

*В. В. Белов, Е. М. Воробьев,
В. Е. Шаталов*

ТЕОРИЯ ГРАФОВ

*Допущено Министерством
высшего и среднего
специального образования СССР
в качестве учебного пособия
для студентов высших
технических учебных заведений*



МОСКВА «ВЫСШАЯ ШКОЛА» 1976

51
Б43
УДК 51 (075)

Рецензенты: кафедра прикладной математики Кишиневского государственного университета и профессор МГУ Л. Н. Королев.

Белов В. В. и др.
Б 43 Теория графов. Учеб. пособие для втузов.
М., «Высш. школа», 1976.

392 с. с. ил.

Перед загл. авт.: В. В. Белов, Е. М. Воробьев,
В. Е. Шаталов.

В учебном пособии рассматриваются некоторые элементы теории графов; выделяются топологические, комбинаторные и прикладные аспекты теории. Большое внимание уделяется алгоритмам решения задач теории графов.

Предназначается для студентов факультетов прикладной математики втузов.

Б $\frac{20204 - 250}{001(01) - 76}$ 79—75

51

ПРЕДИСЛОВИЕ

В основу книги положен материал ряда общих и специальных курсов по теории графов, которые читаются кафедрой прикладной математики МИЭМ. Владение методами теории графов является в настоящее время необходимой составной частью образования специалистов, занимающихся вопросами прикладной математики.

Успех применения графов можно объяснить тем, что они являются удобным языком для формулировки и эффективным инструментом для решения задач, относящихся к весьма широкому кругу научных проблем. Можно упомянуть в этой связи вопросы конструирования систем автоматизированного проектирования программирования для ЭВМ и создания операционных систем, исследования операций и управления, а также ряд задач экономики, статистической и теоретической физики, теории информации, социологии, математической лингвистики и т. д. Возможность приложения теории графов к столь различным областям заложена, в сущности, уже в самом понятии графа, сочетающего в себе теоретико-множественные, комбинаторные и топологические аспекты.

Говорить о единой математической теории, применяющейся в различных прикладных областях со своим кругом идей и методов в каждой, можно только в том случае, если она достигла достаточно высокого уровня формализации и удовлетворяет требованиям математической строгости. Эту сторону вопроса авторы книги стремились подчеркнуть способом изложения теории графов, в котором определения основных понятий даются в наибольшей общности, а рассуждения проводятся, по возможности, математически строго.

Материал книги можно разбить на три группы. К первой группе, которую можно назвать алгоритмическими задачами теории графов, относятся те задачи, где теория графов применяется как удобный язык для описания на содержательном уровне алгоритмов (гл. II, VIII, § 3 гл. IX, X, XI). Ко второй группе относятся вопросы комбинаторного характера (гл. VI, IX, XII—XIV). Наконец, в

третьей группе сосредоточены вопросы применения теории графов к различным математическим задачам и собственно вопросы теории графов (гл. I, III, IV, V, VII, XI, XV).

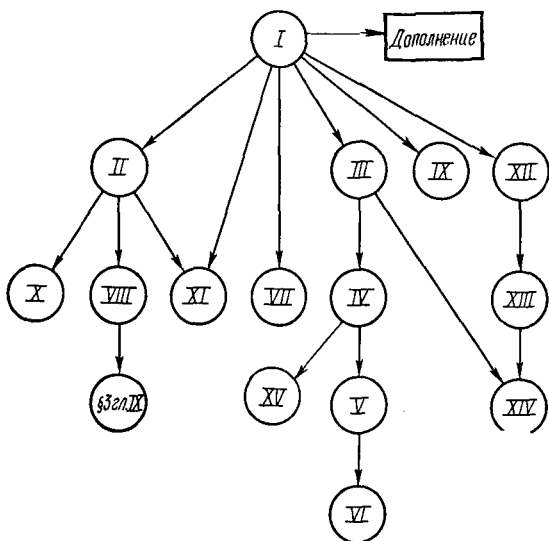


Рис. 1

Это деление материала книги, конечно, в значительной степени условно.

При написании книги авторы существенно использовали источники, указанные в списке литературы.

Дополнение, написанное В. В. Беловым и В. П. Масловым, посвящено приложению теории графов к ряду физических проблем, связанных с нерелятивистской квантовой механикой и квантовой теорией поля. На основе понятия дискретной комплексной марковской цепи устанавливается связь между фейнмановской вероятностной концепцией квантовой механики, методом функционального интегрирования (интегрирования по траекториям) и диаграммами теории возмущений. Этот материал потребовал введения математического аппарата, не связанного непосредственно с теорией графов, и поэтому его включение в основной текст пособия было признано нецелесообразным.

В качестве первого примера использования языка графов приводим граф зависимости глав книги (рис. 1).

Авторы

ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ. ПЕРВЫЕ ЗАДАЧИ ТЕОРИИ ГРАФОВ

§ 1. ПРЕДВАРИТЕЛЬНЫЕ СВЕДЕНИЯ ИЗ ТЕОРИИ МНОЖЕСТВ

1. Множества и операции над ними. Термины «множество» и «элемент» считаются первичными и поэтому не определяются. Предполагается, что для данных элемента a и множества A всегда можно определить, принадлежит элемент a множеству A (пишется $a \in A$) или элемент a не принадлежит множеству A (запись: $a \notin A$, или $a \bar{\in} A$).

Введем некоторые обозначения:

ϕ — множество, не содержащее элементов, т. е. для любого элемента a $a \notin \phi$; ϕ — пустое множество;

$\mathbb{Z}$ — множество целых чисел, т. е. такое множество, что $a \in \mathbb{Z}$ тогда и только тогда, когда a — целое число;

$\mathbb{Z}_+$ — множество неотрицательных целых чисел;

$\mathbb{R}$ — множество действительных чисел.

Определим теперь операции над множествами. Если из $x \in A$ следует $x \in B$, то $A \subseteq B$. В этом случае будем говорить, что A вложено в B (A содержится в B) и множество A называть подмножеством множества B .

Если A и B — множества, то через $A \setminus B$ обозначается множество всех элементов A , не принадлежащих B , т. е. $x \in A \setminus B$ тогда и только тогда, когда $x \in A$ и $x \notin B$.

Через $A \cup B$ обозначается такое множество, что $x \in A \cup B$ тогда и только тогда, когда $x \in A$ или $x \in B$ (при этом не исключено, что одновременно $x \in A$ и $x \in B$). Через $A \cap B$ обозначается множество такое, что $x \in A \cap B$ тогда и только тогда, когда $x \in A$ и $x \in B$.

Множество $A \cup B$ называется *суммой множеств* A и B , а $A \cap B$ — *пересечением* их.

Так, справедливы соотношения

$$A \cap B \subseteq A, \quad A \cap B \subseteq B, \quad A \subseteq A \cup B, \quad B \subseteq A \cup B.$$

Отметим, что если $A \subseteq B$, $B \subseteq C$, то $A \subseteq C$, поэтому из приведенных выше соотношений следует, что $A \cap B \subseteq A \cup B$.

Справедливы также формулы дистрибутивности

$$(A \cup B) \cap C = (A \cap C) \cup (B \cap C), \quad (A \cap B) \cup C = \\ = (A \cup C) \cap (B \cup C)$$

(отметим, что $A = B$ выполнено тогда и только тогда, когда $A \subseteq B$ и $B \subseteq A$).

Определения суммы и пересечения можно несколько обобщить. Пусть I — произвольное множество, каждому элементу i которого поставлено в соответствие множество A_i . Тогда через $\bigcup_{i \in I} A_i$ обозначается такое множество, что $x \in \bigcup_{i \in I} A_i$ тогда и только тогда, когда существует $i_0 \in I$ такой, что $x \in A_{i_0}$, а через $\bigcap_{i \in I} A_i$ — такое, что $x \in \bigcap_{i \in I} A_i$ тогда и только тогда, когда для всех $i \in I$ $x \in A_i$.

Справедливы обобщенные формулы дистрибутивности

$$\left(\bigcup_{i \in I} A_i \right) \cap B = \bigcup_{i \in I} (A_i \cap B), \quad \left(\bigcap_{i \in I} A_i \right) \cup B = \bigcap_{i \in I} (A_i \cup B).$$

Все приведенные соотношения доказываются весьма просто; проверка их предоставляется читателю.

Приведем два примера множеств. Пусть A — некоторое множество. Через 2^A обозначим множество подмножеств A , т. е. $B \in 2^A$ тогда и только тогда, когда $B \subseteq A$.

Множество A называется *конечным*, если в нем содержится конечное число элементов. Если эти элементы есть $a_1, \dots, a_n$, то пишут $A = \{a_1, a_2, \dots, a_n\}$. В частности, множество, состоящее из одного элемента a , обозначается через $\{a\}$. Число n называется *количеством элементов* A и обозначается через $|A|$. Отметим, что $|2^A| = 2^{|A|}$.

Опишем еще один способ построения новых множеств из уже построенных. Пусть A, B — два множества. Через $A \times B$ обозначим множество, состоящее из упорядоченных пар (a, b) , таких, что $a \in A, b \in B$. Иначе говоря $c \in A \times B$ тогда и только тогда, когда c есть пара (a, b) , причем $a \in A, b \in B$. Мы будем обозначать a через $\text{Пр}_1 c$, b — через $\text{Пр}_2 c$ и называть их первой и второй проекциями элемента c .

Множество $A \times B$ называется *декартовым произведе-*

* Приведенные рассуждения нельзя рассматривать как строгое определение конечного множества; такое определение весьма сложно.

нием A и B . С помощью декартова произведения можно ввести операцию дизъюнктивной суммы $\bigvee_{i \in I} A_i$ семейства множеств $A_i (i \in I)$. Именно, по определению

$$\bigvee_{i \in I} A_i = \bigcup_{i \in I} \{i\} \times A_i.$$

В частности, $A \vee B = (\{1\} \times A) \cup (\{2\} \times B)$. При дизъюнктивном сложении, в отличие от операции объединения, не отождествляются одинаковые элементы множеств A и B (или семейства A_i).

2. Отношения. Отношение эквивалентности и разбиение на классы. Введем понятие отношения. Будем говорить, что на A задано отношение R , если задано подмножество R декартова произведения $A \times A$. Элемент $a \in A$ находится в отношении R к элементу $b \in A$ (что записывается как aRb) тогда и только тогда, когда $(a, b) \in R$. При этом существует порядок элементов: может быть, что a находится в отношении R к элементу b , но b не находится в отношении R к a .

Отношение называется а) *симметричным*, если aRb тогда и только тогда, когда bRa ; б) *рефлексивным*, если для любого $a \in A$ aRa ; в) *транзитивным*, если из aRb и bRc следует aRc ; д) *антисимметричным*, если из aRb и bRa следует $a = b$.

Из множества всех отношений выделяется класс отношений, одновременно симметричных, транзитивных и рефлексивных. Такие отношения называются *отношениями эквивалентности**.

Теорема 1.1. Если R есть отношение эквивалентности на множестве A , то существует такое разбиение $A_i (i \in I)$ множества A

$$A = \bigcup_{i \in I} A_i, \quad A_i \cap A_j = \emptyset (i \neq j),$$

что aRb тогда и только тогда, когда существует $i_0 \in I$ такое, что $a \in A_{i_0}$, $b \in A_{i_0}$. Обратно, каждое разбиение A задает на A такое отношение эквивалентности.

Доказательство. Если A пусто, то теорема тривиальна. Если A непусто, то через $\{a\}$ обозначим под-

* Отметим, что это не единственный класс, который может быть выделен. Например, отношения, обладающие одновременно свойствами а), в) и д), называются отношениями порядка. Нам такие отношения понадобятся в гл. IX.

множество элементов b множества A таких, что aRb . Справедливы утверждения:

1) $a \in \{a\}$, так как в силу $b) aRa$;

2) если $\{a\} \cap \{b\} \neq \emptyset$, то $\{a\} = \{b\}$, так как если $x \in \{a\} \cap \{b\}$, то из $y \in \{a\}$ следует aRy . Далее, aRx , следовательно, по $a) xRa$. Используя $c)$, получаем xRy и, используя bRx , аналогично получим bRy , т. е. $y \in \{b\}$. Доказано, что $\{a\} \subset \{b\}$. Симметрично получим $\{b\} \subset \{a\}$.

Из 1) и 2) следует, что множество подмножеств вида $\{a\}$, $a \in A$ образует разбиение множества A . Очевидно, что aRb эквивалентно тому, что a и b лежат в одном классе разбиения. ■

3. Отображения. Образ и прообраз множества при отображении. Рассмотрим теперь понятие отображения. Термины «отображение» и «функция» будут рассматриваться как синонимы. Считаем, что задано отображение $f: X \rightarrow Y$, если каждому элементу $x \in X$ сопоставлен элемент $f(x) \in Y^*$.

Пусть $f: X \rightarrow Y$ — некоторое отображение. Если $A \subset X$, то образом A при отображении f называется множество элементов $B \subset Y$ такое, что $y \in B$ тогда и только тогда, когда существует $x \in A$ такой, что $y = f(x)$. Аналогично, *прообразом* $C \subset Y$ при отображении f называется множество $D \subset X$ такое, что $x \in D$ тогда и только тогда, когда $f(x) \in C$. Образ и прообраз обозначим соответственно $f(A)$ и $f^{-1}(C)$. Операция взятия образа или прообраза множества удовлетворяют соотношениям:

$$f^{-1}(A \cup B) = f^{-1}(A) \cup f^{-1}(B), \quad f(A \cup B) = f(A) \cup f(B);$$

$$f^{-1}(A \cap B) = f^{-1}(A) \cap f^{-1}(B), \quad f(A \cap B) \subset f(A) \cap f(B);$$

$$f^{-1}(Y \setminus A) = X \setminus f^{-1}(A), \quad f(X \setminus A) \supset f(X) \setminus f(A).$$

Однако нельзя утверждать, что $f(A \cap B) = f(A) \cap f(B)$ и $f(X \setminus A) = f(X) \setminus f(A)$. Достаточно рассмотреть постоянное отображение $f: X \rightarrow Y$ (такое, что $f(x) = y_0 \in Y$ для любого $x \in X$).

Отображение называется *взаимно-однозначным*, если $f(X) = Y$ и для любого $y \in Y$ множество $f^{-1}(\{y\})$ состоит из одного элемента. Нетрудно видеть, что определено отображение $f^{-1}: Y \rightarrow X$. Именно, $f^{-1}(y) = a$ тогда и только тогда, когда $f^{-1}(\{y\}) = \{a\}$.

* Приведенное высказывание нельзя считать определением функции; будем считать это понятие первичным.

Пусть $f: X \rightarrow Y$, $g: Y \rightarrow Z$ — отображения. Через $g \circ f: X \rightarrow Z$ обозначается отображение, которое элементу $x \in X$ ставит в соответствие элемент $g(f(x))$. Если f взаимно-однозначно, то $f \circ f^{-1} = 1_X$, $f^{-1} \circ f = 1_Y$, где $1_X(x) = x$ для любого $x \in X$, $1_Y(y) = y$ для любого $y \in Y$.

Пусть X, Y — множества, $X \times Y$ — их декартово произведение. Обозначим через $\text{Пр}_1: X \times Y \rightarrow X$ отображение, сопоставляющее элементу $(x, y) \in X \times Y$ элемент $\text{Пр}_1(x, y) = x \in X$. Аналогично определяется отображение $\text{Пр}_2: X \times Y \rightarrow Y$.

Пусть $A = \{1, 2, \dots, n\}$. Подстановкой порядка n называется взаимно-однозначное отображение A в A . Множество всех подстановок порядка n обозначается S_n (позднее мы увидим, что S_n — группа относительно операции композиции).

§ 2. ОСНОВНЫЕ ОПРЕДЕЛЕНИЯ ТЕОРИИ ГРАФОВ

1. Определение мультиграфа. Прежде чем приступить к точным определениям, поясним их наглядно. Любое конечное множество точек (вершин), некоторые из которых

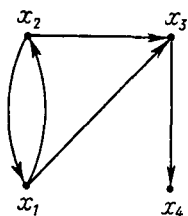


Рис. 2

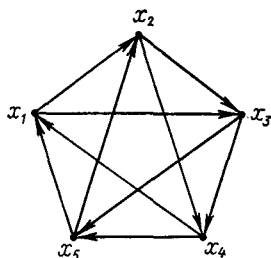


Рис. 3

соединены попарно стрелками (в теории графов эти стрелки называются дугами), можно рассматривать как граф (рис. 2). Например, граф может изображать сеть улиц в городе, вершины графа — перекрестки, стрелками обозначены улицы с разрешенным направлением движения. (Улицы могут быть с односторонним и двусторонним движением.) Отметим здесь два момента. Во-первых, не любой граф можно изобразить на плоскости так, чтобы дуги пересекались только в вершинах (рис. 3). Во-вторых, две вершины могут быть соединены несколькими дугами, идущими

в одном направлении (например, таким свойством может обладать схема железных дорог). Такие дуги будем называть *кратными*, а граф, содержащий кратные дуги, часто называют *мультиграфом* (рис. 4). Если подчеркивается, что в данном графе нет кратных дуг, то он называется графом без кратных дуг.

Определение 2.1. Мультиграфом или просто графом G называется пара объектов $G = (X, \Gamma)$, где X — конечное множество, а Γ — конечное подмножество прямого произведения $X \times X \times \mathbb{Z}_+$. При этом X называется множеством вершин, а Γ — множеством дуг графа G .



Рис. 4

Покажем, каким образом введенное определение формализует понятие графа, интуитивно описанное выше. Дугу (стрелку), соединяющую вершины x_i, x_j , мы обозначим через (x_i, x_j, n) , где x_i — начало дуги, x_j — ее конец, а n — номер дуги. Например, графы, изображенные на рис. 2, 24, 4, могут быть аналитически описаны соотношениями:

$$X = \{x_1, x_2, x_3, x_4\}, \quad \Gamma = \{(x_1, x_2, 1), (x_2, x_1, 1), (x_2, x_3, 1), (x_3, x_4, 1), (x_1, x_3, 1)\}; \quad (2.1)$$

$$X = \{x_1, x_2, x_3, x_4, x_5\}, \quad \Gamma = X \times X \times \{1\} \setminus E \times \{1\}, \quad (2.2)$$

где $E \subset X \times X$ состоит из всех пар вида (x_i, x_i) ;

$$X = \{x_1, x_2\}, \quad \Gamma = \{(x_2, x_1, 1), (x_2, x_1, 2), (x_2, x_1, 3), (x_1, x_1, 1)\}. \quad (2.3)$$

Отметим, что при нумерации дуг не требуется, чтобы были использованы все номера, начиная с единицы и до числа дуг между данными вершинами. Например, граф на рис. 2 может быть описан и так:

$$X = \{x_1, x_2, x_3, x_4\}, \quad \Gamma = \{(x_1, x_2, 101), (x_2, x_1, 1), (x_2, x_3, 4), (x_3, x_4, 5), (x_1, x_3, 10^5)\}. \quad (2.4)$$

Графы, заданные соотношениями (2.1) и (2.4), будем считать различными. Для формализации «идентичности» этих графов вводится ниже понятие изоморфизма.

Как уже сказано в определении 2.1, элементы Γ называются дугами. Дуга вида (x_i, x_i, n) называется петлей.

Несколько сложнее понятие ребра. Иногда на графе, возникающем в какой-либо прикладной задаче, нет необ-

ходимости задавать направление связи между x_i и x_j . При изображении в этом случае связь рисуют без стрелки и говорят о ребре, соединяющем x_i и x_j (рис. 5, а). Будем считать такие ребра объединением дуг (x_i, x_j, n) и (x_j, x_i, n) с одинаковыми номерами. Таким образом, *ребром* называется подмножество вида $\{(x_i, x_j, n), (x_j, x_i, n)\} \subset \Gamma$ (рис. 5, б).

Определение 2.2. Граф $G = (X, \Gamma)$ называется *симметрическим*, если для любых i, j из $(x_i, x_j, n) \in \Gamma$ следует включение $(x_j, x_i, n) \in \Gamma$.

Из определения видно, что множества вида $\{(x_i, x_j, n), (x_j, x_i, n)\}$ для симметрического графа образуют разбиение множества Γ . Множество классов этого разбиения обозначается $\bar{\Gamma}$ и называется множеством ребер симметрического графа G . Подчеркнем, что $\bar{\Gamma}$ определяется только для симметрических графов.

В дальнейшем, допуская вольность речи в случаях, где это не может привести к разночтениям, вместо «ребро $\{(x_i, x_j, n), (x_j, x_i, n)\}$ » будем говорить «ребро (x_i, x_j, n) ». Отметим, что выражения «ребро (x_i, x_j, n) » и «ребро (x_j, x_i, n) » при этом тождественны.

Определение 2.2'. Полностью неориентированным графом называется пара $G = (X, \bar{\Gamma})$, где X — конечное множество, а $\bar{\Gamma}$ — конечное подмножество прямого произведения $2^X \times \mathbb{Z}_+$, состоящее из элементов вида (A, n) , где A содержит ровно два элемента.

З а м е ч а н и е. Каждый симметрический граф однозначно определяет некоторый полностью неориентированный граф. Элементы $\bar{\Gamma}$ называются при этом ребрами.

Формализуем теперь понятие графа без кратных дуг.

Определение 2.3. Графом без кратных дуг называется (мульти) граф $G = (X, \Gamma)$, удовлетворяющий условию: для любых i, j существует самое большее один элемент $n \in \mathbb{Z}_+$ такой, что $(x_i, x_j, n) \in \Gamma$.

Для описания графов без кратных дуг можно воспользоваться упрощающей символикой — дугу (x_i, x_j, n) графа G обозначать (x_i, x_j) . Таким образом, можно дать другое определение графа без кратных дуг, эквивалентное определению 2.3.

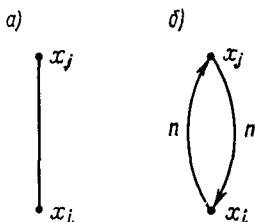


Рис. 5

Определение 2.3'. *Графом без кратных дуг называется пара объектов $G = (X, \Gamma)$, где X — конечное множество, а $\Gamma \subset X \times X$.*

Мы пользуемся обозначением Γ в определениях 2.1 и 2.3, хотя этим символом обозначены различные объекты. В подавляющем большинстве случаев это не вызывает дополнительных затруднений. Отметим еще, что соответствие между графами в смысле определений 2.3 и 2.3' не взаимно-однозначное, поскольку нумерация ребер в определении 2.3' отсутствует.

Покажем, что граф без кратных дуг определяет некоторое отображение (которое также будем обозначать буквой Γ) $\Gamma: X \rightarrow 2^X$. Каждому элементу $x_i \in X$ сопоставим подмножество $\Gamma(x_i) \in 2^X$ следующим образом. Считаем, что $x_j \in \Gamma(x_i)$ тогда и только тогда, когда $(x_i, x_j) \in \Gamma$ (здесь использовано определение 2.3'). Легко видеть, что и обратно, любое отображение $\Gamma: X \rightarrow 2^X$ определяет граф в смысле определения 2.3'. Можно дать определение эквивалентное определению 2.3'.

Определение 2.3''. *Графом без кратных дуг называется пара объектов $G = (X, \Gamma)$, где X — конечное множество, а $\Gamma: X \rightarrow 2^X$ — отображение.*

В дальнейшем будут рассматриваться и полностью неориентированные графы без кратных ребер. Ребра такого графа будем часто обозначать (x_i, x_j) или, эквивалентно (x_j, x_i) и говорить о неупорядоченной паре вершин (x_i, x_j) . Все упомянутые обозначения следует рассматривать как вольности речи (впрочем, сильно облегчающие изложение). Точные определения даны выше.

2. Некоторые общие понятия теории графов. Дугу $u \in \Gamma$ и вершину $x \in X$ назовем инцидентными, если $x = \text{Пр}_1 u$ или $x = \text{Пр}_2 u$. Ребро $\bar{u} \in \Gamma$ и вершину $x \in X$, назовем инцидентными, если инцидентны дуга $u \in \bar{u}$ и x . Дугу u , такую, что $\text{Пр}_1 u = x$, назовем исходящей из x , а если $\text{Пр}_2 u = x$, то — входящей в x . Две вершины x_i и x_j называются смежными, если существует дуга, соединяющая либо x_i с x_j , либо x_j с x_i .

Определение 2.4. *Путем в мультиграфе $G = (X, \Gamma)$ называется такая конечная последовательность дуг $\{u_1, u_2, \dots, u_m\}$, $u_i \in \Gamma$, что $\text{Пр}_2 u_i = \text{Пр}_1 u_{i+1}$ для $i = 1, 2, \dots, m-1$. Если кроме того, $\text{Пр}_2 u_m = \text{Пр}_1 u_1$, то путь $\{u_1, \dots, u_m\}$ называется контуром. Путь называется простым, если из неравенства $i \neq j$ следует $u_i \neq u_j$,*

и элементарным, если последовательность вершин $\{\text{Пр}_1 u_1, \text{Пр}_2 u_1, \dots, \text{Пр}_2 u_m\}$ не содержит совпадающих элементов, кроме, может быть, крайних.

Интерпретация пути (и контура) в мультиграфе очевидна. Путь является простым, если при движении вдоль него одна и та же дуга никогда не приходится дважды, и элементарным, если при этом никакая вершина не встречается дважды.

Пусть теперь $G = (X, \Gamma)$ — симметрический мультиграф.

Определение 2.5. Цепью в симметрическом мультиграфе* G называется такая последовательность $\{\bar{u}_1, \dots, \bar{u}_m\}$, $\bar{u}_i \in \bar{\Gamma}$ ребер графа G , что существует путь $\{u_1, \dots, u_m\}$ в G , удовлетворяющий условию $u_i \in \bar{u}_i$ для $i = 1, 2, \dots, m$. Цепь называется циклом, если соответствующий путь есть контур. Цепь называется простой {элементарной}, если соответствующий путь простой {элементарный}.

Очевидно, что для любой цепи $\{\bar{u}_1, \dots, \bar{u}_m\}$, $m \geq 2$ существует ровно один путь $\{u_1, \dots, u_m\}$, такой, что $\bar{u}_i \in \bar{u}_i$. Это доказывает корректность определений простой и элементарной цепи.

Определение 2.6. Мультиграф $G' = (X', \Gamma')$ называется суграфом мультиграфа $G = (X, \Gamma)$, если $X' \subset X$, $\Gamma' \subset \Gamma$. Граф G' называется подграфом графа G , если $X' \subset X$, $\Gamma' = \Gamma \cap (X' \times X' \times \mathbb{Z}_+)$. Наконец, граф G' называется частью графа G , если $X' \subset X$, $\Gamma' \subset \Gamma$.

Таким образом, суграф получается из графа удалением некоторого числа дуг (с сохранением вершин). Подграф получается из графа удалением некоторого числа вершин вместе со всеми дугами, инцидентными удаленной вершине. Часть графа получается из графа применением конечного числа обеих описанных операций.

Определение 2.7. Наименьший симметрический мультиграф $G_{\text{сим}}$ такой, что G является суграфом $G_{\text{сим}}$, называется симметризованным мультиграфом мультиграфа $G = (X, \Gamma)$.

Определение 2.8. Мультиграф $G = (X, \Gamma)$ называется сильно связным, если для любых двух вершин $x_i, x_j \in X$, $x_i \neq x_j$ существует путь $\{u_1, u_2, \dots, u_m\}$ такой, что $x_i = \text{Пр}_1 u_1$, $x_j = \text{Пр}_2 u_m$.

* В этом определении симметрический мультиграф может быть заменен на отвечающий ему полностью неориентированный граф; в дальнейшем такие замечания будут опускаться.

Определение 2.9. Симметрический мультиграф $G = (X, \Gamma)$ называется связным, если он сильно связан. Произвольный граф называется связным, если связан его симметризованный граф $G_{\text{сим}}$.

Из этих определений прямо следует, что понятия связности и сильной связности совпадают для симметрических мультиграфов.

Определение 2.10. Подграф G' графа G называется компонентой связности графа G , если 1) G' связан, 2) G' обладает свойством максимальности, т. е. если G'' — некоторый другой связный подграф G и $G' \subsetneq G''$, то графы G' и G'' совпадают.

Будем говорить, что ребро $\bar{u} \in \bar{\Gamma}$ является перешейком в симметрическом мультиграфе $G = (X, \Gamma)$, если при удалении этого ребра количество компонент связности графа увеличивается на единицу. Очевидно, что ребро является перешейком тогда и только тогда, когда не существует простого цикла, содержащего это ребро.

Определение 2.11. Полу степенью исхода p вершин x называется число исходящих из нее дуг. Полу степенью захода q вершины x называется число входящих в нее дуг. Степенью вершины $st\ x$ называется число $p + q$.

З а м е ч а н и е. Иногда в симметрических мультиграфах степень вершины x будет называться число инцидентных ей ребер. Из контекста всегда будет ясно значение числа $st\ x$ в каждом случае.

Определение 2.12. Ребро $\bar{u}$ симметрического мультиграфа $G = (X, \Gamma)$ называется тупиком, если степень одной из его вершин равна единице (здесь степень вершины определяется как число инцидентных ей ребер).

Определение 2.13. Пусть $G = (X, \Gamma)$ — мультиграф, $n = |X|$. Матрицей смежности G называется $n \times n$ матрица $A = \|a_{ij}\|$, в которой элемент a_{ij} равен количеству дуг из Γ вида (x_i, x_j, k) .

Определение 2.14. Мультиграфы $G = (X, \Gamma)$ и $G' = (X', \Gamma')$ называются изоморфными, если существуют взаимно-однозначные отображения $\varphi: X \rightarrow X'$, $\psi: \Gamma \rightarrow \Gamma'$, такие, что $\text{Pr}_1 \circ \psi = \varphi \circ \text{Pr}_1$, $\text{Pr}_2 \circ \psi = \varphi \circ \text{Pr}_2$.

§ 3. ЗАДАЧИ, ПОСЛУЖИВШИЕ ОСНОВОЙ ТЕОРИИ ГРАФОВ

1. **Задача о кенигсбергских мостах.** На рис. 6 схематически изображена карта города Кенигсберга, относящаяся к XVIII в. Город был расположен на берегах и двух остро-

вах реки Преголи. Острова между собой и с берегами были связаны семью мостами. Жители города любили размышлять над проблемой: можно ли, выйдя из дома, вернуться обратно, пройдя по каждому мосту только один раз?

Полностью неориентированный мультиграф G , отвечающий задаче, представлен на рис. 7. Вершины x_1, x_4 соответ-

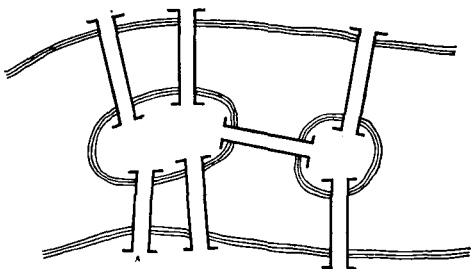


Рис. 6

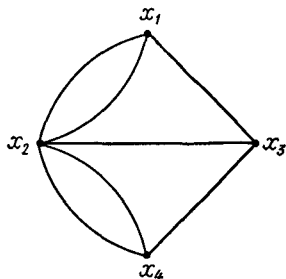


Рис. 7

ствуют правому и левому берегам реки, вершины x_2, x_3 — островам, ребра мультиграфа — мостам. На языке графов, следовательно, задача формулируется следующим образом: *существует ли в мультиграфе простой цикл, содержащий все ребра (эйлеров цикл).*

Л. Эйлер сформулировал и доказал необходимое и достаточное условие того, чтобы в произвольном полностью неориентированном связном мультиграфе существовал эйлеров цикл.

Теорема 3.1. *Эйлеров цикл в симметрическом связном мультиграфе G существует тогда и только тогда, когда степени всех его вершин — четные числа.*

Доказательство. Необходимость условия теоремы очевидна, поскольку при каждом проходе через вершину, используются ровно два ребра.

Достаточность можно доказать индукцией по числу ребер мультиграфа. При числе ребер, равном двум, как нетрудно видеть, теорема справедлива. Пусть утверждение теоремы верно для всех мультиграфов с числом ребер, не превосходящем n . Для мультиграфа с числом ребер $n + 1$ рассмотрим произвольный простой цикл.

Хотя бы один такой цикл обязательно существует для графов с четными степенями вершин. Если в графе не имеется ни одного простого цикла, то любое его ребро — пере-

шеек. Удаляя какое-либо ребро графа, инцидентное, скажем, вершинам x_i и x_j , разбиваем граф на две компоненты связности. Каждое ребро в этих компонентах связности — также перешеек. После удаления ребра степени вершин x_i и x_j стали нечетными.

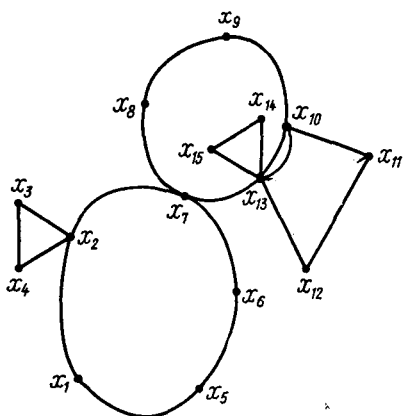


Рис. 8

Склеим компоненты связности по вершинам x_i и x_j . Новая вершина, как и все остальные, в полученном связном графе имеет четную степень. Число ребер в получившемся графе равно n . По предположению индукции в нем существует эйлеров цикл. Но это противоречит предположению, что в каждой компоненте связности все ребра — перешейки.

Если рассматриваемый простой цикл μ содержит все ребра, то теорема доказана. В противном случае найдутся вершины x_{j_i} ($i = 1, 2, \dots, k$), принадлежащие парам смежных ребер цикла, такие, что из каждой из них исходит по крайней мере одно ребро, не входящее в построенный цикл. В силу условия теоремы число ребер, исходящих из x_{j_i} и не принадлежащих циклу μ , обязательно четно. Например, в графе, изображенном на рис. 8, имеем

$$\mu = \{(x_1, x_2), (x_2, x_7), (x_7, x_{13}), (x_{13}, x_{10}), (x_{10}, x_9), (x_9, x_8), (x_8, x_7), (x_7, x_6), (x_6, x_5), (x_5, x_1)\};$$

x_{j_i} есть вершины x_2, x_{10}, x_{13} .

Рассмотрим подграф G' мультиграфа G , построенный с помощью ребер, не входящих в цикл μ . Подграф G' состоит, быть может, из нескольких компонент связности. Степени всех вершин G' — четные числа. Каждая компонента связности G_p подграфа G' удовлетворяет предположению индукции, так как содержит меньше, чем $n + 1$ ребро. Следовательно, в каждой компоненте G_p существует эйлеров цикл, содержащий, по крайней мере, одну из вер-

шин x , Присоединяя к циклу μ эйлеровы циклы подграфов G_p , получаем эйлеров цикл графа G . ■

2. Задача о четырех красках. Формулировка этой задачи чрезвычайно проста и не соответствует всей глубине и сложности проблемы: *можно ли на любой политико-административной карте раскрасить страны так, чтобы никакие две страны, имеющие общую границу, не были раскрашены одинаковой краской, и чтобы были использованы всего четыре краски?* Уточним, что если две страны граничат по точке, то они не считаются имеющими общую границу.

В терминах графов задача может быть поставлена следующим образом. Дан произвольный полностью неориентированный плоский граф G . Можно ли каждую вершину графа G раскрасить с помощью одной из четырех красок так, чтобы никакие две смежные вершины (вершины, соединенные хотя бы одним ребром) не были раскрашены в один цвет. Конечно, нетрудно привести примеры графов, которые раскрашиваются в одну, две, три или четыре краски. В гл. VI будет доказана теорема о том, что любой плоский граф может быть раскрашен с помощью пяти красок. Тем не менее, проблема четырех красок до сих пор не решена. Удалось лишь доказать, что такую раскраску можно осуществить для всех плоских графов с числом вершин, не превосходящим 38.

Задача эта приобрела известность с 1878 г., когда английский математик Кэли привел ее формулировку на заседании английского королевского научного общества; добавив, что не мог ее решить, хотя и размышлял над ней длительное время. С тех пор многие выдающиеся математики пробовали свои силы в решении этой задачи. Удивительно, что для графов, нарисованных на торе, листе Мёбиуса или бутылке Клейна (см. гл. V), соответствующая задача решена, т. е. установлено необходимое и достаточное число красок для раскрашивания.

Глава II

АЛГОРИТМИЧЕСКИЕ ЗАДАЧИ

Рассмотренные в § 3 гл. I задачи были связаны с установлением необходимых или достаточных условий существования некоторого объекта в графе (эйлерова цикла, раскраски графа в n цветов и т. п.). В этой главе познакомимся с новым типом задач, которые назовем алгоритмическими задачами теории графов.

Понятие алгоритма считается в современной математике столь же фундаментальным, как понятие множества и соответствия (функции), и не сводимым к более элементарным. Не будем пытаться формально определить, что такое алгоритм, а обратимся к примерам, чтобы у читателя появилась возможность интуитивно уяснить себе смысл понятия алгоритма решения задач теории графов.

§ 1. ЗАДАЧИ О КРАТЧАЙШИХ ПУТЯХ

1. Путь с наименьшим числом дуг. Пусть задан произвольный граф $G = (X, \Gamma)$. Требуется построить такой путь, соединяющий две заданные вершины a и b , который содержит наименьшее число дуг (путь кратчайшей длины, если считать, что длины всех дуг одинаковы).

Если граф содержит небольшое число вершин и дуг, а задачу решает человек, то искомый путь может быть непосредственно «увиден», т. е. найден без применения каких-либо явно осознаваемых правил. При значительном количестве вершин и дуг в графе возникает необходимость в четком описании способа решения задачи.

А л г о р и т м р е ш е н и я

1°. Присвоить вершине a метку 0.

2°. Если $x \neq a$ и $x \in \Gamma a$, то присвоить каждой такой вершине метку 1.

3°. Пусть $A(m)$ ($m = 0, 1, 2, \dots$) — множество вершин, имеющих метку m . Вершинам множества

$A(m+1) = \{x \in X / x \in \Gamma A(m), x \notin A(k) \text{ при } k \leq m\}$
присвоить метку $m+1$.

4°. Процесс присвоения вершинам меток прекратить, как только вершина b получит некоторую метку n , ($b \in A(n)$).

5°. Рассмотреть вершины $x_{i_1}, x_{i_2}, \dots, x_{i_n}$, такие, что $x_{i_1} \in A(n-1), b \in \Gamma x_{i_1}$,

$x_{i_2} \in A(n-2), x_{i_1} \in \Gamma x_{i_2}$,

.....

$x_{i_n} \in A(0), x_{i_{n-1}} \in \Gamma x_{i_n}$.

Путь $\mu = [a = x_{i_n}, x_{i_{n-1}}, \dots, b]$ дает решение задачи.

З а м е ч а н и е. Если на некотором шаге невозможно присвоение метки $m+1$ вершинам в силу того, что множество $\Gamma A(m)$ пусто, и вершина b не получила метки, то это означает, что в графе G не существует никакого пути, соединяющего вершину a с вершиной b .

Доказательство того, что применение правил алгоритма всегда приводит к решению задачи 1, основывается на том очевидном факте, что вершины множества $A(m)$ — это все те вершины, в которые можно попасть из вершины a по путям, содержащим ровно m дуг, и нельзя попасть по пути длины меньшей, чем m .

2. Путь кратчайшей длины. Рассмотрим теперь случай, когда каждой дуге u графа $G = (X, \Gamma)$ сопоставлено **п о л о ж и т е л ь н о е** число $l(u)$. Это число $l(u)$ можно назвать длиной дуги. *Длиной пути μ* назовем сумму длин дуг, входящих в μ :

$$L[\mu] = \sum_{u \in \mu} l(u).$$

Возникает следующая задача. *Найти в графе G путь μ кратчайшей длины, соединяющий вершину a с вершиной b .*

А л г о р и т м р е ш е н и я

1°. Перенумеровать вершины графа G так, чтобы вершина a получила номер 0. Обозначить вершину a через x_0 . (При этом вершина b совпадет с некоторой вершиной x_n .)

2°. Присвоить каждой вершине x_i метку λ_i так, чтобы $\lambda_0 = 0, \lambda_i = +\infty$ при $i > 0$.

3°. Найти такую дугу (x_i, x_j) , для которой $\lambda_j - \lambda_i > l(x_i, x_j)$. (Полагаем, что $\infty - \infty = 0$.) У вер-

шины x_j заменить метку λ_j на новую, меньшую метку $\lambda_j = \lambda_i + l((x_i, x_j))$.

4°. Применять правило 3° до тех пор, пока для каждой дуги (x_i, x_j) не станет справедливым неравенство $\lambda_j - \lambda_i \leq l((x_i, x_j))$.

5°. В множестве $\Gamma^{-1}x_n$ ($x_n = b$) найти такую вершину x_{p_1} , что

$$\lambda_n = \lambda_{p_1} + l((x_{p_1}, x_n)).$$

Аналогично, в множестве $\Gamma^{-1}x_{p_1}$ найти такую вершину x_{p_2} , чтобы было справедливо равенство

$$\lambda_{p_1} = \lambda_{p_2} + l((x_{p_2}, x_{p_1})) \text{ и т. д.}$$

После некоторого числа шагов вершина x_{p_k} совпадет с вершиной x_0 ($x_0 = a$).

Путь $\mu = [a = x_{p_k}, x_{p_{k-1}}, \dots, x_{p_1}, b]$ — кратчайший, его длина $L[\mu] = \lambda_n$.

Обоснование алгоритма

Докажем, что после конечного числа применений правила 3° для каждой дуги графа станет справедливым неравенство

$$\lambda_j - \lambda_i \leq l((x_i, x_j)).$$

Для этого заметим, что на любом этапе метки $\lambda_j > 0$ при $j \neq 0$, а метка $\lambda_0 = 0$, что можно доказать по индукции. В самом деле, при первом применении правила 3° будет изменена метка λ_k у одной из вершин x_k , смежных с вершиной x_0 ($x_k \in \Gamma x_0$). Эта вершина x_k получит новую метку $\lambda_k = l(x_0, x_k) > 0$.

Предположим, что после того как применено правило 3° n раз ($n > 1$), станет справедливым утверждение, что $\lambda_0 = 0, \lambda_i > 0$ для $i > 0$. На $n + 1$ шаге какая-то вершина x_j получит новую метку $\lambda_j = \lambda_i + l((x_i, x_j))$. Но в силу предположения индукции $\lambda_i \geq 0$, кроме того, $l((x_i, x_j)) > 0$; поэтому $\lambda_j > 0$.

Ясно, что при каждом изменении метка вершины графа уменьшается на положительную величину, не меньшую, чем минимальная разность длин путей графа.

Из этих двух утверждений вытекает, что метка любой вершины графа может изменяться лишь конечное число раз. Так как вершин конечное множество, то правило 3° может применяться лишь конечное число раз.

Докажем теперь утверждения, содержащиеся в правиле 5⁰. Вершина $x_{p_1} \in \Gamma^{-1}x_n$ такая, что $\lambda_n = \lambda_{p_1} + l((x_{p_1}, x_n))$ обязательно найдется в случае, если существует хотя бы один путь, соединяющий x_0 с x_n . Ибо тогда, как нетрудно сообразить, метка $\lambda_n < \infty$. Поэтому x_{p_1} — это, например, та вершина, которая послужила для изменения метки λ_n в последний раз.

Аналогично доказывается существование вершин x_{p_k} ($k > 1$).

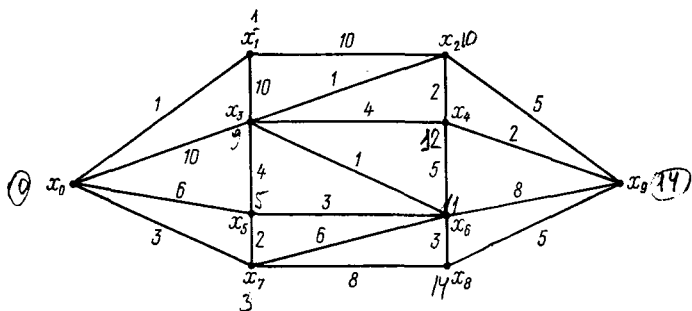


Рис. 9

По условию дуги графа имеют положительную длину, поэтому метки $\lambda_n, \lambda_{p_1}, \dots, \lambda_{p_j}, \dots$ образуют строго убывающую последовательность неотрицательных чисел, отличающихся друг от друга на величину, большую или равную длине кратчайшей дуги графа. Следовательно, какое-то $\lambda_{p_k} = 0, x_{p_k} = x_0$. (Вершина x_0 выделена тем, что ей в силу правила 2⁰ с самого начала присвоена метка $\lambda_0 = 0$ и в формировании этой метки не участвуют дуги графа.)

