Парадигма развития науки

Методологическое обеспечение

А. Е. Кононюк

СИСТЕМОЛОГИЯ

ОБЩАЯ ТЕОРИЯ СИСТЕМ

Книга 2

Общая теория структур

Часть 2

Организация структур

Киев
Освіта України
2014

УДК 51 (075.8)
ББК В161.я7
К213

Рецензент:
Н. К. Печурин — д-р техн. наук, проф. (Национальный авиационный университет).

Кононюк А. Е.
К213 Системология. Общая теория систем. — В 4-х кн. Кн. 2.Ч.2— К.: Освіта України. 2014. — 708 с.

ISBN 978-966-373-693-8 (многотомное издание)
ISBN 978-966-373-694-5 (книга 2)

В предлагаемом многотомном научно-учебном издании предпринята попытка раскрыть с единой идеологической и методологической точки зрения чрезвычайно сложную научную проблему - общую теорию систем. В работе достаточно полно, на взгляд автора, раскрыты методы системного подхода и системного анализа, используемые при анализе и синтезе различного класса систем. Определяются основные (базовые) понятия теории систем, раскрывается содержание системного анализа, его технология.

Для бакалавров, специалистов, магистров, аспирантов, докторантов всех специальностей.

УДК 51 (075.8)
ББК В161.я7

ISBN 978-966-373-693-8 (многотомное издание)
ISBN 978-966-373-694-5 (книга 2,часть 2)

Оглавление

1. Анализа структур систем

1.1. Задачи анализа структур систем

Иследуемые системы характеризуются структурами, определяющими множество их качественных свойств, обусловливаемых множеством элементов и связей между ними. При анализе структур систем и их подсистем в целях выявления их общих свойств конкретные (физические) реализации отдельных элементов и связей не имеют существенного значения и обычно не рассматриваются. Подобное абстрагирование позволяет производить анализ структур в наиболее общем виде — как схему «чистых» отношений ее элементов и связей. Основным свойством структур является однозначное соответствие их некоторому подмножеству K_s качественных свойств структур, причем каждой структуре соответствует определенный вид этого подмножества.

Однако обратного однозначного соответствия подмножества K_s и структуры не существует, т. е. одно и то же подмножество качественных свойств может принадлежать нескольким (в общем случае — множеству) различным структурам, причем, чем меньше мощность подмножества K_s, тем больше мощность множества соответствующих структур s.

Указанное обстоятельство имеет существенное значение в процессе создания структуры, так как им обусловливается необходимость выбора (формирования) рекомендаций оценки качества структур или соответствующих критериев, определяющих метод выбора предпочтительных или оптимальных структур из множества s.

Одним из ***свойств структур*** является также ***их устойчивость к влиянию внешних воздействий***, что может выражаться в нарушении тех или иных связей, вследствие чего уменьшается число их качественных свойств, определяемое мощностью подмножества K_s или наоборот, исключающие или препятствующие использованию конкретной структуры по ее прямому назначению.

Одной из ***главных задач <u>структурного анализа</u> является <u>построение</u> наглядной <u>формальной модели</u>, отображающей существующую систему отношений элементов как между собой, так и с внешней средой. Структурная модель системы*** является ***<u>многоуровневой</u>***, причем конкретизация структуры дается на стольких уровнях, сколько их требуется для создания полного представления об основных свойствах системы (рис. 1).

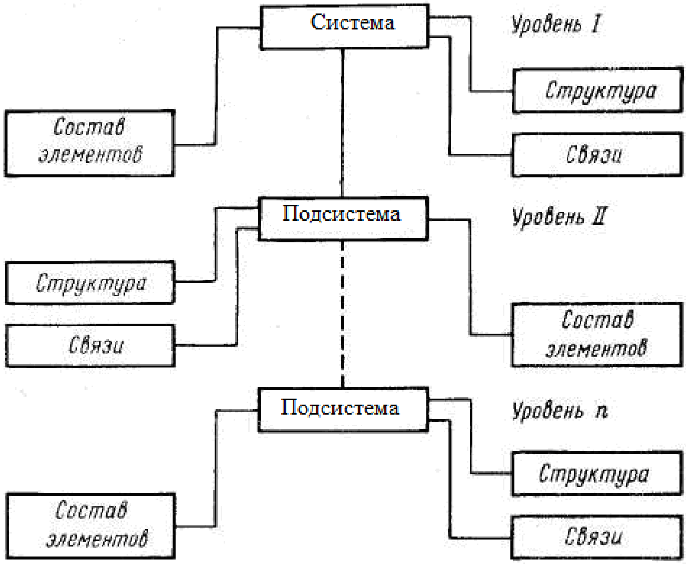


Рис. 1. Структурная модель системы

Рассмотрим структурную модель системы на нескольких крупных уровнях: организационном, функциональном, техническом. При анализе структур системы решаются следующие задачи:

организационная:

1) описание состава системы и построение ее структурной схемы;

2) определение функции отдельных подсистем системы, раскрытие их структурной схемы;

3) описание материальных, вещественных, информационных и других видов связи;

4) построение обобщенной структурной модели системы;

функциональная:

1) изучаются вопросы функционирования каждой подсистемы системы;

2) выбирается состав функций;

3) определяются их взаимосвязи;

4) составляется обобщенная структура задач функционирования системы;

техническая:

1) определяются основные элементы, участвующие в реализации функциональных задач системы;

2) составляется формальная структурная модель системы технических средств с учетом топологии расположения элементов системы и взаимодействия их как между собой, так и с внешней средой.

Независимо от уровня рассмотрения ***общая задача структурного анализа системы состоит в том, чтобы, исходя из заданного описания элементов системы и непосредственных связей между ними, получить заключение о структурных свойствах системы в целом и ее подсистем***. При решении практических задач структурного анализа системы будем принимать ***три уровня описания связей между элементами***:

1) наличие связи;

2) направление связи;

3) вид и направление воздействий, определяющих взаимодействие элементов.

На первом уровне, когда исходят из наличия или отсутствия связей между элементами, структуре системы может соответствовать неориентированный граф, вершинами которого являются элементы системы, а ребрами — связи между элементами. Основные задачи структурного анализа на этом уровне сводятся:

1) к определению связности (целостности) системы; если система не является связной, то ставят задачу выделения изолированных несвязных подсистем со списками входящих в них элементов;

2) к выделению циклов;

3) к определению минимальных и максимальных последовательностей элементов (цепей), разделяющих элементы друг от друга.

На втором уровне, когда задано направление связи, системе соответствует ориентированный граф, направления дуг которого совпадают с направлениями связей. На этом уровне результаты структурного анализа оказываются более содержательными. К задачам структурного анализа в этом случае относят:

1) определение связности системы;

2) топологическую декомпозицию с выделением сильно связных подсистем;

3) выделение уровней в структуре и определение их взаимосвязи;

4) определение максимальных и минимальных длин путей;

5) определение характеристик топологической значимости элементов;

6) получение информации о слабых местах структуры и др.

На третьем уровне описания связей между элементами системы учитывается не только направленность связи, но и раскрываются состав и характер функций взаимодействия элементов (входные, входные и управляющие воздействия).

Далее будут рассмотрены задачи структурного анализа при минимальном объеме априорной информации о структуре системы, когда учитывается только факт наличия связи и ее направленность.

1.2. Введение в анализ функциональных структур

Разработка любых сложных систем обычно начинается с анализа требований к системе и идентификации основных функций и отношений между функциями. Этот процесс может быть представлен диаграммой, изображенной на рис. 2.

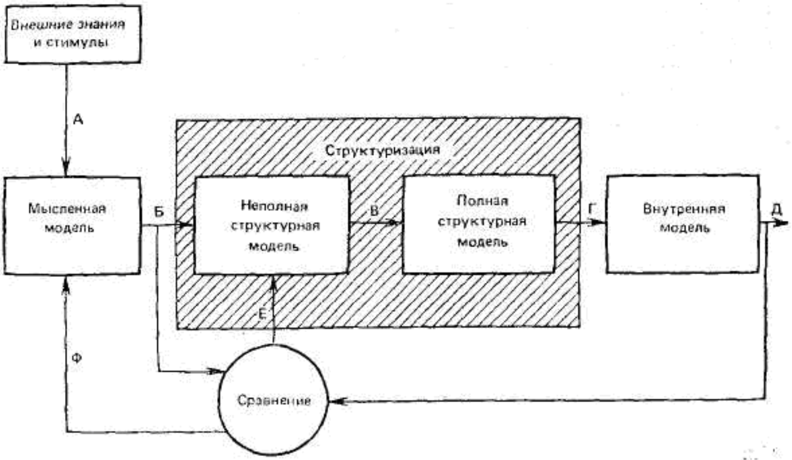


Рис. 2. Диаграмма процесса структурного моделирования системы

Под воздействием внешних стимулов исследователь, используя знания, полученные в результате анализа (А), создает мысленную модель системы. Используя ее, исследователь генерирует неполную структурную модель системы (Б), на базе которой строит полную структурную модель системы путем разбиения и группирования (В). Для получения полной структурной модели используются методы адекватного структурирования частей в целостную систему. Полученная полная модель системы, представленная в виде внутренней модели (Д), сравнивается с мысленной моделью системы. При этом обычно требуется уточнение мысленной модели и внесение изменений и исправлений в ее структуру. Кроме того, в ходе этого этапа обычно достигается более глубокое понимание структуры системы, интерпретируется и корректируется полная структурная модель системы. Надо также отметить большую важность процесса визуализации и документирования структурной модели системы.

Описанный процесс будем называть структурным моделированием (анализом) системы и он позволяет:

— идентифицировать ключевые элементы системы и их отношения;

— показать существование или не существование требуемых элементов и связей;

— определить уровни в системной иерархии;

— специфицировать подсистемы внутри уровней;

— определить критические и сильносвязанные элементы.

Процесс структурного моделирования системы включает в себя эмпирические и интуитивные шаги, поэтому его строгое определение затруднительно.

Пусть $\mathbf{A}$ и $\mathbf{T}$ — множество и мощность множества $\mathbf{A}$, $\mathbf{t}$ — подмножество $\mathbf{T}$, тогда $\mathbf{t} \times \mathbf{t}$ — декартово произведение этого множества на себя. Определим отношение в $\mathbf{t} \times \mathbf{t}$ как подмножество пар в $\mathbf{t} \times \mathbf{t}$, а свойство в отношении — как присвоение информации (направлений, знаков, весов и других описателей) каждой паре в отношении. Если рассматривать множество отношений $\mathbf{R}=\{r_1, ..., r_j, ..., r_n\}$, $r_j \in \mathbf{t} \times \mathbf{t}$ и множество соответствующих свойств $\mathbf{P}=\{p_1, ..., p_j, ... p_n)$, то $\mathbf{S} = (\mathbf{A}, \mathbf{R})$ и $\mathbf{U} = (\mathbf{A}, \mathbf{R}, \mathbf{P})$ обозначают соответственно структуру и нагруженную структуру.

Рассмотрим две нагруженные структуры $\mathbf{U}_1=(\mathbf{A}_1, \mathbf{R}_1, \mathbf{P}_1)$ и $\mathbf{U}_2=(\mathbf{A}_2, \mathbf{R}_2, \mathbf{P}_2)$. Если между $\mathbf{U}_1$, и $\mathbf{U}_2$ имеются следующие отношения: $\mathbf{A}_2 = \mathbf{A}_1 \cup \Delta\mathbf{A}$, $\Delta\mathbf{A}$ — приращение элементов, $\mathbf{R}_2 = \mathbf{R}_1 \cup \Delta\mathbf{R}$, $\Delta\mathbf{R}$ — приращение отношений, $\mathbf{P}_2 = \mathbf{P}_1 \cup \Delta\mathbf{P}$, $\Delta\mathbf{P}$ — приращение свойств, то структура $\mathbf{U}_2$ является ***более расширенной***, чем $\mathbf{U}_1$, или структура $\mathbf{U}_1$ является ***более элементарной***, чем $\mathbf{U}_2$.

Процесс получения более расширенной нагруженной структуры из элементарной будем называть <u>структурированием</u>. Когда структура реализуется для реальной системы, то ее будем называть ***структурной моделью системы***. Может быть получен широкий спектр структурных моделей, формируемых путем структуризации, сгенерированной из первоначальной неполной структурной модели системы. Конечная модель системы должна объединять в себе различные точки зрения на структуру системы.

Процесс структурного моделирования можно формально представить рис. 3.

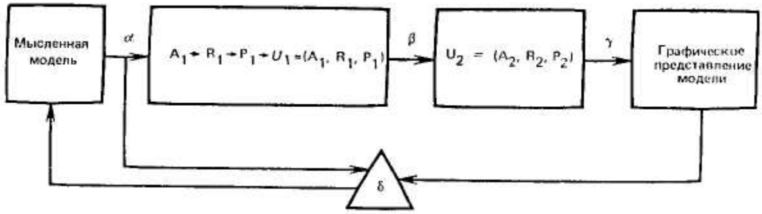


Рис. 3. Формальное представление процесса структурного моделирования

Среди показанных подпроцессов (α — генерация, β — структуризация, γ — визуализация, δ — модификация) α и β являются эмпирическими или интуитивными, а β и γ — формализованными. На данном рисунке не показаны процессы документирования и анализа внешних знаний и стимулов, приведенные на рис. 2.

Для ***представления структурной модели системы*** будем использовать ***таблицы, матрицы и графы*** вместо отношений. При этом обозначения удобно заменить графическим представлением, что облегчает работу со структурной информацией. Графическое представление определяется как двумерная диаграмма структурной модели системы или ее части. На рис. 4 показаны графы, которые рекомендуется широко применять в структурных моделях системы, где *а* — ненаправленный, *б* — направленный, *в* — знаковый, *г* — нагруженный, *д, е* — гиперграфы, *ж* — композиционный граф.

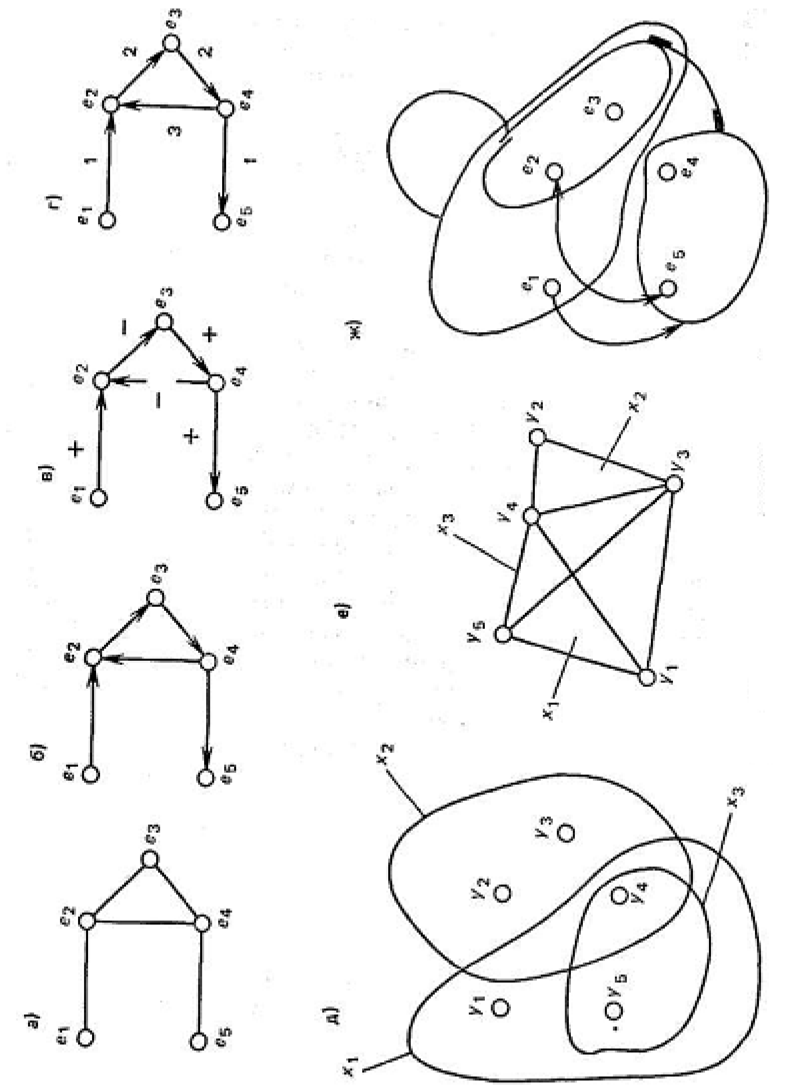


Рис. 4. Графы структурных моделей системы

Применение метода структурного моделирования рассмотрим на примере функциональной структуры. Исходным при этом является список элементов (переменных) и текстовое описание системы. Поэтому для начальной идентификации элементов удобен способ табличного описания структурного элемента, позволяющий иерархически описывать структурный элемент путем последовательного заполнения четырех граф: ***условие, событие, действие и результат***. Таким образом, на каждом уровне иерархии описания в графах «условие» и «результат» идентифицируются элементы-данные, а в графах «событие» и «действие» — элементы-функции. Последовательность описаний внутри каждого уровня и между ними задает структуру управления для данной системы.

На нижнем уровне иерархии процесс проектирования, представляющий собой композицию проектных процедур, описывается в терминах ***элементарных функций и данных*** с точки зрения конкретной предметной области. Рассмотрим пример:
пусть описание функциональной структуры состоит из множества функций

$$F = \{ f_1, f_2, \ldots, f_n \}.$$

Для любых f при $f_i \neq f_j \Rightarrow i \neq j$, тогда можно построить матрицу смежности функций

$$\mathbf{A}=\begin{array}{cc} & \begin{matrix} f_1 & f_2 & \cdots & f_n \end{matrix} \\ \begin{matrix} f_1 \\ \cdot \\ \cdot \\ \cdot \\ f_n \end{matrix} & \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \cdot \\ \cdot & \cdot & & \\ z_{n1} & z_{n2} & \cdots & z_{nn} \end{bmatrix} \end{array},$$

где $z_{ij} = 0$ при $f_i \overline{R} f_j$ и $z_{ij} = 1$ при $f_i \mathbf{R} f_j$.

Символ **R** представляет собой бинарное отношение между функциональными элементами системы.

Зная матрицу **A**, можно определить матрицу достижимости функций путем сложения матрицы смежности и единичной матрицы **I** и умножения полученной матрицы саму на себя (n — 1) раз:

$$\mathbf{D}=(\mathbf{A}+\mathbf{I})^n==(\mathbf{A}+\mathbf{I})^{n-1}\neq(\mathbf{A}+\mathbf{I})^{n-2}.$$

Эта матрица показывает, что длина пути, по которому данные из функционального элемента k достигают функциональный элемент m, находится в диапазоне от 0 до n — 1.

В соответствии с методом структурного анализа далее определяются разбиения множества **F**, уточняющие функциональную структуру путем выполнения следующих процедур:

— вычисления **R*** — рефлексивно-транзитивного замыкания отношения **R**;

— выделения уровней иерархии $\mathbf{L}_i$ на множестве **F** :

$$\mathbf{П}_1 [\mathbf{F}] = [\mathbf{L}_0, \mathbf{L}_1, ..., \mathbf{L}_m],$$

где $\mathbf{П}_1[\mathbf{F}]$ —разбиение $\mathbf{F}$ на уровни иерархии в соответствии с алгоритмом $\mathbf{L}_0=0$;

$$\mathbf{L}_j = \{ f_i \in \bar{\mathbf{L}}_j \mid \mathbf{Q}_j(f_i) = \mathbf{Q}_j(f_i) \mid \mathbf{P}(f_i)\},$$

где

$$\bar{\mathbf{L}}_j = \mathbf{F} - \mathbf{L}_0 - \mathbf{L}_1 - ... - \mathbf{L}_{j-1}; \ \mathbf{Q}_j(f_i) = \{ f_k \in \mathbf{L}_j \mid f_i \mathbf{R}^* f_k\} -$$

множество элементов $\bar{\mathbf{L}}_j$, достижимых из f_i;

$\mathbf{P}_j(f_i) = \{f_k \in \bar{\mathbf{L}}_j \mid f_k \mathbf{R}^* f_i\}$ — множество элементов $\bar{\mathbf{L}}_j$, из которых достижим f_i;

— анализа на достижимость, выполняемого на основе выделения всех компонент связанности графа, задающего анализируемую структуру, и установления недостающих связей в графе;

— анализа на согласованность, выполняемого путем определения циклов внутри каждого уровня иерархии: $\mathbf{П}_2[\mathbf{L}_i] = \{\mathbf{S}_i, \mathbf{C}_i\}$, где $\mathbf{S}_i$— множество не связанных внутри уровня $\mathbf{L}_i$ элементов, $\mathbf{C}_i$ — множество элементов, образующих циклы в $\mathbf{L}_i$, $\mathbf{S}_i = \{f_j \in \mathbf{L}_i \mid \mathbf{Q}_i(f_j) = \{f_j\}\}|$, где

$\mathbf{Q}_i(f_k) = \{f_k \in \mathbf{L}_i \mid f_j \mathbf{R}^* f_k\}$ — множество элементов уровня $\mathbf{L}_i$, достижимых из f_k;

— предметной интерпретации выделенных циклов и сравнения полученной интерпретации каждого цикла с предметными интерпретациями других элементов, принадлежащих тому же уровню, что и рассматриваемый цикл, а также выделения транзитивных связей между элементами несмежных уровней;

— декомпозиции на подсистемы $\mathbf{П}_3[\mathbf{F}] = [\mathbf{U}_1, \mathbf{U}_2, ..., ..., \mathbf{U}_k]$, где

$\mathbf{U}$, — подсистема, которая определяется как последовательность сильно связанных функций, достижимых и предшествующих друг другу, с минимизацией связей с другими подсистемами.

Программная реализация метода интерпретирующего структурного моделирования обеспечивает***: создание и редактирование структурной модели сложной системы, анализ имеющейся структуры, протоколирование хода работ разработчика с интерактивным комплексом программ интерпретирующего структурного моделирования (ИКП ИСМ), создание выходных документов, отражающих текущую структуру моделируемой системы.***

Преимущества ИКП ИСМ заключаются в обеспечении диалогового режима работы и в использовании внешней памяти на магнитных носителях хранения данных о структуре исследуемой системы.

Использование системы меню, а также контроль ответов пользователя позволяют удобно и оперативно «проигрывать» различные варианты структуры моделируемой системы с получением документации на каждом месте.

Метод структурного анализа обеспечивает выполнение следующих функций:

1) создание, модификацию, просмотр матриц взаимодействия элементов системы;

2) вычисление матриц достижимости;

3) вычисление уровней иерархии по рассчитанной ранее матрице достижимости;

4) распечатку всех интересующих пользователя матриц и уровней иерархии в удобной для чтения форме;

5) определение циклов (сильносвязанных множеств) системы внутри каждого уровня иерархии;

6) прекращение сеанса работы с ИКП ИСМ по требованию пользователя с сохранением результатов до следующего сеанса;

7) ведение диалога с неподготовленным пользователем, имеющим общее представление о ИКП ИСМ;

8) выделение транзитивных (избыточных) связей между элементами исследуемой системы.

Входными данными для работы ИКП ИСМ является множество ключевых терминов (словарь), каждый из которых описывает простой функционально законченный на некотором уровне рассмотрения элемент исследуемой системы. Это множество выделяется в процессе анализа предметной области пользователем соответствующей предметной области.

На первом этапе работы пользователь предметной области с помощью ИКП ИСМ в диалоговом режиме специфицирует структуру на заданном множестве элементов системы, определяя связи между элементами. В результате выполнения первого этапа формируется матрица связности исследуемой системы.

На последующих этапах работы с ИКП ИСМ анализируется заданная структура с целью выявления из перечня необходимых свойств, поддерживаемых системой.

Конечным результатом работы с ИКП ИСМ является комплект документов, отражающих результаты анализа исследуемой структуры системы в терминах данной предметной области.

Для обеспечения мобильности все машинно-зависимые функции, использованные в ИКП ИСМ, выделяют в отдельные модули с их полной спецификацией.

В качестве примера рассмотрим построение функциональной структуры подсистемы технологического проектирования процессов и оснащения сборки в машиностроении (рис. 5).

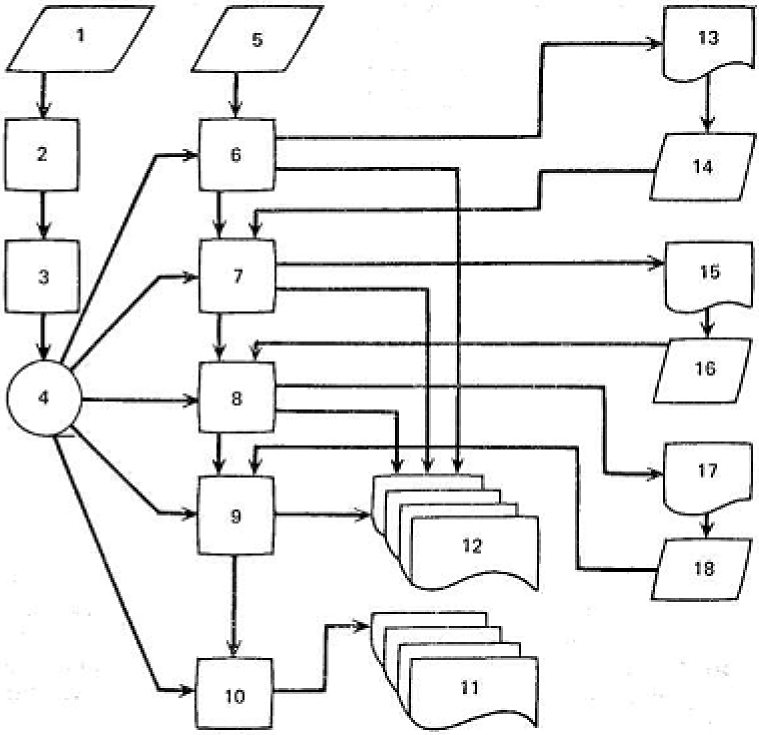


Рис. 5. Функциональная структура подсистемы технологического проектирования

В подсистеме осуществляется автоматизированное проектирование технологических процессов и оснащения сборки, а также оценка технологичности сборочных единиц путем реализации последовательности функциональных преобразований, где первоначальная функциональная декомпозиция рассматриваемой подсистемы состоит из следующих этапов:

1. Разработка модели производственной системы (МПС).
2. Отладка МПС.
3. Печать МПС.
4. Запись МПС на носитель информации.
5. Подготовка исходных данных.

6. Выбор схемы базирования.

7. Определение конструктивной схемы сборочного приспособления.

8. Проектирование технологического процесса.

9. Формирование технологической документации.

10. Определение последовательности установки элементов сборочной единицы.

11. Печать документации.

12. Печать промежуточных результатов проектирования и диагностирования МПС.

13. Выбор вариантов схем базирования.

15. Выбор вариантов конструктивных схем сборочных приспособлений.

17. Выбор вариантов последовательности установки.

14, 16, 18. Дополнение и изменение исходных данных.

Применение метода интерпретирующего структурного анализа позволяет сформировать рекомендации по построению упорядоченной с точки зрения агрегации данных функциональную структуру, из которой следует несогласованность по данным между элементом 4 и элементами 6, 7, 8, 9, 10, а также между элементом 12 и элементами 6, 7, 8, 9 (рис. 6).

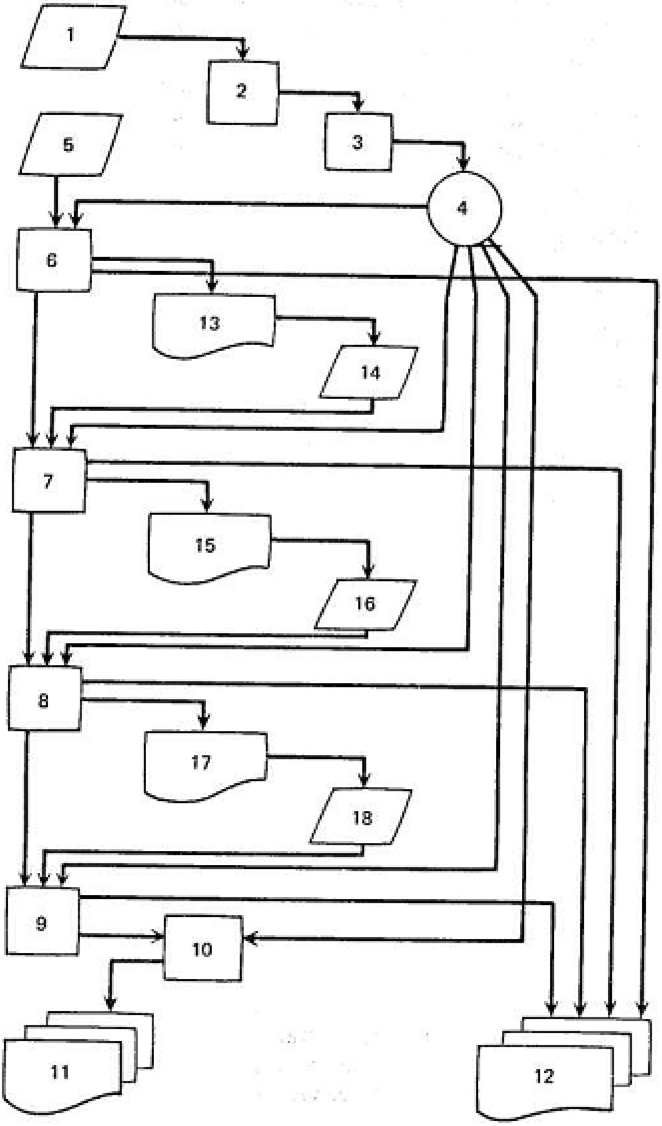


Рис. 6. Иерархически упорядоченная функциональная структура подсистемы технологического проектирования

Предметная интерпретация полученных результатов показывает, что основными функциональными преобразованиями в рассматриваемом процессе проектирования являются преобразования, выполняемые элементами 6, 7, 8 и 9. Анализ текстуального описания элементов 6, 7, 8 и 9 показывает, что они представляются в виде композиции следующих функциональных элементов:

Для элемента 6

6.1. Формирование множества возможных схем.

6.2. Проверка схемы на отсутствие нерациональных сочетаний сборочных баз.

6.3. Предварительная проверка схемы по точности.

6.4. Проверка схемы на существование возможной последовательности установки.

6.5. Выбор из всех возможных схем базирования схемы, оптимальной по заданному критерию, в качестве которого принимается технологическая себестоимость.

Для элемента 7

7.1. Уточнение схемы базирования.

7.2. Определение состава базовых элементов приспособления.

7.3. Укрупненный расчет размерных характеристик базовых элементов.

7.4. Выбор типовой конструкции каркаса сборочного приспособления.

7.5. Выбор конкретных конструктивных элементов сборочного приспособления.

7.6. Определение конкретных конструктивных элементов сборочного приспособления.

7.7. Определение размерных характеристик конкретных конструктивных элементов сборочного приспособления.

7.8. Прочностные расчеты конкретных конструктивных элементов сборочного приспособления.

7.9. Выбор из числа возможных конструктивных схем сборочного приспособления схемы, оптимальной по заданному критерию оптимизации.

Для элемента 8

8.1. Формирование массива перестановок,возможных по условиям базирования и доступа.

8.2. Сокращения пространства возможных престановок.

8.3. Поиск последовательности устоновки, оптимальной по заданному критерию.

Для элемента 9

9.1. Определение состава и последовательности этапов сборки.

9.2. Предварительная оценка рекомендаций по укрупненным показателям.

9.3. Детализация этапов сборки до уровня операций.

9.4. Выбор вида оборудования, инструмента и оснастки.

9.5. Предварительная оценка рекомендаций по укрупненным показателям.

9.6. Нормирование.

9.7. Расчет технико-экономических показателей.

9.8. Выбор оптимального варианта технологического процесса по заданному критерию.

Повторное применение метода интерпретирующего структурного анализа к полученному множеству элементов формирует уточненную функциональную структуру анализируемой подсистемы (рис. 7).

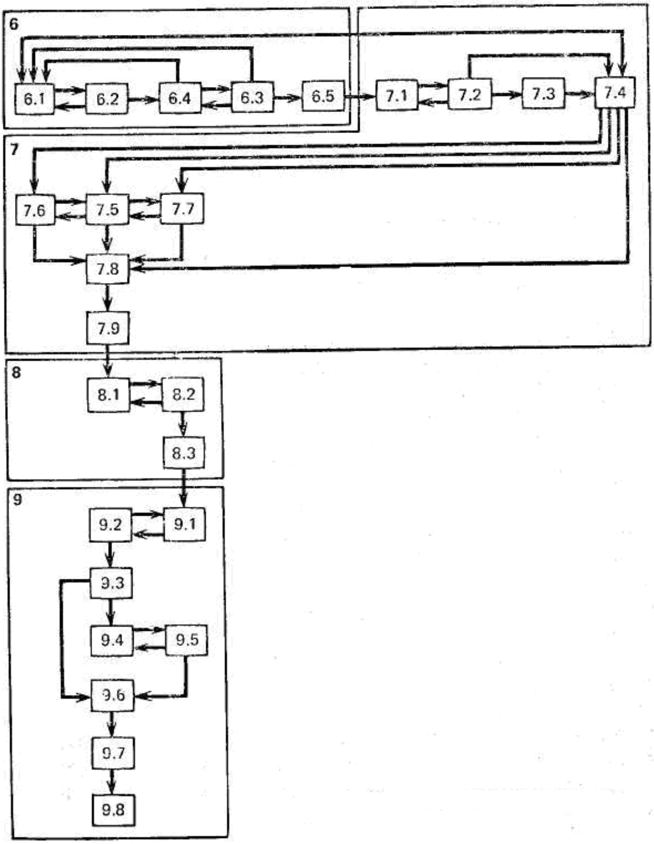


Рис. 7. Уточненная функциональная структура

Анализ структуры на достижимость и сравнение с результатами анализа, полученными на предыдущем этапе, показывают отсутствие связи элемента 7 с элементом 6, что свидетельствует о противоречивости исходной структурной декомпозиции и необходимости ее устранения, например, путем применения метода синтеза функциональной структуры.

1.3. Анализ структур систем

Классификация всего многообразия ***структур систем*** может быть произведена по различным признакам, наиболее употребительными из которых являются: ***этап создания, природа связей, отношение между частями, структурные признаки.***

При использовании ***первого классификационного признака*** все структуры системы будем разделять на подмножества исходных $S_п$, разрабатываемых $S_р$ и производных $S_{пр}$ структур, причем

$$S = S_п \cup S_р \cup S_{пр} .$$

Подмножество $S_п$ составляет структуры, задаваемые или однозначно определяемые множеством требований Z.

Элементами подмножества $S_п$ являются ***структурные, функциональные, электрические, кинематические, пневматические, гидравлические и другие схемы системы***, причем, как правило, ряд свойств этих элементов не известен и подлежит определению в процессе их создания, а следовательно,

$$S_п \cap S_р \neq \varnothing.$$

Подмножество $S_р$ составляет как структуры, разрабатываемые в процессе проектирования, так и некоторые элементы подмножества, исходная информация о которых является недостаточной для их материального воспроизведения.

Элементами подмножества $S_Р$ являются: ***схема связей между элементами структуры, отдельные модули системы, а также геометрические, механические, информационные и другие структуры.***

Подмножество $S_{пр}$ составляют те структуры, которые не разрабатываются в процессе проектирования систем, а являются производными — вытекающими из подмножества исходных и разрабатываемых структур. Очевидно, что для всего подмножества производных структур имеет место соотношение:

$$(S_{\text{п}} \cup S_{\text{р}}) \cap S_{\text{пр}} = \varnothing$$

При использовании ***второго классификационного признака*** все структуры системы будем разделять на подмножества S_k (k =1, 2,…, n) по ***природе связей структуры***, т. е.

$$S = \bigcup_{k=1}^{n} S_k.$$

Элементами S_k (подмножествами) множества S являются ***электрические, кинематические, информационные, функциональные, геометрические и другие структуры***.

При использовании ***третьего классификационного признака*** структуры системы будем разделять на подмножества ***сложных структур*** $S_{\text{с}}$ (элементы которых имеют собственную структуру) и подмножество ***элементарных структур*** $S_{\text{Э}}$ (элементы которых не имеют собственной структуры), причем

$$S_c \cup S_э = S; \; S_c \cap S_э = \varnothing.$$

Следует отметить, что деление структур на сложные и элементарные является условным и определяется задачами и целями проектирования систем.

При использовании ***четвертого классификационного признака*** все структуры системы по ***признаку направления действия связей*** будем разделять на подмножества ***ориентированных*** $S_{\text{о}}$, ***неориентированных*** $S_{\text{но}}$ и ***частично ориентированных*** $S_{\text{чо}}$ структур, структур с обратными связями $S_{\text{ос}}$ и т. д., причем

$$S_o \cup S_{но} \cup S_{чо} = S;$$

$$S_{ос} \subset S.$$

Система в развитии проходит последовательно несколько этапов — от идеи создания до демонтажа (утилизации).

Первый этап — появление идеи созданию системы, когда определяется ее назначение. На этом этапе, как правило, уже известно, к какому классу будет принадлежать система и после формулировки основных функций определяется и семейство, к которому она относится. Если заранее известен и комплекс условий, то неопределенность в выборе вида системы уменьшается.

Второй этап — разработка технического задания (ТЗ) на создание системы. Основная задача этапа состоит в том, чтобы определить различные показатели системы. Определение всего множества характеристик может вызвать определенные трудности, поскольку научно-обоснованных методик для этой процедуры еще практически не разработано. При назначении характеристик обычно руководствуются следующим: основные характеристики системы должны быть лучше, а остальные не хуже, чем у ее прототипа, что является явно недостаточным аргументом и руководящим началом.

Третий этап — проектирование системы. Его основная задача — формирование рекомендаций по определению структуры системы и алгоритма (процесса) ее функционирования. Формализация процесса проектирования пока решена слабо, и в основе деятельности проектировщика лежит, как правило, интуиция, практический опыт и поиск аналогий.

Четвертый этап — конструирование, когда отыскивается конструктивное решение, определяемая структурой системы и множеством элементов.

Пятый этап — изготовление и монтаж системы.

Шестой этап — испытание и сдача в эксплуатацию системы; оценка соответствия полученных характеристик заданным.

Седьмой, заключительный, этап — эксплуатация системы. В процессе эксплуатации системы можно правильно оценить или рассчитать по полученным статистическим данным все эксплуатационные характеристики системы, а после некоторого периода функционирования системы дать правильные выводы об ее эффективности.

Вынашивая идею создания системы, необходимо, прежде всего, принять решение относительно того, создавать или не создавать систему. Выдавая техническое задание на разработку системы, необходимо принять решение относительно значений различных характеристик системы. Приступая к проектированию, следует принять решение относительно структуры будущей системы и т. д.

Всем процессам принятия решения присущи две основные черты:

1) выбор из множества всевозможных решений одного, вполне определенного;

2) решение принимается ради определенной цели.

Из сказанного следует, что каждый ***процесс принятия решения можно описать функцией, аргументами которой являются допустимые варианты решения, а значениями — числа, которые описывают меру достижения поставленной цели.*** Эту ***функцию*** принято называть ***целевой***.

Задача принятия решения сводится тем самым к нахождению максимального (минимального) значения целевой функции, а также к нахождению того конкретного аргумента, при котором это значение достигается.

В результате в каждом процессе принятия решения возникают две проблемы:

1) описать множество допустимых решений и целевую функцию;

2) найти максимум целевой функции и допустимое решение, реализация которого позволит осуществить этот максимум.

Первая проблема является задачей математического описания условий, в которых протекает процесс принятия решения, а также цели, ради которой он проводится.

Различные варианты второй проблемы называют экстремальными задачами.

Следует заметить, что составление математических описаний, математических моделей систем требует, прежде всего, основательных знаний в области функционирования системы, а также опыта математических рассуждений. Наиболее перспективными в этом направлении исследований являются подходы к разработке модели системы с помощью специальных языков моделирования. Однако, использование языков моделирования не исключает создания чисто математических, аналитических моделей, в которых действие различных факторов представляется в некоторой абстрактной форме и учитываются только наиболее существенные черты исследуемой системы.

В таком же положении и вторая проблема оптимизации: для одной части задач существуют хорошо разработанные методы решения, для другой — их вовсе нет.

Наиболее хорошо разработаны методы оптимизации для тех математических моделей систем, которые ***функционируют в евклидовом или целочисленном пространстве***. В этом случае для решения задач оптимизации пригодны: методы вариационного исчисления, принципы оптимальности Понтрягина и Бэллмана и всевозможные их разновидности; классический анализ, линейное, нелинейное и динамическое программирование; целочисленные задачи линейного программирования, некоторые нелинейные задачи и некоторый класс задач на графах.

При анализе системы, используют два типа методик:

1) быстрого анализа системы, цель которого — установить целесообразность создания системы по решению задач сформулированной проблемы, формы ее структуры и др.;

2) подробного обследования системы.

Следует отметить, что при проектировании системы не всегда проводится предварительное ее исследование, поэтому часто имеют случаи, когда неэкономичность и малая эффективность системы выявляются уже после израсходования значительных средств и времени на создание системы. Методика становится методикой лишь в том случае, если метод, положенный в ее основу, содержит в себе алгоритм решения задачи, обеспечен исходными данными и дает возможность решить задачу за конечное (реальное) время. Для начальных этапов проектирования системы в настоящее время еще не разработаны эффективные методики автоматизированного проектрования систем. Такое положение можно объяснить отсутствием четко сформулированного основного направления, вокруг которого должны сгруппироваться все остальные направления исследования системы на ранних этапах их проектирования. Таким стержневым направлением должен явиться ***структурный анализ*** и более высокая его ступень — ***синтез оптимальных структур систем.***

Малое количество исходной информации на начальных этапах проектирования системы требует поиска таких моделей системы, которые были бы обеспечены исходными данными и «работали» бы при минимуме исходной информации. ***Такой моделью является структура системы совместно с совокупностью отношений на ней***.

Проведение структурного анализа позволяет:

- получить информацию о степени «нагруженности» и значимости элементов системы;
- сравнить системы с различными структурами;
- получить информацию о «слабых местах» системы, что дает возможность своевременно произвести доработку структуры системы;
- скорректировать программу обеспечения требуемых характеристик и качества системы.

Таким образом, можно сделать вывод, что ***методика исследования системы, включающая элементы структурного анализа, относится к классу <u>методик предварительного быстрого изучения проблемы</u>, цель которых — установить целесообразность принятого в проекте технического решения.***

1.4. Анализ систем и их структур на начальных стадиях проектирования

Выбирая модели системы на начальных этапах проектирования необходимо руководствоваться практическими соображениями, а именно: модель системы должна быть проще самой системы во всех аспектах, за исключением тех, которые определяют выполнение выбранного отношения эквивалентности.

Выбор модели системы во многом предопределяет методику исследования и получение конечных результатов. При проектировании системы на начальных этапах возникает вопрос, какую абстрактную систему выбрать в качестве ее модели. Этот выбор будет определяться, и обусловливаться следующими обстоятельствами: объемом исходной информации о проектируемой системе; требуемой степенью детализации конечных результатов исследования. Все это в целом определит уровень абстракции при проведении исследования системы. Необходимо иметь в виду, что на каждом этапе познания имеется определенная степень абстракции, а каждый этап познания, имея объектом исследования присущую ему модель, дает вполне определенные, соответствующие именно этому этапу, результаты.

Любая система, любое изделие в процессе своей эволюции, прежде чем стать системой, изделием проходит ряд этапов. Каждый из этапов по-своему важен и по-своему специфичен. Однако представляется, что наибольшее внимание должно быть обращено на ранний этап проектирования, поскольку здесь формируются основные контуры будущей системы. Кроме того, на этом этапе должно быть принято решение о целесообразности построения системы, о путях реализации желаемых свойств на этапе конструктивно-технологической разработки системы. Следовательно, начальный этап определит, что будет заложено в будущую систему. От правильности принятого решения во многом зависит стоимость системы, так как все последующие изменения, доработки и доделки обходятся очень дорого, не принося подчас должного эффекта.

Чтобы принять правильное решение, необходимо рассмотреть и исследовать ряд вариантов, поскольку процесс окончательного получения определенного изделия с желаемыми свойствами включает в себя выбор. Следовательно, модель, являющаяся эквивалентом изучаемой системы, должна быть относительно простой.

Характерной особенностью начального этапа проектирования является ограниченность информации о будущей системе. Исходные данные, которые можно использовать в процессе исследования, обычно содержат: общие требования к характеристикам системы; структурно-функциональной схемы системы с весьма общим описанием принципа ее действия.

В таких условиях стремятся получить максимум возможного из этих минимальных сведений, что заставляет в первую очередь обратиться к структурно-функциональной схеме и содержащейся в ней информации. Основным и общим во всех работах, связанных со структурными исследованиями, является стремление очистить объект исследования от всего вторичного, рассмотреть лишь его наиболее характерные признаки. Так, при структурных исследованиях системы в них не остается иного содержания, кроме связей, их числа, дифференциального порядка, знака и конфигурации. При исследовании систем на ранних этапах их проектирования объект исследования оказывается «очищенным» естественным образом из-за недостатка информации о его свойствах. В подобных условиях исследуются элементы системы и существующие между ними связи, т. е. структурная схема системы.

Анализ системы требует использования основ математической теории отношений, которая оперирует с произвольными по своей природе объектами.

Теория отношений является составной частью теории множеств. Рассмотрим элементы теории отношений, необходимые для более глубокого понимания существа выбранной модели и правил оперирования с нею.

Отношением в множестве M называют закон, который выделяет некоторые $П$ элементов ($П$-парные отношения) из M с указанием, какой из них является первой компонентой, какой второй и т. д.

В дальнейшем будут использоваться лишь бинарные отношения. Под бинарным понимают закон, который выделяет некоторые пары элементов (X, Y) из M с указанием, какой из них считается первой компонентой пары и какой второй компонентой.

Совокупность отношений представимы матрицами отношений. В свою очередь матрицы отношений могут быть истолкованы как аналитический образ графа, или, наоборот, граф может быть истолкован как геометрический образ матрицы отношений, что позволяет совокупности отношений в системе поставить в соответствие граф отношений и изучать в дальнейшем только его свойства. Теория графов позволяет разработать общие формальные приемы исследования конкретной системы, не зависящие от ее сложности.

Информацию, содержащуюся в некотором графе, можно представить в алгебраическом виде матрицей отношений. Эта связь графа и матрицы отношений имеет важное значение, так как позволяет перевести структурные особенности системы на язык чисел, фигурирующих в математических уравнениях, описывающих структуры системы. При изучении структур системы с помощью графов необходимо помнить, что два графа обладают одной и той же структурой, если в одном графе столько же вершин и ребер, как и в другом, и если вершины и ребра одного графа соединены друг с другом так же, как и в другом графе. Равноструктурные графы называются эквивалентными.

Рассмотрим ряд правил оперирования с совокупностью отношений как математическим объектом. Доказано, что бинарные отношения на фиксированном множестве образуют булеву алгебру. В связи с этим операции конъюнкции, дизъюнкции и отрицания определяются обычным способом и все тождества булевой алгебры применимы к алгебре отношений.

Выводы для практики исследования системы, следующие отсюда, позволяют отметить: используя тождественные формулы булевой алгебры, можно отыскать новые конфигурации структурных схем системы, эквивалентные исходной, но более удобные для анализа. Кроме того, поиск эквивалентных структур позволяет найти и указать общие свойства различных типов и видов систем, на первый взгляд, различных по своей структуре.

Известно также, что матрицы отношений можно складывать и умножать как обычные матрицы. Отношения представимы также алгебрами нумероидов.

Нумероидом называют множество A с определенными на нем сложением $x+y$ и умножением xy, если:

1) сложение и умножение суть однозначные операции, применимые к любым элементам из A;

2) сложение и умножение ассоциативны и коммутативны:

$$x+(y+z) = (x+y)+z;\ x(yz) — (xy)z;\ x+y = y+x;\ xy = yx.$$

3) сложение имеет хотя бы один нуль, такой, что $0+x = x+0 = x$ для всех x из A, а умножение — хотя бы одну единицу, такую, что $1x=1x=x$ для всех x из A;

4) умножение дистрибутивно относительно сложения:

$$x(y+z) = xy+xz.$$

Нумероид называют абсорбентом, если сложение в нем абсорбитивно по отношению к умножению, т. е., если для x и y из нумероидов справедлив закон абсорбции (поглощения) $x+xy = x$.

Вывод для практики, следующий отсюда, сводится к тому, что, если рассматривается вопрос только лишь о наличии либо отсутствии отношений между элементами системы, то все арифметические операции в алгебре отношений должны подчиняться правилам проведения операций в абсорбентах, т. е.

$$1+x = 1;\ 0\cdot x = x\cdot 0 = 0;\ xx = x.$$

Это справедливо также и потому, что единичный поднумероид (т. е. содержащий только нули и единицы) есть двухэлементная булева алгебра.

Когда в совокупности отношений устанавливается логическое понятие «r влечет S» (или «имеет следствием»), определенное следующим образом: $r \subset S$, означающее, что из $\alpha r \beta$ следует $\alpha S \beta$, то могут потребоваться другие виды исчислений.

Действительно, если некоторые элементы находятся в отношении r, $ar \subset S$, то эти элементы тем более находятся в отношении S.

В этом случае для исследования совокупности отношений можно привлечь одно из экстремальных исчислений, а именно, максимультипликативное. Это исчисление получается, если на множестве неотрицательных действительных чисел положить

$$x \mathbin{\widehat{+}} y = \max(x,\ y);\ \ x \mathbin{\widehat{\cdot}} y = xy;\ \ \widehat{0} = 0;\ \ \widehat{1} = 1.$$

Здесь знак «$\wedge$» (крышка) означает, что арифметические операции выполняются в одном из экстремальных исчислений.

Другим экстремальным исчислением, находящим применение, является миниаддитивное. Оно получается, если множество действительных чисел, включая и бесконечность, положить:

$$x \mathbin{\widehat{+}} y = \min(x,\ y);\ \ x \mathbin{\widehat{\cdot}} y = x + y;\ \ \min(\infty,\ x) = \min(x,\ \infty) = x;$$
$$\infty + x = x + \infty = \infty.$$

Поскольку все числа, которыми в последующем мы будем оперировать, являются положительными, то введенное в рассмотрение миниаддитивное исчисление будет сильным и абсорбентным.

1.5. Анализ структурно-топологических характеристик систем

Проектирование систем начинается с составления структурных схем, которые характеризуются множеством качественных свойств, обусловливаемых множеством элементов системы и связей между ними. Как уже отмечалось, структурная схема системы — это условное графическое изображение элементов системы и связей между ними, которое удобно анализировать с помощью графов. При таком подходе в ряде случаев можно найти адекватные задачи в теории графов, что позволяет воспользоваться для решения задач проектирования известными математическими методами.

Граф дает наглядное представление порядка и вида отношений между элементами двух множеств, что еще раз указывает на целесообразность его использования для изображения структурных схем систем. При решении одних задач элементам системы можно поставить в соответствие ребра графа, а связям — его вершины; при решении других задач элементам системы ставят в соответствие вершины графа, а связям между ними — ребра. Граф, полученный в первом случае, будем называть реберным, а во втором случае — вершинным. Эти два подхода не исключают друг друга. Даже при исследовании одной и той же проблемы возникает необходимость преобразовывать один граф в другой.

Рассмотрим структурную схему, например, производственной системы (рис. 8). Наиболее общую модель производственной системы можно получить, если расчленить ее на функциональные подсистемы. В нашем примере производственная система расчленена на само производство, систему управления производством и обеспечение ресурсами производства.

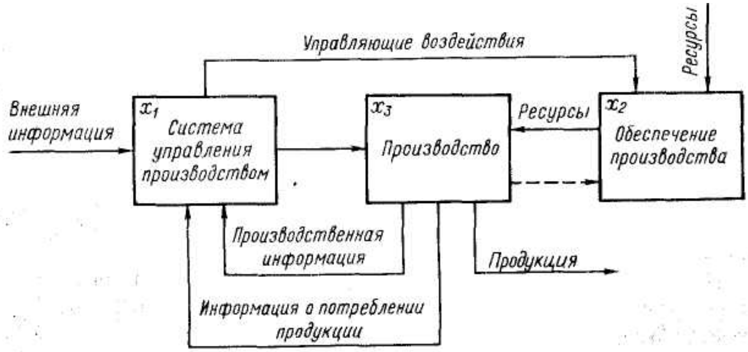


Рис. 8. Структурная схема производственной системы

Представим рассматриваемую структурную схему производственной системы в форме графа: подсистемы этой системы —- вершинами графа, а потоки информации, ресурсов и продукции —- ребрами графа. На рис. 9 изображена графовая модель структурной схемы рассматриваемой производственной системы.

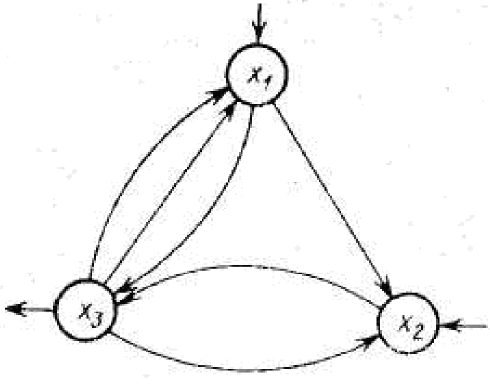


Рис. 9. Граф структурной схемы производственной системы

Данный пример иллюстрирует методику построения вершинного графа.

Модели сложных систем, представляемые в виде вершинных графов, имеют следующий недостаток: как физическое содержание отдельных элементов, так и логические условия их осуществления объединены в одних и тех же элементах — в вершинах графа. Это обстоятельство чрезвычайно затрудняет анализ, делая его индивидуальным для каждой структуры, представленной вершинным графом.

Переход от вершинных графов к реберным дает возможность придать все физические свойства элементов дугам графа, а все логические условия сосредоточить в вершинах. Это существенно упрощает формирование логической структуры системы и позволяет разработать полностью формализованные методы построения структурных схем системы с вершинами, в которых могут реализоваться любые логические функции.

Как известно, в теории графов существуют две классические задачи: задача Эйлера о кенингсбергских мостах и задача Гамильтона. Задача Эйлера состоит в отыскании путей, проходящих по ребрам графа, а задача Гамильтона состоит в отыскании путей, проходящих по вершинам графа.

Таким образом, задачу Эйлера можно рассматривать как задачу упорядочения дуг реберного графа, а задачу Гамильтона — как задачу упорядочения вершин вершинного графа.

Доказанные Эйлером теоремы позволяют однозначно по виду графа решать вопросы о существовании эйлеровых путей. Для гамильтоновых путей такого критерия не существует. Для большинства таких графов не разработаны даже удовлетворительные алгоритмы, позволяющие установить наличие гамильтоновых путей.

Таким образом, возникает противоречие: анализ сложной системы позволяет установить перечень элементов системы и определить систему бинарных отношений на множестве этих элементов, т. е. построить матрицу непосредственных путей. Матрица непосредственных путей всегда позволяет построить ориентированный вершинный граф, анализ которого (т. е. поиск целесообразных путей, проходящих через его вершины) связан с очень большими, иногда непреодолимыми трудностями.

В то же время матрица непосредственных путей в большинстве случаев не позволяет непосредственно построить ориентированный реберный граф, анализ которого в принципе всегда возможен.

Построение реберного графа, эквивалентного вершинному, подчинено более сложным закономерностям. Ниже будут рассмотрены правила и порядок их использования при построении реберного графа. При этом будем рассматривать только ориентированные графы. Если исходный вершинный граф неориентированный, то его необходимо преобразовать в ориентированный с помощью операции удвоения, когда любой неориентированный путь представляется парой противоположно ориентированных путей (рис. 10).

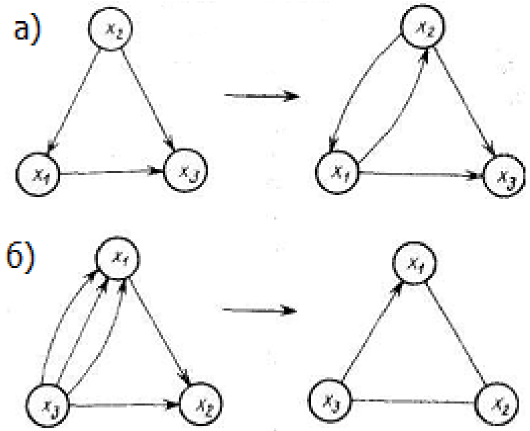


Рис. 10. Преобразование графов

Далее имеющиеся в исходном вершинном графе параллельные пути объединяются в один путь на основании правил операций в адсорбентах.

Построим реберный граф, эквивалентный вершинному, воспользовавшись графом структуры производственной системы (рис. 9).

Объединяя параллельные пути, идущие из третьей вершины в первую, получим преобразованный граф следующего вида (рис. 11).

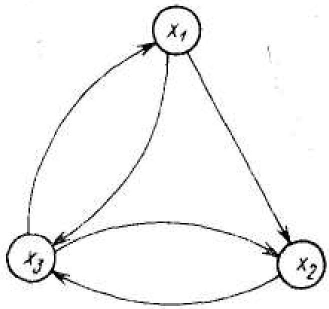


Рис. 11. Преобразованный граф производственной системы

Запишем для него матрицу отношений:

$$\|a_{ij}\|_1^3 = \begin{array}{c c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} & \\ \left\|\begin{array}{c|c|c} & 1 & 1^* \\ \hline & & 1 \\ \hline 1 & 1^* & \end{array}\right\| & \begin{array}{l} x_1\ (2) \\ x_2\ (1) \\ x_3\ (2) \end{array} \\ \begin{array}{ccc} (1) & (2) & (2) \end{array} & \end{array} \qquad (1)$$

В пустых клетках здесь и далее будут подразумеваться нули. По матрице (1) построим еще две матрицы $\|b_{ij}\|$ и $\|c_{ij}\|$. Элементы первой из них определим по формуле

$$b_{ij} = a_{ij}\left(\sum_{i=1}^{n} a_{ij} + \sum_{j=1}^{n} a_{ij}\right) \qquad (2)$$

где $\sum_{i=1}^{n} a_{ij}, \sum_{j=1}^{n} a_{ij}$ — суммы единиц по строкам и столбцам матрицы $\|a_{ij}\|$; их значения проставлены в круглых скобках в (1).

Производя вычисление по (2), получим матрицу $\|b_{ij}\|$ следующего вида:

$$\|b_{ij}\| = \begin{array}{c c} \begin{array}{ccc} x_1 & x_2 & x_3 \end{array} & \\ \left\|\begin{array}{c|c|c} & 4 & 4 \\ \hline & & 3 \\ \hline 3 & 4 & \end{array}\right\| & \begin{array}{l} x_1 \\ x_2 \\ x_3 \end{array} \end{array} \qquad (3)$$

Элементы матрицы c_{ij} вычислим по формуле:

$$c_{ij} = a_{ij}\,(\delta_{j/i}b_{ij} + \delta_{i/j}b_{ij}), \qquad (4)$$

где $\delta_{j/i}b_{ij} = (b_{ij}$ — $\min b_{ij})i$ — превышение суммы полустепеней исхода и захода данного элемента $a_{ij}=1$ в i-й строке по сравнению с элементом, имеющим в этой строке минимальное значение $b_{ij}\neq 0$; $\min b_{ij}$ — минимальное по j в i-й строке значение суммы полустепеней захода и исхода для элементов $a_{ij} = 1$; $\delta_{i/j}b_{ij} = (b_{ij}$ — $\min b_{ij})j$ — превышение суммы полустепеней исхода и захода данного элемента $a_{ij} = 1$ в j-м столбце по сравнению с элементом, имеющим в этой строке минимальное значение $b_{ij}\neq 0$; min b_{ij} — минимальное по i в j-столбце значение суммы полустепеней захода и исхода для элементов $a_{ij}=1$.

Подсчитаем величины элементов матрицы $||c_{ij}||$.

$$c_{12}=a_{12}\left(\delta_{2/1}b_{12}+\delta_{1/2}b_{12}\right);$$

$$\delta_{2/1}b_{12}=\left(b_{12}-\min_{2} b_{12}\right)_1=4-4=0;$$

$$\delta_{1/2}b_{12}=\left(b_{12}-\min_{1} b_{12}\right)_2=4-4=0;$$

$$c_{12}=0;$$

$$c_{13}=a_{13}\left(\delta_{3/1}b_{13}-c_{13}b_{13}\right);$$

$$\delta_{3/1}b_{13}=\left(b_{13}-\min_{3} b_{13}\right)_1=4-3=1;$$

$$\delta_{1/3}b_{13}=\left(b_{13}-\min_{1} b_{13}\right)=4-4=0;\ c_{13}=1.$$

Продолжая вычисления, получим, что $c_{23}=0$, $c_{32}=1$, $c_{31}=0$. Таким образом, окончательно имеем

$$||c_{ij}||_1^3=\begin{array}{c}\begin{array}{ccc}x_1 & x_2 & x_3\end{array}\\ \left\|\begin{array}{c|c|c} & 0 & 1\\ \hline & & 0\\ \hline 0 & 1 & \end{array}\right\|\end{array}\begin{array}{c}\\ x_1\\ x_2\\ x_3\end{array} \quad (5)$$

Пометим в матрице $||a_{ij}||_1^3$ звездочками те элементы, для которых c_{ij} в матрице (5) не равны нулю. Такими элементами будут a_{32} и a_{13}. Расширим матрицу $||a_{ij}||_1^3$, дополнив ее двумя строками и столбцами, а элементы a_{13} и a_{32} заменим парой новых, как это показано штриховой линией в матрице (6).

$$\|a_{ij}\|_1^5 = \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ \left\| \begin{array}{c|c|c|c|c} & 2 & \bullet & 1 & \\ \hline & & & & \\ \hline 1 & \bullet & & & 1 \\ \hline & & 1 & & \\ \hline & 1 & & & \end{array} \right\| \end{array} \begin{array}{l} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{array}. \qquad (6)$$

По матрице $||a_{ij}||^5{}_1$ вновь построим матрицы $\| b_{ij} \|_1^5$ и $\| c_{ij} \|_1^5$. Они будут иметь вид:

$$\|b_{ij}\|_1^5 = \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ \left\| \begin{array}{c|c|c|c|c} & 4 & & & 3 \\ \hline & & 3 & & \\ \hline 3 & & & & 3 \\ \hline & & 3 & & \\ \hline & 3 & & & \end{array} \right\| \end{array} \begin{array}{l} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{array}; \qquad (7)$$

$$\|c_{ij}\|_1^5 = \begin{array}{c} \begin{array}{ccccc} x_1 & x_2 & x_3 & x_4 & x_5 \end{array} \\ \left\| \begin{array}{c|c|c|c|c} & 2 & & & \\ \hline & & & & \\ \hline & & & & \\ \hline & & & & \\ \hline & & & & \end{array} \right\| \end{array} \begin{array}{l} x_1 \\ x_2 \\ x_3 \\ x_4 \\ x_5 \end{array}. \qquad (8)$$

У матрицы (8) только один элемент $c_{ij} = c_{12,}$ соответствующий $a_{ij}= a_{12} = 1$, не равен нулю.

Дополним матрицу (6) еще одной строкой и одним столбцом, а элемент a_{12} заменим парой новых

$$\|a_{ij}\|_1^6 = \begin{array}{c|c|c|c|c|c|c}
x_1 & x_2 & x_3 & x_4 & x_5 & x_6 & \\
\hline
 & \bullet & & 1 & & 1 & x_1 \\
 & & 1 & & & & x_2 \\
1 & & & & 1 & & x_3 \\
\bullet & & 1 & & & & x_4 \\
 & 1 & & & & & x_5 \\
 & 1 & & & & & x_6
\end{array} \qquad (9)$$

Процесс преобразования исходной матрицы (2) заканчивается на (9), так как построенная для нее матрица $\|c\|^6_1$ будет иметь вес $c_{ij} = 0$ для соответствующих им $a_{ij} = 1$.

Таким образом, преобразование и расширение исходной матрицы бинарных отношений следует производить до тех пор, пока в соответствующей ей матрице $||c_{ij}||$ все элементы не станут равными нулю.

Дальнейший этап преобразования сводится к определению нумерации вершин нового реберного графа и распределению между ними ребер, эквивалентных исходному вершинному графу. Рассмотрим алгоритм определения нумерации вершин нового реберного графа.

1. Если в матрице $||a_{ij}||^n_1$ имеется больше, чем один пустой столбец, то к матрице слева и сверху приписываются еще один столбец и строка, причем новый столбец остается пустым, а в новой строке на местах, соответствующих пустым столбцам, проставляются единицы.

2. Если имеется более, чем одна пустая строка, то к матрице приписываются слева и снизу столбец и строка, причем приписанная строка остается пустой, а в новом столбце на местах, соответствующих пустым строкам, проставляются единицы.

3. Пустому столбцу приписывается первый номер. Этот же номер присваивается строке, имеющей тот же индекс, который проставляется справа от матрицы в колонке под начальным номером $N_{нач}$. Если в исходной матрице $||a_{ij}||^n{}_1$ нет пустых столбцов, то первый номер может быть присвоен произвольному столбцу.

4. Отметим первый столбец и опустимся по нему вниз до встречи с элементом, отличным от нуля. Вычеркиваем его. Затем, перемещаясь по строке, где был встречен первый элемент, отличный от нуля, вычеркиваем все элементы $a_{ij} \neq 0$. Если в очередном столбце нет более элементов, отличных от нуля, то ему присваивается тот же номер (индекс). Эти же номера (индексы) присваиваются тем же строкам. Они проставляются в колонке под рубрикой $N_{нач}$. Операция повторяется до тех пор, пока в отмеченном столбце не останется ни одного не вычеркнутого элемента. Затем переходят к следующему столбцу матрицы, присваивая ему очередной номер (индекс). Операция повторяется, пока в матрице $||a_{ij}||^n{}_1$ не останется ни одного невычеркнутого элемента.

5. Полученные номера (индексы) столбцов выписывают в виде дополнительной строки $N_{нач}$ под матрицей $||a_{ij}||^n{}_1$.

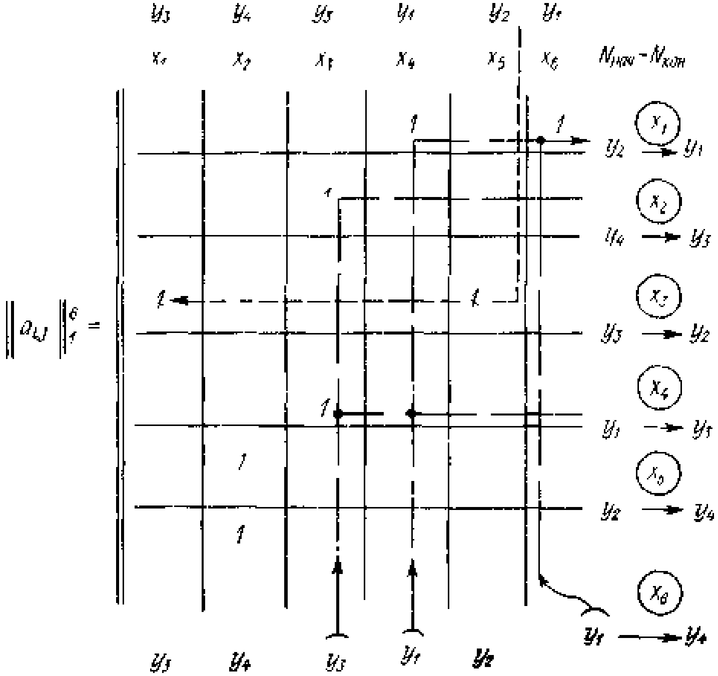
6. Составляют второй дополнительный столбец $N_{кон}$ (номер конечный) путем проектирования $N_{нач}$ на матрицу $||a_{ij}||^n{}_1$. Проектирование осуществляется следующим образом. Номер (индекс) из строки $N_{нач}$ перемещается по столбцу до встречи с $a_{ij} \neq 0$. Затем по этой строке номер (индекс) из строки $N_{нач}$ перемещается в столбец $N_{кон}$.

7. На этом процедура преобразования заканчивается. Номера (индексы) столбцов $N_{нач}$ и $N_{кон}$ указывают на вершины, которые связаны ребром с номером (индексом), соответствующим номеру (индексу) строки, бывшем ранее вершиной в исходном вершинном графе.

8. По полученной матрице вместе со столбцами $N_{нач}$ и $N_{кон}$ строится эквивалентный граф.

Осуществим с помощью описанного алгоритма преобразование матрицы (9). В ней нет строк и столбцов, содержащих только нули, поэтому процедуры, определенные первыми двумя пунктами алгоритма, отпадают. Выбор первого столбца осуществляется произвольно. Пусть им будет столбец 6. Присвоим ему индекс y_1. Строке 6 также присваивается этот же индекс и проставляется в столбце $N_{нач}$. Элемент a_{16} =1. Вычеркиваем его. Двигаясь по первой строке, вычеркиваем элемент a_{14} =1. В столбце 4 нет больше отличных от нуля элементов. Ему также присваивается индекс y_1 и соответственно строке 4 индекс проставляется в столбце $N_{нач}$.

Переходим к следующему столбцу. Пусть им будет столбец 5. Присваиваем ему индекс y_2 (в строке 5). В столбце 5 отличен от нуля элемент a_{35}. Вычеркиваем его. Двигаемся по строке 3 до встречи с отличным от нуля элементом a_{13}. Вычеркиваем его. В столбце 1 нет больше ненулевых элементов. Ему присваивается индекс y_2 (в первой строке). Переходим к следующему столбцу. Пусть им будет столбец 3. Присвоим ему (и строке 3) индекс y_3 и т. д. Получим новую нумерацию в виде столбца $N_{\text{нач}}$. Запишем новую строку $N_{\text{нач}}$ и спроектируем ее на матрицу, что даст столбец $N_{\text{кон}}$.



(10)

По матрице (10) видно, что эквивалентный граф имеет четыре вершины y_1, y_2, y_3, y_4, соединенные шестью ребрами.

Полученный граф эквивалентен исходному (см. рис. 5), поскольку в нем сохраняется та же система бинарных отношений.

Изложенная методика построения графа достаточно сложна для ее ручной реализации. Поэтому существует программа, ориентированная на реализацию ее на ЭВМ.

1.6. Анализ количественных характеристик структур систем

Проводя структурный анализ системы и их подсистем, необходимо располагать методикой, позволяющей определять некоторые структурные характеристики системы и давать им количественную оценку.

Качественная характеристика оценки структур системы и их подсистем необходима для определения на начальной стадии проектирования систем степени пригодности данной структурной схемы для выполнения поставленных перед системой задач.

Весьма важными показателями при оценке структурных схем систем являются установление роли и значения отдельных элементов системы, определение наиболее уязвимых мест и слабых схемных решений. Кроме того, при разработке множества вариантов структурных схем системы и их подсистем, возникает необходимость в качественном сравнении этих схем.

Представляя структурную схему системы или ее подсистемы математической моделью в виде графа, еще ничего нельзя сказать относительно качественных показателей самой системы. Для определения качественных показателей структур системы в количественном выражении требуется разработать методику определения этих количественных выражений. Используя указанные методики, можно разработать алгоритмы и программы реализации этих методик в подсистеме автоматизированного анализа структурных схем систем.

Параметры структурной схемы системы и ее подсистем, определяющие качество схемы, должны быть простыми и не требовать для своего определения данных больше, чем их содержится в самой структурной схеме. Для определения ***качественных характеристик структур систем и их подсистем*** можно ввести ***следующие критерии***.

1. Связность структуры. Этот критерий позволяет определить отсутствие требуемых связей между элементами системы, «висящие» вершины и др.

Связность элементов структуры, если она представлена в форма ориентированного графа, можно определить с помощью матрицы связности $S = ||s_{ij}||$. Элементы этой матрицы вычисляются с использованием матрицы смежности этого графа:

$$A_{\Sigma} = \sum_{k=1}^{n} A^k .$$

В этом случае элемент

$$s_{ij} = \begin{cases} 1, \text{ если } a_{ij}^{\Sigma} \geq 1; \\ 0, \text{ если } a_{ij}^{\Sigma} = 0. \end{cases} \quad (11)$$

В случае представления структуры системы в форме неориентированного графа связность его элементов определяется с помощью следующего выражения:

$$\frac{1}{2}\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij} \geq n-1; i \neq j. \quad (12)$$

В этом выражении правая часть неравенства определяет необходимое минимальное число связей в структуре неориентированного графа, содержащего *n* вершин.

2. **Структурная избыточность**. Под структурной избыточностью будем понимать такой структурный параметр, который отражает превышение общего числа связей над минимально необходимым.

В соответствии с (12) структурную избыточность *N* будем определять из выражения

$$N = \frac{1}{2}\left[\sum_{i=1}^{n}\sum_{j=1}^{n} a_{ij}\right]\frac{1}{n-1} - 1.$$

Эту характеристику структуры системы будем использовать для косвенной оценки экономичности и эффективности системы. Системы, представляемые моделью «полный граф», имеют максимальную избыточность. У таких систем величина параметра структурной избыточности больше нуля *(N*>0). У системы с минимальной избыточностью этот параметр равен нулю *(N*=0). Несвязные структуры характеризуются отрицательной величиной *N (N* < 0).

Отсюда следует, что наибольшей эффенктивностью обладают системы, у которых наибольшее численное значение параметра структурной избыточности.

В ряде задач анализа структурной эффективности системы будем использовать такой ***структурный параметр***, который учитывает ***неравномерность распределения связей*** μ^2.

Как известно, в неориентированном графе, имеющем *m* ребер и *n* вершин, равномерное распределение связей можно охарактеризовать средней степенью вершины $\overline{\rho}$ = 2*m*/*n*. Введем понятие отклонения

ρ_1 — $\overline{\rho}$, где $\overline{\rho}$ — действительная степень *i*-й вершины заданного графа.

Определим квадратичное отклонение заданного распределения степеней вершин от равномерного:

$$\mu^2 = \sum_{i=1}^{n} (\rho_i - \bar{\rho})^2 = \sum_{i=1}^{n} \rho_i^2 - 2\bar{\rho} \sum_{i=1}^{n} \rho_i + 4m^2/n =$$

$$= \sum_{i=1}^{n} \rho_i^2 - 2 \cdot 2m/n2m + 4m^2/n = \sum_{i=1}^{n} \rho_i^2 - 4m^2/n.$$

Структурный параметр μ^2 характеризует уровень «недогрузки» заданной структуры, имеющей m ребер и n вершин, в достижении максимальной связности. Этот показатель в относительном выражении можно использовать для сравнения различных вариантов структур системы.

3. **Структурная компактность**. Под структурной компактностью будем понимать величину D, отражающую общую структурную близость элементов между собой в системе. Близость двух элементов i и j между собой будем определять через минимальную длину пути для ориентированного графа и цепи для неориентированного. Численное значение величины D определим из выражения

$$D = \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij}, (i \neq j).$$

Можно использовать не абсолютное значение параметра D, а относительное

$$D_{\text{отн}} = (D/D_{\min}) - 1,$$

где $D_{\min} = n\ (n - 1)$ — минимальное значение компактности для структуры системы, промоделированной полным графом.

Структурную компактность можно характеризовать и другой характеристикой — диаметром структуры:

$$d = \max_{ij} d_{ij}.$$

Параметры $D_{\text{отн}}$, D и d дают интегральную оценку инерционности различных процессов в системе, а в случае равных значений μ^2 и N их возрастание отражает увеличение числа разделяющих связей в структуре, характеризуя тем самым снижение общей эффективности системы.

4. **Степень централизации в структурах**. Этот параметр характеризует степень распределения связей в структуре системы. Численное значение этого параметра будем определять из выражения:

$$\alpha = (n-1)(2z_{\max} - n)\frac{1}{z_{\max}(n-2)},$$

где z_{max} — максимальное значение величины

$$z_i = \frac{D}{2}\left(\sum_{i=1}^{n} d_{ij}\right)^{-1}, i = 1, 2, ..., n, i \neq j;$$

значение этого параметра колеблется в пределах от 0 до 1.

На рис. 12 изображены основные модели структур системы.

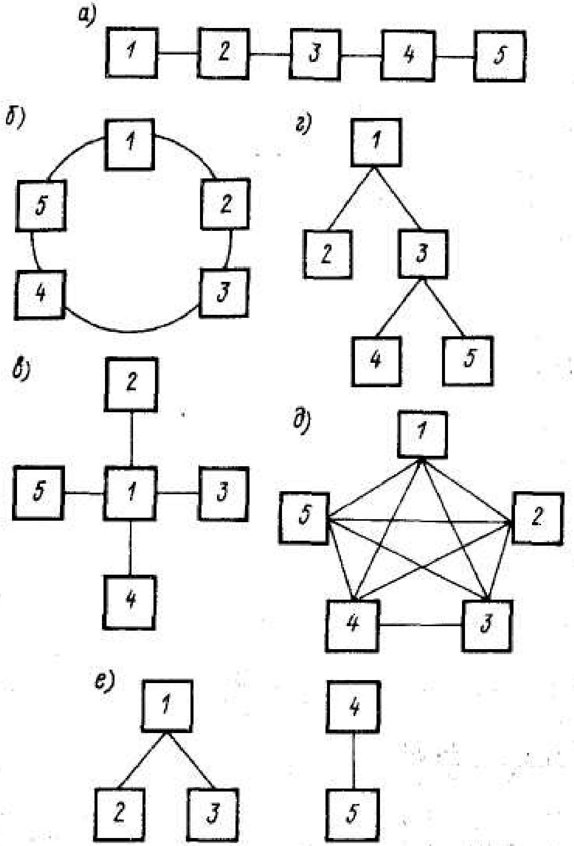


Рис. 12. Модели типовых структур системы

Структура, изображенная на рис. 12, *в,* имеет максимальную степень централизации — α=1. При равномерном распределении связей величина степени централизации равна нулю (рис. 12, *б* и *д)*

5. **Ранг элемента**. Под рангом элемента будем понимать отношение рассматриваемого элемента графа структурной схемы системы с другими элементами этого графа.

Данный критерий позволяет распределить элементы структурной схемы системы в порядке их значимости. Значимость элемента здесь определяется только качеством связей данного элемента с другими.

Величина ранга еще не дает нам полной характеристики главенства того или иного элемента, поскольку не учитывает ряд других характеристик элементов — точностных, скоростных и др. Однако характеризуя значимость элемента рангом — структурным критерием качества, можно высказать предположение, что чем выше ранг элемента, тем сильнее он связан с другими элементами схемы — тем тяжелее будут условия работы системы и ее подсистем в случае выхода из строя элемента, обладающего наивысшей величиной ранга. Подробная методика определения ранга элемента приводится далее.

6. **Множество сочленений**. Под множеством сочленения структуры системы будем понимать такое множество ее элементов, удаление которых делает структуру несвязной. При подобных исследованиях может возникнуть вопрос, удаление каких ребер нарушит связность графа, что равносильно выходу из строя элементной структурной схемы системы.

Значимость элемента в этом случае можно определять и оценивать его принадлежностью к множеству сочленения, т. е. такому множеству, удаления которого из исходного графа делает его несвязным.

В дальнейшем будем говорить не вообще о множестве сочленения, поскольку граф в целом уже есть множество сочленения, а о минимальном множестве сочленения, так как исследователя в первую очередь интересует тот минимум элементов, удаление которых разрушит схему. Подробная методика определения множества сочленения будет приведена далее.

Приведенные критерии оценки качества структурных схем системы и их подсистем позволяют количественно оценить их качество. Это позволяет оценить топологические свойства структур системы. С точки зрения топологии внутренних связей можно выделить следующие основные типы структур (рис. 12): *а* — последовательная; *б* — кольцевая; *в* — радиальная; *г* — древовидная; *д* — типа «полный граф»; *е* — несвязная.

Рассмотрим пример применения количественных характеристик к анализу топологических свойств этих структур. Результаты анализа представлены в табл. 1.

Таблица 1

Вид структуры	N	μ^2	D	α	R
а	0	1,2	1,0	4	0,7
б	0,25	0	0,5	2	0
в	0	7,2	0,6	2	1,0
г	0	3,2	0,7	3	0,7
д	1,50	0	0	1	0
е	—0,25	—	—	—	—

Анализируя данные табл. 1, можно сделать следующие выводы.

1. Для несвязных структур $N < 0$, для структур без избыточности (последовательная, радиальная, древовидная) $N = 0$, для структур с избыточностью по связям (кольцевая, типа «полный граф») — $N > 0$.

2. Структуры (последовательная, радиальная, древовидная) с N=0 различаются по показателю μ^2. Наибольшую неравномерность связей имеет радиальная структура.

3. Наибольшую близость элементов (показатель *D)* имеет структура типа «полный граф», наименьшую — последовательная; радиальная и кольцевая структуры, неразличимые по показателю α, имеют различные значения *D*.

4. Радиальная и древовидная структуры, имеющие одинаковые или близкие значения *N, D,* α*,* значительно отличаются по показателю μ^2 и *R,* что соответствует физическому смыслу, ибо отход от полной централизации в структуре ведет к большой равномерности распределения связей по элементам.

Из приведенных критериев оценки структурных схем систем и их подсистем особый интерес представляет структурный параметр — ранг элемента.

Ниже будет изложена методика определения ранга и полного ранга элемента структурной схемы системы и алгоритм реализации этой методики.

Рассмотренные выше структурные характеристики системы получают на основе информации о составе элементов и их связях. Дальнейшее развитие методологии построения структурных параметров для решения задач структурного анализа может быть основано на неструктурной информации за счет введения числовых функций на графах. Это позволяет наряду с составом элементов и направленностью их взаимодействия учитывать при решении задач структурного анализа другие интересующие стороны взаимодействия (временные, надежностные, стоимостные и др.).

Рассмотрим метод определения величины ранга элемента структуры. При решении задач структурного анализа систем и их подсистем возникает вопрос, какой элемент структуры наиболее значимый. Иначе говоря, необходимо знать, какое влияние оказывает на общую надежность структуры системы выход из строя того или иного ее элемента.

На начальных этапах проектирования достаточно определить значимость элемента некоторыми структурными параметрами, и, наоборот, анализировать и оценивать качество структурной схемы этими структурными параметрами значимости. При дальнейшем исследовании систем и их подсистем необходимо уточнить полученные на ранней стадии формирования схем показатели значимости, дополнив их параметрами, характеризующими функционирование принятых схем систем и определяющими эффективность их функционирования. На начальном этапе для такой оценки обычно данные отсутствуют.

Количественная оценка значимости сформулирована при анализе структуры отношения доминирования в теории математических отношений, которая оперирует с произвольными по своей природе объектами. В связи с этим некоторые ее положения могут быть перенесены и на теорию анализа структур систем.

Одним из критериев теории отношения является ранг элемента структуры системы, которым определяется количество отношений этого элемента структуры с другими ее элементами, определяющее значимость элемента в структуре. Более высокие ранги соответствуют элементам, имеющим большее число связей с другими элементами структуры. Численно ранг элемента равен числу всех одночленных и двучленных отношений доминирования, которые данный элемент может осуществлять, т. е. сумме чисел всех ориентированных, непосредственных и двузвенных транзитных путей, соединяющих данный элемент с остальными элементами структуры. Ранги элемента будем обозначать буквой R с индексом, соответствующим номеру элемента на структурной схеме.

Из теории матричного исчисления известно, что матрица M, определяющая ранги элементов, может быть найдена из матрицы непосредственных путей A из выражения $M = A + A^2$.

Число непосредственных путей R_{in}, связывающих элемент с остальными элементами структуры, определяются по формуле

$$R_{in} = \sum_{i=1}^{n} a_{ij} \qquad (13)$$

или в матричной форме

$$R_{\text{в}} = \left\| \begin{matrix} \sum_{j=1}^{n} a_{1j} \\ \sum_{j=1}^{n} a_{2j} \\ \cdot \\ \cdot \\ \cdot \\ \sum_{j=1}^{n} a_{nj} \end{matrix} \right\|,$$

где n — ранг матрицы А.

Число двузвенных транзитных путей R_{iT2}, связывающих элемент i с остальными элементами структуры, определяется по формуле:

$$R_{iT2} = \sum_{j=1}^{n} \left[a_{ij} \right]^2 \qquad (14)$$

или в матричной форме

$$R_{\text{Т}2} = \left\| \begin{matrix} \sum_{j=1}^{n} [a_{1j}]^2 \\ \sum_{j=1}^{n} [a_{2j}]^2 \\ \cdot \\ \cdot \\ \cdot \\ \sum_{j=1}^{n} [a_{nj}]^2 \end{matrix} \right\|,$$

где $[a_{ij}]^2$ элемент, принадлежащий i-й строке и j-му столбцу матрицы A^2.

На основании выражений (13) и (14) имеем:

$$R_i = R_{in} + R_{iT2} = \sum_{j=1}^{n} \left(a_{ij} + \left[a_{ij} \right]^2 \right)$$

или в матричной форме

$$R = R_n - R_{T2} = \left\| \begin{matrix} \sum_{j=1}^{n} (a_{1j} + [a_{1j}]^2) \\ \sum_{j=1}^{n} (a_{2j} + [a_{2j}]^2) \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \sum_{j=1}^{n} (a_{nj} + [a_{nj}]^2) \end{matrix} \right\|.$$

В тех случаях, когда различные пути между вершинами графа, отображающего рассматриваемую структуру, неравноценны, соотствующим звеньям приписываются некоторые коэффициенты (веса), учитывающие их характеристики по выбранному показателю. Ранг элементов в этом случае определяется аналогично рассмотренным выше приемам, но, в отличие от них, ранги отдельных (или всех) элементов структуры не выражаются целыми числами.

Приведенное выше определение понятия ранга элемента в ряде случаев оказывается недостаточным, так как оно не оценивает существования транзитных путей, содержащих более, чем два звена. В подобных случаях, наиболее часто встречающихся при оценке качества сложных структур, вместо этого понятия будем применять понятие — ***полный ранг элемента***, определяемый как сумма всех ориентированных непосредственных и транзитных путей, соединяющий данных элемент с остальными элементами структуры.

Вычисление полного ранга элемента производится аналогично описанному выше с использованием машинных методов. Для определения полного ранга элементов последовательно находятся матрицы *A*, A^2_1 и т. д. до тех пор, пока все элементы одной из этих матриц не станут равными 0. Полный ранг элемента определим по фррмуле:

$$R_{ni} = \sum_{k=1}^{m} \sum_{j=1}^{n} \left[a_{ij} \right]^k$$

или в матричной форме

$$R_n = \left\| \begin{matrix} \sum_{k=1}^{m} \sum_{j=1}^{n} [a_{1j}]^k \\ \sum_{k=1}^{m} \sum_{j=1}^{n} [a_{2j}]^k \\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot\ \cdot \\ \sum_{k=1}^{m} \sum_{j=1}^{n} [a_{nj}]^k \end{matrix} \right\|,$$

где $[a_{ij}]^k$ — элемент, принадлежащий i-й строке и j-му столбцу матрицы $A^k = ([a_{ij}]^1 = a_{ij})$.

При оценке качества структуры по рассматриваемому показателю следует иметь в виду, что наличие одного или нескольких элементов с высокими рангами заставляет предъявлять к этим элементам повышенные требования по безотказности, поэтому при разработке структур надо стремиться к тому, чтобы все элементы структуры имели близкие друг к другу и относительно невысокие ранги.

Оценивая соотношения элементов в сисиеме рангом, можно определить не только их <u>структурную значимость</u>, но и <u>функциональную</u>. Для этого необходимо знать описание характера последствий изменения параметров того или иного элемента системы. Здесь достаточно иметь лишь качественное описание, примерно в такой форме: «изменение параметра первого элемента вызывает изменение параметров второго и третьего элементов ...» и т. д.

На основании словесного описания можно построить граф (рис. 13), отражающий взаимное влияние элементов в системе, и матрицу отношений.

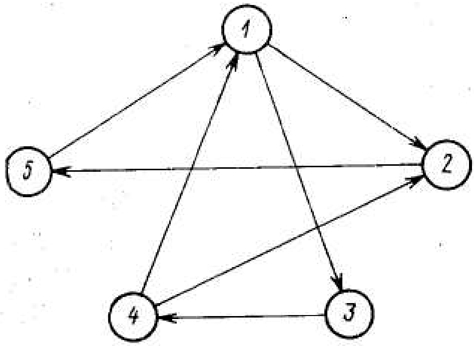


Рис. 13. Граф взаимного влияния элементов системы

В случае, если степень влияния точно количественно неизвестна, то на месте i-го элемента матрицы отношений ставится единица. Это означает лишь только то, что i-й элемент влияет на j-й. В том случае, когда степень влияния можно характеризовать некоторым числом или набором чисел, то в соответствующей клетке матрицы доминирования ставится этот весовой коэффициент. Нуль, как и прежде, означает отсутствие влияния элементов друг на друга. Порядок вычисления ранга, определяющего функциональную значимость элемента, остается тем же, что и при вычислении ранга, определяющего структурную значимость элемента.

При вложении конкретного содержания в понятие «отношение» может образоваться несколько групп рангов, определяющих структурную, функциональную и иные значимости элементов. Для определения уровня ранжировки, на котором требуется остановиться, воспользуемся ***информационным подходом к решению поставленной задачи.***

Каждое ранжирование можно рассматривать как опыт с некоторым исходом (векторным) R. Компоненты этого вектора — ранги элементов системы, сумма которых равна единице.

Известно, что множество положительных дробей, в сумме составляющих единицу, можно рассматривать как соответствующие некоторому множеству элементы, которые обнаруживают многообразие. В этом случае ***энтропия Н может служить средней мерой многообразия на одну составляющую***. Условимся, что до ранжирования многообразие по рангам элементов в системе оценивалось H_0, а после ранжирования — H_1, Тогда ***величиной***

$$J_n = H_0 - H_1, \qquad (15)$$

называемой ***количеством информации***, ***можно характеризовать эффект, даваемый ранжированием как действием, ограничивающим многообразие (устраняющим неопределенность задачи).***

Изменение неопределенности задачи под влиянием пришедшего сигнала можно интерпретировать как ***процесс запасания полезной информации,*** если за нулевой уровень принять запас информации при равномерной значимости элементов (что соответствует равновероятностному исходу опыта). Тогда

$$J_n = \log n — H_1, \qquad (16)$$

где n — число элементов системы, а

$$H_1 = -\sum_{i=1}^{n} R_i \log R_i. \qquad (17)$$

Поскольку количество информации J достигает максимума при $R=1/n$, то любая первая ранжировка даст приращение объема полезной информации $J_n > 0$. Пусть имеются другие ранжировки. Для выявления целесообразности использования их результатов необходимо выполнить следующие действия. По формуле (17) подсчитать для всех k ранжировок значения H_k. Затем расположить H_k в порядке их убывания и подсчитать по формуле

$$\Delta J_n = H_k – (H_k - 1) \qquad (18)$$

прирост полезной информации, даваемый каждой последующей ранжировкой относительно предыдущей. Задаваясь определенным уровнем прироста полезной информации ΔJ_n ранжировки, которые не дают прироста информации (возможен случай «дезинформации» при $\Delta J_n < 0$), из дальнейшего рассмотрения следует исключить.

Сравнивать эффект ранжировок следует по группе сравнимых, эквивалентных факторов, определяющих качество функционирования системы. В подобной ситуации возможно получать конечные формулы, оценивающие прирост полезной информации.

Допустим, что в k-й ранжировке ранги неполно отражают значимость элементов, отличаясь от «истинных» на величину $\pm\Delta R_i$, а в $(k + 1)$-й ранжировке определены истинные значения рангов. Тогда для i-х составляющих будем иметь в k-й ранжировке

$$R_i \log_2 R_1 = \frac{R_{0i}}{\ln 2}(1 \pm \delta_i)\ [\ln R_{0i} + \ln(1 + \delta_i)] \approx$$

$$\approx \frac{R_{0i} \ln R_{0i}}{\ln 2} \pm \frac{R_{0i}}{\ln 2}\, \delta_i,$$

где $\delta_i = \Delta R_i / R_{0i}$;

в $(k + 1)$-й ранжировке —

$$R_{0i} \log_2 R_{0i} = (1/\ln 2)\ R_{0i} \ln R_{0i}.$$

Отсюда, согласно (18), будем иметь

$$\Delta J_n \approx 1{,}5 \sum_{i=1}^{n} (\pm \Delta R_i). \qquad (19)$$

Из (19) следует, что последующее ранжирование будет целесообразным в том случае, когда оно усиливает найденные значимости элементов, если при этом происходит более резкое разграничение элементов по их значимости.

Рассмотрим метод определения множества сочленения. Как отмечалось ранее, под множеством сочленения графа понимают такое множество его элементов, удаление которых делает граф несвязным.

Принадлежностью к множеству сочленения можно оценивать значимость как вершин, так и ребер графа. Однако большая сложность и трудоемкость определения принадлежности элемента к множеству сочленения делает предпочтительным ранжирование при определении значимости вершин графа — элементов системы. Для оценки значимости ребер графа, что соответствует оценке значимости связей в системе, не просматривается ничего иного, кроме множества сочленения. Поскольку оценить значимость ребер не представляется, возможно, ничем иным, кроме множества сочленения, то приходится мириться с неизбежными усложнениями.

Метод определения принадлежности элементов графа к множеству сочленения сводится к следующему. Пусть имеется граф, вершины которого соответствуют элементам системы, а ребра — отношениям (связям) между ними. Строится вспомогательный граф отображением множества вершин во множество ребер.

Построение вспомогательного графа осуществляется следующим образом: вершины вспомогательного графа, число которых равно числу ребер исходного графа, размечаются индексами (буквами), соответствующими разметке ребер исходного графа (рис. 14); если в исходном графе ребра были смежными, т. е. имели общую вершину, то во вспомогательном графе соответствующие вершины соединяются ребром. Например, ребра y_1, y_2, y_3 исходного графа смежны (рис.14, *а*).

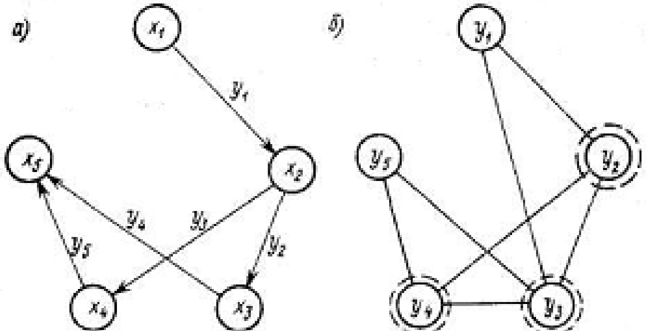


Рис. 14. Граф: *а* — основной; *б* — вспомогательный

Следовательно, во вспомогательном графе вершины y_1, y_2, y_3 следует соединить ребрами (рис. 14, *б)*. Для определения принадлежности элемента к множеству сочленения следует придерживаться следующего порядка.

1. Определить центры и периферийные точки графа.

Рассмотрим некоторые определения. В конечном или бесконечном графе отклонением $\delta(x, y)$ его вершины x от вершины y называется длина кратчайшего пути из x в y. При этом $\delta(x, y)$ удовлетворяет условиям:

а) $\delta (x, x) = 0$;

б) $\delta (x, y) + \delta (x, z) > \delta (x, x)$;

в) для симметричного графа $\delta(x, y) = \delta(y, x)$.

Отклоненностью вершины x называют число:

$$l(x) = \max_{y \in x} \delta(x, y).$$

Если наименьшая из отклоненностей является конечным числом, то вершина, в которой этот минимум достигается, называется ***центром графа***, вершина с наибольшей отклоненностью — ***периферийной точкой графа***. Исключаем из рассмотрения периферийные точки, так как они не могут быть точками сочленения.

2. Строится полная матрица связей. Последовательным перебором оставшихся вершин определяем, удаление каких именно из них нарушит связность графа.

Как видно, операция нахождения множества сочленения весьма трудоемка. Однако, если рассматривать реальные системы, которые имеют ограниченное число входов и выходов, то размерность задачи значительно уменьшится. Поэтому, определяя связность графа и условия ее нарушения, не имеет смысла рассматривать возможность соединения любых двух вершин, а следует ограничиться лишь входами и выходами. Так, в приведенном примере (рис. 14) строго к множеству сочленения относятся элементы y_2, y_3, y_4. Если же входом и выходом являются элементы x_1 и x_5 соответственно, то ко множеству сочленения достаточно отнести элементы y_3 и y_4.

На рис. 15 представлена структура более сложной системы. Здесь две подсистемы *A* и *B* связаны системой различных связей. Необходимо определить, удаление каких связей разрушит структуру системы.

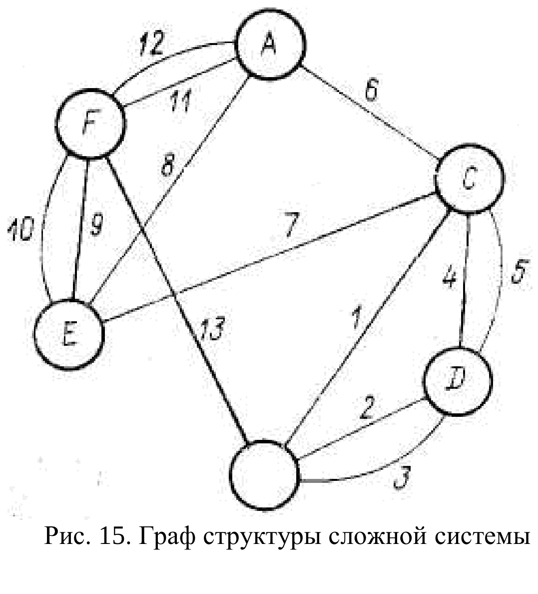


Рис. 15. Граф структуры сложной системы

Придерживаясь описанного порядка и проделав все необходимые операции, определим, что к множеству сочленения во вспомогательном графе (рис. 16) относятся вершины *6, 7, 13.* Следовательно, разрушение связей (ребер основного графа) *8, 7, 13* разорвет и нарушит цепь функциональных связей между подсистемами *А* и *В*.

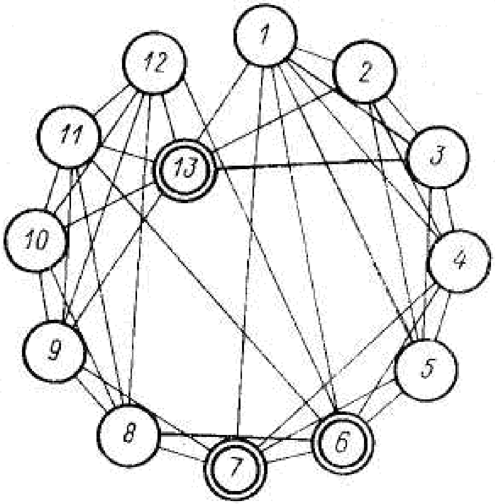


Рис. 16. Вспомогательный граф сложной системы

В заключение следует отметить, что множество сочленения, как структурный параметр, не вводит упорядочение в систему отношений. Принадлежность к нему отвечает лишь на вопрос о возможности разрушения этой системы отношений.

К чему же необходимо стремиться при использовании критерия качества «множество сочленения»? При чисто структурных исследованиях, когда изучаются системы функциональных связей широкого плана, необходимо стремиться к тому, чтобы к множеству сочленения принадлежало как можно большее число элементов системы (ребер графа). Это достигается введением в структуру системы дополнительных связей, установления обходных путей, обычного резервирования. Чем большим будет в этом случае множество сочленения, тем более живуча система, тем труднее нарушить ее связность.

Когда система отношений отражает взаимные влияния элементов друг на друга, то необходимо стремиться, чтобы к множеству сочленения принадлежало как можно меньшее число элементов. Достаточно изъять их, и вся система взаимных влияний нарушится.

Рассмотрение критериев, оценивающих качество структурной схемы системы, убедительно показывает, что применять их следует последовательно, идя от наиболее простого к более сложному, требующему большего количества данных о функционировании системы. Окончательная оценка качества структурной схемы системы, оценка значимости элементов схемы в большинстве случаев явится следствием компромиссного решения применительно к конкретной ситуации.

1.7. Декомпозиция структуры системы

Произвести декомпозицию системы — это значит выделить в ней отдельные сильно связанные подсистемы, т. е. такие подсистемы, все остальные части которых благодаря обратным связям взаимно достижимы. Граф такой системы бисвязен. При декомпозиции выделяются также слабо связанные подсистемы, все составные части которых связаны неориентированным путем. Граф такой подсистемы связен.

Процесс декомпозиции системы можно формализовать сокращением ориентированного графа. Тогда для системы можно рекомендовать следующий алгоритм декомпозиции.

1. Составить матрицу смежности A графа $G(X,U)$.

2. Вычислить матрицу $R_1 = A + E$, где E — единичная матрица; «+» — знак логического сложения; R_1 — матрица первой достижимости, i-я строка которой представляет все ориентированные пути по графу из i-й вершины до всех остальных, если длина пути равна одному ребру.

3. Определить $R_2 = R_1{*}^2$, где знак «*» означает, что при вычислении $R_1 \times R_2$ применяется логическое умножение и суммирование элементов матриц.

Аналогично определяются все матрицы вплоть до $R = R_n = R_1{*}^n$, где R — матрица достижимости графа $G(X,U)$, i-я строка которой представляет все ориентированные пути по графу длиной от одного до n ребер, из i-й вершины ко всем остальным. Матрицы A и R имеют размерность $n \times n$.

При вычислении R не обязательно R_1 возводить в n-ю степень. Если $R_1{*}^k = R_1{*}^{(k-1)}$, то $R = R_1{*}^k$, где $k < n$.

4. Проанализировать матрицу R. Если $R = Q$, где $Q=[q_{ij}]$, такая универсальная матрица, что для всех i и j $q_{ij}= 1$, то граф бисвязен и

декомпозиция системы невозможна, т. е. система состоит из одной сильно связанной подсистемы. Если $R \neq Q$, то необходимо перейти к следующему шагу. В тех случаях, когда известно, что граф связен, следует перейти к шагу 8.

5. Определить матрицу достижимости неориентированного графа $G^0(X,U)$, соответствующего ориентированному графу системы $G(X,U)$. Матрица $R^0=(A+A^{\mathrm{T}}+D)^{*n}$, где знак T означает транспонирование.

6. Определить связные подграфы ориентированного графа $G(X,U)$. Известно, что множество вершин связного подграфа, содержащего вершину i, определено единицами в i -й строке матрицы R^0. Если $R^0=Q$, то граф $G(X,U)$ состоит из одного связного подграфа. В этом случае следует переходить к шагу 8. Если же $R^0 \neq Q$, то надо перейти к следующему шагу.

7. Упорядочить вершины графа $G(X,U)$ (матрицы A) по связным подграфам.

8. Образовать матрицу связности $C=R+R^T$. Здесь сложение обычное, арифметическое.

9. Выделить из матрицы C бисвязные подграфы. Бисвязный подграф, содержащий вершину i, определен двойками в i -й строке матрицы C.

10. Упорядочить A так, чтобы бисвязные подграфы образовали квадратные подматрицы $E_\varphi \subset A$, $\varphi = 1, 2,\ldots,p$.

11. Образовать матрицу $R_+=(A^0{}_+ +A^T{}_+ +E_+)^{*0}$,

где A_+—матрица смежности подграфа с множеством вершин

$V=W/\bigcup_{\varphi=1}^{n} B_\varphi$ и B_φ — подмножество составных частей φ-й сильно связанной подсистемы, и выполнить п. 6 и 7.

12. По упорядоченной матрице A' определить связи (ребра) между подсистемами, которые должны быть разорваны в результате декомпозиции.

Рассмотрим пример реализации описанного алгоритма. Пусть имеется система, структура которой представлена графом (рис. 17).

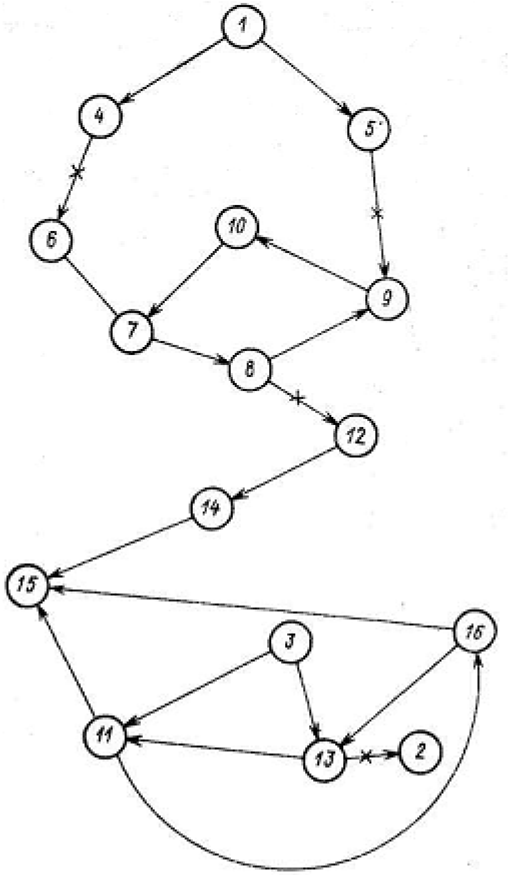


Рис. 17. Граф структурной схемы системы

Составим матрицу смежности *А* этого графа:

$$
A = \begin{array}{||cccccccccccccccc||c}
1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 & 13 & 14 & 15 & 16 & \\
 & & & 1 & 1 & & & & & & & & & & & & 1 \\
 & & & & & & & & & & & & & & & & 2 \\
 & & & & & & & & & & & & 1 & & & & 3 \\
 & & & & & 1 & & & & & & & & & & & 4 \\
 & & & & & & & & 1 & & & & & & & & 5 \\
 & & & & & & 1 & & & & & & & & & & 6 \\
 & & & & & & & 1 & & & & & & & & & 7 \\
 & & & & & & & & 1 & & & 1 & & & & & 8 \\
 & & & & & & & & & 1 & & & & & & & 9 \\
 & & & & & 1 & & & & & & & & & & & 10 \\
 & 1 & & & & & & & & & & & & & 1 & 1 & 11 \\
 & & & & & & & & & & & & & 1 & & & 12 \\
 & 1 & & & & & & & & & 1 & & & & & & 13 \\
 & & & & & & & & & & & & & & 1 & & 14 \\
 & & & & & & & & & & & & & & & 1 & 15 \\
 & & & & & & & & & & & & 1 & & & & 16
\end{array}
\qquad (20)
$$

Далее получим матрицу

$$
R = \begin{array}{c} \\ \left\| \begin{array}{cccccccccccccccc}
1&2&3&4&5&6&7&8&9&10&11&12&13&14&15&16 \\
\hline
1&1&1&1&1&1&1&1&1&1&1&1&1&1&1&1\\
&1&&&&&&&&&&&&&&\\
&1&1&&&&&&&&1&&1&&1&1\\
&1&1&1&&1&1&1&1&1&1&1&1&1&1&1\\
&1&1&&1&1&1&1&1&1&1&1&1&1&1&1\\
&1&1&&&1&1&1&1&1&1&1&1&1&1&1\\
&1&1&&&1&1&1&1&1&1&1&1&1&1&1\\
&1&1&&&1&1&1&1&1&1&1&1&1&1&1\\
&1&1&&&1&1&1&1&1&1&1&1&1&1&1\\
&1&1&&&1&1&1&1&1&1&1&1&1&1&1\\
&1&1&&&&&&&&1&&1&&1&1\\
&1&1&&&&&&&&1&1&1&1&1&1\\
&1&1&&&&&&&&1&&1&&1&1\\
&1&1&&&&&&&&1&&1&1&1&1\\
&1&1&&&&&&&&1&&1&&1&1\\
&1&1&&&&&&&&1&&1&&1&1
\end{array} \right\| \end{array}
\begin{array}{c} \\ 1\\2\\3\\4\\5\\6\\7\\8\\9\\10\\11\\12\\13\\14\\15\\16 \end{array}
$$

(21)

и матрицу

$$
C = \begin{array}{c} \\ \left\| \begin{array}{cccccccccccccccc}
1&2&3&4&5&6&7&8&9&10&11&12&13&14&15&16 \\
\hline
2&&&&&&&&&&&&&&&\\
&2&&&&&&&&&&&&&&\\
&&2&&&&&&&&2&&2&&2&2\\
&&&2&&&&&&&&&&&&\\
&&&&2&&&&&&&&&&&\\
&&&&&2&2&2&2&2&&&&&&\\
&&&&&2&2&2&2&2&&&&&&\\
&&&&&2&2&2&2&2&&&&&&\\
&&&&&2&2&2&2&2&&&&&&\\
&&&&&2&2&2&2&2&&&&&&\\
&&2&&&&&&&&2&&2&&2&2\\
&&&&&&&&&&&2&&&&\\
&&2&&&&&&&&2&&2&&2&2\\
&&&&&&&&&&&&&2&&\\
&&2&&&&&&&&2&&2&&2&2\\
&&2&&&&&&&&2&&2&&2&2
\end{array} \right\| \end{array}
\begin{array}{c} \\ 1\\2\\3\\4\\5\\6\\7\\8\\9\\10\\11\\12\\13\\14\\15\\16 \end{array}
$$

(22)

Следовательно, имеются две сильно связанные подсистемы B_1=6, 7, 8, 9, 10 и B_2=3, 11, 13, 15, 16.

По матрице

$$R_{+}^{0} = \begin{array}{c} \begin{array}{cccccc} 1 & 2 & 4 & 5 & 12 & 14 \end{array} \\ \left\| \begin{array}{cccccc} 1 & & 1 & 1 & & \\ & 1 & & & & \\ 1 & & 1 & 1 & & \\ 1 & & 1 & 1 & & \\ & & & & 1 & 1 \\ & & & & 1 & 1 \end{array} \right\| \end{array} \begin{array}{c} \\ 1 \\ 2 \\ 4 \\ 5 \\ 12 \\ 14 \end{array} \qquad (23)$$

получаем еще три слабо связанные подсистемы S_1={1, 4, 5}, S_2={12, 14} и S_3 ={2}.

Отсюда следует, что в упорядоченной матрице смежности графа системы необходимо разорвать связи, обозначенные на матрице звездочкой (на рисунке крестиками)

$$A' = \begin{array}{c} \\ 1 \\ 4 \\ 5 \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \\ 12 \\ 14 \\ 3 \\ 11 \\ 13 \\ 15 \\ 16 \\ 2 \\ \\ \end{array} \begin{array}{c} \begin{array}{cccccccccccccccc} 1 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 12 & 14 & 3 & 11 & 13 & 15 & 16 & 2 \end{array} \\ \left\| \begin{array}{cccccccccccccccc} & 1 & 1 & & & & & & & & & & & & & \\ & & & 1^* & & & & & & & & & & & & \\ & & & & & & 1^* & & & & & & & & & \\ & & & & 1 & & & & & & & & & & & \\ & & & & & 1 & & & & & & & & & & \\ & & & & & & 1 & & 1^* & & & & & & & \\ & & & & & & & 1 & & & & & & & & \\ & & & 1 & & & & & & & & & & & & \\ & & & & & & & & & 1 & & & & & & \\ & & & & & & & & & & & & & 1^* & & 1 \\ & & & & & & & & & & & & 1 & & & \\ & & & & & & & & & & 1 & & & 1 & 1 & \\ & & & & & & & & & & & 1 & & & & 1^* \\ & & & & & & & & & & & & & & & \\ & & & & & & & & & & & & 1 & & & \\ & & & & & & & & & & & & & & & \end{array} \right\| \\ \underbrace{\quad\quad}_{S_1} \quad \underbrace{\quad\quad\quad\quad\quad}_{B_1} \quad \underbrace{\quad\quad}_{S_2} \quad \underbrace{\quad\quad\quad\quad\quad}_{B_2} \quad \underbrace{\quad}_{S_3} \end{array} \begin{array}{l} \\ \left.\begin{array}{c} \\ \\ \\ \end{array}\right\} S_1 \\ \left.\begin{array}{c} \\ \\ \\ \\ \\ \end{array}\right\} B_1 \\ \left.\begin{array}{c} \\ \\ \end{array}\right\} S_2 \\ \left.\begin{array}{c} \\ \\ \\ \\ \end{array}\right\} B_2 \\ \left.\begin{array}{c} \\ \end{array}\right\} S_3 \\ \\ \end{array}$$

После декомпозиции матрицу A' можно разрушить, сохранив лишь подматрицы смежности D_s и матрицу смежности сокращенного графа M, которые естественным образом получаются из A'.

Замкнутые контуры в подсистемах B_1 и B_2 устраняются с использованием матриц смежности

$$D_{B_1} = \begin{array}{c} \begin{array}{ccccc} 6 & 7 & 8 & 9 & 10 \end{array} \\ \left\| \begin{array}{ccccc} & 1 & & & \\ & & 1 & & \\ & & & 1 & \\ & & & & 1 \\ 1 & & & & \end{array} \right\| \end{array} \begin{array}{c} \\ 6 \\ 7 \\ 8 \\ 9 \\ 10 \end{array} \text{ и } D_{B_2} = $$

$$= \begin{array}{c} \begin{array}{ccccc} 3 & 11 & 13 & 15 & 16 \end{array} \\ \left\| \begin{array}{ccccc} & & 1 & & \\ 1 & & & 1 & 1 \\ & 1 & & & \\ & & & & 1 \\ & & 1 & & \end{array} \right\| \end{array} \begin{array}{c} \\ 3 \\ 11 \\ 13. \\ 15 \\ 16 \end{array}$$

После декомпозиции и устранения замкнутых контуров можно представить систему сокращенным ориентированным графом $Г(\Delta, \Omega)$, где Δ — множество вершин (подсистем); Ω — множество ориентированных ребер (связей между подсистемами). Сокращенный граф является удобной моделью для решения ***динамических, информационных и диагностических систем*** при проектировании систем. Он обладает всеми основными топологическими свойствами исходной системы, поскольку система преобразовывалась таким образом, что топологическое пространство, представленное ориентированным графом *G (X, U),* непрерывно отображалось в топологическое пространство, представленное графом $Г(\Delta, \Omega)$.

Так, для рассмотренного примера матрица смежности сокращенного графа имеет вид:

$$M = \begin{array}{c|ccccc|} & B_1 & B_2 & S_1 & S_2 & S_3 \\ B_1 & & & & 1 & \\ B_2 & & & & & 1 \\ S_1 & 1 & & & & \\ S_2 & & 1 & & & \\ S_3 & & & & & \end{array}, \qquad (24)$$

а сам граф $\Gamma(\Delta, \Omega)$ представлен на рис. 17.

Во всех матрицах рассмотренного примера, за исключением (22), в свободных клетках подразумеваются нули; в матрице (22) в свободных клетках единицы.

Изложенный метод декомпозиции основан только на анализе топологических свойств модели системы, т. е. на анализе свойств графа системы. Функциональная сторона анализа в данном случае не рассматривается, но предполагается, что она должна явиться логическим продолжением топологического анализа. Следует отметить, что прием декомпозиции структурной схемы системы дает возможность осуществить не только преобразование графа системы к виду, удобному для решения последующих задач (появляется возможность существенно уменьшить размерность графа), но и одновременно появляется возможность до некоторой степени косвенно оценить качество структурной схемы. Так, рассмотренный алгоритм дает возможность указать те связи, разрыв которых приводит к распаду системы как единого целого на отдельные подсистемы. Тогда, если смотреть на решение задачи с использованием указанного алгоритма с позиций ***надежности и живучести исследуемой системы***, то становится очевидным, что они позволяют (в первом приближении, поскольку учитываются только топологические свойства) определить наиболее жизненно важные элементы связи системы.

В заключение следует отметить, что структурный анализ является необходимым этапом общего исследования структур систем. Требуя для своего проведения минимум исходной информации — структурную схему — его результаты позволяют в принципе сформулировать рекомендации, обеспечивающие функциональную разгрузку отдельных элементов, усиление слабых звеньев структуры введением структурной и функциональной избыточности и позволяющие, наконец, целесообразно распределить усилия для обеспечения и достижения заданного качества функционирования системы.

2. Анализ структур средствами блочных групп и модуль-графами

Рассмотрим методы анализа структур средствами блочных групп и модуль-графами на примерах структур электрическим цепей.
Анализ пассивных и активных структур цепей можно проводить классическим методом с помощью определителей и матриц, с помощью графов сигналов или сетевых графов (метод деревьев) или с использованием унистора. В данном разделе анализ пассивных активных структур цепей будет производиться методом, который использует средства блочных групп и модуль-графов. Топологические методы анализа, а именно метод графов сигналов и метод деревьев, значительно облегчают анализ электрических цепей. Однако их существенный недостаток заключается в невозможности определения алгоритма расчетов чисто алгебраическим путем. Кроме того, к недостаткам этих методов следует отнести многозначность способов решения данной задачи; особенно при использовании графов сигналов. В зависимости от выбора базисных узлов конечный результат в случае метода графов сигналов может быть представлен в совершенно различных формах. Поэтому возникают большие трудности для программирования на ЭЦВМ.

Достоинство метода блочных групп и модуль-графов состоит в возможности получения результата вычислений в стандартной форме вне зависимости от степени сложности цепи. Кроме того, этот метод по сравнению с методом графов сигналов обладает тем преимуществом, что нет необходимости изображать граф сигналов и можно для расчетов использовать непосредственно электрическую схему. Любую электрическую цепь можно теоретически рассматривать как активную цепь, содержащую зависимые источники. Это означает, что, кроме независимых источников, далее будут рассмотрены также такие источники, э. д. с. в которых зависит от тока или напряжения в другой ветви цепи.

2.1. Анализ структур, представленных пассивными цепями

2.1.1. Анализ пассивного четырехполюсника

Рассмотри структуру цепи в виде *четырехполюсника* (рис. 1).

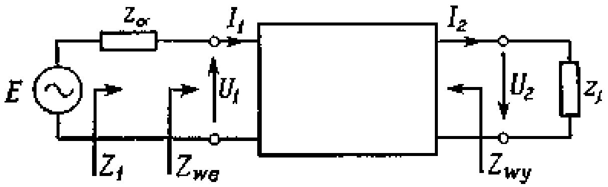


Рис. 1. Структура цепи в виде четырехполюсника.

Для него можно написать следующие характеристические функции:

$$Z_1 = \frac{E}{I_1}, \quad Z_{вх} = \frac{U_1}{I_1}, \quad Z_{вых} = \left.\frac{U_2}{I_2}\right| z_\beta = \infty,$$

$$K_u = \frac{U_2}{E}, \quad K_i = \frac{I_2}{I_1}, \quad \Gamma_s = \ln\left[\frac{E}{2U_2}\sqrt{\frac{Z_\beta}{Z_\alpha}}\right], \qquad (1)$$

где K_u и K_i называются *передаточными функциями* (коэффициентами передачи) *напряжения* и *тока;* Z_1, $Z_{вх}$ и $Z_{вых}$ — *первичным, входным и выходным импедансами;* Γ_s — *рабочим затуханием.*

Средствами алгебры блочных групп и модуль-графами можно создать общие алгоритмы расчета этих характеристических функций независимо от степени сложности рассматриваемой структуры цепи. Используя введенные ранее понятия для блочных групп, можно написать следующие формулы:

$$K_u = \frac{\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial \alpha}, \frac{\partial A}{\partial \beta}\right)}{\underset{Z}{\det} A} Z_\beta, \qquad K_i = \frac{\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial \alpha}, \frac{\partial A}{\partial \beta}\right)}{\underset{Z}{\det}\frac{\partial A}{\partial \alpha}},$$

$$\Gamma_s = \ln\left[\frac{\underset{Z}{\det} A}{2\,\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial \alpha}, \frac{\partial A}{\partial \beta}\right)} \frac{1}{\sqrt{Z_\alpha Z_\beta}}\right], \qquad Z_1 = \frac{\underset{Z}{\det} A}{\underset{Z}{\det}\frac{\partial A}{\partial \alpha}}, \qquad (2)$$

$$Z_{вх} = \frac{\underset{Z}{\det}\frac{\delta A}{\delta \alpha}}{\underset{Z}{\det}\frac{\partial A}{\partial \alpha}}, \qquad Z_{вых} = \frac{\underset{Z}{\det}\frac{\delta A}{\delta \beta}}{\underset{Z}{\det}\frac{\partial A}{\partial \beta}},$$

где A — блочная группа, для которой граф четырехполюсника служит обратным изображением, Z — множество импедансов четырехполюсника.

Докажем формулы для передаточной функции тока K_i и первичного импеданса Z_1 четырехполюсника. Доказательство справедливости формул для остальных функций четырехполюсника проводится аналогично.

Передаточную функцию тока K_i для четырехполюсника можно представить в виде

$$K_i = \frac{I_2}{I_1} = \frac{\sum\limits_{\nu} I_{2\nu} - \sum\limits_{\mu} I_{2\mu}}{I_1}, \qquad (3)$$

где $I_{2\nu}$ — токи контуров четырехполюсника, содержащих одинаково ориентированные ребра α=1 и β=2; $I_{2\nu}$ — токи контуров четырехполюсника, содержащих противоположно ориентированные ребра α=1 и β=2.

Так как столбцы блочной группы A определяют все дополнения деревьев цепи, то столбцы блочных групп $\partial A/\partial\alpha$ и $\partial A/\partial\beta$ определяют те ветви цепи, исключение которых преобразует ее в один контур. Поэтому одинаковые столбцы блочных групп $\partial A/\partial\alpha$ и $\partial A/\partial\beta$ определяют все ветви, исключение которых приводит к образованию контуров, содержащих ребра α и β, т. е. контуров с токами $I_{2\nu}$ и $I_{2\mu}$. Кроме того, на основании уравнений контурных токов имеем

$$\sum_{i=1}^{n} z_{ij} I_j = \begin{cases} E, \ j = 1 = \alpha \\ 0, \ j > 1 \end{cases} \qquad j = 1, 2, \ldots, n. \qquad (4)$$

Отсюда следует. Что

$$K_i = \frac{I_2}{I_1} = \frac{\Delta_{12}}{\Delta_{11}}, \qquad (5)$$

где Δ_{11} и Δ_{12} — соответствующие миноры определителя уравнений (4).

Минор Δ_{11} представляет собой сумму всех величин дополнений деревьев цепи с исключенной ветвью $\alpha = 1$. Поэтому

$$\Delta_{11} = \det_z \frac{\partial A}{\partial \alpha} \qquad (6)$$

где A — блочная группа, обратным изображением которой служит данная цепь (с короткозамкнутым источником).

Минор Δ_{11} есть линейная комбинация с коэффициентами +1 и —1 величин таких дополнений деревьев цепи с исключенной ветвью α или β, при удалении которых $K_i \neq 0$. Поэтому

$$\Delta_{12} = \operatorname*{Sim}_{Z}\left(\frac{\partial A}{\partial \alpha}, \frac{\partial A}{\partial \beta}\right), \qquad (7)$$

где Z — множество импедансов цепи.
Справедливо выражение

$$K_i = \frac{\operatorname*{Sim}_{Z}\left(\frac{\partial A}{\partial \alpha}, \frac{\partial A}{\partial \beta}\right)}{\det\limits_{z} \frac{\partial A}{\partial \alpha}} \quad (\alpha = 1, \beta = 2),$$

Для доказательства формулы для импеданса Z_1 заметим, что

$$Z_i = \frac{\Delta}{\Delta_{11}},$$

где Δ — определитель системы (4), а Δ_{11} — минор, равный сумме величин дополнений деревьев цепи с исключенным ребром α=1.
Таким образом, $\Delta = \det\limits_{Z} A$ и $\Delta_{11} = \det\limits_{Z} \partial A/\partial\alpha$, т. е.

$$Z_1 = \frac{\det\limits_{Z} A}{\det\limits_{Z} \frac{\partial A}{\partial \alpha}}.$$

Справедливость формул для K_u, Γ_s, $Z_{вх}$, $Z_{вых}$ доказывается аналогично.
Применение этих формул поясним на следующих примерах.
Пример 1. Определить передаточные функции K_u и K_i, а также импедансы Z_1, $Z_{вх}$, $Z_{вых}$ мостового четырехполюсника, изображенного на рис. 2, *а*.

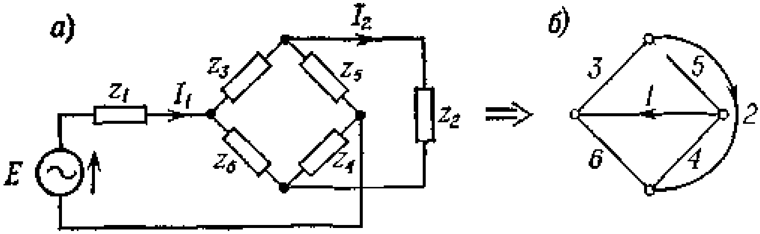


Рис. 2.

Решение. Блочная группа *A,* для которой известно обратное изображение (рис. 2, б), определим на основании рассмотренной ранее теоремы о простых однострочных сомножителях:

$$A = [1\ 3\ 5]\ [1\ 4\ 6]\ [2\ 4\ 5].$$

Умножение первичных блочных групп производим следующим образом:

$$\begin{array}{r} 1\,3\,5 \\ \times\quad 1\,4\,6 \\ 2\,4\,5 \\ \hline \end{array}$$

$$A = \left[\begin{array}{c|cccccccccccc|ccc} 1 & 1 & 1 & 1 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 5 & 5 & 5 & 5 \\ 4 & 6 & 6 & 6 & 1 & 1 & 1 & 4 & 4 & 6 & 6 & 6 & 1 & 4 & 6 & 6 \\ 2 & 2 & 4 & 5 & 2 & 4 & 5 & 2 & 5 & 2 & 4 & 5 & 2 & 2 & 2 & 4 \end{array}\right]$$

При расчете этого произведения вначале опускаем столбцы, в которых повторяется какой-либо элемент, а затем в полученной блочной группе вычеркиваем четное число одинаковых столбцов (в данном случае имеются два одинаковых столбца, содержащих элементы 1, 4, 5). Далее рассчитываем алгебраические производные *∂A/∂1* и *∂A/∂2.* Получаем

$$\frac{\partial A}{\partial 1} = \begin{bmatrix} 4 & 6 & 6 & \boxed{\begin{matrix}6\\5\end{matrix}} & \begin{matrix}3\\2\end{matrix} & \boxed{\begin{matrix}3\\4\end{matrix}} & \begin{matrix}3\\5\end{matrix} & \begin{matrix}5\\2\end{matrix} \\[-1.2em] 2 & 2 & 4 & & & & & \end{bmatrix}, \quad \frac{\partial A}{\partial 2} = \begin{bmatrix} 1 & 1 & 3 & \boxed{\begin{matrix}3\\4\end{matrix}} & \begin{matrix}3\\6\end{matrix} & \begin{matrix}5\\1\end{matrix} & \begin{matrix}5\\4\end{matrix} & \boxed{\begin{matrix}5\\6\end{matrix}} \\[-1.2em] 4 & 6 & 1 & & & & & \end{bmatrix}.$$

Одинаковые столбцы блочных групп *∂A/∂1* и *∂A/∂2* обведены рамкой. Исходя из ориентации ребер α = 1 и β = 2 в графе (рис. 2, б), определяем знаки функции совпадения. В результате находим

$$\operatorname*{Sim}_{z}\left(\frac{\partial A}{\partial 1}, \frac{\partial A}{\partial 2}\right) = z_5 z_6 - z_3 z_4.$$

Определим также обратные производные δ*A*/δ1 и δ*A*/δ2

$$\frac{\delta A}{\delta 1} = \begin{bmatrix} 3 & 3 & 3 & 3 & 3 & 5 & 5 & 5 \\ 4 & 4 & 6 & 6 & 6 & 4 & 6 & 6 \\ 2 & 5 & 2 & 4 & 5 & 2 & 2 & 4 \end{bmatrix}, \quad \frac{\delta A}{\delta 2} = \begin{bmatrix} 1 & 1 & 3 & 3 & 3 & 3 & 3 & 5 \\ 6 & 6 & 1 & 1 & 4 & 6 & 6 & 6 \\ 4 & 5 & 4 & 5 & 5 & 4 & 5 & 4 \end{bmatrix}.$$

На основании формул (2) теперь можно написать все искомые величины

$$K_u = \frac{(z_5z_6 - z_3z_4)\, z_2}{z_1z_4z_2 + z_1z_6z_2 + z_1z_6z_4 + z_1z_6z_5 + z_3z_1z_2 + z_3z_1z_4 + z_3z_1z_5 + z_3z_4z_2 + \ldots + z_5z_6z_4};$$

$$K_i = \frac{z_5z_6 - z_3z_4}{z_4z_2 + z_6z_2 + z_6z_4 + z_6z_5 + z_3z_2 + z_3z_4 + z_3z_5 + z_5z_2};$$

$$Z_1 = \frac{z_1z_4z_2 + z_1z_6z_2 + z_1z_6z_4 + z_1z_6z_5 + z_3z_1z_2 + z_3z_1z_4 + z_3z_1z_5 + \ldots + z_5z_6z_4}{z_4z_2 + z_6z_2 + z_6z_4 + z_6z_5 + z_3z_2 + z_3z_4 + z_3z_5 + z_5z_2};$$

$$Z_{\text{вх}} = \frac{z_3z_4z_2 + z_3z_4z_5 + z_3z_6z_2 + z_3z_6z_4 + z_3z_6z_5 + z_5z_4z_2 + z_5z_6z_2 + z_5z_6z_4}{z_4z_2 + z_6z_2 + z_6z_5 + z_3z_2 + z_3z_4 + z_3z_5 + z_5z_2};$$

$$Z_{\text{вых}} = \frac{z_1z_6z_4 + z_1z_6z_5 + z_3z_1z_4 + z_3z_1z_5 + z_3z_4z_5 + z_3z_6z_4 + z_3z_6z_5 + z_5z_6z_4}{z_1z_2 + z_1z_6 + z_3z_1 + z_3z_4 + z_2z_6 + z_5z_1 + z_5z_4 + z_5z_6}.$$

Из проведенных расчетов можно сформулировать некоторые замечания. Прежде всего данный метод позволяет обойтись без записи уравнений Кирхгофа и их решения. Все операции довольно просты и производятся над индексами ребер графа, благодаря чему в ходе расчета получается сжатая форма записи всех промежуточных формул. Дополнительное преимущество метода заключается в принципиальной возможности записи всех интересующих зависимостей по известной рассчитанной блочной группе A и нескольким ее алгебраическим и обратным производным.

Пример 2. Определить передаточную функцию напряжения K_u и тока K_i для четырехполюсника (перекрытый Т-образный мост, рис. 3, а).

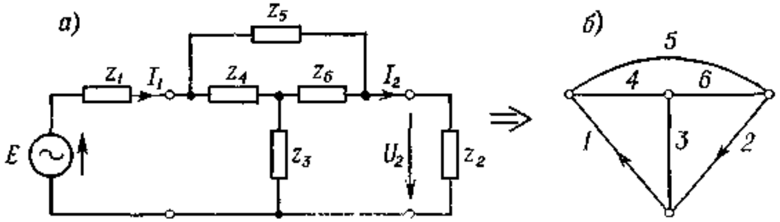


Рис. 3.

Решение. Блочную группу с заданным обратным изображением (рис. 3, б) рассчитаем на основании рассмотренной ранее теоремы о простых однострочных сомножителях::

$$A = [134]\ [236]\ [456].$$

В результате имеем

$$\begin{array}{r} 1\ 3\ 4 \\ \times 2\ 3\ 6 \\ 4\ 5\ 6 \\ \hline \end{array}$$

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 3 & 3 & 3 & 3 & 3 & 4 & 4 & 4 & 4 & 4 \\ 2 & 2 & 2 & 3 & 3 & 3 & 6 & 6 & 2 & 2 & 2 & 6 & 6 & 2 & 2 & 3 & 3 & 6 \\ 4 & 5 & 6 & 4 & 5 & 6 & 4 & 5 & 4 & 5 & 6 & 4 & 5 & 5 & 6 & 5 & 6 & 5 \end{bmatrix},$$

а также

$$\frac{\partial A}{\partial 1}=\left[\begin{array}{ccc|c|c|c|c|c} 2&2&2&3&3&3&6&6\\ 4&5&6&4&5&6&4&5\end{array}\right],\quad \frac{\partial A}{\partial 2}=\left[\begin{array}{ccc|c|c|c|c|c} 1&1&1&3&3&3&4&4\\ 4&5&6&4&5&6&5&6\end{array}\right].$$

Функция совпадения равна

$$\operatorname*{Sim}_{z}\left(\frac{\partial A}{\partial 1},\frac{\partial A}{\partial 2}\right)=z_3z_4+z_3z_5+z_3z_{65}+z_6z_4.$$

Отсюда

$$K_u=\frac{(z_3z_4+z_3z_5+z_3z_6+z_6z_4)\,z_2}{z_1z_2z_4+z_1z_3z_5+z_1z_2z_6+z_1z_3z_4+z_1z_3z_5+z_1z_3z_6+z_1z_6z_4+z_1z_6z_5+ +z_3z_2z_4+\ldots+z_4z_6z_5};$$

$$K_i=\frac{z_3z_4+z_3z_5+z_3z_6+z_6z_4}{z_2z_4+z_2z_5+z_2z_6+z_3z_4+z_3z_5+z_3z_6+z_6z_4+z_6z_5}.$$

Пример 3. Определить передаточную функцию *U/E* лестничного четырехполюсника (рис. 4, *а).*

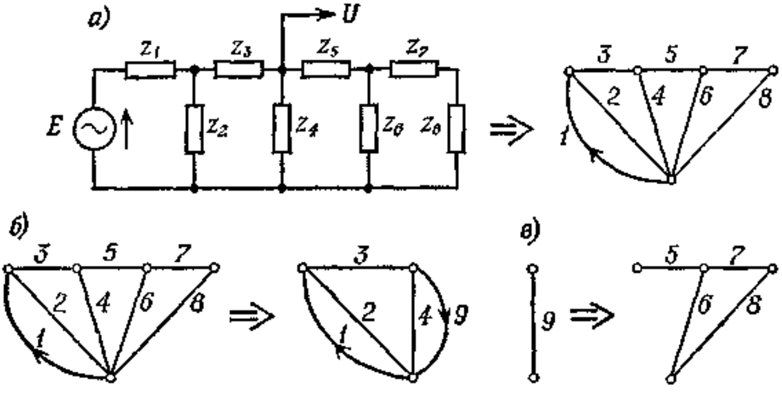


Рис. 4.

Решение. Упростим расчеты, заменив часть графа, расположенную справа от ветви с напряжением *U,* одним новым ребром (рис. 4, *в).* Рассчитаем блочную группу упрощенного графа

$$\begin{array}{r} 1\ 2\\ \times 2\ 3\ 4\\ 4\ 9\\ \hline \end{array},$$

$$A=\begin{bmatrix}1&1&1&1&1&2&2&2\\ 2&2&3&3&4&3&3&4\\ 4&9&4&9&9&4&9&9\end{bmatrix}$$

Откуда

$$\frac{\partial A}{\partial 1}=\left[\begin{bmatrix}\boxed{\begin{matrix}2\\4\end{matrix}} & \begin{matrix}2&3&3&4\\9&4&9&9\end{matrix}\end{bmatrix}\right],\quad \frac{\partial A}{\partial 9}=\left[\begin{bmatrix}\begin{matrix}1&1&1&2\\2&3&4&3\end{matrix} & \boxed{\begin{matrix}2\\4\end{matrix}}\end{bmatrix}\right]$$

и

$$\operatorname*{Sim}_{z}\left(\frac{\partial A}{\partial 1},\frac{\partial A}{\partial 9}\right)=z_2z_4.$$

Выражение искомой передаточной функции запишется в виде

$$\frac{U}{E}=\frac{z_2z_4}{z_1z_2z_4+z_1z_2z_9+z_1z_3z_4+z_1z_3z_9+z_1z_4z_9+z_2z_3z_4+z_2z_3z_9+z_2z_4z_9}.$$

Импеданс z_9 (рис. 4, *в)* рассчитаем по формуле

$$z_9=\frac{\det\limits_{z} A_1}{\det\limits_{z}\dfrac{\partial A_1}{\partial 5}},$$

где

$$A=[5\ 6]\ [6\ 7\ 8]=\begin{bmatrix}5\,5\,5\,6\,6\\6\,7\,8\,7\,8\end{bmatrix},\qquad \frac{\partial A_1}{\partial 5}=[6\ 7\ 8],$$

Поэтому

$$z_9=\frac{z_5z_6+z_5z_7+z_5z_8+z_6z_7+z_6z_8}{z_6+z_7+z_8}.$$

Пример 4. Определим рабочее затухание пассивного четырехполюсника (рис. 5, *а)*.

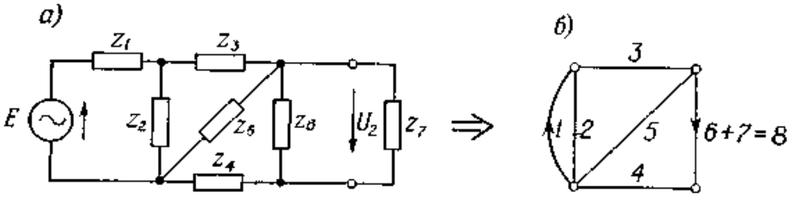


Рис. 5.

Решение. Имеем (рис. 5, б)

$$A=[1\ 2]\ [2\ 3\ 5]\ [4\ 5\ 8],$$

т. е.

$$\begin{array}{r} 1\ 2 \\ \times 2\ 3\ 5 \\ 4\ 5\ 8 \\ \hline \end{array}$$

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 2 \\ 2 & 2 & 2 & 3 & 3 & 3 & 5 & 5 & 3 & 3 & 3 & 5 & 5 \\ 4 & 5 & 8 & 4 & 5 & 8 & 4 & 8 & 4 & 5 & 8 & 4 & 8 \end{bmatrix},$$

а также

$$\frac{\partial A}{\partial 1} = \begin{bmatrix} 2 & \boxed{2} & 2 & 3 & 3 & 3 & 5 & 5 \\ 4 & \boxed{5} & 8 & 4 & 5 & 8 & 4 & 8 \end{bmatrix}, \qquad \frac{\partial A}{\partial 8} = \begin{bmatrix} 1 & 1 & 1 & 2 & \boxed{2} \\ 2 & 3 & 5 & 3 & \boxed{5} \end{bmatrix},$$

$$\operatorname*{Sim}_{z}\left(\frac{\partial A}{\partial 1}, \ \frac{\partial A}{\partial 8}\right) = z_2 z_5;$$

окончательно

$$\Gamma_s = \ln \frac{z_1 z_2 z_4 + z_1 z_2 z_5 + z_1 z_2 z_8 + z_1 z_3 z_4 + z_1 z_3 z_5 + z_1 z_3 z_8 + \ldots + z_2 z_5 z_8}{2 z_2 z_5 \sqrt{z_1 z_7}},$$

где

$$z_8 = \frac{z_6 z_7}{z_6 + z_7}.$$

Пример 5. Определить передаточную функцию K_i, пассивного четырехполюсника (рис. 6, а) с импедансом нагрузки z_9.

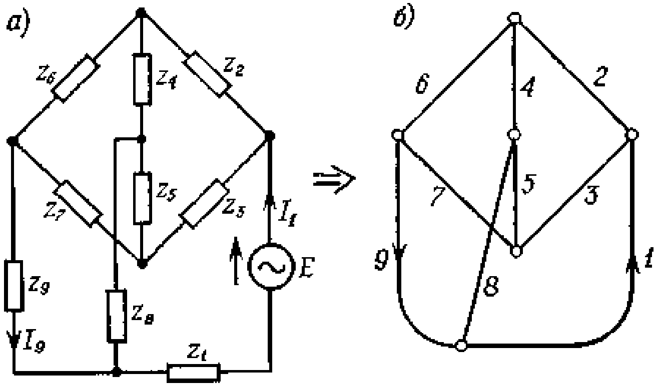


Рис. 6.

Решение. Так как передаточная функция тока определяется по формуле

$$K_i = \frac{\underset{Z}{\mathrm{Sim}}\left(\dfrac{\partial A}{\partial \alpha}, \dfrac{\partial A}{\partial \beta}\right)}{\underset{z}{\det}\dfrac{\partial A}{\partial \alpha}},$$

то, приняв $\alpha = 1$, $\beta = 9$, можно упростить расчеты и рассмотреть граф с исключенным ребром $\alpha = 1$. Блочная группа A_1 такого графа в соответствии с определенным ранее свойством 1 равно производной $\partial A/\partial\alpha$, т. е. $A_1 = \partial A/\partial\alpha$. Это число равно следующему произведению однострочных блочных групп (фиг. 6, *б):*

$$A_1 = [2\;3\;4\;5]\;[4\;5\;6\;7]\;[5\;7\;8\;9].$$

После умножения получим

$$\begin{array}{r}2\;3\;4\;5\\ \times\quad 4\;5\;6\;7\\ 5\;7\;8\;9\end{array}$$

$$A_1 = \left[\begin{array}{cccc|ccc|cccc|ccc|cccc|ccc|cccc|}
2&2&2&2&\cancel{2}&2&2&2&2&2&2&\cancel{2}&2&2&3&3&3&3&\cancel{3}&3&3&3&3&3&3\\
4&4&4&4&\cancel{5}&5&5&6&6&6&6&\cancel{7}&7&7&4&4&4&4&\cancel{5}&5&5&6&6&6&6\\
5&7&8&9&\cancel{7}&8&9&5&7&8&9&\cancel{5}&8&9&5&7&8&9&\cancel{7}&8&9&5&7&8&9
\end{array}\right.$$

$$\left.\begin{array}{|ccc|ccc|cccc|ccc|ccc|ccc|cc}
\cancel{3}&3&3&\cancel{4}&\cancel{4}&\cancel{4}&4&4&4&4&4&4&4&\cancel{5}&\cancel{5}&\cancel{5}&5&5&5&5&5\\
\cancel{7}&7&7&\cancel{5}&\cancel{5}&\cancel{5}&6&6&6&6&7&7&7&\cancel{4}&\cancel{4}&\cancel{4}&6&6&6&7&7\\
\cancel{5}&8&9&\cancel{7}&\cancel{8}&\cancel{9}&5&7&8&9&5&8&9&\cancel{7}&\cancel{8}&\cancel{9}&7&8&9&8&9
\end{array}\right]$$

В этом произведении дважды встречаются столбцы, содержащие числа 2 5 7, 3 7 5, 4 5 8, 5 4 9, которые вычеркиваются. Три столбца этого произведения содержат числа 4 5 7. Согласно определению произведения блочных групп, оставляем из них только один, например 4 7 5. Отыскание одинаковых столбцов можно упростить, поделив столбцы блочной группы на «блоки», состоящие из столбцов, имеющих два первичных идентичных элемента. В нашем примере такие блоки разделены пунктиром. Идентичные столбцы встречаются в разных блоках. После деления на блоки найти идентичные столбцы относительно легко.

Для вычисления производной

$$\frac{\partial A}{\partial \beta} = \frac{\partial A}{\partial 9}$$

заметим, что (рис. 6, *б)*

$$A = A_1[1\;3\;5\;8],$$

поэтому

$$\frac{\partial A}{\partial 9} = \frac{\partial A_1}{\partial 9}\;[1\;3\;5\;8].$$

Так как

$$\frac{\partial A_1}{\partial 9}=\begin{bmatrix} 2\;2\;2\;2\;3\;3\;3\;3\;4\;4\;5\;5 \\ 4\;5\;6\;7\;4\;5\;6\;7\;6\;7\;6\;7 \end{bmatrix},$$

то после элементарных расчетов получим

$$C=\frac{\partial A}{\partial 1}\cap\frac{\partial A}{\partial 9}=\left[\begin{array}{c|c|ccc|c|ccccccccccccc} 3 & 3 & 5 & 5 & 5 & 5 & 5 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 & 8 \\ 4 & 5 & 2 & 2 & 3 & 3 & 4 & 2 & 2 & 2 & 2 & 3 & 3 & 3 & 3 & 4 & 4 & 5 & 5 \\ 7 & 6 & 4 & 6 & 4 & 6 & 6 & 4 & 5 & 6 & 7 & 4 & 5 & 6 & 7 & 6 & 7 & 6 & 7 \end{array}\right].$$

Рассматривая ориентацию ребер *1* и *9* в контурах, оставшихся после исключения из графа ребер, определяемых отдельными столбцами блочной группы *C,* найдем знаки слагаемых функции совпадения. Получим

$$\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial 1},\ \frac{\partial A}{\partial 9}\right)=z_3z_4z_7+z_2z_5z_4+z_5z_2z_6+z_5z_3z_4+z_5z_4z_6+z_8z_2z_4+$$
$$+z_8z_2z_5+z_8z_2z_6+z_8z_2z_7+z_8z_3z_4+z_8z_3z_5+z_8z_3z_6+$$
$$+z_8z_3z_7+z_8z_4z_6+z_8z_4z_7+z_8z_5z_6+z_8z_5z_7=N. \qquad (8)$$

Искомую передаточную функцию тока можно выразить в виде отношения многочленов *N* и *M*

$$K_i = N/M,$$

где *N* определяется выражением (8), а многочлен *M* равен

$$M=z_2z_4z_5+z_2z_4z_7+z_2z_4z_8+z_2z_4z_9+z_2z_5z_8+z_2z_5z_9+$$
$$+z_2z_6z_5+z_2z_6z_7+z_2z_6z_8+z_2z_6z_9+z_2z_7z_8+z_2z_7z_9+z_3z_4z_5+$$
$$+z_3z_4z_7+z_3z_4z_8+z_3z_4z_9+z_3z_5z_8+z_3z_5z_9+z_3z_6z_5+z_3z_6z_7+$$
$$+z_3z_6z_8+z_3z_6z_9+z_3z_7z_8+z_3z_7z_9+z_4z_6z_5+z_4z_6z_7+z_4z_6z_8+$$
$$+z_4z_6z_9+z_4z_7z_5+z_4z_7z_8+z_4z_7z_9+z_5z_6z_7+z_5z_6z_8+z_5z_6z_9+$$
$$+z_5z_7z_8+z_5z_7z_9.$$

Характеристические функции четырехполюсника можно также выразить через проводимости. Тогда формулы передаточных функций напряжения и тока будут иметь вид

$$K_u=\frac{\underset{Y}{\mathrm{Sim}}\left(\frac{\partial A}{\partial\alpha},\ \frac{\partial A}{\partial\beta}\right)}{\underset{Y}{\det}A}\,y_\alpha;\qquad K_i=\frac{\underset{Y}{\mathrm{Sim}}\left(\frac{\partial A}{\partial\alpha},\ \frac{\partial A}{\partial\beta}\right)}{\underset{Y}{\det}\frac{\delta A}{\delta\alpha}}\,y_\beta. \qquad (9)$$

В этом случае геометрическое изображение блочной группы A представляет собой граф цепи. Тогда блочную группу можно определить согласно приведенной ранее теореме о простых однострочных сомножителях. Ребро α представляет собой ветвь источника, ребро β — измерительную ветвь, Y — множество проводимостей цепи.

Знаки слагаемых функции совпадения определяем так же, как и при записи характеристических функций через импедансы для данного обратного геометрического изображения. При этом замыкаем ребра, определяемые отдельными слагаемыми этой функции, и исследуем ориентацию ребер α и β.

Так как имеются *две дуальные формы* записи характеристических функций, то первые из них [формулы (2)] назовем *импедансными,* а вторые [формулы (9)] — *адмитансными.*

Приведем примеры практического применения формул (9).

Пример 6. Определить с помощью адмитансных формул передаточные функции напряжения K_u и тока K_i мостового четырехполюсника (рис. 2).

Решение. Так как в этом случае геометрическое изображение блочной группы представляет собой граф цепи, то на основании теоремы о простых однострочных сомножителях получаем

$$A = [1\ 3\ 6]\ [2\ 3\ 5]\ [2\ 4\ 6],$$

откуда

$$\begin{array}{rc} & 1\,3\,6 \\ \times & 2\,3\,5 \\ & 2\,4\,6 \\ \hline \end{array},$$

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 3 & 3 & 3 & 3 & 3 & 6 & 6 & 6 & 6 & 6 \\ 2 & 2 & 3 & 3 & 3 & 5 & 5 & 5 & 2 & 2 & 5 & 5 & 5 & 2 & 3 & 3 & 5 & 5 \\ 4 & 6 & 2 & 4 & 6 & 2 & 4 & 6 & 4 & 6 & 2 & 4 & 6 & 4 & 2 & 4 & 2 & 4 \end{bmatrix}$$

а также

$$\frac{\partial A}{\partial 1} = \begin{bmatrix} 2 & 2 & 3 & \boxed{\begin{matrix} 3 \\ 4 \end{matrix}} & 3 & 5 & 5 & \boxed{\begin{matrix} 5 \\ 6 \end{matrix}} \\ 4 & 6 & 2 & & 6 & 2 & 4 & \end{bmatrix}; \qquad \frac{\partial A}{\partial 2} = \begin{bmatrix} 1 & 1 & 1 & 1 & \boxed{\begin{matrix} 3 \\ 4 \end{matrix}} & 3 & 6 & \boxed{\begin{matrix} 6 \\ 5 \end{matrix}} \\ 4 & 6 & 3 & 5 & & 5 & 4 & \end{bmatrix}.$$

Далее, рассматривая ориентацию ребер α и β, находим

$$\operatorname*{Sim}_{Y}\left(\frac{\partial A}{\partial 1}, \frac{\partial A}{\partial 2}\right) = y_3 y_4 - y_5 y_6.$$

Искомые передаточные функции имеют вид

$$K_u = \frac{(y_3y_4 - y_5y_6)\, y_1}{y_1y_2y_4 + y_1y_2y_6 + y_1y_3y_2 + y_1y_3y_4 + y_1y_3y_6 + y_1y_5y_2 + \ldots + y_6y_5y_4};$$

$$K_i = \frac{(y_3y_4 - y_5y_6)\, y_2}{y_3y_2y_1 + y_3y_5y_2 + y_3y_5y_4 + y_3y_5y_6 + y_6y_2y_4 + y_6y_3y_4 + y_6y_5y_2 + y_6y_5y_4}.$$

Пример 7. Определить передаточные функции K_u и K_i четырехполюсника (Т-образный мост), изображенного на рис. 7, а, выражая их через проводимости ветвей цепи.

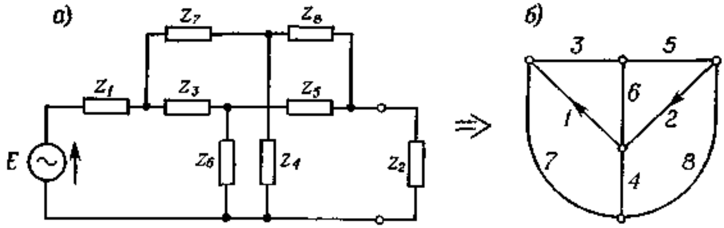


Рис. 7.

Решение. Блочную группу A определим на основании теоремы о простых однострочных сомножителях. Получим (рис. 7, б)

$$A = [3\ 5\ 6]\ [5\ 2\ 8]\ [6\ 1\ 2\ 4]\ [7\ 4\ 8].$$

Отсюда

$$\times\begin{array}{l} 3\ 5\ 6 \\ 5\ 2\ 8 \\ 6\ 1\ 2\ 4 \\ 7\ 4\ 8 \end{array}$$

$$A = \left[\begin{array}{ccccccccccc|cccccccc|ccccccc|ccc|ccccc}
3&3&3&3&3&3&3&3&3&3&3&3&3&3&3&3&3&3&3&3&3&3&3&3&3&3&5&5&5&5&5&5&5&5 \\
5&5&5&5&5&5&5&5&5&5&5&2&2&2&2&2&2&2&2&8&8&8&8&8&8&8&2&2&2&2&2&2&2&2 \\
6&6&6&1&1&1&2&2&2&4&4&6&6&6&1&1&1&4&4&6&6&1&1&2&2&4&6&6&6&1&1&1&4&4 \\
7&4&8&7&4&8&7&4&8&7&8&7&4&8&7&4&8&7&8&7&4&7&4&7&4&7&7&4&8&7&4&8&7&8
\end{array}\right.$$

$$\left.\begin{array}{ccccccc|ccccc|ccc|ccccc|ccccc}
5&5&5&5&5&5&5&6&6&6&6&6&6&6&6&6&6&6&6&6&6&6&6&6&6&6 \\
8&8&8&8&8&8&8&5&5&5&5&5&5&5&5&2&2&2&2&2&8&8&8&8&8 \\
6&6&1&1&2&2&4&1&1&1&2&2&2&4&4&1&1&1&4&4&1&1&2&2&4 \\
7&4&7&4&7&4&7&7&4&8&7&4&8&7&8&7&4&8&7&8&7&4&7&4&7
\end{array}\right]$$

Рассчитаем алгебраические производные *дA/d1* и *дA/д2:*

$$\frac{\partial A}{\partial 1}=\left[\boxed{\begin{matrix}3&3&3\\5&5&5\\7&4&8\end{matrix}}\;\begin{matrix}3&3&3\\2&2&2\\7&4&8\end{matrix}\;\boxed{\begin{matrix}3&3\\8&8\\7&4\end{matrix}}\;\begin{matrix}5&5&5\\2&2&2\\7&4&8\end{matrix}\;\boxed{\begin{matrix}5&5&6&6\\8&8&5&5\\7&4&7&4\end{matrix}}\;\begin{matrix}6&6&6&6\\5&2&2&2\\8&7&4&8\end{matrix}\;\boxed{\begin{matrix}6&6\\8&8\\7&4\end{matrix}}\right],$$

$$\frac{\partial A}{\partial 2}=\left[\boxed{\begin{matrix}3&3&3\\5&5&5\\7&4&8\end{matrix}}\;\begin{matrix}3&3&3&3&3&3&3&3\\6&6&6&1&1&1&4&4\\7&4&8&7&4&8&7&8\end{matrix}\;\boxed{\begin{matrix}3&3\\8&8\\7&4\end{matrix}}\;\begin{matrix}5&5&5&5&5&5&5\\6&6&1&1&1&4&4\\7&4&7&4&8&7&8\end{matrix}\right.$$

$$\left.\boxed{\begin{matrix}5&5&6&6\\8&8&5&5\\7&4&7&4\end{matrix}}\;\begin{matrix}6&6&6&6&6\\1&1&1&4&4\\7&4&8&7&8\end{matrix}\;\boxed{\begin{matrix}6&6\\8&8\\7&4\end{matrix}}\right].$$

Затем, рассматривая ориентацию ребер 1 и 2, находим

$$\operatorname*{Sim}_{Y}\left(\frac{\partial A}{\partial 1},\ \frac{\partial A}{\partial 2}\right)=y_3y_5y_7+y_3y_5y_4+y_3y_5y_8+y_3y_8y_7+y_3y_8y_4+$$
$$+y_5y_8y_7+y_5y_8y_4+y_6y_5y_7+y_6y_5y_4+y_6y_8y_7+y_6y_8y_4. \tag{10}$$

Обратная производная равна

$$\frac{\delta A}{\delta 1}=\left[\begin{matrix}3&3&3&3&3&3&3&3&3&3&3&3&3&3&3&3&3&3&5&5&5&5&5&5&5&5&5\\5&5&5&5&5&5&5&5&2&2&2&2&2&8&8&8&8&8&2&2&2&2&8&8&8&8&8\\6&6&6&2&2&2&4&4&6&6&6&4&4&6&6&2&2&4&6&6&4&4&6&6&2&2&4\\7&4&8&7&4&8&7&8&7&4&8&7&8&7&4&7&4&7&7&4&7&8&7&4&7&4&7\end{matrix}\right.$$

$$\left.\begin{matrix}6&6&6&6&6&6&6&6&6\\5&5&5&5&2&2&8&8&8\\2&2&4&4&4&4&2&2&4\\7&4&7&8&7&8&7&4&7\end{matrix}\right].$$

Поэтому по формуле (9) получим

$$K_u=\frac{N}{M_1}y_1;\quad K_i=\frac{N}{M_2}y_2,$$

где N определяется из выражения (10), а

$$M_1=\det_{Y}A=y_3y_5y_6y_7+y_3y_5y_6y_4+y_3y_5y_6y_8+y_3y_5y_1y_7+$$
$$+y_3y_5y_1y_4+\ldots+y_6y_8y_4y_7,$$

$$M_2=\det_{Y}\frac{\delta A}{\delta 1}=y_3y_5y_6y_7+y_3y_5y_6y_4+y_3y_5y_6y_8+y_3y_5y_2y_7+$$
$$+y_3y_5y_2y_4+\ldots+y_6y_8y_4y_7.$$

Рассмотрим *ненагруженный четырехполюсник* (рис. 8),

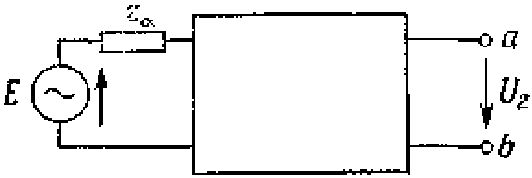


Рис. 8.

Режим холостого хода можно считать предельным случаем работы четырехполюсника, нагруженного импедансом z_β при $z_\beta \to \infty$. Если при этом применить соответствующую расчетную формулу, то решение значительно упрощается. Простые рассуждения приводят к следующей формуле для передаточной функции напряжения ненагруженного четырехполюсника:

$$K_i = \frac{\underset{Z}{\mathrm{Sim}}\left(\dfrac{\partial A}{\partial \alpha} D_\beta, A\right)}{\underset{z}{\det} A}. \qquad (11)$$

В этом выражении D_β — однострочная блочная группа, состоящпя из элементов какого-либо пути, соединяющего выходные зажимы *(a, b)* четырехполюсника и не содержащего элемента α. Знаки слагаемых функции совпадения определяются, как и в предыдущем случае, но с той лишь разницей, что исследуется совпадение ориентации ребра α с направлением стрелки выходного напряжения U_2 четырехполюсника.

Практическое применение формулы (11) иллюстрируют следующие примеры.

Пример 8. Определить K_u мостового четырехполюсника (рис. 9).

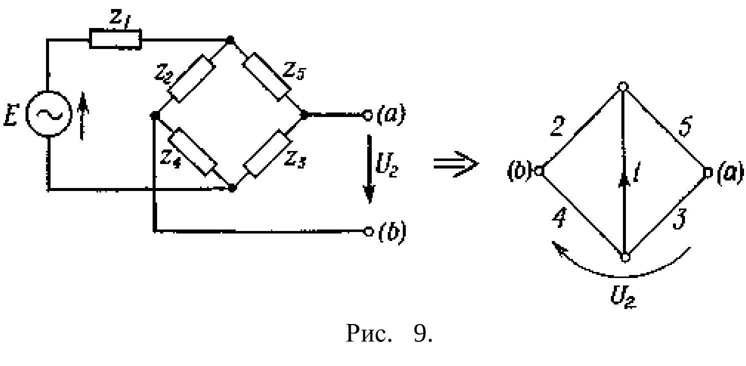


Рис. 9.

Решение. Блочная группа (рис. 9, *б)* равна

$$A = [1\ 2\ 4]\ [1\ 3\ 5],$$

или

$$\frac{\times\begin{matrix}1 & 2 & 4\\ 1 & 3 & 5\end{matrix}}{A = \begin{bmatrix}1 & 1 & 2 & 2 & 2 & 4 & 4 & 4\\ 3 & 5 & 1 & 3 & 5 & 1 & 3 & 5\end{bmatrix}},$$

а алгебраическая производная имеет вид

$$\frac{\partial A}{\partial 1} = [3\ 5\ 2\ 4].$$

Путь, соединяющий вершину (а) с вершиной *(б),* можно записать в виде

$$D^1_\beta = [3\ 4] \qquad \text{или} \qquad D^2_\beta = [2\ 5].$$

Легко проверить, что конечный результат будет одинаков независимо от того, используется в дальнейших расчетах путь D^1_β или D^2_β. Например, произведение равно

$$\frac{\partial A}{\partial 1} D^1_\beta = \begin{bmatrix}5 & 2 & \cancel{4} & \cancel{3} & 5 & 2\\ 3 & 3 & \cancel{3} & \cancel{4} & 4 & 4\end{bmatrix},$$

а

$$\frac{\partial A}{\partial 1} D^1_\beta \cap A = \begin{bmatrix}2 & 5\\ 3 & 4\end{bmatrix}.$$

Рассматривая ориентацию ребра *1* и напряжения U_2, получим

$$\operatorname*{Sim}_{Z}\left(\frac{\partial A}{\partial 1} D^1{}_\beta,\ A\right) = z_1 z_3 — z_5 z_4.$$

Окончательно

$$K_u = \frac{z_2 z_3 - z_5 z_4}{z_1 z_3 + z_1 z_5 + z_2 z_1 + z_2 z_3 + z_2 z_5 + z_4 z_1 + z_4 z_3 + z_4 z_5}.$$

Пример 9. Определить передаточную функцию напряжения K_u четырехполюсника (рис. 10, *а).*

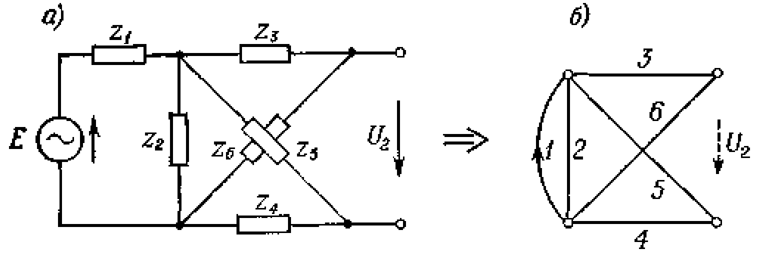


Рис. 10.

Решение. Имеем (рис. 10, б)

$$A = [1\ 2]\ [2\ 4\ 5]\ [2\ 3\ 6],$$

или

$$\begin{array}{r} 1\ 2 \\ \times\ 2\ 4\ 5 \\ 2\ 3\ 6 \\ \hline \end{array}$$

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

Кроме того,

$$\frac{\partial A}{\partial 1} = \begin{bmatrix} 2\ 2\ 4\ 4\ 4\ 5\ 5\ 5 \\ 3\ 6\ 2\ 3\ 6\ 2\ 3\ 6 \end{bmatrix}, \quad D_\beta = [6\ 4]$$

поэтому

$$\frac{\partial A}{\partial 1} D_\beta = \begin{bmatrix} 2 & 4 & 4 & 5 & 5 & 2 & 2 & 5 & 5 & 5 \\ 3 & 2 & 3 & 2 & 3 & 3 & 6 & 2 & 3 & 6 \\ 6 & 6 & 6 & 6 & 6 & 4 & 4 & 4 & 4 & 4 \end{bmatrix},$$

$$\frac{\partial A}{\partial 1} D_\beta \sqcap A = \begin{bmatrix} 5 & 2 \\ 2 & 3 \\ 6 & 4 \end{bmatrix},$$

$$\operatorname*{Sim}_{z} \left(\frac{\partial A}{\partial 1} D_\beta,\ A \right) = z_5 z_2 z_6 - z_2 z_3 z_4.$$

Передаточная функция напряжения запишется в виде

$$K_u = \frac{z_5 z_2 z_6 - z_2 z_3 z_4}{z_1 z_2 z_3 + z_1 z_2 z_6 + z_1 z_4 z_2 + z_1 z_4 z_3 + z_1 z_4 z_6 + z_1 z_5 z_2 + z_1 z_5 z_3 + \ldots + z_2 z_5 z_6}.$$

На практике часто встречается режим работы четырехполюсника с идеальным источником напряжения. При этом передаточная функция напряжения может быть определена по формуле

$$K_u = \frac{\underset{Z}{\mathrm{Sim}}\left(AD_\alpha, AD_\beta\right)}{\underset{z}{\det}(AD_\alpha)}. \tag{12}$$

где D_α — блочная группа, соответствующая пути между узлами, к которым подключен идеальный источник напряжения; D_β — блочная группа, соответствующая пути между измерительными (выходными) узлами четырехполюсника.

Выражение (12) можно использовать как для нагруженного четырехполюсника, так и для четырехполюсника в режиме холостого хода. Приведем пример использования этой формулы.

Пример 10. Рассчитать передаточную функцию напряжения четырехполюсника (2Т-образного моста), изображенного на рис. 11, *а*.

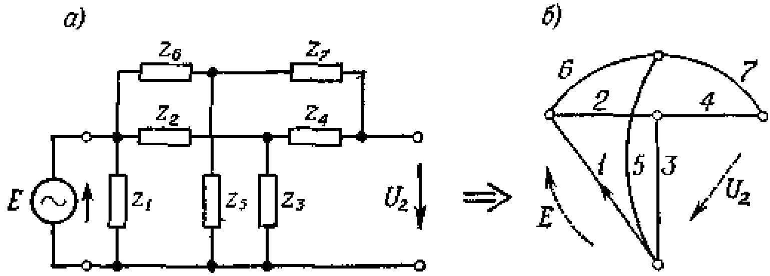


Рис. 11.

Решение. Блочная группа *A* рассматриваемого графа равна (рис. 11, б)

$$\begin{array}{r l} & 1\ 2\ 3 \\ \times & 1\ 5\ 6 \\ & 3\ 4\ 5\ 7 \end{array}$$

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 3 & 3 & 3 & 3 & 3 & 3 & 3 & 3 \\ 5 & 5 & 5 & 6 & 6 & 6 & 6 & 1 & 1 & 1 & 1 & 5 & 5 & 5 & 6 & 6 & 6 & 6 & 1 & 1 & 1 & 5 & 5 & 6 & 6 & 6 \\ 3 & 4 & 7 & 3 & 4 & 5 & 7 & 3 & 4 & 5 & 7 & 3 & 4 & 7 & 3 & 4 & 5 & 7 & 4 & 5 & 7 & 4 & 7 & 4 & 5 & 7 \end{bmatrix}.$$

Кроме того,

$$D_\alpha = [1], \quad D_\beta = [3\ 4],$$

а также

$$AD_\alpha = \left[\begin{array}{ccccccc|c|c|c|c|c|} 1&1&1&1&1&1&1&1&1&1&1&1 \\ 2&2&2&2&2&2&2&3&3&3&3&3 \\ 5&5&5&6&6&6&6&5&5&6&6&6 \\ 3&4&7&3&4&5&7&4&7&4&5&7 \end{array}\right],$$

$$AD_\beta = \left[\begin{array}{|c|c|c|c|c|cccccccccccccccccccccccc} 3&3&3&3&3&3&3&3&3&3&3&3&3&4&4&4&4&4&4&4&4&4&4&4&4&4&4&4&4 \\ 1&1&1&1&1&2&2&2&2&2&2&2&2&1&1&1&1&2&2&2&2&2&2&2&2&3&3&3&3 \\ 5&5&6&6&6&1&1&1&5&5&6&6&6&5&6&6&6&1&1&1&5&5&6&6&6&1&5&6&6 \\ 4&7&4&5&7&4&5&7&4&7&4&5&7&7&3&5&7&3&5&7&3&7&3&5&7&7&7&5&7 \end{array}\right].$$

Поэтому

$$K_u = \frac{(z_3z_5z_4 + z_3z_5z_7 + z_3z_6z_4 + z_3z_6z_5 + z_3z_6z_7)\, z_1}{z_1z_2z_5z_3 + z_1z_2z_5z_4 + z_1z_2z_5z_7 + \ldots + z_1z_3z_6z_7}.$$

2.1.2. Анализ пассивного двухполюсника

Методом блочных групп можно определить входной импеданс или адмитанс пассивной цепи любой сложности. Так как каждый двухполюсник практически всегда работает с источником, имеющим некоторое внутреннее сопротивление или проводимость, то рассмотрим два случая (рис. 12), учитывающих *внутренний иммитанс* источника.

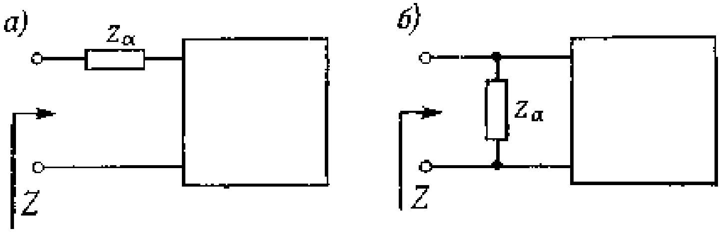


Рис. 12. Двухполюсник с вынесенным импедансом.

Для случая, показанного на рис. 12, *а,* можно получить следующие формулы для входного импеданса *Z:*

$$Z = \frac{\det\limits_{z} A^{d}}{\det\limits_{z} \frac{\partial A^{d}}{\partial \alpha}}, \quad Z = \frac{1}{y_{\alpha}} \frac{\det\limits_{Y} A}{\det\limits_{Y} \frac{\delta A}{\delta \alpha}}. \qquad (12а)$$

Для рис. 12,б имеем

$$Z = z_{\alpha} \frac{\det\limits_{z} \frac{\delta A^{d}}{\delta \alpha}}{\det\limits_{z} A^{d}}, \quad Z = \frac{\det\limits_{Y} \frac{\partial A}{\partial \alpha}}{\det\limits_{Y} A}. \qquad (12б)$$

Во всех приведенных формулах A — блочная группа, геометрическим изображением которого служит граф двухполюсника. В первом случае это граф с *замкнутыми* входными зажимами, во втором случае — с *разомкнутыми.* Вывод формул (12) не представляет особой трудности. В качестве примера докажем первую из формул (12б). Из теории цепей известно, что импеданс, измеренный между зажимами a и b, можно записать

$$Z = \frac{\Delta_{ab}}{\Delta},$$

где Δ и Δ_{ab} — определители матрицы контурных сопротивлений цепи соответственно при разомкнутых и замкнутых вершинах a и *б.* Кроме того,

$$\Delta = \det\limits_{z} A^{d}, \quad \Delta_{ab} = z_{\alpha} \det\limits_{z} \frac{\delta A^{d}}{\delta \alpha},$$

где формула для A_{ab} следует из рассмотренного ранее свойства 2. Проиллюстрируем формулу (12) на примерах.

Пример 11. Определить импеданс мостового двухполюсника (рис.13,*а)*.

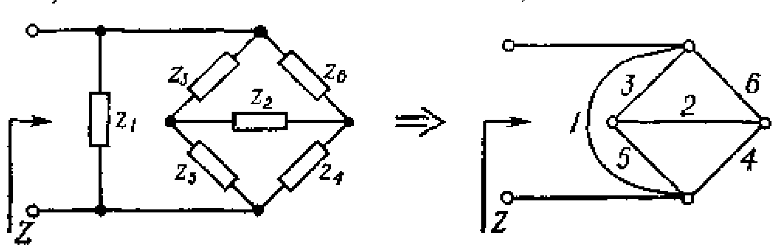


Рис. 13.

Решение. Определим импеданс по первой из формул (12б). Дополнительная блочная группа A^{d} (рис. 13, *б)* равна

$$A^d=[1\ 3\ 5]\ [2\ 3\ 6]\ [2\ 4\ 5],$$

или

$$\times\begin{matrix} 1 & 3 & 5 \\ 2 & 3 & 6 \\ 2 & 4 & 5 \end{matrix}$$

$$A^d=\begin{bmatrix} 1&1&1&1&1&1&1&1&3&3&3&3&3&5&5&5&5&5 \\ 2&2&3&3&3&6&6&6&2&2&6&6&6&2&3&3&6&6 \\ 4&5&2&4&5&2&4&5&4&5&2&4&5&4&2&4&2&4 \end{bmatrix},$$

а производная имеет вид

$$\frac{\delta A^d}{\delta 1}=\begin{bmatrix} 3&3&3&3&3&5&5&5&5&5 \\ 2&2&6&6&6&2&3&3&6&6 \\ 4&5&2&4&5&4&2&4&2&4 \end{bmatrix}.$$

Поэтому

$$Z=z_1\frac{z_3z_2z_4+z_3z_2z_5+z_3z_6z_2+z_3z_6z_4+z_3z_6z_5+z_5z_2z_4+\ldots+z_5z_6z_4}{z_1z_2z_4+z_1z_2z_5+z_1z_3z_2+z_1z_3z_4+z_1z_3z_5+\ldots+z_5z_6z_4}.$$

Пример 12. Определите импеданс двухполюсника (рис. 13), используя вторую из формул (12б).

Решение. В этом случае блочная группа равна

$$A = [1\ 3\ 6]\ [2\ 3\ 5]\ [2\ 4\ 6].$$

Следовательно,

$$\times\begin{matrix} 1 & 3 & 6 \\ 2 & 3 & 5 \\ 2 & 4 & 6 \end{matrix}$$

$$A=\left[\begin{array}{ccccccccc|ccccc|cccc} 1&1&1&1&1&1&1&1&3&3&3&3&3&6&6&6&6&6 \\ 2&2&3&3&3&5&5&5&2&2&5&5&5&2&3&3&5&5 \\ 4&6&2&4&6&2&4&6&4&6&2&4&6&4&2&4&2&4 \end{array}\right],$$

а также

$$\frac{\partial A}{\partial 1}=\begin{bmatrix} 2\,2\,3\,3\,3\,5\,5\,5 \\ 4\,6\,2\,4\,6\,2\,4\,6 \end{bmatrix}.$$

В результате получим

$$Z=\frac{y_2y_4+y_2y_6+y_3y_2+y_3y_4+y_3y_6+y_5y_2+y_5y_4+y_5y_6}{y_1y_2y_4+y_1y_2y_6+y_1y_3y_2+y_1y_3y_4+y_1y_3y_6+\ldots+y_6y_5y_4}.$$

2.1.3. Анализ произвольных цепей

Пусть дана цепь, состоящая из взаимных элементов и источников напряжения. Если в цепи имеются источники тока, то по известным формулам их легко заменить на источники напряжения. На структуру цепи ограничения не наложены. Выделим в цепи ветви с интересующими нас токами $I_{\beta_1}, I_{\beta_2}, ..., I_{\beta_m}$ (рис. 14).

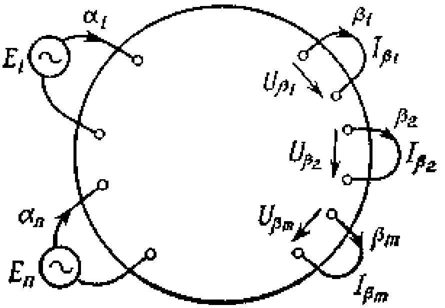


Рис. 14. Электрическая сеть произвольной структуры с вынесенными ветвями.

Рассмотрение этой цепи всегда можно свести к случаю воздействия одного источника напряжения, применив для произвольного числа источников принцип суперпозиции.

Если цепь содержит один источник напряжения, то ток I_β в любой ветви можно представить в виде

$$I_\beta = K_i E, \tag{13}$$

где K_i — передаточная функция тока.

В общем случае имеем

$$I_\beta = \sum_\nu K_{\nu i} E_\nu, \tag{14}$$

причем суммирование проводится по всем источникам.

Напряжение произвольной ветви можно выразить аналогичным образом:

$$U_\beta = \sum_\nu K_{\nu u} E_\nu, \tag{15}$$

где $K_{\nu u}$ — передаточная функция напряжения для ветви ν.

Рассмотренный на примере четырехполюсника метод может быть использован при анализе любых взаимных цепей, не содержащих зависимых источников.

Используя выведенные ранее формулы (2) для передаточных функций тока и напряжения, можно найти распределение токов в цепи и напряжения на отдельных ветвях. Тем же методом можно определить напряжение любой пары вершин цепи. В этом случае достаточно определить

$$K_u = \frac{\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial \alpha} D_\beta, A\right)}{\underset{z}{\det} A}, \tag{16}$$

где D_β — блочная группа, соответствующая пути между интересующими нас вершинами, но не проходящего по ребру α.

Практическое использование этого метода покажем на примерах.

Пример 13. Определить токи I_4 и I_5 цепи (рис. 15, *а)* с двум источниками напряжения.

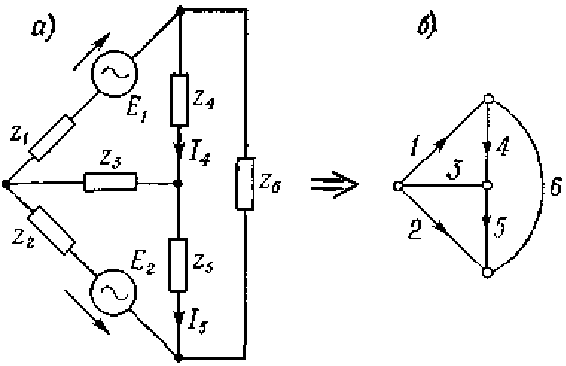


Рис. 15.

Решение. Имеем (рис. 15, *б)*

А = [1 3 41 [2 3 51 [4 5 6],

или

$$\begin{array}{r}
1\;3\;4 \\
\times\quad 2\;3\;5 \\
4\;5\;6 \\
\hline
\end{array},$$

$$A = \left[\begin{array}{ccccccccccc|ccc|cc}
1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 3 & 3 & 3 & 3 & 4 & 4 & 4 & 4 \\
2 & 2 & 2 & 3 & 3 & 3 & 5 & 5 & 2 & 2 & 2 & 5 & 2 & 2 & 3 & 5 \\
4 & 5 & 6 & 4 & 5 & 6 & 4 & 6 & 4 & 5 & 6 & 6 & 5 & 6 & 6 & 6
\end{array}\right]$$

а производные равны

$$\frac{\partial A}{\partial 1}=\begin{bmatrix}2&2&2&3&3&3&5&5\\4&5&6&4&5&6&4&6\end{bmatrix},\qquad \frac{\partial A}{\partial 2}=\begin{bmatrix}1&1&1&3&3&3&4&4\\4&5&6&4&5&6&5&4\end{bmatrix},$$

$$\frac{\partial A}{\partial 4}=\begin{bmatrix}1&1&1&3&2&2&3&5\\2&3&5&2&5&6&6&6\end{bmatrix},\qquad \frac{\partial A}{\partial 5}=\begin{bmatrix}1&1&1&1&3&3&4&4\\2&3&4&6&2&6&2&6\end{bmatrix}.$$

Используя формулу (14), напишем

$$I_4 = K_{14}E_1 + K_{24}E_2; \qquad I_5 = K_{15}E_1 + K_{25}E_2,$$

причем, согласно уравнениям (2), имеем

$$K_{14}=\left.\frac{I_4}{E_1}\right|_{E_2=0}=\frac{\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial 1},\ \frac{\partial A}{\partial 4}\right)}{\underset{Z}{\det} A},$$

$$K_{24}=\left.\frac{I_4}{E_2}\right|_{E_1=0}=\frac{\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial 2},\ \frac{\partial A}{\partial 4}\right)}{\underset{Z}{\det} A},$$

$$K_{15}=\left.\frac{I_5}{E_1}\right|_{E_2=0}=\frac{\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial 1},\ \frac{\partial A}{\partial 5}\right)}{\underset{Z}{\det} A},$$

$$K_{25}=\left.\frac{I_5}{E_2}\right|_{E_1=0}=\frac{\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial 2},\ \frac{\partial A}{\partial 5}\right)}{\underset{Z}{\det} A}.$$

Из приведенных выражений следует

$$K_{14}=\frac{z_2z_5+z_2z_6+z_3z_6+z_5z_6}{z_1z_2z_4+z_1z_2z_5+z_1z_2z_6+z_1z_3z_4+\ldots+z_4z_5z_6}$$

$$K_{24}=\frac{z_1z_5+z_3z_6}{z_1z_2z_4+z_1z_2z_5+z_1z_2z_6+z_1z_3z_4+\ldots+z_4z_5z_6}$$

$$K_{15}=\frac{z_2z_4+z_3z_6}{z_1z_2z_4+z_1z_2z_5+z_1z_2z_6+z_1z_3z_4+\ldots+z_4z_5z_6}$$

$$K_{25}=\frac{-z_1z_4-z_1z_6-z_4z_6-z_3z_6}{z_1z_2z_3+z_1z_2z_5+z_1z_2z_6+z_1z_3z_4+\ldots+z_4z_5z_6}.$$

Пример 14. Определить напряжение U_0 цепи (рис. 16, *а)* с тремя источниками.

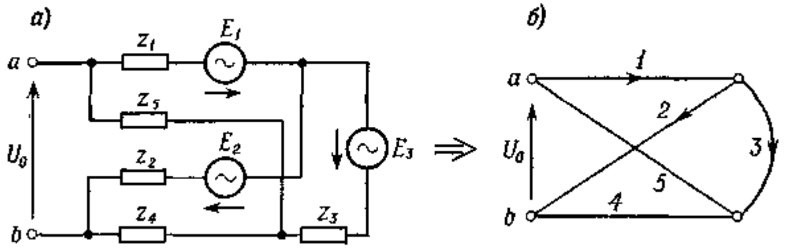


Рис. 16.

Решение. Для рассматриваемого случая блочная группа равна

$$A = [1\ 3\ 51\ [2\ 3\ 4] = \begin{bmatrix} 1\ 1\ 1 3 3\ \ 5 5 5 \\ 2\ 3\ 4\ 2\ 4\ 2\ 3\ 4 \end{bmatrix},$$

следовательно,

$$\frac{\partial A}{\partial 1} = [2\ 3\ 4], \frac{\partial A}{\partial 2} = [1\, 3\, 5], \frac{\partial A}{\partial 3} = [1\ 2\ 4\ 5].$$

Пусть, например, путь D, соединяющий узлы a и б, описывается блочной группой [5 4] (рис. 16, б), т. е.

$$D = [5\ 4].$$

Тогда

$$\frac{\partial A}{\partial 1} D = \begin{bmatrix} 2 & 3 & 4 & 2 & 3 \\ 5 & 5 & 5 & 4 & 4 \end{bmatrix}; \qquad \frac{\partial A}{\partial 2} D = \begin{bmatrix} 1 & 3 & 3 & 1 & 3 & 5 \\ 5 & 5 & 4 & 4 & 4 & 4 \end{bmatrix}; \qquad \frac{\partial A}{\partial 3} D = \begin{bmatrix} 1 & 2 & 4 & 1 & 2 & 5 \\ 5 & 5 & 5 & 4 & 4 & 4 \end{bmatrix}.$$

Отсюда можно записать соответствующие функции совпадения

$$\operatorname*{Sim}_{Z} = \left(\frac{\partial A}{\partial 1} D,\ A\right) = z_2z_5 + z_3z_5 + z_4z_5 + z_3z_4,$$

$$\operatorname*{Sim}_{Z} = \left(\frac{\partial A}{\partial 2} D,\ A\right) = z_3z_5 + z_1z_4 + z_3z_4 + z_5z_4,$$

$$\operatorname*{Sim}_{Z} = \left(\frac{\partial A}{\partial 3} D,\ A\right) = z_2z_5 - z_1z_4.$$

Окончательно получим

$$U_0 = \frac{(z_2z_5 + z_3z_5 + z_4z_5 + z_3z_4)\, E_1}{z_1z_2 + z_1z_3 + z_1z_4 + z_3z_2 + z_3z_4 + z_5z_2 + z_5z_3 + z_5z_4} +$$

$$+ \frac{(z_3z_5 + z_1z_4 + z_3z_4 + z_5z_4)\, E_2 + (z_2z_5 - z_1z_4)\, E_3}{z_1z_2 + z_1z_3 + z_1z_4 + z_3z_2 + z_3z_4 + z_5z_2 + z_5z_3 + z_5z_4}.$$

Так же довольно просто вычислить *импеданс или адмитанс между произвольной парой узлов цепи.* Для этого достаточно воспользоваться формулой

$$Z = \frac{\det\limits_{Z}(AD)}{\det\limits_{Z} A}, \tag{17}$$

где A — блочная группа, для которой граф цепи служит обратным изображением, D — блочная группа произвольного пути, соединяющего узлы, между которыми определяется импеданс **Z.**

Порядок расчета токов, напряжений или передаточных функций цепи можно упростить, если вначале упростить граф цепи, заменяя параллельные и последовательные ребра одним ребром.

Для последовательного соединения ребер имеем

$$Z_i = z_1 + z_2 + \ldots + z_n,$$

для параллельного соединения

$$Z_i = \frac{1}{\frac{1}{z_1} + \frac{1}{z_2} + \ldots + \frac{1}{z_m}}.$$

2.2. Анализ активных цепей

2.2.1 Анализ цепи, содержащей один зависимый источник напряжения

Цепь, содержащая один зависимый источник напряжения, изображена на рис. 17.

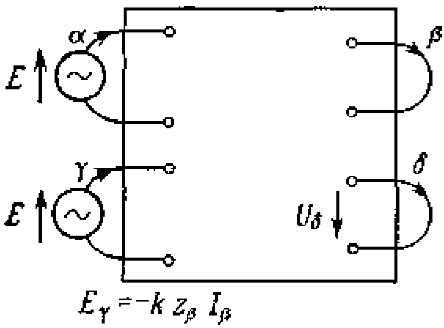


Рис. 17.

Выделим ветви цепи, как показано на рис. 17. Напряжение U_δ на выходе с помощью блочных групп можно записать в виде

$$U_\delta = \frac{\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial\alpha}, \frac{\partial A}{\partial\delta}\right) z_\delta E}{\underset{Z}{\det} A} - \frac{\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial\gamma}, \frac{\partial A}{\partial\delta}\right) z_\beta I_\beta K z_\delta}{\underset{Z}{\det} A}, \quad (18)$$

где ток I_β равен

$$I_\beta = \frac{\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial\alpha}, \frac{\partial A}{\partial\beta}\right) E}{\underset{Z}{\det} A} - \frac{\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial\gamma}, \frac{\partial A}{\partial\beta}\right) K z_\beta I_\beta}{\underset{Z}{\det} A}.$$

В этой формуле ток I_β выражен в неявном виде. После соответствующих преобразований получим

$$I_\beta = \frac{\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial\alpha}, \frac{\partial A}{\partial\beta}\right)}{\underset{Z}{\det} A + K z_\beta \underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial\beta}, \frac{\partial A}{\partial\gamma}\right)} E. \quad (19)$$

Подставив уравнение (19) в уравнение (18) и выполнив преобразования, получим выражение передаточной функции цепи, содержащей один зависимый источник напряжения:

$$\frac{U_\beta}{E} = \frac{\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial\alpha}, \frac{\partial A}{\partial\delta}\right)}{\underset{Z}{\det} A} z_\delta - \frac{\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial\gamma}, \frac{\partial A}{\partial\delta}\right) \underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial\alpha}, \frac{\partial A}{\partial\beta}\right) K z_\beta}{\underset{Z}{\det} A \left[\underset{Z}{\det} A + K z_\beta \underset{Z}{\mathrm{Sim}}\left(\frac{\partial A}{\partial\beta}, \frac{\partial A}{\partial\gamma}\right)\right]} z_\delta. \quad (20)$$

Эта формула определяет коэффициент усиления напряжения цепи, имеющий две составляющие:

$$U_\delta/E = K_0 - K_1. \quad (21)$$

Составляющая K_0 представляет собой усиление цепи при отсутствии активной связи, т. е. когда коэффициент усиления импеданса $K = 0$; составляющая K_1 — дополнительное усиление, обусловленное наличием зависимого источника напряжения.

Выражение (20) можно значительно упростить, представив его в следующем виде:

$$\frac{U_\delta}{E}=\frac{z_\delta}{\det\limits_Z A}\left[\frac{\operatorname*{Sim}\limits_Z\left(\frac{\partial A}{\partial\alpha},\ \frac{\partial A}{\partial\delta}\right)\left[\det\limits_Z A+W\operatorname*{Sim}\limits_Z\left(\frac{\partial A}{\partial\beta},\ \frac{\partial A}{\partial\gamma}\right)\right]-}{\det\limits_Z A+W\operatorname*{Sim}\limits_Z\left(\frac{\partial A}{\partial\beta},\ \frac{\partial A}{\partial\gamma}\right)}\right.$$

$$\left.\frac{-W\operatorname*{Sim}\limits_Z\left(\frac{\partial A}{\partial\alpha},\ \frac{\partial A}{\partial\beta}\right)\operatorname*{Sim}\limits_Z\left(\frac{\partial A}{\partial\gamma},\ \frac{\partial A}{\partial\delta}\right)}{}\right],$$

где $W = z_\beta K$.

Параметр W называется передаточным активным импедансом зависимого источника.

Если через L_1обозначить выражение

$$L_1=\frac{\operatorname*{Sim}\limits_Z\left(\frac{\partial A}{\partial\alpha},\ \frac{\partial A}{\partial\beta}\right)\operatorname*{Sim}\limits_Z\left(\frac{\partial A}{\partial\gamma},\ \frac{\partial A}{\partial\delta}\right)-\operatorname*{Sim}\limits_Z\left(\frac{\partial A}{\partial\alpha},\ \frac{\partial A}{\partial\delta}\right)\operatorname*{Sim}\limits_Z\left(\frac{\partial A}{\partial\beta},\ \frac{\partial A}{\partial\gamma}\right)}{\det\limits_Z A},\tag{22}$$

то выражение (20) примет вид

$$\frac{U_\delta}{E}=\frac{\operatorname*{Sim}\limits_Z\left(\frac{\partial A}{\partial\alpha},\ \frac{\partial A}{\partial\delta}\right)-WL_1}{\det\limits_Z A+W\operatorname*{Sim}\limits_Z\left(\frac{\partial A}{\partial\beta},\ \frac{\partial A}{\partial\gamma}\right)}z_\delta.\tag{23}$$

Выражение (22) можно упростить, учитывая, что функции совпадения определяют соответствующие миноры определителя контурных сопротивлений Δ^0 исследуемой цепи (при $K = 0$) Имеем следующее равенство:

$$\operatorname*{Sim}\limits_Z\left(\frac{\partial A}{\partial\mu},\ \frac{\partial A}{\partial\gamma}\right)=\Delta^0_{\mu\nu},\tag{24}$$

где $\Delta^0_{\mu\nu}$ — минор, полученный из определителя Δ^0 путем вычеркивания μ-строки и ν-столбца.

Если учесть это равенство в формуле (22), то для L_1 получим следующее выражение:

$$L_1=\frac{\Delta^0_{\alpha\beta}\Delta^0_{\gamma\delta}-\Delta^0_{\alpha\delta}\Delta^0_{\beta\gamma}}{\Delta^0}.$$

Из теории определителей известно, что

$$\Delta^0_{\alpha\beta}\Delta^0_{\gamma\delta}-\Delta^0_{\alpha\delta}\Delta^0_{\beta\gamma}=\Delta^0\Delta^0_{\alpha\beta,\gamma\delta},\tag{25}$$

где $\Delta^0_{\alpha\beta,\gamma\delta}$ —минор, полученный вычеркиванием строк α и γ и столбцов β и δ из определителя Δ^0, поэтому величину L_1 МОЖНО определить непосредственно с помощью одной функции совпадения, а именно

$$L_1 = \Delta_{\alpha\beta,\,\gamma\delta} = \underset{Z}{\mathrm{Sim}} \left(\frac{\partial^2 A}{\partial\alpha,\ \partial\gamma},\ \frac{\partial^2 A}{\partial\beta,\ \partial\delta} \right). \qquad (26)$$

Подставив уравнение (26) в (23), окончательно получим следующее выражение для усиления напряжения рассматриваемой цепи:

$$\frac{U_\delta}{E} = \frac{\underset{Z}{\mathrm{Sim}} \left(\frac{\partial A}{\partial\alpha},\ \frac{\partial A}{\partial\delta} \right) - Kz_\beta \underset{Z}{\mathrm{Sim}} \left(\frac{\partial^2 A}{\partial\alpha,\ \partial\gamma},\ \frac{\partial^2 A}{\partial\beta,\ \partial\delta} \right)}{\underset{Z}{\det} A + Kz_\beta \underset{Z}{\mathrm{Sim}} \left(\frac{\partial A}{\partial\beta},\ \frac{\partial A}{\partial\gamma} \right)} z_\delta. \qquad (27)$$

Отметим, что знаменатель этого выражения представляет собой определитель матрицы контурных сопротивлений Δ цепи, содержащей один зависимый источник напряжения. Следовательно,

$$\Delta = \underset{Z}{\det} A + Kz_\beta \underset{Z}{\mathrm{Sim}} \left(\frac{\partial A}{\partial\beta},\ \frac{\partial A}{\partial\gamma} \right). \qquad (28)$$

Во всех приведенных формулах A — блочная группа, для которой граф исследуемой цепи служит обратным изображением.

Если известно усиление напряжения цепи и определитель Δ, то легко найти все остальные параметры, характеризующие эту цепь.

Например, *возвратная разность* равна

$$F = 1 + \frac{Kz_\beta \underset{Z}{\mathrm{Sim}} \left(\frac{\partial A}{\partial\beta},\ \frac{\partial A}{\partial\gamma} \right)}{\underset{Z}{\det} A}. \qquad (29)$$

Эта формула следует из зависимости, полученной Боде, согласно которой

$$F = \Delta/\Delta^0.$$

Рассмотрим два частных случая.

Допустим, что импеданс z_β стремится к бесконечности ($z_\beta \to \infty$). Учитывая соотношение

$$\underset{Z}{\det} A = \underset{Z}{\det} \frac{\delta A}{\delta\beta} + z_\beta \underset{Z}{\det} \frac{\partial A}{\partial\beta}, \qquad (30)$$

непосредственно вытекающее из зависимости

$$A = \frac{\delta A}{\delta\beta} + [\beta] \frac{\partial A}{\partial\beta},$$

получим формулу для усиления напряжения цепи, содержащей один зависимый источник напряжения при условии, что $z_\beta \to \infty$:

$$\left. \frac{U_\delta}{E} \right|_{z_\beta \to \infty} = \frac{-K \underset{Z}{\mathrm{Sim}} \left(\frac{\partial^2 A}{\partial\alpha\, \partial\gamma},\ \frac{\partial^2 A}{\partial\beta\, \partial\delta} \right)}{\underset{Z}{\det} \frac{\partial A}{\partial\beta} + K \underset{Z}{\mathrm{Sim}} \left(\frac{\partial A}{\partial\beta},\ \frac{\partial A}{\partial\gamma} \right)} z_\delta. \qquad (31)$$

Рассмотрим случай, соответствующий очень сильной обратной связи, т. е. когда $K\,z_\beta\to\infty$.

Тогда усиление напряжения выражается формулой

$$\left.\frac{U_\delta}{E}\right|_{Kz_\beta\to\infty}=\frac{\operatorname*{Sim}_{Z}\left(\frac{\partial^2 A}{\partial\alpha\,\partial\gamma},\ \frac{\partial^2 A}{\partial\beta\,\partial\delta}\right)}{\operatorname*{Sim}_{Z}\left(\frac{\partial A}{\partial\beta},\ \frac{\partial A}{\partial\gamma}\right)}\,z_\delta. \qquad (32)$$

Пример 15. Для цепи, изображенной на рис. 118, определить: 1) усиление напряжения в общем случае; 2) усиление напряжения при $z_\beta\to\infty$; 3) усиление напряжения при $Kz_\beta\to\infty$; 4) возвратную разность F.

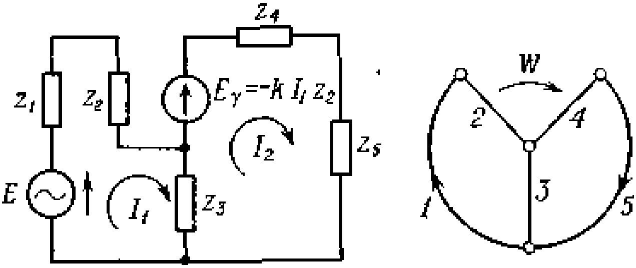


Рис. 18.

1) Для расчета усиления напряжения в общем случае воспользуемся формулой (27). Имеем

$\alpha = 1,\quad \beta = 2,\quad \gamma=4,\quad \delta=5.$

Блочная группа этой цепи равна

$$A = [1\ 2\ 3]\ [3\ 4\ 5]=\begin{bmatrix}11122233\\34534545\end{bmatrix}.$$

поэтому

$$\frac{\partial A}{\partial 1}=[3\ 4\ 5],\qquad \frac{\partial A}{\partial 2}=[3\ 4\ 5],$$

$$\frac{\partial A}{\partial 4}=[1\ 2\ 3],\qquad \frac{\partial A}{\partial 5}=[1\ 2\ 3],$$

$$\frac{\partial^2 A}{\partial 1\,\partial 4}=1,\qquad \frac{\partial^2 A}{\partial 2\,\partial 5}=1,$$

$$\operatorname*{Sim}_{Z}\left(\frac{\partial A}{\partial\alpha},\ \frac{\partial A}{\partial\delta}\right)\quad z_3,\qquad \operatorname*{Sim}_{Z}\left(\frac{\partial A}{\partial\gamma},\ \frac{\partial A}{\partial\beta}\right)=z_3.$$

Подста,вив эти величины в выражение (27), получим

$$\frac{U_5}{E}=\frac{z_3z_5-Kz_2z_5}{z_1z_3+z_1z_4+z_1z_5+z_2z_3+z_2z_4+z_2z_5+z_3z_4+z_3z_5+Kz_2z_3}.$$

2) Усиление напряжения при $z_\beta \to \infty$ определяем в соответствии с формулой (31)

$$\frac{U_5}{E}=\frac{-Kz_5}{z_3+z_4+z_5+Kz_3}.$$

3) Усиление напряжения при $Kz_\beta \to \infty$ находим по формуле (32)

$$\frac{U_5}{E}=-\frac{z_5}{z_3}.$$

4) Возвратную разность по отношению к элементу *К* можно определить по формуле (29):

$$F=1+\frac{Kz_2z_3}{z_1z_3+z_1z_4+z_1z_5+z_2z_3+z_2z_3+z_2z_5+z_3z_4+z_3z_5}.$$

2.2.2. Анализ цепи, содержащей два зависимых источника напряжения

Рассматривая цепь, содержащую два зависимых источника напряжения, поступим, как и в случае одного зависимого источника.

Выделим из линейной цепи зависимые источники напряжения и ветви, содержащие управляющие сигналы. В результате получим схему, изображенную на рис. 19, где обозначено: *E* — независимый источник напряжения; E_{γ_1}, E_{γ_2} — зависимые источники напряжения; I_{β_1} — ток, управляющий источником E_{γ_1}; I_{β_2} — ток, управляющий источником E_{γ_2}; U_δ — выходное напряжение.

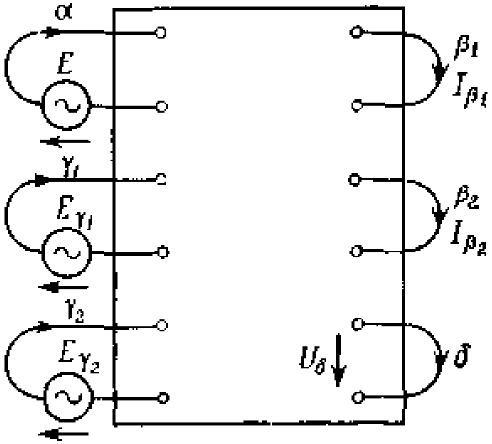


Рис. 19. Цепь с двумя зависимыми источниками э.д.с.

Кроме этого, справедливы равенства

$$E_{\gamma_1} = - K_1\, z_{\beta_1}\, I_{\beta_1}, \qquad E_{\gamma_2} = -K_2\, z_{\beta_2}\, I_{\beta_2}, \tag{33}$$

где z_{β_1} и z_{β_2} — импедансы ребер β_1 и β_2 соответственно.

Напряжение на выходе цепи U_δ можно рассматривать как результат наложения трех источников напряжения, т. е.

$$U_\delta = \frac{\Delta^0_{\alpha\delta}}{\Delta^0} z_\delta E + \frac{\Delta^0_{\gamma_1\delta}}{\Delta^0} z_\delta E_{\gamma_1} + \frac{\Delta^0_{\gamma_2\delta}}{\Delta^0} z_\delta E_{\gamma_2}, \tag{34}$$

где Δ^0 — главный определитель матрицы контурных сопротивлений пассивной цепи без зависимых источников напряжения;

$\Delta^0_{\alpha\delta}$ — минор определителя Δ^0, полученный вычеркиванием строки α и столбца δ;

$\Delta^0_{\gamma_1\delta}$ — минор определителя Δ^0, полученный вычеркиванием строки γ_1 и столбца δ;

$\Delta^0_{\gamma_2\delta}$ — минор определителя Δ^0, полученный вычеркиванием строки γ_2 и столбца δ.

Выражение (34) получено с помощью метода контурных токов Максвелла.

Токи I_{β_1} и I_{β_2} также можно определить, исходя из системы контурных уравнений и принципа наложения:

$$I_{\beta_1}=\frac{\Delta^0_{\alpha\beta_1}}{\Delta^0}E+\frac{\Delta^0_{\beta_1\gamma_1}}{\Delta^0}E_{\gamma_1}+\frac{\Delta^0_{\beta_1\gamma_2}}{\Delta^0}E_{\gamma_2},$$

$$I_{\beta_2}=\frac{\Delta^0_{\alpha\beta_2}}{\Delta^0}E+\frac{\Delta^0_{\beta_2\gamma_1}}{\Delta^0}E_{\gamma_1}+\frac{\Delta^0_{\beta_2\gamma_2}}{\Delta^0}E_{\gamma_2}. \tag{35}$$

Учитывая соотношение (33) и вводя обозначения

$$W_1 = K z_{\beta_1}, \qquad W_2 = K z_{\beta_2}$$

(здесь W_1 и W_2 называются передаточными активными импедансами цепи), получим систему уравнений

$$I_{\beta_1}\left[1+\frac{\Delta^0_{\gamma_1\beta_1}}{\Delta^0}W_1\right]+\frac{\Delta^0_{\gamma_2\beta_1}}{\Delta^0}I_{\beta_2}W_2=\frac{\Delta^0_{\alpha\beta_1}}{\Delta^0}E,$$

$$I_{\beta_2}\left[1+\frac{\Delta^0_{\gamma_2\beta_2}}{\Delta^0}W_2\right]+\frac{\Delta^0_{\gamma_1\beta_2}}{\Delta^0}I_{\beta_1}W_1=\frac{\Delta^0_{\alpha\beta_2}}{\Delta^0}E \tag{36}$$

Определитель этой системы имеет вид

$$\Delta'=\begin{vmatrix}\Delta^0+W_1\Delta^0_{\gamma_1\beta_1} & W_2\Delta^0_{\gamma_2\beta_1}\\ W_1\Delta^0_{\gamma_1\beta_2} & \Delta^0+W_2\Delta^0_{\gamma_2\beta_2}\end{vmatrix},$$

первый минор:

$$\Delta'_1=\begin{vmatrix}\Delta^0_{\alpha\beta_1}E & W_2\Delta^0_{\gamma_2\beta_1}\\ \Delta^0_{\alpha\beta_2}E & \Delta^0+W_2\Delta^0_{\gamma_2\beta_2}\end{vmatrix},$$

второй минор:

$$\Delta'_2=\begin{vmatrix}\Delta^0+W_1\Delta^0_{\gamma_1\beta_1} & \Delta^0_{\alpha\beta_1}E\\ W_1\Delta^0_{\gamma_1\beta_2} & \Delta^0_{\alpha\beta_2}E\end{vmatrix}.$$

Ток можно записать в следующем виде:

$$I_{\beta_1}=\frac{\Delta'_1}{\Delta'}=\frac{\Delta^0_{\alpha\beta_1}(\Delta^0+W_2\Delta^0_{\gamma_2\beta_2})-\Delta^0_{\alpha\beta_2}\Delta^0_{\gamma_2\beta_1}W_2E}{(\Delta^0+W_1\Delta^0_{\gamma_1\beta_1})(\Delta^0+W_2\Delta^0_{\gamma_2\beta_2})-W_1W_2\Delta^0_{\gamma_1\beta_2}\Delta^0_{\gamma_2\beta_1}}. \tag{37}$$

После преобразований окончательно получим

$$I_{\beta_1}=\frac{(\Delta^0_{\alpha\beta_1}+W_2L_1)E}{\Delta^0+W_1\Delta^0_{\gamma_1\beta_1}+W_2\Delta^0_{\gamma_2\beta_2}-W_1W_2L_{12}}, \tag{38}$$

где

$$L_1=\frac{\Delta^0_{\alpha\beta_1}\Delta^0_{\gamma_2\beta_2}-\Delta^0_{\alpha\beta_2}\Delta^0_{\gamma_2\beta_1}}{\Delta^0}=\Delta^0_{\alpha\beta_1,\,\gamma_2\beta_2},$$

$$L_{12}=\frac{\Delta^0_{\gamma_1\beta_1}\Delta^0_{\gamma_2\beta_2}-\Delta^0_{\gamma_1\beta_1}\Delta^0_{\gamma_2\beta_1}}{\Delta^0}=\Delta^0_{\gamma_1\beta_1,\,\gamma_2\beta_2}. \tag{39}$$

Формулы (39) можно получить, раскрыв определитель по методу Лапласа.

Используя метод блочных групп, полученные миноры, можно выразить через функции совпадения:

$$L_1 = \operatorname*{Sim}_Z \left(\frac{\partial^2 A}{\partial\alpha\,\partial\gamma_2},\ \frac{\partial^2 A}{\partial\beta_1\,\partial\beta_2} \right),$$

$$L_{12} = \operatorname*{Sim}_Z \left(\frac{\partial^2 A}{\partial\gamma_1\,\partial\gamma_2},\ \frac{\partial^2 A}{\partial\beta_1\,\partial\beta_2} \right), \qquad (40)$$

где A — блочная группа, для которой граф цепи служит обратным изображением, Z — множество импедансов цепи. Учитывая соотношения (39), ток I_{β_1} можно записать в виде

$$I_{\beta_1} = \frac{(\Delta^0_{\alpha\beta_1} + W_2\Delta^0_{\alpha\beta_1,\ \gamma_2\beta_2})\,E}{\Delta^0 + W_1\Delta^0_{\gamma_1\beta_1} + W_2\Delta^0_{\gamma_2\beta_2} - W_1W_2\Delta^0_{\beta_1\gamma_1,\ \beta_1\gamma_2}} = \frac{N_1}{\Delta}E. \qquad (41)$$

Аналогично для тока I_{β_2} имеем

$$I_{\beta_2} = \frac{\Delta'_2}{\Delta'} = \frac{(\Delta^0_{\alpha\beta_2} - W_1L_2)\,E}{\Delta'}, \qquad (42)$$

где

$$L_2 = \frac{\Delta^0_{\gamma_1\beta_2}\Delta^0_{\alpha_1\beta_1} - \Delta^0_{\alpha\beta_2}\Delta^0_{\beta_1\gamma_1}}{\Delta^0} = \Delta^0{}_{\alpha\beta_1,\ \gamma_1\beta_2}. \qquad (43)$$

В этой формуле $\Delta^0{}_{\alpha\beta_1,\gamma_1\beta_2}$ — минор, полученный вычеркиванием строк α и γ и столбцов β_1 и β_2 из определителя Δ^0. Этот минор можно также выразить с помощью функции совпадения

$$L_2 = \operatorname*{Sim}_Z \left(\frac{\partial^2 A}{\partial\alpha\,\partial\gamma_1},\ \frac{\partial^2 A}{\partial\beta_1\,\partial\beta_2} \right). \qquad (44)$$

В формуле (43), как и выше, A — блочная группа, для которой граф цепи служит обратным изображением, а Z — множество импедансов цепи.

Учитывая формулу (43), ток I_{β_2} можно представить в виде

$$I_{\beta_2} = \frac{(\Delta^0_{\alpha\beta_2} - W_1\Delta^0_{\alpha\beta_1,\gamma_1\beta_2})\,E}{\Delta} = \frac{N_2E}{\Delta}, \qquad (45)$$

где

$$\Delta = \Delta^0 + W_1\,\Delta^0{}_{\alpha\beta_1} + W_2\,\Delta^0{}_{\gamma_2\beta_2} - W_1\,W_{12}\,\Delta^0{}_{\beta_1\gamma_1,\beta_2\gamma_2}.$$

Подставляя соотношения (41) и (45) в (34), получим выражение, определяющее усиление цепи, содержащей два зависимых источника напряжения:

$$\frac{U_\delta}{E} = \left(\frac{\Delta^0_{\alpha\delta}}{\Delta^0} - \frac{\Delta^0_{\gamma_1\delta}}{\Delta^0}\frac{N_1}{\Delta}W_1 - \frac{\Delta^0_{\gamma_2\delta}}{\Delta^0}\frac{N_2}{\Delta}W_2 \right) z_\delta.$$

Эту формулу можно преобразовать к следующему виду:

$$\frac{U_\delta}{E} = \frac{z_\delta}{\Delta^0}\,\frac{\Delta^0_{\alpha\delta}\Delta - \Delta^0_{\gamma_1\delta}N_1W_1 - \Delta^0_{\gamma_2\delta}N_2W_2}{\Delta}. \qquad (46)$$

Подставив в выражение (46) значения коэффициентов N_1 и N_2, получим

$$\frac{U_\delta}{E}=\frac{\Delta^0_{\alpha\delta}-W_1L_1^*-W_2L_2^*+W_1W_2L_{12}^*}{\Delta}, \tag{47}$$

где

$$L_1^*=\frac{\Delta^0_{\gamma_1\delta}\Delta^0_{\alpha\beta_1}-\Delta^0_{\alpha\delta}\Delta^0_{\beta_1\gamma_1}}{\Delta^0}=\Delta^0_{\gamma_1\delta,\ \alpha\beta_1},$$

$$L_2^*=\frac{\Delta^0_{\gamma_2\delta}\Delta^0_{\alpha\beta_2}-\Delta^0_{\alpha\delta}\Delta^0_{\beta_2\gamma_2}}{\Delta^0}=\Delta^0_{\gamma_2\delta,\ \alpha\beta_2}, \tag{48}$$

$$L_{12}^*=\frac{\Delta^0_{\alpha\delta}\Delta^0_{\beta_1\gamma_1,\ \beta_2\gamma_2}-\Delta^0_{\gamma_1\delta}\Delta^0_{\alpha\beta_1,\ \gamma_2\beta_2}+\Delta^0_{\gamma_2\delta}\Delta^0_{\alpha\beta_1,\ \gamma_1\beta_2}}{\Delta^0}=\Delta^0_{\alpha\delta,\ \beta_1\gamma_1,\ \beta_2\gamma_2}.$$

Здесь, как и выше, величины $\Delta^0_{\alpha\beta_1,\gamma_1\beta_2}$, $\Delta^0_{\beta_1\gamma_1,\beta_2\gamma_2}$, $\Delta^0_{\alpha\beta_1,\gamma_2\beta_2}$ — соответствующие миноры определителя Δ^0, которые можно выразить через функции совпадения

$$L_1^*=\operatorname*{Sim}_Z\left(\frac{\partial^2 A}{\partial\alpha\,\partial\gamma_1},\ \frac{\partial^2 A}{\partial\beta_1\,\partial\delta}\right),$$

$$L_2^*=\operatorname*{Sim}_Z\left(\frac{\partial^2 A}{\partial\alpha\,\partial\gamma_2},\ \frac{\partial^2 A}{\partial\beta_2\,\partial\gamma}\right), \tag{49}$$

$$L_{12}^*=\operatorname*{Sim}_Z\left(\frac{\partial^3 A}{\partial\alpha\,\partial\gamma_1\,\partial\gamma_2},\ \frac{\partial^3 A}{\partial\beta_1\,\partial\beta_2\,\partial\delta}\right),$$

причем граф цепи служит обратным изображением блочной группы A, Z — множество импедансов цепи.

Учитывая полученные результаты, окончательно запишем формулу силения цепи, содержащей два зависимых источника напряжения:

$$\frac{U_\delta}{E}=\frac{\begin{gathered}\operatorname*{Sim}_Z\left(\frac{\partial A}{\partial\alpha},\frac{\partial A}{\partial\delta}\right)-W_1\operatorname*{Sim}_Z\left(\frac{\partial^2 A}{\partial\alpha\,\partial\gamma_1},\ \frac{\partial^2 A}{\partial\beta_1\,\partial\delta}\right)-\\-W_2\operatorname*{Sim}_Z\left(\frac{\partial^2 A}{\partial\alpha\,\partial\gamma_2},\frac{\partial^2 A}{\partial\beta_2\,\partial\delta}\right)+W_1W_2\operatorname*{Sim}_Z\left(\frac{\partial^2 A}{\partial\alpha\,\partial\gamma_1\,\partial\gamma_2},\frac{\partial^3 A}{\partial\delta\,\partial\beta_1\,\partial\beta_2}\right)z_\delta\end{gathered}}{\begin{gathered}\det_Z A+W_1\operatorname*{Sim}_Z\left(\frac{\partial A}{\partial\beta_1},\ \frac{\partial A}{\partial\gamma_1}\right)+W_2\operatorname*{Sim}_Z\left(\frac{\partial A}{\partial\beta_2},\frac{\partial A}{\partial\gamma_2}\right)-\\-W_1W_2\operatorname*{Sim}_Z\left(\frac{\partial^2 A}{\partial\beta_1\,\partial\beta_2},\ \frac{\partial^2 A}{\partial\gamma_1\,\partial\gamma_2}\right)\end{gathered}}. \tag{50}$$

Отдельным величинам и выражениям этой формулы можно дать простую физическую интерпретацию.

Первые члены числителя и знаменателя выражения (50) характеризуют непосредственное прохождение сигнала в предположении, что оба коэффициента усиления равны нулю (т. е. оба активных передаточных импеданса W_1 и W_2 равны нулю). Таким образом, непосредственная передача сигнала определяется формулой

$$\frac{U_\delta}{E}=\frac{\operatorname*{Sim}_{Z}\left(\frac{\partial A}{\partial \alpha},\ \frac{\partial A}{\partial \delta}\right)}{\operatorname*{det}_{Z} A}.$$

Два первых члена числителя и два первых члена знаменателя определяют усиление цепи при нулевом коэффициенте усиления другого элемента $(W_2 = 0)$. Аналогично первый и третий члены числителя и знаменателя определяют усиление цепи при нулевом коэффициенте усиления первого элемента $(W_1 = 0)$.

2.2.3. Формулы для расчета цепи, содержащей N зависимых источников напряжения

На основании формул для цепей, содержащих один или два зависимых источника, можно вывести формулы для цепи, содержащей N зависимых источников напряжения.

Формула усиления цепи с N зависимыми источниками напряжения имеет вид

$$\frac{U_\delta}{E}=\frac{\operatorname*{Sim}_{Z}\left(\frac{\partial A}{\partial \alpha},\ \frac{\partial A}{\partial \delta}\right)-\sum\limits_{i=1}^{N} W_i \operatorname*{Sim}_{Z}\left(\frac{\partial^2 A}{\partial \alpha\, \partial \gamma_i},\ \frac{\partial^2 A}{\partial \beta_i\, \partial \delta}\right)+\sum\limits_{i<j} W_i W_j \operatorname*{Sim}_{Z}\left(\frac{\partial^3 A}{\partial \alpha\, \partial \gamma_i\, \partial \gamma_j},\ \frac{\partial^3 A}{\partial \beta_i\, \partial \beta_j\, \partial \delta}\right)-\ldots}{\operatorname*{det}_{Z} A+\sum\limits_{i=1}^{N} W_i \operatorname*{Sim}_{Z}\left(\frac{\partial A}{\partial \beta_i},\ \frac{\partial A}{\partial \gamma_i}\right)-\sum\limits_{i<j} W_i W_j \left(\frac{\partial^2 A}{\partial \beta_i\, \partial \beta_j},\ \frac{\partial^2 A}{\partial \gamma_i\, \partial \gamma_j}\right)+\ldots}\, z_\delta, \tag{51}$$

где $W_i = K_i z_{\beta_i}$.

Если предположить, что все активные передаточные импедансы стремятся к бесконечности (этот случай соответствует очень сильной обратной связи), то формула для усиления упростится и примет вид

$$\left.\frac{U_\delta}{E}\right|_{W_1W_2\ldots W_n\to\infty}=$$

$$=(-1)^N\frac{\underset{Z}{\mathrm{Sim}}\left(\dfrac{\partial^{N+1}A}{\partial\alpha\,\partial\gamma_1\ldots\partial\gamma_N},\ \dfrac{\partial^{N+1}A}{\partial\beta_1\,\partial\beta_2\ldots\partial\beta_N\,\partial\delta}\right)}{\underset{Z}{\mathrm{Sim}}\left(\dfrac{\partial^N A}{\partial\beta_1\,\partial\beta_2\ldots\partial\beta_N},\ \dfrac{\partial^N A}{\partial\gamma_1\,\partial\gamma_2\ldots\partial\gamma_N}\right)}z_\delta. \tag{52}$$

Формулу для возвратной разности можно вывести на основании свойства, полученного Боде, согласно которому возвратная разность выражается отношением определителя матрицы импедансов цепи к ее определителю при нулевых зависимых источниках напряжения. В результате получим

$$F=1+\frac{\sum\limits_{i=1}^{N}W_i\underset{Z}{\mathrm{Sim}}\left(\dfrac{\partial A}{\partial\beta_i},\ \dfrac{\partial A}{\partial\gamma_i}\right)-\sum\limits_{i<j}W_iW_j\underset{Z}{\mathrm{Sim}}\left(\dfrac{\partial^2 A}{\partial\beta_i\,\partial\beta_j},\ \dfrac{\partial^2 A}{\partial\gamma_i\,\partial\gamma_j}\right)+\ldots}{\underset{Z}{\det}A}. \tag{53}$$

Определим усиление цепи при $z_{\beta_N}\to\infty$.

