

Министерство образования Республики Беларусь
Учреждение образования
«Белорусский государственный университет информатики
и радиоэлектроники»

Кафедра систем управления

А. В. Павлова

МАТЕМАТИЧЕСКИЕ ОСНОВЫ ТЕОРИИ СИСТЕМ

Конспект лекций

по курсу «Математические основы теории систем»
для студентов специальности 1-53 01 07 «Информационные технологии
и управление в технических системах»

В 2-х частях

Часть 1

Минск

СОДЕРЖАНИЕ

ВВЕДЕНИЕ. ОБЩИЕ СРЕДСТВА МАТЕМАТИЧЕСКОГО ОПИСАНИЯ СИСТЕМ	4
ТЕМА 1. ОСНОВЫ ТЕОРИИ МНОЖЕСТВ	5
1.1. Основные понятия и определения	5
1.2. Операции над множествами	6
1.3. Законы и тождества алгебры множеств	8
1.4. Принцип двойственности	9
1.5. Уравнение с множествами	10
1.6. Упорядоченное множество. Прямое произведение множеств	11
1.7. Соответствия	12
1.8. Отображения и их виды	13
1.9. Отношения и их свойства	14
1.10. Виды отношений	15
1.11. Нечёткие множества. Способы задания. Понятие лингвистической переменной	16
1.12. Операции над нечёткими множествами	20
1.13. Параметры нечётких множеств	21
1.14. Методы дефазификации нечётких множеств	23
ТЕМА 2. ЭЛЕМЕНТЫ ТЕОРИИ ГРАФОВ и ее приложения	24
2.1. Основные понятия и определения. Способы задания графов	25
2.2. Типы графов	28
2.3. Расстояния и пути в графах. Центры и периферийные вершины	31
2.4. Числовая функция на графе. Сигнальные графы	32
2.6. Операции над графами	37
2.7. Задача о кратчайшем пути связного неориентированного графа	39
2.8. Деревья. Символ дерева	40
2.9. Покрывающее дерево связного графа. Экстремальное дерево	41
2.10. Корневые деревья. Код дерева	42
ТЕМА 3. ТРАНСПОРТНЫЕ СЕТИ	44
3.1. Основные понятия и определения	44
3.2. Задача о максимальном потоке. Алгоритм форда-Фалкерсона	45
3.3. Транспортная задача	48
ТЕМА 4. СЕТИ ПЕТРИ	52
4.1. Особенности сетей Петри и области их применения	52
4.2. Основные определения. Способы задания сетей Петри	53
4.3. Функционирование сетей Петри	55
4.4. Свойства сетей Петри	57
4.5. Анализ сетей Петри	58
4.6. Подклассы и расширения сетей Петри	60
ТЕМА 5. ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ ЛОГИКИ И ТЕОРИИ АВТОМАТОВ	62
5.1. Основные понятия алгебры логики	62
5.2. Элементарные булевы функции	63
5.3. Полнота системы булевых функций	65
5.4. Законы и тождества алгебры логики	66
5.5. Представление булевых функций дизъюнктивными и конъюнктивными нормальными формами	67
5.6. Минимизация функций алгебры логики	69
5.7. Неполностью определенные логические функции и их минимизация	74

5.8. Синтез комбинационных схем	75
5.9. Понятие о конечных автоматах и способы их задания.....	78
5.10. Синтез конечных автоматов	81
ТЕМА 6. МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ ЛИНЕЙНЫХ СИСТЕМ	
И ИХ ЭЛЕМЕНТОВ С ПОСТОЯННЫМИ ПАРАМЕТРАМИ.....	93
6.1. Классификация элементов.....	93
6.2. Уравнения динамики и статики	94
6.3. Понятие передаточной функции.....	95
6.4. Передаточные функции различных соединений звеньев	97
6.5. Временные характеристики систем и их элементов	98
6.6. Понятие о частотных характеристиках систем и их элементов	100
6.7. Понятие о логарифмических частотных характеристиках.....	102
6.8. Построение логарифмических частотных характеристик разомкнутых одноконтурных систем	108
6.9. Математические модели элементов в параметрах пространства состояний.....	109
6.10. Решение уравнений состояния первого порядка.....	111
6.11. Представление уравнений состояния при помощи матриц.....	112
6.12. Решение матричных уравнений состояния, представленных в нормальной форме	116
6.13. Характеристическое уравнение. Модальная матрица. Преобразование подобия	118
6.14. Каноническая форма уравнений состояния	119
6.15. Понятие об устойчивости линейных систем	123
6.16. Математическое описание дискретных систем и их элементов	127
6.17. Уравнения состояния и моделирование дискретных систем.....	131
ЛИТЕРАТУРА.....	134

ВВЕДЕНИЕ. ОБЩИЕ СРЕДСТВА МАТЕМАТИЧЕСКОГО ОПИСАНИЯ СИСТЕМ

Слово система происходит от греческого понятия «составленное» и определяется как совокупность элементов, объединенных так, что они образуют единое целое, функционируют согласованно и подчинены определенной форме управления. Системой может быть и электронный усилитель, и контур регулирования, в котором усилитель является одним из многих элементов, и технологический процесс, и промышленное предприятие и т. д. Отвлекаясь от физического содержания свойств объектов и связей, будем рассматривать систему как абстрактное множество элементов, находящихся друг с другом в некоторых отношениях. Одной из важнейших характеристик всякой системы является ее структура. Под структурой системы понимают совокупность элементов и связей между ними, которые определяются функциональным назначением системы и целями, поставленными перед системой. Современные сложные технические системы широко используют средства вычислительной техники, микроэлектронику, робототехнику и т. д. Важную роль в исследованиях, проектировании и эксплуатации подобных систем играют математические методы описания, анализа и синтеза. Рассмотрение системы как совокупности элементов дает возможность привлечь для ее математического описания аппарат теории множеств. Связи между элементами описываются уравнениями (алгебраическими или дифференциальными), либо с помощью аппарата математической логики. Формализовать описание структуры системы можно также на основе теории графов.

Основы теории множеств, теории графов и сетей, основы математической логики и теории автоматов составляют содержание первой части конспекта лекций по курсу «Математические основы теории систем».

ТЕМА 1. ОСНОВЫ ТЕОРИИ МНОЖЕСТВ

1.1. Основные понятия и определения

Под *множеством*, понимают совокупность определенных объектов, рассматриваемых как единое целое и хорошо различаемых между собой. Отдельные объекты, которые образуют множество, называются элементами множества. Обычно множества обозначаются прописными латинскими буквами X, Y, Z, A, B, C, \dots . Элементы множеств обозначаются соответственно строчными буквами x, y, z, a, b, c, \dots . Если элемент x принадлежит множеству X , то используется запись $x \in X$ (\in – символ принадлежности), в противном случае – $x \notin X$. Множество считается заданным, если перечислены все его элементы или указано характерное свойство, которым обладают все элементы множества. Для обозначения множества используют фигурные скобки $\{ \}$. Например, множество X цифр десятичного алфавита можно задать в виде $X = \{0, 1, 2, \dots, 9\}$ или $X = \{x \mid x - \text{целое}, x = 0, 9\}$, где справа от вертикальной черты указано свойство этого множества. Множество A четных чисел можно записать в виде $A = \{a \mid - \text{четное}\}$.

Множество называется *конечным*, если оно содержит конечное число элементов, и *бесконечным*, если содержит бесконечное число элементов. Множество, не содержащее ни одного элемента, называется пустым и обозначается \emptyset . Любое множество можно характеризовать мощностью (кардинальным числом). *Мощность* конечного множества определяется числом его элементов. Понятие мощности позволяет сравнивать между собой два бесконечных множества, сопоставляя элементы любого из них элементам другого и образуя пары соответствующих элементов.

Множества A и B называются равными ($A = B$), если любой элемент одного из них принадлежит другому, т. е. A и B представляют собой одно и то же множество.

Множество A' называется *подмножеством* множества A , если любой элемент множества A' принадлежит множеству A . Записывается так:

$$A' \subseteq A \leftrightarrow \{\forall a \mid a \in A' \rightarrow a \in A\},$$

где \subseteq – символ включения, читается « A содержит A' »; \leftrightarrow – символ эквивалентности, означающий «если и только если», или «то же самое, что»; \rightarrow – символ следствия, означающий «если ..., то» или «влечет за собой»; \forall – символ общности, означающий «любой», «для всех».

Множество A' строго включено в A или A' есть истинное подмножество A , если $A' \subset A$ и $A' \neq A$, т. е. множество A содержит и другие элементы, кроме элементов A' , \subset – символ строгого включения. Пустое множество и само множество A называются несобственными подмножествами множества A : $\emptyset \subseteq A$, $A \subseteq A$. Все другие подмножества множества A называются его собственными подмножествами.

Множество I , которое включает в себя все множества, участвующие в рассмотрении конкретной задачи, называется *полным*, или *универсальным*, или *единичным*.

Множества часто задают графически с помощью диаграмм Эйлера-Венна. При этом универсальное множество изображается в виде множества точек прямоугольника. Остальные множества изображаются в виде областей в прямоугольнике, которые находятся внутри замкнутых линий, называемых кругами Эйлера.

Например, на рис. 1.1 показаны три множества A , B , C , принадлежащие универсальному множеству I .

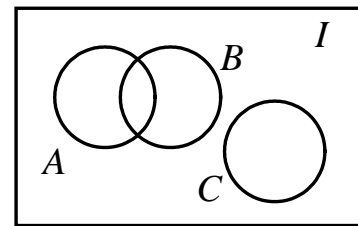


Рис. 1.1

1.2. Операции над множествами

Объединением множеств A и B называется множество C , состоящее из элементов, которые принадлежат хотя бы одному из множеств A или B . Объединение обозначается знаком \cup – «чашка». В соответствии с определением

$$C = A \cup B = \{c \mid c \in A, \text{ или } c \in B\}.$$

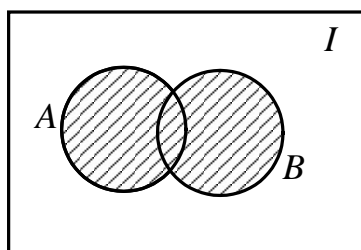
Если A и B имеют общие элементы, то каждый из них берется в C только один раз. Объединение множеств иногда называют суммой множеств.

Пересечением множеств A и B называется множество C , состоящее из элементов, которые принадлежат каждому из множеств A и B . Пересечение обозначается знаком \cap – «крышка»:

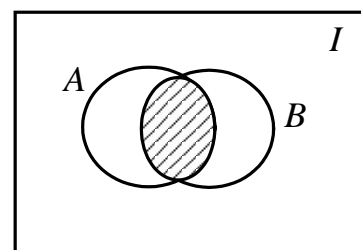
$$C = A \cap B = \{c \mid c \in A \text{ и } c \in B\}.$$

Множества называются **непересекающимися**, если не имеют общих элементов, т.е. если $A \cap B = \emptyset$. Пересечение множеств иногда называют произведением множеств.

На рис. 1.2 заштрихованные области дают геометрическую интерпретацию операций объединения (а) и пересечения (б) множеств A и B .



а) $C = A \cup B$



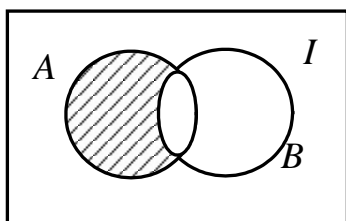
б) $C = A \cap B$

Рис. 1.2

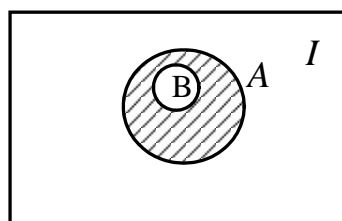
Разностью множеств A и B называется множество C , состоящее из элементов, которые принадлежат множеству A и не принадлежат множеству B . Обозначается $C = A \setminus B = \{c \mid c \in A \text{ и } c \notin B\}$.

Если $B \subset A$, то разность множеств A и B называется **дополнением** множества B до множества A .

На рис. 1.3 заштрихованные области иллюстрируют операции разности $A \setminus B$ (а) и дополнения B до множества A (б).



а) $C = A \setminus B$



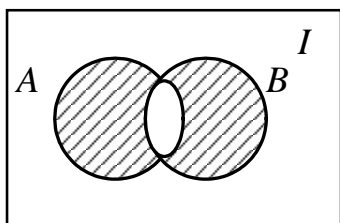
б) $C = A \setminus B, (B \subset A)$

Рис. 1.3

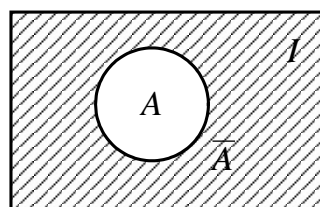
Дополнением множества A до универсального множества I называется множество \bar{A} , определяемое из соотношения $\bar{A} = I \setminus A$.

Симметрической разностью множеств A и B называется множество C , состоящее из элементов, принадлежащих в точности одному из множеств A и B . Обозначается $A \Delta B$, т. е. $C = A \Delta B = \{c \mid c \in (A \setminus B) \cup (B \setminus A)\}$.

Операции симметрической разности и дополнения до универсального множества проиллюстрированы на рис. 1.4, а и б соответственно.



а) $C = A \Delta B$



б) $\bar{A} = I \setminus A$

Рис. 1.4

Операции разности и симметрической разности могут быть выражены через операции $\cap, \cup, -$, так $A \setminus B = A \cap \bar{B}$; $A \Delta B = (A \setminus B) \cup (B \setminus A) = (A \cap \bar{B}) \cup (\bar{A} \cap B)$.

Разбиение множества – это представление его в виде системы подмножеств. Система множеств $\xi = (A_1, A_2, \dots, A_n)$ называется **разбиением множества** C , если она удовлетворяет следующим условиям:

1) любое множество A из ξ является подмножеством множества C :

$$\forall A \in \xi \rightarrow A \subseteq C;$$

2) любые два множества A_i и A_k из ξ являются непересекающимися:

$$\forall A \in \xi \rightarrow A_i \cap A_k = \emptyset, \text{ при } i \neq k;$$

3) объединение всех множеств, входящих в разбиение, дает множество C :

$$\cup A_i = A_1 \cup A_2 \cup \dots \cup A_n = C.$$

1.3. Законы и тождества алгебры множеств

Алгебра множеств представляет собой теоретико-множественный аналог обычной алгебры действительных чисел. В алгебре множеств рассматриваются основные свойства операций объединения \cup , пересечения \cap , дополнения $-$ и связей между ними, которые заданы в универсальном множестве I . Для перенесенных операций справедливы следующие законы [2]:

1) коммутативности объединения и пересечения $A \cup B = B \cup A$, $A \cap B = B \cap A$;

2) ассоциативности объединения и пересечения $A \cup (B \cup C) = (A \cup B) \cup C$,
 $A \cap (B \cap C) = (A \cap B) \cap C$;

3) дистрибутивности пересечения относительно объединения и объединения относительно пересечения $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
 $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$;

4) идемпотентности объединения и пересечения $A \cup A = A$, $A \cap A = A$;

5) де Моргана $\overline{A \cup B} = \bar{A} \cap \bar{B}$, $\overline{A \cap B} = \bar{A} \cup \bar{B}$;

6) двойного дополнения $\overline{\bar{A}} = A$;

7) действия с универсальным I и пустым \emptyset множествами $A \cup I = I$; $A \cap I = A$;
 $A \cup \bar{A} = I$; $A \cup \emptyset = A$; $A \cap \emptyset = \emptyset$; $A \cap A = \emptyset$.

Законы 1) и 2) имеют аналоги в обычной алгебре в виде переместительного и сочетательного законов для операций сложения и умножения. Однако дистрибутивный закон не имеет аналога, так как замена сложения умножением и наоборот привела бы к абсурдному выражению $a + (bc) = (a+b)(a+c)$.

Справедливость любого из этих законов можно проверить, показав, что множество, стоящее по одну сторону от знака равенства, включено в множество, стоящее по другую сторону от этого знака равенства.

Пример 1.1. Доказать, что $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

Пусть x – элемент множества, стоящего в левой части равенства, т. е. $x \in A \cup (B \cap C)$. Тогда $x \in A$ или $x \in B \cap C$. Если $x \in A$, то $x \in A \cup B$ и $x \in A \cup C$, а следовательно, x принадлежит и пересечению этих множеств, т. е. $x \in (A \cup B) \cap (A \cup C)$. Если $x \in B \cap C$, то $x \in B$ и $x \in C$. Следовательно, $x \in A \cup B$ и $x \in A \cup C$, так что и в этом случае $x \in (A \cup B) \cap (A \cup C)$.

Предположим теперь, что x – элемент множества, стоящего в правой части, т. е. $x \in (A \cup B) \cap (A \cup C)$. Тогда $x \in A \cup B$ и $x \in A \cup C$, следовательно, $x \in A$ или же $x \in B$ и $x \in C$. Из этого вытекает, что $x \in A \cup (B \cap C)$.

На рис. 1.5 приведены диаграммы Эйлера-Венна, подтверждающие, что выражения в правой и левой части доказываемого закона определяют одно и то же множество.

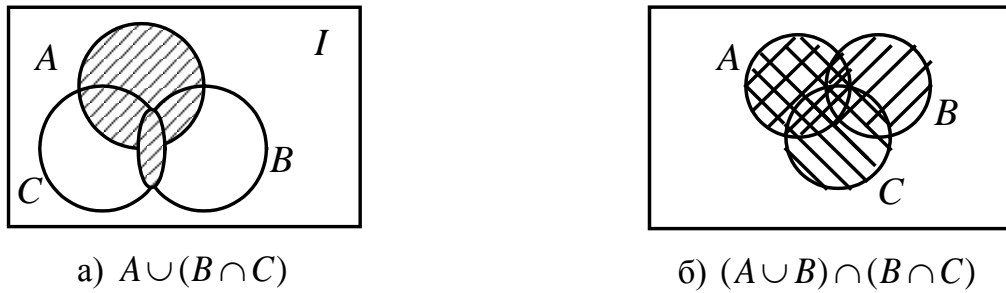


Рис. 1.5

Тождества алгебры множеств – это равенства, справедливые независимо от того, каково универсальное множество I и какие именно конкретные подмножества множества T обозначаются входящими в эти равенства буквами. Законы алгебры множеств позволяют упрощать различные сложные выражения, доказывать тождества подобно тому, как это делается в алгебре. При этом сначала выполняется операция дополнения, затем пересечения и только затем операция объединения (разности). Для изменения этого порядка используют скобки.

Пример 1.2. Доказать тождество $(\bar{X} \cap \bar{Y}) \cup (X \cap \bar{Y} \cap Z \cap L) \cup (\bar{L} \cap \bar{Y}) = \bar{Y}$.

На основании дистрибутивного закона «вынесем» из каждой скобки \bar{Y} , а затем применим коммутативный закон, двойное дополнение и закон де Моргана. Тогда

$$\begin{aligned} \bar{Y} \cap [(X \cap Z \cap L) \cup \bar{X} \cup \bar{Z} \cup \bar{L}] &= \bar{Y} \cap [(X \cap Z \cap L) \cup \overline{(\bar{X} \cup \bar{Z} \cup \bar{L})}] = \\ &= \bar{Y} \cap [(X \cap Z \cap L) \cup \overline{(X \cap Z \cap L)}] = \bar{Y} \cap I = \bar{Y}. \end{aligned}$$

Пример 1.3. Упростить выражение

$$\begin{aligned} [(A \cap B) \setminus C \cup \bar{A} \cup \bar{B} \cup C] \cap D &= [(A \cap B \cap \bar{C}) \cup (\bar{A} \cup \bar{B} \cup C)] \cap D = \\ &= \overline{[(A \cap B \cap \bar{C}) \cup (A \cup B \cup C)]} \cap D = [\overline{(\bar{A} \cup \bar{B} \cup C)} \cup (A \cup \bar{B} \cup C)] \cap D = I \cap D = D. \end{aligned}$$

1.4. Принцип двойственности

Пусть $F(A_1, \dots, A_n)$ – некоторая формула алгебры множеств, написанная с помощью символов \cap , \cup , $-$, пустого \emptyset и универсального I множеств.

Если есть операции \setminus и Δ , то их можно исключить с помощью формул

$$A \setminus B = A \cap \bar{B}; \quad A \Delta B = (A \cap \bar{B}) \cup (B \cap \bar{A}).$$

Формулу $F^*(A_1, \dots, A_n)$ называют **двойственной**, если она может быть получена из F , путем формальной замены символов \cap на \cup , \cup на \cap , \emptyset на I , I на \emptyset и, возможно, последующих преобразований.

Примеры двойственных формул:

$$F = (A \cap B) \cup (\bar{A} \cap I); F^* = (A \cup B) \cap (\bar{A} \cup \emptyset).$$

Если $F(A_1, \dots, A_n) = F^*(A_1, \dots, A_n)$, т. е. двойственные формулы тождественны, то $F(A_1, \dots, A_n)$ называется *самодвойственной*.

Пример самодвойственной формулы:

$$(A \cap B) \cup (A \cap C) \cup (B \cap C) = (A \cup B) \cap (A \cup C) \cap (B \cup C).$$

Докажем ее, используя дистрибутивный закон:

$$\begin{aligned} [A \cap (B \cup C)] \cup (B \cap C) &= [(B \cap C) \cup A] \cap [(B \cap C) \cup (B \cup C)] = \\ &= [(A \cup B) \cap (A \cup C)] \cap (B \cup C). \end{aligned}$$

1.5. Уравнение с множествами

Наряду с тождествами алгебра множеств рассматривает уравнения, которые содержат фиксированные подмножества A_1, A_2, \dots, A_n и неизвестные, подлежащие определению подмножества X_1, X_2, \dots, X_m . В простейшем случае в уравнение входит только одно такое подмножество X .

Решение уравнения осуществляется в следующей последовательности:

1) если в уравнении есть правая часть, то равенство преобразуется в симметрическую разность его левой и правой частей, которая приравнивается пустому множеству; это возможно на основании равенства $A = B$, если $A \Delta B = \emptyset$ или $(A \cap \bar{B}) \cup (\bar{A} \cap B) = \emptyset$;

2) полученное уравнение преобразуется к виду $(M \cap \bar{X}) \cup (N \cap X) = \emptyset$, где M и N – некоторые множества, не содержащие X ;

3) объединение множеств пусто только при условии, что каждое из них также пустое множество, поэтому преобразованное уравнение можно заменить зависимой системой двух уравнений

$$\begin{cases} M \cap X = \emptyset, \\ N \cap \bar{X} = \emptyset; \end{cases}$$

4) пара уравнений, а следовательно, и исходное уравнение имеет смысл, когда $x \subset \bar{M}$ и $\bar{x} \subset N$ или $N \subset X$. Если $N \subset X$ и $X \subset \bar{M}$, то $N \subset \bar{M}$. Решением уравнения является любое множество X , такое, что $N \subset X \subset \bar{M}$.

Пример 1.4. Найти множество X , если $X \cup C = D$.

1. Симметрическая разность имеет вид:

$$(X \cup C) \Delta D = [(X \cup C) \cap \bar{D}] \cup [\overline{(X \cup C)} \cap D] = \emptyset.$$

2. Преобразуем выражение в квадратных скобках:

$$\begin{aligned}
 [(X \cup C) \cap \bar{D}] \cup [(\overline{X \cup C}) \cap D] &= (X \cap \bar{D}) \cup (C \cap \bar{D}) \cup (\bar{X} \cap \bar{C} \cap D) = \\
 &= (X \cap \bar{D}) \cup [(C \cap \bar{D}) \cap (X \cup \bar{X})] \cup (\bar{X} \cap \bar{C} \cap D) = \\
 &= (X \cap \bar{D}) \cup (C \cap \bar{D} \cap X) \cup (C \cap \bar{D} \cap \bar{X}) \cup (\bar{X} \cap \bar{C} \cap D) = \\
 &= X \cap [\bar{D} \cup (C \cap \bar{D})] \cup \bar{X} \cup [(C \cap \bar{D}) \cup (\bar{C} \cap D)] = \\
 &= (X \cap \bar{D}) \cup \bar{X} \cap [(C \Delta D)] = \emptyset.
 \end{aligned}$$

3. Уравнение разбивается на два: $\bar{D} \cap X = \emptyset$ и $(C \Delta D) \cap \bar{X} = \emptyset$.

4. Так как $X \cap \bar{D} = \emptyset$, $x \subset D$. В то же время $C \Delta D \subset X$. Если $C \subset D$, то и $C \cap \bar{D} = \emptyset$ и $C \Delta D = \emptyset \cup [(\bar{C} \cap D)] = D \cap \bar{C} = D \setminus C$. Поэтому $D \setminus C \subset x \subset D$.

Следовательно, любое X , которое входит в D и содержит дополнение множества C до D , является решением уравнения $X \cup C = D$.

Геометрическое решение иллюстрируется рис. 1.6. Множество X обязательно содержит заштрихованную область и может включать любое подмножество из C .

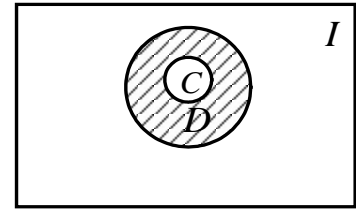


Рис. 1.6

1.6. Упорядоченное множество. Прямое произведение множеств

Упорядоченным множеством, или *кортежем*, называется последовательность элементов, в которой каждый элемент занимает определенное место. Например, множество людей, стоящих в очереди, – множество слов в фразе. Число элементов кортежа называется его длиной. Для обозначения кортежа используются круглые скобки. Так, множество $A = (a_1, a_2, \dots, a_n)$ является кортежем длины n с элементами a_1, \dots, a_n . Кортежи длины 2 называются парами, или упорядоченными парами, кортежи длины 3 – тройками, 4 – четверками и т. д. В общем случае кортежи длины n называются n -ками.

Прямым декартовым произведением множеств A и B называется множество, элементами которого являются упорядоченные пары (a, b) , в которых первая компонента принадлежит множеству A , а вторая принадлежит множеству B [1]. Прямое произведение множеств обозначается $A \times B$. Таким образом, по определению $A \times B = \{(a, b) : a \in A, b \in B\}$. Если $A = \{a_1, a_2, a_3\}$ и $B = \{b_1, b_2, b_3, b_4\}$, то $A \times B = \{(a_1, b_1), (a_1, b_2), (a_1, b_3), (a_1, b_4), \dots, (a_3, b_4)\}$. Операция прямого произведения распространяется и на большее число множеств.

Прямым произведением r множеств называется множество, состоящее из кортежей длины r , первая компонента которых принадлежит первому множеству, вторая – второму, третья – третьему и т. д.

Пусть к предыдущему произведению добавится третий сомножитель $D = \{d_1, d_2, d_3\}$, тогда декартово произведение $A \times B \times D = \{(a_1, b_1, d_1), (a_1, b_1, d_2),$

$(a_1, b_1, d_3), (a_1, b_2, d_1), \dots, (a_3, b_4, d_3)\}$. Частным случаем операции прямого произведения является понятие степеней множества. S -й степенью множества X , обозначаемой X^S , называется прямое произведение S одинаковых множеств, равных X .

Геометрической интерпретацией упорядоченной двойки (a_1, a_2) является точка на плоскости или вектор, проведенный из начала координат в данную точку. Компоненты a_1 и a_2 будут проекциями вектора на оси 1 и 2.

$$\text{Пр}_1(a_1, a_2) = a_1; \text{Пр}_2(a_1, a_2) = a_2 \quad (\text{рис. 1.7}).$$

Пусть $\{X = 1, 2\}, Y = 1, 3, 4\}$, тогда $X \times Y = \{(1, 1), (1, 3), (1, 4), (2, 1), (2, 3), (2, 4)\}$ (рис. 1.8, а). Если X и Y – отрезки вещественной оси, то прямое произведение – заштрихованная область (рис. 1.8, б). Кортеж (a_1, a_2, a_3) может рассматриваться как точка в трехмерном пространстве или трехмерный вектор, проведенный из начала координат в эту точку.

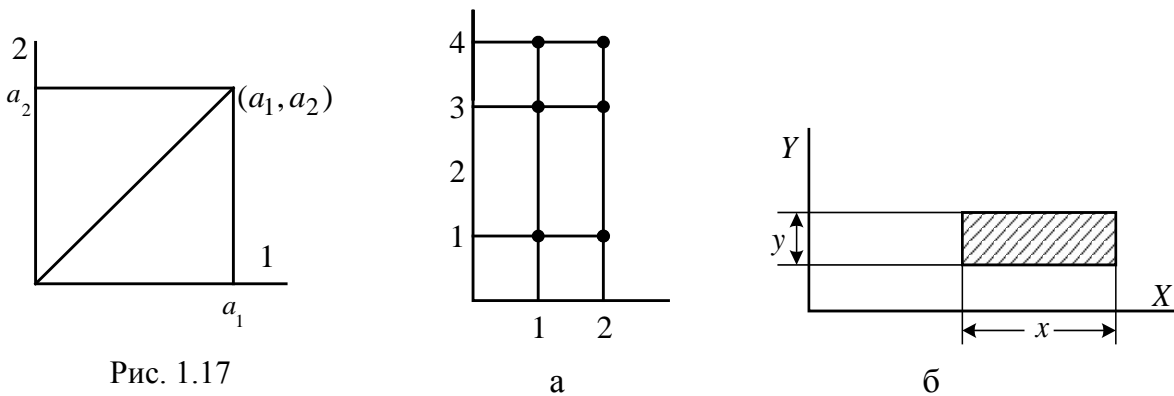


Рис. 1.17

Рис. 1.8

1.7. Соответствия

Рассмотрим два множества $X = \{x_1, x_2, \dots, x_n\}$ и $Y = \{y_1, y_2, \dots, y_m\}$. Если для элементов множества X указаны элементы множества Y , с которыми они сопоставляются, то говорят, что между множествами X и Y установлено соответствие. При этом *не обязательно, чтобы в сопоставлении участвовали все элементы множеств X и Y* . Соответствие q представляет собой тройку множеств $q = (X, Y, Q)$, где X и Y – это множества, элементы которых сопоставляются.

Множество $Q \subseteq X \times Y$ определяет закон, по которому осуществляется соответствие, т. е. перечисляет все пары, участвующие в сопоставлении. Для каждого $q = (X, Y, Q)$ можно указать обратное соответствие $q^{-1} = (X, Y, Q^{-1})$, где $Q^{-1} = Y \times X$.

Композицией соответствий называется последовательное применение двух или более соответствий. Геометрическое представление прямого и обратного соответствий показано на рис. 1.9, а и б соответственно.

Например, это может быть закрепление машин за шоферами и распределение шоферов по машинам в таксопарке. Обратное соответствие обратного соответствия даст прямое соответствие $(q^{-1})^{-1} = q$.

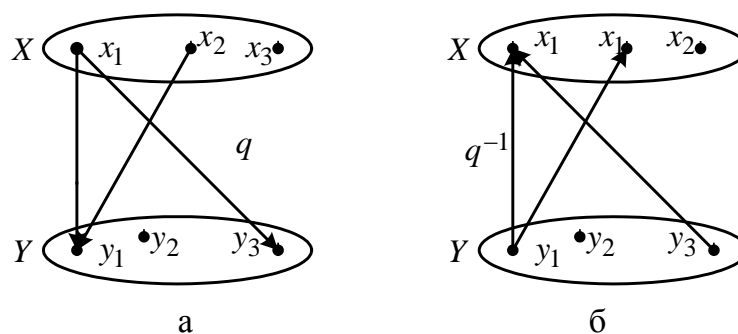


Рис. 1.9

Соответствие называется **взаимно однозначным**, если каждому элементу множества X соответствует (поставлен в пару с ним) единственный элемент множества Y и наоборот. Если между X и Y установлено взаимно однозначное соответствие, то они имеют поровну элементов.

1.8. Отображения и их виды

Отображение является частным случаем соответствия. Соответствие, характеризующее правило, по которому **каждому** элементу множества X сопоставляется один или несколько элементов множества Y , называется отображением и записывается как $\Gamma: X \rightarrow Y$, где множество Γ определяет закон отображения. Пусть $X = \{x_1, x_2, x_3\}$; $Y = \{y_1, y_2, y_3, y_4, y_5, y_6\}$. Каждому элементу $x_i \in X$ отображение Γ ставит в соответствие некоторое подмножество $\Gamma \subseteq Y$, называемое образом элемента x : $\Gamma_{x_1} = \{y_1, y_2\}$, $\Gamma_{x_2} = \{y_3\}$, $\Gamma_{x_3} = \{y_4, y_5, y_6\}$.

На рис. 1.10 показано геометрическое представление этого отображения.

Определяют три вида отображений: сюръективное, инъективное и биективное.

Отображение называется **сюръективным** (или отображением «на»), если образы точек множества X заполняют все множество Y , причем различные точки множества X могут иметь один и тот же образ.

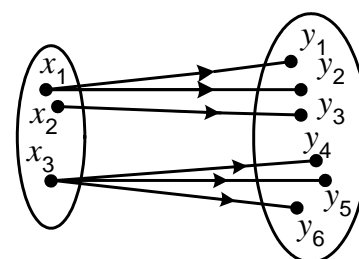


Рис. 1.10

Отображение называется **инъективным** (или отображением «в»), если элементы множества X отображаются не на все множество Y , а в его какую-то часть. При этом каждому элементу $x \in X$ соответствует один элемент $y \in Y$ и обратно, прообразом y является один элемент x .

Геометрическое представление сюръективного и инъективного отображений приведено на рис. 1.11, а и б соответственно.

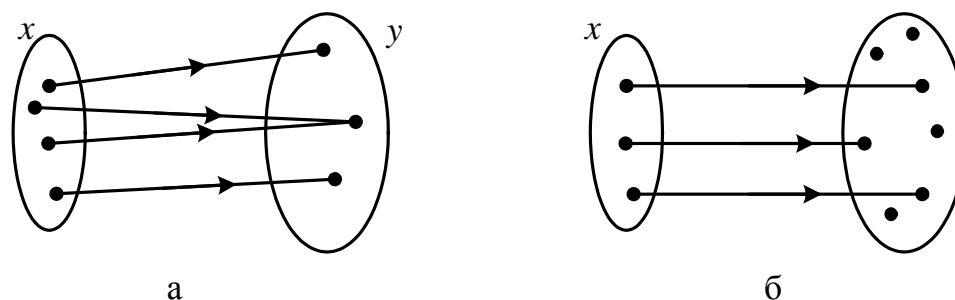


Рис. 1.11

Биективное отображение является одновременно инъективным и сюръективным, т. е. является взаимно однозначным.

Важным случаем отображения является отображение элементов внутри одного множества. При этом отображение $\Gamma: X \rightarrow X$ будет определяться парой (X, Γ) , где $\Gamma \subseteq X \times X$ или $\Gamma \subseteq X^2$. С помощью отображений могут быть даны определения таким понятиям, как функция, функционал, оператор, которые широко используются при математическом описании систем. Если отображение $\Gamma: X \rightarrow Y$ рассматривается как соответствие между множествами X и Y , то множество $f = \{(x, y) \in X \times Y : y = f(x)\}$ называется функцией. Таким образом, f является множеством, элементами которого являются пары (x, y) , участвующие в соответствии, и $f(x)$ является обозначением для $y \in Y$, соответствующего данному $x \in X$ [1].

Функционал устанавливает зависимость между множеством чисел, с одной стороны, и некоторым множеством функций с другой. Примером функционала может служить определенный интеграл вида
$$I(f) = \int_a^b f(x) dx.$$

Оператор устанавливает соответствие между двумя множествами функций. Если обозначить через p оператор дифференцирования, то связь между производной $f'(x)$ и функцией $f(x)$ может быть записана в виде операторного соотношения $f'(x) = p[f(x)]$.

1.9. Отношения и их свойства

Для обозначения некоторых видов отображений, заданных на одном и том же множестве, используется понятие «*отношение*» [1]. Пусть отображение (X, Γ) является отношением. Если элемент x_1 находится в отношении R к элементу x_2 , то это записывается как $x_1 R x_2$ или $(x_1, x_2) \in R$, где R – символ отношения. Примером отношений могут служить такие понятия, как «меньше, чем», «делится на», «включено в», «больше чем» и т. д.

Отношение между двумя элементами называется **бинарным**, или двухместным, между тремя – **тернарным**, или трехместным, между n элементами – **n -нарным**, или n -местным. Различают шесть основных свойств отношений:

рефлексивность, антирефлексивность, симметричность, антисимметричность, тождественность, транзитивность.

Отношение R называется **рефлексным** на множестве X , если для любого $x \in X$ справедливо xRx или $(xx) \in R$ на множестве X . Например, «равенство», «самообслуживание».

Отношение R называется **антирефлексивным**, если для любого $x \in X$ не выполняется xRx , т. е. $(xx) \notin R$. Например, «строгое неравенство», «быть старше», т. е. отношения, которые могут выполняться только для несовпадающих объектов.

Отношение R называется **симметричным** на множестве X , если для любых $x \in X$ справедливо соотношение: если x_1Rx_2 , то x_2Rx_1 или если $(x_1x_2) \in R$, то $(x_2x_1) \in R$. Например, «расстояние между двумя точками», «быть братом».

Отношение R называется **антисимметричным** на множестве X , если для любых $x \in X$ справедливо соотношение: если x_1Rx_2 истинно, то x_2Rx_1 ложно, или если $(x_1x_2) \in R$, то $(x_1x_2) \notin R$. Например, «строгое включение», «быть отцом».

Отношение R называется **тождественным** на множестве X , если для любых $x \in X$ из одновременной истинности x_1Rx_2 и x_2Rx_1 следует, что $x_1 = x_2$. Например, «получают повышенную стипендию» и «сдали сессию на хорошо и отлично» на множестве студентов факультета.

Отношение R называется **транзитивным** на множестве X , если для любых $x \in X$ справедливо соотношение: если x_1Rx_2 и x_2Rx_3 , то x_1Rx_3 .

Например, «параллельность», «больше чем». Для каждого отношения R можно определить обратное R^{-1} , считая, что $x_2R^{-1}x_1$ в том и только в том случае, когда x_1Rx_2 .

1.10. Виды отношений

В зависимости от того, какими свойствами обладает отношения, они делятся на три вида; отношение эквивалентности, отношение порядка и отношение доминирования.

Отношение R на множестве X называется **отношением эквивалентности**, если оно обладает свойствами рефлексивности (xRx для $\forall x \in X$), симметричности ($x_1Rx_2 \rightarrow x_2Rx_1$ для всех $x_1, x_2 \in X$), транзитивности (x_1Rx_2 и $x_2Rx_3 \rightarrow x_1Rx_3$ для всех $x_1, x_2, x_3 \in X$). Элементы множества можно рассматривать как эквивалентные, если любой из этих элементов может быть заменен другим. Для обозначения эквивалентности служит символ \equiv или \sim , т. е. $x_1 \equiv x_2$ или $x_1 \sim x_2$. Примерами отношения эквивалентности являются отношения равенства векторов, фигур, геометрическое отношение подобия, отношение параллельности.

Другим видом отношений является **отношение порядка**, которое характеризуется определенным порядком расположения элементов множества. Раз-

личают отношения нестрогого порядка (квазипорядка) на множестве X и отношения строгого порядка.

Отношение квазипорядка обозначается символом \leq , частным случаем является символ \subseteq . Это отношение, которое обладает свойствами рефлексивности ($x \leq x$), транзитивности (если $x_1 \leq x_2$ и $x_2 \leq x_3$, то $x_1 \leq x_3$) и тождественности (если $x_1 \leq x_2$ и $x_2 \leq x_1$, то $x_1 = x_2$).

Отношение **строгого порядка** обозначается символом $<$, частным случаем его являются символы \rightarrow , \subset . Это отношение, обладающее свойствами транзитивности (если $x_1 < x_2$ и $x_2 < x_3$, то $x_1 < x_3$) и антисимметричности (если $x_1 < x_2$, то $x_2 < x_1$). Это отношение характерно для различного рода иерархий с подчинением одного объекта другому.

При рассмотрении множеств людей, или группы людей, возникает отношение, которое называется отношением доминирования (доминировать \sim это значит преобладать). Обозначается оно символом \gg .

Отношением доминирования на множестве x называется отношение, которое обладает свойствами антирефлексивности ($x \gg x$ ложно) и антисимметричности (если $x_1 \gg x_2$, то $x_2 \gg x_1$, ложно). Эти свойства означают, что никакой индивидуум не может доминировать самого себя и в каждой паре только один элемент доминирует другого. Например, победившая команда или спортсмен; лицо, пользующееся авторитетом.

1.11. Нечёткие множества. Способы задания. Понятие лингвистической переменной

Теория нечётких множеств ведёт своё начало с 1965 года. Основоположителем теории нечётких множеств «Fuzzy Sets» является американский учёный Лотфи Заде, который ввёл понятие о нечётких множествах, как обобщение обычных (чётких) множеств. Прилагательное «fuzzy» переводится на русский язык как нечёткий, размытый.

Введение нечётких множеств – это попытка описать математически некоторую нечёткую информацию для создания математической модели. В нечётких множествах объекты, обладающие общим свойством, могут обладать им в различной степени. Поэтому высказывание типа $x \in X$ («элемент принадлежит данному множеству») теряет смысл. Необходимо указать «насколько сильно» или с какой степенью конкретный элемент удовлетворяет свойствам данного множества.

Нечёткое или фази-множество (ФМ) характеризуется двумя показателями:

- 1) фактом принадлежности объектов к множеству;
- 2) степенью принадлежности объектов к данному множеству.

Для представления элемента x нечёткого множества \tilde{A} используется функция принадлежности $\mu_{\tilde{A}}(x)$, которая равна 1, если этот элемент принадлежит к

множеству \tilde{A} или равна 0, если элемент не принадлежит множеству \tilde{A} . В общем случае $\mu_{\tilde{A}}(x) \in [0,1]$. Значения функции принадлежности являются рациональными числами из интервала $[0,1]$. Конкретное значение функции принадлежности называется *коэффициентом* или *степенью принадлежности*. Эта степень может быть определена явным образом в виде функциональной зависимости, либо дискретно для конечной последовательности значений x .

Нечётким множеством \tilde{A} называется совокупность пар $(\mu_{\tilde{A}}(x)/x)$, где $\mu_{\tilde{A}}(x)$ – степень принадлежности элемента x к нечёткому множеству \tilde{A} .

$$\tilde{A} = \{ \mu_{\tilde{A}}(x_1)/x_1, \mu_{\tilde{A}}(x_2)/x_2, \dots, \mu_{\tilde{A}}(x_N)/x_N \}.$$

Так, например, нечёткое множество целых чисел, определённое понятием «около 10», можно задать следующим образом

$$\tilde{A} = \{0,1/7; 0,3/8; 0,8/9; 1/10; 0,8/11; 0,1/13\}.$$

Другой способ задания этого же множества – задание его функции принадлежности зависимостью $\mu_{\tilde{A}}(x) = \frac{1}{1+(x-10)^2}$. График функции представлен на рис. 1.12.

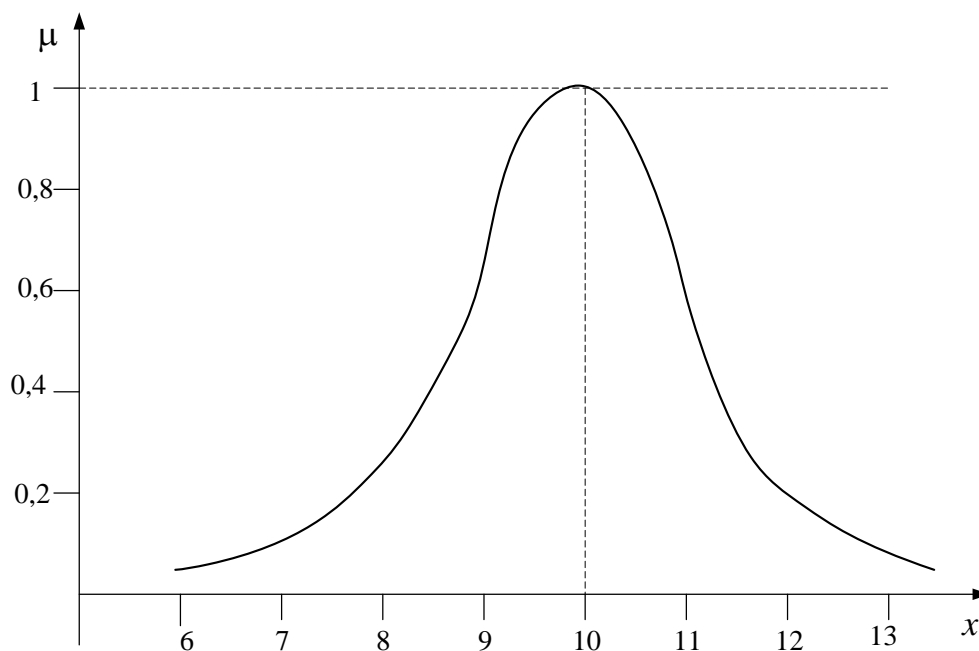


Рис. 1.12

Наибольшей популярностью пользуются функции принадлежности гауссовского типа, а также треугольной и трапецеидальной формы. Гауссовская функция принадлежности переменной x с центром в C и вариацией σ для множества \tilde{A} описывается выражением

$$\mu_{\tilde{A}} = \exp \left[- \left(\frac{x - C}{\sigma} \right)^2 \right].$$

и приведена на рис 1.13.

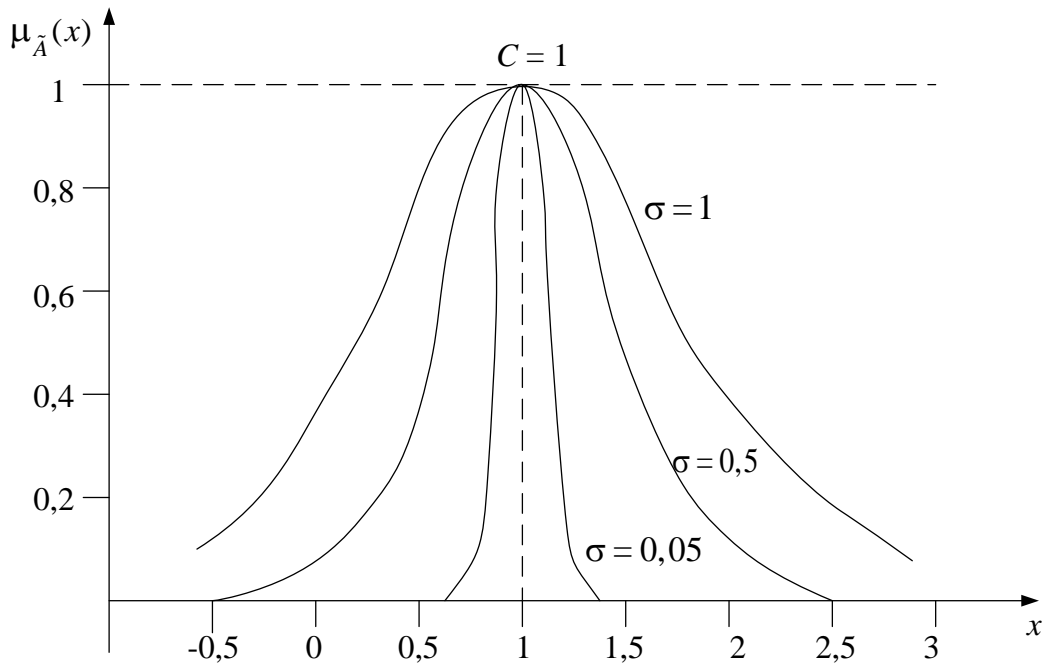


Рис. 1.13

Параметр σ характеризует форму функции. Чем меньше σ , тем больше крутизна функции.

$C = \text{const} = 1$ – центр нечёткого множества, при его изменении функция смещается по горизонтальной оси.

Сходство $\mu(x)$ с функцией распределения только внешнее. Определить форму (не обязательно симметричную) должен специалист в соответствующей области. В общем случае $\mu(x)$ может быть кусочно-непрерывной.

Треугольная симметричная функция принадлежности описывается выражением

$$\mu_{\bar{A}} = \begin{cases} 1 - \frac{|x - C|}{d}, & \text{для } x \in [c - d; c + d], \\ 0, & \text{для остальных } x \end{cases}$$

и приведена на рис 1.14.

Обобщением треугольной функции принадлежности является трапециевидальная функция, которая описывается зависимостью

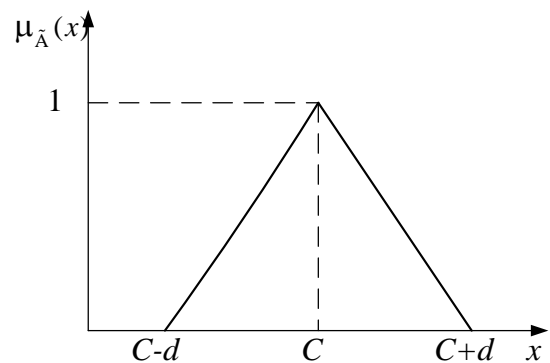


Рис. 1.14

$$\mu_{\bar{A}}(x) = \begin{cases} 0, & \text{для } x > z \text{ или } x < y, \\ 1, & \text{для } C - \frac{t}{2} \leq x \leq C + \frac{t}{2}, \\ S(z - x), & \text{для } C + \frac{t}{2} \leq x \leq z, \\ S(z - y), & \text{для } y \leq x \leq C - \frac{t}{2}. \end{cases}$$

Вид соответствующей функции показан на рис. 1.15. Здесь принято, $y = C - \frac{t}{2} - \frac{1}{S}$, $z = C + \frac{t}{2} + \frac{1}{S}$, где S обозначает угол наклона.

Выбор значения $t = 0$ преобразует трапециевидальную функцию в треугольную. Применительно к техническим системам объектами нечётких множеств являются значения некоторых физических переменных, например, значения температуры, скорости перемещения, электрического напряжения, тока и т. д.

Физическую переменную можно описать словесно (лингвистически), выделив некоторую качественную оценку в лингвистической форме. Так же как обычная переменная может принимать различные значения, лингвистическая переменная, например, « x = температура», может принимать различные значения такие как:

- отрицательная малая OM ,
- нулевая Z ,
- положительная средняя PS ,
- положительная высокая PW и т. п.