Докажем, что путь $\mu = [x_0, x_{p_{k-1}}, \dots, x_{p_1}, x_n]$ — кратчайший. Для этого рассмотрим произвольный путь $\nu = [x_0, x_{i_1}, \dots, x_{i_m}, x_n]$, соединяющий x_0 с x_n . Имеем неравенства

$$\lambda_n - \lambda_{i_m} \leq l((x_{i_m}, x_n)),$$

$$\lambda_{i_m} - \lambda_{i_{m-1}} \leq l((x_{i_{m-1}}, x_{i_m})),$$

.....

$$\lambda_{i_1} - \lambda_0 \leq l((x_0, x_{i_1})).$$

Складывая эти неравенства, получаем соотношение

$$\lambda_n \leq l((x_0, x_{i_1})) + \sum_{j=1}^{m-1} l((x_{i_j}, x_{i_{j+1}})) + l((x_{i_m}, x_n)) = L[v],$$

так как $\lambda_0 = 0$.

В то же время, по построению пути μ имеем:

$$L[\mu] = \lambda_n, \text{ откуда } L[\mu] \leq L[v].$$

П р и м е р. В графе G (рис. 9) легко установить, что кратчайший путь, соединяющий вершины x_0 и x_9 , имеет длину $L[\mu] = 14$.

§ 2. АЛГОРИТМ ПОСТРОЕНИЯ ЭЙЛЕРОВА ЦИКЛА

Обратимся к задаче об эйлеровом цикле, уже рассмотренной нами в предыдущей главе. Пусть G — связный мультиграф G , степени всех вершин которого — четные числа. Теорема 3.1 гл. I гарантирует существование эйлерова цикла.

Как фактически построить его? Оказывается, что достаточно выполнять следующие правила.

А л г о р и т м

1°. Выбрать произвольно некоторую вершину a .

2°. Выбрать произвольно некоторое ребро u , инцидентное a , и присвоить ему номер 1. (Назовем это ребро «пройденным».)

3°. Каждое пройденное ребро вычеркивать и присваивать ему номер, на единицу больший номера предыдущего вычеркнутого ребра.

4°. Находясь в вершине x , не выбирать ребра, соединяющего x с a , если только есть возможность другого выбора.

5°. Находясь в вершине x , не выбирать ребра, которое является «перешейком» (при удалении которого граф, образованный незачеркнутыми ребрами, распадается на две компоненты связности, имеющие хотя бы по одному ребру).

6°. После того как в графе будут занумерованы все ребра, цикл $\mu = [u_{i_1}, u_{i_2}, \dots, u_{i_n}]$, образованный ребрами с номерами от 1 до n , где n — число ребер в графе, есть эйлеров цикл.

Обоснование алгоритма

Пусть мы находимся в некоторой вершине $x \neq a$. В исходном графе G степень вершины x — четное число, поэтому после зачеркивания ребер, по которым мы приходили и уходили из вершины x , ее степень — нечетна. Следовательно, существует, по крайней мере, одно незачеркнутое ребро, инцидентное вершине x . Если это ребро — единственное, инцидентное вершине x , то оно, в силу замечания в п. 5^о алгоритма, не может быть «перешейком», и по нему можно покинуть вершину x .

Пусть ребер, инцидентных вершине x — нечетное число, большее единицы. Докажем, что среди них хотя бы одно ребро не является перешейком. Допустим противное: все ребра, инцидентные вершине x , — перешейки. Удалим одно из этих ребер, такое, чтобы вершина x и a оказались в разных компонентах связности. Такое ребро существует, так как в противном случае вершины x и a были бы связаны более чем одной простой цепью. Это означало бы, что существует простой цикл, содержащий вершины x и a . Но ребра, входящие в простой цикл, не могут быть перешейками.

Рассмотрим компоненту связности G_1 , содержащую вершину x (и не содержащую вершину a). В графе G_1 степени всех вершин, в том числе и вершины x — четные числа. Следовательно, в графе G_1 существует эйлеров цикл. Ребра, входящие в цикл, не могут быть перешейками.

Итак, наше допущение ведет к противоречию. Более того, мы убедились, что среди ребер, инцидентных вершине x в графе, полученном из графа G удалением пройденных ребер, лишь одно может быть перешейком.

Таким образом, доказано, что невозможность выполнить предписания алгоритма может возникнуть только в вершине a , если попасть в нее, по крайней мере, во второй раз. В отличие от других вершин степень вершины a при k -м попадании в нее ($k > 1$) — четна. Если эта степень равна нулю, алгоритм перестает работать.

Докажем, что в этом случае эйлеров цикл уже построен. В самом деле, в силу правила 3^о любое ребро может войти в цикл не более одного раза. В силу правил 4^о, 5^о — пройдены все ребра. Действительно, непройденные ребра определяют в графе компоненты связности. Если эти компоненты можно связать с вершиной a цепью из более чем одного зачеркнутого ребра, то среди этих ребер наверняка одно — перешеек; если одним ребром, то была возможность выбора ребра, не ведущего в вершину a .

§ 3. ЛАБИРИНТЫ

Рассмотрим задачу о поиске выхода из лабиринта, коридоры которого не обязательно находятся на одном уровне. Подобная ситуация возникает, например, при блуждании в пещерах или катакомбах. Условимся, что можно каким-либо образом отмечать, в каком направлении пройден данный коридор и какой коридор приводит на данный перекресток в первый раз. Пусть мы находимся на перекрестке a , тогда поиск выхода (перекрестка b), из лабиринта может быть описан следующим алгоритмом.

А л г о р и т м

1°. Никогда не проходить по одному и тому же коридору в одном и том же направлении дважды.

2°. Находясь на некотором перекрестке, не выбирать коридора, который привел на этот перекресток в первый раз, если только есть возможность другого выбора.

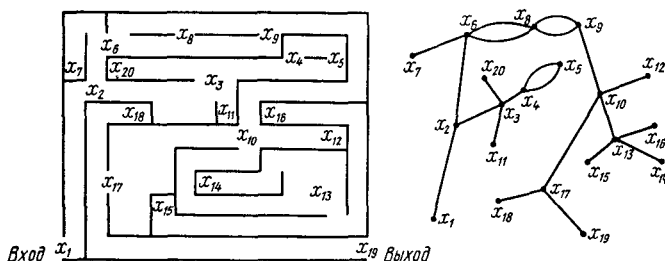


Рис. 10

Прежде чем приступить к обоснованию алгоритма, заметим, что, отождествив коридоры лабиринта с ребрами, а перекрестки, или концы тупиков с вершинами, мы приходим к связному полностью неориентированному графу, представляющему схему лабиринта (рис. 10). Таким образом, на самом деле, поставлена задача о нахождении цепи, соединяющей две заданные вершины графа. Конечно, для решения задачи можно применить алгоритм нахождения кратчайшего пути — алгоритм решения задачи п. 1. Однако этот алгоритм содержит слишком много операций. Применяя его, пришлось бы проходить по некоторым реб-

рам дважды, чтобы присвоить метку всем перекресткам, смежным с данным, если таких перекрестков больше одного.

Обоснование алгоритма

Покажем, что любой перекресток x такой, что $x \neq a$, $x \neq b$, всегда можно покинуть, не нарушая правила 1^о. Пусть мы находимся на перекрестке x в $k + 1$ раз. Тогда k коридоров, инцидентных x , окажутся пройденными в направлении от перекрестка, и $k + 1$ коридоров — по направлению к перекрестку. Следовательно, имеется, по крайней мере, один коридор, по которому не выходили из x .

Предположим теперь, что мы находимся на перекрестке a и не можем его покинуть, не нарушив правила 1^о. Покажем, что в этом случае все перекрестки лабиринта (в том числе и выход) были пройдены хотя бы один раз. Для этого присвоим перекрестку x индекс, равный числу пройденных перекрестков в момент, когда мы в x попали впервые. Вершине a присваиваем индекс 0. Оказывается, каждое ребро, инцидентное вершине с индексом p , было пройдено дважды.

Докажем это утверждение по индукции. Для перекрестка a оно справедливо, так как из a нельзя выйти. Пусть это утверждение справедливо для всех индексов $p \leq m$, докажем его для $p = m + 1$. Пусть (x_i, x_p) — ребро, по которому мы пришли в x_p в первый раз. Очевидно, $i \leq m$. Но любое ребро, инцидентное x_i , по предположению индукции, пройдено в обоих направлениях. Это же ребро инцидентно x_p , поэтому в силу правила 2^о остальные ребра, инцидентные x_p , также пройдены в обоих направлениях. Из этого утверждения вытекает, что каждая вершина, смежная с уже пройденной, также была пройдена. Следовательно, были пройдены все вершины.

§ 4. АЛГОРИТМЫ ТЕОРИИ ГРАФОВ

1. Общие свойства алгоритмов. Рассмотренные задачи позволяют сформулировать следующие свойства алгоритмов.

1. Алгоритм состоит из совокупности конечного числа правил или предписаний. Действия над графом, производимые в соответствии с правилами, должны быть достаточно простыми.

2. Применение алгоритма считается совершающимся в дискретном времени; другими словами, правила алгоритма применяются по шагам — общее число шагов конечно.

3. Какое из правил будет применено на данном шаге или какое действие будет совершено в соответствии с некоторым правилом, зависит только от результатов предыдущих шагов.

4. Некоторые алгоритмы теории графов обладают свойством локальности: действие в соответствии с правилом или установление непротиворечивости некоторого действия правилам алгоритма происходит на основе анализа дуг, инцидентных данной вершине, или вершин, смежных с данной.

5. Алгоритм обладает свойством массовости, т. е. применяется либо для всех, либо для некоторого бесконечного множества графов.

2. Реализация алгоритмов теории графов на ЭВМ. Несмотря на то, что все приведенные в настоящей главе алгоритмы обладают совокупностью свойств 1^0 — 5^0 , каждый из них имеет свою специфику, отражающуюся в его формулировке. Имеется определенное различие в степени подготовленности записи алгоритмов для их реализации на ЭВМ. Электронно-вычислительная машина может быть применена для решения какой-либо задачи теории графов только после того, как алгоритм решения будет записан в виде программы. Программа прежде всего должна предусмотреть, какие данные и в каком виде должны быть введены в память ЭВМ для их последующей обработки. Граф вводится в ЭВМ с помощью матрицы смежности. Для графа $G = (X, \Gamma)$ без кратных дуг матрица смежности — это квадратная матрица $\|a_{ij}\|$ размером $n \times n$, где n — число вершин графа. Элемент $a_{ij} = 1$, если дуга $(x_i, x_j) \in \Gamma$; $a_{ij} = 0$ — в противном случае. (Подробнее матрица смежности графов будет рассмотрена в гл. XI.) Кроме того, может потребоваться введение дополнительной информации о графе, например, матрицы расстояний в задаче п. 2.1. Далее программа должна четко предписать последовательность операций, которые нужно произвести над матрицей смежности графа, другими входными данными или вспомогательными величинами, формируемыми в процессе решения, чтобы получить ответ, а также предусмотреть, в каком виде этот ответ будет выдан ЭВМ.

После этих кратких замечаний нетрудно понять, какая большая дистанция, как правило, имеется между формули-

ровкой алгоритма теории графов и программой для ЭВМ. Из приведенных в этой главе алгоритмов ближе всего к программе первый алгоритм. Это объясняется тем, что шаги алгоритма здесь детально расписаны и даже введены в формулировку алгоритма вспомогательные массивы $A(m)$. По существу, при составлении программы требуется лишь более подробно и формализованно описать операции формирования массива $A(m+1)$, используя матрицу смежности графа и массив $A(m)$.

Алгоритм решения второй задачи менее формализован и более далек от программы, хотя также легко может быть реализован на ЭВМ, так как в его описание входят лишь операции перебора дуг, сравнения чисел и присвоения меток вершинам.

Алгоритм построения эйлерова цикла, приведенный в § 2, уже достаточно труден для записи в виде программы. Положение осложняется еще и тем, что правило 4⁰ алгоритма требует умения определять, является ли данное ребро «перешейком».

Про все алгоритмы, имеющиеся в книге, можно сказать, что они написаны на «содержательном» уровне. Там, где излишняя формализация записи алгоритма не затемняет его содержания, авторы стремились в описанном выше смысле приблизить его к программе.

Глава III

ГОМОМОРФИЗМЫ МУЛЬТИГРАФОВ. ГРУППА АВТОМОРФИЗМОВ

Гомоморфизм графов есть отображение, сохраняющее отношение инцидентности. Частным случаем гомоморфизма является автоморфизм графа, который можно наглядно представить, изобразив граф в виде сети из нитей, узлы которой отвечают вершинам графа. При этом автоморфизм — такая деформация графа, в которой запрещено рвать нити, завязывать новые узлы и развязывать старые. Точное определение этих понятий, которое здесь приводится, использует технику теории групп, поэтому даются основные определения и теоремы теории групп.

§ 1. ВСПОМОГАТЕЛЬНЫЕ СВЕДЕНИЯ ИЗ ТЕОРИИ ГРУПП

1. Определение группы. Общие понятия теории групп. Приведем сначала общее определение группы.

Определение 1.1. Множество g , в котором определено произведение (задана функция $\varphi: g \times g \rightarrow g$; $\varphi(x, y)$ или xy), называется группой, если выполнены свойства

- 1) $x(yz) = (xy)z$ (ассоциативность) и
- 2) уравнения $ax = b$, $ya = b$ однозначно разрешимы для любых $a, b \in g$.

Отметим, что не требуется, чтобы $xy = yx$. Группы, для которых $xy = yx$, называются абелевыми и для них пользуются аддитивной записью для групповой операции ($x + y$ вместо xy).

Докажем несколько свойств группы.

Теорема 1.1. Существует единственный элемент e группы g такой, что для любого $a \in g$ выполняются равенства $ae = ea = a$.

Доказательство. Пусть $a \in g$ — некоторый элемент. Обозначим через e_1 решение уравнения $xa = a$, а через e_2 — решение уравнения $ax = a$.

Пусть b — произвольный элемент g ; y_1, y_2 таковы, что $ay_1 = b$, $y_2a = b$. Тогда имеем равенства

$$e_1b = e_1ay_1 = ay_1 = b, \quad be_2 = y_2ae_2 = y_2a = b,$$

т. е. $e_1 x = x$ для любого $x \in g$, $x e_2 = x$ для любого $x \in g$. Но при этом $e_2 = e_1 e_2 = e_1$. Обозначив элемент $e_1 = e_2$ через e , получаем утверждение теоремы.

Элемент e называется *единицей группы* (для абелевых групп в аддитивной записи он называется *нулем*).

Теорема 1.2. Для любого $a \in g$ существует такой элемент $a^{-1} \in g$, что $aa^{-1} = a^{-1}a = e$.

Д о к а з а т е л ь с т в о. Пусть a^{-1} — решение уравнения $ax = e$. Рассмотрим равенство

$$a(xa) = (ax)a = ea = a.$$

Из этого равенства и из определения e следует равенство $xa = e$. Доказано, что для любого a существует обратный элемент a^{-1} , такой, что $a^{-1}a = aa^{-1} = e$ (для абелевых групп в аддитивной записи говорим об обратном элементе и обозначаем его $-a$). ■

Не продолжая далее изучение подобных свойств групп, перейдем к дальнейшим определениям.

Определение 1.2. Подгруппой в g называется такое подмножество $g' \in g$, что из $x, y \in g'$ следует $xy \in g'$ и $x^{-1} \in g'$.

Определение 1.3. Пусть g, g' — группы. Отображение $\varphi: g \rightarrow g'$ называется гомоморфизмом (групп) тогда и только тогда, когда выполняется равенство $\varphi(xy) = \varphi(x)\varphi(y)$.

Множество элементов, для которых $\varphi(x) = e'$, называется ядром гомоморфизма φ и обозначается через $\text{Ker } \varphi$, образ φ обозначается $\text{Im } \varphi$.

Теорема 1.3. Множество $\text{Ker } \varphi$ есть подгруппа в g , $\text{Im } \varphi$ есть подгруппа в g' .

Д о к а з а т е л ь с т в о. Во-первых, отметим, что справедливы соотношения

$$\varphi(e) = \varphi(ee) = \varphi(e)\varphi(e),$$

откуда $\varphi(e) = e'$ и $e \in \text{Ker } \varphi$. Далее, если $a \in \text{Ker } \varphi$, $b \in \text{Ker } \varphi$, то

$$\varphi(ab) = \varphi(a)\varphi(b) = e'e' = e',$$

$$\varphi(a^{-1}) = e'\varphi(a^{-1}) = \varphi(a)\varphi(a^{-1}) = \varphi(aa^{-1}) = \varphi(e) = e',$$

т. е. ab и a^{-1} принадлежит множеству $\text{Ker } \varphi$. Таким образом, $\text{Ker } \varphi$ есть подгруппа.

Докажем, что множество $\text{Im } \varphi$ есть подгруппа в g' . Так как $\varphi(e) = e'$, то $e' \in \text{Im } \varphi$. Пусть элементы $a', b' \in \text{Im } \varphi$. Это значит, что существуют такие a и b из g , что $a' = \varphi(a)$, $b' = \varphi(b)$. Тогда выполнены равенства

$$a'b' = \varphi(a)\varphi(b) = \varphi(ab),$$

$a' \varphi(a^{-1}) = \varphi(a) \varphi(a^{-1}) = \varphi(aa^{-1}) = \varphi(e) = e'$;
 поэтому $\varphi(a^{-1}) = a'^{-1}$ и, следовательно, $a'b' \in \text{Im } \varphi$,
 $a'^{-1} \in \text{Im } \varphi$. ■

Пусть теперь g' — некоторая подгруппа в g . Определим в g отношение эквивалентности. Считаем элемент x эквивалентным элементу y , т. е. $x \sim y$ тогда и только тогда, когда $xy^{-1} \in g'$. Проверим, что отношение $\sim$ есть отношение эквивалентности. Действительно

1) $x \sim x$, так как $xx^{-1} = e \in g'$ (рефлексивность),

2) пусть $x \sim y$, т. е. $xy^{-1} \in g'$;

но тогда выполняется соотношение $(xy^{-1})^{-1} \in g'$. С другой стороны, $(xy^{-1})(yx^{-1}) = x(y^{-1}y)x^{-1} = xx^{-1} = e$, так что получаем равенство

$$(xy^{-1})^{-1} = yx^{-1}.$$

Отсюда следует, что $yx^{-1} \in g'$, т. е. $y \sim x$ (симметричность),

3) пусть $x \sim y$, $y \sim z$. Это значит, что $xy^{-1} \in g'$, $yz^{-1} \in g'$. Из последних соотношений получим, что $xz^{-1} = (xy^{-1})(yz^{-1}) \in g'$, т. е. $x \sim z$ (транзитивность).

Обозначим класс эквивалентности, который содержит x , через $\{x\}$.

Теорема 1.4. Пусть подгруппа g' обладает свойством

$$x^{-1}g'x = g' \quad (1.1)$$

для любого $x \in g$. Тогда множество g/g' классов эквивалентности по $\sim$ образует группу относительно операции умножения:

$$\{x\}\{y\} = \{xy\}.$$

Такая операция определена корректно.

Доказательство. Опишем сначала класс эквивалентности, содержащий x . Пусть $y \in \{x\}$, или $y \sim x$. Это значит, что $yx^{-1} = h \in g'$, откуда $y = hx$, $h \in g'$. Наоборот, если $y = hx$, $h \in g'$, то $yx^{-1} = h \in g'$, т. е. $y \sim x$. Значит, когда h пробегает g' , hx пробегает $\{x\}$. Это можно символически записать в виде $\{x\} = g'x$. Множество $\{x\}\{y\}$ (множество всех произведений вида x_1y_1 , $x_1 \sim x$, $y_1 \sim y$) должно быть классом эквивалентности $\{xy\}$. Тогда это множество можно назвать произведением классов $\{x\}$ и $\{y\}$. В силу равенства $\{x\}\{y\} = g'xg'y$, для того чтобы множество $\{x\}\{y\}$ было классом эквивалентности, нужно, чтобы выполнялось равенство

$$xg' = g'x^*, \text{ или } g' = x^{-1}g'x,$$

* Для равенства $xg' = g'x$ не нужно, чтобы равенство $xh = hx$ соблюдалось для всех $h \in g'$. Достаточно, чтобы для любого $h' \in g'$ существовал $h \in g'$ такой, что $xh = h'x$

что совпадает с условием (1.1) для любого $x \in g$. В таком случае имеем

$$g'xg'y = g'g'xy = g'xy,$$

так как $g'g' = g'$ (g' — подгруппа). Единицей g/g' является множество $g'e = g'$; все аксиомы группы проверяются тривиально. Проверку предоставляем читателю.

Дадим теперь определение нормального делителя.

Определение 1.4. Подгруппа g' такая, что $x^{-1}g'x = g'$ для любого $x \in g$, называется нормальным делителем группы g .

Построенная выше группа классов эквивалентности называется фактор-группой g по g' .

З а м е ч а н и е. В случае, когда группа g коммутативна, условие $x^{-1}g'x = g'$ выполняется автоматически, ибо $x^{-1}g'x = g'x^{-1}x = g'$. В аддитивной записи $x \sim y$ значит $x - y \in g'$. Любая подгруппа коммутативной группы есть ее нормальный делитель.

Задача 1. Доказать, что множество всех взаимно-однозначных отображений конечного множества X в себя есть группа с композицией отображений в качестве групповой операции. Такая группа обозначается S_n ($n = |X|$).

Задача 2. При каком числе элементов множества группа S_n коммутативна?

Задача 3. Пусть g', g'' — группы. Доказать, что множество $g' \times g''$ с операцией умножения $(x', x'')(y', y'') = (x'y', x''y'')$ образует группу. Эта группа называется прямой суммой g' и g'' и обозначается $g' \oplus g''$.

2. Некоторые сведения из теории абелевых групп. Проведем теперь построение, которое потребуется в гл. IV. Пусть S — произвольное конечное множество. Рассмотрим множество формальных линейных комбинаций элементов из множества S с целыми коэффициентами. Обозначим это множество через $g(S)$. Иными словами, если $S = \{s_1, s_2, \dots, s_n\}$, то элементами $g(S)$ будут формальные выражения вида

$$\sum_{i=1}^n k_i s_i.$$

Определим сумму элементов из множества $g(S)$ формулой

$$\sum_{i=1}^n k_i s_i + \sum_{i=1}^n k'_i s_i = \sum_{i=1}^n (k_i + k'_i) s_i.$$

Очевидно, что $g(S)$ с этой операцией есть абелева группа.

Определение 1.5. Построенная выше группа $g(S)$ называется свободной абелевой группой над множеством S .

Определим теперь понятие изоморфизма групп. В дальнейшем изоморфные группы будут рассматриваться как одинаковые.

Определение 1.6. Группы g и g' называются изоморфными, если существует гомоморфизм $\varphi: g \rightarrow g'$, обладающий обратным.

Теорема 1.5. Любая конечная группа g' изоморфна фактор-группе свободной группы $g(g')$.

Доказательство. Рассмотрим свободную группу $g(g')$. Ее элементы — формальные суммы $\sum k_i x_i$, где $x_i \in g'$. Поставим в соответствие каждой такой сумме соответствующую сумму элементов g' , полагая

$$k_i x_i = x_i + \dots + x_i \text{ } (k_i \text{ раз}), \text{ если } k_i > 0;$$

$$k_i x_i = 0, \text{ если } k_i = 0$$

и

$$k_i x_i = (-x_i) + \dots + (-x_i) \text{ } (-k_i \text{ раз}), \text{ если } k_i < 0.$$

Полученное отображение $\varphi: g(g') \rightarrow g'$, очевидно, есть гомоморфизм, причем $\text{Im } \varphi = g'$.

Покажем, что группа $g(g')/\text{Ker } \varphi$ изоморфна g' . Рассмотрим гомоморфизм $\psi: g(g')/\text{Ker } \varphi \rightarrow g'$, определенный так: $\psi(\{x\}) = \varphi(x)$. Это определение корректно, так как если взять $y \in \{x\}$, так что $\{x\} = \{y\}$, то

$$y - x \in \text{Ker } \varphi, \text{ т. е. } \varphi(x - y) = 0, \text{ или } \varphi(x) = \varphi(y).$$

Покажем, что существует гомоморфизм ψ^{-1} . Если $a \in g'$, то множество всех $x \in g(g')$ таких, что $\varphi(x) = a$ есть класс эквивалентности по $\text{Ker } \varphi$, так как если $\varphi(x_1) = a$, $\varphi(x_2) = a$, то $\varphi(x_1 - x_2) = 0$, т. е. $x_1 - x_2 \in \text{Ker } \varphi$. Такое отображение, очевидно, и будет ψ^{-1} .

Задача 4. Пусть $\mathbb{Z}$ — группа целых чисел, $m\mathbb{Z}$ — подгруппа чисел, делящихся на m . Показать, что группа $\mathbb{Z}_m = \mathbb{Z}/m\mathbb{Z}$ изоморфна группе корней m -й степени из 1 с умножением в качестве групповой операции.

Задача 5. Пусть в последовательности групп и гомоморфизмов

$$g_1 \xrightarrow{\varphi} g_2 \xrightarrow{\psi} g_3,$$

$$\text{Ker } \varphi = \{0_{g_1}\}, \text{ Ker } \psi = \text{Im } \varphi, \text{ Im } \psi = g_3.$$

Доказать, что группа g_3 изоморфна $g_2/\text{Im } \varphi$.

§ 2. ГОМОМОРФИЗМЫ, ИЗОМОРФИЗМЫ И АВТОМОРФИЗМЫ ГРАФОВ

1. Основные определения. Пусть G — мультиграф, т. е. $G = (X, \Gamma)$, Γ — конечное подмножество в $X \times X \times \mathbb{Z}_+$ ($\mathbb{Z}_+$ — множество натуральных чисел). Обозначим через Pr_i проекцию прямого произведения $X \times X \times \mathbb{Z}_+$ на i -й сомножитель. Например,

$$\text{Pr}_1(x_1, x_2, n) = x_1 \in X, \text{Pr}_2(x_1, x_2, n) = x_2 \in X, \text{Pr}_3(x_1, x_2, n) = n \in \mathbb{Z}_+.$$

Определение 2.1. Гомоморфизмом Φ мультиграфа $G = (X, \Gamma)$ в мультиграф $G' = (X', \Gamma')$ называется пара отображений

$$\Phi = (\varphi, \psi); \quad \varphi : X \rightarrow X', \quad \psi : \Gamma \rightarrow \Gamma'$$

таких, что

$$\text{Pr}_1 \circ \psi = \varphi \circ \text{Pr}_1; \quad \text{Pr}_2 \circ \psi = \varphi \circ \text{Pr}_2. \quad (2.1)$$

Наглядно φ есть отображение вершин G в вершины G' , ψ — отображение дуг G в дуги G' . Требования (6.1) сводятся к тому, чтобы образ дуги оставался инцидентным образам его концов. Таким образом, гомоморфизм графов — это «деформация» с допущенными операциями склеивания вершин и дуг.

Определение 2.2. Композицией гомоморфизмов $\Phi: G \rightarrow G'$ и $\Phi': G' \rightarrow G''$ называется гомоморфизм $\Phi' \circ \Phi = (\varphi' \circ \varphi, \psi' \circ \psi)$.

Тривиальным образом проверяется, что $\Phi' \circ \Phi$ — гомоморфизм графов. Этот гомоморфизм отвечает последовательному проведению деформаций, отвечающих гомоморфизмам Φ и Φ' .

Определение 2.3. Гомоморфизм $\Phi^{-1}: G' \rightarrow G$ называется обратным к гомоморфизму $\Phi: G \rightarrow G'$, если

$$\Phi^{-1} \circ \Phi = \text{id}_G = (1_X, 1_\Gamma); \quad \Phi \circ \Phi^{-1} = \text{id}_{G'} = (1_{X'}, 1_{\Gamma'}).$$

Определение 2.4. Гомоморфизм Φ называется изоморфизмом, если он имеет обратный. Изоморфизм Φ графа G в себя называется автоморфизмом графа G .

Очевидно, что если Φ обладает обратным Φ^{-1} и $\Phi = (\varphi, \psi)$, то $\Phi^{-1} = (\varphi^{-1}, \psi^{-1})$, т. е. φ и ψ — взаимно-однозначные отображения. Отсюда ясно, что деформации, от-

вещающие изоморфным отображениям графов, проводятся без склеиваний вершин и ребер.

Приведем несколько примеров гомоморфизмов, изоморфизмов и автоморфизмов.

1. Граф G_2 , изображенный на рис. 11, б, является гомоморфным образом графа G_1 , изображенного на рис. 11, а. Гомоморфизм задается отображениями

$$\begin{array}{ll} x_1 \rightarrow y_2 & (x_1, x_2) \rightarrow (y_2, y_2) \\ \varphi : x_2 \rightarrow y_2 & \psi : (x_2, x_3) \rightarrow (y_2, y_1) \\ x_3 \rightarrow y_1 & (x_3, x_1) \rightarrow (y_1, y_2); \end{array}$$

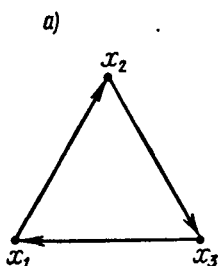


Рис. 11

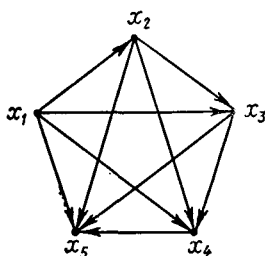
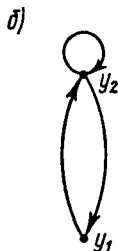


Рис. 12

это — гомоморфизм на. Если G_1 и G_2 — графы без кратных дуг, то гомоморфизм Φ всегда задается отображением φ . Не следует, однако, думать, что любое отображение вершин φ можно продолжить до гомоморфизма. Например, в рассматриваемом случае отображение

$$\varphi : x_1 \rightarrow y_1, x_2 \rightarrow y_1, x_3 \rightarrow y_2$$

не может быть продолжено до гомоморфизма графов (докажите это!).

2. Пусть дан граф G , изображенный на рис. 18. Как указывалось в конце примера 1, для задания гомоморфизма на простом графе достаточно задать отображение на вершинах. Легко проверить, что подстановки

$$\varphi_1 : \begin{pmatrix} x_1 x_2 x_3 x_4 x_5 \\ x_2 x_3 x_4 x_5 x_1 \end{pmatrix} \text{ и } \varphi_2 : \begin{pmatrix} x_1 x_2 x_3 x_4 x_5 \\ x_3 x_4 x_5 x_1 x_2 \end{pmatrix}$$

задают автоморфизмы графа G . Пусть Φ_1 — автоморфизм, отвечающий φ_1 , Φ_2 — автоморфизм, отвечающий φ_2 . Можно доказать (сделайте это!), что любой автоморфизм Φ

графа G может быть представлен в виде $\Phi = \Phi_1^\alpha \circ \Phi_2^\beta$, где $\Phi_1^\alpha = \Phi_1 \circ \dots \circ \Phi_1$ (α сомножителей) при $\alpha > 0$ и $\Phi_1^\alpha = \Phi_1^{-1} \circ \dots \circ \Phi_1^{-1}$ ($|\alpha|$ сомножителей) при $\alpha < 0$, Φ_2^β определяется аналогично, и $\Phi_1 \circ \Phi_2 = \Phi_2 \circ \Phi_1$.

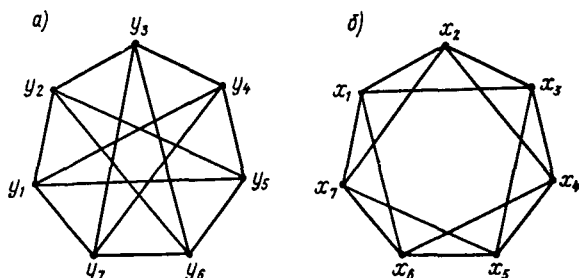


Рис. 13

3. Графы G_1 и G_2 , изображенные на рис. 13, а, б, изоморфны; их изоморфизм определяется подстановкой

$$\varphi : \begin{pmatrix} x_1 x_2 x_3 x_4 x_5 x_6 x_7 \\ y_1 y_4 y_7 y_3 y_6 y_2 y_5 \end{pmatrix}.$$

4. Рассмотрим мультиграф G , изображенный на рис. 14.

Подстановке

$$\varphi : \begin{pmatrix} x_1 x_2 x_3 \\ x_2 x_3 x_1 \end{pmatrix}$$

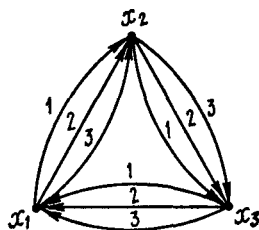


Рис. 14

отвечает не один автоморфизм графа G . Так, подстановка φ вместе с отображениями

$$\begin{aligned} (x_1, x_2, 1) &\rightarrow (x_2, x_3, 3) \\ (x_1, x_2, 2) &\rightarrow (x_2, x_3, 2) \\ \psi_1 : (x_2, x_3, i) &\rightarrow (x_3, x_1, i) \text{ и } \psi_2 : (x_1, x_2, 3) \rightarrow (x_2, x_3, 1) \\ (x_3, x_1, i) &\rightarrow (x_1, x_2, i) \\ (x_2, x_3, i) &\rightarrow (x_3, x_1, i) \\ (x_3, x_1, i) &\rightarrow (x_1, x_2, i) \end{aligned}$$

образует два различных автоморфизма графа G (подсчитайте число различных автоморфизмов с φ на первом месте).

Разобранные примеры подсказывают возможность описания группы автоморфизмов мультиграфа G с помощью отображения ее в группу подстановок на множестве вершин.

Теорема 2.1. *Множество всех автоморфизмов графа G с композицией в качестве групповой операции образует группу*.*

Доказательство. Равенство

$$(\Phi_1 \circ \Phi_2) \circ \Phi_3 = \Phi_1 \circ (\Phi_2 \circ \Phi_3), \quad \text{т. е.} \quad ((\varphi_1 \circ \varphi_2) \circ \varphi_3,$$

$(\psi_1 \circ \psi_2) \circ \psi_3) = (\varphi_1 \circ (\varphi_2 \circ \varphi_3), \psi_1 \circ (\psi_2 \circ \psi_3))$ следует из общих свойств отображений. Далее, очевидно, что

$$\Phi \circ (1_X, 1_\Gamma) = \Phi$$

для любого автоморфизма Φ графа G . По определению, для каждого автоморфизма Φ существует обратный Φ^{-1} , т. е. такой, что

$$\Phi \circ \Phi^{-1} = \Phi^{-1} \circ \Phi = \text{id}_G = (1_X, 1_\Gamma).$$

С помощью этого доказывается однозначная разрешимость уравнений $\Phi \circ X = \Psi$, $Y \circ \Phi = \Psi$ (проведите это доказательство).

Теорема 2.2. *Пусть $G = (X, \Gamma)$ — мультиграф, $n = |X|$. Существует естественный гомоморфизм*

$$f: \text{Aut}(G) \rightarrow S_n,$$

где $\text{Aut}(G)$ — группа автоморфизмов; G, S_n — группа подстановок порядка n . Этот гомоморфизм задается формулой

$$f(\Phi) = f((\Phi, \Psi)) = \varphi.$$

Ядром гомоморфизма является прямая сумма

$$\text{Ker } f = \bigoplus_{u \in X \times X} S_{n(u)}, \quad (2.2)$$

где $n(u) = |N(u)|$, $N(u) = \{a \in Z_+ \mid (u, a) \in \Gamma\}$, т. е.

$N(u)$ — множество номеров дуг, соединяющих x_i с x_j , где $u = (x_i, x_j)$.

Доказательство. Введем еще множество

$$N^*(u) = \{x \in \Gamma \mid x = (u, a), a \in N(u)\},$$

т. е. множество дуг, соединяющих данную пару вершин. Легко видеть, что f — гомоморфизм групп. Нужно доказать лишь последнюю часть теоремы — равенство (2.2).

* В дальнейшем эта группа будет обозначаться $\text{Aut}(G)$.

Пусть $\Phi \in \text{Ker } f$, т. е. $f(\Phi) = 1_X$; тогда $\Phi = (1_X, \psi)$. В силу (2.1) имеем:

$$\text{Pr}_i \circ \psi = 1_X \circ \text{Pr}_i = \text{Pr}_i \quad (i = 1, 2),$$

поэтому, если $x \in N^*(u)$, то $\psi(x) \in N^*(u)$. Отсюда можно рассмотреть сужение ψ на $N^*(u)$:

$$\bar{\psi}_u : N^*(u) \rightarrow N^*(u),$$

причем $\bar{\psi}_u$ — взаимно-однозначное отображение. Таким образом, каждый элемент $\Phi \in \text{Ker } f$ определяет элемент

$$(\bar{\psi}_u \mid u \in X \times X)$$

группы $\bigoplus_{u \in X \times X} S_{n(u)}$, причем каждая группа $S_{n(u)}$ реализована как группа подстановок на множестве $N^*(u)$. Наоборот, если $(\bar{\psi}_u \mid u \in X \times X)$ — элемент этой группы, то $\Phi = (1_X, \psi)$, где $\bar{\psi} \equiv \bar{\psi}_u$ на $N^*(u)$ — элемент $\text{Ker } f$. ■

Следствие 2.1. *Группа автоморфизмов любого графа конечна.*

Следствие 2.2. *Группа автоморфизмов графа без кратных дуг есть подгруппа S_n , где $n = |X|$, $G = (X, \Gamma)$.*

Докажем, что в следствии 2.1 сформулировано единственное характеристическое свойство групп автоморфизмов мультиграфов.

Теорема 2.3. *Для любой конечной группы g существует мультиграф G такой, что $g = \text{Aut } (G)$.*

На самом деле мы докажем, что G можно выбрать полностью неориентированным графом без кратных ребер. Доказательство будет проведено конструктивно. Удобно провести доказательство в два приема. Докажем сначала лемму.

Лемма 2.1. *Для любой конечной группы g существует полностью неориентированный граф G с раскрашенными ребрами такой, что группа g изоморфна группе автоморфизмов графа G , сохраняющих раскраску ребер.*

Доказательство. В качестве графа G возьмем граф

$$G = (g, g \times g \setminus E),$$

где E — диагональ, $E = ((\alpha, \alpha) \mid \alpha \in g)$. Сопоставим каждому элементу группы g свою краску и раскрасим ребро (α, β) в краску, соответствующую элементу $\alpha^{-1}\beta$ группы g . Докажем два утверждения.

1°. Каждый элемент группы g определяет автоморфизм графа G , сохраняющий раскраску ребер.

Рассмотрим некоторый элемент $\alpha_0 \in g$. Отображение $\varphi_{\alpha_0} : g \rightarrow g$, $\varphi_{\alpha_0}(\alpha) = \alpha_0 \alpha$ является взаимно-однозначным, так как левое умножение на α_0^{-1} есть обратное отображение к $\varphi_{\alpha_0} : \varphi_{\alpha_0} \circ \varphi_{\alpha_0}^{-1} = \varphi_{\alpha_0}^{-1} \circ \varphi_{\alpha_0} = 1_g$. Далее, так как G — простой граф, то отображению φ_{α_0} отвечает не более одного автоморфизма графа. Поскольку G — полный граф, то, как нетрудно проверить, любая подстановка вершин φ определяет автоморфизм $\Phi = (\varphi, \psi)$ графа по формуле $\psi((\alpha, \beta)) = (\varphi(\alpha), \varphi(\beta))$.

Проверим, что автоморфизм, определяемый φ_{α_0} , сохраняет раскраску. Отображение ψ переводит ребро (α, β) , раскрашенное в цвет, соответствующий $\alpha^{-1}\beta$, в ребро $(\varphi_{\alpha_0}(\alpha), \varphi_{\alpha_0}(\beta)) = (\alpha_0 \alpha, \alpha_0 \beta)$, раскрашенное в цвет, соответствующий элементу $(\alpha_0 \alpha^{-1})(\alpha_0 \beta) = \alpha^{-1} \alpha_0^{-1} \alpha_0 \beta = \alpha^{-1} \beta$. ■

2°. Каждый автоморфизм графа G , сохраняющий раскраску ребер, определяется единственным элементом группы g .

Пусть $\Phi = (\varphi, \psi)$ — автоморфизм G , сохраняющий раскраску. Пусть e — единица группы g . Покажем, что для фиксированного α_0 существует точно один автоморфизм G , сохраняющий раскраску, и такой, что $\varphi(e) = \alpha_0$. Заметим, что в каждой вершине имеется ровно одно входящее ребро каждого цвета и ровно одно выходящее. Пусть α — произвольная вершина графа. Она лежит на конце выходящего из e ребра цвета α . После преобразования она должна лежать на выходящем из α_0 ребре цвета α , например, на (α_0, β) . Но (α_0, β) помечено цветом $\alpha_0^{-1}\beta$, откуда $\alpha_0^{-1}\beta = \alpha$ и $\beta = \alpha_0 \alpha$. Доказано, что $\varphi(\alpha) = \alpha_0 \alpha$. ■

Из утверждений 1° и 2° вытекает, что автоморфизмы, сохраняющие раскраску, образуют подгруппу группы g . Кроме того, каждому такому автоморфизму отвечает элемент α_0 такой, что $\varphi(\alpha) = \alpha_0 \alpha$. Соответствие $\varphi \rightarrow \alpha_0$ взаимно-однозначно. Это соответствие есть гомоморфизм, так как если $\varphi \rightarrow \alpha_0$, $\psi \rightarrow \beta_0$, то $\varphi_0 \psi \rightarrow \alpha_0 \beta_0$; отсюда следует, что оно есть изоморфизм.

Следующим шагом в доказательстве теоремы 2.3 является построение произвольного конечного числа графов, неизоморфных друг другу и имеющих тривиальную группу автоморфизмов. Это можно сделать многими способами. Для доказательства используем графы G_j , изображенные на рис. 15.

Доказательство теоремы 2.3. Пусть задана группа g . Построим для нее раскрашенный граф, как в лемме 6.1. Перенумеруем элементы группы g произвольным образом и заменим каждое ребро, раскрашенное в цвет α_i , графом G_i , заменяя начало ребра на x_1 , а конец на x_4 . Теперь при любом автоморфизме полученного графа каждый из подграфов G_i должен переходить в подграф G_i с тем же номером. Отсюда вытекает, что каждый такой автоморфизм порождается единственным автоморфизмом G , сохраняющим раскраску. Применение леммы 6.1 завершает доказательство.

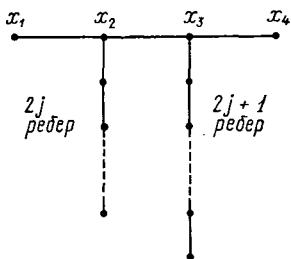


Рис. 15

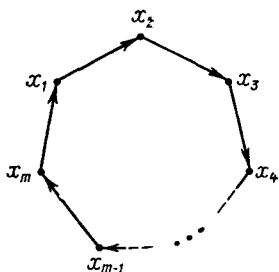


Рис. 16

2. Вычисление групп автоморфизмов для некоторых графов. Приведем теперь примеры вычисления группы $\text{Aut}(G)$ для некоторых конкретных графов G .

Пример 1. Вычислить группу автоморфизмов простого ориентированного цикла длины m (рис. 16).

Пусть имеется некоторый автоморфизм Φ этого графа. Пусть вершина x_1 переходит при этом автоморфизме в вершину x_{k+1} . Поскольку из вершины x_1 исходит ровно одно ребро с концом x_2 , то x_2 должно перейти в x_{k+2} *. Аналогично получаем, что каждая вершина с номером i переходит в вершину с номером j , являющимся остатком от деления $k+i$ на m ($j = k+i$ если $k+i \leq m$). Иными словами, каждый автоморфизм этого графа продолжается циклической подстановкой вершин. Наоборот, очевидно, что каждая циклическая подстановка вершин порождает автоморфизм графа. Любая циклическая подстановка есть степень подстановки

$$\sigma = \begin{pmatrix} x_1 & x_2 & \dots & x_m \\ x_2 & x_3 & \dots & x_1 \end{pmatrix},$$

и при этом σ^m есть тождественная подстановка. Окончательно,
 $\text{Aut}(G) = \mathbb{Z}_m$.

* Это верно, если $k+1 \neq m$. В случае $k+1 = m$ вершина x_2 переходит в x_1 .