Можно показать, что детерминантная функция блочной группы A имеет вид

$$\underset{Z}{\det}A=\underset{Z}{\det}\frac{\delta^N A}{\delta\beta_1\ldots\delta\beta_N}+z_{\beta_1}\ldots z_{\beta_N}\underset{Z}{\det}\frac{\partial^N A}{\partial\beta_1\ldots\partial\beta_N}.$$

В случае, если выполняется условие

$$\lim_{z_{\beta_1}\ldots z_{\beta_N}\to\infty}\frac{\underset{Z}{\mathrm{Sim}}\left(\dfrac{\partial A}{\partial\alpha},\ \dfrac{\partial A}{\partial\delta}\right)}{z_{\beta_1}z_{\beta_2}z_{\beta_3}\ldots z_{\beta_N}}=0, \tag{54}$$

формула усиления напряжения цепи запишется в виде

$$\left.\frac{U_\delta}{E}\right|_{z_{\beta_1}\ldots z_{\beta_N}\to\infty}=$$

$$=(-1)^N\frac{K_1K_2\ldots K_N\underset{Z}{\mathrm{Sim}}\left(\dfrac{\partial^{N+1}A}{\partial\alpha\,\partial\gamma_1\ldots\partial\gamma_N},\ \dfrac{\partial^{N+1}A}{\partial\beta_1\ldots\partial\beta_N\,\partial\delta}\right)}{\underset{Z}{\det}\dfrac{\partial^N A}{\partial\beta_1\ldots\partial\beta_N}+K_1K_2\ldots K_N\underset{Z}{\mathrm{Sim}}\left(\dfrac{\partial^N A}{\partial\gamma_1\ldots\partial\delta_N},\ \dfrac{\partial^N A}{\partial\beta_1\ldots\partial\beta_N}\right)}z_\delta. \tag{55}$$

Эта формула может быть записана в виде

$$k\Big|_{z_{\beta_1}\ldots z_{\beta_N}\to\infty}=\frac{\mathcal{K}}{1+\mathcal{K}\mathcal{B}}, \tag{56}$$

где

$$\mathcal{K}=\frac{(-1)^N K_1K_2\ldots K_N \operatorname{Sim}\limits_Z\left(\dfrac{\partial^{N+1}A}{\partial\alpha_1\ldots\partial\gamma_1\partial\gamma_N},\ \dfrac{\partial^{N+1}A}{\partial\beta_1\ldots\partial\beta_N\,\partial\delta}\right)}{\det\limits_Z\dfrac{\partial^N A}{\partial\beta_1\ldots\partial\beta_N}}\,z_\delta, \tag{57}$$

$$\mathcal{B}=\frac{(-1)^N \operatorname{Sim}\limits_Z\left(\dfrac{\partial^{N}A}{\partial\gamma_1\ldots\partial\gamma_N},\ \dfrac{\partial^{N}A}{\partial\beta_1\ldots\partial\beta_N}\right)}{\operatorname{Sim}\limits_Z\left(\dfrac{\partial^{N+1}A}{\partial\alpha\,\partial\gamma_1\ldots\partial\gamma_N},\ \dfrac{\partial^{N+1}A}{\partial\beta_1\ldots\partial\beta_N\,\partial\delta}\right)}\,\frac{1}{z_\delta}. \tag{58}$$

Утверждение 1. При выполнении условия (54) любую активную цепь с зависимыми источниками напряжения можно представить в виде цепи с одной обратной связью, причем передаточная функция $\mathcal{K}$ главной ветви и передаточная функция ветви обратной связи $\mathcal{B}$ определяются формулами (57) и (58) (рис. 20).

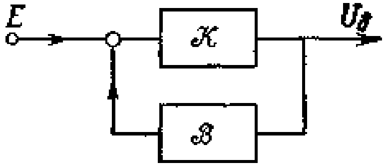


Рис. 20. Цепь с обратной связью.

В случае очень сильной обратной связи, т. е. когда $\mathcal{K}\to\infty$, имеем

$$k\Big|_{\mathcal{K}\to\infty}=\frac{1}{\mathcal{B}}. \tag{59}$$

Из этой формулы, известной из теории обратной связи, следует, что усиление цепи с очень сильной обратной связью не зависит от элементов главной ветви.

Пример 16. Определить усиление напряжения и возвратную разность для схемы с обратной связью (рис. 21, *а).*

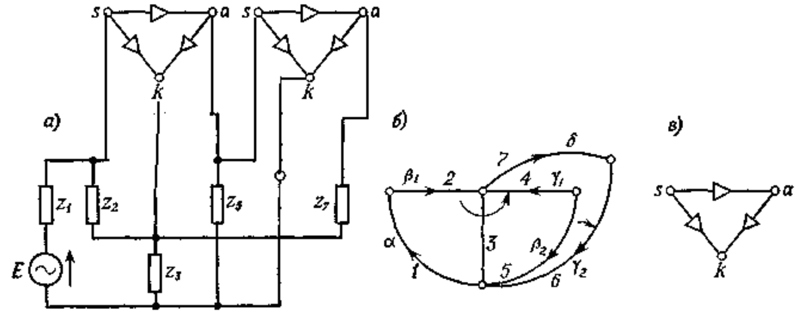


Рис. 21.

Решение. В этой цепи две активные связи (рис. 21, *б):* β_1 с γ_1, а также β_2 с γ_2. Выбираем ориентацию ветвей β и γ, учитывая то, что в схеме используются унисторные треугольники (рис. 21, в).

Рассчитаем блочную группу заданной цепи (граф цепи служит ее обратным изображением), которая представляет собой произведение трех блочных групп

$$P_1 = [1\ 2\ 3], \quad P_2 = [3\ 4\ 5], \quad P_3 = [3\ 6\ 7].$$

В результате получим

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 2 & 3 & 3 & 3 & 3 \\ 3 & 3 & 4 & 4 & 4 & 5 & 5 & 5 & 3 & 3 & 4 & 4 & 4 & 5 & 5 & 5 & 4 & 4 & 5 & 5 \\ 6 & 7 & 3 & 6 & 7 & 3 & 6 & 7 & 6 & 7 & 3 & 6 & 7 & 3 & 6 & 7 & 6 & 7 & 6 & 7 \end{bmatrix}.$$

Усиление напряжения цепи рассчитаем по формуле

$$K = \frac{\operatorname*{Sim}_{Z}\left(\frac{\partial A}{\partial \alpha}, \frac{\partial A}{\partial \delta}\right) - L_1^* W_1 - L_2^* W_2 + L_{12}^* W_1 W_2}{\Delta^0 + L_1 W_1 + L_2 W_2 - L_{12} W_1 W_2} z_\delta.$$

Имеем

$$\frac{\partial A}{\partial \alpha} = \frac{\partial A}{\partial 1} = \begin{bmatrix} 3 & 3 & 4 & 4 & 4 & 5 & 5 & 5 \\ 6 & 7 & 3 & 6 & 7 & 3 & 6 & 7 \end{bmatrix}, \qquad \frac{\partial A}{\partial \delta} = \frac{\partial A}{\partial 7} =$$

$$= \begin{bmatrix} 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 \\ 3 & 4 & 5 & 3 & 4 & 5 & 4 & 5 \end{bmatrix},$$

$$L_1 = \operatorname*{Sim}_{Z}\left(\frac{\partial A}{\partial \beta_1}, \frac{\partial A}{\partial \gamma_1}\right),$$

$$\frac{\partial A}{\partial \beta_1} = \frac{\partial A}{\partial 2} = \begin{bmatrix} 3 & 3 & 4 & 4 & 4 & 5 & 5 & 5 \\ 6 & 7 & 3 & 6 & 7 & 3 & 6 & 7 \end{bmatrix},$$

$$\frac{\partial A}{\partial \gamma_1} = \frac{\partial A}{\partial 4} = \begin{bmatrix} 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 \\ 3 & 6 & 7 & 3 & 6 & 7 & 6 & 7 \end{bmatrix},$$

$$L_2 = \operatorname*{Sim}_Z \left(\frac{\partial A}{\partial \beta_2}, \frac{\partial A}{\partial \gamma_2} \right),$$

$$\frac{\partial A}{\partial \beta_2} = \frac{\partial A}{\partial 5} = \begin{bmatrix} 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 \\ 3 & 6 & 7 & 3 & 6 & 7 & 6 & 7 \end{bmatrix},$$

$$\frac{\partial A}{\partial \gamma_2} = \frac{\partial A}{\partial 6} = \begin{bmatrix} 1 & 1 & 1 & 2 & 2 & 2 & 3 & 3 \\ 3 & 4 & 5 & 3 & 4 & 5 & 4 & 5 \end{bmatrix},$$

$$L_{12} = \operatorname*{Sim}_Z \left(\frac{\partial^2 A}{\partial \beta_1 \partial \beta_2}, \frac{\partial^2 A}{\partial \gamma_1 \partial \gamma_2} \right),$$

$$\frac{\partial^2 A}{\partial \beta_1 \partial \beta_2} = \frac{\partial^2 A}{\partial 2 \partial 5} = [3 \ \ 6 \ \ 7], \quad \frac{\partial^2 A}{\partial \gamma_1 \partial \gamma_2} = \frac{\partial^2 A}{\partial 4 \partial 6} = [1 \ \ 2 \ \ 3],$$

$$L^*_1 = \operatorname*{Sim}_Z \left(\frac{\partial^2 A}{\partial \alpha \partial \gamma_1}, \frac{\partial^2 A}{\partial \beta_1 \partial \delta} \right),$$

$$\frac{\partial^2 A}{\partial \alpha \partial \gamma_1} = \frac{\partial^2 A}{\partial 1 \partial 4} = [3 \ \ 6 \ \ 7], \quad \frac{\partial^2 A}{\partial \beta_1 \partial \delta} = \frac{\partial^2 A}{\partial 2 \partial 7} = [3 \ \ 4 \ \ 5],$$

$$L^*_2 = \operatorname*{Sim}_Z \left(\frac{\partial^2 A}{\partial \alpha \partial \gamma_2}, \frac{\partial^2 A}{\partial \beta_2 \partial \delta} \right),$$

$$\frac{\partial^2 A}{\partial \alpha \partial \gamma_2} = \frac{\partial^2 A}{\partial 1 \partial 6} = [3 \ \ 4 \ \ 5], \quad \frac{\partial^2 A}{\partial \beta_2 \partial \delta} = \frac{\partial^2 A}{\partial 5 \partial 7} = [1 \ \ 2 \ \ 3],$$

$$L^*_{12} = \operatorname*{Sim}_Z \left(\frac{\partial^3 A}{\partial \alpha \partial \gamma_1 \partial \gamma_2}, \frac{\partial^3 A}{\partial \delta \partial \beta_1 \partial \beta_3} \right),$$

$$\frac{\partial^3 A}{\partial \alpha \partial \gamma_1 \partial \gamma_2} = \frac{\partial^3 A}{\partial 1 \partial 4 \partial 6} = 1, \quad \frac{\partial^3 A}{\partial \delta \partial \beta_1 \partial \beta_2} = \frac{\partial^3 A}{\partial 7 \partial 2 \partial 5} = 1,$$

$$\operatorname*{Sim}_Z (1, \ 1) = 1,$$

$$W_1 = K_1 z_2, \ W_2 = K_2 z_5.$$

Подставив эти выражения в формулу усиления напряжения, получим

$$K=\frac{(z_3z_4+z_3z_5-K_1z_2z_3-K_1K_2z_2z_5)\,z_7}{\begin{gathered}z_1z_3z_6+z_1z_3z_7+z_1z_4z_3+z_1z_4z_6+z_1z_4z_7+z_1z_5z_3+z_1z_5z_6+\\+z_1z_5z_7+z_2z_3z_6+z_2z_3z_7+z_2z_4z_3+z_2z_4z_6+z_2z_4z_7+z_2z_5z_3+\\+z_2z_5z_3+z_2z_5z_7+z_3z_4z_6+z_3z_4z_7+z_3z_5z_6+z_3z_5z_7+\\+K_1z_2(z_3z_6+z_3z_7)+K_2z_5(z_1z_3+z_2z_3)+K_1K_2z_2z_3z_5\end{gathered}}.$$

Примечание. Знаки слагаемых функции совпадения можно найти, приняв следующее определение функции ориентации:
Ориентация ($\beta_1, \beta_2, \ldots, \beta_m, \gamma_1 \ldots \gamma_n$) =

$$=\begin{cases}+1, \textit{ если четное число ветвей } \beta_i,\ \gamma_i \textit{ не совпадает}\\ \textit{с положительным направлением контура;}\\ -1, \textit{ если нечетное число ветвей } \beta_i, \gamma_i \textit{ не совпа-}\\ \textit{дает с положительным направлением контура,}\end{cases}$$

согласно формулам (48).
Возвратная разность определяется по формуле $F = \Delta/\Delta^0$, где $\Delta^0 \det_Z A$, т. е.

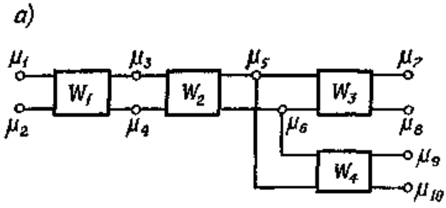
$$F=1+\frac{K_1z_2(z_3z_6+z_3z_7)+K_2(z_1z_3+z_2z_3)+K_1z_2z_3z_5}{z_1z_3z_6+z_1z_3z_7+\ldots+z_3z_5z_7}.$$

2.3. Анализ электрических модуль-схем методом блочных групп

2.3.1. Введение

Электрическая модуль-схема — это схема, составленная из многополюсников, где под *многополюсником* понимается цепь с несколькими выводами (например, двух-, трех-, четырехполюсник и т. д.), содержащая элементы с сосредоточенными или распределенными параметрами и не имеющая индуктивной связи с другими многополюсниками схемы. Многополюсник может иметь внутренние индуктивные связи. В данной работе будем рассматривать линейные электрические модуль-схемы.

На рис. 1. а изображена модуль-схема, состоящая из четырех четырехполюсников, т. е. модулей с четырьмя выводами W_1, W_2, W_3, W_4, и имеющая 10 узлов μ_1, μ_2, μ_3, …, μ_{10}.



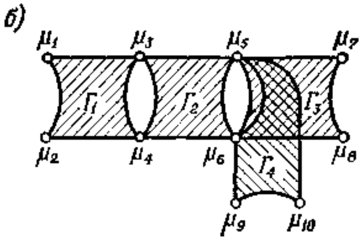


Рис. 1. Модуль-схема: *а)* схема цепи; *б)* модуль-граф цепи.

Считаем, что проводники, соединяющие отдельные блоки, не обладают сопротивлением и не имеют взаимных индуктивных и емкостных связей (в противном случае проводники нужно рассматривать как часть соответствующих блоков).

Граф модуль-схемы будем изображать в виде модуль-графа, отдельные модули которого соответствуют многополюсникам схемы. На рис. 1, *б* представлен модуль-граф для схемы, показанной на рис. 1, *а,* число модулей, и вершин которого равно соответственно числу многополюсников и узлов модуль-схемы, а число выводов — числу выводов соответствующего многополюсника схемы. Модуль-граф не имеет ребер, соединяющих выводы разных модулей, т. е. проводники, соединяющие выводы многополюсников схемы, представляются соответствующими вершинами модуль-графа. Заметим, что число контуров, образованных многополюсниками схемы, равно числу циклов скелета ее модуль-графа. В связи с этим напишем

$$M = \sum_{k=1}^{g} v_k - v - g + 1, \tag{1}$$

где M — число независимых контуров, образованных соединениями модуль-схемы, равное числу независимых циклов скелета ее модуль-графа; v_k — число выводов многополюсника W_k, равное числу вершин модуля Γ_k модуль-графа; g — число многополюсников модуль-схемы, равное числу модулей модуль-графа; v — число узлов модуль-графа, равное числу вершин модуль-графа.

Например, для модуль-схемы, представленной на рис. 1, *а,* а также для ее модуль-графа (рис. 1, б) имеем

$v_1 = 4$, $v_2 = 4$, $v_3 = 4$, $v_4 = 4$, $v = 10$, $g = 4$,

следовательно,

$$M = 16 — 10 — 4 + 1 = 3.$$

При рассмотрении модуль-схемы возможны два случая:
а) структура многополюсников (модулей) схемы известна;
б) структура многополюсников схемы неизвестна.

В первом случае модули модуль-графа можно заменить подграфами отдельных многополюсников схемы, после чего определить блочные группы A_i этих модулей.

Во втором случае можно воспользоваться только внешними параметрами отдельных многополюсников, например передаточными сопротивлениями или проводимостями, коэффициентами передачи напряжения, тока, а также напряжениями и токами зажимов и т. д. Этот случай более интересен по следующим причинам. Во-первых, при анализе цепей с известной структурой многополюсников, особенно когда в них встречаются элементы с распределенными параметрами и индуктивными связями, эти многополюсники можно рассматривать как части цепи, не определяя их схем замещения. Так, например, анализируя электрическую цепь с индуктивными связями, можно выделить в этой цепи элементы связи и рассматривать их как многополюсник (с неизвестной структурой), а затем анализировать цепь как модуль-схему. Во-вторых, при синтезе электрических схем можно рассматривать их как модуль-схемы, учитывая лишь внешние параметры отдельных многополюсников, т. е. не определяя их внутренней структуры, так как внешние параметры многополюсников определяют не один, а целый класс структур.

Модуль-схемы (неэлектрические) находят применение во многих областях: экономике, организации и т. д. Если, например, построить «хозяйственную» модель в виде модуль-схемы, отдельные модули которой представляют собой хозяйственные единицы, а связывающие их линии — пути взаимодействия этих единиц, то такую модель можно анализировать (не изучая структуры отдельных модулей), основываясь только на их внешних характеристиках, определяющих способ отклика хозяйственных единиц на внешние возмущения.

При анализе модуль-схем с неизвестной структурой модулей воспользуемся методом полных блочных групп.

2.3.2. Детерминантная функция модуль-схемы

Как известно, детерминантная функция блочной группы A равна определителю Δ матрицы узловых проводимостей схемы, граф которой служит геометрическим изображением блочной группы A. а проводимости ветвей образуют множество Y комплексных чисел, т. е.

$$\Delta = \det_Y A = \sum_{j=1}^{n} \prod_{i=1}^{m} y_{\alpha_{ij}}, \quad Y = \{y_1, y_2, ..., y_g\}. \qquad (2)$$

Аналогично определим *детерминантную функцию полной блочной группы* A

$$\det_Y \mathsf{A} = \sum_{j=1}^{n} \prod_{i=1}^{m} y_{\alpha_{ij}}, \quad Y = \{y_1, y_2, ..., y_g\} \qquad (3)$$

Например,

$$\det_Y \begin{Bmatrix} 1 & 2 & 2 & 1 \\ 2 & 3 & 1 & 1 \\ 3 & 4 & 3 & 4 \end{Bmatrix} = y_1y_2y_3 + y_2y_3y_4 + y_2y_1y_3 + y_1y_1y_4 =$$

$$= 2y_1y_2y_3 + y_2y_3y_4 + y_1^2y_4.$$

Заметим, что в отличие от детерминантной функции блочной группы A в детерминантную функцию полной блочной группы могут входить коэффициенты и показатели степени, большие 1.

На основании определения равенства блочных групп A и A можно написать

$$(\mathsf{A} \overset{s}{=} A) \Leftrightarrow (\det_Y \mathsf{A} = \det_Y A = \Delta), \quad Y = \{y_{1,} y_{2,}, ..., y_g\} \qquad (4)$$

Аналогичные определения и соотношения имеют место и для детерминантной функции дополнительной блочной группы A^d и полной блочной группы A^d. Эти функции определим на множестве комплексных чисел, равных импедансам ветвей цепи. Поэтому имеем

$$\Delta^d = \det_Z A^d = \sum_{j=1}^{n} \prod_{i=1}^{m} z_{\alpha_{ij}}, \quad Z = \{z_1, z_2, ..., z_g\}, \qquad (5)$$

$$(\mathsf{A} \overset{s}{=} A) \Leftrightarrow (\det_Z \mathsf{A}^d = \det_Z A^d = \Delta^d), \quad Y = \{y_{1,} y_{2,}, ..., y_g\}, \qquad (6)$$

где Δ^d — детерминантная функция дополнительной блочной группы A^d, равная определителю матрицы контурных сопротивлений электрической цепи, граф которой служит геометрическим изображением дополнительной блочной группы A^d, а импедансы ветвей образуют множество Z комплексных чисел.

Можно проверить, что

$$\frac{\Delta}{\Delta^d} = \frac{\det_Y A}{\det_Z A^d} = \prod_{i=1}^{g} y_i, \qquad (7)$$

$$\frac{\Delta^d}{\Delta} = \frac{\det\limits_{Z} A^d}{\det\limits_{Y} A} = \prod_{i=1}^{g} z_i, \qquad (8)$$

где $y_i = z^{-1}_i$ — адмитанс ветви α_i, g — число всех ветвей цени.

Детерминантную функцию блочной группы второго ранга ^2A можно определить двумя способами:

1. $$\det_Y {}^2A = \det_Y A, \quad A \overset{e}{=} {}^2A \;, \qquad (9)$$

2. $$\det_Y {}^2A = \det_{\Delta_{ij} \in \mathsf{D}} \mathsf{A}, \quad \mathsf{A} \overset{e}{=} {}^2A$$

$$\mathsf{D} = \{\Delta_{ij}\}, \quad \Delta_{ij} = \det_Y \mathsf{A}_{ij}, \qquad (10)$$

$$\mathsf{A}_{ij} \overset{e}{=} A_{ij}.$$

Это означает, что в первом случае блочную группу ^2A приводим к замещающей ее блочной группе A первого ранга, а затем определяем детерминантную функцию блочной группы A. Второй способ основан на определении полной блочной группы A, выраженной через полные блочные группы $\mathsf{A}_{ij} \overset{e}{=} A_{ij}$, где A_{ij} — элементы блочной группы ^2A, подстановке вместо блочных групп A_{ij} их детерминантных функций $\Delta_{ij} = \det\limits_Y \mathsf{A}_{ij}$ и проведении соответствующих алгебраических операций.

Вычисление детерминантной функции блочной группы второго ранга ^2A или дополнительной блочной группы $^2A^d$ подобно расчету определителя матриц проводимостей или сопротивлений модуль-схемы. Поясним это на примерах.

Пример 1. Найти определитель матрицы узловых проводимостей модуль-схемы (рис. 2, а), модуль-граф *Г* которой показан на рис. 2, *б,* а его скелет Γ_0 — на рис. 2, *в*.

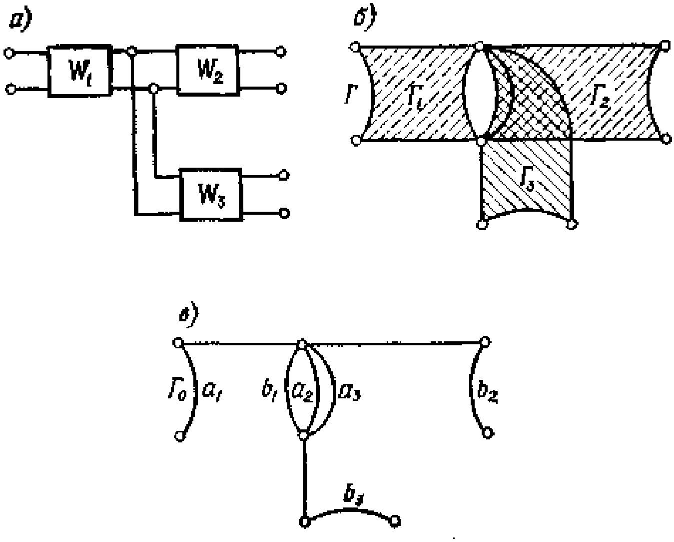


Рис. 2.

Дополнительная блочная группа A^d_0 скелета $Г_0$ равна

$$A_0^d = [b_1\ a_2]\,[b_1\ a_3] = \begin{bmatrix} b_1 & b_1 & a_2 \\ a_3 & a_2 & a_3 \end{bmatrix};$$

следовательно, блочная группа ${}^2A$ графа $Г$ запишется в виде

$$^2A = [A_0^d]\downarrow \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} A_{1b} & A_{1b} & A_1 \\ A_2 & A_{2a} & A_{2a} \\ A_{3a} & A_3 & A_{3a} \end{bmatrix}. \quad \text{(a)}$$

Если известна структура многополюсников (четырехполюсников) рассматриваемой схемы, то определитель матрицы проводимостей этой схемы можно рассчитать первым методом, приравнивая ${}^2A$ к ее замещающей блочной группе $A \overset{e}{=} {}^2A$.

Пусть, например (для .упрощения расчетов), все модули имеют одинаковую структуру (рис. 3, *а)*.

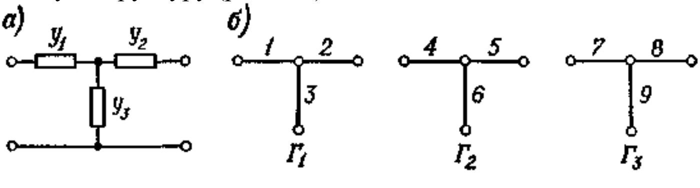


Рис. 3.

Графы отдельных модулей графа *Г* показаны на рис. 3, *б* (одинаковые обозначения ребер этих графов недопустимы). Блочные группы подграфов запишем в виде

$$A_1 = \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix}, \qquad A_2 = \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix}, \qquad A_3 = \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix},$$

$$A_{1b} = \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix}, \qquad A_{2a} = \begin{bmatrix} 4 & 5 \\ 5 & 6 \end{bmatrix}, \qquad A_{3a} = \begin{bmatrix} 7 & 8 \\ 8 & 9 \end{bmatrix}.$$

Подставляя блочные группы подграфов в выражение (а), для блочной группы ${}^2A$ имеем

$${}^2A = \begin{bmatrix} \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} & \begin{bmatrix} 1 & 1 \\ 2 & 3 \end{bmatrix} & \begin{bmatrix} 1 \\ 2 \\ 3 \end{bmatrix} \\ \begin{bmatrix} 4 \\ 5 \\ 6 \end{bmatrix} & \begin{bmatrix} 4 & 5 \\ 5 & 6 \end{bmatrix} & \begin{bmatrix} 4 & 5 \\ 5 & 6 \end{bmatrix} \\ \begin{bmatrix} 7 & 8 \\ 8 & 9 \end{bmatrix} & \begin{bmatrix} 7 \\ 8 \\ 9 \end{bmatrix} & \begin{bmatrix} 7 & 8 \\ 8 & 9 \end{bmatrix} \end{bmatrix} \overset{e}{=} A = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 2 & 2 & 3 & 3 & 2 & 2 & 3 & 3 & 2 & 2 & 2 & 2 \\ 4 & 4 & 4 & 4 & 4 & 5 & 4 & 5 & 3 & 3 & 3 & 3 \\ 5 & 5 & 5 & 5 & 5 & 6 & 5 & 6 & 4 & 4 & 5 & 5 \\ 6 & 6 & 6 & 6 & 7 & 7 & 7 & 7 & 5 & 5 & 6 & 6 \\ 7 & 8 & 7 & 8 & 8 & 8 & 8 & 8 & 7 & 8 & 7 & 8 \\ 8 & 9 & 8 & 9 & 9 & 9 & 9 & 9 & 8 & 9 & 8 & 9 \end{bmatrix}.$$

Таким образом, определитель матрицы проводимостей модуль-схемы равен

$$\Delta = \det_Y A = y_1 y_2 y_4 y_5 y_8 (y_6 y_7 + y_6 y_9 + y_7 y_9 + y_3 y_7 + y_3 y_9) +$$
$$+ y_1 y_3 y_5 y_6 y_8 (y_4 y_7 + y_4 y_9 + y_7 y_9 + y_2 y_7 + y_2 y_9) +$$
$$+ y_1 y_5 y_7 y_8 y_9 (y_2 y_6 + y_3 y_4).$$

Если адмитансы соответствующих ветвей всех многополюсников равны друг другу, т. е.

$$y_1 = y_4 = y_7, \qquad y_2 = y_5 = y_8, \qquad y_3 = y_6 = y_9,$$

то определитель схемы будет иметь вид

$$\Delta = y_1 y_2^2 y_3 (3y_1^2 y_2 + 4y_1 y_2 y_3 + 2y_1^2 y_3 + 2y_1 y_3^2 + y_2 y_3^2).$$

Замечания. Можно заметить, что для данной схемы проще найти определитель матрицы контурных сопротивлений схемы, так как дополнительная блочная группа равна

$$ {}^2A^d = [A_0^d]\downarrow \begin{bmatrix} A_1^d \\ A_2^d \\ A_3^d \end{bmatrix} = \begin{bmatrix} A_{1b}^d & A_{1b}^d & A_1^d \\ A_2^d & A_{2a}^d & A_{2a}^d \\ A_{3a}^d. & A_3^d & A_{3a}^d \end{bmatrix}, \qquad (б)$$

а дополнительные блочные группы отдельных подграфов

$$A^d_1 = 1, \qquad A^d_2 = 1, \qquad A^d_3 = 1,$$

$$A^d_{1b} = [2\ 3], \quad A^d_{2a} = [4\ 6], \quad A^d_{3a} = [7\ 9J.$$

Таким образом

$$ {}^2A^d = \begin{bmatrix} [2\ 3] & [2\ 3] & 1 \\ 1 & [4\ 6] & [4\ 6] \\ [7\ 9] & 1 & [7\ 9] \end{bmatrix} \overset{e}{=} A^d = \begin{bmatrix} 2 & 2 & 3 & 3 & 2 & 2 & 3 & 3 & 4 & 4 & 6 & 6 \\ 7 & 9 & 7 & 9 & 4 & 6 & 4 & 6 & 7 & 9 & 7 & 9 \end{bmatrix},$$

$$\Delta^d = (z_2 + z_3)(z_7 + z_9 + z_4 + z_6) + (z_4 + z_6)(z_7 + z_9).$$

Если

$$z_1 = z_4 = z_7, \qquad z_2 = z_5 = z_8, \qquad z_3 = z_6 = z_9,$$

то

$$\Delta^d = (z_1 + z_3)(z_1 + 2z_2 + 3z_3).$$

Полученные результаты можно проверить с помощью формулы

$$\frac{\Delta}{\Delta^d} = y_1^3 y_2^3 y_3^3$$

либо

$$\Delta = \Delta^d y_1^3 y_2^3 y_3^3$$

Если структура модулей интереса не представляет, то необходилю рассчитать полную блочную группу модуль-графа схемы. С этой целью для ранее найденной блочной группы ${}^2A$ [см. (а)] построим таблицу порядков

$$\mathbf{P} = \begin{pmatrix} 1 & 1 & 0 \\ 0 & 1 & 1 \\ 1 & 0 & 1 \end{pmatrix}.$$

Так как таблица не содержит одинаковых столбцов, то полная блочная группа равна

$$\mathrm{A} = \mathrm{A}_{1b}\mathrm{A}_2\mathrm{A}_{3a} + \mathrm{A}_{1b}\mathrm{A}_{2a}\mathrm{A}_3 + \mathrm{A}_b\mathrm{A}_{2a}\mathrm{A}_{3a}$$

Введем следующие обозначения:

$\Delta_1 = \det\limits_{Y} \mathrm{A}_1$ — определитель матрицы проводимостей многополюсника W_1;

$\Delta_2 = \det\limits_Y \mathsf{A}_2$ —определитель матрицы проводимостей многополюсника *W_2;*

$\Delta_3 = \det\limits_Y \mathsf{A}_3$ —определитель матрицы проводимостей многополюсника *W_3;*

$\Delta_{1b} = \det\limits_Y \mathsf{A}_{1b}$ — определитель матрицы проводимостей многополюсника W_1 с закороченным путем b_1 (замкнутого на выходе);

$\Delta_{2a} = \det\limits_Y \mathsf{A}_{2a}$ — определитель матрицы проводимостей многополюсника W_2 с закороченным путем a_2 (короткое замыкание входа);

$\Delta_{3a} = \det\limits_Y \mathsf{A}_{3a}$ — определитель матрицы проводимостей мпогополюсника W_3 с закороченным путем a_3 (короткое замыкание входа).

При этом определитель матрицы проводимостей рассматриваемой модуль-схемы равен

$$\Delta = \Delta_{1b}\Delta_2\Delta_{3a} + \Delta_{1b}\Delta_{2a}\Delta_3 + \Delta_1\Delta_{2a}\Delta_{3a}. \tag{11}$$

Так как дополнительная блочная группа ${}^2A^d$ графа $Г$ [выражение (б) строится аналогично блочной группы ${}^2A$ [выражение (а)], то определитель матрицы сопротивлений модуль-схемы запишем в виде

$$\Delta^d = \Delta^d{}_{1b}\Delta^d{}_{2}\Delta^d{}_{3a} + \Delta^d{}_{1b}\Delta^d{}_{2a}\Delta^d{}_{3} + \Delta^d{}_{1}\Delta^d{}_{2a}\Delta^d{}_{3a}. \tag{12}$$

Пример 2. Рассчитать определитель Δ матрицы проводимостей модуль-схемы (рис. 4, *а*). Модуль-граф $Г$ этой схемы показан на рис. 4, *б,* а его скелет $Г_0$ (сплошные линии) — на рис. 4. *в.*

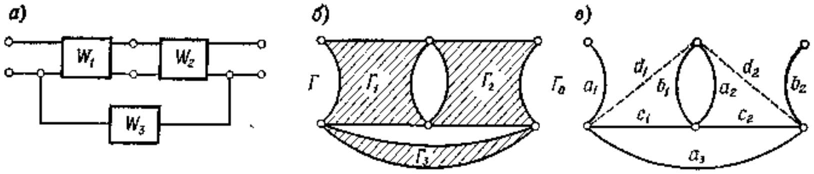


Рис. 4.

Расчет выполним без учета структуры отдельных многополюсников схемы.

Дополнительная блочная группа $\Delta^d{}_0$ скелета $Г_0$ равна

$$A_0^d = [b_1\ a_2]\,[c_1\ c_2\ a_3] = \begin{bmatrix} b_1 & b_1 & b_1 & a_2 & a_2 & a_2 \\ c_1 & c_2 & a_3 & c_1 & c_2 & a_3 \end{bmatrix},$$

а блочная группа ${}^2A$ графа Γ

$$ {}^2A = [A_0^d] \downarrow \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} A_{1bc} & A_{1b} & A_{1b} & A_{1c} & A_1 & A_1 \\ A_2 & A_{2c} & A_2 & A_{2a} & A_{2ac} & A_{2a} \\ A_3 & A_3 & A_{3a} & A_3 & A_3 & A_{3a} \end{bmatrix} . \qquad (a) $$

Таблица порядков **P** имеет следующий вид:

$$ \mathbf{P} = \begin{array}{c} \begin{pmatrix} 2 & 1 & 1 & 1 & 0 & 0 \\ 0 & 1 & 0 & 1 & 2 & 1 \\ 0 & 0 & 1 & 0 & 0 & 1 \end{pmatrix} . \\ \begin{array}{cccccc} & X & & X & & \end{array} \end{array} $$

В таблице имеются два одинаковых столбца (обозначенные внизу *X),* поэтому следует определить дефект суммы столбцов блочной группы

$$ {}^2A^{\mathrm{I}} = A_3 \begin{bmatrix} A_{1b} & A_{1c} \\ A_{2c} & A_{2a} \end{bmatrix} . $$

Для этого блоки Γ_1 и Γ_2 будем считать трехполюсниками (так как они инцидентны остальным модулям только в трех вершинах) и в скелете Γ_0 (рис. 4. *в)* добавим ребра d_1 и d_2 (пунктир). Ребра b_1, c_1 и d_i образуют замещающий граф (полный) для модуля Γ_1, а ребра a_2, c_2 и d_2 — замещающий граф (полный) для модуля Γ_2. Следовательно,

$$ A_{1b} = [c_{1(1)}\ d_{1(1)}], \quad A_{1c} = [b_{1(1)}\ d_{1(1)}], $$

$$ A_{2a} = [c_{2(1)}\ d_{2(1)}], \quad A_{2c} = [a_{2(1)}\ d_{2(1)}] $$

и далее

$$ \begin{bmatrix} A_{1b} & A_{1c} \\ A_{2c} & A_{2a} \end{bmatrix} \stackrel{s}{=} \mathsf{A}_{1b}\mathsf{A}_{2c} + \mathsf{A}_{1c}\mathsf{A}_{2a} - 2\begin{Bmatrix} d_{1(1)} \\ d_{2(1)} \end{Bmatrix} $$

где $\mathsf{A}_{1b} \stackrel{s}{=} A_{1b}$, $\mathsf{A}_{2c} \stackrel{s}{=} A_{2c}$ и т. д., а дендритные веса имеют вид

$$ d_{1(1)} = \frac{1}{2}(\mathsf{A}_{1b} + \mathsf{A}_{1c} - \mathsf{A}_{1d}), $$

$$ d_{2(1)} = \frac{1}{2}(\mathsf{A}_{2a} + \mathsf{A}_{2c} - \mathsf{A}_{2d}). $$

Подставив эти выражения в (а) и выполнив необходимые операции, получим

$$ \mathsf{A} = \mathsf{A}_3(\mathsf{A}_{1bc}\mathsf{A}_2 + \mathsf{A}_{1b}\mathsf{A}_{2c} + \mathsf{A}_{1c}\mathsf{A}_{2a} + \mathsf{A}_1\mathsf{A}_{2ac}) + $$

$$ + \mathsf{A}_{3a}(\mathsf{A}_{1b}\mathsf{A}_2 + \mathsf{A}_1\mathsf{A}_{2a}) - \frac{1}{2}(\mathsf{A}_{1b} + \mathsf{A}_{1c} - \mathsf{A}_{1d})\,(\mathsf{A}_{2a} + \mathsf{A}_{2c} - \mathsf{A}_{2d}) + $$

Таким образом, определитель Δ матрицы проводимостей модуль-схемы (рис. 4, а) имеет вид

$$\Delta = \det_{\mathscr{D}} \mathscr{A} = \Delta_3 (\Delta_{1bc}\Delta_2 + \Delta_{1b}\Delta_{2c} + \Delta_{1c}\Delta_{2a} + \Delta_1\Delta_{2ac}) +$$

$$+ \Delta_{3a} (\Delta_{1b}\Delta_2 + \Delta_1\Delta_{2a}) - \frac{1}{2} (\Delta_{1b} + \Delta_{1c} - \Delta_{1d}) (\Delta_{2a} + \Delta_{2c} - \Delta_{2d}). \quad (13)$$

Смысл отдельных символов в выражении (13) показан на рис. 5, где представлены многополюсники W_1, W_2 и W_3 рассматриваемой цепи при различных режимах на зажимах, а также даны символы определителей матрицы узловых проводимостей.

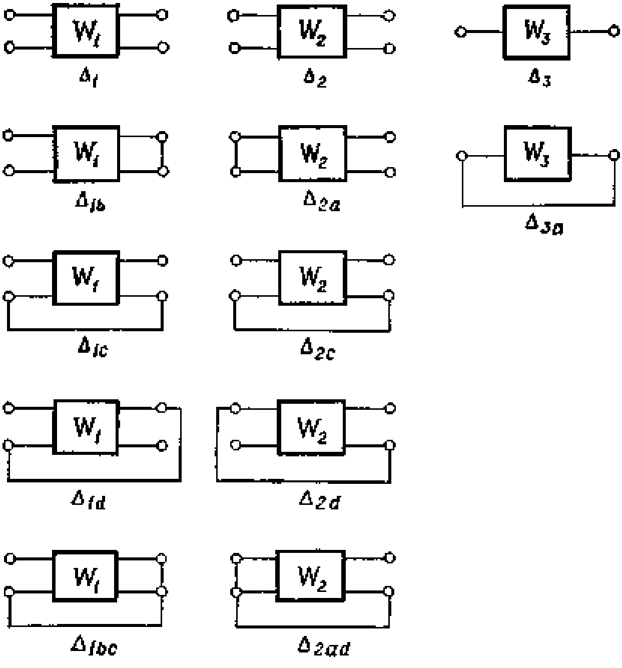


Рис. 5.

Для определителя матрицы контурных сопротивлений рассматриваемой модуль-схемы имеет место формула, аналогичная (13).

Замечания. Чтобы перейти от формулы

$$A_{ab} = \frac{1}{A} \left\{ A_a A_b - \frac{1}{4} (A_a + A_b - A_d)^2 \right\}$$

к (13), нужно исключить определители Δ_{1bc} и Δ_{2ac}.

2.3.3. Входной импеданс модуль-схемы

Для расчета *импеданса* $Z_{\mu_1\mu_2}$ или *адмитанса* $Y_{\mu_1\mu_2}$ между узлами μ_1 и μ_2 модуль-схемы применяем известную формулу

$$Z_{\mu_1\mu_2} = \frac{\Delta_{\mu_1\mu_2}}{\Delta} = \frac{\det\limits_{Y} A_{\mu_1\mu_2}}{\det\limits_{Y} A} = \frac{1}{Y_{\mu_1\mu_2}}, \tag{14}$$

где Δ — определитель матрицы узловых проводимостей модуль-схемы;

$\Delta_{\mu_1\mu_2}$ — определитель матрицы узловых проводимостей модуль-схемы с закороченными узлами μ_1 и μ_2;

$A \overset{e}{=} {}^2A$ — замещающая блочная группа для блочной группы второго ранга ${}^2A$ модуль-графа рассматриваемой модуль-схемы;

$A_{\mu_1\mu_2} \overset{e}{=} {}^2A_{\mu_1\mu_2}$ — замещающая блочная группа для блочной группы второго ранга ${}^2A_{\mu_1\mu_2}$ модуль-графа с закороченными вершинами μ_1 и μ_2. Как известно [формула (10)],

$$\det_{Y} {}^2A = \det_{\Delta_{ij}\in \mathsf{D}} \mathsf{A}, \quad \mathsf{A} \overset{e}{=} {}^2A \ \mathsf{D} = \{\Delta_{ij}\}, \quad \Delta_{ij} = \det_{Y} \mathsf{A}_{ij}, \qquad \mathsf{A}_{ij} \overset{e}{=} A_{ij}.$$

а также

$$\det_{Y} A_{\mu_1\mu_2} = \det_{\mathsf{D}} \mathsf{A}_{\mu_1\mu_2}, \quad \mathsf{A}_{\mu_1\mu_2} \overset{s}{=} {}^2A_{\mu_1\mu_2}$$

Поэтому формулу (14) можно записать в другом виде:

$$Z_{\mu_1\mu_2} = \frac{\Delta_{\mu_1\mu_2}}{\Delta} = \frac{\det\limits_{\mathsf{D}} \mathsf{A}_{\mu_1\mu_2}}{\det\limits_{\mathsf{D}} \mathsf{A}} = \frac{1}{Y_{\mu_1\mu_2}}, \tag{15}$$

где

$$\mathsf{D} = \{\Delta_{ij}\}, \quad \Delta_{ij} = \det_{Y} \mathsf{A}_{ij}, \ \mathsf{A}_{ij} \overset{e}{=} A_{ij}.$$

Импеданс $Z_{\mu_1\mu_2}$ или адмитанс $Y_{\mu_1\mu_2}$ можно также выразить с помощью определителей матрицы контурных сопротивлений модуль-схемы, т. е.

$$Z_{\mu_1\mu_2} = \frac{\Delta^d}{\Delta^d_{\mu_1\mu_2}} = \frac{1}{Y_{\mu_1\mu_2}}. \tag{16}$$

Для определения входного импеданса модуль-схемы с помощью входных импедансов отдельных многополюсников перепишем выражение (14) в следующем виде:

$$Z_{\mu_1\mu_2}=\left(\frac{\Delta_{\mu_1\mu_2}}{\prod\limits_{i=1}^{g}\Delta_i}\right):\left(\frac{\Delta}{\prod\limits_{i=1}^{g}\Delta_i}\right), \tag{17}$$

где Δ_1, Δ_2, . . ., Δ_q — определители матриц проводимостей многополюсников W_1, W_2, . . ., W_q модуль-схемы.

Равенство (17) позволяет выразить импеданс $Z_{\mu_1\mu_2}$ модуль-схемы через входные импедансы его отдельных многополюсников, не учитывая при этом их внутренней структуры. Следовательно, в этих многополюсниках могут встречаться элементы с распределенными и сосредоточенными параметрами, а также индуктивные связи.

При использовании формулы (17) следует учитывать, что

$$\frac{\Delta_{abcd\,\ldots\,k}}{\Delta}=\frac{\Delta_{abcd\,\ldots\,k}}{\Delta_{bcd\,\ldots\,k}}\,\frac{\Delta_{bcd\,\ldots\,k}}{\Delta_{cd\,\ldots\,k}}\,\frac{\Delta_{cd\,\ldots\,k}}{\Delta_{d\,\ldots\,k}}\cdots\frac{\Delta_k}{\Delta}=$$

$$=Z_{abcd\,\ldots\,k}Z_{bcd\,\ldots\,k}Z_{cd\,\ldots\,k}\cdots Z_k. \tag{18}$$

(по этой формуле можно рассчитать определитель матрицы полных проводимостей электрической цепи. Если каждую ветвь этой цепи считать путем, то

$$\Delta_{abcd\cdots k}=1,$$

следовательно,

$$\Delta= Y_{abed...k}Y_{bed...k}Y_{cd...\,k}\ldots Y_k,$$

где, например, $Y_{abcd\cdots k}$ — адмитаис цепи с закороченными путями b, c, d, . . ., k, измеренный между коночными узлами пути a.)

где Δ — определитель матрицы проводимостей электрической цепи; $\Delta_{abcd\cdots ft}$ —определитель матрицы проводимостей цепи с замкнутыми путями a, b, c, d, . . ., k (короткое замыкание пути a означает замыкание коночных узлов пути a);

$Z_{ahcd\cdots k}$ — импеданс электрической цепи с замкнутыми путями b, c, d, . . ., k, измеренный между конечными узлами пути a;

$Z_{bcd...k}$ — импеданс электрической цепи с замкнутыми путями c, d. . ., k, измеренный между конечными узлами пути b;

Z_k — импеданс электрической цепи, измеренный между конечными узлами пути k.

Метод расчета импеданса $Z_{\mu_1\mu_2}$ модуль-схемы без учета внутренней структуры ее многополюсников иллюстрируют следующие примеры.

Пример 3. Рассчитать входной импеданс $Z_{\mu_1\mu_2} = Z_{a1}$ модуль-схемы (рис. 6, *а)*, модуль-граф *Г* которой показан на рис. 6, *б,* а скелет $Г_0$ этого графа — на рис. 6, *в.*

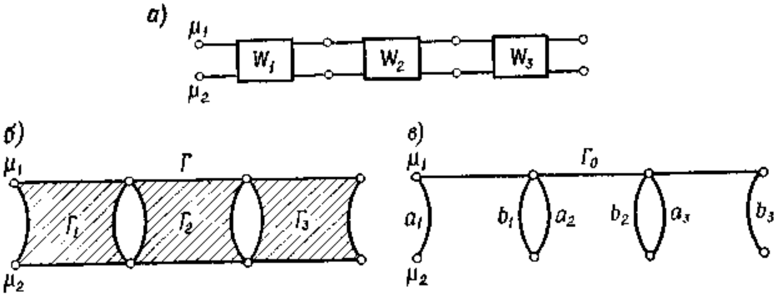


Рис. 6.

Дополнительная блочная группа A^d_0 скелета $Г_0$ равна

$$A_0^d = [b_1\ a_2][b_2\ a_3] = \begin{bmatrix} b_1 & b_1 & a_2 & a_2 \\ b_2 & a_3 & b_2 & a_3 \end{bmatrix}$$

Блочная группа ^2A графа *Г* имеет вид

$$^2A = [A_0^d] \downarrow \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} A_{1b} & A_{1b} & A_1 & A_1 \\ A_{2b} & A_2 & A_{2ab} & A_{2a} \\ A_3 & A_{3a} & A_3 & A_{3a} \end{bmatrix}, \qquad (а)$$

а блочная группа $^2A_{a_1}$ графа с замкнутым путем a_1 —

$$^2A_1 = [a_1] \downarrow [^2A] = \begin{bmatrix} A_{1ab} & A_{1ab} & A_{1a} & A_{1a} \\ A_{2b} & A_2 & A_{2ab} & A_{2a} \\ A_3 & A_{3a} & A_3 & A_{3a} \end{bmatrix}. \qquad (б)$$

Таблицы порядков блочных групп ^2A и $^2A_{a_1}$ запишем в виде

$$\mathbf{P} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}, \qquad \mathbf{P}_{a_1} = \begin{pmatrix} 2 & 2 & 1 & 1 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}.$$

Так как никакая из этих таблиц не имеет повторяющихся столбцов, полные блочные группы A и A_{a_1} выражаются следующим образом:

$$A = A_{1b}(A_{2b}A_3 + A_2A_{3a}) + A_1(A_{2ab}A_3 + A_{2a}A_{3a}),$$
$$A_{a_1} = A_{1ab}(A_{2b}A_3 + A_2A_{3a}) + A_{1a}(A_{2ab}A_3 + A_{2a}A_{3a}).$$

Определители Δ и Δ_{a_1} матриц проводимостей модуль-схемы с неамкнутым и замкнутым путем a_1 соответственно равны

$$\Delta = \Delta_{1b}(\Delta_{2b}\Delta_3 + \Delta_2\Delta_{3a}) + \Delta_1(\Delta_{2ab}\Delta_3 + \Delta_{2a}\Delta_{3a}),$$
$$\Delta_{a_1} = \Delta_{1ab}\cdot(\Delta_{2b}\Delta_3 + \Delta_2\Delta_{3a}) + \Delta_{1a}(\Delta_{2ab}\Delta_3 + \Delta_{2a}\Delta_{3a}).$$

После деления этих определителей на произведение определителей матриц проводимостей всех многополюсников схемы получаем:

$$\frac{\Delta}{\Delta_1\Delta_2\Delta_3} = \frac{\Delta_{1b}}{\Delta_1}\left(\frac{\Delta_{2b}\Delta_3}{\Delta_2\Delta_3} + \frac{\Delta_2}{\Delta_2}\,\frac{\Delta_{3a}}{\Delta_3}\right) + \frac{\Delta_1}{\Delta_1}\left(\frac{\Delta_{2ab}}{\Delta_{2b}}\,\frac{\Delta_{2b}}{\Delta_2}\,\frac{\Delta_3}{\Delta_3} + \frac{\Delta_{3a}}{\Delta_3}\right) =$$
$$= Z_{1b}(Z_{2b} + Z_{3a}) + Z_{2ab}Z_{2b} + Z_{2a}Z_{3a}.$$
$$\frac{\Delta_{a1}}{\Delta_1\Delta_2\Delta_3} = \frac{\Delta_{1ab}}{\Delta_{1b}}\,\frac{\Delta_{1b}}{\Delta_1}\left(\frac{\Delta_{2b}}{\Delta_2}\,\frac{\Delta_3}{\Delta_3} + \frac{\Delta_2}{\Delta_2}\,\frac{\Delta_{3a}}{\Delta_3}\right) + \frac{\Delta_{1a}}{\Delta_1}\left(\frac{\Delta_{2ab}}{\Delta_{2b}}\,\frac{\Delta_{2b}}{\Delta_2}\,\frac{\Delta_3}{\Delta_3} + \frac{\Delta_{2a}}{\Delta_2}\,\frac{\Delta_{3a}}{\Delta_3}\right) =$$
$$= Z_{1ab}Z_{1b}(Z_{2b} + Z_{3a}) + Z_{1a}(Z_{2ab}Z_{2b} + Z_{2a}Z_{3a}).$$

Следовательно, входной импеданс $Z_{\mu_1\mu_2}$ модуль-схемы (рис. 6, а) равен

$$Z_{\mu_1\mu_2} = \left(\frac{\Delta_{\mu_1\mu_2}}{\Delta_1\Delta_2\Delta_3}\right) : \left(\frac{\Delta}{\Delta_1\Delta_2\Delta_3}\right) =$$
$$= \frac{Z_{1ab}Z_{1b}(Z_{2b} + Z_{3a}) + Z_{1a}(Z_{2ab}Z_{2b} + Z_{2a}Z_{3a})}{Z_{1b}(Z_{2b} + Z_{3a}) + Z_{2ab}Z_{2b} + Z_{2a}Z_{3a}}, \tag{19}$$

где Z_{1a}, Z_{2a}, Z_{3a} — входные импедансы четырехполюсников W_1, W_2, W_3; Z_{1b}, Z_{2b} — выходные импедансы четырехполюсников W_1, W_2; Z_{1ab}, Z_{2ab} — входные импедансы четырехполюсников W_1, W_2 с замкнутыми выходами.

Пример 4. Рассчитать входной импеданс $Z_{\mu_1\mu_2}$ схемы (рис. 7, *а)*, модуль-граф *Г* которой показан на рис. 6.7, б, а его скелет $Г_0$ — на рис. 6.7, *в*.

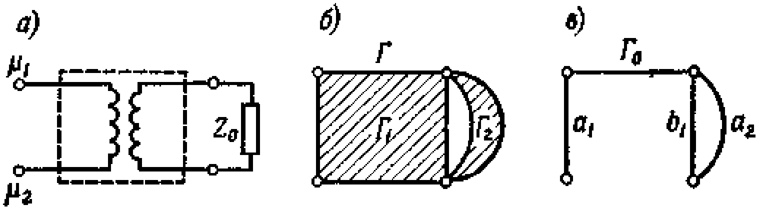


Рис. 7.

В этой схеме трансформатор рассматривается как четырехполюсник.
Дополнительная блочная группа A^d_0 скелета Γ_0 графа равна

$$A^d_0 = [b_1 a_2];$$

следовательно, блочную группу ${}^2A$ графа Γ можно записать в виде

$${}^2A = [A^d_0] \downarrow \begin{bmatrix} A_1 \\ A_2 \end{bmatrix} = \begin{bmatrix} A_{1b} & A_1 \\ A_2 & A_{2a} \end{bmatrix}$$

а блочную группу ${}^2A_{a_1}$ графа Γ_{a_1} с замкнутым путем a_1 в виде

$${}^2A_{a_1} = [a_1] \downarrow \left[{}^2A\right] = \begin{bmatrix} A_{1ab} & A_{1a} \\ A_2 & A_{2a} \end{bmatrix}$$

Так как таблицы порядков обоих блочных групп не содержат повторяющихся столбцов, то

$$\Delta = \Delta_{1b}\Delta_2 + \Delta_1\Delta_{2a},$$
$$\Delta_{a1} = \Delta_{1ab}\Delta_2 + \Delta_{1a}\Delta_{2a},$$

откуда

$$Z_{\mu_1\mu_2} = \frac{\frac{\Delta_{1ab}}{\Delta_{1b}}\frac{\Delta_{1b}}{\Delta_1} + \frac{\Delta_{1a}}{\Delta_1}\frac{\Delta_{2a}}{\Delta_2}}{\frac{\Delta_{1b}}{\Delta_1} + \frac{\Delta_{2a}}{\Delta_2}} = \frac{Z_{1ab}Z_{1b} + Z_{1a}Z_0}{Z_{1b} + Z_0}, \qquad (20)$$

где Z_{1ab}, — входной импеданс трансформатора при коротком замыкании выхода;
Z_{1a} — входной импеданс трансформатора в режиме холостого хода;
Z_{1b} — выходной импеданс трансформатора в режиме холостого хода на входе;
Z_0 — импеданс нагрузки на выходе трансформатора.
Такой же результат можно получить, используя формулу (19), если принять, что $Z_{3a} \to \infty$, $Z_{2a} = Z_{2b} = Z_0$, $Z_{2ab} = 0$.

Пример 5. Рассчитать импеданс $Z_{\mu_1\mu_3}$ между полюсами μ_1 и μ_3 модуль-схемы (рис. 8, а), модуль-граф Γ которой показан на рис. 8, *б,* а его скелет Γ_0 — на рис. 8, *в* (сплошные линии).

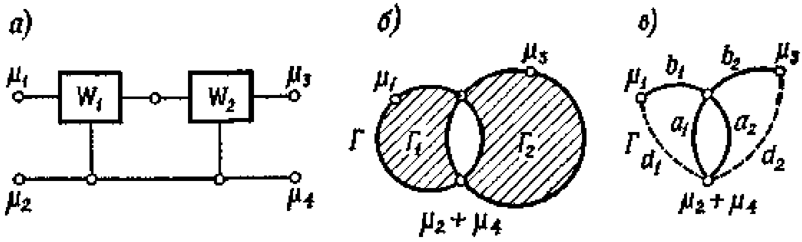


Рис. 8.

Блочная группа ${}^2A$ графа $Г$ равна

$$ {}^2A=[A_0^d]\downarrow\begin{bmatrix}A_1\\A_2\end{bmatrix}=[a_1\ a_2]\downarrow\begin{bmatrix}A_1\\A_2\end{bmatrix}=\begin{bmatrix}A_{1a}\ A_1\\A_2\ \ A_{2a}\end{bmatrix} $$

Блочная группа ${}^2A_{\mu_1\mu_3}$ графа $Г_{\mu_1\mu_3}$, образованного в результате замыкания узлов μ_1 и μ_3 графа Г, имеет вид

$$ {}^2A_{\mu_1\mu_3}=[b_1\ b_2]\downarrow\begin{bmatrix}A_{1a}\ A_1\\A_2\ \ A_{2a}\end{bmatrix}=\begin{bmatrix}A_{1ab}\ A_{1b}A_{1a}\ A_1\\A_2\ \ A_{2a}\ A_{2b}\ A_{2ab}\end{bmatrix} $$

Таблица порядков блочной группы ${}^2A$ не содержит повторяющихся столбцов, но таблица порядков блочной группы ${}^2A_{\mu_1\mu_3}$

$$ \mathbf{P}=\begin{array}{c}\begin{pmatrix}2&1&1&0\\0&1&1&2\end{pmatrix}\\ \quad\ X\ \ X\end{array} $$

имеет два одинаковых столбца (обозначенные *X);* поэтому дополняем скелет $Г_0$ ребрами d_1 и d_2 (указаны на рис. 8, *в* пунктиром). В результате

$$ A_{1b}=[a_{1(1)}\ d_{1(1)}],\qquad A_{2b}=[a_{2(1)}\ d_{2(1)}], $$
$$ A_{1a}=[b_{1(1)}\ d_{2(1)}],\qquad A_{2a}=[b_{2(1)}\ d_{2(1)}]. $$

Блочную группу запишем в виде

$$ {}^2A^I=\begin{bmatrix}A_{1b}\ A_{1a}\\A_{2a}\ A_{2b}\end{bmatrix}^s=\mathrm{A_{1b}A_{2a}}+\mathrm{A_{1a}A}_{2b}-2\begin{Bmatrix}d_{1(1)}\\d_{2(1)}\end{Bmatrix}, $$

где

$$d_{1(1)} = \frac{1}{2}(А_{1a} + А_{1b} - А_{1d}),$$

$$d_{2(2)} = \frac{1}{2}(А_{2a} + А_{2b} - А_{2d}).$$

Следовательно,

$$\Delta = \Delta_{1a}\Delta_2 + \Delta_1\Delta_{2a},$$

$$\Delta_{\mu_1\mu_3} = \Delta_{1ab}\Delta_2 + \Delta_{1b}\Delta_{2a} + \Delta_{1a}\Delta_{2b} + \Delta_1\Delta_{2ab} -$$

$$-\frac{1}{2}(\Delta_{1a} + \Delta_{1b} - \Delta_{1d})(\Delta_{2a} + \Delta_{2b} - \Delta_{2d}).$$

Таким образом, импеданс схемы $Z_{\mu_1\mu_3}$

$$Z_{\mu_1\mu_3} = \left(\frac{\Delta_{\mu_1\mu_3}}{\Delta_1\Delta_2}\right) : \left(\frac{\Delta}{\Delta_1\Delta_2}\right) =$$

$$= \frac{Z_{1ab}Z_{1b} + Z_{1b}Z_{2a} + Z_{1a}Z_{2b} + Z_{2ab}Z_{2b} - \frac{1}{2}(Z_{1a} + Z_{1b} - Z_{1d})(Z_{2a} + Z_{2b} - Z_{2d})}{Z_{1a} + Z_{2a}}. \tag{21}$$

Обозначения отдельных импедансов трехполюсников W_1 и W_2 в выражении (21) указаны на рис. 9.

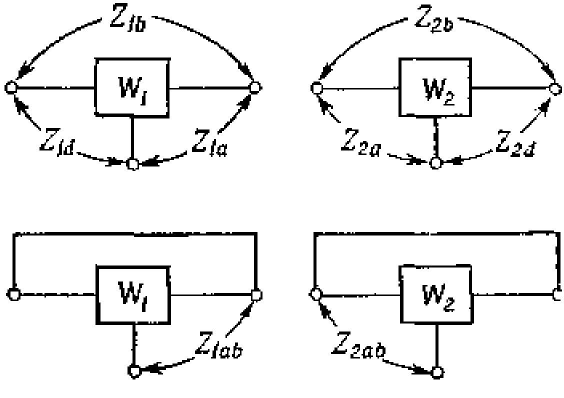


Рис. 9.

Пример 6. Рассчитать входной импеданс цепной схемы с бесконечно большим числом одинаковых четырехполюсников.

Рассмотрим цепную схему с n-одинаковыми четырехполюсниками (рис. 10, *а*). Скелет $Г_0$ модуль-графа $Г$ показан на рис. 10, *б*.

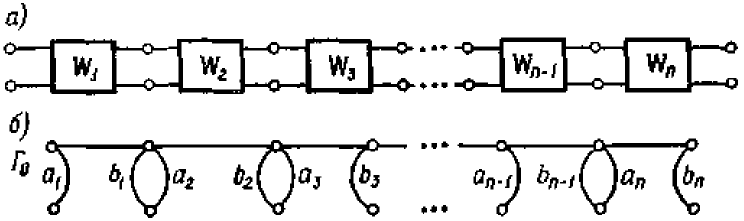


Рис. 10.

Обозначим $A^{(1)}$ — блочную группу первого модуля графа $Г$, а A_{n-t} — блочную группу графа $Г$ без первого модуля.
Блочная группа ^2A_n полного графа равна

$$^2A_n = [b_1 a_2] \downarrow \begin{bmatrix} A_{n-1} \\ A^{(1)} \end{bmatrix} = \begin{bmatrix} A_{n-1} & A_{n-1a_2} \\ A^{(1)}_{b_1} & A^{(1)} \end{bmatrix}.$$

Таблица порядков этой блочной группы второго ранга не содержит одинаковых столбцов; следовательно, определитель матрицы полных проводимостей Δ_n рассматриваемой схемы имеет вид

$$\Delta_n = \Delta_{n-1}\Delta_b + \Delta_{n-1a}\Delta,$$

где Δ_{n-1} — определитель матрицы проводимостей цепной схемы с n — 1 одинаковыми четырехполюсниками; Δ_{n-1a} — определитель матрицы проводимостей схемы с n — 1 четырехполюсниками; Δ — определитель матрицы проводимостей каждого четырехполюсника схемы; Δ_b — определитель матрицы проводимостей каждого четырехполюсника с замкнутым выходом.
Если первый четырехполюсник схемы замкнут на входе, то из приведенного уравнения следует формула для определителя Δ_{na} матрицы проводимостей замкнутой на входе схемы с n модулями

$$\Delta_{na} = \Delta_{n-1}\Delta_{ab} + \Delta_{n-1a}\Delta_a.$$

Напишем входной импеданс Z_{na} схемы

$$Z_{na} = \frac{\Delta_{na}}{\Delta_n} = \frac{\frac{\Delta_{ab}}{\Delta} + \frac{\Delta_{n-1a}}{\Delta_{n-1}} \cdot \frac{\Delta_a}{\Delta}}{\frac{\Delta_b}{\Delta} + \frac{\Delta_{n-1a}}{\Delta_{n-1}}} = \frac{Z_{ab}Z_b + Z_{n-1a}Z_a}{Z_b + Z_{n-1a}}, \qquad (22)$$

где Z_a — входной импеданс четырехполюсника;
Z_b — выходной импеданс четырехполюсника;
Z_{ab} — входной импеданс замкнутого на выходе четырехполюсника;
Z_{n-1a} — входной импеданс цепной схемы n — 1 с одинаковыми модулями (четырехполюсниками).
Для $n \to \infty$ $Z_{n-1a} = Z_{na}$ и, следовательно,

$$Z_{\infty a} = \frac{Z_{ab}Z_b + Z_{\infty a}Z_a}{Z_b + Z_{\infty a}}$$

Решая это квадратное уравнение, получим

$$Z_{\infty a} = \tfrac{1}{2}\{Z_a - Z_b + [(Z_a - Z_b)^2 + 4Z_{ab}Z_b]^{1/2}\}. \qquad (23)$$

2.3.4. Коэффициент передачи напряжения модуль-схемы

Рассмотрим четырехполюсник, питаемый от идеального источника напряжения E (рис. 11).

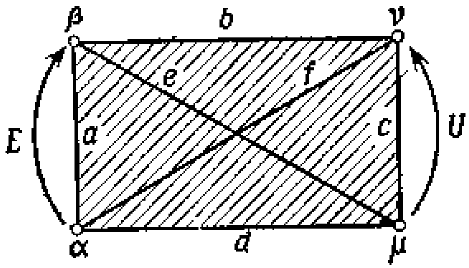


Рис. 11. Четырехполюсник, питаемый от идеального источника напряжения.

Здесь отрезки a, b, c, d, e, f представляют собой произвольные пути четырехполюсника, соединяющие отдельные пары его узлов, а соответствующие пути модуль-графа $Г$ четырехполюсника условимся обозначать аналогичным образом. Заметим, что идеальный источник напряжения E содержит узлы α и β четырехполюсника. Обозначим: A — блочная группа графа рассматриваемой схемы при короткой замыкании зажимов источника напряжения, Δ — детерминантная функция блочной группы A (равная определителю матрицы узловых проводимостей рассматриваемой схемы).

На основании формулы Персиваля коэффициент передачи напряжения K_{ac} четырехполюсника равен

$$K_{ac} = \frac{U}{E} = \frac{Z_e + Z_f - Z_b - Z_d}{2Z_a} \qquad (24)$$

где Z_a, Z_b, . . ., Z_f — импедансы четырехполюсника, измеренные между конечными узлами соответствующих путей. Подставив в выражение (24)

$$Z_a = \frac{\Delta_a}{\Delta}, \quad Z_b = \frac{\Delta_b}{\Delta},$$

где $\Delta_a = \det_Y A_a = \det_Y \partial A/\partial a, \quad \Delta_b = \det_Y A_b = \det_Y \partial A/\partial b$ и т. д., получим у

$$K_{ac} = \frac{\Delta_e + \Delta_f - \Delta_b - \Delta_d}{2\Delta_a} \tag{25}$$

или

$$K_{ac} = \frac{\det_Y(\mathsf{A}_e + \mathsf{A}_f) - \det_Y(\mathsf{A}_b + \mathsf{A}_d)}{2\det_Y \mathsf{A}_a}, \tag{26}$$

где $A_a \overset{s}{=} A_a, A_b \overset{s}{=} A_b$ и т. д.

Числитель выражения (26) содержит четыре блочные группы, которые можно поделить на 2^4 — 1 частей $x_1, x_2, ..., x_{15}$, составив таблицу.

Таблица 1

A_e	x_1	x_2	x_3	x_4		x_6	x_7	x_8				x_{12}			
A_f	x_1	x_2	x_3		x_5	x_6			x_9	x_{10}			x_{13}		
A_b	x_1	x_2		x_4	x_5		x_7		x_9					x_{14}	
A_d	x_1		x_3	x_4	x_5			x_8		x_{10}					x_{15}
	x_1	0	0	0	0	x_6	x_7	x_8	x_9	x_{10}		0	0	0	0

На основании закона для циклов можем написать (см. рис. 11)

$$A_e + A_f + A_b + A_d = 0,$$

тогда

$$x_2 = x_3 = x_4 = x_5 = x_{12} = x_{13} = x_{14} = x_{15} = 0$$

(это видно из нижней строки табл. 1). В соответствии с табл. 1 напишем

$$\mathsf{A}_e + \mathsf{A}_f - \mathsf{A}_b - \mathsf{A}_d = 2(x_6 - x_{11})$$

причем

$$x_6 = (\mathrm{A}_e \cap \mathrm{A}_f) - (\mathrm{A}_e \cap \mathrm{A}_f \cap \mathrm{A}_b \cap \mathrm{A}_d)$$

$$x_{11} = (\mathrm{A}_b \cap \mathrm{A}_d) - (\mathrm{A}_e \cap \mathrm{A}_f \cap \mathrm{A}_b \cap \mathrm{A}_d)$$

следовательно,

$$K_{ac} = \frac{\det_Y(\mathrm{A}_e \cap \mathrm{A}_f) - \det_Y(\mathrm{A}_b \cap \mathrm{A}_d)}{2\det_Y \mathrm{A}_a} \qquad (27)$$

Каждое из выражений (24) — (27) содержит пять разных величин, необходимых для расчета K_{ac}.

Каким образом уменьшить число этих величин?

При выводе формулы (27) было применено правило циклов для контура, состоящего из четырех путей: *b, d, e, f*. Теперь используем это правило для всех независимых контуров, состоящих из трех путей: *(с, d, f)*, *(a, d, e)* и (*a, b, f*) (рис. 11). Имеем

$$A_c + A_d + A_f = 0, \qquad (а)$$

$$A_a + A_d + A_e = 0, \qquad (б)$$

$$A_a + A_b + A_f = 0. \qquad (в)$$

Для этих блочных групп $\mathrm{A}_a, \mathrm{A}_c, \mathrm{A}_d, \mathrm{A}_f$ составим таблицу частей, аналогичную табл. 1.

Таблица 2

A_a	x_1	x_2	x_3	x_4		x_6	x_7	x_8				x_{12}			
A_c	x_1	x_2	x_3		x_5	x_6			x_9	x_{10}			x_{13}		
A_d	x_1	x_2		x_4	x_5		x_7		x_9		x_{11}	.		x_{14}	
A_f	x_1		x_3	x_4	x_5			x_8		x_{10}	x_{11}				x_{15}
	0	x_2	x_3	x_4	0	0	0	0	x_9	x_{10}	x_{11}	x_{12}	0	0	0

Используя уравнение (а), находим

$$x_1 = x_5 = x_6 = x_7 = x_8 = x_{13} = x_{14} = x_{15} = 0$$

(это показано в нижней строке табл. 2).

Из уравнения (б) следует, что

$$A_e = A_a + A_d \overset{s}{=} x_1 + x_9 + x_{11} + x_{12},$$

а из (в) —

$$A_{\text{ь}} = A_a + A_f \overset{s}{=} x_2 + x_{10} + x_{11} + x_{12}.$$

В результате имеем

$$\mathsf{A}_e + \mathsf{A}_f = 2x_3 + x_4 + x_9 + x_{10} + 2x_{11} + x_{12},$$

$$\mathsf{A}_b + \mathsf{A}_d = 2x_2 + x_4 + x_9 + x_{10} + 2x_{11} + x_{12},$$

поэтому

$$(\mathsf{A}_e + \mathsf{A}_f) - (\mathsf{A}_b + \mathsf{A}_d) = 2(x_3 - x_2)$$

Из табл. 2 вытекает

$$x_3 = (\mathsf{A}_a \,\mathrm{I}\, \mathsf{A}_c \,\mathrm{I}\, \mathsf{A}_f)$$

$$x_2 = (\mathsf{A}_a \,\mathrm{I}\, \mathsf{A}_c \,\mathrm{I}\, \mathsf{A}_d)$$

Таким образом,

$$(\mathsf{A}_e + \mathsf{A}_f) - (\mathsf{A}_b + \mathsf{A}_d) = 2\{(\mathsf{A}_a \,\mathrm{I}\, \mathsf{A}_c \,\mathrm{I}\, \mathsf{A}_f) - (\mathsf{A}_a \,\mathrm{I}\, \mathsf{A}_c \,\mathrm{I}\, \mathsf{A}_d)$$

Теперь можно написать следующую формулу для коэффициента передачи:

$$K_{ac} = \frac{\det\limits_Y(\mathsf{A}_a \,\mathrm{I}\, \mathsf{A}_c \,\mathrm{I}\, \mathsf{A}_f) - \det\limits_Y(\mathsf{A}_a \,\mathrm{I}\, \mathsf{A}_c \,\mathrm{I}\, \mathsf{A}_d)}{2\det\limits_Y \mathsf{A}_a}, \qquad (28)$$

причем

$$\mathsf{A}_a \,\mathrm{I}\, \mathsf{A}_c \,\mathrm{I}\, \mathsf{A}_f \,\mathrm{I}\, \mathsf{A}_d = 0. \qquad (29)$$

Выражение (28) содержит четыре разные величины, т. е. на одну меньше, чем предыдущие формулы. Кроме того, согласно формуле (29), это выражение характерно тем, что обе функции в числителе не имеют общих слагаемых.

Числитель в выражении (28) назовем *функцией совпадения* и обозначим

$$\operatorname*{Sim}_Y\left(\frac{\partial A}{\partial a}, \frac{\partial A}{\partial c}\right) = \operatorname*{Sim}_Y(A_a, A_c) = \det(\mathsf{A}_a \,\mathrm{I}\, \mathsf{A}_c \,\mathrm{I}\, \mathsf{A}_f) - \det_Y(\mathsf{A}_a \,\mathrm{I}\, \mathsf{A}_c \,\mathrm{I}\, \mathsf{A}_d) \qquad (30)$$

Легко показать, что

$$(\mathsf{A}_a \,\mathrm{I}\, \mathsf{A}_c \,\mathrm{I}\, \mathsf{A}_f) - (\mathsf{A}_a \,\mathrm{I}\, \mathsf{A}_c \,\mathrm{I}\, \mathsf{A}_d) = (\mathsf{A}_a \,\mathrm{I}\, \mathsf{A}_c \,\mathrm{I}\, \mathsf{A}_e) - (\mathsf{A}_a \,\mathrm{I}\, \mathsf{A}_c \,\mathrm{I}\, \mathsf{A}_d) \qquad (31)$$

а также

$$\mathsf{A}_a \,\mathrm{I}\, \mathsf{A}_c \,\mathrm{I}\, \mathsf{A}_e \mathsf{A}_b = 0 \qquad (32)$$

Поэтому можно написать другую формулу для функции совпадения:

$$\operatorname*{Sim}_Y(A_a, A_c) = \operatorname*{det}_Y(A_a \cap A_c \cap A_e) - \operatorname*{det}_Y(A_a \cap A_c \cap A_b) \quad (33)$$

Из табл. 2 получаем

$$A_a \cap A_c = x_2 + x_3 = (A_a \cap A_c \cap A_f) + (A_a \cap A_c \cap A_d) \quad (34)$$

Таким образом, блочная группа $A_a \cap A_c$ состоит из всех столбцов, входящих в выражение

$$(A_a \cap A_c \cap A_f) - (A_a \cap A_c \cap A_d),$$

и не содержит других столбцов. Отсюда следует, что функцию совпадения можно также найти, рассчитав конъюнкцию

$$\frac{\partial A}{\partial a} \cap \frac{\partial A}{\partial c}$$

и определив знаки ее отдельных слагаемых.

Заметим, что столбцы блочной группы

$$\frac{\partial A}{\partial a} \cap \frac{\partial A}{\partial c}$$

представляют собой деревья, которые соединяют пары вершин ({α, μ} и {β, ν}) или ({α, ν} и {β, μ}) (упрощенно показано на рис. 12).

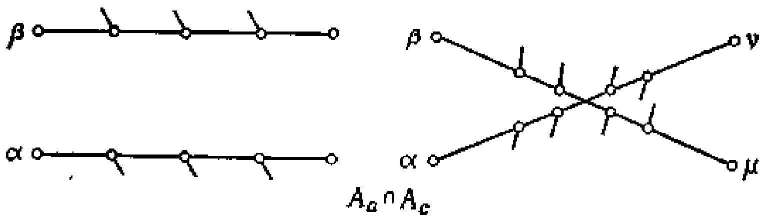


Рис. 12. Упрощенный рисунок 2-деревьев — изображений столбцов блочной группы $A_a \cap A_c$.

Столбцы блочной группы

$$A_a \cap A_c \cap A_f$$

представляют собой 2-деревья, соединяющие пары вершин

{α, μ} и {β, μ };

столбцы блочной группы

$$A_a \cap A_c \cap A_d$$

—2-деревья, соединяющие пары вершин

$$\{\alpha, \nu\} \quad \text{и} \quad \{\beta, \mu\}$$

(упрощенно показано на рис.   13).

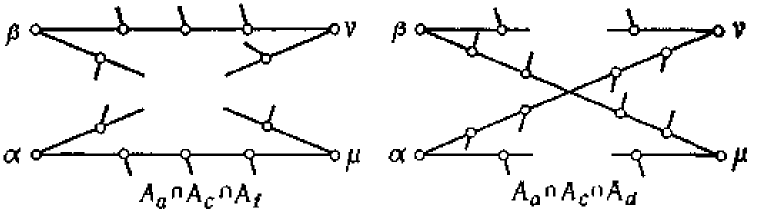


Рис. 13.   Упрощенный   рисунок   2-деревьев — изображений столбцов блочных групп, записанных под рисунком.

Рассматривая в графе *Г* пути, соответствующие определенным столбцам блочной группы

$$\frac{\partial A}{\partial a} \mathrm{I} \frac{\partial A}{\partial b},$$

можно определить знаки их детерминантных функций. Эти знаки зависят от направления напряжений *E* и *U* (рис.   11).

Выражая коэффициент передачи четырехполюсника K_{ac} с помощью функции совпадения, получим

$$K_{ac} = \frac{\operatorname*{Sim}_{Y}(A_a, A_c)}{\operatorname*{det}_{Y} A_a}. \tag{35}$$

Для определения знаков детерминантных функций эта формула требует расчета только двух величин A_a и A_c и их конъюнкции. Сравнивая выражения (25) и (35), получаем

$$\operatorname*{Sim}_{Y}(A_a, A_c) = \frac{\Delta_e + \Delta_f - \Delta_b - \Delta_d}{2} \tag{36}$$

Легко убедиться, что приведенные выше формулы для коэффициента передачи схемы также справедливы для дополнительных блочных групп A^d графа схемы, если положить

$$\Delta^d = \operatorname*{det}_{Z} A^d,$$

а также

$$\operatorname*{Sim}_{Z}(A^d{}_a, A^d{}_c) = \operatorname*{det}_{Z}(A^d{}_a \mathrm{I}\ A^d{}_c \mathrm{I}\ A^d{}_f) - \operatorname*{det}_{Z}(A^d{}_a \mathrm{I}\ A^d{}_c \mathrm{I}\ A^d{}_d) \tag{37}$$

причем

$$A^d{}_a = [a]\, A^d \text{ и т. д.}$$

Заметим также, что

$$\operatorname*{Sim}_{Y}(A_a, A_c) = \operatorname*{Sim}_{Y}(A_c, A_a) \tag{38}$$

Для расчета коэффициента передачи напряжения модуль-схемы можно воспользоваться любой из формул, однако самая удобная из них — формула (35). Проиллюстрируем это на примерах.

Пример 7. Рассчитать коэффициент передачи напряжения K_{12} (рис. 14, *а),* модуль-граф *Г* которой показан на рис. 14, *б,* а скелет $Г_0$ графа *Г* (сплошные линии) — на рис. 14, *в.*

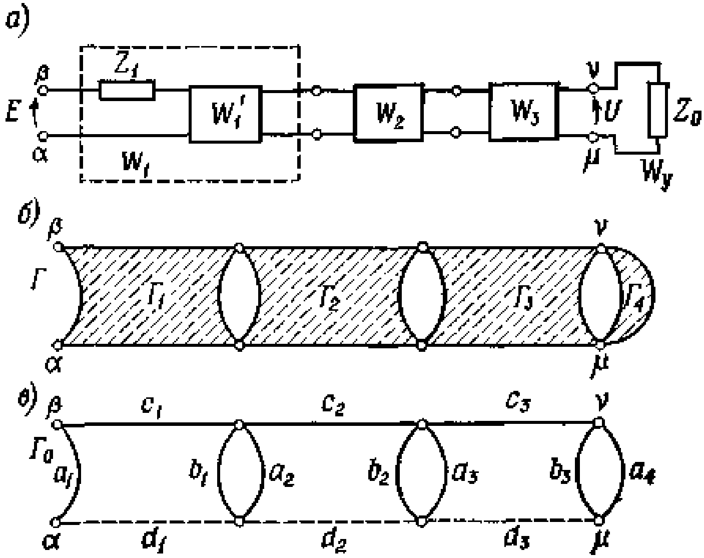


Рис. 14.

В этой схеме входной импеданс Z_1 включен в модуль (многополюсник) W_1. Блочная группа ${}^2A$ графа *Г* равна

$$^2A = [A_0^d] \downarrow \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} = \begin{bmatrix} b_1 & b_1 & a_2 & a_2 & b_1 & b_1 & a_2 & a_2 \\ b_2 & a_3 & b_2 & a_3 & b_2 & a_3 & b_2 & a_3 \\ b_3 & b_3 & b_3 & b_3 & a_4 & a_4 & a_4 & a_4 \end{bmatrix} \downarrow \begin{bmatrix} A_1 \\ A_2 \\ A_3 \\ A_4 \end{bmatrix} =$$

$$= \begin{bmatrix} A_{1b} & A_{1b} & A_1 & A_1 & A_{1b} & A_{1b} & A_1 & A_1 \\ A_{2b} & A_2 & A_{2ab} & A_{2a} & A_{2b} & A_2 & A_{2ab} & A_{2a} \\ A_{3b} & A_{3ab} & A_{3b} & A_{3ab} & A_3 & A_{3a} & A_3 & A_{3a} \\ A_4 & A_4 & A_4 & A_4 & A_{4a} & A_{4a} & A_{4a} & A_{4a} \end{bmatrix}.$$

Чтобы рассчитать $K_{12} = K_{a1a4}$ по формуле (25)

$$K_{a_1a_4} = \frac{\Delta_{\beta\mu} + \Delta_{\alpha\nu} - \Delta_{\beta\nu} - \Delta_{\alpha\mu}}{2\Delta_{\alpha\beta}},$$

необходимо вычислить блочные группы

$$^2A_{\beta\mu} = [c_1\ c_2\ c_3\ a_4] \downarrow [^2A],$$
$$^2A_{\alpha\nu} = [d_1\ d_2\ d_3\ a_4] \downarrow [^2A],$$
$$^2A_{\beta\nu} = [c_1\ c_2\ c_3] \downarrow [^2A],$$
$$^2A_{\alpha\mu} = [d_1\ d_2\ d_3] \downarrow [^2A],$$
$$A_{\alpha\beta} = Aa_1 = [a_1] \downarrow [^2A],$$

где d_1, d_2, d_3 — пути, показанные на рисг. 14, *в* штриховыми линиями (можно также выбрать другие пути или использовать пути, представляемые ребрами скелета Γ_0), а потом преобразовать эти блочные группы в полные блочные группы.

Как видно, этот способ расчета передачи довольно сложен, поэтому K_{a1a4} рассчитаем с помощью функции совпадения по формуле (35)

$$K_{a_1a_4} = \frac{\underset{Y}{\mathrm{Sim}}\,(A_{a_1}, A_{a_4})}{\underset{Y}{\det}\, A_{a_1}}. \qquad \text{(a)}$$

Рассчитаем конъюнкцию

$$A_{a_1} \cap A_{a_4} \overset{e}{=} {}^2A_{a_1} \cap {}^2A_{a_4} = ([a_1] \downarrow [^2A]) \cap ([a_4] \downarrow [^2A]) =$$

$$= \begin{bmatrix} A_{1ab} & A_{1ab} & A_{1a} & A_{1a} & A_{1ab} & A_{1ab} & A_{1a} & A_{1a} \\ A_{2b} & A_2 & A_{2ab} & A_{2a} & A_{2b} & A_2 & A_{2ab} & A_{2a} \\ A_{3b} & A_{3ab} & A_{3b} & A_{3ab} & A_3 & A_{3a} & A_3 & A_{3a} \\ A_4 & A_4 & A_4 & A_4 & A_{4a} & A_{4a} & A_{4a} & A_{4a} \end{bmatrix} \cap$$

$$\cap \begin{bmatrix} A_{1b} & A_{1b} & A_1 & A_1 \\ A_{2b} & A_2 & A_{2ab} & A_{2a} \\ A_{3b} & A_{3ab} & A_{3b} & A_{3ab} \\ A_{4a} & A_{4a} & A_{4a} & A_{4a} \end{bmatrix}.$$

Для этого используем таблицы порядков обоих блочных групп. Очевидно, что конъюнкция будет ненулевой только для тех столбцов блочных групп, которые имеют одинаковые столбцы в таблицах порядков. Определим конъюнкцию таблиц порядков

$$\mathbf{P}_{a_1} \cap \mathbf{P}'_{a_4} = \begin{pmatrix} 2 & 2 & 1 & 1 & 2 & 2 & 1 & 1 \\ 1 & 0 & 2 & 1 & 1 & 0 & 2 & 1 \\ 1 & 2 & 1 & 2 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 \end{pmatrix} \cap \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 1 \\ 1 & 2 & 1 & 2 \\ 1 & 1 & 1 & 1 \end{pmatrix} = \begin{pmatrix} 1 \\ 1 \\ 1 \\ 1 \end{pmatrix},$$

из которой следует

$$A_{a_1} \cap A_{a_4} \overset{e}{=} \begin{bmatrix} A_{1a} \\ A_{2a} \\ A_{3a} \\ A_{4a} \end{bmatrix} \cap \begin{bmatrix} A_{1b} \\ A_{2b} \\ A_{3b} \\ A_{4a} \end{bmatrix} \overset{e}{=} (A_{1a} \cap A_{1b})\,(A_{2a} \cap A_{2b})\,(A_{3a} \cap A_{3b})\,A_{4a}.$$

Так как это произведение не имеет дефекта произведения, то, согласно формуле (а) для K_{a1a4} рассматриваемой схемы, получаем

$$K_{a_1a_4} = \frac{\underset{Y}{\mathrm{Sim}}(A_{1a},\ A_{1b})\,\underset{Y}{\mathrm{Sim}}(A_{2a},\ A_{2b})\,\underset{Y}{\mathrm{Sim}}(A_{3a},\ A_{3b})\,\underset{Y}{\det} A_{4a}}{\underset{Y}{\det} A_{a_1}}. \quad (39)$$

Так как модуль $Г_4$ графа $Г$ имеет две вершины, то детерминантная функция $\underset{Y}{\det} A_{4a}$ одновременно служит функцией совпадения двухполюсника нагрузки Z_0

$$\underset{Y}{\det} A_{4a} = \underset{Y}{\mathrm{Sim}}(A_{4a}, A_{4a})$$

Коэффициент передачи K_{a1a4} рассматриваемой схемы можно записать в виде

$$K_{a_1a_4} = \frac{\prod_{i=1}^{4} \underset{Y}{\mathrm{Sim}}(A_{ia},\ A_{ib})}{\underset{Y}{\det} A_{a_1}}. \quad (40)$$

Эту формулу легко обобщить на случай цепной схемы с n-модулями

$$K_{a_1b_n} = \frac{\prod_{i=1}^{n} \underset{Y}{\mathrm{Sim}}(A_{ia},\ A_{ib})}{\underset{Y}{\det} A_{a_1}} \quad (41)$$

или

$$\underset{Y}{\mathrm{Sim}}(A_{a1}, A_{bn}) = \prod_{i=1}^{n} \underset{Y}{\mathrm{Sim}}(A_{ia}, A_{ib}) \quad (42)$$

Это означает, что функция совпадения цепной схемы с n-модулями равна произведению функций совпадения всех отдельных четырехполюсников схемы и двухполюсника нагрузки. Выражения (41) и (42) справедливы и для дополнительных блочных групп A^d.

Пример 8. Рассчитать функцию совпадения цепной схемы (рис. 15, *а),* состоящей из четырех четырехполюсников, графы которых показаны на рис. 15, *б.*

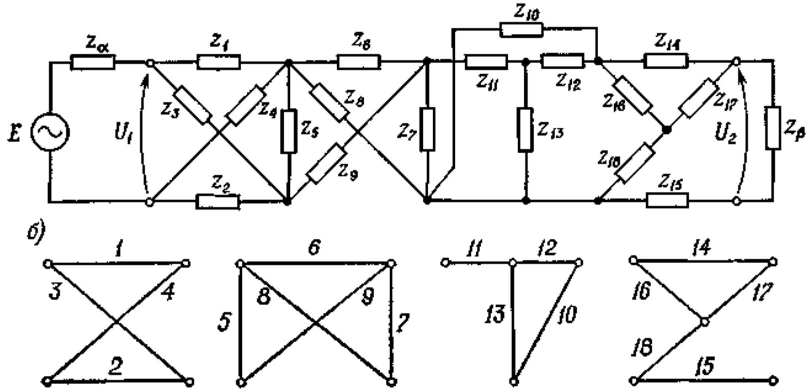


Рис. 15.

Рассчитаем функции совпадения отдельных четырехполюсников:

$$\operatorname*{Sim}_z\left(\begin{bmatrix}1&1&2&3\\2&3&4&4\end{bmatrix},\begin{bmatrix}1&1&2&3\\2&4&3&4\end{bmatrix}\right)=z_3z_4-z_1z_2,$$

$$\operatorname*{Sim}_z\left(\begin{bmatrix}5&5&5&5&5\\6&6&6&7&8\\7&8&9&9&9\end{bmatrix},\begin{bmatrix}5&5&6&6&7\\6&7&7&7&8\\7&8&8&9&9\end{bmatrix}\right)=-z_5z_6z_7,$$

$$\operatorname*{Sim}_z\left(\begin{bmatrix}11&11&10&12&10\\13&12&11&13&13\end{bmatrix},\begin{bmatrix}10&10\\12&13\end{bmatrix}\right)=z_{10}z_{13},$$

$$\operatorname*{Sim}_z\left(\begin{bmatrix}14&16&14&17&16\\16&17&18&18&18\end{bmatrix},\begin{bmatrix}14&14&14&15&16&16&15&17\\15&17&18&16&17&18&17&18\end{bmatrix}\right)=$$

$$=z_{16}z_{17}+z_{14}z_{18}+z_{17}z_{18}+z_{16}z_{18}.$$

Согласно формуле (42), функцию совпадения данной схемы запишем в виде

$$\operatorname*{Sim}_z(A^a_{a_1},\ A^d_{b_4})=(z_1z_2-z_3z_4)\,z_5z_6z_7z_{10}z_{13}\,(z_{16}z_{17}+z_{14}z_{18}+z_{17}z_{18}+z_{16}z_{18})\,z_\beta.$$

Пример 9. Рассчитать коэффициент передачи напряжения модуль-схемы (рис. 16, *а).*

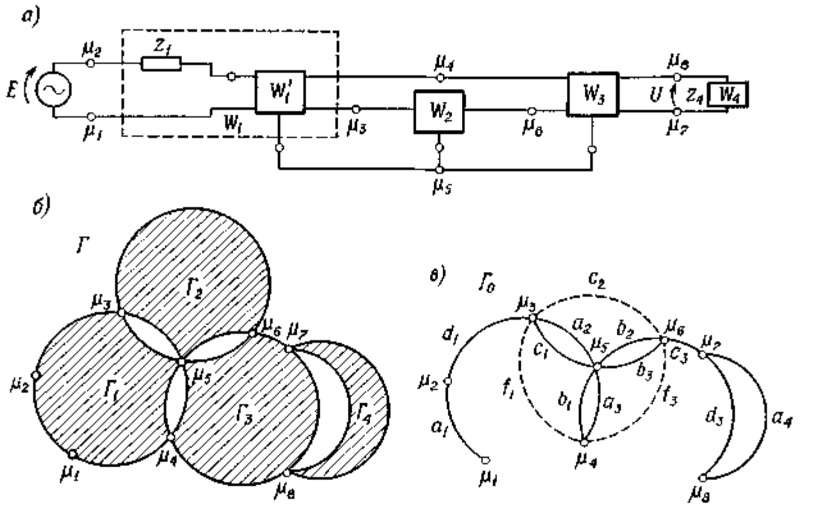


Рис. 16.

Для упрощения расчетов включим входной импеданс Z_1 в модуль (многополюсник) W_1. Модуль-граф *Г* этой схемы изображен на рис. 16, *б,* а его скелет $Г_0$ — на рис. 16, *в* (сплошные линии). Дополнительная блочная группа A^d_0 скелета $Г_0$ равна

$$A_0^d = [c_1a_2]\,[b_1a_3]\,[b_2b_3]\,[d_3a_4] =$$

$$= \begin{bmatrix} c_1 & c_1 & b_1 & a_2 & c_1 & c_1 & b_1 & a_2 & c_1 & c_1 & b_1 & a_2 & c_1 & c_1 & b_1 & a_2 \\ b_1 & b_2 & a_2 & b_2 & b_1 & a_3 & a_2 & a_3 & b_1 & b_2 & a_2 & b_2 & b_1 & a_3 & a_2 & a_3 \\ b_2 & a_3 & b_2 & a_3 & b_3 & b_3 & b_3 & b_3 & b_2 & a_3 & b_2 & a_3 & b_3 & b_3 & b_3 & b_3 \\ d_3 & d_3 & d_3 & d_3 & d_3 & d_3 & d_3 & d_3 & a_4 & a_4 & a_4 & a_4 & a_4 & a_4 & a_4 & a_4 \end{bmatrix}. \quad \text{(a)}$$

Для упрощения записи на основе рассчитанной блочной группы A^d_0 построим таблицу порядков блочной группы ${}^2A$ графа *Г*

$$\mathbf{P} = \begin{pmatrix} 2 & 1 & 1 & 0 & 2 & 1 & 1 & 0 & 2 & 1 & 1 & 0 & 2 & 1 & 1 & 0 \\ 1 & 1 & 2 & 2 & 0 & 0 & 1 & 1 & 1 & 1 & 2 & 2 & 0 & 0 & 1 & 1 \\ 1 & 2 & 1 & 2 & 2 & 3 & 2 & 3 & 0 & 1 & 0 & 1 & 1 & 2 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix},$$

а также таблицы порядков блочных групп ${}^2A_{a1} = [a_1] \downarrow [{}^2A]$ и ${}^2A_{a4} = [a_4] \downarrow [{}^2A]$:

$$\mathbf{P}_{a_1}=\begin{array}{c}\begin{pmatrix} 3 & 2 & 2 & 1 & 3 & 2 & 2 & 1 & 3 & 2 & 2 & 1 & 3 & 2 & 2 & 1 \\ 1 & 1 & 2 & 2 & 0 & 0 & 1 & 1 & 1 & 1 & 2 & 2 & 0 & 0 & 1 & 1 \\ 1 & 2 & 1 & 2 & 2 & 3 & 2 & 3 & 0 & 1 & 0 & 1 & 1 & 2 & 1 & 2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \\ \begin{matrix} & X_5 & & & & & X_5 & & & X_1 & & X_2 & & X_3 & X_1 & X_4 \end{matrix}\end{array}, \qquad (б)$$

$$\mathbf{P}_{a_4}=\begin{array}{c}\begin{pmatrix} 2 & 1 & 1 & 0 & 2 & 1 & 1 & 0 \\ 1 & 1 & 2 & 2 & 0 & 0 & 1 & 1 \\ 1 & 2 & 1 & 2 & 2 & 3 & 2 & 3 \\ 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \end{pmatrix} \\ \begin{matrix} X_1 & X_4 & X_2 & & X_3 & & X_4 & \end{matrix}\end{array}.$$

В таблицах $\mathbf{P}_{a_1}$ и $\mathbf{P}_{a_4}$ одинаковые столбцы обозначаем X_1, X_2, X_3, X_4, X_5. Эти столбцы заменим соответствующими столбцами блочных групп ${}^2A_{a1}$ и ${}^2A_{a4}$ и напишем конъюнкцию

$$ {}^2A_{a_1}\cap{}^2A_{a_4}=\left(\begin{bmatrix} A_{1ac} & A_{1ab} \\ A_{2b} & A_{2a} \\ A_{3a} & A_{3b} \\ A_{4a} & A_{4a} \end{bmatrix}\cap\begin{bmatrix} A_{1bc} \\ A_{2b} \\ A_{3d} \\ A_{4a} \end{bmatrix}\right)+\left(\begin{bmatrix} A_{1a} \\ A_{2ab} \\ A_{3a} \\ A_{4a} \end{bmatrix}\cap\begin{bmatrix} A_{1b} \\ A_{2ab} \\ A_{3d} \\ A_{4a} \end{bmatrix}\right)+$$

$$+\left(\begin{bmatrix} A_{1ac} \\ A_{2} \\ A_{3ab} \\ A_{4a} \end{bmatrix}\cap\begin{bmatrix} A_{1bc} \\ A_{2} \\ A_{3bd} \\ A_{4a} \end{bmatrix}\right)+\left(\begin{bmatrix} A_{1a} \\ A_{2a} \\ A_{3ab} \\ A_{4a} \end{bmatrix}\cap\begin{bmatrix} A_{1c} & A_{1b} \\ A_{2b} & A_{2a} \\ A_{3ad} & A_{3bd} \\ A_{4a} & A_{4a} \end{bmatrix}\right),$$

причем эта сумма не имеет дефекта.
Выносят A_{4a} за скобки, получаем

$$ {}^2A_{a_1}\cap{}^2A_{a_4}=A_{4a}\left\{\left(\begin{bmatrix} A_{1ac} & A_{1ab} \\ A_{2b} & A_{2a} \\ A_{3a} & A_{3a} \end{bmatrix}\cap\begin{bmatrix} A_{1bc} \\ A_{2b} \\ A_{3d} \end{bmatrix}\right)+\left(\begin{bmatrix} A_{1a} \\ A_{2ab} \\ A_{3a} \end{bmatrix}\cap\begin{bmatrix} A_{1b} \\ A_{2ab} \\ A_{3d} \end{bmatrix}\right)+\right.$$

$$\left.+\left(\begin{bmatrix} A_{1ac} \\ A_{2} \\ A_{3ab} \end{bmatrix}\cap\begin{bmatrix} A_{1bc} \\ A_{2} \\ A_{3bd} \end{bmatrix}\right)+\left(\begin{bmatrix} A_{1a} \\ A_{2a} \\ A_{3ab} \end{bmatrix}\cap\begin{bmatrix} A_{1c} & A_{1b} \\ A_{2b} & A_{2a} \\ A_{3ad} & A_{3bd} \end{bmatrix}\right)\right\}. \qquad (в)$$

Заменим блочные группы второго ранга в этом выражении полными блочными группами. Все блочные группы второго ранга, кроме первой и последней, одностолбцовые, поэтому преобразуем их в полные блочные группы путем замены их блочных групп первого ранга полными блочными группами.

Блочная группа с двумя столбцами

$$\begin{bmatrix} A_{1ac} & A_{1ab} \\ A_{2b} & A_{2a} \\ A_{3a} & A_{3b} \end{bmatrix},$$

таблица порядков которого имеет одинаковые столбцы, заменим полной блочной группой, используя при этом замещающие графы модулей Γ_{1a}, Γ_2, Γ_3 и соответственно дополняя скелет Γ_0 ребрами f_1, c_2 и f_3 (пунктирные линии на рис. 16, *в)*. Тогда

$$A_{1ac} = [b_{1(2)} \ \ f_{1(2)}],$$
$$A_{1ab} = [c_{1(2)} \ \ f_{1(2)}],$$
$$A_{2b} = [a_{2(1)} \ \ c_{2(1)}],$$
$$A_{2a} = [b_{2(1)} \ \ c_{2(1)}],$$
$$A_{3a} = [b_{3(1)} \ \ f_{3(1)}],$$
$$A_{3b} = [a_{3(1)} \ \ f_{3(1)}]$$

и

$$\begin{bmatrix} A_{1ac} & A_{1ab} \\ A_{2b} & A_{2a} \\ A_{3a} & A_{3b} \end{bmatrix} = \begin{bmatrix} [b_{1(2)} f_{1(2)}] & [c_{1(2)} f_{1(2)}] \\ [a_{2(1)} c_{2(1)}] & [b_{2(1)} c_{2(1)}] \\ [b_{3(1)} f_{3(1)}] & [a_{3(1)} f_{3(1)}] \end{bmatrix} =$$

$$= \mathcal{A}_{1ac}\mathcal{A}_{2b}\mathcal{A}_{3a} + \mathcal{A}_{1ab}\mathcal{A}_{2a}\mathcal{A}_{3b} - 2\begin{Bmatrix} f_{1(2)} \\ c_{2(1)} \\ f_{3(1)} \end{Bmatrix}. \qquad (г)$$

Отдельные дендритные веса равны

$$\left.\begin{aligned} f_{1(2)} &= \frac{1}{2}(\mathcal{A}_{1ab} + \mathcal{A}_{1ac} - \mathcal{A}_{1af}), \\ c_{2(1)} &= \frac{1}{2}(\mathcal{A}_{2a} + \mathcal{A}_{2b} - \mathcal{A}_{2c}), \\ f_{3(1)} &= \frac{1}{2}(\mathcal{A}_{3a} + \mathcal{A}_{3b} - \mathcal{A}_{3f}). \end{aligned}\right\} \qquad (д)$$

Аналогично другую блочную группу второго ранга с двумя столбцами в выражении (в) заменим полной блочной группой. После перехода к функциям совпадения и к детерминантным функциям получим

$$\underset{Y}{\mathrm{Sim}}(A_{a_1}, A_{a_4}) = \underset{Y}{\det} A_{4a}\Big\{\underset{Y}{\mathrm{Sim}}(A_{1ac}, A_{1bc}) \underset{Y}{\det} A_{2b} \underset{Y}{\mathrm{Sim}}(A_{3a}, A_{3d}) +$$
$$+ \underset{Y}{\mathrm{Sim}}(A_{1ab}, A_{1bc}) \underset{Y}{\mathrm{Sim}}(A_{2a}, A_{2b}) \underset{Y}{\mathrm{Sim}}(A_{3b}, A_{3d}) -$$

$$-\frac{1}{4}[\operatorname*{Sim}_Y(A_{1ab}, A_{1bc})+\operatorname*{Sim}_Y(A_{1ac}, A_{1bc})-\operatorname*{Sim}_Y(A_{1af}, A_{1bc})]\times$$

$$\times[\operatorname*{Sim}_Y(A_{2a}, A_{2b})+\operatorname*{det}_Y A_{2b}-\operatorname*{Sim}_Y(A_{2c}, A_{2b})]\times$$

$$\times[\operatorname*{Sim}_Y(A_{3a}, A_{3d})+\operatorname*{Sim}_Y(A_{3b}, A_{3d})-\operatorname*{Sim}_Y(A_{3f}, A_{3d})]+$$

$$+\operatorname*{Sim}_Y(A_{1a}, A_{1b})\operatorname*{det}_Y A_{2ab}\operatorname*{Sim}_Y(A_{3a}, A_{3d})+$$

$$+\operatorname*{Sim}_Y(A_{1ac}, A_{1bc})\operatorname*{det}_Y A_2\operatorname*{Sim}_Y(A_{3ab}, A_{3bd})+$$

$$+\operatorname*{Sim}_Y(A_{1a}, A_{1c})\operatorname*{Sim}_Y(A_{2a}, A_{2b})\operatorname*{Sim}_Y(A_{3ab}, A_{3ad})+$$

$$+\operatorname*{Sim}_Y(A_{1a}, A_{1b})\operatorname*{det}_Y A_{2a}\operatorname*{Sim}_Y(A_{3ab}, A_{3bd})-$$

$$-\frac{1}{4}[\operatorname*{Sim}_Y(A_{1a}A_{1b})+\operatorname*{Sim}_Y(A_{1a}, A_{1c})-\operatorname*{Sim}_Y(A_{1a}, A_{1f})]\times$$

$$\times[\operatorname*{det}_Y A_{2a}+\operatorname*{Sim}_Y(A_{2a}, A_{2b})-\operatorname*{Sim}_Y(A_{2a}, A_{2c})]\times$$

$$\times[\operatorname*{Sim}_Y(A_{3ab}, A_{3ad})+\operatorname*{Sim}_Y(A_{3ab}, A_{3bd})-\operatorname*{Sim}_Y(A_{3ab}, A_{3fd})]\Big\}. \quad \textbf{(e)}$$

Это выражение представляет собой функцию совпадения модуль-схемы (рис. 16, *а),* выраженную через функции совпадения и детерминантные функции отдельных модулей схемы. Для наглядности на рис. 17 показаны схемы многополюсников для некоторых функций совпадения формулы (е).

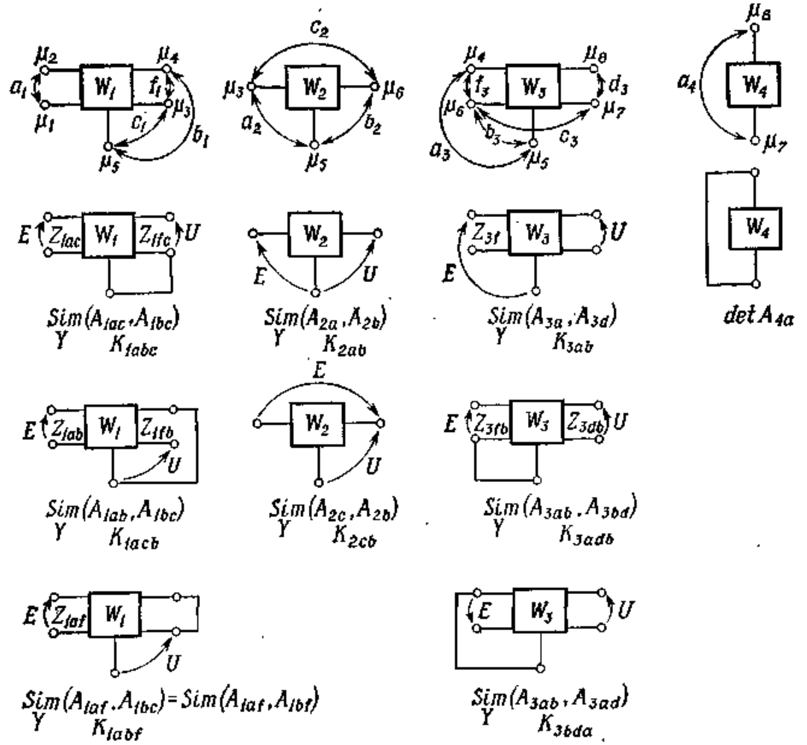


Рис. 17. Обозначения путей, функций совпадения, передаточных функций и входных импедансов многополюсников в различных стадиях замыкания.

Определяя их, нужно использовать обозначения путей, соединяющих отдельные вершины скелета $Г_0$ (рис. 16, *в*). Заметим, что в этой формуле равны следующие слагаемые:

$$\operatorname*{Sim}_Y(A_{1af}, A_{1bc}) = \operatorname*{Sim}_Y(A_{1af}, A_{1bf})\ Y$$

и

$$\operatorname*{Sim}_Y(A_{3ab}, A_{3fd}) = \operatorname*{Sim}_Y(A_{3af}, A_{3df})$$

так как

$$A_{1bc} = A_{1bf} \quad \text{и} \quad A_{3ab} = A_{3af}.$$

Для расчета коэффициента передачи K_{a1a4} рассматриваемой молуль-схемы необходимо определить полную блочную группу A_{a_1}, модуль-графа $Г$. Таблица порядков P_{a1} [выражение (б)] этой блочной группы содержит два одинаковых столбца *(X_1)*, которые в выражениях (г) и (д) заменены на полные блочные группы. Аналогично поступаем с другой парой одинаковых столбцов *(X_5)*. Обозначив через Δ детерминантную функцию $\det\limits_Y A$, можно записать следующее выражение для детерминантной функции блочной группы A_{a_1} :

$$\det_Y A_{a_1} = (\Delta_{1abc}\Delta_{2b}\Delta_{3d} + \Delta_{1ac}\Delta_{2b}\Delta_{3ad} + \Delta_{1ab}\Delta_{2ab}\Delta_{3d} + \Delta_{1a}\Delta_{2ab}\Delta_{3ad} +$$

$$+ \Delta_{1abc}\Delta_{2}\Delta_{3bd} + \Delta_{1ac}\Delta_{2}\Delta_{3abd} + \Delta_{1ab}\Delta_{2a}\Delta_{3bd} + \Delta_{1a}\Delta_{2a}\Delta_{3abd})\,\Delta_4 +$$

$$+ (\Delta_{1abc}\Delta_{2b}\Delta_{3} + \Delta_{1ac}\Delta_{2b}\Delta_{3a} + \Delta_{1ab}\Delta_{2ab}\Delta_{3} + \Delta_{1a}\Delta_{2ab}\Delta_{3a} +$$

$$+ \Delta_{1abc}\Delta_{2}\Delta_{3b} + \Delta_{1ac}\Delta_{2}\Delta_{3ab} + \Delta_{1ab}\Delta_{2a}\Delta_{3b} + \Delta_{1a}\Delta_{2a}\Delta_{3ab})\,\Delta_{4a} -$$

$$- \frac{1}{4}\{(\Delta_{1ab} + \Delta_{1ac} - \Delta_{1af})(\Delta_{2a} + \Delta_{2b} - \Delta_{2c})\,[\Delta_4(\Delta_{3ad} +$$

$$+ \Delta_{3bd} - \Delta_{3df}) + \Delta_{4a}(\Delta_{3a} + \Delta_{3b} - \Delta_{3f})]\}. \qquad (ж)$$

Коэффициент передачи K_{a1a4} рассматриваемой модуль-схемы получаем, поделив выражение (е) для функции совпадения схемы на выражение (ж) для детерминантной функции схемы

$$K_{a_1a_4} = \frac{\operatorname{Sim}\limits_Y (A_{a_1},\ A_{a_4})}{\det\limits_Y A_{a_1}}\,.$$

Поделив в последнем равенстве числитель и знаменатель на произведение

$$\Delta_1\Delta_2\Delta_3\Delta_4 = \det_Y A_1 \det_Y A_2 \det_Y A_3 \det_Y A_4\,,$$

найдем коэффициент передачи напряжения K_{a1a4} рассматриваемой модуль-схемы, выраженный через коэффициенты передачи напряжения и входные импедансы отдельных многополюсников схемы. Числитель коэффициента передачи будет иметь вид

$$Z_{4a}\left\{K_{1abc}Z_{1ac}Z_{1c}Z_{2b}K_{3ad}Z_{3a}+K_{1acb}Z_{1ab}Z_{1b}K_{2ab}Z_{2a}K_{3bd}Z_{3b}-\right.$$
$$-\frac{1}{4}(K_{1acb}Z_{1ab}Z_{1b}+K_{1abc}Z_{1ac}Z_{1c}-K_{1abf}Z_{1af}Z_{1f})\times$$
$$\times(K_{2ab}Z_{2a}+Z_{2b}-K_{2cb}Z_{2c})(K_{3ad}Z_{3a}+K_{3bd}Z_{3b}-K_{3fd}Z_{3f})+$$
$$+K_{1ab}Z_{1a}Z_{2ab}Z_{2b}K_{3ad}Z_{3a}+K_{1abc}Z_{1ac}Z_{1c}K_{3adb}Z_{3ab}Z_{3b}+$$
$$+K_{1ac}Z_{1a}K_{2ab}Z_{2a}K_{3bda}Z_{3ba}Z_{3a}+K_{1ab}Z_{1a}Z_{2a}K_{3adb}Z_{3ab}Z_{3b}-$$
$$-\frac{1}{4}(K_{1ab}Z_{1a}+K_{1ac}Z_{1a}-K_{1af}Z_{1a})(Z_{2a}+K_{2ab}Z_{2a}-K_{2ac}Z_{2a})\times$$
$$\left.\times(K_{3bda}Z_{3ba}Z_{3a}+K_{3adb}Z_{3ad}Z_{3d}-K_{3adf}Z_{3af}Z_{3f})\right\},$$

а знаменатель

$$Z_{1abc}Z_{1bc}Z_{1c}Z_{2b}Z_{3d}+Z_{1ac}Z_{1c}Z_{2b}Z_{3ad}Z_{3d}+Z_{1ab}Z_{1b}Z_{2ab}Z_{2c}Z_{3d}+$$
$$+Z_{1a}Z_{2ab}Z_{2b}Z_{3ad}Z_{3d}+Z_{1abc}Z_{1bc}Z_{1c}Z_{3bd}Z_{3d}+Z_{1ac}Z_{1c}Z_{3abd}Z_{3bd}Z_{3d}+$$
$$+Z_{1ab}Z_{1b}Z_{2a}Z_{3bd}Z_{3d}+Z_{1a}Z_{2a}Z_{3abd}Z_{3bd}Z_{3d}+$$
$$+(Z_{1abc}Z_{1bc}Z_{1c}Z_{2b}+Z_{1ac}Z_{1c}Z_{2b}Z_{3a}+Z_{1ab}Z_{1b}Z_{2ab}Z_{2b}+Z_{1a}Z_{2ab}Z_{2b}Z_{3a}+$$
$$+Z_{1abc}Z_{1bc}Z_{1c}Z_{3b}+Z_{1ac}Z_{1c}Z_{3ab}Z_{3b}+Z_{1ab}Z_{1b}Z_{2a}Z_{3b}+Z_{1a}Z_{2a}Z_{3ab}Z_{3b})Z_{4a}-$$
$$-\frac{1}{4}(Z_{1ab}Z_{1b}+Z_{1ac}Z_{1c}-Z_{1af}Z_{1f})(Z_{2a}+Z_{2b}-Z_{2c})\times$$
$$\times[Z_{3ad}Z_{3d}+Z_{3bd}Z_{3d}-Z_{3df}Z_{3f}+Z_{4a}(Z_{3a}+Z_{3b}-Z_{3f})].$$

В этих выражениях использованы обозначения:
K_{1abc} — коэффициент передачи напряжения многополюсника W_1 между путями a_1 и b_1 с замкнутым путем c_1;
K_{3ad} — коэффициент передачи напряжения многополюсника W_3 между путями a_3 и d_3;
K_{1acb} — коэффициент передачи напряжения многополюсника W_1 между путями a_1 и c_1 с замкнутым путем b_1;

………………………………………………………………

Z_{1abc} — импеданс многополюсника W_1 с замкнутыми путями b_1 и c_1, измеренный относительно пути a_1;
Z_{1ac} — импеданс многополюсника W_1 с замкнутым путем c_1, измеренный относительно пути a_1;
Z_{3b} — импеданс многополюсника W_3, измеренный относительно пути b_3;
Некоторые из этих обозначений поясняет рис. 17.
Последний пример иллюстрировал способ определения коэффициента передачи напряжения модуль-схемы через внешние параметры

отдельных многополюсников, не учитывая при этом их внутренней структуры.

Представим теперь второй способ определения коэффициента передачи напряжения модуль-схемы, который во многих практических случаях может быть более удобным, чем предыдущий.

В общем случае адмитанс $Y_{\alpha\beta}$ и импеданс $Z_{\alpha\beta}$ линейной модуль-схемы можно выразить через адмитанс y_k и импеданс z_k ветви k

$$Y_{\alpha\beta}=\frac{a'y_k+b'}{c'y_k+d'},$$

$$Z_{\alpha\beta}=\frac{a''z_k+b''}{c''z_k+d''},$$

где $a', b', c', d', a'', b'', c'', d''$ — величины, не зависящие от y_k и z_k. Применяя теорему Эйлера для однородных функций, последние выражения можно записать в виде

$$Y_{\alpha\beta}=\sum_{k=1}^{g}\frac{\partial Y_{\alpha\beta}}{\partial y_k}y_k, \tag{43}$$

$$Z_{\alpha\beta}=\sum_{k=1}^{g}\frac{\partial Z_{\alpha\beta}}{\partial Z_k}z_k, \tag{44}$$

где g — число ветвей цепи.

Составляя баланс мощности для пассивной цепи, питаемой от идеального источника напряжения $E_{\alpha\beta}$, и пользуясь формулой (43), получим

$$E_{\alpha\beta}^2\sum_{k=1}^{g}\frac{\partial Y_{\alpha\beta}}{\partial y_k}y_k=\sum_{k=1}^{g}U_k^2y_k,$$

т. е.

$$\sum_{k=1}^{g}\left(E_{\alpha\beta}^2\frac{\partial Y_{\alpha\beta}}{\partial y_k}-U_k^2\right)y_k=0,$$

где U_k — напряжение ветви k, $Y_{\alpha\beta}$ — входной адмитанс цепи.

Чтобы это уравнение было справедливо при произвольных величинах y_k, необходимо выполнение равенства

$$\frac{U_k}{E_{\alpha\beta}}=\pm\sqrt{\frac{\partial Y_{\alpha\beta}}{\partial y_k}}. \tag{45}$$

Эту формулу можно также применять для пары узлов $\{\mu, \nu\}$, не соединенных ветвью k, так как всегда можно предположить, что такая ветвь имеет адмитанс $y_k = 0$.

Обозначим через a и c произвольные пути между парами вершин $\{\alpha, \beta\}$ и $\{\mu, \nu\}$ графа $Г$ электрической цепи (рис. 18).

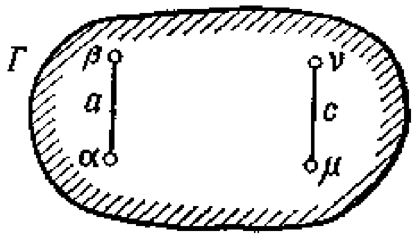


Рис. 18. Модуль-граф четырехполюсника с обозначенными входными и выходными путями.

Если $\mathcal{A}$ — полная блочная группа графа *Г,* то на основании формулы (45) напишем

$$K_{ac}=\frac{U_c}{E_a}=\pm\left\{\det_Y\frac{\partial}{\partial c}\left(\frac{\mathcal{A}}{\partial\mathcal{A}/\partial a}\right)\right\}^{1/2}=$$

$$=\frac{\left\{\det\limits_Y\left(\frac{\partial\mathcal{A}}{\partial a}\frac{\partial\mathcal{A}}{\partial c}\right)-\det\limits_Y\left(\mathcal{A}\frac{\partial^2\mathcal{A}}{\partial a\partial c}\right)\right\}^{1/2}}{\det\limits_Y\frac{\partial\mathcal{A}}{\partial a}}. \qquad (46)$$

Применив операцию перемещения нижних индексов, эту формулу приведем к виду

$$K_{ac}=\pm\frac{\{\det\limits_Y(\mathcal{A}_a\mathcal{A}_c)-\det\limits_Y(\mathcal{A}\mathcal{A}_{ac})\}^{1/2}}{\det\limits_Y\mathcal{A}_a}. \qquad (47)$$

Сравнивая ее с формулами (27) или (28), заметим, что выражение под корнем

$$\sqrt{\mathcal{A}_a\mathcal{A}_c-\mathcal{A}\mathcal{A}_{ac}}$$

равно квадрату разности двух полных блочных групп. Необходимым условием существования геометрического изображения блочной группы $A \overset{s}{=} \mathcal{A}$ будет условие существования действительного корня. Кроме того, из соотношения

$$\mathcal{A}_a\mathcal{A}_c-\mathcal{A}\mathcal{A}_{ac}=\frac{1}{4}(\mathcal{A}_e+\mathcal{A}_f-\mathcal{A}_b-\mathcal{A}_d)^2=$$

$$=\{(\mathcal{A}_e\cap\mathcal{A}_f)-(\mathcal{A}_b\cap\mathcal{A}_d)\}^2=$$

$$=\{(\mathcal{A}_a\cap\mathcal{A}_c\cap\mathcal{A}_f)-(\mathcal{A}_a\cap\mathcal{A}_c\cap\mathcal{A}_d)\}^2 \qquad (48)$$

следуют формулы перехода.

Выражения (47) можно также записать в виде

$$\mathrm{K}_{ac} = \pm \frac{(\Delta_a \Delta_c - \Delta \Delta_{ac})^{\frac{1}{2}}}{\Delta_a} \qquad (49)$$

где

$$\Delta = \det_Y A,\ \Delta_a = \det_Y \frac{\partial A}{\partial a},\ \Delta_c = \det_Y \frac{\partial A}{\partial c},$$

$$\Delta_{ac} = \det_Y \frac{\partial^2 A}{\partial a \partial c}, \quad A \overset{s}{=} \mathsf{A}$$

После элементарных преобразований формула (49) примет вид

$$\mathrm{K}_{ac} = \pm \left(\frac{Z_c - Z_{ca})}{Z_a} \right)^{\frac{1}{2}} \qquad (50)$$

где Z_a — входной импеданс цепи, Z_c — выходной импеданс цепи, Z_{ca} — выходной импеданс цепи с замкнутым входом.

Из формул (49) и (50) невозможно определить знак коэффициента передачи K_{ac}. На практике знак K_{ac} не всегда представляет интерес, или в противном случае его можно определить, например, путем рассмотрения одного 2-дерева в графе *Г,* которое служит изображением столбца одной из двух блочных групп, полученных в результате вычисления выражения

$$(\mathsf{A}_a \mathsf{A}_c - \mathsf{A}\mathsf{A}_{ac})^{\frac{1}{2}}$$

Из сравнения формул (46), (47), (49), (50) следует, что практически самой удобной является формула (50), которая при анализе модуль-схем позволяет рассчитать K_{ac} без рассмотрения структуры отдельных модулей.

Покажем это на примере.

Пример 10. Рассчитать коэффициент передачи K_{a1b3} модуль-схемы (рис. 19, *а),* причем входной импеданс и импеданс нагрузки этой схемы включены в многополюсники W_1 и W_3.

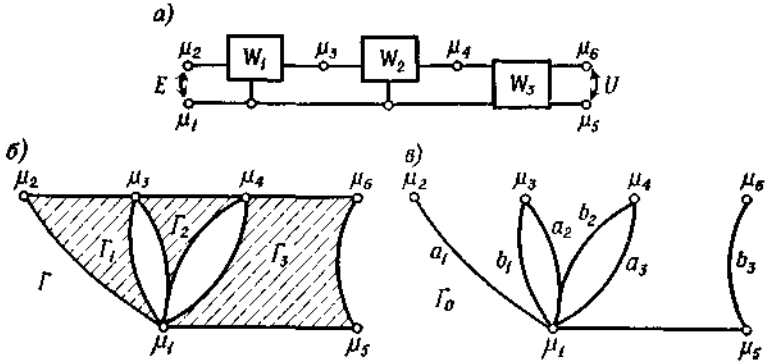


Рис. 19.

Модуль-граф этой схемы изображен на рис. 19, *б,* а его скелет Γ_0 — на рис. 19, *в.*

Блочная группа ${}^2A$ графа Γ равна

$$^2A = [A_0^d] \downarrow \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} b_1 & b_1 & a_2 & a_2 \\ b_2 & a_3 & b_2 & a_3 \end{bmatrix} \begin{bmatrix} A_1 \\ A_2 \\ A_3 \end{bmatrix} = \begin{bmatrix} A_{1b} & A_{1b} & A_1 & A_1 \\ A_{2b} & A_2 & A_{2ab} & A_{2a} \\ A_3 & A_{3a} & A_3 & A_{3a} \end{bmatrix},$$

а таблица порядков блочной группы

$$\mathbf{P} = \begin{pmatrix} 1 & 1 & 0 & 0 \\ 1 & 0 & 2 & 1 \\ 0 & 1 & 0 & 1 \end{pmatrix}$$

не содержит одинаковых столбцов. Поэтому можно непосредственно записать выражение для детерминантной функции Δ рассматриваемой модуль-схемы

$$\Delta = \Delta_{1b} (\Delta_{2b}\Delta_3 + \Delta_2\Delta_{3a}) + \Delta_1 (\Delta_{2ab}\,\Delta_3 + \Delta_{2a}\Delta_{3a}),$$

а затем и выражения для детерминантных функций Δ_{a_1} и Δ_{b_3} схемы, замкнутой на входе и выходе:

$$\Delta_{a_1} = \Delta_{1ab} (\Delta_{2b}\Delta_3 + \Delta_2\Delta_{3a}) + \Delta_{1a} (\Delta_{2ab}\,\Delta_3 + \Delta_{2a}\Delta_{3a}),$$

$$\Delta_{b_3} = \Delta_{1ab} (\Delta_{2b}\Delta_{3b} + \Delta_2\Delta_{3ab}) + \Delta_{1a} (\Delta_{2ab}\,\Delta_{3b} + \Delta_{2a}\Delta_{3ab}).$$

Рассчитаем теперь импедансы $Z_{b_3}, Z_{b_3a_1}, Z_{a_1}$:

$$Z_{b_3}=\left(\frac{\Delta_{b_3}}{\Delta_1\Delta_2\Delta_3}\right):\left(\frac{\Delta}{\Delta_1\Delta_2\Delta_3}\right)=\frac{Z_{1b}Z_{3b}(Z_{2b}+Z_{3ab})+Z_{2a}Z_{3b}(Z_{2ba}+Z_{3ab})}{Z_{1b}(Z_{2b}+Z_{3a})+Z_{2a}(Z_{2ba}+Z_{3a})},$$

$$Z_{b_3a_1}=\frac{Z_{1ba}Z_{3b}(Z_{2b}+Z_{3ab})+Z_{2a}Z_{3b}(Z_{2ba}+Z_{3ab})}{Z_{1ba}(Z_{2b}+Z_{3a})+Z_{2a}(Z_{2ba}+Z_{3a})},$$

$$Z_{a_1}=\left(\frac{\Delta_{a_1}}{\Delta_1\Delta_2\Delta_3}\right):\left(\frac{\Delta}{\Delta_1\Delta_2\Delta_3}\right)=\frac{Z_{1ab}Z_{1b}(Z_{2b}+Z_{3a})+Z_{1a}Z_{2a}(Z_{2ba}+Z_{3a})}{Z_{1b}(Z_{2b}+Z_{3a})+Z_{2a}(Z_{2ba}+Z_{3a})}.$$

Подставив рассчитанные выражения в формулу для коэффициента передачи схемы

$$K_{a_1b_3}=\pm\left(\frac{Z_{b_3}-Z_{b_3a_1}}{Z_{a_1}}\right)^{1/2},$$

получим коэффициент передачи рассматриваемой модуль-схемы, выраженный через импедансы многополюсников этой схемы в разных режимах короткого замыкания.

Заметим, что модуль-схема в этом примере представляет собой цепную схему, поэтому расчет коэффициента передачи можно упростить, используя при этом формулу (41), причем

$$\underset{Y}{\mathrm{Sim}}\,(A_{ia}, A_{ib})=\pm(\Delta_{ia}\Delta_{ib}-\Delta_i\Delta_{iab})^{1/2},$$

откуда

$$K_{a_1b_3}=\frac{\pm\{(\Delta_{1a}\Delta_{1b}-\Delta_1\Delta_{1ab})(\Delta_{2a}\Delta_{2b}-\Delta_2\Delta_{2ab})(\Delta_{3a}\Delta_{3b}-\Delta_3\Delta_{3ab})\}^{1/2}}{\Delta_{1ab}(\Delta_{2b}\Delta_3+\Delta_2\Delta_{3a})+\Delta_{1a}(\Delta_{2ab}\Delta_3+\Delta_{2a}\Delta_{3a})}.$$

Поделив числитель и знаменатель на произведение детерминантных функций AjAaAg многополюсников схемы, получаем следующее выражение:

$$K_{a_1b_3}=\frac{\pm[Z_{1a}(Z_{1b}-Z_{1ba})Z_{2a}(Z_{2b}-Z_{2ba})Z_{3a}(Z_{3b}-Z_{3ba})]^{1/2}}{Z_{1ab}Z_{1b}(Z_{2b}+Z_{3a})+Z_{1a}Z_{2a}(Z_{2ba}+Z_{3a})}.$$

2.3.5. Схемы замещения

На практике часто используются схемы замещения, которые имеют такие же электрические свойства, что и исходные схемы. Рассматривать схемы замещения весьма полезно при анализе модуль-схем, схем с индуктивными связями, с распределенными параметрами, а также с полупроводниковыми приборами. Рассмотрим *метод расчета схем замещения.* Если в исходной схеме (рис. 20, *а)* имеется v_z узлов, то схемой замещения будем называть цепь в виде полного многоугольника с v_z узлами и $\frac{v_z(v_z-1)}{2}$ ветвями (рис. 20, *б).*

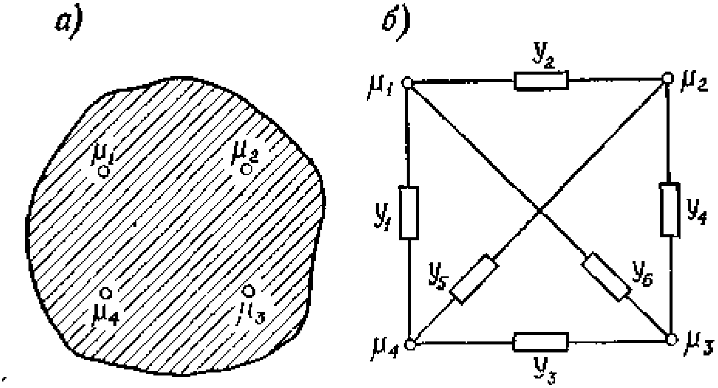


Рис. 20.

Пусть обозначения ребер графа схемы замещения одновременно будут соответствовать путям графа исходной схемы. Заметим, что граф схемы замещения представляет собой замещающий (полный) граф $Г_2$ графа $Г$ исходной схемы.

Допустим, что детерминантные функции графа A исходной схемы и графа A_z схемы замещения равны, т. е.

$$\det_Y A = \det_Y A_z \tag{51}$$

Для выполнения этого равенства необходимо, чтобы адмитанс y_k ветви α_k схемы замещения удовлетворял следующему соотношению:

$$y_k = \frac{\det\limits_Y A_k}{\det\limits_Y A_{k\alpha}} = \frac{\det\limits_Y a_{k\,(v_z-2)}}{\det\limits_Y \dfrac{\partial A}{\partial D}}, \tag{52}$$

где A_k — блочная группа двухполюсного модуля $Г_k$ в модуль-графе $Г^*_z$, эквивалентном графу $Г$;

α_k — обозначение ребра замещающего графа $Г_z$ соответствующей ветви α_k схемы замещения;

$A_{k_\alpha} = \dfrac{\partial A_k}{\partial \alpha_k}$; $a_{k(v_z-1)}$ — дендритный вес ребра α_k замещающего графа F_z;

D — блочная группа произвольного дерева, касающегося всех v_z выделенных вершин графа $Г$.

Дендритный вес $a_{k(v_z-2)}$ ребра α_k замещающего графа T_z рассчитаем по формулам, приведенным в ранее:

$$a_{k(v_z-2)} = \frac{1}{2}\left(\frac{\partial\mathcal{A}}{\partial D^{\mu_{1k}}} + \frac{\partial\mathcal{A}}{\partial D^{\mu_{2k}}} - \frac{\partial\mathcal{A}}{\partial(D^{\mu_{1k}\mu_{2k}}[\alpha_k])}\right) \quad (53)$$

или

$$a_{k(v_z-2)} \overset{s}{=} \frac{\partial A}{\partial D^{\mu_{1k}}} \cap \frac{\partial A}{\partial D^{\mu_{2k}}}, \quad (54)$$

где $\mathsf{A} \overset{s}{=} A$ — блочная группа графа *Г* исходной схемы;

$D^{\mu_{1k}}$ — блочная группа произвольного дерева, касающегося всех выделенных вершин графа *Г,* кроме вершины μ_{1k}, инцидентной ребру α_k;

$D^{\mu_{2k}}$ — блочная группа произвольного дерева, касающегося всех выделенных вершин графа *Г,* кроме вершины μ_{2k}, инцидентной ребру α_k;

$D^{\mu_{1k}\mu_{2k}}$ — блочная группа произвольного дерева, касающегося всех выделенных вершин графа *Г,* кроме вершин μ_{1k} и μ_{2k}, инцидентных ребру α_k. Обозначим

$$\Delta^{\mu_{1k}} = \det_Y \frac{\partial A}{\partial D^{\mu_{1k}}},$$

$$\Delta^{\mu_{2k}} = \det_Y \frac{\partial A}{\partial D^{\mu_{2k}}},$$

$$\Delta_{\alpha_k}^{\mu_{1k}\mu_{2k}} = \det_Y \frac{\partial A}{\partial(D^{\mu_{1k}\mu_{2k}}[\alpha_k])}, \quad (55)$$

$$\Delta' = \det_Y \frac{\partial A}{\partial D},$$

где $\Delta^{\mu_{1k}}(\Delta^{\mu_{2k}})$ — детерминантная функция, равная определителю матрицы проводимостей (узловых) исходной схемы при коротком замыкании всех выделенных узлов, кроме узла $\mu_{1k}(\mu_{2k})$, принадлежащего ветви α_k;

$\Delta_{\alpha k}^{\mu_{1k}\mu_{2k}}$ — детерминантная функция, равная определителю матрицы проводимостей (узловых) исходной схемы при коротком замыкании всех выделенных узлов, кроме узлов μ_{1k} и μ_{2k}, которые в свою очередь замкнуты друг с другом;

Δ' — детерминантная функция, равная определителю матрицы проводимостей исходной схемы при коротком замыкании всех v_z выделенных узлов.

Теперь можно записать равенство

$$y_k = \frac{\Delta^{\mu_{1k}} + \Delta^{\mu_{2k}} - \Delta^{\mu_{1k}\mu_{2k}}}{2\Delta'} \qquad (56)$$

откуда после преобразований получаем

$$y_k = \frac{1}{2}(Y_{\alpha_k}^{\mu_1} + Y_{\alpha_k}^{\mu_2} - Y_{\alpha_{k\pm1}\alpha_k}^{\mu_1\mu_2}),$$

где $Y_{\alpha_k}^{\mu_1}$ — адмитанс исходной схемы при коротком замыкании всех выделенных узлов, кроме узла μ_{1k}, измеренный относительно пути α_k (т. е. между узлом μ_{1k} и общим узлом);

$Y_{\alpha_k}^{\mu_2}$ — адмитанс исходной схемы при коротком замыкании всех выделенных узлов, кроме узла μ_{2k}, измеренный относительно пути α_k;

$Y_{\alpha_{k\pm1}\alpha_k}^{\mu_1\mu_2}$ — адмитанс исходной схемы при коротком замыкании всех выделенных узлов, кроме замкнутых друг с другом узлов μ_{1k} и μ_{2k}, измеренный относительно пути α_{k+1} или α_{k-1} (т. е. между узлом μ_{1k} + μ_{2k} и общим узлом).

Адмитанс y_k ветви α_k графа замещения Γ_z, выраженный в виде конъюнкции (54), рассчитаем по формуле

$$y_k = \frac{\det_{Y}\left(\frac{\partial A}{\partial D^{\mu_{1k}}} \sqcap \frac{\partial A}{\partial D^{\mu_{2k}}}\right)}{\det_{Y}\frac{\partial A}{\partial D}}. \qquad (58)$$

Описанный способ расчета адмитанса ветви схемы замещения иллюстрируют следующие примеры.

Пример 11. Для схемы рис. 21, *а* построить схему замещения с тремя узлами, показанную на рис. 21, б, и рассчитать адмитансы ветвей схемы замещения.

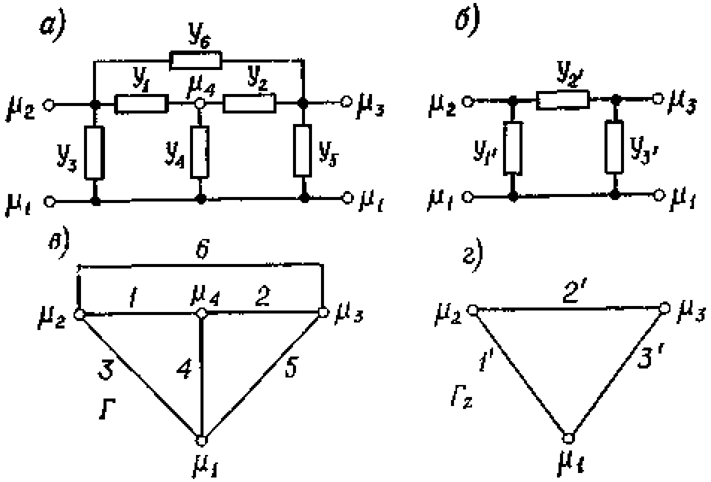


Рис. 21.

Графы обеих схем изображены на рис. 21, *в* и *г.* Для расчетов используем формулу (58).

Блочная группа графа Γ исходной схемы равна

A = [1 3 61 [2 5 6] [3 4 5].

Для определения адмитанса $y_{1'}$ рассчитаем конъюнкцию

$$(A_{2'} \cap A_{3'}) = (A_6 \cap A_5) = \left(\frac{\partial}{\partial 6}\frac{\delta A}{\delta 5}\right) \cap \left(\frac{\partial}{\partial 5}\frac{\delta A}{\delta 6}\right) =$$

$$= ([1\ 2\ 3]\ [3\ 4]) \cap ([1\ 3]\ [2\ 3\ 4]) =$$

$$= \begin{bmatrix} 1 & 1 & 2 & 2 & 3 \\ 3 & 4 & 3 & 4 & 4 \end{bmatrix} \cap \begin{bmatrix} 1 & 1 & 1 & 2 & 3 \\ 2 & 3 & 4 & 3 & 4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 2 & 3 \\ 3 & 4 & 3 & 4 \end{bmatrix};$$

$$A_{2'3'} = A_{65} = \frac{\partial^2 A}{\partial 6\, \partial 5} = [1\ \ 2\ \ 4],$$

следовательно.

$$y'_1 = \frac{\det\limits_{Y} (A_{2'} \cap A_{3'})}{\det\limits_{Y} A_{2'3'}} = \frac{y_1 (y_3 + y_4) + y_3 (y_2 + y_4)}{y_1 + y_2 + y_4}.$$

Для расчета адмитанса $y_{2'}$ рассчитаем конъюнкцию

$$(A_1 \cap A_{3'}) = (A_3 \cap A_5) = \left(\frac{\partial}{\partial 3}\frac{\delta A}{\delta 5}\right) \cap \left(\frac{\partial}{\partial 5}\frac{\delta A}{\delta 3}\right) =$$
$$= ([1\ 4\ 6]\ [2\ 6]) \cap ([2\ 4\ 6]\ [1\ 6]) =$$
$$= \begin{bmatrix} 1 & 1 & 2 & 4 & 2 \\ 2 & 6 & 4 & 6 & 6 \end{bmatrix} \cap \begin{bmatrix} 1 & 1 & 1 & 2 & 4 \\ 2 & 4 & 6 & 6 & 6 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 4 & 2 \\ 2 & 6 & 6 & 6 \end{bmatrix}.$$

Таким образом,

$$y_{2'} = \frac{y_1(y_2 + y_6) + y_6(y_2 + y_4)}{y_1 + y_2 + y_4}$$

Аналогично рассчитаем $y_{3'}$

$$y_{3'} = \frac{y_4(y_2 + y_5) + y_5(y_1 + y_3)}{y_1 + y_2 + y_4}$$

Пример 12. Для модуль-схемы (рис. 22, *а)* построить схему замещения с тремя узлами, показанную на рис. 22, *б,* рассчитать адмитансы ветвей этой схемы замещения, построить схемы замещения отдельных модулей исходной схемы.

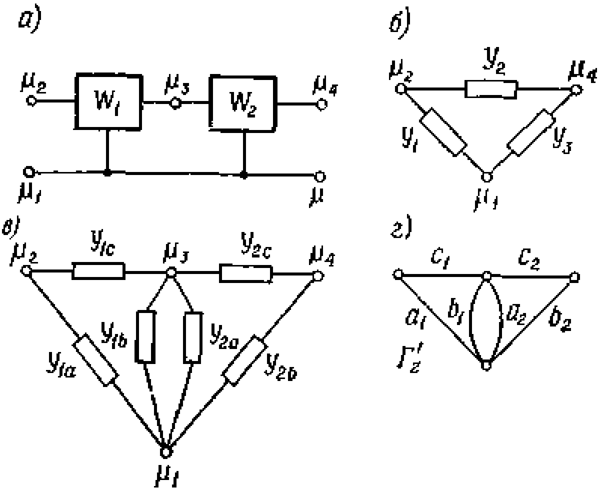


Рис. 22.

Схема замещения изображена на рис. 6.22, *в;* граф Γ'_z ее представлен на рис. 22, *г.*

В соответствии с формулой (57) рассчитаем адмитансы отдельных ветвей схемы замещений (рис. 22, *в)*

$$
\begin{aligned}
y_{1a} &= \frac{1}{2}(Y_{1ab} + Y_{1ac} - Y_{1ca}),\\
y_{1b} &= \frac{1}{2}(Y_{1ba} + Y_{1bc} - Y_{1ab}),\\
y_{1c} &= \frac{1}{2}(Y_{1ca} + Y_{1cb} - Y_{1ac}),\\
y_{2a} &= \frac{1}{2}(Y_{2ab} + Y_{2ac} - Y_{2ca}),\\
y_{2b} &= \frac{1}{2}(Y_{2ba} + Y_{2bc} - Y_{2ab}),\\
y_{2c} &= \frac{1}{2}(Y_{2ca} + Y_{2cb} - Y_{2ac});
\end{aligned}
\qquad \text{(a)}
$$

Y_{1ab} — адмитанс многополюсника W_1 с замкнутым путем b_1, измеренный относительно пути a_1,

…………………………………………………………………………….

Y_{1ab} — адмитанс многополюсника W_2 с замкнутым путем c_2, измеренный относительно пути a_2.

Теперь рассчитаем адмитансы ветвей схемы замещения (рис. 22, б).

Для этого вычислим блочную группу графа $Г'_z$ (рис. 22, г)

$$A'_z = [a_1\ c_1]\ [b_1\ c_1\ a_2\ c_2]\ [b_2\ c_2].$$

Найдем конъюнкцию

$$
\begin{aligned}
\frac{\partial A'_z}{\partial b_2} \cap \left(\frac{\partial A'_z}{\partial b_2} + \frac{\partial A'_z}{\partial a_1}\right) &= \frac{\partial A'_z}{\partial b_2} + \left\{\left(\frac{\partial}{\partial b_2}\frac{\delta A'_z}{\delta a_1}\right) \cap \left(\frac{\partial}{\partial a_1}\frac{\delta A'_z}{\delta b_2}\right)\right\} =\\
&= [a_1\ c_1]\ [b_1\ c_1\ a_2\ c_2] + \{([c_1]\ [b_1\ a_2\ c_2]) \cap\\
&\cap ([c_2]\ [b_1\ c_1\ a_2])\} =\\
&= \begin{bmatrix} a_1 & a_1 & a_1 & a_1 & c_1 & c_1 \\ b_1 & c_1 & a_2 & c_2 & b_1 & a_2 \end{bmatrix},
\end{aligned}
$$

$$\frac{\partial^2 A'_z}{\partial a_1 \partial b_2} = [b_1\ c_1\ a_2\ c_2].$$

В результате получаем

$$y_1 = y_{1a} + \frac{y_{1c}(y_{1b} + y_{2a})}{y_{1b} + y_{1c} + y_{2a} + y_{2c}}$$

Подставив выражение (а) в последнюю формулу, получим

$$y_1=\frac{1}{2}(Y_{1ab}+Y_{1ac}-Y_{1ca})+$$
$$+\frac{(Y_{1ca}+Y_{1cb}-Y_{1bc})(Y_{1ba}+Y_{1bc}-Y_{1ab}+Y_{2ab}+Y_{2ac}-Y_{2ca})}{4(Y_{1ca}+Y_{2ca})}.$$

Аналогично рассчитываются адмитансы y_2 и y_3.

2.3.6. Преобразование активных модуль-схем

Рассмотрим активную блок-схему с выделенными v_z узлами, содержащую независимые источники питания. Эта схема может быть самостоятельной цепью, отдельным многополюсником или частью большой схемы. Не учитывая внутренней структуры схемы, допустим, что известны ее внешние параметры, т. е. межузловые напряжения на ветвях любого дерева, построенного на выделенных узлах этой схемы, токи выделенных узлов схемы, а также входные импедансы или адмитансы схемы при различных вариантах замыкания выделенных узлов.

Заметим, что если известны межузловые напряжения на ветвях произвольного дерева, построенного на выделенных узлах схемы, то можно легко определить все остальные межузловые напряжения этой схемы.

Схемой замещения рассматриваемой модуль-схемы будет схема с v_z узлами, построенная из $v_z(v_z — 1)/2$ ветвей, образующих полный многоугольник.

Адмитансы этих ветвей рассчитываем методами, описанными в предыдущем разделе. Токи узлов схемы замещения должны быть равны токам в узлах рассматриваемой модуль-схемы, также как и межузловые напряжения схемы замещения должны быть равны соответствующим межузловым напряжениям модуль-схемы. Из второго закона Кирхгофа следует, что достаточно, чтобы равенство соответствующих напряжений выполнялось только для $v_z — 1$ межузловых напряжений. Таким образом, для расчета напряжений источников напряжения схемы замещения достаточно составить $v_z — 1$ независимых уравнений. Отсюда вытекает, что схема замещения должна содержать не более $v_z — 1$ источников напряжения. Эти источники не могут быть размещены в ветвях, образующих контур, т. е. они должны находиться в ветвях, образующих произвольное дерево схемы замещения.

Для расчета напряжений этих источников составляем $v_z — 1$ уравнений для токов в $v_z — 1$ узлах схемы замещения. Для любого узла μ схемы замещения (рис. 23) уравнение для токов имеет вид

$$\sum_{i=1}^{v_{z-1}} (E_i - U_i) y_i + I_\mu = 0 \qquad (59)$$

или

$$\sum_{i=1}^{v_{z-1}} E_i y_i = \sum_{i=1}^{v_{z-1}} U_i y_i - I_\mu \qquad (60)$$

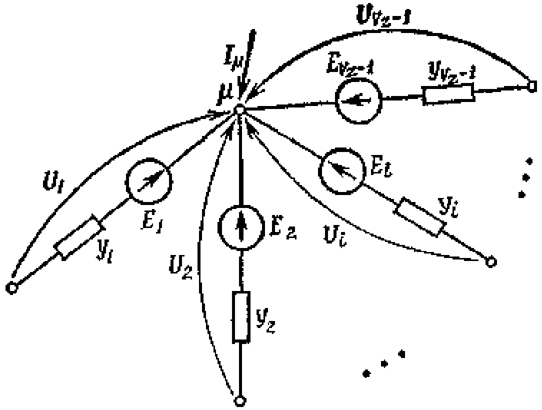


Рис. 23. Схема части замещающей цепи с узлом μ.

Знаки в приведенных формулах справедливы для направлений напряжений и тока I_μ в узел μ (рис. 23). При противоположном направлении напряжений или тока в вышеприведенных формулах нужно изменить знак.

После решения системы этих линейных уравнений получаем искомые напряжения E_i источников схемы замещения.

Замечание. Для упрощения расчета источники следует разместить в ветвях, инцидентных одному узлу схемы замещения.

Тогда можно непосредственно составить выражение для каждого искомого напряжения источника. Способ расчета напряжений источников в схеме замещения иллюстрирует следующий пример.

Пример. 13. Рассчитать напряжения источников схемы замещения (рис. 24, *б)* по известным напряжениям между узлами модуль-схемы рис. 24, *а*.

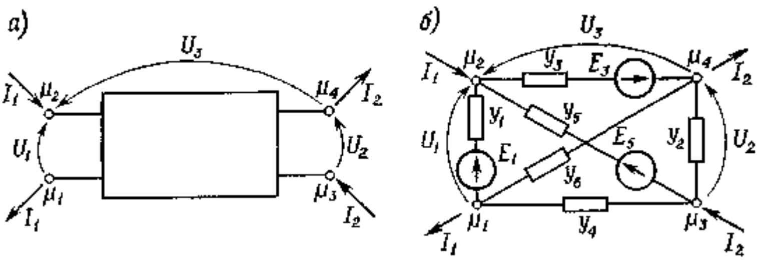


Рис. 24.

Составим уравнение для тока узла μ_1 схемы замещения:

$- E_1 y_1 = - U_1 y_1 + (U_3 – U_1) y_0 + (U_2 + U_3 – U_1) y_4 + I_1.$

Из этого уравнения следует

$E_1 = U_1 - (U_3 – U_1)(y_0/ y_1) -(U_2 + U_3 – U_1) (y_4/ y_1) -(I_1/ y_1).$

Аналогично, составляя уравнения для тока узлов μ_3 и μ_4, получим

$$E_5 = U_2 + U_3 + U_2 \frac{y_2}{y_5} + (U_2 + U_3 - U_1) \frac{y_4}{y_5} + \frac{I_2}{y_5},$$

$$E_3 = -U_3 + U_2 \frac{y_2}{y_3} - (U_3 - U_1) \frac{y_6}{y_3} - \frac{I_2}{y_3}.$$

Значения адмитансов ветвей схемы замещения $y_1, y_2, \ldots, y_6$ рассчитаем любым из методов, описанных в предыдущем разделе. Например, в соответствии с уравнением (57) имеем

$$y_1 = (1/2)(Y_{123} + Y_{124} — Y_{312}),$$

где Y_{123} — адмитанс модуль-схемы (рис. 24, *а)* при коротком замыкании узлов μ_2, μ_3 и μ_4, измеренный между узлами μ_1 и ($\mu_2 + \mu_3 + \mu_4$); Y_{124}— адмитанс модуль-схемы при коротком замыкании узлов μ_1, μ_3 и μ_4, измеренный между узлами μ_2 и ($\mu_1 + \mu_3 + \mu_4$); Y_{312}— адмитанс модуль-схемы при коротком замыкании узловых пар μ_l, μ_2 и μ_3, μ_4 измеренный менеду узлами ($\mu_l + \mu_2$) и ($\mu_3 + \mu_4$).

Аналогично рассчитываются адмитансы других ветвей схемы замещения.

2.4. Анализ электрических цепей с помощью переключающих схем и методом циклов

Изложенные в данном разделе методы анализа электрических схем основаны на соотношениях (14) и (28). Они дают возможность относительно просто определять параметры схем без решения систем уравнений, а также автоматизировать анализ электрических цепей с помощью ЭЦВМ.

2.4.1. Анализ электрических цепей с помощью переключающих схем

2.4.1.1. Переключающая схема

Описываемый ниже метод опирается на понятие переключающей схемы, называемой также нуль-импедансной схемой и кратко обозначаемой ZI.

Определение 1. *Переключающей схемой* будем называть схему, построенную из электрических ключей (контактных и безконтактных), соединенных между собой проводниками, с такой же самой топологической структурой, как рассматриваемая электрическая схема, причем каждой ветви рассматриваемой схемы соответствует один ключ в схеме ZI. На рис. 1 представлены для примера схема электрической цепи и соответствующая переключающая схема ZI.

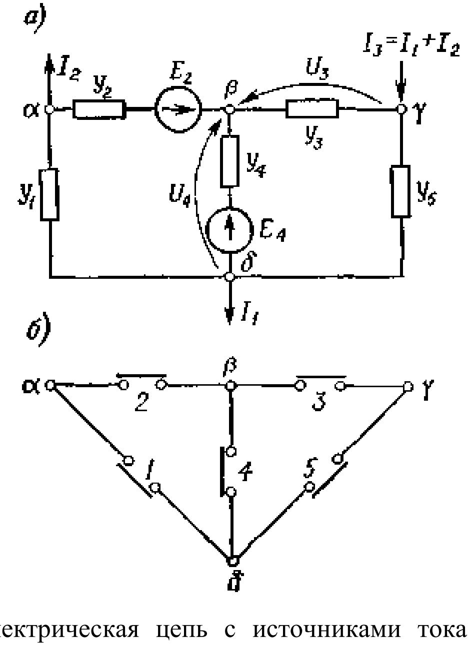


Рис. 1. Электрическая цепь с источниками тока и напряжения: *а)* схема цепи; *б)* схема переключающей цепи ZI.

На схеме рассматриваемой цепи (рис. 1, *а)* $y_1, y_2, \ldots, y_5$ обозначают проводимости отдельных ветвей этой цепи; α, β, γ и δ — узлы; E_2 и E_4 — э. д. с. источников напряжений; I_1, I_2 и $I_3 = I_4 + I_2$ — токи источников тока, а U_3 и U_4 — искомые напряжения ветвей y_3 и y_4.

На схеме цепи ZI (рис. 1, *б)* символы *1,2, . . ., 5* обозначают ключи. Ключ *1* соответствует ветви y_1 рассматриваемой электрической цепи, ключ *2* — ветви y_2 и т. д. Символы ; α, β, γ и δ обозначают узлы цепи ZI, соответствующие также обозначенным узлам рассматриваемой цепи.

С помощью ключей можно создать дерево в цепи ZI, т. е. гальванически соединить все узлы цепи ZI так, чтобы не было ни одного контура.

Например, если в цепи ZI, показанной на рис. 1, замкнуть ключи *2, 3* и *4* и разомкнуть ключи *1* и 5, то образуется дерево, показанное на рис. 2, *а*.

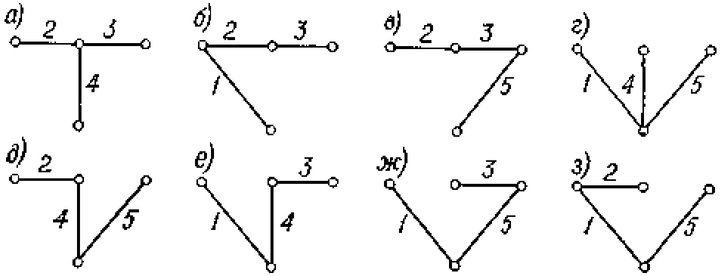


Рис. 2. Схемы деревьев цепи ZI (исходная цепь — рис. 1, *б).*

Это дерево имеет *v* — 1 замкнутых ключей, причем *v* обозначает число узлов цепи ZI.

В цепи ZI можно создать *Г* различных деревьев, причем

$$T = \det A_{v-1}^{v-1}, \tag{1}$$

где A_{v-1}^{v-1} — симметричная узловая матрица цепи ZI порядка *v* — 1, у которой диагональные элементы $a_{\gamma\gamma}$ равны числу проводов, соединяющихся в отдельных узлах, а остальные элементы $a_{\delta\varepsilon} = a_{\varepsilon\delta}$ равны числу ключей, соединяющих узлы δ и ε, имеют отрицательные знаки. Например, в цепи ZI, показанной на рис. 1, *б,* можно образовать

$$T = \det \begin{vmatrix} 2 & -1 & 0 \\ -1 & 3 & -1 \\ 0 & -1 & 2 \end{vmatrix} = 8$$

разных деревьев, схемы которых изображены на рис. 2.

В этих схемах замкнутые ключи представлены в виде отрезков, разомкнутые ключи не показаны.

Определение 2. *Величиной* D_d *дерева схемы* ZI будем называть произведение проводимостей ветвей исходной схемы, представленных замкнутыми ключами дерева d схемы ZI:

$$D_d = \prod_{i=1}^{\nu-1} y_{di}, \qquad (2)$$

где y_{di} — проводимость ветви рассматриваемой схемы, представленной замкнутым ключом в дереве d схемы ZI; ν — число узлов схемы ZI.

Например, величину D_a дерева, показанного на рис. 2, *а,* представим

$$D_a = y_2 y_3 y_4.$$

Если в схеме ZI замкнем два узла, то получим новую схему, которую назовем производной схемой ZI. В производной схеме ZI деревья создаются так же, как в первичной схеме ZI. На рис. 3 представлена производная схема ZI, полученная в результате замыкания узлов α и γ в первичной схеме ZI (рис. 1, б); на рис. 4 представлены изображения всех деревьев производной схемы.

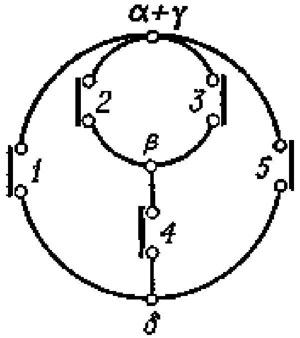


Рис. 3. Схема производной цепи ZI, образованной в результате замыкания узлов α и β исходной цепи рис. 1, *б.*

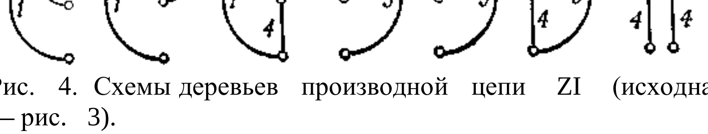


Рис. 4. Схемы деревьев производной цепи ZI (исходная цепь — рис. 3).

2.4.1.2. Анализ пассивных схем

Проводимость $Y_{\mu\nu}$, измеренную между узлами μ и ν рассматриваемой электрической цепи, можно рассчитать по формуле

$$Y_{\mu\nu}=\frac{\sum\limits_{d=1}^{T} D_d}{\sum\limits_{d'=1}^{T'} D'_{d'}}, \tag{3}$$

где D_d — величина дерева d схемы ZI; T — число деревьев схемы ZI; $D'_{d'}$ — величина дерева d' производной схемы ZI, полученной в результате замыкания в первичной схеме ZI узлов μ и ν; T' — число деревьев производной схемы ZI.

Например, проводимость $Y_{\alpha\gamma}$, измеренная между узлами α и γ электрической схемы (рис. 1, а), рассчитывается следующим образом на основании деревьев, изображенных на рис. 2 и 4:

$$Y_{\alpha\gamma}=(y_2y_3y_4+y_1y_2y_3+y_2y_3y_5+y_1y_4y_5+y_2y_4y_5+y_1y_4y_3+$$
$$+y_1y_3y_5+y_1y_2y_5):(y_1y_2+y_1y_3+y_1y_4+y_2y_5+y_3y_5+y_4y_5+y_2y_4+y_3y_4).$$

Для определения величин напряжений ветвей рассматриваемой электрической схемы следует в деревья схемы ZI включать замещающие источники, соответствующие источникам напряжений и токов электрической схемы. Например, на рис. 5 приведены все деревья схемы ZI, показанной на рис. 1, *б,* с включенными источниками.

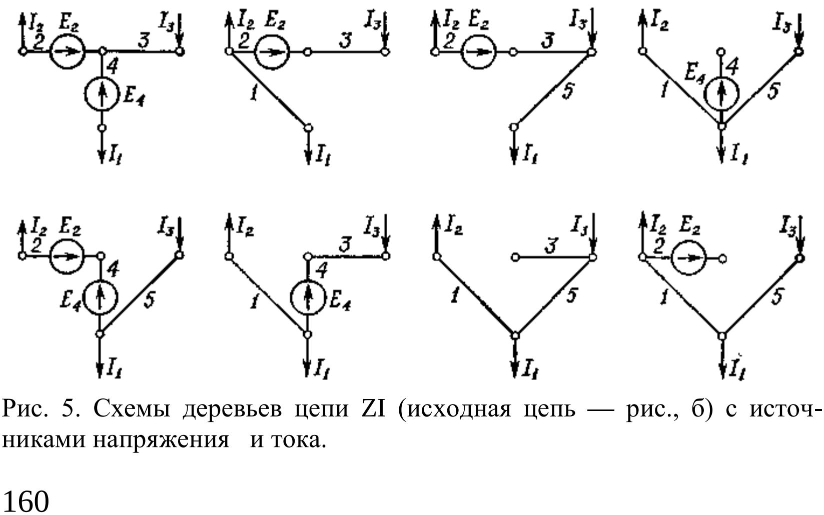


Рис. 5. Схемы деревьев цепи ZI (исходная цепь — рис., б) с источниками напряжения и тока.

На дереве с источниками определяем или измеряем межузловые напряжения и токи в ключах. Заметим, что эти напряжения и токи являются алгебраической суммой соответствующих напряжений и токов источников.

Напряжение U_k ветви k, соединяющей узлы μ и ν, рассматриваемой схемы (рис. 6), рассчитываем по формуле

$$U_k = \frac{\sum_{d=1}^{T}\left[D_d\left(u_{d\mu\nu}+\frac{i_{dk}}{y_{dk}}\right)\right]}{\sum_{d=1}^{T} D_d} = \frac{\sum_{d=1}^{T} D_d u_{d\mu\nu}+\sum_{d=1}^{T}\frac{\partial D_d}{\partial y_{dk}} i_{dk}}{\sum_{d=1}^{T} D_d}, \quad (4)$$

где $D_d = \prod_{i=1}^{\nu-1} y_{di}$ — величина дерева d схемы ZI; y_{di} — проводимость ветви, представленной замкнутым i ключом в дереве d; ν — число узлов схемы ZI; T — число деревьев схемы ZI; $u_{d\mu\nu}$ — напряжение между узлами μ и ν дерева d с включенными источниками; i_{dk} — ток в ключе k дерева d с включенными источниками (если в дереве d нет замкнутого ключа k, то $i_{dk} = 0$); y_{dk} — проводимость ветви k.

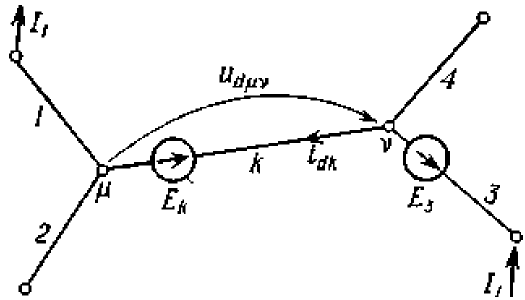


Рис. 6. Схема части дерева D_d с независимыми источниками и обозначением напряжения и тока k-й ветви (выключатель); в этом случае $u_{d\mu\nu} = E_k$; $i_{dk} = I_1$.

Например, напряжение U_3 ветви электрической схемы (рис. 1, *а)* рассчитываем по формуле, полученной из рассмотрения деревьев с источниками, представленных на рис. 5:

$$U_3 = (y_1y_4y_5E_4 + y_2y_4y_5E_4 + y_1y_2y_5E_2 - y_2y_4I_3 - y_1y_2I_3 - \\ - y_2y_5I_2 - y_1y_4I_3) : (y_2y_3y_4 + y_1y_2y_3 + y_2y_3y_5 + y_1y_4y_5 + \\ + y_2y_4y_5 + y_1y_3y_4 + y_1y_3y_5 + y_1y_2y_5).$$

Аналогично формула для напряжения U_4 ветви y_4 (рис. 1, *а),* полученная из рассмотрения деревьев, изображенных на рисг. 5, имеет вид

$$U_4 = (y_2y_3y_4E_4 + y_1y_2y_3E_2 + y_1y_4y_5E_4 + y_2y_4y_5E_4 + \\ + y_1y_3y_4E_4 + y_1y_2y_5E_2 + y_2y_3I_1 - y_2y_5I_2 + y_1y_3I_3) : \\ : (y_2y_3y_4 + y_1y_2y_3 + y_2y_3y_5 + y_1y_4y_5 + y_2y_4y_5 + y_1y_3y_4 + \\ + y_1y_3y_5 + y_1y_2y_5).$$

Если в схеме присутствуют только источники э. д. с, то формула для напряжения между произвольными узлами μ и ν схемы имеет вид

$$U_{\mu\nu} = \frac{\sum\limits_{d=1}^{T} D_d u_{d\mu\nu}}{\sum\limits_{d=1}^{T} D_d}. \qquad (5)$$

Эта формула справедлива для узлов μ и ν, не соединенных непосредственно ветвью. Например, если принять, что в схеме (рис. 1, *а)* токи $I_1 = I_2 = I_3 = 0$, то напряжение между узлами α и γ будет равно (рис. 5)

$$U_{\alpha\gamma} = [(y_2y_3y_4 + y_1y_2y_3 + y_2y_3y_5) E_2 + y_2y_4y_5 (E_2 - E_4) + \\ + y_1y_3y_4 E_4] : (y_2y_3y_4 + + y_1y_2y_3 + y_2y_3y_5 + y_1y_4y_5 + y_2y_4y_5 + \\ + y_1y_3y_4 + y_1y_3y_5 + y_1y_2y_4).$$

2.4.1.3. Анализ активных цепей

Пользуясь переключающей схемой ZI, можно анализировать также активные схемы с управляемыми источниками.

Для схемы с управляемыми напряжением источниками напряжения (рис. 7, *а)* формула для расчета напряжения имеет вид

$$U = L/M, \qquad (6)$$

где

$$L = \sum_{p=1}^{T} D_p \{u_p + (-1)^{n_{pi}} u_p u_{pi} + (-1)^{n_{pj}} u_p u_{pj} + \\ + (-1)^{n_{pk}} u_p u_{pk} + \quad + (-1)^{n_{pij}} u_p u_{pi} u_{pj} + \\ + (-1)^{n_{pik}} u_p u_{pi} u_{pk} + (-1)^{n_{pjk}} u_p u_{pj} u_{pk} + \dots \\ \dots + (-1)^{n_{pijk}} u_p u_{pi} u_{pj} u_{pk} + \dots\}; \qquad (7)$$

$$M = \sum_{p=1}^{T} D_p \{1 + u_{pi} + u_{pj} + u_{pk} + \dots + (-1)^{n'_{pij}} u_{pi} u_{pj} +$$

$$+(-1)^{n'_{pik}}u_{pi}u_{pk}+(-1)^{n'_{pjk}}u_{pj}u_{pk}+\ldots$$
$$\ldots+(-1)^{n'_{pijk}}u_{pi}u_{pj}u_{pk}+\ldots\}; \qquad (8)$$

T — число деревьев схемы ZI; D_p — величина *р*-го дерева; u_p, u_{pi}, u_{pj}, u_{pk} — напряжения между соответствующими узлами дерева *р* с поочередным включением источников (рис. 7, *б)* (следует обратить внимание на противоположное направление источников, имитирующих управляемые источники); n_{pi} — число четных перестановок элементов множества $\{|E|, |s_i|\}$, полученного подстановкой в множество $\{|u_p|, |u_{pi}|\}$ соответствующих абсолютных величин напряжений, измеренных на дереве *р; n_{pij}* — число четных перестановок элементов множества $\{|E|, |s_i|, |s_j|\}$, полученного подстановкой в множество $\{|u_p|, |u_{pi}|, |u_{pj}|\}$ соответствующих абсолютных величин напряжений, измеренных на дереве *р; n'_{pij}* — число четных перестановок элементов множества $\{|s_i|, |s_j|\}$, полученного подстановкой в множество $\{|u_{pi}|, |u_{pj}|\}$ соответствующих абсолютных величин напряжений на дереве *р,* и т. д.

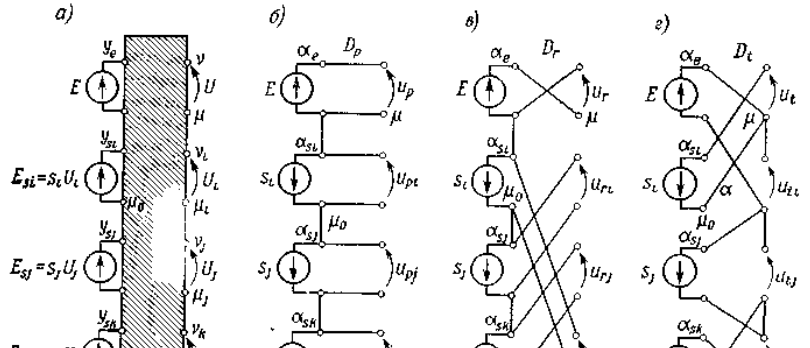


Рис. 7. Активная цепь с управляемыми напряжением зависимыми источниками напряжения: а) молуль-схема цепи; *б)* схема деревьев D_p, D_r, D_t цепи ZI с источниками (зависимые источники с обратной полярностью).

Отдельные слагаемые сумм в выражениях (7) и (8) записываются по приведенным ниже правилам. Введем следующие обозначения (рис. 7):

А — множество деревьев схемы ZI;

$A_{\mu\nu}$ — множество деревьев схемы ZI с замкнутыми узлами μ и ν;

$A_{\mu\nu,\mu_1\nu_1}$ — множество деревьев схемы ZI с замкнутыми парами узлов μ и ν, а также μ_1 и ν_1;

A^{α_e} — множество деревьев схемы ZI, в которой отключен от узла один зажим ключа α_e;

$A^{\alpha_e\alpha_{si}}$ — множество деревьев схемы ZI, в которой отключены от узлов отдельные зажимы ключей α_e и α_{si},

$A^{\alpha_e}{}_{\mu\nu}$ — множество деревьев схемы ZI с замкнутыми вершинами μ и ν, а также с отключенным зажимом ключа α_e и т. д.

Используя эти обозначения, можно доказать, что в выражениях (7) и (8)
слагаемое u_p появляется только для величины деревьев, содержащихся в множестве

$$A \mid A^{\alpha_e}{}_{\mu\nu};$$

слагаемое $u_p u_{pi}$ появляется только для величин деревьев, содержащихся в множестве

$$A \cap A^{\alpha_e\alpha_{si}}_{\mu\nu,\, \mu_i\nu_i};$$

слагаемое $u_p u_{pi} u_{pj}$ появляется только для величин деревьев, содержащихся в множестве

$$A \cap A^{\alpha_e\alpha_{si}\alpha_{sj}}_{\mu\nu,\, \mu_i\nu_i,\, \mu_j\nu_j};$$

слагаемое $u_p u_{pi} u_{pj} u_k$ появляется только для величин деревьев, содержащихся в множестве

$$A \cap A^{\alpha_e\alpha_{si}\alpha_{sj}\alpha_{sk}}_{\mu\nu,\, \mu_i\nu_i,\, \mu_j\nu_j,\, \mu_k\nu_k};$$

слагаемое u_{pi} появляется только для величин деревьев, содержащихся в множестве

$$A \cap A^{\alpha_{si}}_{\mu_i\nu_i};$$

слагаемое $u_{pi} u_{pj}$ появляется только для величин деревьев, содержащихся в множестве

$$A \cap A^{\alpha_{si}\alpha_{sj}}_{\mu_i\nu_i,\, \mu_j\nu_j}.$$

Это означает, что для каждого дерева D_i отдельные слагаемые сумм появляются только в случаях, когда соответствующие преобразования дерева D_i, приведенные выше, дают также деревья схемы ZI.

На основании формулы (6) проводимость $Y_{\mu\nu}$ схемы, измеренная между любыми узлами μ и ν, рассчитывается по формуле

$$Y_{\mu\nu} = \frac{M}{\partial M / \partial y_{\mu\nu}}, \tag{9}$$

где $y_{\mu\nu}$ — проводимость ветви, соединяющей узлы μ и ν.

Замечание. Формула (9) аналогична формуле (3), применяемой для схем, не содержащих управляемых источников, так как

$$\sum_{d'=1}^{T'} D'_{d'} = \frac{\partial}{\partial y_{\mu\nu}} \sum_{d=1}^{T} D_d.$$

Проводимость $Y^0_{\mu\nu}$ схемы с отключенной ветвью $y_{\mu\nu}$, таким образом, равна

$$Y^0_{\mu\nu} = \frac{M|_{y_{\mu\nu}=0}}{\partial M / \partial y_{\mu\nu}}. \tag{9a}$$

Чтобы проиллюстрировать формулы (6) — (8), обратим внимание на деревья *p, r* и *t* схемы ZI (рис. 7, *б — г).* Для этих деревьев (при условии что в схеме имеются только источники *E*, E_{si}, E_{sj} и E_{sk}) получим

$$L = \ldots + D_p\{E - Es_i - Es_j - Es_k + Es_is_j + Es_is_k + Es_js_k - Es_is_js_k\} + D_r\{-E + Es_is_js_k\} + D_t\{Es_i - Es_is_j + Es_is_k - Es_is_js_k\} + \ldots,$$

$$M = \ldots + D_p\{1 - s_i - s_j - s_k + s_is_j + s_is_k + s_js_k - s_is_js_k\} + D_r\{1 - s_is_js_k\} + D_t\{1 - s_j + s_k - s_js_k\} + \ldots.$$

Проводимость схемы между узлами μ_0 и μ (рис. 7, а) или относительно ветви y_α (рис. 7, г), согласно формуле (9), равна

$$Y_{\mu_0\mu} = \frac{M}{\ldots + \frac{\partial D_t}{\partial y_\alpha}\{1 - s_j + s_k - s_js_k\} + \ldots},$$

где, как следует из рис. 7, *б* и *в:*

$$\frac{\partial D_p}{\partial y_\alpha} = 0, \quad \frac{\partial D_r}{\partial y_\alpha} = 0$$

При анализе схем с управляемыми током источниками напряжений в соответствующие ветви деревьев схемы ZI вместо источников с напряжениями s_i, sj, s_k, ... следует включить источники с напряжениями

$$s'_i y_i, \quad s'_j y_j, \quad s'_k y_k, \ldots,$$

где s'_i, s'_j, s'_k, . . .— коэффициенты управления по току ($s'_i = E_{si}/I_i$, …); y_i. y_j, y_k, … проводимости ветвей управляющих токов I_i, I_j, I_k, ...

При анализе схем с управляемыми током источниками токов (рис. 8, *а)* ток *I* рассчитываем по формуле

$$I = \frac{L'}{M'} \tag{10}$$

где

$$L' = \sum_{p=1}^{T} D_p \{ i_p + (-1)^{n_{pi}} i_p i_{pi} + (-1)^{n_{pj}} i_p i_{pj} + (-1)^{n_{pk}} i_p i_{pk} + \dots$$
$$\dots + (-1)^{n_{pij}} i_p i_{pi} i_{pj} + (-1)^{n_{pik}} i_p i_{pi} i_{pk} + (-1)^{n_{pjk}} i_p i_{pj} i_{pk} + \dots$$
$$\dots + (-1)^{n_{pijk}} i_p i_{pi} i_{pj} i_{pk} + \dots \}; \tag{11}$$

$$M' = \sum_{p=1}^{T} D_p \{ 1 + i_{pi} + i_{pj} + i_{pk} + \dots + (-1)^{n'_{pij}} i_{pi} i_{pj} +$$
$$+ (-1)^{n'_{pik}} i_{pi} i_{pk} + (-1)^{n'_{pijk}} i_{pj} i_{pk} + \dots + (-1)^{n'_{pijk}} i_{pi} i_{pj} i_{pk} + \dots \}; \tag{12}$$

i_p, i_{pi}, i_{pj}, i_{pk} ,…… —токи, измеренные в ветвях дерева p (рис. 8, *б)* (следует обратить внимание на обратное направление источников тока s_i, s_j, s_k, . . ., имитирующих управляемые источники I_{si}, I_{sj}, I_{sk}, …); n_{pi} — число четных перестановок элементов множества $\{|I_0|, \ | \ s_i \ |\}$, полученного в результате подстановки в множество $\{| \ i_p \ |, \ | \ i_{pi} \ |\}$ абсолютных значений токов, измеренных в соответствующих ветвях дерева p и т. д. (аналогично формулам (7) и (8)).

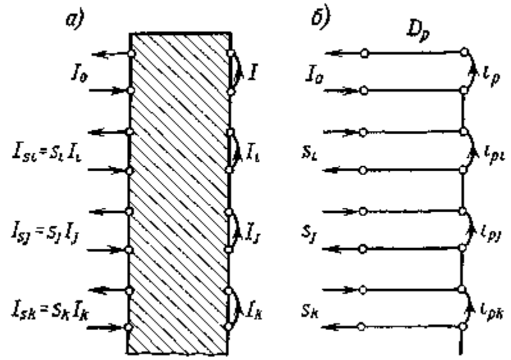


Рис. 8. Активная цепь с управляемыми током зависимыми источниками тока: *а)* модуль-схема; *б)* схема дерева D_p цепи ZI с источниками (зависимые источники с обратной полярностью).

Отдельные слагаемые сумм в выражениях (11) и (12) определяются так же, как в выражениях (7) и (8), но с той разницей, что в пересечениях множеств деревьев верхние указатели заменяем на нижние, а нижние на верхние.

Проводимость $Y_{\mu\nu}$ схемы между ее любыми узлами μ и ν рассчитываем по формуле

$$Y_{\mu\nu} = \frac{M'}{\partial M'/\partial y_{\mu\nu}}, \tag{13}$$

где $y_{\mu\nu}$ — проводимость ветви, соединяющей узлы μ и ν. Проводимость $Y^0_{\mu\nu}$ схемы с отключенной ветвью $y_{\mu\nu}$ равна

$$Y^0_{\mu\nu} = \frac{M'|_{y_{\mu\nu}=0}}{\partial M'/\partial y_{\mu\nu}}. \tag{13a}$$

Анализируя схему с управляемыми напряжением источниками тока, следует в соответствующие узлы деревьев схемы ZI вместо источников тока с токами s_i, s_j, s_k, . . . включать источники с токами

$$\frac{s'_i}{y_i},\quad \frac{s'_j}{y_j},\quad \frac{s_k}{y_k},\ \ldots,$$

где s'_i, s'_j, s'_k, . . .,— коэффициенты источников

$(I_{si}/U_i,\ I_{sj}/U_j,\ I_{sk}/U_k,\ \ldots)$;

y_i, y_j, y_k, . . .— проводимости управляющих ветвей.

Если рассматриваемая схема с управляемыми источниками тока подключена к источникам напряжения, то при применении данного метода такие источники следует заменить источниками тока так, как показано на рис. 9.

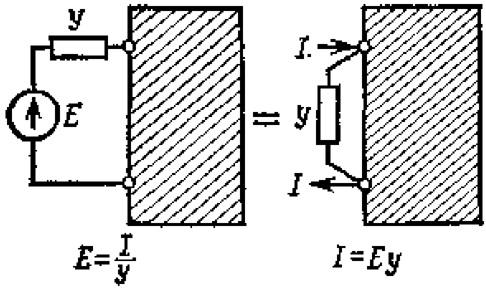


Рис. 9. Замена источника напряжения источником тока и наоборот.

В случае схем с управляемыми источниками напряжений это правило может также служить для замены источника тока источником напряжения.

Величины деревьев *D*,, в формулах (7), (8) и (11), (12) можно заменить значениями их дополнений, выраженными в виде произведений импедансов ветвей цепи, не принадлежащих отдельным деревьям. Если обозначить полученные таким образом знаменатели выражений соответственно через *M** (вместо *M*) и через *M'** (вместо *M'*), то импеданс $Z_{\mu\nu}$, измеренный между любыми двумя узлами μ и ν, равен

$$Z_{\mu\nu} = \frac{M^* |_{z_{\mu\nu}=0} z_{\mu\nu}}{M^*} \quad \text{или} \quad Z_{\mu\nu} = \frac{M'^* |_{z_{\mu\nu}=0} z_{\mu\nu}}{M'^*}, \tag{14}$$

где $z_{\mu\nu}$ — импеданс ветви, соединяющий узлы μ и ν.
Таким образом, импеданс $Z^0_{\mu\nu}$ цепи с отключенной ветвью $z_{\mu\nu}$ равен

$$Z^0_{\mu\nu} = \frac{M^* |_{z_{\mu\nu}=0}}{\partial M^* / \partial z_{\mu\nu}} \quad \text{или} \quad Z^0_{\mu\nu} = \frac{M'^* |_{z_{\mu\nu}=0}}{\partial M'^* / \partial z_{\mu\nu}}. \tag{15}$$

Замечание. Такой способ замены адмитанса на импеданс применим и для пассивных цепей (3 —5).
Проиллюстрируем на примерах метод анализа активных цепей.
Пример 1. Рассчитать параметры транзистора, включенного по схеме с общей базой (рис. 10, *а).* Схема цепи ZI показана на рис. 10, б, а ее деревья — на рис. 10, *в.*

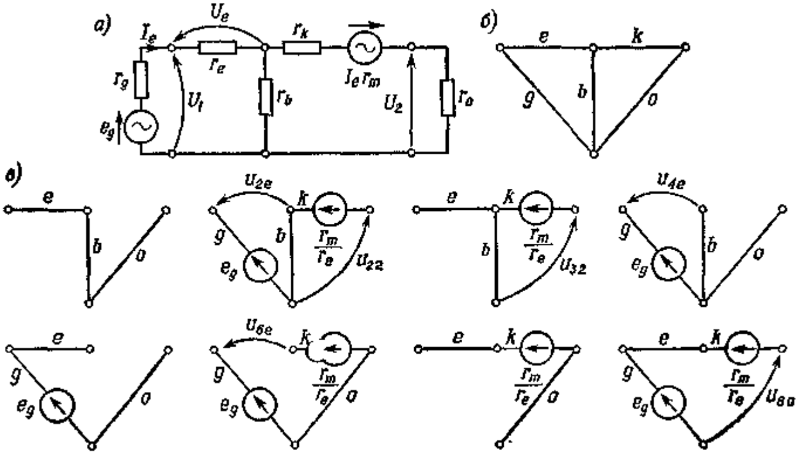


Рис. 10. Схема транзистора с общей базой: а) схема замещения; *б)* схема деревьев цепи ZI с включенными источниками.

Коэффициент управления напряжения равен

$$\frac{I_e r_m}{I_e r_e} = \frac{r_m}{r_e}.$$

Так как используем импедансы (резистансы) ветвей схемы, то вместо величин деревьев будем писать величины их дополнений.
На основании изображений деревьев (6) — (8) напишем

$$M^* = r_k r_g + r_e r_0 + r_g r_0 + r_e r_k + r_k r_b + r_e r_b \left(1 - \frac{r_m}{r_e}\right) + r_g r_b + r_b r_0 =$$

$$= r_k r_g + r_e r_0 + r_g r_0 + r_e r_k + r_k r_b + r_b (r_e - r_m) + r_g r_b + r_b r_0,$$

$$U_1 = \frac{r_e r_0 + r_e r_k + r_k r_b + r_b (r_e - r_m) + r_b r_0}{M^*} e_g,$$

$$U_2 = \frac{r_b r_0 + r_e r_0 (r_m / r_e)}{M^*} e_g = \frac{r_0 (r_b + r_m)}{M^*} e_g,$$

$$R^0_{\text{вх}} = \frac{M^* |_{r_g = 0}}{\partial M^* / \partial r_g} = \frac{r_e r_0 + r_e r_k + r_k r_b + r_b (r_e - r_m) + r_b r_0}{r_k + r_0 + r_b} =$$

$$= r_e + r_b \frac{r_k - r_m + r_0}{r_b + r_k + r_0},$$

$$R^0_{\text{вых}} = \frac{M^* |_{r_0 = 0}}{\partial M^* / \partial r_0} = \frac{r_k r_g + r_e r_k + r_k r_b + r_b (r_e - r_m) + r_g r_b}{r_e + r_g + r_b} =$$

$$= r_k - r_b \frac{r_m - r_g - r_e}{r_e + r_g + r_b}.$$

Из полученных величин можно рассчитать усиление напряжения и тока, мощности, сопротивления согласования и т. д.

Пример 2. Рассчитать напряжение U_3, а также входную $Y_{\text{вх}}$ и выходную $Y_{\text{вых}}$ проводимости схемы с индуктивной связью (рис. 11, *а)*.

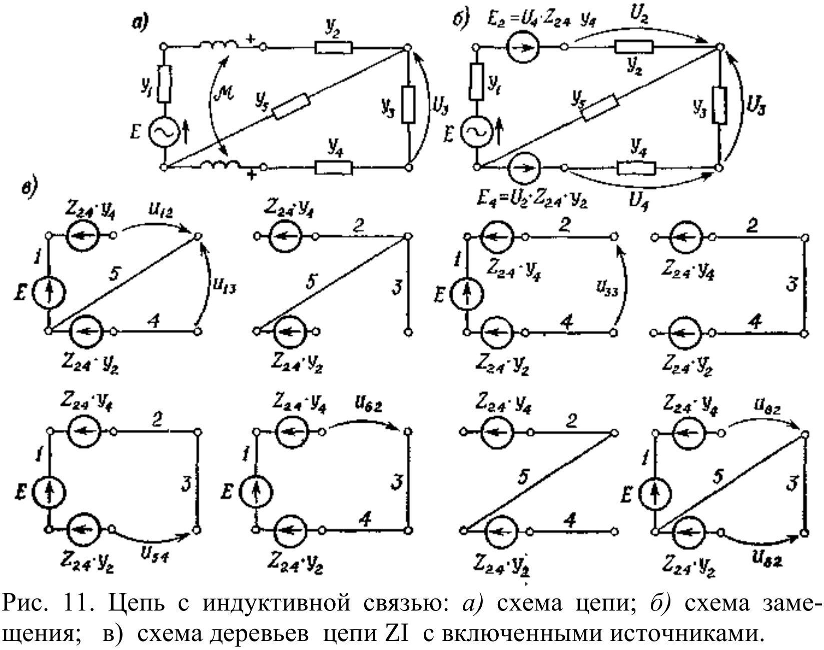


Рис. 11. Цепь с индуктивной связью: *а)* схема цепи; *б)* схема замещения; в) схема деревьев цепи ZI с включенными источниками.

Составим схему замещения с двумя источниками напряжения E_2 и E_4, управляемыми напряжениями U_4 и U_2 (рис. 11, б).
Коэффициенты источников равны

$$\frac{E_2}{U_4} = Z_{24} y_4, \quad \frac{E_4}{U_2} = Z_{24} y_2$$

где $Z_{24} = j\omega M$ (М— коэффициент взаимной индукции).
Из рассмотрения деревьев схемы ZI (рис. 11, в) (источники без сопротивления подключаем к соответствующим узлам схемы) можно написать

$$\begin{aligned} M &= y_1y_4y_5 + y_2y_3y_5 + y_1y_2y_4 + y_2y_3y_4 + y_1y_2y_3 + \\ &+ y_1y_3y_4 + y_2y_4y_5 + y_1y_3y_5 - Z_{24}y_4y_1y_2y_3 - \\ &- Z_{24}y_2y_1y_3y_4 - Z_{24}^2y_2y_4y_1y_3y_5 = (y_1 + y_2)(y_2y_4 + \\ &+ y_3y_5 + y_4y_5) + y_1y_2(y_3 + y_4) - y_1y_2y_3y_4(2 + Z_{24}y_5)Z_{24}, \end{aligned}$$

$$L = y_1y_2y_4E + Z_{24}y_2Ey_1y_4y_5 = y_1y_2y_4(1 + Z_{24}y_5)E,$$

$$U_3 = \frac{L}{M}.$$

Входная проводимость схемы без ветви y_1 источника E равна

$$Y_{\text{вх}} = \frac{M}{\partial M/\partial y_1} - y_1 = \frac{M\,|_{y_1=0}}{\partial M/\partial y_1} =$$

$$= \frac{y_2(y_3y_4 + y_3y_5 + y_4y_5)}{y_3y_4 + y_3y_5 + y_4y_5 + y_2(y_3 + y_4) - y_2y_3y_4(2 + Z_{24}y_5)Z_{24}},$$

а выходная проводимость схемы без ветви y_3

$$Y_{\text{вых}} = \frac{M}{\partial M/\partial y_3} - y_3 = \frac{M\,|_{y_3=0}}{\partial M/\partial y_3} =$$

$$= \frac{(y_1 + y_2)y_4y_5 + y_1y_2y_4}{(y_1 + y_2)(y_4 + y_5) + y_1y_2 - y_1y_2y_4(2 + Z_{24}y_5)Z_{24}}.$$

2.4.1.4 Образование деревьев схемы ZI

Описанный метод анализа пассивных и активных схем основан на образовании деревьев переключающей схемы ZI. При анализе простых схем деревья можно найти алгебраическим методом и после их изображения непосредственно установить зависимости для параметров этих схем. Образование деревьев можно автоматизировать при помощи генератора деревьев, действие которого опирается на следующие свойства дерева:

1) все узлы схемы ZI гальванически соединены замкнутыми ключами дерева;

2) дерево не содержит ни одного контура;

3) число замкнутых ключей в дереве равняется числу узлов схемы ZI, уменьшенному на единицу.

Принципиальная схема генератора деревьев представлена на рис. 12.

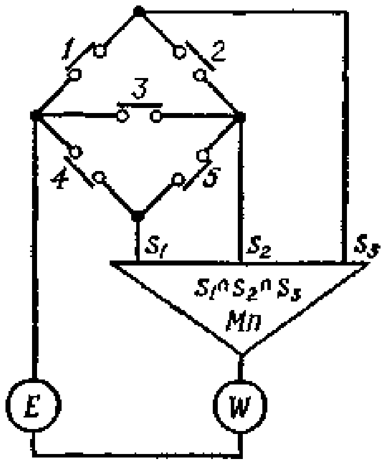


Рис. 12. Принципиальная схема генератора деревьев цепи ZI: *E* — источник питания; *Мп* — блок логического умножения; *W* — индикатор напряжения или тока; 1, 2, 3, 4, 5 — ключи.

Источник питания *Е* включается в произвольный узел схемы ZI. Оставшиеся узлы присоединяются через блок конъюнкции *Мп* сигналов S_i к индикатору напряжения или тока *W*. Условием образования дерева служит отклонение индикатора *W* при разомкнутых *М* ключах в схеме ZI, где *М* — цикломатическое число схемы ZI.

Включая в каждое образованное дерево схемы ZI источники, имитирующие источники питания и управляемые источники (в последнем случае с противоположной полярностью), а также осуществляя измерения соответствующих напряжений между узлами и токов в ключах дерева, можно описанным выше способом найти параметры схемы.

2.4.2. Анализ электрических схем методом циклов

Этот метод служит для расчета распределения тока в анализируемой электрической схеме, содержащей источники напряжения. Он основывается на определении всех элементарных циклов в графе рассматриваемой схемы, содержащих ребро, представляющее ветвь питания схемы, и ребро α_k, представляющее ветвь, в которой определяется значение тока I_k.

Число ребер w_t каждого из этих циклов не должно быть больше числа v вершин графа.

Выражение для тока I_k имеет вид

$$I_k = \frac{E}{\det\limits_{Y} A} \sum_{t=1}^{p} \det_{Y} \left(\prod_{i=1}^{w_t} [\alpha_{ti}] \prod_{r=1}^{v-w_t} [P_{tr}] \right), \qquad (16)$$

где I_k — ток в ветви схемы, представленной ребром α_k в графе схемы; E — э. д. с. источника напряжения; A — блочная группа графа схемы; $p \leq (2^M — 1)$ — количество элементарных циклов, содержащих $w_t \leq v$ ребер, среди них — ребро, представляющее ветвь с источником E, а также ребро α_k; M — цикломатическое число графа; α_{ti} — ребро элементарного цикла t; P_{tr} — блочная группа вершины графа μ_r, не принадлежащей циклу t (блочная группа вершины μ_r графа является однострочной блочной группой, элементы которой — обозначения всех инцидентных ребер с вершиной μ_r и не замкнутых в этой вершине); v — количество вершин графа анализируемой схемы; Y — множество проводимостей всех ветвей схемы.

Знаки отдельных слагаемых суммы (16) определяем следующим образом: если в цикле t направления E и I_k одинаковы, то слагаемое t суммы имеет знак плюс, в противном случае — знак минус.

Если в схеме много источников напряжения, то применяем принцип суперпозиции токов.

Элементарные циклы в планарном графе создаем путем суммирования блочных групп независимых циклов во всевозможных комбинациях (блочная группа цикла — однострочная блочная группа, элементами которой являются обозначения всех ребер этого цикла.). При этом опускаем суммы блочных групп, логическое произведение которых равно нулю.

В результате этого суммирования получаем не больше чем $(2^M — 1)$ блочных групп вместе с M блочными группами независимых циклов (M — цикломатическое число графа). Из найденного множества блочных групп элементарных циклов выбираем в формулу (16) все блочные группы A_t, удовлетворяющие условию

$$\alpha_k, \alpha_E \in A_t, \qquad (17)$$

где α_E — обозначение ребра графа, представляющего ветвь источника *E*.

Для иллюстрации изложенного метода расчета токов ветвей в электрических схемах рассмотрим два примера.

Пример 3. Рассчитать ток I_5 в ветви 5 электрической схемы, граф которой с источником э. д. с. E_1 и током I_5 приведен на рис. 13.

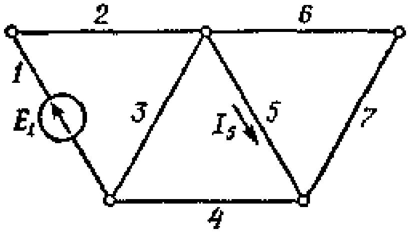


Рис. 13.

Определим блочные группы элементарных независимых циклов графа

A_1 = [1 2 3],

A_2 = [3 4 5],

A_3= [5 6 7],

а затем суммы этих блочных групп, опуская суммы, не содержащие ребра 1 и 5. В данном примере получим только одну сблочную группу

$A_1 + A_2$ = [1 2 4 51,

так как

$$5 \notin A_1, \quad 1 \notin A_2, \quad 1 \notin A_3, \quad A_1 \cap A_3 = 0,$$
$$1,\ 5 \notin (A_2 + A_3), \quad 5 \notin (A_1 + A_2 + A_3).$$

Блочная группа $A_1 + A_2$ представляет элементарный цикл.

Согласно формуле (16), имеем

$$I_5 = \frac{E_1}{\underset{Y}{\det A}} y_1 y_2 y_4 y_5 (y_6 + y_7),$$

так как цикл, составленный из ребер 1, 2, 4, 5, не содержит вершины графа, которой инцидентны ребра 6 и 7. В приведенном выражении *A* обозначает блочную группу графа.

Пример 4. Рассчитать I_7 в ветви 7 электрической схемы, граф которой с источниками напряжения E_1, E_2, . . ., E_6 и током I_7 представлены на рис. 14.

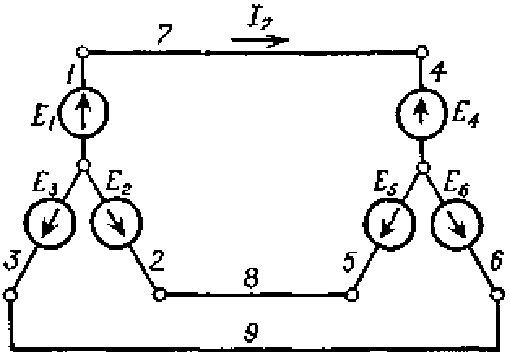


Рис. 14.

Определим блочную группу независимых элементарных циклов графа

A_1 = [1 2 4 5 7 8],

A_2 = [2 3 5 6 8 9].

Образуем сумму

$$A_1 + A_2 = [1\ 3\ 4\ 6\ 7\ 9], \quad A_1 \cap A_2 \neq 0.$$

На основании формулы (16) и принципа суперпозиции получим

$$I_7 = \frac{1}{\det\limits_{Y} A}\left[(E_1 + E_5 - E_2 - E_4)\, y_1y_2y_4y_5y_7y_8(y_3y_6 + y_3y_9 + y_6y_9) + (E_1 + E_6 - E_3 - E_4)\, y_1y_3y_4y_6y_7y_9(y_2y_5 + y_2y_8 + y_5y_8)\right],$$

так как

$$[3\,9]\ [6\,9] = \begin{bmatrix} 3\,3\,6 \\ 6\,9\,9 \end{bmatrix},$$

$$[2\,8]\ [5\,8] = \begin{bmatrix} 2\,2\,5 \\ 5\,8\,8 \end{bmatrix}$$

а блочная группа A_2 не содержит элемента 7.

3. Синтез организации структур

3.1. Постановка задачи синтеза структур

Понятие «синтез» структур в широком смысле слова близко по содержанию к понятию «организация». Задача синтеза структур состоит и том, чтобы по заданному функциональному назначению системы или по закону ее функционирования получить структуру по решению задач системы в виде некоторого описания.

Синтез структур нацелен на создание новых вариантов, а анализ используется для оценки этих вариантов, т. е. синтез и анализ выступают в диалектическом единстве.

При синтезе заданы допустимый набор используемых (функционирующих) элементов (злектрорадиоэлементов при синтезе электронных схем, набор балок и блочных конструкций при синтезе строительных сооружений и т. д.), возможные правила их соединения между собой и способы определения по синтезированной структуре системы функции, которую она реализует.

Напомним, что под *структурой системы* понимается набор составляющих ее элементов и связей между ними. Структура определяет, как устроена система, из каких физических частей она состоит и как эти части связаны друг с другом.

К проблемам синтеза относятся структура и физически реализовання система, причем под *физически реализованной системой* будем понимать материализованную совокупность соединенных между собой элементов, выполняющих заданные функции.

На структуру и физическую реализацию любой системы накладывается множество различных ограничений. При этом одна группа ограничений относится к методу решения задачи синтеза и охватывает такие вопросы, как наличие знаний, сроки и имеющиеся в распоряжении технические средства синтеза. Другая группа ограничений связана с требованиями ТЗ на параметры системы, с требованиями стандартов и процессов синтеза по решению задач системы и их подсистем. Третья группа ограничений формируется физическими принципами реализации закона функционирования системы и получения ее предельно желаемых характеристик. Дополнительные ограничения накладываются способами и формами взаимодействия системы с внешней средой, а также методами организации взаимодействия пользователя с системой в процессе ее функционирования.

При синтезе структуры системы выбирается или задается в ТЗ критерий оптимальности. В зависимости от характера и назначения системы в качестве критерия оптимальности может быть принята ее стоимость, КПД, потребляемая мощность или другой, более сложный критерий.

При синтезе важное значение имеет определение оптимальных вариантов структур и функций системы и подсистем, параметров схем, режимов функционирования системы и т. д. Под *оптимальным* будем понимать такой вариант структуры или функции системы, параметры которой удовлетворяют всем системным, конструктивным, технологическим, электрическим и экономическим требованиям ТЗ, а критерий оптимальности, описывающий качество структуры или функции, принимает наилучшее (минимальное или максимальное) значение.

Формирование оптимальных рекомендаций по решению задач синтеза становится возможным и достижимым как по срокам синтеза, так и по стоимости реализации синтезируемой структуры системы только при автоматизированном синтезе, когда появляется возможность синтезировать и исследовать множество вариантов структур и функций системы, а также проводить количественное изучение синтезируемых структур, которые ранее изучались лишь качественно.

Любому варианту системы соответствуют свои структура и функции. При автоматизированном синтезе для порождения множества альтернативных структур КП, эквивалентных по функциональному назначению, но различных по тактико-техническим характеристикам, необходима разработка математической модели структуры системы и системы в целом, представляющей собой формальное описание системы на принятом уровне детализации.

При оптимальном синтезе необходимо обосновать критерий оптимальности системы и определить множество показателей
$\theta = (\upsilon_1, ..., \upsilon_n)$, на которые наложено множество ограничений

$$\mathbf{V} = (V_1, ..., V_n).$$

Для решения задачи синтеза структуры по решению задач системы будем выделять некоторую совокупность независимых переменных $\mathbf{X} = (x_1, ..., x_m)$, фиксация значений которых определяет один из вариантов синтезируемой структуры и ее количественные характеристики, в том числе, значение критерия оптимальности, а также показателей, принятых в качестве ограничений.

Переменные $x_1, ..., x_m$ будем называть *переменными синтеза* и, в зависимости от физической природы системы, они имеют различную интерпретацию, в частности могут характеризовать количество элементов (подсистем) каждого типа в системе, указывать на включение или на невключение той или иной подсистемы в структуру системы, представлять геометрические размеры системы и т. д.

Критерий оптимальности $F(\mathbf{X}) = F(x_1, ..., x_m)$ и показали
$\theta_i(\mathbf{X}) = \theta_i(x_1, ..., x_m)$, на значения которых наложены ограничения, являются функциями независимых переменных $x_1, ..., x_m$.

В реальных задачах синтеза функция критерия оптимальности $F(\mathbf{X})$, функции ограничений $\theta_i(\mathbf{X})$, как правило, нелинейно зависят от значений множества переменных $\mathbf{X}$. Кроме того, во многих задачах синтеза на значения переменных накладывают условия целочисленности.

В формализованном виде задача синтеза структур заключается в определении значений независимых переменных $x_1,...,x_m$, при которых критерий оптимальности синтеза

$$F(\mathbf{X}) = F(x_1,x_2,...,x_m) \qquad (1)$$

принимает экстремальное (максимальное или минимальное) значение при условиях

$$\theta_i(x_1,...,x_m)\leq 0, \quad i = 1,2,\ldots, m; \qquad (2)$$

$$a_j\leq x_j\leq b_j, \quad j = 1,2,\ldots,n. \qquad (3)$$

Ограничения вида

$$\theta_i(x_1,...,x_m)\geq 0 \qquad (4)$$

путем умножения левой части неравенства на минус единицу могут быть приведены к выражению (2).

Ограничения вида (2) и (4) в своей исходной постановке могут быть заданными в виде уравнений; в случае задания системы неравенств всегда можно перейти от неравенств к уравнениям

$$\theta_i(x_1,...,x_m, x_{m+i})= 0 \qquad (5)$$

путем введения дополнительных переменных x_{m+i}, причем для неравенств вида (2) $x_{m+i}>0$, а для неравенств вида (4) $x_{m+i}<0$. На ряд переменных может быть наложено условие целочисленности, т. е. x_p — целые числа, $p=1,2,\ldots q$, $q\leq m$.

В задаче оптимизации, в которой ограничения имеют вид уравнений, количество ограничений n не может быть больше числа переменных m, т. е. $n\leq m$. Разность $m—n$ определяет число степеней свободы в данной задаче. Только $m—n$ переменных берутся произвольными, значения же остальных переменных определяются из системы ограничений. Если $m=n$, то число степеней свободы равно нулю и задача в этом случае является алгебраической. Оптимизация целевой функции $F(X)$ при этом не требуется.

Задача оптимального синтеза (формирования рекомендаций) в виде (1) — (5) представляет собой ***задачу математического программирования***. При этом если ***целевая функция (1) и все ограничения линейны, то задачу оптимизации называют задачей линейного программирования, если же целевая функция или хотя бы одно ограничение нелинейны, то задача оптимизации является задачей нелинейного программирования***.

Во многих задачах математического программирования некоторые переменные могут принимать лишь определенные дискретные значения (например, диаметр обмоточного провода, выбираемый из определенного сортамента, номиналы конденсаторов и т. д.) либо только целочисленные значения (например, число выпускаемых станков, самолетов и т. д.). ***В этом случае задача синтеза может быть сформулирована в терминах дискретного программирования.***

При синтезе структур и систем в целом с использованием моделей и методов математического программирования оказывается удобной геометрическая иллюстрация процесса формирования оптимальных рекомендаций. Рассмотрим геометрическую интерпретацию задачи математического программирования с линейной целевой функцией и с системой ограничений, образующих выпуклую оболочку области существования задачи оптимизации, т. е. пусть имеется система уравнений

$$\theta_i(x_1,\dots,x_m)\le 0, \quad i = 1,2,\dots, m; \tag{6}$$

$$a_j\le x_j\le b_j, \quad j = 1,2,\dots,n., \tag{7}$$

и целевая функция вида

$$F(\mathbf{X}) = \sum_{j=1}^{n} c_j x_j \tag{8}$$

Требуется найти значения переменных, удовлетворяющих ограничениям (6) и (7) и обращающих в минимум целевую функцию (9.8).

В данной задаче $m—n$ переменных могут быть выбраны произвольно, в частности взяты равными нулю. Это дает возможность проиллюстрировать задачу геометрически в $m—n$-мерном пространстве E^{m-n}. В этом пространстве каждой точке $\mathbf{X}^{(r)}$ соответствует совокупность чисел $x^r_{1},\dots, x^r_{m—n}$ равных проекции вектора, проведенного из начала координат в точку $\mathbf{X}^{(r)}$, на координатные оси пространства E^{m-n}.

Функция $F(x_1, \dots, x_m)$ в каждой точке пространства имеет определенное значение, следовательно, пространство E^{m-n} является скалярным полем критерия оптимальности $F(\mathbf{X})$ и функций ограничений $\theta_i(\mathbf{X})$. Функциям ограничений (6) соответствуют граничные гиперповерхности (в частном случае — гиперплоскости). Ограничениям (7) соответствуют гиперплоскости, выделяющие в пространстве определенную пространственную область. Если ограничения (6) и (7) представляют собой выпуклую область, то решения задачи оптимизации будут соответствовать такой точке этой области со скалярным полем критерия $F(\mathbf{X})$, в которой критерий $F(\mathbf{X})$ принимает минимальное значение. Поясним вышесказанное на примере.

Пример 1. Имеется следующая задача нелинейного программирования: минимизировать целевую функцию

$$F(\mathbf{X}) = x_1-x_2, \tag{9}$$

при ограничениях

$$x_1^2 + x_2^2 \le 4; \quad -2x_1 + x_2 \le 1; \quad x_1 - 2x_2 \le 1 \tag{10}$$

Введя дополнительные переменные x_3, x_4 и x_5, (10) превратим в систему уравнений, которую разрешим относительно этих введенных переменных. При этом получим

$$\begin{aligned} x_3 &= 4 - x_1^2 - x_2^2; \quad x_3 \geq 0; \\ x_4 &= 1 + 2x_1 - x_2; \quad x_4 \geq 0; \\ x_5 &= 1 - x_1 + 2x_2; \quad x_5 \geq 0. \end{aligned} \tag{11}$$

В данной задаче число переменных $m = 5$, число уравнений $n = 3$, тогда $m—n=2$, что дает возможность дать геометрическую интерпретацию задачи в пространстве E^2, т. е. на плоскости.

Поскольку все переменные должны быть положительными:

$$x_j \geq 0, \quad j=1,2,3,4,5 \tag{12}$$

каждое из неравенств (12) определяет некоторую допустимую область в пространстве E^2. Так, неравенство $x_1 \geq 0$ определяет верхнюю полуплоскость, неравенство $x_4 \geq 0$— полуплоскость, лежащую по одну сторону от прямой $2x_1—x_2+1=0$, а именно ту, которая содержит начало координат. Область, соответствующая $x_4<0$, является запрещенной, и ее удобно отметить (на рис. 1, *а* она отмечена штриховкой).

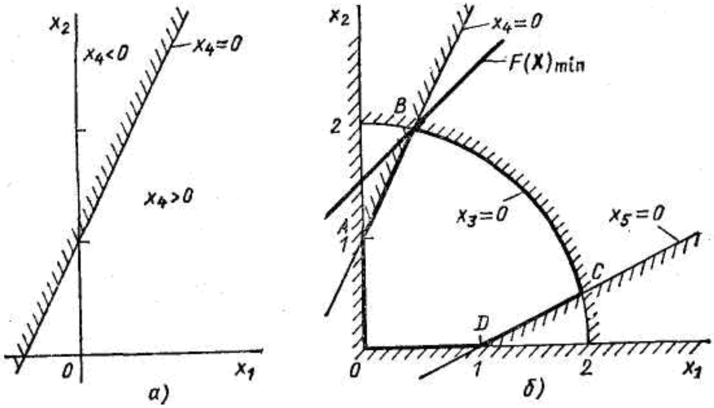


Рис. 1. Допустимая и запрещенная полуплоскости (*а*) и область существования задачи нелинейного программирования (*б*)

Подобные построения для всех x_i приведены на рис. 1, *б*. Областью существования задачи является область *OABCD*. В рассматриваемом примере функция (9) принимает минимальное значение на границе допустимой области в точке *B*.

Отметим ряд особенностей задачи нелинейного программирования с нелинейной целевой функцией. В общем случае заранее нельзя сказать о расположении точки, в которой функция $F(\mathbf{X})$ принимает максимальное или минимальное значение. Эта точка может находиться как на границе допустимой области, так и внутри нее. Функция $F(\mathbf{X})$ может достигнуть экстремального значения как в одной точке, так и на некотором множестве (гиперлинии или гиперповерхности).

Другая особенность задачи математического программирования состоит в том, что в общем случае нелинейная целевая функция $F(\mathbf{X})$ может иметь несколько локальных экстремумов в допустимой области, включая ее границу. Поэтому при оптимальном синтезе возникает проблема определения глобального экстремума задачи оптимизации, что в практических задачах синтеза нередко вызывает значительные трудности при вычислениях.

В общем виде задача нелинейного программирования пока не имеет строгого математического решения. Однако в связи с тем что данный класс задач довольно часто встречается в практических задачах синтезак, разработано большое число методов и эвристических алгоритмов решения конкретных задач нелинейного программирования.

3.2. Структурный синтез и параметрическая оптимизация структур

Задача синтеза структур включает в себя создание структуры системы и расчет ее параметров. Эти две части синтеза будем называть *структурным и параметрическим синтезом систем.*

Структурный синтез. Задача структурного синтеза заключается в поиске оптимальной или рациональной структуры (схемы) системы для реализации заданных функций в рамках выбранного принципа действия. Результаты структурного синтеза могут быть представлены в виде перечня элементов вместе с таблицей связей, схемы расположения элементовв в линейном евклидовом пространстве с указаниями их типов, схемы алгоритмов и т. п.

Задачи структурного синтеза относят к задачам синтеза, наиболее сложным с точки зрения возможностей формализации. Сложность формализованного синтеза структур заключается прежде всего в наличии большого числа факторов, влияющих на разновидности, свойства и параметры синтезируемых структур, а также в трудностях решения задач оптимизации большой размерности при высокой степени детализации описания синтезируемых структур.

Так, при синтезе структур по созданию структур специализированных ЭВМ и микропроцессорных систем возникают трудности решения ряда специфичных проблем, связанных с согласованием структур и алгоритмов функционирования ЭВМ и систем с характеристиками решаемых задач. Результатами структурного синтеза ЭВМ являются рекомендации по номенклатуре входящих в состав ЭВМ блоков, число блоков каждого типа, топология информационных и управляющих взаимосвязей между блоками, а также расписание функционирования каждого блока при организации вычислительного процесса.

Большая размерность задач синтеза структур систем делает целесообразным блочно-иерархический подход, при котором процесс синтеза разбивается на взаимосвязанные иерархические уровни. Формирование рекомендаций по структурному синтезу составляет существенную часть процесса синтеза и также организуется по блочно-иерархическому принципу. Это означает, что синтезируется не вся сложная структура целиком, а на каждом уровне в соответствии с выбранным способом декомпозиции синтезируются определенные функциональные блоки структуры с соответствующим уровнем детализации. Можно предложить различные способы классификации задач синтеза на стадии структурного синтеза. Так, в частности, в зависимости от стадии синтеза будем различать следующие процедуры структурного синтеза: выбор основных принципов функционирования системы, формирование рекомендации в рамках заданных принципов функционирования, выпуск технической документации. В зависимости от типа синтезируемых структур будем различать ***задачи одномерного, схемного и геометрического синтеза***. В зависимости от возможностей формализации будем различать ***задачи, в которых возможен полный перебор вариантов структур, задачи, которые не могут быть решены путем полного перебора за приемлемое время, задачи поиска вариантов синтеза структур в счетном множестве допустимых вариантов и задачи синтеза, решение которых является проблематичным***.

В самом общем случае для решения задачи необходим перебор вариантов синтеза структур, и сокращение перебора является актуальной проблемой. Иерархический подход уменьшает число вариантов синтеза структур на каждом уровне и делает решение задачи определения оптимальной структуры системы реальным.

Отметим существенное различие между задачами синтеза оптимальных структур и задачами анализа качества структур систем. В анализе необходимо убедиться, что решение существует, а численные методы анализа устойчивы. При структурном синтезе не гарантировано даже существование номинальной структуры, удовлетворяющей всем требованиям ТЗ на систему. Существующие и разрабатываемые ММ синтезируемых систем и их структур, как правило, оказываются довольно чувствительными к начальным условиям, к размерности задачи оптимизации, к виду целевых функций и ограничений. Поэтому необходимым условием для решения задач синтеза структур систем различной природы является использование методов и средств автоматизированного синтеза. Естественно, что формализованные модели и методы для САС, с одной стороны, должны характеризоваться высокой степенью общности и достоверности, а с другой стороны, должны быть разрешимыми с вычислительной точки зрения.

Существо получения ММ системы для решения задач структурного синтеза поясним на примерах компоновки, размещения и трассировки, довольно часто встречающихся в задачах конструирования ЭВА, распределения оборудования по производственным цехам, размещения цехов по территории завода, при проектировании линий электропередачи транспортных средств и т. п.

Пример 2. *Задача компоновки*. Под задачами компоновки понимают задачи разбиения множества $\mathbf{D} = \{d_1, ..., d_n\}$ из n элементов на ряд непересекающихся подмножеств $\mathbf{D}_k$, $k=1,2,\ldots,N$, чтобы при этом выполнялись заданные ограничения и достигался экстремум некоторой функции качества $F(\mathbf{X})$.

При заданном числе N подмножеств разбиения задача компоновки формулируется следующим образом:

$$F(\mathbf{X})\to\min;$$

$$\forall k,\ l\in\{1,\ 2,\ldots,\ N\}\ (\mathbf{D}_k\cap\mathbf{D}_l=\varnothing),\ \bigcup_{k=1}^{N}\mathbf{D}_k=\mathbf{D}, \qquad (13)$$

где $\mathbf{D}_k\subset\mathbf{D}$ — множество элементов, принадлежащих k-му подмножеству разбиения.

Общее число разбиений

$$\prod_{k=1}^{N} C^{n_k}_{n-\sum\limits_{q=1}^{k-1} n_q}=\frac{n!}{n_1!n_2!\ldots n_N!}$$

при условии, что мощность каждого подмножества $\mathbf{D}_k$ является заданной, т. е.

$$|\mathbf{D}_k|=n_k,\ \sum_{k=1}^{N} n_k=n.$$

Поэтому при реальном синтезе (при n>100) получить решение задачи компоновки путем перебора всех вариантов разбиения даже с использованием современных ЭВМ практически невозможно. Для уменьшения перебора задачу компоновки можно сформулировать в терминах целочисленного программирования.

Пусть требуется распределить n компонентов электронной схемы между N блоками таким образом, чтобы суммарное число связей между блоками было минимально. Введем вектор **X** переменных синтеза, компоненты x_{ik} (i=1,2,…,n, k=1,2,…,N) которого указывают на включение или невключение элемента $d_i \in \mathbf{D}$ в подмножество $\mathbf{D}_k$, т, е.

$$x_{ik} = \begin{cases} 1, \text{ если компонент } d_i \text{ включается в подмножество } \mathbf{D}_k; \\ 0 \text{ в противном случае.} \end{cases}$$

Содержательно функция качества $F(\mathbf{X})$ может характеризовать число связей между подмножествами $\mathbf{D}_i$ (при заданном числе N подмножеств), число подмножеств N, число типов подмножеств $\mathbf{D}_i$, определяемых данным разбиением. Очевидно, все перечисленные функции качества $F(\mathbf{X})$ следует минимизировать.

Пусть $F(\mathbf{X})$ характеризует общее число связей между подмножестиами $\mathbf{D}_k$, k=1,2,…,N. Тогда задача компоновки формулируется следующим образом: минимизировать целевую функцию

$$F(\mathbf{X}) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sum_{k=1}^{N-1} \sum_{l=k+1}^{N} x_{ik} x_{jl} \pi_{ij} \tag{14}$$

при ограничениях

$$\sum_{k=1}^{N} x_{ik} = 1, i = \overline{1, n}; \sum_{i=1}^{n} x_{ik} = n_k, \ k = \overline{1, N}; \tag{15}$$

$$\sum_{i=1}^{n} \vartheta_i^{(S)} x_{ik} \leqslant V_S^{(k)}. \tag{16}$$

Здесь π_{ij}— число связей между элементами d_i и d_j; $\upsilon^{(S)}_i$—значение параметра S для элемента d_i; $V^{(k)}_S$—ограничение по параметру S, накладываемое на подмножество $\mathbf{D}_k$, причем под параметром $\upsilon^{(S)}_i$ элемента d_i может подразумеваться любой показатель, подчиняющийся свойству аддитивности (объем, масса, стоимость, энергоемкость и т.д.),

Целевая функция (14) является квадратичной, поэтому задача компоновки, сформулированная в виде задачи (14) — (16), является квадратичной задачей дискретного программирования.

Пример 3. *Задача размещения*. После того как решена задача компоновки, требуется определенным образом расположить компоненты структуры, входящие в один блок. От того, как будут размещены микросхемы на определенной печатной плате (или любые элементы на плоскости), зависит длина соединительных проводников (связей), от которой в свою очередь зависят уровень помех и время распространения сигналов. Подобные задачи получили название задач размещения. В общем случае требуется найти такое размещение компонентов $d_1, d_2, \ldots, d_n$ на множестве $q_1, q_2, \ldots, q_m$ $(m>n)$ позиций монтажного пространства, при котором суммарная длина электрических соединений между компонентами была бы минимальной. Введем псевдобулевы переменные

$$x_{ik} = \begin{cases} 1, \text{ если компонент } d_i \text{ назначается на позицию } q_k; \\ 0 \text{ в противном случае.} \end{cases}$$

Тогда задача размещения может быть сформулирована в следующем виде: минимизировать целевую функцию

$$F(X) = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \sum_{k=1}^{m} \sum_{s=1}^{m} x_{ik} x_{js} l_{ks} m_{ij} \tag{17}$$

при ограничениях

$$\sum_{k=1}^{m} x_{ik} = 1, \; i = \overline{1, n}; \; \sum_{i=1}^{n} x_{ik} \leqslant 1, \; k = \overline{1, m}. \tag{18}$$

Здесь l_{ks} — расстояние между позициями q_k и q_s; m_{ij} — число связей между компонентами d_i и d_j.

Первое ограничение гарантирует, что каждая компонента разместится только на одной позиции; второе ограничение гарантирует, что на каждую позицию будет назначено не более одной компоненты.

Следует отметить, что среди известных критериев размещения наибольшее распространение получили минимум суммарной длины соединительных проводников, минимум наибольшей длины из всех длин соединительных проводников, минимум числа пересечений проводников и др.

Пример 4. *Задача трассировки*. Задача заключается в определении трасс соединений между компонентами схемы с учетом заданных ограничений, причем *трассой* называют множество связанных отрезков, соединяющих точки цепи. Задачи трассировки встречаются при конструировании печатных плат, при разработке систем водоснабжения, канализации, электроснабжения и т. д.

Критериями оптимальности в задачах трассировки могут выбираться минимум суммарной длины трасс, минимум числа соединений трасс длины больше заданной, минимум числа переходов между слоями в многослойных структурах и др.

