.4.12) определяет оптимальную систему любого заданного класса. Из него можно вывести конкретные уравнения, определяющие оптимальные системы различных конкретных классов. В частности, из уравнения (13.4.12) можно вывести уравнение, определяющее оптимальную систему в классе линейных систем.

Обозначим через $g(t, \tau)$ весовую функцию оптимальной линейной системы, а через $h(t, \tau)$ весовую функцию произвольной линейной системы. Так как нас интересуют только физически возможные линейные системы, то весовые функции $g(t, \tau)$ и $h(t, \tau)$ равны нулю при $\tau > t$ и в соответствии с общей формулой (2.2.5)

$$\left. \begin{aligned} AZ(t) &= \int_{t_0}^t g(t, \tau) Z(\tau) d\tau, \\ BZ(t) &= \int_{t_0}^t h(t, \sigma) Z(\sigma) d\sigma. \end{aligned} \right\} \quad (13.5.1)$$

Эти формулы выражают выходные переменные рассматриваемых систем в том случае, когда класс допустимых линейных систем R содержит системы, использующие все значения входного сигнала $Z(t)$, начиная с момента начала работы системы t_0 . Однако во многих случаях практики значения входного сигнала в прошедшие моменты времени, далеко отстоящие от данного момента t , не несут никакой информации о полезном сигнале, соответствующем данному моменту t . Более того, эти значения входного сигнала могут содержать неправильную информацию о полезном сигнале. Так, например, следящая система локатора, предназначенного для измерения координат летящих самолетов, закончив сопровождение одного самолета, должна переключиться на другой самолет. В этом случае значения входного сигнала, полученные от первого самолета, будут только мешать правильному измерению координат второго самолета. Чтобы этого не было, следящая система должна «забыть» входной сигнал, полученный от первого самолета, как можно скорее после начала сопровождения второго самолета. Для обеспечения этого достаточно ограничить «память» системы, т. е. потребовать, чтобы ее весовая функция обращалась в нуль для всех значений второго аргумента, достаточно далеких от первого аргумента, т. е. от текущего момента t . Ограничив класс линейных систем R , в котором ищется оптимальная система, только такими системами, память которых имеет длительность T , т. е.

весовые функции которых равны нулю при всех значениях второго аргумента τ , меньших, чем $t - T$, мы можем переписать формулы (13.5.1) в виде

$$AZ(t) = \int_{t-T}^t g(t, \tau) Z(\tau) d\tau, \quad BZ(t) = \int_{t-T}^t h(t, \sigma) Z(\sigma) d\sigma. \quad (13.5.2)$$

В этом случае все допустимые линейные системы, т. е. все системы рассматриваемого класса R , в котором ищется оптимальная система, используют значения входного сигнала $Z(t)$ только в интервале времени $t - T \leq \tau \leq t$ длительностью T , непосредственно предшествующем данному моменту t .

Вопрос о выборе длительности памяти системы должен решаться в зависимости от назначения системы с учетом двух противоречивых требований. С одной стороны, чем больше T , тем больше информации о полезном сигнале содержит входной сигнал $Z(t)$. С другой стороны, чем меньше T , тем скорее система решит задачу после переключения на новую реализацию входного сигнала. В некоторых случаях может оказаться целесообразным сделать память системы T функцией времени $T(t)$.

С математической точки зрения формулы (13.5.1) можно рассматривать как частный случай формул (13.5.2) при $T = t - t_0$. Поэтому, приняв формулы (13.5.2), мы охватываем оба практически интересных случая: когда интервал наблюдения T совпадает с полным временем работы системы и когда интервал наблюдения T совпадает с памятью системы. В некоторых задачах оказывается допустимым считать интервал наблюдения T бесконечным. Это допущение является, конечно, идеализацией, так как в действительности любая система работает лишь конечное время, хотя и может иметь при этом бесконечную память. Однако эта идеализация, как мы увидим дальше, существенно упрощает нахождение оптимальной линейной системы.

Для вывода уравнения, определяющего весовую функцию $g(t, \tau)$ оптимальной линейной системы, подставим выражения (13.5.2) в общее уравнение (13.4.12). Тогда получим

$$M \left[\left\{ \int_{t-T}^t g(t, \tau) Z(\tau) d\tau - W(t) \right\} \int_{t-T}^t h(t, \sigma) Z(\sigma) d\sigma \right] = 0, \quad (13.5.3)$$

и это равенство должно выполняться для любой весовой функции $h(t, \sigma)$. Внося выражение в фигурных скобках под знак второго интеграла, выполняя почленное умножение этого выражения под знаком второго интеграла на $Z(\sigma)$ и меняя местами операции мате-

математического ожидания и интегрирования, получим

$$\begin{aligned}
 M \left[\left\{ \int_{t-T}^t g(t, \tau) Z(\tau) d\tau - W(t) \right\} \int_{t-T}^t h(t, \sigma) Z(\sigma) d\sigma \right] = \\
 = M \left[\int_{t-T}^t h(t, \sigma) \left\{ Z(\sigma) \int_{t-T}^t g(t, \tau) Z(\tau) d\tau - W(t) Z(\sigma) \right\} d\sigma \right] = \\
 = \int_{t-T}^t h(t, \sigma) M \left[\int_{t-T}^t g(t, \tau) Z(\tau) Z(\sigma) d\tau - W(t) Z(\sigma) \right] d\sigma = \\
 = \int_{t-T}^t h(t, \sigma) \left\{ \int_{t-T}^t g(t, \tau) M[Z(\tau) Z(\sigma)] d\tau - M[W(t) Z(\sigma)] \right\} d\sigma.
 \end{aligned}$$

Подставляя это выражение в (13.5.3) и принимая во внимание, что математические ожидания представляют собой соответственно начальный момент второго порядка случайной функции $Z(t)$ и совместный момент второго порядка случайных функций $W(t)$ и $Z(t)$:

$$M[Z(\tau) Z(\sigma)] = \Gamma_z(\tau, \sigma), \quad M[W(t) Z(\sigma)] = \Gamma_{wz}(t, \sigma),$$

приведем уравнение (13.5.3) к виду

$$\int_{t-T}^t h(t, \sigma) \left\{ \int_{t-T}^t \Gamma_z(\tau, \sigma) g(t, \tau) d\tau - \Gamma_{wz}(t, \sigma) \right\} d\sigma = 0. \quad (13.5.4)$$

Это уравнение может удовлетворяться для любой функции $h(t, \sigma)$ только в том случае, если выражение в фигурных скобках тождественно равно нулю в интервале интегрирования $t - T \leq \sigma \leq t$:

$$\int_{t-T}^t \Gamma_z(\tau, \sigma) g(t, \tau) d\tau - \Gamma_{wz}(t, \sigma) = 0 \quad (t - T \leq \sigma \leq t). \quad (13.5.5)$$

Действительно, если это выражение не равно нулю тождественно относительно σ , то, выбирая функцию $h(t, \sigma)$ положительной при тех значениях σ , при которых выражение в фигурных скобках (13.5.4) положительно, и отрицательной при тех значениях σ , при которых выражение в фигурных скобках в (13.5.4) отрицательно, мы получим в левой части равенства (13.5.4) существенно положительную величину. Это убеждает нас в невозможности выполнения равенства (13.5.4) для любой функции $h(t, \sigma)$, если выражение в фигурных скобках в (13.5.4) не равно нулю при всех

значениях σ в интервале $t - T \leq \sigma \leq t$ *). Таким образом, уравнение (13.5.5) является необходимым условием оптимальности линейной системы с весовой функцией $g(t, \sigma)$.

Однако условие (13.5.5) в общем случае недостаточно. Действительно, оптимальные операции над входным сигналом могут включать операции дифференцирования. Предположим, что входной сигнал $Z(t)$ можно дифференцировать p раз. Это означает, что весовые функции $g(t, \tau)$ и $h(t, \sigma)$ могут содержать δ -функции и их производные до порядка p включительно. Полагая в (13.5.4) $h(t, \sigma) = \delta^{(k)}(t - a - \sigma)$, получим на основании (2.1.10)

$$\int_{t-T}^t \left[\frac{\partial^k \Gamma_z(\tau, \sigma)}{\partial \sigma^k} \right]_{\sigma=t-a} g(t, \tau) d\tau - \left[\frac{\partial^k \Gamma_{wz}(t, \sigma)}{\partial \sigma^k} \right]_{\sigma=t-a} = 0. \quad (13.5.6)$$

Здесь величина a может иметь любое значение в пределах $0 \leq a \leq T$. Однако тождественное выполнение равенства (13.5.5) внутри интервала $t - T < \sigma < t$ влечет за собой автоматическое выполнение равенства (13.5.6) при всех k внутри интервала $t - T < \sigma < t$, т. е. при всех a в пределах $0 < a < T$. Следовательно, равенство (13.5.6) может нарушаться, не выполняясь при выполнении равенства (13.5.5) только на концах интервала $[t - T, t]$, т. е. при $a = 0$ и при $a = T$. Таким образом, остается потребовать его выполнения на концах интервала интегрирования, т. е. при $a = 0$ и при $a = T$ для всех допустимых значений k , и мы приходим к выводу, что для оптимальности системы необходимо тождественное относительно σ выполнение равенства (13.5.5) в замкнутом интервале $t - T \leq \sigma \leq t$ и, кроме того, выполнение граничных условий

$$\left. \begin{aligned} \int_{t-T}^t \left[\frac{\partial^k \Gamma_z(\tau, \sigma)}{\partial \sigma^k} \right]_{\sigma=t-T} g(t, \tau) d\tau - \left[\frac{\partial^k \Gamma_{wz}(t, \sigma)}{\partial \sigma^k} \right]_{\sigma=t-T} &= 0, \\ \int_{t-T}^t \left[\frac{\partial^k \Gamma_z(\tau, \sigma)}{\partial \sigma^k} \right]_{\sigma=t} g(t, \tau) d\tau - \left[\frac{\partial^k \Gamma_{wz}(t, \sigma)}{\partial \sigma^k} \right]_{\sigma=t} &= 0. \end{aligned} \right\} \quad (13.5.7)$$

$(k = 1, \dots, p)$

*) Обратим внимание на то, что равенство (13.5.5) должно выполняться и на концах интервала интегрирования, т. е. в замкнутом интервале $[t - T, t]$. Действительно, допустимые линейные системы могут выполнять над сигналом операцию чистого (безынерционного) усиления и запаздывания на время памяти системы T . Этому соответствует наличие слагаемых вида $\delta(t - \tau)$ и $\delta(t - T - \tau)$ в выражениях весовых функций $g(t, \tau)$ и $h(t, \tau)$. Если бы равенство (13.5.5) не выполнялось при $\sigma = t$ или $\sigma = t - T$, то, включив в $h(t, \sigma)$ допустимое слагаемое, пропорциональное $\delta(t - \sigma)$ или соответственно $\delta(t - T - \sigma)$, мы бы получили отличную от нуля левую часть формулы (13.5.4).

Все эти условия в совокупности достаточны для того, чтобы функция $g(t, \tau)$ была весовой функцией оптимальной системы, так как при их выполнении равенство (13.5.4) будет удовлетворено при любой допустимой функции $h(t, \sigma)$, как непрерывной, так и содержащей импульсные функции до соответствующего порядка.

Заметим теперь, что так как весовая функция $g(t, \tau)$ также может содержать допустимые δ -функции и их производные, то уравнения (13.5.7) фактически содержат смешанные производные функции $\Gamma_z(\tau, \sigma)$ до порядка $\frac{\partial^{2p}\Gamma_z(\tau, \sigma)}{\partial \tau^p \partial \sigma^p}$ включительно. Отсюда следует, что уравнения (13.5.7) будут все выполняться в том и только в том случае, если производная $\frac{\partial^{2p}\Gamma_z(\tau, \sigma)}{\partial \tau^p \partial \sigma^p}$ непрерывна всюду,

в том числе и при $\tau = \sigma$. Это условие может служить для определения числа p допустимых дифференцирований входного сигнала. А именно наивысший порядок допустимого дифференцирования p есть наибольшее целое число, при котором производная $\frac{\partial^{2p}\Gamma_z(\tau, \sigma)}{\partial \tau^p \partial \sigma^p}$ непрерывна при $\tau = \sigma$. Если определить число p таким образом, то все равенства (13.5.7) будут следствиями выполнения равенства (13.5.5) в замкнутом интервале $t - T \leq \sigma \leq t$.

Таким образом, мы получаем другую форму условий минимума средней квадратической ошибки. А именно для того чтобы функция $g(t, \tau)$ была весовой функцией оптимальной линейной системы, необходимо и достаточно, чтобы она удовлетворяла уравнению (13.5.5) в замкнутом интервале $t - T \leq \sigma \leq t$ и содержала δ -функции и их производные не выше p -го порядка, если p — наибольшее целое число, при котором производная $\frac{\partial^{2p}\Gamma_z(\tau, \sigma)}{\partial \tau^p \partial \sigma^p}$ непрерывна. На недостаточность условия (13.5.5) без дополнительного условия ограничения порядка производных δ -функций, входящих в выражение весовой функции оптимальной системы, впервые обратил внимание В. М. Семенов.

Заметим, что если сделать замену переменных $\xi = t - \tau$, $\eta = t - \sigma$, то мы приведем интервал интегрирования и интервал изменения η к замкнутому интервалу $[0, T]$. В этом случае величина t будет играть в уравнении (13.5.5) роль несущественного параметра. Вследствие этого величину t иногда опускают при записи уравнения (13.5.5) в переменных ξ, η .

В задачах практики полезный сигнал $S(t)$, содержащийся во входном сигнале $Z(t)$, часто представляет собой сумму линейной комбинации известных функций (регулярной части) и нерегулярной случайной части:

$$S(t) = \sum_{r=1}^N U_r \varphi_r(t) + X_1(t), \quad (13.5.8)$$

где $\varphi_1(t), \dots, \varphi_N(t)$ — известные функции времени, а $X_1(t)$ — случайная функция времени. Коэффициенты $U_1, \dots, U_N$ могут быть случайными величинами или просто неизвестными величинами, которые могут иметь различные значения. В этом случае, предполагая, что преобразование L , которое проектируемая система должна выполнять над полезным сигналом, линейно, получим на основании формулы (13.1.4) следующее выражение требуемого выходного сигнала проектируемой системы:

$$W(t) = \sum_{r=1}^N U_r L\varphi_r(t) + LX_1(t). \quad (13.5.9)$$

Полагая для краткости

$$\psi_r(t) = L\varphi_r(t) \quad (r = 1, \dots, N), \quad Y(t) = LX_1(t), \quad (13.5.10)$$

можем переписать формулу (13.5.9) в виде

$$W(t) = \sum_{r=1}^N U_r \psi_r(t) + Y(t). \quad (13.5.11)$$

Предположение о линейности оператора L означает, что идеальная система, которая теоретически абсолютно точно решает поставленную задачу при полном отсутствии помех, линейна, хотя, может быть, и физически невозможна (например, идеальный экстраполлятор). Поэтому, обозначая ее весовую функцию через $l(t, \tau)$, мы можем переписать формулы (13.5.10) в виде

$$\left. \begin{aligned} \psi_r(t) = L\varphi_r(t) &= \int_{-\infty}^{\infty} l(t, \tau) \varphi_r(\tau) d\tau \quad (r = 1, 2, \dots, N), \\ Y(t) = LX_1(t) &= \int_{-\infty}^{\infty} l(t, \tau) X_1(\tau) d\tau. \end{aligned} \right\} \quad (13.5.12)$$

Подставляя выражение (13.5.8) в (13.1.6) и объединяя для краткости записи нерегулярную часть полезного сигнала $X_1(t)$ с помехой $N(t)$:

$$X(t) = X_1(t) + N(t), \quad (13.5.13)$$

получим следующее выражение для входного сигнала искомой оптимальной системы:

$$Z(t) = \sum_{r=1}^N U_r \psi_r(t) + X(t). \quad (13.5.14)$$

Таким образом, входной сигнал $Z(t)$ и требуемый выходной сигнал $W(t)$ во многих задачах практически имеют форму (13.5.14) и (13.5.11), где $\varphi_1, \dots, \varphi_N, \psi_1, \dots, \psi_N$ — известные функции, а $X(t)$ и $Y(t)$ — некоторые случайные функции.

Мы будем считать $U_1, \dots, U_N$ случайными величинами, не коррелированными со случайными функциями $X(t)$ и $Y(t)$. Случай, когда $U_1, \dots, U_N$ (или некоторые из них) являются неизвестными неслучайными величинами, которые могут принимать любые значения, как показал В. М. Семенов, можно рассматривать как частный случай, когда дисперсии случайных величин $U_1, \dots, U_N$ (или некоторых из них) равны бесконечности [62]. Предположение о некоррелированности случайных величин $U_1, \dots, U_N$ со случайными функциями $X(t), Y(t)$ отражает тот факт, что в задачах практики регулярная часть полезного сигнала, представляющая собой линейную комбинацию известных функций, его нерегулярная часть и помеха обычно имеют различную физическую природу и порождаются независимыми друг от друга источниками.

Для простоты выкладок мы предположим еще, что математические ожидания случайных функций $X(t)$ и $Y(t)$ тождественно равны нулю. Это предположение ни в какой мере не ограничивает общности, так как в случае отличных от нуля математических ожиданий $m_x(t)$ и $m_y(t)$ их можно рассматривать как дополнительные регулярные части входного сигнала $Z(t)$ и требуемого выходного сигнала $W(t)$. Действительно, написав формулы (13.5.14) и (13.5.11) в виде

$$\left. \begin{aligned} Z(t) &= \sum_{r=1}^N U_r \varphi_r(t) + m_x(t) + X^0(t), \\ W(t) &= \sum_{r=1}^N U_r \psi_r(t) + m_y(t) + Y^0(t) \end{aligned} \right\} \quad (13.5.15)$$

и полагая

$$\varphi_{N+1}(t) = m_x(t), \quad \psi_{N+1}(t) = m_y(t), \quad U_{N+1} = 1,$$

приведем формулы (13.5.15) к виду (13.5.14) и (13.5.11) с увеличенным на единицу числом слагаемых в суммах и тождественно равными нулю математическими ожиданиями случайных функций $X(t)$ и $Y(t)$. При этом последний коэффициент в сумме U_{N+1} будет представлять собой неслучайную величину, равную единице, которую всегда можно рассматривать как случайную величину с нулевой дисперсией и математическим ожиданием, равным единице. Таким образом, случай отличных от тождественного нуля математических ожиданий нерегулярных случайных частей входного и требуемого выходного сигналов приводится к случаю нулевых математических ожиданий за счет увеличения на единицу числа слагаемых в регулярных частях входного и требуемого выходного сигналов.

Пример 13.5.1. Рассмотрим следующую систему примера 13.1.1. Для определенности предположим, что эта следящая система является следящей системой радиолокатора, предназначенного для измерения координат

самолетов в декартовой системе координат. При этом для простоты будем рассматривать только одну координату. С достаточной для практики точностью можно считать истинную координату самолета $S_1(t)$ полиномом второй степени относительно времени:

$$S_1(t) = U_1 + U_2 t + U_3 t^2.$$

При этом начальное положение U_1 и начальная скорость U_2 самолета, попавшего в поле зрения радиолокатора, могут изменяться в весьма широких пределах и являются заранее неизвестными величинами. Поэтому их можно считать случайными величинами с бесконечными дисперсиями. Что же касается ускорения $2U_3$, то оно может изменяться лишь в сравнительно узких пределах. Поэтому целесообразно считать его случайной величиной с конечной дисперсией, которую можно определить на основании статистического анализа технических данных самолетов. Функция $S_1(t)$ является регулярной частью полезного сигнала на входе следящей системы.

Помеха $N_2(t)$, действующая на исполнительное устройство следящей системы, в данном случае может быть моментом аэродинамических сил, действующих на антенну радиолокатора. В случае турбулентной атмосферы момент аэродинамических сил является случайной функцией времени. Мы видели в примере 13.1.1, что в данном случае преобразованная исполнительным устройством помеха $N_2(t)$, т. е. выходная переменная исполнительного устройства, соответствующая действию на его входе помехи $N_2(t)$, взятая с обратным знаком, также должна рассматриваться как составная часть полезного сигнала на входе следящей системы:

$$X_1(t) = -B_N N_2(t),$$

где B_N — оператор исполнительного устройства. На основании изложенного в §§ 3.13 и 3.14 передаточная функция исполнительного устройства (от управляющего входного сигнала до угла поворота вала) часто выражается формулой

$$\Phi_N(s) = \frac{k}{s(1 + T_N s)}.$$

При этом зависимость между помехой $N_2(t)$ и соответствующей частью выходного сигнала исполнительного устройства $X_1(t)$ будет равноценна дифференциальному уравнению второго порядка

$$T_N X_1''(t) + X_1'(t) = -k N_2(t)$$

(см. § 2.5). Зная это уравнение или передаточную функцию, можно методами главы 7 найти корреляционную функцию нерегулярной части полезного сигнала $X_1(t)$ по данной корреляционной функции случайной функции $N_2(t)$.

Полный полезный сигнал $S(t)$ на входе следящей системы будет

$$S(t) = S_1(t) + X_1(t) = U_1 + U_2 t + U_3 t^2 + X_1(t).$$

Таким образом, задача проектирования оптимальной следящей системы радиолокатора приводится к рассмотренной выше общей задаче, когда входной сигнал проектируемой системы состоит из регулярной части полезного сигнала $S_1(t)$, представляющей собой в данном случае квадратный трехчлен со случайными коэффициентами, нерегулярной части полезного сигнала $X_1(t)$, представляющей собой в данном случае преобразованный исполнительным устройством случайный аэродинамический момент, взятый с обратным знаком, и помехи $N_1(t)$. Функции $\varphi_r(t)$, $\varphi_r(t)$ и случайная функция $Y(t)$ в данном случае определяются формулами

$$\varphi_1(t) = \psi_1(t) = 1, \quad \varphi_2(t) = \psi_2(t) = t,$$

$$\varphi_3(t) = \psi_3(t) = t^2, \quad Y(t) = X_1(t).$$

Если система должна экстраполировать координату самолета (т. е. определять с наилучшей возможной точностью упрежденную координату), то при тех же функциях $\varphi_1(t)$, $\varphi_2(t)$ и $\varphi_3(t)$ функции $\psi_1(t)$, $\psi_2(t)$, $\psi_3(t)$ и случайная функция $Y(t)$ выражаются формулами

$$\psi_1(t) = 1, \quad \psi_2(t) = t + \Delta, \quad \psi_3(t) = (t + \Delta)^2, \quad Y(t) = X_1(t + \Delta).$$

Если система предназначена для измерения скорости самолета, то при тех же функциях $\varphi_1(t)$, $\varphi_2(t)$ и $\varphi_3(t)$

$$\psi_1(t) = 0, \quad \psi_2(t) = 1, \quad \psi_3(t) = 2t, \quad Y(t) = X'_1(t).$$

Для вывода уравнений, определяющих оптимальную линейную систему в случае, когда входной и требуемый выходной сигналы выражаются формулами (13.5.11) и (13.5.14) вычислим начальные моменты второго порядка $\Gamma_z(\tau, \sigma)$ и $\Gamma_{wz}(t, \sigma)$. Пользуясь свойствами математических ожиданий и учитывая, что по предположению $m_x(t) = m_y(t) = 0$, получим

$$\begin{aligned} \Gamma_z(\tau, \sigma) &= M[Z(\tau)Z(\sigma)] = \\ &= \sum_{p, q=1}^N \varphi_p(\tau) \varphi_q(\sigma) M[U_p U_q] + M[X(\tau)X(\sigma)] + \\ &+ \sum_{r=1}^N \{\varphi_r(\tau) M[U_r X(\sigma)] + \varphi_r(\sigma) M[U_r X(\tau)]\}. \end{aligned} \quad (13.5.16)$$

По математическое ожидание произведения некоррелированных случайных величин равно произведению их математических ожиданий. Следовательно, $M[U_r X(\sigma)] = M[U_r] M[X(\sigma)] = 0$, так как $M[X(\sigma)] = m_x(\sigma) \equiv 0$. Точно так же $M[U_r X(\tau)] = M[U_r] M[X(\tau)] = 0$. Учитывая эти равенства и вводя для начальных моментов второго порядка случайных величин $U_1, \dots, U_N$ обозначение

$$\gamma_{pq} = M[U_p U_q] \quad (p, q = 1, \dots, N), \quad (13.5.17)$$

можем переписать формулу (13.5.16) в виде

$$\Gamma_z(\tau, \sigma) = M[Z(\tau)Z(\sigma)] = \sum_{p, q=1}^N \gamma_{pq} \varphi_p(\tau) \varphi_q(\sigma) + K_x(\tau, \sigma). \quad (13.5.18)$$

Таким же образом найдем смешанный момент второго порядка требуемого выходного сигнала $W(t)$ и входного сигнала $Z(t)$:

$$\Gamma_{wz}(t, \sigma) = M[W(t)Z(\sigma)] = \sum_{p, q=1}^N \gamma_{pq} \psi_p(t) \varphi_q(\sigma) + K_{yx}(t, \sigma), \quad (13.5.19)$$

где $K_{yx}(t, \sigma)$ — взаимная корреляционная функция случайных функций $Y(t)$ и $X(t)$.

Подставляя выражения (13.5.18) и (13.5.19) в исходное уравнение (13.5.5), приведем его к виду

$$\begin{aligned} & \int_{t-T}^t K_x(\tau, \sigma) g(t, \tau) d\tau = \\ & = K_{yx}(t, \sigma) + \sum_{p, q=1}^N \gamma_{pq} \left\{ \psi_p(t) - \int_{t-T}^t g(t, \tau) \varphi_p(\tau) d\tau \right\} \varphi_q(\sigma) \quad (13.5.20) \\ & \quad (t-T \leq \sigma \leq t). \end{aligned}$$

Неравенства в скобках показывают, что это уравнение должно удовлетворяться для всех значений σ в замкнутом интервале $[t-T, t]$.

Уравнение (13.5.20) для решения неудобно, так как неизвестная весовая функция входит под знаками $N+1$ интегралов. Поэтому целесообразно заменить его более простой системой уравнений. Для этого введем новые неизвестные

$$\lambda_q = \sum_{p=1}^N \gamma_{pq} \left[\psi_p(t) - \int_{t-T}^t g(t, \tau) \varphi_p(\tau) d\tau \right] \quad (q=1, \dots, N). \quad (13.5.21)$$

Заметим, что эти новые неизвестные являются функциями времени $\lambda_q = \lambda_q(t)$ ($q=1, \dots, N$), но для краткости мы будем аргумент t опускать. На основании формулы (13.5.21) уравнение (13.5.20) принимает вид

$$\begin{aligned} & \int_{t-T}^t K_x(\tau, \sigma) g(t, \tau) d\tau = \\ & = K_{yx}(t, \sigma) + \sum_{q=1}^N \lambda_q \varphi_q(\sigma) \quad (t-T \leq \sigma \leq t). \quad (13.5.22) \end{aligned}$$

В это уравнение неизвестная весовая функция входит только под знаком одного интеграла, вследствие чего это уравнение проще, чем (13.5.20). Но зато в нем появились N дополнительных неизвестных $\lambda_1, \dots, \lambda_N$. Уравнения (13.5.21) и (13.5.22) вместе образуют систему $N+1$ уравнений с $N+1$ неизвестными $g(t, \tau), \lambda_1(t), \dots, \lambda_N(t)$.

Уравнения (13.5.21) удобно преобразовать, решив их относительно выражений в квадратных скобках. Полагая

$$\eta_p = \psi_p(t) - \int_{t-T}^t g(t, \tau) \varphi_p(\tau) d\tau \quad (p=1, \dots, N), \quad (13.5.23)$$

перепишем уравнения (13.5.21) в виде

$$\sum_{p=1}^N \gamma_{pq} \eta_p = \lambda_q \quad (q=1, \dots, N). \quad (13.5.24)$$

Это — система линейных алгебраических уравнений относительно неизвестных $\eta_1, \dots, \eta_N$. Решая ее при помощи известного правила Крамера, получим

$$\eta_p = \frac{(p)}{\Gamma} \quad (p = 1, \dots, N). \quad (13.5.25)$$

$$(p) = \begin{vmatrix} \gamma_{11} & \gamma_{12} & \dots & \lambda_1 & \dots & \gamma_{1N} \\ \gamma_{21} & \gamma_{22} & \dots & \lambda_2 & \dots & \gamma_{2N} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \gamma_{N1} & \gamma_{N2} & \dots & \lambda_N & \dots & \gamma_{NN} \end{vmatrix}$$

$$\Gamma = \begin{vmatrix} \gamma_{11} & \gamma_{12} & \dots & \dots & \dots & \gamma_{1N} \\ \gamma_{21} & \gamma_{22} & \dots & \dots & \dots & \gamma_{2N} \\ \dots & \dots & \dots & \dots & \dots & \dots \\ \gamma_{N1} & \gamma_{N2} & \dots & \dots & \dots & \gamma_{NN} \end{vmatrix}$$

Раскладывая определитель в числителе по элементам p -го столбца и вводя обозначение

$$c_{pq} = \frac{\Gamma_{pq}}{\Gamma} \quad (p, q = 1, \dots, N), \quad (13.5.26)$$

где Γ — определитель, стоящий в знаменателе формулы (13.5.25), а Γ_{pq} — алгебраическое дополнение элемента γ_{pq} в этом определителе, приведем формулу (13.5.25) к виду

$$\eta_p = \sum_{q=1}^N c_{pq} \lambda_q \quad (p = 1, \dots, N). \quad (13.5.27)$$

Наконец, заменяя здесь вспомогательную величину η_p ее выражением (13.5.23), получим окончательно

$$\psi_p(t) - \int_{t-T}^t g(t, \tau) \varphi_p(\tau) d\tau = \sum_{q=1}^N c_{pq} \lambda_q \quad (p = 1, \dots, N). \quad (13.5.28)$$

Уравнения (13.5.22) и (13.5.28) являются окончательной формой уравнений, определяющих весовую функцию оптимальной линейной системы в случае, когда требуемый выходной сигнал W и входной сигнал Z выражаются формулами (13.5.11) и (13.5.14).

Рассмотрим теперь предельный случай, когда дисперсии величин $U_1, \dots, U_N$ стремятся к бесконечности. В этом случае $\gamma_{pp} \rightarrow \infty$ ($p = 1, \dots, N$), т. е. все диагональные элементы определителя Γ и соответствующие элементы определителей Γ_{pq} стремятся к бесконечности. Чтобы вычислить величины c_{pq} в этом случае, разделим все элементы первой строки и первого столбца определителя Γ на $\sqrt{\gamma_{11}}$, все элементы второй строки и второго столбца на $\sqrt{\gamma_{22}}$ и т. д., все элементы N -й строки и N -го столбца —

на $\sqrt{\gamma_{NN}}$. Тогда получим

$$\Gamma = \gamma_{11}\gamma_{22} \dots \gamma_{NN} \begin{vmatrix} 1 & \rho_{12} & \dots & \rho_{1N} \\ \rho_{21} & 1 & \dots & \rho_{2N} \\ \dots & \dots & \dots & \dots \\ \rho_{N1} & \rho_{N2} & \dots & 1 \end{vmatrix} = \gamma_{11}\gamma_{22} \dots \gamma_{NN} P, \quad (13.5.29)$$

где

$$\rho_{pq} = \frac{\gamma_{pq}}{\sqrt{\gamma_{pp}\gamma_{qq}}} \quad (p, q = 1, \dots, N), \quad (13.5.30)$$

причем вследствие известного из теории вероятностей неравенства для начальных моментов второго порядка (см. [54], § 3.9)

$$|\rho_{pq}| \leq 1.$$

Совершенно таким же образом преобразуем определитель Γ_{pq} . Имея в виду, что этот определитель отличается от определителя Γ тем, что в нем отсутствует p -я строка и q -й столбец, приходим к выводу, что перед определителем появятся те же множители, что и в (13.5.29), за исключением $\sqrt{\gamma_{pp}}$ и $\sqrt{\gamma_{qq}}$. Следовательно,

$$\Gamma_{pq} = \frac{\gamma_{11}\gamma_{22} \dots \gamma_{NN}}{\sqrt{\gamma_{pp}\gamma_{qq}}} P_{pq}, \quad (13.5.31)$$

где P_{pq} — алгебраическое дополнение элемента ρ_{pq} в определителе (13.5.29). Подставляя выражения (13.5.29) и (13.5.31) в (13.5.26), получим

$$c_{pq} = \frac{P_{pq}}{P \sqrt{\gamma_{pp}\gamma_{qq}}}. \quad (13.5.32)$$

Отсюда видно, что если хотя бы одна из величин γ_{pp} или γ_{qq} стремится к бесконечности, то $c_{pq} \rightarrow 0$. Если все величины $\gamma_{11}, \dots, \gamma_{NN}$ стремятся к бесконечности, то все величины c_{pq} стремятся к нулю и уравнения (13.5.28) принимают вид

$$\psi_p(t) = \int_{t-T}^t g(t, \tau) \psi_p(\tau) d\tau \quad (p = 1, \dots, N). \quad (13.5.33)$$

Таким образом, в случае, когда $U_1, \dots, U_N$ являются неизвестными величинами, которые могут иметь любые значения, задача определения оптимальной линейной системы сводится к решению уравнения (13.5.22) при дополнительных условиях (13.5.33).

Если только некоторые из величин $\gamma_{11}, \dots, \gamma_{NN}$, например первые k , стремятся к бесконечности, то все величины c_{pq} , у которых хотя бы один из номеров равен числам $1, \dots, k$, стремятся

к нулю вследствие (13.5.32). Следовательно, в этом случае первые k уравнений (13.5.28) принимают вид (13.5.33), а в остальных $N - k$ суммирование по q в правой части будет производиться от $q = k + 1$ до $q = N$.

Выясним подробнее смысл условий (13.5.33). Для этого найдем фактическую выходную переменную оптимальной системы $W^*(t)$. Вследствие формул (13.5.2) и (13.5.14) получим в данном случае

$$\begin{aligned} W^*(t) &= \int_{t-T}^t g(t, \tau) Z(\tau) d\tau = \\ &= \sum_{r=1}^N U_r \int_{t-T}^t g(t, \tau) \varphi_r(\tau) d\tau + \int_{t-T}^t g(t, \tau) X(\tau) d\tau. \end{aligned} \quad (13.5.34)$$

Сравнивая эту формулу с (13.5.11) и принимая во внимание, что математические ожидания случайных функций $X(t)$ и $Y(t)$ тождественно равны нулю, приходим к заключению, что математическое ожидание ошибки оптимальной системы равно

$$\begin{aligned} M[E(t)] &= M[W^*(t) - W(t)] = \\ &= \sum_{r=1}^N M[U_r] \left\{ \int_{t-T}^t g(t, \tau) \varphi_r(\tau) d\tau - \psi_r(t) \right\}. \end{aligned} \quad (13.5.35)$$

Отсюда видно, что при выполнении, условий (13.5.33) математическое ожидание ошибки системы тождественно равно нулю, и, следовательно, математическое ожидание выходного сигнала $W^*(t)$ совпадает с математическим ожиданием требуемого выходного сигнала $W(t)$. Пользуясь терминологией математической статистики, можно сказать, что $W^*(t)$ является несмещенной оценкой сигнала $W(t)$. Вследствие этого условия (13.5.33) обычно называются *условиями несмещенности*.

Оптимальная линейная система, удовлетворяющая дополнительным условиям несмещенности (13.5.33), имеет ту особенность, что в идеальном случае при отсутствии нерегулярной части полезного сигнала и помехи, т. е. когда $X(t) \equiv Y(t) \equiv 0$, она идеально точно воспроизводит требуемый выходной сигнал $W(t)$. Действительно, из формул (13.5.11), (13.5.34) и (13.5.33) видно, что в этом случае выходной сигнал $W^*(t)$ тождественно равен требуемому выходному сигналу $W(t)$.

Пример 13.5.2. В условиях примера 13.5.1 можно считать, что $\gamma_{11} = \gamma_{22} = \infty$, в то время как величина γ_{33} конечна. Следовательно, в этом случае $c_{11} = c_{12} = c_{13} = c_{21} = c_{22} = c_{23} = c_{31} = c_{32} = 0$ и только величина $c_{33} = 1/\gamma_{33}$ отлична от нуля. Уравнения (13.5.22) и (13.5.28) при этом

принимают вид

$$\left. \begin{aligned} \int_{t-T}^t K_x(\tau, \sigma) g(t, \tau) d\tau &= K_{yx}(t, \sigma) + \lambda_1 + \lambda_2 \sigma + \lambda_3 \sigma^2, \\ \int_{t-T}^t g(t, \tau) d\tau &= 1, \quad \int_{t-T}^t g(t, \tau) \tau d\tau = t, \\ t^2 - \int_{t-T}^t g(t, \tau) \tau^2 d\tau &= c_{22} \lambda_3. \end{aligned} \right\} \quad (13.5.36)$$

Корреляционная функция случайной функции $X(t) = X_1(t) + N_1(t)$ определяется формулой

$$K_x(\tau, \sigma) = K_{x_1}(\tau, \sigma) + K_{n_1}(\tau, \sigma),$$

так как радиопомеху $N_1(t)$ и случайный аэродинамический момент $N_2(t)$ можно считать независимыми. Случайная функция $Y(t)$ в данном случае совпадает с $X_1(t)$. Поэтому

$$\begin{aligned} K_{yx}(t, \sigma) &= M[X_1(t) X(\sigma)] = M[X_1(t) \{X_1(\sigma) + N_1(\sigma)\}] = \\ &= M[X_1(t) X_1(\sigma)] = K_{x_1}(t, \sigma). \end{aligned}$$

Таким образом, зная корреляционные функции помех $N_1(t)$ и $N_2(t)$ и оператор исполнительной части следящей системы, мы можем вычислить корреляционные функции, входящие в первое уравнение (13.5.36). Тогда получим четыре уравнения (13.5.36), определяющие неизвестную весовую функцию оптимальной замкнутой следящей системы $g(t, \tau)$ и три вспомогательные неизвестные функции λ_1 , λ_2 и λ_3 .

Совершенно так же выводятся уравнения, определяющие оптимальную многомерную линейную систему. Если входной сигнал системы $Z(t)$ представляет собой n -мерную векторную случайную функцию, а требуемый выходной сигнал $W(t)$ — m -мерную векторную случайную функцию, то матрица весовых функций оптимальной линейной системы $g(t, \tau)$, обладающей минимальными средними квадратическими ошибками на всех выходах, определяется уравнением

$$\int_{t-T}^t g(t, \tau) \Gamma_z(\tau, \sigma) d\tau = \Gamma_{wz}(t, \sigma) \quad (t-T \leq \sigma \leq t), \quad (13.5.37)$$

где $\Gamma_z(\tau, \sigma)$ — матрица начальных моментов второго порядка составляющих входного сигнала $Z(t)$, а $\Gamma_{wz}(t, \sigma)$ — матрица взаимных начальных моментов второго порядка требуемого выходного и входного сигналов $W(t)$ и $Z(t)$:

$$\Gamma_z(\tau, \sigma) = M[Z(\tau) Z(\sigma)^T], \quad \Gamma_{wz}(t, \sigma) = M[W(t) Z(\sigma)^T] \quad (13.5.38)$$

(как и в конце § 7.2, векторы $Z(t)$ и $W(t)$ представлены в виде матриц-столбцов, а верхний индекс «т» означает операцию транспонирования матрицы). Уравнение (13.5.5), очевидно, является частным случаем уравнения (13.5.37) при $m = n = 1$.

Следует отметить, что решение интегрального уравнения (13.5.5) (или (13.5.37)) весьма чувствительно к малым изменениям исходных данных. Так, например, если $Z(t)$ представляет собой стационарную случайную функцию с дробно-рациональной спектральной плотностью $s_z(\omega) = P(\omega^2)/Q(\omega^2)$, где P и Q — полиномы относительно ω^2 степени m и n , $n > m$, соответственно, то выходной сигнал оптимальной системы, как мы видели, может содержать линейную комбинацию производных входного сигнала до порядка $p = n - m - 1$ включительно. Если при определении спектральной плотности входного сигнала $Z(t)$ была допущена ошибка и его фактическая спектральная плотность равна

$$s_z(\omega) = \frac{P(\omega^2) + \varepsilon \omega^{2m+2}}{Q(\omega^2)},$$

где ε — сколь угодно малое положительное число, то выходной сигнал найденной «оптимальной» системы будет содержать белый шум, вследствие чего средняя квадратическая ошибка этой системы будет бесконечной. Таким образом, при сколь угодно малом изменении спектральной плотности стационарного входного сигнала оптимальная система не только перестает быть оптимальной, но может стать и совсем непригодной для решения поставленной задачи.

Задачи, решение которых сильно изменяется при сколь угодно малом изменении исходных данных, называются в математике *некорректно поставленными* или просто *некорректными*. Таким образом, задача оптимизации линейной системы по критерию минимума средней квадратической ошибки является некорректной. Чтобы сделать ее корректной, необходимо сузить класс допустимых систем, исключив из рассматриваемого класса линейных систем все системы, выходной сигнал которых содержит текущие значения входного сигнала и его производных. Это можно сделать, потребовав, чтобы выходной сигнал системы имел конечную дисперсию при подаче на вход стационарного белого шума. Математически это условие, на основании (7.6.16), выражается неравенством

$$\int_{-T}^T g^2(t, \tau) d\tau < \infty. \quad (13.5.39)$$

Чтобы получить дополнительные возможности практической реализации системы, В. И. Кухтенко предложил, кроме условия (13.5.39), наложить на весовую функцию оптимальной системы

еще условия вида [38]

$$\int_{t-T}^t \left[\frac{\partial^2 g(t, \tau)}{\partial t^2} \right]^2 d\tau < \infty \quad (k = 1, \dots, r). \quad (13.5.40)$$

Физически эти условия означают, что производные выходного сигнала системы до порядка r включительно имеют конечные дисперсии при действии белого шума на входе.

Дополнительные условия (13.5.39) и (13.5.40) «регуляризуют» задачу определения оптимальной системы, делают ее корректной. Общая теория некорректных задач и их регуляризации была разработана А. Н. Тихоновым [94].

Само собой разумеется, что, сузив класс допустимых систем дополнительными условиями (13.5.39) и (13.5.40), мы отходим от оптимума, ухудшаем качество системы при точном знании исходных данных. Уменьшение чувствительности системы к ошибкам в исходных данных покупается ценой ухудшения ее качества при точном совпадении статистических характеристик входного и требуемого выходного сигналов с расчетными. Для оценки потери качества системы по сравнению с истинной оптимальной необходимо уметь находить оптимальные системы и оценивать их точность. Для этого и служит излагаемая статистическая теория оптимальных систем.

§ 13.6. Общий анализ уравнений, определяющих оптимальную линейную систему

Произведем сначала общий анализ системы уравнений (13.5.22) и (13.5.28) с целью определить последовательность операций, которые необходимо выполнить для нахождения весовой функции оптимальной линейной системы. Рассмотрим уравнения

$$\int_{t-T}^t K_x(\tau, \sigma) g^{(0)}(t, \tau) d\tau = K_{yx}(t, \sigma) \quad (t-T \leq \sigma \leq t), \quad (13.6.1)$$

$$\int_{t-T}^t K_x(\tau, \sigma) g^{(q)}(t, \tau) d\tau = \varphi_q(\sigma) \quad (t-T \leq \sigma \leq t; q = 1, \dots, N) \quad (13.6.2)$$

с неизвестными функциями $g^{(0)}, g^{(1)}, \dots, g^{(N)}$. Эти уравнения являются линейными интегральными уравнениями, так как содержат неизвестные функции под знаком интеграла, и притом в первой степени. Каждое из уравнений (13.6.1) и (13.6.2) содержит только одну неизвестную функцию $g^{(0)}(t, \tau), g^{(1)}(t, \tau), \dots, g^{(N)}(t, \tau)$. Следовательно, уравнения (13.6.1) и (13.6.2) могут быть решены по отдельности.

Допустим теперь, что мы решили эти уравнения и нашли все функции $g^{(0)}, g^{(1)}, \dots, g^{(N)}$. Умножим каждое из уравнений (13.6.2) на соответствующую величину λ_q и сложим полученные таким образом равенства с (13.6.1). Тогда получим

$$\begin{aligned} \int_{t-T}^t K_x(\tau, \sigma) \left[g^{(0)}(t, \tau) + \sum_{q=1}^N \lambda_q(t) g^{(q)}(t, \tau) \right] d\tau = \\ = K_{yx}(t, \sigma) + \sum_{q=1}^N \lambda_q \varphi_q(\sigma). \end{aligned} \quad (13.6.3)$$

Сравнивая это равенство с (13.5.22), приходим к заключению, что неизвестная весовая функция оптимальной системы $g(t, \tau)$ выражается формулой

$$g(t, \tau) = g^{(0)}(t, \tau) + \sum_{q=1}^N \lambda_q(t) g^{(q)}(t, \tau). \quad (13.6.4)$$

Таким образом, умея решать интегральные уравнения (13.6.1) и (13.6.2), мы можем выразить неизвестную весовую функцию оптимальной системы $g(t, \tau)$ через решения этих уравнений и вспомогательные неизвестные $\lambda_1, \dots, \lambda_N$ формулой (13.6.4). Если теперь найти вспомогательные неизвестные λ_q , то задача будет полностью решена.

Для нахождения $\lambda_1, \dots, \lambda_N$ подставим выражение (13.6.4) весовой функции $g(t, \tau)$ в уравнения (13.5.28). Тогда получим

$$\begin{aligned} \psi_p(t) - \int_{t-T}^t g^{(0)}(t, \tau) \varphi_p(\tau) d\tau - \sum_{q=1}^N \lambda_q \int_{t-T}^t g^{(q)}(t, \tau) \varphi_p(\tau) d\tau = \\ = \sum_{q=1}^N c_{pq} \lambda_q. \end{aligned} \quad (13.6.5)$$

Если функции $g^{(0)}, g^{(1)}, \dots, g^{(N)}$ известны, то интегралы в левой части уравнения (13.6.5) можно вычислить. Введем для них обозначения

$$b_{p0} = \int_{t-T}^t g^{(0)}(t, \tau) \varphi_p(\tau) d\tau, \quad b_{pq} = \int_{t-T}^t g^{(q)}(t, \tau) \varphi_p(\tau) d\tau \quad (p, q = 1, \dots, N).$$

Тогда уравнения (13.6.5) примут вид

$$\sum_{q=1}^N (b_{pq} + c_{pq}) \lambda_q = \psi_p(t) - b_{p0} \quad (p = 1, \dots, N). \quad (13.6.7)$$

Таким образом, неизвестные $\lambda_1, \dots, \lambda_N$ определяются системой линейных алгебраических уравнений (13.6.7). Решив эту систему

уравнений, мы найдем $\lambda_1, \dots, \lambda_N$, после чего формула (13.6.4) полностью определит весовую функцию оптимальной линейной системы.

Заметим, что величины $b_{10}, \dots, b_{N0}, b_{11}, \dots, b_{N1}, \dots, b_{1N}, \dots, b_{NN}$, определяемые формулами (13.6.6), в общем случае являются функциями времени t . Поэтому определяемые системой уравнений (13.6.7) величины $\lambda_1, \dots, \lambda_N$ также являются функциями времени, как это уже было отмечено раньше.

Легко доказать еще, что матрица коэффициентов уравнений (13.6.7) симметрична. Для этого заметим, во-первых, что $\gamma_{pq} = \gamma_{qp}$, вследствие чего и $c_{pq} = c_{qp}$. Во-вторых, на основании (13.6.6) и (13.6.2), принимая во внимание свойство симметрии корреляционной функции, имеем:

$$\begin{aligned} b_{qp} &= \int_{t-T}^t \int_{t-T}^t K_x(\tau, \sigma) g^{(q)}(t, \tau) g^{(p)}(t, \sigma) d\tau d\sigma = \\ &= \int_{t-T}^t g^{(q)}(t, \tau) \left\{ \int_{t-T}^t K_x(\sigma, \tau) g^{(p)}(t, \sigma) d\sigma \right\} d\tau = \\ &= \int_{t-T}^t g^{(q)}(t, \tau) \varphi_p(\tau) d\tau = b_{pq} \quad (p, q = 1, \dots, N). \end{aligned}$$

Итак, мы доказали, что решение уравнений (13.5.22) и (13.5.28), определяющих весовую функцию оптимальной линейной системы, сводится к последовательному выполнению следующих операций:

1) решение линейных интегральных уравнений (13.6.1) и (13.6.2), определяющих вспомогательные весовые функции

$$g^{(0)}(t, \tau), g^{(1)}(t, \tau), \dots, g^{(N)}(t, \tau);$$

2) вычисление коэффициентов b_{p0} и b_{pq} по формулам (13.6.6);

3) решение системы линейных алгебраических уравнений (13.6.7), определяющих вспомогательные неизвестные функции $\lambda_1, \dots, \lambda_N$ *);

4) вычисление оптимальной весовой функции по формуле (13.6.4).

Наиболее сложной из всех этих операций является решение интегральных уравнений (13.6.1) и (13.6.2). Все эти уравнения имеют одну общую черту. А именно ядром этих уравнений является одна и та же корреляционная функция $K_x(\tau, \sigma)$ суммы нерегулярной части полезного сигнала и помехи, и различаются они только

*) Можно показать, что определитель системы уравнений (13.6.7) всегда отличен от нуля (точнее, всегда положителен). Следовательно, система уравнений (13.6.7) всегда имеет единственное решение.

правыми частями. Таким образом, уравнения (13.6.1) и (13.6.2) являются линейными интегральными уравнениями вида

$$\int_{t-T}^t K_x(\tau, \sigma) g(t, \tau) d\tau = f(t, \sigma) \quad (t-T \leq \sigma \leq t) \quad (13.6.8)$$

с заданной функцией $f(t, \sigma)$ и неизвестной функцией $g(t, \tau)$. К аналогичным операциям сводится решение уравнения (13.5.37), определяющего матрицу весовых функций оптимальной многомерной линейной системы.

Чтобы завершить решение задачи, остается оценить точность оптимальной линейной системы, т. е. найти минимальную среднюю квадратическую ошибку, достижимую с помощью оптимальной линейной системы

$$\eta_{\min} = M[E^2] = M[(W^* - W)^2] = M[(AZ - W)^2] = \\ = M[(AZ - W)AZ] - M[(AZ - W)W]. \quad (13.6.9)$$

Согласно общему условию, которому должен удовлетворять оператор оптимальной системы (13.4.12) и которое должно выполняться для всех операторов B рассматриваемого класса, в том числе и для оптимального оператора A ,

$$M[(AZ - W)AZ] = 0.$$

Следовательно, формула (13.6.9) принимает вид

$$\eta_{\min} = M[W^2] - M[WAZ] = M[W^2] - M[WW^*]. \quad (13.6.10)$$

Эта формула определяет минимальную среднюю квадратическую ошибку, достижимую с помощью любой оптимальной системы, как линейной, так и нелинейной.

Теперь применим формулу (13.6.10) к случаю оптимальной линейной системы. Первое слагаемое в правой части формулы (13.6.10) представляет собой начальный момент второго порядка требуемого выходного сигнала $\Gamma_w(t, t)$. Второе слагаемое есть математическое ожидание произведения случайной функции $W(t)$ на выходной сигнал оптимальной системы $W^*(t)$. Поэтому можно написать

$$\eta_{\min} = \Gamma_w(t, t) - M \left[W(t) \int_{t-T}^t g(t, \tau) Z(\tau) d\tau \right] = \\ = \Gamma_w(t, t) - \int_{t-T}^t \Gamma_{wz}(t, \tau) g(t, \tau) d\tau. \quad (13.6.11)$$

Подставляя в эту формулу выражение (13.5.19) и аналогичное выражение для $\Gamma_w(t, t)$:

$$\Gamma_w(t, t) = K_y(t, t) + \sum_{p, q=1}^N \gamma_{pq} \psi_p(t) \psi_q(t) \quad (13.6.12)$$

и принимая во внимание, что величина $K_y(t, t)$ представляет собой дисперсию случайной функции $Y(t)$, получим

$$\eta_{\min} = D_y - \int_{t-T}^t K_{yx}(t, \tau) g(t, \tau) d\tau + \\ + \sum_{p, q=1}^N \gamma_{pq} \left[\psi_q(t) - \int_{t-T}^t g(t, \tau) \varphi_q(\tau) d\tau \right] \psi_p(t)$$

или, принимая во внимание, что $\gamma_{pq} = \gamma_{qp}$, и учитывая (13.5.21),

$$\eta_{\min} = D_y - \int_{t-T}^t K_{yx}(t, \tau) g(t, \tau) d\tau + \sum_{p=1}^N \lambda_p \psi_p(t). \quad (13.6.13)$$

Теперь остается подставить в эту формулу выражение (13.6.4) весовой функции оптимальной системы. Тогда получим

$$\eta_{\min} = D_y - \int_{t-T}^t K_{yx}(t, \tau) g^{(0)}(t, \tau) d\tau + \\ + \sum_{p=1}^N \lambda_p(t) \left[\psi_p(t) - \int_{t-T}^t K_{yx}(t, \sigma) g^{(p)}(t, \sigma) d\sigma \right]. \quad (13.6.14)$$

Но вследствие уравнений (13.6.1) и (13.6.2) и симметрии корреляционной функции

$$\int_{-T}^t K_{yx}(t, \sigma) g^{(p)}(t, \sigma) d\sigma = \\ = \int_{t-T}^t \int_{t-T}^t K_x(\tau, \sigma) g^{(0)}(t, \tau) g^{(p)}(t, \sigma) d\tau d\sigma = \\ = \int_{t-T}^t g^{(0)}(t, \tau) \left\{ \int_{t-T}^t K_x(\sigma, \tau) g^{(p)}(t, \sigma) d\sigma \right\} d\tau = \int_{t-T}^t g^{(0)}(t, \tau) \varphi_p(\tau) d\tau.$$

Поэтому, принимая во внимание обозначения (13.6.6) и вводя дополнительное обозначение

$$b_{00} = \int_{t-T}^t K_{yx}(t, \tau) g^{(0)}(t, \tau) d\tau, \quad (13.6.15)$$

представим формулу (13.6.14) в виде

$$\eta_{\min} = D_y(t) - b_{00}(t) + \sum_{p=1}^N \lambda_p(t) [\psi_p(t) - b_{p0}(t)]. \quad (13.6.16)$$

Заметим еще, что в случае отсутствия нерегулярной части полезного сигнала $X_1(t) \equiv Y(t) \equiv 0$, $K_{yx}(t, \sigma) \equiv 0$ и вследствие этого весовая функция $g^{(0)}$ и все величины D_y , b_{00} , b_{10} , ..., b_{N0} равны нулю.

Следовательно, в случае отсутствия нерегулярной части полезного сигнала формула (13.6.16) принимает следующий простой вид:

$$\eta_{\min} = \sum_{p=1}^N \lambda_p(t) \psi_p(t). \quad (13.6.17)$$

Пример 13.6.1. Решение уравнений (13.5.36) примера 13.5.2 на основании изложенного сводится к определению вспомогательных весовых функций $g^{(0)}$, $g^{(1)}$, $g^{(2)}$, $g^{(3)}$ путем решения интегральных уравнений

$$\left. \begin{aligned} \int_{t-T}^t K_x(\tau, \sigma) g^{(0)}(t, \tau) d\tau &= K_{yx}(t, \sigma), \\ \int_{t-T}^t K_x(\tau, \sigma) g^{(1)}(t, \tau) d\tau &= 1, \\ \int_{t-T}^t K_x(\tau, \sigma) g^{(2)}(t, \tau) d\tau &= \sigma, \\ \int_{t-T}^t K_x(t, \sigma) g^{(3)}(t, \tau) d\tau &= \sigma^2, \end{aligned} \right\} \quad (13.6.18)$$

вычислению величин b_{pq} по формулам

$$\begin{aligned} b_{00} &= \int_{t-T}^t g^{(0)}(t, \tau) K_{yx}(t, \tau) d\tau, & b_{10} &= \int_{t-T}^t g^{(0)}(t, \tau) d\tau, \\ b_{20} &= \int_{t-T}^t g^{(0)}(t, \tau) \tau d\tau, & b_{30} &= \int_{t-T}^t g^{(0)}(t, \tau) \tau^2 d\tau, & b_{11} &= \int_{t-T}^t g^{(1)}(t, \tau) d\tau, \\ b_{12} &= b_{21} = \int_{t-T}^t g^{(2)}(t, \tau) d\tau = \int_{t-T}^t g^{(1)}(t, \tau) \tau d\tau, \\ b_{13} &= b_{31} = \int_{t-T}^t g^{(3)}(t, \tau) d\tau = \int_{t-T}^t g^{(1)}(t, \tau) \tau^2 d\tau, & b_{22} &= \int_{t-T}^t g^{(2)}(t, \tau) \tau d\tau, \\ b_{23} &= \int_{t-T}^t g^{(3)}(t, \tau) \tau^2 d\tau, & b_{33} &= b_{32} = \int_{t-T}^t g^{(3)}(t, \tau) \tau d\tau = \int_{t-T}^t g^{(2)}(t, \tau) \tau^2 d\tau \end{aligned}$$

и определению величин λ_1 , λ_2 и λ_3 путем решения системы линейных алгебраических уравнений

$$\left. \begin{aligned} b_{11}\lambda_1 + b_{12}\lambda_2 + b_{13}\lambda_3 &= 1 - b_{10}, \\ b_{21}\lambda_1 + b_{22}\lambda_2 + b_{23}\lambda_3 &= t - b_{20}, \\ b_{31}\lambda_1 + b_{32}\lambda_2 + (b_{33} + c_{33})\lambda_3 &= t^2 - b_{30}. \end{aligned} \right\} \quad (13.6.19)$$

После этого формула (13.6.4) дает следующее выражение для весовой функции оптимальной линейной системы:

$$g(t, \tau) = g^{(0)}(t, \tau) + \lambda_1(t) g^{(1)}(t, \tau) + \lambda_2(t) g^{(2)}(t, \tau) + \lambda_3(t) g^{(3)}(t, \tau). \quad (13.6.20)$$

Формула (13.6.16) дает следующее выражение для минимальной средней квадратической ошибки, соответствующей оптимальной системе:

$$\eta_{\min} = Dx_1(t) - b_{00}(t) + \lambda_1(t) [1 - b_{10}(t)] + \lambda_2(t) [t - b_{20}(t)] + \lambda_3(t) [t^2 - b_{30}(t)]. \quad (13.6.21)$$

Пример 13.6.2. В случае, когда случайный аэродинамический момент N_2 не действует на исполнительное устройство следящей системы, $N_2(t) = 0$ и нерегулярная часть полезного сигнала $X_1(t)$ отсутствует. В этом случае $Y = 0$ и, следовательно, $K_{yx}(t, \sigma) = 0$. Поэтому величины b_{00} , b_{10} , b_{20} , b_{30} , а также D_y равны нулю. В этом случае оптимальная система определяется решением последних трех интегральных уравнений (13.6.18), решением системы линейных алгебраических уравнений

$$\left. \begin{aligned} b_{11}\lambda_1 + b_{12}\lambda_2 + b_{13}\lambda_3 &= 1, \\ b_{21}\lambda_1 + b_{22}\lambda_2 + b_{23}\lambda_3 &= t, \\ b_{31}\lambda_1 + b_{32}\lambda_2 + (b_{33} + c_{33})\lambda_3 &= t^2 \end{aligned} \right\} \quad (13.6.22)$$

и вычислением весовой функции оптимальной системы по формуле

$$g(t, \tau) = \lambda_1(t) g^{(1)}(t, \tau) + \lambda_2(t) g^{(2)}(t, \tau) + \lambda_3(t) g^{(3)}(t, \tau). \quad (13.6.23)$$

Формула (13.6.21) для минимальной средней квадратической ошибки, соответствующей оптимальной линейной системе, примет вид

$$\eta_{\min} = \lambda_1(t) + \lambda_2(t) t + \lambda_3(t) t^2. \quad (13.6.24)$$

Пример 13.6.3. В случае оптимальной системы, которая должна с минимальной средней квадратической ошибкой экстраполировать закон движения самолета, т. е. выдавать его упрежденные координаты, в уравнениях (13.6.19) и (13.6.22) и в формулах (13.6.21) и (13.6.24) вместо t и t^2 будут фигурировать соответственно $t + \Delta$, $(t + \Delta)^2$. Кроме того, изменится операция, при помощи которой случайная функция $Y(t)$ формируется из $X_1(t)$, вследствие чего изменится правая часть первого уравнения (13.6.18). В остальном все уравнения и формулы предыдущих двух примеров останутся неизменными.

Пример 13.6.4. В случае оптимальной системы, которая должна с минимальной средней квадратической ошибкой выдавать текущее значение скорости самолета, величины 1 , t , t^2 в уравнениях (13.6.19) и (13.6.22) и в формулах (13.6.21) и (13.6.24) должны быть заменены соответственно величинами 0 , 1 , $2t$. Кроме того, изменится правая часть первого уравнения (13.6.18). Все остальные формулы и уравнения примеров 13.6.1 и 13.6.2 останутся неизменными.

Рассмотренные примеры показывают, что весовые функции $g^{(r)}(t, \tau)$ и интегральные уравнения (13.6.2), которым они удовлетворяют, полностью определяются структурой входного полезного сигнала и ни в какой мере не зависят от назначения проектируемой системы, т. е. от идеальной операции L , которую необходимо произвести над входным полезным сигналом. От этой операции, как показывают формулы (13.5.10), в случае отсутствия нерегулярной части полезного сигнала зависят только функции $\psi_r(t)$, входящие в правые части уравнений (13.6.7), определяющих величины $\lambda_r(t)$, и в формулу (13.6.17) для минимальной средней квадратической ошибки, соответствующей оптимальной системе. Таким образом, в случае отсутствия нерегулярной части полезного сигнала от назначения системы зависят только функции $\psi_r(t)$ и $\lambda_r(t)$.

Аналогичные результаты дает анализ уравнений, определяющих оптимальную многомерную линейную систему. В частности, получаем следующую формулу для матрицы Π начальных моментов второго порядка вектора ошибки оптимальной системы:

$$\Pi = \Gamma_w(t, t) - \int_{t-T}^t g(t, \tau) \Gamma_{wz}(t, \tau)^T d\tau, \quad (13.6.25)$$

где в дополнение к обозначениям конца § 13.5 $\Gamma_w(t, \tau)$ — матрица начальных моментов второго порядка требуемого выходного сигнала:

$$\Gamma_w(t, \tau) = M [W(t) W(\tau)^T], \quad (13.6.26)$$

а верхний индекс « t » по-прежнему означает операцию транспонирования матрицы. Диагональные элементы матрицы Π представляют собой средние квадраты ошибок на выходах оптимальной системы (т. е. минимальные достижимые средние квадраты ошибок на соответствующих выходах системы).

§ 13.7. Уравнения, определяющие оптимальную дискретную линейную систему

Уравнения предыдущего параграфа определяют, в частности, оптимальные дискретные линейные системы. В этом случае входная функция действует на систему только в дискретном ряде моментов времени $t_1 = t - T$, $t_2, \dots, t_h = t$ и весовая функция оптимальной линейной системы согласно (5.2.4) будет представлять собой линейную комбинацию δ -функций:

$$g(t_h, \tau) = \sum_{l=1}^h g_{hl} \delta(\tau - t_l). \quad (13.7.1)$$

Полагая в уравнениях (13.6.1) и (13.6.2) $t = t_h$, $\sigma = t_q$ и заменяя

весовые функции $g^{(0)}, g^{(1)}, \dots, g^{(N)}$ соответствующими линейными комбинациями δ -функций, приведем эти уравнения к виду

$$\sum_{l=1}^h K_x(t_l, t_q) g_{hl}^{(0)} = K_{yx}(t_h, t_q) \quad (q = 1, \dots, h), \quad (13.7.2)$$

$$\sum_{l=1}^h K_x(t_l, t_q) g_{hl}^{(r)} = \varphi_r(t_q) \quad (q = 1, \dots, h; r = 1, \dots, N). \quad (13.7.3)$$

Таким образом, интегральные уравнения (13.6.1) и (13.6.2) для оптимальной дискретной линейной системы заменяются системами линейных алгебраических уравнений (13.7.2) и (13.7.3).

Формулы (13.6.6), определяющие величины b_{pq} , для дискретной системы принимают вид

$$b_{pq} = \sum_{l=1}^h g_{hl}^{(q)} \varphi_p(t_l) \quad (p = 1, \dots, N; q = 0, 1, \dots, N). \quad (13.7.4)$$

После нахождения коэффициентов b_{pq} и решения систем линейных алгебраических уравнений (13.7.2), (13.7.3) и (13.6.7) весовые коэффициенты оптимальной дискретной линейной системы определяются формулой

$$g_{hl} = g_{hl}^{(0)} + \sum_{q=1}^N \lambda_q g_{hl}^{(q)} \quad (l = 1, \dots, h). \quad (13.7.5)$$

которая является следствием общей формулы (13.6.4).

Формулы (13.6.16) и (13.6.17) определяющие минимальную среднюю квадратическую ошибку, соответствующую оптимальной линейной системе, справедливы и для дискретных линейных систем при $t = t_h$. При этом величина b_{00} определяется формулой

$$b_{00} = \sum_{l=1}^h g_{hl}^{(0)} K_{yx}(t_h, t_l), \quad (13.7.6)$$

вытекающей из (13.6.15).

Таким образом, оптимальные дискретные линейные системы всегда находятся принципиально просто путем решения систем линейных алгебраических уравнений и элементарных алгебраических вычислений. Однако при большом числе h моментов наблюдения входного сигнала решение систем уравнений (13.7.2) и (13.7.3) связано с трудоемкими вычислениями. В § 14.7 мы изложим общий метод решения уравнений (13.7.2) и (13.7.3), дающий явные выражения для весовых коэффициентов $g_{hl}^{(q)}$ при любом числе моментов наблюдения h .

Глава 14

МЕТОДЫ ОПРЕДЕЛЕНИЯ ОПТИМАЛЬНЫХ ЛИНЕЙНЫХ СИСТЕМ

§ 14.1. Определение оптимальной линейной системы в случае белого шума на входе

В предыдущей главе было показано, что задача определения оптимальной линейной системы с одним входом и одним выходом по критерию минимума средней квадратической ошибки приводит к решению интегральных уравнений вида

$$\int_{t-T}^t K_x(\tau, \sigma) g(t, \tau) d\tau = f(t, \sigma) \quad (t-T \leq \sigma \leq t), \quad (14.1.1)$$

где $f(t, \sigma)$ — известная функция, а $g(t, \tau)$ — неизвестная весовая функция. Умея решать интегральные уравнения такого типа, мы можем элементарными вычислениями находить оптимальные линейные системы.

В следующей главе мы увидим, что к интегральным уравнениям вида (14.1.1) приводится также задача определения оптимальной системы среди всех возможных систем по различным критериям (см. [53], §§ 141—145). Поэтому интегральное уравнение (14.1.1) является основным уравнением статистической теории оптимальных систем, и методы решения этого уравнения составляют основу теории оптимальных систем.

Нетрудно видеть, что уравнение (14.1.1) легко решается в случае, когда случайная функция $X(t)$ представляет собой белый шум, так как в этом случае корреляционная функция $K_x(\tau, \sigma)$ выражается через δ -функцию и уравнение (14.1.1) превращается в алгебраическое. Действительно, подставляя в уравнение (14.1.1) выражение корреляционной функции белого шума $X(t)$:

$$K_x(\tau, \sigma) = G(\sigma) \delta(\tau - \sigma) \quad (14.1.2)$$

и вспоминая, что интеграл от произведения δ -функции на любую функцию равен значению этой функции в точке разрыва δ -функции, получаем

$$G(\sigma) g(t, \sigma) = f(t, \sigma) \quad (t - T \leq \sigma \leq t). \quad (14.1.3)$$

Таким образом, в случае, когда сумма нерегулярной части полезного сигнала и помехи $X(t)$ является белым шумом, интегральное

уравнение (14.1.1) приводится к линейному алгебраическому уравнению. Решая это уравнение и заменяя букву σ буквой τ , находим

$$g(t, \tau) = \frac{f(t, \tau)}{G(\tau)} \quad (t - T \leq \tau \leq t). \quad (14.1.4)$$

Таким образом, мы имеем выражение функции $g(t, \tau)$ в интересующем нас интервале наблюдения $t - T \leq \tau \leq t$. Вне этого интервала ее следует принять равной нулю, так как значения входного сигнала, предшествующие моменту $t - T$, система не должна использовать (или не может использовать, если момент $t - T$ совпадает с моментом начала работы системы), а при $\tau > t$ весовая функция $g(t, \tau)$ равна нулю вследствие условия физической возможности системы.

Подставляя в (14.1.4) вместо $f(t, \tau)$ по очереди функции $K_{yx}(t, \tau)$, $\varphi_1(\tau)$, ..., $\varphi_N(\tau)$, мы получим соответственно функции $g^{(0)}(t, \tau)$, $g^{(1)}(t, \tau)$, ..., $g^{(N)}(t, \tau)$. После этого весовая функция оптимальной системы определяется элементарными операциями, как было изложено в предыдущем параграфе.

Пример 14.1.1. Найти оптимальную систему для выделения полезного сигнала известной формы и неизвестной случайной амплитуды в случае, когда помеха является белым шумом постоянной интенсивности k .