Пример 2. Вычислить группу автоморфизмов простого неориентированного цикла длины m (рис. 17).

Рассуждения, аналогичные примеру 6.1, показывают, что если

$$\varphi(x_1) = x_{k+1},$$

то $\varphi(x_2)$ может быть равно либо x_k , либо x_{k+2} (x_1 , если $k+1=m$). Образы дальнейших вершин определяются однозначно. Таким образом, любая подстановка, отвечающая автоморфизму графа G , может быть однозначным образом представлена в виде $\sigma_1^n \sigma_2^k$, где

$$\sigma_1 = \begin{pmatrix} x_1 x_2 \cdots x_m \\ x_2 x_3 \cdots x_1 \end{pmatrix}; \quad \sigma_2 = \begin{pmatrix} x_1 x_2 \cdots x_m \\ x_m x_{m-1} \cdots x_1 \end{pmatrix} \quad (n = 0, 1, \dots, m-1; \quad k = 0, 1).$$

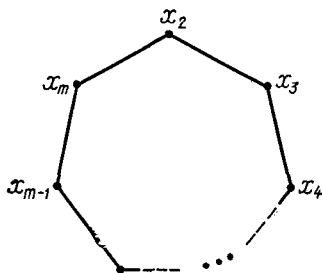


Рис. 17

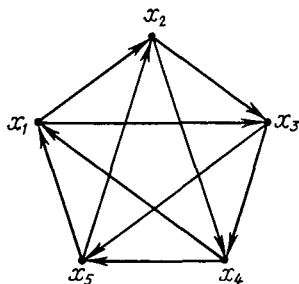


Рис. 18

Однако из этого нельзя заключить, что $\text{Aut}(G) = \mathbb{Z}_m \oplus \mathbb{Z}_2$, ибо $\mathbb{Z}_m \oplus \mathbb{Z}_2$ — абелева группа, а σ_1 и σ_2 не коммутируют: $\sigma_1 \sigma_2 = \sigma_2 \sigma_1^{-1} \neq \sigma_2 \sigma_1$.

Можно описать группу $\text{Aut}(G)$, задав каждый ее элемент парой показателей (n, k) в его разложении $\sigma_1^n \sigma_2^k$. При этом умножение определится формулами:

$$(n_1, 1)(n_2, k) = (n_1 - n_2, 1 + k),$$

$$(n_1, 0)(n_2, k) = (n_1 + n_2, k),$$

где сложение в первой компоненте производится по модулю m , а во втором — по модулю 2.

Пример 3. Вычислить группу автоморфизмов графа G (рис. 18).

Аналогичными рассуждениями (см. пример 2) получим, что любой автоморфизм графа G может быть определен подстановкой вида $\sigma_1^n \sigma_2^k$, где $n, k = 0, 1, 2, 3, 4$ и

$$\sigma_1 = \begin{pmatrix} x_1 x_2 x_3 x_4 x_5 \\ x_2 x_3 x_4 x_5 x_1 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} x_1 x_2 x_3 x_4 x_5 \\ x_3 x_4 x_5 x_1 x_2 \end{pmatrix},$$

причем, в отличие от примера 6 $\sigma_1\sigma_2 = \sigma_2\sigma_1$. Но из соотношения $\sigma_2 = \sigma_1^2$ немедленно следует, что

$$\text{Aut}(G) = \mathbb{Z}_5.$$

Пример 4. Пусть даны графы G_1, G_2 (рис. 19, а, б). Изоморфизм этих графов установлен в примере 2 этой главы. Доказать, что группы $\text{Aut}(G_1)$ и $\text{Aut}(G_2)$ при этом изоморфны и вычислить эти группы.

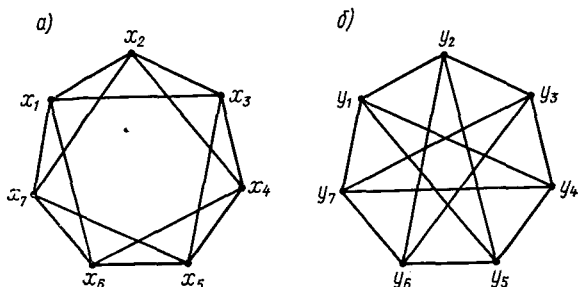


Рис. 19

Пусть $\bar{\Phi}: G_1 \rightarrow G_2$ — изоморфизм. Тогда соответствие

$$\text{Aut}(G_1) \ni \Phi \rightarrow \bar{\Phi}^{-1}\Phi\bar{\Phi} \in \text{Aut}(G_2)$$

устанавливает изоморфизм групп $\text{Aut}(G_1)$ и $\text{Aut}(G_2)$ (докажите это!).

Вычислим теперь группу $\text{Aut}(G_1)$. Как и ранее, любой автоморфизм Φ определяется подстановкой φ . Пусть при φ элемент x_1 переходит в x_k . Тогда x_2 может переходить в вершины, соседние с x_k в одной из последовательностей:

$$x_1x_2x_3x_4x_5x_6x_7x_1,$$

$$x_1x_3x_5x_7x_2x_4x_6x_1.$$

Позиции дальнейших вершин определяются уже однозначно. Таким образом, можно убедиться, что любой автоморфизм Φ определяется одной из подстановок вида $\Phi = \sigma_1^k \sigma_2^e \sigma_3^m (k, e=0, 1, \dots, 6; m=0, 1)$, где

$$\sigma_1 = \begin{pmatrix} x_1x_2x_3x_4x_5x_6x_7 \\ x_2x_3x_4x_5x_6x_7x_1 \end{pmatrix}; \quad \sigma_2 = \begin{pmatrix} x_1x_2x_3x_4x_5x_6x_7 \\ x_3x_4x_5x_6x_7x_1x_2 \end{pmatrix};$$

$$\sigma_3 = \begin{pmatrix} x_1x_2x_3 \dots x_7 \\ x_7x_6x_5 \dots x_1 \end{pmatrix}.$$

Заметим, что $\sigma_1^2 = \sigma_2$. Аналогично примеру 6 получим, что группа может быть представлена элементами (n, k) , где $n = 0, 1, \dots, 6; k = 0, 1$ и умножение задается равенствами

$$(n, 1)(n', k) = (n - n', 1 + k),$$

$$(n, 0)(n', k) = (n + n', k);$$

в первой компоненте сложение производится по модулю 7, а во второй — по модулю 2.

В заключение этого параграфа сформулируем теорему, доказательство которой предоставляется читателю.

Теорема 2.4. *Если граф G состоит из двух несвязных и неизоморфных компонент с группами автоморфизмов g_1 и g_2 , то группа $\text{Aut}(G)$ есть прямая сумма групп g_1 и g_2 :*

$$\text{Aut}(G) = g_1 \oplus g_2.$$

ЦИКЛЫ В ГРАФАХ. ЦИКЛОМАТИЧЕСКОЕ ЧИСЛО ГРАФА

Во многих прикладных задачах существенны свойства графов, связанные с существованием в графе замкнутых цепей (циклов). К рассмотрению этих вопросов мы и приступим. Все графы данной главы предполагаются полностью неориентированными и не содержащими кратных ребер. Условие отсутствия кратных ребер не является существенным и введено, главным образом, для упрощения рассуждений.

§ 1. ЦИКЛОМАТИЧЕСКОЕ ЧИСЛО ГРАФА. ДЕРЕВЬЯ

1. Определение дерева. Прежде чем дать определение дерева, введем важную характеристику графа — цикломатическое число.

Определение 1.1. Пусть граф $G = (X, \Gamma)$ имеет p компонент связности. Цикломатическим числом графа G называется число

$$\nu(G) = S_1 - S_0 + p,$$

где S_0 — число вершин графа, S_1 — число его ребер.

Одним из свойств цикломатического числа является его монотонность. Дадим строгую формулировку этого свойства.

Теорема 1.1. Пусть граф G' является суграфом графа G , тогда выполняется неравенство

$$\nu(G') \leq \nu(G).$$

Доказательство. Очевидно, достаточно рассмотреть случай, когда граф G' получен из графа G удалением одного из ребер. В этом случае $S'_1 = S_1 - 1$, $S'_0 = S_0$. Если выбрасываемое ребро является перешейком, то $p' = p + 1$ и, следовательно,

$$\begin{aligned} \nu(G') &= S'_1 - S'_0 + p' = (S_1 - 1) - S_0 + (p + 1) = \\ &= S_1 - S_0 + p = \nu(G). \end{aligned}$$

Если же выбрасываемое ребро не перешеек, то $p' = p$ и

$$\begin{aligned} \nu(G') &= S_1' - S_0' + p' = (S_1 - 1) - S_0 + p = S_1 - S_0 + \\ &+ p - 1 = \nu(G) - 1 < \nu(G). \blacksquare \end{aligned}$$

Уже из доказательства этой теоремы выясняется смысл цикломатического числа. Видно, что при удалении ребра цикломатическое число уменьшается на единицу в том и только в том случае, когда удаляемое ребро не является перешейком. Но если ребро l не есть перешеек, оно обязательно входит в какой-либо цикл. Действительно, из того,

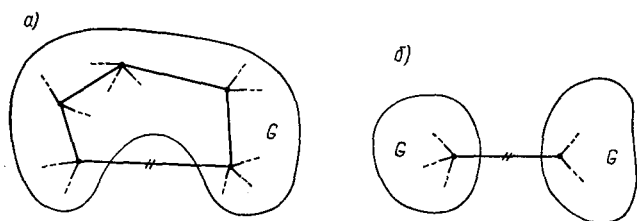


Рис. 20

что удаление этого ребра не нарушает связность графа, следует, что его концы соединены некоторой цепью в графе с удаленным ребром l . Добавляя к этой цепи ребро l , получим цикл (например, на рис. 20, а отмеченное ребро не есть перешеек, а на рис. 20, б — перешеек). Таким образом, цикломатическое число связано в каком-то смысле с количеством циклов в графе. Точное утверждение относительно этой связи будет сформулировано в § 2.

Следствие 1.1. Для любого графа G выполнено неравенство $\nu(G) \geq 0$.

Доказательство. Пусть $G = (X, \Gamma)$ — некоторый граф. Очевидно $\bar{G} = (X, \emptyset)$ (граф с теми же вершинами, что и G , но без ребер) есть суграф графа G . Используя теорему 1.1, имеем

$$\nu(\bar{G}) \leq \nu(G).$$

Но для $\bar{G}$ справедливо соотношение $p = S_0$, откуда следует формула

$$\nu(\bar{G}) = S_1 - S_0 + p = 0 - p + p = 0. \blacksquare$$

Теорема 1.2. Следующие свойства графа G эквивалентны:

1) граф G связан и не имеет циклов; 2) граф G связан и $v(G) = 0$; 3) граф G связан и $S_0 = S_1 + 1$; 4) $S_0 = S_1 + 1$ и $v(G) = 0$.

Доказательство. Заметив, что для связного графа ($p = 1$) равенство $v(G) = 0$ влечет равенство $S_0 = S_1 + 1$; $S_0 = S_1 + 1$ — равенство $v(G) = S_1 - S_0 + 1 = 0$, а из условий $S_0 = S_1 + 1$ и $v(G) = 0$ (без предположения о связности) вытекает, что $p = 1$, так как

$$0 = S_1 - S_0 + p = S_1 - (S_1 + 1) + p = p - 1$$

и, следовательно, граф связан. Таким образом, свойства 2), 3) и 4) эквивалентны.

Покажем, что из свойства 1) следует свойство 3). Если свойство 1) выполнено, то при удалении любого из ребер графа G (без удаления вершин) количество компонент связности увеличивается на единицу. Первоначальное количество компонент связности равно единице. После удаления S_1 ребер число компонент связности станет равным $S_1 + 1$. С другой стороны, после удаления S_1 ребер мы получим граф с тем же количеством вершин S_0 , что и у исходного, но без ребер. Число компонент связности такого графа есть S_0 , откуда $S_0 = S_1 + 1$.

Докажем теперь, что из свойства 3) вытекает 1). Доказательство проведем от противного. Пусть G содержит некоторый цикл. Удаляя ребро этого цикла (любое), мы не нарушаем связности графа. Для полученного графа G'

$$S'_0 = S_0, \quad S'_1 = S_1 - 1, \quad p' = 1,$$

откуда

$$v(G') = S'_1 - S'_0 + p' = (S_1 - 1) - S_0 + 1 = v(G) - 1 = -1,$$

что противоречит следствию из теоремы 1.1.

Определение 1.2. Граф G называется деревом, если он обладает одним (и, следовательно, всеми остальными) из свойств 1) — 4).

Иногда при определении дерева исключается случай $S_0 = 1$. По определению 1.2 такой граф является деревом; в ряде случаев такое определение оказывается более удобным.

Заметим, что если граф G состоит из двух несвязных графов G' и G'' , то справедливо соотношение

$$v(G) = v(G') + v(G'').$$

Докажем теперь характеристическое свойство $\nu(G)$.

Теорема 1.3. *Равенство $\nu(G) = 0$ эквивалентно отсутствию циклов в графе G .*

Доказательство. Пусть граф G состоит из несвязных компонент $G_1, \dots, G_p$. Тогда имеем равенство

$$\nu(G) = \nu(G_1) + \dots + \nu(G_p);$$

кроме того,

$$\nu(G_1) \geq 0, \dots, \nu(G_p) \geq 0; \nu(G) = 0.$$

Из этих соотношений следует равенство

$$\nu(G_i) = 0 \quad (i = 1, 2, \dots, p).$$

Поскольку каждый из графов $G_1, \dots, G_p$ связан, можно применить теорему 1.2, из которой следует, что в этих графах нет циклов. В G нет циклов, так как $G_1, \dots, G_p$ — несвязные компоненты этого графа.

По аналогии с понятием дерева граф без циклов (возможно, несвязный) иногда называется *лесом*.

С помощью понятия дерева можно сформулировать необходимое и достаточное условие связности графа, соответствующее теореме.

2. Условие связности графа. Приведем теорему, дающую необходимое и достаточное условие связности графа.

Теорема 1.4. *(о частичном дереве). Граф G связан тогда и только тогда, когда он содержит некоторое дерево в качестве суграфа.*

Доказательство. Очевидно, что если граф содержит частичное дерево, то он связан.

Докажем обратное. Пусть граф G связан. Тогда, если $\nu(G) = 0$, то G есть дерево (см. определение 1.2). Если $\nu(G) \neq 0$, то по теореме 1.3 в графе G есть хотя бы один цикл. Удалив одно из ребер этого цикла, мы не нарушим связность графа G и уменьшим его цикломатическое число на единицу. Через конечное число таких шагов получим связный суграф с $\nu(G) = 0$, т. е. дерево. ■

З а м е ч а н и е. Суграф графа G , являющийся деревом, называют *частичным деревом графа G* .

§ 2. БАЗА НЕЗАВИСИМЫХ ЦИКЛОВ

1. 1-цепи и 0-цепи графа. Как уже отмечалось, цикломатическое число $\nu(G)$ графа G связано с количеством циклов в графе G , а именно, оно равно количеству независимых

циклов в G . Точному определению этого понятия и доказательству утверждения и посвящен настоящий параграф.

Определение 2.1. Свободная абелева группа, порожденная множеством всех неориентированных ребер графа $G = (X, \Gamma)$, обозначается через $C_1(G)$ и называется группой одномерных цепей графа G . Свободная абелева группа, порожденная множеством вершин графа G , обозначается через $C_0(G)$ и называется группой нульмерных цепей графа G .

В обозначениях § 1 гл. III это определение можно записать так:

$$C_1(G) = g(\bar{\Gamma}), \quad C_0(G) = g(X),$$

где $\bar{\Gamma}$ — множество неориентированных ребер графа G .

Для дальнейших построений нам потребуется ориентировать все ребра графа G (как уже было сказано в начале этой главы, граф G предполагается полностью неориентированным). Поставим в соответствие каждому неориентированному ребру $\{(x_i, x_j), (x_j, x_i)\}$ один из элементов (x_i, x_j) или (x_j, x_i) из $X \times X$ произвольным, но фиксированным образом. При этом мы не изменяем самого графа G . В дальнейшем, на ребрах этого графа могут возникать и другие ориентации. Таким образом, ориентация ребер графа есть отображение

$$\varphi: \bar{\Gamma} \rightarrow X \times X,$$

где $\bar{\Gamma}$ — множество неориентированных ребер G , т. е. элементов вида $\{(x_i, x_j), (x_j, x_i)\}$. Отображение φ должно удовлетворять условию

$$\varphi(u) \in u \text{ для любого } u \in \bar{\Gamma}.$$

Выбранную ориентацию называем *канонической*.

В дальнейшем будет показано, что данные определения не зависят от произвола в выборе ориентации.

Элементы группы $C_1(G)$ называем 1-цепями. Выясним связь между 1-цепями и цепями в смысле определения 2.5 гл. I. Напомним, что цепью называется последовательность ребер $\{u_1, \dots, u_m\}$ такая, что конец u_{i-1} совпадает с началом u_i ($i = 2, 3, \dots, m$).

Задание цепи определяет на ребрах этой цепи ориентацию, которую мы назовем *индуцированной*. Эта ориентация строится следующим образом. Пусть задана цепь $\{u_1, \dots, u_m\}$. По определению цепи существует путь $\{u'_1, \dots, u'_m\}$ такой, что $u'_i \in u_i$. Определим $\varphi_1(u_i) = u'_i$. Отметим, что

индуцированная ориентация φ_1 , вообще говоря, не совпадает с канонической ориентацией φ .

Зададим отображение ψ цепей в 1-цепи следующим образом: если $\{u_1, \dots, u_m\}$ — некоторая цепь, то определим

$$\psi(\{u_1, \dots, u_m\}) = \pm u_1 \pm \dots \pm u_m,$$

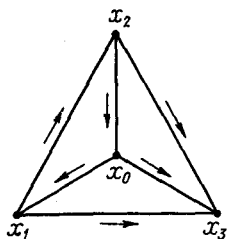


Рис. 21

где знак «+» выбирается для тех i , для которых индуцированная ориентация совпадает с канонической, а знак «—» для всех других i .

Вообще говоря, неверно, что каждая 1-цепь есть образ некоторой цепи при отображении ψ . Однако справедлива следующая лемма.

Лемма 2.1. *Каждая 1-цепь c может быть представлена в виде $c = \sum \psi(c'_i)$, где c'_i — цепи в G .*

Доказательство. Достаточно рассмотреть цепи c'_i , состоящие из одного ребра каждая. ■

Пример 2.1. Пусть G — граф, изображенный на рис. 21. Цепь

$$\{u_1, u_2, u_3\},$$

где

$$u_1 = \{(x_0, x_1), (x_1, x_0)\}, \quad u_2 = \{(x_1, x_3), (x_3, x_1)\}, \quad u_3 = \{(x_3, x_2), (x_2, x_3)\}$$

индуцирует на ребрах, входящих в нее, ориентацию следующим образом:

$$u_1 = \{(x_0, x_1), (x_1, x_0)\} \rightarrow (x_0, x_1),$$

$$u_2 = \{(x_1, x_3), (x_3, x_1)\} \rightarrow (x_1, x_3),$$

$$u_3 = \{(x_3, x_2), (x_2, x_3)\} \rightarrow (x_3, x_2).$$

Ориентация первых двух ребер этой цепи совпадает с канонической, а ориентация последнего не совпадает, поэтому этой цепи отвечает 1-цепь

$$\psi(\{u_1, u_2, u_3\}) = u_1 + u_2 - u_3.$$

С другой стороны, как легко видеть, 1-цепи $u_1 + u_3$ не отвечает никакая цепь в графе G .

Пусть u_1, u_2, u_3 определены, как и выше, а $u_4 = \{(x_1, x_2), (x_2, x_1)\}$. Последовательность

$$\{u_1, u_2, u_3, u_4, u_1\}$$

есть цепь. При этом на ребре u_1 возникают две индуцированные ориентации (так как оно дважды входит в цепь). Этой цепи соответствует 1-цепь

$$\psi(\{u_1, u_2, u_3, u_4, u_1\}) = u_1 + u_2 - u_3 - u_4 - u_1 = u_2 - u_3 - u_4,$$

причем коэффициент при u_1 равен нулю. С другой стороны, если обозначить $u_5 = \{(x_2, x_0), (x_0, x_2)\}$, то цепи

$$\{u_1, u_4, u_5, u_1\}$$

отвечает 1-цепь $\psi(\{u_1, u_4, u_5, u_1\}) = 2u_1 + u_4 + u_5$.

2. Граничный оператор графа. 1-циклы. Определим здесь граничный оператор, сопоставляющий каждой 1-цепи его «обобщенную границу».

Определение 2.2. Дифференциалом, или граничным оператором d , группы $C_1(G)$ называется гомоморфизм

$$d : C_1(G) \rightarrow C_0(G).$$

определенный следующим образом:

1) если $u = \{(x_i, x_j), (x_j, x_i)\}$ и $\varphi(u) = (x_i, x_j)$, то $du = x_j - x_i$

$$2) \text{ если } x = \sum_{i=1}^{|\bar{\Gamma}|} k_i x_i, \text{ то } dx = \sum_{i=1}^{|\bar{\Gamma}|} k_i du_i.$$

По самому определению отображение d есть гомоморфизм групп. Таким образом, для каждого ребра определяется его «граница» или сумма граничных точек с коэффициентами, отвечающими канонической ориентации соответствующего ребра. Например (см. рис. 21), можно написать:

$$d((x_1, x_2) + (x_2, x_3) - (x_1, x_3)) = (x_2 - x_1) + (x_3 - x_2) - (x_3 - x_1) = 0.$$

(В последней формуле допущена вольность в обозначениях. Следовало бы вместо (x_1, x_2) писать $\{(x_1, x_2), (x_2, x_1)\} = u_4$, а $\varphi(u_4) = (x_1, x_2)$. Но такая вольность допустима, если всегда использовать каноническую ориентацию.) Такую 1-цепь естественно назвать 1-циклом. С другой стороны, имеем

$$d((x_1, x_2) + (x_2, x_3) + (x_1, x_3)) = (x_2 - x_1) + (x_3 - x_2) + (x_3 - x_1) = 2x_3 - 2x_1;$$

такая 1-цепь не есть 1-цикл; 1-цикл, отвечающий цепи

$\{(x_1, x_2), (x_2, x_3), (x_3, x_1)\}$ (такой цикл является цепью в смысле определения 2.5 гл. I), есть

$$(x_1, x_2) + (x_2, x_3) - (x_3, x_1).$$

Определение 2.3. Элемент $x \in C_1(G)$ называется 1-циклом, если $dx = 0$. Подгруппа $Z_1 = \text{Ker } d \subset C_1(G)$ называется подгруппой 1-циклов графа G .

Лемма 2.2. Любой 1-цикл $c \neq 0$ можно представить в виде $c = \sum \psi(c_i)$, где c_i — циклы в G .

Доказательство. Пусть $c = \sum_{i=1}^{|\bar{\Gamma}|} a_i u_i = \sum_{i=1}^N \pm u'_i$

(в последовательности $u'_1, \dots, u'_N$ могут встречаться равные ребра). Поскольку c — цикл, то

$$d \sum_{i=1}^N \pm u'_i = \sum_{i=1}^N d(\pm u'_i) = 0. \quad (2.1)$$

Рассмотрим 1-цепь $\pm u'_i$, имеем $d(\pm u'_i) = x_{i_1} - x_{i_2}$. Из (2.1) очевидно, что существует номер j_1 такой, что $1 < j_1 \leq N$ и $d(\pm u'_{j_1}) = x_{i_2} - x_{i_3}$. Если $x_{i_2} \neq x_{i_1}$, то существует j_2 такой, что $d(\pm u'_{j_2}) = x_{i_3} - x_{i_4}$. Продолжим построение последовательностей x_{i_k} и u'_{j_k} до того момента, когда x_{i_k} окажется равным какому-либо $x_{i_{k_0}}$ для $k_0 \leq k-2$. В силу конечности множества X процесс обязательно закончится. Очевидно, полагая для удобства $j_0 = 1$, цепь

$$\{u'_{j_{k_0-1}}, u'_{j_{k_0}}, \dots, u'_{j_{k-2}}\}$$

есть цикл. Нетрудно проверить, что

$$\psi(\{u'_{j_{k_0-1}}, u'_{j_{k_0}}, \dots, u'_{j_{k-2}}\}) = \pm u'_{j_{k_0-1}} \pm \dots \pm u'_{j_{k-2}},$$

причем знаки совпадают со знаками в разложении c . Применяя тот же процесс к $c - \psi(\{u'_{j_{k_0-1}}, \dots, u'_{j_{k-2}}\})$, в конце концов приходим к утверждению леммы. ■

З а м е ч а н и е. В доказательстве не учтена возможность $du'_i = 0$. Но это возможно только, когда u_i — петля. В этом случае u_i как 1-цепь равна $\psi\{u_i\}$. Исключив сначала все такие слагаемые, приходим к утверждению леммы 2.2 и в общем случае.

Докажем, что система

$$x^{(0)}, x^{(1)}, \dots, x^{(n)} \quad (2.3)$$

независима. Пусть существуют $k_0, k_1, \dots, k_n$ не все равные нулю, такие, что

$$k_0 x^{(0)} + k_1 x^{(1)} + \dots + k_n x^{(n)} = 0.$$

Если расписать все $x^{(i)}$ по ребрам графа G , то ясно, что в этой сумме коэффициенты при всех ребрах должны обращаться в ноль. Однако 1-циклы $x^{(1)}, \dots, x^{(n)} \in Z_1(G')$ и, поэтому они не содержат u_0 , а $x^{(0)}$ содержит u_0 . Отсюда следует, что $k_0 = 0$, но при этом один из $k_1, \dots, k_n$ не равен нулю и $x^{(1)}, \dots, x^{(n)}$ зависимы, что невозможно.

Покажем теперь, что система (2.3) полна. Пусть $x \in \in Z_1(G)$.

Тогда

$$x = k_0 u_0 + \dots + k_{S_i} u_{S_i} \text{ и } x \pm k_0 x^{(0)} \in Z_1(G).$$

При правильном выборе знака имеем

$$x \pm k_0 x^{(0)} \in Z_1(G').$$

В силу предположения индукции существуют k'_i такие, что

$$x \pm k_0 x^{(0)} = k'_1 x^{(1)} + \dots + k'_n x^{(n)},$$

поэтому

$$x = \mp k_0 x^{(0)} + k'_1 x^{(1)} + \dots + k'_n x^{(n)},$$

т. е. система (2.3) полна. ■

Сформулируем и докажем теорему, обосновывающую независимость построений от выбора канонической ориентации ребер графа.

Теорема 2.2. Пусть $C_1(G)$ и $C'_1(G)$ — группы, отвечающие графу G с различным выбором канонической ориентации. Тогда существует такой изоморфизм $f: C_1(G) \rightarrow C'_1(G)$, что диаграмма

$$\begin{array}{ccc} C_1(G) & \xrightarrow{f} & C'_1(G) \\ \downarrow d & & \downarrow d' \\ & C_0(G) & \end{array}$$

коммутативна, т. е. $d = d' \circ f$ (и поэтому $d' = d \circ f^{-1}$).

Доказательство. Зададим $f(u_i) = \pm u_i$ в зависимости от того, совпадают или не совпадают ориентации u_i при переориентации ребер. На остальных элементах определим f так, чтобы он был гомоморфизмом, т. е.

$$f\left(\sum_{i=1}^{|\bar{\Gamma}|} k_i u_i\right) = \sum_{i=1}^{|\bar{\Gamma}|} k_i f(u_i).$$

Очевидно, что f — изоморфизм. Остальные утверждения теоремы проверяются тривиально (проверьте все утверждения теоремы самостоятельно). ■

Следствие. В предположениях теоремы f устанавливает изоморфизм групп $Z_1(G)$ и $Z'_1(G)$.

Доказательство. Пусть f — ограничение f на $Z_1(G)$, $\bar{f}^{-1}$ — ограничение f^{-1} на $Z'_1(G)$. Нужно доказать, что

$$\begin{aligned} \bar{f}(Z_1(G)) &= Z'_1(G), \quad \bar{f}^{-1}(Z'_1(G)) = Z_1(G) \text{ и } \bar{f}_0 \bar{f}^{-1} = \\ &= \bar{f}^{-1} \circ \bar{f} = 1, \end{aligned}$$

где 1 — тождественный гомоморфизм. Проверим эти условия.

1) Пусть $x \in Z(G)$, т. е. $dx = 0$, тогда

$$d'(f(x)) = d' \circ f(x) = dx = 0,$$

т. е. $d'(f(x)) = 0$ и $f(x) \in Z'_1(G)$. Отсюда

$$f(Z_1(G)) \subset Z'_1(G).$$

2) Пусть $x' \in Z'_1(G)$, тогда $d'x' = 0$ и $d(f^{-1}(x')) = d \circ f^{-1}(x') = 0$. Отсюда, как и в первом случае, $f^{-1}(Z'_1(G)) \subset Z_1(G)$.

3) Докажем, что $f(Z_1(G)) = Z'_1(G)$. Пусть $x' \in Z'_1(G)$, тогда $\kappa = f^{-1}(x') \in Z_1(G)$ (см. 2) и $f(\kappa) = x'$. Точно так же $f^{-1}(Z'_1(G)) = Z_1(G)$.

4. Последнее утверждение теоремы очевидно, так как на любом $x \in Z_1(G)$ $\bar{f}(x) = f(x) \in Z'_1(G)$ и поэтому $\bar{f}^{-1} \circ \bar{f}(x) = \bar{f}^{-1}(f(x)) = f^{-1}(f(x)) = x$. Аналогично, если $x' \in Z'_1(G)$, то $\bar{f} \circ \bar{f}^{-1}(x') = x'$. ■

С помощью гомоморфизма d можно охарактеризовать также и количество связных компонент графа G .

Теорема 2.3. Количество элементов любой базы группы $C_0(G)/\text{Im } d$ равно количеству связных компонент графа G .

Точнее, группа $C_0(G)/\text{Im } d$ изоморфна группе

$$\mathbb{Z} \oplus \dots \oplus \mathbb{Z}, \quad (2.4)$$

где $\mathbb{Z}$ — группа целых чисел, а количество прямых слагаемых равно p — числу связных компонент графа G .

Доказательство. Очевидно, из второго утверждения следует первое. Докажем второе утверждение. Выделим в каждой связной компоненте G_j графа G по одной (произвольной) вершине x_{i_j} ($j = 1, \dots, p$). Рассмотрим группу $g(\{x_{i_1}, \dots, x_{i_p}\})$, очевидно, изоморфную группе (2.4). Пусть $\{kx_{i_1}\}$ — класс эквивалентности элемента kx_{i_1} по $\text{Im } d$. Если $x_{i'}$ принадлежит той же компоненте, что и x_{i_1} , то существует цепь $\{u_0, u_1, \dots, u_m\}$ такая, что конец u_i совпадает с началом u_{i+1} , начало u_0 есть x_{i_1} , а конец u_m есть $x_{i'}$. Тогда, выбирая знаки соответствующим образом, видим, что

$$d(\pm ku_0 \pm ku_1 \pm \dots \pm ku_m) = kx_{i'} - kx_{i_1} \in \text{Im } d \text{ и } kx_{i'} \sim kx_{i_1}.$$

С другой стороны, эквивалентность $kx_{i'} \sim kx_{i_1}$ невозможна, если $x_{i'}$ принадлежит компоненте связности, не содержащей x_{i_1} и $k'x_{i'} \sim kx_{i_1}$ невозможно при любом $x_{i'}$ и $k' \neq k$ (докажите это!). Таким образом, в каждом классе эквивалентности по $\text{Im } d$ есть один и только один элемент вида $\sum_{j=1}^p k_j x_{i_j}$. Тем самым устанавливается соответствие между классами эквивалентности и элементами группы $g(\{x_{i_1}, \dots, x_{i_p}\})$. Очевидно, что это соответствие есть изоморфизм.

4. Алгоритмы построения базы независимых циклов. Приведем теперь алгоритм отыскания базы независимых 1-циклов для произвольного графа G . Сначала рассмотрим алгоритм нахождения частичного дерева графа G в том случае, когда G связен. Пусть ребра графа перенумерованы произвольным образом.

А л г о р и т м 2.1

1°. Выбрать первое ребро графа G и занести его в список.

2°. Выбрать следующее по номеру ребро графа.

3°. Проверить, образует ли выбранное ребро цикл

с ранее занесенными в список. Если да, перейти к 4°. Если нет — к 5°.

4°. Проверить, что номер ребра меньше $|\bar{G}|$. Если да, перейти к 2°, если нет — к 6°.

5°. Занести данное ребро в список. Перейти к 4°.

6°. Закончить работу.

Теорема 2.4. *Ребра, занесенные в список к концу работы алгоритма 1, образуют частичное дерево графа G .*

Доказательство. Пусть G' — граф, построенный из ребер списка и содержащий все вершины G . Очевидно, что он не содержит циклов. Докажем, что он связан. Пусть, напротив, граф G' не связан. Тогда в G' существует, по крайней мере, две компоненты связности. В силу связности G существует цепь в G , соединяющая эти две компоненты. При добавлении к G' любого ребра этой цепи не образуется циклов, что противоречит тому, что все ребра были проверены согласно с пунктом 3°, так как данное ребро в соответствии с 5° должно было быть занесено в список.

Теорема 2.5. *Пусть G — связный граф, G' — его частичное дерево и $u_1, \dots, u_m$ — ребра G , не входящие в G' . Тогда при добавлении к G' каждого из ребер $u_1, \dots, u_m$ в графе образуется ровно один простой цикл. Соответствующие 1-циклы образуют базис в G .*

Доказательство. Добавим к G' одно из ребер $u_1, \dots, u_m$, например u_1 . Тогда в графе образуется, по крайней мере, один цикл. Так как для полученного графа $\bar{G}$ его цикломатическое число $\nu(\bar{G}) = 1$, то группа $Z_1(G)$ имеет базу, состоящую из одного элемента, поэтому в графе $\bar{G}$ имеется ровно один простой цикл. В силу того что каждый из циклов, полученных таким образом, содержит одно из ребер $u_1, \dots, u_m$ и не содержит остальных, эти циклы независимы. Подсчитаем число m . Для G' выполнено соотношение

$$\nu(G') = S'_1 - S'_0 + 1 = 0,$$

так как G' — дерево, отсюда $S'_1 = S_0 - 1$. Не использованных в графе G' ребер в G содержится $m = S_1 - S'_1 = S_1 - S_0 + 1 = \nu(G)$. Таким образом, система независимых циклов содержит $\nu(G)$ элементов. Если она не является базой, то можно дополнить ее до некоторой базы, количество циклов в которой больше $\nu(G)$, что противоречит теореме 2.1.

Теорема 2.6. *Пусть G — граф, $G_1, \dots, G_p$ — его связные*

компоненты. Тогда для группы $Z_1(G)$ графа G имеем равенство

$$Z_1(G) = Z_1(G_1) \oplus \cdots \oplus Z_1(G_p)$$

(оно понимается в смысле изоморфизма групп).

Доказательство. Пусть $x \in Z_1(G)$. Разложим элемент $x = x^{(1)} + \cdots + x^{(p)}$, где в каждом $x^{(j)}$ содержатся ребра из G_j и только они. Такое разложение единственно. Кроме того, выполнено равенство

$$dx^{(1)} + \cdots + dx^{(p)} = 0,$$

причем $dx^{(j)}$ содержит вершины только из G_j . Отсюда $dx^{(j)} = 0$ ($j = 1, 2, \dots, p$), т. е. $x^{(j)} \in Z_1(G_j)$ ($j = 1, 2, \dots, p$). Определено отображение

$$Z_1(G) \rightarrow Z_1(G_1) \oplus \cdots \oplus Z_1(G_p) : x \rightarrow (x^{(1)}, \dots, x^{(p)}),$$

обратное к которому строится тривиально, а именно:

$$(x^{(1)}, \dots, x^{(p)}) \rightarrow x^{(1)} + \cdots + x^{(p)}.$$

Очевидно, что эти отображения есть гомоморфизмы.

Таким образом, можно выписать алгоритм нахождения базы независимых циклов для произвольного графа G .

А л г о р и т м 2.2

1°. Выделить связные компоненты графа G .

2°. В каждой компоненте построить базу независимых циклов, как в теореме 2.4, пользуясь алгоритмом 2.1.

3°. Объединить полученные таким образом базы.

В остальной части этого параграфа используется понятие плоской реализации графа. Точное определение этого понятия дано в следующей главе. При первом чтении можно ограничиться наглядным представлением о плоском графе.

Сформулируем алгоритм нахождения базы независимых циклов для графа G , обладающего плоской реализацией.

Теорема 2.7. Если связный граф G допускает плоскую реализацию, то границы областей, на которые граф G разбивает плоскость*, образуют базу независимых циклов**.

* Неограниченная область при этом не учитывается.

** Здесь и далее не проводится различие между циклами и соответствующими им 1-циклами.

Доказательство. Как будет доказано ниже, для плоского графа выполняется равенство

$$S_0 - S_1 + S_2 = 2, \text{ откуда } \nu(G) = S_1 - S_0 + p = \\ = S_2 - 2 + 1 = S_2 - 1,$$

и достаточно доказать лишь независимость циклов базы. Доказательство проведем индукцией по S_2 .

Базис индукции. Если $S_2 = 2$, то утверждение очевидно, так как база содержит всего один цикл.

Индукционный шаг. Пусть теорема справедлива для всех графов G' с $S_2 = n$. Пусть G — связный граф с $S_2 = n + 1$. Возьмем одно из ребер, лежащих на границе неограниченной области и не являющееся тупиком или перешейком. После удаления его получаем граф G' с $S_2 = n$. Применяя предположение индукции и замечая, что цикл, содержащий выброшенное ребро, не зависит от остальных, завершаем индукционный шаг. ■

Замечание. В силу теоремы 2.6 ограничение на связность графа несущественно.

Таким образом, определено некоторое обобщение понятия цикла, поскольку 1-циклом мы называем и то, что наглядно следовало бы называть конечным набором циклов. Преимущества этой терминологии будут видны ниже.

Проиллюстрируем это понятие на примерах (рис. 22):

$$x = (x_0, x_2) - (x_1, x_2) - (x_0, x_1)$$

есть 1-цикл, так как

$$dx = d((x_0, x_2) - (x_1, x_2) - (x_0, x_1)) = 0;$$

$$x = (x_0, x_2) - (x_1, x_2) - (x_0, x_1) + (x_6, x_7) + (x_7, x_8) + (x_8, x_6)$$

есть 1-цикл, так как $dx = 0$;

$x = (x_0, x_2) + (x_1, x_2) + (x_0, x_1) + (x_6, x_7)$
не есть 1-цикл, так как

$$dx = x_2 - x_0 + x_2 - x_1 + x_1 - x_0 + x_7 - x_6 = 2x_2 - 2x_0 + x_6 - x_7 \neq 0.$$

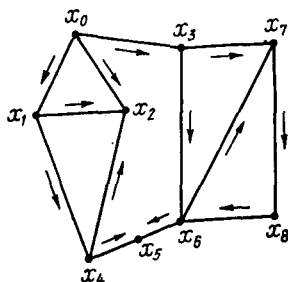


Рис. 22

В этом рассмотрении выясняется группы $C_0(G)$.

3. Независимость и полнота системы 1-циклов. Понятие независимости будем формулировать в терминах 1-циклов. При этом независимость обычных циклов будет пониматься как независимость соответствующих им 1-циклов.

Определение 2.4. В графе G 1-циклы $x^{(1)}, \dots, x^{(m)}$ называются зависимыми, если существуют такие целые числа $k_1, \dots, k_m$ не все равные нулю, что

$$k_1 x^{(1)} + \dots + k_m x^{(m)} = 0. \quad (2.2)$$

Если же из (2.2) следует, что $k_1 = \dots = k_m = 0$, то эти 1-циклы называются независимыми. Если $x^{(1)}, \dots, x^{(m)}$ таковы, что для каждого $x \in Z_1$ существует разложение

$$x = k_1 x^{(1)} + \dots + k_m x^{(m)},$$

то система $x^{(1)}, \dots, x^{(m)}$ называется *полной*. Любая полная система независимых 1-циклов называется *базой независимых 1-циклов*.

Интерпретация зависимости 1-циклов очевидна. Например, 1-цикл

$$x = (x_0, x_2) - (x_4, x_2) - (x_1, x_4) - (x_0, x_1)$$

зависит от 1-циклов

$$x^{(1)} = (x_0, x_2) - (x_0, x_1) - (x_1, x_2),$$

$$x^{(2)} = (x_1, x_2) - (x_4, x_2) - (x_1, x_4),$$

так как $x = x^{(1)} + x^{(2)}$.

Лемма 2.3. Если в графе G нет циклов, то группа $Z_1\{G\}$ тривиальна.

Доказательство. Если $c \in Z_1(G)$ и $c \neq 0$, то по лемме 2.2

$$c = \sum \psi(c_i'),$$

где c_i' — циклы в G . Поскольку в G нет циклов, то $c = 0$.

Теорема 2.1. Число элементов любой базы независимых 1-циклов графа G равно $\nu(G)$.

Доказательство. Воспользуемся методом индукции по $\nu(G)$.

Базис индукции. Если $\nu(G) = 0$, то по теореме 1.3 граф G не содержит циклов, т. е. $\text{Ker } d = 0$.

Индукционный шаг. Пусть дан граф G с $\nu(G) = n + 1$. Пусть теорема доказана для любого графа G с $\nu(G) = n$. По теореме 1.3 в графе G имеется цикл. Удалим одно из ребер этого цикла. Получим граф G' , причем из доказательства теоремы 1.1 очевидно, что $\nu(G') = n$. Применяя предположение индукции, выделим базу независимых 1-циклов

$$x^{(1)}, \dots, x^{(n)} \in Z_1(G').$$

Заметим, что группу $C_1(G')$ можно считать подгруппой $C_1(G)$, $C_0(G')$ — подгруппой $C_0(G)$ и $d(C_1(G')) \subset C_0(G')$. Отсюда ясно, что $Z_1(G')$ есть подгруппа в $Z_1(G)$.

Пусть цикл, из которого мы удалили ребро, есть $\{u_0, u_1, \dots, u_m\}$, причем, ребро u_0 было удалено. Пусть $x^{(0)}$ — элемент $C_1(G)$, отвечающий этому циклу, т. е.

$$x^{(0)} = \psi(\{u_0, \dots, u_m\}) = \pm u_0 \pm u_1 \pm \dots \pm u_m;$$

здесь знаки выбраны, как и ранее; тогда $dx^{(0)} = 0$, или $x^{(0)} \in Z_1(G)$.

ТОПОЛОГИЧЕСКИЕ ИНВАРИАНТЫ
ДВУМЕРНЫХ ПОВЕРХНОСТЕЙ

Теория графов может быть применена для вычисления некоторых инвариантов двумерных поверхностей — эйлеровой характеристики и групп гомологий.

Поясним понятие инварианта. Назовем две поверхности *эквивалентными*, если одну можно перевести в другую с помощью произвольной деформации без разрывов и склеиваний. Например, с этой точки зрения сфера эквивалентна поверхности куба, поверхность тора эквивалентна «сфере с одной ручкой», которая строится следующим образом. Вырезаются из сферы два непересекающихся маленьких диска и к образовавшимся отверстиям приклеивается цилиндр так, как это показано на рис. 23. Пусть задана функция, ставящая в соответствие каждой поверхности некоторое число. Такая функция называется инвариантом в том случае, если числа, сопоставляемые любым двум эквивалентным поверхностям, одинаковы.

Определения эйлеровой характеристики и групп гомологий даются здесь на основе теории графов. Другие (более общие) определения этих инвариантов можно найти в учебниках алгебраической топологии.

Отметим еще, что в дальнейшем не дается строгих определений топологических понятий, таких, как деформация или многообразие, а приводится их наглядное объяснение.

В этой главе все графы предполагаются полностью неориентированными и без кратных дуг.

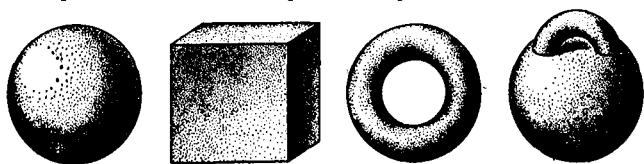


Рис. 23

§ 1. ЭЙЛЕРОВА ХАРАКТЕРИСТИКА

1. Плоские графы. Дадим определение плоского графа.