Рассмотрим формальную постановку одной из разновидностей задачи трассировки, а именно — задачи построения связывающих сетей минимальной длины для цепей α_k. Соединяемые по цепи α_k точки образуют множество $\mathbf{U}_k$ мощностью $|U_k|=n_k$, в котором каждому элементу $u_k \in U_k$ в пространстве соответствует одна точка. Введем псевдобулевы переменные

$$x_{ij} = \begin{cases} 1, \text{ если ребро } (i, j) \text{ длиной } l_{ij} \text{ включается в связывающую сеть;} \\ 0 \text{ в противном случае.} \end{cases}$$

Задача построения минимальной связывающей сети имеет вид: минимизировать целевую функцию

$$L(\mathbf{x}) = \sum_{i=1}^{n_k-1} \sum_{j=i+1}^{n_k} l_{ij} x_{ij} \tag{19}$$

при ограничениях

$$\sum_{i=1}^{n_k} x_{ij} < K_0, \quad i = \overline{1, n_k}; \tag{20}$$

$$y_{ij}^{(s)} = \sum_{r=1}^{n_k} y_{ir}^{(s-1)} x_{rj}; \quad y_{ij}^{(1)} = x_{ij}; \tag{21}$$

$$\sum_{s=1}^{n_k-1} y_{ij}^{(s)} \geqslant 1, \quad i = \overline{1, n_k}, \quad j = \overline{2, n_k}, \quad s = \overline{2, n_k - 1}. \tag{22}$$

Здесь K_0 — максимально допустимое число соединений, инцидентных одной точке цепи; $y^{(s)}_{ij}$—вспомогательные переменные.

Условия (21), (22) гарантируют связность определяемой сети.

Приведенные примеры показывают, что во многих случаях задачи структурного синтеза являются экстремальными комбинаторными задачами, которые могут быть сведены к задачам дискретного программирования. Оценка трудоемкости получения точных решений задач этого класса позволяет сделать вывод, что при реальном синтезе получение точных решений либо невозможно, либо требует больших затрат машинного времени. Поэтому для структурного синтеза каждого класса систем необходима разработка специальных приближенных методов, позволяющих получать эффективные решения, близкие к оптимальным, а точные методы при этом служат для оценки качества синтеза.

Параметрический синтез. ***Задача параметрического синтеза заключается в формировании рекомендаций по определению наилучших значений параметров для выбранной структуры системы с учетом всех требований ТЗ.***

Функционирование любой системы подчиняется определенным физическим законам. Закон функционирования системы описывается аналитическими соотношениями между входными, внутренними и выходными переменными системы. Эти переменные связаны определенными соотношениями с переменными синтеза (формирования рекомендаций) **X**, под которыми понимаются внутренние переменные, допускающие варьирование. В процессе параметрического синтеза варьирование переменных синтеза (формирования рекомендаций) **X** ведет к изменению выходных параметров **Y** системы.

Для пояснения сущности задач параметрического синтеза будем использовать геометрическую интерпретацию, связанную с введением m-мерного пространства E^m пространства параметров синтеза (управляемых параметров) и k-мерного пространства E^k выходных параметров. Каждой точке пространства E^m и E^k соответствуют векторы **X** и **Y** значений переменных синтеза и выходных параметров соответствующего варианта системы.

Для постановки и решения задачи параметрического синтеза необходимо формирование целевой функции $F(\mathbf{X})$, отражающей качество функционирования системы. Векторный характер критериев оптимальности (мпогокритериальность) в задачах синтеза обусловливает сложность проблемы постановки задач оптимизации.

Формально задачу параметрического синтеза можно представить как задачу нахождения вектора $X \in E^m$, которым минимизирует целевую функцию

$$F(\mathbf{X}) \to \min \qquad (23)$$

при ограничениях

$$g(\mathbf{X})=0 \text{ и } h(\mathbf{X}) \leq 0, \qquad (24)$$

где $g(\mathbf{X})$ и $h(\mathbf{X})$ — векторные функции от **X**, описывающие систему ограничений на параметры синтеза **X**.

В качестве целевой функции целесообразно выбирать один из параметров, наиболее полно характеризующих свойства системы. Если частный критерий при синтезе выбрать затруднительно, то будем прибегать к формированию обобщенных критериев. В зависимости от целей синтеза и типов математических моделей систем целевые функции могут задаваться по-разному, для чего в САС должна быть предусмотрена ***библиотека целевых функций***. Примерами целевых функций могут служить целевая функция максимального модуля отклонения характеристик системы от заданных

$$F(\mathbf{X}) = \max_i \left| y_j^{(0)} - y_j^{(p)}(\mathbf{X}) \right|, \quad j = \overline{1,m}, \qquad (25)$$

и среднеквадратичная целевая функция

$$F(\mathbf{X}) = \sqrt{\sum_{j=1}^{m} \left(y_j^{(0)} - y_j^{(p)}(\mathbf{X}) \right)^2}\ , \qquad (26)$$

где $y_i^{(0)}$ — заданное значение параметра y_i; $y_i^{(p)}$ — реальное значение этого параметра.

Очевидно, что функции вида (25) и (26) в задачах оптимизации следует минимизировать. Оптимизация может быть осуществлена различными методами, включающими весьма сложные аналитические и численные математические процедуры.

Если в задачах оптимального синтеза все переменные состояний являются непрерывными, то для формирования рекомендаций по решению задач параметрического синтеза могут быть использованы методы решения задач нелинейного программирования, основанные на хорошо разработанных процедурах поиска экстремума функций. Однако не всегда все элементы в синтезируемых структурах могут принимать любые значения в пределах некоторой допустимой области. Это связано прежде всего со стандартизацией и унификацией объектов в различных проблемных областях. Так, в радиотехнике параметры резисторов и конденсаторов могут принимать только определенные значения из разрешенной шкалы номиналов, в строительстве плиты перекрытия, балки и другие комплектующие изделия имеют ряд определенных стандартных размеров. Кроме того, на параметры разрабатываемых объектов также накладывается ряд ограничений, учитывающих условия стандартизации и унификации. Так, в электротехнике и радиоэлектронике разрешается использовать только определенные значения питающих напряжений, в вычислительной технике существуют стандартные градации емкости устройств памяти и коммутационного оборудования.

Поэтому для решения задач оптимизации при синтезе структур с дискретными значениями параметров методы оптимизации непрерывных объектов непосредственно неприменимы. Эти задачи относятся к задачам дискретного программирования. Если при оптимизации часть параметров дискретна, а часть имеет непрерывный характер, то задача должна решаться методами частично дискретного программирования. Из-за недифференцируемости выходных параметров в задачах дискретного программирования довольно часто возникают трудности при вычислениях.

Рассмотрим пример задачи параметрического синтеза.

Пример 5. Формирование рекомендаций на проектирование трехстержиевой фермы. Цель проектирования — выбор конструкции трехстержневой фермы (рис. 2) минимальной массы.

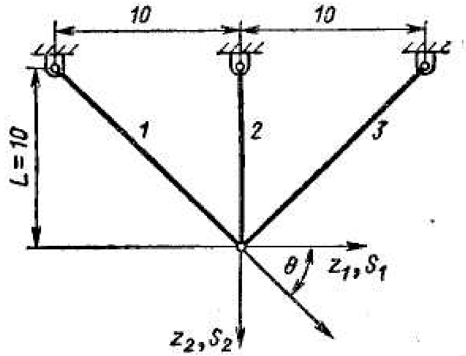


Рис. 2. Трехстержневая ферма

Проектирование сводится к выбору площадей поперечных сечений отдельных стержней x_1, x_2 и x_3 (переменные проектирования), так чтобы ферма была по возможности легкой и удовлетворялись ограничения на напряжение, устойчивость при продольном изгибе, смещение и размеры стержней.

Целевая функция *F*(**X**) представляет собой массу конструкции (при θ = 45°):

$$F(\mathbf{X}) = \rho g\left(10\sqrt{2}x_1 + 10x_2 + 10\sqrt{2}x_3\right), \qquad (27)$$

где *pg* — плотность материала, из которого изготовлена ферма.

Горизонтальное и вертикальное смещения z_1 и z_2 общего узла получены из линейных соотношений теории упругости. Для данной конструкции эти уравнения имеют вид

$$\mathbf{K}(x)\mathbf{Z} - \mathbf{S} = 0, \qquad (28)$$

где $\mathbf{Z}=[z_1, z_2]^T$;

$$\mathbf{K}(x) = \frac{\sqrt{2}E}{40}\begin{bmatrix}(x_1+x_3) & (x_1-x_3)\\ (x_1-x_3) & \left(x_1+x_3+2\sqrt{2}x_2\right)\end{bmatrix}, \mathbf{S} = \begin{bmatrix}P\cos\theta\\ P\sin\theta\end{bmatrix} \qquad (29)$$

— положительно определенная матрица жесткости и вектор нагрузки соответственно; *E* — модуль Юнга; θ — угол приложения нагрузки, отсчитываемый от горизонтали.

Вычисляя деформации ε в зависимости от смещения узла и применяя закон Гука, найдем напряжение в каждом стержне:

$$\sigma_1 = E(z_1+z_2)/20, \ \sigma_2 = Ez_2/10, \ \sigma_3 = E(z_2-z_1)/20. \qquad (30)$$

С помощью допустимых напряжений

$$\sigma_i^{(0)}, \ i=\overline{1, 3},$$

определенных для каждого стержня, ограничения на напряжения примут вид

$$\frac{E}{20}\,|z_1+z_2|-\sigma_1^{(0)}\leqslant 0; \tag{31}$$

$$\frac{E}{10}\,|z_2|-\sigma_2^{(0)}\leqslant 0; \tag{32}$$

$$\frac{E}{20}\,|z_2-z_1|-\sigma_3^{(0)}\leqslant 0. \tag{33}$$

Чтобы наложить ограничения по устойчивости, необходимо задать вид зависимости момента инерции от площади поперечного сечения для каждого стержня. Общей при инженерных расчетах является зависимость вида $I=\beta x^2$, где β — безразмерная постоянная. Подобная зависимость получается, если зафиксировать форму поперечного сечения и все его размеры менять в одинаковой пропорции. Осевые усилия имеют вид $\Phi_i=\sigma_i x_i$, $i=1,2,3$, растяжения стержней считаются положительными. Ограничения по устойчивости имеют вид

$$-\Phi_i\leqslant \pi^2 EI_i/L_i^2,\ i=\overline{1,3}. \tag{34}$$

Используя равенства (28), можно записать:

$$\begin{gathered}-E\,(z_1+z_2)\,x_1/20-\pi^2E\beta x_1^2/200\leqslant 0;\\ -Ez_2x_2/10-\pi^2E\beta x_2^2/100\leqslant 0;\\ -E(z_2-z_1)\,x_3/20-\pi^2E\beta x_3^2/200\leqslant 0.\end{gathered} \tag{35}$$

Ограничения на смещения могут быть сформулированы в виде

$$|z_1|-z_1^{(0)}\leqslant 0,\ |z_2|-z_2^{(0)}\leqslant 0, \tag{36}$$

где $z_1^{(0)}$ и $z_2^{(0)}$—заданные верхние границы для z_1 и z_2 соответственно. Наконец, требуется, чтобы площади поперечных сечений были неотрицательны, так что следует наложить ограничения

$$-x_1\le 0,\ -x_2\le 0,\ -x_3\le 0. \tag{37}$$

Теперь задачу проектирования можно рассматривать как задачу выбора таких переменных проектирования x_1, x_2 и x_3, которые минимизируют $F(\mathbf{X})$ и удовлетворяют ограничениям (28), (31) — (37).

Заметим, что смещения z_1 и z_2 играют в данной формулировке важную роль. Их определяют из уравнения для конструкции (28), коль скоро заданы переменные проектирования. Эти переменные представляют собой отклик системы на приложенную нагрузку, их называют *переменными состояния,* а уравнения (28) — *уравнениями состояния.*

3.3. Разновидности задач оптимизации синтеза структур

В задачах оптимального формирования рекомендаций по решению задач синтеза вектор переменных синтеза **X** = $\{x_1,\ldots \ldots,x_n)$ выбирают в результате определения экстремума целевой функции F(**X**) в допустимой области, заданной системой ограничений на параметры синтеза. В самом общем виде целевая функция и ограничения являются нелинейными функциями переменных синтеза **X**.

Задачи, в которых экстремум ищут в пределах неограниченного пространства переменных синтеза, относятся к задачам *безусловной оптимизации*. Найденные при чтом *экстремумы* называют *безусловными.* Наличие ограничений любого вида приводит к задачам *условной оптимизации*, решение которых дает *условный экстремум.*

При решении задач оптимизации первоначально проверяют условия, которым должен удовлетворять вектор переменных синтеза **X**, минимизирующий (максимизирующий) критерий качества $F(\mathbf{X})$. Эти условия проверяют для отыскания стационарных точек, среди которых находится искомый вектор **X**.

Функция $F(\mathbf{X})$, определенная в E^m, имеет абсолютный минимум в $\mathbf{X}^* \in E^m$, если

$$F(\mathbf{X}^*) \leq F(\mathbf{X}) \qquad (38)$$

для всех $\mathbf{X} \in E^m$. Минимум является строгим, если в (38) стоит знак строгого неравенства для $\mathbf{X} \neq \mathbf{X}^*$.

Функция $F(\mathbf{X})$ имеет в $\mathbf{X}^*$ относительный минимум, если существует $\varepsilon>0$ такое, что для любой точки $\mathbf{X} \in S_\varepsilon(\mathbf{X}^*)$ выполняется неравенство

$$F(\mathbf{X}) - F(\mathbf{X}^*) > 0, \qquad (39)$$

где $S_\varepsilon(\mathbf{X}^*)$ —окрестность точки $\mathbf{X}^*$. При определении максимума $F(\mathbf{X})$ (39) должно быть заменено на неравенство

$$F(\mathbf{X}) - F(\mathbf{X}^*) < 0.$$

Еели $F(\mathbf{X})$ имеет в точке $\mathbf{X}^*$ абсолютный минимум, то эта функция имеет там также и относительный минимум. Обратное не обязательно верно. Часто относительный минимум называют *локальным,* а абсолютный минимум — *глобальным.*

Рассмотрим необходимые и достаточные условия экстремума. Классические методы оптимизации используют тогда, когда известно аналитическое выражение функции $F(\mathbf{X})$ и известно, что она по крайней мере дважды дифференцируема по переменным синтеза. Тогда для определения экстремума будем использовать необходимые и достаточные условия безусловного экстремума. Эти условия легко получить с помощью разложения *F(**X**)* в окрестностях экстремальной точки **X*** в ряд Тейлора:

$$F(\mathbf{X}) - F(\mathbf{X}^*) = \sum_{i=1}^{m} \frac{\partial F(\mathbf{X}^*)}{\partial x_i}(x_i - x_i^*) + \frac{1}{2}\sum_{i=1}^{m}\sum_{k=1}^{m} \frac{\partial^2 F(\mathbf{X}^*)}{\partial x_i\, \partial x_k}(x_i - x_i^*)(x_k - x_k^*). \quad (40)$$

Если **X*** — точка максимума, то линейные члены в (40) равны нулю, тогда равны нулю составляющие вектора—градиента функции *F(**X**)*. Следовательно, необходимым условием экстремума является условие

$$\operatorname{grad} F(\mathbf{X}) = 0,$$

где

$$\operatorname{grad} F(\mathbf{X}) = \frac{\partial F(\mathbf{X})}{\partial \mathbf{X}} = \left(\frac{\partial F(\mathbf{X})}{\partial x_1}, \frac{\partial F(\mathbf{X})}{\partial x_2}, \ldots, \frac{\partial F(\mathbf{X})}{\partial x_m}\right).$$

Все точки, где выполняется это условие, называются *стационарными.*

Достаточное условие максимума на основании (39) и (40) заключается в том, что сумма членов со вторыми частными производными должна быть отрицательной при любых $\mathbf{X} \in S_\varepsilon(\mathbf{X}^*)$, т. е.

$$(\Delta\mathbf{X})^T \mathbf{\Gamma} (\Delta\mathbf{X}) < 0, \quad (41)$$

где $\Delta\mathbf{X} = \mathbf{X} - \mathbf{X}^*$ — вектор-столбец; $((\Delta\mathbf{X}^T = (\mathbf{X} - \mathbf{X}^*)^T$ — вектор-строка);

$$\mathbf{\Gamma} = \begin{bmatrix} \frac{\partial^2 F(\mathbf{X})}{\partial x_1^2}, & \frac{\partial^2 F(\mathbf{X})}{\partial x_1\, \partial x_2}, & \ldots, & \frac{\partial^2 F(\mathbf{X})}{\partial x_1\, \partial x_m} \\ \frac{\partial^2 F(\mathbf{X})}{\partial x_1\, \partial x_2}, & \frac{\partial^2 F(\mathbf{X})}{\partial x_2^2}, & \ldots, & \frac{\partial^2 F(\mathbf{X})}{\partial x_2\, \partial x_m} \\ \cdot & \cdot & \cdot & \cdot \\ \frac{\partial^2 F(\mathbf{X})}{\partial x_1\, \partial x_m}, & \frac{\partial^2 F(\mathbf{X})}{\partial x_2\, \partial x_m}, & \ldots, & \frac{\partial^2 F(\mathbf{X})}{\partial x_m^2} \end{bmatrix}$$

— матрица Гессе; Т — знак транспонирования матрицы.

Условие (41) есть достаточное условие максимума. Матрицу **Г**, удовлетворяющую условию (41) при любых $\Delta\mathbf{X}$, называют *отрицательно определенной,* а в случае $(\Delta\mathbf{X})^T \mathbf{\Gamma}(\Delta\mathbf{X}) > 0$ для любых $\Delta\mathbf{X}$ — *положительно определенной.* Поэтому достаточные условия экстремума можно представить как требование отрицательной определенности матрицы Гессе для максимума или положительной определенности для минимума в экстремальной точке.

Если достаточные условия не выполняются, то имеем не экстремальную, а седловую точку.

При наличии ограничений на переменные синтеза **X** решение задачи оптимизации ищут в некоторой допустимой области *S* согласно следующему направляющему принципу. Если $F(x_1...,x_m)$ является функцией нескольких переменных, определенных на допустимой области *S*, то максимальное значение $F(\mathbf{X})$, если оно существует, достигается в одной или более точках, которые могут принадлежать следующим множествам:

K_1 = {($x_{1,...,}x_m$) | ($x_{1,...,}x_m$) —точка в области *S*, в которой функция *F* стационарна};

K_2={($x_{1,...,}x_m$) | ($x_{1,...,}x_m$)— точка границы области *S*};

K_3={($x_1,...,x_m$) | —функция $F(\mathbf{X})$ не дифференцируема в точке ($x_1,...,x_m$)}.

На рис. 3 приведен пример геометрической интерпретации многоэкстремальной задачи оптимального синтеза.

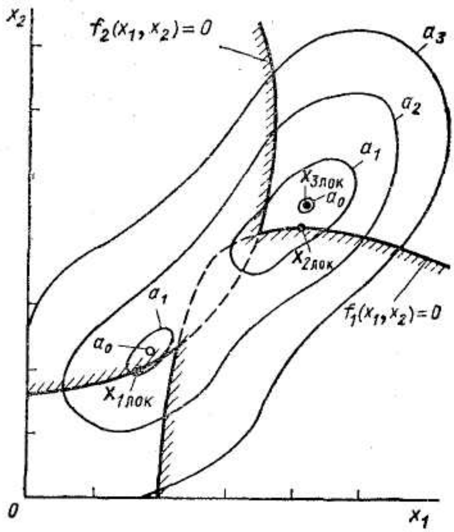


Рис. 3. Многоэкстремальная задача оптимального синтеза

На рисунке показаны линии равного уровня целевой функции $F(\mathbf{X})(a_3>a_2>a_1>a_0)$ и видны три локальных оптимума, которые находятся в областях, определяемых общим направляющим принципом (точки $\mathbf{X}_{1лок}$, $\mathbf{X}_{2лок}$, $\mathbf{X}_{3лок}$ являются точками локальных оптимумов, причем точка $\mathbf{X}_{3лок}$ совпадает с глобальным оптимумом).

К сожалению, отсутствуют формальные признаки многоэкстремальных ситуаций. Исключением являются задачи, где целевая функция выпуклая (в задачах минимизации) или вогнутая (при максимизации).

Напомним, что функцию $F(\mathbf{X})$ с числовыми значениями, определенными на выпуклом множестве S, называют *вогнутой,* если для любой пары точек $\mathbf{X}_1, \mathbf{X}_2 \in S$ и для всех чисел λ $(0 \leq \lambda \leq 1)$ выполняется неравенство $F(\lambda\mathbf{X}_1+(1-\lambda)\mathbf{X}_2) \geq \lambda F(\mathbf{X}_1)+(1-\lambda)F(\mathbf{X}_2)$. Если $F(\lambda\mathbf{X}_1+(1-\lambda)\mathbf{X}_2) \leq \lambda F(\mathbf{X}_1)+(1-\lambda)F(\mathbf{X}_2)$, то функцию $f(\mathbf{X})$ называют *выпуклой.* Если имеют место строгие неравенства, то говорят, что функция *строго вогнута* или *строго выпукла.* Для дважды дифференцируемой функции критерий вогнутости или выпуклости формируется следующим образом. Дифференцируемая функция $F(\mathbf{X})$ строго вогнута в некоторой окрестности точки $\mathbf{X}^{(0)} = (x_1^{(0)},..., x_m^{(0)})$, если выполняются условия

$$F_{11}(\mathbf{X}^{(0)}) < 0, \begin{vmatrix} F_{11}(\mathbf{X}^{(0)}) & F_{12}(\mathbf{X}^{(0)}) \\ F_{21}(\mathbf{X}^{(0)}) & F_{22}(\mathbf{X}^{(0)}) \end{vmatrix} > 0,$$

$$\begin{vmatrix} F_{11}(\mathbf{X}^{(0)}) & F_{12}(\mathbf{X}^{(0)}) & F_{13}(\mathbf{X}^{(0)}) \\ F_{21}(\mathbf{X}^{(0)}) & F_{22}(\mathbf{X}^{(0)}) & F_{23}(\mathbf{X}^{(0)}) \\ F_{31}(\mathbf{X}^{(0)}) & F_{32}(\mathbf{X}^{(0)}) & F_{33}(\mathbf{X}^{(0)}) \end{vmatrix} < 0,...,$$

т. е. если знаки этих определителей чередуются указанным образом. Здесь $F_{ij}(\mathbf{X}^{(0)})$—частная производная второго порядка, вычисленная в точке $\mathbf{X}^{(0)}$.

Функция $F(\mathbf{X})$ строго выпукла в малой окрестности точки $\mathbf{X}^{(0)}$, если все определители, указанные выше, положительны.

Достаточные условия для определения максимума или минимума формулируются следующим образом: для того, чтобы в точке $\mathbf{X}^{(0)}$ достигался внутренний локальный максимум, достаточно равенства нулю всех частных производных и строгой вогнутости функции в некоторой окрестности этой точки; для того чтобы в точке $\mathbf{X}^{(0)}$ достигался внутренний локальный минимум, достаточно, чтобы все частные производные обращались в нуль и чтобы в малой окрестности этой точки функция была строго выпуклой.

В большинстве задач синтеза при отсутствии аналитического задания целевых функций проверка $F(\mathbf{X})$ на выпуклость или вогнутость, как правило, невозможна, поэтому для решения задач оптимального синтеза следует использовать ***методы поисковой оптимизации*, *основанные на исследовании малой окрестности оптимальной точки в допустимой области***. Основные требования, предъявляемые к методу поиска,— высокая алгоритмическая надежность, приемлемые затраты машинного времени и требуемой памяти.

Методы поиска экстремума классифицируются по следующим признакам: в зависимости от характера экстремума существуют методы условной и безусловной, локальной и глобальной оптимизации; по числу переменных синтеза различают методы одномерного и многомерного поиска, а по характеру информации о виде целевой функции — методы нулевого, первого и второго порядков, причем в методах первого порядка используют градиент целевой функции, поэтому эти методы называются градиентными, в методах второго порядка применяют вторые производные, а в методах нулевого порядка производные не используют.

3.4. Показатели эффективности синтеза структур и выбор методов поиска экстремума

В задачах оптимального синтеза структур в самой общей постановке целевая функция и ограничения являются нелинейными функциями переменных синтеза структур.

В общем случае эти задачи многоэкстремальны, поэтому для их решения следует применять методы определения поиска глобального экстремума, которые включают в себя один из способов генерации начальных точек, а также поиска локальных оптимумов.

Поиск локального оптимума состоит из следующих этапов определения: направления движения к оптимуму, длины шага поиска, окончания поиска.

Алгоритмы поиска локального оптимума $\mathbf{X}^*$ являются, как правило, итеративными, т. е. порождают последовательность векторов $\{\mathbf{X}^{(k)}\}=\mathbf{X}_1, \mathbf{X}_2..... \mathbf{X}_k$, сходящуюся к вектору $\mathbf{X}^*$.

Будем говорить, что вектор $\mathbf{X}^*$ является пределом сходящейся последовательности $\{\mathbf{X}^{(k)}\}$, если для любого $\varepsilon>0$ найдется такой номер N, что при $k>N$ выполняется неравенство $|\mathbf{X}_k — \mathbf{X}^*|< \varepsilon$. Отсюда следует, что допустимая область S должна вместе с любой сходящейся последовательностью содержать и ее предел. Такую ***область*** будем называть ***замкнутой***. Примером замкнутой области может служить множество всех точек, удовлетворяющих ограничениям (2) и (3).

Множество всех точек пространства E^n, которые не содержатся в замкнутой области $S(E^n\backslash S)$, будем называть ***открытым***. В замкнутой области S, если она не совпадает со всем пространством E^n, всегда можно найти точки, в ε-окрестности которых имеются точки из $E^n\backslash S$.

Такие точки области будем называть ***граничными***. Множество всех граничных точек образует ***границу области*** *S*. В частности, если область *S* определяется условиями (2) и (3), его границу составляют те точки, в которых хотя бы одно из ограничений выполняется как ***строгое равенство***.

Эффективность методов поиска локального оптимума определяется ***скоростью их сходимости*** к **X***, а критериями оценки качества выбора направления являются:

- улучшение значения критерия оптимальности во вновь выбранной точке по сравнению с его величиной в данной точке;

- наиболее быстрое убывание (возрастание) критерия в окрестности данной точки;

- наиболее вероятное расположение экстремума с учетом кривизны гиперповерхности, представляющей критерий оптимальности.

Использование каждого из трех критериев выбора направления движения к оптимуму требует различного числа обращений к модели системы. Необоснованное усиление критерия выбора направления поиска может привести к резкому возрастанию числа обращений к модели системы, а ослабление — к беспорядочному блужданию в окрестности оптимума. В обоих случаях возрастают затраты машинного времени на синтез структур.

Проведем краткий анализ методов поиска экстремума. Особенности методов будем иллюстрировать примерами их применения к поиску экстремума функции $F(\mathbf{X})$ в двумерном пространстве переменных синтеза структур.

Методы безусловной оптимизации. Для решения задачи безусловной оптимизации будем использовать итерационные процессы вида

$$\mathbf{X}_k = \mathbf{X}_{k-1} + \alpha_k \Delta\mathbf{X}_k, \qquad (42)$$

где $\Delta\mathbf{X}_k$ — вектор, определяющий направление движения из точек $\mathbf{X}_{k-1}$, α_k — числовой множитель, значение которого определяет длину шага в направлении $\Delta\mathbf{X}_k$. Для большинства методов

$$\Delta\mathbf{X}_k = \mathbf{P}_k / \| \mathbf{P}_k \|. \qquad (43)$$

где **P** — вектор, указывающий направление поиска; $\|\mathbf{P}_k\|$— норма вектора **P**; k — индекс, обозначающий номер шага поиска.

Процесс (42) будет определен, если указаны способы построения вектора $\Delta\mathbf{X}_k$. и вычисления величины α_k на каждой итерации. От того, каким образом строится вектор $\Delta\mathbf{X}_k$ и определяется множитель α_k, непосредственно зависят свойства процесса: поведение функции $F(\mathbf{X})$ на элементах последовательности $\{\mathbf{X}^{(k)}\}$, сходимость последовательности к решению, скорость сходимости и др. В то же время различные способы построения вектора $\Delta\mathbf{X}_k$ и множителя α_k требуют различных затрат машинного времени и различной емкости оперативной памяти ЭВМ.

Чтобы приблизиться к точке **X***, естественно двигаться от точки $\mathbf{X}_{k-1}$ в одном из направлений убывания функции $F(\mathbf{X})$ (в направлении спуска). Если точка $\mathbf{X}_{k-1}$ не является точкой минимума или стационарной точкой, то существует бесконечно много векторов $\Delta\mathbf{X}_k$, определяющих направления спуска из точки $\mathbf{X}_{k-1}$, причем каждый из них определяется условием

$$(F'(\mathbf{X}_{k-1}), \Delta\mathbf{X}_k)<0,$$

вытекающим из (40).

Поскольку местонахождение точки **X*** неизвестно, процесс поиска экстремума может быть прекращен в точке $\mathbf{X}_k$, при этом число шагов r, разделяющих точки $\mathbf{X}_k$ и $\mathbf{X}_{k-1}$, определяют из условия $|\mathbf{X}_k — \mathbf{X}_{k-r}| <\varepsilon$. Следовательно, поиск прекращается, если расстояние, на которое продвинулась отображающая точка в пространстве переменных синтеза структур за последние r шагов, оказывается меньше заданного числа ε.

Прямые методы оценки направлений. Наиболее простым является *метод покоординатного спуска* (метод Гаусса — Зейделя). Направление поиска выбирают поочередно вдоль всех координатных осей, т. е. вектор $\mathbf{P}_k$ в (43) состоит из нулевых элементов за исключением одного, равного единице.

Пусть $\mathbf{X}_0 = (x_1^{(0)}, \ldots, x_n^{(0)})$ — начальная точка поиска (рис. 4, *а*).

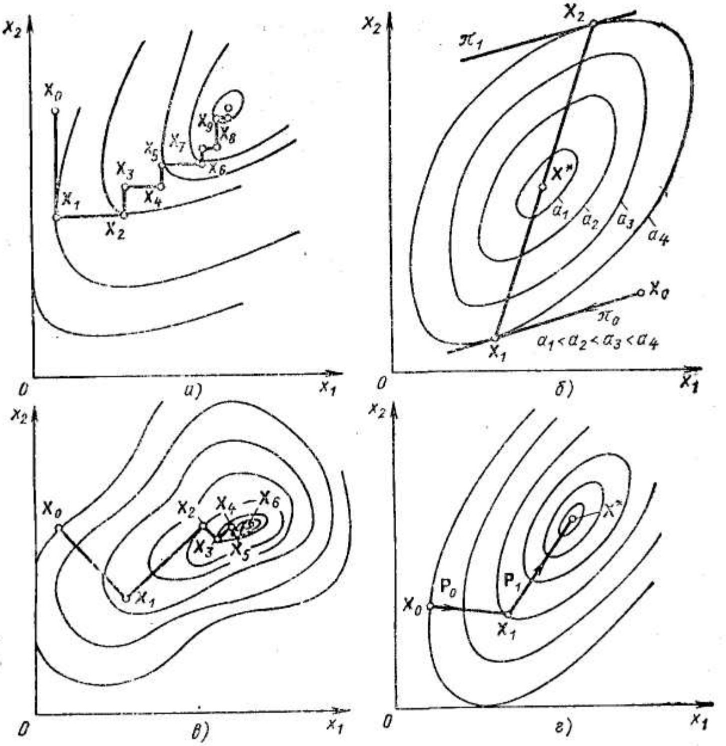


Рис. 4. Поиск экстремума квадратичной функции методом покоординатного спуска (*а*), методом параллельных касательных *(б)*, методом наискорейшего спуска (*в*) и методом сопряженных градиентов (*г*)

Результат первой итерации $\mathbf{X}_1$ получается из $\mathbf{X}_0$ при выполнении однопараметрической оптимизации по параметру x_1, т. е.

$$\mathbf{X}_1 = (x_1^{(1)}, x_2^{(0)} \dots, x_n^{(0)}),$$

причем

$$F\left(x_1^{(1)}, x_2^{(0)}, \dots, x_n^{(0)}\right) = \min_{x_1} F\left(x_1^{(1)}, x_2^{(0)}, \dots, x_n^{(0)}\right).$$

Результат второй итерации $\mathbf{X}_2$ получается из $\mathbf{X}_1$ путем оптимизации $F(\mathbf{X})$ по параметру x_2 и т. д. Далее процесс продолжается из точки $\mathbf{X}_n$ снова путем варьирования параметра x_1 и т. д. Величину шага α_k выбирают по способу оптимального шага.

Величина r за один цикл поиска равна размерности пространства, т. е. $r = n$.

Модификацией алгоритма покоординатного спуска является *метод ортогональных направлений (метод Розенброка),* который основан на вращении системы координат в соответствии с изменением скорости убывания критерия оптимальности. При этом направление одной оси соответствует наиболее вероятному направлению скорейшего убывания на данной итерации критерия оптимальности, а остальные находятся из условия ортогональности.

Если в задаче оптимального синтеза структур поверхность отклика ограничена концентрическими эллипсоидами, то точное местоположение оптимума не более чем за ($2n$—1) одномерных итераций позволяет получить *метод параллельных касательных.* Идея этого метода для n=2 иллюстрируется на рис. 4, *б*.

Метод заключается в поиске центра системы концентрических эллипсов. Первоначально определяют направление касательной π_0 из точки $\mathbf{X}_0$ к линии уровня (точка $\mathbf{X}_1$), затем строят произвольную прямую π_1, параллельную π_0, и на ней точку $\mathbf{X}_2$, соответствующую точке касания π_1 и линии уровня. Местоположение центра $\mathbf{X}^*$ определяется путем поиска вдоль линии $\mathbf{X}_1\mathbf{X}_2$.

Метод параллельных касательных пригоден для поиска оптимума на любой унимодальной функции, требует небольшого количества информации, но сопряжен с громоздкими вычислениями.

Градиентные методы. Методы поиска, в которых направление движения от итерации $\mathbf{X}_{k-1}$ к $\mathbf{X}_k$ определяется градиентом (антиградиентом), вычисленным в точке $\mathbf{X}_{k-1}$, носят название *градиентных методов*.

Наиболее применяемым из методов минимизации является *метод наискорейшего спуска.* В большой степени широкому распространению метода способствуют его сравнительная простота и возможность применения для минимизации весьма широкого класса функций. При определении направления поиска выбирают наибыстрейшее убывание целевой функции $F(\mathbf{X})$, т. е.

$$\mathbf{P}_k=-\mathrm{grad}F(\mathbf{X}_{k-1}).$$

Работа метода заключается в следующем. После определения градиента критерия оптимальности в точке $\mathbf{X}_k$ движутся вдоль направления антиградиента до точки, в которой достигается минимальное значение функции. Затем в этой точке снова определяют градиент и движутся по прямой согласно направлению нового антиградиента и т. д., пока не достигнут точки, имеющей наименьшее значение функции $F(\mathbf{X})$. На рис. 4, *в* приведен пример движения при поиске методом наискорейшего спуска оптимума для критерия оптимальности, зависящего от двух переменных. Направление grad $F(\mathbf{X}_{k-1})$ является касательным к поверхности уровня в точке $\mathbf{X}_k$, и, следовательно, grad $F(\mathbf{X}_k)$ в точке $\mathbf{X}_{k-1}$ ортогонален grad $F(\mathbf{X}_{k-1})$.

Простейшим среди градиентных методов является метод, при котором движение из точки $\mathbf{X}_{k-1}$ осуществляется *по антиградиенту с постоянным шагом* α_k=const.

Если окажется, что $F(\mathbf{X}_k)>F(\mathbf{X}_{k-1})$, то для определения экстремума используют один из возможных алгоритмов *поиска с возвратом*. Так, в частности, поиск может производиться путем возврата в точку $\mathbf{X}_{k-1}$ и дальнейшего перемещения из этой точки вдоль антиградиента с шагом $\alpha/2$. Процедура дробления шага производится до тех пор, пока его значение не станет меньше некоторого малого положительного числа ε.

Описанные варианты реализации градиентного метода отличаются друг от друга способом выбора длины шага. Скорость сходимости этих методов примерно одинакова, а трудоемкость каждой итерации вариантов процесса (42) различна только в способах определения параметра α_k. Как правило, вычисления градиента в меньшем числе точек требует метод наискорейшего спуска.

Градиентные методы эффективны для решения задач минимизации гладких и выпуклых функций. В практике синтеза структур для ускорения сходимости решения задач оптимизации градиентные методы следует использовать в комбинации с другими, более эффективными на начальной стадии решения задачи.

Метод сопряженных градиентов. В градиентных методах для поиска экстремума использовались свойства ортогональности векторов. В методе сопряженных градиентов оптимум целевой функции ищется на основе свойств орготональности приращений вектора градиентов. Для этой цели наряду с градиентом используют матрицу Гессе **Г** критерия оптимальности. С помощью матрицы **Г** удается выбрать направление поиска, наиболее полно учитывающее особенности критерия оптимальности. Напомним, что векторы **А** и **В** называют сопряженными относительно симметричной и положительно определенной матрицы **Г**, если скалярное произведение векторов **А** и **ГВ** равно нулю, т. е. $<\mathbf{A}, \mathbf{ГB}>=0$. Направление поиска $\mathbf{P}_{k+1}$ на k+1-м шаге определяется как

$$\mathbf{P}_{k+1} = -\operatorname{grad} F(\mathbf{X}_k) + \beta_k \mathbf{P}_k,$$

где

$$\beta_k = <\operatorname{grad} F(\mathbf{X}_k),\ \operatorname{grad} F(\mathbf{X}_k)>/<\operatorname{grad} F(\mathbf{X}_{k-1}),\ \operatorname{grad} F(\mathbf{X}_{k-1})>.$$

Если критерий оптимальности представлен квадратичной функцией, то минимум функции достигается ровно за *n* шагов (рис. 4, *г*). В случае критерия оптимальности произвольного вида метод позволяет для заданной погрешности получить приближенное решение быстрее, чем это позволяют сделать методы наискорейшего спуска и параллельных касательных.

Методы Ньютона и переменной метрики. Ускорение поиска экстремума связано с улучшением выбора сопряженных направлений. Довольно эффективным является поиск сопряженных направлений с одновременным накоплением информации о матрице Гессе критерия оптимальности. Используют соотношение

$$\mathbf{X}_{k+1} = \mathbf{X}_k - \alpha_k \mathbf{\Gamma}_k^{-1} \operatorname{grad} F(\mathbf{X}_k), \alpha_k > 0, \; k = 0, \; 1, \ldots,$$

которое является формулой *метода Ньютона с регулировкой шага.* Обычный метод Ньютона соответствует случаю α_k=l.

Метод Ньютона с регулировкой шага сходится к решению независимо от выбора начальной точки $\mathbf{X}_0$ и обладает либо линейной, либо квадратичной скоростью сходимости в зависимости от вида функции $F(\mathbf{X})$. В обычном методе Ньютона сходимость гарантируется лишь при наличии достаточно хорошего начального приближения.

При решении задач минимизации выпуклых функций метод Ньютона обеспечивает более высокую скорость сходимости последовательных приближений к решению по сравнению с градиентными методами, однако количество вычислений на итерации метода Ньютона высоко за счет необходимости вычисления и обращения матрицы вторых производных. Минимизация квадратичных функций происходит за один шаг.

На основе метода Ньютона разработан эффективный метод, получивший название *метода переменной метрики.* Идея метода заключается в использовании информации о градиенте критерия оптимальности для приближенного вычисления матрицы Гессе. Этот метод — итерационный. Поиск в нем ведется по формуле

$$\mathbf{X}_{k+1} = \mathbf{X}_k - h_k \mathbf{H}_k \operatorname{grad} F(\mathbf{X}_k),$$

где $\mathbf{H}_k$ — приближенно вычисляемая обратная матрица Гессе; h_k — величина шага, определяемая одномерной минимизацией целевой функции на луче — $\mathbf{H}_k$ grad $F(\mathbf{X}_k)$.

Главное преимущество метода переменной метрики перед методом Ньютона — отказ от вычислений матрицы Гессе на каждой итерации. Положительно определенная матрица

$$\mathbf{H}_k = \mathbf{H}_{k-1} + \frac{\Delta\mathbf{X}_{k-1}(\Delta\mathbf{X}_{k-1})^{\mathrm{T}}}{\langle \Delta\mathbf{X}_{k-1}, \mathbf{R}_{k-1} \rangle} + \frac{\mathbf{H}_{k-1}\mathbf{R}_{k-1}(\mathbf{R}_{k-1})^{\mathrm{T}}\mathbf{H}_{k-1}}{(\mathbf{R}_{k-1})^{\mathrm{T}}\mathbf{H}_{k-1}\mathbf{R}_{k-1}},$$

где $\Delta\mathbf{X}_{k-1}$ — приращение вектора переменных синтеза структур на предыдущем (k—1)-м шаге, имеет форму вектор-столбца; $(\Delta\mathbf{X}_{k-1})^{\mathrm{T}}$— транспонированный вектор $\Delta\mathbf{X}_{k-1}$, т. е. вектор-строка;

$$\mathbf{R}_{k-1} = \operatorname{grad} F(\mathbf{X}_k) - \operatorname{grad} F(\mathbf{X}_{k-1})$$

— вектор-строка.

В начале вычислений нужно задаться произвольной положительно определенной матрицей $\mathbf{H}_0$, в частности $\mathbf{H}_0$ может быть единичной матрицей. Шаг h_k выбирают по методу одномерной оптимизации.

Ввиду того что в методе переменной метрики достаточно полно учитывается локальная информация, его целесообразно применять в окрестности оптимального решения.

Методы одномерной оптимизации. Эти методы позволяют найти оптимум для функций одной переменной. ***Они не требуют <u>условия дифференцируемости</u> критерия оптимальности, а требуют только его <u>непрерывности</u>.*** Среди этих методов наиболее распространены ***методы чисел Фибоначчи, золотого сечения и полиномиальной аппроксимации***. Для их применения необходимо знать интервал [a, b] переменной x, на котором функция $F(x)$ имеет единственный минимум.

При применении ***метода чисел Фибоначчи*** должно быть зафиксировано число точек N, в которых производится вычисление критерия оптимальности.

Пусть на k-й итерации интервал изменения переменной x сократился до $[a_k, b_k]$, $a \le a_k \le b_k \le b$. Тогда для вычисления следующего интервала $[a_{k+1}, b_{k+1}]$ выбирают точки x_k и x'_k (рис. 5) по формулам

$$x_k = \frac{\Phi_{N-1-k}}{\Phi_{N+1-k}}\left(b_k - a_k\right) + a_k; \quad x'_k = \frac{\Phi_{N-k}}{\Phi_{N+1-k}}\left(b_k - a_k\right) + a_k,$$

где Φ_k — числа Фибоначчи, определяемые с помощью рекуррентных соотношений

$$\Phi_k = \Phi_{k-1} + \Phi_{k-2}; \quad \Phi_0 = \Phi_1 = 1.$$

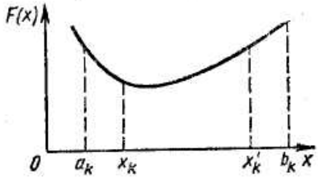


Рис. 5. Расположение интервалов поиска экстремума методом чисел Фибоначчи

Если $F(x_k) < F(x'_k)$, то в качестве следующего интервала выбирают $[a_{k+1}, b_{k+1}]=[a_k, x'_k)$, если $F(x_k)>F(x'_k)$, то выбирают $[a_{k+1}, b_{k+1}] = [x_k, b_k]$, если $F(x_k) =F(x'_k)$, то может быть выбран любой из интервалов.

Последние точки задаются формулами

$$x'_{N-1} = (0{,}5 + \varepsilon)(b_{N-1} - a_{N-1}) + a_{N-1};$$

$$x_{N-1} = 0{,}5(b_{N-1} - a_{N-1}) + a_{N-1},$$

где ε — произвольно малое число, вводимое на последней итерации. Следует отметить, что длина последнего интервала неопределенности определяется как

$$b_N - a_N = (b_0 - a_0)/(2\Phi_N). \tag{44}$$

Выражение (44) позволяет определить количество вычислений критерия оптимальности исходя из требуемой точности поиска.

В ***методе золотого сечения*** сохраняется постоянным отношение длин двух последовательных интервалов неопределенности:

$$\tau = L_k / L_{k+1} = 1{,}61803...$$

По результатам двух экспериментов устанавливают, какую область неопределенности оставить для дальнейших исследований. Процесс поиска оптимума можно продолжать сколь угодно долго. После N испытаний длина интервала неопределенности составляет $L_N = 1/\tau^{N-1}$.

Отметим, что метод золотого сечения требует сравнительно небольшого объема памяти ЭВМ и прост в реализации.

Метод полиномиальной аппроксимации заключается в определении полинома, аппроксимирующего функцию $F(\mathbf{X})$ (чаще всего — квадратичного полинома), и поиске его минимума.

В конкретных задачах оптимального синтеза структур довольно часто зависимость критерия оптимальности F от параметров синтеза структур $\mathbf{X}$ получается слишком сложной. В этих случаях вместо вышеизложенных регулярных методов оптимизации рекомендуется использовать ***методы случайного поиска***. В этих методах направление поиска $\mathbf{P}_k$, выбирают случайно, например, равновероятно в пределах гиперсферы с центром в точке $\mathbf{X}_{k-1}$. Существует огромное количество алгоритмов случайного поиска. Следует отметить, что регулярные алгоритмы поиска являются частным (а точнее, вырожденным) случаем стохастических алгоритмов.

Методы условной оптимизации. ***Задачи условной оптимизации, заключающиеся в минимизации некоторого критерия оптимальности с ограничениями на область существования переменных синтеза структур, относятся к классу задач математического программирования.***

Одним из наиболее простых и широко известных методов решения задачи математического программирования является ***метод штрафных функций***. Основная идея метода состоит в приближенном сведении задачи минимизации функции $F(\mathbf{X})$ при ограничениях $Q_i(\mathbf{X})\leq 0$, i==l, 2,…, n, к задаче минимизации функции

$$\Phi(\mathbf{X},t) = F(\mathbf{X}) + tR(\mathbf{X}) \qquad (45)$$

без ограничений. При этом вспомогательную функцию $\Phi(\mathbf{X}, t)$ подбирают так, чтобы она совпадала с функцией $F(\mathbf{X})$ внутри допустимой области S и быстро возрастала вне ее.

В выражении (45) $R(\mathbf{X})$ — дифференцируемая функция штрафа, удовлетворяет следующим условиям: $R(\mathbf{X})$ =0, если $\mathbf{X}\in S$, и $R(\mathbf{X})>0$, если хотя бы для одного k будет $Q_k,(\mathbf{X})>0$, (k =1,2,…, n), t — некоторое положительное число — коэффициент штрафа.

Примерами функций $R(\mathbf{X})$ могут служить выражения

$$R_1(\mathbf{X}) = \sum_{i=1}^{n} \max R_i \{(\mathbf{X}), 0\}^{\alpha}, \quad \alpha \geqslant 1;$$

или

$$R_2(\mathbf{X}) = \sum_{i=1}^{n} \left[\frac{R_i(\mathbf{X}) + |R_i(\mathbf{X})|}{2} \right]^2.$$

Рис. 6 иллюстрируется метод штрафных функций в одномерном случае.

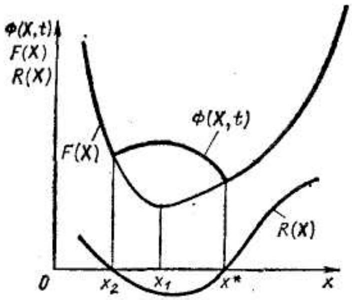


Рис. 6. Метод штрафных функций

Допустимая область S определяется ограничением $R(\mathbf{X})\geq 0$, в этой области $F(\mathbf{X})$ и $\Phi(\mathbf{X}, t)$ совпадают. В области, где $R(\mathbf{X})<0$, функция $\Phi(\mathbf{X}, t)$ резко возрастает. На рисунке $\mathbf{X}_1$ и $\mathbf{X}^*$ — точки безусловного и условного минимумов.

В задачах условной оптимизации, в которых ограничения заданы только в виде неравенств, возможно построение обобщенного критерия оптимальности с помощью ***барьерных функций***. Значения, принимаемые барьерной функцией, неограниченно возрастают при приближении к границе допустимой области.

Примером барьерной функции является

$$R(\mathbf{X}) = -\sum_{i=1}^{n} \ln\left[-R_i\ (\mathbf{X})\right].$$

Эта функция существует только внутри допустимой области S. Вне области S и на ее границе функция $R(\mathbf{X})$ не определена, а при приближении к границе области S она неограниченно возрастает.

Исходя из организации поиска условного оптимума иногда метод штрафных функций называют ***методом внешней точки***, а метод барьерных функции — ***методом внутренней точки***.

3.5. Задачи синтеза по оптимизации допусков на функционирование структур

Качество формируемых рекомендаций по решению задач синтеза структур в значительной мере определяется характером постановки задачи параметрического синтеза структур, реализуемой при синтеза структур, т. е. тем, насколько сформулированные целевая функция и ограничения отражают объективно существующие требования к свойствам структур. При формализации ТЗ такие требования выражаются в виде ***условий функционирования*** структур. ***Условие функционирования структур — это требуемое соотношение между выходным параметром*** y_j***, значения которого зависят от принимаемых сформированных рекомендаций, и предельно допустимым значением — нормой*** $y_j^{(0)}$. Величину $y_j^{(0)}$ будем называть также ***требованием на параметр*** y_j. Условия функционирования структуры могут иметь одну из следующих форм:

$$y_j < y_j^{(0)}, \qquad (46а)$$

$$y_j > y_j^{(0)}, \qquad (46б)$$

$$y_{j\min}^{(0)} < y_j < y_{j\max}^{(0)}. \qquad (46в)$$

Формы (46б) и (46в) могут быть сведены к форме (46а), поэтому в дальнейшем будем считать, что все условия функционирования структуры в ТЗ имеют вид (46а).

Область в пространстве **ХП** управляемых параметров, в которой все условия функционирования структуры, а заданные прямые ограничения на управляемые параметры x_i вида

$$x_{i\min} \leqslant x_i \leqslant x_{i\max} \qquad (9.46г)$$

удовлетворяются, будем называть ***областью функционирования структуры*** $\mathbf{Z}_0$.

Задача формирования рекомендаций по оптимизации ***допусков функционирования структуры*** сводится к определению размеров допусковой области $\mathbf{Z}_д$ и ее расположения в пространстве **ХП**. ***Цель оптимизации допусков* — *максимизация размеров области $Z_д$ при выполнении ограничений на степень несовпадения областей Z_0 и $Z_д$.***

Решение задачи оптимизации допусков выполняется в два этапа: на первом ищут какую-либо точку $\mathbf{X}_э \in \mathbf{Z}_0$, называемую ***опорной***; на втором определяют ***оптимальную допусковую область***.

На первом этапе для определения опорной точки целесообразно использовать постановку задачи оптимизации параметров, известную под названием ***максиминной постановки***. Последняя приводит к получению опорной точки внутри области $\mathbf{Z}_0$ на достаточном удалении от границ, что удобно для реализации алгоритмов второго этапа.

При максиминной постановке вводится количественная оценка s_j степени выполнения j-го условия функционирования структуры. Каждая из оценок s_j может носить детерминированный или статистический характер. При детерминированном подходе используют формулу

$$s_j = (y_j^{(0)} - y_j)/y_j^{(0)},$$

при статистическом — формулу

$$s_j = (y_j^{(0)} - M_j)/\delta_j,$$

где M_j — оценка математического ожидания; δ_j — оценка рассеяния параметра y_j.

В частности, величина δ_j может рассматриваться как некоторый весовой коэффициент, указание его физического смысла упрощает правильное задание его численного значения без трудоемких статистических расчетов.

Так как вероятность надежного функционирования структуры определяется главным образом наименьшей из вероятностей выполнения отдельных условий функционирования структуры, то в первую очередь нужно ***увеличивать наименьший из запасов*** s_j. Поэтому в качестве целевой функции $F(\mathbf{X})$ следует выбрать наименьший из запасов, и задача оптимизации параметров структуры формулируется как максиминная задача нелинейного программирования:

$$\max_{\mathbf{X} \in \mathbf{Z}_0} \min_j s_j(\mathbf{X}).$$

Ограничениями задачи при этом будут прямые ограничения (46г).

Максиминный критерий запаса функционирования структуры применим при наличии у проблемы синткза параметров с условиями функционирования структуры любого вида. Этот критерий в зависимости от конкретной ситуации может рассматриваться либо как детерминированный, либо как статистический.

На втором этапе формулируется ***задача вписывания гиперпараллелепипеда допусков*** $\mathbf{Z}_д$ в область функционирования $\mathbf{Z}_0$. В этой задаче к исходным данным относятся:

1. Область функционирования структуры $\mathbf{Z}_0$, задаваемая прямыми ограничениями (46г) и условиями функционирования структуры, приведенными к виду $y_j\,(\mathbf{X})\leq y_j^{(0)}$.
2. Точка $\mathbf{X}_э \in \mathbf{Z}_0$.
3. Соотношения α_i, между допусками g_i, равноценными с позиций затрат на их получение при производстве.

При решении задачи синтеза структуры сначала нормируют управляемые параметры в соответствии с заданными соотношениями между допусками

$$u_i = \alpha_i\,(x_i - x_{эi})/x_{эi} + 1,$$

где u_i — нормированный параметр x_i; $\alpha_i = g_1/g_i$.

При таком нормировании оптимальной формой допусковой области является ***гиперкуб с ребрами, параллельными координатным осям***. Областям $\mathbf{Z}_0$ и $\mathbf{Z}_д$ в нормированном пространстве управляемых параметров соответствуют области $\mathbf{U}_0$ и $\mathbf{U}_1$.

Задача вписывания формулируется как ***задача определения центра гиперкуба*** $\mathbf{U}^* \in \mathbf{U}_0$, имеющего максимальную длину ребра при условии, что оценка рассогласования положений областей работоспособности структуры $\mathbf{U}_0$ и допусковой области $\mathbf{U}_1$ не превышает заданной величины.

Рассмотрим один из алгоритмов вписывания, основанный на линеаризации зависимостей $y_j(\mathbf{U})$ и ориентированный на случай, когда $\mathbf{U}_1 \subset \mathbf{U}_0$. Этот случай часто называют оптимизацией параметров и допусков в условиях 100%-ного создания структур с заданными параметрами. Принимается также допущение о постоянстве знаков коэффициентов влияния $a_{ji} = \partial y_j / \partial x_i$ т.

$$\operatorname{sgn}(a_{ji}) = \text{const}, \tag{47}$$

в пределах $\mathbf{U}_0$, i=l, 2, ..., n; n — количество управляемых параметров.

Для j-го участка границы области $\mathbf{U}_0$, выражаемого уравнением

$$y_j^{(0)} - y_j\,(\mathbf{U}) = 0, \tag{48}$$

известно положение нормали, проходящей через центр $\mathbf{U}^*$ гиперкуба. Система уравнений этой нормали в параметрическом виде

$$u_i - u_i^* = a_{ji} h_j, \tag{49}$$

где h_j — параметр; a_{ji} определяется в точке пересечения нормали с поверхностью (48).

Запишем систему уравнений для диагонали искомого гиперкуба, имеющей ту же ориентацию, что и j -я нормаль (49):

$$u_i - u_i^* = \operatorname{sgn}(a_{ji})\,\gamma_j, \tag{50}$$

где γ_j — параметр.

В соответствии с (47) вычисление a_{ji} может быть выполнено в точке $\mathbf{U}_э$. Проведем линеаризацию границ (48) области $\mathbf{U}_0$, используя разложение $y_j(\mathbf{U})$ в ряд Тейлора в окрестности опорной точки $\mathbf{U}_э$:

$$y_j^{(0)} - y_j(\mathbf{U}_э) - \sum_{i=1}^{n} a_{ji}(u_i - u_i^*) - \sum_{i=1}^{n} a_{ji}(u_i^* - u_{эi}) = 0. \tag{51}$$

Обозначим

$$c_j(\mathbf{U}^*) = y_j^{(0)} - y_j\ (\mathbf{U}_э) - \sum_{i=1}^{n} a_{ji}\ (u_i^* - u_{эi}).$$

Тогда (51) перепишем в виде

$$c_j(\mathbf{U}^*) - \sum_{i=1}^{n} a_{ji}(u_i - u_i^*) = 0. \tag{52}$$

Точка пересечения диагонали (50) с границей (52) определяется из совместного решения уравнений (50) и (52):

$$c_j(\mathbf{U}^*) - \gamma_j \sum_{i=1}^{n} \operatorname{sgn}\ (a_{ji})\ a_{ji} = 0,$$

$$\gamma_j = c_j(\mathbf{U}^*) / \sum_{i=1}^{n} |a_{ji}|. \tag{53}$$

Здесь γ_j имеет смысл половины длины ребра гиперкуба, имеющего центр в точке $\mathbf{U}^*$ и вершину на гиперповерхности (52). Всего имеем m условий функционирования структуры и, следовательно, не более чем m вершин на границах $\mathbf{U}_0$.

Чтобы выполнить условие $\mathbf{U}_1 \subset \mathbf{U}_0$, нужно контролировать принадлежность области $\mathbf{U}_0$ вершины, которой соответствует минимальная величина γ_q среди величин γ_j. Чтобы область $\mathbf{U}_1$ имела максимально возможные размеры, нужно, максимизировать γ_q, варьируя положение центра $\mathbf{U}^*$.

Таким образом, задача вписывания допусковой области в область функционирования структуры формулируется как задача линейного программирования

$$\max_{\mathbf{U}^* \in \mathbf{U}_0} \gamma_q(\mathbf{U}^*)$$

при ограничениях

$$\gamma_p(\mathbf{U}^*) - \gamma_q(\mathbf{U}^*) \geqslant 0; \quad u_i^* \geqslant 0; \quad \|\mathbf{U}_э - \mathbf{U}^*\| < \gamma_q.$$

Последнее ограничение ставится в том случае, если точка $\mathbf{U}_э$ должна принадлежать допусковой области.

Отметим, что основные затраты машинного времени на реализацию алгоритма связаны с анализом чувствительности. Анализ чувствительности методом приращений требует n+1 раз обращаться к математической модели структуры. Первое обращение производится при значении вектора управляемых параметров $\mathbf{U}_э$ и позволяет вычислить $y_j(\mathbf{U}_э)$, фигурирующие в (51). Каждое последующее обращение позволяет вычислить очередную строку матрицы чувствительности и в итоге дает значения a_{ji}. Теперь полностью определена линеаризованная модель структуры (53). Манипулирование ею при решении задач линейного программирования не требует заметных затрат машинного времени.

При большом числе n управляемых параметров (несколько сотен и более) применение алгоритма, включающего анализ чувствительности методом приращений, становится нерациональным. С ростом n более предпочтительными оказываются алгоритмы, основанные на ***статистических испытаниях***.

Одним из таких алгоритмов является ***алгоритм центрирования по методу статистического градиента***. На каждом шаге алгоритма выполняются N статистических испытаний с выбором случайных точек в пределах некоторой области $\mathbf{U}_{СТ}$. По результатам испытаний выделяются те точки $\mathbf{U}_p \in \mathbf{U}_{СТ}$, которые оказались в области функционирования структуры $\mathbf{U}_0$. Для следующего шага в качестве координат u_i* центра $\mathbf{U}$* допусковой области $\mathbf{U}_{СТ}$ принимаются средние арифметические значения координат u_{ip} выделенных точек $\mathbf{U}_p$.

Если в процессе испытаний область $\mathbf{U}_{СТ}$ выбирается в соответствии с некоторыми условиями (например, $\mathbf{U}_0 \in \mathbf{U}_{СТ}$ и $\mathbf{U}_{СТ}$ имеет форму гиперсферы) и если выполнено достаточное количество шагов, то точка $\mathbf{U}$* может быть принята в качестве центра области $\mathbf{U}_0$, а окончательная допусковая область $\mathbf{U}_д$ устанавливается в соответствии с характером распределения точек $\mathbf{U}_p$ на последнем шаге центрирования. Задача оптимизации допусков обычно решается на том иерархическом уровне синтеза структур, на котором в качестве управляемых параметров фигурируют ***параметры базовых элементов***.

Сформированные рекомендации по рассчету допусков используются для выбора унифицированных элементов по справочникам и каталогам либо служат непосредственными исходными данными для последующего технологического проектирования.

На промежуточных иерархических уровнях нисходящего функционального или конструкторского проектирования структур также возникают задачи, подобные задаче оптимизации допусков. Предположим, что на k-м иерархическом уровне управляемыми параметрами структуры являются параметры y_j. На следующем, (k+1)-м иерархическом уровне эти же параметры рассматриваются уже как выходные параметры подпроблем, а управляемыми параметрами здесь оказываются другие параметры x_i. Для выполнения синтеза структуры на (k +1)-м иерархическом уровне на выходные параметры y_j нужно задать условия функционирования структуры. Очевидно, что эти условия должны быть результатом сформированных рекомендаций на k -м уровне, т. е. должны быть определены не только некоторая оптимальная точка **Y*** в пространстве параметров y_j, но и требования синтеза структуры $y_j^{(0)}$ на эти параметры.

Задача оптимального расчета требований синтеза $y_j^{(0)}$ по своей постановке и методам решения близка к рассмотренной выше задаче оптимизации допусков. Исходными данными при ее решении являются условия функционирования структуры, задаваемые на параметры структуры, а результатом должны быть условия функционирования структуры для подпроблем. При оптимизации требований синтеза структуры на последующем (k +1)-м уровне исходными данными будут условия функционирования структуры на параметры подпроблем, а результатом — условия функционирования структуры для элементов подпроблем и т. д.

Таким образом, блочно-иерархический подход приводит к формулировке основных оптимизационных задач нисходящего синтеза структур как задач оптимального преобразования технического задания на структуру k-го иерархического уровня в техническое задание на структуру (k+1)-го иерархического уровня. ***Эти задачи решаются с помощью статистических или детерминированных алгоритмов вписывания гиперфигур в заданную область n-мерного пространства параметров.***

3.6. Методы формирования рекомендаций по оптимизации различных процессов синтеза структур

Комплексные САК охватывают все этапы консультирования проблем. Автоматизация формирования рекомендаций по решению консультационных задач различных процессов включает в себя разработку принципиальных схем процессов, маршрутной технологии, операционной технологии и получение управляющей информации на машинных носителях для программно-управляемого процессоного оборудования.

Любой процесс независимо от его физической природы всегда можно представить в виде некоторой КП, а следовательно, для его организации следует применять системный подход, сущность которого заключается в комплексном, едином рассмотрении всех частей технологических систем и в гармоническом их сочетании.

Постановку задачи оптимизации рекомендаций по решению задач проектирования заданного процесса - как консультируемой проблемы, можно представить следующим образом (рис. 9.7).

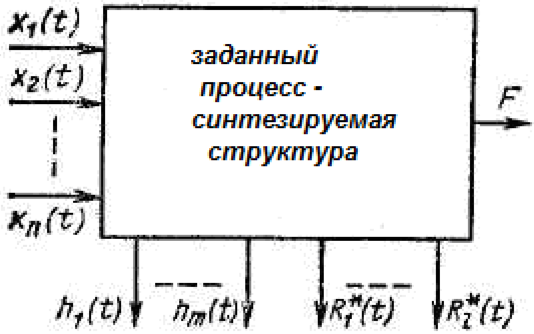


Рис. 7. Постановка задачи оптимизации рекомендаций по решению задач проектирования заданного процесса - как проблемы синтеза структур

Заданный процесс рассматривается как система, на вход которой поступает вектор **X**$(t)=(x_1(t),...,x_n(t))$ входных переменных, а скалярный выход F характеризует качество заданного процесса. Требуется сформировать рекомендации по такому вектору управляющих воздействий

$$\mathbf{R}^*(t) = (R_1^*(t), \ldots, R_l^{(*)}(t)),$$

который минимизировал бы значение показателя *F*, т. е.

$$F(\mathbf{X}(t), R^*(t)) = \min_{\mathbf{R}(t)} \{F_k(\mathbf{X}(t), \mathbf{R}(t))\}.$$

Выходы $h_1(t), \ldots, h_m(t)$ характеризуют состояние заданного процесса и индицируют нежелательные режимы работы процессных средств или выход контролируемых параметров заданного процесса за установленные пределы:

$$h_j(t) \leqslant b_j, \quad j = \overline{1,m},$$

где b_j — требуемые или допустимые значения соответствующих параметров.

Таким образом, оптимизация рекомендаций по решению задач проектирования заданного процесса рассматривается как задача определения оптимального вектора управления $\mathbf{R}^*_{\text{ит}}(t)$, минимизирующего целевую функцию $F(\mathbf{X}(t), \mathbf{R}(t)$ при условии выполнения заданных ограничений.

Подобные задачи оптимизации решают в два этапа. На первом этапе определяют идеальный вектор управления $\mathbf{R}^*_{\text{ит}}(t)$, обеспечивающий оптимизацию заданного процесса. Практически реализовать это не представляется возможным, и вектор $\mathbf{R}^*_{\text{ит}}(t)$ является эталоном, к которому надо стремиться. Зная $\mathbf{R}^*_{\text{ит}}(t)$, на втором этапе выбирают реализуемый квазиоптимальный вектор управления, с помощью которого стараются получить рекомендацию, наименее отличающуюся от идеальной и в то же время реализуемую наиболее просто.

Следует отметить, что в большинстве практических случаев оптимизированные рекомендации по решению задач проектирования заданных процессов синтеза структуы дополнительно подвергаются наладке и корректировке, поскольку при построении математических моделей процессов синтеза структуы невозможно учесть все влияющие на процесс факторы.

На любой заданный процесс оказывает влияние множество случайных факторов. Так, например, на технологический процесс изготовления механических изделий влияет неточность тенических средств, неточности режущего инструмента и приспособлений, внутренние напряжения обрабатываемой детали, разброс параметров у электрорадиоэлементов при монтаже электронных схем и т. д. Поэтому параметры изготовляемых изделий являются случайными величинами, причем влияние действующих случайных факторов на изменение параметров изделий можно, как правило, определить исходя из статистического анализа. В подобных случаях очень важно учитывать характер взаимосвязи между случайными величинами. Для количественного выражения этой взаимосвязи служат регрессия и корреляция. Остановимся более подробно на этих понятиях.

Пусть x и y — случайные величины, характеризующие параметры некоторой структуры, причем упорядоченная пара (x, y) характеризует параметры одного варианта рекомендаций синтеза структуы и может быть изображена точкой на плоскости. Полная совокупность вариантов рекомендаций изображается множеством точек, показанных на рис. 8. Математические ожидания случайных величин x и y равны соответственно $M_{(x)}$ и $M_{(y)}$, и среднеквадратичные отклонения σ_x и σ_y характеризуют рассеивание величин x и y относительно их математических ожиданий.

Рассмотрим зависимость $\overline{y}(x)$, являющуюся условным математическим ожиданием $M_{(x|y)}$. Используя выражение для условного математического ожидания и обозначая через $p(x, y)$ совместную вероятность данных значений x и y, находим

$$\bar{y}(x) = M_{(y\,|\,x)} \sum_{y} yp(x, y) / \sum_{y} p(x, y). \tag{54}$$

Определяя $\overline{y}(x)$ при различных x, можно построить линию, графически выражающую эту зависимость и называемую линией регрессии y по x (рис..8).

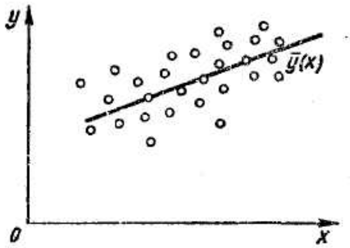


Рис. 8. Линия регрессии

Аналогично может быть получена зависимость $\overline{x}(y)$, называемая регрессией x по y.

На практике наиболее часто встречается случай линейной регрессии, уравнение которой записывается в виде

$$\bar{y}(x) = a + b(x - M_{(x)}). \tag{55}$$

Коэффициенты a и b выбирают такими, чтобы получить наибольшую концентрацию точек *(x, y)* вблизи прямой $\overline{y}\,(x)$, что выражается условием

$$\varphi(a,b) = M\{[y - \bar{y}(x)]^2\} = \min. \qquad (56)$$

Выражение (56) с учетом (55) дает следующую систему уравнений для определения коэффициентов a и b:

$$M_{(y)} - a = 0; \quad M\left[y\left(x - M_{(x)}\right)\right] - b\sigma_x^2 = 0. \qquad (57)$$

Величину $\mu_{xy} = M[y(x—M_{(x)})]$ называют ковариацией между x и y. Она служит мерой взаимной связи между случайными величинами a и b.

Из (57) находят значения

$$a = M_{(y)} \text{ и } b = \mu_{xy}/\sigma_x^2,$$

определяющие линию регрессии.

Ковариация μ_{xy} зависит от дисперсий самих случайных величин, поэтому для оценки взаимосвязи между случайными величинами более удобен коэффициент корреляции $r_{xy}=\mu_{xy}/(\sigma_x\sigma_y)$, который может меняться от нуля для независимых случайных величин до единицы, если случайные величины связаны линейной функциональной зависимостью. При технологическом проектировании в качестве критериев оптимальности рекомендаций могут рассматриваться такие показатели эффективности, как себестоимость производства изделий, производительность технологических процессов, основное технологическое время и т. д.

Методы оптимизации формирования рекомендаций по решению задач проектирования заданных процессов рассмотрим на примерах управления технологическим процессом производства магнитоуправляемых контактов и управления технологической установкой.

Пример 6. ***Формирование рекомендаций по управлению технологическим процессом на основе текущего регрессионного анализа.*** Рассмотрим технологический процесс производства магнитоуправляемых контактов (МК). Основная задача производства МК — получение изделий с заданными величинами напряженности магнитного поля Θ и зазора S между контактами.

Одной из задач автоматизации формирования рекомендаций по проектированию технологического процесса производства МК является определение функциональной связи между величинами Θ и S с последующей реализацией математической модели процесса управления заварки лепестков МК на управляющей ЭВМ.

Схема управления заваркой лепестков МК приведена на рис. 9.

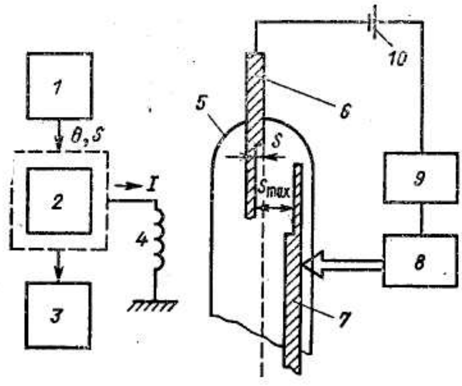


Рис. 9. Схема управления заваркой лепестков МК:
1— пульты статистического контроля Θ и *S; 2* — управляющая ЭВМ; *3*—блок выдачи рекомендаций технологу; *4* — катушка МК; *5*— стеклянный баллон; *6, 7* — верхний и нижний пружинные контакты; *8* — исполнительный механизм; *9* — регистрирующее устройство; *10* — источник питания

Технологический процесс установки зазора по заданной напряженности магнитного поля состоит в следующем: заваривается верхний пружинный контакт *6;* нижний пружинный контакт 7 отводится исполнительным механизмом *8* на максимальное расстояние; в катушку *4* подается ток, значение которого обеспечивает заданную величину напряженности магнитного поля для срабатывания МК; начинается движение контакта *7* к контакту *6.*

Движение контактов происходит до момента их замыкания под действием поля катушки *4.* Значение поля пропорционально силе тока *I*, протекающего в катушке. При замыкании регистрирующее устройство *9* срабатывает и останавливает исполнительный механизм 8; нижний пружинный контакт 7 заваривается, зазор становится равным величине *S*. Исследования показывают, что величины Θ и *S* после заварки не связаны жесткой функциональной зависимостью. Характер геометрического места точек Θ(S) зависит от многих факторов — качества исходного материала, режимов операции заварки и т. д.

С пультов статистического контроля *1* данные экспериментов поступают в управляющую ЭВМ *2,* в которой непрерывно строятся регрессионные модели и вырабатываются рекомендации по реализации управляющих воздействий на технологический процесс (корректируется сила тока в катушке, выдаются различные рекомендации).

Математически процесс управления заваркой МК можно свести к следующей модели. Пусть Θ и S — требуемые значения напряженности магнитного поля и зазора между контактами. Для решения задачи оптимизации сформулируем целевую функцию вида

$$F(\Theta,\ S)=\left(\frac{S_т-S}{S_т}\right)^2+\left(\frac{\Theta_т-\Theta}{\Theta_т}\right)^2,$$

которую следует минимизировать с учетом ограничения $\Theta = kS + b$, определяемого уравнением регрессии.

Сформулированная задача является ***задачей квадратичного программирования, которую можно решить с использованием неопределенных множителей Лагранжа.***

Функция Лагранжа имеет вид

$$F(\Theta,\ S,\ \lambda)=\left(\frac{S_т-S}{S_т}\right)^2+\left(\frac{\Theta_т-\Theta}{\Theta_т}\right)^2+\lambda(\Theta-kS-b).$$

Беря частные производные, получаем

$$\frac{\partial F}{\partial S}=-2\left(\frac{S_т-S}{S_т}\right)-k\lambda=0;\ \frac{\partial F}{\partial\Theta}=$$
$$=-2\left(\frac{\Theta_т-\Theta}{\Theta_т}\right)+\lambda=0;\ \frac{\partial F}{\partial\lambda}=\Theta-kS-b=0. \qquad (58)$$

Решением системы будут значения S^* и Θ^*, при этом Θ^* косвенно определяет силу тока I в катушке.

Пусть $\Theta_т$ = 40±2 А/м; $S_Т$ = 10±2 мкм; $\Theta = 3S + 6$; число витков в катушке ω = 50.

Решая систему уравнений (58), получаем Θ^*=40,44 А/м,
S^*=10,48 мкм.

Оптимальные значения Θ и S находятся в поле допуска. Необходимый ток в катушке I =0,8 А. После установки I =0,8 А строят новую регрессионную зависимость и рассчитывают новое значение I. В результате многократной коррекции силы тока I находят такое значение, при котором процент выхода МК, отвечающих требованиям консультационных условий, будет максимальным.

Пример 7. ***Формирование рекомендаций по оптимизации управления технологической установкой***.

При конструкторском проектировании и при технологической подготовке производства узлов ЭВА для формирования графической информации используют приборы последовательного действия (ППД), в частности координатографы, сверлильные станки с числовым программным управлением и др.

Управление ППД производится подачей на их входы совокупности команд *A,* которые формируют изображение *M.* Требуется сформировать рекомендации для получения такой совокупности команд, которая обеспечивает функционирование ППД (получение изображения *M)* с минимальным числом переходов от фрагмента к фрагменту.

Допустим, выработана последовательность команд $A_s = (\alpha_1, \ldots \ldots, \alpha_n)$ с индексом *s*. Тогда время работы ППД при формировании полного изображения *M* определяется соотношением

$$T^{(s)} = T_{\text{п}}^{(s)} + T_{\text{м}} = \sum_{i=1}^{n-1} t_{\alpha_i, \alpha_{i+1}} + \sum_{i=1}^{n} \Delta t_i,$$

где $T_{\text{п}}^{(s)}$ — общее время переходов между фрагментами изображения; $T_{\text{м}}$ — время выполнения команд ($T_{\text{м}}$ = const); $t_{\alpha_i}, t_{\alpha_{i+1}}$ — время перехода от команды α_i к команде α_{i+1}; Δt_i — время формирования *i*-го фрагмента изображения *M.*

Стоимость формирования полного изображения *M* последовательностью команд *A,* с учетом затрат на оптимизацию последовательности будет

$$C^{(s)} = C_{\text{опт}}^{(s)} + C_{\text{ППД}}^{(s)} = C_{\text{ЭВМ}} T_{\text{ЭВМ}}^{(s)} + C_{\text{ППД}} T^{(s)} q,$$

где $C_{\text{эвм}}$, $C_{\text{ппд}}$ —соответственно удельная стоимость работы ЭВМ, производящей оптимизацию, и ППД; $T_{\text{эвм}}^{(s)}$ и $T_{\text{ппд}}^{(s)}$ — соответственно времени работы ЭВМ и ППД; *q* — необходимое число копий изображения *M.*

Для оптимизации последовательностей *A* обычно выбирают такие методы, для которых

$$C_{\text{опт}}^{(s)} \ll C_{\text{ППД}}^{(s)}.$$

Таким образом, оптимизация последовательных команд *A* заключается в минимизации переходов T_n.

Для получения решения сформируем матрицу $\|t_{ij}\|_{n\times n}$, в которой элемент t_{ij} равен времени перехода от выполнения команды α_i к команде α_{ij}.

Введем псевдобулевы переменные

$$x_{ij} = \begin{cases} 1, & \text{если } i = \alpha_k, \ j = \alpha_{k+1}; \\ 0, & \text{если } i = \alpha_k, \ j \neq \alpha_{k+1} \end{cases}$$

и предположим, что последовательность команд образует замкнутый контур (начинается и заканчивается командой α_i). Тогда наша задача совпадает с классической задачей о коммивояжере

$$F(x)=\sum_{1\leqslant i+j\leqslant n}\sum t_{ij}\,x_{ij}\to\min$$

при ограничениях

$$\sum_{\substack{i=1\\ i\neq j}}^{n}x_{ij}=1,\ j=\overline{1,\ n};\ \sum_{\substack{j=1\\ j\neq i}}^{n}x_{ij}=1,\ i=\overline{1,\ n}, \qquad (59)$$

$$u_i-u_j+nx_{ij}\leqslant n-1,\ i,\ j=\overline{1,\ n}, \qquad (60)$$

где u_i, u_j — произвольные действительные числа.

Условия (59) гарантируют непрерывность маршрута, а условие (60) — получение последовательности команд, не распадающейся на ряд замкнутых циклов.

Задача может быть решена методом ветвей и границ.

3.7. Формирование рекомендаций при решении задач функционольного и структурного синтеза

При синтезе структуры разработчик, как правило, определяет совокупность функциональных компонент, из которых состоит система, вводит отношения между определенными компонентами и принимает это как начальное представление о системе.

Далее перед разработчиком возникает задача полноты определенного множества функциональных структур, их неизбыточности и согласованности по параметрам. В терминах общей теории структур это сводится к ***решению проблемы синтеза структуры системы***, которая представляется следующим образом: ***задана система данных, включающая в себя множество компонент структуры и отношения между ними; необходимо идентифицировать наилучшее представление этой системы с помощью ее структуры, элементы которой связаны с подмножествами переменных, описывающих компоненты.***

Процедура решения состоит из генерации наиболее приемлемых с точки зрения начального поведения проблемы структурных предположений рассматриваемой проблемы, анализа каждого из них на основе задайной системы данных и начальной структуры и сравнения результатов анализа с исходными данными с использованием различных критериев оценки полученного структурного предположения.

Например, при проектировании системы автоматизированного конструирования (САК) можно выделить четыре рекомендуемые компоненты, обеспечивающие ***управление, диалог, поиск, вычисления.*** Для каждой конкретной САК, которую будем рассматривать как проблему синтеза, эти компоненты могут быть различными способами реализованы.

Всего для этих компонент можно получить 2^{2^n} структур САК. Естественно, не все структуры этого множества приемлемы для более детального анализа (уточнение параметров выделенных компонент системы и определение требований к связям между этими компонентами).

Задачей синтеза рекомендаций в данном случае является формирование рекомендаций по выбору структуры из множества возможных структур, которая обеспечивала бы устойчивое и эффективное функционирование системы. Для проведения синтеза структур системы необходимо сформировать некоторую начальную структуру, т. е. выделить компоненты системы и определить отношения между ними. Далее, используя метод, изложенный ниже, можно синтезировать рекомендации для получения новых структур системы. Однако проводить синтез структур, не зная, как оценивать синтезируемые структуры, бессмысленно. Следовательно, необходимо иметь рекомендованный метод, который позволял бы оценивать структуры и выбирать наиболее приемлемые для последующего анализа и синтеза. Чтобы этот метод использовался при решении задачи синтеза структур систем, он должен обладать инвариантными свойствами и обеспечивать формальное описание системы.

Для структурных предположений введем частичный порядок. В этом случае множество структурных предположений вместе с частичным порядком образуют решетку, которая позволяет проводить процесс поиска «наилучшего» структурного предположения не по всему множеству возможных структурных предположений, а только по ограниченной части его.

Определим структуру как семейство подмножеств множества переменных **V**. Множеством всех возможных структур по отношению к множеству **V** будет множество

$$\mathbf{S} = \{\mathbf{S}_i | \mathbf{S}_i \in \mathbf{P}(\mathbf{V})\},$$

где **P**(**V**) — мощность множества **V**.

Проведем классификацию введенного множества S. Определим отображение

$$\mathbf{r}{:}\mathbf{S} \to \mathbf{R},$$

где **R** — множество всех симметричных бинарных отношений, определенных на множестве **V**, и $\mathbf{r}(\mathbf{S}_i)$ — бинарное отношение, в котором переменные v_i и v_j связаны тогда и только тогда, когда они обе принадлежат, по крайней мере, одному из подмножеств множества **V**, через которое определена $\mathbf{S}_i$;

$$\mathbf{r}(\mathbf{S}_i)=\{(v_i, v_j)|\ v_i, v_j \in \mathbf{V},\ \exists\ (\mathbf{E}_a \in \mathbf{S}_i)(v_j \in \mathbf{E}_a,\ v_i \in \mathbf{E}_a)\},$$

где $\mathbf{E}_a$ — подсистема структуры $\mathbf{S}_i$.

Отображение **r** невзаимооднозначное, так как $\mathbf{S}_i \equiv \mathbf{S}_j$, если $\mathbf{r}(\mathbf{S}_i)=\mathbf{r}(\mathbf{S}_j)$. Будем использовать символ **S/r** для обозначения эквивалентных классов в **S** посредством **r**.

Некоторые структуры из множества **S** могут не сохранять первоначального понятия об объекте исследования, следовательно, необходимо, введя некоторые ограничения на множество **S**, сузить множество всех возможных структур до множества значимых структур, дающих полное представление о проблеме. Введем эти ограничения:

— множество связанных переменных структуры должно быть полным;

— каждый элемент структуры должен быть определен непустым множеством переменных;

— не должно существовать элемента, который состоял бы только из переменных, включенных полностью в некоторый другой элемент этой же структуры. В результате получается множество $\boldsymbol{S}_g$, называемое множеством приемлемых структур:

$$\mathbf{S}_G = \left\{ \mathbf{S}^i{}_G \,|\, \mathbf{S}^j{}_G \subset \mathbf{P}(\mathbf{V}),\ \cup \mathbf{E}_a = \mathbf{V}, \mathbf{E}_a \not\subset \mathbf{E}_b, \forall \mathbf{E}_a, \mathbf{E}_b \in \mathbf{S}^i{}_G\ (a \neq b) \right\}$$

Пусть требуется разработать комплекс программ, который позволял бы генерировать рекомендации по созданию всего множества приемлемых структур и оценивать их для заданного множества переменных. Как отмечалось ранее, множество приемлемых структур вместе с определенным на этом множестве частичным порядком образуют решетку. Рассмотрим подробно понятие «частичный порядок» для множества приемлемых структур.

Предположим, что $\mathbf{S}^i{}_G$ и $\mathbf{S}^j{}_G$ принадлежат множеству $\mathbf{S}_G$, где $\mathbf{S}^i{}_G \leq \mathbf{S}^j{}_G$ тогда и только тогда, когда для каждого множества $\mathbf{E}_a \in \mathbf{S}^i{}_G$ существует множество $\mathbf{E}_b \in \mathbf{S}^j{}_G$ такое, что $\mathbf{E}_a \subseteq \mathbf{E}_b$. Если $\mathbf{S}^i{}_G \leq \mathbf{S}^j{}_G$, то будем называть $\mathbf{S}^i{}_G$ «уточнением» структуры $\mathbf{S}^j{}_G$ и, наоборот, $\mathbf{S}^j{}_G$ «агрегатом» $\mathbf{S}^i{}_G$. Кроме того, $\mathbf{S}^i{}_G$ — непосредственное «уточнение» $\mathbf{S}^j{}_G$ тогда и только тогда, когда $\mathbf{S}^i{}_G \leq \mathbf{S}^j{}_G$ и не существует такой $\mathbf{S}^k{}_G$, что $\mathbf{S}^i{}_G \leq \mathbf{S}^k{}_G$ и $\mathbf{S}^k{}_G < \mathbf{S}^j{}_G$. Аналогично определяется непосредственный «агрегат».

Определим отображение как

$$\mathbf{r}_G : \mathbf{S}_G \rightarrow \mathbf{R}^*,$$

где $\mathbf{R}^*$ — множество всех симметричных и рефлексивных бинарных отношений, определенных на множестве **V**. Это множество можно интерпретировать как множество ненаправленных графов с петлями. Символом $\mathbf{S}_G/\mathbf{r}_G$ обозначим множество классов эквивалентности, определенных на множестве приемлемых структур посредством $\mathbf{r}_G$. Каждому классу эквивалентности соответствует единственная каноническая структура, содержащая минимальное число элементов. Между множеством канонических структур $\mathbf{S}_C$ и множеством ненаправленных графов, определенных на данном множестве переменных, существует взаимно однозначное соответствие. Каноническая структура всегда будет агрегатом для структур, принадлежащих классу эквивалентности, определяемому этой структурой. Введем также структуры, которые будут являться уточнением для любой структуры из определенного класса эквивалентности. Они построены так же, как пары переменных, которые связаны в графе, представляющем класс эквивалентности, и как единичные переменные, которые изолированы в графе. Будем называть введенные структуры соответственно **С**- и **Р**-структуры. Очевидно, что $\mathbf{C} \geq \mathbf{P}$ и все остальные структуры (из определяемого этими структурами класса эквивалентности) будут всегда находиться между **С** и **Р**.

Из этого следует, что структурные предположения, не входящие в указанное множество $\mathbf{S}_G$, могут быть игнорированы без каких-либо потерь для множества анализируемых синтезированных структурных предположений.

Процесс синтеза структурных предположений системы можно проводить различными способами в зависимости от ограничений, накладываемых консультантом на начальную структуру системы.

Во-первых, если разработчик считает, что все связи между функциональными компонентами, заданными в начальной структуре, «жесткие», т. е. удаление какой-либо связи повлечет за собой коренное изменение в функциональной принадлежности начально определенных агрегатов компонент структуры, то процесс синтеза новых структурных предположений необходимо проводить в рамках выделенного класса эквивалентности, соответствующего начальной структуре. В этом случае механизм синтеза структурных предположений заключается в агрегировании либо в декомпозиции элементов начальной структуры системы.

Во-вторых, если разработчик считает, что удаление связей возможно и даже необходимо, то процесс синтеза основывается на переходе из одного класса эквивалентности структурных предположений в другой, причем связи можно и удалять и добавлять.

Для правильного понимания физического смысла таких понятий, как «удаление связей», «добавление связей», «агрегирование и декомпозиция элементов» и т. д., необходимо уяснить одно из центральных понятий структурного синтеза — поведение системы, которое описывается следующим образом:

— определяется множество переменных, через которые наблюдается система — множество **V** (например, множество функциональных компонент, необходимых для функционирования системы);

— каждой переменной задается определенное состояние из множества состояний данной переменной. Совокупность состояний множества переменных — множество **X** (например, конкретные реализации функциональных модулей);

— на множестве **V** определяется функция *f*, ставящая в соответствие каждой переменной (функциональной компоненте) определенное состояние (конкретную реализацию) ;

— определяется множество агрегатов состояний **A** (в данном случае — множество возможных конфигураций структуры);

— вводится отображение *b*, ставящее в соответствие каждому агрегату состояний некоторое число их [0, 1], в данном случае — назначение каждой конфигурации синтезируемой структуры определенной вероятности (если сложно выделить приоритеты различных конфигураций системы, то возможные конфигурации принимаются равновероятными).

Описав таким образом поведение системы, нетрудно проследить качественные изменения, происходящие в процессе синтеза новых структурных предположений.

В качестве исходных данных для процедур синтеза имеется начальная структура системы, а также сформулированное описанным способом поведение системы. Используя процедуры генерации, можно получить множество новых структурных предположений в рассматриваемой проблеме синтеза и оценить эти сгенерированные структурные предположения.

Для проведения оценки необходимо выяснить, как соотносятся эти сгенерированные структурные предположения с начальным поведением структуры. Предлагается использовать структурное предположение (совокупность подмножеств множества переменных или функциональных модулей системы) в качестве некоторой маски, накладываемой на начальное поведение системы (множество реализаций рекомендаций) для получения нового поведения, соответствующего сгенерированному структурному предположению. Далее эти оба поведения (новое и начальное) можно сравнить по множеству агрегатов состояний **A** и значению вероятностей появления данного конкретного агрегата в поведении системы, определенном структурным предположением. Естественно, новое поведение, соответствующее сгенерированному структурному предположению, может иметь агрегаты состояний (возможные конфигурации структуры), которые не были описаны в начальном поведении. В этом случае к рассмотрению этих агрегатов состояний необходимо подойти с особой внимательностью, и если полученные агрегаты состояний представляют собой допустимые конфигурации структуры, то целесообразно включить эти агрегаты состояний в начальное поведение и последующие итерации проводить уже с модифицированным начальным поведением. ***Таким образом, структура и начальное поведение определяют поведение системы, имеющей эту структуру***. Следовательно, одним из критериев выбора структуры системы может являться факт сохранения начального поведения системы для сгенерированного структурного предположения. Сохранение означает совпадение у начального и нового поведения агрегатов состояний и значений вероятностей появления этих агрегатов.

В этом случае структурные предположения, полученные в результате первой итерации и сохранившие начальное поведение, выбираются в качестве начальных структур для последующей итерации и т. д.

В конкретных реализациях сформированных рекомендаций, естественно, кроме критерия выбора структур для последующих итераций используется и ряд других критериев, например минимизация количества элементов в структуре и т. п.

Учитывая тот факт, что совокупность структурных предположений образует решетку, можно определить направление синтеза новых структурных предположений.

С чисто практической точки зрения наиболее интересным является синтез структурных предположений в рамках определенного класса эквивалентности. Как правило, разработчик считает свою модель системы (начальное поведение и структуру) «жесткой», т. е. установленные им связи (отношения между состояниями переменных) неизменными. На самом деле, начальная структура в большинстве случаев избыточна и содержит различного рода противоречия. Для того, чтобы корректно задать начальное поведение и структуру системы, необходимо на первом шаге промоделировать исходное структурное предположение. Используя метод интерпретирующего структурного моделирования, можно выделить компоненты связанности в проблеме синтеза, проверить структуру системы на достижимость переменных, определить уровни иерархии системы и т. д. Только после проведенного таким образом анализа начальной структуры можно приступать к синтезу новых структурных предположений системы.

Например, формируя рекомендации для проектирования системы документирования, на вход которой поступает информация от других систем, и необходимо сформировать на выходе этой системы описания свойств, спецификацию компонент в виде таблицы, а также построить структурную схему. Представим, что состояние информации в системе описывается четырьмя переменными **V** = (v_1, v_2, v_3, v_4):

$$v_1 = \begin{cases} 1 \text{ — наличие входной информации;} \\ 0 \text{ — отсутствие входной информации;} \end{cases}$$

$$v_2 = \begin{cases} 0 \text{ — отсутствие описания;} \\ 1 \text{ — описание свойств;} \\ 2 \text{ — описание свойств и элементов;} \\ 3 \text{ — описание свойств и внутренних связей;} \\ 4 \text{ — описание свойств, элементов и внутренних связей;} \end{cases}$$

$$v_3 = \begin{cases} 0 \text{ — отсутствие спецификации компонентов;} \\ 1 \text{ — наличие таблицы спецификации;} \end{cases}$$

$$v_4 = \begin{cases} 0 \text{ — отсутствие структурной схемы;} \\ 1 \text{ — наличие структурной схемы.} \end{cases}$$

Пусть в результате предварительного анализа было получено требуемое поведение системы, представленное табл. 1.

Таблица 1

v_1	v_2	v_3	v_4
0	0	0	0
1	1	0	0
1	2	1	0
1	3	0	1
1	4	1	1

Появление каждого агрегата в этом поведении принимается равновероятным с вероятностью, равной 0,2.

Пусть в качестве начальной структуры выбрана структура вида

$$\mathbf{S}_0=(v_1, v_2, v_3, v_4).$$

Это означает, что $\mathbf{S}_0$ состоит из одного элемента и все переменные попарно связаны. Далее, применяя процедуры декомпозиции данной структуры на подсистемы, получаем новые структурные предположения.

В соответствии с начальным поведением для каждой синтезированной структуры вычисляются поведение, а также степень отличия вычисленного поведения от начального. Например, для структур

$$\mathbf{S}_1 = (v_1, v_2, v_3)(v_2, v_3, v_4);$$

$$\mathbf{S}_2 = (v_1, v_2, v_4)(v_2, v_3, v_4);$$

$$\mathbf{S}_3 = (v_1, v_2, v_3)(v_1, v_2, v_4)$$

вычисленное поведение совпадает с начальным и степень отличия равна нулю. Взяв эти структуры в качестве исходных для последующих итераций синтеза и применив процедуру декомпозиции, получаем множество структур, в которых изменились связи между переменными по сравнению с исходными структурами. Для полученных структур также вычисляется поведение и сравнивается с исходным.

Структуры, выделенные из множества синтезированных структур, согласно критерию сохранения поведения систем показаны на рис. 10.

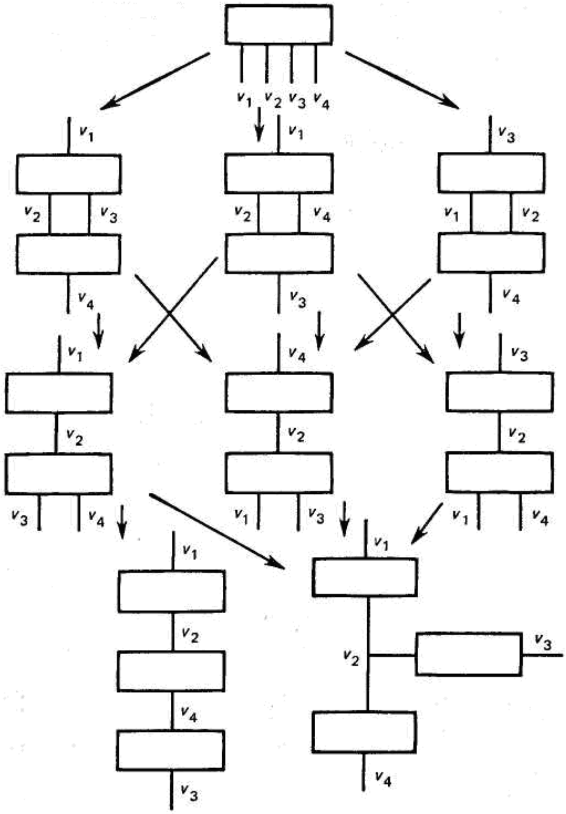


Рис. 10. Возможные структуры системы документирования

В качестве структуры проектируемой системы можно принять любую из этих структур, но необходимо также учитывать некоторые другие критерии, по которым можно сузить это множество, например критерий минимума подсистем в системе или минимума связей между подсистемами. Проведенный анализ синтезированных структур показал, что начальному поведению системы удовлетворяют и наиболее приемлемы структуры, изображенные на рис. 11.

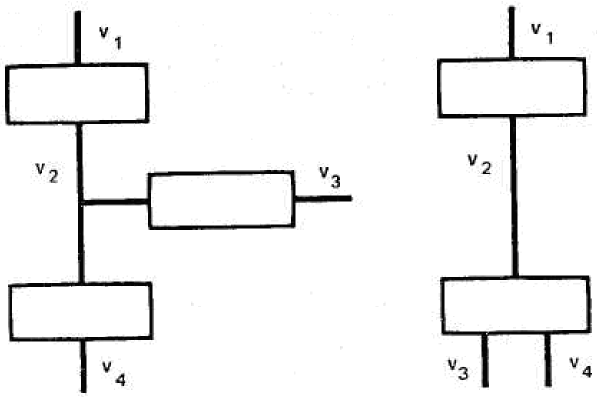


Рис. 11. Структуры системы, удовлетворяющие начальному поведению

4. Методы синтеза систем и их структур

4.1. Постановка задачи синтеза систем и их структур

Чтобы определить функции и структуры системы, необходимо:

1) составить задачу функционирования системы;
2) выбрать алгоритм их решения;
3) сформировать общую структуру системы и распределить задачи ее функционирования по уровням системы;
4) скомпоновать комплекс технических средств системы и ее подсистем.

Перечисленные выше этапы синтеза системы взаимно связаны, и задачи каждого из них решаются с учетом ресурсов, выделяемых на создание системы. Синтезированная системы считается оптимальной, если достигается максимум (минимум) выбранного показателя эффективности, отражающего основные свойства системы с точки зрения выполнения поставленных задач.

На начальном этапе синтеза системы функции системы представляются в виде совокупности взаимосвязанных задач, которые, в свою очередь, могут быть разбиты на совокупности операций. По известным характеристикам операций и их взаимосвязей могут быть получены соответствующие характеристики задач.

При формализации взаимосвязей между функциями обычно учитываются порядок следований операций и их длительности (временные связи). Формализация взаимосвязей обычно производится на основе теории графов. Постановки задачи синтеза системы будем осуществлять, исходя из следующих условий:

1) задано или подлежит выбору множество реализуемых функций;

2) заданы или подлежат выбору взаимосвязи между функциями, в том числе и временные;

3) заданы или подлежат выбору элементы комплекса технических средств (КТС);

4) учитывается или не учитывается расположение элементов на плоскости (в пространстве);

5) заданы или подлежат выбору связи между элементами системы;

6) учитывается или не учитывается возможность выполнения задачи несколькими элементами.

Одновременно с этими условиями задачи синтеза системы могут различаться:

а) видом показателей эффективности;

б) типом учитываемых характеристик элементов;

в) видом ограничений, накладываемых на учитываемые ресурсы (временные, сырьевые, технико-экономические и т. д.).

При решении задач синтеза системы находят применение различные модели и методы. Широкое распространение, как известно, получили модели математического программирования. Рассмотрим постановку задачи синтеза системы на этапе синтеза структуры системы по изготовлению объектов.

Считаем, что предполагаемая топология размещения модулей системы известна; функции, реализуемые модулями, перечислены в виде последовательности задач, которые необходимо распределить между модулями системы. При построении структуры системы необходимо распределить функции (задачи) ($i = 1,2, ...,I$) между модулями системы (j=1,2,...,J), выбрать алгоритмы реализации функции ($k = 1, 2, ..., K$), типы технических средств модуля ($l = 1,2,...,L$) и вариант организации функциональной связи между модулями. При этом должны учитываться такие технико-экономические характеристики системы, как затраты на создание A и эксплуатацию B, длительность выполнения цикла реализации функции (задачи) T, надежность системы P, масса W, энергопотребление E и др. Для формализации постановки задачи введем переменные:

x_{ikjl} = 1, если i-*я* функция (операция) реализуется k-м вариантом в j-м модуле при помощи l-го технического средства;

x_{jl} = 1, если j-й модуль имеет в своем составе l-е техническое средство;

$x_{jj'}$=1, если необходимо создать функциональную связь между модулями j и j';

$x_{ikjl} = x_{jl} = x_{jj'} = 0$ — в противном случае.

Тот факт, что каждый вариант построения структуры должен включать в себя лишь один способ распределения функций (операций) по модулям системы и один способ выполнения каждой задачи, учитывает ограничение.

$$\sum_{k,j,l} x_{ikjl} = N, \quad i = 1, 2, \ldots, I$$

Переменные x_{jl}, x_{jj}' зависят от x_{ikjl} и используются для удобства записи аналитических выражений качества различных характеристик вариантов структуры:

$$x_{jl} = \begin{cases} 1, & \text{если} \quad \sum\limits_{i,\,k} x_{ikjl} \geqslant 1; \\ 0, & \text{если} \quad \sum\limits_{i,\,k} x_{ikjl} = 0; \end{cases}$$

$$x_{jj'} = \begin{cases} 1, & \text{если} \quad \sum\limits_{ikl'k'll'} x_{ijkl} x_{i'k'j'l'} \geqslant 1; \\ 0, & \text{если} \quad \sum\limits_{ikl'k'll'} x_{ijkl} x_{i'k'j'l'} = 0. \end{cases}$$

Для определения характеристик вариантов структуры введем следующие обозначения: A_l — стоимость технического средства либо затраты на разработку и изготовление перспективных средств; $A_{jj'}$ — стоимость создания функциональной связи; $A_{ikj'l}$ — затраты на разработку необходимых видов обеспечения решения (реализации) l-й функции (операции) в k-м варианте в j'-м модуле при наличии l-го технического средства; $B_{ij(i+1)j'}$ — затраты, понесенные в результате перехода от i-й функции (операции), реализуемой в j-м модуле, к (i + 1)-й функции (операции), реализуемой j'-м модуле; t_{ikjl} — время реализации i-й функции k-м способом в j-м модуле при наличии l-го технического средства; $t_{ij(i+1)j'}$ — время, затрачиваемое на переход от i-й функции, реализуемой в j-м модуле, к (i + 1)-й функции, реализуемой в j'-м модуле; P_{ikjl} — надежность реализации функции (технического средства), определяемая либо вероятностью безотказной работы, либо ее функцией, $P_{ij(i+1)j'}$ — надежность функциональной связи, определяемая как и P_{ikjl} ; W_e — масса технического средства; E_e—энергопотребление технического средства.

Используя введенные обозначения, запишем математические выражения для определения характеристик различных вариантов структуры системы.

Капитальные затраты *A* включают в себя стоимость технических средств в модуле, стоимость создания функциональных связей между модулями системы, затраты на разработку алгоритмов выполнения функций (операций):

$$A=\sum_{l,\, l} A_l x_{lj}+\sum_{l,\, l'} A_{ll'} x_{ll'}+\sum_{i,\, k,\, j,\, l} A_{ikjl} x_{ikjl}.$$

Эксплуатационные затраты *B* включают в себя затраты на реализацию функционирования системы и затраты на транспортирование полуфабрикатов между модулями:

$$B=\sum_{i,\, k,\, j,\, l} B_{ikjl} x_{ijkl}+\sum_{k',\, i,\, k,\, j,\, l,\, j',\, l'} B_{ij\,(i+1)\,j'} x_{ikjl} x_{(i+1)\,k'j'l'}.$$

Время выполнения цикла функционирования системы вычисляется аналогично:

$$T=\sum_{i,\, k,\, j,\, l} t_{ikjl} x_{ijkl}+\sum_{k',\, i,\, k,\, j,\, l,\, j'l'} t_{ij\,(i+1)j}\, x_{ijkl} x_{(i+l)\,k'j'l'}$$

При определении надежности системы будем считать систему невосстанавливаемой, в результате чего ее надежность, характеризуемая вероятностью безотказной работы элементов, равна произведению вероятностей безотказной работы элементов. Поэтому

$$\bar{P}=\sum_{i,\, j,\, k,\, l} \bar{p}_{ikjl} x_{ikjl}+\sum_{k',\, i,\, j,\, k,\, l,\, j',\, l'} \bar{p}_{ij\,(i+1)\,j'} x_{ikjl} x_{(i+1)\,k'j'l'},$$

(1)

где

$$\bar{p},\ \bar{p}_{ikjl},\ \bar{p}_{ij\,(i+1)\,j'}$$

— логарифмы соответствующих величин. Масса и энергопотребление технических средств в модулях системы:

$$W_j=\sum_{i,\, k,\, l} W_l x_{ikjl};$$

$$E_j=\sum_{i,\, k,\, l} E_l x_{i\,kjl},\quad j=1,\ 2,\ \ldots,\ J.$$

(2)

При решении задачи синтеза структуры любая из указанных характеристик может быть выбрана в качестве показателя эффективности (в зависимости от цели и назначения системы), который оптимизируется, а другие учитываются в ограничениях. Поэтому в общем виде задача оптимизации структуры системы выглядит следующим образом:

$$\min(\max) f(x_{ijkl}, x_{jl}, x_{jj'}). \tag{3}$$

При ограничениях:

$$\begin{gathered} f_s(x_{ikjl}, x_{jl}, x_{jj'}) \leqslant D_s, \quad s = 1, 2, \ldots, S; \\ \sum_{k,j,l} x_{ijkl} = N, \quad i = 1, 2, \ldots, I; \\ x_{ijkl} = (0; 1); \quad x_{jl} = (0; 1); \quad x_{jj'} = (0; 1). \end{gathered} \tag{4}$$

Из приведенной постановки задачи следует, что она является задачей дискретного программирования и носит комбинаторный характер. При решении таких задач возникают трудности принципиального характера, а именно, необходимо исключить явный перебор всех допустимых решений и стремиться к эффективному частичному перебору сравнительно малого числа допустимых вариантов решения соответствующей задачи и неявному перебору остальных.

4.2. Критерии синтеза

Выбор путей использования системы будем производить в два этапа. На первом этапе решим вопрос о целесообразности использования системы по результатам оценки экономической эффективности, на втором, если принято решение о целесообразности использования системы, произведем выбор рациональной структуры и оптимального варианта системы с учетом определенного числа факторов и условий.

Рассмотрим основное содержание первого этапа. Прежде всего следует выполнить прикидочную оценку целесообразности использования системы, установить ее рентабельность. Разработка и внедрение системы представляет практический интерес, если суммарный экономический выигрыш за весь срок ее функционирования будет превышать суммарные расходы на ее разработку, внедрение и эксплуатацию. Это означает, что в начальный период времени система окупает те затраты, которые в нее вложены, а в дальнейшем до конца ее жизненного цикла дает экономию (прибыль).

Основной критерий экономической эффективности (абсолютная прибыль от использования системы) выразится соотношением

$$F(x) = \sum_{i=1}^{n} F_i x_i, \tag{5}$$

где F_i — абсолютная прибыль от внедрения i-го варианта системы.

Значение F_i определим по формуле:

$$F_i=u_iT-C_i(T). \qquad (6)$$

Здесь u_i — среднегодовой экономический выигрыш при внедрении i-го варианта системы; C_i — суммарная стоимость

$$C_i(T)= C_{0i\ i} + s_iT, \qquad (7)$$

где C_{0i} — стоимость капиталовложений при реализации i-го варианта системы; s_t — годовые эксплуатационные расходы от реализации i-го варианта системы.

При $F_i \leq 0$ реализация i-го варианта системы является нецелесообразной, так как такой вариант системы является нерентабельным.

Оценка целесообразности разработки системы производится по наилучшему варианту, удовлетворяющему соотношение

$$F^* = F(X^*) = \max \sum_{i=1}^{n} F_i x_i. \qquad (8)$$

Можно оценку целесообразности производить по совокупности критериев (5), (8), пользуясь ***теорией аддитивной последовательности.*** Тогда, обозначив:

$\overline{F}_1(X)$ — относительное значение абсолютного дохода от реализации системы;

$\overline{F}_2(X)$ — относительное время, в течение которого реализованная системы дает прибыль, окупив полную ее стоимость.

Запишем

$$\overline{F} = \overline{F}_1(X)g_l + \overline{F}_2(X)g_2. \qquad (9)$$

где g_1 и g_2 — вес каждого из критериев (в частности, можно положить, что $g_1 = g_2 = 1$).

Наилучший вариант системы, используемый для оценки целесообразности ее разработки, определим на основании соотношения:

$$\overline{F}^* = \overline{F}(X^*) = \max \sum_{i=1}^{n} \left(\overline{F}_{i1}g_1 + \overline{F}_{i2}g_2\right) x_i, \qquad (10)$$

где

$$F_{i1} = u_i/u_0 - c_i(T)/u_0T; \quad F_{i2} = 1 - c_i(T)/u_iT_i; \quad u_0$$

— предельное значение среднегодового экономического дохода (при внедрении некоторой идеальной системы).

Максимальное значение критерия $F(X^*)$ или $\overline{F}(X^*)$ по формуле (8) или (10) находят при выполнении ряда условий или ограничений на переменные x_i.

Ограничение на численные значения переменных X отражает тот факт, что $x_i=1$, если 1-й вариант системы выбирается для оценки, и $x_i = 0$, если не выбирается, а также то, что для оценки выбирается наилучший вариант:

$$\sum_{i=1}^{n} x_i = 1. \tag{11}$$

При нахождении решения (8) и при использовании в качестве критерия абсолютного экономического дохода должно выполняться ограничение:

$$F_i x_i \geq F^{(u)}. \tag{12}$$

Наложим ограничение на критерий времени

$$F_i x_i \geq F^{(t)}, \tag{13}$$

где $F^{(t)}$ — минимально допустимое значение времени, в течение которого система дает прибыль.

При нахождении (10) должны выполняться условия:

$$\overline{F}_i^{(1)} x_i \geqslant \overline{F}^{(u)}; \tag{14}$$

$$\overline{F}_i^{(2)} x_i \geqslant \overline{F}^{(t)}. \tag{15}$$

Ограничение по максимально допустимой стоимости на разработку, внедрение и эксплуатацию можно записать в следующем виде:

$$\sum_{i=1}^{n} c_i x_i \leqslant c_m.$$

Может быть наложено ограничение по максимально допустимой стоимости на разработку, внедрение и эксплуатацию системы, по условию допустимой стоимости

$$c'_m < c_m.$$

Следовательно, в общем случае, допустимые затраты на разработку, внедрение и эксплуатацию системы определяются из соотношения $c = \min (c_m, c'_m)$, и соответствующее ограничение запишутся в виде

$$\sum_{i=1}^{n} c_i x_i \leqslant c. \tag{16}$$

Кроме ограничения сверху на суммарную стоимость системы бывает необходимым введение ограничения по минимально возможной стоимости системы. Это ограничение отражает то обстоятельство, что для системы данного назначения существует минимальное значение стоимости, уменьшение которой приводит к принципиальной невозможности ее создания. Иными словами, существует тот минимум средств, который позволяет решить задачу наиболее простой системы данного назначения. Указанное ограничение запишем в виде

$$\sum_{i=1}^{n} c_i x_i \leqslant \underline{c}, \qquad (17)$$

где $\underline{c}$ — минимально возможная стоимость системы

Объединяя неравенства (16) и (17), получим ограничение по суммарной стоимости системы

$$\sum_{i=1}^{n} c_i x_i \leqslant \bar{c}. \qquad (18)$$

Таким образом, оценка целесообразности разработки и внедрения системы производится:

- при выборе в качестве критерия абсолютного дохода от внедрения системы — с помощью решения задач (8), (11), (12), (18);
- при выборе в качестве критерия времени, в течение которого система дает прибыль — путем решения задач (8), (11), (13), (18);
- при использовании двухкритериальной оценки — решением задач (10), (11), (14), (15), (18).

В этом случае, если решение существует, то разработка системы целесообразна; если же указанного решения не существует, то разработка системы нецелесообразно.

4.3. Особенности решения задач структурного синтеза

Задача структурного синтеза заключается в выборе принципа действия системы и в определении оптимальной структуры системы для реализации заданных функций. Структура системы определяется природой входящих в нее элементов и физической реализацией связей между ними в составе системы. Структуры наглядно изображают, как устроена системы — из каких частей она состоит и как эти части связаны друг с другом. Одной из форм математического описания структуры, как мы знаем, является граф. Напомним, что вершины графа отождествляются с элементами системы (на принятом уровне детализации), а дуги и ребра графа — со связями между соответствующими элементами. Форма отображения структуры — схема. Схема и граф тождественны по своему содержанию, но различаются по форме. В схемах для обозначения элементов используют различные геометрические фигуры, разнообразие форм которых облегчает чтение схем.

Принцип, по которому объединение элементов приводит к появлению новых свойств, отличных от свойств элементов, называют *принципом организации*. ***Организация*** ***— понятие более высокого***

ранга, чем функция и структура. Различные принципы организации могут приводить к построению системы, различающихся своими структурами, но тождественным по своему функциональному назначению.

Сложность *— свойство системы, заключающееся в том, что функция, реализуемая системой, не может быть представлена в виде композиции функций, реализуемых элементами системы.* Например, при структурном синтезе ЭВМ рассматривается как система, состоящая из взаимосвязанных функциональных блоков и узлов, организованных таким образом, чтобы их функционирование приводило к реализации заданных функций — вычислениям на основе алгоритмов. Одни и те же функции могут быть реализованы различными структурами, обеспечивающими производительность решения задач при различных затратах оборудования. Закон функционирования ЭВМ невозможно рассмотреть только с точки зрения электрических процессов, происходящих в цепях ЭВМ. Функции ЭВМ выявляются лишь при рассмотрении процессов в ЭВМ в информационном и алгоритмическом аспектах. ***Это объясняется эффектом организации, порождающим в совокупностях элементов новые свойства.***

Процесс формирования структуры сложной системы представляет собой многоэтапную процедуру, осуществляемую по блочно-иерархическому принципу:

Этап 1. Построение допустимого множества (каталога) принципиально возможных типов, входящих в систему элементов. Примерами могут служить: каталоги типовых рекомендаций в САК больших систем; каталог допустимых проектных модулей в САК; каталог типовых проектов в САК и др. Каталоги входят в состав информационного обеспечения САК.

Этап 2. Выбор множества допустимых элементов, используемых при структурном синтезе. Поскольку различные элементы, входящие в каталог применяемых элементов, могут быть реализованы на основе разнообразных физических принципов и с помощью различных решений, не все элементы могут стыковаться друг с другом. Поэтому основной задачей данного этапа является определение требований к параметрам и принципам функционирования отдельных элементов, входящих в состав системы.

Этап 3. Построение вариантов структур системы с учетом их функционирования. Здесь должны быть известны состав элементов, правила их соединения между собой и способ определения по структуре синтезируемого объекта функцию, которую он реализует.

При структурном синтезе небольшой сложности возможно построение полного множества допустимых структур (для реализации полного перебора вариантов). При этом в ЭВМ должны быть заложены правила генерации всех вариантов структур системы.

При синтезе сложных структур прямой перебор уже невозможен и необходима разработка процедур и алгоритмов направленного поиска оптимальной структуры системы. Эти процедуры обычно базируются на использовании методов математического программирования (в основном — дискретного программирования), последовательных и итерационных алгоритмов синтеза, сетевых и графовых моделей системы, а также методов теории эвристических решений и методов решений проектных задач.

Этап 4. Оценка вариантов и выбор компромиссной структуры системы. Как правило, оценка варианта структуры требует формирования и анализа математической модели синтезированной структуры и выполнения параметрической оптимизации, так как для объективной оценки сравнивать варианты структуры имеет смысл при оптимальных значениях параметров. Эти процедуры сложны и громоздки, в связи с чем полный перебор вариантов при таком подходе практически неосуществим.

Для уменьшения сложности этого этапа целесообразно либо использовать косвенные критерии предпочтения вариантов, либо искать оценки варианта структуры без исследования громоздких математических моделей. При таком подходе вводят параметр, характеризующий качество системы. Это может быть число элементов в системе, ее стоимость, занимаемый объем, максимальное число элементов, находящихся в активном состоянии (мощность), вероятность нерешения проблемы, максимальная длина путей (в задачах размещения и трассировки) и т.д.

Другим путем уменьшения сложности решения задачи структурного синтеза является организация диалогового режима пользователя с ЭВМ на 3-м и 4-м этапах синтеза. При этом пользователь сам решает, какие программы анализа и оптимизации будет использовать для оценки вариантов. Сокращение времени на формирование рекомендации в диалоговом режиме происходит за счет эвристических способностей человека, за счет возможности прерывания построения заведомо бесперспективного варианта структуры и за счет поиска не оптимального, а допустимого варианта синтезируемой структуры.

Этап 5. Коррекции технического задания. Под коррекцией ТЗ понимают изменения заданных ограничений в тех случаях, когда не существует вариантов рекомендаций, обеспечивающих требуемое ограничение. Если коррекция произведена, то соответствующие этапы проектирования повторяют при новых значениях ограничений.

Объем решаемых на каждом этапе задач настолько велик, что проведение исследований в полном объеме невозможно без средств автоматизации проектирования. При этом на различных этапах проектирования привлекаются различные специализированные коллективы научных работников и инженеров. Следует отметить, что при использовании отдельных элементов, устройств и подсистем, входящих в состав сложной системы, следует проводить их анализ и синтез на различных уровнях. В процессе проектирования приходится учитывать существующие многократные перекрестные связи между элементами, что существенно усложняет задачу структурного синтеза.

Исходя из вышесказанного следует, что особенность задач структурного синтеза заключается в том, что для получения оптимального варианта структуры системы необходимо наличие ее математической модели, представляющей собой формальное описание множества структур системы на принятом уровне детализации. В этом случае задача структурного синтеза сводится к выбору компромиссного варианта в счетном множестве.

4.4. Методы выбора структуры

Методы выбора структуры системы рассмотрим на примере системы автоматизированного конструирования (САК)

Структура САК характеризуется прежде всего числом уровней функционирования, наличием в составе САК различных модулей САК, числом и типами комплексов технических средств различного функционального назначения, схемой организации потоков конструктивных требований (запросов) между элементами и модулями САК и т. д. Смысл критериев выбора (5), (10) в том, что возможные их значения определяют возможные варианты САК с присущей им структурой. При таком представлении критериев выбора будем говорить о вариантном методе выбора структуры САК. Если же задан вектор параметров выбора

$$X = (x_1, ..., x_{\bar{k}}), \tag{19}$$

где компоненты x_k, $(k = 1, ..., \bar{k})$ означают число комплектов технических средств каждого k-го типа, входящих в САК, то возможные значения вектора X характеризуют различные структуры САК.

Изменение вектора (19) в общем случае приводит к изменению как номенклатуры, так и количества средств САК.

Метод решения задачи выбора структуры САК в терминах переменных вектора (19) назовем ***методом вариации состава средств САК.*** Задачи выбора структуры САК в методологическом плане аналогичны задачам оценки целесообразности ее разработки по полной совокупности вариантов САК. Различие состоит лишь в том, что на этапе выбора структуры САК учитывается существенно больше ограничений, более детально и глубже отражающих возможность реализации той или иной структуры, того или иного варианта САК. Ниже изложены два метода выбора рациональных структур САК: ***вариантный и метод вариации состава средств САК.***

Вариантный метод выбора структур САК изложим для случая использования в качестве критерия выбора абсолютного дохода от внедрения САК за весь срок ее функционирования. Показатели экономической эффективности (8), (10) запишем:

для стохастического принципа

$$F(X) = T \sum_{i=1}^{n} \sum_{y \in G} u_i(y)\, P(y) - \sum_{i=1}^{n} c_i x_i; \tag{20}$$

для детерминированного принципа

$$F(X) = T \sum_{i=1}^{n} u_i x_i - \sum_{i=1}^{n} c_i x_i, \tag{21}$$

где u_i *(у)* и u_i — среднегодовой экономический выигрыш за счет внедрения i-го варианта САК соответственно при случайном векторе состояний внешней среды и при осредненном его значении; $P(y)$ — плотность вероятности; c_i — суммарная стоимость разработки и внедрения i-го варианта САК.

Переменные x_i представленных выше показателей экономической эффективности должны быть подчинены ряду условий (ограничений), учитываемых при решении задачи. Правильная формализация и учет достаточно большого числа противоречивых ограничений является важной основой решения задач выбора рационального варианта структуры САК. Следует учитывать, что система ограничений зависит также от выбора показателя экономической эффективности. В качестве показателя экономической эффективности может быть принят один из приведенных в предыдущем пункте показателей, тогда некоторые из оставшихся функций могут быть переведены в разряд ограничений. Например, можно, взяв за показатель среднегодовой экономический выигрыш, стремиться его максимизировать при заданных ограничениях сверху на различные виды стоимости или, наоборот, выбрав стоимость, стремиться минимизировать затраты на САК при ограничениях снизу на среднегодовой экономический выигрыш от внедрения САК. Выше было определено, что за показатель экономической эффективности принимается абсолютный экономический доход от внедрения САК.

Сформулируем основные ограничения на переменные показателей (20) или (21) экономической эффективности внедрения САК применительно к вариантному методу выбора:

- по численным значениям переменных

$$\sum_{i=1}^{n} x_i = 1; \quad x_i = 0 \quad \text{или} \quad 1; \tag{22}$$

- по стоимости капиталовложений на создание и внедрение САК

$$\sum_{i=1}^{n} C_{0_i} x_i \leqslant C_0, \tag{23}$$

где C_{0_i} — капиталовложения на создание i-го варианта САК; C_0 — допустимые капиталовложения при внедрении САК.

Ограничения по различным составляющим капиталовложений:

- по стоимости технологической части САК

$$\sum_{i=1}^{n} \overline{C}_{0_i} x_i \leqslant \overline{C}_0 \tag{24}$$

($\overline{C}_{0_i}$ — стоимость технологической части i-го варианта САК; $\overline{C}_0$ — допустимые затраты на технологическую часть САК);

- по стоимости капитального строительства

$$\sum_{i=1}^{n} \widehat{C}_{0_i} x_i \leqslant \widehat{C}_0 \tag{25}$$

($\widehat{C}_{0_i}$ — стоимость строительства сооружений i-го варианта САК;

$\widehat{C}_0$ — допустимые затраты на капитальное строительство).

Ограничение по стоимости эксплуатационных расходов при внедрении САК можно определить из следующего выражения:

$$\sum_{i=1}^{n} C_{э_i} x_i \leqslant C_э, \tag{26}$$

где $C_{э_i}$ — эксплуатационные затраты при внедрении i-го варианта САК; $C_э$ — допустимые эксплуатационные затраты.

Ограничение по суммарной стоимости внедрения САК определяется как

$$\sum_{i=1}^{n} C_i x_i \leqslant C, \tag{27}$$

где C — допустимые суммарные затраты на разработку и внедрение САК.

Ограничение по абсолютной окупаемости САК. Отношение суммарной стоимости системы к среднегодовому доходу за счет внедрения САК представляет собой период (срок) абсолютной окупаемости внедряемых САК. Очевидно, для того, чтобы внедрение САК было экономически целесообразно, должны выполняться условия:

$$\sum_{i=1}^{n} [C_i x_i / \sum_{y \in G} u_i(y)\, p(y)] \leqslant T^{(A)} \tag{28}$$

или

$$\sum_{i=1}^{n} (C_i x_i / u_i) \leqslant T^{(A)}, \tag{29}$$

где $T^{(A)}$ — допустимый срок абсолютной окупаемости САК.

Ограничение по окупаемости САК. Часто наряду с ограничением по абсолютной окупаемости используют ограничение по окупаемости. Указанное ограничение выражается в виде:

$$\sum_{i=1}^{n} [C_{0_i} x_i / \sum_{y \in G} u_i(y)\, p(y)] \leqslant T^{(0)} \tag{30}$$

или

$$\sum_{i=1}^{n} (C_{0_i} x_i / u_i) \leqslant T^{(0)}, \tag{31}$$

где $T^{(0)}$ — допустимый срок окупаемости САК.

Ограничение общей численности личного состава, занятого эксплуатацией САК. Это условие записывается в форме следующего неравенства:

$$\sum_{i=1}^{n} N_i x_i \leqslant N, \tag{32}$$

где N_i — численность личного состава, занятого обслуживанием i-го варианта САК; N — допустимая общая численность личного состава на эксплуатацию САК.

Ограничения по числу специалистов того или иного профиля, занятых эксплуатацией САК. Пусть $\bar{\mu}$ — общее число профилей специалистов, необходимых для привлечения к эксплуатации САК; N_{μ_i} — число специалистов μ-го профиля ($\mu=1, \ldots, \bar{\mu}$), эксплуатирующих i-й вариант САК; N_μ — допустимое число специалистов μ-го профиля, тогда

$$\sum_{i=1}^{n} N_{\mu_i} x_i \leqslant N_\mu, \quad \mu = 1, \ \ldots, \ \bar{\mu}. \tag{33}$$

Ограничения производственных возможностей промышленных предприятий по поставке серийных средств автоматизации. Пусть для внедрения i-го варианта САК необходимо осуществить в определенные сроки поставки m_{ω_i} серийных средств автоматизации (комплектующих изделий) ω-го типа, тогда

$$\sum_{i=1}^{n} m_{\omega_i} x_i \leqslant m_\omega, \quad \omega = 1, \ \ldots, \ \bar{\omega}, \tag{34}$$

где m_ω — число серийных средств автоматизации ω-го типа, которое может выпустить промышленность за требуемый период поставки этих средств для внедрения САК с учетом уже имеющихся заказов на поставку аналогичных изделий по другим работам; величина m_ω характеризует свободные производственные мощности промышленных предприятий по выпуску изделий данного вида.

Ограничения по возможности разрабатывающих организаций.

Если для внедрения i-го варианта САК требуется разработка некоторых средств (изделий) автоматизации заново, то необходимо учитывать возможности таких разработок. По аналогии с предыдущим указанное ограничение запишется

$$\sum_{i=1}^{n} \alpha_{\psi_i} x_i \leqslant \alpha_\psi, \quad \psi = 1, \ \ldots, \ \bar{\psi}, \tag{35}$$

где α_{ψ_i} — число специалистов (или производственных мощностей) ψ-го вида, которых необходимо привлечь для разработки новых изделий в определенные сроки при создании i-го варианта САК; α_ψ — число свободных специалистов (свободные производственные мощности) разрабатывающих организаций.

Ограничения по срокам внедрения САК определяются как

$$\sum_{i=1}^{n} \bar{t}_i x_i \leqslant T_0, \tag{36}$$

где $\bar{t}_i$ — срок создания и внедрения i-го варианта САК; T_0 — максимально допустимый срок внедрения и ввода в эксплуатацию САК.

Изложенные задачи выбора оптимального (исходя из принятого показателя экономической эффективности) варианта САК заключаются в нахождении такого решения системы линейных неравенств (22)—(28), (30)—(36), которое обращало бы в максимум линейную функцию (20), либо в нахождении решения системы линейных неравенств (22)— (27), (29), (31)—(36), обращающего в максимум линейную функцию (22).

Рассмотрим кратко ***метод решения*** сформулированной задачи для более общего случая, когда требуется определить максимум функции (20).

Рассмотрим произвольный i-й шаг. К началу этого шага известны векторы

$$\| x_i^{(t-1)} \| = 0 \quad \text{и} \quad \| r_i^{(t-1)} \|.$$

Для первого шага

$$\| r_i^{(0)} \| = \| T \sum_{y \in G} u_i(y)\, p(y) - C_i \|;$$

находим

$$r_{i_t}^{(r-1)} = \max_i r_i^{(r-1)}$$

Затем проверяем неравенства:

$$C_{0_{i_t}} \leqslant C_0, \quad \overline{C}_{0_{i_t}} \leqslant \overline{C}_0; \quad \widehat{C}_{0_{i_t}} \leqslant \widehat{C}_0, \quad C_{*_{i_t}} \leqslant C_*, \quad C_{i_t} \leqslant C;$$

$$\frac{C_{i_t}}{\sum\limits_{y \in G} u_{i_t}(y)\, p(y)} \leqslant T^{(A)}, \quad \frac{C_{0_{i_t}}}{\sum\limits_{y \in G} u_{i_t}(y)\, p(y)} \leqslant T^{(0)};$$

$$N_{i_t} \leqslant N; \quad N_{\mu_{i_t}} \leqslant N_\mu, \quad \mu = 1, \ldots, \bar{\mu};$$

$$m_{\omega_{i_t}} \leqslant m_\omega; \quad \omega = 1, \ldots, \bar{\omega}; \quad \alpha_{\psi_{i_t}} \leqslant \alpha_\psi; \quad \psi = 1, \ldots, \bar{\psi}; \quad \bar{t}_{i_t} \leqslant T_0.$$

Если все неравенства удовлетворяются, то полагаем $x_{i_t} = 1$, $r^{(t)}_{i_t} = r^{(t-1)}_{i_t}$ и процесс нахождения максимума функции (20) заканчивается. Если хотя бы одно из приведенных выше неравенств нарушено, то полагаем x_{i_t} и $r^{(t)}_{i_t}=0$, после чего переходим к следующему шагу, повторяя изложенную процедуру. Процесс поиска заканчивается через число шагов $t \leq m$.

Рассмотрим принцип выбора структурных схем САК ***методом вариации состава средств*** САК. С учетом смыслового содержания введенного выше вектора (19) среднегодовой экономический доход от внедрения САК приближенно может быть представлен в виде

$$u(X) = u_0 [1 - \exp(-\sum_k \beta_k x_k)], \tag{37}$$

где u_0 — предельный среднегодовой экономический доход при внедрении наиболее совершенной (идеальной) САК; β_k — коэффициент, учитывающий прирост среднегодового экономического дохода при введении в состав САК одного КТС i-го типа.

Из приведенной зависимости видно, что искомая величина u зависит от переменных x_k, значения которых и подлежат определению. Выражение (37) можно записать в виде

$$u = u_0 \{1 - \exp[-\sum_{\delta \neq k} \rho_\delta x_\delta] \exp(\beta_k x_k)\}$$

или

$$u = u_0 [-A(x) \exp(-\beta_k x_k)],$$

где

$$A(X) = \exp(-\sum_{\delta = k} \beta_\delta x_\delta).$$

При $x_k = 1$ имеем

$$\overline{u} = u_0 [1 - A(X) \exp(-\beta_k x_k),$$

откуда

$$\beta_k = -\ln [(u_0 - u)/A(X) u_0].$$

Полученное выражение показывает, что, строго говоря, коэффициент $\beta_k = \beta_k(X)$. Однако с достаточной для практики точностью можно полагать β_k некоторым средним значением по множеству X или рассчитывать β_k для одного из вариантов структурной схемы САК. Вычисления показывают, что зависимость β_k от X не является резко выраженной, поэтому изложенный выше прием оправдан.

Выражение для β_k можно преобразовать следующим образом. При $x_k = 0$ имеем $u = u_0 [1 - A(X)]$, откуда $A(X) = (u_0 - \vartheta)/u_0$.

Подставляя значение $A(X)$ в формулу для β_k, получаем окончательное выражение для β_k:

$$\beta_k = -\ln [(u_0 - \bar{u})/(u_0 - \tilde{u})], \tag{38}$$

где $\overline{u}$ и ϑ рассчитываются для некоторого фиксированного варианта структурной схемы САК.

Точность приближения β_k тем выше, чем ближе этот фиксированный вариант САК к оптимальному в смысле принятого показателя экономической эффективности.

Применительно к рассматриваемому методу выбора структуры САК показатели экономической эффективности (20) и (21) записываются в форме

$$F(x) = T\,[1 - \exp(-\sum_k \beta_k x_k)] \sum_{y \in G} u_0(y)\, p(y) - \sum_k C_k x_k \tag{39}$$

и

$$F(X) = u_0 T \left[1 - \exp\left(-\sum_k \beta_k x_k\right)\right] - \sum_{k=1}^{\bar{k}} C_k x_k. \tag{40}$$

Рассмотрим ограничения на переменные показатели экономической эффективности (39), (40).

<u>Ограничения по численным значениям переменных</u>. Пусть I — множество КТС ξ-го функционального назначения ($\xi = 1, ..., \bar{\xi}$):

$$\sum_{k \in I_\xi} x_k \leqslant M_\xi, \quad \xi = 1, \ \ldots, \ \bar{\xi}, \tag{41}$$

где M_ξ — максимальное число КТС ξ -функционального назначения, которые могут входить в САК; x_k = 0, 1, 2, ... — целочисленные значения.

<u>Ограничения по стоимости капитальных вложений на создание САК</u>

$$\sum_{k=1}^{\bar{k}} C_{0_k} x_k \leqslant C_0. \tag{42}$$

<u>Ограничения по различным составляющим капиталовложений</u>:

- по стоимости технологической части САК

$$\sum_{k=1}^{\bar{k}} \overline{C}_{0_k} x_k \leqslant \overline{C}_0 \tag{43}$$

($\overline{C}_{0_k}$ — стоимость k-го элемента САК; $\overline{C}_0$ — допустимые затраты на технологическую часть САК);

- по стоимости строительства

$$\sum_{k=1}^{\bar{k}} \widehat{C}_{0_k} x_k \leqslant \widehat{C}_0 \tag{44}$$

($\tilde{C}_{0_k}$ — стоимость строительства сооружений для k-го элемента САК; $\tilde{C}_0$ — допустимые затраты на строительство).

<u>Ограничения по стоимости эксплуатационных расходов</u>

$$\sum_{k=1}^{\bar{k}} C_{\text{э}_k} x_k \leqslant C_{\text{э}}, \tag{45}$$

где $C_{\text{э}_k}$ — стоимость эксплуатационных расходов k-го элемента САК; $C_{\text{э}}$ — допустимая стоимость эксплуатационных расходов при внедрении САК.

<u>Ограничение по суммарной стоимости САК</u>

$$\sum_{k=1}^{\bar{k}} C_k x \leqslant C, \tag{46}$$

где C — допускаемая суммарная стоимость САК.

<u>Ограничения по абсолютной окупаемости САК</u>

$$\frac{\sum\limits_{k=1}^{\bar{k}} C_k x_k}{\left[1-\exp\left(-\sum\limits_{k} \beta_k x_k\right)\right] \sum\limits_{y \in G} u_0(y)\, p(y)} \leqslant T^{(A)} \tag{47}$$

или

$$\frac{\sum\limits_{k=1}^{\bar{k}} C_k x_k}{u_0 \left[1-\exp\left(-\sum\limits_{k} \beta_k x_k\right)\right]} \leqslant T^{(A)}. \tag{48}$$

<u>Ограничения по окупаемости САК</u>

$$\frac{\sum\limits_{k}^{\bar{k}} C_{0_k} x_k}{\left[1-\exp\left(-\sum\limits_{k} \beta_k x_k\right)\right] \sum\limits_{y \in G} u_0(y)\, p(y)} \leqslant T^{(0)} \tag{49}$$

или

$$\frac{\sum\limits_{k=1}^{\bar{k}} C_{0_k} x_k}{u_0 \left[1-\exp\left(-\sum\limits_{k} \beta_k x_k\right)\right]} \leqslant T^{(0)}. \tag{50}$$

<u>Ограничения по общей численности эксплуатационного состава САК</u>

$$\sum_{k=1}^{\bar{k}} N_k x_k \leqslant N, \tag{51}$$

где N_k — число людей, занятых эксплуатацией k-го элемента САК; N — допустимая общая численность эксплуатационного состава САК.

Ограничение по числу специалистов различного профиля, занятых эксплуатацией САК:

$$\sum_{k=1}^{\bar{k}} N_{\mu_k} x_k \leqslant N_\mu, \quad \mu = 1, \ldots, \bar{\mu}, \tag{52}$$

где N_{μ_k} — число специалистов μ-го профиля, занятых эксплуатацией k-го элемента САК; N_μ — допустимое число специалистов μ-го профиля.

Ограничения по производственным возможностям промышленных предприятий по поставке серийных технических средств:

$$\sum_{k=1}^{\bar{k}} m_{\omega_k} x_k \leqslant m_\omega, \quad \omega = 1, \ldots, \bar{\omega}, \tag{53}$$

где m_{ω_k} — число серийных технических средств (комплектующих изделий) ω-го типа для k-го элемента САК; m_ω — число серийных технических средств ω-го типа, которое может выпустить промышленность за требуемый период поставки этих средств для внедрения САК с учетом уже имеющихся заказов на поставку аналогичных изделий по другим работам.

Ограничения по трудовым ресурсам разрабатывающих организаций.

$$\sum_{p=1}^{\bar{p}} \alpha_{\psi_p} l_k \Big(\sum_{k \in K_p} \min\{x_k, 1\} \Big) + \sum_{k \in K_0} \alpha_{\psi_k} \min\{x_k, 1\} \leqslant \alpha_\psi, \quad \psi = 1, \ldots, \bar{\psi}, \tag{54}$$

где K_p $(p = 1, ..., \bar{p})$ — подмножества типов КТС, на которые распространяются требования по унификации; K_0 — подмножество типов КТС, на которые не распространяется требование унификации; $l_k(x_k)$ — коэффициент, характеризующий уменьшение трудоемкостей при унификации разработки КТС и представляющей собой слабо возрастающую функцию от суммы

$$\sum_{k \in K_p} \min\{x_k, 1\}$$

в области

$$\sum_{k \in K_v} \min\{x_k,\ 1\} > 1$$

и l_k (0), l_k (1) = 1; α_{ψ_k} α_{ψ_k} — число специалистов ψ-го профиля, которых необходимо привлечь для разработки или доработки k-го элемента САК,

$$\alpha_{\psi_p} = \max_{k \in K_p} \{\alpha_{\psi_k}\};$$

α_ψ— число свободных специалистов ψ–го профиля в разрабатывающих организациях.

Ограничение по срокам разработки и внедрения САК. Для рассматриваемого метода выбора структуры САК это ограничение зависит от способов организации внедрения, принятого плана работ по внедрению и в общем случае выражается нелинейными функциями от переменных. Запишем это ограничение для самого простого случая, когда внедрение САК осуществляется последовательно:

$$\sum_{k=1}^{k} \tilde{t}_k x_k \leqslant T_0, \tag{55}$$

где $\tilde{t}_k$— время, необходимое для внедрения комплекса (устройства) k-го вида.

Ограничения, связанные с выполнением требований по унификации КТС. При выборе структур комбинированным методом требования унификации учитываются на уровне формирования вариантов САК. В излагаемом здесь методе выбора структур рассматриваются множества I_{ξ_i} (ξ = 1, ..., $\overline{\xi}$) КТС, которые различаются как типом выполнения, так и функциональным назначением, так что в ходе решения задачи в структуру САК могут быть включены КТС различных типов одного функционального назначения. В ряде случаев это может быть неприемлемым, и поэтому требуется ввести специальное условие, исключающее применение разнотипных средств.

Пусть K_p, (p = 1, ..., $\overline{p}$) — подмножества типов КТС, на которые распространяются требования по унификации. На каждое подмножество K_p типов КТС распространяется унификация p-го вида. Унификация p-го вида означает, что в K_p-группе КТС к реализации может быть принят только один тип КТС. Указанные ограничения записываются в виде равенств нулю попарных произведений x' и x'' для любых пар $(k', k'') \in K_p$, т. е.

$$x_{k'} x_{k''} = 0; \quad k' < k''; \quad (k',\ k'') \in K_p; \quad p = 1,\ \ldots,\ \bar{p}. \tag{56}$$

Таким образом, сформулированная задача выбора варианта структурной схемы САК состоит в нахождении максимального значения функционала (39) при ограничениях (41)—(47), (49), (51), (55), (56) на переменные задачи. Решение такой экстремальной задачи обычно проводится известными численными методами. При усреднении случайного вектора *у,* т. е. при детерминированном подходе задача сводится к нахождению максимума функционала (40) при ограничениях (41)—(46), (50)—(56). Функционалы (39), (40) являются выпуклыми от переменных, а ограничения (41)—(55), за исключением ограничений (47)—(51), (56), представляют собой линейные неравенства.

Сформулированная задача приближенно может быть решена так называемым ***методом максимального элемента***. Изложим суть этого метода применительно к максимизации функционала (40).

Отыскание максимума производится за ряд последовательных шагов. Рассмотрим произвольный t-й шаг. К началу t-го шага известны векторы

$$\|x_k^{(t-1)}\|, \quad \|r_k^{(t-1)}\| = \left\| \frac{\partial F\,(X)^{(t-1)}}{\partial X_k} \right\|, \quad \|M_\xi^{(t-1)}\|,$$
$$\|N_\mu^{(t-1)}\|, \ \|m_\omega^{(t-1)}\|, \ \|\alpha_\psi^{(t-1)}\|$$

и величины

$$C_0^{(t-1)}, \ \overline{C}^{(t-1)}, \ \widehat{C}_0^{(t-1)}, \ C_э^{(t-1)}, \ C^{(t-1)};$$
$$T_0^{(A,\ t-1)}, \ T^{(0,\ t-1)}, \ N^{(t-1)}, \ T_0^{(t-1)}.$$

Для 1-го шага

$$x_k^{(0)} = 0; \quad r_k^{(0)} = u_0 T\beta_k - C_k; \quad M_\xi^{(0)} = M_\xi;$$
$$C_0^{(0)} = C_0; \quad \overline{C}_0^{(0)} = \overline{C}_0; \quad \widehat{C}_0^{(0)} = \widehat{C}_0; \quad C_э^{(0)} = C_э; \quad C^{(0)} = C;$$
$$T^{(A,\ 0)} = T^{(A)}; \quad T^{(0,\ 0)} = T^{(0)}; \quad N^{(0)} = N; \quad N_\mu^{(0)} = N_\mu;$$
$$m_\omega^{(0)} = m_\omega; \quad \alpha_\varphi^{(0)} = \alpha_\psi; \quad T_0^{(0)} = T_0.$$

Находим

$$r_k^t = \max_k r_k^{(t-1)}.$$

Определяем новые значения переменных

$$x_k^{(t)} = \begin{cases} X_k^{(t-1)}+1 & \text{при } k=k_t \notin K_p, \text{ или } k=k_t \in K_p, \\ K_p \cap \{k_1, \dots, k_{t-1}\} \neq 0, \ k_t \in \{k_1, \dots, k_{t-1}\}, & \text{или} \\ k=k_t \in K_p, \quad K_p \cap \{k_1, \dots, k_{t-1}\} = 0; \\ X_k^{(t-1)} & \text{при } k \neq k_t \text{ или } k=k_t \in K_p, \\ K_p \cap \{k_1, \dots, k_{t-1}\} \neq 0, \quad k_t \in \{k_t, \dots, k_{t-1}\}. \end{cases}$$

Если

$$k = k_t \in K_p, \quad K_p \cap \{k_1, \dots, k_{t-1}\} \neq 0,$$

$k_t \notin \{k_1,\dots,k_{t-1}\}$ то полагаем $r_{k_t}^{(t-1)}$ -й элемент вектора $\|r_k^{(t-1)}\|$ равным нулю, получая таким образом вектор $\| r_k^{(t)} \|$, и переходим к следующему шагу. Затем определяем

$$r_k^{(t)} = \begin{cases} r_k^{(t-1)} e^{\beta_k} - C_k(1-e^{-\beta_k}) & \text{при } k=k_t; \\ r_k^{(t-1)} & \text{при } k \neq k_t; \end{cases}$$

$$M_\xi^{(t)} = \begin{cases} M_\xi^{(t-1)} - 1 & \text{при } k_t \in J_\xi; \\ M_\xi^{(t-1)} & \text{при } k_t \notin J_\xi; \end{cases}$$

$$C_0^{(t)} = C_0^{(t-1)} - C_{0k_t},$$

$$\overline{C}_0^{(t)} = \overline{C}_0^{(t-1)} - \overline{C}_{0k_t},$$

$$\widehat{C}_0^{(t)} = \widehat{C}_0^{(t-1)} - \widehat{C}_{0k_t}, \quad C_э^{(t)} = C_э^{(t-1)} - C_{э k_t},$$

$$C^{(t)} = C^{(t-1)} - C_{k_t}, \quad T_0^{(t)} = T_0^{(t-1)} - t_{k_t};$$

$T^{(A,t)} = T^{(A)}$ при выполнении соотношения (44) и $T^{(A,t)} = 0$ в противном случае; $T^{(0,t)} = T^{(0)}$ при выполнении соотношения (46) и $T^{(0,t)} = 0$ в противном случае.

Процесс поиска максимального значения $F(x)$ заканчивается, если вектор $\| r_k^{(t)} \| = 0$. Найденные значения вектора X^* представляют собой оптимальный с точки зрения принятого критерия выбора варианта структурной схемы САК.

Рассмотрим пример решения задачи выбора структуры ***методом вариации состава технических средств САК***. Предположим, что нужно определить рациональную структуру подсистемы САК при следующих исходных данных.

В составе подсистемы САК рассматриваются КТС трех функциональных назначений: КТС-1, $\xi = 1$; КТС-2, $\xi = 2$; КТС-3, $\xi = 3$. Выбор рациональной структуры осуществляется на базе восьми типов КТС ($\bar{k} = 8$), при этом подмножество КТС-1 включает $J_1=\{k = 1, 2, 3\}$; КТС-2 $J_2=\{k = 4, 5, 6\}$; КТС-3 $J_3=\{k = 7, 8\}$. Стоимость КТС характеризуется вектором $C = \| C_k \| = \| 5, 8, 15, 5, 8, 10, 2, 1{,}5\|$ (C_k измеряется в миллионах рублей). Ограничение по суммарной стоимости $C = 60$ млн. руб. Коэффициенты β_k заданы вектором
$\beta =\| \beta_k\| = \| 0{,}1,\ 0{,}2,\ 0{,}3,\ 0{,}2,\ 0{,}5,\ 0{,}1,\ 0{,}05 \|$. Численность состава обслуживающего персонала КТС каждого типа
$\| N_k \| = \| 20, 22, 25, 20, 22, 22, 2, 2 \|$, а ограничение по общей численности обслуживающего персонала $N = 115$ чел. Число специалистов высшей квалификации для эксплуатации КТС каждого типа ($\mu = 1$)
$\| N_{1k} \| = \| 7, 10, 10, 7, 8, 8, 1, 1, \|$, а ограничение по общему числу специалистов высшей квалификации $N_1 = 43$ чел. Число специалистов средней квалификации для обслуживания КТС каждого типа ($\mu = 2$)
$\| N_{2k} \| = \| 13, 12, 15, 13, 14, 14, 1, 1 \|$, а ограничение по общему числу специалистов средней квалификации $N_2 = 77$ чел. Предельный среднегодовой экономический доход от внедрения подсистемы САК
$u_0 = 15$ млн. руб., срок службы $T = 10$ лет, допустимый период абсолютной окупаемости $T^{(A)} =4{,}1$ года.

Максимальное число КТС первого функционального назначения $M_1=1$, второго — $M_2=3$, третьего — $M_3=9$. Унификация распространяется на два подмножества КТС: $k_1 = \{k = 4, 5, 6\}$,
$k_2 =\{k =7, 8\}$.

Математически сформулированная задача запишется в следующем виде.

Критерии выбора:

$$F(X) = 150\,[1 - \exp \}- (0{,}1x_1 + 0{,}2x_2 + 0{,}3x_3 + 0{,}2x_4 + 0{,}5x_5 + \\ + 0{,}5x_6 + 0{,}1x_7 + 0{,}05x_8)\}] - (5x_1 + \bar{8}x_2 + 15x_3 + 5x_4 + 8x_5 + \\ + 10x_6 + 2x_7 + 1{,}5x_8).$$

Ограничения:

$$\sum_{k \in J_1} X_k = x_1 + x_2 + x_3 \leqslant 1 \ (\xi = 1,\ J_1 = \{1,\ 2,\ 3\});$$

$$\sum_{k \in J_2} X_k = x_4 + x_5 + x_6 \leqslant 3\,(\xi = 2,\ J_2 = \{4,\ 5,\ 6\});$$

$$\sum_{k \in J_3} X_k = x_7 + x_8 \leqslant 9\,(\xi = 3,\ J_3 = \{7,\ 8\};$$

$$x_4 x_5 = 0, \quad x_4 x_6 = 0,$$

$$x_5 x_6 = 0 \quad (p = 1,\ k_1 = \{4,\ 5,\ 6\});$$

$$x_7 x_8 = 0 \quad (p = 2,\ k_2 = \{7,\ 8\});$$

$$\sum_{k=1}^{8} C_k X_k = 5x_1 + 8x_2 + 15x_3 + 5x_4 + 8x_5 + 10x_6 + 2x_7 + 1{,}5x_8 \leqslant 60;$$

$$[5x_1 + 8x_2 + 15x_3 + 5x_4 + 8x_5 + 10x_6 + 2x_7 + 1{,}5x_8] \cdot 15^{-1} \times$$
$$\times\,[1 - \exp\{-(0{,}1x_1 + 0{,}2x_2 + 0{,}3x_3 + 0{,}2x_4 + 0{,}3x_5 +$$
$$+\,0{,}5x_6 + 0{,}1x_7 + 0{,}05x_8\}]^{-1} \leqslant 41;$$

$$\sum_{k=1}^{8} N_k X_k = 20x_1 + 22x_2 + 25x_3 + 20x_4 + 22x_5 + 22x_6 +$$
$$+\,2x_7 + 2x_8 \leqslant 115;$$

$$\sum_{k=1}^{8} N_{1k} X_k = 7x_1 + 10x_2 + 10x_3 + 7x_4 + 8x_5 + 8x_6 +$$
$$+\,x_7 + x_8 \leqslant 43\,(\mu = 1);$$

$$\sum_{k=1}^{8} N_{2k} X_k = 13x_1 + 12x_2 + 15x_3 + 13x_4 + 14x_5 +$$
$$14x_6 + x_7 + x_8 \leqslant 77.$$

Первый шаг:

$$\| r_k^{(0)} \| = \| 10;\ 22;\ 30;\ 25;\ 37;\ 65;\ 13;\ 6 \|;$$

$$M_1^{(0)} = 1;\ M_2^{(0)} = 3;\ M_3^{(0)} = 9;\ C^{(0)} = 60;\ T^{(A,\,0)} = 0;\ N^{(0)} = 115;$$

$$N_1^{(0)} = 43;\ N_2^{(0)} = 77;\ r_{k1}^{(0)} = r_6^{(0)} = \max_k r_k^{(0)};\ k_1 = 6;$$

$$x_k^{(1)} = \begin{cases} 0 + 1 \text{ для } k = k_1 = 6 \in k_1,\ k_1 \cap \{0\} = 0; \\ 0 \text{ для } k \neq 6; \end{cases}$$

$$X^{(1)} = \| 0;\ 0;\ 0;\ 0;\ 0;\ 1;\ 0;\ 0 \|;$$

$$\tilde{r}_k^{(1)} = \begin{cases} 67e^{-0,5} - 4 = 35 \text{ при } k = 6; \\ r_k^{(0)} \text{ при } k \neq 6; \end{cases}$$

$$\| \tilde{r}_k^{(1)} \| = \| 10;\ 22;\ 30;\ 25;\ 37;\ 35;\ 13;\ 6 \|;$$

$$M_\xi^{(1)} = \begin{cases} M_\xi^{(0)} - 1 = 3 - 1 = 2 \text{ при } \xi = 1; \\ M_\xi^{(0)} \text{ при } \xi \neq 1; \end{cases}$$

$$\| M_\xi^{(1)} \| = \| 1;\ 2;\ 9 \|;\ C^{(1)} = C^{(0)} - C_{k1} = 60 - 10 = 50;$$

$$\frac{\sum\limits_k C_k X_k}{u_0\left[1-\exp\left(-\sum\limits_k \beta_k x_k\right)\right]}=\frac{10}{15\,(1-0{,}6)}=1{,}67\leqslant 4{,}1;$$

$$N^{(1)}=N^{(0)}-N_{k1}=N^{(0)}-N_6=115-22=93;$$

$$N_1^{(1)}=N^{(0)}-N_{1k_1}=43-8=35;$$

$$N_2^{(1)}=N_2^{(0)}-N_{2k_1}=77-14=63.$$

Второй шаг:

$$r_{k_2}^{(1)}=r_5^{(1)}=\max_k r_k^{(1)},\;\; k=5;$$

$$x_k^{(2)}=X_k^{(1)},\; \text{так как }\; k_2\in k_1,\;\; k_1\cap\{k_1\}\neq 0,\;\; k_2\notin\{k_1\};$$

$$\|r_k^{(2)}\|=\|10;\;22;\;30;\;25;\;0;\;35;\;13;\;6\|.$$

Третий шаг:

$$r_{k_3}^{(2)}=r_6^{(2)}=\max r_k^{(2)},\;\; k_3=6;$$

$$x_k^{(3)}=\begin{cases} x_k^{(2)}+1=2\;\text{ где }\; k=k_3=6,\;\; k_3\in k_1, \\ \quad k_1\cap\{k_1,\;k_2\}\neq,\;\; k_3\in\{k_1,\;k_2\}; \\ \quad\quad x_k^{(2)}\;\text{ для }\; k\neq 6; \end{cases}$$

$$x^{(3)}=\|0;\;0;\;0;\;0;\;0;\;2;\;0;\;0\|;$$

$$\tilde r_k^{(3)}=\begin{cases} 35\mathrm{e}^{-0,5}-15=2-1-4\;\text{ для }\; k=k_3=6; \\ \quad r_k^{(2)}\;\text{ для }\; k\neq 6; \end{cases}$$

$$\|\tilde r^{(3)}\|=\|10;\;22;\;30;\;25;\;0;\;17;\;13;\;6\|;$$

$$M_\xi^{(3)}=\begin{cases} M_\xi^{(2)}-1=2-1=1\;\text{ при }\;\xi=1; \\ \quad M_\xi^{(2)}\;\text{ при }\;\xi\neq 1; \end{cases}$$

$$\|M_\xi^{(3)}\|=\|1;\;1;\;9\|;\;\; C^{(3)}=C^{(2)}-C_{k3}=50-10=40;$$

$$20/15\,(1-0{,}45)=2{,}42\leqslant 4{,}1;\;\; T^{(A,\,3)}=4{,}1;$$

$$N^{(3)}=N^{(2)}-N_{k3}=93-22=71;\;\; N_1^{(3)}=N_1^{(2)}-N_{1k3}=35-8= \\ =27;$$

$$N_2^{(3)}=N_2^{(2)\prime}-N_{2k3}=63-14=49;$$

$$\|r_k^{(3)}\|=\|10;\;22;\;3;\;25;\;0;\;17;\;13;\;6\|.$$

На 12-м шаге процесс заканчивается, так как наступает ограничение по абсолютной окупаемости подсистемы САК. Получаем решение

$x^{(12)} = x^* = \| 0, 0, 1, 0, 0, 3, 6, 0\|$, из которого следует, что структура подсистемы САК включает в себя: один КТС, используемый в САК-1 третьего типа $\{k = 6)$; шесть КТС, используемых в САК-3 седьмого типа $(k = 7)$.

В заключение отметим, что, в отличие от комбинированного метода, где все особенности структуры органически содержатся в каждом оцениваемом варианте, в рассмотренном методе особенности структуры формируются не только в ходе решения задачи выбора, но и на этапе формирования исходных данных для задачи, а также после получения решения, когда устанавливается схема функциональной взаимосвязи, исходя из совокупности КТС данной САК.

4.5. Методы и алгоритмы оптимизации структур

Из множества формализуемых задач структурного синтеза значительная их часть может быть сведена к определению экстремального значения целевой функции

$$F(\mathrm{X}) = \sum_{j=1}^{m} c_j x_j \to \max(\min) \tag{57}$$

при ограничениях

$$\sum_{j=1}^{m} a_{ij} x_j \leqslant b_i \left(\sum_{j=1}^{m} a_{ij} x_j \geqslant b_i\right), \quad i = \overline{1,n}, \tag{58}$$

$$x_j \geqslant 0, \quad j = \overline{1,m}; \tag{59}$$

x_j — целые числа, $j = 1, 2, \ldots, p$ $(p \leq m)$. (60)

Если в сформулированной задаче ограничения (60) отсутствуют, то имеет место классическая задача линейного программирования, если ограничения (60) имеются и $p=m$, то данная задача является полностью целочисленной, при $p<m$ задача является частично целочисленной.

Задача линейного программирования. В настоящее время теория линейного программирования хорошо разработана и имеется целый арсенал методов решения задач линейного программирования — это, например симплекс-метод, реализующий последовательную процедуру направленного поиска оптимального значения целевой функции (57) при существующих ограничениях вида (58) и (59).

Симплекс-метод. Первоначально система неравенств (58) путем введения дополнительных переменных $x_{m+i} \geq Q$ преобразуется в систему уравнений таким образом, чтобы имело место одно из двух выражений

$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{im}x_m + x_{m+i} = b_i;$$
$$a_{i1}x_1 + a_{i2}x_2 + \ldots + a_{im}x_m - x_{m+i} = b_i, \quad i = \overline{1,n}.$$

Такое изменение приводит просто к увеличению числа переменных, что не меняет существа задачи. Введем ряд понятий, широко применяемых в задачах линейного программирования.

Базисом называют любой набор из *n* таких переменных, что определитель, составленный из коэффициентов при этих переменных, не равен нулю. Отметим, что остальные *т—п* переменных называют *свободными.*

Не уменьшая общности, рассмотрим сущность симплекс-метода на примере задачи максимизации целевой функции (57) при наличии ограничений (58) и (59). Для определения первоначального базисного решения какие-либо *т—п* переменные принимают за свободные, т.е. приравнивают нулю, при этом все базисные переменные выражают через свободные, после чего решают систему полученных уравнений. Если некоторые из базисных переменных окажутся отрицательными, то полученное базисное решение является *недопустимым* и производится переход к новому базису путем выбора новой совокупности свободных переменных. Базисное решение, в котором отсутствуют отрицательные переменные, называют *допустимым.*

После того как найдено допустимое базисное решение, проверяют, не достигнут ли максимум целевой функции $F(\mathbf{X})$. Если нет, то ищут новое допустимое базисное решение, но не любое, а такое, которое увеличивает значение целевой функции $F(\mathbf{X})$. Затем процедуру повторяют. Данный метод довольно быстро приводит к цели, так как позволяет исключить из рассмотрения большое число базисных решений, заведомо не обращающих в максимум целевую функцию $F(\mathbf{X})$.

Проверку того, не достигнут ли при найденном решении максимум целевой функции, можно сделать путем поиска нового базисного решения, при котором значение целевой функции $F(\mathbf{X})$ будет больше предыдущего. Для прихода к новому допустимому базисному решению одну из свободных переменных следует сделать базисной, при этом она будет отличной от нуля, т. е. возрастет. Следовательно, если какая-либо из свободных переменных входит в выражение для целевой функции со знаком « + », а значит, при ее увеличении целевая функция увеличивается, то максимум целевой функции не достигнут и данную свободную переменную следует перевести в базисную.

Однако при возрастании свободной переменной некоторые из базисных переменных начнут уменьшаться. Так как отрицательные значения переменных недопустимы, в качестве новой свободной переменной следует взять ту из базисных, которая раньше других обратится в нуль.

Задачи целочисленного программирования. В общем случае условие целочисленности накладывает дополнительные ограничения, вследствие которых максимальное значение целевой функции (в задачах максимизации) оказывается, как правило, меньше максимального значения целевой функции соответствующей задачи линейного программирования; в последней отсутствуют условия целочисленности переменных.

Комбинаторные методы. Методы решения задач целочисленного линейного программирования в классе комбинаторных методов базируются на максимальном учете характера задач и конечности множества вариантов их решения. В их вычислительных схемах используется идея частичного перебора вариантов решения задач. Это достигается путем отбрасывания некоторых подмножеств вариантов, которые согласно известным свойствам оптимального решения заведомо могут считаться неперспективными.

Отличительной особенностью комбинаторных алгоритмов является и то, что во многих из них вообще не используется решение задач линейного программирования, соответствующей исходной целочисленной задаче линейного программирования.

Таким образом, для большинства комбинаторных методов не требуется специальных доказательств их конечности, что и способствовало широкому применению методов этого типа на практике. Большая группа комбинаторных методов базируется на достаточно общей схеме методов, которые объединены под названием — методы «ветвей и границ». Схема метода «ветвей и границ» широко используется на практике.

Приближенные методы. Одной из основных причин бурного развития этой группы методом следует назвать сравнительную сложность реализации точных методов, а зачастую и просто невозможность применения последних для задач большой размерности и для задач, время решения которых ограничено. Другим фактором, способствующим на практике развитию приближенных методов, является то, что для многих практических задач в значительной мере точные решения не обеспечиваются из-за малой достоверности исходных данных.

Среди приближенных методов решения целочисленных задач линейного программирования следует выделить применение различного рода эвристических алгоритмов, основанных на:

а) методах, построенных на использовании случайного поиска;

б) методах, сочетающих случайный поиск с идеей локальной оптимизации;

в) методах, вычислительные схемы которых строятся на максимальном учете специфики конкретных типов задач и др.

Рассмотрим два основных подхода к отысканию точного оптимального решения задач целочисленного программирования, базирующихся на методах отсекающих плоскостей и методах возврата.

Методы отсекающих плоскостей (методы отсечения). Исходным моментом решения задачи целочисленного программирования является оптимальное решение соответствующей задачи линейного программирования, полученной после отбрасывания условий целочисленности. На каждой итерации добавляется линейное ограничение, удовлетворяющее целочисленному решению исходной задачи, но исключающее текущее нецелочисленное решение. Вычислительный процесс прекращается, как только будет достигнуто любое целочисленное решение. Сходимость обеспечивается за конечное, но иногда очень большое число итераций.

Сущность алгоритмов, основанных на методе отсечения, легко уяснить, обратившись к геометрическим представлениям в пространстве решений. Определим выпуклую оболочку множества допустимых целочисленных точек (решений) как минимальное выпуклое множество, содержащее все эти точки. Допустимыми решениями будет не вся область допустимых решений, находящаяся внутри и на границе выпуклой оболочки, а лишь отдельные дискретные точки этой области, имеющие все целочисленные координаты. Целевая функция достигает оптимального значения в одной из вершин этой выпуклой оболочки, которая представляет собой одно из допустимых целочисленных решений.

Выпуклую оболочку можно представить конечным множеством линейных ограничений (58), как изображено на рис. 1.

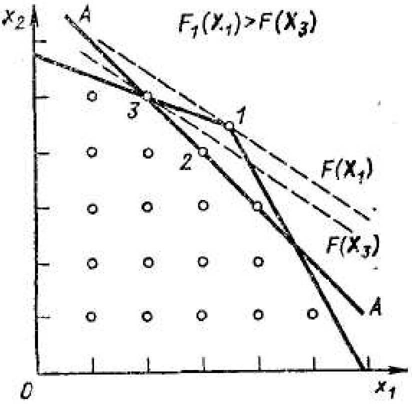


Рис. 1. Геометрическая иллюстрация принципа отсечения

Можно, не считаясь с условиями целочисленности, найти решение, определяемое точкой *1*, а затем, округлив это решение до ближайших целых чисел, получить целочисленное решение в точке *2*. Однако при этом может получиться решение, далекое от оптимального. Оптимальным целочисленным решением будет точка *3*.

Для определения оптимального решения в алгоритмах отсечения вначале рассмотрим выпуклую оболочку, определенную линейными ограничениями (58) и условиями неотрицательности переменных исходной задачи, и отыщем экстремальную точку этой оболочки (точка *1* на рис. 1). Если такое решение оказывается нецелочислеппым, то добавляют ограничение, отсекающее текущую экстремальную точку и уменьшающее «объем» выпуклой оболочки (прямая *А—А)*. Однако новое ограничение не отсекает ни одной экстремальной точки выпуклой оболочки, принадлежащей допустимым целочисленным решениям. В конечном итоге вводится такое число дополнительных ограничений, что экстремальная точка усеченной выпуклой оболочки представляет собой целочисленное решение исходной задачи.

Покажем, как формально построить дополнительные линейные ограничения, которым должно удовлетворить любое решение задачи (57) — (60). Рассмотрим ограничение

$$\sum_{j=1}^{m} a_{ij} x_j = b_i. \tag{61}$$

Возможно, одна или несколько величин a_{ij} и b_i являются дробными. Обозначим символом $[a_{ij}]$ целую часть a_{ij}, т. е. наибольшее целое число, меньшее действительного числа a_{ij} или равное ему. Поскольку на величины x_j наложено ограничение (59), любые значения x_j, соответствующие условию (61), должны удовлетворять более слабому ограничению

$$\sum_{j=1}^{m} [a_{ij}]\, x_j \leqslant b_i. \tag{62}$$

Поскольку сумма в левой части неравенства (62) должна быть целочисленной, (62) можно усилить следующим образом:

$$\sum_{j=1}^{m} [a_{ij}]\, x_j \leqslant [b_i]. \tag{63}$$

Наконец, неравенство (63) можно преобразовать в равенство, добавив дополнительную переменную x_{m+i}:

$$\sum_{j=1}^{m} [a_{ij}]\, x_j + x_{m+i} = [b_i], \tag{64}$$

где x_{m+i} — целое, положительное.

Таким образом, если в исходную задачу добавить линейное ограничение (64), задача все равно останется полностью целочисленной.

Уравнение отсекающей плоскости получается в результате вычитания из (64) уравнения (61):

$$\sum_{j=1}^{m} - f(a_{ij})\, x_j + x_{m+i} = - f(b_i), \tag{65}$$

где $f(a_{ij})$, $f(b_i)$ —дробная часть чисел a_{ij} и b_i.

Заметим, что текущее решение задачи линейного программирования не удовлетворяет ограничению (65), поскольку значение $x_{m+i} = -f(b_i)$ строго отрицательно.

Алгоритм отсечения (алгоритм Гомори) состоит из следующих шагов.

1. Найти оптимальное решение задачи линейного программирования (57) — (59) без условия целочисленности (60).
2. Прекратить вычисления, если текущее решение задачи является целочисленным. В противном случае выбрать какую-либо дробную базисную переменную. Составить ограничение (65) из уравнения, содержащего эту базисную переменную в текущем оптимальном решении задачи линейного программирования.

3. Добавить к исходной задаче линейного программирования новое ограничение (65), найти оптимальное решение задачи с дополнительным ограничением и вернуться к шагу 2.

Введение на шаге 3 отсекающего ограничения (65) наряду с условием x_{m+i}>0 делает текущее решение задачи линейного программирования недопустимым. Отсечение текущего оптимального решения означает, что на шаге 3 значение целевой функции в задаче максимизации не может возрасти, а может лишь уменьшиться. Отметим, что алгоритмы отсечения не гарантируют получения допустимого целочисленного решения до самой последней итерации. Вследствие этого при преждевременном прекращении вычислений можно допустимое решение и не получить.

Методы возврата. В этой группе методов имеются различные модификации. Наиболее распространенным среди них является ***метод ветвей и границ***, который предназначен для решения частично целочисленных задач. Как и в методе отсечения, решение задачи начинается с отыскания оптимального решения задачи линейного программирования без учета условия целочисленности. Затем формируется семейство связанных, но различных задач линейного программирования. Термин «возврат» определяет специфический способ формирования и решения последовательности задач.

Рассмотрим задачу максимизации (57) — (60). Допустим, что для каждой целочисленной переменной можно задать верхнюю и нижнюю границы, в пределах которых безусловно содержатся ее оптимальные значения

$$L_j \leqslant x_j \leqslant U_j, \quad j=\overline{1,p}. \tag{66}$$

Идея метода ветвей и границ основана на следующем элементарном факте. Рассмотрим любую переменную x_j и примем, что R есть некоторое целое число, где $L_j \leq R \leq U_{j-1}$. Тогда оптимальное решение задачи будет удовлетворять одному из ограничений

$$x_j \geq R+1 \text{ или } x_j \leq R. \tag{67}$$

К примеру, если при решении задачи линейного программирования получено x_i = 2,6, то можно поставить и решить две задачи линейного программирования, причем в одну из них вводится согласно (67) условие $3 \leq x_1 \leq U_1$, а к другую — условие $L_1 \leq x_1 \leq 2$. Предположим, что каждая из этих задач имеет оптимальное решение, удовлетворяющее условию целочисленности (60). Тогда решение, доставляющее большее значение целевой функции, является оптимальным решением исходной целочисленной задачи.

Метод ветвей и границ основан на решении некоторого множества задач линейного программирования. Границы (66) на каждую переменную x_j служат для конкретизации диапазона изменений переменных в задаче целочисленного программирования, что в свою очередь определяет мощность множества задач линейного программирования, используемых в процедуре реализации метода ветвей к границ. Поэтому трудоемкость вычислений определяется числом целочисленных переменных, содержащихся в задаче.

Алгоритм решения задачи целочисленного программирования методом ветвей и границ заключается в следующем. На каждой итерации (обозначим номер итерации через *t)* имеются нижняя оценка $F^*_t(\mathbf{X})$ оптимального значения целевой функции и список задач линейного программирования, подлежащих решению. Процедура решения состоит в последовательном улучшении оценки $F^*(\mathbf{X})$ и приближении ее к оптимальному значению $F_{\text{opt}}(\mathbf{X})$.

На итерации 1 список задач содержит одну задачу (57) — (59). На итерации t из списка выбирают и решают задачу линейного программирования. Если она не имеет допустимого решения или если полученное оптимальное значение целевой функции $F_{t\text{ opt}}(\mathbf{X}) \leq F_t^*(\mathbf{X})$, то нижняя оценка остается прежней и из списка выбирают очередную задачу для решения. Если полученное решение удовлетворяет условию целочисленности (60) и $F_{t\text{ opt}}(\mathbf{X}) > F_t^*(\mathbf{X})$, то полученное оптимальное решение $F_{t\text{ opt}}(\mathbf{X})$ на итерации t принимают в качестве нижней оценки для последующих итераций. Если полученное оптимальное решение задачи линейного программирования не удовлетворяет условиям целочисленности (60), то выбирают нецелочисленную переменную x_j и решаемую задачу разбивают на две новые задачи линейного программирования путем введения в каждую из них по одному ограничению (67).

При остановке алгоритма в случае, если допустимому решению соответствует значение целевой функции $F_{t\text{ opt}}(\mathbf{X}) = F_t^*(\mathbf{X})$, полученное решение оптимально, в противном случае допустимого решения не существует. Метод ветвей и границ особенно эффективен для решения комбинаторных задач, в частности задачи коммивояжера.

В общем виде процедура реализации этого метода заключается в следующем (рассмотрим ее применительно к задаче минимизации). Предположим, что имеется возможность получить нижнюю оценку качества решения $F^*_V(\boldsymbol{X})$ как для всего множества V возможных решений, так и для его различных подмножеств. Разобьем множество на два непересекающихся подмножества A и B с точными нижними границами критерия качества $F_{A\text{ opt}}(\mathbf{X})$ и $F_{B\text{ opt}}(\mathbf{X})$, связанных с $F^*(\mathbf{X})$ очевидным соотношением

$$F_{\text{opt (}}(\mathbf{X}) = \min\{F_{A\text{ opt}}(\mathbf{X}), F_{B\text{ opt}}(\mathbf{X})\}.$$

Поскольку множества A и B имеют меньшее число элементов, чем множество V, т.е. возможность получить нижние оценки качества $F^*_A(\mathbf{X})$ и F^*_B $(\mathbf{X})$ более близкие к $F_{A\text{opt}}(\mathbf{X})$ и $F_{B\text{opt}}(\mathbf{X})$, чем в первоначальном множестве V, это означает, что оценка
$\min\{F^*_A(\mathbf{X}), F^*_B(\mathbf{X})\}$ будет более близка к $F_{\text{opt}}(\mathbf{X})$, чем оценка $F^*_V(\mathbf{X})$.

Можно предположить, что оптимальное решение будет с большой вероятностью принадлежать тому из подмножеств A и B, которое имеет меньшую нижнюю оценку. Если оказалось, что $F^*_B(\mathbf{X})<F^*_A(\mathbf{X})$, то есть все основания для детального исследования подмножества B. Последнее разбивают на два подмножества C и D с соответствующими нижними оценками F^*_C $(\mathbf{X})$ и F^*_D $(\mathbf{X})$, и если оказалось, что $F^*_D(\mathbf{X})<F^*_C(\mathbf{X})$, то подобному разбиению подвергается подмножество D. Эту процедуру продолжают до тех пор, пока не придут к подмножеству, состоящему из одного элемента с оценкой $F^*(\mathbf{X})=F_0(\mathbf{X})$. Процедура разбиения обычно представляется в виде дерева решений (рис. 2).

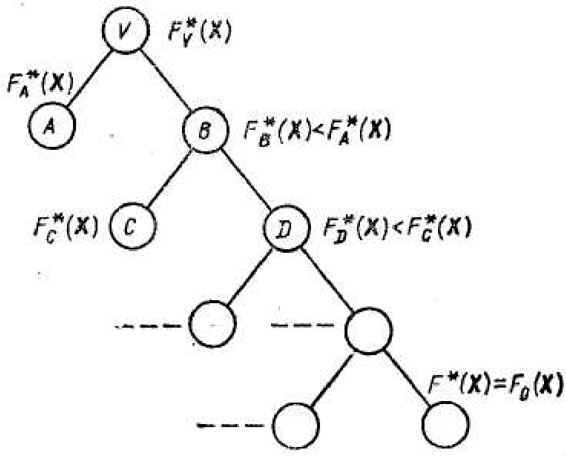


Рис. 2. Геометрическая иллюстрация метода ветвей и границ

Найденное решение необходимо проверить на оптимальность. Для этого проверяют, нет ли среди нерассмотренных множеств A, C и т. д. элемента со значением нижней оценки $F^*(\mathbf{X})<F_0(\mathbf{X})$. При этом множества, у которых $F^*(\mathbf{X})\geq F_0(\mathbf{X})$, дальше не рассматриваются.

Разбиение перспективных множеств либо приведет к решению с оценкой $F^*(\mathbf{X})<F_0(\mathbf{X})$, которое должно быть принято за оптимальное, либо позволит убедиться в оптимальности полученного ранее решения

с оценкой $F_0(\mathbf{X})$. Из процедуры реализации метода ветвей и границ становится ясным, почему этот метод получил название метода возврата.

Метод ветвей и границ наряду с методами отсечения обладает существенными достоинствами с вычислительной точки зрения. Алгоритмы, построенные на этих методах, сравнительно легко программируются на ЭВМ и реализуются на любой итерации без вмешательства человека, однако их эффективность резко снижается при увеличении размерности решаемой задачи.

Ниже приводятся примеры постановки типовых задач структурного синтеза в терминах математического программирования.

Пример. ***Формирование рекомендаций по решению задачи оптимизация структуры сети электросвязи***. Процесс консультирования по решению задачи проектирования региональных сетей электросвязи состоит из ряда взаимосвязанных этапов:

- построения структуры первичной сети связи, обеспечивающей связь между заданными пунктами региона по критерию минимальной стоимости;

- корректировки полученной структуры сети путем введения дополнительных связей между пунктами с целью получения живучей сети электросвязи;

- построения вторичной сети связи, заключающейся в определении конфигурации сетки пучков каналов электросвязи для обеспечения необходимой пропускной способности каналов между различными пунктами региона.

Формально консультационную задачу по формированю рекомендаций синтеза структуры первичной сети связи можно представить в виде следующей задачи математического программирования. Задана матрица расстояний $D=|d_{ij}|$ размерности $n \times n$ между всеми n пунктами данного региона. Необходимо определить такую структуру сети, которая обеспечивала бы связь между всеми пунктами региона по критерию минимальной стоимости. При этом будем считать, что стоимость канала связи между пунктами i и j пропорциональна расстоянию d_{ij} между ними.

Введем псевдобулевы переменные:

$$x_{ij} = \begin{cases} 1, \text{ если имеется прямой канал связи между пунктами } i \text{ и } j; \\ 0 \text{ в противном случае.} \end{cases}$$

Задача синтеза оптимальной структуры сети связей по критерию минимальной стоимости заключается в определении таких переменных $x_{ij} \in \{0, 1\}$, которые обращали бы в минимум целевую функцию

$$F(\mathbf{X}) = \sum_{i=1}^{n} \sum_{j=1}^{n} d_{ij} x_{ij}$$

при ограничениях

$$\sum_{i=1}^{n} x_{ij} \geqslant 1, \sum_{j=1}^{n} x_{ij} \geqslant 1, \sum_{i=1}^{n} \sum_{j=1}^{n} x_{ij} \geqslant n-1.$$

Решение сформулированной задачи математического программирования удобно определять методом ветвей и границ.

Пример. ***Консультационная задача формирования рекомендаций осуществления синтеза структуры памяти специализированной ЭВМ.*** Требуется синтезировать структуру внутренней памяти (СОЗУ—ОЗУ—ПЗУ) специализированной ЭВМ, работающей в режиме реального времени, совместно с определением необходимого пакета объектных программ и их размещением в различных блоках системы памяти.

Пусть ЭВМ ориентирована на решение множества задач (заявок) $C=\{c_1, ..., c_n\}$, причем каждая задача $c_i \in C$ встречается с относительной частотой g_i, и пусть на множестве C задано его разбиение на подмножества C_l, l=l, 2,…,u, $(u \leq n)$, так что в подмножество $C_l \subset C$ входят только те задачи $c \in C_l$, которые имеют l- й приоритет, причем чем меньше l, тем выше приоритет.

Дли решения поступающих в ЭВМ задач имеется множество алгоритмов $\mathcal{L} = \{L_1, ..., L_p), p \leq n$, по отношению к которым множество C разбито на подмножества $C(L_i) \subseteq C$ таким образом, что если $c \in C\ (L_i)$. то задача может быть решена с помощью алгоритма $L_i \in \mathcal{L}$. Каждый из алгоритмов $L_i \in \mathcal{L}$ может быть реализован m_i методами, т. е. для каждой c_l можно записать множество упорядоченных пар $L_l = \{(i, j)\}$, i=1, 2,…, p, j =1, , 2,…, m_l, где любая пара (i, j) описывает метод решения задачи c_l. Каждый из методов $(i, j) \in L_l$ характеризуется: способом хранения программы (в ОЗУ или ПЗУ); числом требующихся ячеек a_{ij} ОЗУ или ПЗУ для записи программ; числом b_{ij} ячеек СОЗУ для хранения часто встречающихся данных и промежуточных результатов и числом h_{ij} ячеек ОЗУ для хранения исходных данных при обработке заявки c_l в реальном времени; количеством $r_{ij}^{(k)}$ ресурса k, где под ресурсами будем понимать габариты, массу, стоимость, число используемых ячеек или панелей, потребляемую мощность и т. д. Поскольку программы могут храниться как в ПЗУ, так и в ОЗУ, это влияет на время реализации программы, надежность и оперативность в смене программ. Часть параметров, например, время решения задачи, стоимость, габариты, потребляемая мощность и др., подчиняется условиям аддитивности, в то время как такие параметры, как надежность, диагностическая возможность, ремонтопригодность и др., являются неаддитивными.

Задача заключается в формировании рекомендаций по определению структуры внутренней памяти и такого набора программ **П** решения поступающих задач, чтобы удовлетворялись все ограничения на параметры ЭВМ, а выбранный критерий оптимальности достигал своего экстремального значения.

В частности, в качестве критерия оптимальности можно выбрать суммарный объем ОЗУ и ПЗУ, который требуется для хранения программ и данных и который следует минимизировать.

Дадим математическое описание задачи. Введем псевдобулевы переменные:

$$x_{ij} = \begin{cases} 1, \text{ если алгоритм } L_i \in \widetilde{L} \text{ обработки заявки } c_l \in C \text{ реализуется } j\text{-й программой;} \\ 0 \text{ в противном случае.} \end{cases}$$

Для ЭВМ, работающих в реальном времени (время реализации алгоритмов строго ограничено и задаче соответствует свой алгоритм решения $L_i \in \overset{\rho}{L}$, *(p=n),* задачу выбора набора программ для реализации множества алгоритмов $\overset{\rho}{L}$ можно сформулировать следующим образом (при этом заменим соответствующие обозначения L_i на индекс *i):* минимизировать целевую функцию

$$\Phi(X) = \sum_{i=1}^{n} \sum_{j=1}^{m_i} (a_{ij} + h_{ij}) x_{ij} \tag{68}$$

при временных ограничениях на реализацию алгоритмов

$$\sum_{i \in C_1} \sum_{j=1}^{m_i} \bar{t}_{ij} g_i x_{ij} \leqslant T_1;$$

$$\sum_{i \in C_1} \sum_{j=1}^{m_i} \bar{t}_{ij} g_i x_{ij} + \sum_{i \in C_2} \sum_{j=1}^{m_i} \bar{t}_{ij} g_i x_{ij} \leqslant T_2;$$

$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$

$$\sum_{l=1}^{u} \sum_{i \in C_l} \sum_{j=1}^{m_i} \bar{t}_{ij} g_i x_{ij} \leqslant T_l;$$

$$\max_{L_i \in L} \left\{ \sum_{j=1}^{m_i} b_{ij} x_{ij} \right\} \leqslant B; \quad \sum_{i=1}^{n} \sum_{j=1}^{m_i} h_{ij} x_{ij} \leqslant H; \tag{69}$$

при ограничениях на аддитивные параметры

$$\sum_{i=1}^{n}\sum_{j=1}^{m_i} r_{ij}^{(k)}x_{ij} \leqslant R_k,\ k=\overline{1,\ q};$$ (70)

при логических условиях существования единственной программы реализации каждого алгоритма

$$\sum_{j=1}^{m_i} x_{ij} = 1,\ i=\overline{1,\ n}.$$ (71)

Здесь $\overline{t_{ij}}$ — математическое ожидание времени реализации алгоритма L_i j-м методом; g_i — относительная частота появления алгоритма в общей программе работы ЭВМ; R_k, B, H — ограничения на k-й ресурс, объем СОЗУ и ОЗУ соответственно; T_l— максимально допустимое время реализации задач, имеющих l-й приоритет.

Задача (68) — (61) также является задачей дискретного программирования с псевдобулевыми переменными. Подставляя в целевую функцию задачи оптимизации другие параметры, в частности стоимость, время решения задач, энергетические и другие параметры системы памяти, можно оптимизировать структуру системы памяти ЭВМ по соответствующим критериям.

Рассмотренная задача может быть использована для синтеза технических и программных средств САК.

4.6. Последовательные методы в задачах синтеза структур

Выбор оптимального варианта структуры методами, базирующимися на полном переборе вариантов, является дорогостоящей, трудоемкой и, как правило, неосуществимой процедурой. Использование методов математического программирования для формирования рекомендаций по решению задач структурного синтеза требует большой предварительной подготовки для исследования пространства рекомендаций и не всегда оправдано из-за больших трудностей учета многочисленных факторов, влияющих на корректность постановки задачи оптимального синтеза структурі, и из-за существенных вычислительных трудностей решения задач математического программирования большой размерности.

При проектировании сложных систем довольно эффективными оказываются последовательные методы анализа и синтеза.

Последовательные методы анализа основаны на направленной генерации множества вариантов формируемых рекомендаций и

осуществлении процедуры анализа вариантов с целью выбора наилучшего путем последовательного отсеивания неперспективных вариантов.

При решении задач проектирования структур можно рекомендовать метод последовательного анализа вариантов, основанный на обобщении идей теории последовательных статистических решений А. Вальда.

Последовательные алгоритмы синтеза основаны на наращивании структуры путем добавления по определенным правилам элементов к некоторому начальному элементу.

Метод последовательного формирования рекомендаций, анализа и отсеивания вариантов. В основе этого метода лежит идея процесса формирования рекомендаций в виде многоступенчатой структуры. Каждая ступень связана с проверкой наличия определенных свойств у подмножества вариантов и либо ведет к непосредственному сокращению исходного множества вариантов, либо подготавливает возможность такого сокращения в будущем. Для решения задачи необходимо определить отличительные свойства, которыми должен обладать искомый вариант. Первоначально из множества признаков выбирают наиболее легко проверяемые и присущие одновременно возможно большему числу вариантов. После этого выбор численной схемы решения состоит в выборе рационального порядка проверки признаков, позволяющего провести отсев неконкурентоспособных вариантов и найти оптимальный.

Алгоритмы последовательного анализа вариантов основаны на принципе оптимальности, который представляет собой естественное обобщение принципа оптимальности динамического программирования для решения многошаговых задач оптимизации.

Напомним, что в методе динамического программирования выбор управления на отдельном шаге производится не с точки зрения интересов данного шага, выражающихся в минимизации потерь на данном шаге, а с точки зрения всего многошагового процесса формирования рекомендаций в целом, выражающихся в минимизации суммарных потерь на всех последующих шагах. Отсюда следует основное свойство оптимального процесса формирования рекомендаций, заключающееся в том, что каковы бы ни были начальное состояние и начальная рекомендация, последующие рекомендации на каждом шаге должны быть оптимальными относительно состояния, являющегося результатом применения первой рекомендации. Из этого свойства следует, что ***оптимизация выбора рекомендации для многошагового процесса формирования рекомендаций заключается в выборе рекомендаций только на последующих шагах процесса***.

***<u>Основное правило отсева бесперспективных вариантов</u>* — *монотонная рекурсивность, идейно родственная критерию оптимальности динамического программирования*.**

Пусть имеются множество $W=\{w\}$ вариантов рекомендаций и множество опытов $\Pi=\{\pi_\alpha\}$. Каждый вариант $w \in W$ описывается некоторым множеством признаков. Задача состоит в определении подмножества $W^* \subseteq W$, инвариантного относительно любого π_α и содержащего оптимальную рекомендацию $w_0 \in W^*$. Для определения подмножества W^* необходимо поставить опыты по анализу и оценке свойств элементов $w \in W$. Исходы опытов позволяют отбросить неперспективные варианты w, которые не имеют общих частей с элементами подмножества W^*, и сделать заключение о целесообразности постановки последующих опытов с целью определения элементов, входящих в подмножество W^*.

Приведем описание правила отсева. Пусть задано некоторое базовое множество **X**. Обозначим множество конечных последовательностей вида

$$\mathbf{p} = (x_1, \ldots, x_l, \ldots, x_{k_p}), \quad x_l \in X, \quad 1 \leqslant l \leqslant k_p,$$

через $P(X)$. В этом множестве выделено некоторое подмножество допустимых последовательностей $W(X) \subseteq P(X)$, в свою очередь во множестве $W(X)$ выделено подмножество $W_0(X) \subseteq W(X)$ полных допустимых последовательностей. Пусть задана последовательность **p**. Последовательность

$$\mathbf{p}_l = (x_1, \ldots, x_l), \quad 1 \leqslant l \leqslant k_p,$$

называют l-м начальным отрезком последовательности **p**, а последовательность

$$\mathbf{p}^{(q)} = (x_q, x_{q+1}, \ldots, x_{k_p}), \quad 1 \leqslant q \leqslant k_p,$$

называют q-м конечным отрезком. Если $q=l+1$, то соответствующие части последовательности **p** называют сопряженными. Рассмотрим две допустимые последовательности $\mathbf{p}_1$ и $\mathbf{p}_2$. В $\mathbf{p}_1$ выделены l_1-й начальный отрезок $\mathbf{p}_{1_{l_1}}$ и $(l_1 + 1)$-й конечный отрезок $\mathbf{p}_{1_{l_1+1}}$, в $\mathbf{p}_2$ выделены l_2-й начальный отрезок $\mathbf{p}_{2_{l_2}}$ и (l_2+1)-й конечный отрезок $\mathbf{p}_{2_{l_2+2}}$. Если функционал Φ, определенный на множестве $W(X)$, обладает тем свойством, что из

$$\mathbf{p}_{1l_1} \in W(X); \quad \mathbf{p}_{2l_2} \in W(X); \quad \mathbf{p}_{1l_1+1} \equiv \mathbf{p}_{2l_2+1};$$
$$\Phi(\mathbf{p}_{1l_1}) < \Phi(\mathbf{p}_{2l_2})$$

следует $\Phi(\mathbf{p}_1) < \Phi(\mathbf{p}_2)$, то его называют ***монотонно-рекурсивным***.

Вариант $w_0 \in W^*$, характеризующийся максимальным значением функционала Ф, определяется согласно обобщенному принципу оптимальности: если заданы монотонно-рекурсивный функционал Ф и две допустимые последовательности $\mathbf{p}_1$ и $\mathbf{p}_2$, причем

$$\Phi(\mathrm{p}_1) < \Phi(\mathrm{p}_2), \quad P(\mathrm{p})_1 \subseteq P(\mathrm{p}_2),$$

то последовательности, у которых начальным является отрезок $\mathbf{p}_1$ будут неперспективными и не входят в подмножество W^*. Здесь $P(\mathbf{p}_i)$ — множество конечных отрезков, сопряженных с последовательностью $\mathbf{p}_i$.

Таким образом, схема последовательного выбора оптимального варианта сводится к следующим повторяющимся процедурам:

1. Множество вариантов рекомендаций W разбивают на несколько подмножеств, каждое из которых обладает специфическими свойствами.
2. На основании обобщенного принципа оптимальности отсеивают те варианты $w \in W$, значение функционала Ф для которых заведомо не может быть максимальным.
3. Последовательно формируются и анализируются не отсеянные ранее варианты рекомендаций.
4. Для множества, состоящего из вновь образованных в п. 3 допустимых последовательностей, и последовательностей не исключенных и не продолженных ранее, выполняют операции п. 2.

Далее операции п. 2, 3 и 4 циклически повторяются. Если на каком-то этапе процесса формирования рекомендаций не останется ни одной последовательности, требующей своего развития до получения полной последовательности, то процесс завершен и в качестве рекомендации берется одна из рассмотренных полных допустимых последовательностей с наибольшим значением функционала Ф($\mathbf{p}$).

Основным достоинством метода последовательного анализа вариантов является значительная экономия в вычислительной процедуре за счет отсеивания бесперспективных начальных частей вариантов до их полного построения. Экономия является тем существенней, чем больше определяющих свойств задачи использовано для построения процедур анализа и отсева.

Последовательные алгоритмы структурного синтеза. ***Последовательные алгоритмы синтеза относятся к классу <u>эвристкческих.</u>*** Их достоинством является сравнительно высокая экономичность по затратам машинного времени и требуемому объему оперативной памяти за счет отсутствия процедуры многоразового анализа вариантов структуры. Однако, как правило, последовательные алгоритмы дают не оптимальные, а близкие к оптимальным рекомендации.

Развитие диалоговых средств общения пользователя с ЭВМ способствует широкому применению последовательных методов и алгоритмов в структурном синтезе из разнообразных областей. В качестве иллюстрации рассмотрим идею формирования последовательных алгоритмов для решения консультационных задач формирования рекомендаций по осуществлению конструкторского проектирования ЭВА — задач компоновки, размещения и трассировки.

При иерархической организации конструкции ЭВА под *компоновкой* понимают определение состава типовых конструкций каждого уровня. Задача компоновки обычно решается «снизу — вверх», т.е. известные схемы i—1-го уровня необходимо распределить по конструкциям i-го уровня. Так, например, на самом низшем уровне элементами могут выступать корпуса элементов, а конструкциями (блоками)— типовые элементы замены, связанные друг с другом путем разъемных соединений.

В качестве критериев оптимальности при решении задач компоновки наиболее часто используют критерии либо минимума суммарного числа N_i типов модулей

$$N_i = \sum_j x_{ij},$$

где x_{ij} — число модулей j-го типа i-го уровня схемы, либо минимума межблочных соединений

$$R_i = 0{,}5 \sum_{k=1}^{N_i} R_{ik},$$

где R_{ik} — число внешних связей каждого модуля i-го уровня.

Первый критерий непосредственно связан с конструктивными характеристиками аппаратуры и показателем технологичности стоимости, второй критерий ведет к повышению надежности конструктивной реализации схемы за счет сокращения числа разъемных соединений, уменьшению помех и задержек сигналов благодаря снижению числа межблочных соединений.

Для иллюстрации рассмотрим принцип построения последовательных алгоритмов компоновки по критерию минимума межблочной связности. Этот критерий широко используется при компоновке оборудования в различных технических приложениях. ***Идея алгоритмов заключается в следующем***. Первоначально выбирают исходный элемент (модуль) схемы. Выбор начального элемента основывается на схемотехнических соображениях.

В первый компонуемый узел включены все элементы, смежные с начальным, и сам начальный элемент. Если полученное число элементов равно максимально допустимому числу элементов в первом узле, то компоновка этого узла заканчивается. Если это число больше или меньше максимально допустимого, то выполняются соответствующие операции по устранению лишних или добавлению недостающих элементов, причем из нескомпонованных элементов выбирают такой, который имеет наибольшее число связей с элементами, уже вошедшими в состав компонуемого узла. Далее сформированный узел удаляют из схемы и компонуют новые узлы. Процесс повторяется до тех пор, пока схема не будет разбита на требуемое число частей или не будет выяснена невозможность этого.

Сформулируем описанный алгоритм в терминах теории графов. Пусть задан граф схемы $G=(X, U)$, который необходимо разбить на l частей $G_1, G_2,..., G_l$ с числом вершин в каждом соответственно

$$n_1, n_2, \ldots, n_l \ \left(\sum_{i=1}^{l} n_i = n, \quad |\mathbf{X}| = n\right).$$

Первоначально в графе G определяют вершину $x_i \in X$ с наибольшей локальной степенью $\rho(x_i)$ [напомним, что локальной степенью $\rho(x_i)$ вершины $x_i \in X$ называют число ребер, инцидентных этой вершине графа]. Если таких вершин несколько, то предпочтение отдается той, которая имеет большее число кратных ребер. Вершина x_i и все смежные с ней вершины включаются в граф G_1. Обозначим это множество вершин через Γx_i. Если $|\Gamma x_i| = n_1$ то G_1 образован, если же $|\Gamma x_i| > n_1,$ то из графа G_1 удаляют вершины, связанные с остающимися вершинами графа G меньшим числом ребер. Когда $|\Gamma x_i| < n$, то выбирают вершину $x_j \in \Gamma x_i |$, удовлетворяющую условию

$$\sigma(x_j) = \max_{x_k \in \Gamma x_i} \{\sigma(x_k)\} = \max_{x_k \in \Gamma x_i} \{\rho(x_k) - a_k\},$$

где a_k — число ребер, соединяющих вершину x_k со всеми невыбранными вершинами графа G.

Строят множество вершин Γx_j, смежных x_j, и процесс выборки вершин G_1 повторяют. Образованный подграф G_1 исключают из исходного и получают граф $G^*=(X^*, U^*)$, где X^* $X \backslash X_1$, $U^* = U \backslash U_1$. Далее в графе G^* выбирают вершину с наибольшей локальной степенью, включают ее в G_2 и процесс повторяют до тех пор, пока граф G не будет разрезан на l частей.

Первоначальную компоновку можно улучшить с помощью итерационных алгоритмов, основанных на реализации методов парных или групповых перестановок элементов из одной части схемы в другую таким образом, чтобы улучшилось значение целевой функции с учетом заданных ограничений.

Задача *размещения* заключается в определении оптимального (с точки зрения выбранного критерия оптимальности) положения элементов и связей между ними в монтажном пространстве типовой конструкции с учетом заданных конструктивно-технологических ограничений. Исходными данными в задаче решения являются принципиальная электрическая схема узла или устройства, метрические параметры и топологические свойства монтажного пространства.

Главная цель размещения — создание наилучших условий для трассировки с учетом обеспечения тепловых режимов и электромагнитной совместимости электрорадиоэлементов. Несмотря на обилие существующих критериев размещения (минимума пересечений, минимума суммарной длины соединений и т.д.) истинной целью размещения компонентов является максимальное упрощение процесса трассировки соединений, т. е. достижение минимального числа непроведенных трасс. При размещении n электрорадиоэлементов в регулярном монтажном пространстве с числом позиций m общее число размещений $N(n, m)$ определяется как

$$N(n,m) = n!C^n_m = m!1(m — n)!.$$

В связи с этим поиск оптимального размещения с помощью перебора нецелесообразен уже при $n>15$.

Имеется много разновидностей последовательных алгоритмов размещения. Основной идеей этих алгоритмов является идея упорядочения электрорадиоэлементов по определенным признакам. Сначала устанавливают очередность электрорадиоэлементов, а затем для каждого из них определяют наилучшую позицию по выбранному критерию, например по суммарной длине связей с уже размещенными компонентами. Затем процесс повторяют для оставшихся компонентов и свободных позиций. Связность размещаемых элементов задается матрицей смежности R графа $G=(X, U)$. Для выбора размещаемого элемента используют различные оценки степени связности.

Пусть на k-м шаге алгоритма размещено $I_k \subset I$ элементов, тогда $I'_k = I \backslash I_k$ — множество еще не размещенных элементов. Основными правилами для выбора элемента на (k+1)-м шаге алгоритма являются:

а) максимум суммарной связности h_i со всеми размещенными элементами

$$h_i = \max_{i \in I'_k} \Big\{ \sum_{j \in I_k} r_{ij} \Big\}, \quad i \neq j;$$

б) максимум разности связей f_i между размещенными и неразмещенными элементами

$$f_i = \max_{i \in I'_k} \left\{ \sum_{j \in I_k} r_{ij} - \sum_{j \in I'_k} r_{ij} \right\}, \quad i \neq j.$$

Выбранный для размещения элемент устанавливают в такую позицию среди оставшихся незаполненных, при которой будет иметь наименьшее значение некоторая целевая функция. Для многих задач размещения в качестве такой функции может быть выбрана суммарная длина связей с уже размещенными элементами.

Последовательные алгоритмы размещения требуют небольших затрат машинного времени, относят их к классу полиномиальных алгоритмов со сложностью $O(n)$, приводящих к неоптимальным решениям. Улучшить решение можно путем применения итерационных алгоритмов компоновки, основанных на изменении позиций одиночных элементов или групп элементов. Итерационные алгоритмы также относятся к классу полиномиальных со сложностью порядка $O(n^2)$ — $O(n^4)$.

Задача *трассировки* заключается в определении конкретной геометрии печатного или проводного монтажа, реализующего соединения между элементами схемы. Исходными данными для трассировки являются список цепей, метрические параметры и топологические свойства типовой конструкции и ее элементов, а также результаты решения задачи размещения, по которым находят координаты выводов элементов.

При решении задачи трассировки строят множество трасс, соединяющих выводы элементов соответствующих цепей схемы. Разработка отдельной трассы представляет собой построение на фиксированных вершинах минимального покрывающего или связывающего дерева, а разработка множества трасс сводится к построению леса непересекающихся минимально покрывающих или связывающих деревьев. Известно, что на n вершинах можно построить n^{n-2} различных деревьев, поэтому точное решение задачи трассировки методом полного перебора практически нереализуемо.

В последовательных алгоритмах трассировки трассы цепей проводятся в определенном порядке одна за другой, при этом каждая проложенная трасса становится препятствием для всех последующих цепей. В последовательных алгоритмах производят локальную оптимизацию качества трассировки каждой отдельной трассы без учета влияния размещения данной трассы на возможность проведения последующих. Это приводит к тому, что некоторые участки платы могут оказаться заблокированными.

Большинство известных алгоритмов трассировки основывается на <u>волновом алгоритме (алгоритм Ли).</u> Основные принципы волнового алгоритма Ли заключаются в следующем. Плоскость трассировки разбивают на прямоугольные площадки — дискреты заданного размера. Размер дискретной площадки определяется допустимыми размерами проводников и расстояниями между ними. Задача проведения трасс сводится к получению последовательности дискретов, соединяющих элементы a и b, соответствующие началу и концу проводимой трассы.

Вводим целевую функцию $F=F(f_1,...,f_r)$ как критерий качества пути. Начиная с элемента a дискретам, соседним с ранее просмотренным, присваивают определенное значение целевой функции $F_{ij}=m\ (i, j)$. Этот этап проводится итерационно до элемента b, которому присваивают некоторое значение веса $m(i_b,\ j_b)$. Затем, начиная от элемента b, перемещаются к элементу a по пройденным дискретам таким образом, чтобы значения целевых функций дискретов монотонно убывали. В результате получается трасса, соединяющая элементы a и b.

Обычно работа алгоритма Ли реализуется следующим образом. На трассируемой плоскости из источника a моделируется распространение волны до тех пор, пока не будет достигнута точка b или пока на некотором шаге фронт волны не сможет включить ни одного незанятого дискрета. Эту часть алгоритма называют распространением волны. После этого проводят трассу, начиная от конечной точки b, по дискретам с последовательно уменьшающимися весами. На рис. 3 цифры в квадратах соответствуют весам дискретов, занятые дискреты заштрихованы, а построенная трасса показана штриховой линией.

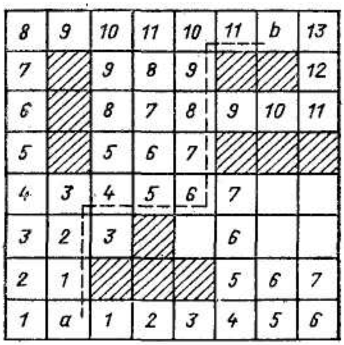


Рис.3. Пример соединения элементов a и b с помощью волнового алгоритма

Существует несколько вариантов проведения пути, из которых конструктор (или ЭВМ) выбирает один, наиболее удовлетворяющий заданным требованиям. Имеется многообразие модификаций волновых алгоритмов, направленных на повышение быстродействия трассировки, уменьшение объема требуемой оперативной памяти ЭВМ и т. д. Волновые алгоритмы широко применяют в различных областях техники, в частности при разработке сетей связи и сетей ЭВМ.

4.7. Алгоритм распределения реализуемых структурой и ее элементами функций по структурным модулям

Пусть задано множество технологических функций, реализуемых в системе (i = 1, 2, ...,), задано множество функциональных модулей структуры (j = 1, 2.....). Необходимо так распределить функции по модулям структуры, чтобы достигнуть максимального значения эффективности, не выходя из области допустимых ограничений. Пусть C_{ij} — затраты на реализацию i-й функции (технологических операций) в j-м модуле, t_{ij} — время реализации функции в j-м модуле.
Вводим дополнительную переменную

$$x_{ij} = \begin{cases} 1, & \text{если } i \text{ — 1-я функция реализуется в } j\text{-м модуле,} \\ 0 & \text{— в противном случае.} \end{cases}$$

Оптимизация распределения технологических функций по модулям может производиться по одной из следующих целевых функций:

$$\min_{x_{ij}} \sum_{i=1}^{I} \sum_{j=1}^{J} C_{ij} x_{ij}; \tag{72}$$

$$\min_{x_{ij}} \sum_{i=1}^{I} \sum_{j=1}^{J} t_{ij} x_{ij}; \tag{73}$$

$$\min_{x_{ij}} \left[\max_{j} \left(\sum_{i=1}^{J} t_{ij} x_{ij} \right) \right]. \tag{74}$$

Функция (72) соответствует минимизации затрат, функция (73) — минимизации общего времени реализации технологических функций, а функций (74) — минимизации максимального времени, реализации технологической функции в модуле. При этом могут учитываться следующие ограничения:

а) связи между функциями (обычно задаются графом G *(J),* где *J* — множество функций;

б) связи между модулями, в состав которых входят технические средства (элементы) системы (обычно задаются графом *G (J),* где *J* — множество модулей);

в) ограничение на общее время реализации технологических функций

$$\sum_{i=1}^{I}\sum_{j=1}^{J} t_{ij}x_{ij} \leqslant T_{з},$$

если минимизируются затраты, либо ограничение на общие затраты по реализации функций в системе

$$\sum_{i=1}^{I}\sum_{j=1}^{J} C_{ij}x_{ij} \leqslant C_{з},$$

если минимизируется время реализации всех функций;

г) ограничение на загрузку технологическими объектами каждого модуля

$$\sum_{i=1}^{I} \lambda_i t_{ij}x_{ij} \leqslant \rho_{зj},\ j = 1,\ 2,\ \ldots,\ J,$$

где λ_i — интенсивность поступления *i*-го объекта на изготовление; ρ_{sj} — допустимая загрузка *j*-го модуля;

д) если каждая функция реализуется только в одном модуле системы, то

$$\sum_{j=1}^{J} x_{ij} = 1,\ i = 1,\ 2,\ \ldots,\ J.$$

В зависимости от того, как выбрана целевая функция и какие ограничения учитываются, возникает ряд частных постановок задачи оптимального распределения технологических функций по модулям.

Миимизация общих затрат (общего времени) при ограничениях на загрузку каждого из модулей и при условии, что каждая функция реализуется только в одном модуле системы, т. е.

$$\min \sum_{i=1}^{I}\sum_{j=1}^{J} C_{ij}x_{ij} \text{ либо } \min \sum_{i=1}^{I}\sum_{j=1}^{J} t_{ij}x_{ij} \text{ при } \sum_{i=1}^{I} \lambda_i t_{ij}x_{ij} \leqslant \rho_{зj};$$

$$\sum_{j=1}^{J} x_{ij} = 1;\ j = 1,\ 2,\ \ldots,\ J;\ i = 1,\ 2,\ \ldots,\ J.$$

Минимизация общих затрат (общего времени) при ограничениях на общее время (общие затраты) и при условии, что каждая функция реализуется только в одном модуле системы, т. е.

$$\min \sum_{i=1}^{I} \sum_{j=1}^{J} C_{ij} x_{ij}; \tag{75}$$

$$\sum_{i=1}^{I} \sum_{j=1}^{J} t_{ij} x_{ij} \leqslant T_3;$$

$$\sum_{j=1}^{J} x_{ij} = 1; \; i = 1, 2, \ldots, J; \; x_{ij} = 0; \; 1.$$

В дальнейшем будем рассматривать только задачу (75). Дерево ветвления будем строить следующим образом. Подмножество первого уровня разбиения формируем, фиксируя возможность реализации первой функции различным модулям *(x_1, x_2, ..., x_{j_1}, ..., x_j).* Множество x_{j_1} включает в себя варианты, где первая функция реализуется в модуле j_1; а распределение остальных функций по модулям произвольное. Аналогично, множеством второго уровня формируем, фиксируя соответствие второй функции различным модулям. Множество $x_{j_1 j_2}$ включает в себя все варианты реализаций, где первая функция реализуется в модуле j_1, вторая функция — в модуле j_2, а остальные функции имеют произвольное распределение по модулям и т. д. Для каждого из подмножеств (вершин дерева) необходимо построить оценки целевой функции и ограничения. Общее выражение оценки целевой функции для множества вариантов $x_{j_1, j_2, \ldots, j_i, \ldots, j_e}$ в данной задаче может быть построено следующим образом:

$$V_c \left(x_{j_1, j_2, \ldots, j_i, \ldots, j_e}\right) = \sum_{i \leqslant e} C_{ij_i} + \sum_{i > e} \min_j C_{ij}, \tag{76}$$

а общее выражение для оценки ограничения может быть построено аналогично:

$$V_t \left(x_{j_1, j_2, \ldots, j_i, \ldots, j_e}\right) = \sum_{i \leqslant e} t_{ij_i} + \sum_{i > e} \min_j t_{ij}, \tag{77}$$

где j_1, j_2 , ..., j_i , ..., j_e — множество модулей системы, закрепленных за соответствующими функциями 1, 2, ..., *i*, ..., *e*.

Оценка для функции ограничения (77) необходима в данном случае для исключения из процесса ветвления множества заведомо неподходящих вариантов с учетом принятого ограничения на общее время реализации функций. Для ускорения процесса поиска используются всякие дополнительные приемы, учитывающие специфику задачи. Рассмотрим их на численном примере. Пусть заданы: матрицы стоимости

$$C = \begin{Vmatrix} 3 & 7 & 2 & 2 \\ 4 & 8 & 1 & 3 \\ 5 & 9 & 6 & 2 \\ 6 & 10 & 7 & 1 \\ 7 & 5 & 3 & 1 \end{Vmatrix};$$

временных затрат

$$T = \begin{Vmatrix} 1{,}5 & 3 & 2 & 9 \\ 2 & 6 & 5 & 10 \\ 3 & 7 & 6 & 11 \\ 4 & 8 & 7 & 12 \\ 4 & 9 & 8 & 5 \end{Vmatrix};$$

значение $T_s = 20$.

Так как каждая функция может быть реализована только одним модулем системы, то

$$\min_{x_{ij}} \sum_{i=1}^{I} \sum_{j=1}^{J} C_{ij} x_{ij} = \sum_{i=1}^{I} \min_{x_{ij}} \sum_{j=1}^{J} C_{ij} x_{ij}.$$

Используя это условие и принимая во внимание ограничение по времени, можно исключить из матриц C и T элементы, которые не влияют на выбор оптимального решения. Такими элементами в каждой строке являются:

$$C_{ij} > \min C_{ij} = C_{i\rho};\ t_{ij} > t_{i\rho}.$$

После проведения такого преобразования матрицы $C^{(0)}$ и $T^{(0)}$ будут выглядеть следующим образом:

$$C^{(0)} = \begin{Vmatrix} 3 & — & 2 & — \\ 4 & — & 1 & — \\ 5 & 9 & 6 & 2 \\ 6 & 10 & 7 & 1 \\ 4 & — & — & — \end{Vmatrix};$$

$$T^{(0)} = \begin{Vmatrix} 1{,}5 & - & 2 & - \\ 2 & - & 5 & - \\ 3 & 7 & 6 & 11 \\ 4 & 8 & 7 & 12 \\ 4 & - & - & 5 \end{Vmatrix}.$$

Принимая во внимание ограничение по времени $T_8 = 20$, можно из матриц $C^{(0)}$ и $T^{(0)}$ также вычеркнуть элементы, при которых всегда будет нарушено ограничение по времени Распишем для выделения таких элементов данное ограничение в следующем виде:

$$\sum_{i=1}^{r-1} \min_j t_{ij} + t_{rj} + \sum_{i=r+1}^{J} \min_j t_{ij} \leqslant T_8. \tag{78}$$

Изменяя индекс строки $r(r = 1, 2, ..., i, ..., J)$ и просматривая все элементы этой строки в соответствии с условием (78), можно еще раз преобразовать матрицы $C^{(0)}$ и $T^{(0)}$ и вычеркнуть такие элементы, если они имеются. Проведем такое преобразование матриц $C^{(0)}$ и $T^{(0)}$. В результате получим

$$C^{(1)} = \begin{Vmatrix} 3 & - & 2 & - \\ 4 & - & 1 & - \\ 5 & 9 & 6 & - \\ 6 & 10 & 7 & - \\ 7 & - & - & 1 \end{Vmatrix};$$

$$T^{(1)} = \begin{Vmatrix} 1{,}5 & - & 2 & - \\ 2 & - & 5 & - \\ 3 & 7 & 6 & - \\ 4 & 8 & 7 & - \\ 4 & - & - & 5 \end{Vmatrix}.$$

Из матриц $C^{(1)}$ и $T^{(1)}$ видно, что исходящие матрицы значительно упростились после проведения дополнительных преобразований. Перейдем теперь непосредственно к решению задачи на основе алгоритма ветвей и границ. Дерево решения с учетом построенных оценок (76) и (77) представлено на рис 4.

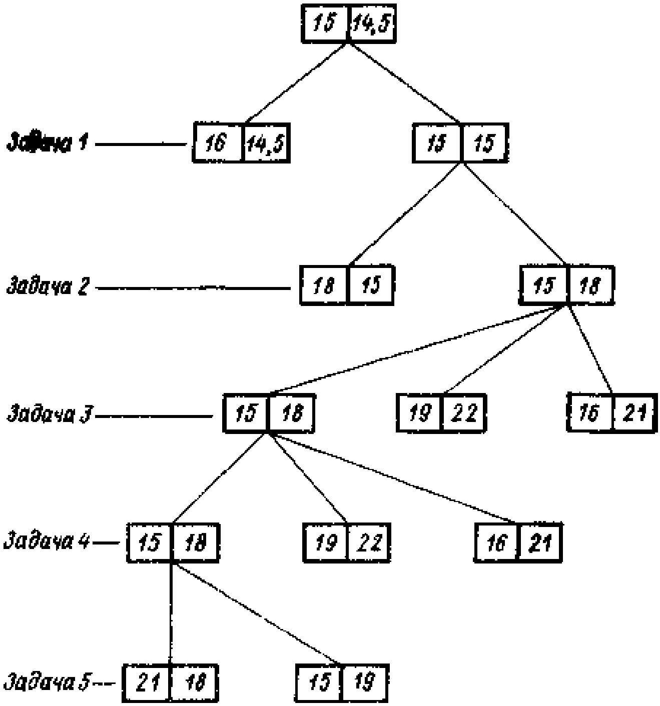


Рис. 4. Дерево решений с оценками (76) и (77)

В узлах дерева слева представлена оценка целевой функции, справа — оценка ограничения; зачеркнутые узлы, не удовлетворяющие ограничению. В результате решения получили следующее оптимальное распределение технологических функций по модулям: *1—3, 2—3, 3—1, 4—1, 5—4* (функция —модуль). Жирной чертой на рис. 4 представлен путь поиска оптимального решения на дереве вариантов.

Из примера видно, что учет специфики задачи значительно повышает эффективность вычислительной процедуры метода «ветвей и границ».

В заключение еледует отметить, что в ряде случаев структурный синтез хотя и может быть формализован в виде оптимальных задач дискретного программирования, однако он трудно разрешим в силу большой размерности задач, наличия большого числа случайных факторов, из-за векторного характера показателя эффективности. В таких ситуациях структурный синтез обычно понимают как выбор структуры элементов и принципов их взаимодействия из некоторого числа вариантов, перспективных для дальнейшего использования. Предварительный выбор таких вариантов по совокупности свойств осуществляется группой квалифицированных специалистов, имеющих опыт проектирования подобных систем, а детальное исследование количественных показателей основных свойств этих вариантов проводится для каждого из них в отдельности. После этого проводится сравнительный анализ и окончательный выбор основного варианта построения структуры системы. В этом случае будем говорить, что структурный синтез ведется на основе методов анализа.

4.8. Синтез комплекса технических средств системы

Задача выбора элементов КТС системы является наиболее важной при проектировании системы и их подсистем. В составе элементов КТС системы будем различать функционально законченные устройства, комплексы и подсистемы, такие как промышленные роботы, металлорежущие станки, накопители, координатные столы и др.

Задачу выбора элементов КТС системы сформулируем следующим образом. Пусть имеется совокупность $\{i\} = \{1,\ldots, n\}$ объектов одинакового функционального назначения, которые можно использовать для создания ГПС и ее подсистем. Каждый i-й объект имеет множество $\{k\} = \{1,...., k\}$ характеристик. Известны требования к каждой k-й характеристике объекта. Необходимо из совокупности объектов выбрать такой, который бы наиболее полно удовлетворял предъявленным к нему требованиям.

Все характеристики того или иного объекта можно разделить на три вида. Характеристики первого вида могут задаваться и определяться различными численными размерными или безразмерными физическими величинами, и чем больше численное значение характеристики объекта, тем он лучше при прочих равных условиях (например, производительность станков, скорость перемещения захвата промышленного робота, коэффициент готовности, средняя наработка на отказ, срок службы изделия).

Характеристики второго вида также измеряются численными физическими величинами, но чем меньше значение этой характеристики, тем лучше объект при прочих равных условиях (например, величиной точности позиционирования ПР, стоимость, потребляемая мощность, время установки детали в станок, период окупаемости изделия и т. д.).

Требования к характеристикам первого и второго видов задаются в виде численных значений физических величин.

Характеристики третьего вида не измеряются численными значениями физических величин, а выражаются интегральной возможностью реализации определенных материальных свойств и качеств данного объекта (например, возможность перевозки объекта в неработающем состоянии на тех или иных видах транспортных средств демонтажа и повторного монтажа оборудования, функционирования изделия в подвижном состоянии, возможность удовлетворения той или иной конструктивной группе). Характеристики этого вида могут формулироваться с использованием ряда частных характеристик первого и второго видов, однако интегральная оценка общей характеристики выражается утверждением о наличии или отсутствии у рассматриваемого объекта определенного свойства или качества. Требования к характеристикам третьего вида задаются в виде словесной формулировки о необходимости реализации объектом того или иного свойства или качества. Оценка факта реализации требования объектом производится в виде логического утверждения: «Требование выполнено» или «Требование не выполнено».

Обозначим: K_1, K_2 и K_3 — соответственно подмножества характеристик первого, второго и третьего видов. Численные значения k-й характеристики первого или второго вида для i-го объекта задаются величинами y_{ik} для всех $k \in K_1 \cup K_2$, а численные значения соответствующих требований — величиной Y_k.

Для характеристик третьего вида i-го изделия условно будем полагать, что если k-е требование выполняется объектом, то $y_{ik} = 1$, если не выполняется, то $y_{ik} = 0$ для всех $k \in K_3$. Само же k-е требование при формально-логическом описании задается в виде значения $y_k = 1$ для всех $k \in K$.

Величины y_{ik} и y_k, $k = 1,.., k$ характеризуют условия выбора. Вектор условий выбора запишем в виде

$$Y = \| y_{1\bar{1}}, \ldots, y_{n\bar{k}}, y_1, \ldots, y_{\bar{k}} \|. \qquad (79)$$

Разность $y_{ik} — y_k$ характеризует степень (полноту) выполнения (реализации) k -го требования i-м объектом. Отрицательные значения разностей говорят о том, что соответствующие характеристики i-го изделия меньше заданного требования на полученные разности, положительные — характеризуют превышение этого требования. Казалось бы $\sum_k (y_{ik} - y_k)$ должна характеризовать суммарную полноту реализации i-м объектом заданных требований по всей их совокупности. Но это справедливо лишь в том случае, когда все требования имеют одинаковые размерности отражающих их физических величин, что на практике практически не встречается. Обычно требования задаются физическими величинами различных измерений. Таким образом, указанная сумма не имеет физического смысла. Для того, чтобы характеризовать определенным образом выполнение совокупности требований i-м объектом, нужно эту сумму представить в безразмерных величинах или с одинаковой размерностью.

Пусть α_k — коэффициент приведения разностей y_{ik} — y_k к одинаковой размерности. Тогда $\sum_k \alpha_k \left(y_{ik} - y_k \right)$ в определенной мере характеризуют суммарную полноту реализации заданных требований, которое дает i-й объект. Однако и эта сумма не является исчерпывающей характеристикой, так как вычисляя ее, можно получить значение, равное нулю или близкое к нему, что характеризует факт полной реализации заданных требований, т. е. создается впечатление весьма благоприятной картины при выборе i-го объекта. На самом деле может оказаться, что ряд требований не выполняется, т. е. разности y_{ik} — y_k имеют отрицательные для всех $k \in K_1$ и положительные для $k \in K_2$ значения, а ряд требований перевыполняется, и указанные разности имеют положительные для $k \in K_1$ и отрицательные для $k \in K_2$ значения. Во избежание указанного положения на величины y_{ik} для $k = K_1 \cup K_1$ должны быть наложены соответствующие ограничения.

Рассмотрим особенности задания ограничения к характеристикам первого и второго видов. Подмножество характеристик первого вида $(k \in K_1)$ в свою очередь может содержать четыре подмножества.

Первое подмножество K_1^1 характерно тем, что требования к характеристикам этого подмножества задаются в виде ограничения снизу y_k.

Второе подмножество K_1^2 характеризуется тем, что требования к характеристикам этого подмножества задаются в виде ограничения снизу y_k и ограничения сверху A_k. Ограничение сверху вызвано тем, чтобы в ряде случаев при задании требований к наиболее важным характеристикам не допускать большой избыточности в реализации этих характеристик, что может повлечь за собой удорожание оборудования, увеличение габаритов, потребляемой мощности, усложнение и удорожание эксплуатации и т. д. Можно потребовать ограничить сверху производительность станка или требуемый диаметр обрабатываемой детали.