Пусть известная форма полезного сигнала характеризуется функцией $\varphi_1(t)$. Случайную амплитуду полезного сигнала обозначим через U_1 . Тогда задача сведется к решению одного интегрального уравнения (13.6.2). Пользуясь формулой (14.1.4), находим весовую функцию $g^{(1)}$:

$$g^{(1)}(t, \tau) = \frac{\varphi_1(\tau)}{k}.$$

После этого вычисляем по второй формуле (13.6.6) единственную в данном случае величину b_{11} :

$$b_{11} = \int_{t-T}^t g^{(1)}(t, \tau) \varphi_1(\tau) d\tau = \frac{1}{k} \int_{t-T}^t \varphi_1^2(\tau) d\tau.$$

Решение единственного в данном случае уравнения (13.6.7) дает функцию $\lambda_1(t)$:

$$\lambda_1(t) = \frac{\varphi_1(t)}{b_{11} + c_{11}},$$

где $c_{11} = 1/\gamma_{11}$ — величина, обратная по отношению ко второму начальному моменту случайной амплитуды сигнала. Наконец, формула (13.6.4) дает для весовой функции оптимальной системы следующее выражение:

$$g(t, \tau) = \lambda_1(t) g^{(1)}(t, \tau) = \frac{\varphi_1(t) \varphi_1(\tau)}{k(b_{11} + c_{11})}.$$

Выясним подробнее смысл оптимальных операций над входной функцией в данном случае. На основании формулы (13.5.34) выходная переменная

оптимальной системы выражается формулой

$$W^*(t) = \frac{\lambda_1(t)}{k} \left\{ U_1 \int_{t-T}^t \varphi_1^2(\tau) d\tau + \int_{t-T}^t \varphi_1(\tau) X(\tau) d\tau \right\}.$$

Выражению в фигурных скобках получается путем умножения входного сигнала $Z(\tau)$ на известную форму полезного сигнала $\varphi_1(\tau)$ и интегрирования по интервалу наблюдения. При этом полезная часть входного сигнала возводится в квадрат и дает существенно положительную величину, при интегрировании которой происходит как бы накопление полезного сигнала, в то время как шумовая часть при интегрировании не должна накапливаться. К этой идее накопления сигнала путем умножения входной функции на известную форму полезного сигнала и интегрирования легко можно прийти интуитивно, без применения теории оптимальных систем. Однако такое очевидное решение задачи можно получить только при полностью известной форме сигнала. В случае сигнала более сложной структуры найти оптимальную систему интуитивным путем затруднительно.

Простота решения уравнения (14.1.1) в случае, когда $X(t)$ представляет собой белый шум, приводит к следующей идее решения уравнения (14.1.1) в общем случае, когда $X(t)$ не является белым шумом. Если найти линейную систему, преобразующую данную случайную функцию $X(t)$, т. е. сумму нерегулярной части полезного сигнала и помехи, в белый шум $V(t)$, а

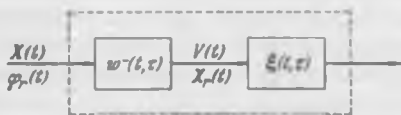


Рис. 14.1.1.

затем найти оптимальную систему для случая белого шума $V(t)$ на входе, то последовательное соединение этих двух систем и будет представлять собой оптимальную линейную систему для случайной функции $X(t)$ на входе (обведена пунктиром на рис. 14.1.1). Эта идея лежит по существу в основе всех известных методов решения интегрального уравнения (14.1.1).

Таким образом, решение интегрального уравнения (14.1.1) сводится к задаче канонического представления случайной функции $X(t)$. Действительно, мы знаем, что проблема канонического представления случайной функции есть проблема ее выражения через белый шум — дискретный в случае канонического разложения и непрерывный в случае интегрального канонического представления. Вследствие этого очевидно, что теоретической основой, на которой может быть построена общая теория оптимальных систем, является метод канонических представлений случайных функций. Далее мы рассмотрим применение обоих типов канонических представлений, когда случайная функция $X(t)$ выражается через дискретный белый шум, т. е. через последовательность некоррелированных импульсов, и когда она выражается через непрерывный белый шум. Применяя интегральное каноническое представление,

мы получаем решение интегрального уравнения (14.1.1) в конечной форме. Однако, к сожалению, пока не существует общих методов получения интегральных канонических представлений для произвольных случайных функций. С другой стороны, любую случайную функцию можно, пользуясь стандартным приемом, представить каноническим разложением. Однако в этом случае решение интегрального уравнения (14.1.1) получается в форме бесконечно-го ряда.

Заметим, что формулой (14.1.4) можно пользоваться как приближенной и в том случае, когда случайная функция $X(t)$ не является белым шумом, но близка к белому шуму. А именно формула (14.1.4) дает приближенное решение уравнения (14.1.1), если интервал корреляции τ_0 случайной функции $X(t)$ (т. е. максимальное значение разности $|\tau - \sigma|$, при котором корреляционная функция $K_x(\tau, \sigma)$ существенно отличается от нуля) достаточно мал для того, чтобы функцию $f(t, \tau)/G(\tau)$ можно было считать практически постоянной в любом интервале изменения τ длины $2\tau_0$. Для доказательства подставим выражение (14.1.4) в левую часть уравнения (14.1.1), оставив пока функцию $G(\tau)$ неопределенной. Тогда, принимая во внимание, что в данном случае при любом значении σ в интервале $t - T + \tau_0 \leq \sigma \leq t - \tau_0$ подынтегральная функция практически равна нулю вне интервала $(\sigma - \tau_0, \sigma + \tau_0)$, и вынося за знак интеграла значение функции $f(t, \tau)/G(\tau)$ в середине σ этого интервала, будем иметь

$$\begin{aligned} \int_{t-T}^t K_x(\tau, \sigma) \frac{f(t, \tau)}{G(\tau)} d\tau &\approx \int_{\sigma-\tau_0}^{\sigma+\tau_0} K_x(\tau, \sigma) \frac{f(t, \tau)}{G(\tau)} d\tau \approx \\ &\approx \frac{f(t, \sigma)}{G(\sigma)} \int_{-\tau_0}^{\tau_0} K_x(\sigma + \eta, \sigma) d\eta. \end{aligned}$$

Отсюда видно, что если определить функцию $G(\sigma)$ формулой

$$G(\sigma) = \int_{-\tau_0}^{\tau_0} K_x(\sigma + \eta, \sigma) d\eta,$$

то формула (14.1.4) даст приближенное решение уравнения (14.1.1).

§ 14.2. Определение оптимальной линейной системы в случае стационарных сигнала и помехи и бесконечного интервала наблюдения

Применим изложенную общую идею для нахождения оптимальной линейной системы в частном случае, когда требуемый выходной сигнал W и входной сигнал Z содержат только нерегулярные части, представляющие собой стационарные и стационарно связанные

случайные функции времени:

$$Z(t) = X(t), \quad W(t) = Y(t), \quad (14.2.1)$$

а интервал наблюдения бесконечен: $T = \infty$. Для нахождения системы, преобразующей стационарную случайную функцию $X(t)$ в белый шум, согласно изложенному в § 7.6, достаточно выразить ее спектральную плотность в виде

$$s_x(\omega) = |\Phi(i\omega)|^2, \quad (14.2.2)$$

где $\Phi(i\omega)$ — функция, все нули и полюсы которой лежат в верхней полуплоскости. Тогда устойчивая стационарная линейная система с передаточной функцией $1/\Phi(s)$ и будет системой, преобразующей случайную функцию $X(t)$ в белый шум $V(t)$ с единичной спектральной плотностью и интенсивностью, равной 2π .

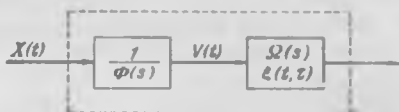


Рис. 14.2.1.

Обозначим через $\xi(t, \tau)$ неизвестную весовую функцию оптимальной системы для случая белого шума V на входе. Тогда оптимальной системой для входной случайной функции X будет последовательное соединение системы с передаточной функцией $1/\Phi(s)$ и системы с весовой функцией $\xi(t, \tau)$ (рис. 14.2.1).

Для определения весовой функции $\xi(t, \tau)$ оптимальной линейной системы для случая белого шума V на входе имеем в данном случае одно уравнение (13.6.1) при $T = \infty$, правая часть которого представляет собой взаимную корреляционную функцию $k_{y\nu}(t - \sigma)$ требуемого выходного сигнала $Y(t)$ системы и белого шума $V(\sigma)$. Применяя для решения этого уравнения формулу (14.1.4), находим

$$\xi(t, \tau) = \frac{1}{2\pi} k_{y\nu}(t - \tau) \quad (-\infty < \tau \leq t). \quad (14.2.3)$$

При $\tau > t$ функцию $\xi(t, \tau)$ следует положить равной нулю, так как мы ищем физически возможную оптимальную систему. Из (14.2.3) видно, что оптимальная линейная система для белого шума V на входе стационарна. Ее передаточная функция $\Omega(s)$ на основании (2.4.6) определяется формулой

$$\Omega(s) = \frac{1}{2\pi} \int_0^{\infty} k_{y\nu}(\xi) e^{-s\xi} d\xi. \quad (14.2.4)$$

Следовательно, и оптимальная система для входной случайной функции $X(t)$, представляющая собой последовательное соединение двух стационарных линейных систем, стационарна. Так как передаточная функция последовательного соединения линейных систем равна произведению передаточных функций соединяемых

систем, то передаточная функция оптимальной линейной системы для входной случайной функции X определяется формулой

$$\Psi(s) = \frac{\Omega(s)}{\Phi(s)}. \quad (14.2.5)$$

Для полного решения задачи осталось теперь только найти взаимную корреляционную функцию $k_{yx}(\xi)$ требуемого выходного сигнала системы Y и белого шума V . Так как белый шум V есть результат преобразования случайной функции X устойчивой стационарной линейной системой с передаточной функцией $1/\Phi(s)$, то для определения взаимной спектральной плотности $s_{yv}(\omega)$ случайной функции Y и белого шума V можно применить формулу (7.4.21). В результате получим

$$s_{yv}(\omega) = \frac{s_{yx}(\omega)}{\Phi(-i\omega)}. \quad (14.2.6)$$

Зная взаимную спектральную плотность случайных функций Y и V , можно найти их взаимную корреляционную функцию по известной формуле теории вероятностей:

$$k_{yv}(\xi) = \int_{-\infty}^{\infty} s_{yv}(\mu) e^{i\mu\xi} d\mu = \int_{-\infty}^{\infty} \frac{s_{yx}(\mu)}{\Phi(-i\mu)} e^{i\mu\xi} d\mu. \quad (14.2.7)$$

Здесь мы обозначили переменную интегрирования через μ , так как определенный интеграл не зависит от обозначения переменной интегрирования.

Подставляя выражение (14.2.7) в (14.2.4) и после этого подставляя полученное выражение в (14.2.5), получим следующую формулу для передаточной функции оптимальной линейной системы:

$$\Psi(s) = \frac{1}{2\pi\Phi(s)} \int_0^{\infty} e^{-s\xi} d\xi \int_{-\infty}^{\infty} \frac{s_{yx}(\mu)}{\Phi(-i\mu)} e^{i\mu\xi} d\mu. \quad (14.2.8)$$

Эта формула является основной формулой теории экстраполяции, интерполяции и сглаживания стационарных случайных функций, разработанной Винером [12].

Двойной интеграл в формуле (14.2.8) во многих случаях легко вычисляется. Так, например, этот интеграл легко вычисляется, если отношение $s_{yx}(\mu)/\Phi(-i\mu)$ является дробно-рациональной функцией. В этом случае, полагая в (14.2.8) $s = i\omega$ и выполняя интегрирование, получим следующую формулу для частотной характеристики оптимальной линейной системы (см. [53], § 129):

$$\Psi(i\omega) = \frac{1}{\Phi(i\omega)} \left[\frac{s_{yx}(\omega)}{\Phi(-i\omega)} \right]_+, \quad (14.2.9)$$

где квадратные скобки со знаком $+$ внизу означают, что рациональная дробь, заключенная в эти скобки, должна быть разложе-

на простые дроби и в полученном разложении должны быть отброшены все простые дроби, соответствующие корням знаменателя, расположенным в нижней полуплоскости комплексной переменной ω . Иными словами, второй множитель в формуле (14.2.9) представляет собой сумму всех простых дробей в разложении на простые дроби рациональной дроби, заключенной в квадратные скобки, которым соответствуют полюсы, расположенные в верхней полуплоскости.

Пример 14.2.1. Найти оптимальную линейную систему (фильтр) для отделения полезного сигнала, представляющего собой стационарную случайную функцию $X_1(t)$ с нулевым математическим ожиданием и корреляционной функцией $k_{x1}(\tau) = De^{-\alpha|\tau|}$, от некоррелированной с ним помехи $X_2(t)$, которая является белым шумом постоянной интенсивности k .

По условиям задачи входной сигнал и требуемый выходной сигнал системы определяются формулами

$$Z(t) = X(t) = X_1(t) + X_2(t), \quad W(t) = Y(t) = X_1(t). \quad (14.2.10)$$

Находим сначала спектральную плотность случайной функции $X(t)$. Так как $X_1(t)$ и $X_2(t)$ не коррелированы, то $s_x(\omega) = s_{x1}(\omega) + s_{x2}(\omega)$. Подставляя сюда выражение спектральной плотности случайной функции $X_1(t)$:

$$s_{x1}(\omega) = \frac{D}{\pi} \cdot \frac{\alpha}{\alpha^2 + \omega^2}$$

и спектральной плотности $k/2\pi$ белого шума $X_2(t)$, получим

$$s_x(\omega) = \frac{D}{\pi} \cdot \frac{\alpha}{\alpha^2 + \omega^2} + \frac{k}{2\pi} = \frac{k}{2\pi} \cdot \frac{\beta^2 + \omega^2}{\alpha^2 + \omega^2}, \quad \beta^2 = \frac{2D\alpha}{k} + \alpha^2.$$

Эту спектральную плотность можно представить в виде квадрата модуля некоторой функции, имеющей нули и полюсы в верхней полуплоскости:

$$s_x(\omega) = \left| \sqrt{\frac{k}{2\pi}} \cdot \frac{\beta + i\omega}{\alpha + i\omega} \right|^2.$$

Следовательно, передаточная функция стационарной линейной системы, формирующей случайную функцию $X(t)$ из белого шума с единичной спектральной плотностью, выражается формулой

$$\Phi(s) = \sqrt{\frac{k}{2\pi}} \cdot \frac{\beta + s}{\alpha + s}. \quad (14.2.11)$$

Величина $1/\Phi(i\omega)$ представляет собой частотную характеристику обратной системы, преобразующей данную случайную функцию $X(t)$ в белый шум.

Для решения задачи нам нужно найти еще взаимную спектральную плотность случайных функций $Y(t)$ и $X(t)$. Учитывая, что случайные функции $X_1(t)$ и $X_2(t)$ не коррелированы, можем написать

$$\begin{aligned} k_{yx}(\tau) &= M[Y(t+\tau)X(t)] = M[X_1(t+\tau)\{X_1(t) + X_2(t)\}] = \\ &= M[X_1(t+\tau)X_1(t)] = k_{x1}(\tau). \end{aligned}$$

Таким образом, случайные функции $Y(t)$ и $X(t)$ в данном случае стационарно связаны и их взаимная корреляционная функция совпадает с корреляционной функцией случайной функции $Y(t) = X_1(t)$. Следовательно, и взаимная спектральная плотность случайных функций $Y(t)$ и $X(t)$

совпадает со спектральной плотностью случайной функции $X_1(t)$:

$$s_{yx}(\omega) = s_{x_1}(\omega) = \frac{D}{\pi} \cdot \frac{\alpha}{\alpha^2 + \omega^2}. \quad (14.2.12)$$

На основании формул (14.2.11) и (14.2.12) имеем

$$\begin{aligned} \frac{s_{yx}(\omega)}{\Phi(-i\omega)} &= D\alpha \sqrt{\frac{2}{k\pi}} \cdot \frac{\alpha - i\omega}{(\beta - i\omega)(\alpha^2 + \omega^2)} = \\ &= D\alpha \sqrt{\frac{2}{k\pi}} \cdot \frac{1}{(-i\omega + \beta)(i\omega + \alpha)} = D\alpha \sqrt{\frac{2}{k\pi}} \cdot \frac{1}{(\omega + i\beta)(\omega - i\alpha)}. \end{aligned}$$

Разложив это выражение на простые дроби, получим

$$\frac{s_{yx}(\omega)}{\Phi(-i\omega)} = \frac{D\alpha}{i(\alpha + \beta)} \sqrt{\frac{2}{k\pi}} \left(\frac{1}{\omega - i\alpha} - \frac{1}{\omega + i\beta} \right).$$

В полученном разложении полюс первой дроби $i\alpha$ лежит в верхней полуплоскости, а полюс второй дроби $-i\beta$ — в нижней полуплоскости. Поэтому вторую дробь следует отбросить. В результате получим

$$\left[\frac{s_{yx}(\omega)}{\Phi(-i\omega)} \right]_+ = \frac{D\alpha}{i(\alpha + \beta)} \sqrt{\frac{2}{k\pi}} \cdot \frac{1}{\omega - i\alpha}.$$

Подставляя это выражение и выражение (14.2.11) в формулу (14.2.9), находим частотную характеристику оптимальной линейной системы:

$$\begin{aligned} \Psi(i\omega) &= \frac{D\alpha}{i(\alpha + \beta)} \sqrt{\frac{2}{k\pi}} \cdot \frac{1}{\omega - i\alpha} \sqrt{\frac{2\pi}{k}} \cdot \frac{\omega - i\alpha}{\omega - i\beta} = \\ &= \frac{2D\alpha}{ik(\alpha + \beta)} \cdot \frac{1}{\omega - i\beta} = \frac{2D\alpha}{k(\alpha + \beta)} \cdot \frac{1}{i\omega + \beta}. \end{aligned}$$

Но $2D\alpha/k = \beta^2 - \alpha^2$. Подставив это выражение в предыдущую формулу, получим окончательно

$$\Psi(i\omega) = \frac{\beta - \alpha}{i\omega + \beta}.$$

Эта формула показывает, что в данном случае оптимальный фильтр представляет собой апериодическое звено с постоянной времени $T = 1/\beta$ и коэффициентом усиления $(\beta - \alpha)/\beta$. Следовательно, в данном случае оптимальный фильтр можно точно реализовать в виде последовательного соединения усилителя и цепочки RC.

Заметим, что точная реализации оптимального фильтра оказалась в рассмотренном примере возможной только благодаря тому, что помеха является белым шумом, а полезный сигнал не содержит компонентов в виде белого шума. В таких случаях формула (14.2.9) всегда выражает передаточную функцию оптимальной системы в виде дробно-рациональной функции, у которой степень числителя меньше степени знаменателя. Такие передаточные функции в принципе всегда можно точно реализовать при помощи сопротивлений, емкостей и индуктивностей. Если помеха не является белым шумом и полезный сигнал не содержит компоненты, представляющей собой белый шум, то степень числителя передаточной функции оптимальной системы в общем случае будет по меньшей мере равна степени знаменателя, вследствие чего при реализации оптимальной системы возникают серьезные трудности. Однако, как мы уже знаем, в точной реализации оптимальных систем нет практической необходимости (см. пример 13.3.1).

§ 14.3. Общая формула для определения оптимальной линейной системы

Предположим, что мы умеем найти физически возможную линейную систему, весовую функцию которой обозначим $w^-(t, \tau)$, преобразующую данную случайную функцию $X(t)$ в белый шум $V(t)$. Предположим, что нам известна весовая функция обратной системы $w(t, \tau)$, формирующей случайную функцию $X(t)$ из белого шума $V(t)$. При этом мы вначале рассмотрим случай бесконечно-го интервала наблюдения $T = \infty$. В следующем параграфе мы покажем, как можно применить полученные результаты к случаю конечного интервала наблюдения.

В данном случае мы можем написать следующие равенства:

$$V(t) = \int_{-\infty}^t w^-(t, \tau) X(\tau) d\tau, \quad (14.3.1)$$

$$X(t) = \int_{-\infty}^t w(t, \tau) V(\tau) d\tau. \quad (14.3.2)$$

Теперь заметим, что на входе искомой оптимальной системы обычно действует не только сумма нерегулярной части полезного сигнала и помехи $X(t)$, но также и регулярная часть полезного сигнала, т. е. сумма известных функций $\varphi_r(t)$, умноженных на случайные коэффициенты U_r . Эти функции $\varphi_r(t)$ также будут преобразовываться системой, которая преобразует $X(t)$ в белый шум. Результат этого преобразования функций $\varphi_1(t), \dots, \varphi_N(t)$ мы обозначим $\chi_1(t), \dots, \chi_N(t)$ (рис. 14.1.1). Тогда будем иметь следующие соотношения:

$$\chi_r(t) = \int_{-\infty}^t w^-(t, \tau) \varphi_r(\tau) d\tau \quad (r = 1, \dots, N). \quad (14.3.3)$$

$$\varphi_r(t) = \int_{-\infty}^t w(t, \tau) \chi_r(\tau) d\tau \quad (r = 1, \dots, N). \quad (14.3.4)$$

Таким образом, решая задачу определения оптимальной системы для белого шума на входе, мы должны вместо функций $\varphi_r(t)$ взять функции $\chi_r(t)$. Кроме того, вместо взаимной корреляционной функции $K_{yx}(t, \sigma)$ мы должны взять взаимную корреляционную функцию случайной функции $Y(t)$ и белого шума $V(t)$:

$$K_{yv}(t, \sigma) = M[Y(t)V(\sigma)] =$$

$$= M\left[Y(t) \int_{-\infty}^{\sigma} w^-(\sigma, \tau) X(\tau) d\tau\right] = \int_{-\infty}^{\sigma} w^-(\sigma, \tau) K_{yx}(t, \tau) d\tau. \quad (14.3.5)$$

Точно так же из (14.3.2) находим

$$\begin{aligned} K_{yx}(t, \tau) &= M[Y(t)X(\tau)] = M\left[Y(t) \int_{-\infty}^{\tau} w(\tau, \sigma) V(\sigma) d\sigma\right] = \\ &= \int_{-\infty}^{\tau} w(\tau, \sigma) K_{vv}(t, \sigma) d\sigma. \end{aligned} \quad (14.3.6)$$

Сравнивая формулы (14.3.3) и (14.3.5), мы приходим к выводу, что при преобразовании случайной функции $X(t)$ в белый шум $V(t)$ необходимо подвергнуть такому же преобразованию правую часть интегрального уравнения (14.1.1), т. е. вместо функции $f(t, \sigma)$ нужно взять функцию

$$f_1(t, \sigma) = \int_{-\infty}^{\sigma} w^-(\sigma, \tau) f(t, \tau) d\tau. \quad (14.3.7)$$

Функции $f(t, \tau)$ и $f_1(t, \sigma)$ связаны, как показывают формулы (14.3.4) и (14.3.6), и обратным преобразованием:

$$f(t, \tau) = \int_{-\infty}^{\tau} w(\tau, \sigma) f_1(t, \sigma) d\sigma. \quad (14.3.8)$$

Это необходимо учитывать, как мы увидим в дальнейшем, при решении уравнения (14.1.1) в случае конечного интервала наблюдения.

Обозначим весовую функцию оптимальной линейной системы для белого шума $V(t)$ на входе через $\xi(t, \tau)$. Пользуясь формулой (14.1.4), находим

$$\xi(t, \tau) = \frac{f_1(t, \tau)}{G(\tau)} \quad (-\infty < \tau \leq t). \quad (14.3.9)$$

Искомая оптимальная система представляет собой последовательное соединение линейных систем с весовыми функциями $w^-(t, \tau)$ и $\xi(t, \tau)$ (рис. 14.1.1). Поэтому, пользуясь формулой (4.2.5), получим

$$g(t, \tau) = \int_{\tau}^t \xi(t, \sigma) w^-(\sigma, \tau) d\sigma. \quad (14.3.10)$$

Таким образом, последовательность формул (14.3.7), (14.3.9) и (14.3.10) дает решение поставленной задачи. Эти три формулы можно объединить. Для этого подставим выражение (14.3.7) в (14.3.9) и полученное таким образом выражение подставим

в (14.3.10). В результате получим

$$g(t, \tau) = \int_{\tau}^t \frac{w^{-}(\sigma, \tau) d\sigma}{G(\sigma)} \int_{-\infty}^{\sigma} w^{-}(\sigma, \lambda) f(t, \lambda) d\lambda. \quad (14.3.11)$$

Полагая в (14.3.11) последовательно $f(t, \lambda) = K_{yx}(t, \lambda)$, $\varphi_1(\lambda)$, ..., $\varphi_N(\lambda)$, мы найдем функции $g^{(0)}(t, \tau)$, $g^{(1)}(t, \tau)$, ..., $g^{(N)}(t, \tau)$, после чего весовая функция оптимальной системы $g(t, \tau)$ находится, как показано в § 13.6.

Формулу (14.3.11) можно, конечно, получить чисто формальным путем, пользуясь интегральным каноническим представлением случайной функции $X(t)$. Однако изложенное решение более наглядно и вскрывает физический смысл применения метода интегральных канонических представлений в данном случае.

Подставляя полученное решение (14.3.11) в интегральное уравнение (14.1.1) при $T = \infty$ и принимая во внимание, что при любом $t > \sigma, \tau$

$$\begin{aligned} K_x(\sigma, \tau) &= \int_{-\infty}^t G(u) w(\sigma, u) w(\tau, u) du, \\ \int_{-\infty}^t w(t, \sigma) w^{-}(\sigma, \tau) d\sigma &= \delta(t - \tau), \\ \int_{-\infty}^t w^{-}(t, \sigma) w(\sigma, \tau) d\sigma &= \delta(t - \tau) \end{aligned}$$

(см. формулы (7.6.13), (7.6.7) и (7.6.8)), можно убедиться в том, что выражение (14.3.11) действительно удовлетворяет интегральному уравнению (14.1.1) при $T = \infty$:

$$\int_{-\infty}^t K_x(\tau, \sigma) g(t, \tau) d\tau = f(t, \sigma) \quad (-\infty < \sigma \leq t). \quad (14.3.12)$$

Применим формулу (14.3.11) к частному случаю стационарной случайной функции X с дробно-рациональной спектральной плотностью. Определив в этом случае передаточную функцию устойчивой стационарной линейной системы, преобразующей случайную функцию X в белый шум, путем представления спектральной плотности формулой (14.2.2), можем найти ее весовую функцию по формуле (7.6.27) или (7.6.30). Имея в виду, что в данном случае система, преобразующая случайную функцию X в белый шум V , стационарна, пользуясь правом произвольно расширять пределы интегрирования в случае нулевой подынтегральной функции и учитывая, что интенсивность белого шума $V(t)$ равна в данном случае

$G(t) = 2\pi$, перепишем формулу (14.3.11) в виде

$$g(t, \tau) = \frac{1}{2\pi} \int_{-\infty}^t w^-(\sigma - \tau) d\sigma \int_{-\infty}^{\infty} w^-(\sigma - \lambda) f(t, \lambda) d\lambda. \quad (14.3.13)$$

После замены переменных $\eta = t - \lambda$, $\xi = t - \sigma$ эта формула примет вид

$$g(t, \tau) = \frac{1}{2\pi} \int_0^{\infty} w^-(t - \tau - \xi) d\xi \int_{-\infty}^{\infty} w^-(\eta - \xi) f(t, t - \eta) d\eta. \quad (14.3.14)$$

Выразив весовую функцию $w^-(\eta - \xi)$ через соответствующую частотную характеристику $1/\Phi(i\omega)$ (см. формулу (7.6.30)):

$$w^-(\eta - \xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{-i\mu(\eta - \xi)}}{\Phi(-i\mu)} d\mu,$$

можем написать

$$\begin{aligned} \int_{-\infty}^{\infty} w^-(\eta - \xi) f(t, t - \eta) d\eta &= \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \frac{f(t, t - \eta) e^{-i\mu(\eta - \xi)}}{\Phi(-i\mu)} d\mu d\eta. \end{aligned} \quad (14.3.15)$$

Изменим в полученном интеграле порядок интегрирования и обратим внимание на то, что интеграл

$$\chi(t, \mu) = \frac{1}{2\pi} \int_{-\infty}^{\infty} f(t, t - \eta) e^{-i\mu\eta} d\eta \quad (14.3.16)$$

представляет собой преобразование Фурье функции $f(t, t - \eta)$, рассматриваемой как функция η . Тогда получим

$$\int_{-\infty}^{\infty} w^-(\eta - \xi) f(t, t - \eta) d\eta = \int_{-\infty}^{\infty} \frac{\chi(t, \mu) e^{i\mu\xi}}{\Phi(-i\mu)} d\mu. \quad (14.3.17)$$

Подставив это выражение в (14.3.14) и заменив весовую функцию $w^-(t - \tau - \xi)$ ее выражением через частотную характеристику $1/\Phi(i\omega)$ (формула (7.6.27)):

$$w^-(t - \tau - \xi) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{e^{i\omega(t - \tau - \xi)}}{\Phi(i\omega)} d\omega,$$

получим

$$g(t, \tau) = \frac{1}{4\pi^2} \int_0^{\infty} d\xi \int_{-\infty}^{\infty} \frac{e^{i\omega(t - \tau - \xi)}}{\Phi(i\omega)} d\omega \int_{-\infty}^{\infty} \frac{\chi(t, \mu) e^{i\mu\xi}}{\Phi(-i\mu)} d\mu. \quad (14.3.18)$$

Наконец, изменив порядок интегрирования по переменным ξ и ω , получим

$$g(t, \tau) = \frac{1}{4\pi^2} \int_{-\infty}^{\infty} \frac{e^{i\omega(t-\tau)}}{\Phi(i\omega)} d\omega \int_0^{\infty} e^{-i\omega\xi} d\xi \int_{-\infty}^{\infty} \frac{\chi(t, \mu) e^{i\mu\xi}}{\Phi(-i\mu)} d\mu. \quad (14.3.19)$$

Сравним формулу (14.3.19) с формулой (2.3.20), выражающей весовую функцию линейной системы через ее частотную характеристику. Тогда получим следующую формулу для частотной характеристики системы, весовая функция которой определяется уравнением (14.3.12):

$$z(t, i\omega) = \frac{1}{2\pi\Phi(i\omega)} \int_0^{\infty} e^{-i\omega\xi} d\xi \int_{-\infty}^{\infty} \frac{\chi(t, \mu) e^{i\mu\xi}}{\Phi(-i\mu)} d\mu. \quad (14.3.20)$$

Эта формула показывает, что в общем случае, несмотря на стационарность входного случайного возмущения X , оптимальная система нестационарна, так как ее частотная характеристика зависит от времени.

Заметим теперь, что для вычисления функции χ по формуле (14.3.16) необходимо задать функцию $f(t, \lambda)$, рассматриваемую как функция λ , на всей числовой оси, в то время как для пахождения оптимальной системы достаточно задать ее в интервале $-\infty < \lambda \leq t$. Это кажущееся противоречие не возникает при применении формулы (14.3.11), из которой была выведена формула (14.3.20). Дело в том, что формальное расширение пределов интегрирования по λ в (14.3.13) позволяет продолжить функцию $f(t, \lambda)$ в область $t < \lambda < \infty$ произвольно, так как $w^-(\sigma - \lambda) = 0$ при $\lambda > \sigma$, а $\sigma \leq t$ в области интегрирования по σ . Например, можно принять $f(t, \lambda) = 0$ при $\lambda > t$. Однако при этом могут получиться разрывы функции $f(t, \lambda)$ и ее производных по λ в точке $\lambda = t$. А так как весовая функция $w^-(\sigma - \lambda)$ обычно содержит линейную комбинацию δ -функции и ее производных, то внутренний интеграл в (14.3.13) содержит линейную комбинацию производных функции $f(t, \lambda)$ по λ . Вследствие дифференцирования разрывной функции внутренний интеграл в (14.3.13) будет содержать лишние члены с δ -функцией и ее производными. Эти лишние члены возникают только в результате произвольно допускаемого разрыва функции $f(t, \lambda)$ и ее производных по λ в точке $\lambda = t$. Их можно избежать, если продолжить функцию $f(t, \lambda)$ в область $\lambda > t$ таким образом, чтобы она была непрерывна вместе со своими производными по λ до соответствующего порядка. Поэтому все лишние члены с δ -функцией и ее производными, полученные при вычислении внутреннего интеграла в формуле (14.3.20), равного на основании

(14.3.17) внутреннему интегралу в (14.3.13) и (14.3.14), следует отбросить. Но члены с $\delta(\xi)$, $\delta'(\xi)$, $\delta''(\xi)$, ... во внутреннем интеграле дают в выражении двойного интеграла в (14.3.20) соответственно 1, $i\omega$, $(i\omega)^2$, ..., т. е. полином относительно ω . Поэтому при вычислении двойного интеграла в формуле (14.3.20) целую часть, т. е. полином относительно ω , всегда следует отбрасывать, так как она может возникнуть только вследствие разрывов функции $f(t, \lambda)$ и ее производных по λ , вводимых при произвольном ее продолжении в область $\lambda > t$. Это обстоятельство всегда следует учитывать при вычислении интеграла в формуле (14.3.20).

В частном случае, когда правая часть интегрального уравнения (14.3.12) зависит только от разности аргументов:

$$f(t, \sigma) = h(t - \sigma), \quad (14.3.21)$$

функция

$$\chi(t, \mu) = \frac{1}{2\pi} \int_{-\infty}^{\infty} h(\eta) e^{-i\mu\eta} d\eta = \psi(\mu) \quad (14.3.22)$$

не зависит от времени. Поэтому правая часть формулы (14.3.20) не зависит от времени, и, следовательно, оптимальная линейная система стационарна. Формула (14.3.20) дает в этом случае следующую формулу для определения частотной характеристики стационарной оптимальной линейной системы:

$$\Psi(i\omega) = \frac{1}{2\pi\Phi(i\omega)} \int_0^{\infty} e^{-i\omega\xi} d\xi \int_{-\infty}^{\infty} \frac{\psi(\mu)}{\Phi(-i\mu)} e^{i\mu\xi} d\mu. \quad (14.3.23)$$

При $\psi(\mu) = s_{yx}(\mu)$ эта формула совпадает с формулой Винера (14.2.8), выведенной в предыдущем параграфе. Таким образом, формула Винера (14.2.8) является частным случаем формулы (14.3.20), которая в свою очередь является следствием общей формулы (14.3.11). Формула (14.3.20) была впервые получена Бутоном для случая $f(t, \sigma) = K_{yx}(t, \sigma)$.

Двойные интегралы в формулах (14.3.20) и (14.3.23) вычисляются совершенно так же, как интеграл в (14.2.8). Если отношение $\psi(\mu)/\Phi(-i\mu)$ является дробно-рациональной функцией, то формула (14.3.23) принимает вид

$$\Psi(i\omega) = \frac{1}{\Phi(i\omega)} \left[\frac{\psi(\omega)}{\Phi(-i\omega)} \right]_+, \quad (14.3.24)$$

где символ $[]_+$ означает, что рациональная дробь, заключенная в квадратные скобки, должна быть представлена в виде суммы полинома (целой части) и простых дробей, соответствующих всем

корням знаменателя, и в полученном разложении должны быть отброшены целая часть и все простые дроби, соответствующие корням знаменателя, расположенным в нижней полуплоскости комплексной переменной ω .

§ 14.4. Определение оптимальной линейной системы в случае стационарной помехи и конечного интервала наблюдения

Формулу (14.3.11) можно в некоторых случаях применить и для определения оптимальной линейной системы в случае конечного интервала наблюдения. Легко понять, что если интегральное уравнение с бесконечным нижним пределом

$$\int_{-\infty}^t K_{\kappa}(\tau, \sigma) g(t, \tau) d\tau = f(t, \sigma) \quad (-\infty < \sigma \leq t) \quad (14.4.1)$$

имеет решение $g(t, \tau)$, равное нулю при $\tau < t - T$:

$$g(t, \tau) = 0 \quad \text{при } \tau < t - T, \quad (14.4.2)$$

то это решение является также решением интегрального уравнения с конечными пределами

$$\int_{t-T}^t K_{\kappa}(\tau, \sigma) g(t, \tau) d\tau = f(t, \sigma) \quad (-\infty < \sigma \leq t). \quad (14.4.3)$$

Следовательно, для решения уравнения (14.4.3) достаточно найти такое решение уравнения (14.4.1), которое равно нулю при всех $\tau < t - T$. Однако уравнение (14.4.1) далеко не всегда, не при всяком виде функции $f(t, \sigma)$, имеет такое решение. Чтобы обеспечить существование такого решения уравнения (14.4.1), воспользуемся тем обстоятельством, что в случае конечного интервала наблюдения T достаточно, чтобы уравнение (14.4.3) удовлетворялось только для значений σ в интервале наблюдения $t - T \leq \sigma \leq t$. Иными словами, правая часть интегрального уравнения (14.4.3) должна быть задана только в интервале наблюдения $t - T \leq \sigma \leq t$. Ее значения в этом интервале полностью определяют интересующее нас решение уравнения (14.4.3). Поэтому значениями функции $f(t, \sigma)$ при $\sigma < t - T$ можно распорядиться совершенно произвольно. Этим обстоятельством можно воспользоваться и выбрать функцию $f(t, \sigma)$ при $\sigma < t - T$ таким образом, чтобы обеспечить существование решения уравнения (14.4.1), удовлетворяющего условию (14.4.2). Это решение будет удовлетворять и уравнению (14.4.3) при всех $\sigma \leq t$. В частности, оно будет удовлетворять уравнению (14.4.3) при $t - T \leq \sigma \leq t$, т. е. будет удовлетворять интересующему нас уравнению (14.1.1).

Таким образом, задача определения оптимальной линейной системы в случае конечного интервала наблюдения сводится к такому продолжению заданной правой части $f(t, \sigma)$ интегрального уравнения (14.1.1) в область $\sigma < t - T$, чтобы обеспечить выполнение условия (14.4.2). После этого решение интегрального уравнения (14.1.1) определяется по общей формуле (14.3.11). Такова общая идея метода решения интегрального уравнения (14.1.1) в случае конечного интервала наблюдения T .

Найдем решение интегрального уравнения (14.1.1), основываясь на изложенной общей идее, ограничиваясь частным случаем, когда случайная функция $X(t)$ стационарна и имеет дробно-рациональную спектральную плотность. Мы видели в § 7.6, что любую дробно-рациональную спектральную плотность можно представить в виде

$$s_X(\omega) = \frac{1}{2\pi} \frac{H(i\omega) H(-i\omega)}{F(i\omega) F(-i\omega)} \\ \left(= \frac{1}{2\pi} \left| \frac{H(i\omega)}{F(i\omega)} \right|^2 \text{ при действительных } \omega \right). \quad (14.4.4)$$

Представление спектральной плотности в таком виде дает основание трактовать случайную функцию $X(t)$ как результат прохождения стационарного белого шума со спектральной плотностью $1/2\pi$ и, следовательно, интенсивностью, равной единице, через устойчивую стационарную линейную систему с передаточной функцией $\Phi(s) = H(s)/F(s)$. Этой передаточной функции соответствует линейное дифференциальное уравнение с постоянными коэффициентами, связывающее случайную функцию $X(t)$ с белым шумом $V(t)$,

$$F(D) X = H(D) V. \quad (14.4.5)$$

Учитывая, что система, преобразующая случайную функцию $X(t)$ в белый шум $V(t)$, в данном случае стационарна и интенсивность белого шума $V(t)$ тождественно равна единице, перепишем цепочку формул (14.3.7), (14.3.9) и (14.3.10) в виде

$$g(t, \tau) = \int_{-\infty}^t \xi(t, \sigma) w^-(\sigma - \tau) d\sigma, \quad (14.4.6)$$

$$\xi(t, \tau) = \int_{-\infty}^{\tau} w^-(\tau - \lambda) f(t, \lambda) d\lambda. \quad (14.4.7)$$

Мы расширили интервал интегрирования в формуле (14.4.6), имея в виду, что $w^-(\sigma - \tau) = 0$ при $\sigma < \tau$.

Для определения весовой функции w^- воспользуемся формулой (4.4.26), которая в данном случае дает

$$w^-(t - \tau) = F\left(-\frac{d}{d\tau}\right) p(t - \tau), \quad (14.4.8)$$

где $p(t - \tau)$ — весовая функция физически возможной системы, определяемая линейным дифференциальным уравнением

$$H\left(\frac{d}{dt}\right)p(t - \tau) = \delta(t - \tau). \quad (14.4.9)$$

Предположим теперь, что полином $H(\mu)$ имеет только простые корни, которые мы обозначим через $\mu_1, \dots, \mu_m$. Все эти корни на основании изложенного в § 7.6 имеют отрицательные действительные части. В этом случае для определения весовой функции $p(t - \tau)$ можно применить формулу (4.4.49). В результате получим

$$p(t - \tau) = \sum_{q=1}^m \frac{e^{\mu_q(t-\tau)}}{H'(\mu_q)} \quad (\tau < t). \quad (14.4.10)$$

Подставляя выражение (14.4.8) в (14.4.6) и изменяя порядок операций дифференцирования и интегрирования, получим

$$\begin{aligned} g(t, \tau) &= \int_{-\infty}^t \xi(t, \sigma) F\left(-\frac{d}{d\tau}\right) p(\sigma - \tau) d\sigma = \\ &= F\left(-\frac{d}{d\tau}\right) \int_{-\infty}^t p(\sigma - \tau) \xi(t, \sigma) d\sigma. \end{aligned} \quad (14.4.11)$$

Вводя промежуточную весовую функцию

$$\eta(t, \tau) = \int_{-\infty}^t p(\sigma - \tau) \xi(t, \sigma) d\sigma = \int_{\tau}^t p(\sigma - \tau) \xi(t, \sigma) d\sigma, \quad (14.4.12)$$

перепишем формулу (14.4.11) в виде

$$g(t, \tau) = F\left(-\frac{d}{d\tau}\right) \eta(t, \tau). \quad (14.4.13)$$

Таким образом, решение интегрального уравнения (14.4.1) в данном случае определяется последовательностью формул (14.4.7), (14.4.12) и (14.4.13).

Чтобы применить формулы (14.4.7), (14.4.12) и (14.4.13) к случаю конечного интервала наблюдения, мы должны согласно изложенной общей идее определить функцию $f(t, \tau)$ при $\tau < t - T$ таким образом, чтобы удовлетворить условию (14.4.2). На основании формулы (14.4.13) условие (14.4.2) может быть переписано в виде

$$F\left(-\frac{d}{d\tau}\right) \eta(t, \tau) = 0 \quad (\tau < t - T). \quad (14.4.14)$$

Таким образом, весовая функция $\eta(t, \tau)$ определяется при $\tau < t - T$ линейным дифференциальным уравнением с постоянными коэффициентами (14.4.14).

Как известно из теории дифференциальных уравнений, общий интеграл линейного дифференциального уравнения (14.4.14) выражается формулой

$$\eta(t, \tau) = c_1 e^{k_1 \tau} + c_2 e^{k_2 \tau} + \dots + c_n e^{k_n \tau}, \quad (14.4.15)$$

где $k_1, \dots, k_n$ — корни характеристического уравнения

$$F(-k) = 0, \quad (14.4.16)$$

а $c_1, \dots, c_n$ — произвольные постоянные. Обозначим корни полинома $F(v)$ через $v_1, \dots, v_n$ и для простоты предположим, что все они различны. Кроме того, ни один из корней $v_1, \dots, v_n$ не совпадает ни с одним из корней $\mu_1, \dots, \mu_m$ полинома $\Pi(\mu)$, так как полиномы $F(s)$ и $\Pi(s)$ в формуле (14.4.4) не имеют общих множителей (если бы числитель и знаменатель выражения (14.4.4) имели общий множитель, то на него можно было бы дробь сократить). Наконец, на основании изложенного в § 7.6 все корни $v_1, \dots, v_n$ полинома $F(v)$ имеют отрицательные действительные части. Очевидно, что корни характеристического уравнения (14.4.16) равны корням полинома $F(v)$, взятым с обратным знаком:

$$k_1 = -v_1, \quad k_2 = -v_2, \quad \dots, \quad k_n = -v_n.$$

Подставляя эти значения в (14.4.15) получим следующую формулу, определяющую функцию $\eta(t, \tau)$ в области $\tau < t - T$:

$$\eta(t, \tau) = c_1 e^{-v_1 \tau} + c_2 e^{-v_2 \tau} + \dots + c_n e^{-v_n \tau} \quad (\tau < t - T). \quad (14.4.17)$$

Для определения функции $\xi(t, \tau)$ при $\tau < t - T$ воспользуемся формулой (14.4.12). На основании изложенного в §§ 4.3 и 4.4 формула (14.4.12) определяет $\eta(t, \tau)$ как интеграл дифференциального уравнения, сопряженного с (14.4.9):

$$H\left(-\frac{d}{d\tau}\right) \eta(t, \tau) = \xi(t, \tau). \quad (14.4.18)$$

Подставляя сюда выражение (14.4.17) функции $\eta(t, \tau)$ и учитывая, что при дифференцировании показательной функции $e^{k\tau}$ оператор дифференцирования $d/d\tau$ везде заменяется множителем k , получим

$$\xi(t, \tau) = c_1 \Pi(v_1) e^{-v_1 \tau} + c_2 \Pi(v_2) e^{-v_2 \tau} + \dots + c_n \Pi(v_n) e^{-v_n \tau} \quad (\tau < t - T). \quad (14.4.19)$$

Наконец, для определения функции $f(t, \tau)$ при $\tau < t - T$ заметим, что согласно общей теории § 4.4 формула (14.4.7) опре-

делает функцию $\xi(t, \tau)$ как интеграл линейного дифференциального уравнения с постоянными коэффициентами

$$H\left(\frac{d}{d\tau}\right)\xi(t, \tau) = F\left(\frac{d}{d\tau}\right)f(t, \tau). \quad (14.4.20)$$

Это уравнение является также дифференциальным относительно функции $f(t, \tau)$ и может быть использовано для определения функции $f(t, \tau)$, если известна функция $\xi(t, \tau)$. Подставляя в уравнение (14.4.20) выражение (14.4.19) функции $\xi(t, \tau)$, получим следующее дифференциальное уравнение, определяющее функцию $f(t, \tau)$ при $\tau < t - T$:

$$F\left(\frac{d}{d\tau}\right)f(t, \tau) = \sum_{r=1}^n c_r H(\nu_r) H(-\nu_r) e^{-\nu_r \tau}. \quad (14.4.21)$$

При этом вследствие формул (14.3.8) и (14.3.9) функция $f(t, \tau)$ представляет собой такой интеграл уравнения (14.4.20), который тождественно равен нулю, если функция $\xi(t, \tau)$ тождественно равна нулю, т. е. частный интеграл уравнения (14.4.20) или, что то же, уравнения (14.4.21), не содержащий свободных колебаний. Согласно общей теории линейных дифференциальных уравнений с постоянными коэффициентами, такой частный интеграл уравнения (14.4.21) следует искать в форме

$$f(t, \tau) = \sum_{r=1}^n A_r e^{-\nu_r \tau}, \quad (14.4.22)$$

где $A_1, \dots, A_n$ — неопределенные коэффициенты. Для определения этих коэффициентов следует подставить выражение (14.4.22) в уравнение (14.4.21) и сравнить коэффициенты при одинаковых показательных функциях в левой и правой частях полученного равенства. Тогда получим

$$\sum_{r=1}^n A_r F(-\nu_r) e^{-\nu_r \tau} = \sum_{r=1}^n c_r H(\nu_r) H(-\nu_r) e^{-\nu_r \tau},$$

$$A_r F(-\nu_r) = c_r H(\nu_r) H(-\nu_r),$$

откуда

$$A_r = c_r \frac{H(\nu_r) H(-\nu_r)}{F(-\nu_r)} \quad (r = 1, \dots, n). \quad (14.4.23)$$

Таким образом, функцию $f(t, \tau)$ для значений $\tau < t - T$ необходимо определить формулой

$$f(t, \tau) = \sum_{r=1}^n c_r \frac{H(\nu_r) H(-\nu_r)}{F(-\nu_r)} e^{-\nu_r \tau} \quad (\tau < t - T). \quad (14.4.24)$$

Это выражение содержит n произвольных постоянных $c_1, \dots, c_n$, и пока мы еще не имеем никаких данных для их определения. Поэтому оставим их пока неопределенными и будем считать функцию $f(t, \tau)$ заданной в интервале $-\infty < \tau < t - T$ формулой (14.4.24), а в интервале $t - T \leq \tau \leq t$ — совпадающей с известной правой частью интегрального уравнения (14.1.1).

Перейдем теперь к последовательному определению функций ξ , η и g в интервале $t - T \leq \tau \leq t$, считая функцию f известной во всем интервале $-\infty < \tau \leq t$. Разбивая при $t - T \leq \tau \leq t$ интервал интегрирования в формуле (14.4.7) на две части $(-\infty, t - T)$ и $(t - T, \tau)$ и подставляя в первый интеграл выражение (14.4.24) функции $f(t, \tau)$ при $\tau < t - T$, получим

$$\begin{aligned} \xi(t, \tau) = \sum_{r=1}^n c_r \frac{H(\nu_r) H(-\nu_r)}{F(-\nu_r)} \int_{-\infty}^{t-T} w^-(\tau - \lambda) e^{-\nu_r \lambda} d\lambda + \\ + \int_{t-T}^{\tau} w^-(\tau - \lambda) f(t, \lambda) d\lambda. \end{aligned} \quad (14.4.25)$$

Заметим теперь, что весовая функция $w^-(\tau - \lambda)$ при $\lambda < \tau$ может быть определена по формуле (4.4.50):

$$w^-(\tau - \lambda) = \sum_{q=1}^m \frac{F(\mu_q)}{H'(\mu_q)} e^{\mu_q(\tau - \lambda)} \quad (\lambda < \tau). \quad (14.4.26)$$

Этот же результат получается, если выражение (14.4.10) подставить в формулу (14.4.8) и изменить соответственным образом обозначение аргументов. На основании формулы (14.4.26) можем написать

$$\begin{aligned} \int_{-\infty}^{t-T} w^-(\tau - \lambda) e^{-\nu_r \lambda} d\lambda &= \sum_{q=1}^m \frac{F(\mu_q)}{H'(\mu_q)} e^{\mu_q \tau} \int_{-\infty}^{t-T} e^{-(\nu_r + \mu_q)\lambda} d\lambda = \\ &= -e^{-\nu_r(t-T)} \sum_{q=1}^m \frac{F(\mu_q)}{(\nu_r + \mu_q) H'(\mu_q)} e^{\mu_q(\tau - t + T)}. \end{aligned} \quad (14.4.27)$$

Подставляя выражение (14.4.8) в последний интеграл в формуле (14.4.25), приведем его к виду

$$\int_{t-T}^{\tau} w^-(\tau - \lambda) f(t, \lambda) d\lambda = \int_{t-T}^{\tau} f(t, \lambda) F\left(-\frac{d}{d\lambda}\right) p(\tau - \lambda) d\lambda.$$

Для дальнейшего преобразования этого интеграла воспользуемся формулой Грина

$$\int_{t_1}^{t_2} (vFu - uF^*v) dt = \left[\sum_{h=0}^{n-1} u^{(h)} \sum_{k=0}^{n-h-1} \sum_{l=0}^h (-1)^k C_{h+k+1}^l a_{h+k+1}^{(h-l)} v^{(l)} \right]_{t_1}^{t_2},$$

справедливой для любого линейного дифференциального оператора

$$F = \sum_{k=0}^n a_k D^k,$$

соответствующего сопряженного оператора

$$F^* = \sum_{k=0}^n (-1)^k D^k (a_k \cdot)$$

и любых функций u и v , производные которых могут иметь разрывы первого рода в некоторых точках интервала интегрирования *). Применяя формулу Грина, учитывая, что для линейных дифференциальных операторов с постоянными коэффициентами

$$F^* \left(\frac{d}{d\lambda} \right) = F \left(-\frac{d}{d\lambda} \right),$$

и принимая во внимание, что весовая функция $p(\tau - \lambda)$ равна тождественно нулю при $\lambda > \tau$, получим

$$\begin{aligned} \int_{t-T}^{\tau} \omega^-(\tau - \lambda) f(t, \lambda) d\lambda &= \int_{t-T}^{\tau} p(\tau - \lambda) F \left(\frac{d}{d\lambda} \right) f(t, \lambda) d\lambda + \\ &+ \sum_{k=0}^{n-1} f_{\tau}^{(k)}(t, t-T) \sum_{h=0}^{n-k-1} a_{k+h+1} p^{(h)}(\tau - t + T). \end{aligned}$$

Подставляя сюда выражение (14.4.10) весовой функции p , будем иметь

$$\begin{aligned} \int_{t-T}^{\tau} \omega^-(\tau - \lambda) f(t, \lambda) d\lambda &= \\ &= \sum_{q=1}^m \frac{e^{\mu_q \tau}}{H'(\mu_q)} \left\{ \int_{t-T}^{\tau} e^{-\mu_q \lambda} F \left(\frac{d}{d\lambda} \right) f(t, \lambda) d\lambda + \right. \\ &\quad \left. + e^{-\mu_q(t-T)} \sum_{k=0}^{n-1} f_{\tau}^{(k)}(t, t-T) \sum_{h=0}^{n-k-1} a_{k+h+1} \mu_q^h \right\}. \quad (14.4.28) \end{aligned}$$

*) Читатель может найти вывод формулы Грина, например, в [53], § 84.

Дальнейшее вычисление этого интеграла в общем виде невозможно. Можно, однако, заметить, что этот интеграл легко выражается через элементарные функции во всех случаях, когда функция $f(t, \tau)$ представляет собой полином относительно τ , показательную функцию, или тригонометрическую функцию, или сумму произведений показательных и тригонометрических функций на полиномы.

Подставляя выражения (14.4.27) и (14.4.28) в формулу (14.4.25), получим

$$\begin{aligned} \xi(t, \tau) = & - \sum_{r=1}^n c_r \frac{H(v_r) H(-v_r)}{F(-v_r)} e^{-v_r(t-T)} \sum_{q=1}^m \frac{F(\mu_q)}{(v_r + \mu_q) H'(\mu_q)} e^{\mu_q(\tau-t+T)} + \\ & + \sum_{q=1}^m \frac{e^{\mu_q \tau}}{H'(\mu_q)} \left\{ \int_{t-T}^{\tau} e^{-\mu_q \lambda} F\left(\frac{d}{d\lambda}\right) f(t, \lambda) d\lambda + \right. \\ & \left. + e^{-\mu_q(t-T)} \sum_{h=0}^{n-1} f_{\tau}^{(h)}(t, t-T) \sum_{h=0}^{n-h-1} a_{h+h+1} \mu_q^h \right\}. \quad (14.4.29) \end{aligned}$$

Вводя для краткости обозначения

$$b_r = c_r \frac{H(v_r) H(-v_r)}{F(-v_r)} e^{-v_r(t-T)} \quad (r = 1, \dots, n), \quad (14.4.30)$$

$$\begin{aligned} \xi_0(t, \tau) = & \sum_{q=1}^m \frac{e^{\mu_q \tau}}{H'(\mu_q)} \left\{ \int_{t-T}^{\tau} e^{-\mu_q \lambda} F\left(\frac{d}{d\lambda}\right) f(t, \lambda) d\lambda + \right. \\ & \left. + e^{-\mu_q(t-T)} \sum_{h=0}^{n-1} f_{\tau}^{(h)}(t, t-T) \sum_{h=0}^{n-h-1} a_{h+h+1} \mu_q^h \right\}, \quad (14.4.31) \end{aligned}$$

можем представить формулу (14.4.29) в более компактном виде

$$\xi(t, \tau) = - \sum_{r=1}^n b_r \sum_{q=1}^m \frac{F(\mu_q)}{(v_r + \mu_q) H'(\mu_q)} e^{\mu_q(\tau-t+T)} + \xi_0(t, \tau). \quad (14.4.32)$$

Переходим к определению функции $\eta(t, \tau)$ в интервале $t - T \leq \tau \leq t$. Для этого подставим выражение (14.4.32) в формулу (14.4.12). Тогда получим

$$\begin{aligned} \eta(t, \tau) = & - \sum_{r=1}^n b_r \sum_{q=1}^m \frac{F(\mu_q)}{(v_r + \mu_q) H'(\mu_q)} e^{-\mu_q(t-T)} \int_{\tau}^t p(\sigma - \tau) e^{\mu_q \sigma} d\sigma + \\ & + \int_{\tau}^t p(\sigma - \tau) \xi_0(t, \sigma) d\sigma. \quad (14.4.33) \end{aligned}$$

Выразив весовую функцию $p(\sigma - \tau)$ формулой (14.4.10), заменив в ней предварительно индекс суммирования q на p , можем написать

$$\int_{\tau}^t p(\sigma - \tau) e^{\mu_q \sigma} d\sigma = \sum_{p=1}^m \frac{e^{-\mu_p \tau}}{H'(\mu_p)} \int_{\tau}^t e^{(\mu_p + \mu_q) \sigma} d\sigma = \\ = - \sum_{p=1}^m \frac{e^{\mu_q \tau} - e^{-\mu_p \tau + (\mu_p + \mu_q) t}}{(\mu_p + \mu_q) H'(\mu_p)}.$$

Подставляя это выражение и выражение (14.4.10) весовой функции $p(\sigma - \tau)$ в (14.4.33), получим

$$\eta(t, \tau) = \sum_{r=1}^n b_r \sum_{p, q=1}^m \frac{F(\mu_q) e^{\mu_q \tau} [e^{-\mu_q(t-\tau)} - e^{\mu_p(t-\tau)}]}{(\nu_r + \mu_q)(\mu_p + \mu_q) H'(\mu_p) H'(\mu_q)} + \\ + \sum_{q=1}^m \frac{e^{\mu_q \tau}}{H'(\mu_q)} \int_{\tau}^t e^{\mu_q \sigma} \xi_0(t, \sigma) d\sigma. \quad (14.4.34)$$

Последний интеграл легко выражается через элементарные функции во всех случаях, когда функция $f(t, \tau)$ представляет собой полином относительно τ , или показательную функцию, или тригонометрическую функцию, или сумму произведений показательных и тригонометрических функций на полиномы. В подобных случаях и функция $\xi_0(t, \tau)$, как показывает формула (14.4.31), имеет такую же форму.

Итак, мы определили функцию $\eta(t, \tau)$ внутри интервала $t - T < \tau < t$. Однако в выражении этой функции остались неизвестными величины $b_1, \dots, b_n$, так как в определяющих их формулах (14.4.30) неизвестны величины $c_1, \dots, c_n$. Для нахождения этих величин воспользуемся доказанным в § 13.5 положением, что для определения весовой функции оптимальной системы необходимо не любое решение интегрального уравнения (14.1.1), а только такое его решение, которое содержит производные δ -функции не выше p -го порядка, если p — наибольшее целое число, для кото-

рого производная $\frac{\partial^2 p \Gamma_x(\tau, \sigma)}{\partial \tau^p \partial \sigma^p}$, а следовательно на основании формулы (13.5.18) и производная корреляционной функции $\frac{\partial^2 p K_x(\tau, \sigma)}{\partial \tau^p \partial \sigma^p}$, непрерывна. Но $\frac{\partial^2 p K_x(\tau, \sigma)}{\partial \tau^p \partial \sigma^p}$ является корреляционной функцией производной $X^{(p)}(t)$ случайной функции $X(t)$. Для определения p вернемся к уравнению (14.4.5). Это уравнение содержит производные белого шума $V(t)$ до порядка m включительно и производные случайной функции $X(t)$ до порядка n включительно, причем $n > m$. Следовательно, случайная функция $X(t)$ содержит результат интегрирования белого шума $V(t)$ по

меньшей мере $n - m$ раз и ее производная порядка $n - m$ содержит в виде слагаемого белый шум $V(t)$. Поэтому производная

$$\frac{\partial^{2n-2m} K_x(\tau, \sigma)}{\partial \tau^{n-m} \partial \sigma^{n-m}} = (-1)^{n-m} k_x^{(2n-2m)}(\tau - \sigma)$$

содержит слагаемое типа $\delta(\tau - \sigma)$. Но при этом производная $k_x^{(2n-2m-1)}(\tau - \sigma)$ имеет разрыв первого рода при $\sigma = \tau$, а производная

$$\frac{\partial^{2n-2m-2} K_x(\tau, \sigma)}{\partial \tau^{n-m-1} \partial \sigma^{n-m-1}} = (-1)^{n-m-1} k_x^{(2n-2m-2)}(\tau - \sigma)$$

непрерывна при $\sigma = \tau$. Таким образом, в данном случае $p = n - m - 1$ и весовая функция $g(t, \tau)$ может содержать δ -функцию и ее производные порядка не выше $n - m - 1$.

Формула (14.4.13) показывает, что выражение весовой функции $g(t, \tau)$ содержит производные функции $\eta(t, \tau)$ по τ до порядка n включительно. Следовательно, n -я производная функции $\eta(t, \tau)$ может содержать производные δ -функции порядка не выше $n - m - 1$. Отсюда следует, что $n - k$ -я производная функции $\eta(t, \tau)$ может содержать производные δ -функции порядка не выше $n - m - k - 1$. А так как $n - m - k - 1 \geq 0$ только при $k < n - m$, то m -я производная функции $\eta(t, \tau)$ может иметь только разрыв первого рода. Таким образом, функция $\eta(t, \tau)$ должна быть непрерывна вместе со своими производными $\eta_\tau^{(l)}(t, \tau), \dots, \eta_\tau^{(m-1)}(t, \tau)$ при $\tau \leq t$. Это условие автоматически выполняется в точке $\tau = t$ в силу формулы (14.4.12) и вытекающих из (4.4.15) условий

$$p(0) = p'(0) = \dots = p^{(m-1)}(0) = 0, \quad (14.4.35)$$

которым должна удовлетворить весовая функция $p(\tau - \sigma)$. Для того чтобы удовлетворить условиям непрерывности функций $\eta(t, \tau), \eta_\tau^{(1)}(t, \tau), \dots, \eta_\tau^{(m-1)}(t, \tau)$ в точке $\tau = t - T$, необходимо, чтобы значения этих функций в точке $\tau = t - T$, вычисленные по формулам (14.4.17) и (14.4.34), совпадали. Из этого условия, исключая величины $c_1, \dots, c_n$ при помощи формул (14.4.30), получаем m уравнений

$$\sum_{r=1}^n b_r \left\{ \sum_{p,q=1}^m \frac{F(\mu_q) [\mu_q^k - (-\mu_p)^k e^{(\mu_p + \mu_q)T}]}{(\nu_r + \mu_q)(\mu_p + \mu_q) H'(\mu_p) H'(\mu_q)} - \frac{(-\nu_r)^k F(-\nu_r)}{H(\nu_r) H(-\nu_r)} \right\} =$$

$$= -\eta_{0\tau}^{(k)}(t, t-T) \quad (k=0, 1, \dots, m-1), \quad (14.4.36)$$

где для краткости введена функция

$$\eta_0(t, \tau) = \sum_{p=1}^m \frac{e^{-\mu_p \tau}}{H'(\mu_p)} \int_{\tau}^t e^{\mu_p \sigma} \xi_0(t, \sigma) d\sigma. \quad (14.4.37)$$

Чтобы найти недостающие $n - m$ уравнений для определения величин $b_1, \dots, b_n$, обратимся к уравнениям (14.4.18) и (14.4.20). В уравнение (14.4.18) входят производные функции $\eta(t, \tau)$ до порядка m включительно. А так как функция $\eta(t, \tau)$ непрерывна вместе со своими производными до порядка $m - 1$ включительно, а ее производная $\eta_t^{(m)}(t, \tau)$ может иметь лишь разрывы первого рода, то функция $\xi(t, \tau)$ может иметь самое большее разрыв первого рода и не может содержать δ -функцию и ее производные. Но в уравнение (14.4.20) входят производные функции $\xi(t, \tau)$ до порядка m включительно. Следовательно, так как $\xi_t(t, \tau)$ может содержать δ -функцию, $\xi_t^{(m)}(t, \tau)$ — первую производную δ -функции и т. д., $\xi_t^{(m)}(t, \tau)$ — $(m - 1)$ -ю производную δ -функции, то левая часть уравнения (14.4.20) не может содержать производных δ -функции выше $m - 1$ -го порядка. Но в уравнение (14.4.20) входят производные функции $f(t, \tau)$ до порядка n включительно. Таким образом, n -я производная $f_t^{(n)}(t, \tau)$ может содержать производные δ -функции до порядка $m - 1$ включительно и вообще $n - k$ -я производная $f_t^{(n-k)}(t, \tau)$ — производные δ -функции не выше $m - k - 1$ -го порядка. А так как $m - k - 1 \geq 0$ только при $k < m$, то производная $f_t^{(n-m)}(t, \tau)$ уже не может содержать δ -функций, а может иметь только разрывы первого рода. Следовательно, функция $f(t, \tau)$ должна быть непрерывна вместе со своими производными по τ до порядка $n - m - 1$ включительно. Это дает еще $n - m$ уравнений для определения величин $b_1, \dots, b_n$. Для получения этих уравнений достаточно приравнять значения функций $f(t, \tau), f_t'(t, \tau), \dots, f_t^{(n-m-1)}(t, \tau)$ в точке $\tau = t - T$, вычисленные по формуле (14.4.24), значениям этих же величин для заданной в интервале $t - T \leq \tau \leq t$ функции $f(t, \tau)$. В результате, принимая во внимание (14.4.30), получим

$$\sum_{r=1}^n (-v_r)^k b_r = f_t^{(k)}(t, t - T) \quad (k = 0, 1, \dots, n - m - 1). \quad (14.4.38)$$

Система n линейных алгебраических уравнений (14.4.36) и (14.4.38) полностью определяет n величин $b_1, \dots, b_n$.

Из предыдущего рассуждения на первый взгляд следует вывод, что непрерывность функций $f(t, \tau), f_t'(t, \tau), \dots, f_t^{(n-m-1)}(t, \tau)$ влечет за собой и непрерывность функций $\eta(t, \tau), \eta_t'(t, \tau), \dots, \eta_t^{(m-1)}(t, \tau)$. Это действительно так, функция $\eta(t, \tau)$, определяемая формулами (14.4.7) и (14.4.12), непрерывна вместе со своими производными по τ до порядка $m - 1$ включительно, если функции $f(t, \tau), f_t'(t, \tau), \dots, f_t^{(n-m-1)}(t, \tau)$ непрерывны. Однако в общем случае она не совпадает в интервале $-\infty < \tau \leq t - T$ с функцией $\eta(t, \tau)$, определяемой формулой (14.4.17). Поэтому, пользуясь формулой (14.4.17) для вычисления $\eta(t, t - T), \eta_t'(t, t - T), \dots, \eta_t^{(m-1)}(t, t - T)$ и требуя совпадения этих

величин с соответствующими величинами, вычисленными при помощи формулы (14.4.34), мы получаем условия, независимые от условий непрерывности функций $f(t, \tau)$, $f'_\tau(t, \tau)$, ..., $f^{(n-m-1)}_\tau(t, \tau)$. Можно показать, что эти новые условия представляют собой не что иное, как условия совпадения функции $\eta(t, \tau)$, определяемой формулами (14.4.7) и (14.4.12), с функцией $\eta(t, \tau)$ определяемой формулой (14.4.17), в интервале $-\infty < \tau \leq t - T$ ([53], § 130).