Определение 1.1. Пусть задан набор $\{l_1, l_2, \dots, l_n\}$ отрезков гладких кривых на плоскости, причем

1) отрезки l_i и l_j не имеют общих точек, кроме, быть может, концевых (в частности l_i не самопересекается, хотя может быть замкнутым);

2) граф G изоморфен $G[l_1, \dots, l_n]$, где граф $G[l_1, \dots, l_n]$ определен равенством

$$G[l_1, \dots, l_n] = (X', \bar{\Gamma}').$$

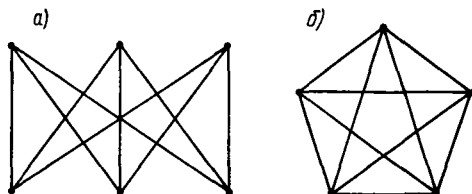


Рис. 24

X' — множество концевых точек отрезков $l_1, \dots, l_n$ на плоскости, $\bar{\Gamma}' = \{(x, y, i), (y, x, i) \mid x, y \text{ — концы } l_i\}$.

Набор $\{l_1, l_2, \dots, l_n\}$ называется плоской реализацией графа $G = (X, \bar{\Gamma})$.

Разумеется, не каждый граф допускает плоскую реализацию. Например, графы, изображенный на рис. 24, как будет доказано ниже, не допускают плоской реализации.

Отметим, что рис. 24, а не является плоской реализацией графа, так как не выполнено требование 1) определения 1.1.

Граф называется *плоским*, если он допускает плоскую реализацию.

Дадим теперь определение наложения двух плоских реализаций графов.

Определение 1.2. Пусть $\{l_i\}$ ($i = 1, 2, \dots, n$) — плоская реализация графа G , а $\{l'_j\}$ ($j = 1, 2, \dots, m$) — графа G' . Пусть $l_{i_1}, \dots, l_{i_{k_i}}$ — части, на которые кривые набора $\{l'_j\}$ делят отрезок l_i ; $l'_{j_1}, \dots, l'_{j_{k'_j}}$ определены аналогично. Набор

$$l_{ik}, l'_{jk'} \quad (r = 1, 2, \dots, n, k = 1, 2, \dots, k_i; j = 1, 2, \dots, m, k' = 1, 2, \dots, k'_j)$$

удовлетворяет требованию 1) определения 1.1 и является плоской реализацией графа

$$G'' [l_{ik}, l'_{jk'} \ (i = 1, \dots, n, \ k = 1, \dots, k_i, \ j = 1, \dots, m; \\ k' = 1, \dots, k'_j)],$$

построенного по этому набору. Граф G'' называется наложением графов G и G' (относительно их заданных плоских реализаций).

В этом параграфе считается, что каждый граф плоский и снабжен фиксированной плоской реализацией. Поэтому различие между графом и его плоской реализацией не проводится.

Определение 1.3. Эйлеровой характеристикой связного плоского графа G называется число

$$\chi(G) = S_0 - S_1 + S_2,$$

где S_0 — число вершин графа G , S_1 — число его ребер, S_2 — число областей, на которые G делит плоскость (включая и бесконечную область).

Сформулируем теперь основную теорему этого параграфа — теорему Эйлера.

Теорема 1.1. Для любого плоского связного графа G справедливо равенство $\chi(G) = 2$.

Доказательству этой теоремы предположим доказательство двух лемм.

Лемма 1.1. Пусть связный плоский граф G' получен из связного плоского графа G наложением на последний ребра l с несовпадающими концами. Тогда справедливо соотношение $\chi(G) = \chi(G')$.

Доказательство. Из условия очевидно, что пересечение графа G и ребра l не пусто (иначе G' не мог бы быть связным). Пусть a, b — концы отрезка l , а $x_1, \dots, x_n$ — точки пересечения l и G , рассмотрим несколько случаев.

Случай 1. Ни одна из точек $a, b, x_1, \dots, x_n$ не совпадает ни с одной вершиной графа G . Очевидно, что число вершин в графе G' равно $S'_0 = S_0 + n + 2$. Если на граф G наложить отрезок (a, x_1) , то число ребер увеличится на два, а число областей не изменится (рис. 25, а). При последовательном наложении отрезков $(x_1, x_2), \dots, (x_{n-1}, x_n)$ число ребер увеличивается на два, а число областей — на единицу (рис. 25, б). При добавлении же отрезка (x_n, b) число ребер увеличивается на единицу, а число областей

не изменяется (рис. 25, в). Это рассуждение показывает, что числа ребер и областей графа G' равны

$$S'_1 = S_1 + 2n + 1; \quad S'_2 = S_2 + n - 1.$$

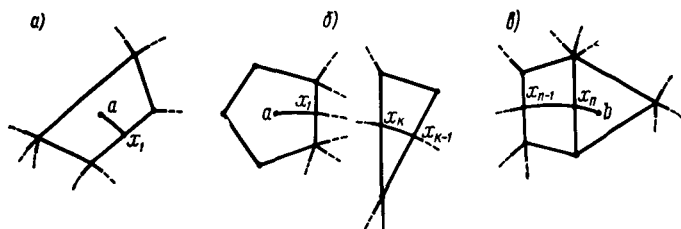


Рис. 25

Складывая эти равенства, получаем утверждение леммы.

С л у ч а й 2. Некоторые из точек $x_1, \dots, x_n$ являются вершинами графа G . Пусть число таких точек k ($k \leq n$). Тогда справедливо равенство

$$S'_0 = S_0 + n - k + 2.$$

Рассуждая аналогично случаю 1, заметим, что если конечная точка отрезка (x_i, x_{i+1}) есть вершина графа G , то S_1 увеличивается на единицу, если же конечная точка не совпадает с вершиной, то — на два. При этом для числа ребер графа G' получим выражение

$$S'_1 = S_1 + 2n + 1 - k$$

и, замечая, что S_2 меняется так же, как и в случае 1, — соотношние

$$S'_2 = S_2 + n - 1.$$

В результате сложения S'_0 , S'_1 и S'_2 получаем утверждение леммы.

При разборе случаев 1 и 2 предполагалось, что точки a и b не принадлежат графу G . Если, например, точка a принадлежит графу G , то доказательство можно свести к уже разобранным следующим образом.

Добавим к отрезку l малый кусок (a', a) так, чтобы он пересекался с G только в точке a (рис. 26). Нетрудно видеть, что эйлерова характеристика не изменилась, так как

$S'_0 = S'_0 + 1$, $S'_1 = S'_1 + 1$, $S'_2 = S'_2$. При этом задача сводится к уже разобранный. Подобным образом поступим и с концом b в том случае, если $b \in X$.

Легкое видоизменение доказательства позволяет получить тот же результат и при $a = b$. Детали доказательства предоставляются читателю.

Лемма 1.2. Пусть связный граф G' получен наложением графа G' на граф G , тогда $\chi(G') = \chi(G)$.

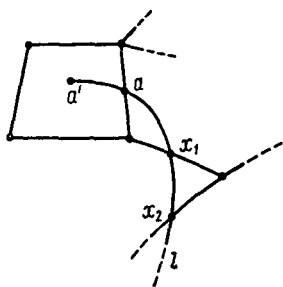


Рис. 26

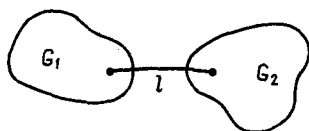


Рис. 27

Доказательство. Проведем доказательство методом индукции по числу ребер n графа G_1 . Базисом индукции является лемма 1.1. Проведем индукционный шаг. Различаем два случая.

С л у ч а й 1. В графе G_1 имеется, по крайней мере, два ребра, которые пересекаются с ребрами графа G . В этом случае, удаляя ребро графа G_1 , не являющееся перешейком, получим граф G'_1 с n ребрами такой, что наложение G'' графа G'_1 на граф G связно. Пользуясь индукционным предположением, получаем равенство $\chi(G'') = \chi(G)$. Добавляя выкинутое ребро по лемме 1.1, заключаем, что $\chi(G') = \chi(G)$.

С л у ч а й 2. Граф G_1 пересекается с графом G по одному ребру l . Если граф G_1 — не дерево, то в нем есть цикл. Удаляя ребро этого цикла, не совпадающее с l и рассуждая аналогично случаю 1, получим равенство $\chi(G') = \chi(G)$. Если же G_1 — дерево, то в нем имеется, по крайней мере, два тупика. Удаляя не совпадающий с l тупик, теми же рассуждениями получим, что эйлеровы характеристики графов G и G' совпадают. ■

Доказательство теоремы 1.1. Рассмотрим два плоских связных графа G_1 и G_2 (рис. 27). Пусть l — дуга, одна вер-

шина которой принадлежит G_1 , а другая G_2 . Пусть G' — наложение l на G_1 . По лемме 1.2 справедливо равенство $\chi(G_1) = \chi(G')$, поскольку G' связан. Кроме того, наложение G'' графа G' на граф G_2 также связано, а поэтому $\chi(G'') = \chi(G_1)$. Аналогично получим равенство $\chi(G_2) = \chi(G'')$ и поэтому эйлеровы характеристики графов G_1 и G_2 совпадают. Из произвольности графов G_1 и G_2 получаем, что любые два плоских связанных графа имеют одну и ту же эйлерову характеристику. Достаточно подсчитать ее поэтому для простейшего плоского связанного графа. В качестве такого графа выберем граф $(\{x\}, \emptyset)$. Для него $S_0 = 1$, $S_1 = 0$, $S_2 = 1$; отсюда $\chi(G) = 2$. ■

З а м е ч а н и я 1. Эта теорема дает основание называть число 2 эйлеровой характеристикой плоскости. Подробнее об этом см. в следующем параграфе.

2. Условие связности графа существенно для доказательства теоремы, так как уже для графа $(\{x_1, x_2\}, \emptyset)$ имеем равенство $S_0 = 2$, $S_1 = 0$, $S_2 = 1$ и $\chi(G) = 3$.

В качестве элементарного приложения теоремы Эйлера докажем, что графы, представленные на рис. 24, не допускают плоской реализации.

Теорема 1.2. Графы

$$G_1 = (\{x_1, x_2, x_3, y_1, y_2, y_3\}, \{\{x_i, y_j\}, (y_j, x_i)\}, \\ i, j = 1, 2, 3\})$$

и

$$G_2 = (\{x_1, x_2, x_3, x_4, x_5\}, = X, X \times X \setminus E),$$

где множество E состоит из элементов вида (x, x) , не допускают плоской реализации.

Д о к а з а т е л ь с т в о. Отметим, что в графе G_1 нет циклов длины 3, так как любое ребро ведет из группы вершин (x_1, x_2, x_3) в группу (y_1, y_2, y_3) . Любой цикл в G_1 имеет поэтому длину большую или равную четырем. Предположим теперь, что G_1 допускает плоскую реализацию, тогда для чисел S_0, S_1, S_2 получаем равенства $S_0 = 6$, $S_1 = 9$, а из $\chi(G_1) = 2$ следует $S_2 = 5$.

Пусть n_i ($i = 1, \dots, 5$) — число ребер, ограничивающих грань с номером i ; имеем $n_i \geq 4$ и

$$S_1 = \frac{1}{2} \sum_{i=1}^5 n_i \geq 10, \text{ т. е. } 9 \geq 10.$$

Полученное противоречие доказывает первую часть теоремы.

Допустим, что G_2 допускает плоскую реализацию. Тогда из теоремы 1.1 имеем равенства

$$S_0 = 5, S_1 = C_5^2 = 10, S_2 = 7.$$

Пусть n_i ($i = 1, \dots, 7$) имеют тот же смысл, что и раньше. Очевидно, $n_i \geq 3$ и поэтому выполнено неравенство

$$S_1 = \frac{1}{2} \sum_{i=1}^7 n_i \geq \frac{3 \cdot 7}{2} > 10, \quad 10 > 10. \blacksquare$$

Доказанная теорема допускает обращение. Именно, она представляет доказательство необходимости в следующем критерии, позволяющем решить вопрос о том, допускает ли граф плоскую реализацию.

Теорема 1.3. (Понтрягин — Куратовский). *Для того чтобы граф был плоским, необходимо и достаточно, чтобы он не содержал в качестве подграфов графы G_1 или G_2 из условия теоремы 1.2.*

Доказательство достаточности в этой теореме здесь не приводится. Отметим, что аналогичных критериев реализуемости графа уже для тора не найдено (определение графа, реализуемого на поверхности, см. ниже).

2. Эйлерова характеристика двумерных поверхностей. Как было показано в пункте 1, для любого плоского связного графа его эйлерова характеристика $\chi(G) = S_0 - S_1 + S_2$ равна двум. Число два можно назвать поэтому *эйлеровой характеристикой плоскости*. В этом параграфе обобщается понятие эйлеровой характеристики на произвольную двумерную поверхность без края. Как указывалось в замечании 2 к теореме 1.1, условие связности графа существенно в формулировке этой теоремы. Из доказательства теоремы видно, что это условие обеспечивает следующее свойство областей, на которые граф G делит плоскость.

А. Любая гладкая кривая без самопересечений, концы которой лежат на границе области, делит эту область на две части.

Из рис. 28 видно, что в случае произвольной поверхности в трехмерном пространстве это условие уже не обеспечивается только связностью графа. Например, граф, состоящий из вершины x_0 и петли l_1 (рис. 28, а), ограничивает

на торе область, показанную на рис. 28, б, и является связным. Петля l_2 , оба конца которой лежат на границе области, не делит, однако, эту область на две части. Таким образом, условие связности в этом более общем случае должно быть заменено на некоторое более сильное.

Существует несколько вариантов такой замены; остановимся на следующем.

Б. Любая область из числа тех, на которые граф G делит поверхность S , эквивалентна (в смысле, поясненном в начале данной главы) единичному кругу плоскости.

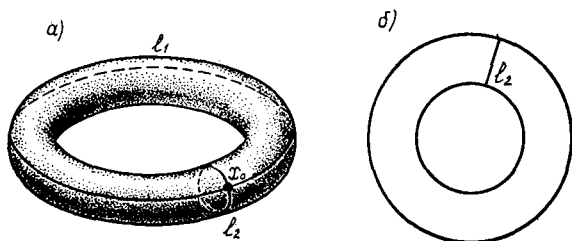


Рис. 28

Очевидно, это условие влечет за собой условие А (хотя строгое доказательство этого факта не просто). Перейдем теперь к точным формулировкам.

Определение 1.4. Реализацией графа $G = (X, \Gamma)$ на поверхности называется такой конечный набор отрезков гладких кривых $l_1, \dots, l_n$, заданных на S , что выполняются условия 1) и 2) определения 1.1.

Если у данного графа G существует реализация на поверхности S , то граф G будем называть реализуемым на S , или, короче, S -графом. Допуская вольность речи, в этой главе зачастую будем отождествлять граф G с его реализацией на поверхности S . Отметим еще тот факт, что один и тот же граф может быть реализован на различных поверхностях.

Введем теперь условие, о котором говорилось в начале этого параграфа.

Определение 1.5. Реализация графа G на поверхности S называется правильной, если выполнено условие Б этого параграфа.

Отметим, что понятие правильности относится не к самому графу G , а к его конкретной реализации на поверхнос-

ти S . Один и тот же граф может иметь на данной поверхности S и правильную, и неправильную реализации. Например, граф

$$G = (\{x_0, x_1, x_2\}, \{(x_0, x_0), (x_0, x_1), (x_0, x_2), (x_1, x_2)\})$$

имеет две реализации на торе (рис. 29, а, б), из которых первая является правильной, а вторая — неправильной. Допуская вольность речи, будем говорить о правильных S -графах.

Определение 1.6. Для любого правильного S -графа определим его эйлерову характеристику равенством $\chi(G) =$

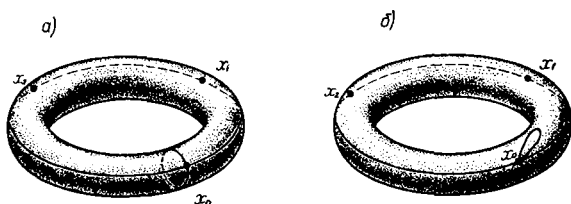


Рис. 29

$= S_0 - S_1 + S_2$, где S_0 — число вершин графа, S_1 — число его ребер, а S_2 — число областей, на которые граф делит S .

Сформулируем теперь теорему, аналогичную теореме 1.1 для S -графов.

Теорема 1.4. Для любых двух правильных S -графов G_1 и G_2 имеем равенство $\chi(G_1) = \chi(G_2)$.

Доказательство этой теоремы проводится дословно так же, как и доказательство теоремы 1.1, и поэтому предоставляется читателю.

Приведем формулировки лемм, аналогичные леммам 1.1 и 1.2. Сначала дадим определение наложения.

Определение 1.7. Пусть l_i ($i = 1, 2, \dots, n$) — S -реализация графа G , l'_j ($j = 1, 2, \dots, m$) — S -реализация графа G' . Пусть l_{ik} ($i = 1, 2, \dots, n; k = 1, 2, \dots, k_i$) части l_i на которые его делят отрезки l'_j , а $l'_{jk'}$ ($j = 1, 2, \dots, m; k' = 1, 2, \dots, k'_j$) определяется аналогично. Граф G'' , определяемый отрезками $l_{ik}, l'_{jk'}$ как своей S -реализацией, называется наложением G на G' (или G' на G).

Лемма 1.3. Пусть S -граф G' получен из правильного S -графа G наложением на последний ребра l . Тогда, если G' связан, то он правилен, и $\chi(G') = \chi(G)$.

Лемма 1.4. Пусть S -граф G' получен наложением правильного S -графа G_1 на правильный S -граф G . Тогда, если G' связан, то он правилен, и $\chi(G) = \chi(G')$.

Рассмотрим здесь только утверждения последних двух лемм, касающиеся правильности графа G' . По условию Б каждая область графа G эквивалентна единичному кругу; так как построение G' из G можно провести добавлением одного (или нескольких) новых ребер, пересекающихся с G , то доказательство этих утверждений эквивалентно следующему факту.

Любая гладкая несамопересекающаяся кривая с концами на границе единичного круга делит его на две области, каждая из которых эквивалентна единичному кругу. Доказательство этого утверждения, относящееся к топологии, проводить не будем, а примем его в качестве очевидного.

Дадим теперь определение эйлеровой характеристики поверхности S . Корректность такого определения обосновывается теоремой 1.4.

Определение 1.8. Эйлеровой характеристикой данной двумерной замкнутой ограниченной поверхности S без границы называется эйлерова характеристика любого правильного S -графа.

Обсудим теперь (не приводя, впрочем, точных доказательств) инвариантный характер эйлеровой характеристики S относительно эквивалентности, определенной в начале настоящей главы. Пусть поверхности S' и S'' эквивалентны. Это означает, что существует некоторая деформация без разрывов и склеиваний, переводящая S' в S'' . Пусть G — некоторый правильный S' -граф. В процессе деформации граф G остается правильным. Это следует из того, что каждая область из числа тех, на которые G делит S' , деформируется при этой деформации без разрывов и склеиваний и поэтому остается эквивалентной единичному кругу в плоскости. Кроме того, из условия отсутствия разрывов и склеиваний следует, что числа S_0, S_1, S_2 в процессе этой деформации не изменяются. Граф G перейдет в результате в правильный S'' -граф G' , причем выполняется соотношение

$$\chi(G) = \chi(G'),$$

так как $\chi(G)$ в процессе деформации постоянно. Эти рас-

суждения показывают, что эйлерова характеристика поверхности есть инвариант.

Вычислим теперь эйлерову характеристику так называемой «поверхности рода p » в трехмерном евклидовом пространстве $\mathbb{R}^3$. Поверхность рода p в $\mathbb{R}^3$ строится следующим образом. Рассмотрим в $\mathbb{R}^3$ сферу S^2 . Вырежем из S^2 два непересекающихся малых диска и приклеим к получившимся отверстиям цилиндр (рис. 30). Результатом явится «сфе-

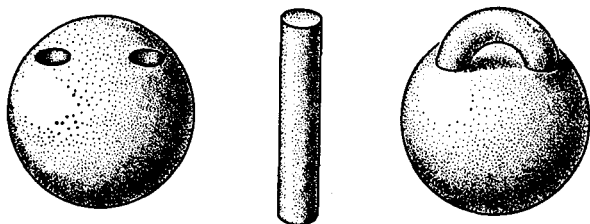


Рис. 30

ра с ручкой» или поверхность рода I . Поверхность рода p может быть описана такими же преобразованиями и представляет собой сферу с p ручками (поверхностью рода ноль по определению называют сферу S^2). Важность этих поверхностей в $\mathbb{R}^3$ обуславливается следующей теоремой, которая здесь приводится без доказательства.

Теорема 1.5. Любая замкнутая ограниченная поверхность без границы в $\mathbb{R}^3$ эквивалентна поверхности рода p при некотором p .

Теорема 1.6. Эйлерова характеристика поверхности рода p равен $2-2p$.

Доказательство. В силу теоремы 1.4 достаточно вычислить эйлерову характеристику некоторого S -графа на поверхности S рода p . В качестве такого графа возьмем граф G , строящийся следующим образом. Вершины G находятся на основаниях «ручек», по две на каждом основании (рис. 31). Опишем ребра графа G (рис. 32, на котором изображена одна из «ручек»). Каждое осно-

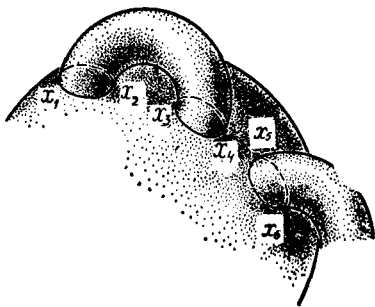


Рис. 31

вание ручки разделяется вершинами на две дуги ($l_1, l_2; l_3, l_4$ на рис. 32). Эти дуги есть ребра G . Кроме того, выбрав на каждом основании ручки по одной точке (x_i и x_{i+2} на рис.

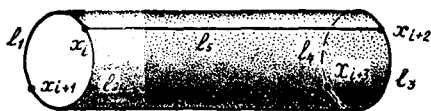


Рис. 32

32), соединим их кривой l_5 . Эту кривую также причислим к ребрам графа. Наконец, рассмотрев сферу с $2p$ «дырками», соединим все вершины графа G на сфере так, как показано на рис. 33.

Из построения G очевидно, что это — правильный S -граф. Вычисление эйлеровой характеристики $\chi(G)$ дается теперь прямым подсчетом:

$$S_0 = 4p, S_1 = 7p, S_2 = p + 2, \chi(G) = 4p - 7p + p + 2 = 2 - 2p. \blacksquare$$

Следствие. Две поверхности родов p_1, p_2 соответственно эквивалентны тогда и только тогда, когда $p_1 = p_2$.

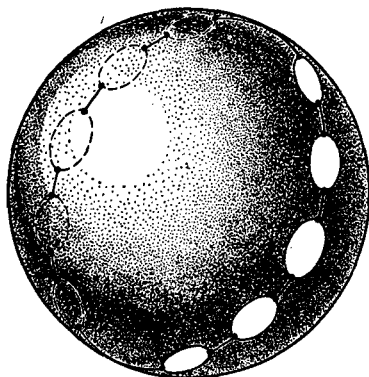


Рис. 33

Эквивалентность при $p_1 = p_2$ очевидна. Если две поверхности эквивалентны, то $\chi(S') = \chi(S'')$, так что эквивалентность поверхности рода p_1 , поверхности рода p_2 при $p_1 \neq p_2$ невозможна. Если учесть еще теорему 9.2, то можно утверждать, что две замкнутые ограниченные поверхности S' и S'' без границы в $\mathbb{R}^3$ эквивалентны тогда и только тогда, когда $\chi(S') = \chi(S'')$.

Отметим, что все требования последнего заме-

чания существенны.

3. Эйлерова характеристика многообразия размерности два. Рассмотрим понятие двумерного топологического многообразия, обобщающего понятие поверхности в $\mathbb{R}^3$.

Наглядно можно представлять себе многообразие как «поверхность», склеенную из некоторого числа треугольников. Говоря точнее, в качестве элементарных компонент для построения многообразия будем использовать части двумерных поверхностей, полученные деформацией треугольника. В силу этого понятие «многообразия» оказывается весьма широким. В частности все поверхности, рассмотренные выше, представляют собой многообразия. Например, разбиение сферы на треугольники можно получить, вписав в шар правильный тетраэдр и спроектировав его ребра из центра на поверхность сферы (рис. 34).

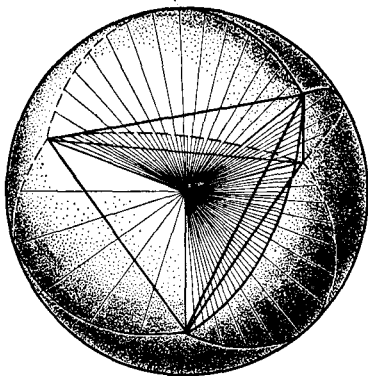


Рис. 34

В качестве первого шага в построении многообразий определим понятие алгебраического многообразия.

Определение 1.9. Алгебраическим многообразием размерности 2 назовем конечное множество $X = \{x_1, \dots, x_n\}$ с выделенными системами двухэлементных подмножеств $\mathcal{A}_1 = \{a_1, \dots, a_m\}$ ($a_i = \{x_{j_i}, x_{k_i}\}$) и трехэлементных подмножеств $\mathcal{A}_2 = \{b_1, \dots, b_l\}$ ($b_i = \{x_{r_i}, x_{s_i}, x_{t_i}\}$), причем выполнены условия:

- 1) каждый элемент множества X входит хотя бы в одно из множеств a_i ;
- 2) каждое подмножество a_i вложено точно в два множества из $\mathcal{A}_2$.

Элементы из $\mathcal{A}_1$ называются ребрами, элементы из $\mathcal{A}_2$ — гранями.

Наглядно элементы X есть вершины треугольников, $\mathcal{A}_1$ — стороны, а $\mathcal{A}_2$ — сами треугольники. Если сторона a_i принадлежит треугольникам b_j и b_k , то их нужно представлять себе склеенными по этой стороне.

Приведем примеры.

1. Рассмотрим проективную плоскость $\mathbb{RP}^2$ (рис. 35).

Множество X состоит из тринадцати элементов $\{x_1, \dots, x_{13}\}$. Не будем выписывать множества $\mathcal{A}_1$ и $\mathcal{A}_2$ —

читатель может проделать это самостоятельно. Отметим только, что из рисунка следует, что диаметрально противоположные точки квадрата должны быть отождествлены,

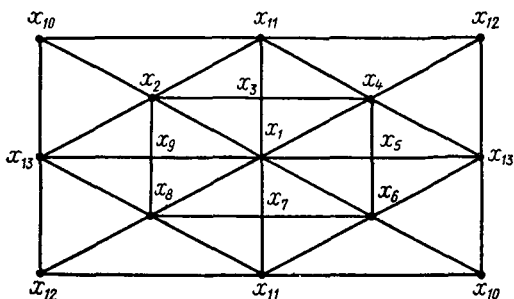


Рис. 35

иными словами, $\mathbb{R}P^2$ есть квадрат (рис. 36), у которого отождествлены стороны l_1 и l_2 ; l_3 и l_4 в направлениях, указанных стрелками. Отождествление одной из пар сторон превращает квадрат в лист Мебиуса. Второе отождествление без самопересечений в $\mathbb{R}^3$ уже невозможно, что оправдывает введение нового определения для таких объектов.

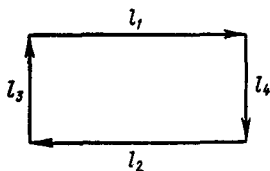


Рис. 36

В дальнейшем будем пользоваться рисунками типа рис. 36 для представления многообразий, не

выписывая явно множеств X , $\mathcal{A}_1$ и $\mathcal{A}_2$.

2. Рассмотрим многообразие, изображенное на рис. 37 (бутылка Кляйна).

а)



б)

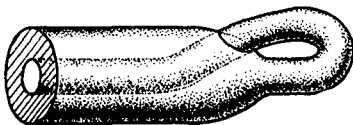


Рис. 37

Если отождествлять стороны квадрата в направлениях, показанных на рис. 37, а, то получим многообразие, называемое *бутылкой Кляйна*. Ее реализация в $\mathbb{R}^3$ (с самопересечениями) показана на рис. 37, б.

Задача 1. Начертить разбиение бутылки Кляйна на треугольники и выписать множества X , $\mathcal{A}_1$ и $\mathcal{A}_2$.

Определим теперь понятие многообразия. Рассмотрим m равносторонних непересекающихся треугольников на плоскости; $m = |\mathcal{A}_2|$. Поставим в соответствие каждому элементу из $\mathcal{A}_2$ один и только один из этих треугольников. Рассмотрим какой-либо из этих треугольников. Пусть ему отвечает элемент вида $\{x_{i_1}, x_{i_2}, x_{i_3}\}$ из $\mathcal{A}_2$.

Маркируем символами x_{i_1} , x_{i_2} и x_{i_3} вершины треугольника. Прделаем то же самое со всеми m треугольниками. Треугольники будем обозначать символами $b_1, \dots, b_m$ (в соответствии с отвечающими им элементами множества $\mathcal{A}_2$). отождествим теперь ребра треугольников, концы которых маркированы одинаково так, как это показано на рис. 38. Точками многообразия будем считать внутрен-

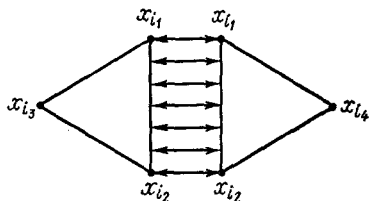


Рис. 38

ренние точки треугольников и пары точек, отвечающих друг другу при таком отождествлении. Покажем теперь, что каждая точка такого многообразия обладает окрестностью, устроенной так же, как часть плоскости. Окрестностью внутренних точек данного треугольника будем считать внутренность этого треугольника. Рассмотрим теперь точку, лежащую на стороне треугольника и не совпадающую с вершиной (рис. 39). Для наглядности изобразим эти стороны слившимися. Окрестностью отмеченной на рисунке точки считаем внутренность полученного при склейке четырехугольника. Наконец, рассмотрим вершину κ_{i_0} . Очевидно, что (сдеформировав, если нужно, треугольники) можно изобразить на плоскости совокупность треугольников, содержащих κ_{i_0} как вершину, как это сделано на рис. 40. На этом рисунке уже проведены необходимые отождествления. Внутренность изображенной на рисунке фигуры будем считать окрестностью κ_{i_0} . Сформулируем определение.

Определение 1.10. Многообразием S , отвечающим алгебраическому многообразию $(X, \mathcal{A}_1, \mathcal{A}_2)$, называется описанное выше множество точек. При этом последовательность точек $y_1, \dots, y_n \in S$ называется сходящейся, если начиная с некоторого номера N все они принадлежат одной из окрестностей вида, описанного выше, и в этой окрест-

ности последовательность $y_N, y_{N+1}, \dots$ сходится как последовательность точек на плоскости.

Это определение отвечает, естественно, интуитивному понятию склеивания треугольников.

Теперь можно обычным образом определить непрерывность отображения S в S' , где S и S' — многообразия.

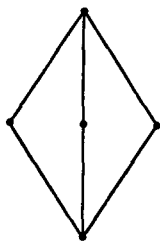


Рис. 39

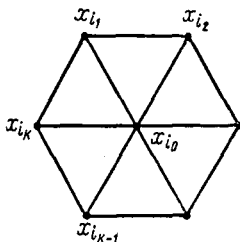


Рис. 40

Именно, отображение $f : S \rightarrow S'$ непрерывно тогда и только тогда, когда из сходимости последовательности $x_1, \dots, x_n, \dots \in S$ к x_0 следует сходимость последовательности $f(x_1), \dots, f(x_n), \dots \in S'$ к $f(x_0)$.

Наконец, дадим определение эквивалентных многообразий.

Определение 1.11. Два многообразия S и S' называются эквивалентными, если существует такое взаимно-однозначное отображение $f : S \rightarrow S'$, что f и f^{-1} непрерывны.

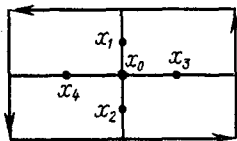


Рис. 41

Теперь не представляет трудности определить понятие реализации графа на данном многообразии, эйлерову характеристику многообразия и доказать теорему, аналогичную теореме 1.4. Предоставляем читателю провести соответствующие вычисления.

Вычислим эйлеровы характеристики $\mathbb{R}P^2$ и бутылки Кляйна. Для этого достаточно вычислить $\chi(G)$ для фиксированных правильных графов на этих многообразиях. Для проективной плоскости $\mathbb{R}P^2$ правильный $\mathbb{R}P^2$ -граф показан на рис. 41; это граф

$$G = (\{x_0, \dots, x_4\}, \{(x_0, x_1), (x_1, x_2), (x_2, x_0), (x_0, x_3), (x_3, x_4), (x_4, x_0)\}).$$

Этот граф отсекает область, эквивалентную единичному кругу (учтите склейку), и поэтому справедливы равенства $S_0 = 5$, $S_1 = 6$, $S_2 = 1$, $\chi(\mathbb{RP}^2) = 0$.

Задача 2. Вычислить эйлерову характеристику бутылки Кляйна.

Обсудим результаты вычислений. Как видим, эйлеровы характеристики $\mathbb{RP}^2$, бутылки Кляйна и тора (поверхность рода 1) одинаковы. Из этого, однако, нельзя вывести заключения о эквивалентности этих многообразий. В действительности же они неэквивалентны. Таким образом, в общем случае, если эйлеровы характеристики двух многообразий не равны, то можно утверждать, что они неэквивалентны. Но если эйлеровы характеристики равны, то многообразия могут оказаться и неэквивалентными.

§ 2. ГОМОЛОГИИ ДВУМЕРНЫХ ПОВЕРХНОСТЕЙ

1. Граничные гомоморфизмы и группа H_1 . В этом параграфе приведем рассуждения, обобщающие проведенные в § 2 гл. IV и позволяющие построить еще один инвариант плоских поверхностей.

Рассмотрим поверхность S и G — правильный S -граф на этой поверхности. Пусть группы $C_1(G)$, $C_0(G)$ определены так же, как и ранее, а $C_2(G)$ — свободная абелева группа, порожденная множеством областей, на которые G разбивает поверхность.

Предположим, что граф G не имеет тупиков (будем рассматривать лишь такие графы). Это не является ограничением, так как если из правильного S -графа удалить тупики, то результатом будет снова правильный S -граф.

Зададим ориентацию каждого ребра и каждой области. Под ориентацией области понимается упорядочение множества вершин, лежащих на ее границе так, чтобы каждая вершина следовала за вершиной, соединенной с ней ребром непосредственно. Для каждой области возможно ровно две такие ориентации. Например, область, изображенная на рис. 42, допускает ориентации $(x_0, x_1, x_2, x_3, x_4)$ и $(x_0, x_4, x_3, x_2, x_1)$. Будем говорить, что ориентация ребра согласована с ориентацией области s , если ребро u лежит на границе s и $u = (x, x')$, а в ориентации s вершина x' следует за x . Например, ориентация ребра (x_1, x_0) на рис. 42 не согласована с первой ориентацией и согласована со второй.

Теперь зададим гомоморфизмы

$$C_2(G) \xrightarrow{d_2} C_1(G) \xrightarrow{d_1} C_0(G). \quad (2.1)$$

Гомоморфизм d_1 задается так же, как и в гл. IV. Гомоморфизм d_2 зададим так:

$$d_2 s_i = \sum_{j=1}^m \pm u_j \text{ и } d_2 \left(\sum_{i=1}^{S_2} k_i s_i \right) = \sum_{i=1}^{S_2} k_i d_2 s_i,$$

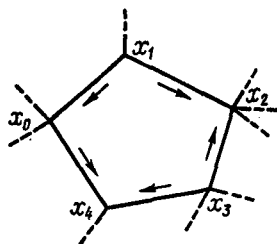


Рис. 42

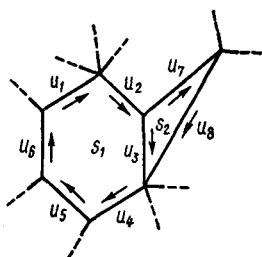


Рис. 43

где $u_1, \dots, u_m$ — все ребра, ограничивающие s_i ; знак «+» выбирается, если ориентация u_j согласована с ориентацией s_i , «—» — в противном случае.

Так же как и ранее доказывается, что все построения не зависят от ориентации.

Теорема 2.1. *Справедливо равенство $d_1 \circ d_2 = 0$.*

Доказательство. Очевидно, достаточно доказать равенство $d_1 \circ d_2(s_i) = 0$ ($i=1, \dots, S_2$). Пусть вершины s есть $x_0, \dots, x_m$ в порядке, задаваемом ориентацией. Тогда нетрудно видеть, что композиция $d_1 \circ d_2(s_i)$ равна нулю:

$$d_1 \circ d_2(s_i) = (x_1 - x_0) + (x_2 - x_1) + \dots + (x_m - x_{m-1}) + \\ + (x_0 - x_m) = 0. \blacksquare$$

Эта теорема означает, что выполнено включение $\text{Im } d_2 \subset \text{Ker } d_1$.

Определение 2.1. *Элементы из $\text{Im } d_2$ называются границами, элементы из $\text{Ker } d_1$ — циклами. Группа $\text{Ker } d_1 / \text{Im } d_2$ называется группой одномерных гомологий G и обозначается $H_1(G)$.*

Поясним смысл этого определения. По определению фактор-группы H_1 мы считаем два цикла $x^{(1)}, x^{(2)}$ эквива-

лентными, если существует такой элемент $s \in C_2(G)$, что $x^{(1)} - x^{(2)} = d_2 s$.

Рассмотрим в качестве примера граф, изображенный на рис. 43. Цикл

$$x^{(1)} = u_1 + u_2 + u_7 + u_8 + u_4 + u_5 + u_6$$

эквивалентен циклу

$$x^{(2)} = u_1 + u_2 + u_3 + u_4 + u_5 + u_6,$$

так как справедливо соотношение

$$x^{(1)} - x^{(2)} = u_7 + u_8 - u_3 = \pm ds_2$$

в зависимости от ориентации s_2 . Это отвечает тому факту, что поскольку 1-цепь $u_7 + u_8 - u_3$ является границей элемента s_2 , то цикл $x^{(1)}$ можно стянуть в цикл $x^{(2)}$, переводя $u_7 + u_8$ по s_2 в u_3 . На самом деле циклы $x^{(1)}$ и $x^{(2)}$ эквивалентны нулю, так как выполнены соотношения

$$x^{(1)} = \pm ds_1 \pm ds_2, \quad x^{(2)} = \pm ds_1.$$

2. Инвариантность группы H_1 . Докажем, что группа H_1 есть характеристика поверхности, а не графа.

Теорема 2.2. Пусть G_1, G_2 — два правильных S -графа. Тогда группы $H_1(G_1)$ и $H_1(G_2)$ изоморфны.

Доказательство. Докажем, что если G — наложение G_1 и G_2 и если G — правильный S -граф, то $H_1(G_1) = H_1(G)$. Граф G можно получить из G_1 применением двух операций: 1) делением ребра u с помощью новой вершины на две части и 2) делением области s с помощью нового ребра на две части*.

Покажем, что при выполнении каждой из этих операций группа $H_1(G)$ не изменяется.

1°. Пусть граф G' получен из G применением операции 1) к ребру u . Тогда $C_0(G)$ есть подгруппа $C_0(G')$. Группу $C_1(G)$ можно отобразить в $C_1(G')$ так. Ориентируем ребра u' и u'' , полученные делением u вдоль u . Отобразим каждое ребро G , кроме u в него само, как ребро G' . Ребро u отобразим в сумму $u' + u''$. По аддитивности продолжим это соответствие до гомоморфизма групп $C_1(G) \rightarrow C_1(G')$. Очевидно, что ядро этого гомоморфизма есть нуль и поэтому $C_1(G)$ можно считать подгруппой $C_1(G')$. Группы $C_2(G)$ и $C_2(G')$ совпадают.

* Концы этого ребра — вершины в G .

Пусть x — цикл в $C_1(G')$. Тогда, очевидно, что коэффициенты в x при u' и u'' совпадают и, следовательно, x — цикл в $C_1(G)$. Отсюда следует, что справедливо включение

$$\text{Ker } d_1 \subset C_1(G).$$

Так как множество $\text{Im } d_2$ вложено в множество $\text{Ker } d_1$, то $\text{Im } d_2 \subset C_1(G)$. Отсюда ясно, что группа H_1 не изменилась при применении преобразования 1) к ребру u .

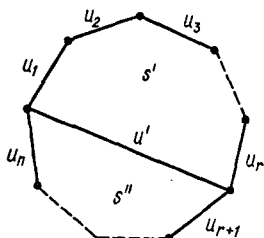


Рис. 44

2°. Пусть граф G' получен из G преобразованием 2). Очевидно, что $C_0(G)$ есть подгруппа в $C_0(G')$, $C_1(G)$ — подгруппа в $C_1(G')$. То, что $C_2(G)$ есть подгруппа в $C_2(G')$, доказывается так же, как и в 1°. При этом области s , подвергшейся делению, сопоставляется сумма $s' + s''$ областей, получившихся в результате деления. Ориентации s' и s''

выбираются согласованными с ориентацией s .

Докажем следующие утверждения*:

а) если $x^{(1)} \in C_1(G')$, то существует такой элемент $x' \in C_1(G)$, что $x^{(1)} \sim x'$ относительно $\text{Im } d_2'$;

б) если $x^{(1)} \sim x^{(2)}$ ($x^{(1)} \in C_1(G)$, $x^{(2)} \in C_1(G')$) относительно $\text{Im } d_2'$, то $x^{(1)} \sim x^{(2)}$ относительно $\text{Im } d_2$;

в) если $x^{(1)} \sim x^{(2)}$ ($x^{(1)} \in C_1(G)$, $x^{(2)} \in C_1(G')$) относительно $\text{Im } d_2'$, то $x^{(1)} \sim x^{(2)}$ относительно $\text{Im } d_2$.

Доказательство. а) Если $x^{(1)}$ содержит не равный нулю коэффициент при новом ребре, то, очевидно, справедливо разложение

$$x^{(1)} = \sum_{i=1}^{S_1} k_i u_i + k' u',$$

где u' — добавленное ребро (рис. 44). С другой стороны, справедливо равенство

$$d_2 s' = u' + \sum_{i=1}^r \pm u_i.$$

* Все обозначения со штрихом относятся к G' , без штриха — к G . Так x_1 — вершина G , x'_1 — вершина G' (может быть та же самая, что x_1 , но рассматриваемая как вершина G'). Аналогично d_1 , d_2 — граничные гомоморфизмы в $C_2(G)$, $C_1(G)$, а d'_1 , d'_2 — в $C_2(G')$, $C_1(G')$.

Вычитая из $x^{(1)}$ 1-цепь $k'd_2s'$, получим цикл x' такой, что $x^{(1)} \sim x'$.

б) Пусть $x^{(1)} \sim x^{(2)}$ относительно $\text{Im } d_2$. Это означает, что существует элемент $y = \sum_{i=1}^{S_2} k_i s_i \in C_2(G)$, для которого выполнено равенство $x^{(1)} - x^{(2)} = dy$. Заменяя s на $s' + s''$, получаем, что разность $x^{(1)} - x^{(2)}$ можно представить в виде

$$x^{(1)} - x^{(2)} = dy', \quad y' = \sum_{i=1}^{S_2+1} k_i s'_i$$

($\{s'_i\}$ получается из $\{s_i\}$ заменой s на s', s'').