Лингвистические переменные называются **термами**.

Терм-множеством называется множество всех значений лингвистической переменной.

Лингвистические переменные (термы) количественно оцениваются не числами, а числовыми множествами, перекрывающимися друг друга. Физическая переменная на заданном интервале X её изменения от x_{\min} до x_{\max} характеризуется рядом фаззи-множеств $\tilde{A}_1, \tilde{A}_2, \dots, \tilde{A}_n$ с функциями принадлежности $\mu_1(x), \mu_2(x), \dots, \mu_n(x)$. По существу определение формы, места расположения центра C и границ интервала x возлагается на эксперта в соответствующей области.

На рис. 1.16 переменная температура помещения представлена термами

- положительная низкая PN для $x = 4 \div 16^\circ\text{C}$;
- положительная средняя PS для $x = 12 \div 28^\circ\text{C}$;
- положительная высокая PW для $x > 23^\circ\text{C}$.

На участках перекрытия термов нарушается однозначность принадлежности значений переменной x только одному терму. Ширина участков перекрытия может быть различной, в пределе и нулевой, однако там, где для одного терма $\mu_i(x) = 1$, для соседних $0 \leq \mu_{i\pm 1}(x) < 1$.

Значение $\mu(\alpha_i) = 1$ означает беспорную принадлежность значения $x = \alpha_i$ к соответствующему терму.

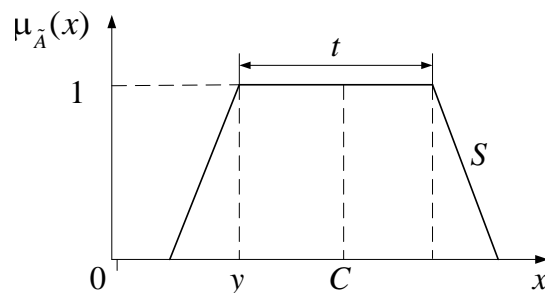


Рис. 1.15

$T = 20^{\circ}\text{C}$ – наиболее комфортная для самочувствия человека принимается за среднюю.

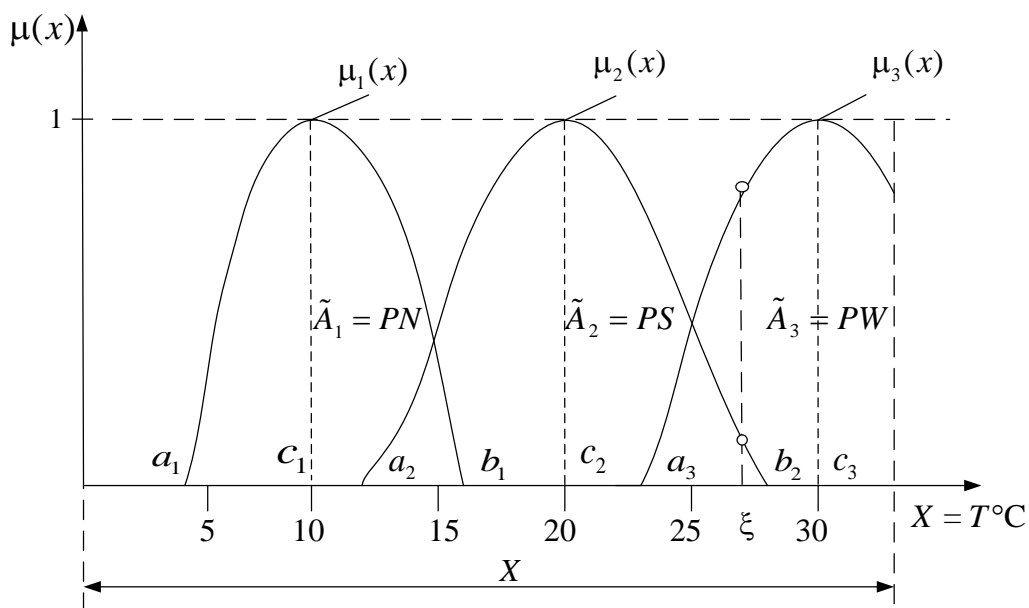


Рис. 1.16

$T = 10^{\circ}\text{C}$ – несомненно низкая, когда следует включить обогреватель.

$T = 30^{\circ}\text{C}$ – слишком высокая, когда требуется включить кондиционер.

При $T = \xi$ температура относится одновременно к двум термам и возникает проблема, какое решение принять.

1.12. Операции над нечёткими множествами

Дополнением нечёткого множества \tilde{A} заданного на X называется нечёткое множество \tilde{A}^c с функцией принадлежности $\mu_{\tilde{A}^c}(x) = 1 - \mu_{\tilde{A}}(x)$.

На рис. 1.17 приведен пример выполнения операции нечёткого дополнения.

Пересечением нечётких множеств \tilde{A} и \tilde{B} , заданных на X , называется нечёткое множество $\tilde{C} = \tilde{A} \cap \tilde{B}$, с функцией принадлежности $\mu_{\tilde{C}}(x) = \min\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}$, для всех $x \in X$. Операция нахождения минимума также обозначается знаком \wedge , г. е. $\mu_{\tilde{C}}(x) = \mu_{\tilde{A}}(x) \wedge \mu_{\tilde{B}}(x)$.

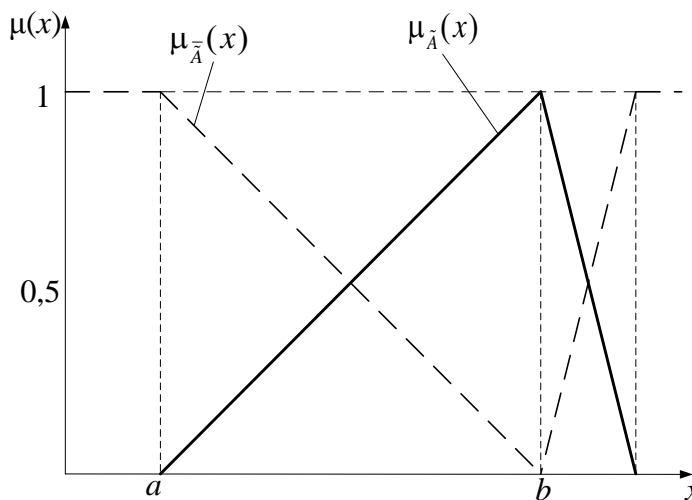


Рис. 1.17

Объединением нечётких множеств \tilde{A} и \tilde{B} , заданных на X , называется нечёткое множество $\tilde{D} = \tilde{A} \cup \tilde{B}$, с функцией принадлежности $\mu_{\tilde{D}}(x) = \max\{\mu_{\tilde{A}}(x), \mu_{\tilde{B}}(x)\}$ для всех $x \in X$. Операция нахождения максимума также обозначается знаком \vee , т.е. $\mu_{\tilde{D}}(x) = \mu_{\tilde{A}}(x) \vee \mu_{\tilde{B}}(x)$.

На рис. 1.18 приведён результат операций \cup и \cap фаззи множеств \tilde{A} и \tilde{B} одной переменной x .

Линия, выделенная более чётко и идущая выше, соответствует дизъюнкции двух функций принадлежности $(\mu_{\tilde{A}}(x) \vee \mu_{\tilde{B}}(x))$, а другая линия – конъюнкции двух функций принадлежности $(\mu_{\tilde{A}}(x) \wedge \mu_{\tilde{B}}(x))$.

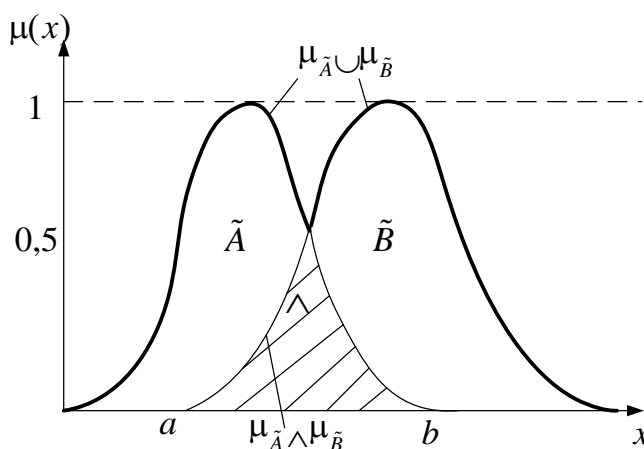


Рис. 1.18

Включение нечётких множеств. Пусть заданы \tilde{A} и \tilde{B} на множестве X . \tilde{A} называется подмножеством нечёткого множества \tilde{B} и обозначается $\tilde{A} \subseteq \tilde{B}$, если $\mu_{\tilde{A}}(x) \leq \mu_{\tilde{B}}(x)$, $\forall x \in X$. Нечёткие множества \tilde{A} и \tilde{B} разны, если $\mu_{\tilde{A}}(x) = \mu_{\tilde{B}}(x)$, $\forall x \in X$.

1.13. Параметры нечётких множеств

Высотой нечёткого множества \tilde{A} называется верхняя граница его функции принадлежности $h(\tilde{A}') = \sup_{x \in X} \mu_{\tilde{A}}(x)$. Для дискретного множества X супремум соответствует максимуму степеней принадлежности его элементов.

Нечёткое множество называется **нормальным**, если его высота равна единице, в противном случае множества называются субнормальными.

Нормализация – преобразование субнормального нечёткого множества \tilde{A}' в нормальное \tilde{A} . Определяется следующим образом

$$\mu_{\tilde{A}}(x) = \frac{\mu_{\tilde{A}'}(x)}{h(\tilde{A}')}, \quad \forall x \in X.$$

Пусть $\mu_{\tilde{A}'}(x) = \frac{0,6}{1+(10-x)^2}$, высота $h = 0,6$.

На рис 1.19 показана нормализация нечёткого множества \tilde{A}' .

Ядром нечёткого множества \tilde{A} называется чёткое подмножество универсального множества X , элементы которого имеют степени принадлежности, равные 1.

$$\text{Core}(\tilde{A}) = \{x | \mu_{\tilde{A}}(x) = 1\}.$$

Ядро субнормального нечёткого множества пустое.

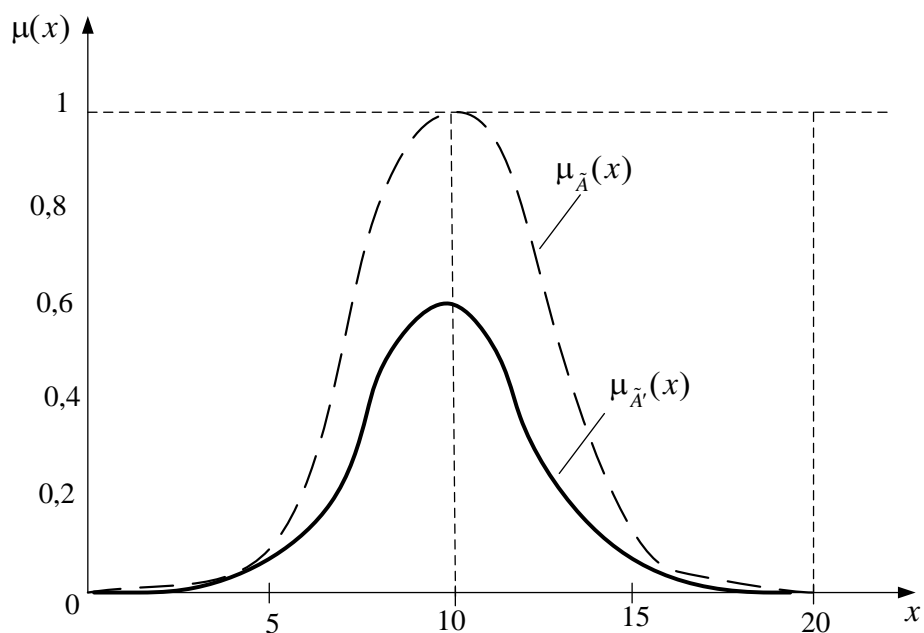


Рис. 1.19

α -сечением нечёткого множества \tilde{A} называется чёткое подмножество универсального множества X , элементы которого имеют степени принадлежности большие или равные α

$$A_\alpha = \{x | \mu_{\tilde{A}}(x) \geq \alpha\}, \text{ при } \alpha \in [0, 1].$$

Носителем нечёткого множества \tilde{A} называется чётное подмножество универсального множества X , элементы которого имеют ненулевые степени принадлежности. (англ: support – носитель).

Ядро, α -сечение и носитель показаны на рис. 1.20.

$$\text{supp}(\tilde{A}) = \{x | \mu_{\tilde{A}}(x) > 0\}.$$

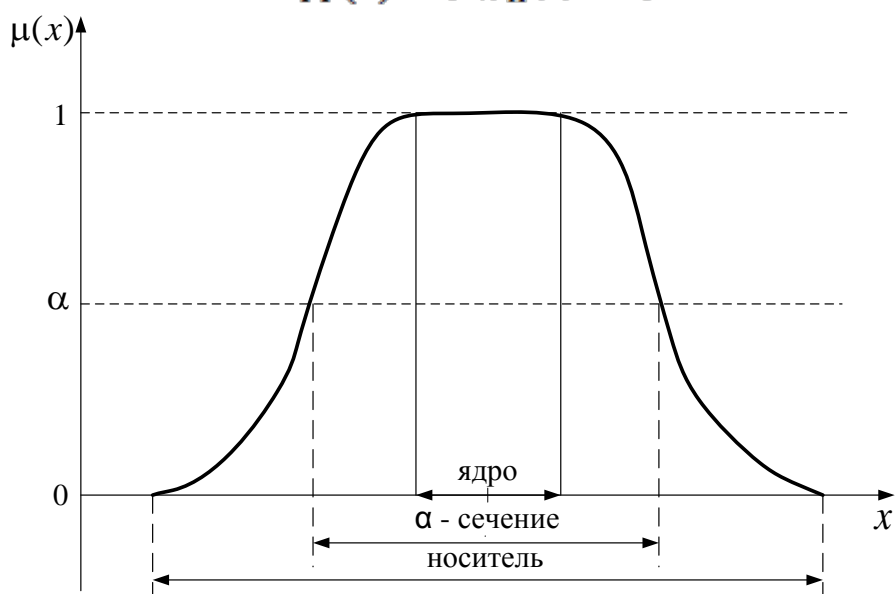


Рис. 1.20

Кардинальное число M (или мощность) нечёткого множества \tilde{A} равно сумме степеней принадлежности всех элементов к этому множеству

$$M(\tilde{A}) = \sum \mu_{\tilde{A}}(x).$$

1.14. Методы дефаззификации нечётких множеств

Дефаззификация – процедура преобразования нечёткого множества в чёткое число. Простейший способ, пригодный только для одноэкстремальных функций принадлежности, заключается в выборе чёткого числа ξ , соответствующего максимуму функции принадлежности.

Метод центра тяжести. Дефаззификация \tilde{A} осуществляется по формуле (для дискретного универсального множества X).

$$\xi = \frac{\sum_{i=1}^n x_i \cdot \mu_{\tilde{A}}(x_i)}{\sum_{i=1}^n \mu_{\tilde{A}}(x_i)}.$$

Пример: Провести дефаззификацию нечёткого множества «мужчина среднего роста», заданного следующим образом

$$\tilde{A} = \{0/155; 0,1/160; 0,3/165; 0,8/170; 1/175; 1/180; 0,5/185; 0/190\}.$$

Результат дефаззификации имеет вид

$$\xi = \frac{0 \cdot 155 + 0,1 \cdot 160 + 0,3 \cdot 165 + 0,8 \cdot 170 + 1 \cdot 175 + 1 \cdot 180 + 0,5 \cdot 185 + 0 \cdot 190}{0 + 0,1 + 0,3 + 0,8 + 1 + 1 + 0,5 + 0}.$$

Для непрерывного универсального множества X – процедура находит центр тяжести плоской фигуры, ограниченной осями координат и графиком функции принадлежности $\mu_{\tilde{A}}(x)$ в соответствии с выражением

$$\xi = \frac{\int_a^b x \cdot \mu_{\tilde{A}}(x) dx}{\int_a^b \mu_{\tilde{A}}(x) dx}.$$

Метод медианы. Дефаззификация нечёткого множества \tilde{A} по методу медианы состоит в нахождении такого числа ξ , что

$$\int_a^{\xi} \mu_{\tilde{A}}(x) dx = \int_{\xi}^b \mu_{\tilde{A}}(x) dx.$$

Геометрически это означает нахождение такой точки на оси абсцисс, что перпендикуляр, восстановленный в этой точке, делит площадь под кривой функции принадлежности на две равные части.

Метод центра максимумов. Дефаззификация нечёткого множества \tilde{A} осуществляется по формуле

$$\xi = \frac{\int_G x dx}{\int_G dx},$$

где G – множество всех элементов из интервала $[a, b]$, имеющих максимальную степень принадлежности нечёткому множеству \tilde{A} .

Таким образом, находится среднее арифметическое значение элементов множества X , имеющих максимальные степени принадлежности. Если множество X конечно, то формула приводится к виду

$$\xi = \frac{\sum_{x_i \in G} x_i}{|G|},$$

где $|G|$ – мощность множества G .

Теория нечётких множеств и её приложения используется при управлении сложными и трудно описываемыми математически технологическими процессами. Фаззи управление не требует знания точной модели объекта. Оно организует приближённую стратегию управления, моделируя способ мышления человека и задавая алгоритм решения в лингвистических переменных. Реализация управления осуществляется фаззи – контроллером с применением фаззи – логики.

ТЕМА 2. ЭЛЕМЕНТЫ ТЕОРИИ ГРАФОВ И ЕЕ ПРИЛОЖЕНИЯ

Первая работа по графам была опубликована математиком Эйлером в 1736 году. Она содержала решение задачи о кенигсбергских мостах: можно ли совершить прогулку таким образом, чтобы, выйдя из любого места города, вернуться в него, пройдя в точности один раз по каждому мосту (рис. 2.1, а).

По условию задачи не имеет значения, как проходит путь по частям суши a, b, c, d , поэтому их можно представить точками или вершинами. А так как связи осуществляются через семь мостов, то каждый из них можно изобразить линией, соединяющей эти вершины (рис. 2.1, б). Эйлер дал отрицательный ответ на вопрос задачи.

Начало развития теории графов как самостоятельной математической дисциплины положено Д. Кенигом, выпустившим в 1936 году книгу «Теория конечных и бесконечных графов». В настоящее время круг задач, решаемых с помощью аппарата теории графов, очень разнообразен: анализ и синтез цепей и систем, проектирование каналов связи и исследование процессов передачи информации, автоматизация проектирования, теория кодирования и теория игр, выбор оптимальных маршрутов и потоков в сетях и т. д.

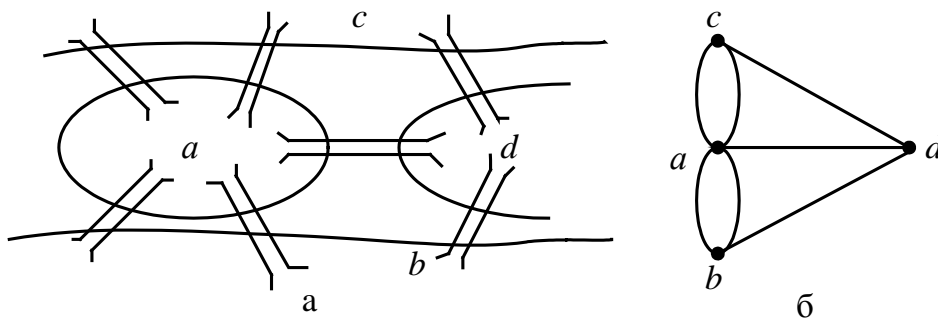


Рис. 2.1

2.1. Основные понятия и определения. Способы задания графов

Ориентированный граф G представляет собой множество элементов с их отображениями в этом множестве и обозначается символом $G = (X, \Gamma)$, где $X = \{x_1, x_2, \dots, x_n\}$ – множество элементов, а $\Gamma: X \rightarrow X$ – множество, определяющее закон отображения. Поясним данное определение посредством описания различных способов задания графа: аналитического, геометрического и матричного.

При **аналитическом способе** задания для каждого элемента x_i множества X должно быть определено отображение Γ_{x_i} . Пусть, например, $X = \{x_1, x_2, x_3, x_4\}$, а $\Gamma_{x_1} = \{x_2, x_3\}$, $\Gamma_{x_2} = \{x_3\}$, $\Gamma_{x_3} = \{x_2, x_4\}$, $\Gamma_{x_4} = \{x_1, x_4\}$. Эти множества однозначно определяют ориентированный граф G .

При **геометрическом способе** задания графа элементы множества X изображаются точками плоскости и называются вершинами графа. Линии, соединяющие любые пары точек x_i, x_j , из которых x_j является отображением x_i , называются дугами, или ориентированными ребрами. Дуги графа имеют направление, обозначаемое стрелкой в направлении от x_i к x_j .

На рис. 2.2 приведен пример графа, который выше был задан аналитически. Вершины графа могут располагаться в произвольном порядке и соединяться прямыми или кривыми линиями. Если $x_i = x_j$, то дуга изображается линией без стрелки и называется петлей. Каждую дугу (x_i, x_j) можно обозначить буквой $v_k \in V$, где V – множество упорядоченных дуг рассматриваемого графа. Тогда граф G можно определить также как $G = (X, V)$, где V – множество упорядоченных пар $(x_i, x_j) = v_k$. Две вершины x_i и x_j

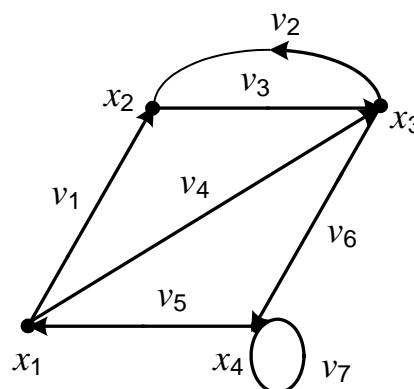


Рис. 2.2

называются **смежными**, если существует соединяющая их дуга $v_k = (x_i, x_j)$. Если вершина x_j является одним из концов дуги v_k , то говорят, что они **инцидентны**, т. е. вершина инцидентна дуге, а дуга инцидентна вершине. Таким образом, смежность – отношение между однородными объектами, инцидентность – между разнородными.

При **матричном способе** задания ориентированный граф можно описать матрицей смежности, или матрицей инцидентности. **Матрица смежности** R_G ориентированного графа $G(X, \Gamma)$ с n вершинами – это квадратная матрица порядка n , элементы r_{ij} которой определяются следующим образом:

$$r_{ij} = \begin{cases} 1, & \text{если есть дуга, идущая из } x_i \text{ в } x_j; \\ 0, & \text{в противном случае.} \end{cases}$$

Матрица инцидентности A_G ориентированного графа $G(X, \Gamma)$ – это прямоугольная матрица размером $n \times$

m (n – число вершин, m – число дуг), элементы a_{ij} которой определяются следующим образом:

$$a_{ij} = \begin{cases} 1, & \text{если } j\text{-я дуга выходит из } i\text{-й вершины;} \\ -1, & \text{если } j\text{-я дуга заходит в } i\text{-ю вершину;} \\ 0, & \text{если } i\text{-я вершина не инцидентна } j\text{-й дуге.} \end{cases}$$

Для графа, изображенного на рис. 2.2, матрицы смежности и инцидентности имеют вид

$$R_G = \begin{matrix} & \begin{matrix} x_1 & x_2 & x_3 & x_4 \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{matrix} & \begin{bmatrix} 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 \end{bmatrix} \end{matrix}, \quad A_G = \begin{matrix} & \begin{matrix} v_1 & v_2 & v_3 & v_4 & v_5 & v_6 & v_7 \end{matrix} \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 1 & -1 & 0 & 0 \\ -1 & -1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & -1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & -1 & 1^* \end{bmatrix} \end{matrix}.$$

Знаком * в матрице A_G помечена петля v_7 при вершине x_4 .

Полустепень исхода вершины x_i – число дуг, исходящих из вершины x_i , обозначается $\rho^+(x_i)$.

Полустепень захода вершины x_i – число дуг, заходящих в вершину x_i , обозначается $\rho^-(x_i)$.

Иногда граф рассматривают без учета ориентации его дуг, в этом случае граф называют неориентированным.

На рис. 2.3 изображен неориентированный граф, который получен из графа на рис. 2.2 удалением стрелок с дуг.

Такой неориентированный граф называется соотнесенным данному ориентированному. Будем обозначать неориентированные графы символом $D = (X, V)$, где V – множество неупорядоченных пар $(x_i, x_j) = v_k$.

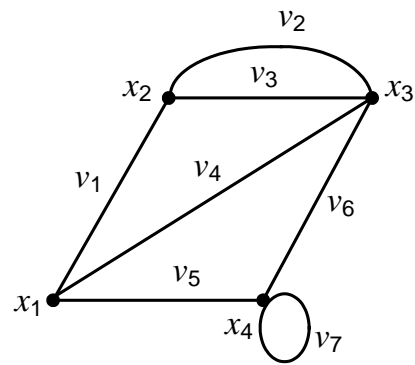


Рис. 2.3

Матрица смежности R_D неориентированного графа $D = (X, V)$ с n вершинами – это квадратная матрица порядка n , элементы r_{ij} которой определяются следующим образом:

$$r_{ij} = \begin{cases} \ell, & \text{если } i\text{-я и } j\text{-я вершины соединены } \ell \text{ ребрами;} \\ 0, & \text{в противном случае.} \end{cases}$$

Матрица смежности неориентированного графа всегда симметрична относительно главной диагонали.

Матрица инцидентности A_D неориентированного графа $D = (X, V)$ представляет собой прямоугольную матрицу размером $n \times m$ (n – число вершин, m – число ребер), элементы a_{ij} которой определяются следующим образом:

$$a_{ij} = \begin{cases} 1, & \text{если } i\text{-я вершина инцидентна } j\text{-му ребру;} \\ 0, & \text{в противном случае.} \end{cases}$$

Для графа, изображенного на рис. 2.3, матрицы смежности и инцидентности имеют вид

$$R_D = \begin{matrix} & x_1 & x_2 & x_3 & x_4 \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{matrix} & \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 2 & 0 \\ 1 & 2 & 0 & 1 \\ 1 & 0 & 1 & 1 \end{bmatrix}, \end{matrix}$$

$$A_D = \begin{matrix} & v_1 & v_2 & v_3 & v_4 & v_5 & v_6 & v_7 \\ \begin{matrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 1 & 1 & 0 & 0 \\ 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 1 & 1 & 1 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 \end{bmatrix}. \end{matrix}$$

Степенью вершины x_i в графе D называется число ребер, инцидентных вершине x_i . Обозначается $\rho_i = \rho(x_i)$. Вершина x_k степени $\rho_k = 1$ называется **концевой, или висячей** вершиной. Вершина степени нуль – **изолированная** вершина. Сумма степеней вершин графа равна удвоенному числу его ребер: $\sum \rho_i = 2m$.

2.2. Типы графов

Граф без петель и кратных ребер называется **простым**. Граф без петель, но с кратными ребрами называется **мультиграфом** (рис. 2.4, а). Наибольшее число ребер образует мультичисло и называется **кратностью**. Простой граф, в котором две любые вершины соединены ребром, называется **полным** и обозначается K_n , где n – число вершин.

На рис. 2.4, б приведен пример полного графа K с пятью вершинами.

Граф называется **двудольным** (биграфом), если множество его вершин X может быть разбито на два таких подмножества X_1 и X_2 , что каждое ребро имеет один конец в подмножестве X_1 , а другой в подмножестве X_2 , при этом $X_1 \cap X_2 = \emptyset$, $X_1 \cup X_2 = X$. Пример двудольного графа приведен на рис. 2.4, в, где $X_1 = \{x_1, x_3, x_5, x_7\}$, $X_2 = \{x_2, x_4, x_6\}$. Полный двудольный граф обозначается $K_{m,n}$, где m и n – количество вершин в подмножествах X_1 и X_2 соответственно.

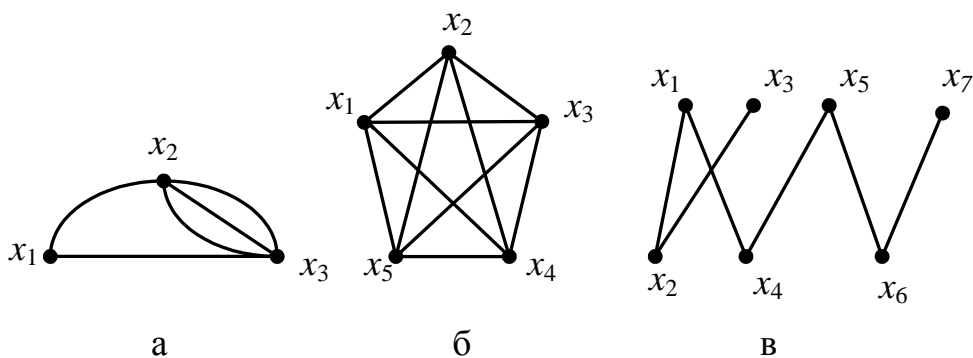


Рис. 2.4

Подграфом графа D (или G) называется граф, в которой входит лишь часть вершин графа D (или G) вместе с ребрами, соединяющими эти вершины. **Частичным** графом по отношению к графу D (или G) называется граф, содержащий только часть ребер графа.

На рис. 2.5 показаны граф D , его подграф и частичный граф.

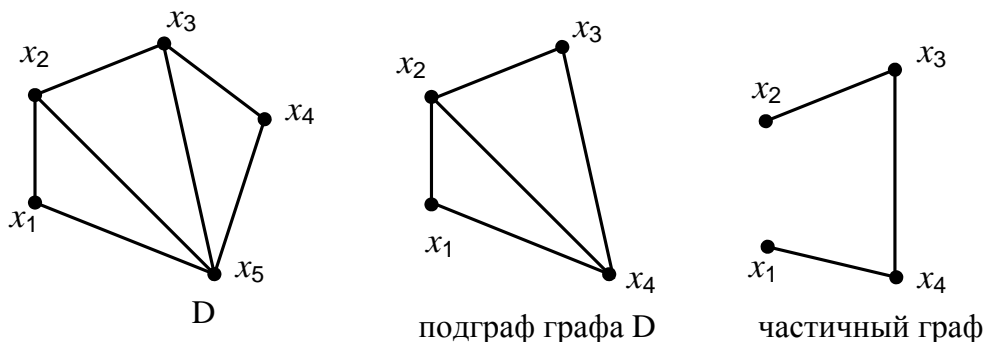


Рис. 2.5

Граф называется **связным**, если каждую его вершину можно соединить с любой другой его вершиной некоторой последовательностью ребер. Если граф не связан, то его можно разбить на подграфы так, что все его вершины в каждом подграфе связны. Такие подграфы называются компонентами связности графа.

На рис. 2.6, а приведен пример связного графа, а на рис 2.6, б – несвязного графа с тремя компонентами связности.

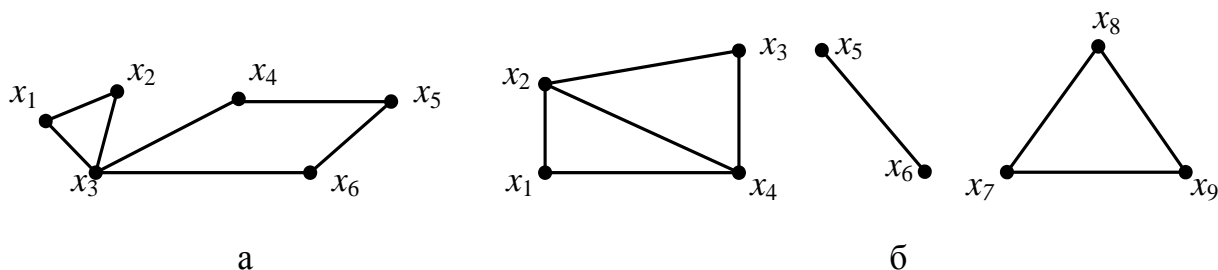


Рис. 2.6

Графы называются **изоморфными**, если между множествами их вершин существует взаимнооднозначное соответствие, такое, что вершины соединены ребрами в одном из графов в том и только в том случае, если соединены соответствующие им вершины в другом графе.

Изоморфные графы, приведенные на рис. 2.7, по существу различаются лишь начертанием. Если существенные свойства графа не связаны со способом его изображения на плоскости или нумерацией его вершин и ребер, то изоморфные графы, как правило, не различают между собой.

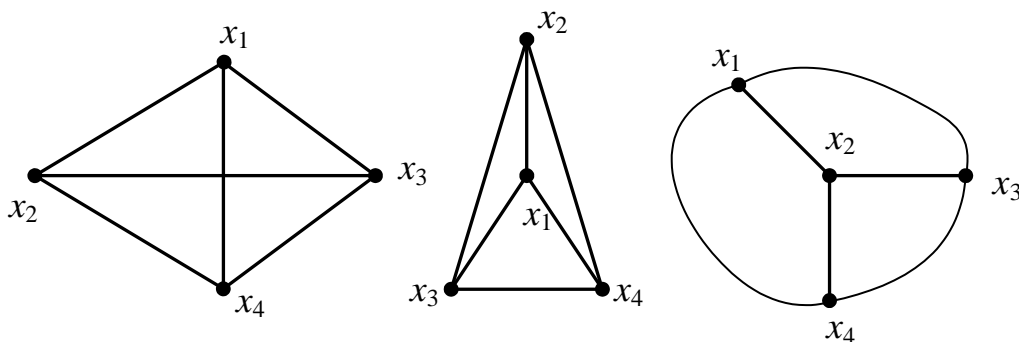


Рис. 2.7

Изоморфизм – это отношение эквивалентности на графах. Граф $D = (X, V)$ называется плоским (планарным), если существует изоморфный ему граф, который может быть изображен на плоскости без пересечения ребер.

Пример планарного графа приведен на рис. 2.8, а, б. Хотя в одном из графов ребра пересекаются, изоморфный ему граф не имеет пересечений, а следовательно, он плоский.

На рис. 2.8, в, г показаны два не плоских графа, играющих фундаментальную роль в теории планарности и называемых графами Понтрягина-Куратовского.

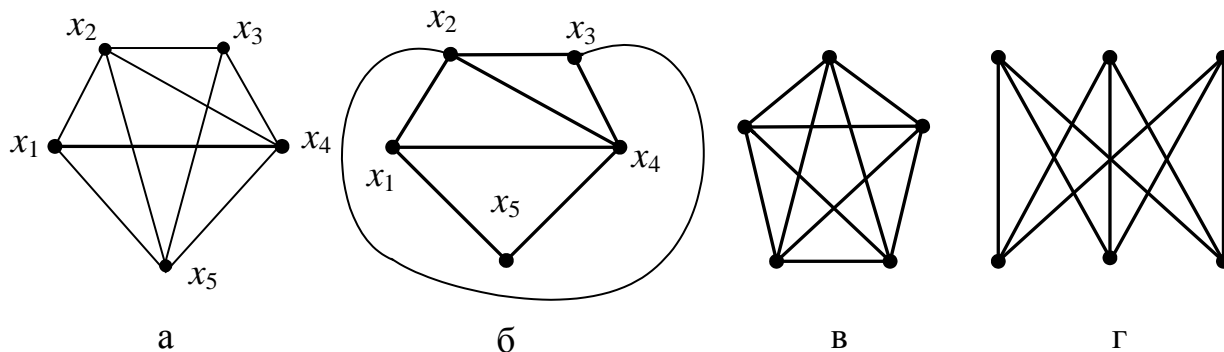


Рис. 2.8

Это полный граф K_5 и двудольный граф $K_{3,3}$. Доказано, что граф является плоским тогда и только тогда, когда он не содержит подграфов, стягиваемых к графам K_5 и $K_{3,3}$. Элементарное стягивание графа осуществляется следующим образом. В графе $D = (X, V)$ выбирается ребро (x_i, x_j) и удаляется из графа. При этом вершины x_i и x_j сохраняются и отождествляются, т. е. $x_i = x_j$. Отождествленная вершина будет смежна в новом графе с теми вершинами, с которыми были смежны как x_i , так и x_j в графе G . Из получившихся в результате процедуры кратных ребер оставляется одно. Граф D называется стягиваемым к графу K , если K можно получить из D путем последовательных элементарных стягиваний.

На рис. 2.9 показана процедура стягивания графа K_5 к графу K_4 .

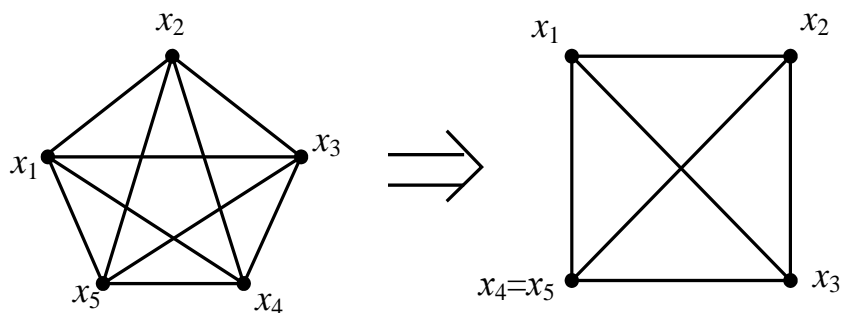


Рис. 2.9

Следует заметить, что при операции удаления ребра из графа все вершины сохраняются, вершина удаляется из графа вместе с инцидентными ей ребрами.

Планарность является существенным свойством графов, которые моделируют коммуникации и связи между объектами на плоскости (линии передач электроэнергии, межсоединения на печатных платах электронных устройств и кристаллах интегральных схем, водопроводные и газовые сети и т. д.).

2.3. Расстояния и пути в графах. Центры и периферийные вершины

Путь в ориентированном графе G – это последовательность дуг, в которой конец каждой предыдущей дуги совпадает с началом последующей. Путь μ обозначается последовательностью вершин, которые в него входят, например, $\mu = (x_1, x_2, \dots, x_n)$. Длина ℓ пути μ определяется числом дуг, составляющих этот путь $\ell(\mu) = k$. Путь, в котором никакая дуга не встречается дважды, называется **простым**. Путь, в котором никакая вершина не встречается дважды, называется **элементарным**. **Контур** – это конечный путь $\mu = (x_1, x_2, \dots, x_k)$, у которого начальная вершина x_1 совпадает с конечной x_k . Контур называется элементарным, если все его вершины различны.

Для графа, изображенного на рис. 2.10, можно указать: $\mu_1 = (x_1, x_2, x_5, x_4, x_2, x_3, x_6)$ – простой путь; $\mu_2 = (x_1, x_2, x_3, x_6)$ – элементарный путь; $\mu_3 = (x_2, x_5, x_4, x_2)$ – контур.

Отклонением $\mu d(x_i, x_j)$ вершины x_i от вершины x_j называется длина кратчайшего пути из x_i в x_j : $d(x_i, x_j) = \min \{ \ell(x_i, x_j) \}$.

Отклоненностью вершины x_i называется число $d(x_i) = \max d(x_i, x_j)$, т.е. это наибольшее из отклонений вершины x_i от всех остальных.

Матрица отклонений $d(x_i, x_j)$ для графа на рис. 2.10 представлена табл. 2.1, а вектор отклоненностей $d(x_j)$ – табл. 2.2.

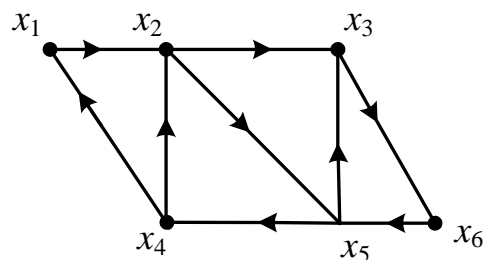


Рис. 2.10

Таблица 2.1

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	0	1	2	3	2	3
x_2	3	0	1	2	1	2
x_3	4	4	0	3	2	1
x_4	1	1	2	0	2	3
x_5	2	2	1	1	0	2
x_6	3	3	2	2	1	0

Таблица 2.2

	$d(x_i)$
x_1	3
x_2	3
x_3	4
x_4	3
x_5	2
x_6	3

Вершина графа с наименьшей отклоненностью называется **центром графа**. В графе может быть несколько центров. Вершина с наибольшей отклоненностью называется **периферийной вершиной**. **Радиусом $\rho(G)$** ориентированного графа называется отклоненность центра. **Диаметром $D(G)$** ориентированного графа называется отклоненность периферийной вершины. В рассматриваемом графе вершина x_5 является центром, а вершина x_3 является периферийной вершиной, соответственно $\rho(G) = 2$; $D(G) = 4$. В неориентиро-

ванных графах перемещаться можно в любом направлении, здесь вместо понятий «путь», «отклонение» и «отклоненность» используются понятия «цепь», «расстояние» и «удаленность». Замкнутая цепь называется циклом. **Расстоянием** $d(x_i, x_j)$ между двумя вершинами x_i и x_j неориентированного графа G называется длина кратчайшей простой цепи, соединяющей эти вершины: $d(x_i, x_j) = \min \{ \ell(x_i, \dots, x_j) \}$. **Удаленностью вершины** x_i называется число $d(x_i) = \max d(x_i, x_j)$, соответствующее наибольшему из расстояний от вершины x_i до всех остальных.

На рис. 2.11 представлен соотнесенный не ориентированный граф, соответствующий графу, изображенному на рис. 2.10. Матрица расстояний $d(x_i, x_j)$ и вектор удаленностей $d(x_i)$ представлены табл. 2.3 и 2.4. Центрами графа на рис. 2.11 в соответствии с табл. 2.4 будут являться вершины x_2, x_3, x_4 и x_5 с наименьшей удаленностью. Радиус $\rho(G) = 2$.

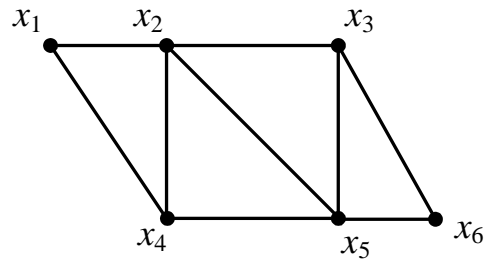


Рис. 2.11

Периферийными вершинами являются вершины x_1 и x_6 с наибольшей удаленностью. Диаметр графа $D(G) = 3$.

Таблица 2.3

	x_1	x_2	x_3	x_4	x_5	x_6
x_1	0	1	2	1	2	3
x_2	1	0	1	1	1	2
x_3	2	1	0	2	1	1
x_4	1	1	2	0	1	2
x_5	3	2	1	2	1	0
x_6	3	2	1	2	1	0

Таблица 2.4

	$d(x_i)$
x_1	3
x_2	2
x_3	2
x_4	2
x_5	2
x_6	3

2.4. Числовая функция на графе. Сигнальные графы

Числовую функцию на графе задают либо на вершинах, либо на дугах графа. **Числовая функция на вершинах** графа G считается заданной, если каждой вершине x_i ставится в соответствие некоторое число q_i из некоторого множества Q . **Числовая функция на дугах** графа G считается заданной, если каждой дуге $v = (x_i, x_j)$ ставится в соответствие число $l(v)$ из некоторого множества L . Количественные значения, приписываемые вершинам или дугам, называются **весами**. В некоторых случаях числовая функция на графе задается комбинированным способом, как на вершинах, так и на дугах.

Для моделирования физических систем используются взвешенные ориентированные графы, называемые **сигнальными графами**, или **графами потоков сигналов**. Вершины сигнального графа отождествляются с некоторыми пере-

менными x_i , называемыми **сигналом вершины**. Дуги отображают связи между переменными, и каждая дуга (x_i, x_j) характеризуется величиной k_{ij} , называемой **передачей дуги**. Величина k_{ij} представляет собой численное или функциональное отношение, характеризующее передачу сигнала от одной вершины к другой. Для одиночной дуги (x_i, x_j) $k_{ij} = x_j / x_i$ или $x_j = k_{ij} \cdot x_i$. Если в одну вершину x_j сходится n направленных к ней дуг, то сигнал вершины определяется суммой

$$x_j = k_{1j} \cdot x_1 + k_{2j} \cdot x_2 + \dots + k_{nj} \cdot x_n = \sum_{i=1}^n k_{ij} \cdot x_i.$$

Наличие выходящих дуг не влияет на сигнал вершины x_j , эти дуги влияют на сигналы других вершин. Приведенное равенство указывает на способ построения графа по заданной системе линейных алгебраических уравнений и, наоборот, на способ записи алгебраических уравнений, соответствующих данному графу.

Пример сигнального графа представлен на рис. 2.12. Соответствующая ему система уравнений имеет вид

$$\begin{cases} x_2 = k_{12}x_1 + k_{22}x_2, \\ x_3 = k_{13}x_1 + k_{23}x_2 + k_{43}x_4, \\ x_4 = k_{24}x_2 + k_{34}x_3. \end{cases}$$

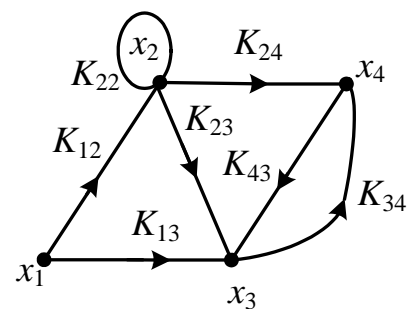


Рис. 2.12

Вершина, имеющая только выходящие дуги, называется **источником**.

Вершина, имеющая только входящие дуги, называется **стоком**. В целом граф топологически отображает передачу сигнала от источников к стокам. Рассмотрим правила преобразования сигнальных графов [5], пользуясь эквивалентными преобразованиями простейших подграфов

1. Последовательное соединение двух одинаково направленных дуг с передачами a и b может быть заменено одной эквивалентной дугой, передача которой равна произведению передач исходных дуг (рис. 2.13, а).

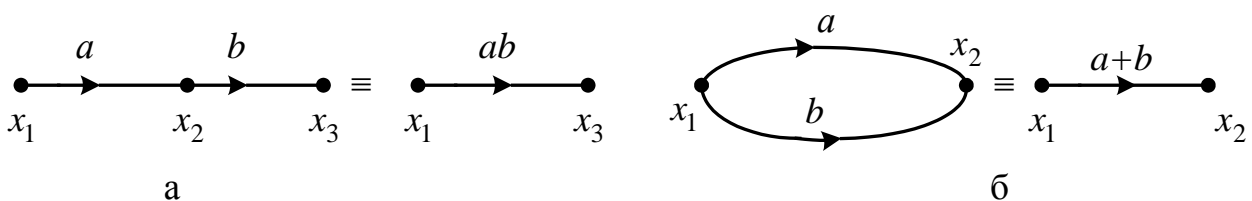


Рис. 2.13

2. При параллельном соединении двух одинаково направленных дуг его можно заменить одной дугой с передачей, равной сумме передач исходных дуг (рис. 2.13, б).

3. Пусть в некоторую вершину x_j входит несколько дуг и несколько дуг выходит из нее (рис. 2.14). При отсутствии петель на вершине x_j после ее устранения каждая входная вершина (x_1, x_2, x_3) будет связана с каждой выходной

вершиной (x_4, x_5) дугой, передача которой равна произведению передач дуг, расположенных между ними и внутренней вершиной.

Справедливость этого вытекает из рассмотрения системы уравнений, соответствующей данному подграфу:

$$\begin{cases} x_j = ax_1 + bx_2 + cx_3, \\ x_4 = dx_j, \\ x_5 = ex_j. \end{cases}$$

Исключая из этих уравнений переменную x_j , имеем

$$\begin{cases} x_4 = adx_1 + bdx_2 + cdx_3, \\ x_5 = aex_2 + bex_2 + cex_3. \end{cases}$$

4. Если удаляемая вершина имеет петлю с передачей f (рис. 2.15), то передача эквивалентной ветви умножается на $1/(1-f)$. Это следует из системы уравнений

$$\begin{cases} x_2 = ax_1 + fx_2, \\ x_3 = bx_2. \end{cases}$$

Устраняя переменную x_2 , получим $x_3 = (ab/(1-f))x_1$.

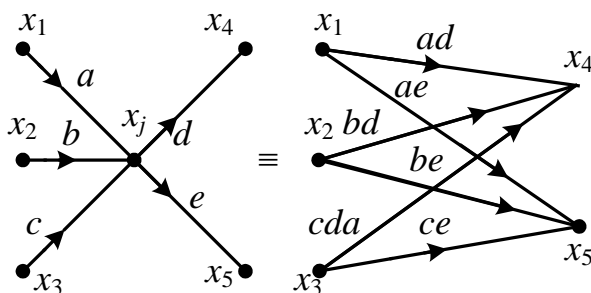


Рис. 2.14

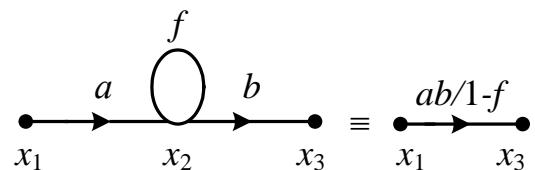
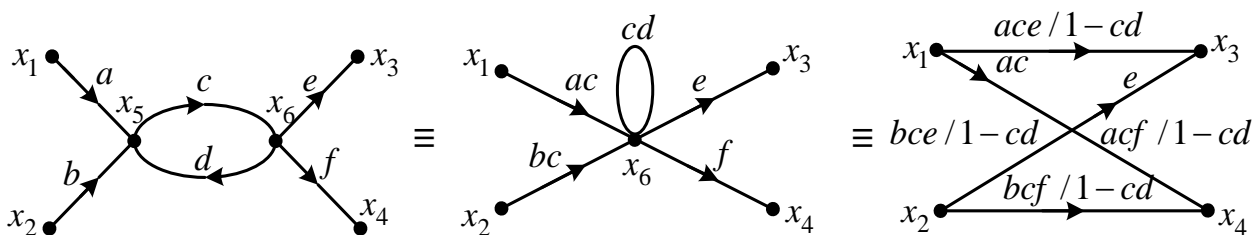


Рис. 2.15

5. Граф имеет в наличии дуги, соединяющие две вершины и направленные в различные стороны относительно друг друга (рис. 2.16). При удалении вершины x_5 образуется петля при вершине x_6 с передачей cd . Применяя затем правило устранения вершины с петлей, получаем эквивалентный подграф с четырьмя вершинами, показанный в цепочке преобразований на рис. 2.16.