Для определения весовой функции $g(t, \tau)$, удовлетворяющей интегральному уравнению (14.4.3), теперь остается только подставить выражение (14.4.34) функции $\eta(t, \tau)$ в интервале $t - T < \tau < t$ в формулу (14.4.13). При этом разрывы производных $\eta^{(m)}_\tau$, ..., $\eta^{(n)}_\tau$ в точках $\tau = t - T$ и $\tau = t$ дадут линейную комбинацию δ -функций $\delta(\tau - t + T)$ и $\delta(\tau - t)$ и их производных до порядка $n - m - 1$ включительно. В результате, принимая во внимание обозначение (14.4.37), перепишем формулу (14.4.13) в виде

$$g(t, \tau) = F\left(-\frac{d}{d\tau}\right) \eta_0(t, \tau) + \\ + \sum_{r=1}^n b_r \sum_{p, q=1}^m \frac{F(\mu_q) e^{\mu_q T} [F(-\mu_q) e^{-\mu_q(t-\tau)} - F(\mu_p) e^{\mu_p(t-\tau)}]}{(\nu_r + \mu_q)(\mu_p + \mu_q) H'(\mu_p) H'(\mu_q)} + \\ + \sum_{k=0}^{n-m-1} [A_k \delta^{(k)}(\tau - t + T) + B_k \delta^{(k)}(\tau - t)]. \quad (14.4.39)$$

Теперь остается только определить коэффициенты $A_0, A_1, \dots, A_{n-m-1}, B_0, B_1, \dots, B_{n-m-1}$ при δ -функциях. Для определения этих коэффициентов заметим, что δ -функция получается в выражении $\eta^{(m+1)}_\tau$ в результате разрыва производной $\eta^{(m)}_\tau$, в выражении $\eta^{(m+2)}_\tau$ — в результате разрыва производной $\eta^{(m+1)}_\tau$ и т. д., в выражении $\eta^{(n)}_\tau$ — в результате разрыва производной $\eta^{(n-1)}_\tau$. А так как

$$F\left(-\frac{d}{d\tau}\right) \eta(t, \tau) = \sum_{h=0}^n (-1)^h a_h \eta^{(h)}_\tau(t, \tau), \quad (14.4.40)$$

то коэффициент при $\delta(\tau - t + T)$ в выражении весовой функции $g(t, \tau)$ получается равным

$$A_0 = (-1)^{m+1} a_{m+1} \Delta_0 \eta^{(m)} + (-1)^{m+2} a_{m+2} \Delta_0 \eta^{(m+1)} + \dots \\ \dots + (-1)^n a_n \Delta_0 \eta^{(n-1)}, \quad (14.4.41)$$

где через $\Delta_0 \eta^{(l)}$ обозначен разрыв производной $\eta^{(l)}_\tau(t, \tau)$ в точке $\tau = t - T$. Для вычисления $\Delta_0 \eta^{(l)}$ продифференцируем выражения (14.4.17) и (14.4.34) l раз и положим после дифференцирова-

ния $\tau = t - T$, после чего вычтем первый результат из второго. Тогда, принимая во внимание (14.4.30), получим

$$\begin{aligned} \Delta_0 \eta^{(l)} &= \eta_{\tau}^{(l)}(t, t - T + 0) - \eta_{\tau}^{(l)}(t, t - T - 0) = \\ &= \sum_{r=1}^n b_r \left\{ \sum_{p, q=1}^m \frac{F(\mu_q) [\mu_q^l - (-\mu_p)^l] e^{(\mu_p + \mu_q)T}}{(v_r + \mu_q)(\mu_p + \mu_q) H'(\mu_p) H'(\mu_q)} - \frac{(-v_r)^l F(-v_r)}{H(v_r) H(-v_r)} \right\} + \\ &\quad + \eta_{0\tau}^{(l)}(t, t - T) \quad (l = m, \dots, n-1). \end{aligned} \quad (14.4.42)$$

Совершенно так же вычисляются коэффициенты при производных δ -функции. Замечая, что производная $\delta^{(k)}$ получается в выражении $\eta_{\tau}^{(m+k+1)}$ в результате разрыва производной $\eta_{\tau}^{(m)}$, в выражении $\eta_{\tau}^{(m+k+2)}$ — в результате разрыва производной $\eta_{\tau}^{(m+1)}$ и т. д., в выражении $\eta_{\tau}^{(n)}$ — в результате разрыва производной $\eta_{\tau}^{(n-k-1)}$, получаем на основании (14.4.40):

$$A_k = (-1)^{m+k+1} a_{m+k+1} \Delta_0 \eta^{(m)} + \dots + (-1)^n a_n \Delta_0 \eta^{(n-k-1)} \quad (k = 1, \dots, n - m - 1). \quad (14.4.43)$$

Формулы (14.4.41) и (14.4.43) могут быть написаны в виде

$$A_k = \sum_{l=m}^{n-k-1} (-1)^{l+k+1} a_{l+k+1} \Delta_0 \eta^{(l)} \quad (k = 0, 1, \dots, n - m - 1). \quad (14.4.44)$$

Аналогичную формулу получим для коэффициентов $B_0, B_1, \dots, B_{n-m-1}$:

$$B_k = \sum_{l=m}^{n-k-1} (-1)^{l+k+1} a_{l+k+1} \Delta_1 \eta^{(l)} \quad (k = 0, 1, \dots, n - m - 1), \quad (14.4.45)$$

где $\Delta_1 \eta^{(l)}$ — разрыв производной $\eta_{\tau}^{(l)}(t, \tau)$ в точке $\tau = t$. Для вычисления $\Delta_1 \eta^{(l)}$ заметим, что $\eta(t, \tau)$, как весовая функция физически возможной системы, равна нулю при $\tau > t$. Следовательно, пользуясь формулами (14.4.34) и (14.4.37), получим

$$\begin{aligned} \Delta_1 \eta^{(l)} &= - \sum_{r=1}^n b_r \sum_{p, q=1}^m \frac{F(\mu_q) e^{\mu_q T} [\mu_q^l - (-\mu_p)^l]}{(v_r + \mu_q)(\mu_p + \mu_q) H'(\mu_p) H'(\mu_q)} - \eta_{0\tau}^{(l)}(t, t) \\ &\quad (l = m, \dots, n-1). \end{aligned} \quad (14.4.46)$$

Итак, для нахождения решения интегрального уравнения (14.1.1) в случае стационарной случайной функции $X(t)$ на входе в конечного интервала наблюдения T необходимо представить дробно-рациональную спектральную плотность стационарной случайной функции $X(t)$ в форме (14.4.4), найти корни полиномов $F(v)$ и $H(\mu)$ (которые, согласно изложенному в § 7.6 методу определения этих полиномов, все лежат в левой полуплоскости),

определить величины $b_1, \dots, b_n$ путем решения системы линейных алгебраических уравнений (14.4.36) и (14.4.38), определить функцию $\eta_0(t, \tau)$ по формулам (14.4.31) и (14.4.37) и вычислить коэффициенты $A_0, A_1, \dots, A_{n-m-1}, B_0, B_1, \dots, B_{n-m-1}$ по формулам (14.4.42), (14.4.44), (14.4.45) и (14.4.46). После этого решение интегрального уравнения (14.1.1) определится по формуле (14.4.39).

Проанализируем подробнее структуру решения (14.4.39) интегрального уравнения (14.1.1) в рассматриваемом случае. Изменяя порядок суммирования, можем представить формулу (14.4.39) в виде

$$g(t, \tau) = F\left(-\frac{d}{d\tau}\right) \eta_0(t, \tau) + \\ + \sum_{q=1}^m \left\{ \sum_{r=1}^n b_r \sum_{p=1}^m \frac{F(\mu_q) F(-\mu_q) e^{\mu_q T}}{(v_r + \mu_q)(\mu_p + \mu_q) H'(\mu_p) H'(\mu_q)} \right\} e^{-\mu_q(t-\tau)} - \\ - \sum_{p=1}^m \left\{ \sum_{r=1}^n b_r \sum_{q=1}^m \frac{F(\mu_q) F(\mu_p) e^{\mu_q T}}{(v_r + \mu_q)(\mu_p + \mu_q) H'(\mu_p) H'(\mu_q)} \right\} e^{\mu_p(t-\tau)} + \\ + \sum_{h=0}^{n-m-1} [A_h \delta^{(h)}(\tau - t + T) + B_h \delta^{(h)}(\tau - t)].$$

Заменяя в первой сумме индексы суммирования q на h , p на l , а во второй, наоборот, p на h , q на l , получим

$$g(t, \tau) = F\left(-\frac{d}{d\tau}\right) \eta_0(t, \tau) + \\ + \sum_{h=1}^m \left\{ \sum_{r=1}^n b_r \sum_{l=1}^m \frac{F(\mu_h) F(-\mu_h) e^{\mu_h T}}{(v_r + \mu_h)(\mu_h + \mu_l) H'(\mu_h) H'(\mu_l)} \right\} e^{-\mu_h(t-\tau)} - \\ - \sum_{h=1}^m \left\{ \sum_{r=1}^n b_r \sum_{l=1}^m \frac{F(\mu_h) F(\mu_l) e^{\mu_l T}}{(v_r + \mu_l)(\mu_h + \mu_l) H'(\mu_h) H'(\mu_l)} \right\} e^{\mu_h(t-\tau)} + \\ + \sum_{h=0}^{n-m-1} [A_h \delta^{(h)}(\tau - t + T) + B_h \delta^{(h)}(\tau - t)].$$

Полагая для краткости

$$\begin{aligned} x_h &= \sum_{r=1}^n b_r \sum_{l=1}^m \frac{F(\mu_h) F(-\mu_h) e^{\mu_h T}}{(v_r + \mu_h)(\mu_h + \mu_l) H'(\mu_h) H'(\mu_l)}, \\ x_{m+h} &= - \sum_{r=1}^n b_r \sum_{l=1}^m \frac{F(\mu_h) F(\mu_l) e^{\mu_l T}}{(v_r + \mu_l)(\mu_h + \mu_l) H'(\mu_h) H'(\mu_l)} \end{aligned} \quad \left. \begin{array}{l} \\ \\ \end{array} \right\} (14.4.47)$$

($h = 1, \dots, m$),

ПОЛУЧАЕМ

$$g(t, \tau) = F\left(-\frac{d}{d\tau}\right) \eta_0(t, \tau) + \sum_{h=1}^m [\chi_h e^{-\mu_h(t-\tau)} + \chi_{m+h} e^{\mu_h(t-\tau)}] + \\ + \sum_{k=0}^{n-m-1} [A_k \delta^{(k)}(\tau - t + T) + B_k \delta^{(k)}(\tau - t)]. \quad (14.4.48)$$

Таким образом, решение интегрального уравнения (14.1.1) в рассматриваемом случае состоит из линейной комбинации показательных функций $e^{\mu_h(t-\tau)}$ и $e^{-\mu_h(t-\tau)}$, соответствующих всем корням полинома $H(\mu)$ (по предположению различным), линейной комбинации δ -функций и их производных до порядка $n - m - 1$ включительно с особенностями на концах интервала наблюдения $t = T$ и t и линейной комбинации функций, вид которых определяется формой правой части $f(t, \tau)$ интегрального уравнения (14.1.1).

Исследуем теперь структуру выходного сигнала оптимальной системы. Для этого заметим, что, поскольку решения всех интегральных уравнений (13.6.1) и (13.6.2) выражаются формулами вида (14.4.48), весовая функция оптимальной линейной системы вследствие (13.6.4) также выражается формулой вида (14.4.48). Поэтому выходной сигнал оптимальной системы

$$W^*(t) = AZ(t) = \\ = \int_{t-T}^t \left\{ F\left(-\frac{d}{d\tau}\right) \eta_0(t, \tau) + \sum_{h=1}^m [\chi_h e^{-\mu_h(t-\tau)} + \chi_{m+h} e^{\mu_h(t-\tau)}] \right\} Z(\tau) d\tau + \\ + A_0 Z(t-T) + B_0 Z(t) - A_1 Z'(t-T) - \\ - B_1 Z'(t) + \dots + (-1)^{n-m-1} A_{n-m-1} Z^{(n-m-1)}(t-T) + \\ + (-1)^{n-m-1} B_{n-m-1} Z^{(n-m-1)}(t). \quad (14.4.49)$$

Знание аналитической формы решения интегрального уравнения (14.1.1) в рассматриваемом случае позволяет применить иной прием для нахождения этого решения. А именно можно сразу написать решение в форме (14.4.48) с неопределенными коэффициентами $\chi_1, \dots, \chi_{2m}, A_0, A_1, \dots, A_{n-m-1}, B_0, B_1, \dots, B_{n-m-1}$ и коэффициентами при известных функциях в выражении $F\left(-\frac{d}{d\tau}\right) \eta_0(t, \tau)$, подставить выражение (14.4.48) в уравнение (14.1.1) и потребовать, чтобы оно обратилось в тождество. В результате получим систему линейных алгебраических уравнений для определения неизвестных коэффициентов. После решения этой системы уравнений формула (14.4.48) полностью определит решение интегрального уравнения (14.1.1). Такой метод неопределенных коэффициентов был предложен для решения интегрального уравнения (14.1.1) в рассматриваемом случае для функции f , представляющей собой полином, Заде и Рагадзини, которые

впервые папли аналитическую форму (14.4.48) решения [25] (см. также [53], § 132).

Заметим, что при решении интегрального уравнения (14.1.1) в рассматриваемом случае методом неопределенных коэффициентов число линейных алгебраических уравнений, которые приходится решать для определения неизвестных коэффициентов, всегда больше, чем $2m + 2(n - m) = 2n$. Поэтому метод неопределенных коэффициентов практически целесообразно применять только при небольших значениях n . При больших n целесообразнее применить изложенный выше метод, который требует только решения n линейных алгебраических уравнений (14.4.36) и (14.4.38) и дает для коэффициентов $\kappa_1, \dots, \kappa_{2m}$ формулы (14.4.47), а для коэффициентов $A_0, A_1, \dots, A_{n-m-1}, B_0, B_1, \dots, B_{n-m-1}$ — формулы (14.4.44) и (14.4.45).

В частном случае, когда степень $2m$ числителя спектральной плотности равна нулю, т. е. полином $\Pi(\mu)$ является постоянной, которая без ущерба для общности может быть принята равной единице, формулы (14.4.18) и (14.4.20) дают

$$\xi(t, \tau) = \eta(t, \tau) = F\left(\frac{d}{d\tau}\right) f(t, \tau). \quad (14.4.50)$$

Подставим это выражение в (14.4.13). При этом необходимо учесть, что функция $f(t, \tau)$ непрерывна вместе со своими производными по τ до $n - 1$ -го порядка, а ее производная n -го порядка может иметь разрыв первого рода. В точке $\tau = t$ функция $\xi(t, \tau)$ согласно изложенному может иметь разрыв первого рода, который даст разрыв первого рода производной $f_{\tau}^{(n)}(t, \tau)$. Вследствие этих разрывов появятся δ -функции $\delta(\tau - t + T)$ и $\delta(\tau - t)$ в $f_{\tau}^{(n+1)}$, $\delta'(\tau - t + T)$ и $\delta'(\tau - t)$ в $f_{\tau}^{(n+2)}(t, \tau)$ и т. д., $\delta^{(n-1)}(\tau - t + T)$ и $\delta^{(n-1)}(\tau - t)$ в $f_{\tau}^{(2n)}(t, \tau)$. В результате получим для решения интегрального уравнения (14.1.1) следующую формулу:

$$g(t, \tau) = F\left(-\frac{d}{d\tau}\right) F\left(\frac{d}{d\tau}\right) f(t, \tau) + \sum_{h=0}^{n-1} [A_h \delta^{(h)}(\tau - t + T) + B_h \delta^{(h)}(\tau - t)]. \quad (14.4.51)$$

Коэффициенты при δ -функциях здесь определяются формулами (14.4.44) и (14.4.45) при $m = 0$, в которых разрывы функции η и ее производных на основании (14.4.17), (14.4.30) и (14.4.50) определяются формулами

$$\left. \begin{aligned} \Delta_0 \eta^{(l)} &= \frac{d^l}{d\tau^l} F\left(\frac{d}{d\tau}\right) f(t, t - T) - \sum_{r=1}^n b_r (-v_r)^l F(-v_r), \\ \Delta_l \eta^{(l)} &= -\frac{d^l}{d\tau^l} F\left(\frac{d}{d\tau}\right) f(t, t), \quad (l = 0, 1, \dots, n-1), \end{aligned} \right\} \quad (14.4.52)$$

где для краткости положено

$$\left. \begin{aligned} \frac{d^l}{d\tau^l} F\left(\frac{d}{d\tau}\right) f(t, t-T) &= \left[\frac{d^l}{d\tau^l} F\left(\frac{d}{d\tau}\right) f(t, \tau) \right]_{\tau=t-T}, \\ \frac{d^l}{d\tau^l} F\left(\frac{d}{d\tau}\right) f(t, t) &= \left[\frac{d^l}{d\tau^l} F\left(\frac{d}{d\tau}\right) f(t, \tau) \right]_{\tau=t}. \end{aligned} \right\} \quad (14.4.53)$$

Коэффициенты $b_1, \dots, b_n$ определяются в данном случае системой n линейных алгебраических уравнений (14.4.38) при $m = 0$.

Таким образом, при $m = 0$ решение интегрального уравнения (14.1.1) сводится к решению системы линейных алгебраических уравнений (14.4.38), определению разрывов функции $\eta(t, \tau)$ и ее производных по τ в точках $\tau = t - T$ и $\tau = t$ по формулам (14.4.52) и (14.4.53), вычислению коэффициентов при δ -функциях по формулам (14.4.44), (14.4.45) при $m = 0$ и определению весовой функции $g(t, \tau)$ по формуле (14.4.51).

Все предыдущие результаты получены в предположении, что полиномы $F(v)$ и $H(\mu)$ имеют только простые корни. Однако это предположение несущественно. Совершенно так же, применяя известные правила интегрирования линейных дифференциальных уравнений с постоянными коэффициентами в случае кратных корней характеристического уравнения, можно найти решение интегрального уравнения (14.1.1) в случае кратных корней полиномов $F(v)$ и $H(\mu)$. При этом все вычисления доводятся до конца в элементарных функциях.

Пример 14.4.1. Найти решение интегрального уравнения

$$D \int_{t-T}^t e^{-\alpha|\tau-t|} g(t, \tau) d\tau = a_0 + a_1 \sigma \quad (t-T \leq \sigma \leq t). \quad (14.4.54)$$

Спектральная плотность стационарной случайной функции $X(t)$ в данном случае может быть представлена в виде

$$s_x(\omega) = \frac{D}{\pi} \frac{\alpha}{\alpha^2 + \omega^2} = \frac{1}{2\pi} \left| \frac{\sqrt{2D\alpha}}{\alpha + i\omega} \right|^2.$$

Отсюда видно, что в данном случае можно принять

$$H(\mu) = 1, \quad F(v) = (\alpha + v)/\sqrt{2D\alpha}.$$

Полином $F(v)$ имеет единственный корень $v_1 = -\alpha$. Так как в данном случае полином H отсутствует, то следует применить формулу (14.4.51). Для этого предварительно находим скачки функции η по формулам (14.4.52):

$$\begin{aligned} \Delta_0 \eta &= \left[F\left(\frac{d}{d\tau}\right) (a_0 + a_1 \tau) \right]_{\tau=t-T} - b_1 F(-v_1) = \\ &= \frac{1}{\sqrt{2D\alpha}} \left[\left(\alpha + \frac{d}{d\tau} \right) (a_0 + a_1 \tau) \right]_{\tau=t-T} - \sqrt{\frac{2\alpha}{D}} b_1 = \\ &= \frac{\alpha a_0 + a_1 + \alpha a_1 (t-T)}{\sqrt{2D\alpha}} - \sqrt{\frac{2\alpha}{D}} b_1, \\ \Delta_1 \eta &= - \left[F\left(\frac{d}{d\tau}\right) (a_0 + a_1 \tau) \right]_{\tau=t} = - \frac{\alpha a_0 + a_1 + \alpha a_1 t}{\sqrt{2D\alpha}}. \end{aligned}$$

Для определения коэффициента b_1 имеем в данном случае одно уравнение (14.4.38), которое имеет вид

$$b_1 = f(t, t - T) = a_0 + a_1(t - T).$$

Подставляя это выражение в формулу для $\Delta_0\eta$, получим

$$\Delta_0\eta = \frac{a_1 - \alpha a_0 - \alpha a_1(t - T)}{\sqrt{2D\alpha}}.$$

Подставляя полученные выражения $\Delta_0\eta$ и $\Delta_1\eta$ в формулы (14.4.44) и (14.4.45), находим коэффициенты при δ -функциях

$$A_0 = -\frac{a_1}{2D\alpha} + \frac{a_0 + a_1(t - T)}{2D}, \quad B_0 = \frac{a_1}{2D\alpha} + \frac{a_0 + a_1t}{2D}.$$

Далее имеем

$$\begin{aligned} F\left(-\frac{d}{d\tau}\right) F\left(\frac{d}{d\tau}\right) f(t, \tau) &= \frac{\left(\alpha - \frac{d}{d\tau}\right) \left(\alpha + \frac{d}{d\tau}\right)}{2D\alpha} (a_0 + a_1\tau) = \\ &= \left(\frac{\alpha}{2D} - \frac{1}{2D\alpha} \frac{d^2}{d\tau^2}\right) (a_0 + a_1\tau) = \frac{\alpha}{2D} (a_0 + a_1\tau). \end{aligned}$$

Подставляя полученные выражения в (14.4.51), найдем решение интегрального уравнения (14.4.54):

$$\begin{aligned} g(t, \tau) &= \frac{\alpha}{2D} (a_0 + a_1\tau) + \frac{1}{2D} \left[a_0 + a_1(t - T) - \frac{a_1}{\alpha} \right] \delta(\tau - t + T) + \\ &+ \frac{1}{2D} \left(a_0 + a_1t + \frac{a_1}{\alpha} \right) \delta(\tau - t). \end{aligned} \quad (14.4.55)$$

Для иллюстрации характера продолжения функции $f(t, \tau)$ в область $\tau < t - T$ в данном случае найдем функцию $f(t, \tau)$ при $\tau < t - T$, хотя для решения интегрального уравнения нам это и не понадобилось. Формулы (14.4.24) и (14.4.30) в данном случае дают

$$\begin{aligned} f(t, \tau) &= b_1 e^{-\nu_1(\tau - t + T)} = \\ &= [a_0 + a_1(t - T)] e^{\alpha(\tau - t + T)} \quad \text{при } \tau < t - T. \end{aligned}$$

На рис. 14.4.1 приведен график функции $f(t, \tau)$, определяемой выведенной формулой в интервале $-\infty < \tau < t - T$ и равной правой части интегрального уравнения $a_0 + a_1\tau$ в интервале $t - T \leq \tau \leq t$. Мы видим, что решение интегрального уравнения изложенным методом требует такого

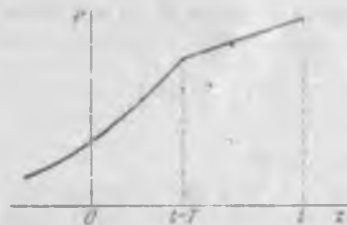


Рис. 14.4.1.

продолжения функции $f(t, \tau)$ в область $\tau < t - T$, при котором ее аналитическое выражение не сохраняется, а совершенно меняет свой характер. В данном случае функция $f(t, \tau)$ имеет угловую точку при $\tau = t - T$.

Пример 14.4.2. Найти оптимальную линейную систему с памятью длительности T для экстраполяции полезного сигнала, представляющего собой линейную функцию времени со случайными коэффициентами, если помеха $X(t)$ имеет математическое ожидание, тождественно равное нулю, и корреляционную функцию $k_x(\tau) = D e^{-\alpha|\tau|}$.

Согласно изложенному в § 13.6 необходимо прежде всего решить интегральные уравнения

$$\left. \begin{aligned} D \int_{t-T}^t e^{-\alpha|\tau-\sigma|} g^{(1)}(t, \tau) d\tau &= 1, \\ D \int_{t-T}^t e^{-\alpha|\tau-\sigma|} g^{(2)}(t, \tau) d\tau &= \sigma. \end{aligned} \right\} \quad (t-T \leq \sigma \leq t) \quad (14.4.56)$$

Эти уравнения являются частными случаями уравнения (14.4.54). Поэтому решение первого из этих уравнений дается формулой (14.4.55) при $a_0 = 1$, $a_1 = 0$, а решение второго — формулой (14.4.55) при $a_0 = 0$, $a_1 = 1$:

$$\left. \begin{aligned} g^{(1)}(t, \tau) &= \frac{1}{2D} [\alpha + \delta(\tau - t + T) + \delta(\tau - t)], \\ g^{(2)}(t, \tau) &= \frac{1}{2D} \left[\alpha\tau + \frac{\alpha t - \alpha T - 1}{\alpha} \delta(\tau - t + T) + \right. \\ &\quad \left. + \frac{\alpha t + 1}{\alpha} \delta(\tau - t) \right]. \end{aligned} \right\} \quad (14.4.57)$$

Подставляя эти выражения в (13.6.6) и принимая во внимание, что в данном случае $\varphi_1(\tau) = 1$, $\varphi_2(\tau) = \tau$, находим величины b_{pq} :

$$\left. \begin{aligned} b_{11} &= \frac{2 + \alpha T}{2D}, \quad b_{12} = b_{21} = \frac{(2 + \alpha T)(2t - T)}{4D}, \\ b_{22} &= \frac{1}{6D\alpha} [3T + 3\alpha t(t - T)(2 + \alpha T) + \alpha T^2(3 + \alpha T)]. \end{aligned} \right\} \quad (14.4.58)$$

Определив коэффициенты b_{pq} , находим λ_1 , λ_2 путем решения линейных алгебраических уравнений (13.6.7):

$$\left. \begin{aligned} (b_{11} + c_{11})\lambda_1 + (b_{12} + c_{12})\lambda_2 &= 1, \\ (b_{12} + c_{12})\lambda_1 + (b_{22} + c_{22})\lambda_2 &= t + \Delta, \end{aligned} \right\} \quad (14.4.59)$$

где Δ — время экстраполяции, а c_{pq} определяются формулой (13.5.26) при $N = 2$.

После нахождения величин λ_1 и λ_2 весовая функция оптимальной линейной системы определится формулой (13.6.4):

$$g(t, \tau) = \lambda_1(t) g^{(1)}(t, \tau) + \lambda_2(t) g^{(2)}(t, \tau). \quad (14.4.60)$$

Так как нерегулярная часть полезного сигнала в данном случае отсутствует, то для определения минимальной средней квадратической ошибки, соответствующей оптимальной системе, можно воспользоваться формулой (13.6.17). В результате, принимая во внимание, что в данном случае $\varphi_1(t) = 1$, $\varphi_2(t) = t + \Delta$, получим

$$\eta_{\min} = \lambda_1(t) + \lambda_2(t)(t + \Delta). \quad (14.4.61)$$

Пример 14.4.3. Решить интегральное уравнение

$$D \int_{t-T}^t e^{-\alpha|\tau-\sigma|} \cos \omega_0(\tau - \sigma) g(t, \tau) d\tau = e^{\alpha\sigma} \quad (t - T \leq \sigma \leq t). \quad (14.4.62)$$

В данном случае спектральная плотность стационарной случайной функции $X(t)$ была представлена в виде (14.4.4) в примере 7.6.3, и при этом было найдено

$$H(\mu) = \sqrt{2Da} (b + \mu), \quad \mu_1 = -b,$$

$$F(v) = b^2 + 2av + v^2, \quad v_1 = -a - i\omega_0, \quad v_2 = -a + i\omega_0,$$

где $b = \sqrt{a^2 + \omega_0^2}$, и

$$\frac{H(v_1)H(-v_1)}{F(-v_1)} = -Di\omega_0, \quad \frac{H(v_2)H(-v_2)}{F(-v_2)} = Di\omega_0,$$

$$F(\mu_1) = 2b(b-a), \quad F(-\mu_1) = 2b(b+a).$$

Формулы (14.4.31) и (14.4.37) дают

$$\begin{aligned} \xi_0(t, \tau) &= \frac{e^{-b\tau}}{\sqrt{2Da}} \left\{ \int_{t-T}^{\tau} e^{b\lambda} \left(b^2 + 2a \frac{d}{d\lambda} + \frac{d^2}{d\lambda^2} \right) e^{a\lambda} d\lambda + \right. \\ &\quad \left. + e^{b(t-T)} [(2a-b)e^{a(t-T)} + ae^{a(t-T)}] \right\} = \\ &= \frac{1}{\sqrt{2Da}} \left[\frac{b^2 + 2a\alpha + \alpha^2}{b + \alpha} e^{\alpha\tau} - \frac{2b(b-a)}{b + \alpha} e^{-b(\tau-t+T) + \alpha(t-T)} \right], \end{aligned}$$

$$\begin{aligned} \eta_0(t, \tau) &= \frac{e^{b\tau}}{2Da} \int_{\tau}^t e^{-b\sigma} \left[\frac{b^2 + 2a\alpha + \alpha^2}{b + \alpha} e^{\alpha\tau} - \frac{2b(b-a)}{b + \alpha} e^{-b(\tau-t+T) + \alpha(t-T)} \right] d\sigma = \\ &= \frac{\alpha^2 + 2a\alpha + b^2}{2Da(b^2 - \alpha^2)} [e^{\alpha\tau} - e^{-(b-\alpha)t + b\tau}] + \frac{b-a}{2Da(b+\alpha)} e^{(b+\alpha)(t-T)} [e^{-2bt + b\tau} - e^{-b\tau}]. \end{aligned}$$

Формулы (14.4.42), (14.4.44), (14.4.45) и (14.4.46) в данном случае дают

$$\begin{aligned} A_0 = \Delta_0 \eta' &= \frac{\alpha^2 + 2a\alpha + b^2}{2Da(b^2 - \alpha^2)} e^{\alpha t} (\alpha e^{-\alpha T} - b e^{-bT}) + \\ &+ \frac{b(b-a)}{2Da(b+\alpha)} e^{\alpha(t-T)} (1 + e^{-2bT}) + b_1 \left\{ \frac{a + i\omega_0}{Di\omega_0} - \frac{b(b-a)(1 + e^{-2bT})}{2Da(b+a+i\omega_0)} \right\} - \\ &- b_2 \left\{ \frac{a - i\omega_0}{Di\omega_0} + \frac{b(b-a)(1 + e^{-2bT})}{2Da(b+a-i\omega_0)} \right\}, \\ B_0 = \Delta_1 \eta' &= \frac{\alpha^2 + 2a\alpha + b^2}{2Da(b+\alpha)} e^{\alpha t} - \frac{b(b-a)}{Da(b+\alpha)} e^{\alpha t - (b+\alpha)T} + \\ &+ \frac{(b-a)e^{-bT}}{2Da(b+\alpha)} [(b+a-i\omega_0)b_1 + (b+a+i\omega_0)b_2]. \end{aligned}$$

Для определения величин b_1 и b_2 имеем одно уравнение (14.4.36) и одно уравнение (14.4.38):

$$\begin{aligned} -b_1 \left\{ \frac{1}{Di\omega_0} + \frac{(b-a)(1 - e^{-2bT})}{2Da(b+a+i\omega_0)} \right\} + b_2 \left\{ \frac{1}{Di\omega_0} - \frac{(b-a)(1 - e^{-2bT})}{2Da(b+a-i\omega_0)} \right\} = \\ = \frac{\alpha^2 + 2a\alpha + b^2}{2Da(b^2 - \alpha^2)} e^{\alpha t} (e^{-\alpha T} - e^{-bT}) - \frac{b-a}{2Da(b+\alpha)} e^{\alpha(t-T)} (1 - e^{-2bT}), \\ b_1 + b_2 = e^{\alpha(t-T)}. \end{aligned}$$

Формулы (14.4.47) в данном случае дают

$$x_1 = b_1 \frac{b(b^2 - a^2) e^{-bT}}{Da(a + b + i\omega_0)} + b_2 \frac{b(b^2 - a^2) e^{-bT}}{Da(a + b - i\omega_0)},$$

$$x_2 = -b_1 \frac{b(b-a)^2 e^{-bT}}{Da(a + b + i\omega_0)} - b_2 \frac{b(b-a)^2 e^{-bT}}{Da(a + b - i\omega_0)}.$$

Наконец, подставляя найденное выражение функции $\eta_0(t, \tau)$ в формулу (14.4.48), получим окончательную формулу для решения интегрального уравнения (14.4.62):

$$g(t, \tau) = \frac{(\alpha^2 + b^2)^2 - 4a^2\alpha^2}{2Da(b^2 - \alpha^2)} e^{\alpha\tau} + \frac{b(b-a)^2}{Da(b+\alpha)} e^{\alpha(t-T)-bT-b(t-\tau)} -$$

$$- \frac{b(b-a)(\alpha^2 + 2a\alpha + b^2)}{Da(b^2 - \alpha^2)} e^{\alpha t - b(t-\tau)} - \frac{b\omega_0^2}{Da(b+\alpha)} e^{\alpha(t-T)-bT+b(t-\tau)} +$$

$$+ x_1 e^{b(t-\tau)} + x_2 e^{-b(t-\tau)} + A_0 \delta(\tau - t + T) + B_0 \delta(\tau - t). \quad (14.4.63)$$

Уравнение (14.4.62) можно решить также методом неопределенных коэффициентов. Решение этого уравнения согласно изложенному представляет собой линейную комбинацию показательных функций $e^{b(t-\tau)}$ и $e^{-b(t-\tau)}$, δ -функций $\delta(\tau - t + T)$ и $\delta(\tau - T)$ и функций, входящих в состав выражения $F\left(-\frac{d}{d\tau}\right) \eta_0(t, \tau)$. Так как в данном случае правая часть интегрального

уравнения есть показательная функция, то в результате любых дифференциальных и интегральных операций над ней получится та же показательная функция, умноженная на постоянный (точнее, не зависящий от τ) коэффициент. Дополнительные члены, которые появились у нас при интегрировании в конечных пределах в формуле (14.4.31), представляют собой постоянные, умноженные на те же показательные функции $e^{b(t-\tau)}$ и $e^{-b(t-\tau)}$. Поэтому их можно объединить с членами $x_1 e^{b(t-\tau)}$ и $x_2 e^{-b(t-\tau)}$ соответственно. Это равноценно объединению четвертого члена правой части формулы (14.4.63) с членом $x_1 e^{b(t-\tau)}$, а второго и третьего членов — с членом $x_2 e^{-b(t-\tau)}$. В результате, обозначая полученные путем такого объединения коэффициенты при $e^{b(t-\tau)}$ и $e^{-b(t-\tau)}$ соответственно через x'_1 и x'_2 , приходим к следующей форме искомого решения:

$$g(t, \tau) = x'_1 e^{b(t-\tau)} + x'_2 e^{-b(t-\tau)} + x'_3 e^{\alpha\tau} + A_0 \delta(\tau - t + T) + B_0 \delta(\tau - t).$$

Подставляя это выражение в левую часть уравнения (14.4.62) и выполняя интегрирование, получим в левой части линейную комбинацию показательных функций $e^{\alpha\sigma}$, $e^{b(t-\sigma)}$, $e^{-b(t-\sigma)}$, $e^{-(a+i\omega_0)\sigma}$ и $e^{-(a-i\omega_0)\sigma}$. Для того чтобы получить тождество, необходимо приравнять друг другу коэффициенты при $e^{\alpha\sigma}$ в левой и правой частях полученного равенства и потребовать равенства нулю коэффициентов при $e^{b(t-\sigma)}$, $e^{-b(t-\sigma)}$, $e^{-(a+i\omega_0)\sigma}$ и $e^{-(a-i\omega_0)\sigma}$. В результате получим систему пяти линейных алгебраических уравнений для определения пяти неизвестных коэффициентов x'_1 , x'_2 , x'_3 , A_0 , B_0 . Предоставляем читателю самостоятельно проделать все эти выкладки и убедиться в том, что в результате получится выражение функции $g(t, \tau)$, совпадающее с (14.4.63).

§ 14.5. Синтез оптимальной стационарной линейной системы методом логарифмических частотных характеристик

Определение оптимальной системы является необходимым этапом проектирования автоматической системы управления, работающей в условиях помех. Без знания оптимальной системы нельзя обоснованно предъявлять требования к проектируемой системе. Однако определение оптимальной системы еще не решает основной задачи, стоящей перед конструктором, так как оптимальная система определяет только последовательность математических операций, которые необходимо произвести над входным сигналом для

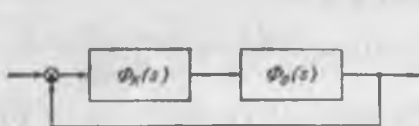


Рис. 14.5.1.

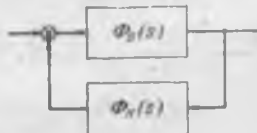


Рис. 14.5.2.

обеспечения наибольшей точности приближения к требуемому выходному сигналу. Иными словами, оптимальная система определяет динамические характеристики, которыми должна обладать проектируемая система. Конструктор должен эти динамические характеристики реализовать в проектируемой системе. При этом он должен учитывать, что большая часть элементов системы управления и главная часть ее структурной схемы определены назначением системы и изменять их невозможно.

Для приближения динамических характеристик проектируемой системы к оптимальным конструктор может вводить в структурную схему корректирующие цепи, подобрав для них соответствующую структуру и параметры. Корректирующая цепь может быть последовательной или параллельной. В первом случае она устанавливается в прямой цепи системы, а во втором — в цепи обратной связи. На рис. 14.5.1 и 14.5.2 часть системы с передаточной функцией $\Phi_0(s)$ задана (например, является объектом управления), а система с передаточной функцией $\Phi_K(s)$ является корректирующей цепью, которую требуется спроектировать. Определение структурной схемы и параметров корректирующей цепи из условия близости динамических характеристик проектируемой системы к оптимальным будем называть *синтезом* системы. Общих методов синтеза автоматических систем управления в настоящее время не существует. Сравнительно просто задача синтеза решается лишь в случае стационарной линейной системы.

Основная идея синтеза заключается в том, что по найденной передаточной функции оптимальной стационарной линейной системы и передаточным функциям известных звеньев проектируемой

системы определяется оптимальная передаточная функция корректирующей цепи. Обозначим передаточную функцию оптимальной системы через $\Psi(s)$. Передаточную функцию последовательной корректирующей цепи можно определить по формуле (4.6.8), положив в ней в соответствии со схемой на рис. 14.5.1 $\Phi_1(s) = \Phi_0(s) \Phi_K(s)$. Тогда, решая полученное уравнение относительно $\Phi_K(s)$, получим

$$\Phi_K(s) = \frac{\Psi(s)}{\Phi_0(s) [1 - \Psi(s)]}. \quad (14.5.1)$$

Для нахождения передаточной функции корректирующей цепи в обратной связи, можно воспользоваться формулой (4.6.7), положив в ней в соответствии со схемой на рис. 14.5.2 $\Phi_1(s) = \Phi_0(s)$, $\Phi_2(s) = \Phi_K(s)$. В результате получим

$$\Phi_K(s) = \frac{1}{\Psi(s)} - \frac{1}{\Phi_0(s)} = \frac{\Phi_0(s) - \Psi(s)}{\Psi(s) \Phi_0(s)}. \quad (14.5.2)$$

Не представляет труда также выразить оптимальную передаточную функцию корректирующей цепи через оптимальную передаточную функцию системы и передаточные функции заданных звеньев системы в случае более сложных автоматических систем.

Передаточная функция $\Psi(s)$ оптимальной системы в зависимости от статистических характеристик входных сигналов и интервала наблюдения может быть либо дробно-рациональной, либо трансцендентной. Соответственно передаточная функция корректирующей цепи также будет дробно-рациональной или трансцендентной. В последнем случае реализовать ее элементарными звеньями не представляется возможным. Для приближенной реализации передаточной функции корректирующей цепи необходимо аппроксимировать ее дробно рациональной функцией. Для этого можно воспользоваться методом логарифмических частотных характеристик.

Для приближенного представления передаточной функции корректирующей цепи дробно-рациональной функцией с помощью метода логарифмических частотных характеристик необходимо по найденной передаточной функции оптимальной системы построить ее амплитудную и фазовую логарифмические частотные характеристики. Затем в случае последовательной корректирующей цепи определяются логарифмические частотные характеристики разомкнутой системы. Это можно сделать с помощью номограммы связи между логарифмическими частотными характеристиками разомкнутой системы и замкнутой системы (приложение 2). В случае корректирующей цепи в обратной связи структурная схема на рис. 14.5.2 преобразуется, как показано на рис. 14.5.3 (см. § 4.6). Затем логарифмические частотные характеристики заданной части системы, соответствующие передаточной функции

$\Phi_0(s)$, вычитаются из соответствующих характеристик оптимальной системы. После этого с помощью номограммы находятся логарифмические частотные характеристики разомкнутой системы. Полученная таким образом логарифмическая амплитудная характеристика разомкнутой системы аппроксимируется отрезками прямых линий с наклонами, кратными 20 дБ/дек. При этом необходимо стремиться обеспечить, чтобы всем сопрягающим частотам элементарных звеньев, входящих в состав заданной части системы, имеющей передаточную функцию $\Phi_0(s)$, соответствовали угловые

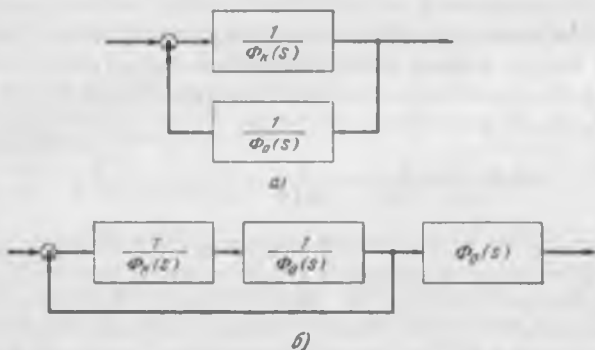


Рис. 14.5.3.

точки ломаной. В результате определяются типы элементарных звеньев, входящих в состав разомкнутой системы, их сопрягающие частоты и ее общий коэффициент усиления. Одновременно следует строить фазовую характеристику полученной разомкнутой системы и следить, чтобы она была достаточно близка к оптимальной. При этом в случае последовательной корректирующей цепи следует добиваться хорошего приближения только в диапазонах частот, в которых логарифмическая амплитудная характеристика разомкнутой системы превышает — 20 дБ. В случае корректирующей цепи в обратной связи следует добиваться хорошего приближения только в тех диапазонах частот, в которых логарифмическая амплитудная характеристика разомкнутой системы меньше 20 дБ, так как в приведенную ниже формулу (14.5.4) для передаточной функции корректирующей цепи в обратной связи передаточная функция разомкнутой системы входит в знаменатель.

Определив передаточную функцию разомкнутой системы $\Phi_p(s)$ и зная передаточную функцию заданной части системы $\Phi_0(s)$, можно найти передаточную функцию корректирующей цепи. В случае последовательной корректирующей цепи ее передаточная функция определяется формулой (рис. 14.5.1).

$$\Phi_k(s) = \frac{\Phi_p(s)}{\Phi_0(s)}, \quad (14.5.3)$$

В случае корректирующей цепи в обратной связи ее передаточная функция в соответствии со схемой на рис. 14.5.3, б определяется формулой

$$\Phi_K(s) = \frac{1}{\Phi_0(s) \Phi_P(s)}. \quad (14.5.4)$$

Вычисления по формулам (14.5.3) и (14.5.4) практически выполняются, конечно, простым суммированием и вычитанием соответствующих логарифмических частотных характеристик. При этом определяются типы элементарных звеньев, входящих в состав корректирующей цепи, и их параметры.

Передаточная функция корректирующей цепи $\Phi_K(s)$, найденная изложенным методом, будет дробно-рациональной функцией. Однако степень числителя этой функции часто оказывается больше степени знаменателя. Такие передаточные функции практически невозможно реализовать. Если в состав входной функции оптимальной системы входит нерегулярная случайная часть полезного сигнала и помеха является белым шумом, то передаточная функция разомкнутой системы в случае последовательной корректирующей цепи будет иметь степень числителя на единицу меньше степени знаменателя, а в случае корректирующей цепи в обратной связи, как правило, степень числителя будет по меньшей мере на единицу больше степени знаменателя. Имея в виду, что качество приближения передаточной функции разомкнутой системы несущественно в диапазонах частот, в которых ее логарифмическая амплитудная характеристика меньше — 20 дБ в случае последовательной корректирующей цепи и больше 20 дБ в случае корректирующей цепи в обратной связи, можно в этих диапазонах произвольно увеличить абсолютную величину наклона логарифмической амплитудной характеристики на величину, кратную 20 дБ, с таким расчетом, чтобы получить степень числителя передаточной функции корректирующей цепи меньшей степени знаменателя.

В общем случае, когда помеха не является белым шумом, чтобы сделать степень числителя передаточной функции корректирующей цепи меньше степени знаменателя, необходимо ограничить класс линейных систем, в котором ищется оптимальная система, дополнительным условием, чтобы при действии на входе белого шума производные выходной переменной оптимальной системы до определенного порядка имели конечные дисперсии. А именно, если числитель и знаменатель передаточной функции заданной части системы являются полиномами степени m и n соответственно, то производная порядка $n - m - 1$ выходной переменной оптимальной системы должна иметь конечную дисперсию при действии на входе белого шума [38].

Пример 14.5.1. Спроектировать следящую систему с конечной памятью T для воспроизведения полезного сигнала, представляющего собой линейную функцию времени $U_1 + U_2 t$ с неизвестными коэффициентами,

если помеха $X(t)$ является белым шумом интенсивности k , а исполнительное устройство имеет передаточную функцию

$$\Phi_0(s) = \frac{k_1}{s(T_1 s + 1)}. \quad (14.5.5)$$

В этом случае входная функция системы $Z(t)$ и ее требуемый выходной сигнал $W(t)$ выражаются формулами

$$Z(t) = U_1 + U_2 t + X(t), \quad W(t) = U_1 + U_2 t, \quad (14.5.6)$$

где U_1 и U_2 — неизвестные величины, которые могут иметь любые значения (или, что согласно изложенному в § 13.5 то же самое, случайные величины

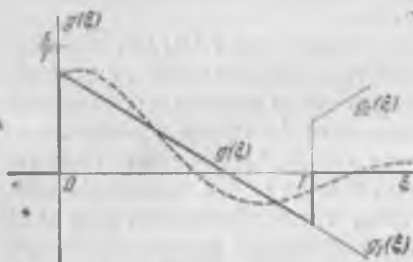


Рис. 14.5.4.

с бесконечными моментами второго порядка $\gamma_{11} = \gamma_{22} = \infty$). Предоставляем читателю, применяя метод § 14.1, самостоятельно решить уравнение § 13.6 и убедиться в том, что оптимальная линейная система в данном случае стационарна и ее весовая функция определяется формулой

$$g(t-\tau) = \frac{4}{T} - \frac{6}{T^2}(t-\tau) \quad (0 \leq t-\tau \leq T). \quad (14.5.7)$$

При $t-\tau < 0$ и при $t-\tau > T$ весовая функция оптимальной системы равна нулю. Таким образом, весовая функция оптимальной линейной системы $g(\xi)$ (где $\xi = t-\tau$) графически изображается отрезком прямой в интервале $0 \leq \xi \leq T$ и осью абсцисс вне этого интервала (жирная линия на рис. 14.5.4).

Для нахождения передаточной функции оптимальной линейной системы по известной весовой функции можно воспользоваться формулой (2.4.6). В результате получим

$$\begin{aligned} \Psi(s) &= \frac{4}{T} \int_0^T e^{-s\xi} d\xi - \frac{6}{T^2} \int_0^T \xi e^{-s\xi} d\xi = \\ &= \frac{4}{Ts} - \frac{6}{T^2 s^2} + \left(\frac{2}{Ts} + \frac{6}{T^2 s^2} \right) e^{-Ts}. \end{aligned} \quad (14.5.8)$$

Этот же результат можно получить другим способом. А именно из рис. 14.5.4 видно, что весовую функцию оптимальной системы можно представить как сумму двух весовых функций:

$$g_1(\xi) = \left(\frac{4}{T} - \frac{6}{T^2} \xi \right) 1(\xi), \quad g_2(\xi) = \left[\frac{2}{T} + \frac{6}{T^2} (\xi - T) \right] 1(\xi - T).$$

Графики этих двух весовых функций изображены на рис. 14.5.4 тонкими прямыми. Но $1(\xi)$ есть весовая функция интегратора, а $1(\xi - T)$ — весовая функция двойного интегратора. Сдвиг графиков весовых функций интегратора и двойного интегратора по оси ξ на величину T означает последовательное присоединение к ним запаздывающего звена с временем запаздывания T . Таким образом, оптимальную систему можно в данном случае представить как параллельное соединение интегратора с коэффициентом усиления $4/T$, двойного интегратора с коэффициентом усиления $-6/T^2$, интегратора с коэффициентом усиления $2/T$ с подключенным к нему последовательно запаздывающим звеном и двойного интегратора с коэффициентом усиления $6/T^2$ с подключенным к нему последовательно запаздывающим звеном (рис. 14.5.5). Из структурной схемы на рис. 14.5.5 немедленно вытекает формула (14.5.8) для передаточной функции оптимальной системы.

Для построения логарифмических частотных характеристик оптимальной системы представим ее структурной схемой, изображенной на рис. 14.5.6, и воспользуемся методом нахождения частотной характеристики параллельного соединения звеньев, изложенным в § 4.6. Согласно этому методу находим обратную систему (рис. 14.5.7) и преобразуем ее к системе с жесткой обратной связью (рис. 14.5.8). В результате мы представим систему, обратную по отношению к оптимальной системе, в виде последовательного соединения форсирующего звена первого порядка с постоянной времени $-2T/3$ и коэффициентом усиления -1 , аperiodического звена с постоянной времени $T/3$ и коэффициентом усиления 1 , запаздывающего звена, двойного дифференциатора и неустойчивого аperiodического звена с постоянной времени $-2T/3$ и коэффициентом усиления $-T^2/6$ с отрицательной жесткой обратной связью, охватывающей первые три звена. Следовательно, логарифмические частотные характеристики оптимальной системы можно построить, пользуясь характеристиками элементарных звеньев (§ 2.6) и номограммой, связывающей логарифмические

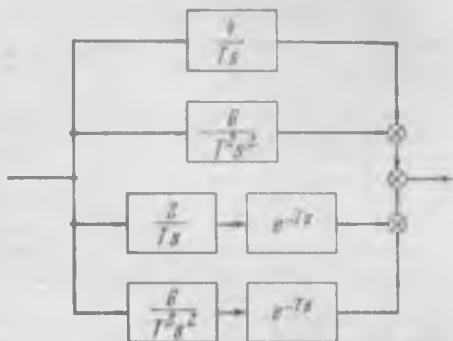


Рис. 14.5.5.

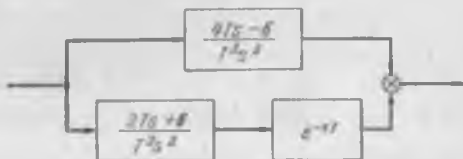


Рис. 14.5.6.

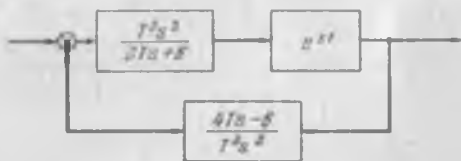


Рис. 14.5.7.

частотные характеристики разомкнутой и замкнутой систем. Полученные в результате логарифмические частотные характеристики оптимальной системы представлены на рис. 14.5.9.

Согласно изложенному методу для проектировании последовательной корректирующей цепи необходимо найти логарифмические частотные характеристики разомкнутой системы. Это можно сделать с помощью номограммы. Однако в данном случае номограмма не дает удовлетворительной точности при $T\omega < 3$, так как логарифмическая амплитудная характеристика замкнутой системы близка к нулю при $T\omega < 3$. Практически это означает, что значительное изменение частотных характеристик разомкнутой системы при

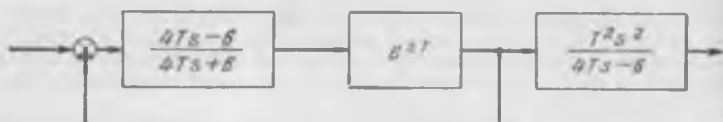


Рис. 14.5.8.

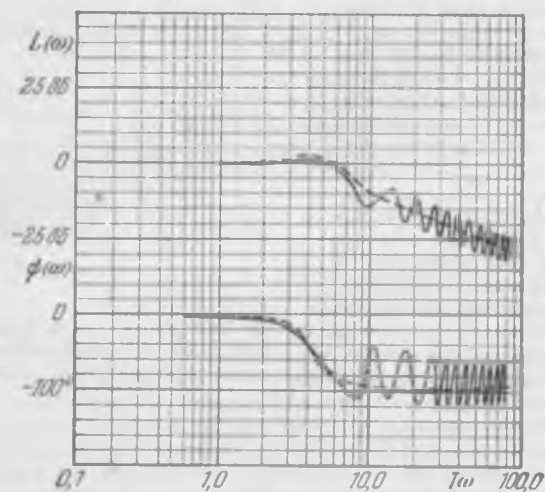


Рис. 14.5.9.

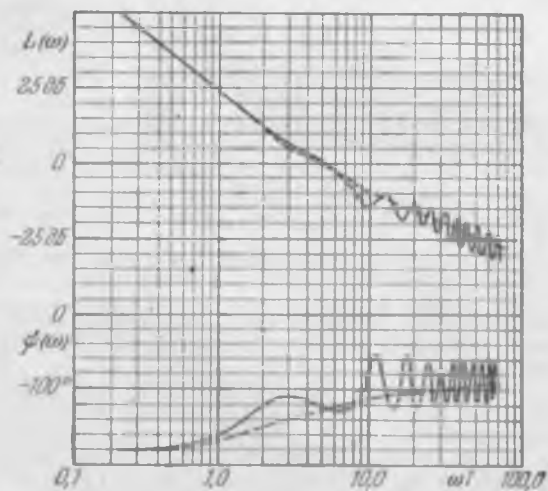


Рис. 14.5.10.

$T\omega < 3$ в данном случае не оказывает существенного влияния на характеристики замкнутой системы. Это дает возможность выбрать частотные характеристики разомкнутой системы при $T\omega < 3$ произвольно, чтобы удовлетворить другим требованиям к системе. В данном случае дисперсии случайных величин U_1 и U_2 бесконечны. Согласно изложенному в § 13.5 оптимальная система должна в этом случае идеально точно воспроизводить любые линейные функции времени при отсутствии помех по окончании переходного процесса. В § 7.3 мы видели, что для выполнения этого условия необходимо, чтобы система была астатической второго порядка. Исходя из этого условия, вводим в разомкнутую систему двойной интегратор. При этом логарифмическая амплитудная характеристика разомкнутой системы при $T\omega < 3$ будет представлять собой прямую с наклоном -40 дБ/дек, а фазовая характеристика будет близка к горизонтальной прямой на уровне -180° (рис. 14.5.10).

Из рис. 14.5.10 видно, что частотные характеристики разомкнутой системы достаточно хорошо аппроксимируются, если логарифмическую амплитудную характеристику аппроксимировать отрезком прямой с наклоном -20 дБ/дек на участке $3 < T\omega < 5$, отрезком прямой с наклоном -40 дБ/дек на участке $5 < T\omega < 7$ и прямой с наклоном -20 дБ/дек на участке $T\omega > 7$ (пунктирные линии на рис. 14.5.10). При этом получается хорошее приближение и к частотным характеристикам замкнутой оптимальной системы (пунктирные линии на рис. 14.5.9). Таким образом, мы можем приблизительно представить оптимальную разомкнутую систему в виде последовательного соединения двойного интегратора, двух форсирующих звеньев первого порядка с постоянными времени $T/3$ и $T/7$ и аperiodического звена с постоянной времени $T/5$. Для определения общего коэффициента усиления k_0 разомкнутой системы вспомним, что ордината низкочастотной асимптоты логарифмической амплитудной характеристики при $\omega = 1$ равна $20 \lg k_0$ (§ 4.7). В данном случае из рис. 14.5.10 видно, что логарифмическая амплитудная характеристика имеет значение 24 дБ при $T\omega = 1$. Следовательно, ее значение при $\omega = 1$ равно $24 + 40 \lg (1/T) = 24 - 40 \lg T$, т. е.

$$20 \lg k_0 = 24 - 40 \lg T \approx 20 \lg 16 - 40 \lg T.$$

Отсюда следует, что $k_0 = 16/T^2$. Таким образом, передаточная функция найденной разомкнутой системы выражается формулой

$$\Phi_p(s) = \frac{16(1 + 0,333Ts)(1 + 0,143Ts)}{T^2s^2(1 + 0,2Ts)}. \quad (14.5.9)$$

Из (14.5.5) и (14.5.9) находим выражение передаточной функции последовательной корректирующей цепи:

$$\Phi_k(s) = \frac{\Phi_p(s)}{\Phi_0(s)} = \frac{16(1 + 0,333Ts)(1 + 0,143Ts)(1 + T_1s)}{k_1T^2s(1 + 0,2Ts)}. \quad (14.5.10)$$

Если постоянная времени T_1 незначительно отличается от $0,2T$, то при практической реализации корректирующей цепи целесообразно заменить постоянную времени $0,2T$ величиной T_1 и таким образом упростить корректирующую цепь. Если T_1 значительно отличается от $0,2T$, то существенное упрощение корректирующей цепи будет невозможно без значительного отклонения от оптимальной характеристики.

Для завершения проектирования системы необходимо проверить, во-первых, насколько память спроектированной системы отличается от заданной и, во-вторых, насколько точность спроектированной системы ниже точности оптимальной системы. Для решения этих двух вопросов находим передаточную функцию спроектированной замкнутой системы:

$$\Phi_z(s) = \frac{\Phi_p(s)}{1 + \Phi_p(s)} = \frac{16(0,0476T^2s^2 + 0,476Ts + 1)}{0,2T^3s^3 + 1,762T^2s^2 + 7,62Ts + 16}. \quad (14.5.11)$$

По заданию требуется спроектировать систему с памятью длительности T . Точное выполнение этого условия невозможно, так как на основании общего теорема теории дифференциальных уравнений система с дробно-рациональной передаточной функцией всегда имеет бесконечную память. Однако практически можно считать память системы равной времени переходного процесса. Для определения времени переходного процесса находим по формулам (4.4.50) и (6.5.4) весовую и переходную функции спроектированной системы. Мы предоставляем читателю проделать необходимые выкладки самостоятельно и приведем только графики весовой функции (пунктирная кривая на рис. 14.5.4) и переходной функции (рис. 14.5.11). Из рис. 14.5.4 и 14.5.11 видно, что мы получили хорошее приближение к весовой функции оптимальной системы и время переходного процесса (т. е. время, начиная с которого

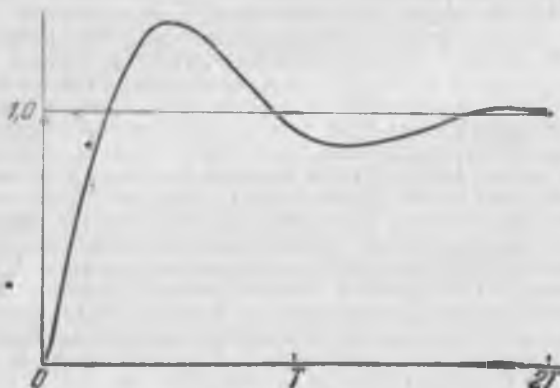


Рис. 14.5.11.

переходная функция $h(t)$ не отклоняется от ее установившегося значения $h(\infty)$ больше чем на 5%) около $1,55T$, что также можно считать удовлетворительным.

Для проверки точности спроектированной системы находим по формуле (7.4.4) установившуюся дисперсию ее выходной переменной. Имея в виду что спектральная плотность стационарного белого шума интенсивности k равна $k/2\pi$, получим

$$\eta = D_w = \frac{k}{2\pi} \int_{-\infty}^{\infty} \frac{0,581T^4 (\omega)^4 - 32,68T^2 (\omega)^2 + 256}{|0,2T^3 (\omega)^3 + 1,762T^2 (\omega)^2 + 7,62T (\omega) + 16|^2} d\omega \quad (14.5.12)$$

Проводя замену переменных $T\omega = \mu$, получаем

$$\eta = D_w = \frac{k}{T} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{0,581 (\mu)^4 - 32,68 (\mu)^2 + 256}{|0,2 (\mu)^3 + 1,762 (\mu)^2 + 7,62 (\mu) + 16|^2} d\mu. \quad (14.5.13)$$

Пользуясь для вычисления этого интеграла формулой (7.7.6) находим

$$\eta = D_w \approx \frac{4,10k}{T}.$$

Формула (13.6.17) дает в данном случае для дисперсии ошибки оптимальной системы выражение

$$\eta_{\min} = \frac{4k}{T}.$$

Таким образом, спроектированная система и по точности достаточно близка к оптимальной.

Предоставляем читателю самостоятельно проверить, насколько близка к оптимальной система, которая получится, если аппроксимировать логарифмическую амплитудную характеристику оптимальной разомкнутой системы двумя полупрямыми с наклонами -40 дБ/дек и -20 дБ/дек, сопрягающимися при $\omega T = 4$.

§ 14.6. Определение оптимальной линейной системы методом канонических разложений

Уравнение (14.1.1) может быть в общем случае решено методом канонических разложений. При этом решение получается в форме бесконечного ряда, конечным отрезком которого можно пользоваться для приближенного определения оптимальной линейной системы.

Представим случайную функцию $X(\tau)$ которая, как мы помним является суммой нерегулярной части полезного сигнала и помехи, каким-нибудь ее каноническим разложением в интервале наблюдения $t - T \leq \tau \leq t$:

$$X(\tau) = \sum_v V_v x_v(\tau) \quad (t - T \leq \tau \leq t). \quad (14.6.1)$$

Как известно, это всегда можно сделать для любой случайной функции с конечной дисперсией. Для этого достаточно найти систему пар функций $a_v(\tau)$ и $x_v(\tau)$, удовлетворяющих условиям (см. [54], § 6.2 или [53], § 60)

$$\int_{t-T}^t x_v(\tau) \overline{a_\mu(\tau)} d\tau = \delta_{v\mu}, \quad (14.6.2)$$

$$x_v(\tau) = \frac{1}{D_v} \int_{t-T}^t K_x(\tau, \sigma) a_v(\sigma) d\sigma, \quad (14.6.3)$$

где для краткости через $\delta_{v\mu}$ обозначены числа, равные единице при одинаковых индексах и нулю при различных индексах ($\delta_{vv} = 1$, $\delta_{v\mu} = 0$ при $\mu \neq v$). Предположим, что функции $a_v(\tau)$ и $x_v(\tau)$, удовлетворяющие условиям (14.6.2) и (14.6.3), найдены, и будем искать решение интегрального уравнения (14.1.1) в форме ряда

$$g(t, \tau) = \sum_v \xi_v \overline{a_v(\tau)}, \quad (14.6.4)$$

где ξ_v — неопределенные коэффициенты.

Подставляя выражение (14.6.4) в уравнение (14.1.1) и выполняя почленное интегрирование ряда, получим

$$\sum_v \xi_v \int_{t-T}^t K_x(\tau, \sigma) \overline{a_v(\tau)} d\tau = f(t, \sigma) \quad (t - T \leq \sigma \leq t). \quad (14.6.5)$$

По заданию требуется спроектировать систему с памятью длительности T . Точное выполнение этого условия невозможно, так как на основании общих теорем теории дифференциальных уравнений система с дробно-рациональной передаточной функцией всегда имеет бесконечную память. Однако практически можно считать память системы равной времени переходного процесса. Для определения времени переходного процесса находим по формулам (4.4.50) и (6.5.4) весовую и переходную функции спроектированной системы. Мы предоставляем читателю проделать необходимые выкладки самостоятельно и приведем только графики весовой функции (пунктирная кривая на рис. 14.5.4) и переходной функции (рис. 14.5.11). Из рис. 14.5.4 и 14.5.11 видно, что мы получили хорошее приближение к весовой функции оптимальной системы и время переходного процесса (т. е. время, начиная с которого

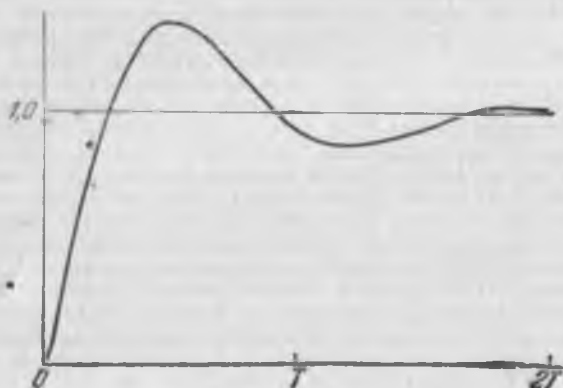


Рис. 14.5.11.

переходная функция $h(t)$ не отклоняется от ее установившегося значения $h(\infty)$ больше чем на 5%) около $1,55T$, что также можно считать удовлетворительным.

Для проверки точности спроектированной системы находим по формуле (7.4.4) установившуюся дисперсию ее выходной переменной. Имея в виду что спектральная плотность стационарного белого шума интенсивности k равна $k/2\pi$, получим

$$\eta = D_w = \frac{k}{2\pi} \int_{-\infty}^{\infty} \frac{0,581T^4 (\omega)^4 - 32,68T^2 (\omega)^2 + 256}{|0,2T^3 (\omega)^3 + 1,762T^2 (\omega)^2 + 7,62T (\omega) + 16|^2} d\omega \quad (14.5.12)$$

Производя замену переменных $T\omega = \mu$, получаем

$$\eta = D_w = \frac{k}{T} \frac{1}{2\pi} \int_{-\infty}^{\infty} \frac{0,581 (\mu)^4 - 32,68 (\mu)^2 + 256}{|0,2 (\mu)^3 + 1,762 (\mu)^2 + 7,62 (\mu) + 16|^2} d\mu. \quad (14.5.13)$$

Пользуясь для вычисления этого интеграла формулой (7.7.6) находим

$$\eta = D_w \approx \frac{4,10k}{T}.$$

Формула (13.6.17) дает в данном случае для дисперсии ошибки оптимальной системы выражение

$$\eta_{\text{lim}} = \frac{4k}{T}.$$

Таким образом, спроектированная система и по точности достаточно близка к оптимальной.

Предоставляем читателю самостоятельно проверить, насколько близка оптимальная система, которая получится, если аппроксимировать логарифмическую амплитудную характеристику оптимальной разомкнутой системы двумя полупрямыми с наклонами -40 дБ/дек и -20 дБ/дек, сопрягающимися при $\omega T = 4$.

§ 14.6. Определение оптимальной линейной системы методом канонических разложений

Уравнение (14.1.1) может быть в общем случае решено методом канонических разложений. При этом решение получается в форме бесконечного ряда, конечным отрезком которого можно пользоваться для приближенного определения оптимальной линейной системы.

Представим случайную функцию $X(\tau)$ которая, как мы помним является суммой нерегулярной части полезного сигнала и помехи, каким-нибудь ее каноническим разложением в интервале наблюдения $t - T \leq \tau \leq t$:

$$X(\tau) = \sum_{\nu} V_{\nu} x_{\nu}(\tau) \quad (t - T \leq \tau \leq t). \quad (14.6.1)$$

Как известно, это всегда можно сделать для любой случайной функции с конечной дисперсией. Для этого достаточно найти систему пар функций $a_{\nu}(\tau)$ и $x_{\nu}(\tau)$, удовлетворяющих условиям (см. [54], § 6.2 или [53], § 60)

$$\int_{t-T}^t x_{\nu}(\tau) \overline{\alpha_{\mu}(\tau)} d\tau = \delta_{\nu\mu}, \quad (14.6.2)$$

$$x_{\nu}(\tau) = \frac{1}{D_{\nu}} \int_{t-T}^t K_x(\tau, \sigma) a_{\nu}(\sigma) d\sigma, \quad (14.6.3)$$

где для краткости через $\delta_{\nu\mu}$ обозначены числа, равные единице при одинаковых индексах и нулю при различных индексах ($\delta_{\nu\nu} = 1$, $\delta_{\nu\mu} = 0$ при $\mu \neq \nu$). Предположим, что функции $a_{\nu}(\tau)$ и $x_{\nu}(\tau)$, удовлетворяющие условиям (14.6.2) и (14.6.3), найдены, и будем искать решение интегрального уравнения (14.1.1) в форме ряда

$$g(t, \tau) = \sum_{\nu} \xi_{\nu} \overline{a_{\nu}(\tau)}, \quad (14.6.4)$$

где ξ_{ν} — неопределенные коэффициенты.

Подставляя выражение (14.6.4) в уравнение (14.1.1) и выполняя почленное интегрирование ряда, получим

$$\sum_{\nu} \xi_{\nu} \int_{t-T}^t K_x(\tau, \sigma) \overline{a_{\nu}(\tau)} d\tau = f(t, \sigma) \quad (t - T \leq \sigma \leq t). \quad (14.6.5)$$

Но согласно формуле (14.6.3)

$$\int_{t-T}^t K_x(\tau, \sigma) \overline{a_v(\tau)} d\tau = \int_{t-T}^t K_x(\sigma, \tau) \overline{a_v(\tau)} d\tau = D_v \overline{x_v(\sigma)}.$$

Подставляя это выражение в (14.6.5), получим

$$\sum_{\nu} D_{\nu} \overline{\xi_{\nu} x_{\nu}(\sigma)} = f(t, \sigma) \quad (t-T \leq \sigma \leq t). \quad (14.6.6)$$

Таким образом, для определения неизвестных коэффициентов ξ_{ν} достаточно представить правую часть интегрального уравнения (14.1.1) $f(t, \sigma)$, рассматриваемую как функция второго аргумента, разложением по сопряженным координатным функциям $x_{\nu}(\sigma)$ в интервале наблюдения $t-T \leq \sigma \leq t$.

Для определения коэффициентов ξ_{ν} умножим равенство (14.6.6) на $a_{\mu}(\sigma)$ и проинтегрируем результат по σ в пределах от $t-T$ до t . Тогда получим

$$\sum_{\nu} D_{\nu} \xi_{\nu} \int_{t-T}^t a_{\mu}(\sigma) \overline{x_{\nu}(\sigma)} d\sigma = \int_{t-T}^t a_{\mu}(\sigma) f(t, \sigma) d\sigma. \quad (14.6.7)$$

Все члены ряда в левой части этого равенства, кроме члена, для которого индекс суммирования ν равен μ , равны нулю вследствие формулы (14.6.2). Для члена ряда, соответствующего $\nu = \mu$, интеграл на основании формулы (14.6.2) равен единице. Следовательно, равенство (14.6.7) дает

$$D_{\mu} \xi_{\mu} = \int_{t-T}^t a_{\mu}(\sigma) f(t, \sigma) d\sigma. \quad (14.6.8)$$

Отсюда находим

$$\xi_{\mu} = \frac{1}{D_{\mu}} \int_{t-T}^t a_{\mu}(\sigma) f(t, \sigma) d\sigma. \quad (14.6.9)$$

Таким образом, неизвестные коэффициенты ряда (14.6.4) определяются весьма просто.