Третье подмножество характеристик K_1^3 требования к которым задаются в виде ограничений снизу y_k, но при этом допускается его снижение до величины B_k. Так, требуемая пропускная способность тракта накопителя деталей может быть задана величиной, например, $y_k \geq 200$ шт./ч, но допускается снижение ее до $B_k = 160$ шт./ч.

Четвертое подмножество характеристик K_1^4, требования к которым задаются в виде ограничений снизу y_k и ограничений сверху, но при этом допускается снижение ограничения снизу до величины B_k. Например, мощность, расходуемая станком, задается величиной $y_k = 10$ кВт, но эта величина не должна превосходить значения $A_k = 15$ кВт. Однако допускается снижение мощности до значения $B_k = 8$ кВт.

Характеристики второго вида, как и первого, могут также разбиваться на четыре подмножества.

Первое подмножество характеристик K_2^1, для которых требования задаются в виде ограничения сверху y_k. Например, вероятность брака обработанных деталей не должна превышать определенной наперед заданной величины.

Второе подмножество характеристик K_2^2, требования к которым задаются в виде ограничений сверху y_k и ограничений B_k. Так же как и для характеристик первого вида, это связано с необходимостью избежать громоздких и дорогостоящих решений при существенно лучших характеристиках, чем требуемые. Например, требуемая вероятность потерь от брака не должна превышать величины $y_k = 10^{-2}$, но в то же время в системе нет необходимости применения станков, обеспечивающих вероятность получения брака меньше $B_k = 10^{-3}$.

Третье подмножество характеристик K_2^3, для которых требования задаются в виде ограничений сверху y_k, но при этом допускается его увеличение до A_k. К примеру, среднее время восстановления изделия ограничено величиной $y_k = 0{,}5$*ч*, однако может быть увеличено до значения $A_k = 0{,}6$*ч*.

Четвертое подмножество характеристик K_2^4, требования к которым задаются в виде ограничений сверху y_k и ограничения снизу B_k, но при этом допускается его увеличение до значения A_k.

Если в примере для подмножества характеристик K_2^2 допустить возможность увеличения требуемой вероятности брака до значения $A_k = 10^{-1}$, то получим пример для подмножества характеристик K_2^4. Второй пример для этого подмножества состоит в том, что требуемая величина среднего времени восстановления вышедшего из строя изделия ограничивается значением y_k=0,8 ч, но при этом, исходя из функционального назначения рассматриваемого изделия, нет надобности стремиться к снижению этого времени ниже B = 0,3 ч, а при необходимости значение y_k можно увеличить до A_k=1 ч. Величины A_k и B_k должны содержаться в ТЗ на разработку системы.

С учетом указанных ограничений, представленная сумма с достаточной для лица, принимающего решение степенью строгости характеризует суммарную полноту реализации заданных требований i-м изделием. Однако, для принятия более обоснованного решения и выработки более аргументированного суждения о качествах i-го изделия целесообразно учитывать важность или вес g_k каждого k-го требования.

Тогда сумма

$$\sum g_k L_k (y_{ik} - y_k) \tag{80}$$

в учетом введенных ограничений наиболее полно характеризует суммарную полноту реализации заданных требований i-м объектом.

Если попытаться определить смысл введенных коэффициентов g_k и L_k, то обратимся к понятию качества функционирования данного комплекса, подсистемы или системы в целом, куда входят рассматриваемые объекты, т. е. к качеству функционирования объекта оснащения системы.

Существует функциональная зависимость качества функционирования объекта оснащения системы, включающего объекты рассматриваемого функционального назначения, от величин y_{ik}:

$$M = M(y_{i1}, \ldots, y_{ik})$$

откуда

$$M = (y_{i1}, \ldots, y_{ik}) - M(y_1, \ldots, y_k) \approx \sum_{k=1,\bar{k}} \frac{\partial M(y_k)}{\partial y_k}(y_{ik} - y_k), \tag{81}$$

где $\partial M(y_k)/\partial y_k$. — частная производная $\partial M(y_{ik})/\partial y_{ik}$ в точке $y_{ik} = y_k$. Для всех $k = 1,...., k$.

Сравнивая выражения (80) с (81), получаем

$$g_k = \frac{\partial M(y_k)}{\partial y_k} y_k; \; L_k = 1/y_k. \tag{82}$$

Таким образом, все требования g_k представляют собой изменение качества функционирования комплекса, включающего рассматриваемый объект при отклонении k-й характеристики объекта на единицу, а коэффициент n_k приводит разность y_{ik} — y_k к относительным безразмерным величинам. Полученные выводы следуют и из чисто физических соображений.

Вектор параметров выбора представим в виде

$$X = \| x_1, \ldots, x_n \|, \tag{83}$$

где его компоненты x_i= 1, если из совокупности n объектов выбирается i-й объект, и x_i=0 — в противном случае. Тогда критерий эффективности выбора с учетом приведенных выше рассуждений запишется в виде

$$F(X,\ Y) = \sum_{i=1,\ \bar{n}} \sum_{k=1,\ k} g_k \frac{y_{ik} - y_k}{y_k} x_i, \qquad (84)$$

где g_k определяется» по формуле (82). Эквивалентной записью критерия (84) является выражение

$$F(X,\ Y) = \sum_{i=1,\ \bar{n}} \sum_{k=1,\ k} g_k \frac{y_{ik}}{y_k} x_i. \qquad (85)$$

Отношения y_{ik}/y_k, входящие в формулу (852), выражают также полноту реализации k-го требования i-м объектом, а $\sum_k g_k \frac{y_{ik}}{y_k}$ характеризует суммарную полноту реализации всей совокупности требований i-м объектом с учетом веса каждого требования. Примем следующие ограничения на переменные x_i, которые должны выполняться при использовании критериев (84) и (85). Ограничения на численные значения переменных:

$$\sum_{i=1,\ \bar{n}} x_i = 1. \qquad (86)$$

Ограничения на пределы изменения величин y_{ik} для подмножеств характеристик следующие:

для K^1_1

$$\sum_{i=1,\ \bar{n}} y_{ik} x_i \geqslant y_k,\ k \in K_1^1; \qquad (87)$$

для K^2_1

$$\sum_{i=1,\ \bar{n}} y_{ik} x_i \geqslant y_k, \quad \sum_{i=1,\ \bar{n}} y_{ik} x_i \leqslant A_k,\ k \in K_1^2; \qquad (88)$$

для K^3_1

$$\sum_{i=1,\ \bar{n}} y_{ik} x_i \geqslant B_k,\ k \in K_1^3; \qquad (89)$$

для K^4_1

$$\sum_{i=1,\ \bar{n}} y_{ik} x_i \geqslant B_k, \quad \sum_{i=1,\ \bar{n}} y_{ik} x_i \leqslant A_k,\ k \in K_1^4; \qquad (90)$$

для K^1_2

$$\sum_{i=1,\ \bar{n}} y_{ik} x_i \leqslant y_k,\ k \in K_2^1; \qquad (91)$$

для K^2_2

$$\sum_{i=1,\ \bar{n}} y_{ik} x_i \leqslant y_k, \quad \sum_{i=1,\ \bar{n}} y_{ik} x_i \geqslant B_k, \quad k \in K_2^2; \qquad (92)$$

для K^3_2

$$\sum_{i=1,\ n} y_{ik}x_i \leqslant A_k, \quad k \in K^3_2; \tag{93}$$

для K^4_2

$$\sum_{i=1,\ n} y_{ik}x_i \leqslant A_k, \quad \sum_{i=1,\ n} y_{ik}x_i \geqslant B_k, \quad k \in K^4_2; \tag{94}$$

для подмножества характеристик $k \in K_3$ должно выполняться условие вида

$$\sum_{i=1,\ n} y_{ik}x_i = 1. \tag{95}$$

Кроме того, требования к некоторым характеристикам могут быть заданы только в виде ограничений (87)—(95), а не входить непосредственно в выражение для критерия эффективности выбора. Это целесообразно делать в том случае если не представляется возможность установить количественную связь между качеством функционирования технического средству и характеристикой объекта, т. е. определить вес или важность данной характеристики. Такими характеристиками объекта могут быть, например, мощность потребляемой электроэнергии,, занимаемая объектом площадь, его объем, весовые характеристики, стоимость, некоторые характеристики третьего вида $(k \in K_3)$ и т. д. В этом случае в формулах (84) и (85) сумма по k распространяется только на те характеристики, которые вводятся в выражение для критерия эффективности выбора. Если обозначить через K_0 подмножество характеристик, которые входят в выражение для критерия эффективности выбора, формулы (84) и (85) для этого критерия соответственно принимает вид

$$F(X,\ Y) = \sum_{i=1,\ n} \sum_{k \in K_0} g_k \frac{y_{ik} - y_k}{y_k} x_i; \tag{96}$$

$$F(X,\ Y) = \sum_{i=1,\ n} \sum_{k \in K_0} g_k \frac{y_{ik}}{y_k} x_i. \tag{97}$$

Эта запись критериев является более общей.

Задача выбора объекта того или иного функционального назначения из совокупности n объектов сводится к отысканию такого решения системы (86)—(95) которое обращало бы в максимум критерии эффективности выбора (96) либо (97), т. е. суммарную полноту выполнения требований с учетом веса каждого требования. Объекты, соответствующие решению этой задачи, наиболее полно удовлетворяют поставленным требованиям.

Эквивалентной формой записи ограничений на компоненты вектора условий выбора (79) являются ограничения на величины отклонения характеристик y_{ik} от задаваемых требований y_k. Для рассматриваемого случая эти ограничения (вместо 87)— (95) записываются в виде:

$$\sum_{i=1,\bar{n}} (y_{ik} - y_k)\, x_i \geqslant 0, \quad k \in K_1^1; \tag{98}$$

$$\sum_{i=1,\bar{n}} (y_{ik} - y_k)\, x_i \geqslant 0, \quad \sum_{i=1,\bar{n}} (y_{ik} - A_k)\, x_i \leqslant 0, \quad k \in K_1^2; \tag{99}$$

$$\sum_{i=1,\bar{n}} (y_{ik} - B_k)\, x_i \geqslant 0, \quad k \in K_1^3; \tag{100}$$

$$\sum_{i=1,\bar{n}} (y_{ik} - B_k)\, x_i \geqslant 0, \quad \sum_{i=1,\bar{n}} (y_{ik} - A_k)\, x_i \leqslant 0, \quad k \in K_1^4; \tag{101}$$

$$\sum_{i=1,\bar{n}} (y_{ik} - y_k)\, x_i \leqslant 0, \quad k \in K_2^1; \tag{102}$$

$$\sum_{i=1,\bar{n}} (y_{ik} - y_k)\, x_i \leqslant 0, \quad \sum_{i=1,\bar{n}} (y_{ik} - B_k)\, x_i \geqslant 0, \quad k \in K_2^2; \tag{103}$$

$$\sum_{i=1,\bar{n}} (y_{ik} - A_k)\, x_i \leqslant 0, \quad k \in K_2^3; \tag{104}$$

$$\sum_{i=1,\bar{n}} (y_{ik} - A_k)\, x_i \leqslant 0, \quad \sum_{i=1,\bar{n}} (y_{ik} - B_k)\, x_i \geqslant 0, \quad k \in K_2^4; \tag{105}$$

$$\sum_{i=1,\bar{n}} (y_{ik} - y_k)\, x_i = 0, \quad k \in K_3. \tag{106}$$

Второй эквивалентной формой записи ограничений на y_{ik} и y_k являются ограничения на относительные отклонения характеристик от задаваемых требований.

При этом в рассматриваемом случае получим ограничения:

$$\sum_{i=1,\bar{n}} \frac{y_{ik}}{y_k} x_i \geqslant 1, \quad k \in K_1^1; \tag{107}$$

$$\sum_{i=1,\bar{n}} \frac{y_{ik}}{y_k} x_i \geqslant 1, \quad \sum_{i=1,\bar{n}} \frac{y_{ik}}{y_k} x_i \leqslant \bar{A}_k, \quad k \in K_1^2, \tag{108}$$

где $\bar{A}_k = A_k / y_k$;

$$\sum_{i=1,\bar{n}} \frac{y_{ik}}{y_k} x_i \geqslant \bar{B}_k, \quad k \in K_1^3, \tag{109}$$

где $\bar{B}_k = B_k / y_k$;

$$\sum_{i=1,\bar{n}} \frac{y_{ik}}{y_k} x_i \geqslant \bar{B}_k, \quad \sum_{i=1,\bar{n}} \frac{y_{ik}}{y_k} x_i \leqslant \bar{A}_k, \quad k \in K_1^4; \tag{110}$$

$$\sum_{i=1,\bar{n}} \frac{y_{ik}}{y_k} x_i \leqslant 1, \quad k \in K_2^1; \tag{111}$$

$$\sum_{i=1,\bar{n}} \frac{y_{ik}}{y_k} x_i \leqslant 1, \quad \sum_{i=1,\bar{n}} \frac{y_{ik}}{y_k} x_i \geqslant \overline{B}_k, \quad k \in K_2^2; \tag{112}$$

$$\sum_{i=1,\bar{n}} \frac{y_{ik}}{y_k} x_i \leqslant \overline{A}_k, \quad k \in K_2^3; \tag{113}$$

$$\sum_{i=1,\bar{n}} \frac{y_{ik}}{y_k} x_i \leqslant \overline{A}_k,$$

$$\sum_{i=1,\bar{n}} \frac{y_{ik}}{y_k} x_i \geqslant \overline{B}_k, \quad k \in K_2^4; \tag{114}$$

$$\sum_{i=1,\bar{n}} \frac{y_{ik}}{y_i} x_i = 1, \quad k \in K_{в}. \tag{115}$$

Сформулированные выше задачи записаны для детерминированного случая, когда вектор условий выбора (79) не является случайной величиной. Это наиболее простой случай. Рассмотрим еще три случая, когда величины y_{ik} и y_k в отдельности или одновременно могут быть случайными. При этом случайные величины y_{ik} и y_k заменяются их математическими ожиданиями a_{ik} и b_k.

Необходимо отметить, что случайные величины y_{ik} и y_k в полученных моделях могут быть как независимыми, так и зависимыми. В большинстве случаев численные значения как характеристик y_{ik}, так и требований y_k, являются взаимно независимыми для различных номеров характеристик и требований (различных значений k). Если же ряд требований и характеристик является функционально связанными (зависимыми) величинами, то из последовательности этих зависимых величин выделяется последовательность независимых величин, которые и вводятся в условия выбора. Например, коэффициент готовности, средняя наработка на отказ и среднее время восстановления объекта есть последовательность зависимых величин, а любая пара из них является независимыми величинами. Следовательно, в условиях выбора должны вводиться только две из перечисленных выше характеристик и требований к ним. Однако, даже в том случае когда все же случайные величины y_{ik} и y_k по тем или иным причинам являются зависимыми, то целесообразно прибегать к допущению об их независимости, так как в противном случае получение выражений для критериев и некоторых видов ограничений может вызвать существенные трудности.

Приведенная выше совокупность моделей выбора содержит модели двух классов. Модели первого класса, в которых компоненты y_{ik} и y_k вектора условий выбора *Y* являются детерминированными величинами, соответствуют детерминированному принципу выбора в условиях определенности. Модели второго класса, в которых компоненты y_{ik} или y_k, либо обе вместе являются случайными величинами с известными распределениями вероятностей, соответствуют вероятностному принципу выбора в условиях риска.

В обоих случаях весьма существенным является нахождение подходящей формы критериев в эффективности выбора.

Форма записи критериев эффективности выбора (96)—(97), является универсальной и наиболее общей. В некоторых случаях можно предложить более удобную запись. Например, как следует из выражений (80)—(81), более естественной записью критерия эффективности (96) является

$$F(X,\ Y) = \sum_{i=1,\ \bar{n}} \ \sum_{k \in K_0} g_k\,(y_{ik} - y_k)\,x_i,$$

где $g_k = \partial M\ (y_{11},\ ...,\ y_n k_0) / \ \partial y_{ik}$ в точках $y_{ik} = y_k$ для всех $k \in K_0$, а $\alpha_k = 1$.

В этом случае размерность критерия *F (X, Y)* совпадает с размерностью функции $M(y_{11},...,y_{n\bar{k}_0})$.Однако такая запись была бы оправдана, если бы всегда представлялась возможность получить функцию $M(y_{11},...,y_{n\bar{k}_0})$. Но в ряде случаев функциональную зависимость качества функционирования объекта оснащения от его характеристик получить не представляется возможным. Тогда будем прибегать к некоторым способам определения весов в виде безразмерных величин и нормированию значений y_{ik} или $y_{ik} - y_k$, т. е. представление в виде y_{ik}/y_k или $(y_{ik} - y_k)/y_k$, является необходимым. Отсюда следует, что форма записи критериев эффективности выбора, приведенная в предыдущем параграфе, является наиболее общей и позволяет унифицировать вычислительные процедуры при любых способах определения весов характеристик g_k.

Рассмотрим некоторые приемы определения весов g_k при отсутствии зависимости между характеристиками объекта и качеством его функционирования.

Одним из распространенных приемов является определение весов характеристик с помощью экспертов или, как принято говорить, с помощью экспертных оценок.

Приведенные выше рассуждения позволяют сформулировать четыре характерных случая вынесения экспертных оценок.

1. Экспертные оценки при разнородном составе экспертов, т. е., когда эксперты различны по своей квалификации и степени объективности. В этом случае веса характеристик определяются из соотношения

$$g_k = \sum_{\beta=1,\,\bar{\beta}} W_{k\beta} q_{k\beta} \Big/ \sum_{\beta=1,\,\bar{\beta}} W_{k\beta} q_{k\beta}, \qquad (116)$$

где $\bar{\beta}$ — число экспертов, участвующих в оценке изделий; $W_{k\beta}$ — коэффициент, выражающий количественно степень квалификации β-го эксперта ($\beta = 1,\ldots, \bar{\beta}$) при оценке *k*-й характеристики, $0 \le W_{k\beta} \le 1$ (чем выше квалификация эксперта, тем больше значение коэффициента $W_{k\beta}$); $q_{\beta k}$ — коэффициент, отражающий степень объективности β-го эксперта, при оценке *k*-й характеристики; $q_{k\beta}$ — вес *k*-й характеристики, назначаемой β-м экспертом.

Веса характеристик целесообразно назначать в интервале от нуля до некоторого максимального значения, например в интервалах (0; 1), (0; 10), (0; 100), хотя верхняя граница интервала не обязательно должна быть единицей или кратной десяти. В принципе она может быть любым положительным числом. Однако, удобнее всего работать в указанных выше интервалах и предпочтительнее в интервале (0; 1). Что касается оценки степени -, объективности экспертов, то вопрос определения численного значения коэффициента $q_{k\beta}$ решается более сложно, чем простое приписывание тому или иному эксперту определенного значения этого коэффициента. Эксперт при проведении оценки характеристик изделия по тем или иным причинам может обнаружить тенденции к завышению весов одних характеристик и к занижению весов других характеристик. Поэтому с учетом упомянутых тенденций можно предложить следующую формулу для расчета величины $q_{k\beta}$:

$$q_{k\beta} = \begin{cases} \bar{q}_{k\beta} & \text{при тенденции к завышению;} \\ \min\left\{\dfrac{1}{\bar{q}_{k\beta}},\ \dfrac{q_{\max}}{q_{k\beta}}\right\} & \text{при тенденции к занижению,} \end{cases} \qquad (117)$$

где $q_{k\beta}$ — некоторый коэффициент ($0 \le q_{k\beta} \le 1$), характеризующий степень объективности β-го эксперта (чем больше значение $q_{k\beta}$, тем больше объективность β-го эксперта при оценке *k*-й характеристики); $q_{шах}$ — верхняя граница интервала, на котором оцениваются веса характеристик объекта

Например, пусть веса характеристик оцениваются в интервале (0; 10), т. е. $q_{шах} = 10$ и есть основания полагать, что β_1-й эксперт попытался завысить *k*-ю характеристику, но не очень сильно, и ему приписывается значение $\bar{q}_{k\beta_1} = 0{,}8$, а β_2-й эксперт обнаруживает усиленную тенденцию к занижению этой характеристики и ему приписывается значение $\overline{q}_{k\beta_2}$ = 0,4. Первый эксперт дает $q_{k\beta_1}$ = 9, а второй полагает $q_{k\beta_2}$ = 3. Тогда по формуле (116) находим: $q_{k\beta_1}$ = $\overline{q}_{k\beta_1}$ = 0,8; $q_{k\beta_2}$ = min {1/0,4; 0/3} = min {2,5; 3,33} = 2,5 и с учетом степени объективности экспертов получим $q_{k\beta_1}\ g_{k\beta_1}$ =0,8·9= 7,2; $q_{k\beta_2}\ g_{k\beta_2}$ = 2,5·3 = 7,5.

2. Экспертные оценки при разнородном составе экспертов. Разнородность характеризуется тем, что эксперты различны по своей квалификации, но одинаковы по степени объективности.

Тогда вес k-я характеристики определим по формуле:

$$g_k = \sum_{\beta=1,\ \bar{\beta}} W_{k\beta} g_{k\beta} \Big/ \sum_{\beta=1,\ \bar{\beta}} W_{k\beta}. \tag{118}$$

Этот случай может также иметь место тогда, когда эксперты хотя и различны по своей объективности, но степень объективности каждого установить не представляется возможным.

3. Экспертные оценки при разнородном составе экспертов. Разнородность характеризуется тем, что эксперты различны по степени объективности, но одинаковы по своей квалификации. В этом случае:

$$g_k = \sum_{\beta=1,\ \bar{\beta}} g_{k\beta} q_{k\beta} \Big/ \sum_{\beta=1,\ \bar{\beta}} q_{k\beta}. \tag{119}$$

По этой формуле может производиться расчет весов характеристик и тогда, когда эксперты хотя и различны по своей квалификации, но коэффициенты $W_{k\beta}$ определить не представляется возможным.

4. Экспертные оценки при однородном составе экспертов, т. е., когда эксперты одинаковы по своей квалификации и степени объективности. Тогда веса характеристик g_k определяют по формуле:

$$g_k = \Big(\sum_{\beta=1,\ \bar{\beta}} g_{k\beta} \Big) \Big/ \bar{\beta}. \tag{120}$$

Аналогичным образом можно поступать и тогда, когда: состав экспертов различен по квалификации и объективности, но приписать экспертам значения $W_{k\beta}$ и $q_{k\beta}$ не представляется возможным; состав экспертов различен по квалификации и одинаков по объективности, но определение $W_{k\beta}$ затруднительно или невозможно по объективности, определение $q_{k\beta}$ затруднительно или невозможно; состав экспертов одинаков по квалификации и различен по объективности, но определить $q_{k\beta}$ не представляется возможным.

5. Синтез структур средствами блочных групп и модуль-графами

Существует множество методов синтеза линейных пассивных электрических цепей. Большинство их названий связывают с фамилиями их создателей. Иногда эти методы отличаются оригинальностью и показывают творческую мысль авторов. Однако у них есть один основной недостаток. Как правило, это рецептурные методы, требующие различного подхода к каждой конкретной проблеме синтеза. Эти методы накладывают резкие ограничения на структуру синтезируемой схемы, а также на величины и род используемых элементов. Например, они ограничивают проблему синтеза схемами лестничной или мостовой структуры, четырехполюсником в виде перекрытого или двойного Т-образного моста и т. д. Такие большие ограничения не позволяют выделить среди этих схем лучшую. Конструктор принимает решение, но может оказаться, что такое решение не существует вообще.

Отсутствие в настоящее время метода, решающего проблему синтеза электрических цепей каким-то общим методом, можно объяснить следующими причинами:

1. Большой сложностью расчета, связанной с определением множества схем, удовлетворяющих всем условиям синтеза.
2. Отсутствием простого и одновременно достаточно общего расчетного алгоритма.
3. Отсутствием достаточно просто сформулированных условий реализации схемы.

По мере развития вычислительной техники трудности расчета можно полностью преодолеть, нужен только соответствующий расчетный алгоритм, хорошо приспособленный к технике машинного расчета и учитывающий условия физической реализации схем.

Таким алгоритмом могла бы быть алгебра блочных групп, которая непосредственно связывает геометрические свойства графа цепи с расчетным методом. Преимущество алгебры блочных групп заключается в том, что она позволяет алгебраическим методом записывать структуру схемы, а также дает простые связи всех изменений топологии схемы с операциями над блочными группами. По-видимому, это свойство представляет собой одно из основных преимуществ использования алгебры блочных групп для синтеза схем.

В настоящем разделе описано использование алгебры блочных групп для синтеза пассивных двух- и четырехполюсников. Напомним, что под термином «синтез» нужно понимать совокупность операций, необходимых для определения параметров электрической цепи или множества цепей, выполняющих поставленные требования. Поэтому синтез представляет собой понятие, противоположное понятию анализа, при котором имеется заданный объект, подлежащий анализу, т. е. изучению его свойств. Синтез в принципе должен быть инженерным методом, в котором проектировщик стремится к получению возможно лучшего решения.

Первой работой, посвященной проблеме синтеза электрических цепей, была работа Вильгельма Кауэра «Способ реализации двухполюсника с заданным импедансом», опубликованная им в 1926 г.

В данном разделе описана методика синтеза двухполюсника, основанная на разложении функции импеданса в цепную дробь. Следует обратить внимание на тот факт, что она была опубликована ровно через 100 лет после работы Ома, считающейся первой работой в области теоретической электротехники. В качестве одной из первых работ в области синтеза нужно также отметить и работу Отто Вруне (1931 г.). В этой работе автор представил метод синтеза RLC-двухполюсника, позволяющий реализовать произвольную конечную положительную действительную функцию импеданса. Кроме Кауэра и Вруне значительные результаты в области методов синтеза были получены Баттервортом (1930г.), Воде (1934 г.), Дарлингтоном (1939 г.), Боттом, Даффином (1949 г.), Гиллемином (1949 г.), Мията (1952 г.), Озаки (1953 г.), Реза (1954 г.) и другими. Работы этих авторов служат своеобразными вехами на пути развития методов синтеза линейных электрических цепей.

Представленный в данном разделе метод синтеза электрических цепей опирается на алгебру блочных групп, служит общим методом, не накладывающим никаких принципиальных ограничений на структуру схемы, и требует применения ЭВМ. Метод в представленном здесь виде не завершен до конца и требует дальнейших исследований для различных конкретных приложений.

5.1. Синтез структур, представленных пассивными двухполюсниками

Метод блочных групп позволяет решать проблему синтеза электрической цепи в общем виде без каких-либо ограничений, накладываемых на структуру проектируемой цепи.

Задачу синтеза электрической цепи можно разбить па два этапа:
1) топологический синтез графа цепи,
2) расчет величин отдельных элементов схемы.

Под термином «топологический синтез графа» понимаем совокупность операций, связанных с определением класса структур графов, реализующих поставленную проблему синтеза. Так как определение этих структур связано с определением соответствующих изображений блочных групп, то при решении этой проблемы важны условия существования геометрического изображения. Напомним эти условия:

1) блочная группа *A* должна иметь разложение на простые множители

$$A = P_1 P_2 \ldots P_m; \tag{1}$$

2) любой элемент α_{i_k} может присутствовать не более чем в двух числах блочных группах P_i, P_j произведения (1).

Кроме того, должны выполняться следующие дополнительные условия:

3) $P_i \neq P_j$; $i, j = 1, 2, \ldots, m$ $(i \neq j)$;

4) $P_i \neq \sum P_k$; $k = 1,2,\ldots,m$ $\{k \neq i)$, (2)

обеспечивающие $A \neq 0$.

Рассмотрим пассивный RLC-двухполюсник (рис. 1), в котором выделен импеданс z_α источника.

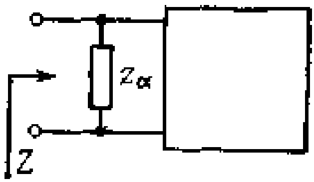


Рис. 1. Пассивный RLC-двухполюсник с вынесенным импедансом z_α.

С помощью теории блочных групп входной импеданс рассматриваемого двухполюсника можно выразить следующим образом

$$Z = \frac{\det\limits_{Y} \dfrac{\partial A}{\partial \alpha}}{\det\limits_{Y} A} \quad или \quad Z = z_\alpha \frac{\det\limits_{Z} \dfrac{\delta A^d}{\delta \alpha}}{\det\limits_{Z} A^d} \tag{3}$$

где *A* — блочная группа, изображением которой служит граф двухполюсника.

В дальнейшем будем пользоваться второй из приведенных выше формул, которая выражает входной импеданс двухполюсника через импедансы его элементов.

Рассмотрим наиболее общий случай, когда в каждой ветви двухполюсника присутствуют последовательно соединенные резистор, индуктивность и конденсатор. В этом случае импеданс каждой ветви выражается формулой

$$z_i = sL_i + R_i + s^{-1}C_i^{-1}. \tag{4}$$

Входной импеданс синтезируемого двухполюсника представляет собой положительную вещественную функцию

$$Z(s) = \frac{d_{\bar{n}+1}s^{\bar{n}+1} + d_{\bar{n}}s^{\bar{n}} + \ldots + d_1 s + d_0}{c_{\bar{n}}s^{\bar{n}} + c_{\bar{n}-1}s^{\bar{n}-1} + \ldots + c_1 s + c_0}.$$

Умножая числитель и знаменатель этой функции на множитель s^{-l}, где

$$l = E\left(\frac{\bar{n}}{2}\right)$$

(E (x) обозначает целую часть *x),* приводим нашу функцию к следующему виду:

$$Z(s) = \frac{a_{n+1}s^{n+1} + a_n s^n + \ldots + a_0 + \ldots + a_{-(n+1)}s^{-(n+1)}}{b_n s^n + b_{n-1}s^{n-1} + \ldots + b_0 + \ldots + b_{-n}s^{-n}} = \frac{P(s)}{Q(s)}. \tag{5}$$

Некоторые коэффициенты a_i, b_j полученной функции могут быть равны нулю.

Обозначим: b — число ветвей цепи, w — число узлов цепи, m — число независимых контуров цепи (цикломатическое число цепи).

Величины b, w и m связаны зависимостью

$$b - w + 1 = m.$$

С другой стороны, из теории блочных групп известно, что степень $\det_Z A^d$ равна цикломатическому числу m графа, служащего изображением блочной групп A. Отсюда следует, что степень полинома $Q(s)$ должна быть равна цикломатическому числу, т. е. $m = n$, и, кроме того,

$$b - w + 1 = n. \tag{6}$$

Так как для определения $3b$ неизвестных можем располагать $4n + 4$ уравнениями, то должно выполняться следующее неравенство:

$$3b \geq 4n + 4. \tag{7}$$

Если неизвестных больше, чем уравнений, то можно задаться значениями параметров некоторых элементов цепи. Формулы (6) и (7) дают оценку числа узлов цепи, т. е.

$$w \geq \frac{n+4}{3} + 1 \qquad (8)$$

Так как количество множителей в произведении (1) равно w— 1, то

$$m = E\left(\frac{n+4}{3}\right) + k \qquad (9)$$

В связи с этим схема будет содержать следующее число ветвей:

$$b = E\left(\frac{n+4}{3}\right) + k + n; \quad k = 1, 2, ..., \qquad (10)$$

Приведем операции, необходимые для синтеза двухполюсника.

1. Положим $m = E(n + 4)/3 + 1$ и рассмотрим множество ветвей

$$B = \{1, 2, ..., b\}; \quad b = E\left(\frac{n+4}{3}\right) + n + 1,$$

на основе которого строим произведения однострочных блочных групп

$$A = P_1 P_2 \dots P_m$$

с учетом условий реализуемости (1) и (2) и выбирая по крайней мере двухэлементные блочные группы P_i. Например,

$$A = [12]\,[2\ 3]\,[3\ 4\ 51 \quad \text{и т. д.}$$

Выполнив умножение, получаем всевозможные блочные группы, а значит, и всевозможные графы, соответствующие случаю $k = 1$.

2. Найдем все дополнительные блочные группы

$$A_1^d\, A_2^d\, ..., A_i^d\, ..., A_r^d.$$

3. Подсчитаем обратные алгебраические производные

$$\frac{\delta A_i^d}{\delta 1}, \frac{\delta A_i^d}{\delta 2}, ..., \frac{\delta A_i^d}{\delta \alpha}, ..., \frac{\delta A_i^d}{\delta b}; \quad i = 1, 2, ..., r$$

4. Вычислим детерминантные функции

$$\det_Z A^d{}_i, \det_Z \frac{\delta A_i^d}{\delta \alpha}, \quad \alpha = 1, 2, ..., b_i; \quad i = 1, 2, ..., r$$

Приравняем коэффициенты рациональных функций

$$\frac{P(s)}{Q(s)} = \frac{\det_Z \frac{\delta A_i^d}{\delta \alpha}}{\det_Z A_i^d} z_\alpha$$

и в результате получим систему нелинейных уравнений

$$f^{\alpha,i}_{n+1}(R_1, \ldots, L_1, \ldots, C_1, \ldots) = a_{n+1},$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$f^{\alpha,i}_{-(n+1)}(R_1, \ldots, L_1, \ldots, C_1, \ldots) = a_{-(n+1)},$$
$$\varphi^i_n(R_1, \ldots, L_1, \ldots, C_1, \ldots) = b_n, \tag{11}$$
$$\cdots\cdots\cdots\cdots\cdots\cdots\cdots\cdots$$
$$\varphi^i_{-n}(R_1, \ldots, L_1, \ldots, C_1, \ldots) = b_{-n}.$$

Эту систему нелинейных уравнений можно решить с помощью ЭВМ, например, методом итераций. Нетрудно заметить, что функции f_k и φ_k будут суммами следующих произведений:

$$R_{\alpha_1}\ldots R_{\alpha_m}L_{\alpha_1}\ldots L_{\alpha_m}C_{\alpha_1}\ldots C_{\alpha_l}\,. \tag{12}$$

Для первой итерации, например, можно принять

$$R^0_{\alpha_i} = 1\ \textit{ом},\quad L^0_{\alpha_i} = 1\ \textit{гн},\quad C^0_{\alpha_i} = 1\ \textit{ф},\quad i = 1,2, \ldots, r \tag{13}$$

и исследовать разность

$$f^0_v - b^0_v = \varepsilon^0_v;\qquad v = -(n+1), \ldots, (n+1),$$
$$\varphi^0_v - a^0_v = \delta^0_v;\qquad v = -n, \ldots, n. \tag{14}$$

В результате процесса итерации получим

$$R^n_{\alpha_i} \simeq R_{\alpha_i};\quad L^n_{\alpha_i} \simeq L_{\alpha_i};\quad C^n_{\alpha_i} \simeq C_{\alpha_i};\quad n \geq 1. \tag{15}$$

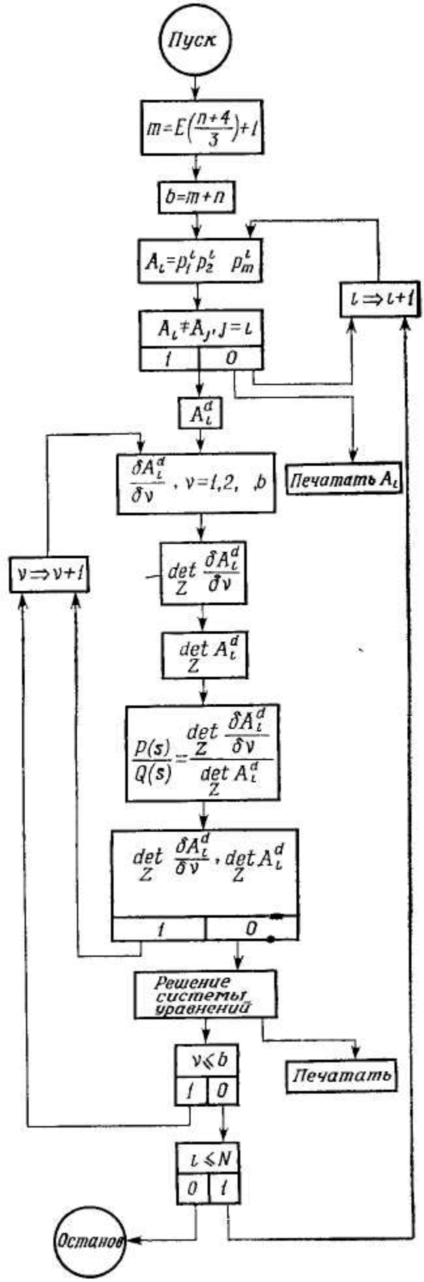
Если при аппроксимации были нормированы шкалы частот и импедансов, то для первой итерации вполне приемлемы значения (13).
В результате решения приведенных уравнений получим значения элементов цепи, соответствующие блочным группам A_1, A_2, …., A_i, ... Можно заметить, что число неизвестных в системе (11) больше, чем число уравнений, на величину

$$\delta = 3E\left(\frac{n+4}{3}\right) - n - 1 \tag{16}$$

В связи с этим можно задаться значениями б элементов или принять их равными нулю.
6. В случае необходимости повторяем весь цикл расчета для

$$m = E\left(\frac{n+4}{3}\right) + 2,\quad m = E\left(\frac{n+4}{3}\right) + 3\quad \textit{и т.д.}$$

Очевидно, что в этих случаях число элементов будет больше, чем для $k = 0$. Следует подчеркнуть, что в соответствии с изложенными ранее способами идентификации слабосвязных структур при расчетах необходимо сразу же исключить все неправильные структуры.
Ниже дана упрощенная блок-схема программы синтеза двухполюсника приведенным методом.


Пуск
Печатать A_i
Решение системы уравнений
Печатать
Останов

5.2. Синтез пассивного RLC-четырехполюсника

5.2.1. Предварительные сведения

Пусть дан пассивный RLC-четырехполюсник (рис. 2).

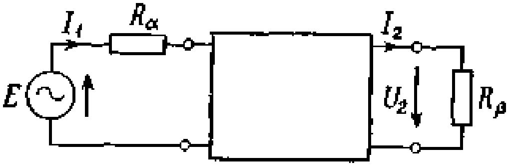


Рис. 2. Пассивный RLC-четырехполюсник.

Характеристики четырехполюсника обычно определяются с помощью передаточных функций или затухания

$$K_u = \frac{U_2}{E}, \quad K_i = \frac{I_2}{I_1}, \quad K_s = \frac{2U_2}{E}\sqrt{\frac{R_\beta}{R_\alpha}}, \quad \Gamma_s = \ln\frac{1}{K_s}$$

В методе блочных групп передача напряжения K_u, тока K_i, передача K_s, а также рабочее затухание Γ_s выражаются в следующем виде:

$$K_u = \frac{\underset{z}{\mathrm{Sim}}\left(\frac{\partial A^d}{\partial\alpha}, \frac{\partial A^d}{\partial\beta}\right)}{\underset{z}{\det} A^d} z_\beta, \qquad K_i = \frac{\underset{z}{\mathrm{Sim}}\left(\frac{\partial A^d}{\partial\alpha}, \frac{\partial A^d}{\partial\beta}\right)}{\underset{z}{\det} \frac{\partial A^d}{\partial\alpha}},$$

$$K_s = \frac{\underset{z}{\mathrm{Sim}}\left(\frac{\partial A^d}{\partial\alpha}, \frac{\partial A^d}{\partial\beta}\right)}{\frac{1}{2}\underset{z}{\det} A^d} \sqrt{R_\alpha R_\beta}, \qquad \Gamma_s = \ln\frac{1}{K_s}, \tag{18}$$

где A — блочная группа, геометрическим изображением которой служит граф цепи.

Рассмотрим функцию пассивного четырехполюсника (из условий физической реализации четырехполюсника следует, что знаменатель функции K_s представляет собой полином Гурвица, а числитель — произвольный полином комплексной частоты $s = \sigma + j\omega$ с вещественными коэффициентами.)

$$K_s = \frac{d_{\bar m} s^{\bar m} + d_{\bar m - 1} s^{\bar m - 1} + \ldots + d_1 s + d_0}{c_{\bar n} s^{\bar n} + c_{\bar n - 1} s^{\bar n - 1} + \ldots + c_1 s + c_0}, \qquad \bar m \leqslant \bar n. \tag{19}$$

Умножая числитель и знаменатель этой функции на s^{-1}, получим

$$K_s = \frac{d_{\overline{m}} s^{\overline{m}-l} + \ldots + d_0 s^{-l}}{c_{\overline{n}} s^{\overline{n}-l} + \ldots + c_0 s^{l}} = \frac{b_{n-1} s^{n-1} + \ldots + b_{-(n-1)} s^{-(n-1)}}{a_n s^n + \ldots + a_{-n} s^{-n}}. \quad (20)$$

Выражение (20) будем называть стандартной передачей. Всегда справедливы следующие условия:

$$c_{\overline{n}} \neq 0 \quad и \quad d_{\overline{m}} \neq 0$$

Кроме того, допустим, что

$$d_0 = d_1 = \ldots = d_{\upsilon-1} = 0; \qquad d_\upsilon \neq 0. \quad (21)$$

Докажем следующую вспомогательную теорему.

Теорема 1. Граф, реализующий заданную передачу, степень числителя которой равна *m,* а знаменателя *n,* имеет цикломатическое число *M* не меньшее, чем

$$M_{\text{мин}} = \begin{cases} E\left(\dfrac{\overline{n}}{2}\right)+1 & \text{для } \overline{m} < \overline{n}, \\ E\left(\dfrac{\overline{n}+1}{2}\right)+1 & \text{для } \overline{m} = \overline{n} \end{cases} \quad (22)$$

Доказательство. Так как в выражение стандартной передачи всегда можно ввести дополнительные слагаемые с коэффициентами, равными нулю, то должны выполняться следующие неравенства:

$n - l \geq \overline{m} - l, \quad n - l \geq l - \upsilon$ (из выражения для числителя),

$n \geq \overline{n} - l, \quad n \geq l$ (из выражения для знаменателя).

Таким образом, должно выполняться неравенство

$$n \geq \text{Max}\{\overline{m} - l + 1;\ l - \upsilon + 1;\ \overline{n} - l;\ l\}. \quad (23)$$

Могут иметь место два частных случая:

а) $\overline{m} = \overline{n}$. Тогда

$$n_{\text{мин}} = \min_{l} \text{Max}\{\overline{n} - l + l,\ l - \upsilon + 1,\ l\}.$$

В этом случае выражение (23) будет минимальным для $\upsilon = 0$ при $\overline{n} - l + 1 = l + 1$, а для $\upsilon > 0$ — при $\overline{n} - l + 1 = l$. Обозначив l_0 значение l, при котором $n = \min$, получим

$$l_0 = \frac{\overline{n}}{2} \qquad для\ \vartheta = 0,$$

$$l_0 = \frac{\overline{n}+1}{2} \quad для\ \vartheta > 0$$

Очевидно, l_0 должно быть целым числом. Можно заметить, что

$$l_0 = E\left(\frac{\overline{n}}{2}\right) \quad и \quad l_0 = E\left(\frac{\overline{n}+1}{2}\right)$$

приводят к одинаковой величине

$$n_{мин} = E\left(\frac{\overline{n}+1}{2}\right)+1.$$

б) $\overline{m} < \overline{n}$.

Тогда

$$n_{мин} = \min_l \operatorname{Max}\{l — \upsilon + 1,\ \overline{n} — l,\ l\}.$$

В этом случае выражение (23) для $\upsilon = 0$ будет минимальным при $l + 1 = \overline{n} — 1$, а для $\upsilon > 0$ — при $l = \overline{n} — 1$.

Следовательно, имеем

$$l_0 = \frac{\overline{n}-1}{2} \quad для\ \vartheta = 0,$$

$$l_0 = \frac{\overline{n}}{2} \quad для\ \vartheta > 0$$

Легко заметить, что как для четных, так и нечетных $\overline{n}$ минимальное значение n равно

$$n_{мин} = E\left(\frac{\overline{n}}{2}\right)+1.$$

Так как степень знаменателя приведенной к стандартному виду передачи должна быть равна числу множителей в каждом из слагаемых $\det\limits_{Z} A^d$, а с другой стороны, это число равно числу m простых

множителей, на которые разлагается A^d

$$A^d = P_1P_2 \ldots P_m,$$

то на основе приведенной ранее теоремы заключаем, что n равно цикломатическому числу графа, реализующего заданную передачу, т. е.

$n = M$ и

$$M_{мин} = \begin{cases} E\left(\frac{\overline{n}}{2}\right)+1 & для\ \overline{m} < \overline{n}, \\ E\left(\frac{\overline{n}+1}{2}\right)+1 & для\ \overline{m} = \overline{n} \end{cases}$$

Заметим, что M имеет минимум при $l = E\left(\dfrac{\overline{n}}{2}\right)$. Доказанная теорема может быть использована при синтезе четырехполюсников методом блочных групп, так как она позволяет определить класс графов, реализующих заданную передачу четырехполюсника.

Теорема 2. В произвольном пассивном BLC-четырехполюснике число X_L индуктивностей и число X_R резисторов удовлетворяют следующим неравенствам:

$$2X_L + X_R \geq \overline{n} + 1$$

или (24)

$$2X_L + X_R \geq \overline{n},$$

где $\overline{n}$ — степень знаменателя передаточной функции четырехполюсника.

Первое неравенство относится к случаю, когда полином в числителе функции содержит свободный член (υ = 0), второе — когда в числителе отсутствует свободный член (υ > 0).

Доказательство. Допустим, что передаточная функция записана в виде (19) и пусть справедливо равенство (21). Умножая числитель и знаменатель выражения (19) на множитель s^{-l}, приведем его к стандартному виду

$$K_s = \frac{0\cdot s^{n-1} + 0\cdot s^{n-2} + \ldots + b_\nu s + \ldots + b_0 + \ldots + b_{-(n-1)} s^{-(n-1)}}{0\cdot s^{n} + 0\cdot s^{n-1} + \ldots + a_\mu s + \ldots + a_0 + \ldots + a_{-n} s^{-n}}, \tag{25}$$

где b_ν и a_μ — ненулевые коэффициенты при наибольших степенях числителя и знаменателя K_S: имеющие следующий вид:

$$a_\mu = \sum_k \prod_{i=1}^{M_k} R_{\alpha_i k} \prod_{i=1}^{\mu} L_{\alpha_i k} \prod_{j=1}^{N_k} L_{\alpha_j k} C^{-1}_{\alpha_j k},$$
$$b_\nu = \sum \prod_{i=1}^{M_k^*} R_{\beta_i k} \prod_{i=1}^{\nu} L_{\beta_i k} \prod_{j=1}^{N_k^*} L_{\beta_j k} C^{-1}_{\beta_j k}, \tag{26}$$

где

$$M_k + \mu + 2N_k = n,$$
$$M^*_k + \nu + 2N^*_k = n - 1. \tag{27}$$

В выражении (26) мы приняли, что μ, ν > 0. Можно показать, что случай, когда μ, ν < 0, не изменяет условия теоремы. Из доказательства теоремы 1 следует, что

$$n \geq \mathrm{Max}\{\overline{m} - l - 1;\ l - \upsilon + 1;\ \overline{n} - l;\ l\},\ \overline{m} \leq \overline{n}. \tag{28}$$

Рассмотрим случай $\overline{m} < \overline{n}$. Тогда для n можно написать

а) при υ = 0

$$n \geq \begin{cases} l+1 & для\ l+1 \geq \bar{n}-l, \\ \bar{n}-1 & для\ l+1 < \bar{n}-l \end{cases} \qquad (29а)$$

откуда

$$n\text{-}l=1, 2, 3, ...;$$

б) при $\upsilon > 0$

$$n \geq \begin{cases} l & для\ l \geq \bar{n}-l, \\ \bar{n}-l & для\ l < \bar{n}-l, \end{cases} \qquad (29б)$$

откуда

$$n — l = 0, 1, 2, ...$$

Обозначим

$$\mu + N_k = X_{Lk}, \quad \nu + N^*_k = X^*_{Lk}. \qquad (30)$$

Подставляя (30) в (27), получим

$$2X_{Lk} + M_k = n + \mu, \qquad 2X^*_{Lk} + M^*_k = n — 1 + \nu. \qquad (31)$$

Так как

$$\mu = \bar{n} — l, \qquad \nu = \bar{m} — l,$$

то можно написать

$$2X_{Lk} + M_k = \bar{n} + (n\text{-}l),$$
$$2X^*_{Lk} + M^*_k = \bar{m} + (n\text{-}l)\text{-}1. \qquad (32)$$

Принимая во внимание выражения (29а) и (29б), запишем

а) для $\upsilon=0$, $\bar{m} < \bar{n}$

$$2X_{Lk} + M_k \geq \bar{n} + 1, \qquad 2X^*_{Lk} + M^*_k \geq \bar{m}; \qquad (33)$$

б) для $\upsilon>0$, $\bar{m} < \bar{n}$

$$2X_{Lk} + M_k \geq \bar{n}, \qquad 2X^*_{Lk} + M^*_k \geq \bar{m} - 1. \qquad (34)$$

Если теперь X_L — число индуктивностей, а X_R — число резисторов четырехполюсника, то легко заметить, что

$$X_{Lk}, X^*_{Lk} \leq X_L, \qquad M_k, M^*_k \leq X_R. \qquad (35)$$

Учитывая это в выражениях (33) и (34). получим

а) для $\upsilon=0$, $\bar{m} < \bar{n}$

$$2X_L + X_R \geq \bar{n} + 1, \qquad 2X_L + X_R \geq \bar{m}; \qquad (33а)$$

б) для $\upsilon>0$, $\bar{m} < \bar{n}$

$$2X_L + X_R \geq \bar{n}, \qquad 2X_L + X_R \geq \bar{m} - 1. \qquad (34а)$$

Из неравенств (33а) и (34а) окончательно следует, что для $\bar{m} < \bar{n}$

$$2X_L + X_R \geq \bar{n} + 1 \qquad для \quad \upsilon = 0,$$
$$2X_L + X_R \geq \bar{n} \qquad для \quad \upsilon > 0. \qquad (36)$$

Рассмотрим случай $\bar{m} = \bar{n}$. При этом

$$n \geq \mathrm{Max}\{\bar{n} — l + 1, l — \upsilon + 1, \bar{n} — l, l\},$$

и можно написать

а) при υ=0

$$n \geq \begin{cases} l+1 & \text{для } l+1 \geq \overline{n}-l+1, \\ \overline{n}-l+1 & \text{для } l+1 < \overline{n}-l+1, \end{cases}$$

откуда

$$n — l = 1, 2, 3, \ldots;$$

б) при υ > 0

$$n \geq \begin{cases} l & \text{для } l \geq \overline{n}-l+1, \\ \overline{n}-l+1 & \text{для } l < \overline{n}-l+1, \end{cases}$$

откуда

$$n — l = 0, 1, 2, \ldots .$$

Для определения разности $n — l$ имеем условия, аналогичные случаю $\overline{m} < \overline{n}$. Учитывая их в соотношении (32), получим

$$\left.\begin{aligned} 2X_L + X_R &\geq \overline{n}+1 \\ 2X_L + X_R &\geq \overline{m} \end{aligned}\right\} \quad \text{для} \quad \vartheta = 0,$$

а также

$$\left.\begin{aligned} 2X_L + X_R &\geq \overline{n} \\ 2X_L + X_R &\geq \overline{m}-1 \end{aligned}\right\} \quad \text{для} \quad \vartheta > 0.$$

Последние неравенства также приводят к формулам (36), доказанным ранее для $\overline{m} < \overline{n}$.

Из теоремы 2 вытекают следствия.

Следствие 1. Если в пассивном RLC-четырехполюснике число резисторов равно числу индуктивностей, т. е. $X_R = X_L$, то

$$X_L \geq \begin{cases} \dfrac{\overline{n}+1}{3}, & \vartheta = 0, \\ \dfrac{\overline{n}}{3}, & \vartheta > 0. \end{cases} \tag{37}$$

Следствие 2. Если в пассивном RLC-четырехполюснике, нагруженном активными сопротивлениями R_α, R_β, последовательно с каждой индуктивностью включен резистор, т. е. если $X_R = X_L + 2$, то

$$X_L \geq \begin{cases} \dfrac{\overline{n}-1}{3}, & \vartheta = 0, \\ \dfrac{\overline{n}-2}{3}, & \vartheta > 0. \end{cases} \tag{38}$$

Следствие 3. Если принять, что каждый резистор, за исключением R_α, R_β, включен последовательно с индуктивностью, т. е. $X_R = X_L + 2$, то число индуктивностей минимально, если все индуктивности содержатся в одном дополнении дерева с резисторами R_α, R_β.

Выражение (35) имеет знак равенства, если в выражениях (26) существует такое слагаемое, в котором сгруппированы все индуктивности и резисторы, что имеет место тогда, когда все индуктивности и резисторы содержатся в одном дополнении дерева с резисторами R_α, R_β. Следует подчеркнуть, что только в этом случае выполняется равенство и в выражениях (33а) и (34а), а значит, и достигаются минимальные значения X_L и X_R, например при $X_R = X_L + 2$.

Пример 1. Передачу

$$K_s = \frac{P(S)}{Q(S)},$$

степень числителя которой равна 9 ($\overline{n}$ = 9), можно реализовать при помощи схемы (рис. 3, *а)*.

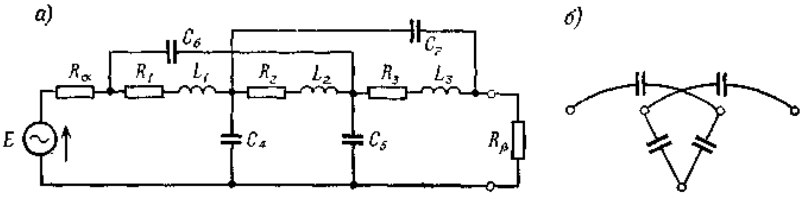


Рис. 3. *а)* пример четырехполюсника с минимальным числом индуктивностей; *б)* противодерево, не содержащее катушек индуктивности.

Эта цепь содержит минимальное число индуктивностей, так как все они содержатся в дополнении дерева с резисторами R_α, R_β.

Действительно, при исключении ветвей, содержащих индуктивности и резисторы R_α, R_β, схема будет иметь структуру (рис. 3, б), т. е. представляет собой дерево. Исключенные ветви служат хордами дерева (рис. 3, *б)* и, следовательно, образуют дополнение дерева с элементами R_α, R_β.

Рассмотренное свойство, сформулированное в следствии 3, может быть использовано при синтезе RLC-четырехполюсника с помощью ЭВМ. Согласно этому свойству, элементы *R, L* можно размещать в одном из выбранных дополнений деревьев, а все остальные ветви четырехполюсника оставить для емкостных элементов. Такой способ размещения элементов в графе легко запрограммировать для ЭВМ.

5.2.2. Определение знаков слагаемых функции совпадения

Знаки функции совпадения

$$\operatorname*{Sim}_{Z}\left(\frac{\partial A^d}{\partial \alpha}, \frac{\partial A^d}{\partial \beta}\right)$$

при анализе схем методом блочных групп без вычислительной машины определяются простым рассмотрением ориентации ветвей α и β в графе. Понятно, что вычислительная машина непосредственно не может определить ориентации ветвей в графе. Поэтому при разработке алгоритма синтеза четырехполюсников с применением ЭВМ важно найти чисто алгебраический метод нахождения знаков слагаемых функции совпадения. Такой метод можно получить следующим образом. Слагаемые функции совпадения, имеющие знак плюс, соответствуют контурам, показанным на рис. 4, *а,* а слагаемые со знаком минус — контурам, показанным на рис. 4, *б.*

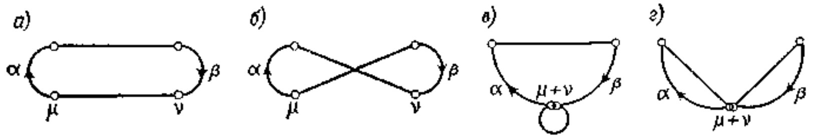


Рис. 4. Иллюстрация алгебраического метода определения знаков функции совпадения.

Короткое замыкание узлов μ и ν в этих контурах приводит к графам (рис. 4, *в* и *г).* Из рис. 4 видно, что в случае *б,* соответствующем знаку плюс, существует контур, содержащий ветви α и β, в то время как в случае *г* такого контура нет. Это служит основой алгебраического метода определения знаков функции совпадения.

Допустим, что блочная группа *А* равна произведению однострочных простых блочных групп

$$A = P_1P_2 \dots P_m.$$

Введем некоторые определения.

Определение 1. Простые блочные группы P_i, P_j будем называть сгруппированными блочными группами, если они содержат по крайней мере один одинаковый элемент, т. е.

$$P_i, P_j \text{ сгруппированная } \underset{\alpha}{\Leftrightarrow} \exists(\alpha \in P_i) \wedge (\alpha \in P_j). \qquad (39)$$

Например, однострочные блочные группы

$$P_i = [1\ 2\ 4\ 7], \qquad \text{а также } P_j = [3\ 4\ 6\ 8]$$

будут сгруппированными блочными группами, так как обе эти блочные группы содержат элемент 4.

Если блочные группы P_i, P_j сгруппированы, то пишем $P_i \leftrightarrow P_j$. Если существует последовательность однострочных блочных групп

$$P_1 \leftrightarrow P_2 \leftrightarrow P_3 \leftrightarrow \ldots \leftrightarrow P_{n-1} \leftrightarrow P_n: \qquad (40)$$

соответствующая набору вершин 1, 2, . . ., n в графе, то тогда существует путь, соединяющий вершину 1 с вершиной n.

Если данный граф (мультиграф) служит геометрическим изображением блочной группы, то блочная группа, соответствующая графу, в котором все ветви пути, соединяющего вершину μ с вершиной ν, замкнуты, обозначим через $A^*_{\mu\nu}$. Получим

$$A = P_1 P_2 \ldots P_{w-1}; \quad A^*_{\mu\nu} = (P_\mu + P_{\mu 1} + P_{\mu 2} + \ldots + P_\nu) \prod_{i=1}^{w-1} P_i, \qquad (41)$$

где w — число вершин (узлов) графа, а P_μ, $P_{\mu 1}$, $P_{\mu 2}, \ldots, P_\nu$ — однострочные сгруппированные блочные группы, соответствующие вершинам μ, μ_1, μ_2, . . ., ν, которые принадлежат пути, соединяющему вершину μ с вершиной ν, т. е.

$$P_\mu \leftrightarrow P_{\mu 1} \leftrightarrow P_{\mu 2} \leftrightarrow \ldots \leftrightarrow P_\nu$$

Сформулируем следующую теорему.

Теорема 3. Если A — блочная группа, геометрическим изображением которой служит граф четырехполюсника со структурой, показанной на рис. 5, а вершины μ и ν инцидентны соответственно элементам α и β (например, $\alpha \in P_\mu$, $\beta \in P_\nu$), то столбцы блочной группі

$$C = \frac{A^{*d}_{\mu\nu}}{\partial\alpha} \cap \frac{A^{*d}_{\mu\nu}}{\partial\beta} \qquad (42)$$

определяют все слагаемые функции совпадения $\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A^d}{\partial\alpha}, \frac{\partial A^d}{\partial\beta}\right)$, которые имеют знак плюс.

Справедливость теоремы 3 непосредственно вытекает из предыдущего рассуждения. Исключение всех ветвей графа, определенных одним из столбцов конъюнкции

$$\frac{A^{*d}_{\mu\nu}}{\partial\alpha} \cap \frac{A^{*d}_{\mu\nu}}{\partial\beta},$$

приводит к графу с одним циклом (контуром) (рис. 4, *а).* Такой контур соответствует положительному члену функции совпадения.

Метод определения знаков функции совпадения основан на теореме 3 и применим для четырехполюсника со структурой, показанной на рис. 5.

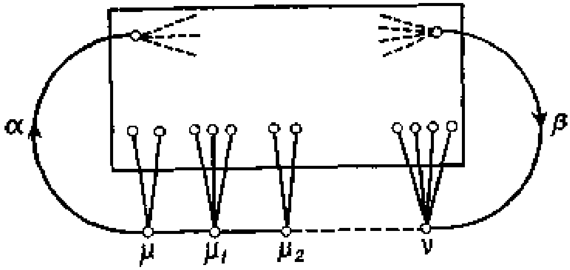


Рис. 5.

Такая структура довольно часто встречается в практике. Однако можно сформулировать более общий алгоритм, позволяющий определять знаки функции совпадения любой схемы. Такой алгоритм приведен ниже.

Блочную группу, соответствующую графу (мультиграфу), в котором закорочены два узла μ и *v*, обозначим $A_{\mu\nu}$. Тогда

$$A = P_1P_2 \ldots P_{w-1}, \quad A_{\mu\nu} = (P_\mu + P_\nu)\prod_{i=1}^{w-1} P_i \quad (i \neq \mu, \nu), \tag{43}$$

где *w* — число вершин графа (узлов). Обозначим $A^{\gamma_1\gamma_2\ldots\gamma_i}$ блочную группу, полученную из блочной группы *А* путем исключения в ней всех столбцов, содержащих любой из элементов γ_1, γ_2,. . ., γ_i. Очевидно, что

$$A^{\gamma_1\gamma_2\ldots\gamma_i} = \frac{\delta^i A}{\delta\gamma_1\delta\gamma_2\ldots\delta\gamma_i}. \tag{44}$$

Теорема 4. Каждый член $z_{\gamma_1}, z_{\gamma_2}, \ldots, z_{\gamma_i}$ функции совпадения $\operatorname*{Sim}_{Z}\left(\frac{\partial A^d}{\partial\alpha}, \frac{\partial A^d}{\partial\beta}\right)$ будет иметь знак плюс, если *z*

$$\frac{\partial A^{\gamma_1\gamma_2\ldots\gamma_i}_{\mu\nu}}{\partial\alpha} \cap \frac{\partial A^{\gamma_1\gamma_2\ldots\gamma_i}_{\mu\nu}}{\partial\beta} \neq 0, \tag{45a}$$

или знак минус, если

$$\frac{\partial A_{\mu\nu}^{\gamma_1\gamma_2\cdots\gamma_i}}{\partial\alpha} \cap \frac{\partial A_{\mu\nu}^{\gamma_1\gamma_2\cdots\gamma_i}}{\partial\beta} = 0. \qquad (45б)$$

Доказательство этой теоремы основано на следующих положениях:

1. После исключения ветвей $\gamma_1, \gamma_2, \ldots, \gamma_i$ в графе, изображающем блочную группу *А,* этот граф переходит в граф, содержащий только один цикл, проходящий через ветви α и β. В зависимости от знака члена $z_{\gamma_1}, z_{\gamma_2}, \ldots, z_{\gamma_i}$ в графе блочной группы $A_{\mu\nu}^{\gamma_1\gamma_2\cdots\gamma_i}$ этот цикл сохраняется или исчезает (рис. 4, *в* и *г).*

2. Для любой блочной группы *В* имеет место следующая формула

$$\left(\frac{\partial B^d}{\partial\alpha} \cap \frac{\partial B^d}{\partial\beta} = 0\right) \Leftrightarrow \left(\frac{\partial B}{\partial\alpha} \cap \frac{\partial B}{\partial\beta} = 0\right).$$

Практическое применение теоремы 4 проиллюстрируем на следующем примере.

Пример 2. Определить знаки функции совпадения $\operatorname*{Sim}_{z}\left(\frac{\partial A^d}{\partial 1}, \frac{\partial A^d}{\partial 2}\right)$ для схемы, приведенной на рис. 6.

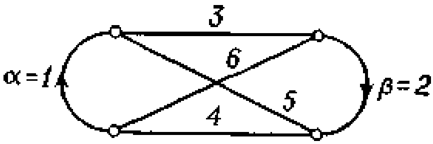


Рис. 6.

Имеем

A = [1 4 6] 12 4 5] [1 3 5],

$A_{\mu\nu}$ = ([1 4 6] + [2 4 51) [1 3 51 = [1 2 5 6] [1 3 5],

откуда

$$\begin{array}{r} 1\;4\;6 \\ \times\, 2\;4\;5 \\ 1\;3\;5 \\ \hline \end{array}$$

$$A = \begin{bmatrix} 1 & 1 & 1 & 1 & 4 & 4 & 4 & 6 & 6 & 6 & 6 & 6 & 6 & 6 & 6 \\ 2 & 2 & 4 & 5 & 2 & 2 & 5 & 2 & 2 & 2 & 4 & 4 & 4 & 5 & 5 \\ 3 & 5 & 3 & 3 & 1 & 3 & 3 & 1 & 3 & 5 & 1 & 3 & 5 & 1 & 3 \end{bmatrix},$$

$$\begin{array}{l} \quad 1\ 2\ 5\ 6 \\ \times \\ \quad 1\ 3\ 5 \\ \hline \end{array}$$

$$A_{\mu\nu} = \begin{bmatrix} 1 & 2 & 2 & 2 & 5 & 6 & 6 & 6 \\ 3 & 1 & 3 & 5 & 3 & 1 & 3 & 5 \end{bmatrix}.$$

После определения дополнительной блочной группы A^d находим

$$\frac{\partial A^d}{\partial 1} \cap \frac{\partial A^d}{\partial 2} = \begin{bmatrix} 5\ 4 \\ 6\ 4 \end{bmatrix}.$$

Далее определяем производные

$$\frac{\partial A^{56}_{\mu\nu}}{\partial 1}, \quad \frac{\partial A^{56}_{\mu\nu}}{\partial 2}, \quad \frac{\partial A^{34}_{\mu\nu}}{\partial 1}, \quad \frac{\partial A^{34}_{\mu\nu}}{\partial 2}.$$

Получаем

$$A^{56}_{\mu\nu} = \begin{bmatrix} 1 & 2 & 2 \\ 3 & 1 & 3 \end{bmatrix}, \qquad A^{34}_{\mu\nu} = \begin{bmatrix} 2 & 2 & 6 & 6 \\ 1 & 5 & 1 & 5 \end{bmatrix}, \qquad \frac{\partial A^{56}_{\mu\nu}}{\partial 1} = [2\ \ 3],$$

$$\frac{\partial A^{56}_{\mu\nu}}{\partial 2} = [1\ \ 3], \quad \frac{\partial A^{34}_{\mu\nu}}{\partial 1} = [2\ \ 6], \quad \frac{\partial A^{34}_{\mu\nu}}{\partial 2} = [1\ \ 5],$$

$$\frac{\partial A^{56}_{\mu\nu}}{\partial 1} \cap \frac{\partial A^{56}_{\mu\nu}}{\partial 2} = [3] \neq 0, \qquad \frac{\partial A^{34}_{\mu\nu}}{\partial 1} \cap \frac{\partial A^{34}_{\mu\nu}}{\partial 2} = 0,$$

а также

$$\operatorname*{Sim}_{z} \frac{\partial A^d}{\partial 1} \cap \frac{\partial A^d}{\partial 2} = z_5 z_6 - z_3 z_4.$$

Теперь, пользуясь формулами (18), можно определить характеристические функции рассматриваемого четырехполюсника, например рабочее затухание:

$$\Gamma_S = \ln \left[\frac{z_4 z_5 z_6 + z_3 z_4 z_6 + z_2 z_5 z_6 + z_2 z_3 z_6 + z_3 z_5 z_6 + \ldots + z_1 z_3 z_4}{2 (z_5 z_6 - z_3 z_4) \sqrt{z_1 z_2}} \right].$$

Очевидно, что при анализе схемы знаки функции совпадения можно легко определить путем рассмотрения ориентации ребер α и β в цепи. Однако при синтезе схемы при помощи ЭВМ такой подход не применим.

5.2.3. Синтез четырехполюсника средствами ЭВМ

Метод синтеза RLC-четырехполюсника при помощи ЭВМ поясним следующим примером.

Пример 3. Пусть дан четырехполюсник (рис. 7).

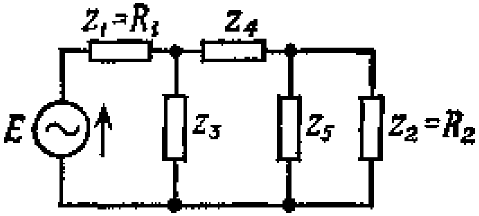


Рис. 7.

Его элементы определим таким образом, чтобы передаточная функция комплексной частоты s = σ + *j*ω определялась выражением

$$K_s = \frac{1}{s^3 + 2s^2 + 2s + 1}. \tag{46}$$

На основе приведенных формул имеем

$$K_s = \frac{\underset{Z}{\mathrm{Sim}} \left(\frac{\partial A^d}{\partial \alpha}, \ \frac{\partial A^d}{\partial \beta} \right)}{\frac{1}{2} \underset{Z}{\det} A^d} \sqrt{R_1 R_2}.$$

После простых расчетов получаем

$$\underset{Z}{\det} A^d = z_1 z_2 z_3 + z_1 z_3 z_5 + z_1 z_2 z_4 + z_1 z_4 z_5 + z_1 z_2 z_5 + z_2 z_3 z_4 + z_3 z_4 z_5 + z_2 z_3 z_5,$$

а также

$$\underset{Z}{\mathrm{Sim}} \frac{\partial A^d}{\partial \alpha} \mid \frac{\partial A^d}{\partial \beta} = z_3 z_5$$

Умножая числитель и знаменатель формулы (46) на s^{-2} и дополняя их нулевыми членами таким образом, чтобы передаточная функция стала симметричной относительно степеней комплексной частоты *s*, получим

$$K_s = \frac{0 \cdot s^2 + 0 \cdot s + 0 \cdot s^{-1} + 0 \cdot s^{-2}}{0 \cdot s^3 + 0 \cdot s^2 + s + 2 + 2 \cdot s^{-1} + s^{-2} + 0 \cdot s^{-3}}.$$

Если принять

$z_i = R_i$, $z_2 = R_2$, $z_3 = \frac{1}{sC_3}$, $z_4 = sL_4$, $z_5 = \frac{1}{sC_5}$,

то

$$\underset{Z}{\det} A^d = R_1 R_2 C_3^{-1} s^{-1} + R_1 C_3^{-1} s^{-2} + R_1 R_2 L_4 s +$$
$$+ R_1 L_4 C_5^{-1} + R_1 R_2 C_5^{-1} s^{-1} + R_2 C_3^{-1} L_4 + C_3^{-1} L_4 C_5^{-1} s^{-1} + R_2 C_3^{-1} C_5^{-1} s^{-2}.$$

Функция совпадения равна

$$\operatorname*{Sim}_{Z}\left(\frac{\partial A^d}{\partial\alpha},\ \frac{\partial A^d}{\partial\beta}\right)\Big|_{\substack{\alpha=1\\ \beta=2}} = z_3 z_5 = C_3^{-1} C_5^{-1} s^{-2}.$$

С другой стороны,

$$\frac{1}{2}\det_{Z} A^d = 0\cdot s^3 + 0\cdot s^2 + s + 2 + 2s^{-1} + s^{-2} + 0\cdot s^{-3},$$

$$\operatorname*{Sim}_{Z}\left(\frac{\partial A^d}{\partial\beta},\ \frac{\partial A^d}{\partial\beta}\right) = 0\cdot s^2 + 0\cdot s + 1 + 0\cdot s^{-1} + s^{-2},$$

откуда для определения значений неизвестных величин элементов четырехполюсника получаем уравнения

$R_1R_2R_4=2$,

$R_1L_4C^{-1}_5+ R_2L_4C^{-1}_3=4$,

$R_1R_2\, C^{-1}_3+ R_1R_2\, C^{-1}_5+ L_4\, C^{-1}_3\, C^{-1}_5== 4$,

$R_1\, C^{-1}_3\, C^{-1}_5+ R_2C^{-1}_3\, C^{-1}_5=2$.

$C^{-1}_3\, C^{-1}_5=1$.

Решение последних уравнений дает

$R_1= R_2 = 1$ *ом*, $C_3 = C_5 = 1$ *ф*, $L_4 = 2$ *гн*.

Таким образом, мы определили схему с передаточной функцией (46) и структурой, показанной на рис. 7.

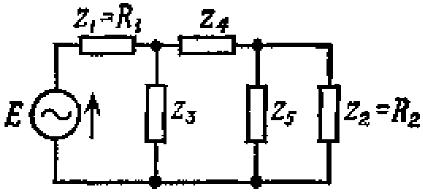


Рис. 7.

Этот пример поясняет способ синтеза четырехполюсника методом блочных групп. Далее проведем обобщение этого метода. С этой целью рассмотрим передачу стандартного вида (20)

$$K_s = \frac{b_{n-1}s^{n-1} + \ldots + b_{-(n-1)}s^{-(n-1)}}{a_n s^n + \ldots + a_{-n}s^{-n}} = \frac{P(s)}{Q(s)}.$$

Очевидно, некоторые коэффициенты a_i, а также b_i функции K_s могут быть равными нулю, как это было показано в последнем примере. Число нулевых коэффициентов можно определить с помощью следующей формулы:

$$\Theta = \begin{cases} 2+2E\left(\dfrac{\overline{n}}{2}\right)-\overline{m} & \text{для четного } \overline{n}, \\ 1+2E\left(\dfrac{\overline{n}}{2}\right)-\overline{m} & \text{для нечетного } \overline{n}, \end{cases} \tag{47}$$

где $\overline{m}$ и $\overline{n}$ — числа, определяющие порядок числителя и знаменателя функции (19). В этом рассуждении мы приняли допущение, что все коэффициенты c_i и d_i функции (19) отличны от нуля.

Допустим, что каждая (внутренняя) ветвь четырехполюсника представляет собой последовательное соединение резистора, индуктивности и конденсатора. Тогда импеданс z_i произвольной внутренней ветви четырехполюсника запишется в виде

$$z_i = sL_i + R_i + s^{-1}C_i^{-1} \qquad (i \neq \alpha,\ i \neq \beta). \tag{48}$$

Для ветвей α и β имеем

$$z_\alpha = R_1,\ \ z_\beta = R_2,$$

где величины резисторов R_1 и R_2 заранее известны. Введем следующие обозначения: b — число ветвей графа, w — число вершин графа, M — цикломатическое число графа, равное количеству линейно независимых контуров.

В соответствии с формулой Эйлера для произвольного графа или мультиграфа

$$b - w + 1 = M. \tag{49}$$

При этом заметим, что порядок функции $\det_z A^d$ равен n. В результате

тате получаем

$$M = n \quad \text{и} \quad b - w + 1 = n. \tag{50}$$

Так как число уравнений, которые можно составить для определения элементов схемы, равно количеству коэффициентов рациональной функции (20), т. е.

$$(2n + 1) + (2n - 1) = 4n,$$

а число неизвестных элементов равно $3(b - 2)$, то должно выполняться следующее неравенство:

$$3\,(b - 2) \geq 4n. \tag{51}$$

Учитывая это неравенство в равенствах (50), получим формулу для оценки числа вершин графа:

$$w \geq (n/3) + 3. \tag{52}$$

Так как количество простых множителей блочной группы A равно $m = w - 1$, имеем

$$m = E(n/3) + 3 + k; \qquad k = 0, 1, 2, \ \ldots, \tag{53}$$

где $E\,(x)$ обозначает целую часть от x.

Число ветвей b графа может быть определено с помощью выражения

$$b = E(n/3) + 3 + n + k; \qquad k = 0, 1, 2, \ ..., \tag{54}$$

Если в последней формуле принять $k = 0$, то получим минимальное количество ветвей графа, которое зависит от числа *n.* Вводя такое допущение, оптимизируем схему по числу ветвей.

Самые общие приемы, используемые при синтезе четырехполюсника на ЭВМ с помощью описанного метода, можно сформулировать в следующем виде:

1. Принимаем, что число простых множителей блочной группы A равно $n = E(n/3) + 3$, и рассматриваем множество ветвей

$$B = \{1, 2, \ldots, b\}; \quad b = m + n,$$

на основании которого (в соответствии с определенной программой ЭВМ) находим все возможные блочные группы

$$A_i = P^i{}_1 P^i{}_2 \ldots . \ P^i{}_m; \quad i = 1, 2, \ ..., \tag{55}$$

удовлетворяющие условиям реализации (2) и соответствующие сильносвязным графам. К тому же надо принять, что каждая однострочная блочная группа $P^i{}_j$ содержит по крайней мере два элемента, например

$$A_i = [1\ 2]\ [1\ 3]\ [2\ 3\ 4] \quad \text{и т. д.}$$

Таким образом, получаем блочные группы $A_1, A_2, \ . \ . \ ., A_i, \ . \ . \ .$, определяющие алгебраическим путем различные структуры графа.

2. Рассчитываем дополнительные блочные группы

$$A^d{}_1, A^d{}_2, \ldots, A^d{}_i, \ldots,$$

3. Находим детерминантные функции

$$\det_Z A^d{}_1, \ \det_Z A^d{}_2, \ \ldots, \det_Z A^d{}_i, \ \ldots,$$

4. Ищем алгебраические производные

$$\frac{\partial A^d{}_i}{\partial 1}, \frac{\partial A^d{}_i}{\partial 2}, \ldots, \frac{\partial A^d{}_i}{\partial b}; \quad i = 1, 2, \ldots$$

5. Исследуем все заземленные четырехполюсники, которые соответствуют блочным группам A_1, $A_2, \ldots, A_i, \ \ldots$ (рис. 8).

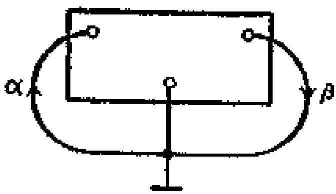


Рис. 8. Четырехполюсник с тремя узлами.

Структуру получаем путем поочередного выбора входной ветви α и выходной ветви β таким образом, чтобы α и β принадлежали одной и той же простой блочной группе P^i_j произведения (55). В результате такой операции получаем функцию совпадения

$$\operatorname*{Sim}_Z\left(\frac{\partial A^d}{\partial \alpha_k}, \frac{\partial A^d}{\partial \beta_k}\right)$$

для которой α_k и β_k принадлежат одному сомножителю P^i_j

$$\alpha_k, \beta_k \in P^i_j \qquad t, j, k=1, 2, \ldots .$$

Расчет такой функции совпадения прост в случае синтеза трехполюсника, так как все члены этой функции положительны.

6. Рассчитываем значения элементов четырехполюсника аналогично расчету элементов двухполюсника. Из сравнения коэффициентов рациональных функций

$$\frac{P(s)}{Q(s)} = \frac{\operatorname*{Sim}_Z\left(\frac{\partial A^d}{\partial \alpha_k}, \frac{\partial A^d}{\partial \beta_k}\right)}{\operatorname*{det}_Z A^d} z_\beta \tag{56}$$

получаем следующую систему нелинейных уравнений, которые можно решить средствами ЭВМ

$$\begin{gathered} f_{n-1}(R_1, \ldots, L_1, \ldots, C_1, \ldots) = b_{n-1}, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ f_{-(n-1)}(R_1, \ldots, L_1, \ldots, C_1, \ldots) = b_{-(n-1)}, \\ \varphi_n(R_1, \ldots, L_1, \ldots, C_1, \ldots) = a_n, \\ \ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots \\ \varphi_{-n}(R_1, \ldots, L_1, \ldots, C_1, \ldots) = a_{-n}. \end{gathered} \tag{57}$$

Функции f_k и φ_k будут суммами произведений

$$R_{\alpha_1} \ldots R_{\alpha_m} L_{\alpha_1} \ldots L_{\alpha_n} C_{\alpha_1} \ldots C_{\alpha_l} . \tag{58}$$

Метод итерации решения этих уравнений средствами ЭВМ приводится ниже. Он основан на том, что входными величинами служат неизвестные $R^0{}_{\alpha_i}, L^0{}_{\alpha_i}, C^0{}_{\alpha_i}$, для которых, например, можно принять значения

$$R^0{}_{\alpha_i} = 1 \text{ ом}, \quad L^0{}_{\alpha_i} = 1 \text{ гн}, \quad C^0{}_{\alpha_i} = 1 \text{ ф}, \quad i = 1, 2, \ \ldots \tag{59}$$

Средствами ЭВМ исследуем разности

$$\begin{gathered} f^0{}_\nu - b^0{}_\nu = \delta^0{}_\nu, \qquad \nu = -(n-1), \ \ldots, (n-1), \\ \varphi^0{}_\nu - a_\nu = \delta^0{}_\nu, \qquad \nu = -n, \ \ldots, n. \end{gathered} \tag{60}$$

Итерационный расчет проводим таким образом, чтобы уменьшить значения Max | $\delta^{\mu}{}_{v}$ |, а также Max | $\delta^{\mu}{}_{v}$ |. Тогда получаем сходящийся ряд итераций, приближенно определяющий величины неизвестных $R_{\alpha_i}, L_{\alpha_i}, C_{\alpha_i}$

$$R^n{}_{\alpha_i} \simeq R_{\alpha_i}; \quad L^n{}_{\alpha_i} \simeq L_{\alpha_i}; \quad C^n{}_{\alpha_i} \simeq C_{\alpha_i}; \quad n \geq 1. \tag{61}$$

Если при аппроксимации было введено нормирование частоты к полосе $w \in$ [0 11 и нормирование сопротивлений к $R_1 = 1$ *ом,* то тогда полностью оправданы принятые ранее значения входных величин (59).

7. В случае необходимости исследуем также структуры с большим числом узлов, соответствующие блочным группам $A_1, A_1,\ldots, A_q$.

Эти структуры можно получить путем поочередного выбора в множестве

$$B = \{1, 2, 3, \ldots, b\}$$

двух таких элементов α_k и β_k, что α_k и β_k принадлежат двум различным блочным группам $P^i{}_j$ произведения (55), т. е.

$$\alpha_k \in P^i{}_j, \ \beta_k \in P^i{}_j; \quad i, j, k, l = 1, 2, \ \ldots \quad (j \neq l).$$

В этом случае функция

$$\operatorname*{Sim}_{Z}\left(\frac{\partial A^d}{\partial \alpha_k}, \frac{\partial A^d}{\partial \beta_k}\right)$$

может содержать отрицательные слагаемые. Положительные и отрицательные слагаемые можно различить с помощью метода, описанного ранее.

8. Рассчитываем значения элементов в случае четырехполюсника произвольной структуры аналогично п. 6.

Представленный способ синтеза четырехполюсника средствами ЭВМ позволяет учесть различные критерии оптимизации схемы, что служит одним из достоинств описываемого метода. Если, например, принять $k > 0$, а значит, $k = 1, 2, 3, 4$ и т. д., то получим другие значения элементов схемы, благодаря чему можно ввести новые связи, определяющие, например, равенство всех индуктивностей или конденсаторов, заданный разброс параметров составных элементов четырехполюсника, собственные потери индуктивностей и конденсаторов, минимальную суммарную индуктивность схемы и т. д.

Часто удобно принять

$$l > E\left(\frac{\overline{n}}{2}\right),$$

Тогда получим граф (мультиграф) с минимальным цикломатическим числом. Благодаря этому допущению можно создать новые, часто более удобные структуры графов. Можно также рассмотреть и те схемы, которые в результате проведенного синтеза имеют некоторые отрицательные элементы. Такая схема может быть выполнена только при использовании активных элементов.

Достоинства представленного метода синтеза четырехполюсника средствами ЭВМ можно кратко сформулировать следующим образом:

1. Общность метода, заключающаяся прежде всего в том, что он не накладывает никаких ограничений на структуру схемы.
2. Возможность определения множества схем, реализующих поставленную задачу синтеза.
3. Возможность учета разных критериев оптимизации схемы (например, минимального цикломатического числа, минимального числа ветвей, равенства всех индуктивностей, собственной добротности элементов, учет разброса параметров элементов и т. д.).

Данный метод дает также возможность для дальнейших обобщений. Например, с помощью этого метода можно упростить процесс синтеза сложной цепи путем ее разбиения на отдельные подсхемы. С этой целью можно использовать результаты, представленные ранее. Эта проблема, однако, требует дальнейших исследований и здесь не рассматривается.

Остановимся на проблеме исключения цепей, графы которых не имеют сильносвязной структуры. Сильносвязные (или собственные) графы были рассмотрены ранее. С помощью метода блочных групп граф такого типа можно определить, пользуясь известным свойством. Согласно этому свойству, граф имеет сильносвязную структуру тогда и только тогда, когда выполняется следующее условие:

$$\underset{\alpha\in A}{\exists}\ \underset{\beta\in A}{\forall}\left[\frac{\partial A}{\partial\alpha}\mid\frac{\partial A}{\partial\beta}\right]\neq 0$$

где A — блочная группа, для которой рассматриваемый граф служит геометрическим изображением.

На рис. 9 дан пример четырехполюсников, структуры которых не сильносвязны и которые должны быть исключены в процессе синтеза.

Рис. 9. Четырехполюсники не сильносвязной структуры.

Нетрудно проверить, что графы (рис. 9) не удовлетворяют условию (62). Очевидно, что это условие может быть легко запрограммировано на ЭВМ. Поэтому уже в процессе поиска блочных групп можно исключить ненужные графы.

Процесс синтеза четырехполюсника можно значительно упростить, если некоторые операции, предусмотренные представленным алгоритмом, будут выполнены не машиной, а человеком. Так, например, если определить схему четырехполюсника, сразу ориентируясь на определенную структуру, например лестничную или мостовую, то процесс машинного синтеза значшельно упрощается, так как не нужно находить блочные группы и исключать ненужные графы.

Такой упрощенный полуавтоматический метод синтеза может иметь большое применение, например, для проектирования электрических фильтров. Ниже приведены результаты расчетов трех простых RLC-фильтров, спроектированных при помощи ЭВМ. Синтез этих фильтров был приведен упрощенным методом, в котором разработчик сразу определил структуру проектируемой схемы.

При разработке программы для ЭВМ было принято допущение, что в каждой ветви схемы находится только один элемент, благодаря чему получено некоторое упрощение по сравнению с общим методом синтеза RLC-четырехполюсников. Для получения минимального числа индуктивностей фильтров использовано свойство, описанное ранее. Произведение простых блочных групп также было рассчитано непосредственно без помощи машины. Естественно, что при более сложной схеме можно запрограммировать операцию умножения блочных групп цифровой машиной, и тогда расчет фильтра сводится к очень простой операции, так как почти весь процесс проектирования будет выполнен машиной.

Пример 4. Фильтр с равноволновой характеристикой, передаточная функция которого имеет седьмую степень:

$$K = \frac{s^2+4}{7,9168s^5+14,159s^4+22,782s^3+20,365s^2+12,754s+4}. \quad (63)$$

Расчет фильтра был проведен на ЭВМ. В результате расчетов получена схема фильтра (рис. 10).

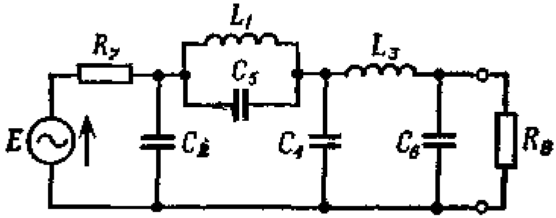


Рис. 10. Реактивный фильтр, передаточная функция которого определяется по формуле (63).

Средствами ЭВМ в результате 10 итераций определены следующие значения элементов:
L_1 = 1,154722 *гн,*
L_3 = 1,675625 *гн,*
C_2 = 1,112944 *ф,*
C_4 = 0,615818 *ф,*
C_5 = 4,6188886 *ф,*
C_6 = 1,11286 *ф,*
R_7 = 1 *ом,*
R_8 = 1,13542 *ом.*

Пример 5. Фильтр с оптимально монотонной характеристикой (Папулиса) с передаточной функцией пятого порядка

$$K_s = \frac{0,224}{s^5 + 1,551s^4 + 2,203s^3 + 1,693s^2 + 0,898s + 0,224} \cdot \quad (64)$$

Расчет этого фильтра также был проведен на ЭВМ. Значения элементов получены в результате 5 итераций. Схема фильтра показана на рис. 11.

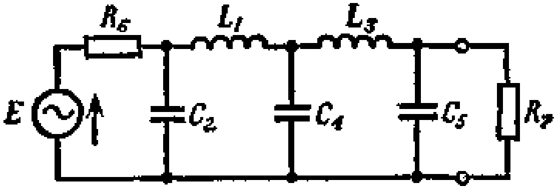


Рис. 11. Реактивный фильтр, передаточная функция которого определяется по формуле (64).

Значения элементов фильтра равны
$L_1 = L_3$ = 1,466371 *гн,*
$C_2 = C_5$ = 0,7479164 *ф,*
C_4 = 0,4263531 *ф,*

$R_6 = R_7 = 0{,}9644312$ *ом.*

Пример 6. Фильтр на индуктивностях с потерями, с максимально плоской характеристикой (Баттерворта), передаточная функция пятого порядка которого имеет вид

$$K_s = \frac{0{,}914166}{s^5+3{,}2360s^4+5{,}236068s^3+5{,}236068s^2+3{,}236068s+1}.$$

Расчет этого фильтра был проведен в два этапа. Вначале был рассчитан фильтр без потерь с передачей

$$K_s = \frac{1}{s^5+3{,}2360s^4+5{,}236068s^3+5{,}236068s^2+3{,}236068s+1}.$$

Расчет этого фильтра был проведен за 6 итераций.

На втором этапе были учтены потери в индуктивностях, причем в качестве начальных значений параметров элементов принимались величины, полученные в конце первого этапа. В результате 3 итераций были получены следующие значения элементов фильтра:

$L_1 = L_2 = 1{,}2595972$ *гн,*

$r_1 = r_2 = 7{,}4542268 \cdot 10^{-2}$ *ом,*

$C_1 = C_3 = 1{,}2375818$ *ф,*

$C_2 = 0{,}37590573$ *ф,*

$R_1 = R_2 = 0{,}7939046$ *ом.*

Схема рассчитанного фильтра изображена на рис. 12.

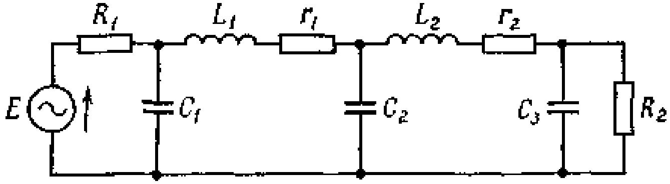


Рис. 12. Реактивный фильтр, передаточная функция которого определяется по формуле (65).

В заключение следует обратить внимание еще на одно достоинство представленного здесь метода синтеза фильтров. Этот метод не требует особой точности расчета, что необходимо при других методах синтеза реактансных фильтров, как, например, в методах Бруне или Кауэра. Это положение следует из вида линейных уравнений (57), на основании которых рассчитываются значения элементов фильтра, так как в этих уравнениях производится только умножение и сложение неизвестных и не требуется выполнения операций вычитания и деления. Это серьезное достоинство, обеспечивающее достаточную точность расчета величин элементов без необходимости применения особых методов увеличения точности расчета.

В приведенных выше примерах расчеты проводились для передач с нормированием по частоте и сопротивлению. Для расчета элементов фильтров, удовлетворяющих определенным техническим условиям, необходимо еще провести денормирование значения элементов. Так как это элементарная операция, она не программируется. Преобразование полученных прототипов фильтров низких частот в фильтры других типов, например полосовых фильтров, фильтров высоких частот или режекторных фильтров, можно легко провести при помощи преобразования частоты.

Приведенные расчеты касались фильтров с очень простой структурой. Однако метод расчета фильтров, представленный ниже, позволяет проводить расчет фильтра произвольной сложности.

5.2.4. Метод расчета элементов четырехполюсников

Задаются следующие исходные данные: блочная группа графа заземленного четырехполюсника, функция совпадения этой блочной группы, а также коэффициент передачи напряжения четырех-полюсника. Алгоритм метода определяет систему нелинейных уравне-ний (алгебраических), из которых находятся элементы четырех-полюсника: величины сопротивлений, емкостей, индуктивностей. Расчет состоит в определении решений линеаризованных систем уравнений методом Ньютона (путем последовательных приближений) .

5.2.4.1. Система уравнений, определяющих коэффициент передачи напряжения четырехполюсника

Система уравнений для определения параметров четырехполюсника при заданном коэффициенте передачи напряжения составляется с помощью формулы

$$\frac{\sqrt{R_\alpha R_\beta}\,\underset{Z}{\mathrm{Sim}}\left(\frac{\partial A^d}{\partial \alpha},\ \frac{\partial A^d}{\partial \beta}\right)}{\underset{Z}{\det} A^d} = K_u(s), \qquad (65)$$

где A^d —дополнительная блочная группа, связанная с графом четырехполюсника; α — номер входной ветви; β — номер выходной ветви; R_α — сопротивление входной ветви; R_β — сопротивление выходной ветви; K_u(s) — коэффициент передачи напряжения четырехполюсника, являющийся рациональной функцией переменной s. В дальнейшем приняты следующие ограничения.

а) блочная группа A^d описывает заземленную структуру четырехполюсника;
б) в каждой ветви четырехполюсника находится один элемент определенного типа (емкость, сопротивление или индуктивность). Ограничение б) не сужает класс решаемых задач и значительно увеличивает быстроту расчетов.

Система уравнений получается из формулы (65) путем приравнивания коэффициентов при переменной s в правой и левой сторонах числителя и знаменателя. Получается зависимость между величинами C^{-1}, R или L ветви четырехполюсника и коэффициентами числителя и знаменателя коэффициента передачи.

Из однородности полученной системы уравнений следует, что параметры элементов четырехполюсника можно определить с точностью до коэффициента пропорциональности.

Знаменатель левой дроби формулы (65) — однородный многочлен относительно элементов, имеющий степень, равную числу ветвей в деревьях A^d (N [5]); числитель — однородный многочлен степени N [5] — 1 = N [7], умноженный на $2\sqrt{R_\alpha R_\beta}$.

Блочная группа A^d и ее функция совпадения выражаются через номера ветвей. Коэффициент передачи напряжения задан коэффициентами знаменателя и числителя, представляющими действительные числа, и степенями переменной s (целые числа).

5.2.4.2. Расчет величин, определяющих систему уравнений

Столбец блочной группы A^d (дерево) определяет произведение элементов, находящихся в этом дереве (ограничение б)), и представляет собой одночлен относительно элементов. Все одночлены с одинаковыми показателями степеней s суммируются и приравниваются коэффициентам многочлена передаточной функции с той же степенью s.

Уравнение для числителя получается аналогично, но с той разницей, что полученный многочлен вначале умножается на $2\sqrt{R_\alpha R_\beta}$.

Система уравнений имеет вид

$$F(x_1, \ldots, x_n) = 0,$$
$$i = 1, 2, \ldots, nr = N\,[3] + N[4], \quad n = N[1] — N\,[2]. \qquad (66)$$

N [2] переменных исключается, поскольку приравниваются величины некоторых элементов.

5.2.4.3. Расчет якобиана уравнений (66)

Метод решения, применяемый к системе уравнений (66), предполагает расчет якобиана в любой точке Х = $(x_1 .. ., x_n)$. Обозначим

$$I_{i,\,k}\,(\mathrm{X}_{\,1},\, x_1 .. .,\, x_n) = \frac{\partial F_i(x_1 .. .,\ x_n)}{\partial \mathrm{X}_k},$$

$$i = 1, \ldots, nr; \quad k = 1,2, \ldots, n. \tag{67}$$

Расчет аналогичен расчету F_i, изложенному в разд. 5.2.4.2. Для величин F_i, определяемых многочленом знаменателя (65), составляются одночлены из переменных, номера которых находятся в столбцах A^d, при этом по очереди опускается по одной переменной.

Сумма одночленов, образующих F_i, в которых пропущена переменная x_k, создает $I_{i,\,k}$ (Х). Аналогично поступают и для F_i, определяемых числителем, только одночлены умножаются на $2\sqrt{R_\alpha R_\beta}$.

Дифференцирование относительно R_α и R_β (присутствующих в формуле только в множителе $2\sqrt{R_\alpha R_\beta}$) более сложно.

5.2.4.4. Метод Ньютона

Система (66) заменяется линейным приближением

$$F(\mathrm{X}) + I(\mathrm{X})\Delta\mathrm{X} = 0, \tag{68}$$

где

$$\mathrm{X} \text{ — матрица-столбец } \begin{Vmatrix} x_1 \\ x_n \end{Vmatrix};$$

$$F\,(\mathrm{X}) \text{ — матрица-столбец } \begin{Vmatrix} F_1(\mathrm{X}\,) \\ F_{nr}(\mathrm{X}\,) \end{Vmatrix};$$

$$I\,(\mathrm{X}) \text{ — матричный якобиан } I_{i,\,k}\,(\mathrm{X}) = \begin{Vmatrix} \dfrac{\partial F_i(\mathrm{X}\,)}{\partial \mathrm{X}_k} \end{Vmatrix};$$

$$\Delta(\mathrm{X}) \text{ — матрица-столбец приращений } \begin{Vmatrix} \Delta\mathrm{X}_1 \\ \Delta\mathrm{X}_m \end{Vmatrix} \quad .$$

Систему (68) можно решить относительно Δ(X), если она непротиворечива. Достаточным условием этого является равенство рядов матриц

$$I\,(\mathrm{X}),\ \text{а также}\quad \left\|\begin{matrix} I_1 & I_{1n} & F_1 \\ I_{nr,1} & I_{nr,n} & F_{nr} \end{matrix}\right\|.$$

Когда это условие выполняется, то для каждого Δ(X), удовлетворяющего (68), найдется такое $0 \le \lambda \le 1$, что

$$\sum_{i/1}^{nr} [F_i(\mathscr{X} + \mu\Delta\mathscr{X})]^2 < \sum_{i/1}^{nr} [F_i(\mathscr{X})]^2 \ \text{для}\ \mu \leqslant \lambda. \qquad (69)$$

Программа выполняется следующим образом. Для начального приближения X подсчитываются значения функции F и матрица I. Методом ортогонализации находится одно из Δ(X), удовлетворяющих уравнению (68). Проверяется условие (69) при μ=1. Если оно выполняется, то Δ(X) уменьшается в два раза и проверка производится снова. Так поступают до тех пор, пока условие (69) не будет удовлетворяться. Этим способом достигается уменьшение $\sum_{i/1}^{nr}[F_i(\mathrm{X})]^2$. Новое приближение X получается путем сложения предыдущего Δ(X), при этом могут быть два случая:

1) $\sum_{i/1}^{nr}[F_i(\mathrm{X})]^2 <$ epsilon — тогда система уравнений (66) считается решенной;

2) $\sum_{i/1}^{nr}[F_i(\mathrm{X})]^2 <$ delta и не выполняется 1) — тогда система уравнений считается нерешенной и ищется решение из нового начального приближения.

В результате применения метода ортогонализации через несколько итераций противоречивость системы (68) приведет к выполнению условия 2.

5.2.4.5. Решение системы линейных уравнений методом ортогонализации

Метод ортогонализации делает возможным нахождение решения только непротиворечивой системы линейных уравнений. Этот метод основывается на преобразовании матрицы коэффициентов систем уравнений в матрицу, у которой строки являются ортогональными. Это означает, что скалярные произведения различных строк равны нулю, а скалярные квадраты строк — нулю или единице.

Вместе с матрицей коэффициентов преобразуются свободные члены. Если какая-либо строка имеет скалярный квадрат, равный нулю, а свободный член этой строки не равен нулю, то система противоречива. Умножение транспонированной ортогонализованной матрицы на столбец свободных членов дает решение. В случае противоречивой системы этот метод определяет множество независимых уравнений, для которых решение ищется таким же способом.

Геометрически этот метод можно интерпретировать как прямоугольную проекцию начала координат на $(n — r)$-мерную гиперплоскость решений, расположенную в n-мерном эвклидовом пространстве (r — порядок системы уравнений).

5.2.4.6. Случаи симметрии в уравнениях (66)

Функции F_i (X) иногда симметричны, это значит существуют нетождественные преобразования переменных $x_1, \ldots, x_n$, приводящие систему уравнений (66) в себя.

Эту симметрию можно обнаружить на этапе проектирования четырехполюсника, так как она появляется в графе и через действия над блочными группами переходит в систему уравнений (65), (66), (68).

Если имеется одно решение системы (66), то, применяя к нему упомянутые преобразования, можно получить множество других решений.

С другой стороны, существование множества преобразований для системы (66) и начального приближения не должно изменять решения при этих преобразованиях, поскольку этот метод не выделяет ни одну из переменных. Число переменных можно уменьшить, если отождествить те элементы, которые переходят в самих себя. Это значительно ускоряет расчеты.

6. Основы математической теории организации структур

Математическая теория структур — стремительно развивающаяся ветвь математики, оформившаяся в качестве самостоятельной научной дисциплины в 40-х года XX в. Даже само основное понятие, которое является предметом изучения этой теории, еще не имеет общепринятого названия — в американской литературе преимущественно употребляется термин «lattice», в немецкой— «Verband». В русской литературе используется, как правило, термин «структура»; он принят и в настоящей работе.

Понятие структуры является одним из небольшого числа тех математических понятий, возникновение которых вполне оправдано и подготовлено предшествующим развитием науки. Оно возникло не как простое обобщение одного из ранее существовавших понятий, не путем поверхностной перестройки системы аксиом: до своего оформления в качестве самостоятельного объекта изучения понятие структуры на протяжении ряда десятилетий вызревало в недрах более старых ветвей математики, таких, как математическая логика, алгебра, позже теория множеств, теория вероятностей, топология, функциональный анализ. В тридцатых годах XX в., к которым относится начало самостоятельного существования теории структур, были найдены связи понятия структуры и с другими отделами математики; так, например, теория проективных геометрий оказалась просто частью теории дедекиндовых структур. Вообще, **понятие <u>структуры</u> появляется всюду, где используются <u>частично упорядоченные множества</u>**. Через теорию булевых алгебр — теория структур нашла первые выходы и за пределы математики, а именно, в теорию релейно-контактных схем.

В развитии теории структур можно отметить три периода.

После опубликования в 1847 г. «Математического анализа логики» Буля некоторые математики-логики подвергли анализу постулаты булевой алгебры и алгебры отношений. Их основные результаты изложены в «Алгебре логики» Шредера (1890—1895 гг.).

Второй период в развитии теории структур наступил спустя два года после выхода в свет «Современной алгебры» Ван-дер-Вардена. Ряд статей, опубликованных в 1933—1937 гг. Дж. Нейманом, Стоном и Канторовичем, показал, что обобщения булевой алгебры до подходящей «структуры» имеют фундаментальное значение для современной алгебры, проективной геометрии, теории точечных множеств и функционального анализа, равно как и для логики и теории вероятностей. Некоторые из этих работ были предвосхищены в двух статьях Дедекинда, оставшихся почти незамеченными, и в трудах Менгера, Тарского и Фр. Клейна. В первом издании «Теории структур», написанном в 1937—1939 гг., была сделана попытка объединить и расширить результаты этих статей.

В итоге всех этих работ теория структур стала рассматриваться как существенная отрасль математики. Теория структур непрерывно пополнялась новыми исследованиями, которые ознаменовали третий период ее развития. Этот период продолжается и в настоящее время.

6.1. Начальные сведения из некоторых разделов математики

6.1.1. Начальные сведения из алгебры

Будем предполагат, что читатель знаком с некоторыми, по крайней мере, со следующими разновидностями алгебры: группа, кольцо, поле, векторное пространство, линейная алгебра. Они не будут здесь определяться. С другой стороны, чтобы формулировать в дальнейшем многие теоремы с достаточной степенью общности, нам нужно будет несколько общих определений, применимых ко *всем* таким алгебрам по способу, который мы кратко изложим.

Под *алгеброй А* будем понимать в дальнейшем множество элементов вместе с некоторым множеством операций f_α. Каждое f_α будет *однозначной* (одновалентной) функцией, определяющей при некотором *конечном* n — $n(\alpha)$ для каждой последовательности $(x_1, \ldots, x_n)$ элементов из *А значение* $f_\alpha (x_1, \ldots, x_n)$ в *А*. Следует подчеркнуть, что хотя число различных операций f_α может быть бесконечным, каждое отдельное f_α является финитарным, т. е. применяется только к конечным последовательностям фиксированной длины, зависящей от α.

Под *подалгеброй* абстрактной алгебры понимается подмножество, содержащее в себе любую алгебраическую комбинацию своих собственных элементов — это определение включает в себя, как частные случаи, обычные определения подгруппы, подкольца, подполя, подпространства, подалгебры и т. д.

Под *изоморфизмом* между двумя алгебрами, допускающими одни и те же операции (например, между двумя группами или двумя кольцами) понимается взаимно однозначное соответствие элементов, которое сохраняет все операции. Под *гомоморфизмом* между двумя алгебрами понимается однозначное, но не обязательно взаимно однозначное соответствие, обладающее тем же свойством. Изоморфизм алгебры с самой собой называется *автоморфизмом;* гомоморфизм алгебры в себя (или на свою подалгебру) называется *эндоморфизмом.* Под *отношением конгруентности* в алгебре A понимается отношение эквивалентности $x \equiv y$ (θ) со свойством замены для каждого f_α.

(*Отношение эквивалентности* есть бинарное отношение, удовлетворяющее условию тождественности $x \equiv x$, условию, что $x \equiv y$ влечет за собой $y \equiv x$, и условию, что $x \equiv y$ и $y \equiv z$ влечет за собой $x \equiv z$.).

Если

$$x_i \equiv y_i \ (\theta) \text{ для } \ i = 1,...,n,$$

то (1)

$$f_\alpha \, (x_1, ..., x_n) \equiv f_\alpha (y_1, . . ., y_n) \ (\theta).$$

Это означает в действительности, что если разбить A на подклассы элементов, «эквивалентных» по modθ, то для каждой последовательности $X_1, . .., X_n$ таких подклассов при $n=n$ (α) множество элементов $f_\alpha(x_1, ..., x_n)$ при $x_k \in X_k$, находится в одном и том же подклассе Y, который будем обозначать через $f_\alpha(X_1, ..., X_n)$. Это определяет гомоморфный образ A_θ алгебры A с теми же однозначными операциями, что и в A. Обратно, если Ф отображает гомоморфно A на алгебру B и мы определяем $x \equiv y(\Phi)$ в A в том смысле, что $\Phi(x) = \Phi(y)$ в В, то мы приходим к отношению конгруентности. Этим устанавливается однозначное в одну сторону соответствие между отношениями конгруентности на алгебре A и ее гомоморфными образами.

Под *прямым произведением* $X \times Y$ двух алгебр X и Y, имеющих одни и те же операции f_α, понимают алгебру, элементами которой являются пары $[x, y]$ с $x \in X$ и $y \in Y$ в которой алгебраическая операция совершается покомпонентно, т. е.,

$$f_\alpha \, ([x_1, y_1], ..., [x_n, y_n]) = [f_\alpha \, (x_1,..., x_n), f_\alpha \, (y_1, ... y_n)]. \qquad (2)$$

Аналогично определяется прямое произведение n алгебр.

С произвольным классом A алгебр, имеющих одни и те же операции f_α, и с произвольным кардинальным числом n (конечным или бесконечным) ассоциируется *свободная* A *-алгебра F с n образующими a_k*, такая, что любое однозначное отображение Ф элементов a_k в элементы из $A \in \mathrm{A}$ можно продолжить до *гомоморфизма* F в A. (Подмножество G из алгебры Q называется «множеством образующих» для Q, если наименьшая подалгебра из Q, содержащая G, есть само Q.)

Очевидно, что $F_n(\mathrm{A})$ определяется однозначно с точностью до изоморфизма. В самом деле, если F и F', с образующими a_k и a'_k, обе удовлетворяют определению, то соответствие $a_k \leftrightarrow a'_k$ можно продолжить до двустороннего гомоморфизма, который есть изоморфизм. Далее, если A определено как множество всех алгебр, удовлетворяющих множеству $\mathfrak{S}$ тождественных соотношений, то мы можем описать $F_n(\mathrm{A})$ как алгебру с n образующими, удовлетворяющую $\mathfrak{S}$, для которой всякая другая алгебра с n образующими, удовлетворяющая $\mathfrak{S}$, является гомоморфным образом.

Дадим *конструктивное* определение $F_n(\mathrm{A})$, доказывающее существование $F_n(\mathrm{A})$ во всех случаях. Для этого берем, прежде всего, n символов a_k и называем «значением» символов a_k всякое их однозначное отображение Φ в элементы из $A \in \mathrm{A}$. Далее, назовем f_α «функциями порядка один» и определим «функцию порядка m» по индукции, как операцию вида

$$f_\alpha\,(g_1\,(x'_1, \ldots, x'_{r\,(1)}), \ldots, g_n\,(x^n_1, \ldots, x^n_{r\,(n)})), \tag{3}$$

где g_k — функция порядка, не превышающего m—1. Очевидно, что любая функция порядка m определяет при всяком значении символов a_k и при всякой подстановке их в (3) вместо x^i_j элемент $g\,(x'_1, \ldots, x^n_{r\,(n)})$ алгебры A, содержащей a_k. Так, например, если мы рассматриваем группу как алгебру с операциями x^{-1} и xy, то $(x^{-1}y)\,x$ будет функцией порядка три.

Теперь мы можем легко определить $F_n(\mathrm{A})$. Элементы из $F_n(\mathrm{A})$ представляют собой как раз символы $g\,(x'_1, \ldots, x^n_{r\,(n)})$, в которых каждое x^i_j замещается некоторым a_k. Два таких символа считаются равными тогда и только тогда, если они дают один и тот же результат для каждого значения Φ. По определению, эти равенства являются в точности *тождественными соотношениями,* справедливыми в каждом $A \in \mathrm{A}$.

Отсюда следует, что $F_n(\mathrm{A})$есть подалгебра прямого произведения изоморфных образов алгебр $A \in \mathrm{A}$, по одному множителю для каждого значения Φ. Кроме того, любая алгебра B с n образующими, удовлетворяющая всем тождественным соотношениям, справедливым во всяком A, является гомоморфным образом $F_n(\mathrm{A})$; следовательно. $F_n(\mathrm{A})$ есть свободная алгебра; с n образующими, определенная для A в смысле описательного определения.

6.1.2. Некоторые сведения из теории множеств и отношений

Как уже отмечалось, составление программы для ЭВМ заключается, собственно, в записи соответствующего алгоритма обработки данных на языке, понятном ЭВМ. Естественно, что алгоритм, а значит и программа будут зависеть от того, как организованы данные, относящиеся к решаемой задаче. В связи с этим остановимся детально на вопросах представления данных и их структур в ЭВМ.

Наиболее простая структура данных соответствует случаю, когда между отдельными данными отсутствуют какие-либо внутренние связи. Их набор можно рассматривать как множество.

Множество есть любое собрание определенных и различных между собой объектов, которые можно рассматривать как единое целое. Так, например, можно рассмотреть множества людей, машин, телевизоров. Отдельные объекты множества называются элементами.

Для любых объектов a_1, a_2,...a_n множество обозначается через $\{a_1, a_2,...,a_n\}$, т. е. путем перечисления всех объектов, входящих в множество.

Другая форма обозначения множества состоит в указании общего свойства объектов, из которых оно образуется $M = \{x|P(x)\}$.

Эта запись читается «множество элементов x таких, что имеет место $P(x)$». Вертикальная черта в записи расшифровывается как фраза «таких, что», а $P(x)$ обозначает свойство, характеризующее все элементы данного множества. Например, множество четных чисел М можно записать так:

$$М\{ x{:}|\, x{:} \text{ — целое число, делящееся на } 2\}.$$

К основным операциям над множествами относятся операции объединения, пересечения, разности.

Объединением множеств $М_1$ и $М_2$ ($М_1 \cup М_2$) называют множество тех элементов, которые содержатся по крайней мере в одном из множеств $М_1$ или $М_2$:

$$М_1 \cup M_2 = \{x|x \in М_1 \text{ или } x \in М_2\},$$

Пересечением множеств $М_1$ и $М_2$ ($М_1 \cap М_2$) называют такое множество, которое содержит элементы, одновременно принадлежащие как $М_1$ так и $М_2$:

$$M_1 \cap M_2 = \{x|x \in M_1 \text{ и } x \in М_2\},$$

Разностью множеств $М_1$ и $М_2$ ($М_1 \backslash М_2$) называют множество тех элементов, которые принадлежат $М_1$ и не принадлежат $М_2$:

$$М_1 \backslash М_2 = \{ x|x \in M_1 \text{ и } x \notin М_2\}.$$

Прежде чем перейти к отношениям, рассмотрим еще одно важное понятие — упорядочение элементов. Если будем рассматривать только два элемента, то интуитивно ясно, что упорядоченная пара элементов — это просто совокупность, состоящая из двух предметов, расположенных в некотором определенном порядке. Когда это понятие используют в математике, то говорят, что упорядоченной парой элементов множества М называется объект $<x_1, x_2>$, состоящий из двух (не обязательно различных) элементов $x_1, x_2 \in$ М с указанием, какой из них следует считать первым, а какой — вторым. Так, если рассмотреть множество М=$\{x_1, x_2, x_3, x_4, x_5\}$, то упорядочение пары $<x_1, x_2>$, $<x_2, x_1>$ следует считать различными. К упорядоченным парам элементов из множества М относят также пары $<x_1, x_1>$, $<x_2, x_2>$ и т. д. Множество всех упорядоченных пар $<x_i, x_j>$ из М называют декартовым (или прямым) произведением множества на себя и обозначают М×М:

$$M\times M = \{<x_i, x_j> \mid x_i \in M, x_j \in M\}.$$

Аналогичным образом вводится и прямое произведение множества M_1 на M_2:

$$M_1\times M_2 = \{<x, y> \mid x \in M_1, y \in M_2\}.$$

По аналогии с упорядоченными парами можно ввести понятие упорядоченной тройки $<x_1, x_2, x_3>$, упорядоченной четверки $<x_1, x_2, x_3, x_4>$ и, вооообще, упорядоченной n-ки, где n — произвольное число.

Упорядоченная n-ка элементов из М — это л (не обязательно различных) элементов множества М, данных в определенной последовательности $<a_1, a_2,..., a_n>$. Упорядочение n-ки элементов из множества М называют иногда кортежами над М (длиной n).

Кортежи как раз и используются для формализации отношений. Вообще термин «отношения» применяется для обозначения какой-либо связи между объектами или понятиями. Например, мы говорим «объект *х* лучше объекта у», «*х* меньше, чем *у*», «*а* равно *b*», «ДИОД Д2 включен в схему R15». Здесь примерами отношений являются выражения: «лучше», «меньше», «чем», «равно», «включен в», которые показывают, в каком отношении друг к другу находятся элементы.

Отношения бывают двуместные (или бинарные), трехместные (или тернарные) и вообще n-местные или n-арные, где n — произвольное число. Термин «бинарное (двуместное) отношение» говорит о том, что для его представления используется упорядоченная пара элементов.

Формально бинарное отношение (обозначим его R) между двумя множествами M_1 и M_2 может быть определено как подмножество прямого произведения $M_1\times M_2$:

$R(M_1, M_2) = \{<x, y> \mid x \in M_1, y \in M_2\}$.

Другими словами, бинарное отношение R представляет собой множество упорядоченных пар (x, y). Если $x \in M_1$, $y \in M_2$, и данная упорядоченная пара находится в отношении R, то это записывается в виде xRy.

n-местным отношением на множествах $M_1\times M_2\times \ldots \times M_n$ называется подмножество декартового произведения $M_1\times M_2\times \ldots \times M_n$:

$R(M_1, M_2,..., M_n) = \{<x_1, x_2,..., x_n> | x_1 \in M_1, x_2 \in M_2,..., x_n \in M_n \}$.
Множества M_1, M_2,..., M_n называют областями определения (доменами), т. е. домены — множества, из которых черпаются конкретные значения элементов кортежа, n-местное отношение $R(M_1,...,M_n)$ удобно представлять в виде таблицы, где каждая i-я строка есть кортеж $<x^i_1, x^i_2,..., x^i_n>$, а каждый столбец соответствует одному и тому же компоненту декартова произведения, т. е. в нем могут появляться только элементы из соответствующей области определения (соответствующего домена).

Очень важное требование к отношениям заключается в том, что каждая строка (кортеж) рассматривается как элемент множества R, а в множествах дубликаты не имеют смысла, т. е., например, множества {1, 2, 3} и {3, 2, 3, 1, 2} считаются эквивалентными. Поэтому если в результате операций над отношениями появляются одинаковые строки, то в результирующем отношении должна остаться только одна из них.

Можно отметить, что n-местное отношение может быть использовано для представления набора объектов. Рассмотрим отношение, названное ТЕЛЕВИЗОР (табл. 1).

Таблица 1

Наименование телевизора	Индекс	Диаметр кинескопа, см	Цена, руб.
Горизонт	Ц-355	51	610
Рубин	Ц-266д	67	1040
Темп	Ц-280д	61	755
Электрон	Ц-283	61	755

Каждая строка в табл. 1 выполняет роль описания отдельного объекта, а в данном случае атрибуты объекта интерпретируются столбцами отношений. Множество допустимых значений атрибута интерпретируется соответствующим доменом. Поэтому столбцы отношений называют также атрибутами и присваивают им имена. В этом случае можно говорить об отображении имен атрибутов в значения, принадлежащие соответствующим доменам. В общем случае при разработке структуры хранения данных используют различные виды таблиц.

Важную роль в процессе обработки данных играет упорядоченность строк таблиц. Иногда к таблице подсоединяют дополнительный «внешний» столбец, содержащий последовательность целых чисел от 1 до n, который отражает упорядоченность строк таблицы. При этом пользователь не имеет доступа к этому столбцу. Такая таблица

называется последовательной таблицей. В ней каждой текущей строчке можно поставить в соответствие предыдущую строку и последующую.
Упорядочение таблицы может быть произведено и по значениям одного из столбцов, соответствующему арбитру. В этом случае она называется упорядоченной.
В системе обработки данных атрибут или совокупность атрибутов, позволяющие уникально определять каждую строчку таблицы, называют ключом. В общем случае в таблице может быть несколько ключей, один из которых назначается первичным ключом, а остальные будут представлять возможные ключи.
Следует помнить, что существует понятие — вторичный ключ или ключ поиска. Это ключ, посредством которого можно выделять из таблицы все кортежи, имеющие определенные значения этих атрибутов. Таким образом, вторичный ключ позволяет выделять из таблицы кортежи, обладающие интересующими нас свойствами.
Список имен атрибутов отношения называется схемой отношения. Если отношение называется R и его схема имеет атрибуты с именами $A_1, A_2,\dots,A_n$, то схема отношений будет

$$R(A_1, A_2,\dots,A_n).$$

Для табл. 1 схема отношения будет иметь вид: ТЕЛЕВИЗОР (наименование, индекс, диаметр кинескопа, цена).
Можно заметить, что существует аналогия между отношением и файлом, между кортежем и логической записью, между схемой отношения и форматом записи.
Графы. В математике хорошо разработано теоретическое направление, позволяющее выражать отношениями между объектами в виде специальных рисунков — графов. Граф представляет собой набор некоторых точек на плоскости X, которые называют вершинами, и множество отрезков U, соединяющих все или некоторые вершины. Отрезки U называют дугами или ребрами графа. Например, на рис. 1 изображен граф, состоящий из пяти вершин $X=\{x|x=a, b, c, d, e)$ и трех ребер $U=\{u|u=(a,b), (a,d), (c,e)\}$.

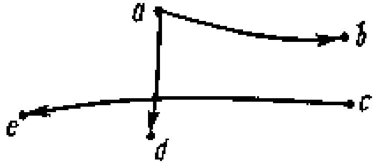


Рис. 1. Пример графа

Математически граф G можно определить как пару множеств X и U: $G(X, U)$.

Форма представления отношения R в виде графа, заданного на множестве М×М, позволяет элементы множеств М выбрать в качестве вершин, а в качестве дуг изобразить упорядоченные пары элементов $<x_i, x_j> \in$ М×М, характеризующие само отношение R. Так как дуга изображается стрелкой, то началу стрелки ставится в соответствие первый элемент кортежа $<x_i, x_j>$, а окончанию стрелки — второй. Так, графу (рис. 1) можно поставить в соответствие некоторое отношение ($<a, b>$, $<a, d>$, $<c, e>$), заданное на прямом произведении множества М на себя, где М=$\{a, b, c, d, e)$. Таким образом, при отображении с помощью графов некоторого отношения R между объектами вершины графа будут представлять объекты, а дуги — отношения между объектами.

Поскольку граф представляет собой абстрактное геометрическое отображение связей между объектами, дуги не обязательно должны быть прямыми, а вершины на плоскости могут быть расположены произвольно (рис. 6.2).

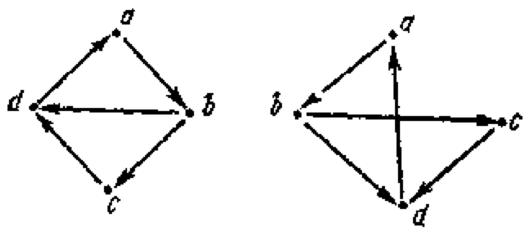


Рис. 2. Эквивалентные графы

В практических задачах типична ситуация, когда связи между объектами являются симметричными, т. е. отношение R выполняется как между объектами $<x_i, x_j>$, так и между $< x_j, x_i>$. В этом случае вместо двух дуг, соединяющих объекты (рис. 3, а), можно нарисовать одну дугу (рис. 3, *б).*

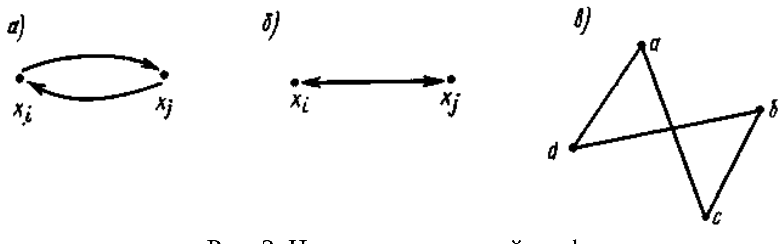


Рис. 3. Неориентированный граф

Когда все дуги в графе изображаются двойной стрелкой, стрелки можно не рисовать (рис. 3, *в).*

Дуги в этом случае называют ребрами, а сам граф — неориентированным графом. Граф, в котором задана ориентация дуг, называется соответственно ориентированным графом или орграфом.

Некоторые свойства графов. Иногда удобно представлять графы в виде некоторых матриц, в частности смежности и инциденций.

Две вершины x_i и x_j являются смежными, если они различны и существует дуга, идущая из x_i в x_j. Дуга называется инцидентной вершине *x,* если она заходит в эту вершину или исходит из нее.

Обозначим через x_1, x_2.,..., x_i,..., x_j,..., x_n вершины графа, а через u_1,..., u_k,..., u_l его дуги.

Введем числа:

$$r_{ij} = \begin{cases} 1,\ \text{если имеется дуга, соединяющая} \\ \text{вершины } x_i \subset x_j; \\ 0,\ \text{если такой дуги нет.} \end{cases}$$

Квадратная матрица $\|r_{ij}\|$ порядка $n \times k$ называется матрицей смежности графа.

Для построения матрицы инциденций введем числа s_{il}:

$$s_{il} = \begin{cases} +1,\ \text{если } u_l \text{ исходит из } x_i; \\ —1,\ \text{если } u_l \text{ заходит в } x_i; \\ 0,\ \text{если } u_l \text{ не инцидентна } x_i. \end{cases}$$

Матрица S = $\|s_{il}\|$ порядка $n \times k$ называется матрицей инциденций для дуг графа.

Подграфом *GA* графа *G(X, U)* называется граф, в который входит лишь часть вершин графа G, образующих множество $A_x \subseteq X$ вместе с дугами, соединяющими эти вершины.

Путем в графе С называется такая последовательность дуг μ= (u_1,..., u_k)*,* в которой конец каждой предыдущей дуги совпадает с началом следующей.

Контур — это замкнутый путь μ= (x_1,..., x_k)*,* у которого начальная вершина x_1 совпадает с конечной x_k.

Контур единичной длины, образованный дугой вида (x_i, x_i) называется петлей.

6.1.3. Некоторые сведения из теории топологии

Основные идеи теории точечных множеств (т. е. общей топологии) можно ввести наиболее простым способом с помощью понятия метрического пространства.

Метрическое пространство есть совокупность M элементов (точек) вместе с определенным в нем *расстоянием* $\delta\ (x,\ y)$ между любыми двумя точками $x,\ y \in M,$ удовлетворяющего условиям:

(I) $\delta(x,\ x) = 0$, в то время как $\delta\ (x,\ y) > 0$, если $x \neq y,$

(II) $\delta\ (x,y) = \delta\ (y,x),$

(III) $\delta\ (x,\ y) + \delta\ (y,\ z) \geq \delta\ (x,\ z)$ (неравенство треугольника).

Подмножество S метрического пространства называется *открытым* тогда и только тогда, если для любого $a \in S$ можно найти такое малое число $\varepsilon > 0$, что из $\delta\ (x,\ a) < \varepsilon$ вытекает $x \in S$. Подмножество S из M называется *замкнутым* тогда и только тогда, если его дополнение S' открыто. Легко видеть, что

(1) сумма любого числа открытых множеств есть открытое множество и

(2) пересечение любых двух открытых множеств открыто. Двойственно этому, поскольку при переходе к дополнениям суммы и произведения множеств меняются местами,

(1') пересечение любого числа замкнутых множеств замкнуто и

(2') сумма любых двух замкнутых множеств, есть замкнутое множество.

Легко также убедиться, что

(3) само пространство M и любая точка являются замкнутыми множествами.

Различные более общие классы абстрактных пространств можно определить следующим образом. Любое пространство с семейством «замкнутых множеств», удовлетворяющих (1') — (2') — (3), называется *T_1-простраиством.* Пересечение всех замкнутых множеств, содержащих заданное множество X, называется *замыканием* множества X и обозначается через $\overline{X}$ *;* легко проверить, что

C1. $X \leq \overline{X}$ C2. $\overline{\overline{X}} = \overline{X}$

C3*. $\overline{X + Y} = \overline{X} + \overline{Y}$ C4. Если p — точка, то $p = \overline{p}$.

Обратно, C1 — C4 влекут за собой (1') — (2') — (3), если считать множество X замкнутым тогда и только тогда, когда $X = \overline{X}$. В сущности, C1 вытекает из C3* — C4. Если заменить C4 более слабым условием, что $\overline{p} = \overline{q}$ влечет за собой $p = q$, мы получаем определение *T_0-пространства.*

Семейство Г замкнутых множеств в T_0 -пространстве, обладающее тем свойством, что любое замкнутое подмножество из T есть пересечение множеств из Г, называется *базисом* замкнутых множеств. Двойственно этому, семейство Δ открытых множеств, таких, что каждое открытое подмножество из T есть сумма множеств из Δ, называется базисом открытых множеств. *Псевдобазис* замкнутых множеств есть семейство ∑ замкнутых множеств, такое, что всякое множество некоторого базиса является суммой конечного числа множеств из ∑, следовательно, такое, что каждое замкнутое множество является пересечением конечных объединений множеств из ∑. Можно, что пересечения конечных объединений множеств *любого* семейства ∑ удовлетворяют (1') —(2'). Понятие псевдобазиса открытых множеств определяется двойственным образом. *Окрестность* $U(x)$ точки x есть открытое множество, содержащее x.

Топологическое произведение (или декартово произведение) $X=X^1\times X^2\times X^3\times\ldots$ топологических пространств X^1, X^2, X^3, ... определяется как пространство, точки которого суть $x=(x^1, x^2, x^3, \ldots)$ с $x^i\in X^i$, а псевдобазис замкнутых множеств состоит из множеств, которые мы получаем, беря для каждого замкнутого множества C из псевдобазиса каждого X^i совокупность всех x с $x^i\in$ С. Так, например, замкнутые интервалы образуют псевдобазис для прямолинейного отрезка, а прямоугольники—для квадрата. В силу (1') псевдобазис можно также составить из множеств всех x с $x^i\in C^i$ при различных выборах замкнутого множества C^i из псевдобазиса каждого X^i.

Бесконечную последовательность x_1, x_2, x_3, ... точек метрического пространства называют *сходящейся* к пределу a тогда и только тогда, если $\delta(x_n, a)\to 0$ при $n\to\infty$. Если последовательность $\{x_n\}$ сходится, то она удовлетворяет условию Коши: $\mathrm{Lim}_{m,\ n\to\infty}\delta(x_m, x_n)=0$; такие последовательности называются *последовательностями Коши.* Метрическое пространство, в котором сходится любая последовательность Коши, называется *полным.* Каждое метрическое пространство M определяет однозначным образом *полное* метрическое пространство, представляющее собой множество всех последовательностей Коши $\{x_n\}$, $\{y_n\}$, ...; $\delta(\{x_n\}, \{y_n\})= \mathrm{Lim}_{m,\ n\to\infty}\delta(x_n, y_n)$, причем точки, расстояние между которыми равно нулю, отождествляются между собой.

В более общих топологических пространствах следует рассмотреть также сходимость направленных множеств.

Определяемая ниже сходимость будет иметь свой обычный смысл, тогда как «конфинальное подмножество» будет просто подпоследовательностью.

Направленное множество индексов есть класс A индексов α *с* транзитивным отношением $\alpha \geq \beta$ (словами: α следует за β), обладающий так называемым свойством Мура — Смита:

(4) Для любых α, $\beta \in A$ существует такое $\gamma \in A$, что $\gamma \geq \alpha$ и $\gamma \geq \beta$; словами: у любых двух элементов из A имеется общий последующий.

Направленное множество $\{x_\alpha\}$ точек есть функция, относящая каждому индексу α направленного множества A точку x_α.

Если M пространство с системой Δ «открытых» множеств, то *сходимость* направленного множества $\{x_\alpha\}$ точек из M определяется согласно правилу

(5) $x_\alpha \to a$ означает, что для каждого открытого множества U, содержащего a, существует такое $\beta(U)$, что $x_\alpha \in U$ для всех $\alpha \geq \beta(U)$.

Легко доказывается следующий закон сходимости:

(6) если $x_\alpha = a$ для всех α, то $x_\alpha \to a$.

Можно также показать, что если пространство X есть топологическое произведение пространств X^1, X^2, X^3, ..., то $x_\alpha \to a$ в X, где $x_\alpha = (x_\alpha^1, x_\alpha^2, x_\alpha^3, \ldots)$ и $a = (a^1, a^2, a^3, \ldots)$, эквивалентно тому, что $x_\alpha \to a^i$ в каждом X^i.

Кроме того, мы определим *конфиналъное подмножество* направленного множества $\{x_\alpha\}$ как такое подмножество $\{x_\sigma\}$ множества $\{x_\alpha\}$, что у каждого члена из $\{x_\alpha\}$ имеется последующий в $\{x_\sigma\}$. Отсюда непосредственно следует

(7) если $x_\alpha \to a$ и $\{x_\sigma\}$ есть конфинальное подмножество множества $\{x_\alpha\}$, то $x_\sigma \to a$.

Если подмножество $\{x_\sigma\}$ множества $\{x_\alpha\}$ не конфинально, то у некоторого x_α нет последующего в $\{x_\sigma\}$. Следовательно, если $x_\beta \in \{x_\alpha\}$, то любой общий последующий для x_α и x_β должен находиться в дополнении множества x_σ. Но в силу (4) такой последующий существует, и мы заключаем:

(8) если подмножество направленного множества не конфинально, то его дополнение конфинально.

Если семейство Δ открытых множеств удовлетворяет (1) — (2), то

(9) подмножество X из M замкнуто тогда и только тогда, если X содержит предел любого сходящегося направленного множества своих точек.

Доказательство. Если дополнение X' множества X открыто, то из $x_\alpha \to a$, $a \in X'$, вытекает, что у каждого x_α есть последующий в X'; следовательно, $\{x_\alpha\}$ не может лежать в X. Поэтому, если $\{x_\alpha\}$ лежит в X и $x_\alpha \to a$, то $a \in X$. Обратно, если X не замкнуто, то существует точка a, не лежащая в X, но лежащая в каждом замкнутом множестве, содержащем X (так как в силу (1) пересечение $\overline{X}$ всех замкнутых множеств, содержащих *X,* замкнуто, а потому не совпадает с *X).* Следовательно, каждое открытое множество *U,* содержащее *a,* имеет замкнутое дополнение *U',* которое не содержит *X,* так что мы можем выбрать точку x_U, лежащую в *X* и в *U.* Но если теперь мы определим последующие для *U* как окрестности, содержащиеся в нем, то в силу (2) свойство Мура — Смита имеет место, а потому $\{x_U\}$ будет направленным множеством точек из *X,* сходящимся к *a* в смысле (5).

Пусть снова *M* некоторое «пространство» и пусть дано правило, устанавливающее для каждого направленного множества $\{ x_\alpha \}$ и точки *a,* будет или нет $x_\alpha \to a$. Если мы определим «замкнутые» множества согласно (9), то, очевидно, будут иметь место (1)— (1').

Далее, если предположить (7), то можно, используя (8), показать, что имеют место (2) — (2'): *если* $x_\alpha \to a$, где x_α лежат в Y+Z, а Y и Z замкнуты, то в силу (8) подмножество элементов x_α, лежащее в *Y* или в *Z,* будет конфинальным; следовательно, в силу (7) *a* лежит в Y или в *Z* (т. е. $a \in Y+Z$). Наконец, если $x_\alpha \to a$, и *U* открытое множество, содержащее *a*, то дополнение *U'* множества *U* замкнуто, и $a \notin U'$. Следовательно, подмножество из $\{ x_\alpha \}$, лежащее в *U',* не конфинально, и существует такое β*(U),* что ни одно x_α [α≥β] не лежит в *U';* значит, $x \in U$ для всех α≥β (*U*).

Это показывает, что если мы применим снова (5), то получим по крайней мере столько же «сходящихся» направленных множеств, сколько мы имели первоначально. Точные условия, при которых сходимость сохраняет свой первоначальный смысл, т. е. абстрактная характеристика свойств сходимости, являются, более сложными.

6.2. Частично упорядоченные множества

6.2.1. Основное определение

Характерной особенностью математики как дедуктивной науки является то, что она использует очень мало неопределенных терминов. Это особенно верно для теории структур, все понятия которой можно определить в терминах одного неопределенного отношения: *«х* содержит *у»* в смысле *«у* есть часть ж». Имеется бесчисленное множество примеров таких отношений, смысл некоторых мы кратко изложим.

Определение. *Под «частично упорядоченным множеством» (*от немецкого «teilweise geordnete Menge»*) понимают систему X, в которой определено. бинарное отношение x≥y, удовлетворяющее*

P1. *Для всех x x≥x (рефлексивность).*

P2. *Если* $x \geq y$ *и* $y \geq x$, *то* $x = y$ *(антисимметричность).*
P3. *Если* $x \geq y$ *и* $y \geq z$, *то* $x \geq z$ *(транзитивность).*
Символ $\geq$ читается «содержит», «включает» или «больше или равно». Если $x \geq y$, но $x \neq y$, то пишут $x > y$ и говорят: x «больше, чем» или «собственно включает» y. Отношение $x \geq y$ пишут также в виде $y \leq x$ и говорят, что y содержится в x (или включен в x, или есть часть x). Аналогично $x > y$ пишут также в виде $y < x$.
Эти обозначения и терминология приняты повсюду.
Если $a > b$ в частично упорядоченном множеотве P, то множество элементов x, удовлетворяющих $a \geq x \geq b$, называется замкнутым *интервалом* $[b, a]$. Об элементах x, удовлетворяющих $a \geq x \geq b$, говорят, что они находятся *между* a и b.
Под *порядком* $n(P)$ частично упорядоченного множества P мы понимаем число его элементов. Порядок может быть конечным или бесконечным.

6.2.2. Примеры

Окружающий нас мир изобилует примерами частично упорядоченных множеств. Всякая иерархия по существу образует частично упорядоченное множество; это же относится и к различным частям любого целого. Мы дадим теперь несколько простых математических примеров частично упорядоченных множеств.
Пример 1. P состоит из всех подмножеств некоторого множества I (включая само I и пустое множество $\varnothing$); $x \geq y$ означает, что x содержит y как подмножество.
Пример 2. P состоит из целых положительных чисел, а $x \geq y$ означает, что x делит y.
Пример 3. P состоит из идеалов $H, J, K, \ldots$ произвольного кольца R, а $H \geq K$ означает, что K есть подмножество в H.
Идеал является, как известно, подмножеством H кольца R, отличающимся от других подмножеств в R свойствами: а) из $a \in H$ и $b \in H$ следует $a \pm b \in H$ и б) из $a \in H$ и $b \in R$ следует $ab \in H$ и $ba \in H$. Более общо, класс подмножеств, «отличающихся» *некоторым* специальным свойством, образует частично упорядоченное множество относительно теоретико-множественного включения. Так, например, это справедливо для подгрупп некоторой группы, для подпространств некоторого векторного пространства, для «измеримых» подмножеств на прямой, для топологически замкнутых подмножеств некоторого T_0-пространства и т. д.. Эти примеры наглядно показывают, что частично упорядоченное множество

содержит, вообще говоря, пары элементов *Н, К,* которые *несравнимы* в том смысле, что не имеет места ни $H \geq K$, ни $K \geq H$.

Более общо, пусть *А* частично упорядоченная система и *X* подмножество в *А*. Тогда в *X* можно определить отношение $x \geq y$, означающее, что $x \geq y$ в *А*. Если Р1 — Р3 выполняются для $\geq$ в *А,* то они *тем более* выполняются для $\geq$ в *X*. Мы получаем, таким образом, следующий результат.

Теорема 1. *Любое подмнсжество частично упорядоченного множества является само частично упорядоченным относительно того же отношения включения.*

Пример 4. *R* — совокупность вещественных чисел и $x \geq y$ имеет свое обычное значение для вещественных чисел.

Пример 5. *F* состоит из всех вещественных однозначных функций $f(x)$, определенных для $-1 \leq x \leq 1$; $f \geq g$ обозначает, что $f(x) \geq g(x)$ для всех *x*.

Пример 6. *Р* состоит из разбиений множества *S,* т. е. делений π, π', π'', ... множества *S* на неперекрывающиеся подклассы; $\pi \leq \pi'$ означает, что π есть *подразбиение* π'.

Пример 7. Пусть *Р* состоит из тех разбиений θ, θ', θ'', ... абстрактной алгебры *А,* которые являются «отношениями конгруентности», т. е. обладают свойством замены, описанным в алгебраическом предисловии. Определяем $\theta \leq \theta'$ так же, как и в примере 6.

6.2.3. Изоморфизм и двойственность

Под *изоморфизмом* между двумя частично упорядоченными множествами *S* и *S** понимают взаимно однозначное соответствие θ между *S* и *S** такое, что

$$x \geq y \text{ влечет за собой } \theta(x) \geq \theta(y) \tag{1}$$

и

$$\theta(x) \geq \theta(y) \text{ влечет за собой } x \geq y. \tag{1'}$$

Два частично упорядоченных множества называются *изоморфными* тогда и только тогда, если между ними существует изоморфизм; изоморфизм частично упорядоченного множества с самим собой называется автоморфизмом. Соответствие, однозначное в одну сторону и удовлетворяющее (1), называется *изотопным* или *сохраняющим порядок.*

Под отношением, *обратным* отношению ρ, понимают такое отношение $\breve{\rho}$, что $x \breve{\rho} y$ (словами: «*x* находится в отношении $\breve{\rho}$ к *y*») тогда и только тогда, когда $y \rho x$. Так, например, обратным отношению «содержит» будет отношение «содержится»; обратным для «больше» будет «меньше». Из рассмотрения условий Р1—Р3 имеет место

Теорема 2 (принцип двойственности). *Отношение, обратное частичной упорядоченности, является само частичной упорядоченностью.*

Определение. *«Двойственной» к частично упорядоченной системе X называют частично упорядоченную систему* $\breve{X}$ *, определенную на тех же самых элементах с помощью обратного отношения.*

Эта терминология законна, поскольку $\breve{\breve{X}}$ *=X;* мы будем часто говорить о системах, изоморфных с $\breve{X}$, как о двойственных *X*, хотя было бы лучше говорить о них, как об «антиизоморфных» *X*. Очевидно, что частично упорядоченные множества попарно двойственны (или антиизоморфны), если только они не двойственны самим себе. Подобным же образом определения и теоремы, относящиеся к частично упорядоченным множествам, попарно двойственны, если только они не двойственны самим себе; если какая-нибудь теорема справедлива для частично упорядоченных множеств, то "справедлива и двойственная ей теорема.

Как мы увидим позже, этот «принцип двойственности» распространяется на алгебру, проективную геометрию, логику и др..

Очевидно, что если θ есть антиизоморфизм, то

$$x \geq y \text{ влечет за собой } \theta(x) \geq \theta(y) \qquad (2)$$

и

$$\theta(x) \leq \theta(y) \text{ влечет за собой } x \geq y. \qquad (2')$$

Более обще, однозначное в одну сторону соответствие называется *антитонным,* если оно удовлетворяет условию (2).

Многие частично упорядоченные множества двойствственны самим себе (т. е. антиизоморфны самим себе). Так, например, множество примера 1, 6.2 2, двойственно самому себе; соответствие, переводящее каждое подмножество в свое дополнение, взаимно однозначно и переворачивает отношение включения. Аналогично, множество всех линейных подпространств *n*-мерного эвклидового пространства, содержащих начало координат, двойственно самому себе; соответствие, переводящее каждое подпространство в свое ортогональное дополнение, взаимно однозначно и переворачивает отношение включения.

В этих случаях двойственность самому себе имеет период два: трансформация $(x')'$ трансформации x' произвольного элемента x есть x. Мы будем называть двойственности самим себе (или антиавтоморфизмы) периода два *инволюциями.*

6.2 4. Квази-упорядоченпость

Многие математические системы обладают отношениями, которые удовлетворяют P1 и P3, но не удовлетворяют P2. Назовем такие отношения *квази-упорядоченностями.*

Например, R — кольцо целостности с единицей и $a\rho b$ означает, что $ax=b$ для некоторого x в R. Или, Г —класс топологических пространств и $S\rho T$ означает, что T гомеоморфно подмножеству на S. Далее, пусть Ф состоит из вещественных функций, определенных на единичном квадрате, $0\leq x,\ y\leq 1$, и $f\rho g$ означает, что $f(x,\ y)\geq g\ (x,\ y)$, за исключением множества меры нуль («почти всюду»).

Каждая такая квази-упорядоченность связана естественным образом с отношением эквивалентности (т. е. рефлексивным, симметричным и транзитивным отношением) и с частичной упорядоченностью.

Теорема 3. *Пусть* ρ *некоторая квази-пуорядоченностъ множества X. Отношение x~y, означающее, что xρy и yρx, есть отношение эквивалентности. Если отождествить «эквивалентные» элементы, то ρ превратится в частичную упорядоченность.*

Доказательство. В силу P1 $x\sim x$. Очевидно, что иэ $x\sim y$ следует $y\sim x$. Наконец, если $x\sim y$ и $y\sim z$, то $x\sim z$, ибо $x\rho y$ и $y\rho z$; в то же время $z\rho x$, ибо $z\rho y$ и $y\rho x$ (а также в силу P3); следовательно, $x\sim z$. Этим показано, что ~ есть отношение эквивалентности, или что это есть разбиение X на неперекрывающиеся подмножества, такое, что $x\sim y$ тогда и только тогда, если x и y лежат в одном и том же подклассе. Кроме того, если $x\rho y$, $x\sim x^*$ и $y\sim y^*$, то $x^*\rho x\rho y\rho y^*$, откуда $x^*\rho y^*$, т. е. отношение ρ определено, по существу, на целых подклассах. Отсюда непосредственно следует, что в системе, образованной этими подклассами, ρ удовлетворяет P1 и P3 и что $x\rho y$ и $y\rho x$ влекут за собой $x\sim y$ (т. е. $x = y$ в системе, образованной подклассами).

Предыдущая конструкция имеет много приложений. Так, например, она приводит к обычному определению ассоциированных чисел в первом примере, к определению Фреше топологической размерности во втором и к обычному определению эквивалентных функций в третьем приведенном выше примере.

6.2.5. Диаграммы

В любой иерархии важно знать, в каком случае один человек является непосредственным старшим для другого. Понятие непосредственного старшего можно определить абстрактно в теории частично упорядоченных множеств следующим образом.

Определение. *Выражение «а покрывает* b*» означает, что* $a > b$ *и что* $a > x > b$ *не выполняется ни для какого x.*

Это приводит к графическому представлению любого конечного частично упорядоченного множества X. Для изображения элементов из X рисуются малые кружочки так, что a лежит выше, чем b, если $a > b$. a и b соединяют отрезком тогда, если a покрывает b. Любая полученная так фигура называется *диаграммой X* (часто также «диаграммами Хассе» ввиду их эффективного использования Хассе); примеры изображены на рис. 1, *а* —1, *д*.

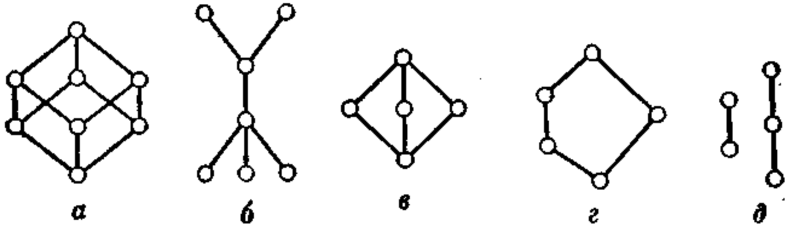


Рис. 1.

Легко показать, что любое конечное частично упорядоченное множество определяется с точностью до изоморфизма своей диаграммой; $a > b$ тогда и только тогда, когда существует такая последовательность $x_0, x_1, ..., x_n$, что $a=x_0$, $b = x_n$ и x_{i-1} покрывает x_i для $i=1, ..., n$. Графически это означает, что из a можно перейти вниз к b вдоль ломаной линии. Этот принцип облегчает построение абстрактных примеров частично упорядоченных множеств.

Наличие или отсутствие изоморфизма частично упорядоченных множеств, имеющих малое число элементов, можно установить совсем просто, наблюдая их диаграммы. Всякий изоморфизм должен быть взаимно однозначным соответствием между самыми низкими элементами, между элементами, лежащими непосредственно над самыми низкими элементами, и так все выше. Соответствующие элементы должны покрываться равным числом различных элементов и т. д.; при небольшом воображении можно быстро довести испытание до конца.

Множество $\check{X}$, двойственное частично упорядоченному множеству X, получается просто путем поворота диаграммы X вверх ногами.

Наконец, под *графом* частично упорядоченного множества понимают граф (одномерный комплекс), определенный диаграммой этого множества; это понятие, однако, менее важно.

6.2.6. Наибольший и наименьший элементы

Легко показать, что частично упорядоченное множество X может содержать самое большее один элемент a, удовлетворяющий соотношению $a \leq x$ для всех $x \in X$. Действительно, если a и b два таких элемента, то, по предположению, $a \leq b$ и $b \leq a$, откуда в силу P2 $a = b$. Такой элемент, если он существует, обозначают через 0 и называют наименьшим элементом (или нулевым элементом) в X.

Более общо, под *наименьшим* элементом некоторого подмножества X в P мы понимаем такой элемент $a \in X$, что $a \leq x$ для всех $x \in X$. Двойственно этому, под *наибольшим* элементом в X мы понимаем элемент $b \in X$ такой, что $b \geq x$ для всех $x \in X$. Наибольший элемент в P, если он существует, обозначают через I; его называют также единичным элементом в P.

Предыдущие понятия не следует смешивать с понятиями *минимального* и *максимального* элементов. Минимальный элемент подмножества X частично упорядоченного множества P есть такой элемент a, что ни для какого $x \in X$ не имеет места $a > x$; максимальные элементы определяются двойственным образом. Очевидно, что наименьший элемент должен быть минимальным, а наибольший элемент — максимальным, но обратное неверно (см. рис. 1, *б)*.

Теорема 4. *Конечное подмножество X частично упорядоченного множества обладает минимальным и максимальным членами.*

Доказательство. Пусть элементы X суть $x_1, \ldots, x_n$; определим $m_1 = x_1$ и положим m_k равным x_k, если $x_k < m_{k-1}$ и m_{k-1} в противном случае. Тогда m_n будет минимальным. Двойственным образом X будет иметь максимальные элементы.

Элемент, покрывающий 0 в частично упорядоченном множестве P (т. е. минимальный элемент в подмножестве множества P, полученном исключением 0), называется *атомом* или *точкой* (Эвклид: «Точка есть то, что не имеет частей»).

6.2.7. Кардинальные арифметические операции

Многие важные частично упорядоченные множества можно представить наиболее удобным образом как арифметические комбинации более простых множеств, используя шесть обобщений обычных арифметических операций сложения, умножения и возвышения в степень. Первые две из этих операций можно определить для произвольных отношений; мы сохраним обозначение $\geq$.

Определение. *Пусть X и Y— множества, каждое с отношением* $\geq$. *Под кардинальной суммой X + Y множеств X и Y мы понимаем множество всех элементов, лежащих в X или в Y, где* $\geq$ *сохраняет свое значение внутри X и Y,* $x \geq y[x \in X, y \in Y]$ *никогда не имеет места и X и Y рассматриваются как разделенные. Под кардинальным произведением XY мы понимаем множество всех пар (x, y)* $[x \in X, y \in Y]$, *где* $(x, y) \geq (x_1, y_1)$ *означает, что* $x \geq x_1$ *в X и* $y \geq y_1$ *в Y.*

Графически сложение двух частично упорядоченных множеств сводится просто к расположению их диаграмм одной рядом с другой. Умножение сводится к образованию их комбинаторного произведения (см. ниже, 6.2.10). Легко показать, что кардинальные сумма и произведение любых двух частично упорядоченных множеств частично упорядочены.

Легко также видеть, что если *m, n*, ... представляют собой конечные или бесконечные неупорядоченные множества из *m, n,* ... элементов, в которых $\geq$ означает равенство, то $m + n$ и mn имеют свой обычный смысл. Далее, $1A = A$, $24 = A+A$ и т. д. для всех *A*. В дополнение, легко доказать следующие обобщения обычных арифметических тождеств:

$$X + Y=Y + X, \qquad X+(Y + Z) = (X + Y) + Z, \tag{3}$$

$$XY = YX, \qquad X(YZ) = (XY)Z, \tag{4}$$

$$X(Y + Z)=XY + XZ, \quad (X + Y)Z = XZ + YZ. \tag{5}$$

Здесь знак равенства означает «изоморфно с».

Определение. *Пусть X и Y — частично упорядоченные множества. Кардинальная степень* Y^X *с основанием Y и показателем X есть множество всех изотопных функций* $y = f(x)$, *отображающих X в Y; это множество частично упорядочивается отношением* $f \geq g$, *означающим, что* $f(x) \geq g\{x)$ *для всех* $x \in X$.

Приводим без доказательства того, что Y^X является всегда частично упорядоченным множеством, так же как и без доказательства тождеств

$$X^{Y+Z} = X^Y X^Z, \quad (XY)^Z = X^Z Y^Z, \quad (X^Y)^Z = X^{YZ}. \tag{6}$$

Приводим без доказательства также отношения двойственности

$$\breve{\overbrace{A+B}} = \breve{A} + \breve{B}, \quad \breve{\overbrace{AB}} = \breve{A}\breve{B}, \quad \breve{\overbrace{B^A}} = \breve{B}^{\breve{A}}. \tag{7}$$

6.2.8. Ординальная арифметика

Под конечным *ординальным* (порядковым) числом *n,* выписываемым жирным шрифтом, мы будем понимать повсюду множество целых чисел 1,2, ... , *n* в их естественном порядке или какое-нибудь изоморфное множество. Определим три операции на частично упорядоченных множествах, которые в случае порядковых чисел сводятся к обычному сложению, умножению и возведению в степень, исключая то, что степени порядковых чисел с бесконечными показателями не обязаны быть порядковыми числами и степени бесконечных частично упорядоченных множеств не обязаны быть частично упор ядоченными. (Бесконечное порядковое число означает вполне упорядоченное множество; трудность возведения в степень бесконечных порядковых чисел и частично упорядоченных множеств связана с проблемой полного упорядочивания несчетного множества).

Определение. *Ординальная сумма* $X \oplus Y$ *двух (не перекрывающихся) частично упорядоченных множеств* X *и* Y *есть множество всех* $x \in X$ *и* $y \in Y$, *где* $x \leq x_1$ *в* X *и* $y \leq y_1$ *в* Y *сохраняют свое первоначальное значение и* $x \leq y$ *для всех* $x \in X$, $y \in Y$. *Ординальное произведение* $X \circ Y$ *есть множество всех пар* (x, y), *где* $(x, y) \leq (x_1, y_1)$ *определено в том смысле, что или* $x \leq x_1$, *или* $x = x_1$ *и* $y \leq y_1$. *Ординальная степень* ${}^{X}Y$ *состоит из всех функций* $y = f(x)$, *отображающих* X *в* Y, *где* $f \leq g$ *означает: для каждого* x *такого, что* $f(x) > g(x)$, *существует* $x_1 < x$ *такое, что* $f(x_1) < g(x_1)$.

Хотя ординальное сложение и умножение в применении к конечным порядковым числам коммутативны, они не будут таковыми для бесконечных порядковых чисел. Однако они являются ассоциативными:

$$X \oplus (Y \oplus Z) = (X \oplus Y) \oplus Z \text{ и } X \circ (Y \circ Z) = (X \circ Y) \circ Z. \quad (8)$$

Умножение и возведение в степень являются также полудистрибутивными на суммах. В формулах это означает, что для любых частично упорядоченных множеств X, Y, Z

$$(X \oplus Y) \circ Z = (X \circ Z) \oplus (Y \circ Z) \text{ и } {}^{X \oplus Y}Z = {}^{X}Z \circ {}^{Y}Z. \quad (9)$$

Нужно также отметить, что существует много тождественных соотношений, относящихся к ординальным и к кардинальным операциям. Наконец, нужно заметить, что три ординальные и две из трех кардинальных операций являются частными случаями двух обобщенных операций. Определим лексикографическую сумму $\sum_X Y_x$ семейства (не перекрывающихся) частично упорядоченных множеств Y_x, взятых по одному для каждого индекса x в базисном частично упорядоченном множестве X, как множество всех элементов y_x из всех Y_x, частично упорядоченное отношением $y_x \leq y^*_1$, означающим, что $x < x^*$ или $x = x^*$ и $y \leq y^*$. Подобно этому, определим лексикографическое произведение $\Pi_X Y_x$ как множество всех функций $f(x)$, относящих каждому $x \in X$ $y = f(x) \in Y_x$, частично упорядоченное отношением $f \leq g$, означающим, что если $f(x) > g(x)$ для некоторого x, то $f(x_1) < g(x_1)$ для некоторого $x_1 < x$.

Если X — неупорядоченное множество 2 из двух элементов, то $\sum_2 Y_x = Y_1 + Y_2$; если X—упорядоченное множество 2 из двух элементов, то $\sum_2 Y_x = Y_1 \oplus Y_2$; если все Y_x изоморфны некоторому фиксированному Y, то $\sum_X Y_x = X \circ Y$. Графически $\sum_X Y_x$ можно построить путем замены каждого элемента x в диаграмме X диаграммой множества Y_x.
Аналогично, если $X = 2$, то $\Pi_2 Y_x = Y_1 Y_2$; если $X = 2$, то $\Pi_2 Y_x = Y_1 \circ Y_2$; если все Y_x изоморфны некоторому фиксированному Y, то $\Pi_X Y_x = {}^X Y$.

6.2.9. Цепи

Отношение включения во многих частично упорядоченных множествах удовлетворяет условию
Р4. Для любых заданных x и y или $x \geq y$, или $y \geq x$.
Иными словами, из любых двух элементов один является меньшим, а другой большим.
Определение. *О частично упорядоченном множестве, удовлетворяющем условию* Р4, *говорят, что оно «просто упорядочено», и называют его «цепью».*
Положительные целые числа, упорядоченные по величине, образуют цепь; то же относится к вещественным числам. Очевидно, что любое подмножество цепи есть цепь; множество, двойственное цепи, есть цепь. Кроме того, справедлива
Теорема 5. *В цепях понятия минимального и наименьшего (максимального и наибольшего) элементов существенно эквивалентны. Следовательно, любая конечная цепъ имеет наименьший (первый) и наибольший (последний) элемент.*
Доказательство. Если ни для какого $x \in X$ не выполняется $x < a$, то в силу Р4 $x \geq a$ для всех $x \in X$.
Записывая теперь по порядку первый (наименьший) элемент конечной цепи, затем первый из оставшихся элементов и т. д., мы видим, что каждый элемент содержится во всех последующих элементах и (в силу Р2) ни в каких других. Таким образом, имеет место
Теорема 6. *Каждая конечная цепь из n элементов изоморфна цепи целых чисел* 1, ..., n.
Под *размерностью* или *высотой* $d[x]$ элемента x в частично упорядоченном множестве P мы понимаем максимум «длины» d цепей $x_0 < x_1 < ... < x_d = x$ в P, имеющих x наибольшим элементом,—в случае, когда d конечно. Аналогично, под *размерностью* или *длиной* $d[P]$ множества P мы понимаем максимум длин цепей в P.
Очевидно, что $d[P]$ есть максимум чисел $d[x]$; очевидно также, что при определении размерностей нужно рассматривать лишь цепи $x_0 < x_1 < ...$

$< x_d = x$, являющиеся «максимальными» или «связными» в том смысле, что x_i покрывает x_{i-1} для всех i. Понятие размерности особенно важно, если P имеет 0 и выполняется следующее

Условие Жордана—Дедекинда. *Все конечные связные цепи между двумя фиксированными конечными точками имеют одну и ту же длину.*

Это условие выполняется в примерах 1—3, п. 6.2.2, для неприводимых алгебраических многообразий в n-мерном пространстве и во многих других примерах. Если оно выполнено в P, то удобно изображать каждое $x \in P$ маленьким кружочком на $d\,[x]$ единиц выше, чем 0. Если x покрывает y, то существует связная цепь $x > y > y_1 > \ldots > 0$ от x до 0 длины $d[y]+1$, и обратно. Следовательно, x покрывает y тогда и только тогда, если $x>y$ и $d[x] = d[y]+1$, и отрезки в диаграмме множества P соединяют только элементы на соседних уровнях.

Лемма. *Пусть P произвольное частично упорядоченное множество, имеющее 0 и I, в котором все цепи конечны. Тогда P удовлетворяет условию Жордана—Дедекинда в том и только в том случае, если существует целочисленная функция $d\,[x]$ такая, что*

x покрывает y тогда и только тогда, если $x> y$ и $d\,[x] = d\,[y] + 1$. (*)

6.2.10. Абстрактные конфигурации и топологические комплексы

Понятие абстрактной конфигурации, определявшееся по-разному различными авторами, является в некоторой степени неясным. Однако обычно соглашаются с тем, что это понятие должно включать в себя понятия комбинаторного комплекса (в топологическом смысле) и проективной геометрии и само входить в содержание понятия частично упорядоченного множества. Мы введем поэтому следующее определение.

Определение. *Конфигурация есть частично упорядоченное множество конечной длины с I, удовлетворяющее условию Жордана—Дедекинда. Два элемента $x,y \in P$ называются инцидентными, если $x \geq y$ или $y \geq x$.*

Для всех практических приложений верно также, что $x > y$ означает наличие «точки», содержащейся в x, но не содержащейся в y.

Мы увидим в дальнешем, что точки, прямые и плоскости в эвклидовой, проективной и гиперболической геометриях (вместе с пустым множеством или без него) составляют «конфигурации» в этом смысле. «Точки» имеют размерность один (нуль, если пустое множество исключено); «прямые» имеют в этой абстрактной конфигурации размерность два. Это очевидно всякому, имеющему необходимые геометрические представления; очевидно также, что подобные примеры можно найти для размерностей *n* и (за исключением гиперболического случая) над произвольными полями.

Нас интересуют здесь предпочтительно топологические комплексы как абстрактные конфигурации. Любой топологический комплекс есть конфигурация, и топологические «*р*-клетки» (или *р*-мерные элементы) суть элементы с $d\ [x] = p$. Коэффициент инциденции $[x : y]$ отличен от нуля тогда и только тогда, если *х* «покрывает» *у*. Далее, «произведения» в смысле п.6.2.7 соответствуют декартовым произведениям в обычном смысле; это иллюстрируется на рис. 2,а — 2,в, где нарисованы диаграммы абстрактных конфигураций, изображающих отрезок S_1 (1—симплекс), квадрат $S_1 \times S_1$ и треугольник (2—симплекс) S_2.

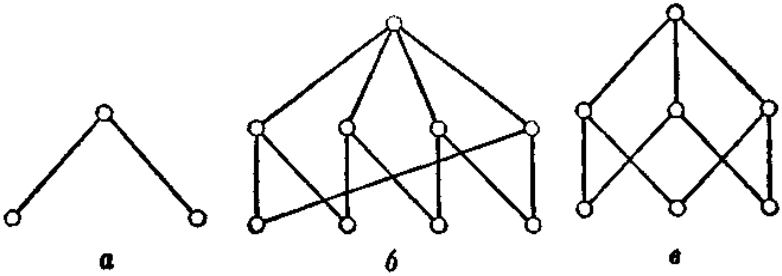


Рис. 2.

Топологическая «сумма» двух многообразий также соответствует кардинальной сумме (см. п.6.2.7) соответствующих конфигураций. Наконец, если мы включим в рассмотрение пустое множество, то конфигурация, определенная комплексом, топологически «дуальным» комплексу *С,* будет двойственной (в смысле п.6.2.3) конфигурации, определенной комплексом *С.*

Более того, конфигурация, определенная *n*-симплексом, есть просто 2^{n+1} (2 обозначает упорядоченное множество из двух элементов) с вычеркнутым 0. Следовательно, *можно легко охарактеризовать те конфигурации, которые изображают симплициальные комплексы.* Однако абстрактная характеристика полиэдра (подразделения *n*-сферы) не известна; неясно, следовательно, как охарактеризовать конфигурации, изображающие полиэдральные комплексы.

Лефшец предложил называть абстрактную конфигурацию «комплексом» тогда и только тогда, если она удовлетворяет следующему условию ориентируемости: значения ± 1 можно так сопоставить отношениям покрытия (иначе, отношениям инцидентности), что для любых x и x'' с $d\,[x] = d[x^n]+2$ будет $\sum_{x'} [x : x'] \, [x' : x'']=0$. Он показал, следуя идеям Ньюмена, Тэкера и Мейера, что многие из наиболее важных понятий топологии применимы к таким «комплексам».

6.2.11. Частично упорядоченные множества и *T_0*- пространства

Другая важная простая связь частично упорядоченных множеств с топологией существует на основе понятия T_0-пространства, определенного ранее. Определим *М-замыкание* любого подмножества S частично упорядоченного множества X как множество $\overline{S}$ всех t таких, что $t \leq s$ для одного или нескольких $s \in S$.

В силу Р1 $\overline{S} \geq S$ и в силу Р3 $\overline{\overline{S}} = \overline{S}$. Ясно, далее, что $\overline{S+T} = \overline{S} + \overline{T}$. Наконец, в силу Р2 из $\overline{p} = \overline{q}$ следует $p = q$.

Следовательно, X есть T_0-пространство. Кроме того, в этом пространстве $q \in \overline{p}$ тогда и только тогда, если $q \leq p$. Но, обратно, определение $q \leq p$ в смысле $q \in \overline{p}$ частично упорядочивает точки любого T_0-пространства. (Доказательство: $p \in \overline{p}$; далее, из $p \in \overline{q}$ и $q \in \overline{p}$ вытекает $\overline{p} = \overline{q}$, а потому $p = q$; наконец, из $p \in \overline{q}$ и $q \in \overline{r}$ вытекает $q \in \overline{\overline{r}} = \overline{r}$.)

Если мы теперь рассмотрим только конечные множества $S = \sum_{i=1}^{n} p_i$, то $\overline{S} = \sum_{i=1}^{n} \overline{p}_i$ является множеством элементов $q \leq p_i$ для одного или нескольких $p_i \in S$ и эти соответствия взаимны. Отсюда вытекает

Теорема 7. *Существует взаимно однозначное соответствие между конечными частично упорядоченными множествами X и конечными T_0 -пространствами $\mathfrak{X}$; $q \leq p$ в X означает, что $q \in \overline{p}$ в $\mathfrak{X}$.*

Теорема 8. *В обозначениях предыдущей теоремы 2^X изоморфно «кольцу» всех открытых подмножеств множества $\mathfrak{X}$ и, следовательно, антиизоморфно «кольцу» всех замкнутых подмножеств множества $\mathfrak{X}$.*

Замечание. Система подмножеств называется «кольцом», если она вместе с любыми двумя множествами S и T содержит их объединение $S \cup T$ и пересечение $S \cap T$.

6.2.12. Численные методы

Пусть P частично упорядоченная система с 0, у которой все интервалы [0, a] (множества элементов x, удовлетворяющих соотношению $0 \leq x \leq a$) конечны. Различные авторы дали рекуррентное определение *функции Мёбиуса* μ на *P:*

$$\mu\,[0] = 1, \text{ и } \mu\,[x] = -\sum_{y<x} \mu[y], \text{ если } x > 0. \qquad (10)$$

(Существует связь между функцией Мёбиуса топологического комплекса и его характеристикой Эйлера—Пуанкаре).

Основанием для названия ее функцией Мёбиуса является то, что она для случая положительных чисел сводится к функции Мёбиуса теории чисел, если $x \leq y$ значает, что x делит y (двойственно примеру 2, п.6.2.2

Ф. Холл показал, что если обозначить через λ *(х; п)* число цепей длины *п,* которые можно построить между 0 и x, то

$$-\mu\,[x] = \lambda\,(x;1) - \lambda\,(x;2) + \ldots \qquad (11)$$

Он доказал также, что если P есть структура, то μ [x]=0, если только x не является объединением точек.

Если P удовлетворяет условию Жордана—Дедекинда и $d[x]$ обозначает размерность элемента $x \in P,$ то можно связать с каждым $x \in P$ «характеристический многочлен», определенный посредством

$$p_x[\lambda] = \lambda^{d[x]+1} - \sum_{y<x} p_y[\lambda]. \qquad (12)$$

Этот многочлен связан с функцией Мёбиуса в силу того обстоятельства, что если $\mu_y[x]$ обозначает функцию Мёбиуса для x в подмножестве P_y элементов $x \geq y,$ то

$$p_x[\lambda] = \sum_{y \leq x} \lambda^{d[y]+1} \mu_y[x] = \mu[x]\,\lambda + \ldots \qquad (13)$$

Рассмотрим, например, проблему раскрашивания карты:

Назовем «подкартой» карты Ω карту, полученную из Ω стиранием границ. Подкарты любого Ω образуют частично упорядоченное множество с 0, удовлетворяющее условию цепей Жордана—Дедекинда; размерность любой подкарты на единицу меньше числа ее областей. Кроме того, имеется λ^n способов раскрашивания n областей в λ цветов, каждый из которых дает такое раскрашивание самого Ω или однозначно определяемой подкарты Ω, что никакие две смежные области не имеют одной и той же окраски. Отсюда в силу (12) получаем по индукции, что имеется точно p_Ω [λ] способов раскрашивания Ω в λ цветов таким образом, что никакие две смежные области не окрашиваются в один и тот же цвет.

6.3. Структуры

6.3.1. Определение

Общая теория частично упорядоченных множеств основывается на одном неопределенном отношении. Общая теория структур также основывается косвенным образом на этом отношении; однако непосредственно в ее основе лежат две бинарные операции, аналогичные во многих отношениях обычному сложению и умножению. Это та аналогия, которая делает теорию структур ветвью алгебры.

Под *верхней гранью* подмножества X частично упорядоченного множества P понимают элемент $a \in P$, содержащий каждое $x \in X$. *Наименьшая* верхняя грань есть верхняя грань, содержащаяся во всякой другой верхней грани. Понятия нижней грани и наибольшей нижней грани определяются двойственным образом. Очевидно в силу Р2, что подмножество частично упорядоченного множества может иметь самое большее одну наим. в. г. (наименьшую верхнюю грань) и одну наиб. н. г. (наибольшую нижнюю грань).

Определение. *Структура есть частично упорядоченное множество P, любые два элемента которого имеют наиб. н. г., или «пересечение»* $x \cap y$ *и наим. в. г., или «.объединение»* $x \cup y$.

Очевидно, что множество $\check{L}$, двойственное структуре L, есть снова структура, в которой пересечения и объединения меняются ролями. Легко также показать по индукции, что любое конечное подмножество имеет наиб. н. г. и наим. в. г. Частично упорядоченные множества, в которых *любое* подмножество имеет наиб. н. г. и наим. в. г., называются *полными* структурами.

Пусть X произвольное подмножество структуры L. Выберем $a_0 \in X$. Тогда или a_0 есть верхняя грань для X, или существует $a_1 \in X$ такое, что $a_1 \not\leq a_0$, в этом случае определим $b_1 = a_0 \cup a_1$. Далее, или b_1 есть верхняя грань для X, или существует $a_2 \not\leq b_1$; в этом случае определим $b_2 = b_1 \cup a_2 = (a_0 \cup a_1) \cup a_2$. Тогда или составляемая так цепочка $a_0 < b_1 < b_2 < \ldots$ закончится после конечного числа шагов на верхней грани b_n множества X, которая является *наименьшей* верхней гранью, или L содержит бесконечную возрастающую цепочку, $a_0 < b_1 < b_2 < \ldots < b_n < \ldots$ Следовательно, *X имеет наименьшую верхнюю грань или же L имеет бесконечную возрастающую цепочку.*

Отсюда вытекает, что если в структуре L все цепи конечны, то всякое подмножество в L имеет наим. в. г. и наиб. н. г. (L есть полная структура). В частности, само L имеет верхнюю грань I и нижнюю грань 0.
Мы будем употреблять термины «supremum» и «sup.» как синонимы понятия наим. в. г. и термины «infimum» и «inf.» как синонимы понятия наиб. н. г.

6.3.2. Примеры

Большая часть наиболее важных рассматриваемых в математике частично упорядоченных систем — структуры. Кроме того, в этих системах операции $\cap$ и $\cup$ соответствуют обычно хорошо известным и важным построениям.
Пример 1. Пусть $\sum$ состоит из всех подмножеств некоторой совокупности I и пусть включение означает теоретико-множественное включение. Тогда «объединения» являются суммами множеств, а «пересечения» — произведениями множеств.
Пример 2. Пусть J есть множество целых положительных чисел и пусть $m \leq n$ означает: m делит n. Тогда $m \cap n$ = общ. наиб. дел. (m, n) и $m \cup n$ = общ. наим. кр. (m,n).
Пример 3. Пусть $\sum$ состоит из подгрупп некоторой группы и пусть включение означает теоретико-множественное включение. Тогда термины «объединение» и «пересечение» имеют свое обычное значение.
Пример 4. Пусть H состоит из отношений конгруентуости на абстрактной алгебре A; пусть $\theta \leq \theta'$ означает, что θ есть подразбиение θ', т. е. что $x \equiv y (\mathrm{mod}\ \theta)$ влечет за собой $x \equiv y\ (\mathrm{mod}\ \theta')$. Тогда H является структурой. Кроме того, $\theta \cap \theta'$ определяет произведение разбиений θ и θ' в обычном смысле: $x \equiv y\ (\mathrm{mod}\ \theta \cap \theta')$ означает, что $x \equiv y\ (\mathrm{mod}\ \theta)$ и $x \equiv y\ (\mathrm{mod}\ \theta')$. В вырожденном случае, когда на A не определено ни одной операции, H является структурой всех разбиений множества A.

6.3.3. Структуры как абстрактные алгебры

Существует аналогия между выражениями $x \cup y$, $x \cap (y \cup z)$, встречающимися в теории структур, и подобными выражениями: $x+y$, $x(y+z)$, относящимися к обычной алгебре.
Легко, например, проверить как непосредственные следствия основного определения следующие тождества:

L1. $x \cap x = x$ и $x \cup x = x$,

L2. $x \cap y = y \cap x$ и $x \cup y = y \cup x$,

L3. $x \cap (y \cap z) = (x \cap y) \cap z$ и $x \cup (y \cup z) = (x \cup y) \cup z$,

L4. $x \cap (x \cup y) = x$ и $x \cup (x \cap y) = x$.

Эти тождества называются соответственно законами идемпотентности, коммутативности, ассоциативности и поглощения (названия в основном принадлежат Булю).

Теорема 1. *Тождества* LI — L4 *полностью характеризуют структуры*.

Доказательство. В любой системе, удовлетворяющей LI —L4, $x \cap y = y$ тогда и только тогда, если $x \cup y = x \cup (x \cap y) = x$; кроме того, если определить $x \geq y$ в смысле $x \cap y = y$, то получим *структуру,* в которой $x \cap y$ и $x \cup y$ являются соответственно наибольшей нижней и наименьшей верхней гранями для x и y. Например, из $x \cap x = x$ вытекает P1. Далее, если $x \geq y$ и $y \geq x$, то в силу предположения и L2 $y = x \cap y = y \cup x = x$, чем доказывается P2. В то же время если $x \geq y$ и $y \geq z$, то в силу L3 $x \cap z = x \cap (y \cap z) = (x \cap y) \cap z = y \cap z = z$, откуда $x \geq z$. Наконец, так как в силу LI — L3 $x \cap (x \cap y) = (x \cap x) \cap y = x \cap y$, то $x \cap y$ является нижней гранью для x; в силу L2 $x \cap y$ является также нижней гранью для y. Но $x \cap y$ есть *наибольшая* нижняя грань, так как из $x \geq z$ и $y \geq z$ вытекает, согласно L3, $(x \cap y) \cap z = x \cap (y \cap z) = x \cap z = z$. Двойственно этому, $x \cup y$ является наименьшей верхней гранью для x и y, что и завершает доказательство.

Заметим, что элементы 0 и I удовлетворяют соотношениям

$$0 \cap x = 0,\ 0 \cup x = x,\ x \cap I = x,\ x \cup I = I \text{ для всех } x. \quad (1)$$

Первые три из этих тождеств аналогичны законам $0x = 0$, $0+x = x$ и $x1=x$ обычной арифметики.

6.3.4. Подструктуры и многочлены

Аналогия между описанием структур при помощи LI — L4 и обычными определениями групп, колец и т. д. позволяет применить к структурам общую терминологию абстрактной алгебры.

Таким образом, нам следует определить *подструктуру* структуры L как подмножество, содержащее вместе с любыми двумя элементами их объединение и их пересечение. Подмножество структуры L может быть структурой по отношению к закону включения на L, но не быть подструктурой структуры L.

Определим также *структурный многочлен* как функцию переменных x_1, ..., x_n, являющуюся либо одним из x_i либо (по индукции) объединением или пересечением других структурных многочленов. Таким образом, структурные многочлены являются сложными функциями, полученными из первоначальных операций объединения и пересечения, и подструктура, порожденная некоторым подмножеством *X* структуры *L*, состоит из всех структурных многочленов, образованных из элементов множества *X*.

Так как структурные операции ассоциативны, мы можем определить по индукции, по аналогии с обычной П — ∑ системой обозначений,

$$\bigwedge_{i=1}^{n} x_i = x_1 \cap \dots \cap x_n \text{ и } \bigvee_{i=1}^{n} x_i = x_1 \cup \dots \cup x_n. \quad (2)$$

Эта система обозначений принадлежит Пирсу.

Рассмотрим некоторые правила, которым удовлетворяют структурные многочлены. Так, например, имеет место

Теорема 2. *Структурные многочлены являются изотопными функциями своих переменных: если $f(x_1, \dots, x_n)$ есть некоторый структурный многочлен, и $a_i \leq b_i$ для всех i, то $f(a_1, \dots, a_n) \leq f(b_1, \dots, b_n)$.*

Доказательство. Следуя методу индукции, достаточно показать, что из $x \leq x'$ вытекает $x \cap y \leq x' \cap y$ и $x \cup y \leq x' \cup y$. Но если $x \leq x'$, то

$$x \cap y = (x \cap x') \cap y = x \cap (x' \cap y) \leq x' \cap y$$

и двойственно этому

$$x' \cup y = (x' \cup x) \cup y = x' \cup (x \cup y) \geq x \cup y.$$

Как частный случай, получаем односторонние дистрибутивные законы

$$x \cap (y \cup z) \geq (x \cap y) \cup (x \cap z) \text{ и } x \cup (y \cap z) \leq (x \cup y) \cap (x \cup z). \quad (3)$$

В самом деле, по теореме 2, $x \cap (y \cup z)$ является верхней гранью как для $x \cap y$, так и для $x \cap z$, а потому содержит их объединение. Второй закон вытекает из двойственных соображений. Предполагая в (3) $x \geq z$, мы приходим к одностороннему модулярному закону:

$$\text{Если } x \geq z, \text{ то } x \cap (y \cup z) \geq (x \cap y) \cup z. \quad (4)$$

Более общим следствием является следующее неравенство минимакса:

$$\bigwedge_{i=1}^{m} \left(\bigvee_{j=1}^{n} a_{ij}\right) \geq \bigvee_{j=1}^{n} \left(\bigwedge_{i=1}^{m} a_{ij}\right). \quad (5)$$

Формулируя более полно, пусть a_{ij} определено для $i = 1, \dots, m$ и $j=1, \dots, n$, тогда наибольшая нижняя грань (когда изменяется i) наименьших верхних граней $\vee_j a_{ij}$ элементов a_{ij} (когда i фиксировано, а j изменяется) содержит в себе наименьшую верхнюю грань (когда изменяется j) наибольших нижних граней $\wedge_i a_{ij}$ элементов a_{ij} (когда j фиксировано, а i изменяется). Доказательство — непосредственное, ибо для всех i, j

$$b_i = \vee_j a_{ij} \geqslant a_{ij} \geqslant \bigwedge_i a_{ij} = c_j,$$

в силу чего $\wedge_i b_i$ является верхней гранью для c_j, а потому и для $\vee_j c_j$.

6.3.5. Гомоморфизмы и идеалы

Изоморфизм двух частично упорядоченных множеств P и P^* был определен ранее как взаимно однозначное соответствие, сохраняющее порядок, так что

$$x \geq y \text{ в } P \text{ тогда и только тогда, если } \theta(x) \geq \theta(y) \text{ в } P^*. \qquad (6)$$

Такое соответствие, если оно существует, должно сохранять объединения и пересечения, так что если P и P^* — структуры, то

$$\theta(x \cap y) = \theta(x) \cap \theta(y), \qquad (6')$$

$$\theta(x \cup y) = \theta(x) \cup \theta(y). \qquad (6'')$$

Автоморфизм структуры P есть ее изоморфизм с самой собой. Рассмотрим теперь однозначные в одну сторону соответствия θ: $L \rightarrow L^*$ между структурами. Как было впервые отмечено Орэ, эти соответствия могут удовлетворять условию

$$x \geq y \text{ влечет за собой } \theta(x) \geq \theta(y), \qquad (7)$$

но не удовлетворять ни условию (6'), ни условию (6″); они могут удовлетворять условию (6'), но не удовлетворять условию (6"); или удовлетворять условию (6"), но не удовлетворять условию (6'); или, наконец, удовлетворять обоим условиям. Мы будем называть такие соответствия соответственно изотонными, гомоморфизмами по пересечениям, гомоморфизмами по объединениям и структурными гомоморфизмами.

Легко показать, что из условия (6') или (6") вытекает (7). Так, например, поскольку $x \geq y$ означает, что $x \cap y = y$, получаем, используя (6'), $\theta(x) \cap \theta(y) = \theta(x \cap y) = \theta(y)$, откуда вытекает (7). Легко также показать, что множество прообразов элемента 0 при любом структурном гомоморфизме или гомоморфизме по объединениям является идеалом в следующем смысле.

Определение. *Подмножество J элементов структуры L есть идеал тогда и только тогда, если*

$$\text{из } x \in J \text{ и } y \in J \text{ вытекает, что } x \cup y \in J, \qquad (8)$$

$$\text{из } x \in J \text{ и } t \leq x \text{ вытекает, что } t \in J. \qquad (8')$$

Множество, двойственное идеалу, есть дуальный идеал.

Доказательство. Если $\theta(x) = \theta(y) = 0$, то $\theta(x \cup y) = \theta(x) \cup \theta(y) = 0 \cup 0 = 0$, и если $t \leq x$, то $\theta(t) \leq \theta(x) = 0$.

Легко также показать, что условие (8') эквивалентно условию

из $x \in J$ и $a \in L$ вытекает, что $a \cap x = x \cap a \in J$. (8")

Действительно, $x \cap a \leq x$ для всех a, а потому из (8') следует (8"); обратно, если $t \leq x$, то $t = x \cap)t$, а потому из (8") следует (8').

Если продолжать рассматривать $x \cap y$ как аналог xy и $x \cup y$ как аналог $x + y$, то условия (8), (8") будут в точности аналогами обычного определения идеала в кольце. Но в то время как в кольце каждый гомоморфный образ определяется с точностью до изоморфизма идеалом элементов, отображающихся в нуль, это уже неверно для многих структур. Например, цепь из трех элементов можно структурно-гомоморфно отобразить на самое себя и на цепь из двух элементов так, что единственным элементом, отображающимся в 0, будет 0.

Чтобы сделать ясной общую ситуацию, мы определим (точно так же, как в алгебраическом предисловии) *отношение конгруентности* на структуре L как отношение эквивалентности θ со свойством замены *(S)*. **Если** $x \equiv x^* (\theta)$ **и** $y \equiv y^* (\theta)$, **то** $x \cap y \equiv x^* \cap y^* (\theta)$ **и** $x \cup y \equiv x^* \cup y^* (\theta)$.

Отметим, что, как и всегда, отождествление конгруентных элементов приводит к структурно-гомоморфному образу структуры L и все структурно-гомоморфные образы L могут быть получены по такому способу.

В дальнейшем нам будет нужен следующий результат.

Лемма. *Если* $u == \upsilon \pmod\theta$ *в структуре* L, *то* $x \equiv y \pmod\theta$ *для всех* x, y *в интервале* $u \cap \upsilon \leq x, y \leq u \cup \upsilon$.

Доказательство. В силу предположений, $x = x \cup (u \cap v) \equiv$
$\equiv x \cup (u \cap u) \equiv x \cup u \pmod\theta$ и двойственно этому

$$x = x \cap (u \cup v) \equiv x \cap (u \cup u) \equiv x \cap u \pmod\theta.$$

Следовательно, $u = u \cap (u \cup x) \equiv u \cap x \equiv x \pmod\theta$. Подобным же образом $u \equiv y \pmod\theta$, откуда $x \equiv y \pmod\theta$.

Итак, прообразы любого элемента при структурном гомоморфизме образуют *выпуклую подструктуру,* т. е. подструктуру, содержащую вместе с любыми a и b все элементы между a и b.

Следствие. *Отношение конгруентности* θ *на структуре* L *конечной длины вполне определено, если всякий раз, когда* x *покрывает* y, *нам известно, имеет ли место* $x \equiv y \pmod\theta$ *или нет.*

Доказательство. Для любых $u, \upsilon \in L$ образуем цепь $u \cap \upsilon = a_0 < a_1 < \ldots < a_n = u \cup \upsilon$ такую, что a_i покрывает a_{i-1}.

Если $u \equiv \upsilon$ (mode), то в силу леммы все $a_i \equiv a_{i-1}$ (mod θ); обратно, если все $a_i \equiv a_{i-1}$ (mod θ), то $u \cap \upsilon \equiv u \cup \upsilon$(mod θ), откуда в силу леммы $u \equiv \upsilon$ (mod θ).

6.3.6. Структуры с дополнениями

В некоторых структурах всякое отношение конгруентности определяется полностью идеалом из элементов, конгруентных с 0.

Определение. *Под дополнением элемента x структуры L с O и I понимают элемент $y \in L$ такой, что $x \cap y = O$ и $x \cup y = I$; L называется структурой с дополнениями тогда и только тогда, если все ее элементы обладают дополнениями. Структура L называется структурой* с *относительными дополнениями, если все ее замкнутые интервалы являются структурами с дополнениями.*

Таким образом, L является структурой с относительными дополнениями тогда и только тогда, если для любых заданных $a \leq x \leq b$ существует такой элемент y, что $x \cap y = a$ и $x \cup y = b$. Этот элемент является «относительным дополнением» элемента x в замкнутом интервале $[a, b]$.

Теорема 3. *В структуре с относительными дополнениями, обладающей O, или даже в структуре с O, все замкнутые интервалы* $[0, a]$ *которой являются структурами с дополнениями, любое отношение конгруентности определяется идеалом из элементов* $x = 0(\text{mod}\theta)$.

Доказательство. Легко показать, что элементы $t \equiv 0(\theta)$ образуют идеал J_θ. В силу леммы 6.3.5 $x \equiv y(\theta)$ тогда и только тогда, если $x \cap y \equiv x \cup y$ (θ). Пусть $x \cap y \equiv v$, $x \cup y = u$; по предположению, существует такой элемент w, что $v \cap w = 0$, $v \cup w = u$. Кроме того, $u \equiv v$ (θ) влечет за собой $w = u \cap w \equiv v \cap w = 0(\theta)$ и $w \equiv 0(\theta)$ влечет за собой $u = v \cup w \equiv v \cup 0 = v(\theta)$. Таким образом, $x \equiv y(\theta)$ тогда и только тогда, если существует $w \in J_\theta$ такое, что $x \cap y \cap w = 0$ и $(x \cap y) \cup w = x \cup y$; следовательно, J_θ определяет θ.

Следующий результат имеет место не только для любой структуры, но даже для любой алгебры в смысле, изложенном в алгебраическом предисловии.

Теорема 4. *Пусть A — некоторая алгебра и пусть C — некоторое множество отношений конгруентности* θ *на A. Определим новые отношения* ξ и *η посредством* (I) $a \equiv b$ (ξ) *означает, что* $a \equiv b(\theta)$ *для всех* $\theta \in C$, (II) $a \equiv b(\eta)$ *означает, что для некоторой конечной* *последовательности* $a = x_0, x_1, \ldots, x_m = b$, $x_{i-1} \equiv x_i(\theta_i)$ *для некоторого* $\theta_i \in C$. *Тогда* ξ *и* η *будут отношениями конгруентности; кроме того,* ξ *есть наиб.н. г., а* η *есть наим. в. г. для элементов* $\theta \in C$.

Замечание. Отношения конгруентности на A считаются частично упорядоченными.

Доказательство. Очевидно, что $\xi \leq \theta \leq \eta$ для всех $\theta \in C$; очевидно, далее, что если $\xi' \leq \theta' \leq \eta'$ для всех $\theta \in C$ и ξ', η' суть отношения конгруентности, то $\xi' \leq \xi$ и $\eta' \geq \eta$. Легко, кроме того, проверить, что ξ и η являются отношениями эквивалентности: рефлексивность и симметричность проверяются непосредственно; транзитивность для ξ также проверяется непосредственно в предположении, что она имеет место для всех $\theta \in C$; транзитивность η вытекает из определения, так как если $a \equiv b(\eta)$ и $b \equiv c(\eta)$, то существует конечная последовательность $a = x_0, x_1, \ldots, x_m = b = y_0, y_1, \ldots, y_n = c$ такая, что $x_{i-1} \equiv x_i(\theta_i)$ для некоторого $\theta_i \in C$ и $y_{j-1} \equiv y_j(\theta'_j)$ для некоторого $\theta'_j \in C$ $[i = 1, \ldots, m;\ j = 1, \ldots, n]$. Остается доказать свойство замены для ξ и η.

Очевидно, далее, что ξ обладает свойством замены: если $x_i \equiv y_i(\xi)$ для всех i и f есть некоторая n-арная операция на A, то $f(x_1, \ldots, x_n) \equiv f(y_1, \ldots, y_n)$ (ξ), так как для всех i $x_i \equiv y_i(\theta)$, а потому $f(x_1, \ldots, x_n) \equiv f(y_1, \ldots, y_n)$ для всех $\theta \in C$. Это верно даже в том случае, если операции на A не являются финитарными. Наконец, чтобы показать, что η обладает свойством замены для f, заметим следующее: *если* $a_i = b_i(\eta)$ для $i = 1, \ldots, n$, то, согласно (II), мы можем образовать цепи $a_i = x_{i0}, x_{i1}, \ldots, x_{im} = b_i$ так, что $x_{i,j-1} \equiv x_{i,j}$ $(\theta_{i,j})$. Мы можем заменять тогда каждое $x_{i,j-1}$ на $x_{i,j}$ не изменяя при этом значения $f(a_1, \ldots, a_n)$ по mod η; после mn шагов мы получим $f(b_1, \ldots, b_n)$.

Теорема 5. *Отношения конгруентности на произвольной структуре L удовлетворяют бесконечному дистрибутивному закону,*

$$\theta \cap \bigvee_C \theta_\gamma = \bigvee_C (\theta \cap \theta_\gamma). \tag{9}$$

Доказательство. В силу (3'), и Р2 достаточно доказать, что

$$\text{из } a \equiv b(\text{mod}\,\theta \cap \bigvee_C \theta_\gamma) \text{ следует } a \equiv b\ (\text{mod} \bigvee_C (\theta \cap \theta_\gamma)). \tag{9'}$$

В обозначениях $\eta = \vee\theta_\gamma$ теоремы 4, это есть утверждение, что из $a \equiv b$ (θ) и $a \equiv b$ (η) вытекает существование последовательности $a = y_0, y_1, \ldots, y_m = b$ с $y_{i-1} \equiv y_i(\theta)$ и $y_{i-1} \equiv y_i(\theta_i)$ для некоторого $\theta_i \in C$. Но $a \equiv b$ (η) означает, что существует последовательность $a = x_0, x_1, \ldots, x_m = b$ с $x_{i-1} \equiv x_i(\theta_i)$. Положим теперь $y_i = [(a \cap b) \cup x_i] \cap (a \cup b)$: очевидно, что $y_0 = a$, $y_m = b$, $y_{i-1} \equiv y_i(\theta)$ в силу леммы 6.3.5 и что $y_{i-1} \equiv y_i(\theta_i)$, так как θ обладает свойством замены.

6.3.7. Произведения и степени

В приведенной ранее терминологии, «кардинальная сумма» двух структур не может быть структурой, ибо элементы x и y, взятые из различных слагаемых, не обладают верхней гранью. С другой стороны, «ординальная сумма» $L \oplus M$ двух структур является всегда структурой.

Теорема 6. *Кардинальное «произведение» LM двух структур L и M есть прямое произведение структур, рассматриваемых как абстрактные алгебры; следовательно, оно всегда является структурой. Если M — структура, а X — произволъное частично упорядоченное множество, то M^X также является структурой.*

Доказательство. Пусть $[x, y]$ и $[x', y']$—два произвольных элемента из LM. Легко показать тогда, что $[x \cup x', y \cup y']$, где $x \cup x' \in L$ и $y \cup y' \in M$, будет верхней гранью для $[x, y]$ и $[x', y']$, содержащейся во всякой другой верхней грани. Доказательство того, что LM есть структура, завершается путем двойственных рассуждений. Рассмотрим теперь множество M^X. Пусть $y = f(x)$ и $y = g(x)$ — два произвольных изотонных однозначных отображения X в M. Если мы для всех x определим $h(x) = f(x) \cup g(x) \in M$, то получим изотонную однозначную функцию h, являющуюся наименьшей верхней гранью для f и g в M^X. Двойственным образом $h^*(x) = f(x) \cap g(x)$ будет наибольшей нижней гранью, что и завершает доказательство.

6.3.8. Теорема однозначности разложения на множители

Пусть P — какое-нибудь произведение частично упорядоченных множеств X_i; при этом, мы предполагаем, что P имеет 0 и I. Очевидно, что каждая компонента элемента $I \in P$ есть I для X_i и что для 0 справедливо двойственное утверждение; следовательно, все X_i имеют 0 и I.

Определим e_i как элемент из P, у которого компонента в X_i есть I, а все остальные компоненты суть 0. Тогда элементы $t \le e_i$ из P будут образовывать подмножество X^*_i множества P, изоморфное X_i. Кроме того, для любого элемента $a=[a_1, a_2, a_3,...] \in P$ его компонента в X_i есть $a_i = a \cap e_i$; таким образом, это пересечение существует и является компонентой a_i в X_i. Остальные компоненты элемента at суть 0. Следовательно, $a_i \in X^*_i$ и $a=\sup_i a_i$. Далее, элементы $t \le a$ в P представляют собой такие элементы, каждая компонента t_i которых содержится в соответствующей компоненте элемента a. Следовательно,

Лемма. *Множество A элементов $t \le a$ в P является произведением сомножителей, изоморфных множествам A_i элементов $t \le a_i$.*

Предположим теперь, что P можно разложить в произведение множителей X_i, а также в произведение множителей Y_j. Определим элементы e_i так, как это было сделано выше, и аналогичным образом определим элементы e'_j так, что $x \le e'_j$ образуют множество Y_j*, изоморфное Y_j. Обозначим, наконец, через Z^i_j множество элементов $t \le e_i \cap e'_j$ в P. Тогда в силу леммы каждое X^*_i является произведением Z^i_j с верхним индексом i, а каждое Y_j*—произведением Z^i_j с нижним индексом j. Мы получаем

Теорема 7. *Любым двум разложениям частично упорядоченного множества P с 0 и I соответственно на множители X_i и Y_i можно сопоставить разложение P на множители Z^i_j такое, что произведение Z^i_j при фиксированном i есть X_i и произведение Z^i_j при фиксированном j есть Y_j.*

Следствие. *Если P можно разложить на неразложимые множители, то это разложение единственно в том точном смысле, что любое разложение P на множители можно получить путем группировки этих неразложимый множителей в подсемейства.*

Теорема 8 (Дилуорс). *Всякая структура L с относительными дополнениями, имеющая конечную длину, является кардинальным произведением «простых» структур с относительными дополнениями.*

Доказательство. Пусть θ — какое-нибудь отношение конгруентности; J — идеал элементов $x \equiv 0 \pmod{\theta}$ и J'—дуальный идеал элементов $x \equiv I \pmod{\theta}$. Так как L имеет конечную длину, J есть главный идеал с наибольшим элементом a, а J' — главный идеал с наименьшим элементом a'.

Рассмотрим соответствие $x \to x \cup a$. Если $x \equiv y \pmod{\theta}$, то в силу теоремы 3 $x \cap y \cap w = 0$ и $(x \cap y) \cup w = x \cup y$ для некоторого $w \le a$. Следовательно, $(x \cap y) \cup a = (x \cap y) \cup w \cup a = x \cup y \cup a$, откуда $x \cup a = y \cup a$. Обратно, если $x \cup a = y \cup a$, то $x \equiv x \cup a = y \cup a = y \pmod{\theta}$. Поэтому соответствие $x \to x \cup a$ отображает x в наибольший элемент класса вычетов, содержащего x, и является идемпотентным *структурным эндоморфизмом.* Двойственно этому, $x \to x \cup a'$ является идемпотентным структурным эндоморфизмом, отображающим каждое x в наименьший элемент класса вычетов, содержащего x. Отсюда мы получаем, в частности, $a' \cup a = I$ и $a \cap a'=0$; таким образом, a и a' являются дополнениями.

Далее, если $x \cap a' = 0$, то $a = 0 \cup a = (x \cap a') \cup a = (x \cup a) \cap (a' \cup a) = x \cup a$, а потому $x \leq a$. (В этом смысле a является *сильным* дополнением a'.) Но теперь для *любого* x элемент $a \cap x$ имеет относительное дополнение t в x такое, что $(a \cap x) \cap t = 0$ и $(a \cap x) \cup t = x$. Очевидно, что $a \cap (x \cap t)=0$, откуда $t = x \cap t \leq a'$; кроме того, $t \leq x$. Следовательно,

$$(a \cap x) \cup (a' \cap x) \geqslant (a \cap x) \cup t = x.$$

Отсюда следует, что соответствие $x \rightarrow (a \cap x,\ a' \cap x)$ есть изоморфизм L с кардинальным произведением J и идеала элементов $x \leq a'$. Следовательно, если L не просто, оно является кардинальным произведением, что и требовалось доказать.

6.3.9. Центр структуры

Предшествующее доказательство позволяет сформулировать определение *центра* частично упорядоченного множества P с O и I как множества элементов $e \in P$, у которых при некотором разложении P в произведение одна X_i-я компонента есть I, а остальные — O. Так как кардинальное умножение структур коммутативно и ассоциативно, мы видим, что для этого необходимо и достаточно, чтобы при некотором представлении $P = X_i Z$ множества P в виде произведения двух сомножителей имело место $e = [I,O]$. Используя это, сформулирем лемму

Лемма. *Каждый элемент центра имеет одно единственное дополнение, которое также принадлежит центру. Центр любого частично упорядоченного множества сохраняется при дуальных автоморфизмах.*

Доказательство. Очевидно, что $[I, O] \cap [x, y]=[x, O]$ и $[I, O] \cup [x\ y] = [I, y]$, следовательно, $[x, y]$ является дополнением элемента $[I, O]$ тогда и только тогда, если $[x, y] = [O, I]$, которое, таким образом, единственно. Далее, P есть произведение множества элементов $x \leq [I, O]$ и множества элементов $y \leq [O, I]$ тогда и только тогда, когда оно есть произведение множества элементов $s \geq [I, O]$ и множества элементов $t \geq [O, I]$; следовательно, если элемент $[I, O]$ лежит в центре множества L, то он лежит также в центре множества, двойственного L.

Заметим, что доказательство теоремы 7 показывает, что если e_i и e'_j лежат в центре множества L, то $e_i \cap e'_j$ существует и также лежит в центре. По двойственности (т. е. в силу предшествующей леммы) то же самое справедливо для $e_i \cup e'_j$. Отсюда можно сформулировать теорему:

Теорема 9. *Центр всякой структуры с O* и *I является подструктурой, инвариантной при дуальных изоморфизмах.*

6.3.10. Нейтральные элементы

Структура L называется *дистрибутивной* тогда и только тогда, если для любых $x, y, z \in L$

$$x \cap (y \cup z) = (x \cap y) \cup (x \cap z) \text{ и } x \cup (y \cap z) = (x \cup y) \cap (x \cup z). \quad (10)$$

Дистрибутивные структуры рассматриваются позже.

Определение. *Элемент a структуры L называется нейтральным тогда и только тогда, если каждая тройка элементов {a, x, y} порождает дистрибутивную подструктуру структуры L.*

Лемма. *Если элемент a нейтрален, то* а) *двойственные отображения $x \to x \cap a$ и $x \to x \cup a$ являются эндоморфизмами структуры L и* б) *из $x \cap a = y \cap a$ и $x \cup a = y \cup a$ вытекает $x=y$.*

Доказательство. Для доказательства «а» заметим, что в силу L1 —L3 $(x \cap y) \cap a = (x \cap a) \cap (y \cap a)$ и в силу (10) $(x \cup y) \cap a = (x \cap a) \cup (y \cup a)$ — и двойственным образом. Для доказательства «б» заметим, что в силу непосредственных выкладок, при использовании условий (10) дважды, $x = x \cap (x \cup a) = x \cap (y \cup a) = (x \cap y) \cup (x \cap a) = (x \cap y) \cup (y \cap a) = y \cap (x \cup a) = y \cap (y \cup a) = y$.

Из леммы вытекает, что отображение $x \to [x \cap a, x \cup a]$ структуры L есть изоморфизм между L и подструктурой кардинального произведения AB структуры (идеала) A элементов $x \leq a$. и структуры (дуального идеала) B элементов $x \geq a$. Кроме того, при этом отображении $a \to [a, a] = [I, 0]$, так как $a = I$ в A и $a = 0$ в B. Обратно, элемент $[I, 0]$ нейтрален в AB или в некоторой подструктуре структуры AB, так как $\{I, x, y\}$ и $\{0, x, y\}$ удовлетворяют всегда условию (10); мы предоставляем читателю проделать двенадцать необходимых выкладок; в силу двойственности и симметрии они сводятся к четырем. Следовательно,

Теорема 10. *Элемент структуры L нейтрален тогда и только тогда, если при изоморфизме между L и подструктурой некоторого (кардинального) произведения AB он переводится в элемент [I, O].*

Следствие. *Элемент принадлежит центру структуры L тогда и только тогда, если он нейтрален и обладает дополнением.*

Доказательство. В силу леммы 6.3.9 всякий элемент из центра структуры L обладает дополнением; в силу теоремы 10 он нейтрален. Обратно, в силу теоремы 9 если нейтральный элемент $[I,0]=a$ обладает дополнением $a' = [x, y]$, то $[0,0] = = [I,0] \cap [x, y] — [x,0]$, откуда $x = 0$; двойственным образом, $y = I$. Следовательно, $[x, y] = [0, I]$, и (в силу теоремы 9) этот элемент нейтрален. Ясно теперь, что соответствие $z \to (z \cap a,\ z \cap a')$ представляет L в виде кардинального произведения идеала A элементов $t \leq a$ и идеала A' элементов $t \leq a'$ и что, следовательно, элементы $a = [I,0]$ и $a' = [0, I]$ принадлежат центру структуры L.

Теорема 11. *Множество нейтральных элементов структуры L есть пересечение ее максимальных дистрибутивных подструктур.*

Доказательство. Если элемент a не нейтрален, то некоторая тройка элементов $\{a, x, y\}$ не дистрибутивна. Следовательно, ни одна максимальная дистрибутивная подструктура, получаемая путем расширения дистрибутивной подструктуры, порожденной $\{x, y\}$, не может содержать a.

Отсюда вытекает, что пересечение максимальных дистрибутивных подструктур структуры L' не содержит ненейтральных элементов.

Обратно, пусть S — произвольная максимальная дистрибутивная подструктура структуры L; пусть элемент a нейтрален. Легко показать, используя теорему 10, что подструктура $\{a, S\}$ структуры $L \leq AB$, порожденная элементом $a = [I,0]$ и S, дистрибутивна. Так как подструктура S максимальна, отсюда следует, что $a \in S$

Следствие. *Во всякой структуре нейтральные элементы образуют дистрибутивную подструктуру.*

6.3.11. Свободные структуры

«Свободная» структура FL(к) с к образующими была определена Уитменом для произвольного кардинального числа к. Мы строим ее следующим образом.

Прежде всего, FL(к) содержит «образующие» x_α, где индексы α пробегают множество из к элементов; так, если к конечно, то можно писать $x_1, ..., x_n$. Эти образующие будем называть также «выражениями веса ноль». Определим тогда по индукции «выражение веса w» как выражение вида $p \cap q$ или $p \cup q$, где сумма весов p и q есть $w—1$. Таким образом, «вес» любого выражения есть общее число связок $\cap$ и $\cup$ встречающихся в этом выражении. Элементами структуры FL(к) являются все «выражения», однако не предполагается, что различные выражения представляют собой различные элементы.

Теперь мы квазиупорядочим множество FL(к) путем систематического повторения следующих четырех основных правил:

$$p \cup q \leq a, \text{ если } p \leq a \text{ и } q \leq a, \tag{11}$$

$$b \leq p \cap q> \text{ если } b<p \text{ и } b \leq q, \tag{11'}$$

$$p \cap q \leq a, \text{ если } p \leq a \text{ или } q \leq a, \tag{12}$$

$$b \leq p \cup q \text{ если } b \leq p \text{ или } b \leq q. \tag{12'}$$

Очевидно, что (11) и (11') двойственны и что (12) и (12') также двойственны. Если, например, мы хотим установить, будет ли $p \cup q \leq r \cup s$, мы производим сперва проверку (11) для $p \cup q$, рассматривая $r \cup s$ как *а,* а затем проверку (12') для $r \cup s$, рассматривая $p \cup q$ как *b*. Следовательно, $p \cup q \leq r \cup s$ тогда и только тогда, если а) $p \leq r \cup s$ и $q \leq r \cup s$ или б) $p \cup q \leq r$ или в) $p \cup q \leq s$. Таким способом проверка соотношения $p \cup q \leq r \cup s$ сводится к проверке комбинации из четырех или меньшего числа соотношений, в каждом из которых полный вес входящих в него выражений на единицу меньше. Следовательно, продолжая этот процесс сокращения, *мы можем проверить справедливость соотношения а ≤ b, если сумма весов а и b есть w+w', путем испытания* $4^{w+w'}$ *или меньшего числа элементарных соотношений* вида $x_i \leq x_j$ (которые верны тогда и только тогда, если $i = j$).

Лемма 1. FL(к) есть квазиупорядоченноt множество.

Доказательство. Из $p \leq p$ и $q \leq q$ вытекает в силу (12') $p \leq p \cup q$ и $q \leq p \cup q$, откуда в силу (11) $p \cup q \leq p \cup y$. Следовательно, по соображениям двойственности и индукции, отношение рефлексивно. Докажем теперь транзитивность: из $a \leq b$ и $b \leq c$ вытекает $a \leq c$. Рассмотрим сперва случай, когда одно из крайних выражений *а, с* участвует в приведении одного или обоих неравенств.

Если $p \cup q \leq b$, согласно (11), и $b \leq c$, то $p \leq b \leq c$ и $q \leq b \leq c$, откуда по индукции $p \leq c$ и $q \leq c$, а потому в силу (11) $p \cup q \leq c$. Далее, если $p \cap q \leq b$, согласно (12), и $b \leq c$, то $p \leq b \leq c$ или $q \leq b \leq c$, откуда, по индукции, $p \leq c$ или $q \leq c$, а потому в силу (12) $p \cap q \leq c$. Двойственно этому, если $a \leq b \leq p \cap q$ или $a \leq b \leq p \cup q$ вытекает из разложения $p \cap q$, соответственно $p \cup q$, то мы можем доказать, что $a \leq p \cap q$, соответственно $a \leq p \cup q$.

Остается случай, когда оба неравенства $a \leq b$ и $b \leq c$ приводятся путем записи $b = p \cup q$ (или $b = p \cap q$). Если $a \leq p \cup q$, согласно (12'), и $p \cup q \leq c$, согласно (11), то $a \leq p$ или $a \leq q$ и $p \leq c$ и $q \leq c$. Следовательно, $a \leq p \leq c$ или $a \leq q \leq c$; в каждом из этих случаев в силу индуктивных соображений $a \leq c$. Случай $a \leq p \cap q \leq c$ является двойственным предыдущему, чем и завершается доказательство. Используя теперь теорему 3 гл. I, мы выводим следующее

Следствие. *Если мы определим $a=b$ в* FL*(к) в том смысле, что $a \leq b$ и $b \leq a$, то* FL*(к) будет частично упорядоченным множеством.*

Лемма 2. FL*(к) является структурой, в которой $a \cap b$ есть наиб. н. г. (a, b), а $a \cup b$ есть наим. в. г. (a, b).*

Доказательство. В силу (12) и леммы 1 $a \cap b \leq a$ и $a \cap b \leq b$. В силу (11') из $x \leq a$ и $x \leq b$ вытекает, что $x \leq a \cap b$. Следовательно, $a \cap b$ есть наиб. н. г. (a, b). Двойственно этому $a \cup b$ есть наим. в. г. (a, b).

Теорема 12. FL*(к) есть свободная структура, порожденная символами x_α.*

Доказательство. Выберем в произвольной структуре L элементы g_α, соответствующие элементам x_α. Посредством прямой подстановки каждое $a \in$ FL*(к)* определяет единственный элемент $a^* \in L$. В силу леммы 2 это соответствие есть гомоморфизм FL*(к)* на подструктуру структуры L, порожденную элементами g_α.

6.4. Цепи

6.4.1. Цепи вещественных чисел

Ранее *цепь* была определена как частично упорядоченное множество, удовлетворяющее P4. Для любых заданных x и y или $x \leq y$, или $y \leq x$. Возможны и другие эквивалентные определения.

Каждая конечная цепь из n элементов изоморфна в силу теоремы 6 гл. I с цепочкой целых положительных чисел 1, ...,n. Эту цепь мы будем обозначать жирным шрифтом $\boldsymbol{n}$. Мы будем обозначать также через $\boldsymbol{\omega}$ цепь всех положительных целых чисел, через $\boldsymbol{J}$ цепь всех целых чисел, через $\boldsymbol{R}$ цепь всех рациональных чисел и через $\boldsymbol{R}$# цепь всех вещественных чисел. Многие цепи могут быть представлены как подцепи последней, а следовательно, могут быть изображены как множества точек на прямой. Такое графическое изображение очень удобно. Оно позволяет, например, наглядно представить себе смысл следующего определения:

Определение. *Цепь C называется плотной в себе тогда и только тогда, если для любых заданных $a < b$ из C существует $c \in C$, удовлетворяющее соотношению $a < c < b$. Подмножество S цепи C называется плотным по упорядоченности в C тогда и только тогда,* *если для любых a < b, принадлежащих C, но не принадлежащих S, можно найти элемент s в S, удовлетворяющий соотношению a < s < b.*

Теорема 1. *Любая счетная цепь изоморфна подцепи цепи **R** рациональных чисел. Любая плотная в себе счетная цепь изоморфна или **R** или **R** $\oplus$ **1**, или **1** $\oplus$ **R** или же **1** $\oplus$ **R** $\oplus$ **1**.*

Доказательство. Рациональные числа можно пронумеровать в виде $r_1, r_2, r_3, \ldots;$ аналогичным образом элементы заданной цепи можно пронумеровать в виде $a_1, a_2, a_3, \ldots$ Положим $f(a_1) = 0$ и определим по индукции $f(a_n)$ для любого n. Мы знаем, что или а) a_n превышает $a_1, \ldots, a_{n-1}$, или б) a_n меньше, чем $a_1, \ldots, a_{n-1}$; или в) a_n лежит между a_i и a_j для некоторого наибольшего $a_i < a_n$ и наименьшего $a_j > a_n$ $[i, j < n]$. В случае «а» положим $f(a_n) = n;$ в случае «б» положим $f(a_n) = —n$; в случае «в» определим $f(a_n)$ как первое r_i, расположенное в смысле порядка относительно $f(a_1), \ldots, f(a_{n-1})$ так же, как элемент a_n относительно $a_1; \ldots, a_n$. Так как **R** «плотно в себе», такое $f(a_n)$ всегда существует. Очевидно, что соответствие изотонно и что $f(a_n)$ отлично от $f(a_1), \ldots, f(a_{n-1})$. Кроме того, если **C** плотно в себе, то образ **C** содержит все рациональные числа между некоторыми $—m$ и n: он не может не содержать какого-нибудь (первого) r_i между двумя рациональными числами вида $f(a_n)$. Здесь $—m$ или n может быть бесконечным, и в соответствии с этим имеют место четыре случая.

Заметим, далее, что **R** плотно по упорядоченности в **R**#. Этот факт является типичным в следующем смысле.

Теорема 2. *Цепь **C** изоморфна подцепи цепи **R**# тогда и только тогда, если **C** содержит плотное по упорядоченности счетное подмножество.*

Доказательство. Предположим, что **C** содержит плотное по упорядоченности счетное подмножество $A = \{a_1, a_2, a_3, \ldots\}$; не нарушая общности, мы можем предположить, что A содержит наибольший и наименьший элементы множества C, если только они существуют. В силу теоремы 1 мы можем отобразить A на подцепь цепи рациональных чисел. В силу предположения и определения 1 каждое c, принадлежащее $C — A$ (каждое c, принадлежащее C, но не принадлежащее A), определяется однозначно *сечением,* которое оно задает на множестве A, т. е. разбиением A на множество элементов $a_i < c$ и множество элементов $a_i > c$. Пусть r_1 = наим. в. г. $f(a_i)$ для $a_i < c$ и r_2= наиб. н. г. $f(a_i)$ для $a_i > c$; мы определяем тогда $2f(c) = r_1 + r_2$. Это определяет изоморфизм между **C** и подмножеством множества **R**#..

Обратно, если **C** изоморфно подмножеству множества **R**#, то мы можем найти счетное плотное подмножество множества C, перенумеровывая интервалы I_i: $m_i / n_i \leq x \leq m_i' / n_i'$ множества **R**#, имеющие рациональные концы, и выбирая по одному c_i из каждого I_i, исключая те случаи, когда ни один элемент из C не соответствует элементу из I_i

6.4.2. Полная упорядоченность

Фундаментальным понятием в теории бесконечных цепей является понятие вполне упорядоченного множества.

Определение. Частично упорядоченное множество $\boldsymbol{W}$ называется вполне упорядоченным, или порядковым, тогда и только тогда, если каждое не пустое его подмножество имеет первый (т. е. наименьший) элемент.

В применении к подмножествам из 2 элементов это означает что $\boldsymbol{W}$ должно быть цепью (при заданных $a \neq b$ — $a < b$ или $b < a$). Кроме того, в силу теоремы 6 гл. 2 для случая конечных множеств понятие вполне упорядоченного множества и понятие цепи эквивалентны между собой. Непосредственно видно, что любое подмножество вполне упорядоченного множества само вполне упорядочено.

Свойство быть вполне упорядоченным не двойственно себе в бесконечном случае; в самом деле, множество $\boldsymbol{W}$ и двойственное ему множество $\breve{W}$ являются оба вполне упорядоченными тогда и только тогда, если $\boldsymbol{W}$ есть конечная цепь.

Простейшее бесконечное вполне упорядоченное множество дается последовательностью ω целых,положительных чисел 1, 2, 3, ... Утверждение, что это есть полная упорядоченность, является одной из форм принципа конечной индукции. Этот принцип можно расширить до следующего принципа.

Принцип трансфинитной индукции. *Пусть* $\{P_\alpha\}$— *вполне упорядоченное множество утверждений. Если мы можем доказать для каждого* α, *что из справедливости всех* P_β *при* $\beta < \alpha$ *вытекает справедливость* P_α, *то каждое* P_α *справедливо.*

Доказательство. Не может быть первого неверного P_α; следовательно, множество всех неверных P_α пусто.

Мы можем сформулировать этот принцип в иной форме. Под *предельным числом* во вполне упорядоченной последовательности мы понимаем элемент α такой, что для любого заданного $\beta < \alpha$ существует γ, удовлетворяющее соотношению $\beta < \gamma < \alpha$. Тогда, по определению, мы получаем как следствие

Второй принцип трансфинитной индукции. *Пусть* $\{P_\alpha\}$—*вполне упорядоченное множество утверждений. Чтобы доказать справедливость всех* P_α, *достаточно доказать, что* а) P_1 *справедливо,* б) *если справедливо P_α, то справедливо $P_{\alpha+1}$*, в) *если* α—n *предельное число и все P_β при* $\beta < \alpha$ *справедливы, то P_α справедливо.*

Другие вполне упорядоченные множества можно построить путем выполнения ординальных операций.

Теорема 3. *Ординальные сумма и произведение двух любых вполне упорядоченных множеств V, W сами вполне упорядочены.*

Доказательство. Если *S*—какое-нибудь непустое подмножество множества $V\oplus W$, то или $S\leq W$ и содержит первый элемент, или пересечение $S \cap V$ множеств *S* и *V* непусто и содержит первый элемент, который является (так как $\upsilon<w$ для любых $\upsilon\in V$, $w\in W$) первым элементом в *S*. Далее, если *S*— какое-нибудь непустое подмножество множества $V\circ W$, то множество *T* элементов $\upsilon \in V$ таких, что $(\upsilon,w) \in S$ для некоторого *w* непусто и должно иметь первый элемент υ_1. Множество элементов $w\in W$, таких, что $(\upsilon_1,w)\in S$ непусто по построению; пусть w_1—первый элемент этого множества. Тогда (υ_1, w_1) есть первый элемент в *S*.

Вследствие этого мы можем строить много вполне упорядоченных можеств: $\omega\oplus 1$, $\omega\oplus 2$, ..., $\omega\oplus\omega = 2\circ\omega$, $(2\circ\omega)\oplus 1$ и т. д. Эти множества принято записывать, используя обычные (т. е. кардинальные) символы сложения и умножения, например (ω +1, ω + 2, ..., 2 ω, 2 ω +1, ..., но мы будем придерживаться наших общих обозначений.

Пусть *W*—какое-нибудь вполне упорядоченное множество. Тогда *W* должно содержать первый элемент w_1;, W — w_1 должно содержать первый элемент w_2 и т. д. Следовательно, *если W бесконечно, то оно должно содержать начальный интервал, изоморфный* ω, состоящий из элементов w_l, w_2, w_3, ... Следовательно, мы можем писать $W=\omega \oplus R$, где *R* есть остаточный интервал. Повторяя процесс, мы будем иметь или $W = \omega \oplus \boldsymbol{n}$ для некоторого конечного упорядоченного множества $\boldsymbol{n}$, или же $W= \omega \oplus \omega \oplus S= (2\circ\omega)\oplus S$, где S — остаточный интервал.

Этот процесс мы можем продолжить, чтобы показать, что $\boldsymbol{W}$ или имеет вид $(\boldsymbol{m}\circ\omega) \oplus \boldsymbol{n}$, или же содержит начальный отрезок, изоморфный ${}^2\omega$ (т. е. вполне упорядоченному множеству ω, возведенному в ординальную степень, с показателем —порядковым числом два). Продолжая дальше, мы получим *порядковые полиномы* относительно ω

$$\boldsymbol{W} = \boldsymbol{a}_n \circ {}^n\omega \oplus \boldsymbol{a}_{n-1} \circ {}^{n-1}\omega \oplus \dots \oplus \boldsymbol{a}_0, \qquad (1)$$

где $\boldsymbol{a}_0, \ldots, \boldsymbol{a}_n$ — конечные порядковые множества. Построение показывает, что каждое вполне упорядоченное множество, не изоморфное порядковому полиному относительно ω, содержит начальные отрезки, изоморфные каждому такому полиному и в этом смысле «больше» чем все эти полиномы. Построение указывает далее на возможность того, что сами вполне упорядоченные множества образуют вполне упорядоченное множество

$\mathbf{1}, \mathbf{2}, \mathbf{3}, \ldots; \omega, \omega \oplus \mathbf{1}, \omega \oplus \mathbf{2}, \ldots; \mathbf{2} \circ \omega, (\mathbf{2} \circ \omega) \oplus \mathbf{1}, \ldots;$ $\mathbf{3} \circ \omega, (\mathbf{3} \circ \omega) \oplus \mathbf{1}, \ldots$ Мы докажем это формально в 6.4.3.