в) Если $x^{(1)} - x^{(2)} = dy'$, то так как в 1-цепи $x^{(1)} - x^{(2)}$ коэффициент при новом ребре есть нуль, то, очевидно, коэффициенты в y' при s' и s'' совпадают. Объединяя области s' и s'' , получаем равенство $x^{(1)} - x^{(2)} = dy$, где $y \in C_2(G)$. ■

Выберем теперь в каждом классе эквивалентности по подгруппе $\text{Im } d_2$ по одному элементу и зафиксируем их. Покажем, что в каждом классе эквивалентности по $\text{Im } d'_2$ содержится точно по одному фиксированному элементу. Пусть $x' \in C_1(G)$ и $\{x'\}$ — класс эквивалентности элемента x' по $\text{Im } d'_2$. В силу утверждения а) в $\{x'\}$ содержится элемент x'' из $C_1(G)$. Для некоторого x из числа выбранных элементов $x'' \sim x$ относительно $\text{Im } d_2$. Но тогда $x'' \sim x$ по $\text{Im } d'_2$ по утверждению б) и $x \in \{x'\}$. Далее, если $x^{(1)}, x^{(2)}$ — выбранные представители, то они не могут содержаться в одном и том же классе эквивалентности по $\text{Im } d'_2$, так как тогда, в силу утверждения в) имелась бы эквивалентность $x^{(1)} \sim x^{(2)}$ относительно $\text{Im } d_2$, что невозможно. Таким образом, построено взаимно-однозначное соответствие $H_1(G)$ и $H_1(G')$. Это соответствие есть гомоморфизм, так как если

$$\{x\} \rightarrow \{x'\}, \quad \{y\} \rightarrow \{y'\},$$

то (считаем, что x — фиксированный представитель $\{x\}$, y — представитель $\{y\}$, а z — представитель $\{x\} + \{y\}$) $z \sim x + y$ относительно $\text{Im } d'_2$ и по утверждению б) относительно $\text{Im } d'_2$; из включений $x \in \{x'\}$, $y \in \{y'\}$ следует включение $x + y \in \{x'\} + \{y'\}$ и поэтому $z \in \{x'\} + \{y'\}$,

т. е.

$$\{x\} + \{y\} \rightarrow \{x'\} + \{y'\}.$$

Для окончания доказательства теоремы 2.2 заметим, что если наложение G_1 на G_2 несвязно, то G_1 содержится в одной из областей G_2 (рис. 45). Дополнив G_2 цепью, имеющей общие точки с G_1 , сделаем наложение G_1 на G_2 связным, а

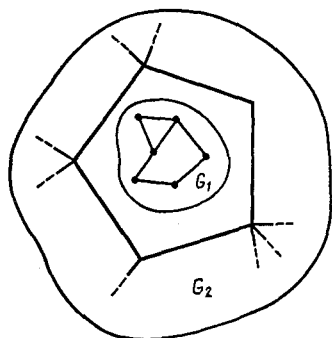


Рис. 45

следовательно, правильным S -графом. С другой стороны, по 2⁰ мы не изменим группы H_1 . Из этих рассуждений следует равенство $H_1(G_1) = H_1(G_2)$. ■

Эта теорема позволяет считать H_1 характеристикой поверхности. Очевидно, если две поверхности эквивалентны, то их группы гомологий изоморфны.

Заметим еще, что естественно считать нулем границу точки. Это дает нам продол-

жение диаграммы (2.1):

$$C_2(G) \xrightarrow{d_2} C_1(G) \xrightarrow{d_1} C_0(G) \xrightarrow{d_0} 0, \quad (2.1')$$

где через ноль обозначена нулевая группа, состоящая из одного нуля. Далее, поскольку в рассмотрении не присутствует элементов трех измерений, то естественно продолжить эту диаграмму так:

$$0 \xrightarrow{d_3} C_2(G) \xrightarrow{d_2} C_1(G) \xrightarrow{d_1} C_0(G) \xrightarrow{d_0} 0. \quad (2.1'')$$

Очевидно, что равенство $d_{i+1} \circ d_i = 0$ выполнено для $i = 0, 1, 2$, и поэтому можно определить группы гомологий $H_i(G) = \text{Ker } d_i / \text{Im } d_{i+1}$.

Нульмерная группа гомологий графа фактически вычислена в теореме 2.3 гл. IV. Поскольку $\text{Ker } d_0 = C_0(G)$, то $H_0(G) = C_0(G) / \text{Im } d_1$, а эта группа, как доказано в теореме 2.3 гл. IV, изоморфна группе $\mathbb{Z} \oplus \dots \oplus \mathbb{Z}$, где число прямых слагаемых равно числу компонент связности. Правильный S -граф на связной поверхности S всегда связан*,

* Как уже говорилось выше, такие утверждения топологического характера считаются здесь очевидными; строгие доказательства увели бы слишком далеко за рамки этой книги.

поэтому число связных компонент любого связного S -графа равно числу связных компонент S . Это доказывает независимость $H_0(G)$ от G . В дальнейшем группа $H_0(G)$ для некоторого правильного S -графа G обозначается $H_0(S)$.

Вычислим $H_1(S)$ для трех многообразий.

Т о р T^2 . Тор можно представить как квадрат с противоположными сторонами, склеенными так, как показано на рис. 46, а. На этом же рисунке изображен правильный T^2 -граф. Он же изображен непосредственно на торе (рис. 46, б).

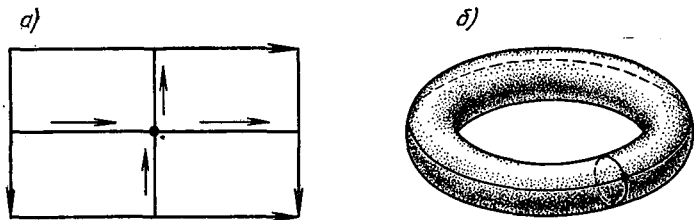


Рис. 46

Вычислим группу $H_1(G)$:

- 1) $C_2(G) = \mathbb{Z}$, общий элемент $C_2(G)$ имеет вид ks_1 ;
- 2) $C_1(G) = \mathbb{Z} \oplus \mathbb{Z}$, общий элемент $C_1(G)$ имеет вид $k_1u_1 + k_2u_2$;
- 3) $C_0(G) = \mathbb{Z}$, общий элемент $C_0(G)$ имеет вид kx_1 .

Вычислим $\text{Ker } d_1$:

$$d_1(k_1u_1 + k_2u_2) = (k_1 + k_2)(x_1 - x_1) = 0; \quad \text{Ker } d_1 = C_1(G).$$

Вычислим $\text{Im } d_2$:

$$d_2(ks_1) = k(u_1 - u_2 - u_1 + u_2) = 0; \quad \text{Im } d_2 = 0,$$

отсюда

$$H_1(T^2) = H_1(G) = \text{Ker } d_1 / \text{Im } d_2 = \mathbb{Z} \oplus \mathbb{Z}.$$

Б у т ы л к а К л я й н а. Здесь отождествление сторон квадрата нужно проводить иначе, чем в случае T^2 (соответствующий чертеж вместе с правильным S -графом дан на рис. 47).

Проведем вычисления:

- 1) $C_2(G) = \mathbb{Z}$, общий элемент группы $C_2(G)$ есть ks_1 ;
- 2) $C_1(G) = \mathbb{Z} \oplus \mathbb{Z}$, общий элемент группы $C_1(G)$ есть $k_1u_1 + k_2u_2$;
- 3) $C_0(G) = \mathbb{Z}$, общий элемент группы $C_0(G)$ есть kx_1 .

Вычислим ядро d_1 :

$$d_1(k_1u_1 + k_2u_2) = k_1(x_1 - x_1) + k_2(x_1 - x_1) = 0,$$

$$\text{Ker } d_1 = \mathbb{Z} \oplus \mathbb{Z} = C_1(G); \quad \text{далее,} \quad d_2(ks_1) = k(u_2 + u_1 - u_2 + u_1) = 2ku_1.$$

Таким образом, образ гомоморфизма d_2 есть

$$\text{Im } d_2 = 2\mathbb{Z} \oplus 0 \subset \mathbb{Z} \oplus \mathbb{Z} = \text{Ker } d_1.$$

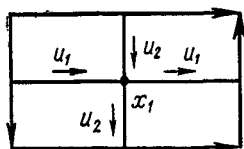


Рис. 47

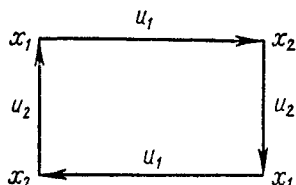


Рис. 48

Нетрудно видеть, что группа $H_1(S)$ определяется равенством

$$H_1(S) = \text{Ker } d_1 / \text{Im } d_2 = \mathbb{Z} \oplus \mathbb{Z}.$$

Проективная плоскость $\mathbb{RP}^2$. Вычисления проведем для двух различных правильных $\mathbb{RP}^2$ -графов.

Способ 1. Правильный $\mathbb{RP}^2$ -граф показан на рис. 48.

Вычислим группы $C_i(G)$:

- 1) $C_2(G) = \mathbb{Z}$, общий элемент $C_2(G)$ есть ks_1 ;
- 2) $C_1(G) = \mathbb{Z} \oplus \mathbb{Z}$, общий элемент $C_1(G)$ есть $k_1u_1 + k_2u_2$;
- 3) $C_0(G) = \mathbb{Z} \oplus \mathbb{Z}$, общий элемент $C_0(G)$ есть $k_1x_1 + k_2x_2$.

Найдем ядро d_1 :

$$d_1(k_1u_1 + k_2u_2) = k_1(x_2 - x_1) + k_2(x_1 - x_2) = (k_1 - k_2)(x_2 - x_1).$$

Отсюда ядро d_1 есть $\mathbb{Z}$ с общим элементом $k(u_1 + u_2)$. Вычислим образ d_2 :

$$d_2(ks_1) = k(u_1 + u_2 + u_1 + u_2) = 2k(u_1 + u_2).$$

Отсюда получим соотношение

$$\text{Im } d_2 = 2\mathbb{Z} \subset \mathbb{Z} = \text{Ker } d_1$$

и окончательно имеем

$$H_1(\mathbb{RP}^2) = \text{Ker } d_1 / \text{Im } d_2 = \mathbb{Z}_2.$$

Способ 2. Правильный $\mathbb{RP}^2$ -граф показан на рис. 49, а. На рис. 49, б, в указаны ориентации областей s_1 и s_2 , которые также указаны на рис. 49, а.

Проведем вычисления:

1) $C_2(G) = \mathbb{Z} \oplus \mathbb{Z}$, общий элемент $C_2(G)$ есть $k_1 s_1 + k_2 s_2$;

2) $C_1(G) = \mathbb{Z} \oplus \mathbb{Z}$, общий элемент $C_1(G)$ есть $k_1 u_1 + k_2 u_2$;

3) $C_0(G) = \mathbb{Z}$, общий элемент $C_0(G)$ есть kx_1 .

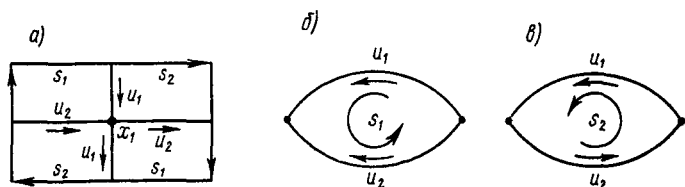


Рис. 49

Вычислим образ d_1 :

$$d_1(k_1 u_1 + k_2 u_2) = k_1(x_1 - x_1) + k_2(x_1 - x_1) = 0;$$

$$\text{Ker } d_2 = \mathbb{Z} \oplus \mathbb{Z} = C_1(G);$$

$$d_2(k_1 s_1 + k_2 s_2) = k_1(u_1 - u_2) + k_2(u_1 + u_2) = (k_1 + k_2) u_1 + (k_1 - k_2) u_2.$$

Таким образом, группа $\text{Im } d_2$ есть подгруппа группы $\mathbb{Z} \oplus \mathbb{Z} = C_1(G)$, состоящая из таких элементов $t_1 u_1 + t_2 u_2$, в которых t_1 и t_2 имеют одинаковую четность. Очевидно, что это — подгруппа $\mathbb{Z} \oplus \mathbb{Z}$. Существует ровно два класса эквивалентности по этой подгруппе, а именно:

$$\{0\} = \{t_1 u_1 + t_2 u_2 \mid t_1, t_2 \text{ одинаковой четности}\};$$

$$\{1\} = \{t_1 u_1 + t_2 u_2 \mid t_1, t_2 \text{ разной четности}\}.$$

Очевидно, что $\{0\} + \{0\} = \{0\}$, $\{0\} + \{1\}$, $\{1\} + \{1\} = \{0\}$, так что фактор-группа есть $\mathbb{Z}_2$:

$$H_1(\mathbb{RP}^2) = \mathbb{Z}_2,$$

что согласуется с результатами случая 1.

Группы $\mathbb{Z}_2$, $\mathbb{Z}_2 \oplus \mathbb{Z}$ и $\mathbb{Z} \oplus \mathbb{Z}$ неизоморфны (докажите это!), поэтому поверхности $\mathbb{RP}^2$, T^2 и S неэквивалентны.

В примерах 1⁰—3⁰ рассматривались не простые графы, а мультиграфы. Но доказательство теорем 2.1 и 2.2 дословно переносится на случай правильных S -мультиграфов.

Глава VI

РАСКРАСКИ ГРАФОВ. ХРОМАТИЧЕСКОЕ ЧИСЛО

В гл. I уже встречалась задача о раскраске графов. Напомним ее постановку. Пусть дана географическая карта страны, в которой имеется несколько областей. Требуется окрасить каждую область так, чтобы любые две области, граничащие между собой, были окрашены в различные цвета. При этом «граничащими» областями называются области, которые граничат по некоторой линии (а не в точке). Во введении было показано, что эта задача может быть сведена к задаче о раскраске плоского графа: *имея некоторое число красок, раскрасить каждую вершину одной из этих красок таким образом, чтобы любые две смежные вершины имели различную окраску.*

Задача об определении минимального числа красок, нужных для правильной раскраски графа, которую мы здесь рассмотрим, является одной из первых задач теории графов. В этой главе слово «граф» будет обозначать полностью неориентированный граф без кратных ребер.

§ 1. ВЫЧИСЛЕНИЕ ХРОМАТИЧЕСКОГО ЧИСЛА ЗАДАННОГО ГРАФА

1. Полином $\pi(G, \lambda)$ и хроматическое число. Дадим сначала точные формулировки. Отметим некоторую вольность в обозначениях. Говоря, что для графа $G = (X, \bar{\Gamma})$ $(x_i, x_j) \in \bar{\Gamma}$, мы будем подразумевать, что в графе G существует неориентированное ребро, соединяющее вершины x_i и x_j , т. е. $\{(x_i, x_j), (x_j, x_i)\} \in \bar{\Gamma}$. Заметим, что задача о раскраске мультиграфа сводится к задаче о раскраске графа, полученного склеиванием всех кратных ребер.

Дадим теперь определения правильной раскраски и хроматического числа.

Определение 1.1. Граф $G = (X, \Gamma)$ называется правильно раскрашенным λ красками, если каждая его вершина раскрашена одной из λ красок и если из $(x_i, x_j) \in \Gamma$ следует, что x_i и x_j раскрашены разными красками. Если все эти требования выполнены, то мы говорим также, что задана правильная раскраска графа G λ красками.

Определение 1.2. Граф $G = (X, \Gamma)$ называется p -хроматическим, если существует правильная раскраска графа G p красками. Минимальное из таких чисел p называется хроматическим числом данного графа.

Вычисление хроматического числа заданного графа может быть проведено с использованием вводимой ниже функции $\pi(G, \lambda)$.

Определение 1.3. Для заданного графа G и натурального числа λ через $\pi(G, \lambda)$ обозначается число всех различных правильных раскрасок графа G с помощью λ красок.

Выведем теперь формулы, связывающие число $\pi(G, \lambda)$ с другими характеристиками графа. Первые две из них дают возможность вычислять $\pi(G, \lambda)$ только для сравнительно простых графов. Остальные формулы связывают $\pi(G, \lambda)$ с $\pi(G_i, \lambda)$ для его более простых компонент G_i . Такие формулы дают возможность вычислять $\pi(G, \lambda)$ для данного графа, последовательно строя его из более простых частей («разбирая» его на более простые части). Примеры вычисления $\pi(G, \lambda)$ для некоторых графов будут даны ниже.

Теорема 1.1. Справедлива формула

$$\pi(G, \lambda) = \lambda^{S_0} - \sum_i \mu(u_i) + \sum_{i \neq j} \mu(u_i, u_j) - \sum_{i \neq j \neq k} \mu(u_i, u_j, u_k) \dots, \quad (1.1)$$

где S_0 — количество вершин графа G ; u_i ($i = 1, 2, \dots, S_1$) — ребра графа G , перенумерованные некоторым способом, а $\mu(u_i, u_j, \dots, u_l)$ — число таких раскрасок графа G , при которых концы каждого из ребер $u_i, u_j, \dots, u_l$ закрашены одним цветом.

Доказательство. Очевидно, что число всех возможных (не обязательно правильных) раскрасок графа G равно λ^{S_0} . Из этого числа нужно вычесть количество тех раскрасок, в которых концы хотя бы одного ребра закрашены одинаково. Пусть число раскрасок, в которых ровно k

ребер имеют концы одного цвета, есть v_k . Имеем соотношение

$$\pi(G, \lambda) = \lambda^{S_0} - \sum_{k=1}^{S_1} v_k. \quad (1.2)$$

Подсчитаем число v_1 . Рассмотрим сумму $\sum_i \mu(u_i)$. В этой сумме раскраски, входящие в v_2 , учтены дважды. Именно, если концы ребер u_{i_1} и u_{i_2} закрашены одинаково, то эта раскраска входит в $\mu(u_{i_1})$ и $\mu(u_{i_2})$. Аналогично, раскраски, входящие в v_k , учтены k раз; получим

$$v_1 = \sum_i \mu(u_i) - \sum_{i=2}^{S_1} i v_i.$$

Учитывая, что $i = C_i^1$, перепишем эту формулу в виде

$$\sum_i \mu(u_i) = \sum_{i=1}^{S_1} C_i^1 v_i. \quad (1.3)$$

Аналогичными рассуждениями получим формулу

$$v_k = \sum_{i_1 \neq \dots \neq i_k} \mu(u_{i_1}, \dots, u_{i_k}) - \sum_{i=k+1}^{S_1} C_i^k v_i.$$

В самом деле, в сумме $\sum_{i_1 \neq \dots \neq i_k} \mu(u_{i_1}, \dots, u_{i_k})$ раскраска, у которой закрашены одинаково концы ребер $u_{i_1}, \dots, u_{i_l}$ ($l > k$), учтена C_l^k раз, что и дает последнюю формулу. Учитывая, что $C_k^k = 1$, перепишем эту формулу в виде

$$\sum_{i_1 \neq \dots \neq i_k} \mu(u_{i_1}, \dots, u_{i_k}) = \sum_{i=k}^{S_1} C_i^k v_i. \quad (1.4)$$

Используя формулы (1.3) и (1.4), составим сумму

$$- \sum_i \mu(u_i) + \sum_{i_1 \neq i_2} \mu(u_{i_1}, u_{i_2}) \dots = \sum_{k=1}^{S_1} (-1)^k \sum_{i=k}^{S_1} C_i^k v_i.$$

Меняя в этом выражении порядок суммирования [все вхо-

дящие в сумму индексы (i, k) заключены (рис. 50) в заштрихованный треугольник], получим:

$$\begin{aligned}
 & - \sum_i \mu(u_i) + \sum_{i_1 \neq i_2} \mu(u_{i_1}, u_{i_2}) - \sum_{i_1 \neq i_2 \neq i_3} \mu(u_{i_1}, u_{i_2}, u_{i_3}) + \\
 & + \dots = \sum_{i=1}^{S_1} v_i \sum_{k=1}^i C_i^k (-1)^k.
 \end{aligned}$$

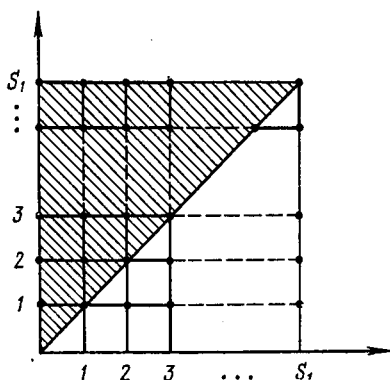


Рис. 50

Теперь, воспользовавшись формулой бинома Ньютона

$$0 = (1 - 1)^i = \sum_{k=0}^i C_i^k (-1)^k = 1 + \sum_{k=1}^i C_i^k (-1)^k,$$

напишем

$$- \sum_i \mu(u_i) + \sum_{i_1 \neq i_2} \mu(u_{i_1}, u_{i_2}) - \dots = - \sum_{i=1}^{S_1} v_i.$$

Подставляя последнюю формулу в соотношение (1.2), получим формулу (1.1). ■

Отметим, что формула (1.1), несмотря на свой общий характер, непригодна для фактических вычислений $\pi(G, \lambda)$ для конкретных графов и представляет скорее теоретический интерес. Следующая теорема дает формулу, удобную для вычислений при достаточно малых S_0 . Чтобы выписать эту формулу, введем обозначение. Обозначим через $p(s, c)$ число суграфов графа G , включающих s ребер и имею-

щих s компонент связности; при этом, если при данных s и c таких суграфов не существует, то полагаем $p(s, c) = 0$.

Теорема 1.2. *Справедлива формула*

$$\pi(G, \lambda) = \sum_{s, c} (-1)^s p(s, c) \lambda^c. \quad (1.5)$$

Доказательство. Вычислим отдельно каждый член, входящий в формулу (1.1). Пусть дан набор ребер

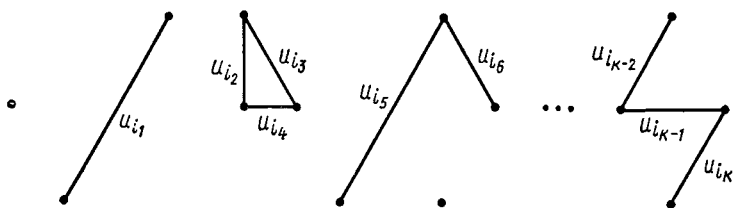


Рис. 51

$u_{i_1}, \dots, u_{i_k}$. Рассмотрим суграф графа G , включающий ребра $u_{i_1}, \dots, u_{i_k}$ и только их. Пусть число компонент связности этого суграфа* равно s . Если дана раскраска графа G , у которого концы ребер $u_{i_1}, \dots, u_{i_k}$ раскрашены попарно одинаково, то в рассматриваемом суграфе одинаково раскрашенными оказываются его компоненты связности. Обратно, раскрасив каждую компоненту связности суграфа (рис. 51) в свой цвет, получим раскраску графа, у которой концы ребер $u_{i_1}, \dots, u_{i_k}$ раскрашены одинаково. Число таких раскрасок, очевидно, равно λ^c отсюда следует равенство

$$\sum_{i_1 \neq \dots \neq i_k} \mu(u_{i_1}, \dots, u_{i_k}) = \sum_c p(k, c) \lambda^c,$$

где суммирование распространено на все c , для которых $p(k, c) \neq 0$. Очевидно, что таких c — конечное число, так как, например, количество компонент связности любого суграфа меньше или равно S_0 . Подставляя последнюю формулу в (1.1), получим (1.5).

* Напомним, что, по определению суграфа, он включает все вершины графа G .

Выведенная формула (1.5) описывает структуру $\pi(G, \lambda)$ как функции переменной λ . Непосредственно очевидны следующие свойства функции $\pi(G, \lambda)$:

- 1) $\pi(G, \lambda)$ есть многочлен степени S_0 с коэффициентом при старшем члене, равном единице;
- 2) коэффициент при λ^0 в $\pi(G, \lambda)$ равен нулю, откуда видно, что $\pi(G, \lambda)$ как многочлен делится на λ .

Как уже говорилось выше, эта формула дает возможность вычислять $\pi(G, \lambda)$ при не слишком большом S_0 . В ка-

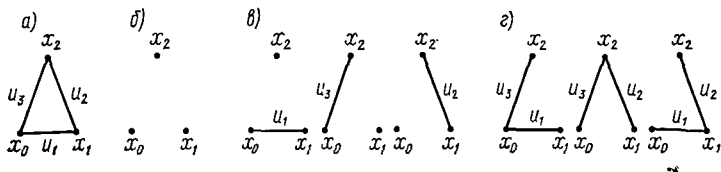


Рис. 52

честве примера вычислим $\pi(G, \lambda)$ для полного графа с тремя вершинами (рис. 52, а). Все суграфы этого графа показаны на рис. 52, б, в, г. На рис. 52, б показан суграф с $s = 0$, $c = 3$; на рис. 52, в — три суграфа с $s = 1$, $c = 2$; на рис. 52, г — три суграфа с $s = 2$, $c = 1$; на рис. 52, а — сам граф, который можно рассматривать как суграф с $s = 3$ и $c = 1$. Числа $p(s, c)$ для этого графа поэтому есть $p(0, 3) = 1$; $p(1, 2) = 3$, $p(2, 1) = 3$, $p(3, 1) = 1$ и в остальных случаях $p(s, c) = 0$. Пользуясь формулой (1.5), имеем соотношение

$$\pi(\Delta, \lambda) = \lambda^3 - 3\lambda^2 + 3\lambda - \lambda = \lambda(\lambda - 1)(\lambda - 2).$$

Из последней формулы следует, что хроматическое число этого графа есть 3, так как из равенств

$$\pi(G, 1) = \pi(G, 2) = 0; \pi(G, 3) = 3! = 6$$

следует, что граф G можно закрасить тремя красками, но нельзя закрасить одной или двумя.

Следующие четыре леммы позволяют свести подсчет функции $\pi(G, \lambda)$ к подсчету этой функции для более простых графов.

Лемма 1.1. Если граф G состоит из двух несвязных частей G_1 и G_2 , то справедлива формула

$$\pi(G, \lambda) = \pi(G_1, \lambda) \cdot \pi(G_2, \lambda).$$

Доказательство. Эта формула — очевидное следствие того факта, что раскраски G_1 и G_2 можно выбирать независимо. ■

Лемма 1.2. Если граф G получен из двух графов склеиванием по одной точке x_0 (рис. 53), то

$$\pi(G, \lambda) = \frac{1}{\lambda} \pi(G_1, \lambda) \cdot \pi(G_2, \lambda).$$

Доказательство. Очевидно, что эта формула будет доказана, если число раскрасок графа G_1 , в которых данная вершина имеет предписанный цвет, равно $\frac{1}{\lambda} \pi(G_1, \lambda)$.

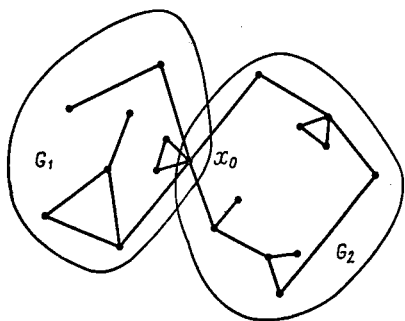


Рис. 53

λ). Докажем это.

Разобьем множество всех раскрасок графа G_1 на классы, относя к одному и тому же классу все раскраски, в которых вершина x_0 окрашена в данный цвет. Очевидно, что число таких классов равно λ . Докажем, что число элементов каждого класса одинаково. Пусть $\alpha_1, \dots, \alpha_\lambda$ — краски. Пусть в

классе A вершина x_0 окрашена краской α_1 , в классе B — краской α_2 . Перекрасив каждую раскраску из A , заменяя краски в соответствии с подстановкой

$$\alpha_1 \rightarrow \alpha_2 \rightarrow \alpha_3 \rightarrow \dots \rightarrow \alpha_\lambda \rightarrow \alpha_1,$$

получим раскраски класса B . При этом различные раскраски переходят в различные, и мы получим класс B целиком, так как обратное перекрашивание переводит B в A . Получено взаимно-однозначное соответствие между элементами A и B . Эти классы конечны, поэтому их количества элементов равны. Из произвольности классов A и B следует, что число элементов каждого класса есть $\frac{1}{\lambda} \pi(G_1, \lambda)$. Раскра-

шивая произвольно граф G_2 и сочетая эти раскраски с теми раскрасками G_1 , у которых x_0 раскрашена в тот же цвет, получим $\pi(G_2, \lambda) \cdot \frac{1}{\lambda} \pi(G_1, \lambda)$ различных раскрасок.

Очевидно, что все раскраски учтены. ■

Лемма 1.3. Пусть граф G получен из двух несвязных графов G_1 и G_2 склейкой с помощью ребра l , внешнего для обоих графов (рис. 54). Тогда справедливо соотношение

$$\pi(G, \lambda) = \left(1 - \frac{1}{\lambda}\right) \pi(G_1, \lambda) \pi(G_2, \lambda).$$

Доказательство. Ясно, что если G' — граф, полученный добавлением l к G_2 (см. рис. 54), то

$$\pi(G', \lambda) = (\lambda - 1) \pi(G_2, \lambda).$$

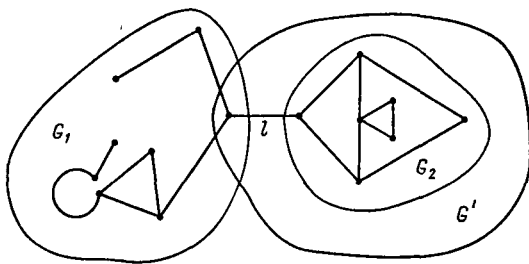


Рис. 54

Пользуясь леммой 1.2, получаем

$$\pi(G, \lambda) = \frac{1}{\lambda} \pi(G_1, \lambda) (\lambda - 1) \pi(G_2, \lambda). \blacksquare$$

Лемма 1.4. Пусть граф G получен из графа G_1 добавлением ребра без изменения числа вершин (рис. 55, а). Тогда для $\pi(G, \lambda)$ справедлива формула

$$\pi(G, \lambda) = \pi(G_1, \lambda) - \pi(G', \lambda),$$

где граф G' (рис. 55, б) получен из G_1 склеиванием вершин, инцидентных ребру l и преобразованием полученного мультиграфа в простой граф склеиванием кратных ребер.

Доказательство. Очевидно, что число раскрасок графа G равно $\pi(G_1, \lambda)$ минус число раскрасок графа G_1 , в которых концы ребра l одинаково раскрашены. Но последнее число есть $\pi(G', \lambda)$.

2. Примеры вычисления числа $\pi(G, \lambda)$ для различных графов. Покажем, как соотношения, доказанные в леммах 1.1—1.4, применяются для вычисления функции $\pi(G, \lambda)$.

1. Рассмотрим n -мерный симплекс

$$S_n = (X, \Gamma).$$

где $X = \{x_0, \dots, x_n\}$, а $\Gamma = X \times X \setminus E$, E — множество элементов вида (a, a) (диагональ).

Так как каждые две вершины соединены между собой, то S_n нельзя раскрасить менее чем $n + 1$ красками. Отсюда ясно, что

$$\pi(G, \lambda) = 0 \text{ при } \lambda \leq n \text{ } (\lambda \text{ — целое}).$$

Кроме того, как уже отмечалось выше (свойства (1) и (2)), $\pi(G, \lambda)$ есть многочлен степени $n+1$, делящийся на λ .

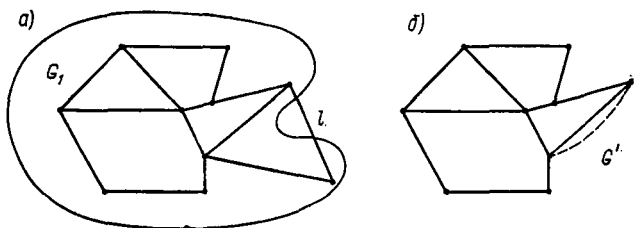


Рис. 55

Его корнями должны поэтому быть числа $0, 1, 2, \dots, n$. Учитывая еще, что коэффициент при старшем члене равен 1, имеем:

$$\pi(G, \lambda) = \lambda(\lambda - 1) \dots (\lambda - n). \quad (1.6)$$

В частности, $\pi(G, n + 1) = (n + 1)!$

Дадим доказательство этой формулы с использованием леммы 1.4. Выделим некоторую вершину графа и перенумеруем ребра, исходящие из нее, числами $1, \dots, n$. Обозначим $G^{(k)}$ граф S_n с выкинутыми ребрами $u_1, \dots, u_k$. Тогда $G^{(n)}$ есть граф, состоящий из изолированной точки и S_{n-1} . Выкидываем ребра в соответствии с леммой 1.4 по очереди. Отмечаем, что на каждом этапе $G' = S_{n-1}$; равенства

$$\pi(S_n, \lambda) = \pi(G^{(1)}, \lambda) - \pi(S_{n-1}, \lambda),$$

$$\pi(G^{(1)}, \lambda) = \pi(G^{(2)}, \lambda) - \pi(S_{n-1}, \lambda),$$

.....

$$\pi(G^{(n-1)}, \lambda) = \pi(G^{(n)}, \lambda) - \pi(S_{n-1}, \lambda),$$

сумма которых есть

$$\pi(S_n, \lambda) = \pi(G^{(n)}, \lambda) - n\pi(S_{n-1}, \lambda).$$

Заметим теперь, что в силу леммы 1.1 справедливо равенство $\pi(G^{(n)}, \lambda) = \lambda\pi(S_{n-1}, \lambda)$, из которого следует:

$$\pi(S_n, \lambda) = (\lambda - n)\pi(S_{n-1}, \lambda).$$

Отмечая, что $\pi(S_0, \lambda) = \lambda$, так как S_0 есть граф, состоящий из одной вершины и без ребер, получаем формулу (1.6) по индукции. ■

2. Пусть d_n — дерево с n вершинами. Вычислим функцию $\pi(d_n, \lambda)$.

Удалим из дерева одно из ребер, являющееся тупиком. При применении леммы 1.4 заметим, что граф G' в этом случае есть d_{n-1} , а G_1 есть граф, состоящий из d_{n-1} и изолированной вершины. Отсюда следуют равенства

$$\begin{aligned}\pi(d_n, \lambda) &= \pi(G_1, \lambda) - \pi(G', \lambda) = \lambda\pi(d_{n-1}, \lambda) - \pi(d_{n-1}, \lambda) = \\ &= (\lambda - 1)\pi(d_{n-1}, \lambda).\end{aligned}$$

Заметив, что $\pi(d_1, \lambda) = \lambda$ (d_1 — изолированная вершина), по индукции получим:

$$\pi(d_n, \lambda) = \lambda(\lambda - 1)^{n-1}. \quad (1.7)$$

3. Пусть c_n — цикл с n вершинами. Для вычисления функции $\pi(G, \lambda)$ используем лемму 1.4. При удалении любого из ребер цикла будем иметь равенства

$$\pi(c_n, \lambda) = \pi(d_n, \lambda) - \pi(c_{n-1}, \lambda) = \lambda(\lambda - 1)^{n-1} - \pi(c_{n-1}, \lambda).$$

Запишем эти равенства при $n = 4, 5, \dots, n_0$

$$\pi(c_4, \lambda) = \lambda(\lambda - 1)^3 - \pi(c_3, \lambda),$$

$$\pi(c_5, \lambda) = \lambda(\lambda - 1)^4 - \pi(c_4, \lambda),$$

$$\dots \dots \dots$$

$$\pi(c_{n_0}, \lambda) = \lambda(\lambda - 1)^{n_0-1} - \pi(c_{n_0-1}, \lambda)$$

и умножим их на $(-1)^k$ ($4 \leq k \leq n_0$):

$$(-1)^k \pi(c_k, \lambda) = [\lambda(\lambda - 1)^{k-1} - \pi(c_{k-1}, \lambda)] (-1)^k.$$

Просуммируем эти равенства по k от 4 до n_0 :

$$\begin{aligned} \sum_{k=4}^{n_0} (-1)^k \pi(c_k, \lambda) &= \sum_{k=4}^{n_0} (-1)^k \lambda (\lambda - 1)^{k-1} - \\ &- \sum_{k=4}^{n_0} (-1)^k \pi(c_{k-1}, \lambda) = -\lambda \frac{(-1)^{n_0} (\lambda - 1)^{n_0} + (\lambda - 1)^3}{-\lambda} - \\ &- \pi_3(c_3, \lambda) + \sum_{k=4}^{n_0-1} (-1)^k \pi(c_k, \lambda). \end{aligned}$$

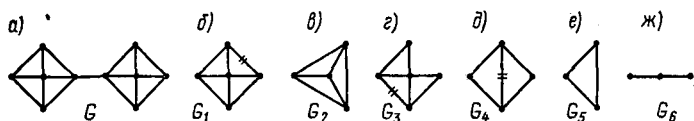


Рис. 56

Отсюда следует формула

$$\pi(c_{n_0}, \lambda) = (-1)^{n_0} [(-1)^{n_0} (\lambda - 1)^{n_0} + (\lambda - 1)^3 - \lambda(\lambda - 1)(\lambda - 2)] \quad (1.8)$$

4. Пусть G — граф, показанный на рис. 56, а. Вычислим функцию $\pi(G, \lambda)$.

Используя лемму 1.3, получаем

$$\pi(G, \lambda) = \left(1 - \frac{1}{\lambda}\right) (\pi(G_1, \lambda))^2.$$

Вычисление $\pi(G_1, \lambda)$ проводим с помощью леммы 1.4, выкидывая отмеченное на рис. 56, б ребро. В результате получим:

$$\begin{aligned} \pi(G_1, \lambda) &= \pi(G_3, \lambda) - \pi(G_2, \lambda) = \pi(G_3, \lambda) - \\ &- \lambda(\lambda - 1)(\lambda - 2)(\lambda - 3), \end{aligned}$$

так как граф G_2 (рис. 56, в) есть симплекс S_3 . Вычисляем $\pi(G_3, \lambda)$ с помощью леммы 1.4, выбрасывая отмеченное на рис. 56, г ребро:

$$\pi(G_3, \lambda) = \frac{1}{\lambda} [\pi(S_2, \lambda)]^2 - \pi(G_4, \lambda).$$

Аналогично получим соотношение

$$\pi(G_4, \lambda) = \pi(G_5, \lambda) - \pi(G_6, \lambda) = \pi(c_4, \lambda) - \pi(d_3, \lambda).$$

(Графы G_4 , G_5 и G_6 изображены на рис. 56, *д*, *е*, *ж*.) Собирая все эти формулы и пользуясь (1.7) и (1.8), можно получить выражение для $\pi(G, \lambda)$. Это выражение мы здесь приводить не будем и вывод его предоставляем сделать читателю в качестве упражнения.

Возникает вопрос: любой ли многочлен с целыми коэффициентами, делящийся на λ , есть $\pi(G, \lambda)$ для некоторого графа G . Чтобы ответить на это, выведем еще некоторые свойства $\pi(G, \lambda)$, ясные уже из определения этого числа:

- 1) если $\pi(G, n) = 0$, то $\pi(G, \lambda) = 0$ при $\lambda < n$;
- 2) если $\pi(G, n) \neq 0$, то $\pi(G, \lambda) \neq 0$ при $\lambda > n$;
- 3) значения $\pi(G, \lambda)$ неотрицательны при целых положительных λ .

Уже из этих трех свойств ясно, что не любой многочлен представим в виде $\pi(G, \lambda)$. Например, многочлен

$$\lambda(\lambda - 2) = P(\lambda)$$

уже не есть $\pi(G, \lambda)$ ни для какого графа G , так как $P(2) = 0$, $P(1) \neq 0$.

С помощью $\pi(G, \lambda)$ можно дать критерий p -хроматичности и метод вычисления хроматического числа для заданного графа G . Именно, граф G является p -хроматическим тогда и только тогда, когда $\pi(G, p) \neq 0$. Хроматическое число можно вычислить по формуле

$$\min\{p : p \text{ целое, } \pi(G, p) \neq 0\}.$$

Заметим, что для определения хроматического числа графа можно использовать методы линейного программирования.

§ 2. ОЦЕНКА ХРОМАТИЧЕСКОГО ЧИСЛА ПЛОСКОГО ГРАФА

1. Теорема о пяти красках. Теорема утверждает, что любой граф, обладающий плоской реализацией, может быть правильно раскрашен пятью красками. Вспоминая задачу, сформулированную в начале этой главы, видим, что из этой теоремы следует возможность раскрашивания любой карты пятью цветами.

Вопрос о том, достаточно ли для любой карты четырех красок, до сих пор не решен. Это так называемая проблема четырех цветов. Примеры плоских графов, не раскра-

шивающихся тремя цветами, строятся очень легко и будут приведены ниже.

Теорема 2.1. *Любой плоский граф 5-хроматичен.*

Прежде чем доказывать теорему 2.1, докажем лемму, которая имеет чисто технический характер и будет использоваться при доказательстве теоремы.

Лемма 2.1. *В любом простом связном плоском графе без тупиков и перешейков существует вершина степени меньшей или равной пяти.*

Доказательство. Предположим, что степень любой вершины графа $G = (X, \Gamma)$ больше пяти. Из этого предположения вытекает следующая оценка для числа ребер графа S_1 :

$$2S_1 = \sum_{x_i \in X} \text{st } x_i \geq 6x_0.$$

По условию теоремы граф G является плоским. Как и ранее, будем отождествлять G и его плоскую реализацию. *Гранью графа G* поэтому называется любая связная часть плоскости, ограниченная ребрами, принадлежащими к плоской реализации G . Количество ребер, ограничивающих каждую грань графа не меньше трех, так как граф не содержит кратных ребер. Отсюда имеем неравенство

$$3S_2 \leq 2S_1.$$

Подставим теперь оценки

$$S_0 \leq \frac{1}{3} S_1; \quad S_2 \leq \frac{2}{3} S_1$$

в выражение для эйлеровой характеристики графа G :

$$\chi(G) = S_0 - S_1 + S_2 \leq \frac{1}{3} S_1 - S_1 + \frac{2}{3} S_1 = 0.$$

Это противоречит тому, что, по теореме 1.1 гл. V,

$$\chi(G) = 2. \blacksquare$$

Доказательство теоремы 2.1. Рассмотрим следующие операции над мультиграфом:

- 1) удаление тупика (вместе с вершиной);
- 2) склеивание двух ребер, инцидентных одной и той же паре вершин x_i, x_j ;
- 3) удаление перешейка.

Очевидно, что если после проведения любой из указанных операций над графом G его хроматическое число стало

равным p , то хроматическое число исходного графа G есть $\max(2, p)$. Кроме того, после применения некоторого количества операций 1) — 3) получается простой граф, состоящий из конечного числа связных компонент, в каждой из которых нет тупиков и перешейков. Если хроматические числа этих компонент есть $p_1, \dots, p_m$, то хроматическое число графа G есть $\max_{1 \leq j \leq m} (p_j)$. Из этих рассуждений

видно, что достаточно доказать следующее утверждение: *хроматическое число любого простого связного плоского графа без тупиков и перешейков меньше или равно пяти, т. е. такой граф всегда 5-хроматичен.*

Докажем это утверждение. Оно будет доказываться по индукции по числу S_1 ребер графа G . При этом будет использована лемма 2.1.

Б а з и с и н д у к ц и и. При $S_1 = 3$ теорема верна, так как единственным связным графом без тупиков и перешейков с $S_1 = 3$ является треугольник. Треугольник 3-хроматичен, а следовательно, и 5-хроматичен.

И н д у к ц и о н н ы й ш а г. Пусть теорема верна для любого графа с количеством ребер $S_1 \leq n$. Рассмотрим (простой плоский без тупиков и перешейков) граф G с $S_1 = n + 1$. По лемме 2.1 в этом графе существует по меньшей мере одна вершина степени меньшей или равной пяти. Зафиксируем одну из таких вершин x_{i_0} . Будем различать три случая.

Случай 1. Пусть $\text{st } x_{i_0} \leq 4$. Рассмотрим граф G' , полученный из G удалением вершины x_{i_0} и инцидентных ей ребер. Число ребер такого графа, очевидно, меньше n . Применим индукционное предположение, согласно которому этот граф 5-хроматичен. Рассмотрим какую-либо правильную раскраску G' с помощью пяти красок.

Возвращаясь к графу G , видим, что раскраска графа G' задает раскраску графа G , в которой, однако, не закрашена вершина x_{i_0} . Но эта вершина соединена ребрами не более чем с четырьмя уже закрашенными вершинами. Имеется возможность закрасить ее так, чтобы не нарушить правильность раскраски. В предположениях этого случая индукционный шаг завершен.

Случай 2. Пусть $\text{st } x_{i_0} = 5$, и не все соседние ребра, выходящие из x_{i_0} , образуют треугольники (рис. 57).

В этом случае склеим две вершины, смежные с x_{i_0} , инцидентные двум соседним ребрам и не соединенные

между собой ребром (на рисунке $x_{i'}$ и $x_{i''}$). Применяя преобразование 2), получим из получившегося мультиграфа простой граф, в котором число ребер меньше или равно n . Рассмотрим существующую по предположению индукции

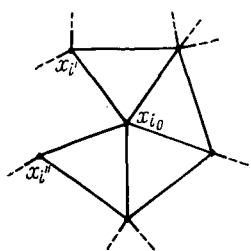


Рис. 57

правильную раскраску полученного графа. Она индуцирует некоторую раскраску графа, в которой вершины, ранее склеенные ($x_{i'}$ и $x_{i''}$) имеют, однако, одинаковую окраску. Полученная раскраска будет правильной, так как эти вершины в графе непосредственно не связаны ребром, и в предположениях случая 2 индукционный шаг также завершен.