Подставляя выражение (14.6.9) коэффициентов ξ_{ν} (сменив в нем, конечно, индекс μ на ν) в (14.6.4), получим следующую окончательную формулу для решения интегрального уравнения (14.1.1):

$$g(t, \tau) = \sum_{\nu} \frac{\overline{a_{\nu}(\tau)}}{D_{\nu}} \int_{t-T}^t a_{\nu}(\sigma) f(t, \sigma) d\sigma. \quad (14.6.10)$$

Таким образом, метод канонических разложений случайных функций дает возможность весьма просто найти решение основного интегрального уравнения статистической теории оптимальных систем (14.1.1) в форме бесконечного ряда (14.6.10). Этот ряд определяет решение интегрального уравнения (14.1.1) во всех случаях, когда его правая часть $f(t, \sigma)$ может быть представлена разложением (14.6.6) по сопряженным координатным функциям $\overline{x_v(\sigma)}$ случайной функции $X(t)$ в интервале наблюдения $t - T \leq \sigma \leq t$. В случае, когда функция $f(t, \sigma)$ не может быть представлена разложением (14.6.6), интегральное уравнение (14.1.1) не имеет решения ([53], §§ 134, 135).

В предыдущих выкладках мы выполнили почленное интегрирование ряда в формуле (14.6.5). Можно доказать, что эта операция законна и что конечным отрезком ряда (14.6.10) всегда можно пользоваться для приближенного определения весовой функции оптимальной линейной системы ([53], § 135). При этом, несмотря на то, что найденная таким образом весовая функция может довольно значительно отличаться от весовой функции оптимальной системы, соответствующая средняя квадратическая ошибка, как правило, лишь незначительно превосходит минимальную среднюю квадратическую ошибку оптимальной системы, если взять $10 \div 20$ членов ряда (14.6.10).

Изложенный метод решения интегрального уравнения (14.1.1) в форме бесконечного ряда гораздо проще методов, изложенных в предыдущих параграфах. Однако для применения этого метода необходимо предварительно определить систему пар функций $a_v(\tau)$, $x_v(\tau)$, определяющих каноническое разложение суммы нерегулярной части полезного сигнала и помехи. Кроме того, изложенным методом можно получить, как правило, лишь приближенное решение интегрального уравнения (14.1.1), в то время как методы, изложенные в предыдущих параграфах, дают точное решение.

Достоинством изложенного метода является то обстоятельство, что он даст возможность не только приближенно определить весовую функцию оптимальной системы, но и найти допустимые отклонения от нее, не приводящие к существенному ухудшению точности. А именно, если данный конечный отрезок ряда (14.6.10) определяет такую весовую функцию, для которой средняя квадратическая ошибка достаточно близка к минимальной, то любая линейная комбинация не использованных в этом отрезке ряда (14.6.10) функций $a_v(\tau)$ является допустимым отклонением от найденной весовой функции, не приводящим к существенному увеличению средней квадратической ошибки.

Полагая в формуле (14.6.10) последовательно

$$f(t, \sigma) = K_{jx}(t, \sigma), \varphi_1(\sigma), \dots, \varphi_N(\sigma),$$

найдем решения всех интегральных уравнений (13.6.1) и (13.6.2):

$$g^{(0)}(t, \tau) = \sum_v \frac{\overline{a_v(\tau)}}{D_v} \int_{t-T}^t a_v(\sigma) K_{yx}(t, \sigma) d\sigma, \quad (14.6.11)$$

$$g^{(r)}(t, \tau) = \sum_v \frac{\overline{a_v(\tau)}}{D_v} \int_{t-T}^t a_v(\sigma) \varphi_r(\sigma) d\sigma \quad (r=1, \dots, N). \quad (14.6.12)$$

Полагая для краткости

$$y_v(t) = \frac{1}{D_v} \int_{t-T}^t a_v(\sigma) K_{yx}(t, \sigma) d\sigma, \quad (14.6.13)$$

$$\alpha_{vr}(t) = \frac{1}{D_v} \int_{t-T}^t a_v(\sigma) \varphi_r(\sigma) d\sigma \quad (r=1, \dots, N), \quad (14.6.14)$$

можем представить формулы (14.6.11) и (14.6.12) в виде

$$g^{(0)}(t, \tau) = \sum_v y_v(t) \overline{a_v(\tau)}, \quad (14.6.15)$$

$$g^{(r)}(t, \tau) = \sum_v \alpha_{vr}(t) \overline{a_v(\tau)} \quad (r=1, \dots, N). \quad (14.6.16)$$

Подставляя эти выражения в формулы (13.6.6) и выполняя почленное интегрирование рядов, получим

$$\begin{aligned} b_{p0} &= \int_{t-T}^t g^{(0)}(t, \tau) \varphi_p(\tau) d\tau = \\ &= \sum_v y_v(t) \int_{t-T}^t \overline{a_v(\tau)} \varphi_p(\tau) d\tau \quad (p=1, \dots, N), \end{aligned} \quad (14.6.17)$$

$$\begin{aligned} b_{pq} &= \int_{t-T}^t g^{(q)}(t, \tau) \varphi_p(\tau) d\tau = \\ &= \sum_v \alpha_{vq}(t) \int_{t-T}^t \overline{a_v(\tau)} \varphi_p(\tau) d\tau \quad (p, q=1, \dots, N). \end{aligned} \quad (14.6.18)$$

Но на основании формулы (14.6.14)

$$\int_{t-T}^t \overline{a_v(\tau)} \varphi_p(\tau) d\tau = \overline{\int_{t-T}^t a_v(\sigma) \varphi_p(\sigma) d\sigma} = D_v \overline{\alpha_{vp}(t)}.$$

Следовательно, формулы (14.6.17) и (14.6.18) могут быть переписаны в виде

$$b_{p0} = \sum_v D_v y_v(t) \overline{\alpha_{vp}(t)} \quad (p = 1, \dots, N), \quad (14.6.19)$$

$$b_{pq} = \sum_v D_v \alpha_{vq}(t) \overline{\alpha_{vp}(t)} \quad (p, q = 1, \dots, N). \quad (14.6.20)$$

Точно так же, подставляя выражение (14.6.15) весовой функции $g^{(0)}$ в формулу (13.6.15), выполняя почленное интегрирование ряда и принимая во внимание (14.6.13), получим

$$b_{00} = \sum_v D_v |y_v(t)|^2. \quad (14.6.21)$$

Таким образом, выражая решения интегральных уравнений (13.6.1) и (13.6.2) в форме рядов (14.6.15) и (14.6.16), мы получаем выражения величин b_{pq} ($p, q = 0, 1, \dots, N$) также в форме рядов (14.6.19), (14.6.20) и (14.6.21). Можно доказать, что для сходимости всех этих рядов достаточно, чтобы случайная функция $Y(t)$ имела конечную дисперсию и чтобы сходились все ряды ([53], § 135)

$$b_{pp} = \sum_v D_v |\alpha_{vp}(t)|^2 \quad (p = 1, \dots, N). \quad (14.6.22)$$

Сходимость этих рядов также и необходима для того, чтобы линейные системы с весовыми функциями $g^{(0)}(t, \tau)$ преобразовывали случайную функцию $X(\tau)$ в случайную функцию, имеющую конечную дисперсию, т. е. необходима для существования требуемых решений интегральных уравнений (13.6.2).

§ 14.7. Определение оптимальной дискретной линейной системы

Изложенный в предыдущем параграфе общий метод можно, в частности, применить для определения оптимальной дискретной линейной системы. В этом случае система наблюдает входной сигнал $Z(t)$ в дискретном ряде точек $t_1, \dots, t_h$ и для нахождения оптимальной системы необходимо представить случайную функцию X (т. е. сумму помехи и нерегулярной части полезного сигнала, если она имеется) каноническим разложением в этом дискретном ряде точек:

$$X(t_k) = \sum_{v=1}^h V_v x_v(t_k) \quad (k = 1, \dots, h). \quad (14.7.1)$$

Число членов этого разложения не превосходит числа точек h в интервале наблюдения. Следовательно, при конечном h сумма в (14.7.1) конечна, а при $h = \infty$ эта сумма представляет собой в общем случае бесконечный ряд.

Для нахождения оптимальной линейной системы в этом случае достаточно во всех формулах предыдущего параграфа взять функции $a_v(\tau)$ в виде линейных комбинаций δ -функций:

$$a_v(\tau) = \sum_{l=1}^h a_{vl} \delta(\tau - t_l) \quad (v = 1, \dots, h), \quad (14.7.2)$$

удовлетворяющих условиям (14.6.2) и (14.6.3). Подстановка выражения (14.7.2) в (14.6.2) и (14.6.3) дает следующие условия, которым должны удовлетворять коэффициенты a_{vl} *):

$$\sum_{l=1}^h a_{vl} x_v(t_l) = \delta_{v\mu} \quad (v, \mu = 1, \dots, h), \quad (14.7.3)$$

$$x_v(\tau) = \frac{1}{D_v} \sum_{l=1}^h a_{vl} K_x(\tau, t_l) \quad (v = 1, \dots, h). \quad (14.7.4)$$

Как известно, коэффициенты a_{vl} , удовлетворяющие этим условиям, всегда могут быть найдены, и притом бесчисленным множеством способов (см. [54], § 6.3).

Подставляя выражение (14.7.2) в формулы (14.6.15) и (14.6.16), получим следующие формулы для весовых коэффициентов дискретных линейных систем, удовлетворяющих системам алгебраических уравнений (13.7.2) и (13.7.3):

$$g_{hk}^{(0)} = \sum_{v=1}^h a_{vh} y_v(t_h) \quad (k = 1, \dots, h), \quad (14.7.5)$$

$$g_{hk}^{(r)} = \sum_{v=1}^h a_{vh} \alpha_{vr} \quad (k = 1, \dots, h; r = 1, \dots, N). \quad (14.7.6)$$

Для определения коэффициентов $y_v(t_h)$ и α_{vr} в этих формулах подставляем выражение (14.7.2) в формулы (14.6.13) и (14.6.14). Тогда получим

$$y_v(t_h) = \frac{1}{D_v} \sum_{l=1}^h a_{vl} K_{yx}(t_h, t_l) \quad (v = 1, \dots, h), \quad (14.7.7)$$

$$\alpha_{vr} = \frac{1}{D_v} \sum_{l=1}^h a_{vl} \Psi_r(t_l) \quad (v = 1, \dots, h; r = 1, \dots, N). \quad (14.7.8)$$

После определения весовых коэффициентов $g_{hk}^{(0)}, g_{hk}^{(1)}, \dots, g_{hk}^{(N)}$ по формулам (14.7.5) — (14.7.8), вычисления величин b_{pq} ($p, q = 0, 1, \dots, N$) по формулам (14.6.19), (14.6.20) и (14.6.21) или по формулам (13.7.4) и (13.7.6) и решения систем

*) Каноническое разложение случайной функции в дискретном ряде точек всегда выгодно брать в действительной форме.

линейных алгебраических уравнений (13.6.7) весовые коэффициенты оптимальной дискретной линейной системы находятся по формуле (13.7.5).

Легко видеть, что изложенный метод дает точные формулы для весовых коэффициентов оптимальной дискретной линейной системы в случае конечного числа точек наблюдения входной функции, что всегда и бывает в задачах практики. Поэтому метод канонических разложений является простейшим и самым удобным методом нахождения оптимальных дискретных линейных систем. Решение систем линейных алгебраических уравнений (13.7.2) и (13.7.3) обычными методами при большом h очень громоздко и трудоемко.

Пример 14.7.1. Найти оптимальную дискретную линейную систему для условий примера 14.4.2, предполагая, что входная случайная функция Z действует на систему в равноотстоящие моменты времени $t_k = t_h - (h - k) T_n$ ($k = 1, \dots, h$; $T_n = T/(h - 1)$).

Согласно изложенному в § 13.7 решение задачи приводится в данном случае к решению двух систем линейных алгебраических уравнений (13.7.3) для $f(t, \sigma) = \varphi_1(\sigma) = 1$ и $f(t, \sigma) = \varphi_2(\sigma) = \sigma$. Воспользуемся каноническим разложением случайной функции X в точках $t_1, \dots, t_h$. Легко проверить, что в данном случае уравнениям (14.7.3) и (14.7.4) удовлетворяют

$$\left. \begin{aligned} \alpha_{11} &= 1, & \alpha_{1l} &= 0 & \text{при } l > 1, \\ \alpha_{vv} &= 1, & \alpha_{v, v-1} &= -q, & \alpha_{vl} &= 0 \text{ при } l < v-1 \text{ и при } l > v, \\ D_1 &= D, & D_v &= D(1 - q^2) & (v = 2, \dots, h), \\ x_1(\tau) &= e^{-\alpha|\tau - t_1|}, \\ x_v(\tau) &= \frac{1}{1 - q^2} (e^{-\alpha|\tau - t_v|} - e^{-\alpha|\tau - t_{v-1}|}) & (v = 2, \dots, h), \end{aligned} \right\} \quad (14.7.9)$$

где $q = e^{-\alpha T_n}$.

Подставляя выражения (14.7.9) коэффициентов α_{vh} и выражения функций φ_1 и φ_2 в (14.7.8), находим величины α_{vr} :

$$\begin{aligned} \alpha_{11} &= \frac{1}{D}, & \alpha_{v1} &= \frac{1}{D(1+q)}, & (v = 2, \dots, h), \\ \alpha_{12} &= \frac{t_1}{D}, & \alpha_{v2} &= \frac{t_v}{D(1+q)} + \frac{qT_n}{D(1-q^2)} & (v = 2, \dots, h). \end{aligned}$$

Подставляя выражения (14.7.9) коэффициентов α_{vh} и полученные выражения величин α_{vr} в (14.7.6), получим

$$\left. \begin{aligned} s_{hh}^{(1)} &= s_{hh}^{(1)} = \frac{1}{D(1+q)}, & s_{hh}^{(1)} &= \frac{1-q}{D(1+q)}, & (k = 2, \dots, h-1), \\ s_{h1}^{(2)} &= \frac{t_1}{D(1+q)} - \frac{qT_n}{D(1-q^2)}, & s_{hh}^{(2)} &= \frac{t_h}{D(1+q)} + \frac{qT_n}{D(1-q^2)}, \\ s_{hh}^{(2)} &= \frac{t_h(1-q)}{D(1+q)} & (k = 2, \dots, h-1), \end{aligned} \right\} \quad (14.7.10)$$

Формулы (14.6.19) — (14.6.21) или (13.7.4) в данном случае дают

$$\left. \begin{aligned} b_{11} &= \frac{h - (h-2)q}{D(1+q)}, & b_{12} &= b_{21} = b_{11} \left(t_1 + \frac{T}{2} \right), \\ b_{22} &= b_{11} \left(t_1 + \frac{T}{2} \right)^2 + \frac{TT_{\Pi}}{12D(1-q^2)} [h(h+1) - \\ &\quad - 2(h-3)(h+1)q + (h-3)(h-2)q^2]. \end{aligned} \right\} \quad (14.7.11)$$

После определения величин λ_1, λ_2 путем решения системы линейных алгебраических уравнений (14.4.59) при найденных значениях коэффициентов b_{jk} весовые коэффициенты оптимальной дискретной линейной системы, согласно (13.7.5), определяются формулой

$$g_{hk} = \lambda_1 g_{hk}^{(1)} + \lambda_2 g_{hk}^{(2)} \quad (k = 1, \dots, h). \quad (14.7.12)$$

Формулы (14.7.10) и (14.7.12) полностью и точно определяют оптимальную дискретную линейную систему при любом числе точек наблюдения h . Средняя квадратическая ошибка оптимальной дискретной линейной системы определяется в данном случае формулой (14.4.61) при $t = t_h$ при значениях λ_1, λ_2 , соответствующих условиям этого примера.

§ 14.8. Оптимальные системы, описываемые дифференциальными уравнениями

Ввиду того, что системы, описываемые обыкновенными дифференциальными уравнениями, сравнительно просто моделируются на аналоговых вычислительных машинах, естественно, возникает вопрос, в каких случаях оптимальную линейную систему можно описать конечным числом обыкновенных дифференциальных уравнений и как получить эти уравнения. Прежде всего, ясно, что никакая система с конечной памятью не может описываться дифференциальными уравнениями. Это следует непосредственно из теоремы о единственности решения линейных дифференциальных уравнений. Поэтому оптимальные системы, описываемые дифференциальными уравнениями, могут существовать только в частном классе линейных систем с бесконечной памятью. У таких систем интервал наблюдения T всегда совпадает с полным временем работы системы, $T = t - t_0$, и нижний предел во всех интегралах предыдущей главы постоянен и равен t_0 .

Чтобы найти подход к решению поставленного вопроса, рассмотрим сначала пример.

Пример 14.8.1. В примере 14.1.1 была решена задача выделения полезного сигнала известной формы, характеризуемой функцией $\varphi_1(t)$, в случае, когда помеха представляет собой белый шум постоянной интенсивности k . Полученные в примере 14.1.1 результаты относятся к общему случаю произвольной памяти системы. Покажем, что найденная в этом примере оптимальная система в частном случае, когда ее память бесконечна, описывается обыкновенным дифференциальным уравнением первого порядка. Положив $T = t - t_0$, перепишем для этого частного случая формулу для

весовой функции оптимальной системы:

$$g(t, \tau) = \frac{\Phi_1(t) \Phi_1(\tau)}{k(b_{11} + c_{11})} \quad (-\infty < \tau \leq t), \quad (14.8.1)$$

где

$$b_{11} = \frac{1}{k} \int_{t_0}^t \Phi_1^2(\tau) d\tau \quad (14.8.2)$$

и формулу для минимального среднего квадрата ошибки, реализуемого оптимальной системой:

$$\eta = \frac{\Phi_1^2(t)}{b_{11} + c_{11}} \quad (14.8.3)$$

(для удобства последующих выкладок мы пишем здесь просто η вместо $\eta_{\min}$). Логарифмируем формулу (14.8.1), а затем дифференцируем по t с учетом (14.8.2). В результате получим

$$\frac{\dot{g}(t, \tau)}{g(t, \tau)} = \frac{\dot{\Phi}_1(t)}{\Phi_1(t)} - \frac{\Phi_1^2(t)}{k(b_{11} + c_{11})},$$

где точкой, как и везде в дальнейшем, обозначено дифференцирование по t . Приняв во внимание (14.8.3) и положив для краткости

$$a = \frac{\dot{\Phi}_1(t)}{\Phi_1(t)}, \quad (14.8.4)$$

приведем полученное уравнение к виду

$$\dot{g}(t, \tau) = \left(a - \frac{\eta}{k}\right) g(t, \tau). \quad (14.8.5)$$

Таким образом, весовая функция оптимальной системы в данном случае удовлетворяет дифференциальному уравнению первого порядка. Следовательно, оптимальная система описывается уравнением первого порядка. Чтобы найти уравнение, связывающее входной и выходной сигналы оптимальной системы, достаточно заменить начальное условие $g(\tau, \tau) = \Phi_1^2(\tau)/k(b_{11} + c_{11})$ соответствующим слагаемым в (14.8.5), содержащим $\delta(t - \tau)$ (см. §§ 4.4 и 4.5). Так как при любом τ весовая функция $g(t, \tau)$ при переходе t через τ изменяется скачком от 0 до $\Phi_1^2(\tau)/k(b_{11} + c_{11})$, то $\delta(t - \tau)$ должна войти в правую часть (14.8.5) с коэффициентом $\Phi_1^2(t)/k(b_{11} + c_{11})$. Заменяя в полученном таким путем уравнении $\delta(t - \tau)$ входным сигналом $Z = Z(t)$, а весовую функцию $g(t, \tau)$ выходным сигналом оптимальной системы $W^* = W^*(t)$, получим дифференциальное уравнение оптимальной системы:

$$\dot{W}^* = \left(a - \frac{\eta}{k}\right) W^* + \frac{\eta}{k} Z. \quad (14.8.6)$$

Анализируем структуру оптимальной системы. Так как требуемый выходной сигнал $W = W(t)$ в данном случае равен $U_1 \Phi_1(t)$, то из (14.8.4) вытекает следующее уравнение для требуемого выходного сигнала:

$$\dot{W} = aW. \quad (14.8.7)$$

Сравнивая это уравнение с (14.8.6), видим, что оптимальная система получается из системы формирования требуемого выходного сигнала путем

замыкая ее отрицательной жесткой обратной связью и включения на входе усилителя с коэффициентом усиления η/k . На рис. 14.8.1 показана оптимальная система. Пунктирной линией обведена система формирования требуемого выходного сигнала.

Величина η , входящая в коэффициент усиления входного сигнала, может быть вычислена заранее и введена в систему с помощью соответствующего программного устройства. Однако часто ее предпочитают вычислять непосредственно в системе в процессе ее работы (или в модели системы при ее моделировании). При этом вычисление η можно осуществить путем интегрирования дифференциального уравнения, которому она удовлетворяет. Чтобы

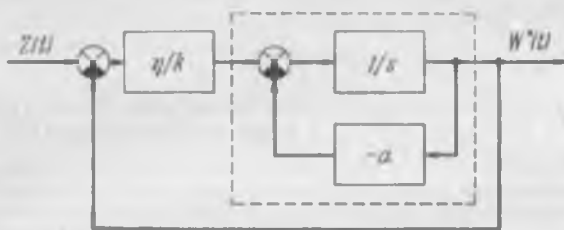


Рис. 14.8.1.

получить это уравнение, прологарифмируем формулу (14.8.3), а затем продифференцируем с учетом (14.8.2) и (14.8.4). В результате получим

$$\frac{\dot{\eta}}{\eta} = \frac{\dot{\Sigma}\varphi_1(t)}{\varphi_1(t)} - \frac{\varphi_1^2(t)}{k(b_{11} + c_{11})} = 2a - \frac{\eta}{k}.$$

Отсюда вытекает дифференциальное уравнение для η :

$$\dot{\eta} = 2a\eta - \frac{\eta^2}{k}. \quad (14.8.8)$$

Интегрируя совместно уравнения (14.8.6) и (14.8.8) (уже нелинейные), можно получить в каждый момент времени t одновременно выходной сигнал оптимальной системы W^* и соответствующий средний квадрат ошибки η .

Мы видим, что в данном случае оптимальная система строится очень просто и наглядно.

Рассмотрим общую задачу воспроизведения регулярного требуемого выходного сигнала

$$W(t) = \sum_{r=1}^N U_r \psi_r(t) \quad (14.8.9)$$

в случае, когда полезный входной сигнал определяется формулой

$$S(t) = \sum_{k=0}^{N-1} \beta_k W^{(k)}(t), \quad (14.8.10)$$

а помеха $N(t) = X(t)$ представляет собой белый шум интенсивности $G(t)$. Подставив выражение (14.8.9) в (14.8.10) и изменив порядок суммирования, представим входной сигнал в привычном нам

виде (13.5.14):

$$Z(t) = \sum_{r=1}^N U_r \varphi_r(t) + X(t), \quad (14.8.11)$$

где

$$\varphi_r(t) = \sum_{k=0}^{N-1} \beta_k \psi_r^{(k)}(t) \quad (r = 1, \dots, N). \quad (14.8.12)$$

В данном случае согласно (13.6.4) при $g^{(0)}(t, \tau) \equiv 0$ и (14.1.4) весовая функция оптимальной системы с бесконечной памятью ($T = t - t_0$) определяется формулой

$$g(t, \tau) = \frac{1}{G(\tau)} \sum_{r=1}^N \lambda_r(t) \varphi_r(\tau) \quad (-\infty < \tau \leq t), \quad (14.8.13)$$

где $\lambda_1(t), \dots, \lambda_N(t)$ — функции, определяемые системой уравнений (13.6.7), которая в данном случае, когда $b_{p0} = 0$ ($p = 1, \dots, N$), принимает вид

$$\sum_{q=1}^N (b_{pq} + c_{pq}) \lambda_q(t) = \psi_p(t) \quad (p = 1, \dots, N). \quad (14.8.14)$$

Формула (13.6.6), определяющая величины b_{pq} , в данном случае принимает вид

$$b_{pq} = \int_{t_0}^t \frac{\varphi_p(\tau) \varphi_q(\tau)}{G(\tau)} d\tau \quad (p, q = 1, \dots, N). \quad (14.8.15)$$

Легко убедиться непосредственной подстановкой, что требуемый выходной сигнал $W(t)$ удовлетворяет линейному дифференциальному уравнению

$$\begin{vmatrix} W & \psi_1(t) & \dots & \psi_N(t) \\ W' & \psi_1'(t) & \dots & \psi_N'(t) \\ \dots & \dots & \dots & \dots \\ W^{(N)} & \psi_1^{(N)}(t) & \dots & \psi_N^{(N)}(t) \end{vmatrix} = 0. \quad (14.8.16)$$

Совершенно так же, как в примере 14.8.1, дифференцируя формулу (14.8.13) N раз с учетом (14.8.14) и (14.8.15), можно убедиться в том, что весовая функция оптимальной системы в данном случае определяется линейным дифференциальным уравнением N -го порядка, достаточно просто связанным с уравнением (14.8.16). Однако необходимые для этого выкладки очень громоздки, утомительны и мешают пониманию существа дела. Чтобы упростить задачу, целесообразно, во-первых, заменить ее задачей одновременного воспроизведения сигнала W и его производных

$W', \dots, W(N-1)$, во-вторых, заменить уравнение (14.8.16) равноценной системой уравнений первого порядка и в-третьих, вместо формул (14.8.13), (14.8.15) и уравнений (14.8.14) исходить непосредственно из уравнения (13.5.37) и формулы (13.6.25) и произвести все выкладки в матричной форме. Заметим еще, что выкладки почти не изменятся, если в правые части системы уравнений формирования требуемых выходных сигналов добавить произвольные белые шумы. Это равносильно добавлению нерегулярных частей полезных сигналов, формируемых из белых шумов той же системой дифференциальных уравнений. Если нерегулярные части требуемых выходных сигналов формируются другой системой дифференциальных уравнений, то ее следует просто добавить к уравнениям, определяющим регулярные части этих сигналов.

В соответствии со сказанным рассмотрим задачу определения оптимальной многомерной линейной системы в случае, когда векторный требуемый выходной сигнал $W(t)$ определяется матричным дифференциальным уравнением

$$\dot{W} = A W + V(t), \quad (14.8.17)$$

а векторный входной сигнал определяется формулой

$$\dot{Z}(t) = B W(t) + V_1(t), \quad (14.8.18)$$

где A и B — матрицы коэффициентов (в общем случае переменных), $V(t)$ и $V_1(t)$ — некоррелированные векторные белые шумы, в общем случае нестационарные. На основании сказанного в § 13.5 можно без потери общности считать математические ожидания белых шумов $V(t)$ и $V_1(t)$ равными нулю. Матрицу интенсивностей и взаимных интенсивностей составляющих белого шума $V(t)$ обозначим G . Матрицу интенсивностей и взаимных интенсивностей составляющих помехи $V_1(t)$ обозначим G_1 . Тогда матрицы корреляционных и взаимных корреляционных функций белых шумов $V(t)$ и $V_1(t)$ выразятся соответственно формулами

$$K_v(t, \tau) = M[V(t) V(\tau)^T] = G \delta(t - \tau), \quad (14.8.19)$$

$$K_{v_1}(t, \tau) = M[V_1(t) V_1(\tau)^T] = G_1 \delta(t - \tau). \quad (14.8.20)$$

Для краткости мы не показываем явно зависимость матриц A , B , G и G_1 от времени t . Однако в случае, когда их аргумент обозначен другой буквой, мы будем указывать его соответствующим нижним индексом.

Для вывода дифференциальных уравнений оптимальной системы воспользуемся уравнением (13.5.37), определяющим матрицу весовых функций оптимальной системы, и формулой (13.6.25), определяющей матрицу вторых начальных моментов вектора ошиб-

ки оптимальной системы *). Положив в (13.5.37) и (13.6.25) интервал наблюдения T равным полному времени работы системы, $T = t - t_0$, перепишем эти равенства в виде

$$\int_{t_0}^t g(t, \tau) \Gamma_z(\tau, \sigma) d\tau = \Gamma_{wz}(t, \sigma) \quad (t_0 \leq \sigma \leq t), \quad (14.8.21)$$

$$H = \Gamma_w(t, t) - \int_{t_0}^t g(t, \tau) \Gamma_{wz}(t, \tau)^\tau d\tau. \quad (14.8.22)$$

Кроме этих соотношений нам понадобятся еще некоторые вспомогательные соотношения. Умножив уравнение (14.8.17) справа на $Z(\sigma)^\tau$ и имея в виду, что требуемый выходной сигнал $W(\sigma)$, а следовательно, и входной сигнал $Z(\sigma)$ в момент $\sigma < t$ не коррелирован со значением белого шума $V(t)$ в последующий момент времени t , получим

$$M[W(t) Z(\sigma)^\tau] = AM[W(t) Z(\sigma)^\tau], \quad \sigma < t,$$

или

$$\dot{\Gamma}_{wz}(t, \sigma) = A\Gamma_{wz}(t, \sigma), \quad \sigma < t. \quad (14.8.23)$$

Из (14.8.18) ввиду некоррелированности $Z(\sigma)$ с $V_1(t)$ при $\sigma < t$ следует

$$\Gamma_z(t, \sigma) = M[Z(t) Z(\sigma)^\tau] = BM[W(t) Z(\sigma)^\tau] = B\Gamma_{wz}(t, \sigma).$$

Подставив сюда выражение $\Gamma_{wz}(t, \sigma)$ из (14.8.21), получим

$$\Gamma_z(t, \sigma) = B \int_{t_0}^t g(t, \tau) \Gamma_z(\tau, \sigma) d\tau \quad (t_0 \leq \sigma < t). \quad (14.8.24)$$

Аналогично из той же формулы (14.8.18) следует

$$\Gamma_{zw}(\tau, \sigma) = M[Z(\tau) W(\sigma)^\tau] = B_\tau M[W(\tau) W(\sigma)^\tau],$$

или

$$\Gamma_{zw}(\tau, \sigma) = B_\tau \Gamma_w(\tau, \sigma), \quad (14.8.25)$$

где, в соответствии с нашим условием, B_τ — значение матрицы B при $t = \tau$.

Теперь мы можем перейти к выводу дифференциального уравнения оптимальной системы. Дифференцируя уравнение (14.8.21) по t и имея в виду, что от t зависит и подынтегральная функция

* Читатель, не владеющий основами матричной алгебры, может пропустить этот вывод и прочитать сразу первый абзац после уравнения (14.8.31), в котором формулируется окончательный результат, и после этого перейти к чтению первого абзаца после формулы (14.8.37).

и верхний предел интеграла, получим при $\sigma < t$

$$\int_{t_0}^t g(t, \tau) \Gamma_z(\tau, \sigma) d\tau + g(t, t) \Gamma_z(t, \sigma) = \dot{\Gamma}_{wz}(t, \sigma).$$

Подставив сюда выражения $\Gamma_z(t, \sigma)$ и $\dot{\Gamma}_{wz}(t, \sigma)$ из (14.8.24) и (14.8.23), будем иметь при $\sigma < t$

$$\int_{t_0}^t [g(t, \tau) + g(t, t) B g(t, \tau)] \Gamma_z(\tau, \sigma) d\tau = A \Gamma_{wz}(t, \sigma).$$

Вычтя из этого равенства уравнение (14.8.21), умноженное слева на матрицу A , получим

$$\int_{t_0}^t \{g(t, \tau) - [A - g(t, t) B] g(t, \tau)\} \Gamma_z(\tau, \sigma) d\tau = 0 \quad (t_0 \leq \sigma < t).$$

Таким образом, функция

$$h(t, \tau) = g(t, \tau) - [A - g(t, t) B] g(t, \tau) \quad (14.8.26)$$

удовлетворяет уравнению

$$\int_{t_0}^t h(t, \tau) \Gamma_z(\tau, \sigma) d\tau = 0 \quad (t_0 \leq \sigma < t). \quad (14.8.27)$$

Умножив это уравнение справа на $h(t, \sigma)^T$ и проинтегрировав по σ от t_0 до t , получим

$$D_t = \int_{t_0}^t \int_{t_0}^t h(t, \tau) \Gamma_z(\tau, \sigma) h(t, \sigma)^T d\tau d\sigma = 0.$$

Но на основании (7.2.63) D_t представляет собой матрицу вторых начальных моментов случайного вектора

$$U_t = \int_{t_0}^t h(t, \tau) Z(\tau) d\tau$$

при любом фиксированном t . Поэтому диагональные элементы матрицы D_t не могут быть отрицательными. Более того, если между составляющими белого шума $V_1(t)$ нет тождественных относительно t линейных зависимостей, то дисперсии составляющих случайного вектора U_t не могут быть равны нулю. Поэтому равенство (14.8.27) возможно при всех σ , $t_0 \leq \sigma < t$, только в том случае,

когда $h(t, \tau) = 0$. Учитывая это, получаем из (14.8.26)

$$\dot{g}(t, \tau) = [A - g(t, t) B] g(t, \tau). \quad (14.8.28)$$

Но в задачах практики между составляющими белого шума $V_1(t)$ не может быть тождественных линейных зависимостей, так как это означало бы возможность абсолютно точного, без помех, измерения некоторых линейных комбинаций составляющих требуемого выходного сигнала $W(t)$. Следовательно, из интегрального уравнения (14.8.21) вытекает в рассматриваемом случае дифференциальное уравнение (14.8.28). Таким образом, весовая функция оптимальной системы в данном случае определяется матричным дифференциальным уравнением (системой дифференциальных уравнений) (14.8.28). Чтобы получить дифференциальное уравнение оптимальной системы, следует заменить начальное условие для уравнения (14.8.28) соответствующим слагаемым $g(t, t) \delta(t - \tau)$ в правой части и после этого заменить δ -функцию $\delta(t - \tau)$ входным сигналом $Z = Z(t)$, а матрицу весовых функций $g(t, \tau)$ выходным сигналом оптимальной системы $W^* = W^*(t)$. В результате получим

$$\dot{W}^* = [A - g(t, t) B] W^* + g(t, t) Z. \quad (14.8.29)$$

Осталось определить коэффициент усиления $g(t, t)$ входного сигнала. Для этого подставим в первую формулу (13.5.38),

$$\Gamma_z(\tau, \sigma) = M[Z(\tau) Z(\sigma)^T],$$

выражения $Z(\tau)$ и $Z(\sigma)$ из (14.8.18). Тогда, учитывая, что транспонированное произведение матриц равно произведению соответствующих транспонированных матриц, взятых в обратном порядке, [83] и принимая во внимание некоррелированность $W(t)$ с $V_1(t)$, получим

$$\begin{aligned} \Gamma_z(\tau, \sigma) &= B_\tau M[W(\tau) W(\sigma)^T] B_\sigma^T + M[V_1(\tau) V_1(\sigma)^T] = \\ &= B_\tau \Gamma_w(\tau, \sigma) B_\sigma^T + G_{10} \delta(\tau - \sigma). \end{aligned}$$

Подставив это выражение в уравнение (14.8.21), приведем его к виду

$$g(t, \sigma) G_{10} + \int_0^t g(t, \tau) B_\tau \Gamma_w(\tau, \sigma) B_\sigma^T d\tau = \Gamma_{wz}(t, \sigma).$$

Положив здесь $\sigma = t$ и приняв во внимание, что на основании (14.8.25) и вследствие симметричности матрицы $\Gamma_w(t, t)$

$$\Gamma_{wz}(t, t) = \Gamma_{zw}(t, t)^T = \Gamma_w(t, t) B^T,$$

получим

$$g(t, t) G_1 = \left[\Gamma_w(t, t) - \int_{t_0}^t g(t, \tau) B \Gamma_w'(\tau, t) d\tau \right] B^T.$$

Но согласно (14.8.25)

$$B \Gamma_w(\tau, t) = \Gamma_{zw}(\tau, t) = \Gamma_{wz}(t, \tau)^T.$$

Подставив это выражение в предыдущую формулу, на основании (14.8.22) получим

$$g(t, t) G_1 = \left[\Gamma_w(t, t) - \int_{t_0}^t g(t, \tau) \Gamma_{wz}(t, \tau)^T d\tau \right] B^T = H B^T.$$

Отсюда находим *)

$$g(t, t) = H B^T G_1^{-1}. \quad (14.8.30)$$

Подставив это выражение в уравнение (14.8.29), приведем его к виду

$$\dot{W}^* = (A - H B^T G_1^{-1} B) W^* + H B^T G_1^{-1} Z. \quad (14.8.31)$$

Сравнивая уравнение (14.8.31) с (14.8.17), видим, что, как и в примере 14.8.1, оптимальная система получается из системы формирования требуемого выходного сигнала путем замыкания ее отрицательной обратной связью, содержащей усилитель с матричным коэффициентом усиления B , и установки перед ее входом усилителя с матричным коэффициентом $K = H B^T G_1^{-1}$ **). На рис. 14.8.2 представлена оптимальная система. Пунктирной линией обведена система формирования требуемого выходного сигнала.

Для вывода дифференциального уравнения для матрицы моментов второго порядка ошибки оптимальной системы H продифференцируем формулу (14.8.22) по t :

$$\begin{aligned} \dot{H} = & \frac{d}{dt} \Gamma_w(t, t) - \int_{t_0}^t \dot{g}(t, \tau) \Gamma_{wz}(t, \tau)^T d\tau - \\ & - \int_{t_0}^t g(t, \tau) \dot{\Gamma}_{wz}(t, \tau)^T d\tau - g(t, t) \Gamma_{wz}(t, t)^T. \end{aligned} \quad (14.8.32)$$

*) Так как между составляющими белого шума $V_1(t)$ в задачах практики не может быть линейных зависимостей, то обратная матрица G_1^{-1} всегда существует.

**) Мы называем для краткости *усилителем с матричным коэффициентом усиления* многомерную систему, состоящую из соответственно соединенных усилителей и сумматоров, вырабатывающую произведение мгновенного значения векторного сигнала на заданную матрицу. Примеры матричных усилителей обведены пунктирной линией слева и внизу справа на схеме оптимальной системы, представленной на рис. 14.8.4 (пример 14.8.4).

Но согласно (14.8.23)

$$\dot{\Gamma}_{wz}(t, \tau)^T = \Gamma_{wz}(t, \tau)^T A^T. \quad (14.8.33)$$

Матрица вторых начальных моментов требуемого выходного сигнала $\Gamma_w(t, t)$ определяется матричным дифференциальным уравнением (7.7.18):

$$-\frac{d}{dt} \Gamma_w(t, t) = A \Gamma_w(t, t) + \Gamma_w(t, t) A^T + G. \quad (14.8.34)$$

Далее, из (14.8.25) следует, что

$$\Gamma_{wz}(t, t)^T = \Gamma_{zw}(t, t) = B \Gamma_w(t, t). \quad (14.8.35)$$

Подставив в (14.8.32) выражения $d\Gamma_w(t, t)/dt$, $\dot{g}(t, \tau)$, $\dot{\Gamma}_{wz}(t, \tau)^T$

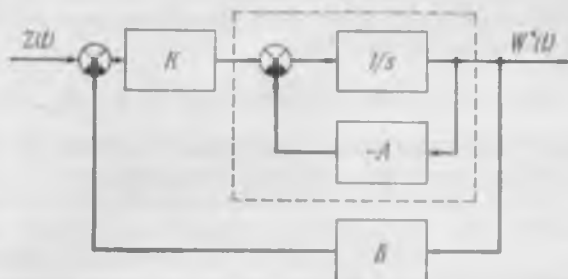


Рис. 14.8.2.

и $\Gamma_{wz}(t, t)^T$ соответственно из (14.8.34), (14.8.28), (14.8.33) и (14.8.35), получим

$$\begin{aligned} \dot{\Pi} = [A - g(t, t) B] & \left[\Gamma_w(t, t) - \int_{t_0}^t g(t, \tau) \Gamma_{wz}(t, \tau)^T d\tau \right] + \\ & + \left[\Gamma_w(t, t) - \int_{t_0}^t g(t, \tau) \Gamma_{wz}(t, \tau)^T d\tau \right] A^T + G \end{aligned}$$

или, на основании (14.8.22) и (14.8.30),

$$\dot{\Pi} = A\Pi + \Pi A^T - \Pi B^T G^{-1} B \Pi + G. \quad (14.8.36)$$

Уравнения (14.8.31) и (14.8.36) следует интегрировать при начальных условиях $W^*(t_0) = m_w(t_0)$, $\Pi(t_0) = K_w(t_0, t_0)$. При этом выходной сигнал оптимальной системы будет оптимальной несмещенной оценкой требуемого выходного сигнала в любой момент времени $t \geq t_0$. Действительно, так как $m_z = Bm_w$, то математическое ожидание m_w требуемого выходного сигнала,

удовлетворяющее уравнению

$$\dot{m}_w = A m_w,$$

является вместе с тем интегралом уравнения

$$\dot{m}_{w*} = A m_{w*} + H B^T G_1^{-1} (m_z - B m_{w*})$$

при начальном условии $m_{w*}(t_0) = m_w(t_0)$. Следовательно, при этом начальном условии $m_{w*}(t) = m_w(t)$ при $t \geq t_0$. Матрица H при таких начальных условиях представляет собой корреляционную матрицу вектора ошибки системы в любой момент времени.

Заметим еще, что если математическое ожидание белого шума $V(t)$ в (14.8.17) отлично от нуля, то для получения оптимальной несмещенной оценки требуемого выходного сигнала в любой момент времени $t \geq t_0$ достаточно добавить к правым частям уравнений (14.8.29) и (14.8.31) функцию m_v . В результате уравнение оптимальной системы примет вид

$$\dot{W}^* = (A - H B^T G_1^{-1} B) W^* + H B^T G_1^{-1} Z + m_v. \quad (14.8.37)$$

Действительно, математическое ожидание требуемого выходного сигнала, определяемое в этом случае уравнением

$$\dot{m}_w = A m_w + m_v,$$

в силу того, что $m_z = B m_w$, является также интегралом уравнения

$$\dot{m}_{w*} = A m_{w*} + H B^T G_1^{-1} (m_z - B m_{w*}) + m_v$$

при начальном условии $m_{w*}(t_0) = m_w(t_0)$.

В случае постоянных матриц A , B , G и G_1 в оптимальной системе существует установившийся режим, $H = \text{const}$. В этом случае уравнение (14.8.36) превращается в алгебраическое уравнение

$$A H + H A^T - H B^T G_1^{-1} B H + G = 0 \quad (14.8.38)$$

и оптимальная система, описываемая уравнением (14.8.31), стационарна.

Итак, мы показали, что в случае, когда помеха является белым шумом, нерегулярная часть требуемого выходного сигнала связана с белым шумом обыкновенным дифференциальным уравнением, а полезный входной сигнал представляет собой результат линейного алгебраического преобразования мгновенного значения требуемого выходного сигнала, оптимальная линейная система с бесконечной памятью описывается системой обыкновенных линейных дифференциальных уравнений. Этот результат был впервые получен Калманом и Бьюси [87, 88]. Поэтому оптимальные системы такого типа обычно называются *фильтрами Калмана*.

Калмэн показал также, что оптимальная система для экстраполяции сигнала $W(t)$ получается путем последовательного соединения системы, описываемой уравнением (14.8.31), с усилителем с матричным коэффициентом усиления $\Psi(t + \Delta)\Psi^{-1}(t)$, где $\Psi(t)$ — матрица фундаментальных решений однородного уравнения

$$\dot{\Psi} = A\Psi. \quad (14.8.39)$$

Результаты Калмэна и Бьюси были распространены на случай произвольной помехи, связанной с белым шумом обыкновенным дифференциальным уравнением, Брайсоном и Нохансеном [81] и Гулько и Новосельцевой [84, 85].

Изложим коротко метод Гулько и Новосельцевой. Предположим, что помеха $N(t)$ формируется из белого шума фильтром, описываемым системой конечного числа обыкновенных дифференциальных уравнений. Тогда входной сигнал $Z(t)$ определится формулой

$$Z(t) = BW(t) + N(t). \quad (14.8.40)$$

При этом возможны два случая. В первом случае оператор системы, обратной формирующему фильтру, является дифференциальным. Во втором случае он представляет собой сумму дифференциального оператора и интегрального оператора, весовая функция которого не содержит δ -функций и ее производных.

В первом случае, преобразуя $Z(t)$ дифференциальным оператором системы, обратной формирующему фильтру, и заменяя производные функции $W(t)$ их выражениями, вытекающими из (14.8.17), мы получим функцию $Z_1(t)$ вида (14.8.18) с белым шумом $V_1(t)$ в правой части (но, конечно, с другой матрицей B). При этом, само собой разумеется, полезный входной сигнал должен быть дифференцируемым большее число раз, чем помеха, вследствие чего белый шум $V(t)$ и его производные не войдут в сигнал $Z_1(t)$. Следовательно, в этом случае задача непосредственно сводится к применению изложенного метода. Оптимальная система в этом случае представляет собой последовательное соединение системы, обратной формирующему фильтру, и оптимальной системы для входного сигнала Z_1 [84].

Во втором случае, согласно идее Ф. Б. Гулько и Ж. А. Новосельцевой [85], систему, обратную формирующему фильтру, следует представить в виде параллельного соединения двух систем, одна из которых — непрерывная часть обратной системы — имеет весовую функцию, не содержащую δ -функций и их производных, а другая имеет чисто дифференциальный оператор (т. е. ее весовая функция представляет собой линейную комбинацию δ -функций и ее производных). Обозначив полезный выходной сигнал первой системы при входном сигнале Z через W' , получим для него

дифференциальное уравнение

$$\dot{W}' = A_1 B W + A_2 W' \quad (14.8.41)$$

(дифференциальное уравнение непрерывной части системы, обратной формирующему фильтру, при входном сигнале $BW(t)$). Тогда в результате преобразования входного сигнала $Z(t)$ системой, обратной формирующему фильтру, получим сигнал $Z_1(t)$ вида

$$Z_1(t) = B_1 W(t) + W'(t) + V_1(t),$$

где $V_1(t)$ — белый шум, из которого формируется помеха $N(t)$, а матрица B_1 получается в результате применения к произведению $BW(t)$ дифференциального оператора дифференцирующей части системы, обратной формирующему фильтру, и исключения производных функции $W(t)$ с помощью уравнения (14.8.17). Таким образом, и во втором случае задача сводится к применению метода Калмана путем ввода дополнительных компонент W' требуемого выходного сигнала и соответствующего расширения системы формирования требуемого выходного сигнала за счет добавления уравнения (14.8.41). Оптимальная система и в этом случае представляет собой последовательное соединение системы, обратной формирующему фильтру, и оптимальной системы для входного сигнала $Z_1(t)$, полученной изложенным методом.

Напомним еще раз, что оптимальные системы рассмотренного типа (калмановские фильтры) получаются лишь в частном (правда, широко распространенном) случае, когда оптимальная система ищется в классе систем с бесконечной памятью, а входной сигнал представляет собой сумму помехи, формируемой из белого шума, с результатом линейного преобразования мгновенного значения требуемого выходного сигнала.

Пример 14.8.2. В условиях примера 14.2.1 требуемый выходной сигнал $W(t) = X_1(t)$ определяется дифференциальным уравнением

$$\dot{W} = -\alpha W + V(t),$$

где $V(t)$ — стационарный белый шум со спектральной плотностью $D\alpha/\pi$ и, следовательно, с интенсивностью $G = 2D\alpha$. Помеха представляет собой белый шум $V_1(t) = X_2(t)$ интенсивности $G_1 = k$. Входным сигналом служит сумма $Z(t) = W(t) + V_1(t)$. Следовательно, в данном случае все векторы и матрицы представляют собой скаляры, $A = -\alpha$, $B = 1$, $N = \eta$. Уравнения (14.8.31) и (14.8.36) принимают соответственно вид

$$\dot{W}^* = -\left(\alpha + \frac{\eta}{k}\right) W^* + \frac{\eta}{k} Z(t),$$

$$\dot{\eta} = -2\alpha\eta - \frac{\eta^2}{k} + 2D\alpha.$$

Так как $t_0 = -\infty$, то система работает с установившимся средним квадратом ошибки $\eta = \text{const}$. Поэтому уравнение для η превращается в квадратное

уравнение

$$\eta^2 + 2\alpha k\eta - 2D\alpha k = 0.$$

Решая это уравнение, находим

$$\eta = \sqrt{\alpha^2 k^2 + 2D\alpha k} - \alpha k = (\beta - \alpha) k.$$

Подставив это выражение в уравнение оптимальной системы, получим

$$\dot{W}^* = -\beta W^* + (\beta - \alpha) Z(t).$$

Отсюда находим передаточную функцию оптимальной системы:

$$\Psi(s) = \frac{\beta - \alpha}{s + \beta}.$$

Этот результат, как и следовало ожидать, совпадает с полученным в примере 14.2.1.

Пример 14.8.3. Так как в условиях предыдущего примера уравнение (14.8.39) имеет решение $\Psi(t) = e^{-\alpha t}$, то оптимальная система для экстраполяции сигнала $W(t)$ на время Δ получается последовательным соединением

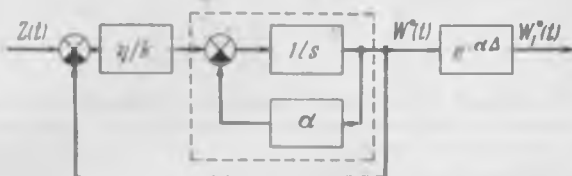


Рис. 14.8.3.

оптимальной системы предыдущего примера и усилителя с коэффициентом усиления $e^{-\alpha\Delta}$ (рис. 14.8.3). Выходной сигнал этого оптимального экстраполятора обозначен $W_1^*(t)$.

Пример 14.8.4. Полезный входной сигнал представляет собой сумму синусоиды заданной частоты ω_0 со случайными амплитудой и фазой и стационарной случайной функции $X_1(t)$ с корреляционной функцией

$$k(\tau) = D e^{-\alpha|\tau|} \left(\cos \omega_1 \tau + \frac{\alpha}{\omega_1} \sin \omega_1 |\tau| \right),$$

а помеха является стационарной случайной функцией $N(t)$ с корреляционной функцией $k_1(\tau) = D_1 e^{-\alpha_1|\tau|}$. Требуется найти оптимальную систему для воспроизведения (фильтрации) полезного сигнала.

Обозначим регулярную часть полезного сигнала W_1 ,

$$W_1(t) = U_1 \cos \omega_0 t + U_2 \sin \omega_0 t,$$

где U_1 и U_2 — случайные величины, а нерегулярную часть полезного сигнала $W_2 = W_2(t) = X_1(t)$. Легко проверить непосредственной подстановкой, что сигнал W_1 и его производная $W_2 = \dot{W}_1$ определяются системой дифференциальных уравнений

$$\dot{W}_1 = W_2, \quad \dot{W}_2 = -\omega_0^2 W_1. \quad (14.8.42)$$

Спектральная плотность случайной функции $W_3(t) = X_1(t)$ определяется формулой ([54], пример 5.3.3)

$$S_{X_1}(\omega) = \frac{2Da}{\pi} \frac{b^2}{b^4 + 2(a^2 - \omega^2)\omega^2 + \omega^4} = \frac{2Dab^2}{\pi} \frac{1}{|(i\omega)^2 + 2a(i\omega) + b^2|^2},$$

где $b^2 = a^2 + \omega_0^2$. Отсюда видно, что сигнал W_3 связан с белым шумом $V(t)$, имеющим спектральную плотность $s_0 = 2Dab^2/\pi$, дифференциальным уравнением

$$\ddot{W}_3 + 2a\dot{W}_3 + b^2W_3 = V(t).$$

Положив $W_4 = \dot{W}_3$, приведем это уравнение к системе двух уравнений первого порядка

$$\dot{W}_3 = W_4, \quad \dot{W}_4 = -b^2W_3 - 2aW_4 + V(t). \quad (14.8.43)$$

Таким образом, вектор W требуемых выходных сигналов с составляющими W_1 , W_2 , W_3 и W_4 определяется системой уравнений (14.8.42) и (14.8.43), которая может быть записана в матричной форме (14.8.17), где матрица A определяется формулой

$$A = \begin{bmatrix} 0 & 1 & 0 & 0 \\ -\omega_0^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & -b^2 & -2a \end{bmatrix}. \quad (14.8.44)$$

Векторный белый шум $V(t)$ в (14.8.17) в данном случае имеет нулевые первые три составляющие и только четвертая составляющая представляет собой скалярный белый шум интенсивности $2\pi s_0 = 4Dab^2$. Следовательно, матрица G имеет в данном случае все нулевые составляющие, кроме одной $G_{44} = 4Dab^2$.

По условиям задачи оптимальная система должна выделять полезный сигнал

$$S(t) = U_1 \cos \omega_0 t + U_2 \sin \omega_0 t + X_1(t) = W_1(t) + W_3(t)$$

из аддитивной смеси его с шумом $N(t)$. Таким образом, входной сигнал определяется в данном случае формулой

$$Z(t) = S(t) + N(t) = W_1(t) + W_3(t) + N(t).$$

Так как помеха $N(t)$ не является белым шумом, то изложенная теория для решения данной задачи непосредственно неприменима. Поэтому выполним предварительно такое преобразование входного сигнала $Z(t)$, при котором помеха преобразуется в белый шум. На основании результатов примера 7.6.2 такое преобразование осуществляется форсирующим звеном первого порядка с передаточной функцией $s + \alpha$. При подаче на вход такого звена сигнала $Z = Z(t)$ получится выходной сигнал

$$Z_1 = \dot{Z} + \alpha Z = \dot{W}_1 + \alpha W_1 + \dot{W}_3 + \alpha W_3 + V_1,$$

где $V_1 = \dot{N} + \alpha N$ — белый шум интенсивности $2D_1\alpha$. Подставив в эту формулу выражения $\dot{W}_1$ и $\dot{W}_3$ из (14.8.42) и (14.8.43), получим

$$Z_1 = \alpha W_1 + W_2 + \alpha W_3 + W_4 + V_1.$$

Эту формулу можно представить в матричном виде (14.8.18), где матрица B определяется формулой

$$B = \|\alpha \ 1 \ \alpha \ 1\|. \quad (14.8.45)$$

Предоставляем читателю самостоятельно написать уравнения (14.8.36) и (14.8.38) и вытекающие из них системы уравнений, определяющие элементы η_{pq} матрицы $\mathbf{H}$. Средний квадрат ошибки оптимальной системы равен в данном случае второму начальному моменту суммы ошибок на первом и третьем выходах калмановского фильтра, т. е. $\eta_{11} + 2\eta_{13} + \eta_{33}$.

Пример 14.8.5. Рассмотрим задачу предыдущего примера в случае помехи $N(t)$ с показательной-косинусной корреляционной функцией $k_1(\tau) =$

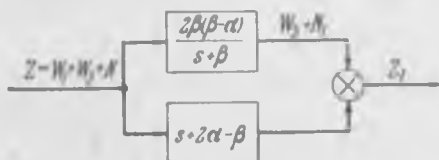


Рис. 14.8.5.

через обратную систему. Весовая функция этой обратной системы определяется формулой (4.4.40), полученной в примере 4.4.2, при $k = 1$, $\gamma = 0$, $a = \alpha$, $b = \beta$:

$$w^-(t, \tau) = 2\beta(\beta - \alpha)e^{-\beta(t-\tau)} + (2\alpha - \beta)\delta(t - \tau) + \delta'(t - \tau).$$

Согласно изложенному методу, представим систему с такой весовой функцией в виде параллельного соединения системы с весовой функцией

$$w_1^-(t, \tau) = 2\beta(\beta - \alpha)e^{-\beta(t-\tau)}$$

и системы с весовой функцией

$$w_2^-(t, \tau) = (2\alpha - \beta)\delta(t - \tau) + \delta'(t - \tau).$$

Первая система представляет собой аperiodическое звено с передаточной функцией $2\beta(\beta - \alpha)/(s + \beta)$, а вторая — форсирующее звено с передаточной функцией $s + 2\alpha - \beta$. Обозначим через W_5 результат преобразования первой системой полезного входного сигнала $W_1 + W_2$. Тогда при подаче на вход системы, обратной формирующему фильтру, входного сигнала $Z = Z(t)$ ее выходным сигналом будет

$$\begin{aligned} Z_1 &= (2\alpha - \beta)(W_1 + W_2) + \dot{W}_1 + \dot{W}_2 + W_3 + V_1 = \\ &= (2\alpha - \beta)(W_1 + W_2) + W_3 + W_4 + W_5 + V_1, \end{aligned}$$

где $V_1 = V_1(t)$ — белый шум, в который преобразуется помеха $N(t)$ (рис. 14.8.5). Сигнал W_5 , очевидно, определяется дифференциальным уравнением

$$\dot{W}_5 = 2\alpha(\beta - \alpha)(W_1 + W_2) - \beta W_5. \quad (14.8.46)$$

Таким образом, чтобы применить изложенный метод в данном случае, следует ввести дополнительную компоненту полезного сигнала W_5 , представляющую собой результат преобразования полезного входного сигнала $W_1 + W_2$ непрерывной частью системы, обратной формирующему фильтру, и таким путем расширить систему формирования требуемого выходного сигнала. Расширенная система формирования требуемого выходного сигнала описывается уравнениями (14.8.42), (14.8.43) и (14.8.46). В результате матричный коэффициент усиления B в цепи обратной связи будет равен

$$B = \begin{bmatrix} 2\alpha - \beta & 1 & 2\alpha - \beta & 1 & 1 \end{bmatrix}.$$

ОПРЕДЕЛЕНИЕ ОПТИМАЛЬНЫХ НЕЛИНЕЙНЫХ СИСТЕМ

§ 15.1. Общий метод определения оптимальной системы

Рассмотренные в предыдущих главах методы определения оптимальных линейных систем по критерию минимума средней квадратической ошибки дают возможность решать большое количество практических задач. Однако далеко не всегда оптимальная линейная система обеспечивает наилучшее решение задачи. Критерий же минимума средней квадратической ошибки, как было показано в § 13.2, является лишь одним из возможных, но не единственным критерием. Во многих задачах, как, например, в задаче обнаружения сигналов, приходится пользоваться другими критериями. Мы изложим здесь общий метод определения оптимальных систем, который позволяет при весьма общих условиях находить оптимальные системы, как линейные, так и нелинейные, практически по любым критериям.

Большое количество задач автоматики, теории информации и других областей науки приводит к следующей общей задаче. На вход системы поступает сигнал $Z(t)$, представляющий собой сумму полезного сигнала $S(t)$ и помехи $X(t)$, которую без потери общности можно считать имеющей математическое ожидание, тождественно равное нулю. Полезный сигнал представляет собой известную функцию времени t и некоторых неизвестных параметров $U_1, \dots, U_N$, которые могут быть случайными или просто неизвестными величинами. Для краткости записи мы будем обозначать все эти параметры одной буквой U . Тогда полезный сигнал выразится формулой

$$S(t) = \varphi(t, U) \quad (15.1.1)$$

На выходе системы требуется получить результат заданного преобразования полезного сигнала:

$$W(t) = \psi(t, U) = P_t \varphi(t, U), \quad (15.1.2)$$

где P — некоторый оператор, в общем случае нелинейный. Требуется найти оптимальную систему, выходной сигнал которой $W^*(t)$ возможно более точно воспроизводит требуемый выходной сигнал $W(t)$. Эта задача является частным случаем основной задачи теории оптимальных систем, сформулированной в § 13.1.

Для решения поставленной задачи в различных задачах практики приходится пользоваться различными вероятностными критериями точности. В § 13.2 мы видели, что большинство критериев, которыми приходится пользоваться на практике, могут быть представлены в виде

$$M[l(W, W^*)] = \min, \quad (15.1.3)$$

где $l(W, W^*)$ — некоторая заданная функция требуемого выходного сигнала W и действительного выходного сигнала W^* (функция потерь).

Оказывается, что все задачи такого рода можно решать одним общим методом при единственном условии, что помеха $X(t)$ распределена нормально и независима от параметров сигнала. Этот общий метод дает возможность находить оптимальные системы различных назначений по любым критериям вида (15.1.3). В отличие от задач §§ 13.5—13.7 и предыдущей главы, мы будем искать оптимальную систему среди всех возможных систем, имеющих в качестве входного сигнала случайную функцию $Z(t)$, не ограничиваясь классом линейных систем или каким-либо другим частным видом систем. Такая постановка задачи дает возможность находить системы, обеспечивающие абсолютный минимум математического ожидания функции потерь $l(W, W^*)$, определяющий предельную точность системы, которая теоретически может быть достигнута при данных условиях (потенциальную точность).

Будем производить вероятностное осреднение в формуле (15.1.3) в два приема: сначала по всем возможным значениям требуемого выходного сигнала при данной фиксированной реализации входного сигнала $Z(t)$, а затем по всем возможным реализациям случайной функции $Z(t)$. В результате первого осреднения мы получим условное математическое ожидание функции потерь относительно случайной функции $Z(t)$. В результате второго осреднения получим безусловное математическое ожидание функции потерь:

$$M[l(W, W^*)] = M[M[l(W, W^*) | Z]]. \quad (15.1.4)$$

Очевидно, что если мы найдем систему, обеспечивающую минимум условного математического ожидания функции потерь для каждой данной реализации z случайной функции Z :

$$M[l(W, W^*) | z] = \min, \quad (15.1.5)$$

то эта система будет обеспечивать и абсолютный минимум безусловного математического ожидания функции потерь, т. е. будет искомой оптимальной системой среди всех возможных систем.

Для нахождения условного математического ожидания в формуле (15.1.5) необходимо найти сначала условную плотность вероятности требуемого выходного сигнала $W(t)$ или, что одно и то же,

параметров сигнала U относительно входной случайной функции Z . Для этого предположим, как и в главах 13 и 14, что система должна вырабатывать выходной сигнал W^* в момент t по результатам наблюдения входной случайной функции $Z(\tau)$ в интервале времени $[t - T, t]$, и представим помеху $X(\tau)$ в интервале $[t - T, t]$ каким-либо каноническим разложением в действительной форме

$$X(\tau) = \sum_{v=1}^{\infty} V_v x_v(\tau) \quad (t - T \leq \tau \leq t), \quad (15.1.6)$$

где $V_1, V_2, \dots$ — некоррелированные действительные случайные величины, математические ожидания которых равны нулю, а $x_1(\tau), x_2(\tau), \dots$ — действительные координатные функции. Случайные величины V_v выражаются через случайную функцию $X(\tau)$ формулой ([54], § 6.2 или [53], § 60)

$$V_v = \int_{t-T}^t a_v(\tau) X(\tau) d\tau \quad (v = 1, 2, \dots), \quad (15.1.7)$$

где $a_1(\tau), a_2(\tau), \dots$ — действительные функции, удовлетворяющие совместно с координатными функциями $x_v(\tau)$ условию биортогональности

$$\int_{t-T}^t a_\mu(\tau) x_v(\tau) d\tau = \delta_{\mu v} \quad (\delta_{vv} = 1, \delta_{\mu v} = 0 \text{ при } \mu \neq v) \quad (15.1.8)$$

Так как помеха $X(\tau)$ по предположению распределена нормально, а случайные величины V_v представляют собой результат линейного преобразования случайной функции $X(\tau)$, то и случайные величины V_v распределены нормально и, следовательно, не только не коррелированы, но и независимы. Поэтому совместная плотность вероятности случайных величин $V_1, \dots, V_n$ при любом n выражается формулой

$$f_v(v_1, \dots, v_n) = \frac{1}{V(2\pi)^n D_1 \dots D_n} \exp \left\{ -\frac{1}{2} \sum_{v=1}^n \frac{v_v^2}{D_v} \right\}, \quad (15.1.9)$$

где $D_1, D_2, \dots$ — соответственно дисперсии случайных величин $V_1, V_2, \dots$.

Рассмотрим теперь случайные величины

$$Z_v = \int_{t-T}^t a_v(\tau) Z(\tau) d\tau \quad (v = 1, 2, \dots). \quad (15.1.10)$$

По условию входная случайная функция Z представляет собой сумму полезного сигнала и помехи:

$$Z(\tau) = \varphi(\tau, U) + X(\tau). \quad (15.1.11)$$

Подставляя это выражение в (15.1.10) и принимая во внимание (15.1.7), получим

$$Z_v = \int_{t-T}^t a_v(\tau) \varphi(\tau, U) d\tau + V_v \quad (v = 1, 2, \dots). \quad (15.1.12)$$

Полагая

$$\alpha_v(U) = \frac{1}{D_v} \int_{t-T}^t a_v(\tau) \varphi(\tau, U) d\tau \quad (v = 1, 2, \dots), \quad (15.1.13)$$

можем переписать формулу (15.1.12) в виде

$$Z_v = D_v \alpha_v(U) + V_v \quad (v = 1, 2, \dots). \quad (15.1.14)$$

Таким образом, случайные величины Z_v являются функциями случайных параметров сигнала U и соответствующих случайных величин V_v .

На основании (15.1.14) и (15.1.6) имеем

$$\sum_{v=1}^{\infty} Z_v x_v(\tau) = \sum_{v=1}^{\infty} D_v \alpha_v(U) x_v(\tau) + X(\tau) \quad (t-T \leq \tau \leq t). \quad (15.1.15)$$

Сравнивая эту формулу с (15.1.11), видим, что если функция $\varphi(\tau, U)$ представима в интервале наблюдения $t-T \leq \tau \leq t$ разложением по координатным функциям:

$$\varphi(\tau, U) = \sum_{v=1}^{\infty} D_v \alpha_v(U) x_v(\tau) \quad (t-T \leq \tau \leq t), \quad (15.1.16)$$

то входной сигнал $Z(\tau)$ также представим в интервале $t-T \leq \tau \leq t$ разложением по координатным функциям:

$$Z(\tau) = \sum_{v=1}^{\infty} Z_v x_v(\tau) \quad (t-T \leq \tau \leq t). \quad (15.1.17)$$

Формулы (15.1.10) и (15.1.17) устанавливают взаимно однозначное соответствие между случайной функцией $Z(\tau)$ и совокупностью случайных величин $Z_1, Z_2, \dots$. Каждой данной реализации случайной функции $Z(\tau)$ в силу формулы (15.1.10) соответствуют определенные возможные значения всех случайных величин $Z_1, Z_2, \dots$. Наоборот, каждой совокупности возможных значений случайных величин $Z_1, Z_2, \dots$ на основании формулы (15.1.17) соответствует определенная реализация случайной функции $Z(\tau)$. Поэтому условный закон распределения случайных параметров сигнала U

относительно случайной функции $Z(\tau)$ тождественно совпадает с условным законом распределения случайных параметров сигнала U относительно совокупности случайных величин $Z_1, Z_2, \dots$. Найдем сначала условную плотность вероятности $f_1(u | z_1, \dots, z_n)$ параметров сигнала U относительно конечного числа n случайных величин $Z_1, \dots, Z_n$. На основании известных формул теории вероятностей

$$f_1(u | z_1, \dots, z_n) = \frac{f(u) f_2(z_1, \dots, z_n | u)}{\int_{-\infty}^{\infty} f(u) f_2(z_1, \dots, z_n | u) du}, \quad (15.1.18)$$

где $f(u)$ — безусловная (*априорная*) плотность вероятности параметров сигнала U , которую мы будем считать заданной, $f_2(z_1, \dots, z_n | u)$ — условная плотность вероятности случайных величин $Z_1, \dots, Z_n$ относительно случайных параметров сигнала U . Для определения условной плотности вероятности случайных величин $Z_1, \dots, Z_n$ относительно случайных величин U заменим в (15.1.14) случайные величины U их возможными значениями u . Тогда случайные величины Z_v будут простыми линейными функциями соответствующих случайных величин V_v , и вследствие (15.1.9) мы будем иметь

$$\begin{aligned} f_2(z_1, \dots, z_n | u) &= \\ &= \frac{1}{V(2\pi)^n D_1 \dots D_n} \exp \left\{ -\frac{1}{2} \sum_{v=1}^n \frac{1}{D_v} [z_v - D_v \alpha_v(u)]^2 \right\} = \\ &= \frac{1}{V(2\pi)^n D_1 \dots D_n} \exp \left\{ -\sum_{v=1}^n \frac{z_v^2}{2D_v} + \right. \\ &\quad \left. + \sum_{v=1}^n \alpha_v(u) z_v - \frac{1}{2} \sum_{v=1}^n D_v \alpha_v^2(u) \right\}. \quad (15.1.19) \end{aligned}$$

Подставляя это выражение в (15.1.18) и производя сокращения, находим условную плотность вероятности случайных величин U относительно случайных величин $Z_1, \dots, Z_n$:

$$\begin{aligned} f_1(u | z_1, \dots, z_n) &= \\ &= \frac{f(u) \exp \left\{ \sum_{v=1}^n \alpha_v(u) z_v - \frac{1}{2} \sum_{v=1}^n D_v \alpha_v^2(u) \right\}}{\int_{-\infty}^{\infty} f(u) \exp \left\{ \sum_{v=1}^n \alpha_v(u) z_v - \frac{1}{2} \sum_{v=1}^n D_v \alpha_v^2(u) \right\} du}. \quad (15.1.20) \end{aligned}$$

Переходя к пределу при $n \rightarrow \infty$, находим условную плотность вероятности случайных величин U относительно совокупности всех случайных величин Z_v^* :

$$f_1(u | z_1, z_2, \dots) = \frac{f(u) \exp \left\{ \sum_{v=1}^{\infty} \alpha_v(u) z_v - \frac{1}{2} \sum_{v=1}^{\infty} D_v \alpha_v^2(u) \right\}}{\int_{-\infty}^{\infty} f(u) \exp \left\{ \sum_{v=1}^{\infty} \alpha_v(u) z_v - \frac{1}{2} \sum_{v=1}^{\infty} D_v \alpha_v^2(u) \right\} du} \quad (15.1.21)$$

Заметим теперь, что на основании (15.1.10)

$$\begin{aligned} \sum_{v=1}^{\infty} \alpha_v(u) Z_v &= \int_{t-T}^t \sum_{v=1}^{\infty} \alpha_v(u) a_v(\tau) Z(\tau) d\tau = \\ &= \int_{t-T}^t g(t, \tau, u) Z(\tau) d\tau, \end{aligned} \quad (15.1.22)$$

где для краткости положено

$$g(t, \tau, u) = \sum_{v=1}^{\infty} \alpha_v(u) a_v(\tau). \quad (15.1.23)$$

При этом мы указываем явно зависимость функции g от конца интервала наблюдения t (на основании формулы (15.1.13) величины $\alpha_v(u)$ зависят также от t).