Учтем сперва, что было нами проделано до сих пор. Мы продемонстрировали во вполне конкретной форме большое число счетных вполне упорядоченных множеств. Однако мы были не в состоянии построить какое-нибудь несчетное вполне упорядоченное множество. *Проблема «построения» несчетного вполне упорядоченного множества является, повидимому, наиболее важной проблемой теории множеств.*

6.4.3. Фундаментальная теорема о вполне упорядоченных множествах

Получим теперь некоторые свойства вполне упорядоченных множеств, основываясь на построениях, включающих в себя *идеалы.* Идеал цепи *С* есть подцепь цепи С, содержащая вместе с любым $c \in$ С все $x < c$, т. е. идеал есть начальный интервал. Пустое множество не исключается.

Лемма 1. *Идеалами вполне упорядоченного множества W являются само W и множества элементов x <* а *для каждого фиксированного* $a \in W$.

Доказательство. Для каждого $a \in W$ множество элементов $x<a$ является идеалом. Обратно, если J — идеал в W, отличный от W, то должен существовать первый элемент a из W, не принадлежащий J; очевидно, что J содержит все $x < a$, но не содержит a и элементов $c>a$.

Лемма 2. *Существует самое большее один изоморфизм между идеалами J и К вполне упорядоченных множеств V, W.*

Доказательство. Предположим, что их имеется два, скажем, θ и θ*. Так как J вполне упорядочено, должно иметься первое $a \in J$, такое, что $\theta(a) \neq \theta^*(a)$, скажем, $\theta(a) > \theta^*(a)$. Тогда образ $\theta(J)$ идеала J не может содержать $\theta^*(a)$, так как $\theta(b)>\theta^*(a)$, если $b \geq a$, и $\theta(b)>\theta^*(a)<\theta^*(a)$, если $b < a$. Однако $\theta(a) > \theta^*(a)$; следовательно, $\theta(J)$ не является идеалом, что противоречит предположению.

Отметим теперь, что если θ_α является для каждого α однозначным отображением подмножества S_α. множества X в множество Y и если для любых α и β $\theta_\alpha(x) = \theta_\beta(x)$ всякий раз, когда они оба определены, то θ_α имеют наибольшее однозначное *общее расширение* θ*,* которое определим, полагая $\theta(x) = \theta_\alpha(x)$, если $\theta_\alpha(x)$ определено, для одного или большего числа индексова α. Из леммы 2 вытекает, что существует максимальное расширение изоморфизма между идеалами вполне упорядоченных множеств V и W. В силу леммы 1 оно включает в себя или все V, или все W, или же имеет место между множеством элементов $x < a$ и множеством элементов $y < b$ для некоторых $a \in V$ и $b \in W$. Но в последнем случае его можно было бы расширить дальше, полагая $\theta(a) = b$. Таким образом, имеет место

Лемма 3. *Из двух любых вполне упорядоченных множеств V, W одно изоморфно идеалу другого.*

Следовательно, если мы положим $V \leq W$ в том смысле, что V изоморфно идеалу W, то будет иметь место Р4. Но Р1, Р3 очевидны, а Р2 (с равенством, означающим изоморфизм) является следствием леммы 2. Наконец, если Ω — какое-нибудь непустое множество вполне упорядоченных множеств, то выберем некоторое $W \in \Omega$. Система множеств $V \not> W$ состоит в силу леммы 3 из вполне упорядоченных множеств, изоморфных с идеалами множества W, а последние в силу леммы 1 образуют вполне упорядоченное множество. Мы заключаем:

Теорема 4. *Любая система вполне упорядоченных множеств является вполне упорядоченным множеством, если мы определим $V \leq W$ в том смысле, что V изоморфно идеалу множества W; здесь $V = W$ означает, что V и W изоморфны.*

В силу леммы 1 каждое вполне упорядоченное множество W изоморфно вполне упорядоченному множеству вполне упорядоченных множеств $V < W$.

Лемма 4. *Пусть W—вполне упорядоченное множество и пусть ψ— некоторое правило, которое для каждого $a \in W$ и при выбранных $\psi(b)$ для всех $b < a$ однозначно определяет $\psi(a)$. Тогда $\psi(a)$ однозначно и последовательно определено для всех $a \in W$.*

Доказательство. Пусть для каждого начального интервала A множества W $P(A)$ обозначает утверждение, что $\psi(a)$ однозначно и последовательно определено для всех $a \in A$. Так как A вполне упорядочены, достаточно показать, что нет первого A, для которого $P(A)$- неверно. Но очевидно, что A не может быть таковым, если A обладает последним элементом a. Если же A есть объединение начальных отрезков $B<A$ и ψ однозначно и последовательно определено для всех $B < A$, то мы можем также составить (однозначно и последовательно определенное) общее расширение ψ на все A.

6.4.4. Условия обрыва цепей

Гильберт показал, что классическая теория инвариантов существенным образом опирается на тот факт, что идеалы во многих полиномиальных кольцах образуют структуру, удовлетворяющую следующему условию для возрастающих цепей.

Определение. *Частично упорядоченное множество P удовлетворяет условию обрыва возрастающих цепей (соответственно условию обрыва убывающих цепей) тогда и только тогда, если каждое непустое подмножество множества P имеет максимальный (минимальный) элемент.*

Если *P* удовлетворяет условию обрыва возрастающих цепей, то это же, очевидно, справедливо для каждого подмножества множества *P;* если *P* и *Q* удовлетворяют оба этому условию, то ему же удовлетворяют $P+Q$, $P \oplus Q$, PQ и $P \circ Q$. Далее, так как условия быть минимальным и быть наименьшим эквивалентны для цепей, то мы видим, что цепь удовлетворяет условию обрыва *убывающих* цепей тогда и только тогда, если она вполне упорядочена. Вообще, справедлива

Теорема 5. *Следующие утверждения относительно частично упорядоченного множества P эквивалентны:* а) *P удовлетворяет условию обрыва возрастающих цепей, б) во множестве* $\breve{P}$, *двойственном множеству P, каждая цепь вполне упорядочена,* в) *все возрастающие цепи в P конечны.*

Доказательство. Если *P* не удовлетворяет условию обрыва возрастающих цепей, то оно содержит непустое подмножество *X,* не имеющее максимального элемента. Выберем $a_1 \in X$; так как a_1 не максимально, мы можем найти в *X* $a_2>a_1$; затем, по тем же соображениям, мы найдем в *X* $a_3>a_2$ и т. д. Таким образом, мы получаем бесконечную возрастающую цепь $a_1 < a_2 < a_3 < ...$, изоморфную вполне упорядоченному множеству ω, и «в» не имеет места. Если не имеет места «в», то не имеет места и «б», ибо множество, двойственное ω, не является вполне упорядоченным. Наконец, если не имеет места «б», то, так как в цепях минимальные элементы являются наименьшими, «а» также не имеет места, чем и завершается доказательство.

Если *P* удовлетворяет условию обрыва возрастающих цепей, то мы можем доказывать по индукции следующим образом. Если свойство Ф верно не для всех элементов из *P,* то должен быть некоторый максимальный элемент *m,* для которого это свойство неверно. Все элементы $x > m$ будут тогда обладать этим свойством. Мы заключаем:

Обобщенный принцип индукции. *Пусть частично упорядоченное множество P удовлетворяет условию обрыва возрастающих цепей и пусть* Ф *— некоторое свойство. При доказательстве того, что каждое* $a \in P$ *обладает свойством* Ф, *мы можем предполагать для каждого a, что все* $x > a$ *обладают свойством* Ф.

Может случиться, что не все цепи частично упорядоченного множества P являются вполне упорядоченными, но, тем не менее, любые $a < b$ можно связать по крайней мере одной вполне упорядоченной максимальной цепью. Это имеет место, например, для системы всех подмножеств некоторого класса; используя аксиому выбора (см. ниже, 6.4.6), мы можем добавлять к заданному множеству отдельные точки до тех пор, пока мы не исчерпаем некоторое предписанное нам множество, содержащее в себе первоначальное множество. Это имеет место также для системы всех замкнутых подмножеств некоторого топологического пространства и для алгебраически замкнутых подполей некоторого поля (ибо мы можем присоединять последовательно точки, соответственно трансцендентные элементы).
Во всех случаях, когда имеет место указанное свойство, множество P будем называть ↑-атомным, если же ↑-атомным является множество, двойственное P, мы будем называть множество P ↓-атомным. Система всех подмножеств континуума (будучи двойственной себе) является как ↑-атомной, так и ↓-атомной. Из систем, не являющихся ни ↑-атомными, ни ↓-атомными, мы отметим вещественные числа; измеримые множества, определяемые с точностью до множеств меры нуль (все максимальные цепи изоморфны отрезкам прямой), и систему всех множеств целых положительных чисел, определяемых с точностью до конечных подмножеств.

6.4.5. Топология цепей

Известно, что топология вещественного континуума может быть определена в терминах упорядоченности; это может быть обобщено на произвольные частично упорядоченные множества. Исторически первым было сделано обобщение на цепи; его мы и рассмотрим.
Определение. *Пусть C — какая-нибудь цепь. Открытыми интервалами в C являются* а) *само C, обозначаемое через* $(-\infty, +\infty)$, б) *для каждого $a \in C$ множество $(a, +\infty)$ всех $x > a$,* в) *для каждого $a \in C$ множество $(-\infty, a)$ всех $x < a$ и г) для любых $a < b$ в C множество (a, b) всех x, удовлетворяющих соотношению $a < x < b$. Замкнутые интервалы $[-\infty, +\infty]$, $[a, +\infty]$ и т. д. в цепи C получаются из открытых интервалов путем замены значка $<$ значком $\leq$. Под окрестностью элемента $p \in C$ понимают произвольный открытый интервал, содержащий p.*
Теорема 6. *Всякая цепь является нормальным хаусдорфовым пространством относительно присущей ей топологии, и эта топология инвариантна относительно автоморфизмов и дуальных автоморфизмов.*

Доказательство. Определение означает в первую очередь, что любое $p \in C$ имеет окрестность (например, *C);* что пересечение двух любых окрестностей *p* есть снова окрестность *p;* что любая окрестность *p* является также окрестностью всех своих точек, т. е. что «открытые интервалы» могут быть взяты в качестве базиса окрестностей. Далее, если $p \neq q$ в *C,* то или *p* покрывает *q*, или $p > a > q$ для некоторого *a,* или же имеет место двойственное соотношение. В первом случае $(-\infty, q)$ и $(p, +\infty)$ образуют пару непересекающихся окрестностей; во втором случае такую пару образуют $(-\infty, a)$ и $(a, +\infty)$; двойственные выражения имеют место в двух последних случаях. Следовательно, *C* является хаусдорфовым пространством и, в частности, T_1 — пространством. Доказательство нормальности будет приведено после теоремы 7.

В любом хаусдорфовом пространстве непустое *открытое* множество определяется как объединение окрестностей (следовательно, в нашем случае, как объединение открытых интервалов), а *замкнутое* множество — как множество, дополнение которого открыто. Легко показать, что любое объединение открытых множеств открыто и что любое пересечение замкнутых множеств замкнуто. *Замыкание* $\overline{S}$ множества *S* определяется как пересечение всех замкнутых множеств, содержащих *S; открытое ядро* множества *S* определяется как объединение всех открытых множеств, содержащихся в *S*.

Наиболее общее замкутое множество в цепи является, таким образом, пересечением дополнений открытых интервалов. Но дополнение любого открытого интервала есть сумма одного или двух замкнутых интервалов. Следовательно, наиболее общее замкнутое множество является пересечением конечных сумм замкнутых интервалов. Этим доказана

Теорема 7. *В любой цепи С замкнутые интервалы образуют псевдобазис семейства замкнутых множеств.*

Пусть теперь *S* и *T*— произвольные, непересекающиеся, замкнутые множества из С; тогда множество $S \cup T$ будет замкнутым, а потому его дополнение будет объединением открытых интервалов I_α. В каждом I_α выберем элемент x_α (это требует аксиомы выбора). Каждая точка *p* из *S,* не принадлежащая открытому ядру *S,* должна с каждой стороны отделяться от *T* некоторым интервалом I_α, ибо *T* замкнуто и $S \cap T = 0$. В каждом таком случае мы присоединяем к *p* открытый интервал (p, x_α) или (x_α, p). После аналогичных рассуждений в отношении множества *T* мы получаем вложение *S* и *T* в непересекающиеся открытые множества; но это есть определение нормальности. Тем самым мы завершили доказательство теоремы 7.

Предположим теперь, что C есть *полная цепь,* т. е., что каждое непустое подмножество в C имеет наим. в. г. и наиб. н. г. Каждая конечная цепь полна, полной является также система вещественных чисел ***R#,*** если к ней присоединены — ∞ и +∞ (т. е. 0 и I). Вообще, если X и Y— две полные цепи, то полными являются также $X \oplus Y$ и $X \circ Y$. По такому способу можно построить большое многообразие неизоморфных полных цепей.

Определение. *Пусть* $\{x_\alpha\}$ *— некоторое направленное множество элементов полной цепа. Мы определяем*

$$\operatorname{Lim\,inf}\{x_\alpha\} = \sup_\beta \{\inf_{\alpha \geqslant \beta} x_\alpha\}, \quad \operatorname{Lim\,sup}\{x_\alpha\} = \inf_\beta \{\sup_{\alpha \geqslant \beta} x_\alpha\}. \quad (2)$$

Из неравенства минимакса (5) главы II следует, что

$$\operatorname{Lim\,inf}\{x_\alpha\} \leqslant \operatorname{Lim\,sup}\{x_\alpha\}. \quad (3)$$

Теорема 8. *В полной цепи* $x_\alpha \to a$ *тогда и только тогда, если*

$$\operatorname{Lim\,inf}\{x_\alpha\} = \operatorname{Lim\,sup}\{x_\alpha\} = a. \quad (4)$$

Доказательство. Выражение $x_\alpha \to a$, в соответствии с топологическим предисловием, означает, что каждый открытый интервал (b, c), содержащий a, содержит для некоторого β все x_α с $\alpha \geq \beta$. Но если это имеет место, то $\operatorname{Inf}_{\alpha \geqslant \beta} x_\alpha > b$, и, так как это справедливо для всех $b < a$, $\operatorname{Lim\,inf}\{x_\alpha\} \geq a$. Двойственным образом из $x_\alpha \to a$ следует, что $\operatorname{Lim\,sup}\{x_\alpha\} \leq a$, откуда в силу (3) $a = \operatorname{Lim\,inf}\{x_\alpha\} = \operatorname{Lim\,sup}\{x_\alpha\}$. Обратно, если $\operatorname{Lim\,inf}\{x_\alpha\} = a = \operatorname{Lim\,sup}\{x_\alpha\}$ и $a \in (b, c)$, то $\operatorname{Inf}_{\alpha \geqslant \beta} x_\alpha > b$ и $\operatorname{Sup}_{X \geq \gamma} x_\alpha < c$ для некоторых β, γ; следовательно, если $\delta \geq \beta, \gamma$, то $x_\alpha \in (b, c)$ для всех $\alpha \geq \delta$, а потому $x_\alpha \to a$.

Следствие. *Если в полной цепи* $x_\alpha \to a$, *то существуют направленные множества* $t_\alpha \uparrow a$ *и* $u_\alpha \downarrow a$ *с* $t_\alpha \leq x_\alpha \leq u_\alpha$.

(Обозначение $t_\alpha \uparrow a$ означает, что из $\alpha \leq \alpha'$ следует $t_\alpha \leq t_{\alpha'}$ и что $Vt_\alpha = a$.) Этот результат сохраняет силу и для произвольных цепей. Однако он не будет уже иметь места для последовательностей в произвольной структуре.

Теорема 9. *Цепь С полна тогда и только тогда, если она топологически компактна.*

Доказательство. Пусть $\{W_\alpha\}$—произвольное семейство открытых множеств, покрывающее полную цепь C. Так как каждое W_α является объединением открытых интервалов, то C покрывается также семейством открытых интервалов $V_{\alpha,\beta} \leq W_\alpha$. Для доказательства компактности С достаточно показать, что C можно покрыть конечным подсемейством интервалов $V_{\alpha,\beta}$. Пусть A—множество всех a таких, что $[0, a]$ может быть покрыто конечной подсистемой интервалов $V_{\alpha,\beta}$. Заметим сперва, что $\sup A = b \in A$, ибо некоторое $V_{\alpha,\beta}$ содержит b и, следовательно, некоторое $a \in A$, а потому существует конечное покрытие интервала $[0, b] \leq [0, a] \cup V_{\alpha,\beta}$. Покажем теперь, что $b = I$, в силу чего $C = [0, I]$ может быть покрыто требуемым образом. В самом деле, пусть c — наиб. н. г. элементов $x > b$; если множество элементов $x > b$ не пусто, т. е. если b отлично от I, то, добавляя какую-нибудь окрестность элемента c к конечному покрытию интервала $[0, b]$, мы получили бы конечное покрытие большего интервала. Но это противоречило бы определению $b = \sup A$.

Обратно, пусть X—произвольное подмножество компактной цепи C. Для каждого конечного подмножества F множества X определим Y_F как наименьший элемент среди $x \in F$; очевидно, что ни одно из конечных пересечений $[-\infty, Y_F]$ замкнутых интервалов $[-\infty, x]$ не будет пустым; следовательно, $\wedge_x[-\infty, x]$ непусто и X имеет нижнюю грань a. Рассмотрим теперь множество всех замкнутых интервалов $[a, x]$, где $x \in X$, а a — переменная нижняя грань для X. Никакое конечное пересечение таких интервалов непусто; следовательно, $\wedge[a, x]$ содержит элемент b. Но любое такое b является нижней гранью для X, содержащей каждую нижнюю грань для X; следовательно, $b = \inf X$, которое, таким образом, существует. Двойственным образом существует $\sup X$, чем и завершается доказательство.

6.4.6. Аксиома выбора

Рассмотрим теперь предположение, которое кажется не зависимым от обычных логических предположений относительно классов и соответствий, но, тем не менее, согласовано с ними. Это —так называемая аксиома выбора.

Ее можно формулировать различными способами, и она означает, что каждое множество может быть вполне упорядочено. Особенно простой формулировкой является следующая.

(АС1) Каждая пепь C в частично упорядоченном множестве P содержится в *максимальной* цепи M из P.

Таким образом, существует множество $M \leq P$, которое является цепью и которое не содержится ни в какой большей цепи N, $N \leq P$. Легко показать, что из (*АС*1) вытекает следующее условие.

(АС2) Если каждая цепь C частично упорядоченного множества P имеет в P верхнюю грань $U(C)$, то P содержит максимальный элемент.

В самом деле, любая верхняя грань U в любой максимальной цепи M из P будет максимальным элементом в P, и такой элемент существует, если (АС1) имеет место.

Тэкей обобщил (АС2). Он называет свойство Ф множеств S «свойством конечного характера», если $S \in Ф$ эквивалентно тому, что $F \in Ф$ для всех конечных подмножеств множества S. В этом случае из $S \in Ф$ следует, что $T \in Ф$ для каждого $T \leq S$; следовательно, свойства конечного характера суть то, что называют «свойствами независимости». Условие Тэкея следующее:

(АС3) Пусть Ф—свойство конечного характера, определенное для подмножеств множества S. Тогда существует максимальное подмножество M множества S, обладающее свойством Ф (т. е. максимальное $M \in Ф$).

Покажем теперь, что из (АС2) следует (АС3). Пусть C— максимальная цепь множеств $X_c \in Ф[X_c \leq S]$. Если тогда F— любое конечное подмножество из объединения U множеств $X_c \in C$, то, так как каждое $a_i \in F$ содержится в некотором $X_i \in C$, наибольшее из этих X_i будет содержать F; следовательно, $F \in Ф$. Отсюда следует, что $U \in Ф$, в силу чего Ф удовлетворяет предположению в (АС2), а потому [в предположении, что справедливо (АС2)] имеет максимальный член.

Условие, что множество T удовлетворяет всем свойствам $Ф_\alpha$ из совокупности $\mathfrak{F}$ свойств конечного характера, является само свойством конечного характера. Следовательно, мы имеем также

(АС3'). Пусть $\mathfrak{F}$ —произвольная совокупность свойств $Ф_\alpha$ конечного характера на множестве S. Тогда существует максимальное подмножество M множества S, удовлетворяющее каждому $Ф_\alpha \in \mathfrak{F}$.

Заметим, что свойство быть цепью есть свойство конечного характера. (Оно является свойством характера два: оно требует, чтобы для заданных x, $y \in T$ было бы или $x \geq y$ или $y \geq x$). Следовательно, из (АС3') вытекает (АС1). Тем самым завершается цикл включений: (АС1)→(АС2)→(АС3)→(АС3')→(АС1). *Каждое из этих условий эквивалентно, таким образом, всем остальным.*

Мы покажем теперь, что все эти условия эквивалентны также первоначальной аксиоме выбора Цермело:

(АС) Пусть I—произвольное множество. Существует однозначная функция, которая относит каждому непустому подмножеству T множества I вполне определенный элемент $g_M(T)$ из T.

Рассмотрим класс Г всех однозначных функций g: $g(T) \in T$, определенных на некоторых (но не обязательно на всех) подмножествах множества I. Мы будем писать $g \leq h$, если $h(S)$ существует и равно $g(S)$ всякий раз, когда $g(S)$ определено; тем самым Г становится частично упорядоченным множеством. Предполагая справедливым (АС1), мы выберем теперь в Г *максимальную цепь М* и образуем *общее* *расширение* g_M функций $g \in M$. Функция g_M будет тогда определена для всех непустых *T;* в противном случае существовало бы непустое T_0 с некоторым элементом t_0, для которого g_M (T_0) не было бы определено. Мы могли бы определить

$$g^* \begin{cases} g^*(T) = g_M(T) \in T, & \text{если } g_M(T) \text{ определено,} \\ g^*(T_0) = t_0 \in T_0. \end{cases}$$

Добавляя g^* к *M'*, мы получили бы цепь большую, чем *M,* придя тем самым к противоречию. Следовательно, из (АС1) вытекает (АС). Теорема Цермело утверждает, что (АС) эквивалентно (АС4). Каждое множество *I* может быть вполне упорядочено.

Доказательство. Рассмотрим отношения ρ, которые вполне упорядочивают *подмножества* *S*(ρ) множества *I*. Мы пишем $\rho \leq \rho^*$, если *S*(ρ) является начальным интервалом множества *S*(ρ*). Рассмотрим лишь отношения полного упорядочения, «совместимые» с g_M из (АС) в том смысле, что g_M относит дополнению *I* — *A* каждого начального интервала множества *S(ρ),* определенного отношением ρ, первый элемент, следующий за *A*. Можно показать, так же как и в лемме 3, 6.4.3, что эти ρ образуют *цепь.* Следовательно, их общее расширение $\bar{\rho}$ является однозначным, вполне упорядочивающим (любое подмножество объединения *V* множеств *S* (ρ) содержит первый элемент для некоторого ρ, а потому и для всех ρ) и совместимым с *gM* отношением. Очевидно теперь, что $\bar{\rho}$ максимально; мы утверждаем, что оно вполне упорядочивает множество *I*. В противном случае мы могли бы определять $\rho_1 \geq \bar{\rho}$, полагая, что ρ_1 имеет *V* в качестве начального интервала, и определяя $g_M(I—V) = \upsilon$ как последний элемент множества *V,* дополненного элементом υ.

Мы доказали уже, что (АС1)→(АС2)→(АС3)→(АС3')→(АС1) и (АС1)→(АС)→(АС4); если мы сможем доказать, что (АС4) →(АС1), то мы докажем тем самым теорему:

Теорема 10. *Утверждения* (АС), (АС1), (АС2), (АС3), (АС3') *и* (АС4) *являются эквивалентными формулировками одной и той же гипотезы.*

Докажем теперь, что (АС4)→(АС1). Если заданы *Р* и *С,* фигурирующие в формулировке (АС1), то мы сперва представим в виде вполне упорядоченной трансфинитной последовательности *W* элементы из *Р,* не принадлежащие *С.* Определим теперь функцию, относящую каждому $a \in W$ значение $\psi(a) = M$, если для всех $b<a$ таких, что $\psi(b) = M$, и для всех $b \in C$ или $a \geq b$, или $a \geq b$, и значение $\psi(a) = M'$ в противном случае. В силу леммы 4, 6.4.3, эта функция однозначно и последовательно определена; множество всех *a* с $\psi(a) = M$ является

цепью и притом максимальной цепью, чем и завершается доказательство.

6.4.7. Проблема континуума

Используя аксиому выбора, мы можем доказать многие результаты, имеющие высокую степень общности. Эти результаты не могут быть, по всей вероятности, получены иным способом.

Теорема 11. *Пусть, по определению, $S\leq T$ означает, что между S и подмножеством множества T имеет место взаимно однозначное соответствие. Тогда любое множество кардинальных чисел является вполне упорядоченным.*

Доказательство. Каждое порядковое число *W* является вполне упорядоченным множеством, которому отвечает кардинальное число ***n***(W}; при этом соответствие *W→**n**(W)* изотонно. Кроме того, в силу (АС4) каждое кардинальное число есть ***n**(W)* для некоторого *W*. Но любое множество порядковых чисел вполне упорядочено; следовательно, то же самое справедливо для кардинальных чисел.

Используя аксиому выбора, мы можем также доказать для кардинальных чисел, что если α или β бесконечно, то

$$\alpha+\beta = \alpha\beta = \max(\alpha, \beta). \tag{2}$$

Как следствие, мы получаем для случая, когда α бесконечно,

$$\alpha+\alpha = \alpha^2 = \alpha. \tag{3}$$

Теорема 12. *T_1-пространство С компактно, если оно содержит псевдобазис В замкнутых множеств* С7р, *удовлетворяющих условию:* (НВ) *если любая конечная система множеств Uf имеет непустое пересечение, то Uf имеют общую точку[1]).*

Доказательство. Пусть *К*—произвольная совокупность базисных замкнутых подмножеств T_α, обладающая «свойством конечных пересечений», т. е. любая конечная подсистема *F* множеств $T_\alpha\in K$ имеет непустое пересечение. Пусть *М* есть *максимальное* расширение *К;* так как свойство конечных пересечений является свойством конечного характера, то в силу (АСЗ) *М* существует. Поскольку *В* — псевдобазис, то любое $S_\mu\in M$ является объединением конечной совокупности F_μ множеств U_μ, из *В*. Справедливость для всех $U_{\mu,i}\in F_\mu$ и для подходяще выбранных $S_{i,j}$ соотношений $U_{\mu,i}\cap S_{i,1}\cap \ldots \cap S_{i,r(i)} = 0$ означала бы, что

$$S_\mu \cap (\wedge S_{i,j}) = \vee (U_{\mu,i}) \cap (\wedge S_{i,j}) = \vee [U_{\mu,i} \cap \wedge S_{i,j}] \leqslant \vee 0 = 0.$$

Поэтому для каждого S_μ, мы можем добавить к *М* некоторое $U_{\mu,i}$, не нарушая свойства конечных пересечений. Но *М* максимально; следовательно, для каждого $S_\mu\in M$ некоторое $U_\mu\leq S_\mu$. из *В* также лежит в M; в частности, это справедливо для T_α. Но так как множества U_μ удовлетворяют (НВ) и обладают свойством конечных пересечений, то $\wedge\ U_\mu > 0$. Тогда

$$\wedge T_\alpha \geqslant \wedge S_\mu \geqslant \wedge U_\mu > 0.$$

Следствие. *Топологическое произведение компактных T_1-пространств компактно.*

Мы видели уже в 6.4.2, что наименьшим бесконечным кардинальным числом является счетная бесконечность, κ_0. Но известно, что вещественные числа имеют большее кардинальное число c. Поскольку не существует конструктивного полного упорядочения несчетных множеств, нам неизвестно место каждого несчетного множества во вполне упорядоченной последовательности κ_0, κ_1, κ_2,…бесконечных кардинальных чисел. Важным предположением является так называемая

Гипотеза континуума. *Мощность континуума с является вторым бесконечным кардинальным числом;* Серпинский показал, что это предположение эквивалентно большому числу других не доказанных (и не опровергнутых) утверждений теории множеств.

6.4.8. Однородный континуум

Система вещественных чисел ***R#*** является фундаментальной в математике, и много усилий было затрачено на то, чтобы дать абстрактную характеристику отношения упорядоченности в этой системе.

Теорема 13. *Система вещественных чисел **R#** и множество J всех целых чисел являются двумя единственными цепями, которые* а) *для каждого ограниченного подмножества имеют наим. в. г. и наиб. н. г.,* б) *имеют счетное плотное подмножество,* в) *являются «однородными» в том смысле, что они обладают транзитивной группой автоморфизмов.*

Доказательство. В силу теоремы 2 условие «б» означает, что цепь C изоморфна подмножеству множества ***R#***; так мы и будем ее представлять. Дополнение множества $\overline{C}$ открыто и состоит из счетного семейства неперекрывающихся открытых интервалов. Из условия «а» следует, что по крайней мере один из концов каждого такого конечного интервала должен принадлежать C. Если в каком-либо случае оба конца a и b принадлежат C, то тем самым некоторый элемент $a \in C$ покрывается другим элементом $b \in C$; отсюда следует в силу «в», что *каждое* $x \in C$ должно покрываться некоторым $y \in C$ и, аналогично, что каждое x *покрывает* некоторое $z \in C$. Мы можем, следовательно, найти частичный интервал в *C,* изоморфный с *J*. Наконец, если бы некоторое $a \in C$ не принадлежало *J*, то мы могли бы в силу «а» образовать sup*J* или inf*J;* но такие элементы не могли бы покрывать другие элементы и покрываться другими элементами; следовательно, *J = C.*

В другом возможном случае множеству *C* принадлежит в точности один из концов каждого открытого интервала. Если мы в этом случае сдвинем каждую точку *p* из *C* по направлению к началу отсчета на расстояние, равное в точности сумме длин интервалов между *p* и началом отсчета, то мы получим изоморфное отображение множества С на ***R#*** или на интервал множества ***R#***. Если бы некоторое $x \in C$ было наибольшим или наименьшим, то таковым был бы (в силу «в») и каждый элемент из *C.* Поэтому или *C* состоит из одного элемента 0 (исключительный случай), или *C* изоморфно открытому интервалу множества ***R#*** и, следовательно, самому ***R#.***

Возможны различные видоизменения предшествующих условий. Так, например, мы можем заменить «в» гипотезой «в'», что *не* имеет места отношение покрытия (т. е. что цепь плотна в себе), и получить ***R#***; или мы можем заменить «в» гипотезой «в"», что каждое $a \in C$ покрывает другой элемент и покрывается другим элементом, и получить *J*.

Весьма важным видоизменением «б» является условие Суслина:

«б'» *всякая система неперекрывающихся открытых интервалов множества C является счетной*

Легко показать, используя теорему 2, что из «б» следует «б'». Для каждого целого положительного числа *n* имеется самое болшее *n* + 1 неперекрывающихся интервалов длины большей или равной 1/*n,* середины которых удовлетворяют соотношению $m \le x \le m+1$; следовательно, в ***R#*** возможно только счетное множество неперекрывающихся интервалов положительной длины.

Проблема Суслина состоит в том, чтобы определить, следует ли, обратно, или нет в произвольной цепи из условия «б'» условие «б». Это представляется крайне трудным. Наиболее естественный подход включает в себя построение максимальной трансфинитной последовательности замкнутых множеств $S_1, S_2, S_3, \dots S_\omega, S_{\omega+1}, S_{\omega+2}, \dots$ точек из подразбиений множества *C* по следующему способу. Не нарушая общности, мы можем предполагать, что *C* полно (компактно). Пусть S_1—пустое множество. Множество S_{n+1} образуем из множества S_n путем добавления одной точки Ф*(J)* из каждого открытого интервала *J* между последующими точками множества S_n; это возможно, так как дополнение множества S_n состоит из счетного семейства неперекрывающихся открытых интервалов. Если α—предельное число, то S_α есть замыкание теоретико-множественного объединения всех S_β с

$\beta < \alpha$. Получаемые при этом процессе открытые интервалы образуют частично упорядоченное множество *P,* которое определяет *C* с точностью до изоморфизма (гомеоморфизма), и проблема Суслина может быть сведена к изучению *P*.

Из условия «б'» вытекает, что определенное так частично упорядоченное множество *P* интервалов удовлетворяет условию обрыва возрастающих цепей, (S1) для любого $a \in P$ множество всех $x \geq a$ в *P* образует цепь, (S2) каждая цепь счетна (так как несчетная убывающая последовательность интервалов $I_1 > I_2 > I_3 > \ldots$ давала бы несчетное множество неперекрывающихся интервалов $I_i — I_{i+1}$.) и (S3) каждое множество несравнимых между собой элементов счетно. Утверждение, что подмножество из *C* плотно, равносильно утверждению, что оно имеет представителя в каждом интервале. Это приводит к следующей форме проблемы Суслина.

Пусть *P*—произвольная *ветвистость,* т. е. частично упорядоченное множество, удовлетворяющее условию (S1). Если каждая цепь и каждое множество несравнимых между собой элементов из *P* имеют не более, чем счетную мощность, то само *P* счетно.

Другая система условий для цепей была предложена Дж. Биркгофом. Они эквивалентны условиям:

«а**» объединение $\bigvee_{n=1}^{\infty} I$ любой возрастающей последовательности $I_1 < I_2 < I_3 < \ldots$ открытых интервалов есть открытый интервал, условию Суслина «б'» и «в**» все интервалы изоморфны между собой.

Цепи, удовлетворяющие условиям «а**» и «в**», называются *однородными линейными континуумами.* Они представляют собой естественные обобщения системы ***R#*** вещественных чисел. Примеры однородных линейных континуумов, не изоморфных ***R#,*** были найдены Васкесом и Субиета, а также Аренсом.

6.5. Полные структуры

6.5.1. Определение; операции замыкания

Ранее мы определили *полную структуру* как частично упорядоченное множество, в котором каждое подмножество имеет наименьшую верхнюю грань и наибольшую нижнюю грань. Очевидно, что всякая конечная структура, а также всякая структура, цепи которой конечны, являются полными. Любая полная цепь, как, например, система вещественных чисел с $-\infty$ и $+\infty$, есть полная структура;

кардинальные произведения и степени полных структур суть полные структуры; цолными структурами являются также ординальные сумма и произведение двух полных структур.

Однако наиболее типичные полные структуры получаются не по этому способу, а с помощью *операций замыкания,* имеющих следующий смысл.

Определение. *Под «операцией замыкания» на подмножествах X совокупности I понимаем операцию* $X \to \overline{X}$ *такую, что*

С1. $\overline{X} \geq X$ *(экстенсивность),*

С2. $\overline{\overline{X}} = \overline{X}$ *(идемпотентность),*

С3. *из* $X \geq Y$ *следует* $\overline{X} \geq \overline{Y}$ *(изотонностъ).*

Под «замкнутым» множеством мы понимаем множество X, равное своему «замыканию» $\overline{X}$.

Теорема 1. *Подмножества, «замкнутые» относительно какой-нибудь операции замыкания, образуют полную структуру, в которой наиб. н. г. означает пересечение* (Мур).

Доказательство. Теоретико-множественное произведение *Р* произвольного семейства Ф замкнутых подмножеств $X_Ф$ замкнуто, ибо в силу С3 $\overline{P} \leq \overline{X}_Ф = X_Ф$, для всех $X_Ф \in Ф$, откуда $\overline{P} \leq P$. Следовательно, *Р* есть наиб. н. г. для $X_Ф$. Далее, замыкание $\overline{S}$ теоретико-множественного объединения *S* множеств $X_Ф$ содержит в силу С1 каждое $X_Ф$ и содержится в каждом замкнутом множестве *Т,* содержащем каждое $X_Ф$, ибо $T \geq S$, а потому в силу С3 $T = \overline{T} \geq \overline{S}$ для каждого такого *Т.* Следовательно, $\overline{S}$ есть наим. в. г. для $X_Ф$.

Существует обобщение утверждения Мура, получаемое из того факта, что если всегда существует наиб. н. г. или наим. в. г., то всегда существует и то и другое. Справедлива

Теорема 2. *Если Р частично упорядоченное множество с I и каждое непустое подмножество множества Р имеет наиб. н. г. (или двойственно этому), то Р есть полная структура.*

Доказательство. Пусть *X*—произвольное подмножество элементов $x_Ф$ из *Р* и пусть *U*—множество верхних граней элементов $x_Ф$. Множество *U* непусто, ибо оно содержит *I*; пусть $a = \inf U$. Так как каждое $x_Ф$ является нижней гранью для *U, а* содержит каждое $x_Ф$; так как *а* есть нижняя грань для *U, а* содержится в каждом элементе, содержащем каждое $x_Ф$; следовательно, $a = \sup X$. Таким образом, $\sup X$ существует.

Мур рассмотрел *абстрактно* свойства множеств, которые, подобно свойству быть «замкнутым» по отношению к оператору замыкания, а) справедливы для *I*, б) справедливы для любого пересечения $\wedge X_Ф$

множеств, удовлетворяющих этим свойствам. Он назвал такие свойства множеств «экстенсивно достижимыми» и показал, что каждое экстенсивно достижимое свойство связано со свойством замыкания.
Если S произвольное подмножество в полной структуре L, содержащее I и содержащее вместе с любым подмножеством X из L также inf X, то S также является полной структурой. Если S содержит всегда inf$X$ и sup$X$, то S называется *замкнутой подструктурой.*

6.5.2. Примеры

Топологическое замыкание в произвольном T_0-пространстве есть «операция замыкания»; условия С1 — С2 постулируются, а С3 следует из того, что $X \geq Y$ означает $X \cup Y = X$,откуда $\overline{X} = \overline{X \cup Y} = \overline{X} \cup \overline{Y} \geq \overline{Y}$. Следовательно.
Теорема 3. *Замкнутые (и, двойственным образом, открытые) подмножества топологического T_0-пространства образуют полную структуру.*
Отношения конгруентности в произвольной алгебре с финитарными операциями образуют полную структуру. Поскольку в той части доказательства, которая относится к ξ, не требуется финитарности операций, мы получаем из приведенной выше теоремы 2 результат: *отношения конгруентности на произвольной абстрактной алгебре образуют полную структуру.* В частности, множество всех разбиений некоторой совокупности образует структуру; эта структура рассматривается дальше.
Теорема 4. *Подалгебры любой абстрактной алгебры A образуют полную структуру.*
Доказательство. Естественно, что A является своей собственной подалгеброй; очевидно также (определения см. в предисловии), что пересечение любого множества подалгебр есть подалгебра. Отсюда следует, что подгруппы произвольной группы, подкольца произвольного кольца, подпространства произвольного векторного пространства и т. д. образуют полные структуры.
Выпуклые тела в пространстве также образуют полную структуру; «выпуклая оболочка» произвольного множества X может быть определена как множество всех $\lambda_1 x_1 + ... + \lambda_n x_n$ с $x_i \in X$ и $\lambda_i \geq 0$ для $i = 1,...,n$ и $\sum \lambda_i = 1$.

6.5.3. Условная полнота; σ-структуры

Многие важные структуры, хотя и не являющиеся полными, обладают тем свойством, что *каждое непустое ограниченное подмножество имеет* наиб. н. г. и наим. в. г. Такие структуры называются *условно полными.* Так, например, система ***R#*** вещественных чисел условно полна; условно полным является также множество всех функций, субгармонических в заданной области и принимающих заданные граничные значения.

Докажем следующий частичный аналог теоремы 2.

Теорема 5. *Для того, чтобы структура L была условно полной, достаточно, чтобы каждое ограниченное непустое множество имело наиб. н. г. В структуре L каждое непустое подмножество, имеющее нижнюю грань, имеет наиб. н. г.*

Доказательство. Предположим, что каждое ограниченное подмножество X в L имеет наиб. н. г.; рассмотрим множество U верхних граней множества X. Так как X ограничено, U непусто; выберем $a \in U$; пусть V обозначает множество элементов $a \cap u[u \in U]$. Тогда V ограничено сверху элементом a, а снизу элементами множества X; следовательно, оно имеет наиб. н. г. b. Очевидно, что $b \geq x$ для всех $x \in X$, так как все $x \in X$ являются нижними гранями для V; очевидно также, что $b \leq a \cap u \leq u$ для всех $u \in U$; следовательно, $b = \sup X$.

Далее, если Y непустое подмножество в L, имеющее нижнюю грань, то образуем аналогичным образом для некоторого фиксированного $c \in Y$ множество элементов $c \cap y[y \in Y]$. Оно будет ограниченным множеством, и его наиб. н. г. будет, аналогичным образом, $\inf Y$.

Фактически, как это было отмечено Реньи, достаточно, чтобы каждое вполне упорядоченное подмножество имело наим. в. г. и двойственно этому; но доказательство этого требует аксиомы выбора.

Утверждения, двойственные двум утверждениям теоремы 5, доказываются аналогичным образом.

Покажем теперь, что единственным отличием условно полных структур от полных структур является отсутствие 0 и *I*. Это обобщает тот факт, что мы можем превратить ***R#*** в полную структуру путем добавления элементов $-\infty$ и $+\infty$.

Теорема 6. *Пусть P — произвольное частично упорядоченное множество, в котором каждое непустое подмножество, имеющее нижнюю грань, имеет наиб. н. г. Тогда, если к P добавить* 0, *I, то получим полную структуру* $\overline{P}$.

Доказательство. Пусть X—произвольное непустое подмножество в $\overline{P}$; сводим все к случаю подмножества множества P. Если X имеет нижнюю грань в P, то наиб. н. г. множества X в P будет также наиб. н. г. в $\overline{P}$; в противном случае 0=inf$X$, являясь единственной нижней гранью для X. Далее применяем теорему 2 п.6.5.1.

Пример, изображенный на рис. 1, *е*, показывает, что недостаточно просто предполагать, что каждое ограниченное множество имеет наиб. н. г. и наим. в. г.

В теории множеств очень полезно составление «σ-колец» множеств, т. е. семейств Ф, которые содержат объединение $\vee_{i=1}^{\infty} X_i$ и пересечение Л $\wedge_{i=1}^{\infty}$ любого счетного подсемейства множеств X_1, X_2, X_3, ... из Ф. Например, измеримые подмножества и борелевские подмножества на прямой, плоскости и в пространстве являются σ-кольцами множеств. Каждое σ-кольцо множеств является σ-структурой в следующем смысле.

Определение. Структура L, в которой каждое счетное подмножество имеет наиб. н. г. и наим. в. г., называется *σ-структурой.*

Другой совокупностью свойств частично упорядоченных множеств являются *(m, n)* интерполяционные свойства.

Определение. *Частично упорядоченное множество P называют имеющим (m, n) интерполяционное свойство тогда и только тогда, если всякий раз, когда* $a_1, \ldots, a_m \in P$ *и* $b_1, \ldots, b_n \in P$ *удовлетворяют соотношениям* $a_i \le b_j$ *для всех i, j, существует элемент* $c \in P$ *такой, что* $a_i \le c \le b_j$*- для всех i, j.*

6.5.4. Обобщенные законы; теорема о неподвижных точках

Идемпотентный, коммутативный и ассоциативный законы могут быть объединены в закон L*, характеризующий вместе с L4 полные структуры.

Теорема 7. *Во всякой полной структуре каждое подмножество S элементов* x_α *имеет* $\wedge_S x_\alpha$ *и* $\vee_S x_\alpha$; *кроме того,* L*. *Если* Ф *произвольное семейство множеств* S_φ *и S обозначает теоретико-множественное объединение множеств* S_φ, *то* $\wedge_\Phi \{\wedge_{S_\varphi} x_\alpha\} = \wedge_S x_\alpha$ *и двойственно,* $\vee_\Phi \{\vee_{S_\varphi} x_\alpha\} = \vee_S x_\alpha$ -

Обратно, любая система с операциями $\vee_S x_\alpha$ *и* $\wedge_S x_\alpha$*, определенными для любого подмножества S элементов, и такими, что справедливы* L* *и* L4, *является полной структурой.*

Доказательство. Нижняя грань для S_φ есть то же самое, что и нижняя грань для $\wedge_{S_\varphi}$; нижняя грань для S есть то же самое, что и нижняя грань для каждого S_φ, и, следовательно, для каждого $\wedge_{S_\varphi}$, откуда $\wedge_\Phi\{\wedge_{S_\varphi} x_\alpha\} = \wedge_S x_\alpha$. Остающаяся часть утверждения L* следует из двойственных соображений. Обратно, легко показать, что LI — L3 являются частными случаями утверждения L*. В то же время $\wedge_S x_\alpha \le x_\alpha$ для всех $x_\alpha \in S$ (так как объединение x_α и S есть S, то $x_\alpha \cap \wedge_S x_\alpha = \wedge_S x_\alpha$), и если $t \cap x_\alpha, = t$ для всех $x_\alpha \in$ S, то

$t \cap \wedge_S x_\alpha = \wedge(x_\alpha \cap t) = \wedge\ t = t$; следовательно, $\wedge_S x_\alpha = \inf S$.

Докажем теорему о неподвижных точках.

Теорема 8. *Пусть* $y = f(x)$ *— произвольное изотопное отображение полной структуры L на себя. Тогда для некоторого* $a \in L$ $a = f(a)$.

Доказательство. Определим а как наименьшую верхнюю грань множества S элементов $x \in L$ таких, что $x \le f(x)$. Множество S непусто, так как $0 \le f(0)$. Так как $f(x)$ изотонно и $a > x$ для всех $x \in S$, то $f(a) \ge f(x) \ge x$ для всех $x \in S$; следовательно, $f(a) \ge \sup S = a$. Так как $f(x)$ изотонно, отсюда следует, что $f(f(a)) \ge f(a)$; следовательно, $f(a) \in S$. Но это означает, что $f(a) \le a$, ибо $a = \sup S$. Мы заключаем, что $a = f(a)$.

6.5.5. Полярность

Пусть ρ — некоторое бинарное отношение, определенное между членами двух совокупностей I и I^*. Если $x \in I$, $x^* \in I^*$, то мы пишем $x\rho x^*$, когда x находится в отношении ρ к x^*, и $x\rho' x^*$, когда это не имеет места; таким образом, ρ' является дополнением ρ. Мы не требуем, чтобы I и I^* были различными.

Если X некоторое подмножество множества I, то мы обозначим через X^* множество элементов x^* таких, что $x\rho x^*$ для всех $x \in X$; обратно, если Y некоторое подмножество множества I^*, то мы обозначим через Y^* множество элементов $x \in I$ таких, что $x\rho y$ для всех $y \in Y$. Имеет место

Лемма 1. *Для любого отношения ρ имеют место неравенства:*

Если $X \ge X_1$ *в* I, *то* $X^* \le X^*_1$ *в* I^*, (1)

Если $Y \ge Y_1$ *в* /*, *то* $Y^+ \le Y^+_1$ *в* I, (1*)

$$X \leq (X^*)^+ \text{ и } Y \leq (Y^+)^* \text{ для всех } X \text{ в } I,\ Y \text{ в } I^*. \tag{2}$$

Так, например, из того, что $x\rho x^*$ для всех $x \in X$, $x^* \in X^*$, следует $X \leq (X^*)^+$.

Следствие. *Для любых подмножеств X из I и Y из I* имеем*

$$((X^*)^+)^* = X^* \text{ и } ((Y^+)^*)^+ = Y^+. \tag{3}$$

Доказательство. В силу (2) $X^* \leq ((X^*)^+)^*$; точно так жо и силу (2) $X \leq (X^*)^+$ откуда в силу 1 $X^* \geq ((X^*)^+)^*$; следовательно в силу P2 $((X^*)^+)^* = X^*$.

Теорема 9. *Операции* $X \to (X^*)_+$ *и* $Y \to (Y^+)^*$ *являются операциями замыкания; кроме того, соответствия* $X \to X^*$ *и* $Y \to Y^+$ *определяют, дуальный изоморфизм между полными структурами «замкнутых» подмножеств множества I и множества I*.*

Доказательство. В силу (2) соответствие $X \to (X^*)^*$ удовлетворяет С1; в силу (1) и (1*) из $X \geq X_1$ следует $(X^*)^+ \geq (X_1^*)^+$, что доказывает С2; в силу (3) $(((X^*)^+)^*)^+ = (X^*)^+$, что доказывает С3. Следовательно, $X \to (X^*)^+$ есть операция замыкания; аналогично для $Y \to (Y^*)^*$. В силу (3) соответствия $X^* \to (X^*)^+$ и $Y \to (Y^+)^*$ взаимно обратны в том смысле, что произведение их, взятое в любом порядке, есть тождественное соответствие; следовательно, оба эти соответствия взаимно однозначны. Наконец, в силу (1) они обращают отношение включения, чем и завершается доказательство.

Пример 1. Пусть $I = I^*$— некоторое кольцо и пусть $x\rho y^*$ означает, что $xy = 0$. Тогда каждое X^* есть правый идеал, каждое X^+ — левый идеал. В случае «регулярных» колец (включающем в себя полупростые линейные ассоциативные алгебры) справедливо также и обратное, а потому структуры правых и левых идеалов дуально изоморфны.

Пример 2. Пусть I—некоторое поле или кольцо с делением (тело) и пусть I^*—какая-нибудь конечная группа автоморфизмов α в I. Если мы определим $x\rho\alpha$. в том смысле, что $\alpha(x) = x$, то мы придем к хорошо известному дуальному изоморфизму между подгруппами в I^* и некоторыми подполями в I.

Имеются другие примеры, в которых $I = I^*$ и ρ симметрично. Следующий случай является типичным.

Пример 3.Пусть $A = \|a_{ij}\|$ — какая-нибудь симметричная $(n \times n)$-матрица; тогда $x_i a_{ij} x_j = 0$ определяет коническое сечение, квадратичную поверхность или квадратичную гиперповерхность Q в зависимости от того, будет ли $n=3$, $n = 4$ или $n > 4$. Для двух векторов $\xi = (x_0, .. ., x_n)$ и $\eta = (y_0, . .., y_n)$ проективного n-пространства I мы определим $\xi\rho\eta$ в том смысле, что $x_i a_{ij} y_j = 0$. «Замкнутыми» подмножествами пространства I являются тогда его точки, прямые, плоскости и другие подпространства; если X— какое-нибудь такое подпространство, то X^* есть его *поляра* относительно Q.

Последний пример наводит на мысль называть в общем случае множество X^* *полярой* X по отношению к ρ; он приводит также к следующему результату.

Следствие. *Если* I — I^* *и отношение* ρ *симметрично, то* $X^+ = X^*$ *и в полной структуре замкнутых множеств* $X = (X^*)^*$ *соответствие* $X \to X^*$ *является инволюцией. Символически*

$$(X^*)^* = X, \quad (4)$$

$$(X \cap Y)^* = X^* \cup Y^*, \ (X \cup Y)^* = X^* \cap Y^*. \quad (5)$$

Если ρ *также антирефлективно (если* $x\rho x$ *не имеет места ни для какого* x *или если из* $x\rho x$ *следует* $x\rho y$ *для всех* y*), то*

$$X \cap X^* = 0 \text{ и } X \cup X^* = I. \quad (6)$$

Пример 4. Пусть I—некоторая группа и пусть $x\rho y$ означает, что $xy = yx$. Тогда имеют место (4) — (5); замкнутые множества суть некоторые подгруппы, а соответствие $X \to X^*$ переводит каждую подгруппу в ее «централизатор».

Пример 5. Пусть I—некоторая совокупность и пусть $x\rho y$ означает, что $x \neq y$. Тогда имеют место (4) — (6); каждое множество замкнуто, а инволюция $X \to X^*$ переводит каждое множество в его теоретико-множественное дополнение.

Пример 6. Пусть I—линейное n-пространство и пусть $x\rho y$ означает, что $x \perp y$ (x ортогонально y). Тогда имеют место равенства (4) — (6); «замкнутые» подмножества суть линейные подпространства, и инволюция переводит каждое подпространство в его ортогональное дополнение. (Это есть частный случай примера 3.)

Доказательство. Так как $X^* = X^+$, то (4) следует из (3), а (5) из теоремы 9, ибо дуальный изоморфизм переставляет местами объединения и пересечения. Далее, в (6) множество $X \cap X^*$ содержит только такие элементы x, что $x\rho x$, т. е. только элементы из 0 (которое в примере 5 представляет собой пустое множество, а в примере 6—начало координат).

6.5.6. Связи Галуа

Предшествующие результаты могут быть следующим образом абстрактно обобщены на частично упорядоченные множества.

Определение. *Пусть* P, Q — *частично упорядоченные множества и пусть* $x \to x^*$, $y \to y^*$—*соответствия между ними, определенные для всех* $x \in P$, $y \in Q$ — *такие, что*

если $x \geq x_1$ *в* P, *то* $x^* \leq x^*_1$ *в* Q, (1)

если $y \geq y_1$ *в* Q, *то* $y^+ \leq y^+_1$ *в* P, (1*)

$x \leq (x^*)^+$ *и* $y \leq (y^+)^*$ *для всех* $x \in P$, $y \in Q$. (2)

Тогда говорят, что соответствия $x \to x^*$, $y \to y^*$ определяют связь Галуа между P и Q.

Формальные доказательства соотношения (3), приведенные в п.6.5. 5, и того факта, что соответствия $x \to (x^*)^+$ и $y \to (y^+)^*$ являются операциями замыкания, проходят тогда от начала до конца без изменений. Применением теоремы 1 получается

Теорема 10. *Пусть $x \to x^*$, $y \to y^*$ определяют связь Галуа между некоторыми двумя полными структурами L и M. Тогда связь Галуа дает дуальный изоморфизм между полными структурами S и T «замкнутых» подмножеств в L и в M.*

Мы получаем также перефразировку других результатов, установленных выше; так, например, для симметричных и антирефлексивных отношений имеют место равенства (4) —(6).

Хотя большинство связей Галуа имеет место между множествами, на что указывают приведенные выше примеры 1 — 6, но имеются также некоторые важные связи Галуа между элементами абстрактных структур. Вызывает интерес следующий пример.

Пример 7. Пусть L—некоторая конечная дистрибутивная структура и пусть $x^* = x^+$—объединение элементов y_α таких, что $x \cap y_\alpha = 0$. Тогда имеют место правила (1) —(5) и соотношение $x \cap x^* = 0$; кроме того, элементы $x^+ = x^*$ образуют двойственную себе структуру.

6.5.7. Теорема о представлении

Пусть P — произвольное частично упорядоченное множество и пусть $x\rho y$ в смысле п.6.5.5 означает, что $x \leq x$ в P. Тогда в силу определения X^* есть множество верхних граней, а X^+ — множество нижних граней для X; следовательно, $(X^*)^+$ есть множество всех нижних граней множества всех верхних граней для X.

В частности, если x—некоторый элемент из P, то x^* есть множество элементов $u > x$, а $(x^*)^+$ множество (главный идеал) элементов $t \leq x$. Следовательно, если $x > y$, то $(x^*)^+ \geq (y^*)^+$. Кроме того, если $a = \inf X$, то $t \leq x$ тогда и только тогда, если $t \leq x_\alpha$ для всех $x_\alpha \in X$; следовательно, $(a^*)^+$ является пересечением всех $(x_\alpha^*)^+$. Мы приходим к следующей теореме о представлении.

Теорема 11. *Любое частично упорядоченное множество P изоморфно совокупности Ф (P) подмножеств множества P таким образом, что наиб. н. г. в P переходят в пересечения.*

Следствие 1. *Отношение включения полностью характеризуется постулатами* P1 — P3.

Неверно, однако, что операции теоретико-множественного объединения и пересечения характеризуются постулатами LI — L4; эти операции удовлетворяют дистрибутивному закону L6, который не имеет места в общих структурах. Рассмотрим утверждение, обратное теореме 4.

Следствие 2. *Любая конечная структура L изоморфна структуре всех подалгебр некоторой абстрактной алгебры A.*

Доказательство. Пусть A состоит из элементов структуры L, на которой заданы: нулевая операция с постоянным значением 0, одинарные операторы α, по одному для каждого $a \in L$, определяемые посредством соотношения $\alpha(x) = x \cap a$, и бинарная операция $x \cup y$. «Подалгебрами» алгебры A являются тогда непустые идеалы структуры L; поскольку эти идеалы главные, они соответствуют изоморфным образом элементам структуры L.

Принадлежащее Макнилу видоизменение предшествующей конструкции показывает, что дедекиндово построение иррациональных чисел с помощью «сечений» может быть выполнено в действительности в произвольных частично упорядоченных множествах.

Теорема 12. *Любое частично упорядоченное множество P может быть вложено в полную структуру L так, что при этом сохраняется включение вместе со всеми наиб. н. г. и наим. в. г., существующими в P.*

Доказательство. Сперва мы присоединяем к P элемент 0, если только P не имеет уже наименьшего элемента. Пусть тогда L состоит из всех непустых «замкнутых» подмножеств $X = (X^*)^+$ из P, в силу теоремы 9 L есть полная структура. В силу теоремы 11 соответствие $x \to {}^*(x^*)^+$ вкладывает P в L с сохранением включения и наиб. н. г. Предположим теперь, что $a = \sup X$ в P. Тогда в силу (3) $(T^*)^+ \geq (X^*)^+$ в L в том и только в том случае, когда $T^* \leq X^*$; но в силу определения наим. в. г. $X^* = a^*$. Следовательно, $(T^*)^+ \geq (X^*)^+$ тогда и только тогда, если $T^* \leq a^*$, или в силу (1*) если $(T^*)^+ \geq (a^*)^+$. Следовательно, $(a^*)^+$ есть наим. в. г. для $(X^*)^+$, а потому наим. в. г. сохраняются, что и требовалось доказать.

Замечание. Предыдущая конструкция не является единственным логически возможным вложением частично упорядоченных множеств в полные структуры.

Если L структура и X произвольное подмножество в L, то $(X^*)^+$ является *идеалом.* В самом деле, если $a, b \in (X^*)^+$, то a, b суть нижние грани для X; следовательно, X^* будет верхней гранью для a и b, а потому и для $a \cup b$, откуда $a \cup b \in (X^*)^+$. Далее, если $a \in (X^*)^+$ и $t \leq a$, то, $t \in (X^*)^+$. Приведем следующее определение.

Определение. *Замкнутый идеал частично упорядоченного множества P есть подмножество множества P, содержащее (на самом деле, состоящее из) все нижние грани множества своих верхних граней.*

6.5.8. Внутренние топологии

Пусть $\{x_\alpha\}$ — произвольное направленное множество элементов полной структуры L. Пусть по определению $x_\alpha \to a$ (словами, x_α сходится по упорядоченности к a) означает, что

$$a = \bigwedge_{\alpha} \{ \bigvee_{\beta \geq \alpha} x_\beta\} = \bigvee_{\alpha} \{ \bigwedge_{\beta \geq \alpha} x_\beta\}, \tag{7}$$

т. е. что Lim sup $\{x_\alpha\}$ = Lim inf $\{x_\alpha\}$ $=a$. Полагая $u_\alpha = \vee_{\beta\geq\alpha} x_\beta$ и $\upsilon_\alpha = \wedge_{\beta\geq\alpha} x_\beta$, мы видим, что $x_\alpha \to a$ тогда и только тогда, если существуют монотонные направленные множества $u_\alpha \downarrow a$ и $\upsilon_\alpha \uparrow a$ такие, что $u_\alpha \geq x_\alpha \geq \upsilon_\alpha$, для всех α. Можно доказать соответствующие соотношения для обычных счетных последовательностей в любой σ-структуре. Если P произвольное частично упорядоченное множество, то мы можем определить $x_\alpha \to a$ в P в том смысле, что $x_\alpha \to a$ в пополнении P с помощью сечений; эта «относительная топология» делает возможным дальнейшее исследование. Мы знаем, что даже в цепи из $x_\alpha \to a$ не обязано вытекать существование $u_\alpha \downarrow a$ и $\upsilon_\alpha \uparrow a$ с $u_\alpha \geq x_\alpha \geq \upsilon_\alpha$ для всех α, хотя обратное справедливо.

Легко проверить первые две из приводимых ниже формул:

если $x_\alpha = a$ для всех α, то $x_\alpha \to a$, (8)

если $x_\alpha \to a$ и $x_\alpha \to b$, то $a = b$, (9)

если $x_\alpha \to a$ и $\{ x_\alpha \}$ конфинальное подмножество множества $\{ x_\alpha \}$, то $x_\sigma \to a$. (10)

Для доказательства (10) заметим, что для любого α существует $\sigma \geq \alpha$; тогда $\vee_{\tau\geq\sigma} x_\tau$ в $\{ x_\sigma \}$ содержится в $\vee_{\beta\geq\alpha} x_\beta$ из $\{ x_\alpha \}$, ибо каждое x_τ есть некоторое x_β. Следовательно, используя двойственность, мы получаем

$$a = \text{Lim sup}\,\{ x_\alpha \} \geq \text{Lim sup}\,\{ x_\sigma \} \geq \text{Lim inf}\,\{ x_\sigma \} \geq \text{Lim inf}\,\{ x_\alpha \} = a,$$

откуда непосредственно вытекает (10).

Пример 1. Пусть L—полная структура всех подмножеств бесконечной совокупности I. Пусть индексами F являются различные конечные подмножества из I и пусть $F \geq G$ означает, что F содержит G. Пусть x_F обозначает множество F; тогда $\{x_F\}$ есть монотонное направленное множество. Легко показать, что $x_F \uparrow I$, несмотря на то, что в случае несчетного I никакое вполне упорядоченное направленное множество (т. е. трансфинитная последовательность) элементов x_F не может сходиться к I. (В самом деле, для всех τ infimum элементов x_α $(\alpha \geq \tau)$ должен быть конечным; но объединение любой возрастающей вполне упорядоченной последовательности конечных множеств пе более чем счетно.)

Мы называем подмножество X частично упорядоченного множеств P замкнутым в топологии упорядоченности тогда и только тогда, если из $\{ x_\alpha \} \leq X$ и $x_\alpha \to a$ следует $a \in X$; словами, тогда и только тогда, если предел любого сходящегося по упорядоченности направленного множества элементов из X сам принадлежит X. Из (8) — (10) вытекает

Теорема 13. *Каждое частично упорядоченное множество P является хаусдорфовым пространством относительно своей топологии упорядоченности.*

Другим результатом является следующая

Лемма 2. *Любой замкнутый интервал* $[a, b]$, $[-\infty, b]$, $[a, +\infty]$ *или* $[-\infty, +\infty]$ *замкнут в своей топологии упорядоченности.*

Доказательство. Если $x_\alpha \to c$, где $a \leq x_\alpha \leq b$ для всех α, то каждое $\vee_{\beta \geq \alpha} x_\beta \geq a$; следовательно, $c = \wedge_\alpha \{ \vee_{\beta \geq \alpha} x_\beta \} \geq a$. Двойственным образом $c \leq b$, что и завершает доказательство.

Определение. *Под интервальной топологией частично упроядоченного множества P мы понимаем топологию, определенную путем взятия замкнутых интервалов множества P в качестве псевдобазиса замкнутых множеств.*

Заметим, что, так как любой конечный интервал является пересечением двух полубесконечных интервалов, мы можем в качестве псевдобазиса замкнутых множеств взять замкнутые полубесконечные интервалы. Заметим также, что мы не можем выбирать «открытые интервалы» в качестве базиса открытых множеств. Рассмотрим, например, координатную плоскость и предположим, что $(x, y) \leq (x', y')$ означает $x \leq x'$ и $y \leq y'$. Тогда $(-1, 0) < (x, y) < (1, 0)$ не будет открытым множеством ни в каком естественном смысле.

Поскольку каждая точка есть замкнутый интервал, имеет место

Теорема 14. *Каждое частично упорядоченное множество P является T_1-пространством относительно своей интервальной топологии.*

В силу леммы 1 интервальная топология «слабее» топологии упорядоченности или же гомеоморфна ей. Приведем пример, где она слабее.

Пример 2. Пусть L—полная структура, состоящая из 0, I и счетного множества элементов $x_1, x_2, x_3, \ldots$, причем $0 < x_i < I$ для всех i, но отношение $x_i < x_j$ не имеет места ни для каких $i = j$. Рассмотрим последовательность $\{x_n\}$. В интервальной топологии $x_n \to 0$ и $x_n \to I$; в топологии упорядоченности L является дискретным пространством.

Этот пример иллюстрирует два факта: L есть хаусдорфово пространство в своей топологии упорядоченности, но не является таковым в своей интервальной топологии; L есть компактное пространство в своей интервальной топологии, но не является таковым в своей топологии упорядоченности.

Теорема 15. *Каждая полная структура компактна в своей интервальной топологии.*

Доказательство. В силу теоремы 12, гл. III, достаточно показать, что замкнутые интервалы обладают «свойством конечных пересечений». Пусть C—какой-нибудь класс замкнутых интервалов $[a_\gamma, b_\gamma]$ такой, что пересечение $[a_\gamma \cup a_{\gamma'}, b_\gamma \cap b_{\gamma'}]$ любых *двух* интервалов из C непусто. Тогда $a_\gamma \leq b_{\gamma'}$ тождественным образом; следовательно, $a = \wedge a_\gamma \leq \wedge b_\gamma = b$, а потому *все* $[a_\gamma, b_\gamma]$ из C содержат $[a, b]$ и имеют непустое пересечение.

6.5.9. Звездная сходимость; полунепрерывные функции

Ранее мы видели, что между совокупностями сходящихся направленных множеств из множества S и совокупностями замкнутых (или открытых) подмножеств множества S имеет место связь Галуа. В этом соответствии совокупность замкнутых множеств является замкнутым множеством тогда и только тогда, если оно есть T_1-пространство; мы не знаем необходимых и достаточных условий замкнутости множества сходящихся направленных множеств. Мы знаем, однако, что условия (8) — (10) являются необходимыми. Справедливо также условие:

если из точек каждого конфинального подмножества $\{x_\beta\}$ направленного множества $\{x_\alpha\}$ можно построить направленное множество, сходящееся к a, то $x_\alpha \to a$. (11)

Доказательство. Мы знаем, что имеем дело с T_1-пространством. Если $x_\alpha \to a$ не имеет места, то некоторое открытое множество U содержит a и не содержит *ни одного* из последующих некоторого x_α. Однако множество последующих элемента x_α конфинально в $\{x_\alpha\}$. Никакое направленное подмножество точек этого конфинального подмножества, целиком лежащего вне окрестности U, не сходится к a.

Определение. *Звездная топология на структуре есть T_1-топология, определенная сходимостями по упорядоченности.*

В частности, для обычных последовательностей $\{x_n\}$ мы будем писать $x_n \to^* a$ и говорить, что $\{x_n\}$ звездно сходится к a, если каждая подпоследовательность $\{x_{n(k)}\}$ последовательности $\{x_n\}$ содержит подподпоследовательность $\{x_{n(k(i))}\}$, сходящуюся по упорядоченности к a при $i\to\infty$. В этом случае $\{x_n\}$ сходится к a в звездной топологии;

кроме того, этот случай достаточен для приложений звездной сходимости.

Различные интересные проблемы, связанные со звездной сходимостью, упоминаются ниже в упражнениях.

В общем случае топологическая алгебра может быть определена как топологическое пространство с алгебраическими операциями, которые непрерывны в топологии. Руководствуясь этой идеей (хотя мы могли бы использовать также звездную сходимость или интервальную топологию), мы определим *топологическую структуру* как структуру с топологией, при которой

из $x_\alpha \to x$ и $y_\beta \to y$ следует $x_\alpha \cap y_\beta \to x \cap y$ и $x_\alpha \cup y_\beta \to x \cup y$. (12)

В случае обычных последовательностей это эквивалентно условию

из $x_n \to x$ и $y_n \to y$ следует $x_n \cap y_n \to x \cap y$ и $x_n \cup y_n \to x \cup y$. (12')

Это условие не имеет места в общих структурах .

Лемма 3. *Для того, чтобы* (12) *имело место в полной структуре при ее топологии упорядоченности, достаточно, чтобы выполнялось условие*

из $x_\alpha \uparrow x$ *следует* $a \cap x_\alpha \uparrow a \cap x$ *и двойственно.* (13)

Доказательство. (Условие необходимо.) Пусть в обозначениях (12) u_α — пересечение элементов, следующих за x_α, и v_β—пересечение элементов, следующих за y_β; мы должны доказать, что $u_\alpha \cap v_\beta \uparrow x \cap y$; в силу двойственных соображений этим утверждение будет доказано. Но u_α, v_β и $u_\alpha \cap v_\beta$ изотонны; следовательно, достаточно доказать, что $\sup(u_\alpha \cap v_\beta) = x \cap y$; это мы теперь и проделаем. Так как $x \geq u_\alpha$ и $y \geq v_\beta$, то $x \cap y \geq u_\alpha \cap v_\beta$ для всех α, β, а потому $x \cap y \geq \sup(u_\alpha \cap v_\beta)$. Обратно, для всех a в силу (13) $\sup(u_a \cap v_\beta) \geq u_\alpha \cap \sup v_\beta = u_\alpha \cap y$; следовательно, еще раз применяя (13), имеем $\sup(u_a \cap v_\beta) \geq \sup(u_\alpha \cap y) = x \cap y$.

Мы получим теперь результат, напоминающий теорему 8; для его доказательства нам нужно понятие полунепрерывной функции.

Определение. *Функция* $y=f(x)$, *задающая отображение* T_1*-пространства на частично упорядоченное множество, называется полунепрерывной снизу, если из* $x_\alpha \to a$ *и* $f(x_\alpha) \leq c$ *для всех* x_α *следует* $f(a) \leq c$.

Полунепрерывные снизу вещественнозначные функции полезны не только в общем анализе; они играют центральную роль в теоремах существования вариационного исчисления. Это показал Тонелли, следуя Фреше.

Теорема 16. *Полунепрерывная снизу функция, задающая отображение компактного пространства S в частично упорядоченное множество P, принимает минимальное значение.*

Доказательство. Образуем максимальную цепь M элементов $y \in P$, таких, что $f(x) \leq y$ для некоторого $x \in S$. Такая цепь должна существовать, ибо свойство быть цепью и свойство иметь $f(x) \leq y$ Для некоторого $x \in S$ и для всех $y \in M$ суть свойства конечного характера (соответственно, характера два и характера один). Пусть для каждого $y \in M$ S_y есть непустое множество элементов $x \in S$, для которых $f(x) \leq y$. Так как $f(x)$ полунепрерывно снизу, то каждое S_y замкнуто. Так как S_y образуют просто упорядоченное множество относительно теоретико-множественного включения, то любая конечная система множеств S_y имеет непустое пересечение. Так как S компактно, то существует, следовательно, точка a, содержащаяся в каждом $S_y[y \in M]$. Соотношение $f(x) < f(a)$ $[x \in S]$ означало бы тогда, что $f(x)$ можно присоединить к M, а это противоречит предположению максимальности M. Мы заключаем отсюда, что $f(a)$ есть требуемое минимальное значение.

Следствие. *Полунепрерывная снизу функция, задающая отображение компактного пространства в цепь, принимает наименьшее значение.*

Двойственные этим результаты также справедливы.

6.6. Дедекиндовы структуры

6.6.1. Определение и примеры

Подобно тому, как элементы многих групп удовлетворяют тождеству $xy = yx$, так и элементы многих структур удовлетворяют следующему «модулярному тождеству»

L5. Если $x \leq z$, то $x \cup (y \cap z) = (x \cup y) \cap z$.

Структура называется *дедекиндовой* тогда и только тогда, если ее элементы удовлетворяют модулярному тождеству L5. Дедекиндовы структуры называют также дедекиндовыми решетками (Dedekind lattices, Dedckindische Verbande) и модулярными структурами (Modular lattices).

Теорема 1. *Нормальные делители произвольной группы образуют дедекипдову структуру.*

Доказательство. В любой структуре мы имеем односторонний модулярный закон; следовательно, нужно лишь показать, что из $x \leq z$ следует $x \cup (y \cap z) \geq (x \cup y) \cap z$. Предположим, что X, Y, Z—нормальные делители группы, причем $X \leq Z$. Тогда $X \cup Y$ есть множество XY всех произведений $xy[x \in X, y \in Y]$, как это доказывается в теории групп. (Очевидно, что каждое такое $xy \in X \cup Y$; в то же время, поскольку $(xy)(x'y') = (xx')((x'^{-1}yx')y')$ и $(xy)^{-1} = (y^{-1}x^{-1}y)y^{-1}$ принадлежат XY, множество XY есть вся подгруппа $X \cup Y$.) Следовательно, $(X \cup Y) \cap Z$ состоит из элементов xy, лежащих в Z. Но в этом случае $y = x^{-1}(xy) \in Z$, ибо $X \leq Z$; следовательно, xy лежит в $X \cup (Y \cap Z)$ и $(X \cup Y) \cap Z \leq X \cup (Y \cap Z)$, что и требовалось доказать.

Ниже будет доказано более общее утверждение, что Ω-подгруппы любой группы с системой Ω операторов, включающей в себя все внутренние автоморфизмы, образуют дедекиндову структуру.

Будет показано, что элементы любой проективной геометрии образуют дедекиндову структуру. Будут рассмотрены другие примеры структур, удовлетворяющих более сильному *дистрибутивному* закону $x \cap (y \cup z) = (x \cap y) \cup (x \cap z)$. В этом разделе мы ограничимся рассмотрением абстрактных дедекиндовых структур и их свойств.

Отметим, что структура, двойственная дедекиндовой структуре, есть дедекиндова структура и что любые подструктура, гомоморфный образ или произведение дедекиндовых структур суть дедекиндовы структуры.

6.6.2. Альтернативные характеристики

Пусть L—произвольная структура и пусть в L выбраны $z > x$ и y. В силу одностороннего модулярного закона имеем

$$x \cap y \leqslant x \leqslant x \cup (y \cap z) \leqslant (x \cup y) \cap z \leqslant x \cup z \leqslant y \cup z,$$
$$x \cap y \leqslant y \cap z \leqslant z \leqslant x \cup z \leqslant y \cup z.$$

Используя двойственность, видим, что справедливы все отношения включения в диаграмме на рис. 3, *а*.

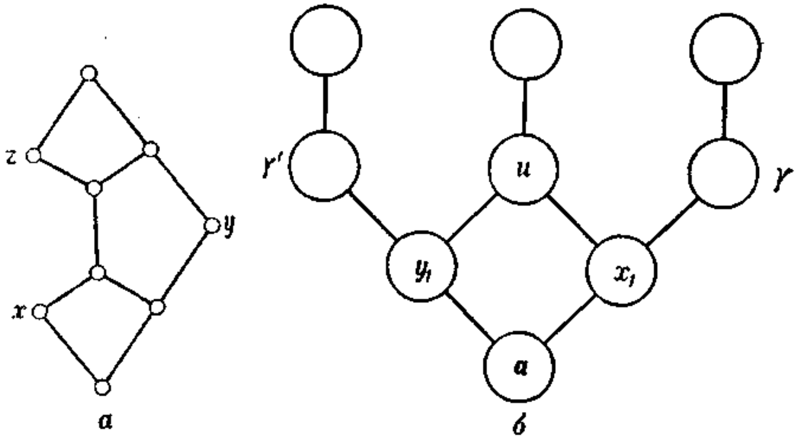


Рис.3

Кроме того, $[(x \cup y) \cap z] \cup y \geq x \cup y$ ибо $(x \cup y) \cap z \geq x$; однако $x \cup y$ является верхней гранью для $(x \cup y) \cap z$ и y. Следовательно, $[(x \cup y) \cap z] \cup y = x \cup y$ и двойственно этому. Таким образом, *подструктура произвольной структуры, порожденная элементами* $z > x$ *и* y*, является гомоморфным образом структуры, изображенной на рис. 3, а.*

Теорема 2. *Структура L не является дедекиндовой тогда и только тогда, если она содержит подструктуру, изоморфную пятиэлементной структуре, изображенной на рис, 1, г.*

Действительно, эта структура не дедекиндова; если, с другой стороны, для элементов $z > x, y$ не выполняется L5, то подструктура рис. 3, *а,* состоящая из элементов y, $x \cup y$, $y \cap z$, $(x \cup y) \cap z$ и $x \cup (y \cap z)$, удовлетворяет нашим условиям.

Следствие 1. *В дедекиндовой структуре отношение* $v>u$ *несовместимо с* $y \cup u = y \cup v$ *и* $y \cap u = y \cap v$; *обратно, из этого условия следует дедекиндовость.*

Следствие 2. *В дедекиндовой структуре*

(ξ1'). *Если* y *покрывает* a *и* $a<x \not\leq y$, *то* $x \cup y$ *покрывает* x *и двойственно этому,*

(ξ1''). *Если* a *покрывает* y *и* $a > x \not\leq y$, *то* x *покрывает* $x \cap y$..

Доказательство. Если бы $x \cup y$ не покрывало x, то подструктура, порожденная элементами x, y и каким-нибудь z, лежащим между x и $x \cup y$, была бы запрещенной не дедекиндовой пятиэлементной структурой.

В частности, из следствия 2 мы получаем двойственные *условия покрытия:*

(ξ'). Если x и y покрывают a и $x \neq y$, то $x \cap y$ покрывает x и y,

(ξ"). Если a покрывает x и y и $x \neq y$, *то* x и y покрывают $x \cap y$

Условимся под выражением «x самое большее покрывает a» понимать: «a покрыват a или $x = a$». Тогда из *(ξ')* вытекает

(ξ2'). Если x и y самое большее покрывают a, то $x \cup y$ самое большее покрывает x.

Действительно, случаи $x = a$, $x = y$ и $y = a$ являются тривиальными. Из *(ξ")* может быть выведено двойственное условие (ξ2").

Пусть теперь L—произвольная структура, в которой все ограниченные цепи конечны. Покажем, что из *(ξ')* вытекает условие Жордана — Дедекинда. Пусть для каждого целого положительного числа m $P(m)$ представляем собой утверждение, что если одна максимальная цепь γ: $a = x_0 < x_1 < \ldots < x_m = b$ имеет длину m, то и каждая максимальная цепь между a и b имеет длину m.(Под «максимальной» или «связной» цепью мы подразумеваем, что x_i покрывает x_{i+1} для всех $i = 1, \ldots, m$.) P (1) травиально в любой структуре; мы покажем, что из P (m— 1) следует $P(m)$. В самом деле, пусть γ': $a = y_0 < y_1 < \ldots < y_n = b$— какая-нибудь другая (конечная) максимальная цепь, связывающая a и b; положим $u = x_1 \cup y_1$ (рис. 3,*б*). Случай $x_1 = y_1$ тривиален. Если же $x_1 \neq y_1$, то в силу (ξ') u покрывает x_1 и y_1; образуем какую-нибудь максимальную цепь γ'', связывающую u и b. В силу $P(m$ — 1) цепь x_1, γ'' имеет длину m—1; например, $x_1 < x_2 < \ldots \ldots < x_m$; следовательно, γ'' имеет длину m —2; поэтому y_1, γ'' имеет длину m — 1; следовательно, снова в силу P (m — 1) $y_1 < y_2 < \ldots \ldots < y_m$ имеет длину m — 1, а потому $m = n$.

Отсюда следует, что мы можем определить функцию *размерности* $d[x]$, такую, что x покрывает y тогда и только тогда, если $x > y$ и $d\,[x] = d\,[y]+1$. Так как условие Жордана — Дедекинда двойственно себе, то оно вытекает равным образом из условия (ξ"), двойственного *(ξ')*. Кроме того, мы можем доказать следующую лемму:

Лемма 1. *Из* (ξ') *вытекает неравенство* $d\,[x]+d\,[y] \geq d\,[x \cup y]+d\,[x \cap]$; *обратное неравенство вытекает из* (ξ").

Доказательство. Образуем произвольные связные цепи

$$x \cap y = x_0 < x_1 < \ldots < x_m = x, \quad x \cap y = y_0 < y_1 < \ldots < y_n = y.$$

Предполагая тогда по индукции, что $x_{i-1} \cup y_j$ и $x_i \cup y_{j-1}$ самое большее покрывают $x_{i-1} \cup y_{j-1}$, заключаем из (ξ2'), что $x_i \cup y_j = (x_{i-1} \cup y_j) \cup (x_i \cup y_{j-1})$ самое большее покрывает $x_{i-1} \cup y_j$ и $x_i \cup y_{j-1}$. Мы заключаем отсюда, что $d[x \cup y_j] — d[x \cup y_{j-1}] \leq 1$ и, следовательно,

$d\,[x \cup y] - d\,[x] \le n \le d[y] - d[x \cap y]$, что и доказывает первое утверждение. Второе утверждение вытекает из двойственных соображений.

Отсюда следует, что во всякой структуре *L,* у которой все ограниченные цепи конечны и для которой справедливы (ξ') — (ξ''), имеет место условие Жордана — Дедекинда и

$$d\,[x] + d\,[y] = d\,[x \cap y] + d\,[x \cup y]. \qquad (1)$$

Если *L*—произвольная структура с числовой функцией, удовлетворяющей (1), то мы видим, что при $x \le z$

$$d\,[x \cup (y \cap z)] - d\,[(x \cup y) \cap z] = d\,[x] + d\,[y \cap z] - \\ - d\,[x \cap y] - d\,[x \cup y] - d\,[z] + d\,[y \cup z].$$

Это сводится после перестановки слагаемых к $d\,[y] - d\,[y] = 0$, ибо $d\,[x] - d\,[x \cap y] - d\,[x \cup y] = -d\,[y]$ и аналогично для трех остальных членов. Следовательно, если из $u > v$ следует, как и в данном случае, $d[u] > d[v]$, то мы не можем иметь $x \cup (y \cap z) < (x \cup y) \cap z$. В силу одностороннего модулярного закона заключаем, что должно иметь место- L5. В результате доказана

Теорема 3. *Пусть L—произвольная структура, в которой все ограниченные цепи конечны. Следующие условия являются эквивалентными:*

а) *модулярное тождество,*

б) *условия покрытия (ξ') — (ξ''),*

в) *условие Жордана—Дедекинда вместе с* (1).

Заметим, что дедекиндовы структуры можно определить как структуры, в которых выполняется тождественное соотношение:

$$(x \cap z) \cup (y \cap z) = [(x \cap z) \cup y] \cap z.$$

Условия покрытия дают легко применимый графический критерий. Рассмотрим некоторые более чистые алгебраические конструкции.

6.6.3. Свободная дедекиндова структура с тремя образующими

Определим свободную дедекиндову структуру с тремя образующими.

Теорема 4. *Свободная дедекиндова структура с тремя образующими имеет двадцать восемь элементов и диаграмму, изображенную на рис. 4.*

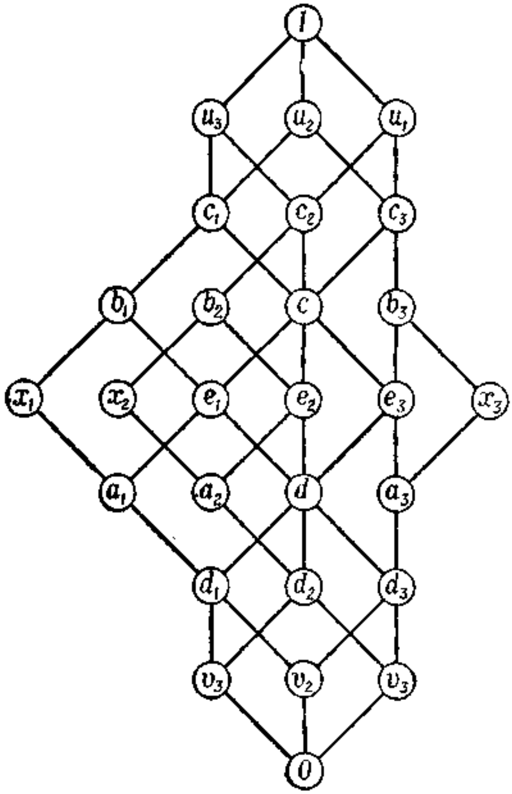


Рис. 4.

Доказательство. Поскольку частично упорядоченное множество L_{28}, изображенное на рис. 4, имеет двенадцать симметрий (все перестановки индексов и двойственность), то нетрудно доказать путем рассмотрения ограниченного числа характерных случаев, что каждая пара элементов имеет объединение (и пересечение). Модулярный закон можно тогда легко проверить, используя (ξ') и двойственность. Далее, эта структура порождается элементами x_1, x_2, x_3. В самом деле, $u_1 = x_2 \cup x_3$ и циклически, $\upsilon_1 = x_2 \cap x_3$ и циклически, $I = x_1 \cup x_2 \cup x_3$, $0 = x_1 \cap x_2 \cap x_3$, $a_i = x_i \cap u_i$, $b_i = x_i \cup \upsilon_i$, $c_i = u_2 \cap u_3$ и циклически, $d_1 = \upsilon_2 \cup \upsilon_3$ и циклически, $c = u_1 \cap u_2 \cap u_3$, $d_1 = \upsilon_1 \cup \upsilon_2 \cup \upsilon_3$ и $e_i = u_i \cap (x_i \cup \upsilon_i) = (u_i \cap x_i) \cup \upsilon_i$.

Обратно, все соотношения, изображенные на рис. 4, являются следствиями условий LI — L5. Необходимые для доказательства этого выкладки можно значятейьно сократить по соображениям симметрии. Мы дадим только два образчика,

$$a_1 \bigcup a_2 = (x_1 \bigcap u_1) \bigcup (x_2 \bigcap u_2) = ((x_1 \bigcap u_1) \bigcup x_2) \bigcap u_2$$

$$\text{(в силу L5)} = (u_1 \bigcap (x_1 \bigcup x_2)) \bigcap u_2 = u_1 \bigcap u_3 \bigcap u_2 = c.$$

$$b_2 \bigcup e_3 = (x_2 \bigcup v_2) \bigcup (a_3 \bigcup v_3) = x_2 \bigcup a_3 \text{ (в силу закона}$$

$$\text{поглощения) (так как } v_3 \leqslant x_2 \text{ и } v_2 = x_3 \bigcap x_1 \leqslant x_3 \bigcap u_3 = a_3)$$

$$= x_2 \bigcup (x_3 \bigcap (x_1 \bigcup x_2)) = (x_2 \bigcup x_3) \bigcap (x_1 \bigcup x_2) = c_2.$$

Эти образчики типичны тем, что в них используется L5 в качестве *смешанного ассоциативного закона.*

С другой стороны, свободная дедекиндова структура, порожденная четырьмя образующими, бесконечна. В самом деле, рассмотрим структуру, образованную точками и прямыми проективной плоскости I, вместе с I и пустым множеством 0. Легко проверить справедливость условия «в» теоремы 3; следовательно, эта структура дедекиндова. Однако подструктура, порожденная четырьмя точками (0,0), (1,0), (0,1) и (1,1), является бесконечной.

Этот пример в действительности сводится к напоминанию, что гармоническая сетка, порожденная полным четырехугольником, бесконечна (Веблен и Юнг). Свободная дедекиндова структура с четырьмя образующими имеет даже бесконечную длину. Пусть I— абелева группа, порожденная элементами $x_1, x_2, \ldots, x_n, \ldots$ Пусть S_1— подгруппа, порожденная элементами x_{2k+1}, S_2 — элементами $x_{2k}+x_{2k+1}$, S_3—элементами x_{2k+1}, S_4—элементами $x_{2k+1}+x_{2k+2}$, где $k = 0, 1, 2, \ldots$ Пусть

$$T_1 = S_1,\ T_{4n+1} = T_{4n} \cap S_1,\ T_{4n+2} = T_{4n+1} \cup S_2,$$

$$T_{4n+3} = T_{4n+2} \cap S_3,\ T_{4n+4} = T_{4n+3} \cup S_2.$$

Тогда, как легко проверить $T_4 > T_8 > T_{12} > \ldots$

6.6.4. Свободная дедекиндова структура, порожденная двумя цепями

Пусть L—произвольная структура и пусть $0 = x_0 < x_1 < \ldots \ldots < x_m = I$ и $0 = y_0 < y_1 < \ldots < y_n = I$—две цепи в L между 0 и I. Очевидно, что множество элементов $u^i_j = x_i \cap y_j$, включает в себя все x_i и y_j (ибо $x_i \cap y_n = x_i$ и $x_m \cap y_j = y_j$); двойственно этому множество элементов $\upsilon^i_j = x_i \cup y_j$; также включает в себя x_i и y_j. Следовательно, то же самое справедливо для множества *объединений* элементов u^i_j и для множества *пересечений* элементов υ^i_j.

Лемма 2. *Всякое объединение элементов* u^i_j *может быть записано в виде* $(x_{i(1)} \cap y_{j(1)}) \cup \ldots \cup (x_{i(r)} \cap y_{j(r)})$, *где* $i(1) > \ldots$ $\ldots > i(r)$ *и* $j(1) < \ldots < j(r)$.

Доказательство. Если два элемента u^i_j имеют один и тот же верхний индекс, то, так как y_j образуют цепь, одно u^i_j должно содержаться в другом и, следовательно, в силу L4 может быть отброшено. Таким образом, мы можем сделать все *i(k)* и аналогично все *j(k)* различными. Кроме того, если $i > i'$ и $j > j'$, то $x_i \cap y_j$ будет поглощать $x'_i \cap y'_j$. Следовательно, после того как мы отбросим столько элементов, сколько возможно, и используем L2 для представления *i(k)* в порядке убывания, будем иметь также $j(1) < \ldots < j(r)$.

Л ем м а 3. *Если в дедекиндовой структуре* $a_i \geq a_{i+1}$ и $b_i \leq b_{i+1}$ *для всех i, то*

$$(a_1 \cap b_1) \cup \ldots \cup (a_r \cap b_r) = a_1 \cap (b_1 \cup a_2) \cap \ldots \cap (b_{r-1} \cup a_r) \cap b_r,$$

$$(b_1 \cup a_1) \cap \ldots \cap (b_r \cup a_r) = b_1 \cup (a_1 \cap b_2) \cup \ldots \cup (a_{r-1} \cap b_r) \cup a_r.$$

Доказательство. Используя двойственность и индукцик> по *r,* мы должны доказать лишь первое тождество в предположении, что второе справедливо для меньшего, чем *г,* числа слагаемых. Но в силу L5, применяемого дважды, $(a_1 \cap b_1) \cup \ldots \cup (a_r \cap b_r)$ может быть переписано в виде

$$a_1 \cap [b_1 \cup (a_2 \cap b_2) \cup \ldots \cup (a_{r-1} \cap b_{r-1}) \cup a_r] \cap b_r.$$

В силу второго тождества для случая $(r — 1)$ два выражения

$$(b_1 \cup a_2) \cap (b_2 \cup a_3) \cap \ldots \cap (b_{r-1} \cup a_r),$$

$$b_1 \cup (a_2 \cap b_2) \cup (a_3 \cap b_3) \cup \ldots \cup (a_{r-1} \cap b_{r-1}) \cup a_r$$

равны между собой. Если мы подставим первое из них вместо последнего в квадратные скобки написанного выше выражения, то получим правую часть первого тождества, что и требовалось доказать.

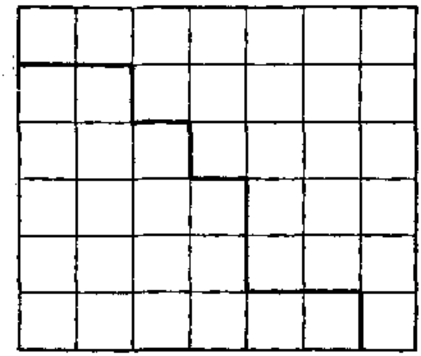
Лемма 4. *Объединения элементов* u^i_j *образуют подструктуру.*

Доказательство. Очевидно, что объединение объединений элементов u^i_j есть объединение элементов u^i_j; всякое же пересечение объединений элементов u^i_j является в силу лемм 1—2 пересечением пересечений элементов v^i_j и, следовательно, пересечением элементов v^i_j, а потому в силу лемм 1—2 объединением элементов u^i_j.

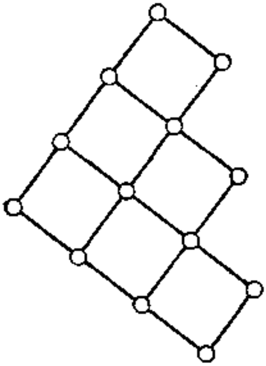
Отметим теперь, что если X_i обозначает множество точек (x, y) прямоугольника $0 \leq x \leq m$, $0 \leq y \leq n$, удовлетворяющих условию $x \leq i$, а Y_j обозначает множество точек, удовлетворяющих условию $y \leq j$, и если объединения и пересечения понимаются как теоретико-множественные объединения и произведения, то все выражения, допускаемые в лемме

1, описывают различные множества (пилообразной формы). Например, рис. 5, *а* изображает

$$(x_2 \cap y_5) \cup (x_3 \cap y_4) \cup (x_4 \cap y_2) \cup (x_6 \cap y_1) \quad [m = 6,\ n = 5].$$



а



б

Рис. 5.

В силу леммы 3 эти множества образуют *кольцо* множеств (т. е. подструктуру структуры всех подмножеств квадрата), которое, таким образом, представляет изоморфно *свободную* дедекиндову структуру, порожденную двумя цепями. Мы заключаем, что эта структура дистрибутивна и конечна. В результате доказана

Теорема 5. *Свободная дедекиндова структура, порожденная двумя конечными цепями, является конечной дистрибутивной структурой.*

Следствие. *Две любые конечные цепи с одними и теми же конечными точками обладают уплотнениями, подинтервалы которых попарно проективны* (см. 6.6.5).

6.6.5. Принцип транспозиции Дедекинда

Как мы увидим, понятие *проективности* применимо в общем случае к дедекиндовым структурам.

Определение. *(Замкнутый) интервал [x, y] называется «простым», если y покрывает x. Интервалы, которые могут быть записаны в виде $[x \cap y, x]$ и $[y, x \cup y]$, называются транспонированными; в то же время два фактора, [x, y] и [x', y'] называются проективными (что обозначается в виде [x, y]~[x'', y']), если существует конечная последовательность [x, y], $[x_1, y_1]$, $[x_2, y_2]$…, в которой любые два последовательных фактора транспонированы.*

Теорема 6. *Пусть L—произвольная дедекиндова структура и пусть и и υ — два каких-нибудь элемента из L. Тогда соответствия* $x \to u \cup x$ *и* $y \to v \cap y$ *являются инверсными изоморфизмами между* $[u \cap v, v]$ *и* $[u, u \cup v]$. *Кроме того, они переводят факторы в этих интервалах в транспонированные факторы.* .

Доказательство. Если $u \cap v \le x \le v$, то $u = u \cup (u \cap v) \le u \cup x \le u \cup v$, и если $x \le x'$, то $u \cap x \le u \cup x'$. Следовательно, первое соответствие есть изотонное отображение $[u \cap v, v]$ на подмножество множества $[u, u \cup v]$. Двойственным образом второе соответствие есть изотонное отображение $[u, u \cup v]$ на подмножество множества $[u \cap v, v]$. Но из $u \cap v \le x \le v$ вытекает, в силу L5, что $v \cap (u \cup x) = (v \cap u) \cup x = x$; двойственным образом, $u \cup (v \cap y) = y$ для всех $y \in [u, u \cup v]$. Таким образом, эти два соответствия взаимно обратны; следовательно, оба они взаимно однозначны, а потому являются изоморфизмами. Покажем, что они производят транспозиции.

Для этого достаточно заметить, что если $u \cap v \le x \le x' \le v$, то $[x, x']$ и $[u \cup x, u \cup x']$ транспонированы и двойственно этому. В самом деле

$$x' \cup (x \cup u) = (x' \cup x) \cup u = x' \cup u, \text{ в то же время}$$
$$x' \cap (u \cup x) = (x' \cap u) \cup x = [(x' \cap v) \cap u] \cup x =$$
$$= [x' \cap (v \cap u)] \cup x = (v \cap u) \cup x = x.$$

Следствие. *В любой дедекиндовой структуре проективные интервалы изоморфны.*

Теорема 7. *Структура, порожденная* $[u \cap v, u] = X$ *и* $[u \cap v, v] = Y$, *является кардинальным произведением XY.*

Доказательство. Пусть (x, y) обозначает $x \cup y$ для любых $x \in X$, $y \in Y$. Тогда в силу L2—1.3

$$(x, y) \cup (x', y') = x \cup y \cup x' \cup y' =$$
$$= (x \cup x') \cup (y \cup y') = (x \cup x', y \cup y').$$

Далее, в силу теоремы 6 и L5 соответственно мы имеем

$$(x, y) = x \cup y = [(v \cup x) \cap u] \cup y = (v \cup x) \cap (u \cup y).$$

Следовательно, в силу L2—L3 $(x, y) \cap (x', y') = [(v \cup x) \cap \cap (v \cap x')] \cap [(u \cup y) \cap (u \cup y)]$. В силу теоремы 6 $(v \cup x) \cap \cap (v \cup x') = v \cup (x \cap x')$ и $(u \cup y) \cap (u \cup y') = u \cup (y \cap y')$.

Следовательно,

$$(x, y) \cap (x', y') = [v \cup (x \cap x')] \cap [u \cup (y \cap y')] = (x \cap x', y \cap y').$$

Этим завершается доказательство. По индукции мы получаем

Следствие. *Если* $(x_1 \cup \ldots \cup x_k) \cap x_{k+1} = a$ *для* $k = 1, \ldots \ldots, n - 1$, *то подструктура, порожденная интервалами* $X_k = [a, x_k]$, *есть* $X_1X_2 \ldots X_n$. Но определение произведения $X_1X_2 \ldots X_n$ симметрично относительно нижних индексов. Следовательно, условие предыдущего следствия инвариантно при всех перестановках нижних, индексов, и мы можем ввести

Определение. *Если выполняются предположения предыдущего следствия, то элементы* $x_1, \ldots, x_n$ *называются независимыми над а.*

Понятие независимости над 0 содержит в себе как частные случаи обычные понятия «неперекрываемости» в теории множеств и линейной независимости подпространств линейного пространства. «Независимость под *а*» может быть, очевидно, определена двойственным образом.

6.6.6. Оценки

В общем случае под *оценкой* на структуре L понимают вещественную функцию $v[x]$, определенную на L, которая удовлетворяет условию

VI. $\quad v[x] + v[y] = v[x \cap y] + v[x \cup y].$

Оценка называется *изотопной* тогда и только тогда, если

V2. $\quad$ Из $x \geq y$ следует $v[x] \geq v[y]$,

и *положительной* тогда и только тогда, если из $x > y$ следует $v[x] > v[y]$. Она называется *ограниченной вариацией* тогда и только тогда, если для некоторого конечного K и для всех цепей

$$\gamma: \qquad x_0 < x_1 < \ldots < x_n, \quad \sum_{i=1}^{n} |v[x_i] - v[x_{i-1}]| < K.$$

Вероятности и меры множеств суть оценки; оценками являются размерность в проективной геометрии, произвольная вещественная функция на цепи. Предыдущее определение содержит в себе как частные случаи обычные определения (Сакс) ограниченной вариации как для обычных вещественных функций, так и для функций множеств. Далее, справедлива

Лемма 5. *Вещественный функционал* $v[x]$ *на структуре с относительными дополнениями есть оценка при условии, что*

VI*. $\quad v[x \cup y] = v[x] + v[y]$ всякий раз, когда $x \cap y = 0$.

Доказательство. Пусть для произвольных x, y элемент t является относительным дополнением $x \cap y$ в $[0, y]$. По определению, $(x \cap y) \cap t = 0$ и $(x \cap y) \cup t = y$; следовательно, $v[y] = v[x \cap y] + v[t]$. Кроме того, поскольку $t \leq y$, имеем $x \cap t =$

$= x \cap (y \cap t) = (x \cap y) \cap t = 0$, в то время как $x \cup t =$
$= [x \cup (x \cap y)] \cup t = x \cup [(x \cap y) \cup t] = x \cup y$; следовательно,
$v[x \cup y] = v[x] + v[t]$. Вычитая из второго равенства первое, имеем
$v[x \cup y] - v[y] = v[x] + v[t] - v[x \cap y] - v[t]$, т. е. соотношение VI.

Определим теперь все оценки $v[x]$ на произвольной дедекиндовой структуре конечной длины.

Если для интервала $[a, b]$ структуры L мы определим $v[a, b] = v[b] - v[a]$, то в силу VI, транспонированные интервалы и, следовательно, *проективные интервалы имеют одно и то оке значение.* Но отношение проективности между (простыми) интервалами рефлексивно, симметрично и транзитивно, т. е. является отношением эквивалентрости. Следовательно, каждая оценка относит единственное значение λ_p каждому классу проективных простых интервалов.

Кроме того, если $\gamma : 0 = x_0 < x_1 < ... < x_n = x$ произвольная цепь, соединяющая 0 и x, то $v[x] = v[0] + \sum v[x_{i-1}, x_i]$. Следовательно, если $p[x, \gamma]$ обозначает, сколько раз встречается в γ простой интервал, проективный p, то мы имеем

$$v[x] = v[0] + \sum \lambda_p \, p[x, \gamma]. \tag{2}$$

Основываясь на рис. 3, *б* из п.6.6 2, покажем теперь, что $p[x, \gamma]$ *одно и то же для всех* γ, *соединяющих* 0 *и* x. В самом деле, пусть $P(m)$ есть утверждение, что если одна максимальная цепь, связывающая 0 и x, содержит m простых интервалов, проективных с p, то и каждая такая цепь содержит m таких интервалов. Предполагая справедливым $P(m - 1)$, мы, так же как и в п.6.6.2, доказываем $P(m)$. Единственным отличием является замечание, что так как $[y_0, y_1]$ и $[y_1, u]$ суть транспозиции $[x_1, u]$ и $[x_0, x_1]$ соответственно, то $p[x, \gamma]$ не меняется, когда мы переходим от x_0, x_1, γ'' к y_0, y_1, γ''.

Таким образом, мы определяем $p[x]$ как общее значение всех $p[x, \gamma]$. Из теоремы 6 мы непосредственно выводим

$$p[x \cup y] - p[x] = p[y] - p[x \cap y], \tag{2'}$$

в силу чего $p[x]$ является оценкой. Очевидна далее

Лемма 6. *Любая линейная комбинация* $\lambda_0 + \sum \lambda_k \upsilon_k[x]$ *оценок сама является оценкой.*

Отсюда следует, что при любом выборе коэффициентов λ_p (2) представляет собой оценку в L. Нами доказана

Теорема 8. *Различные оценки на дедекиндовой структуре L конечной длины взаимно однозначно характеризуются заданием* $\upsilon[0]$ *и значений* λ_p, *относимых классам проективных простых интервалов в L:*

$$\upsilon[x] = \upsilon[0] + \sum \lambda_p p[x], \tag{2''}$$

где p [x] есть число простых интервалов, проективных p в произвольной максимальной цепи, соединяющей 0 *с x.*

6.6.7. Метрические структуры

Структура *L* с изотонной оценкой называется *квазиметрической;* если оценка положительна, то *L* называется *метрической структурой* Рассуждение, непосредственно предшествующее теореме 3, показывает в действительности, что *любая метрическая структура является дедекиндовой.* Мы повторим это рассуждение. Если $x \leq z$, то из VI следует

$$v[x \cup (y \cap z)] - v[(x \cup y) \cap z] = v[x] + v[y \cap z] - v[x \cap y] - v[x \cup y] - v[z] + v[y \cup z] = -v[y] + v[y] = 0$$ **(применяя V1 четыре раза).**

Это делает невозможным отношение $x \cup (y \cap z) < (x \cup y) \cap z$, ввиду чего (в силу одностороннего модулярного закона) имеет место L5.

Более обще, мы определим в произвольной квазиметрической структуре

$$\partial(x, y) = v[x \cup y] - v[x \cap y] \tag{3}$$

как *расстояние* от *x* до *y*. Мы применим теперь метрические концепции, принадлежащие в основном Фреше.

Лемма 7. *Преобразования* $x \to a \cup x$ *и* $x \to a \cap x$ *являются сжатиями; в самом деле,*

$$\partial(a \cup x, a \cup y) + \partial(a \cap x, a \cap y) \leqslant \partial(x, y). \tag{4}$$

Доказательство. По определению левая часть соотношения (4) есть

$$v[a \cup x \cup y] - v[(a \cup x) \cap (a \cup y)] + v[(a \cap x) \cup (a \cap y)] - v[a \cap x \cap y].$$

В силу одностороннего дистрибутивного закона это есть самое большее

$$v[a \cup x \cup y] - v[a \cup (x \cap y)] + v[a \cap (x \cup y)] - v[a \cap x \cap y].$$

Переставляя местами средние члены и используя дважды VI, мы получаем

$$v[a] + v[x \cup y] - v[a] - v[x \cap y] = \partial(x, y).$$

Теорема 9. *Всякая квазиметрическая структура является квазиметрическим пространством, в котором объединения и пересечения равномерно непрерывны. Отношение ∂(x, y)=*0 *есть отношение* *конгруентности, отображающее L изометрически и структурно-гомоморфно на метрическую структуру.*

Пояснение. Квазиметрическое пространство есть пространство, в котором

$$\partial(x, x)=0,\ \partial(x, y)\geqslant 0,\ \partial(x, y)=\partial(y, x) \tag{5}$$

и

$$\partial(x, y)+\partial(y, z)\geqslant\partial(x, z)\ \textbf{(неравенство треугольника).} \tag{6}$$

Доказательство. Условия (5) очевидны. Далее,

$$\partial(x, y)+\partial(y, z)=\partial(x\bigcup y, y)+\partial(y, x\bigcap y)+\partial(y\bigcup z, y)+\partial(y, y\bigcup z)\geqslant$$

$$\geqslant\partial(x\bigcup y\bigcup z, y\bigcup z)+\partial(y\bigcup z, y)+\partial(y, x\bigcap y)+\partial(x\bigcap y, x\bigcap y\bigcap z),$$

так как в силу (4) $\partial(x\bigcup y\bigcup z,\ y\bigcup z)\leqslant\partial(x\bigcup y,\ y)$ и $\partial(x\bigcap y\bigcap z, y\bigcap z)\leqslant\partial(x\bigcap y, y)$. Но последняя сумма есть

$$\partial(x\bigcup y\bigcup z, x\bigcap y\bigcap z)\geqslant\partial(x\bigcup y, x\bigcap y)=\partial(x, y),$$

чем доказано (6). Докажем теперь равномерную непрерывность объединений в отчетливой форме

$$\partial(a\bigcup b, c\bigcup d)\leqslant\partial(a\bigcup b, c\bigcup b)+\partial(c\bigcup b, c\bigcup d)\ \text{в силу} \tag{5}$$

$$\partial(a, c)+\partial(b, d)\ \text{в силу (4) и L2.} \tag{7}$$

Из двойственных соображений получаем то же самое для пересечений. Можно показать, что

$$\partial(a\bigcup b, c\bigcup d)+\partial(a\bigcap b, c\bigcap d)\leqslant\partial(a, c)+\partial(b, d). \tag{7'}$$

Наконец, соотношения (5) показывают, что отношение $\partial(x, y)=0$ рефлексивно и симметрично, в то же время (6) и $\partial(x, z)\geq 0$ показывают, что это отношение транзитивно; следовательно, это есть отношение эквивалентности. В силу (7') оно обладает свойством замены для объединений и пересечений; следовательно, оно является отношением конгруентности, по модулю которого *L* есть структурно-гомоморфный образ *L**. В силу (6) конгруентные пары элементов находятся на равных расстояниях от других элементов; следовательно, соответствие $L\to L^*$ изометрично и (5) —(6) имеют место в L^*. Но в структуре L^* из $x\neq y$ следует по определению $\partial(x, y)>0$; следовательно, L^* есть метрическое пространство.

Теорема 10. *Пусть L—произвольная дедекиндова структура конечной длины. Отношения конгруентности в L соответствуют взаимно однозначно множествам классов проективных простых факторов, которые они аннулируют. Следовательно, отношения конгруентности в L образуют булеву алгебру.*

Доказательство. Пусть θ—какое-нибудь отношение конгруентности на *L*. Если $x\equiv x\bigcup y \pmod\theta$, то $x\cap y\equiv(x\bigcup y)\cap y\equiv y \pmod\theta$ и двойственно этому; следовательно, если θ аннулирует (простой) фактор, то оно аннулирует все проективные ему (простые) факторы (мы используем тот факт, что θ транзитивно). Обратно, пусть $S = S$ (θ)—какое-нибудь множество классов проективных простых факторов; определим υ_S *[x]* как сумму чисел *p [x]* для *p, не* принадлежащих *S*. В силу теоремы 8 это есть оценка; очевидно, что она изотонна. В силу теоремы 9 эта оценка определяет отношение конгруентности θ(*S*). Кроме того, $x \equiv y$ (mod θ *(S))* тогда и только тогда, если $x \cup y \equiv x \cap y$ (mod θ(*S*)), что случается, тогда и только тогда, если каждый простой фактор в любой максимальной цепи, соединяющей $x \cap y$ с $x \cup y$, принадлежит *S*. Следовательно, θ определяется множеством *S* (θ) и соответствие θ → *S* (θ) взаимно однозначно. Очевидно, однако, что оно изотонно; следовательно, это соответствие есть изоморфизм.

Следствие. *Дедекиндова структура конечной длины является «простой» (т. е. не имеет собственных отношений конгруентности) тогда и только тогда, если все ее простые факторы проективны.*

Рассмотрим, далее, свободную дедекиндову структуру *L* с тремя образующими, изображенную в п.6.6.3 на рис. 4. Непосредственный подсчет простых факторов показывает, что *L* может быть шестью способами гомоморфно отображена на дистрибутивную структуру 2 и одним способом на прямую *М* с тремя точками, изображенную на рис. 1,*в*. Кроме того, *М* «просто». Отсюда следует, что если *x, y, z* какие-нибудь три элемента дедекиндовой структуры и если любые два элемента, отличные друг от друга в *М,* равны в подструктуре, порожденной элементами *x, y,* z, то *x, y, z* порождают дистрибутивную подструктуру. В результате доказана

Теорема 11. *Элементы x, y, z дедекиндовой структуры порождают дистрибутивную подструктуру, если имеет место одно из равенств* $x \cap (y \cup z) = (x \cap y) \cup (x \cap z)$ *или* $x \cup (y \cap z) = (x \cup y) \cap (x \cup z)$.

Отсюда следует, что каждое из этих равенств является в действительности двойственным себе и симметричным условием относительно трех переменных *x, y,* z.

6.6.8. Идеалы, нейтральные элементы

Рассмотрим следующий результат, принадлежащий Дилуорсу.

Теорема 12. *Идеалы произвольной дедекиндовой структуры сами образуют дедекиндову структуру по отношению теоретико-множественного включения.*

Замечание. Для случая структур конечной длины, где каждый идеал является главным, этот результат тривиален.

Доказательство. Предположим, что $X \leq Z$, Y идеалы; в силу одностороннего модулярного закона нам нужно лишь показать, что каждое $t \in (X \cup Y) \cap Z$ лежит в $X \cup (Y \cap Z)$. Но $t \in (X \cup Y) \cap Z$ означает для идеалов в общем случае, что $t \leq (x \cup y) \cap z$ для некоторых $x \in X$, $y \in Y$, $z \in Z$; поскольку $X \leq Z$, элемент $z_1 = x \cup z$ также лежит в Z и, очевидно, $t \leq (x \cup y) \cap z_1$, где $x \leq z_1$. В силу L5 отсюда следует, что $t \leq x \cup (y \cap z_1)$; следовательно, $t \in X \cup (Y \cap Z)$.

Поставим теперь вопрос о том, какие идеалы в дедекиндовой структуре являются множествами элементов, конгруентных 0 при подходящем отношении конгруентности.

Теорема 13. *Элемент а дедекиндовой структуры нейтрален тогда и только тогда, если соответствие* $x \to x \cap a$ *или соответствие* $x \to x \cup a$ *является эндоморфизмом структуры L.*

Доказательство. Если соответствие $x \to x \cap a$ есть эндоморфизм, то $(x \cup y) \cap a = (x \cap a) \cup (y \cap a)$, а потому в силу теоремы 11 каждая тройка элементов *a, x, y* порождает дистрибутивную подструктуру. Следовательно, по определению, элемент *a* нейтрален. Доказательство завершается применением принципа двойственности.

Отсюда следует, что если элемент *a* нейтралеа, то идеал [0, *a*] состоит из всех элементов, отображающихся в 0 при подходящем структурном гомоморфизме. Однако, как это можно видеть на рис. 6,*б,* существуют также главные идеалы [0, *a*], состоящие из всех элементов, отображающихся в 0 при структурном гомоморфизме, у которых элемент *a* не нейтрален.

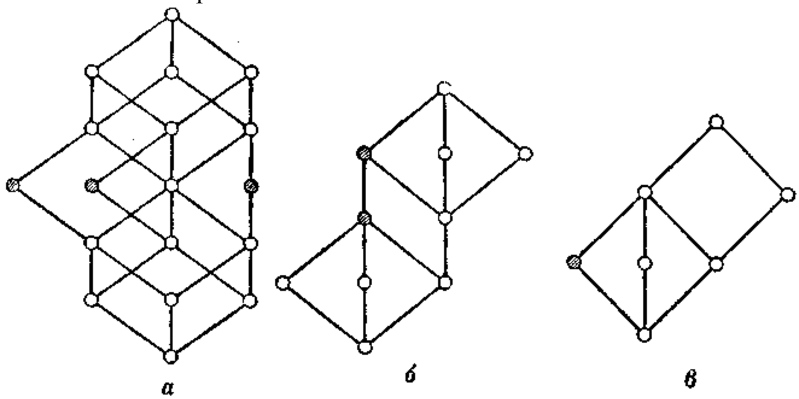


Рис. 6.

Если элемент a нейтрален, то легко показать, что a имеет самое большее одно дополнение. Обратное, однако, неверно; так, например, Маршалл Холл отметил, что заштрихованный элемент на рис. 6,*в* имеет одно единственное дополнение, но не является нейтральным.

6.6.9. Метрическая топология в сопоставлении с топологией упорядоченности

Пусть M—произвольная метрическая структура; мы можем пополнить M как метрическое пространство подобно тому, как это было описано в «Предварительных сведениях из топологии». Далее, если $\{x_n\}$ и $\{y_n\}$ последовательности Кошн, то в силу (7) $\{x_n \cup y_n\}$ и $\{x_n \cap y_n\}$ также будут последовательностями Коши; мы обозначим их, соответственно, как $\{x_n\} \cup \{y_n\}$ и $\{x_n\} \cap \{y_n\}$. Тогда путем перехода к пределу могут быть доказаны LI —L4; кроме того, $\{v[x_n]\}$ будет сходиться к пределу, который мы можем определить как $v[\{x_n\}]$ — тогда доказывается, что это есть «положительная оценка» в смысле п.6.6.6. Тем самым для M однозначным образом определяется (метрически) *полная метрическая структура.* Следовательно, из теоремы 9 получается непосредственно

Теорема 14. *Любая метрическая структура М имеет одну-единственную метрически полную оболочку, в которой она (метрически) плотна.*

Вообще говоря, метрическое пополнение M не изоморфно пополнению M с помощью сечений. Однако оно является обычно полной структурой.

Теорема 15. *Любая метрически полная метрическая структура М является условно полной и удовлетворяет условию*

из $x_n \uparrow x$ *следует* $v[x_n] \uparrow v[x]$ *и двойственно этому;* (8)

обратно, σ-полная метрическая структура, удовлетворяющая (8), метрически полна. В метрически полной метрической структуре метрическая сходимость и звездная сходимость эквивалентны.

Доказательство. Пусть M полная метрическая структура и пусть S какое-нибудь ограниченное подмножество в M. Рассмотрим объединения $\sup X$ конечных подмножеств X из S; так как S ограничено и $v[x]$ изотонно, то множество вещественных чисел $v[\sup X]$ будет ограниченным; следовательно, оно имеет наименьшую верхнюю грань v_s. Поэтому мы можем найти такие $X_1, X_2, X_3, \ldots$, что $v[X_n] \geq v_s - 2^{-n}$. Обозначая тогда через U (конечное) теоретико-множественное объединение X_m и X_n, мы имеем в силу (6)

$$\partial(\sup X_m, \sup X_n) \leqslant v[\sup U] - v[\sup X_m] + v[\sup U] - v[\sup X_n] \leqslant 2^{-m} + 2^{-n} \text{ по построению.}$$

Следовательно, в силу метрической полноты $\sup X_n$ сходятся метрически к некоторому $s \in M$. Пусть теперь дано $x \in S$; в силу леммы из п.6.6.7

$$v[x \cup s] - v[s] = \operatorname*{Lim}_{n\to\infty} \{v[x \cup \sup X_n] - v[\sup X_n]\} \leqslant 2^{-n}$$

для всех n. Следовательно, $x \cup s = s$ и s является верхней гранью для S. В то же время, если u произвольная верхняя грань для S, то $u \geq \sup X_n$ для всех n, а потому в силу непрерывности $u \cap s = s$. Следовательно, s есть наименьшая верхняя грань для S. Существование $\inf S$ следует из двойственных соображений; следовательно, M условно полно.

Для доказательства (8) заметим, что если $x_n \uparrow x$, то $v[x_n]\uparrow$ и $v[x_n] \leq v[x]$ для всех n; следовательно, $v[x_n]\uparrow c$ для некоторого вещественного числа c. Отсюда следует, что $\partial(x_m, x_n) = |v[x_m] - v[x_n]| \to 0$ при $m, n \to \infty$; следовательно, в силу метрической полноты $\partial(x_m, y) \to 0$ при $m \to \infty$ для некоторого y; очевидно, что $v[x_m] \uparrow v[y]$; кроме того, $y \cap x_n = \operatorname{Lim} x_m \cap x_n = \operatorname{Lim} x_n = x_n$ (используя метрические пределы); следовательно, y является верхней гранью для x_n. Но по определению, $x = \sup\{x_n\}$; следовательно, $x \leq y$ и $v[x] \leq v[y]$. Но мы уже показали, что $v[x_n] \uparrow v[y]$ и $v[x_n] \leq v[x]$ для всех n; следовательно, $v[x_n] \uparrow v[x]$ и условие (8) доказано.

Обратно, пусть M σ-полная метрическая структура, о которой предполагается лишь, что она удовлетворяет условию (8). Из всякой подпоследовательности последовательности Коши можно выделить подподпоследовательность $\{x_n\}$, удовлетворяющую условию $\partial(x_n, x_{n+1}) < 2^{-n}$ для всех $n = 1, 2, 3, \ldots$ Мы покажем, что а) для некоторого y $\partial(x_n, y) \to 0$ — этим будет доказана метрическая полнота —и б) $\{x_n\}$ сходится по упорядоченности к y —этим будет доказано, что из метрической сходимости вытекает звездная сходимость.

В самом деле, образуем $y_{n,r} = x_n \cup x_{n+1} \cup \ldots \cup x_{n+r}$. Для фиксированного n $y_{n,r}\uparrow$. В силу σ-полноты $y_{n,r} \uparrow y_n$, где, согласно (8), $\partial(y_{n,r}, y_n) \to 0$. Следовательно,

$$\partial(x_n, y_n) \leqslant \sum_{r=0}^{\infty} \partial(y_{n,r}, y_{n,r+1}) = \sum_{r=0}^{\infty} \partial(y_{n,r} \cup x_{n+r}, y_{n,r} \cup x_{n+r+1}) <$$

$$< \sum_{r=0}^{\infty} 2^{-n-r} = 2 \cdot 2^{-n}$$ в силу (4) из р.6.6.7.

Двойственно этому определим $z_{n,r} = x_n \cap x_{n+1} \cap \ldots \cap x_{n+r}$; тогда $z_{n,r} \downarrow z_n$, где $\partial(x_n, z_n) \leq 2 \cdot 2^{-n}$. Кроме того, $y_n \geq x_n \geq z_n$; поскольку $y_n \geq y_{n,r+1} \geq y_{n+1,r}$ для всех r, то $y_n \geq y_{n+1}$; двойственно этому $z_n \leq z_{n+1}$. Снова в силу σ-полноты имеем $y_n \downarrow y$ и $z_n \uparrow z$, где

$$\partial(y,z)=\operatorname*{Lim}_{n\to\infty}\partial(y_n,z_n)\leqslant\operatorname*{Lim}_{n\to\infty}\{\partial(y_n,x_n)+\partial(x_n,z_n)\}=0,$$

ибо $\partial(y_n, x_n)+\partial(x_n, z_n) < 4 \cdot 2^{-n}$ для всех n. Следовательно, $y = z$ и «б» $\{x_n\}$ сходится по упорядоченности к y. Кроме того, мы имеем «а», так как

$$\partial(x_n,y)=v[x_n\bigcup y]-v[x_n\bigcap y]\leqslant v[y_n]-v[z_n]<4\cdot 2^{-n}.$$

Чтобы закончить доказательство теоремы 15, мы должны показать, что из звездной сходимости вытекает метрическая сходимость. Но так как последовательность сходится метрически, если все ее подпоследовательности содержат метрически сходящиеся подподпоследовательности, то нам нужно лишь показать, что из сходимости по упорядоченности вытекает метрическая сходимость. Это следует непосредственно из (8), ибо если $u_n{\downarrow}x$ и $w_n{\uparrow}x$, где $u_n\geq x_n\geq w_n$, то в силу (8)

$$\partial(x_n,x)=v[x_n\bigcup x]-v[x_n\bigcap x]\leqslant v[u_n]-v[w_n]\downarrow 0.$$

Доказательство теоремы закончено. Так как в любой структуре с 0 и I из условной полноты вытекает σ-полнота, имеет место

Следствие. *В метрической структуре с* 0 *и I, удовлетворяющей условию* (8), *метрическая полнота, полнота (по упорядоченности) и* σ-*полнота эквивалентны.*

Теорема 16. *Всякая полная метрическая структура является топологической структурой.*

Доказательство. Предположим, что $x_\alpha{\uparrow}x$, и пусть $s = \sup\upsilon[x_\alpha]$. Мы можем выбрать $t_n = x_{\alpha(n)}$ $(n = 1, 2, 3, ...)$ с $\alpha(n) \leq\alpha(n+ 1)$ и $\partial(t_n, a) <s+2^{-n}$. Так как $\alpha(n) \leq\alpha(n+ 1)$, то $\partial(t_n, t_{n+1}) < 2^{-n}$; следовательно, $t_n\to t$ метрически для некоторого t, где $v[t] = s$. Мы покажем теперь, что $t\geq x_\alpha$ для всех α. В противном случае $v[t \cup x_\alpha]- v[t] > 2^{-n}$ для некоторого n; следовательно, $v[t_m \cup x_\alpha]-s[t] > 2^{-n}$ для некоторых m, n; следовательно, $v[x_\beta] >s+2^{-n}$ для любого общего последующего элементов $x_{\alpha(m)} = t_m$ и x_α; это противоречит определению s. Отсюда следует, что $t\geq x$; но $\upsilon[t] = s\leq\upsilon[x]$; следовательно, $t = x$ и $\partial(x_\alpha, x)\downarrow 0$. В силу (4) из 6.6.7 $\partial(a \cap x_\alpha, a \cap x)\downarrow 0$ для всех a; отсюда легко показать, что элемент $a \cap x$, являющийся верхней гранью для изотонного направленного множества $\{a \cap x_\alpha\}$, будет наименьшей верхней гранью и, следовательно, пределом этого множества. Далее мы используем двойственность.

6.6.10. Жорданово разложение

Жорданово разложение функций ограниченной вариации на монотонные составляющие было обобщено Риссом на аддитивные функции множеств. Оно может быть обобщено на оценки в структурах. Рассмотрим оценки $v[x]$ на структуре L, удовлетворяющие условию $v[a] = 0$ для некоторого фиксированного a. Они образуют векторное пространство, поскольку любая линейная комбинация оценок есть оценка. Мы определим

$v \geq v_1$ означает, что $v[x] - v_1[x]$ изотонно. (9}

Тогда справедливость P1 и P3 очевидна, а P2 следует из того, что мы предполагаем $v[a] = 0$. Очевидно, далее, что

из $v \geq v_1$ вытекает $v + v_2 \geq v_1 + v_2$ для всех v_2. (10)

Следовательно, оценки на L с $v[a] = 0$ образуют *частично упорядоченное векторное пространство*.

Теперь мы отождествим оценки с *функциями интервалов*. Если $x \leq y$, то мы определим $v[x, y] = v[y] - v[x]$. Тогда V1 утверждает, что перспективные и *проективные интервалы имеют одно и то оке значение.*

С произвольной цепью γ: $x = x_0 < x_1 < \dots < x_n = y$ в $[x, y]$, мы свяжем *положительную вариацию*

$$v^+[x, y; \gamma] = \sum_{i=1}^{n} \sup \{v[x_{i-1}, x_i], 0\} \qquad (11).$$

оценки $v[x]$ на γ определяемую как сумма положительных приращений $v[x]$ вдоль γ. Мы определим

$$v^+[x, y] = \sup_\gamma v^+[x, y; \gamma] \qquad (12).$$

как положительную вариацию оценки v на $[x, y]$. Очевидно, что если γ'' произвольное уплотнение цепи γ, то $v^+[x, y; \gamma''] \geq v^+[x, y; \gamma]$. Далее, если γ' произвольная цепь $x = x'_0 < < x'_1 < \dots < x'_m = y$, разбивающая $[x, y]$, то в силу следствия .из теоремы 5 γ будет иметь уплотнение γ'' подынтервалы которого проективны соответствующим интервалам некоторого уплотнения γ''' цени; γ', следовательно, такое уплотнение, что $v^+[x, y; \gamma''] = v^+[x, y; \gamma'''] \geqslant v^+[x, y; \gamma']$. Мы выводим отсюда, что

$$v^+[x, y] = \sup_{\gamma''} v^+[x, y; \gamma''] \text{ для уплотнений } \gamma'' \text{ цепи } \gamma. \qquad (13)$$

В силу определения $v[x]$ имеет *ограниченную вариацию* тогда и только тогда, если каждое $v^+[x, y]$ и двойственное ему $v^-[x, y]$ конечны. Мы назовем их *положительной и отрицательной вариациями* оценки $v[x]$ на $[x, y]$.

В этом случае проективные интервалы имеют равные $\upsilon^+[x, y]$, ибо они имеют соответствующие цепи и (как отмечено выше) равные $\upsilon^+[x, y; \gamma]$ для соответствующих цепей. Отсюда следует, что

$$\upsilon^+[x] = \upsilon^+[a, a \cup x] - \upsilon^+[a \cap x, a] \qquad (14)$$

является оценкой в L. Далее, $\upsilon^+ \geq 0$, поскольку подстановка x_1 вместо x расширяет $[a, a \cup x]$ и сужает $[a \cap x, a]$; в силу построения $\upsilon^+ \geq \upsilon$. Обратно, если $\upsilon_1 \geq \upsilon$, 0, то для каждого γ $\upsilon_1[x, y] \geq \upsilon^+[x, y; \gamma]$; следовательно, $\upsilon_1[x, y] \geq \upsilon^+[x, y]$, а потому $\upsilon_1 \geq \upsilon^+$. Таким образом, имеет место

Теорема 17. *Если $\upsilon[x]$ имеет ограниченную вариацию, то* $\upsilon \cup 0$ *существует и совпадает с υ^+, определенным формулами* (11) —(14).

Отсюда следует, что оценки ограниченной вариации с $\upsilon[a]=0$ образуют *векторную структуру*. Из этого можно вывести много других следствий.

6.7. Приложения к алгебре

6.7.1. Нормальные и перестановочные отношения конгруентности

Квазигруппа есть система с бинарной операцией умножения, такой, что любые два из трех членов в выражении $ab = c$ однозначно определяют третий. Это означает, что умножение однозначно и что $ax = c$ и $yb = c$ имеют и притом единственные решения. *Лупа* есть квазигруппа с двусторонней «единицей» 1, удовлетворяющей соотношениям $1x = x1 = x$ для всех x. Общую теорию квазигрупп и луп можно найти в других работах.

Определение. *Отношение конгруентности* θ *на лупе* G *называется нормальным, если*

$$ux \equiv x(\theta) \text{ или } xu \equiv x(\theta) \text{ влечет за собой, } u \equiv 1(\theta). \qquad (1)$$

Теорема 1. *Все отношения конгруентности на лупе G нормальны, если выполняется одно из следующих условий:* а) *G есть группа,* б) $(ux)x^{-1} = u = x^{-1}(xu)$ *для всех* $u, x \in G$ *и некоторого* x^{-1}, в) *G конечна.*

Доказательство. Случай «б» содержит в себе случай «а»; поэтому мы начинаем с него. Но, используя «б», мы получаем $u \equiv (ux)x^{-1} \equiv (1x)x^{-1} \equiv x\,x^{-1} = 1 \pmod\theta$ и симметрично этому. Предположим, что имеет место «в». Во всякой лупе из $x \equiv y(\theta)$ вытекает $xb \equiv yb(\theta)$. Поскольку $x \to xb$ переводит различные элементы в различные элементы, число элементов в классе вычетов, содержащем xb, не меньше, чем число элементов в классе вычетов, содержащем x. Но для любых x, y уравнение $xb = y$ имеет решение; следовательно, все классы вычетов состоят из одного и того же числа элементов. Если лупа G конечна, то отсюда вытекает (1); в противном случае соответствие $t \to tx$ вводило бы новые отношения конгруентности.

Определение. *Два отношения конгруентности* θ *и* θ' *называются перестановочными тогда и только тогда, если*

$$\theta\theta' = \theta'\theta \qquad (2)$$

в общем смысле умножения отношений. Иными словами, записывая $a \equiv x \pmod{\theta}$ *в виде* $a\theta x$, *мы требуем, чтобы из* $a\theta x$ *и* $x\theta' b$ *для некоторого* x *следовало* $a\theta' y$ *и* $y\theta b$ *для некоторого* y, *и обратно*.

Теорема 2. *Нормальное отношение конгруентности на лупе перестановочно со всяким отношением конгруентности.*

Доказательство. Предположим, что $a\theta x$ и $x\theta' b$; положим $a = ux$, $b = xv$ и $y = u(xv) = ub$. Поскольку $x\theta' b$, имеем $y = ub \equiv ux = a$ (θ'). Далее, так как $ux = a \equiv x = 1x$ (θ) и θ нормально, $u \equiv 1(\theta)$, а потому $y = ub \equiv 1b = b(\theta)$. Этим доказано, что $\theta\theta' \leq \theta'\theta$; чтобы доказать обратное включение, мы просто изменяем порядок умножения.

Перенесение теоремы 2 на квазигруппы было сделано Киокемейстером.

Следствие. *В любой группе* (с *операторами или без операторов), а также в любой лупе, удовлетворяющей* «б» или «в», *все отношения конгруентности перестановочны.*

Действительно, если мы присоединяем к алгебре новые операции, то структура отношений конгруентности становится меньше. Следовательно, если (1) имеет место в группе (или другой лупе) G, то это условие имеет место и в той же группе, как группе с операторами.

6.7.2. Прямые разложения

Установим два фундаментальных свойства алгебр с перестановочными отношениями конгруентности.

Теорема 3. *Отношения конгруентности на произвольной алгебре* с *перестановочными отношениями конгруентности образуют дедекиндову структуру, в которой* $\theta \cup \theta' = \theta\theta' = \theta'\theta$.

Доказательство. Так как $a \equiv b(\theta \cup \theta')$ означает, что для некоторой конечной цепи $a = x_0\theta x_1\theta' x_2\theta x_3\theta' \ldots x_n = b$, то очевидно, что $\theta \cup \theta'$ есть объединение конечных произведений $\theta\theta'\theta\theta' \ldots$ Но в силу (1) и соотношений $\theta^2 = \theta$, $\theta'^2 = \theta'$ это есть $\theta\theta' = \theta'\theta$; следовательно, $\theta \cup \theta' = \theta\theta' = \theta'\theta$.

Используя односторонний модулярный закон, остается доказать, что если $\theta_1 \geq \theta_2$, то $\theta_1 \cap (\theta\theta_2) \leq (\theta_1 \cap \theta)\theta_2$. Предположим, что $a \equiv b(\theta_1)$ и $a \equiv b(\theta\theta_2)$, т. е. $a \equiv b$ $(\theta_1 \cap (\theta\,\theta_2))$. Тогда для некоторого $xa \equiv x(\theta)$ и $1 \equiv b(\theta_2)$. Но $\theta_1 \geq \theta_2$; следовательно, $x \equiv b(\theta_1)$. Так как также $a \equiv b(\theta_1)$ и θ_1 транзитивно, мы заключаем, что $x\ a(\theta_1)$. Следовательно, $x \equiv a(\theta \cap \theta_1)$. Так как $x \equiv b(\theta_2)$, мы получаем теперь непосредственно, что $a \equiv b((\theta_1 \cap \theta)\ \theta_2)$.

Лемма. *Пусть* θ_1, θ_2 *некоторые перестановочные отношения конгруентности на алгебре А такие, что* $\theta_1 \cap \theta_2 = 0$, $\theta_1 \cup \theta_2 = I$. *Тогда А изоморфно прямому произведению* $A_1 \times A_2$, *где* A_i *есть гомоморфный образ А по* mod θ_i (i=l, 2).

Доказательство. Пусть для элементов $x, y, ...$ из A и i=l, 2 символы x_i y_i ... обозначают классы вычетов A по mod θ_i, содержащие соответственно x, y, ... Очевидно тогда, что соответствие $x \rightarrow (x_1, x_2)$ является гомоморфизмом A на подалгебру S алгебры $A_1 \times A_2$. Так как $\theta_1 \cap \theta_2 = 0$, из $x_1 = y_1$ и $x_2 = y_2$ следует $(x_1, x_2) = (y_1, y_2)$, т. е. гомоморфизм является изоморфизмом. Так как $\theta_1\theta_2 = \theta_1 \cup \theta_2 = I$ для любых (x_1, x_2) и (y_1, y_2) существует такое (z_1, z_2), что $(x_1, x_2) \equiv (z_1, z_2)$ (θ_1) и $(z_1, z_2) \equiv (y_1, y_2)$ (θ_2), т. е. такое, что $x_1 = z_1$ и $z_2 = y_2$. Это означает, что для любых x_1 и y_2 (x_1, y_2) есть элемент из S; следовательно, $S = A_1 \times A_2$, чем и завершается доказательство.

Теорема 4. *Представления алгебры А в виде прямого произведения* $A = A_1 \times \ldots \times A_r$ *находятся во взаимно однозначном соответствии с множествами перестановочных отношений конгруентности* $\theta_1, ..., \theta_r$ *на А, удовлетворяющих условию*

$$\theta_1 \cap \ldots \cap \theta_r = 0 \ \text{и}\ (\theta_1 \cap \ldots \cap \theta_{i-1}) \cup \theta_i = I \quad [i = 2, \ldots, r]. \tag{3}$$

Замечание. Эти отношения конгруентности являются независимыми в двойственном смысле элементами, пересечение которых есть 0.

Доказательство. Предположим, что $A = A_1 \times \ldots \times A_r$; пусть $x \equiv y(\theta_i)$ означает, что $x = [x_1, ..., x_r]$ и $y = [y_1, ..., y_r]$ имеют одинаковую i-ю компоненту $x_i = y_i$. Тогда (3) является очевидным. Предположим, что, обратно, имеет место (3). Тогда в силу леммы $A = B_r \times A_r$, где B_r есть A по mod $\theta_1 \cap ... \cap \theta_r$. Индукцией по r получаем $B_r = A_1 \times ... \times A_{r-1}$, чем и завершается доказательство.

6.7.3. Теорема Жордана—Гельдера

Для получения теоремы Жордана — Гельдера в ее многочисленных формах нам требуется не только, чтобы отношения конгруентности на алгебре были перестановочны, но также чтобы алгебра содержала одноэлементную подалгебру.

Определение. *Пусть А — алгебра* с *выделенной одноэлементной подалгеброй* 1, *причем все отношения конгруентности* θ, θ', ... *на алгебре А перестановочны. Для любого* θ *мы определим* $S(\theta)$ *как* *подалгебру элементов* $x \equiv 1(\theta)$ *в А. Если* $\theta \geq \theta'$, *то фактор-символом* θ/θ' *мы будем обозначать гомоморфный образ* $S(\theta)$ *по* mod θ'.

Теорема 5. *При отмеченных выше предположениях относительно алгебры А имеем для всех* θ_1, θ_2

$$(\theta_1 \bigcup \theta_2) / \theta_2 \text{ и } \theta_1 / \theta_1 \bigcap \theta_2 \text{ изоморфны.} \qquad (4)$$

Замечание. Эта теорема обобщает то, что называют первой теоремой об изоморфизмах.

Доказательство. В силу второй теоремы об изоморфизмах мы можем отождествить элементы, конгруентные по mod $\theta_1 \cap \theta_2$; следовательно, не нарушая общности, мы можем предполагать, что $\theta_1 \cap \theta_2 = 0$. Далее, в (4) фигурируют лишь элементы, содержащиеся в $S(\theta_1 \cup \theta_2)$; в этой подалгебре $\theta_1 \cup \theta_2 = I$. Таким образом, мы свели доказательство к случаю, рассмотренному в лемме 6.7.2. В этом случае $S(\theta_1 \cup \theta_2)$ по mod θ_2 изоморфно A_2; $S(\theta_1)$ состоит из элементов интервала $[1, a_2]$, изоморфного A_2, чем и завершается доказательство.

Следствие 1. *Пусть А—произвольная алгебра с одноэлементной подалгеброй* 1 *и перестановочными отношениями конгруентности. Тогда проективные интервалы* $[\theta', \theta]$ *в (дедекиндовой) структуре отношений конгруентности на А определяют изоморфные фактор-символы.*

Доказательство вытекает непосредственно из определения проективности и теоремы 3.

Используем теперь доказательство теоремы 3 п.6.5 со специальной ссылкой на рис. 3, *б.* Усилим $P(m)$ до утверждения $Q(m)$: *если* $a = x_0 < x_1 < \ldots < x_m = b$ и $a = y_0 < y_2 < \ldots < y_m = b$ любые две максимальные цепи, то существует взаимно однозначное соответствие между $[x_{i-1}, x_i]$ и $[y_{i-1}, y_i]$ такое, что соответствующие интервалы проективны Поскольку $[x_1, u]$ и $[a, y_1]$, а также $[a, x_1]$ и $[y_1, u]$ проективны, мы можем доказать $Q(m)$, исходя из $Q(m— 1)$, в точности так же, как мы раньше доказывали $P(m)$, исходя из $P(m—1)$. В силу предыдущего следствия из $Q(m)$ вытекает следующая обобщенная теорема Жордана — Гельдера для *главных* рядов.

Теорема 6. *Пусть* $0 = \theta_0 < \theta_1 < \ldots < \theta_m = I$ *и* $0 < \theta'_0 < \theta'_1 < \ldots < \theta'_n = I$*—две произвольные конечные максимальные цепи отношений конгруентности на алгебре А с одноэлементной подалгеброй, все отношения конгруентности на которой перестановочны. Тогда* $m = n$ *и* θ_i / θ_{i-1} *попарно изоморфны с* $\theta'_j / \theta'_{j-1}$.

То же самое рассуждение приводит к теореме Жордана — Гельдера для *композиционных* рядов. В этом случае θ_{i-1} предполагается максимальным отношением конгруентности на подалгебре $S(\theta_i)$. Для того, чтобы доказательство прошло целиком, мы должны потребовать, чтобы отношения конгруентности на S (θ_i) были перестановочны, так же как и на самом A. Это имеет место для классов алгебр, рассмотренных в теореме 1, ибо эти классы замкнуты относительно перехода к подалгебрам и гомоморфным образам. Мы получаем

Следствие 2. *Если A группа* (с *операторами или без операторов) или лупа, удовлетворяющая одному из условий теоремы 1, то фактор-группы, встречающиеся в любых двух конечных композиционных рядах, попарно изоморфны.*

В связи с этим переходом от произвольных алгебр к лупам следует отметить, что в теореме 6 не утверждается, что S (θ) однозначно определяет θ (т. е. что отношение конгруентности определяется подалгеброй элементов, конгруентных 1) или что S ($\theta_1 \cup \theta_2$) является наименьшей верхней гранью для S (θ_1) и S (θ_2) в структуре подалгебр алгебры A. Эти факты справедливы, однако, в лупах, где S (θ) суть так называемые *нормальные подлупы;* следовательно, они справедливы также для групп.

Было сделано два рода обобщений теоремы Жордана — Гельдера на алгебры, не удовлетворяющие условиям, налагаемым на цепи. Первое обобщение принадлежит Шрейеру и Цасенхаузу и является следствием теоремы 6 п.6.5.

Следствие 3. *Пусть* $0 = \theta_0 < \theta_1 < \ldots < \theta_m = I$ *и* $0 = \theta'_0 < \theta'_1 < \ldots < \theta'_n = I$ *любые две конечные цепи отношений конгруентности на A. Тогда эти цепи можно уплотнить путем включения членов* $\theta_{i,j} = \theta_{i-1} \cup (\theta'_j \cap \theta_i)$ *и* $\theta'_{i,j} = \theta'_{j-1} \cup (\theta_i \cap \theta'_j)$ *так, что соответствующие факторы* $\theta'_{i,j} / \theta'_{i,j-1}$ и $\theta'_{i,j} / \theta'_{i-1,j}$ *будут изоморфны.*

В самом деле, мы можем доказать непосредственно, не используя индукцию, что как $\theta'_{i,j} / \theta'_{i,j-1}$, так и $\theta'_{i,j} / \theta'_{i-1,j}$ перспективны фактору $(\theta'_i \cup \theta'_{j-1}) \cap (\theta_{i-1} \cup \theta'_j) / \theta_{i-1} \cup \theta'_{j-1}$. Это доказательство приводит к теореме 6, как к следствию.

Наконец, используя трансфинитную индукцию, можно показать, что теорема 6 справедлива также для *возрастающих* макимальных трансфинитных последовательностей отношений конгруентности на алгебрах, имеющих одноэлементную подалгебру и перестановочные отношения конгруентности, но несправедлива для убывающих последовательностей.

6.7.4. Группы с операторами

Эндоморфизм группы или лупы G обычно называют *оператором. Группа с операторами* (или лупа с операторами) есть, таким образом, система (G, Ω), состоящая из группы (или лупы) G и множества Ω эндоморфизмов в (операторов на) G, которые рассматриваются как одинарные операции в алгебре (G, Ω). Ω-гомоморфизм (G,Ω) в (H, Ω) есть, таким образом, однозначное отображение α: $x \to \alpha(x)$ группы G в H такое, что

$$\alpha(x\omega) = [\alpha(x)]\omega \text{ для всех } \omega \in \Omega. \qquad (5)$$

Если α взаимно однозначно, то это есть Ω-*изоморфизм.*

Результаты 6.7.3 применимы к любой группе с операторами; они применимы также к любой лупе при условии, что число элементов конечно или $x^{-1}(xu) = (ux)x^{-1} = u$ для всех x, u. В самом деле, при этих предположениях, как было замечено в 6.7.1, все отношения Ω-конгруентности (будучи отношениями конгруентности) перестановочны между собой. Далее, поскольку $ax = a$ имеет 1 своим единственным решением и в силу (5) $1\omega = (11)\omega = (1\omega)(1\omega)$, то $1\omega = 1$ для всех ω. Следовательно, любая группа или лупа с операторами имеет одноэлементную подалгебру. Отсюда следует, что теорема Жордана — Гельдера применима также к группам с операторами, причем соответствующие факторы не только просто изоморфны, но даже Ω-изоморфны.

Используя понятие оператора, можно уточнить теорему Жордана — Гельдера и ее следствия для групп *без* операторов. В самом деле, каждое отношение конгруентнсоти на группе является также отношением конгруентности для всех операторов «внутренних автоморфизмов» $x \to a^{-1}xa$. Следовательно, изоморфизмы, о которых идет речь в теореме Жордана —Гельдера, сохраняются при всех внутренних автоморфизмах. Такие изоморфизмы α называются в теории групп «центральными изоморфизмами». Они удовлетворяют соотношению

$$a^{-1}\alpha(x)\,a = \alpha(a^{-1}xa) = \alpha(a)^{-1}\alpha(x)\,\alpha(a) \text{ для всех } a,\ x,$$

откуда $\alpha(a)\,a^{-1}$ перестановочно со всеми $\alpha(x)$, и $\alpha(a) = az$, где z лежит в центре.

Это может быть обобщено на лупы следующим образом. В силу свойства замены всякое отношение конгруентности на лупе G является также отношением конгруентности для одинарных операций $x \to a(xb)$ и $x \to (ax)b$. Если $b = a^{-1}$, то 1 есть одноэлементная подалгебра, хотя соответствие $x \to a(xa^{-1})$ не обязано быть изоморфизмом. Если G удовлетворяет предположениям теоремы 1, то все отношения конгруентности перестановочны. Следовательно, если мы определим «центральный изоморфизм» между двумя фактор-лупами одной и той же лупы с операторами как изоморфизм, сохраняющийся при всех отображениях $x \to a(xa^{-1})$ и $x \to (a^{-1}x)a$, то имеет место

Теорема 7. *В теореме Жордана — Гельдера для луп соответствующие факторы центрально изоморфны при условии, если существуют обратные элементы.*

6.7.5. Операторы изоморфизмов

Пусть x_1, x_2, x_3 три каких-нибудь отношения конгруентности на алгебре *A* с одноэлементной подалгеброй и перестановочными отношениями конгруентности. Сославшись на рис. 4 мы найдем, что

$$e_1 \bigcap e_2 = e_2 \bigcap e_3 = e_3 \bigcap e_1 = d, \quad e_1 \bigcup e_2 = e_3 \bigcup e_3 = e_3 \bigcup e_1 = c. \quad (6)$$

Следовательно, e_1/d и e_3/d перспективны оба фактору c/e_2, а потому в силу следствия из теоремы 5 они изоморфны. Аналогично, e_1/d и e_2/d изоморфны, и по теореме 3 c/d является прямым произведением любых двух из этих факторов. Мы заключаем

Теорема 8. *Пусть* θ_1, θ_2, θ_3 *отношения конгруентности на алгебре A с одноэлементной подалгеброй и перестановочными отношениями конгруентности. Пусть,* $\alpha_1 = (\theta_2 \bigcup \theta_3) \bigcap [\theta_1 \bigcup (\theta_2 \bigcap \theta_3)]$ *и циклически,* $\beta = (\theta_1 \bigcap \theta_2) \bigcup (\theta_2 \bigcap \theta_3) \bigcup (\theta_3 \bigcap \theta_1)$ *и пусть* γ *двойственно* β. *Тогда шесть факторов* α_i/β *и* γ/α_i *изоморфны* и γ/β *является прямым произведением любых двух из них.*

Отметим теперь, что в случае групп или в случае луп, удовлетворяющих второму достаточному условию $(ux)x^{-1}=u=x^{-1}(xu)$ теоремы 1, мы имеем в силу теоремы 7 центральный изоморфизм. Но в $(\gamma/\beta) = (\alpha_1/\beta) \times (\alpha_3/\beta)$ для любых $[a, 1]$ из α_1/β и $[1, y]$ из α_3/β имеем

$$([a,\ 1]^{-1}[1,\ y])[a,\ 1] = [(a^{-1}x)a,\ 1^{-1}y] = [1,\ y].$$

В силу центрального изоморфизма $[(a^{-1}x)a,\ 1^{-1}1] = [x,\ 1]$ для элемента $[x,\ 1]$ из α_1/β, соответствующего элементу $[1, y]$ в α_3/β. Но каждое $[x,\ 1]$ соответствует некоторому $[1, y]$; следовательно, $(a^{-1}x)a = x$ для всех *x, a*. Умножая справа на a^{-1}, получим $a^{-1}x = xa^{-1}$. Мы заключаем:

Следствие. *Если A есть группа или лупа, в которой* $(ux)x^{-1} = u = x^{-1}(xu)$ *для всех x, u, то* α_i/β, *определенные в теореме 8, коммутативны*

.

6.7.6. Полупрямые произведения

Пусть *S* некоторая подалгебра конечного прямого произведения $A_1 \times \ldots \times A_r$ абстрактных алгебр A_i. Тогда *S* есть подалгебра прямого произведения $S_1 \times \ldots \times S_r$ подалгебр S_i элементов из A_i, являющихся *i*-ми компонентами элементов из *S*. Мы будем говорить, что *S* есть *полупрямое произведение* алгебр S_i.

Отображение каждого элемента $a = [a_1, \dots, a_r]$ из S в его i-ю компоненту a_i является, очевидно, гомоморфизмом θ_i: $S \to S_i$. Кроме того, если рассматривать θ_i как отношение конгруентности на S, то два элемента конгруентны по mod каждого θ_i тогда и только тогда, если они совпадают. Следовательно,

$$\wedge\, \theta_i = \theta_1 \cap \dots \cap \theta_r = 0.$$

Обратно, пусть $\theta_1, \dots, \theta_r$—отношения конгруентности на абстрактной алгебре S. Если мы определим класс вычетов по mod θ_i, содержащий произвольно заданное $a \in S$, как i-ю компоненту *сц* элемента a, то соответствие $a \to [a_1, \dots, a_r]$ будет гомоморфизмом S на подалгебру прямого произведения $S_1 \times \dots \times S_r$ алгебр S_i классов вычетов алгебры S по mod θ_i. Кроме того, каждый элемент из S_i является i-й компонентой некоторого $a \in S$. Наконец, гомоморфизм $a \to [a_1, \dots, a_r]$ есть изоморфизм в том и только в том случае, если $\wedge\, \theta_i = 0$.

Используя несколько более разработанную систему обозначений, то же самое доказательство можно применить к полупрямым произведениям, содержащим бесконечное число сомножителей. Следовательно.

Теорема 9. *Представления абстрактной алгебры А в виде полупрямого произведения находятся во взаимно однозначном соответствии с множествами отношений конгруентности на А, удовлетворяющими условию* $\wedge\, \theta_i = 0$.

Рассмотрим теперь для любых $a \neq b$ частично упорядоченное множество $C(a, b)$ всех отношений конгруентности θ на A таких, что $a \not\equiv b (\mathrm{mod}\ \theta)$. Если T какая-нибудь подцепь в $C(a, b)$, то мы определим объединение τ элементов $\theta \in T$ по правилу: $x \equiv y(\tau)$ означает, что $x \equiv y\ (\theta)$ для некоторго $\theta \in T$. Очевидно, что $a \not\equiv b(\tau)$ и что (так как A имеет финитарные операции) τ есть отношение конгруентности. Следовательно, в $C(a, b)$ каждая подцепь имеет верхнюю грань τ. Используя (АС2) из гл. 6.3.6, мы устанавливаем, что $C(a, b)$ имеет *максимальный* элемент σ.

Рассмотрим теперь гомоморфный образ H алгебры A по modσ. Если φ есть любое отношение конгруентности на H, то $a \equiv b(\varphi)$, ибо в противном случае σ не было бы максимальным. Следовательно, наименьшее отношение конгруентности на H, при котором $a \equiv b\ (\theta)$, есть отношение конгруентности θ_0, содержащееся во всяком другом отношении конгруентности за исключением 0. Следовательно, если отношения конгруентности $\varphi_i \neq 0$, то $\wedge\, \varphi_i \geq \varphi_0 > 0$.

Напоминаем, что быть рефлексивным, симметричным, транзитивным и обладать свойством замены—все это экстенсивно достижимые свойства (т. е. свойства замыкания); следовательно, таковым же является свойство быть отношением конгруентности, а потому φ_0 существует.

Имеет место

Теорема 10. *Всякая алгебра А может быть представлена в виде полупрямого произведения не разложимых в полупрямое произведение алгебр.*

В самом деле, пересечение элементов σ есть 0, ибо при заданных $a \neq b$, $a \not\equiv b(\sigma)$ для некоторого σ. С другой стороны, каждое $H = H(\sigma)$ неразложимо в том строгом смысле, что в любом изоморфном представлении *Н* в виде полупрямого произведения имеется множитель, изоморфный *Н*.

6.7.7. Теорема Куроша — Орэ

Теорема 9 представляет интерес в вопросе о представлениях элемента (а именно, 0) дедекиндовой структуры в виде пересечения бóльших элементов. Теория идеалов также включает в себя такие представления, отчасти потому, что любой «неразложимый» идеал является «примарным» в том смысле, что он содержит степень «простого» идеала.

Назовем элемент *а* дедекиндовой структуры *L* «разложимым в пересечение» (короче, «разложимым»), если этот элемент есть пересечение $x \cap y$ элементов $x > a$ и $y > a$, больших, чем он сам; в противном случае будем называть его «неразложимым». Простое индуктивное рассуждение показывает, что если *L* удовлетворяет условию обрыва возрастающих цепей, то каждое $a \in L$ имеет представление в виде пересечения неразложимых элементов. В самом деле, в этом случае мы можем, согласно обобщенному принципу индукции, предположить, что всякие $x, y > a$ имеют такое представление. Естественно, далее, назвать компоненту x_k в «разложении» $a = x_1 \cap \ldots \cap x_r$ элемента *а* «сократимой», если $a = x_1 \cap \ldots \cap x_{k-1} \cap x_{k+1} \cap \ldots \cap x_h$; и очевидно, что любому разложению *а* соответствует разложение, в котором ни одна из компонент не сократима. Такие разложения будем называть «несократимыми».

Лемма 2. *Пусть* $a = x_1 \cap \ldots \cap x_r = x_1^* \cap \ldots \cap x_s^*$ *—два любых несократимых разложения элемента а на неразложимые компоненты. Тогда вместо любого* x_i *можно подставить подходящее* x^*_j *и получить новое разложение элемента а.*

Доказательство. Положим $y_i = x_1 \cap \ldots \cap x_{i-1} \cap x_{i+1} \cap \ldots \cap x_r$. Тогда в силу несократимости $y_i > a$, но $x_i \cap y_i = a$. Образуем теперь

$z_j=y_i \cap x^*_j$; очевидно, что $y_i \geq z_j \geq a$, и, так как $z_j \leqslant x^*_j$, $a \leqslant z_1 \cap \ldots \cap z_s \leqslant x^*_1 \cap \ldots \cap x^*_s \leqslant a$. Но в силу теоремы 6 п.6.5 подструктура между $a = x_i \cap y_i$ и y_i изоморфна подструктуре между x_i и $x_i \cup y_i$, и так как x_i неразложимо в последней, то a неразложимо в первой. Следовательно, некоторое z_j есть a, и $x_1 \cap \ldots \cap x_{i-1} \cap x^*_j \cap x_{i+1} \cap \ldots \cap x_r = a$.

Теорема 11. *Число компонент в несократимых разложениях любого элемента не зависит от разложения: в предыдущей лемме* $r = s$.

Доказательство. Выберем r минимальным и будем заменять один за другим элементы x_i элементами x^*_j. Мы получим в конце концов $a = x^*_{j(1)} \cap \ldots \cap x^*_{j(r)}$, откуда в силу несократимости элементов x^*_j индексов $j(i)$ будет самое меньшее s, а потому $s \leq r$. Следовательно, ввиду минимальности $r, s = r$.

Следствие. *Пусть А—произвольная алгебра, у которой отношения конгруентности перестановочны и удовлетворяют условию обрыва возрастающих цепей. Число множителей в любом несократимом представлении А в виде полупрямого произведения неразложимых в прямое произведение алгебр является одинаковым для всех таких представлений.*

Доказательство. Конечность числа сомножителей в любом несократимом представлении, а также существование таких представлений вытекает из условия обрыва возрастающих цепей и из теоремы 9. Доказательство завершается ссылкой на на теоремы 11 и 2. В силу теоремы 1 этот результат справедлив для групп с операторами, а также для конечных луп и для луп, удовлетворяющих соотношению $x^{-1}(xu) = (ux)x^{-1} = u$.

6.7.8. Теорема Орэ

Рассмотрим следующую теорему, принадлежащую Орэ. Назовем элемент e дедекиндовой структуры *прямым объединением* элементов $a_1, \ldots, a_n$ (в обозначениях $e = a_1 \times \ldots \times a_n$), если элементы a_i независимы и имеют своим объединением e.

Теорема 12. *Пусть L произвольная дедекиндова структура конечной длины. Если I имеет два представления* $a_1 \times \ldots \times a_m$ *и* $b_1 \times \ldots \times b_n$ *в виде прямого объединения неразложимых элементов, то m=n и элементы* a_i *и* b_j *попарно проективны.*

Доказательство. Пусть $\bar{a}_i = a_1 \times \ldots \times a_{i-1} \times a_{i+1} \times \ldots \times a_m$ и $\bar{b}_j = b_1 \times \ldots \times b_{j-1} \times b_{j+1} \times \ldots \times b_n$. Тогда $I = a_i \times \bar{a}_i = b_j \times \bar{b}_j$ для всех i, j. Если $I = b_j \times \overline{a}_i$ для некоторых i, j, то мы скажем, что a_i *замещаемо* элементом b_j. Покажем, что каждое a_i замещаемо некоторым b_j; не нарушая общности, предположим, что $i = 1$.

Случай I. Если $a_1 \cup \overline{b_j} = \overline{a_1} \cup b_j = I$ для некоторого j, то по размерности

$$d\,[a_1] = d\,[I] - d\,[\overline{b_j}] + d\,[a_1 \cap \overline{b_j}] = d\,[b_j] + d\,[a_1 \cap \overline{b_j}] \geqslant d\,[b_j].$$

Аналогично, $d\,[b_j] \geq d\,[a_1]$. Этим доказано, что $d\,[a_1] = d[b_j]$, откуда $d\,[a_1 \cap \overline{b_j}] = d\,[\overline{a_1} \cap b_j] = 0.$ Мы заключаем, что $a_1 \cap \overline{b}_j = \overline{a}_1 \cap b_j = 0$, а потому a_1 и b_j *взаимно* замещаемы.

Случай II. Предположим, что $a_1 \cup \overline{b}_j < I$ для некоторого j, скажем, для $j == 1$. Обозначим $(a_1 \cup \overline{b_k}) \cap b_k$ через q_k. Поскольку из $a_1 \cup \overline{b}_j \geq b_1$ вытекало бы $a_1 \cup \overline{b}_1 \geqslant b_1 \cup \overline{b}_1 = I$, что противоречит предположению, то, очевидно, $q_1 = (a_1 \cup \overline{b}_1) \cap b_1 < b_1$. Кроме того, так как элементы b_k независимы и $q_k \leq b_k$, элемент $c = \bigvee\limits_{k=1}^{n} q_k$ есть *прямое* объединение q_k, а потому $d\,[c] = \sum d\,[q_k] < \sum d\,[b_k] = d\,[I]$, откуда $c < I$.

Далее, полагая $d_k = a_1 \cup \overline{b}_k$, получаем индукцией по n

$$\begin{aligned} c = \bigvee_{k=1}^{n} q_k &= (b_1 \cap d_1) \cup [\bigvee_{k=2}^{n} b_k \cap \bigwedge_{k=2}^{n} d_k] \\ &= d_1 \cap (b_1 \cup [\bigvee_{k=2}^{n} b_k \cap \bigwedge_{k=2}^{n} d_k]) \text{ в силу L5} \\ &= d_1 \cap (\bigvee_{k=1}^{n} b_k) \cap \bigwedge_{k=2}^{n} d_k \text{ в силу L5, ибо } \bigwedge_{k \neq 1} d_k \geqslant b_1 \\ &= d_1 \cap I \cap \bigwedge_{k=2}^{n} d_k = \bigwedge_{k=1}^{n} (a_1 \cup \overline{b}_k) \geqslant a_1. \end{aligned} \tag{7}$$

Поэтому в силу L5

$$(c \cap \overline{a}_1) \cup a_1 = c \cap (\overline{a}_1 \cup a_1) = c \text{ и } c \cap \overline{a}_1 \cap a_1 = 0,$$

откуда $c = a_1 \times (c \cap \overline{a}_1)$.

Отсюда в силу индукции по длине структуры следует, что в любом представлении элемента $c = q_1 \times \ldots \times q_n = a_1 \times (c \cap \overline{a}_1)$ в виде прямого объединения неразложимых множителей e_{kk} элементов q_k элемент a_1 может быть замещен некоторым множителем $e_{k1} > 0$

некоторого q_k; для краткости мы пишем $e=e_{k1}$. Тогда всилу определения $c = e\times (c \cap \overline{a}_1)$, откуда

$$e \cup \overline{a}_1 = e \cup (c \cap \overline{a}_1) \cup \overline{a}_1 \text{ в силу L4}$$
$$= a_1 \cup (c \cap \overline{a}_1) \cup \overline{a}_1 = a_1 \cup \overline{a}_1 = I \text{ в силу (7).}$$

Но $d[e]=d[a_1]$ в силу свойства заменяемости в $c=a_1\times(c\cap \overline{a}_1)$; следовательно, $e \cap \overline{a}_1 = 0$ и $I = e\times \overline{a}_1$. Кроме того, $e= b_k$. В самом деле, $e \cap (b_h \cup \overline{a}_1) \leqslant e \cap \overline{a}_1 = 0$ и в силу L5

$$e \cup (b_h \cap \overline{a}_1) = (e \cup \overline{a}_1) \cap b_h = I \cap b_h = b_h,$$

откуда e является прямым множителем в $b_h = e \times (b_h \cap \overline{a}_h)$. Но по предположению b_k неразложимо; следовательно, $e = b_k$ и $I=b_k\times \overline{a}_k$. Следовательно, a_1 замещаемо элементом b_k, где $k\neq 1$.

Случай III. Предположим, что $a_1 \cup \overline{b}_j = I$ для всех j, однако $\overline{a}_1 \cup b_j < I$ для всех j—единственная остающаяся возможность. Тогда, так же как и в случае II, но меняя ролями a_1 и b_1, элемент b_1 можно заместить некоторым $a_k \neq a_1$, скажем (при изменении нумерации), a_m. Тогда отображение $x\to(x \cup a_m) \cap \overline{a}_m$ есть перспективность $[0, \overline{b}_1]$ на $[0, \overline{a}_m]$. Следовательно,если обозначить $(b_j \cup a_m)\cap \overline{a}_m$ через b^*_j, то $\overline{a}_m = a_1\times \ldots \times a_{m-1} = b^*_2 \times \ldots \times b^*_n$. В силу индукции по длине a_1 замещаемо некоторым b^*_j в $[0, \overline{a}_m]$. Кроме того, $b_j \cup \overline{a}_1$ содержит

$$b_j \cup a_m = (b_j \cup a_m) \cap (\overline{a}_m \cup a_m) = [(b_j \cup a_m) \cap \overline{a}_m] \cup a_m = b^*_j \cup a_m.$$

Следовательно, $b_j \cup \overline{a}_1$ содержит $b^*_j \cup a_2 \cup \ldots \cup a_{m-1} \cup a_m =$

$= \overline{a}_m \cup a_m$ (ибо a_1 замещаемо в $\overline{a}_m$ элементом b^*_j; таким образом, $b_j \cup \overline{a}_1 = I$. Но $d[a_1] = d[b^*_j] = d[b_j]$, так как все три элемента проективны. Следовательно, $I=b_j\times \overline{a}_1$, и элемент a_1 замещаем элементом b_j, что и требовалось доказать.

Используя теорему 4 и следствие из теоремы 5, выводим

Следствие 1. *Пусть $A = A_1 \times \ldots \times A_m = B_1 \times \ldots \times B_n$ два любых представления алгебры A в виде прямого произведения неразложимых сомножителей, где* а) *A имеет одноэлементную подалгебру,* б) *все отношения, конгруентности на A перестановочны и* в) *структура отношений конгруентности на A имеет конечную длину. Тогда $m = n$ и алгебры A и B попарно изоморфны.*

Следствие 2. *Пусть G произвольная группа с операторами или лупа, удовлетворяющая условию* $x^{-1}(xu)=(ux)x^{-1}=u$; *предположим также, что структура отношений конгруентности на G имеет конечную длину. Тогда в любых двух представлениях G в виде прямого произведения неразложимых в прямое произведение сомножителей сомножители являются попарно центрально-изоморфными.*

6.7.9. Структуры подгрупп

Хотя дедекиндова структура отношений конгруентности (нормальных делителей) на группе *G* наиболее ясно характеризует строение группы *G*, структура *L(G) всех* подгрупп группы *G* также представляет интерес. Эти две структуры изоморфны, если группа *G* коммутативна.
Следуя Прюферу, будем называть группу *G обобщенной циклической,* если любые два элемента a, $b \in G$ являются степенями подходящего третьего элемента $c = c(a, b)$ из *G*.
Теорема 13. *Структура L(G) всех подгрупп группы G дистрибутивна тогда и только тогда, если G является обобщенной циклической группой.*
Доказательство. Пусть *G*—обобщенная циклическая группа; в силу одностороннего дистрибутивного закона структура *L(G)* дистрибутивна, если $A \cap (B \cup C) \leqslant (A \cap B) \cup (A \cap C)$ для любых трех подгрупп *A, B, C* группы *G*. Предположим, что $a = bc\ [a \in A,\ b \in B,\ c \in C]$; в силу определения элементы b и c можно представить как степени $b = d^m$, $c = d^n$ некоторого третьего элемента *d;* следовательно, $d^{m+n} = a$. Если m'= общ. наим. кр. $[m, m+n]$ и n' = общ. наим. кр. $[n, m + n]$, то, очевидно, $d^{m'} \in A \cap B$ и $d^{n'} \in A \cap C$; следовательно, если $h = \lambda m' + \mu n'$ = общ. наиб. дел. (m', n'), то $d^h \in (A \cap B) \cup (A \cap C)$. Но, согласно элементарной теории чисел,
$$h = (m', n') = ([m, m+n],$$
$$[n, m+n]) = [(m, n), m+n] = m+n.$$
Следовательно, $a = d^{m+n} =$
$$= d^h \in (A \cap B) \cup (A \cap C).$$

Обратно, предположим, что структура *L(G)* дистрибутивна. Пусть a, b—два любых элемента из *G* и пусть *m, n*—два любых целых числа. Если *A, B, C* обозначают подгруппы, порожденные соответственно элементами a, b и c-=ab, то $(A \cup B) \cap C =$
$= (A \cap C) \cup (B \cap C)$. Но $A \cap C$ порождено наименьшей степенью $a^r = c^k$ элемента a, которая является также степенью элемента c, а $B \cap C$ порождено равным образом наименьшей степенью $b^s = c^k$. Очевидно, что a^r и b^s перестановочны; следовательно, подгруппа, порожденная

ими, состоит из элементов $a^{mr}b^{ns}=c^{mh+nk}$ и в силу предположения $a^{mr}b^{ns}=ab$ для некоторых m, n; следовательно, $a^{mr-1} = b^{1-ns}$. Поэтому a^{mr-1} перестановочно с b, а следовательно, и с b^s; но $a^{mr}=(a^r)^m$ перестановочно с b^s; следовательно, само a перестановочно с b^s. Далее, $b^{1-ns}=a^{mr-1}$ перестановочно с a; следовательно, с a перестановочно $b=(b^s)^n b^{1-ns}$ и $ab=ba$. Мы заключаем, что элементы a и b порождают в G *коммутативную* подгруппу. Согласно основной теореме для (возможно и бесконечных) абелевых групп с двумя образующими, эта подгруппа или является циклической, или имеет два *независимых* образующих a', b'. В последнем случае, как можно проверить непосредственно, подгруппы, порожденные элементами a', b' и $a'b'$, не удовлетворяют дистрибутивному закону.

Можно определить также все конечные группы G, для которых структура $L(G)$ является дедекиндовой. Это имеет место в случае, если все подгруппы группы G являются нормальными делителями; таким образом, достаточно, чтобы группа G порождалась своим центром Z и двумя другими элементами a и b, удовлетворяющими соотношениям

$$a^2=b^2=c,\ c^2=1,\ a^{-1}b^{-1}ab=c. \tag{8}$$

Таким образом, a, b порождают так называемую группу *кватернионов* Гамильтона. Дедекинд показал, что «гамильтоновы» группы только что описанного типа суть единственные группы, все подгруппы которых являются нормальными делителями.

Достаточно также, чтобы все подгруппы группы G были *перестановочны*. Это условие выполняется для «модулярных p-групп» $G=\{Y,\ b\}$, порождаемых коммутативной подгруппой Y, элементы которой удовлетворяют тождеству $y^{p^m} = 1$, где p простое число, и элементом b порядка p^r, где для некоторого $n = ps\geq p$

$$b^{-1}yb=y^n \text{ для всех } y\in Y, \text{ где } n^{p^r}\equiv 1 \pmod{p^m} \text{ и } p^s\geqslant 4, \text{ если } p=2. \tag{9}$$

Модулярный закон выполняется также для подгрупп так называемых «модулярных pq-групп» $G = \{X,\ a\}$, порождаемых коммутативной подгруппой X, элементы которой удовлетворяют тождеству $x^p = 1$ для некоторого простого p, и элементом a, порядок которого — степень простого числа q^β, удовлетворяющим соотношению

$$a^{-1}xa=x^n \text{ для всех } x\in X, \text{ где } n^q\equiv 1 \pmod{p}. \tag{10}$$

Однако подгруппы модулярных pq-групп не обязаны быть перестановочными. Обратно, Ивасава показал, что каждая конечная группа G, у которой структура $L(G)$ дедекиндова, является прямым произведением определенных выше гамильтоновых групп, модулярных pq -групп и модулярных p-групп.

6.7.10. Классификация групп при помощи структур подгрупп

Джонсом было показано, что структуры подгрупп модулярных p-групп и модулярных pq-групп изоморфны структурам подгрупп соответствующих коммутативных групп. Этим подчеркивается тот факт, что *не изоморфные группы могут обладать изоморфными структурами подгрупп.* Однако тот факт, что все группы, структуры подгрупп которых изоморфны структурам подгрупп абелевых групп и, следовательно, дедекиндовы, являются группами метабелевыми, показывает, что группы с изоморфными структурами подгрупп обычно очень схожи.

Так, для групп G, порядок которых n может быть разложен на $k \leq 4$ простых множителей (считая и повторяющиеся множители), длина структуры $L(G)$ равна k. Это можно показать, заметив, что знакопеременная группа пятой степени является единственной не разрешимой из этих групп. Далее, структура $L(G)$ является цепью тогда и только тогда, когда G циклическая группа, порядок которой есть степень простого числа. Представляется вероятным, что если группа G проста, то она с точностью до изоморфизма определяется структурой $L(G)$.

Если ограничиться рассмотрением коммутативных групп, то о них можно сказать больше. Исключая те случаи, когда силовские подгруппы циклические, структура $L(G)$ полностью определяет группу G. Более того, каждый автоморфизм структуры $L(G)$ индуцируется групповым автоморфизмом группы G. Этот результат связан для случая «элементарных» абелевых групп порядка p^n, которые могут рассматриваться как векторные пространства размерности n над конечными полями, с хорошо известными свойствами коллинеаций в проективных геометриях.

Далее, можно доказать, используя основную теорему и характеры, что *структура всех подгрупп любой конечной коммутативной группы A является себе двойственной.*

6.8. Полудедекиндовы структуры

6.8.1. Определение

Структура L конечной длины называется *полудедекиндовой* тогда и только тогда, если ее элементы удовлетворяют условию

(ξ'). Если x и y покрывают a и $x \neq y$, то $x \cup y$ покрывает x и y.

Мы видели уже, что любая полудедекиндова структура удовлетворяет условию Жордана — Дедекинда и что функция размерности удовлетворяет неравенству

$$d[x \cap y] + d[x \cup y] \leqslant d[x] + d[y]. \qquad (1)$$

Из этого следует, что

если x покрывает $x \cap y$, то $x \cup y$ покрывает y. (2)

Действительно, в силу (1), поскольку $d[x] = d[x \cap y] + 1$, мы получаем после сокращений $d[x \cup y] \leq d[y] + 1$. Но из равенства $d[x \cup y] = d[y]$ вытекало бы $x \leq y$, вследствие чего $x = x \cap y$ не могло бы покрывать $x \cap y$. Обратно, из (2) следует (1).

Следуя Вилкоксу, будем говорить, что (x, y) есть *модулярная пара,* если из $x \geq z$ следует $x \cap (y \cup z) = (x \cap y) \cup z$. В дедекиндовой структуре все пары модулярны.

Лемма 1. *Элементы x, y полудедекиндовой структуры конечной длины образуют модулярную пару тогда и только тогда, если в* (1) *имеет место равенство.*

Доказательство. Предположим, что $u = \cap (y \cup z) > (x \cap y) \cup z = t$ для некоторого $z \leq x$. Тогда (см. рис. 3, *а*) $x \cap y \leq t < u \leq x$, однако $t \cup y = u \cup y$. Это означает, что любая связная цепь, соединяющая $x \cap y$ и x и проходящая через t, u, *укорачивается* при отображении $s \rightarrow s \cup y$. Но в силу (2) эта цепь переходит в связную цепь; следовательно, в (1) имеет место неравенство. Обратно, если x и y образуют модулярную пару, то преобразование $s \rightarrow s \cup y$ интервала $[x \cap y, x]$ обладает однозначным обратным преобразованием $s \cup y \rightarrow (s \cup y) \cap x$; следовательно, связные цепи переходят в связные цепи той же длины; поэтому $d[x \cup y] - d[y] = d[x] - d[x \cap y]$, и в (1) имеет место равенство.

Однако условие, что в (1) имеет место равенство, является симметричным относительно x и y; следовательно, если в полудедекиндовой структуре x, y есть модулярная пара, то y, x также будет модулярной парой. Напротив, если структура L конечной длины не полудедекиндова, то для некоторых x, y, z мы имеем: y и z покрывают

$a = y \cap z$, однако $y \cup z > x > z$. Тогда x, y не есть модулярная пара; однако y, x является модулярной парой. Действительно, если $t \leq y$, то элемент $t \cup (x \cap y) = t \cup a$ удовлетворяет соотношению $a \leq t \cup a \leq y$. Если $t \cup a = y$, то $t \cup (x \cap y) = y \geq (t \cup x) \cap y$; если $t \cup a = a$, то $t \leq a$, а потому $t \cup (x \cap y) = a = x \cap y = (t \cup x) \cap y$.

Выведим следующий результат.

Теорема 1. *Структура L конечной длины является полудедекиндовой тогда и только тогда, если отношение модулярности для пар элементов в L симметрично.*

Это условие имеет то преимущество, что оно применимо к структурам с непрерывными цепями равно как и с дискретными цепями. Имеет место также следующее условие.

Теорема 2. *Структура L конечной длины является полудедекиндовой тогда и только тогда, если выполняется условие:*

(γ) *если (x_0, y_0), (x_1, y_1), ..., (x_n, y_n) произвольная максимальная цепь в кардинальном произведении $[x \cap y,\ x] \times [x \cap y,\ y]$, то $x_0 \cup y_0,\ x_1 \cup y_1,\ \dots,\ x_n \cup y_n$ является максимальной цепью в $[x \cap y,\ x \cup y]$.*

Доказательство. В силу определения кардинального произведения если (x_i, y_i) покрывает (x_{i-1}, y_{i-1}) в L, то или $x_i = x_{i-1}$ и y_i покрывает y_{i-1} или x_i покрывает x_{i-1} и $y_i = y_{i-1}$; в каждом из этих случаев если структура L полудедекиндова, то в силу (2) $x_i \cup y_i$ покрывает элемент $x_{i-1} \cup y_{i-1}$ или равен ему; следовательно, элементы $x_i \cup y_i$ образуют максимальную цепь. Обратно, если структура L не полудедекиндова, то условие *(ξ')* не имеет места, и $a \cup a,\ a \cup y,\ x \cup y$ или $a \cup a,\ x \cup a,\ x \cup y$ не будет максимальной цепью интервала $[x \cap y,\ x \cup y]$.

Покажем теперь, что (γ) имеет место в *любой* дедекиндовой структуре M. Пусть $\{x_\alpha\}$ и $\{y_\alpha\}$—максимальные цепи соответственно в $[x \cap y, x]$ и $[x \cap y, y]$; тогда $\{x_\alpha \cup y_\alpha\}$ будет цепью в $[x \cap y,\ x \cup y]$; пусть c— произвольное сечение в $\{x_\alpha \cup y_\alpha\}$; положим $x_c = (c \cup y) \cap x$ и $y_c = (c \cup x) \cap y$. В силу L5 мы имеем

$$[(x_\alpha \cup y_\alpha) \cup y] \cap x = (x_\alpha \cup y) \cap x = x_\alpha \cup (y \cap x) = x_\alpha.$$

Следовательно, если $x_\alpha \cup y_\alpha \leq c$, то $x_\alpha \leq (c \cup y) \cap x = x_c$, и , если $x_\alpha \cup y_\alpha \geq c$, то $x_\alpha \geq (c \cup y) \cap x = x_c$, ибо структурные операции изотонны. Имеем также

$$x_\alpha = x_\alpha \cap x \leq (x_\alpha \cup y_\alpha) \cap x = x_\alpha \cup (y_\alpha \cap x) \leq x_\alpha \cup (y \cap x) = x_\alpha;$$

поэтому из $x_\alpha \cup y_\alpha \leq c$ вытекает $x_\alpha = (x_\alpha \cup y_\alpha) \cap x \leq c \cap x$ и из

$x_\alpha \cup y_\alpha \geq c$ вытекает $x_\alpha = (x_\alpha \cup y_\alpha) \cap x \geq c \cap x$. Поскольку $x_c \in \{ x_\alpha \}$ и $y_c \in \{ y_\alpha \}$, то мы заключаем, что $x_c \cup y_c = x_\alpha \cup y_\alpha$. Следовательно, одно и то же сечение c во множестве индексов α определяет x_c и $c \cap x$; и, поскольку $\{ x_\alpha \}$ максимально, $x_c = c \cap x \in \{ x_\alpha \}$. Аналогично, $y_c = c \cap y \in \{ y_\alpha \}$, где c есть один из, индексов α. Отсюда

$$c \leqslant (c \cup y) \cap (x \cup c) = [(c \cup y) \cap (x \cup y)] \cap (c \cup x),$$ так как

$$c \leqslant x \cup y, \ = [\{(c \cup y) \cap x\} \cup y] \cap (c \cup x) = [(c \cup y) \cap x] \cup [y \cap (c \cup x)]$$

в силу L5, применяемого дважды,

$$= x_c \cup y_c = (c \cap x) \cup (c \cap y) \leqslant c.$$

Мы заключаем, что $c = x_c \cup y_c \in \{x_\alpha \cup y_\alpha\}$; таким образом, эта цепь максимальна, что и требовалось доказать.

6.8.2. Примеры

Пусть теперь R—произвольное ассоциативное кольцо с единицей 1; рассмотрим множество n-векторов $\xi = (x_1, \ldots, x_n)$, $\eta=(y_1, \ldots, y_n)$ с компонентами из R. Мы определим сложение n-векторов и умножение векторов на скаляры $c \in R$ по обычным формулам

$$\xi + \eta = (x_1 + y_1, \ \ldots, \ x_n + y_n), \qquad c\xi = (cx_1, \ \ldots, \ cx_n).$$

Тем самым определена алгебраическая система, которую обозначают через $V(R; n)$ и называют n-мерным векторным пространством над R.

Линейная комбинация векторов $\xi_1, \ldots, \xi_r$ определяется как произвольная конечная сумма $c_1\xi_1 + \ldots + c_r \xi_r$ векторов $c_i\xi_i$, представляющих собой произведения векторов ξ_i на скаляры. Совокупность линейных комбинаций множества X векторов называется *векторным подпространством,* порожденным множеством X, и может быть обозначена через $\overline{X}$; векторное подпространство, порожденное пустым множеством, определяется как (0, ..., 0). Легко проверяется, что операция $X \to \overline{X}$ является операцией замыкания; следовательно, векторные подпространства пространства $V(R; n)$ образуют полную структуру. В силу теоремы 1 из п. 6.6. эта структура дедекиндова; она обозначается через $PG(R; n—1)$ и называется *(n — 1)-мерной проективной геометрией над R.*

Наиболее важным является случай, когда R кольцо с делением; в этом случае каждый ненулевой вектор ξ порождает минимальное собственное подпространство, состоящее из всевозможных векторов $c\xi$. Действительно, любое $c\xi \neq 0$ порождает подпространство, содержащее вектор $c^{-1}(c\xi)=\xi$. Эти минимальные собственные подпространства называются *проективными точками;* они соответствуют проективным точкам в однородных координатах в обычном смысле.

Каждое подпространство пространства *V (R; n)* является объединением проективных точек, содержащихся в нем; кроме того, если $\overline{\xi} \neq \overline{\eta}$, то $\overline{\xi} \cup \overline{\eta}$ содержит $\overline{\eta + \xi}$, определяющее собой третью проективную точку; следовательно,

Теорема 3. *Для любого кольца с делением R и любого n PG(R; n* — 1) *является дедекиндовой структурой, в которой объединение двух различных точек содержит третью точку.*

Аффинное подпространство пространства *V (R; n)* определяется как подмножество, содержащее вместе с любыми ξ, η также все «аффинные комбинации» ξ + *c*(η — ξ); пустое подмножество рассматривается как аффинное подпространство. Легко проверить, что любое непустое аффинное подпространство является трансформацией векторного подпространства при преобразовании τ→τ+ξ; кроме того, в соответствии с п. 6.5.1, аффинные подпространства пространства *V (R; n)* образуют полную структуру. Она обозначается через *AG (R; n)* и называется *n-мерной аффинной геометрией над H.*

Далее, пусть *R*—кольцо с делением. Пусть в *PG (R; n) S* есть векторное подпространство, состоящее из всех (n + 1)-векторов (x_1, ...,x_n, O); мы называем его «бесконечно удаленной гиперплоскостью». Каждый вектор $\xi \in V(R; n+1) - S$ определяет в точности один вектор $(x_{n+1}^{-1}x_1, \ldots, x_{n+1}^{-1}x_n, 1)$ в аффинном подпространстве *A* векторов {y_1, ...,y_n, 1); это есть обычное соответствие между однородными и неоднородными координатами. Кроме того, это соответствие переводит линейные комбинации в аффинные комбинации и обратно; следовательно, оно переводит векторные подпространства, не содержащиеся в *S*, в аффинные подпространства из *A* и обратно. Таким способом можно показать, что

Теорема 4. *Для любого кольца с делением R и любого n AG(R; n) изоморфно подмножеству тех элементов из PG(R; n), которые не содержатся в максимальном собственном элементе S<I.*

Можно тогда показать, что *AG(R; n)* является всегда полудедекиндовой структурой. Можно показать, что любое кардинальное произведение, а также гомоморфный по объединениям образ и выпуклая подструктура полудедекиндовых структур конечной длины сами являются полудедекиндовыми структурами.

Пусть $L \to L_1$—такой гомоморфизм и пусть x_1 и y_1 покрывают a_1 в L_1. Пусть *a* наибольший из прообразов элемента a_1 и *x, у* произвольные прообразы элементов x_1, y_1. Образуем цепи от *а* до $a \cup x$ и до $a \cup y$;

элементы, покрывающие a в каждой из них, будут соответственно прообразом x^* элемента x_1 и прообразом y^* элемента y_1. Следовательно, $x^* \cup y^*$ будет прообразом элемента $x_1 \cup y_1$; таким образом, элемент $x_1 \cup y_1$ покрывает x_1 и y_1.

6.8.3. Зависимость и ранг

Далее будем рассматривать только полудедекиндовы структуры конечной длины. Последовательность $x_1, ..., x_r$ элементов полудедекиндовой структуры называется *независимой* тогда и только тогда, если она удовлетворяет симметричному условию

$$d[x_1 \cup \ldots \cup x_r] = d[x_1] \cup \ldots \cup d[x_r]. \tag{3}$$

В силу (1) $d[x_1 \cup \ldots \cup x_r] \leqslant d[x_1 \cup \ldots \cup x_{r-1}] +$
$+ d[x_r] - d[(x_1 \cup \ldots \cup x_{r-1}) \cap x_r]$;

следовательно, в любом случае (в силу индуктивных соображений) $d[x_1 \cup \ldots \cup x_r] \leqslant d[x_1] + \ldots + d[x_r]$. Равенство имеет место тогда и только тогда, если а) $(x_1 \cup \ldots \cup x_{r-1}) \cap x_r = 0$, б) $x_1 \cup \ldots \cup x_{r-1}$ и x_r являются «модулярной парой» и в) элементы $x_1, ..., x_{r-1}$ независимы: В частности, *любое подмножество независимого множества является независимым.*

Наиболее существенным является случай, когда x_i — точки. В этом случаемы можем назвать число $d[\sup X]$ *рангом* X. Так как элемент $x_1 \cup \ldots \cup x_{k+i}$ самое большее покрывает $x_1 \cup \ldots \cup x_k$, то мы имеем $d[x_1 \cup \ldots \cup x_{k+1}] = d[x_1 \cup \ldots \cup x_k] + 1$ или $d[x_1 \cup \ldots \cup x_k]$ в соответствии с тем, имеет или не имеет места равенство $(x_1 \cup \ldots \cup x_k) \cap x_{k+1} = 0$. Отсюда следует

Лемма 1. *Последовательность $x_1, ..., x_r$ точек независима тогда и только тогда, если* $(x_1 \cup \ldots \cup x_k) \cap x_{k+1} = 0$ *для* $k = 1, \ldots, r-1$.

Поскольку условие (3) симметрично, отсюда вытекает условие, что $(x_1 \cup \ldots \cup x_k) \cap x_{k+1} = 0$ для $k = 1, \ldots, r-1$, инвариантно при всех перестановках элементов x_i. (4)

Далее, пусть $x_1, ..., x_r$ —любая последовательность точек, зависимая или независимая. Мы можем построить, элемент за элементом, подпоследовательность, ни один из элементов которой не содержится в объединении предшествующих, но объединение которой есть $x_1 \cup \ldots \cup x$, — это осуществляется просто вычеркиванием x_k, если $x_k \leqslant x^1 \cup \ldots \cup x_{k-1}$. Мы заключаем:

Лемма 2. *Любое множество точек содержит независимое подмножество, имеющее тот же самый ранг (т. е. то же самое объединение).*

Теорема 5. *Пусть X—множество независимых элементов* $x_i>0$ *полудедекиндовой структуры L конечной длины. Тогда* x_i *порождают подструктуру, изоморфную телу всех подмножеств множества X.*

Доказательство. Свяжем с каждым подмножеством *S* из *X* элемент sup *S*. Множество всех sup *S* замкнуто, очевидно, относительно операции объединения. В силу неравенства мини-макса $\sup(S \cap T) \leqslant \sup S \cap \sup T$, в то время как в силу LI—L3 $\sup(S \cup T) = \sup S \cup \sup T$. В силу этого и в силу (1)

$$d[\sup S \cap \sup T] = d[\sup S] + d[\sup T] - d[\sup S \cup T] =$$
$$= \sum_{S} d[x_i] + \sum_{T} d[x_i] - \sum_{S \cup T} d[x_i] = \sum_{S \cap T} d[x_i] = d[\sup S \cap T],$$

как может быть непосредственно подсчитано. Следовательно, неравенство минимакса является равенством, и пересечения соответствуют теоретико-множественным произведениям в такой же мере, как объединения теоретико-множественным суммам. Условие $x_i > 0$ обеспечивает то, что мы имеем не только гомоморфизм, но и изоморфизм.

Дилуорс получил, используя абстрактную теорию линейной зависимости, результат, что *каждая конечная структура изоморфна подструктуре полудедекиндовой структуры.*

6.8.4. *М*-структуры

Структура *с дополнениями* есть структура, удовлетворяющая условию

L7. Для любого заданного *x* существует такое *x′,* что $x \cap x' = 0$, $x \cup x' = I$;

структура *с относительными дополнениями* есть структура, удовлетворяющая условию

L7R. Для любых заданных $a \leq x \leq b$ существует такое *y,* что $x \cap y = a$, $x \cup y = b$.

Любая структура конечной длины с относительными дополнениями является также структурой с дополнениями.

Теорема 6. *Полудедекиндова структура конечной длины является структурой с дополнениями тогда и только тогда, если она удовлетворяет условию*

L7'. *I есть объединение точек.*

Она является структурой с относительными дополнениями тогда и только тогда, если

L7R'. *Каждый элемент x есть объединение точек.*

Доказательство. Во всякой структуре конечной длины из L7 вытекает L7'. Действительно, пусть *a* объединение всех точек; *a'*=0, ибо этот элемент не может содержать точку. Обратно, пусть *L* полудедекиндова структура (конечной длины), удовлетворяющая условию L7'; пусть дано $a \in L$. Существует последовательный ряд точек $p_1 \not\leq a$, $p_2 \not\leq a \cup p_1$ ..., до тех пор, пока не будет $a \cup p_1 \cup \ldots \cup p_r = I$. Так же, как и в лемме 1 р.6.8.3, $d[a \cup (p_1 \cup \ldots \cup p_r)] = d[a] + r$; следовательно,

$$d[a \cap (p_1 \cup \ldots \cup p_r)] \leqslant d[a] + d[p_1 \cup \ldots \cup p_r] -$$
$$-d[a \cup p_1 \cup \ldots \cup p_r] \leqslant d[a] + r - (d[a] + r) = 0.$$

Отсюда следует, что $a' = p_1 \cup \ldots \cup p_r$ является дополнением *a* и даже что *a, a'* образуют модулярную пару. Рис. 7, *а* показывает, что в общей структуре ни условие L7', ни условие L7R' не влекут за собой L7, не говоря уже об L7R.

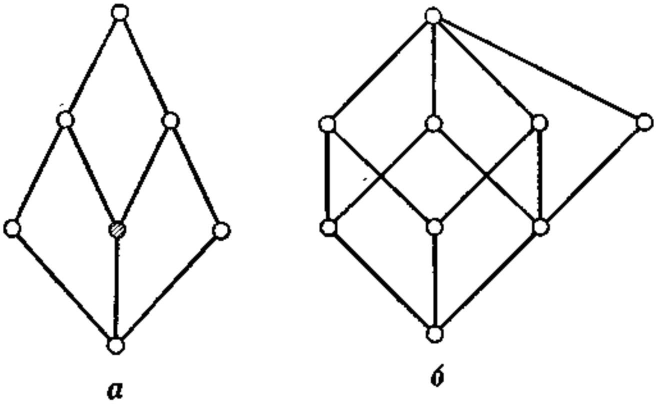


Рис. 7.

Далее, в любой структуре конечной длины из L7R вытекает в силу предыдущего соображения, примененного к структурес дополнениями [0, *x*], условие L7R'. Обратно, пусть *L* полудедекиндова структура, удовлетворяющая условию I 7R'; пусть даны $a \leq x \leq b$. Существует последовательный ряд точек $p_1, \ldots, p_r \leq b$ такой, что $(x \cup p_1 \cup \ldots \cup p_k) \cap p_{k+1} = 0$, $x \cup p_1 \cup \ldots \cup p_r = b$.

Так же, как в лемме 1, р.6.8.3 и в предыдущем параграфе, если мы положим $z = p_1 \cup \ldots \cup p_r$, то $d[a \cup z] = d[a] + r$ и $d[x \cup z] = d[x] + r = d[b]$. Пусть $y = a \cup z$. Тогда $x \cup y =$

$= x \bigcup \bar{a} \bigcup z = \bar{x} \bigcup z = b;$ в то же время $x \cap y \geq a$ и

$$d[x \cap y] \leqslant d[x] + d[a \cup z] - d[x \cup y] =$$
$$= d[x] + d[a] + r - (d[x] + r) = d[a].$$

Следовательно, $x \cap y = a$ и условия L7R и L7R' *эквивалентны, если имеет место* (ξ').

Рис. 7, *б* показывает, что существуют полудедекиндовы структуры, являющиеся структурами с дополнениями, но не являющиеся структурами с относительными дополнениями.

Определение. *Полудедекиндова структура с относительными дополнениями называется М-структурой.*

Любая полудедекиндова структура удовлетворяет «аксиомам замены» Штейница и Маклейна, как определено, соответственно, в (5) и (6),

если p, q точки и $a \leq a \cup q \leq a \cup p$, *то* $a \cup p = a \cup q$, (5)

если p точка, то или $p \leq a$, или $a \cup p$ покрывает a. (6)

Теорема 7. *Структура L конечной длины является М-структурой тогда и только тогда, если она удовлетворяет* а) L7R *или* L7R' и б) *одному из условий* (ξ'), (4), (5), (6). ,

Доказательство. Поскольку в любой структуре конечной длины из L7R вытекает L7R', то достаточно показать, что если имеет место L7R', то условия (ξ'), (4), (5) и (6) вытекают в циклическом порядке одно из другого. Действительно, тогда любое из этих условий вместе с L7R' повлечет за собой (ξ') и L7R', следовательно, также (ξ') и L7R, а потому *L* будет *М*-структуро.

Прежде всего, как и в п.6.8.3, лемма 1, из (ξ') вытекает (4).

Далее, из (4) вытекает (5). Используя L7R', легко получить независимые точки $p_1, ..., p_r$, для которых $p_1 \cup ... \cup p_r = 0$. Кроме того, так как $a \cup q \leq a \cup p$, последовательность $p_1, ..., p_r, p, q$ не является независимой; следовательно в силу (4) последовательность $p_1, ..., p_r, q, p$ также не является независимой; поэтому $p \geq a \cup q) \cap p > 0$; следовательно, $p \leq a \cup q$ и $a \cup p \leq a \cup q$, что дает в силу предположения и P2 $a \cup p = a \cup q$

Далее, из (5) вытекает (6). Если $a < b \leq a \cup p$, то существует в силу L7R' точка q, содержащаяся в b, но не содержащаяся в a; очевидно, что $a < a \cup q \leq b$. В силу (5) $a \cup p = a \cup q \leq b$; следовательно, $a \cup p = b$ и $a \cup p$ покрывает a. Наконец, из (6) вытекает *(ξ')*. Если x, y покрывают a, то в силу L7R' существуют точки p, q, для которых $a < a \cup p \leq x$, $a < a \cup q \leq y$, откуда $a \cup p = x$ и $a \cup q = y$. Следовательно, $x \cup y = a \cup p \cup q$, и этот элемент, согласно (6), покрывает $x = a \cup p$ и $y = a \cup q$.

6.8.5. Структуры разбиений; степень трансцендентности

Проективные геометрии и аффинные геометрии в р.6.8.2 являются *М*-структурами: каждый элемент есть объединение точек. (Даже 0 есть объединение пустого множества точек!) Далее, любое кардинальное произведение, а также структурно-гомоморфный образ и интервальная подструктура *М*-структур сами являются *М*-структурами.

Теорема 8. *Множество всех разбиений* π *любой конечной совокупности X является М-структурой.*

Доказательство. Пусть X имеет n элементов $x_1, ..., x_n$. Пусть $d[\pi] = n — n(\pi)$, где $n(\pi)$ есть число неперекрывающихся подклассов, на которые π разбивает X. Легко показать, что разбиения $\pi'\geq\pi$ образуют интервальную подструктуру $[\pi, I]$, изоморфную структуре всех разбиений системы $n(\pi)$ предметов; следовательно, для доказательства (ξ') нам нужно доказать лишь, что если π' и π'' различные минимальные разбиения (отношения эквивалентности), то $\pi' \cup \pi''$ покрывает π' и π''. Но минимальное разбиение просто отождествляет два неравных элемента; следовательно, $\pi' \cup \pi''$ либо отождествляет три элемента, либо отождествляет элементы двух различных пар. В обоих случаях $\pi' \cup \pi''$ покрывает π' и π''. Чтобы доказать L7R', мы заметим только, что если $x_i\equiv x_j(\mathrm{mod}\ \pi)$, то существует минимальное разбиение π_{ij}, отождествляющее x_i и x_j, но не отождествляющее другие пары элементов; отсюда $\pi_{ij} \leq\pi$. Ясно тогда, что для таких π_{ij} имеет место $\pi = \vee\ \pi_{ij}$, далее применяем теорему 7.

Орэ указал, что только что определенная *симметричная структура разбиений степени п* является до некоторой степени аналогом аффинной геометрии.

Уитмен показал, что *каждая структура изоморфна подструктуре структуры всех разбиений некоторого-(бесконечного) множества.*

Теорема 9. *Алгебраически замкнутые подполя произвольного поля I образуют М-структуру.*

Доказательство. Пусть *F*—любое алгебраически замкнутое подполе и пусть *Н* и *К* покрывают *F*. Выберем элементы *х* в *Н, у* в *К,* не принадлежащие *F*. Тогда *Н* есть множество чисел, алгебраических над кольцом *{F, х}, К* есть множество чисел, алгебраических над *{F, у}* и $H \cup K$ — множество чисел, алгебраических над *{F, х, у}*. Если теперь z принадлежит $H \cup K$, то некоторый полином $p(x, y, z) = 0$; если z не принадлежит *Н*, то этот полином должен включать в себя положительные степени y; следовательно, y алгебраично над $\{F, x, z\}$. Отсюда следует, что если $H<Z\leq H \cup K$, то Z содержит y, а потому $Z=H \cup K$; этим доказана справедливость (ξ') и «аксиомы замены» Штейница-Маклейна. Наконец, структура удовлетворяет условию L7R': если $G>F$, то выберем максимальное множество чисел x_α из G, алгебраически независимых над F; тогда алгебраические замыкания множеств $\{F, x_\alpha\}$ будут точками, объединение которых есть G.

Мы можем определить степень трансцендентности G над F как теоретико-структурную размерность G над F. Тогда приведенные выше теоремы о зависимости, ранге и т. д. включают в себя как следствия большую часть теории алгебраической зависимости Штейница; например, число элементов x_α есть $d[G/F]$.

6.8.6. Геометрии на плоскости

Другие важные примеры M-структур дают нам аффинные и проективные геометрии на плоскости.

Аффинная геометрия на плоскости может быть определена комбинаторно как совокупность точек и множеств точек, называемых прямыми, удовлетворяющих условиям:

APG1. Любые две различные точки p, q находятся на одной и только одной прямой $p \cup q$.

APG2. Для заданных $p \cup q$ и $r \not< p \cup q$ существует в точности одна прямая $r \cup s$ такая, что $p \cup q$ и $r \cup s$ не имеют общих точек.

APG3. Существуют три точки, не лежащие на одной прямой.

Легко показать, что если присоединить пустое множество 0 и множество всех точек I, то мы получим M-структуру A.

Две прямые в A называются *параллельными* тогда и только тогда, если они тождественны или не имеют общих точек. Легко показать, что отношение параллелизма есть отношение эквивалентности. Действительно, отношение параллелизма рефлексивно и симметрично. В то же время, если μ и $\mu'\neq\mu$ параллельны λ, то они не могут иметь общую точку p, — в противном случае имелись бы две прямые, проходящие через p и параллельные λ, что противоречило бы APG2.

Отсюда следует, что мы можем присоединить к A «бесконечно удаленную прямую» λ_∞, содержащую в точности по одной точке для каждого множества параллельных прямых. Две конечные прямые будут тогда пересекаться точно один раз —или на конечной части плоскости, или (если они параллельны) на λ_∞; кроме того, каждая конечная прямая принадлежит в точности одному множеству параллельных прямых и имеет, следовательно, в точности одну общую точку с λ_∞. Следовательно, после расширения *A* условие APG2 может быть заменено условиями:

PG2'. Каждая пара различных прямых имеет в точности одну общую точку.

PG2". Каждая прямая содержит три или большее число точек. (Действительно, если некоторая «прямая» в *A* содержит в точности одну точку, то в силу APG2 имеется точно одна прямая, проходящая через каждую другую точку, что невозможно в силу APol и APG3.)

Множество точек и прямых, удовлетворяющее условиям APG1, PG2', PG2" и APG3, называется *проективной геометрией на плоскости;* нами доказана

Теорема 10. *Каждая аффинная геометрия на плоскости может быть расширена путем присоединения «бесконечно удаленной прямой до проективной геометрии на плоскости.*

Покажем теперь, что аффинные геометрии на плоскости происходят от систем координат более общих, нежели кольца с делением.

Определяем *точку* как пару чисел (x, y). В качестве прямых мы берем а) прямые $x = c$ с бесконечно большим наклоном и б) прямые $y = a + sx$ с «наклоном» s и ««y-пересечением» a. Это определение включает в себя тернарную операцию над нашими «числами». Очевидно, что прямые с бесконечно большим наклоном являются «параллельными». Если $(x, a+sx) = (x', a' + sx')$, то $x = x'$, $a + sx = a' + sx$, т. е., прямые с одинаковым конечным «наклоном» параллельны тогда и только тогда, если

(I) из $a + sx = a' + sx$ вытекает $a = a'$.

Далее, $x = c$ и $(x, a+sx)$ имеют, очевидно, в качестве общей точки $(c, a+sc)$. Следовательно, для выполнимости условия APG2 достаточно, чтобы выполнялось:

(II) если $s \neq s'$, то $a + sx = a' + s'x$ имеет в точности одно решение.

Кроме того, APG3 выполняется тривиальным образом. Следовательно, наша тернарная операция $a+bc$ будет определять аффинную геометрию на плоскости тогда и только тогда, если выполняется APG1, а именно,

(III) при заданных $x, y, x' \neq x, y$ мы имеем в точности для одной пары a, s $a+sx = y$ и $a+sx' = y'$.

Наименьшая из известных систем R, удовлетворяющая условиям (I) — (III), но не являющаяся кольцом с делением, была открыта Вебленом и Веддербарном. Ее тернарная операция $a+sx$ может быть определена косвенным образом через бинарное сложение и умножение. По сложению R является элементарной абелевой группой порядка 9; ее общий элемент имеет вид $m + na$ (m и n целые числа по mod3). Здесь $1=1+0a$ есть единица в умножении и $(x+y)z = xz+yz$; следовательно, умножение полностью определяется соотношениями $aa = 2$, $a(a+1) = =2a+1$, $a(a+2)=a+1$, $a(2x) = 2(ax)$.

Пусть P—конечная проективная плоскость с $n + 1$ точкой на некоторой прямой λ. Если p—любая точка, не принадлежащая λ, то каждая прямая, проходящая через p, пересекает λ точно в одной точке; следовательно, через p проходит n+1 прямая. Это верно также для *любой* прямой λ', не проходящей через p; но никакое λ' не исчерпывает собой $P — \lambda$; следовательно, *каждая прямая в P содержит в точности $n + 1$ точку.* Аналогично, каждая точка лежит в точности на $n + 1$ прямой. Поскольку каждая отличная от p точка лежит в точности на одной прямой, проходящей через p, и каждая из $n + 1$ таких прямых содержит n точек, отличных от p, то P должно содержать в точности n^2+n+1 точку.

Поэтому *аффинная* плоскость из n^2 точек, получаемая из P отбрасыванием «бесконечно удаленной прямой», обладает следующим свойством. Имеется n+1 последовательных способов $\pi_1, ...,\pi_{n+1}$ разбиения этих n^2 точек в ряды («параллельные прямые») из n элементов таких, что каждая пара точек находится в одном и том же ряду при одном и только одном π_i.

Более общей комбинаторной проблемой является следующая. Пусть m и n—целые числа, такие, что $(n — 1)|(mn—1)$; пусть $r=(mn—1)/(n — 1)$. Пусть мы стараемся найти такие разбиения $\pi_1, \pi_2, \pi_3,\ldots$ mn предметов в m рядов по n предметов, что никакие два предмета не встречаются в одном и том же ряду более, чем один раз. После r разбиений каждый предмет будет тогда находиться в общем ряду с $r(n — 1) = mn — 1$ предметами, что равно числу всех отличных от него предметов; следовательно, r есть наибольшее возможное число таких разбиений. Мы определим (m, n)-проблему как проблему отыскания такого множества разбиений; это множество называется (m, n)-системой.

Для четырех разбиений $\pi_1, \pi_2, \pi_3, \pi_4$ описанного типа множества из n^2 элементов можно построить так называемый $(n\times n)$-эйлеров квадрат следующим образом. Строки и столбцы квадрата ассоциируются с рядами в разбиениях π_1 и π_2 соответственно; таким образом, (i, j)-пространство в квадрате ассоциируется с точкой p_{ij}, попадающей при разбиении π_1 в i-й ряд, а при разбиении π_2 в j-й ряд. Ряды в π_3, π_4 нумеруются числами $k = 1,...,n$ и $l=1, ...,n$; (i, j)-пространство снабжается тогда символом (k, l), если p_{ij} попадает в k-й ряд разбиения π_3 и в l-й ряд разбиения π_4.

Поскольку $\pi_i \cap \pi_j = 0$ при $i \neq j$, то каждое k и каждое l встречаются точно один раз в каждой строке и в каждом столбце; кроме того, никакая пара *(k, l)* не встречается более чем один раз. Такая конструкция называется эйлеровым (или греко-латинским) квадратом. Если число *n* нечетное, то его легко построить.

Используя такой квадрат, легко построить *магический квадрат* путем вписывания числа $(k—1)n+1$ в клетку, имеющую вход *(k, l)*.

6.9. Дедекиндовы структуры с дополнениями

6.9.1. Определение

Настоящий раздел будет касаться *дедекиндовых структур с дополнениями,* т. е. дедекиндовых структур L с 0 и I, удовлетворяющих условию L7. Каждое $x \in L$ имеет «дополнение» x' такое, что $x \cap x' = 0$ и $z \cup x' = I$. Мы покажем сперва, что дедекиндова структура с дополнениями является необходимо структурой с относительными дополнениями в смысле р. 6.8.4. В самом деле, пусть заданы $a \leq x \leq b$ и пусть x'—произвольное дополнение элемента x. Тогда

$$(a \cup x') \cap b \cap x = (a \cup x') \cap x = a \cup (x' \cap x) = a \cup 0 = a,$$
$$x \cup a \cup (x' \cap b) = x \cup (x' \cap b) = (x \cup x') \cap b = I \cap b = b.$$

Следовательно, $(a \cup x') \cap b = a \cup (x' \cap b)$ является относительным дополнением элемента x в интервале $[a, b]$. Этим доказана

Теорема 1. *Всякая дедекиндова структура с дополнениями является структурой с относительными дополнениями.*

Следствие. *Для дедекиндовых структур конечной длины каждое из условий* L7, L7R, L7', L7R' р. 6.8.4 *влечет, все остальные.*

Доказательство. В силу теоремы 6 из р. 6.8 условия L7 и L7' эквивалентны и условия L7R и L7R' также эквивалентны. Мы показали, что L7 влечет L7R; обратно, любая структура конечной длины имеет 0 и I; в такой структуре L7R тривиальным образом влечет L7.

Отметим, что из следующей леммы непосредственно вытекает, что L7→L7R'.

Лемма. *«Разности»* $d_k = x'_k \cap x_{k+i}$ *для двух последовательных членов произвольной цепи* $0 = x_0 < x_1 < \ldots < x_s$ *являются независимыми элементами, объединение которых есть* x_s.

Доказательство. Индукцией по s мы можем показать, что $(d_1 \cup \ldots \cup d_{s-1}) \cup d_s$ есть $x_{s-1} \cup (x'_{s-1} \cap x_s)$; в силу L5 это есть $(x_{s-1} \cup x'_{s-1}) \cap x_s = x_s$. Элементы d_k являются независимыми в смысле р. 6.6.5, ибо

$$(d_1 \bigcup \ldots \bigcup d_{s-1}) \bigcap d_s = x_{s-1} \bigcap d_s = x_{s-1} \bigcap x'_{s-1} \bigcap x_s = 0.$$

Теорема 2. *Пусть L—метрическая структура с дополнениями и пусть x'—произвольное дополнение элемента x в L. Если* $|x — y| < s$, *то у имеет дополнение у', удовлетворяющее условию* $|x' — y'| < \varepsilon$.

Доказательство. Пусть $t = x \cap y \leq x$. Образуем цепь $0 \leqslant x' \leqslant t \bigcup x' \leqslant x \bigcup x' = I$. В силу леммы, примененной к цепи $0 \leq t \leq x \leq I$, элементы t, $t' \cap x$, x' являются независимыми элементами, объединение которых есть I; следовательно, если t' какое-нибудь дополнение элемента t, то $(t' \cap x) \cup x' = t^*$ другое дополнение этого элемента. Но в силу р. 6.5.7, (7) при замене x на t в $(t' \cap x) \cup x'$ произойдет сдвиг самое большее на расстояние $|t — x|$; при этой замене мы получаем $(t' \cap t) \cup x' x'$; следовательно, $|t^* — x'| \leq |t — x|$. Аналогично, мы можем найти дополнение y^* элемента $у$ такое, что $|y^* — t^*| \leq |y — t|$. В силу неравенства треугольника

$$|y^* - x'| \leqslant |x - x \bigcap y| + |y - x \bigcap y|$$

$$= v[x] - v[x \bigcap y] + v[y] - v[x \bigcap y], \text{ так как } x, y \geqslant x \bigcap y$$

$$= v[x \bigcup y] - v[y] + v[y] - v[x \bigcap y] = |x - y|.$$

6.9.2. Примеры

Поскольку дополнения множеств удовлетворяют условию L7, алгебра всех подмножеств некоторой совокупности *I* образует дедекиндову структуру с дополнениями; по той же причине любое тело подмножеств из *I* образует дедекиндову структуру с дополнениями. Далее, мы видим, используя L7', что линейные подпространства любого конечномерного векторного пространства [т. е. проективные геометрии *PG (R; n)*] образуют дедекиндову структуру с дополнениями. Аналогично, нормальные делители любого конечного прямого произведения простых групп, а также идеалы (соответственно инвариантные подалгебры) прямой суммы произвольного конечного множества простых колец (соответственно линейных алгебр) образуют дедекиндовы структуры с дополнениями. Более обще, Ω-подгруппы любого конечного прямого произведения «простых» групп с системой Ω операторов образуют дедекиндову структуру с дополнениями. Далее мы увидим, что эти результаты применимы равным образом к бесконечномерным векторным пространствам и к произвольным прямым произведениям групп с операторами.

Дедекиндовы структуры с дополнениями важны также в теории представлений. Пусть Ω-произвольное множество линейных операторов, действующих на линейном пространстве *I*. Тогда условие того, что дедекиндова структура *M* Ω-инвариантных подпространств (т. е. таких подпространств *X*, что $x\omega \in X$, если $x \in X$ и $\omega \in \Omega$. Известно, что *M* является структурой с дополнениями, если Ω— конечная группа или полупростая гиперкомплексная алгебра, за исключением, быть может, случая, когда Ω есть группа, порядок которой кратен характеристике поля скаляров для *I*) пространства *I* является структурой с дополнениями, означает по определению, что если Ω может быть «полуприведено» Ω-инвариантным подпространством *X*, то Ω может быть «вполне приведено» подпространством *X* и некоторым дополнительным инвариантным подпространством *X'*. Это есть условие того, что *M* — структура с дополнениями, и эквивалентность этого условия другим условиям следствия из теоремы 1 играет важную роль в теории представлений.

Можно показать, что любая структура, двойственная дедекиндовой структуре с дополнениями, а также кардинальное произведение, структурно-гомоморфные образы и интервальные подструктуры дедекиндовых структур с дополнениями сами являются дедекиндовыми структурами с дополнениями.

6.9.3. Проективные геометрии как структуры

Пусть *M* является дедекиндовой структурой с дополнениями конечной длины; мы определяем «точки» в *M* обычным образом и «прямые» как элементы из *M,* покрывающие точки. Тогда, очевидно,

PG1. Две различные точки лежат на одной и только одной прямой.

Легко, далее, доказать

PG2. Если прямая λ пересекает две стороны треугольника (не в их точке пересечения), то она пересекает также третью сторону.

Доказательство. Пусть эти две стороны суть $\cup\, q$, $p \cup r$ и пусть точки пересечения — $q^* \leq p \cup q$, $r^* \leq p \cup r$. Поскольку $q^* \neq p$, то $\lambda = q^* \cup r^*$. Для завершения доказательства мы подсчитываем

$$d[(q^* \cup r^*) \cap (q \cup r)] = d[q^* \cup r^*] + d[q \cup r] - d[q^* \cup r^* \cup q \cup r] \geqslant$$
$$\geqslant 2 + 2 - d[p \cup q \cup r] = 4 - 3 = 1,$$

неравенство здесь очевидно, поскольку $q^* \leq p \cup q \cup r$, $r^* \leq p \cup q \cup r$.

Определение. *Проективная геометрия есть система точек и прямых, удовлетворяющая условиям* PG1, PG2 и

PG3. *Каждая прямая содержит самое меньшее три точки. Она называется конечномерной тогда и только тогда, если*

PG4. *Существует конечное множество точек, такое, что любая «плоскость», содержащая это множество, содержит все пространство. Плоскость есть множество точек, содержащее вместе с точками p и q прямую, проходящую через p и q.*

Очевидно, что для любого $a \in M$ множество $S(a)$ всех точек $p \leq a$ есть плоскость. Обратно, пусть S — любая плоскость в M; мы можем представить $s = V_S P_i$ как объединение конечного подмножества точек $p_i \in S$, $s = p_1 \bigcup \ldots \bigcup p_r$; пусть $t = p_1 \cup \ldots \cup p_{r-1}$. Если точка $q \leq s = t \cup p_r$, то или $q \leq t$, откуда в силу индукции по r мы можем предположить, что $q \in S$, или $t < s$, и в этом случае $p_r \leqslant (q \bigcup p_r) \bigcap (t \bigcup p_r) = [(q \bigcup p_r) \bigcap t] \bigcup p_r$. Здесь $(q \cup p_r) \cap t$ есть точка (или 0, что является тривиальным случаем), ибо

$$d[(q \bigcup p_r) \bigcap t] = d[q \bigcup p_r] + d[t] - d[q \bigcup p_r \bigcup t] = d[q \bigcup p_r] - 1 \leqslant 1;$$

следовательно, в силу индукции этот элемент (будучи в t) входит в S. Следовательно, $q \in S$, так что $S = S(s)$, и структура *M изоморфна структуре плоскостей в ассоциированной с ней проективной геометрии.*

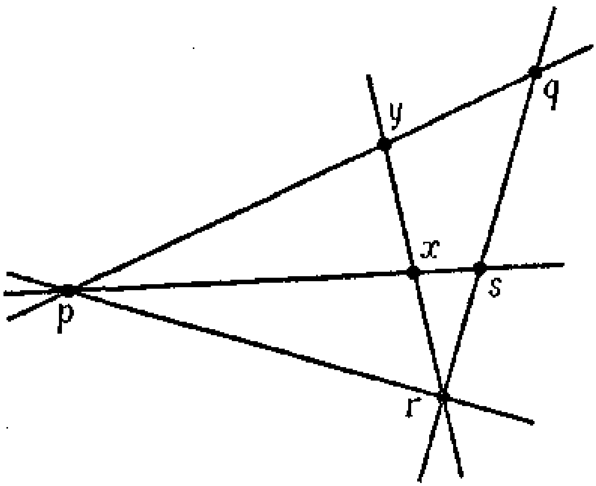


Рис. 8.

Обратно, пусть P любая система, удовлетворяющая условиям PG1 — PG2. Обозначим через $p \cup q$ (единственную) прямую, проходящую через p и q. Постулат PG1 утверждает, что $p \cup q = q \cup p$ и, далее, что если *r* лежит на $p \cup q$, то (поскольку $p \cup r$ и $p \cup q$ ба содержат *p* и *r*) $p \cup r = p \cup q$.

Теорема 3. *Плоскости любого пространства, удовлетворяющего условиям* PG1 — PG2, *образуют полную дедекиндову структуру. В этой структуре «объединение» двух любых плоскостей S и T есть множество всех точек на прямых* $s \cup t$, *соединяющих S и T при условии, если только ни S, ни T не являются пустыми множествами и если они оба не являются одной и той же точкой.*

Доказательство. В силу теоремы 1, р.5, плоскости образуют полную структуру. Рассмотрим теперь $S \cup T$. Предположим, что $p \leqslant s \bigcup t,\ q \leqslant s' \bigcup t'\ [s, s' \in S;\ t, t' \in T]$ и $x \leq p \cup q$. Если мы сможем показать, что $x \leqslant s'' \bigcup t''\ [s'' \in S,\ t'' \in T]$, то будет ясно, что множество точек на прямых, соединяющих *S* п *T*, есть плоскость. Но это множество содержит (за исключением отмеченных случаев), множества *S* и *T* и содержится в любой плоскости, содержащей *S* и *T*; следовательно, оно будет теоретико-структурным объединением *S* и *T*. Теорема доказана.

Вообще, пусть $X \vee Y$ означает множество точек на прямых, соединяющих *X* и *Y*; мы покажем, что для различных точек имеет место не только $p \vee q = q \vee p$ (что является очевидным), но также $p \vee (q \vee r) = (p \vee q) \vee r$; в силу симметрии и P2 достаточно показать, что $p \vee (q \vee r) \leqslant (p \vee q) \vee r$. Но если $x \in p \vee (q \vee r)$, то $x \in p \vee s$, где $s \in q \vee r$. Образуем теперь прямую $r \vee x$; в силу PG2 она будет пересекать $q \vee r$ в некоторой точке *y*: так как $r \vee x$ пересекает две стороны треугольника *pqs*, она должна пересекать и третью. Но $y \in p \vee q$, $x \in y \vee r \leq (p \vee q) \vee r$, что и завершает доказательство. (Случай, когда *p, q, r* коллинеарны, тривиален.)

Используя эти законы L2 —L3, мы имеем для рассмотренного в первом абзаце доказательства элемента *x*

$$x \in (s \vee t) \vee (s' \vee t') = (s \vee s') \vee (t \vee t'),$$

так что $x \leq s'' \vee t''$ для некоторых $s'' \in s \bigcup s',\ t'' \in t \bigcup t'$, что мы и хотели показать.

Наконец, структура плоскостей является дедекиндовой. Предположим, что $S \leq U$, и пусть *x*—любая точка из $(S \cup T) \cap U$; тогда $x \in S \bigcup T$ и $x \in U$. В силу предыдущего результата *x* лежит на некоторой прямой $s \bigcup t\ [s \in S,\ t \in T]$, но $x \in U$; следовательно, $x \bigcup s \in S \bigcup U = U$, а потому $t \in U$. Мы заключаем, что *x* принадлежит $S \cup (T \cap U)$. Тем самым доказано, что

$(S \cup T) \cap U \leq S \cup (T \cap U)$. В силу одностороннего модулярного закона получаем L5. Теорема 3 доказана.
Далее, L7' является очевидным, а потому условие L7 выполняется, если имеет место PG4. Этим доказано

Следствие 1. *Класс конечномерных дедекиндовых структур с дополнениями идентичен с классом структур «плоскостей» в системах, удовлетворяющих условиям* PGl, PG2 и PG4.