Справедливость этого преобразования вытекает из системы уравнений подграфа:

$$\begin{cases} x_5 = ax_1 + bx_2 + dx_3, \\ x_6 = cx_5, \\ x_3 = ex_6, \\ x_4 = fx_6. \end{cases}$$

Подставив x_5 из первого уравнения во второе, получим: $x_6 = acx_1 + bcx_2 + dcx_3$, откуда $x_6 = (ac/1-cd)x_1 + (bc/1-cd)x_2$, тогда $x_3 = ace/1-cd)x_1 + (bce/1-cd)x_2$; $x_4 = (acf/1-cd)x_1 + (bcf/1-cd)x_2$.

Применяя правила преобразования к графу произвольной конфигурации, его можно свести к одной эквивалентной дуге, передача которой будет характеризовать передачу графа от начальной до конечной вершины. В ряде случаев этот метод оказывается громоздким.

2.5. Правило Мэзона

Передача сигнального графа от источника до некоторой вершины может быть определена по правилу Мэзона, которое известно так же, как правило несоприкасающихся контуров. Контур не соприкасаются, если они не имеют общих вершин.

На предварительном этапе в исследуемом графе выделяются все элементарные пути от источника до рассматриваемой вершины и все контуры графа. Далее определяются передачи путей и контуров. **Передача пути P** равна произведению передач дуг вдоль этого пути. **Передача контура L** равна произведению передач дуг, входящих в этот контур. В соответствии с правилом Мэзона передача графа от i -го источника до j -й вершины определяется по формуле

$$K_{ij} = (1/D) \sum_s P_s D_s,$$

где P_s – передача S -го элементарного пути; D – определитель графа, который зависит только от передач контуров; D_s – алгебраическое дополнение для S -го пути; S – число элементарных путей.

Определитель графа D вычисляется в соответствии с выражением

$$D = 1 - \sum_r L_r^{(1)} + \sum_r L_r^{(2)} - \sum_r L_r^{(3)} + \dots = 1 + \sum_r \sum_r (-1)^q L_r^{(q)},$$

где $L_r^{(q)}$ – произведение передач r -й комбинации из q несоприкасающихся контуров.

Таким образом, из единицы вычитается сумма передач всех контуров плюс сумма передач всех парных комбинаций несоприкасающихся контуров минус сумма произведений передач всех комбинаций по три несоприкасающихся контура и т. д. Отсюда, в частности, следует, что определитель графа, не имеющего контуров, равен единице. Величина D_s равна определителю той части графа, которая не соприкасается с S -м путем.

Пример 2.1. Определить передачу графа (рис. 2.17) от вершины x_1 до вершины x_6 . В графе можно выделить два элементарных пути от x_1 до x_6 : $\mu = (x_1, x_2, x_5, x_6)$ и $\mu = (x_1, x_2, x_4, x_5, x_6)$.

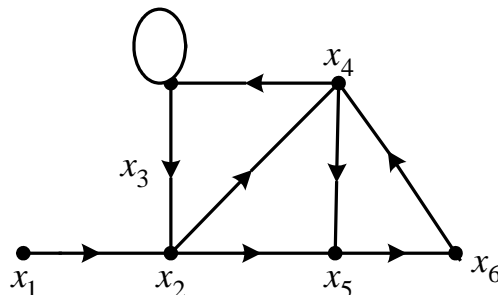


Рис. 2.17

Одним из контуров графа является петля при вершине x_3 с передачей $L_1 = K_{33}$, кроме того есть контур (x_2, x_4, x_3, x_2) с передачей $L_2 = K_{24} \cdot K_{43} \cdot K_{32}$, контур (x_4, x_5, x_6, x_4) с передачей $L_3 = K_{45} \cdot K_{56} \cdot K_{64}$ и контур $(x_2, x_5, x_6, x_4, x_3, x_2)$ с передачей $L_4 = K_{25} \cdot K_{56} \cdot K_{64} \cdot K_{43} \cdot K_{32}$. Среди указанных контуров есть только одна пара несоприкасающихся, это контуры с передачами L_1 и L_3 . $D = 1 - (L_1 + L_2 + L_3 + L_4) + L_1 \cdot L_3$, $D_1 = 1 - L_1$, $D_2 = 1 - L_1$, так как и первый, и второй пути не соприкасаются только с контуром L_1 . В результате

$$K_{16} = \frac{P_1(1 - L_1) + P_2(1 - L_1)}{1 - (L_1 + L_2 + L_3 + L_4) + L_1 \cdot L_3}$$

Правило Мэзона целесообразно использовать при определении передаточных функций сложных многоконтурных автоматических систем. Передаточная функция характеризует свойства системы в целом и определяется структурной схемой. Структурную схему автоматической системы можно рассматривать как один из видов сигнального графа. При этом вершинами считают точки приложения воздействий как внешних, так и внутренних, а дугами заменяют соединения звеньев, входящих в систему. Передача дуги определяется передаточной функцией звена.

Пример 2.2. На рис. 2.18 изображена структурная схема автоматической системы. Определить передаточную функцию системы между входом и выходом.

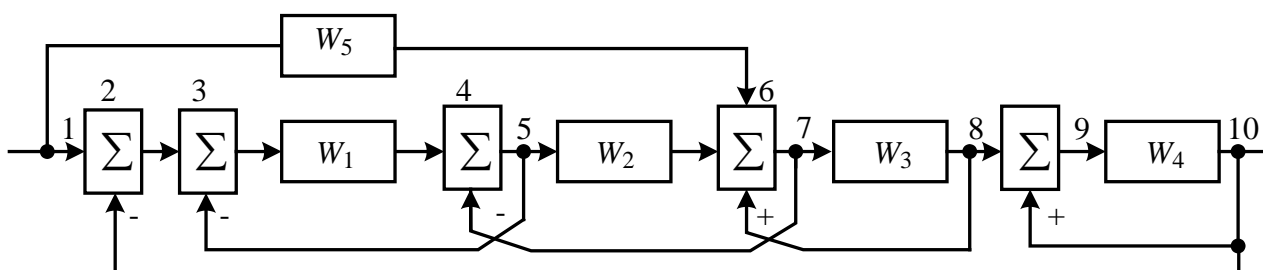


Рис. 2.18

Точки приложения воздействий на структурной схеме обозначены цифрами от 1 до 10, они соответствуют вершинам графа, изображенного на рис. 2.19.

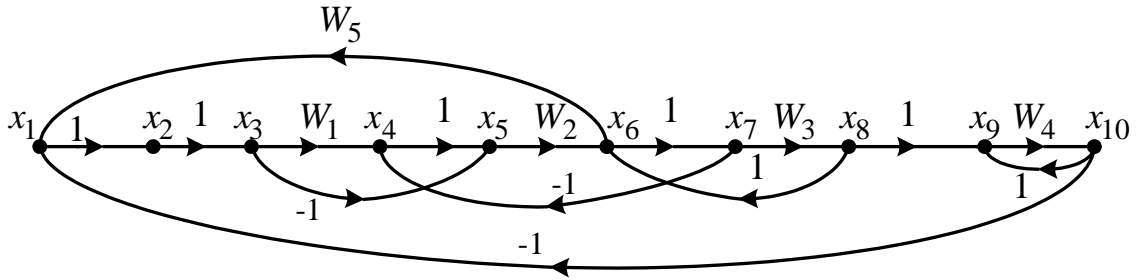


Рис. 2.19

Дугам графа, не обозначенным W , присваиваются вес, равный единице. При определении передаточных функций контуров нужно учитывать знак обратной связи (+ или -). Существует два пути из x_1 в x_{10} с передачами $P_1 = W_1 W_2 W_3 W_4$ и $P_2 = W_5 W_3 W_4$. Выпишем передачи всех контуров графа: $L_1 = -W_1$, $L_2 = -W_2$, $L_3 = W_3$, $L_4 = W_4$, $L_5 = -W_1 W_2 W_3 W_4$. Попарно не соприкасаются контуры L_1 и L_3 , L_1 и L_4 , L_2 и L_4 , L_3 и L_4 . В соответствии с этим передача графа между вершинами x_1 и x_{10} определяется выражением

$$K_{1,10} = \frac{P_1 D_1 + P_2 D_2}{D} = \frac{P_1 \cdot 1 + P_2 (1 - L_1)}{1 - (L_1 + L_2 + L_3 + L_4) + L_1 L_3 + L_1 L_4 + L_2 L_4 + L_3 L_4 - L_1 L_3 L_4}.$$

Подставив сюда значения P_i и L_j , получим передаточную функцию исходной системы:

$$W = \frac{W_3 W_4 [W_1 W_2 + W_5 (1 + W_1)]}{1 + W_1 + W_2 - W_3 - W_4 + W_1 W_2 W_3 W_4 - W_1 W_3 - W_1 W_4 - W_2 W_4 + W_3 W_4 + W_1 W_3 W_4}.$$

2.6. Операции над графами

Теоретико-множественные свойства графов определяют операции объединения, пересечения, дополнения до насыщенного графа и разности графов.

Пусть $G_1 = (X_1, \Gamma_1)$ и $G_2 = (X_2, \Gamma_2)$ – произвольные подграфы некоторого графа. Граф $G = (X, \Gamma)$ называется **объединением графов** G_1 и G_2 и обозначается $G = G_1 \cup G_2$, если $X = X_1 \cup X_2$ и $\Gamma = \Gamma_{1x} \cup \Gamma_{2x}$. Граф $G = (X, \Gamma)$ называется **пересечением** графов G_1 и G_2 и обозначается $G = G_1 \cap G_2$, если $X = X_1 \cap X_2$ и $\Gamma = \Gamma_{1x} \cap \Gamma_{2x}$.

Пример 2.3. Заданы два графа: $G_1: X_1 = \{x_1, x_2, x_3\}$, $\Gamma_{1x_1} = \{x_2\}$, $\Gamma_{1x_2} = \Gamma_{1x_3} = \{x_1, x_2\}$ и $G_2: X_2 = \{x_1, x_2\}$, $\Gamma_{2x_1} = \{x_1\}$, $\Gamma_{2x_2} = \{x_1, x_2\}$. Найти объединение и пересечение графов G_1 и G_2 .

Графы G_1 и G_2 приведены на рис. 2.20, а. Найдем объединение графов $G_1(X_1, \Gamma_1)$ и $G_2(X_2, \Gamma_2)$: $X = X_1 \cup X_2 = \{x_1, x_2, x_3\}$, $\Gamma_{x_1} = \Gamma_{1x_1} \cup \Gamma_{2x_1} = \{x_1, x_2\}$, $\Gamma_{x_2} = \Gamma_{1x_2} \cup \Gamma_{2x_2} = \{x_1, x_2\}$, $\Gamma_{x_3} = \Gamma_{1x_3} \cup \Gamma_{2x_3} = \{x_1, x_2\}$. Найти пересечение графов $G_1(X_1, \Gamma_1)$ и $G_2(X_2, \Gamma_2)$: $X = X_1 \cap X_2 = \{x_1, x_2\}$, $\Gamma_{x_1} = \Gamma_{1x_1} \cap \Gamma_{2x_1} = \emptyset$, $\Gamma_{x_2} = \Gamma_{1x_2} \cap \Gamma_{2x_2} = \{x_1, x_2\}$. Геометрическая интерпретация результатов объединения и пересечения показана на рис. 2.20, б и в соответственно.

Насыщенным называется граф, матрица смежности которого содержит только единичные элементы. Это значит, что в насыщенном графе в каждой вершине есть петля и каждые две вершины связаны дугами. **Дополнением по отображению** графа G до насыщенного графа G_x называется граф $\bar{G} = (X, \bar{\Gamma})$, у которого $\bar{\Gamma}_x = X \setminus \Gamma_x$. Дополнение по отображению можно определять не только до насыщенного графа, но и до любого графа, в который включается данный граф.

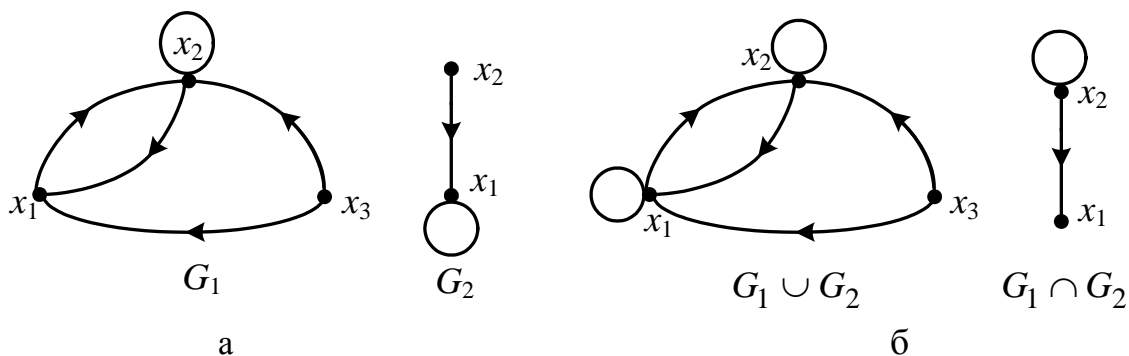


Рис. 2.20

Разностью графов $G_1(X_1, \Gamma_1)$ и $G_2(X_2, \Gamma_2)$ называется граф $G(X, \Gamma) = G_1 \setminus G_2 = G_1 \cap \bar{G}_2$, где \bar{G}_2 – дополнение по отображению графа G_2 до насыщенного.

Пример 2.4. Пусть G_1 и G_2 – графы, заданные в примере 2.3. Найти разность графов $G = G_1 \setminus G_2$.

Найдем дополнение графа G_2 до насыщенного графа с тремя вершинами. $\bar{G}_2 = (X, \bar{\Gamma}_2)$, где $X = \{x_1, x_2, x_3\}$, $\bar{\Gamma}_{2x_1} = X \setminus \Gamma_{2x_1} = \{x_2, x_3\}$, $\bar{\Gamma}_{2x_2} = X \setminus \Gamma_{2x_2} = \{x_3\}$, $\bar{\Gamma}_{2x_3} = X \setminus \Gamma_{2x_3} = \{x_1, x_2, x_3\}$.

Граф G_2 изображен на рис. 2.21, а. Граф $G = G_1 \setminus G_2 = G_1 \cap \bar{G}_2$ изображен на рис. 2.21, б.

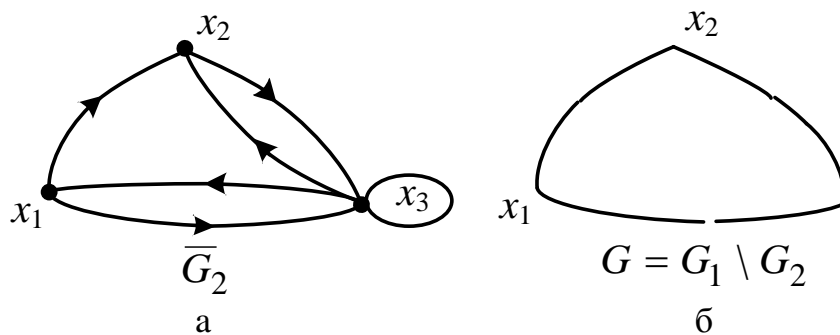


Рис. 2.21

2.7. Задача о кратчайшем пути связного неориентированного графа

Дан неориентированный граф $G(X, \Gamma)$, каждому ребру $v = (x_i, x_j)$ которого приписано некоторое число $l(v) \geq 0$, называемое длиной ребра. При этом любая цепь μ будет характеризоваться длиной $l(\mu) = \sum_{v \in V} l(v)$. Требуется среди всех возможных путей между произвольными вершинами x_i и x_j найти такой, чтобы его полная длина была наименьшей.

Рассмотрим алгоритм определения кратчайшего пути, использующий индексацию вершин графа.

1. Каждая вершина x_k рассматриваемого графа должна быть помечена индексом λ_k . Конечной вершине x_j первоначально присваивается индекс $\lambda_j = 0$.

2. На следующем шаге двигаемся от конечной вершины по инцидентным ребрам в сторону начальной вершины x_i . Вторым вершинам этих ребер присваивается индекс λ_j , численно равный длине ребра $l(x_j, x_i)$ от конечной точки до данной.

3. От каждой оцифрованной вершины двигаемся по инцидентным ребрам в сторону начальной вершины и вторым их концам присваиваем индексы, численно равные сумме величины индекса предыдущей вершины и длины данного ребра графа: $\lambda_n = \lambda_1 + l(x_1, x_n)$. К каждой вершине можно подойти несколькими путями, поэтому осуществляется процесс замены индексов, т. е. для каждой вершины следует найти такой индекс, который был бы наименьшим. Процедура продолжается, пока не будет оцифрована начальная вершина x_i . Индекс λ_i начальной вершины будет равен длине кратчайшего пути.

4. Кратчайший путь будет проходить через вершины x_k, x_n , начиная с x_i , разность индексов которых $\lambda_k - \lambda_n$ численно равна длине ребра.

В качестве примера рассмотрим граф, приведенный на рис. 2.22.

Ребрам графа приписаны числа, которые в условных единицах соответствуют их длине. Нужно найти кратчайший путь между вершинами A и B . Индексы вершин записаны в кружочках. Кратчайший путь выделен на рисунке жирной линией.

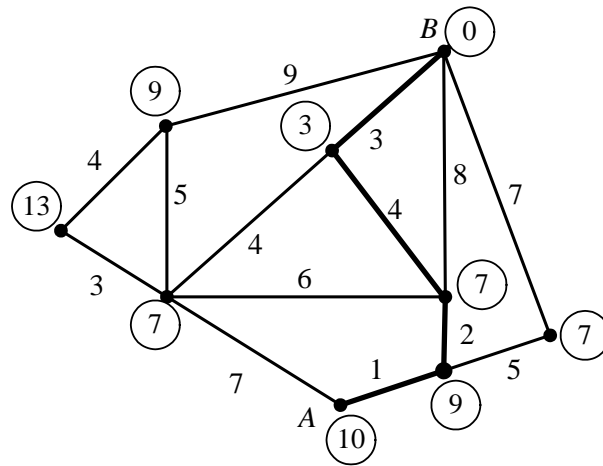


Рис. 2.22

2.8. Деревья. Символ дерева

Деревом называется конечный связный неориентированный граф, не имеющий циклов. Деревья обладают следующими свойствами:

- 1) любые две вершины дерева связаны простой цепью;
- 2) дерево с n вершинами имеет $n-1$ ребер;
- 3) число N различных деревьев, которые можно построить на множестве n вершин, определяется как $N=n^{n-1}$.

При $n = 10$ имеем 10^8 различных деревьев, но из них только 106 неизоморфны. Неизоморфные деревья считаются существенно различными. Примеры деревьев приведены на рис. 2.23.

Любое дерево G с n вершинами можно описать упорядоченной последовательностью $n-1$ номеров вершин $\alpha(G) = (\alpha_1, \alpha_2, \dots, \alpha_{n-1})$, которая называется **символом дерева**. Запись символа для данного дерева осуществляется следующим образом:

1. Осуществляется нумерация вершин дерева.
2. Выбирается концевая (или висющаяся) вершина с наименьшим номером и в последовательность $\alpha(G)$ записывается номер α_1 смежной с ней вершины, а сама концевая вершина удаляется из последовательности вместе с ребром.
3. Процесс повторяется, причем через $n-2$ шагов от дерева остается одно ребро, положение которого определяется парой номеров вершин.

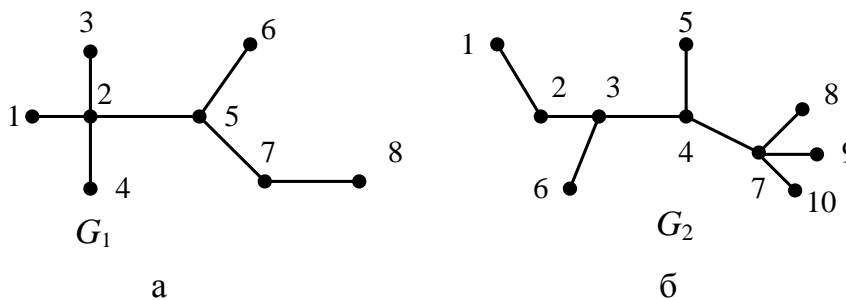


Рис. 2.23

Для дерева G_1 , изображенного на рис. 2.23, а, $\alpha(G_1) = (2, 2, 2, 5, 5, 7)$, а для дерева G_2 на рис. 2.23, б, $\alpha(G_2) = (2, 3, 3, 4, 4, 7, 7, 7)$.

Если известен символ дерева, то **построение дерева по его символу** осуществляется в следующей последовательности:

1. Записывается множество номеров вершин дерева $N = \{1, 2, \dots, n\}$.
2. Из множества N выбирается наименьший номер α_{\min} , который отсутствует в $\alpha(G)$ и строится ребро $(\alpha_{\min}, \alpha_1)$.
3. Из множества N удаляется номер α_{\min} , а из множества $\alpha(G)$ удаляется номер α_1 , и процесс продолжается до исчерпывания символа $\alpha(G)$. В последовательности N остается пара вершин, которая определяет последнее ребро дерева.

2.9. Покрывающее дерево связного графа. Экстремальное дерево

Покрывающим деревом связного графа называется любое дерево, связывающее все его вершины и имеющее в качестве ветвей ребра этого графа.

На рис. 2.24 приведен связный граф и одно из его покрывающих деревьев. Число различных покрывающих деревьев графа без петель определяется теоремой Трента. В графе G без петель с n вершинами число L различных деревьев равно минору любого из элементов главной диагонали квадратной матрицы B порядка n .

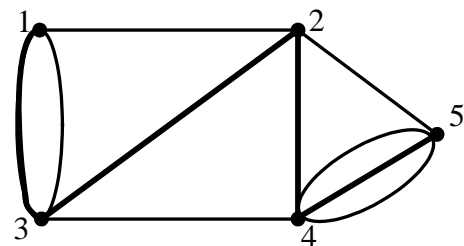


Рис. 2.24

Матрица B строится следующим образом: на главной диагонали записываются степени вершин графа, а ij -й и ji -й элементы равны взятому со знаком минус числу ребер, связывающих вершины i и j . Для графа на рис. 2.24 матрица B имеет вид

$$B = \begin{bmatrix} 3 & -1 & -2 & 0 & 0 \\ -1 & 4 & -1 & -1 & -1 \\ -2 & -1 & 4 & -1 & 0 \\ 0 & -1 & -1 & 5 & -3 \\ 0 & -1 & 0 & -3 & 4 \end{bmatrix}, \quad L = \Delta_{22} = \begin{bmatrix} 3 & -2 & 0 & 0 \\ -2 & 4 & -1 & 0 \\ 0 & -1 & 5 & -3 \\ 0 & 0 & -3 & 4 \end{bmatrix} = 76.$$

В ряде случаев, например, при анализе цепей и систем возникает необходимость получить все покрывающие деревья графа. Рассмотрим один из алгоритмов решения этой задачи на примере.

Пример 2.5. Найти все покрывающие деревья графа на рис. 2.25.

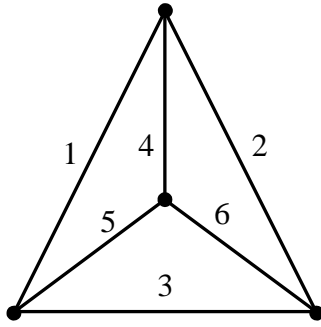


Рис. 2.25

Осуществляется нумерация ребер графа.

Записывается $n-1$ множеств номеров ребер, инцидентных $n-1$ вершинам графа (кроме одной из n вершин): $\Omega_1 = \{1,3,5\}$, $\Omega_2 = \{1,2,4\}$, $\Omega_3 = \{2,3,6\}$. Образуется таблица, столбцы которой представляют собой все возможные сочетания элементов множеств $\Omega_1, \Omega_2, \Omega_3$:

$$\begin{bmatrix} \Omega_1 \\ \Omega_2 \\ \Omega_3 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 3 & 3 & 3 & 3 & 3 & 5 & 5 & 5 & 5 & 5 & 5 & 5 & 5 \\ 2 & 2 & 4 & 4 & 4 & 1 & 1 & 2 & 4 & 4 & 1 & 1 & 1 & 2 & 2 & 4 & 4 & 4 \\ 3 & 6 & 2 & 3 & 6 & 2 & 6 & 6 & 2 & 6 & 2 & 3 & 6 & 3 & 6 & 2 & 3 & 6 \end{bmatrix}.$$

4. Столбцы, содержащие одинаковые ребра, попарно вычеркиваются независимо от их порядка (первый и шестой столбцы). Каждый столбец оставшейся таблицы соответствует ребрам одного из покрывающих деревьев.

Экстремальное дерево графа – это покрывающее дерево, связывающее его вершины наиболее экономичным образом (линии связи, автомобильные дороги и т. д.). Задача построения экстремального дерева формируется следующим образом. Каждому ребру $x_i x_j$ полного графа с n вершинами приписывается вес q_{ij} , численно выражающий длину, стоимость или другую величину, характеризующую любую пару вершин. Требуется построить экстремальное дерево, связывающее все вершины так, чтобы был минимальным суммарный вес всех ветвей дерева $\sum q_{ij} \rightarrow \min$. Способ построения экстремального дерева основан на последовательном введении в него ребер с приоритетом по минимуму их весов. Сначала для дерева выбирается ребро с наименьшим весом. Затем на каждом следующем шаге рассматривается минимальное по весу ребро и, если оно не образует цикла с ранее выбранными ветвями, вводится в дерево. Построение заканчивается после отбора для дерева $n-1$ ребер.

2.10. Корневые деревья. Код дерева

Любое дерево можно рассматривать как корневое, если одну из вершин выбрать в качестве корня, а остальные расположить на соответствующих уровнях (ярусах).

Пример корневого дерева приведен на рис. 2.26.

Корень располагается на нулевом уровне. Припишем каждой вершине некоторое число, называемое весом. Веса концевых вершин равны единице, вес корня дерева равен числу всех его вершин. Вес произвольной вершины равен общему числу вершин поддерева, корнем которого она является. Возле каждой вершины дерева на рис. 2.26 указан ее вес.

При таком представлении корневое дерево однозначно определяется упорядоченной последовательностью $\beta(G)$ весов его вершин, в которой на первом

месте стоит вес корня дерева, а затем следуют соответствующие последовательности для поддеревьев: $\beta(G) = (17, 1, 4, 1, 1, 1, 11, 4, 1, 2, 1, 6, 1, 1, 3, 1, 1)$.

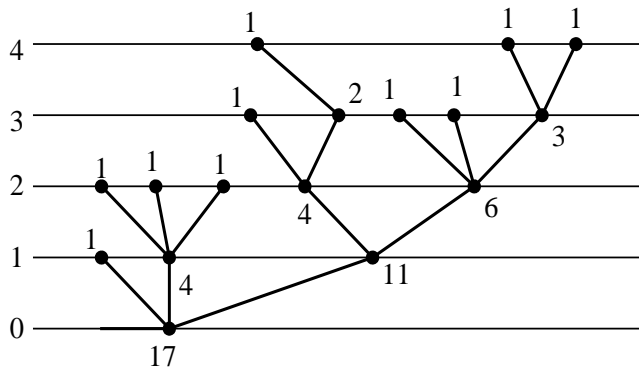


Рис. 2.26

Одна из стандартных процедур выбора корня состоит в следующем: из дерева удаляются все концевые вершины и ребра, затем в полученном дереве снова удаляются все вершины и ребра. В первом случае оставшаяся вершина выбирается в качестве корня и называется центром, во втором случае две вершины и связывающее их ребро образуют бицентр.

При этом за корень принимается та вершина, из которой вырастает дерево с меньшим числом вершин (если число вершин одинаково, то за корень принимается любая из вершин бицентра).

На рис. 2.27, а приведено дерево, в котором выделен бицентр, корневая форма этого дерева изображена на рис. 2.27, б.

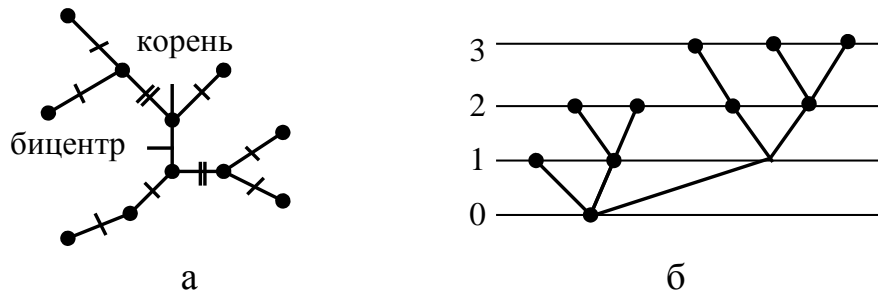


Рис. 2.27

Аналитически корневое дерево можно задать также с помощью кода $\gamma(G)$, представляющего собой последовательность 0 и 1, записанную в определенном порядке. Осуществляется обход дерева по всем ребрам по одному разу в каждом из противоположных направлений. При движении по ребру от корневой вершины в последовательность $\gamma(G)$ записывается 0, а при обратном движении – 1. Обход начинается и заканчивается в корне дерева. Длина последовательности $\gamma(G)$ равна удвоенному количеству ребер дерева. Для графа, приведенного на рис. 2.27, б, $\gamma(G) = (01001011000110010111)$.

ТЕМА 3. ТРАНСПОРТНЫЕ СЕТИ

3.1. Основные понятия и определения

К анализу транспортных сетей сводятся многие задачи, возникающие при планировании поставок, распределении грузопотоков между потребителями и т. д.

Транспортной сетью называется ориентированный связный мультиграф без петель, в котором:

существует одна и только одна вершина x_0 , называемая входом сети, что

$$\Gamma_{x_0}^{-1} = \emptyset;$$

существует одна и только одна вершина x_n , называемая выходом сети, что

$$\Gamma_{x_n} = \emptyset;$$

каждой дуге (x_i, x_j) графа отнесено целое число C_{ij} называемое пропускной способностью дуги [1].

Очевидно, что вершина x_0 является источником, а вершина x_n – стоком. Величина C_{ij} характеризует максимальное количество груза, который соответствующая коммуникация может пропустить за единицу времени. В общем случае $C_{ij} \neq C_{ji}$. Если $C_{ij} = C_{ji}$, то соответствующие дуги могут быть заменены ребром. В силу этого рассматриваемые вопросы будут справедливы также для смешанных и неориентированных графов.

Поток φ_{ij} характеризует количество груза, перемещаемого в единицу времени из пункта i в пункт j . Считается, что $\varphi_{ij} = -\varphi_{ji}$. Из физических соображений возникают следующие ограничения на числа φ_{ij} :

1) поток по любому ребру не может быть больше его пропускной способности, т. е.

$$0 \leq \varphi_{ij} \leq C_{ij}; \quad (3.1)$$

2) в каждой вершине, не являющейся ни источником, ни стоком, количество прибывающего вещества должно быть равно количеству отправляющегося:

$$\sum_i \varphi_{ik} - \sum_i \varphi_{kj} = 0, \quad \forall k \neq 0, n; \quad (3.2)$$

3) поток, выходящий из источника x_0 , в точности равен потоку, входящему в сток x_n , т. е.

$$\sum_i \varphi_{0i} = \sum_j \varphi_{jn} = \varphi. \quad (3.3)$$

Величина φ называется величиной потока транспортной сети. Если поток по дуге (x_i, x_j) равен ее пропускной способности, т. е. $\varphi_{ij} = C_{ij}$, то такая дуга называется **насыщенной**. Путь из x_0 в x_n называется насыщенным, если

он содержит хотя бы одну насыщенную дугу. Поток, насыщающий все пути из x_0 в x_n , называется полным. Наибольший из полных потоков называется максимальным.

На рис. 3.1 приведен пример транспортной сети, в которой вершина x_0 является входом сети (источник), а вершина x_5 является выходом сети (сток). Числа, приписанные дугам и ребрам, означают их пропускные способности. Для исследования распределения потока по дугам сети вводят понятие **разреза транспортной сети**.

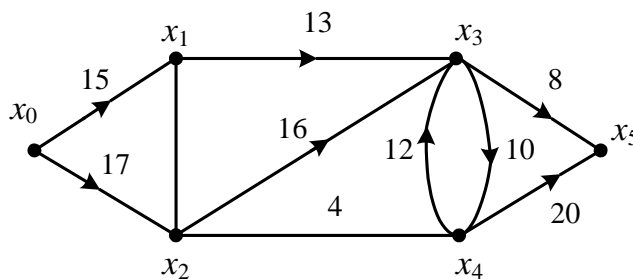


Рис. 3.1

Пусть все вершины транспортной сети разбиты на два непересекающихся подмножества R и \bar{R} ($R \cup \bar{R} = X$) так, что $x_0 \in R$, а $x_n \in \bar{R}$. Разрезом, отделяющим x_0 от x_n , называется совокупность всех дуг (x_i, x_j) , которые исходят из вершин $x_i \in R$, заканчиваются в вершинах $x_j \in \bar{R}$ и при удалении которых из сети блокируются все пути из источника в сток. Пропускная способность разреза (R, \bar{R}) численно равна сумме пропускных способностей дуг, его образующих, т. е. $C(R, \bar{R}) = \sum_{x_i \in R, x_j \in \bar{R}} C_{ij}$.

Любой путь из источника в сток содержит хотя бы одну дугу разреза (R, \bar{R}) , именно поэтому при удалении дуг какого-либо разреза, все пути из источника в сток прерываются. В качестве примера можно привести разрез транспортной сети (см. рис. 3.1): $(x_1x_3; x_2x_3; x_2x_4)$, при этом $x_0, x_1, x_2 \in R$, а $x_3, x_4, x_5 \in \bar{R}$. В каждой сети существует некоторое множество разрезов.

3.2. Задача о максимальном потоке. Алгоритм Форда-Фалкерсона

При анализе транспортных сетей часто возникает задача определения максимального потока, который может пропустить данная сеть, а также задача распределения этого потока по дугам сети.

С математической точки зрения задача о максимальном потоке формулируется следующим образом: при заданной конфигурации сети и известной пропускной способности C_{ij} найти неотрицательные значения φ_{ij} , удовлетворяющие условиям (3.1)–(3.3) и максимизирующие функцию φ , т. е.

$$\varphi = \sum_i \varphi_{0i} = \sum_j \varphi_{jn} \rightarrow \max.$$

Алгоритм для нахождения максимального потока был предложен Фордом и Фалкерсоном и заключается в постепенном увеличении потока, пропускаемого по сети, до тех пор, пока он не станет наибольшим. Алгоритм основан на теореме Форда-Фалкерсона: в любой транспортной сети максимальный поток из источника x_0 в сток x_n , равен минимальной пропускной способности разреза, отделяющего x_0 от x_n .

Рассмотрим табличный алгоритм нахождения максимального потока в сети из x_0 в x_n на примере сети, изображенной на рис. 3.1. Алгоритм состоит из подготовительного этапа и конечного числа шагов, на каждом из которых происходит допустимое увеличение потока. На подготовительном этапе формируется матрица пропускных способностей дуг сети (табл. 3.1). Если пропускная способность дуги (x_i, x_j) больше нуля, а пропускная способность симметричной ей дуги (x_j, x_i) равна нулю, то в клетку (j, i) заносится элемент C_{ij} , а в клетку $(i, j) - (j, i)$ не заполняются.

Таблица 3.1

	X_0	$X_{1(0)}$	$X_{2(0)}$	$X_{3(1)}$	$X_{4(2)}$	$X_{5(3)}$
X_0		15 ⁻	17			
X_1	0 ⁺		9	13 ⁻		
X_2	0	9		16	4	
X_3		0 ⁺	0		10	8 ⁻
X_4			4	12		20
X_5				0 ⁺	0	

Первый шаг. 1. По табл. 3.1 находится какой-либо путь из x_0 в x_5 с положительной пропускной способностью. Просматривается первая строка и столбцы, содержащие элементы $C_{0i} > 0$, помечается цифрой 0.

При этом окажутся выделенными первые дуги возможных путей из x_0 в x_5 . В рассматриваемом примере это дуги (x_0, x_1) и (x_0, x_2) . Далее про-

сматриваются те строки, номера которых совпадают с номерами помеченных столбцов. В каждой такой i -й строке отыскиваются элементы $C_{ij} > 0$, находящиеся в непомеченных столбцах, и помечаются эти столбцы номером просматриваемой строки. Так выделяются вторые дуги путей из x_0 в x_5 . Это дуги (x_1, x_3) и (x_2, x_4) . Просмотр строк продолжается до тех пор, пока не будет помечен столбец, соответствующий стоку x_5 . Это означает, что существует путь с положительной пропускной способностью из источника в сток. Столбец, соответствующий последней вершине x_5 , помечен номером (3). По нему находим предшествующую вершину x_3 , столбец которой помечен в свою очередь номером (1), приводящим к столбцу x_1 , помеченному номером (0). Достигли истока. В результате получен путь $l_1 = (x_0, x_1, x_3, x_5)$. Элементы этого пути C_{ij} помечаем знаком минус, а симметричные элементы C_{ji} – знаком плюс. Если при просмотре таблицы столбец, соответствующий стоку, не будет помечен, то путь x_0 в x_n будет отсутствовать.

2. Определяется пропускная способность найденного пути, которая равна наименьшей из пропускных способностей дуг:

$$C_1 = \min\{C_{01}^-, C_{13}^-, C_{35}^-\} = \min\{15, 13, 8\} = 8.$$

3. Определяются остаточные пропускные способности дуг найденного пути и симметричных ему дуг. Для этого из элементов C_{ij}^- табл. 3.1 вычитаем C_1 , а к элементам C_{ij}^+ прибавляем C_1 . В результате получим новую табл. 3.2 с измененными пропускными способностями.

Второй шаг.

1. Помечаем столбцы табл. 3.2, находим второй путь $l_2 = (x_0, x_2, x_4, x_5)$ и расставляем знаки. Элементы C_{02}, C_{24}, C_{54} помечают знаком минус, а симметричные им элементы C_{20}, C_{42}, C_{54} – знаком плюс.

2. Пропускная способность пути l_2 $C_2 = \min\{C_{02}^-, C_{24}^-, C_{54}^-\} = \min\{17, 4, 20\} = 4$. Изменим пропускные способности помеченных дуг на C_2 (табл. 3.3).

Таблица 3.2

	X_0	$X_{1(0)}$	$X_{2(0)}$	$X_{3(1)}$	$X_{4(2)}$	$X_{5(4)}$
X_0		7	17 ⁻			
X_1	8		9	5		
X_2	0 ⁺	9		16	4 ⁻	
X_3		8	0		10	0
X_4			4 ⁺	12		20 ⁻
X_5				8	0 ⁺	

Таблица 3.3

	X_0	$X_{1(0)}$	$X_{2(0)}$	$X_{3(1)}$	$X_{4(3)}$	$X_{5(4)}$
X_0		7 ⁻	13			
X_1	8 ⁺		9	5 ⁻		
X_2	4	9		16	0	
X_3		8 ⁺	0		10 ⁻	0
X_4			8	12 ⁺		16 ⁻
X_5				8	4 ⁺	

Третий шаг.

1. Пометив столбцы, находим $l_3 = (x_0, x_1, x_3, x_4, x_5)$.

2. Величина потока по пути l_3 $C_3 = \min\{7, 5, 10, 16\} = 5$.

Вычислив новые пропускные способности дуг, приходим к табл. 3.4.

Четвертый шаг.

1. Находим путь $l_4 = (x_0, x_2, x_3, x_4, x_5)$.

2. Величина потока по пути l_4 $C_4 = \min\{13, 16, 5, 11\} = 5$.

Вычислив новые пропускные способности дуг, приходим к табл. 3.5.

Таблица 3.4

	X_0	$X_{1(0)}$	$X_{2(0)}$	$X_{3(2)}$	$X_{4(3)}$	$X_{5(4)}$
X_0		2	13 ⁻			
X_1	13		9	0		
X_2	4 ⁺	9		16 ⁻	0	
X_3		13	0 ⁺		5 ⁻	0
X_4			8	17 ⁺		11 ⁻
X_5				8	9 ⁺	

Таблица 3.5

	X_0	$X_{1(0)}$	$X_{2(0)}$	$X_{3(2)}$	X_4	X_5
X_0		2	8			
X_1	13		9	0		
X_2	9	9		11	0	
X_3		13	5		0	0
X_4			8	22		6
X_5				8	14	

Пятый шаг. 1. Просматривая строки и помечая столбцы, убеждаемся в том, что столбцы x_4 и x_5 пометить невозможно. Следовательно, больше не существует ни одного пути с положительной пропускной способностью из вершины x_0 в вершину x_5 . Подмножество R образуют помеченные вершины x_0, x_1, x_2, x_3 (см. табл.3.5), а подмножество \bar{R} в данном случае – непомеченные вершины x_4 и x_5 . Разрез с минимальной пропускной способностью образуют дуги, начальные вершины которых принадлежат подмножеству R , а конечные – \bar{R} . Таким образом, разрез с минимальной пропускной способностью $(\bar{R}, R^*) = \{(x_3, x_5)(x_3, x_4)(x_2, x_4)\}$. Удалив дуги этого разреза, блокируем все пути из источника в сток. Пропускная способность разреза

$$C(R^*, \bar{R}^*) = C_{35}, C_{34}, C_{24} = 8 + 10 + 4 = 22.$$

Заключительный шаг. Вычитая из элементов табл.3.1 соответствующие элементы табл. 3.5, получим табл. 3.6.

Таблица 3.6

	X_0	$X_{1(0)}$	$X_{2(0)}$	$X_{3(2)}$	X_4	X_5
X_0		13	9			
X_1	-13		0	13		
X_2	-9	0		5	4	
X_3		-13	-5		10	8
X_4			-4	-10		14
X_5				-8	-14	

Положительные элементы этой таблицы характеризуют величины дуговых потоков. Следовательно, $\varphi_{01} = 13, \varphi_{02} = 9, \varphi_{13} = 13, \varphi_{23} = 5, \varphi_{24} = 4, \varphi_{34} = 10, \varphi_{35} = 8, \varphi_{45} = 14$, а потоки всех остальных дуг равны нулю.

Величина максимального потока равна сумме элементов x_0 -й строки табл. 3.6 или сумме элементов x_5 -го столбца. Все дуги разреза (R^*, \bar{R}^*) являются насыщенными.

3.3. Транспортная задача

Транспортная задача – это задача отыскания наиболее экономичного распределения потока по дугам транспортной сети. Решение транспортной задачи позволяет определить такое распределение маршрутов, которое обеспечивает минимальную стоимость перевозок или доставку грузов к потребителю в кратчайшее время [1,4].

3.3.1. Транспортная задача по критерию стоимости формулируется следующим образом. Задана транспортная сеть, по которой должен быть пропущен поток φ , каждой дуге которой поставлены в соответствие две величины: $C_{ij} = C(x_i, x_j)$ – пропускная способность дуги (x_i, x_j) ; $d_{ij} = d(x_i, x_j)$ – стоимость доставки единицы потока по дуге (x_i, x_j) . Требуется найти такое распределение

потока φ по дугам сети, которое обеспечивает минимальную стоимость прохождения потока.

Если рассматривать величины d_{ij} как длины соответствующих дуг, то задача минимизации стоимости прохождения потока φ по сети сведется к решению рассмотренной ранее задачи нахождения кратчайшего пути в графе от x_0 до x_n . При этом стоимость прохождения потока φ по найденному пути будет равна произведению длины этого пути на величину потока φ . Итак, если нет ограничений на пропускную способность дуг, кратчайший путь является путем, который обеспечивает минимальную стоимость прохождения потока.

При наличии ограничений на пропускные способности дуг нужно последовательно находить различные пути минимальной стоимости и пропускать потоки по ним до тех пор, пока суммарная величина потока по всем путям не будет равна заданной величине потока.

Рассмотрим алгоритм решения этой задачи при наличии ограничений на пропускные способности дуг.

Первый шаг. В исходной сети S_1 с длинами дуг d_{ij} устанавливается кратчайший путь l_1 от x_0 до x_n .

Второй шаг. Определяется пропускная способность пути l_1 $C_1 = \min C_{ij}$.

По этому пути пропускается поток $\varphi = \begin{cases} \varphi & \text{если } \varphi \leq C_1, \\ C_1 & \text{если } \varphi > C_1. \end{cases}$

Если $\varphi \leq C_1$, то задача решена и путь l_1 является наиболее экономичным для потока φ . Переходим к четвертому шагу.

Если $\varphi > C_1$, то поток рассматривается как частичный и осуществляется переход к третьему шагу.

Третий шаг. Переходим к сети S_2 , которая получается из S , путем замены пропускных способностей дуг C_{ij} на C'_{ij} , где $C'_{ij} = \begin{cases} C_{ij} - C_1 & \text{для } x \in l_1, \\ C_{ij} & \text{для } x \notin l_1. \end{cases}$

При этом дуги, у которых $C = 0$, исключаются из рассмотрения. Поток, распределение которого ищется в сети S_2 , принимается равным $\varphi' = \varphi - \varphi_1$. Теперь возникает первоначальная задача отыскания наиболее экономичного распределения потока φ' в сети S_2 . Возвращаемся к шагам 1 и 2.

Четвертый шаг. Стоимость перевозки грузов по транспортной сети определяется выражением $Z = \sum_{i=1}^m \sum_{j=1}^k d_{ij} \varphi_{ij}$.

Разобранный метод решения транспортной задачи дает способ нахождения величин частичных потоков, минимизирующих сумму Z .

Пример 3.1. [4]. В сети, представленной на рис. 3.2, найти такое распределение потока $\varphi = 3$ по дугам, которое обеспечивает минимальную стоимость прохождения потока. Первое число, приписанное каждому ребру, соответствует

пропускной способности c , а второе – стоимости d доставки единицы потока по ребру (φ) из одной вершины в другую в любом направлении.

Итерация 1.

Первый шаг. Применяя алгоритм нахождения кратчайшего пути, определяем, что минимальная стоимость доставки единицы потока из x_0 в x_5 , равная $d_1=4$, достигается сразу по двум путям $l_1 = (x_0, x_1, x_3, x_5)$ и $l_2 = (x_0, x_2, x_3, x_5)$.

Второй шаг. Определяем пропускные способности найденных путей $C_1 = \min\{2, 1, 1\} = 1$, $C_2 = \min\{1, 1, 1\} = 1$. Дуга (x_3, x_5) с пропускной способностью $C_{35} = 1$ является общей для обоих путей. Два потока она не пропустит. Выбираем первый путь. По этому пути протекает поток $\varphi_1 = 1$. Будем рассматривать его как частичный.

Третий шаг. Переходим к сети S_2 . Для этого определим пропускные способности дуг: $C'_{01} = C_{01} - C_1 = 2 - 1 = 1$; $C'_{13} = C_{13} - C_1 = 1 - 1 = 0$; $C'_{35} = C_{35} - C_1 = 1 - 1 = 0$. Все остальные $C'_{ij} = C_{ij}$, так как соответствующие им дуги не принадлежат рассматриваемому пути. Дуги C_{13} и C_{35} исключаем из рассмотрения (штриховые линии на рис. 3.2, б). Поток, распределение которого ищется в сети S_2 , $\varphi' = \varphi - \varphi_1 = 3 - 1 = 2$.

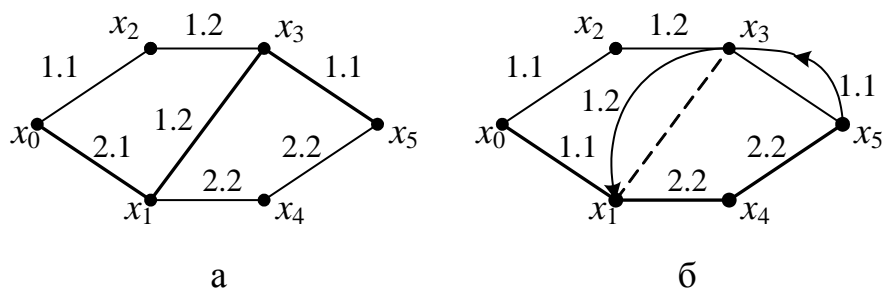


Рис. 3.2

Итерация 2.

Первый шаг. Находим кратчайший путь в сети S_2 . Это будет $l_2 = (x_0, x_1, x_4, x_5)$ с $d_2 = 5$.

Второй шаг. Пропускная способность этого пути $C_2 = \min\{1, 2, 2\} = 1$. По пути l_2 пропускается поток $\varphi_2 = 1$.

Третий шаг. Переходим к сети S_3 , $C'_{01} = C_{01} - C_2 = 1 - 1 = 0$; $C'_{14} = C_{14} - C_2 = 2 - 1 = 1$; $C'_{45} = C_{45} - C_1 = 2 - 1 = 1$. Все остальные $C'_{ij} = C_{ij}$. Останется пропустить поток $\varphi'' = \varphi' - \varphi_2 = 2 - 1 = 1$.

Итерация 3. Будем рассматривать путь $l_3 = (x_0, x_2, x_3, x_1, x_4, x_5)$, $d = 5$. $C_3 = \min\{1, 1, 1, 1, 1\} = 1$. Осталось пропустить поток, который также равен единице. Возвращаемся к исходному графу и суммируем частичные потоки. По дугам (x_0, x_1) , (x_1, x_4) и (x_4, x_5) два раза пропускали потоки, равные единице. Поток по дугам (x_3, x_1) и (x_1, x_3) компенсировался.

На рис. 3.3 в кружочках показано распределение потока по дугам сети.