Вводя функцию

$$\beta(u) = \sum_{v=1}^{\infty} D_v \alpha_v^2(u), \quad (15.1.24)$$

получим на основании (15.1.22) следующее выражение условной плотности вероятности случайных величин U относительно случайной функции $Z(\tau)$:

$$f_1(u | z) = \kappa(z) f(u) \exp \left\{ \int_{t-T}^t g(t, \tau, u) z(\tau) d\tau - \frac{1}{2} \beta(u) \right\}, \quad (15.1.25)$$

где

$$\begin{aligned} \kappa(z) &= \left[\int_{-\infty}^{\infty} f(u) \exp \left\{ \int_{t-T}^t g(t, \tau, u) z(\tau) d\tau - \right. \right. \\ &\quad \left. \left. - \frac{1}{2} \beta(u) \right\} du \right]^{-1} \end{aligned} \quad (15.1.26)$$

*) Доказательство существования предела читатель может найти в [53].

— постоянный нормирующий множитель, зависящий от реализации z случайной функции Z .

На основании (15.1.13) и (15.1.23) формулу (15.1.24) для функции $\beta(u)$ можно переписать в виде

$$\begin{aligned}\beta(u) &= \sum_{v=1}^{\infty} \alpha_v(u) \int_{t-T}^t a_v(\tau) \varphi(\tau, u) d\tau = \\ &= \int_{t-T}^t g(t, \tau, u) \varphi(\tau, u) d\tau. \quad (15.1.27)\end{aligned}$$

Формула (15.1.25) определяет условную плотность вероятности случайных параметров сигнала U относительно входного сигнала системы Z , т. е. *апостериорную* плотность вероятности параметров сигнала U , которая вполне определяется после наблюдения, когда становится известной реализация z случайной функции Z . Зная апостериорную плотность вероятности параметров сигнала U , можно вычислить апостериорные математические ожидания (условные математические ожидания относительно входного сигнала Z) любых функций параметров сигнала U . В частности, можно вычислить и условное математическое ожидание функции потерь $l(W, W^*)$, которая вследствие (15.1.2) является функцией U и выходного сигнала системы W^* . А именно на основании формул (15.1.2) и (15.1.25)

$$\begin{aligned}\rho(Z, W^*) &= M[l(W, W^*) | Z] = \\ &= \kappa(Z) \int_{-\infty}^{\infty} l(\psi(t, u), W^*) f(u) \times \\ &\times \exp \left\{ \int_{t-T}^t g(t, \tau, u) Z(\tau) d\tau - \frac{1}{2} \beta(u) \right\} du. \quad (15.1.28)\end{aligned}$$

Для полного решения задачи остается определить выходной сигнал оптимальной системы $W^* = W^*(t)$ при каждом данном значении t из условия минимума интеграла в (15.1.28). Этим условием полностью определяется оператор оптимальной системы. Действительно, условие минимума интеграла (15.1.28) полностью определяет последовательность математических действий, которые должна выполнить оптимальная система над входной функцией $Z(\tau)$, чтобы получить выходную функцию $W^*(t)$.

Таким образом, для нахождения оператора оптимальной системы необходимо определить весовую функцию $g(t, \tau, u)$ по формуле (15.1.23) и функцию $\beta(u)$ (зависящую также от t) по формуле (15.1.24) или (15.1.27). После этого оператор оптимальной системы определится как совокупность математических действий, которые

следует выполнить, чтобы при данном входном сигнале $Z(t)$ найти такую функцию $W^*(t)$, при которой интеграл в (15.1.28) достигает наименьшего значения в любой момент времени t .

Подставим теперь выражение (15.1.13) функции $a_v(u)$ в формулу (15.1.23). Тогда получим

$$g(t, \tau, u) = \sum_{v=1}^{\infty} \frac{a_v(\tau)}{D_v} \int_{t-T}^t a_v(\sigma) \varphi(\sigma, u) d\sigma. \quad (15.1.29)$$

Сравнивая эту формулу с (14.6.10) и принимая во внимание, что функции $a_v(\tau)$ в рассматриваемом случае действительны, убеждаемся в том, что весовая функция $g(t, \tau, u)$ является решением интегрального уравнения

$$\int_{t-T}^t K_x(\tau, \sigma) g(t, \tau, u) d\tau = \varphi(\sigma, u) \quad (t-T \leq \sigma \leq t), \quad (15.1.30)$$

где $K_x(\tau, \sigma)$ — корреляционная функция помехи $X(t)$. Формулы (15.1.13) и (15.1.23), равноценные формуле (15.1.29), дают решение интегрального уравнения (15.1.30) в виде бесконечного ряда. Однако в некоторых случаях может оказаться возможным определение весовой функции $g(t, \tau, u)$ непосредственным решением уравнения (15.1.30), например, одним из методов, изложенных в предыдущей главе.

Линейная система с весовой функцией $g(t, \tau, u)$ обладает интересным свойством: она осуществляет такое линейное преобразование входного сигнала $Z(\tau)$, что отношение сигнал/шум на выходе имеет наибольшее возможное значение в любой момент времени t . Действительно, на основании (15.1.1), (15.1.11) и (15.1.27) полезный сигнал и шум на выходе системы с весовой функцией $g(t, \tau, u)$ при данном значении u параметра сигнала U определяются соответственно формулами

$$s = \int_{t-T}^t g(t, \tau, u) \varphi(\tau, u) d\tau = \beta(u),$$

$$N = \int_{t-T}^t g(t, \tau, u) X(\tau) d\tau.$$

Дисперсия шума на выходе равна

$$D_n = \int_{t-T}^t \int_{t-T}^t g(t, \tau, u) g(t, \tau', u) K_x(\tau, \tau') d\tau d\tau'.$$

Но на основании (15.1.30) при любом τ' , $t-T \leq \tau' \leq t$,

$$\int_{t-T}^t g(t, \tau, u) K_x(\tau, \tau') d\tau = \varphi(\tau', u).$$

— постоянный нормирующий множитель, зависящий от реализации z случайной функции Z .

На основании (15.1.13) и (15.1.23) формулу (15.1.24) для функции $\beta(u)$ можно переписать в виде

$$\begin{aligned}\beta(u) &= \sum_{v=1}^{\infty} \alpha_v(u) \int_{t-T}^t a_v(\tau) \varphi(\tau, u) d\tau = \\ &= \int_{t-T}^t g(t, \tau, u) \varphi(\tau, u) d\tau. \quad (15.1.27)\end{aligned}$$

Формула (15.1.25) определяет условную плотность вероятности случайных параметров сигнала U относительно входного сигнала системы Z , т. е. *апостериорную* плотность вероятности параметров сигнала U , которая вполне определяется после наблюдения, когда становится известной реализация z случайной функции Z . Зная апостериорную плотность вероятности параметров сигнала U , можно вычислить апостериорные математические ожидания (условные математические ожидания относительно входного сигнала Z) любых функций параметров сигнала U . В частности, можно вычислить и условное математическое ожидание функции потерь $l(W, W^*)$, которая вследствие (15.1.2) является функцией U и выходного сигнала системы W^* . А именно на основании формул (15.1.2) и (15.1.25)

$$\rho(Z, W^*) = M[l(W, W^*) | Z] =$$

$$\begin{aligned}&= \kappa(Z) \int_{-\infty}^{\infty} l(\psi(t, u), W^*) f(u) \times \\ &\times \exp \left\{ \int_{t-T}^t g(t, \tau, u) Z(\tau) d\tau - \frac{1}{2} \beta(u) \right\} du. \quad (15.1.28)\end{aligned}$$

Для полного решения задачи остается определить выходной сигнал оптимальной системы $W^* = W^*(t)$ при каждом данном значении t из условия минимума интеграла в (15.1.28). Этим условием полностью определяется оператор оптимальной системы. Действительно, условие минимума интеграла (15.1.28) полностью определяет последовательность математических действий, которые должна выполнить оптимальная система над входной функцией $Z(\tau)$, чтобы получить выходную функцию $W^*(t)$.

Таким образом, для нахождения оператора оптимальной системы необходимо определить весовую функцию $g(t, \tau, u)$ по формуле (15.1.23) и функцию $\beta(u)$ (зависящую также от t) по формуле (15.1.24) или (15.1.27). После этого оператор оптимальной системы определится как совокупность математических действий, которые

следует выполнить, чтобы при данном входном сигнале $Z(\tau)$ найти такую функцию $W^*(t)$, при которой интеграл в (15.1.28) достигает наименьшего значения в любой момент времени t .

Подставим теперь выражение (15.1.13) функции $a_v(u)$ в формулу (15.1.23). Тогда получим

$$g(t, \tau, u) = \sum_{v=1}^{\infty} \frac{a_v(\tau)}{D_v} \int_{t-T}^t a_v(\sigma) \varphi(\sigma, u) d\sigma. \quad (15.1.29)$$

Сравнивая эту формулу с (14.6.10) и принимая во внимание, что функции $a_v(\tau)$ в рассматриваемом случае действительны, убеждаемся в том, что весовая функция $g(t, \tau, u)$ является решением интегрального уравнения

$$\int_{t-T}^t K_x(\tau, \sigma) g(t, \tau, u) d\tau = \varphi(\sigma, u) \quad (t-T \leq \sigma \leq t), \quad (15.1.30)$$

где $K_x(\tau, \sigma)$ — корреляционная функция помехи $X(t)$. Формулы (15.1.13) и (15.1.23), равноценные формуле (15.1.29), дают решение интегрального уравнения (15.1.30) в виде бесконечного ряда. Однако в некоторых случаях может оказаться возможным определение весовой функции $g(t, \tau, u)$ непосредственным решением уравнения (15.1.30), например, одним из методов, изложенных в предыдущей главе.

Линейная система с весовой функцией $g(t, \tau, u)$ обладает интересным свойством: она осуществляет такое линейное преобразование входного сигнала $Z(\tau)$, что отношение сигнал/шум на выходе имеет наибольшее возможное значение в любой момент времени t . Действительно, на основании (15.1.1), (15.1.11) и (15.1.27) полезный сигнал и шум на выходе системы с весовой функцией $g(t, \tau, u)$ при данном значении u параметра сигнала U определяются соответственно формулами

$$s = \int_{t-T}^t g(t, \tau, u) \varphi(\tau, u) d\tau = \beta(u),$$

$$N = \int_{t-T}^t g(t, \tau, u) X(\tau) d\tau.$$

Дисперсия шума на выходе равна

$$D_n = \int_{t-T}^t \int_{t-T}^t g(t, \tau, u) g(t, \tau', u) K_x(\tau, \tau') d\tau d\tau'.$$

Но на основании (15.1.30) при любом τ' , $t-T \leq \tau' \leq t$,

$$\int_{t-T}^t g(t, \tau, u) K_x(\tau, \tau') d\tau = \varphi(\tau', u).$$

Следовательно,

$$D_n = \int_{t-T}^t g(t, \tau', u) \varphi(\tau', u) d\tau' = \beta(u). \quad (15.1.31)$$

За отношение сигнал/шум можно принять отношение полезного сигнала к среднему квадратическому отклонению шума. Поэтому отношение сигнал/шум на выходе системы с весовой функцией $g(t, \tau, u)$ на основании полученных формул равно

$$\gamma = \frac{s}{\sqrt{D_n}} = \frac{\beta(u)}{\sqrt{\beta(u)}} = \sqrt{\beta(u)}. \quad (15.1.32)$$

Пусть теперь $h(t, \tau)$ — весовая функция произвольной линейной системы. Полезный сигнал и шум на выходе этой системы при том же входном сигнале $Z(\tau)$ определяются формулами

$$s_1 = \int_{t-T}^t h(t, \tau) \varphi(\tau, u) d\tau,$$

$$N_1 = \int_{t-T}^t h(t, \tau) X(\tau) d\tau.$$

Корреляционный момент выходных шумов N и N_1 систем с весовыми функциями $g(t, \tau, u)$ и $h(t, \tau)$ равен

$$k_{nn_1} = \int_{t-T}^t \int_{t-T}^t g(t, \tau, u) h(t, \tau') K_x(\tau, \tau') d\tau d\tau'$$

или, вследствие (15.1.30),

$$k_{nn_1} = \int_{t-T}^t h(t, \tau') \varphi(\tau', u) d\tau' = s_1. \quad (15.1.33)$$

Но в силу известного из теории вероятностей неравенства $|k_{nn_1}| \leq \sqrt{D_n D_{n_1}}$ (см. [54], § 3.9). Отсюда вследствие (15.1.33), (15.1.31) и (15.1.32) вытекает

$$|s_1| \leq \sqrt{D_n D_{n_1}} = \sqrt{\beta(u) D_{n_1}} = \gamma_1 \sqrt{D_{n_1}}.$$

Разделив это неравенство на положительную величину $\sqrt{D_{n_1}}$, получаем для отношения сигнал/шум на выходе системы с весовой функцией $h(t, \tau)$ неравенство

$$\gamma_1 = \frac{|s_1|}{\sqrt{D_{n_1}}} \leq \gamma.$$

Это неравенство показывает, что ни для какой линейной системы отношение сигнал/шум на выходе не может быть больше, чем у системы с весовой функцией $g(t, \tau, u)$, что и требовалось доказать. Равенство $\gamma_1 = \gamma$ достигается тогда и только тогда, когда шумы N и N_1 связаны при любом t линейной зависимостью, что возможно только в том случае, когда $h(t, \tau) = \theta(t) g(t, \tau, u)$, где $\theta(t)$ — произвольная неслучайная функция. Множитель $\theta(t)$, очевидно, не имеет существенного значения, так как умножение весовой функции линейной системы на $\theta(t)$ равноценно присоединению к выходу этой системы безынерционного линейного усилителя с коэффициентом усиления $\theta(t)$, а безынерционный линейный усилитель не изменяет отношения сигнал/шум.

Таким образом, оптимальная обработка входного сигнала при любом статистическом критерии оптимальности включает оптимальное линейное преобразование, обеспечивающее максимальное отношение сигнал/шум для любой реализации u параметра сигнала U . Фактически это означает, что оптимальная система всегда содержит множество оптимальных линейных систем, соответствующих всем возможным значениям u параметра сигнала U , каждая из которых максимизирует отношение сигнал/шум для соответствующего значения u параметра U .

Минимизацию интеграла (15.1.28) во многих задачах практики можно выполнить аналитически. В подобных случаях мы получим явную формулу, определяющую оператор оптимальной системы, т. е. явную зависимость выходной переменной оптимальной системы W^* от ее входной переменной Z . Если пайти минимум интеграла (15.1.28) в аналитической форме не удастся, то его можно найти приближенно, применяя численные методы. В таких случаях явной зависимости выходной переменной оптимальной системы W^* от входной переменной Z мы не получим и оператор оптимальной системы будет выражен в форме алгоритма.

Изложенный метод можно применять для нахождения как непрерывных, так и дискретных оптимальных систем. Для определения оптимальной дискретной системы достаточно, так же как в § 14.7, взять функции $a_v(\tau)$ в виде линейных комбинаций δ -функций:

$$a_v(\tau) = \sum_{h=1}^h a_{vh} \delta(\tau - t_h) \quad (v = 1, \dots, h). \quad (15.1.34)$$

Подставляя это выражение в (15.1.23) и полагая $t = t_h$,

$$g_{kh}(u) = \sum_{v=1}^h a_{vh} \alpha_v(u) \quad (k = 1, \dots, h), \quad (15.1.35)$$

получим

$$g(t_h, \tau, u) = \sum_{h=1}^h g_{kh}(u) \delta(\tau - t_h) \quad (15.1.36)$$

и

$$\int_{t_h-T}^{t_h} g(t_h, \tau, u) Z(\tau) d\tau = \sum_{h=1}^h g_{hh}(u) Z(t_h). \quad (15.1.37)$$

Интегральное уравнение (15.1.30) в этом случае превращается в систему линейных алгебраических уравнений, определяющих весовые коэффициенты $g_{hh}(u)$:

$$\sum_{h=1}^h K_x(t_h, t_q) g_{hh}(u) = \varphi(t_q, u) \quad (q = 1, \dots, h). \quad (15.1.38)$$

Изложенный метод обладает большой общностью и позволяет решать большое число различных задач практики. Для простоты изложения мы ограничились здесь лишь простейшим вариантом этого метода. Однако он легко распространяется на значительно более общие и сложные задачи, например на задачи определения оптимальных многомерных систем, имеющих любое число входов и выходов (см. [53], §§ 141—145). В последнем случае задача также сводится к минимизации интеграла (15.1.28), в котором $Z(\tau)$, W , W^* , $\psi(t, u)$ и $g(t, \tau, u)$ представляют собой векторы, причем вектор $g(t, \tau, u)$ определяется матричным уравнением (15.1.30), в котором векторы $g(t, \tau, u)$ и $\varphi(\sigma, u)$ представлены в виде матриц-столбцов, а $K_x(t, t')$ — матрица корреляционных и взаимных корреляционных функций составляющих векторной случайной функции $X(t)$. При этом произведение векторов $g(t, \tau, u)$ и $Z(\tau)$ в (15.1.28) и произведение векторов $g(t, \tau, u)$ и $\varphi(\tau, u)$ в (15.1.27), так же как и произведение векторов во всех предыдущих формулах, следует понимать как скалярные произведения соответствующих векторов.

Применяя изложенный метод, следует помнить, что буквой U во всех предыдущих выкладках обозначена в общем случае совокупность всех параметров сигнала $U_1, \dots, U_N$ и соответственно u представляет собой совокупность возможных значений $u_1, \dots, u_N$ параметров $U_1, \dots, U_N$, а все интегралы по u являются кратными интегралами по переменным $u_1, \dots, u_N$.

Случай, когда некоторые из параметров $U_1, \dots, U_N$ (или все) не являются случайными, а представляют собой неизвестные величины, которые могут иметь любые значения, можно рассматривать как предельный, когда эти параметры распределены равномерно в некоторых пределах, при неограниченном расширении этих пределов. В этом случае постоянная плотность вероятности равномерно распределенных параметров U в формулах (15.1.21) и (15.1.25) сокращается, а $f(u)$ представляет собой плотность вероятности одних только случайных параметров U . Таким образом, изложенный метод применим и в том случае, когда некоторые из парамет-

ров $U_1, \dots, U_N$ (или все) являются неслучайными величинами, которые могут иметь любые значения. При этом $f(u)$ в (15.1.25), (15.1.26) и (15.1.28) следует понимать как плотность вероятности тех параметров U , которые являются случайными величинами. Если все параметры $U_1, \dots, U_N$ являются неслучайными величинами, которые могут принимать любые значения, то $f(u) \equiv 1$.

Пример 15.1.1. Во многих задачах практики, в частности в задачах радиолокации, возникает необходимость определения фазы синусоидального сигнала, который принимается совместно с помехой. В таких случаях входной сигнал выражается формулой

$$Z(t) = U_1 \sin(\omega t + U_2) + X(t), \quad (15.1.39)$$

где U_1 — амплитуда полезного сигнала, случайная вследствие различных физических явлений, например флуктуаций отраженного от цели сигнала в задачах радиолокации, U_2 — неизвестная фаза, которую можно считать случайной величиной, равномерно распределенной в пределах периода $-\pi < u_2 < \pi$. Найдем оптимальную систему измерения фазы U_2 по критерию минимума вероятности того, что ошибка превзойдет по абсолютной величине некоторую данную величину a , предполагая, что случайные амплитуда U_1 и фаза U_2 полезного сигнала независимы, а помеха $X(t)$ представляет собой нормально распределенный белый шум интенсивности k , независимый от U_1 и U_2 .

Для решения поставленной задачи можно применить изложенный метод. В этом случае требуемый выходной сигнал W представляет собой неизвестную фазу U_2 , а функция потерь определяется формулой (13.2.8):

$$l(W, W^*) = \begin{cases} 0 & \text{при } |W^* - W| \leq a, \\ 1 & \text{при } |W^* - W| > a. \end{cases} \quad (15.1.40)$$

Решение интегрального уравнения (15.1.30) в данном случае может быть найдено по формуле (14.1.4), которая дает

$$g(t, \tau, u) = \frac{1}{k} \varphi(\tau, u) = \frac{u_1}{k} \sin(\omega\tau + u_2). \quad (15.1.41)$$

Следовательно, в данном случае

$$\int_{t-T}^t g(t, \tau, u) Z(\tau) d\tau = \frac{u_1}{k} \int_{t-T}^t Z(\tau) \sin(\omega\tau + u_2) d\tau \quad (15.1.42)$$

и

$$\beta(u) = \int_{t-T}^t g(t, \tau, u) \varphi(\tau, u) d\tau = \frac{u_1^2}{k} \int_{t-T}^t \sin^2(\omega\tau + u_2) d\tau.$$

Полагая интервал наблюдения T кратным полупериоду колебаний сигнала π/ω , получим после интегрирования

$$\beta(u) = \frac{u_1^2 T}{2k}. \quad (15.1.43)$$

Подставляя это выражение и выражение (15.1.42) в (15.1.28), принимая во внимание (15.1.40) и учитывая, что в данном случае $W = U_2$, $W^* = U_2^*$,

получим

$$\rho(Z, U_2^*) = M[l(U_2, U_2^*) | Z] = P(|U_2^* - U_2| > a | Z) =$$

$$= \frac{\kappa(Z)}{2\pi} \int_0^\infty h(u_1) \left[\int_{-\pi}^{U_2^* - a} \exp \left\{ \frac{u_1}{k} \int_{-T}^t Z(\tau) \sin(\omega\tau + u_2) d\tau - \frac{u_1^2 T}{4k} \right\} du_2 + \right. \\ \left. + \int_{U_2^* + a}^\pi \exp \left\{ \frac{u_1}{k} \int_{-T}^t Z(\tau) \sin(\omega\tau + u_2) d\tau - \frac{u_1^2 T}{4k} \right\} du_2 \right] du_1, \quad (15.1.44)$$

где $h(u_1)$ — плотность вероятности амплитуды полезного сигнала, представляющей собой существенно положительную случайную величину, вследствие чего $h(u_1) = 0$ при $u_1 < 0$. Дифференцируя выражение (15.1.44) по U_2^* и приравняв результат нулю, получаем уравнение для определения оптимальной оценки фазы U_2^* :

$$\frac{\partial \rho}{\partial U_2^*} = \frac{\kappa(Z)}{2\pi} \int_0^\infty h(u_1) \left[\exp \left\{ \frac{u_1}{k} \int_{-T}^t Z(\tau) \sin(\omega\tau + U_2^* - a) d\tau - \frac{u_1^2 T}{4k} \right\} - \right. \\ \left. - \exp \left\{ \frac{u_1}{k} \int_{-T}^t Z(\tau) \sin(\omega\tau + U_2^* + a) d\tau - \frac{u_1^2 T}{4k} \right\} \right] du_1 = 0. \quad (15.1.45)$$

Заметим теперь, что при любой плотности вероятности $h(u_1)$ интеграл

$$q(\eta) = \int_0^\infty h(u_1) \exp \left\{ \frac{u_1}{k} \eta - \frac{u_1^2 T}{4k} \right\} du_1 \quad (15.1.46)$$

является монотонно возрастающей функцией параметра η , так как при возрастании η подынтегральная функция при всех значениях u_1 увеличивается. Следовательно, значения интеграла (15.1.46) при двух значениях η_1, η_2 параметра η равны друг другу, $q(\eta_1) = q(\eta_2)$, тогда и только тогда, когда $\eta_1 = \eta_2$. Поэтому из (15.1.45) следует уравнение

$$\int_{-T}^t Z(\tau) \sin(\omega\tau + U_2^* - a) d\tau = \int_{-T}^t Z(\tau) \sin(\omega\tau + U_2^* + a) d\tau. \quad (15.1.47)$$

Решая это уравнение, находим оптимальную оценку фазы U_2^* :

$$\operatorname{tg} U_2^* = \frac{\int_{-T}^t Z(\tau) \cos \omega\tau d\tau}{\int_{-T}^t Z(\tau) \sin \omega\tau d\tau}. \quad (15.1.48)$$

Предоставляем читателю самостоятельно убедиться в том, что при $X(t) = 0$ формула (15.1.48) дает оценку фазы, совпадающую с истинной фазой U_2 . При $X(t) \neq 0$ оценка фазы U_2^* , определяемая формулой (15.1.48), является случайной величиной и вероятность ее совпадения с истинной фазой равна нулю. Однако в этом случае формула (15.1.48) обеспечивает определение фазы сигнала с минимальной вероятностью ошибки, превосходящей a .

Найденную оптимальную систему измерения фазы легко реализовать практически. Она состоит из устройства, в котором принимаемый сигнал

умножается на синусоидальный сигнал той же частоты с нулевой начальной фазой и на тот же сигнал, сдвинутый по фазе на $\pi/2$ (рис. 15.1.1). Выходные сигналы этого устройства поступают в интеграторы. Выходные переменные интеграторов подаются в преобразователь декартовых координат в полярные.

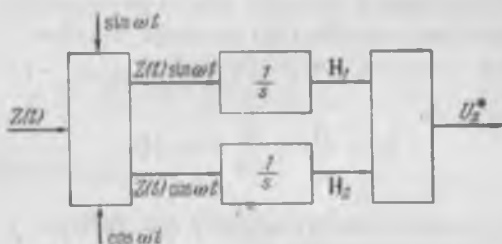


Рис. 15.1.1.

В этом преобразователе по данным интегралам в формуле (15.1.48), рассматриваемым как декартовы координаты точки плоскости, определяется полярный угол этой точки, который и представляет собой искомую оптимальную оценку фазы полезного сигнала U_1^* .

Заметим, что из (15.1.48) следует уравнение

$$\int_{t-T}^t Z(\tau) \cos(\omega\tau + U_1^*) d\tau = 0, \quad (15.1.49)$$

которое можно рассматривать как условие максимума выражения

$$k\left(\frac{U_1^*}{\omega}\right) = \frac{1}{T} \int_{t-T}^t Z(\tau) \sin \omega\left(\tau + \frac{U_1^*}{\omega}\right) d\tau. \quad (15.1.50)$$

Это выражение представляет собой временную взаимную корреляционную функцию *) принимаемого сигнала $Z(\tau)$ и функции $\sin \omega\tau$, определяющей форму полезного сигнала. Таким образом, оптимальная система определяет фазу из условия максимума временной взаимной корреляционной функции входного сигнала и полезного сигнала. Такой метод выделения сигналов, называемый обычно *корреляционным*, получил в радиотехнике широкое применение. Приемники, осуществляющие выделение сигнала корреляционным методом, обычно называются *корреляционными приемниками*.

*) Это название объясняется тем, что для эргодических стационарных и стационарно связанных случайных функций $X(\tau)$ и $Y(\tau)$, имеющих равные нулю математические ожидания, выражение

$$k(\sigma) = \frac{1}{T} \int_{t-T}^t X(\tau) Y(\tau + \sigma) d\tau$$

при достаточно большом T с вероятностью, как угодно близкой к единице, будет при любом t сколь угодно мало отличаться от взаимной корреляционной функции случайных функций $X(\tau)$ и $Y(\tau)$.

§ 15.2. Определение оптимальной системы в случае линейной зависимости входной функции от параметров сигнала

Оператор оптимальной системы легко находится в явной форме в случае, когда содержащийся во входном сигнале $Z(t)$ полезный сигнал линейно зависит от параметров $U_1, \dots, U_N$:

$$\varphi(t, U) = \sum_{r=1}^N U_r \varphi_r(t). \quad (15.2.1)$$

В этом случае решение интегрального уравнения (15.1.30) также линейно зависит от параметров $u_1, \dots, u_N$:

$$g(t, \tau, u) = \sum_{r=1}^N u_r g^{(r)}(t, \tau), \quad (15.2.2)$$

где $g^{(r)}(t, \tau)$ ($r = 1, \dots, N$) — весовые функции, определяемые интегральными уравнениями

$$\int_{t-T}^t K_x(\tau, \sigma) g^{(r)}(t, \tau) d\tau = \varphi_r(\sigma) (t-T \leq \sigma \leq t; r = 1, \dots, N). \quad (15.2.3)$$

Эти уравнения совпадают с уравнениями (13.6.2).

Подставляя выражение (15.2.1), в котором случайные величины $U_1, \dots, U_N$ заменены их возможными значениями $u_1, \dots, u_N$, и выражение (15.2.2) в (15.1.27), получим

$$\beta(u) = \sum_{p, q=1}^N b_{pq} u_p u_q, \quad (15.2.4)$$

где величины b_{pq} определяются формулой (13.6.6).

Подставляя выражение (15.2.1) в (15.1.13), получим

$$\alpha_v(U) = \sum_{r=1}^N U_r \frac{1}{D_v} \int_{t-T}^t a_v(\tau) \varphi_r(\tau) d\tau = \sum_{r=1}^N U_r \alpha_{vr}, \quad (15.2.5)$$

где α_{vr} — величины, определяемые формулой (14.6.14).

Подставляя в (15.1.23) выражение (15.2.5), в котором случайные величины $U_1, \dots, U_N$ заменены их возможными значениями $u_1, \dots, u_N$, и сравнивая полученное выражение с (15.2.2), получим для весовых функций $g^{(1)}(t, \tau), \dots, g^{(N)}(t, \tau)$ формулы (14.6.14), выведенные в § 14.6 методом канонических разложений.

На основании (15.2.2) и (15.2.4) выражение (15.1.28) принимает вид

$$\rho(Z, W^*) = \kappa(Z) \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} I(\psi(t, u_1, \dots, u_N), W^*) f(u_1, \dots, u_N) \times \\ \times \exp \left\{ \sum_{r=1}^N u_r \int_{t-T}^t g^{(r)}(t, \tau) Z(\tau) d\tau - \frac{1}{2} \sum_{p, q=1}^N b_{pq} u_p u_q \right\} du_1 \dots du_N. \quad (15.2.6)$$

Рассмотрим интеграл

$$I(t, \chi, \eta_1, \dots, \eta_N) = \int_{-\infty}^{\infty} \dots \int_{-\infty}^{\infty} I(\psi(t, u_1, \dots, u_N), \chi) f(u_1, \dots, u_N) \times \\ \times \exp \left\{ \sum_{r=1}^N \eta_r u_r - \frac{1}{2} \sum_{p, q=1}^N b_{pq} u_p u_q \right\} du_1 \dots du_N. \quad (15.2.7)$$

Этот интеграл является функцией параметров $\chi, \eta_1, \dots, \eta_N$ и времени t . Рассматривая его как функцию величины χ при фиксированных $t, \eta_1, \dots, \eta_N$, мы можем найти значение χ_0 переменной χ , при котором интеграл (15.2.7) достигает наименьшего значения:

$$I(t, \chi_0, \eta_1, \dots, \eta_N) \leq I(t, \chi, \eta_1, \dots, \eta_N).$$

Очевидно, что величина χ_0 зависит от значений величин $t, \eta_1, \dots, \eta_N$, т. е. является определенной функцией $t, \eta_1, \dots, \eta_N$:

$$\chi_0 = \chi_0(t, \eta_1, \dots, \eta_N). \quad (15.2.8)$$

Интеграл в (15.2.6) отличается от интеграла (15.2.7) только тем, что вместо $\chi, \eta_1, \dots, \eta_N$ в (15.2.6) входят соответственно величины

$$W^*, \int_{t-T}^t g^{(1)}(t, \tau) Z(\tau) d\tau, \dots, \int_{t-T}^t g^{(N)}(t, \tau) Z(\tau) d\tau.$$

Поэтому величина (15.2.6) при каждой данной реализации случайной функции $Z(\tau)$ будет иметь наименьшее возможное значение, если определить W^* формулой

$$W^*(t) = \chi_0 \left(t, \int_{t-T}^t g^{(1)}(t, \tau) Z(\tau) d\tau, \dots, \int_{t-T}^t g^{(N)}(t, \tau) Z(\tau) d\tau \right). \quad (15.2.9)$$

Эта формула дает явное выражение оператора оптимальной системы.

Оптимальная система в данном случае может быть представлена в виде N линейных систем с общим входом, выходные переменные которых подаются на входы нелинейного безынерционного устройства, вырабатывающего определенную функцию χ_0 (рис. 15.2.1).

Так как метод предыдущего параграфа дает оптимальную систему только в случае, когда входной полезный сигнал $\varphi(\tau, U)$

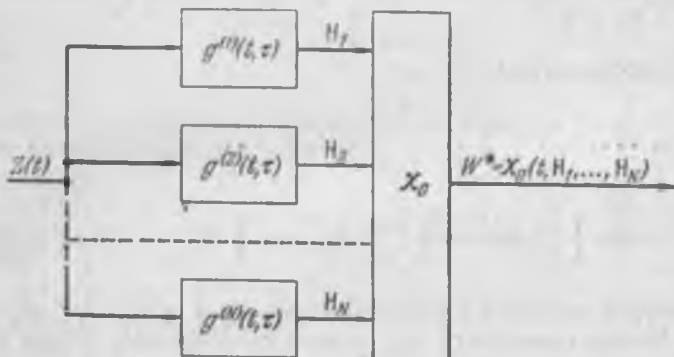


Рис. 15.2.1.

может быть представлен в интервале наблюдения разложением (15.1.16) по координатным функциям помехи $X(\tau)$, то формула (15.2.9) определяет оператор оптимальной системы только в случае, когда все функции $\varphi_1(\tau), \dots, \varphi_N(\tau)$ могут быть представлены в интервале наблюдения разложениями по координатным функциям помехи

$$\varphi_r(\tau) = \sum_{v=1}^{\infty} D_v \alpha_{vr} x_v(\tau) \quad (t - T \leq \tau \leq t; r = 1, \dots, N), \quad (15.2.10)$$

которые вытекают из (15.1.16), (15.2.1) и (15.2.5). Условия (15.1.16) и (15.2.10) обычно всегда выполняются в задачах практики.

Пример 15.2.1. Найти оптимальную систему для экстраполяции полезного сигнала, представляющего собой линейную функцию времени, по критерию минимума вероятности того, что ошибка превзойдет по абсолютной величине данную величину a , если помеха $X(t)$ независима от полезного сигнала, распределена нормально и имеет тождественно равное нулю математическое ожидание и корреляционную функцию

$$K_x(\tau, \sigma) = k_x(\tau - \sigma) = D e^{-\alpha|\tau - \sigma|}. \quad (15.2.11)$$

Начальное значение U_1 сигнала в момент $t = 0$ и скорость его изменения U_2 являются нормально распределенными случайными величинами с математическими ожиданиями, равными нулю.

В данном случае

$$Z(\tau) = U_1 + U_2\tau + X(\tau), \quad W(t) = U_1 + U_2(t + \Delta).$$

Интегральные уравнения (15.2.3), определяющие весовые функции $g^{(1)}$ и $g^{(2)}$, имеют вид (14.4.56). Эти уравнения были решены в примере 14.4.2. Их решение определяется формулами (14.4.57). Величины b_{11} , b_{12} и b_{22} для этого случая определяются формулами (14.4.58). Таким образом, для полного решения задачи остается найти минимум интеграла (15.2.7). Так как в данном случае функция потерь на основании изложенного в § 13.2 определяется формулой (15.1.40) и случайные величины U_1 и U_2 распределены нормально, то интеграл (15.2.7) имеет вид

$$I(\chi, \eta_1, \eta_2) = \frac{\sqrt{c_{11}c_{22} - c_{12}^2}}{2\pi} \int \int_{|u_1 + u_2(t+\Delta) - \chi| > a} \exp \left\{ \eta_1 u_1 + \eta_2 u_2 - \right. \\ \left. - \frac{1}{2} (c_{11} + b_{11}) u_1^2 - (c_{12} + b_{12}) u_1 u_2 - \frac{1}{2} (c_{22} + b_{22}) u_2^2 \right\} du_1 du_2. \quad (15.2.12)$$

Для нахождения минимума интеграла (15.2.12) можно применить обычный метод нахождения минимума функций. Для этого следует определить явно

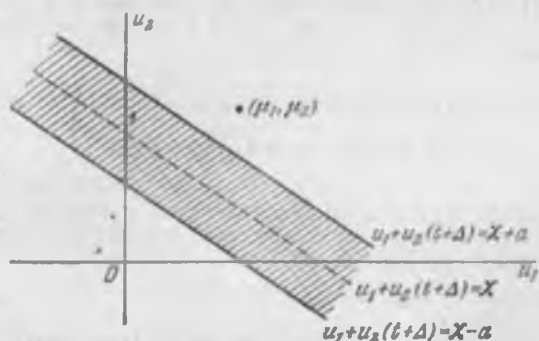


Рис. 15.2.2.

пределы интегрирования и приравнять нулю производную полученного интеграла по величине χ . Полученное в результате уравнение и определяет величину χ_0 как функцию t , η_1 и η_2 . Однако в данном случае минимум интеграла (15.2.12) можно найти значительно проще, пользуясь чисто вероятностными соображениями. А именно интеграл (15.2.12) пропорционален вероятности попадания случайной точки (U_1, U_2) в полосу, определяемую неравенствами (рис. 15.2.2)

$$-a < u_1 + u_2(t + \Delta) - \chi < a \quad (15.2.13)$$

при нормальном законе рассеивания. Но эта вероятность имеет минимальное значение в случае, когда центр рассеивания лежит в середине полосы. Таким образом, значение χ_0 величины χ , при котором интеграл (15.2.12) имеет минимальное значение, определяется условием, чтобы центр рассеивания лежал на прямой

$$u_1 + u_2(t + \Delta) - \chi_0 = 0. \quad (15.2.14)$$

Но центр рассеивания есть точка, в которой нормальная плотность вероятности имеет максимум. Следовательно, для определения координат центра

рассеивания μ_1, μ_2 достаточно приравнять нулю частные производные показателя степени в (15.2.12) по u_1 и u_2 при $u_1 = \mu_1, u_2 = \mu_2$. В результате получим уравнения

$$\left. \begin{aligned} (c_{11} + b_{11}) \mu_1 + (c_{12} + b_{12}) \mu_2 &= \eta_1, \\ (c_{12} + b_{12}) \mu_1 + (c_{22} + b_{22}) \mu_2 &= \eta_2. \end{aligned} \right\} \quad (15.2.15)$$

Решив эти уравнения и подставив полученные выражения в (15.2.14) вместо u_1, u_2 , найдем величину χ_0 :

$$\chi_0 = \mu_1 + \mu_2 (t + \Delta). \quad (15.2.16)$$

Так как уравнения (15.2.15) линейны относительно μ_1 и μ_2 , то они определяют μ_1 и μ_2 как линейные комбинации величин η_1, η_2 . Подставляя эти линейные комбинации в (15.2.16), получим χ_0 как линейную комбинацию η_1, η_2 :

$$\chi_0 = \lambda_1 \eta_1 + \lambda_2 \eta_2. \quad (15.2.17)$$

Для определения величин λ_1 и λ_2 достаточно подставить в (15.2.17) выражения η_1, η_2 из (15.2.15) и сравнить коэффициенты при μ_1 и μ_2 в полученном выражении и в выражении (15.2.16). Тогда получим следующие уравнения для определения величин λ_1 и λ_2 :

$$\left. \begin{aligned} (c_{11} + b_{11}) \lambda_1 + (c_{12} + b_{12}) \lambda_2 &= 1, \\ (c_{12} + b_{12}) \lambda_1 + (c_{22} + b_{22}) \lambda_2 &= t + \Delta. \end{aligned} \right\} \quad (15.2.18)$$

Итак, функция $\chi_0(t, \eta_1, \eta_2)$ определяется в данном случае формулой (15.2.17). Следовательно, выходной сигнал оптимальной системы согласно (15.2.9) определяется формулой

$$W^*(t) = \int_{t-T}^t [\lambda_1(t) g^{(1)}(t, \tau) + \lambda_2(t) g^{(2)}(t, \tau)] Z(\tau) d\tau. \quad (15.2.19)$$

Эта формула показывает, что оптимальная система в данном случае линейна и имеет весовую функцию

$$g(t, \tau) = \lambda_1(t) g^{(1)}(t, \tau) + \lambda_2(t) g^{(2)}(t, \tau). \quad (15.2.20)$$

Величины λ_1 и λ_2 здесь зависят от времени t , вследствие того что коэффициенты b_{11}, b_{12}, b_{22} уравнений (15.2.18), определяемые формулами (14.4.58), и правая часть второго уравнения (15.2.18) являются функциями времени. Сравнивая полученные результаты с результатами примера 14.4.2, убеждаемся в том, что найденная оптимальная система совпадает с оптимальной линейной системой, найденной по критерию минимума средней квадратической ошибки в примере 14.4.2. Таким образом, при нормальном законе распределения полезного сигнала и помехи оптимальная линейная система, найденная по критерию минимума средней квадратической ошибки в примере 14.4.2, является в то же время наилучшей и с точки зрения обеспечения минимума вероятности ошибки, превосходящей по абсолютной величине данную величину a , среди всех возможных систем.

§ 15.3. Определение оптимальных систем по критерию минимума средней квадратической ошибки

В случае критерия минимума средней квадратической ошибки функция потерь определяется формулой (см. § 13.2)

$$l(W, W^*) = (W^* - W)^2 \quad (15.3.1)$$

и формула (15.1.28) принимает вид

$$\rho(Z, W^*) = \kappa(Z) \int_{-\infty}^{\infty} [W^* - \psi(t, u)]^2 f(u) \times \\ \times \exp \left\{ \int_{t-T}^t g(t, \tau, u) Z(\tau) d\tau - \frac{1}{2} \beta(u) \right\} du. \quad (15.3.2)$$

Приравнявая нулю производную этого интеграла по W^* , получаем уравнение для определения выходного сигнала оптимальной системы W^* :

$$\int_{-\infty}^{\infty} [W^* - \psi(t, u)] f(u) \exp \left\{ \int_{t-T}^t g(t, \tau, u) Z(\tau) d\tau - \frac{1}{2} \beta(u) \right\} du = 0. \quad (15.3.3)$$

Решая это уравнение, находим

$$W^*(t) = \frac{\int_{-\infty}^{\infty} \psi(t, u) f(u) \exp \left\{ \int_{t-T}^t g(t, \tau, u) Z(\tau) d\tau - \frac{1}{2} \beta(u) \right\} du}{\int_{-\infty}^{\infty} f(u) \exp \left\{ \int_{t-T}^t g(t, \tau, u) Z(\tau) d\tau - \frac{1}{2} \beta(u) \right\} du}. \quad (15.3.4)$$

Таким образом, в случае критерия минимума средней квадратической ошибки метод § 15.1 дает явное выражение оператора оптимальной системы.

Правая часть формулы (15.3.4) на основании (15.1.25) и (15.1.26) представляет собой условное математическое ожидание функции $\psi(t, U)$, т. е. требуемого выходного сигнала $W(t)$ относительно входного сигнала $Z(t)$. Следовательно, формула (15.3.4) может быть переписана в виде

$$W^*(t) = M[W(t) | Z]. \quad (15.3.5)$$

Таким образом, оптимальной системой среди всех возможных систем в случае критерия минимума средней квадратической ошибки является система, которая дает на выходе условное математическое ожидание требуемого выходного сигнала W относительно входной функции Z , т. е. апостериорное математическое

ожидание требуемого выходного сигнала W . Вычисление этого условного математического ожидания в каждом конкретном случае производится по формуле (15.3.4).

Если требуемый выходной сигнал $W(t)$ является линейной функцией параметров $U_1, \dots, U_N$:

$$W(t) = \psi(t, U) = \sum_{r=1}^N U_r \psi_r(t), \quad (15.3.6)$$

то формула (15.3.5) принимает вид

$$W^*(t) = \sum_{r=1}^N M[U_r | Z] \psi_r(t). \quad (15.3.7)$$

Но величины

$$U_r^* = M[U_r | Z] \quad (r = 1, \dots, N) \quad (15.3.8)$$

согласно той же формуле (15.3.5) представляют собой оптимальные оценки параметров $U_1, \dots, U_N$ по критерию минимума средней квадратической ошибки. Следовательно, для нахождения оценки сигнала, представляющего собой линейную функцию параметров $U_1, \dots, U_N$, по критерию минимума средней квадратической ошибки достаточно найти оценки $U_1^*, \dots, U_N^*$ параметров $U_1, \dots, U_N$ и заменить этими оценками параметры $U_1, \dots, U_N$ в выражении (15.3.6) требуемого выходного сигнала:

$$W^*(t) = \sum_{r=1}^N U_r^* \psi_r(t). \quad (15.3.9)$$

Если и входная функция Z зависит линейно от параметров $U_1, \dots, U_N$ и априорное распределение этих параметров нормально, то вследствие (15.2.2) и (15.2.4) формула (15.1.25) дает нормальное апостериорное распределение параметров $U_1, \dots, U_N$, вследствие чего их апостериорные математические ожидания в формуле (15.3.8) могут быть определены из условия максимума апостериорной плотности вероятности параметров $U_1, \dots, U_N$:

$$f_1(u | Z) = \kappa(Z) \exp \left\{ \sum_{r=1}^N u_r \int_{t-T}^t g^{(r)}(t, \tau) Z(\tau) d\tau - \right. \\ \left. - \frac{1}{2} \sum_{p, q=1}^N b_{pq} u_p u_q - \frac{1}{2} \sum_{p, q=1}^N c_{pq} (u_p - m_p)(u_q - m_q) \right\}, \quad (15.3.10)$$

где $m_1, \dots, m_N$ — априорные математические ожидания случайных величин $U_1, \dots, U_N$ соответственно.

Приравнявая нулю частные производные показателя степени в (15.3.10) по $u_1, \dots, u_N$, получим следующую систему алгебраич-

ческих уравнений для определения условных математических ожиданий $U_1^*, \dots, U_N^*$ параметров $U_1, \dots, U_N$:

$$\sum_{q=1}^N (b_{pq} + c_{pq}) U_q^* = \sum_{q=1}^N c_{pq} m_q + \int_{t-T}^t g^{(p)}(t, \tau) Z(\tau) d\tau \quad (15.3.11)$$

$$(p = 1, \dots, N).$$

Таким образом, в случае линейной зависимости входного сигнала Z и требуемого выходного сигнала W от параметров сигнала и нормального априорного распределения параметров сигнала оценки $U_1^*, \dots, U_N^*$ параметров сигнала определяются системой линейных алгебраических уравнений (15.3.11). В результате решения этой системы уравнений величины $U_1^*, \dots, U_N^*$ получаются как линейные комбинации правых частей уравнений (15.3.11). Подставляя эти линейные комбинации в (15.3.9), получим оценку W^* требуемого выходного сигнала W в виде линейной комбинации правых частей уравнений (15.3.11):

$$W^*(t) = \int_{t-T}^t \sum_{p=1}^N \lambda_p(t) g^{(p)}(t, \tau) Z(\tau) d\tau + \sum_{p, q=1}^N c_{pq} m_q \lambda_p(t). \quad (15.3.12)$$

Для определения величин $\lambda_1(t), \dots, \lambda_N(t)$ достаточно подставить в формулу (15.3.12) вместо правых частей уравнений (15.3.11) их левые части и сравнить полученную формулу с (15.3.9). Тогда, приравнявая друг другу коэффициенты при $U_1^*, \dots, U_N^*$ в двух полученных выражениях $W^*(t)$ и учитывая, что $b_{pq} = b_{qp}$, $c_{pq} = c_{qp}$, получим для определения $\lambda_1, \dots, \lambda_N$ систему линейных алгебраических уравнений

$$\sum_{p=1}^N (b_{qp} + c_{qp}) \lambda_p = \psi_q(t) \quad (q = 1, \dots, N). \quad (15.3.13)$$

Формула (15.3.12) показывает, что в случае нормально распределенных сигнала и помехи критерий минимума средней квадратической ошибки всегда дает линейную оптимальную систему. Иными словами, при нормальном распределении сигнала и помехи оптимальная система среди всех возможных систем линейна. Однако в общем случае эта линейная система неоднородна и отличается от оптимальной однородной системы, которую дают методы, изложенные в предыдущих главах, тем, что в выражение ее выходного сигнала $W^*(t)$, кроме выходного сигнала линейной системы, соответствующей случаю нулевых математических ожиданий случайных величин $U_1, \dots, U_N$, входит еще независимое от входной функции Z слагаемое. Легко проверить, что это слагаемое представляет собой поправку на смещение, которая ликвидирует смещение и делает математическое ожидание

выходного сигнала системы W^* равным математическому ожиданию требуемого выходного сигнала W . Действительно, на основании (15.3.12), (15.2.1), (13.6.6), (15.3.13) и (15.3.6)

$$\begin{aligned} M[W^*(t)] &= \int_{t-T}^t \sum_{p=1}^N \lambda_p(t) g^{(p)}(t, \tau) \sum_{q=1}^N M[U_q] \varphi_q(\tau) d\tau + \\ &+ \sum_{p, q=1}^N c_{pq} m_q \lambda_p(t) = \sum_{q=1}^N m_q \sum_{p=1}^N b_{qp} \lambda_p(t) + \sum_{q=1}^N m_q \sum_{p=1}^N c_{qp} \lambda_p(t) = \\ &= \sum_{q=1}^N m_q \sum_{p=1}^N (b_{qp} + c_{qp}) \lambda_p(t) = \sum_{q=1}^N m_q \psi_q(t) = M[W(t)]. \end{aligned} \quad (15.3.14)$$

Доказанное предложение имеет весьма общий характер и справедливо не только для критерия минимума средней квадратической ошибки. А именно в случае линейной зависимости входного сигнала Z и требуемого выходного сигнала W от параметров $U_1, \dots, U_N$ и нормального распределения параметров $U_1, \dots, U_N$ и помехи $X(t)$ любой критерий вида (15.1.3), соответствующий функции потерь, зависящей только от ошибки системы $W^* - W$, дает линейную оптимальную систему. При этом для любой функции потерь, представляющей собой неубывающую функцию модуля ошибки $|W^* - W|$, получается та же оптимальная система, которую дает критерий минимума средней квадратической ошибки (см. [53], §§ 143 и 145).

§ 15.4. Оптимальные системы для обнаружения сигналов в шумах

В § 15.2 было показано, что задача обнаружения сигнала $\varphi(t, U)$ в шумах представляет собой частный случай задачи, решенной в § 15.1, когда функция потерь определяется формулой (13.2.14):

$$l(W, W^*) = \begin{cases} 1 & \text{при } W \neq 0, W^* \leq c, \\ \lambda & \text{при } W = 0, W^* > c, \\ 0 & \text{при } W = 0, W^* \leq c \text{ и при } W \neq 0, W^* > c. \end{cases} \quad (15.4.1)$$

Очевидно, что начало отсчета параметров сигнала U всегда можно выбрать так, чтобы нулевому значению одного из параметров сигнала соответствовало отсутствие сигнала. Тогда будем иметь $\varphi(\tau, 0) \equiv 0$. При этом вследствие (15.1.30), (15.1.27) будем иметь также

$$g(t, \tau, 0) \equiv 0, \quad \beta(0) = 0. \quad (15.4.2)$$

Обозначим через p вероятность присутствия сигнала, через $q = 1 - p$ вероятность отсутствия сигнала во входной функции Z , а через $h(u)$ условную плотность вероятности параметров сигнала при его наличии. Тогда априорная плотность вероятности параметров сигнала U выразится формулой

$$f(u) = q\delta(u) + ph(u). \quad (15.4.3)$$

Подставляя выражение (15.4.3) в формулу (15.1.28) и учитывая (15.4.2), получим

$$\begin{aligned} \rho(Z, W^*) = & \kappa(Z) \left[ql(0, W^*) + p \int_{-\infty}^{\infty} l(\varphi(t, u), W^*) h(u) \times \right. \\ & \left. \times \exp \left\{ \int_{t-T}^t g(t, \tau, u) Z(\tau) d\tau - \frac{1}{2} \beta(u) \right\} du \right]. \end{aligned} \quad (15.4.4)$$

Отсюда на основании (15.4.1) находим

$$\rho(Z, W^*) = \begin{cases} \kappa(Z) pQ(Z) & \text{при } W^* \leq c, \\ \kappa(Z) q\lambda & \text{при } W^* > c, \end{cases} \quad (15.4.5)$$

где для краткости через $Q(Z)$ обозначена величина

$$Q(Z) = \int_{-\infty}^{\infty} h(u) \exp \left\{ \int_{t-T}^t g(t, \tau, u) Z(\tau) d\tau - \frac{1}{2} \beta(u) \right\} du. \quad (15.4.6)$$

Для определения $\kappa(Z)$ подставим выражение (15.4.3) в (15.1.26). Тогда, принимая во внимание (15.4.2) и (15.4.6), найдем

$$\kappa(Z) = [q + pQ(Z)]^{-1}. \quad (15.4.7)$$

Подставляя это выражение в (15.4.5), получим

$$\rho(Z, W^*) = \begin{cases} \frac{pQ(Z)}{q + pQ(Z)} & \text{при } W^* \leq c, \\ \frac{q\lambda}{q + pQ(Z)} & \text{при } W^* > c. \end{cases} \quad (15.4.8)$$

Отсюда видно, что при любой реализации входной случайной функции Z величина $\rho(Z, W^*)$, рассматриваемая как функция W^* , представляет собой ступенчатую функцию с разными ординатами при $W^* \leq c$ и при $W^* > c$ (рис. 15.4.1). Наименьшее значение величины $\rho(Z, W^*)$ соответствует наименьшей из двух ординат. Поэтому, если при данной реализации случайной функции Z первая ступенька ниже второй (рис. 15.4.1, а), то система должна принять решение, что сигнал отсутствует, т. е. выдать выходной сигнал W^* ниже порогового уровня c . Если же при данной реализации случайной функции Z вторая ступенька ниже первой

(рис. 15.4.1, б), то система должна принять решение, что сигнал есть, т. е. выдать выходной сигнал W^* , превосходящий пороговый уровень c . Для обеспечения автоматического выбора наименьшей ступеньки достаточно принять за выходной сигнал системы W^*

любую возрастающую функцию величины $Q(Z)$. При этом в качестве порогового уровня c необходимо принять значение этой функции при аргументе, равном $\lambda q/p$:

$$W^* = \chi(Q(Z)), \quad c = \chi\left(\frac{\lambda q}{p}\right), \quad (15.4.9)$$

где $\chi(\xi)$ — произвольная возрастающая функция. Действительно, если $Q(Z) \leq \lambda q/p$, то первое значение функции $\rho(Z, W^*)$, определяемой формулой (15.4.8) (первая ступенька), будет меньше второго. При этом будет $\chi(Q(Z)) \leq \chi(\lambda q/p)$, т. е. $W^* \leq c$. Если $Q(Z) > \lambda q/p$, то второе значение функции $\rho(Z, W^*)$, определяемой формулой (15.4.8) (вторая ступенька), будет меньше первого. При этом будет $\chi(Q(Z)) > \chi(\lambda q/p)$, т. е. $W^* > c$.

Для оценки вероятности ошибки вспомним, что на основании сказанного в § 13.2 величина $\rho(Z, W^*)$ при $\lambda = 1$

совпадает с условной вероятностью ошибки при данной реализации входной случайной функции Z . Следовательно, условная вероятность ошибки при данной реализации входной случайной функции Z определяется формулой

$$p_{\text{ош}}(Z) = \begin{cases} \frac{pQ(Z)}{q + pQ(Z)} & \text{при } pQ(Z) \leq q\lambda, \\ \frac{q}{q + pQ(Z)} & \text{при } pQ(Z) > q\lambda. \end{cases} \quad (15.4.10)$$

Для решения задачи обнаружения сигнала по критерию минимума вероятности ошибки (критерий «идеального наблюдателя») следует на основании сказанного в § 13.2 положить во всех предыдущих формулах $\lambda = 1$. Для решения задачи обнаружения сигнала по критерию Неймана — Пирсона величину λ следует определить из условия равенства вероятности ложной тревоги (т. е. вероятности того, что выходной сигнал системы превысит пороговый уровень при отсутствии полезного сигнала во входной функции Z) заданной величине α . В общем виде вероятность ложной тревоги вычислить затруднительно, вследствие чего и определение λ



Рис. 15.4.1.

в общем виде затруднительно. Однако практически всегда полезный сигнал представляет собой линейную комбинацию известных функций с коэффициентами, зависящими от параметров U . В таких случаях вероятность ложной тревоги может быть вычислена сравнительно просто. Мы увидим, как это делается, на примерах.

Изложенный способ определения оптимальной обработки принимаемых сигналов с целью обнаружения в них полезных сигналов позволяет оценивать предельные возможности систем обнаружения, их максимальную теоретически достижимую помехоустойчивость, которая не может быть улучшена никакими техническими средствами при данных статистических свойствах сигналов и помех. Такая предельно возможная помехоустойчивость систем обнаружения сигналов обычно называется *потенциальной помехоустойчивостью*. Помехоустойчивость любой реальной системы обнаружения сигналов не может превзойти потенциальную помехоустойчивость. Основы теории потенциальной помехоустойчивости в радиотехнике были заложены В. А. Котельниковым [33]. Ее значение заключается главным образом в том, что она позволяет оценивать предельные качества приемных систем при данном характере полезных сигналов и помех и определяет возможные требования к проектируемым системам.

Пример 15.4.1. Найти оптимальную систему для обнаружения синусоидального сигнала со случайной амплитудой U_1 и известной фазой, если помеха представляет собой нормально распределенный белый шум интенсивности k . В данном случае

$$Z(t) = U_1 \sin \omega t + X(t), \quad W(t) = U_1 \sin \omega t. \quad (15.4.11)$$

Так как полезный сигнал является линейной функцией параметра U_1 , то весовая функция $g(t, \tau, u)$, согласно (15.2.2), определяется формулой

$$g(t, \tau, u) = u_1 g^{(1)}(t, \tau).$$

Для определения $g^{(1)}(t, \tau)$ можно воспользоваться формулой (14.1.4). В результате получим

$$g^{(1)}(t, \tau) = \frac{1}{k} \sin \omega \tau.$$

На основании (15.1.27) функция $\beta(u)$ определится формулой

$$\beta(u) = \frac{u_1^2}{k} \int_{t-T}^t \sin^2 \omega \tau d\tau = \frac{u_1^2 T}{2k},$$

если считать интервал наблюдения T кратным полупериоду π/ω . После этих предварительных вычислений, учитывая, что амплитуда U_1 существенно положительна, по формуле (15.4.6) находим

$$Q(Z) = \int_0^\infty h(u_1) \exp \left\{ \frac{u_1}{k} \int_{t-T}^t Z(\tau) \sin \omega \tau d\tau - \frac{u_1^2 T}{4k} \right\} du_1.$$

Очевидно, что в данном случае величина $Q(Z)$ при любой плотности вероятности $h(u_1)$ является возрастающей функцией интеграла в показателе степени. Следовательно, обратная функция определяет интеграл в показателе степени как возрастающую функцию $Q(Z)$. Поэтому выходной сигнал оптимальной системы можно определить формулой

$$W^*(t) = \int_{t-T}^t Z(\tau) \sin \omega \tau d\tau. \quad (15.4.12)$$

Таким образом, в рассматриваемом случае всегда существует оптимальная линейная система, независимая от закона распределения амплитуды сигнала. Весовая функция оптимальной системы обнаружения в данном случае равна $g(t - \tau) = \sin \omega \tau$ или $g(\xi) = \sin \omega(t - \xi)$.

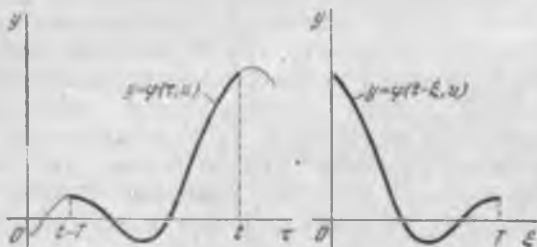


Рис. 15.4.2.

Стационарный линейный фильтр с весовой функцией $g(\xi)$, пропорциональной входному полезному сигналу $\phi(\tau, u)$, в котором аргумент τ заменен разностью $t - \xi$, а момент t фиксирован, называется в радиотехнике *согласованным* фильтром. На рис. 15.4.2 показаны графики функций $\phi(\tau, u)$ ($t - T \leq \tau \leq t$) и $g(t, \xi) = \phi(t - \xi, u)$ ($0 \leq \xi \leq T$). Из этого рисунка видно, что график весовой функции согласованного фильтра представляет собой отражение графика полезного сигнала от вертикальной прямой в конце интервала наблюдения t , смещенное влево на t . В общем случае линейный фильтр с весовой функцией $g(t, \tau, u)$, максимизирующий отношение сигнал/шум на выходе в любой момент времени t , очевидно, является обобщением согласованного фильтра на случай произвольных нестационарных входных сигналов и шумов.

Если применить для решения задачи обнаружения критерий минимума вероятности ошибки (критерий «идеального наблюдателя»), то $\lambda = 1$ и пороговый уровень сигнала s на основании второй формулы (15.4.9) определится уравнением

$$\frac{q}{p} = \int_0^{\infty} h(u_1) \exp \left\{ \frac{u_1}{k} c - \frac{u_1^2 T}{4k} \right\} du_1. \quad (15.4.13)$$

Таким образом, пороговый уровень сигнала в случае критерия «идеального наблюдателя» существенно зависит от априорной вероятности наличия сигнала p . А так как, к сожалению, априорная вероятность присутствия сигнала, как правило, неизвестна, то практическое применение полученного решения оказывается очень ограниченным.

В случае критерия Неймана — Пирсона пороговый уровень s можно определить непосредственно из условия равенства вероятности ложной тревоги, заданной величине α . Для вычисления вероятности ложной тревоги заметим,

что в случае отсутствия полезного сигнала $Z(\tau) = X(\tau)$ и

$$W(t) = \int_{t-T}^t X(\tau) \sin \omega \tau d\tau.$$

А так как $X(\tau)$ — нормально распределенный белый шум, то выходной сигнал W^* , как результат линейного преобразования случайной функции $X(\tau)$, также распределен нормально, причем его математическое ожидание равно нулю, а дисперсия определяется формулой

$$D[W^*] = \int_{t-T}^t \int_{t-T}^t k \delta(\sigma - \tau) \sin \omega \sigma \sin \omega \tau d\sigma d\tau = k \int_{t-T}^t \sin^2 \omega \tau d\tau = \frac{kT}{2}.$$

Зная условный закон распределения выходного сигнала W^* при отсутствии полезного сигнала, можно вычислить вероятность ложной тревоги по известным формулам теории вероятностей:

$$\begin{aligned} P_{\text{лт}} = P(W^* > c | W=0) &= \frac{1}{\sqrt{\pi kT}} \int_c^{\infty} e^{-\frac{v^2}{kT}} dv = \\ &= \frac{1}{\sqrt{2\pi}} \int_{c\sqrt{\frac{2}{kT}}}^{\infty} e^{-\frac{z^2}{2}} dz = \frac{1}{2} - \Phi\left(c\sqrt{\frac{2}{kT}}\right). \end{aligned} \quad (15.4.14)$$

Приравнявая вероятность ложной тревоги заданной величине α , получаем следующее уравнение для нахождения порогового уровня c :

$$\Phi\left(c\sqrt{\frac{2}{kT}}\right) = \frac{1}{2} - \alpha. \quad (15.4.15)$$

Эта формула показывает, что в случае критерия Неймана — Пирсона пороговый уровень сигнала c не зависит от априорной вероятности присутствия сигнала p . Вследствие этого критерий Неймана — Пирсона получил широкое распространение при решении практических задач обнаружения различных сигналов.

Формула (15.4.14) может также служить для вычисления вероятности ложной тревоги в случае критерия «идеального наблюдателя», если подставить в нее значение c , полученное из уравнения (15.4.13).

Для вычисления вероятности пропуска сигнала заметим, что при наличии полезного сигнала выходной сигнал оптимальной системы W^* , согласно (15.4.11) и (15.4.12), равен

$$W^*(t) = \frac{U_1 T}{2} + \int_{t-T}^t X(\tau) \sin \omega \tau d\tau.$$

Отсюда видно, что условный закон распределения выходного сигнала W^* при данном значении u_1 амплитуды полезного сигнала U_1 отличается от условного закона распределения при отсутствии полезного сигнала только тем, что при $U_1 = u_1$ математическое ожидание W^* равно $u_1 T/2$, а не нулю.

Очевидно, что в данном случае величина $Q(Z)$ при любой плотности вероятности $h(u_1)$ является возрастающей функцией интеграла в показателе степени. Следовательно, обратная функция определяет интеграл в показателе степени как возрастающую функцию $Q(Z)$. Поэтому выходной сигнал оптимальной системы можно определить формулой

$$W^*(t) = \int_{t-T}^t Z(\tau) \sin \omega \tau d\tau. \quad (15.4.12)$$

Таким образом, в рассматриваемом случае всегда существует оптимальная линейная система, независимая от закона распределения амплитуды сигнала. Весовая функция оптимальной системы обнаружения в данном случае равна $g(t - \tau) = \sin \omega \tau$ или $g(\xi) = \sin \omega(t - \xi)$.

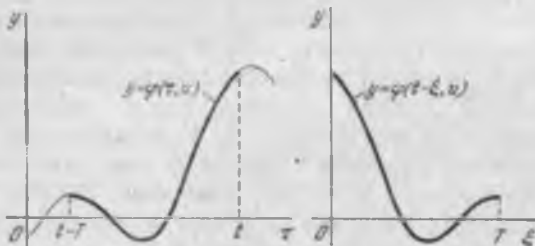


Рис. 15.4.2.

Стационарный линейный фильтр с весовой функцией $g(\xi)$, пропорциональной входному полезному сигналу $\varphi(\tau, u)$, в котором аргумент τ заменен разностью $t - \xi$, а момент t фиксирован, называется в радиотехнике *согласованным* фильтром. На рис. 15.4.2 показаны графики функций $\varphi(\tau, u)$ ($0 \leq \tau \leq t$) и $g(t - \xi) = \varphi(t - \xi, u)$ ($0 \leq \xi \leq T$). Из этого рисунка видно, что график весовой функции согласованного фильтра представляет собой отражение графика полезного сигнала от вертикальной прямой в конце интервала наблюдения t , смещенное влево на t . В общем случае линейный фильтр с весовой функцией $g(t, \tau, u)$, максимизирующий отношение сигнал/шум на выходе в любой момент времени t , очевидно, является обобщением согласованного фильтра на случай произвольных нестационарных входных сигналов и шумов.

Если применить для решения задачи обнаружения критерий минимума вероятности ошибки (критерий «идеального наблюдателя»), то $\lambda = 1$ и пороговый уровень сигнала c на основании второй формулы (15.4.9) определится уравнением

$$\frac{q}{p} = \int_0^{\infty} h(u_1) \exp \left\{ \frac{u_1}{k} c - \frac{u_1^2 T}{4k} \right\} du_1. \quad (15.4.13)$$

Таким образом, пороговый уровень сигнала в случае критерия «идеального наблюдателя» существенно зависит от априорной вероятности наличия сигнала p . А так как, к сожалению, априорная вероятность присутствия сигнала, как правило, неизвестна, то практическое применение полученного решения оказывается очень ограниченным.

В случае критерия Неймана — Пирсона пороговый уровень c можно определить непосредственно из условия равенства вероятности ложной тревоги, заданной величине α . Для вычисления вероятности ложной тревоги заметим,

что в случае отсутствия полезного сигнала $Z(\tau) = X(\tau)$ и

$$W(t) = \int_{t-T}^t X(\tau) \sin \omega \tau d\tau.$$

А так как $X(\tau)$ — нормально распределенный белый шум, то выходной сигнал W^* , как результат линейного преобразования случайной функции $X(\tau)$, также распределен нормально, причем его математическое ожидание равно нулю, а дисперсия определяется формулой

$$D[W^*] = \int_{t-T}^t \int_{t-T}^t k\delta(\sigma - \tau) \sin \omega \sigma \sin \omega \tau d\sigma d\tau = k \int_{t-T}^t \sin^2 \omega \tau d\tau = \frac{kT}{2}.$$

Зная условный закон распределения выходного сигнала W^* при отсутствии полезного сигнала, можно вычислить вероятность ложной тревоги по известным формулам теории вероятностей:

$$\begin{aligned} P_{\text{лт}} = P(W^* > c | W = 0) &= \frac{1}{\sqrt{\pi kT}} \int_c^{\infty} e^{-\frac{v^2}{kT}} dv = \\ &= \frac{1}{\sqrt{2\pi}} \int_{c\sqrt{\frac{2}{kT}}}^{\infty} e^{-\frac{z^2}{2}} dz = \frac{1}{2} - \Phi\left(c\sqrt{\frac{2}{kT}}\right). \end{aligned} \quad (15.4.14)$$

Приравнявая вероятность ложной тревоги заданной величине α , получаем следующее уравнение для нахождения порогового уровня c :

$$\Phi\left(c\sqrt{\frac{2}{kT}}\right) = \frac{1}{2} - \alpha. \quad (15.4.15)$$

Эта формула показывает, что в случае критерия Неймана — Пирсона пороговый уровень сигнала c не зависит от априорной вероятности присутствия сигнала p . Вследствие этого критерий Неймана — Пирсона получил широкое распространение при решении практических задач обнаружения различных сигналов.