Следствие 2. *Конечномерные дедекиндовы структуры с дополнениями, удовлетворяющие* PG3, *являются структурами, плоскостей в различных конечномерных проективных геометриях.*

6.9.4. Перспективность и проективность

Как было впервые Дж. Нейманом, понятие «транспонированного интервала», введенное в р. 6.6 с целью получения теорем типа теоремы Жордана — Гельдера, эквивалентно понятию «перспективности», определяемому в классической проективной геометрии. Кроме того, последнее понятие естественным образом применимо к любой дедекиндовой структуре с дополнениями, как это приводится ниже.

Определение. *Два элемента а и b дедекиндовой структуры с дополнениями являются «перспективными» тогда и только тогда, если они имеют общее дополнение с, называемое «осью перспективности» для а и b.*

Известно, что в проективных геометриях любые два элемента, имеющие одинаковую размерность (точки, прямые и т. д.), являются перспективными; далее мы увидим, что это свойство выделяет в точности проективные геометрии среди других дедекиндовых структур с дополнениями, имеющих конечную длину.

Теорема 4. *Два элемента а и b связаны последовательностью перспективных тогда и только тогда, если идеалы* $[0, a]$ *и* $[0, b]$ *проективны.*

Доказательство. Если a и b перспективны посредством c, то $[0, a]$, $[c, I]$ и $[0, b]$ транспонированы в написанном порядке; следовательно, $[0, a]$ и $[0, b]$ проективны. Поскольку проективность есть транзитивное отношение, то $[0, a]$ и $[0, b]$ проективны (в смысле р. 6.6), если они связаны последовательностью перспективных элементов. Обратно, пусть $[0, a]$ и $[0, b]$ проективны и пусть $[u \cap v, v]$ и $[u, u \cup c]$ любые смежные интервалы в последовательности транспозиций, связывающей (в силу определения) $[0, a]$ и $[0, b]$. Тогда в силу леммы из р.6.9.1 относительное дополнение w элемента $u\cap\upsilon$ в υ и относительное дополнение w_1 элемента u в $u\cup\upsilon$ перспективны посредством любого $(u \cup \upsilon)' \cup t$, где t — любое относительное дополнение элемента $u\cap\upsilon$ в u. (Доказательство. Рассматриваем цепи $0\leqslant u\cap v\leqslant u\leqslant u\cup v\leqslant I$ и $\leqslant u\leqslant u\cup v\leqslant I$ $0\leqslant u\cap v\leqslant v\leqslant u\cup v\leqslant I$.) Следовательно, по индукции любые относительные дополнения элемента 0 в a и элемента 0 в b связаны последовательностью перспективных. Но a и b являются единственными такими относительными дополнениями, что и требовалось доказать.

6.9.5. Отношения конгруентности и эндоморфизмы

Если a нейтральный элемент в произвольной структуре, то главный идеал $[0, a]$ является модулем конгруентности для эндоморфизма $x \to \cup\ a$. Обратно, если L структура с относительными дополнениями, удовлетворяющая условию обрыва возрастающих цепей, то для любого отношения конгруентности θ идеал, состоящий из элементов $x\equiv 0(\theta)$, является главным идеалом $[0, a]$. Кроме того, в силу р. 6.3.6, $x \equiv y(\theta)$ в L тогда и только тогда, если некоторое $w\leq a$ является относительным дополнением элемента $x \cap y$ п $x \cup y$. Но отсюда вытекает

$$x\cup a = x\cup w\cup a \geqslant (x\cap y)\cup w\cup a = x\cup y\cup a \geqslant y\cup a$$

и, аналогично, $y \cup a \geq x \cup a$, откуда $x \cup a = y \cup a$. Обратно, из $x\cup a=y\cup a$ и $a\equiv 0(\theta)$ вытекает $x = x\cup 0 \equiv y\cup 0 = y\,(\theta)$. Этим доказан следующий результат.

Лемма 1. *Во всякой структуре с относительными дополнениями, удовлетворяющей условию обрыва возрастающих цепей, каждое отношение конгруентности ассоциируется со структурным эндоморфизмом* $x\to x\cup a$.

В соединении с теоремой 13 из р. 6.6 и с замечанием, что любая дедекиндова структура с дополнениями является структурой с относительными дополнениями, получается

Теорема 5. *Отношения конгруентности на дедекиндовой структуре L с дополнениями, имеющей конечную длину, находятся во взаимно однозначном соответствии с нейтральными элементами a из L.*

Соответствие имеет место с эндоморфизмом $x \to x \cup a$, индуцируемым элементом *a.* Заметим здесь также, что из следствия теоремы 9 р. 6.3 вытекает

Следствие. *Следующие условия для элемента a дедекиндовой структуры с дополнениями являются эквивалентными:* а) *элемент a нейтрален,* б) *элемент a принадлежит центру-*

Сравнивая с р. 6.2.9, получаем непосредственно следующий результат.

Лемма 2. *Любая дедекиндова структура с дополнениями, имеющая конечную длину, является кардинальным произведением дедекиндовых структур с дополнениями, не имеющих собственных отгсошений конгруентности.*

Такие структуры могут быть названы *простыми* в том смысле, что любая алгебра является «простой», если она не имеет нетривиальных отношений конгруентности (т. е. никаких отношений, кроме 0 и *I*). Это определение, относящееся к универсальной алгебре, включает в себя как частные случаи обычные определения простых групп, простых колец и простых линейных алгебр.

6.9.6. Теорема о разложении на множители

Покажем, что единственными простыми дедекиндовыми структурами с дополнениями конечной длины являются 2 и проективные геометрии (более точно, структуры, состоящие из плоскостей на подходящей проективной геометрии, которые мы будем называть ради сокращения «проективными геометриями»).

Лемма 3. *Две точки p и q перспективны тогда и только тогда, если $p \cup q$ содержит третью точку s.*

Доказательство. Пусть *a*—ось перспективности; тогда

$$d[a \cap (p \cup q)] = d[a] + d[p \cup q] - d[a \cup p \cup q] = \\ = d[I] - 1 + 2 - d[I],$$

а потому $a \cap (p \cup q)$ является третьей точкой на $p \cup q$. Обратно, если *s* третья точка на $p \cup q$, то, поскольку *p, s* и $(p \cup q')$ являются независимыми элементами с объединением *I*, элемент $e = (p \cup q)' \cup s$ есть дополнение для *p;* аналогично, он является дополнением для *q*, а потому *p* и *q* перспективны.

Пусть теперь *a*—ось перспективности для *p* и *q* и *b*—ось перспективности для *p* и *r*. Пусть *s* и *t*—третьи точки соответственно на $p \cup q$ и на $p \cup r$. Тогда $u = (q \cup t) \cap (r \cup s)$ является в силу V1 точкой; аналогично, точкой будет $(p \cup u) \cap (q \cup r)$, являясь на самом деле третьей точкой на $q \cup r$. Следовательно, в силу леммы 3 перспективность между точками есть транзитивное отношение. Итак, две точки проективны тогда и только тогда, если они перспективны.

Отсюда в силу теоремы 10 из р. 6.6 следует, что в «простой» дедекиндовой структуре с дополнениями L, имеющей конечную длину, все точки перспективны. Следовательно, в силу леммы 3 каждая прямая в L содержит самое меньшее три точки, т. е. выполняется PG3. Следовательно, в силу леммы 2 мы имеем

Теорема 6. *Всякая дедекиндова структура* с *дополнениями, имеющая конечную длину, является (однозначным образом) прямым произведением (т. е. кардинальным произведением) проективных геометрий.*

Случай структуры 2, состоящей из пустого множества и одной единственной точки, является исключительным.

Следствие. *Каждое из нижеследующих условий на дедекиндову структуру с дополнениями L, имеющую конечную длину, является необходимым и достаточным для того, чтобы она была проективной геометрией:* а) *структура L проста,* б) *структура L неразложима в прямое произведение,* в) *все точки из L перспективны,* г) *все точки из L проективны,* д) *любые две оценки на L являются линейно зависимыми.*

Детали доказательства мы предоставляем читателю. Несколько менее тривиальным является следующий результат.

Теорема 7. *В проективной геометрии эквивалентны следующие условия:* а) $d[x] = d[y]$, б) x и *y перспективны,* в) *x и y проективны.*

Доказательство. Если мы используем V1, то очевидно, что «б»→«в» →«а». Предположим теперь, что $d\ [x] = d\ [y] = s$, и пусть $p_1, ..., p_r$ и $q_1, ... , q_r$ являются базисами соответственно для $x — x \cap y$ и $y — x \cap y$. Пусть $t_1, .. ., t_r$—третьи точки на $p_1 \cup q_1, ..., p_r \cup q_r$. Тогда, очевидно,

$$x \cup t_1 \cup \dots \cup t_r = (x \cap y) \cup p_1 \cup t_1 \cup \dots \cup p_r \cup t_r =$$
$$= (x \cap y) \cup p_1 \cup q_1 \cup \dots \cup p_r \cup q_r = x \cup y,$$

а потому, по соображениям размерности, $x \cap y$, элементы p_1, элементы t_j и $(x \cup y)'$ являются независимыми элементами, объединение которых есть I. Следовательно, $(x \cup y)' \cup t_1 \cup \dots \cup t_r$ и $x = (x \cap y) \cup p_1 \cup \dots \cup p_r$ являются взаимными дополнениями; подобным же образом первый элемент и y являются взаимными дополнениями. Таким образом, x и y перспективны, откуда и следует, что «а»→«б». Теорема доказана.

6.9.7. Автоморфизмы как коллинсации; проективность

Рассмотрим *PG (R; n)* как структуру всех подпространств (n + 1)-мерного векторного пространства *V(R;+*1*)* (n+1)-векторов $k=(x_0, x_1,...,x_n)$ с однородными координатами в кольце с делением (т. е. в поле или в теле) *R*. Очевидно, что любой структурный *автоморфизм* структуры *PG(R;n)* переводит точки в точки, прямые в прямые и коллинеарные точки в коллинеарные точки (все термины употребляются в их обычном проективном смысле). Такие преобразования будем называть *коллинеациями;* обратно, любая коллинеация на *PG(R;n)* есть структурный автоморфизм.

Пусть $x \to \omega(x)$—какой-нибудь автоморфизм в *R* и $A = \| a_{ij} \|$ любая (n+1)×(n+1)-неособая матрица с коэффициентами в *R*. Тогда соответствие $\xi \to \eta$ на *V (R; n*+1*),* определяемое посредством

$$y_i = a_{i0}\,\omega(x_0) + a_{i1}\omega(x_1) + \ldots + a_{in}\omega(x_n), \tag{1}$$

называют *полулинейным* преобразованием. Поскольку это соответствие сохраняет линейную зависимость, хотя и с измененными коэффициентами, формула (1) определяет коллинеацию (структурный автоморфизм) на *PG(R;n).* Известным результатом является

Теорема 8. *Каждая коллинеация на PG(R;n) задается преобразованием вида* (1)*, если только n* > 1.

Набросок доказательства. Пусть $\alpha:\xi\to\alpha(\xi)$ — какая-нибудь коллинеация на *PG(R;n);* используя однородные координаты, пусть $\varepsilon_0, ...,\varepsilon_n$ обозначают единичные векторы (1, 0, ..., 0).....(0, 0, . . ., 1) в $V_{n+1}(R)$. С помощью «проективности» $\eta\to\zeta$, где $\zeta = (z_0, z_1, ..., z_n)$ дается формулой

$$z_i = b_{i0}\, y_0 + b_{i1}\, y_1 + \ldots + b_{in}\, y_n, \tag{2}$$

мы можем свести все к случаю, когда $\alpha(\varepsilon_i) = \varepsilon_i$ и (1, 1, ..., 1) переходит в себя. Но с помощью ε_i однородные координаты ζ однозначно определены; таким образом, $(\zeta \bigcup \varepsilon_2 \bigcup \ldots \bigcup \varepsilon_n) \bigcap (\varepsilon_0 \bigcup \varepsilon_1)$ проектирует η однозначным образом на прямую $\varepsilon_0 \cup \varepsilon_1$. Сравнивая с $(\xi \bigcup \varepsilon_2 \bigcup \ldots \bigcup \varepsilon_n) \bigcap (\varepsilon_0 \bigcup \varepsilon_1)$, устанавливаем взаимно однозначные соответствия $x_i \leftrightarrow z_i$ между координатами ξ и координатами ζ Кроме того, поскольку это соответствие есть коллинеация, алгебра вурфов Штаудта сохраняется; следовательно, соответствия дают кольцевые автоморфизмы ω_i с $z_i = \omega(x_i)$. Так как вектор (1, 1, ..., 1) инвариантен, все ω_i равны и $z_i = \omega(x_i)$. Доказательство завершается замечанием, что если *A* — матрица, обратная матрице *B,* определяемой формулами (2), то (1) получается вполне определенным образом.

Поскольку $\omega^{-1}\left(\sum a_{ki}\omega(x_i)\right)=\sum\omega(a_{ki})\,x_i$, можно видеть, что проективности на *PG (R; n)* образуют нормальный делитель группы всех коллинеаций. Его можно характеризовать теоретико-структурным образом как подгруппу, порожденную перспективностями.

6.9.8. Корреляции и полярности; ортодополнение

Определим теперь *корреляцию* в *PG (R; n)* как любой *дуальный* автоморфизм соответствующей структуры. Если *R** — кольцо антиизоморфное[2]) *R,* то соответствие

$$(x_0,\, x_1,\, \ldots,\, x_n) \longleftrightarrow y_0\,\omega(x_0)+y_1\,\omega(x_1)+\ldots+y_n\,\omega(x_n)=0 \qquad (3)$$

является дуальным изоморфизмом между *PG(R;n)* и *PG(R*;n).* Но если и α и β любые два дуальных изоморфизма, то $\alpha^{-1}\beta$ есть изоморфизм; следовательно, наиболее общий дуальный изоморфизм *PG (R; n)* на *PG(R*;n)* дается посредством формулы

$$(x_0,\, x_1,\, \ldots,\, x_n) \longleftrightarrow \sum_{i,\,j} y_i\,a_{ij}\,\omega(x_j)=0. \qquad (4)$$

Далее, поскольку *R** может быть определено теоретико-структурно (или «комбинаторно») алгеброй вурфов, то имеет место

Теорема 9. *Проективная геометрия PG (R; n)* [$n\geq2$] *имеет дуальный автоморфизм тогда и только тогда, если кольцо с делением R антиизоморфно себе. В этом случае имеет место взаимно однозначное соответствие (4) между корреляциями и неособыми билинейными формами.*

Мы можем писать $\eta A\xi^*=0$ с целью обозначения того факта, что ξ лежит на *(n* — 1)-плоскости, определяемой посредством η выражением (4); это есть род перпендикулярности между ξ и η. В случае, когда это отношение симметрично, для корреляции употребляется термин *полярность;* это есть условие

L9. $(x')'=x$ для любой плоскости *x*.

Записывая

$$0=\omega^{-1}(0)=\sum x_j\,b_{ji}\,\omega^{-1}(y_i) \quad [b_{ji}=\omega^{-1}(a_{ij})]$$

и отмечая, что роли индексов *i* и *j* симметричны, мы видим, что это является достаточным (и существенно необходимым) для того, чтобы ω было *инволюцией,* так что $\omega(\omega(x))=x$, и чтобы матрица *A* была «эрмитовой» в том смысле, что $\omega(a_{ij})=a_{ji}$. Далее, если матрица *дефинитна* в том смысле, что $\sum x_i\,a_{ij}\omega(x_j)=0$ влечет $x_0=x_1=\ldots=x_n=0$, то имеем условие

L10. $x\cap x'=0$ и $x\cup x'=I$.

Определение. *Дедекиндова структура называется структурой с ортодополнениями тогда и только тогда, если она допускает дуальный автоморфизм $x \to x'$, удовлетворяющий условиям* L9, L10.

(Любой дуальный автоморфизм обязательно удовлетворяем условию

L8. $(x \cap y)' = x' \cup y'$ и $(x \cup y)' = x' \cap y'$.)

Например, если Q кольцо кватернионов и если $\omega(x)$ есть кватернион, сопряженный x, то «дефинитная диагональная эрмитова форма» $\sum y_i \omega(x_i)=0$ определяет $PG(Q;n)$ как дедекиндову структуру с ортодополнениями. Вообще, можно доказать, что *каждое* $PG(R; n)$ с ортодополнениями может быть получено (с точностью до изоморфизма) как и выше из инволюции $x \to \omega(x)$ в R и дефинитной диагональной эрмитовой формы $\sum y_i a_i \omega(x_i) = 0$ с $\omega(a_i) = a_i$.

Геометрический интерес представляют также «нуль-системы», или полярности, такие, что

$$\xi \le \xi' \text{ для каждой проективной точки I.} \qquad (5)$$

В заключение отметим, что *свободная* дедекиндова структура с ортодополнениями, имеющая два образующих, есть $^4 2 \times PG(J_3,1)$ и состоит из 96 элементов;. $PG(J_3; 1)=1 \oplus 4 \oplus 1$.

6.9.9. Нейтральные элементы и идеалы

Мы видели (теорема 5, следствие), что условие для элемента быть *нейтральным* эквивалентно условию принадлежности *центру* структуры L. Кроме того, имеет место

Теорема 10. *Элемент принадлежит центру дедекиндовой структуры с дополнениями L тогда и только тогда, если он имеет единственное дополнение.*

Доказательство. В любой структуре $L = M \times N$ элемент $[I, 0]$ может иметь дополнением только $[0, I]$; следовательно, никакой элемент центра не может иметь более одного дополнения. Обратно, предположим, что элемент a имеет единственное дополнение a' и что $u \cap a = 0$. Тогда, по предположению, элементы u, a и любое дополнение $(u \cup a)'$ элемента $u \cup a$ являются независимыми элементами с объединением I; следовательно, $u \cup (u \cup a)' = a'$ и $u \le a'$. Таким образом, a' содержит каждое u, для которого $u \cap a = 0$. В частности, поскольку $[(a \cap x)' \cap x] \cap a = 0$ независимо от x, $(a \cap x)' \cap x$ содержится в a' так же, как и в x, и

$$x = (a \cap x) \cup [(a \cap x)' \cap x] \leqslant (a \cap x) \cup (a' \cap x) \leqslant x \cup x = x.$$

Следовательно, $x = (a \cap x) \cup (a' \cap x)$, и каждое $x \in L$ может быть записано в виде $y \cup z$ $[y \leq a, z \leq a']$. Из теоремы 7, р.6.6, вытекает, что L ость кардинальное произведение $[0, a] \times [0, a']$.

Теперь обобщим теорему 5.

Теорема 11. *Отношения конгруентности на L находятся во взаимно однозначном соответствии с нейтральными идеалами N в L, т. е. идеалами, содержащими вместе с любым элементом a все перспективные элементы.*

Доказательство. Предположим, что N определяет отношение конгруентности на L, и что $a \in N$ и b перспективны с осью перспективности c. Тогда

$$b = b \cap I = b \cap (a \cup c) \equiv b \cap c = 0 \pmod N).$$

Следовательно, $b \in N$ и идеал N нейтрален. Обратно, пусть идеал N нейтрален и пусть $x \equiv y$ означает, что $x \cup a = y \cup a$ для некоторого $a \in N$. Для любого идеала в любой структуре это является отношением эквивалентности и даже гомоморфизмом по объединениям. Мы хотим показать, что в дедекиндовой структуре с дополнениями из $x \equiv y$ следует $x \cap u \equiv y \cap u$, если идеал N нейтрален. Поскольку мы имеем дело с отношением эквивалентности и $x \equiv x \cup a = y \cup a \equiv y$, мы можем свести все к случаю $y = x \cup a$; следовательно, *достаточно показать, что в любой дедекиндовой структуре*

$$(x \cap u) \cup a \equiv [(x \cup a) \cap u] \cup a = (x \cup a) \cap (u \cup a).$$

Но при обозначениях $x = x_1$, $u = x_2$, $a = x_3$ это эквивалентно утверждению, что $c_3 \equiv b_3$ в свободной дедекиндовой структуре с тремя образующими. Однако при внимательном рассмотрении можно заметить, что $[b_3, c_3]$ проективно части x_3. Следовательно, в силу теоремы 4, $x_3 = a$ и некоторый элемент d, содержащий $c_3 \cap b'_3$, являются перспективными. Следовательно,

$$c_3 \cap b'_3 \in N \text{ и } b_3 \equiv b_3 \cup (b'_3 \cap c_3) = c_3,$$

чем и завершается доказательство.

6.9.10. Непрерывномерные проективные геометрии

Класс метрических дедекиндовых структур с дополнениями, функция размерности которых непрерывно изменяется между нулем и единицей, был открыт Дж. Нейманом.

Пусть F какое-нибудь поле или тело. Если a и a' взаимно дополнительные $(n — 1)$-мерные элементы из $PG(F; 2n —1)$, то, подструктуры X и Y элементов $x \leq a$ и $y \leq a'$ изоморфны с $PG(F; n—1)$;

следовательно, они изоморфны между собой. Типичный случай получим, принимая за a и a' скрещивающиеся прямые в проективном пространстве. Кроме того, поскольку a и a' независимы, объединения $x \cup y[x \in X, y \in Y]$ образуют в силу теоремы 7, р. 6.6, *подструктуру,* изоморфную $XX = X^2$ в $PG(F; 2n—1)$. Очевиден следующий вывод:

Лемма 1. *Пары $[x, x]$ из X^2 образуют подструктуру X^2, изоморфную X при изоморфизме, переводящем* 0 в 0, *I в I и увеличивающем структурную размерность* $d\ [x]$ *в два раза.*

(Заметим, что $d\ [x]$ превышает обычную геометрическую размерность элемента x на единицу.) Мы заключаем, что $PG\ (F;\ n — 1)$ может быть изоморфно вложено в $PG\ (F;\ 2n — 1)$ так, что «нормированная функция размерности» $d[x]\ /\ d\ [I]$ сохраняется. Повторяя этот процесс, мы получаем последовательность расширений $PG(F;\ 1)$:

$$PG(F;\ 1) \leqslant PG(F;\ 3) \leqslant PG(F;\ 7) \leqslant \ldots \leqslant PG(F;\ 2^n - 1) \leqslant \ldots,$$

в которой структурные операции, а также «нормированная функция размерности» $d\ [x]\ /\ d\ [I]$ сохраняются.

Беря объединение этих расширений, мы получаем объемлющую *метрическую структуру,* содержащую элементы любой мыслимой ДВОИЧНОЙ структурной размерности $k/2^n$. По теореме 14 из р. 6.6 она является метрически плотной подструктурой *полной* метрической структуры, которая будет поэтому содержать элементы любой размерности d, $0 \leq d \leq 1$. Эта последняя структура есть «непрерывная геометрия» $CG(F)$ Дж. Неймана, имеющая F в качестве основного поля.

Теорема 12. *Для любого тела F структура $CG(F)$ есть полная метрическая дедекиндова структура с дополнениями, в которой любые два элемента одинаковой размерности перспективны.*

Набросок доказательства. Покажем сперва, что метрическое пополнение $\overline{M}$ любой метрической дедекиндовой структуры M с дополнениями есть структура с дополнениями. В самом деле, для заданного $a \in \overline{M}$ мы можем найти последовательность $\{a_n\}$ в M такую, что $\delta\ (a_n,\ a) < 2^{n+1}$, откуда $\delta\ (a_n,\ a_{n+1}) < 2^n$. В силу теоремы 2 любому дополнению a'_n элемента a_n соответствует дополнение a'_{n+1} элемента a_{n+1}, для которого $\delta\ (a'_{n+1},\ a'_n) < 2^n$; отсюда следует, что a'_n есть сходящаяся последовательность, предел которой a' будет [в силу формул (7) —(7') р. 6.6.7] удовлетворять соотношениям $a \cap a' = 0$, $a \cup a' = 1$.

Доказательство того, что два любых элемента одинаковой размерности перспективны, является аналогичным; мы строим сходящуюся последовательность осей перспективности c_n для a_n, b_n. Как и в теореме

2, мы сводим все к случаю $a_{n+1}<a_n$ и ему двойственному. Предположим, что a, b перспективны в M посредством с и что $a > a_1$; $b > b_1$, причем $\delta(a, a_1) = \delta(b, b_1) = \delta_1$. Тогда, по предположению, $c \cup a_1$, $c \cup b_1$ будут иметь общее дополнение q с d $[q]=\delta_1$; поэтому a_1, b_1 будут перспективны посредством $c \cup q$, где $\delta (c \cup q, c) = \delta_1$.

Двойственно этому предположим, что $a_2>a$, $b_2>b$, причем $\delta (a, a_2) = \delta (b, b_2) = \delta_2$. Тогда $d[a_2 \cap c] = d[b_2 \cap c] = \delta_2$ и любое общее *относительное* дополнение r элементов $a_2 \cap c$ и $b_2 \cap c$ в с будет удовлетворять условию $\delta(c, r)=\delta_2$. Далее, элементы a_2 и b_2 будут перспективны посредством r.

6.9.11. Обратный результат: координаты в «регулярных кольцах»

Мы знаем в силу теоремы 15 из р. 6.6, что $CG (F)$ топологическая структура.

Дж. Нейман доказал, что, обратно, в любой полной, топологической, дедекиндовой структуре с дополнениями, центр которой состоит только из 0 и I, можно ввести функцию размерности, которая, если выполнена «нормировка» $d[0] = 0$, $d[I] = 1$, является единственной.

Чтобы сделать это, он определяет сперва $d[x] = d[y]$ как отношение, которое имеет место тогда и только тогда, если x и y перспективны. Затем он доказывает, что перспективность транзитивна, что и оправдывает употребление символа равенства. После этого он определяет $d[x] \leq d[y]$ в том смысле, что x перспективно части y. Затем он доказывает, что это отношение просто упорядочивает его абстрактные «элементы размерности». Вслед за этим он определяет сложение «элементов размерности», полагая, что $x \cap y =0$ влечет $d [x \cup y] =d [x] +d [y]$. Он показывает, что эта операция коммутативна и ассоциативна. Наконец, он показывает, что построенная так упорядоченная аддитивная система изоморфна или системе 0, $1/n$, $2/n$, ..., $n— 1/n$, 1 или континууму $0\leq x\leq 1$ (непрерывномерный случай). Это оправдывает название такой структуры в последнем случае «абстрактная непрерывная геометрия».

Мы видели, что любая проективная геометрия конечной длины $n > 3$ изоморфна структуре $PG (n — 1; D)$ всех подпространств n-мерного линейного пространства с координатами в подходящем теле D. Дж. Нейман показал, что любая абстрактная непрерывная геометрия может быть аналогично задана «координатами из кольца». Опишем эту аналогию для конечномерного случая.

Легко показать, что $PG(n—1; D)$ изоморфно структуре всех левых идеалов «простого» кольца всех $(n \times n)$-матриц над D. Кроме того, если R любое «полупростое» кольцо в обычном смысле, то структура $L(R)$ всех его левых идеалов будет все еще структурой с дополнениями и дедекиндовой. Обратно, этим путем можно получить все дедекиндовы структуры с дополнениями, удовлетворяющие теореме Дезарга; кардинальные произведения (прямые произведения) структур соответствуют прямым суммам колец. Таким образом, «координатизация с помощью кольца» Дж. Неймана имеет существенное преимущество перед «координатизацией с помощью поля» Штаудта в том, что она координатизирует *разложимые* (дезарговы) дедекиндовы структуры с дополнениями.

Линейная ассоциативная алгебра R с конечным базисом является «полупростой» в обычном смысле тогда и только тогда, если она «регулярна» в том смысле, что каждое $a \in R$ имеет «относительно обратный» элемент x такой, что

$$axa = a. \tag{6}$$

Кроме того, это условие применимо также к кольцам с бесконечным базисом или не имеющим базиса над центром, давая этим важное обобщение.

Можно показать, что два регулярных кольца изоморфны тогда и только тогда, если изоморфны дедекиндовы структуры с дополнениями их левых идеалов, и антиизоморфны тогда и только тогда, если последние дуально изоморфны. Точно так же двусторонние идеалы, которые в обычном «полупростом» случае соответствуют идемпотентам в центре кольца, образуют «центр» структуры левых идеалов.

Можно показать, что если фактор $I/0$ абстрактной непрерывной геометрии может быть разделен на четыре взаимно перспективные части, то могут быть введены координаты в регулярном кольце; добавленная гипотеза нужна для того, чтобы исключить недезаргов случай.

6.9.12. Атомные дедекиндовы структуры

Рассмотрим полную дедекиндову структуру $PG(R; d)$ всех векторных подпространств векторного пространства бесконечной размерности d над телом R. («Размерность» конечномерного или бесконечномерного векторного пространства V есть число элементов в базисе V.) Очевидно, что каждый ненулевой вектор ξ порождает минимальное собственное подпространство, множество всех скалярных кратных ξ. Следовательно, в структуре $PG(R;d)$ I является объединением «точек» (фактически, проективных точек), т. е. имеет место L7'. Далее, как и всегда для подалгебр, мы имеем непрерывность по пересечениям,

$$y_\alpha \uparrow y \text{ влечет } x \cap y_\alpha \uparrow x \cap y. \qquad (7)$$

Мы можем теперь привести некоторые теоретико-структурные выводы.

Теорема 13. *В полной дедекиндовой структуре M, удовлетворяющей условию* (7), L7' *влечет* L7, L7R *и* L7R'.

Доказательство. Предположим, что точки структуры M вполне упорядочены: $p_1, p_2, p_3, ..., p_\omega, ...$ Положим $a_\alpha = \vee_{\tau<\alpha} p_\tau$; тогда, очевидно, $a_1 = 0$, $a_{\alpha+1} = a_\alpha \cup p_\alpha$, и если λ —предельное порядковое число, то $a_\lambda = \vee_{\alpha<\lambda} a_\alpha$. Мы докажем сперва L7; в силу теоремы 1 оно влечет L7R. Определяем $y_1 = 0$ и y_α рекуррентно по трансфинитной индукции следующим образом: I) если $x \cup a_{\alpha-1} = x \cup a_\alpha$, то полагаем $y_\alpha = y_{\alpha-1}$; II) если $x \cup a_\alpha \neq x \cup a_{\alpha-1}$, то полагаем $y_\alpha = y_{\alpha-1} \cup p_\alpha$; III) если α —предельное порядковое число, то полагаем $y_\alpha = \bigvee\limits_{\lambda<\alpha} y_\lambda$.

Мы докажем по индукции, что $x \cap y_\alpha = 0$, $x \cup y_\alpha = x \cup a_\alpha$; поскольку в силу предположения L7' некоторое $a_\alpha = I$, это даст нам дополнение для x. В случае I $x \cap y_\alpha = x \cap y_{\alpha-1} = 0$, $x \cup y_\alpha = x \cup y_{\alpha-1} = x \cup a_{\alpha-1} = x \cup a_\alpha$ в силу предположения и индукции. В случае III $0 = x \cap y_\lambda$ для всех $\lambda < \alpha$ в силу индукции, но $x \cap y_\lambda \uparrow x \cap y_\alpha$ в силу предположения, что $y_\lambda \uparrow y_\alpha$, и (7); следовательно, $x \cap y_. = 0$. Аналогично, $x \cup a_\lambda = x \cup y_\lambda \uparrow x \cap y_\alpha$, но $a_\lambda \uparrow a_\alpha$; следовательно, $x \cup y_\alpha = x \cup a_\alpha$.. В случае II отметим, что $x, y_{\lambda-1}, p_\alpha$ *независимы* в смысле р. 6.6.5, так как $x \cap y_{\alpha-1} = 0$ по индукции и так как из $(x \cup y_{\alpha-1}) \cap p_\alpha = (x \cup a_{\alpha-1}) \cap p_\alpha$(по индукции) $\neq 0$ вытекало бы $p_\alpha \leq x \cup a_{\alpha-1}$ и $x \cup a_\alpha = x \cup a_{\alpha-1} \cup p_\alpha = x \cup a_{\alpha-1}$ в противоречии с предположением. Следовательно, $x \cap y_\alpha = x \cap (y_{\lambda-1} \cup p_\lambda) = 0$. Наконец, $x \cup y_\alpha = x \cup y_{\alpha-1} \cup p_\alpha = x \cup a_{\alpha-1} \cup$ $\cup\ p_\alpha$ *(по* индукции) $= x \cup a_\alpha$, чем и завершается доказательство L7 и, следовательно, L7R.

Остается доказать L7R'; мы покажем путем трансфинитной индукции, что каждый интервал [0, a] удовлетворяет L7R'; мы можем теперь предполагать справедливость L7 и L7R для всех $a_\alpha = \vee_{\tau<\alpha} p_\tau$ и L7R' для всех a_λ, предшествующих a_α.

Случай I: $\alpha = 1$. Он является тривиальным; равным образом для конечного α нам нужно лишь применить следствие из теоремы 1.

Случай II: α—1 существует. Пусть $x \le a_\alpha$ и $x' = x \cap a_{\alpha-1}$ и пусть y— любое относительное дополнение x' в $[0, a_{\alpha-1}]$, так что $x' \cap y = 0$, $x' \cup y = a_{\alpha-1}$. Мы будем писать для краткости p вместо $p_{\alpha-1}$.

П о д с л у ч а й IIa: $x \cap y > 0$. Обозначаем $x \cap y = q$; это есть точка, поскольку x покрывает x' и $x' \cap y = 0$. Следовательно, $x' < x' \cup q \le x$; но x покрывает x', откуда $x' \cup q = x$. Но x' есть объединение точек в силу индукции; следовательно, x также является объединением точек.

Подслучай IIб: $x \cap y = 0$. Если $x \le a_{\alpha-1}$, то x есть объединение точек в силу индукции по α. В противоположном случае $x \cup a_{\alpha-1} > a_{\alpha-1}$; но, так как $a_\alpha \geqslant x \cup y \geqslant x \cup x' \cup y = x \cup a_{\alpha-1} > a_{\alpha-1}$ и a_α самое большее покрывает $a_{\alpha-1}$, отсюда вытекает $x \cup y = a_\alpha$ и (так как $a_{\alpha-1} \cup p > a_{\alpha-1}$) $p \cap a_{\alpha-1} = 0$. Положим $r = (p \cup y) \cap x$. Если бы r было 0, то (так как $p \cap y \leqslant p \cap a_{\alpha-1} = 0$) p, y, x были бы независимыми, что невозможно, поскольку $(x \cup y) \cap p = a_\alpha \cap p > 0$. Следовательно, $r > 0$; но, так как $(p \cup y) \cap x$ самое большее покрывает $(0 \cup y) \cap x = 0$, r есть точка. Кроме того,

$$r \cup x' = [(p \cup y) \cap x] \cup x' = x \cap [p \cup y \cup x'] \text{ (в силу L5) } =$$
$$= x \cap (p \cup a_{\alpha-1}) = x.$$

Следовательно, x есть объединение точек (x' является таковым по индукции).

Случай III: α —предельное число. Каждое $x < a_\alpha$ удовлетворяет в силу (7) условию $x \cap a_\lambda \uparrow x \cap a_\alpha = x$. Но каждое $x \cap a_\lambda$ есть в силу индукции объединение точек. Беря объединение всех этих точек, мы выражаем $x = \sup(x \cap a_\lambda)$ как объединение точек. Этим завершается доказательство L7R'.

Определение. *Полная дедекиндова структура, удовлетворяющая условиям* (7) *и* L7' *(а следовательно, и условиям* L7, L7R, L7R'), *будет называться атомной.*

Следствие (Фринк). *Структура всех подпространств любого векторного пространства является полной атомной дедекиндовой структурой.*

6.9.13. Теоремы Фринка о разложениях

Теперь кратко изложим доказательства некоторых фундаментальных результатов, принадлежащих Фринку; подробности читатель может найти в работе Фринка.

Прежде всего, поскольку L7' и (7) выполняются, мы отмечаем, как непосредственное следствие теоремы 13, что плоскости в любой системе точек и прямых, удовлетворяющей условиям PG1 — PG2, образуют полную атомную дедекиндову структуру. Но и обратно, в точности как в р.6.9.3, точки и прямые любой полной атомной дедекиндовой структуры *M* удовлетворяют условиям PG1 —PG2. Кроме того, структура *M* изоморфна структуре всех «плоскостей» этой системы. Это обобщает теорему 3, устанавливая взаимно однозначное соответствие между системами, удовлетворяющими PG1 —PG2 и полными атомными дедекиндовыми структурами *M*.

Далее, как и в р.6.9.6, точки из *M* можно разбить единственным способом на неперекрывающиеся классы S_1, S_2, S_3, . . . взаимно перспективных точек. Плоскости в любом классе S_i взаимно перспективных точек удовлетворяют условию PG3 и образуют, как можно показать, *атомную проективную геометрию.* Можно показать, что *M* — подструктура прямого произведения (кардинального произведения) этих проективных геометрий. Отсюда вытекает

Теорема 14. *Любая полная атомная дедекиндова структура является подструктурой прямого произведения атомных проективных геометрий, и обратно.*

Далее, за исключением некоторых конечных проективных прямых и недезарговых проективных плоскостей, в любую атомную проективную геометрию могут быть введены координаты из тела *R* посредством «алгебры вурфов» Штаудта. Пусть теперь *d* — мощность максимального независимого подмножества в S_i (существование такого подмножества вытекает из того, что независимость есть свойство конечного характера). Фринк показал, что если *d* бесконечно, то атомная проективная геометрия есть просто *PG (R; d)* из р.6.9.12; конечномерный случай был уже рассмотрен. Так как обратное мы также рассмотрели в следствии из теоремы 13, то нами доказана

Теорема 15. *За исключением хорошо известных исключительных проективных прямых и недезарговых проективных плоскостей, каждая атомная проективная геометрия изоморфна структуре всех векторных подпространств векторного пространства V (R; d) подходящей размерности d над подходящим телом R.*

Пусть, наконец, *L любая* дедекиндова структура с дополнениями не обязательно атомная или полная. Напомним, что в силу теоремы 12 из р. 6.6 структура всех идеалов в *L* является дедекиндовой — следовательно, и структура всех дуальных идеалов в *L* дедекиндова. Заметим теперь, что свойство быть собственным дуальным идеалом в *L* эквивалентно свойству не содержать 0 и что это есть условие конечного характера. Следовательно, каждый собственный дуальный идеал может быть расширен до максимального собственного дуального идеала, который «покрывается» несобственным дуальным идеалом L. В частности, для любого $a \neq I$ из L главный дуальный идеал элементов $x \geq a'$ может быть расширен до максимального собственного идеала Q, который не может содержать a (в противном случае он содержал бы $a \cap a' = 0$). Следовательно, пересечение всех максимальных идеалов есть I, наименьший дуальный идеал. Вследствие этого мы заключаем, что структура M дуальных идеалов есть дедекиндова структура, в которой каждый элемент есть объединение точек, *при условии,* что M упорядочено отношением, *двойственным* включению.

Мы не можем, однако, заключить, что имеет место (7), хотя двойственное ему условие справедливо. Но если мы определим «точку» как максимальный дуальный идеал и «прямую» как пересечение двух различных максимальных идеалов, то легко доказать PG1 — PG2. «Плоскости» образуют полную атомную дедекиндову структуру.

Сопоставим теперь каждому элементу $a \in L$ множество $S(a)$ всех максимальных дуальных идеалов P таких, что $a \in P$. Легко показать, что $S(a)$ является всегда «плоскостью», что $a > b$ влечет, $S(a) > S(b)$ и что $S(a (\cap b)$ есть пересечение $S(a)$ и $S(b)$. Фринк показал далее, что $S(a \cup b)$ есть объединение $S(a)$ *и* $S(b)$; трудное место — показать, что если $a \cup b \in P$, то существуют Q, R, такие, что $a \in Q$, $b \in R$ и P содержит теоретико-множественное пересечение Q и R.

Поскольку «плоскости» образуют прямое произведение атомных проективных геометрий, мы получаем, в силу того, что было уже показано, следующий конечный результат.

Теорема 16. *Любая дедекиндова структура с дополнениями изоморфна подструктуре прямого произведения атомных проективных геометрий.*

6.10. Дистрибутивные структуры

6.10.1. Определение

Многие важные структуры удовлетворяют трем следующим тождествам:

L6. $(x \cap y) \cup (y \cap z) \cup (z \cap x) = (x \cup y) \cap (y \cup z) \cap (z \cup x)$,

L6′. $x \cap (y \cup z) = (x \cap y) \cup (x \cap z)$,

L6″. $x \cup (y \cap z) = (x \cup y) \cap (x \cup z)$.

Определение. *Структура будет называться «дистрибутивной» тогда и только тогда, если она тождественно удовлетворяет условиям* L6, L6', L6".

Теорема 1. *Каждое из тождеств* L6, L6', L6" *влечет* L5 *и оба остальных.*

Доказательство. L6' влечет L6", ибо путем непосредственных выкладок с использованием L6', L4 и L6' вместе с L2 имеем

$$(x \bigcup y) \bigcap (x \bigcup z) = [(x \bigcup y) \bigcap x] \bigcup [(x \bigcup y) \bigcap z] =$$
$$= x \bigcup [(x \bigcap z) \bigcup (y \bigcap z)] = [x \bigcup (x \bigcap z)] \bigcup (y \bigcap z) = x \bigcup (y \bigcap z).$$

Далее, L6" влечет L6. Действительно, путем непосредственного раскрытия скобок получаем

$$[(x \bigcap y) \bigcup (y \bigcap z)] \bigcup (z \bigcap x) =$$
$$= [(x \bigcap y) \bigcup (y \bigcap z) \bigcup z] \bigcap [(x \bigcap y) \bigcup (y \bigcap z) \bigcup x] =$$
$$\text{в силу } \mathbf{L3-L4} = [(x \bigcap y) \bigcup z] \bigcap [(y \bigcap z) \bigcup x] =$$
$$\text{снова в силу } \mathbf{L6''} = (x \bigcup z) \bigcap (y \bigcup z) \bigcap (y \bigcup x) \bigcap (z \bigcup x) =$$
$$\text{в силу } \mathbf{L1-L3} = (x \bigcup y) \bigcap (y \bigcup z) \bigcap (z \bigcup x),$$

чем доказано L6. Наконец, L6 влечет L5 и L6'. L5 мы получаем, полагая просто в L6 $x \geq z$, что дает нам слева $(x \bigcap y) \bigcup [(y \bigcap z) \bigcup z] = (x \bigcap y) \bigcup z$, а справа, двойственным образом, $(x \bigcup y) \bigcap (y \bigcup z) \bigcap x = x \bigcap (y \bigcup z)$. Затем, записывая L6 сокращенно в форме $u = v$, мы получаем из равенства $x \cap u = x \cap v$ с правой стороны в силу L4 и L2 — L3 $x \cap (y \cup z)$. С левой стороны мы получаем, используя L5,

$$x \bigcap ([y \bigcap z] \bigcup [(x \bigcap y) \bigcup (x \bigcap z)]) = (x \bigcap y \bigcap z) \bigcup (x \bigcap y) \bigcup (x \bigcap z).$$

Но в силу L3 —L4 это есть $(x \cap y) \cup (x \cap z)$, что и завершает доказательство.

Следствие. *Любая дистрибутивная структура является дедекиндовой.*

Из рис. 4 ясно, что недистрибутивная дедекиндова структура содержит подструктуру из 5 элементов c, d, e_1, e_2, e_3, изоморфную структуре, изображенной на рис. *1,в.* В сочетании с теоремой 2 р.6.4, мы выводим

Теорему 2. *Структура, не являющаяся дистрибутивной, содержит одну из изображенных на рис. 1,в—1,г структур в качестве подструктуры.*

Следствие 1. *Структура дистрибутивна тогда и только тогда, если относительные дополнения в ней являются, самое большее, однозначно определенными.*

Это означает, что для заданных $a \leq x \leq b$ существует самое большее один элемент y, удовлетворяющий соотношениям $x \cap y = a$ и $x \cup y = b$.

Доказательство. В дистрибутивной структуре из $x \cap u = x \cap v$ и $x \cup u = x \cup v$ следует $u = u \cap (x \cup v) =$
$= (u \cap x) \cup (u \cap v) = (v \cap x) \cup (u \cap v) = v \cap (u \cup x) = v.$
Обратно, на каждой из рис. 1,*в* и 1,*г* один элемент имеет два относительных дополнения.

Следствие 2. *Любая недистрибутивпая дедекиндова структура, удовлетворяющая одному из двух условий обрыва цепей, содержит структуру, изображенную на рис. 1,в, в качестве подструктуры, в которой x, u, v покрывают* 0.

Доказательство. Если имеет место условие обрыва убывающих цепей, то мы находим элемент $x^* \leq x$, который покрывает 0. Тогда $(x^* \cup u) \cap v$ и $(x^* \cap v) \cup u$ будут обеспечивать нужный пример; роль *е* будет играть $[(x^* \cup u) \cap v] \cup x^* =$
$= (x^* \cup u) \cap (v \cup x^*)$. Условия покрытия гарантируются в силу следствия 2 из теоремы 2 р.6.4. В случае обрыва возрастающих цепей можно поступить двойственным образом.

6.10.2. Примеры

Любая цепь есть дистрибутивная структура: для любой цепи каждая сторона в L6 есть средний из числа элементов *x, y, z*. Далее, подмножества любой совокупности *I* образуют дистрибутивную структуру и, более общо, любое кольцо множеств есть дистрибутивная структура. Так, открытые подмножества любого топологического пространства образуют (полную) дистрибутивную структуру; то же самое справедливо для замкнутых множеств.

Кроме того, любая подструктура или кардинальное произведение дистрибутивных структур есть снова дистрибутивная структура. Следовательно, дистрибутивной структурой является любая степень D^X любой дистрибутивной структуры *D* с частично упорядоченным множеством *X* в качестве показателя; в частности, 2^X есть дистрибутивная структура для любого частично упорядоченного множества *X*.

Поэтому структура натуральных целых чисел, упорядоченных по делимости, представляет собой дистрибутивную структуру: она является подструктурой прямого произведения счетного множества цепей (состоящих из степеней различных простых чисел 2, 3, 5, ...). Следовательно, то же можно сказать о (изоморфной) структуре идеалов в кольце всех целых алгебраических чисел любого конечного расширения поля рациональных чисел.

Нейтральные элементы (см. р.6.3.10) любой структуры образуют дистрибутивную подструктуру. Структура, двойственная дистрибутивной, является дистрибутивной. Булевы алгебры являются дистрибутивными структурами. Наконец, структурно упорядоченные группы и векторные структуры.

6.10.3. Альтернативные системы постулатов

Следуя идеям Ньюмена, Дж. Биркгоф развили очень сжатую систему постулатов для дистрибутивных структур.

Теорема 3. *Любая алгебраическия система, удовлетворяющая условиям*

$$a \cap a = a \text{ для всех } a, \tag{1}$$

$$a \cup I = I \cup a = I \text{ для некоторого } I \text{ и для всех } a, \tag{2}$$

$$a \cap I = I \cap a = a \text{ для некоторого } I \text{ и для всех } a, \tag{3}$$

$$a \cap (b \cup c) = (a \cap b) \cup (a \cap c) \text{ и } (b \cup c) \cap a = (b \cap a) \cup (c \cap a) \tag{4}$$

для всех a, b, c, есть дистрибутивная структура с единицей I.
Доказательство. Сперва мы показываем, что для всех *a, b*

$$a = a \cap I = a \cap (a \cup I) = (a \cap a) \cup (a \cap I) = a \cup a, \tag{5}$$

$$(a \cap b) \cup a = (a \cap b) \cup (a \cap I) = a \cap (b \cup I) = a \cap I = a \tag{6}$$

и аналогично

$$a \cup (a \cap b) = a \cup (b \cap a) = (b \cap a) \cup a = a. \tag{6'}$$

Теперь мы в состоявии доказать, используя (4), (1) и (6'),

$$a \cap (a \cup b) = (a \cap a) \cup (a \cap b) = a \cup (a \cap b) = a \tag{7}$$

и аналогично

$$a \cap (b \cup a) = (a \cup b) \cap a = (b \cup a) \cap a = a. \tag{7'}$$

Мы можем теперь доказать коммутативный закон

$$\begin{aligned} a \cup b &= [a \cap (b \cup a)] \cup [b \cap (b \cup a)] = (a \cup b) \cap (b \cup a) \\ &\qquad \text{в силу } (7)-(7'),\ (4), \\ &= [(a \cup b) \cap b] \cup [(a \cup b) \cap a] = b \cup a \text{ в силу } (4),\ (7)-(7'). \end{aligned} \tag{8}$$

Прежде чем доказать ассоциативный закон для объединении, мы показываем

$$\begin{aligned} &a \cap [(a \cup b) \cup c] = [a \cap (a \cup b)] \cup (a \cap c) = a \cup (a \cap c) = a \\ &\qquad \text{в силу } (4),\ (7),\ (6'), \\ &b \cap [(a \cup b) \cup c] = [b \cap (a \cup b)] \cup (b \cap c) = b \cup (b \cap c) = b \\ &\qquad \text{аналогичным образом}, \end{aligned} \tag{9}$$

$$c \cap [(a \cup b) \cup c] = [c \cap (a \cup b)] \cup (c \cap c) = [c \cap (a \cup b)] \cup c = c$$

в силу (4), (1), (6). Теперь мы доказываем ассоциативный закон для объединений,

$$a \cup (b \cup c) = \{a \cap [(a \cup b) \cup c]\} \cup (\{b \cap [(a \cup b) \cup c]\} \cup \cup \{c \cap [(a \cup b) \cup c]\}) = \text{ в силу } (9), \quad (10)$$

$= [a \cup (b \cup c)] \cap [(a \cup b) \cup c]$ в силу (4), примененного дважды, $=(a \cup b) \cup c$ аналогичным образом, в силу левой симметрии. Теперь мы доказываем соотношения, *двойственные* (4), а именно

$$(a \cup b) \cap (a \cup c) = [a \cap (a \cup c)] \cup [b \cap (a \cup c)] = \\ = a \cup [(b \cap a) \cup (b \cap c)] = \text{ в силу } (4), \quad (11). \\ = [a \cup (b \cap a)] \cup (b \cap c) = a \cup (b \cap c) \\ \text{в силу } (10),\ (6'),$$

$$(a \cap b) \cup c = (a \cup c) \cap (b \cup c) \text{ в силу левой симметрии.} \quad (11')$$

Мы доказали уже соотношение (5), двойственное (1), в то время как (6) и (7) являются двойственными. Но это были единственные законы, использованные при доказательстве (8), (10); следовательно, доказательства, в точности двойственные доказательствам (8), (10), приводят к коммутативному и ассоциативному законам для пересечений,

$$a \cap b = b \cap a \text{ и } a \cap (b \cap c) = (a \cap b) \cap c. \quad (12)$$

Этим завершается доказательство законов LI, L2, L3, L4, L6, используемых обычно в определении дистрибутивной структуры!

Мы рассмотрим теперь некоторые свойства двойственной себе тернарной операции, содержащейся в L6,

$$(a,\ b,\ c) = (a \cap b) \cup (b \cap c) \cup (c \cap a) = \\ = (a \cup b) \cap (b \cup c) \cap (c \cup a). \quad (13)$$

Мы называем *(a, b, c) медианой* элементов *a, b, c,* ибо *(a, b, c)* сводится к таковой в случае цепей.

Лемма 1. *В любой дистрибутивной структуре (a, x, b) = x тогда и только тогда, если*

$$a \cap b \leqslant x \leqslant a \cup b.$$

Доказательство. Если $a \cap b \leqslant x \leqslant a \cup b$, то, очевидно,

$$(a \cap b) \cup (b \cap x) \cup (x \cap a) = (a \cap b) \cup [x \cap (b \cup a)] = x.$$

Обратное может быть доказано читателем. В действительности, если мы определим вместе с Даси *отрезок* $<a, b>$, соединяющий *a* и *b* как множество всех *x,* удовлетворяющих соотношению $a \cap b \leq x \leq a \cup b$, то имеет место

Лемма 2. Элемент *(a, b, c)* есть пересечение трех множеств $<a, b>$, $<b, c>$, $<c, a>$.

Доказательство. В силу леммы 1 *(a, b, c)* принадлежит всем трем множествам. Обратно, если x принадлежит всем трем множествам, то, очевидно,

$$(a, b, c) = (a \cap b) \cup (b \cap c) \cup (c \cap a) \leqslant x \leqslant$$
$$\leqslant (a \cup b) \cap (b \cup c) \cap (c \cup a) = (a, b, c).$$

Можно определить дистрибутивную структуру с 0 и I в терминах одной только тернарной операции (a, b, c).

Теорема 4. *Пусть A—любая алгебраическая система с тернарной операцией (a, b, c) и элементами* 0, *I, такими, что*

$$(0, a, I) = a, \tag{14}$$

$$(a, b, a) = a, \tag{15}$$

$$(a, b, c) = (b, a, c) = (b, c, a) \text{ (симметрия)}, \tag{16}$$

$$((a, b, c), d, e) = ((a, d, e), b, (c, b, e)) \tag{17}$$

выполнено тождественно. Тогда, если мы определим

$$a \cup b = (a, I, b) \text{ и } a \cap b = (a, 0, b), \tag{18}$$

то A будет дистрибутивной структурой, в которой имеет место (13).

Доказательство. Докажем сперва, используя теорему 3, что A дистрибутивная структура. В самом деле, из (15) и (18) вытекает (1); из (15) и (18) вместе с (16) вытекает (2); из (14) и (18) вместе с (16) вытекает (3). В то же время (17) влечет второе тождество из (4), если мы положим $b = I$ и $d = 0$; первое вытекает из соображений симметрии. [Заметим, что из условий в (16) вытекает инвариантность выражения (a, b, c) при *всех* перестановках его членов.]

Используем теперь лемму 2, чтобы показать, что в этой дистрибутивной структуре *(a, b, c)* имеет свое обычное значение. В самом деле, используя несколько раз (16) вместе с (17) и (15), имеем

$$(a, (a, b, c), b) = ((a, c, b), a, b) = ((a, a, b), c, (b, a, b)) =$$
$$= (a, c, b) = (a, b, c).$$

Аналогично, $(a, (a, b, c), c) = (b, (a, b, c), c) = (a, b, c)$, откуда в силу леммы 1 (a, b, c) является пересечением $<a, b>$, $<b, c>$ и $<a, c>$, а потому имеет свое обычное значение.

6.10.4. Теория представлений: конечный случай

Мы отмечали, что любое кольцо множеств есть дистрибутивная структура: в п.6.10.4 — 6.10.5 докажем обратный этому результат. Обратимся сперва к случаю конечной длины.

Лемма 1. *Если в дистрибутивной структуре L элемент p неразложим в объединение, то* $p \leq \vee_{i=1}^{k} x_i$ *влечет* $p \leq x_i$ *для некоторого* i.

Доказательство. Из предположения следует

$$p = p \bigcap \vee_{i=1}^{k} x_i = \vee_{i=1}^{k} (p \bigcap x_i).$$

Следовательно, так как p неразложимо в объединение, $p = p \cap x_i$ для некоторого i, что равносильно $p \leq x_i$ для некоторого i.

Лемма 2. *Если L содержит n неразложимых в объединение элементов* $p_1, \ldots, p_n$, *то* $d[L] \geq n$.

Доказательство. Перенумеруем p_i так, чтобы из $p_i < p_j$ следовало $i < j$; это возможно, поскольку частичная упорядоченность не является цикличной. Тогда цепь $0 < p_1 < p_1 \cup p_2 < \ldots \vee_{i=1}^{n} p_i$ имеет в силу леммы 1 длину n.

Теорема 5. *Пусть L любая дистрибутивная структура конечной длины n. Тогда частично упорядоченное подмножество X неразложимых в объединение элементов* $p_i > 0$ *имеет порядок n,* в $L = 2^{\breve{X}}$.

Доказательство. В силу конечной индукции каждое a в L является объединением $\vee_A p_i$ множества A неразложимых в объединение элементов $p_i > 0$, которые оно содержит. При этом если $p_i \leq a$ и $p_j \leq p_i$, то $p_j \leq a$, т. е. каждое множество $A \xi$ M-замкнуто в X. Но обратно в силу леммы 1 если A M-замкнуто, то $\vee_A p_i$ не содержит ни одного p_k, не принадлежащего A. Следовательно, соответствие $a \leftrightarrow A$, которое изотонно, является взаимно однозначным соответствием, а потому есть изоморфизм. Но система множеств A изоморфна кольцу «замкнутых» подмножеств T_0-пространства, определяемого множеством $\breve{X}$, двойственным X, т. е. изоморфна $2^{\breve{X}}$. Но $d[2^{\breve{X}}]$ равно порядку $\breve{X}$, чем и завершается доказательство.

Следствие. *Число (неизоморфных) дистрибутивных структур длины n равно числу частично упорядоченных множеств из n элементов.*

6.10.5. Общая теорема о представлении

Доказжем следующий общий результат.

Теорема 6. *Любая дистрибутивная структура L (за исключением* 1) *изоморфна полупрямому произведению структур, изоморфных одной и той же структуре* 2.

Доказательство. В силу теоремы 10 р.6.7 L является полупрямым произведением неразложимых в полупрямое произведенне дистрибутивных структур. Следовательно, достаточно показать, что имеет место

Лемма. *Единственная неразложимая в полупрямое произведение дистрибутивная структура D есть (за исключением* 1) *структура* 2.

Доказательство. Предположим, что D содержит элемент x, отличный от 0 и I. Тогда существовал бы элемент $r \in D$, не содержащийся в ж; следовательно, существовало бы $s = x \cup r > x$; двойственно этому, в D существовало бы $t < x$. Следовательно, эндоморфизмы в D, $u \to u \cap x$ и $u \to u \cup x$ отождествляли бы различные элементы. Но, как и в следствии 1 теоремы 2, из $u \cap x = v \cap x$ и $u \cup x = v \cup x$ следует $u = v$; следовательно, соответствие $d \to (d \cap x, d \cup x)$ давало бы полупрямое разложение D. Мы заключаем, что D не может содержать элемента, отличного от 0 и I, откуда, очевидно, и вытекает лемма.

Следствие. *Любая дистрибутивная структура L изоморфна кольцу множеств.*

Доказательство. Случай L=l тривиален. Если $L \neq 1$, то L представимо как полупрямое произведение структур, изоморфных 2, каждую из которых мы будем называть *точкой* и обозначать через p_α. Каждому $a \in L$ мы сопоставим функцию f такую, что $p(p_\alpha)$=l, если p_α -компонента элемента a есть I, и $f(p_\alpha)=0$ в противном случае. Это есть характеристическая функция множества A точек p_α, таких, что p_α-компонента элемента a есть I. Поскольку A определяет все p_α-компоненты элемента a, соответствие является взаимно однозначным. Кроме того, объединение и пересечение в L соответствуют, очевидно, для каждой p_α-компоненты, объединению и пересечению для соответствующих множеств; это хорошо известное исчисление характеристических функций. Следовательно, мы установили изоморфизм, чем и завершается доказательство.

6.10.6. Идеалы

Мы видели уже, что идеалы любой структуры L сами по себе образуют структуру. Кроме того, можно показать, что в этой структуре $J \cap K$ есть просто множество всех $s \cap t$, где $s \in J$, $t \in K$. Если структура L *дистрибутивна,* то верно также, что $J \cup K$ есть множество всех $s \cup t[s \in J, \ t \in K]$. Действительно, пересечение всех идеалов в L, содержащих J и K, содержиткаждое такое $s \cup t$. Но, обратно, поскольку $(s \cup t) \cup (s_1 \cup t_1) = (s \cup s_1) \cup (t \cup t_1)$, в то время как $x \leq s \cup t$ влечет $x = x \cap (s \cup t) = (x \cap s) \cup (x \cap t) = s_1 \cup t_1 \, [s_1 \in J, \ t_1 \in K]$,
множество всех $s \cup t[s \in J, t \in K]$ является идеалом, содержащим J и K.

Резюмируем: если L—дистрибутивная структура, то операции I и U можно толковать в смысле исчисления комплексов.

Это — понятие универсальной алгебры, впервые определенное Фробепиусом для групп. Если $S_1, \ldots, S_n$ любые непустые подмножества алгебры с n-арной операцией f_i, то $f_i \ (S_1, \ldots, S_n)$ есть множество всех $x = f_i \ (s_1, \ldots, s_n)$, где $s_1 \in S_1, \ldots, s_n \in S_n$.

Мы видели в р. 6.6, теорема 12, что идеалы любой дедекиндовой структуры сами образуют дедекиндову структуру. Подобно этому мы теперь докажем

Теорему 7. *Идеалы любой дистрибутивной структуры сами образуют дистрибутивную структуру относительно теоретико-множественного включения (или, что изоморфно, относительно исчисления комплексов).*

Доказательство. В силу одностороннего дистрибутивного закона (3) из р.6.3.4, $H \cap (J \cup K) \geqslant (H \cap J) \cup (H \cap K)$.

Но каждый элемент $s \cap (t \cup u) \; [s \in H, \ t \in J, \ u \in K]$ из $H \cap (J \cup K)$ есть также элемент $(s \cap t) \cup (s \cap u)$ из $(H \cap J) \cup [H \cap K)$; следовательно, имеет место обратное неравенство.

Мы также определили уже главные идеалы и замкнутые идеалы. Главные идеалы любой структуры L тривиальным образом образуют подструктуру структуры всех идеалов в L, изоморфную с самим L. С другой стороны, замкнутые идеалы в L не образуют подструктуру. В самом деле, Котлар, Фунаяма и Дилуорс дали независимо друг от друга примеры дистрибутивных структур, пополнения которых при помощи сечений даже не являются дедекиндовыми.

Мы определим теперь новый тип идеала: простые идеалы. Это есть точный аналог понятия простого идеала в теории колец, если мы рассматриваем $\cap$ в смысле умножения.

Определение. *Идеал P структуры L является простым тогда и только тогда, если*

$$x \cap y \in P \text{ влечет } x \in P \quad \text{или} \quad y \in P. \tag{19}$$

Теорема 8. *Каждое из следующих условий является необходимым и достаточным для того, чтобы идеал J структуры L был простым:* а) *дополнение L — J идеала J есть дуальный идеал, б) J есть множество прообразов* 0 *при структурном гомоморфизме* $L \to 2$.

Доказательство. В любом случае $x \in J$, $y \in J$, $u \in L — J$, $v \in L — J$ влечет: $x \cap y$, $x \cup y$ и $x \cap u$ принадлежат J, в то время как $x \cup u$ и $u \cup v$ принадлежат $L — J$. Единственный вопрос — как обстоит дело с элементом $u \cap y$. Утверждение $u \cap v \in L — J$ эквивалентно (19), поскольку оно исключает возможность $u \cap v \in J$, $u \notin J$ $v \notin J$, а в силу определений оно эквивалентно «а» и «б». (Случаи, когда P пусто и когда $P=L$, являются исключительными в предшествующих рассмотрениях.) Отсюда следует, что «точки» в любом представлении (дистрибутивной) структуры L кольцом множеств соответствуют однозначным образом «простым идеалам» в L и что две точки соответствуют одному и тому же простому идеалу тогда и только тогда, если одна является сократимой. Кроме того, «точное» представление, в котором каждый простой идеал встречается один и только один раз, играет особую роль.

6.10.7. Теорема единственности разложения

Пусть L—любая дистрибутивная структура и пусть

$$a = x_1 \cup \ldots \cup x_r = y_1 \cup \ldots \cup y_s$$

любые два представления элемента $a \in L$ в виде объединения неразложимых в объединение элементов. Тогда в силу леммы 1 из р.6.10.4 для заданного x_i некоторые $y_j \geq x_i$ и аналогично некоторое $x_{k\geq} > y_j$. Следовательно, если только x_i не является *сократимым* в том смысле, что $x = x_1 \cup \ldots \cup x_{i-1} \cup x_{i+1} \cup \ldots \cup x_r$, мы имеем $x_i = y_j = x_k$, откуда $k=i$. Таким образом, если разложения не сократимы, то элементы x_i и y_j попарно равны, $r = s$, и отсюда следует

Лемма 1. *В дистрибутивной структуре представление элемента в виде несократимого объединения неразложимых в объединение элементов является единственным (и двойственно этому).*

Но если имеет место условие обрыва убывающих цепей, то легко доказать существование такого представления, откуда вытекает

Теорема 9. *В дистрибутивной структуре L, удовлетворяющей условию обрыва убывающих цепей, каждый элемент имеет одно и* *только одно представление в виде несократимого объединения неразложимых в объединение элементов. (И двойственно этому, если L удовлетворяет условию обрыва возрастающих цепей.)*

Пусть теперь *P* — частично упорядоченное множество всех неразложимых в объединение элементов из L; очевидно, что *P* само удовлетворяет условию обрыва убывающих цепей. Мы определим «крону» в *P* как конечное подмножество $X = (x_1, ..., x_r)$ в *P* такое, что $x_i \leq x_j$ в *X* влечет $x_i = x_j$. Тогда в силу теоремы 9 имеет место взаимно однозначное соответствие между «кронами» и *P* и элементами в L, даваемое посредством $(x_1, ..., x_r) \leftrightarrow x_1 \cup ... \ ... \cup x_r$. Кроме того, $x_1 \cup \ldots \cup x_r \leqslant y_1 \cup \ldots \cup y_s$ тогда и только тогда, если каждое x_i содержится в некотором y_j.

Обратно, если *P* — абстрактное частично упорядоченное множество, удовлетворяющее условию обрыва убывающих цепей, то конечные «кроны» в *P* образуют дистрибутивную структуру *L*, удовлетворяющую условию обрыва убывающих цепей, если мы определим $(x_1, ..., x_r) \leq (y_1, . .., y_s)$ в том смысле, что каждое x_i содержится в некотором y_j. Кроме того, неразложимые в объединение элементы из *L* суть одноэлементные кроны. Можно, таким образом, доказать следующий результат, обобщающий теорему 5.

Теорема 10. *Имеет место взаимно однозначное соответствие между дистрибутивными структурами L, удовлетворяющими условию обрыва убывающих цепей, и частично упорядоченными множествами P, удовлетворяющими условию обрыва убывающих цепей. При этом соответствии P изоморфно подмножеству неразложимых в объединение элементов из L (и двойственно этому).*

6.10.8. Приложения к алгебре и алгебраической геометрии

Комбинируя теорему 9, приведенную выше, с теоремой 9 из р.. 6.7, мы получаем непосредственно следующий результат.

Теорема 11. *Пусть A—любая абстрактная алгебра, отношения конгруентности которой образуют дистрибутивную структуру, удовлетворяющую условию обрыва возрастающих цепей. Тогда A и ее гомоморфные образы обладают единственными представлениями в виде полупрямого произведения неразложимых в полупрямое произведение сомножителей.*

Этот результат имеет место, если *A* — любая конечная структура, полупростая группа, группа порядка, не делящегося на квадраты чисел, или полупростая гиперкомплексная алгебра.

Более интересное приложение мы имеем к алгебраической геометрии. *Алгебраическое многообразие* V в аффинном n-пространстве над полем F определяется как множество всех точек $(x_1, \ldots , x_n)$, удовлетворяющих подходящей системе $J(V)$ алгебраических уравнений $p_i\ (x_1, \ldots , x_n) = 0$ с коэффициентами в F. Если V и W алгебраические многообразия, определенные соответственно r уравнениями $p_i = 0$ и s уравнениями $q_i = 0$, то легко показать, что rs уравнений $p_i q_i = 0$ удовлетворяются для точек из V или W и ни для каких других, в то время как $r+s$ уравнений $p_i = 0$, $q_i = 0$ удовлетворяются для точек из V и W и ни для каких других. Следовательно, алгебраические многообразия образуют кольцо множеств (дистрибутивную структуру).

Легко показать, кроме того, что идеалы кольца полиномов $F\ (x_1, \ldots , x_n)$ удовлетворяют условию обрыва возрастающих цепей; кроме того, соответствие между многообразиями и уравнениями, которым они удовлетворяют, есть полярность. Следовательно, как и в р. 6.5, алгебраические многообразия удовлетворяют условию обрыва убывающих цепей. Мы заключаем, что здесь применимы предположения теоремы 9, откуда *каждое алгебраическое многообразие имеет единственное выражение в виде несократимой суммы конечного числа неразложимых компонент.*

Наконец, пусть L — любая дедекиндова структура, удовлетворяющая условию обрыва убывающих цепей. Если мы будем отождествлять элементы x и y из L всякий раз, когда интервал $[x \cap y,\ x \cup y]$ имеет конечную длину, то мы получим собственный гомоморфный образ L_1 структуры L. Повторяя этот процесс, мы получим последовательность гомоморфных образов L f L_1 f L_2 f ... , которая впервые была абстрактно рассмотрена Орэ; пусть $0 < \theta < \theta_2 < \ldots$ обозначает соответствующую последовательность отношений конгруентности на L.

В случае алгебраических многообразий можно показать, что *размерность* многообразия V есть наименьшее k, для которого $V \equiv \mathrm{mod}\,\theta_k$. Двойственно этому если $J(V)$ идеал алгебраических уравнений, которые удовлетворяются для всех $(x_1, \ldots , x_n) \in V$, и $F\ (x_1, \ldots , x_n)$ кольцо всех полиномов относительно $x_1, \ldots , x_n$ с коэффициентами в F, то k есть *степень трансцендентности* фактор-кольца (кольца классов вычетов) $F\ (x_1, \ldots , x_n) / J(V)$.

6.10. 9. Общая конечная дистрибутивность

Доказательство по индукции обычного обобщенного дистрибутивного закона в обычной алгебре применимо и к дистрибутивным структурам при условии, если суммы понимаются как объединения, а произведения как пересечения, или двойственно этому. Например, применяя индукцию по *n,* мы имеем

$$x \bigcap \bigvee_{j=1}^{n} y_j = x \bigcap \left(\bigvee_{j=1}^{n-1} y_j \bigcup y_n\right) = \left(x \bigcap \bigvee_{j=1}^{n-1} y_j\right) \bigcup (x \bigcap y_n) =$$

$$= \bigvee_{j=1}^{n-1} (x \bigcap y_j) \bigcup (x \bigcap y_n) = \bigvee_{j=1}^{n} (x \bigcap y_j) \text{ и двойственно.} \quad (20)$$

Аналогично, индукцией по *m* мы получаем

$$\bigvee_{i=1}^{m} x_i \bigcap \bigvee_{j=1}^{n} y_j = \bigvee_{i,\, j} (x_i \bigcap y_j) \quad \text{и двойственно.} \quad (21)$$

Наконец, применяя индукцию по числу *r* членов, мы можем вывести *обобщенный* дистрибутивный закон

$$\bigwedge_{h=1}^{r} \left[\bigvee_{h=1}^{n(h)} u_{h,\,k}\right] = \bigvee_{f} \left[\bigwedge_{h=1}^{r} u_{h,\,f(h)}\right] \quad \text{и двойственно.} \quad (22)$$

Здесь *f* пробегает все функции, сопоставляющие каждому *h* определенное $k=1, \ldots, n(h)$.

Теорема 12. *Любой структурный полином* φ *относительно элементов* $x_1, \ldots, x_m$ *дистрибутивной структуры* L *может быть записан в форме*

$$\bigwedge_{h=1}^{r} \left[\bigvee_{k=1}^{n(h)} x_{i(h,\,k)}\right],$$

а также двойственно.

Доказательство. Используя обобщенный дистрибутивный закон, мы можем заменить $\wedge\ \vee$ на $\vee\ \wedge$ и обратно. Следовательно, мы можем заменить любое $\wedge\ \vee\ \wedge\ \vee\ \ldots$ на $\wedge\ \wedge\ \vee\ \vee\ \ldots$, а это, в свою очередь, используя ассоциативный закон, на $\wedge\vee$. Окончание доказательства теперь очевидно в силу индуктивных соображений.

6.10.10. Свободные дистрибутивные структуры

Мы определим теперь свободную дистрибутивную структуру *FD(n),* порожденную *n* символами $x_1, \ldots, x_n$. Кроме того, мы присоединим 0 и *I*.

Во-первых, в силу теоремы 12 мы можем записать каждый элемент из *FD(n)* в виде $\vee_{\sigma\in F}\ (\wedge_{i\in\sigma}\ x_i)$, где σ обозначает множества элементов x_i, а *F* обозначает систему множеств σ. Кроме того, в силу LI — L3, поскольку пересечения и объединения определяются *множествами* элементов, из которых они образованы, каждый элемент вполне определяется написанным выражением. Далее, если σ*≥ σ, то

$$\bigwedge_{i\in\sigma} x_i \bigcup \bigwedge_{j\in\sigma^*} x_j = \bigwedge_{i\in\sigma} x_i,$$

следовательно, каждый элемент из *FD(n)* может быть записан в форме $\bigvee_{\sigma\in F^*}(\bigwedge_{i\in\sigma} x_i)$, где *F** содержит вместе с каждым подмножеством σ все σ*≥ σ. Мы будем называть такое семейство *J-замкнутым.*

Этим устанавливается однозначное в одну сторону соответствие (которое сохраняет порядок) между множеством всех *F** и элементами из *FD(n).* Мы введем *I*, соответствующее случаю, когда *F** пусто, и 0 для случая, когда *F** содержит пустое множество σ_0 и никаких других. Следовательно, соответствие должно быть изоморфизмом, если при некоторой реализации различные *F** соответствуют различным элементам. Такую реализацию мы теперь установим.

В самом деле, пусть $L = 2^n$ семейство всех подмножеств с элементов $x_1, . . . , x_n$ и пусть X_i обозначает подмножество в *L,* которое содержит σ тогда и только тогда, если $x_i \in \sigma$. Тогда X_i будет *J*-замкнутым; следовательно, это будет верно для всех подмножеств кольца, порожденного X_i. Кроме того, $\wedge_{i\in\tau} X_i$ будет содержать σ тогда и только тогда, если каждое $x_i[i \in \tau]$ принадлежит σ, т. е. тогда и только тогда, если σ ≥τ. Следовательно, $\bigvee_{\tau\in G}(\bigwedge_{i\in\tau} X_i)$ содержит σ тогда и только тогда, если σ ≥τ для некоторого τ, — следовательно (если *G* является *J*-замкнутым) тогда и только тогда, если $\sigma \in G$. Таким образом, различные *J*-замкнутые *G* определяют различные системы множеств σ, и соответствие есть изоморфизм. Этим доказана

Теорема 13. *Свободная дистрибутивная структура, порожденная n символами, с присоединенными 0 и I, изоморфна кольцу всех J-замкнутых подмножеств структуры* 2^n *всех подмножеств множества n точек.*

Поскольку 2^n двойственно себе, мы выводим непосредственно в силу теоремы 8 из р. 6.2

Следствие. *Свободная дистрибутивная структура, порожденная n символами, с присоединенными* 0 *и I, есть* 2^{2^n}.

6.10.11. Бесконечная дистрибутивность

Естественными обобщениями дистрибутивных законов (20) — (22) из р.6.10.9 являются соотношения

$$x \bigcap \bigvee_B y_\beta = \bigvee_B (x \bigcap y_\beta) \text{ и двойственно,} \tag{20'}$$

$$\bigvee_A x_\alpha \bigcap \bigvee_B y_\beta = \bigvee_{A,B} (x_\alpha \bigcap y_\beta) \text{ и двойственно} \tag{21'}$$

и

$$\bigwedge_C [\bigvee_{A_\gamma} u_{\gamma,\alpha}] = \bigvee_F [\bigwedge_C u_{\gamma,\varphi(\gamma)}] \text{ и двойственно.} \tag{22'}$$

В (22') F обозначает класс всех однозначных функций φ, относящих каждому $\gamma \in C$ значение $\varphi(\gamma) \in A_\gamma$. Эти «бесконечные дистрибутивные законы» справедливы *не для каждой* дистрибутивной структуры. Однако их справедливость связана с ранее введенными понятиями.

Теорема 14. *Во всякой полной дистрибутивной структуре L тождество* (20') *эквивалентно тождеству* (21') *и условию, что L является топологической структурой*.

Доказательство. Используя принцип двойственности, мы уменьшаем процесс доказательства наполовину. Используя (20') дважды, мы получаем

$$\bigvee_A x_\alpha \bigcap \bigvee_B y_\beta = \bigvee_A (x_\alpha \bigcap \bigvee_B y_\beta) =$$
$$= \bigvee_A (\bigvee_B (x_\alpha \bigcap y_\beta)) = \bigvee_{A,B} (x_\alpha \bigcap y_\beta),$$

где последнее соотношение вытекает из обобщенного ассоциативного закона. Следовательно, (20') влечет (21'); обратное очевидно. Далее, в силу леммы р. 6.5.9, для того, чтобы L было топологической структурой, достаточно, чтобы $x_\alpha \uparrow x$ следовало $a \cap x_\alpha \uparrow a \cap x$ и двойственно, а это есть непосредственное следствие тождества (20'). Обратно, пусть в полной топологической структуре Г есть множество конечных подмножеств G множества B. Тогда $z_G = \bigvee_G y_\beta \to \bigvee_B y_\beta$. Следовательно, $x \bigcap z_G \to x \bigcap \bigvee_B y_\beta$. Но для каждого G мы имеем в силу (20) $x \bigcap z_G = \bigvee_G (x \bigcap y_\beta) = u_G$. В силу непрерывности $u_G \to \bigvee_B (x \bigcap y_\beta)$. Мы заключаем, что оба предела суть одно и то же.

6.10.12. Структуры с псевдодополнениями

Топологические дистрибутивные структуры и многие другие являются структурами с относительными псевдодополнениями в следующем смысле.

Определение. *Под псевдодополнением a*b элемента a относительно элемента b в структуре L понимают такой элемент c, что $a \cap x \leq b$ тогда и только тогда, если $x \leq c$. Структура, в которой a*b существует для всех a и b, называется структурой с относительными псевдодополнениями. Элемент a**0 *называется псевдодополнением* *элемента a и обозначается через* a^*.Таким образом, a^* есть попросту наибольший элемент, имеющий 0 в пересечении с a, в случае, если таковой существует.

Теорема 15. *Полная структура L является структурой с относительными псевдодополнениями тогда и только тогда, если она удовлетворяет первому тождеству из* (20').

Доказательство. Предположим, что L удовлетворяет первому тождеству из (20'); пусть заданы a и b. Пусть далее B обозначает множество всех y_β, удовлетворяющих соотношению $a \cap y_\beta \leq b$. Тогда $a \vee_{By_\beta} = \vee_B (a \cap y_\beta) \leq b$; следовательно, элемент $\vee_{By_\beta} = c$ удовлетворяет соотношению $a \cap c \leq b$, и $a \cap y_\beta \leq b$ влечет $y_\beta \leq c$; следовательно, $a * b$ существует. Обратно, предположим, что L — структура с тносительными псевдодополнениями, и пусть $b = \vee_B (a \cap y_\beta)$. Очевидно, что $a \cap y_\beta \leq b$ для каждого y_β, следовательно, каждое $y_\beta \leq a*b$ и $\vee_{By_\beta} \leq a* b$. Подставляя в тождественное отношение $a \cap (a* b) \leq b$, мы получаем $a \cap \vee_{By_\beta} \leq b = \vee_B (a \cap y_\beta)$; но обратное неравенство имеет место в любой полной структуре, чем и завершается доказательство.

Следствие. *Любая структура* с *относительными псевдодополнениями является дистрибутивной и любая топологическая дистрибутивная структура является структурой с относительными псевдодополнениями.*

Так как подалгебры любой абстрактной алгебры удовлетворяют первому тождеству из (20'), т. е. имеют непрерывную операцию объединения, идеалы любой дистрибутивной структуры образуют структуру с относительными псевдодополнениями. Аналогично, открытые подмножества любого топологического пространства образуют структуру с относительными псевдодополнениями. Далее), логики Брауэра тесно связаны со структурами с относительными псевдодополнениями.

Соответствие $a \to a^*$ является связью Галуа в любой структуре с относительными псевдодополнениями L. Поскольку также отношение $x \cap y = 0$ симметрично, мы заключаем в силу р. 6.5. 6,

$$a \leqslant a^{**},\ a^* = a^{***},\ a \leqslant b \text{ влечет } a^* \geqslant b^*,$$
$$(a \cup b)^* = a^* \cap b^* \text{ и } (a \cap b)^* \geqslant a^* \cup b^*. \tag{23}$$

Далее, «замкнутые» элементы из L, удовлетворяющие соотношению $a = a^{**}$, образуют полную структуру A, в которой объединения задаются новой операцией $a \vee b$ $(a \cup b)^{**}$, в то время как операция пересечения та же самая, что и в L. Поскольку

$$(a \cup a^*)^* = a^* \cap a^{**} = 0,\ a \vee a^* = 0^* = I$$

и

$$a \cap a^* = 0 \text{ и } a \vee a^* = I. \qquad (23')$$

Предположим теперь, что $x \cap a \cap b = 0$ в L; рассмотрим элемент $y = x \cap a^{**} \cap b^{**} \leq x$. Очевидно, что $y \cap a \cap b = 0$; но отсюда следует: $y \cap a \leq b^*$; так как $y \cap a \leq b^{**}$, мы заключаем, что $y \cap a = 0$. Аналогично, из $y \cap a = 0$ мы заключаем, что $y \leq a^* \cap a^{**} = 0$. Таким образом, $x \cap a \cap b = 0$ влечет $x \cap a^{**} \cap b^{**} = 0$, или $(a \cap b)^* \leqslant (a^{**} \cap b^{**})^*$. Но обратное неравенство очевидно в силу (23.). Используя снова (23) в получаемом равенстве, мы имеем

$$(a \cap b)^* = (a^{**} \cap b^{**})^* = ((a^* \cup b^*)^*)^* = a^* \vee b^*. \qquad (23'')$$

В силу (23) — (23") соответствие $a \to a^{**}$ есть гомоморфизм L на A. Таким образом, A является дистрибутивной структурой с дополнениями. Мы опишем теперь соответствующее отношение конгруентности на L.

Пусть D множество (дуальный идеал) всех $d \in L$, удовлетворяющих условию $d^* = 0$; такие элементы мы будем называть «плотными». Очевидно, что условие $d \in D$ эквивалентно условию $d = a \cup a^*$ для некоторого a и условию $d^{**} = I$. Мы установим теперь конечный результат.

Теорема 16 (Гливенко)). *В любой дистрибутивной структуре с псевдодополнениями L соответствие $a \to a^{**}$ представляет собой операцию замыкания в L и структурный гомоморфизм L на полную булеву алгебру «замкнутых» элементов. Кроме того, $a^{**} = b^{**}$ тогда и только тогда, если $a \cap d = b \cap d$ для некоторого «плотного» d, удовлетворяющего условию $d^{**} = I$.*

Доказательство. Мы доказали уже все, кроме последнего, утверждения. Предположим, что $a \cap d = b \cap d$, где $d^{**} = I$. Тогда

$$a^{**} = a^{**} \cap I = a^{**} \cap d^{**} = (a \cap d)^{**} = (b \cap d)^{**} = \ \ldots \ = b^{**}.$$

Обратно, предположим, что $a^{**} = b^{**}$; положим $d = (a \cup b^*) \cap$ $\cap (a^* \cup b)$. Тогда в силу уже доказанного $d^{**} = (a^{**} \vee b^*) \cap$ $\cap (a^* \vee b^{**}) = (b^{**} \vee b^*) \cap (a^* \vee a^{**}) = I$. Кроме того,

$$a \cap d = a \cap (a \cup b^*) \cap (a^* \cup b) = a \cap (a^* \cup b) =$$
$$= (a \cap a^*) \cup (a \cap b) = a \cap b.$$

Аналогично, $b \cap d = a \cap b$, чем и завершается доказательство.

6.10.13. Дистрибутивные функционалы

Мы охарактеризуем теперь те оценки, которые определяют *дистрибутивные* метрические структуры.

Это весьма просто; поскольку $x \bigcup (y \bigcap z) \leqslant (x \bigcup y) \bigcap (x \bigcup z)$ в любой структуре, метрическая структура будет дистрибутивной тогда и только тогда, если $v\,[x \bigcup (y \bigcap z)] = v\,\{(x \bigcup y)\} \bigcap (x \bigcup z)]$ тождественным образом. Но во всякой метрической структуре мы имеем в силу VI из р. 6.6.6,

$$v\,[x \bigcup (y \bigcap z)] = v\,[x] + v\,[y] + v\,[z] - v\,[y \bigcup z] - v\,[x \bigcap y \bigcap z],$$

$$v\,[(x \bigcup y) \bigcap (x \bigcup z)] = v[x \bigcup y] + v[x \bigcup z] - v\,[x \bigcup y \bigcup z].$$

Произведя подстановку и переставив члены, мы получаем эквивалентное симметричное условие

$$v\,[x \bigcup y \bigcup z] - v\,[x \bigcap y \bigcap z] =$$
$$= v\,[x \bigcup y] + v\,[y \bigcup z] + v\,[z \bigcup x] - v\,[x] - v\,[y] - v\,[z]. \qquad (24)$$

Это условие не двойственно себе, но в силу VI имеем

$$v\,[x \bigcup y] + v\,[y \bigcup z] + v\,[z \bigcup x] - v\,[x] - v\,[y] - v\,[z] =$$
$$= v\,[x] + v\,[y] + v\,[z] - v\,[x \bigcap y] - v\,[y \bigcap z] - v\,[z \bigcap x].$$

Следовательно, (24) эквивалентно двойственному себе симметричному условию

$$2\,\{v\,[x \bigcup y \bigcup z] - v\,[x \bigcap y \bigcap z]\} = v\,[x \bigcup y] + v\,[y \bigcup z] + v\,[z \bigcup x] -$$
$$- v\,[x \bigcap y] - v\,[y \bigcap z] - v\,[z \bigcap x]. \qquad (25)$$

Оценки, удовлетворяющие условию (25), будем называть *дистрибутивными* оценками. Отсюда вытекает

Теорема 17. *Метрическая структура дистрибутивна тогда и только тогда, если ее оценка дистрибутивна.*

Весьма интересной является связь понятия оценки с двумерным волновым уравнением $\partial^2 v \,/\, \partial x^2 = \partial^2 v \,/\, \partial t^2$. Хорошо известно, что если принять характеристики $r = x + t$, $s = x - t$ в качестве новых независимых переменных, то волновое уравнение примет простую форму $\partial^2 v \,/\, \partial r\, \partial s = 0$. Общее решение дается тогда функцией

$$v\,(r, s) = f\,(r) + g\,(s),$$

подчиненной соответствующим условиям дифференцируемости.

Но мы напомним теперь, что характеристики частично упорядочивают пространство-время независимых переменных. При этой частичной упорядоченности $(r, s) \le (r_1, s_1)$ тогда и только тогда, если $r \le r_1$ и $s \le s_1$. Следовательно, $(r, s) \bigcup (r_1, s_1$ (max (r, r_1), max(s, s_1)) и двойственно, давая этим дистрибутивную структуру R^{*2}, где R^* обозначает цепь вещественных чисел. Теперь мы можем установить связь.

Теорема 18. *Вещественнозначная функция удовлетворяет волновому уравнению в двумерном пространстве-времени тогда и только тогда, если она является оценкой для структуры, определяемой путем обычного частичного упорядочивания пространства теории относительности.*

В самом деле, предположим, что v есть оценка. Тогда $(r, s) = (r, 0) \bigcup (0, s)$ и $(0, 0) = (r, 0) \bigcap (0, s)$. Следовательно $v(r, s) = v(r, 0) + [v(0, s) - v(0, 0)] = f(r) + g(s)$. Обратно, если $v(r, s) = f(r) + g(s)$, то

$$
\begin{aligned}
v &= v[(r, s) \bigcup (r_1, s_1)] + v[(r, s) \bigcap (r_1, s_1)] = \\
&= v[\max(r, r_1),\ \max(s, s_1)] + v[\min(r, r_1),\ \min(s, s_1)] = \\
&= f(\max(r, r_1)) + f(\min(r, r_1)) + g(\max(s, s_1)) + g(\min(s, s_1)) = \\
&= f(r) + f(r_1) + g(s) + g(s_1) = v(r, s) + v(r_1, s_1).
\end{aligned}
$$

Этим завершается доказательство.

6.11 Булевы алгебры

6.11.1. Определение

Исторически первыми рассмотренными структурами были дистрибутивные структуры с дополнениями, изучавшиеся Булем и названные его именем. Очевидно, что все результаты р. 6.9 применимы к булевым алгебрам.

Имеет место тесная связь между дистрибутивным законом и свойством единственности дополнений, отмеченная уже в р. 6.10, теорема 2, следствие 1. Имеет место

Теорема 1. *В дистрибутивной структуре дополнения единственны и являются ортодополнениями. Следовательно, в любой булевой алгебре мы имеем*

L8. $(x \cap y)' = x' \cup y'$ и $(x \cup y)' = x' \cap y'$.

L9. $(x')' = x$, а также

L10. $x \cap x' = 0$ и $x \cup x' = I$.

Доказательство. Если $a \bigcup x = I$ и $a \bigcap y = 0$, то

$$x = 0 \bigcup x = (a \bigcap y) \bigcup x = (a \bigcup x) \bigcap (y \bigcup x) = I \bigcap (y \bigcup x) = y \bigcup x.$$

Если также $a \cap x = 0$ и $a \cup y = I$, то аналогично $y = y \cup x$, откуда $x = y$. Этим доказана единственность. Но в силу L2 отношение дополнительности симметрично, а потому $(a')'$ a. Далее, если $a \leq b$, то $a \cup a' = I$ и $a \cap b' \leq b \cap b' = 0$. Следовательно, как и выше, $a' = b' \cup a'$ и $b' \leq a'$. Таким образом, соответствие $a \to a'$ есть дуальный автоморфизм, чем и завершается доказательство.

Из этой теоремы непосредственно следует, что булева алгебра дуально изоморфна себе. Вторым следствием является то, что структурный автоморфизм булевой алгебры сохраняет дополнения: операция взятия дополнения является *внутренней.*

Теорема 2. *Элементы любой дистрибутивной структуры, обладающие дополнениями, образуют подструктуру.*

Доказательство. Если элементы x и y дополняемы, то

$$(x \cap y) \cap (x' \cup y') = (x \cap y \cap x') \cup (x \cap y \cap y') = 0 \cup 0 = 0$$

и двойственно. Следовательно, $x \cap y$ имеет дополнение $x' \cup y'$, как и в L8. Двойственным образом $x \cup y$ имеет дополнение $x' \cap y'$, чем и завершается доказательство.

6.11.2. Примеры

Подмножества любой совокупности I образуют булеву алгебру: теоретико-множественные операции суммы, произведения и дополнения становятся теоретико-структурными операциями объединения, пересечения и дополнения. Более общо, любое тело множеств является булевой алгеброй при том же самом толковании операций. Булевы алгебры возникают также во многих других вопросах. Из них мы выбираем несколько примеров.

Пример 1. Центр любой структуры, есть булева алгебра.

Пример 2. Рассмотрим совокупность всех бинарных отношений между элементами двух классов I и J. Имеет место взаимно однозначное соответствие между отношениями ρ и подмножествами R произведения классов $I \times J$: $x \rho y$ тогда и только тогда, если $(x,y) \in R$. Следовательно, бинарные отношения между I и J образуют булеву алгебру.

Эта булева алгебра подробно изучается в р.6.14. Другие примеры булевых алгебр, встречающихся в теории множеств и логике, изучаются в р. 6.12.6.12. Пока мы отметим лишь следующие дополнительные абстрактно определяемые примеры.

Любое кардинальное произведение булевых алгебр есть булева алгебра. Таковой же является и любая «подалгебра» булевой алгебры, т. е. любое подмножество, замкнутое относительно всех трех операций в булевой алгебре. Любой структурно-гомоморфный образ булевой алгебры есть также булева алгебра: этот гомоморфный образ является дистрибутивной структурой, и равенства $x\cap x'=0$, $x\cup x'=I$ сохраняются.

6.11.3. Булевы кольца

Стон показал, что теорию булевых алгебр можно свести к общей теории колец, фактически даже к коммутативным кольцам характеристики два в обычном смысле.

Чтобы обосновать это, мы напомним понятие «характеристической функции» f_X множества X в пространстве I: это есть функция, определенная на точках $p \in I$, и такая, что $f_X(p) = 1$ или 0 в соответствии с тем, будет ли $p \in X$ или $p \in X'$. Тогда $f_X f_Y = f_{X\cap Y}$ и $f_X + f_Y = f_{(X\cap Y')\cup(X'\cap Y)} \bmod 2$. Таким образом, элементы булевой алгебры A, определенной каким-нибудь телом множеств, образуют по отношению к *пересечениям* $XY = X\cap Y$ и *симметричным разностям* $X+Y=(X\cap Y')\cup(X'\cap Y)$ кольцо $R(A)$ характеристики два с единицей I, удовлетворяющей соотношению $IX — XI = X$ для всех X. Кроме того, в $R(A)$ $XY = YX$ и $XX = X$ тождественным образом: кольцо является коммутативным, и все его элементы идемпотентны.

Фактически в любом кольце тождество xx x влечет $x+y= =(x+y)(x+y)=xx+yx+xy+yy=x+y+yx+xy$, а потому оно влечет $xy+yx = 0$. Полагая $x = y$, мы получаем $x+x = 0$, откуда $x = — x$. Используя это, мы получаем $xy — yx = 0$, а потому $xy = yx$. Следовательно, из $xx = x$ следуют оба тождества $xy = yx$ и $x + x = 0$. Это подсказывает дать следующее

Определение. *«.Булево кольцо» есть такое кольцо, все элементы которого идемпотентны.*

Теорема 3. *Имеет место взаимно однозначное соответствие между булевыми алгебрами и булевыми кольцами с единицей. При этом отношение включения соответствует делимости, теоретико-структурные пересечения — произведениям в кольцах,* $0 = 0$, I=1 *и*

$$x+y=(x\cap y')\cup(x'\cap y) \quad u \quad x\cup y=x+y-xy. \tag{1}$$

Доказательство. Пусть R—произвольное булево кольцо с единицей 1. Если определить $x \geq y$ в смысле $xy = y$, то, очевидно, а) $1\geq x\geq 0$ для всех x; б) $x\geq x$ в силу идемпотентности; в) если $x\geq y$ и $y\geq x$, то $x = yx = xy = y$ в силу предположения и коммутативности; г) если $x\geq y$ и $y\geq z$, то $x = xy = x(yz) = (xy)z = xz$, а потому $x\geq z$; д) $x\geq xy$, так как $x(xy) = (xx)y = xy$; е) аналогично, путем использования коммутативности, $y \geq = xy$; ж) если $x \geq z$ и $y \geq z$, то $xyz = xz = z$, а потому $xy \geq z$; з) соответствие $x \longleftrightarrow 1 - x$ является, очевидно, взаимно однозначным; кроме того, поскольку $xy=y$ влечет $(1-y)(1-x) = 1 - y - x + xy = 1 - x$, оно переворачивает отношение включения; следовательно, это соответствие есть дуальный автоморфизм.

Следовательно, наше определение превращает R в частично упорядоченное множество с 0 и I (в силу «а» — «г»), в котором $x \cap y$ существует и совпадает с xy (в силу «д» —«ж»), откуда (в силу «з») $x \cup y$ также существует и совпадает с $1 - (1 - x)(1 - y) = x + y - xy$, как в (1). Кроме того, если мы положим, по определению $x' = 1 - x$, то $x \cap x' = x(1 - x) = 0$ и $x \cup x' = x + (1 - x) + x(1 - x) = 1$; следовательно, наше определение превращает R в структуру с дополнениями. Наконеп,

$$x \cap (y \cup z) = x(y + z - yz) = xy + xz - xyz =$$
$$= xy + xz - xyxz = (x \cap y) \cup (x \cap z),$$

а потому R есть булева алгебра.

Обратно, в произвольной булевой алгебре A определим в соответствии с (1)

$$x + y = (x \cap y') \cup (x' \cap y) = (x \cup y) \cap (x' \cup y').$$

Очевидно, что $0 + x = x$, $x + y = y + x$ и $x + x = 0$. Далее, легко подсчитать, что $(x + y)' = (x' \cap y') \cup (x \cap y)$ и, используя это,

$$(x + y) + z =$$
$$= (x \cap y' \cap z') \cup (x' \cap y \cap z') \cup (x' \cap y' \cap z) \cup (x \cap y \cap z).$$

В силу левой симметрии это равно также $x+(y+z)$. Следовательно, сложение коммутативно и ассоциативно, 0 есть нуль, и каждый элемент является обратным самому себе; следовательно, A есть коммутативная группа по сложению. Далее, умножение, определяемое по формуле $xy = x \cap y$ является идемпотентным, коммутативным и ассоциативным, и I есть мультипликативная единица.

Остается проверить дистрибутивный закон, что делается следующим образом:

$$xz + yz = [x \cap z \cap (y' \cup z')] \cup [(x' \cup z') \cap y \cap z] =$$
$$= [x \cap z \cap y'] \cup [x' \cap y \cap z] =$$
$$= [(x \cap y') \cup (x' \cap y)] \cap z = (x + y) z.$$

Этим завершается доказательство.

6.11.4. Теория постулатов

Булева алгебра может быть определена с помощью разнообразных систем постулатов. Мы упомянем лишь некоторые из них, хотя их сравнительное изучение дало бы нам представление о силе теории постулатов.

Привлекательная система постулатов принадлежит Ньюмену.   Пусть *А* — любая алгебра с двумя бинарными операциями, удовлетворяющая следующим постулатам:

N1. $a(b+c)=ab+ac.$

N1'. $(a+b)c=ac+bc.$

N2. Существует такой элемент 1, что $a1=a$ для всех a.

N3. Существует такой элемент 0, что $a+0=a=0+a$ для всех a.

N4. Каждому a соответствует по крайней мере одно такое a', что $aa'=0$ и $a+a'=1$.

Отметим, что идемпотентность, коммутативность и ассоциативность этих двух операций не предполагается. Мы доказываем теперь

(T1) $$aa=aa+0=aa+aa'=a(a+a')=a1=a.$$

Путем несколько больших выкладок мы доказываем

(T2) $(a')'=a$ для всех a и всех $(a')'$.

Доказательство.

$$\begin{aligned}(a')' &= 0+(a')'(a')'=\\ &=a'(a')'+(a')'(a')'=\\ &=(a'+(a')')(a')'=1(a')'=\\ &=(a+a')(a')'=a(a')'+0=\\ &=0+a(a')'=aa'+a(a')'=\\ &=a(a'+(a')')=a1=a.\end{aligned}$$

Отсюда следует, что

(N4') $a'a=0$ *и* $a'+a=1$,

а также, что дополнения единственны. [Действительно, если a' любое дополнение a, то $a'=((a')')'$ для любого $((a')')'$]- Далее,

(T3) $a0=0=0a$ для всех a.

Доказательство.

$$\begin{aligned}0&=aa'=a(a'+0)=\\ &=aa'+a0=0+a0=a0\end{aligned}$$

и

$$0=bb'=(0+b)b'=0b'+bb'=0b'+0=0b'.$$

Но в силу T2 каждое $a=b'$, где $b=a'$.

Отсюда следует, что если 0=1, то

$$0=0+0=a+a\cdot 0=0+a\cdot 1=a\cdot 1=a \text{ для всех } a.$$

Следовательно, за исключением этого тривиального одноэлементного случая, 0≠1. Отныне мы будем предполагать, что 0≠1. Другим следствием является

(N2'). $$1a=(a+a')\,a=aa+a'a=a+0=a \text{ для всех } a.$$

Следовательно, имеет место *полная левая симметрия в свойствах сложения и умножения.*

Определим теперь $1+1=2,\ (1+1)+(1+1)=2+2=4$ и назовем левые кратные $y2$ элемента 2 *четными* элементами. Заметим, что в силу T1, $4=2+2=2\cdot 1+2\cdot 1=2\,(1+1)=2\cdot 2=2.$ Далее,

(T4) x четно тогда и только тогда, если $x+x=x$.

Действительно, очевидно, что $y2+y2=y\,(2+2)=y2;$ обратно, если $x=x+x,$ то $x=x1+x1=x\,(1+1)=x2.$

(T5) Любое кратное xt или ux четного элемента x есть четный элемент.

Действительно, если $x=x+x,$ то $xt=(x+x)\,t=xt+xt$ и $ux=u\,(x+x)=ux+ux$ для всех t, u.

(T6) Соответствие $x\to x+x=x2$ есть идемпотентный эндоморфизм:

$$(x+y)\,2=x2+y2,\ (xy)\,2=(x2)\,(y2) \text{ и } (x2)\,2=x2.$$

Доказательство, $(x+y)\,2=x2+y2;\ (x2)\,2=x2+x2=$

$$=x\,(2+2)=x2$$

и

$$(x2)\,(y2)=(x+x)\,(y+y)=(x+x)\,y+(x+x)\,y=$$
$$=(xy+xy)+(xy+xy)=(xy)\,2+(xy)\,2=(xy)\,2 \text{ (в силу (T4))}.$$

Отсюда следует, что *четные элементы образуют подалгебру B,* в которой сложение идемпотентно. В этой подалгебре мы имеем

$$a+1=(a+1)\,(a+a')=(aa+a)+(aa'+1a')=$$
$$=(a+a)+(0+a')=a+a'=1$$

для всех a. В силу левой симметрии $1+a=a$. Но из этого результата, (T1), N1 —N4 и (N4'), (N2'), вытекает в силу теоремы 3 р. 6.10, что подалгебра B есть булева алгебра, в которой 2 играет роль I. Дополнения в B являются относительными дополнениями в A.

Аналогично, можно показать, что «нечетные» элементы, удовлетворяющие эквивалентным условиям $x=y2'$, $x=x2'$, $x+x=0$, образуют подалгебру, являющуюся коммутативным кольцом с единицей, удовлетворяющей условию 1+1 = 0, — т. е. неасссоциативное булево кольцо. Кроме того, A является прямым произведением этих подалгебр. Отсюда вытекает

Теорема 4. *Любая «ньюменова алгебра», удовлетворяющая условиям* N1 — N1', N2, N3, N4, *является прямым произведением булевой алгебры и (возможно, неассоциативного) булева кольца.*

Из того, что было нами доказано подробно, вытекает

Следствие. *Тождества* $a \bigcup a = a,\ a \bigcup 0 = 0 \bigcup a = a,\ I \bigcap a = a,$ L6'— L6'' *и* L10 *являются постулатами для булевой алгебры.*

Другой системой постулатов для булевой алгебры является система постулатов в терминах *операции с чертой* Шеффера

$$x \mid y = x' \cap y' \text{ (бинарное отклонение).} \qquad (2)$$

Все булевы операции могут быть выражены в терминах этой единственной бинарной операции; так,

$$x' = x \mid x, \quad x \bigcap y = (x \mid x) \mid (y \mid y), \quad x \bigcup y = (x \mid y) \mid (x \mid y). \qquad (3)$$

Однако системы постулатов для этой операции не очень изящны. Так, например, Шеффер предложил определить $a' = a|a$ и постулировать $(a')' = a,$

$$(a \mid (b \mid b')) = a', \ (a \mid (b \mid c))' = (b' \mid a) \mid (c' \mid a).$$

Тернарная операция *медианы*

$$(x,\ y,\ z) = (x \bigcap y) \bigcup (y \bigcap z) \bigcup (z \bigcap x) = (x \bigcup y) \bigcap (y \bigcup z) \bigcap (z \bigcup x) \qquad (4)$$

из р. 6.10.3, может быть также использована в качестве основы для булевой алгебры. Отметим здесь следующий результат, необходимый для р. 15.

Теорема 5. *Каждый групповой сдвиг* $x \to x+a$ *в булевой алгебре A является автоморфизмом относительио операции медианы* (4).

Доказательство. Путем прямых выкладок мы получаем из (1)

$$\begin{aligned}(x+a) \bigcap (y+a) &= [a' \bigcap (x \bigcap y)] \bigcup [a \bigcap (x \bigcup y)'],\\ (x+a) \bigcup (y+a) &= [a' \bigcap (x \bigcup y)] \bigcup [a \bigcap (x \bigcap y)'].\end{aligned} \qquad (5)$$

Путем индукции мы устанавливаем, что для каждого структурного полинома $p\ (x_1, \ldots, x_n)$

$$p\,(x_1 + a,\ \ldots,\ x_n + a) = [a' \bigcap p\,(x_1,\ \ldots,\ x_n)] \bigcup [a \bigcap g\,(x_1,\ \ldots,\ x_n)],$$

где g есть дополнение элемента, двойственного *p.* Следовательно, если *p двойственно себе,* подобно (x, y, z), то мы имеем

$$p\,(x_1 + a,\ \ldots,\ x_n + a) = p\,(x_1,\ \ldots,\ x_n) + a.$$

Этим завершается доказательство.

6.11.5. Теория представлений

Теория представлений дистрибутивных структур, приведенная в п.4 — 5 р. 6.10, имеет существенное уточнение для случая булевых алгебр. Мы знаем, что единственными неразложимыми в объединение элементами $p_i > 0$ дедекиндовой структуры с дополнениями являются ее точки. Следовательно, если L — структура с дополнениями, то множество X из теоремы 5, р. 6.10, является совершенно неупорядоченным. Отсюда вытекает

Теорема 6. *Каждая булева алгебра конечной длины n изоморфна телу всех подмножеств множества из n элементов.* Таким образом, в частности, имеется в точности одна булева алгебра длины n; это есть 2^n.

Далее, в следствии из теоремы 6, р. 6.10, мы можем удалить точки, не принадлежащие I и точки, принадлежащие 0, и получить тем самым изоморфное представление, в котором 0 соответствует пустому множеству, а I — всему пространству. Но в таком представлении дополнения должны в силу L10 соответствовать дополнениям. Мы получаем результат Стона:

Теорема 7. *Любая булева алгебра изоморфна телу множеств.*

Таким образом, любая система постулатов для булевой алгебры является тем самым полной системой постулатов для алгебры классов с операциями конечного объединения, конечного пересечения и взятия дополнения.

6.11.6. Теория идеалов

В дистрибутивной структуре каждый идеал нейтрален. Следовательно, как непосредственное следствие теоремы 11 р. 6.9, мы имеем следующий результат:

Теорема 8. *Идеалы любой булевой алгебры A находятся во взаимно однозначном соответствии с ее отношениями конгруентности* θ; *каждое* θ *соответствует идеалу J элементов* $x \equiv 0 \pmod{\theta}$.

Можно также вывести этот результат из теории колец. Но он является столь важным, что мы дадим специальное, более естественное доказательство того факта, что каждый идеал J определяет в A отношение конгруентности. В сочетании с теоремой 3 р. 6.3 это будет давать независимое доказательство теоремы 8.

Доказательство. Пусть для заданного J выражение $x \equiv y$ (J) означает, что $x \cup t = y \cup t$ для некоторого $t \in J$; это отношение является рефлексивным и симметричным. Оно является транзитивным, ибо если $x \cup t = y \cup t$ и $y \cup u = z \cup u$ $[t, u \in J]$, то $x \cup (t \cup u) = y \cup t \cup u = y \cup u \cup t = z \cup u \cup t = z \cup (t \cup u)$ для некоторого $t \cup u \in J$. Следовательно, это отношоние есть отношение эквивалентности. Далее, предположим, что $x \equiv y (J)$. Тогда для любого a $a \cup x \cup t = a \cup y \cup t$, следовательно, $a \cup x \equiv a \cup y (J)$, а потому справедливо свойство замены для объединений. Отметим, что все это применимо к любой структуре. Далее, если $x \equiv y(J)$, то $(a \cap x) \cup t = (a \cup t) \cap (x \cup t) = (a \cup t) \cap (y \cup t) = (a \cap y) \cup t$, и мы имеем тем самым отношение конгруентности для пересечений и объединений. Наконец, $x \equiv 0(J)$ влечет $x \leqslant x \cup t = 0 \cup t = t$ для некоторого $t \in J$, следовательно, $x \in J$ и обратно; таким образом, J есть множество во всех $x \equiv 0(J)$, что и требовалось доказать.

Отметим, что любой структурный гомоморфизм переводит в силу L10 дополнения в дополнения. Следовательно, $x \equiv y(J)$ влечет $x' \equiv y'(J)$, а потому структурные гомоморфизмы обладают свойством замены для всех трех булевых операций.

Напомним, что идеалы любой дистрибутивной структуры сами образуют дистрибутивную структуру; кроме того, объединения и пересечения идеалов можно истолковывать в смысле исчисления комплексов. Эти и все прочие результаты р. 6.10.6, *заведомо* применимы к булевым алгебрам. Отметим здесь те свойства идеалов в булевых алгебрах, которые неверны в случае произвольных дистрибутивных структур.

Прежде всего, идеал булевой алгебры A является *простым* тогда и только тогда, если он *максимален*. В самом деле, если P простой идеал в булевой алгебре, то для любого $a \notin P$ мы имеем $a' \in P$, ибо $a \cap a' = 0 \in P$; следовательно, любой идеал $J>P$ содержит некоторое $a \notin P$ и a', а потому совпадает с I. Обратно, предположим, что идеал M максимален, и пусть $x \cap y \in M$, $x \notin M$. Тогда $x \cup M$ содержит I; следовательно, для некоторого $z \in M$ $x \cup z = I$ и $y = y \cap I = y \cap (x \cup z) = (y \cap x) \cup (y \cap z) \in M \cup M = M$. Следовательно, идеал M прост, что и требовалось доказать.

Пусть для любого подмножества X булевой алгебры A X^* обозначает множество всех $a \in A$ таких, что $a \cap x = 0$для всех $x \in X$. Очевидно, что $a \cap x' = 0$ для всех $x \in X$ тогда и только тогда, если $a \leq x$ для всех $x \in X$, т. е. тогда и только тогда, если a принадлежит *замкнутому идеалу,* верхнее сечение которого есть X (см. р. 6.5.7). Далее, если J идеал в A, то J^* является псевдодополнением J в дистрибутивной структуре всех идеалов, в смысле р. 6.10.12. Следовательно, идеал булевой алгебры замкнут тогда и только тогда, если он является псевдодополнением некоторого другого идеала.

Мы будем называть идеал J *плотным* тогда и только тогда, если $J^* =0$. Это означает, что для любого $a > 0$ из A мы можем найти такое $x > 0$ в J, что $x \cap a > 0$,—следовательно, можно найти $y = x \cap a > 0$ в J, причем $y \leq a$. Используя это условие, легко показать, что пересечение любых двух плотных идеалов есть плотный идеал. Далее, если J плотно и

$K \geq J$, то, очевидно, K плотно. Следовательно, *плотные идеалы образуют дуальный идеал в дистрибутивной структуре всех идеалов из А.*

Поскольку $(K \bigcup K^*)^* = K^* \bigcap (K^*)^* = 0$ для любого идеала K, $K \cup K^*$ является всегда плотным. Кроме того,

$(K^*)^* \bigcap (K \bigcup K^*) = ((K^*)^* \bigcap K) \bigcup ((K^*)^* \bigcap K^*) = K;$

следовательно, по модулю этого дуального идеала каждый идеал конгруентен своему замыканию $(K^*)^*$. Покажем теперь, что, обратно, любой идеал $K \cap D$, конгруентный замкнутому идеалу K по модулю «плотного» идеала, имеет K в качестве замыкания. В самом деле, очевидно, что $((K \bigcap D)^*)^* \leqslant (K^*)^* = K$. В то же время для любых $b \in (K \cap D)^*$ и $a \in K$ из $a \cap b > 0$ вытекало бы $0 < d \leqslant a \bigcap b \leqslant a \in K$ для некоторого $d \in D$, следовательно, $d \in K \cap D$; но $d \leqslant a \bigcap b \leqslant b \in (K \bigcap D)^*$; следовательно, $d = 0$— противоречие. Таким образом, каждое $a \in K$ удовлетворяет соотношению $a \cap b = 0$ для всех $b \in (K \bigcap D)^*$ и $K \leqslant (K \bigcap D^*)^*$.

Следовательно, имеет место структурный гомоморфизм дистрибутивной структуры всех идеалов на структуру замкнутых идеалов, в силу чего последняя также является дистрибутивной структурой. В этой структуре K^* является дополнением для K, поскольку $K \cap K^* = 0$ и объединение $((K \bigcup K^*)^*)^* = 0^* = I$ идеалов K и K^* в структуре (не подструктуре) замкнутых идеалов есть I. Следовательно, имеет место

Теорема 9 (Гливенко и Стон). *Пополнение с помощью сечений А любой булевой алгебры А есть булева алгебра; кроме того, соответствие* $J \longrightarrow (J^*)^*$ *является структурным гомоморфизмом структуры всех идеалов в А на* $\overline{A}$.

6.11.7. Подалгебры

Подалгебра булевой алгебры есть подструктура, содержащая вместе с любым x также x', и, следовательно, содержащая 0 и I. Легко определить подалгебры конечной булевой алгебры $A=2^n$. В самом деле, любая подалгебра S алгебры A будет сама булевой алгеброй конечной длины, «точки» которой независимые элементы в 2^n (неперекрывающиеся подмножества в I), имеющие объединением I, т. е. эти «точки» являются подмножествами, на которые распадается I при некотором *разбиении.* Обратно, подмножества, на которые любое разбиение делит множество I, являются неделимыми, или «атомными» элементами тела множеств, так что соответствие между подалгебрами алгебры 2^n и разбиениями множества I взаимно однозначно. Наконец,

$S \leq T$ тогда и только тогда, если «точки» из S являются объединениями «точек» из T, т. е. тогда и только тогда, если T осуществляет подразбиение разбиения, производимого S. Таким образом, имеет место.

Теорема 10. *Структура всех подалгебр любой конечной булевой алгебры* 2^n *дуально изоморфна структуре всех разбиений множества из n элементов.*

Соответствующий результат справедлив для структуры всех *замкнутых* подалгебр алгебры $2^к$ в случае любого кардинального числа к; эти подалгебры являются «полными телами множеств» множества из к точек.

Далее, для любого n всякая перестановка n точек алгебры $A=2^n$ индуцирует вполне определенный автоморфизм A. Кроме того, всякий автоморфизм алгебры A переставляет ее точки. Следовательно, *группа автоморфизмов алгебры* 2^n *является симметрической группой* Г, *составленной из n букв.* В частности 1 и 2 — естественные конечные булевы алгебры, не имеющие собственных автоморфизмов.

Обратно, пусть $\varphi(x_1, \ldots, x_r)$ —любая операция на подмножествах x_i класса I из n элементов, перестановочная с каждым $\gamma \in$ Г, так что

$$\varphi(x_1\gamma, \ldots, x_r\gamma) = \gamma[\varphi(x_1, \ldots, x_r)] \text{ для всех } \gamma \in \Gamma. \qquad (6)$$

Булева подалгебра S алгебры 2^n, порожденная подмножествами $x_1, \ldots, x_r$, определяет в соответствии с теоремой 10 разбиение π класса I на подклассы $a_1, \ldots, a_s$. Кроме того, подгруппа $\sum$ всех $\gamma \in \Gamma$, удовлетворяющих соотношениям $x_i\gamma = x_i$ для всех i, является наибольшей группой, имеющей a_i в качестве множеств транзитивности (в теоретико-групповом смысле). Следовательно, если $y = \varphi(x_1, \ldots, x_r)$, так что в силу (6) $y\gamma = y$ для всех $\gamma \in \sum$, то y должно быть объединением множеств a_i, а потому $y \in S$ является булевой функцией от x_i. Отсюда вытекает, что *булевыми операциями на конечных множествах являются в точности те операции, которые инвариантны относительно сим.метрической группы перестановок точек.*

6.11.8. Свободные булевы алгебры

Определим теперь свободную булеву алгебру, порожденную n символами. Задача может быть сформулирована иначе следующим образом. Под «булевой функцией» переменных $x_1, \ldots, x_n$ понимают функцию, построенную на основе трех базисных операций: объединения, пересечения и взятия дополнения. Мы спрашиваем: каковы различные булевы функции переменных x и каким образом их можно комбинировать?

Для получения ответа образуем сперва 2^n «элементарных» булевых функций $f_i : x_{i1} \cap \ldots \cap x_{in}$, где x_{ij} есть либо x_j, либо x'_j, в зависимости от i. Например, если $n = 2$, мы образуем

$$f_1 = x_1 \cap x_2, \quad f_2 = x_1 \cap x'_2, \quad f_3 = x'_1 \cap x_2, \quad f_4 = x'_1 \cap x'_2.$$

Отметим теперь, что различные f_i имеют в пересечении 0: если $i \neq k$, то $f_i \cap f_k \leq x_j \cap x'_j$ для некоторого j, откуда $f_i \cap f_k = 0$.

Сопоставим затем каждому непустому множеству S функций f_i функцию $g_S = \vee_{i\in S} f_i$ и определим g_0 как 0. Мы докажем, что

1) $g_0 = 0$ и $g_1 = I$, 2) $g_{S\cap T} = g_S \cap g_T$, $g_{S\cup T} = g_S \cup g_T$ и $g_{S'} = (g_S)'$, 3) каждая булева функция переменных x_i есть g_S.

Доказательство п. 1. В силу определения $g_0 = 0$; в то же время в силу общего дистрибутивного закона

$$I = \bigwedge_{j=1}^{n} (x_j \cup x'_j) = \bigvee_{i=1}^{2^n} \bigwedge_{j=1}^{n} x_{ij} = \bigvee_{i\in I} f_i.$$

Доказательство п. 2. В силу определения и LI — L3 $g_S \cup g_T = g_{S\cup T}$. Далее, $g_S \cap g_T$ является в силу общего дистрибутивного закона объединением $f_i \cap f_k \, [i \in S,\ k \in T]$; с другой стороны, $f_i \cap f_k$ есть 0, если $i \neq k$. Следовательно, $g_S \cap g_T = \vee f_i \cap f_i \, [i \in S,\ i \in T]$, а это в силу L1 есть $\vee\ f_i[i \in S \cap T]$, т. е. $g_{S\cap T}$, что и требовалось доказать. Отсюда следует, что $g_S \cup g_{S'} = I$ и $g_S \cap g_{S'} = 0$, откуда $g_{S'} = (g_S)'$.

Доказательство п. 3. В силу (2) всякая булева функция от g_S сама является $g_{S'}$, следовательно, достаточно показать, что x_i суть g_S, или в силу соображений симметрии, что x_1 есть g_S. Но в силу общего дистрибутивного закона, если X_1 обозначает множество таких i, что $x_{i1} = x_1$, имеем

$$x_1 = x_1 \cap \bigwedge_{j=2}^{n} (x_j \cap x'_j) = \bigvee_{i\in X_1} f_i.$$

Отсюда следует, что каждая булева функция есть g_S; покажем теперь, что различные g_S представляют собой различные функции. Пусть I состоит из всех точек с n координатами, каждая из которых есть 0 или 1 (вершины n-мерного куба). Обозначим через x_j множество точек, j-я координата которых есть единица. Тогда различные f_i будут представлять различные точки: f_i представляет точку, j-я координата которой есть 0 или 1 в зависимости от того, будет ли x_{ij} равно x'_j или x_j. Следовательно, для этого случая, а потому и в случае свободной булевой алгебры различные множества S определяют различные g_S. Отсюда вытекает

Теорема 11. *Свободная булева алгебра с n образующими изоморфна алгебре* 2^{2^n}.

Хотя свободная структура с двумя образующими есть булева алгебра, она не является свободной булевой алгеброй с двумя образующими.

6.11. 9. Булевы уравнения

Как будет показано в р. 6.13.3, любая система утверждений классической логики может быть записана в форме системы уравнений с коэффициентами в булевой алгебре. Обсудим теперь вопрос о решении таких уравнений или, что равносильно этому, вопрос о сведении их к возможно более простой канонической форме. Эта проблема привлекала много внимания математиков логиков девятнадцатого столетия, поскольку она соответствует наибольшему логическому упрощению системы утверждений.

Одним из приемов является следующий. Любое уравнение $p = q$ может быть приведено к форме $r = (p' \cap q) \cup (p \cap q') = 0$. Система двух уравнений $r = 0$ и $s = 0$ эквивалентна одному единственному уравнению $r \cup s = 0$. Следовательно, может быть составлено в точности $2^{2^{m+n}}$ неэквивалентных систем утверждений относительно m классов, включающих в себя n свойств. Если мы используем $a_1, \ldots, a_n$ для обозначения свойств, а $x_1, \ldots, x_m$ для обозначения классов, которые мы пытаемся описать, то мы можем систематизировать описание в соответствии с п.6.11.8, выписывая *произвольное подмножество множества всех* 2^{m+n} *элементарных утверждений*

$$
\begin{gathered}
a_1 \cap \ldots \cap a_n \cap x_1 \cap \ldots \cap x_m = 0, \\
a_1 \cap \ldots \cap a_n \cap x_1 \cap \ldots \cap x'_m = 0, \\
\ldots\ldots\ldots\ldots\ldots\ldots\ldots\ldots, \\
a'_1 \cap \ldots \cap a'_n \cap x'_1 \cap \ldots \cap x'_m = 0.
\end{gathered}
$$

Это является вполне регулярным приемом. Согласно п.6.11.3, мы можем также записать каждое уравнение, используя систему обозначений для *колец;* это особенно полезно для *линейных* уравнений. Рассмотрим таким образом

$$ax = b, \quad \text{что эквивалентно} \quad ax + b = 0. \qquad (7)$$

Умножая на $1 + a = 1 — a$, мы получаем $0 = b + ab$; следовательно, (7) не имеет решений, кроме случая, когда $b=ab$. В этом случае оно сводится к $a(x + b) = 0$. Умножая на b, мы получаем $abx+ab = 0$, или $x \geq ab$. Далее, $0=a(x+ab)$, откуда $x+ab \leq 1+a$. Записывая $x = ab+t$, мы получаем $t \leq 1+a$. Обратно, любое $x = ab+t$ с $t = (1+a)t$ удовлетворяет соотношению $a(ab+(1 + a) t) = aab= b$, если $b = ab$. Мы заключаем, что (7) *не* имеет решений, кроме случая, когда $b \leq a$, и если $b \leq a$, то общим решением является $x = ab+t$, где $t \leq a'$.

Для системы линейных уравнений мы можем осуществить процесс *исключения,* используя алгоритм Эвклида. Мы проиллюстрируем это на случае двух уравнений:

$$ax + by = c, \quad dx + ey = f. \tag{8}$$

Прибавляя *d* раз первое уравнение ко второму, мы заменяем (8) эквивалентной системой (8')

$$ax + by = c, \quad (d+ad)x+(c + bd)y = f+cd. \tag{8'}$$

Прибавляя теперь второе из них к первому, получаем

$$(a \bigcup d)\, x + (b + c + bd)\, y = c + f + cd. \tag{9}$$

Прибавляя $d+ad$ раз (9) ко второму уравнению из (8'), мы получаем

$$(e + bd + de + ade)\, y = f + cd + df + adf. \tag{9'}$$

Уравнения (9) —(9') эквивалентны системе (8), которую мы таким образом «привели» к треугольной форме.

6.11.10. Бесконечная дистрибутивность

Покажем теперь, что бесконечный дистрибутивный закон $x \bigcap \vee_B y_\beta = \vee_B (x \bigcap y_\beta)$ справедлив в любой булевой алгебре. В самом деле, $x \cap \vee_B y_\beta$ является верхней гранью для каждого $x \cap y_\beta$; следовательно, нам нужно лишь показать, что $x \cap \vee_B y_\beta$ содержится во всякой верхней грани *u* для $x \cap y_\beta$. Но если $x \cap y_\beta \le u$ Для всех β, то $y_\beta = (y_\beta \bigcap x) \bigcup (y_\beta \bigcap x') \leqslant u \bigcup x'$ для всех β, а потому

$$x \bigcap \vee_B y_\beta \leqslant x \bigcap (u \bigcup x') = (x \bigcap u) \bigcup (x \bigcap x') = x \bigcap u \leqslant u.$$

Этим завершается доказательство. На основании теоремы 14 из р. 6.10.11, и принципа двойственности мы устанавливаем следующий результат.

Теорема 12. *Любая полная булева алгебра является топологической структурой, удовлетворяющей бесконечным дистрибутивным законам*

$$x \bigcap \vee_B y_\beta = \bigvee_B (x \bigcap y_\beta), \quad \bigvee_A x_\alpha \bigcap \bigvee_B y_\beta = \bigvee_{A, B} (x_\alpha \bigcap y_\beta)$$

$$\text{и двойственно.} \tag{10}$$

Поскольку пополнение булевой алгебры с помощью сечений есть булева алгебра (теорема 9) и так как пополнение с помощью сечений сохраняет наиб. н. г. и наим. в. г., если только они определены, мы заключаем, что бесконечные дистрибутивные законы (10) справедливы в любой булевой алгебре, когда входящие в них выражения определены.

Это неверно для дистрибутивного закона (22') из р. 6.10.11. В действительности конструкция из п. 8, трансфинитно продолженная, приводит к следующему результату Тарского.

Теорема 13. *Если булева алгебра A вполне дистрибутивна, то она изоморфна алгебре* $2^{к}$ *всех подмножеств некоторого множества.*

Замечание. Структура называется «вполне дистрибутивной» тогда и только тогда, если для нее справедлив закон (22') из р. 6.10.11, без ограничения на число входящих в него членов.

Доказательство. Пусть x_α обозначает общий элемент в *A;* в силу (22') мы можем писать $I = \wedge_A (x_\alpha \bigcup x'_\alpha) = \vee\, \theta_\varphi$, где θ_φ обозначает общий элемент $\wedge_A x_{\varphi\alpha}$. ($x_{\varphi\alpha} = x_\alpha$ или x'_α). Покажем сперва, что каждое θ_φ, отличное от 0, покрывает 0. В самом деле, если $x_\alpha < \theta_\varphi$, то (так как из $x_{\varphi\alpha}=x_\alpha$ вытекало бы $\theta_\varphi \le x_\alpha$ в противоречие тому, что $x_\alpha < \theta_\varphi$) $x_{\varphi\alpha} = x'_\alpha$, а потому $x_\alpha = x_\alpha \bigcap \theta_\varphi \leqslant x_\alpha \bigcap x'_\alpha = 0$. Следовательно, элементы $\theta_\varphi > 0$ являются точками; это — основная идея Тарского.

Кроме того, каждое $x_\alpha = x_\alpha \bigcap I = x_\alpha \bigcap \vee\, \theta_\varphi = \vee\, (x_\alpha \bigcap \theta_\varphi)$ является объединением «точек» $\theta_\varphi > 0$, которые оно содержит (ибо или $\theta_\varphi \leqslant x_\alpha$ и $x_\alpha \bigcap \theta_\varphi = \theta_\varphi$, или $x_\alpha \bigcap \theta_\varphi < \theta_\varphi$ и $x_\alpha \bigcap \theta_\varphi = 0$).

Далее, объединение $g_S = \vee_S \theta_\varphi$ точек не содержит точек *p,* не принадлежащих *S,* ибо, если $p \notin S$, то в силу (10)

$$g_S \bigcap p = \bigvee_S (\theta_\varphi \bigcap p) = \bigvee_{\varphi \in S} 0 = 0.$$

Этим устанавливается взаимно однозначное и сохраняющее упорядоченность соответствие между множествами *S* точек и различными элементами g_S из *A,* что и требовалось доказать. Заметим, что утверждение «и двойственно» в (22') является ненужным для этого доказательства.

6.11.11. Булевы σ-алгебры

Мы докажем теперь теорему о представлениях *булевых σ-алгебр,* т. е. σ -полных булевых алгебр.

Теорема 14 (Люмис). *Любая булева σ-алгебра A является σ-гомоморфным образом σ-тела множеств.*

Доказательство. Пусть *I* — класс всех функций *P,* которые отбирают из каждой пары *a, a'* дополнений в *A* какой-нибудь один элемент; мы назовем эти функции «точками». Каждому $a \in A$ мы сопоставим множество τ (*a*) всех таких точек *P,* что $a \in P$. Эти множества замкнуты относительно взятия дополнения, ибо τ *(a')* = [τ *(a)*]'. Следовательно, они порождают σ-тело Φ подмножеств множества *I* со счетными объединениями и пересечениями. Пусть *N* — семейство конечных или счетных пересечений $\wedge\tau(a_i)$, для которых $\wedge a_i=0$ в A. Пусть J σ-идеал, порожденный множеством N; он будет состоять из всех $t\leq\vee n_i$ для счетных подмножеств $\{ n_i \}$ из N. Мы покажем, что Φ/J изоморфно A при соответствии τ: $a\to\tau(a)$.

Прежде всего, τ есть σ-гомоморфизм A на все Φ/J. Действительно, если $a=\vee a_i$, то $a' \sqcap a_i \leqslant a' \sqcap a = 0$ для всех i и $0 = a \sqcap a' = a \sqcap (\vee a_i)' = a \sqcap \wedge a'_i$. Следовательно, в силу определения N и J каждое $\tau(a') \sqcap \tau(a_i) \in N$ и $\tau(a) \sqcap \wedge \tau(a'_i) \in N$. В силу первого

$$\tau(a') \sqcap [\vee \tau(a_i)] = \vee [\tau(a') \sqcap \tau(a_i)] \in J.$$

В силу последнего $\tau(a) \sqcap [\vee \tau(a_i)]' = \tau(a) \sqcap \wedge \tau(a'_i) \in N$. Следовательно, симметричная разность между $\tau(a) = \tau(\vee a_i)$ и $\vee\tau(a_i)$ лежит в J. Двойственное рассуждение приводит к аналогичному результату для пересечений.

Этот σ-гомоморфизм будет изоморфизмом, если из $\tau(a)\in J$ следует, что $a = 0$. Это мы и покажем теперь с помощью «диагонального процесса», завершая тем самым доказательство.

Если $\tau(a) \in J$, то, как и выше, $\tau(a) \leqslant \vee_i (\wedge_j (a_{ij}))$, где $\wedge_j a_{ij} =0$ для всех i. Следовательно, для любой функции $j(i)$ и силу закона изотонности

(A) $\tau(a) \leqslant \bigvee\limits_i \tau(a_{i,\, j(i)})$, **откуда** $\tau(a') \geqslant \bigwedge\limits_i \tau(a'_{i,\, j(i)})$.

Предположим теперь, что $a \neq 0$. Тогда $a' < I$, а потому

$$I > a' = a' \sqcup 0 = a' \sqcup \wedge a_{1j} = \wedge (a' \sqcup a_{1j}).$$

Следовательно; некоторое $a' \sqcup a_{1,\, j(1)} < I$. Повторяя процесс, имеем

$$I > a' \sqcup a_{1,\, j(1)} = a' \sqcup a_{1,\, j(1)} \sqcup 0 =$$
$$= (a' \sqcup a_{1,\, j(1)}) \sqcup \wedge a_{2,\, j} = \wedge (a' \sqcup a_{1,\, j(1)} \sqcup a_{2,\, j}).$$

Этим «диагональным процессом» мы получаем функцию $j(i)$ такую, что для всех n

$$I > a' \sqcup a_{1,\, j(1)} \sqcup \ldots \sqcup a_{n,\, j(n)}.$$

Эта последовательность не может содержать ни элемента a, ни пары взаимных дополнений; следовательно, мы можем образовать точку P, содержащую a и каждое $a'_{i,j(i)}$. Очевидно, что $P\in\tau(a)$, однако $P \notin \vee_i \tau(a_{i,\, j(i)})$ (ибо никакое из множеств $\tau(a_{i,j(i)})$ не содержит P и объединение в Φ есть теоретико-множественное объединение). Но это противоречит (А), чем и завершается доказательство.

6.11.12. Алгебры с мерой

В теории меры рассматривают обычно метрические булевы алгебры, в которых $v\ [0] = 0$. Их определяют обычно как булевы алгебры с положительным функционалом $v\ [x]$, удовлетворяющим условию

$$v[x \bigcup y] = v[x] + v[y], \text{ если } x \bigcap y = 0. \qquad (11)$$

Полагая в (11) $x = y = 0$, мы получаем (после выкладок) $v\ [0] = 0$. Кроме того, мы получаем по индукции

$$v[\bigvee_{i=1}^{n} x_i] = \sum_{i=1}^{n} v[x_i], \text{ если все } (x_1 \bigcup \ldots \bigcup x_{k-1}) \bigcap x_k = 0. \qquad (11')$$

Булева алгебра с мерой, обладающей счетной аддитивностью:

$$v[\bigvee_{i=1}^{\infty} x_i] = \sum_{i=1}^{\infty} v[x_i], \text{ если все } (x_1 \bigcup \ldots \bigcup x_{k-1}) \bigcap x_k = 0, \qquad (11'')$$

называется обычно *алгеброй с мерой.* Для этого необходимо и достаточно, чтобы $v\ [x]$ было непрерывным.

Любая конечная алгебра с мерой M может быть легко реализована геометрически. Если точки M суть $p_1, ..., p_r$, мы ставим в соответствие каждой точке p_k полузамкнутый интервал

$$v[p_1] + \ldots + v[p_{k-1}] \leqslant x < v[p_1] + \ldots + v[p_k]$$

на вещественной оси. Используя (11), это соответствие можно продолжить до сохраняющего меру изоморфизма («изометрического изоморфизма») между M и телом элементарных подмножеств оси x. Таким образом, *полная основа для теории меры в отношении конечного объединения, конечного пересечения и взятия дополнения, дается любой системой постулатов для булевой алгебры, если к ней добавлено* (11).

Мы построим теперь в высшей степени однородную, универсальную сепарабельную алгебру с мерой, играющую центральную роль в теории меры и являющуюся прототипом рассмотренной уже непрерывномерной проективной геометрии.

Теорема 15. *Существует с точностью до изоморфизма только одна полная сепарабельпая алгебра с мерой, не имеющая точек и удовлетворяющая условию* $v\ [I] = 1$; *обозначим ее через* $\bar{M}$. *Каждая сепарабельная алгебра с мерой является после изменения масштаба изометрически изоморфной подалгебре алгебры* $\bar{M}$. *Автоморфизмы алгебры* $\bar{M}$ *транзитивны на элементах x, отличных от* 0 *и I; для любого вещественного числа с азометричные автоморфизмы транзитивны на элементах, удовлетворяющих условию* $v\ [x] = c$.

Доказательство. Пусть A—любая сепарабельная алгебра с мерой, имеющая счетное плотное подмножество x_1, x_2, x_3, . .. Тогда A удовлетворяет условию $w[I]$ = 1 для новой оценки $w\ [x]= v[x]/v[I]$, что равносильно просто изменению масштаба. Пусть X_n обозначает конечную подалгебру алгебры A, порожденную x_1, . .., x_n; представим ее, как и выше, конечными суммами интервалов прямой линии. Объединение U алгебр X_n также может быть представлено таким же образом; кроме того, за исключением случая, когда A содержит точку, *каждая* конечная сумма интервалов прямой (с обычной мерой) является метрическим пределом элементов из U. Следовательно, если мы обозначим через M метрическую булеву алгебру конечных сумм полуоткрытых интервалов интервала [0, 1], то мы видим, что метрическое пополнение $\bar{M}$ алгебры M, не только содержит U, но и является изометрически изоморфным с $\bar{U}$ = A, если A полно и не имеет точек. То, что $\bar{M}$ есть булева алгебра, следует из р. 6.9, теорема 12; в действительности приведенное там доказательство может быть упрощено путем использования L6.

Так как замкнутые идеалы [0, x] из M являются полными и не содержат точек, они являются всегда изометрически изоморфными после изменения масштаба. Остающаяся часть теоремы 15 может быть легко доказана путем использования этого факта и того, что любой изоморфизм [0, x] с [0, y] и [0, x'] с [0, y'] может быть продолжен до автоморфизма алгебры M.

Алгебры с мерой и теория меры будут рассмотрены дальше.

6.11.13. Структуры с единственными дополнениями

Используя методы теории свободных структур, Дилуорс доказал, что *всякая структура является подструктурой структуры с единственными дополнениями.* (Структура «с единственными дополнениями» есть структура, в которой каждый элемент имеет одно и только одно дополнение.) Отсюда вытекает, что структуры с единственными дополнениями не обязаны быть дистрибутивными и даже дедекиндовыми. Тем не менее недистрибутивные структуры с единственными дополнениями не могут быть конечными.

Теорема 16. *Пусть L —произвольная полная атомная структура с единственными дополнениями. Тогда структура L изоморфна булевой алгебре всех подмножеств своих точек.*

Доказательство. Каждому множеству S точек из L сопоставим объединение $x(S)$ точек $p \in S$ и пересечение $y(S)$ дополнений p' точек $p\in S$. Из обобщенной ассоциативности будет следовать, что $x(S \cup T = x(S) \cup x(T)$ и $y(S \cup T) = y(S) \cap y(T)$. Далее, дополнение для $x(I)$ не может содержать точки (полагаем, что I обозначает как множество всех p, так и наибольший элемент в I). Следовательно, $[x(I)]'$ есть 0 и $x(I) = [[x(I)]']' = 0' = I$. Двойственно этому $y(I) = 0$.

Далее, дополнение p' любой точки p покрывается элементом I, т. е. $x\geq p'$ влечет $x= p'$ или $x=I$. Действительно, если $x \geq p$, то $x\geq p \cup p' =I$; и если $x \not\geq p$, *то* $x \cap p < p$ должно быть равно 0, в то время как $x \cup p>p' \cup p\geq = I$ откуда $x= p'$. Следовательно, а) если p и q различные точки, то $p\leq q'$. Действительно, если $p \not\leq q'$, то $p\cap q' = 0$ и $p\cup q' =I$ (так как p покрывает 0 и I покрывает q'), откуда следует, что $q' = p'$, а потому $q = p$- Из свойства «а» следует, что $x(S)\leq y(S')$.

Из этого неравенства мы выводим: б) $x(S) \cap x(S') \leqslant \leqslant y(S') \cap y(S) = y(S \cup S') = y(I) = 0.$ Таким образом, x (S) не содержит точек, не принадлежащих S, в силу чего различные множества S определяют различные $x(S)$, а потому частично упорядоченная система элементов $x(S)$ изоморфна атомной булевой алгебре всех множеств S.

Остается показать, что каждое $a\in L$ есть $x(S)$. Обозначим через S множество точек $p\in L$ таких, что $p\leq a$. Очевидно, что $a\cap x(S')$ должно содержать только точки, принадлежащие как S так и S'; следовательно, $a\cap x(S')=0$. С другой стороны, $a \cup x(S') \geqslant x(S) \cup x(S') = x(S \cup S') = I$; следовательно, a есть единственное дополнение $[x(S')]'$ элемента $x(S')$. Но таковым дополнением является в силу «б» и равенства $x(S) \cup x(S') = I$ элемент $x(S)$, чем и завершается доказательство.

Далее, структура с ортодополнениями, обладающая единственными дополнениями, является булевой алгеброй.

Теорема 17. *Если каждый элемент а в структуре L имеет единственное дополнение а' и если а→а' есть дуальный автоморфизм, то L есть булева алгебра.*

Доказательство. Покажем сперва, что $b \geq a$ влечет $(b \cap a') \cup a= b$ и двойственно, $(a \cup b') \cap b = a$. В самом деле, полагая $c = b \cap a'$, имеем $c\cap a$ 0. Точно так же $(c \cup a)' \cap b = [(b' \cup a) \cap a'] \cap b = (b' \cup a) \cap (a' \cap b) = 0,$ ибо $b' \cup a$ и $(b' \cup a)' = b \cap a'$ являются ортодополнениями. Далее, так как

$c \bigcup a = (b \bigcap a') \bigcup a \leqslant b \bigcup b = b,$ имеем, $(c \bigcup a)' \bigcup b \geqslant b' \bigcup b = I.$
Мы заключаем, что $(c \cup a)' = b'$, откуда $c \cup a = b$. Это первый результат; двойственным образом $(a \bigcup b') \bigcap b = a.$
Предположим теперь, что $x \bigcap y = a,\ x \bigcup y = b,$ и образуем $e = b' \cup (x \cap a')$. Используя дважды только что полученный результат, а также то, что $y = a \cup y = b \cap y$, имеем

$$e \bigcup y = b' \bigcup (x \bigcap a') \bigcup a \bigcup y = b' \bigcup x \bigcup y = b' \bigcup b = I,$$

$$e \bigcap y = [b' \bigcup (x \bigcap a')] \bigcap b \bigcap y = x \bigcap a' \bigcap y = a' \bigcap a = 0.$$

Отсюда следует, что y есть (единственное) дополнение элемента e, так что даже *относительные* дополнения являются единственными. Доказательство завершается ссылкой на следствие 1 из теоремы 2, р. 6.10

6.12. Приложения к теории множеств

6.12.1. Элементарные и борелевские множества

Пусть I_n обозначает замкнутый единичный n-мерный куб, т. е. множество всех точек $x=(x_1, ..., x_n)$ с n вещественными координатами x_i, $0 \leq x_i \leq 1$. Пусть $B\ (I_n)$ обозначает полную, вполне дистрибутивную булеву алгебру всех подмножеств множества 1_n.
Простейшими фигурами в 1_n являются замкнутые «n-мерные параллелепипеды» т. е. множества точек $(x_1, ..., x_n)$, удовлетворяющих соотношениям $a_k \leq x_k \leq b_k$ $[k = 1, ..., n]$ для некоторой фиксированной системы $2n$ вещественных чисел a_k, b_k.
Определение. *Булева подалгебра алгебры $B(I_n)$, порожденная n-мерными параллелепипедами, состоит из элементарных подмножеств множества I_n; σ-подалгебра алгебры, $B(I_n)$, порожденная n-мерными параллелепипедами, состоит из борелевских подмножеств множества I_n.*
Понятия элементарных и борелевских множеств играют фундаментальную роль в теории множеств.
Удаляя из замкнутого параллелепипеда грани, ребра, вершины и т. д., мы получаем его внутреннюю часть, выражаемую как разность замкнутых параллелепипедов, следовательно, любой открытый параллелепипед является элементарным множеством. Но любое открытое множество в I_n может быть представлено как объединение счетного числа открытых параллелепипедов, а именно открытых

параллелепипедов, определяемых рациональными координатами a_k, b_k. Отсюда следует, что всякое открытое или замкнутое множество есть борелевское множество. Следовательно, если мы определим борелевское множество в произвольном топологическом пространстве как множество в σ-подалгебре множеств, порожденной открытыми и замкнутыми множествами, то мы получим законное обобщение предыдущего определения.

Эвклидово n-мерное пространство не требует отдельного рассмотрения, поскольку оно гомеоморфно открытому параллелепипеду в I_n при соответствии, сохраняющем понятие координатного параллелепипеда.

6.12.2. Компактные пространства и дистрибутивные структуры

Установим соотношение между топологическими пространствами и дистрибутивными структурами; эта теория принадлежит Уолмену. Пусть I—произвольное T_1-пространство. Замкнутые подмножества из I образуют полную дистрибутивную структуру.

Теорема 1. *Всякое топологическое пространство I определяется с точностью до гомеоморфизма дистрибутивной структурой L (I) всех своих замкнутых подмножеств.*

Действительно, в силу С4 точки из I являются элементами, покрывающими 0 («атомы») в *L(I).* Замыкание *S* любого подмножества из I состоит из точек $p \in I$, которые соответствуют атомам, содержащимся в объединении (в *L(I))* атомов, соответствующих точкам из *S*.

Теорема 2. *Компактное* T_1*-пространство I определяется с точностью до гомеоморфизма любой подструктурой S структуры его замкнутых множеств, являющейся базисом.*

Набросок доказательства. Каждую точку в *I* мы можем определить как максимальный дуальный идеал в *S*, т. е. как максимальное семейство Ф базиса, никакое конечное подсемейство которого не имеет пустого пересечения. То, что пересечение такого семейства замкнутых подмножеств должно быть точкой, есть одно из определений компактности. Каждый элемент $a \in S$ должен определять «замкнутое» подмножество всех «точек» Ф, для которых $a \in \Phi$, и топология в *I* определяется тогда взятием этих подмножеств в качестве базиса замкнутых множеств.

Уолмен показал, что абстрактная дистрибутивная структура с 0 и *I* изоморфна базису замкнутых подмножеств бикомпактного

T_1-пространства тогда и только тогда, если она обладает «свойством отделимости»: для любых заданных $S > T$ существует такое X, что $S \cap X > 0$, но $T \cap X = 0$.

Кроме того, если S любая подструктура структуры всех замкнутых подмножеств T_1-пространства I, являющаяся базисом замкнутых подмножеств, то мы можем определить, исходя из S, компактное пространство I^*, как и в теореме 2. Оно будет бикомпактным пополнением Чеха пространства I и, следовательно, должно иметь ту же самую размерность (Чеха) и группу гомологии, что и пространство I.

При наличии счетного базиса мы можем выбрать S счетным и рассматривать его как предел возрастающей последовательности конечных подструктур. Это замечание приводит к данной П. С. Александровым характеристике компактных пространств со счетной базой при помощи последовательностей «конечных покрытий» (абстрактных комплексов), где каждое покрытие является подразделением предыдущего. Используя направленные множества конечных покрытий, мы получаем аналогичным образом все компактные пространства.

6.12.3. Булевы пространства и булевы алгебры

В силу теоремы 2 любое компактное T_1-пространство топологически характеризуется любым «базисом» замкнутых множеств. Но замкнутое множество имеет замкнутое дополнение тогда и только тогда, если оно открыто. Следовательно, T_1-пространство имеет булеву алгебру в качестве базиса замкнутых множеств тогда и только тогда, если оно *вполне несвязно.*

Лемма. *Компактное пространство X является вполне несвязным тогда и только тогда, если оно нульмерно.*

Доказательство. Пусть X вполне несвязно и пусть U— иобое открытое множество, содержащее точку p из X. Для каждого $q \in U'$ существует открытое и замкнутое множество $S(q)$, содержащее q, но не содержащее p. Но U' компактно (так как компактно X); следовательно, U' должно содержаться в некотором *конечном* объединении $S(q_1) \cup \ldots \cup = S(q_n)$. Его дополнение $V= \wedge S'(q_i)$ будет окрестностью точки p, содержащейся в U и имеющей границей пустое множество. Существование такого V означает в силу определения, что пространство X нульмерно. Обратно, любое нульмерное пространство является тривиальным образом вполне несвязным.

Кроме того, любая булева алгебра обладает свойством отделимости Уолмена. Следовательно, ее можно рассматривать как базис замкнутых множеств вполне несвязного компактного T_1-пространства. Обратно, тело всех открытых и замкнутых подмножеств любого такого пространства *I* образует булеву алгебру *A(1),* определяющую *I* с точностью до гомеоморфизма. Нами доказана.

Теорема 3 (Стон). *Имеет место однозначное в одну сторону соответствие между булевыми алгебрами А и вполне несвязными (т. е. нульмерными) компактными T_1-пространствами I, при котором элементы из А соответствуют подмножествам из I, являющимся одновременно открытыми и замкнутыми, а точки из I—простым идеалам в А.*

В силу этого обстоятельства Стон назвал нульмерные компактные пространства булевыми пространствами.

6.12.4. Теорема Капланского

Рассмотрим теперь структуру *L* всех вещественнозначных непрерывных функций, определенных на компактном хаусдорфовом пространстве *X*. Мы будем говорить, что простой идеал *P* в *L связан* с точкой $x \in X$, если из $f \in P$ и $g(x)<f(x)$ следует $g \in P$.

Лемма 1. *Каждый простой идеал P связан с одной и только одной точкой из X.*

Доказательство. Предположим, что *P* не связано ни с одной точкой из *X*. Тогда для каждого $x \in P$ *X* мы имели бы функции *f, g,* для которых $g(x)<f(x)$, однако $f \in P$, $g \in L — P$. Поскольку *X* компактно, а множество *Y(f,g)* элементов *y,* на которых $g(y) < f(y)$, открыто, мы имели бы конечное число пар $f_1, g_1, \ldots, f_n, g_n$ таких, что

$$X = \vee Y(f_i, g_i) \leqslant Y(\vee f_i, \wedge g_i).$$

Отсюда вытекало бы $\vee f_i > \wedge g_i$, что невозможно, так как (поскольку *P* — идеал, а *L* — *P* — дуальный идеал) $\vee f_i \in P$, а $\wedge g_i \in L\text{-}P$.

Следовательно, *P* связано по крайней мере с одной точкой $x \in X$; предположим, что *P* связано с двумя точками, скажем, *x* и *y*. Выберем $f \in P$, $g \in L — P$. Так как всякое компактное хаусдорфово пространство нормально, мы можем построить такое *h*, что $h(x) < f(x)$, $h(y) > g(y)$. Но это означало бы, что *h* принадлежит как *P*, так и *L* — *P*.

Лемма 2. *Два простых идеала P и Q связаны с одной и той же точкой $x \in X$ тогда и только тогда, если $P \cap Q$ содержит простой идеал.*

Доказательство. Предположим, что *P* и *Q* оба связаны *с x.* Выберем $f \in P$, $g \in Q$, и пусть *a*—произвольное число, меньшее чем *f(x)* и *g(x)*.

Тогда множество всех h, удовлетворяющих условию $h(x)\leq a$, образует простой идеал R, содержащийся в $P \cap Q$. Обратно, пусть P, Q, R — простые идеалы и пусть $P \cap Q \geq R$- Предположим, что они связаны с точками x, y, z. Выберем произвольные $f \in R$, $g \in L — P$; если $x= z$, то будет существовать h, для которого $h(z) < f(z)$, $h(x) > g(x)$. Тогда $h \in R \cap (L — P)$, что невозможно; следовательно, $x = z$. Аналогично, $y = z$, а потому $x=y$.

Лемма 3. *Пусть f_0 — фиксированная функция в L и S—произвольное подмножество в X. Тогда точка x принадлежит замыканию $\overline{S}$ множества S в том и только в том случае, если некоторый простой идеал $P(x)$, связанный с x, содержит пересечение $A(S)$ простых идеалов, содержащих f_0, которые связаны с точками из S.*

Доказательство. Если $x \in \overline{S}$, мы выберем $\alpha > f_0(x)$; пусть P состоит из всех f, для которых $f(x)\leq\alpha$; очевидно, что P есть простой идеал, связанный с x. Далее, $g \in A(S)$ влечет $g(y)\leq f_0(y)$ для всех $y \in S$ (ибо каждое такое неравенство определяет простой идеал, содержащий f_0), а потому $g(x)\leq f_0(x)$, откуда $g(x) \in P$. Следовательно, $A(S) \leq P$. Обратно, предположим, что $x \notin \overline{S}$; пусть P—любой простой идеал, связанный с x. Для любого $f \in L — P$ мы можем найти (снова в силу нормальности X) такую функцию g, что $g(y)$ $\inf_{z \in S} f_0(z) — 1$ на S, однако $g(x) > f_0(x)$. Тогда g принадлежит $A(S)$, но не принадлежит P.

Теорема 4 (Капланский). *Любое компактное хаусдорфово пространство X определяется с точностью до гомеоморфизма структурой L своих непрерывных функций.*

Доказательство. Назовем два простых идеала в L «эквивалентными», если их пересечение содержит третий простой идеал. Леммы 1 — 2 показывают, что классы «эквивалентных» простых идеалов в L можно толковать как «точки» пространства X. Лемма 3 показывает, что топология в X может быть выражена в терминах отношений включения между этими простыми идеалами.

Следствие 1. *X определяется кольцом R непрерывных вещественных функций на X.*

Доказательство. $f \geq g$ тогда и только тогда, если $f-g = h^2$ в R.

Следствие 2. *X определяется банаховым пространством $C(X)$ непрерывных вещественных функций на X.*

Доказательство. Пусть e—«экстремальная точка» на единичной сфере S в $C(X)$, т. е. точка, не являющаяся внутренней точкой отрезка, лежащего на S. Тогда $e(x)=1$ на S и $e(x)= —1$ на S', где S и S' дополнительные замкнутые множества. Следовательно, если положим по определению $f > g$, когда $f \neq g$ и $\| (f-g)/\|f-g\| - e \| \leqslant 1$, то мы сохраним упорядоченность на S, заменим ее обратной на S' и получим тем самым структуру, изоморфную рассмотренной выше структуре L, определив, таким образом, X.

6.12.5. Псевдодополнения и регулярные открытые множества

Пусть X — произвольное топологическое пространство и пусть — L полная дистрибутивная структура всех его открытых подмножеств. Мы установим топологический смысл результатов р. 6.10.12.

Если T открыто и $S \cap T = O$, то $\bar{S} \cap T = 0$, а потому $T \leq \bar{S}'$; в то же время $S \cap \bar{S}' \leqslant \bar{S} \cap \bar{S}' = 0$; следовательно, для любого $S \in L$ S' является псевдодополнением S^* множества S в L. Далее, так как $S^{**} = \bar{S}^{\prime -\prime}$ и $\bar{S}^{\prime -\prime}$ является в силу определения открытым ядром множества $\bar{S}$, мы видим, что $S = S^{**}$ тогда и только тогда, если S есть открытое ядро своего замыкания, т. е. так называемое «регулярное» открытое множество. В противном случае $S < S^{**}$ и граница множества S содержит внутренние точки этого множества.

Как следствие этого результата и теоремы 16 р. 6.10 получается

Теорема 5. *Пусть X—произвольное T_0-пространство. Соответствие $S \to \bar{S}^{\prime -\prime}$ есть структурный гомоморфизм структуры всех открытых множеств пространства X на полную булеву алгебру B «регулярных» открытых множеств пространства X.*

Далее, $S^{**} = I$ тогда и только тогда, если $\bar{S}^{\prime -\prime}=I$, а это эквивалентно тому, что $\bar{S}^{\prime -} = 0$, или $\bar{S}' = 0$, или $\bar{S} = I$, т. е. тогда и только тогда, если S—*плотное* открытое множество. Так как $N = S'$ замкнуто, отсюда следует, что $N^{-\prime} = (S')^{-\prime} = S^{-} = \bar{S} = I$.

Определение. *Множество T, удовлетворяющее условию $\bar{T}^{\prime -}=I$, называется нигде не плотным.*

В силу теоремы 16 из р. 6.10 плотные открытые множества, удовлетворяющие условию $S^{**} = I$, образуют дуальный идеал. Но замкнутое множество C является нигде не плотным тогда и только тогда, если $C^{\prime -} = \bar{C}' = I$, т. е. тогда и только тогда, если его открытое дополнение C' плотно. Следовательно, нигде не плотные замкнутые множества образуют идеал. Очевидно далее, что множество T нигде не плотно тогда и только тогда, если нигде не плотно его замыкание $\bar{T}$; следовательно, *нигде не плотные множества образуют идеал.*

Лемма 1. *Если X и Y открытые множества, то X** = Y** тогда и только тогда, если разность между X и Y есть нигде не плотное множество.*

Доказательство. В силу теоремы 16 из р. 6.10 $X^{**} = Y^{**}$ тогда и только тогда, если $X \cap D = Y \cap D$, где D плотное открытое множество. Но для булевой алгебры это означает, что $0 = X \cap D + Y \cap D = (X + Y) \cap D$, т. е. что (симметричная) разность $X+Y$ между X и Y лежит внутри нигде не плотного дополнения плотного открытого множества D, а это эквивалентно утверждению, что $X+Y$ нигде не плотно.

Поскольку $X^{**} \geq X > 0 = 0^{**}$, мы получаем

Следствие. *Не существует нигде не плотного открытого множества X, отличного от пустого множества.*

Теорема 6. *Если X—подмножество эвклидова пространства, не содержащее изолированных точек, то булева алгебра A «регулярных» открытых множеств в X является пополнением с помощью сечений свободной булевой алгебры B_∞ со счетным множеством образующих.*

Доказательство. Доказательство можно применить вообще к любому T_1- пространству со счетным базисом регулярных открытых множеств a_i. В самом деле, a_i и их псевдодополнения a_i^* порождают булеву алгебру, изоморфную свободной булевой алгебре B_∞ со счетным множеством образующих, которую можно также определить как предел последовательности $2 \leq 2^2 \leq 2^4 \leq \ldots \leq 2^{2^n} \leq \ldots$ Очевидно, что каждое $a \in A$ определяет сечение в B_∞; x лежит в нижней половине сечения тогда и только тогда, если $x \leq a$, и в верхней половине тогда и только тогда, если $x \geq a$. Далее, за исключением случая, когда $b \leq a$ в A, $b — \overline{a}$ является непустым (поскольку a и b регулярны) открытым множеством, которое содержит некоторое $a_i > 0$, лежащее в b, но не лежащее в a; следовательно, различные элементы в A соответствуют различным сечениям в B_∞,. Наконец, каждое сечение L, U в B_∞ соответствует некоторому $a \in A$. Действительно, возьмем множество всех $x_i \leq b$ для всех $b \in U$, где U — верхняя половина сечения. Тогда регулярная оболочка $(\vee x_i)^{**} \leq b^{**} = b$ для всех $b \in U$. Но $x \in L$ в B_∞ тогда и только тогда, если $x \leq b$ для всех $b \in U$, т. е. тогда и только тогда, если $x_i \leq b$ $(\vee x_i)^{**} \in A$, что и завершает доказательство.

Лемма 2. *Пусть X—булево пространство. Тогда структура открытых подмножеств в X изоморфна структуре идеалов булевой алгебры A подмножеств из X, являющихся открытыми и замкнутыми.*

Доказательство. Если J идеал в A, то образуем открытое объединение $S(J) = \vee S_\alpha$, замкнутых, открытых подмножеств S_α, из X, соответствующих элементам $s_\alpha \in J$. Предположим, что T открыто и замкнуто, но не является множеством S_α. Тогда никакое конечное объединение $\vee_F S_\alpha$. множеств S_α не содержит *T;* следовательно, никакое конечное замкнутое пересечение $T \cap (\vee_F S_\alpha)' = T \cap \wedge_F S'_\alpha$ не является пустым; но *X* компактно; следовательно, $T \cap \wedge_J S_\alpha'$ непусто и *T не* содержится в *S(J).* Следовательно, *J* равно идеалу *К* (*S* (*J*)) всех $T \leq S(J)$. Обратно, пусть *U*—любое открытое множество в *X; К (U)* является идеалом в *А*. Кроме того, так как *X* нульмерно, каждая точка $p \in U$ имеет открытую и замкнутую окрестность $S_\alpha (p) \leq U$. Следовательно, множество *S (К (U)),* которое содержится в *U*, совпадает с *U*.

6.12.6. Множества первой категории

Говорят, что подмножество топологического пространства *X* есть (одмножество *первой категории* (Бэр), если оно может быть представлено как сумма счетной системы нигде не плотных множеств. Очевидно, что подмножества первой категории в *X* образуют σ-идеал *J* в булевой алгебре *A* всех подмножеств *X*. Кроме того, поскольку борелевские множества порождаются открытыми множествами, из леммы 1 и индуктивных соображений следует, что каждое борелевское множество конгруентно «регулярному» открытому множеству по mod *J*. (Любое открытое множество *S* конгруентно S^{**}, в то время как замкнутое множество *С* конгруентно $C'^{*} = C^{r\prime}$ по модулю нигде не плотной границы первой категории.)

Лемма. *В полном метрическом пространстве или в локально компактном пространстве никакие два различных «регулярных» открытых множества не конгруентны по модулю множества первой категории.*

Доказательство. Мы легко сведем все к случаю регулярных открытых множеств $S > R$, $R = A \cap B$ и $S = A \vee B$. Тогда множество $S \cap R^* > 0$ будет регулярным и открытым, что позволяет нам свести рассмотрение к случаю $S > 0$. Предположим, что $S = N_1 \cup N_2 \cup N_3 \cup \ldots$, где N_i — нигде не плотные множества. Тогда $S \cap N'_1$ содержит замкнутую сферу S_i радиуса, не превышающего 1/2; по индукции $S_k \cap N'_{k+1}$ содержит замкнутую сферу S_{k+1} радиуса, не превышающего $1/2^{k+1}$. Метрический предел множеств S_k принадлежит *S,* но не принадлежит $\vee N_k$, чем завершается доказательство для случая полного метрического пространства. В локально компактном случае мы используем при доказательстве компактные S_k.

Объединяя предыдущие результаты, нами получена

Теорема 7. *Во всяком полном метрическом или локально компактном пространстве X булева σ-алгебра борелевских множеств в X по модулю множеств первой категории и структура открытых множеств по модулю нигде не плотных множеств являются обе изоморфными полной булевой алгебре «регулярных» открытых множеств пространства X.*

Следствие. *Если X—(локально) компактное метрическое пространство, удовлетворяющее второй аксиоме счетности, то фактор-алгебры теоремы* 7 *изоморфны пополнению с помощью сечений алгебры* B_∞

6.12.7. Абстрактные алгебры с замыканиями

Маккинси и Тарский провели исследование постулатов операции замыкания в топологических пространствах. Очевидно, что подмножества любого T_0-пространства образуют алгебру с замыканиями в следующем смысле.

Определение. *Алгебра с замыканиями есть булева алгебра с операцией замыкания, удовлетворяющей условиям* C1, C2, C3*, *а также условию* $\overline{0}=0$.

В любой булевой алгебре *B* соответствие $x\to x\cap a$ является для любого фиксированного $a\in B$ структурным гомоморфизмом *B* на идеал *A* всех элементов $y\leq a$. Кроме того, *A* становится алгеброй с замыканиями, если определить *относительное* замыкание каждого $y\leq a$ по формуле

$$\overline{y}_a=\overline{y}\cap a. \tag{1}$$

Лемма 1. *Соответствие* $x\to x\cap a$ *является эндоморфизмом B на идеал A, рассматриваемый как алгебра с замыканиями, тогда и только тогда, если a открыто.*

Доказательство. Условие, которое подвергается проверке, есть $\overline{x}\cap a=((x\cap a)^-)_a$, где выражение справа есть $(x\cap a)^-\cap a$. При $x=a'$ это равенство дает $a'^-\cap a=\overline{0}\cap a=0$, или $a'^-\leq a'$. Но это означает, что a' замкнуто, т. е., что a «открыто». Обратно, если a открыто, то $(x\cap a')^-\cap a\leqslant a'^-\cap a=a'\cap a=0$, откуда

$$\begin{aligned}\overline{x}\cap a&=[(x\cap a)\cup(x\cap a')]^-\cap a=[(x\cap a)^-\cup(x\cap a')^-]\cap a\\&=[(x\cap a)^-\cap a]\cup[(x\cap a')^-\cap a]=(x\cap a)^-\cap a.\end{aligned}$$

Следствие. *Если a открыто, то элементы* $x\in B$, *удовлетворяющие условию* $x\cap a=0$ *или* $x\geq a$, *образуют подалгебру с замыканиями алгебры B, изоморфную прямому произведению* 2 • A_1 *где* A_1 *алгебра с относительными замыканиями всех x, для которых* $x\leq a'$, *т. е.* $x\cap a=0$.

Действительно, это суть элементы, отображающиеся на 0 и *a* в предыдущей лемме.

Мы можем теперь построить *свободную алгебру с замыканиями с одним образующим x.* Она лежит в пересечении подалгебр, определенных в соответствии с предыдущим следствием «открытым ядром» $x^{-'}$ элемента x и «открытым ядром» $\overline{x}'$ элемента x'. Не считая компонент в 2^2, каждый ее элемент определяется поэтому своей компонентой на замкнутой *границе* $b=(x^{-'}\bigcup x^{-'})'=\overline{x}\bigcap x'^{-}$ элемента x. Но $x\bar{b}=x'\bar{b}=b$; следовательно, компоненты определяют *четырехэлементную* алгебру с замыканиями С, определяемую двумя точками p, q, для которых $\overline{p}=\overline{q}=I$. Таким образом,

Теорема 8 (Куратовский). *Свободная алгебра с замыканиями, имеющая один образующий, есть* 2^2C *и содержит шестнадцать элементов.*

Теорема 9. *Открытые элементы любой алгебры с замыканиями образуют структуру с относительными псевдодополнениями (брауэрову логику).*

«Регулярные» открытые множества образуют булеву алгебру и т. д. и т. п.

6.12. 8. Теория меры

Многое из сущности теории меры содержится в абстрактных рассмотрениях, приводимых ниже. Формулировка принадлежит Каратеодори, рассмотрение приложений к специальным случаям мы откладываем до п.9.

Предположим, что даны булева алгебра *A,* подалгебра *S* алгебры *A* и неотрицательная оценка (мера) $v[x]$, определенная на *S*. Мы допускаем $+\infty$ как возможное значение, но требуем, чтобы было $v[x]=0$. Например, *A* состоит из всех подмножеств бесконечной прямой, *S* включает в себя все конечные суммы интервалов (с концами или без них) и $v[x]$есть сумма длин интервалов, из которых состоит *x*.

Мы определим *внешнюю меру* на *A* следующим образом

$$m^*[a]=\inf\sum_{i=1}^{n}v[x_i] \text{ для } x_i\in S,\text{ где } \bigvee_{i=1}^{n}x_i\geqslant a. \tag{2}$$

Такие конечные системы элементов x_i будут называться *покрытиями a.* Независимо от того, аддитивно $v[x]$ или нет, мы имеем

$$a\leqslant b \text{ влечет } m^*[a]\leqslant m^*[b], \tag{3}$$

$$m^*[0]=0, \tag{4}$$

$$m^*[a \bigcup b] \leqslant m^*[a] + m^*[b]. \quad (5)$$

Вообще, любая вещественнозначная функция, определенная на булевой алгебре, будет называться *внешней мерой,* если она удовлетворяет условиям (3) — (5).

В такой системе мы будем называть *a измеримым,* если

$$m^*[b] = m^*[a \bigcap b] + m^*[a' \bigcap b] \text{ для всех } b \in A. \quad (6)$$

Лемма 1. *Измеримые элементы из A образуют булеву подалгебру M алгебры A, на которой m* аддитивно.*

Доказательство. В силу (6), поскольку *(a')' = a,* если $a \in M$, то $a' \in M$. Далее, если $a,c \in M$, то для всех $b \in A$ имеем

$$m^*[b] = m^*[b \bigcap a] + m^*[b \bigcap a']$$

$$= m^*[b \bigcap a \bigcap c] + m^*[b \bigcap a \bigcap c'] + m^*[b \bigcap a' \bigcap c] + m^*[b \bigcap a' \bigcap c']$$

$$\geqslant m^*[b \bigcap (a \bigcap c)] + m^*[(b \bigcap a \bigcap c') \bigcup (b \bigcap a' \bigcap c) \bigcup (b \bigcap a' \bigcap c')]$$

$$= m^*[b \bigcap (a \bigcap c)] + m^*[b \bigcap (a \bigcap c)'].$$

Следовательно, $a \cap c \in M$ и M есть подалгебра алгебры A. Аддитивность m^* для элементов $ab \in M$ следует из того факта, что если $a \cap b = 0$, то

$$m^*[a \bigcup b] = m^*[(a \bigcup b) \bigcap a] + m^*[(a \bigcup b) \bigcap a'] = m^*[a] + m^*[b].$$

В M мы называем $m^*[a]$ *мерой a* и пишем $m[a]$. Мы показали, что

$$\text{если } a \bigcap b = 0, \text{ то } m[a \bigcup b] = m[a] + m[b]. \quad (7)$$

Далее, если m^* определено посредством (2), исходя из оценки, то можно показать, что все $a \in S$ являются измеримыми (см. ниже лемму 2) и что $m[a] = \upsilon[a]$.

Мы можем определить вторую внешнюю меру на *A,* заменяя конечные суммы в (2) *счетными* суммами,

$$m^{**}[a] = \inf \sum_{i=1}^{\infty} \upsilon[x_i] \text{ для } x_i \in S, \text{ где } \bigvee_{i=1}^{\infty} x_i \geqslant a. \quad (2')$$

В точности так же, как и выше, мы устанавливаем (3) — (4); кроме того, (5) может быть усилено:

$$m^{**}[\bigvee_{i=1}^{\infty} a_i] \leqslant \sum_{i=1}^{\infty} m^{**}[a_i]. \quad (5')$$

Следовательно, к этой мере применимы (6) и лемма 1.

Мы используем теперь предположение, что $\upsilon[a]$ аддитивно.

Лемма 2. *Любое* $a \in S$ *измеримо.*

Доказательство. Для любого заданного покрытия $\vee x_i \geq b$ любого $b \in A$ мы можем построить покрытия элементов $b \cap a$ и $b \cap a'$ с помощью

$x_i \cap a$ и $x_i \cap a'$ соответственно, где по предположению $\sum v[x_i] = \sum v[x_i \cap a] + \sum v[x_i \cap a'] \geqslant$
$\geqslant m^{**}[b \cap a] + m^{**}[b \cap a']$. Так как $m^{**}[b] = \inf \sum v[x_i]$ для $\vee x_i \geq b$, выводим

$$m^{**}[b] \geqslant m^{**}[b \cap a] + m^{**}[b \cap a'].$$

Комбинируя с (5), получаем (6),
В отличие от внешней меры с конечной аддитивностью, даваемой с помощью (2), внешняя мера, даваемая с помощью (2'), но обязана удовлетворять условию $m^{**}[a] = \upsilon[a]$ для всех $a \in S$. Имеет место
Лемма 3. *Если A произвольная булева σ-алгебра с внешней мерой, удовлетворяющей условию* (5'), *то измеримые элементы образуют σ -подалгебру, в которой*

$$\text{если все } (a_1 \cup \ldots \cup a_{n-1}) \cap a_n = 0, \text{ то } m[\bigvee_{k=1}^{\infty} a_k] = \sum_{k=1}^{\infty} m[a_k]. \quad (7')$$

Доказательство. В силу леммы 1 измеримые элементы образуют подалгебру *M;* предположим, что $c_1, c_2, c_3, \ldots \in M$; если мы сможем доказать, что $\vee c_i \in M$, то этим будет показано, что *M* есть σ -подалгебра. Определим

$$d_n = c_1 \cup \ldots \cup c_n, \; c = \bigvee_{i=1}^{\infty} c_i, \; r_n = c \cap d'_n \text{ и } a_n = d_n \cap d'_{n-1};$$

таким образом, a_n являются неперекрывающимися элементами и $\bigvee_{i=1}^{n} a_i = d_n = \bigvee_{i=1}^{n} c_i$. Докажем, что $c \in M$, т. е. $m^{**}[b] = m^{**}[c \cap b] + m^*[c' \cap b]$. Случай $m^{**}[c \cap b] = \infty$ является тривиальным, поскольку тогда $m^{**}[b] \geq m^{**}[c \cap b] = \infty$. Если $m^{**}[c \cap b] \neq \infty$, то

$$\sum_{k=1}^{\infty} m^{**}[b \cap a_k] = \sup m^{**}[b \cap d_n] < \infty,$$

так что для заданногое ε>0 мы можем сделать путем выбора достаточно большого *n*

$$\sum_{k=n}^{\infty} m^{**}[b \cap a_k] < \varepsilon.$$

Следовательно, в силу теоремы 12 из р. 11 и (5')

$$m^{**}[b \cap r_n] = m^{**}[b \cap \bigvee_{k=n}^{\infty} a_k] =$$

$$= m^{**}[\bigvee_{k=n}^{\infty}(b \cap a_k)] \leqslant \sum_{k=n}^{\infty} m^{**}[b \cap a_k] < \varepsilon.$$

В силу леммы 1 каждое d_n измеримо, а потому

$$m^{**}[b] + \varepsilon > m^{**}[b \cap d_n] + m^{**}[b \cap d'_n] + m^{**}[b \cap r_n].$$

В силу (3) $m^{**}[b \cap c'] \leqslant m^{**}[b \cap d'_n]$ и в силу (5)

$$m^{**}[b \cap c] = m^{**}[b \cap (d_n \cup r_n)] \leqslant m^{**}[b \cap d_n] + m^{**}[b \cap r_n].$$

Подставляя, получаем $m^{**}[b \cap c'] + m^{**}[b \cap c] < m^{**}[b] + \varepsilon$ для всех $\varepsilon > 0$; этим доказано, что $c \in M$.

Для доказательства (7') нужно лишь заметить, что в силу (3)

$$m^{**}[\bigvee_{k=1}^{\infty} a_k] \geqslant \sup m^{**}[\bigvee_{k=1}^{n} a_k] = \sup\left(\sum_{k=1}^{n} m^{**}[a_k]\right) = \sum_{k=1}^{\infty} m^{**}[a_k],$$

в то время как обратное неравенство обеспечивается условием (5').

Лемма 4. *Если в булевой σ-алгебре внешняя мера определена посредством* (2')*, то каждый измеримый элемент конгруентен счетному пересечению счетных объединений элементов из S по модулю элемента меры нуль. Конечные объединения элементов из S плотны в M.*

Доказательство. Пусть задано $a \in M$; мы можем найти для каждого m счетное множество элементов $x_{mn} \in S$ таких, что $y_m = \bigvee_{n=1}^{\infty} x_{mn} \geqslant a$, однако

$$m^{**}[y_m] \leqslant \sum_{n=1}^{\infty} m^{**}[x_{mn}] \leqslant \sum_{n=1}^{\infty} v[x_{mn}] \leqslant m^{**}[a] + 2^{-m}.$$

Положим

$$z = \bigwedge_{m=1}^{\infty} y_m \in M;$$

очевидно, что $z \geq a$; однако $m^{**}[z] \leq m^{**}[a]$. Отсюда вытекает, что $m^{**}[(z \cap a') \cup (z' \cap a)] = 0.$

Первое утверждение этим доказано. Так как $\bigvee_{n=1}^{r} x_{mn}$ мы можем сделать сколь угодно близким к y_m, мы получаем и второй результат.

Следствие. *Если S счетно, то алгебра с мерой, определенная алгеброй M по модулю элементов меры нуль, является евпарабельной.*

6.12.9. Приложения

Наиболее важным приложением теории, изложенной в п. 8, является приложение к случаю, когда A состоит из подмножеств эвклидова n-пространства, S состоит из «элементарных» подмножеств в смысле п. 1 и $\upsilon[x]$ есть n-мерный объем множества x в элементарном смысле.

В этом случае построение, даваемое с помощью (2) и леммы 1, определяет то, что обычно называют «мерой Жордана». Построение, даваемое при помощи (2'), дает классическую меру Бореля— Лебега для n измерений. Поскольку (в силу компактности) любое покрытие замкнутого параллелепипеда элементарными множествами определяет покрытие открытыми параллелепипедами, полный объем которого больше объема данного параллелепипеда на произвольно малую величину, мы видим, что в этом случае $m^{**}[x] = \upsilon[x]$ для элементарных множеств. Таким образом, *мера Бореля—Лебега является прямым продолжением элементарного объема.*

Далее, мы могли бы заменить S *счетным* телом множеств (булевой подалгеброй алгебры A), порожденным n-параллелепипедами, ограниченными рациональными координатами, не изменяя при этом $m^{**}[x]$ в A. Следовательно, в силу следствия в конце п. 8 алгебра с мерой, определяемая измеримыми по Лебегу множествами в эвклидовом n-пространстве, является *сепарабельной.* Поскольку она не содержит неделимых элементов положительной меры, доказана

Теорема 10. *«Алгебра с мерой» измеримых подмножеств единичного куба в эвклидовом n-пространстве* $[n\geq 1]$ *по модулю множеств меры нуль есть алгебра M из теоремы* 15, *р.* 11.

Отметим, что лемма 3 из п. 8 показывает, что всякое борелевское множество измеримо, в то время как лемма 4 показывает, что, обратно, всякое измеримое множество конгруентно по модулю нуль-множества пересечению счетного числа открытых множеств, которое является борелевским множеством класса C_δ.

Следствие. *Борелевские подмножества единичного куба в n-пространстве по модулю множеств меры нуль образуют метрическое пополнение алгебры* B_∞.

Рассмотрим далее теорию Даниэля, касающуюся меры на счетномерном торе, состоящем из точек $x = (x_1, x_2, x_3, \ldots)$ $[x_i \bmod 1]$. Пусть π_n обозначает разбиения этого тора на $2^{n(n-1)/2}$ «обобщенных параллелепипедов» R:

$$k(R)/2^{n-i} \leqslant x_i < (k(R)+1)/2^{n-i} \quad [i=1, 2, \ldots, n].$$

Разбиения π_n определяют последовательность конечных булевых алгебр с мерой. Объединение их счетно; применяя теперь (2'), мы получаем меру Даниэля. Кроме тою, поскольку тор компактен, мы имеем, как и выше, $m^{**}[x]= \upsilon[x]$. Следовательно, как и выше, алгебра измеримых по Даниэлю подмножеств счетно-мерного тора по модулю множеств с мерой Даниэля, равной нулю, является изометрически изоморфной алгебре $\overline{M}$.

С другой стороны, в теории меры, связанной с теорией интегрирования Радона — Стильтьеса, получается подалгебра алгебры $\overline{M}$, в которой допускаются «атомы» положительной меры.

6.12.10. О невозможности задания меры

Мы покажем теперь, что, хотя мера с конечной аддитивностью может быть построена в любой булевой алгебре, существует много важных (бесконечных) булевых алгебр, в которых невозможно ввести меру со счетной аддитивностью. Невероятно, например, чтобы на интервале $0 \leq x \leq 1$ можно было построить меру со счетной аддитивностью, для которой все точки были бы меры нуль.

Теорема 11. *Любая мера с конечной аддитивностью, определенная на подалгебре S булевой алгебры A, может быть продолжена до меры, определенной на всем A.*

Доказательство. Поскольку наши условия имеют конечный характер, достаточно показать, что возможно продолжение с любого $S < A$ на более широкую подалгебру. Пусть задано $a \notin S;$ элементы $(a \cap b) \cup (a' \cap c)$ $[b, c \in S]$ образуют такую более широкую подалгебру. Мы определяем

$$\begin{aligned} &m_1[a \cap b] = \inf m[s] \quad \text{для} \quad s \geqslant a \cap b \quad \text{в} \quad S, \\ &m_1[a' \cap c] = \sup m[t] \quad \text{для} \quad t \leqslant a' \cap c \quad \text{в} \quad S, \\ &m_1[(a \cap b) \cup (a' \cap c)] = m_1[a \cap b] + m_1[a' \cap c]. \end{aligned} \tag{8}$$

Покажем, что это есть требуемое продолжение. Имеем

$$m_1[a \cap b] + m_1[a' \cap b] = \inf_s \{m[s] + \sup_t m[t]\}.$$

Но для заданного $s \geq a \cap b$ элемент $s' \cap b \leqslant (a' \cup b') \cap b = a' \cap b$ есть элемент t требуемого рода. Следовательно, для любого допустимого s мы имеем

$$\begin{aligned} m[s] + \sup_t m[t] \geqslant m[s] + m[s' \cap b] \geqslant m[s \cup (s' \cap b)] = \\ = m[s \cup b] \geqslant m[b]. \end{aligned}$$

Следовательно, $m_1[a \cap b] + m_1[a' \cap b] \geqslant m[b]$ для любого $b \in S$.

Двойственным путем мы получаем обратное неравенство; этим доказано, что $m_1[b]=m\,[b]$.

Остается доказать, что m_1 аддитивно; достаточно показать, что a-компоненты аддитивны, т. е. что для всех $b,c \in S$

$$\text{из } (a \cap b) \cap (a \cap c) = 0 \text{ следует } m_1[(a \cap b) \cup (a \cap c)] = \\ = m_1[a \cap b] + m_1[a \cap c].$$

Для любого $\varepsilon > 0$ мы можем найти в S $t \geq a \cap b$, $u \geq a \cap c$, для которых $m[t] < m_1[a \cap b] + \varepsilon$, $m[u] < m_1[a \cap c] + \varepsilon$, следовательно,

$$m[t \cup u] \leqslant m[t] + m[u] \leqslant m_1[a \cap b] + m_1[a \cap c] + 2\varepsilon.$$

Так как $t \cup u \geqslant (a \cap b) \cup (a \cap c)$, $m_1[(a \cap b) \cup (a \cap c)] \leqslant m_1[a \cap b] + m_1[a \cap c]$. Обратно, для любого $v \geqslant (a \cap b) \cup (a \cap c) = a \cap (b \cup c)$ в S мы имеем $m[v] \geqslant m[v \cap b \cap c'] + m[v \cap c]$, так как элементы $v \cap b \cap c'$ и $v \cap c$ имеют в пересечении нуль. Но так как $a \cap c \cap c' = 0 = a \cap b \cap c$, имеем

$$v \cap b \cap c' \geqslant a \cap (b \cup c) \cap (b \cap c') = \\ = (a \cap b \cap c') \cup (a \cap b \cap c) = a \cap b.$$

Следовательно, $m[v \cap b \cap c'] \geqslant m_1[a \cap b]$; аналогично, $m[v \cap c] \geq m_1[a \cap c]$. Следовательно, каждое $m[v] \geqslant m_1[a \cap b] + m_1[a \cap c]$ и

$$m_1[(a \cap b) \cup (a \cap c)] = \inf m[v] \geqslant m_1[a \cap b] + m_1[a \cap c],$$

что и требовалось доказать. Мы переходим теперь к отрицательным результатам.

Теорема 12. *Во всякой алгебре с мерой мы имеем для любого счетного множества элементов* x_i [i = l, 2, 3, ...]

$$\inf_{0 < y_i \leqslant x_i} \left(\bigvee_{i=1}^{\infty} y_i\right) = 0,$$ *если не существует точек (положительной меры).* (9)

Кроме того, если для всех i $\vee_i a^i_j = I$ *и* $m[I] > 0$, *то*

$$m\left[\wedge_i \left(\bigvee_{j=1}^{r(i)} a^i_j\right)\right] > 0$$ *для некоторой функции* $r(i)$ (10)

Доказательство формулы (9). Если точек не существует, то мы можем найти для любого $\varepsilon > 0$ такие y_i что $0 < y_i \leq x_i$ и $m[y_i] < \varepsilon/2^i$.

Отсюда следует, что $m\left[\bigvee_{i=1}^{\infty} y_i\right] \leqslant \sum_{i=1}^{\infty} m[y_i] < \varepsilon$. Следовательно,

меняя выбор элементов y_i, имеем $m\left[\wedge \left(\bigvee_{i=1}^{\infty} y_i\right)\right] < \varepsilon$ для всех $\varepsilon > 0$, а

потому $\wedge \left(\bigvee_{i=1}^{\infty} y_i\right) = 0.$

Доказательство формулы (10). Для каждого $\varepsilon>0$ и i мы можем выбрать столь большое $r(i)$, что $m[\bigvee_{j=1}^{r(i)} a_j^i] > m[I] - \varepsilon/2^i$. Отсюда следует, что

$$m[\wedge_i \bigvee_{j=1}^{r(i)} a_j^i] > m[I] - \varepsilon.$$

Этим доказано (10), если $m[I] > 0$; доказано даже, что

$$\sup_{r(i)} \{m[\wedge_i \bigvee_{j=1}^{r(i)} a_j^i]\} = m[I]. \qquad (10')$$

Предыдущие результаты могут быть получены более просто, если мы отождествим элементы, симметричная разность которых имеет меру нуль.

Следствие 1. *Алгебра борелевских множеств по модулю множеств первой категории не изоморфна никакой алгебре с мерой.*

Действительно, если мы возьмем в качестве x_i интервалы с рациональными концами, то $\vee y_i = I$ при любом выборе элементов y_i таких, что $0< y_i \le x_i$. Сравнивая со следствиями из теорем 7, 10, мы получаем.

Следствие 2. *Пополнение с помощью сечений метризуемой свободной булевой алгебры со счетным множеством образующих не изоморфно ее метрическому пополнению.*

Теорема 13. *Если гипотеза континуума верна, то невозможно определить нетривиальную меру со счетной аддитивностью для всех подмножеств континуума такую, что каждая точка имеет меру нуль.*

Доказательство. Образуем класс всех однозначных функций α: $\alpha(i) = j$, определенных для целых положительных чисел и имеющих своими значениями целые положительные числа. Таких функций имеется $2^{\aleph_0}$; следовательно, если гипотеза континуума верна, то мы можем так вполне упорядочить множество A функций α, что каждое α имеет счетное множество предшествующих. Отбросим теперь все такие α, что для некоторого $\beta<\alpha$ в A $\alpha(i)<\beta(i)$ для всех i. Остающееся множество B неотбрасываемых β обладает тем свойством, что каждому $\alpha \in A$ соответствует $\beta \in B$ такое, что $\beta\le \alpha$ и $\alpha(i)\le\beta(i)$ для всех i (так как отношение транзитивно и A вполне упорядочено). Далее, если бы B было счетным, то мы могли бы перенумеровать его члены: $\beta_1, \beta_2, \beta_3, \ldots$; определив $\beta^*(k) = \beta_k(k) + 1$, мы получили бы новое неотбрасываемое β^* в противоречие с предположением. Следовательно, B несчетно. Поэтому в силу гипотезы континуума имеет место взаимно однозначное соответствие между B и континуумом, так что точки континуума можно обозначать через $p_\beta[\beta \in B]$. Пусть теперь X^i_k обозначает множество всех точек p_β с $\beta(i)=k$; очевидно, что $\vee_{k=1}^{\infty} X^i_k = I$ для всех i, ибо $\beta(i)$ принимает некоторое значение. Следовательно, в силу (10) для некоторого $r(i)$ имеем $m[\wedge_i(\vee_{j=1}^{r(i)} X^i_j)] > 0$. Но некоторое $\beta_0 \in B$ удовлетворяет неравенству $\beta_0(i) \geq r(i)$ для всех i. Но из $\beta(i) \leqslant r(i) \leqslant \beta(i)$ для всех i должно следовать $\beta \leq \beta_0$ в B; множество этих элементов, предшествующих β_0, является всегда счетным. Следовательно, $\wedge_i(\vee_{j=1}^{r(i)} X^i_j)$ содержит только *счетное* множество точек p_β и должно иметь меру нуль. Это противоречие и доказывает теорему.

Улам доказал невозможность введения такой меры при несколько более слабых предположениях.

6.13.Приложение к логике и к теории вероятностей

6.13.1. Алгебра свойств

Настоящий раздел будет посвящен лишь алгебраической стороне логики и теории вероятностей. Мы не будем пытаться толковать здесь более спорные стороны этих областей, а также разрешать парадоксы теории множеств. Поскольку лишь два наших результата являются теоремами теории структур, многое из приводимого материала будет дано только в общих чертах.

Мы начнем с описания булевой алгебры свойств. Понятие *свойства* (называемого также «атрибутом» или «качеством») объекта является столь основным, что невозможно никакое его определение в терминах более фундаментальных. Свойства обычно обозначаются именами прилагательными (а именно, красный, жидкий, мертвый и т. д.) или родовыми именами существительными (а именно, животное, дерево, океан и т. д.). Легко видеть, что в своем логическом эффекте имена прилагательные и родовые имена существительные эквивалентны, — сказать «вода жидкая» все равно, что сказать «вода есть жидкость».

Свойства можно *комбинировать* с помощью союзов *и, или* (а именно, красный и жидкий, красный или жидкий); можно также построить отрицание любого свойства путем использования слова *не* (а именно, не красный). Можно также установить между свойствами отношение включения; так, например, свойство быть океаном включает в себя свойство быть жидким («каждый океан жидкий»).

Первый закон Буля. Пусть свойства обозначаются буквами, «и», «или», «не» — соответственно символами $\cup$, $\cap$, $'$, и «каждое *x* есть *y*» — посредством *«x≤y»*. Тогда свойства образуют булеву алгебру.

Это утверждение может быть проверено путем сравнения LI — L7 с правилами выводов, даваемыми в учебниках формальной логики, или же путем непосредственного анализа. Читатель без труда убедится, что это не есть дедуктивное «доказательство» математической теоремы. В действительности мы лишь переформулировали некоторые из основных предложений классической логики, приспособив их для математического анализа.

Выведем теперь некоторые следствия из первого закона Буля и предыдущих результатов. Так, теория постулатов показывает, что для формальной логики могут быть найдены эквивалентные системы аксиом, значительно различающиеся по внешнему виду. Например, «или» и «не» можно определить в терминах одного только «и».

Далее, из заданных *n* свойств может быть составлено путем использования «и», «или» и «не» в точности 2^{2^n} различных свойств. Наконец, *всякое* тождество, имеющее место для конечных «булевых комбинаций» свойств с помощью «и», «или» и «не», будет справедливо в булевой алгебре. Действительно, свободная булева алгебра с *n* образующими может быть реализована множествами и, следовательно (см. ниже, п. 2), при помощи свойств, относящихся к соответствующим множествам. Следовательно, любое (конечное) тождество, неверное для какой-нибудь булевой алгебры, не будет верным для свойств вообще.

На самом деле, легко показать, используя р. 6.11.8, что конечные тождества булевой алгебры являются максимальными в том смысле, что если присоединить любое независимое тождество, то *каждое* тождество может быть доказано. Этот результат можно истолковать следующим образом. Классическую логику свойств невозможно усилить, не приходя при этом к абсурду; ее можно только ослабить.

6.13.2. Булевский дуальный изоморфизм

Естественно отождествить каждое свойство *x* с классом € всех объектов (или «предметов»), обладающих этим свойством. Кроме того, в силу определения теоретико-множественного произведения множество объектов, обладающих как свойством *x,* так и свойством *y,* является теоретико-множественным произведением множества € и множества €. Аналогичными рассуждениями для двух других случаев мы докажем

$$\widehat{x\bigcup y}=\hat{x}\bigcap\hat{y},\quad \widehat{x\bigcap y}=\hat{x}\bigcup\hat{y},\quad (\widehat{x'})=(\hat{x})'; \qquad (1)$$

соответствие $x\to$ € есть дуальный гомоморфизм.

Обратно, каждому классу объектов X мы можем сопоставить свойство a_X «быть членом класса X»; кроме того, $€_X = X$ для всех X. Мы можем говорить, что два свойства «объектно-эквивалентны», если они определяют один и тот же класс; это приводит к следующему результату Буля .

Второй закон Буля. Соответствие $x\to$ € есть дуальный изоморфизм между булевой алгеброй классов и булевой алгеброй свойств при условии, что объектно-эквивалентные свойства отождествлены.

Этот дуальный изоморфизм, доведенный до его логического конца, содержит в себе умозаключения, отвергаемые здравым смыслом, равно как и хорошо известные математические парадоксы.

Так, например, понятие «класс всех объектов» явно не имеет смысла. Не будучи всеведущими, мы попросту не знаем, что это такое; например включал ли он в себя мезоны в 1850 г.? Далее, «класс всех порядковых чисел» является вполне упорядоченным; обозначим через τ его порядковый тип; τ не может существовать (в силу теоремы 4 из р. 6.4 и следующего за ней рассуждения); это является математическим парадоксом, по крайней мере если допускать аксиому выбора.

Математиками-логиками предлагались различные правдоподобные допущения с целью устранения таких парадоксов; каждое допущение имеет свои достоинства, но ни для одного из них не была доказана его необходимость. Эти допущения ниже рассматриваться не будут.

Вместо этого мы рассмотрим в п. 4 —9 различные альтернативные модели для алгебры логики, которые могут быть обоснованы с помощью физических или физиологических аргументов. Вполне возможно, что различные физические теории могут допускать различные алгебры логики. В самом деле, такая возможность широко допускалась по отношению к теории относительности и квантовой механике. Окончательное суждение по поводу относительной пригодности этих небулевых алгебр логики не установится еще, возможно, в течение ряда лет.

6.13.3. Исчисление высказываний

Прежде чем сосредоточить свое внимание на небулевых алгебрах логики, мы опишем булеву алгебру высказываний.

Синонимами для высказывания являются: «суждение», «утверждение», «теорема». Исчисление высказываний имеет дело с составными суждениями подобно: «Иван спит и Петр гуляет», «Иван спит или Петр гуляет», «Иван не спит».

Для любых двух высказываний x и y высказывания «x и y», «или x или y (или оба)» и «не x» можно обозначать, соответственно, через $x \cup y$, $x \cap y$ и x'. При этой системе обозначений тождества булевой алгебры устанавливают логически эквивалентные утверждения. Так, например, L2: утверждение «Иван спит и Петр гуляет» верно или не верно в соответствии с тем, верно или не верно утверждение «Петр гуляет и Иван спит». В итоге мы получаем

Третий закон Буля. Высказывания образуют булеву алгебру.

Приложения булевой алгебры к исчислению высказываний можно получить аналогично описанию, приведенному в конце п. 1.

В классической логике всякое высказывание либо верно, либо неверно. Кроме того, $x \cup y$ верно тогда и только тогда, если верны оба высказывания x и y; $x \cap y$ верно, если верно x или y; из двух высказываний x и x' одно верно, а другое неверно. Следовательно, мы получаем

Четвертый закон Буля. Верные высказывания образуют (собственный) *простой идеал* в булевой алгебре всех высказываний: дополнительный дуальный идеал состоит из неверных высказываний.

Получаемое в результате исчисление высказываний может быть применено *только* к таким системам, в которых верность или неверность каждого высказывания доказуемы; таким образом, оно не должно применяться к явлениям, где очевидность ограничена. В таких системах сложное высказывание «x влечет y» («если x, то y»), что обозначается через $x \to y$, имеет весьма специальный смысл. Очевидно, что высказывание $x \to y$ верно или неверно в соответствии с тем, верно или неверно высказывание «y или не x»; таким образом, $x \to y$ можно отождествить с $x' \cap y$, *если* отождествлять высказывания, являющиеся оба верными или оба неверными. Это приемлемо, если идет речь лишь об определении класса всех верных высказываний в отдельной системе.

В таких системах высказывание *«x* эквивалентно *y»* можно обозначить через $x \sim y$ и отождествить его с высказыванием *«x* влечет *y* и *y* влечет *x»,* т. е. с симметричной разностью $(x \to y) \bigcup (y \to x) = x \dotplus y$ между x и y. Можно показать также, что многие сложные высказывания p являются «тавтологиями», т. е. верны благодаря одной их логической структуре. Алгебраически это равносильно тому, что сказать $p \sim 0$. Простейшей тавтологией является $x \mid x'$ («*x* или не *x*). Рассматривая булеву алгебру, легко показать, что следующие высказывания также являются тавтологиями:

$$I \to x,\ \ x \to 0,\ \ x \to x,\ \ x \sim x,\ \ x \to (y \to x),$$
$$(x \sim y) \sim (y \sim x),\ \ (x \sim 0) \bigcap (x \sim I),\ \ (x \to y) \bigcap (y \to x), \quad (2)$$
$$[(x \to y) \bigcup (y \to z)] \to (x \to z),\ \ (x' \to I) \to x.$$

Вопрос о законности принятия этих тавтологий в математической логике будет обсуждаться в **п.** 6.

6.13.4. Модель из классической механики

В силу теоремы 13 из р. 11 отождествление свойств с классами, принимаемое в классической формальной логике (второй закон Буля), равносильно предположению справедливости обобщенного дистрибутивного закона (22') р. 10. В силу теоремы 16 р. 11 это равносильно предположению, что существуют *максимальные* свойства, верные для одного единственного элемента (следовательно, соответствующие его собственному наименованию); это предположение мы можем назвать гипотезой атомистичности.

Как мы теперь увидим, гипотезу атомистичности можно, однако, избежать в типичных физических приложениях. Рассмотрим проблему n тел, как ее обычно трактуют (эти тела могут представлять собой либо планеты и их спутников (небесная механика), либо молекулы газа (статистическая механика). «Состояние» системы $\sum$ n тел в любой момент времени t_0 может быть выражено с помощью $6n$ вещественных чисел: каждое тело имеет три координаты положения и три координаты скорости. Кроме того, это описание является полным в том смысле, что состояние системы $\sum$ в любой последующий (или предшествующий!) момент времени t определяется этими числами и законами взаимного притяжения и отталкивания.

Таким образом, «состояние» (или фаза) системы $\sum$ может быть представлено точкой в $6n$-мерном координатном пространстве; пространство всех таких точек называется *фазовым пространством I* системы $\sum$. Каждое свойство системы $\sum$ определяет множество в этом фазовом пространстве: множество всех «состояний», в которых $\sum$ имеет заданное свойство. Таким образом, в соответствии с булевой логикой должно иметь место взаимно однозначное соответствие между подмножествами пространства I и свойствами системы $\sum$.

Но это является физически абсурдным. Так, поскольку точность измерений ограничена, мы никогда не можем экспериментально установить, является ли кинетическая энергия системы $\sum$, измеренная в эргах, рациональным числом.

Это соответствие может быть также математически неудобным, поскольку оно, повидимому (теорема 13 р. 12), несовместимо с принципом, существенным в статистической механике, что *каждое* важное свойство имеет вероятность, где вероятность счетно-аддитивна.

Не продолжая обсуждения, мы можем утверждать, что требования математической последовательности и физической правдоподобности заставляют рассматривать «свойства», соответствующие измеримым подмножествам, игнорируя множества меры нуль.

(Дж. Нейман отметил, что, поскольку множество измеримо тогда и только тогда, если оно почти всюду имеет плотность 0 или 1, мы отбираем в точности те свойства, о верности или лживости которых можно высказаться сколь угодно близко к достоверности, производя достаточно точные измерения. Точно так же все известные свойства (а именно, свойства иметь в заданных пределах температуру, давление и т. д.) соответствуют борелевским (и, следовательно, измеримым) множествам.)

Получаемая в результате система является универсальной сепарабельной алгеброй с мерой из теоремы 15, р. 11. Она является *непрерывной* булевой алгеброй, чем и избегается внушающая сомнения гипотеза атомистичности. Однако в конечной области она является строго булевой логикой; она удовлетворяет даже бесконечным дистрибутивным законам р. 11, теорема 12. Эта система допускает нетривиальную счетно-аддитивную функцию вероятности, будучи в действительности моделью, на которой базируется большая часть теорий вероятностей (см. п. 9).

6.13.5. Модель из квантовой механики

Логика квантовой механики является более сложной. В соответствии с общепринятой математической теорией «состояние» системы может быть представлено точкой ψ в комплексном гильбертовом пространстве, выступающем, таким образом, как «фазовое пространство» I. Кроме того, наблюдаемые величины соответствуют самосопряженным операторам A на I.

Если A имеет *дискретный* спектр, то теория свойств, определимых в терминах оператора A, является простой. Имеется счетное множество ортогональных собственных состояний ψ_1, ψ_2, ψ_3,…, таких, что $\psi_i A= \lambda_i\psi_i$. Любое заданное «состояние» ψ можно представить в виде $\psi=\sum c_i\psi_i$, где ψ_i характеристические состояния, связанные с различными собственными значениями λ_i. *Априорная* вероятность того, что наблюдение ψ даст величину λ_i, есть тогда $|c_i|^2 = c_ic_i^*$. Наиболее общее физическое свойство, определимое в терминах оператора A, имеет вид

$$\lambda \in L,\ L \text{ —любое подмножество спектра.} \qquad (3)$$

Это означает «верное или неверное» *апостериорное* утверждение. Соответствующая *априорная* вероятность того, что наблюдение ψ даст нам $\lambda \in L$, есть тогда

$$P_L = \sum_{\lambda_i \in L} |c_i|^2.$$

Это выражение обладает формальными свойствами вероятности (см. п.9).

Очевидно, что *утверждения,* определимые в терминах оператора A, образуют атомную булеву алгебру всех подмножеств счетного множества; в то же время *предсказания* в отношении A являются чисто статистическими.

Предположим теперь, что A имеет непрерывный спектр; случай x-координаты положения является типичным. Каждому элементарному утверждению вида $x> \alpha$ (α— любое вещественное число) соответствует *замкнутое подпространство* S_α гильбертова пространства; пусть E_α обозначает оператор ортогонального проектирования на S_α. Для заданного ψ *предсказание*

$$x \in L_\alpha, \text{ где } L_\alpha \text{ есть множество чисел } x > \alpha, \qquad (3')$$

имеет вероятность быть выполненным, даваемую выражением

$$P_\alpha = |\psi E_\alpha|^2 = \psi E_\alpha E_\alpha^* \psi^*. \qquad (4')$$

Относительно операций счетного объединения, пересечения и взятия дополнения множества L_α порождают σ-тело Φ всех борелевских подмножеств L спектра (см. р. 6.13.1). Соответствующие операции на полной структуре замкнутых подпространств гильбертова пространства дают булеву алгебру замкнутых подпространств S_L, которая есть фактически Φ по модулю множеств меры нуль (как и в п. 4) Ортогональные проектирования E_L на S_L порождаются, в свою очередь, операторами E_α посредством умножения, сложения по модулю два («булево сложение», см. р. 6.12.3) и перехода к пределу;

получаемое в результате булево кольцо изоморфно также неатомной булевой алгебре из п. 4.

Следовательно, если *A* имеет непрерывный спектр, то *апостериорные* утверждения относительно *A* соответствуют борелевским подмножествам *L* спектра по модулю множеств меры нуль. Априорная вероятность предсказания $\lambda \in L$ дается в квантовой теории посредством

$$P_L = |\psi E_L|^2 = \psi E_L E_L^* \psi^*. \qquad (4'')$$

Более обще, если *A, B, C, ...* любое множество перестановочных наблюдаемых, то мы знаем, что каждое утверждение *L* соответствует замкнутому подпространству S_L гильбертова пространства.

При отсутствии физической очевидности мы составим следующую гипотезу.

Гипотеза. Линейная сумма любых двух замкнутых подпространств, соответствующих наблюдаемым свойствам, сама соответствует наблюдаемому свойству.

Отсюда будет следовать, что любая линейная сумма $X \cup Y$, ортогональное дополнение X' и пересечение $X \cap Y = (X' \cup Y')'$ таких замкнутых подпространств соответствуют наблюдаемым свойствам. Таким образом, *логика квантовой механики является дедекиндовой структурой с ортодополнениями.*

Итак, дистрибутивный закон логики теряет силу даже в конечной области. (Бесконечная дистрибутивность отсутствовала уже в модели п. 4.) Характерно, что *дистрибутивный закон выполняется для одновременно наблюдаемых свойств, но ни для каких других.* (Две величины являются «одновременно наблюдаемыми в квантовом механике тогда и только тогда, если соответствующие линейные операторы перестановочны. Заслуживает внимания то, что эти результаты остались бы справедливыми, если бы мы заменили предыдущую гипотезу гипотезой, что каждое замкнутое подпространство гильбертова пространства соответствует наблюдаемому свойству. В любом случае априорные вероятности даются, по-видимому, с помощью (4).

6.13.6. Возражения в отношении булевой логики

В п. 4—5 мы рассмотрели алгебры свойств, подсказанные нуждами теоретической физики. Мы рассмотрим теперь исчисления высказываний, подсказываемое чисто метафизическими возражениями в отношении алгебры логики Буля—Уайтхеда, описанной в п. 1—3.

Так, например, принцип, что «ложное утверждение влечет всякое утверждение», выражаемый в виде тавтологии $I \rightarrow x$ (см. п. 3), представляется неудовлетворительным. То же можно сказать в отношении тавтологии $(p \rightarrow q) \cap (q \rightarrow p)$, которая утверждает, что «из двух любых высказываний p и q —либо p влечет q, либо q влечет p».

В том же скептическом направлении можно подвергнуть сомнению справедливость доказательств путем сведения к противоречию. Почему из опровержения (или «сведения к абсурду») высказывания «не *р»* должна следовать справедливость высказывания «*р*»? Такие доказательства от противного представляются, в частности, неудовлетворительными, когда «*р*» утверждает существование числа, но опровержение высказывания «не *р»* не указывает никакой процедуры для его отыскания.

Брауэр и его «интуиционистская» школа отвергают все такие «неконструктивные» доказательства. И в самом деле, невозможно отстаивать доказательства от противного, если допускать существование «неразрешимых» высказываний, для которых ни *«р»,* ни «не *р»* не являются доказуемыми. Имеется некоторое основание для существования неразрешимых высказываний.

Так, Сколем и Гедель построили правдоподобную и последовательную логическую систему, в которой такое неразрешимое высказывание существует в отношении обычных целых чисел! Доказательство существования само является, однако, неконструктивным и основывается на допущении существования несчетного множества «высказываний», но только счетного множества «доказательств».

Хотя и неизвестны специфически неразрешимые «подлинно математические» высказывания, вполне возможно, что гипотеза континуума (и аксиома выбора) являются неразрешимыми в следующем точном смысле. Может существовать одна, вполне согласованная с логикой, система трансфинитных чисел, в которой гипотеза континуума верна, и другая система, в которой она ложна. Иными словами, обычная логика может не быть *категоричной* — она может не определять арифметики трансфинитных чисел с точностью до изоморфизма.

Однако можно высказать мнение, что в конечном счете не будет принята никакая из логических систем, допускающих неразрешимые высказывания.

6.13.7. Логика Брауэра: отношение «строго влечет» Левиса

Исходя из этих метафизических рассмотрений, была развита заслуживающая внимания алгебра логики, в которой тождество $(x')'=x$ заменено более слабым условием $(x')'\leq x$. Это значит, что допускается $x\to(x')'$, но не требуется $(x')'\to x$. А именно, имеем данную Рейтингом формулировку интуиционистской или брауэровой логики, которая может быть выражена следующим образом.

Определение. *Брауэрова логика есть структура, дуальная структуре с относительными псевдодополнениями (см. p.* 6.10.12).

В такой структуре мы обозначим выражение, двойственное $a*b$, через $a\to b$ и отношение $x\to 0$ (т. е. тавтологии) через $\vdash x$. Поскольку $x\to y$ есть наименьший элемент t такой, что $x \cup t = y$, то, очевидно, $\vdash x \to y$ тогда и только тогда, если $x\geq y$. Используя это отношение, легко получить постулаты Рейтинга:

$$\vdash a \to (a \cup a),\ \vdash a \cup b \to b \cup a,\ \vdash (a \to b) \to (a \cup c \to b \cup c),$$
$$\vdash [(a \to b) \cup (b \to c)] \to (a \to c),\ \vdash b \to (a \to b),$$
$$\vdash [a \cup (a \to b)] \to b,\ \vdash a \to (a \cap b),\ \vdash a \cap b \to b \cap a,$$
$$\vdash [(a \to c) \cup (b \to c)] \to (a \cap b \to c),\ \vdash a^* \to (a \to b),$$
$$\vdash [(a \to b) \cup (a \to b^*)] \to a^*.$$

Отсюда следует, что систему замкнутых подмножеств любого топологического пространства (технически, T_0-пространства) можно рассматривать как полную брауэрову логику. Обратно, любую полную атомистичную брауэрову логику можно было бы представить как изоморфную системе замкнутых подмножеств подходящего T_0-пространства.

В силу теоремы 16 В. И. Гливенко из п. 10 можно также любой брауэровой логике сопоставить булеву логику.

Другая алгебра логики возникает в связи с исчислением отношения «строго влечет» Левиса, которое может быть определено следующим образом.

В любой булевой алгебре A мы можем определить, как и выше, $p \to q = p' \cap q$ (словами, p влечет q «материально»). По определению, $p\ \mathrm{p}\ q$ (словами, p «строго» влечет q) есть 0, если $p\geq q$ (т. е. если $p'\cap q=0$), и I в противном случае. Таким образом, элементы $p\to q$ и $p\,\mathrm{p}\,q$ эквивалентны в двуэлементной булевой алгебре, но они не эквивалентны в общем случае даже если $p\ \mathrm{p}\ q = 0$. Мы можем также определить $\Diamond p$ как элемент p р I и получить алгебраическую систему.

6.13.8. Модальная логика

Вообще говоря, исчисление высказываний Буля — Уайтхеда относится к двузначной дедуктивной логике, в которой каждое высказывание либо доказуемо верно, либо доказуемо ложно. Альтернативное исчисление, рассмотренное в п. 5 и 7, может иметь место только тогда, когда допускаются и другие категории высказываний. Теории логики, допускающие более чем две категории «истинных» и «ложных» высказываний, составляют то, что обычно называют «модальной» логикой, а допускаемые ими категории называются «модусами» или «степенями правдоподобности».

Модальная логика является очень древней. Так, аристотелевская логика признавала четыре модуса: неизбежный, условный, возможный, невозможный. Современная логика придает значение трем: верный, неразрешимый, ложный. В теории вероятностей шкала степеней правдоподобности пробегает все числа от нуля до единицы.

Алгебраические системы, описывающие исчисление высказываний с просто упорядоченными множествами модусов, были предложены Лукашевичем и Тарским, а также Постом. Для подробного их изучения мы отсылаем читателя к обширной литературе.

Большинство систем модусов, изучавшихся в прошлом, были просто упорядоченными по степени правдоподобности, которую они приписывали высказываниям. Все остальные образовывали дистрибутивные структуры и, следовательно, полупрямые произведения двузначных логик. Представлялось бы желательным достроить исчисление высказываний, базирующееся на недистрибутивных структурах степеней правдоподобности, — скажем, на двух недистрибутивных структурах из пяти элементов. При попытках это сделать возникает проблема о том, каким образом степени правдоподобности высказываний p и q могут определять степени правдоподобности $p \to q$ и p'.

6.13. 9. Вероятность и мера

Все говорят о вероятности, но никто не может сказать в удовлетворительной для других форме, что это такое. Мы пудем следовать наиболее влиятельной математической школе и отождествим вероятность с мерой на базе постулатов.

Определение. *Алгебра вероятности есть алгебра с мерой* (р. 6.11.12), *в которой* $p[I]$ = 1.

Из любой алгебры с мерой, в которой $m[I] \neq 0$, можно получить алгебру вероятности, полагая $p\ [x] = m\ [x]/m\ [I]$. Следовательно, обе теории имеют одинаковое распространение с чисто алгебраической точки зрения. Используя «геометрические вероятности», можно дать даже непосредственное отождествление с помощью следующего примера.

Пример 1. Пусть круглый диск радиуса $1/2\pi$ приводится во вращение. Для каждого подмножества S края диска пусть $p\ [S]$ есть вероятность того, что, когда диск остановится, радиус, проведенный в фиксированном (скажем, восточном) направлении, будет пересекать периферию в точке множества S. Очевидно, что $p\ [S]$ есть просто мера множества S.

Можно было бы возразить, что вследствие ограниченной точности измерений этот эксперимент не может быть физически выполнен. Однако с помощью конструкции, приведенной в р. 6.11.12, второй абзац, мы можем реализовать по такому способу с произвольной точностью любую конечную алгебру вероятности. Физический же смысл бесконечного с точки зрения, опыта является крайне сомнительным. С точки зрения строго эмпирической наиболее общим примером является, повидимому, следующий.

Пример 2. Пусть повторяющийся эксперимент E имеет в качестве возможных результатов $h_1,..., h_m$. Пусть A—булева алгебра всех свойств X, определяемых с помощью булевых операций, примененных к h_i. Наконец, пусть $p_n[X]$ обозначает ту часть, которую среди первых n повторений эксперимента E составляют повторения, обладающие свойством X. Тогда для каждого n $p_n\ [X]$ является вероятностью.

Представим теперь себе эксперимент, повторяющийся последовательно бесконечное число раз. Мы можем назвать свойство X *статистически регулярным,* если при $n\to\infty$ $p_n\ [X]$ стремится к конечному пределу $p_\infty[X]$, называемому *частотой* X. Можно показать тогда, что статистически регулярные свойства образуют булеву подалгебру, на которой p_∞ обладает *конечной* аддитивностью. Однако в общем случае p_∞ может и не обладать счетной аддитивностью.

Вероятность того, что «выбранное наугад» целое положительное число будет принадлежать множеству S, определяется вполне аналогично. Пусть $p_n[S]$ обозначает часть, которую составляют среди первых n целых чисел числа, принадлежащие S. Тогда предел $p_\infty\ [S]$ чисел $p_n\ [S]$ определяет конечно-аддитивную меру.

(Это есть основа теории вероятностей согласно школе Мизеса**.** Приводимые примеры могут объяснить, почему некоторые ученые не включают в рассмотрение постулата счетной аддитивности.)

Используя чезаровские средние, можно значительно расширить класс измеримых множеств. Является ли это «вероятностью»? По мнению ряда авторов, это дело вкуса; но данное определение дает отрицательный ответ.

Во многих случаях мы можем отождествить понятие вероятности как частоты с понятием вероятности как меры. Эти случаи подобны следующему, хорошо известному примеру, принадлежащему Борелю.

Пример 3. Пусть E состоит в выборе наугад одной из цифр 0, . . ., 9. Представим себе, что этот эксперимент повторяется бесконечное число раз. Возможный результат такой бесконечной последовательности экспериментов можно записать в виде бесконечной десятичной дроби. Исключая (совершенно невероятные) десятичные дроби, оканчивающиеся бесконечной последовательностью цифр 9, мы можем поэтому установить взаимно однозначное соответствие между последовательностями экспериментов и точками интервала $0\leq x< 1$. Следовательно, для этих последовательностей экспериментов мы можем определить счетно-аддитивную *меру* $m\ [X]$.

Мы отождествим теперь эту меру с вероятностью. Под «элементарным утверждением» мы понимаем утверждение $S_{i,\ j}$ вида: i-е повторение эксперимента E дает значение j. Оно соответствует множеству $X_{i,\ j}$ экспериментов с i-й цифрой, равной j; очевидно, что $p\ [S_{i,\ j}]\ = 1/10$. Следовательно, конечные или счетные булевы комбинации «элементарных утверждений» соответствуют борелевским подмножествам интервала и единственная «вероятность», определяемая для каждого такого утверждения, равна мере соответствующего подмножества интервала.

В силу теоремы 15 р. 6.11 предыдущий пример является универсальным для алгебр вероятности, порождаемых конечным или счетным множеством утверждений.

Алгебра вероятности включает в себя еще одну операцию — взятия *весового значения,* как мы можем это показать следующим простым примером.

Пусть большой квадрат разбит на площадки $a_1, . .., a_n$; пусть, далее, P и Q обозначают эксперименты, заключающиеся в бросании монеты в квадрат с противоположных сторон. Пусть $p[a_i]$ и $q[a_i]$ — вероятности того, что монета попадет в a_i соответственно, при эксперименте P и Q. Тогда любое весовое значение $\lambda p+\mu q$ вероятностей p и $q[\lambda>0, \mu=1 —\lambda>0]$ может быть реализовано путем следующего эксперимента. Диск, разбитый на сектора в $2\pi\lambda$ и $2\pi\mu$ радиан, вращается в фиксированной горизонтальной плоскости, и монета бросается тогда со стороны P или со стороны Q в зависимости от того, какой из секторов окажется после остановки напротив фиксированной точки.

Математические свойства этой операции рассматриваются с технической точки зрения в р. 6.16. 13; она играет фундаментальную роль при рассмотрении абстрактных марковских процессов и в эргодической теории (см. р. 6.17). Мы заметим здесь только, что при этой операции алгебра вероятности становится нормированной векторной структурой.

6.14. Структурно упорядоченные полугруппы

6.14.1. Определение; интерпретация в теории идеалов

Понятие структурно упорядоченной полугруппы, или l-полугруппы, возникло естественным образом в теории идеалов.

Пример 1. Пусть F—нормальное поле расширения, образованное присоединением к рациональным числам корней алгебраического уравнения; пусть E—подкольцо всех *целых* алгебраических чисел из F. Дедекиндов *идеал* в F есть непустое подмножество H из F такое, что если $a, b \in H$ и $r \in E$, то $a \pm b \in H$ и $ra \in H$.

Так как это есть свойство замыкания, идеалы из F образуют полную структуру относительно теоретико-множественного включения. Если H и K два идеала, то $H \cap K$ есть их пересечение (которое никогда не пусто, поскольку нуль принадлежит каждому идеалу), а множество всех сумм $x + y$ [$x \in H, y \in K$] является объединением $H \cup K$ идеалов H и K. Произведение HK идеалов H и K определяется как наименьший идеал, содержащий все произведения $xy[x \in H, y \in K]$.

Относительно этих трех операций дедекиндовы идеалы в F удовлетворяют всем свойствам, которые будут абстрактно определены.

Определение. *Под мультипликативной структурой, или m-структурой, мы понимаем структуру L с бинарным умножением, удовлетворяющим условию*

$$a(b \cup c) = ab \cup ac \quad \text{и} \quad (a \cup b)c = ac \cup bc. \tag{1}$$

Нуль m-структуры L есть элемент 0, *удовлетворяющий условию*

$$0 \cap x = 0x = x0 = 0 \ \textit{для всех} \ x \in L. \tag{2}$$

Единица структуры L есть элемент e, удовлетворяющий условию

$$ex = xe = x \ \textit{для всех} \ x \in L. \tag{3}$$

Бесконечность в L есть элемент I, удовлетворяющий условию

$$I \cup x = Ix = xI = I \ \textit{для всех} \ x \in L. \tag{4}$$

Структура L называется коммутативной или ассоциативной, если, соответственно, выполняются условия

$$xy = yx, \tag{5}$$

$$x\,(yz) = (xy)\,z \tag{6}$$

для всех $x,y,z \in L$. Если структура L условно полная и удовлетворяет неограниченным дистрибутивным законам

$$a \vee b_\alpha = \vee\,(ab_\alpha) \text{ и } (\vee a_\alpha)\, b = \vee\,(a_\alpha b),$$

то она называется полной т-структурой, или ст-структурой. Ассоциативная т-структура с единицей называется структурно-упорядоченной полугруппой, или l-полугруппой; а если она полна, то ее называют cl-полугруппой.

Легко проверить, что в примере 1 дедекиндовы идеалы образуют коммутативную *cl*-полугруппу с нулем. Так, (1') следует из того, что $\vee K_\alpha$ является теоретико-множественным объединением объединений (в смысле теории идеалов) конечного числа идеалов K_α; идеал, состоящий из одного только числа 0, удовлетворяет условию (2).

Теорема 1. *В любой т-структуре мы имеем:*

$$\text{из } a \le b \text{ следует } xa \le xb \text{ и } ay \le by \text{ для всех } x, y, \tag{7}$$

$$(a \cap b)\,(a \cup b) \leqslant ba \cup ab \text{ для всех } a, b. \tag{8}$$

Если т-структура имеет единицу е, то

$$\text{из } a \cup b = e \text{ следует } a \cap b = ba \cup ab, \tag{9}$$

$$\text{из } a \cup b = a \cup c = e \text{ следует } a \cup bc = a \cup (b \cap c) = e. \tag{10}$$

Если она имеет элемент $z \le e$, удовлетворяющий условию $zx = xz = z$ для всех x, то этот элемент есть нуль.