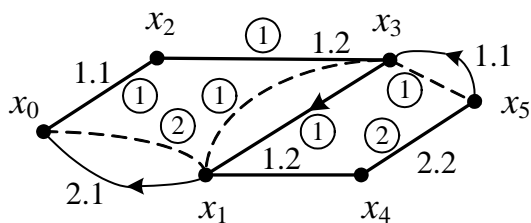


Рис. 3.3

Минимальная стоимость перевозки грузов по сети такого вида определится по формуле

$$Z = \sum d_{ij}\varphi_{ij} = d_{01}\varphi_{01} + d_{02}\varphi_{02} + d_{14}\varphi_{14} + d_{23}\varphi_{23} + d_{35}\varphi_{35} + d_{45}\varphi_{45} = 14.$$

Рассмотрим наиболее часто встречающийся вариант транспортной задачи

по критерию стоимости [1]. Однородный груз имеется в пунктах x_1, x_2, \dots, x_n в количествах a_1, a_2, \dots, a_m . Его требуется доставить в пункты y_1, y_2, \dots, y_r в количествах b_1, b_2, \dots, b_e . Предполагается, что общее количество требуемого груза равно имеющимся запасам:

Стоимость перевозки единицы потока из пункта x_i в пункт y_j равна d_{ij} . Требуется найти наиболее экономичные маршруты перевозки грузов.

Транспортная сеть, соответствующая этой задаче, строится следующим образом. Вход x_0 соединяется с каждой из вершин x_i дугой с пропускной способностью $C_{01} = a_i$. Каждая из вершин y_j соединяется с выходом x_n дугой с пропускной способностью $C_{jn} = b_j$. Стоимость прохождения потока по введенным дугам (x_0, x_i) и (y_j, x_n) считается равной нулю. Далее каждая вершина x_i соединяется с каждой вершиной y_j дугой с бесконечной пропускной способностью, стоимость прохождения единицы потока по которой равна d_{ij} .

Пример 3.2. [1]. Найти наиболее экономичные маршруты для транспортной задачи, заданной табл. 3.7.

Таблица 3.7

b_j	y_1	y_2	y_3	y_4
a_i	5	10	20	15
x_1	10	8	3	5
x_2	15	4	1	6
x_3	25	1	9	4

Таблица соответствует схеме дорог, связывающих три завода-изготовителя (x_1, x_2, x_3) некоторой продукции с заводами (y_1-y_4), потребляющими эту продукцию. В клетках таблицы записаны стоимости соответствующих перевозок d_{ij} .

Транспортная сеть, соответствующая данным табл. 3.7, приведена на рис. 3.4.

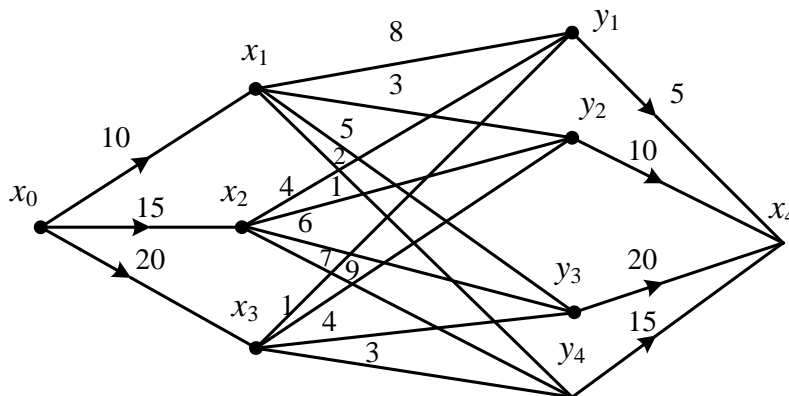


Рис. 3.4

Применяя описанный выше метод, найдем путь из источника x_0 в сток x_4 минимальной стоимости. Это будет путь $x_0 x_2 y_2 x_4$, или путь от изготовителя x_2 к потребителю y_2 . По этому пути можно пропустить частичный поток, равный 10 единицам, стоимость перевозки по этому пути также равна 10. Данные этого шага и пяти последующих занесены в табл. 3.8.

Таблица 3.8

№ пп	Маршрут	Частичный поток φ_k $X_{2(0)}$	d_{ij}	Стоимость перевозки
1	$x_2 y_2$	10	1	10
2	$x_3 y_1$	5	1	5
3	$x_1 y_4$	10	2	20
4	$x_3 y_4$	5	3	15
5	$x_3 y_3$	15	4	60
6	$x_2 y_3$	5	6	30
Итого		50		140

3.3.2. Транспортная задача по критерию времени. При решении транспортной задачи по критерию времени ребрам сети ставятся в соответствие величины t_{ij} , трактуемые как время, необходимое на перевозку груза из пункта x_i в пункт x_j .

Очевидно, что время, необходимое на перевозку всех грузов из x_0 в x_n , будет опре-

деляться путем, имеющим наибольшую продолжительность прохождения потока, так как перевозка грузов по остальным путям закончится раньше. Следовательно, время T , необходимое на перевозку всех грузов, будет равно $T = \max t_{ij}$. Для решения задачи нужно распределить весь поток φ таким образом, чтобы длительность наиболее продолжительного пути была бы минимальной.

ТЕМА 4. СЕТИ ПЕТРИ

4.1. Особенности сетей Петри и области их применения

Теория сетей Петри зародилась в 1962 году. В настоящее время сети Петри являются распространенной моделью, позволяющей описывать структуру и взаимодействие параллельно действующих объектов и процессов.

Можно отметить следующие достоинства аппарата сетей Петри:

1. Сети Петри позволяют моделировать асинхронность и недетерминизм параллельных, независимых событий, параллелизм конвейерного типа, конфликтные ситуации между процессами.

2. Сети Петри позволяют описывать как типовые ситуации в дискретных подсистемах, так и общую динамику работы сложной асинхронной системы.

3. Сети Петри позволяют производить иерархическую детализацию программных и аппаратных подсистем модели, производить совместное отображение структуры управления и потоков данных.

Знание основ сетей Петри является обязательным для специалистов по системному анализу, вычислительным сетям и т. д.

4.2. Основные определения. Способы задания сетей Петри

Сеть Петри – это двудольный ориентированный мультиграф, все множество X вершин которого разбито на два класса так, что дуги соединяют вершины только из разных классов. Эти классы вершин образуются:

множеством позиций $P = \{p_1, p_2, \dots, p_n\}$, которые обозначаются кружочками \circ ;

множеством переходов $T = \{t_1, t_2, \dots, t_m\}$, которые обозначаются планками —, при этом $P \cup T = X$, $P \cap T = \emptyset$.

Как и любой граф, сеть Петри может быть задана графическим, аналитическим и матричным способами.

На рис. 4.1 приведено графическое представление сети Петри, в которой $P = \{p_1, p_2, p_3, p_4\}$ и $T = \{t_1, t_2, t_3, t_4\}$.

Дуги могут идти только из позиции в переход или из перехода в позицию. Если кратность всех дуг равна 1, то сети называются ординарными. Сети с кратными дугами называются обобщенными.

При моделировании отражаются два аспекта систем: события и условия. Возможность наступления событий обеспечивается выполнением определенных условий. *Условиям* соответствуют *позиции* сети Петри, а *событиям*, происходящим в системе, соответствуют *переходы* (см. рис. 4.2).

Для представления динамики функционирования позициям могут присваиваться фишки, которые изображаются точками внутри вершин-позиций.

Присвоение фишек позициям сети Петри называется маркировкой, или разметкой. Сети Петри функционируют, переходя от одной маркировки к другой.

Начальная маркировка сети обозначается вектором $\mu_0 = [\mu_1, \mu_2, \dots, \mu_n]$, составляющие $\mu_i = \mu(p_i)$ которого определяют для каждой позиции p_i количество фишек в этой позиции. Таким образом, составляющими вектора μ являются неотрицательные целые числа.

При аналитическом способе задания сеть Петри задается как $C = (P, T, F, H, \mu^0)$, где, кроме множеств позиций P и переходов T , задаются входная F и выходная H функции. Через $F(t_j)$ обозначается множество входных позиций, а через $H(t_j)$ – множество выходных позиций перехода t_j ; μ^0 – начальная маркировка сети.

Так, например, сеть, изображенная на рис. 4.1, может быть представлена как $P = \{p_1, p_2, p_3, p_4\}$, $T = \{t_1, t_2, t_3, t_4\}$,

$$\begin{aligned}
F(t_1) &= \{p_1\}, \\
F(t_2) &= \{p_3\}, \\
F(t_3) &= \{p_2\}, \\
F(t_4) &= \{p_3, p_4\},
\end{aligned}$$

$$\begin{aligned}
H(t_1) &= \{p_2\}, \\
H(t_2) &= \{p_1\}, \\
H(t_3) &= \{p_4\}, \\
H(t_4) &= \{p_1, p_2, p_2\}, \quad \mu^0 = [1021].
\end{aligned}$$

Здесь переход t_4 отображен в комплект позиций. Комплект – это обобщение понятия множества, в него могут быть включены многократно повторяющиеся элементы.

Матричная форма определения сети Петри эквивалентна аналитическому способу задания $C = (P, T, D^-, D^+, \mu^0)$. Здесь D^- и D^+ – матрицы входных и выходных инцидентий соответственно размером $m \times n$, где m – число переходов и n – число позиций.

Элемент d_{ij}^- матрицы D^- равен кратности дуг, входящих в i -й переход из j -й позиции.

Элемент d_{ij}^+ матрицы D^+ равен кратности дуг, выходящих из i -го перехода в j -ю позицию.

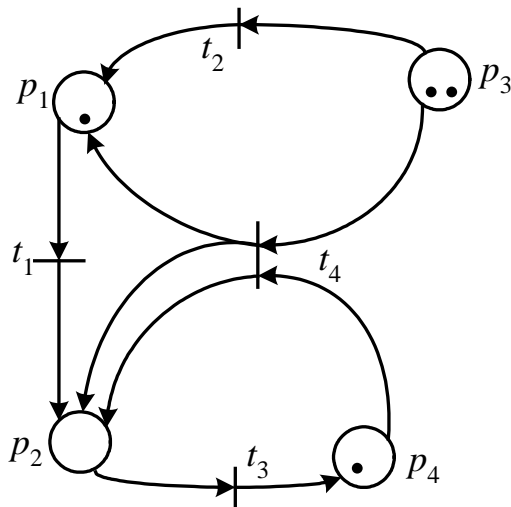


Рис. 4.1

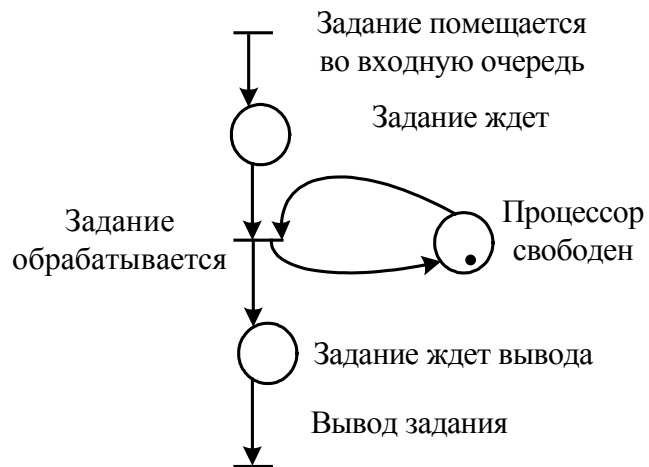


Рис. 4.2

Для рассматриваемой сети Петри

$$D^- = \begin{matrix} & p_1 & p_2 & p_3 & p_4 \\ \begin{matrix} t_1 \\ t_2 \\ t_3 \\ t_4 \end{matrix} & \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} \end{matrix},$$

$$D^+ = \begin{matrix} & p_1 & p_2 & p_3 & p_4 \\ \begin{matrix} t_1 \\ t_2 \\ t_3 \\ t_4 \end{matrix} & \begin{bmatrix} 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 1 & 2 & 0 & 0 \end{bmatrix} \end{matrix}.$$

Матрица $D = D^+ - D^-$ называется матрицей инцидентности сети Петри,

$$D = \begin{matrix} & p_1 & p_2 & p_3 & p_4 \\ \begin{matrix} t_1 \\ t_2 \\ t_3 \\ t_4 \end{matrix} & \begin{bmatrix} -1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 2 & -1 & -1 \end{bmatrix} \end{matrix},$$

в которой входящие в переход t_i дуги помеченные знаком « \leftarrow ».

4.3. Функционирование сетей Петри

Выполнение определенных условий связано с появлением меток в соответствующих этим условиям позициях. Последовательность событий, происходящих в моделируемой системе, отображается срабатыванием переходов. В результате срабатывания одного из переходов $t_i \in T$ сети происходит перераспределение фишек между позициями, и маркировка сети изменяется. Сеть Петри функционирует, переходя от одной маркировки к другой.

Необходимое условие срабатывания перехода t_i : каждая из его входных позиций должна иметь не меньше фишек, чем число дуг из этой позиции в переход, т. е. $\forall p_j \in F(t_i) : \mu(p_j) \geq 1$. Переход t_i , для которого выполняется данное условие, называется разрешенным.

В результате срабатывания перехода t_i из всякой входной позиции p_j перехода t_i удаляется столько фишек, сколько дуг ведет из p_j в t_i , а в каждую выходную позицию p_k помещается столько фишек, сколько дуг ведет из t_i в p_k .

При начальной маркировке $\mu_0 = [1021]$ сети Петри на рис. 4.1 разрешенными являются переходы t_1, t_2 и t_4 . Переход t_3 не разрешен, так как входящая в него дуга из позиции p_2 не подкреплена фишкой ($\mu_2 = 0$). Переходы t_1, t_2 и t_4 могут сработать в любом порядке, возникающий порядок появления событий не однозначен. Разрешенные переходы не влияют друг на друга.

Срабатывание перехода t_1 приводит к новой маркировке $\mu' = [0121]$, перехода t_2 – к маркировке $\mu' = [2011]$, перехода t_4 – к маркировке $\mu' = [2210]$.

Условие срабатывания перехода t_1 в имеющейся маркировке μ записывается следующим образом: $\mu^T \geq e^T[i] \cdot D^-$, $i = \overline{1, m}$,

где $e^T[i]$ – вектор-строка, компоненты которого соответствуют переходам и все равны нулю, за исключением i -й, которая равна единице.

Срабатывание перехода рассматривается как мгновенное событие, занимающее нулевое время.

Проверим условие срабатывания перехода t_4 при $\mu_0 = [1021]$:

$$[1021] \geq [0001] \cdot \begin{bmatrix} 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \end{bmatrix} = [0011].$$

Векторы сравниваются каждой из составляющих. Условие выполняется.

В результате срабатывания t_i возникает новая маркировка $\mu'^T = \mu^T + e^T[i] \cdot D$. При срабатывании перехода t_4 получим вместо μ_0 маркировку μ'^T

$$\mu'^T = [1021] + [0001] \cdot \begin{bmatrix} -1 & 1 & 0 & 0 \\ 1 & 0 & -1 & 0 \\ 0 & -1 & 0 & 1 \\ 1 & 2 & -1 & -1 \end{bmatrix} = [1021] + [12 -1 -1] = [2210].$$

После срабатывания перехода, имеющего несколько выходных позиций (рис. 4.3, а), все позиции получают метки, т.е. происходит **распараллеливание процесса**, и активизированные параллельные участки могут выполняться независимо.

Переход, имеющий несколько входных позиций (рис. 4.3, б), может быть реализован, если только все эти позиции содержат метки. Это соответствует выполнению всех соединяющихся в данном переходе параллельных участков процесса.

Сети Петри предусматривают также **конфликтные** состояния, когда необходимо запретить одновременное развитие нескольких процессов. На рис. 4.3, в переходы t_1 и t_2 находятся в конфликте: запуск одного из них удаляет фишку из p_i и тем самым запрещает другой.

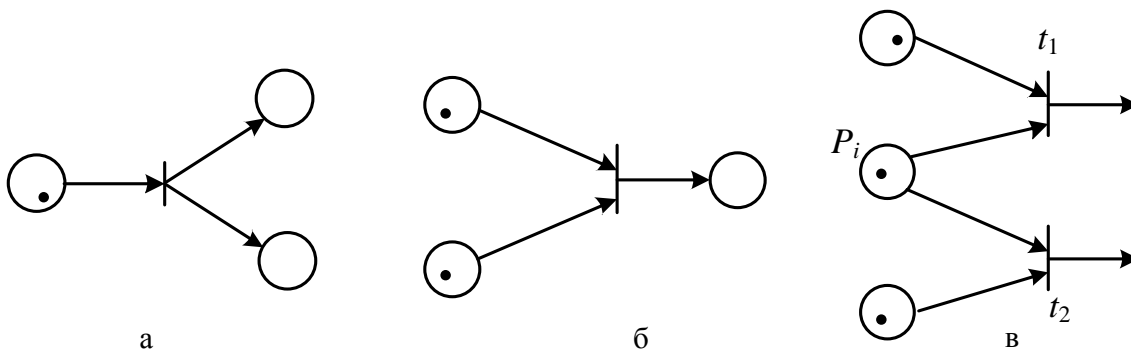


Рис. 4.3

Вопрос о разрешении конфликта, при котором одновременно несколько переходов готовы использовать метки одной и той же позиции, должен решаться дополнительными средствами.

Реализация переходов осуществляется до тех пор, пока существует хотя бы один разрешенный переход.

4.4. Свойства сетей Петри

Модель на основе сети Петри позволяет анализировать систему на наличие желательных или нежелательных свойств.

Ограниченность. Это свойство связано с введением ограничений на число меток в позициях.

Позиция p_i называется k -ограниченной, если количество фишек в ней не может превышать целого числа k , т. е. $\mu(p_i) \leq k$. Сеть Петри называется ограниченной, если все ее позиции ограничены. Ограниченную сеть Петри можно реализовать аппаратно, а неограниченную нельзя.

Безопасность. Позиция сети Петри называется безопасной, если число фишек в ней никогда не превышает единицы. Сеть Петри безопасна, если безопасны все ее позиции. Это свойство важно при интерпретации позиций как простых условий: если в позиции есть фишка, то условие выполняется, если нет, то не выполняется. Безопасную позицию можно реализовать одним триггером.

Сохраняемость. Сеть Петри $C = (P, T, F, H, \mu^0)$ называется строго сохраняющей, если сумма фишек по всем позициям остается строго постоянной в процессе выполнения сети, т. е. для всех возможных маркировок μ'

$$\sum_{i=1}^n \mu'(p_i) = \sum_{i=1}^n \mu_0(p_i).$$

Это свойство особенно присуще сетям, в которых фишки интерпретируются как ресурсы; фишки не должны ни создаваться, ни уничтожаться, следовательно, число входов в каждый переход должно равняться числу выходов.

Если распределение ресурсов моделируют только некоторые позиции, то вводится понятие взвешивания позиций. Вектор $W = (w_1, w_2, \dots, w_n)$ определяет вес $w_i \geq 0$ для каждой позиции p_i . Сеть Петри $C = (P, T, F, H, \mu^0)$ называется сохраняющей по отношению к вектору взвешивания W , если для всех возможных маркировок μ'

$$\sum_{i=1}^n w_i \mu'(p_i) = \sum_{i=1}^n w_i \mu_0(p_i),$$

т. е. скалярное произведение векторов W и μ постоянно.

Если $W = (0, 2, 1)$, то в условиях сохранения первая позиция не учитывается ($w_1 = 0$), число фишек во второй позиции умножается на 2 ($w_2 = 2$), а число фишек в третьей позиции берется неизменным ($w_3 = 1$).

Живость. Под живостью перехода t_i понимают принципиальную возможность его срабатывания при функционировании сети Петри. Тупик в сети Петри – это переход (или множество переходов), который в имеющейся маркировке μ' и в последующих достижимых из μ' маркировках не разрешен.

Достижимость. Свойство достижимости используется при установлении возможности возникновения некоторой ситуации в системе. Пусть проверяемая ситуация описывается разметкой μ' . Возникает задача: достижима ли

маркировка μ' из начальной маркировки μ^0 данной сети Петри. Задача достижимости является одной из наиболее важных задач анализа сетей Петри.

4.5. Анализ сетей Петри

Основная задача анализа сетей Петри – *задача достижимости*: достижима ли маркировка μ' из начальной маркировки μ_0 данной сети Петри.

Для ее решения имеется два основных подхода. Первый основан на построении дерева достижимости. Дерево достижимости – это ориентированное корневое дерево, вершинам которого, соответствуют возможные маркировки, а дугам – переходы. Начальная маркировка соответствует корню дерева. Из него исходят дуги, соответствующие разрешенным переходам. На каждом шаге строится очередной ярус дерева. Например, дерево достижимости для сети Петри, изображенной на рис. 4.4, после четырех шагов имеет вид, приведенный на рис. 4.5 (маркировки представлены векторами).

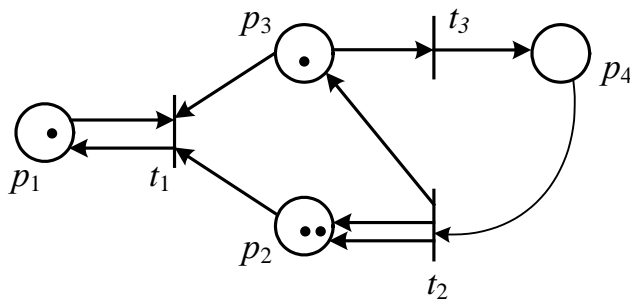


Рис. 4.4

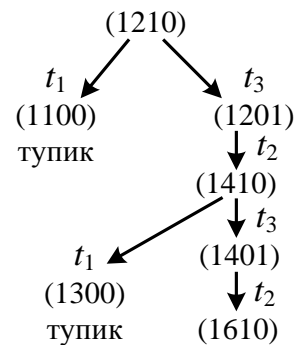


Рис. 4.5

Дерево представляет все возможные последовательности запусков переходов. Всякий путь, начинающийся в корне, соответствует допустимой последовательности переходов.

Другой подход к анализу сетей Петри называется матричным и основан на их матричном представлении. Ранее отмечалось, что переход t_i разрешен, если $\mu^T \geq e(i)D^-$, а результатом его срабатывания из маркировки μ_0 является

$$\mu'^T = \mu_0^T + e^T(i) \cdot D.$$

Пусть осуществляется последовательность срабатываний переходов $\sigma' = t_{i_1}, t_{i_2}, \dots, t_{i_k}$. В результате получится маркировка

$$\mu'^T = \mu_0^T + e^T(i_1) \cdot D + e^T(i_2) \cdot D + \dots + e^T(i_k) \cdot D = \mu_0^T + f^T(\sigma) \cdot D,$$

где $f^T(\sigma) = e^T(i_1) + \dots + e^T(i_k)$ – вектор целых положительных чисел, r -й элемент которого показывает сколько раз сработал переход t_{ir} в цепочке срабатываний, переводящей из μ_0 в μ' .

Для того чтобы существовала последовательность срабатываний σ , которая приводит из μ_0 в μ' , необходимо, чтобы вектор $f(\sigma)$ являлся неотрицательным целым решением матричного уравнения

$$\mu'^T = \mu_0^T + x^T D.$$

Если это уравнение не имеет решений, разметка μ' является недостижимой на данной сети из μ_0 . Пусть $x = f(\sigma) = (2, 3, 1)$, это означает, что переход t_1 сработал два раза, переход t_2 сработал три раза, а переход t_3 – один раз.

Недостатки матричного анализа в том, что вектор срабатывания $f(\sigma)$ не дает информации о порядке срабатывания переходов, кроме того, возможны недействительные решения уравнения, т. е. получаем некоторую последовательность переходов, которая не возможна.

Пример 4.1. Исследовать задачу достижимости для сети Петри, изображенной на рис. 4.4, с начальной маркировкой (1210) для маркировки (1610).

Матрицы входных инцидентностей D^- и выходных инцидентностей D^+ имеют вид

$$D^- = \begin{matrix} & p_1 & p_2 & p_3 & p_4 \\ t_1 & \begin{bmatrix} 1 & 1 & 1 & 0 \end{bmatrix} \\ t_2 & \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \\ t_3 & \begin{bmatrix} 0 & 0 & 1 & 0 \end{bmatrix} \end{matrix}, \quad D^+ = \begin{matrix} & p_1 & p_2 & p_3 & p_4 \\ t_1 & \begin{bmatrix} 1 & 0 & 0 & 0 \end{bmatrix} \\ t_2 & \begin{bmatrix} 0 & 2 & 1 & 0 \end{bmatrix} \\ t_3 & \begin{bmatrix} 0 & 0 & 0 & 1 \end{bmatrix} \end{matrix}.$$

Матрица $D = D^+ - D^-$ называется матрицей инцидентности сети Петри,

$$\text{Матрица инцидентности } D = D^+ - D^- = \begin{bmatrix} 0 & -1 & -1 & 0 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

Уравнение $\mu'^T = \mu_0^T + x^T D$ принимает вид

$$[1610] = [1210] + [x_1 x_2 x_3] = \begin{bmatrix} 0 & -1 & -1 & 0 \\ 0 & 2 & 1 & -1 \\ 0 & 0 & -1 & 1 \end{bmatrix}.$$

Перенесем μ_0 в левую часть и выполним умножение, тогда

$$[0400] = [0 \ (-x_1+2x_2) \ (-x_1+x_2-x_3) \ (-x_2+x_3)].$$

Приравняем составляющие векторов

$$\begin{cases} -x_1 + 2x_2 = 4, \\ -x_1 + x_2 - x_3 = 0, \\ -x_2 + x_3 = 0. \end{cases}$$

Система имеет решение $x_1 = 0; x_2 = 2; x_3 = 2$.

Это значит, что исследуемая маркировка достижима и в последовательности срабатываний переход t_2 срабатывает два раза, переход t_3 также два раза, а переход t_1 не срабатывает.

Другой важной задачей анализа сетей Петри является *задача сохранения*: является ли данная сеть Петри сохраняющей.

Если сеть Петри является сохраняющей по отношению к некоторому вектору взвешивания W , то

$$\sum_{i=1}^n \mu'(p_i)w_i = \sum_{i=1}^n \mu_0(p_i)w_i$$

или $\mu_0^T W = \mu'^T W$ для любой достижимой маркировки μ' . Так как $\mu'^T = \mu_0^T + f^T(\sigma)^* D$, то $\mu_0^T W = \mu_0^T W + f^T(\sigma) \cdot D \cdot W$. Следовательно, $f^T(\sigma) \cdot D \cdot W = 0$, что возможно только при $DW = 0$. Это уравнение позволяет найти ненулевой вектор взвешивания W , для которого взвешенная сумма по всем достижимым маркировкам постоянна.

Таким образом, сеть Петри является сохраняющей тогда и только тогда, когда существует такой положительный вектор W , что $D \cdot W = 0$.

4.6. Подклассы и расширения сетей Петри

Анализ сетей большой размерности является достаточно трудоемким. В связи с этим разработаны подклассы сетей Петри, в которых вводятся определенные ограничения на структуру сети, что позволяет использовать более простые алгоритмы для ее анализа.

К подклассу *автоматных графов* относят сети Петри, в которых каждый переход имеет одну входную и одну выходную позиции. Такие сети описывают последовательные процессы и как математическая модель эквивалентны конечным автоматам. В автоматных графах легко представить конфликтные ситуации, но нельзя моделировать создание и уничтожение фишек, необходимых для моделирования параллельных процессов или ожидания.

К подклассу *маркированных графов* относятся сети Петри, в которых каждая позиция имеет только один вход и один выход. Маркированные графы являются двойственными по отношению к автоматным графам. Они позволяют моделировать параллельность и синхронизацию, но не могут моделировать конфликты или принятие решений, зависящих от данных. Наиболее интересными структурными компонентами маркированных графов являются циклы. Пример маркированного графа приведен на рис. 4.6.

К подклассу *устойчивых сетей Петри* относятся сети, которые обладают следующим свойством: если при любой маркировке μ два любых перехода t_i и t_j оказываются разрешенными, то срабатывание одного из них не исключает возможности срабатывания другого перехода.

В теории сетей Петри предложены также несколько расширений, ориентированных на увеличение моделирующих возможностей сетей.

Временные сети Петри позволяют отразить в модели временные параметры системы. Если моделируемое событие имеет отличную от нуля длительность, как например, событие «задание обрабатывается», то оно представляется в виде двух мгновенных событий типа «начало события», «конец события» и условия «событие происходит» (рис. 4.7). Считается, что события происходят одновременно. Позиции во временных сетях взвешиваются временем выполнения.

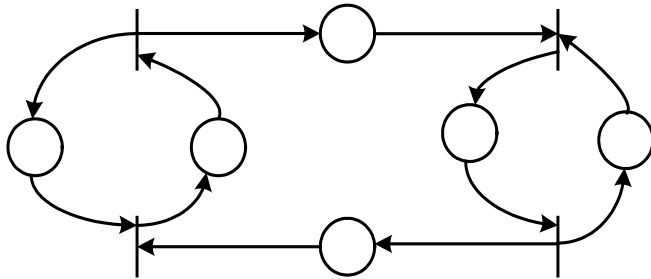


Рис. 4.6

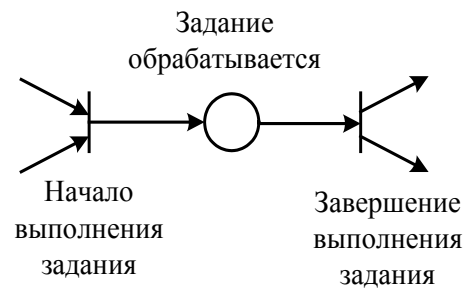


Рис. 4.7

Раскрашенные сети Петри характеризуются тем, что каждой фишке в позициях сети сопоставляется определенный признак (цвет). Это позволяет задавать различные типы условий, объектов или ресурсов, которые характеризуют состояние системы. Для срабатывания перехода t_i его входная позиция должна содержать метки определенного цвета, которым помечается дуга, направленная от позиции к переходу t_i . Раскрашенные сети Петри позволяют уменьшить размерность графа при моделировании сложных систем.

Приоритетные сети Петри и сети с проверкой на нуль позволяют учитывать приоритетность событий в модели. В сетях с проверкой на нуль вводится дополнительное множество дуг запрета (сдерживающих дуг). На графе сети Петри такие дуги на конце имеют не стрелку, а маленький кружок (рис. 4.8).

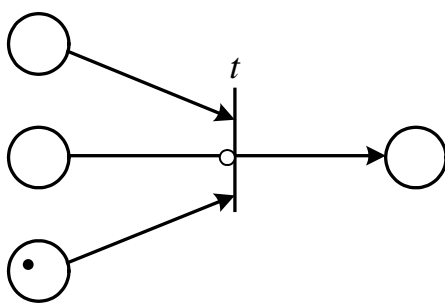


Рис. 4.8

Переход сети Петри с проверкой на нуль является разрешенным, если фишки присутствуют во всех его обычных входных позициях и отсутствуют в позициях запрета. В приоритетных сетях Петри вводится специальная функция приоритетности, которая задает соотношение приоритетов срабатывания для двух подмножеств переходов T_1 и T_2 , где $T_1, T_2 \subset T, T_1 \cap T_2 = \emptyset, T_1 \vee T_2 = T$.

E-сети, или оценочные сети – наиболее мощное расширение сетей Петри, являющееся средством описания моделей функционирования вычислительных систем. В E-сетях учитывается фактор времени, усложнена логика работы переходов, введены различные операции над метками.

ТЕМА 5. ЭЛЕМЕНТЫ МАТЕМАТИЧЕСКОЙ ЛОГИКИ И ТЕОРИИ АВТОМАТОВ

5.1. Основные понятия алгебры логики

Математический аппарат, базирующийся на алгебре логики, широко используется для описания функционирования, анализа и синтеза цифровых схем. Основным понятием алгебры логики является высказывание. **Высказыванием** называется всякое суждение (утверждение), которое либо истинно, либо ложно. Одновременно истинным и ложным высказывание быть не может. Истинность высказывания обозначается единицей, а ложность – нулем. **Простое высказывание** не зависит от значений других высказываний. Значение истинности **сложного высказывания** зависит от истинности других высказываний, составляющих его. Любое сложное высказывание можно считать логической функцией от простых высказываний (аргументов). Логическая функция, как и ее аргументы, принимает только два значения: единица или нуль.

Множество символов $X = \{x_1, x_2, \dots, x_n\}$, каждый из которых принимает значения единица или нуль, называется множеством переменных или аргументов. Функция $f(x_1, x_2, \dots, x_n)$, определенная на множестве всевозможных наборов аргументов из X и принимающая значения единица или нуль, называется функцией алгебры логики или булевой функцией. Областью определения булевой функции служит совокупность всевозможных n -мерных наборов из единиц и нулей.

Приняты три способа задания булевых функций:

1. Формула, указывающая в явном виде последовательность операций, производимых над переменными:

$$F = f(x_1, x_2, \dots, x_n).$$

2. Таблица истинности, в левой части которой перечисляются все возможные комбинации значений аргументов x_1, x_2, \dots, x_n , а в правой – значения функции. При n переменных число строк таблицы равно 2^n .

3. Логическая схема или условное графическое изображение логической функции.

Число различных функций алгебры логики, зависящих от n аргументов, конечно и равно 2^{2^n} . Значения функции могут быть заданы не на всех возможных наборах аргументов. Функции, значения которых на некоторых наборах не определены, называются **не полностью определенными**.

Функция $f(x_1, \dots, x_{i-1}, x_i, x_{i+1}, \dots, x_n)$ **существенно зависит** от аргумента x_i , если имеет место соотношение

$$f(x_1, \dots, x_{i-1}, 0, x_{i+1}, \dots, x_n) \neq f(x_1, \dots, x_{i-1}, 1, x_{i+1}, \dots, x_n).$$

В противном случае функция зависит от x_i несущественно и x_i является ее **фиктивным аргументом**. Функция не изменится, если к ее аргументам дописать любое число фиктивных аргументов или зачеркнуть те аргументы, которые для данной функции являются фиктивными.

Число всех функций алгебры логики A_n , существенно зависящих от n аргументов, определяется следующим рекуррентным соотношением:

$$A_n = 2^{2^n} - C_n^{n-1} A_{n-1} - C_n^{n-2} A_{n-2} - \dots - C_n^1 A_1 - A_0, \quad (5.1)$$

где A_i – число функций алгебры логики, существенно зависящих от i аргументов, C_n^m – число сочетаний из n элементов по m , $C_n^m = \frac{n!}{m!(n-m)!}$.

5.2. Элементарные булевы функции

Элементарные булевы функции образуются путем использования однородных связей между двоичными переменными. Рассмотрим одиннадцать элементарных функций, которые часто употребляются в алгебре логики и ее приложениях.

Имеется две функции, которые не зависят ни от одного аргумента ($n=0$ в выражении (5.1)). Это $f_1 = 0$ – **константа нуль** и $f_2 = 1$ – **константа единица**.

При $n = 1$ имеем две функции, существенно зависящие от одного аргумента x . Одна из них $f_3 = x$ называется **функцией прямой передачи сигнала**, другая $f_4 = \bar{x}$ (читается «не x »), называется **функцией отрицания или инверсии**.

Функции f_3, f_4 представлены табл. 5.1.

Таблица 5.1

x	$f_3(x)$	$f_4(x)$
0	0	1
1	1	0

Значения $f_3(x)$ совпадают со значением переменной x , а $f_4(x)$ принимает значения, противоположные значениям переменной x .

Устройства, реализующие элементарные булевы функции, называются логическими элементами. Их входы соответствуют булевым переменным, а выход – реализуемой функции. Для обозначения логических элементов используют упрощенные изображения в виде прямоугольников, внутри которых помещаются условные названия или символы соответствующей функции.

На рис. 5.1, а и б в виде логических элементов изображены функции $f_3(x)$ и $f_4(x)$, где инверсный выход обозначается кружком.

В соответствии с выражением (5.1) должно быть 10 функций, существенно зависящих от двух аргументов x_1 и x_2 .

К элементарным относят семь функций, представленных табл. истинности 5.2.

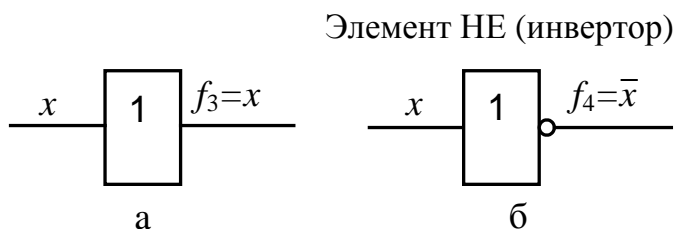


Рис. 5.1

Таблица 5.2

x_1	x_2	f_5	f_6	f_7	f_8	f_9	f_{10}	f_{11}
0	0	0	0	1	1	1	1	0
0	1	1	0	0	1	0	1	1
1	0	1	0	0	0	0	1	1
1	1	1	1	1	1	0	0	0

Функция $f_5(x_1, x_2) = x_1 \vee x_2$ называется **дизъюнкцией**, или логическим **сложением** x_1 и x_2 . Читается « x_1 или x_2 ». Функция f_3 принимает значение 1, если хотя бы одно из слагаемых x_1 или x_2 равно 1. Изображение функции дизъюнкции в виде логического элемента приведено на рис. 5.2, а, его называют элементом ИЛИ.

Функция $f_6(x_1, x_2) = x_1 \wedge x_2$ называется **конъюнкцией**, или **логическим умножением** x_1 и x_2 . Читается « x_1 и x_2 ». Функция f_6 принимает значение 1 только в том случае, если оба сомножителя x_1 и x_2 равны единице. Изображение функции конъюнкции в виде логического элемента приведено на рис. 5.2, б, его называют элементом И. В дальнейшем логическое умножение x_1 и x_2 будем записывать $x_1 x_2$.

Функции f_6, f_5, f_4 определяют логические операции конъюнкции \wedge , дизъюнкции \vee и отрицания \neg (И, ИЛИ, НЕ).

Функция $f_7(x_1, x_2) = x_1 \sim x_2$ называется **функцией эквивалентности**, или **функцией равнозначности**. Читается « x_1 эквивалентно x_2 ». Функция f_7 принимает значение 1, когда значения переменных x_1 и x_2 одинаковы, и значение 0 в противном случае. Изображение функции эквивалентности в виде логической схемы приведено на рис. 5.2, в.

Функция $f_8(x_1, x_2) = x_1 \rightarrow x_2$ называется **функцией импликации**. Читается «если x_1 , то x_2 ». Функция f_6 принимает значение 0 только в том случае, если $x_1 = 1$ и $x_2 = 0$, и значение 1 в противном случае. Изображение функции импликации приведено на рис. 5.2, г.

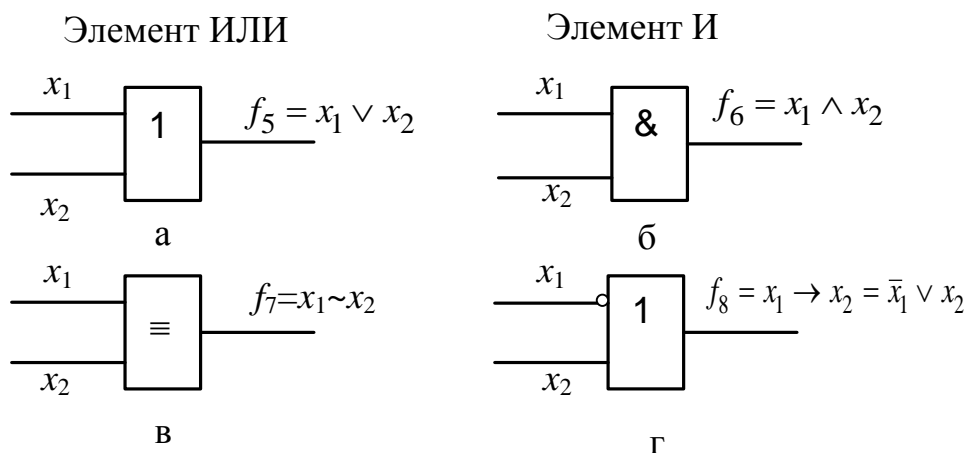


Рис. 5.2

Функция $f_9(x_1, x_2) = x_1 \downarrow x_2$ называется **функцией Вебба, или стрелкой Пирса**. Читается «ни x_1 ни x_2 ». Функция f_9 принимает значение 1 только в том случае, если $x_1 = 0$ и $x_2 = 0$, и значение 0 в противном случае.

Функция $f_{10}(x_1, x_2) = x_1 | x_2$ называется **функцией Шеффера**. Читается «неверно, что x_1 и x_2 ». Функция f_{10} принимает значение 0 только в том случае, если $x_1 = 1$ и $x_2 = 1$, и значение 1 в противном случае.

Функция $f_{11}(x_1, x_2) = x_1 \oplus x_2$ называется **функцией сложения по модулю 2**. Читается « x_1 неравнозначно x_2 ». Функция f_{11} принимает значение 1 только в том случае, если переменные x_1 и x_2 имеют различные значения, и значение 0 в противном случае.

Изображение функций f_9 , f_{10} и f_{11} в виде логических элементов приведено на рис. 5.3, а, б и в соответственно.

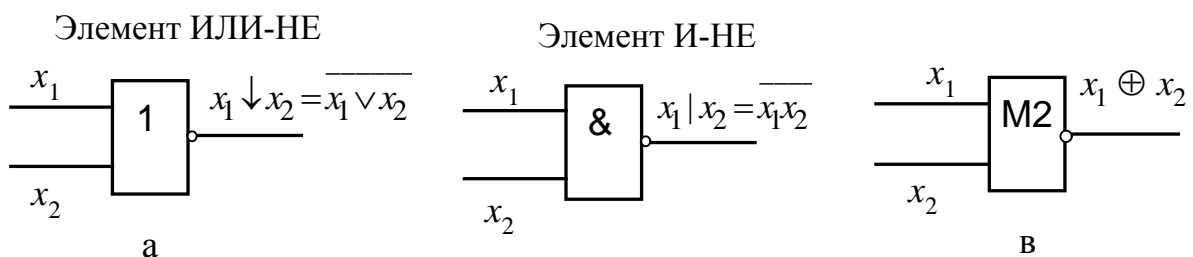


Рис. 5.3

5.3. Полнота системы булевых функций

Одно из основных понятий алгебры логики - понятие функциональной **полноты системы булевых функций**. Система булевых функций называется функционально полной, если она позволяет представить любую булеву функцию. Так, рассмотренные ранее функции от двух переменных $f_7 - f_{11}$ могут быть представлены с помощью трех функций: отрицания, дизъюнкции и конъюнкции.

$$\begin{aligned}
 f_7 &= x_1 \sim x_2 = (\bar{x}_1 \vee x_2)(x_1 \vee \bar{x}_2) = \bar{x}_1 \bar{x}_2 \vee x_1 x_2; \\
 f_8 &= x_1 \rightarrow x_2 = (\bar{x}_1 \vee x_2) = \overline{x_1 \bar{x}_2}; \\
 f_9 &= x_1 \downarrow x_2 = \overline{x_1 \vee x_2} = \bar{x}_1 \bar{x}_2; \\
 f_{10} &= x_1 | x_2 = \bar{x}_1 \vee \bar{x}_2 \vee \overline{x_1 x_2}; \\
 f_{11} &= x_1 \oplus x_2 = (\bar{x}_1 \vee x_2)(x_1 \vee \bar{x}_2) = x_1 \bar{x}_2 \vee \bar{x}_1 x_2.
 \end{aligned}$$

Полным набором служит также единственная функция – функция Шеффера x_1/x_2 . В этом легко убедиться, выразив через нее три выше упомянутых функции полного набора:

$$\begin{aligned}\bar{x} &= \overline{\overline{x}} = x/x; \\ x_1x_2 &= \overline{\overline{x_1x_2}} = \overline{x_1/x_2} = (x_1/x_2)/(x_1/x_2); \\ x_1 \vee x_2 &= \overline{\overline{x_1 \vee x_2}} = \overline{\overline{\bar{x}_1\bar{x}_2}} = \bar{x}_1/\bar{x}_2 = (x_1/x_1)/(x_2/x_2).\end{aligned}$$

Полный набор представляет собой также и функция Пирса.

Логические элементы, соответствующие функционально полным наборам булевых функций, образуют так называемый базис и позволяют построить любую сколь угодно сложную логическую схему. Наиболее распространенными являются базисы И-ИЛИ-НЕ, ИЛИ-НЕ, И-НЕ.

5.4. Законы и тождества алгебры логики

Законы алгебры логики устанавливают эквивалентность логических формул, образованных с помощью полного набора логических операций И, ИЛИ, НЕ. Приведем основные законы, определяющие эти операции:

1) коммутативности дизъюнкции и конъюнкции

$$x_1 \vee x_2 = x_2 \vee x_1, \quad x_1x_2 = x_2x_1;$$

2) ассоциативности дизъюнкции и конъюнкции

$$x_1 \vee (x_2 \vee x_3) = (x_1 \vee x_2) \vee x_3, \quad x_1(x_2x_3) = (x_1x_2)x_3;$$

3) дистрибутивности конъюнкции относительно дизъюнкции и дизъюнкции относительно конъюнкции

$$x_1(x_2 \vee x_3) = x_1x_2 \vee x_1x_3, \quad x_1 \vee x_2x_3 = (x_1 \vee x_2)(x_1 \vee x_3);$$

4) идемпотентности дизъюнкции и конъюнкции

$$x \vee x = x, \quad xx = x;$$

5) де Моргана

$$\overline{x_1 \vee x_2} = \bar{x}_1\bar{x}_2, \quad \overline{x_1x_2} = \bar{x}_1 \vee \bar{x}_2;$$

6) двойного отрицания

$$\overline{\bar{x}} = x;$$

7) склеивания

$$(x_1 \vee x_2)(x_1 \vee \bar{x}_2) = x_1, \quad x_1x_2 \vee x_1\bar{x}_2 = x_1;$$

8) поглощения

$$x_1 \vee x_1x_2 = x_1, \quad x_1(x_1 \vee x_2) = x_1;$$

9) действия с константами 0 и 1

$$\begin{aligned}x \vee 0 &= x, & x \cdot 0 &= 0, & x \vee 1 &= 1; \\ x \cdot 1 &= x, & x \vee \bar{x} &= 1, & x \cdot \bar{x} &= 0.\end{aligned}$$

Из последнего закона вытекают следующие правила.

Правило 1. Если логическая сумма двоичных переменных содержит хотя бы одну пару слагаемых, из которых одно есть некоторая переменная, а другое – ее отрицание, то она является тождественно истинной:

$$x_1 \vee \bar{x}_5 \vee x_4 \vee x_3 \vee \bar{x}_4 \equiv 1.$$

Правило 2. Если логическое произведение двоичных переменных содержит хотя бы одну пару сомножителей, из которых один есть некоторая переменная, а другой – ее отрицание, то оно является тождественно ложным

$$x_1 \bar{x}_2 \bar{x}_4 x_3 x_2 \equiv 0.$$

Запишем ряд тождеств, часто используемых для упрощения сложных логических функций:

$$x_1 \vee \bar{x}_1 x_2 = x_1 \vee x_2,$$

$$\bar{x}_1 \vee x_1 x_2 = \bar{x}_1 \vee x_2,$$

$$x_1 \vee \bar{x}_1 x_2 \vee \bar{x}_1 \bar{x}_2 x_3 \vee \bar{x}_1 \bar{x}_2 \bar{x}_3 x_4 \vee \dots = x_1 \vee x_2 \vee x_3 \vee x_4 + \dots,$$

$$x_1 (\bar{x}_1 \vee x_2) (\bar{x}_1 \vee \bar{x}_2 \vee x_3) (\bar{x}_1 \vee \bar{x}_2 \vee \bar{x}_3 \vee x_4) = x_1 x_2 x_3 x_4,$$

$$x_1 \vee x_2 \vee x_3 = x_1 \vee x_2 \vee x_3 \vee x_k \bar{x}_k.$$

Последнее соотношение показывает возможность дополнения исходной формулы фиктивными членами, совокупность которых является тождественно ложной.

Следует отметить, что законы де Моргана справедливы для любого числа переменных:

$$\overline{x_1 \vee x_2 \vee \dots \vee x_n} = \bar{x}_1 \bar{x}_2 \dots \bar{x}_n;$$

$$\overline{x_1 x_2 \dots x_n} = \bar{x}_1 \vee \bar{x}_2 \vee \dots \vee \bar{x}_n.$$

5.5. Представление булевых функций дизъюнктивными и конъюнктивными нормальными формами

Любая логическая функция может выражаться различными логическими формулами, являющимися эквивалентными. Наиболее удобными для практического использования являются нормальные формы представления сложных логических функций.

Элементарной конъюнкцией Q называется логическое произведение любого конечного числа переменных и их отрицаний, причем каждая переменная встречается только один раз. Число переменных, составляющих элементарную конъюнкцию, называется ее рангом. Так, выражение

$$Q = x_1 \bar{x}_2 x_3 \bar{x}_4 x_5$$

является элементарной конъюнкцией ранга 5.

Дизъюнктивной нормальной формой (ДНФ) называется дизъюнкция элементарных конъюнкций:

$$N = Q_1 \vee Q_2 \vee \dots \vee Q_k.$$

Любая булева функция может быть представлена в ДНФ, например,

$$f(x_1, x_2, x_3) = x_1 x_2 \bar{x}_3 \vee \bar{x}_1 x_3 \vee \bar{x}_2.$$

Элементарной дизъюнкцией D называется логическая сумма конечного числа переменных и их отрицаний, причем каждая переменная встречается в сумме один раз. Число переменных, составляющих элементарную дизъюнкцию, называется ее рангом. Так, выражение $D = x_1 \vee \bar{x}_2 \vee \bar{x}_3 \vee x_4$ является элементарной дизъюнкцией четвертого ранга.

Конъюнктивной нормальной формой (КНФ) называется конъюнкция элементарных дизъюнкций:

$$K = D_1 \wedge D_2 \wedge \dots \wedge D_n.$$

Любую булеву функцию можно представить в КНФ, например,

$$f(x_1, x_2, x_3) = (x_1 \vee \bar{x}_2 \vee x_3)(\bar{x}_1 \vee x_3)(x_2 \vee \bar{x}_3).$$

Одна и та же логическая функция путем эквивалентных преобразований может быть представлена различными ДНФ или КНФ. Единственность представления обеспечивают совершенные нормальные формы.