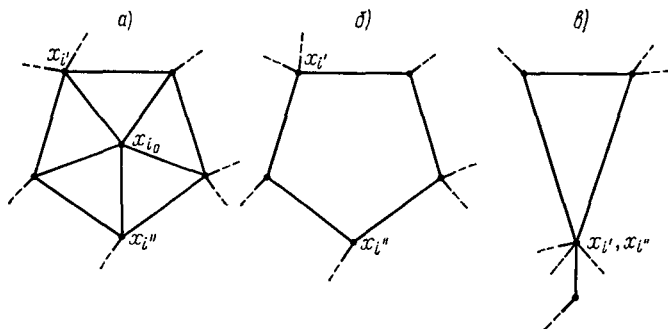


Рис. 58

Случай 3. Если $\text{st } x_{i_0} = 5$ и вершина x_{i_0} есть центр пятиугольника (рис. 58, а). В этом случае обязательно должны существовать две вершины, не соединенные непосредственно ребром, из числа пяти, окружающих x_{i_0} (на рис. $x_{i'}$ и $x_{i''}$), так как в противном случае эти пять вершин образовывали бы второй из графов, упомянутых в теореме 1.2 гл. V, и, следовательно, граф не был бы плоским. Удалим вершину x_{i_0} с инцидентными ей ребрами (рис. 58, б) и склеим $x_{i'}$ и $x_{i''}$ (рис. 58, в). Число ребер полученного графа меньше n . Рассмотрим его правильную раскраску с помощью пяти красок. Она индуцирует правильную (так как $x_{i'}$ и $x_{i''}$ не соединены ребром непосредственно) раскраску графа G , в которой, однако, не раскра-

шена вершина x_{i_0} . Заметив, что из пяти вершин, соединенных с x_{i_0} непосредственно, две (x_{i_1} и x_{i_n}) раскрашены одинаково, выводим отсюда, что эта раскраска может быть дополнена до правильной раскраски графа G . В предположениях случая 3 индукционный шаг завершен.

Приведем теперь пример плоского графа, для правильной раскраски которого недостаточно трех красок. Такой граф изображен на рис. 59. Доказательство того, что для его раскраски недостаточно трех красок, вполне тривиально и предоставляется читателю.

2. Необходимые и достаточные условия 1- и 2-хроматичности. Необходимые и достаточные условия даются следующей теоремой.

Теорема 2.2. *Граф G 1-хроматичен тогда и только тогда, когда он не содержит ни одного ребра. Граф G 2-хроматичен тогда и только тогда, когда он не содержит циклов нечетной длины.*

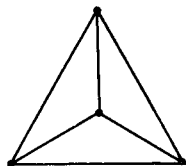


Рис. 59

Доказательство. Первое утверждение теоремы вполне очевидно. Для доказательства второго, очевидно, можно ограничиться случаем связного графа G .

Рассмотрим связный граф G и рассмотрим частичное дерево этого графа, полученное удалением из него ребер $e_1, \dots, e_n$. Очевидно, что оно может быть раскрашено двумя красками. Если в начальном графе G был некоторый цикл нечетной длины, то концы того из ребер e_{i_0} , которое участвует в этом цикле, окажутся раскрашенными одинаково, ибо путь по дереву, соединяющий эти две вершины, состоит из четного количества ребер. Таким образом, если бы граф G мог быть правильно раскрашен двумя красками, то эта раскраска была бы правильной и для частичного дерева; но тогда концы ребра e_{i_0} были бы раскрашены одинаково, что невозможно.

Обратно, если граф не содержит циклов нечетной длины, то для любой правильной раскраски частичного дерева концы любого из ребер e_i закрашены по-разному, а следовательно, индуцированная раскраска графа G правильна. ■

Глава VII

ЯДРО ГРАФА.

ПРИМЕНЕНИЕ ГРАФОВ В ТЕОРИИ ИГР

Строгое определение дискретной игры (с полной информацией) существенным образом связано с понятием графа. Здесь кратко рассматривается связь некоторых характеристик графа, например, ядра графа с теорией игр.

Для определения и изучения свойств ядра графа вводятся такие понятия, как «внешне устойчивое множество» и «внутренне устойчивое множество» графа и «база» графа, которые представляют самостоятельный интерес и находят применение в различных конкретных задачах.

§ 1. ЯДРО ГРАФА И ЕГО СВОЙСТВА

1. Определения устойчивых множеств. Ядро графа.

Определение 1.1. Множество $S \subseteq X$ называется ядром графа $G = (X, \Gamma)$, если выполнены следующие два условия, что

$$1) \forall x \in S \text{ справедливо равенство } \Gamma x \cap S = \emptyset, \quad (1.1)$$

$$2) \forall x \notin S \text{ справедливо соотношение } \Gamma x \cap S \neq \emptyset. \quad (1.2)$$

Из условия (1.1) вытекает, что ядро S не содержит вершин, в которых имеется петля, а из условия (1.2) — что S содержит все такие вершины, для которых $\Gamma x = \emptyset$; пустое множество не может быть ядром.

Условие (1.1) выделяет из всех подмножеств множества X внутренне устойчивые множества, т. е. такие множества $J \subseteq X$, любые две вершины которых не смежны.

Определение 1.2. Число $\alpha(G) = \max_{J \in \Sigma} |J|$, где Σ — семейство всех внутренне устойчивых множеств, называется числом внутренней устойчивости графа.

Определение 1.3. Множество $e \subset X$, обладающее тем

свойством, что для любой вершины $x \notin \varepsilon$ выполняется условие

$$\Gamma(x) \cap \varepsilon \neq \emptyset,$$

называется внешне устойчивым, а число

$$\beta(G) = \min_{\varepsilon \in \Gamma} |\varepsilon|,$$

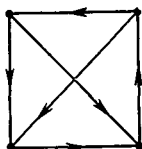


Рис. 60

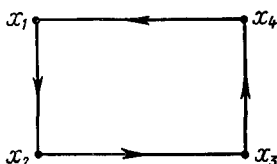


Рис. 61

где Γ — семейство всех внешне устойчивых множеств, — числом внешней устойчивости графа.

Из сказанного следует, что ядро графа S представляет собой множество, устойчивое как внутренне, так и внешне.

Числа внутренней $\alpha(G)$ и внешней $\beta(G)$ устойчивости графа, имеющего ядро S , удовлетворяют соотношению

$$\alpha(G) = \max_{J \in \Sigma} |J| \geq |S| \geq \min_{\varepsilon \in \Gamma} |\varepsilon| = \beta(G).$$

Как показывают примеры, не всякий граф обладает ядром. Например, граф G_1 (рис. 60) ядра не имеет, так как $\alpha(G_1) = 1 < 2 = \beta(G_1)$. С другой стороны, у графа может быть несколько ядер. Так, граф, изображенный на рис. 61, имеет два различных ядра: $S_1 = \{x_1, x_3\}$, $S_2 = \{x_2, x_4\}$.

Понятия внутренне и внешне устойчивого множества удобно проиллюстрировать на примерах, связанных с шахматной игрой.

Пример 1. Можно ли на шахматной доске расставить восемь ферзей так, чтобы ни один из них не находился под ударом какого-либо другого?

Эта задача сводится к нахождению наибольшего внутренне устойчивого множества в графе $G = (X, \Gamma)$, который определяется следующим образом.

Множество вершин $X = \{x_1, x_2, \dots, x_{64}\}$ графа взаимно-однозначно соответствует множеству клеток шахматной доски, а отображение $\Gamma: X \rightarrow 2^X$ определяется условием: если $x_i \in X$, то Γx_i — множество тех и только тех вершин графа, которые соответствуют полям шахматной доски, находящимся под ударом ферзя, распо-

ложенного на клетке с номером i . Очевидно, что если $x_j \in \Gamma x_i$, то при перестановке ферзя на клетку с номером j поле i будет под его ударом, иными словами, построенный граф G -симметрический. Любые две клетки шахматной доски, не лежащие на одной горизонтали, вертикали или диагонали, определяют в графе G несмежные вершины. Только на таких двух клетках можно расставить два ферзя так, чтобы они не находились под ударом друг друга. Отсюда следует, что наибольшее количество попарно несмежных вершин в графе G , т. е. число внутренней устойчивости $\alpha(G)$, равно наибольшему количеству ферзей, размещенных на шахматной доске так, что ни один из них не находится под ударом другого.

Отметим, что ответ на вопрос, поставленный в задаче, можно получить без привлечения теории графов, причем задача допускает 92 решения.

Задача 1. Показать, что для графа задачи, рассмотренной в примере 1 (о восьми ферзях), число внутренней устойчивости $\alpha(G) = 8$.

Пример 2. Определить число ферзей, которых достаточно расставить на шахматной доске так, чтобы каждая клетка доски находилась под ударом хотя бы одного из них? При этом считается, что клетка, занимаемая фигурой, находится под ударом.

Эта задача эквивалентна задаче нахождения наименьшего внешне устойчивого множества в графе $G = (X, \Gamma)$, рассмотренного в предыдущем примере. В самом деле, условие, определяющее внешне устойчивое множество $\varepsilon: x_j \notin \varepsilon \Rightarrow \Gamma x_j \cap \varepsilon \neq \emptyset$ в силу симметричности графа означает, что найдется такая вершина $x_k \in \varepsilon$, что $\Gamma x_k = x_j$, т. е. при положении ферзя на клетке с номером k ($x_k \in \varepsilon$) поле с номером j находится под его ударом (см. пример 1).

Задача 2. Показать, что число внешней устойчивости $\beta(G)$ графа G из примера 2 равно 5.

2. Внешне устойчивое множество. Прежде чем сформулировать общие условия существования ядра для произвольного графа, укажем алгоритм построения внешне устойчивого множества. Алгоритм построения внутренне устойчивого множества для графов специального вида, так называемых двудольных, будет рассмотрен в гл. IX.

Предварительно рассмотрим конструкцию, связанную с внешне устойчивым множеством графа, необходимую для обоснования алгоритма. По заданному графу $G = (X, \Gamma)$ без петель построим граф $G_1 = (X \cup \bar{X}, \Gamma_1)$, где $\bar{X} = \{\bar{x}_1, \bar{x}_2, \dots, \bar{x}_n\}$ — второй экземпляр множества $X = \{x_1, x_2, \dots, x_n\}$, а отображение $\Gamma_1: X \rightarrow 2^{\bar{X}}$ определим следующим условием:

$$\bar{x}_i \in \Gamma_1 x_j \text{ тогда и только тогда, когда } i=j \text{ или } \Gamma x_i \ni x_j. \quad (1.3)$$

Лемма 1.1. Если ε — внешне устойчивое множество графа $G(X, \Gamma)$, то $\Gamma_1 \varepsilon = \bar{X}$, и наоборот, если $\Gamma_1 \varepsilon = \bar{X}$, то множество ε внешне устойчиво.

Доказательство. Пусть ε — внешне устойчивое множество; покажем, что $\Gamma_1 \varepsilon = \bar{X}$. Действительно, если x_j — произвольно выбранная вершина множества X , то или $x_j \notin \varepsilon$ и в этом случае из определения (1.3) следует, что существует вершина $x_k \in \varepsilon$ такая, что $\Gamma x_j \ni x_k$; или $x_j \in \varepsilon$. Из условия (1.3) в первом случае получаем, что вершина $\bar{x}_j \in \Gamma_1 x_k$ ($x_k \in \varepsilon$), во втором случае — $\bar{x}_j \in \Gamma_1 x_j$ ($x_j \in \varepsilon$). Таким образом, и в том и другом случае для произвольной вершины $\bar{x}_j \in \bar{X}$ существует вершина $\bar{x}$ из ε ($x = x_k$, или $x = x_j$), такая, что $\Gamma_1 x \ni \bar{x}_j$, т. е. $\Gamma_1 \varepsilon = \bar{X}$.

Покажем, что в случае, когда $\Gamma_1 \varepsilon = \bar{X}$, множество ε — внешне устойчивое. Действительно, пусть x_j — произвольная вершина, не принадлежащая множеству ε , $x_j \notin \varepsilon$; тогда по условию леммы в графе G_1 для вершины $\bar{x}_j$ найдется такая вершина $x_k \in \varepsilon$, что $\Gamma_1 x_k \ni \bar{x}_j$. Из этого равенства, на основании условия (1.3), следует, что $\Gamma x_j \ni x_k$, $x_k \in \varepsilon$, $x_j \notin \varepsilon$, т. е. пересечение $\Gamma x_j \cap \varepsilon \neq \emptyset$, и по определению (1.3), множество ε — внешне устойчивое. ■

Для описания алгоритма построения внешне устойчивого множества удобно ввести следующие определения.

Определение 1.4. Вершина $\bar{x}_i \in \bar{X}$ в графе $G_1 = (X \cup \bar{X}, \Gamma_1)$ называется *висячей*, если $\Gamma_1^{-1} \bar{x}_i = \{x_j\}$, $x_j \in X$.

Определение 1.5. Граф $G_1 = (X \cup \bar{X}, \Gamma_1)$ называется *неприводимым*, если 1) для любой пары вершин $x, y \in X$ не выполняется включение $\Gamma_1 x \subseteq \Gamma_1 y$; 2) в графе G_1 нет висячих вершин.

Алгоритм построения наименьшего внешне устойчивого множества

1°. Удалить из X все такие вершины x , для каждой из которых существует вершина $y \in X$ ($x \neq y$) такая, что $\Gamma_1 x \subseteq \Gamma_1 y$. Из множества вершин $\{x_{i_1}, \dots, x_{i_k}\} \subseteq X$ таких, что $\Gamma_1 x_{i_1} = \Gamma_1 x_{i_2} = \dots = \Gamma_1 x_{i_k}$, оставить в X только одну.

2°. Удалить из $\bar{X}$ все висячие вершины.

3°. Удалить из X вершины x , смежные с висячими, и занести их в список T .

4°. Удалить из $\bar{X}$ вершины, смежные с вершинами, удаленными в 3°.

5°. К полученному графу применять п. 1° — 4° до тех пор, пока не прекратится удаление вершин.

(Процесс, описанный в п. 1⁰—5⁰, будем называть упрощением графа G_1 до неприводимого $G'_1 = (X' \cup \bar{X}', \Gamma'_1)$.)

6⁰. Если $\bar{X}' = \phi$, то вершины, занесенные в список T , образуют наименьшее внешне устойчивое множество.

Если $\bar{X}' \neq \phi$, то перейти к 7⁰.

7⁰. По списку T и графу $G'_1 = (X' \cup \bar{X}', \Gamma'_1)$ строить внешне устойчивые множества $\varepsilon_{j_1}(T), \varepsilon_{j_2}(T), \dots, \varepsilon_{j_m} \varepsilon_{j_{m-1}} \dots \varepsilon_{j_1}(T)$ в соответствии с п. 8⁰—10⁰.

8⁰. Удалить из X' произвольную вершину x_{i_1} , а из $\bar{X}'$ — все вершины $\bar{x} \in \Gamma'_1 x_{i_1}$ (полученный граф обозначить через $G'_{x_{i_1}}$). Составить список $\varepsilon_{j_1}(T)$, занеся в список T вершину x_{i_1} . Заменить в 3⁰ список T на $\varepsilon_{j_1}(T)$.

9⁰. Упростить граф $G'_{x_{i_1}}$ до неприводимого $\varepsilon_{j_1}(G')$, применяя п. 1⁰—5⁰.

10⁰. Применяя п. 7⁰—9⁰, строить графы

$\varepsilon_{j_1}(G'), \varepsilon_{j_2} \varepsilon_{j_1}(G'), \dots, \varepsilon_{j_m} \varepsilon_{j_{m-1}} \dots \varepsilon_{j_1}(G') (1 \leq j_1 \leq |X'|, \dots, 1 \leq j_m \leq |\varepsilon_{j_{m-1}} \dots \varepsilon_{j_1}(X')|)$ такие, что

$$\varepsilon_{j_k} \varepsilon_{j_{k-1}} \dots \varepsilon_{j_1}(\bar{X}') = \phi, \quad k \in (1, 2, \dots, m).$$

11⁰. Среди списков $\varepsilon_{j_k} \varepsilon_{j_{k-1}} \dots \varepsilon_{j_1}(T), k \in (1, 2, \dots, m)$ отвечающих графам $\varepsilon_{j_k} \varepsilon_{j_{k-1}} \dots \varepsilon_{j_1}(G')$, построенным в п. 10⁰, найти список с наименьшим числом элементов. Вершины графа $G_1 = (X \cup \bar{X}, \Gamma_1)$, внесенные в этот список, образуют наименьшее внешне устойчивое множество графа $G = (X, \Gamma)$.

Работу алгоритма проиллюстрируем для графа $G = (X, \Gamma)$, изображенного на рис. 62. Отвечающий ему граф $G_1 = (X \cup \bar{X}, \Gamma_1)$ представлен на рис. 63.

1⁰. В графе G_1 удалить вершины x_1 и x_6 , поскольку $\Gamma_1 x_1 \subset \Gamma_1 x_2$, $\Gamma_1 x_6 \subset \Gamma_1 x_4$. Получившийся граф изображен на рис. 64.

2⁰. Вершина $\bar{x}_5$ висячая. Удаляя вершины $\bar{x}_5, x_5$, а также вершины $\bar{x}_3, \bar{x}_4, (\Gamma x_5 = \{\bar{x}_3, \bar{x}_4, \bar{x}_5\})$, получим не-

приводимый граф G'_1 , представленный на рис. 65

$$|\bar{X}| \neq \emptyset, T = \{x_5\}.$$

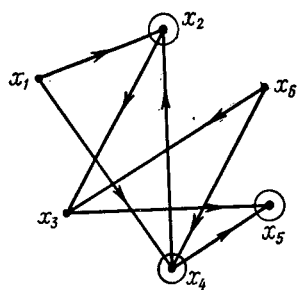


Рис. 62

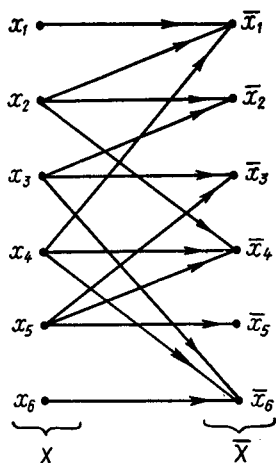


Рис. 63

3°. Удалим, например, вершину x_4 вместе с вершинами $\bar{x}_1, \bar{x}_6 \in \Gamma_1 x_4$. Составим список $\varepsilon_{j_1}(T) = \{x_5, x_4\}$ ($j_1 = 4$) и построим граф G'_4 (рис. 66).

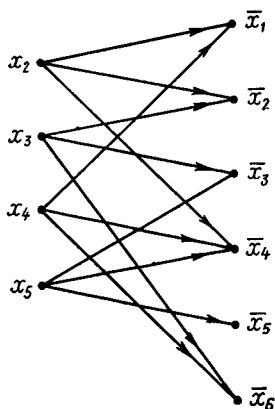


Рис. 64

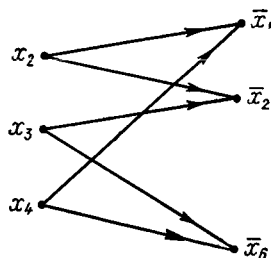


Рис. 65

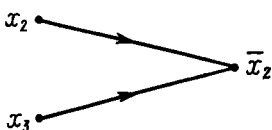


Рис. 66

4°. Удаляя в соответствии с п. 1° вершину x_3 и занося вершину x_2 в список, получаем внешне устойчивое множество $\{x_5, x_4, x_2\} = \varepsilon_4(T)$.

5°. Если вместо вершины x_4 удалить вершину x_2 или x_3 , списки $\varepsilon_2(T)$ и $\varepsilon_3(T)$ не будут содержать меньшего числа вершин. Так что множество $\{x_4, x_5, x_2\}$ — наименьшее внешне устойчивое множество.

Обоснование алгоритма построения внешне устойчивого множества

Ввиду конечности графа $G_1 = (X \cup \bar{X}, \Gamma_1)$ и того, что при применении правил алгоритма в п. 2°, 4°, 8° число вершин графа, остающихся в множестве $\bar{X}$, монотонно убывает. очевидно, что либо $\bar{X}' = \phi$, либо найдется такое целое неотрицательное число m , что

$$\varepsilon_{j_k} \varepsilon_{j_{k-1}} \dots \varepsilon_{j_1}(\bar{X}') = \phi \text{ при } 0 \leq k \leq m, 1 \leq j_1 \leq |\bar{X}'|,$$

$$1 \leq j_2 \leq |\varepsilon_{j_1}(\bar{X}')|, \dots, 1 \leq j_k \leq |\varepsilon_{j_{k-1}} \dots \varepsilon_{j_1}(\bar{X}')|.$$

Далее, заменим, что применение правил п. 1°—5° к графу G_1 приводит к построению множества X' и множества вершин в списке T таких, что выполняется условие

$$\Gamma_1 X' \cup \Gamma_1 T = \bar{X}, \quad X' \cap T = \phi. \quad (1.4)$$

В силу леммы 1.1 отсюда следует, что множество $T \cup X'$ — внешне устойчивое в графе G . Более того, из п. 1° и 3° алгоритма, определения 1.4 висячей вершины и условия (1.3) следует, что любое внешне устойчивое множество в графа G содержит все вершины T .

Если теперь выполняется условие $\bar{X}' = \phi$, из которого вытекает, в силу п. 3°, 4°, условие $X' = \phi$, то из равенства (1.4) получаем, что множество вершин списка T — наименьшее внешне устойчивое множество.

Если $\bar{X}' \neq \phi$, то списки вершин $\varepsilon_{j_k} \varepsilon_{j_{k-1}} \dots \varepsilon_{j_1}(T)$ при условии, что $\varepsilon_{j_k} \varepsilon_{j_{k-1}} \dots \varepsilon_{j_1}(\bar{X}') = \phi$, $k \in (1, 2, \dots, m)$ по тем же соображениям, что и выше, определяют внешне устойчивые множества графа G . Так как по построению (п. 8°) эти списки вершин являются расширением списка T , причем любое расширение множества вершин T до внешне устойчивого может быть получено путем применения пра-

вил алгоритма в п. 8⁰—10⁰, то наименьший из списков $\varepsilon_{j_k} \varepsilon_{j_{k-1}} \dots \varepsilon_{j_1}(T)$, $k \in (1, \dots, m)$ является наименьшим внешне устойчивым множеством.

3. Максимальность ядра графа. Для доказательства свойства максимальной ядра графа приведем определение.

Определение 1.5. Множество A называется максимальным множеством семейства множеств Σ , если из включения $B \supset A$, где $B \in \Sigma$, следует равенство $B = A$.

Теорема 1.1. Если S — ядро графа $G = (X, \Gamma)$, то множество S является максимальным элементом в семействе Σ всех внутренне устойчивых множеств.

Доказательство. Пусть A — внутренне устойчивое множество, содержащее ядро S , т. е. $A \supset S$. Тогда существует вершина $\tilde{a} \in A$, $\tilde{a} \notin S$. Так как S — внешне устойчивое множество, то пересечение $\Gamma\tilde{a} \cap S \neq \emptyset$, а следовательно, $\Gamma\tilde{a} \cap A \neq \emptyset$, что противоречит внутренней устойчивости множества $A \in \Sigma$. ■

Теорема 1.2. В симметрическом графе без петель каждое максимальное множество семейства Σ всех внутренне устойчивых множеств является ядром графа.

Доказательство. Пусть $S \in \Sigma$ — максимальное множество семейства Σ . Покажем, что S устойчиво внешне. Действительно, если $\Gamma x \cap S = \emptyset$ для некоторой вершины $x \notin S$, то множество $A = S \cup \{x\}$ в симметрическом графе без петель внутренне устойчиво. Тогда включение $A \supset S$ противоречит максимальнойности множества S в Σ . ■

Как следствие теорем 1.1 и 1.2 получаем следующий результат.

Теорема 1.3. Симметрический граф без петель обладает ядром.

Доказательство. Укажем способ построения множества S — максимального в множестве Σ всех внутренне устойчивых множеств и тем самым, в силу теоремы 1.2, ядра графа.

Возьмем произвольную вершину $x_0 \in X$ и положим $S_0 = \{x_0\}$; затем выберем некоторую вершину $x_1 \notin \Gamma x_0$ и положим $S_1 = \{x_0, x_1\}$, далее выберем вершину $x_2 \notin \Gamma S_1$ и т. д. Граф конечен, поэтому на некотором шаге получим множество S_n такое, что $\Gamma S_n = X$. Множество S_n — внутренне устойчивое по построению, и очевидно, что оно максимально в Σ , поэтому S_n — ядро графа. ■

4. База графа. Для доказательства следующей теоремы существования ядра понадобится понятие базы графа.

Определение 1.6. Множество $B \subset X$ называется базой графа $G = (X, \Gamma)$, если оно удовлетворяет двум условиям:

1) для любых двух вершин $b_1 \in B$, $b_2 \in B$, $b_1 \neq b_2$ в графе G не существует пути ни из b_1 в b_2 , ни из b_2 в b_1 ;

2) для каждой вершины $x \notin B$ существует такая вершина $b_x \in B$, что найдется путь из x в $b_x \in B$.

Например, граф, изображенный на рис. 67, имеет базу $B = \{b_1, b_2, b_3\}$.

З а м е ч а н и е. База графа может содержать один элемент. Тогда условие 1) считается выполненным по определению.

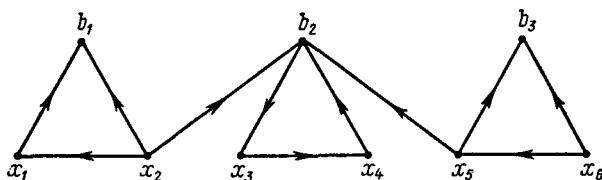


Рис. 67

Определение 1.7. Вершина x имеет строго последующую вершину y (вершина y строго следует за x), если

1) существует путь из x в y ;

2) x и y не лежат на одном контуре.

Так, на рис. 68 вершины x_5 , x_1 , x_2 , x_0 строго следуют за x_4 . Ни одна из вершин контура $\{x_0, x_1, x_2\}$ не имеет строго последующих.

Если вершина x графа $G = (X, \Gamma)$ является вершиной, имеющей строго последующие, то найдется вершина z_x , не имеющая строго последующих, и такая, что существует путь из x в z_x . Это утверждение очевидно в силу конечности графа и определения 1.7. Для графа, изображенного на рис. 69, для вершины x , например, вершина $x_j = z_x$, для вершины x_0 соответствующая вершина z_{x_0} определяется неоднозначно: $z_{x_0} = x_j$ или $z_{x_0} = x_n$.

Лемма 1.2. Всякий конечный граф $G = (X, \Gamma)$ имеет базу.

Д о к а з а т е л ь с т в о. Рассмотрим множество K всех вершин графа $G = (X, \Gamma)$, которые не имеют строго последующих. Это множество вершин K разобьем на по-

парно непересекающиеся группы, объединяя в одну группу те и только те вершины из K , которые лежат на одном контуре (в том числе контуре нулевой длины для изолированных вершин). В силу конечности графа G таких групп конечное число; обозначим их через

$$B_1, \dots, B_N : K = \bigcup_{i=1}^N B_i, B_i \cap B_j = \emptyset$$

и выберем в каждой группе B_i ($i = 1, 2, \dots, N$) по элементу $b_i \in B_i$. Утверждается, что множество $\{b_i (i = 1, 2, \dots, N)\} = \tilde{B}_0$ — база графа G . В самом деле, во-первых, если

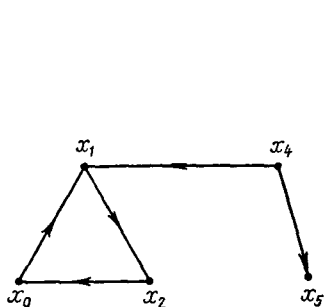


Рис. 68

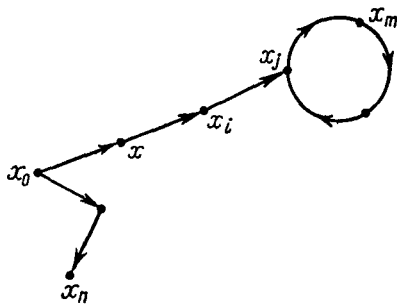


Рис. 69

$b_i \neq b_j$, то не существуют пути ни из b_i в b_j , ни из b_j в b_i , так как ни вершина b_i , ни b_j не имеют строго последующих, а с другой стороны, b_i, b_j не лежат на одном контуре, поскольку $b_i \in B_i, b_j \in B_j$, а пересечение $B_i \cap B_j = \emptyset$. Во-вторых, если $x \notin \tilde{B}_0$, то в случае $x \in K$, имеем $x \in B_i$ при некотором $i = 1, 2, \dots, N$, и поэтому существует путь из x в $b_i \in B_i$ (все элементы из B_i лежат на одном контуре); в случае, когда $x \notin K$, очевидно найдется путь из вершины x в вершину z_x , не имеющую строго последующих, и поэтому $z_x \in K$, т. е. $z_x \in B_{i_0}$ при некотором $i_0 \in (1, 2, \dots, N)$, а тогда существует путь из z_x в b_{i_0} , т. е. окончательно найдем путь из $x \notin K$ в $b_{i_0} \in \tilde{B}_0$. По определению 1.6, $\tilde{B}_0$ — база. ■

Введем на множестве K вершин графа $G = (X, \Gamma)$ отношение R между любыми двумя элементами $k_1 \in K, k_2 \in K$ следующим образом: $k_1 R k_2$ тогда и только тогда, когда k_1 и k_2 лежат на одном контуре.

Задача 3. Доказать, что отношение R есть отношение эквивалентности.

Рассмотренная в доказательстве леммы 1.2 конструкция построения разбиения $K = \bigcup_{i=1}^N B_i$ представляет собой разбиение множества K на классы смежности (по отношению R), а базу $\tilde{B}_0$ образуют представители всех классов смежности (по одному представителю из каждого класса).

Теорема 1.4. *Граф, не содержащий контуров нечетной длины, обладает ядром.*

Доказательство. Пусть $G = (X, \Gamma)$ — граф без контуров нечетной длины. Определим последовательно множества $Y_0, Y_1, \dots \subset X$ следующим образом:

1. Множество Y_0 — пусто. Пусть B_0 — база подграфа, порождаемого множеством $X \setminus Y_0$; эта база в силу леммы 1.2 существует. Положим $Y_1 = B_0 \cup \Gamma^{-1}B_0$.

Если множество Y_n определено, то обозначим через B_n какую-либо базу подграфа, порождаемого множеством $X \setminus Y_n$ и удовлетворяющую условию $B_n \in \Gamma^{-1}(\Gamma^{-1}B_{n-1} \setminus Y_{n-1})$. Нетрудно показать, что такая база существует. Полагаем $Y_{n+1} = Y_n \cup B_n \cup \Gamma^{-1}B_n$.

Тогда справедливы включения

$$\phi = Y_0 \subset Y_1 \subset \dots$$

Исходный граф конечен, поэтому при некотором m выполняется равенство $Y_m = X$, и множество

$$S = \bigcup_{n=0}^{m-1} B_n$$

является ядром графа.

Действительно, S внешне устойчиво, так как если $x \notin S$, то из построения S и определения базы B_n следует, что $x \in \Gamma^{-1}B_n \setminus Y_n$ при некотором $n \in (1, 2, \dots, m)$, а следовательно, $\Gamma x \cap B_n \neq \phi$ и $\Gamma x \cap S \neq \phi$.

Покажем, что S — внутренне устойчиво. Предположим, обратное, что существует две смежные вершины, принадлежащие S . Так как B_n $n \in (1, 2, \dots, m)$ — база некоторого подграфа, то эти вершины не могут принадлежать одному и тому же множеству B_n . Пусть одна из указанных вершин принадлежит B_n , а другая B_r ($r < n$, $r, n \in (1, 2, \dots, m)$). Тогда имеем $B_n \cap \Gamma^{-1}B_r = \phi$, ибо $\Gamma^{-1}B_r \subset Y_{r+1}$, $Y_{r+1} \subset Y_n$, $B_n \subset X \setminus Y_n \subset X \setminus Y_{r+1}$. С другой

стороны, $B_r \cap \Gamma^{-1} \bar{B}_n = \emptyset$, так как в противном случае можно было бы построить путь

$$\begin{aligned} \mu = \{ & (x_0, x_n), (x_n, y_{n-1}), (y_{n-1}, x_{n-1}), (x_{n-1}, y_{n-2}), \dots, \\ & (y_r, x_r) \}, x_0 \in B_r, x_0 \in \Gamma^{-1} B_n; \\ x_n \in B_n \subset & \Gamma^{-1} (\Gamma^{-1} B_{n-1} \setminus Y_{n-1}); y_{n-1} \in \Gamma^{-1} B_{n-1} \setminus Y_{n-1}; x_{n-1} \in B_{n-1}; \\ & \dots, y_r \in \Gamma^{-1} B_r \setminus Y_r; x_r \in B_r. \end{aligned}$$

Пусть μ соединяет вершину $x_0 \in B_r$ с вершиной $x_r \in B_r$, причем все остальные вершины пути также лежат в $X \setminus Y_r$. Так как B_r — база подграфа, порожденного вершинами $X \setminus Y_r$, то $x_0 = x_r$. Таким образом, путь μ представляет собой контур нечетной длины, равной $2(n-r)-1$, что противоречит условию теоремы. ■

§ 2. ГРАФЫ В ТЕОРИИ ИГР

1. Игра «ним». На графе естественным образом можно определить некоторую игру для двух игроков A и B , называемую игрой «ним».

Пусть (X, Γ) — конечный граф. Рассмотрим игру, положениями которой служат вершины графа. Начальная вершина x_0 выбирается жеребьевкой, и противники играют поочередно: сперва игрок A выбирает вершину x_1 в множестве Γx_0 и затем B выбирает вершину x_2 в множестве Γx_1 и т. д. Если один из игроков выбрал вершину x_k , для которой $\Gamma x_k = \emptyset$, то партия на этом заканчивается и этот игрок считается выигравшим.

Введенное в § 1 понятие ядра графа позволяет охарактеризовать все выигрышные положения игры, т. е. те вершины графа, выбор которых обеспечивает выигрыш партии независимо от ответов противника. Основным результатом является следующая теорема.

Теорема 2.1. *Если граф (X, Γ) игры «ним» имеет ядро и если один из игроков выбрал вершину в ядре, то этот выбор обеспечивает ему выигрыш или ничью.*

Д о к а з а т е л ь с т в о. Действительно, если игрок A выбрал вершину $x_1 \in S$, то либо $\Gamma x_1 = \emptyset$, тогда он выиграл партию, либо его противник B обязан выбрать вершину x_2 в множестве $X \setminus S$, а значит следующим ходом игрок A сможет выбрать вершину $x_3 \in S$ и т. д. Если в какой-либо момент один из игроков выиграл, выбрав вершину x_m , для

которой $\Gamma x_m = \phi$, то $x_m \in S$, и выигравшим партнером необходимо является A .

2. Общее определение игры. Ограничимся играми с двумя игроками A и B .

Определение 2.1. *Игра определена, если задано:*

1) множество Y , элементы которого называются положениями игры;

2) отображение P множества Y в множество 2^Y всех подмножеств множества Y , $P: Y \rightarrow 2^Y$, называемое правилом игры;

3) отображение Δ множества Y в множество $\{0, 1, 2\}$, называемое кодом, причем $\Delta(y) = 0$ тогда и только тогда, когда $P_y = \phi$; при $\Delta(y) = 1$ очередной ход в положении y должен сделать игрок A , при $\Delta(y) = 2$ — игрок B ;

4) две вещественные функции $f_A(y)$ и $f_B(y)$, определенные на Y , называемые функциями предпочтения соответственно для игроков A и B .

Считается, что оба игрока в каждый момент времени знают положение игры $y \in Y$, поэтому такую игру называют игрой с полной информацией.

Партия осуществляется следующим образом. Начальное положение $y_0 \in Y$ выбирается по жребию. Если, например, $\Delta(y_0) = 1$, то первый ход делает игрок A , причем он должен выбрать $y_1 \in P_{y_0}$; затем игрок, который должен сделать следующий ход (игрок A , если $\Delta(y_1) = 1$, или игрок B , если $\Delta(y_1) = 2$), выбирает положение y_2 в множестве P_{y_1} , и так далее.

Если один из игроков выбрал такое положение y_h , для которого $P_{y_h} = \phi$, то партия заканчивается.

Свяжем с каждым из игроков один из эпитетов: активный или пассивный. Если $S \subset Y$ — множество положений, встретившихся в течение партии, то выигрыш игрока A , по определению, есть:

$$f_A(S) = \sup_{y \in S} f_A(y), \text{ если игрок } A \text{ активный,}$$

$$f_A(S) = \inf_{y \in S} f_A(y), \text{ если игрок } A \text{ пассивный.}$$

Аналогично, с помощью функции $f_B(y)$ определяется выигрыш игрока B . Если $f_A(S) > f_B(S)$, то игрок A считается выигравшим; если $f_A(S) < f_B(S)$ — проигравшим.

Таким образом, можно сказать, что игра представляет собой абстрактное множество Y с заданной на нем «структурой игры», определяемой:

- 1) структурой графа (Y, P) на множестве Y ,
- 2) вектор-функцией $(\Delta(y), f_A(y), f_B(y))$, заданной на Y .

Рассмотрим два примера, иллюстрирующих введенное определение. Пусть задана на графе $G = (X, \Gamma)$ игра «ним», рассмотренная в п. 1. Тогда множество положений игры определяется так:

$$Y = X \times \{1, 2\} = \{(x, i) \mid x \in X, i \in \{1, 2\}\}.$$

Так как игроки A и B делают ходы поочередно, то отображения P и Δ задаются соотношениями:

$$P(x, 1) = \Gamma x \times \{2\}, \quad P(x, 2) = \Gamma x \times \{1\},$$

$$\Delta(x, i) = i, \text{ если } \Gamma x \neq \emptyset; \quad \Delta(x, i) = 0, \text{ если } \Gamma x = \emptyset.$$

Функции $f_A(y)$ и $f_B(y)$ определяются следующим образом:

$$f_A(x, 2) = f_B(x, 1) = 1 \text{ при } \Gamma x = \emptyset$$

$$f_A(x, 2) = f_B(x, 1) = 0 \text{ в остальных случаях.}$$

Оба игрока считаются активными.

Второй пример. Два игрока A и B играют в шахматы.

Пусть $t_f \in \{1, 2, \dots, 65\}$ — номер клетки шахматной доски занимаемой фигурой f , причем $t_f = 65$ означает, что f отсутствует на доске, тогда вектор $x = (t_1, \dots, t_f, \dots)$ называют *диаграммой положения*. Положение игры $y = (x, i)$, где индекс i , равный 1 или 2, соответствует очередному ходу в положении x , который делает игрок A или B соответственно. Оба игрока активны; $f_A(x) = 1$, когда x — матовое положение и $\Gamma(x) = \emptyset$, $f_A(x) = 0$ во всех остальных случаях. Соответственно определяется и функция $f_B(x)$ для игрока B .

Глава VIII

ТРАНСПОРТНЫЕ СЕТИ

Теория транспортных сетей возникла при решении задач, связанных с организацией перевозки грузов. Тем не менее, понятие потока на транспортной сети, алгоритм нахождения потока наибольшей величины и критерий существования потока, насыщающего выходные дуги сети, оказались плодотворными для многих других прикладных и теоретических вопросов комбинаторного характера.

§ 1. ПОТОКИ НА ТРАНСПОРТНЫХ СЕТЯХ

1. Основная задача теории транспортных сетей. Введем основные понятия теории.

Определение 1.1. *Транспортная сеть T есть совокупность двух объектов: 1. Связного графа $G = (X, \Gamma)$ со свойствами 1°) в графе G отсутствуют петли, 2°) в графе G существует одна и только одна вершина x_0 такая, что множество $\Gamma^{-1}x_0 = \emptyset$, 3°) в графе G существует одна и только одна вершина z , такая, что множество $\Gamma z = \emptyset$.*

2. Целочисленной неотрицательной функции $c(u)$, заданной на множестве Γ дуг графа G .

Вершина x_0 называется *входом сети*, вершина z — *выходом*. Значение функции $c(u)$ на дуге u называется *пропускной способностью дуги*.

Определение 1.2. Пусть U_x^- — множество дуг, заходящих в вершину x , а U_x^+ — множество дуг, выходящих из вершины x . Целочисленная неотрицательная функция $\varphi(u)$, заданная на множестве Γ дуг графа G , называется *поток*ом, если она удовлетворяет следующим свойствам:

$$1) \sum_{u \in U_x^-} \varphi(u) = \sum_{u \in U_x^+} \varphi(u) \quad (x \neq x_0, x \neq z),$$

$$2) \varphi(u) \leq c(u).$$

Термины, входящие в определения 1.1 и 1.2, наводят на мысль, что при рассуждениях относительно транспортных сетей очень удобно представлять, что по дугам транспортной сети движется некоторая несжимаемая «жидкость». Дуги могут пропускать «жидкость» только в одном направлении и в количестве, не превышающем их пропускной способности. Свойство 1) определения 1.2 можно понимать как закон сохранения количества жидкости. Вся жидкость, движущаяся по транспортной сети, вытекает из входа сети и стекает в выход. Сколько жидкости поступает из входа сети, столько же стекает в выход, так как из свойства 1) определения 1.2 следует равенство

$$\sum_{u \in U_{x_0}^+} \varphi(u) = \sum_{u \in U_z^-} \varphi(u), \quad (1.1)$$

Определение 1.3. Величина $\sum_{u \in U_z^-} \varphi(u)$ называется величиной потока φ и обозначается φ_z .

Задача. На данной транспортной сети T построить поток наибольшей величины.

Прежде чем изложить алгоритм решения задачи, введем еще два термина. Дуга u называется *насыщенной*, если $\varphi(u) = c(u)$. Поток $\varphi(u)$ называется *полным*, если каждый путь из x_0 в z содержит хотя бы одну насыщенную дугу.

2. Алгоритм Форда — Фалкерсона для нахождения потока наибольшей величины

1°. Перенумеровать произвольным образом вершины сети T , отличные от x_0 и z .

2°. Построить произвольный поток φ на транспортной сети T (например, положить $\varphi(u) = 0$, $\forall u \in \Gamma$).

3°. Просмотреть пути, соединяющие вход сети x_0 с выходом z . Если поток φ полный — перейти к пункту

4°. В противном случае рассмотреть путь μ , соединяющий x_0 с z , все дуги которого не насыщены. Построить новый поток φ' :

$$\varphi'(u) = \begin{cases} \varphi(u), & \text{если } u \notin \mu, \\ \varphi(u) + k, & \text{если } u \in \mu, \end{cases}$$

где $k = \min_{u \in \mu} (c(u) - \varphi(u))$. Повторить этот процесс до получения полного потока φ .

4⁰. Присвоить целочисленные метки вершинам сети T и знаки «+» или «—» дугам по правилам а) входу x_0 присвоить метку 0; б) если вершина x_i получила некоторую метку, а y — еще непомеченная вершина, то вершине $y \in \Gamma x_i$, такой что $\varphi((x_i, y)) < c((x_i, y))$ присвоить метку i , а дуге (x_i, y) — знак «+»; вершине $y \in \Gamma^{-1} x_i$, такой что $\varphi((y, x_i)) > 0$, присвоить метку i , а дуге (y, x_i) — «—». Остальные непомеченные вершины и дуги метки и знака не получают;

в) повторить процесс, описанный в пункте 4⁰ б) до тех пор, пока не прекратится появление новых отмеченных вершин и дуг. Если в результате процесса 4⁰ б) вершина z не получит метки, то поток обладает наибольшей величиной. В противном случае перейти к пункту 5⁰.

5⁰. Рассмотреть последовательность отмеченных вершин $\lambda = (z, x_{i_1}, x_{i_2}, \dots, x_0)$, каждая из которых имеет метку, равную номеру последующей вершины, и последовательность дуг μ (не обязательно путь), соединяющих последовательные вершины из λ . Построить новый поток φ' :

$$\varphi'(u) = \begin{cases} \varphi(u), & \text{если } u \notin \mu, \\ \varphi(u) + 1, & \text{если } u \in \mu \text{ и имеет знак «+»,} \\ \varphi(u) - 1, & \text{если } u \in \mu \text{ и имеет знак «—».} \end{cases}$$

Перейти к пункту 4⁰.

2. Соотношение между величиной потока и пропускной способностью разреза сети. Введем новые понятия теории транспортных сетей.

Определение 1.4. Пусть $A \subset X$ такое множество вершин графа, что $x_0 \notin A$, $z \in A$. Множество U_A^- дуг, входящих в A , т. е. соединяющих вершины $y \notin A$ с вершинами $x \in A$, называется разрезом сети T .