Формула (15.4.14) может также служить для вычисления вероятности ложной тревоги в случае критерия «идеального наблюдателя», если подставить в нее значение c , полученное из уравнения (15.4.13).

Для вычисления вероятности пропуска сигнала заметим, что при наличии полезного сигнала выходной сигнал оптимальной системы W^* , согласно (15.4.11) и (15.4.12), равен

$$W^*(t) = \frac{U_1 T}{2} + \int_{t-T}^t X(\tau) \sin \omega \tau d\tau.$$

Отсюда видно, что условный закон распределения выходного сигнала W^* при данном значении u_1 амплитуды полезного сигнала U_1 отличается от условного закона распределения при отсутствии полезного сигнала только тем, что при $U_1 = u_1$ математическое ожидание W^* равно $u_1 T/2$, а не нулю.

Следовательно, вероятность пропуска сигнала определяется формулой

$$P_{\text{пр}} = P(W^* \leq c | W \neq 0) = \int_0^\infty \frac{h(u_1) du_1}{\sqrt{\pi k T}} \int_{-\infty}^c e^{-\frac{(2v - u_1 T)^2}{4kT}} dv = \\ = \int_0^\infty h(u_1) \left[\frac{1}{2} + \Phi\left(\frac{2c - u_1 T}{\sqrt{2kT}}\right) \right] du_1. \quad (15.4.16)$$

Формула (15.4.15) показывает, что величина $v = c \sqrt{2/kT}$ зависит только от заданной вероятности ложной тревоги α . С другой стороны, отношение значения амплитуды принимаемого сигнала u_1 к $\sqrt{k}$ может быть принято за меру отношения сигнала к шуму $\gamma = u_1/\sqrt{k}$. Поэтому формулу (15.4.16) для вероятности пропуска сигнала можно переписать в виде

$$P_{\text{пр}} = \sqrt{k} \int_0^\infty h(\gamma \sqrt{k}) \left[\frac{1}{2} + \Phi\left(v - \gamma \sqrt{\frac{T}{2}}\right) \right] d\gamma. \quad (15.4.17)$$

Таким образом, при данной вероятности ложной тревоги вероятность пропуска сигнала зависит от закона распределения отношения сигнала к шуму γ и от времени наблюдения сигнала T . Графики условной вероятности пропуска сигнала при данном γ в зависимости от γ и T при данной вероятности ложной тревоги α и в зависимости от γ и α при данном времени наблюдения T позволяют весьма просто и наглядно оценивать потенциальную помехоустойчивость систем обнаружения и их предельные возможности.

Пример 15.4.2. В условиях предыдущего примера найти оптимальную систему для обнаружения синусоидального сигнала со случайной амплитудой U_1 и неизвестной фазой. В данном случае

$$Z(t) = U_1 \sin(\omega t + U_2) + X(t), \quad W(t) = U_1 \sin(\omega t + U_2), \quad (15.4.18)$$

где U_2 — неизвестная начальная фаза, которая может иметь любое значение в пределах периода $-\pi < U_2 < \pi$. Это все равно, что U_2 является случайной величиной, равномерно распределенной в интервале $-\pi < u_2 < \pi$. Поэтому в данном случае можно принять

$$f(u) = q \delta(u_1) \delta(u_2) + \frac{p}{2\pi} h(u_1) \quad (-\pi < u_2 < \pi),$$

где $h(u_1)$ — условная плотность вероятности амплитуды сигнала U_1 при его наличии. Весовая функция $g(t, \tau, u)$ и функция $\beta(u)$ в случае интервала наблюдения T , кратного полупериоду π/ω , определяются в данном случае формулами (15.1.41) и (15.1.43), полученными в примере 15.1.1. Формула (15.4.6) в этом случае дает

$$Q(Z) = \int_0^\infty \frac{h(u_1) du_1}{2\pi} \int_{-\pi}^{\pi} \exp \left\{ \frac{u_1}{k} \int_{t-T}^t Z(\tau) \sin(\omega\tau + u_2) d\tau - \frac{u_1^2 T}{4k} \right\} du_2,$$

или

$$Q(Z) = \frac{1}{2\pi} \int_0^\infty h(u_1) du_1 \int_{-\pi}^{\pi} \exp \left\{ \frac{u_1}{k} (H_1 \cos u_2 + H_2 \sin u_2) - \frac{u_1^2 T}{4k} \right\} du_2, \quad (15.4.19)$$

где для краткости положено

$$H_1 = \int_{-T}^t Z(\tau) \sin \omega \tau d\tau, \quad H_2 = \int_{-T}^t Z(\tau) \cos \omega \tau d\tau. \quad (15.4.20)$$

Пологая в (15.4.19)

$$H_1 = \sqrt{H_1^2 + H_2^2} \cos \Phi, \quad H_2 = \sqrt{H_1^2 + H_2^2} \sin \Phi,$$

приведем формулу (15.4.19) к виду

$$Q(Z) = \frac{1}{2\pi} \int_0^\infty h(u_1) du_1 \int_{-\pi}^\pi \exp \left\{ \frac{u_1}{k} \sqrt{H_1^2 + H_2^2} \cos(u_2 - \Phi) \right\} du_2.$$

Далее, принимая во внимание, что подынтегральная функция является периодической с периодом 2π , можем при замене переменных $\theta = u_2 - \Phi$ оставить пределы интегрирования неизменными. В результате получим

$$Q(Z) = \frac{1}{2\pi} \int_0^\infty h(u_1) du_1 \int_{-\pi}^\pi \exp \left\{ \frac{u_1}{k} \sqrt{H_1^2 + H_2^2} \cos \theta \right\} d\theta. \quad (15.4.21)$$

Отсюда видно, что $Q(Z)$ является возрастающей функцией $H_1^2 + H_2^2$ и, наоборот, величина $H_1^2 + H_2^2$ является возрастающей функцией величины $Q(Z)$. А так как за выходной сигнал оптимальной системы можно принять любую возрастающую функцию величины $Q(Z)$, то в данном случае можно принять

$$W^* = H_1^2 + H_2^2. \quad (15.4.22)$$

Таким образом, оптимальная обработка сигнала в случае обнаружения синусоидального сигнала с неизвестной фазой состоит в умножении принимаемого сигнала на $\sin \omega t$ и на $\cos \omega t$, интегрировании полученных выражений и суммировании квадратов полученных интегралов.

В случае критерия идеального наблюдателя $\lambda = 1$ и вторая формула (15.4.9) дает для определения порогового уровня c уравнение

$$\frac{q}{p} = \frac{1}{2\pi} \int_0^\infty h(u_1) du_1 \int_{-\pi}^\pi e^{\frac{u_1}{k} \sqrt{c} \cos \theta} d\theta. \quad (15.4.23)$$

В случае критерия Неймана — Пирсона находим вероятность ложной тревоги. Для этого, так же как и в предыдущем примере, убеждаемся в том, что при отсутствии сигнала, когда $Z(\tau) = X(\tau)$, величины H_1 и H_2 являются независимыми нормально распределенными случайными величинами с математическими ожиданиями, равными нулю, и одинаковыми дисперсиями $kT/2$. Таким образом, если H_1 и H_2 рассматривать как декартовы координаты случайной точки на плоскости, то рассеивание на этой плоскости будет круговым нормальным с центром рассеивания в начале координат. Вероятность ложной тревоги определится как вероятность неравенства $H_1^2 + H_2^2 > c$, т. е. как вероятность непадения случайной точки (H_1, H_2) в круг радиуса $\sqrt{c}$:

$$P_{\text{лт}} = P(W^* > c | W = 0) = \frac{1}{\pi kT} \iint_{H_1^2 + H_2^2 > c} e^{-\frac{H_1^2 + H_2^2}{kT}} dH_1 dH_2.$$

Переходя к полярным координатам, получим

$$P_{\text{лт}} = \frac{1}{\pi k T} \int_0^{2\pi} d\varphi \int_{\sqrt{c}}^{\infty} e^{-\frac{r^2}{kT}} r dr = e^{-\frac{c}{kT}}. \quad (15.4.24)$$

Приравнивая вероятность ложной тревоги заданной величине α и решая полученное уравнение относительно c , находим пороговый уровень:

$$c = -kT \ln \alpha = kT \ln \frac{1}{\alpha}. \quad (15.4.25)$$

Предоставляем читателю самостоятельно вывести формулу для вероятности пропуска сигнала в этом случае.

Пример 15.4.3. Совершенно так же решаются задачи обнаружения импульсных сигналов. Если в условиях примера 15.4.1 импульсы помехи распределены нормально, независимы и имеют дисперсии k (т. е. помеха является импульсным белым шумом), то система линейных алгебраических уравнений (15.1.38) принимает вид

$$kg_{hl}(u) = u_l s_l \quad (l = 1, \dots, h), \quad (15.4.26)$$

где для краткости положено

$$s_l = \sin \omega t_l \quad (l = 1, \dots, h). \quad (15.4.27)$$

Решая полученные уравнения и подставляя результат в формулу (15.1.37), находим

$$\int_{t_h - T}^{t_h} g(t_h, \tau, u) Z(\tau) d\tau = \frac{u_l}{k} \sum_{l=1}^h s_l Z_l, \quad Z_l = Z(t_l) \quad (l = 1, \dots, h).$$

Формула (15.1.27) в данном случае дает

$$\beta(u) = \sum_{l=1}^h g_{hl}(u) u_l \sin \omega t_l = \frac{u_l^2}{k} \sum_{l=1}^h s_l^2 = \frac{u_l^2}{k} S. \quad (15.4.28)$$

После этого формула (15.4.6) дает

$$Q(Z) = \int_0^{\infty} h(u_l) \exp \left\{ \frac{u_l}{k} \sum_{l=1}^h s_l Z_l - \frac{u_l^2}{2k} S \right\} du_l. \quad (15.4.29)$$

Отсюда видно, совершенно так же, как в примере 15.4.1, что за выходной сигнал W^* оптимальной системы можно принять

$$W^* = \sum_{l=1}^h s_l Z_l. \quad (15.4.30)$$

Пороговый уровень c в случае критерия «идеального наблюдателя» определяется как значение W^* , соответствующее $Q(Z) = q/p$. Это дает для c уравнение

$$\frac{q}{p} = \int_0^{\infty} h(u_l) \exp \left\{ \frac{u_l}{k} c - \frac{u_l^2}{2k} S \right\} du_l. \quad (15.4.31)$$

Для определения ϵ в случае критерия Неймана — Пирсона находим вероятность ложной тревоги. При отсутствии полезного сигнала принимаемые импульсы Z_l совпадают с импульсами помехи X_l . Следовательно, условный закон распределения величины W^* при отсутствии сигнала является нормальным, причем математическое ожидание величины W^* равно нулю, а дисперсия определяется формулой

$$D[W^*] = k \sum_{l=1}^h s_l^2 = kS,$$

так как импульсы помехи X_l по условию независимы и их дисперсии равны k . Поэтому вероятность ложной тревоги определяется в данном случае формулой

$$P_{\text{лт}} = \frac{1}{\sqrt{2\pi kS}} \int_c^{\infty} e^{-\frac{v^2}{2kS}} dv = \frac{1}{2\pi} \int_c^{\infty} e^{-\frac{z^2}{2}} dz = \frac{1}{2} - \Phi\left(\frac{c}{\sqrt{kS}}\right). \quad (15.4.32)$$

Приравнивая эту вероятность заданной величине α , получим уравнение для определения порогового уровня:

$$\Phi\left(\frac{c}{\sqrt{kS}}\right) = \frac{1}{2} - \alpha. \quad (15.4.33)$$

Для определения вероятности пропуска сигнала заметим, что в случае наличия сигнала амплитуды u_1

$$W^* = u_1 \sum_{l=1}^h s_l^2 + \sum_{l=1}^h s_l X_l = u_1 S + \sum_{l=1}^h s_l X_l.$$

Следовательно, в случае наличия полезного сигнала амплитуды u_1 выходной сигнал оптимальной системы обнаружения W^* является нормально распределенной случайной величиной с математическим ожиданием $u_1 S$ и дисперсией kS . Поэтому вероятность пропуска сигнала определяется формулой

$$\begin{aligned} P_{\text{пр}} &= \frac{1}{\sqrt{2\pi kS}} \int_0^{\infty} h(u_1) du_1 \int_{-\infty}^c e^{-\frac{(v-u_1 S)^2}{2kS}} dv = \\ &= \int_0^{\infty} h(u_1) \left[\frac{1}{2} + \Phi\left(\frac{c-u_1 S}{\sqrt{kS}}\right) \right] du_1. \end{aligned} \quad (15.4.34)$$

Замечая, что согласно (15.4.33) величина $v = c/\sqrt{kS}$ зависит только от заданной вероятности ложной тревоги α , можем переписать полученную формулу для вероятности пропуска сигнала в виде

$$P_{\text{пр}} = \sqrt{k} \int_0^{\infty} h(\gamma \sqrt{k}) \left[\frac{1}{2} + \Phi(v - \gamma \sqrt{S}) \right] d\gamma. \quad (15.4.35)$$

§ 15.5. Определение оптимальных систем для селекции сигналов

В задачах практики требуемый выходной сигнал часто сопровождается не только искусственной или естественной помехой, представляющей собой случайную функцию времени, отличающуюся по своим свойствам от полезного сигнала, но и помехой

Переходя к полярным координатам, получим

$$P_{\text{лт}} = \frac{1}{\pi k T} \int_0^{2\pi} d\varphi \int_{\sqrt{c}}^{\infty} e^{-\frac{r^2}{kT}} r dr = e^{-\frac{c}{kT}}. \quad (15.4.24)$$

Приравнивая вероятность ложной тревоги заданной величине α и решая полученное уравнение относительно c , находим пороговый уровень:

$$c = -kT \ln \alpha = kT \ln \frac{1}{\alpha}. \quad (15.4.25)$$

Предоставляем читателю самостоятельно вывести формулу для вероятности пропуска сигнала в этом случае.

Пример 15.4.3. Совершенно так же решаются задачи обнаружения импульсных сигналов. Если в условиях примера 15.4.1 импульсы помехи распределены нормально, независимы и имеют дисперсии k (т. е. помеха является импульсным белым шумом), то система линейных алгебраических уравнений (15.1.38) принимает вид

$$kg_{hl}(u) = u_l s_l \quad (l = 1, \dots, h), \quad (15.4.26)$$

где для краткости положено

$$s_l = \sin \omega t_l \quad (l = 1, \dots, h). \quad (15.4.27)$$

Решая полученные уравнения и подставляя результат в формулу (15.1.37), находим

$$\int_{t_h - T}^{t_h} g(t_h, \tau, u) Z(\tau) d\tau = \frac{u_l}{k} \sum_{l=1}^h s_l Z_l, \quad Z_l = Z(t_l) \quad (l = 1, \dots, h).$$

Формула (15.1.27) в данном случае дает

$$\beta(u) = \sum_{l=1}^h g_{hl}(u) u_l \sin \omega t_l = \frac{u_l^2}{k} \sum_{l=1}^h s_l^2 = \frac{u_l^2}{k} S. \quad (15.4.28)$$

После этого формула (15.4.6) дает

$$Q(Z) = \int_0^{\infty} h(u_l) \exp \left\{ \frac{u_l}{k} \sum_{l=1}^h s_l Z_l - \frac{u_l^2}{2k} S \right\} du_l. \quad (15.4.29)$$

Отсюда видно, совершенно так же, как в примере 15.4.1, что за выходной сигнал W^* оптимальной системы можно принять

$$W^* = \sum_{l=1}^h s_l Z_l. \quad (15.4.30)$$

Пороговый уровень c в случае критерия «идеального наблюдателя» определяется как значение W^* , соответствующее $Q(Z) = q/p$. Это дает для c уравнение

$$\frac{q}{p} = \int_0^{\infty} h(u_l) \exp \left\{ \frac{u_l}{k} c - \frac{u_l^2}{2k} S \right\} du_l. \quad (15.4.31)$$

Для определения c в случае критерия Неймана — Пирсона находим вероятность ложной тревоги. При отсутствии полезного сигнала принимаемые импульсы Z_l совпадают с импульсами помехи X_l . Следовательно, условный закон распределения величины W^* при отсутствии сигнала является нормальным, причем математическое ожидание величины W^* равно нулю, а дисперсия определяется формулой

$$D[W^*] = k \sum_{l=1}^h s_l^2 = kS,$$

так как импульсы помехи X_l по условию независимы и их дисперсии равны k . Поэтому вероятность ложной тревоги определяется в данном случае формулой

$$P_{лт} = \frac{1}{\sqrt{2\pi kS}} \int_c^{\infty} e^{-\frac{v^2}{2kS}} dv = \frac{1}{2\pi} \int_{\frac{c}{\sqrt{kS}}}^{\infty} e^{-\frac{z^2}{2}} dz = \frac{1}{2} - \Phi\left(\frac{c}{\sqrt{kS}}\right). \quad (15.4.32)$$

Приравнявая эту вероятность заданной величине α , получим уравнение для определения порогового уровня:

$$\Phi\left(\frac{c}{\sqrt{kS}}\right) = \frac{1}{2} - \alpha. \quad (15.4.33)$$

Для определения вероятности пропуска сигнала заметим, что в случае наличия сигнала амплитуды u_1

$$W^* = u_1 \sum_{l=1}^h s_l^2 + \sum_{l=1}^h s_l X_l = u_1 S + \sum_{l=1}^h s_l X_l.$$

Следовательно, в случае наличия полезного сигнала амплитуды u_1 выходной сигнал оптимальной системы обнаружения W^* является нормально распределенной случайной величиной с математическим ожиданием $u_1 S$ и дисперсией kS . Поэтому вероятность пропуска сигнала определяется формулой

$$\begin{aligned} P_{пр} &= \frac{1}{\sqrt{2\pi kS}} \int_0^{\infty} h(u_1) du_1 \int_{-\infty}^c e^{-\frac{(v-u_1S)^2}{2kS}} dv = \\ &= \int_0^{\infty} h(u_1) \left[\frac{1}{2} + \Phi\left(\frac{c-u_1S}{\sqrt{kS}}\right) \right] du_1. \end{aligned} \quad (15.4.34)$$

Замечая, что согласно (15.4.33) величина $v = c/\sqrt{kS}$ зависит только от заданной вероятности ложной тревоги α , можем переписать полученную формулу для вероятности пропуска сигнала в виде

$$P_{пр} = \sqrt{k} \int_0^{\infty} h(\gamma \sqrt{k}) \left[\frac{1}{2} + \Phi(v - \gamma \sqrt{S}) \right] d\gamma. \quad (15.4.35)$$

§ 15.5. Определение оптимальных систем для селекции сигналов

В задачах практики требуемый выходной сигнал часто сопровождается не только искусственной или естественной помехой, представляющей собой случайную функцию времени, отличающуюся по своим свойствам от полезного сигнала, но и помехой

в виде ложных сигналов. Так, например, противник может создавать помехи работе локационных станций в виде ложных целей, чтобы отраженные от них сигналы смешались с сигналом от действительной цели и нарушили работу локационных станций. В таких случаях возникает задача выбора действительного сигнала из группы сигналов или, как говорят, задача *селекции* сигнала.

Ясно, что для возможности селекции сигнала необходимо, чтобы этот сигнал по каким-либо физическим свойствам отличался от ложных сигналов. Так обычно и бывает на практике вследствие практической невозможности создать ложные цели, не имеющие никаких физических отличий от действительной. Математически это отличие выражается в том, что функциональная форма действительного сигнала и его зависимость от параметров отличаются от функциональной формы ложных сигналов и их зависимости от параметров. Но в таком случае для решения проблемы оптимальной селекции сигнала можно применить общий метод § 15.1.

Предположим, что сигнал $\varphi(t, U)$ в (15.1.11) является суммой нескольких сигналов $\varphi_0(t, U), \varphi_1(t, U), \dots, \varphi_m(t, U)$, представляющих собой известные функции времени t и неизвестных (случайных) параметров $U_1, \dots, U_N$:

$$Z(t) = \varphi_0(t, U) + \sum_{r=1}^m \varphi_r(t, U) + X(t). \quad (15.5.1)$$

Предположим, что $\varphi_0(t, U)$ представляет собой действительный сигнал, а остальные сигналы являются ложными. Тогда задача селекции сигнала будет состоять в наилучшем воспроизведении сигнала $\varphi_0(t, U)$ по результатам наблюдения случайной функции $Z(t)$ (т. е. всех принимаемых сигналов вместе с помехой). Требуемый выходной сигнал $W(t)$ определяется в данном случае формулой

$$W(t) = \psi(t, U) = \varphi_0(t, U). \quad (15.5.2)$$

В более общем случае, когда система должна не только выделить истинный сигнал от ложных и от помехи, но и осуществить некоторые преобразования истинного сигнала, например экстраполяция или дифференцирование, требуемый выходной сигнал определяется формулой

$$W(t) = \psi(t, U) = P_1 \varphi_0(t, U), \quad (15.5.3)$$

где P — некоторый оператор, в общем случае нелинейный.

Подставляя выражения функций $\varphi(t, U)$ и $\psi(t, U)$ из (15.5.1), (15.5.2) или (15.5.3) в формулы § 15.1, получим общий метод решения задач селекции сигнала по любым статистическим критериям. Для практического применения этого общего метода к различным конкретным задачам селекции сигнала необходимо определить форму истинных и ложных сигналов и их зависимость

от неизвестных параметров для используемых в проектируемой системе физических принципов приема сигнала. Тогда, выбрав подходящий критерий, например критерий минимума средней квадратической ошибки, можно общим методом § 15.1 найти оптимальное решение задачи селекции сигнала и оценить потенциальные качества системы селекции, основанной на данных физических принципах.

Вследствие того что изложенный общий метод определения оптимальных способов обработки принимаемых сигналов, как было отмечено в § 15.1, применим к нахождению оптимальных многомерных систем, имеющих любое число входов и выходов, он дает принципиальную возможность находить оптимальное решение задачи селекции сигнала и для многоканальных систем, основанных на одновременной обработке нескольких сигналов, получаемых от различных источников информации, работающих на различных физических принципах.

МЕТОДЫ ОПРЕДЕЛЕНИЯ ОПТИМАЛЬНОГО УПРАВЛЕНИЯ

§ 16.1. Принцип максимума

Среди задач теории оптимальных систем, перечисленных в § 13.1, следует выделить такие, в которых объект управления задан и требуется определить управляющие действия или, короче, *управления*, обеспечивающие в известном смысле наилучшее, оптимальное протекание процесса управления в заданных условиях. При этом информация об объекте управления задается обычно в виде уравнений, описывающих его поведение. Должен быть задан также критерий оптимальности системы, в качестве которого могут быть выбраны различные технические и экономические показатели или вероятностные характеристики ошибки. Задача состоит в определении алгоритма управления (закона управления), доставляющего экстремум критерию качества при условии, что на сами управления и на координаты объекта могут быть наложены ограничения. Для решения этой задачи применяются методы вариационного исчисления, принцип максимума Л. С. Понтрягина [48], являющийся развитием вариационного метода, метод динамического программирования Р. Беллмана [77], также являющийся развитием вариационного метода и связанный с принципом максимума. Из них последние два более адекватны задачам теории оптимального управления, так как дают возможность учесть реальные ограничения и получить решение в тех случаях, когда классические методы вариационного исчисления неприменимы. С этим направлением связаны методы аналитического конструирования регуляторов, изложенные в [92], а также методы, развитые в [89].

Рассмотрим типовые задачи оптимального управления объектом, поведение которого описывается конечным числом обыкновенных дифференциальных уравнений. Примерами таких объектов могут служить летательный аппарат, двигатель и т. п. Состояние такого объекта характеризуется n выходными переменными $y_1, \dots, y_n$. Удобно эти величины считать координатами вектора y . Объект управления при функционировании подвергается действию управляющих сигналов $u_1, \dots, u_r$ через органы управления. Если таких управлений несколько, то их также можно объединить в вектор управления u .

Предположим, что поведение объекта управления описывается дифференциальными уравнениями и начальными условиями вида

$$\dot{y}_i = f_i(y_1, \dots, y_n, u_1, \dots, u_r, t), \quad y_i(t_0) = y_{i0} \quad (i = 1, \dots, n), \quad (16.1.1)$$

где f_i — в общем случае нелинейные функции, y_{i0} — начальные значения переменных y_i . Считая производные $\dot{y}_i$ и функции f_i координатами соответствующих векторов $\dot{\mathbf{y}}$, $\mathbf{f}$, можно переписать уравнения (16.1.1) в более компактном векторном виде:

$$\dot{\mathbf{y}} = \mathbf{f}(\mathbf{y}, \mathbf{u}, t), \quad \mathbf{y}(t_0) = \mathbf{y}_0. \quad (16.1.2)$$

Практически управления u_i не могут быть любыми вследствие физических свойств объекта или ограниченной мощности органов управления. Они должны удовлетворять определенным ограничениям, которые часто имеют вид

$$|u_i| \leq R_i \quad (i = 1, \dots, n). \quad (16.1.3)$$

В общем случае эти ограничения характеризуют область R допустимых управлений, т. е. область R , которой должен принадлежать вектор управления $\mathbf{u}$:

$$\mathbf{u} \in R. \quad (16.1.4)$$

Кроме того, ограничения могут быть наложены и на координаты объекта. Учет ограничений на управления чрезвычайно важен при проектировании управляющих устройств, так как во многих случаях приходится искать оптимальное управление при ограниченных ресурсах.

Цель управления можно рассматривать как достижение экстремума критерия качества I за счет выбора оптимального вектора допустимого управления $\mathbf{u}$. Применяемые при определении оптимального управления критерии качества можно свести к следующим трем основным:

1) Критерий минимума времени перехода объекта из заданного начального состояния $\mathbf{y}_0$ в заданное конечное состояние $\mathbf{y}_K$ (критерий максимального быстродействия):

$$I(t_K) = \int_{t_0}^{t_K} dt = t_K - t_0, \quad (16.1.5)$$

где t_K — момент времени окончания управления.

2) Критерий минимума функции от конечного состояния объекта при заданном времени управления $t_k - t_0 = T^*$):

$$I(t_k) = F(y(t_k)), \quad (16.1.6)$$

где F — некоторая дифференцируемая функция.

3) Критерий минимума интеграла от функции состояния объекта и управления при заданном времени управления $t_k - t_0$:

$$I(t_k) = \int_{t_0}^{t_k} F(y, u, t) dt. \quad (16.1.7)$$

В тех случаях, когда некоторые параметры системы и действующие на нее возмущения случайны, применяют статистический критерий оптимальности — математическое ожидание некоторого функционала качества.

Рассмотренные три типовые задачи сводятся к одной обобщенной задаче. Для этого достаточно рассматривать функционал качества не только в конечный момент времени t_k , но и в любой момент времени t и принять его текущее значение за дополнительную координату $y_{n+1}(t)$. Тогда задача сведется к обеспечению минимума величины $y_{n+1}(t_k)$. Так, приняв время t за дополнительную координату y_{n+1} , определяемую уравнением

$$\dot{y}_{n+1} = 1, \quad y_{n+1,0} = y_{n+1}(t_0) = t_0, \quad (16.1.8)$$

сведем задачу минимизации времени управления (16.1.5) к минимизации $y_{n+1}(t_k)$. Приняв текущее значение функционала (16.1.6) за новую переменную $y_{n+1} = F(y)$, получим в силу (16.1.1) дополнительное уравнение

$$\dot{y}_{n+1} = \sum_{h=1}^n \frac{\partial F}{\partial y_h} f_h(y_1, \dots, y_n, u_1, \dots, u_r, t), \quad (16.1.9)$$

и задача сведется к нахождению минимума $y_{n+1}(t_k)$. Наконец, приняв текущее значение функционала (16.1.7) за новую переменную y_{n+1} ,

$$\dot{y}_{n+1}(t) = \int_{t_0}^t F(y, u, t) dt,$$

получим для этой переменной уравнение

$$\dot{y}_{n+1} = F(y_1, \dots, y_n, u_1, \dots, u_r, t), \quad y_{n+1}(t_0) = 0, \quad (16.1.10)$$

и задача снова сведется к нахождению минимума $y_{n+1}(t_k)$.

*) В дальнейшем везде будем рассматривать минимум соответствующего критерия, так как переменной знака максимум приводится к минимуму.

Следовательно, задача отыскания минимума критерия качества I любого из рассмотренных трех типов сводится к отысканию минимума координаты y_{n+1} в момент времени t_k по отношению к управлению u . При этом на конечное положение объекта управления в некоторых случаях могут накладываться ограничения. В первой из рассмотренных задач такое ограничение состоит в задании конечных значений координат объекта управления. В остальных двух задачах ограничения на конечное состояние могут иметь вид

$$q_p(y_1(t_k), \dots, y_n(t_k)) = 0 \quad (p = 1, \dots, m), \quad (16.1.11)$$

где q_p — некоторые дифференцируемые функции. Тогда, согласно известному методу нахождения условного экстремума, задача сведется к отысканию экстремума функционала вида

$$\pi = y_{n+1}(t_k) + \sum_{j=1}^m \lambda_j q_j(y_1(t_k), \dots, y_n(t_k)), \quad (16.1.12)$$

где λ_j — неопределенные множители Лагранжа, определяемые из условий (16.1.11).

В общей постановке задача оптимального управления объектом, поведение которого описывается уравнениями (16.1.1) и одним из уравнений (16.1.8), (16.1.9), (16.1.10), состоит в определении управления u , удовлетворяющего ограничениям (16.1.4) и доставляющего минимум функционалу π .

Простое и изящное решение задачи отыскания оптимального управления и минимизации функционала π впервые было получено Л. С. Понтрягиным и его учениками [79] благодаря открытому ими принципу максимума, связавшему оптимизируемый функционал состояния π с динамикой процесса. Эта связь была установлена через функцию Гамильтона вида

$$H(y, u, \psi, t) = \sum_{i=1}^{n+1} \psi_i f_i, \quad (16.1.13)$$

где функции f_i — правые части уравнений (16.1.1) и одного из уравнений (16.1.8), (16.1.9), (16.1.10), а функции ψ_i удовлетворяют уравнениям

$$\dot{\psi}_i = - \sum_{j=1}^{n+1} \psi_j \frac{\partial f_j}{\partial y_i} \quad (i = 1, \dots, n+1) \quad (16.1.14)$$

при конечных условиях

$$\left. \begin{aligned} \psi_i(t_k) &= - \frac{\partial \pi}{\partial y_i} = - \sum_{j=1}^m \lambda_j \frac{\partial q_j}{\partial y_i} \quad (i = 1, \dots, n), \\ \psi_{n+1}(t_k) &= -1. \end{aligned} \right\} \quad (16.1.15)$$

Принцип максимума состоит в том, что для оптимального управления u и соответствующих координат y_i , для которых функционал J имеет минимальное значение, функция Гамильтона H имеет максимум или в общем случае свое наибольшее значение (supremum) по отношению к управлениям на всем интервале (t_0, t_n) . Принцип максимума был обоснован как необходимое условие оптимальности для нелинейных объектов и необходимое и достаточное условие оптимальности для линейных объектов [48, 79, 93].

Из сформулированного утверждения вытекает, что необходимым условием для того, чтобы управление u обеспечивало минимум функционалу J , является выполнение условия максимума функции Гамильтона по отношению к управлению в течение всего времени управления. Следовательно, определение оптимального управления u состоит в нахождении максимума функции $H(y, u, \psi, t)$ относительно управления u для любого момента времени в интервале $t_0 \leq t \leq t_n$.

Используя выражение (16.1.13) для функции H , можно переписать уравнения объекта управления (16.1.1) для координат y_i и уравнения (16.1.14) для вспомогательных функций ψ_i в другой форме. Для этого продифференцируем выражение (16.1.13) по ψ_i :

$$\frac{\partial H}{\partial \psi_i} = f_i(y, u, t) \quad (i = 1, \dots, n+1). \quad (16.1.16)$$

Дифференцируя (16.1.13) по y_i , получим

$$\frac{\partial H}{\partial y_i} = \sum_{j=1}^{n+1} \psi_j \frac{\partial f_j}{\partial y_i} \quad (i = 1, \dots, n+1). \quad (16.1.17)$$

При помощи (16.1.16) и (16.1.17) уравнения (16.1.1) и (16.1.14) приводятся к следующей канонической форме уравнений Гамильтона:

$$\dot{y}_i = \frac{\partial H}{\partial \psi_i} \quad (i = 1, \dots, n+1), \quad (16.1.18)$$

$$\dot{\psi}_i = -\frac{\partial H}{\partial y_i} \quad (i = 1, \dots, n+1). \quad (16.1.19)$$

Процедура определения оптимального управления заключается в нахождении из условия максимума функции H вектора управления $u(\psi)$ как функции вектора ψ , подстановке $u(\psi)$ в уравнения (16.1.1) и (16.1.14) и в решении этих уравнений относительно $y(t)$, $\psi(t)$. Как следует из изложенного, при отыскании оптимального управления необходимо решать систему дифференциальных уравнений (16.1.1) при начальных условиях и (16.1.14) при конечных условиях (так называемую «двухточечную» задачу). Это обстоятельство существенно осложняет достаточно ясный алгоритм решения и приводит к тому, что практические сложные задачи

должны решаться численно при подборе подходящих начальных условий в уравнениях (16.1.14) методом проб.

Докажем необходимость максимума функции H для оптимальности управления, ограничиваясь случаем, когда конечное состояние объекта не подчинено никаким условиям, а время управления $t_1 - t_0 = T$ задано. В этом случае конечные условия (16.1.15) имеют вид

$$\psi_i(t_k) = 0 \quad (i = 1, \dots, n), \quad \psi_{n+1}(t_k) = -1. \quad (16.1.20)$$

Л. С. Понтрягин предложил эффективный путь решения задачи с помощью понятия *игольчатой вариации* управления. Это понятие является центральным в доказательстве принципа максимума. При этом допустимыми управлениями $u_j(t)$ будем считать все кусочно-непрерывные функции времени t , удовлетворяющие ограничениям (16.1.4).

Пусть $u^*(t)$ — допустимое оптимальное управление, а $y^*(t)$ — соответствующая фазовая траектория. Оптимальное управление $u^*(t)$ является вектором, координаты которого $u_j^*(t)$ — кусочно-непрерывные функции на отрезке (t_0, t_k) . Следуя

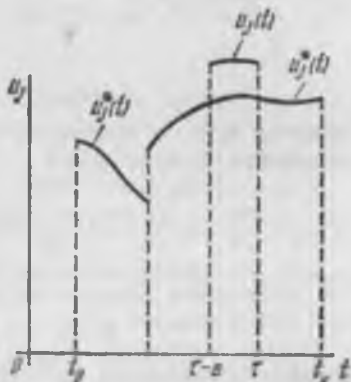


Рис. 16.1.1.

Л. С. Понтрягину [79], рассмотрим бесконечно малый промежуток времени $\tau - \varepsilon < t < \tau$ и проварьируем управление $u^*(t)$, заменив его на этом малом интервале величиной $u(t)$, оставив его неизменным вне этого интервала. Такое варьирование называется игольчатым (рис. 16.1.1). Отметим, что мы не требуем, чтобы вариации $\delta u_j(t) = u_j(t) - u_j^*(t)$ были бесконечно малыми величинами. Важно только, чтобы они удовлетворяли ограничению (16.1.4). При этом, несмотря на конечную величину разности $u_j(t) - u_j^*(t)$, импульс приращения управления $\delta u_{j\varepsilon}$ и влияние этой вариации на последующее поведение объекта бесконечно малы. Для доказательства вычислим вызываемое вариацией управления изменение поведения объекта. В соответствии с уравнениями (16.1.1) и (16.1.8) (или (16.1.9), или (16.1.10)) вариация $\delta y_i(t) = y_i(t) - y_i^*(t)$ в момент $t = \tau$ равна с точностью до малых высшего порядка разности скоростей изменения координат, умноженной на промежуток времени ε :

$$\begin{aligned} \delta y_i(\tau) &= \left[\left(\frac{dy_i}{dt} \right)_{t=\tau} - \left(\frac{dy_i^*}{dt} \right)_{t=\tau} \right] \varepsilon = \\ &= [f_i(y(\tau), u(\tau), \tau) - f_i(y^*(\tau), u^*(\tau), \tau)] \varepsilon \quad (i = 1, \dots, n+1). \end{aligned} \quad (16.1.21)$$

Эти разности бесконечно малы и являются начальными значениями вариаций переменных на интервале (τ, t_k) . В силу этого переменные $y_i(t)$ будут бесконечно мало отличаться от оптимальных $y_i^*(t)$ в интервале времени (τ, t_k) . Положив в уравнениях (16.1.1) $y_i(t) = y_i^*(t) + \delta y_i(t)$, $u(t) = u^*(t)$ при $t > \tau$, получим

$$\frac{dy_i^*}{dt} + \frac{d\delta y_i}{dt} = f_i(y_i^* + \delta y_i, \dots, y_n^* + \delta y_n, u_1^*, \dots, u_r^*, t). \quad (16.1.22)$$

Раскладывая правые части в ряд Тейлора в окрестности $y_i^*(t), \dots, y_n^*(t)$ и отбрасывая члены второго порядка малости, получим

$$\begin{aligned} \frac{dy_i^*}{dt} + \frac{d\delta y_i}{dt} = f_i(y_i^*, \dots, y_n^*, u_1^*, \dots, u_r^*, t) + \\ + \sum_{p=1}^{n+1} \frac{\partial f_i(y_i^*, \dots, y_n^*, u_1^*, \dots, u_r^*, t)}{\partial y_p^*} \delta y_p. \end{aligned}$$

Вычитая отсюда соответствующие уравнения (16.1.1), получим уравнения в вариациях

$$\frac{d\delta y_i}{dt} = \sum_{p=1}^{n+1} \frac{\partial f_i(y^*, u^*, t)}{\partial y_p^*} \delta y_p \quad (i=1, \dots, n+1). \quad (16.1.23)$$

Эти уравнения следует интегрировать при начальных условиях (16.1.21) в момент τ .

Теперь вычислим вариацию функционала J . Так как оптимальное управление $u^*(t)$ должно обеспечивать минимальное значение J , то при замене $u^*(t)$ любым другим управлением $u(t)$ значение J может лишь увеличиваться. Следовательно, необходимым условием минимума J будет $\delta J = \delta y_{n+1}(t_k) \geq 0$. Вариацию $\delta y_{n+1}(t_k)$ можно вычислить, не интегрируя уравнения (16.1.23). Действительно, вычислим производную по времени от суммы

$$\sum_{i=1}^{n+1} \psi_i \delta y_i$$

$$\frac{d}{dt} \sum_{i=1}^{n+1} \psi_i \delta y_i = \sum_{i=1}^{n+1} \dot{\psi}_i \delta y_i + \sum_{i=1}^{n+1} \psi_i \dot{\delta y}_i.$$

Подставляя сюда выражения $\dot{\psi}_i$, $\dot{\delta y}_i$ из (16.1.14) и (16.1.23), получим

$$\frac{d}{dt} \sum_{i=1}^{n+1} \psi_i \delta y_i = - \sum_{i,j=1}^{n+1} \psi_j \frac{\partial f_j}{\partial y_i} \delta y_i + \sum_{i,j=1}^{n+1} \psi_j \frac{\partial f_j}{\partial y_i} \delta y_i = 0.$$

Отсюда следует, что на интервале $\tau \leq t \leq t_k$

$$\sum_{i=1}^{n+1} \psi_i(t) \delta y_i(t) = \text{const.}$$

Отсюда, учитывая (16.1.20), находим

$$\sum_{i=1}^{n+1} \psi_i(\tau) \delta y_i(\tau) = \sum_{i=1}^{n+1} \psi_i(t_k) \delta y_i(t_k) = -\delta y_{n+1}(t_k).$$

Таким образом,

$$\delta \pi = -\delta y_{n+1}(t_k) = -\sum_{i=1}^{n+1} \psi_i(\tau) \delta y_i(\tau).$$

Подставив сюда выражение $\delta y_i(\tau)$ из (16.1.21), получим

$$\delta \pi = -\varepsilon \sum_{i=1}^{n+1} [\psi_i(\tau) f_i(y(\tau), u(\tau), \tau) - \psi_i(\tau) f_i(y^*(\tau), u^*(\tau), \tau)]. \quad (16.1.24)$$

Так как в качестве момента τ можно выбрать любой текущий момент времени t в интервале (t_0, t_k) , то равенство (16.1.24) справедливо для любого момента времени t . Следовательно, учитывая (16.1.13), равенство (16.1.24) можно записать в виде

$$\delta \pi = -\varepsilon [H(y, u, \psi, t) - H(y^*, u^*, \psi, t)]. \quad (16.1.25)$$

Отсюда следует, что $\delta \pi \geq 0$ при допустимой вариации δu только при

$$H(y^*, u^*, \psi, t) \geq H(y, u, \psi, t).$$

Таким образом, доказано, что для минимума функционала π необходимо, чтобы функция Гамильтона H имела максимальное значение относительно управления u на всем интервале управления.

В частном случае, когда ограничений на управления нет, оптимальное управление является стационарной точкой функции Гамильтона при любом t , т. е. доставляет ей максимум. В этом случае оптимальные управления определяются из уравнений

$$\frac{\partial H}{\partial u_i} = 0 \quad (i = 1, \dots, r). \quad (16.1.26)$$

Доказательство принципа максимума для общей задачи оптимального управления, а также его достаточности для линейного объекта более сложно. Его можно найти в [48, 93].

§ 16.2. Применение принципа максимума для решения типовых задач

Пусть движение объекта управления описывается уравнениями

$$\dot{y}_i = \varphi_i(y_1, \dots, y_n, t) + u_i \quad (i = 1, \dots, n). \quad (16.2.1)$$

Требуется определить оптимальные управления u_i , обеспечивающие перевод объекта из заданного начального состояния $y_{i0} = y_i(t_0)$ в заданное конечное состояние $y_i(t_k)$ за минимальное

время $t_k - t_0$ с учетом ограничений $|u_i| \leq R_i$. Добавим к (16.2.1) уравнение (16.1.8) и, учитывая, что конечное состояние объекта задано, получим для функционала π формулу

$$\pi = y_{n+1}(t_k) + \sum_{j=1}^n \lambda_j y_j(t_k). \quad (16.2.2)$$

Функция Гамильтона имеет в данном случае вид

$$H = \sum_{i=1}^n \psi_i [\varphi_i(y, t) + u_i] + \psi_{n+1}. \quad (16.2.3)$$

Это выражение достигает максимума при

$$u_i = R_i \operatorname{sgn} \psi_i \quad (i = 1, \dots, n). \quad (16.2.4)$$

Таким образом, оптимальное управление является релейным, причем переключения определяются знаками функций ψ_i .

Для определения вспомогательных переменных ψ_i составим уравнения (16.1.14) и конечные условия (16.1.15):

$$\left. \begin{aligned} \dot{\psi}_i &= - \sum_{j=1}^n \psi_j \frac{\partial \varphi_j}{\partial y_i}, & \psi_i(t_k) &= -\lambda_i \\ & & (i &= 1, \dots, n), \\ \dot{\psi}_{n+1} &= 0, & \psi_{n+1}(t_k) &= -1. \end{aligned} \right\} \quad (16.2.5)$$

Решая уравнения (16.2.1), (16.2.5), определим окончательно закон управления. Однако в общем случае это удастся сделать только численными методами, применяя метод проб, так как заданы начальные условия для координат y_i и конечные условия для вспомогательных функций ψ_i .

В частном случае линейного стационарного объекта уравнения (16.2.1) имеют вид

$$\dot{y}_i = \sum_{j=1}^n a_{ij} y_j + u_i, \quad y_i(t_0) = y_{i0} \quad (i = 1, \dots, n). \quad (16.2.6)$$

Уравнения (16.2.5) для функций ψ_i при этом принимают форму

$$\dot{\psi}_i = - \sum_{j=1}^n a_{ij} \psi_j, \quad \psi_j(t_k) = -\lambda_j \quad (i = 1, \dots, n). \quad (16.2.7)$$

Решение уравнений (16.2.7) имеет вид

$$\psi_i(t) = \sum_{j=1}^n c_{ij} e^{\nu_j t},$$

где ν_j — корни характеристического уравнения. Эта сумма переходит через нуль не более чем $n - 1$ раз и имеет, следовательно,

не более n интервалов постоянства знака. Поэтому каждое из управлений u_i имеет в общем случае $n - 1$ переключение, т. е. n интервалов постоянства знака, на которых $u_i = \pm R_i$.

Пример 16.2.1. Рассмотрим процесс управления объектом, поведение которого описывается уравнениями

$$\dot{y}_1 = y_2, \quad \dot{y}_2 = -2ay_2 - b^2y_1 + u_1$$

при начальных и конечных условиях $y_1(t_0) = y_{10}$, $y_2(t_0) = y_{20}$, $y_1(t_K) = 0$, $y_2(t_K) = 0$. Требуется определить управление $u_1(t)$, ограниченное по модулю, $|u_1| \leq R$, и обеспечивающее минимальное время перевода объекта из начального в конечное состояние.

Добавим к написанным уравнениям уравнение для y_3 :

$$\dot{y}_3 = 1, \quad y_3(t_0) = t_0.$$

Минимизируемый функционал для этого случая имеет вид

$$\pi = y_3(t_K) + \lambda_1 y_1(t_K) + \lambda_2 y_2(t_K),$$

а функция Гамильтона выражается формулой

$$H = \psi_1 y_2 + \psi_2 (-2ay_2 - b^2y_1 + u_1) + \psi_3.$$

Отсюда следует, что функция Гамильтона имеет максимум при

$$u_1 = R \operatorname{sgn} \psi_2.$$

Вспомогательная переменная ψ_2 определяется из уравнений $\dot{\psi}_1 = b^2\psi_2$, $\dot{\psi}_2 = 2a\psi_2 - \psi_1$, $\dot{\psi}_3 = 0$ при конечных условиях $\psi_1(t_K) = -\lambda_1$, $\psi_2(t_K) = -\lambda_2$, $\psi_3(t_K) = -1$. Исключая ψ_1 , получим

$$\ddot{\psi}_2 - 2a\dot{\psi}_2 + b^2\psi_2 = 0.$$

Решение этого уравнения при $a^2 - b^2 < 0$ имеет вид

$$\psi_2(t) = A_1 e^{at} \sin(\sqrt{b^2 - a^2} t + A_2),$$

где A_1 и A_2 определяются из конечных условий. Оптимальное управление определяется формулой

$$u_1 = R \operatorname{sgn} [A_1 e^{at} \sin(\sqrt{b^2 - a^2} t + A_2)].$$

Для определения постоянных A_1 и A_2 необходимо подставить найденное управление в уравнения для y_1 и y_2 и проинтегрировать их при заданных начальных условиях. Тогда A_1 , A_2 , λ_1 , λ_2 и момент окончания управления определяется из заданных конечных условий для y_1 , y_2 , ψ_1 и ψ_2 . Эту часть задачи можно решать только численно.

Рассмотрим задачу управления конечным состоянием объекта, описываемого уравнениями (16.2.1), при заданном начальном состоянии $y(t_0) = y_0$. Требуется определить оптимальные управления u_i , ограниченные по модулю, $|u_i| \leq R_i$, обеспечивающие минимум заданной функции конечного состояния:

$$I(t_K) = F(y_1(t_K), \dots, y_n(t_K)). \quad (16.2.8)$$

Добавим к (16.2.1) уравнение (16.1.9), которое в данном случае имеет вид

$$\dot{y}_{n+1} = \sum_{j=1}^n \frac{\partial F}{\partial y_j} [\varphi_j(y, t) + u_j], \quad y_{n+1}(t_0) = F(y(t_0)). \quad (16.2.9)$$

Так как никаких ограничений на конечное состояние объекта не наложено, то $\pi = y_{n+1}(t_K)$. Учитывая (16.2.1) и (16.2.9), находим функцию Гамильтона (16.1.13);

$$H = \sum_{i=1}^n \left(\psi_i + \psi_{n+1} \frac{\partial F}{\partial y_i} \right) \varphi_i(y, t) + \sum_{i=1}^n \left(\psi_i + \psi_{n+1} \frac{\partial F}{\partial y_i} \right) u_i.$$

Отсюда видно, что H будет иметь наибольшее возможное значение в любой момент t , если

$$u_i = R_i \operatorname{sgn} \left(\psi_i + \psi_{n+1} \frac{\partial F}{\partial y_i} \right) \quad (i = 1, \dots, n). \quad (16.2.10)$$

Таким образом, управления и в этом случае имеют релейный характер с функциями переключения $\psi_i + \psi_{n+1} \frac{\partial F}{\partial y_i}$. Для определения этих функций переключения необходимо проинтегрировать уравнения (16.1.14), которые в данном случае имеют вид

$$\left. \begin{aligned} \dot{\psi}_i &= - \sum_{j=1}^n \left[\left(\psi_j + \psi_{n+1} \frac{\partial F}{\partial y_j} \right) \frac{\partial \varphi_j}{\partial y_i} + \psi_{n+1} \frac{\partial^2 F}{\partial y_j \partial y_i} \dot{y}_j \right], \\ \dot{\psi}_{n+1} &= 0 \end{aligned} \right\} \quad (16.2.11)$$

с конечными условиями $\psi_i(t_K) = 0$ ($i = 1, \dots, n$), $\psi_{n+1}(t_K) = -1$. Из этих уравнений следует, что $\psi_{n+1}(t) = -1$. Присоединяя к уравнениям (16.2.11) систему (16.2.1) и формулу (16.2.10), получим полную систему уравнений, определяющих ψ_i и y_i . Численное их решение дает возможность найти функции ψ_i и y_i и окончательно определить оптимальные управления u_i . Отметим, что только в случае линейного объекта (функции φ_i линейно зависят от y_i) уравнения (16.2.11) интегрируются отдельно.

Пример 16.2.2. Объект управления описывается линейными уравнениями

$$\dot{y}_1 = a_{11}(t)y_1 + a_{12}(t)y_2 + u_1, \quad \dot{y}_2 = a_{21}(t)y_1 + a_{22}(t)y_2$$

при начальных условиях $y_1(0) = y_{10}$, $y_2(0) = y_{20}$. Требуется определить оптимальное управление, ограниченное по модулю, $|u_i| \leq R$, минимизирующее модуль координаты y_1 в момент времени t_K : $I(t_K) = F(y_1(t_K)) = |y_1(t_K)|$. Дополнительное уравнение для y_3 в данном случае имеет вид

$$\dot{y}_3 = |y_1|, \quad y_3(0) = |y_{10}|.$$

В соответствии с (16.2.10) оптимальное управление определяется формулой

$$u_1 = R \operatorname{sgn}(\psi_1 - \operatorname{sgn} y_1).$$

Для определения ψ_1 служат уравнения

$$\dot{\psi}_1 = -a_{11}(t)\psi_1 - a_{21}(t)\psi_2 + a_{11}\operatorname{sgn} y_1, \quad \psi_1(t_k) = 0,$$

$$\dot{\psi}_2 = -a_{12}(t)\psi_1 - a_{22}(t)\psi_2 + a_{12}(t)\operatorname{sgn} y_1, \quad \psi_2(t_k) = 0.$$

Численное решение этих уравнений совместно с уравнениями для y_1, y_2 при учете формулы для u_1 дает возможность определить оптимальное управление $u_1(t)$.

Рассмотрим третью типовую задачу определения оптимального управления для объекта, поведение которого описывается теми же уравнениями (16.2.1), при заданном начальном состоянии $y_0 = y(t_0)$. Требуется определить оптимальные управления u_i , ограниченные по модулю, $|u_i| \leq R_i$, и переводящие объект из начального состояния $y_i(t_0)$ в заданное конечное состояние $y_i(t_k)$ за заданное время $t_k - t_0$ при минимизации интеграла (16.1.7).

Добавляя к уравнениям (16.2.1) уравнение (16.1.10) и учитывая, что конечное состояние объекта задано, запишем минимизируемый функционал в виде

$$\pi = y_{n+1}(t_k) + \sum_{j=1}^n \lambda_j y_j(t_k). \quad (16.2.12)$$

Используя уравнения (16.2.1) и (16.1.10), находим функцию Гамильтона (16.1.13):

$$H = \sum_{i=1}^n \psi_i [\varphi_i(y, t) + u_i] + \psi_{n+1} F(y, u, t). \quad (16.2.13)$$

Условие максимума этого выражения с учетом ограничений $|u_i| \leq R_i$ определяет оптимальное управление u как функцию величин ψ_i/ψ_{n+1} , y и t . Для определения вспомогательных переменных ψ_i , ψ_{n+1} составим уравнения (16.1.14) при конечных условиях (16.1.15):

$$\left. \begin{aligned} \dot{\psi}_i &= - \sum_{j=1}^n \psi_j \frac{\partial \varphi_j}{\partial y_i} - \psi_{n+1} \frac{\partial F}{\partial y_i}, & \psi_i(t_k) &= 0 \\ (i &= 1, \dots, n), \\ \psi_{n+1} &= 0, & \psi_{n+1}(t_k) &= -1. \end{aligned} \right\} \quad (16.2.14)$$

Из последнего уравнения следует, что $\psi_{n+1}(t) \equiv -1$. Объединяя уравнения (16.2.1), (16.2.14) и формулу для u , полученную из условия максимума H , и учитывая начальные и конечные условия, получаем полную систему уравнений, численное решение которой дает возможность определить все переменные и вычислить оптимальные управления $u_i(t)$.

Пример 16.2.3. Рассмотрим систему, поведение которой описывается уравнениями

$$\dot{y}_1 = y_2, \quad \dot{y}_2 = -ay_1 + u_1 \quad (a > 0).$$

Требуется найти оптимальное управление u_1 , ограниченное по модулю, $|u_1| \leq R$, переводящее объект из начального состояния y_{10}, y_{20} в состояние y_{1K}, y_{2K} за заданное время t_K и минимизирующее интеграл

$$I(t_K) = \int_0^{t_K} |u_1(t)| dt.$$

Дополнительное уравнение (16.1.10) в данном случае имеет вид

$$\dot{y}_3 = |u_1|, \quad y_3(0) = 0.$$

С учетом заданного конечного положения объекта минимизируемый функционал принимает вид (16.2.12). Функция Гамильтона определяется формулой

$$H = \psi_1 y_2 + \psi_2 (-ay_1 + u_1) + \psi_3 |u_1|.$$

Уравнения (16.2.14), определяющие функции ψ_1 и ψ_2 , имеют вид

$$\dot{\psi}_1 = 0, \quad \dot{\psi}_2 = -\psi_2 + a\psi_3, \quad \psi_1(t_K) = -\lambda_1, \quad \psi_2(t_K) = -\lambda_2.$$

а $\psi_3(t) = -1$. Оптимальное управление, доставляющее максимум функции H , имеет вид $u_1 = 0$ при $|\psi_2(t)| < 1$, $u_1 = R \operatorname{sgn} \psi_2(t)$ при $|\psi_2(t)| \geq 1$. Оптимальный процесс и оптимальное управление $u_1(t)$ определяются путем численного решения уравнений для y_1, y_2, ψ_1, ψ_2 при учете полученной формулы для u_1 , а также начальных и конечных условий.

Пример 16.2.4. Объект управления описывается уравнением

$$\dot{y} = -0,6y + u, \quad y(t_0) = 0, \quad y(t_K) = 10.$$

Требуется определить оптимальное управление u , ограниченное по модулю, $|u| \leq 25$, при котором объект за заданное время $t_K - t_0 = 1$ с переводится из начального состояния $y_0 = 0$ в конечное $y_K = 10$ и при этом функционал

$$I(t_K) = \int_{t_0}^{t_K} [y^2(t) + 1,25^2 u^2(t)] dt$$

принимает минимальное значение.

Минимизируемый функционал (16.2.12) имеет вид

$$\pi = y_2(t_K) + \lambda_1 y_1(t_K),$$

где $y_2 = y_1^2 + 1,25^2 u^2$, $y_1 = y$. Функция Гамильтона в данном случае имеет вид

$$H = \psi(-0,6y_1 + u) - y_1^2 - 1,25^2 u^2.$$

Оптимальное управление, минимизирующее функцию H , имеет вид

$$u = \frac{1}{2k^2} \psi \quad \text{при} \quad \left| \frac{1}{2k^2} \psi \right| < R, \quad u = R \operatorname{sgn} \psi \quad \text{при} \quad \left| \frac{1}{2k^2} \psi \right| > R.$$

где $k = 1, 2, 5$, $R = 25$. Для определения функции ψ составим уравнение (16.2.14):

$$\dot{\psi} = 0,6\psi + 2u_1, \quad \psi(t_k) = 0.$$

Далее решение задачи можно получить путем совместного численного интегрирования уравнений для y_1 и ψ при учете найденного вида управления u .

§ 16.3. Динамическое программирование

Метод динамического программирования был разработан практически одновременно с принципом максимума и представляет собой другой способ решения тех же вариационных задач с ограничениями на управления и на фазовые координаты объекта. Этот метод достаточно полно обоснован для оптимизации дискретных систем, в области которых лежит основная сфера его применения. При этом вариационная задача рассматривается как многоэтапный процесс решения более простых задач, а оптимальное управление отыскивается последовательно шаг за шагом.

При применении динамического программирования к вариационным задачам теории управления снимается трудность решения граничной двухточечной задачи, к которой приводит применение принципа максимума. Однако возникают свои вычислительные трудности, связанные с возрастанием размерности решаемых задач.

Метод динамического программирования обладает большой общностью и применим для решения широкого круга задач, в которых связи между координатами, управлениями и сами критерии могут быть заданы в виде уравнений произвольного вида, графиков или таблиц.

При некоторых допущениях метод динамического программирования применим непосредственно и к непрерывным системам, описываемым дифференциальными уравнениями. В этом случае вариационная задача также сводится к некоторой граничной задаче дифференциальных уравнений в частных производных, решение которой, однако, связано с вычислительными трудностями.

В основе метода динамического программирования лежит принцип оптимальности, доказанный Р. Беллманом, как общее необходимое условие оптимального процесса, справедливое как для дискретных, так и для непрерывных систем. Этот принцип формулируется следующим образом: оптимальное управление обладает тем свойством, что, каково бы ни было начальное состояние и управление до данного момента, последующее управление должно быть оптимальным при том состоянии, в которое пришла система в результате первого этапа управления [77, 78]. Из принципа оптимальности следует, что оптимальное управление определяется лишь состоянием системы в каждый данный момент

времени и целью и не зависит от состояний в предыдущие моменты времени.

Для изложения существа метода и получения рекуррентных расчетных соотношений рассмотрим простейшую одномерную дискретную систему, поведение которой описывается разностным уравнением на заданном интервале $t_k - t_0 = T$:

$$y(k+1) = y(k) + f(y(k), u(k), k), \quad (16.3.1)$$

где $y(k)$ — координата системы, $u(k)$ — управление в момент времени kT_n ($k = 0, 1, \dots, N$), $T_n = (t_k - t_0)/N$ — интервал дискретности. Пусть начальная координата объекта равна $y(t_0) = y_0$, а управление ограничено, $|u| \leq R$. Требуется определить последовательность оптимальных управлений $u(0), \dots, u(N-1)$, минимизирующую сумму

$$I = \sum_{k=0}^{N-1} F(y(k), u(k)) + Q(y(N)), \quad (16.3.2)$$

где F и Q — некоторые известные функции. Задача минимизации функции (16.3.2) многих переменных $u(0), u(1), \dots, u(N-1)$ сложна. Однако, воспользовавшись принципом оптимальности, можно свести общую операцию минимизации функции многих переменных к последовательной минимизации функций одной переменной. Действительно, рассмотрим последний интервал $(N-1)T_n \leq t \leq NT_n$, предполагая, что состояние $y(N-1)$ известно. На основании принципа оптимальности управление $u(N-1)$ на этом интервале зависит от $y(N-1)$, должно удовлетворять ограничению и минимизировать на этом интервале сумму, входящую в (16.3.2):

$$I_{N-1} = F(y(N-1), u(N-1)) + Q(y(N)). \quad (16.3.3)$$

Заменив $y(N)$ его выражением из уравнения (16.3.1), получим

$$I_{N-1} = F(y(N-1), u(N-1)) + Q(y(N-1) + f(y(N-1), u(N-1), N-1)). \quad (16.3.4)$$

Сумма (16.3.4) зависит от одной неизвестной $u(N-1)$, которую следует выбрать из условия минимума I_{N-1} при учете ограничения на u . Обозначим это оптимальное значение $u^*(N-1)$, а минимальное значение I_{N-1} , зависящее от $y(N-1)$, через $S_{N-1}(y(N-1))$:

$$\begin{aligned} S_{N-1}(y(N-1)) &= \min_{|u(N-1)| \leq R} I_{N-1} = \\ &= \min_{|u(N-1)| \leq R} \{F(y(N-1), u(N-1)) + Q(y(N-1) + \\ &\quad + f(y(N-1), u(N-1), N-1))\}. \end{aligned} \quad (16.3.5)$$

Таким образом, минимизируя (16.3.4) по переменной $u(N-1)$, определяем функцию $S_{N-1}(y(N-1), (N-1))$.

Перейдем теперь к интервалу времени $(N-2)T_{\Pi} \leq t \leq NT_{\Pi}$, состоящему из последнего и предпоследнего участков. Этому интервалу соответствует сумма, входящая в (16.3.2),

$$I_{N-2} = F(y(N-2), u(N-2)) + I_{N-1}. \quad (16.3.6)$$

Так же, как и ранее, состояние $y(N-2)$ считаем известным. Тогда на основании принципа оптимальности состояние $y(N-2)$ и цель управления — минимизация I_{N-2} — определяют оптимальное управление на интервале $((N-2)T_{\Pi}, NT_{\Pi})$. Определим минимум I_{N-2} по $u(N-2)$ и $u(N-1)$, обозначив его через S_{N-2} :

$$S_{N-2}(y(N-2)) = \min_{\substack{|u(N-2)| \leq R \\ |u(N-1)| \leq R}} I_{N-2}. \quad (16.3.7)$$

Учитывая, что минимум I_{N-1} по $u(N-1)$ уже определен и равен $S_{N-1}(y(N-1), (N-1))$, а $y(N-1)$ зависит от $y(N-2)$ и $u(N-2)$ согласно уравнению (16.3.1), выражение (16.3.7) запишем в виде

$$S_{N-2}(y(N-2)) = \min_{|u(N-2)| \leq R} \{F(y(N-2), u(N-2)) + \\ + S_{N-1}(y(N-2) + f(y(N-2), u(N-2), (N-2)))\}. \quad (16.3.8)$$

Минимизация здесь производится также по одной переменной $u(N-2)$ при заданной координате $y(N-2)$. В результате определяем оптимальное значение $u^*(N-2)$ и $S_{N-2}(y(N-2))$ как функции $y(N-2)$.

Продолжая описанную процедуру оптимизации по шагам от конечного момента времени N к начальному, получим рекуррентную формулу

$$S_{N-k}(y(N-k)) = \min_{|u(N-k)| \leq R} \{F(y(N-k), u(N-k)) + \\ + S_{N-k+1}(y(N-k) + f(y(N-k), u(N-k), (N-k)))\}, \quad (16.3.9)$$

в которой минимизация производится по одной переменной $u(N-k)$. В результате оптимальное управление $u^*(N-k)$ в момент времени $t = (N-k)T_{\Pi}$ определяется как функция $y(N-k)$.

Применяя формулу (16.3.9) последовательно для $k = 1, \dots, N$, определяем управления $u^*(N-k)$ и, наконец, приходим к значению $k = N$, при котором определяем $u^*(N-N) = u^*(0)$ — управление в начальный момент времени, зависящее от начального известного состояния объекта $y(0) = y_0$. В результате полностью определяется оптимальное управление на заданном промежутке времени.

Весь изложенный алгоритм справедлив и для многомерных систем с любым числом управлений $u_1, \dots, u_n$. Для применения его необходимо заменить скаляры y, u, f в приведенных формулах соответственно векторами y, u, f . В результате потребуется на каждом этапе минимизировать функцию r переменных $u_1 (N - k), \dots, u_r (N - k)$. При этом на каждом этапе вычислений необходимо определять и запоминать две функции $S_{N-k} (y (N - k))$ и $S_{N-k+1} (y (N - k + 1))$ n переменных. Это требует большого объема памяти при вычислениях на машине. В этом состоят вычислительные трудности метода динамического программирования.

Пример 16.3.1. Решим задачу, рассмотренную в примере 16.2.4, применив метод динамического программирования. Для этого заменим дифференциальное уравнение разностным, а интегральный критерий — суммой. В результате приходим к задаче отыскания оптимального управления в дискретной системе, описываемой уравнением

$$y_{k+1} = y_k + (u_k - 0,6y_k) \Delta t \quad (k = 0, 1, \dots, N).$$

Допустим, что объект управления имеет 6 дискретных состояний: 0, 2, 4, 6, 8, 10. Требуется определить оптимальные управления u_k , ограниченные по модулю, $|u_k| \leq 25$, при которых объект переходит из начального состояния $y_0 = 0$ в состояние $y_N = 10$ за время $T = 1$, так, чтобы минимизировать величину

$$I = \sum_{k=0}^{N-1} (y_k^2 + 1,25^2 u_k^2) \Delta t.$$

Для решения задачи методом динамического программирования примем постоянный интервал дискретности $\Delta t = 0,2$ и, значит, $N = 5$. Определим сначала траектории, ведущие в конечную точку $y_5 = 10$ из любого положения y_k , в котором система может быть к моменту t_k , и вычислим для них значения критерия по формуле

$$I_k = (y_k^2 + 1,25^2 u_k^2) \Delta t.$$

Вычисленные значения I_k приведены в таблице 16.3.1. Там же показаны значения соответствующих управлений u_k .

ТАБЛИЦА 16.3.1

y_k	0	2	4	6	8	10
I_k	782	529	332	177	81	31
u_k	50	41,2	32,4	23,6	14,8	6

На рис. 16.3.1 изображены точки, характеризующие возможные состояния объекта. На этом же рисунке прямыми линиями изображены траектории перехода объекта из одного состояния в другое. Над отрезками прямых даны значения критерия I_k для соответствующих переходов. Из результатов вычис-

лений и таблицы 16.3.1 следует, что возможны только три траектории перехода, для которых $|u_k| < 25$. Эти траектории отмечены сплошными отрезками прямых, а остальные — пунктиром. Следовательно, в момент времени $t_k = 0,8$ объект может находиться в трех положениях: 6, 8 или 10.

На следующем этапе вычисления

$$I_3 = (y_3^2 + 1,25u_3^2) \Delta t + I_4$$

и определяем траектории перехода в состояние $y_5 = 10$ из различных состояний в момент t_3 . Надо найти для каждого исходного положения этого этапа единственный оптимальный путь, имеющий наименьшее значение I_3 и допустимое u_3 . На рис. 16.3.1 сплошными линиями показаны эти оптимальные пути с обозначением соответствующих значений $\Delta I_3 = I_3 - I_4$.

На следующем и вообще на каждом текущем участке тоже для каждого исходного положения выбирается единственная оптимальная траектория с наименьшим значением критерия, удовлетворяющая ограничению $|u| < 25$. В результате такого целенаправленного перебора возможных решений, на каждом следующем расчетном этапе для каждого исходного состояния исключаются неоптимальные варианты и выбирается только один — оптимальный вариант. Продолжая таким образом вычисления и построения оптимальных траекторий из каждой исходной точки, доходим до последнего этапа, на котором рассматривается одна исходная точка $y_0 = 0$. Для этой точки строятся траектории в точки 0, 2, 4, 6, 8, 10, и для них вычисляются управления и критерий

$$I_0 = (y_0^2 + 1,25u_0^2) \Delta t + I_1.$$

После этого определяется, на какой траектории суммарный критерий минимален. Минимальное значение критерия в данном случае $S_0 = 268$ достигается для траектории, показанной на рис. 16.3.1 жирной линией. В таблице 16.3.2 приведены отрезки оптимальной траектории, соответствующие оптимальные управления u_k^* и значения критерия ΔI_k на этих отрезках.

ТАБЛИЦА 16.3.2

y_k	0	2	4	6	8	10
t	0	0,2	0,4	0,6	0,8	1
u_k^*	10	11,2	12,4	13,6	14,8	
ΔI_k	31	40	51	65	81	$\Sigma = 268$

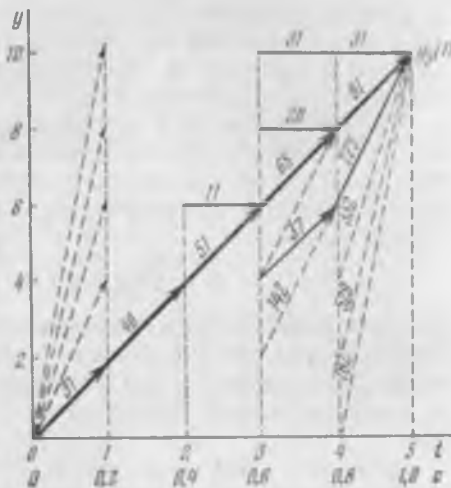


Рис. 16.3.1.