Совершенной ДНФ (СДНФ) логической функции $f(x_1, x_2, x_n)$ от n различных переменных называется ДНФ, которая содержит только конъюнкции ранга n и не содержит одинаковых конъюнкций.

Произвольная логическая функция $f(x_1, x_2, x_n)$ приводится к СДНФ в следующей последовательности: 1) функция f приводится к какой-либо ДНФ; 2) конъюнкции, не содержащие всех двоичных переменных, дополняются до конъюнкций n -го ранга; 3) из полученной ДНФ с конъюнкциями n -го ранга удаляются повторяющиеся друг друга конъюнкции.

Пример 5.1. Привести функцию $f(x_1, x_2, x_3) = \bar{x}_1 x_2 \vee x_1 \bar{x}_2 x_3 \vee \bar{x}_1 \bar{x}_3$ к СДНФ.

Дополним конъюнкции второго ранга до конъюнкций третьего ранга, используя закон склеивания:

$$\bar{x}_1 x_2 = \bar{x}_1 x_2 x_3 \vee \bar{x}_1 x_2 \bar{x}_3 \quad \text{и} \quad \bar{x}_1 \bar{x}_3 = \bar{x}_1 x_2 \bar{x}_3 \vee \bar{x}_1 \bar{x}_2 \bar{x}_3.$$

Просуммируем конъюнкции:

$$\begin{aligned} f(x_1, x_2, x_3) &= \bar{x}_1 x_2 x_3 \vee \bar{x}_1 x_2 \bar{x}_3 \vee x_1 \bar{x}_2 x_3 \vee \bar{x}_1 x_2 \bar{x}_3 \vee \bar{x}_1 \bar{x}_2 \bar{x}_3 = \\ &= \bar{x}_1 x_2 x_3 \vee \bar{x}_1 x_2 \bar{x}_3 \vee x_1 \bar{x}_2 x_3 \vee \bar{x}_1 \bar{x}_2 \bar{x}_3. \end{aligned}$$

Если логическая функция задана таблицей истинности, то построение СДНФ осуществляется по следующему алгоритму: 1) выбираются наборы аргументов, на которых функция обращается в единицу; 2) выписываются конъюнкции, соответствующие этим наборам, причем если аргумент x_i входит в набор как единица, то в конъюнкцию он вписывается без изменения. Если же аргумент x_i входит в данный набор как нуль, то в соответствующую конъюнкцию вписывается его отрицание; 3) все выписанные конъюнкции соединяют знаком дизъюнкции. Элементарные конъюнкции СДНФ называют *конституэнтами единицы*.

Пример 5.2. Построить СДНФ для функции $f(x_1, x_2, x_3)$, заданной табл. 5.3.

Таблица 5.3

X_1	X_2	X_3	$f(x_1, x_2, x_3)$
0	0	0	0
0	0	1	1
0	1	0	1
0	1	1	0
1	0	0	1
1	0	1	1
1	1	0	0
1	1	1	1

Функция f принимает значение единица пять раз, поэтому ее СДНФ представляет собой логическую сумму пяти элементарных конъюнкций третьего ранга

$$f(x_1, x_2, x_3) = \bar{x}_1 \bar{x}_2 x_3 \vee \bar{x}_1 x_2 \bar{x}_3 \vee x_1 \bar{x}_2 \bar{x}_3 \vee x_1 \bar{x}_2 x_3 \vee x_1 x_2 x_3.$$

Совершенной КНФ (СКНФ) логической функции $f(x_1, x_2, \dots, x_n)$ от n различных переменных называется КНФ, которая содержит только дизъюнкции ранга n и не содержит одинаковых дизъюнкций. Построение СКНФ по

таблично заданной функции осуществляется в следующей последовательности 1) выбираются наборы аргументов, на которых функция обращается в нуль; 2) выписываются дизъюнкции, соответствующие этим наборам, причем если аргумент x_i входит в набор как нуль, то в дизъюнкцию он вписывается без изменения. Если же аргумент x_i входит в данный набор как единица, то в соответствующую дизъюнкцию вписывается его отрицание; 3) все выписанные дизъюнкции соединяют знаком конъюнкции.

Элементарные дизъюнкции СКНФ называют *конституэнтами нуля*.

Пример 5.3. Построить СКНФ для функции $f(x_1, x_2, x_3)$, заданной табл. 5.3.

Функция f принимает значение нуль три раза, поэтому ее СКНФ представляет собой логическое произведение трех элементарных дизъюнкций третьего ранга $f(x_1, x_2, x_3) = (x_1 \vee x_2 \vee x_3)(x_1 \vee \bar{x}_2 \vee \bar{x}_3)(\bar{x}_1 \vee \bar{x}_2 \vee x_3)$.

5.6. Минимизация функций алгебры логики

Минимизация функций алгебры логики (ФАЛ) является одним из основных этапов анализа и синтеза цифровых устройств. Основной целью миними-

зации логических функций является получение их минимальных дизъюнктивных или конъюнктивных форм.

ДНФ (КНФ) функции $f(x_1, x_2, \dots, x_n)$ называется минимальной, если она содержит наименьшее число переменных x_i по сравнению со всеми другими эквивалентными ДНФ (КНФ).

Существуют различные аналитические и табличные методы минимизации, рассмотрим некоторые из них.

5.6.1. Метод непосредственных преобразований. Сущность метода непосредственных преобразований заключается в том, что минимизация исходной ФАЛ осуществляется путем применения основных законов и тождеств алгебры логики.

Сокращенной ДНФ называется форма представления ФАЛ, которая получается из СНДФ путем склеивания вначале конъюнкты единицы между собой по всем переменным, а затем конъюнкций ранга $n-1, n-2$ и т. д.

Простая импликанта – это конъюнкция, которая не склеивается ни с какой другой конъюнкцией, входящей в данную ФАЛ. Используя понятие импликанты, сокращенную ДНФ можно определить как дизъюнкцию простых импликант.

Пример 5.4. Минимизировать функцию, заданную в СНДФ.

$$f(x_1, x_2, x_3) = \overset{\textcircled{1}}{\bar{x}_1 \bar{x}_2 \bar{x}_3} \vee \overset{\textcircled{2}}{x_1 \bar{x}_2 x_3} \vee \overset{\textcircled{3}}{\bar{x}_1 x_2 x_3} \vee \overset{\textcircled{4}}{x_1 \bar{x}_2 \bar{x}_3} \vee \overset{\textcircled{5}}{\bar{x}_1 x_2 \bar{x}_3} \vee \overset{\textcircled{6}}{x_1 x_2 \bar{x}_3}.$$

Используем законы склеивания и поглощения. При этом учтем, что одно и то же слагаемое СНДФ может склеиваться с несколькими другими. Первое слагаемое $\bar{x}_1 \bar{x}_2 \bar{x}_3$ склеивается с пятым $\bar{x}_1 x_2 \bar{x}_3$ слагаемым.

Запишем это следующим образом:

$$1 \text{ и } 5: \bar{x}_1 \bar{x}_2 \bar{x}_3 \vee \bar{x}_1 x_2 \bar{x}_3 = \bar{x}_1 \bar{x}_2 \quad (\text{по } x_3).$$

Кроме того, склеиваются следующие пары:

$$2 \text{ и } 4: x_1 \bar{x}_2 x_3 \vee x_1 \bar{x}_2 \bar{x}_3 = x_1 \bar{x}_2 \quad (\text{по } x_3).$$

$$4 \text{ и } 6: x_1 \bar{x}_2 \bar{x}_3 \vee x_1 x_2 \bar{x}_3 = x_1 \bar{x}_2 \quad (\text{по } x_3).$$

$$3 \text{ и } 5: \bar{x}_1 x_2 x_3 \vee \bar{x}_1 x_2 \bar{x}_3 = \bar{x}_1 x_2 \quad (\text{по } x_3).$$

В результате получим сумму четырех конъюнкций второго ранга

$$f = \overset{\textcircled{1}}{\bar{x}_1 \bar{x}_3} \vee \overset{\textcircled{2}}{x_1 \bar{x}_2} \vee \overset{\textcircled{3}}{x_1 \bar{x}_3} \vee \overset{\textcircled{4}}{\bar{x}_1 x_2},$$

в которой склеиваются первое и третье слагаемые:

$$1 \text{ и } 3: \bar{x}_1 \bar{x}_3 \vee x_1 \bar{x}_3 = \bar{x}_3 \quad (\text{по } x_1).$$

В результате минимальная ДНФ имеет вид

$$f(x_1, x_2, x_3) = \bar{x}_3 \vee x_1 \bar{x}_2 \vee \bar{x}_1 x_2.$$

5.6.2. Метод карт Карно. Логическая функция, записанная в СНДФ, может быть представлена в виде специальных таблиц, известных под названием карт Карно или диаграмм Вейча. Каждая клетка таблицы соответствует одному из наборов таблицы истинности. Клетки карты обозначаются таким образом, что любой соседней паре клеток соответствуют склеивающиеся слагаемые. Для логической функции двух переменных карта Карно изображается в виде горизонтального ряда из четырех клеток (рис. 5.4, а). Пусть $f(x_1, x_2, x_3) = \bar{x}_1x_2 \vee x_1x_2 \vee x_1\bar{x}_2$.

В каждую клетку записывается значение функции единица или нуль, на соответствующем этой клетке наборе переменных.

Единицы в клетках карты Карно объединяются в группы и обводятся контуром. Любая пара единиц, расположенных в соседних клетках, выражается одной переменной, той, которая присутствует в каждом из наборов, объединенных в группу. Одна и та же клетка может входить в несколько групп. На карте Карно рис. 5.4, а отмечены две группы единиц, соответствующие склеивающимся слагаемым: $\bar{x}_1x_2 \vee x_1x_2 = x_2$ и $x_1x_2 \vee x_1\bar{x}_2 = x_1$, в результате минимизации $f(x_1x_2) = x_1 \vee x_2$. Карта Карно для функции трех переменных содержит восемь клеток (совпадает с числом строк таблицы истинности равным 2^3) и приведена на рис. 5.4, б. Ее следует рассматривать не как плоскостную, а как свернутую в трубку (в виде цилиндра) соединением первого и последнего столбца. При этом соседними оказываются клетки на противоположных границах карты.

Для минимизации образуются группы из двух или четырех единиц, расположенных в соседних клетках. Две единицы, расположенные в соседних клетках, выражаются двумя переменными, а четыре единицы – одной переменной, той, которая присутствует во всех наборах, объединенных в группу.

На рис. 5.4, б отмечены 3 группы единиц, соответствующих склеивающимся слагаемым.

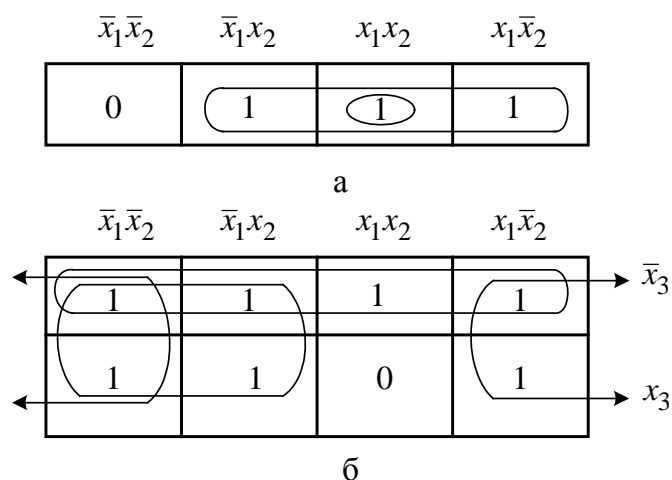


Рис. 5.4

Минимальная ДНФ соответствующей функции имеет вид

$$f(x_1, x_2, x_3) = \bar{x}_3 \vee \bar{x}_2 \vee \bar{x}_1.$$

Следует помнить, что количество единиц, объединяемых в группу, должно быть целой степенью двойки, т. е. может быть равно 1,2,4,8,... и т. д. Контур должен быть прямоугольным или квадратным. Каждый контур должен включать как можно больше единиц, а общее число контуров должно быть как можно меньше. Все единицы карты должны быть охвачены контурами.

Карта Карно логической функции четырех переменных приведена на рис. 5.5. Она содержит $2^4 = 16$ клеток.

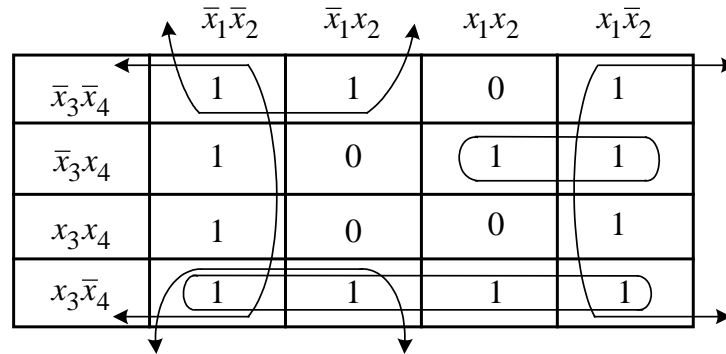


Рис. 5.5

Каждой клетке соответствует один из наборов аргументов и в нем записано соответствующее значение функции (один или ноль):

$$f = \bar{x}_2 \vee \bar{x}_4\bar{x}_1 \vee x_3\bar{x}_4 \vee x_1\bar{x}_3x_4.$$

Здесь сохраняются предыдущие правила склеивания, но добавляется склеивание по тороиду, т. е. первую и последнюю колонку диаграммы, а также верхнюю и нижнюю строки следует считать соседними.

На этой диаграмме одной переменной соответствует восемь единиц, расположенных в соседних клетках, произведению, включающему две переменные – четыре соседних единицы; произведению трех переменных – две и произведению четырех переменных – одна единица. Одна и та же клетка может входить в несколько групп.

При минимизации функции пяти переменных пользуются картой из 32 клеток, изображенной на рис. 5.6.

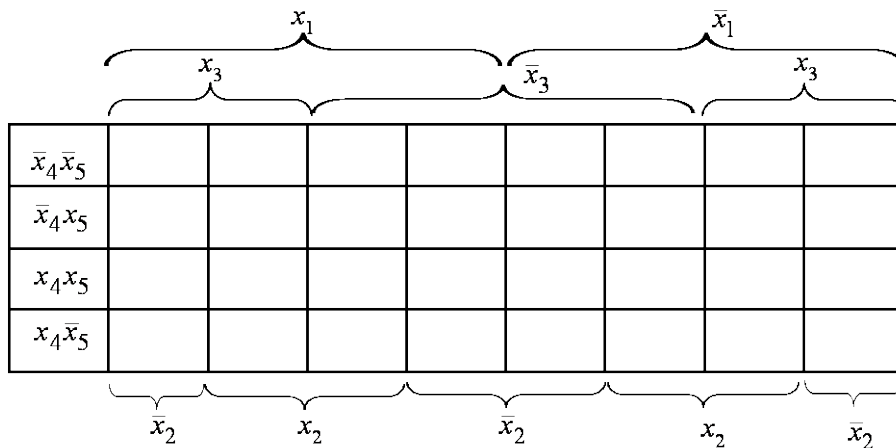


Рис. 5.6

На этой диаграмме одной переменной соответствует 16 единиц, расположенных в смежных клетках, произведению двух переменных – восемь, единиц, произведению трех переменных – четыре, произведению четырех переменных – две единицы. При работе с этой таблицей следует помнить, что для переменных x_2, \bar{x}_2 и x_3 «соседние» клетки оказываются разнесенными.

Рассмотрим способ минимизации функции пяти переменных с помощью карты Карно из 16 клеток.

Идея минимизации основывается на использовании представления функции пяти переменных в следующем виде:

$$F(x_1, x_2, x_3, x_4, x_5) = x_5 F(x_1, x_2, x_3, x_4, 1) \vee \bar{x}_5 F(x_1, x_2, x_3, x_4, 0).$$

Отсюда вытекает, что функция пяти аргументов при задании любых значений x_1, x_2, x_3, x_4 может быть представлена одним из четырех значений: 0, 1, x_5, \bar{x}_5 .

В клетке записывается:

$$0 \text{ при } f(x_1, x_2, x_3, x_4, 1) = f(x_1, x_2, x_3, x_4, 0) = 0,$$

$$1 \text{ при } f(x_1, x_2, x_3, x_4, 1) = f(x_1, x_2, x_3, x_4, 0) = 1,$$

$$\bar{x}_5 \text{ при } f(x_1, x_2, x_3, x_4, 1) = \bar{f}(x_1, x_2, x_3, x_4, 0) = \bar{x}_5,$$

$$x_5 \text{ при } f(x_1, x_2, x_3, x_4, 1) = \bar{f}(x_1, x_2, x_3, x_4, 0) = x_5.$$

При комбинации клеток следует помнить правила:

1) для комбинаций единичных клеток правила объединения те же, что и для функций четырех переменных, при этом в результате склеивания в соответствующем слагаемом будем отсутствовать пятая переменная;

2) комбинации клеток, содержащих только x_5 (или только \bar{x}_5), соответствует слагаемое минимизированной формы, которое получается в случае функции четырех переменных с единичными значениями в тех же клетках, умноженное на x_5 (или на \bar{x}_5);

3) комбинации клеток, содержащих единицы и x_5 (или \bar{x}_5), соответствуют два слагаемых минимизированной формы: одно из них получается при рассмотрении клеток, содержащих единицы согласно правилу 1), второе получается с помощью правила 2), примененного ко всей комбинации клеток при условии мысленной замены единиц на x_5 (или \bar{x}_5) во всех единичных клетках рассматриваемой комбинации.

Пример 5.5. Минимизировать функцию

$$f(x_1, x_2, x_3, x_4, x_5) = \bar{x}_1 x_2 \bar{x}_3 \bar{x}_4 \bar{x}_5 \vee x_1 x_2 \bar{x}_3 \bar{x}_4 \bar{x}_5 \vee \bar{x}_1 \bar{x}_2 x_3 \bar{x}_4 \bar{x}_5 \vee x_1 \bar{x}_2 x_3 \bar{x}_4 \bar{x}_5 \vee x_1 \bar{x}_2 x_3 \bar{x}_4 x_5 \vee x_1 x_2 \bar{x}_3 x_4 \bar{x}_5 \vee x_1 x_2 x_3 x_4 x_5 \vee \bar{x}_1 \bar{x}_2 x_3 \bar{x}_4 x_5 \vee x_1 x_2 \bar{x}_3 x_4 x_5.$$

Сгруппируем слагаемые, содержащие x_5 и \bar{x}_5 , тогда

$$f(x) = x_5 (x_1 \bar{x}_2 x_3 \bar{x}_4 \vee x_1 x_2 \bar{x}_3 x_4 \vee x_1 x_2 x_3 x_4 \vee \bar{x}_1 \bar{x}_2 x_3 \bar{x}_4) \vee \bar{x}_5 (\bar{x}_1 x_2 \bar{x}_3 \bar{x}_4 \vee x_1 x_2 \bar{x}_3 \bar{x}_4 \vee \bar{x}_1 \bar{x}_2 x_3 \bar{x}_4 \vee x_1 \bar{x}_2 x_3 \bar{x}_4 \vee x_1 x_2 \bar{x}_3 x_4).$$

Соответствующая карта Карно имеет вид, приведенный на рис. 5.7.

	$\bar{x}_1\bar{x}_2$	\bar{x}_1x_2	x_1x_2	$x_1\bar{x}_2$
$\bar{x}_3\bar{x}_4$	0	\bar{x}_5	\bar{x}_5	0
\bar{x}_3x_4	0	0	1	0
x_3x_4	0	0	x_5	0
$x_3\bar{x}_4$	1	0	0	1

Объединение 1 дает $\bar{x}_2 x_3 \bar{x}_4$.

Объединение \bar{x}_5 дает $x_2 \bar{x}_3 \bar{x}_4 \bar{x}_5$.

Объединение 1 и x_5 дает два слагаемых

$$x_1x_2\bar{x}_3x_4 \vee x_1x_2x_4x_5.$$

Рис. 5.7

Минимизированная функция имеет вид

$$f(x) = \bar{x}_2x_3\bar{x}_4 \vee x_2\bar{x}_3\bar{x}_4\bar{x}_5 \vee x_1x_2\bar{x}_3x_4 \vee x_1x_2x_4x_5.$$

Аналогично строятся карты Карно и для ФАЛ с большим числом аргументов, однако работа с картами затрудняется.

Картами Карно можно пользоваться и для представления функций в минимальной конъюнктивной форме. Процесс склеивания определяется расположением нулей в карте Карно. В группы объединяются нулевые клетки.

Правила объединения и запись групп остаются теми же. Минимизированная функция представляет собой логическое произведение, каждый сомножитель которого записывается на основе группы карты.

Пример 5.6. Записать минимальную КНФ функции, представленной картой Карно на рис. 5.8.

	$\bar{x}_1\bar{x}_2$	\bar{x}_1x_2	x_1x_2	$x_1\bar{x}_2$
x_3	1	1	0	1
\bar{x}_3	0	0	1	1

$$f(x_1, x_2, x_3) = (x_1 \vee x_3)(\bar{x}_1 \vee \bar{x}_2 \vee \bar{x}_3).$$

Рис. 5.8

5.7. Неполностью определенные логические функции и их минимизация

До сих пор рассматривались логические функции n переменных, значения которых были заданы на каждом из возможных 2^n наборов. На практике часто на ряде наборов значения логической функций не заданы, поскольку на этих наборах значение функции для проектировщика цифрового устройства не представляет интереса. Такие функции принято называть неполностью определенными. Их обычно доопределяют таким образом, чтобы максимально упростить соответствующие ФАЛ. Для этой цели удобно применять карты Карно.

Рассмотрим неполностью определенную функцию, представленную табл. 5.4. Звездочки соответствуют незадаваемым значениям функции и в

табл. 5.4, и в соответствующей карте Карно (рис. 5.9.) На карте Карно сформированы три группы по четыре клетки в предположении, что вместо звездочек стоят единицы. В этом случае минимизированная форма функции имеет вид

$$f(x_1x_2x_3x_4) = \bar{x}_1\bar{x}_2 \vee \bar{x}_1x_3 \vee x_1x_4.$$

Таблица 5.4

	$\bar{x}_1\bar{x}_2$	\bar{x}_1x_2	x_1x_2	$x_1\bar{x}_2$
$\bar{x}_3\bar{x}_4$	*	0	0	0
\bar{x}_3x_4	1	0	1	1
x_3x_4	1	1	*	1
$x_3\bar{x}_4$	1	*	0	0

Рис. 5.9

x_1	x_2	x_3	x_4	F
0	0	0	0	*
0	0	0	1	1
0	0	1	0	1
0	0	1	1	1
0	1	0	0	0
0	1	0	1	1
0	1	1	0	0
0	1	1	1	1
1	0	0	0	0
1	0	0	1	*
1	0	1	0	0
1	0	1	1	1
1	1	0	0	0
1	1	0	1	1
1	1	1	0	0
1	1	1	1	*

При минимизации функции пяти переменных символ неопределенности * может принимать значения 0, 1, x_5 , \bar{x}_5 , целесообразные с точки зрения процесса минимизации.

5.8. Синтез комбинационных схем

Под комбинационной схемой понимается техническое устройство, предназначенное для преобразования дискретной информации, причем значения выходных сигналов однозначно определяются значениями входных сигналов в данный момент времени. Предполагается, что в комбинационных схемах не происходит задержки сигнала, а входные и выходные сигналы могут принимать только значения единица и нуль (это могут быть высокий и низкий уровни напряжения).

Синтезировать комбинационную схему – это означает на основе заданного алгоритма работы построить структурную схему минимальной сложности из логических элементов заданного базиса.

Синтез комбинационных схем осуществляется в три этапа:

1) запись условий функционирования устройства (эти условия могут быть заданы словесно, с помощью таблицы истинности, либо с помощью логической функции);

2) минимизация логической функции и приведение ее к заданному базису;

3) составление структурной схемы устройства.

Пример 5.7. Синтезировать комбинационную схему, реализующую булеву функцию $f(x_1, x_2, x_3) = (x_1 / x_2) \oplus (\bar{x}_3 \rightarrow x_1)$ в базисе И-ИЛИ-НЕ. Рассмотреть переход к базисам И-НЕ и ИЛИ-НЕ.

Представим функцию в ДНФ. Для этого используем формулы

$$x_1 / x_2 = \overline{x_1 x_2}; \quad x_1 \rightarrow x_2 = \bar{x}_1 \vee x_2; \quad x_1 \oplus x_2 = x_1 \bar{x}_2 \vee \bar{x}_1 x_2.$$

Тогда

$$\begin{aligned} f(x_1, x_2, x_3) &= \overline{x_1 x_2} \oplus (x_3 \vee x_1) = \overline{x_1 x_2} (\overline{x_3 \vee x_1}) \vee \overline{\overline{x_1 x_2} (x_3 \vee x_1)} = \\ &= (\bar{x}_1 \vee \bar{x}_2) \bar{x}_3 \bar{x}_1 \vee x_1 x_2 (x_3 \vee x_1) = \bar{x}_1 \bar{x}_3 \vee \bar{x}_1 \bar{x}_2 \bar{x}_3 \vee x_1 x_2 x_3 \vee x_1 x_2 = \bar{x}_1 \bar{x}_3 \vee x_1 x_2. \end{aligned}$$

Логическая схема, реализующая эту функцию в базисе И-ИЛИ-НЕ, приведена на рис. 5.10, а.

Преобразуем $f(x_1, x_2, x_3)$ к базису И-НЕ:

$$\bar{x}_1 \bar{x}_3 \vee x_1 x_2 = \overline{\overline{\bar{x}_1 \bar{x}_3} \cdot \overline{x_1 x_2}} = \overline{\bar{x}_1 \bar{x}_3} \cdot \overline{x_1 x_2}.$$

Реализация функции в базисе И-НЕ приведена на рис. 5.10, б.

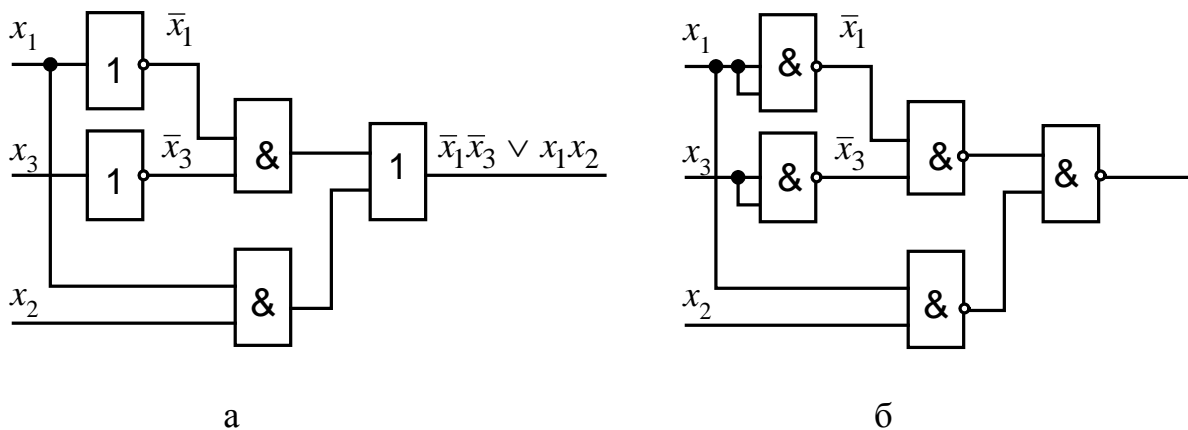


Рис. 5.10

Преобразуем $f(x_1, x_2, x_3)$ к базису ИЛИ-НЕ:

$$\bar{x}_1 \bar{x}_3 \vee x_1 x_2 = \overline{\overline{\bar{x}_1 \bar{x}_3} \cdot \overline{x_1 x_2}} = \overline{x_1 \vee x_3} \vee \overline{\bar{x}_1 \vee \bar{x}_2} = \overline{x_1 \vee x_3} \vee \overline{\bar{x}_1 \vee \bar{x}_2}.$$

Реализация функции в базисе ИЛИ-НЕ приведена на рис. 5.11.

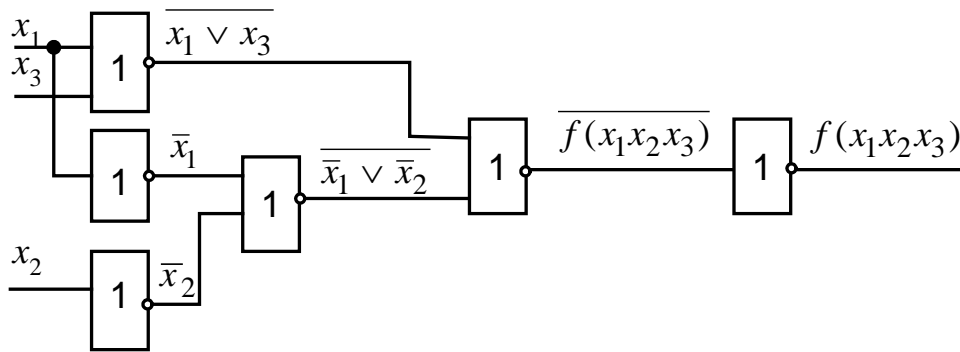


Рис. 5.11

В серийно выпускаемых интегральных микросхемах в одном корпусе могут быть объединены несколько логических схем, например, элемент 4И-НЕ, элемент 2И-ИЛИ-НЕ, элемент 2-2-2-3И-4ИЛИ-НЕ. Эти элементы приведены на рис. 5.12, а, б, в соответственно.

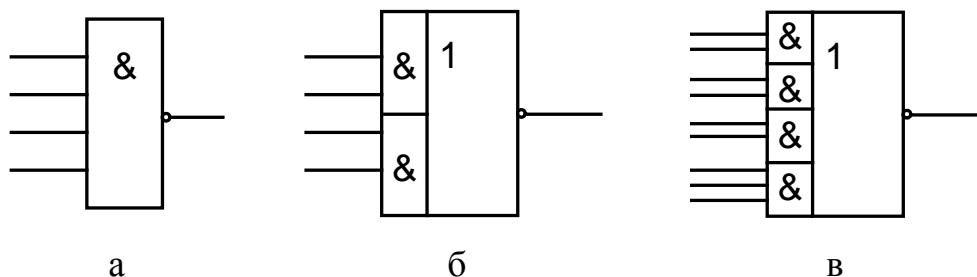


Рис. 5.12

Многовходовый элемент можно настраивать на выполнение нескольких различных функций за счет объединения некоторых входов, либо фиксации сигнала некоторых входов, как показано на рис. 5.13 на примере элемента 2И-ИЛИ-НЕ.

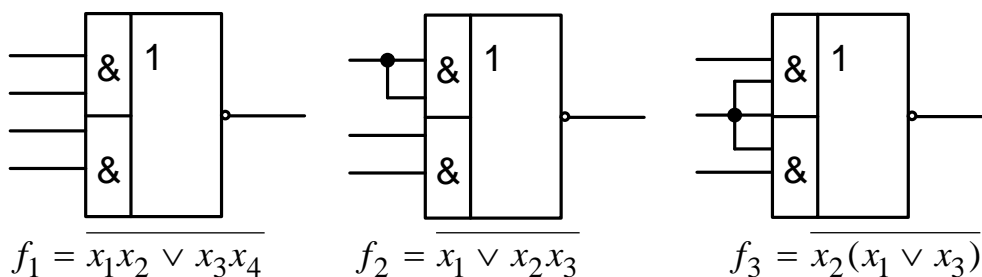


Рис. 5.13

Сложность логической схемы определяется суммарным числом входов логических элементов, условно выражающих цену схемы. Чем меньше входов, тем проще логическая схема и меньше ее цена.

5.9. Понятие о конечных автоматах и способы их задания

Термин «конечный автомат» используется для обозначения одного класса цифровых устройств, находящих применение в автоматике, телемеханике, вычислительной технике. В отличие от комбинационных схем эти устройства содержат память. Выходные сигналы конечного автомата (КА) зависят от значений на входах не только в данный момент времени, но и от предыдущих значений входных сигналов. Необходимая информация о сигналах, поступивших на входы раньше, может быть учтена посредством введения промежуточных сигналов, которые связаны с внутренней структурой автомата и называются состояниями автомата.

Используют два типа моделей КА – абстрактная и структурная. Абстрактный автомат – это математическая модель, в которой абстрагируются от реальной физической природы сигналов и рассматривают их как буквы некоторого алфавита. Абстрактный автомат (АА) имеет один вход и один выход и работает в дискретном времени, принимающем целые неотрицательные значения $t = 0, 1, 2, \dots$. Эти моменты времени называются тактами. В момент t АА, находясь в состоянии $q(t)$, способен воспринять на входе в этот же момент букву выходного алфавита $y(t)$ и перейти в следующее состояние $q(t+1)$. Если на вход АА подавать буква за буквой некоторую последовательность букв входного алфавита x_1, x_2, x_3, \dots – входное слово, то на выходе АА будут последовательно появляться буквы выходного алфавита y_1, y_2, y_3, \dots – выходное слово.

АА может быть задан аналитическим, табличным или матричным и графическим способами. При аналитическом способе задания АА задается множеством из пяти элементов: $A = \{X, Y, Q, \Gamma q, q_1\}$, где $X = \{x_1, x_2, \dots, x_n\}$ – множество входных сигналов (входной алфавит); $Y = \{y_1, y_2, \dots, y_m\}$ – множество выходных сигналов (выходной алфавит); $Q = \{q_1, q_2, \dots, q_r\}$ – множество возможных внутренних состояний (алфавит состояний); $\Gamma q = \{\Gamma q_1, \Gamma q_2, \dots, \Gamma q_r\}$ – отображение множества Q в себя, которое любому $q \in Q$ и каждой входной букве $x \in X$ сопоставляет состояние $q_k \in Q$, определяющее функцию переходов $\varphi(q, x)$, и выходную букву $y \in Y$, определяющую функцию выходов $\psi(q, x)$; $q_1 \in Q$ – начальное состояние автомата.

Пусть $X = \{x_1, x_2, x_3\}$, $Y = \{y_1, y_2, y_3, y_4, y_5, y_6\}$, $Q = \{q_1, q_2, q_3, q_4, q_5\}$;

$\Gamma q_1 = \{q_2(x_1/y_1), q_4(x_2/y_3), q_5(x_3/y_4)\}$,

$\Gamma q_2 = \{q_3(x_1/y_6), q_1(x_2/y_1)\}$,

$\Gamma q_3 = \{q_1(x_2/y_5), q_4(x_3/y_3)\}$,

$\Gamma q_4 = \{q_4(x_1/y_4), q_5(x_2/y_2)\}$,

$\Gamma q_5 = \{q_1(x_1/y_5), q_2(x_3/y_1)\}$.

Запись для отображения Γq_1 читается следующим образом: АА переходит из состояния q_1 в состояние q_2 , если на входе x_1 , при этом на выходе появляется y_1 ; АА переходит из состояния q_1 в q_4 , если на входе x_2 , при этом на выходе появ-

ляется y_3 ; АА переходит из состояния q_1 в q_5 , если на входе x_3 , при этом на выходе появляется y_4 . Аналогичный смысл имеют отображения $\Gamma q_2, \Gamma q_3, \Gamma q_4$ и Γq_5 .

Автомат называется конечным, если конечны множества X, Y, Q .

Функция переходов $\varphi(q, x)$ и функция выходов $\psi(q, x)$ определяют закон функционирования конечного автомата в дискретные моменты времени. На практике наибольшее распространение получили автоматы Мили и Мура. Закон функционирования автомата Мили задается уравнениями

$$\begin{cases} q(t+1) = \varphi[q(t), x(t)], \\ y(t) = \psi[q(t), x(t)], \end{cases}$$

которые показывают, что $q(t+1)$ и $y(t)$ однозначно определяются состоянием $q(t)$ и входным сигналом $x(t)$. В автомате Мура выходные сигналы зависят только от состояний автомата в рассматриваемый момент времени и не зависят от значений входных сигналов. Закон функционирования автомата Мура описывается следующими уравнениями:

$$\begin{cases} q(t+1) = \varphi[q(t), x(t)], \\ y(t) = \psi[q(t)]. \end{cases}$$

Два автомата называются эквивалентными, если любую одну и ту же входную последовательность они перерабатывают в одну и ту же выходную последовательность, но могут иметь различные состояния. Для каждого автомата Мили существует эквивалентный ему автомат Мура, т. е. модели Мили и Мура обладают равными функциональными возможностями. Наиболее общей является модель Мили, ее и будем рассматривать. Если функции φ и ψ определены на всех значениях $q(t)$ и $x(t)$, то такие автоматы называются **полными или полностью определенными**. Если φ и ψ определены не на всех значениях $q(t)$ и $x(t)$, то такие автоматы называются **частичным**.

Табличный способ предполагает задание АА с помощью обобщенной таблицы переходов и выходов. Строки таблицы соответствуют возможным значениям входного сигнала, а столбцы – внутренним состояниям автомата. На пересечении строки и столбца указывается очередное состояние автомата и через косую черту соответствующее значение выходного сигнала. Так автомату A , заданному ранее аналитическим способом, соответствует табл. 5.5.

Таблица 5.5

X	Q	q_1	q_2	q_3	q_4	q_5
x_1		q_2/y_1	q_3/y_6	q_1/y_5	q_4/y_4	q_1/y_5
x_2		q_4/y_3	q_1/y_1	–	q_5/y_2	–
x_3		q_5/y_4	–	q_4/y_3	–	q_2/y_1

Из табл. 5.5 видно, что автомат является частичным.

Иногда используют две отдельные таблицы: таблицу переходов и таблицу выходов. АА можно задать также *матрицей соединений автомата*. Строки и столбцы этой матрицы соответствуют различным состояниям автомата. На пересечении q_k -строки и q_e -столбца записывается буква входного алфавита $x_j \in X$, вызывающая переход автомата из состояния q_k в q_e , а через косую черту – буква выходного алфавита $y_j \in Y$, которая появляется на выходе автомата. Если ни одна из букв входного алфавита не переводит автомат из состояния q_k в q_e , то на соответствующем пересечении ставится нуль. Для рассматриваемого здесь примера матрица соединений автомата имеет вид, приведенный на рис. 5.14, а.

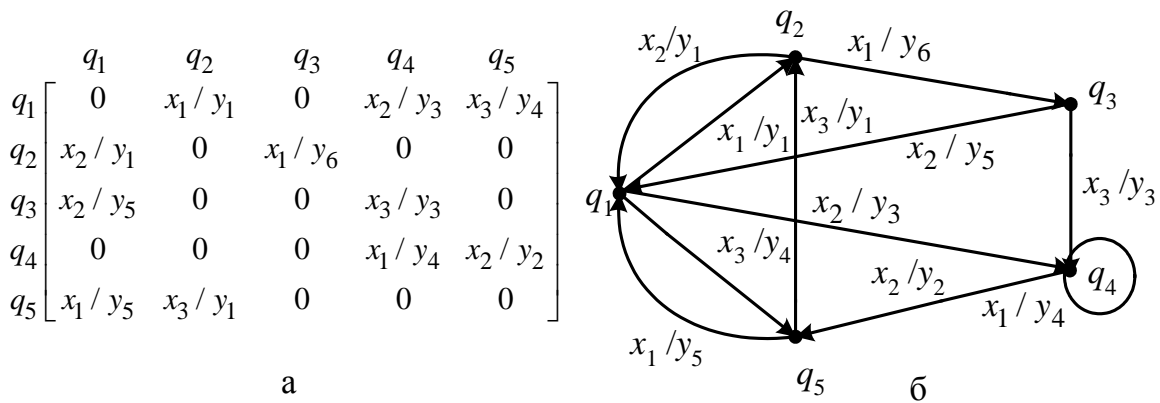


Рис. 5.14

При *графическом способе* задания АА изображается в виде ориентированного графа. Вершины графа отождествляются с внутренними состояниями автомата. Каждая дуга отличается входным сигналом, вызвавшим в автомате соответствующий переход, и выходным сигналом, который возникает при этом переходе. Для рассматриваемого ранее автомата граф будет иметь вид, представленный на рис. 5.14, б. Все рассмотренные способы задания АА равнозначны, однако графический способ обладает большей наглядностью.

При описании КА различают также понятие *структурного автомата*. В отличие от АА (см. рис. 5.15, а), имеющего один вход и один выход, структурный автомат (СА) имеет p входов (u_1, u_2, \dots, u_p) и ℓ выходов (v_1, v_2, \dots, v_e), на каждом из которых сигнал может принимать два значения – 0 или 1 (см. рис. 5.15, б).

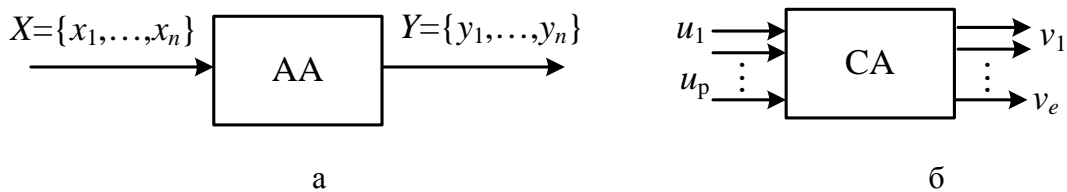


Рис. 5.15

Таким образом, букве x_i входного алфавита АА соответствует вектор, компонентами которого являются нули и единицы – сигналы на входах СА. Для кодирования входных сигналов АА различными векторами должно быть выполнено условие $p \geq \log_2 n$, т. е. p выбирается равным ближайшему целому числу, не меньшему чем $\log_2 n$. Точно так же букве y_i выходного алфавита АА соответствует вектор из 0 и 1, число компонентов e которого определяется выражением $e \geq \log_2 m$.

5.10. Синтез конечных автоматов

Используемый на практике метод синтеза КА предполагает, что общая структура автомата имеет вид, представленный на рис. 5.16.

Первое комбинационное устройство (КУ1) вырабатывает входные сигналы (сигналы возбуждения) для элементов памяти (ЭП).

Второе комбинационное устройство (КУ2) вырабатывает выходные сигналы автомата. Синтез КА сводится к определению количества элементов памяти и выбору их типов, а также к построению схем КУ1 и КУ2 в выбранном базисе.

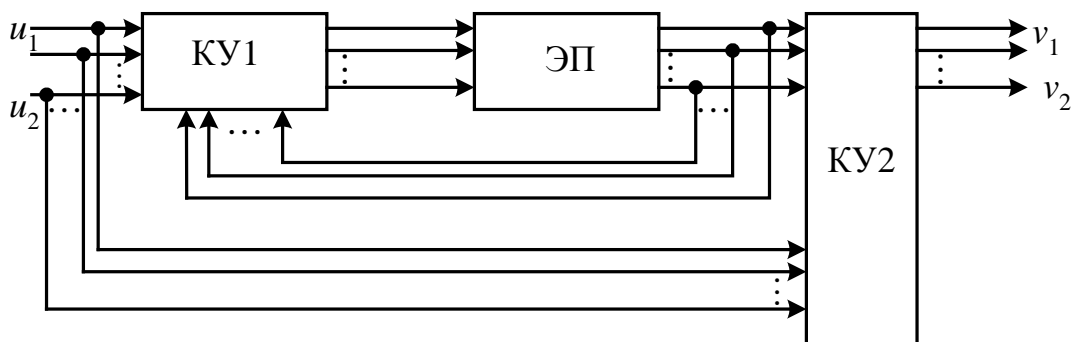


Рис. 5.16

5.10.1. Элементарные автоматы. В качестве ЭП, обеспечивающих временную задержку сигналов на один такт, используются серийно выпускаемые триггеры. Триггер – это двоичный запоминающий элемент, имеющий один или несколько входов и два выхода. Под действием входных сигналов триггер может переключаться в любое из двух устойчивых состояний (0 или 1) и сохранять это состояние в течение заданного времени'. Так как триггеры имеют только два устойчивых состояния, их называют элементарными автоматами. Выходные сигналы триггера совпадают с его состоянием.

Описать работу триггера можно таблицей переходов, в которой указываются значения 0 или 1 входных сигналов, вызывающих один из четырех возможных типов переходов: $0 \rightarrow 0$; $0 \rightarrow 1$; $1 \rightarrow 0$; $1 \rightarrow 1$.

Рассмотрим несколько типов триггеров.

1. *D*-триггер. Функциональная схема *D*-триггера приведена на рис. 5.17, а.

Триггер имеет один вход D и два выхода, обозначенных w и \bar{w} .
Табл. 5.6 определяет переходы D -триггера $w(t) \rightarrow w(t+1)$.

Таблица 5.6

D	$w(t)$	$w(t+1)$
0	0	0
1	0	1
0	1	0
1	1	1

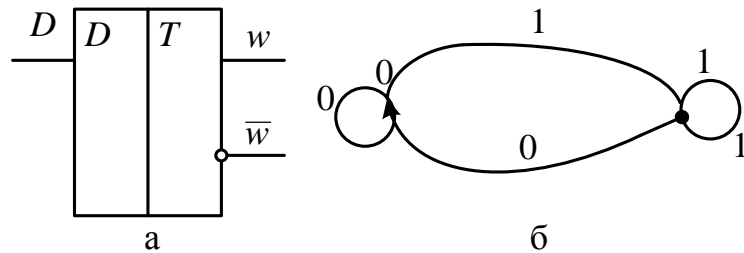


Рис. 5.17

Название D -триггера произошло от английского слова Delay (задержка), так как его следующее состояние равно сигналу на входе D , задержанному на один такт. Как уже отмечалось, выходные сигналы триггеров совпадают с состоянием, в которое они переходят, поэтому для описания элементарных автоматов достаточно задать таблицу переходов. Граф D -триггера приведен на рис. 5.17, б. Вершины графа соответствуют состояниям 0 и 1, а дуги отмечены сигналами, вызывающими соответствующие переходы.

Функцию переходов D -триггера можно также представить в аналитической форме $w(t+1) = D\bar{w}(t) \vee Dw(t) = D$.

2. T -триггер или триггер со счетным входом. Функциональная схема T -триггера приведена на рис. 5.18, а, а таблица переходов задана табл. 5.7.

Таблица 5.6

D	$w(t)$	$w(t+1)$
0	0	0
1	0	1
0	1	0
1	1	1

Таблица 5.7

T	$w(t)$	$w(t+1)$
0	0	0
1	0	1
1	1	0
0	1	1

3. При $T = 0$ триггер находится в состоянии хранения информации, сигнал $T = 1$ вызывает переключение триггера в противоположное состояние. Граф T -триггера приведен на рис. 5.18, б.

Таблица 5.7

T	$w(t)$	$w(t+1)$
0	0	0
1	0	1
1	1	0
0	1	1

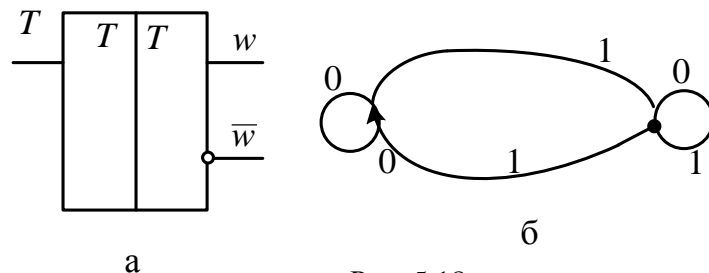


Рис. 5.18

Функция переходов T -триггера имеет вид

$$w(t + 1) = T\bar{w}(t) \vee \bar{T}w(t).$$

4. RS -триггер. Функциональная схема RS -триггера приведена на рис. 5.19, а. Схема имеет два входа S и R и два выхода, обозначенных w и \bar{w} . Таблица переходов RS -триггера задана табл. 5.8. Сигнал S (от англ. set – установка) переключает триггер в единичное состояние, а сигнал R (англ. reset – переустановка) вызывает переключение триггера в нулевое состояние. Вход S называется единичным установочным, а вход R – нулевым установочным.

Таблица 5.8

s	R	$w(t)$	$w(t+1)$
0	*	0	0
1	0	0	1
0	1	1	0
*	0	1	1

Таблица 5.9

J	K	$w(t)$	$w(t+1)$
0	*	0	0
1	*	0	1
*	1	1	0
*	0	1	1

Символ * в табл. 5.8 означает, что подача сигналов ноль или единица на соответствующие входы S и R не влияет на данный переход триггера. Граф RS -триггера приведен на рис. 5.19, б.

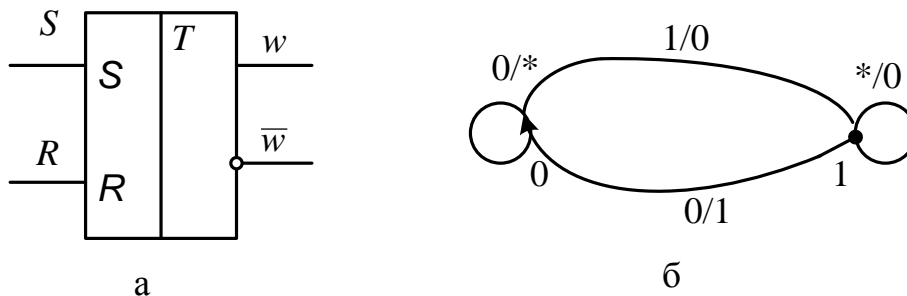


Рис. 5.19

Функция переходов RS -триггера имеет вид

$$w(t + 1) = S\bar{R}\bar{w}(t) \vee \bar{R}w(t).$$

5. J - K триггер. Функциональная схема J - K триггера приведена на рис. 5.20, а. Вход J называется единичным установочным входом, а вход K – нулевым установочным. В J - K триггере допускается одновременная подача входных сигналов $J=1$ и $K=1$. Таблица переходов J - K триггера задана табл. 5.9. Как и в предыдущем случае, символ * означает, что значение сигнала 0 или 1 на отмеченном входе не влияет на данный переход триггера.