Доказательство. Утверждение (7): из $a \le b$ следует

$$by = (a \cup b)\,y = ay \cup by.$$

Утверждение (8): в силу (7) $(a \cap b)\,(a \cup b) = (a \cap b)\,a \cup (a \cap b)\,b \leqslant$ $\leqslant ba \cup ab.$

Утверждение (9): если $a \cup b = e$, то в силу (8) $a \cap b \leqslant ba \cup ab.$

Но в силу (7) $ba \leqslant ea = a,\ ba \leqslant b,\ ab \leqslant a,\ ab \leqslant b;$ следовательно, $ba \cup ab \leqslant a \cap b,$ чем и доказано (9).

Утверждение (10): очевидно, что $e \ge a,\ b,\ c,$ откуда $a \cup bc \geqslant a \geqslant aa,$ $ba,\ ac.$ Следовательно,

$$e = e \cup ee \geqslant a \cup bc \geqslant aa \cup ba \cup ac \cup bc = (a \cup b)\,(a \cup c) = ee = e.$$

Далее, так как $b \ge bc$ и $c \ge bc$, мы имеем $b \cap c \ge bc$, так что $a \cup (b \cap c) \geqslant a \cup bc = e,$ чем и доказано (10). Последнее утверждение доказывается непосредственно, так как для всех $x \in L$ $z = zx \le ex = x$, откуда $z = z \cap x = x \cap z$.

Определение. *Пусть G—любая m-структура. Мы определим правое частное h:k элемента h по k как наибольший элемент x (если он существует), удовлетворяющий условию xk≤h; левое частное h::k элемента h по k есть наибольший элемент y, удовлетворяющий условию ky≤h. m-Структура, в которой такие частные всегда существуют, называется структурой* с *делением.*

m-Структура примера 1 является структурой с делением; *Н:К* есть множество всех $x \in F$ таких, что $xk \in H$ для всех $k \in K$. Это есть идеал, поскольку $(x \pm y)k = xk \pm yk$ и $(rx)k = r(xk)$ $[r \in E]$.

Это несколько видоизменяет терминологию Уорда—Дилуорса, где структура с делением была определена как коммутативная *l*-полугруппа, в которой *e≥x* для всех *x.* Последнее предположение соответствует обычному в теории идеалов различию между *целыми* идеалами *К≤Е* из *F* и *дробными* идеалами, для которых такое включение не имеет места. Целые дедекиндовы идеалы в *F* суть обычные идеалы в *Е* в том смысле, как это теперь принято.

Определение. *Элементы x≤e m-структуры будут называться целыми. Если каждый элемент содержится в е, то m-структра будет называться целой.*

Мы покажем теперь, что целые идеалы в примере 1 образуют целую структуру с делением. Более общо, мы имеем

Теорема 2. *В любой cm-структуре если некоторое х удовлетворяет условию xk≤h, то h:k существует.*

Доказательство. Пусть *и* объединение всех x_α таких, что $x_\alpha k \le h$. Тогда

$$uk = (\vee x_\alpha) k = \vee (x_\alpha k) \leq h.$$

Следствие 1. *Целые элементы любой cm-структуры образуют «целую» структуру с делением.*

Следствие 2. *Любая cm-структура с нулем есть структура с делением.*

Доказательство. $0k = 0 \le h$ для всех *h, k.*

Мы приводим без доказательства следующий результат.

Теорема 3. *Во всякой структуре с делением мы имеем*

а) $(\vee a_\alpha) : b = \wedge (a_\alpha : b)$ *и симметрично,*

б) $a : (\vee b_\beta) = \wedge (a : b_\beta)$ *и симметрично,*

в) *условия* $a \geq cb$, $a:b \geq c$ и $a::c \geq b$ *эквивалентны,*

г) $(ab) : b \geq a$ *и* $(ba) :: b \geq a$.

Если умножение ассоциативно, то также

д) $(a :: b) : c = (a : c) :: b$ *есть наибольшее x, для которого* $bxc \leq a$,

е) $a : (bc) = (a : c) : b$ *и симметрично.*

В «а» — «б» из существования левой части следует существование правой части.

Заметим, что, хотя *l*-полугруппы удовлетворяют принципу левой и правой симметрии, принцип двойственности к ним не применим.

6.14. 2. Родственные интерпретации

Можно обобщить пример 1 в следующем направлении.

Пример 2. Пусть *R*—любая система с бинарным умножением *ху* и, возможно, другими бинарными операциями *х+у, х — у,...,* по отношению к которым умножение дистрибутивно, так что $a(b\pm c)=ab\pm ac$, $(a\pm b)c=ac\pm bc$. Ввиду отсутствия более подходящего термина мы будем называть такую систему *кольцоидом;* обычные кольца и дистрибутивные структуры включаются сюда как специальные случаи. Как и в случае колец, мы будем называть *модулем* в *R* такое подмножество *Н,* что $a,b\in H$ влечет $a\pm b\in H$; как и в случаях колец и дистрибутивных структур, мы будем называть *идеалом* такой модуль *Н,* что $a\in H$ влечет $xa\in H$ и $ay\in H$ для всех $x, y\in R$.

Теорема 4. *Модули любого кольцоида R образуют cm-структуру с делением, в которой идеалы образуют cm-подструктуру с делением.*

Доказательство. Поскольку свойство быть модулем есть свойство замыкания, модули в *R* образуют полную структуру, в которой пересечения являются теоретико-множественными произведениями. Далее, объединение $K\cup L$ двух модулей состоит из всех комбинаций с помощью ± элементов $y_i\in K$ и $z_i\in L$ таких, как, например, $[(y_1-z_1)+(y_2-z_2)]-y_3$. Мы определим *произведение НК* двух модулей как модуль, порожденный произведениями *ху* $[x\in H,\ y\in K]$. Очевидно, что $HK\leq H(K\cup L)$ и $HL\leq H(K\cup L)$; следовательно, $HK\cup HL\leqslant H(K\cup L)$. Но, обратно, $H(K\cup L)$ порождается комбинациями вида

$$x\{[(y_1-z_1)+(y_2-z_2)]-y_3\}=[(xy_1-xz_1)+(xy_2-xz_2)]-xy_3,$$

принадлежащими $HK\cup HL$; следовательно, $H(K\cup L)\leqslant HK\cup HL$. В силу симметрии мы получаем (1). Отсюда следует (1'), ибо $\vee K_\alpha$ является теоретико-множественным объединением объединений конечного числа модулей K_α.; следовательно, модули образуют *cm*-структуру. В этой *cm* -структуре пустое множество является нулем; следовательно, в силу следствия 2 из теоремы 2 эта *cm*-структура является структурой с делением.

В этой *cm*-структуре любое объединение или пересечение модулей, являющихся идеалами, само является идеалом. Это утверждение тривиально для пересечений (т. е. теоретико-множественных произведений); для объединений оно вытекает из тождества, приведенного выше. В силу этого факта доказательства тождеств (1) — (1'), приведенные выше для произведений модулей, применимы в равной мере и к произведениям идеалов, если произведение двух идеалов H и K есть идеал, порожденный произведениями $xy[x \cup H,\ y \cup K]$. (Если умножение ассоциативно, то это есть также произведение модулей, определенное выше.)

Если R не имеет никаких других операций кроме умножения, то модули образуют по отношению к кронекерову умножению комплексов булеву алгебру всех подмножеств («комплексов») системы с бинарным умножением. В случае групп или луп непустые подмножества образуют *ст*-структуру, которая обычно не является структурой с делением.

Если R—кольцо, то непустые модули и идеалы образуют *ст-структуры с делением, обладающие нулем,* который состоит из нулевого элемента кольца. То же самое замечание применимо к подалгебрам и инвариантным подалгебрам любой линейной алгебры. В этих случаях, если умножение ассоциативно, правые частные являются левыми идеалами и левые частные являются правыми идеалами.

Если R есть дистрибутивная структура, то мы выводим, что идеалы образуют *ст*-структуру с делением. Кроме того, в этом случае, как и в случае колец с единицей, R является единицей для *ст*-структуры, в силу чего эта структура есть *целая ст*-структура с делением, обладающая нулем (Уорд).

Но мы знаем, что идеалы любой дистрибутивной структуры образуют дистрибутивную структуру с относительными псевдодополнениями. Мы можем обобщить этот результат следующим образом.

Теорема 5. *Если в структуре ху определено как $x \cap y$, то она является т-структурой тогда и только тогда, если она дистрибутивна. Она является структурой с делением тогда и только тогда, если она есть структура с относительными псевдодополнениями.*

Следствие. *Брауэрова логика может быть определена как структура, двойственная структуре с 0 и I, являющейся структурой с делением по отношению к структурному пересечению.*

Этот результат дает связь между теорией идеалов, топологией (алгебра открытых множеств) и математической логикой.

Хорошо известно, что умножение в алгебрах Ли является аналогом операции образования коммутатора $[a,\ b] = a^{-1}b^{-1}ab$ в группах. Это может основываться на следующем аналоге теоремы 4.

Пример 3. Пусть L структура всех непустых нормальных делителей какой-нибудь группы G; определим $[M,\ N]$ как подгруппу, порожденную всеми коммутаторами $x^{-1}y^{-1}xy\ [x \in M,\ y \in N]$. Тогда

L является коммутативной *cm*-структурой с делением, имеющей в качестве нуля 1 (единица группы).

Доказательство. Так как подгруппа *[M, N]* инвариантна относительно всех внутренних автоморфизмов, она является нормальным делителем. Так как $y^{-1}x^{-1}yx = y^{-1}(x^{-1}(y^{-1})^{-1}xy^{-1})y$, имеем *[M, N] = [N, M]*. Так как

$$x^{-1}(yz)^{-1}x(yz) = (x^{-1}z^{-1}xz)z^{-1}(x^{-1}y^{-1}xy)z,$$

мы имеем $[L, M \bigcup N] = [L, M] \bigcup [L, N]$. В силу симметрии мы имеем (1), откуда, как и в теореме 4, следует (1').

Понятия нильпотентных колец, нильпотентных групп и т. д. переносятся на любую структуру с делением *L*, имеющую 0 и *I*; для простоты мы ограничимся коммутативным случаем. Мы определим $I^1 = I$ и $I^{n+1} = II^n$; ${}^0I = 0$ и ${}^{n+1}I = ({}^nI) : I$. Будем говорить, что структура *L нильпотентна,* если некоторое $I^n = 0$; в этом случае *L* содержит «нильпотентную» цепь вида

$$I = a_0 > a_1 > a_2 > \ldots > a_r = 0, \text{ где } a_i I \geqslant a_{i+1}. \qquad (*)$$

Но для любой такой цепи мы можем доказать с помощью (7) и индукции, что $I^{i+1} \leqslant a_i \leqslant {}^{r-i}I$. Отсюда следует, что $I^r+^1 = 0$ и ${}^rI = I$. Следовательно, если мы назовем последовательность

$$I > I^2 > I^3 > \ldots > I^n = 0$$

нижним центральным рядом в *L*, а дуальную последовательность

$$0 = {}^0I < {}^1I < {}^2I < \ldots < {}^mI = I$$

верхним центральным рядом в *L*, то мы имеем $n \leq r+1$ и $m \leq r$. Следовательно, оба числа n—1 и m равны длине r кратчайшей нильпотентной цепи. Мы заключаем

Теорема 6. *Во всякой нильпотентной структуре с делением верхний и нижний центральные ряды имеют одну и ту же длину* r. *Если* $I = a_0 > a_1 > \ldots > a_r = 0$ *любая нилъпотентная цепь, то*

$$I^{i+1} \leqslant a_i \leqslant {}^{r-i}I \text{ для всех } i. \qquad (11)$$

6.14.3. Целые *m*-структуры

Мы видели, что идеалы любого кольца *E* с единицей 1 образуют целую *m*-структуру с единицей *e* = *E* и нулем 0. В любой такой структуре можно доказать некоторые свойства относительно простых или «взаимно простых» идеалов, хорошо известные в теории коммутативных колец.

Определение. *Максимальный элемент целой m-структуры L есть элемент, покрываемый элементом e*; *простой элемент есть такой* *элемент р, что $xy \leq p$ влечет $x \leq p$ или $y \leq p$. Два элемента* a, $b \in L$ *являются взаимно простыми, если* $a \cup b = e$.

Лемма 1. *Если а и b взаимно просты, то*

$$x = xa \cup xb = (x \cap a) \cup (x \cap b) \text{ для всех } x. \tag{12}$$

Доказательство. Заметим прежде всего, что, поскольку $xy \leq xe = x$ и $xy \leq ey = y$, мы имеем

$$xy \leq x \cap y \qquad \text{для всех } x, y. \tag{13}$$

Мы выводим затем (12) из последовательности неравенств

$$x = xe = x(a \cup b) = xa \cup xb \leqslant (x \cap a) \cup (x \cap b) \leqslant x.$$

Лемма 2. *Если а и b взаимно просты и* $a \cap b \leq x$, *то*

$$x = (x \cup a) \cap (x \cup b) = (x \cap a) \cup (x \cap b). \tag{14}$$

Доказательство. В любой структуре $(x \cap a) \cup (x \cap b) \leqslant \leqslant x \leqslant (x \cup a) \cap (x \cup b)$. При наших предположениях имеем в силу (12) и (1)

$$(x \cup a) \cap (x \cup b) = a[(x \cup a) \cap (x \cup b)] \cup b[(x \cup a) \cap (x \cup b)]$$
$$\leqslant a(x \cup b) \cup b(x \cup a) = (ax \cup ab) \cup (bx \cup ba).$$

Далее, поскольку $ax \leqslant a \cap x$ и $ab < a \cap b = a \cap (a \cap b) \leqslant a \cap x$, имеем $ax \cup ab \leq a \cap x$. Аналогично, $bx \cup ba \leq b \cap x$. Подставляя, получим $(x \cup a) \cap (x \cup b) \leqslant (a \cap x) \cup (b \cap x)$, чем и завершается доказательство.

Теорема 7. *Если a и b взаимно просты, то интервал* $[a \cap b, e]$ *структурно изоморфен кардинальному произведению* $[a, e] \times [b, e]$.

Доказательство. Для заданного $a \cap b \leq x \leq e$ положим $t = x \cup a$ и $u = x \cup b$; очевидно, что $t \in [a, e]$ и $u \in [b, e]$. Обратно, для заданных $t \in [a, e]$ и $u \in [b, e]$ образуем $\varphi(t, u) = t \cap u$. Это суть однозначные отображения $[a \cap b, e]$ в $[a, e] \times [b, e]$ и обратно. В силу леммы 2 $\varphi(x \cup a, x \cup b) = x$; обратно, так как $a \leq t$,

$$t \leqslant (t \cap u) \cup a \leqslant (t \cup a) \cap (u \cup a) = t \cup a \ (\text{так как } u \cup a = e) = t;$$

аналогично, $u = (t \cap u) \cup b$; следовательно, эти отображения взаимно обратные. Но они являются изотонными; следовательно, оба они суть изоморфизмы.

Теорема 8. *Каждая целая m-структура с дополнениями L есть булева алгебра с* $xy = x \cap y$.

Доказательство. Пусть $a \in L$—произвольный элемент и a' любое дополнение элемента a. Тогда в силу теоремы 7 a и a' лежат в центре структуры L; следовательно, L есть булева алгебра. Далее, в силу (12) и в силу дистрибутивности

$$x \cap a = (xa \cup xa') \cap (xa \cup x'a) =$$
$$= xa \cup (x \cap x')\, a \cup x\,(a' \cap a) \cup (xa' \cap x'a).$$

В последнем выражении все члены, кроме первого, равны нулю, а потому $x \cap a = xa$, что и требовалось доказать.

Предположим теперь, что выполняется условие обрыва возрастающих цепей. Каждому максимальному элементу m_i из L мы сопоставим множество M_i всех *единообразных* элементов, удовлетворяющих условию $x \cup m_j = e$ для всех максимальных m_j, кроме, быть может, m_i. Так как в силу условия обрыва цепей каждое $x < e$ содержится по крайней мере в одном максимальном элементе, каждое M_i состоит из e и всех x, содержащихся в m, и ни *в каком другом* максимальном элементе.

В случае теории алгебраических чисел максимальные идеалы являются простыми идеалами и единообразные идеалы являются степенями простых идеалов. В самом деле, в любой целой m-структуре каждый максимальный элемент m является простым, так какиз $xy \leq m$ и $x \nleq m$ следует $x \cup m = e$, откуда $y = ey =$
$= (x \cup m)\, y = xy \cup my \leqslant m.$

Ввиду этого следующий результат может быть рассматриваема как частичное обобщение основной теоремы теории идеалов[3]).

Теорема 9. *Пусть L—произвольная целая m-структура, удовлетворяющая условию обрыва возрастающих цепей. Подструктура S, порожденная множествами M_i единообразных элементов, является кардинальным произведением $M_1 \times M_2 \times M_3 \times \ldots$ подструктур M_i.*

Доказательство. В силу (10) каждое M_i является интервальной подструктурой (дуальным идеалом), замкнутой относительно умножения. Далее, если $x \in M_i$ и $y \in M_j$ $[i \neq j]$, то $x \cup y$ не содержится ни в каком максимальном элементе, а потому $x \cup y = e$. В силу (10) то же самое замечание справедливо, если $\bar{x} = x_1 \cap \ldots \cap x_n\, [x_i \in M_i]$ и $y \in M_{n+1}$. Следовательно, в силу теоремы 7 и индукции интервальная подструктура, порожденная подструктурами M_i, является их кардинальным произведением. При наличии условия обрыва возрастающих цепей мы можем распространить этот результат на бесконечное число множителей при условии, что все «компоненты», кроме конечного числа, равны e.

Замечание. Хотя M_i замкнуты относительно умножения, их кардинальное произведение может и не быть таковым.

6.14.4. Изотопные, полунепрерывные и субгармонические функции

Рассмотрим класс N всех вещественных функций $f(x)$ действительного переменного x, таких, что $x \geq y$ влечет $f(x) \geq f(y)$. Такие функции могут быть названы *изотопными* или *неубывающими.* Если мы определим $f \geq g$ в том смысле, что $f(x) \geq g(x)$ для всех x, то мы получим структуру R^R (см. р. 6.2.7), где R есть система вещественных чисел. Мы можем превратить $N = R^R$ в m-структуру одним из следующих двух способов.

Мы можем принять в качестве fg *сложную* функцию $g(f(x))$, что даст нам ассоциативное, некоммутативное умножение. Тогда обе функции $fg \cup h$ и $f(g \cup h)$ относят каждому x наибольшее из чисел $g(f(x))$ и $h(f(x))$, а $fh \cup gh$ и $(f \cup g)h$ относят обе наибольшее из чисел $h(f(x))$ и $h(g(x))$; следовательно, мы определяем тем самым m-структуру.

Эта конструкция дает для любой цепи C m-структуру C^C. Заметим, что для цепей изотонная функция есть то же самое, что эндоморфизм по объединениям. Это есть тот класс операторов, который дает m-структуру в общем случае.

Пример 4. Эндоморфизмы по объединениям любой структуры L образуют l-полугруппу M.

В самом деле, мы имеем $x(\theta \cup \theta_1) = x\theta \cup x\theta_1$, что является гомоморфизмом по объединениям, так как

$$(x \cup y)(\theta \cup \theta_1) = (x \cup y)\theta \cup (x \cup y)\theta_1 = x\theta \cup y\theta \cup x\theta_1 \cup y\theta_1 = \\ = x(\theta \cup \theta_1) \cup y(\theta \cup \theta_1).$$

Умножение есть обычное умножение эндоморфизмов. Поэтому

$$x(\theta \cup \theta_1)\theta_2 = (x\theta \cup x\theta_1)\theta_2 = x\theta\theta_2 \cup x\theta_1\theta_2 = x(\theta\theta_2 \cup \theta_1\theta_2),$$

$$x\theta(\theta_1 \cup \theta_2) = x\theta\theta_1 \cup x\theta\theta_2 = x(\theta\theta_1 \cup \theta\theta_2).$$

Этим доказательство завершается.

Мы можем также в качестве нашей вспомогательной операции принять операцию образования *суммы* $h = f+g$ двух функций, определяемую по формуле $h(x) = f(x) + g(x)$ для всех x. Класс всех изотонных вещественных функций является замкнутым относительно сложения и объединения. Тот факт, что он образует l-полугруппу, является поэтому следствием приводимого ниже принципа.

Пример 5. Пусть G—произвольная полная l-группа; групповую операцию мы обозначаем через +, а нулевой элемент— через 0. Пусть S — любое подмножество в G, содержащее 0, содержащее вместе с любыми двумя элементами x и y их сумму $x+y$ и содержащее $\vee x_\alpha$

всякий раз, когда все x_α принадлежат S и имеют верхнюю грань. Тогда S есть полная l-полугруппа.

В самом деле, (1) —(1') имеют место в любой l-группе.

По той же самой причине класс всех *полунепрерывных сверху* вещественных функций образует полную коммутативную l-полугруппу, если в качестве групповой операции взято сложение. То же можно сказать в отношении класса всех *субгармонических* функций одного, двух или большего числа переменных. Это же справедливо для класса всех нигде не положительных субгармонических функций и т. д., и т. д.

В случае субгармонических функций частное $0{:}a$ имеет интересное истолкование. Это есть объединение всех субгармонических функций, имеющих те же граничные значения, что и супергармоническая функция — a; следовательно, эта функция *гармоническая.* Вообще, гармонические функции суть субгармонические функции h, аддитивные обратные которым являются гармоническими, и все $h{:}a$ также являются гармоническими (имеют обратные).

6.14.5. Алгебра отношений

Пусть Г—произвольный класс элементов α, β, γ, ... Под бинарным (или диадическим) *отношением* на Г понимают любое правило r, которое указывает для каждой упорядоченной пары (α, β) элементов из Г, что либо отнощение r имеет место между α и β (символически, $\alpha r\beta$), либо оно не имеет места (символически, $\alpha r\beta$).

Если Г состоит из n элементов $\alpha_1, ..., \alpha_n$, то можно установить взаимно однозначное представление отношений r на Г «матрицами отношений» $\|r_{ij}\|$ состоящими из нулей и единиц. Оно определяется посредством

$$r_{ij}=\begin{cases}1,\ если\ \alpha_i r\alpha_j,\\ 0,\ если\ \alpha_i r' a_j\ \left(т.\ е.\ в\ противном\ случае\right).\end{cases}$$

Так, отношение *равенства* e соответствует единичной матрице e_{ii}=1 и e_{ij} = 0, если $i\neq j$. *Нуль-отнощение* 0 соответствует нулевой матрице 0_{ij} = 0 для всех i, j; *универсальное* отношение I матрице I_{ij} = 1 для всех i, j.

Мы определим $r\leq s$ в соответствии с логическим понятием «r влечет s» следующим образом:

$r\leq s$ означает, что из $\alpha r\beta$ следует $\alpha s\beta$. (16)

Таким образом, в конечном случае это означает, что $r_{ij}\leq s_{ij}$ для всех i, j. Используя представление матрицами отношений (которое легко распространить на бесконечный случай), мы видим:

Лемма 1. Бинарные отношения на фиксированном множестве Г, мощности к образуют атомную булеву алгебру 2^{N^2} относительно включения. Мы можем поэтому *определить* $r \cap s$, $\cup s$ и r' *обычным способом. Таким образом,* r' *имеет тот же смысл, что и в первом абзаце, а* $\alpha(r \cup s)\beta$ *означает, что либо* $\alpha r\beta$, *либо* $\alpha s\beta$. *Далее, отношение r рефлексивно тогда и только тогда, если* $r \geq e$.

Помимо этих булевых операций, мы имеем операцию (относительного) *умножения* отношений, определенную посредством

$$\alpha(rs)\,\beta \text{ означает, что } \alpha r\gamma \text{ и } \gamma s\beta \text{ для некоторого } \gamma \in \Gamma. \qquad (*)$$

Таким образом, отношение *r транзитивно* тогда и только тогда, если $r^2 \leq r$. Оно является взаимно однозначным преобразованием тогда и только тогда, если $rx=xr=e$ для соответствующего «обратного» элемента x; кроме того, в случае преобразований (взаимно однозначных или невзаимно однозначных) (*) сводится к обычному умножению преобразований.

Мы перечислим теперь некоторые основные законы алгебры отношений, дополняющие собой булевы тождества, установленные леммой 1.

Теорема 10. *Отношения на* Г *образуют относительно включения и умножения булеву cl-полугруппу R* (Г) *с нулем* 0, *единицей e и наибольшим элементом I. Кроме того,*

$$rs \leq e' \quad \text{влечет} \quad sr \leq e', \qquad (17)$$

$$e' : (e' : r) = r \quad \text{для всех} \quad r, \qquad (18)$$

$$e' : (rs)' = (e' : s')\,(e' : r'), \qquad (19)$$

$$\text{если } r > 0, \text{ то } IrI = I. \qquad (20)$$

Доказательство. Все утверждения, кроме (17)—(20), являются очевидными. Отсюда следует, что *все тождества в теореме* 1 *и в теореме* 3 *применимы к алгебре отношений.* Мы докажем теперь (17)—(20).

Если $rs \leq e'$, то $r_{ik}s_{ki} = 0$ для всех i, k; это означает, что $r_{ik}= 1$ влечет $s_{ki} = 0$. Отсюда следует, что $s_{ki}r_{ik} = 0 = (sr)_{kk}$ для всех k, i, откуда $sr \leq e'$; этим доказано (17).

Далее, если мы положим $s_{ki} = 0$, когда $r_{ik}=1$ и $s_{ki}=1$ *в* противном случае, то $sr \leq e'$; следовательно, наибольшее $s= e':r$, удовлетворяющее условию $sr \leq e'$, получается из r путем перестановки индексов и подстановки нулей вместо единиц и единиц вместо нулей. Если мы проделаем это дважды, то, очевидно, снова получим r, чем и доказано (18). Мы отложим на время доказательство свойства (19) и докажем (20), замечая, что если $r > 0$, то некоторое $r_{ij}=1$, откуда каждое

$$(I\,r\,I)_{hk} \geqslant I_{hi}\,r_{ij}\,I_{jk} = 1.$$

Закон (18) тесно связан со свойствами отношения $\breve{r}$ *обратного* отношению *r*, которое определяется следующим образом:

$$\alpha \breve{r} \beta \text{ тогда и только тогда, если } \beta r \alpha. \qquad (21')$$

Так как $e' : r$ получается из *r* перестановкой индексов и подстановкой нулей вместо единиц и единиц вместо нулей, мы имеем

$$\breve{r} = e' : r'. \qquad (21)$$

Следовательно, *нам не нужно рассматривать взятие обратного отношения как неопределенную операцию,* хотя ее обычно и рассматривают так. Мы получаем теперь (19) как другой способ записи соотношения

$$\breve{rs} = \breve{r}\,\breve{s}. \qquad (19')$$

Теорема 10 этим доказана; мы можем определить *алгебру отношений* как любую систему, удовлетворяющую условиям теоремы 10.

В любой структуре с делением для любого фиксированного элемента *a* соответствие $r \to a : r$ изменяет включение на противоположное. В силу (18) соответствие $r \to e' : r$ обратно самому себе, а потому оно взаимно однозначно; следовательно, это соответствие есть дуальный автоморфизм. Следовательно,

$$\breve{r} = e' : r' = (e' : r)'. \qquad (22)$$

Далее, $r \to \breve{r}$ есть произведение двух дуальных автоморфизмов, а потому является структурным автоморфизмом, дающим

$$\breve{(r \cap s)} = \breve{r} \cap \breve{s}, \quad \breve{(r \cup s)} = \breve{r} \cup \breve{s}, \quad \breve{r}' = (\breve{r'}). \qquad (23)$$

Остановимся теперь на включениях, даваемых в (17). Из этих включений следует, что $e' : r = e' :: r$, т. е., что левое и правое частные *e*' по *r* совпадают. Из них следует также, что

$$\text{если } (r_1 r_2 \ldots r_{n-1}) r_n \leq e', \text{ то } r_n (r_1 r_2 \ldots r_{n-1}) \leq e'; \qquad (24)$$

иными словами. *n-арное отношение* $r_1 r_2 \ldots r_n \leq e'$ *инвариантно относительно циклических перестановок.*

6.14. 6. Строение и теория представлений

Мы знаем (см. р.6.11, теорема 7), что любая булева алгебра изоморфна подалгебре булевой алгебры всех подмножеств подходящим образом выбранного множества точек. Кроме того, любой гомоморфный образ, подалгебра и прямое произведение булевых алгебр сами являются булевыми алгебрами. Мы покажем теперь, что нельзя надеяться на получение аналогичных результатов в отношении алгебр отношений.

Лемма 2. *Каждая булева т-структура R с* 0, *e и I, которая удовлетворяет условию* (20), *является простой, т. е. она не имеет* *собственных гомоморфных образов, кроме себя самой и одноэлементной алгебры* $0 = e = I$.

Доказательство. Пусть θ—любое отношение конгруентности на R. Если θ не является отношением равенства, то в силу теоремы 8 р. 6.11 некоторое $r > 0$ удовлетворяет условию $r \equiv 0(\mathrm{mod}\,\theta)$. Следовательно, в силу (20) $I = IrI \equiv I0I = 0(\mathrm{mod}\ \theta)$, а потому (снова в силу теоремы 8 р. 6.11) каждое $x \equiv 0(\mathrm{mod}\ \theta)$.

Но (20) имеет место в каждой подалгебре алгебры отношений, содержащей 0, e и I, т. е. является замкнутым относительно этих нулъарных операций. С другой стороны, (20) не имеет места для прямого произведения двух алгебр отношений; следовательно, *прямое произведение двух алгебр отношений не изоморфно подалгебре алгебры отношений,* несмотря на то, что соответствующий факт справедлив для булевой алгебры и для матричной алгебры.

Отсюда следует, что понятие «алгебры отношений» является не вполне ясным. Все законы теоремы 10 сохраняются при переходе к подалгебре и (тривиальным образом) к гомоморфному образу; но если мы желаем рассматривать класс алгебр, образованных из $R(\Gamma)$ с помощью этих операций, *а также* операции взятия прямого произведения, то мы должны использовать иные постулаты.

Неизвестно, будет ли каждая абстрактная алгебра, удовлетворяющая условиям теоремы 10, изоморфна подалгебре некоторого $R(\Gamma)$. (Хотя известно, что каждая «простая» матричная алгебра с конечным базисом является полной матричной алгеброй над телом.) Многие определяют «подалгебру» как подмножество, содержащее 0, e, I и замкнутое относительно булевых операций, умножения и взятия обратного отношения. Результаты п.5 наводят на мысль потребовать замкнутость относительно операции взятия частных, для которой операция взятия обратного является частным случаем.

Мы проиллюстрируем возможную сложность подалгебр алгебр отношений на 32-элементной подалгебре 512-элементной алгебры всех отношений для трех элементов, порожденной элементами 0, e, I и первой матрицей на рис. 9. В качестве минимальных

$$\begin{matrix}0&1&1\\0&0&0\\0&0&0\end{matrix}\qquad\begin{matrix}1&0&0\\0&0&0\\0&0&0\end{matrix}\qquad\begin{matrix}0&0&0\\0&1&0\\0&0&1\end{matrix}\qquad\begin{matrix}0&0&0\\1&0&0\\1&0&0\end{matrix}\qquad\begin{matrix}0&0&0\\0&0&1\\0&1&0\end{matrix}$$

Рис. 9.

элементов она имеет остальные матрицы отношений рис. 9. Более симметричным примером является следующий.

Пример 6. Подмножества $S, T, U, \ldots$ произвольной группы G образуют алгебру отношений, если в качестве S мы принимаем множество всех элементов, обратных элементам из S, а в качестве ST — множество всех произведений $st[s \in S,\ t \in T]$. Эту систему мы представим как подалгебру алгебры $R(G)$ всех отношений на G следующим образом. Каждое S мы отождествляем с отношением σ: $x\sigma y$ означает, что $y \in xS$, иными словами, что $xs = y$ для некоторого $s \in S$. Можно проверить, что $S \cap T$, $S \cup T$, $\breve{S}$ и имеют верный смысл.

Хотя мы и не можем получить теорему о представлении для общих алгебр отношений, мы можем перестроить $R(\Gamma)$ как алгебру отношений, используя один прием из теории полупростых алгебр Веддербарна.

Определение. *Левоидеальный элемент алгебры отношений R есть такой элемент x, что* $Ix \le x$.

Поскольку $I(Ix)=(II)x \le I(x)$, каждый элемент $I(x)$ является левоидеальным; если x и y левоидеальные элементы, то, так как $I(x \cup y) = Ix \cup Iy \le x \cup y$, $x \cup y$ есть также левоидеальный элемент.

Лемма 3. *Левоидеалъные элементы произвольной булевой т-структуры R с I, удовлетворяющей законам* (17) —(19), *образуют булеву подалгебру, замкнутую относительно умножения.*

Доказательство. Если Ix левоидеальный элемент, то из $Ix \cap t = 0$ следует $x\breve{t} \cap I = 0$; следовательно, $0 = x\breve{t} = x(\breve{t}\breve{I}) \cap I = Ix \cap It.$ Поэтому дополнение элемента Ix, являющееся наибольшим элементом, имеющим 0 в пересечении с Ix, будет левоидеальным элементом. Поскольку $x \cup y$ есть левоидеальный элемент, если элементы x и y левоидеальны, отсюда следует, что левоидеальные элементы составляют булеву подалгебру. Наконец, любое правое кратное xy левоидеального элемента x удовлетворяет соотношению $I(xy) = (Ix)y \le xy$.

Пусть теперь $m_1, \ldots, m_s$—*минимальные* левоидеальные элементы в R. Для любого $r \in R$ определим $r_{ij} = 1$, если $m_i r \ge m_j$, и $r_{ij} = 0$ в противном случае. (В последнем случае $m_i r \cap m_j = 0$.) Можно доказать, что $r_{ji} = 0$ тогда и только тогда, если $\breve{r}_{ij} = 0$. Кроме того, $(r \cup s)_{ij} = r_{ij} \cup s_{ij}$. Для доказательства того, что представление гомоморфно, мы должны *предполагать,* что каждое $m_i m_j$ покрывает 0 (является атомом); это, однако, справедливо, если $R = R(\Gamma)$. Это не имеет места в примере 6.

6.14.7. Булевы матрицы

Алгебра отношений R(Г) может также рассматриваться как алгебра всех *эндоморфизмов по объединениям* булевой алгебры A всех подмножеств множества Г. Для получения этой интерпретации удобно представлять элементы алгебры A как характеристические функции $f = (f_1, \ldots, f_n)$. Тогда трансформация $g = fr$ функции f посредством эндоморфизма, соответствующего $r = \|r_{ij}\|$, определяется по формулам $g_j = \vee f_i r_{ij}$. Это наводит на мысль рассмотреть, как аналоги алгебр отношений, m-структуры эндоморфизмов по объединениям заданных дистрибутивных структур и, возможно, других структур, хотя не известно, когда эндоморфизмы по объединениям недистрибутивных структур образуют структуры.

Другой аналог алгебр отношений дается квадратными матрицами порядка n с коэффициентами r_{ij} в фиксированной булевой алгебре A. Эти матрицы образуют булеву алгебру, если мы определим

$$(r \bigcup s)_{ij} = r_{ij} \bigcup s_{ij}, \quad (r \bigcap s)_{ij} = r_{ij} \bigcap s_{ij}, \quad (r')_{ij} = r'_{ij}. \qquad (25)$$

Мы можем также определить (относительное) *умножение* по формуле

$$(rs)_{ij} = \bigvee_k (r_{ik} \bigcap s_{kj}). \qquad (26)$$

Мы можем определить 0, I, e согласно

$$0_{ij} = 0, \text{ и } I_{ij} = I \text{ для всех } i, j; \qquad (27)$$

$$e_{ii} = I, \text{ и } e_{ij} = 0, \text{ если } i \neq j. \qquad (27')$$

Это определяет всегда l-полугруппу, а в случае, когда A есть двухэлементная булева алгебра 2, это определяет рассмотренную уже алгебру отношений.

6.15. Структурно упорядоченные группы

6.15.1. Определение: положительные элементы

Ниже мы будем заниматься изучением структурно упорядоченных групп, или *l-групп,* в следующем смысле.

Определение. *l-группа G есть* а) *структура,* б) *группа, в которой* в) *отношение включения инвариантно относительно всех групповых сдвигов* $x \to a + x + b$. *Если G есть частично упорядоченное множество, удовлетворяющее условиям* «б» *и* «в», *то G называется частично упорядоченной группой.*

Мы будем ниже использовать для групповой операции аддитивную запись вместо принятой в р. 6.14 мультипликативной записи. В соответствии с этим мы будем через —*a* обозначать элемент, обратный *a,* через $a - b$ элемент $a + (-b)$, через $-a+b$ элемент $(-a) + b$ и через *na* элемент $a+ ...+a$ (*n* слагаемых).

Во всякой частично упорядоченной группе условие «в» эквивалентно условию

$$x \geqslant y \text{ влечет } a + x + b \geqslant a + y + b \text{ для всех } a, b \in G. \quad (1)$$

В *l*-группе это эквивалентно дистрибутивным законам

$$a + (x \cup y) = (a + x) \cup (a + y) \text{ и } (x \cup y) + b = (x + b) \cup (y + b) \quad (2)$$

или законам, им двойственным.

Следовательно, *любая l-группа является l-полугруппой* в смысле р. 6.14 и все результаты того раздела имеют здесь силу. Однако почти всегда легче получить непосредственно результат относительно *l*-групп.

Определение. *Элемент a l-группы G называется положительным, если* $a \geq 0$. *Множество всех положительных элементов группы G будет обозначаться через* G^+.

Теорема 1. *Любая частично упорядоченная группа G определяется с точностью до изоморфизма множеством* G^+ *своих положительных элементов, так как*

$a \leq b$, $b - a \in G^+$ *и* $-a + b \in G^+$ *суть эквивалентные условия.* (3)

Кроме того, а) $0 \in G^+$, б) *если* $x, y \in G^+$, *то* $x + y \in G^+$, в) *если* $x, y \in G^+$ *и* $x + y = 0$, *то* $x = y = 0$, г) *для всех* $a \in G$ *имеет место* $a + G^+ = G^+ + a$. *Обратно, если G любая группа и* G^+ *подмножество в G, удовлетворяющее условиям* «а» — «г», *то условие* (3) *определяет G как частично упорядоченную группу.*

Доказательство. Условие (3) мы легко получаем из (1); так, $a \leq b$ влечет $0 = a - a \leq b - a$. Условие а) является тривиальным; кроме того, поскольку из $s \geq 0$, $t \geq 0$ следует $s+t \geq s+0 \geq 0 + 0 = 0$, то имеет место «б». При сделанных в «в» предположениях имеем в силу (3) $0 = x+y \geq x, y \geq 0$, откуда $x = y = 0$, чем и доказано «в». Наконец, в силу (3) оба множества $a + G^+$ и $G^+ + a$ определяют множество всех $x \geq a$, чем и доказано «г».

Обратно, пусть *G*—группа и G^+—любое подмножество в *G,* которое удовлетворяет перечисленным условиям; определим $x \geq y$ в соответствии с (3). Тогда из условий «а», «б», «в» следуют соответственно правила Р1, Р3, Р2, определяющие частично упорядоченное множество. Далее, предположим, что $x - y \in G^+$, и используем «г» в форме «г'» $G^+ = -a + G^+ + a$. Тогда

$$(a + x + b) - (a + y + b) = a + x + b - b - y - a =$$
$$= a + (x - y) - a \in G^+;$$

следовательно, имеет место (1), так что *G* есть частично упорядоченная группа.

Замечание. Условия «а» — «б» утверждают, что G^+ есть полугруппа. Условие «г» в форме «г'» утверждает, что G^+ инвариантно относительно всех внутренних автоморфизмов группы G. Наиболее легкий способ показать, что заданная система является l-группой, состоит обычно в проверке этих условий и в использовании затем следующего результата.

Теорема 2. *Частично упорядоченная группа G тогда и только тогда является l-группой, если для всех* $a \in G$ *существует* $a \cup 0 = a^+$.

Доказательство. Если G есть l-группа, то $a \cup 0$ существует для всех a. Обратно, пусть G—любая частично упорядоченная группа, в которой a^+ существует для всех a. В силу (2) мы имеем

$$(a-b)^+ + b = [(a-b) \cup 0] + b = a \cup b \text{ для всех } a, b.$$

Но по предположению $(a — b)^+ + b$ существует; следовательно, $a \cup b$ всегда существует. Двойственно этому, поскольку из $x \geq 0$ следует $0 = x - x \geqslant 0 - x = -x$ и обратно, соответствие $x \to -x$ есть дуальный автоморфизм. Следовательно,

$$-(-a \cup -b) = a \cap b \text{ всегда существует.} \qquad (5)$$

6.15. 2. Примеры

Мы проиллюстрируем теперь на нескольких простых примерах широкую распространенность l-групп в математике. Во всех рассматриваемых случаях теоремы 1 — 2 дают наиболее удобный способ проверки условий, определяющих l-группы. Мы начинаем с коммутативных или «абелевых» l-групп.

Пример 1. G — аддитивная группа вещественных чисел; G^+ состоит из всех неотрицательных вещественных чисел.

Пример 2. G — группа всех положительных рациональных чисел по умножению (целое число один является единицей группы); G^+ есть множество всех целых положительных чисел. В этом случае x^+ есть числитель числа x, представленного несократимой дробью.

Пример 3. G — группа по сложению всех векторов $x = (x', x'')$ с двумя вещественными компонентами; G^+ содержит x тогда и только тогда, если $x' > 0$, или $x' = 0$ и $x'' \geq 0$.

Все рассматриваемые в р. 6.16 «векторные структуры» являются коммутативными lгруппами. Мы выбираем два примера.

Пример 4. G— аддитивная группа всех непрерывных вещественных функций, определенных на интервале $0 \leq x \leq 1$; G^+ включает в себя те из указанных функций, которые неотрицательны (удовлетворяют условию $f(x) \geq 0$ для всех x). Здесь $f^{+}(x)$ есть для всех x наибольшее из чисел $f(x)$ и 0.

Пример 5. G — аддитивная группа функций ограниченной вариации на отрезке $0 \leq x \leq 1$ таких, что $f(0) = 0$. G^{+} состоит из всех таких «неубывйющих» функций (функций, для которых из $x \geq y$ следует $f(x) \geq f(y)$). Здесь $f^{+}(x)$ — положительная вариация функции $f(x)$ в обычном смысле.

Примером дискретной некоммутативной l-группы является следующий. (Такая группа должна быть бесконечной в силу теоремы 7.)

Пример 6. Группа G имеет три образующих a, b, c бесконечного порядка и определяющие соотношения $a + b = c + a$, $a + c = b + a$, $b + c = c + b$. G^* содержит $ma+nb+n'c$ тогда и только-тогда, если $m > 0$, или $m = 0$, в то время как $n \geq 0$ и $n' \geq 0$.

Мы дадим теперь два примера l-групп Ли.

Пример 7. Группа G состоит из всех пар (x, y) вещественных чисел, и сложение определено по формуле

$$(x, y) + (x', y') = (x + x', y + y').$$

G^{+} состоит из тех пар, у которых $x > 0$ или $x = 0$, $y \geq 0$.

Пример 8. G есть группа по умножению всех матриц; следующего вида:

$$X = \begin{pmatrix} 1 & a & c \\ 0 & 1 & b \\ 0 & 0 & 1 \end{pmatrix}.$$

$X \geq$ означает, что $a > 0$, или $a = 0$ и $b > 0$ или $a = b = 0$ и $c \geq 0$.

Следующий пример некоммутативной бесконечной непрерывной l-группы был рассмотрен Эверетом и Уламом.

Пример 9. G есть группа с естественным умножением всех гомеоморфизмов интервала $0 \leq x \leq 1$. G^{+} состоит из тех гомеоморфизмов $f(x)$, которые удовлетворяют условию $f(x) \geq x$ дла всех x. Как структура, G является, таким образом, подструктурой приведенной в примере 4 структуры, но групповая операция здесь иная. Таким образом, $f^{+}(x)$ есть наибольшее из чисел x и $f(x)$.

6.15.3 Направленные группы как полугруппы

Под *направленной группой* мы понимаем частично упорядоченную группу, обладающую свойством Мура —Смита.

Для любых заданных $a, b \in G$ существует такое $c \in G$, что

$$c \geq a \text{ и } c \geq b. \tag{6}$$

Очевидно, что любая l-группа, будучи структурой, является направленной группой.

Лемма 1 (Клиффорд). *Во всякой частично упорядоченной группе свойство Мура—Смита эквивалентно утверждению, что*

каждый элемент является разностью положительных элементов. (7)

Доказательство. Предполагая условие (6) с $b = 0$, мы получаем $a = c —(—a + c)$, где $c \geq 0$ и $—a+c=—a + (c —a)+ a \geq — a + 0 + a = 0$. Обратно, если $a=-a' — a''$ и $b=b' — b''$, где a', a'', b', b'' положительны, то $c = a'+ b'$ является верхней гранью для a и b.

Мы покажем теперь, что любая направленная группа определяется с точностью до изоморфизма полугруппой своих положительных элементов, упорядоченной при помощи отношения делимости. Отсюда следует, что l-группы можно рассматривать как естественные расширения до групп достаточно хорошо определенного класса полугрупп. Построение упорядоченной аддитивной группы всех целых чисел, исходя из полугруппы целых положительных чисел, включается сюда как специальный случай. Заметим сперва, что множество S всех положительных элементов произвольной направленной группы G удовлетворяет условиям «а» — «г» теоремы 1, а также закону сокращения

из $x+a = y+a$ или $b+x = b+y$ следует $x = y$. (8)

В силу (8) и «г», предполагаемых только для S, мы видим, что,

если заданы $a, x \in S$, то существует в точности одно $x_\alpha \in S$, удовлетворяющее соотношению $x +a = a + x_\alpha$. (9)

Эти условия мы можем переформулировать следующим образом. S есть полугруппа, удовлетворяющая законам сокращения. В этой полугруппе S $a \leq b$ означает, что b является левым кратным элемента a, или, эквивалентно этому, что b является правым кратным элемента a. Таким образом, $a \leq b$ означает, что $b \in S+a=a+S$.

В любой такой полугруппе $x+a+b=a+x_a+b=a+b+(x_a)_b$ и $x+y+a=x+a+y_a=a+x_a+y_a$. Поэтому мы имеем

$$(x_a)_b = x_{a+b} \text{ и } (x+y)_a = x_a + y_a. \qquad (10)$$

Эти правила очевидны для групп, где $x_\alpha=—a+x+a$. Результатом является тот факт, что они справедливы в любой полугруппе S, удовлетворяющей закону сокращения (8) и слабому коммутативному закону $S + a = a + S$ для всех $a \in S$. Мы можем построить группу G из формальных разностей $(b, a) = (b — a)$, элементов из S, полагая

$(b — a) = (d — c)$ в G означает, что $b+c = d+ a_c$ в S, (11)

$(b — a) + (d — c)$ в G есть $((b + d_a) — (c + a))$, (12)

где $b+d_a$ и $c+a$ принадлежат S. Простые выкладки показывают, что формулы (11) — (12) определяют всегда группу, в которой пары

(b — 0) образуют полугруппу, изоморфную S, и что в группе G (b — a) = (b — 0) — (a — 0). С использованием теоремы 1 получается

Теорема 3. *Направленная группа может быть определена как расширение до группы G полугруппы S, в которой* а) *имеет место закон сокращения,* б) $S+a=a+S$ *для всех* $a\in S$, в) $a+b==0$ *влечет* $a=b=0$.

Следствие 1. *l-группа может быть определена в соответствии с теоремой* 3, *если предположить дополнительно, что* г) *любые два элемента из S имеют наименьшее общее кратное в S.*

В действительности, условие «в» можно было бы отбросить, если бы мы пожелали ввести отношение эквивалентности.

6.15.4. Основные алгебраические правила

Поскольку соответствие $x\to -x$ есть дуальный автоморфизм, то дуальным автоморфизмом будет и соответствие $x \to a - x + b$, и мы имеем в любой l-группе

$$a-(x\cap y)+b=(a-x+b)\cup(a-y+b). \qquad (13)$$

Полагая в (13) $x = a$ и $y=b$, получается

Теорема 4. *Во всякой l-группе мы имеем для всех a и b*

$$a-(a\cap b)+b=b\cup a. \qquad (13')$$

Следствие 1. *Во всякой коммутативной l-группе имеем*

$$a+b=(a\cup b)+(a\cap b) \text{ для всех } a \text{ и } b. \qquad (13'').$$

В примере 2, п. 2, *модулярный закон* (13") сводится к известному тождеству $ab = (a, b)\,[a, b]$ теории чисел.

Определение. *Положительной частью a^+ элемента a называется элемент $a\cup 0$; отрицательной частью a^- элемента a называется элемент $a\cap 0$; модулем $|a|$ элемента a называется элемент $a\cup -a$.*

Полагая в следствии 1 $b = 0$, мы получаем

Следствие 2. *Для любого a $a = a^+ + a^-$.*

Словами, каждый элемент l-группы есть сумма своей положительной и отрицательной части (так называемое жорданово разложение).

Теорема 5. *Всякая l-гуппа является дистрибутивной структурой.*

Доказательство. В силу следствия 1 Бергмана из теоремы 1 р. 6.10 достаточно показать, что из $a\cap x = a\cap y$ и $a\cup x = a\cup y$ следует $x = y$. Но в силу (13') из этих соотношений следует

$$x=(a\cap x)-a+(x\cup a)=(a\cap y)-a+(y\cup a)=y.$$

Теорема 6. *Во всякой l-группе мы имеем*

$$\text{из } a\cap b = 0 \text{ и } a\cap c =0 \text{ следует } a\cap(b + c) = 0, \qquad (14)$$

из $a \cup b = 0$ и $a \cup c = 0$ *следует* $a \cup (b + c) = 0$. (14')

Доказательство. Если использовать двойственность, то зто является специальным случаем теоремы 7 р. 6.14. Однако мы дадим новое доказательство свойства (14). Так как *a, b* и *c* положительны, то, очевидно, $a \cap (b + c) \geq 0$. Но в силу соотношений, двойственных (2),

$$0 = 0 + 0 = (a \cap b) + (a \cap c) = (a + a \cap c) \cap (b + a \cap c) =$$
$$= (a + a) \cap (a + c) \cap (b + a) \cap (b + c) \geqslant a \cap (b + c).$$

Мы можем перефразировать теорему 6 в терминах понятия дизъюнктивности.

Определение. *Два положительных элемента a и b называются дизъюнктными — что обозначается в виде $a \perp b$ — тогда и только тогда, если $a \cap b = 0$.*

В примере 2 п. 2 это сводится к понятию взаимной простоты. Теорема 6 утверждает, что множество положительных элементов, дизъюнктных любому *a,* замкнуто относительно сложения. Далее, если мы предположим в (13') $a \cap b = 0$, то мы получим, поскольку $b \cup a = a \cup b$:

Лемма 1. *Дизъюнктные элементы перестановочны,*

если $a \cap b = 0$, *то* $a + b = b + a$. (15)

Лемма 2. *Если* $b \cap c = 0$, *то* $(b - c)^+ = b$ *и* $(b - c)^- = -c$.

Доказательство. В силу предыдущих формул
$(b - c) \cup 0 = (b \cup c) - c = b - (b \cap c) + c - c = b$ и двойственно.

Лемма 3. *Если* $na \geq 0$, то $a \geq 0$.

Доказательство. Применяя соотношения, двойственные (2), имеем
$$n(a \cap 0) = na \cap (n-1)a \cap (n-2)a \cap \ldots \cap a \cap 0.$$
Но если $na \cap 0 = 0$, то это равно
$(n-1)a \cap (n-2)a \cap \ldots \cap a \cap 0 = (n-1)(a \cap 0)$. Производя новые сокращения, мы получаем $a \cap 0 = 0$, что и требовалось.

Комбинируя лемму 3 с двойственным утверждением, получается

Теорема 7. *В l-группе каждый элемент, за исключением нуля, имеет бесконечный порядок.*

Другим следствием является тот факт, что во всякой *коммутативной l*-группе из $na \geq nb$ следует $n(a - b) \geq 0$, а потому $a \geq b$.

Лемма 4. *Положительная и отрицательная части любого элемента дизъюнктны между собой,*

$$\textit{для любого } a \quad (a \cup 0) \cap (-a \cup 0) = a^+ \cap (-a^-) = 0. \qquad (16)$$

Доказательство. Очевидно что $-(a \cup -a) = -a \cap a$.

Следовательно, $$2(a \cup -a) = (a \cup -a) - (a \cap -a) \geqslant 0.$$

Следовательно, в силу леммы 3 $a \cup -a \geq 0$, так что (переходя к двойственному выражению) $a \cap -a \leq 0$. Используя это и теорему 5, мы получаем $(a \cup 0) \cap (-a \cup 0) = (a \cap -a) \cup 0 = 0$. Из $|a|=0$ следует $a \cup -a = a \cap -a$, так что $a = -a$, $2a = 0$ и $a = 0$.

Теорема 8. *Во всякой l-группе модули удовлетворяют условиям*

$$\text{если } a \neq 0, \text{ то } |a| > 0, \text{ в то время как } |0| = 0, \tag{17}$$

$$|na| = |n| \cdot |a| \text{ для любого целого числа } n, \tag{18}$$

$$|a-b| = (a \cup b) - (a \cap b), \tag{19}$$

$$|(a \cup b) - (a^* \cup b)| \leqslant |a - a^*| \text{ и двойственно.} \tag{20}$$

Доказательство. Формула (17) вытекает из (19) при $b = 0$. Формула (19) при $b = 0$ дает также $|a| = a^+ - a^-$, но в силу лемм 1 и 4 элементы a^+ и a^- дизъюнктны и перестановочны. Следовательно, $na = n(a^+) - n(a^-)$ и в силу теоремы 6 и индукции $n(a^+) \cap n(-a^-) = 0$. Следовательно, в силу леммы 2 $(na)^+ = n(a^+)$ и $(na)^- = n(a^-)$. Отсюда вытекает формула (18) для положительных n. Для отрицательных n она вытекает из того факта, что $|-x| = |x|$.

Докажем теперь формулу (19); мы не можем предполагать доказанными формулы (17) и (18). Применяя последовательно следствие 2 теоремы 4, свойства групповых операций и (2) вместе с двойственными соотношениями, мы подсчитываем

$$|a-b| = (a-b)^+ - (a-b)^- = [(a-b) \cup 0 + b] - [(a-b) \cap 0 + b] =$$
$$= (a \cup b) - (a \cap b).$$

Наконец, для доказательства (20) представим левую часть этого неравенства, используя (19) и дистрибутивный закон (см. теорема 5), в виде $a \cup b \cup a^* - (a \cup b) \cap (a^* \cup b) =$ $= a \cup a^* \cup b - (a \cap a^*) \cup b$. Рассмотрение сводится к случаю $a \geq a^+$, так как в силу (19) правая часть неравенства (20) есть $a \cup a^+ - a \cap a^+$. В этом случае, записывая $a = t + a^*$ $[t \geqslant 0]$, мы имеем

$$a \cup b = (t + a^*) \cup b = t + [a^* \cup (-t + b)] \leqslant t + (a^* \cup b).$$

Вычитая справа $a^+ \cup b$, мы получаем (20), поскольку $t = |a - a^+|$.

6.15.5. Идеалы

Хорошо известно, что отношения конгруентности на любой группе G суть разбиения G на смежные классы по ее различным нормальным делителям N. Поэтому отношениями конгруептности на l-группе являются те разбиения на смежные классы по нормальным делителям, которые обладают свойством замены для двух структурных операций или, эквивалентно этому в силу (4) —(5), характеризуются тем, что из $a \equiv b(\theta)$ следует $a^+ \equiv b^+$ (θ).

Определение. *Под l-идеалом l-группы G понимают нормальный делитель группы G, содержащий вместе с любым a также все элементы x, для которых* $|x|\leq|a|$.

Само G и 0 являются l-идеалами в G; они называются несобственными l- идеалами; все остальные l-идеалы в G называются собственными l-идеалами. Пусть, далее, N—произвольный l-идеал в G и предположим, что $a, b\in N$. Если $a \cap b \leqslant x \leqslant a \cup b$, то

$$|x| = x \cup -x \leqslant (a \cup b) \cup -(a \cap b) = a \cup b \cup -b \cup -a = \\ = |a| \cup |b| \leqslant |a| + |b|.$$

Следовательно, $x\in N$, а потому любой l-идеал в l-группе является *выпуклой* l-подгруппой.

Теорема 9. *Отношения конгруентности на произвольной l-группе G являются разбиениями G на смежные классы по ее различным l-идеалам.*

Доказательство. Если N есть множество элементов, конгруентных 0 при заданном отношении конгруентности, то из $a\in N$ и $|\ x\ |\leq|\ a\ |$ следует $a \cap -a \leqslant x \leqslant a \cup -a$; следовательно, $0 \cap 0 \leqslant x \leqslant 0 \cup 0 \bmod N$, а потому $x\in N$. Обратно, если N — l-идеал, то из $x\equiv x^* \bmod N$ следует в силу (20) $|(x \cup y) - (x^* \cup y)| \leqslant |x - x^*|$, а поэтому $x \cup y \equiv x^* \cup y \bmod N$. Используя симметрию и двойственность, мы видим, что разбиение G на смежные классы по N обладает свойством замены для обеих структурных операций, чем и завершается доказательство.

Далее, отношения конгруентности на произвольной алгебре A образуют полную подструктуру структуры всех разбиений множества элементов алгебры A. Поэтому отношения конгруентности на любой l-группе G образуют подструктуру структуры всех отношений конгруентности на G, когда последняя рассматривается исключительно как структура. Так как она дистрибутивна, то имеет место

Теорема 10. *l-идеалы любой l-группы G образуют полную дистрибутивную структуру; следовательно, то же справедливо для отношений конгруентности на G.*

Комбинируя с гл. VI, § 8, упр. 2 (а), или перефразируя соответствующим образом рассуждение гл. II, § 8, получаем следующий результат.

Теорема 11. *Любые два представления l-группы G в виде прямого произведения (кардинального произведения) имеют общее продолжение. Если структура всех отношений конгруентности на G имеет конечную длину, то G имеет единственное представление в виде прямого произведения неразложимых множителей.*

6.15.6. Единицы

Идея единицы является очень полезной, особенно в случае коммутативных *l*-групп.

Определение. *Под сильной единицей l-группы G понимают такой элемент* $e \in G$*, что для любого* $a \in G$ *имеет место* $ne>a$ *при некотором целом положительном числе n. Два элемента a* и *b из G называются дизъюнктными тогда и только тогда, если* $|a| \cap |b| = 0$/ *Положительный элемент* $c \in G$ *называется слабой единицей, если единственным элементом, дизъюнктным ему, является* 0.

Лемма 1. *Любая сильная единица является слабой единицей.*

Доказательство. В силу леммы 3, п. 4 любая сильная единица положительна. В силу теоремы 6 для любого *e* из $e \cap a = 0$ следует $ne \cap a = 0$ для всех *n.* Но если *e* сильная единица, то $ne \geq a$ для некоторого *n,* а потому из $e \cap a = 0$ следует $a = ne \cap a = 0$, в силу чего *e* — слабая единица.

Не все слабые единицы являются сильными единицами. Например, аддитивная *l*-группа всех вещественных непрерывных функций, заданных на полупрямой $0 \leq x < +\infty$, имеет слабую единицу $f(x) = 1$, которая не является сильной единицей, ибо ни для какого *n* и $f(x) \geq 0$ неравенство $nf(x) \geqslant [f(x)+x]^2$ не может иметь места сразу для всех *x*. С другой стороны, в *l*-группе всех ограниченных вещественных функций, заданных в произвольной области, функция $f(x) = 1$ является сильной единицей.

Даже в *l*-группах без слабых единиц *l*-идеалы могут иметь сильные единицы; такие *l*-идеалы называются *главными* *l*-идеалами. В действительности, как мы сейчас покажем, в любой коммутативной *l*-группе положительный элемент является сильной единицей для подходящего *l*-идеала.

Теорема 12. В любой коммутативной l-группе для любого $a>0$ *множество* $J(a)$ *всех b таких, что* $|b|<na$ *для некоторого целого положительного n, образует l-идеал, имеющий элемент a в качестве сильной единицы. Кроме того, J(a) есть наименьший l-идеал, содержащий a.*

Доказательство. Если $|b| < ma$ и $|c| < na$, то, очевидно, $|b \pm c| < (m+n)a$; в то же время если $|b| < ma$ и $|x| \leq |b|$, то $|x| < ma$; следовательно, $J(a)$ есть *l*-идеал. Очевидно, что *a* является сильной единицей в $J(a)$. Наконец, любой *l*-идеал, содержащий *a,* должен содержать каждое *na,* а потому и все *b*, для которых $|b| < na$.

Пусть теперь *A*—любая абелева *l*-группа. Если она не является просто упорядоченной, то в силу инвариантности упорядоченности относительно группового сдвига она должна содержать такой элемент *а,* что не имеет места ни $a \geq 0$, ни $0 \geq a$. Для этого элемента *а* элементы $a^{+} > 0$ и — $a^{-} > 0$ будут в силу леммы 4, п. 4 дизъюнктными. Следовательно, в силу теорем 6 и 12 $J(a^{+})$ и $J(—a^{-})$ будут дизъюнктными *l* -идеалами. В силу теоремы 9 р. 6.7 имеет место

Лемма 2. *Коммутативная l-группа является либо просто упорядоченной, либо разложима в полупрямое произведение.*

Применяя теперь теорему 10 гл. VI, получаем следующий результат

Теорема 13. *Любая коммутативная l-группа является полупрямым произведением просто упорядоченных l-групп.*

6.15.7. Просто упорядоченные группы; архимедов случай

Просто упорядоченная группа, или *упорядоченная группа,* определяется обычно как частично упорядоченная группа, удовлетворяющая условию

Р4. Для любых заданных *х* и *у* либо $x \geq y$, либо $y \geq x$.

Таким образом, это есть частично упорядоченная группа, являющаяся цепью в смысле р. 6.4, или, эквивалентно этому, частично упорядоченная группа, в которой каждый элемент либо положителен, либо отрицателен, либо нуль. Будучи цепью, упорядоченная группа является структурой, а потому *l*-группой.

Мы видели (см. теорема 7), что во всякой *l*-группе каждый элемент, отличный от нуля, имеет бесконечный порядок. Обратно, имеет место

Теорема 14. *Любая абстрактная коммутативная группа А, все элементы которой, за исключением нуля, имеют бесконечный порядок, может быть просто упорядочена.*

Доказательство. В такой группе уравнение $nx=ma$ имеет самое большее одно решение. Действительно, из $nx=ny$ следует $n(x—y) = 0$, откуда $x = y$. Если $nx = ma$ имеет решение, то мы обозначим его через $\frac{m}{n}a$ и заметим, что все законы векторной алгебры справедливы для определенного так умножения на *рациональные* скаляры.

Под вполне упорядоченным *рациональным базисом* в *А* мы понимаем вполне упорядоченное (конечное или бесконечное) подмножество элементов a_{α} из *А* такое, что каждый ненулевой элемент из *А* является конечной рациональной комбинацией

$$n_1 a_{\alpha(1)} + \ldots + n_r a_{\alpha(r)} \quad [\alpha(1) < \ldots < \alpha(r)]$$

элементов a_α, в то время как из $\sum n_i a_{\alpha(i)} = 0$ следует, что каждое $n_i = 0$, или, эквивалентно этому, из $\sum \frac{m_i}{n_i} a_{\alpha(i)} = 0$ следует, что каждое $m_i/n_i=0$. Существование вполне упорядоченного рационального базиса может быть доказано непосредственно, так как любое максимальное вполне упорядоченное рационально независимое подмножество является базисом. Наконец, относительно такого базиса любой, отличный от 0, элемент из A может быть назван положительным или отрицательным в соответствии с тем, является ли его первый ненулевой коэффициент m_i/n_i положительным или отрицательным. Это «лексикографическое» упорядочивание A превращает A в просто упорядоченную группу.

Следствие. *Коммутативная группа является аддитивной группой некоторой l-группы тогда и только тогда, если она не имеет элементов конечного порядка, кроме нуля.*

В логических основах анализа особенно важную роль играют группы, являющиеся «архимедовыми» в следующем смысле.

Определение. *Элемент a частично упорядоченной группы называется несравненно меньшим, чем второй элемент b (что обозначается в виде $a \ll b$), тогда и только тогда, если $na \leq b$ для любого целого n. Частично упорядоченная группа называется архимедовой тогда и только тогда, если из $a \ll b$ следует $a = 0$.*

Иначе говоря, $a \ll b$ означает, что b является верхней гранью для всей циклической подгруппы, порожденной элементом a. Так, например, в примере 3 $(0, 1) \ll (1, 0)$; в построении теоремы 14 $a_1 \gg a_2 \gg a_3 \gg \ldots$ Легко проверить, что отношение $\ll$ антисимметрично и транзитивно и что любая подгруппа архимедовой l-группы сама является архимедовой относительно того же самого отношения упорядоченности независимо от того, является ли она l-подгруппой или нет. Далее, имеет место

Лемма 1. *l-Группа G тогда и только тогда является архимедовой, если*

$$\text{из } na \leq b \text{ для всех } n=1, 2, 3, \ldots \text{ следует } a \leq 0. \qquad (*)$$

Доказательство. Если G удовлетворяет условию (*) и $na \leq b$ для $n=0, \pm1, \pm2, \ldots$, то $na \leq b$ и $n(-a) \leq b$ для $n=1, 2, 3, \ldots$ Следовательно, в силу (*) $a \leq 0$ и $-a \leq 0$, а потому $a = 0$. Таким образом, l-группа G архимедова; доказательство справедливо для любой частично упорядоченной группы.

Обратно, предположим, что l-группа G архимедова и что $na \leq b$ для всех $n = 1, 2, 3, \ldots$ Тогда, как и в доказательстве теоремы 8, $na^+ = (na)^+ = na \cup 0 \leqslant b \cup 0$ для $n = 1, 2, 3, \ldots$ Но $na^+ \leq 0 \leq b^+$ для $na = 0, -1, -2, \ldots$; следовательно, $na^+ \leq b^+$ для $n = 0, \pm1, \pm2, \ldots$ Так как G

архимедова, мы выводим отсюда, что $a^+ = 0$ и $a = a^+ + a^- \leq 0$, что и требовалось доказать.

Лемма 2. *Просто упорядоченная группа G является архимедовой тогда и только тогда, если любое e > 0 есть сильная единица.* **Доказательство.** Если любое $e>0$ есть сильная единица, то из $a \neq 0$ следует $n|a| > b$ для всех b и подходящих n; но $|a|=\pm a$; следовательно, $(\pm n)a>b$ для некоторого n, а потому группа G архимедова. Обратно, если группа G архимедова, то в силу леммы 1 неравенство $ne\leq b$ не может иметь места для всех n; следовательно, для любого $e > 0$ некоторое $ne > b$, а потому e является сильной единицей.

Лемма 3. *В просто упорядоченной группе $na \geq nb$ влечет $a\geq b$, а потому $na = nb$ влечет $a = b$.*

Доказательство. Если $a \ngeq b$, то $a<b$, откуда $na < nb$, так как сложение сохраняет упорядоченность. Но это противоречит предположению. Мы заметим, что в l-группе примера 9 из $na = nb$ не следует $a = b$.

Теорема 15. *Любая просто упорядоченная архимедова группа G изоморфна подгруппе аддитивной группы всех вещественных чисел, а потому является коммутативной.*

Доказательство. Мы покажем сперва, что архимедова группа G обязательно коммутативна. Достаточно рассмотреть для этого положительные элементы. Возможны два случая: случай, когда в G имеется наименьший положительный элемент, и случай, когда такого элемента не существует. Предположим, что μ — наименьший положительный элемент. Если $\alpha>0$, то $\mu\leq\alpha$. Следовательно, существует целое положительное число n, для которого $n\mu\leq\alpha < (n+1)\mu$. Но тогда $0\leq\alpha - n\mu<\mu$, откуда $\alpha= n\mu$. Поэтому группа G изоморфна аддитивной группе целых чисел. Предположим теперь, что в группе G не существует наименьшего положительного элемента. Мы установим прежде всего, что для каждого $\xi > 0$ существует такой элемент η, что $0 < \eta < \xi$ и $2\eta\leq \xi$. Так как $\xi > 0$, существует такой элемент ζ, что $0 < \zeta< \xi$. Тогда либо $2\zeta\leq \xi$, либо $2\zeta> \xi$. В первом случае мы полагаем $\eta = \zeta$. Во втором случае $\zeta > \xi-\zeta > 0$, следовательно, $0 > -\zeta + \xi-\zeta$, откуда $\xi > \xi-\zeta+\xi-\zeta = 2(\xi-\zeta)$, и мы полагаем $\eta = \xi - \zeta$.

Предположим теперь, что для некоторых элементов $\alpha, \beta > 0$ $\alpha+\beta \neq \beta+\alpha$. Мы можем считать, что $(\alpha+\beta)-(\beta+\alpha)=\zeta > 0$. Выберем тогда такой элемент ξ, что $0 < \xi$, $\xi < \alpha$, $\xi < \beta$ и $2\xi \leqslant \zeta$. В силу архимедовости группы существуют такие целые положительные числа m и n, что $m\xi \leqslant \alpha < (m+1)\xi$ и $n\xi \leqslant \beta < (n+1)\xi$. Следовательно, $(m+n)\xi \leqslant \alpha+\beta < (m+n+2)\xi$ и $-(m+n+2)\xi <$

$< -(\beta+\alpha) \leqslant -(m+n)\xi$. Отсюда мы получаем $-2\xi < \zeta < 2\xi \leqslant \zeta$, т. е. $\zeta < \zeta$, что невозможно. Коммутативность группы G тем самым доказана.

Для того, чтобы осуществить изоморфизм между группой G и подгруппой группы всех вещественных чисел, мы поставим в соответствие нулю группы G вещественное число 0 и произвольному фиксированному элементу $\alpha>0$ число 1. Пусть теперь β— произвольный элемент из G. Определим дедекиндово сечение $\bar{\beta}$ в группе всех вещественных чисел, принимая, что рациональное число m/n $(n > 0)$ принадлежит нижнему классу сечения $\bar{\beta}$, если $m\alpha \leq n\beta$, и верхнему классу сечения, если $m\alpha > n\beta$. Из архимедовости группы G вытекает, что ни один из этих классов не может быть пустым. Для доказательства того, что этим действительно определяется сечение, положим, что m/n элемент верхнего класса, а p/q элемент нижнего класса ($n > 0$, $q > 0$). Тогда из $m\alpha>n\beta$ и $p\alpha \leq q\beta$ следует, что $mq\alpha>nq\beta$ и $pn\alpha \leq qn\beta$. Следовательно, $mq\alpha>pn\beta$, а потому $mq > pn$, т. е. $m/n > p/q$, что и требовалось доказать. Вещественное число $\bar{\beta}$, определяемое этим сечением, мы примем в качестве образа элемента β. Из построения β и коммутативности группы G следует, что соответствие, $\beta \to \bar{\beta}$ является изоморфизмом, сохраняющим упорядоченность!.

6.15.8. Упорядоченные тела

Упорядоченное кольцо может быть определено как кольцо, являющееся просто упорядоченной группой по сложению, в которой произведение двух любых положительных элементов само положительно. Так как $a(-b)=-ab$, то этим гарантируется, что произведение любого положительного элемента и отрицательного элемента отрицательно и что произведение двух отрицательных элементов всегда положительно. Упорядоченное кольцо, являющееся полем, называется *упорядоченным полем.*

Пример 10. Пусть F состоит из 0 и обычных полубесконечных формальных *степенных рядов* относительно x с вещественными коэффициентами a_k и с (положительными или отрицательными) целыми показателями

$$p(x) = a_n x^n + a_{n+1}x^{n+1} + a_{n+2}x^{n+2} + \ldots \quad [a_n \neq 0]. \tag{21}$$

Если мы определим

$$\sum a_k x^k + \sum b_k x^k = \sum (a_k + b_k)\, x^k$$

и

$$\left(\sum a_h x^h\right)\left(\sum b_h x^h\right) = \sum_k \left(\sum_{i+j=k} a_i b_j\right) x^k,$$

то мы получим хорошо известное поле. Если мы определим $f > 0$ в том смысле, что начальный коэффициент $a_n>0$, то мы получим упорядоченное поле. В примере 10 мы могли бы брать коэффициенты из любого упорядоченного поля, не меняя конструкции. Таким образом, мы можем по индукции составить упорядоченное поле из степенных рядов относительно n-переменных. Можно построить также *некоммутативные* упорядоченные поля (называемые иначе упорядоченными телами или упорядоченными кольцами с делением) следующим образом.

Пример 11. Пусть G состоит из 0 и степенных рядов

$$g(y) = p_m y^m + p_{m+1} y^{m+1} + p_{m+2} y^{m+2} + \dots \quad [p_m \neq 0] \qquad (21')$$

с коэффициентами p_k в упорядоченном поле примера 10. Пусть сложение и порядок определены так же, как и в примере 10. Но пусть определение умножения исходит из основной формулы $yx = 2xy$, так что если $p = \sum_n^{\infty} a_k x^k$, то

$$yp = y\left(\sum_n^{\infty} a_k x^k\right) = \left(\sum_n^{\infty} 2^k a_k x^k\right) y = \tau(p)\, y$$

и, вообще,

$$\left(\sum a_{hi} x^h y^i\right)\left(\sum b_{jk} x^j y^k\right) = \sum 2^{ij} a_{hi} b_{jk} x^{h+j} y^{i+k}.$$

Так как $p \to \tau(p)$ есть автоморфизм поля коэффициентов, то это определяет кольцо R. (Это подобно основной конструкции циклических алгебр и «скрещенных произведений».) Кроме того, в G можно выделить обратные элементы, производя последовательно деления начальных коэффициентов на начальные коэффициенты и получая при этом в остатке более высокие степени. Мы предоставляем читателю доказать, что всегда возможно точное деление с любой из сторон любого одночлена на любой другой (не равный нулю) одночлен.

Вообще, пусть F—любое упорядоченное поле или тело. Для любых заданных $a, b \in F$ будет в силу определения либо $a \ll b$, либо $b \ll a$, либо $na \leq b$ и $nb \leq a$ для достаточно большого n. В последнем случае мы пишем $a \sim b$ и говорим, что a и b имеют *один и тот же порядок величины.* Так как $\sim$ есть отношение эквивалентности, то эта терминология вполне законна; мы можем даже определить порядок величины как порядок величины классов эквивалентности, определяемых с помощью $\sim$.

Теорема 16. *«Порядки величин» любого упорядоченного поля или тела сами образуют упорядоченную группу.*

Доказательство. Так как $nx \leq b$ тогда и только тогда, если $anx=n\ (ax) \leq ab$ и симметрично этому, то «упорядоченность» с помощью << сохраняется при всех групповых сдвигах $x \rightarrow -axa_1$, а потому она является отношением иметь один и тот же порядок величины.

Так, в примере 10 эта упорядоченная группа есть аддитивная группа целых чисел; степень первого члена ряда $f(x)$ представляет собой порядок величины $f\ (x)$.

6.15.9. Полные *l*-группы

Вещественные числа по сложению, а также многие другие *l*-группы, важные для анализа, удовлетворяют следующему условию.

Определение. *Частично упорядоченная группа G называется полной (σ-полной) тогда и только тогда, если в ней каждое непустое {соответственно, счетное) ограниченное множество имеет наиб. н. г. и наим. в. г.*

Пусть x любой элемент произвольной частично упорядоченной группы G. Если s произвольная верхняя грань для множества $\{nx\}$, то, в силу инвариантности упорядоченности относительно группового сдвига верхними гранями будут также $s+x$ и $s - x$. Поэтому, за исключением случая $x = 0$, множество $\{nx\}$ не может иметь наименьшей верхней грани; этим доказана

Лемма 1. *Множество всех целых кратных (степеней) ненулевого элемента частично упорядоченной группы не может иметь наим. в. г.*

Поэтому, за исключением тривиального случая $G = 0$, никакая частично упорядоченная группа не может быть полной или даже σ-полной в смысле рл. 6.5; в этом причина того, что мы видоизменили нашу терминологию.

Теорема 17. *Любая подгруппа* σ*-полной частично упорядоченной группы G является архимедовой* и *даже «целозамкнутой» в том смысле, что*

из $na \leq b$ *для всех* $n = 1, 2, 3, \ldots$ *и* (*)
фиксированного b *следует* $a \leq 0$.

Обратно, пополнение непустыми сечениями любой целозамкнутой частично упорядоченной группы A есть полная частично упорядоченная группа.

Доказательство. В силу определения σ-полноты из $na \leq b$ в G для всех $n = 1, 2, 3, \ldots$ следует, что $c = \vee\, na$ существует. Но отсюда вытекает, поскольку множество элементов $(n+1)a$ есть подмножество множества всех na, что $c+a=\vee\ (n+1)a\leq\ \vee\ na=c$; это же означает, что $a\leq 0$.

Обратно, пусть A—любая частично упорядоченная группа; мы образуем пополнение $\overline{A}$ группы A непустыми сечениями следующим образом. Любое непустое подмножество X из A, множество $U(X)$ верхних граней которого непусто, определяет единственным образом сечение $(L(U(X)), U(X))$ в A —следовательно, единственный элемент в $\overline{A}$. Кроме того, множества X и X_1 определяют одно и то же сечение тогда и только тогда, если $U(X) = U(X_1)$. Обратно, каждый элемент из $\overline{A}$ определяется по такому способу одним или более чем одним множеством X.

Сложение в $\overline{A}$ мы определим, принимая в качестве $X+Y$ множество всех $x+y[x\in X,\ y\in Y]$. Оно всегда определено, поскольку $U(X)+U(Y)$ непусто. Это сложение является ассоциативным; кроме того, можно легко доказать, что множество A^- элементов $x\leq 0$ из A является нулем относительно сложения в $\overline{A}$, т. е. $U(X+A^-)=U(A^-+X)=U(X)$ для всех X.

Наконец, сложение *однозначно:* подстановка в сумму равных слагаемых оставляет сумму неизменной. Действительно, если $U(X) = U(X_1)$, то $u\in U(X+Y)$ означает, что $u\geq x+y$ для всех $x\in X$, Но это значит, что $u-y\geq x$, или $u - y\in U(X)$, или $u - y\in U(X_1)$, или $u - y\geq x_1$ для всех $x_1\in X_1$, $y\in Y$, т. е. $u\in U(X_1+Y)$. Следовательно, из $U(X) = U(X_1)$ следует $U(X+Y)=U(X_1+Y)$ и аналогично для правых слагаемых.

Далее, из $X\leq X_1$ (что интерпретируется эквивалентным образом как теоретико-множественное включение или включение по упорядоченности) следует $X+Y\leq X_1+Y$; следовательно, сложение сохраняет порядок. Остается поэтому лишь показать, что существуют обратные элементы. Это мы теперь и сделаем.

Пусть $-U$ обозначает множество элементов, обратных по сложению элементам множества $U(X)$; рассмотрим $Z = X + (-U)$. Так как $x\leq u$ для всех $u\in U(X)$, $z = x+(-u)\leq 0$ для всех $z\in Z$; следовательно, $0\in U(Z)$ и $U(Z)\geq A^+$. Чтобы убедиться в том, что Z есть нулевой элемент, достаточно поэтому показать, что каждое $c\in U(Z)$ положительно. В самом деле, предположим, что $c\geq x+(-u)$, т. е. $c+u\geq x$ для всех $x\in X$, $u\in U$. Фиксируем u_1. Тогда, поскольку $c+u_1\geq x$ для всех $x\in X,\quad u_2=c+u_1\in U(X)$. Поэтому для всех $x\in X\quad c+u_2=2c+u_1\geq x$; следовательно, $c+u_2\in U(X)$. Повторяя этот процесс, мы получаем последовательность u_1, u_2, u_3, ... такую, что $c+u_n=nc+u_1\geq -x$ для любого $x\in X$. Следовательно, для u_1, любого $x \in X$ и всех n = 1, 2, 3, ... имеем $n(-c) = -nc \leqslant -x + u_1$. Но в силу (*) отсюда следует, что $-c \leq 0$, или $c \leq 0$, что и требовалось доказать.

Следствие. *Направленная группа тогда и только тогда является подгруппой полной l-группы, если она «целозамкнута».*

Действительно, в направленной группе G каждая пара x, y имеет верхнюю грань, а следовательно, и наим. в. г., если группа G условно полная.

6.15.10. Бесконечная дистрибутивность

В силу определения любой групповой сдвиг в l-группе является структурным автоморфизмом. Следовательно, он переводит бесконечные объединения и пересечения, соответственно, в бесконечные объединения и пересечения. Этот факт может быть выражен формально в виде бесконечных дистрибутивных законов

$$a + \bigvee x_\alpha = \bigvee (a + x_\alpha), \quad a + \bigwedge x_\alpha = \bigwedge (a + x_\alpha),$$
$$\bigvee x_\alpha + b = \bigvee (x_\alpha + b), \quad \bigwedge x_\alpha + b = \bigwedge (x_\alpha + b). \tag{22}$$

Аналогично, поскольку всякое соответствие вида $x \to a - x + b$ есть дуальный автоморфизм, мы имеем закон

$$a - \bigvee x_\alpha + b = \bigwedge (a - x_\alpha + b) \text{ и двойственно.} \tag{22'}$$

Пусть теперь $v = \vee x_\alpha$. Тогда в силу (20) для всех a и α

$$0 \leqslant (a \cap v) - (a \cap x_\alpha) \leqslant v - x_\alpha.$$

Но в силу (22')

$$\bigwedge (v - x_\alpha) = v - \bigvee x_\alpha = v - v = 0,$$

и в силу предыдущего соотношения

$$0 \leqslant \bigwedge [(a \cap v) - (a \cap x_\alpha)] \leqslant \bigwedge (v - x_\alpha) = 0,$$

следовательно,

$$0 = \bigwedge [(a \cap v) - (a \cap x_\alpha)] = (a \cap v) - \bigwedge (a \cap x_\alpha).$$

Перенося второе слагаемое в другую часть равенства, мы получаем первый из бесконечных дистрибутивных законов

$$a \cap \bigvee x_\alpha = \bigvee (a \cap x_\alpha) \text{ и } a \cup \bigwedge x_\alpha = \bigwedge (a \cup x_\alpha); \tag{23}$$

второй вытекает из двойственных соображений. Резюмируя, получаем следующий результат.

Теорема 18. *Бесконечные дистрибутивные законы* (22) — (23) *справедливы во всякой полной l-группе.*

Следствие. *Любая σ-полная l-группа G является топологической группой и топологической структурой в своей топологии упорядоченности: из $x_n \to x$ и $y_n \to y$ следует* $x_n + y_n \to x + y$,

$x_n \sqcap y_n \to x \sqcap y,\ x_n \sqcup y_n \to x \sqcup y.$

Доказательство. Случай групповой операции является типичным. Мы полагаем $u_n = \vee_{h>n} x_h$ и $v_n = \vee_{k>n} y_k$; тогда в силу определения (см. р. 6.5.8) $\wedge u_n = x$ и $\wedge v_n = y$. В силу определения имеем также, используя дважды (22),

$$\operatorname{Lim\,sup}\{x_n + y_n\} = \bigwedge_n \{\bigvee_{m>n} (x_m + y_m)\} = \bigwedge_n \{\bigvee_{h,\,k>n} (x_h + y_k)\} =$$

$$= \bigwedge_n \{\bigvee_{h>n} [\bigvee_{k>n} (x_h + y_k)]\} = \bigwedge_n \{\bigvee_{h>n} [x_h + v_n]\} = \bigwedge_n \{u_n + v_n\}.$$

Далее, для любых заданных *m* и *n*, $u_m + v_n \geqslant u_{m+n} + v_{m+n}$; следовательно, используя дважды (22), имеем

$$\bigwedge_n \{u_n + v_n\} \leqslant \bigwedge_{m,\,n} \{u_m + v_n\} = \bigwedge_m \{\bigwedge_n (u_m + v_n)\} = \bigwedge_m \{u_m + y\} = x + y.$$

Двойственным образом $\operatorname{Lim\,inf}\{x_n + y_n\} \geqslant x + y$, чем и завершается доказательство.

6.15.11. Замкнутые *l*-идеалы

Пусть *J*—любой *l*-идеал *l*-группы *G*, обладающий дополнением *J*'. Тогда любое $a \in G$ имеет одну «компоненту» a' в *G/J* и одну a'' в *G/J';* кроме того, так как $J \cap J' = 0$, элемент *a* вполне определяется парой (a', a''). Далее, *J* состоит из пар (0, a''), а *J'* из пар (a', 0); так как $J+J'=G$, (a',0) и (0, a'') существуют для любого $(a',a'') \in G$. Этим доказано, что *обладающие дополнениями l-идеалы соответствуют прямым множителям группы G.*

Мы расширим теперь наше прежнее определение дизъюнктности, полагая, что $a \perp b$ (словами, *a* и *b* дизъюнктны) означает $|a| \cap |b| = 0$. Используя терминологию и систему обозначений р. 6.5.5, мы определим тогда J^* как множество всех $b = (b', b'') \perp (0, a'')$ для всех $a = (0, a'') \in J$. Это означает что $|b''| \cap |a''| = 0$ для всех a'' (включая $a'' = b''$); следовательно, $b'' = 0$. Обратно, всегда $(b', 0) \perp (0, a'')$. Поэтому, если идеал *J* имеет дополнение, то $J' = J^*$ и $J = J'' = J^{**}$.

Более общо, пусть S^* «поляра» *любого* подмножества *S полной* *l*-группы *G*. Если $\{b_\alpha\}$ любое подмножество в S^*, т. е. $|b_\alpha| \cap |s| = 0$ для всех $s \in S$, и если $\vee |b_\alpha|$ существует (т. е. если множество всех b_α ограничено), то в силу теоремы 18 $\vee |b_\alpha| \cap |s| = 0$, откуда, поскольку $0 \leqslant |\vee b_\alpha| \leqslant \vee |b_\alpha|$, $|\vee b_\alpha| \cap |s| = 0$, а потому $\vee b_\alpha \in S^*$. Далее, если $b \in S^*$ и $c \in S^*$, то

$$|b+c| \leqslant |b|+|c|+|b|, \tag{24}$$

так как $-|b|-|c|-|b| \leqslant b+c \leqslant |b|+|c|+|b|$; следовательно, в силу теоремы 6 $b + c \in S^*$. Точно так же, если $b \in S^*$ и $|c|\leq|b|$, то $c \in S^*$, и, в частности, $-b \in S^*$. Следовательно, S^* есть *замкнутая абсолютная подгруппа* в G в том смысле, что а) S^* есть подгруппа, б) если $b \in S^*$ и $|c|\leq|b|$, то $c \in S^*$, в) если каждое $b_\alpha \in S^*$ и $\vee b_\alpha$. существует в G, то $\vee b_\alpha \in S^*$. Мы изучим теперь замкнутые абсолютные подгруппы и покажем в конечном итоге, что они являются l-идеалами, обладающими дополнениями.

Любая замкнутая абсолютная (т. е. выпуклая) подгруппа, являющаяся нормальным делителем с теоретико-групповой точки зрения, есть замкнутый l-идеал в следующем смысле.

Определение. *l-идеал J полной l-группы G называется замкнутым тогда и только тогда, если он содержит вместе с любым ограниченным подмножеством $\{x_\alpha\}$ также $\vee x_\alpha$.*

Замечание. Так как соответствие $x \to -x$ сохраняет J, как множество, неизменным и изменяет порядок на обратный, отсюда вытекает, что J содержит также $\wedge x_\alpha$. Далее, так как любой l-идеал является выпуклым, замыкание в приведенном выше смысле эквивалентно топологическому замыканию в топологии упорядоченности.

Для любого подмножества T полной l-группы G $T \cap T^* = 0$. Мы покажем теперь, что если T замкнутая абсолютная подгруппа, то также $T + T^* = G$. Чтобы это показать, мы определим T-компоненту x_T любого элемента $x\geq 0$ из G по формуле $x_T=\sup_{t \in T} x \cap t$ в силу предположения $x_T \in T$. Очевидно, что $0\leq x_T \leq x$; следовательно, записывая $x = x_T+u$, имеем $u\geq 0$. Кроме того, для любого $t_1 \in T$, поскольку $x_T +|t_1| \in T$, имеем

$$x_T \geq x \cap (x_T +|t_1|) = (x_T + u) \cap (x_T + |t_1|) = x_T + (u \cap |t_1|).$$

откуда $u \cap |t_1|= 0$ для всех $t_1 \in T$, а потому $u \in T^*$. Таким образом, каждое *положительное* $x \in G$ может быть записано в виде $t + u$ $[t \in T, u \in T^*]$. Но положительные элементы из T и T^*, будучи дизъюнктными, являются перестановочными (см. п.4, лемма 1). Следовательно, для любого $x \in G$ мы имеем

$$x = x^+ - (-x^-) = (t + u) - (t_1+u_1) = t + u - u_1 - t_1 =(t-t_1) + (u - u_1),$$

так что $T + T^* == G$. Кроме того, так как элементы $t - t_1$ из T перестановочны с элементами $u - u_1$ из T^*, T и T^* являются *нормальными делителями* в $G=T+T^*$; следовательно, T и T^* являются (замкнутыми) l-идеалами. Так как они взаимно дополнительны, то имеет место

Теорема 19. *Для любого l-идеала J полной l-группы G следующие утверждения являются эквивалентными:* а) *J обладает дополнением,* б) *$J==J^{**}$,* в) *J замкнуто. Если эти утверждения справедливы, то J^* есть дополнение J.*

Поскольку пересечение замкнутых l-идеалов замкнуто, мы получаем следующее

Следствие. *Замкнутые (т. е. обладающие дополнениями) l-идеалы любой полной l-группы образуют полную булеву алгебру.*

6.15.12. Полные *l*-группы коммутативны

Пусть G—любая полная l-группа. Для любого элемента $c \in G$ имеем $c^- \in (c^+)^*$; следовательно, разложение G на два взаимно дополнительных l-идеала $(c^+)^{**}$ и $(c^+)^*$ позволяет свести любой вопрос относительно любого $c \in G$ к двум частным случаям $c \geq 0$ $c \leq 0$. Этот принцип мы используем несколько раа для установления следующего результата.

Теорема 20. *Любая полная l-группа G коммутативна.*

Лемма 1. *Если все положительные элементы l-группы G перестановочны, то группа G коммутативна.*

Доказательство. $a+b=a^+ + a^- + b^+ + b^-$, $b+a=b^+ + b^- + a^+ + a^-$. Если положительные элементы перестановочны, то a^+ и a^- перестановочны с b^+ и b^-; следовательно,

$$a^+ + a^- + b^+ + b^- = b^+ + a^+ + a^- + b^- = b^+ + b^- + a^+ + a^-,$$

чем и завершается доказательство.

Следовательно, в теореме 20 нужно рассмотреть только случай положительных a и b. Далее, в силу сделанного выше замечания, мы можем ограничиться рассмотрением случаев $a - b \geq 0$ и $b - a \geq 0$ и случаев $-a-b+a-+b \geq 0$ и $-a-b+a+b \leq 0$. Это дает нам четыре случая, из которых случаи $a \geq b \geq 0$ и (в силу симметрии) $b+a \geq a+b$ является типичными. Мы определим t с помощью соотношения $-a+b+a = b+t$, так что $t \geq 0$. Определим далее для каждого целого числа n $b_n = -na+b+na$, $t_n = -na+t+na$; таким образом, b_n и t_n суть преобразования b и t с помощью группы внутренних автоморфизмов, порожденной автоморфизмом $x \to -a+x+a$. Отметим, что, так как $0 \leq b \leq a$ и интервал $[0, a]$ инвариантен относительно всех этих внутренних автоморфизмов, мы имеем $0 \leq b_n \leq a$ для всех n. Кроме того, мы можем доказать по индукции, что $b_{n+1} = b+t+t_1+\ldots+t_n$, ибо

$$-a+(b+t+t_1+\ldots+t_{n-1})+a=$$

$$=-a+b+a+\sum_{i=0}^{n-1}(-a+t_i+a)=(b+t)+t_1+t_2+\ldots+t_n.$$

Предположим теперь, что $t_1 = -a+t+a \geqslant t$. Тогда для любого целого n $t_{n+1}-t_n=-na+(t_1-t)+na\geqslant 0$, ибо внутренние автоморфизмы сохраняют порядок. Следовательно, $0\leqslant t\leqslant t_1\leqslant t_2\leqslant t_3\leqslant\ldots$, и для любого целого положительного n мы имеем $nt\leqslant t+t_1+t_2+\ldots+t_{n-1}\leqslant -b+b_{n+1}\leqslant a$.

Следовательно, в силу теоремы 17, $t\leq 0$; но по предположению $t\geq 0$; следовательно, $t = 0$.

Аналогично, рассмотрим случай $t_1\leq t$. Мы имеем тогда, как и выше, $0\leqslant t\leqslant t_{-1}\leqslant t_{-2}\leqslant t_{-3}\leqslant\ldots$ Кроме того,

$$b=a+(b+t)-a=a+b-a+a+t-a=a+b-a+t_{-1},$$

а потому $a+b-a=b-t_{-1}$. Индукцией по n мы можем показать, что

$$b_{-n-1}=a+(b-t_{-1}-t_{-2}-\ldots-t_{-n})-a=$$
$$=b-t_{-1}-t_{-2}-\ldots-t_{-n-1}.$$

Следовательно, для любого целого положительного числа n

$$nt\leqslant t_{-n}+t_{-n-1}+\ldots+t_{-2}+t_{-1}=$$
$$=-a-b_{-n-1}+a+b=-b_n+b\leqslant b.$$

Отсюда, если $t_t\leq t$, то $t\leq 0$, а потому, как и выше, $t = 0$.

Но путем использования компонент мы можем свести общий случай к двум случаям $t_1\geq t$ и $t_1\leq t$, проектируя на $(t_1 - t)^{+}$* и $(t_1 - t)^{+}$**. Следовательно, в любом случае $t = 0$, т. е. $a+b= b+a$, чем и завершается доказательство.

Следствие. *Любая архимедова направленная группа коммутативна.*

6.15.13. Условие обрыва цепей для элементов

Мы рассмотрим теперь l-грушш, элементы которых удовлетворяют следующему условию обрыва цепей: всякое непустое множество положительных элементов содержит в себе минимальный член. В такой l-группе G всякий элемент, покрывающий 0, будет называться *простым.*

Легко доказать, что G есть полная l-группа, а потому коммутативна. Однако мы предпочитаем дать независимое исследование основных положений, предполагая только алгебраические формулы из п. 4.

Лемма 1. *Любые два простых элемента перестановочны между собой.*

Это вытекает из леммы 1 п. 4. Отсюда следует, что простые элементы порождают абелеву подгруппу, состоящую из всех элементов, представимых в виде суммы $n_1p_1+\ldots+ n_sp_s$ кратных конечного числа различных простых с положительными или отрицательными целыми числами n_i в качестве коэффициентов.

Пусть теперь задано $a > 0$; рассмотрим все те разности

$a - \sum n_i p_i$ которые являются положительными. В силу условия обрыва цепей одна из них должна быть минимальной, а потому не может содержать никакого простого q ($a - (\sum n_i p_i + q)$ было бы меньше, чем эта разность). Далее, в силу условия вбрыва цепей каждый положительный, отличный от 0, элемент b содержит простой элемент, а именно некоторый минимальный элемент x такой, что $0 < x \leq b$. Следовательно, наша минимальная разность должна совпадать с 0, так что $a = \sum n_i p_i$.

Но каждый элемент может быть представлен как разность положительных элементов: для всех c $c = c^+ - (-c)^+$, следовательно,

Лемма 2. *Любой отличный от* 0 *элемент может быть представлен как сумма целых кратных конечного числа различных простых элементов:* $a = n_1 p_1 + \ldots + n_s p_s$.

Сопоставляя леммы 1 и 2, мы заключаем, что группа G абелева. Далее, если все n_i положительны, то, очевидно, элемент a положителен. Обратно, если $a = m_1 p_1 + \ldots + m_r p_r - n_1 q_1 - \ldots - n_s q_s$, где m_i и n_i положительны, то $m_1 p_1 + \ldots + m_r p_r \geqslant n_s q_s + \ldots + n_1 q_1$; но в силу теоремы 6, поскольку простые элементы дизъюнктны, отсюда следует, что $n_s q_s + \ldots \ldots + n_1 q_1 = 0$. Следовательно,

Лемма 3. *В лемме* 2 *элемент а положителен тогда и только тогда, если каждое* n_i *положительно.*

Отсюда вытекает, что элемент a является нулем (положителен и отрицателен) тогда и только тогда, если каждое n_i положительно и отрицательно, что абсурдно. Вследствие этогопредставление, даваемое леммой 2, единственно; действительно, если бы элемент a имел два различных представления, то их разность представляла бы собой 0, в силу чего все коэффициенты были бы нулями. Нами доказана

Теорема 21. *Пусть G—любая l-группа, удовлетворяющая условию обрыва цепей. Тогда группа G коммутативна и каждый ненулевой элемент из G может быть единственным способом представлен в виде суммы целых кратных различных простых элементов. Такая сумма положительна тогда и только тогда, если все коэффициенты положительны.*

6.15.14. Строение неархимедовых коммутативных *l* -групп

Мы можем получить анализ строения любой коммутативной *l* -группы, (дистрибутивная) структура *l* -идеалов которой имеет конечную длину. В самом деле, в силу теоремы 13 любая коммутативная *l*-группа, в которой l-идеал 0 неразложим в пересечение, является просто упорядоченной. Из этого вытекает

Лемма 1. *Структура l-идеалов коммутативной l-группы, в которой l-идеал* 0 *неразложим в пересечение, является цепью.*

Но теперь, если *J* любой l-идеал в l-группе *A,* то l-идеалы из *A,* содержащие *J*, образуют структуру, изоморфную структуре всех l-идеалов (отношений конгруентности) в *A/J;* в самом деле, эта «вторая теорема об изоморфизмах» является принципом универсальной алгебры. Комбинируя это замечание с леммой 1, получаем следующий результат

Лемма 2. *В дистрибутивной структуре всех l-идеалов коммутативной l-группы элементы, содержащие любой неразложимый в пересечение элемент, образуют цепь.*

Следовательно, множества неразложимых в пересечение l-идеалов, содержащиеся в различных максимальных l-идеалах l-группы, не имеют общих элементов и не связаны отношениями включения. Отсюда следует, что либо имеется только один максимальный собственный l-идеал, либо частично упорядоченное множество неразложимых в пересечение l-идеалов является кардинальной суммой двух или большего числа компонент. В последнем случае в силу теоремы 5 р. 6.10 и формулы (6) р. 6.2 структура l-идеалов является прямым произведением. Следовательно, она содержит два собственных взаимно дополнительных l-идеала, и в силу первого абзаца п.11 l-группа является прямым произведением. Мы пришли, таким образом, к следующему результату.

Теорема 22. *Пусть G—любая коммутативная l-группа, структура l-идеалов которой имеет конечную длину. Тогда либо G есть прямое произведение (т. е. кардинальное произведение), либо G содержит максимальный l-идеал M, содержащий все остальные собственные l-идеалы.*

6.16 Векторные структуры

6.16.1. Введение

Абстрактные функциональные пространства пролили столько света на теорию функций, что их изучение признается как важная отрасль математики. Обычно, как в случаях гильбертова пространства и банаховых пространств, это изучение основывалось на анализе свойств расстояния в векторных пространствах.

Анализ свойств упорядоченности в функциональных пространствах в такой же мере заслуживает внимания и может оказаться столь же плодотворным. Этому анализу посвящен настоящий раздел; кроме того, здесь применимы все результаты относительно *l*-групп и коммутативных *l*-групп, полученные в р. 6.15. Начинаем с определения частично упорядоченных векторных пространств и векторных структур.

Определение. *Частично упорядоченное векторное пространство есть векторное пространство V с вещественными скалярами, которое является частично упорядоченной группой по сложению и в котором для любого положительного скаляра к соответствие* $x \to \lambda x$ *есть автоморфизм. Если частично упорядоченное векторное пространство V есть структура, то оно называется векторной структурой; если V условно полно или σ-полно, то оно называется соответственно полной или σ -полной векторной структурой*.

Таким образом, векторные структуры, полные векторные структуры и σ -полные векторные структуры суть соответственно *l*-группы, полные *l*-группы и σ-полные *l*-группы, в которых возможно умножение на скаляры, подчиненное обычным правилам векторной алгебры, а также правилу, что соответствие $x \to \lambda x$ сохраняет упорядоченность при $\lambda > 0$ и заменяет ее на обратную при $\lambda < 0$. Мы будем предполагать ниже знакомство с содержанием р. 6.15, в которой такие *l*-группы были рассмотрены. Напоминаем также, что любое векторное пространство является *коммутативной* группой по сложению.

Можно рассматривать также векторные структуры над упорядоченными полями, отличными от поля вещественных чисел.

Лемма 1. *В векторной структуре всякий l-идеал является подпространством и, следовательно, модулем конгруентности для всех алгебраических операций.*

Доказательство. Пусть J —произвольный *l*-идеал. Если $a \in J$, то $a^+ \in J$ и $a^- \in J$. Следовательно, для любого целого положительного числа n $na^+ \in J$ и $-na^- \in J$. Следовательно, для любого положительного λ, а потому и для любого λ $\lambda a^+ \in J$ и $\lambda a^- \in J$. Поэтому $\lambda a = \lambda (a^+ + a^-) = \lambda a^+ + \lambda a^- \in J.$

В [LT] /-идеалы назывались «нормальными подпространствами» .

6.16.2. Примеры

Простейшей векторной структурой и прототипом для всех векторных структур является система вещественных чисел R. Это есть одномерное векторное пространство, имеющее положительные числа и нуль в качестве своих неотрицательных элементов. На основании леммы 2 р. 6.15.6, в сочетании с теоремой 15 р. 6.15, можно показать, что R является *единственной* векторной структурой без собственных l-идеалов (т. е. единственной «простой» векторной структурой в смысле универсальной алгебры). Далее, имеет место

Лемма 2. *Любое прямое произведение векторных структур является векторной структурой.*

Можно показать, что если исходные векторные структуры являются полными или σ-полными, то таковым же будет их прямое произведение. Отсюда следует, что пространство R^n всех векторов $[x_1, ..., x_n]$ с n вещественными компонентами, пространство R^d всех бесконечных последовательностей $x_l, x_2, x_3, ...$ вещественных чисел и пространство R^c всех вещественных функций $f(x)$, определенных на интервале $0 \leq x \leq 1$, являются полными векторными структурами. Эти обозначения ясны, если принимать систему обозначений со степенями р. 6.2 и полагать, что d и с обозначают, соответственно, счетную бесконечность и степень континуума.

Ясно также, что любое подпространство векторной структуры, которое, кроме того, является подструктурой, есть векторная структура. Это значит, что если подмножество содержит вместе с любыми f и g также $f+g$, $f \bigcup g$, $f \bigcap g$ и каждое λf, то оно является векторной структурой относительно тех же самых операций.

Эти предположения выполняются для следующих множеств функций: подмножество B в R^c, состоящее из всех ограниченных функций; цодмножество (b) в R^d, состоящее из всех ограниченных последовательностей; подмножество C в R^c, состоящее из всех непрерывных функций; множество (c) в R^d, состоящее из всех сходящихся последовательностей; более обще, они выполняются для непрерывных функций $C(X)$ на любом топологическом пространстве X. Они выполняются также для множества $D(X)$ функций на любом таком X, допускающих лишь конечное число точек разрыва, для множества M измеримых (по Лебегу) функций на вещественной прямой, для подмножества M^p функций с суммируемыми p-ми степенями, для множества (l^p) в R^d всех последовательностей, для которых сумма $\sum_{k=1}^{\infty} |x_k|^p$ конечна.

Наконец, пусть Г—любая (бесконечная) группа взаимно однозначных отображений пространства X в себя; назовем функцию «почти периодической» (в смысле Бохнера), если любая бесконечная последовательность ее трансформаций при помощи Г содержит равномерно сходящуюся подпоследовательность. Тогда вещественные функции на X, почти периодические относительно Г, удовлетворяют нашим предположениям.

Если $h = f \bigcup g$ и $h(aT_n)$ любая последовательность трансформаций функции h, то мы можем выбрать подпоследовательность и подподпоследовательность трансформаций, для которых как f, так и g равномерно сходятся, вследствие чего то же самое имеет место и для h.

Большая часть из этих примеров рассматривается как метрические или топологические векторные пространства у С. Банаха, систему обозначений которого мы принимали, когда это представлялось удобным.

Мы можем определить все векторные структуры конечной размерности следующим образом. В силу теоремы 22 р. 6.15 любая такая векторная структура V есть либо прямое произведение векторных структур меньшей размерности, либо она содержит единственный максимальный l-идеал M. Это M будет само векторной структурой меньшей размерности и может быть исследовано по индукции; кроме того, в силу сказанного в первом абзаце этого параграфа фактор-алгебра V/M есть система вещественных чисел R. Легко теперь показать, что $V = R \circ M$. Таким образом, имеет место

Теорема 1. *Всякая конечномерная векторная структура может быть построена из* R *при помощи прямого и лексикографического произведения.*

Следствие. *Всякая архимедова конечномерная векторная структура изоморфна* R^n *при некотором* n.

Бесконечномерные архимедовы векторные структуры не обязаны даже быть изоморфными подалгебре из R^κ при любом кардинальном числе к. Как было замечено Накаямой, векторная структура V может быть тогда и только тогда представлена как такая подалгебра, если пересечение ее максимальных l-идеалов есть 0. Капланский заметил, что вещественные функции на $0 \leq x \leq 1$, непрерывные за исключением конечного числа простых полюсов второго порядка, $\alpha_1/(x - a_1)^2$, образуют архимедову векторную структуру, пересечение максимальных l-идеалов которой состоит из всех непрерывных функций, а потому отлично от 0.

6.16.3. Полнота

Другие векторные структуры могут быть построены путем образования фактор-пространств (т. е. гомоморфных образов) V/J известных векторных структур V по модулю l-идеалов J.

Теория функций изобилует примерами l-идеалов векторных структур. Например, множество функций, равных всюду нули, за исключением конечного множества, является l-идеалом в R^d , R_c и в их векторных подструктурах. То же можно сказать в отношении множества B ограниченных функций. Кроме того, если M и M^p определены в соответствии с п. 2, то каждое M^p является l-идеалом в M; точно так же каждое (l^P) является l-идеалом в R^d.

Аналогично, множество N функций, равных нулю вне множества меры нуль по Лебегу (так называемых «нуль-функций»), есть l-идеал в R^c и в векторных подструктурах R^c. Оно является даже σ-идеалом в том смысле, что оно содержит вместе с любым счетным множеством f_i наименьшую верхнюю грань $\vee f_i$.

Кроме того, фактор-алгебры («векторные фактор-структуры») M/N и M^p/N определяют обычные «пространства» S измеримых функций и L^p как векторные структуры. Другой векторной структурой является пространство $(M \cap B)/(N \cap B)$ ограниченных измеримых функций по модулю нуль-функций.

Хотя L^2 и (l^2) эквивалентны как метрические линейные пространства (гильбертову пространству), они не изоморфны как векторные структуры. Действительно, в (l^2) существуют положительные элементы, положительные подэлементы которых просто упорядочены; L^2 не содержит таких элементов.

Мы отметили уже, что R^n, R^d и R^c являются полными векторными структурами. Кроме того, поскольку l-идеалы выпуклы, имеет место

Теорема 2. *Любой l-идеал полной (σ-полной) векторной структуры сам является полным (соответственно, σ-полным).*

Отсюда следует, что пространства B, (b), (1^P) и N суть полные векторные структуры.

Рассмотрим теперь пространство S измеримых функций $f(x)$ по модулю нуль-функций. Пусть $X(f, a)$ обозначает множество, на котором $f(x) \leq a$. Тогда $X(f, a)$ является сохраняющим упорядоченность отображением бесконечного интервала $R : (-\infty, +\infty)$ в полную булеву алгебру $\overline{M}$ теоремы 15 р. 6.11. Кроме того, если $X(f, a)=X(g, a)$ для всех a, то множество, на котором $|f(x)-g(x)|>1/n$ является нуль-множеством для всех n; следовательно $f(x) - g(x)$ есть нуль-функция и $f = g$. Далее, для любого $f \in S$ $\inf_a X(f, a) = 0$ и $\sup_a X(f, a) = I$. Обратно, любое $X(a)$, имеющее только что описанные свойства, есть $X(f, a)$, и $f(x)$ может быть построено с помощью аппроксимации ступенчатыми функциями, принимающими счетное множество различных значений. Так как, наконец, $f \geq g$ эквивалентно тому, что $X(f, a) \leq X(g, a)$ для всех a, имеет место

Лемма 3. *Пространство S структурно-изоморфно множеству отображений R в* $\overline{M}$ *, которые изотопны и удовлетворяют условиям* $\inf_a X(a) = 0$ *и* $\sup_a X(a)=I$.

Но это есть выпуклое подмножество структуры $\overline{M}^{\,R}$, в системе обозначений р. 6.2.7, а последняя является полной, поскольку $\overline{M}$ полно. Следовательно, имеет место

Теорема 3. *Пространство S измеримых функций по модулю нуль-функций является полной векторной структурой.*

Отсюда следует, что векторные структуры L^p и *SB* полны, ибо они являются *l*-идеалами (выпуклыми подмножествами) пространства *S*. С другой стороны, пространство *C* даже не σ-полно.

6.16.4. Топология упорядоченности и звездная топология

Для понимания этого раздела необходимо помнить из р. 6.5.8 — 9, определения сходимости по упорядоченности и звездной сходимости для последовательностей в общих σ-структурах. Мы рассмотрим теперь специальные свойства этих топологий, имеющие место для векторных структур.

Мы показали уже в р. 6.15.10, что алгебраические операции $x+y$, $x\cap y$, $x\cup y$ непрерывны в топологии упорядоченности любой σ-полной *l*-группы. В действительности доказательства можно распространить на произвольные *l*-группы. Таким образом, в частности, в любой *l*-группе

$$\text{если } f_n\downarrow 0 \text{ и } g_n\downarrow 0, \text{ то } (f_n+g_n)\downarrow 0. \tag{1}$$

Мы докажем теперь непрерывность умножения на скаляр.

Лемма 4. *Пусть V—любая а-полная векторная структура. Тогда*

$$\text{если } \lambda_n\to 0 \text{ и } f \text{ фиксировано, то } \lambda_n f\to 0, \tag{2}$$

$$\text{если } \lambda_n\to\lambda \text{ и } f_n\to f, \text{ то } \lambda_n f_n\to\lambda f. \tag{3}$$

Доказательство свойства (2). Покажем, что $|\lambda_n f| = |\lambda_n|\cdot|f|\to 0$. Но в этом случае мы можем показать, что $\operatorname{Lim\,sup}|2\lambda_n f| = 2\operatorname{Lim\,sup}|\lambda_n f| = \operatorname{Lim\,sup}|\lambda_n f| \geqslant 0$, откуда $\operatorname{Lim\,sup}|\lambda_n f|=0$. Этим показано, что $|\lambda_n f|\to 0$, **откуда** $\lambda_n f\to 0$.

Доказательство свойства (3). $|\lambda_n f_n-\lambda f|\leqslant(\sup|\lambda_n|)\times$ $\times|f_n-f|+|\lambda_n-\lambda|\cdot|f|$, и правая часть неравенства сходится по упорядоченности к нулю в силу (2).

В соединении с предыдущими результатами получается

Теорема 4. *Относительно своей топологии упорядоченности любая σ-полная векторная структура является топологическим линейным пространством и топологической, структурой.*

В любом случае функции, непрерывные относительно топологии сходимости, являются непрерывными в связанной с ней звездной топологии (нужно лишь выбрать подпоследовательности из произвольных последовательностей). Это дает

Следствие. *Относительно звездной топологии любая σ-полная векторная структура является топологическим линейным пространством и топологической структурой.*

6.16.5. Относительная равномерная сходимость

Следуя идее Мура, будем говорить, что последовательность { f_n } элементов векторной структуры сходится, «относительно равномерно» к элементу f, если для некоторого u и $\lambda_n\downarrow 0$ $|f_n — f| \leq \lambda_n u$. Прежде всего установим связь между этой относительно равномерной топологией и топологией упорядоченности.

Лемма 5. *В любой σ-полной векторной структуре последовательность* $|f_n|$ *сходится по упорядоченности к* f *тогда и только тогда, если* $|f_n — f| \leq w_n$ *для некоторой последовательности* $w_n\downarrow 0$.

Доказательство. Так как $x \to x — f$ есть структурный автоморфизм, сохраняющий неизменными модули разностей, мы можем предполагать, что $f= 0$. В этом случае $w_n = \bigvee_{k>n} |f_k|$ представляет собой требуемую последовательность. Обратное тривиально, поскольку $\mathrm{Lim\,sup}\,|f_n| \leqslant \mathrm{Lim}\, w_n = 0$ и двойственно.

Теорема 5. *В любой σ-полной векторной структуре относительно равномерная сходимость влечет сходимость по упорядоченности. Обратное имеет место, если из* $u_n\downarrow 0$ *следует, что некоторая последовательность* $\{ku_{n(k)}\}$ *ограничена.*

Доказательство. Если $|f_n — f| \leq \lambda_n u$ и $\lambda_n\downarrow 0$, то $\{f_n\}$ сходится по упорядоченности к f в силу (2) и леммы 5. Обратно, если $|f_n — f| \leq u_n$, где $u_n\downarrow 0$, и если u есть верхняя грань для $ku_{n(k)}$, то $|f_n — f| \leq u /k$ для всех $n>n(к)$.

Приведенное выше предположение имеет место в «регулярном» случае Л. В. Канторовича. Оно имеет место для пространства S, так как $|\, u_n\,(x)\,| < 1/k^2$, за исключением множества меры, не большей, чем $1/k^2$, для всех $n > N(k)$, если только $N\,(k)$ достаточно велико. Аналогично, оно имеет место для пространств L^P. Оно не имеет места в пространствах B и (b).

6.16.6 Аддитивные отображения векторных структур в векторные структуры

Далее мы рассмотрим отображения $T: f \to fT$ векторной структуры F в полную векторную структуру X. Мы ограничимся рассмотрением «аддитивных» операций — отображений, которые удовлетворяют тождеству $(f + g)T = fT + gT$.

Известно, что аддитивные операции, отображающие любое векторное пространство в любое другое, сами образуют векторное пространство, если λT и $T+U$ определить посредством тождеств $f(\lambda T) = \lambda\ (fT)$ и $f(T + U)=fT+fU$. В данном случае это новое векторное пространство может быть частично упорядочено следующим образом.

Определение. *Аддитивная операция T неотрицательна тогда и только тогда, если она изотопна в том смысле, что*

$$f \geq g \quad \text{в } F \quad \text{влечет} \quad fT \geq gT \quad \text{в } X. \tag{4}$$

Это эквивалентно требованию, что $f - g \geq 0$ влечет $(f-g)T \geq 0$, —словами, что T переводит положительные элементы в положительные элементы (или элемент 0, который мы называем положительным).

Теорема 6. *В силу предшествующего определения аддитивные операции, отображающие F в X, образуют частично упорядоченное векторное пространство.*

Доказательство того, что они образуют частично упорядоченную группу, является следствием теоремы 1 р. 6.15: условие «г» вытекает из коммутативности; условия «а» и «б» тривиальны; мы докажем теперь условие «в». Предположим, что $T+U = 0$, $T \geq 0$ и $U \geq 0$. Тогда $g \geq 0$ влечет $gT \geq 0$, а потому $gU = -gT \leq 0$; но $gU \geq 0$ в силу предположения, что $U \geq 0$; следовательно, $gU=0$ для всех $g \geq 0$. Но это влечет $fU = (f^+ - (-f)^+)U = f^+U - (-f)^+U = 0$, т. е. $U = 0$, чем и завершается доказательство условия «г». Теорема 6 следует теперь из того, что если $\lambda \geq 0$ и $U \geq 0$, то $\lambda U \geq 0$.

С другой стороны, эти операции не обязательно образуют векторную структуру. Чтобы получить векторную структуру, мы должны ограничиться рассмотрением того, что может быть названо «ограниченными» аддитивными операциями.

6.16.7. Ограниченные аддитивные операции

Если аддитивная операция T, отображающая F в X, лежит в произвольной векторной структуре, то множество $\{T, 0\}$ должно быть ограниченным, т. е. должны существовать такие аддитивные, операции U и V, что $V \leq 0$, $T \leq U$.

Определение. *Аддитивная операция T будет называться ограниченной тогда и только тогда, если множество* {*T*, 0} *ограниченно сверху и снизу.*

Лемма 1. *Если T ограниченно, то оно переводит ограниченные множества (т. е. множества H, удовлетворяющие неравенству a≤H≤b) в F в ограниченные множества из X.*

Доказательство. Если *a≤h≤b и V≤T*, 0≤*U*, то

$$hT = aT + (h-a)\,T \leqslant aT + (h-a)\,U \leqslant aT + (b-a)\,U$$

и двойственно, $hT \geqslant bT + (b-a)\,U$. Следовательно, множество элементов $hT\ [h \in H]$ ограниченно.

Лемма 2. *Если T переводит ограниченные множества в ограниченные множества, то* $T^{+}=T\cup 0$ *существует. В действительности,* $fT^{+} = \sup_{0 \leqslant x \leqslant f} xT$, *если* $f \geq 0$, и $fT^{+} = f^{+}T^{+} - (f^{-})\,T^{+}$ *в общем случае.*

Лемма 3. *Если в l-группе* $x \leqslant g + h\ \ [x \geqslant 0,\ g \geqslant 0,\ h \geqslant 0]$, *то* $x = s+t$, *где* $0 \leq s \leq g$ *и* $0 \leq t \leq h$.

Доказательство леммы 3. Пусть $s = x \cap g$; тогда

$$t = -x \cap g + x = -g + x \cup g \leqslant -g + (g+h) = h.$$

Очевидно, что $0 \leqslant s \leqslant g,\ 0 \leqslant t \leqslant h,\ x = s + t.$

Доказательство леммы 2. Прежде всего, определенное в лемме T^{+} аддитивно. В самом деле, в силу леммы 3 sup xT для $0 \leq x \leq f + g$ есть $\sup\,(y+z)\,T = \sup\,(yT + zT)$, для $0 \leq y \leq f$, $0 \leq z \leq g$. Но это есть (sup yT) + (sup zT), ибо, с одной стороны, последнее выражение является верхней гранью для $yT+zT$, а, с другой стороны, любая верхняя грань для всех $yT+zT$ содержит yT + sup zT для каждого фиксированного y и, следовательно, sup yT + sup zT.

Остается показать, что $T^{+}=T\cup 0$. Но из $f \geq 0$ следует, что

$$f\,(T^{+} - T) \geqslant fT - fT = 0 \quad \text{и} \quad f\,(T^{+} - 0) \geqslant 0T - f0 = 0;$$

следовательно, T^{+} является верхней гранью для 0 и T. Обратно, если U любая (аддитивная) верхняя грань для 0 и T, то для всех x между 0 и f

$$fU = xU + (f-x)\,U \geqslant xT + (f-x)\,0 = xT,$$

откуда $f\,(U - T^{+}) = fU - fT^{+} \geqslant 0$ для всех $f \geq 0$, и $U \geq T^{+}$.

Теорема 7. *Ограниченные аддитивные операции, отображающие F в X, образуют векторную структуру. Она вкладывается в частично упорядоченное векторное пространство теоремы* 6 *и содержит каждую векторную структуру, вложимую в это пространство-*

Доказательство. Если $V \leqslant 0,\ T \leqslant U$ и $V^{*} \leqslant 0,\ T^{*} \leqslant U^{*}$, то $V + V^{*} \leqslant 0,\quad T + T^{*} \leqslant U + U^{*}$ и $V \leqslant 0,\ \lambda T \leqslant \lambda U$

(или обратное неравенство) для всех λ; следовательно, охраниченные аддитивные операции образуют подпространство векторного пространства теоремы 6. В силу теоремы 2 р. 6.15 это подпространство есть векторная структура. Но в силу первого замечания настоящего раздела каждая векторная структура, вложимая в частично упорядоченное векторное пространство теоремы 6, содержится в этом подпространстве.

Теорема 8. *Каждое из нижеследующих условии эквивалентно условию ограниченности:* а) *T переводит ограниченные множества в ограниченные множества,* б) *T есть разность неотрицательных операций.*

Доказательство. В силу леммы 1 из ограниченности следует «а»; в силу доказанной выше теоремы 7 и следствия 2 теоремы 4 р. 6.15 из «а» следует «б». Наконец, если $T \geq 0$ и $U \geq 0$, то $-T-U \leq T-U \leq T+U;$ следовательно, из «б» вытекает ограниченность.

6.16.8. Функционалы и сопряженные пространства

Ограничимся случаем, когда X есть система вещественных чисел, напомнив, что анализ вещественных функций или «функционалов» на абстрактных пространствах привлекал внимание математиков еще со времени первоначальной работы Вольтерра и Фреше

Мы определим *сопряженное пространство* F^* векторной структуры F как векторную структуру (см. теорему 7) всех ограниченных аддитивных функционалов на F. В теореме 11 покажем, что это определение сводится к понятию «сопряженного пространства», как оно трактуется у Банаха. Это определение включает в себя как другой частный случай понятие «двойственного пространства», введенное Кете и Теплицем.

Докажем соотношение $F \leq (F^*)^*$. В самом деле, каждое $f \in F$ определяет аддитивный функционал Φ на F^*: $\Phi(T)=fT$. Кроме того, каждое такое Φ ограниченно; так как $f=f^+-(-f^-)$, нам нужно доказать это только для случая $f \geq 0$. Но если $f \geq 0$ и $V \leq T \leq U$, то $fV \leq fT \leq fU$; но отсюда следует, что $\Phi(T)$ на любом ограниченном множестве (замкнутом интервале) ограниченны.

6.16.9. Банаховы структуры

Каждый из известных примеров банахова пространства (или «пространства (В)» в смысле С. Банаха является также векторной структурой в естественном смысле. Кроме того, в любом случае упорядоченность $f \leq g$ и норма $\| f\|$ связаны следующим условием

$$|f| \leq |g| \text{ влечет } \|f\| \leq \|g\|. \tag{5}$$

Векторная структура, являющаяся банаховым пространством и удовлетворяющая условию (5), будет называться *банаховой структурой.* В банаховой структуре

$$||f|| = ||\,|f|\,|| \quad \text{для всех } f. \qquad (6)$$

Мы покажем теперь, что в банаховой структуре L метрическая сходимость эквивалентна относительно равномерной звездной сходимости и что непрерывные аддитивные функционалы на L суть просто ограниченные аддитивные функционалы. Следовательно, во всех известных случаях *метрические понятия могут быть заменены понятиями, связанными с простой упорядоченностью*.

Лемма 1. *Никакая векторная структура L, содержащая такую последовательность* $\{f_n\}$, *что все последовательности* $\{\lambda_n f_n\}$ *ограниченны по упорядоченности, не может быть превращена в банахову структуру.*

Доказательство. Предположим, что L можно превратить в банахову структуру; положим $\lambda_n=n/||\,f_n\,||$. Если бы u было верхней гранью для всех $\lambda_n f_n$, то $||u||$ превышало бы каждое n, поскольку $||\lambda_n f_n\,|| = n$.

Лемма 2. *В любой банаховой структуре операции* $f+g$, $f \cap g$ *и* $f \cup g$ *метрически равномерно непрерывны.*

Доказательство. В силу определения и формулы (20) р. 6.16 имеем

$$|(f \circ g)-(f^* \circ g)| \leqslant |f-f^*|,$$

где $\circ$ обозначает любую из операций $+$, $\cap$ и $\cup$. В силу (5) — (6) отсюда вытекает соответствующее метрическое неравенство; следовательно, все три операции равномерно непрерывны с модулем непрерывности единица в обычной метрике $||\,f — g||$.

Следствие. *Если* $\sum_{n=1}^{\infty} a_n = a$, $\sum_{n=1}^{\infty} b_n = b$ и $\sum_{n=1}^{\infty} x_n = x$ *в метрическом смысле и* $a_n \leq x_n \leq b_n$ *для всех n, то* $a \leq x \leq b$

Лемма 3. *Любая сепарабельная банахова структура L обладает слабой единицей e.*

Доказательство. Пусть $\{f_n\}$—любое счетное, метрически полное подмножество в L. Определим $e=\sum_{n=1}^{\infty}\frac{1}{2^n\,||\,f\,||}|\,f_n\,|$. Для любого $g>0$ в L мы имеем $||\,g-f_n\,|| < ||\,g\,||/10$ для некоторого n.

Кроме того, из $e \cap g=0$ следовало бы, что $g \cap |f_n|/2^n\,||\,f_n\,|| \leqslant$ $\leqslant g \cap e=0$, а потому $g \cap |\,f_n\,| = 0$. Но $g \cap |\,f_n\,| = g \cap [(f_n \cup 0)-$ $-(f_n \cap 0)]$, и подстановка в это выражение g вместо f_n смещает его в силу леммы 2 не более чем на $||g||/5$. Поскольку $g \cap [(g \cup 0)-(g \cap 0)]=g$, отсюда следует, что $||\,(g \cap |\,f_n\,|)-g\,|| \leqslant$ $\leqslant ||\,g\,||/5$, а потому $g\cap |\,f_n| > 0$. Следовательно, $g \cap e \geqslant$

$\geqslant (g \cap |f_n|) / 2^n \| f_n \| > 0$ и e является слабой единицей. Этот результат принадлежит Фрейденталю.

Теорема 9. *В любой банаховой структуре метрическая сходимость эквивалентна относительно равномерной звездной сходимости.*

Доказательство. В силу однородности обеих топологий достаточно рассмотреть лишь сходимость к 0. Но если $|f_n| \leq \lambda_n u$ и $\lambda_n \downarrow 0$, то $\| f_n \| \leqslant \| \lambda_n u \| = |\lambda_n| \cdot \|u\| \downarrow 0$.

Таким образом, относительно равномерная звездная сходимость влечет метрическую звездную сходимость, а потому и метрическую сходимость. Обратно, если $\| f_n \| \to 0$, то мы можем так выбрать $n(k)$, что $k^3 \| f_{n(k)} \| \to 0$, и построить затем $v = \sum_{k=1}^{\infty} k f_{n(k)}$, для которого $|f_{n(k)}| \leqslant \frac{1}{k} |v|$.

Поскольку $a \leq x \leq b$ влечет $\|x\| \leq \|a\| + \|b - a\|$, из ограниченности по упорядоченности подмножества банаховой структуры вытекает метрическая ограниченность. Однако, обратное обычно не имеет места. Так, в L^P и (l^p) не существует верхней грани для элементов нормы единица — хотя в пространствах B и (b) таковая имеется. Но во всех случаях мы имеем следующий результат.

Теорема 10. *Для аддитивных функционалов T на банаховой структуре метрическая «ограниченность» и ограниченность по упорядоченности* п. 7 *эквивалентны.*

Таким образом, условие, что T «переводит ограниченные множества в ограниченные множества», является одним и тем же независимо от того, понимается ли оно в терминах метрической ограниченности или в терминах ограниченности по упорядоченности.

Доказательство. Если T метрически ограниченно, то для любого $a>0$ из $|x|<a$ следует $\|x\| \leq \|a\|$; следовательно, $|xT| \leq \|a\| \cdot \|T\|$ и элементы xT образуют ограниченное множество. Обратно, если T метрически неограниченно, то существует последовательность $\{x_n\}$, для которой $\| x_n \| \leq 2^{-n}$, однако, $| x_n T | \uparrow +\infty$. Следовательно, в силу следствия леммы 3 элементы $a = \sum_{n=1}^{\infty} x_n^-$ и $b = \sum_{n=1}^{\infty} x_n^+$ ограничивают множество элементов y, для которого множество элементов yT неограниченно; таким образом, функционал T неограничен по упорядоченности.

6.16.10. Равномерно монотонная норма

Будем называть норму в банаховой структуре «равномерно монотонной», если для любого заданного $\varepsilon>0$ можно найти столь малое $\delta > 0$, что если $f \geq 0$, $g \geq 0$ и $\|f\| = 1$, то из $\| f + g\| \leq \|f\| + \delta$ следует $\|g\| \leq \varepsilon$. Мы будем называть банахову структуру с равномерно монотонной нормой UMB-*структурой.*

Так, пространство L имеет равномерно монотонную норму, поскольку $\| f + g\| = \|f\| + \|g\|$. Аналогично, пространства L^p и (l^p) имеют равномерно монотонную норму, а потому они являются UMB-структурами. В действительности, любое метрически замкнутое подпространство и подструктура любой UMB-структуры само является UMB-структурой.

С другой стороны, нормы в пространствах *B, (b)* и *C* не являются равномерно монотонными.

Теорема 11. *В UMB-структуре любое метрически ограниченное множество элементов, являющееся «направленным множеством» в структурной упорядоченности, сходится метрически.*

Доказательство. Мы ограничиваемся рассмотрением элементов, следующих за фиксированным элементом *а,* а потому, поскольку $f \to f - a$ есть изометрический структурный автоморфизм, мы можем предполагать, что все элементы f_α множества неотрицательны. Далее, путем умножения на скаляр мы можем принять, что $\| f_\alpha \| < 1$ для всех α. Но в этом случае, если α выбрано так, что $\| f_\alpha \| \geq \sup \| f_\alpha \| - \delta$, мы будем иметь $\| f_\beta - f_\alpha \| \leq \varepsilon$ для всех $f_\beta > f_\alpha$, чем и доказан наш результат.

Следствие. *Любая UMB-структура является полной векторной структурой.*

Действительно, если множество имеет верхнюю грань, то верхнюю грань имеют и объединения конечных подмножеств элементов этого множества, а они образуют метрически ограниченное направленное множество.

Теорема 12. *В UMB-структуре сходимость по упорядоченности и относительно равномерная сходимость эквивалентны.*

Доказательство. В силу теоремы 5 достаточно показать, что

в любой UMB-структуре $u_n \downarrow 0$ влечет $\|u_n\| \downarrow 0$. (7)

Действительно, из (7) вытекает, что для некоторого $n(k)$ $\| u_{n(k)}\| \leq 1/k2^k$, следовательно, бесконечная линейная сумма $\sum_{k=1}^{\infty} k u_{n(k)}$ существует и является (см. лемма 3 п. 9) верхней гранью для последовательности $\{k u_{n(k)}\}$. В силу теоремы 5 отсюда вытекает эквивалентность сходимости по упорядоченности и относительно равномерной сходимости.

Мы докажем теперь свойство (7). В силу теоремы 11 $\{u_n\}$ сходитсн метрически к некоторому *а.* В силу леммы 3 п. 9 $a \cap 0 = 0$ и $a \cup u_n = u_n$

для всех *n;* следовательно, $0 \le a \le u_n$ для всех *n,* а потому $a = 0$, что и утверждалось.

Следствие. *В любой UMB-структуре звездная сходимость и метрическая сходимость эквивалентны.*

Наконец, можно заметить, что теорема 10 может быть распространена на операции T со значениями в произвольной UMB -структуре. Вторая половина доказательства не нуждается в изменении. Что касается первой половины, то заметим, что, так уау два любых разбиения интервала $0 \le x \le f$ имеют общее уплотнение, xT $[0 \le x \le f]$ образует направленное множество. Следовательно, если T метрически ограниченно, то метрически ограниченны и $xT,$ и в силу теоремы 11 они сходятся к верхней грани fT^+. Двойственным образом xT имеет нижнюю грань, а потому T ограниченно по упорядоченности.

6.16.11. Теорема о разложении

Пусть L—любая UMB-структура и $\lambda(f)$—любой ограниченный аддитивный функционал на L. Покажем, что $\lambda(f)$ разлагает L на компоненты, на которых $\lambda(f)$, соответственно, положительно, отрицательно и равно нулю. Справедлива

Лемма 1. *Пусть N^+, N^- и N^0 определены как множества элементов f таких, что из* $0<x<|f|$ *следует, соответственно,* $\lambda(f)>0$, $\lambda(f)<0$ *и* $\lambda(f)= 0$. *Тогда N^+, N^- и N^0 являются независимыми l-идеалами.*

Доказательство. Предположим, что $f \in N^+$ и $g \in N^+$. Тогда $0< x\le|f+g|$ влечет $0 < x \le|f|+|g|$, откуда $x = y+z$, где $0\le y\le|f|$, $0\le z\le|g|$ и $y>0$ или $z > 0$; следовательно, $\lambda(f)= \lambda(y)+\lambda(z)>0$. Это означает, что N^+ замкнуто относительно сложения. Далее, если $f\in N^+$ и $|h| \le |f|$, то $0<x\le|h|$ влечет $0<x\le |f|$, откуда $\lambda(f)>0$ и $h\in N^+$. Мы заключаем, что N^+ есть l-идеал. Аналогично, N^- и N^0 являются l-идеалами.

Далее, поскольку l-идеалы образуют дистрибутивную структуру, l-идеалы N^+, N^- и N^0 должны быть независимыми, если $N^+ \cap N^- = N^- \cap N^0 = N^0 \cap N^+ = 0$. Но последнее почти тривиально: если $f=0$, то $f\in N^+$ влечет $\lambda(|f|)>0$, $f \in N^-$ влечет $\lambda(|f|)< 0$ и $f \in N^0$ влечет $\lambda(|f|)= 0$; эти условия друг друга взаимно исключают.

Теорема 13. *L есть прямое произведение N^+, N^- и N^0.*

Доказательство. В силу набранного курсивом утверждения р. 6.15. 11, первый абзац, и доказанной выше леммы 1 достаточно показать, что L вляется линейной суммой подпространств (см. лемму 1 п. 1) N^+, N^- , N^0, т. е., что $L = N^+ + N^- + N^0$. Сперва нам потребуется равномерная монотонность для доказательства следующей леммы.

Лемма 2. *Функционал $\lambda(f)$ достигает своего супремума M на любом интервале* $0\le x\le f$.

Доказательство. Выберем x_i так, что $\lambda(f_i) > M — 3^{-i}$; покажем, что $g = \lim\sup \{ x_i \}$ — которое существует в силу следствия теоремы 11 — удовлетворяет условию $\lambda (g) = M$. В самом деле,

$$\lambda(x_i \bigcup x_j) = \lambda(x_i) + \lambda(x_j) - \lambda(x_i \bigcap x_j) > M - 3^{-i} - 3^{-j},$$

откуда $\lambda(x_i \bigcup x_{i+1} \bigcup \ldots \bigcup x_{i+n}) > M - (3^{-i} + \ldots + 3^{-i-n})$ превосходит $M — 2 \cdot 3^{-i}$. Переходя один раз к пределу, получаем $M \geqslant \lambda(\sup_{k \geqslant i}\{x_k\}) \geqslant M - 2\cdot 3^{-i}$. Переходя снова аналогичным образом к пределу, мы в силу непрерывности в метрической топологии, которая (теорема 12, следствие) влечет непрерывность в топологии упорядоченности, получаем $\lambda (g) = M$.

Соотношение $L = N^+ + N^- + N^0$ вытекает теперь непосредственно из леммы:

Лемма 3. *Пусть в лемме* 2 *и есть infimum элементов x между* 0 *и f таких, что* $\lambda(x) = M$; *определим v двойственным образом; пусть* $w = f — u — v$. *Тогда* $u \in N^+$, $v \in N^-$ *и* $w \in N^0$.

Доказательство. Существование u и v (и, таким образом, w) следует из полноты L. Далее, если $\lambda (x) = \lambda (y) = M$, то

$\lambda(x \bigcap y) + \lambda(x \bigcup y) = 2M$. Но $\lambda(x \bigcap y) \leqslant M$ и $\lambda(x \bigcup y) \leqslant M$,

а потому $\lambda(x \bigcap y) = \lambda(x \bigcup y) = M$. Отсюда следует в силу теоремы 11 и непрерывности, что $\lambda (u) = M$. Кроме того, $0<x\leq u$ влечет $\lambda (u — x) < \lambda (u)$, а потому $\lambda (x) = \lambda (u) — \lambda (u — x) > 0$; следовательно, $u\in N^+$. Двойственным образом $v\in N^-$. Следовательно, в силу леммы 1 $u\cap v=0$, откуда $u + v = u \bigcup v \leqslant f$ и $0 \leqslant w \leqslant f$. Наконец, $0< x \leq w$ влечет $\lambda(x) + \lambda(u) = \lambda(x + u) \leqslant \lambda(u)$, откуда $\lambda (x) \leq 0$.

Двойственным образом $0<x\leq w$ влечет $\lambda (x)\geq 0$, а потому $\lambda (x) = 0$. Следовательно, $w\in N^0$. Но равенство $f = u+v+w$ очевидно, чем 11 завершается доказательство.

6.16.12. Интегральное представление

Пусть—V любая σ-полная векторная структура со слабой единицей e. Под *разложением* e мы понимаем монотонное возрастающее семейство компонент e_λ элемента e, такое что

$$\lambda < \mu \text{ влечет } e_\lambda \leqslant e_\mu, \quad \mathrm{Lim}_{\lambda\to\infty} e_\lambda = e, \quad \mathrm{Lim}_{\lambda\to-\infty} e_\lambda = 0. \quad (8)$$

По отношению к любому такому «разложению единицы» мы можем определить *интегралы* следующим образом.

Для любого конечного M и любого подразбиения π интервала $—M\leq\lambda\leq M$ на подинтервалы $\lambda_{i-1}\leq\lambda\leq\lambda_i$ мы можем образовать $u_\pi = \sum_i \lambda_{i-1}(e_{\lambda_i} - e_{\lambda_{i-1}})$ и $v_\pi = \sum_i \lambda_i (e_{\lambda_i} - e_{\lambda_{i-1}})$; рассматриваем их, соответственно, как апроксимации интеграла *снизу* и *сверху*. Если π и π'

два любых подразбиения, то легко показать, что $u_\pi \leqslant u_{\pi\cap\pi'} \leqslant v_{\pi\cap\pi'} \leqslant v_{\pi'}$; кроме того, $|v_{\pi'} - u_\pi| \leqslant 2\varepsilon e$, где ε есть длина наибольшего подинтервала подразбиения π или π'. Следовательно, так как в силу (2) $2\varepsilon e \downarrow 0$ при $\varepsilon \downarrow 0$, u_π и v_π стремятся к одному и тому же пределу g. Этот предел мы *определим* как значение символа

$$g = \int_{-M}^{M} \lambda\, de_\lambda. \tag{9}$$

Этим определяются *ограниченные* интегралы. Мы определим

$$f = \int_{-\infty}^{\infty} \lambda\, de_\lambda = \operatorname*{Lim}_{n\to\infty} \left\{ \int_{-n}^{n} \lambda\, de_\lambda \right\} \tag{10}$$

в смысле сходимости по упорядоченности, в случае когда этот предел существует.

Мы только что видели, что каждое разложение e определяет на ограниченном интервале ограниченный интеграл g; очевидно, что $|g| \leq Me$ для некоторого конечного M. Обратно, пусть g любой элемент в V, который *ограничен* в этом смысле по отношению к e, т. е. для которого e является сильной единицей. Тогда мы можем определить e_λ как компоненту элемента e в замкнутом l-идеале, имеющем $(\lambda e - g)^+$ в качество сильной единицы. Нетрудно показать, что по отношению к этому разложению имеет место (9). Более обще, имеет место

Теорема 14. *Любой элемент f σ-полной векторной структуры V со слабой единицей e может быть представлен как интеграл* (10).

Доказательство. Для любого целого положительного числа n мы строим ограниченное *усечение a_n* элемента a посредством

$$f_n = -ne \cup (f \cap ne) = (-ne \cup f) \cap ne \text{ (см. L5).} \tag{11}$$

Тогда последовательность элементов $f_n^+ = 0 \cup (f \cap ne)$, которая ограничена сверху элементом f^+, сходится в силу σ-полноты к пределу $f^* < f^+$. Но мы можем показать, что $(f^+ - f^*) \cap e = 0$, откуда $f^+ - f^* = 0$, а потому $f_n^+ \uparrow f^+$.

Двойственным образом $f_n^- \downarrow f^-$, откуда

$$f_n = f_n^+ + f_n^- \to f^+ + f^- = f \tag{12}$$

в смысле сходимости по упорядоченности. Следовательно, имеет место (10), где $f_n = \int_{-n}^{n} \lambda\, de_\lambda$.

6.16.13. Аддитивные функции множеств и (*L*)-пространства

Рассмотрим класс *L (A)* всех ограниченных оценок *f[x]* на фиксированной булевой алгебре *A*, которые также удовлетворяют условию $f[0] = 0$. При любом изоморфном представлении *A* телом Ф множеств *L* (*A*) может быть рассмотрено как класс всех *ограниченных, аддитивных функций множеств на* Ф.

Лемма. *Если f [x] есть оценка на произвольной булевой алгебре, то вариация f[x] на любой цепи* $a = x_0 < x_1 < \ldots < x_n$ b *равна вариации f [x] на некоторой цепи* $a \le y \le b$ *длины два.*

Доказательство. Полагаем *у* равным объединению *а* и таких разностей $x_i - x_{i-1} = x_i \cap x'_{i-1}$, для которых $f[x_i] > f[x_{i-1}]$.

Отсюда следует, что условие быть ограниченным эквивалентно условию иметь ограниченную вариацию. Мы знаем уже, что оценки ограниченной вариации на любой структуре образуют векторную структуру; следовательно, *L(A)* есть векторная структура. Мы покажем теперь, что если мы примем полную вариацию в качестве нормы, то *b(A)* превратится в банахову структуру. Но в силу леммы полная вариация *f[x]* есть

$$\|f\| = \sup\{|f[x] - f[0]| + |f[I] - f[x]|\} = \sup\{f[x] - f[x']\}. \quad (14)$$

Кроме того, так как бóльшим является *f[x]*, а меньшим $f[x'] = f[I] — f[x]$, мы имеем далее

$$\|f\| = \sup_x f[x] - \inf_x f[x] = \sup\{2f[x] - f[I]\}. \quad (14')$$

Теорема 15. *Если мы определим норму в L(A) как полную вариацию, т. е. по формуле* (14), *то L(A) превратится в банахову структуру, в которой*

$$\text{если } f > 0 \text{ и } g > 0, \text{ то } \|f + g\| = \|f\| + \|g\|. \quad (15)$$

Доказательство. Для структур с относительными дополнениями условие быть ограниченным эквивалентно условию иметь ограниченную вариацию. Но мы знаем, что оценки ограниченной вариации на *любой* структуре образуют векторную структуру.

Далее, *L(A)* есть банахово пространство. Действительно, очевидно, что $\|cf\| = |c| \cdot \|f\|$. Кроме того,

$$\|f + g\| = \sup_x\{f[x] - f[x'] + g[x] - g[x']\}$$
$$\leqslant \sup_{x,y}\{f[x] - f[x'] + g[y] - g[y']\} = \|f\| + \|g\|.$$

Если $f \neq 0$, то $f[x] \neq 0$ для некоторого *x*. Следовательно, либо $f[x'] = f[x]$ и $\|f\| \geq f[I] = 2f[x] > 0$, либо $f[x'] \neq f[x]$ и $\|f\| \geq \max\{f[x]-f[x'], f[x']-f[x-]\} = |f[x]-f[x']| > 0$. Таким образом, $f \neq 0$ влечет $\|f\| > 0$ и *L (A)* есть метрическое векторное пространство. Наконец, если $\|f_m — f_n\| \to 0$

при $m,n\to\infty$, то для любого x | $f_m[x]$ — $f_n[x]$ $\to 0$; следовательно, мы можем определить $f[x] = \mathrm{Lim}_{n\to\infty} f_n[x]$. Это $f[x]$ является аддитивным, ибо все f_n аддитивны; кроме того, что $f[0] = 0$ и $\|f_n - f\| \to 0$. Если $f > 0$, то $f[x] - f[x'] \leq f[I] - f[0]$; следовательно,

$$\text{если } f>0, \text{ то } \|f\| = f[I]. \qquad (14'')$$

Отсюда непосредственно вытекает (15).

Остается доказать, что $|f| \leq |g|$ влечет $\|f\| \leq \|g\|$. Но в силу (15), записывая $h = |g| - |f| \geq 0$, достаточно доказать, что $\|f\| = \||f|\|$. Это мы теперь и покажем. В силу р. 6.6.10 $f^+[I] = \sup f[x]$ и $f^-[I] = \inf f[x]$. Кроме того, $f[x'] = f[I] - f[x]$. Следовательно, в силу (14')

$$\|f\| = \sup\{2f[x] - f[I]\} = 2f^+[I] - f[I] =$$
$$= f^+[I] - f^-[I] = |f|[I] = \||f|\|. \qquad (16)$$

Определение. *Банахова структура, удовлетворяющая условию* (15), *будет называться абстрактным (b)-пространством.*

Следствие. *Ограниченные оценки на произвольной булевой алгебре, удовлетворяющие условию* $f[0]=0$, *образуют абстрактное (L)-пространство. Любое абстрактное (b)-пространство является UMB-струптурои.*

Имеется тесная связь между абстрактными (L-пространствами и понятиями меры и вероятности. Это мы теперь и сформулируем.

Определение. *Под «распределением» или «вероятностным функционалом» на булевой алгебре A понимается положительная оценка* $p[x]$, *удовлетворяющая условиям* $p[0]=0$ *и* $p[1] = 1$. *Если A булева σ-алгебра и если*

$$p\left[\bigvee_{i=1}^{\infty} x_i\right] \leqslant \sum_{i=1}^{\infty} p[x_i],$$

то p[x] называется «σ-распределением».

Теорема 16. *Множество D(A) распределений на A состоит из положительных элементов в L(A), имеющих нормой единицу. Поэтому оно метрически замкнуто, выпукло и имеет самое большее диаметр два.*

Доказательство. В силу определения *D(A)* состоит из положительных элементов, удовлетворяющих условию $f[I] = 1$. Но если $f > 0$, то $\|f\| = f[I]$, чем и доказано первое утверждение. Далее, как множество элементов $f \geq 0$ (в силу леммы 3, п.9), так и «единичная сфера», состоящая из элементов f с единичной нормой, являются метрически замкнутыми; следовательно, их пересечение *D (A)* также метрически замкнуто. В то же время, если $p > 0,\ q > 0,\ \lambda > 0,\ \mu > 0,\ \lambda + \mu = 1,$ то $\lambda p + \mu q > 0;$ если, кроме того, $\|p\| = \|q\| = 1$, то $\|\lambda p + \mu q\| = \lambda\|p\| + \mu\|q\| = 1.$ Наконец, если $p,\ q \in D(A)$, то $\|p - q\| \leqslant \|p\| + \|-q\| = 2.$

6.16.14. σ-Распределения

Рассмотрим теперь подмножество $L_Z(A)$ σ-аддитивных оценок иа булевой σ-алгебре *A,* удовлетворяющих условию $f[0]=0$, и соответствующее подмножество $D_\sigma(A)$ σ-распределений.

Лемма. *Каждое из нижеследующих условий непрерывности эквивалентно σ-аддитивности:* а) $x_n\downarrow 0$ *влечет* $f(x_n)\to 0$ *и* б) $x_n\to x$ *в топологии упорядоченности влечет* $f[x_n]\to] [x]$.

Доказательство. Поскольку $y_n\to y$ тогда и только тогда, если существует последовательность $x_n\downarrow 0$, удовлетворяющая условию $|y_n - y| \le x_n$, где $|y_n - y|$ обозначает симметричную разность между y_n и y, условия «а» и «б» являются эквивалентными. Но $\bigvee_{i=1}^{n} x_i \to \bigvee_{i=1}^{\infty} x_i$; следовательно, если f непрерывно, то $f[\bigvee_{i=1}^{\infty} x_i] = \mathrm{Lim}[\bigvee_{i=1}^{n} x_i] \leqslant \sum_{i=1}^{\infty} f[x_i]$. Обратно, поскольку $x_n\downarrow 0$ влечет $x_1 = \bigvee_{i=1}^{\infty}(x_i - x_{i+1})$, где $(x_i - x_{i+1}) \cap (x_j - x_{j+1}) = 0$, если $i \neq j$, $f[x_1] = \sum_{i=1}^{\infty} f[x_i - x_{i+1}]$, если f σ-аддитивно. Но отсюда следует, что $f[x_n] = \sum_{i=n}^{\infty} f[x_i - x_{i+1}]$, чем и завершается доказательство.

Теорема 17. *Если A есть булева σ-алгебра, то* $L_\sigma(A)$ *есть метрически замкнутый 1-идеал в* $L(A)$.

Доказательство. Имеем $|f[x_n]| \leqslant \| f_m - f \| + f_m[x_n]$. Поэтому, если каждое f_m непрерывно и $\|f_m - f\|\to 0$, то $x_n\downarrow 0$ влечет $f[x_n]\to 0$; таким образом, $L_\sigma(A)$ метрически замкнуто. То, что $L_\sigma(A)$ есть векторное подпространство в $L(A)$, тривиально; остается доказать условие «нормальности»: что из $f \in L_\sigma(A)$ и $0<g\le|f|$ следует $g\in L_\sigma(A)$. Но при этих предположениях из $x_n\downarrow 0$ и $0\le y_n\le x_n$ следует $y_n\downarrow 0$, а потому $f[y_n]\to 0$; следовательно, $\sup_{0\leqslant y_n \leqslant x_n} | f[y_n] || \to 0$. Но в силу предположения и леммы из § 13 $g[x_n]| \leqslant |f|[x_n] \leqslant 2\sup_{0\leqslant y_n \leqslant x_n} \times |f[y_n]|$; поэтому $g[x_n]\to 0$. В силу доказанной выше леммы отсюда следует, что $g\in L_\sigma(A)$, чем и завершается доказательство.

Следствие 1. *Для любой булевой σ-алгебры A множество σ-аддитивных оценок f [x] на A, удовлетворяющих условии* $f[0] = 0$, *является абстрактным* (L)*-пространством. Следовательно, σ-аддитивные функции множеств на любом σ-теле множеств образуют абстрактное* (L)*-пространство.*

Следствие 2. σ-*Распределения на A образуют метрически замкнутое, выпуклое подмножество в L(A).*

6.16.15. Представление сепарабельных абстрактных (*L*)-пространств

Пусть *L* любое сепарабельное абстрактное (*L*)-пространство со счетным метрически плотным подмножестром элементов $f_1, f_2, f_3, \ldots$ Приведем набросок доказательства того факта, что *L* может быть вложено в пространство (L), определенное в п. 2.

Как и в лемме 3 п. 9, определим слабую единицу посредством

$$e=\sum_{n=1}^{\infty}\frac{1}{2^n\|f_n\|}|f_n|>0. \tag{17}$$

Пусть *A*—полная булева алгебра всех замкнутых *l*-идеалов в *L*. Если *J*—любой такой замкнутый *l*-идеал и если *J'*—его дополнение, то мы можем в силу теоремы 11 и метрической полноты образовать объединение e_J всех элементов $x\le e$ таких, что $x\cap y=0$ для всех $y\in J'$. Подобным образом мы можем образовать $e_{J'}$; кроме того. $e_J\cap e_{J'}=0,\quad e_J+e_{J'}=e.$ Таким образом, алгебра *A* изоморфна полной булевой алгебре всех компонент элемента *e*. Поскольку это есть сепарабельиая, метрическая булева алгебра, она изоморфна подалгебре алгебры с мерой всех измеримых подмножеств интервала $0\le x\le 1$ по модулю множеств меры нуль.

Аналогично, для любого $f\in L$ и любого замкнутого *l*-идеала *J* мы можем образовать $(f^+)_J$ и $(f^-)_J$; следовательно, $f_J=(f^+)_J+(f^-)_J$. Это значит, что каждое $f\in L$ имеет компоненты на каждом *J*. Для того, чтобы представить *f* вполне аддитивными оценками, понадобится еще одно замечание. Функционал

$$\lambda(f)=\|f^+\|-\|f^-\|=\|f^+\|-\|(-f)^+\| \tag{18}$$

является линейным функционалом на *L* с нормой единица; доказательство этого следующее. Мы желаем доказать, что

$\lambda(f+g)=\lambda(f)+\lambda(g)$, т. е., что

$$\|(f+g)^+\|-\|(-f-g)^+\|=\|f^+\|-\|(-f)^+\|+\|g^+\|-\|(-g)^+\|.$$

Это получается в результате перестановки из следующих тождеств:

$$\begin{aligned}\mu&=\|(f+g)^+\|+\|(-f)^+\|+\|(-g)^+\|=\\&=\|(f+g)^++(-f)^++(-g)^+\|=\|(f+g)^++(f^+-f)+(g^+-g)\|=\\&=\|[(f+g)\cup 0]-(f+g)+f^++g^+\|=\|[0\cup(-f-g)]+f^++g^+\|=\\&=\|(-f-g)^+\|+\|f^+\|+\|g^+\|.\end{aligned}$$

Тривиальным образом $\lambda(cf) = c\lambda(f)$. Мы сопоставим теперь каждому $f\in L$ следующий функционал на *A:*

$$f(J) = \lambda(f_J). \qquad (19)$$

В силу теоремы 11 $\lambda(f_J) = f(J)$ является σ-аддитивным. Кроме того, так как $(f + g)_J = f_J + g_J$ и $(cf)_J = cf_J$, сложение в L соответствует сложению оценок. Мы покажем теперь, что $\|f\|$ есть норма, введенная в теореме 15 для функций множеств: $\|f\| = \sup|\lambda(f_J) - \lambda(f_{J'})|$. Для всех J имеем

$$|\lambda(f_J) - \lambda(f_{J'})| \leqslant |\,\|f_J^+\| - \|f_J^-\| - \|f_{J'}^+\| + \|f_{J'}^-\|\,|$$

$$\leqslant \|f_J^+\| + \|f_{J'}^+\| + \|f_J^-\| + \|f_{J'}^-\| = \|f^+\| + \|f^-\| = \|f\|.$$

Но для конкретного J, порожденного элементом f^+, как слабой единицей, имеем

$$f_J = f^+,\ f_{J'} = f - f_J = f^-,\ \text{а потому}\ \lambda(f_J) = \|f^+\|,$$

$$\lambda(f_{J'}) = -\|f^-\|\ \text{и}\ |\lambda(f_J) - \lambda(f_{J'})| \geqslant \|f\|.$$

Объединяя эти неравенства, приходим к нашему результату. Наконец, остается показать, что $f^+(J) = \sup_{K\leq J} f(K)$, так что структурные операции имеют тот же самый смысл, что и для оценок. Но так как $f(K) \leqslant f^+(K) \leqslant f^+(J)$ для всех $K\leq J$, неравенство в одну сторону очевидно. Обратное неравенство получим, проектируя на замкнутый l-идеал H в L, порожденный элементом f^+, и полагая $K = J\cap H$, что дает $f(f\cap H) = f^+(f)$. Заметим далее, что, поскольку норма и структурные операции сохраняются, поскольку бесконечные объединения суть просто метрические пределы конечных объединений и поскольку любое абстрактное *(L)*-пространство есть полная структура, множества *f(J)* являются *замкнутыми* подструктурами полной структуры всех σ-аддитивных оценок на *A*. Мы резюмируем:

Теорема 18. *Любое сепарабельное абстрактное (L)-пространство изоморфно замкнутой подструктуре и метрически замкнутому линейному подпространству «конкретного» (L)-пространства всех интегрируемых по Лебегу функций f(x), определенных на единичном интервале* $0\leq x\leq 1$, *по модулю функций, равных нулю почти всюду (т. е. всюду, за исключением множества меры нуль).*

6.17. Эргодическая теория

6.17.1. Циклические полугруппы операторов перехода

Ниже мы будем иметь дело с циклическими полугруппами операторов перехода на (*L*)-пространствах в следующем точном математическом смысле.

Определение. *«Оператор перехода» на (абстрактном) (*L*)-пространстве* L *есть аддитивный оператор, переводящий распределения (т. е. положительные элементы нормы единица) в другие распределения. «Циклическая полугруппа» операторов на пространстве* S *состоит иг оператора* T *и его степеней* T^2, T^3, ... *(дискретный случай) или же из множества операторов* T^r, *определенных для всех положительных вещественных* r, *непрерывных по* r *и удовлетворяющих соотношению* $T^r\ T^s = T^{r+s}$ *для всех* r *и* s *(непрерывный случай).*

Если *T* оператор перехода, то и все его степени являются операторами перехода; следовательно, наиболее общая дискретная циклическая полугруппа операторов перехода дается степенями произвольного оператора перехода *T*.

Случай операторов перехода на (*L*)-пространствах отличается от общего случая полугрупп линейных операторов на банаховых пространствах следующим результатом.

Теорема 1. *Операторы перехода являются либо изометриями, либо сжатиями: они удовлетворяют соотношению*

$$\|fT-gT\|\leqslant\|f-g\|.$$

Доказательство. Так как $fT-gT=(f-g)\,T$, нужно показать лишь, что $\|hT\|\leqslant\|h\|$. Но $\|h^+T\|=\|h^+\|$ и $\|h^-T\|=\|h^-\|$, ибо *T* переводит положительные (а потому и отрицательные) элементы нормы единица в подобные же элементы. Следовательно,

$$\|hT\|=\|h^+T+h^-T\|\leqslant\|h^+T\|+\|h^-T\|=\|h^+\|+\|h^-\|=\|h\|.$$

6.17.2. Интерпретация: конечномерный случай

В случае, когда *L* конечномерно, легко определить все возможные операторы перехода. Действительно в силу теоремы 1 р. 6.16 *L* является тогда пространством всех *n*-векторов $f=[f_1,\ldots,f_n]$ для которых $f\geq 0$, тогда и только тогда, если каждое $f_i\geq 0$. Следовательно, операторы порехода на L суть квадратные матрицы $T = || t_{ij} ||$ порядка n, такие, что а) каждое $t_{ij} \geq 0$ и б) $t_{i1} + \ldots + t_{in} = 1$ для всех i. Иными словами, в конечномерном случае операторы перехода суть то, что обычно называют «матрицами вероятностей перехода». Именно их поведение исследуется в теории конечных «вероятностных зависимостей», называемых иначе «марковскими процессами».

Из многочисленных реализаций таких матриц «стохастическими процессами» мы выбираем одну иллюстрацию.

Пример 1. Пусть i = l, 2, ..., 52! обозначают все возможные порядки расположения карт в колоде. Опишем особенности тасовки сдающего карты, считая, что при одной тасовке он переводит колоду из состояния i в состояние j с вероятностью t_{ij}. Пусть S и T— последовательные тасовки карт различными игроками. Полная вероятность u_{ij} перехода из состояния i в состояние j будет суммой вероятностей u (i, k, j) того, что колода переходит из i в j через различные промежуточные состояния k. И если S и T *независимы,* то мы будем иметь $u(i, k, j)=s_{ik}t_{kj}$, откуда $\| u_{ij} \|$ есть матричное произведение ST.

Следовательно, если один игрок тасует карты несколько раз, не меняя характера тасовки, то результат n тасовок есть T^n, соответствующее дискретной циклической полугруппе операторов перехода.

6.17.3. Интерпретация: общий случай

Циклические полугруппы операторов перехода возникают в теории вероятностей при общих условиях, которые будут описаны.

Вероятность была определена математически в р. 6.13.9 как «оценка» (или мера) $p\ [X]$ на булевой алгебре $\mathfrak{A}$, удовлетворяющая условиям $p\ [X] \geq 0$ для всех $X \in \mathfrak{A}$ и $p\ [I] = 1$. В р. 6.16.13, было показано, что класс всех функций вероятности на любом фиксированном $\mathfrak{A}$, есть замкнутое, выпуклое множество P в абстрактном (L)-пространстве и что диаметр P равен двум, за исключением тривиальных случаев. В р. 6.16. 14, было показано, что то же самое справедливо для множества всех σ-распределений на любой булевой σ-алгебре.

Обычно элементы X, Y, ... алгебры $\mathfrak{A}$ представляют собой свойства системы $\sum$, наблюдаемые *апостериорно,* в то время как для каждого $t>0$ $p_i[X] \in P$ представляет собой наилучшее *априорное* предсказание относительно того, каково будет состояние системы $\sum$ через t единиц времени, считая от некоторого начального момента времени. Предположим, что выполняется также следующее дополнительное условие:

если $s<t$, то любое $p_s \in P$ определяет $p_t = p_sT_{s,t}$ единственным

образом, независимо от значений p_r для $r < s$. (*)

Мы можем тогда аргументировать следующим образом.

Предположение p_s с частотой λ_s и предположение q_s с частотой μ_s привели бы к заключению $p_sT_{s,t}$ в момент времени t с частотой λ_s и к заключению $q_sT_{s,t}$ с частотой μ_s. Следовательно $(\lambda_s p_s + \mu_s q_s) T_{s,t} = \lambda_s p_s T_{s,t} + \mu_s q_s T_{s,t}$. Поскольку (допустимые) распределения должны переходить в допустимые распределения, мы заключаем, что каждое $T_{s,t}$ является оператором перехода.

Далее, если $r < s < t$, то для каждого допустимого p_r

$$p_r T_{r,t} = p_t = p_s T_{s,t} = (p_r T_{r,s}) T_{s,t} = p_r (T_{r,s} T_{s,t}).$$

Следовательно, $T_{r,s}T_{s,t} = T_{r,t}$. Пусть T^u обозначает оператор $T_{0,u}$. ЕСЛИ операторы $T_{r,s}$ *однородны по времени* в том смысле, что для всех r и u $T_{0,u}=T_{r,r+u}$, то мы получаем непосредственно $T^uT^v= T^{u+v}$. Следовательно, *всякий раз, когда мы имеем однородность по времени и справедливо условие* (*), *мы получаем непрерывную циклическую полугруппу операторов перехода.*

В этой связи мы отметим непосредственное соотношение между понятием оператора перехода и законами термодинамики, совершенно не связанное с дифференциальными уравнениями теплопроводности Фурье. Пусть тепловая энергия в каждой области R изолированного твердого тела Iесть $h(R)$ и пусть тепловая энергия через пять минут есть $h^*)$. Тогда закон сохранения энергии дает нам, что из $h(I) = 1$ следует $h^*(I) = 1$, а из второго закона термодинамики вытекает, что повсюду сохраняется положительная температура по отношению к какому-нибудь нулю.

Дискретные циклические полугруппы операторов перехода возникают аналогичным образом всякий раз, когда (как и в примере 1) состояние системы $\sum$ меняется через посредство любой последовательности переходов, обусловленных повторением одной и той же причины, при условии, что $\sum$ имеет время «забыть» всю свою предыдущую историю между двумя последовательными переходами.

6.17.4. Примеры

Ниже приводятся типичные примеры «стохастических процессов», представимых циклическими полугруппами операторов перехода.

Пример 2. Счетчик Гейгера регистрирует космические излучения при средней норме λ в минуту. Пусть $\mathfrak{A}$—булева алгебра всех множеств X, Y, ... целых чисел n = 0, 1, 2, ... Пусть p_i *[X]—априорная* вероятность того, что полное число n космических излучений, сосчитанных в любой t-минутный промежуток времени, принадлежит X. Это есть стохастический процесс. Точная формула может быть раскрыта вкратце следующим образом. Для промежутка времени в Δt минут вероятности перехода выражаются приближенно инфинитезимальной квадратной матрицей $T^{\Delta t} = (I + \lambda\Delta t\, J)/(1 + \lambda\Delta t)$, где I—единичная матрица, а J имеет единицы непосредственно над главной диагональю и нули в остальных местах, откуда J^n имеет единицы на диагонали, лежащей на n рядов выше главной диагонали, и нули в остальных местах. Переходя к пределу, получаем

$$T^t = \operatorname*{Lim}_{n\to\infty} (T^{t/n})^n = \operatorname*{Lim}_{n\to\infty} \frac{1}{\left(1+\frac{\lambda t}{n}\right)^n}\left(I + \frac{\lambda t}{n} J\right)^n =$$

$$= e^{-\lambda t}\left(I + \lambda t\, J + \frac{\lambda^2 t^2}{2!} J^2 + \ldots + \frac{\lambda^n t^n}{n!} J^n + \ldots\right).$$

Отсюда вытекает известная формула

$$p_t[X] = e^{-\lambda t} \sum_{n\in X} \frac{\lambda^n t^n}{n!}. \qquad (1)$$

П р и м е р 3. Пусть ∑—положение мельчайшей частипы, взвешенной в жидкости и совершающей «броуновское движение» вследствие беспорядочных молекулярных толчков. Пусть ее положение в момент времени t = 0 есть (a_1, a_2, a_3). Пусть $\mathfrak{A}$ булева алгебра всех «элементарных» подмножеств пространства. Для каждого $X \in \mathfrak{A}$ и положительного вещественного числа t пусть $p_t[X]$ обозначает вероятность того, что частица будет в момент времени t находиться в X, в той мере, в какой эта вероятность может быть определена знанием положения частицы в момент t = 0.

Аналитическая трактовка этого случая является более трудной. Однако А. Н. Колмогоров показал, что если принять, что мы имеем циклическую полугруппу операторов перехода[2]), что они могут быть выражены в форме интегральных уравнений и явление пространственно однородно и изотропно, то

$$p_t[X] = \frac{1}{\sqrt{\pi\lambda t}} \iiint_X e^{-\frac{1}{\lambda t}\sum (x_i - a_i)^2}\, dx_1\, dx_2\, dx_3 \qquad (2)$$

для подходящего коэффициента диффузии λ.

Предположение независимости последовательных операторов перехода не выполняется для очень коротких промежутков времени. Если бы оно выполнялось, то в силу хорошо известной теоремы Вивера частица имела бы бесконечную скорость.

Следует заметить, что возникающие в примере 3 операторы перехода не являются степенями в обычном смысле инфинитезимального оператора перехода; в действительности, они не имеют даже обратных.

Следовательно, возникающая полугруппа не может быть расширена до группы.

Пример 4. Пусть $v(t)$ обозначает мгновенный вектор скорости жидкости с известными характеристиками турбулентности в фиксированной точке P и при изменяющемся времени t; пусть известны средняя скорость и $v(0)$. Тогда для каждого элементарного подмножества R скоростей, или пространства «годографа», пусть $p_t\ [R]$ обозначает вероятность того, что $v \in R$ в момент времени t, в той мере, в какой эта вероятность может быть определена. (Это может не быть полугруппой.)

Пример 5. Рассмотрим проблему n тел в теории тяготения Ньютона. Если наше знание о системе $\sum n$ тел в момент времени $t = 0$ выражается вероятностной функцией $p[X]$, то в момент времени t_1 оно выражается функцией $p_{t1}\ [X]\ p\ [X_\tau^{-1}]$, где τ есть преобразование фазового пространства, индуцируемое в момент времени t_1 дифференциальными уравнениями движения.

Наше знание не является точным вследствие ограниченной точности инструментов; в теории ошибок часто принимают, что ошибки даются гауссовым распределением.

6.17.5. Типы операторов перехода

При обычных рассмотрениях марковских процессов необходимо давать отдельные формулировки для дискретных и непрерывных физических систем, а также часто и третью формулировку для «детерминистского» случая. Наш более геометрический подход избегает необходимости давать различные формулировки и имеет дополнительное преимущество в том, что дает легко обозримые способы для выявления существенно различных случаев. Чтобы показать это, мы рассмотрим сперва случаи, в которых теорема 1 принимает специальные формы; мы используем систему обозначений п. 3.

Рассмотрим «тихистический» случай «независимых вероятностей», при котором pT не зависит от T, как и в случае последовательных бросаний пары игральных костей. Это есть случай, когда P стягивается посредством T в точку, а L проектируется на ось, проходящую через эту точку. С другой стороны, в противоположном, детерминистском случае классической механики, мы имеем *изометрическое* преобразование L.

Имеется промежуточный «стохастический» случай, типичный для теории зависимых вероятностей: здесь множество P несколько стягивается, но не превращается в точку. Случаем особой важности является тот, в котором выполняется следующее условие, впервые отмеченное Марковым.

Гипотеза Маркова. Для некоторого r в L существует положительная нижняя грань d для преобразований pT^r распределений p. Это означает, что $d > 0$ и что $d \leq pT^r$ для всех $p \in P$.

Эта гипотеза выполняется в примере 1, поскольку тасовки карт, различающиеся транспозицией всегда возможны на практике, а симметрическая группа порождается транспозициями. Она выполняется также в примере 3 при условии, если жидкость находится в *ограниченном* сосуде, а также в примере 4. Она не выполняется в примерах 2 и 5. Мы изучим теперь следствия из этой гипотезы.

6.17.6. Устойчивые распределения; теорема Маркова

Мы определим сперва распределение p как «устойчивое» относительно оператора перехода T тогда и только тогда, если оно является *неподвижной точкой* для T, т. е. тогда и только тогда, если $pT = p$. В этом случае в силу теоремы 1 распределения, первоначально близкие к p, остаются близкими к p после преобразования T и его итераций.

Теорема 2. *Множество точек произвольного (L)-пространства, оставляемых неподвижными каким-нибудь оператором перехода T, является метрически замкнутым, подпространством и подструктурой.*

Доказательство. Так как T непрерывно, множество метрически замкнуто в L. Так как T переводит верхние (нижние) грани в верхние (нижние) грани и является стягиванием, оно переводит единственную верхнюю грань $x = f \cup g$ для f и g, удовлетворяющую соотношению

$$\|f - x\| + \|x - g\| \leqslant \|f - g\|,$$

всебя.

Отсюда следует, что «устойчивые распределения» (пересечение подпространства, описанного в теореме 2, с P) образуют замкнутое выпуклое множество. Следовательно, число устойчивых распределений есть либо нуль, либо единица («метрически транзитивный» случай), либо бесконечвость. Отметим следующий пример.

Пример 6. Пусть в (L)-пространстве (lI) оператор T переводит $\chi = [x_1, x_2, x_3, \dots]$ в $[0, x_1, x_2, \dots]$. Тогда T есть изометрический оператор перехода; однако $\chi T = \chi$ только если $\chi = 0$.

Во многих случаях устойчивые распределения могут быть явно обнаружены специальными методами. Часто полезны рассмотрения симметрии. Далее, если pT^n ограниченны сверху, то $p^* = \mathrm{Lim\,sup}_{n\to\infty} pT^n$ является положительной неподвижной точкой в L и $p^*/\|p^*\|$ есть устойчивое распределение. В конечномерном случае может быть использован характеристический полином матрицы $\|t_{ij}\|$ вероятностей перехода; существование положительного характеристического вектора вытекает из предыдущего замечания и теоремы 1.

Мы покажем теперь, что если гипотеза Маркова справедлива, то мы всегда имеем метрически транзитивный случай.

Теорема 3. *Если гипотеза Маркова выполняется, то имеется единственное устойчивое распределение p_0. Кроме того, pT^k равномерно стремятся к p_0 с такой оке скоростью, с какой члены сходящейся геометрической прогрессии стремятся* к *нулю.*

Доказательство. Пусть заданы $p,q\in P$. Положим $h = p \cap q$, $f = p - h,\ g = q - h,\ \mu = 1 - \|h\|$. Тогда в силу (19), р. 6.15 а также (6) и (15), р. 6.16 $f, g \geqslant 0,\ \|f\| = \|g\| = \mu$ и $\|p - q\| = \|f\| + \|g\| = 2\mu$. Точно так же, поскольку оператор T аддитивен, $pT^r - qT^r = fT^r - gT^r$. В силу аддитивности нормы [см. р. 6.16, (15)] имеем

$$\|fT^r - gT^r\| = \|fT^r\| + \|gT^r\| - 2\|fT^r \cap gT^r\| = 2\mu - 2\|fT^r \cap gT^r\|.$$

Очевидно, что $f=\mu p_1$, для некоторого распределения p_1;. следовательно, в силу гипотезы Маркова $fT^r \geqslant \mu d$; аналогично, $gT^r \geqslant \mu d$; следовательно $fT^r \cap gT^r \geqslant \mu d$. Произведя очевидные замещения в предыдущих равенствах, выводим

$$\|pT^r - qT^r\| = \|fT^r - gT^r\| \leqslant 2\mu - 2\mu\|d\| = (1 - \|d\|)\,2\mu.$$

Так как $2\mu = \|p - q\|$, мы получаем следующую лемму.

Лемма. *T^r сжимает P равномерным образом; в обозначениях*

$$\|pT^r - qT^r\| \leqslant (1 - \|d\|)\,\|p - q\|. \tag{3}$$

Теорема 3 вытекает из (3) и из простых геометрических рассмотрений. В самом деле, если $k > nr$, то $pT^k = (pT^{k-nr})(T^r)^n$ лежит в преобразовании P путем n-кратного повторения T^r. Эти преобразования образуют убывающую последовательность $P \geqslant PT^r \geqslant PT^{2r} \geqslant \ldots \geqslant PT^{nr} \geqslant \ldots$, и в силу (3) диаметр множества PT^{nr} есть самое большее $2(1 - \|d\|)^n$. Следовательно, для любого $p\in P$ $\|pT^h - pT^k\| \leqslant 2(1 - \|d\|)^n$, если $h,k>nr$. Таким образом, pT^h удовлетворяют признаку сходимости Коти, и, так как P есть полное метрическое пространство, они сходятся к пределу p_0.

Кроме того, p_0 *устойчиво.* Действительно, так как $pT^s\to p_0$ при $s\to\infty$ и T^r непрерывно, $pT^sT^r \to p_0T^r$; но $pT^{s+r}\to p_0$ при $s\to\infty$, а потому

$p_0T^r=-p_0$. *Единственность* p_0 и тот факт, что оно является одним и тем же для всех p, вытекают теперь из (3).

Следствие 1. *В случае дискретной полугруппы для любого фиксированного* $p\ \sup pT^n \leqslant p + \sum_{k=0}^{\infty}(pT-p)T^k$ *является конечным. Для непрерывных полугрупп* sup pT^r *является подобным же образом конечным, если* pT^r *ограничены на* $0 \le r \le r_0$ *для некоторого* $r_0 > 0$.

Замечание. Это вытекает из того, что T^r непрерывно по r.

Следствие 2. *Пусть* $T_1, \ldots, T_n$ *любая последовательность операторов перехода; обозначим* $\inf_{p\in P} pT_i$ *через* d_i. *Тогда*

$$\| p(T_1T_2\ldots T_n) - q(T_1T_2\ldots T_n)\| \leqslant \|p-q\| \cdot \prod_{i=1}^{n}(1-\|d_i\|). \quad (4)$$

6.17.7. Эргодические элементы

В детерминистском (изометрическом) случае классической механики преобразования pT^r неустойчивых распределений не могут сходиться, поскольку $\|pT^{r+1}-pT^r\| = \|pT-p\|$ тождественным образом. Тем не менее их *средние* p_N $(\sum_{k=0}^{N-1} pT^k/N$ в дискретном и $\int_0^N pT^s ds/N$ в непрерывном случае) часто сходятся; теоремы, в которых доказывается это утверждение, обычно называются «эргодическими теоремами».

Эргодические теоремы имеют важное значение для статистической механики, но связь между этими теоремами и механикой все-таки еще не полная. Доказательства эргодической теоремы в классической механике были даны впервые Дж. Нейманом и Дж. Биркгофом. Эргодическая теорема последнего применима к отдельным траекториям и, следовательно, более пригодна для динамики; доказательство ее комбинаторное, а потому и теоретико-структурное по своему духу.

Поскольку теорема 3 Маркова много сильнее, чем эргодичность, представляется естественным попытаться доказать эргодические теоремы, которые применимы одновременно к детерминистскому, тихистическому и смешанному случаям. К этой идее приходили независимым образом различные авторы.

Пусть сперва $\{T^r\}$ любая дискретная или непрерывная циклическая полугруппа линейных операторов на банаховом пространстве B. Мы будем называть элемент $f \in B$ «эргодическим» тогда и только тогда, если *средние* g_s элементов fT^r, определенные выше, сходятся метрически к неподвижной точке. Пример 6 из п. 6 показывает, что даже в случае операторов перехода на (*L*)-пространствах может не существовать неподвижных точек, кроме 0; следовательно, мы не можем иметь во всех случаях эргодическую теорему в точном смысле слова. Мы можем, однако, доказать следующий результат.

Теорема 4. *Множество E элементов, эргодических относительно какой-нибудь циклической полугруппы $\{T^r\}$ линейных изометрий или сжатий любого банахова пространства, является метрически замкнутым подпространством. Оно содержит все свои образы и прообразы относительно преобразований из $\{T^r\}$.*

Доказательство. Если средние g_s для f и средние g^*_s для f^* сходятся, соответственно, к a и a^*, то средние $g_s+g^*_s$ и λg_s, соответственно для $f+f^*$ и λf, сходятся, соответственно, к $a+a^*$ и λa; следовательно, E есть подпространство. Столь же легко показать, что E метрически замкнуто. Действительно, $\|f — f^*\| < \varepsilon$ влечет $\|g_s - g^*_s\| < \varepsilon$ для всех s. Следовательно, если $\| f^{(n)} - f^* \| \to 0$ при $n \to \infty$ и каждый элемент $f^{(n)}$ эргодичен, то элементы $a^{(n)} = \lim_{s\to\infty} g_s^{(n)}$ сходятся метрически к пределу a.

Кроме того,

$$\| g_s^* - a \| \leqslant \| g_s^* - g_s^{(n)} \| + \| g_s^{(n)} - a^{(n)} \| + \| a^{(n)} - a \| .$$

Но первый и третий из этих членов ограничены числом $\| f^* — f^{(n)} \|$, а потому произвольно малы, когда n велико. В то же время при фиксированном n второй член стремится к нулю, когда $s\to\infty$. Следовательно, $g^*_s \to a$ при $s\to\infty$ и элемент f^* эргодичен. Наконец, то, что E содержит все свои образы и прообразы относительно преобразований T^r, является очевидным, поскольку каждое fT^r имеет тот же самый предел средних, что и f.

Но любое метрически замкнутое подпространство банахова пространства является слабо замкнутым; таким образом [x]) имеет место

Следствие. *Множество E замкнуто относительно слабой топологии пространства.*

Согласно С. Банаху, $x_i \to x$ «слабо» тогда и только тогда, если для каждого (ограниченного) линейного функционала λ $\lambda(x_i) \to \lambda(x)$.

6.17.8. Эргодическая теорема

Получим теперь общее достаточное условие эргодичности элемента банахова пространства.

Теорема 5 (Эргодическая теорема). *Если средние g_s, для fT^r принадлежат слабо компактному множеству, то элемент f эргодичен.*

Доказательство. В силу предположения некоторая подпоследовательность $\{g_{s(i)}\}$ из $\{g_s\}$ сходится слабо к пределу *a.* Но каждый элемент $f — fT^r$ эргодичен, ибо, если $s > r$, то

$$\left\| \frac{1}{s} \sum_{k=0}^{s-1} (f - fT^r) T^k \right\| = \frac{1}{s} \left\| \sum_{k=0}^{r-1} fT^k - \sum_{k=s}^{s+r-1} fT^k \right\| \leqslant \frac{2r}{s},$$

так что средние преобразований этого элемента сходятся метрически к 0. (Аналогичная формула имеет место для непрерывных полугрупп.) Следовательно, всякая разность $f — g_{s(i)}$, будучи средним для $f — fT^r$, принадлежит множеству *E* эргодических элементов. Но это множество слабо замкнуто (в силу следствия теоремы 4); следовательно, $f—a$ принадлежит *E.* Остается показать, что *a* принадлежит *E,* ибо отсюда будет вытекать, что $f=(f — a)+a$ принадлежит *E.* Но это является очевидным следствием леммы.

Лемма. *Если любая подпоследовательность* $\{g_{s(i)}\}$ *элементов* g_s *сходится метрически или слабо к пределу a, то a является неподвижной точкой.*

Доказательство. Не нарушая общности, можно предположить, что $\|f\| = 1$. Но в этом случае для любого целого положительного s и любого $k<s$ $s\,(g_s — g_sT^k)$ равно

$$\sum_{r=0}^{k-1} fT^r - \sum_{r=s-k}^{s-1} fT^r$$

или

$$\int_0^k fT^r dr - \int_{s-k}^{s} fT^r dr,$$

соответственно, в дискретном и непрерывном случае. В любом из этих случаев

$$\| g_s - g_sT^k \| \leqslant \frac{2k}{s}. \tag{5}$$

При метрической сходимости из $\| g_{s(i)} - a \| \to 0$ следует $\| g_{s(i)}T^k - aT^k \| \to 0$ для любого k. Но в силу (5) $\| g_{s(i)} — g_{S(i)}T^k \| \to 0$; следовательно, $\|a — aT^k\|$ меньше любой положительной константы, а потому $a=aT^k$. Аналогично, при слабой сходимости $\lambda\,(g_{s(i)} - g_{s(i)}T^k) \to 0$ для каждого линейного функционала λ и каждого *k;* следовательно, для всех λ $\lambda(a — aT^k)$ меньше любой положительной константы. Следовательно, $\lambda\,(a) = \lambda\,(aT^k)$ для каждого λ, откуда $a=aT^h$.

Из теоремы 5 можно вывести непосредственно

Следствие 1. *Пусть $\{T^r\}$ любая циклическая полугруппа линейных изометрий или сжатий банахова пространства, которое конечномерно, или же является пространством (L^p), или пространством (l^P) ($p>1$). Тогда каждый элемент эргодичен относительно $\{T^r\}$.*

Действительно, в конечномерном случае единичный шар метрически компактен, а в остальных случаях он слабо компактен. Может быть дано также прямое доказательство для пространств (L^p) и (l^p) ($p > 1$), применимое к любому «равномерно выпуклому» банахову пространству.

Следствие 2. *Если $a\leq fT^r \leq b$ для всех r и банахово пространство есть пространство (L) или (l), то элемент f эргодичен.*

Словами, элемент f эргодичен, если его преобразования ограничены в теоретико-структурном смысле. Это следует непосредственно из известного факта, что множество $a\leq f\leq b$ является слабо компактным. **Следствие 3.** *Если функция $f(x)=1$ инвариантна относительно операторов перехода T^r на пространстве (L), то каждый элемент эргодичен.*

Действительно, условие $a\leq f(x)\leq b$ для всех x сохраняется относительно преобразований T^r, поскольку операторы перехода линейны и сохраняют упорядоченность; следовательно, в силу следствия 2 все ограниченные функции являются эргодичными элементами пространства (L). Но множество ограниченных функций топологически плотно; поэтому рассуждение завершается на основании теоремы 4.

В частности, следствие 3 применимо к сохраняющим меру точечным преобразованиям любого пространства Ω конечной меры. Но, как было замечено Лиувиллем, для любой динамической системы Лагранжа потоки T^t в фазовом пространстве, соответствующие, как и в примере 5, изменению времени t, сохраняют меру в гамильтоновых координатах, так как относительно них $\dot{p}_i=\partial H/\partial q_i$ и $\dot{q}_i=\partial H/\partial p_i$. (Инфинитезимальное преобразование $\dot{x}_i = X_i$ сохраняет объем тогда и только тогда, если $\operatorname{Div} X=\sum \partial X/\partial X_i=0$. В данном случае дивергенция равна $\sum \partial^2 H/\partial p_i\,\partial q_i-\sum \partial^2 H/\partial q_i\,\partial p_i=0$.)

Поэтому, следствие 3 применимо к любой лагранжевой системе, полная мера которой конечна (т. е., фазовое пространство которой компактно). Рассмотрим теперь кинетическую теорию газов. Мы можем найти устойчивое распределение. Следовательно, если бы мы могли доказать, что поток с заданной энергией является «метрически транзитивным», то мы могли прийти к заключению, что среднее по времени распределений скорости всегда приближается к этому распределению. Но это предположение не было никогда ни доказано, ни опровергнуто.

6.17. 9. Обобщения

Распространим следствие 2 на произвольные (L)-пространства; результат принадлежит Какутани. Пусть $\{T^n\}$—дискретная полугруппа операторов перехода на абстрактном (L)- пространстве L. Для любого $f \in L$ элементы fT^n, комбинируемые линейно и теоретико-структурно, порождают *сепарабельное* замкнутое подпространство S пространства L. В силу теоремы 18 р. 6.16 оно может быть вложено в обычное пространство *(L)*. Следовательно, элементы fT^n, будучи ограничены в S элементами inf $\{fT^n\}$ и sup $\{fT^n\}$, лежат в слабо компактном множестве. Ввиду этого следствие 2 применимо также к дискретным полугруппам на абстрактных (L)-пространствах. Чтобы перейти к случаю непрерывных полугрупп, нужно лишь ограничиться рассмотрением средних элемента $g = \int_0^1 fT^r\,dr$ относительно дискретной полугруппы $T^1, T^2, T^3, \ldots$ Мы заключаем

Теорема 6. *Любой элемент f абстрактного (L)-пространства, преобразования которого при помощи циклической полугруппы операторов перехода ограничены в теоретико-структурном смысле, является эргодичным.*

Теорема 7. *Если элемент f абстрактного (L)-пространства эргодичен относительно циклической полугруппы операторов перехода и* $0 \le x \le f$, *то элемент x эргодичен.*

Доказательство. Пусть y_s и g_s обозначают s-е средние соответственно для x и f; так как операторы перехода положительны, $0 \le y_s \le g_s$ для всех s. Но в силу предположения $\{g_s\}$ сходится к устойчивому распределению a при $s \to \infty$; следовательно,

$$\| y_s - (y_s \cap a) \| \leqslant \| g_s - (g_s \cap a) \| \to 0.$$

В силу теоремы 6 элемент $y_s \cap a$ эргодичен для всех s; следовательно, «эргодическое отклонение» $\limsup_{s, s' \to 0} \| y_s - y_{s'} \|$ элемента x ограничено любым положительным числом. Действительно, поскольку каждый элемент $x — y_s$ эргодичен, как и в доказательстве теоремы 5, оно равно эргодическому отклонению каждого y_s, а последнее ограничено каждым $\| y_s - (y_s \cap a) \|$, так как элемент $y_s \cap a$ эргодичен. Следовательно, оно равно нулю и элемент x эргодичен.

Однако, вообще говоря, эргодические элементы (L)-пространства не образуют l-идеала (нормального подпространства). Чтобы это заметить, рассмотрим пример 6 с f= [1, — 1, 0, 0, ...]. Тогда элемент f эргодичен, в то время как f^+ и f^- неэргодичны.

6.17.10. Теорема Пуанкаре о возвращении

С эргодической теоремой тесно связана теорема Пуанкаре о возвращении для детерминистских процессов в фазовом пространстве с конечной полной мерой. Мы посмотрим теперь, насколько далеко этот результат может быть распространен на стохастические процессы. Чтобы это выяснить, мы примем сперва гипотезу.

Гипотеза А. Для некоторого e>0 и f такого, что 0≤f≤e, мы имеем $jT = f$.

Связь этой гипотезы с гипотезой об инвариантной мере (обязывающей исключить пример 6) очевидна. В детерминистском случае, изложенном Пуанкаре, гипотеза А влечет

$$f \cap \limsup_{n\to\infty} \{fT^n\} = f. \tag{6}$$

В этом случае, где мы имеем дело с характеристическими функциями множеств, условие (6) эквивалентно, кроме того, условию

$$\operatorname*{Lim}_{m\to\infty} f \cap m \limsup_{n\to\infty} \{fT^n\} = f. \tag{6'}$$

Основным выводом является следующий.

Теорема 8. *Для любого оператора перехода T гипотеза А влечет свойство* (6').

Доказательство. Обозначим $\limsup_{n\to\infty}\{fT^n\}$ через f*; на языке р. 6.15.11, теорема 8 утверждает, что f лежит в замлнутом l-идеале J (f *), порожденном элементом f*. Докажем это утверждение. Пусть $g = f - \lim_{m\to\infty} f \cap mf^*$ компонента элемента f в дополнении J' l-идеала J(f*) и пусть $g^* = \limsup_{n\to\infty}\{gT^n\}$. Тогда, так как g≤$f$, мы имеем g*≤f*. Кроме того, g ∩ f* = 0, так как J(f*) и J' дизъюнктные l-идеалы; следовательно, g ∩ g* = 0. Утверждение теоремы равносильно тому, что g = 0; это мы теперь и докажем.

Очевидно, что $g \leqslant e = g^* \dotplus (e - g^*)$, откуда

$$g \leqslant [g \cap g^*] + [g \cap (e - g^*)] = 0 + g \cap (e - g^*) \leqslant e - g^*.$$

В силу определения g* очевидно также, что $g^*T = g^*$. Поэтому гипотеза А будет все еще выполняться, если мы подставим $e_1 = e - g^*$ вместо e и g вместо f. Далее, так как $g \le e_1$ и $e_1T = e_1$, каждое $gT^n \le e_1$, а потому $g^* \le e_1$. Следовательно, мы можем рассуждать, как и выше, получив в результате $g \le e_x - g^* = e - 2g = e_2$. Повторяя рассуждение k-раз, мы имеем $e - kg^* \ge g \ge 0$ для всех k. Следовательно, $g^* = 0$ и $\|g\| \le \|g^*\| = 0$, чем и завершается доказательство.

7. Методы организации моделей структур

7.1. Модель и отношения

Одним из основных в математической теории структур является понятие модели. *Моделью* Ψ называется совокупность множества *M* с заданными в нем отношениями

$$S = \{R_{11}, R_{12}, ..., R_{1n_1}, R_{21}, R_{22}, ..., R_{2n_2}, ...R_{m1}, R_{m2}, ..., R_{mn_m}\},$$

где множество *M* — *носитель модели,* а заданные отношения R_{ia}, $R_{ia} \subset M^i$ образуют *сигнатуру модели*

$$\Psi = \langle M,\ S \rangle.$$

Степень носителя определяет *арность отношения*. Два отношения R_α и R_β, имеющие одну и ту же степень, называются *совместимыми по объединению* или просто *совместимыми.*

Очевидно, что *n*-местную операцию $f_n(m_1, m_1, \ldots, m_n) = m_{n+1}$ можно рассматривать как $(n+1)$-арне отношение R_{n+1}.

Совокупность множества *M* с заданными в нем операциями и отношениями, будем называть *алгебраической системой.*

Частным случаем алгебраической системы есть *алгебра отношений* и ее расширение — *реляционня алгебра.*

Рассмотрим алгебру отношений, носитель которой — множество отношений, а сигнатура - операции объединения, пересечение, разности и расширенного декартова произведения отношений.

Объединением $R_\alpha \cup R_\beta$ *двух совместимых отношений* R_α и R_β является множество всех кортежей, каждый из которых принадлежит хотя бы одному з этих отношений. Объединением отношений

$$R_\alpha = \{(a,b,c),(a,b,d),(b,c,e)\} \text{ и } R_\beta = \{(a,b,d),(b,d,e),(c,d,e)\}$$

является

$$R_\alpha \cup R_\beta = \{(a,b,c),(a,b,d),(b,c,e),(b,d,e),(c,d,e)\}.$$

Рассмотренные отношения являются совместимыми, так как их степени равны:

$$s(R_a) = s(R_b) = 3,\ R_a, R_b \subset M^3,\ M = \{a,b,c,d,e\}.$$

Пересечением $R_\alpha \cap R_\beta$ *двух совместимых отношений* R_α и R_β является множество всех кортежей, которые принадлежат как отношению R_α, так и отношению R_β. Пересечением отношений R_α и R_β является

$$R_\alpha \cap R_\beta = \{(a,b,c),(a,b,d),(b,c,e)\} \cap \{(a,b,d),(b,d,e),(c,d,e)\} = \{(a,b,d)\}.$$

Разностью $R_\alpha \backslash R_\beta$ *двух совместимых отношений* R_a и R_β является множество всех кортежей, которые принадлежат отношению R_a и не приналежать отношению R_β. Так, например,

$$R_\alpha \backslash R_\beta = \{(a,b,c),(a,b,d),(b,c,e)\}/ \{(a,b,d),(b,d,e),(c,d,e)\}= \{(a,b,c),(b,c,e)\}.$$

Расширенным декартовым произведением $R_a \times R_\beta$ *двух отношений* R_a и R_β является множество всех кортежей π таких, что π — конкатенация кортежа $a \in R_a$ и кортежа $b \in R_\beta$ *(конкатенация* кортежей $(a_1, a_2,...,a_n)$ и $(b_1, b_2,...,b_m)$ - кортеж $(a_1, a_2,...,a_n, b_1, b_2,...,b_m)$). Например, для рассмотренных отношений R_a и R_β расширенное декартово произведение

$$R_a \times R_\beta = \{(a,b),(c,d),(a,e)\} \times \{(a,b,c),(b,d,e)\}= \{(a,b,a,b,c),(a,b,b,d,e),$$
$$(c,d,a,b,c),(c,d,b,d,e),(a,e,a,b,c),(a,e,b,d,e)\}.$$

Понятие модели и алгебры отношений находят широкое применение при формализации реальных структурных объектов. Рассмотрим, как используется алгебра отношений при создании информационного обеспечения — разработке реляционной базы данных.

Основой построения реляционной базы данных является двумерная таблица, каждый столбец которой соответствует домену (или атрибуту, соответствуещему части домена), строка - кортежу значений атрибутов, находящихся в отношении *R*.

Рассмотрим 5-арне отношение R_5 (испытания) (табл. 1).

Таблица 1.

R_5	D_1	D_2	D_3	D_4	D_5
1	К5-01	ТЕОРИЯ АВТОМАТОВ	ПРОФ. ПЕТРЕНКО	03. ЯНВ.	АУД. 210
2	К5-02	МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА	ПРОФ. ИВЧЕНКО	03. ЯНВ.	АУД. 211
3	К5-03	ФИЗИКА	ПРОФ. КРАВЧЕНКО	03. ЯНВ.	АУД. 211
4	К5-04	АЛГОРИТМИЧЕСКИЕ ЯЗЫКА	ПРОФ. ИВЧЕНКО	05. ЯНВ.	АУД. 210
5	К5-01	ФИЗИКА	ПРОФ. КРАВЧЕНКО	09. ЯНВ.	АУД. 210
6	К5-	ТЕОРИЯ АВТОМАТОВ	ПРОФ. ПЕТРЕНКО	09. ЯНВ.	АУД. 211

	02				
7	К5-03	АЛГОРИТМИЧЕСКИЕ ЯЗЫКА	ПРОФ. ИВЧЕНКО	10. ЯНВ.	АУД. 211
8	К5-04	МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА	ПРОФ. ПЕТРЕНКО	10. ЯНВ.	АУД. 210

Таблица 1 определяет отношение реляционной модели данных. Отношение R_5 является подмножеством декартова произведения $D_1 \times D_2 \times D_3 \times D_4 \times D_5$, в котором сомножитель является доменом D_i. Элементами домена D_i служат значения атрибутов.

Домен D_1 (группа) содержит значение К5-01, К5-02, К5-03, К5-04:

$$D_1 = \{К5\text{-}01, К5\text{-}02, К5\text{-}03, К5\text{-}04\};$$

аналогично имеем домены:

D_2 (дисциплина):

D_2={ТЕОРИЯ АВТОМАТОВ, МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА, ФИЗИКА, АЛГОРИТМИЧЕСКИЕ ЯЗЫКИ};

D_3 (экзаменатор):

D_3={ПРОФ. ПЕТРЕНКО, ПРОФ. ИВЧЕНКО, ПРОФ. КРАВЧЕНКО};

D_4 (дата):

D_4 = {03 ЯНВ., 05 ЯНВ., 09 ЯНВ., 10 ЯНВ.};

D_5 (аудитория):

D_5 = {АУД. 210, АУД. 211}.

Порядок столбцов в таблице фиксирован, строки в общем случае могут располагаться произвольно. Цифры первого столбца 1, 2,..., 8 идентифицируют элементы отношения D_5.

Для преобразования отношений определим *реляционную алгебру*. Носитель реляционной алгебры прелставляет собой множество отношений, сигнатура кроме введенных операций (объединение, пересечение, разность и расширенное декартово произведение отношений) включает специальные операции над отношениями: выбор, проекцию и соединение.

Операция *выбора* представляет собой процедуру построения «горизонтального» подмножества отношений, т.е. подмножества кортежей, обладающих заданным свойством.

Пример 2. С помощью операции выбора построить отношение R'_5 (расписание экзаменов проф. Петренка). Результатом операции выбора являются строки, у которых домен D_3 представлен значениям, ПРОФ. ПЕТРЕНКО; это 1, 6, 8-я строки (табл. 2).

Таблица 2.

R_5	D_1	D_2	D_3	D_4	D5
1	К5-01	ТЕОРИЯ АВТОМАТОВ	ПРОФ. ПЕТРЕНКО	03.ЯНВ.	АУД. 210
2	К5-02	ТЕОРИЯ АВТОМАТОВ	ПРОФ. ПЕТРЕНКО	09. ЯНВ.	АУД. 211
3	К5-04	МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА	ПРОФ. ПЕТРЕНКО	10. ЯНВ.	АУД. 210

Для определения проекций отношений множество в реляционной алгебре разбивается на два подмножества в случае бинарного отношения и на n подмножеств в случае n-арного отношения:

$$R_2 \subset M^2,\ M=A \square B, A \mathrm{I}\ B=\varnothing, R_2 \subset A\times B;$$

$$R_n \subset M^2, M=\bigcup_{i=1}^{n} A_i;\ A_{i_a}\ \mathrm{I}\ A_{i_b}=\varnothing,$$

$$i_a, i_b\ (i_a \neq i_b) \in \{i_1, i_2,\ldots,i_n\}, R_n \subset A_1\times A_2\times\ldots\times A_n\ldots$$

Проекцией Пр(R_2/A) *бинарного отношения* R_2, $R_2 \subset A\times B$ на A называется множество элементов $\{a_i/(a_i,b_i) \in R_2\}$.

Проекцией Пр$(R_n/ A_{i_1}, A_{i_2},\ldots, A_{i_m})$ n-арного отношения

$R_n \subset A_1\times A_2\times\ldots\times A_n$, $n\geq m$, на $A_{i_1}, A_{i_2},\ldots, A_{i_m}$ называется множество кортежей $(a_{i_1}, a_{i_2},\ldots, a_{i_m})$, где $a_{i_1} \in A_{i_1}, a_{i_2} \in A_{i_1},\ldots, a_{i_m} \in A_{i_1}$, каждый из которых является частью элемента n-арного отношения R_n.

Операция проекции определяет построение «вертикального» подмножества отношения, т.е. множества подмножества кортежей, получаемого выбором одних и исключением других доменов.

Пример 3. Проекция Пр $(R_5/D_2, D_3)$ порождает множество пар**,** каждая из которых определяет дисциплину и экзаменатора (табл. 3).

Таблица 3

R_2	D_2	D3
	ТЕОРИЯ АВТОМАТОВ	ПРОФ. ПЕТРЕНКО

	МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА	ПРОФ. ИВЧЕНКО
	ФИЗИКА	ПРОФ. КРАВЧЕНКО
	АЛГОРИТМИЧЕСКИЕ ЯЗЫКИ	ПРОФ. ИВЧЕНКО
	МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА	ПРОФ. ПЕТРЕНКО

Одинаковые строки в табл. 3 объединены в одну.

Операция соединения по двум таблицам, имеющим общий домен, позволяет построить одну таблицу, каждая строка которой образуется соединением двух строк исходных таблиц. Из заданных таблиц берут строки, содержащие одно и то же значение из общего домена; общему домену сопоставляется один столбец.

Пример 4. Найдем по двум заданным таблицам (табл. 4,*а*, 4,*б*) результат операции соединения по домену D_1 (табл. 4,*в*).

Таблица 4, *а*

R_5	D_1	D_2	D_3	D_4	D_5
	К5-01	ТЕОРИЯ АВТОМАТОВ	ПРОФ. ПЕТРЕНКО	03.ЯНВ.	АУД. 210
	К5-02	МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА	ПРОФ. ИВЧЕНКО	03. ЯНВ.	АУД. 211
	К5-03	ФИЗИКА	ПРОФ. КРАВЧЕНКО	05. ЯНВ.	АУД. 211
	К5-04	АЛГОРИТМИЧЕСКИЕ ЯЗЫКА	ПРОФ. ИВЧЕНКО	05. ЯНВ.	АУД. 210

Таблица 4, *б*

R_5	D_1	D_2	D_3	D_4	D_5
	К5-01	ФИЗИКА	ПРОФ. КРАВЧЕНКО	09. ЯНВ.	АУД. 210
	К5-04	МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА	ПРОФ. ИВЧЕНКО	10. ЯНВ.	АУД. 210
	К5-02	ТЕОРИЯ АВТОМАТОВ	ПРОФ. КРАВЧЕНКО	09. ЯНВ.	АУД. 211
	К5-03	АЛГОРИТМИЧЕСКИЕ	ПРОФ.	10.	АУД.

	ЯЗЫКА	ИВЧЕНКО	ЯНВ.	211

Таблица 4, *в*

R_5	D_1	D_2	D_3	D_4	D_5	D'_2	D'_3	D'_4	D'_5
	К5-04	ТЕОРИЯ АВТОМАТОВ	ПРОФ. ПЕТРЕНКО	03 ЯНВ.	АУД. 210	ФИЗИКА	ПРОФ. КРАВЧЕНКО	09 ЯНВ.	АУД. 210
	К5-02	МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА	ПРОФ. ИВЧЕНКО	03 ЯНВ.	АУД. 211	ТЕОРИЯ АВТОМАТОВ	ПРОФ. ПЕТРЕНКО	09 ЯНВ.	АУД. 211
	К5-03	ФИЗИКА	ПРОФ. КРАВЧЕНКО	05 ЯНВ.	АУД. 211	АЛГОРИТМИЧЕСКИЕ ЯЗЫКИ	ПРОФ. ИВЧЕНКО	10 ЯНВ.	АУД. 211
	К5-04	АЛГОРИТМИЧЕСКИЕ ЯЗЫКИ	ПРОФ. ИВЧЕНКО	05 ЯНВ.	АУД. 210	МАТЕМАТИЧЕСКАЯ ЛИНГВИСТИКА	ПРОФ. ПЕТРЕНКО	10 ЯНВ.	АУД. 210

Аналогично можно определить операцию соединения не только по условию «равенства», но и по другим условиям сравнения: $>$, $\geqslant$, $\neq$, $<$, $\leqslant$.

Определим, например, операцию соединения по условию «больше, чем» ($>$).

Соединением по условию «больше, чем» отношения R_a по атрибуту X и отношение R_b по атрибуту Y (атрибуты X и Y являются атрибутами одного и того же домена, общего для отношений R_a и R_b), $X > Y$, называется множество всех кортежей π_i таких, что π_i — конкатенация кортежа a_i, принадлежащего R_a, и кортежа b_i принадлежащего R_b, где X — часть a_i, Y — часть b_i и $X > Y$.

Запрос в реляционной базе данных будет выполнен тем быстрее, чем меньше операций над отношениями он содержит.

7.2. Отношения на базах данных и структурах данных

Как уже нами установлено, все вокруг определяется отношениями. Довольно лишь взять отношение s на переменных $(x_1, ..., x_n)$ так, чтобы можно было построить множество

$$\{(x_1, x_2, ..., x_n): s\ (x_1, x_2, ..., x_n) \text{ истинно}\}.$$

Пусть задан набор элементов $(x_1, x_2, ..., x_n)$. Отношение $s(x_1, ..., x_n)$ можно *разрешить,* т.е. выяснить, $s(x_1, ..., x_n)$ истинно или ложно. Конечно, s не обязательно будет представлено «хорошей» формулой. Нетрудно показать, что вместо отношения s, определяющего множество наборов длины n, любое множество таких наборов также определяет отношение (и содержательные свойства s).

Эти два подхода эквивалентны.

Определение. При обработке данных наборы из n элементов называют *записями;* элементы этих наборов называют *полями.* Записи, определяющие отношение, обычно содержатся в файле. Если потребовать, чтобы несколько файлов содержали совокупность записей, которые удовлетворяют некоторым отношением, то мы получим (относительную) *базу данных.*

Замечание. Для случая обработки данных мы сейчас употребили термин «поле». В следующих разделах мы будем употреблять этот термин в математическом смысле, однако в данном случае это не приводит к недоразумению.

Таким образом, это дает нам первый реальный пример отношений, которые в большей мере связаны с вычислениями, в частности, с прикладными задачами. Тем не менее краткое обсуждение некоторых простейших свойств баз данных не только обеспечивает основу дальнейшего математического исследования отношений, но и проясняет некоторые факторы, понимание которых необходимо для эффективного управления системами баз данных.

Теория баз данных как организованных структур систем включает в себя изучение так называемых нормальных форм, однако обоснование некоторых из них очевидно лишь в простых случаях. Мы рассмотрим только три формы для следующих задач:

— вставить в структуру новый набор из n элементов;

— удалить из структуры набор из n элементов;

— модифицировать набор (структуру), который содержит n элементов.

Начнем с простейшей нормальной формы.

Определение. *Файлы в первой нормальной форме* (INF), или - более просто — *нормализованные файлы,* имеют записи фиксированной длины, которые состоят из элементов, взятых из множеств, чьи элементы далее не могут быть разбиты, и в каждый момент времени этот файл может быть представлен как массив значений $M \times N$. Каждая запись, будучи набором из n элементов, может быть записана как строка массива.

Пример 1. Рассмотрим отношение FAM1, в котором мы собрали вместе родителей и детей. Каждая запись содержит в указанном порядке фамилию и имена отца, матери и детей. Итак, мы имеем записи

(Шевчук, Иван, Нина , (Ира, Борис)) $\in$ FAM l,

(Кравчук, Петр, Надежда, (Люда)) $\in$ FAM l.

Теперь, если мы обозначим через F и M множества родителей и матерей, то из определения следует, что

Иван (Шевчук) $\in F$, Нина (Шевчук) $\in M$,

Петр (Кравчук) $\in F$, Надежда (Кравчук) $\in M$,

Таким образом, Люда является членом семьи Кравчук, но Ира и Борис не являются детьми семьи Кравчук. Так как в этой семье больше одного ребенка, то соответствующая запись больше, и, следовательно, нарушены условия первой нормальной формы.

Из FAM1 мы можем получить отношение FAM2, построив его из S, F, M и C, где S — множество фамилий, а C — множество детей, путем конструирования записей:

(Шевчук, Иван, Нина , Ира),

(Шевчук, Иван, Нина, Борис),

(Кравчук, Петр, Надежда, Люда).

Отношение FAM2 находится в 1NF и может быть представлено с помощью табл. 1.

Таблица 1

Фамилия	Отец	Мать	Ребенок
Шевчук	Иван	Нина	Ира
Шевчук	Иван	Нина	Борис
Кравчук	Петр	Надежда	Люда

Однако не совсем ясно, что будет, если, например, мужчина и женщина Юрчук не имеют детей? Если мы хотим иметь в файле запись о них, то следует пересмотреть структуру файла. Это означает, что все следует начать сначала. Введем следующую терминологию.

Определение. При использовании таблицы для изображения отношения (файла с n-мерными наборами/записями, которые записываются в виде строк) столбцы называются *атрибутами.*

Слкдовательно, ФАМИЛИЯ, ОТЕЦ, МАТЬ и РЕБЕНОК есть атрибутами различных полей в FAM2. Для получения доступа к записям в файле используются так называемые ключи. Более точно это может быть определено в сроках атрибутов.

Определение. Атрибут или (упорядоченное) множество атрибутов, чьи значения однозначно определяют запись в файле, называются *ключом* этого файла. (Заметим, что в файле может быть много различных ключей).

Каждый ключ отношения/файла FAM2 должен включать атрибут РЕБЕНОК.

Перейдем к другим примерам.

Пример 2. Каждый владелец компьютера должен покупать к нему запасные части. Поэтому мы можем рассмотреть файл SUP1, структура которого показана в табл. 2.

Таблица 2

КОМПАНИЯ	ОТДЕЛЕНИЯ	МЕНЕДЖЕР
ACCE	КИЕВ	КОРЖ
IBL	КИЕВ	ИВАНОВ
DATAMETZ	ЖИТОМИР	ИВАНОВ
PRINTACO	ОДЕССА	КОЧАН

WOOLIES	ЖИТОМИР	КОЧАН
RTX	КИЕВ	КОРЖ
OXONDATA	ЛЬВОВ	КАЗАК

Атрибут КОМПАНИЯ является ключом в SUP1; вся другая информация в файле доступна с помощью ключа. Таким образом, например, можно извлечь атрибут ОТДЕЛЕНИЕ с помощью ключа WOOLIES или же МЕНЕДЖЕР из RTX.

Определение. Если запись локализована с помощью некоторого ключа, то поле, которое выделяется из этой записи, называется *проекцией.* В данном контексте проекцией является «из». Будем также говорить, что эти атрибуты *зависят от ключа.*

На рис. 1 представлен графический пример зависимостей в SUP1.

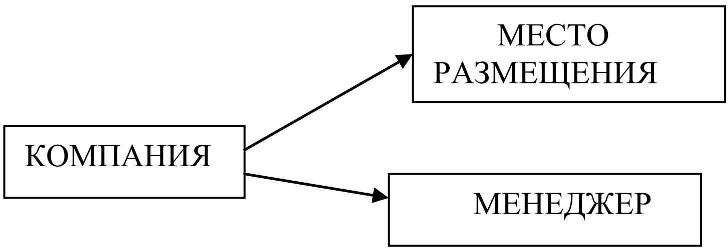


Рис. 1.

Пример 3. Модифицируем файл SUP1 с целью включения туда информации об имеющихся на складе запасных частях и об их количестве, которые отдельный поставщик желает продать в отдельной сделке. Включим в файл также код поставки, из которого мы можем выяснить скорость и частоту поставок. Во избежание лишней детализации введены номер компании и номер запасной части (табл. 3).

Таблица 3

КОМПАНИЯ	МЕСТО ПОЛО-ЖЕНИЯ	ПОСТАВЩИК	ЗАП. ЧАСТЬ	КОЛИЧЕСТВО
1	КИЕВ	2	1	10
1	КИЕВ	2	2	1
1	КИЕВ	2	3	10

2	КИЕВ	2	2	2
2	КИЕВ	2	4	5
3	РОССИЯ	4	2	2
3	РОССИЯ	4	3	10
3	РОССИЯ	4	4	4
5	РОССИЯ	4	1	10
5	РОССИЯ	4	3	10
6	КИЕВ	2	1	10

Из табл. 3 выясним, какие преобразования мы можем делать с SUP2, а какие нет:

а) *вставка*. Например, мы не можем прибавить в файл запись, которая указывает, что компания 4 (PRINTACO) находится в Одессе, без указания деталей, которые она может поставлять;

б) *удаление.* Если компания 6(RTX) прекратила поставки запчастей 1, тогда мы обязаны удалить все записи, которые относятся к этой компании и имеющие в поле ЗАП. ЧАСТЬ код 1;

в) *модификация*. Если код поставщика для Киева изменился, например, из-за транспорта, то соответствующее поле должно быть изменено в каждой записи, где есть код Киев в поле МЕСТО РАСПОЛОЖЕНИЕ. Что можно сделать для того, чтобы уменьшить или отодвинуть эти проблемы? С практической точки зрения мы должны выделить информацию в SUP2 так, чтобы по возможности избежать повторений. Таким образом, мы получаем возможность вставки/удаления части записи в SUP2. Возможнаое и разумное разделение дается в SUP3 (табл. 4). Тогда остаток информации в SUP2 может содержаться в поле ЗАП.ЧАСТЬ.

Таблица 4

SUP3

КОМПАНИЯ	*МЕСТО РАСПОЛОЖЕНИЯ*	*ПОСТАВЩИК*
1	КИЕВ	2
2	КИЕВ	2
3	РОССИЯ	4
5	РОССИЯ	4
6	КИЕВ	2

Таблица 5

КОМПАНИЯ	*МЕСТО РАСПОЛОЖЕНИЯ*	ПОСТАВЩИК
1	КИЕВ	7
2	КИЕВ	7
3	РОССИЯ	4
4	НИВКИ	3
5	РОССИЯ	4

	6 КИЕВ	7

ЗАП. ЧАСТЬ

КОМ-ПАНИЯ	*ЗАП. ЧАСТЬ*	*КОЛИ-ЧЕСТВО*
1	1	10
1	2	1
1	3	10
2	2	2
2	4	5
3	2	2
3	3	10
3	4	4
5	1	10
5	3	10
6	1	10

ЗАП. ЧАСТЬ

КОМ-ПАНІЯ	*ЗАП. ЧАСТЬ*	КОЛИ-ЧЕСТВО
1	1	10
1	2	1
1	3	10
2	2	2
2	4	5
3	2	2
3	3	10
3	4	4
5	1	10
5	3	10

Используя эту конфигурацию, можно, например:

а) включить в SUP3 запись, которая означает, что компания 4 находится на Нивках (код поставщика 3);

б) удалить ссылку на компанию 6 как на поставщика запчасти 1, но оставить соответствующий код в SUP3;

в) изменить код поставщика для КИЕВ на 7 путем замены только трех входов, которые отвечают компаниям с кодом 1, а не всех шести.

Результаты этих изменений приведены в табл. 5. Это уже значительно лучше, однако все же может быть еще усовершенствовано. Чтобы увидеть, в каком направлении продолжать исследование, отделим ключи и подчиненные части (рис. 2). Заметим, что ЗАП. ЧАСТЬ требует объединенного ключа.

Все не-ключи непосредственно связаны с ключом. Это дает нам следующее свойство нормальной формы.

Определение. Файл имеет *вторую нормальную форму* (2NF), если он имеет форму 1NF и неключевые атрибуты целиком независимы от ключа.

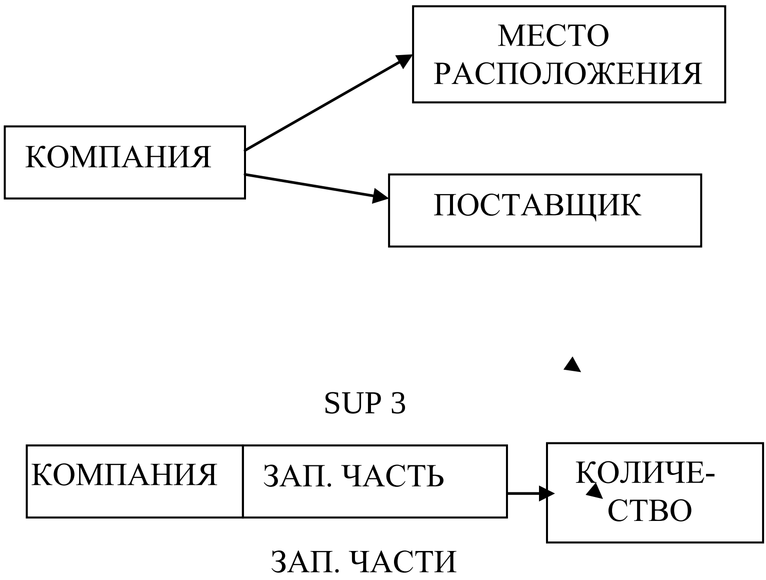


Рис. 2.

Пример 3 (продолжение). Файл SUP3 все еще является довольно сложным в том смысле, что для данной записи ПОСТАВЩИК может быть установлен с помощью исследования поля КОМПАНИЯ или же поля МЕСТО РАСПОЛОЖЕНИЯ. Это является причиной того, что в требовании а) код поставщика для Нивок должен быть вставлен перед записью кода 4 компании, а требование б), возможно, потребует изменить более одной записи для модификации единого поля данных, относящегося к коду поставщика. На практике мы можем убрать эту проблему путем проектирования SUP3 в SUP4 и DEL (табл. 6). (Заметим, что коды поставщика, изменяемые таким образом, препятствуют той возможности, что любые другие записи в файле вызывают противоречивую информацию. В SUP3 можно иметь запись вида «ПОСТАВЩИК КОМПАНИИ 6-2» и «ПОСТАВЩИК КОМПАНИИ 1-7» на некотором этапе модификации SUP2, несмотря на тот факт, что обе компании находятся в КИЕВЕ.)

Зависимость отношений в SUP4 и DEL изображена на рис. 3.

КОМПАНИЯ → МЕСТО РАСПОЛОЖЕНИЯ

SUP4

МЕСТО РАСПОЛОЖЕНИЯ → ПОСТАВЩИК

DEL

Рис. 3.

Нетранзитивность отношения зависимости является внутренним свойством, из которого возникает понятие третьей нормальной формы.

Таблица 6.

КОМПАНИЯ	МЕСТО РАС-ПОЛОЖЕНИЯ
1	КИЕВ
2	КИЕВ
3	РОССИЯ
4	НИВКИ
5	РОССИЯ
6	КИЕВ

МЕСТО РАС-ПОЛОЖЕНИЯ	ПОСТАВЩИК
КИЕВ	7
РОССИЯ	4
НИВКИ	3

Определение. Файл находится в *третий нормальной форме* (3NF), если он является файлом 2NF, и каждый атрибут, который не является ключом, нетранзитивным образом зависит от ключа. Возможен и другой путь - каждый атрибут, который не является ключом, зависит только от ключа и ни от чего другого.

Как было отмечено ранее, существует много других «нормальных» форм, но мы не ставим изучения файлов своей целью в дальнейшем.

Достаточно лишь иметь в виду, что информация в файлах является одной из реализаций математического понятия отношения. Практическое использование отношений SUP4 и DEL требует явной связи атрибута МЕСТО РАСПОЛОЖЕНИЯ файла SUP4 с атрибутом МЕСТО РАСПОЛОЖЕНИЯ файла DEL. Это - отношение эквивалентности (между компонентами различных файлов, имеющих одно и то же имя). Подобные отношения эквивалентности могут быть использованы для определения внутренних связей и других структурных элементов. В качестве иллюстрации рассмотрим рис. 4.