§ 16.4. Применение динамического программирования к непрерывным системам

Метод динамического программирования может быть при некоторых допущениях распространен также на непрерывные системы и приводит к своеобразному дифференциальному уравнению в частных производных.

Рассмотрим задачу оптимизации управления при интегральном критерии и заданном начальном состоянии объекта. Конечное состояние объекта может быть фиксированным или произвольным. Пусть поведение объекта описывается системой уравнений в векторной форме

$$\dot{y} = f(y, u, t), \quad y(t_0) = y_0, \quad (16.4.1)$$

где y и u — векторы, размерности n , r соответственно, $r \leq n$. Предположим, что управление объектом производится на заданном интервале времени (t_0, t_K) и требуется выбрать вектор управления u , принадлежащий области R , $u \in R$, минимизирующий функционал

$$I(t_K) = \int_{t_0}^{t_K} F(y(t), u(t), t) dt + Q(y(t_K)), \quad (16.4.2)$$

где F — заданная интегрируемая функция, Q — заданная функция конечного состояния объекта. Обозначим через $u^*(t)$ вектор оптимального управления, а через $y(t)$ траекторию объекта при оптимальном управлении. Для этой оптимальной траектории введем функцию Беллмана

$$S(y(t), t) = \min_{u \in R} \left[\int_t^{t_K} F(y(\tau), u(\tau), \tau) d\tau + Q(y(t_K)) \right]. \quad (16.4.3)$$

Она в общем случае зависит от координат оптимизируемого интервала и самого момента времени t .

Выведем уравнение, которому удовлетворяет функция S , применив принцип оптимальности. Для этого возьмем бесконечно малый интервал $(t, t + \Delta t)$ и предположим, что оптимальное управление для интервала $(t + \Delta t, t_K)$ найдено. Перепишем формулу (16.4.3) в виде

$$S(y(t), t) = \min_{u \in R} \left[\int_t^{t+\Delta t} F(y(\tau), u(\tau), \tau) d\tau + \right. \\ \left. + \int_{t+\Delta t}^{t_K} F(y(\tau), u(\tau), \tau) d\tau + Q(y(t_K)) \right]$$

или, с точностью до бесконечно малых высшего порядка,

$$S(y(t), t) = \min_{u \in R} [F(y(t), u(t), t) \Delta t + S(y(t + \Delta t), t + \Delta t)]. \quad (16.4.4)$$

Предположим, что функция S непрерывна и имеет всюду непрерывные частные производные по всем своим аргументам y_i, t . Это основное допущение при применении метода динамического программирования к непрерывным системам, которое теоретически пока полностью не обосновано [80]. Разложим второй член в квадратных скобках в (16.4.4) в ряд Тейлора в окрестности t и ограничимся членами не выше первого порядка относительно Δt . Тогда, учитывая (16.4.1), получим

$$S(y(t + \Delta t), t + \Delta t) = S(y, t) + \sum_{i=1}^n \frac{\partial S(y, t)}{\partial y_i} f_i(y, u, t) \Delta t + \frac{\partial S(y, t)}{\partial t} \Delta t.$$

Подставим это выражение в (16.4.4) и учтем, что $S(y, t)$ и $\partial S(y, t)/\partial t$ не зависят от u , вследствие чего их можно вынести за знак минимума по u . В результате получим

$$S(y, t) = S(y, t) + \frac{\partial S(y, t)}{\partial t} \Delta t + \min_{u \in R} \left\{ F(y(t), u(t), t) \Delta t + \sum \frac{\partial S(y, t)}{\partial y_i} f_i(y, u, t) \Delta t \right\}.$$

Разделив это уравнение на Δt после сокращения $S(y, t)$ и устремив Δt к нулю, получим

$$-\frac{\partial S(y, t)}{\partial t} = \min_{u \in R} \left\{ F(y, u, t) + \sum_{i=1}^n \frac{\partial S(y, t)}{\partial y_i} f_i(y, u, t) \right\}. \quad (16.4.5)$$

Это уравнение называется уравнением Р. Беллмана. Его можно переписать в векторной форме

$$-\frac{\partial S(y, t)}{\partial t} = \min_{u \in R} \{ F(y, u, t) + (\text{grad } S(y, t), f(y, u, t)) \}, \quad (16.4.6)$$

где второй член в правой части представляет собой скалярное произведение вектора $\text{grad } S(y, t)$ с компонентами $\partial S(y, t)/\partial y_i$ и вектора $f(y, u, t)$ с компонентами $f_i(y, u, t)$. Граничное условие для уравнения Беллмана имеет вид

$$S(y, t_k) = Q(y(t_k)). \quad (16.4.7)$$

В частном случае, когда в минимизируемом функционале (16.4.2) отсутствует второй член, $Q(y(t_k)) = 0$, условие (16.4.7) принимает вид $S(y, t_k) = 0$.

Оптимальный вектор управления u определяется путем минимизации правой части в (16.4.6). При этом оптимальное управление u будет зависеть от $\text{grad } S(y, t)$. Подставляя найденное выражение для u в (16.4.6), получаем уравнение в частных производных, в которое не входят u . Решив это уравнение и определив $S(y, u, t)$, найдем явный вид оптимального управления u .

Решение уравнения Беллмана само по себе связано с серьезными трудностями. Решить его аналитически можно только в очень немногих частных случаях. Поэтому, как правило, его приходится решать численными методами.

Пример 16.4.1. Пусть объект описывается уравнениями

$$\dot{y}_1 = u y_1 + y_2, \quad \dot{y}_2 = u^2.$$

Требуется выбрать управление u , минимизирующее функционал

$$I(t_n) = \int_{t_0}^{t_n} F(y_1(t), y_2(t), t) dt$$

на заданном интервале (t_0, t_n) при заданных начальных значениях $y_1(t_0) = y_{10}$, $y_2(t_0) = y_{20}$.

Уравнение Беллмана (16.4.6) для данного случая имеет вид

$$-\frac{\partial S(y, t)}{\partial t} = \min_u \left[F(y_1, y_2, t) + \frac{\partial S(y, t)}{\partial y_1} (u y_1 + y_2) + \frac{\partial S(y, t)}{\partial y_2} u^2 \right]. \quad (16.4.8)$$

Выражение в квадратных скобках принимает минимальное значение при u , равном

$$u = -\frac{y_1}{2} \frac{\partial S(y, t)}{\partial y_1} / \frac{\partial S(y, t)}{\partial y_2}.$$

Подставляя это выражение в правую часть уравнения Беллмана (16.4.8) и опуская символ $\min$, получим уравнение в частных производных для определения функции $S(y, t)$:

$$-\frac{\partial S}{\partial t} = F(y_1, y_2, t) + \frac{\partial S}{\partial y_1} y_2 - \frac{1}{4} y_1^2 \left(\frac{\partial S}{\partial y_1} \right)^2 / \frac{\partial S}{\partial y_2}.$$

Это уравнение надо решить при граничном условии $S(y_1, y_2, t_n) = 0$.

Пример 16.4.2. Определим оптимальное управление для условия примера 16.4.1 при ограничении на управление $|u| \leq 1$. В этом случае оптимальное управление u , минимизирующее выражение в квадратных скобках в (16.4.8), имеет вид

$$u = v(t) \quad \text{при } |v(t)| < 1, \\ u = \text{sgn } v(t) \quad \text{при } |v(t)| > 1,$$

где

$$v(t) = -\frac{y_1}{2} \frac{\partial S}{\partial y_1} / \frac{\partial S}{\partial y_2}.$$

Подставляя полученное выражение для управления в правую часть уравнения (16.4.8), получаем уравнение в частных производных для функции S с тем же граничным условием. Решение этого уравнения, как и в предыдущем примере, может быть получено только численно.

Пример 16.4.3. Пусть поведение объекта управления описывается линейными уравнениями

$$\dot{y}_i = \sum_{j=1}^n a_{ij} y_j + m_i u, \quad y_i(t_0) = y_{i0} \quad (i = 1, \dots, n),$$

где a_{ij}, m_i — заданные постоянные *). Требуется найти управление u , которое за бесконечное время переводит объект из начального состояния $y_{i0} \neq 0$ при $t_0 = 0$ в нулевое положение $y_{ik} = 0$, минимизируя функционал

$$I = \int_0^{\infty} \sum_{i,j=1}^n \left(c_{ij} y_i y_j + \frac{1}{k} u^2 \right) d\tau,$$

где c_{ij}, k — заданные постоянные. На управление дополнительных ограничений не наложено. Принимая во внимание, что система стационарна, составим уравнение Беллмана, которое в данном случае имеет вид

$$\min_u \left[\sum_{i,j=1}^n c_{ij} y_i y_j + \frac{1}{k} u^2 + \sum_{i=1}^n \frac{\partial S(y)}{\partial y_i} \left(\sum_{j=1}^n a_{ij} y_j + m_i u \right) \right] = 0. \quad (16.4.9)$$

Для определения u , доставляющего минимум выражению в квадратных скобках, служит уравнение

$$\frac{2}{k} u + \sum_{i=1}^n m_i \frac{\partial S(y)}{\partial y_i} = 0.$$

Отсюда получаем оптимальное управление в форме

$$u = -\frac{1}{2} k \sum_{i=1}^n m_i \frac{\partial S(y)}{\partial y_i}.$$

Подставляя это выражение в (16.4.9), получаем уравнение для определения $S(y)$:

$$\sum_{i,j=1}^n c_{ij} y_i y_j + \frac{k}{4} \left(\sum_{i=1}^n m_i \frac{\partial S}{\partial y_i} \right)^2 + \sum_{i=1}^n \frac{\partial S}{\partial y_i} \left(\sum_{j=1}^n a_{ij} y_j - \frac{k}{2} m_i \sum_{j=1}^n m_j \frac{\partial S}{\partial y_j} \right) = 0. \quad (16.4.10)$$

Решение этого нелинейного уравнения можно искать в виде квадратичной формы

$$S = \sum_{r,p=1}^n b_{rp} y_r y_p.$$

Подставляя это выражение в (16.4.10) и приравнявая нулю коэффициенты при одинаковых произведениях переменных $y_r y_p$, получаем для неизвестных

*) Этот пример заимствован из книги А. М. Летова [92].

коэффициентов b_{rp} следующие нелинейные уравнения:

$$-k \sum_{i,j=1}^n m_i m_j b_{ip} b_{jq} + \sum_{i=1}^n a_{iq} b_{ip} = -c_{pq} \quad (p, q = 1, \dots, n). \quad (16.4.11)$$

Решая эти уравнения, определяем искомые коэффициенты b_{pq} ($p, q = 1, \dots, n$). После их определения окончательно находим

$$u = - \sum_{p=1}^n d_p y_p, \quad d_p = k \sum_{i=1}^n m_i b_{ip} \quad (p = 1, \dots, n).$$

Таким образом, в данном случае управление может быть определено аналитически. Однако уравнения (16.4.11) приходится решать численными методами.

САМОНАСТРАИВАЮЩИЕСЯ СИСТЕМЫ

§ 17.1. Управление при неполной информации

Проектируя автоматическую систему, конструктор всегда старается сделать ее возможно проще, обойтись по возможности простыми корректирующими цепями с постоянными параметрами. При этом расчет системы управления ведется в предположении, что динамические характеристики объекта управления известны. Однако в действительности динамические характеристики объектов управления бывают известны лишь приближенно. Кроме того, характеристики объектов управления обычно изменяются с изменением внешних условий. Так, например, динамические характеристики самолета или крылатой ракеты зависят от скорости полета и плотности воздуха (т. е. от высоты полета) вследствие зависимости от этих факторов аэродинамических сил и моментов. Кроме того, динамические характеристики самолета или крылатой ракеты могут измениться вследствие повреждений, полученных в результате действий противника. Благодаря запасам устойчивости система управления, спроектированная для определенных расчетных характеристик объекта управления, будет успешно осуществлять управление объектом и в том случае, когда его характеристики несколько отличаются от расчетных. Методы, изложенные в предыдущих главах, дают возможность определять допустимые пределы изменений характеристик объекта управления, не приводящих к нарушению работы системы управления.

В некоторых случаях диапазон изменения характеристик объекта управления вследствие изменения внешних условий бывает настолько большим, что управление объектом при помощи простейшей системы управления, имеющей постоянные параметры, становится невозможным. Например, для современных самолетов и крылатых ракет часто оказывается невозможным спроектировать систему управления с постоянными параметрами, которая могла бы осуществлять управление полетом на всех высотах и при всех скоростях полета. В таких случаях можно применить систему управления с изменяемыми параметрами (чаще всего оказывается достаточным изменять коэффициенты усиления корректирующих

цепей) и ввести в нее дополнительные цепи настройки изменяемых параметров в зависимости от внешних условий.

Вместо цепей настройки параметров системы управления можно также применить методы стабилизации динамических характеристик объекта управления с помощью дополнительных обратных связей. Это позволяет уменьшить диапазон изменения характеристик объекта при изменении внешних условий и, следовательно, расширить диапазон изменения внешних условий, при которых возможно применение системы управления с постоянными параметрами.

Однако и системы управления с изменяемыми в зависимости от внешних условий параметрами не могут обеспечить управление объектом при значительном изменении характеристик объекта вследствие не предусмотренных в конструкции системы управления причин. Так, например, динамические характеристики самолета или крылатой ракеты могут значительно измениться вследствие обледенения или значительных повреждений в результате действий противника, не приводящих, однако, к невозможности продолжения управляемого полета. Кроме того, в промышленности часто встречаются объекты управления, динамические характеристики которых совершенно неизвестны и не могут быть достаточно просто определены. Примерами таких объектов могут служить доменная печь и различные агрегаты химической промышленности. В таких случаях обычно приходится применять систему управления со сложными цепями настройки параметров, осуществляющим автоматический поиск подходящих значений параметров на основе анализа результатов управления и поведения объекта управления.

Такие системы управления принято называть *самонастраивающимися*.

Оператор самонастраивающейся системы, обеспечивающий удовлетворительное качество ее работы, устанавливается не сразу после начала ее работы, а лишь после некоторого периода времени, необходимого на поиск подходящих значений ее параметров. В течение времени поиска такая система может работать совершенно неудовлетворительно. Это обстоятельство вообще характерно для задач управления при отсутствии достаточно полной информации.

На практике часто встречаются случаи, когда не только характеристики объектов управления, но и статистические характеристики полезных сигналов и помех оказываются неизвестными. В таких случаях развитие в предыдущих главах методы расчета автоматических систем оказываются бессильными. Мы не можем тогда спроектировать не только оптимальную систему, но и вообще систему, которая могла бы обеспечить удовлетворительное решение задачи. Для решения задач управ-

ления в таких случаях надо предварительно получить необходимую информацию о неизвестных характеристиках. Организовав соответствующие эксперименты и обработав их результаты, мы можем после этого спроектировать систему управления изложенными в предыдущих главах методами. Так и поступали раньше. Однако сейчас техника автоматического управления достигла такого уровня, что обработка результатов экспериментов может быть возложена на саму проектируемую автоматическую систему. Для этой цели можно использовать вычислительные устройства, которые неизбежно приходится вводить в состав сложных систем управления. При таком подходе сначала проектируется система, а затем производятся эксперименты, результаты которых вводятся в систему и должны быть обработаны самой системой. Такие системы получают существенно новые качества по сравнению с обычными системами управления, спроектированными для управления известными объектами при известных характеристиках полезных сигналов и помех. А именно такие системы обладают способностью совершенствоваться на основе получаемой ими информации — «учиться». Будучи вначале совершенно неприспособленными (или плохо приспособленными) для решения поставленной задачи, они постепенно «учатся» решать эту задачу, используя для этого информацию, получаемую в начальном периоде их работы (во время проведения экспериментов).

Процесс совершенствования автоматической системы на основе обработки вводимой в нее информации естественно называть *обучением* автоматической системы, а сами автоматические системы, обладающие способностью совершенствоваться на основе накопления опыта и изучения получаемой информации, — *обучающимися* или *обучаемыми* системами.

Если обучающаяся система получает в процессе обучения лишь ту информацию, которая содержится в ее нормальных входных сигналах, с которыми она должна работать и после окончания обучения, то такая система называется *самообучающейся*. Если же в процессе обучения в систему вводится дополнительная внешняя информация, кроме содержащейся в ее естественных входных сигналах, то мы говорим, что такая система *обучается учителем*. При этом *учителем* мы называем любой источник дополнительной внешней информации независимо от его физической природы. С этой точки зрения учителем автоматической системы может быть как человек, так и другая автоматическая система.

В некоторых случаях обучение автоматической системы может быть осуществлено в течение одного цикла или даже в течение сравнительно небольшого начального периода одного цикла ее работы. В других случаях обучение системы может быть осуществлено только за несколько, иногда за большое число циклов ее работы. Самонастраивающаяся система, согласно данному выше опреде-

лению, представляет собой такую самообучающуюся автоматическую систему, у которой процесс самообучения протекает и накапливается в течение сравнительно небольшого начального периода одного цикла ее работы. Таким образом, самонастраивающиеся автоматические системы представляют собой частный вид самообучающихся систем.

В этой главе мы ограничимся изложением некоторых основных принципов построения самонастраивающихся систем. Более общие типы обучающихся систем рассматривать не будем. Читатели, желающие ознакомиться с началами теории обучающихся систем, могут обратиться к работам [57, 61].

Для организации процессов обучения в автоматических системах можно применить изложенные в предыдущих трех главах методы статистической теории оптимальных систем. В частности, эти методы могут быть с успехом использованы для обработки информации в самонастраивающихся системах с целью улучшить и ускорить процесс самонастройки.

Итак, для проектирования обычных систем управления с постоянными или изменяемыми в зависимости от внешних условий параметрами необходима достаточно полная предварительная (априорная) информация об объекте управления и действующих на него возмущениях. А именно необходимо знать уравнения объекта и пределы возможного изменения его параметров, а также их зависимость от внешних условий. Самонастраивающиеся системы и более общие типы обучающихся систем не требуют полной или даже вообще никакой априорной информации об объекте управления и действующих на него возмущениях. Однако самонастраивающиеся системы, как правило, значительно более сложны, чем обычные системы с постоянными параметрами или системы с настройкой параметров в зависимости от внешних условий. Кроме того, анализ результатов управления и поведения объекта управления и поиск наилучших в определенном смысле значений параметров системы управления требуют времени, причем это время тем больше, чем меньше объем априорной информации об объекте управления.

Из изложенного можно сделать два вывода. Во-первых, применять самонастраивающиеся системы целесообразно только в тех случаях, когда нельзя обойтись обычными простейшими системами с постоянными или изменяемыми в зависимости от внешних условий параметрами. Во-вторых, при проектировании самонастраивающихся систем всегда необходимо использовать всю имеющуюся априорную информацию об объекте управления и действующих на него возмущениях, так как только максимальное использование всей имеющейся информации дает возможность свести к минимуму время поиска подходящих значений параметров системы (время обучения системы).

§ 17.2. Принципы построения самонастраивающихся систем

Рассмотрим основные принципы организации процессов самонастройки в автоматических системах.

В основу устройств, обеспечивающих автоматический анализ результатов управления и поведения объекта управления, должен быть положен какой-нибудь подходящий критерий качества системы. За такой критерий качества в зависимости от назначения системы можно принимать различные величины, например: точность управления, какой-нибудь показатель эффективности управления, выходную мощность, быстродействие, какую-нибудь динамическую характеристику замкнутой системы и т. д. Мы будем обозначать любой критерий качества системы, положенный в основу устройства цепи самонастройки, через I .

На основе анализа измеренного или вычисленного критерия качества I система настройки должна определить отклонение поведения системы от желаемого и в соответствии с этим отклонением выработать сигналы управления для исполнительных устройств цепи настройки, осуществляющих изменение параметров системы. По аналогии с параметрами управления мы будем называть величины, служащие мерой отклонения поведения системы от желаемого, *параметрами настройки*. Параметр настройки для каждого настраиваемого параметра системы должен быть скалярной величиной, определяющей интенсивность и направление необходимого изменения этого параметра. Параметры настройки мы будем обозначать буквой ∇ , придавая им в случае необходимости соответствующие индексы.

В зависимости от того, как формируются параметры настройки по критерию качества управления I , получаются различные принципы устройства самонастраивающихся систем. В настоящее время самонастраивающиеся системы еще не получили широкого распространения. Их развитие еще только начинается. Поэтому мы ограничимся здесь рассмотрением двух наиболее характерных типов самонастраивающихся систем: систем с автоматической настройкой параметров по эталону и систем с экстремальной настройкой параметров.

В системах с настройкой параметров по эталону за параметр настройки для всех изменяемых параметров принимается отклонение критерия качества I от его значения I_0 , соответствующего эталонной системе:

$$\nabla = I - I_0. \quad (17.2.1)$$

В системах с экстремальной настройкой параметры системы настраиваются так, чтобы критерий качества I имел экстремальное значение. В этом случае за параметры настройки естественно принять частные производные критерия качества I по соответ-

ствующим параметрам системы $k_1, \dots, k_n$:

$$\nabla I = \frac{\partial I}{\partial k_i} \quad (i = 1, \dots, n). \quad (17.2.2)$$

Таким образом, параметры настройки в системах с экстремальной настройкой обычно представляют собой составляющие вектора градиента функции $I(k_1, \dots, k_n)$ в пространстве настраиваемых параметров.

Изложенное показывает, что цепь настройки в любой самонастраивающейся системе должна содержать измерительные и вычислительные устройства, определяющие в процессе работы системы критерий качества I и параметры настройки, исполнительные устройства, осуществляющие изменение параметров системы, и функциональные преобразователи, вырабатывающие сигналы управления исполнительными устройствами в зависимости от параметров настройки. Функциональные преобразователи предназначены для обеспечения необходимого качества процесса настройки.

Таким образом, самонастраивающиеся системы, помимо неполной начальной информации о свойствах объекта управления, используют дополнительную информацию, получаемую в процессе управления. В этом отношении самонастраивающиеся системы являются первым шагом по пути использования принципов самосовершенствования и самообучения, широко распространенных в живой природе.

Корректирующие цепи и цепи настройки в самонастраивающихся системах могут быть непрерывными, дискретными или смешанными (например, непрерывное корректирующее устройство с дискретной цепью настройки параметров). Применение непрерывных или дискретных устройств в каждом конкретном случае определяется необходимостью достижения наибольшей простоты и надежности проектируемой системы.

К самонастраивающимся системам полностью применимы общие методы исследования автоматических систем, изложенные в предыдущих главах. Для применения этих методов необходимо к уравнениям, описывающим поведение объекта и системы управления, добавить уравнения, описывающие работу цепей настройки параметров системы. Полученная в результате система уравнений всегда будет содержать столько уравнений, сколько имеется неизвестных функций времени, определяющих состояние объекта и всей системы управления. Эта система уравнений полностью определяет оператор самонастраивающейся автоматической системы.

Характерной особенностью полной системы уравнений, описывающих поведение самонастраивающейся системы, является то, что она всегда нелинейна, вследствие того что в уравнения, описывающие работу системы управления (закон управления), входят

произведения неизвестных — настраиваемых параметров системы управления и координат объекта управления, входящих в состав параметров управления. Вследствие этого в большинстве случаев приходится ограничиваться приближенным исследованием самонастраивающихся систем с помощью методов линеаризации, изложенных в главах 8, 10 и 11.

§ 17.3. Стабилизация динамических характеристик объекта управления с помощью корректирующих цепей

В предыдущем параграфе было отмечено, что расширения диапазона внешних условий работы системы управления в некоторых случаях удастся добиться путем стабилизации динамических характеристик объекта, уменьшения их чувствительности к внешним условиям. В таких случаях можно обойтись без применения настройки параметров системы управления. А так как всегда желательно обойтись более простой системой с постоянными параметрами, если это возможно, то целесообразно исследовать, что может дать стабилизация динамических характеристик объекта управления, прежде чем принимать решение о вводе в систему тех или иных цепей настройки параметров.

Для пояснения основной идеи стабилизации характеристик объекта управления рассмотрим стационарный линейный объект управления с передаточной функцией $\Phi(s)$, параметры которой, в зависимости от условий применения объекта, могут иметь различные значения *). Для изменения передаточной функции объекта управления таким образом, чтобы сделать ее при любых внешних условиях близкой к заданной эталонной передаточной функции $\Phi_0(s)$, можно ввести в систему управления модель с эталонной передаточной функцией $\Phi_0(s)$, подать на ее вход сигнал управления объектом x и разность выходных переменных модели y_0 и объекта y добавить с большим коэффициентом усиления k к сигналу управления x . Иными словами, к сигналу управления объектом x следует добавить разность между желаемой выходной переменной y_0 и действительной выходной переменной объекта управления y с большим коэффициентом усиления, а для формирования желаемой выходной переменной y_0 использовать модель с заданной передаточной функцией $\Phi_0(s)$. Структура такой системы управления представлена на рис. 17.3.1. Выполняя элементарные структурные преобразования методами § 4.6, убеждаемся в том, что

*) Строго говоря, при изменении условий применения такой объект уже не будет стационарной системой и, следовательно, понятие передаточной функции теряет для него смысл. Однако если параметры объекта изменяются достаточно медленно, чтобы их можно было считать постоянными в течение отрезков времени, значительно превосходящих время переходного процесса объекта, то объект можно приближенно считать стационарным.

эти операции равноценны охвату объекта управления отрицательной обратной связью с большим коэффициентом усиления k и введению дополнительного последовательного корректирующего устройства с передаточной функцией $1 + k\Phi_0(s)$ (пунктирный прямоугольник на рис. 17.3.2). Таким образом, изложенный способ стабилизации характеристик объекта основан на применении обычных корректирующих цепей и не приводит к самонастраивающимся системам.

Чтобы доказать, что изложенный способ действительно приближает передаточную функцию объекта к желаемой и делает ее слабо зависимой от внешних условий, найдем передаточную функцию $\Phi_1(s)$ полученной в результате

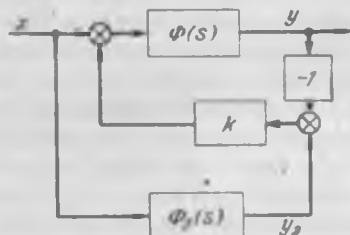


Рис. 17.3.1.

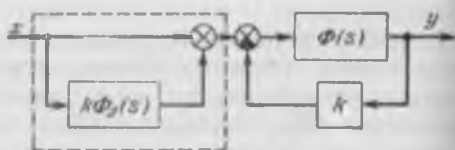


Рис. 17.3.2.

системы (рис. 17.3.2). Для нахождения этой передаточной функции воспользуемся формулами (4.5.1) и (4.5.7). В результате получим

$$\Phi_1(s) = \frac{[1 + k\Phi_0(s)] \Phi(s)}{1 + k\Phi(s)} = \frac{\left[\frac{1}{k} + \Phi_0(s)\right] \Phi(s)}{\frac{1}{k} + \Phi(s)}. \quad (17.3.1)$$

При $k \rightarrow \infty$ правая часть этой формулы стремится к $\Phi_0(s)$. Поэтому при достаточно больших значениях k передаточная функция объекта с введенными корректирующими цепями будет близка к $\Phi_0(s)$ и, следовательно, будет слабо зависеть от внешних условий.

Пример 17.3.1. Для иллюстрации рассмотрим управление самолетом или крылатой ракетой в вертикальной плоскости. На основании изложенного в § 3.16 движение самолета вокруг центра масс в вертикальной плоскости описывается приближенным линейным уравнением

$$\ddot{\alpha} + c_{\dot{\alpha}} \dot{\alpha} + c_{\alpha} \alpha = c_0 - c_0 \delta. \quad (17.3.2)$$

При горизонтальном полете с постоянной скоростью все коэффициенты уравнения (17.3.2) постоянны и, следовательно, динамические свойства самолета можно охарактеризовать передаточной функцией

$$\Phi_c(s) = -\frac{c_{\delta}}{s^2 + c_{\dot{\alpha}} s + c_{\alpha}}. \quad (17.3.3)$$

Для автоматического управления самолетом необходимо поставить на него рулевую машину, обеспечивающую управление отклонением руля δ

в зависимости от сигнала управления x по заданному закону. Считая рулевую машину приближенно линейной системой, обозначим ее передаточную функцию через $\Phi_p(s)$. Результатом работы системы управления является создание управляющей силы, вызывающей нормальное ускорение самолета, изменяющее направление его полета в соответствии с сигналом управления x . Поэтому интересующей нас выходной переменной объекта управления — самолета — является его нормальное ускорение w_N , вызываемое управляющей аэродинамической силой. Это ускорение при малых углах атаки α

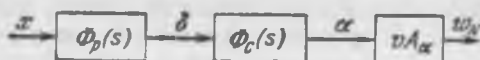


Рис. 17.3.3.

можно считать равным $vA_\alpha\alpha$, где A_α — коэффициент, зависящий от скорости полета и плотности воздуха (см. § 3.16). Таким образом, система, состоящая из самолета и рулевой машины (рис. 17.3.3), при горизонтальном равномерном полете является приближенно стационарной линейной системой с передаточной функцией

$$\Phi(s) = -\frac{c_\delta A_\alpha \Phi_p(s) v}{s^2 + c_\alpha s + c_\alpha} \quad (17.3.4)$$

Эта передаточная функция зависит от скорости и высоты полета вследствие зависимости коэффициентов c_δ , c_α , c_α и A_α от скорости и плотности воздуха,

Если самолет предназначен для полета с различными скоростями в большом диапазоне высот, то коэффициенты c_δ , c_α , c_α и A_α могут изменяться в сотни раз. Вследствие этого одному и тому же сигналу управления x при различных условиях полета будут соответствовать разные нормальные ускорения

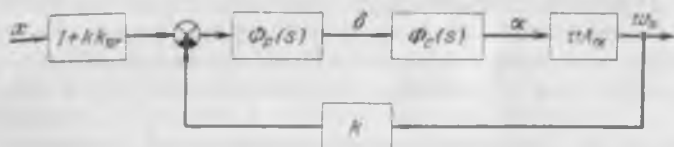


Рис. 17.3.4.

самолета, изменяющиеся в сотни раз при изменении условий полета. Между тем желательно, чтобы самолет реагировал на данный сигнал управления x по возможности всегда одним и тем же нормальным ускорением w_N независимо от условий полета. Поэтому примем в качестве желаемой зависимости между сигналом управления x и нормальным ускорением w_N простейшую зависимость вида $w_N = k_w x$ и применим изложенный способ стабилизации характеристик самолета. Так как эталонная передаточная функция в данном случае равна просто постоянной $\Phi_p(s) = k_w$, то для стабилизации характеристик самолета достаточно поставить на него измеритель поперечного ускорения, которое при малом угле атаки приближенно равно w_N , и ввести в систему управления отрицательную обратную связь по поперечному ускорению с большим коэффициентом усиления k с одновременным вводом усиления с коэффициентом $1 + kk_w$ сигнала управления x (рис. 17.3.4).

Изложенный метод стабилизации характеристик объекта управления применим и к нестационарным линейным объектам, а также к нелинейным объектам с достаточно гладкими характеристиками.

Практические возможности применения изложенного метода стабилизации характеристик объекта управления ограничиваются двумя обстоятельствами. Во-первых, коэффициент усиления k нельзя брать произвольно большим, не нарушая устойчивости системы. Поэтому практически допустимо взять лишь максимальное значение k , при котором вся система в целом имеет достаточные запасы устойчивости. Во-вторых, любой объект содержит нелинейности. Так, например, в рассмотренном выше примере отклонение руля самолета имеет ограничения. Рулевая машина также является нелинейной системой, так как ее скорость ограничена. С увеличением усиления сигнала управления пропорционально растет уровень шумов на входах нелинейных звеньев, что приводит к уменьшению эффективных коэффициентов усиления этих звеньев. При больших значениях k эффективные коэффициенты усиления нелинейных звеньев по полезному сигналу (т. е. статистические коэффициенты усиления по полезному сигналу) будут близки к нулю и вследствие этого нелинейные звенья практически не будут работать. Поэтому возможности увеличения коэффициента усиления k ограничиваются требованием, чтобы уровень шумов на входах нелинейных звеньев системы управления был достаточно мал. Отсюда следует, что, выбирая значение коэффициента усиления k , необходимо производить статистический расчет системы управления методом статистической линеаризации и вычислять передаточную функцию $\Phi_1(s)$ объекта управления с корректирующими цепями с учетом эффективных статистических коэффициентов усиления нелинейных звеньев системы управления. Только так можно будет выбрать оптимальное значение k , при котором достигается наилучшее приближение передаточной функции $\Phi_1(s)$ объекта с корректирующими цепями к желаемой передаточной функции $\Phi_0(s)$. При этом в систему желательно ввести фильтры, подавляющие, насколько возможно, помехи во входных сигналах нелинейных звеньев.

Методы определения оптимальных фильтров и оценки возможностей отделения сигналов от помех изложены в главах 13, 14 и 15.

Вследствие рассмотренных причин невозможно добиться полной стабилизации динамических характеристик объекта управления и приближения их к заданным с любой степенью точности. Однако изложенный метод все же позволяет во много раз сократить диапазон изменения характеристик объекта управления при заданном диапазоне внешних условий и вследствие этого значительно расширить возможности применения простейших систем управления с постоянными параметрами.

§ 17.4. Самонастраивающиеся системы с настройкой параметров по эталону

Самонастраивающаяся система с настройкой параметров по эталону содержит модель, динамические характеристики которой принимаются за эталон, к которому желательно приблизить

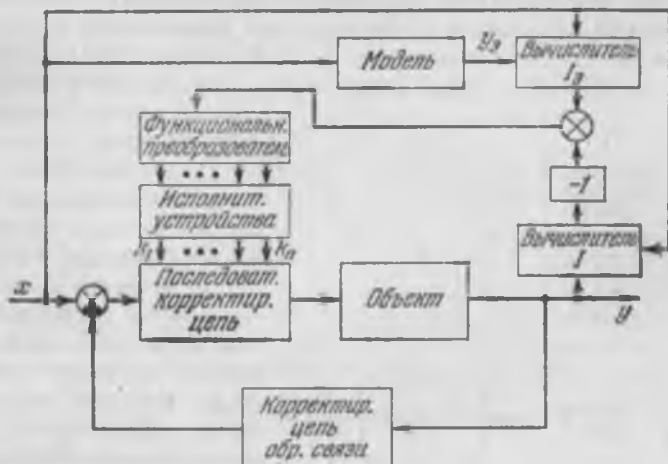


Рис. 17.4.1.

характеристики системы, вычислители критерия качества управления для системы и для модели, устройство сравнения, вырабатывающее параметр настройки ∇ , и функциональные преобразователи, вырабатывающие сигналы управления для исполнительных устройств, осуществляющих изменение параметров системы управления $k_1, \dots, k_n$.

На рис. 17.4.1 представлена структурная схема системы в случае, когда критерий качества управления для системы I и для модели I_0 могут быть вычислены непосредственно по входным и выходным сигналам системы и модели. Очевидно, что такая система далеко не всегда может быть отнесена к классу самонастраивающихся. Так, например, если критерий качества управления I зависит только от текущих значений выходной переменной y , ее производных и интегралов от нее, то такая система не является самонастраивающейся, а представляет собой обычную систему со сложной нелинейной корректирующей цепью. Чтобы сделать это ясным, рассмотрим частный случай следящей системы, для которой в качестве модели с эталонными характеристиками естественно принять идеальную следящую систему с передаточной функцией $\Phi_0(s) = 1$. В этом случае за критерий качества системы

можно принять ее ошибку $x - y$. Тогда для модели критерий качества I будет равен нулю и параметр настройки совпадет с параметром управления $\nabla = \Delta = x - y$ (рис. 17.4.2). В этом случае корректирующая цепь с цепью настройки осуществляет сложное нелинейное преобразование параметра управления и никаких элементов самонастройки не содержит.

Заметим, что такие системы управления с нелинейными корректирующими цепями часто бывают весьма эффективны и обеспечивают, как показали исследования Л. Г. Евланова, высокое быстродействие и качество управления даже при быстром изменении

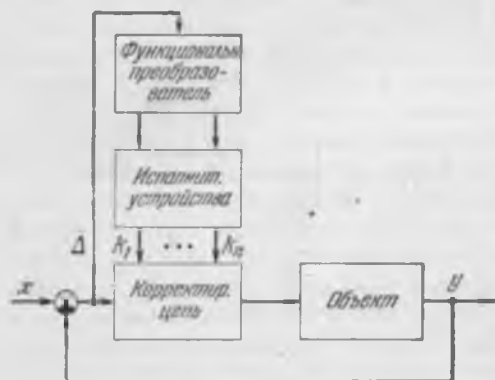


Рис. 17.4.2.

динамических характеристик объекта управления.

Для того чтобы система рассматриваемого типа была самонастраивающейся, необходимо, чтобы критерий качества I содержал некоторую информацию о динамических характеристиках объекта управления, которая могла бы быть использована для целенаправленного изменения параметров системы управления, а не была бы просто функцией текущих координат объекта управ-

ления. Если критерий качества управления I содержит информацию о динамических характеристиках объекта управления, то для его вычисления необходим сложный анализ поведения объекта, включающий операции не только над выходными, но и над входными сигналами объекта. В этом случае система управления с настройкой параметров по эталону получает новое качество и может быть отнесена к классу самонастраивающихся.

В некоторых случаях критерий качества управления I можно определить с помощью дополнительных сигналов поиска (зондирующих сигналов), вводимых в сигнал управления. К системам такого типа мы приходим в случае, когда за критерий качества управления I принята реакция системы на какой-нибудь типовой сигнал: весовая функция, переходная функция, частотная характеристика и т. д. В этом случае для определения критерия качества управления I можно периодически вводить в сигнал управления поисковые сигналы соответствующей формы. Соответственно цепь настройки должна содержать генератор поисковых сигналов, устройства для определения возмущений выходных сигналов объекта, вызываемых поисковыми сигналами, устройства для

вычисления параметра настройки ∇ и исполнительные устройства (рис. 17.4.3).

В качестве поисковых сигналов для определения весовой функции можно применить кратковременные импульсы или случайные сигналы, близкие к белому шуму. Для определения переходной функции целесообразно применить в качестве поисковых сигналов ступенчатые сигналы. Для определения частотной

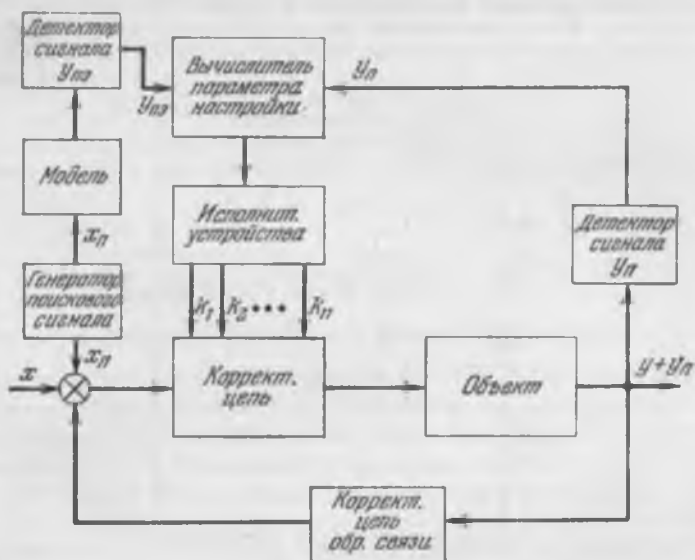


Рис. 17.4.3.

характеристики можно применить в качестве поисковых сигналов гармонические колебания или случайные сигналы, близкие к белому шуму.

При любом выборе критерия качества управления и типа поисковых сигналов поисковые сигналы не должны заметно возмущать систему и нарушать ее работу. Поэтому поисковые сигналы должны быть достаточно малыми. Вследствие этого для выделения результатов влияния поисковых сигналов на выходные сигналы системы и модели в большинстве случаев необходимо применять статистические методы как наиболее тонкие методы, позволяющие с максимально возможной точностью отделять слабые сигналы от помех и других посторонних сигналов (эти методы изложены в главах 13, 14 и 15). Соответственно и вычислительные устройства, определяющие параметр настройки, должны быть в большинстве случаев основаны на принципах статистической теории оптимального отделения сигналов от помех.

Пример 17.4.1. В качестве примера на рис. 17.4.4 дана структурная схема системы с автоматической настройкой коэффициента усиления последовательной корректирующей цепи с помощью синусоидального поискового сигнала небольшой амплитуды $x_n = a_n \sin \omega_n t$. Этот сигнал подается на вход системы вместе с сигналом управления x . Одновременно поисковый сигнал подается в вычислитель параметра настройки V через линейную модель, обладающую эталонной передаточной функцией $\Phi_0(s)$. В этот же вычислитель подается выходной сигнал системы. Резонансный фильтр выделяет из выходного сигнала системы ее реакцию на поисковый сигнал, равную $y_n = a_y \sin(\omega_n t + \psi_y)$. Из этого сигнала формируется параметр настройки $V = a_{y0} - a_y$, где a_{y0} — эталонная амплитуда колебаний выходной переменной. Сигнал, пропорциональный параметру настройки V , действует на

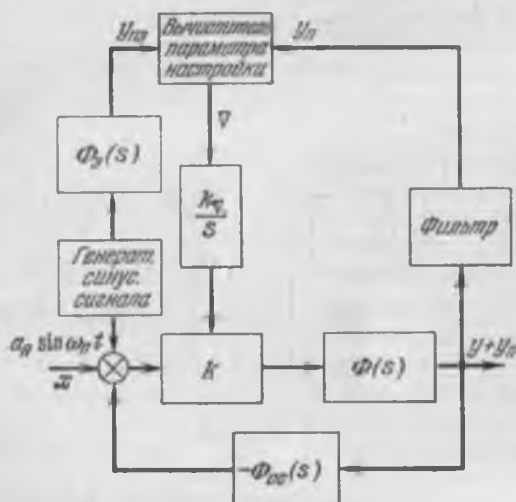


Рис. 17.4.4.

исполнительное устройство, изменяющее коэффициент усиления корректирующей цепи. Считая для простоты, что исполнительное устройство обеспечивает безынерционное слежение скорости изменения параметра k за управляющим сигналом, пропорциональным параметру настройки V , можем написать уравнение работы исполнительного устройства в виде

$$\dot{k} = k_v V. \quad (17.4.1)$$

Таким образом, исполнительное устройство в этом случае является интегратором с коэффициентом усиления k_v .

Чтобы пояснить работу такой самонастраивающейся системы, исследуем приближенно работу ее цепи настройки. Для этого предположим, что все элементы, определяющие амплитудную характеристику системы и параметр настройки, работают идеально точно. Тогда будем иметь

$$a_y = a_n \left| \frac{k\Phi(i\omega_n)}{1 + k\Phi(i\omega_n)\Phi_{oc}(i\omega_n)} \right|. \quad (17.4.2)$$

Так как эталонный коэффициент усиления амплитуды на частоте поискового сигнала ω_n является постоянной величиной $|\Phi_0(i\omega_n)|$, то нет необходимости вводить в рассматриваемую систему модель. Вычитая величину (17.4.2) из

$a_{\text{ув}} = a_{\text{п}} |\Phi_{\text{п}}(i\omega_{\text{п}})|$ и подставляя полученное выражение параметра настройки ∇ в уравнение (17.4.1), получим приближенное уравнение работы цепи настройки в виде

$$\dot{k} = k_{\nabla} a_{\text{п}} \left\{ |\Phi_{\text{э}}(i\omega_{\text{п}})| - \left| \frac{k\Phi(i\omega_{\text{п}})}{1 + k\Phi(i\omega_{\text{п}})\Phi_{\text{ос}}(i\omega_{\text{п}})} \right| \right\}. \quad (17.4.3)$$

Это — нелинейное дифференциальное уравнение первого порядка относительно k .

Параметры передаточной функции объекта могут быть неизвестными. Однако при любых значениях этих параметров существует такое значение коэффициента усиления k , при котором усиление амплитуды на частоте поиска $\omega_{\text{п}}$ в точности равно требуемому усилению $|\Phi_{\text{э}}(i\omega_{\text{п}})|$. Обозначим это значение k через $k_{\text{э}}$ и выполним линеаризацию уравнения (17.4.3) относительно $k - k_{\text{э}}$. Так как правая часть уравнения (17.4.3) является непрерывной гладкой (т. е. с непрерывной производной) функцией k , то можно применить метод обычной линеаризации при помощи ряда Тейлора. Тогда уравнение (17.4.3) заменится приближенным линейным уравнением

$$\dot{k} = -c(k - k_{\text{э}}), \quad (17.4.4)$$

где

$$c = k_{\nabla} a_{\text{п}} \left[\frac{\partial}{\partial k} \left| \frac{k\Phi(i\omega_{\text{п}})}{1 + k\Phi(i\omega_{\text{п}})\Phi_{\text{ос}}(i\omega_{\text{п}})} \right| \right]_{k=k_{\text{э}}}. \quad (17.4.5)$$

Процесс настройки устойчив при $c > 0$ и неустойчив при $c < 0$.

Заметим, что весовую функцию и частотную характеристику системы можно определить и без поисковых сигналов с помощью соответствующей статистической обработки входного и выходного сигналов объекта управления или всей системы. В этом случае в систему управления должны быть введены специальные вычислительные устройства, содержащие в качестве одного из элементов коррелятор — вычислитель корреляционных и взаимных корреляционных функций входных и выходных сигналов.

§ 17.5. Самонастраивающиеся системы с экстремальной настройкой параметров

Самонастраивающиеся системы с экстремальной настройкой параметров всегда содержат устройства для поиска экстремума критерия качества управления I . Как было отмечено в § 17.2, в качестве величин, характеризующих отклонения от экстремума величины I — параметров настройки, — можно принять частные производные величины I по настраиваемым параметрам системы $k_1, \dots, k_n$, т. е. составляющие вектора градиента функции I в пространстве параметров $k_1, \dots, k_n$. Следует, однако, заметить, что критерий качества управления I является обычной функцией параметров системы $k_1, \dots, k_n$ только в том случае, когда эти параметры в процессе управления остаются постоянными. Если же параметры $k_1, \dots, k_n$ изменяются в процессе управления, то критерий качества управления I зависит не только от текущих значе-

ний параметров $k_1, \dots, k_n$, но и от закона их изменения в течение всего времени управления, т. е. является функционалом от $k_1, \dots, k_n$. И лишь в том случае, если параметры $k_1, \dots, k_n$ изменяются достаточно медленно, чтобы можно было считать их постоянными в течение любого промежутка времени, в несколько раз превосходящего время переходного процесса в системе, можно считать I функцией текущих значений параметров $k_1, \dots, k_n$. Поэтому, для возможности оценки качества управления при изменении

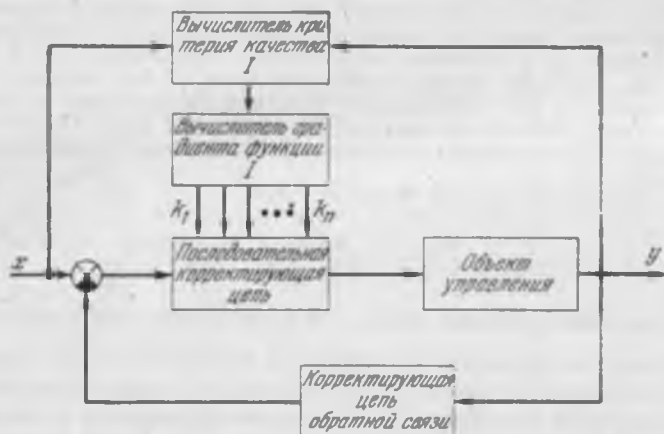


Рис. 17.5.1.

параметров $k_1, \dots, k_n$, необходимо, чтобы эти параметры изменялись достаточно медленно в процессе поиска экстремума величины I .

Общая схема самонастраивающейся системы с экстремальной настройкой параметров в случае, когда за параметры настройки приняты составляющие вектора градиента функции $I(k_1, \dots, k_n)$, представлена на рис. 17.5.1.

Для вычисления составляющих градиента функции I можно применить различные методы. Мы рассмотрим здесь четыре метода: метод конечных приращений, метод производной по времени, метод синхронного детектирования и метод интегрирования вспомогательной системы уравнений, определяющей производные функции I по параметрам $k_1, k_2, \dots, k_n$.

Метод конечных приращений основан на замене частных производных отношениями конечных приращений

$$\begin{aligned} \nabla_i I &= \frac{\partial I}{\partial k_i} \approx \frac{I(k_1, \dots, k_{i-1}, k_i + \Delta k_i, k_{i+1}, \dots, k_n) - I(k_1, \dots, k_n)}{\Delta k_i} = \\ &= \frac{\Delta_i I}{\Delta k_i} \quad (i = 1, \dots, n). \end{aligned} \quad (17.5.1)$$

В этом случае поисковые изменения параметров $k_1, \dots, k_n$ представляют собой поочередные небольшие скачкообразные изменения параметров $k_1, \dots, k_n$ с последующим вычислением соответствующих приращений ΔI . Точность этого метода зависит от характера функции I и величины поисковых приращений Δk_i .

Метод производной по времени основан на поочередном медленном изменении каждого из параметров системы с постоянной скоростью сначала в одну, а потом в другую сторону при постоянных значениях остальных параметров. Из формулы для полной производной функции I по времени

$$\dot{I} = \frac{\partial I}{\partial k_1} \dot{k}_1 + \frac{\partial I}{\partial k_2} \dot{k}_2 + \dots + \frac{\partial I}{\partial k_n} \dot{k}_n. \quad (17.5.2)$$

следует, что если все параметры, кроме k_i , постоянны, а параметр k_i изменяется с постоянной скоростью $\dot{k}_i$, то

$$\nabla I = \frac{\partial I}{\partial k_i} = \frac{1}{\dot{k}_i} \dot{I} \quad (i = 1, \dots, n). \quad (17.5.3)$$

Обычно параметры изменяются при этом методе по треугольному закону $|\dot{k}_i| = \text{const}$.

Недостатком обоих рассмотренных методов определения градиента функции I является то, что они позволяют определять производные $\partial I / \partial k_i$ только поочередно, вследствие чего время определения градиента функции I зависит от числа настраиваемых параметров. От этого недостатка свободен метод синхронного детектирования.

При определении частных производных $\partial I / \partial k_i$ методом синхронного детектирования настраиваемые параметры модулируются небольшими по амплитуде гармоническими колебаниями различных частот, т. е. к параметрам $k_1, \dots, k_n$ добавляются синусоидальные сигналы

$$\Delta k_i = a_i \sin \omega_i t \quad (i = 1, \dots, n) \quad (17.5.4)$$

различных частот $\omega_1, \dots, \omega_n$. Для определения влияния колебаний каждого параметра на критерий качества I используются синхронные детекторы. Каждый детектор выполняет умножение величины I на соответствующий гармонический сигнал $a_i \sin \omega_i t$ и осреднению полученного произведения по времени. Покажем, что выходные сигналы синхронных детекторов приблизительно пропорциональны частным производным функции I по соответствующим параметрам. Для этого заметим, что при малых амплитудах $a_1, \dots, a_n$ колебаний параметров $k_1, \dots, k_n$ функцию I можно разложить в ряд Тейлора по приращениям параметров $\Delta k_1, \dots, \Delta k_n$ и ограничиться в этом разложении членами первой степени относительно

$\Delta k_1, \dots, \Delta k_n$. В результате получим

$$I = I(k_1 + \Delta k_1, \dots, k_n + \Delta k_n) \approx I(k_1, \dots, k_n) + \sum_{l=1}^n \frac{\partial I}{\partial k_l} a_l \sin \omega_l t. \quad (17.5.5)$$

Умножая это равенство на $a_l \sin \omega_l t$ и осредняя результат по времени, получим

$$\begin{aligned} \frac{1}{T} \int_0^T I a_l \sin \omega_l t dt &\approx I(k_1, \dots, k_n) a_l \cdot \frac{1}{T} \int_0^T \sin \omega_l t dt + \\ &+ \sum_{l=1}^n \frac{\partial I}{\partial k_l} a_l a_l \cdot \frac{1}{T} \int_0^T \sin \omega_l t \sin \omega_l t dt. \end{aligned} \quad (17.5.6)$$

Но

$$\frac{1}{T} \int_0^T \sin \omega_l t dt = \frac{1 - \cos \omega_l T}{\omega_l T},$$

$$\frac{1}{T} \int_0^T \sin \omega_l t \sin \omega_l t dt = \frac{1}{2T} \left[\frac{\sin(\omega_l - \omega_l) T}{\omega_l - \omega_l} - \frac{\sin(\omega_l + \omega_l) T}{\omega_l + \omega_l} \right] \text{ при } l \neq i,$$

$$\frac{1}{T} \int_0^T \sin^2 \omega_l t dt = \frac{1}{2} - \frac{\sin 2\omega_l T}{4\omega_l T}.$$

Если взять время осреднения T кратным всем периодам $2\pi/\omega_1, \dots, 2\pi/\omega_n$, то первые два интеграла будут равны нулю, а третий — половине. При любом достаточно большом времени осреднения T первые два интеграла будут как угодно малы, а третий как угодно близок к $1/2$. Следовательно, формула (17.5.6) может быть переписана в виде

$$\frac{1}{T} \int_0^T I a_l \sin \omega_l t dt \approx \frac{a_l^2}{2} \frac{\partial I}{\partial k_l} \quad (l = 1, 2, \dots, n). \quad (17.5.7)$$

Заметим, что в качестве поисковых колебаний параметров системы можно брать не только синусоидальные колебания, но и любые другие типы колебаний, удовлетворяющие единственному условию, чтобы средние значения произведений колебаний различных параметров за время осреднения были достаточно близки к нулю. Можно также использовать в качестве поисковых колебаний независимые стационарные эргодические случайные колебания. В некоторых случаях можно использовать в качестве поисковых колебаний естественные автоколебания, возникающие в системе.

Метод синхронного детектирования позволяет вычислять сразу все составляющие вектора градиента функции I , вследствие чего время, потребное на определение градиента, не зависит от числа настраиваемых параметров. Однако это время все же в несколько раз превосходит время переходного процесса в системе.

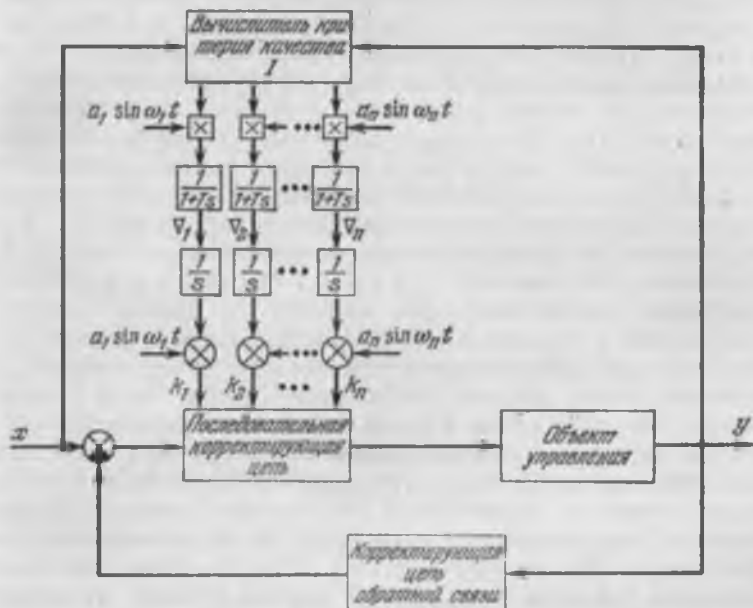


Рис. 17.5.2.

Для непрерывного осреднения по времени произведений $I a_i \sin \omega_i t$ можно использовать апериодические звенья с большими постоянными времени T , например цепочки RC . В этом случае интеграл в (17.5.7) заменяется интегралом в пределах от t_0 до t от произведения $I a_i \sin \omega_i t$ на весовую функцию апериодического звена:

$$\nabla_i = \frac{\partial I}{\partial k_i} \approx \frac{2}{a_i^2 T} \int_{t_0}^t e^{-\frac{t-\tau}{T}} I a_i \sin \omega_i \tau d\tau \quad (i = 1, \dots, n). \quad (17.5.8)$$

Общая схема самонастраивающейся системы, в которой составляющие вектора градиента функции I вычисляются методом синхронного детектирования с помощью апериодических звеньев, представлена на рис. 17.5.2.

Заметим, что для определения составляющих градиента функции I всеми изложенными способами можно также применить

методы оптимальной статистической обработки информации, изложенные в предыдущих трех главах.

Рассмотренные методы определения градиента критерия качества управления используют очень небольшой объем априорной информации об объекте управления и вследствие этого требуют значительного времени на поиск экстремума. Время, необходимое для определения градиента критерия качества, а следовательно и на поиск его экстремума, можно существенно сократить, если использовать значительную априорную информацию об объекте управления, во многих случаях имеющуюся в распоряжении конструктора. Так, например, если известны дифференциальные уравнения объекта, можно легко составить другие, вспомогательные уравнения, определяющие частные производные выходных переменных системы по параметрам системы управления $k_1, \dots, k_n$, решая которые, можно вычислить и производные критерия качества управления по параметрам $k_1, \dots, k_n$. Вводя в систему быстродействующую вычислительную машину, например используя моделирование в сокращенном масштабе времени, можно существенно ускорить вычисление градиента критерия качества. Для применения этого метода необходимо, чтобы были известны не только уравнения объекта управления, но и все входящие в эти уравнения параметры, которые заранее не могут быть известными. Поэтому, используя этот метод ускоренного определения градиента критерия качества управления, необходимо ввести в систему управления дополнительные устройства, определяющие неизвестные параметры объекта управления $l_1, \dots, l_m$. Для этого можно использовать в системе управления модель объекта управления с автоматической настройкой ее параметров с таким расчетом, чтобы эти параметры были возможно более близкими к параметрам объекта управления [44, 22] (так называемая *подслеживающая модель*). При этом для наилучшего приближения параметров модели к параметрам объекта управления можно использовать экстремальную настройку, обеспечивающую экстремум некоторого критерия I_1 близости поведения модели к поведению объекта. Общая структурная схема такой системы приведена на рис. 17.5.3.

Для поиска экстремума критерия качества I можно также использовать различные другие методы. В частности, можно использовать статистические методы нахождения экстремума или градиента I по случайному набору значений параметров $k_1, \dots, k_n$.

В качестве критерия качества управления в системах с экстремальной настройкой часто принимается среднее значение квадрата параметра управления (сигнала ошибки), характеризующего ошибку управления:

$$I = \frac{1}{T} \int_{t-T}^t \Delta^2(\tau) d\tau. \quad (17.5.9)$$

Ввиду того что организовать непрерывное вычисление по формуле (17.5.9) затруднительно, так же как и при вычислении средних значений произведений $Ia_1 \sin \omega_1 t$ при определении градиента методом синхронного детектирования, для осреднения величины

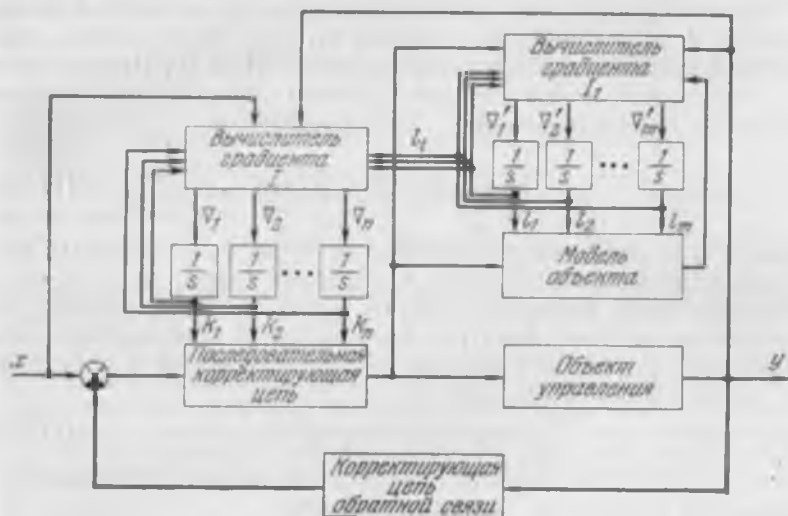


Рис. 17.5.3.

$\Delta^2(\tau)$ часто используют аperiodическое звено с большой постоянной времени T . Соответственно формула (17.5.9) заменяется формулой

$$I = \frac{1}{T} \int_{t_0}^t e^{-\frac{t-\tau}{T}} \Delta^2(\tau) d\tau. \quad (17.5.10)$$

Аналогично при определении градиента критерия качества методом ускоренного вычисления производных с использованием известных уравнений объекта управления за критерий приближения модели к объекту обычно принимают среднее значение квадрата ошибки модели, т. е. разности между ее выходным сигналом y_m и выходным сигналом объекта y :

$$I_1 = \frac{1}{T} \int_{t-T}^t [y_m(\tau) - y(\tau)]^2 d\tau. \quad (17.5.11)$$

При использовании для осреднения аperiodического звена с большой постоянной времени T формула (17.5.11) заменяется

формулой

$$I_1 = \frac{1}{T} \int_{t_0}^t e^{-\frac{t-\tau}{T}} [y_n(\tau) - y(\tau)]^2 d\tau. \quad (17.5.12)$$

Покажем конкретнее, как вычисляются производные критерия качества I по параметрам системы $k_1, \dots, k_n$ в случае, когда критерий качества определяется формулой (17.5.9). Предположим, что поведение автоматической системы описывается системой линейных дифференциальных уравнений

$$\dot{y}_l = \sum_{k=1}^N a_{lk}(t) y_k + F_l(t) \quad (l=1, \dots, N), \quad (17.5.13)$$

где $y_1 = y$ — выходная переменная системы, y_2 — выходная переменная корректирующей цепи в обратной связи, $y_3, \dots, y_N$ — вспомогательные переменные, $F_l(t)$ — внешние возмущения, действующие на систему, $a_{lk}(t)$ — коэффициенты системы, зависящие от времени и параметров системы $k_1, \dots, k_n$. Параметр управления $\Delta(t)$ выражается формулой

$$\Delta(t) = x(t) - y_2(t), \quad (17.5.14)$$

где $x(t) = F_1(t)$ — входной сигнал управления. Подставляя это выражение в формулу (17.5.9), получим

$$I = \frac{1}{T} \int_{t-T}^t [y_2(\tau) - x(\tau)]^2 d\tau. \quad (17.5.15)$$

Продифференцируем это выражение по параметру k_i :

$$\frac{\partial I}{\partial k_i} = \frac{2}{T} \int_{t-T}^t [y_2(\tau) - x(\tau)] \frac{\partial y_2(\tau)}{\partial k_i} d\tau \quad (i=1, \dots, n). \quad (17.5.16)$$

Для определения производных $\partial y_2 / \partial k_i$ продифференцируем по параметру k_i уравнения (17.5.13). В результате получим систему уравнений, определяющую частные производные переменных y_l по параметрам k_i :

$$\frac{d}{dt} \frac{\partial y_l}{\partial k_i} = \sum_{k=1}^n a_{lk}(t) \frac{\partial y_k}{\partial k_i} + \sum_{k=1}^n \frac{\partial a_{lk}(t)}{\partial k_i} y_k \quad \begin{matrix} (l=1, \dots, N) \\ (i=1, \dots, n) \end{matrix}. \quad (17.5.17)$$

Эта система уравнений совместно с формулой (17.5.16) определяет частные производные критерия качества по настраиваемым параметрам. При вычислении частных производных $\partial y_l / \partial k_i$ используется априорная информация о системе в виде ее дифференциальных уравнений (17.5.13). Система уравнений (17.5.17) должна решаться при нулевых начальных условиях.

Пример 17.5.1. Рассмотрим самонастраивающуюся систему стабилизации продольной оси самолета или ракеты в вертикальной плоскости с экстремальной настройкой коэффициентов усиления сигналов от свободного и демпфирующего гироскопов. Движение самолета в прямолинейном горизонтальном полете при малых углах атаки α и малых отклонениях вектора скорости от горизонта θ без учета случайных возмущений описывается уравнениями (3.16.10):

$$\ddot{\alpha} + c_{\alpha} \dot{\alpha} + c_{\alpha\alpha} \alpha = c_0 - c_{\theta} \delta, \quad \dot{\theta} = A_{\alpha} \alpha - \frac{g}{v}, \quad \dot{\theta} = \theta + \alpha. \quad (17.5.18)$$

Свободный гироскоп измеряет отклонение оси самолета от горизонта (угол тангажа) θ , а демпфирующий гироскоп измеряет угловую скорость самолета $\dot{\theta}$. Предполагая для простоты, что эти измерения осуществляются абсолютно точно и что рулевая машина безынерционна, примем уравнение работы рулевой машины в виде

$$\delta = k_1 \theta + k_2 \dot{\theta}, \quad (17.5.19)$$

где k_1, k_2 — коэффициенты усиления. При полете с различными скоростями в большом диапазоне высот коэффициенты уравнений (17.5.18), как было отмечено в примере 17.3.1, изменяются в сотни раз. Чтобы система управления (автопилот) хорошо стабилизировала ось самолета при любых условиях полета, целесообразно применить экстремальную настройку параметров k_1, k_2 . При этом качество управления естественно считать тем лучшим, чем меньше вектор скорости самолета отклоняется от горизонта. В этом случае за критерий качества управления естественно принять среднее значение квадрата угла отклонения вектора скорости самолета от горизонта:

$$I = \frac{1}{T} \int_0^T \theta^2(t) dt. \quad (17.5.20)$$

Для определения градиента величины I применим метод вычисления ее производных по k_1 и k_2 с помощью уравнений движения самолета (17.5.18). Дифференцируя формулу (17.5.20) по k_1 и k_2 , получим

$$\nabla_i = \frac{\partial I}{\partial k_i} = \frac{2}{T} \int_0^T \theta \frac{\partial \theta}{\partial k_i} dt \quad (i=1, 2). \quad (17.5.21)$$

Для вычисления производных $\partial \theta / \partial k_i$ продифференцируем по k_1 и k_2 уравнения (17.5.18) и (17.5.19). В результате получим

$$\left. \begin{aligned} \frac{d^2}{dt^2} \frac{\partial \alpha}{\partial k_i} + c_{\alpha} \frac{d}{dt} \frac{\partial \alpha}{\partial k_i} + c_{\alpha\alpha} \frac{\partial \alpha}{\partial k_i} &= -c_{\theta} \frac{\partial \delta}{\partial k_i}, \\ \frac{d}{dt} \frac{\partial \theta}{\partial k_i} &= A_{\alpha} \frac{\partial \alpha}{\partial k_i}, \quad \frac{\partial \theta}{\partial k_i} = \frac{\partial \theta}{\partial k_i} + \frac{\partial \alpha}{\partial k_i} \quad (i=1, 2), \\ \frac{\partial \delta}{\partial k_1} &= \theta + k_1 \frac{\partial \theta}{\partial k_1} + k_2 \frac{d}{dt} \frac{\partial \theta}{\partial k_1}, \\ \frac{\partial \delta}{\partial k_2} &= \dot{\theta} + k_1 \frac{d \theta}{d k_2} + k_2 \frac{d}{dt} \frac{\partial \theta}{\partial k_2}. \end{aligned} \right\} \quad (17.5.22)$$

Таким образом, в системе управления должны быть вычислительные устройства, интегрирующие уравнения (17.5.22) и вычисляющие производные критерия качества I по формуле (17.5.21).

Заметим, что систему уравнений (17.5.22) можно упростить, если не учитывать инерционности движения самолета вокруг центра массы. Тогда первое уравнение (17.5.22) заменится приближенной формулой

$$\frac{\partial \alpha}{\partial k_i} = - \frac{c_\delta}{c_\alpha} \frac{\partial \delta}{\partial k_i} \quad (i=1, 2). \quad (17.5.23)$$

За счет этого упрощения составляющие градиента будут вычисляться с ошибками, однако это не окажет существенного влияния на точность настройки.

Так как коэффициенты c_δ , c_α , c_α и A_α в уравнениях движения самолета не могут считаться заранее известными, то для их определения необходимо ввести в систему управления модель, поведение которой тоже описывается уравнениями (17.5.18), но с другими значениями коэффициентов c'_δ , c'_α , c'_α , A'_α .

Для настройки этих коэффициентов с расчетом сделать их близкими к коэффициентам c_δ , c_α , c_α , A_α можно применить метод экстремальной

настройки по условию минимума среднего значения квадрата разности углов α , θ или $\bar{\theta}$ для самолета и для модели, например принять

$$I_1 = \frac{1}{T} \int_0^T (\theta_1 - \bar{\theta})^2 dt, \quad (17.5.24)$$

где θ_1 — значение угла θ для модели, а $\bar{\theta}$ — по-прежнему угол наклона оси самолета к горизонту. Коэффициенты c_δ , c_α , c_α , A_α уравнений движения самолета следует заменить в уравнениях (17.5.22) соответствующими коэффициентами c'_δ , c'_α , c'_α , A'_α уравнений модели.

Для составления уравнений, определяющих производные величины I_1 , по коэффициентам c'_δ , c'_α , c'_α , A'_α , следует продифференцировать по этим коэффициентам формулу (17.5.24) и уравнения модели. Мы предоставляем читателю вывести эти уравнения самостоятельно.