Граф J - K триггера приведен на рис. 5.20, б. Дуги отмечены значениями сигналов, приводящими к соответствующим переходам. Функция переходов JK -триггеров имеет вид

$$w(t+1) = J\bar{w}(t) \vee \bar{K}w(t).$$

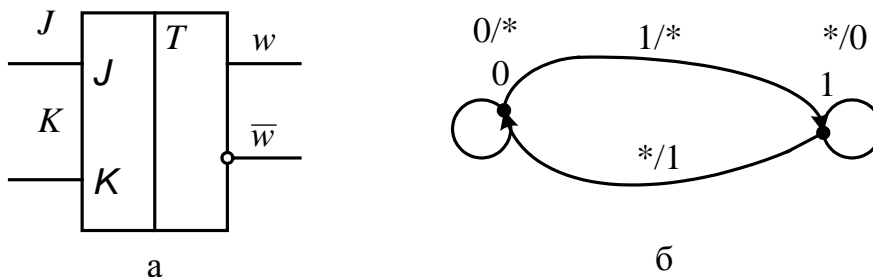


Рис. 5.20

5.10.2. Переход от абстрактного автомата к структурной схеме.

Структурный синтез автоматов заключается в составлении системы логических функций, на основании которой строятся комбинационные устройства, формирующие выходные сигналы и сигналы возбуждения элементов памяти (триггеров). Выделяют пять основных этапов структурного синтеза.

1. Кодирование входного и выходного алфавитов AA , кодирование состояний AA . Чтобы закодировать входные сигналы AA , нужно каждой букве $x_i = (i = \overline{1, n})$ входного алфавита поставить в соответствие совокупность значений двоичных сигналов u_1, u_2, \dots, u_p на входах CA . При этом количество p физических входов CA определяют из условия $p \geq \log_2 n$, выбирая ближайшее целое число.

При кодировании выходных сигналов AA каждой букве $y_j = (j = \overline{1, m})$ выходного алфавита ставится в соответствие совокупность значений двоичных сигналов v_1, v_2, \dots, v_e на выходах CA . Количество e физических выходов CA определяют из условия $e \geq \log_2 m$, выбирая ближайшее целое число. Аналогично кодированию входных и выходных сигналов каждой букве $q_k (k = \overline{1, r})$ алфавита состояний абстрактного автомата ставится в соответствие совокупность значений двоичных сигналов w_1, w_2, \dots, w_z состояний (выходов) элементов памяти. Количество элементов памяти определяют из условия $z \geq \log_2 r$, выбирая ближайшее целое число.

1. Выбор типа элементарных автоматов (элементов памяти). При выборе элементов памяти ориентируются на имеющуюся элементную базу. Для выбранного типа триггеров составляют таблицу переходов, в которой для каждого возможного типа переходов указана комбинация сигналов на входах (сигналов возбуждения триггеров).

2. Составление обобщенной таблицы переходов и выходов для закодированных переменных.

3. Определение функций возбуждения элементарных автоматов и выходных функций СА. Минимизация этих функций.

4. Составление структурной схемы синтезируемого автомата, т. е. составление комбинационной схемы, реализующей функции возбуждения ЭА и выходные функции СА.

Пример 5.8. Осуществить структурный синтез АА, заданного табл. 5.10. В качестве элементов памяти использовать *D*-триггеры, в качестве элементной базы использовать логические элементы И, ИЛИ, НЕ.

В качестве элементов памяти использовать *D*-триггеры, в качестве элементной базы использовать логические элементы И, ИЛИ, НЕ.

Таблица 5.10

<i>X</i>	<i>Q</i>	q_1	q_2	q_3	q_4	q_5
x_1		q_2/y_1	q_3/y_6	q_1/y_5	q_4/y_4	q_1/y_5
x_2		q_4/y_3	q_1/y_1	–	q_5/y_2	–
x_3		q_5/y_4	–	q_4/y_3	–	q_2/y_1

В соответствии с табл. 5.10 количество букв входного алфавита АА $n=3$, количество букв выходного алфавита $m = 6$, количество состояний $r = 5$. Определим количество входов СА: $p \geq \log_2 3$, принимаем $p = 2$. Количество выходов СА: $e \geq \log_2 6$, принимаем $e = 3$. Количество элементов памяти, т. е. необходимое количество *D*-триггеров: $z \geq \log_2 5$, принимаем $z = 3$.

Принцип кодирования переменных будет определять сложность схем комбинационных устройств, формирующих сигналы возбуждения D_i ($i = 1,2,3$) триггеров и выходные сигналы V_j ($j = 1, 2, 3$). Минимальное число слагаемых в ДСНФ для D_i и V_j получается при следующем алгоритме кодирования:

1) упорядочить кодируемые переменные в порядке уменьшения числа их появлений в таблице переходов-выходов;

2) первая из них, т. е. наиболее часто встречающаяся, кодируется нулевым кодом, затем используются коды, содержащие по одной единице, затем по две и т. д., до тех пор, пока все состояния не будут закодированы.

Закодируем переменные x_j, v_j, q_k . Переменные x_1, x_2, x_3 встречаются в табл. 5.10 по одному разу, результаты кодирования занесены в табл. 5.11.

В табл. 5.10. переменная y_1 встречается чаще всего (три раза), кодируем ее нулевым кодом; переменные y_3, y_4, y_5 встречаются по два раза, используем для них коды содержащие по одной единице; переменные y_2, y_6 встречаются в табл. 5.10 по 1 разу, используем для них коды содержащие по две единицы.

Результаты кодирования занесены в табл. 5.12.

При кодировании переменных q_k ($k = \overline{1,5}$) учтено, что q_1 и q_4 встречаются в табл. 5.10 по три раза, q_2 и q_5 по 2 раза, q_3 встречается один раз. Результаты кодирования занесены в табл. 5.13.

Таблица 5.11

X	U	u_1	u_2
x_1		0	0
x_2		0	1
x_3		1	0

Таблица 5.12

Y	V	v_1	v_2	v_3
y_1		0	0	0
y_2		0	1	1
y_3		0	0	1
y_4		0	1	0
y_5		1	0	0
y_6		1	0	1

Таблица 5.13

Q	W	w_1	w_2	w_3
q_1		0	0	0
q_2		0	0	1
q_3		0	1	1
q_4		0	1	0
q_5		1	0	0

На основании результатов кодирования строим обобщенную таблицу переходов и выходов СА (табл. 5.14), заменяя состояния, входные и выходные переменные их кодами. В клетках таблицы записаны состояния $w_1(t+1)w_2(t+1)w_3(t+1)$ и через черту выходы $v_1v_2v_3$ структурного автомата, которые возникнут при появлении на входах комбинаций u_1u_2 и исходном состоянии триггеров $w_1(t)w_2(t)w_3(t)$.

Таблица 5.14

u_1u_2	$w_1w_2w_3$	000	001	011	010	100
00		001/000	011/101	000/100	010/010	000/100
01		010/001	000/000	—	100/011	—
10		100/010	—	010/001	—	001/000

Используя таблицу переходов D -триггера (табл. 5.6) и данные табл. 5.14, составим обобщенную таблицу функционирования СА (табл. 5.15). Функции возбуждения трех триггеров обозначены через D_1 , D_2 и D_3 соответственно. $D_k=1$, если k -й триггер на данном переходе $w_k(t) \rightarrow w_k(t+1)$ переключается из состояния 0 в 1 или наоборот. $D_k=0$, если k -й триггер на рассматриваемом переходе $w_k(t) \rightarrow w_k(t+1)$ не переключается.

По табл. 5.15 можно записать СДНФ выходных функций V_1 , V_2 и V_3 и функций возбуждения триггеров D_1 , D_2 и D_3 , зависящих от набора переменных u_1 , u_2 , $w_1(t)$, $w_2(t)$, $w_3(t)$. В результате получим систему логических функций для построения комбинационной части автомата:

$$\begin{aligned}
V_1 &= \bar{u}_1 \bar{u}_2 \bar{w}_1 \bar{w}_2 w_3 \vee \bar{u}_1 \bar{u}_2 \bar{w}_1 w_2 w_3 \vee \bar{u}_1 \bar{u}_2 w_1 \bar{w}_2 \bar{w}_3, \\
V_2 &= \bar{u}_1 \bar{u}_2 \bar{w}_1 \bar{w}_2 \bar{w}_3 \vee \bar{u}_1 \bar{u}_2 \bar{w}_1 w_2 \bar{w}_3 \vee \bar{u}_1 u_2 \bar{w}_1 w_2 \bar{w}_3, \\
V_3 &= \bar{u}_1 u_2 \bar{w}_1 \bar{w}_2 \bar{w}_3 \vee \bar{u}_1 \bar{u}_2 \bar{w}_1 \bar{w}_2 w_3 \vee u_1 \bar{u}_2 \bar{w}_1 w_2 w_3 \vee \bar{u}_1 u_2 \bar{w}_1 w_2 \bar{w}_3, \\
D_1 &= u_1 \bar{u}_2 \bar{w}_1 \bar{w}_2 w_3 \vee \bar{u}_1 u_2 \bar{w}_1 w_2 \bar{w}_3, \\
D_2 &= \bar{u}_1 u_2 \bar{w}_1 \bar{w}_2 \bar{w}_3 \vee \bar{u}_1 \bar{u}_2 \bar{w}_1 \bar{w}_2 w_3 \vee u_1 \bar{u}_2 \bar{w}_1 w_2 w_3 \vee \bar{u}_1 \bar{u}_2 \bar{w}_1 w_2 \bar{w}_3, \\
D_3 &= \bar{u}_1 \bar{u}_2 \bar{w}_1 \bar{w}_2 \bar{w}_3 \vee \bar{u}_1 \bar{u}_2 \bar{w}_1 \bar{w}_2 w_3 \vee u_1 \bar{u}_2 w_1 \bar{w}_2 \bar{w}_3.
\end{aligned}$$

Таблица 5.15

u_1	u_2	$w_1(t)$	$w_2(t)$	$w_3(t)$	$w_1(t+1)$	$w_2(t+1)$	$w_3(t+1)$	V_1	V_2	V_3	D_1	D_2	D_3
0	0	0	0	0	0	0	1	0	0	0	0	0	1
0	1	0	0	0	0	1	0	0	0	1	0	1	0
1	0	0	0	0	1		0	0	1	0	1	0	0
0	0	0	0	1	0	1	1	1	0	1	0	1	1
0	1	0	0	1	0	0	0	0	0	0	0	0	0
1	0	0	0	1	*	*	*	*	*	*	*	*	*
0	0	0	1	1	0	0	0	1	0	0	0	0	0
0	1	0	1	1	*	*	*	*	*	*	*	*	*
1	0	0	1	1	0	1	0	0	0	1	0	1	0
0	0	0	1	0	0	1	0	0	1	0	0	1	0
0	1	0	1	0	1	0	0	0	1	1	1	0	0
1	0	0	1	0	*	*	*	*	*	*	*	*	*
0	0	1	0	0	0	0	0	1	0	0	0	0	0
0	1	1	0	0	*	*	*	*	*	*	*	*	*
1	0	1	0	0	0	0	1	0	0	0	0	0	1

В записанной системе некоторые конъюнкции встречаются несколько раз. Каждую конъюнкцию обозначим определенным числом, в результате система логических функций примет вид:

$$\begin{aligned}
V_1 &= 1 \vee 2 \vee 3; \\
V_2 &= 4 \vee 5 \vee 6; \\
V_3 &= 7 \vee 1 \vee 8 \vee 6; \\
D_2 &= 7 \vee 1 \vee 8 \vee 5; \\
D_3 &= 9 \vee 1 \vee 10.
\end{aligned}$$

Осуществить минимизацию функций $V_i D_j$.

Карта Карно для V_i представлена на рис. 5.21

При кодировании букв входного алфавита была не задействована комбинация 11, поэтому вторая строка карты заполнена звездочками.

При кодировании состояний автоматов не задействованы кодовые комбинации 110, 101, 111, которые по обозначениям в карте Карно соответствуют

первым трем столбцам. Эти столбцы также заполнены звездочками, что означает независимость возникновения таких сигналов при работе автомата. В карте Карно для всех описываемых функций эти звездочки будут присутствовать.

Остальные клетки карт заполняются в соответствии с табл. 5.15. Звездочки, соответствующие строкам этой таблице, имеют увеличенный размер, чтобы отличать их от предыдущих. Единицы и нули рассматриваются обычным способом.

$W_1W_2W_3$ U_1U_2	101	111	110	100	000	010	011	001
00	*	*	*	1	0	0	1	1
01	*	*	*	☆	0	0	☆	0
11	*	*	*	*	*	*	*	*
10	*	*	*	0	0	☆	0	☆

Рис. 5.21

При минимизации V_1 сформированы две группы. Первая из 8 клеток описывается конъюнкцией двух переменных \bar{U}_1W_1 (произведение букв, которые являются общими для всех клеток карт, объединенных в группу). Вторая из 4 клеток описывается конъюнкцией трех переменных $\bar{U}_1\bar{U}_2W_3$, тогда $U_{1\min} = \bar{U}_1W_1 \vee \bar{U}_1\bar{U}_2W_3$.

Карта Карно для функции V_2 приведена на рис. 5.22.

$W_1W_2W_3$ U_1U_2	101	111	110	100	000	010	011	001
00	*	*	*	0	0	1	0	0
01	*	*	*	☆	0	1	☆	0
11	*	*	*	*	*	*	*	*
10	*	*	*	0	1	☆	0	☆

Рис. 5.22

При минимизации V_2 в группу из 8 клеток объединены 3-й и 6-й столбцы карты, которые зрительно не являются рядом стоящими, но имеют буквы W_2 и \bar{W}_3 , что позволяет описать группу конъюнкцией $W_2\bar{W}_3$. Вторая группа из четырех клеток описывается конъюнкцией $U_1\bar{W}_1\bar{W}_3$, тогда $U_{2\min} = W_2\bar{W}_3 \vee U_1\bar{W}_1\bar{W}_3$. Следует отметить, что сформированная группа, затрагивающая обе половины

карты, должна быть симметрична относительно вертикального центра симметрии (граница 3-го и 4-го столбцов).

Карта Карно для функции V_3 приведена на рис. 5.23.

$W_1W_2W_3$	101	111	110	100	000	010	011	001
U_1U_2								
00	*	*	*	0	0	0	0	1
01	*	*	*	☆	1	1	☆	0
11	*	*	*	*	*	*	*	*
10	*	*	*	0	0	☆	1	☆

Рис. 5.23

При минимизации V_3 в одну группу объединены четыре угловые клетки, которые имеют одну общую букву \bar{U}_2 по строкам и две общие буквы $\bar{W}_2 W_3$ по столбцам и могут быть описаны конъюнкцией $\bar{U}_2 \bar{W}_2 W_3$. Еще сформированы две группы по 8 клеток, показанные на карте. В результате $U_{2\min} = \bar{U}_2 \bar{W}_2 W_3 \vee U_1 \bar{W}_1 \bar{W}_3 \vee U_2 \bar{W}_3 \vee U_1 W_3$. Аналогичным образом осуществим минимизацию функций возбуждения триггеров D_1, D_2 и D_3 .

Карта Карно для функции D_1 приведена на рис. 5.24.

$W_1W_2W_3$	101	111	110	100	000	010	011	001
U_1U_2								
00	*	*	*	0	0	0	0	0
01	*	*	*	☆	0	1	☆	0
11	*	*	*	*	*	*	*	*
10	*	*	*	0	1	☆	0	☆

Рис. 5.24

В результате минимизации получено следующее выражение

$$D_{1\min} = U_2 W_2 \vee U_1 \bar{W}_1 \bar{W}_3.$$

Карта Карно для функции D_2 приведена на рис. 5.25.

$W_1W_2W_3$ U_1U_2	101	111	110	100	000	010	011	001
00	* ↕	*	* ↕	0	0	1 ↕	0	1 ↕
01	* ←	*	*	☆	1	0	☆	0
11	* ↙	*	*	*	*	*	*	*
10	* ↘	*	* ↘	0	0	☆ ↘	1	☆ ↘

Рис. 5.25

В результате минимизации получено следующее выражение

$$D_{2\min} = U_1 W_3 \vee \bar{U}_2 W_2 \bar{W}_3 \vee U_2 \bar{W}_2 \bar{W}_3 \vee \bar{U}_2 \bar{W}_2 W_3.$$

Крат Карно для функции D_3 приведена на рис. 5.26.

$W_1W_2W_3$ U_1U_2	101	111	110	100	000	010	011	001
00	* ↕	*	*	0	1	0	0	1 ↕
01	* ←	*	*	☆	0	0	☆	0
11	* ↙	*	*	*	*	*	*	*
10	* ↘	*	* ↘	1	0	☆	0	☆ ↘

Рис. 5.26

В результате минимизации получено следующее выражение

$$D_{3\min} = U_1 W_3 \vee \bar{U}_2 W_2 \bar{W}_3 \vee \bar{U}_2 \bar{W}_2 W_3 \vee \bar{U}_1 \bar{U}_2 \bar{W}_2 \bar{W}_3.$$

Прежде чем изображать структурную схему автомата, составим таблицы функционирования шифратора и дешифратора, опишем их аналитически с помощью функций алгебры логики и осуществим реализацию, добавив к общей схеме. Шифратор и дешифратор являются комбинационными схемами и реализуются в том же базисе, который задан для автомата.

Шифратор должен обеспечить переход от букв входного алфавита к соответствующим кодам.

Табл. 5.16 соответствует результатам кодирования и позволяет записать U_1 и U_2 как логические функции трех переменных.

$$U_1 = \bar{X}_1 \bar{X}_2 X_3 \text{ и } U_2 = \bar{X}_1 X_2 \bar{X}_3.$$

Таблица 5.16

X_1	X_2	X_3	U_1	U_2
1	0	0	0	0
0	1	0	0	1
0	0	1	1	0

Дешифратор должен обеспечить переход от кодов выходного алфавита к самим буквам.

Табл. 5.17 является истинности дешифратора с тремя входами и шестью выходами.

Таблица 5.17

V_1	V_2	V_3	y_1	y_2	y_3	y_4	y_5	y_6
0	0	0	1	0	0	0	0	0
0	1	1	0	1	0	0	0	0
0	0	1	0	0	1	0	0	0
0	1	0	0	0	0	1	0	0
1	0	0	0	0	0	0	1	0
1	0	1	0	0	0	0	0	1

$$y_1 = \bar{V}_1 \bar{V}_2 \bar{V}_3;$$

$$y_2 = \bar{V}_1 V_2 V_3;$$

$$y_3 = \bar{V}_1 \bar{V}_2 V_3;$$

$$y_4 = \bar{V}_1 V_2 \bar{V}_3;$$

$$y_5 = V_1 \bar{V}_2 \bar{V}_3;$$

$$y_6 = V_1 \bar{V}_2 V_3.$$

Каждому предусмотренному набору входных сигналов соответствует сигнал на одном из выходов. Входные сигналы $y_1 - y_6$ представлены как функции переменных V_1, V_2 и V_3 .

Общая функциональная схема автомата приведена на рис. 5.27.

В основу построения схемы положены минимизированные выражения для $V_1, V_2, V_3, D_1, D_2, D_3$ и выходные функции шифратора и дешифратора.

Переменные с выходов шифратора $U_1, \bar{U}_1, U_2, \bar{U}_2$ пронумерованы и объединены в одну шину, тогда необходимый сигнал можно взять в любой точке шины. Аналогичным образом пронумерованы и объединены в одну шину выходы триггеров $W_1, \bar{W}_1, W_2, \bar{W}_2, W_3, \bar{W}_3$, которые рассматриваются как сигналами

лы обратной связи и являются выходными для комбинационных схем. При построении дешифратора в одну шину объединены выходные сигналы автомата $V_1, \bar{V}_1, V_2, \bar{V}_2, V_3, \bar{V}_3$.

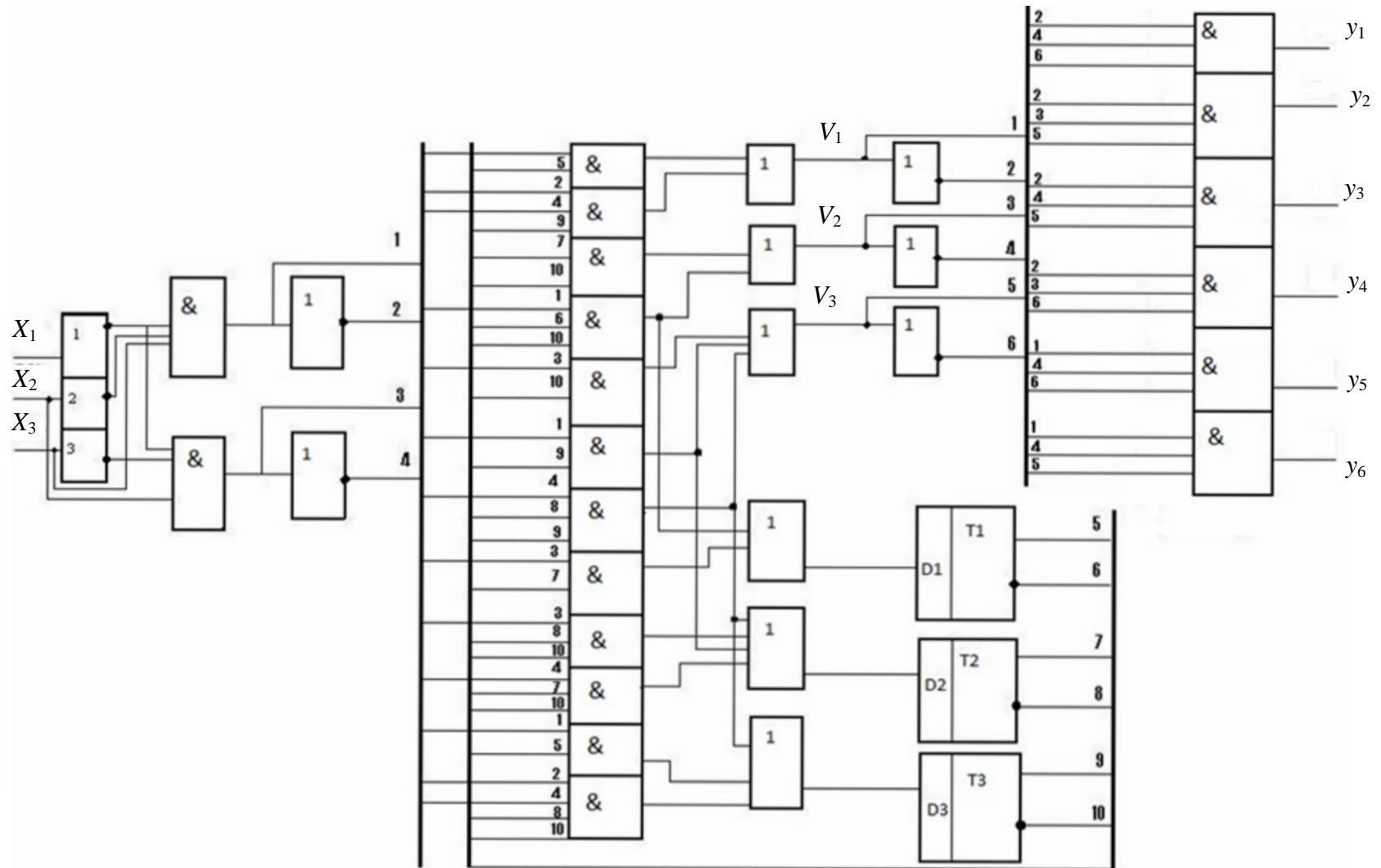


Рис. 5.27

ТЕМА 6. МАТЕМАТИЧЕСКОЕ ОПИСАНИЕ ЛИНЕЙНЫХ СИСТЕМ И ИХ ЭЛЕМЕНТОВ С ПОСТОЯННЫМИ ПАРАМЕТРАМИ

6.1. Классификация элементов

Система автоматического управления – это совокупность элементов, соединенных в замкнутый контур, которые функционируют согласованно и подчинены определенной форме управления.

Элементы можно классифицировать следующим образом:

По функциональному назначению:

1. Измерительные
2. Усилительно-преобразовательные
3. Исполнительные
4. Корректирующие

По виду энергии, используемой для работы:

1. Электрические
2. Механические
3. Гидравлические
4. Пневматические
5. Комбинированные

По характеру математического соответствия между входным и выходным сигналами.

При математическом описании элементы называются звеньями САУ. Несмотря на многообразие различного рода элементов (устройств) и независимо от физических принципов их работы, поведение каждого из них может быть описано дифференциальным уравнением, связывающим входную и выходную переменные. Элементы описываются, как правило, дифференциальными уравнениями первого или второго порядка. При объединении их в систему порядок дифференциального уравнения повышается.

Будем рассматривать одномерную модель с одним входом и одним выходом, и обозначим входную величину звена через $u(t)$, а выходную через $y(t)$. При рассмотрении линейных систем статическая характеристика $y = f(t)$ любого звена может быть изображена прямой линией.

В позиционном (или усилительном) звене линейной зависимостью $y = Ku$ связаны входная и выходная величина в установившемся режиме (рис. 6.1, а). Здесь K – коэффициент передачи или коэффициент усиления звена.

В интегрирующих звеньях линейной зависимостью $\frac{du}{dt} = Ku$ (рис. 6.1, б) связаны производная выходной величины и входная величина в установившемся режиме, или $y = K \int u dt$ откуда и произошло название звена. Если входная и выходная величины имеют одинаковую размерность, то K имеет размерность $[c^{-1}]$.

В дифференцирующих звеньях линейной зависимостью $y = K \frac{du}{dt}$

(рис. 6.1, в) связаны в установившемся режиме выходная величина и производная входной величины. Если входная и выходная величины имеют одинаковую размерность, то K имеет размерность [с].

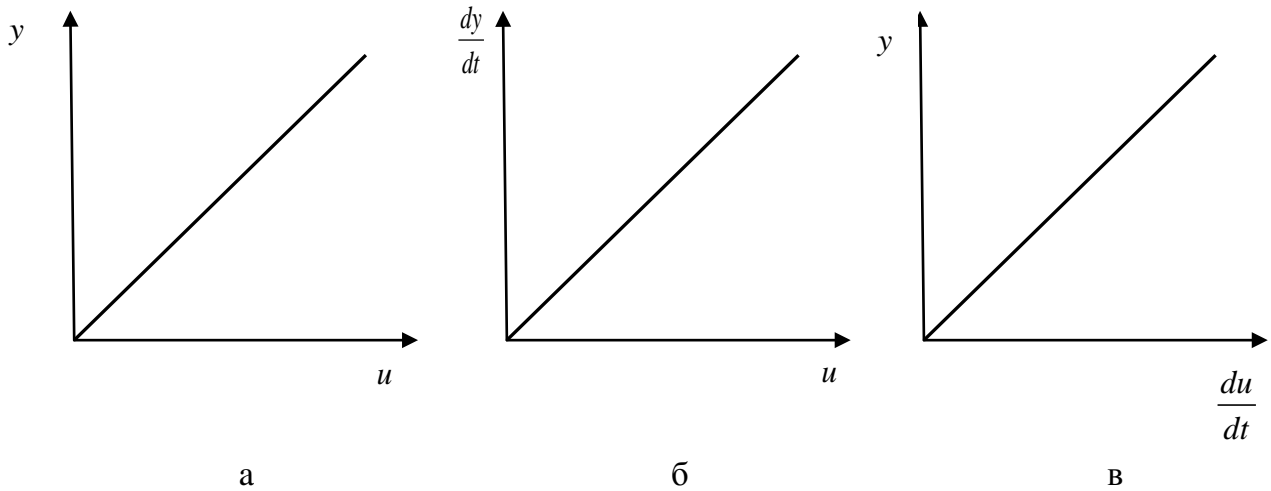


Рис. 6.1 Статические характеристики звеньев: а – усилительного, б – интегрирующего, в – дифференцирующего

Рассмотренные звенья идеализированы, так как не учитывается их инерционность.

6.2. Уравнения динамики и статики

В общем случае линейная система описывается линейным дифференциальным уравнением, представленным в стандартной форме

$$a_n \cdot \frac{d^n}{dt^n} + a_{n-1} \frac{d^{n-1}y}{dt^{n-1}} + \dots + a_1 \frac{dy}{dt} + a_0 y = b_m \frac{d^m u}{dt^m} + \dots + b_1 \frac{du}{dt} + b_0 u. \quad (6.1)$$

где t – текущее время; n – порядок дифференциального уравнения; $m \geq n$; u – входное воздействие (сигнал); y – выходное воздействие (сигнал); a_i, b_j – коэффициенты, определяемые параметрами системы.

Если эти коэффициенты не зависят от времени, то система называется **стационарной**.

Дифференциальные уравнения называют уравнениями динамики, они описывают переходные режимы в системах. Переходной режим возникает при подаче на вход сигнала $u(t)$ (включение устройства) и существует до тех пор, пока на выходе не устанавливается определенная величина сигнала $y(t)$.

Переходной процесс – это процесс изменения сигнала $y(t)$ на выходе от момента подачи входного сигнала $u(t)$ до установления процесса на выходе.

С математической точки зрения $y(t)$ – решение дифференциального уравнения (6.1).

Уравнение статики – уравнение установившегося режима, когда все производные равны нулю. Для позиционного элемента $a_0y = b_0u$; $y = Ku$.

Сигналы, действующие в системе, могут быть непрерывными и дискретными, детерминированными и случайными.

Одной из **основных задач** анализа является определение реакции системы на внешние воздействия. Для оценки свойств системы достаточно знать как она реагирует на некоторые типовые сигналы. В любом случае нужно находить решение дифференциального уравнения.

6.3. Понятие передаточной функции

В инженерной практике широко используется метод решения дифференциальных уравнений, основанный на интегральном преобразовании Лапласа и позволяющий свести задачу к алгебраическим действиям.

Для сигнала $f(t)$ преобразование Лапласа $F(s) = \int_0^{\infty} f(t)e^{-st} dt$, где s – оператор Лапласа, а $F(s)$ называется изображением функции $f(t)$.

При нулевых начальных условиях, т. е. в том случае, если при $t < 0$ (до момента подачи сигнала) входная и выходная величины, а так же их производные, тождественно равны нулю, от уравнения (6.1) формально можно перейти к выражению:

$$\begin{aligned} a_n s^n Y(s) + a_{n-1} s^{n-1} Y(s) + \dots + a_1 s Y(s) + a_0 Y(s) = \\ = b_m s^m U(s) + \dots + b_1 s U(s) + b_0 U(s). \end{aligned} \quad (6.2)$$

Здесь учтено, что дифференцирование функции $f(t)$ во временной области соответствует умножению ее изображения $F(s)$ на оператор s в области изображений по Лапласу. Взятию производной второго порядка $\frac{d^2 f(t)}{dt^2}$ соответствует умножению на квадрат оператора s и так далее.

Передаточная функция звена (системы) $W(s)$ – это отношение изображения по Лапласу выходного сигнала $Y(s)$ к изображению по Лапласу входного сигнала $U(s)$ при нулевых начальных условиях, тогда из (6.2) следует:

$$W(s) = \frac{Y(s)}{U(s)} = \frac{b_m s^m + b_{m-1} s^{m-1} + \dots + b_1 s + b_0}{a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0}, n \geq m. \quad (6.3)$$

Передаточная функция является дробно-рациональной функцией комплексной величины s . Передаточная функция элемента не зависит от того, какой функцией времени является его входное воздействие. Она зависит лишь от вида дифференциального уравнения и от значений параметров элемента, кото-

рые определяют коэффициенты уравнения. Передаточные функции рассмотренных ранее звеньев имеют вид:

для позиционного или усилительного звена $W(s) = K$;

для интегрирующего звена $W(s) = \frac{K}{s}$;

для дифференцирующего звена $W(s) = Ks$.

Передаточная функция – это одна из форм математических моделей элементов. Зная $W(s)$ и $U(s)$ можно найти $Y(s)$ – изображение по Лапласу выходного сигнала: $Y(s) = U(s) \cdot W(s)$, тогда $y(t)$ можно найти как обратное преобразование Лапласа:

$$y(t) = L^{-1}\{Y(s)\} \quad (6.4)$$

не решая дифференциального уравнения (6.1).

Передаточные функции типовых звеньев известны, а для сигналов существуют таблицы соответствия между описанием их во временной области и в области изображений по Лапласу. Поэтому решение (6.4) не представляет сложности.

$W(s)$, заданную в форме (6.3), можно представить следующим образом:

$$W(s) = K \frac{(s - z_1)(s - z_2) \dots (s - z_m)}{(s - p_1)(s - p_2) \dots (s - p_n)} = K \frac{\prod_{j=1}^m (s - z_j)}{\prod_{i=1}^n (s - p_i)}, \quad (6.5)$$

где K – коэффициент усиления, z_j – нули системы, т. е. корни многочлена числителя, p_i – полюсы системы, т. е. корни многочлена знаменателя.

Для описания моделей систем и действий над ними широко используется система MATLAB и пакет прикладных программ *Control System Toolbox*. В пакете введен класс объектов, называемый *lti* объекты – линейные с постоянными параметрами. При создании *lti* объекта ему присваивается имя. Передаточная функция имеет несколько форм представления:

tf-форма, в которой передаточная функция (6.3) задается двумя векторами строками, составленными из коэффициентов многочленов числителя и знаменателя в порядке убывания степеней s . Например, оператор $W = tf([2 \ 1], [1 \ 3 \ 7])$ создает объект W подкласса *tf*, соответствующий передаточной функции $W(s) = \frac{2s+1}{s^2+3s+7}$;

zpk-форма нулей, полюсов и коэффициента усиления, в которой передаточная функция (6.5) описывается двумя векторами-строками и одним числом, задающим нули, полюсы и коэффициент усиления системы. Например оператор $W = zpk([-1, -2], [-1 - j, -1 + j, -2], 5)$ формирует *zpk*-объект, соответствующий передаточной функции $W(s) = \frac{5(s+1)(s+2)}{(s+2)(s^2+2s+2)}$.

При отсутствии нулей на их место записывается знак пусто [].

zpk-форма информативней формы *tf* для человека и удобней для вычислений на ЭВМ. Нули и полюса отрицательны, если они вещественные, либо имеют отрицательную действительную часть.

При описании элементов и систем кроме входных $u(t)$ и выходных $y(t)$ переменных можно выделить некоторые промежуточные переменные $x(t)$, которые связаны с внутренней структурой системы и называются переменными состояния. В параметрах пространства состояний система n -го порядка с одним входом и одним выходом описывается системой уравнений

$$\begin{cases} \dot{x} = Ax + Bu, \\ y = Cx + Du. \end{cases} \quad (6.6)$$

Здесь A – квадратная матрица порядка n , элементы которой определяются коэффициентами дифференциального уравнения, B – вектор-столбец $[n \cdot 1]$ постоянных коэффициентов, C – вектор-строка $[1 \cdot n]$ постоянных коэффициентов, D – одноэлементная матрица.

ss-форма представляет передаточную функцию в параметрах пространства состояний. Если объект с именем W представлен в *tf*-форме, то переход к *zpk*- или *ss*-форме осуществляется командами $zpk(W)$ или $ss(W)$.

6.4. Передаточные функции различных соединений звеньев

При последовательном соединении звеньев с известными передаточными функциями $W_1(s), W_2(s), \dots, W_n(s)$ (рис. 6.2) передаточные функции перемножаются:

$$W(s) = \frac{Y(s)}{U(s)} = W_1(s) \cdot W_2(s) \cdot \dots \cdot W_n(s) = \prod_{i=1}^n W_i(s).$$

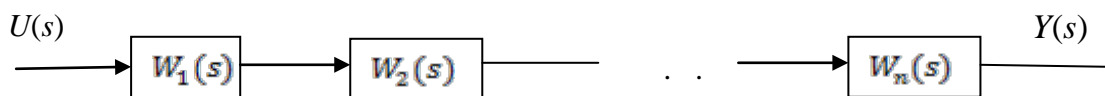


Рис. 6.2. Последовательное соединение звеньев

При параллельном соединении звеньев с передаточными функциями $W_1(s), W_2(s), \dots, W_n(s)$ (рис. 6.3) передаточные функции складываются

$$W(s) = \frac{Y(s)}{U(s)} = W_1(s) + W_2(s) + \dots + W_n(s) = \sum_{i=1}^n W_i(s).$$

При охвате звена с передаточной функцией $W_1(s)$ обратной связью с передаточной функцией $W_2(s)$ (рис. 6.4)

$$W(s) = \frac{Y(s)}{W(s)} = \frac{W_1(s)}{1 \pm W_1(s) \cdot W_2(s)}$$

где знак плюс соответствует отрицательной, а знак минус – положительной обратной связи. Если $W_2(s) = 1$, то обратная связь называется единичной.

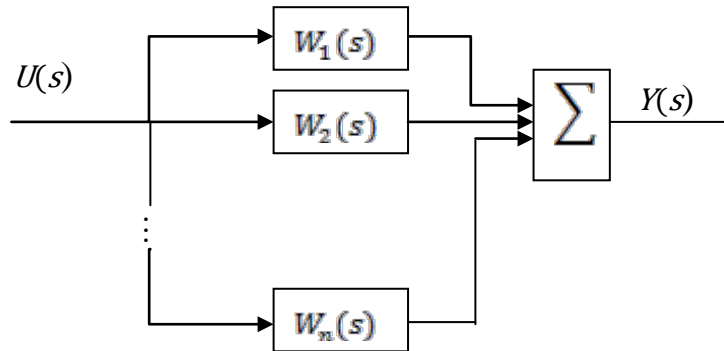


Рис. 6.3. Параллельное соединение звеньев

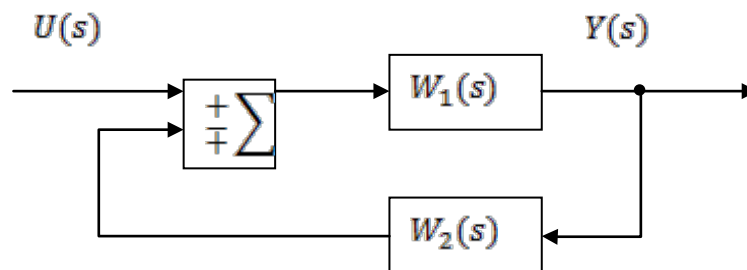


Рис. 6.4. Охват звена обратной связью

6.5. Временные характеристики систем и их элементов

К временным характеристикам линейных элементов и систем относятся переходная и импульсная переходная (весовая) функции.

Переходная функция $h(t)$ – это функция, определяющая изменение выходной величины системы (или отдельного элемента) при воздействии на входе единичного ступенчатого сигнала $1(t)$ при нулевых начальных условиях.

Пусть сигнал $1(t)$ подается на устройство с передаточной функцией $W(s)$ (рис. 6.5). Характер изменения $h(t)$ определяется видом передаточной функции.

Изображение по Лапласу сигнала $1(t)$ имеет вид $L = \{1(t)\} = \frac{1}{s}$, тогда изображение по Лапласу сигнала на выходе $H(s) = \frac{1}{s} W(s)$. Переходная характеристика может быть найдена как обратное преобразование Лапласа от изображения выходного сигнала:

$$h(t) = L^{-1}\{H(s)\} = L^{-1}\left\{\frac{W(s)}{s}\right\}.$$

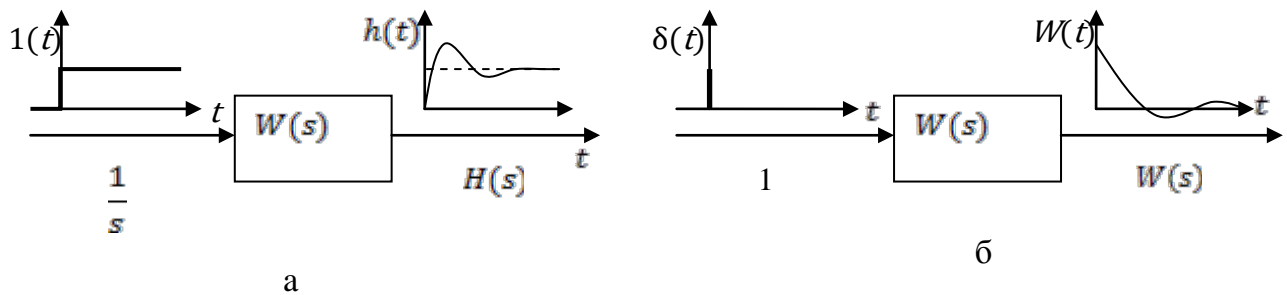


Рис. 6.5. Получение переходной (а) и весовой (б) функции.

Импульсная переходная (или весовая) функция $w(t)$ – это функция, определяющая изменение выходной величины системы (или отдельного элемента) при воздействии на входе дельта функции $\delta(t)$ при нулевых начальных условиях.

Пусть сигнал $\delta(t)$ подается на устройство с передаточной функцией $W(s)$ (рис. 6.5, б). Изображение по Лапласу сигнала $\delta(t)$ равно единице, $L\{\delta(t)\} = 1$. Изображение по Лапласу сигнала на выходе $L\{w(t)\} = W(s)$, таким образом передаточная функция $W(s)$ является изображением весовой функции $w(t)$:

$$W(s) = L\{w(t)\} = \int_0^{\infty} w(t)e^{-st} dt,$$

а весовая функция $w(t)$ может быть найдена как обратное преобразование Лапласа от передаточной функции $W(s)$, т. е.

$$w(t) = L^{-1}\{W(s)\} = \frac{1}{2\pi j} \int_{c-j\infty}^{c+j\infty} W(s)e^{st} ds,$$

Весовая функция позволяет определить реакцию системы на произвольный входной сигнал $u(t)$, являясь универсальной динамической характеристикой. Согласно теореме свертки в вещественной области выражению $Y(s) = U(s) \cdot W(s)$ соответствует свертка оригиналов:

$$y(t) = u(t) \cdot w(t) = \int_0^t u(\tau)w(t - \tau)d\tau = \int_0^t u(t - \tau)w(\tau)d\tau.$$

Переходная и импульсная переходная характеристики связаны между собой следующим образом:

$$w(t) = \frac{dh(t)}{dt}; \quad h(t) = \int_0^t w(\tau)d\tau.$$

Пример 6.1. Определить передаточную, переходную и весовую функции звена, которое описывается дифференциальным уравнением $T \frac{dy}{dt} + y = ku$.

Переходя в область изображений по Лапласу $Ts \cdot Y(s) + Y(s) = K \cdot U(s)$, найдем передаточную функцию $W(s) = \frac{Y(s)}{U(s)} = \frac{K}{Ts+1}$. Весовая или импульсная переходная характеристика: $W(t) = L^{-1} \left\{ \frac{K}{Ts+1} \right\} = \frac{K}{T} e^{-\frac{t}{T}}$. Переходная характеристика: $h(t) = L^{-1} \left\{ \frac{K}{s(Ts+1)} \right\} = K(1 - e^{-\frac{t}{T}})$.

Определим сигнал на выходе $y(t)$, если на входе действует линейно изменяющийся сигнал $u(t) = at$:

$$y(t) = \int_0^t u(\tau)W(t-\tau)d\tau = \int_0^t a\tau \cdot \frac{K}{T} e^{-\frac{(t-\tau)}{T}} d\tau = aK \left[t - T(1 - e^{-\frac{t}{T}}) \right].$$

6.6. Понятие о частотных характеристиках систем и их элементов

Частотные характеристики определяют динамические свойства звеньев при воздействии на них гармонических сигналов. Если входное воздействие линейной стационарной системы (отдельного элемента) является гармонической функцией вида $u(t) = A \cos(\omega t)$, то выходной сигнал после окончания переходного процесса представляет собой гармоническую функцию той же частоты, но отличающуюся в общем случае по амплитуде и фазе, т. е. $y(t) = B \cos(\omega t + \varphi)$. Комплексная передаточная функция $W(j\omega)$ – это отношение преобразования Фурье выходного гармонического сигнала к преобразованию Фурье входного гармонического сигнала при нулевых начальных условиях:

$$W(j\omega) = \frac{Y(j\omega)}{U(j\omega)}, \quad Y(j\omega) = \int_0^{\infty} y(t)e^{-j\omega t} dt, \quad U(j\omega) = \int_0^{\infty} u(t)e^{-j\omega t} dt.$$

Формально частотные характеристики получаются из передаточной функции $W(s)$, при замене $s = j\omega$, где ω – угловая частота, имеющая размерность [рад/с].

Функция $W(j\omega)$, зависящая от комплексной переменной, называется амплитудно-фазочастотной характеристикой (АФЧХ), она определяет изменение амплитуды и фазы выходного сигнала системы (отдельного элемента) в установившемся режиме по отношению к входному гармоническому воздействию.

Комплексная передаточная функция $W(j\omega)$ может быть представлена в виде:

$$W(j\omega) = P(\omega) + jQ(\omega) = A(\omega)e^{j\varphi(\omega)},$$

где $A(\omega) = |W(j\omega)| = \sqrt{P^2(\omega) + Q^2(\omega)}$, $\varphi(\omega) = \text{arctg} \frac{Q(\omega)}{P(\omega)}$.

$A(\omega) = |W(j\omega)|$ называется амплитудно-частотной характеристикой (АЧХ), она показывает, во сколько раз амплитуда выходного гармонического сигнала отличается от амплитуды входного в зависимости от частоты сигнала.

$\varphi(\omega)$ называется фазовой частотной характеристикой (ФЧХ) и определяет сдвиг фазы между выходным и входным сигналами в зависимости от частоты.

Пример 6.2. Найти и построить частотные характеристики апериодического звена с передаточной функцией $W(s) = \frac{K}{Ts+1}$, где K – коэффициент усиления, T – постоянная времени.

Перейдем от $W(s)$ в частотную область, делая подстановку $s = j\omega$:

$$W(j\omega) = \frac{K}{Tj\omega + 1} = \frac{K(1 - j\omega T)}{(1 + j\omega T)(1 - j\omega T)} = \frac{K}{1 + \omega^2 T^2} - j \frac{K\omega T}{1 + \omega^2 T^2} = P(\omega) - jQ(\omega).$$

Соответствующая АФЧХ изображена на рис. 6.6, а. Каждой точке характеристики соответствует определенное значение частоты. Частота $\omega = \frac{1}{T}$ называется сопрягающей частотой, на этой частоте $P(\omega) = \frac{K}{2}$, $Q(\omega) = -\frac{K}{2}$. Длина вектора, проведенного из начала координат в любую точку характеристики, определяет величину коэффициента усиления на соответствующей частоте. Угол φ характеризует сдвиг фаз между входным и выходным сигналом на этой же частоте. АЧХ определяется выражением:

$$|W(j\omega)| = \sqrt{P(\omega)^2 + Q(\omega)^2} = \frac{K\sqrt{1 + \omega^2 T^2}}{1 + \omega^2 T^2} = \frac{K}{\sqrt{1 + \omega^2 T^2}}.$$

Ее вид представлен на рис. 6.6, б.

ФЧХ определяется выражением:

$$\varphi(\omega) = \text{arctg} \frac{Q(\omega)}{P(\omega)} = \text{arctg}(-\omega T) = -\text{arctg}(\omega T).$$

Ее вид показан на рис. 6.6, в.

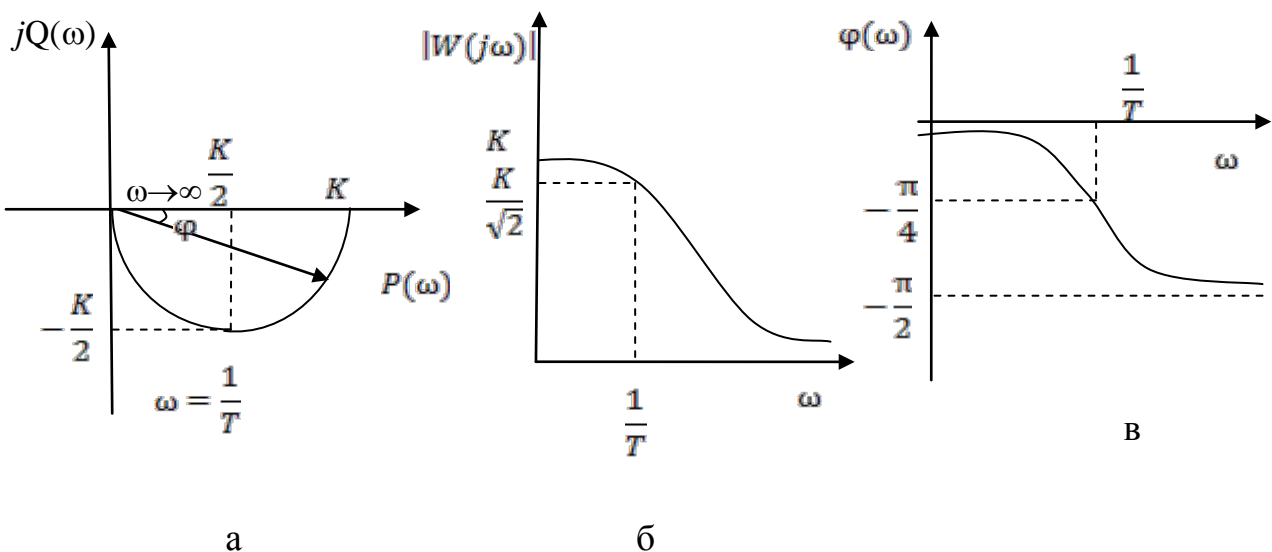


Рис. 6.6. Частотные характеристики апериодического звена

6.7. Понятие о логарифмических частотных характеристиках

Частотные характеристики удобно строить в логарифмическом масштабе, так как существенно сокращается объем вычислений. При этом все степенные функции становятся линейными, а наклон линии определяется показателем степени соответствующей характеристики. При построении логарифмических частотных характеристик (ЛЧХ) по оси абсцисс откладываются значения частот ω в логарифмическом масштабе. Отрезок этой оси, соответствующий изменению частоты ω в десять раз, называется декадой. Длина отрезка, равного декаде, не зависит от частоты и определяется выражением:

$$\lg 10\omega - \lg \omega = 1.$$

На рис. 6.7 отложены 3 декады, которые соответствуют изменениям частоты от 0,1 до 1, от 1 до 10 и от 10 до 100.

Для разметки логарифмической шкалы внутри каждой декады следует пользоваться значениями десятичных логарифмов чисел, которые приведены в таблице 6.1.