Определение 1.5. Пропускной способностью $c(U_A^-)$ разреза U_A^- называется сумма пропускных способностей дуг, входящих в разрез,

$$c(U_A^-) = \sum_{u \in U_A^-} c(u).$$

Лемма 1.1. Для любого потока φ и любого разреза U_A^- справедливо соотношение $\varphi_z \leq c(U_A^-)$.

Доказательство. В силу того, что выход сети $z \in A$, для величины потока φ_z справедливы соотношения

$$\varphi_z = \sum_{u \in U_A^-} \varphi(u) - \sum_{u \in U_A^+} \varphi(u) \leq \sum_{u \in U_A^-} c(u) = c(U_A^-). \blacksquare$$

Следствие. Если для некоторого потока φ и некоторого разреза U_A^- выполняется равенство $\varphi_z = c(U_A^-)$, то поток φ обладает наибольшей величиной.

Лемма 1.1 и следствие из нее необходимы для обоснования рассмотренного алгоритма.

Обоснование алгоритма

Прежде всего заметим, что реализация алгоритма состоит из конечного числа шагов. В самом деле, п. 3⁰ может применяться лишь конечное число раз, так как на каждом шаге величина потока увеличивается по крайней мере на единицу. Вместе с тем величина любого потока не может превзойти суммарной пропускной способности дуг, инцидентных выходу сети z , $\sum_{u \in U_z^-} c(u)$.

Процесс присвоения меток в силу условия, что каждый раз получают метку еще неотмеченные вершины, конечен. И, наконец, в п. 5⁰ поток φ' обладает большей величиной, чем φ . Это вытекает из того, что по определению вершины z и правил п. 4⁰ б) дугам, входящим в z , может быть присвоен только знак «+».

Получаемая по правилу 3⁰ функция φ' — поток. Это непосредственно следует из того, что μ — путь.

В соответствии с правилом п. 5⁰ строится также поток. Чтобы доказать это утверждение, рассмотрим три соседние вершины последовательности λ : $x_{i_{k-1}}$, x_{i_k} , $x_{i_{k+1}}$. В силу правила 4⁰ б) возможны только ситуации, представленные на рис. 70. Но в этих ситуациях φ' — поток.

Пусть процесс, описанный в п. 4⁰ б), приводит к тому, что вершина z не получает метки. Обозначим через A множество непомеченных вершин. В силу того, что $x_0 \notin A$, $z \in A$ множество A определяет разрез сети U_A^- . Каждая дуга $u \in U_A^-$ соединяет помеченную вершину y с непомеченной $x \in A$. Вершина x может остаться непомеченной при условии, что $\varphi((y, x)) = c((y, x))$. Аналогично, на каж-

дой дуге $u \in U_A^+$ (выходящей из A) выполняется равенство $\varphi(u) = 0$, поэтому справедливы соотношения

$$\varphi_z = \sum_{u \in U_A^-} \varphi(u) - \sum_{u \in U_A^+} \varphi(u) = \sum_{u \in U_A^-} c(u) = c(U_A^-).$$

В силу леммы 1.1 это означает, что поток φ обладает наибольшей величиной.

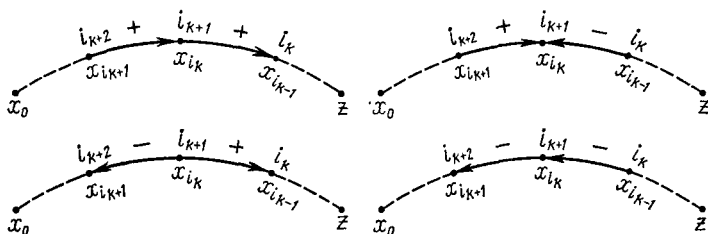


Рис. 70

Проведенное рассуждение совместно с леммой 1.1 составляют доказательство следующей теоремы.

Теорема 1.1. Для заданной транспортной сети величина наибольшего потока равна наименьшей пропускной способности разрезом, т. е.

$$\max_{\varphi} \varphi_z = \min_{U_A^-} c(U_A^-).$$

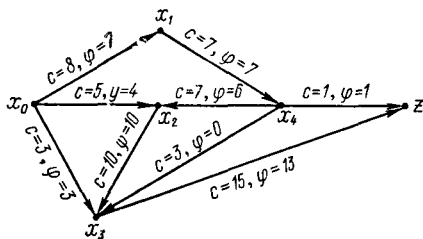


Рис. 71

В качестве примера применения алгоритма Форда—Фалкерсона рассмотрим транспортную сеть T и полный поток φ , для которого $\varphi_z = 14$ (рис. 71). Применяя правила 4⁰ и 5⁰ алгоритма, можно получить поток с $\varphi_z = 15$.

1. Задачи об «источниках» и «потребителях». В задачах этого типа встречаются три рода объектов: «источники», «потребители», «промежуточные пункты». Источниками могут быть электростанции, заводы, контрольно-измерительные приборы, железнодорожные станции и т. д., которые поставляют соответственно электроэнергию, изделия, информацию, грузы и т. д. для потребителей. В промежуточных пунктах может происходить, например, преобразование электроэнергии, упаковка или комплектация изделий, кодирование информации, перевалка грузов, причем, что весьма существенно, без изменения количества. Источники, промежуточные пункты и потребители связаны сетью линий передач или автомобильными, железнодорожными, морскими линиями или каналами связи с заданными пропускными способностями.

Пусть в некоторую единицу времени источники x_j ($j = 1, 2, \dots, k$) вырабатывают a_j единиц продукции, а потребность потребителя y_i ($i = 1, 2, \dots, m$) в этой продукции равна b_i . Пропускные способности линий считаются также отнесенными к выбранной единице времени. Таким образом, в задачах этого типа речь идет о неизменном и однородном по времени (стационарном) процессе. С каждой такой задачей свяжем транспортную сеть T в смысле определения 1.1 Вершинами сети T служат источники, потребители, промежуточные пункты и еще две вспомогательные вершины: вход x_0 и выход z . Вершину x_0 свяжем дугами (x_0, x_j) ($j = 1, 2, \dots, k$) с источниками и припишем им пропускные способности a_j . Потребителей y_i свяжем с выходом z дугами (y_i, z) с пропускными способностями b_i . Остальные вершины соединим между собой дугами в соответствии с реальным наличием связей между ними с соответствующими пропускными способностями.

Возникает задача о нахождении такого распределения потока энергии (грузов, информации) по линиям, чтобы в максимальной степени (желательно полностью) удовлетворить потребителей. Очевидно, что это задача об отыскании потока наибольшей величины на транспортной сети T . В § 3 будут установлены необходимые и достаточные условия, при которых существует поток, насыщающий выходные дуги, что соответствует в этих задачах полному удовлетворению нужд всех потребителей.

2. Динамические транспортные сети. Пусть требуется за заданное время t_0 перевезти максимальное количество грузов, первоначально сосредоточенных в городах $x_1, x_2, \dots, x_{N-1}$ в количествах $a_1, a_2, \dots, a_{N-1}$ в город x_N . Города $x_1, x_2, \dots, x_N$ связаны некоторой сетью железных до-

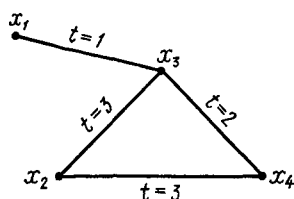


Рис. 72

рог. Обозначим через t_{ij} время, которое затрачивает поезд, чтобы достичь x_j из x_i , если x_i и x_j — соседние города, а через c_{ij} — пропускную способность связывающей их дороги; вообще говоря, $t_{ij} \neq t_{ji}$, $c_{ij} \neq c_{ji}$. Пусть c_{ii} — максимальное количество грузов, которые могут одновременно находиться на станции города x_i ($i < N$). Будем считать, что $t_0, t_{ij}, c_{ij}, c_{ii}$ — целые числа.

Введем множество

$$X = \{x_1, x_2, \dots, x_N\} \times \{0, 1, 2, \dots, t_0\} \cup \{x_0\} \cup \{z\},$$

где x_0, z — вход и выход транспортной сети, а элементы (x_j, t) декартова произведения $\{x_1, x_2, \dots, x_N\} \times \{0, 1, 2, \dots, t_0\}$ соответствуют городам x_j в момент времени t . Множеством вершин транспортной сети будет некоторое подмножество множества X . Формальное описание этого подмножества весьма громоздко, а суть дела легко усмотреть из рис. 72 и 73. На рис. 72 изображена схема дорог, связывающих города x_1, x_2, x_3, x_4 . Рис. 73 отвечает динамической транспортной сети для $t_0 = 5$.

Вершину (x_i, t) соединим дугами $((x_i, t), (x_i, t+1))$ пропускной способности c_{ii} ($i < N$) с вершинами $(x_i, t+1)$ и дугами $((x_i, t), (x_j, t+t_{ij}))$ пропускной способности c_{ij} с вершинами, соответствующими соседним с x_i городами x_j . Вход x_0 соединим дугами $(x_0, (x_i, 0))$ ($i < N$) пропускной способности a_i с вершинами $(x_i, 0)$. Вершины $(x_N, t_1), (x_N, t_1+1), \dots, (x_N, t_0)$, где $t_1 = \min_{1 \leq j \leq N-1} t_{jN}$, соединим с выходом дугами бесконечной пропускной способности. (Снова в сеть T могут войти не все описанные дуги (см. рис. 73)).

Если φ — некоторый поток на сети T , то значение потока φ на дуге $((x_i, t), (x_j, t+t_{ij}))$ будем понимать как количество груза, которое следует отправить из города

x_i в момент времени t на соседнюю станцию x_j , а величину $\varphi((x_i, t), (x_i, t + 1))$ — как количество груза, которое следует оставить на станции x_i . Ясно, что величина потока φ_z равна общему количеству груза, перевезенному на станцию x_N за время t_0 .

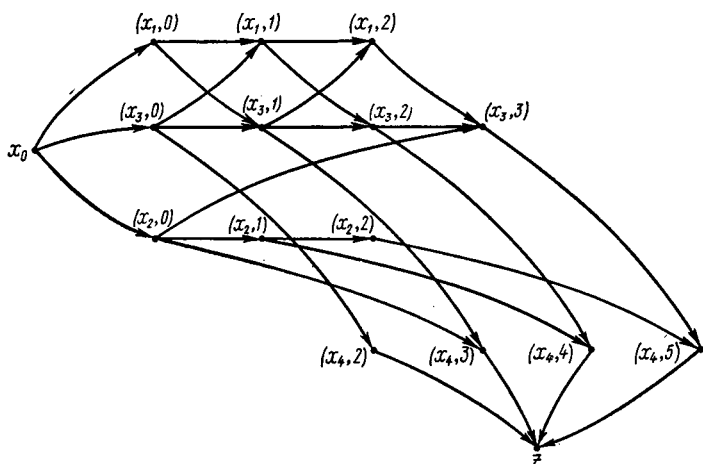


Рис. 73

3. Комбинаторные задачи. Рассмотрим задачу о назначении на должности. Пусть в некотором учреждении имеется 6 вакантных должностей $y_1, y_2, \dots, y_6$ и 6 работников $x_1, x_2, \dots, x_6$. Граф G , изображенный на рис. 74, иллюстрирует, какие должности y может в силу своей квалификации занимать работник x . Пусть выполнено условие: каждый работник может занимать хотя бы одну должность и на каждую должность претендует хотя бы один работник. Возможно ли произвести назначение на должности так, чтобы все шесть должностей заняли работники соответствующей квалификации?

Один из способов решения этой задачи состоит в рассмотрении

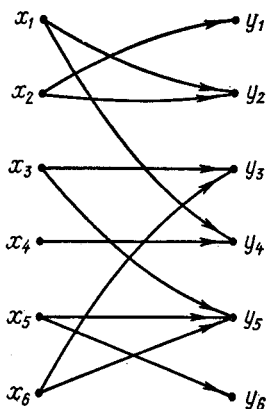


Рис. 74

вспомогательной транспортной сети T . Для ее получения добавим к множеству $X \cup Y$ вершин графа G еще две вершины: вход x_0 и выход z . Соединим x_0 с каждой вершиной x_j дугой пропускной способности $c((x_0, x_j)) = 1$ и каждую вершину y_i с z дугой пропускной способности $c((x_i, z)) = 1$. Припишем дугам исходного графа G бесконечные пропускные способности. Получим транспортную сеть T , изображенную на рис. 75.

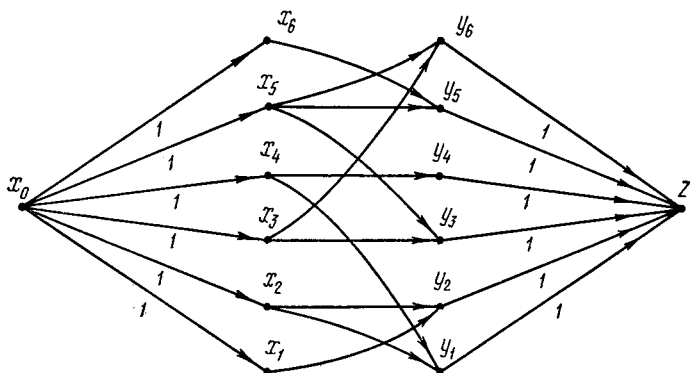


Рис. 75

Ясно, что если на сети T будет существовать поток такой, что $\varphi((y_j, z)) = 1$, т. е. насыщающий выходные дуги (одновременно он будет насыщать и входные дуги (x_0, x_i)), то требуемое назначение будет возможно произвести. Нужно назначить работника x_i на должность y_j в том и только том случае, когда $\varphi((x_i, y_j)) = 1$.

§ 3. ПОТОК, НАСЫЩАЮЩИЙ ВЫХОДНЫЕ ДУГИ

1. Потребность и поглощающая способность множества. Для того чтобы сформулировать условия существования на транспортной сети T потока, насыщающего выходные дуги, введем некоторые новые понятия.

Определение 3.1. Функция $d(y)$, определенная на вершинах транспортной сети T такая, что

$$d(y) = \begin{cases} c((y, z)), & \text{если } y \in \Gamma^- z, \\ 0, & \text{если } y \notin \Gamma^- z, \end{cases}$$

называется потребностью вершины y .

Определение 3.2. Функция $d(A)$ множеств $A \subset X$, определенная равенством

$$d(A) = \sum_{y \in A} d(y),$$

называется потребностью множества A .

Определение 3.3. Пусть множество $A \subset X$, T_A — транспортная сеть, полученная из исходной сети T соединением всех вершин $y \in A$ с выходом сети z дугами бесконечной пропускной способности. Тогда величина

$$F(A) = \max_{\varphi} \sum_{u \in U_A} \varphi(u), \quad (3.1)$$

где максимум берется по всем потокам на сети T_A , называется поглощающей способностью множества A .

Термин «поглощающая способность» становится более ясным, если учесть, что справедлива следующая лемма.

Лемма 3.1. Для любого множества $A \subset X$ существует поток φ_0 на сети T_A , обладающий свойствами:

$$1) F(A) = \sum_{u \in U_A} \varphi_0(u),$$

2) $\varphi_0(u) = 0$ для дуги $u \in U_A^+$ такой, что $u \neq (x, z)$, где $x \in A$.

Доказательство. Рассмотрим поток φ на сети T_A , для которого сумма $\sum_{u \in U_A} \varphi(u)$ принимает максималь-

ное значение. Пусть на некоторой дуге (x, y) , где $x \in A$, $y \notin A$, $y \neq z$ значение потока $\varphi((x, y)) > 0$. Построим новый поток φ' на сети T_A . Для этого положим $\varphi'((x, y)) = 0$, $\varphi'((x, z)) = \varphi((x, y)) + \varphi((x, z))$ и $\varphi'(u) = \varphi(u) - \varphi((x, y))$ на всех дугах, принадлежащих одному из путей, соединяющих вершину y с выходом сети z .

Проделав описанное построение для всех дуг $u \in U_A^+$, не являющихся выходными, получаем поток φ_0 , для которого справедливы утверждения леммы. ■

2. Основная теорема. Критерий существования потока, насыщающего выходные дуги, составляет содержание следующего утверждения.

Теорема (о насыщении) 3.1. Поток φ , насыщающий выходные дуги, существует тогда и только тогда, когда для

любого множества $Y \subseteq \Gamma^{-1}z$ справедливо соотношение между поглощающей способностью $F(Y)$ и потребностью $d(Y)$:

$$F(Y) \geq d(Y). \quad (3.2)$$

Доказательство теоремы будет приведено позднее. Применение критерия (3.2) рассмотрим на примере. Для транспортной сети T (рис. 76) подсчитаем величины $F(Y)$ и $d(Y)$.

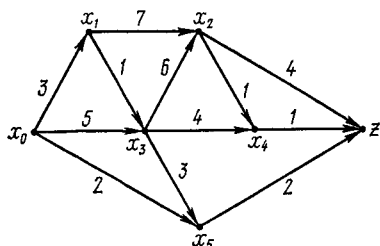


Рис. 76

Множество $\Gamma^{-1}z$ совпадает с множеством $\{x_2, x_4, x_5\}$. Чтобы решить вопрос о существовании потока, насыщающего выходные дуги, последовательно просматриваем семь подмножеств множества $\Gamma^{-1}z$:

$$Y_1 = \{x_2\}, \quad Y_2 = \{x_4\}, \quad Y_3 = \{x_5\}, \quad Y_4 = \{x_2, x_4\},$$

$$Y_5 = \{x_4, x_5\}, \quad Y_6 = \{x_2, x_5\}, \quad Y_7 = \{x_2, x_4, x_5\}.$$

Получаем следующую таблицу (табл. 1).

ТАБЛИЦА 1

Y	Y_1	Y_2	Y_3	Y_4	Y_5	Y_6	Y_7
d	4	1	2	5	3	6	7
F	8	5	5	8	9	10	10

Критерий (3.2) выполнен, следовательно, поток, насыщающий выходные дуги сети T , существует.

Лемма 3.2. Пусть A_0 — некоторое множество вершин сети T , не содержащее входа x_0 и выхода z ($A \subset X \setminus \{x_0, z\}$). Если для любого множества A такого, что $A_0 \subset A \subset X \setminus \{x_0, z\}$, справедливо неравенство

$$c(U_A^-) \geq d(A), \quad (3.3)$$

то для множества A_0 выполняется соотношение

$$F(A_0) \geq d(A_0). \quad (3.4)$$

Доказательство. По заданной транспортной сети T и множеству A_0 построим вспомогательную транспортную сеть T_1 . Сеть T_1 получается из сети T с помощью двух операций:

а) удаления всех дуг $u \in U_{A_0}^+$, а затем всех вершин и всех дуг, не принадлежащих никакому пути, ведущему из x_0 в A_0 (следовательно, выход z сети T всегда удаляется);

б) склеивания всех вершин множества A_0 в одну вершину z_0 . Очевидно, что вершина z_0 будет выходом сети T_1 (например, для сети (см. рис. 76) при выборе множества $A_0 = \{x_2, x_3, x_4\}$ сеть T_1 будет иметь вид, показанный на рис. 77).

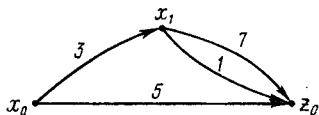


Рис. 77

Пусть ψ — поток наибольшей величины на сети T_1 . По потоку ψ можно построить поток φ_0 на сети T_{A_0} , описанный в лемме 3.1, положив

$$\varphi_0(u) = \begin{cases} \psi(u) & \text{для дуг, общих сетям } T_{A_0} \text{ и } T_1; \\ \sum \psi(y, x) & \text{для дуг вида } u = (x, z), \text{ где } x \in A_0, \\ & \text{а сумма берется по всем вершинам } y \in T_1 \\ & \text{таким, что } y \notin A_0, y \in \Gamma^{-1}x; \\ 0 & \text{в противном случае.} \end{cases}$$

По построению потока φ_0 справедливо соотношение

$$\psi_{z_0} = \sum_{u \in U_{A_0}^-} \varphi_0(u). \quad (3.5)$$

В силу того, что ψ — поток наибольшей величины, из (3.5) вытекает равенство

$$F(A_0) = \psi_{z_0}. \quad (3.6)$$

Для потока ψ по теореме 1.1 существует такой разрез U_A^- сети T_1 , что для пропускной способности разреза U_A^- выполняется равенство

$$c(U_A^-) = \psi_{z_0}. \quad (3.7)$$

Разрез U_A^- порождается некоторым множеством вершин A , общих для сетей T и T_1 , не содержащим входа x_0 и выхода z сети T и таким, что справедливо включение $A_0 \subset A$. Следовательно, при переходе от сети T_1 к сети T новых дуг

к множеству U_A^- не добавляется, и разрез U_A^- сети T_1 можно рассматривать как множество дуг сети T . Это позволяет воспользоваться условием (3.3) леммы и получить из (3.6) и (3.7) соотношения

$$F(A_0) = c(U_A^-) \geq d(A) \geq d(A_0). \quad (3.8)$$

В соотношении (3.8) второе неравенство есть следствие включения $A_0 \subset A$. ■

Следствие. Если неравенство (3.3) выполняется для любого множества $A \subset X \setminus \{x_0, z\}$, то для всех множеств $A \subset X \setminus \{x_0, z\}$ справедливо неравенство (3.4).

Лемма 3.3. Поток насыщающий дуги, инцидентные выходу z , существует в том и только том случае, если для любого множества вершин $A \subset X \setminus \{x_0, z\}$ выполняется неравенство (3.3).

Доказательство. Необходимость. Для любого множества $A \subset X \setminus \{x_0, z\}$ и для любого потока φ справедливо неравенства

$$F(A) \geq \sum_{u \in U_A^-} \varphi(u) \geq \sum_{y \in A} \varphi((y, z)). \quad (3.9)$$

Если существует поток φ , насыщающий выходные дуги, то из неравенств (3.9) и определения 3.1 вытекает неравенство

$$F(A) \geq d(A). \quad (3.10)$$

Поглощающая способность $F(A)$ любого множества A не превосходит пропускной способности разреза U_A^- , т. е.

$$F(A) \leq c(U_A^-), \quad (3.11)$$

поэтому из (3.10) и (3.11) вытекает утверждение леммы.

Достаточность. Допустим, что верно неравенство (3.3). Пусть U_B^- — произвольный разрез сети T . Рассмотрим множество $A = B \setminus \{z\}$. В силу следствия из леммы 3.2 и неравенства (3.3) имеем неравенство

$$d(A) \leq F(A). \quad (3.12)$$

Пропускные способности разрезов U_A^- и U_B^- связаны соотношениями

$$c(U_B^-) = c(U_A^-) + \sum_{x \in A} c((x, z)) = c(U_A^-) + d(X \setminus A). \quad (3.13)$$

Из (3.13), (3.12) и (3.11) вытекает цепочка неравенств, справедливая для любого разреза U_B^- :

$$c(U_B^-) \geq F(A) + d(X \setminus A) \geq d(A) + d(X \setminus A) = d(X) \quad (3.14)$$

Из (3.14) по теореме 1.1 для потока φ наибольшей величины вытекает соотношение

$$\varphi_z = \min_B c(U_B^-) \geq d(X).$$

Следовательно, этот поток насыщает все дуги, идущие в выход z сети T . ■

Доказательство теоремы 3.1. Докажем эквивалентность условий теоремы 3.1 и леммы 3.3.

Покажем, что из неравенства (3.2) следует неравенство (3.3) для любого множества $A \subset X \setminus \{x_0, z\}$. Для этого рассмотрим множество $B = A \cap \Gamma^{-1}z \subseteq \Gamma^{-1}z$. Очевидно, что справедливы соотношения

$$c(U_A^-) \geq F(A), \quad F(A) \geq F(B), \quad d(A) = d(Y). \quad (3.15)$$

Из соотношений (3.15) вытекает неравенство (3.3). Следовательно, в силу леммы 3.3 поток, насыщающий выходные дуги, существует.

Пусть поток φ , насыщающий выходные дуги, существует. По лемме 3.3 для любого множества $A \subset X \setminus \{x_0, z\}$ справедливо неравенство (3.3). В силу следствия из леммы 3.2 неравенства (3.3) вытекает неравенство (3.2). ■

Глава IX

ТЕОРЕМА О СИСТЕМЕ РАЗЛИЧНЫХ ПРЕДСТАВИТЕЛЕЙ И ЕЕ ПРИЛОЖЕНИЕ В ТЕОРИИ ГРАФОВ

В теории графов существуют специальные так называемые двудольные графы, являющиеся удобной моделью для ряда важных проблем, возникающих в абстрактной теории конечных множеств.

Формулировка той или иной задачи дискретной (конечной) математики на языке теории графов часто облегчает ее решение. В этом случае использование эффективных алгоритмов, существующих в теории графов, позволяет найти конструктивное решение рассматриваемой задачи. Примером может служить известная в теории множества теорема о системе различных представителей (СРП) [7]. Доказательство этой теоремы приводится ниже с помощью теории транспортных сетей.

В настоящей главе будет приведен ряд результатов (теоремы 1.2—1.5), не обязательно имеющих теоретико-графовую формулировку, эквивалентных теореме о СРП.

Доказательство эквивалентности всех шести теорем является нетривиальной задачей и требует для своего решения введения ряда важных понятий теории графов, таких, как «опора или вершинное покрытие» графа, «дефицит» и «паросочетание» графа. Все эти понятия представляют самостоятельный интерес и находят применение в прикладных задачах.

§ 1. ТЕОРЕМА О СИСТЕМЕ РАЗЛИЧНЫХ ПРЕДСТАВИТЕЛЕЙ И ДРУГИЕ ЭКВИВАЛЕНТНЫЕ ЕЙ УТВЕРЖДЕНИЯ

1. Теорема о системе различных представителей. Пусть S_i ($i = 1, 2, \dots, n$) — подмножества некоторого множества S .

Определение 1.1. Совокупность элементов $\{x_i\}$ ($i = 1, 2, \dots, n$) множества S называется системой различ-

ных представителей семейства множеств S_i ($i=1, 2, \dots, n$), если выполняются следующие два условия: а) $x_i \in S_i$ при каждом $i \in (1, 2, \dots, n)$, б) $x_i \neq x_j$, если $i \neq j$.

Теорема 1.1. Необходимым и достаточным условием существования системы различных представителей семейства множеств S_i ($i=1, 2, \dots, n$) является выполнение следующего условия:

$$\left. \begin{array}{l} \text{для каждой последовательности различных} \\ \text{индексов } i_1, i_2, \dots, i_k, \text{ где } k \in (1, 2, \dots, n) \\ \text{объединение множеств } \bigcup_{j=1}^k S_{i_j} \text{ содержит не ме-} \\ \text{нее чем } k \text{ различных элементов.} \end{array} \right\} \quad (1.1)$$

Заметим, что необходимость условия (1.1) для существования СРП вполне очевидна. В качестве примера рассмотрим совокупность множеств $S_1 = \{1, 3, 1\}$, $S_2 = \{3, 3, 1\}$, $S_3 = \{1, 1, 3\}$, $S_4 = \{2, 3, 1, 3\}$, для которых условие (1.1) не выполняется, так как объединение $\bigcup_{i=1}^3 S_i$ содержит только два различных элемента: 3 и 1. Поэтому нельзя найти трех различных представителей множеств S_1 , S_2 и S_3 .

Занимательной по форме является «теорема о сватовстве», эквивалентная теореме 1.1.

Теорема 1.1'. Пусть n юношей дружат с девушками и пусть для каждой группы, состоящей из k юношей ($k = 1, 2, \dots, n$), имеется, по крайней мере, k девушек, имеющих друзей среди этих юношей. Тогда каждого юношу можно женить на девушке, с которой он дружит.

2. Теорема о цепном разложении частично упорядоченных множеств. Для формулировки теоремы понадобится следующее определение.

Определение 1.2. Множество $X = \{x, y, \dots\}$ называется частично упорядоченным, если в нем введено бинарное отношение для некоторых пар элементов — отношение порядка $x \leq y$, которое подчиняется условиям:

- 1) $x \leq x$ для всякого $x \in X$ (рефлексивность),
- 2) если $x \leq y$ и $y \leq z$, то $x \leq z$ (транзитивность),
- 3) если $x \leq y$ и $y \leq x$, то $x = y$ (антисимметричность).

Если два элемента x и y множества X сравнимы между собой и $x \leq y$, то говорят, что элемент x не превосходит элемента y , или что он предшествует y ; можно сказать, что элемент y превосходит элемент x или что y следует за x .

Примером частично упорядоченного множества является множество целых чисел, причем отношение порядка $x \leq y$ означает, что « y делится без остатка на x ». Другим важным примером является множество вершин X графа $G = (X, \Gamma)$, если отношение порядка $x \leq y$ ввести следующим образом: для двух вершин x и y , не лежащих на одном контуре, задано отношение $x \leq y$, если x совпадает с y или существует путь из вершины x в вершину y .

Определение 1.3. *Всякое подмножество A частично упорядоченного множества X , в котором любые два элемента сравнимы между собой, называется цепью.*

Пусть X — конечное частично упорядоченное множество. Тогда каждое подмножество вида $\{x\}$ в силу рефлексивности отношения порядка представляет собой одноэлементную цепь. Таким образом, множество X имеет очевидное разложение на n ($n = |X|$), не пересекающихся одноэлементных цепей.

Теорема 1.2. *Минимальное число непересекающихся цепей, составляющих разбиение частично упорядоченного множества X , равно максимальному числу взаимно несравнимых элементов множества X [7].*

3. Теорема о числе независимых элементов бинарной матрицы. Рассмотрим матрицу A порядка $n \times t$, элементы a_{ij} , которой равны единице или нулю. *Линией матрицы* называется ее строка или ее столбец. Два элемента матрицы называются *независимыми*, если они не лежат на одной линии.

Теорема 1.3. *Максимальное число независимых единиц любой бинарной матрицы равно минимальному числу линий, содержащих все единицы матрицы [8].*

В качестве примера, иллюстрирующего эту теорему, рассмотрим матрицу

$$A = \begin{vmatrix} 0 & 0 & 1 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 \end{vmatrix}.$$

Все единичные элементы этой матрицы содержатся во второй и третьей строке и третьем и четвертом столбце. Меньшего числа линий, содержащих все единицы в матрице A , нет, поэтому максимальное число независимых единичных элементов A равно четырем.

Следующие ниже теоремы имеют уже теоретико-графовую формулировку.

4. Теорема о числе дуг паросочетания двудольного графа в случае, когда $d(\Pi) = |X|$. Понятие паросочетания графа было фактически введено в гл. VIII в задаче о назначении на должность. У возникающего в этой задаче графа (B, Γ) множество вершин B определялось двумя непересекающимися множествами X и Y , а отображение Γ действовало из X в 2^Y .

Определение 1.4. Граф (B, Γ) , обладающий свойствами:

- 1) $B = X \cup Y$, $X \cap Y = \emptyset$, 2) $\Gamma X \subset 2^Y$, $\Gamma Y = \emptyset$,

называется **двудольным графом** (обозначается как $(X, Y, \Gamma) = G$).

Определение 1.5. Паросочетанием двудольного графа называется такое множество Пего дуг, в котором никакие две дуги не смежны. Дуги, входящие в паросочетание Π , называются **независимыми**.

Через $d(\Pi)$ будем обозначать число дуг паросочетания Π . Всякое паросочетание Π в двудольном графе $G = (X, Y, \Gamma)$ определяет взаимно-однозначное отображение множества A_Π вершин графа, являющихся началами дуг паросочетания Π , ($A_\Pi \subset X$), на множество B_Π концов дуг из Π ($B_\Pi \subset Y$). Это отображение будем обозначать через $\Pi: X \rightarrow Y$. Таким образом, паросочетание можно рассматривать одновременно и как множество дуг графа, и как отображение. Понимая паросочетание как отображение, будем употреблять соответствующие термины, например, «паросочетание из X в Y ».

На рис. 78 выделено паросочетание $\Pi = \{(x_1, y_2), (x_2, y_3), (x_4, y_1)\}$, $A_\Pi = (x_1, x_2, x_4)$, $B_\Pi = (y_1, y_2, y_3)$.

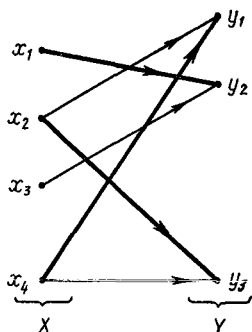


Рис. 78

Задача 1. Доказать, что если в двудольном графе $(X, Y, \Gamma) = G$ выполняются неравенства $|\Gamma x| \geq m$, m — натуральное число, для всех $x \in X$, $|\Gamma^{-1}y| \leq m$, для всех $y \in Y$, то существует паросочетание множества X на множество Y .

Задача 2. Доказать, что если граф $G = (X, Y, \Gamma)$ допускает хотя бы одно максимальное паросочетание, в котором $d(\Pi) =$

$= |X| = n$, и если $\min_{x \in X} |\Gamma x| = r$, то при $r \geq n$ существует не менее чем $r(r-1) \dots (r-n+1) = A_r^n$ различных максимальных паросочетаний, при $r < n$ это число не менее чем $n!$

Из определения паросочетания следует очевидная оценка для числа дуг $d(\Pi)$ произвольного паросочетания Π

$$d(\Pi) \leq |X|. \quad (1.2)$$

В следующей теореме сформулированы условия, при которых достигается оценка (1.2).

Теорема 1.4. [2]. Паросочетание, отображающее X в Y , существует тогда и только тогда, когда для любого множества $A \subseteq X$ выполняется соотношение

$$|A| \leq |\Gamma A|. \quad (1.3)$$

5. Теорема о максимальном (наибольшем) паросочетании двудольного графа при $d(\Pi) < |X|$. Если условия теоремы (1.4) не выполняются, то максимальное число дуг паросочетания определяется дефицитом графа.

Определение 1.6. Дефицитом двудольного графа $G = (X, Y, \Gamma)$ называется число $\delta_0 = \max_{A \subseteq X} (|A| - |\Gamma A|)$. По-

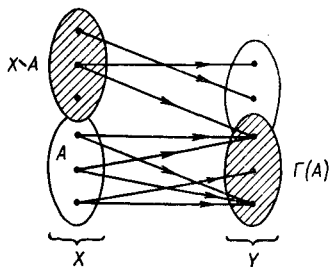


Рис. 79

скольку $\emptyset \subset X$, $|\emptyset| = 0$ и $|\Gamma \emptyset| = 0$, то дефицит двудольного графа $\delta_0 \geq 0$. Наглядный смысл дефицита двудольного графа связан с понятием опоры.

Определение 1.7. Опорой двудольного графа $G = (X, Y, \Gamma)$ называется множество его вершин $S \subset X \cup Y$ такое, что каждая дуга графа G инцидентна по крайней мере

одной вершине из S .

Например, если $A \subset X$, то множество $(X \setminus A) \cup (\Gamma A)$ является опорой графа (рис. 79).

Нетрудно показать, что всякая наименьшая опора S графа $G = (X, Y, \Gamma)$ имеет вид $S = (X \setminus A) \cup (\Gamma A)$.

Действительно, пусть множество $S \subset X \cup Y$ — наименьшая опора графа; справедливы равенства

$$S = S \cap (X \cup Y) = (S \cap X) \cup (S \cap Y). \quad (1.4)$$

Введем обозначение $A = \Gamma^{-1}(C \cap Y)$, тогда в силу того что C — опора графа, имеем равенство

$$(C \cap X) \cup A = X;$$

а поскольку C — наименьшая опора, то $(C \cap X) \cap A = \emptyset$, и поэтому

$$C \cap X = X \setminus A. \quad (1.5)$$

Подставляя (1.5) в (1.4) и учитывая, что $C \cap Y = \Gamma A$, получаем $C = (X \setminus A) \cup (\Gamma A)$. ■

Количественная связь между числом элементов опоры и дефицитом δ_0 устанавливается следующей леммой.

Лемма 1.1. *В двудольном графе $G = (X, Y, \Gamma)$ наименьшее число элементов опоры C равно $|X| - \delta_0$.*

Доказательство. Из определений 1.6 и 1.7 вытекает равенство

$$\begin{aligned} |X| - \delta_0 &= |X| - \max_{A \in X} (|A| - |\Gamma A|) = \\ &= \min_{A \in X} (|X| - |A| + |\Gamma A|). \end{aligned}$$

Введя обозначение $B = X \setminus A$, получим

$$\begin{aligned} \min (|X| - |A| + |\Gamma A|) &= \min_{B \in X} (|B| + \\ &+ |\Gamma(X \setminus B)|) = \min_C |C|, \end{aligned}$$

где $B \cup \Gamma(X \setminus B) = C$ — опора графа G . ■

Теорема 1.5. *В двудольном графе максимальное число дуг паросочетания равно минимальному числу элементов опоры.*

6. Теорема о максимальном целочисленном потоке в «двудольной» транспортной сети. К множеству вершин $X \cup Y$ двудольного графа $G = (X, Y, \Gamma)$ присоединим две вершины x_0 и z и добавим к множеству Γ дуг графа G дуги вида (x_0, x) , $x \in X$ и (y, z) , $y \in Y$. Обозначим через $\tilde{\Gamma}$ новое множество дуг. Определим на множестве $\tilde{\Gamma}$ дуг графа $G \dot{=} (X \cup \{x_0\} \cup Y \cup \{z\}, \tilde{\Gamma})$ целочисленную функцию $C: \tilde{\Gamma} \rightarrow \mathbb{Z}_+$ такую, что $c(x_0, x) = c(y, z) = 1$ для любого $x \in X$ и любого $y \in Y$, $c(x, y) = \infty$.

Получившуюся транспортную сеть будем называть «двудольной» транспортной сетью и обозначать через T_G .

Теорема 1.6. *Максимальный поток в двудольной сети равен минимальной пропускной способности разрезов данной сети.*

Эта теорема для произвольной транспортной сети была доказана в гл. VIII. В § 2 будет показано, что теорема о максимальном потоке для двудольной транспортной сети следует из теоремы 1.5 о максимальном паросочетании в двудольном графе.

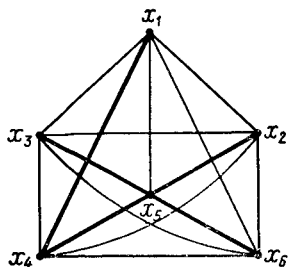


Рис. 80

Задача 3. Пусть $G = (X, Y, \Gamma)$ — двудольный граф. Два элементарных пути графа G , соединяющие вершины x_0 и z графа G , называются непересекающимися, если у них нет общих дуг. Множество дуг M разделяет вершины x_0 и z в графе G , если вершины x_0 и z принадлежат

различным компонентам графа $G_M = (X, Y, \Gamma \setminus M)$.

Доказать, что наибольшее число непересекающихся путей, соединяющих вершины x_0 и z двудольной транспортной сети, равно наименьшему числу дуг, разделяющих вершины x_0 и z .

З а м е ч а н и е. Это утверждение на самом деле справедливо для любого графа и любой пары его вершин (см. [8]). В этом случае утверждение о числе непересекающихся путей эквивалентно теореме о максимальном целочисленном потоке (см. [7]).

На рис. 80 представлен полный шестивершинный граф S_6 , отображающий эквивалентность друг другу всех шести теорем, сформулированных в § 1. Выделенное на рис. 80 частичное дерево

$$(X = \{x_1, x_2, \dots, x_6\}, \Gamma = \{(x_1, x_4), (x_4, x_5), (x_5, x_3), (x_5, x_2), (x_5, x_6)\})$$

графа S_6 является «наикратчайшим» частичным деревом, приводящим к доказательству эквивалентности всех рассматриваемых утверждений.

§ 2. ЭКВИВАЛЕНТНОСТЬ ТЕОРЕМ 1.1. — 1.6

1. Эквивалентность теоремы 1.1. и теоремы 1.4. Для доказательства эквивалентности теорем 1.1 и 1.4 достаточно заметить, что любой совокупности множеств S_i , $i \in \{1, 2, \dots, n\}$ естественным образом можно сопоставить двудольный граф $G_S = (X, Y, \Gamma)$ инцидентий множеств

и элементов. Множеством вершин X графа G является совокупность $\{S_i\}$, $x_i = S_i$, $|X| = n$, а множество вершин Y — есть $\bigcup_{i=1}^n S_i$. Вершина $y \in \Gamma x_i$, если $y \in S_i$. Тогда, очевидно, что условие (1.1) теоремы (1.1) означает не что иное, как условие (1.3) теоремы (1.4) для графа G_S . Наоборот, любой двудольный граф $G = (X, Y, \Gamma)$ можно рассматривать как граф инциденций G_S множеств $S_i = \{y, y \in \Gamma x_i\}$ и элементов $y \in Y$. Таким образом, условие (1.3) теоремы (1.4) влечет за собой условие (1.1) теоремы (1.1). ■

2. Эквивалентность теоремы 1.4. и теоремы 1.5. Для доказательства понадобится следующая лемма.

Лемма 2.1. Положим в графе $G = (X, Y, \Gamma)$ для любого множества $A \subseteq X$ $\delta(A) = |A| - |\Gamma A|$, тогда справедливы утверждения:

$$а) \delta(A_1 \cup A_2) + \delta(A_1 \cap A_2) \geq \delta(A_1) + \delta(A_2);$$

б) семейство множеств $\mathcal{A} = \{A \subseteq X, \delta(A) = \delta_0\}$ замкнуто относительно операций $\cup$ и $\cap$ в конечном числе, т. е. если $(A_1, A_2) \in \mathcal{A}$, то $(A_1 \cup A_2) \in \mathcal{A}$, $(A_1 \cap A_2) \in \mathcal{A}$;

$$в) \text{ множество } A_0 = \bigcap_{A \in \mathcal{A}} A \neq \emptyset, \text{ если } \delta_0 > 0.$$

Доказательство. Неравенство а) получим как разность соотношений:

$$\begin{aligned} & |A_1 \cup A_2| + |A_1 \cap A_2| = |A_1| + |A_2| \\ \text{и} \quad & |\Gamma(A_1 \cup A_2)| + |\Gamma(A_1 \cap A_2)| \leq |\Gamma A_1 \cup \Gamma A_2| + \\ & + |\Gamma A_1 \cap \Gamma A_2| = |\Gamma A_1| + |\Gamma A_2|, \end{aligned}$$

вычитая второе соотношение из первого.

В силу свойства а) имеем соотношение

$$\delta(A_1 \cup A_2) + \delta(A_1 \cap A_2) \geq \delta(A_1) + \delta(A_2) = 2\delta_0.$$

Тогда из неравенств

$$\delta(A_1 \cup A_2) \leq \delta_0, \quad \delta(A_1 \cap A_2) \leq \delta_0$$

следует, что

$$\delta(A_1 \cup A_2) = \delta_0, \quad \delta(A_1 \cap A_2) = \delta_0,$$

т. е. свойство б) доказано.

Поскольку граф G конечен, то A_0 есть пересечение конечного числа множеств из $\mathcal{A}$, а поэтому по свойству б) имеем включение $A_0 \in \mathcal{A}$.

Если $A_0 = \emptyset$, то дефицит $\delta_0 = \delta(A_0) = \delta(\emptyset) = 0$, что противоречит предположению $\delta_0 > 0$. ■

Лемма 2.2. Пусть справедлива теорема 1.4, тогда число дуг максимального паросочетания двудольного графа равно $|X| - \delta_0$.

Доказательство. Рассмотрим множество $A_0 = \bigcap_{A \in \mathcal{A}} A$, где

$\mathcal{A} = \{A, \delta(A) = \delta_0\}$. Если удалить из графа вершину $x_0 \notin A_0$, то дефицит не изменится. Действительно, дефицит δ'_0 графа $(X \setminus \{x_0\}, Y, \Gamma)$ удовлетворяет условию $\delta'_0 \leq \delta_0$, а так как множество $X \setminus \{x_0\} \in A_0$, $\delta(A_0) = \delta_0$, то $\delta'_0 = \delta_0$.

Если удалить из графа вершину $x_0 \in A_0$, то дефицит графа уменьшится ровно на единицу.