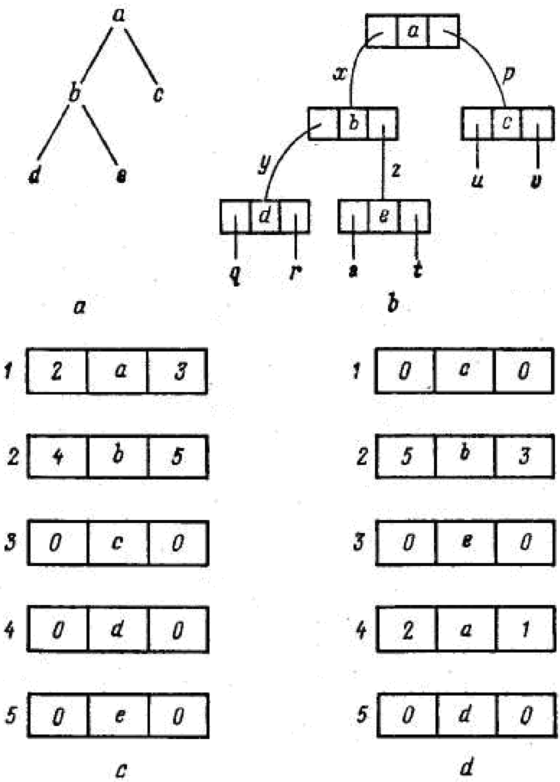


Рис. 4.

Диаграмма на рис. 4, *a* изображает дерево, диаграмма на рис. 4, *b* подобна диаграмме структурных данных, а диаграммы на рис. 4, *c* и *d* описывают их возможные применения. Отметим, что отношение

эквивалентности различны, но результирующие структурные связи сохраняются.
С математической точки зрения это является разбиением на классы эквивалентности. Следовательно, мы можем определить произвольное представление этого дерева как элемент множества $T = D/E$, где

$$D=\{a=(x, a, p), b = (y, b, z), c = (u, c, v), d=(q, d, r), e=(s, e,t)\},$$
$$E = \{(x,b), (y,d), (p,c), (z, e)).$$

Эти вопросы обсуждались нами раньше.

7.3. Составные отношения

Подобно тому как мы устанавливали внутренние связи в файлах, для выделения некоторых данных из информации (например, ПОСТАВЩИК и МЕСТО РАСПОЛОЖЕНИЯ 3 посредством файлов DEL и SUP4 в примере 3) при образовании структур часто приходится связывать бинарные отношения друг с другом. Руководствуясь предыдущими рассуждениями, можно определить это понятие следующим образом.

Определение. Пусть заданы множества *A, B* и *C* и отношение σ между *A* и *B* и ρ между *B* и *C.* Определим отношение между *A* и *C* следующим образом: оно действует из *A* в *B* посредством σ, а потом из *B* в *C* посредством ρ. Такое отношение называют *составным* и обозначают $\rho \circ \sigma$, т.е.

$$(\rho \circ \sigma)(a)=\rho(\sigma(a)).$$

Следовательно, $(x, y)\in (\rho \circ \sigma)$, если существует $z\in B$ такое, что $(x, z)\in\sigma$ и $(z, y)\in\rho$. Отсюда следует, что $\mathcal{D}_{\rho\circ\sigma}=\sigma^{-1}\mathcal{D}_{\rho}$. Чтобы проиллюстрировать ситуацию, рассмотрим рис. 5.

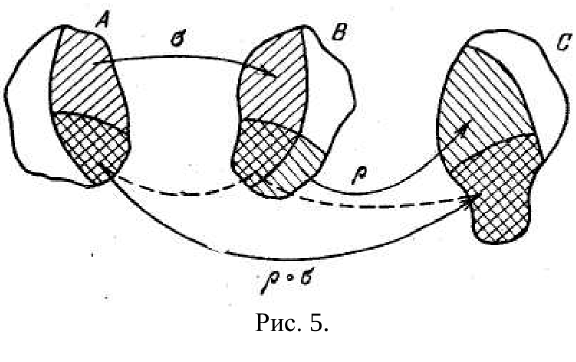


Рис. 5.

Области определения и значений σ и ρ заштрихованы в разных направлениях. Следовательно, сегменты с двойной штриховкой на *A, B* и *C* представляют собой $\mathcal{D}_{\rho\circ\sigma}$, $\mathcal{D}_\rho$ □$\mathcal{R}_\sigma$ и $\mathcal{R}_{\rho\circ\sigma}$, соответственно.

Замечание. Из записи отношений σ и ρ следует, что они применяются справа налево. Следовательно, $(\rho\circ\sigma)(a)$ означает, что вначале берется *a* и преобразуется посредством σ, а затем преобразуется посредством ρ. В алгебре это иногда записывают в виде $a\sigma\rho$. Следует обращать внимание при чтении других математических книг на то, какой порядок выполнения отношений приняты в той книге.

Пример 4. Пусть σ и ρ — отношения на *N* такие, что

$$\sigma = \{(x, x+1): x\in N\},\quad \rho = \{(x^2, x): x\in N\}.$$

Тогда

$$\mathcal{D}_\rho = \{x^2: x\in N,\quad \mathcal{D}_\sigma = \{x: x, x+1\ x\in N = N,$$

$$\mathcal{D}_{\rho\circ\sigma}=\sigma^{-1}\mathcal{D}_\rho=\{x: x\in N \text{ и } x+1 = y^2\},\ \text{где } y\in N = \{3, 8, 15, 24, ...\}$$

(рис. 6).

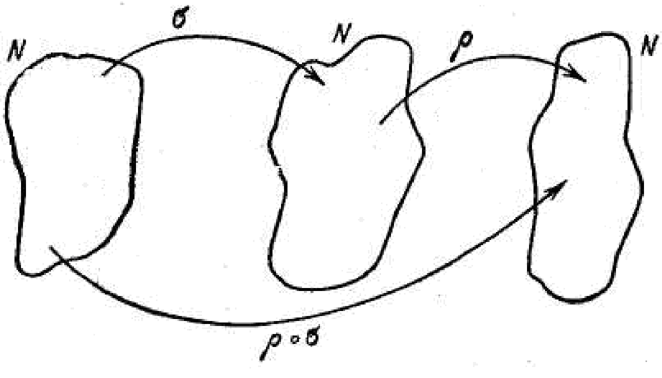


Рис. 6.

В случае, когда мы рассматриваем отношение на множестве, она может быть скомбинировано само с собой.

Например, используя отношение из примера 4, имеем

$$\sigma\circ\sigma = \{(x, x+2): x\in N\} \text{ и } \rho\circ\rho = \{(x^4, x): x\in N\}.$$

Эти отношения можно также обозначать соответственно σ^2 и ρ^2. В общем-то эти обозначения не совсем законы для множеств, однако их легко можно обосновать, поскольку если $(x, y)\in\sigma\circ\sigma$, то $((x, z), (z, y))\in\sigma\times\sigma$ при некоторому z; никакого недоразумения при этом не возникает, поскольку известна структура получаемого результата.

Используя это обозначение, мы можем определить σ^n для любого $n\in N$, $n > 1$, следующим образом:

$$\Sigma^n = \{(x, y): x\sigma z \text{ и } z\sigma^{n-1}y \text{ для некоторого } z\}.$$

Если мы вновь возьмем отношение σ и ρ из примера 4, то получим

$$\Sigma^n = \{(x, x + n), x\in N\} \text{ и } \rho^n = \{\{x^{2^n}, x): x\in N\}.$$

Рассмотрим вопрос о том, насколько в этих случаях может быть применена аналогия с умножением. Пусть A — множество, а R — отношение на A. Тогда отношения $I_A\circ R$, R и $R\circ I_A$ эквивалентны; поэтому I_A является тождественным отношением на A, которое ведет себя подобно числу 1 по отношению к умножению чисел. Чтобы дополнить аналогию, желательно было бы иметь возможность писать $R^{-1}\circ R=I_A=R\circ R^{-1}$. Однако в общем случае этого делать нельзя. Для того чтобы иметь такую возможность, необходимо наложить дополнительные условия.

Замыкание отношений.

Понятие замыкания есть фундаментальным математическим понятием и используется в большинстве разделов математики. Чтобы проиллюстрировать это понятие, рассмотрим следующий пример.

Возьмем объект x_0 и процесс p, который порождает множество и определяет последовательность $x_1, x_2, \dots \dots, x_n, \dots$ такую, что

$$x_1\in p(x_0),$$

$$x_2\in p(x_1),$$

....

$$x_n\in p(x_{n-1}),$$

....

Множество, которое содержит все элементы всех последовательностей, которые могут быть выведены при помощью p, и начинающиеся с x_0, называется *замыканием процесса p* относительно x_0. Поэтому «ответ» будет содержаться в $p^n(x_0)$ при каком-то n. Однако мы не знаем заранее значения n. Более того, если мы возьмем произвольный элемент y из этого замыкания и выполним процесс p, начиная с y, то не получим ничего нового. Результат уже содержится в замыкании. Множество не может быть расширено таким путем (оно замкнуто).

Пример 5. Возьмем квадрат *S*, размеченный, как это показано на рис. 7, и определим процесс *r* следующим образом. Из заданного положения *S* процесс *r* порождает множество всех положений, получаемых в результате поворота по часовой стрелке на прямой угол.

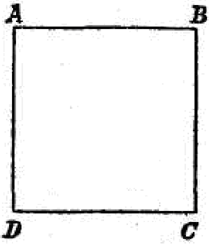


Рис. 7.

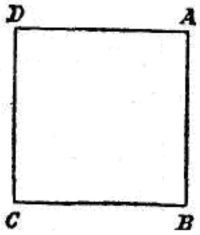
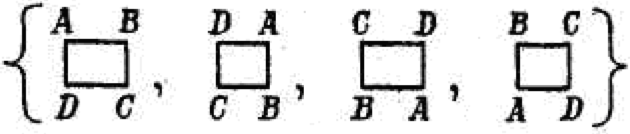


Рис. 8.

Таким образом, *r(S)* дает конфигурацию, которая изображена на рис. 8. После применения *r* четыре раза, мы вернемся к положению, с которого начали, и, следовательно, замыкание в данном случае есть множество из четырех позиций.



Рассмотрим теперь, что произойдет, если процесс определить с помощью отношения. (В действительности это всегда возможно, потому что мы можем определить подходящее отношение при помощи множества $\{(x, y) : y \in p(x)$, где p — исследуемый процесс).) Для построения замыкания отношения A достаточно иметь составные отношения $A, A^2, \ldots, A^n, \ldots$, которые затем комбинируются обычным теоретико-множественным путем.

Определение. *Транзитивным замыканием* (или просто *замыканием)* отношения *A* на множестве называется бесконечное объединение

$$\bigcup_{n=1}^{\infty} A^n = A \cup A^2 \cup A^3 \cup \ldots$$

Транзитивность замыкания отношения следует, очевидно, из его определение, однако слово «транзитивное» часто включают, чтобы подчеркнуть различие между этой и подобной ей операцией, которая в скором времени будет определена. Транзитивное замыкание отношения A обозначают A^+.

Пример 6.

1. Пусть R— отношение на N такое, что

$$R = \{(x, y): y = x+1\}.$$

Тогда

$$R^+ = \{(x, y): x < y\}.$$

2. Пусть σ — отношение на Q такое, что

$$\sigma = \{(x, y): x<y\}.$$

Тогда

$$\sigma^+ = \sigma.$$

3. Пусть ρ — отношение на Q такое, что

$$\rho = \{(x, y): x * y = 1\}.$$

Тогда

$$\rho^+\{(x, x): x \neq 0\} \square \rho.$$

4. Пусть L — множество станций Киевского метро и a, b и c - последовательные станции. Если отношение N на L определено как $N=\{(x, y)$: x есть следующей за y станцией), то (a, b) и $(b, c) \in N$ и $(a, a), (b, b), (c, c)$ и $(a, c) \in N^2$. Следовательно $N^+=U_L=L\times L$.

Из этих примеров легко видеть, что замыкание отношения в общем случае не является рефлексивным. Однако иногда удобно сделать его таким. Это можно легко сделать. Вначале мы примем допущение, что тождественное отношение на X, $I = \{(x, x): x\in X\}$ является нулевой степенью произвольного отношения на X. Таким образом, $A^0=I$ для любого A.

Определение. *Рефлексивным замыканием* A^* отношения A называют множество

$$A^* = \bigcup_{n=0}^{\infty} A^n$$

Замыкания отношений связаны между собой очевидным соотношением

$$A^* = A^+ \square И.$$

Пример 7. Используя отношения, которые определены в предыдущих примерах, получаем

$$R^*=\{(x, y): x\leq y\},\ \ \sigma^* = \{(x, y): x\leq y),$$

$$\rho^* = \rho^+ \cup \{(0, 0)\},\ \ N^*=N^+.$$

7.4. Типы и модели структур

Изучение типов и моделей структур будем иллюстрировать примерами из организации структур баз данных.

Напомним, что когда речь идет о представлении данных в ЭВМ, следует различать физическую и логическую организацию данных (9).

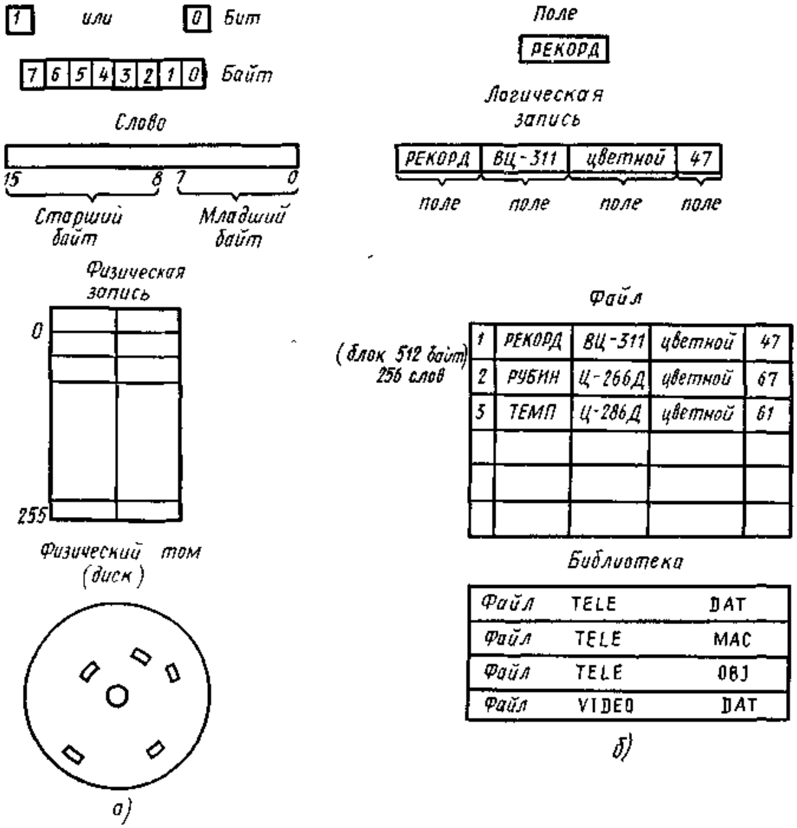


Рис. 9. Физическая *(а)* и логическая (б) организации данных

С другой стороны, когда говорится о данных применительно к какой-то конкретной задаче, следует помнить, что данные, закодированные в машине в виде нулей и единиц, несут в себе информацию о реальном объекте.
Если составляется алгоритм решения задачи, то программист пытается изобразить предметную область путем некоторого набора символов в виде структуры данных. При этом, естественно, его интересуют не все объекты, а лишь некоторые, которые имеют непосредственное отношение к решаемой задаче.
Совокупность таких объектов называется **предметной областью,** а сами объекты — **объектами предметной области.**
Например, при решении задачи проектирования РЭА объектами могут быть: материалы; предметы, элементы, агрегаты, изделия; станки, оборудование; технологические процессы.
Для каждого объекта предметной области должны быть описаны характеристики, являющиеся наиболее важными для решаемой задачи. Эти характеристики называют **атрибутами.** Каждый атрибут может принимать определенные значения.
Например, когда речь идет о телевизоре как объекте предметной области, в качестве атрибутов (в зависимости от решаемой задачи) могут быть использованы характеристики (марка, индекс, изображение, диаметр кинескопа, цена). Для определенного типа телевизора они примут конкретные значения. Например, «Рекорд», ВЦ-311, цветной, 47 см, 640 руб.
Для любого объекта существует совокупность его параметров, которую будем называть **записью.** Она, в свою очередь, содержит **поля.** Семейство записей образует **файл.** Несколько файлов образуют **библиотеку.** При этом можно отметить следующие соответствия между объектами предметной области и логическим представлением:

число объектов равно числу записей в файле;

число атрибутов, описывающих объект, равно числу полей в каждой записи.

Очевидно, что при описании атрибутов могут использоваться числовые величины, строки символов, значения проводимых измерений. Чтобы осуществить обработку этих данных с помощью ЭВМ, необходимо как-то учитывать тот факт, что внутренне представление информации в ЭВМ сводится в конечном счете к последовательности нулей и единиц.
Действительно, пусть в ходе некоторой операции в регистре центрального процессора оказалась двоичная величина 1101010001010011. Как интерпретировать ее?

Для того чтобы машина могла однозначно интерпретировать данные, необходимо снабдить ее некоторой дополнительной информацией, сообщающей о «типе» обрабатываемой величины. Поэтому данные, закладываемые в ЭВМ, классифицируются по типам. Понятие типа играет центральную роль в языках программирования высокого уровня, поскольку при программировании на ассемблере программист должен сам контролировать, с какими данными он имеет дело.

В языках высокого уровня тип данных задается путем явного описания в тексте программ. В различных языках, в общем случае, имеются различные типы данных, но основные классифицируются следующим образом: простые; структурированные; ссылочные.

Простые типы не обладают внутренней структурой и представляют собой конечный набор типов, называемых также базисными. К простым типам данных относятся: целый; плавающий; символьный; логический.

Целый и плавающий типы данных предназначены для представления числовых значений, символьный используется для образования текстов из символов (кода КОИ-7 или ДКОИ). Логический тип состоит из логических значений «истина» и «ложь» и применяется для выражения значений логических условий.

Данные любого простого типа характеризуются тем, что в любой момент времени каждому данному соответствует только одно значение.

Но для решения большинства задач с помощью программ требуется возможность использования таких типов данных, которые могут содержать много значений. Поэтому в рассмотрение кроме базовых типов данных вводят структурированные.

Структурированные типы данных предназначены для конструирования из конечного набора базисных типов сложных структурных данных.

К этим типам данных относят: массивы (одно- и многомерные), последовательности (файлы, стеки).

Под **массивом** понимают группировку набора данных идентичного типа. Массиву присваивается имя, обозначающее всю группу данных. Причем к каждому элементу группы возможен индивидуальный доступ с помощью целого индекса, указывающего позицию элемента в группе. В n-мерных массивах позиция элемента задается с помощью n индексов.

Ссылочный тип данных предназначен для обеспечения ссылок на другие данные и называется указателем. Этот тип применяется для динамического построения сложных структур данных.

Таким образом, из сказанного видно, что при написании программ на языке высокого уровня для того, чтобы ЭВМ правильно интерпретировало получаемую информацию, следует самым внимательным образом описывать типы используемых в программе данных.

Одним из основных способов структуризации данных является использование абстракций.

Абстракция предполагает, что несущественные детали должны быть опущены, а внимание должно быть сконцентрировано на основных общих свойствах множества объектов. Так, общее понятие ТЕЛЕВИЗОР — есть абстракция, отражающая множество наших представлений о конкретных телевизорах. Абстракция может быть многоуровневой т. е. объект абстракции одного уровня может рассматриваться как объект абстракции другого уровня и т. д. Таким образом, абстракция может использоваться для формирования нового типа из других типов. Например, БЫТОВАЯ РАДИОАППАРАТУРА определяется как абстракция типов ТЕЛЕВИЗОР, МАГНИТОФОН, ПРОИГРЫВАТЕЛЬ и т. д. Механизм абстракций широко используется при построении моделей данных.

Модели данных. Разработка САПР любых типов предполагает, что данными, заложенными в систему, будет пользоваться не один человек, а группы проектировщиков, имеющих доступ к системе. Возможность совместного использования одних и тех же данных требует создания в системе проектирования единой информационной базы — базы данных. Под **базой данных** будем понимать массив данных, хранимый в ЭВМ и предназначенный для совместного использования группой людей. Использование баз данных позволяет:

представлять в памяти ЭВМ сложные структуры информации, когда объектом хранения являются не только данные, но и структуры, в которые они организованы;

сокращать дублирование информации за счет структурирования данных, что приводит к экономии памяти на внешних носителях и повышает надежность информации;

повышать сохранность данных от несанкционированного доступа;

обеспечивать независимость прикладных программ от изменений данных;

повышать достоверность информации и сокращать затраты на обслуживание систем.

Если в случае создания отдельной программы пользователь может сам проводить подготовку данных, определять их типы и предъявлять ЭВМ, то при создании базы данных сразу для нескольких пользователей требования к организации данных в ЭВМ несколько усложняются.

Для того чтобы некоторую предметную область представить в базе данных, прежде всего необходимо выделить те понятия, которые интересны с точки зрения всех потенциальных пользователей базы. В силу этих причин создание базы данных требует широкого использования механизма абстрагирования. На абстрактном уровне проводится классификация понятий и объектов предметной области. Сам процесс абстрагирования информации о предметной области, требуемой для создания базы данных, называют **логическим проектированием базы данных (проектирование данных).** При логическом проектировании физические характеристики ЭВМ и особенности программ отдельных пользователей не учитываются. Логическое проектирование осуществляется в рамках некоторой заранее принятой модели абстрагирования, которую называют **модель данных.** Точнее говоря, под ней понимают основные понятия и способы, используемые при выполнении абстрагирования.

Описание предметной области в терминах некоторой модели данных называют **концептуальной схемой,** или **моделью.** Связь между предметной областью и концептуальной моделью может быть отражена определенным образом (рис. 10).

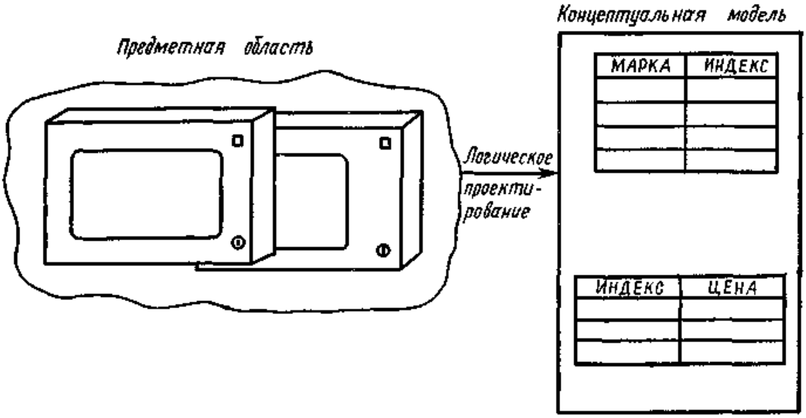


Рис. 10. Связь между предметной областью и концептуальной моделью

Концептуальная модель характеризует логическое представление данных и является тем мостиком, посредством которого связываются возможности физического представления данных в ЭВМ и в то же время их независимого использования в различных прикладных программах.

Концептуальную модель, описанную и скорректированную с точки зрения представления данных в памяти ЭВМ, называют **внутренней моделью.** Поскольку база данных (БД) рассчитана на то, что каждый пользователь имеет доступ к определенным данным, концептуальная модель должна отражать еще один уровень логического представления данных для каждого конкретного пользователя. Такое представление концептуальной модели, непосредственно связанное с применением, называют **внешней моделью** (схемой). Так как каждый отдельный пользователь в большинстве случаев имеет отношение лишь к небольшой, вполне определенной части данных, хранимых в базе, то внешняя модель у него может быть своя. Таким образом, архитектура БД может быть упрощенно представлена, как на рис. 11.

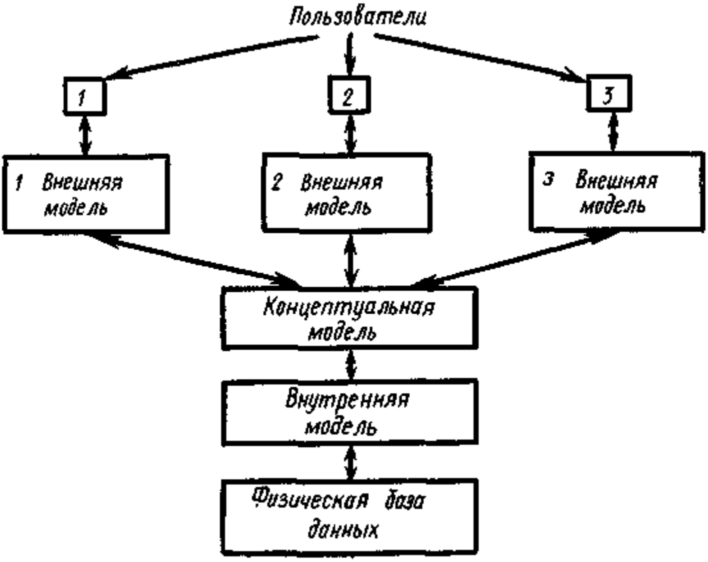


Рис. 11. Архитектура базы данных ,

Можно отметить, что на концептуальном уровне осуществляется формализованное описание информационной базы в терминах конкретной СУБД. Это означает, что на концептуальном уровне одна и та же информационная база описывается средствами различных СУБД. Поэтому при проектировании базы вводится еще один уровень описания. Модель этого уровня должна выражать информацию о предметной области в виде, независимом от конкретно используемой СУБД.

Этот уровень называют информационно-логическим или инфологическим (рис. 12).

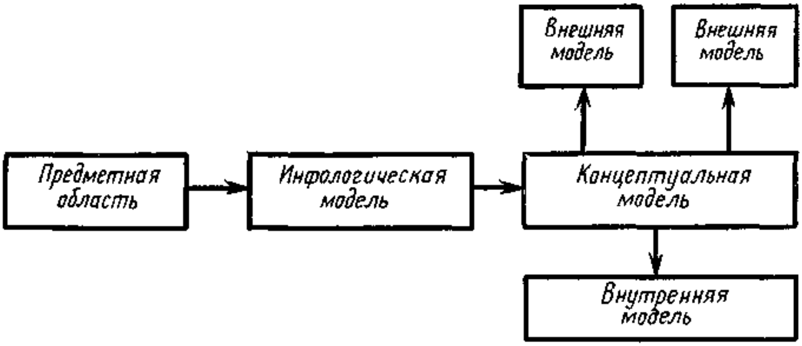


Рис. 12. Место инфологической модели в процессе разработки моделей данных

Инфологическая модель служит для формализации представления о предметной области до привязки к конкретным средствам реализации баз знаний.

Таким образом, при моделировании данных возникает проблема многоуровневого представления данных. Поэтому при проектировании БД уровни представления данных подразделяют на инфологический и даталогический. Причем, если инфологические модели используются для построения семантических (смысловых) моделей, отражающий информационное содержание конкретной предметной области, **даталогические модели** служат для реализации информационных баз в определенной вычислительной среде.

Целостность данных. При создании БД всегда следует учитывать логические ограничения на значения данных и их соотношения. Они обычно представляют собой условия, при которых имеют смысл те или иные данные. Так, например, если будем рассматривать значение стоимость детали, то оно не может превышать значение СТОИМОСТИ ИЗДЕЛИЯ, в которое детали входят составной частью. Или, например, значение ДАТА ИЗГОТОВЛЕНИЯ телевизора не может принять значение 1912 г.

Логические ограничения, накладываемые на данные, рассматриваются как свойства, присущие данным и обеспечивающие адекватное отображение предметной области в БД. Если эти ограничения записать в БД, то их можно использовать для контроля целостности содержимого БД. Отсюда возникает понятие **целостности данных,** т. е. данные, хранимые в БД, не должны противоречить заданным логическим ограничениям, которые называют **ограничениями целостности.** Они обычно задаются для множества объектов. Их можно разбить на два основных типа: внутренние и явные.

Внутренние ограничения обусловлены самой структурой принятой модели данных. Так, например, в **реляционных моделях** данных дубликаты записей не размещаются. Или в **иерархической**модели данных связи ограничены древовидной иерархической структурой.

В некоторых моделях данных вводятся ограничения, которые описываются в явном виде, с помощью специальных конструкций языка описания данных. К явным ограничениям целостности можно отнести ограничения на значение атрибутов объекта. Естественно, что ограничения в явном виде задаются не только для атрибутов, но и для типов объектов (сущностей) и связей. Так, например, если рассматривать сущность СТУДЕНТ, то может быть ограничено число студентов, обучающихся в одной группе. Для того чтобы уточнить, какие бывают ограничения на связи, необходимо рассмотреть основные типы связей.

Связь один к одному (1:1). Она определяет такой вид связи между двумя типами сущностей *А* и *В,* при котором каждому экземпляру сущности *А* соответствует только один *В,* и наоборот. Например, связь студент курса — номер зачетной книжки (рис. 6.8).

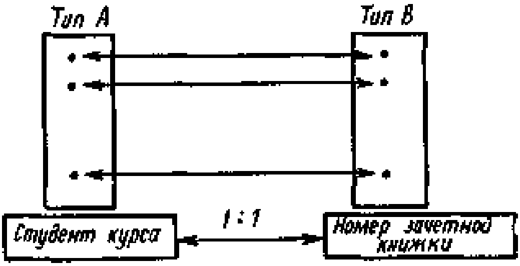


Рис. 13. Связь один к одному (1:)

Здесь каждый экземпляр одного типа сущности однозначно определяет другой.

Связь один ко многим (1:М). Соответствует случаю, когда для двух типов сущностей *А* и *В,* одному экземпляру сущности *А* соответствует несколько (0, 1, 2,...,*М)* экземпляров сущности *В.* Однако каждому *В* соответствует только один экземпляр сущности *А,* например, связь группа — фамилия, имя, отчество студента (рис. 14).

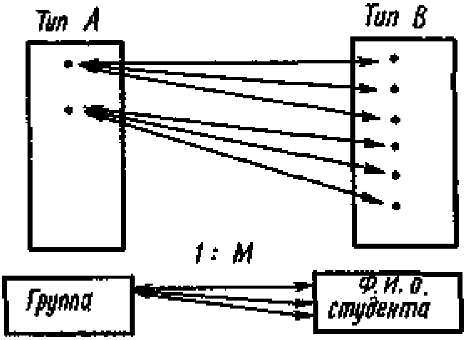


Рис. 14. Связь один ко многим (1 :М)

Связь многие к одному (М: 1). Является вариантом связи, обратных к связи 1 : *М*, т. е. в этом случае многим экземплярам сущности типа *А* соответствует только один *В.* Например, Ф. И. О. студента — группа (рис. 15).

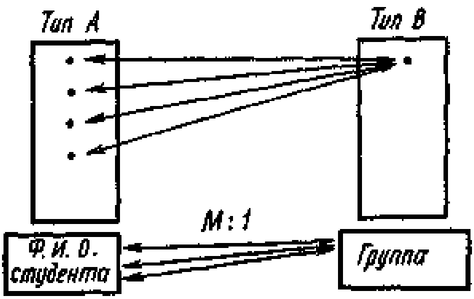


Рис. 15. Связь многих к одному (М:1)

Связи типа 1:1, 1 *:М, М:* 1 называют функциональными.

Связь многие ко многим (М:М). Соответствует случаю, когда каждому экземпляру сущности *А* может соответствовать несколько экземпляров сущности *В,* и наоборот. Например, телевизор — резистор (рис. 16).

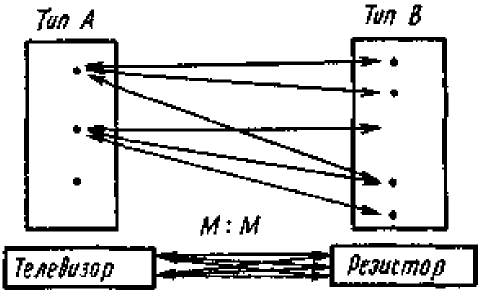


Рис. 16. Связь многих ко многим (М: М)

До сих пор рассматривались двусторонние связи между типами сущностей. Иногда целесообразно рассматривать одностороннюю связь от сущности *А* к сущности *В* (такие связи называют ассоциациями). Ассоциации, в свою очередь, подразделяют на три типа: простая, сложная и условная.

Ассоциация простая (тип 1). Соответствует случаю, когда экземпляр сущности *А* определяет один и только один экземпляр сущности *В,* т. е. идентификация экземпляра сущности *В* является уникальной. Например, связь типа группа —фамилия старосты группы (рис. 17).

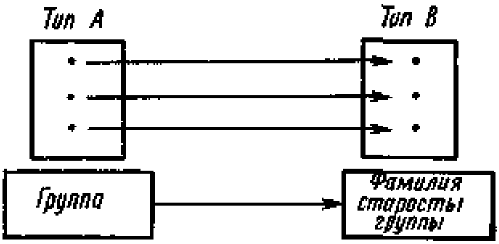


Рис. 17. Ассоциация простая по типу *I*

Естественно, что в частном случае может оказаться, что в разных группах могут оказаться старосты с одинаковыми фамилиями, и поэтому обратная связь (от *В* к *А)* здесь не рассматриваются.

Сложная ассоциация (тип М). Соответствует случаю, когда каждый экземпляр сущности *А* определяет несколько (нуль, один, два и **т** д.) экземпляров сущности *В.* При этом идентификация экземпляра типа *В* не обязательно является уникальной: например, связь между сущностями УЗЛЫ ИЗДЕЛИЯ и ПОСТАВЩИК (рис. 18).

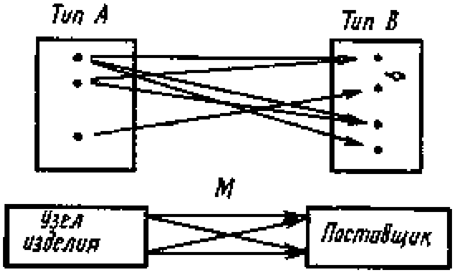


Рис. 18. Сложная ассоциация по типу *М*

Условная ассоциация (тип С). Описывает связь, когда для двух типов сущностей *А* и *В* может не существовать экземпляра типа сущности *В,* но если существуют, то он относится к единственному экземпляру сущности *А.* Например, связь между сущностями служащий — дата увольнения (рис. 19).

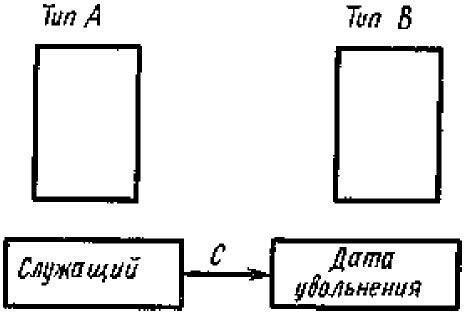


Рис. 19. Условная ассоциация по типу *С*

Приведенные типы связей не исчерпывают все множество видов ограничений, которые могут иметь место. Поэтому контроль выполнения ограничений в явной форме в конкретных моделях данных представляет собой сложную задачу и требует от СУБД целого набора актов доступа к БД и использования средств реляционного исчисления.

Операции над данными. В процессе обработки данных на ЭВМ приходится осуществлять над ними некоторые операции, которые необходимы для заданной конкретной операции.

Селекцию можно осуществить на основе использования логической позиции данного, его значений и связей между ними.

Использование для селекции **логической позиции данного** базируется на определенной упорядоченности данных в памяти системы.

Упорядоченность данных позволяет использовать для селекции такие понятия, как первый элемент, последний, текущий, *n*-й. Относительно текущего элемента можно указать предыдущий элемент и последующий. Этот тип селекции называют также селекцией **посредством текущий.**

Использование для селекции **значений данных** требует задания в явном виде значений атрибутов, являющихся критериями селекции. Простое условие определяется на одном атрибуте и одном его значении и обычно имеет вид

Имя атрибута — оператор условия — значения атрибута.

Под оператором-условием понимается один из арифметических операторов (= , ≠, <;, >, ≤, ≥). Приведем пример простого условия селекции по значению данных:

средний балл студента ≥4,0.

Это условие означает, что из всего списка студентов следует отобрать студентов, имеющих средний балл не меньше 4,0.

Более сложные критерии селекции могут быть заданы выражениями, построенными на простых условиях с помощью логических операторов (И, ИЛИ, И). Например, условие

средний балл ≥4,0 и физика =5

означает селекцию тех студентов, у которых средний балл не менее 4,0 и которые сдали физику на 5.

Селекцию данных может производить не только по логической позиции или по значению, но и в соответствии с логическими связями между ними. Например, при наличии связи типа РАБОТАЕТ между сущностями ПРЕПОДАВАТЕЛЬ и КАФЕДРА можно осуществить селекцию всех лиц, работающих преподавателями на кафедрах университета и всех кафедр, имеющихся в университете.

Селекция такого рода называется селекцией по связности данных.

По характеру производимого действия над данными различают следующие виды операций:

1. Идентификация данного и определение его позиции в базе;
2. Выборка (получение требуемых данных из базы);
3. Включение (запись новых данных в базу);
4. Удаление данных;
5. Обновление (модификация) данных.

Эти действия могут быть применены как к атрибутам, так и к типам сущностей и связям.

По характеру способа получения результата различают навигационные и спецификационные операции. Если результат представлен единичным объектом (значением атрибута, реализацией сущности или связи), полученным при прохождении по логическому пути (т. е. при

навигации) в структуре БД, то соответствующую операцию называют **навигационной.**

Навигационные операции всегда предполагают селекцию посредством текущей. Отсюда возникает задача манипулирования текущими. Для определения текущих могут использоваться специальные индексы, которые могут совпадать с ключами.

Если в операции определяются только требования к результату, но не задается способ его получения, то операции называется **спецификационными. В** спецификационных операциях текущие на пользовательском уровне не видны. Они могут быть определены в терминах операций теории множеств: объединение, пересечение, разность. Поэтому результату спецификационной операции в общем случае соответствует некоторое множество объектов.

В процессе обработки данных возникает необходимость использования более сложных действий над данными, чем позволяют навигационные и спецификационные операции. Например, при выполнении операций селекции необходимо автоматически поддерживать целостность данных. Этого можно добиться путем создания специальной программы, которая будет осуществлять проверку большого количества данных. Эта программа, естественно, потребует использования локальных операций манипулирования над данными. С другой стороны, она может рассматриваться как глобальная операция проверки целостности данных.

Такие глобальные операции называют **процедурами** базы **данных.** Другими словами — это обобщенные операции изменения состояния базы данных. Процедура базы данных рассматривается как единая макрооперация, при выполнении которой ни одна другая процедура или программа не может обратиться к данным. Поэтому такие процедуры или операции называют еще транзакциями.

Один из видов процедур БД — вычисление значений, которые непосредственно в ней не хранятся, например вычисление сумм, подсчет числа экземпляров, определение минимума, максимума. Процедуры этого вида называют **функциями агрегации.**

Важный вид процедур БД — вычисление значений атрибута, например, вычисление возраста студента по дате его рождения и текущей календарной дате. Процедура базы данных, обеспечивающая вычисление значения какого-либо атрибута, называется **виртуальным атрибутом.** Для пользователя он представляется как обычный атрибут, обладающий теми же свойствами, что и любой другой атрибут.

Процедуры БД применяются также для контроля целостности, контроля доступа к данным, расширения языка данных операциями, первоначально в них не предусмотренными.

Особый вид процедур БД составляют процедуры, активизирующиеся при определенных условиях и выполняющих одну или более операций включения, удаления или модификации. Такие процедуры называют запускаемыми включением, удалением, обновлением.

Анализируя различия между процедурами БД и операциями, можно отметить следующее:

1. Процедуры при своем выполнении могут захватывать обширные области данных;
2. Процедуры могут реализовать широкий круг действий;
3. Вызовы процедур не выполняются пользователем;
4. Процедуры обычно описываются в схеме данных, в то время как операции включаются в пользовательскую программу.

Появление и использование БД привело к необходимости создания специальных программ, обеспечивающих управление данными, хранящимися в базе, т. е. систем управления базой данных (СУБД).

Таким образом, *СУБД —это специальный пакет программ, посредством которого реализуется централизованное управление БД и обеспечивается доступ к данным.* База данных вместе с системой управления его являются составными частями банка данных.

Развитие СУБД привело к появлению ряда языков описания состояния предметной области, которое в СУБД интерпретируется состоянием БД. Множество допустимых состояний БД определяется схемой БД, задаваемой на языке **определения данных (ЯОД).**

Изменение состояния БД и извлечение данных из базы для последующей обработки обеспечивается средствами **языка манипулирования данными (ЯМД).**

Описание структуры данного некоторого типа на формализованном языке называют схемой этого данного. Язык описания данных — это язык высокого уровня, предназначенный для схемы БД. С его помощью описываются типы данных, хранящихся в базе, и их структура.

Язык манипулирования данными (ЯМД) используется при написании прикладных программ, обращающихся к данным, хранящимся в базе. Основной функцией ЯМД является выполнение операций ввода-вывода при обработке информационной базы. Собственно говоря, совокупность ЯОД и ЯМД и определяет ***<u>модель данных</u>*, понимаемую как совокупность методов и средств определения логической структуры БД.**

В качестве моделей данных, наиболее широко используемых при создании БД, обычно применяются: реляционная; сетевая; иерархическая.

7.5. Реляционная модель организации структур

Основные понятия. В основе реляционной модели организации структур (РМОС) лежит математическая теория отношений. Этим определяется и название модели (RELATION - отношение). Отношение служит средством структуризации сущностей, которые в нашем рассмотрении будут представляться данными.

Таким образом, представив n-местное отношение в виде таблицы, тем самым определенным образом структурируем данные. Поэтому, естественно, подобные преобразования называют реляционной структурой или реляционным типом. Массив данных, представленный набором реляционных структур, образует реляционную БД, и схема реляционной БД будет представлена набором схем отношений:

$$R_1(A^1{}_1, A^1{}_2,\ldots, A^1{}_k);$$
$$R_2(A^2{}_1, A^2{}_2,\ldots, A^2{}_l);$$
$$\ldots\ldots\ldots\ldots\ldots\ldots$$
$$R_m(A^m{}_1, A^m{}_2,\ldots,A^m{}_n),$$

где A_i^j — имя атрибута, R_j— имя отношения.

Ограничение модели. Число ограничений в реляционной модели невелико, что обеспечивает достаточную свободу в выборе представления типов связей и сущностей. Основным ограничением является невозможность представления в отношении дубликатов строк. Это ограничение позволяет уточнить понятие ключа отношения. В реляционной модели данных (РМД) ключ определяется как подмножество атрибутов, которые позволяют однозначно идентифицировать кортеж. Так как дубликаты строк в отношении запрещаются, то это означает, что каждое отношение имеет, по крайней мере, один ключ (состоящий из всех атрибутов). Отношение может иметь и несколько ключей, называемых **возможными ключами**. Один из возможных ключей выбирается в качестве **первичного.** Следует иметь в виду, что первичный ключ не разрешается обновлять и никакой из его компонентов не может принимать значения «неопределено».

Второе ограничение модели состоит в том, что порядок столбцов в таблице является значимым. Пренебрегать упорядочением столбцов можно только в том случае, если каждому столбцу присвоено уникальное имя.

На значения атрибутов в модели можно задавать ограничения в явном виде. Большинство явных ограничений, которые встречаются на практике, это ограничения на зависимости между атрибутами. Явное задание ограничений обеспечивает возможность самостоятельного исследования зависимостей между атрибутами.

Зависимость отражает тот факт, что один объект зависит от другого. В РБД в качестве таких объектов рассматриваются либо аргументы, либо кортежи. Одним из основных типов зависимостей, рассматриваемых в реляционных БД, являются **функциональные зависимости.**

Рассмотрим формальное определение функциональной зависимости, принимая следующие обозначения. Большими буквами $A,B,C,\dots$ обозначим одиночные атрибуты, X, Y, Z — множество атрибутов, а a, b, c,... и x, y, z,... — соответствующие им значение. Большие буквы R, S будем применять для обозначения отношений.

Предположим, что существует некоторое **универсальное отношение** U, в котором каждый атрибут имеет уникальное имя. При этом будем считать, что множество атрибутов любого другого отношения представляет собой некоторое подмножество атрибутов универсального отношения U.

Пусть A и B атрибуты отношения R. Тогда говорят, что атрибут B отношения R функционально зависит от атрибута A, если в каждый момент времени каждому значению a соответствует не более одного значения b. Функциональную зависимость f атрибута B от атрибута A обозначают: $f{:}A\rightarrow B$.

Эту зависимость f можно также представить множеством упорядоченных пар $\{<a, b>|a\in A, b\in B\}$, в которых каждому значению a соответствует только одно значение b. При этом говорят, что B **функционально зависит** (или просто зависит) от A, а A **функционально определяет** (или **определяет**) B.

Если существует единственная функциональная зависимость B от A, то ее обозначают просто $A\rightarrow B$. В случае отсутствия между ними функциональной зависимости вводят обозначения $A\neq B$.

Если $A\rightarrow B$ и одновременно $B\rightarrow A$, то между A и B существует взаимно-однозначное соответствие, которое записывается как $A\leftrightarrow B$.

Пусть: $f{:}A_1, A_2,\dots ,A_n\rightarrow B$ и $g{:}A_1, A_2,\dots, A_m\rightarrow B$, где $m<n$. Так как атрибуты $A_1, A_2,\dots, A_m$ функционально определяют B, то $A_{m+1}, A_{m+2},\dots ,A_n$ называют посторонними в f. В этом случае B неполно зависит от $A_1, A_2,\dots ,A_n$. Если для данного f не существует g с вышеуказанными свойствами, т.е. левая часть f не содержит посторонних атрибутов, то говорят, что реализуется полная функциональная зависимость.

Пусть есть множество атрибутов $A_1, A_2,\dots ,A_n$ отношения R, а также множество F функциональных зависимостей $X\rightarrow Y$, где X и Y — подмножества атрибутов множества $A_1, A_2,\dots ,A_n$. Тогда из функциональных зависимостей, которые входят в F, могут быть выведены другие функциональные зависимости, присущие отношению R.

Обозначим через F^+ замыкание множества функциональных зависимостей F, т.е. полное множество зависимостей, которую можно получить из F. Множество зависимостей F^+ можно построить из F на основе следующих правил вывода функциональных зависимостей (ФЗ):

1. ФЗ1 (свойство рефлексивности).
2. ФЗ2 (свойство пополнения).
3. ФЗ3 (свойство транзитивности).

Это полный набор правил, т.е. он позволяет по заданному множеству F определить полное множество функциональных зависимостей F^+, присущий рассматриваемой схеме отношений $R(A_1, A_2, \ldots, A_n)$.

Рассмотрим эти правила более подробно.

Правило ФЗ1 (свойство рефлективности). Если $X \subseteq U$, $Y \subseteq U$ и $Y \subseteq X$, то имеет место функциональная зависимость $X \rightarrow Y$. Например, задано отношения $U(A_1, A_2, A_3, A_4, A_5)$.

Рассмотрим два множества атрибутов:

$$X = \{A_1, A_2, A_3, A_4\} \text{ и } Y = \{A_1, A_4\}.$$

Исходя из свойства транзитивности, можно сказать, что существует функциональная зависимость $X \rightarrow Y \in F^+$. Данное правило говорит о том, что, имея исходную функциональную зависимость $A_iA_jA_kA_n \rightarrow A_iA_j$, можно в состав множества атрибутов левой части выражения вводить любые атрибуты из множества U. При этом функциональная зависимость будет сохраняться. Можно также в правую часть включать атрибуты которые уже расположенны в левой части.

Правило ФЗ2 (свойство пополнения). Если $X \subseteq U$, $Y \subseteq U$, $Z \subseteq U$ и имеет место функциональная зависимость $X \rightarrow Y$, то $X \cup Z \rightarrow Y \cup Z$. В отличие от правила ФЗ1 данное говорит о том, что для его применения не существенно выполнение условий $Y \subseteq X$. Т.е. любые атрибуты из множества U можно одновременно подставлять в левую и правую части выражения функциональной зависимости F.

Например, имеется универсальное отношение $U(A_1, A_2, A_3, A_4, A_5)$ и заданы наборы атрибутов $X = \{A_1, A_3\}$, $Y=\{A_2, A_4\}$, $Z=\{A_5\}$. Тогда из условия, что существует функциональная зависимость $X \rightarrow Y$:

$$\{A_1, A_3\} \rightarrow \{A_2, A_4\}$$

следует, что имеет место зависимость

$$\{A_1, A_3, A_5\} \rightarrow \{A_2, A_4, A_5\}.$$

Правило ФЗ3 (свойство транзитивности). Если $X \subseteq U$, $Y \subseteq U$, $Z \subseteq U$ и имеют место зависимости $X \rightarrow Y$ и $Y \rightarrow Z$, то $X \rightarrow Z$. Например, имеются подмножества атрибутов $X=\{A_1, A_3\}$, $Y=\{A_2, A_4\}$, $Z=\{A_5\}$. Тогда из условия существования зависимостей $\{A_1, A_3\} \rightarrow \{A_2, A_4\}$, $\{A_2, A_4\} \rightarrow \{A_5\}$ следует, что имеет место зависимость $\{A_1, A_3\} \rightarrow \{A_5\}$.

Кроме этих правил часто используют дополнительные правила следствия Ф31, Ф32 и Ф33.

Правило Ф34 (свойство расширения). Если $X \subseteq U$, $Y \subseteq U$ и задана функциональная зависимость $X \to Y$, то тогда для любого $Z \subseteq U$ имеет место функциональная зависимость $X \cup Z \to Y$.

Правило Ф35 (свойство продолжения). Если $X \subseteq U$, $Y \subseteq U$, $W \subseteq U$, $Z \subseteq U$ и задана функциональная зависимость $X \to Y$, то для любых $W \subseteq Z$ имеет место зависимость $X \cup Z \to Y \cup W$.

Правило Ф36 (свойство псевдотранзитивности). Если $X \subseteq U$, $Y \subseteq U$, $W \subseteq U$, $Z \subseteq U$ и заданы функциональные зависимости $X \to Y$, $Y \cup W \to Z$, то имеет место функциональная зависимость $X \cup W \to Z$.

Правило Ф37 (свойство аддитивности). Если $X \subseteq U$, $Y \subseteq U$, $Z \subseteq U$ и заданы функциональные зависимости $X \to Y$, $X \to Z$, то имеет место функциональная зависимость $X \to Y \cup Z$.

Правило Ф38 (свойство декомпозиции). Если $X \subseteq U$, $Y \subseteq U$, $Z \subseteq U$ и при этом $Z \subseteq Y$ и задана функциональная зависимость $X \to Y$, то будет иметь место функциональная зависимость $X \to Z$.

Очевидно, что даже при небольшом числе зависимостей F число функциональных F^+ должно быть довольно большое.

Функциональная зависимость играет важную роль в моделировании данных. Вместе с тем можно отметить, что это отнюдь не общий вид зависимости. Важную роль играет также многозначная зависимость.

Многозначные зависимости. Рассмотрим отношение $R(X, Y, Z)$, где X, Y, Z — множество атрибутов. Кортеж отношения $R(X, Y, Z)$ обозначим через $<x, y, z>$. Если в отношении $R(X, Y, Z)$ присутствуют кортежи $<x, y, z>$; $<x, y', z>$;... $(x, y, z'>$; $<x, y', z'>$, то говорят, что существует многозначная зависимость атрибутов Y и Z от X. Эта зависимость обозначается $X \to\to Y$.

Пример. Рассмотрим таблицу, в которой указанные школьники-победители олимпиад по предметам:

№ п/п	Ф. И. О. школьника	Школа №	Предмет
1	Давыдов В. А.	154	Математика
2	Давыдов В. А.	154	Физика
3	Кулешов К. Г.	820	Биология
4	Жукова А. М.	154	Математика
5	Жукова А. М.	154	Физика

В нашем примере имеют место две многозначные зависимости:

ШКОЛА №→→ПРЕДМЕТ (например, № 154→ → (математика, физика), № 820→→ (биология)) и ШКОЛА № →→Ф. И. ШКОЛЬНИКА

Другими словами, *X* многозначно определяет *Y*, если и только если множество $Y=\{y|(x, y, z)\in R\}$ определяется только *X*.

Многозначную зависимость определяют путем следующей проверки.

Если для двух кортежей *t* и *s* отношения *R(X, Y, Z)* справедливо первое условие:

$$t[X]= s\ [X]$$

т.е. *t* и *s* совпадают по значениям атрибутов *X* и существует третий кортеж *u,* такой, что выполняется второе условие:

$$u[X, Y]=t[X, Y]; \quad u[Z]=s[Z],$$

то существует многозначная зависимость.

Пример. Рассмотрим отношение ПОБЕДИТЕЛИ ОЛИМПИАДЫ (Ф. И. О. ШКОЛЬНИКА, ШКОЛА №, ПРЕДМЕТ).

Проверим на основании вышеприведенного условия многозначную зависимость ШКОЛА №→→ПРЕДМЕТ

1) *t* [№154]= *s* [№154].

Можно отметить, что существует четыре кортежа, у которых значение атрибута ШКОЛА И (№ 154) совпадает, т.е. первое условие выполняется. В качестве *t* и *s* возьмем, например, кортежи

t [Давыдов В. А. № 154 математика],

s [Жукова А. М. № 154 физика].

Рассмотрим, существует ли такой кортеж *u,* для которого будет выполняться и второе условие, определяющнн многозначную зависимость.

Можно выделить кортежи, для которых выполняется соотношение

u[№ 154, математика] =*t* [№ 154, математика]; это следующие кортежи:

t [Давыдов В. А., № 154, математика],

u [Жукова А. М., № 154, математика],

т.е. первому уравнению во втором условии - кортежи удовлетворяют. Анализируя таблицу, заметим, что для того, чтобы существовала многозначная зависимость ШКОЛА №→→ПРЕДМЕТ, должно выполняться условие

u[Жукова А. М.]=*s* [Жукова А. М.],

что действительно имеет место.

Другой способ проверки многозначной зависимости $X\to\to Y$ на отношение *R(X, Y, Z)* может быть осуществлен на основе проверки существования кортежа $<x, y, z>$ при условии, что существуют кортежи $<x, y, z'>$ и $<x, y', z>$.

Следует иметь в виду, что в общем случае формальная проверка должна выполняться на множестве всех возможных экземпляров кортежей отношения.

Правила вывода многозначных зависимостей сходны с правилами вывода функциональных зависимостей:

МЗ1 (дополнение). Если $X \subseteq U$, $Y \subseteq U$, $X \to\to Y$, то имеет место многозначительная зависимость $X \to\to U - X - Y$.

МЗ2 (присоединение). Если $X \subseteq U$, $Y \subseteq U$, $Z \subseteq W$, $X \to\to Y$, то имеет место зависимость $XW \to\to YZ$.

МЗ3 (транзитивность). Если $X \subseteq U$, $Y \subseteq U$, $X \to\to Y$, $Y \to\to Z$, то $X \to\to Z — Y$.

Существуют также правила вывода для совокупности ФЗ и МЗ:

ФМ1. Если $X \subseteq U$, $Y \subseteq U$, $X \to Y$, то $X \to\to Y$.

ФМ2. Если $X \subseteq U$, $Y \subseteq U$, $Z \subseteq U$, $W \subseteq U$ и W не пересекается с Y (т.е. $W \mid \; Y=0$), и $X \to\to Y$, $W \to Z$, то имеет место зависимость $X \to\to Z$.

ФМ3. Если $X \to\to Y$, $XY \to Z$, то $X \to Z — Y$.

Функциональная и многозначная зависимости являются свойствами, существующими между двумя атрибутами (множествами). С их помощью осуществляют декомпозиции (разбиение) отношения на два или восстанавливают исходное отношение, соединяя два отношения (рис. 20).

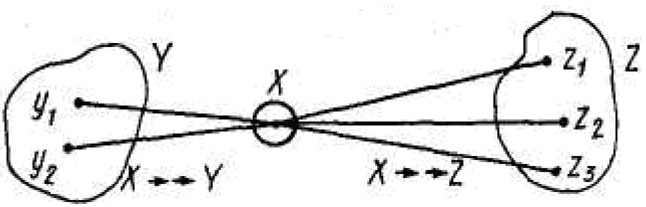


Рис. 20. Пример изображения многозначных зависимостей

Одним из основных вопросов любой организации данных есть вопросы сохранения жизнеспособности прикладных программ при изменениях в базе данных для того, чтобы избежать трудностей поддержания базы данных и назначения ключей для каждого отношения. Рассмотрим отношение $R\{A_1, A_2, \ldots, A_n\}$. Возможный ключ k отношения R — это комбинация атрибутов (возможно, состоящих из одного атрибута), обладающих следующими свойствами.

1. В каждом кортеже отношения R величина k единственным образом определяет этот кортеж.

2. Не существует атрибута в возможном ключе *k*, который мог бы быть удален без нарушения свойства 1.

Всегда существует, по крайней мере, один возможный ключ, т.е. комбинация всех атрибутов *R* удовлетворяет свойству 1.

Если в отношении *R* имеется несколько возможных ключей, то один из них выбирается в качестве первичного.

Атрибут A_i отношения *R* называется также первичным, если он входит в состав любого ключа (возможного или первичного) отношения.

Нормальные формы схем отношений. Рассмотрим четыре уровня нормализации схем отношений и соответственно четыре нормальные формы отношений: 1НФ, 2НФ, 3НФ, 4НФ. Их взаимное отношение можно представить в виде, представленном на рис. 21.

Из рис. 21 следует, что отношения являются как бы вложенными друг в друга по возрастанию номеров.

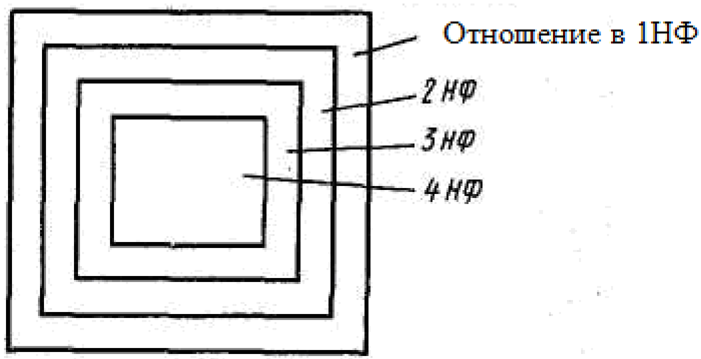


Рис. 21. Взаимное отношение нормальных форм

Так, например, если отношение находится в 4НФ, то оно будет удовлетворять и условиям 3НФ, 2НФ, 1НФ.

Отношение находится в **первой нормальной форме** (1НФ), если каждый атрибут отношения есть простым (атомарным) атрибутом, т.е. отсутствуют составные. Рассмотрим схему отношения:

АВТОМОБИЛЬ (МОДЕЛЬ, МАРКА, МОЩНОСТЬ, СТОИМОСТЬ, ИЗГОТОВИТЕЛЬ (НАЗВАНИЕ ЗАВОДА, ГОРОД)).

В этом случае атрибут ИЗГОТОВИТЕЛЬ составной. Для приведения к 1НФ отношения необходимо избавиться от составного отношения - ИЗГОТОВИТЕЛЬ. Этого можно добиться, рассматривая вместо составного атрибута его составляющие:

АВТОМОБИЛЬ (МОДЕЛЬ, МАРКА МОЩНОСТЬ, СТОИМОСТЬ, НАЗВАНИЕ ЗАВОДА ИЗГОТОВИТЕЛЯ, ГОРОД).

Приведение отношения к 1НФ достаточно для реализации языков запросов.

Чтобы рассмотреть 2НФ, введем понятие **полной зависимости**. Пусть X и Y — подмножества атрибутов отношения R и $X \to Y$. Если Y функционально не зависит от любого подмножества A множества X (причем A не совпадает с X), то Y называется полностью зависимым от X в R. Тогда говорят, что отношение R находится **во второй нормальной форме** (2НФ), если оно нормализовано, т.е. находится в первой нормальной форме, и каждый непервичный атрибут полностью зависит от первичного ключа.

Отношение R находится **в третий нормальной форме** (3НФ), если оно находится во второй нормальной форме и каждый непервичный атрибут в отношении R не содержит транзитивных зависимостей от первичного ключа. Транзитивная зависимость наблюдается в R, если существует такой атрибут A, что $X \to Y$ и $Y \to A$, $Y \leftrightarrow X$, где X — ключ.

Можно заметить, что 2НФ и 3НФ накладывают ограничения на зависимости только в связи с непервичными атрибутами. Поэтому в рассмотрение вводят **усиленную третью нормальную форму** (или нормальную форму Бойса-Кодда).

Отношение R находится в усиленной третий нормальной форме, если для всех зависимостей $X \to A$, когда A не принадлежит X, X является возможным ключом отношения R. Обычно атрибут, от которого функционально полно зависит другой, называют **детерминантой**. Поэтому говорят, что отношение R находится в усиленной третий нормализованной форме, если все детерминанты являются ключами.

Можно отметить следующее отличие этих нормальных форм по информационному содержанию. Вторая нормальная форма более информативна, чем первая, а третья более информативна, чем вторая.

Поэтому 3НФ и усиленная 3НФ больше понятны пользователю и приспособленны к реализации.

Существует также нормальная форма, которая учитывает многозначные зависимости, ее называют **четвертой нормальной формой** (4НФ). Отношение R находится в четвертой нормальной форме, если каждый раз, когда существует многозначная зависимость $X \to\to Y$ (где $Y \neq \varnothing$, $Y \not\subset X$ и XY состоит не из всех атрибутов R), также существует зависимость $X \to A$ для любого атрибута A в R.

Получение отношений нормальной формы достигается декомпозицией их схем. Под декомпозицией схемы отношения $R=\{A_1, A_2,..., A_n\}$ понимается замена схемы совокупностью отдельных схем $\rho = \{ R_1, R_2,...,R_k\}$ таких, что $R_1 \cup R_2 \cup ... \cup R_n=R=\{A_1, A_2,..., A_n\}$.

При этом не накладывается никаких ограничений на пересечение любых двух элементов совокупности ρ.

Для получения нужных данных из базы разрабатываются специальные языки манипулирования данными, обеспечивающие выполнение необходимых операций. Для разработки и исследования языков манипулирования данными используют математический аппарат, основанный на следующих трех подходах:

реляцинная алгебра;

реляцинное исчисление с переменными-кортежами;

реляцинное исчисление с переменными-атрибутами.

Реляционная алгебра. В ней определяются основные операции над данными реляцинного типа. Все операции, вводимые в реляцинной алгебре, можно разделить на традиционные над множествами и специализированные, вводимые для удобства поиска в БД.

К операциям первой группы относятся: объединение, пересечение, разность, декартово произведение. К операциям второй группы следует отнести: проекцию, ограничение, соединение, деление.

Объединение. В результате применения этой операции получается отношение, объединяющее кортежи, содержащиеся в исходных отношениях. Пусть имеем два исходных отношения R_1 и R_2. Операция объединения этих отношений будет обозначаться $R_1 \cup R_2$:

$$R_1 \cup R_2 = \{r \mid r \in R_1 \quad или \quad r \in R_2\}.$$

Объединяемые отношения должны иметь одинаковые атрибуты (должны быть объединимыми):

Пример.

R_1:

A	B	C
a_1	b_1	c_1
a_2	b_2	c_2
a_3	b_3	c_3

R_2:

A	B	C
a_3	b_4	c_2
a_4	b_5	c_3
a_2	b_2	c_2

$R_1 \cup R_2$

A	B	C
a_1	b_1	c_1
a_2	b_2	c_2
a_3	b_3	c_3
a_3	b_4	c_2
a_4	b_5	c_3
a_2	b_2	c_2

Пересечение. В данной операции (которая обозначается $\cap$) получают отношение, включающее кортежи, общие для R_1 и R_2:

$$R_1 \cap R_2 = \{r \mid r \in R_1 \quad и \quad r \in R_2\}.$$

Пример. Для отношения R_1 и R_2 из предыдущего примера получим

$R_1 \cap R_2$:

A	B	C
a_2	b_2	c_2

Разность. В результате применения этой операции *($R_1 \backslash R_2$)* получается отношение, которое содержит кортежи, являющиеся кортежами отношения R_1 и не являющиеся кортежами отношения R_2:

$$R_1 \setminus R_2 = \{r \mid r \in R_1 \quad и \quad r \notin R_2\}.$$

Пример.

$R_1 \setminus R_2$:

A	B	C
a_1	b_1	c_1
a_3	b_3	c_3

Декартово (прямое) произведение. В этой операции ($R_1 \times R_2$) из m-местного отношения R_1 и n-местного отношения R_2 получают $(m+n)$-местное отношение. Причем первые m элементов представляют кортежи из отношения R_1, а последние n элементов - кортежи из отношения R_2:

$$R_1 \times R_2 = \{\langle r_1, r_2 \rangle \mid r_1 \notin R_1 \quad и \quad r_2 \in R_2\}.$$

Пример.

$R_1 \times R_2$:

A	B	C	A	B	C
a_1	b_1	c_1	a_3	b_4	c_2
a_2	b_2	c_2	a_3	b_4	c_2
a_3	b_3	c_3	a_3	b_4	c_2
a_1	b_1	c_1	a_4	b_5	c_3
a_2	b_2	c_2	a_4	b_5	c_3
a_3	b_3	c_3	a_4	b_5	c_3
a_1	b_1	c_1	a_2	b_2	c_2
a_2	b_2	c_2	a_2	b_2	c_2
a_3	b_3	c_3	a_2	b_2	c_2

Проекция. Операция проекции предназначена для изменения числа столбцов в отношении, т.е. в том случае, когда из строк-кортежей требуется исключить какие-либо атрибуты.

Обозначим через $j_1, j_2, ..., j_n$ — номера столбцов n-местного отношения R. Операцию определения проекции отношения R обозначим через $\pi_{j_1, j_2, ..., j_n}(R)$, а сама операция заключается в том, что из отношения R выбираются столбцы и компонуются в указанном порядке $j_1, j_2, ..., j_n$.

Пример. Рассмотрим отношение ТЕЛЕВИЗОР.

Наименование	*Индекс*	*Диагональ кинескопа, см*	*Цена, руб*
Рекорд	ВЦ-311	47	640
Рубин	Ц-266Д	67	1040
Темп	Ц-280Д	61	755
Электроника	Ц-283	61	755
Горизонт	Ц-355	57	610

Тогда можно получить проекцию $\pi_{1,4}$ (ТЕЛЕВИЗОР):

Наименование	Цена, руб.
Рекорд	640
Рубин	1040
Темп	755
Электроника	755
Горизонт	610

$\pi_{4,1,3}$ (ТЕЛЕВИЗОР):

Цена, руб.	Наименование	Диагональ кинескопа, см
640	Рекорд	47
1040	Рубин	67
755	Темп	61
755	Электроника	61
610	Горизонт	51

π_4 (ТЕЛЕВИЗОР):

Цена, руб.
640
1040
755
755
610

Можно отметить, что в последней проекции оказались одинаковые строки.

Ограничение. Ограничением называют такую операцию, в которой отношение исследуют по строкам и выделяют множество строк, которые удовлетворяют заданным условиям.

Обозначим через r строку (кортеж) отношения R, а через θ определим одно из отношений: $=$, $\neq$, $<$, $\leq$, $>$, $\geq$. Тогда θ-ограничение между избранным атрибутом A в отношении R и некоторой величиной C определяют следующим образом: $R[A\theta C]=\{r|r\in R$ и $r[A]\theta C\}$, где $r[A]$ соответствует значению атрибута A в строке r.

Таким образом, θ-ограничение обеспечивает получение среди строк отношения R только тех строк, в которых значение атрибута A и значение величины C удовлетворяют условию сравнения θ.

Например, нужно из отношения ТЕЛЕВИЗОР выделить те марки телевизоров, которые имеют стоимость меньше 700 руб.

ТЕЛЕВИЗОР [цена < 700]:				
	РЕКОРД	ВЦ-311	47	640
	ГОРИЗОНТ	Ц-355	51	610

Следует отметить, что в частном случае в операции сравнения вместо величины C можно использовать другой домен B. Тогда операция ограничения определяется как

$$R(A\theta B)=\{r\,|\,r\in R_1 \quad и \quad r(A)\theta r(B)\}.$$

Пример. Пусть отношение R имеет вид:

$$R: \quad \begin{array}{cccc} (A, & B, & D, & E) \end{array}$$
$$\begin{bmatrix} a & 15 & 17 & p \\ k & 22 & 28 & q \\ e & 39 & 11 & e \\ m & 16 & 43 & m \end{bmatrix}$$

Так как операция θ-ограничение предусматривает выбор среди строк отношения R только тех, в которых значения атрибутов A и B удовлетворяют условию сравнения θ, то для условий ($B<D$) и $(A= E)$ получим следующие отношения:

$$R\,[B<D]: \quad \begin{array}{cccc} A & B & D & E \end{array}$$
$$\begin{bmatrix} a & 15 & 17 & p \\ k & 22 & 28 & q \\ m & 16 & 43 & m \end{bmatrix}$$

$$R\,[A=E]: \quad \begin{array}{cccc} A & B & D & E \end{array}$$
$$\begin{bmatrix} e & 39 & 11 & e \\ m & 16 & 43 & m \end{bmatrix}$$

Соединение. Операция соединения обратна операции проекции. Рассмотрим для простоты два бинарных отношения $R_1(A, B)$ и $R_2(B, C)$. Соединением (обозначается $><$) отношений R_1 и R_2 $(R_1 >< R_2)$ называют операцию, при которой соединяют два отношения, используя в качестве признака соединения общий атрибут Y:

$$R_1 \bowtie R_2=\{\langle A,\ \mathrm{B},\ C\rangle\,|\,\langle A,\ B,\rangle\in R_1 \text{ и } \langle B,\ C,\rangle\in R_2\}.$$

Отношение $R_1 >< R_2$ является отношением с атрибутами $<A, B, C>$.

Например:

R_1:

Наименование	Индекс
Рекорд	ВЦ-311
Темп	Ц-280Д
Электроника	Ц-283
Горизонт	Ц-355

R_2:

Индекс	Цена
ВЦ-311	640
Ц-280Д	755
Ц-283	755
Ц-355	610

Отношение $R_1 >< R_2$ будет иметь вид:

Наименование	Индекс	Цена
Рекорд	ВЦ-311	640
Темп	Ц-280	755
Электроника	Ц-283	755
Горизонт	Ц-355	610

Операция соединения соответствует случаю, когда просто стыкуются таблицы отношений. Но поскольку в получающейся таблице атрибуты с одинаковым содержанием присутствуют дважды, один из одинаковых столбцов исключают. Такую операцию называют естественным соединением. Операцию соединения можно использовать не только для бинарных, но и для n-местных отношений:

$R^1 \bowtie R_2 \bowtie ,..., \bowtie R_n=\{M|M$ — кортежи, которые образованы соединениям атрибутов отношений $R_1, R_2,..., R_n\}$.

Отметим, что рассматривались в основному операции соединения фактически по условию равенства двух атрибутов в двух отношениях. В общем случае соединение можно осуществлять не только по условию равенства, но и по любой другой θ: =, ≠, <, ≤, >, ≥:

$$R_1 \underset{A\theta B}{\bowtie} R_2=\{(r,\ s)\,|\,r\in R_1 \text{ и } s\in R_2 \text{и } (r(A)\theta r(B))\},$$

когда θ - оператор равенства, то операцию называют эквисоединением.

Если сравнение θ возможно, то для атрибутов A и B одинаковые имена необязательны.

Пример. Пусть имеются отношения $R_1(A, B, C)$ и $R_2(D, E, F)$

$$R_1: \begin{array}{c} A\ \ B\ \ C \\ \begin{bmatrix} 8 & 4 \\ 2 & 7 \\ 5 & 5 \\ 8 & 3 \end{bmatrix} \end{array} \qquad R_2: \begin{array}{ccc} D & E & F \\ \end{array} \begin{bmatrix} d_1 & 8 & 3 \\ d_2 & 8 & 2 \\ d_3 & 8 & 10 \\ d_4 & 5 & 12 \end{bmatrix}$$

Для условий ($C=E$), $(B>F)$ получим

$$R_1 \underset{C=E}{\bowtie} R_2: \quad \begin{array}{c} \begin{array}{cccccc} A & B & C & D & E & F \end{array} \\ \begin{bmatrix} a_2 & 5 & 8 & d_1 & 8 & 3 \\ a_2 & 5 & 8 & d_2 & 8 & 2 \\ a_2 & 7 & 8 & d_1 & 8 & 3 \\ a_2 & 7 & 8 & d_2 & 8 & 2 \\ a_3 & 2 & 5 & d_4 & 5 & 12 \end{bmatrix} \end{array}$$

$$R_1 \underset{B>F}{\bowtie} R_2: \quad \begin{array}{c} \begin{array}{cccccc} A & B & C & D & E & F \end{array} \\ \begin{bmatrix} a_1 & 3 & 4 & d_2 & 8 & 2 \\ a_2 & 5 & 8 & d_1 & 8 & 3 \\ a_2 & 5 & 8 & d_2 & 8 & 2 \\ a_2 & 7 & 8 & d_1 & 8 & 3 \\ a_2 & 7 & 8 & d_2 & 8 & 2 \end{bmatrix} \end{array}$$

Таким образом, с помощью операции соединения можно объединять строки разных отношений по критерию сравнения значений каких-нибудь двух атрибутов.

Деление. Рассмотрим деление m-местного отношение R_1 на n-местное отношение R_2.

Пусть из общего количества m атрибутов отношения R_1 выделим несколько атрибутов: $A, B,...,F$ и из них составим список, обозначив его через M. Набор значений атрибутов из M столбцов можно рассматривать как проекцию отношения R на список атрибутов M, т.е. $\pi_M(R_1)$. Тогда через $\bar{M}$ будут обозначаться атрибуты дополнительного к M, т.е. атрибуты отношения R_1, не вошедшие в список M, и соответственно значения атрибутов из списка $\bar{M}$ определяются $\pi_{\bar{M}}(R_1)$. Будем полагать, что делимое (отношение R_1) может быть представлено таким образом, что его атрибуты сгруппированы в порядке $R_1(\bar{M}, M)$. В отношении R_2 также выделим несколько атрибутов $G, K,...,P$ и составим из них список, обозначив его через N (формально этот список представляет собой проекцию $\pi_N(R_2)$). Если проекции $\pi_M(R_1)$ и $\pi_N(R_2)$ объединимы , т.е. имеют одинаковое количество атрибутов, то можно рассматривать операцию деления R_1 по M на R_2 по N (что обозначается как $R_1[M ч N]R_2$). Операция деления $R_1[M ч N]R_2$ представляет собой операцию, которая определяет такое наибольшее множество значений атрибутов с $\pi_{\bar{M}}(R_1)$ (обозначим его через $r_1(\bar{M})$), что прямое произведение этого множества $r_1(\bar{M})$ с $\pi_N(R_2)$ содержится в R_1.

Операции деления можно определить с помощью уже ранее введенных операций таким образом:

$$R_1[M \div N]R_2 = \pi_{\bar{M}}(R_1) \setminus \pi_{\bar{M}}((\pi_{\bar{M}}(R_1) \times \pi_N(R_2)) \setminus R_1),$$

где $\pi_{\bar{M}}$ - это проекция отношения на атрибуты списка $\bar{M}$.

Пример.

R_1:

A	B	C	D
a_1	b_1	c_3	d_1
a_2	b_1	c_1	d_1
a_3	b_2	c_2	d_2
a_1	b_3	c_1	d_3
a_3	b_2	c_1	d_1
a_3	b_1	c_2	d_2

R_2:

C	D
c_1	d_1
c_2	d_2

Включим в список матрицы M атрибуты C и D (из отношения R_1), тогда в список $\bar{M}$ войдут атрибуты A и B, а в список N — атрибуты C и D из отношения R_2:

$$R_1[(C,\ D) \div (C,\ D)]\,R_2 \qquad \begin{array}{cc} A & B \\ \hline a_3 & b_2 \end{array}$$

Действительно,

$\pi_{\overline{M}}(R_1)$:

A	B
a_1	b_1
a_2	b_1
a_3	b_2
a_1	b_3
a_3	b_1

$\pi_{\overline{M}}(R_1) \times \pi_N(R_2)$:

A	B	C	D
a_1	b_1	c_1	d_1
a_1	b_1	c_2	d_2
a_2	b_1	c_1	d_1
a_2	b_1	c_2	d_2
a_3	b_2	c_1	d_1
a_3	b_2	c_2	d_2
a_1	b_3	c_1	d_1
a_1	b_3	c_2	d_2
a_3	b_1	c_1	d_1
a_3	b_1	c_2	d_2

$\pi_{\overline{M}}(R_1) \times \pi_N(R_2) \setminus R_1$:

A	B	C	D
a_1	b_1	c_1	d_1
a_1	b_1	c_2	d_2
a_2	b_1	c_2	d_2
a_1	b_3	c_1	d_1
a_1	b_3	c_2	d_2
a_3	b_1	c_1	d_1

$$\pi_{\overline{M}}\left(\left(\pi_{\overline{M}}(R_1) \times \pi_N(R_2)\right) \setminus R_1\right):$$

A	B
a_1	b_1
a_2	b_1
a_1	b_3
a_3	b_1

$$R_1[(C,\ D) \div (C,\ D)]\, R_2 = \pi_{\overline{M}}(R_1) \setminus \pi_{\overline{M}}((\pi_{\overline{M}}(R_1) \times \pi_N(R_2)) \setminus R_1):$$

A	B
a_3	b_2

При построении языка манипулирования данными на основе реляционной алгебры каждый оператор языка реализует некоторый набор алгебраических операций, в результате выполнения которых получается желательное исходное отношение.

Другой подход к построению языка манипулирования данными основан на использовании моделей реляционного исчисления.

Суть этого подхода заключается в том, что желательный результат определяется не заданием набора операций над отношениями, а путем описания требований, которым должно удовлетворять результирующее отношение. СУБД представляется самой подобрать последовательность операций, которые ведут к поставленной цели. Естественно, что языки, которые основаны на использовании моделей реляционного исчисления, представляют собой языка очень высокого уровня.

Реляционное исчисление. В реляционном исчислении, равно как и в реляционной алгебре, имеется набор понятий и операций, которые позволяют записывать любое отношение в виде некоторой формулы или формального выражения (α-выражения).

Формулы в реляционном исчислении кроме арифметических операций (= , ≠, <, ≤, >, ≥) включают дополнительные **логические** операции. К ним относят операции квантификации:

$\forall$ - квантор общности и $\exists$ - квантор существования.

Квантор общности $\forall$ читается как: «все для всех», «для каждого», «который бы не был».

Например, запись $\forall x$ читается «для любого *x*» или «для всех *x*».

Квантор существования $\exists$ читается «для некоторого», «существует хотя бы один». Например, $\exists y$ читается «для некоторого *y*» или «существует *y* такой, что». Кроме кванторов в реляционном исчислении используются логические операции: $\vee$, $\wedge$, $\urcorner$.

Операция $\vee$ носит название **логическое сложение,** или **дизъюнкция,** и по смыслу соответствует слову «ИЛИ». Операция $\wedge$ **логическое умножение, или конъюнкция,** и отвечает слову «И». Операция $\neg$ — операция отрицания, отвечает «НЕ», т.е. запись $R_1 \wedge R_2$ будет читаться как «отношение R_1 И отношение R_2», запись $R_1 \vee R_2$ — «отношение R_1 ИЛИ отношение R_2».

При записи выражения в реляционном исчислении используется понятие **свободных или связных переменных.**

Вхождение переменных *x* в формулу реляционнго исчисления $\psi(x)$ связано, если в ψ она находится в части формулы, которая начинается квантором $\forall$ или $\exists$, за которым непосредственно следует переменная *x*. В таких случаях говорят, что квантор связывает переменную *x*. В остальных случаях вхождение переменной *x* в формулу ψ свободно.

Например, в формуле

$$\forall \mathbf{x} (R_1(x,\ y) \vee (\exists y) R_2(x,\ y,\ z))$$

переменная *x* связана, переменная *y* в отношении R_1 свободна, а в R_2 — связана, переменная *z* свободна.

Формулы в реляционном исчислении строятся из атомов и совокупности арифметических и логических операторов. Атомы в реляционном исчислении представляются по-разному. В зависимости от того, что используется в качестве переменной - кортеж (строка) или атрибут (столбец). Поэтому различают реляционное исчисление с переменными-кортежами и переменными-доменами.

Реляционное исчисление с переменными-кортежами. Выражение такого исчисления может иметь следующий вид:

$$\{r \mid \psi(r)\},$$

где *r*-кортеж: ψ — некоторая формула исчисления. Например, выражение $\{r \mid R_1(r) \wedge R_2(r)\}$, где в качестве формулы $\psi(r)$ используется выражение $R_1(r) \wedge R_2(r)$, означает, что необходимо получить множество всех кортежей *r*, таких, что они принадлежат одновременно как отношению R_1, так и отношению R_2. В том случае, когда в качестве переменных в реляционном исчислении используются кортежи, атомы, из которых конструируются формулы, они могут быть следующих типов:

1. Атом — отношение $R(r)$, где *r* — кортеж в отношении *R*.

2. Атом — конструкция типа $s[i]\theta v[j]$ или $s[i]\theta c$, где *c* — некоторая константа; θ — арифметический оператор $(=, \neq, <, \leq, >, \geq)$;
s, *v* — кортежи; *i*, *j* — номера (или имена) атрибутов (столбцов) в соответствующих кортежах. Например, атом $(s[1] = q[7])$ означает, что первая компонента кортежа *s* равна седьмому кортежу *q*, а атом $s[5]<10$ означает, что пятая компонента меньше 10. Два

вышеперечисленных типа вида атомов являются единственными в реляционном исчислении.

Сами формулы рекурсивно конструируются из атомов по следующим правилам.

1. Любой атом - это формула. Все вхождения переменных-кортежей, которые упомянуты в атоме, являются свободными.

2. Если ψ формула, то $\neg\psi$ — тоже формула ($\neg$ - символ логического отрицания).

3. Если ψ_1 и ψ_2 формулы, то и выражения $\psi_1 \wedge \psi_2$ и $\psi_1 \vee \psi_2$ также являются формулами. Причем свободными (связными) являются те и только те вхождения переменных, которые происходят от свободных (связных) вхождений переменных ψ_1 и ψ_2.

4. Если ψ формула и r - свободная переменная-кортеж этой формулы, то $\forall r(\psi))$ и $\exists r(\psi)$ также формулы, переменная в этом случае становится связанной.

5. Формулы могут при необходимости заключаться в скобки.

6. Ничто другое не является формулой.

Реляционное исчисление с переменными-доменами. В этом случае в качестве переменных вместо кортежей используются домены. Реляционное исчисление с переменными на доменах строится с использованием тех же операторов, которые и с переменными кортежами. Но в качестве атомов формул исчисления используются следующие типы:

1. $R(x_1, x_2, ..., x_n)$, где R — n-арное отношение, x_i — константа или переменная на некотором домене.

Атом $R(x_1, x_2, ..., x_n)$ указывает, что значения x_i, которые являются переменными, должны быть выбраны так, чтобы $(x_1, x_2, ..., x_n)$ было кортежем отношения.

2. $x\theta в$, где x и y — константы или переменные на некотором домене, θ - арифметический оператор сравнения.

В остальном формулы реляционного исчисления с переменными-доменами строятся аналогично формулам исчисления с переменными-кортежами.

Выражение реляционного исчисления с переменными на доменах имеет вид $\{x_1, x_2.....x_n|\psi(x_1, x_2, ..., x_n)\}$, где ψ — формула, обладающая тем свойством, что только ее свободные переменные на доменах являются различными переменными $x_1, x_2, ..., x_n$.

Выражение реляционного исчисления $\{r|\psi(r)\}$ или $\{x_2, x_2, ..., x_n|\psi(x_1, x_2, ..., x_n)\}$ служит основой реальных языков манипулирования данными.

В реляционном исчислении доказано, что для любого простого выражения исчисления существует эквивалентное ему выражение реляционной алгебры. Поэтому может быть построена универсальная процедура для перевода выражений реляционного исчисления в эквивалентное по смыслу алгебраическое выражение.

7.6. Сетевая организация моделей структур

В основе разработки организации сетевых моделей структур лежит возможность представления связей между элементами структур в графической форме.

В основе сетевой модели лежат понятия «сущность» и «связь», а к основным типам структур модели относят: элемент данных, агрегат, запись, набор.

Сущность — это собирательное понятие, некоторая абстракция реально существующего объекта предметной области, процесса или явления. Набор однородных объектов или явлений определяет **тип** сущности, а каждый конкретный объект в наборе представляет **экземпляр** сущности. Связи между сущностями фиксируются заданием множества отношений. При анализе связей между сущностями наиболее часто используются бинарные связи, т.е. связи между двумя сущностями. По характеру бинарные связи между типами сущностей различают:

один к одному (1:1);

один ко многим $(1:M)$;

много к одному $(M : 1)$;

много ко многим $(M:M)$.

Элемент данных - это наименьшая единица данных, которой можно оперировать в БД и выполнять построение всех других структур. Можно отметить, что элемент данных представляет собой аналог поля в файловых системах. Элемент данных имеет имя, которое хранится в БД как часть описания базы. Именами элементов данных могут быть, например, ИНДЕКС ИЗДЕЛИЯ, ДАТА ВЫПУСКА, СТОИМОСТЬ и т.д. В сетевых моделях элементы данных используются для представления атрибутов сущности.

Агрегат данных — совокупность элементов данных, имеющих общее имя, которую можно рассматривать как единое целое. Например, агрегат данных ДАТА состоит из элементов данных: ЧИСЛО, МЕСЯЦ, ГОД.

Запись - совокупность элементов данных, которые описывают конкретный экземпляр объекта (сущности). Предположим, что сущность ТЕЛЕВИЗОР описывается элементами данных: МАРКА; ИНДЕКС, ЦЕНА. Тогда запись в этом объекте для конкретного изделия может быть: РЕКОРД, ВЦ-311, 640. Можно отметить, что запись эквивалентна кортежу в реляционных моделях данных.

Сетевая модель в качестве базовых использует понятие «экземпляр» и «набор».

Тип — это общее понятие, которое представляет собой собрание экземпляров записи. Каждый тип записи состоит из некоторого числа элементов данных, значения которых размещаются в экземплярах записи данного типа. В качестве связей между типами записей используются наборы. Каждый набор представляет собой отношение (связь) между двумя или несколькими типами записей. Он отображает множество связей между экземплярами записей типа «владелец» и «член». Для каждого типа набора один тип записи может быть объявлен «владельцем», а остальные — его «члены». При этом любой экземпляр записи типа «член» может быть связан не больше чем с одним экземпляром типа «владелец».

Графическая интерпретация сетевой модели данных представляет собой ориентированный граф без петель. Причем, вершинам графа соответстуют типы записей, а дугам - наборы, отражающие связи между соответствующими типами записей. Направленные стрелки на дуге ориентированы от записи типа «владелец» к записи типа «член».

Подмножество дуг, соединяющих один запись - владельца с несколькими записями- членами, называется **экземпляром набора**.

Рассмотрим, например, граф, отражающий упрощенную БД комплектующих деталей телевизора (рис. 22).

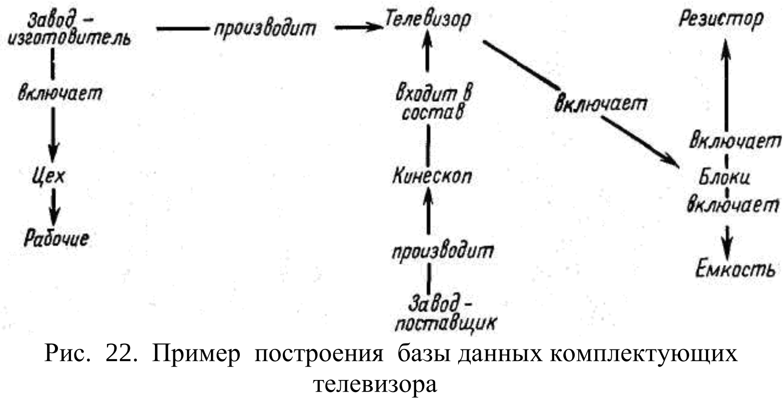


Рис. 22. Пример построения базы данных комплектующих телевизора

Стрелки между вершинами отвечают наборам данных, отражающих связи между записями, а надписи над стрелками - именам наборов.

Как тип записи, так и набор данных в общем случае могут быть представленны таблицами. Но в отличие от таблиц реляционных моделей, в сетевых моделях данных они могут допускать дубликаты строк или записей.

В некоторых сетевых моделях вводится особый тип набора данных, называемый **сингулярным** и имеющий только один экземпляр набора этого типа. В нем запись «владелец» отсутствует (владельцем является система управления базой данных). Этот тип набора обычно используется для создания традиционного файла, состоящего из однотипных записей.

В сетевой модели данных такой как КОДАСИЛ есть несколько ограничений, которые необходимо учитывать при построении модели. Основным внутренним ограничением является функциональность связей, так как нельзя реализовать связи типа *М:М.* В модели это ограничение соответствует положению: в конкретном экземпляре набора экземпляр записи числа может иметь не больше одного экземпляра записи владельца.

Для того чтобы отобразить принятую схему данных в памяти ЭВМ, требуется описать все таблицы, соответствующие записям и наборам. Для этого группой КОДАСИЛ был предложен язык описания данных, позволяющий задать схему данных с помощью четырех типов статей.

Статья схемы задает имя схемы БД. Статья запишется как:

SCHEMA NAME IS имя схемы.

Статья области характеризует область памяти, в которой размещаются экземпляры записей БД. С помощью этой статьи можно в случае необходимости распределять БД по разным ЗУ. Поскольку в общем случае можно выделить под БД несколько разных областей, каждая из них должна иметь собственное уникальное имя и описываться следующим образом:

AREA NAME IS имя области;

Статья записи содержит описание типа записи, которая включает имя записи и характеризующее все элементы данных, которые входят в ее состав. Каждому типа записи соответствует своя статья.

Статья записи начинается предложением:

RECORD NAME IS имя записи;

Следующий за этим предложением текст зависит от варианта реализованного языка КОДАСИЛ. Наиболее распространенными есть ЯОД КОДАСИЛ. Поскольку ряд действующих СУБД реализуют ЯОД, дальше будем рассматривать случай, когда статья записи формируется на его основе. Тогда вторым предложением в схеме записи будет идти предложение:

$$\text{LOCATION MODE IS} \left\{ \begin{array}{l} \text{DIRECT} \\ \text{CALC (имя процедуры) USING (имя calc-элемента)} \\ \text{DUPLICATES ARE (NOT) ALLOWED} \\ \text{VIA имя набора SET} \\ \text{SYSTEM} \end{array} \right.$$

где в фигурных скобках указано одно из возможных описаний.

DIRECT используется в том случае, когда предполагается, что номер страницы области для размещения записи будет определяться программой, которая выполняет включение записи в БД.

CALC применяется, когда предполагается, что специальная программа будет использовать значение ключа БД.

Ключ базы данных — это идентификатор, который уникально определяет запись, помещенную в БД.

Для того чтобы можно было различать отдельные экземпляры записей, которые сохраняются в БД, каждому экземпляру записи присваивается значение ключа, который играет роль внутрисистемного идентификатора

Зная значение ключа, можно быстро отыскать соответствующую запись. В формате CALC USING имя calc-элемента, в качестве имен Calc-элемента используются имена элементов данных записи.

В том случае если ключи не имеют дубликатов значений, то в варианте CALC следует указать DUPLICATES ARE NOT ALLOWED. Если ключ имеет дубликаты значений, например в качестве ключа задан элемент данных ФАМИЛИЯ в записи типа СТУДЕНТ, то следует указать DUPLICATES ARE ALLOWED.

VIA SET используется в том случае, когда нужно запись разместить физически как можно ближе к соответствующему экземпляру набора, в который он будет включен.

SYSTEM применяется в случае, если размещение записей возлагается на саму СУБД (в соответствии с заложенными в нее алгоритмами).

Для того чтобы приписать рассматриваемый тип записи к некоторой области, используется предложение

WITHIN имя области AREA

Для описания внутренней структуры записи в статью записи включается подсистема данных, которая имеет вид

$$\left\{\begin{array}{l}\left\{\begin{array}{l}\texttt{BINARY}\\ \texttt{DECIMAL}\end{array}\right\}\\ \left\{\begin{array}{l}\texttt{FIXED}\\ \texttt{FLOAT}\end{array}\right\}\\ \texttt{REAL}\\ \texttt{CHARACTER}\\ \texttt{DATA-BASE-KEY}\end{array}\right\}$$

С помощью этой подсистемы каждому элемента данных приписывается тип значений данных: двоичное, десятичное, фиксированное, плавающее, натуральное, символьное, ключ БД.

Статья набора позволяет описать наборы БД. Формат задачи набора имеет вид

$$\texttt{SET NAME IS имя набора;}$$

$$\texttt{OWNER IS}\left\{\begin{array}{l}\texttt{имя записи}\\ \texttt{SYSTEM}\end{array}\right\}$$

$$\texttt{ORDER IS PERMANENT INSERTION IS}\left\{\begin{array}{l}\texttt{SORTED}\\ \texttt{PRIOR}\\ \texttt{NEXT}\\ \texttt{LAST}\\ \texttt{FIRST}\end{array}\right\}$$

Первое предложение статьи задает имя описываемого набора. Второе указывает имя типа записи, являющейся записью владельца набора. И последнее - на способ включения экземпляров записей-членов в экземпляры описываемого типа набора.

Обычно предполагается, что на внутреннем уровне экземпляры набора, которые включают несколько членов записей, организованы в виде цепочки.

Цепочка записей реализуется с помощью специального служебного элемента — **указателя,** который содержит (указывает на) адрес записи, логически связанного с рассматриваемой. Посредством указателей можно организовать обращение не только к следующим записям, но и к предыдущим.

Команды, которые входят в фигурные скобки третьего предложения схемы, как раз и позволяют организовать цепочку необходимым образом и означают:

FIRST - запись включается первым в цепочку перебора записей, т.е. сразу же после записи владельца (рис. 23).

LAST - запись включается последней (рис. 23).

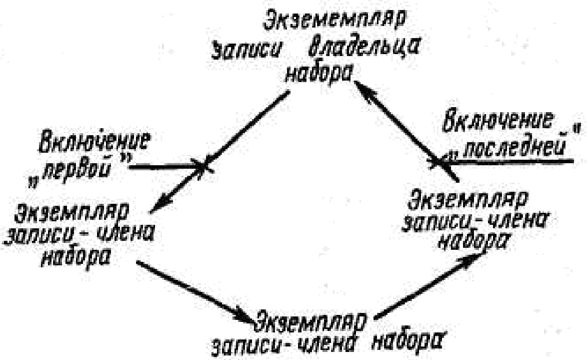


Рис. 23. Реализация вариантов FIRST и LAST

NEXT - включение следующей записи, т.е. записи, которая следует за текущей записью набора (рис. 24). Под текущей записью набора понимается тот экземпляр записи конкретного экземпляра набора, на который указывает «текущая» запись, в типе набора.

PRIOR - включение предыдущей записи, т.е. перед текущей записью набора (рис. 24).

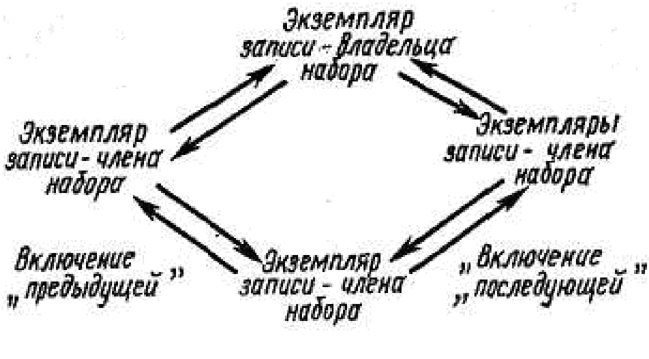


Рис. 24. Реализация вариантов NEXT и PRIOR

SORTED - СУБД поддерживает упорядоченность экземпляров однотипных записей, которые входят в экземпляр набора, в соответствии с возрастанием и убыванием ключевого данного.

Описание числа набора проводится с помощью подстатьи члена набора, имеющей вид

$$\text{MEMBER IS имя записи} \begin{Bmatrix} \text{AUTOMATIC} & \text{MANDATORE} \\ \text{MANUAL} & \text{OPTIONAL} \end{Bmatrix}$$

В зависимости от способа включения экземпляра записи-члена в экземпляр набора различают два типа членства в наборе:

AUTOMATIC - автоматический;

MANUAL - ручной.

Чтобы поместить в БД новую запись, необходимо сформировать ее в рабочей области программ и только затем с помощью операторов языка манипулирования данными записать в БД. Если необходимо экземпляр записи-члена поместить в БД, но при этом не соединять ни с каким экземпляром набора, то указывают оператор MANUAL. Последующее включение этой записи в набор должно быть осуществлено программистом явным образом с помощью соответствующих средств языка манипулирования данными. Тип членства AUTOMATIC предусматривает, что в момент помещения в БД запись должна быть автоматически включена в набор. Операторы MANDATORY (обязательный) и OPTIONAL (необходимый) характеризуют вид исключения экземпляра записи из экземпляра набора описываемого типа. Так тип OPTIONAL означает, что запись может исключаться из набора данных в произвольный момент времени.

Поскольку БД предполагает соответствующее манипулирование данными, рабочей группой КОДАСИЛ был также предложен соответствующий язык, позволяющий осуществлять запись в базе данных: STORE (запомнить), ERASE (стереть), MODIFY (изменить), CONNECT (присоединить), DISCONNECT (исключить) и др.

Оператор STORE (запомнить) - помещает данные, подготовленные в рабочей области программы, в базу.

Оператор ERASE (стереть) - удаляет из БД экземпляр записи.

Оператор MODIFY (изменить) - обновляет текущую запись банка данных.

Оператор CONNECT (присоединить) - включает текущую запись в набор.

Оператор DISCONNECT (исключить) - исключает текущую запись из набора, но при этом она остается в БД.

Можно отметить, что в языке манипулирования данных существуют и другие операторы, но приведенный набор операторов ЯМД является функционально полным, т.е. позволяет образовывать в БД произвольное допустимое схемой базы данных состояние.

7.7. Гиперсети как высшая форма организация сетевых моделей структур

В настоящем разделе рассматривается классификация гиперсетей, приведены определения маршрутов, отделимости и соединимости вершин в гиперсетях и алгоритмы их вычисления в том случае, когда указанные задачи не *NP*-полные. Исследуются методы синтеза некоторых классов гиперсетей с заданной связностью. Гиперсети в качестве математических моделей связности сложных систем — наиболее адекватные модели структур сетей связи, энергетических, транспортных, нефте- и газопроводных и т. п. В процессе синтеза их структур важно учитывать не только оптимизацию, но и структурную надежность систем. Этим объясняется актуальность исследования характеристик связности гиперсетей.

7.7.1. Основные понятия и определения

1.1. Абстрактные гиперсети. Шестерку $S = (X, V, R; P, F, W)$ назовем абстрактной гиперсетью, если

$X = (x_1, x_2, ..., x_n)$— множество вершин;

$V = v_1, v_u ..., v_q)$— множество ветвей;

$R = (r_1, r_2, .. ., r_m)$— множество ребер;

$P : V \rightarrow 2^X$ — отображение, сопоставляющее каждому элементу $v \in V$ множество $P(v) \in X$ его вершин. Тем самым отображение P определяет гиперграф $PS = (X, V; P)$;

$F: R \rightarrow 2^V_{PS}$ — отображение, сопоставляющее каждому элементу $r \in R$ множество $F(r) \in V$ его ветвей. Причем семейство подмножеств ветвей 2^V_{PS} содержит только такие, ветви которых составляют связную часть гиперграфа PS. Отображение F определяет гиперграф $FS = (V, R; F)$;

$\forall r \in R \;\; W : r \rightarrow 2^{P(F(r))}$ — отображение, сопоставляющее каждому элементу $r \in R$ подмножество $W(r) \subseteq P(F(r)) \subseteq X$ его вершин, где $P(F(r))$—множество вершин в PS, инцидентных ветвям $F(r) \in V$. Таким образом, отображение W определяет гиперграф $WS = (X,R; W)$.

Гиперграф PS назовем первичной сетью гиперсети S, а гиперграф WS — вторичной сетью гиперсети S. Обратные отображения определяются следующим образом:

$$P^{-1}(x) = \{v : x \in P(v)\}, \; F^{-1}(v) = \{r : v \in F(r)\},$$
$$W^{-1}(x) = \{r : x \in W(r)\}.$$

В некоторых случаях для удобства будем использовать следующие обозначения гиперсетей: S, $S=(X, V, R)$, $S=(PS, WS; \Phi)$. В первом случае указывается только «имя» абстрактной гиперсети, во втором—образующие множества, а в третьем—«имена» первичной и вторичной сетей гиперсети S и отображение $\Phi : WS\rightarrow PS$.

Абстрактные гиперсети допускают одно важное для моделирования систем сетевой структуры обобщение. Пусть заданы гиперграфы $PS=(X, V; P)=WS_0=(X, R_0; W_0)$, $WS_1=(Y_1, R_1; W_1), \ldots$ $\ldots, WS_k=(Y_k, R_k; W_k)$. Тогда последовательность отображений $\Phi_i : WS_k \xrightarrow{\Phi_k} WS_{k-1} \xrightarrow{\Phi_{k-1}} \ldots \xrightarrow{\Phi_2} WS_1 \xrightarrow{\Phi_1} WS$ определяет иерархическую абстрактную k-гиперсеть $S=(PS, WS_1, \ldots, WS_k; \Phi_1, \ldots, \Phi_k)$, если $Y_k \subseteq Y_{k-1} \subseteq \ldots \subseteq Y_1 \subseteq X$ и $\forall i \in 1, \ldots, k$, $\forall r^i \in R_i$ $P_i(r^i)=\{r^{i-1}\} \subseteq R_{i-1}$, $W_i(r^i) \subseteq W_{i-1}(\{r^{i-1}\})$ и $\{r^{i-1}\}$ образуют связную часть в гиперграфе WS_{i-1}.

1.2. Инцидентность. Смежность. Отображения P, F, W вместе с обратными отображениями являются отношениями инциденции в соответствующих гиперграфах PS, FS, WS и, следовательно, они определяют инцидентность элементов в абстрактной гиперсети S.

Фундаментальным понятием в теории гиперсетей является отношение слабой инциденции. Два элемента из различных множеств слабо инцидентны, если найдется элемент из третьего множества, инцидентный им обоим. Например, вершина $x\in X$ слабо инцидентна ребру $r\in R$ (т. е. $x \in W^*(r)$), если только существует ветвь $v\in V$ такая, что $x\in P(v)$ и $v\in F(r)$, т. е. $x \in P(F(r))$.

Слабо инцидентные элементы могут оказаться инцидентными. Поэтому в некоторых случаях необходимо рассматривать понятие строгой слабой инцидентности, т. е. такие слабо инцидентные элементы, которые не могут быть инцидентными.

Для элементов абстрактных гиперсетей можно определить шесть понятий смежности и столько же — слабой смежности. Действительно, по аналогии с графами и гиперграфами два элемента из одного множества смежны тогда и только тогда, когда найдется элемент из другого множества, инцидентный им обоим. Но так как в абстрактных гиперсетях для пары элементов из любого множества можно найти инцидентные элементы из двух различных множеств, то в этом случае для конкретных элементов имеют место соответственно два понятия смежности. Например, вершины x и y из X V-смежны, если $\exists v\in V : x \in P(v)$ и $y\in P(v)$, и R-смежны, если $\exists r\in R : x \in W(r)$ и $y\in W(r)$. Аналогично определяются смежность других элементов S и слабая смежность этих элементов. Степень (слабая степень) элементов из S равна числу смежных (слабо смежных) ему элементов, т. е. для каждого элемента имеем четыре определения его степени.

7.7.2. Классификация гиперсетей

2.1. Из определений непосредственно следует, что абстрактная гиперсеть обобщает такие математические объекты, как ультраграфы, гиперграфы, гиперсхемы (по зарубежным публикациям — гиперсети), графы, графы связей и т. п. Вводя различные ограничения на множества *X, V, R* и отображения *P, F, W,* можно получить различные классы гиперсетей и перечисленные математические модели связности. Характерной особенностью гиперсетей является то, что рассматриваются три независимых множества элементов и отношения инциденции между ними. Поэтому при классификации гиперсетей разумным подходом будет рассмотрение подклассов, инициируемых только ограничениями на *P, F, W* и типом связной части гиперграфа *PS* при отображении $F : R \rightarrow 2^{V}_{PS}$. Иерархические гиперсети не рассматриваются.

2.2. Если предположить, что первичная *PS = (X, V; P)* и вторичная *WS = (X, R; W)* сети абстрактной гиперсети *S* являются графами, а связные части 2^{V}_{PS} — всевозможными маршрутами в *PS,* то имеет место обычное определение гиперсети. Когда маршрут есть цепь в *PS,* то гиперсеть называется обыкновенной, в случае простой цепи имеем простую гиперсеть. Основное внимание в данном разделе уделено простым гиперсетям.

2.3. В определении абстрактной гиперсети рассматриваются два понятия: слабая инцидентность и инцидентность. Таким образом, вводя ограничения на соответствующие отношения, будем получать различные классы абстрактных гиперсетей. Например, пусть отображение $P: V \rightarrow 2^{X}$ двузначно, тогда первичная сеть *PS* абстрактной гиперсети является графом. Этот факт будем обозначать $\{X \rightarrow V\}$, т. е. любому элементу $v \in V$ инцидентны не более чем две вершины $x, y \in X$. В том случае, когда накладывается ограничение на слабую инцидентность двух элементов, $\{R > X\}$ означает, что любой вершине $x \in X$ слабо инцидентно не более двух ребер, причем здесь рассматривается строгая слабая инцидентность (см. п. 1.2).

2.4. Число инцидентных или слабо инцидентных элементов для заданного элемента из другого множества — основной параметр классификации абстрактных гиперсетей. Обозначим скобками гиперсеть, в которой соответствующий элемент имеет не более чем один инцидентный (слабо инцидентный) элемент. Фигурные скобки обозначают инцидентность двум элементам, а $\{\ \}^k$— инцидентность k элементам (k≥3). Например, в *(R → V)* -гиперсети каждой ветви $v \in V$ инцидентно не более чем одно ребро $r \in R$. Классификация гиперсетей завершена, если для каждой упорядоченной пары множеств из *X, V, R* определены численные ограничения на отношения инциденции (слабой инциденции) и тип связной структуры в первичной сети *PS=(X, V; P)* при отображении *F* (см. п. 2.2).

Легко заметить, что некоторые классы абстрактных гиперсетей включают в себя другие классы (например, *(R> X)* -гиперсеть простая, а *(R > V)* -гиперсеть обыкновенная; обратное неверно) или абстрактная гиперсеть может принадлежать различным классам. Некоторые классы могут иметь пустое пересечение. Из сказанного следует, что система классов абстрактных гиперсетей имеет нетривиальную структуру, установить зависимость между классами в некоторых случаях сложно.

2.5. Так как в коммуникационных сетях, моделируемых абстрактными гиперсетями, первичная сеть в основном является графом, то исследование характеристик связности будет осуществляться только для гиперсетей (вторичная сеть может быть гиперграфом). В п. 2.2 перечислены некоторые типы связных частей в 2^V_{PS}. Необходимо добавить такие важные структуры, как деревья (*T*-гиперсети), циклы (*C*-гиперсети) и *k*-полюсники (*H*-гиперсети). В последнем случае в 2^V_{PS} рассматриваются всевозможные части графа *PS = (X, V; P)* с выделенными k*к* вершинами (например, подграф с заданной связностью между парами заданных вершин).

7.7.3. Маршруты и метрика в гиперсетях

Понятие «маршруты» играет фундаментальную роль в анализе связности гиперсетей и исследовании их метрических свойств. То что маршруты в гиперсетях задаются по-разному, способствует расширению изобразительных и операционных средств теории гиперсетей. В этом разделе решены некоторые задачи, связанные с поиском маршрутов в гиперсетях.

3.1. Маршрутом в гиперсети *S=(X, V, R)* называется конечная последовательность $\mu=(x_1, r_1, x_2, \ldots, x_{k-1}, r_{k-1}, x_k)$, составленная из элементов *X, R* таким образом, что вершины и ребра чередуются, а всякие два соседних элемента инцидентны. Квазимаршрут в гиперсети *S =(X, V, R*)—это конечная последовательность μ, в которой пара соседних элементов x_i, r_i инцидентна, а r_i, x_{i+1} слабо инцидентна. Если в определении маршрута заменить «инцидентность» на «слабую инцидентность», то получим определение слабого маршрута. Ориентированные маршруты определяются аналогично, с учетом ориентации ребер.

3.2. Рангом $\nu(\mu)$ маршрута μ (квазимаршрута, слабого маршрута) называется число ребер (или их частей), принадлежащих этому маршруту. Отдаленность (квазиотдаленность, слабая отдаленность) между вершинами $x, y \in X$ численно равна рангу кратчайшего маршрута (квазимаршрута, слабого маршрута), соединяющего эти вершины, и обозначается через $\nu(x, y)(\overline{\nu}(x, y), \tilde{\nu}(x, y))$. Отдаленность и слабая отдаленность удовлетворяют аксиомам метрики, квазиотдалешюсть — аксиомам орметрикп.

Длиной ребра (или его части) называется число ветвей, инцидентных этому ребру (части ребра). Длина $\rho(\mu)$ маршрута μ (квазимаршрута, слабого маршрута) равна суммарной длине ребер (их частей), входящих в маршрут μ. Расстояние (квазирасстояние, слабое расстояние) между вершинами $x, y \in X$ в гиперсети S равно длине кратчайшего маршрута (квазимаршрута, слабого маршрута), соединяющего эти вершины, и обозначается через $\rho(x,y)(\overline{\rho}(x,y), \tilde{\rho}(x,y))$. Расстояние и слабое расстояние удовлетворяют аксиомам метрики, квазирасстояние — аксиомам орметрики. Аналогично определяются эти понятия для ориентированных гиперсетей.

Рассмотрим теперь методы сводимости задач поиска кратчайших маршрутов в гиперсетях к аналогичным задачам на графах и гиперграфах.

3.3. Кратчайшие маршруты. Из определения маршрута в гиперсети $S = (X, V, R)$ непосредственно следует, что каждому маршруту μ в S между вершинами x и y соответствует маршрут μ в графе $WS =(X, R)$.

Теорема 1. *Пусть $S =(X, V, R)$—простая гиперсеть, заданная матрицами M^{xv}, M^{xr} и M^{rv}. Тогда (i, j)-й элемент матрицы A^p равен числу маршрутов ранга p из x_i в x_j, где*

$$A= M^{xr} \text{ " } M^{rx}, \text{ а } M^{xr} =(M^{rx})^T.$$

Следствие. *Если S — связная гиперсеть, то $\nu(x_i, x_j)$, $i \neq j$, равно наименьшему из целых чисел p, для которых (i, j)-й элемент матрицы A^p отличен от нуля.*

Для ориентированных маршрутов следуют аналогичные утверждения.

Расстояние $\rho(x_i, x_j)$между парой вершии x_i и x_j в гиперсети S находится по взвешенному графу $WS^* =(X, R)$. Каждому ребру r_k графа WS^* ставится в соответствие длина, равная $|\{v_i\}|$, где $\{v_i\}=F(r_k)$. Таким образом, кратчаншие маршруты в гиперсетях можно найти с помощью известных алгоритмов по графу *WS**. Для ориентированной гиперсети граф *WS** ориентируется.

3.4. Кратчайшие квазимаршруты. Отношении квазиотдаленности и квазирасстояния несимметричны для неориентированных гиперсетей, тем более для ориентированных. Рассмотрим способы вычисления кратчайших квазимаршрутов между различными упорядоченными парами вершин $x_i, x_j \in X$.

Теорема 2. *Пусть $S=(X, V, R)$—простая гиперсеть, заданная матрицами M^{xv}, M^{xr} и M^{rv}. Тогда (i, j)-й элемент матрицы B^p равен числу квазимаршрутов ранга p из x_i в x_j, где $B = M^{xr}$ " M^{rv} " M^{vx}.*

Следствие. *Если S — связная гиперсеть, то $\overline{\nu}$ (x_i, x_j), $i \neq j$, равно наименьшему из целых чисел p, для которых (i, j)-й элемент матрицы B^p отличен от нуля.*

Доказательства теоремы 2 и следствия следуют из того факта, что матрица B является матрицей смежности некоторого ориентированного графа $BS = (X, U)$, в котором из вершины x_i идет дуга $u \in U$ в вершину x_j, если и только если в гиперсети S *существует* ребро r_k, инцидентное вершине x_i и слабо инцидентное x_j.

Для ориентированных гиперсетей также можно вычислить квазиотдаленность, если построить ориентированный граф $\overrightarrow{BS} = (X, E)$, в котором из x_i идет дуга $e \in E$ в вершину x_j, если и только если в гиперсети S существует ориентированное ребро r_k, исходящее из x_i и слабо инцидентное x_j. Для матрицы смежности $\vec{B}$ графа $\overrightarrow{BS}$ справедливы теорема 2 и следствие этой теоремы.

Квазиотдаленность $\overline{\nu}$ (x_i, x_j) между парой вершин $x_i, x_j \in X$ можно найти с помощью ультраграфа $US = (X, R; f, g)$, который строится по гиперсети S. Каждой вершине (ребру) гиперсети S ставится в соответствие вершина (ребро) ультраграфа $US=(X, R; f, g)$. Отображения $f{:}X \rightarrow R$ и $g : R \rightarrow X$ для US определяются следующим образом:

$\forall x_i \in X$ $r_j \in f(x_i)$ тогда и только тогда, когда вершина x_i инцидентна ребру r_j в гиперсети S;

$\forall r_j \in R$ $x_i \in g(r_j)$ тогда и только тогда, когда кершина x_i слабо инцидентна ребру r_j.

Нетрудно показать, что $\overline{\nu}$ (x_i, x_j) в гиперсети S равно длине кратчайшего сильного маршрута в ультраграфе US между соответствующими вершинами.

Ориентированной гиперсети, очевидно, ставится в соответствие ориентированный ультраграф $\overrightarrow{US} = (X, R; f, g)$, в котором

$\forall\, x_i \in X\ r_j \in f(x_i)$тогда и только тогда, когда ориентированное ребро $r_j \in S$ выходит из вершины $x_i \in S;$

$\forall\, r_j \in R\ \ x_i \in g(r_j)$ тогда п только тогда, когда вершина x_i слабо инцидентна ребру r_j и $r_j \in f(x_i)$.

Очевидно, что между парой вершин x_j и $x_i \in S$ существует кратчайший ориентированный квазимаршрут тогда и только тогда, когда существует кратчайший сильный маршрут между соответствующими вершинами в ориентированном ультраграфе $\overleftrightarrow{US}$.

Квазирасстояние $\bar{\rho}$ (x_i, x_j) между парой вершин x_i и x_j гиперсети S можно найти по взвешенному орграфу $BS^* = (X, U)$, в котором длина дуги u равна числу ветвей в части ребра r_k, соединяющего вершины x_i и x_j в гиперсети S, т. е. числу элементов в соответствующей части кортежа $F(r_k)=\{x_i, .. ., x_j; ...\}$. Для ориентированных гиперсетей квазирасстояние от вершины x_i к вершине x_j находится с помощью взвешенного орграфа $\overleftrightarrow{BS} = (X, E)$, в котором длины дуг определяются аналогично.

Заметим, что вычисление кратчайших квазимаршрутов с помощью специально построенных графов и ультраграфов не накладывает ограничений на класс гиперсетей.

3.5. Кратчайшие слабые маршруты. Отношения слабой отдаленности и слабого расстояния играют большую роль при решении прикладных задач на гиперсетях. Так же как и в предыдущих случаях, справедлива следующая теорема.

Теорема 3. *Пусть $S = (X, V, R)$ — простая гиперсеть, заданная матрицами M^{xv}, M^{xr} и M^{rv}. Тогда (i, j)-й элемент матрицы D^p равен числу слабых маршрутов ранга p из x_i в x_j, где*

$$D = M^{xv} \,"\, M^{vr} \,"\, M^{rv} \,"\, M^{vx}.$$

Следствие. *Если S — связная гиперсеть, то $\vartheta(x_i, x_j)$, $i \neq j$, равно наименьшему из целых чисел p, для которых (i, j)-й элемент матрицы D^p отличен от нуля.*

Доказательства теоремы 3 и следствия вытекают из свойств графа $DS = (X, U)$, соответствующего матрице смежности D.

Множество вершин графа DS совпадает с множеством вершин гиперсети S и x_i смежна с x_j в DS, если и только если в гиперсети S найдется ребро r_k, слабо инцидентное x_i и x_j. В графе DS любому маршруту взаимно однозначно соответствует слабый маршрут в гиперсети S. Матрице D соответствует также гиперграф $HS = (X, R)$, который задается матрицей инциденций $N^{xr}=M^{xv} \,"\, M^{vr}$, следовательно, кратчайшие слабые маршруты в гиперсети можно находить по гиперграфу *HS*. Кроме того, слабые маршруты в ультраграфе *US* соответствуют таковым в *S*.

Если гиперсеть *S* ориентированная, то ей можно поставить в соответствие орграф $\vec{DS} = (X, \vec{U})$, $|\vec{U}| = |U|$. Из вершины x_i идет дуга u в вершину x_j, если и только если существует ребро $r_k \in S$, слабо инцидентное x_i и x_j и ориентированное от x_i к x_j. Любому ориентированному маршруту в $\vec{DS}$ взаимно однозначно соответствует ориентированный слабый маршрут в *S*.

Слабое расстояние $\delta(x_i, x_j)$ в гиперсетях (ориентированных гиперсетях) между вершинами x_i и x_j, вычисляется с помощью взвешенного графа $DS^*=(X, U)$ ($\vec{DS}^* = (X, \vec{U})$), в котором ребру (дуге) u ставится в соответствие длина, равная числу ветвей в части ребра r_k, соединяющего вершины x_i и x_j в гиперсети *S*, т. е. числу элементов в соответствующей части кортежа $F(r_k) = \{\ldots;\ x_i,\ldots;\ x_j;\ldots\}$.

3.6. О поиске цепей в гиперсетях. Маршрут μ в гиперсети $S= (X, V, R)$ называется *r*-ценью, если каждое ребро используется не более одного раза, и *v*-цепыо, если каждая ветвь используется не более одного раза. Отсюда следует, что всякая *v*-цепь одновременно является *r*-цепью. Обратное неверно. Маршрут μ называется простой цепью, если в нем все вершины различны.

Лемма 1. *Всякий маршрут (квазимаршрут, слабый маршрут) гиперсети S содержит r-цепь (r-квазицепъ, слабую r-цепь), соединяющую ту же пару вершин.*

Теорема 4. *Всякий кратчайший маршрут (квазимаршрут, слабый маршрут) между двумя вершинами гиперсети S является r-цепъю (r-квазицепъю, слабой r-цепъю).*

Следствие. *Задачи поиска r-цепей (r-квазщепей, слабых r-цепей) между двумя вершинами в гиперсети S полиномиально вычислимы.*

Легко показать, что задачи поиска слабой простой цепи и слабой *v*-цепи эквивалентны соответственно задачам поиска простой цепи и цепи в графе *PS*.

ПЦ: пусть задана гиперсеть $S = (X, V, R)$ и пара выделенных вершин $s,\ t \in X$. Существует ли простая цепь между вершинами *s* и *t* в гиперсети *S*?

Теорема 5. *Задача ПЦ является NP-полной.*

Следствие. *Задача поиска простой квазицепи в гиперсети S между вершинами s и t является NP-полной.*

ВЦ: пусть задана гиперсеть $S = (X, V, R)$ и пара выделенных вершин $s, t \in X$. Существует ли *v*-цепь между вершинами *s* и й гиперсети *S*?

Теорема 6 . *Задача ВЦ является NP-полной.*

Следствие. *Задача поиска v-квазицепи в гиперсети S между вершинами s и t является NP-полной.*
Для ориентированных гиперсетей справедливы аналогичные результаты.

7.7.4. Независимость и соединимость

4.1. Два маршрута, соединяющих пару вершин $x, y \in X$, называются внутренне независимыми (внешне независимыми), если не существует вершины $z \neq x, y$, инцидентной (строго слабо инцидентной) ребрам этих маршрутов. Маршруты называются независимыми, если они одновременно внутренне и внешне независимы. Два маршрута, соединяющих пару вершин, V-независимы (R-независимы), если не существует ветви (ребра), принадлежащей обоим маршрутам.
Аналогичным образом определяется независимость для квази- и слабых маршрутов. Поскольку в них отдельные ребра включены частично, имеет смысл ввести еще одно определение независимости этих маршрутов. Два квазимаршрута (слабых маршрута) частично R-независимы, если не существует участка ребра (части ребра, инцидентной одной ветви), принадлежащего обоим квазимаршрутам (слабым маршрутам).
4.2. Две вершины $x, y \in X$ в гинерсети S k-соединимы (k-квазисоединимы, слабо k-соединимы), если эти вершины соединены k-независимыми по вершинам маршрутами (квазимаршрутами, слабыми маршрутами). Аналогично определяются внутренняя и внешняя k-соединимость, k — V-соединимость и k — R-соединимость (маршруты должны быть соответственно внутренне, внешне независимы по вершинам, незанисимы по ветвям или ребрам).
Для квазимаршрутов и слабых маршрутов k-квазисоединимость и слабая k-соединимость по ветвям и ребрам определяются так же, как и в предыдущем случае, а частичная k-квазисоединимость и слабая k-соединимость по ребрам определяются следующим образом. Две вершины $x, y \in X$ в гиперсети $S = (X, V, R)$ k — R-квазисоединимы (слабо k — R-соединимы), если эти вершины соединены k частично R-независпмыми по ребрам квазимаршрутами (слабыми маршрутами).
4.3. Из определения соединимости непосредственно следует существование 17 задач вычисления k-соединимости пары вершин в гиперсети S. В табл. 1 приведена классификация задач поиска k-независимых (x, y)-маршрутов в смысле их принадлежности к полиномиально-вычислимым или NP-полным задачам.

Таблица 1.
Классификация задач вычисления k -соединимости в гиперсетях

Подклассы задач вычисления k-соединимости	Маршруты	Квазимаршруты	Слабые маршруты
Соединимость	NP-полная	NP-полная	P
Внутренняя соединимость	P	P	P
Внешняя соединимость	NP-полная	NP-полная	P
V-соединимость	NP-полная	NP-полная	P
R-соединимость	P	P	P
Частичная R-соединимость	—	P	P

4.4. Задачи вычисления k-соедини-мости и k — V-соединимости пары вершин в произвольной гиперсети $S = (X, V, R)$ являются NP-полными. Покажем, что остальные принадлежат классам (по сложности вычисления), указанным в табл. 1. Попутно отметим, что задача вычисления частичной k — R-соединимости для маршрутов не имеет смысла, так как в маршруте каждое ребро полностью ему принадлежит.

Из определения внутренней k-соединимости следует независимость соответствующих маршрутов на первичной сети $PS = (X, V)$ гиперсети $S=(X, V, R)$, т. е. задача вычисления внутренней k-соединимости принадлежит классу P. NP-полная задача вычисления k-соединимости тривиально сводится к задаче вычисления внешней k-соединимости пары вершин в гиперсети S. Действительно, по гиперсети $S = (X, V, R)$ построим гиперсеть $S' = (X' \cup X, V' \cup V, R)$, применив к каждой вершине $x \in X$ операцию выделения инцидентной вершины (см. рис. 25).

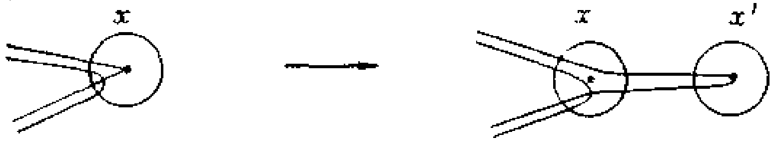


Рис. 25. Выделение инцидентной вершины.

Легко увидеть, что любым k-независимым (s, t)-маршрутам в гиперсети S соответствуют внешне k-независимые (s, t-маршруты, и наоборот. Поскольку задача поиска внешне k-независимых маршрутов принадлежит классу NP, тем самым доказана NP-полнота указанной задачи. Очевидно, что задача вычисления k — R-соединимости полиномиально вычислима, так как непосредственно сводится к задаче поиска k-независимых по ребрам (s, t)-маршрутов во вторичной сети $WS = (X, R)$.

4.5. Для квазимаршрутов имеют место следующие результаты.

Задача $XK(VK)$. Пусть заданы гиперсеть $S = (X, V, R)$, пара вершин $x, y \in X$ и целое положительное число k. Верно ли, что вершины x и y k-квазисоединимы (k— V-квазисоединимы) ?

Теорема 7. *Задачи XK и VK являются NP-полными.*

Доказательство. Очевидно, что задачи XK и VK принадлежат классу NP. Осталось показать, что NP-полная задача полиномиально сводима к задаче XK (для задачи VK доказательство строится аналогично).

Пусть заданы гиперсеть $S=(X, V, R)$ и пара вершин $x,y \in X$. Преобразуем гиперсеть $S = (X, V, R)$ в $S' = (X', V', R)$ следующим образом. К каждой ветви $v=(a, b) \in V$ добавим две инцидентные вершины a' и b', и каждую из вершин a и b разобьем на две a^+, a^-; b^+, b^- (рис. 26) так, чтобы ребра, инцидентные вершинам a и b, стали инцидентными a^+ и b^+, а ребра, слабо инцидентные a и b, оказались слабо инцидентными a_- и b^-.

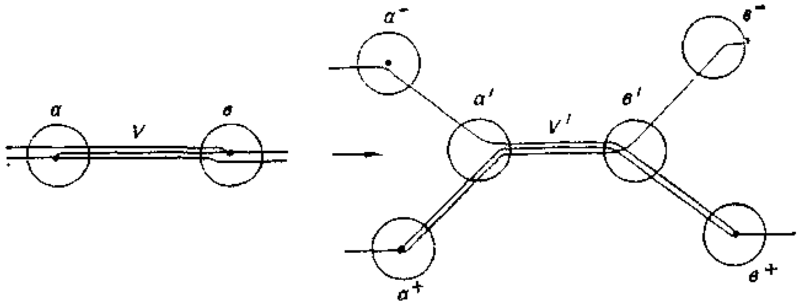


Рис. 26. Подразбиение вершины на инцидентную и слабо инцидентную.

В полученной гиперсети $S' = (X', V', R)$ каждому квазимаршруту в точности соответствует один маршрут в S, и наоборот. Действительно, в гиперсети S' содержится только два типа вершин: (a^-, a', b^-, b') слабо инцидентны ребрам, (a^+, b^+) инцидентны ребрам. Поэтому других квазимарпгрутов (кроме совпадающих с маршрутами) в этой гиперсети нет.

Таким образом, решив задачу XK, этим решим и NP-полную задачу вычисления k-соединимости вершин в гиперсети S. Аналогично доказывается NP-полнота задачи VK, так как каждой ветви v в гиперсети S сопоставлена в точности одна ветвь v' в гиперсети S'. Теорема доказана.

Теорема 8. *Задача вычисления k — R-квазисоединимосги пары вершин $x, y \in X$ в гиперсети $S = (X, V, R)$ полиномиально вычислима.*

Доказательство. Гиперсети $S = (X, V, R)$ сопоставим ультраграф $US = (X, R; f, g)$ (см. п. 3.4), тогда k-независимым по ребрам квазимаршрутам между вершинами x и y в гиперсети S будут взаимно однозначно соответствовать k-независимые по ребрам сильные маршруты в ультраграфе US. Но задача k-соединимости по ребрам в ультраграфе полиномиально вычислима. Теорема доказана.

Теорема 9. *Задача вычисления частичной k — R-квазисоединимости пары вершин в гиперсети $S =(X, V, R)$ полиномиально вычислима.*

Доказательство. Гиперсети $S = (X, V, R)$ сопоставим смешанный граф $G=(X \cup Y, E)$, полученный из данной гиперсети по следующему правилу. К каждой вершине $x \in X$ добавляются $y_1, ..., y_{k_x}$, где k_x — слабая k-степень вершины x. Каждая вершина y_i подразбивает соответствующее ребро, слабо инцидентное x. От y_i к вершине x добавляются две дуги (y_i, x). Пример такого преобразования показан на рис. 27.

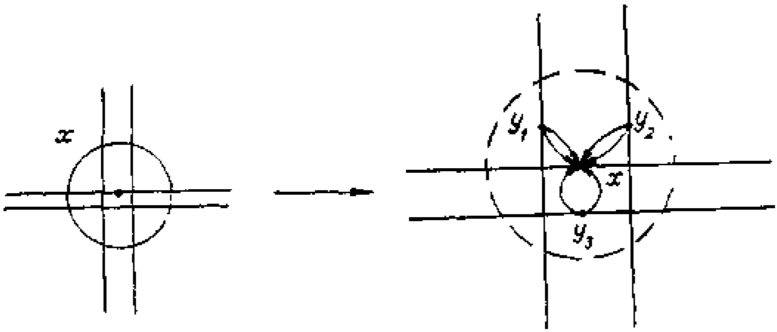


Рис.27. Локальное преобразование гиперсети в смешанный граф.

Легко показать, что любому квазимаршруту в гиперсети $S =(X, V, R)$ взаимно однозначно соответствует ориентированный маршрут в смешанном графе $G=(X \cup Y, E)$ между вершинами из множества X. Но в смешанном графе задача поиска k-независимых по ребрам и дугам маршрутов между парами вершин полиномиально вычислима. Теорема доказана.

Используя преобразование гиперсети S, изложенное в доказательстве теоремы 9, легко показать, что задача поиска внутренне k-независимых (s, t) -квазимаршрутов полиномиально вычислима.

В п. 4.4 гиперсеть *S'*, построенная для доказательства *NP*-полноты задачи внешней *k*-соединимости, может быть использована для доказательства *VP*-полноты задачи внешней *k*-квазисоединимости. Действительно, в гиперсети *S'* каждый квазимаршрут является маршрутом, и наоборот.

Теперь перейдем к рассмотрению слабых маршрутов и соответствующих задач соединимости.

4.6. Для слабых маршрутов справедливы следующие теоремы.

Теорема 10. *Задачи вычисления слабой k-соединимости и слабой k — V-соединимости пары вершин x, y* $\in X$ *в гиперсети S = (X, V, R) полиномиально вычислимы.*

Доказательство. Из определения слабых маршрутов ясно, что любому из них в гиперсети однозначно соответствует маршрут в первичной сети *PS' =(X, V'),* которая получается из *PS* удалением ветвей, которым не инцидентно ни одно ребро. Отсюда и следует утверждение теоремы.

Теорема 11. *Задача вычисления слабой k — R-соединимости пары вершин x, y* $\in$ *X в гиперсети S = (X, V, R) полиномиально вычислима.*

Доказательство. Из определения слабой *k* — *R*-соединимости следует, что любое ребро $r \in R$ входит в слабые маршруты не более одного раза. Поэтому гиперсети *S*=(*X, V, R)* сопоставили гиперграф *HS=(X, R)* (см. п. 3.5), в котором любым *k*-независимым по ребрам *(x, y)* -маршрутам в точности соответствуют *k*-независимые по ребрам *(x, y)* -маршруты в гиперсети *S.* Так как задача *k*-соединимости по ребрам в любом гиперграфе полиномиально вычислима, то из вышесказанного следует доказательство теоремы.

Теорема 12. *Задача вычисления частичной слабой k — R-соединимости пары вершин x, y*$\in$X *в гиперсети S = (X, V, R) полиномиально вычислима.*

Доказательство. Так как в частично *R*-независимых слабых маршрутах любое ребро может по частям войти в различные слабые маршруты, то для доказательства теоремы достаточно построить такой граф *L=(X, E),* в котором любому маршруту взаимно однозначно соответствует слабый маршрут в гиперсети *S.* Такой граф строится по гиперсети *S = (X, V, R)* следующим образом. Все слабо инцидентные вершины для любого ребра сделать инцидентными, в полученной гиперсети *S** рассмотреть вторичную сеть *WS* = (X, E)* и отождествить граф *L* со вторичной сетью *WS*.* Очевидно, что *L* является мультиграфом и *k*-соединимость по ребрам между вершинами *x, y* $\in$ *X* в *L* равна частичной *k* — *R*-соединимости этой пары вершин в гиперсети *S.* Теорема доказана.

Теорема 13. *Задача вычисления слабой внутренней k-соединимости пары вершин x, y ∈ X в гиперсети S=(X, V, R) полиномиально вычислима.*

Доказательство. Преобразуем гиперсеть *S = (X, V, R)* в граф *G(S)* = (*X* ∪ *X', E)* по следующему правилу. Каждой вершине $x \in X$ сопоставим $2\sigma_R(x)+1$ вершин *Y(x),* где $\sigma_R(x)$ - слабая *R*-степень вершины *x.* На множестве *Y(x)* строится полный граф. Все ребра, инцидентные вершине *x* ∈ *X,* остаются инцидентными одной и той же вершине $z \in Y(x)$. Ребра гиперсети, слабо инцидентные *x,* подразбиваются парой смежных вершин из *Y(x).* Пример такой операции приведен на рис. 28.

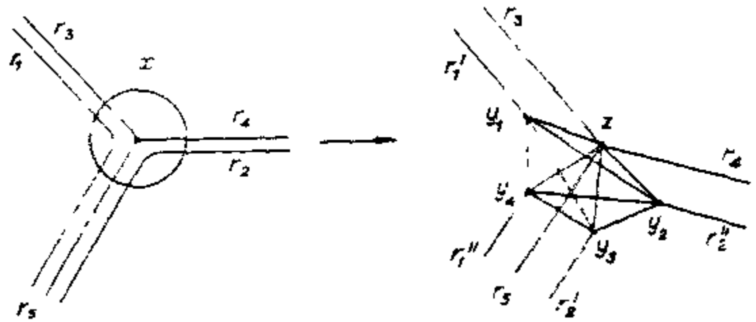


Рис. 28. Преобразование вершин гиперсети в клики слабой инциденпии.

В построенном графе *G(S)* между любой парой вершин z_i и z_j любому множеству непересекающихся цепей взаимно однозначно соответствует множество слабых внутренне независимых маршрутов между соответствующими вершинами в гиперсети *S.* Следовательно, по графу *G(S)* за полиномиальное время можно вычислить слабую внутреннюю *k*-соединимость пары вершин в гиперсети *S,* Теорема доказана.

Аналогичным образом доказывается полиномиальная вычислимость задачи поиска слабой внешней *k*-соединимости пары вершин в гиперсети *S.* В этом случае гиперсеть *S* преобразуется в граф *L(S)* с помощью операции, показанной на рис. 29.

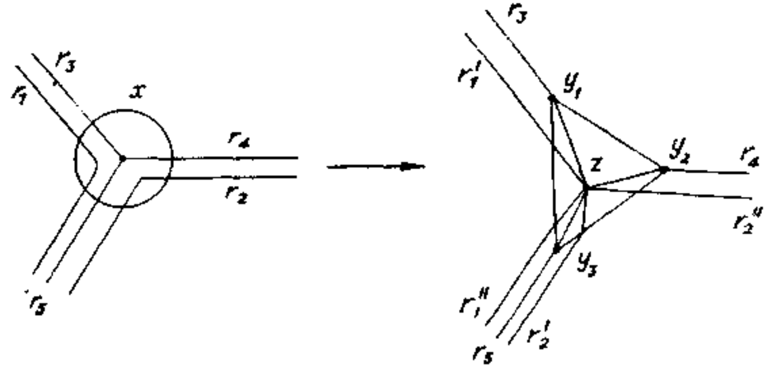


Рис. 29. Преобразование вершин гиперсети в клики инциденции.

4.7. Некоторые *N*P-полные задачи поиска *k*-соединимости пары вершин в гиперсети *S* можно решить за полиномиальное время на важных и нетривиальных классах гиперсетей. Рассмотрим *(R>X)*-гиперсети, в которых каждой вершине строго слабо инцидентно не более одного ребра. Тогда для таких гиперсетей справедлива следующая теорема.

Теорема 14. *В (R>X)-гиперсетях задача поиска внешней k-соединимости (k-квазисоединимости) пары вершин полиномиально вычислима.*

Доказательство. Так как каждой вершине $x \in X$ слабо инцидентно не более одного ребра, то любые независимые по ребрам *(s, t)*-маршруты будут также внешне независимы и, следовательно, для решения задачи можно воспользоваться графом вторичной сети *WS* гнперсети *S*. В случае квазимаршрутов по гиперсети *S* можно построить смешанный граф *G* (теорема 9) и тогда любым непересекающимся по ребрам (*s, t)*-цепям в *G* взаимно однозначно соответствуют внешне независимые (*s, t)*-маршруты в гиперсети *S*. Теорема доказана.

Можно предположить, что для *(R > X)* -гиперсетей задачи поиска *k*-соединимости и *k*-квазисоединимости также полиномиально вычислимы. Используя преобразования гиперсети *S,* приведенные в теореме 14, легко можно показать, что задачи вычисления *k* — *V*-соединимости и *k* — *V*-квазисоединимости пары вершин в *(R→ V)* -гиперсети *S* решаются за полиномиальное время.

7.7.5. Отделимость и связность гиперсетей

Связность любой структурной модели определяется, с одной стороны, способом достижимости вершин (т. е. типом маршрута), а с другой — типом и характером удаления элементов из структурной модели.

5.1. Гиперсеть называется связной, если и только если между любой парой вершив гиперсети S существует соединяющий их маршрут. Отсюда следует, что графы PS и WS связны. Гиперсеть называется односторонне квазисвязной, если и только если любая пара вершин из S соединима хотя бы одним квазимаршрутом. Гиперсеть квазисвязна, если и только если для любых вершин $x, y \in X$ существуют квазимаршруты $\mu(x, y)$ и $\mu(y, x)$. Гиперсеть называется слабо связной, если и только если любая пара вершин S соединима слабым маршрутом.

Рассмотренные отношения связности упорядочиваются по включению, т. е. из связности → квазисвязность → слабая связность. Обратное неверно.

Гиперсеть S называется насыщенной, если гиперграф FS связен, и ненасыщенной — в противном случае. Гиперсеть S называется полносвязной, если и только если связны графы PS, WS и гиперграф FS. Очевидно, что несвязность S влечет несвязность WS.

5.2. Понятия k-отделимости в гиперсетях связаны со способом удаления элементов. Разумеется, можно ввести различные способы удаления элементов. Здесь рассмотрим лишь те, которые отражают возможные преобразования структур систем, моделируемых гиперсетями.

1. Удаление ребер. Удаляются ребра в гиперсети $S =(X, V, R)$ без инцидентных вершин и ветвей.
2. Частичное удаление ребер заключается в удалении части ребра, инцидентной некоторой ветви. При этом ребро разделяется по крайней мере на две части.
3. Удаление ветвей. Удаляются ветви с инцидентными ребрами без инцидентных вершин.
4. Частичное удаление ветвей определяется аналогично, но ребра удаляются частично.
5. Внешнее удаление вершин. Удаляются ребра, слабо инцидентные данным вершинам.
6. Внутреннее удаление вершин. Удаляются ребра, инцидентные данным вершинам.
7. Удаление вершин. Удаляются вершины вместе с инцидентными ветвями (ребра, инцидентные удаленным ветвям, также удаляются).

Из введенных способов удаления вершин можно рассмотреть частичное; остановимся на одном способе удаления.

8. Частичное удаление вершины. Определяется аналогично предыдущему способу, но ребра, инцидентные удаляемым ветвям, удаляются частично.

5.3. Теперь можно привести возможные определения k-отделимости пары вершин в гиперсети $S = (X, V, R)$. Две вершины $x, y \in X$ в гиперсети $S=(X, V, R)$ $k_R(l_R, m_R)$-отделимы (квази-, слабо отделимы), если $k_R(l_R, m_R)$ — наименьшее число ребер, удаление которых (для других способов удаления элементов определения отделимости аналогичны) приводит к разрушению всех маршрутов (квази-, слабых маршрутов) между вершинами x и y в гиперсети S. Таким образом, даны определения 24 понятий отделимости пары вершин.

k-связность гиперсетей определяется через приведенные понятия отделимости.

Пример. Гиперсеть $S = (X, V, R)$ k_x-связна, если k_x принимает наименьшее значение из k_x-отделимости всех пар вершин гиперсети S.

Часть гиперсети S' назовем остовной, если из исходной гиперсети $S = (X, V, R)$ удалено некоторое подмножество ребер R', т. е. $S' =(X, V, R-R')$. Часть гиперсети S' назовем усеченной остовной, если из исходной гиперсети $S = (X, V, R)$ удалено некоторое множество ветвей V' (а следовательно, и инцидентные им ребра R'), т. е. $S' =(X, V- V', R — R')$. При внешнем или внутреннем удалении вершин, очевидно, будет получена остовная гиперсеть. Часть гиперсети S' назовем подгиперсетью (или просто подсетью), если из исходной гиперсети $S = (X, V, R)$ удалено некоторое множество вершин X' (а следовательно, и инцидентные им ветви V' и ребра R'), т.е. $S'=(X-X',V-V', R-R')$.

5.4. Для получения оценок связности гиперсети $S = (X, V, R)$| будут полезны некоторые неравенства, частично упорядочивающие численные значения характеристик связности.

Обозначим: $\Theta:= k, l, m,$ тогда Θ_R (S)—реберная связность гиперсети S при $\Theta = k$, реберная квазисвязность гиперсети S при $\Theta = l$, реберная слабая связность гиперсети S при $\Theta = m$. Для остальных характеристик связности Θ также принимает три значения. $H := X, V, R,$ тогда $k_H(S)$ — связность гиперсети S при $H=X$, связность по ветвям гиперсети S при $H = V$, связность по ребрам при $H = R$. σ= 1, 2, ..., 8 — номер способа удаления элементов из гиперсети S (см. п. 5.2). Тогда, например, $l_3(S)$ означает l_V-квазисвязность гиперсети S по ветвям, а m_6 (S)— внутреннюю слабую связность гиперсети S. Черта над буквой означает частичную отделимость (см. также табл. 2).

Справедливы следующие неравенства:

$$\forall \sigma = \overline{1,8} \qquad m_\sigma(S) \geqslant l_\sigma(S) \geqslant k_\sigma(S), \tag{1}$$

$$\forall \Theta = k,\, l,\, m \quad \Theta_R(S) \geqslant \Theta_V(S) \geqslant \Theta_X(S), \tag{2}$$

$$\overline{\Theta}_R(S) \geqslant \overline{\Theta}_V(S) \geqslant \overline{\Theta}_X(S), \tag{3}$$

$$\Theta_X^0(S) \geqslant \Theta_X(S), \tag{4}$$

$$\Theta_X^I(S) \geqslant \Theta_X(S), \tag{5}$$

$$\Theta_R(S) \geqslant \Theta_X^I(S), \tag{6}$$

$$\forall\, H = X,\, V,\, R;\ \Theta = k,\, l,\, m\ \overline{\Theta}_H(S) \geqslant \Theta_H(S)$$

$$(\text{при } \Theta = k \text{ имеет место равенство}). \tag{7}$$

На рис. 30 значения характеристик связности гиперсети частично упорядочены согласно приведенным неравенствам.

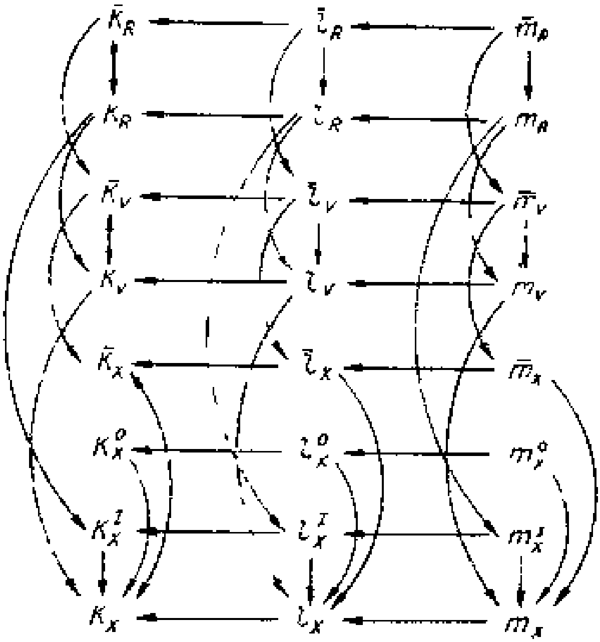


Рис. 30. Граф частичной упорядоченности характеристик связности.

7.7.6. О сложности вычисления отделимости в гиперсетях

Так же как в случае соединимости пары вершин в гиперсети *S,* задачи их отделимости имеют различную сложность. Некоторые из этих задач решаются за полиномиальное время, другие принадлежат к классу *NP*-полных задач. В табл. 2 указаны идентификаторы задач отделимости и их соответствующий класс сложности вычисления.

Таблица 2.

Класификация задач вычисления *k*-отделимости в гиперсетях

Удаление элементов	Маршрут		Квазимаршрут		Слабый маршрут	
Ребра	k_R	*P*	l_R	*P*	m_R	*P*
Ребра, частичное	$\bar{k}_R$	*P*	$\bar{l}_R$	*P*	$\bar{m}_R$	*P*
Ветви	k_V	*NP*-полная	l_V	*NP*-полная	m_V	*NP*-полная
Ветви, частичное	$\bar{k}_V$	*NP*-полная	$\bar{l}_V$	*NP*-полная	$\bar{m}_V$	*P*
Вершин, внешнее	k^o_X	*NP*-полная	l^o_X	*NP*-полная	m^o_X	*NP*-полная
Вершин, внутреннее	k^I_X	*P*	l^I_X	*NP*-полная	m^I_X	*NP*-полная
Вершин	k_X	*NP*-полная	l_X	*NP*-полная	m_X	*NP*-полная
Вершин, частичное	$\bar{k}_X$	*NP*-полная	$\bar{l}_X$	*NP*-полная	$\bar{m}_X$	*P*

6.1. Отделимость. Напомним, две вершины x, $y \in X$ в гиперсети *S* отделимы, если при удалении некоторых элементов гиперсети между вершинами *x* и *y* отсутствует маршрут, их соединяющий.

Задача k_R-отделимости решается за полиномиальное время, к этому же классу относится задача внутренней k^I_X - отделимости. Так же показана *NP*-полнота задач: k_V-, k^0_X-, k_X-отделимости. Действительно, полиномиальность первых двух задач следует из равенств

$k_R(S) = \lambda(WS)$, $k^I_X = \omega(WS)$ соответственно, где $\lambda(WS)$—отделимость по ребрам соответствующей пары вершин в графе вторичной сети *WS* гиперсети *S*, а $\omega(WS)$—отделимость по вершинам этой же пары вершин в *WS*. Справедливость утверждений (см. табл. 2) о сложности вычисления частичной отделимости пары вершин в гиперсети *S* следует из справедливости равенства (7) для $\Theta = k$.

6.2. Квазиотдолимость. Задача вычисления l_R-квазиотделимости может быть решена за полиномиальное время. По гиперсети $S=(X,V, R)$ строится ультраграф $US = (X, R; f, g)$ (см. п. 3.4), в этом ультраграфе любому сильному маршруту взаимно однозначно соответствует квазимаршрут в гиперсети $S = (X, V, R)$. Но k-соединимость по ребрам в ультраграфе US равна k-отделимости по ребрам (при слабом удалении ребер). Отсюда следует полиномиальность вычисления l_R-квазиотделимости пары вершин в гиперсети S. Аналогично доказывается полиномиальная вычислимость частичной $\overline{l}_R$-квазиотделимости, но вместо ультраграфа US рассматривается граф $G = (X \cup Y, E)$ (см. теорему 9).

Доказана NP-полнота задач вычисления l_V-, $l^0{}_X$-, $l^I{}_X$- l_X -отделимости пары вершин в гиперсети $S = (X, V, R)$. Доказана NP-полнота задач частичной $\overline{l}_V$- и $\overline{l}_X$-квазиотделимости. Действительно, при доказательстве NP-полноты задач l_V- и l_X-квазиотделимости x и y фактически можно рассматривать удаление ребер, хотя и частичное, но разрушающее все квазимаршруты из вершины x в вершину y.

6.3. Слабая отделимость. Для задач слабой m_V-, $m^0{}_X$- ,$m^I{}_X$-, m_X-отделимости двух вершин в гиперсети $S = (X, V, R)$ их NP-полнота доказана В.К.Попковым. Им же показано, что задача слабой m_R-отделимости решается за полиномиальное время путем преобразования гиперсети S в гиперграф $HS = (X, R)$ (см. п. 3 5). Легко показать, что задача слабой $\overline{m}_R$-отделимости также решается за полиномиальное время. Для этой цели достаточно перейти от гиперсети S к графу L (см. п. 4.6, теорема 12). Если рассмотреть первичную сеть $PS = (X, V)$ гиперсети S и удалить из нее ветви, не инцидентные ребрам, то в полученном графе $PS' = (X, V')$ задачи определения реберной λ-отделимости и вершинной ω-отделимости пары вершин $x, y \in X$ эквивалентны задачам слабой $\overline{m}_V$-отделимости и слабой $\overline{m}_X$-отделимости соответствующих вершин в гиперсети S. Таким образом, эти задача также полиномиально вычислимы.

6.4. В теории графов справедливы теоремы менгеровского типа, т. е. численное значение отделимости пары вершин равно численному значению их соединимости. Для гиперсетей это положение не всегда выполняется. Например, на рис. 31 приведена гиперсеть $S = (X, V, R)$, в которой вершины s и t 1-соединимы, но 2-отделимы.

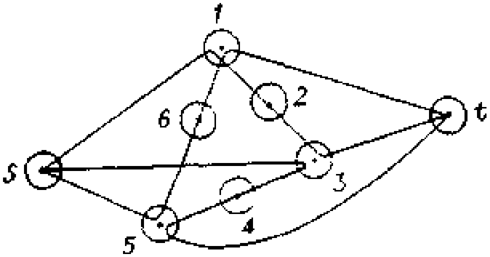


Рис. 31. Пример двухполюсной гиперсети.

Из примера ясно, что квази- и слабая соединимость отличаются от квази- и слабой отделимости. Действительно, s и t слабо 3-соединимы, но слабо 2-отделимы. Легко показать, что для полиномиально вычислимых задач соединимости и отделимости в гиперсетях теорема Менгера справедлива. Причем при исследовании $\overline{\Theta}_H$ (S)-отделимости вершин необходимо рассматривать при $H = X$ k-соединимость, при $H=V$ k — V-соединимость, при $H = R$ частичную k — R-соединимость. Покажем, что для NP-полных задач k_V-, k^0_X-, k_X-отделимости справедливы следующие неравенства:

$$k_V \geqslant z_V, \quad k_X^0 \geqslant z_X^0, \quad k_X \geqslant z_X, \tag{8}$$

где z_V равно k — V-соединимости, z^0_X— внешней k-соединимости, z_X — k-соединимости. В самом деле, так как при удалении элементов в гиперсети S должны быть разрушены (s, t)-маршруты, а по определению k-соединимости пары вершин элементы маршрутов полностью не пересекаются, то очевидно, что k-отделимость не меньше k-соединимости.

В случае слабой связности пары вершин в гиперсети $S =(X, V, R)$ имеют место обратные неравенства:

$$m_V \leqslant \tilde{z}_V, \quad m_X^0 \leqslant \tilde{z}_X^0, \quad m_X^I \leqslant \tilde{z}_X^I, \quad m_X \leqslant \tilde{z}_X, \tag{9}$$

где m — слабая отделимость, а $\tilde{z}$ — слабая соединимость одной и той же пары вершин s, t в гиперсети S. Справедливость этих неравенств следует из того факта, что при удалении одного элемента в гиперсети S может быть разрушено несколько независимых слабых (s, t)-маршрутов.

Соотношения квазиотделимости и квазисоединимости почти не исследованы, но легко можно показать, что l_V-, l^0_X-, l_X-квазиотделимость не оценивается через соответствующую квазисоединимость. Так как слабая соединимость легко вычисляется, то с учетом неравенств (9) и (1) — (7) можно оценивать сверху различные типы отделимости пары вершин в гиперсети *S*.

6.5. Сложность вычисления большинства задач *k*-отделимости в гиперсетях не позволяет исследовать характеристики связности гиперсетей большой размерности уже при $k > 3$. Однако для некоторых классов гиперсетей задачи вычисления *k*-отделимости пары вершин решаются за полиномиальное время.

Теорема 15. *Задача вычисления внутренней k-квазиотделимости пары вершин в $(X \to R)$ -гиперсети $S = (X, V, R)$ полиномиально вычислима.*

Доказательство. По $(X\to R)$-гиперсети $S=(X, V, R)$ построим ориентированный граф $GS=(X, E)$ следующим образом. Множество вершин орграфа *GS* совпадает с множеством вершин в *S*. Пара вершин *x* и *y* в *GS* соединяется дугой *(x, y)* в *GS*, если и только если в гиперсети *S* существует ребро $r \in R$, сильно инцидентное вершине *x* и слабо инцидентное вершине *y*. Так как в $(X\to R)$-гиперсети каждому ребру сильно инцидентна не более чем одна вершина, то при внутреннем удалении произвольной вершины *x* удалению инцидентных ей ребер в *S* соответствует удаление дуг, исходящих из вершины *x* в орграфе *GS*. С другой стороны, любому квазимаршруту в *S* взаимно однозначно соответствует ормаршрут в *GS* между той же парой вершин. Следовательно, разделяющее множество вершин в *GS* является разделяющим множеством (внутреннее удаление вершин) для тех же квазидостижимых вершин в гиперсети *S*. Теорема доказана.

Теорема 16. *Задача вычисления k — V-отделимости пары вершин в $(R \to V)$-гиперсети $S=(X, V, R)$ полиномиально вычислима.*

Доказательство. В $(R \to V)$-гиперсети *S* каждой ветви инцидентно не более чем одно ребро, поэтому $k_V(S)\geq \lambda(WS)$. С другой стороны, пусть $\lambda(WS)$—минимальное разделяющее множество ребер вторичной сети $WS = (X, B)$ $(R \to V)$-гиперсети $S = (X, V, R)$. Тогда в силу утверждения теоремы каждому ребру *r* из разделяющего множества *R'* можно сопоставить единственную инцидентную ему ветвь $v \in V$, удаление которой разрушает данное ребро *r*. Отсюда следует $\lambda(WS)\geq k_V(S)$. Теорема доказана.

Можно предположить, что задачи вычисления *k* — *V*-квазиотделимости и слабой *k* — *V* -отделимости пары вершин в $(R \to V)$-гиперсети также решаются за полиномиальное время. Например, если $(R \to V)$-гиперсеть *S* одновременно является $\{V \to R\}$-гиперсетью, то полиномиальность вычисления *k* — *V* -квазиотделимости пары вершин в *S* можно показать, переходя от гиперсети *S* к смешанному графу *G* из п. 4.5.

Теорема 17. *Задача вычисления внешней k-отделимоста пары вершин в $(R > X)$ -гиперсети решается за полиномиальное время.*

Доказательство. Аналогично доказательству теоремы 16 за тем лишь исключением, что разделяющему (внешнее удаление вершины) множеству вершин в S взаимно однозначно соответствует множество разделяющих ребер в графе $\overline{WS}=(X', R')$, который получается из вторичной сети $WS = (X, R)$ гиперсети S отождествлением вершин, инцидентных такому ребру r, для которого не существует слабо инцидентных вершин из X. Очевидно, что в общем случае некоторые вершины внешне k-неотделимы при любом k. В этом случае в графе $\overline{WS}$ эти вершины будут отождествлены. Теорема доказана.

Легко показать, что для некоторых тривиальных классов гиперсетей (например, для $(V \to R)$-гиперсетей) все задачи вычисления k-отделимости решаются за полиномиальное время.

7.7.7. Задачи синтеза оптимальных гиперсетей с заданной связностью

Эти задачи имеют большое практическое применение. Поэтому, несмотря на NP-полноту большинства задач синтеза гиперсетей, в работе приводятся некоторые эвристические и точные алгоритмы синтеза гиперсетей, удовлетворяющих заданным характеристикам связности. Кроме того, анализируются наиболее общие схемы методов, используемых при составлении различных алгоритмов синтеза.

7.1. Рассмотрим основные параметры задач синтеза.

1°. Исходные данные. В задачах синтеза могут быть заданы: множества вершин X и Y, первичная сеть $PS = (X, V)$, вторичная сеть $WS = (Y, R)$, однозначное отображение f: $Y \to X$ $(|X| \geq |Y|)$. Каждое из приведенных данных может либо присутствовать в постановке задачи, либо отсутствовать. Наличие множеств X и Y обязательно. Кроме того, ветви, ребра и вершины гиперсети могут быть взвешены. Последние данные обычно фигурируют в задачах оптимального синтеза.

2°. Характеристики связности (отделимости) описаны в п. 5. Таким образом, синтезируемая гиперсеть должна обладать заданными значениями одной или нескольких характеристик связности.

3°. Отделимость вершин устанавливается выделенными вершинами или парой вершин.

4°. Значение связности устанавливается максимальное или заданное.

5°. В том случае, когда заданы первичная сеть $PS=(X, V)$ и вторичная $WS=(Y, R)$, для достижения необходимого значения связности синтезируемой гиперсети возможны добавления ребер, ветвей или распараллеливание ребер.

6°. В тех случаях, когда неизвестны первичная или вторичная сети, возможны ограничения на число их ветвей, ребер, диаметр и другие характеристики. Кроме того, в синтезируемой гиперсети могут быть ограничения на число элементов, инцидентных другим элементам гиперсети. Например, число ребер, инцидентных любой ветви, не превосходит заданной величины.

7°. В процессе синтеза гиперсети с заданной связностью часто возникает необходимость в оптимизации значения некоторой целевой функции, в частности, она может отсутствовать. Таким образом, предложенная классификация задач синтеза гиперсетей позволяет сформулировать весь круг проблем, связанных с оптимизацией гиперсетей, удовлетворяющих данным характеристикам связности.

7.2. Для любой задачи синтеза необходим собственный алгоритм, однако можно предложить общие схемы решения поставленных задач, которые могут оказаться полезными при составлении частных алгоритмов. Все методы синтеза оптимальных структур гиперсетей можно подразделить на точные и приближенные. К приближенным методам прибегают в том случае, когда точные методы недостаточно эффективны (не дают решение за приемлемое время).

Точные методы:

1°. Метод ветвей и границ. Применительно к синтезу гиперсетей ветвление можно осуществлять, задавая для каждой ветви v либо $v \in F(R)$, либо $v \notin F(R)$. Такая схема может применяться в тех случаях, когда знание множества $F(R)$-ветвей, вошедших в оптимальную гиперсеть, позволяет эффективно построить ее самое.

2°. Метод направленного перебора. Суть метода направленного перебора заключается в следующем. Для каждого ребра $r \in R$ фиксируется множество цепей, которые могут фигурировать в оптимальной гиперсети в качестве $F(r)$. После этого осуществляется перебор всех вариантов структур гиперсетей. Применимость метода в значительной степени ограничивается необходимостью хранения в памяти ЭВМ множества допустимых цепей.

3°. Метод сведения задачи к уже известной. В некоторых случаях задачу синтеза оптимальной структуры гиперсети удается свести к уже известной, хорошо исследованной задаче. Ограниченность сферы применения этого метода очевидна.

Приближенные методы:

1°. Метод кратчайшего пути. Для каждого ребра $r \in R$ в качестве $F(r)$ выбирается кратчайший путь в графе PS между концевыми вершинами ребра r.

2°. Метод независимых путей. Если в графе вторичной сети *WS* имеется *k* кратных ребер между вершинами *x, y* и, кроме того, локальная связность вершин *x, y* в графе первичной сети *PS* не меньше *k*, то эти *k* кратных ребер можно реализовать по *k* независимым путям. Для того чтобы суммарная длина этих путей была минимальна, необходимо каждой ветви графа *PS* приписать «пропускную способность», равную длине этой ветви; затем воспользоваться алгоритмом нахождения потока минимальной стоимости.

3[е]. Метод кратчайшего пути с адаптацией. От метода кратчайшего пути он отличается тем, что после нахождения очередной трассы *F(r)* длины ветвей, входящих в *F(r),* изменяются по некоторому правилу. Если длины ветвей уменьшаются, то получающаяся гиперсеть имеет тенденцию к «сжатию» (уменьшению числа ветвей, по которым реализованы ребра). Если же длины ветней увеличиваются, то получающаяся гиперсеть имеет тенденцию к «расширению» (каждое последующее ребро реализуется по пути, который должен иметь как можно меньше общих ветвей с путями реализации предыдущих ребер).

4°. Метод случайного поиска с адаптацией. Каждой ветви первичной сети *PS = (X, V)* ставится в соответствие некоторая вероятность удаления данной ветви из *PS.* Разыгрывается *t* вариантов удаления ветвей из *PS* согласно заданным вероятностям. В полученных *t* графах *{PS}* производится трассировка ребер *WS* (например, методом кратчайшего пути). Найденные гиперсети $\{S_i\}$ оцениваются с помощью целевой функции. Ветви PS_i, для которых значение целевой функции наилучшее, поощряются, т. е. вероятность удаления соответствующих ветвей из *PS* уменьшается, а для ветвей, входящих в PS_j с наихудшим значением целевой функции, она увеличивается. Процедура построения *t* случайных суграфов *PS* повторяется до тех пор, пока относительная погрешность не станет меньше заданной величины или процесс не стабилизируется.

5°. Метод гомеоморфного отображения. Сущность метода заключается в том, что отдельные фрагменты *WS* или вторичная сеть в целом отображаются в *PS* так, чтобы часть первичной сети *PS* (ветви этой части инцидентны ребрам *WS)* была гомеоморфной вторичной сети *WS.* Если такая трассировка возможна, то ω (*S*) = ω *(WS),* причем здесь имеется в виду связность вершин, принадлежащих только *WS*.

6°. Метод замены трасс. Сначала трассировка ребер из *WS* осуществляется в *PS* произвольным способом (например, методом кратчайших путей), затем осуществляется перебор по всем ребрам вторичной сети *WS,* т. е. берется произвольное ребро, соответствующая ему трасса уничтожается и осуществляется новая трассировка с учетом оптимизации целевой функции и заданных ограничений. При такой трассировке можно использовать различные эвристические приемы или точные методы.

7°. Метод локальной оптимизации. Сначала строится некоторое приближенное решение задачи синтеза оптимальной структуры гиперсети. Затем по некоторым правилам перестраиваются отдельные фрагменты гиперсети. Если получающаяся гиперсеть оптимальнее старой, то она берется в качестве текущего приближенного решения и процедура повторяется заново уже по отношению к ней. Процесс прекращается, когда не удается построить более оптимальную гиперсеть.

8°. Усеченные точные методы. На базе точных методов (метод ветвей и границ, метод направленного перебора) можно строить приближенные методы решения задачи синтеза оптимальной структуры гиперсети, если в точном методе проводить не весь объем вычислений, а некоторую заранее обусловленную его часть. Например, в методе ветвей и границ можно осуществлять ветвления до тех пор, пока наилучшая из построенных гиперсетей не окажется достаточно близкой к оптимальной.

7.3. Об одном алгоритме синтеза оптимальной гиперсети. Перейдем к описанию эвристического алгоритма решения задачи синтеза, дающего неплохие результаты на достаточно широком классе задач. Сформулируем необходимые определения. Пусть известны графы первичной сети $PS=(X, V)$, вторичной сегя $WS=(Y, R)$. Обозначим: $\rho(v)$ или $\rho_{PS}(v)$ — длина ветви $v \in V$, $\rho(r)$ или $\rho_{WS}(r)$— длина ребра $r \in R$, $\alpha(r)$—пропускная способность (емкость) ветви $v \in V$, $\alpha(r)$—емкость ребра $r \in R$, $\omega(S)$ — вершинная связность гиперсети S.

Требуется построить гиперсеть $S = (PS, WS; \Phi)$, для которой выполняются условия

$$\omega(S) \geqslant k, \tag{10}$$

$$\forall v \in V \sum_{r \in \Phi^{-1}(v)} \alpha(r) \leqslant \alpha(v) \tag{11}$$

и которая минимизирует функционал $\varphi(S)$—стоимость гиперсети S.

Относительно вида функционала φ(S) не формулируется никаких ограничений, поскольку он явным образом не фигурирует в предлагаемом ниже эвристическом алгоритме. Фактически алгоритм осуществляет поиск некоторого допустимого (удовлетворяющего выше описанным ограничениям) решения. При практическом использовании рекомендуется убедиться в том, что для данного конкретного функционала $\varphi(S)$ алгоритм дает приемлемые результаты.

Алгоритм А1:

п. 1. Если $\omega(PS)\geq k$ и $\omega(WS)\geq k$, то к п. 2, иначе к п. 15.

п. 2. Между всеми парами вершин $x,\ y\ \in\ X$ найти максимальные потоки в графах *PS* и *WS*, т. е. вычислить $\mu_{PS}(x,\ y)$ и $\mu_{WS}(x,\ y)$. Если существует пара вершин $x,\ y\ \in\ X$, для которой $\mu_{PS}(x,\ y) < \mu_{WS}(x,\ y)$, то к п. 15, иначе к п. 3.

п. 3. Реализуем ребра графа *WS* по кратчайшим путям в графе *PS*. Если выполняется условие (11), то к п. 8, иначе к п. 4.

п. 4. Для всех ветвей v_i вычислить коэффициент $\Delta_i = \sum\limits_{r\in\Phi^{-1}(v_i)} \alpha(r)/\alpha(v_i)\cdot$

Если для некоторой ветви $v_i\Delta_i>1$, то она называется перенасыщенной.

п. 5. Среди перенасыщенных ветвей выберем ветвь v_j с максимальным значением Δj и найдем ребро $r=(x,\ y)\in\Phi^{-1}(v_j)$ с минимальным значением $\alpha(r)$. Для каждой ветви v_j вычислим значение δ_j по формуле

$$\delta_j=\begin{cases}\Delta_j, \text{ если } v_j\in\Phi(r) \text{ и } \sum\limits_{r'\in\Phi^{-1}(v_j)}\alpha(r')\leqslant\alpha(v_j)-\alpha(r), \quad r'\neq r;\\ 1, \text{ если } v_j\notin\Phi(r) \text{ и } \sum\limits_{r'\in\Phi^{-1}(v_j)}\alpha(r')>\alpha(v_j)-\alpha(r), \quad r'\neq r;\\ \Delta_j-\dfrac{\alpha(r)}{\alpha(v_j)}, \text{ если } v_j\in\Phi(r) \text{ и ветвь } v_j \text{ не перенасыщена};\\ 1, \text{ если } v_j\in\Phi(r) \text{ и } v_j \text{ перенасыщена.}\end{cases}$$

Также для каждой ветви v_j положим $\rho^*(v_j)=\rho(v_j)/(1-\delta_j)$. Если $\delta_j = 1$, то очевидно, $\rho^*(v_j) = \infty$.

п. 6. Ищем кратчайший путь между вершинами $x,\ y$ в графе *PS* с весами ветвей $\rho^*(v)$. Если этот путь имеет конечный вес, то к п. 7, иначе к п. 15.

п. 7. Осуществляем перетрассировку ребра r по найденному в п. 6 кратчайшему пути. Заново вычисляем все Δ_j по формуле

$$\Delta_j=\sum_{r\in\Phi^{-1}(v_j)}\alpha(r)/\alpha(v_j).$$

Если в получившейся гиперсети отсутствуют перенасыщенные ветви, то к п. 8, иначе к п. 5.

п. 8. Если $\omega(S)\geq k$ (проверка осуществляется с помощью переборного алгоритма), то к п. 14, иначе к п. 9.

п. 9. С помощью переборного алгоритма находим минимальное сечение $\{x_1,\ ...,\ x_p\}$, $p<k$, гиперсети по вершинам.

п. 10. Найдем все ребра r, трассы которых содержат вершины из сечения $\{x_1,\ ...,\ x_p\}$, а концевые вершины лежат в разных компонентах связности S_1 и S_2 гиперсети $S\backslash\{x_1,\ ...,\ x_p\}$. Хотя бы одно такое ребро найдется, так как *WS* k-связен. Выберем среди них ребро $r=(x,\ y)$ с наибольшим значением $\alpha(r)$.

п. 11. Для каждой ветви v_j вычислим δ_j по формуле

$$\delta_j = \begin{cases} 1, \text{ если } v_j \notin \Phi(r) \text{ и } \sum\limits_{r' \in \Phi^{-1}(v_j)} \alpha(r') > \alpha(v_j) - \alpha(r), \quad r' \neq r; \\ \dfrac{1}{\alpha(v_j)} \sum\limits_{r' \in \Phi^{-1}(v_j)} \alpha(r) \text{ во всех остальных случаях.} \end{cases}$$

Также для каждой ветви δ_j положим $\rho^*(v_j) = \rho(v_j)/(1-\delta_j)$. Если $\delta_j = 1$, то $\rho^*(\delta_j) = \infty$.

п. 12. Найдем кратчайший путь между вершинами x, y в графе $PS \setminus \{x_1, ..., x_p\}$ с весами ветвей $\rho^*(v)$. Если этот путь имеет конечный вес, то к п. 13, иначе к п. 15.

п. 13. Осуществляем перетрассировку ребра связност r по найденному в п. 12 кратчайшему пути. Если исчерпаны все ребра, трассы которых содержат вершины из сечения $\{x_1, ..., x_p\}$, а концевые вершины лежат в разных компонентах связности S_1 и S_2 гиперсети $S \setminus \{x_1, ..., x_p\}$, то к п. 9, иначе к п. 10. Повторяем п. 9—13 до тех пор, пока не будет найдена гиперсеть, удовлетворяющая условиям (10), (11), или не повторится минимальное сечение. В первом случае к п. 14, во втором —к п. 15.

п. 14. Гиперсеть, удовлетворяющая условиям (10), (11), найдена.

п. 15. Допустимое решение не найдено. Применение переборного алгоритма здесь оправдано тем, что связность гиперсетей для практических целей небольшая.

7.7.8. Алгоритмы синтеза гиперсетей с заданной вершинной связностью

Задачи синтеза гиперсетей имеют различную сложность решения. Покажем на примерах, как незначительное изменение в постановке задачи влияет па сложность ее решения. Пусть заданы первичная сеть $PS = (X, V)$, для которой $\omega(PS) \geq 2$, и вторичная сеть $WS = (Y, R)$— простой цикл, и пусть $|X|=|Y|$. Необходимо найти гиперсеть $S=(PS, WS; \Phi)$, у которой вершинная связность равна 2. Если известно отображение $f: Y \to X$, то легко решается вопрос о наличии допустимого решения, в противном случае к поставленной задаче сводится задача поиска гамильтонова цикла, т. е. задача синтеза становится NP-полной.

Очевидным фактом является то, что k-связность пары вершин в гиперсети S не может превосходить связности этой же пары вершин в первичной и вторичной сетях гиперсети $S = (PS, WS; \Phi)$. Например, на рис. 31 вершины s и t 2-связны в S, но 3-связны в PS и WS. Возникает вопрос, всегда ли можно найти гиперсеть S', для которой $\omega(S') = \min(\omega(PS), (\omega(WS))$? Оказывается, что такая гиперссть не всегда существует (рис. 32).

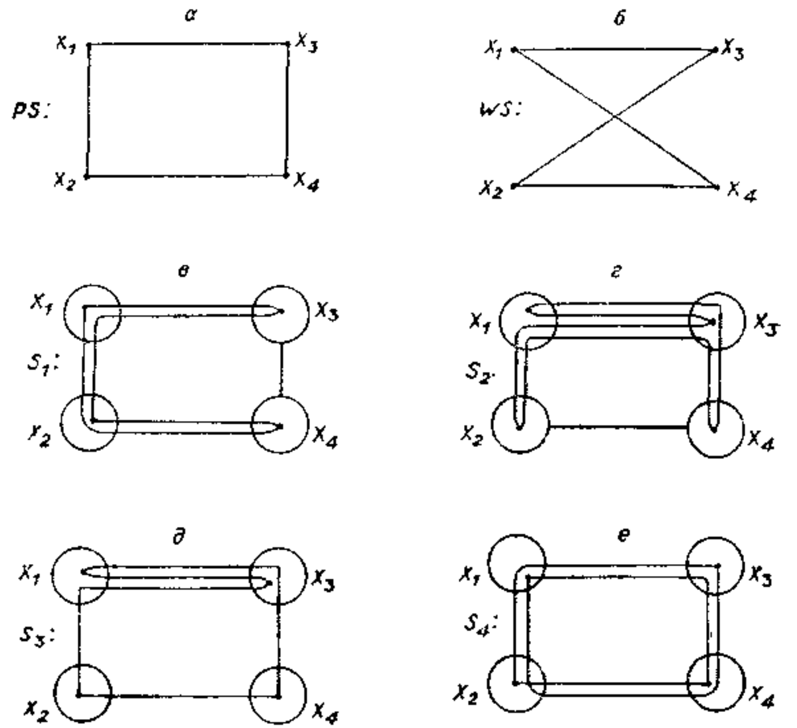


Рис. 32. Односвязные гиперсети.

Гиперсети S_1, S_2 очевидно, односвязны; S_3 разрушается при удалении вершины x_1 или x_3, а S_4 становится несвязной при удалении любой вершины. Из приведенных примеров видно, что для данных *PS* и *WS* не существует гиперсети, имеющей связность, равную двум. В связи с этим возникает следующая задача синтеза гиперсети *S*.

Пусть заданы первичная сеть $PS = (X, V)$ и вторичная $WS =(Y, R)$. Требуется найти гиперсеть $S=(X, V, R')$ такую, что
$\omega(S)= \min(\alpha(PS), (\omega(WS))$. Причем ко вторичной сети *WS* разрешается добавлять ребра. Критерий эффективности — минимум суммарной длины ребер. На рис. 33 приведены примеры гиперсетей S'_1 и S'_3, которые становятся двусвязными при добавлении новых ребер к S_1 и S_3.

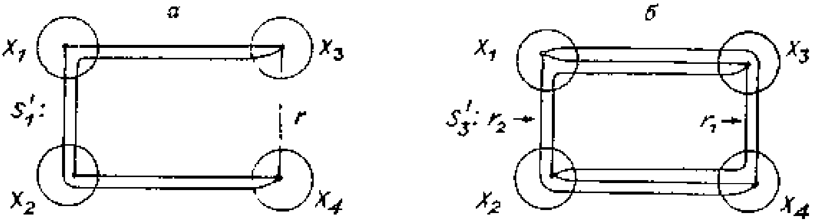


Рис. 33. Двухсвязные гиперсети.

Причем к S_1 (см. рис. 32, *в)* добавлено ребро между вершинами x_3, x_4, несмежными в *WS;* к S_3 (см. рис. 32,*д)* — ребра (x_1, x_4) и (x_2, x_3), соединяющие уже смежные в *WS* вершины, т. е. некоторые ребра в *WS* распараллеливаются и реализуются по независимым маршрутам в *PS.*
Поставленная задача, очевидно, имеет решение, так как, добавив ребра ко всем парам *V*-смежпых вершин в гиперсети *S,* получим гиперсеть *S',* для которой $\omega(S')=\omega(PS)$, но такое решение будет, видимо, неоптимальпым. Прием повышения связности гиперсети за счет добавления новых ребер ие всегда пригоден, так как изменяется структура вторичной сети. Поэтому возникает задача увеличения связности гиперсети за счет распараллеливания ребер вторичной сети. В этом случае структура последней остается неизменной.
8.1. Пусть заданы первичная сеть $PS = (X, V)$ и вторичная $WS = (Y, R)$. Предположим, что при оптимальной (с точки зрения максимума связности) реализации *WS* в *PS* получится гиперсеть *S,* для которой $\omega(S) < \min(\omega (PS), (\omega(WS))$. Тогда имеет место

Теорема 18. *Для любых PS и WS всегда можно построить гиперсетъ S (распараллеливая некоторые ребра WS) такую, что*
$\omega(S)= \min(\omega (PS), (\omega(WS))$.

Доказательство. Предположим противное. Тогда в построенной гиперсети *S* найдется минимальное сечение $\{x_i\}$такое, что
$|\{x_i\}|< \min(\omega (PS), (\omega(WS))= k$. Так как $\omega(WS)\geq k$, то среди вершин $\{x_i\}$ найдется такая, которая слабо инцидентна некоторому ребру из *WS.* Причем концевые вершины x_l, x_t этого ребра будут лежать в разных компонентах связности гиперсети $S\backslash\{x_i\}$. Распараллелим найденное ребро (x_l, x_t) на два: $(x_l, x_t)^1$ и $(x_l, x_t)^2$. Первое из них реализуется по старой трассе, а второе — в графе $PS\backslash\{x_i\}$. Очевидно, что вершины x_l, x_t будут связны в $PS\backslash\{x_i\}$, так как $\omega(PS)\geq k$. Следовательно, во вновь полученной гиперсети *S'* вершины $\{x_i\}$ не являются минимальным сечением,— противоречие. Тем самым теорема доказана.
Теперь рассмотрим следующую задачу. Для заданных первичной *PS* и вторичной *WS* сетей найти гинерсеть $S' = (PS, WS; Ф)$, которая

удовлетворяет условиям (10), (11), а граф WS' получается из WS в результате распараллеливания некоторых ребер.

Алгоритм $A2$:

п. 1. Реализовать алгоритм $A1$. Если полученная гиперсеть удовлетворяет условиям (10), (11), то на п. 6, иначе на п. 2.

п. 2. Если гиперсеть S удовлетворяет условию (11), но не удовлетворяет условию (10) или оба условия нарушаются, то на п. 3, иначе на п. 4.

п. 3. В построенной гиперсети S найти минимальное сечение $\{x_i\}$. По множеству $\{x_i\}$ найти слабо инцидентное одной из вершин $\{x_i\}$ ребро $(x_l, x_t)\in R$ и с помощью известных алгоритмов найти в PS две независимые цепи между вершинами x_l, x_t. Емкость $\alpha(r_{l,t})$ данного ребра распределяется по цепям в PS так, чтобы удовлетворялось условие (11). Если это невозможно, то на п. 5, иначе повторить п. 3. Как только связность S станет равной k, то на п. 4.

п. 4. В гиперсети S найти перенасыщенную ветвь v (для которой $\alpha(v) > \sum_{r\in\Phi^{-1}(v)} \alpha(r)$). Найти инцидентные ребра $\Phi^{-1}(v)$ данной ветви v и выбрать среди них такое подмножество $R''\subset\Phi^{-1}(v)$, чтобы

$$\alpha(v) \leqslant \sum_{r\in\Phi^{-1}(v)\setminus R''} \alpha(r).$$

Последовательно распараллеливая эти ребра и распределяя их емкости по дубликатам, попытаемся удовлетворить условию (11). Если это не удается, то на п. 5, иначе п. 4. Повторяется до тех пор, пока не будут проверены все ветви, переход на п. 6.

п. 5. Допустимое решение не найдено.

п. 6. Гиперсеть S построена.

Сходимость алгоритма очевидна. Если ограничение (11) не является критическим, то гииерсеть с заданной связностью будет обязательно найдена. Кроме того, алгоритм позволяет находить гиперсеть с квазиминимальной стоимостью. Поставленная задача, вообще говоря, не всегда позволяет построить гиперсеть S со связностью, равной $\omega(PS)$, т. е. связность вторичной сети существенно влияет на связность S. Поэтому актуальна задача синтеза гиперсети S, для которой $\omega(S) = \omega(PS)$. Этого можно добиться добавлением новых ребер в WS.

8.2. Пусть заданы первичная сеть $PS = (X, V)$ и вторичная $WS = (Y, R)$. Требуется найти гиперсеть $S = (PS, WS; \Phi)$ такую, чтобы $\omega(S)=k\leq\omega(PS)$ и выполнялось условие (11), где $WS' = (Y, R \cup U)$, а U — множество добавленных ребер. Причем при добавлении очередного ребра емкости ребер WS' пересчиты-ваются.

Алгоритм А3:

и. 1. Реализовать алгоритм А1. Если полученная гиперсеть удовлетворяет условию (11) и $\omega(S) = k$, то на п. 5 иначе на п. 2.

п. 2. Если гиперсеть S не удовлетворяет условию (11), то на п. 4, иначе на п. 3.

п. 3. В S найти минимальное сечение $\{x_i\}$. Выделить в $S\backslash\{x_i\}$ компоненты связности WS^1 и WS^2. Найти в них пару ближайших несмежных вершин $x_l \in WS^1$ и $x_t \in WS^2$. Соединить вершины x_l и x_t ребром. Пересчитать емкости ребер с помощью заданной процедуры расчета емкостей. Перейти на п. 1.

п. 4. Допустимое решение не существует.

п. 5. Гиперсеть S найдена.

Сходимость алгоритма очевидна. Отсутствие допустимого решения определяется только условием (11).

В заключение раздела рассмотрим быстрый алгоритм синтеза гиперсети с достаточно большой связностью.

8.3. Пусть заданы первичная сеть $PS = (X, V)$ п вторичная $WS= (Y, R)$. Требуется найти гиперсеть $S= (PS, WS; Ф)$ с максимально возможной связностью. Вес каждой ветви и вершины равен единице.

Алгоритм А4:

п. 1. Упорядочим ребра $r= (x_i, x_j) \in R$ по убыванию расстояния $\rho(x_i, x_j)$ в PS. Получим список ребер $\{r_i\}$, r=1, 2, ..., m; i:=l.

п. 2. Реализуем r, в PS по кратчайшему пути (кратчайший по суммарному весу вершин и ребер).

п. 3. Увеличим вес вершин и ребер найденной трассы ребра r_i. Вес увеличивается на некоторую величину $z(S)$, которая зависит от текущего состояния гиперсети S.

п. 4. Если все ребра WS' реализованы, то при $\omega(S)$=min(ω (PS), $(\omega(WS))$—на п. 10, иначе на п. 5, если не реализованы — $i:= i + 1$, переход на п. 2.

п. 5. С помощью переборного алгоритма найти минимальное сечение по вершинам $\{x_l\}$.

п. 6. Найти все ребра r, слабо инцидентные вершинам из $\{x_l\}$, такие, что их концевые вершины лежат в разных компонентах связности S_1 и S_2 гиперсети $S\backslash\{x_l\}$. Хотя бы одно такое ребро обязательно найдется, так как WS k-связен.

п. 7. В первичной сети $PS\backslash\{x_l\}$ между всеми парами концевых вершин x_i, x_j найдем кратчайшие пути. (Путь обязательно найдется, так как $\omega(PS) \geq k$.)

8. Выберем такую пару x_i и x_j для которой следующая разность минимальна:

$$\rho_{PS\setminus\{x_l\}}(x_i, x_j) - \rho_{PS}(x_i, x_j).$$

п. 9. Реализуем ребро (x_i, x_j) по новой трассе, тогда сечение, разделяющее компоненты S_1 и S_2, увеличится. Повторяем п. 3—6 до тех пор, пока связность ω (S) увеличивается, переход на п. 10.
п. 10. Гиперсеть с квазимаксимальной связностью найдена.

7.7.9. О построении гиперсетей с заданной квазисвязностью

Рассматриваются задачи синтеза в классе $(X \to R)$-гиперсетей с заданной частичной квазисвязностью и ограничением на ранг связывающих квазимаршрутов. Во всех этих задачах считается заданной первичная сеть $PS=(X, V)$ синтезируемой гиперсети.
9.1. Рассмотрим T-гиперсеть, в первичной сети $PS = (X, V)$ которой выделен корень $x_0 \in X$. Вторичная сеть $WS = (X, R)$ содержит то же множество вершин и $|R|$ ребер, причем каждое ребро $r \in R$ инцидентно x_0 и слабо инцидентно остальным вершинам $X\setminus\{x_0\}$. Из определения T-гиперсети следует, что каждому ребру $r \in R$ сопоставлено некоторое корневое дерево T_r в графе $PS = (X, V)$.
Обозначим $\lambda_i(\omega_i)$— реберную (вершинную) связность вершины $x_i \in X$ с вершиной x_0 в графе PS. Тогда корневая реберная и вершинная связности графа PS определяются следующим образом:

$$\lambda = \max_{x_i \in X - x_0} \lambda_i, \quad \omega = \max_{x_i \in X - x_0} \omega_i. \qquad (12)$$

Два ребра r_1 и r_2 гиперсети $S=(X, V, R)$ называются T_λ-независимыми, если для любой вершины $x \in X$ существуют два независимых по ветвям квазимаршрута из x_0 в x, принадлежащих r_1 и r_2. Соответствующие квазимаршруты назовем T_λ-независимыми. T_ω - независимость определяется аналогично, т. е. с учетом независимости квазимаршрутов по вершинам. В дальнейшем будем говорить просто о T-независимости, если соответствующие утверждения справедливы для частичной квазисвязности T-гиперсети как по ветвям, так и по вершинам.
Множество ребер T-гиперсети, инцидентных вершине x_0, будем называть T-покрытием в $S = (X, V, R)$, если они попарно T -независимы. Полным T-покрытием гиперсети S будем называть такое T-покрытие $\Theta(S)$, что $|R|$ равно корневой связности PS и каждая вершина $x \in X$

слабо инцидентна *k* ребрам, где *k* равно числу связности этой вершины с корнем x_0 в *PS*.

Теорема 19. *Для любой первичной сети PS = (X, V) с корнем* x_0 *существует T-гиперсетъ S = (X, V, R), в которой множество R является полным T-покрытием.*

Существуют доказательство теоремы как для T_ω-, так и для T_λ-покрытия. В последнем случае приведены два алгоритма синтеза. Здесь рассмотрен алгоритм, который позволяет находить полное *T*-покрытие, т. е. алгоритм синтеза *T* -гинерсети, где частичная квазиотделимость (как по ветвям, так и по вершинам) любой вершины *x* от вершины x_0 будет максимально возможной, а ранг любого (x_0, *x*)-квазимаршрута равен единице.

Алгоритм А5.

п. 1. Для каждой вершины $i \in X$, $x \neq x_0$ вычислим $k(x)$-связность вершины *x* с x_0, $k(x_0) := \max_{x \in X - x_0} k(x)$, $i = 1$, $X_i := x_0$, $V_i := \varnothing$, $T_i := (X_i,\ T_i)$,

$\Theta(S) := \{T_1, \ldots, T_{k(x0)}\}$, $\forall x \quad l(x) := 0$. Граф $PS = (X, V)$ преобразуется в орграф $PS' = (X, V')$ заменой каждой ветви на пару противоположно ориентированных дуг, а ветви, инцидентной x_0, на дуги, ориентированные из x_0.

п. 2. В графе *PS'* выбираем дугу $v = (x, \overline{x})$ из X_i в $X — X_i$ такую, что дерево $T_i = (X_i \cup \overline{x}, V_i \cup v)$ не нарушает *T*-покрытие $\Theta(S)$. Причем количество независимых цепей из x_0 в $\overline{x}$ в графе $PS'' = (X, V' — V_i)$ уменьшится по сравнению с графом *PS'* ровно на единицу и степень вершины x_0 в графе *PS''* станет не меньше $k(x_0) — 1$. Если такая дуга не найдена, то на п. 4, иначе на п. 3.

п. 3. $V' := V' - v$, $T_i := (X_i \cup \overline{x}, V_i \cup v)$, переход на п. 2.

п. 4. Проверяем последовательно все существующие и вновь получаемые висячие вершины дерева T_i на выполнение неравенства

$$k(x)—l(x)<k(x_0)—i. \qquad (13)$$

Если это неравенство для некоторой висячей в T_i вершины *x* выполняется, то $V' := V' \cup (y, x)$, $((y, x) \in T_i)$, $T_i := (X_i — x, V_i — (y, x))$, иначе переходим к следующей висячей вершине. Процесс заканчивается, если висячих вершин в T_i, удовлетворяющих неравенству (13), не остается.

п. 5. $\forall x \in T_i\ l(x) := l(x) + 1$, $i := i + 1$, если $i > k(x_0)$, то на п. 6, иначе на п. 2.

п. 6. Полное *T*-покрытие найдено.

Проверка правильности подключения новой вершины в *T,* в п. 2 осуществляется путем проверки числа независимых путей из x_0 в *x* в графе *PS''*. После того как найдено $\Theta(S)$-покрытие, каждому ребру

$r \in R \subset WS$ будет сопоставлено дерево из *Θ(S)*. Таким образом, найдена *Т*-гиперсеть, удовлетворяющая условиям задачи синтеза.

9.2. Применение теории гиперсетей в задачах синтеза коммуникационных сетей инициирует еще несколько важных постановок задач синтеза гиперсетей с заданной частичной квазисвязностью. Пусть задана первичная сеть $PS = (X, V)$ и некоторое подмножество вершин $x_0 \cup X'$, где $X' = (x_1, .. ., x_k) \subset X$. Требуется найти гиперсеть $S= (X, V, R)$ такую, чтобы между вершинами x_0 и $x_i \in X'$ имела место частичная k_i-квазинеотделимость по ветвям (вершинам), где $k_i \leq \lambda_i$ ($k_i \leq \omega$,). Ясно, что решение поставленной задачи легко получить с помощью алгоритма А5. Действительно, положим $k(x_0) = \max_{x \in X'} k(x)$ и найдем $\Theta(S)$-покрытие. Затем, поочередно рассматривая висячие вершины $x_3 \in T_i$, будем их удалять из T_i, если они не принадлежат X' или число слабо инцидентных ребер превышает k_j.

Часто может возникнуть ограничение на число слабо инцидентных вершин любому ребру $r \in R$. Данное ограничение естественным образом может быть учтено в предыдущих задачах путем разбиения каждого ребра r на две части r' и r'' по некоторой вершине x так, что x_0 инцидентно r', а x инцидентно r'' и слабо инцидентно r'. Затем процесс разбиения можно продолжить.

В предыдущих задачах синтеза *Т*-гиперсетей предполагалась частичная k-квазинеотделимость корня x_0 от других вершин гиперсети S. Если потребовать попарную частичную k-квазинеотделимость в синтезируемой гиперсети, то задача решается тривиально. Достаточно потребовать, чтобы гиперссть одновременно принадлежала к классам $(V \rightarrow R)$- и $\{X — R\}$-гиперсетей, т. е. вторичная сеть WS будет изоморфна некоторому суграфу PS' графа PS, причем PS' должен удовлетворять заданным характеристикам связности. Однако данная задача представляет особый интерес, когда есть ограничение на ранг независимых квазимарштрутов, соединяющих различные вершины гиперсети. В этом случае, поочередно объявляя каждую вершину гиперсети корнем, решаем первую задачу (алгоритм А5), затем последовательно удаляем висячие ребра в покрывающих деревьях, отслеживая выполнение ограничения и условия задачи.

7.7.10. Заключение

Развитие теории гиперсетей имеет фундаментальное значение для исследования и проектирования иерархических систем сетевой структуры, т. е. таких, в которых иерархия определяется вложением одной подсистемы в другую. Перечислим некоторые важные направления развития гиперсетей.

10.1. Маршруты в гиперсетях. В разд. 3 сформулированы и решены некоторые задачи поиска оптимальных маршрутов. Представляют интерес задачи, связанные с поиском обходов элементов гиперсетей, в частности бесповторных обходов ребер, ветвей или вершил, а также обобщения указанных задач. Важны также задачи, связанные с метрикой гиперсетей (поиск центров, медиан и т. п.). По-видимому, задачи данного направления в основном разрешимы за полиномиальное время.

10.2. Упаковки и покрытия. Так же как и в теории гиперграфов, для гиперсетей определяются упаковки и покрытия, т. е. для любой пары множеств из *X, V, R* и соответствующих отношений инцидепции или смежности можно определить понятия упаковки и покрытия. Кроме того, в гиперсетях возможно определение этих понятий, когда одно множество сопоставляется двум другим. В этом направлении возникают принципиально новые задачи.

10.3. Частичные гиперсети. Удаляя различные элементы из гиперсетей, можно получать всевозможные частичные гиперсети с теми или иными свойствами. Например, возможны следующие задачи: найти минимальную (по числу вершин, ветвей или ребер) частичную гиперсеть, связывающую заданное множество вершин или ветвей; найти компоненты связности, блоки и части гиперсетей с заданной связностью.

10.4. Эквивалентные гиперсети. Две гиперсети $S' = (X', V', R')$ и $S'' = (X'', V'', R'')$ эквивалентны, если и только если изоморфны PS', PS'' и WS', WS'', кроме того, $X' \leftrightarrow X''$. Две гиперсети S' и S'' равносильны, если WS' изоморфен WS'' при $X' \leftrightarrow X''$, и равнозначны, если PS' изоморфен PS''. Основные задачи по этому направлению связаны с поиском эквивалентных (равносильных, равнозначных) гиперсетей с заданным свойством (набором характеристик).

10.5. Планарность. Гиперсеть *S* называется *R*-планарной, если ее можно расположить на плоскости так, чтобы ребра не пересекались (пересечение ребер запрещается как по ветвям, так и внутри вершин; см. геометрическое представление гиперсети на рис. 30. Здесь ребра пересекаются в вершине *1* и не пересекаются в других вершинах. Кроме того, ребро $(s, 4)$ пересекается с ребром $(s, 6)$ по ветви $(5, 6)$).

С понятием планарности гиперсетей связан ряд задач: найти необходимые и достаточные условия *R*-планарности гиперсети; сформулировать алгоритм укладки *R*-планарной гиперсети; найти эквивалентную данной гиперсеть, которая была бы *R*-планарной, если планарны *PS* и *WS*. Возможны и другие задачи, связанные с минимизацией числа пересечений, и т. п.

7.8. Иерархическая организация моделей структур

Иерархическая модель структур, так же как и сетевая, основаны на возможности представления структур в виде графов. Но в отличие от сетевой на иерархическую модель накладываются более жесткие ограничения. Граф иерархической структуры имеет древовидную структуру связей.

Древовидная структура или дерево - это граф, не содержащий циклов. Дерево представляет собой связанный граф, так как каждая вершина в нем завершает по крайней мере одно ребро (рис. 34).

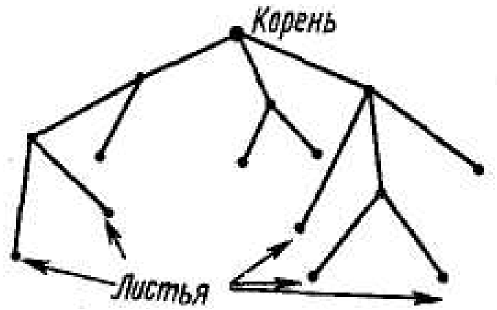


Рис. 34. Граф дерева

При работе с иерархической моделью граф, описывающий структуру, является направленным или ориентированным, т.е. дерево в этом случае будет ориентированным.

В зависимости от направления дуг в ориентированном графе выделяют такие типы вершин, как корень и лист.

Корень — это вершина, которая имеет одну или несколько исходящих дуг и ни одной входящей.

Лист - это вершина, которая имеет одну или несколько входящих дуг и ни одной исходящей.

Корень можно рассматривать как источник направленной древовидной структуры. От корня до любого листа легко проследить элементарную последовательность дуг. Число дуг между корнем и листом называется уровнем диста. Вершину графа, который не является ни корнем, ни листом, называют **узел ветвления.**

Примером иерархической структуры в виде дерева служит схема управления большинства организаций. Корнем схемы является директор. От него исходят одна или несколько дуг к его заместителям. От них отходят дуги к более низким уровням управления и т.д. к непосредственных исполнителей, которые на графе соответствуют листьям.

Основными понятиями в иерархической модели есть **тип записи** и **иерархические отношения.** Вершины в дереве соответствуют типу сущности и называются типом записи. Тип записи состоит из одного или более **элементов.** Во многих иерархических моделях вместо понятия «тип записи» используют эквивалентное понятие «тип сегмента».

Иерархическое отношение (ветка дерева) соединяет два типа записей и представляет собой множество связей между экземплярами записей этих двух типов.

Дуги (ветки) дерева соответствуют функциональному типу связи, т.е. типам 1:1, 1:M, M:1, и их называют **связью исходной порожденной.** Дуга выходит из **типа родительской записи** и заходит в **тип порожденной записи.** Таким образом, рассматривая последовательность связей — исходной-порожденной, можно выделить типы родительских и порожденных записей. Каждый экземпляр родительского типа записей может иметь связь с несколькими (в том числе и с нулем) экземплярами порожденных записей. В свою очередь, каждый экземпляр записи порожденного типа подчинен ровно одному экземпляру записи родительского типа. Другими словами, **<u>иерархическое отношение</u> можно рассматривать как функцию, признаком которой служит экземпляр порожденного типа записи, а ее значением является экземпляр родительского типа.**

Иерархическая модель накладывает жесткие ограничения на иерархические отношения между записями. Поскольку любые два типа записей могут быть связаны не более чем одним иерархическим отношениям, то иерархическим связям не требуются собственные имена. Каждая из иерархических связей может быть однозначно идентифицирована указанием родительской и порожденной записи.

Модели структур «сущность-связь». Они появились как обобщение и развитие иерархических и сетевых моделей. Модель «сущность-связь» была задумана как средство представления предметной области, не зависимого от особенностей среды хранения и не связанного соображениями физической эффективности. Базовыми структурами в этих моделях являются типы сущностей и связей. Тип сущности в модели носит название множество сущностей. **Каждая сущность при этом идентифицирует объект предметной области**. Связь между ними фиксируется заданием множества отношений. Различают два вида отношений - **слабые** и **стандартные**. Слабые связи идентифицируют иерархические отношения.

Множество связей (МС) в данной модели можно представить как математическое отношение n типов сущностей. Если МС есть множество, то его можно определить следующим образом:

$$МС=\{<m_1, m_2,...,m_n>|m_1 \in M_1,\ \ m_2 \in M_2,..., m_n \in M_n\},$$

где m_1 — сущность, которая принадлежит множеству сущностей M_i, а кортеж $<m_1, m_2,...,m_n>$- связь, которая принадлежит множеству МС.

Домен в модели «сущность-связь» называют **множеством значений.** В этом случае значения представляет собой конкретный экземпляр множества значений. Для изображений множеств сущностей и отношений могут быть использованы диаграммы, аналогичные графам. МС изображается прямоугольником, множество отношений - ромбом. Множествам обоих типов присваиваются имена; множества сущностей соединяются с множествами отношений, в котором они участвуют с помощью ненаправленных линий. Множество значений представляется овалом (рис. 35).

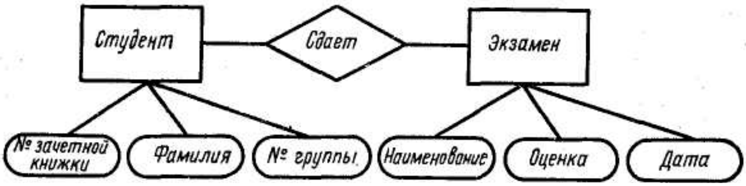


Рис. 35. Модель структур «сущность-связь»

Использование моделей «сущность-связь» является удобным средством для представления концептуальной информации.

Бинарные модели. Отношения, используемые для описания предметной области, в общем случае могут быть n-мерными. Интенсивно идут исследование по представлению информации с помощью, как правило, бинарных отношений, поскольку они позволяют лаконично представить сложные связи.

Бинарные модели, основанные на использовании бинарных отношений, имеют простые базовые структуры, способные обеспечить эффективное представление предметной области. Поскольку графы позволяют наглядно задавать отношения, их использование в бинарных моделях также позволяет лучше уяснить особенности модели. Вершины графа в бинарных моделях соответствуют классификационному обобщению элементов структур в типы и называются *категориями,* а дуги — *бинарным отношением* категорий. Граф, который удовлетворяет этим структурным представление, носит название *графа типов.* Каждому бинарному отношению ставится в соответствие отношение, имеющее противоположное направление. Например:

СТУДЕНТ -УЧИТСЯ У - ПРЕПОДАВАТЕЛЬ - ОБУЧАЕТ -

Обеим направлениям бинарного отношения присваиваются уникальные имена, которые называются *функциями доступа.*

В бинарном отношении категорий СТУДЕНТ и ПРЕПОДАВАТЕЛЬ направление от категории СТУДЕНТ к категории ПРЕПОДАВАТЕЛЬ есть функция доступа УЧИТСЯ У. Представленное отношение может быть охарактеризовано следующим образом:

СТУДЕНТ УЧИТСЯ У ПРЕПОДАВАТЕЛЯ

Функция доступа для противоположного направления будет

ПРЕПОДАВАТЕЛЬ УЧИТ СТУДЕНТА

Расширение бинарного графа типов позволяет рассматривать понятие двух родов-объектов и поименованных бинарных отношений (связей). Взаимосвязи, в которых принимает участие более двух объектов, в свою очередь, интерпретируются как объекты.

Объект - это реализация категории. Например, объект - это конкретные студенты и преподаватели. Объекты подразделяются на **абстрактные и конкретные**. Причем имеется в виду, что абстрактные всегда существуют, в то время как конкретные появляются и исчезают в описании реального объекта. Так, абстрактные объекты используются для представления чисел, дат и др.

Объекты соединены связями, которые отображают реализацию бинарного отношения. Каждой связи в обоих направлениях присваиваются имена, соответствующие функциям доступа.

Для создания категорий в бинарной модели используют оператор CATEGORY. Например, условие

СТУДЕНТ = CATEGORY

говорит о создании новой категории с именем СТУДЕНТ.

Для создания бинарного отношения следует задать его имя, функции доступа и категории соответствующему данному отношению объектов. Например:

СТУДЕНТ - ПРЕПОДАВАТЕЛЬ = RELATION (СТУДЕНТ, ПРЕПОДАВАТЕЛЬ, УЧИТСЯ У, УЧИТ) определяет бинарное отношение СТУДЕНТ - ПРЕПОДАВАТЕЛЬ категорий СТУДЕНТ и ПРЕПОДАВАТЕЛЬ с функциями доступа УЧИТЬСЯ У, ОБУЧАЕТ.

В бинарных моделях отсутствует явно выраженное понятие «свойство объекта». Свойства объектов могут быть определены с помощью бинарного отношения, заданного на множестве объектов и других элементарных объектов, с помощью которых задаются значения свойств.

Как уже отмечалось, отношение характеризуется двумя функциями доступа, каждая из которых определяет минимальное и максимальное число объектов в присоединяемой категории. Ограничение на число объектов задается оператором AFN, позволяющим задать краничные величины подмножеств значений функций доступа. Так, например, если для функции доступа ОБУЧАЕТ запишем:

$$\text{ОБУЧАЕТ} = \text{AFN}\ (0, \infty),$$

то это будет означать, что преподаватель может вообще никого не обучать, или обучать любое число студентов.

Логический доступ к данным бинарной модели обеспечивается с помощью программ, которые реализуют элементарные операции доступа.

Семантические сети. Для представления семантических (смысловых) текстов, задаваемых на естественном языком, разработаны семантические сетевые модели.

Семантическая сеть представляет собой ориентированный граф с намеченными вершинами и дугами. При этом если вершины обозначаются только в целях ссылок к ним, то метки дуг содержат сведения о некоторых их семантических свойствах и значениях.

Для представления моделей элементов структур используются четыре типа вершин: **концепты** (или понятия), **события, характеристики** (свойства) и **значения.**

Концепты - константы или параметры, которые специфицируют физические или абстрактные объекты.

События - отвечают действиям, наблюдаемым в представляемой области.

Характеристики - вершины, соответствующие свойствам концепты.

Значения - вершины, соотносящиеся с областями значений, которые могут принимать характеристики.

Поскольку имеется четыре типа вершин, необходима соответствующая зависимая от этих типов интерпретация дуг, которые соединяют различные вершины. Модели семантической сети предусматривают возможность распределения вершин по типам. В этом случае следует различать вершины-концепты и вершины-классы, которые собственно и представляют определенные типы вершин. Например, КУЛЕШОВ - концепт, СТУДЕНТ- класс.

Различие между классом и концептом весьма близко к тому же, что между типом и экземпляром в других моделях. Отличие заключается в том, что граф семантической сети включает как классы, так и концепты. Кроме того, концепт может быть соотнесен с несколькими классами.

Поскольку семантические сети предусматривают задавать на графе в явном виде различие между вершинами-концептами и вершинами-классами, то в рассмотрение вводятся три вида дуг: утверждение; порождение экземпляра; бинарное отношение.

Утверждение — дуга, которая соединяет два концепта.

Порождение экземпляра — дуга, между классом и концептом.

Бинарное отношение — дуга, которая связывает два класса.

Рассмотрим пример семантической сети (рис. 36), иллюстрирующей сказанное.

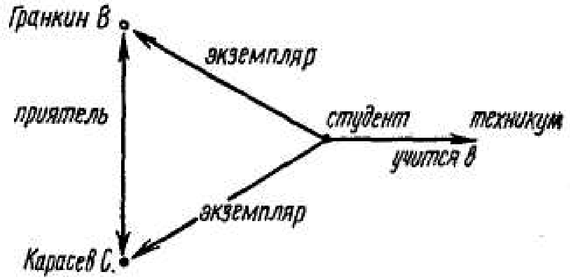


Рис. 36. Пример семантической сети

В данной семантической сети, вершины ГРАНКИН В. и КАРАСЕВ С. - концепты, СТУДЕНТ и ТЕХНИКУМ - классы. Дуга ПРИЯТЕЛЬ - утверждение. Дуги, связывающие вершину СТУДЕНТ с вершинами КАРАСЕВ С. и ГРАНКИН В., отображают связь экземпляров с классами. Дуга УЧИТСЯ В отображает бинарное отношение между классами СТУДЕНТ и ТЕХНИКУМ.

Классы могут быть связаны в иерархию в соответствии со связями ЕСТЬ НЕК и ЕСТЬ - ЧАСТЬ. Эти связи позволяют из отдельных понятий и классов строить более общие понятия и классы.

Для представления в семантической сети некоторых событий и действий вводится набор простых отношений, которые характеризуют основные компоненты события. Для построения с помощью семантической сети структуры события в первую очередь выделяют из него само действие, которое описывается обычно глаголом. После этого выделяют лиц, совершающих действие и объекты, над которыми оно осуществляется. Лицо, которое осуществляет действие, называется **агентом.** Вещи, над которыми действие осуществляется, называют

объектами. Лицо, получающееся результатом действия или испытывающее его, называется **адресат**.

Рассмотрим предложение: «Мастер починил телевизора». В этом предложении выделим действие: ПОЧИНИЛ. Очевидно, что объектом является МАСТЕР (рис. 37).

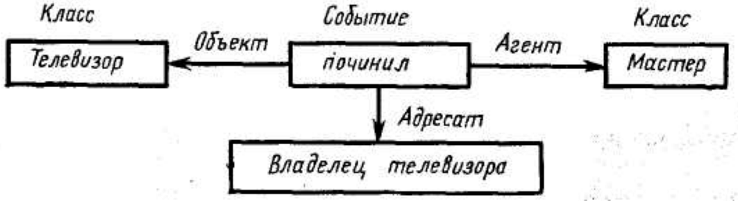


Рис. 37. Сеть предложения «Мастер починил телевизора»

В данном предложении адресат явно не указан, но его наличие можно предположить. Рассмотрим более сложное предложение:

«Вчера мастер Босин П. А. починил магнитофон «Юность», принадлежащий студенту Афонину К. А.».

Семантическая сеть для данного предложения представлена на рис. 38.

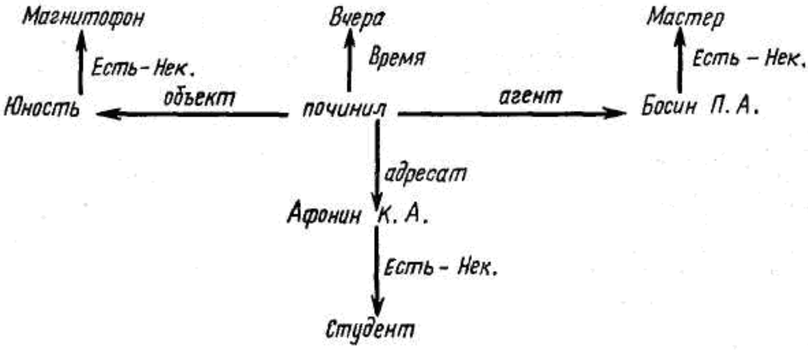


Рис. 38. Пример семантической сети

В этом предложении появилась новая дуга ВРЕМЯ, которое указывает на то, когда происходит событие. В общем случае в семантических сетях кроме простых отношений: агент, объект, адресат выделяют и другие, позволяющие описать широкий класс событий. К ним относят: время, место, инструмент, цель и др.

Операции, совершаемые над элементами структур, задаваемые семантической сетью, разбиваются на два подмножества: операции над классами и над бинарными отношениями.

Над классами могут быть соверщены четыре операции:

- создание экземпляра некоторого класса или установление принадлежности существующего экземпляра некоторого класса к еще одному;

- устранение принадлежности экземпляра к некоторому классу или полное его исключение;

- выборка экземпляров, которые принадлежат к одного класса;

- определение принадлежности экземпляра указанному классу.

Над бинарными отношениями могут быть совершены три операции:

- установление связи между классами;

- выборка всех экземпляров, связанных в данном бинарном отношении с указанным экземпляром;

- установление наличия связей между двумя экземплярами.

В общем случае реализация всех вышеперечисленных операций требует создания специальных программ, которые учитывают рассматриваемую предметную область.

Разработка семантических сетевых моделей явилась следствием повышения требований к интегрированному представлению элементов структур, которое включает не только элементы, но и их категории, свойства категорий и операции над данными. Важную роль в развитии моделей структур этого класса сыграли проблемы алгоритмизации процессов естественного языка. Модели семантических сетей широко используются при разработке систем искусственного интеллекта.

Инфологические модели структур. Ифологическое проедставленне полностью независимо от физических параметров среды хранения. Эта модель в качестве базовой использует понятие: объект, свойства и связи.

Под объектом понимается нечто представляющее интерес для решаемой задачи. Предполагается, что существование объекта связано с такими событиями, как появление (возникновение), изменение и исчезновение.

Объекты подразделяются на атомарные и составные.

Атомарный объект, — это любой объект, дальнейшее разложение которого на другие объекты невозможно.

Составной объект включает в себя множество объектов.

С каждым объектом связывается определенный набор свойств. Важным свойством существования объекта является время (время его возникновения и исчезновения).

С помощью базовых концепций объектов, свойств связей и времени формируется элементарный факт. Элементарный факт задается формально как тройка $<x, y, z>$ или $(<x_1, x_2, ..., x_n> rz)$, где $<x_1, x_2, ..., x_n>$ - кортежи объектов, y — свойство, r — отношение, z — время. Такую базовую структуру называют элементарной плеядой.

Так как в рамках инфологической модели все может быть объявлено объектом, то свойства и отношения, в свою очередь, также могут рассматриваться как объект.

В данной модели типы объектов вводятся путем группирования объектов и их свойств. Объектная группа $O(p)$, соотнесенная со свойством p, определяется как совокупность объектов, потенциально (вне связи со временем) имеющих свойство p. Объекты, обладающие свойством p, в определенный момент времени образуют подмножество $O_t(p)$ множества $O(p)$, называемое временным срезом объектной группы.

В инфологической модели данных понятия атрибута вводится с помощью понятий объекта, свойства, связи и времени. Атрибут определяется как множество свойств $A= \{p_i\}$ объектной группы $O(p)$, такое, что в каждый момент времени каждый объект x, принадлежащий $O_t(p)$, содержится, по крайней мере, в одной $O_t(p)$. Свойства p — это значение атрибута A объектной группы $O(p)$.

Базовые операции включения, обновления и удаления в инфологической модели связаны с вводом новых элементарных сообщений и предполагают наличие механизмов интерпретации элементарного сообщения.

Процедуры обращения к БД **(транзакции)** могут быть представлеы парой (оператор, параметр). К основным операциям, которые изменяют БД, относят <ДОБАВИТЬ, сообщение> <УБРАТЬ сообщение> <ЗАМЕНИТЬ сообщение 1, сообщение 2>.

На основе базовых процедур возможно конструирование более универсальных запросных средств.

Следует иметь в виду, что возможности инфологической модели существенно меньше возможностей естественного языка и их удобно использовать лишь в тех случаях, когда модель предметной области позволяет описать только очевидные факты.

Литература

1. Кононюк А.Е. Системология. Общая теория систем. К. 1 «Начала», - Киев: Освіта України, 2013.

2. Кононюк А.Е. Системология. Общая теория систем. К. 2, Ч.1 «Общая теория структур» , - Киев: Освіта України, 2014, 558 с.

3. Кононюк А.Е. Дискретная математика. К. 1 «Начала (четкие и нечеткие)», - Киев: Освіта України, 2013, 576 с.

4. Кононюк А.Е. Дискретная математика. К. 2 «Множества (четкие)», - Киев: Освіта України, 2013, 522 с.

5. Кононюк А.Е. Дискретная математика. К. 3 «Множества (нечеткие)», - Киев: Освіта України, 2013, 452 с.

6. Кононюк А.Е. Дискретная математика. К. 4 «Отношения (четкие)», - Киев: Освіта України, 2013, 506 с.

7. Кононюк А.Е. Дискретно-непрерывная математика. К. 5, Ч.1 «Алгебры (четкие и нечеткие)», - Киев: Освіта України, 2011, 452 с.

8. Кононюк А.Е. Дискретно-непрерывная математика. К. 5, Ч.2 «Алгебры (четкие и нечеткие)», - Киев: Освіта України, 2011, 612 с.

9. Кононюк А.Е. Дискретно-непрерывная математика. К. 6, Ч.1 «Матрицы (четкие и нечеткие)», - Киев: Освіта України, 2011, 612 с.

10. Кононюк А.Е. Дискретно-непрерывная математика. К. 6, Ч.2 «Матрицы (четкие и нечеткие)», - Киев: Освіта України, 2011, 500 с.

11. Кононюк А.Е. Дискретно-непрерывная математика. К. 6, Ч.3 «Матрицы (четкие и нечеткие)», - Киев: Освіта України, 2012, 520 с.

12. Кононюк А.Е. Дискретно-непрерывная математика. К. 6, Ч.4 «Матрицы (четкие и нечеткие)», - Киев: Освіта України, 2012, 508 с.

13. Кононюк А.Е. Дискретно-непрерывная математика. К. 6, Ч.5 «Матрицы (четкие и нечеткие)», - Киев: Освіта України, 2012, 672 с.

14. Кононюк А.Е. Дискретно-непрерывная математика. К. 7, Ч.1 «Графы (четкие и нечеткие)», - Киев: Освіта України, 2014, 524 с.

15. Кононюк А.Е. Дискретно-непрерывная математика. К. 7, Ч.2 «Графы (четкие и нечеткие)», - Киев: Освіта України, 2014, 584 с.

16. Кононюк А.Е. Информациология. Общая теория информации. К. 1, - Киев: Освіта України, 2011, 476 с.

17. Кононюк А.Е. Информациология. Общая теория информации. К. 2, - Киев: Освіта України, 2011, 476 с.

18. Кононюк А.Е. Информациология. Общая теория информации. К. 3, - Киев: Освіта України, 2011, 412 с.

19. Кононюк А.Е. Информациология. Общая теория информации. К. 4, - Киев: Освіта України, 2011, 488 с.

20. Кононюк А.Е. Обобщенная теория моделирования. Начала. К.1.Ч.1, - Киев: "Освіта України", 2012. - 602 с.

21. Кононюк А.Е. Обобщенная теория моделирования. Начала. К.1.Ч.2, - Киев: "Освіта України", 2012. - 708 с.

22. Кононюк А.Е. Обобщенная теория моделирования. Начала. К.1.Ч.3, - Киев: "Освіта України", 2012. - 568 с.

23. Кононюк А.Е. Обобщенная теория моделирования. Числа (количественные оценки параметров моделей). К.2.- Киев:"Освіта України", 2012. - 548с.

24. Кононюк А.Е. Обобщенная теория моделирования. Величины и размерности. К.3.Ч.1, - Киев:"Освіта України", 2012. - 636 с.

25. Кононюк А.Е. Обобщенная теория моделирования. Величины (количественные характеристики моделей). К.3. Ч. 2. Физические величины (Начало), - Киев:"Освіта України", 2012. - 476 с.

26. Кононюк А.Е. Основы теории оптимизации. К.1. Киев:"Освіта України", 2011. - 692 с.

27. Кононюк А.Е. Основы теории оптимизации. Безусловная оптимизация. К.2. Ч.1. Киев:"Освіта України", 2011. - 552 с.

28. Кононюк А.Е. Основы теории оптимизации. Безусловная оптимизация. К.2. Ч.2. Киев:"Освіта України", 2011. - 616 с.

29. Кононюк А.Е. Основы теории оптимизации. Безусловная оптимизация. К.2. Ч.3. Киев:"Освіта України", 2011. - 456 с.

30. Кононюк А.Е. Основы теории оптимизации. Безусловная оптимизация. К.2. Ч.4. Киев:"Освіта України", 2011. - 512 с.

31. К. Берж. Теория графов и ее применение. М.,Инностр. Лит., 1962

32.Чемоданов Б.К., «Математические основы теории автоматического регулирования», 1977

33.Воронов А.А., «Введение в динамику сложных систем», 1980

34. Заде Л., Дезоер Ч., «Теория линейных систем», 1970

\. Зыков А. А. Гиперграфы.—Успехи мат. наук, 1974, т. XXIX, вып. 6 (180),
с. 89—154.
2. Нечепуренко М. И. Модели структурного резервирования систем,— В кн.: Прикладные задачи на графах и сетях. Новосибирск: ВЦ СО АН СССР, 1981, с. 57—86.
3. Gattinger M., Wensel G. Ilupemetworks as toos for modeling multiport systems— J, of the Franklin Inst., v. 317, N 1, 1984.