Таблица 6.1

X	1	2	3	4	5	6	7	8	9	10
lg	0	0,301	0,47	0,602	0,7	0,78	0,85	0,9	0,95	1

Значение логарифма произвольного числа показывает какую часть декады следует взять от ее начала, чтобы обозначит данные числа на оси ω (рис. 6.7).

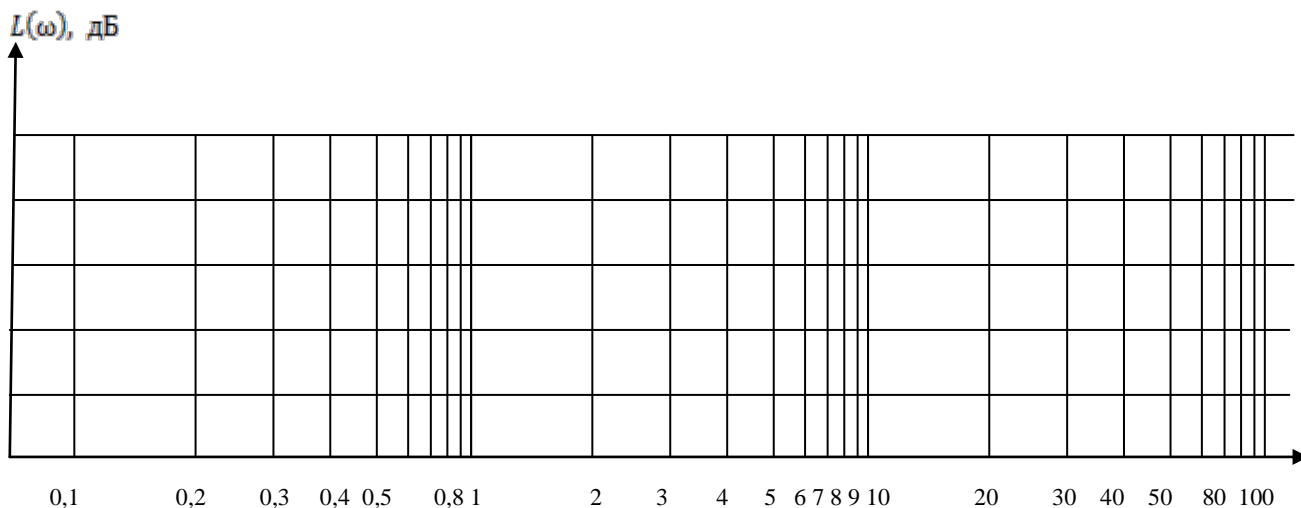


Рис. 6.7. Типовая масштабная сетка

По оси ординат при построении логарифмической амплитудно-частотной характеристики (ЛАЧХ) откладывают в равномерном масштабе значение коэффициента усиления, выраженное в децибелах и вычисляется по формуле

$L(\omega) = 20\lg|W(j\omega)| = 20\lg A(\omega)$. Так как логарифма нуля не существует, нуль на оси ω отсутствует и ось ординат является плавающей. Она может быть проведена через любое значение ω . Наклон отрезков прямых линий, из которых состоят ЛАЧХ, определяется по изменению усиления на интервале частот, соответствующем одной декаде.

Интегрирующее звено с передаточной функцией $W(s) = \frac{K}{s}$ имеет следующие частотные характеристики:

$$W(j\omega) = \frac{K}{j\omega} = \frac{K}{\omega} e^{-j\frac{\pi}{2}}; \quad A(\omega) = \frac{K}{\omega}; \quad \varphi(\omega) = -\frac{\pi}{2}.$$

Тогда при переходе к ЛАЧХ:

$$L(\omega) = 20\lg K - 20\lg\omega.$$

Вычислим значения $L(\omega)$ при $\omega = 10$ и $\omega = 1$ и найдем как измениться $L(\omega)$ при изменении частоты на одну декаду. При $\omega = 1$ $L(\omega) = 20\lg K$; при $\omega = 10$ $L(\omega) = 20\lg K - 20$, т. е. усиление уменьшится на 20 дБ и наклон характеристики будет равен -20 дБ/дек во всем диапазоне частот, как показано на рис. 6.8.

Фазовая характеристика $\varphi(\omega)$ интегрирующего звена не зависит от частоты и равна $\varphi(\omega) = -\frac{\pi}{2}$.

Дифференцирующее звено с передаточной функцией $W(s) = Ks$ имеет следующие частотные характеристики:

$$W(j\omega) = K \cdot j\omega = K\omega e^{j\frac{\pi}{2}}; \quad A(\omega) = K\omega; \quad \varphi(\omega) = \frac{\pi}{2}.$$

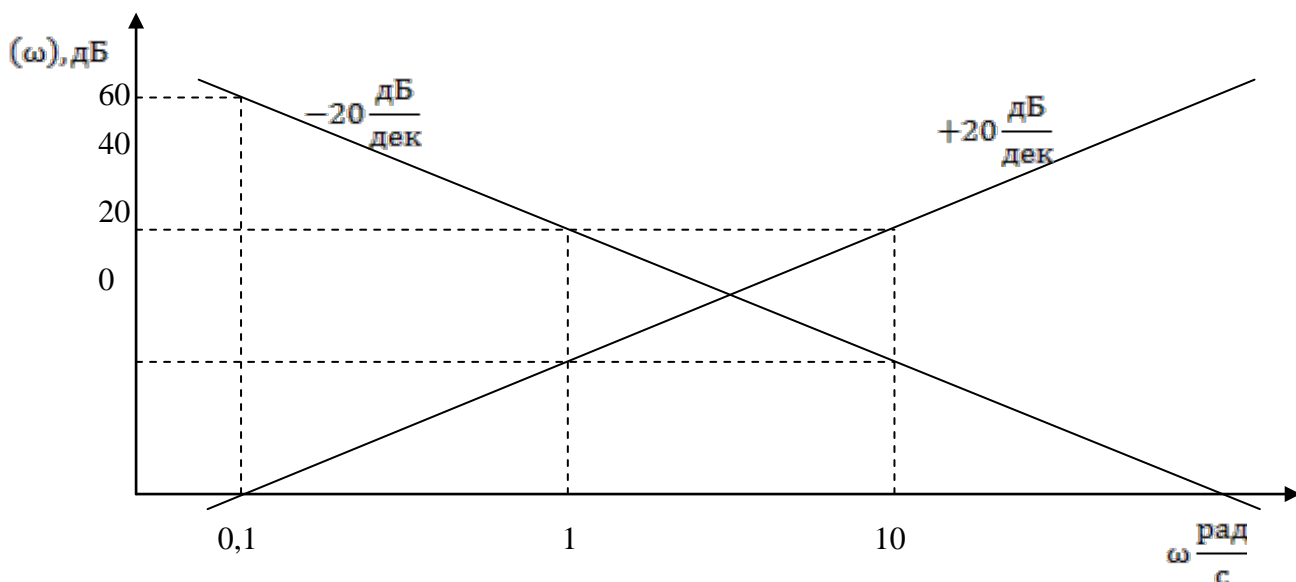


Рис. 6.8. ЛАЧХ интегрирующего звена с $W(s) = \frac{100}{s}$ и дифференцирующего звена с $W(s) = 10 \cdot s$

При переходе к ЛАЧХ $L(\omega) = 20\lg A(\omega) = 20\lg K + 20\lg\omega$.

Вычислим изменение $L(\omega)$ при изменении частоты на одну декаду, задавая $\omega = 1$ и $\omega = 10$. При $\omega = 1$ $L(\omega) = 20\lg K$; при $\omega = 10$ $L(\omega) = 20\lg K + 20$, т. е. усиление увеличится на 20 дБ и наклон ЛАЧХ будет равен -20 дБ/дек во всем диапазоне частот, как показано на рисунке 6.8.

Апериодическое звено с передаточной функцией $W(s) = \frac{K}{Ts+1}$ имеет следующие характеристики:

$$W(j\omega) = \frac{K}{Tj\omega + 1} = \frac{K}{\sqrt{T^2\omega^2 + 1}} e^{j\varphi(\omega)}; \quad A(\omega) = \frac{K}{\sqrt{T^2\omega^2 + 1}};$$

$$\varphi(\omega) = -\arctg\omega T.$$

Здесь T – постоянная времени звена.

При переходе к ЛАЧХ:

$$L(\omega) = 20\lg A(\omega) = 20\lg K - 20\lg\sqrt{T^2\omega^2 + 1} = 20\lg K - 10\lg(T^2\omega^2 + 1).$$

Частота $\omega = \frac{1}{T}$ называется сопрягающей частотой. Для построения асимптотической (приближенной) ЛАЧХ рассмотрим два диапазона частот: $\omega \ll \frac{1}{T}$ и $\omega \gg \frac{1}{T}$. В первом случае $\omega T \ll 1$ и слагаемым $\omega^2 T^2$ в выражении для $L(\omega)$ можно пренебречь, тогда $L(\omega) = 20\lg(K)$. Во втором случае $\omega T \gg 1$ и, пренебрегая единицей в скобках в выражении для $L(\omega)$, получим $L(\omega) = 20\lg(K) - 10\lg(T^2\omega^2) = 20\lg(K) - 20\lg(T\omega)$. Таким образом, до сопрягающей частоты $\omega = \frac{1}{T}$ ЛАЧХ идет параллельно оси частот на уровне $20\lg(K)$, а начиная с $\omega = \frac{1}{T}$ характеристика идет под наклоном -20 дБ/дек, как показано на рисунке 6.9, а. Для упрощения обозначения наклоны ± 20 дБ/дек будут обозначаться через ± 1 , а наклоны $\pm m \cdot 20$ дБ/дек через $\pm m$.

Для построения асимптотической ФЧХ рассмотрим ее значения при $\omega = 0$, $\omega = \frac{1}{T}$ и $\omega \rightarrow \infty$. При $\omega = 0$ $\varphi(\omega) = 0$; при $\omega = \frac{1}{T}$ $\varphi(\omega) = -\arctg 1 = -\frac{\pi}{4}$; при $\omega \rightarrow \infty$ $\varphi(\omega) \rightarrow -\frac{\pi}{2}$. По этим характерным точкам и строится ФЧХ, как показано на рисунке 6.9, б. При построении ФЧХ ось частот размечается так же, а по оси ординат откладывается фаза в градусах или в радианах в линейном масштабе.

Частота ω_c , при которой $L(\omega) = 0$ носит название частоты среза, она соответствует коэффициенту усиления $K = 1$. Следовательно, при $\omega < \omega_c$ происходит усиление ($K > 1$), а при $\omega > \omega_c$ ослабление ($K < 1$) гармонического входного сигнала.

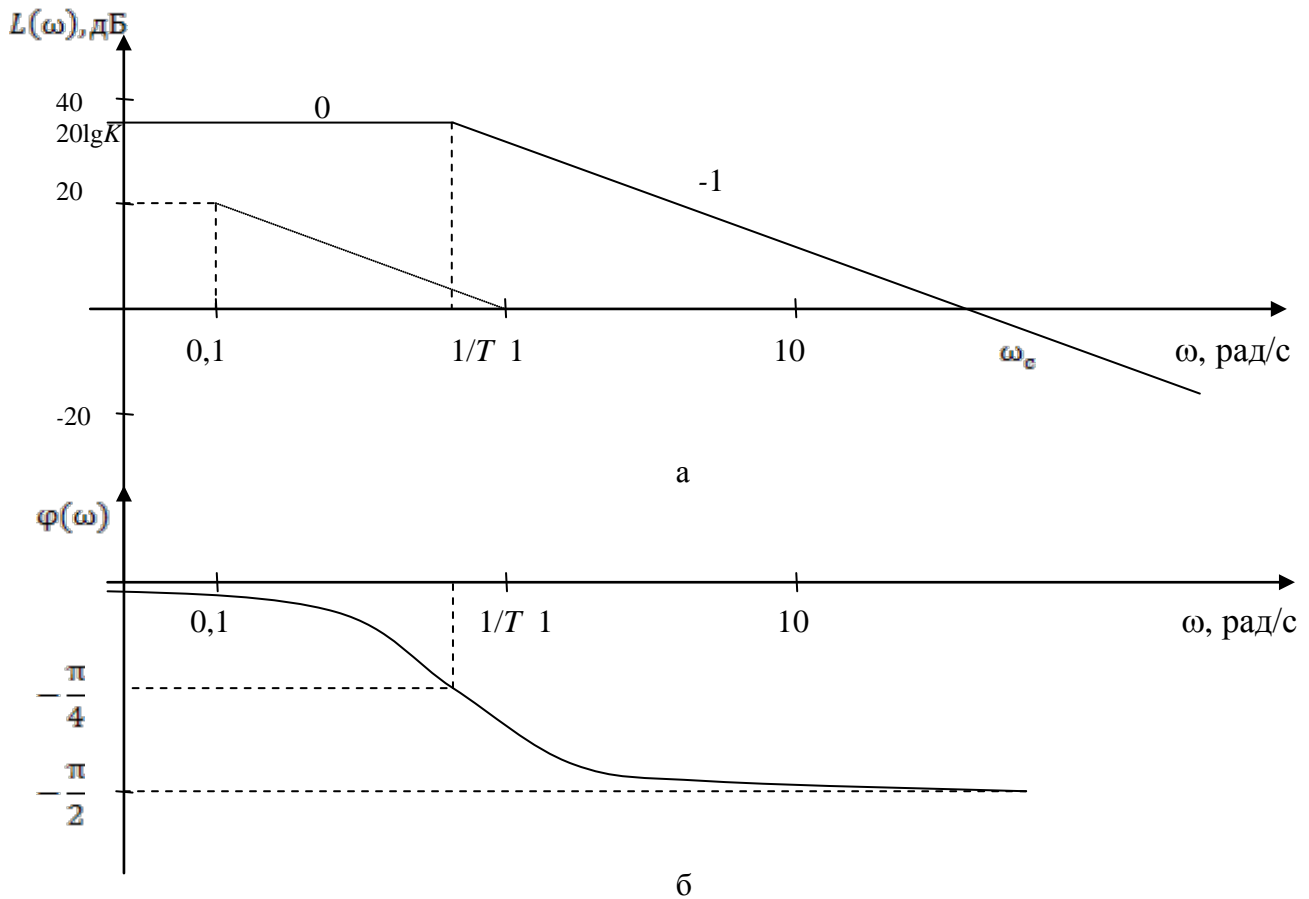


Рис. 6.9. ЛЧХ аperiodического звена: а – ЛАЧХ, б – ЛФЧХ.

Форсирующее звено первого порядка с передаточной функцией $W(s) = K(Ts + 1)$ имеет следующие частотные характеристики:

$$W(j\omega) = K(Tj\omega + 1), \quad A(\omega) = K\sqrt{T^2\omega^2 + 1}, \quad \varphi(\omega) = \text{arctg}(\omega T).$$

При переходе к ЛАЧХ:

$$L(\omega) = 20\lg K + 20\lg\sqrt{T^2\omega^2 + 1} = 20\lg K + 10\lg(T^2\omega^2 + 1).$$

Как и в предыдущем примере рассмотрим диапазон частот справа и слева от сопрягающей частоты $\omega = \frac{1}{T}$, где T – постоянная времени звена. При $\omega \ll \frac{1}{T}$ имеем $T^2\omega^2 \ll 1$ и слагаемым $T^2\omega^2$ в выражении для $L(\omega)$ можно пренебречь, тогда $L(\omega) = 20\lg K$. При $\omega \gg \frac{1}{T}$ значение $T^2\omega^2 \gg 1$, тогда пренебрегая единицей в скобках в выражении для $L(\omega)$ получим $L(\omega) = 20\lg K + 20\lg \omega T$. Таким образом асимптотическая ЛАЧХ представляет собой две прямые линии, сопрягающиеся в точке $\omega = \frac{1}{T}$ (рис. 6.10, а). До сопрягающей частоты $\omega = \frac{1}{T}$ ЛАЧХ

идет параллельно оси частот на уровне $20\lg K$. После этой частоты характеристика идет под наклоном $+20$ дБ/дек, это означает, что при изменении частоты на одну декаду усиление увеличится на 20 дБ, $\varphi(\omega)$ изменяется от 0 до $\frac{\pi}{2}$, при $\omega = \frac{1}{T}$ $\varphi(\omega) = \arctg 1 = \frac{\pi}{4}$ (рис. 6.10, б).

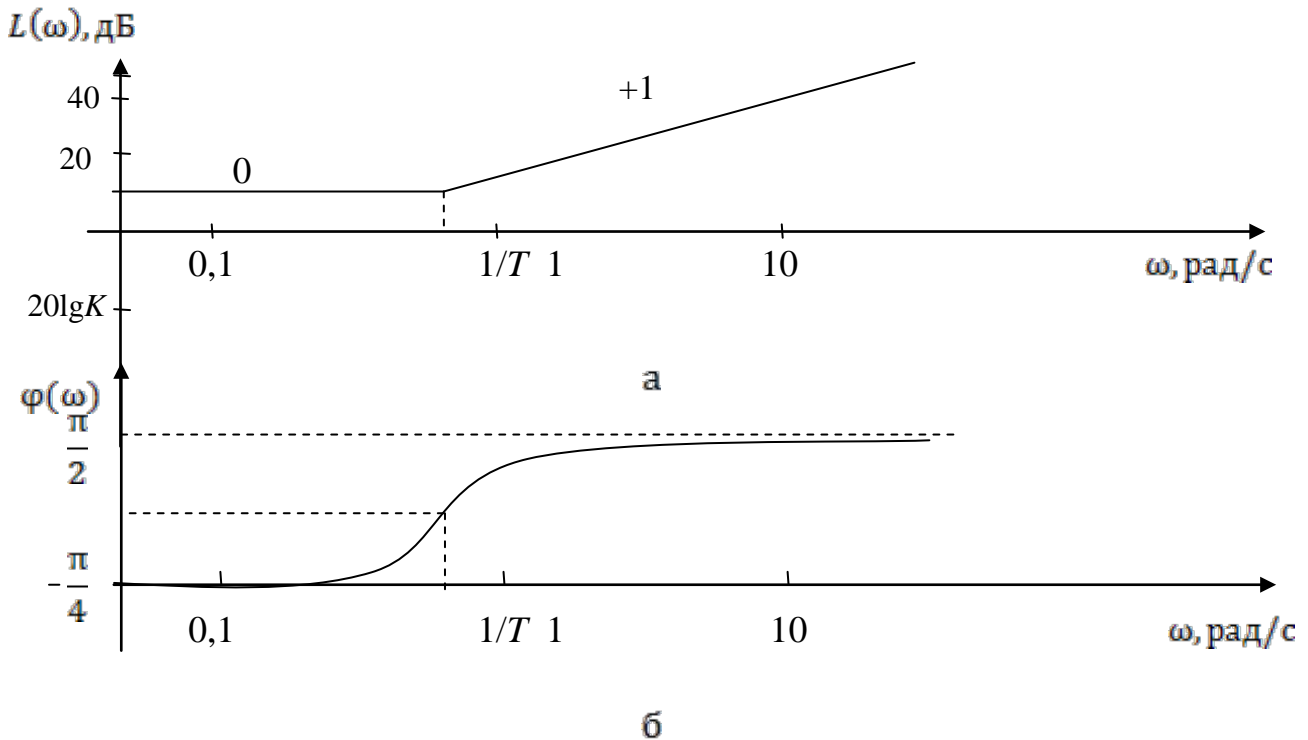


Рис. 6.10. ЛЧХ форсирующего звена: а – ЛАЧХ, б – ЛФЧХ.

ЛАЧХ, изображенные на рис. 6.9, а и 6.10, а, являются асимптотически. Максимальная погрешность Δ между точной и асимптотической ЛАЧХ соответствует частоте $\omega = \frac{1}{T}$ и равна $\Delta = 20\lg\sqrt{T^2\omega^2 + 1} = 20\lg\sqrt{2} \cong 3,03$ дБ, так как именно под знаком корня поочередно пренебрегали одним из слагаемых при построении асимптотической ЛАЧХ.

Колебательное звено с передаточной функцией $W(s) = \frac{K}{T^2s^2 + 2\xi Ts + 1}$.

Это звено второго порядка. Здесь ξ – показатель колебательности, $0 < \xi < 1$.

Частотные характеристики имеют вид:

$$W(j\omega) = \frac{K}{T^2(j\omega)^2 + 2\xi Tj\omega + 1}; \quad A(\omega) = \frac{K}{\sqrt{(1 - T^2\omega^2)^2 + 4\xi^2 T^2\omega^2}};$$

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

При переходе к ЛАЧХ получим

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

Удобно строить ЛЧХ для $\xi = 1$, при этом передаточная функция $W(s) = \frac{K}{(Ts+1)^2}$ соответствует последовательному соединению двух апериодических звеньев, а ЛАЧХ этих звеньев будут складываться. Результирующая ЛАЧХ приведена на рис. 6.11, а. Переход от прямой проведенной на уровне $20\lg K$ с наклоном 0 дБ/дек на прямую с наклоном -40 дБ/дек происходит на сопрягающей частоте $\omega = \frac{1}{T}$. Фазовая характеристика изменяется в пределах от 0 до $-\pi$, при $\omega = \frac{1}{T}$ $\varphi(\omega) = -\frac{\pi}{2}$ (рис. 6.11, б).

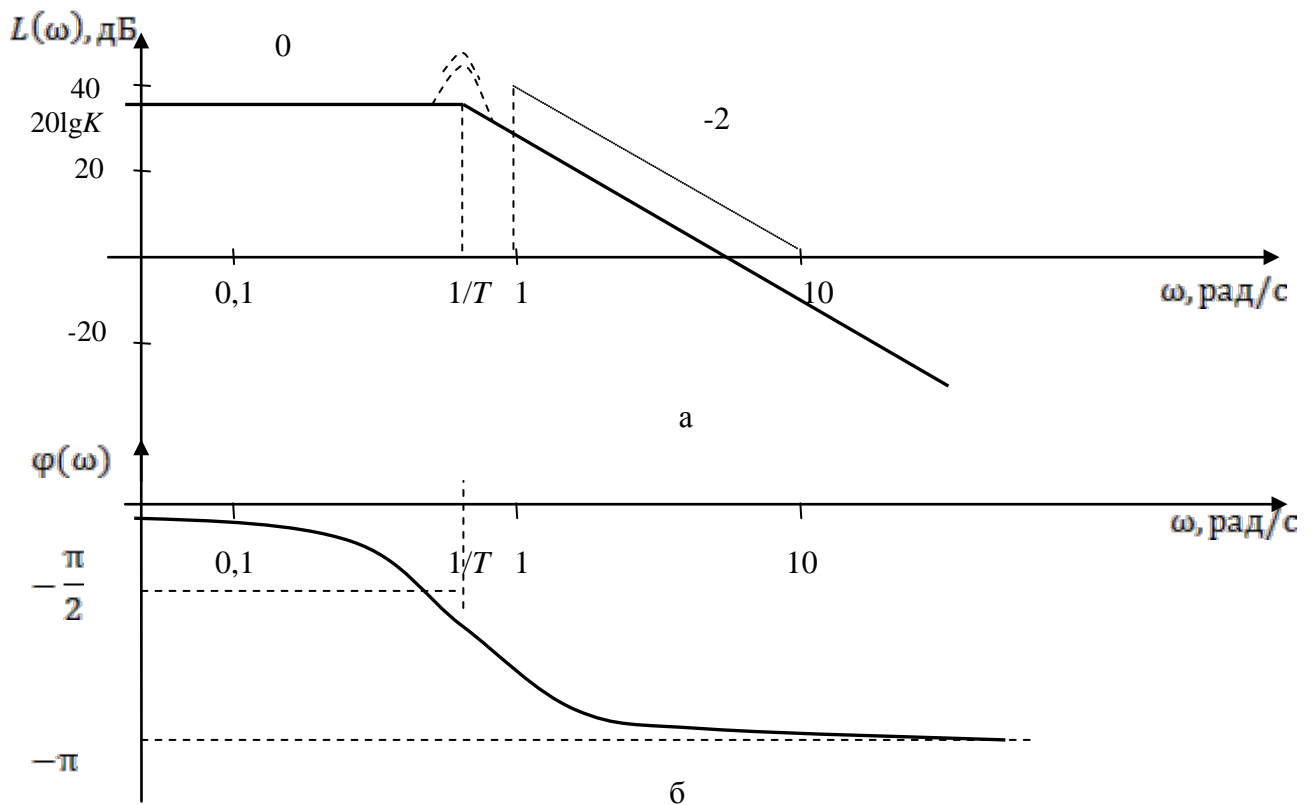


Рис. 6.11. ЛЧХ колебательного звена: а – ЛАЧХ, б – ЛФЧХ

При точном построении ЛАЧХ колебательного звена на частоте $\omega = \frac{1}{T}$ возникает «всплеск», величина которого зависит от ξ (штриховые линии на рис. 6.11, а). Существуют номограммы, которые позволяют уточнить вид ЛАЧХ с учетом значения ξ .

6.8. Построение логарифмических частотных характеристик разомкнутых одноконтурных систем

Предварительно передаточную функцию системы представляют в виде произведения передаточных функций отдельных звеньев. Тогда ЛЧХ системы формируется суммированием ЛЧХ отдельных звеньев.

Рассмотрим систему с передаточной функцией

$$W(s) = \frac{K_1(1 + T_1s)(1 + T_2s) \dots (1 + T_js)}{s^v(1 + T_{j+1}s)(1 + T_{j+2}s) \dots (1 + T_{j+i}s)}, \quad (6.7)$$

которую можно рассматривать как последовательное соединение v – интегрирующих звеньев, j – форсирующих звеньев и i – апериодических звеньев. Если есть идеальные дифференцирующие звенья, то v будет отрицательным числом.

При переходе к логарифмическому масштабу логарифмические частотные характеристики отдельных звеньев складываются, на этом основано **правило построения ЛАЧХ разомкнутой системы**.

Последовательность действий следующая:

1. Определить все сопрягающие частоты $\omega_i = \frac{1}{T_i}$ и отметить их на оси абсцисс.
2. Первая линия в области низких частот проводится через точку с координатами $[\omega = 1; L = 20 \lg K]$ с наклоном $(-20 \text{ дБ/дек}) \cdot v$. Линия проводится до первой сопрягающей частоты (минимальная на оси частот).
3. Продолжить построение, изменяя наклон после каждой из сопрягающих частот в зависимости от того, какому звену эта частота принадлежит. Каждое апериодическое звено изменяет наклон на -20 дБ/дек , а каждое форсирующее на $+20 \text{ дБ/дек}$.
4. Выражение для фазочастотной характеристики имеет вид:

$$\varphi(\omega) = -90^\circ \cdot v + \sum_{k=1}^j \text{arctg} \omega T_k - \sum_{l=j+1}^{j+i} \text{arctg} \omega T_l.$$

Оно определяется как сумма значений ФЧХ каждого из элементов системы на фиксированной частоте.

Пример 6.3. Построить ЛАЧХ и ЛФЧХ для системы с передаточной функцией

$$W(s) = \frac{100(1 + 0,5s)}{s(1 + 10s)}.$$

Сопрягающие частоты: $\omega_1 = \frac{1}{10} = 0,1$; $\omega_2 = \frac{1}{0,5} = 2$. Первый наклон (-1) определяется наличием идеального интегрирующего звена и проведен через точку с координатами $\omega = 1$, $L(\omega) = 20\lg 100 = 40$. На частоте $\omega = 0,1$ наклон ЛАЧХ изменяется до (-2), так как $\omega = 0,1$ является сопрягающей частотой для аperiodического звена. На частоте $\omega = 2$ наклон ЛАЧХ становится равным (-1), так как $\omega = 2$ является сопрягающей частотой принадлежащей форсирующему звену. ЛФЧХ получена в результате сложения характеристик отдельных звеньев: интегрирующего, аperiodического и форсирующего.

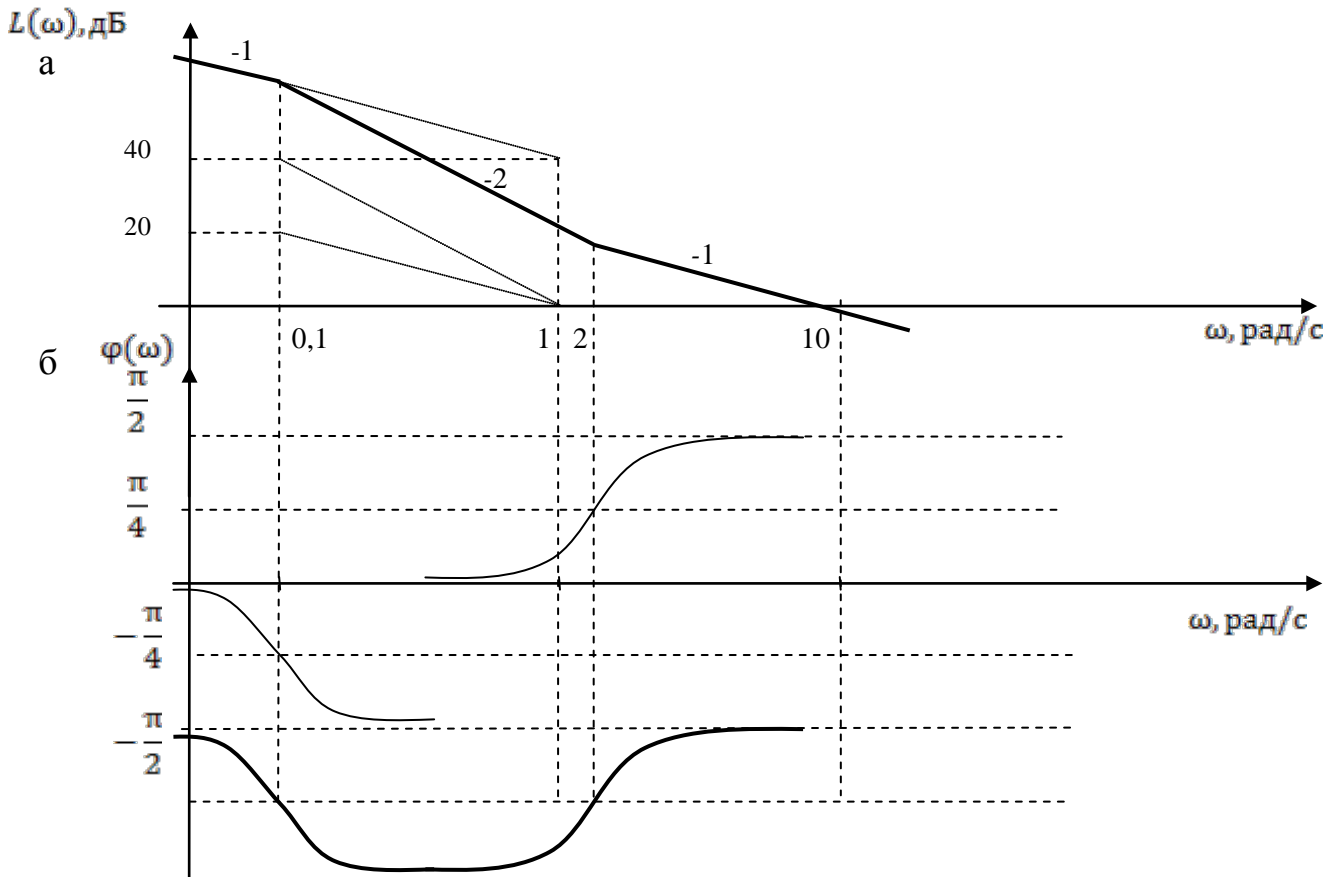


Рис. 6.12. Пример построения ЛЧХ разомкнутой системы.

6.9. Математические модели элементов в параметрах пространства состояний

Для отдельного элемента или системы кроме входных и выходных переменных можно выделить некоторую совокупность промежуточных переменных, которые связаны с внутренней структурой устройства, обозначаются $x(t)$ и называются переменными состояния.

Описание при помощи переменных состояния позволяет представить многие физические системы совокупностями дифференциальных уравнений

первого порядка вида $x(t) = f[\dot{x}(t), u(t), t]$, которые называются **дифференциальными уравнениями состояния**, и совокупностями алгебраических уравнений вида $y(t) = g[x(t), u(t), t]$, которые называются **уравнениями вход-состояние-выход**. В этих уравнениях $u(t)$ – входные переменные, $y(t)$ – выходные переменные, $x(t)$ – переменные состояния. В общем случае подразумевается, что система имеет несколько входов и несколько выходов. В дальнейшем рассматриваются линейные стационарные устройства и системы с одним входом и одним выходом. Количество переменных состояния n определяется порядком системы. Все состояния системы образуют вектор состояний, множество которых образует **пространство состояний**.

Математическая модель элементов, описываемых уравнениями первого порядка, имеет вид:

$$\begin{aligned} \dot{x}(t) &= ax(t) + bu(t), \\ y(t) &= cx(t) + du(t), \end{aligned} \quad (6.8)$$

где a , b , c и d – постоянные числа.

Уравнения (6.8) удобно представить в виде структурной схемы или схемы моделирования. Построение схемы (рис. 6.13) начинается с изображения интегратора, который обозначается соответствующим знаком.

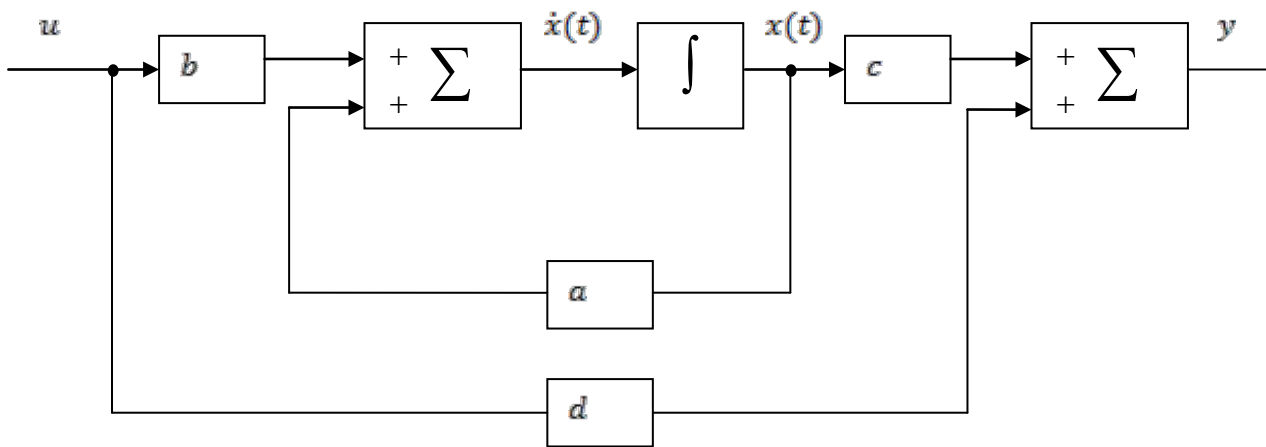


Рис. 6.13. Схема моделирования элемента первого порядка

Переменная состояния $x(t)$ ассоциируется с выходом интегратора, тогда сигнал на входе интегратора можно обозначить через $\dot{x}(t)$. Для формирования сигнала $x(t)$ нужно проинтегрировать его производную $\dot{x}(t)$ и замкнуть контур из условия удовлетворения дифференциальному уравнению $\dot{x}(t) = ax(t) + by(t)$. Контур с блоком a называется контуром обратной связи. Выходной сигнал $y(t)$ является выходом сумматора, на который подаются сигналы $cx(t)$ и $du(t)$. Эле-

менты схемы, внутри которых записаны коэффициенты a, b, c и d моделируют блоки преобразования сигналов с соответствующими коэффициентами усиления.

6.10. Решение уравнений состояния первого порядка

Вначале найдем решение уравнений состояния первого порядка с постоянными коэффициентами

$$\begin{cases} \dot{x} = ax + bu, \\ y = cx + du. \end{cases}$$

Решение первого уравнения складывается из общего решения $x_1(t)$ однородного уравнения $\dot{x} = ax$ (с нулевым входным воздействием, т. е. $u = 0$) и частного решения $x_2(t)$ неоднородного уравнения ($u \neq 0$).

$$x(t) = x_1(t) + x_2(t). \quad (6.9)$$

Решение однородного уравнения найдем методом разделения переменных

$$\frac{dx}{dt} = ax; \quad \frac{dx}{x} = a dt,$$

тогда $\ln x = at + \xi_1$, $x = \xi \cdot e^{at}$, где ξ – постоянная интегрирования, которая находится из начальных условий.

Пусть в момент времени t_0 имеем $x(t_0)$, тогда

$$x(t_0) = \xi \cdot e^{at_0}, \quad (6.10)$$

отсюда $\xi = x(t_0)e^{-at_0}$ и $x_1(t) = x(t_0)e^{a(t-t_0)}$ – общее решение однородного уравнения, оно не зависит от входного воздействия и определяется только параметрами системы. Так как $x_1(t)$ характеризует естественное поведение системы при отсутствии внешних сигналов (возмущений), его называют свободным или невынужденным движением.

Частное решение можно найти методом *вариации постоянной*. Для этого постоянную ξ в (6.10) заменяют неизвестной функцией $g(t)$, тогда $x(t) = g(t)e^{at}$.

Подставив $x(t)$ в уравнение $\dot{x} = ax + bu$, получим

$$g(t) = \int_{t_0}^t bu(\tau)e^{a(t-\tau)} d\tau,$$

тогда

$$x_2(t) = \int_{t_0}^t bu(\tau)e^{a(t-\tau)} d\tau,$$

где $x_2(t)$ – частное решение, оно зависит как от параметров системы, так и от входного воздействия, его называют вынужденным движением.

Решение уравнения (6.9) имеет вид

$$x(t) = x_1(t) + x_2(t) = x(t_0)e^{a(t-t_0)} + \int_{t_0}^t bu(\tau)e^{a(t-\tau)}d\tau. \quad (6.11)$$

Выходной сигнал $y = Cx + du$ находится как решение алгебраического уравнения.

6.11. Представление уравнений состояния при помощи матриц

Если линейная стационарная система описывается дифференциальным уравнением n -го порядка, то ее математическая модель в параметрах пространства состояний имеет вид

$$\begin{cases} \dot{x}(t) = Ax(t) + Bu(t), \\ y(t) = Cx(t) + Du(t), \end{cases}$$

где u – вход; y – выход; x – переменные состояния; A, B, C, D – матрицы постоянных коэффициентов.

Будем рассматривать линейные системы с одним входом и одним выходом. Для системы n -го порядка $x^T = [x_1, x_2, \dots, x_n]$ вектор состояния, т. е. количество переменных состояния равно порядку системы. Рассмотрим систему, описываемую дифференциальным уравнением второго порядка

$$\frac{d^2y}{dt^2} + \alpha_1 \frac{dy}{dt} + \alpha_0 y = u.$$

Необходимо описать систему в параметрах пространства состояний.

Для построения схемы моделирования линейного дифференциального уравнения любого порядка, необходимо последовательно проинтегрировать наивысшие производные уравнения, получить все производные низшего порядка и саму переменную y .

Выразим

$$\frac{d^2y}{dt^2} = u - \alpha_1 \frac{dy}{dt} - \alpha_0 y.$$

Количество интеграторов равно порядку системы и равно двум. Система замыкается исходя из условия удовлетворения дифференциальному уравнению (рис. 6.14).

В качестве переменных состояния удобно выбрать выходы интеграторов: $x_1 = y$; $x_2 = \dot{y}$. Тогда можно перейти к следующей системе уравнений:

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = u - \alpha_1 x_2 - \alpha_0 x_1, \\ y = x_1. \end{cases} \quad (6.12)$$

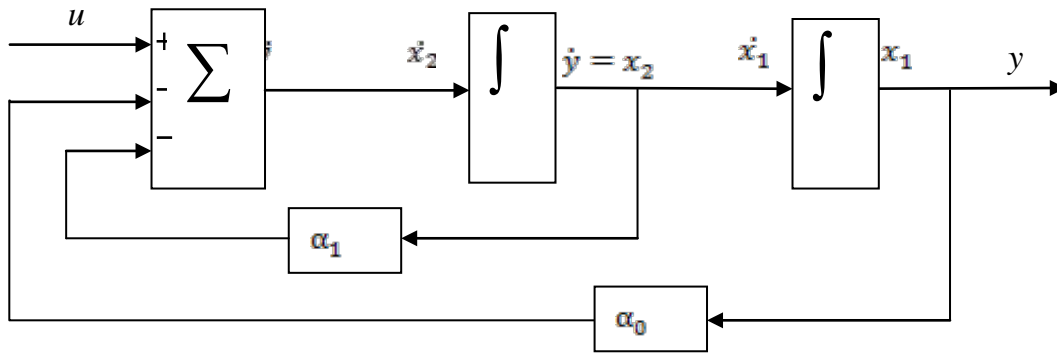


Рис. 6.14. Схема моделирования системы второго порядка.

Систему (6.12) можно записать в матричном виде:

$$\begin{cases} \begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -\alpha_0 & -\alpha_1 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u; \\ y = [1 \ 0] \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} + [0]u. \end{cases} \quad (6.13)$$

где $A = \begin{bmatrix} 0 & 1 \\ -\alpha_0 & -\alpha_1 \end{bmatrix}$, $B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}$, $C = [1 \ 0]$, $D = 0$.

Таким образом, введение переменных состояния позволяет линейное дифференциальное уравнение второго порядка представить в виде системы двух дифференциальных уравнений первого порядка.

Рассмотрим обобщение методики на систему, описываемую дифференциальным уравнением n -го порядка

$$\frac{d^n y}{dt^n} + \alpha_{n-1} \frac{d^{n-1} y}{dt^{n-1}} + \dots + \alpha_1 \frac{dy}{dt} + \alpha_0 y = \beta_m \frac{d^m u}{dt^m} + \beta_n \frac{d^{m-1} u}{dt^{m-1}} + \dots + \beta_0 u. \quad (6.14)$$

Коэффициент при высшей производной по y всегда можно привести к единице. Предположим вначале, что все $\beta_i = 0$, $i = \overline{1, n}$ и только $\beta_0 \neq 0$. Введем переменные состояния следующим образом:

$$x_1 = y; \quad x_2 = \frac{dy}{dt}; \quad \dots; \quad x_n = \frac{d^{n-1} y}{dt^{n-1}}$$

и выразим высшую производную из (6.14). Тогда

$$\frac{d^n y}{dt^n} = \beta_0 u - \alpha_0 x_1 - \alpha_1 x_2 - \dots - \alpha_{n-1} x_n.$$

Переходя к системе уравнений первого порядка, получим

$$\begin{cases} \dot{x}_1 = x_2, \\ \dot{x}_2 = x_3, \\ \dots \dots \dots \\ \dot{x}_{n-1} = x_n, \\ \dot{x}_n = \beta_0 u - \alpha_0 x_1 - \alpha_1 x_2 - \dots - \alpha_{n-1} x_n, \\ y = x_1. \end{cases} \quad (6.15)$$

Схема моделирования, соответствующая системе уравнений (6.15) изображена на рис. 6.15.

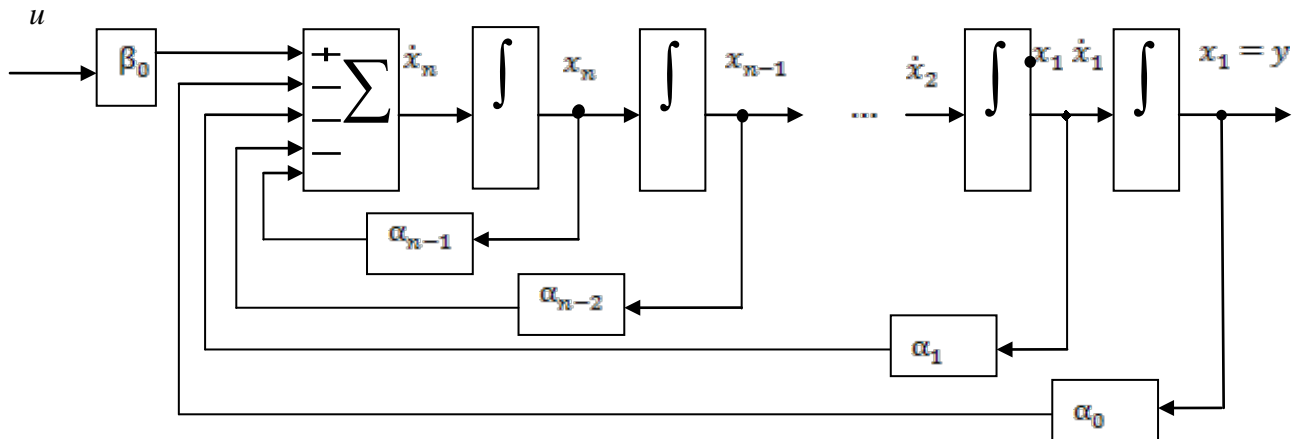


Рис. 6.15 . Схема моделирования системы n -го порядка

Систему уравнений (6.15) можно представить в матричном виде следующим образом

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \vdots \\ \dot{x}_{n-1} \\ \dot{x}_n \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \dots & -\alpha_{n-1} \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_{n-1} \\ x_n \end{bmatrix} + \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ \beta_0 \end{bmatrix} u,$$

$$y = [1 \ 0 \ 0 \ \dots \ 0] \begin{bmatrix} x_1 \\ x_2 \\ \vdots \\ x_n \end{bmatrix} + [0]u.$$

Здесь матрица A коэффициентов при x в дифференциальных уравнениях имеет вид

$$A = \begin{bmatrix} 0 & 1 & 0 & \dots & 0 \\ 0 & 0 & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \dots & 1 \\ -\alpha_0 & -\alpha_1 & -\alpha_2 & \dots & -\alpha_{n-1} \end{bmatrix}$$

и называется матрицей Фробениуса. Элементы этой матрицы, расположенные над главной диагональю – единицы, элементы нижней строки – коэффициенты левой части дифференциального уравнения, взятые с противоположным знаком, все остальные элементы – нули.

Вектор строка C имеет только один единичный элемент, остальные нули. Если в уравнении (6.14) все β_i отличны от нуля и порядок n в левой и правой части одинаковый (т.е. $m = n$), то вид матриц A и C остается таким же, а матрицы

$$D = b_0 \text{ и } B = \begin{bmatrix} b_1 \\ b_2 \\ \vdots \\ b_n \end{bmatrix} \text{ вычисляются по следующим рекуррентным соотношениям}$$

$$\begin{aligned} b_0 &= \beta_n, \\ b_1 &= \beta_{n-1} - \alpha_{n-1} b_0, \\ b_2 &= \beta_{n-2} - \alpha_{n-1} b_1 - \alpha_{n-2} b_0, \\ &\dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \dots \\ b_n &= \beta_0 - \alpha_{n-1} b_{n-1} - \alpha_{n-2} b_{n-2} - \dots - \alpha_0 b_0. \end{aligned}$$

Рассмотренная форма уравнений состояния, в которой матрица A является матрицей Фробениуса, называется **нормальной формой уравнения состояния**.

Пример 6.4. Дифференциальное уравнение системы имеет вид

$$\frac{d^3 y}{dt^3} + 3 \frac{d^2 y}{dt^2} + 4 \frac{dy}{dt} + y = 2 \frac{d^3 u}{dt^3} + 3 \frac{d^2 u}{dt^2} + \frac{du}{dt} + 2u.$$

Привести уравнение к нормальной форме уравнений состояния, изобразить схему моделирования.

Матрица A не требует вычислений и имеет вид $A = \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -3 & -4 \end{pmatrix}$, матрица $C = (1 \ 0 \ 0)$.

Найдем элементы матриц B и D :

$$\begin{aligned} b_0 &= \beta_3 = 2, \\ b_1 &= \beta_2 - \alpha_2 b_0 = 3 - 3 \cdot 2 = -3, \\ b_2 &= \beta_1 - \alpha_2 b_1 - \alpha_1 b_0 = 1 - 3 \cdot (-3) - 4 \cdot 2 = 2, \\ b_3 &= \beta_0 - \alpha_2 b_2 - \alpha_1 b_1 - \alpha_0 b_0 = 2 - 3 \cdot 2 - 4 \cdot (-3) - 1 \cdot 2 = 6. \end{aligned}$$

Соответствующие уравнения состояния и имеют вид:

$$\begin{bmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \end{bmatrix} = \begin{bmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ -1 & -4 & -3 \end{bmatrix} \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + \begin{bmatrix} -3 \\ 2 \\ 6 \end{bmatrix} u,$$

$$y = [1 \ 0 \ 0] \begin{bmatrix} x_1 \\ x_2 \\ x_3 \end{bmatrix} + 2u.$$

или в скалярной форме:

$$\begin{cases} \dot{x}_1 = x_2 - 3u, \\ \dot{x}_2 = x_3 - 2u, \\ \dot{x}_3 = 6u - x_1 - 4x_2 - 3x_3, \\ y = x_1 + 2u. \end{cases}$$

Схема моделирования приведена на рис. 6.16.