§ 17.6. О некоторых перспективах

Из изложенного в книге можно сделать вывод, что теория автоматического управления в настоящее время находится в состоянии быстрого развития, формирования новых идей и понятий, исследования возможностей практического использования ее теоретических достижений и прокладывания новых путей в технике. Темпы развития теории автоматического управления и ее достижения за последнее десятилетие дают основание предвидеть быстрый прогресс автоматике и ее широкое использование во всех областях человеческой деятельности, в первую очередь освобождение человеческого мозга от большого числа сравнительно элементарных стандартных функций, которые будут переданы автоматическим системам. Изложенные в книге методы теории автоматического управления представляют собой лишь первоначальную основу для ее дальнейшего развития и практических применений.

Мы видим, какую большую роль в организации автоматического или автоматизированного управления играет статистическая обработка информации, в результате которой должно быть принято определенное решение. Можно смело сказать, что проблема статистического решения является основной, кардинальной проблемой теории управления сложными системами (так называемыми «большими» системами) и сложными процессами. По существу, любой вопрос о передаче сигналов от одного элемента системы к другому и переработке информации представляет собой задачу статистического решения. Поэтому естественно стремление применять к задачам управления современные методы математической статистики и в первую очередь новейшего и наиболее совершенного ее раздела — теории статистических решений. Статистические методы теории автоматического управления, зародившиеся всего три десятка лет назад, получили сейчас широкое распространение, проникли буквально во все разделы теории процессов управления и практики проектирования автоматических систем. Современная статистическая теория оптимальных систем, изложенная в главах 13, 14 и 15, по существу, является прикладной теорией статистических решений и дает эффективные методы практического нахождения оптимальных решений применительно к задачам автоматического управления. Несомненно, что теория сложных систем, в частности систем одновременного управления большим количеством объектов, каждый из которых имеет возможность до некоторой степени самостоятельно определять свое поведение и принимать решения, может быть построена только на основе статистических методов. Развивающаяся в последнее время общая теория обучающихся систем различного назначения также основана на статистических методах, в частности на методах статистической теории оптимальных систем, изложенных в главах 13, 14 и 15 [57]. Нет сомнения, что и теория более сложных типов самоорганизующихся систем может быть построена только на основе статистических методов. Для этого, само собой разумеется, потребуется дальнейшее развитие статистических методов как в части теории, так и в части ее практического применения к разработке новых принципов устройства автоматических систем и их элементов.

Развитая в предыдущих главах статистическая теория оптимальных систем дает принципиальную возможность решать различные практические задачи обнаружения и выделения сигналов на фоне помех, возникающие в теории информации и в теории проектирования автоматических систем, одним общим методом. Таким образом, изложенная теория обладает большими возможностями. Она применима также к широкому классу задач, возникающих в других областях науки и техники. Так, например, эта теория применима к задаче автоматического обнаружения и определения характера неисправностей и разладок в автома-

тических производственных линиях, в различных машинах и в объектах живой природы. В частности, она принципиально применима к задачам обнаружения и определения характера функциональных расстройств человеческого организма, т. е. к задачам диагностики в медицине. Опыт применения этой теории в технике показывает, что она дает объективные научные способы обработки записей различных процессов, позволяющие обнаруживать и выделять из помех такие слабые сигналы, которые не могут быть замечены ни человеческими органами чувств, ни с помощью каких-либо других способов обработки сигналов. Это дает основание считать, что применение изложенной теории и в различных других областях науки будет успешным. Однако в настоящее время практика значительно отстает от уровня развития теории, и эта теория пока еще далеко не достаточно применяется для решения практических задач даже в области техники, а в области естественных наук и вообще пока почти не находит применения.

Даваемые развитой в главах 13—16 теорией алгоритмы оптимальных систем могут быть положены в основу устройства реальных автоматических систем, имеющих в своем составе быстродействующие математические машины, особенно цифровые, так как цифровые машины принципиально могут производить любые вычисления. Это открывает перспективы значительного роста техники автоматического управления и расширения возможностей автоматических систем. Однако эта теория пока еще главным образом раскрывает перспективы развития техники, прокладывает путь в будущее, так как быстродействие существующих математических машин недостаточно для массового практического применения систем, основанных на оптимальных алгоритмах обработки информации. Вследствие этого изложенная теория в настоящее время должна быть использована в первую очередь в поисковых работах по исследованию принципов построения систем, основанных на оптимальной обработке информации, и возможностей создания на основе этой теории различных типов автоматических систем.

ПРИЛОЖЕНИЯ

1. Основные характеристики

Наименование звена	Уравнение $y = Ax$	Передаточная функция $\Phi(s)$
Усилительное	$y = kx$	k
Запаздывающее	$y(t) = x(t - \tau)$	$e^{-s\tau}$
Интегрирующее	$y(t) = \int_{t_0}^t x(\tau) d\tau$	$\frac{1}{s}$
Дифференцирующее	$y = \dot{x}$	s
Апериодическое	$T\dot{y} + y = kx$	$\frac{k}{Ts + 1}$
Форсирующее первого порядка	$y = T\dot{x} + x$	$Ts + 1$
Неустойчивое апериодическое	$T\dot{y} - y = kx$	$\frac{k}{Ts - 1}$
Колебательное	$T^2\ddot{y} + 2T\zeta\dot{y} + y = kx$	$\frac{k}{T^2s^2 + 2T\zeta s + 1}$
Форсирующее второго порядка	$y = T^2\ddot{x} + 2T\zeta\dot{x} + x$	$T^2s^2 + 2T\zeta s + 1$
Резонансное	$T^2\ddot{y} + y = kx$	$\frac{k}{T^2s^2 + 1}$
Неустойчивое колебательное	$T^2\ddot{y} - 2T\zeta\dot{y} + y = kx$	$\frac{k}{T^2s^2 - 2T\zeta s + 1}$

элементарных стационарных звеньев

Амплитудная частотная характеристика $A(\omega)$	Фазовая частотная характеристика $\varphi(\omega)$	Весовая функция $g(\xi)$
k	0	$k\delta(\xi)$
1	$-\tau\omega$	$\delta(\xi - \tau)$
$\frac{1}{\omega}$	$-\frac{\pi}{2}$	$1(\xi)$
ω	$\frac{\pi}{2}$	$\delta'(\xi)$
$\frac{k}{\sqrt{T^2\omega^2 + 1}}$	$-\text{arctg } T\omega$	$\frac{k}{T} e^{-\frac{\xi}{T}}$
$\sqrt{T^2\omega^2 + 1}$	$\text{arctg } T\omega$	$T\delta'(\xi) + \delta(\xi)$
$\frac{k}{\sqrt{T^2\omega^2 + 1}}$	$\text{arctg } T\omega - \pi$	$\frac{k}{T} e^{\frac{\xi}{T}}$
$\frac{k}{\sqrt{(1 - T^2\omega^2)^2 + 4T^2\zeta^2\omega^2}}$	$-\text{arctg } \frac{2T\zeta\omega}{1 - T^2\omega^2}$	$\frac{k}{T\sqrt{1 - \zeta^2}} e^{-\frac{\xi}{T} \sin \frac{\xi}{T}} \sqrt{1 - \zeta^2}$
$\sqrt{(1 - T^2\omega^2)^2 + 4T^2\zeta^2\omega^2}$	$\text{arctg } \frac{2T\zeta\omega}{1 - T^2\omega^2}$	$T^2\delta''(\xi) + 2T\zeta\delta'(\xi) + \delta(\xi)$
$\frac{k}{ 1 - T^2\omega^2 }$	0 при $\omega < \frac{1}{T}$ π при $\omega > \frac{1}{T}$	$\frac{k}{T} \sin \frac{\xi}{T}$
$\frac{k}{\sqrt{(1 - T^2\omega^2)^2 + 4T^2\zeta^2\omega^2}}$	$\text{arctg } \frac{2T\zeta\omega}{1 - T^2\omega^2}$	$\frac{k}{T\sqrt{1 - \zeta^2}} e^{\frac{\xi}{T} \sin \frac{\xi}{T}} \sqrt{1 - \zeta^2}$

2. Номограммы для определения логарифмических частотных характеристик системы, замкнутой отрицательной жесткой обратной связью, по логарифмическим частотным характеристикам соответствующей разомкнутой системы, и наоборот

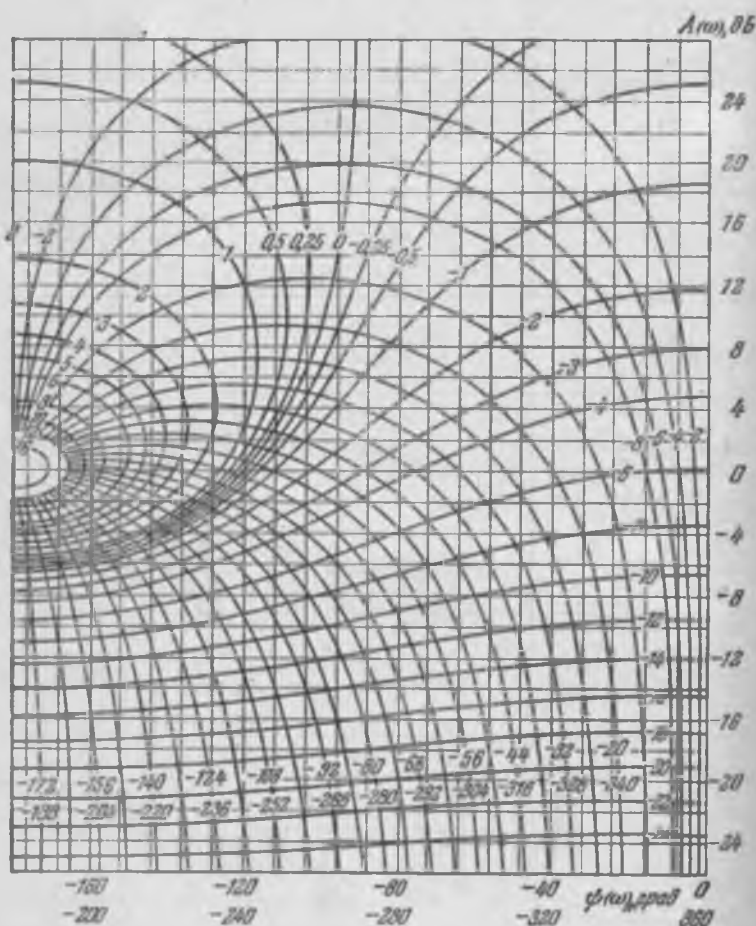
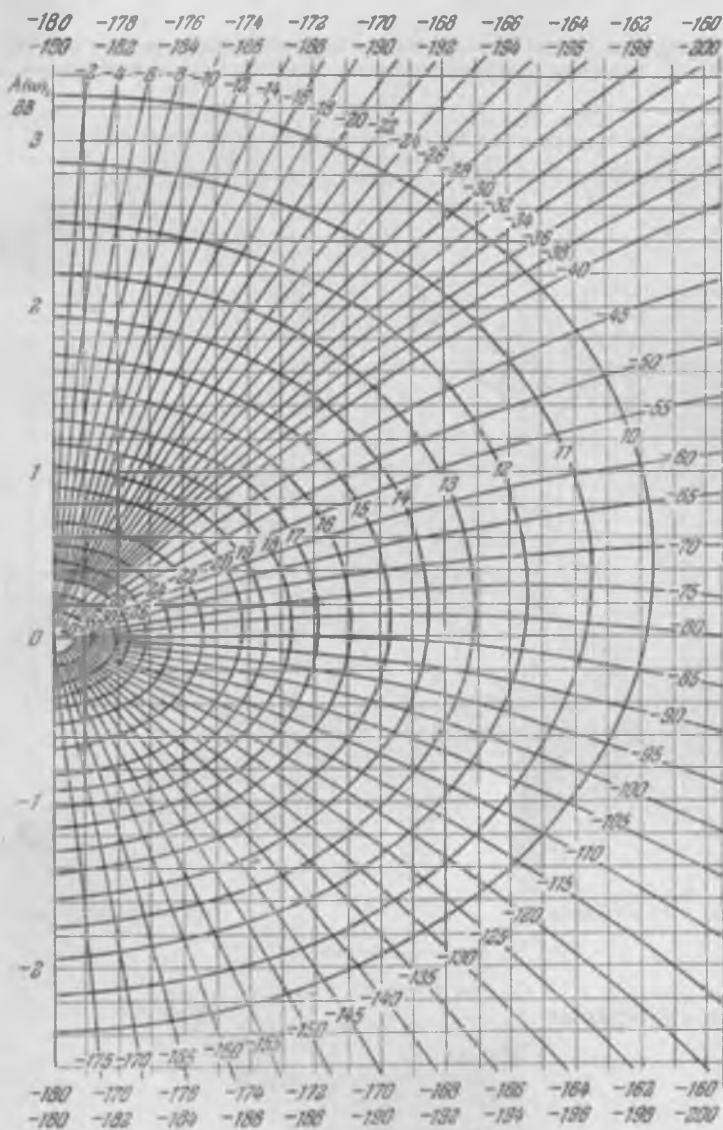


Рис. II.1 (а).

Продолжение



Фаза $\phi(\omega)$ в градусах

Рис. П.1 (б).

3. Номограммы для определения основных характеристик систем, имеющих типовые логарифмические амплитудные характеристики

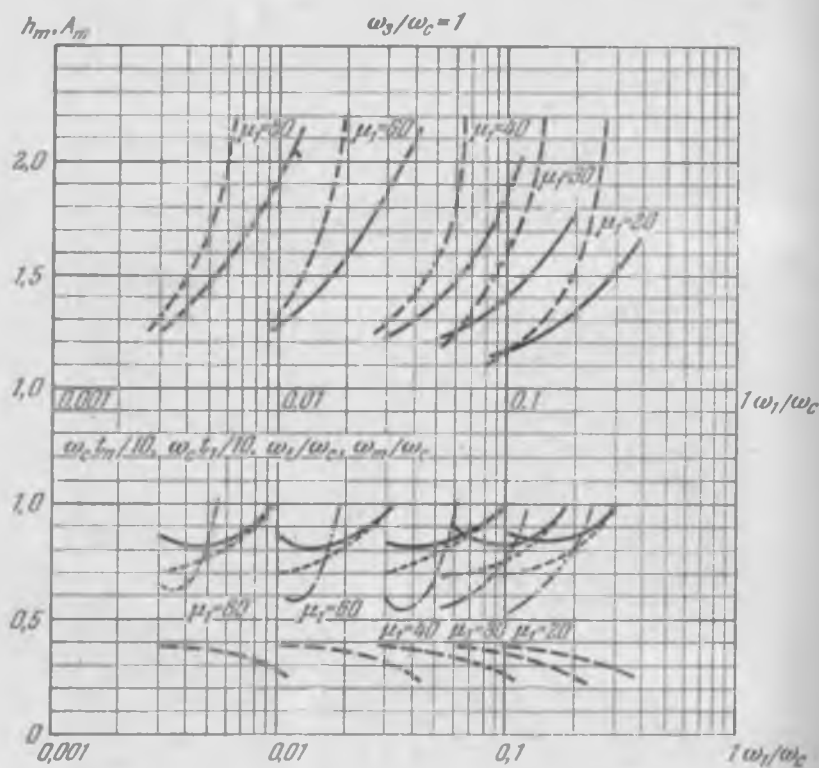


Рис. П.2.

Верхняя номограмма

h_m —————

A_m - - - - -

Нижняя номограмма

$\omega_c t_p/10$ - . - . - .

$\omega_c t_l/10$ — — — — —

ω_l/ω_c — — — — —

ω_m/ω_c - - - - -

Продолжение

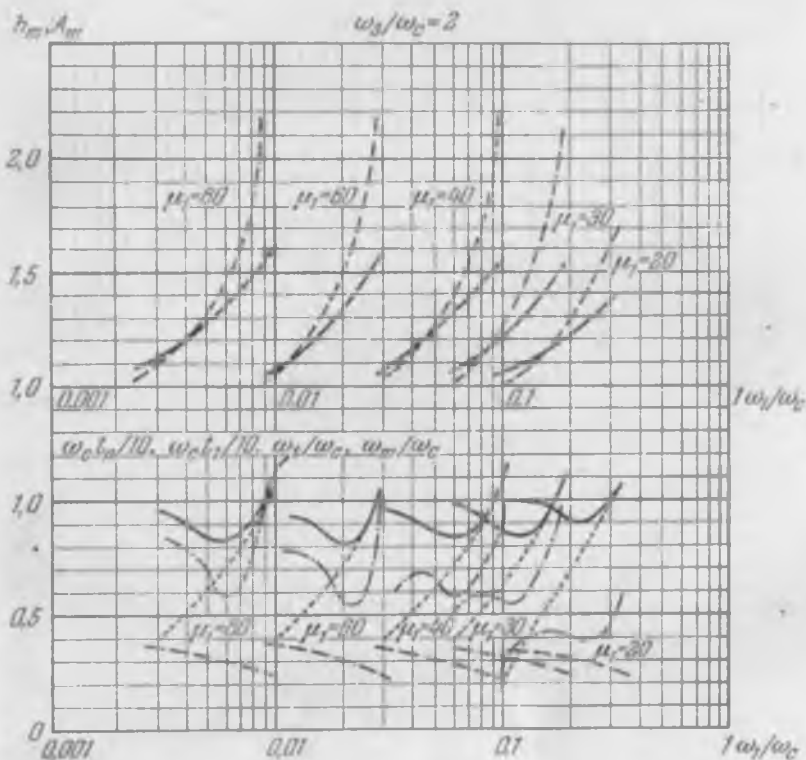


Рис. П.3.

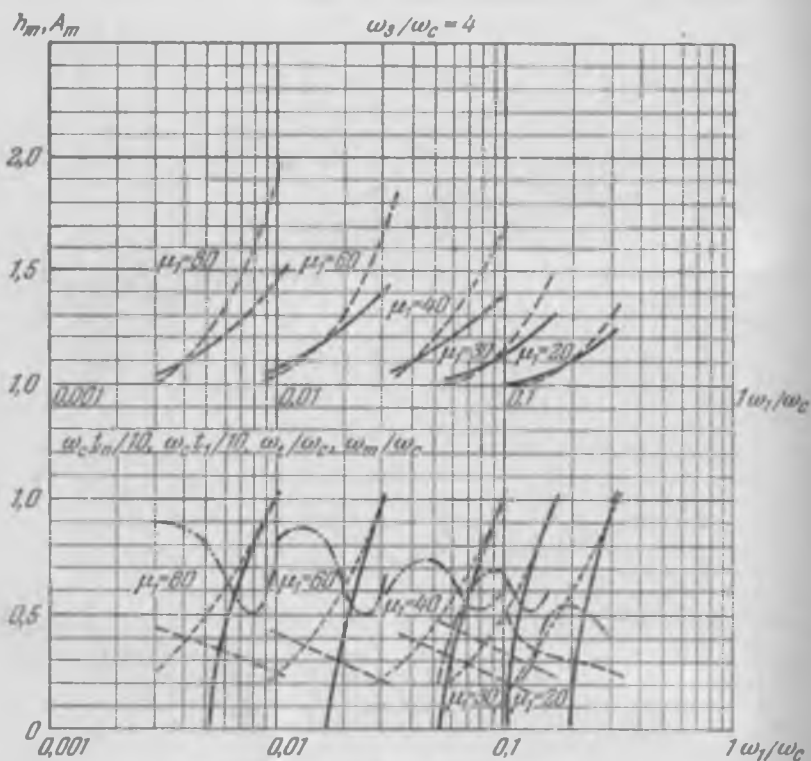
Верхняя номограмма

h_m —————
 A_m - - - - -

Нижняя номограмма

$\omega_c t_n/10$ — · — · —
 $\omega_c t_1/10$ - - - - -
 ω_l/ω_c —————
 ω_m/ω_c - - - - -

Продолжение



Верхняя номограмма

h_m —————

A_m - - - - -

Нижняя номограмма

$\omega_c t_n/10$ - . - . - .

$\omega_c t_1/10$ - - - - -

ω_t/ω_c —————

ω_m/ω_c - - - - -

Продолжение



Рис. П.5.

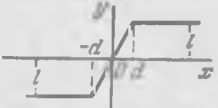
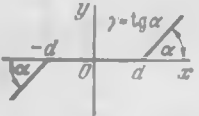
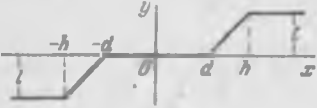
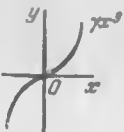
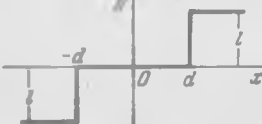
Верхняя номограмма

h_m —————
 A_m - - - - -

Нижняя номограмма

$\omega_c t_\pi/10$ —
 $\omega_c t_1/10$ — - - - -
 ω_1/ω_c —————
 ω_m/ω_c - - - - -

4. Типовые характеристики нелинейных звеньев

№	Вид нелинейности	$y = \varphi(x)$
1		$y = \frac{l}{d} x, \quad x \leq d$ $y = l, \quad x > d$ $y = -l, \quad x < -d$
2		$y = \gamma(x+d), \quad x < -d$ $y = 0, \quad x \leq d$ $y = \gamma(x-d), \quad x > d$
3		$y = -l, \quad x < -h$ $y = \frac{l}{h-d}(x+d), \quad -h \leq x < -d$ $y = 0, \quad x \leq d$ $y = \frac{l}{h-d}(x-d), \quad d \leq x < h$ $y = l, \quad x > h$
4		$y = \gamma x^3$
5		$y = -l, \quad x < 0$ $y = l, \quad x > 0$
6		$y = -l, \quad x < -d$ $y = 0, \quad x \leq d$ $y = l, \quad x > d$

№	Вид нелинейности	$y = \varphi(x)$
7		$\left. \begin{aligned} y &= l, & x &> h \\ y &= 0, & -d &\leq x \leq d \\ y &= -l, & x &< -d \end{aligned} \right\} \dot{x} > 0$ $\left. \begin{aligned} y &= l, & x &> d \\ y &= 0, & -h &\leq x \leq d \\ y &= -l, & x &< -h \end{aligned} \right\} \dot{x} < 0$
8		$\left. \begin{aligned} y &= l, & x &> d \\ y &= -l, & x &< d \end{aligned} \right\} \dot{x} > 0$ $\left. \begin{aligned} y &= l, & x &> -d \\ y &= -l, & x &< -d \end{aligned} \right\} \dot{x} < 0$
9		$y = \gamma x^2$
10		$y = 0, \quad x \leq 0$ $y = \gamma x, \quad x > 0$
11		$y = 0, \quad x \leq 0$ $y = \gamma \epsilon^2, \quad x > 0$
12		$y = 0, \quad x \leq 0$ $y = l, \quad x > 0$
13		$y = \gamma(x-d), \quad \dot{x} > 0$ $y = \gamma(x+d), \quad \dot{x} < 0$

5. Формулы для статистических и гармонических коэффициентов усиления нелинейных звеньев

1) Статистическая линеаризация

- Φ_0 — статистическая характеристика нелинейного звена,
 k_0 — статистический коэффициент усиления математического ожидания,
 $k_1^{(1)}$ — статистический коэффициент усиления флуктуаций, вычисленный по первому способу,
 $k_1^{(2)}$ — статистический коэффициент усиления флуктуаций, вычисленный по второму способу.

2) Гармоническая линеаризация

- Φ^* — средняя характеристика нелинейного звена,
 q — гармонический коэффициент усиления (первый),
 q' — гармонический коэффициент усиления (второй).

3) Совместная статистическая и гармоническая линеаризация

- Φ_0^* — средняя статистическая характеристика нелинейного звена,
 k — статистический гармонический коэффициент усиления синусоидального сигнала,
 k_1 — статистический гармонический коэффициент усиления флуктуаций.

4) Формулы для статистических, гармонических и статистических гармонических коэффициентов усиления (при постоянной составляющей входного сигнала, равной нулю: $x_0 = 0$).

В формулах введены обозначения:

$$m_1 = \frac{m_x}{d}, \quad \sigma_1 = \frac{\sigma_x}{d}, \quad v = \frac{d}{h},$$

$$x_1 = \frac{d}{\sqrt{2}\sigma_x}, \quad x_2 = \frac{h}{\sqrt{2}\sigma_x}, \quad \lambda = \frac{a}{\sqrt{2}\sigma_x},$$

$$\Phi(x) = \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt.$$

Нелинейность 1*)

$$k_0 = \frac{1}{m_x} \left\{ (1+m_1) \Phi\left(\frac{1+m_1}{\sigma_1}\right) - (1-m_1) \Phi\left(\frac{1-m_1}{\sigma_1}\right) + \right. \\ \left. + \frac{\sigma_1}{\sqrt{2\pi}} \left[e^{-\frac{1}{2}\left(\frac{1+m_1}{\sigma_1}\right)^2} - e^{-\frac{1}{2}\left(\frac{1-m_1}{\sigma_1}\right)^2} \right] \right\},$$

$$k_1^{(1)} = \frac{1}{\sigma_x} \left\{ 1 - \frac{k_0^2 m_x^2}{j^2} + (m_1^2 + \sigma_1^2 - 1) \left[\Phi\left(\frac{1+m_1}{\sigma_1}\right) + \Phi\left(\frac{1-m_1}{\sigma_1}\right) \right] - \right. \\ \left. - \frac{\sigma_1}{\sqrt{2\pi}} \left[(1-m_1) e^{-\frac{1}{2}\left(\frac{1+m_1}{\sigma_1}\right)^2} + (1+m_1) e^{-\frac{1}{2}\left(\frac{1-m_1}{\sigma_1}\right)^2} \right] \right\}^{1/2},$$

$$k_1^{(2)} = \frac{1}{\sigma_x} \sigma_1 \left[\Phi\left(\frac{1+m_1}{\sigma_1}\right) + \Phi\left(\frac{1-m_1}{\sigma_1}\right) \right],$$

*) Нелинейности пронумерованы в соответствии с приложением 4.

$$q = \frac{2l}{d\pi} \left[\arcsin \frac{d}{a} + \frac{d}{a} \sqrt{1 - \frac{d^2}{a^2}} \right] \text{ при } a > d, \quad q' = 0,$$

$$\kappa = \frac{l}{a} \sum_{n=0}^{\infty} \frac{x_1^{2n}}{2n+1} B_n(\lambda), \quad \kappa_1 = \frac{l}{\sigma_x} \sum_{n=0}^{\infty} \frac{x_1^{2n}}{2n+1} C_n(\lambda).$$

Нелинейность 2

$$k_0 = \gamma \left\{ 1 - \frac{1}{m_1} \left[(1+m_1) \Phi \left(\frac{1+m_1}{\sigma_1} \right) - (1-m_1) \Phi \left(\frac{1-m_1}{\sigma_1} \right) \right] + \right. \\ \left. + \frac{\sigma_1}{m_1 \sqrt{2\pi}} \left[e^{-\frac{1}{2} \left(\frac{1-m_1}{\sigma_1} \right)^2} - e^{-\frac{1}{2} \left(\frac{1+m_1}{\sigma_1} \right)^2} \right] \right\},$$

$$k_1^{(1)} = \gamma \left\{ 1 + \frac{1+m_1^2}{\sigma_1^2} - \frac{k_0^2 m_1^2}{\gamma^2 \sigma_1^2} - \left[1 + \left(\frac{1+m_1}{\sigma_1} \right)^2 \right] \Phi \left(\frac{1+m_1}{\sigma_1} \right) - \right. \\ \left. - \left[1 + \left(\frac{1-m_1}{\sigma_1} \right)^2 \right] \Phi \left(\frac{1-m_1}{\sigma_1} \right) - \frac{1}{\sigma_1 \sqrt{2\pi}} \left[(1+m_1) e^{-\frac{1}{2} \left(\frac{1+m_1}{\sigma_1} \right)^2} + \right. \right. \\ \left. \left. + (1-m_1) e^{-\frac{1}{2} \left(\frac{1-m_1}{\sigma_1} \right)^2} \right] \right\}^{1/2},$$

$$k_1^{(2)} = \gamma \left[1 - \Phi \left(\frac{1+m_1}{\sigma_1} \right) - \Phi \left(\frac{1-m_1}{\sigma_1} \right) \right],$$

$$q = \gamma - \frac{2\gamma}{\pi} \left[\arcsin \frac{d}{a} + \frac{d}{a} \sqrt{1 - \frac{d^2}{a^2}} \right], \quad a > d, \quad q' = 0,$$

$$\kappa = \gamma \left[1 - \frac{d}{a} \sum_{n=0}^{\infty} \frac{x_1^{2n+1}}{2n+1} B_n(\lambda) \right], \quad \kappa_1 = \gamma \left[1 - \frac{d}{\sigma_x} \sum_{n=0}^{\infty} \frac{x_1^{2n+1}}{2n+1} C_n(\lambda) \right].$$

Нелинейность 3

$$k_0 = \frac{l}{m_x} \frac{1}{1-\nu} \left\{ (1+m_1) \Phi \left(\frac{1+m_1}{\sigma_1} \right) - (1-m_1) \Phi \left(\frac{1-m_1}{\sigma_1} \right) - \right. \\ \left. - (m_1+\nu) \Phi \left(\frac{m_1+\nu}{\sigma_1} \right) + (m_1-\nu) \Phi \left(\frac{m_1-\nu}{\sigma_1} \right) + \right. \\ \left. + \frac{\sigma_1}{\sqrt{2\pi}} \left[e^{-\frac{1}{2} \left(\frac{1-m_1}{\sigma_1} \right)^2} - e^{-\frac{1}{2} \left(\frac{1+m_1}{\sigma_1} \right)^2} - e^{-\frac{1}{2} \left(\frac{m_1+\nu}{\sigma_1} \right)^2} + e^{-\frac{1}{2} \left(\frac{m_1-\nu}{\sigma_1} \right)^2} \right] \right\},$$

$$k_1^{(1)} = \frac{l}{\sigma_x} \left\{ 1 - \frac{k_0^2 m_x^2}{l^2} - \frac{(1+m_1)(1-m_1-2\nu) - \sigma_1^2}{(1-\nu)^3} \Phi \left(\frac{1+m_1}{\sigma_1} \right) - \right. \\ \left. - \frac{(1-m_1)(1+m_1-2\nu) - \sigma_1^2}{(1-\nu)^2} \Phi \left(\frac{1-m_1}{\sigma_1} \right) - \frac{(m_1+\nu)^2 + \sigma_1^2}{(1-\nu)^2} \Phi \left(\frac{m_1+\nu}{\sigma_1} \right) + \right. \\ \left. + \frac{(m_1-\nu)^2 + \sigma_1^2}{(1-\nu)^2} \Phi \left(\frac{m_1-\nu}{\sigma_1} \right) - \frac{\sigma_1}{(1-\nu)^2 \sqrt{2\pi}} \left[(1+m_1-2\nu) e^{-\frac{1}{2} \left(\frac{1-m_1}{\sigma_1} \right)^2} + \right. \right. \\ \left. \left. + (1-m_1-2\nu) e^{-\frac{1}{2} \left(\frac{1+m_1}{\sigma_1} \right)^2} + (m_1+\nu) e^{-\frac{1}{2} \left(\frac{m_1+\nu}{\sigma_1} \right)^2} - (m_1-\nu) e^{-\frac{1}{2} \left(\frac{m_1-\nu}{\sigma_1} \right)^2} \right] \right\}^{1/2},$$

$$+ (1 - m_1 - 2\nu) e^{-\frac{1}{2} \left(\frac{1+m_1}{\sigma_1} \right)^2} + (m_1 + \nu) e^{-\frac{1}{2} \left(\frac{m_1+\nu}{\sigma_1} \right)^2} - \\ - (m_1 - \nu) e^{-\frac{1}{2} \left(\frac{m_1-\nu}{\sigma_1} \right)^2} \Big|^{1/2}.$$

$$k_1^{(1)} = \frac{l}{\sigma_x} \frac{\sigma_1}{1-\nu} \left[\Phi \left(\frac{1+m_1}{\sigma_1} \right) + \Phi \left(\frac{1-m_1}{\sigma_1} \right) - \Phi \left(\frac{m_1+\nu}{\sigma_1} \right) + \Phi \left(\frac{m_1-\nu}{\sigma_1} \right) \right],$$

$$q = \frac{2l}{(h-d)\pi} \left(\arcsin \frac{h}{a} - \arcsin \frac{d}{a} - \frac{d}{a} \sqrt{1 - \frac{d^2}{a^2}} + \frac{h}{a} \sqrt{1 - \frac{h^2}{a^2}} \right), \quad a \geq h,$$

$$q' = 0.$$

Нелинейность 4

$$k_0 = \gamma \sigma_x^2 \left[3 + \left(\frac{m_x}{\sigma_x} \right)^2 \right], \quad k_1^{(1)} = \gamma \sigma_x^3 \left\{ 15 + 36 \left(\frac{m_x}{\sigma_x} \right)^2 + 9 \left(\frac{m_x}{\sigma_x} \right)^4 \right\}^{1/2},$$

$$k_1^{(2)} = 3\gamma \sigma_x^3 \left[1 + \left(\frac{m_x}{\sigma_x} \right)^2 \right],$$

$$q = \frac{3\gamma a^3}{4}, \quad q' = 0,$$

$$\kappa = \frac{3}{4} \gamma (a^2 + 4\sigma_x^2), \quad \kappa_1 = \frac{3}{2} \gamma (a^2 + 2\sigma_x^2).$$

Нелинейность 5

$$k_0 = \frac{2l}{m_x} \Phi \left(\frac{m_x}{\sigma_x} \right), \quad k_1^{(1)} = \frac{l}{\sigma_x} \left[1 - 4\Phi^2 \left(\frac{m_x}{\sigma_x} \right) \right]^{1/2},$$

$$k_1^{(2)} = \frac{l}{\sigma_x} \frac{2}{\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{m_x}{\sigma_x} \right)^2},$$

$$q = \frac{4l}{\pi a}, \quad q' = 0, \quad \kappa = \frac{l}{a} B_0(\lambda), \quad \kappa_1 = \frac{l}{\sigma_x} C_0(\lambda).$$

Нелинейность 6

$$k_0 = \frac{l}{m_x} \left[\Phi \left(\frac{1+m_1}{\sigma_1} \right) - \Phi \left(\frac{1-m_1}{\sigma_1} \right) \right],$$

$$k_1^{(1)} = \frac{l}{\sigma_x} \left\{ 1 - \frac{k_0^2 m_x^3}{l^2} - \Phi \left(\frac{1+m_1}{\sigma_1} \right) - \Phi \left(\frac{1-m_1}{\sigma_1} \right) \right\}^{1/2},$$

$$k_1^{(2)} = \frac{l}{\sigma_x} \frac{1}{\sqrt{2\pi}} \left[e^{-\frac{1}{2} \left(\frac{1+m_1}{\sigma_1} \right)^2} + e^{-\frac{1}{2} \left(\frac{1-m_1}{\sigma_1} \right)^2} \right],$$

$$q = \frac{4l}{\pi a} \sqrt{1 - \frac{d^2}{a^2}}, \quad a \geq d, \quad q' = 0,$$

$$\kappa = \frac{l}{a} \sum_{n=0}^{\infty} x_1^{2n} B_n(\lambda), \quad \kappa_1 = \frac{l}{\sigma_x} \sum_{n=0}^{\infty} x_1^{2n} C_n(\lambda).$$

Нелинейность 7

$$\begin{aligned}
 k_0 &= \frac{l}{m_x} \frac{1}{2} \left[\Phi \left(\frac{1+m_1}{\sigma_1} \right) - \Phi \left(\frac{1-m_1}{\sigma_1} \right) + \Phi \left(\frac{v+m_1}{\sigma_1} \right) - \Phi \left(\frac{v-m_1}{\sigma_1} \right) \right], \\
 k_1^{(1)} &= \frac{l}{\sigma_x} \left\{ 1 - \frac{k_0^2 m_x^2}{l^2} - \frac{1}{2} \left[\Phi \left(\frac{1+m_1}{\sigma_1} \right) + \Phi \left(\frac{1-m_1}{\sigma_1} \right) + \right. \right. \\
 &\quad \left. \left. + \Phi \left(\frac{v+m_1}{\sigma_1} \right) + \Phi \left(\frac{v-m_1}{\sigma_1} \right) \right] \right\}^{1/2}, \\
 k_1^{(2)} &= \frac{l}{\sigma_x} \frac{1}{2\sqrt{2\pi}} \left[e^{-\frac{1}{2} \left(\frac{1+m_1}{\sigma_1} \right)^2} + e^{-\frac{1}{2} \left(\frac{1-m_1}{\sigma_1} \right)^2} + e^{-\frac{1}{2} \left(\frac{v+m_1}{\sigma_1} \right)^2} + \right. \\
 &\quad \left. + e^{-\frac{1}{2} \left(\frac{v-m_1}{\sigma_1} \right)^2} \right],
 \end{aligned}$$

$$q = \frac{2l}{\pi a} \left[\sqrt{1 - \frac{d^2}{a^2}} + \sqrt{1 - \frac{h^2}{a^2}} \right], \quad a > h,$$

$$q' = -\frac{2ld}{\pi a} \left(1 - \frac{d}{h} \right), \quad a > h,$$

$$\kappa = \frac{l}{a} \sum_{n=0}^{\infty} \frac{x_1^{2n} + x_2^{2n}}{2} B_n(\lambda), \quad \kappa_1 = \frac{l}{\sigma_x} \sum_{n=0}^{\infty} \frac{x_1^{2n} + x_2^{2n}}{2} C_n(\lambda).$$

Нелинейность 8

$$\begin{aligned}
 k_0 &= \frac{l}{m_x} \left[\Phi \left(\frac{1+m_1}{\sigma_1} \right) - \Phi \left(\frac{1-m_1}{\sigma_1} \right) \right], \quad k_1^{(1)} = \frac{l}{\sigma_x} \left(1 - \frac{k_0^2 m_x^2}{l^2} \right)^{1/2}, \\
 k_1^{(2)} &= \frac{l}{\sigma_x} \frac{1}{\sqrt{2\pi}} \left[e^{-\frac{1}{2} \left(\frac{1+m_1}{\sigma_1} \right)^2} + e^{-\frac{1}{2} \left(\frac{1-m_1}{\sigma_1} \right)^2} \right], \\
 q &= \frac{4l}{\pi a} \sqrt{1 - \frac{d^2}{a^2}}, \quad q' = -\frac{4ld}{\pi a^2}, \quad a > d,
 \end{aligned}$$

$$\kappa = \frac{l}{a} \sum_{n=0}^{\infty} x_1^{2n} B_n(\lambda), \quad \kappa_1 = \frac{l}{\sigma_x} \sum_{n=0}^{\infty} x_1^{2n} C_n(\lambda).$$

Нелинейность 9

$$\varphi_0 = \gamma \sigma_x^2 \left(1 + \frac{m_x^2}{\sigma_x^2} \right),$$

$$k_1^{(1)} = \gamma \sigma_x \left[2 + 4 \frac{m_x^2}{\sigma_x^2} \right]^{1/2} \cdot \operatorname{sgn} m_x, \quad k_1^{(2)} = 2\gamma m_x,$$

$$\varphi^* = \frac{\gamma a^2}{2}, \quad q = q' = 0.$$

Нелинейность 10

$$\varphi_0 = \gamma \sigma_x \left\{ \frac{m_x}{\sigma_x} \left[\frac{1}{2} + \Phi \left(\frac{m_x}{\sigma_x} \right) \right] + \frac{1}{\sqrt{2\pi}} e^{-\frac{m_x^2}{2\sigma_x^2}} \right\},$$

$$k_1^{(1)} = \gamma \left\{ \left[\frac{1}{2} + \Phi \left(\frac{m_x}{\sigma_x} \right) \right] \left(1 + \frac{m_x^2}{\sigma_x^2} \right) + \frac{m_x}{\sigma_x \sqrt{2\pi}} e^{-\frac{m_x^2}{2\sigma_x^2}} - \frac{\varphi_0^2}{4\sigma_x^2} \right\}^{1/2},$$

$$k_1^{(2)} = \gamma \left[\frac{1}{2} + \Phi \left(\frac{m_x}{\sigma_x} \right) \right], \quad \varphi^* = \frac{\gamma a}{\pi}, \quad q = \frac{\gamma}{2}, \quad q' = 0.$$

Нелинейность 11

$$\varphi_0 = \gamma \sigma_x^2 \left\{ \left(1 + \frac{m_x^2}{\sigma_x^2} \right) \left[\frac{1}{2} + \Phi \left(\frac{m_x}{\sigma_x} \right) \right] + \frac{m_x}{\sigma_x \sqrt{2\pi}} e^{-\frac{m_x^2}{2\sigma_x^2}} \right\},$$

$$k_1^{(1)} = \gamma \sigma_x \left\{ \left(3 + \frac{m_x^2}{\sigma_x^2} \right) \left[\frac{1}{2} + \Phi \left(\frac{m_x}{\sigma_x} \right) \right] - \frac{\varphi_0^2}{\sigma_x^2 \gamma^2} + \right.$$

$$\left. + \frac{1}{\sqrt{2\pi}} e^{-\frac{m_x^2}{2\sigma_x^2}} \left(5 \frac{m_x}{\sigma_x} + \frac{m_x^2}{\sigma_x^2} \right) \right\}^{1/2},$$

$$k_1^{(2)} = \gamma \sigma_x \left\{ \left[1 + 2\Phi \left(\frac{m_x}{\sigma_x} \right) \right] \frac{m_x}{\sigma_x} + \frac{2}{\sqrt{2\pi}} e^{-\frac{m_x^2}{2\sigma_x^2}} \right\},$$

$$\varphi^* = \frac{\gamma a^2}{4\pi}, \quad q = \frac{4}{3} \frac{\gamma a}{\pi}, \quad q' = 0.$$

Нелинейность 12

$$\varphi_0 = l \left[\frac{1}{2} + \Phi \left(\frac{m_x}{\sigma_x} \right) \right],$$

$$k_1^{(1)} = \frac{l}{\sigma_x} \left[\frac{1}{4} - \Phi^2 \left(\frac{m_x}{\sigma_x} \right) \right]^{1/2}, \quad k_1^{(2)} = \frac{l}{\sigma_x \sqrt{2\pi}} e^{-\frac{m_x^2}{2\sigma_x^2}},$$

$$\varphi^* = \frac{l}{2}, \quad q = \frac{2l}{\pi a}, \quad q' = 0.$$

Нелинейность 13

$$q = \frac{\gamma}{\pi} \left[\frac{\pi}{2} + \arcsin \left(1 - \frac{2d}{a} \right) + 2 \left(1 - \frac{2d}{a} \right) \sqrt{\frac{d}{a} \left(1 - \frac{d}{a} \right)} \right], \quad a \geq d,$$

$$q' = -\frac{4\gamma d}{\pi a} \left(1 - \frac{d}{a} \right), \quad a \geq d.$$

Входящие в приведенные формулы функции $B_n(\lambda)$ и $C_n(\lambda)$ определяются формулами

$$B_n(\lambda) = \frac{4}{\sqrt{\pi}} \sum_{k=0}^{\infty} \frac{(-1)^{k+n} [2(k+n)]!}{(2n)! (k!)^2 (k+n)! (k+1)!} \left(\frac{\lambda}{2}\right)^{2k+1},$$

$$C_n(\lambda) = \sqrt{\frac{2}{\pi}} \sum_{k=0}^{\infty} \frac{(-1)^{k+n} [2(k+n)]!}{(2n)! (k!)^2 (k+n)!} \left(\frac{\lambda}{2}\right)^{2k} \quad (n=0, 1, 2, \dots).$$

При вычислении коэффициентов k и k_1 достаточно ограничиться первыми

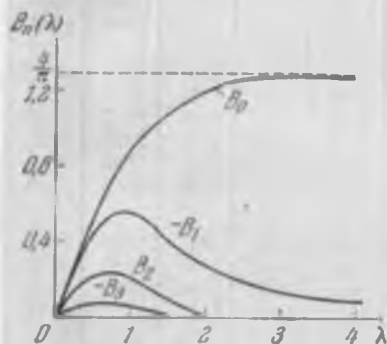


Рис. П.6.

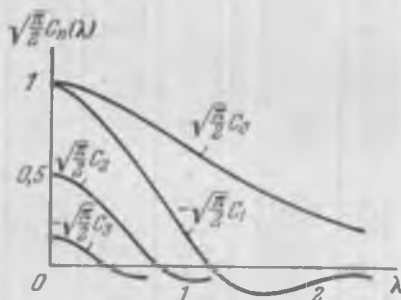


Рис. П.7.

четырьмя членами соответствующих рядов для $B_n(\lambda)$ и $C_n(\lambda)$. Графики этих функций приводятся на рис П.6 и П.7.

6. Значения функции $\Phi(u) = \frac{1}{\sqrt{2\pi}} \int_0^u e^{-\frac{t^2}{2}} dt$

[illegible]

7. Значения функции $\Phi'(u) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}}$

ЛИТЕРАТУРА

1. Александров В. М., Батков А. М., Староверов А. Н., Шукин Б. А., Исследование точности нелинейных нестационарных систем при помощи метода статистической линеаризации. Автоматика и телемеханика, т. XXVI, № 3 (1965), стр. 492—499.
2. Андреев И. И., Определение оптимальной линейной динамической системы по критерию экстремума функционала частного вида. Автоматика и телемеханика, т. XVIII, № 7 (1957), стр. 615—619.
3. Андреев И. И., Общее условие экстремума заданной функции средне-квадратической ошибки и квадрата математического ожидания ошибки динамической системы. Автоматика и телемеханика, т. XX, № 7 (1959), стр. 833—838.
4. Андронов А. А., Витт А. А., Понтрягин Л. С., О статистическом рассмотрении динамических систем. Журнал экспериментальной и теоретической физики, т. 3, вып. 3 (1933), стр. 165—180. (См. также Собрание трудов А. А. Андропова, Изд. АН СССР, 1956, стр. 142—160.)
5. Андронов А. А., Витт А. А., Хайкин С. Э., Теория колебаний. Физматгиз, 1959.
6. Артоболевский И. И., Герц Е. В., Кобринский А. Е., Раевский И. П., К динамике пневматических устройств. Труды семинара по теории машин и механизмов. Изд. АН СССР, т. XIII, вып. 51 (1953), стр. 60—85.
7. Баранов Г. Г., Курс теории механизмов и машин. Машгиз, 1959.
8. Боде Г. В., Теория цепей и проектирование усилителей с обратной связью. ИЛ, 1949.
9. Бусленко Н. П., Шрейдер Ю. А., Метод статистических испытаний. Физматгиз, 1961.
10. Бутон Р. (Booton R. C.), Nonlinear control systems with random inputs. Trans. IRE, vol. CT-1 (1954), pp. 9—18.
11. Вентцель Е. С., Теория вероятностей. Физматгиз, 1962.
12. Винер Н. (Winer N.), Extrapolation, interpolation and smoothing of stationary time series. John Wiley, N.Y., 1949.
13. Власов В. Ф., Курс радиотехники. Госэнергоиздат, 1962.
14. Габлер М., Гашковец И., Томонек Е., Магнитные усилители. «Сов. радио», 1961.
15. Гольдфарб Л. С., О некоторых нелинейностях в системах регулирования. Автоматика и телемеханика, т. VIII, № 5 (1947), стр. 349—383.
16. Гурвиц А. (Hurwitz A.), Über die Bedingungen, unter welchen eine Gleichung nur Wurzeln mit negativen reellen Teilen besitzen. Math. Ann., Bd. 46 (1895), pp. 273—284.
17. Дёч Г., Руководство к практическому применению преобразования Лапласа. Физматгиз, 1958.
18. Джеме Х. М., Никольс Н. Б., Филиппс Р. С. (ред.), Теория следящих систем. ИЛ, 1951.
19. Доброзурский С. О., Казаков В. А., Титов В. К., Счетно-решающие устройства. Оборонгиз, 1959.
20. Доступов В. Г., Приближенное определение вероятностных характеристик решений систем обыкновенных дифференциальных уравнений,

- содержащих случайные параметры. Труды ВВИА им. Н. Е. Жуковского, вып. 592 (1956), стр. 32—46.
21. *Доступов Б. Г.*, Приближенное определение вероятностных характеристик выходных координат нелинейных систем автоматического регулирования. Автоматика и телемеханика, т. XVIII, № 11 (1957), стр. 999—1009.
 22. *Евдокимов Л. Г.*, Самонастраивающаяся система с поиском градиента методом вспомогательного оператора. Изв. АН СССР, ОТН, Техническая кибернетика, № 1 (1963), стр. 13—20.
 23. *Жузе В. П.*, Применение эффекта Холла в полупроводниках. Труды конференции по применению полупроводников в приборостроении. Машгиз (1958), стр. 135—158.
 24. *Жуковский Н. Е.*, О прочности движения. Ученые записки Московского университета, вып. 4 (1882). См. также Полное собрание сочинений, т. 1, 1937, стр. 110—208.
 25. *Zadeh L. A., Ragazzini J. R.*, Extension of Wiener's theory of prediction. Journ. Appl. Physics, vol. 21, № 7 (1950), pp. 645—655.
 26. *Казаков И. Е.*, Приближенный метод статистического исследования нелинейных систем. Труды ВВИА им. Н. Е. Жуковского, вып. 394 (1954).
 27. *Казаков И. Е.*, Некоторые вопросы теории статистической линеаризации и ее приложений. Труды I Международного конгресса по автоматике. Изд. АН СССР, т. 3 (1961), стр. 68—82.
 28. *Казаков И. Е., Доступов Б. Г.*, Статистическая динамика нелинейных автоматических систем. Физматгиз, 1962.
 29. *Казаков И. Е.*, К статистической теории непрерывных самонастраивающихся систем. Изв. АН СССР, ОТН, Энергетика и автоматика, № 6 (1962), стр. 76—81.
 30. *Касаткин А. С., Перекалин М. А., Сергеев П. С.*, Электротехника. Госэнергоиздат, 1953.
 31. *Корректирующие цепи в автоматике.* ИЛ, 1954.
 32. *Котельников В. А.*, О пропускной способности «эфира» и проволоки в электросвязи. Материалы к I Всесоюзному съезду по вопросам реконструкции дела связи. Изд. Управления связи РККА, 1933.
 33. *Котельников В. А.*, Теория потенциальной помехоустойчивости. Госэнергоиздат, 1956.
 34. *Красовский А. А., Поспелов Г. С.*, Основы автоматки и технической кибернетики. Машгиз, 1962.
 35. *Крылов Н. М., Боголюбов Н. Н.*, Исследование продольной устойчивости аэроплана. Гос. авиац. и автотракторное издательство, М.—Л., 1932.
 36. *Крылов Н. М., Боголюбов Н. Н.*, Новые методы нелинейной механики. ОНТИ, 1934.
 37. *Крылов Н. М., Боголюбов Н. Н.*, Введение в нелинейную механику, Киев, 1937.
 38. *Кухтенко В. И.*, К расчету корректирующих цепей систем автоматического управления по критерию минимума среднеквадратической ошибки. Автоматика и телемеханика, т. XX, № 9 (1959), стр. 1180—1187.
 39. *Летов А. М.*, Устойчивость нелинейных регулируемых систем. Физматгиз, 1962.
 40. *Литвин-Седой М. З.*, Гидравлический двигатель в системах автоматки. Машгиз, 1956.
 41. *Лурье А. И.*, Некоторые нелинейные задачи теории автоматического регулирования. Гостехиздат, 1951.
 42. *Ляпунов А. М.*, Общая задача об устойчивости движения. ОНТИ, 1935.
 43. *Малкин И. Г.*, Теория устойчивости движения, Гостехиздат, 1952.
 44. *Марголис М., Леондес К. Т.*, О теории самонастраивающейся системы регулирования; метод обучающейся модели. Труды I Международного конгресса по автоматике. Изд. АН СССР, т. 2 (1961), стр. 683—699.

45. Нейквист Х. (Nyquist H.), Regeneration theory. Bell System Techn. Journ., vol. 11 (1932), pp. 126—147.
46. Ньютон Д. К., Гулд Л. А., Кайзер Д. Ф., Теория линейных следящих систем. Физматгиз, 1961.
47. Основы автоматического регулирования. Машгиз, т. 1, 1954; ч. 1 и ч. 2, 1959.
48. Понтрягин Л. С., Болтянский В. Г., Гамкрелидзе Р. В., Мищенко Е. Ф., Математическая теория оптимальных процессов. Физматгиз, 1961.
49. Попов Е. П., Пальтов И. П., Приближенные методы исследования нелинейных автоматических систем. Физматгиз, 1960.
50. Привалов И. И., Ряды Фурье, ГТТИ, 1934.
51. Пуанкаре А., О кривых, определяемых дифференциальными уравнениями. Гостехиздат, 1947.
52. Пугачев В. С., Случайные функции, определяемые обыкновенными дифференциальными уравнениями. Труды ВВИА им. Н. Е. Жуковского, № 118 (1944).
53. Пугачев В. С., Теория случайных функций и ее применение к задачам автоматического управления. Физматгиз, изд. 3-е, 1962.
54. Пугачев В. С., Введение в теорию вероятностей. «Наука», 1967.
55. Пугачев В. С., Доступов Б. Г., Казаков И. Е., Статистическая теория нелинейных систем (лекции). Изд. ВВИА им. Н. Е. Жуковского, вып. 1, 1961, вып. 2, 3, 1962.
56. Пугачев В. С., Применение теории марковских процессов к анализу точности автоматических систем. Изв. АН СССР, ОТН, Энергетика и автоматика, № 3 (1961), стр. 46—57.
57. Пугачев В. С., Оптимальное обучение автоматических систем в изменяющихся условиях. Автоматика и телемеханика, т. XXVIII, № 10 (1967), стр. 153—167.
58. Растрюгин Л. А., Статистические методы поиска. «Наука», 1968.
59. Раус Э. Дж. (Routh E. J.), A treatise on the stability of a given state of motion. MacMillan, London, 1877.
60. Розенблат М. А., Бесконтактные магнитные устройства автоматики. Изд. АН СССР, 1961.
61. Себестиян Г. С., Процессы принятия решений при распознавании образов. Изд. «Техника», Киев, 1965.
62. Семенов В. М., К теории экстраполирования случайных процессов. Сборник научных трудов ВВИА им. Н. Е. Жуковского, т. I (1954), стр. 67—94.
63. Смирнов В. И., Курс высшей математики. Гостехиздат (Физматгиз), т. I, 1957; т. II, 1957; т. III, ч. 1 и 2, 1954; т. IV, 1957; т. V, 1959.
64. Соколов В. И., Применение методов структурного анализа к исследованию нестационарных линейных динамических систем. Сборник научных трудов ВВИА им. Н. Е. Жуковского, т. II (1957), стр. 53—67.
65. Соколов В. И., Линейка для облегчения построения логарифмических частотных характеристик систем автоматического регулирования. Информационный сборник НИО ВВИА им. Н. Е. Жуковского, вып. V (1957), стр. 25—26.
66. Солодов А. В., Линейные системы автоматического управления с переменными параметрами. Физматгиз, 1962.
67. Сотсков Б. С., Элементы автоматической и телемеханической аппаратуры. Гостехиздат, 1950.
68. Степанов В. В., Курс дифференциальных уравнений, изд. 8-е, Физматгиз, 1959.
69. Файнштейн А., Основы теории информации. ИЛ, 1960.
70. Фельдбаум А. А., Вычислительные устройства в автоматических системах. Физматгиз, 1959.
71. Фельдбаум А. А., О синтезе оптимальных систем с помощью фазового пространства. Автоматика и телемеханика, т. XVI, № 2 (1955), стр. 129—149.

72. Цыпкин Я. З., Теория импульсных систем. Физматгиз, 1958.
73. Цыпкин Я. З., Теория релейных систем автоматического регулирования. Гостехиздат, 1955.
74. Четаев Н. Г., Устойчивость движения. Гостехиздат, 1955.
75. Шеннон К. (Shannon C. E.), A mathematical theory of communication. Bell System Techn. Journ., vol. 27 (1948), pp. 379—423, 623—656.
76. Шляхдин В. М., Элементы автоматики и телемеханики. Оборонгиз, 1954.

Дополнение к списку литературы

77. Беллман Р., Динамическое программирование. ИЛ, 1960.
78. Беллман Р., Калаба Р., Динамическое программирование и современная теория управления. «Наука», 1969.
79. Болтянский В. Г., Гамкрелидзе Р. В., Понтрягин Л. С., К теории оптимальных процессов. Доклады АН СССР, т. 110, № 1 (1956), стр. 7—10.
80. Болтянский В. Г., Достаточные условия оптимальности и обоснование метода динамического программирования. Изв. АН СССР, сер. матем., т. 28, № 3 (1964), стр. 481—514.
81. Брайсон А., Йохансен Д. (Bryson A. E., Johansen D. E.), Linear filtering for time-varying systems using measurements containing colored noise. IEEE Trans. on Automatic Control, vol. AC-10, № 1 (1965), pp. 4—10.
82. Воронов А. А., Основы теории автоматического управления. «Энергия», ч. 1, 1965; ч. 2, 1966; ч. 3, 1970.
83. Гантмахер Ф. Р., Теория матриц. «Наука», 1967.
84. Гулько Ф. Б., Новосельцева Ж. А., Решение нестационарных задач фильтрации и упреждения методами моделирования. Автоматика и телемеханика, т. XXVII, № 4 (1966), стр. 122—141.
85. Гулько Ф. Б., Новосельцева Ж. А., Решение нестационарных задач фильтрации и упреждения при произвольной помехе методами моделирования. Автоматика и телемеханика, т. XXVII, № 10 (1966), стр. 153—168.
86. Заде Л., Дезоер Ч., Теория линейных систем. «Наука», 1970.
87. Калман Р. (Kalman R. E.), A new approach to linear filtering and prediction problems. Trans. of the ASME, Journal of Basic Engineering, vol. 82D (1960), pp. 35—45.
88. Калман Р., Бьюси Р. С. (Kalman R. E., Bucy R. S.), New results in linear filtering and prediction theory. Trans. of the ASME, Journ. of Basic Engineering, vol. 83D (1961), pp. 95—100.
89. Красовский А. А., Статистическая теория переходных процессов в системах управления. «Наука», 1968.
90. Красовский Н. Н., Теория управления движением. «Наука», 1968.
91. Красовский Н. Н., Игровые задачи о встрече движений. «Наука», 1970.
92. Летов А. М., Динамика полета и управление. «Наука», 1969.
93. Моисеев Н. Н., Численные методы в теории оптимальных систем. «Наука», 1971.
94. Тихонов А. Н., О решении некорректно поставленной задачи и методе регуляризации. Доклады АН СССР, т. 151, № 3 (1963), стр. 501—504.

ПРЕДМЕТНЫЙ УКАЗАТЕЛЬ

- Автоколебания 313, 405, 424
Азбука Морзе 473
Акселерометр 99
- Баланс гармонический 321, 322
Бел 69
«Болтанка» 259
- Варнация игольчатая 647
— переменных 411
Влияние нелинейностей 314
— случайных возмущений на авто-
колебания 444
Возмущение 250
Время отпущания 367
— переходного процесса 252
— срабатывания 367
— трогания 367
Вход системы 26
Выход системы 26
- Гамильтона функция 645
Гирискоскоп 97
Годограф 70
Граница области устойчивости 246
Гурвица критерий 230
- Давление контактное 366
Датчик 25, 86
— давления 86
— индуктивный 92
— индукционный 95
— потенциометрический 89
— трансформаторный 95
— электрический 86
— электролитический 338
Двигатель гидравлический 392
— переменного тока 129
— пневматический 397
— постоянного тока 126
Движение 250
— асимптотически устойчивое 406
— возмущенное 250, 406
— невозмущенное 250, 406
— неустойчивое 406
- Движение установившееся 411
— устойчивое 406, 409
Действие направленное 144
Декада 70
Декодирование 472
Декремент затухания 82
Демодулятор 88, 381
Демодуляция 382
Детектирование 382
Детектор 359
Децибел 70
Динамика статистическая 261
Диод вакуумный 358
— полупроводниковый 358
Дискриминатор фазовый 88, 382
Дисперсия установившаяся 284
Дифференциатор 65, 158, 492
Длительность импульсов относи-
тельная 203
Доступова метод эквивалентных воз-
мущений 451
Дрейф нуля 111
Дроссель насыщения 343
- Задача «двухточечная» 646
— некорректная 529
Закон распределения нормальный
330
Законы Кирхгофа 118
«Замораживание» коэффициентов 412
Запас устойчивости 235
— — по амплитуде 247
— — — фазе 246
Звено апериодическое 79, 692
— безынерционное элементарное 315
— дифференцирующее 51, 692
— запаздывающее 51, 65, 159, 692
— инерционное первого порядка 80
— интегрирующее 51, 692
— колебательное 80, 147, 692
— консервативное 83
— линейное стационарное 79
— неустойчивое апериодическое 692
— — колебательное 692
— резонансное 83, 692

- Звено усилительное 692
 - форсирующее 79
 - — второго порядка 692
 - — первого порядка 692
- Избыток фазы 246
- Измеритель 86
 - угловой скорости гироскопический 97
- Импульс единичный 41
- Интеграл Фурье 47, 56
- Интегратор 65, 159
- Интервал наблюдения бесконечный 542
 - — конечный 553
- Информация 19, 471
- Искатель шаговый 126
- Источник информации 23
- Исчисление операционное 69
- Канал (связи) 471
 - с шумами 483
- Качество линейных систем 226
- Квантование сигнала 476
- Кодирование 472
- Количество информации 473
- Контакт нормально замкнутый 366
 - — разомкнутый 366
- Коррелятор 679
- Котельникова теорема 478
- Коэффициент весовой 204
 - колебательности 82
 - модуляции 381
 - ошибки 279
 - — по положению 280
 - — — скорости 280
 - — — ускорению 280
 - прямоугольности 376
 - статической ошибки 280
 - усиления 81, 82
 - — гармонический 321, 702
 - — комплексный 322
 - — по постоянной составляющей 323
 - — статистический 702
 - — по математическому ожиданию 327
 - — — случайной составляющей 327
- Критерий байесов 500
 - Гурвица 230
 - «идеального наблюдателя» 632
 - качества 494, 669
 - максимального быстродействия 643
 - минимаксный 501
 - минимума вероятности ошибки 496
- Критерий минимума среднего риска 498
 - — средней квадратической ошибки 508, 627
 - Найквиста 235
 - Неймана — Пирсона 497
 - оптимальности 489
 - — статистический 495
 - Рауса 230
- Линеаризация 36, 410
 - гармоническая 319, 702
 - нелинейных характеристик 318
 - совместная гармоническая и статистическая 333
 - статистическая 325, 463, 702
- Линия длинная 131
 - задержки 132
 - неоднородная 132
 - однородная 132
 - переключения 427
 - согласованная 136
- Люфт 391
- Маркова цепь стационарная 474
- Метод Бутона статистической линеаризации 332
 - выделения сигналов корреляционный 621
 - гармонического баланса 321, 322
 - гармонической линеаризации 321, 437
 - Гулько и Новосельцевой 601
 - Дупупова эквивалентных возмущений 451
 - «замораживания» коэффициентов 412
 - интегрирования вспомогательной системы 680
 - исследования устойчивости прямой 413
 - Казакова статистической линеаризации 332
 - конечных приращений 680
 - линеаризации 410
 - логарифмических частотных характеристик 195
 - Ляпунова второй 413
 - — первый 410
 - малого параметра 436
 - Монте-Карло 467
 - наискорейшего спуска 504
 - производной по времени 680
 - синхронного детектирования 680
 - сопряженных систем 155
 - статистических испытаний 466
 - статистической линеаризации 326

- Соединение элементов 144
- Сопротивление лобовое 140
- Способность пропускная 472, 483
- Стабилизация динамических характеристик 671
- Статистика математическая 493
- Структура автоматической системы 23
- Суперпозиция 33
- Схема структурная 144, 147
- Такт дискретной системы 203
- Тахогенератор (ТГ) 96
- Тахометр 96
- Теорема Котельникова 478
 - Лежен-Дирихле 413
 - отсчетов 478
 - Шеннона 478
 - — основная 485, 486
- Теоремы Ляпунова 411, 414
- Теория игр 502
 - информации 471
 - оптимальных систем 489
 - — статистическая 489
 - систем управления статистическая 261
 - управления 18
- Термистор 343
- Термосопротивление 341
- Тиратрон 377
- Ток отпускания 367
 - срабатывания 367
- Точка изображающая 407
 - особая 416
 - равновесия 417
- Точность линейных систем 259
 - нелинейных систем 451
 - стационарных систем 456
- Траектория фазовая 407, 416
- Трансформатор вращающийся 360
 - — синусно-косинусный 360
- Трение сухое 390
- Триод полупроводниковый 113
- Угол атаки 140
 - тангажа 139
- Узел 189, 420
- Управление 17, 642
 - дискретное 201
 - импульсное 201
 - оптимальное 642
 - по замкнутому циклу 20
 - по разомкнутому циклу 20
 - при неполной информации 665
 - программное 20
- Уравнение Беллмана 661
 - в вариациях 411
- Уравнение в конечных разностях 219
 - разностное 219
 - сопряженное 168
 - теплопроводности 52
- Уравнения движения самолета 139
 - линейные первого приближения 411
 - оптимальной дискретной линейной системы 537
- Усилитель 87, 158
 - безынерционный 51, 65
 - ламповый 110
 - магнитный 343
 - — дифференциальный 353
 - — мостовой 353
 - на кристаллических триодах 113
 - операционный 122
 - с матричным коэффициентом усиления 598
 - электромашинный (ЭМУ) 114
 - электронный 110
- Условие устойчивости 226
 - физической возможности 49
- Условия несмещенности 527
- Устойчивость асимптотическая 251
 - дискретных линейных систем 247
 - линейных систем 226
 - нелинейных систем 405
 - стационарных линейных систем 230
- Устройство исполнительное 23, 87
- Учитель 667
- Фазоинвертор 88
- Фиксатор 125
- Фильтр Калмана 600
 - согласованный 634
 - формирующий 297, 470
- Флуктуация входной переменной 262
- Фокус 419
- Форма операторная 76
- Формула Бутона 552
 - Винера 544, 552
- Функционал 31
- Функция 31
 - Беллмана 660
 - весовая 48, 155, 693
 - — соединения 151
 - входная 26
 - выходная 26
 - Гамильтона 645
 - знакоопределенная 414
 - знакостоянная 413
 - импульсная единичная 41
 - инвариантная 63
 - корреляционная 295
 - — белого шума 295

Функция корреляционная взаимная

- Ляпунова 415
- определено-отрицательная 414
- определено-положительная 414
- передаточная 61, 63, 692
- — стационарной дискретной линейной системы 216
- переходная 251
- — импульсная 48
- потеря 498, 609
- ступенчатая единичная 45
- Фурье интеграл 47, 56

Характеристика амплитудно-фазовая

- 70, 322
- вольтамперная 357
- импульсного элемента 203
- механическая двигателя постоянного тока 127
- реакции на показательное возмущение 57
- реле временная 368
- — тяговая 369, 371
- статистическая нелинейного звена 327
- статическая магнитного усилителя 347
- холостого хода ЭМУ 116
- частотная 56, 81
- — амплитудная 58, 693
- — — логарифмическая 70
- — — асимптотическая 195
- — действительная 71
- — логарифмическая 69, 195, 574
- — мнимая 71
- — фазовая 58, 693
- — — логарифмическая 70
- элементарного безынерционного звена 345
- ЭМУ внешняя 116

Характеристики дискретных линейных систем 203

Характеристики линейных систем 41, 144

- нелинейных систем 313
- реле 375
- типовые нелинейных звеньев 700
- элементарных стационарных звеньев 692
- Холла поле 363
- эффект 362

Центр 421

Ценочка RC дифференцирующая 119

- — интегрирующая 120
- Цепь корректирующая 24, 671
- Маркова стационарная 474
- Цикл предельный 423

Частота повторения импульсов 203

- резонансная 68
- сопрягающая 81, 83
- среза 68

Число переграничирования 253

Шум белый 272, 539

Щетки ЭМУ 115

Экстраполятор 50, 64, 492

Электродвигатель с редуктором 390

- — фрикционной муфтой 386
- Элемент гальваномагнитный 362

- импульсный 201
- линейный 86
- — импульсный 124
- нелинейный 338
- с распределенными параметрами 131, 132
- управляющий 87
- чувствительный 25, 86
- — магнитоэлектрический 101

Энтропия 473

Эффект Холла 362

z-преобразование 217

δ -функция импульсная 41

ОСНОВЫ АВТОМАТИЧЕСКОГО УПРАВЛЕНИЯ

Под редакцией В. С. ПУГАЧЕВА

М., 1974 г., 720 стр. с илл.

Редактор Л. А. Чульский

Технический редактор В. Н. Кондакова

Корректоры Т. С. Плетнева, Л. С. Сомова

Сдано в набор 24/IX 1973 г. Подписано к печати 7/II 1974 г. Бумага 60×90^{1/16}.
Физ. печ. л. 45. Условн. печ. л. 45. Уч.-изд. л. 46,5. Тираж 16 000 экз. Т-01158.

Цена книги 3 р. 17 и. Заказ 01276

Издательство «Наука»

Главная редакция физико-математической литературы
117071, Москва, В-71, Ленинский проспект, 15

Ордена Трудового Красного знамени Московская типография № 7
«Искра революции» Союзполиграфпрома при Государственном комитете
Совета Министров СССР по делам издательства, полиграфии
и книжной торговли
Москва, К-1, Трехпрудный пер., 9