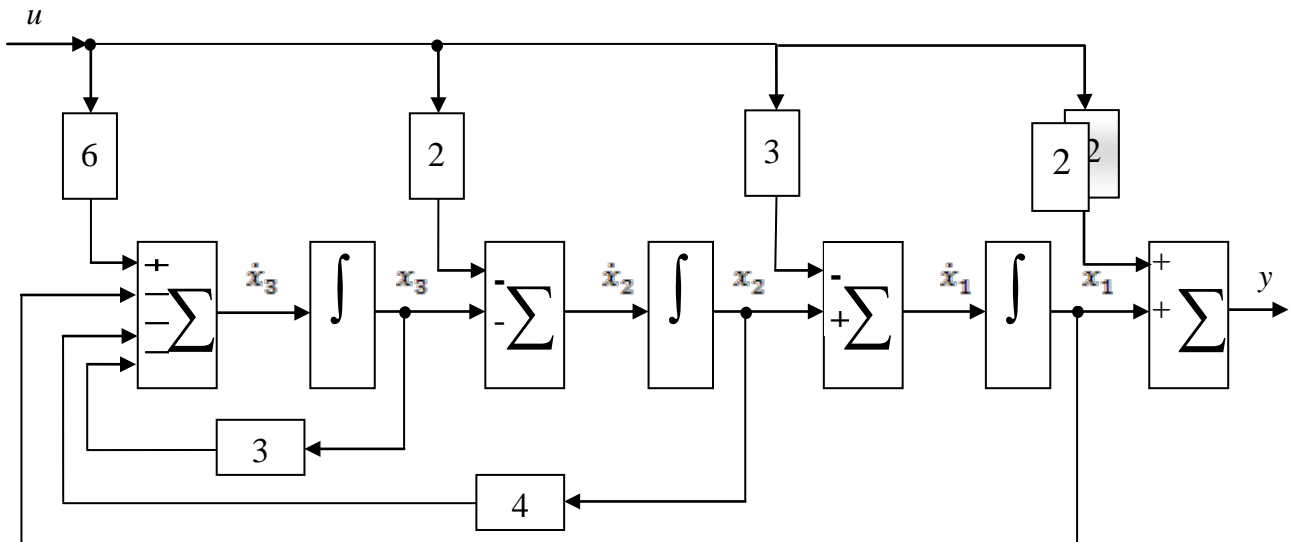


Рис. 6.16. Схема моделирования к примеру 6.16

6.12. Решение матричных уравнений состояния, представленных в нормальной форме

Рассмотрим матричное уравнение $\dot{x} = Ax + Bu$, где A – квадратичная матрица порядка n , B – вектор столбец, x – вектор переменных состояния. Решение этого уравнения запишем по аналогии с выражением (6.11) для системы первого порядка

$$x^T(t) = x^T(t_0) \cdot e^{A(t-t_0)} + \int_{t_0}^t B^T u(\tau) \cdot e^{A(t-\tau)} d\tau,$$

Здесь e^{At} – матричная экспоненциальная функция, ее принято обозначать как $e^{At} = \Phi(t)$. $\Phi(t)$ – квадратная матрица порядка n , которая называется фундаментальной матрицей или переходной матрицей состояния. Матрица $\Phi(t)$ обладает следующими свойствами:

1. $\Phi(0) = E$ (единичная матрица),
2. $\Phi^{-1}(t) = \Phi(-t)$,
3. $\Phi(t_2 - t_0) = \Phi(t_2 - t_1) \cdot \Phi(t_1 - t_0)$, если $x(t_1)$ является промежуточным состоянием между $x(t_0)$ и $x(t_2)$.

Полагая $t_0 = 0$, выражение (6.16) перепишем следующим образом

$$x^T(t) = x^T(0) \cdot \Phi(t) + \int_0^t B^T u(\tau) \Phi(t - \tau) d\tau, \quad (6.16)$$

где первое слагаемое определяет свободную составляющую решения, зависящую от начальных условий, а второе – вынужденную составляющую, зависящую от входного сигнала.

Одним из способов вычисления матрицы $\Phi(t) = e^{At}$ является ее представление в виде ряда по степеням матрицы A :

$$e^{At} = a_0(t) \cdot E + a_1(t) \cdot A + a_2(t) \cdot A^2 + \dots$$

Число членов ряда при решении конкретной задачи определяется порядком системы n . Неизвестные коэффициенты $a_0(t), a_1(t), a_2(t), \dots$ определяются путем решения матричного уравнения

$$\begin{bmatrix} 1 & \lambda_1 & \lambda_1^2 & \dots & \lambda_1^{n-1} \\ 1 & \lambda_2 & \lambda_2^2 & \dots & \lambda_2^{n-1} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & \lambda_n & \lambda_n^2 & \dots & \lambda_n^{n-1} \end{bmatrix} \begin{bmatrix} a_0(t) \\ a_1(t) \\ \vdots \\ a_{n-1}(t) \end{bmatrix} = \begin{bmatrix} e^{\lambda_1 t} \\ e^{\lambda_2 t} \\ \vdots \\ e^{\lambda_n t} \end{bmatrix}, \quad (6.17)$$

где $\lambda_1, \lambda_2, \dots, \lambda_n$ – характеристические числа матрицы A , которые вычисляются путем решения уравнения $|\lambda E - A| = 0$.

Определитель матрицы коэффициентов в левой части уравнения (6.17) носит название определителя Вандермонда.

Пример 6.5. Вычислить $\Phi(t) = e^{At}$, если $A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}$. Размер матрицы A определяет количество членов ряда $e^{At} = a_0(t)E + a_1(t)A$.

Найдем характеристические числа матрицы A , как корни уравнения

$$|\lambda E - A| = \begin{vmatrix} \lambda & -1 \\ 2 & \lambda + 3 \end{vmatrix} = \lambda^2 + 3\lambda + 2 = 0, \quad \lambda_1 = -1; \lambda_2 = -2.$$

Далее решаем матричное уравнение

$$\begin{bmatrix} 1 & -1 \\ 1 & -2 \end{bmatrix} \begin{bmatrix} a_0(t) \\ a_1(t) \end{bmatrix} = \begin{bmatrix} e^{-t} \\ e^{-2t} \end{bmatrix}.$$

Перемножив матрицы в левой части и учитывая равенство векторов, получим систему уравнений

$$\begin{cases} a_0(t) - a_1(t) = e^{-t}, \\ a_0(t) - 2a_1(t) = e^{-2t}. \end{cases}$$

Решение системы имеет вид: $a_0(t) = 2e^{-t} - e^{-2t}$, $a_1(t) = e^{-t} - e^{-2t}$.

Следовательно

$$\begin{aligned}
 e^{At} &= (2e^{-t} - e^{-2t}) \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + (e^t - e^{-2t}) \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} = \\
 &= e^{-t} \left\{ \begin{bmatrix} 2 & 0 \\ 0 & 2 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \right\} - e^{-2t} \left\{ \begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \right\} = \\
 &= e^{-t} \begin{bmatrix} 2 & 1 \\ -2 & -1 \end{bmatrix} + e^{-2t} \begin{bmatrix} -1 & -1 \\ 2 & 2 \end{bmatrix}.
 \end{aligned}$$

6.13. Характеристическое уравнение. Модальная матрица. Преобразование подобия

Если разомкнутая система n -го порядка задана передаточной функцией вида

$$W(s) = \frac{N(s)}{D(s)},$$

то полином $D(s)$ называется характеристическим полиномом разомкнутой системы, а алгебраическое уравнение n -й степени $D(\lambda) = 0$, где λ некоторая переменная, называется характеристическим уравнением разомкнутой системы

$$\begin{aligned}
 D(s) &= s^n + \alpha_{n-1}s^{n-1} + \dots + \alpha_1s + \alpha_0 = 0, \\
 D(\lambda) &= \lambda^n + \alpha_{n-1}\lambda^{n-1} + \dots + \alpha_1\lambda + \alpha_0 = 0.
 \end{aligned} \tag{6.18}$$

Обозначим через $\lambda_i, i = \overline{1, n}$ корни характеристического уравнения (6.18). Если та же система задана уравнениями состояния

$$\begin{aligned}
 \dot{x} &= Ax + Bu, \\
 y &= Cx + Du,
 \end{aligned}$$

то характеристическое уравнение для такой модели можно записать как $|\lambda E - A| = 0$, где E – единичная матрица, λ – скалярная величина.

Раскладывая определитель, получим уравнение вида

$$\lambda^n + \alpha_{n-1}\lambda^{n-1} + \dots + \alpha_1\lambda + \alpha_0 = 0. \tag{6.19}$$

Значения λ_i , удовлетворяющие этому уравнению, называются характеристическими числами матрицы A , или ее собственными значениями. Уравнения (6.18) и (6.19) совпадают, так как речь идет об одной и той же системе.

Если матрица A имеет форму Фробениуса, которая была рассмотрена ранее, и все корни λ_i характеристического уравнения различны, то модальной матрицей называется квадратная матрица порядка n , следующего вида

$$M = \begin{bmatrix} 1 & 1 & 1 & \dots & 1 \\ \lambda_1 & \lambda_2 & \lambda_3 & \dots & \lambda_n \\ \lambda_1^2 & \lambda_2^2 & \lambda_3^2 & \dots & \lambda_n^2 \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ \lambda_1^{n-1} & \lambda_2^{n-1} & \lambda_3^{n-1} & \dots & \lambda_n^{n-1} \end{bmatrix},$$

где λ_i – собственные значения матрицы A (или корни характеристического уравнения).

Любая квадратная матрица A , все собственные значения которой различны, преобразуется в диагональную матрицу Λ , элементами которой являются эти собственные значения следующим образом

$$\Lambda = M^{-1} \cdot A \cdot M. \quad (6.20)$$

Преобразование вида (6.20), где Λ и A квадратные матрицы, а M – неособенная квадратная матрица (т. е. $\det M \neq 0$) называется **преобразованием подобия**.

6.14. Каноническая форма уравнений состояния

Будем считать, что нормальная форма известна

$$\begin{cases} \dot{x} = Ax + Bu, \\ y = Cx + Du. \end{cases} \quad (6.21)$$

Введем новую переменную q , которая связана с x следующим образом: $x = Mq$, $\dot{x} = M\dot{q}$ и подставим в (6.21), тогда

$$\begin{cases} M\dot{q} = AMq + Bu, \\ y = CMq + Du. \end{cases}$$

Умножим первое уравнение слева на M^{-1} , тогда $\dot{q} = M^{-1}AMq + M^{-1}Bu$. С учетом (6.21) получим

$$\begin{cases} \dot{q} = \Lambda q + B_1 u, \\ y = C_1 q + D_1 u. \end{cases} \quad (6.22)$$

Уравнения (6.22) известны как **каноническая форма уравнений состояния**.

Здесь $\Lambda = \begin{bmatrix} \lambda_1 & 0 & \dots & 0 \\ 0 & \lambda_1 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \lambda_1 \end{bmatrix}$ – диагональная матрица коэффициентов, $B_1 = M^{-1}B$, $C_1 = CM$, $D_1 = D$. Вектор C_1 содержит n единиц, что обусловлено видом матриц C и M .

Перейдем к скалярной форме записи в виде системы дифференциальных уравнений первого порядка

$$\begin{cases} \dot{q}_i = \lambda_i q_i + b_{1i} u, i = \overline{1, n}, \\ y = \sum_{i=1}^n q_i + d_1 u. \end{cases} \quad (6.23)$$

Схема моделирования, соответствующая системе (6.23) приведена на рис. 6.17. Для нее характерно параллельное соединение интеграторов.

Преимущество канонической формы в том, что каждое дифференциальное уравнение зависит только от одной переменной и решается просто. Выход системы y складывается из суммы решений отдельных уравнений.

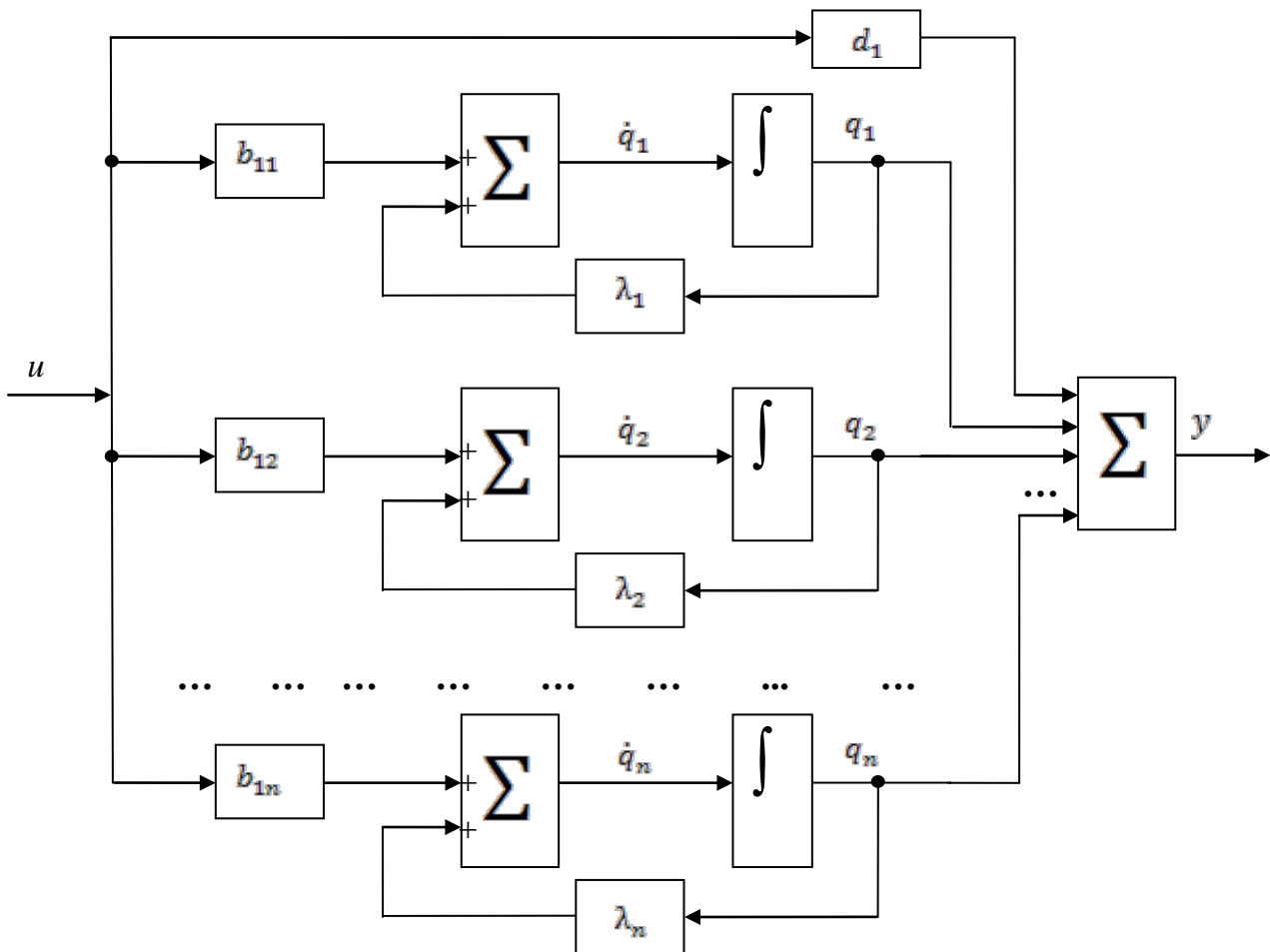


Рис. 6.17. Схема моделирования канонической формы уравнений состояния

Пример 6.6. Передаточная функция системы имеет вид

$$W(s) = \frac{2s^2 + 8s + 11}{s^2 + 3s + 2}.$$

Записать каноническую форму уравнений состояния, изобразить схему моделирования, определить реакцию системы на сигнал $u = u_0 \cdot 1(t)$. Вначале найдем матрицы A, B, C и D , соответствующие нормальной форме уравнений состояния

$$\begin{cases} \dot{x} = Ax + Bu, \\ y = Cx + Du. \end{cases}$$

Дифференциальное уравнение, описывающее исходную систему имеет вид

$$\frac{d^2 y}{dt^2} + 3 \frac{dy}{dt} + 2y = 2 \frac{d^2 u}{dt^2} + 8 \frac{du}{dt} + 11u,$$

тогда

$$A = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix}, C = [1 \ 0], D = b_0 = \beta_2 = 2.$$

Элементы матрицы $B = \begin{bmatrix} b_1 \\ b_2 \end{bmatrix}$ вычислим по следующим соотношениям

$$b_1 = \beta_1 - \alpha_1 b_0 = 8 - 3 \cdot 2 = 2,$$

$$b_2 = \beta_0 - \alpha_1 b_1 - \alpha_0 b_0 = 11 - 3 \cdot 2 - 2 \cdot 2 = 1.$$

Каноническая форма уравнений состояния имеет вид

$$\dot{q} = \Lambda q + B_1 u,$$

$$y = C_1 q + D_1 u.$$

Элементами диагональной матрицы Λ являются характеристические числа матрицы A , которые определяются решением уравнения

$$|\lambda E - A| = \begin{vmatrix} \lambda & -1 \\ 2 & \lambda + 3 \end{vmatrix} = \lambda^2 + 3\lambda + 2 = 0,$$

корни которого равны $\lambda_1 = -1$, $\lambda_2 = -2$.

Тогда $\Lambda = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix}$, модальная матрица $M = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix}$, а обратная ей $M^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix}$. Матрицы D_1 и D совпадают, т. е. $D_1 = D = 2$.

Матрицы B_1 и C_1 определяются следующим образом

$$B_1 = M^{-1} B = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 1 \end{bmatrix} = \begin{bmatrix} 5 \\ -3 \end{bmatrix},$$

$$C_1 = C M = [1 \ 0] \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} = [1 \ 1].$$

Запишем соответствующее уравнения

$$\begin{bmatrix} \dot{q}_1 \\ \dot{q}_2 \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 0 & -2 \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} + \begin{bmatrix} 5 \\ -3 \end{bmatrix} u,$$

$$y = [1 \ 1] \begin{bmatrix} q_1 \\ q_2 \end{bmatrix} + 2u.$$

Переходя к скалярной форме записи, получим

$$\dot{q}_1 = -q_1 + 5u,$$

$$\dot{q}_2 = -2q_2 - 3u,$$

$$y = q_1 + q_2 + 2u.$$

Соответствующая схема моделирования приведена на рис. 6.18.

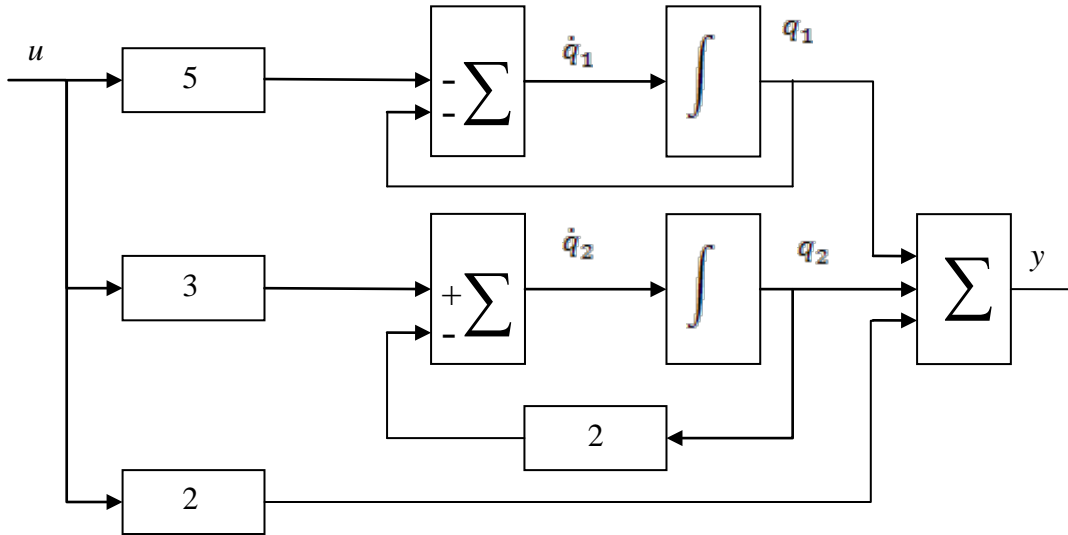


Рис. 6.18. Схема моделирования к примеру 6.6

Определим реакцию системы на входной сигнал $u(t) = u_0 \cdot 1(t)$, если $y(0) = 2$, $\dot{y}(0) = \ddot{y}(0) = 0$.

Для этого необходимо знать начальное состояние для вектора $q(t)$, т. е. $q(0)$. Исходя из ранее введенных обозначений $x_1(0) = y(0) = 2$, $x_2(0) = \dot{y}(0) = 0$. Так как $q = M^{-1}x$, то начальные условия для вектора $q(t)$ будут выглядеть следующим образом

$$\begin{bmatrix} q_1(0) \\ q_2(0) \end{bmatrix} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} \begin{bmatrix} 2 \\ 0 \end{bmatrix} = \begin{bmatrix} 4 \\ -2 \end{bmatrix},$$

следовательно $q_1(0) = 4$; $q_2(0) = -2$.

Решение уравнения $\dot{q} = aq + bu$ записывается следующим образом

$$q(t) = q(0)e^{at} + b \int_0^t e^{a(t-\tau)} u(\tau) d\tau,$$

тогда решение уравнения $\dot{q}_1 = -q_1 + 5u$ имеет вид

$$q_1(t) = 4e^{-t} + 5u_0 \int_0^t e^{-(t-\tau)} d\tau = 4e^{-t} + 5u_0(1 - e^{-t}),$$

а решение уравнения $\dot{q}_2 = -2q_2 - 3u$ имеет вид

$$q_2(t) = -2e^{-2t} - 3u_0 \int_0^t e^{-2(t-\tau)} d\tau = -2e^{-2t} - \frac{3}{2}u_0(1 - e^{-2t}).$$

Полагая $u_0 = 1$ определим выходной сигнал, как

$$y(t) = q_1 + q_2 + 2u = \frac{11}{2} - e^{-t} - \frac{1}{2}e^{-2t}.$$

6.15. Понятие об устойчивости линейных систем

Система является устойчивой, если после прекращения внешнего воздействия, она через некоторое время возвращается к тому состоянию равновесия или вынужденного движения, в котором находилась до начала воздействия.

Устойчивость линейной системы определяется ее структурой и параметрами, и не зависит от поступающих на систему сигналов.

Пусть система описывается дифференциальным уравнением

$$\sum_{i=0}^n a_i \frac{d^i y(t)}{dt^i} = \sum_{j=0}^m b_j \frac{d^j u(t)}{dt^j}.$$

Решение его складывается из двух составляющих $y(t) = y_1(t) + y_2(t) = y_{\text{св}}(t) + y_{\text{вын}}(t)$, где $y_1(t)$ – общее решение однородного дифференциального уравнения без правой части, которое называют свободной или переходной составляющей, характеризует собственное движение системы под влиянием начальных условий $y_2(t)$ – частное решение неоднородного дифференциального уравнения с правой частью, это вынужденная составляющая.

Система называется асимптотически устойчивой, если с течением времени (при $t \rightarrow \infty$) свободная составляющая будет стремиться к нулю, т. е. если

$$\lim_{t \rightarrow \infty} y_{\text{св}}(t) = 0$$

$y_{\text{св}}(t)$ является решением однородного дифференциального уравнения

$$\sum_{i=0}^n a_i \frac{d^i y(t)}{dt^i} = 0,$$

переходя к преобразованию Лапласа получим выражение

$$D(s) = a_n s^n + a_{n-1} s^{n-1} + \dots + a_1 s + a_0 = 0,$$

которое называют характеристическим уравнением системы.

Будем считать, что кратных корней нет, а корни характеристического уравнения либо вещественные λ_i , либо это комплексно-сопряженные пары $\lambda_k \pm j\gamma_k$. Тогда решение уравнения

$$y_{\text{св}}(t) = \sum_i A_i e^{\lambda_i t} + \sum_k C_k e^{\lambda_k t} \cdot \sin(\gamma_k t + \psi_k),$$

где A_i, C_k, ψ_k постоянные интегрирования, которые зависят от начальных условий и всегда ограничены по абсолютной величине.

Первая группа слагаемых соответствует вещественным корням λ_i , а вторая группа соответствует парам комплексно-сопряженных корней $\lambda_k \pm j\gamma_k$

Суммарная величина $\sum_{i=1}^n A_i e^{\lambda_i t}$ может стремиться к нулю только тогда, когда корни будут отрицательными числами. Графическое изображение отдельных составляющих представлено на рис. 6.19, а, где видно, что при $\lambda_i > 0$ процесс расходится.

Вторая группа слагаемых соответствует колебательным составляющим, которые также могут быть сходящимися (при $\lambda < 0$) или расходящимися (при $\lambda > 0$), как показано на рис. 6.19, б.

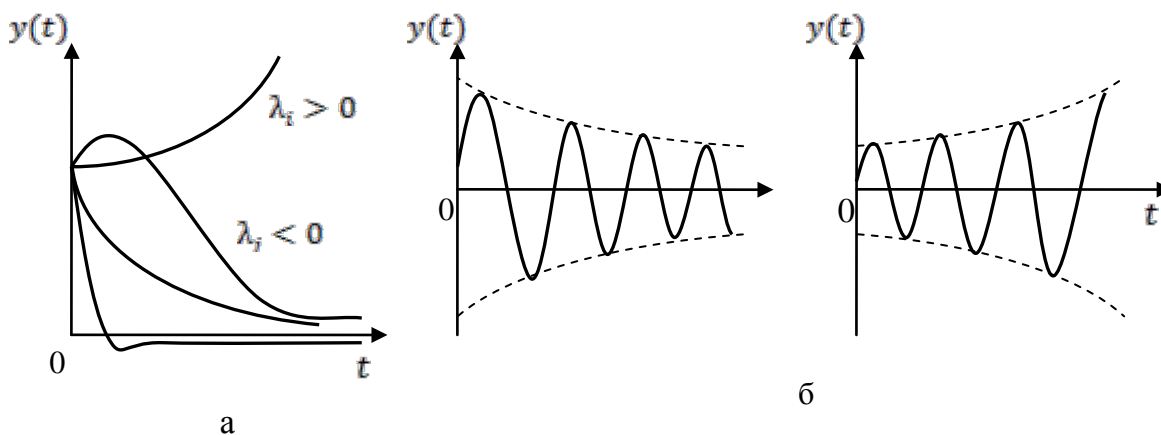


Рис. 6.19

Таким образом, для устойчивости линейной системы, необходимо и достаточно, чтобы все корни ее характеристического уравнения имели отрицательную вещественную часть. При наличии хотя бы одного корня с положительной вещественной частью система неустойчива. Система будет находиться на границе устойчивости при наличии нулевого корня или пары чисто мнимых корней, а остальные с отрицательной вещественной частью. Если присутствует пара мнимых корней, то будут незатухающие гармонические колебания с постоянной амплитудой.

Корни характеристического уравнения, как и всякие комплексные числа, удобно представлять в виде точек на комплексной плоскости (рис. 6.20).

Если действительные части корней отрицательны, то соответствующие им точки на комплексной плоскости P лежат слева от мнимой оси. Мнимая ось является границей устойчивости. На практике для упрощения вычислений устойчивость систем определяют с помощью некоторых критериев без вычисления корней характеристического уравнения.

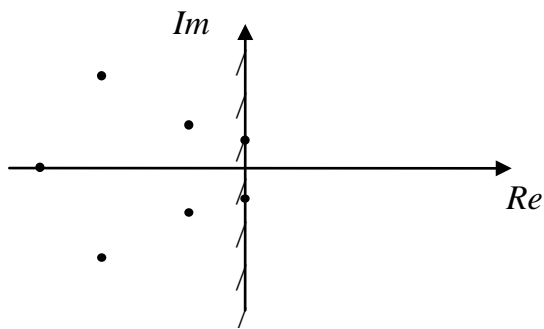


Рис. 6.20

Критерий устойчивости – это математическая формулировка условий, которым удовлетворяют коэффициенты характеристического уравнения устойчивой системы.

Системы первого и второго порядка устойчивы, если все коэффициенты характеристического уравнения положительны. Для систем более высокого порядка положительность коэффициентов является необходимым, но недостаточным условием устойчивости.

Критерий Гурвица. Пусть характеристическое уравнение системы имеет вид

$$a_n S^n + a_{n-1} S^{n-1} + \dots + a_1 S + a_0 = 0.$$

Составим квадратную $n \times n$ матрицу из коэффициентов этого уравнения. Приведенная матрица называется матрицей Гурвица.

$$\begin{bmatrix} a_{n-1} & a_{n-3} & a_{n-5} & \dots & 0 \\ a_n & a_{n-2} & a_{n-4} & \dots & 0 \\ 0 & a_{n-1} & a_{n-3} & \dots & 0 \\ \dots & \dots & \dots & \dots & \dots \\ 0 & 0 & 0 & \dots & a_0 \end{bmatrix}.$$

По главной диагонали располагаются все коэффициенты от a_{n-1} до a_0 , вверх от диагонали индексы коэффициентов уменьшаются, а вниз – увеличиваются. Система устойчива, если при $a_n > 0$ положительны n диагональных миноров Гурвица, получаемых из матрицы: $\Delta_1 = a_{n-1} > 0$; $\Delta_2 = \begin{vmatrix} a_{n-1} & a_{n-3} \\ a_n & a_{n-2} \end{vmatrix} > 0$, $\Delta_3 > 0 \dots$, $\Delta_n = a_0 \cdot \Delta_{n-1} > 0$.

Это необходимое и достаточное условие устойчивости.

Если $\Delta_n = 0$, а все предыдущие определители Гурвица положительны, то система находится на границе устойчивости. $\Delta_n = 0$ в том случае, если $a_0 = 0$ или $\Delta_{n-1} = 0$.

Первое условие соответствует наличию нулевого корня у характеристического уравнения (апериодическая граница устойчивости, нейтральная устойчивость), а второе — пары чисто мнимых корней (колебательная граница устойчивости).

Пример 6.7. Система состоит из трех последовательно соединенных апериодических звеньев. Определить условия, при которых система, замкнутая единичной обратной связью будет устойчивой. Структура системы приведена на рис. 6.21.

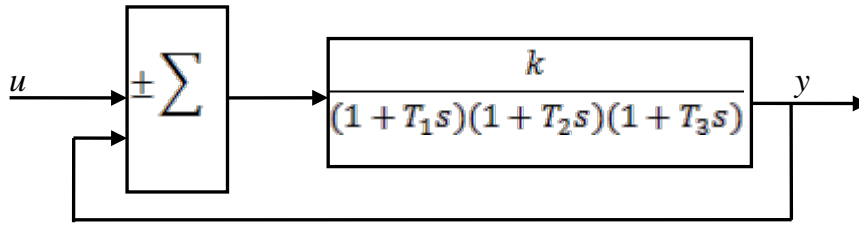


Рис. 6.21

Передаточная функция замкнутой системы имеет вид

$$W_{\text{замкн.}} = \frac{W_p(s)}{1 + W_p(s)}$$

Выполняя преобразования, найдем характеристическое уравнение замкнутой системы $T_1 T_2 T_3 s^3 + (T_1 T_2 + T_1 T_3 + T_2 T_3) s^2 + (T_1 + T_2 + T_3) s + 1 + k = 0$.

Составим матрицу Гурвица и запишем условие устойчивости.

$$\begin{bmatrix} a_2 & a_0 & 0 \\ a_3 & a_1 & 0 \\ 0 & a_2 & a_0 \end{bmatrix}.$$

Очевидно, что $a_3 > 0$ и $\Delta_2 = a_2 > 0$. Необходимо проверить выполнение условия $\Delta_2 = a_2 a_1 - a_3 a_0 > 0$. Подставим в него значения коэффициентов, тогда $(T_1 T_2 + T_1 T_3 + T_2 T_3)(T_1 + T_2 + T_3) - T_1 T_2 T_3 (1 + k) \geq 0$.

Разделим обе части неравенства на произведение $T_1 T_2 T_3$

$$\left(\frac{1}{T_3} + \frac{1}{T_2} + \frac{1}{T_1} \right) (T_1 + T_2 + T_3) - (k + 1) > 0.$$

Каждое слагаемое в первой скобке умножим на T_1 , а во второй скобке разделим на T_1 . Выполняя преобразования, получим соотношение между параметрами системы, которое определяет устойчивость

$$k < \left(\frac{T_1}{T_3} + \frac{T_1}{T_2} + 1 \right) \left(1 + \frac{T_2}{T_1} + \frac{T_3}{T_1} \right) - 1.$$

Коэффициент усиления, при котором система теряет устойчивость называется граничным или предельным. Он определяется не абсолютной величиной постоянных времени T_i , а их отношением.

Если $T_1 = T_2 = T_3$, то $k < 8$; $k_{\text{гр}} = 8$.

Если $T_1 = 10T_3$; $T_2 = 100T_3$, то $k_{\text{гр}} \approx 121$.

Для обеспечения устойчивости системы при возможно большем коэффициенте усиления необходимо разносить постоянные времени входящих в нее звеньев.

6.16. Математическое описание дискретных систем и их элементов

Дискретные системы автоматического управления включают в себя цифровые и импульсные системы.

В цифровых системах осуществляется квантование сигналов по уровню и по времени. В импульсных системах осуществляется только квантование по времени, для таких систем характерна амплитудно-импульсная модуляция сигналов.

Понятие о решетчатых функциях. Процессы, происходящие в импульсных системах, описываются функциями дискретного аргумента, так называемыми решетчатыми функциями. Решетчатые функции образуются из соответствующих непрерывных функций при дискретизации их в равноотстоящие моменты времени. Обозначают их как $f(KT)$, где T – расстояние между соседними дискретными значениями аргумента, а K – целое число ($K = 0, 1, 2 \dots$), рис. 6.22.

Если положить $T = 1$, то можно перейти к нормированному времени $\bar{t} = \frac{t}{T}$ и решетчатую функцию обозначить сокращенно $f(K)$. Если интервал дискретизации задан, то по функции $f(t)$ решетчатая функция определяется однозначным образом. Обратное положение несправедливо: по решетчатой функции $f(K)$ нельзя восстановить функцию $f(t)$ без дополнительных сведений о поведении функции $f(t)$ в интервалах между точками $t = KT$. Вводят понятие смещенной решетчатой функции $f(KT, \sigma) = f(KT + \sigma T)$ или $f(K, \sigma) = f(K + \sigma)$, где функции $0 \leq \sigma \leq 1$. Здесь числовая последовательность образуется в результате выборки значений функции $f(t)$ в точках, смещенных относительно значений KT на некоторый параметр σ .

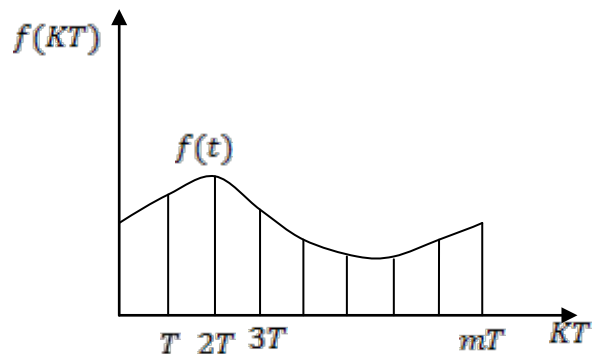


Рис. 6.22

Задавая различные значения σ , можно проследить изменение $f(t)$ на интервале дискретизации. Очевидно, что решетчатая функция $f(KT)$ является частным случаем смещенной решетчатой функции при $\sigma = 0$.

Скорость изменения дискретной нормированной функции характеризуется ее первой разностью $\Delta f(K)$, которая является аналогом производной для непрерывных функций и определяется выражением

$$\Delta f(K) = f(K + 1) - f(K).$$

Оператор Δ для дискретных функций является аналогом дифференциального оператора $s = \frac{d}{dt}$ для непрерывных функций.

Разность второго порядка $\Delta^2 f(K)$ определяется как

$$\begin{aligned}\Delta^2 f(K) &= \Delta f(K+1) - \Delta f(K) = f(K+2) - f(K+1) - f(K+1) + f(K) = \\ &= f(K+2) - 2f(K+1) + f(K).\end{aligned}$$

Разность третьего порядка равна

$$\Delta^3 f(K) = \Delta^2 f(K+1) - \Delta^2 f(K) = f(K+3) - 3f(K+2) + 3f(K+1) - f(K).$$

Соответственно разность произвольного порядка n можно представить следующим образом

$$\Delta^n f(K) = \sum_{v=0}^n (-1)^v \frac{n!}{v!(n-v)!} \cdot f(K+n-v).$$

Системы, в которых взаимосвязь между входом и выходом определяется только в равностоящие моменты времени $t = KT$ описываются разностными уравнениями.

Уравнение в конечных разностях имеет вид

$$\begin{aligned}a_n \Delta^n y(K) + a_{n-1} \Delta^{n-1} y(K) + \dots + a_1 \Delta y(K) + a_0 y(K) = \\ = b_m \Delta^m u(K) + b_{m-1} \Delta^{m-1} u(K) + \dots + b_1 \Delta u(K) + b_0 u(K),\end{aligned}\quad (6.24)$$

где n – порядок уравнения, причем $n \geq m$.

Если в (6.24) выразить разности через значения решетчатой функции, то разностное уравнение будет иметь вид

$$\begin{aligned}c_n y(K+n) + c_{n-1} y(K+n-1) + \dots + c_1 y(K+1) + c_0 y(K) = \\ = d_m u(K+m) + d_{m-1} u(K+m-1) + \dots + d_1 u(K+1) + d_0 u(K).\end{aligned}\quad (6.25)$$

Выражение (6.24) является более близким аналогом дифференциального уравнения, однако выражение (6.25) легче использовать и такая форма разностного уравнения более распространена.

Z-преобразование и его свойства. Подобно тому, как применение преобразования Лапласа к дифференциальным уравнениям дает возможность перейти к алгебраическим уравнениям и получить удобную инженерную методику анализа, так и для дискретных систем был разработан ряд специальных преобразований. Наибольшее распространение получили дискретное преобразование Лапласа и Z-преобразование, которое является дискретным аналогом преобразования Лапласа.

Пусть значения функции $f(t)$ рассматриваются в дискретные моменты времени KT , где $K = 0, 1, 2, \dots$.

Z-преобразованием функции $f(KT)$ называется функция комплексного аргумента z , определяемая следующим выражением

$$F(z) = Z \{f(KT)\} = \sum_{K=0}^{\infty} f(KT) z^{-K}.$$

Если $f(K) = 1(K)$ – единичная ступенчатая дискретная функция, то $F(z) = 1 + z^{-1} + z^{-2} + \dots = \frac{z}{z-1}$ для $|z| > 1$.

Если $f(K) = e^{\alpha K}$ (дискретная экспонента $f(t) = e^{\alpha t}$), то $F(z) = 1 + z^{-1}e^{\alpha} + z^{-2}e^{2\alpha} + \dots = \frac{z}{z-e^{\alpha}}$ для $|z| > e^{\alpha}$.

Если $z = e^{sT}$, где s – оператор Лапласа, то Z -преобразование после перехода к переменной s представляет собой обычное преобразование Лапласа, примененное к последовательности δ -функций, площадь каждой из которых определяется соответствующим значением $f(t)$. Это и есть дискретное преобразование Лапласа

$$L\{f(KT)\} = \sum_{K=0}^{\infty} f(KT)e^{-sKT};$$

Z -преобразование для смещенной решетчатой функции $f(KT + \sigma T)$ называется модифицированным

$$F(z, \sigma) = \sum_{K=0}^{\infty} f(KT + \sigma T) \cdot z^{-K}.$$

Рассмотрим основные свойства Z -преобразования

1. Теорема сдвига. Сдвиг функции-оригинала вправо по оси времени на m интервалов дискретности соответствует умножению изображения на Z^{-m}

$$Z\{f[KT - mT]\} = Z^{-m} \cdot F(z, \sigma).$$

Если сдвиг оригинала по оси времени происходит влево на целое число m интервалов дискретности

$$Z\{f[KT + mT]\} = z^m \cdot F(z) - \sum_{n=0}^{m-1} z^{m-n} \cdot f[KT].$$

Если $f(n, \sigma) = 0$ для $n = 0, 1, \dots, m-1$, то второй член в правой части уравнения обращается в нуль

$$Z\{f[KT + mT]\} = z^m \cdot F(z, \sigma).$$

2. Линейность. Z -преобразование алгебраической суммы функций равно алгебраической сумме их Z -преобразований и постоянный множитель можно выносить за знак Z -преобразования.

$$Z\{\alpha_1 f_1(KT) + \alpha_2 f_2(KT)\} = \alpha_1 F_1(z) + \alpha_2 F_2(z).$$

3. Теорема о предельных значениях

$$\lim_{K \rightarrow \infty} f(KT) = \lim_{z \rightarrow 1} (1 - z^{-1})F(z)$$

$$f(0, \sigma) = \lim_{z \rightarrow \infty} (1 - z^{-1})F(z, \sigma).$$

4. Изображение конечных разностей. Если $\Delta f(KT) = f[(K+1)T] - f(KT)$, то $Z\{\Delta f(KT)\} = (z-1)F(z) - zf(0)$.

5. Теорема свертки

$$F_1(z)F_2(z) = Z \left\{ \sum_{K=0}^n f_1(KT)f_2[(n-K)T] \right\}.$$

Передаточная функция дискретной системы $W(z)$. Передаточная функция $W(z)$ разомкнутой дискретной системы с одним входом и одним выходом равна отношению Z -преобразования выходного сигнала к Z -преобразованию входного сигнала при нулевых начальных условиях.

$$W(z) = \frac{Y(z)}{U(z)}.$$

Передаточная функция $W(z)$ может быть записана непосредственно по разностному уравнению с учетом теоремы сдвига. Например, если разностное уравнение имеет вид

$$y(K+3) + 6y(K+2) + 11y(K+1) + 6y(K) = u(K+1) + u(K),$$

то, переходя к Z -преобразованию, получим

$$(z^3 + 6z^2 + 11z + 6) \cdot Y(z) = (z + 1) \cdot U(z).$$

Тогда

$$W(z) = \frac{Y(z)}{U(z)} = \frac{z + 1}{z^3 + 6z^2 + 11z + 6}.$$

Передаточная функция $W(z)$ является дробно-рациональной функцией Z -преобразование выходного сигнала $Y(z) = U(z) \cdot W(z)$.

Знаменатель передаточной функции, приравненный к нулю, называется характеристическим уравнением разомкнутой импульсной системы

$$a_n z^n + a_{n-1} z^{n-1} + \dots + a_1 z + a_0 = 0.$$

Для устойчивости дискретных систем необходимо и достаточно, чтобы корни характеристического уравнения замкнутой системы были расположены внутри единичного радиуса с центром в начале координат, $|z| < 1$. Так как $z = e^{sT}$, то левая часть s -плоскости для непрерывных систем, отображается на плоскости внутрь единичного круга (в полярных координатах $z = \rho e^{j\theta}$).

Обратное Z -преобразование. Для нахождения дискретной функции (оригинала) по ее заданному Z -преобразованию используется обратное Z -преобразование. Обратное Z -преобразование может быть получено двумя способами:

- разложение заданного преобразования в ряд по степеням z^{-1} ;
- с помощью интегральных вычетов.

Из определения Z -преобразования ясно, что коэффициент при z^{-K} разложения будет равен значению искомой функции

$$F(z) = f(0) + f(T)z^{-1} + \dots + f(KT)z^{-K} + \dots \quad (6.26)$$

Умножим $F(z)$ на z^{K-1} и проинтегрируем в плоскости z по любому замкнутому контуру, на котором и вне которого $F(z)$ не имеет полюсов и является аналитической функцией. Тогда

$$f[KT] = \frac{1}{2\pi j} \oint F(z) z^{K-1} dz = \sum \text{Res} |F(z) z^{K-1}|_{z=z_n},$$

где z_n – точки, в которых функция $F(z)$ имеет полюсы.

Значение интеграла равно сумме вычетов подинтегрального выражения $F(z)z^{K-1}$ в его особых точках. В технических приложениях $F(z, \sigma)$ является дробно-рациональной функцией z . Особыми точками дробно-рациональных функций являются их полюсы, т. е. те точки, в которых дробь обращается в бесконечность.

Для перехода к оригиналу часто прибегают к разложению $W(z)$ на простые дроби и пользуются затем таблицами соответствия для нахождения оригиналов для каждой дроби.

6.17. Уравнения состояния и моделирование дискретных систем

Разностное уравнение n -го порядка

$$\begin{aligned} c_n y(K+n) + c_{n-1} y(K+n-1) + \dots + c_1 y(K+1) + c_0 y(K) = \\ = d_m u(K+m) + \dots + d_1 u(K+1) + d_0 u(K). \end{aligned}$$

может быть представлено в виде системы n разностных уравнений 1-го порядка в матричном виде

$$x(K+1) = A[K]x(K) + B(K)u(K)$$

и матричным уравнением типа вход–состояние–выход

$$y(K) = C(K)x(K) + D(K)u(K).$$

Правила нахождения матриц A , B , C , D те же, что и в случае непрерывных систем.

Рассмотрим однородное нестационарное дискретное разностное уравнение

$$x(K+1) = A[K]x(K).$$

Если начальные условия заданы, то имеем

$$x(1) = A(0)x(0),$$

$$x(2) = A(1)x(1) = A(1)A(0)x(0).$$

и так далее, так что

$$x(K) = \prod_{i=0}^{K-1} A(i)x(0).$$

Определим дискретную переходную матрицу состояния с помощью соотношений:

1) $x(K) = \Phi(K, m)x(m)$. Она описывает процесс перехода системы из m -го состояния в состояние K ;

2) $\Phi(K, K) = E$.

Тогда по определению $x(K) = \Phi(K, 0)x(0)$.

Сравнивая (1) и (2), найдем, что

$$\Phi(K, 0) = \prod_{i=0}^{K-1} A(i).$$

Если $A(i) = A$, т. е. постоянная матрица, то $\Phi(K, 0) = \Phi(K) = A^K$ и для ее вычисления можно использовать теорему Кэли-Гамильтона.

Дискретная матрица перехода удовлетворяет следующим свойствам

$$\Phi(K_3, K_2) \cdot \Phi(K_2, K_1) = \Phi(K_3, K_1), \quad \Phi(K_1, K_2) = \Phi^{-1}(K_2, K_1).$$

Зная матрицу перехода, можно записать полное решение дискретного уравнения состояния

$$x(K) = \Phi(K)x(0) + \sum_{i=0}^{K-1} \Phi(K-i-1)B \cdot u(i).$$

Рассмотрим принципы моделирования линейных разностных уравнений.

Основными звеньями, необходимыми для построения схемы моделирования служат сумматор, усилитель и звено задержки. Сумматор и усилитель – те же звенья, что и в непрерывных системах. Блок задержки для разностных уравнений в некотором смысле аналогичен интегрирующему звену для дифференциальных уравнений, в нем осуществляется временная задержка на интервал дискретизации T .

Пример 6.8. Записать уравнение состояния и построить схему моделирования системы, описываемой разностным уравнением

$$y(K+2) + ay(K+1) + by(K) = u(K).$$

Выразим $y(K+2) = u(K) - ay(K+1) - by(K)$ и изобразим схему моделирования, представленную на рис. 6.23.

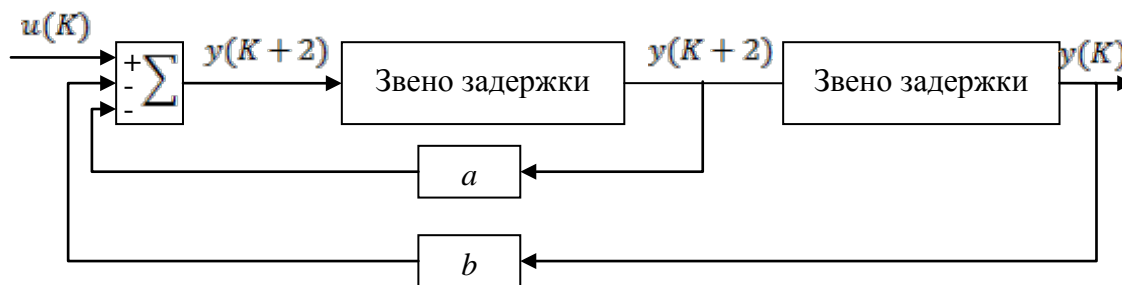


Рис. 6.23. Схема моделирования к примеру 6.8

Количество последовательно включенных звеньев задержки определяется порядком системы, в данном случае их две. Входные и выходные сигналы этих звеньев подписаны на рисунке, схема замкнута, исходя из условия **удовлетворения разностному уравнению**.

В качестве переменных состояния $x_1(K)$ и $x_2(K)$ обозначим выходы звеньев задержки, т. е. $x_1(K) = y(K)$; $x_2(K) = y(K + 1)$. Тогда $x_1(K + 1) = x_2(K)$, $x_2(K + 1) = -bx_1(K) - ax_2(K) + u(K)$, $y(K) = x_1(K)$ или

$$\begin{bmatrix} x_1(K + 1) \\ x_2(K + 1) \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -b & -a \end{bmatrix} \begin{bmatrix} x_1(K) \\ x_2(K) \end{bmatrix} + \begin{bmatrix} 0 \\ 1 \end{bmatrix} u(K),$$

$$y(K) = [1 \quad 0] \begin{bmatrix} x_1(K) \\ x_2(K) \end{bmatrix} + 0 \cdot u(K),$$

$$\text{где } A = \begin{bmatrix} 0 & 1 \\ -b & -a \end{bmatrix}, B = \begin{bmatrix} 0 \\ 1 \end{bmatrix}, C = [1 \quad 0], D = 0.$$

ЛИТЕРАТУРА

1. Горбатов, А. В. Дискретная математика : Учебник для студентов вузов / А. В. Горбатов, В. А. Горбатова, М. В. Горбатова. – Астраль, АСТ, 2006. – 448 с.
2. Палий, И. А. Дискретная математика. Курс лекций. – Эксмо, 2008. – 352 с.
3. Тишин, В. В. Дискретная математика в примерах и задачах. БХВ-Петербург, 2008. – 352 с : ил.
4. Гоноровский, И. С. Радиотехнические цепи и сигналы: Учебное пособие для вузов. 5-е изд. / И. С. Гоноровский. – М. : Дрофа, 2006. – 719 с.
5. Макоха, А. Н. Дискретная математика : учебник / А. Н. Макоха, П. А. Сахнюк, Н. И. Червяков. – Физматлит, 2005. – 368 с.
6. Певзнер, Л. Д. Математические основы теории систем / Л. Д. Певзнер, Е. П. Чураков. – М. : Высшая школа, 2009. – 504 с.
7. Баскаков, С. И. Радиотехнические цепи и сигналы. – М. : Высшая школа, 2000. – 448 с.
8. Павлова, А. В. Математические основы теории систем: Конспект лекций для студентов специальности «Информационные технологии и управление в технических системах». Ч. 1. [Электронный ресурс]. – Минск : БГУИР, 2010. – 171 с. Режим доступа: http://www.bsuir.by/m/12_100229_1_62540.pdf.