В этом случае удаление x_0 «разрушает» в графе G все те множества A , для которых $\delta(A) = \delta_0$ и, поэтому дефицит нового графа $\delta'_0 < \delta_0$. Для множества $A'_0 = A_0 \setminus \{x_0\}$ имеем

$$\delta(A'_0) = |A_0| - 1 - |\Gamma A'_0| < \delta_0 = |A_0| - |\Gamma A_0|,$$

откуда следует, что $|\Gamma A'_0| \geq |\Gamma A'_0|$. Но в силу того что $A'_0 \subset A$, имеем $|\Gamma A_0| \geq |\Gamma A'_0|$, а значит $|\Gamma A'_0| = |\Gamma A_0|$. Отсюда получим равенство

$$\begin{aligned} \delta(A'_0) &= |A'_0| - |\Gamma A'_0| = (|A_0| - 1) - |\Gamma A_0| = \\ &= \delta_0 - 1. \end{aligned}$$

Тогда из справедливости соотношений $\delta(A'_0) = \delta_0 - 1 \leq \delta'_0 < \delta_0$ следует, что

$$\delta'_0 = \delta_0 - 1.$$

Таким образом, удаляя из множества A_0 графа G количество вершин, равное числу δ_0 , получим граф G' с δ_0 , в котором, в силу теоремы 1.4, существует паросочетание Π из X' в Y , $d(\Pi) = |X'| = |X| - \delta_0$, являющееся наибольшим. ■

Из лемм 2.1 и 2.2 следует, что справедливость теоремы 1.4 влечет справедливость теоремы 1.5.

Пусть теперь справедлива теорема 1.5. Тогда, в силу леммы 1.1, $\max_{\Pi} d(\Pi) = |X| - \delta_0$. Отсюда вытекает необходимость и достаточность условия (1.3), ибо равенство $\delta_0 = 0$ эквивалентно условию (1.3).

3. Эквивалентность теорем 1.5 и 1.3. Эквивалентность теорем 1.5 и 1.3 нетрудно установить, если для двудоль-

ного графа $G = (X, Y, \Gamma)$ с дефицитом $\delta_0 \neq 0$ рассмотреть матрицу: $A_G = ((a_{ij}))$ ($i \in (1, 2, \dots, |X|), j \in (1, 2, \dots, |Y|)$), где $a_{ij} = 1$ в том и только том случае, когда $(x_i, y_j) \in \Gamma$. Независимость двух единичных элементов матрицы A_G означает, что соответствующая им пара дуг графа G несмежна. Поэтому максимальное число дуг паросочетания равно максимальному числу независимых единичных элементов.

С другой стороны, каждой линии матрицы A_G отвечает вершина графа G , вершина $x \in X$, если линия — строка матрицы и вершина $y \in Y$, если линия — столбец матрицы. Причем, число единичных элементов линии равно числу дуг графа G , инцидентных вершине x либо вершине y . Отсюда следует, что множество линий, содержащих все единичные элементы матрицы A_G взаимно-однозначно определяет множество вершин графа G , которое по определению 1.7 является опорой графа. Отсюда следует, что минимальное число элементов опоры графа равно минимальному числу линий, содержащих все единичные элементы матрицы A_G . Поскольку любая бинарная матрица является матрицей A_G двудольного графа, то эквивалентность теорем 1.5, 1.3 становится очевидной. ■

4. Эквивалентность теоремы 1.2 теореме 1.5. Пусть $X = \{x_1, x_2, \dots, x_n\}$ — частично упорядоченное множество. Рассмотрим двудольный граф $G_X = (X, X', \Gamma)$, у которого множество вершин составляет два экземпляра множества X , т. е. $X' = X$, а отображение $\Gamma: X \rightarrow 2^{X'}$ определено правилом: пара вершин $x_i \in X$ и $x'_j \in X'$ определяет дугу $(x_i, x'_j) \in \Gamma$, если в частично упорядоченном множестве элемент $x_i \leq x_j$. При любом $i \in (1, 2, \dots, n)$ дуга $(x_i, x'_i) \in \Gamma$, так как в силу рефлексивности отношения порядка всегда $x_i \leq x_i$. Рассмотрим граф на рис. 81. В нем элементы x_1 и x_2 эквивалентны ($x_1 = x_2$), а последовательность (x_n, x_k, x_i) образует цепь множества X . Каждой цепи длины $l \leq n$ множества X соответствует паросочетание с числом дуг, равным $l - 1$ в графе G , на-

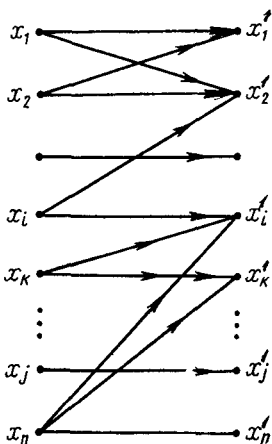


Рис. 81

пример, цепи (x_n, x_k, x_i) отвечает паросочетание $\{(x_n, x'_k), (x_k, x'_i)\}$.

Если две цепи не пересекаются, то при объединении дуг паросочетаний, отвечающих этим цепям, снова получается паросочетание. Тогда если через r обозначить произвольное разбиение множества X на цепи: $X = \sum_{i=1}^{N(r)} X_i$, где $N(r)$ — число элементов разбиения, то это разбиение однозначно определяет паросочетание (обозначим его через $\Pi(r)$, для которого выполняется равенство

$$d(\Pi(r)) = |\Pi(r)| = \sum_{i=1}^{N(r)} |X_i - 1| = \sum_{i=1}^{N(r)} |X_i| - N(r) = |X| - N(r) = n - N(r). \quad (2.1)$$

Справедливо и обратное утверждение.

Лемма 2.2. Для любого паросочетания Π существует разбиение $r = r(\Pi)$ множества X такое, что

$$N(r(\Pi)) = n - d(\Pi). \quad (2.2)$$

Доказательство. Если множество дуг $\{(x_{i_1}, x'_{i_2}), (x_{i_3}, x'_{i_4}) \dots (x_{i_m}, x'_{i_{m+1}})\}$ — произвольное паросочетание Π , то $x_{i_1} \leq x_{i_2}, x_{i_3} \leq x_{i_4}, \dots, x_{i_m} \leq x_{i_{m+1}}$. Если объединить элементы $\{x_{i_1}, x_{i_2}, \dots, x_{i_m}, x_{i_{m+1}}\}$ в цепи множества X (эти цепи не будут пересекаться в силу независимости дуг паросочетания Π) и добавить в качестве одноэлементных цепей все еще не вошедшие в паросочетание Π элементы множества X , то получится разложение $r(\Pi)$ множества X на непересекающиеся цепи $\tilde{X}_i$ ($i = 1, 2, \dots, N(r(\Pi))$); причем

$$\begin{aligned} |X| &= \left| \sum_{i=1}^{N(r)} \tilde{X}_i \right| = \sum_{i=1}^{N(r)} |\tilde{X}_i - 1| + N(r(\Pi)) = \\ &= d(\Pi) + N(r(\Pi)). \quad \blacksquare \end{aligned}$$

Из равенства (2.1), утверждения леммы (2.2) и теоремы 1.5 следует, что минимальное число цепей в разбиении множества X равно дефициту δ_0 графа G_X , так как

$$\min_r N(r) = n - \max_r d(\Pi(r)) = n - \max_{\Pi} d(\Pi) = \delta_0. \quad (2.3)$$

Число несравнимых друг с другом элементов множества X естественно связать с опорой C графа G_X .

Лемма 2.3. Если C_0 — минимальная опора графа G_X , то существует множество $\gamma(C_0) \subset X$ попарно несравнимых элементов, для которого

$$|\gamma(C_0)| = n - |C_0|, \quad |X| = n. \quad (2.4)$$

Доказательство. Пусть C_0 — некоторая минимальная опора графа G и пусть

$$C_0 = C_X \cup C_{X'} = \{x_{i_1}, \dots, x_{i_k}, x'_{j_1}, \dots, x'_{j_m}\} \quad (2.5)$$

$$(0 \leq k, m \leq n, k + m \leq n).$$

Все индексы в (2.5) различны, т. е. $C_X \cap C_{X'} = \emptyset$. Предположим противное, например, что $i_1 = j_1$. Поскольку C_0 — минимальная опора, то найдется вершина $x \notin C_X$ такая, что $(x, x'_{j_1}) \in \Gamma$. Аналогично, найдется вершина $x' \in C_{X'}$, $(x_{i_1}, x') \in \Gamma$. Тогда из транзитивности отношения порядка и определения графа G_X следует, что дуга $(x, x') \in \Gamma$ и ни одним из своих концов не смежна с множеством вершин из C_0 , а это противоречит тому, что C_0 — опора. Таким образом, все индексы $\{i_1, \dots, i_k, j_1, \dots, j_m\}$ попарно различны. Если теперь рассмотреть в частично упорядоченном множестве X множество

$\gamma(C_0) = X \setminus \{x_{i_1}, \dots, x_{i_k}, x'_{j_1}, \dots, x'_{j_m}\}$, то из определения опоры следует, что элементы множества $\gamma(C_0)$ несравнимы и $|\gamma(C_0)| = n - |C_0|$. ■

Используя результаты лемм 2.2 и 2.3, легко получить, что из теоремы 1.5 следует теорема 1.2.

Действительно, из (2.3) и (2.4) имеем соотношения:

$$n - \min_C |C| \equiv n - |C_0| = |\gamma(C_0)| \leq \max_C |\gamma(C)| \leq \max_r |\gamma(r)|, \quad (2.6a)$$

$$n - \max_{\Pi} d(\Pi) = \min_r N(r). \quad (2.6b)$$

Но поскольку всегда для разбиения r частично упорядоченного множества X выполняется неравенство

$$|\gamma(r)| \leq N(r),$$

то из (2.6а) и (2.6б) получаем утверждением теоремы 1.2, при условии, что верна теорема 1.5, т. е. что $\min_C C = \max_{\Pi} d(\Pi)$.

Верно и обратное, что из теоремы 1.2 следует теорема 1.5. Для доказательства этого сопоставим двудольному графу $G = (X, Y, \Gamma)$ частично упорядоченное множество $X \cup Y$, считая, по определению, что элементы x_i и y_j находятся в отношении $x_i \leq y_j$, если дуга $(x_i, y_j) \in \Gamma$, $x_i \in X$, $y_j \in Y$. Тогда если r — произвольное разбиение множества $X \cup Y$ с числом элементов разбиения, равным $N(r)$, то для него существует паросочетание $\Pi(r)$ такое, что

$$d(\Pi(r)) = |X \cup Y| - N(r). \quad (2.7)$$

Паросочетание $\Pi(r)$ состоит из дуг, отвечающих двухэлементным цепям разбиения r .

И, наконец, если $\gamma(r)$ — множество несравнимых элементов, то дополнение $(X \cup Y) \setminus \gamma(r)$ содержит опору $C(\gamma)$ графа G , и поэтому

$$|C(\gamma)| \leq |X \cup Y| - |\gamma(r)|.$$

Отсюда с учетом (2.7) и при условии, что верна теорема 1.2 ($\max_r \gamma(r) = \min_r N(r)$), получаем неравенства

$$\begin{aligned} \min_C |C| &\leq \min_{\gamma} C(\gamma) \leq |X \cup Y| - \max_{\gamma} |\gamma| = \\ &= |X \cup Y| - \min_r N(r) = \max_{\Pi} d(\Pi(r)), \end{aligned}$$

из которых следует утверждение теоремы 1.5 ($\max_{\Pi} d(\Pi) = \min_C |C|$), поскольку $d(\Pi) \leq |C|$ для всех Π и C .

5. Эквивалентность теорем 1.5 и 1.6. Для доказательства эквивалентности теорем достаточно проверить выполнение следующих равенств:

$$\max_{\Pi} d(\Pi) = \max_{\varphi} \varphi, \quad (2.8)$$

$$\min_C |C| = \min_A C(U_A^-), \quad (2.8a)$$

где через Π и C обозначаются паросочетание Π и опора C в двудольном графе $G = (X, Y, \Gamma)$, а через φ и A — поток и множество вершин, определяющих разрез U_A^- в транспортной сети $T_{\tilde{G}}$ вида:

$$1) \tilde{G} = (X \cup \{x_0\} \cup Y \cup \{z\}, \tilde{\Gamma}), \quad \tilde{\Gamma} = \Gamma \cup \{(x_0, x), x \in X\} \cup \{(y, z), y \in Y\},$$

$$2) c(x_0, x) = 1, x \in \tilde{X}; c(y, z) = 1, y \in Y; c(x, y) = \infty, \\ x \in X, y \in Y.$$

Докажем равенство (2.8). Если $\varphi(u), u \in \tilde{\Gamma}$ — поток в транспортной сети $T_{\tilde{G}}$, то множество дуг $\Pi(\varphi) = \{(x, y) \in \tilde{\Gamma}, \varphi((x, y)) = 1\}$ является паросочетанием в графе G . С другой стороны, очевидно, что всякое паросочетание Π в графе G определяет однозначно поток $\varphi(\Pi)$ в сети $T_{\tilde{G}}$.

При указанном взаимно-однозначном соответствии между множеством паросочетаний $\{\Pi\}$ графа G и множеством потоков $\{\varphi\}$ в сети $T_{\tilde{G}}$ выполняется равенство

$$d(\Pi(\varphi)) = \varphi_z,$$

из которого следует соотношение (2.8).

Перейдем к доказательству соотношения (2.8a). Пусть U_A^- — произвольный разрез транспортной сети $T_{\tilde{G}}$, определяемый множеством вершин A , $A \ni z$, $A \not\ni x_0$.

Множество A представимо в виде

$$A = X_A \cup Y_A \cup \{z\}, \quad X_A = X \cap A, \quad Y_A = Y \cap A, \quad (2.9)$$

причем пропускная способность разреза $C(U_A^-)$ равна

$$C(U_A^-) = 1 \cdot |X_A| + 1 \cdot |Y \setminus Y_A| + \infty \cdot |\{(x, y), \\ x \in X \setminus X_A, y \in Y_A\}|. \quad (2.10)$$

Иллюстрацией к соотношениям (2.9) и (2.10) служит рис. 82, в котором множество $X_A = \{x_1, x_2\}$, $Y_A = \{y_6, y_5, y_4\}$, дуга (x_5, y_6) имеет пропускную способность $c = \infty$.

Очевидно, что для разреза U_A^- с минимальной пропускной способностью множество дуг $\{(x, y), x \in X \setminus X_A, y \in Y_A\} = \emptyset$, т. е. для такого разреза выполняется условие

$$\Gamma(X \setminus X_A) \cap Y_A = \emptyset. \quad (2.11)$$

Если теперь U_A^- — разрез, удовлетворяющий условию (2.11) (на рис. 82 такой разрез, например, определяется множеством дуг, заходящих в множество $A = \{x_1, x_2, y_6, y_5,$

$y_4 z\}$, если из графа G удалить дугу (x_5, y_6) , то множество вершин $C(A)$, где

$$C(A) = X_A \cup \Gamma(X \setminus X_A),$$

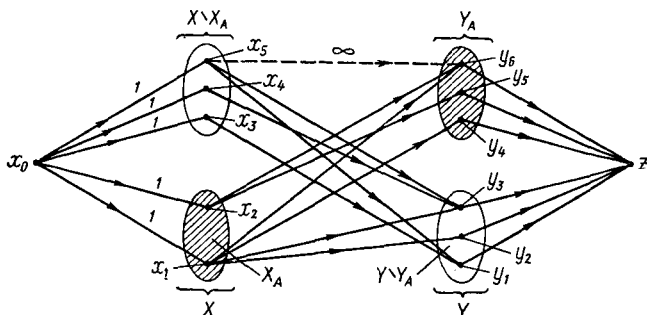


Рис. 82

очевидно, является опорой графа $G = (X, Y, \Gamma)$. Причем в силу условия (2.11), для величин $C(A)$ и $C(U_A^-)$ справедливо неравенство

$$|C(A)| = |X_A| + |\Gamma(X \setminus X_A)| \leq |X_A| + |Y \setminus Y_A| = C(U_A^-). \quad (2.12)$$

Отсюда следует, что

$$\min_A |C(A)| \leq \min_A C(U_A^-), \quad (2.13)$$

а значит

$$\min_C |C| \leq \min_A |C(A)| \leq \min_A C(U_A^-). \quad (2.14)$$

На самом деле, в соотношениях (2.14) всегда выполняются равенства. Действительно, если C_0 — минимальная опора, то по лемме 1.1 она представляется в виде

$$C_0 = B_0 \cup \Gamma(X \setminus B_0).$$

Полагая

$$X_{A_0} = B_0, \quad Y_{A_0} = Y \setminus \Gamma(X \setminus B_0),$$

определим множество A_0

$$A_0 = X_{A_0} \cup Y_{A_0} \cup \{z\},$$

а по нему разрез $U_{A_0}^-$ с пропускной способностью

$$C(U_{A_0}^-) = |X_{A_0}| + |Y \setminus Y_{A_0}| = |B_0| + |\Gamma(X \setminus B_0)| = |C_0|.$$

Отсюда следует, в силу (2.14), равенство

$$|C_0| = \min_C |C| = C(U_{A_0}^-) = \min_C C(U_A^-),$$

доказывающее соотношение (2.8а).

З а м е ч а н и е. Использование при доказательстве эквивалентности теоремы 1.5 и 1.6 двудольной сети с пропускной способностью $C(u) = \infty$, $u = (x_i, y_j)$ $x_i \in X$, $y_j \in Y$ не ограничивает общности рассуждений. Нетрудно видеть, что величина потока в двудольной транспортной сети не зависит от величины пропускных способностей дуг $u \in \Gamma$, если только $c(u) \neq 0$.

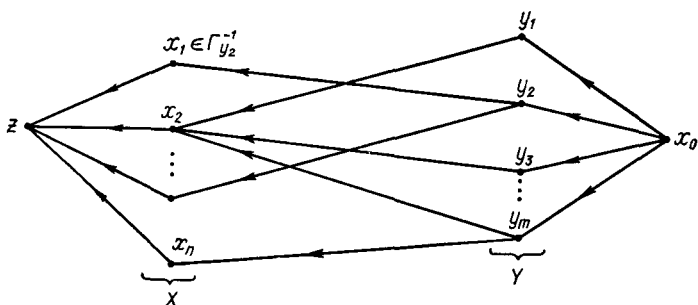


Рис. 83

6. Доказательство теоремы о системе различных представителей. В силу эквивалентности теорем 1.1 — 1.6 § 1 достаточно доказать любое утверждение, эквивалентное теореме о СРП. Для этой цели удобно выбрать теорему 1.4, дающую теоретико-графовую формулировку теоремы о СРП.

Пусть $G = (X, Y, \Gamma)$ — двудольный граф. Сопоставим графу G транспортную сеть $\tilde{G}$ (рис. 83), дополнив двудольный граф (X, Y, Γ^{-1}) входом x_0 и выходом z , причем выберем пропускные способности:

$$\begin{aligned} c(x_0, y_k) &= 1 \quad (k = 1, 2, \dots, m), \\ c(y_k, x_j) &= \infty, \text{ если } x_j \in \Gamma^{-1}y_k, \\ c(x_j, z) &= 1 \quad (j = 1, 2, \dots, n). \end{aligned} \tag{2.15}$$

Тогда полная потребность $d(A)$ некоторого множества $A \subset X$ будет равна $|A|$, а наибольшее количество потока $F(A)$, которое может пройти через A , будет равно $|GA|$. Произвольный поток φ в построенной сети определяет некоторое паросочетание P . Множество P в этом случае состоит из тех и только тех дуг (x_i, y_j) , для которых $\varphi(y_j, x_i) = 1$. Очевидно, что паросочетание P отображает X в Y тогда и только тогда, когда поток насыщает все выходные дуги. Такой поток существует в силу теоремы о насыщении (гл. VIII), если

$F(A) \geq d(A)$, или $|GA| \geq |A|$ для любого $A \subset X$. ■

Как следует из доказательства, для нахождения максимального паросочетания можно применить алгоритм Форда — Фалкерсона (см. § 1 гл. VIII).

В случае, когда условие (1.1) или, что то же самое, условие (1.4) не выполняется, для выделения совокупности множеств, имеющих СРП, из данного семейства множеств $\{S_i, i \in (1, 2, \dots, n)\}$ можно использовать алгоритм теории транспортных сетей.

Существует, однако, более прямой путь решения этой задачи, основанный на идее «чередующихся» цепей двудольного графа и известный под названием «венгерский алгоритм».

§ 3. ВЕНГЕРСКИЙ АЛГОРИТМ

Общая идея алгоритма чрезвычайно проста [2].

Определение 3.1. Пусть для двудольного графа $G = (X, Y, \Gamma)$ построено некоторое паросочетание P . Дуги P будем называть «жирными», остальные дуги — «тонкими». Вершина графа, инцидентная «жирной» дуге, называется насыщенной, в противном случае — ненасыщенной.

Определение 3.2. Цепь*, состоящая из поочередно меняющихся «жирных» и «тонких» дуг, называется «чередующейся цепью».

На рис. 84, а «чередующейся цепью», которая соединяет ненасыщенные вершины x_1 и y_2 , будет цепь $\mu = \{(x_1, y_1), (y_1, x_2), (x_2, y_2)\}$.

* Здесь и на протяжении всего параграфа под цепью $\mu = \{(x_1, y_1), (y_1, x_2), \dots\}$ в двудольном графе понимается последовательность неупорядоченных пар (x_i, y_k) вершин графа.

Определение 3.3. Чередующиеся цепи, которые соединяют ненасыщенные вершины, называются «тонкими».

Наличие в графе тонкой чередующейся цепи позволяет построить вместо паросочетания $\Pi = \{(x_2, y_1)\}$ ($d(\Pi) = 1$) новое паросочетание $\tilde{\Pi} = \{(x_1, y_1), (x_2, y_2)\}$ ($d(\tilde{\Pi}) = 2$), которое содержит на одну дугу больше (рис. 84, б). Для этого необходимо все тонкие дуги цепи μ сделать жирными, а жирные — тонкими, не затрагивая дуг вне цепи μ . К полученному графу G с новым паросочетанием Π при наличии тонкой чередующейся цепи можно применить аналогичную операцию «перекраски» дуг и получить граф с паросочета-

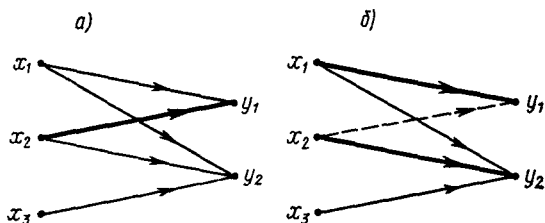


Рис. 84

нием $\tilde{\Pi}$, имеющим на две дуги больше, чем исходное паросочетание Π .

Как видно будет из дальнейшего, отсутствие тонких чередующихся цепей означает, что построенное паросочетание является максимальным. Таким образом, алгоритм построения максимального паросочетания фактически сводится к процессу нахождения тонких чередующихся цепей.

Для компактного описания такого алгоритма, а также для его обоснования удобно ввести следующие обозначения. Пусть $\Pi = \{\pi_1, \pi_2, \dots, \pi_{d(\Pi)}\}$ — множество дуг паросочетания Π , упорядоченное произвольным образом. Обозначим через Π^+ — множество тех жирных дуг, концы которых смежны хотя бы с одной ненасыщенной вершиной, а через Π^- — множество жирных дуг, начала которых смежны хотя бы с одной ненасыщенной вершиной (рис. 85). Заметим, что множества

$$\Pi = \{(x_2, y_1), (x_3, y_4), (x_4, y_2)\} = \{\pi_1, \pi_2, \pi_3\},$$

$\Pi^+ = \{\pi_1, \pi_2, \pi_3\}$, $\Pi^- = \{\pi_1, \pi_2\}$ удовлетворяют включению $\Pi \supseteq \Pi^+ \cup \Pi^-$.

Множества P^+ и P^- не образуют, вообще говоря, разбиения множества P . Паросочетание P , представленное на рис. 85, *полное*, в том смысле, что каждая тонкая дуга смежна, по крайней мере, с одной жирной. Впредь будут рассматриваться полные паросочетания, поскольку от произвольного паросочетания графа к полному легко перейти, добавляя к паросочетанию тонкие дуги, не смежные с жирными.

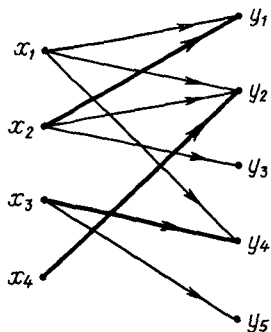


Рис. 85

Поиск тонких чередующихся цепей удобно проводить на вспомогательном графе (P, Γ^*) , вершины которого соответствуют дугам паросочетания P , а отображение Γ^* определяется следующим образом: при $i \neq j$, $\Gamma^* \pi_i \in \pi_j$, если только конец дуги паросочетания π_j является смежным с началом дуги π_i (рис. 86) и всегда $\Gamma^* \pi_i \in \pi_i$. Так, для рассматриваемого примера (рис. 85) граф (P, Γ^*) имеет вид, представленный на рис. 87.

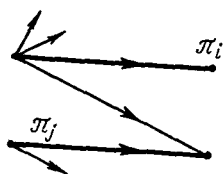


Рис. 86

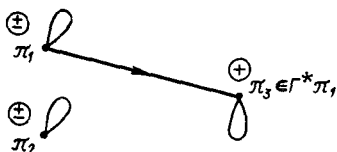


Рис. 87

Из определения множеств P^+ , P^- следует, что тонкая чередующаяся цепь существует, если только найдется хотя бы один путь от вершины, принадлежащей множеству P^+ к вершине из множества P^- на графе (P, Γ^*) . В этом случае возможна замена паросочетания P на $\tilde{P}$ так, что $d(\tilde{P}) = d(P) + 1$. Выбирая, например, на рис. 87 путь $\{\pi_2, \pi_3\}$ ($\pi_2 \in P^+$, $\pi_3 \in P^-$), получаем одно из максимальных паросочетаний

$\tilde{P} = \{(x_1, y_4), (x_2, y_1), (x_3, y_5), (x_4, y_2)\} = \{\pi_1, \pi_2, \pi_3, \pi_4\}$ (рис. 88), для которого $d(\tilde{P}) = |X| = 4$, а соответствующий граф $(\tilde{P}, \Gamma^*)$ изображен на рис. 89. Другое макси-

мальное паросочетание можно получить, если на графе (Π, Γ^*) выбрать путь $\{\pi_1, \pi_1\}$, ведущий из Π^+ в Π^- (рис. 87).

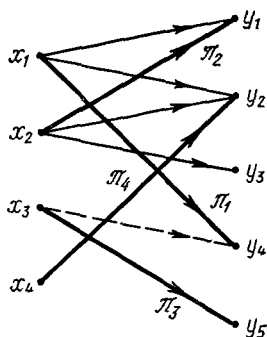


Рис. 88

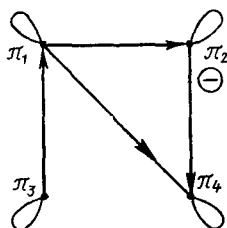


Рис. 89

Алгоритм построения максимального паросочетания в двудольном графе

1°. Построить произвольное полное паросочетание Π . Если Π отображает X в Y , то Π — максимальное паросочетание, если нет, то перейти к 2°.

2°. Построить граф (Π, Γ^*) , отметить на нем вершины множеств Π^+ и Π^- и найти путь, ведущий из Π^+ в Π^- , например, с помощью алгоритма выхода из лабиринта (см. гл. II). Если такого пути нет, то паросочетание максимальное, в противном случае заменить паросочетание Π на $\tilde{\Pi}(d(\tilde{\Pi}) = d(\Pi) + 1)$ и возвратиться к 1°.

Обратимся теперь к обоснованию алгоритма и докажем следующую основную теорему метода чередующихся цепей.

Теорема 3.1. Следующие четыре условия равносильны:

1) в двудольном графе (X, Y, Γ) паросочетание Π максимальное;

2) в двудольном графе (X, Y, Γ) не существует тонких чередующихся цепей;

3) в графе (Π, Γ^*) не существует пути, идущего из точки множества Π^+ в точку множества Π^- ;

4) в графе (Π, Γ^*) возможна такая маркировка вершин знаками «+» или «-», что все вершины множества Π^+ име-

ют знак «+», все вершины множества Π^- имеют знак «—» и не существует дуги, идущей из «+» в «—».

Доказательство. Покажем, что справедлива следующая цепочка импликаций $(1) \Rightarrow (2) \Rightarrow (3) \Rightarrow (4) \Rightarrow (1)$:

$(1) \Rightarrow (2)$, так как если бы существовала тонкая чередующаяся цепь μ_0 , то паросочетание $\tilde{P} = (P \setminus \mu_0) \cup (\mu_0 \setminus P)$ имело бы на одну дугу больше, чем P , что невозможно ввиду максимальности паросочетания P ;

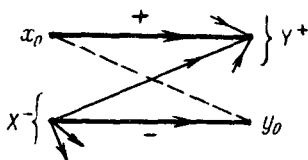


Рис. 90

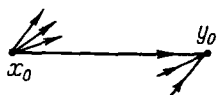


Рис. 91

$(2) \Rightarrow (3)$, так как всякий путь графа (P, Γ^*) из P^+ в P^- определяет хотя бы одну чередующуюся цепь, соединяющую ненасыщенные вершины $x_0 \in X$ и $y_0 \in Y$;

$(3) \Rightarrow (4)$, поскольку справедливо включение $P^+ \subset \hat{\Gamma}^* P^+$ и (по условию) не существует пути из P^+ в P^- , то $P^- \subset P \setminus \hat{\Gamma}^* P^+$. Пометим знаком «+» все вершины множества $\hat{\Gamma}^* P^+$ — концов всех путей с началами в P^+ , а знаком «—» — вершины множества $P \setminus \hat{\Gamma}^* P^+$. Тогда условие 4) будет выполняться, ибо если бы существовала дуга из $\hat{\Gamma}^* P^+$ в P^- , то существовал бы путь из P^+ в P^- ;

$(4) \Rightarrow (1)$. Пометим в графе (X, Y, Γ) дуги паросочетания P знаками «+» или «—» в соответствии с условием 4) и обозначим через X^- множество начальных точек дуг паросочетания, помеченных знаком «—», а через Y^+ — множество конечных точек дуг паросочетания, помеченных знаком «+» (рис. 90). Множество $X^- \cup Y^+ = C$ является опорой графа, т. е. не существует дуги (x_0, y_0) такой, что $x_0 \notin X^-$, $y_0 \notin Y^+$ (последнее утверждение будет доказано ниже). Для любого паросочетания P и любой опоры C имеем: $d(P) = |P| \leq |C|$, а построенные паросочетание P и опора C удовлетворяют условию $|P| = |C|$, поэтому паросочетание P — наибольшее, а опора C — наименьшее.

Предположим теперь, что существует дуга (x_0, y_0) , $x_0 \notin X^-$, $y_0 \notin Y^+$ и рассмотрим четыре возможных случая:

1) вершины x_0 и y_0 не насыщены. Этот случай невозможен, так как P — полное паросочетание (рис. 91);

2) вершина x_0 не насыщена, вершина y_0 насыщена (рис. 92). Тогда существует дуга $\pi \in \Pi$ с концом в y_0 . Поскольку $y_0 \notin Y^+$, то дуга π должна быть помечена знаком «—», но с другой стороны, по определению множества Π^+ , она должна принадлежать этому множеству и быть помеченной поэтому знаком «+». Таким образом, из условия 4) следует, что этот случай невозможен;

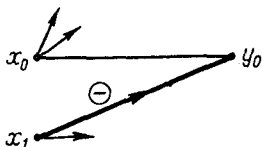


Рис. 92

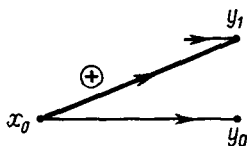


Рис. 93

3) вершина x_0 насыщена, вершина y_0 не насыщена (рис. 93). Тогда существует дуга $\pi \in \Pi$ с началом в x_0 . Поскольку $x_0 \notin X^-$, то дуга π должна быть помечена знаком «+», но так как $\pi \in \Pi^-$, то она должна быть помечена знаком «—», что невозможно;

4) вершины x_0 и y_0 насыщены (см. рис. 90), причем x_0 — начало $\pi_1 \in \Pi^+$, y_0 — конец $\pi_2 \in \Pi^-$. Тогда по определению графа (Π, Γ^*) , $\pi_2 \in \Gamma^* \pi_1$, что противоречит условию 4), по

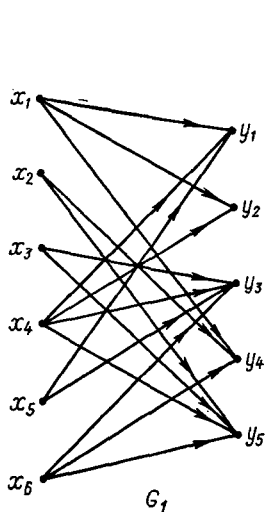


Рис. 94

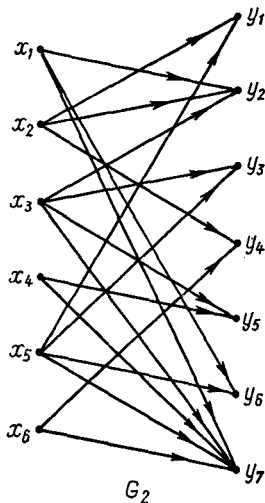


Рис. 95

которому никакая дуга не должна идти из «+» в «—». Таким образом, и этот случай невозможен. ■

Задача. Построить максимальное паросочетание для графов G_1 и G_2 , изображенных на рис. 94 и 95.

§ 4. ПРИКЛАДНЫЕ ЗАДАЧИ НА ПАРОСОЧЕТАНИЯ

1. Нахождение хроматического класса двудольного графа.

Определение 4.1. Хроматическим классом графа G называется натуральное число $\chi(G)$, удовлетворяющее следующим условиям: 1) каждую дугу графа можно окрасить в какую-нибудь из $\chi(G)$ красок таким образом, что никакие две смежные дуги не будут окрашены одинаково;

2) такой раскраски дуг графа нельзя добиться с помощью $\chi(G) - 1$ красок.

Для двудольного графа $G = (X, Y, \Gamma)$ нетрудно вычислить его хроматический класс $\chi(G)$, используя понятие степени двудольного графа $st(G)$.

Определение 4.2. Степень $st(G)$ двудольного графа $G = (X, Y, \Gamma)$ определяется формулой

$$st\,G = \max \{ |\Gamma x|, |\Gamma^{-1}y| \mid x \in X, y \in Y \}. \quad (4.1)$$

Для двудольного графа $G = (X, Y, \Gamma)$ справедливо равенство

$$\chi(G) = st(G). \quad (4.2)$$

Доказательство равенства (4.2) основано на существовании в двудольном графе специального паросочетания, содержащего все вершины графа, полустепень которых максимальна и равна $st(G)$. Если такое паросочетание построено, то окрашиваем первой краской все дуги этого паросочетания. Далее, второй краской окрашиваем дуги паросочетания в суграфе G' , состоящем из еще не окрашенных дуг, которые содержат вершины наибольшей степени $st(G') = st(G) - 1$.

Поскольку, при каждом шаге степень $st(G)$ уменьшается на единицу, то красок в количестве $m = st(G)$ достаточно для раскрашивания всех дуг графа (X, Y, Γ) .

Доказательство равенства (4.2) будет завершено, если доказать следующее утверждение.

Лемма 4.1. В двудольном графе (X, Y, Γ) всегда существует паросочетание, использующее все те вершины $x \in X$,

для которых $|Gx| = \text{st}(G)$, и все те вершины $y \in Y$, для которых $|\Gamma^{-1}y| = \text{st}(G)$.

Доказательство. В силу теоремы 1.4 в графе $G' = (X', Y', \Gamma')$, полустепень вершин которого одна и та же и равна $\text{st}(G')$, всегда существует паросочетание, отображающее X' в Y' .

Построим для графа $G = (X, Y, \Gamma)$, $\text{st}(G) = m$, граф $G' = (X', Y', \Gamma')$ такой, что 1) $\text{st}(G') = m$; 2) G' содержит G в качестве подграфа. Для этого удобно взять m экземпляров графа G :

$$(X^1, Y^1, \Gamma), (X^2, Y^2, \Gamma), \dots, (X^m, Y^m, \Gamma), \quad (4.3)$$

и, если вершина $x_i \in X$ такова, что $m - |Gx_i| = m_i > 0$, $i \in (1, 2, \dots, |X|)$, то рассмотреть соответствующие вершины $x_i^1 \in X^1, x_i^2 \in X^2, \dots, x_i^m \in X^m$ и соединить каждую из них дугами со всеми вершинами нового множества $M(x_i) = \{1, 2, \dots, m_i\}$, состоящего из m_i точек. Рассмотрев таким образом все точки $x_i \in X$, $i \in (1, 2, \dots, |X|)$, перейти к вершинам $y_j \in Y$, $j \in (1, 2, \dots, |Y|)$ и в случае, когда $m - |\Gamma^{-1}y_j| > m_j$, построить множество $M(y_j) = \{1, 2, \dots, m_j\}$ и соединить каждую из вершин $y_j^1 \in Y^1, y_j^2 \in Y^2, \dots, y_j^{m_j} \in Y^{m_j}$ со всеми вершинами множества $M(y_j)$.

Если рассмотреть теперь двудольный граф $G' = (X', Y', \Gamma')$ с множеством вершин

$$X' = \bigcup_{1 \leq i \leq |X|} X^i \cup M(y_j), \quad Y' = \bigcup_{1 \leq j \leq |Y|} Y^j \cup M(x_i)$$

и множеством построенных дуг, то, очевидно (рис. 96), он удовлетворяет требуемому условию. Если теперь $\Pi' : X' \rightarrow Y'$ — максимальное паросочетание в G' , то ограничение отображения Π' на множество X^1 , например, определяет паросочетание в графе $G^{(1)} \equiv G$, содержащее все вершины, полустепени m .

Задача 1. Указать хроматический класс графов G_1 и G_2 (см. рис. 94, 95).

2. Наибольшее внутренне устойчивое множество двудольного графа. Структура графа $G = (X, Y, \Gamma)$ и понятие внутренне устойчивого множества делают почти очевидным результат следующей теоремы.

Теорема 4.1. Если $G = (X, Y, \Gamma)$ — двудольный граф, то 1) множество $C \subset X \cup Y$ является опорой графа тогда и только тогда, когда его дополнение $(X \cup Y) \setminus C = S$ — внутренне устойчивое множество;

2) $\alpha(G) = |Y| + \delta_0$, где δ_0 — дефицит графа G .

Первое утверждение теоремы следует из сопоставления определений множеств C и S . Также просто получить второй результат теоремы:

$$\alpha(G) = \max_S |S| = \max_C (|X| + |Y| - |C|) = |X| + |Y| - \min_C |C| = |X| + |Y| - (|X| - \delta_0) = |Y| + \delta_0.$$

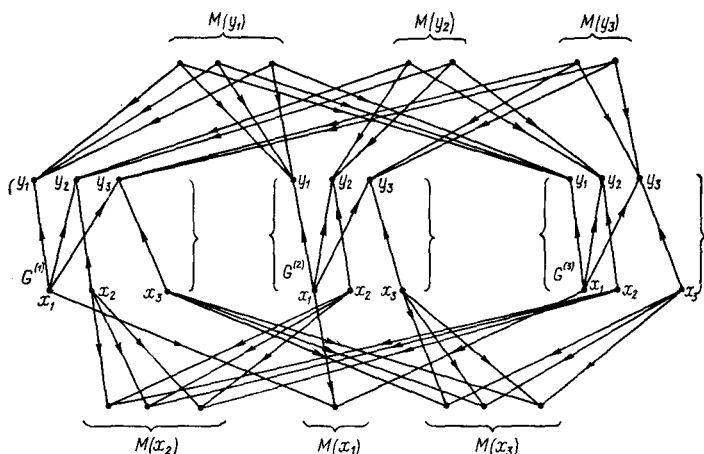


Рис. 96

В этой выкладке мы использовали теорему 1.4 этой главы.

Теорема 4.1 сводит задачу о нахождении наибольшего внутренне устойчивого множества к построению наименьшей опоры графа. Число вершин наименьшей опоры, как известно, равно числу дуг наибольшего паросочетания. Поскольку для нахождения наибольшего паросочетания мы имеем эффективный венгерский алгоритм, то фактически он и является основной частью алгоритма для построения минимальной опоры.

Алгоритм построения наибольшего
внутренне устойчивого множества
двудольного графа

1°. Построить какое-либо наибольшее паросочетание Π графа $G = (X, Y, \Gamma)$.

2°. Пометить в графе G дуги знаками «+» и «—» в соответствии с 4) теоремы 3.1. Тогда наименьшая опора C будет состоять из начальных точек жирных дуг, помеченных знаком «—», и конечных точек жирных дуг, помеченных знаком «+», т. е. $C = X^- \cup Y^+$, где множества X^- и Y^+ были введены при доказательстве теоремы 3.1.

3°. Перейти к дополнению $S = (X \cup Y) \setminus C$.

S — наибольшее внутренне устойчивое множество.

Задача 2. Найти наибольшее внутренне устойчивое множество для графов G_1 и G_2 (см. рис. 94, 95) и определить их число внутренней устойчивости.

Обоснование алгоритма

Для графа G и построенного в п. 1° максимального паросочетания P в силу теоремы 3.1 всегда возможна маркировка дуг, указанная в п. 2° алгоритма.

Дополнение $S = (X \cup Y) \setminus C$, где C — опора является внутренне устойчивым множеством, так как в противном случае найдутся вершины $x \in S$ и $y \in S$, $Gx \ni y$, а следовательно, и дуга $(x, y) \notin C$, что противоречит определению опоры C . Если C — минимальная опора, то $S = (X \cup Y) \setminus C$ — наибольшее внутренне устойчивое множество.

3. Задача об оптимальном назначении (задача 1). Неформальная постановка задачи состоит в следующем. Пусть имеется n рабочих мест на производстве и m кандидатов на эти места. Целые числа $a_{ij} \geq 0$ ($i = 1, 2, \dots, m$, $j = 1, 2, \dots, n$) определяют меру эффективности использования i -го человека на j -м месте. Необходимо произвести назначение m человек на n рабочих мест так, чтобы:

1) суммарный эффект использования рабочих был максимален;

2) на каждое рабочее место назначено не более одного человека;

3) никто не назначен более чем на одно рабочее место. Отметим, что можно ограничиться случаем, когда $m = n$, так как, если $m > n$, всегда можно ввести $(m - n)$ фиктивных работ, для которых $a_{ij} = 0$ при всех $i = 1, 2, \dots, m$, $j = n + 1, \dots, m - n$. Далее, в силу неотрицательности a_{ij} ($i, j = 1, 2, \dots, m$) каждый кандидат должен быть назначен на работу, поскольку в противном случае всегда

можно увеличить суммарную эффективность использования рабочего персонала.

Таким образом, любое назначение определяется подстановкой

$$\pi = \begin{pmatrix} 1 & 2 & \dots & i & \dots & m \\ j(1) & j(2) & \dots & j(i) & \dots & j(m) \end{pmatrix}, \quad (j(i) = \pi(i)), \quad (4.4)$$

и задача состоит в том, чтобы найти такую подстановку $\pi_0 \in S_m$, при которой суммарная эффективность использо-

вания рабочих $\sum_{i=1}^m a_{i \pi(i)}$ ($\pi \in S_m$) была бы максимальной

т. е. требуется найти

$$\max_{\pi \in S_m} \sum_{i=1}^m a_{i \pi(i)}. \quad (4.5)$$

Определим задачу 1' двойственную задаче 1: найти

$$\min \left(\sum_{i=1}^m u_i + \sum_{j=1}^m v_j \right),$$

где минимум берется по всем парам чисел u_i, v_j ($i, j = 1, 2, \dots, m$), удовлетворяющих неравенству

$$u_i + v_j \geq a_{ij} \quad (i, j = 1, 2, \dots, m). \quad (4.6)$$

Важность двойственной задачи определяется следующей теоремой.

Теорема 4.2*. Максимум суммы $\sum_{i=1}^m a_{i, \pi(i)}$ задачи (1) равен

* Теорема 4.2 является частным проявлением общего принципа «двойственности» для линейных программ. Для задачи 1 линейная программа, соответствующая требованиям 1) — 3), будет иметь

вид: 1) $\max \sum_{i,j=1}^m a_{ij} \xi_{ij}$; 2) $\sum_{i=1}^m \xi_{ij} \leq 1$ ($j = 1, \dots, m$); 3) $\sum_{j=1}^m \xi_{ij} \leq 1$

($i = 1, \dots, m$), где ξ_{ij} принимает значения 0 или 1 в зависимости от того, назначен i -й кандидат на j -е место или нет.

минимуму суммы $\sum_{i=1}^m u_i + \sum_{j=1}^m v_j$ задачи 1'. Пусть это общее значение для сумм достигается на парах $u_i, v_{j(i)}$, тогда

$u_i + v_{j(i)} = a_{ij(i)} \quad (i = 1, 2, \dots, m)$
и соответствующая подстановка

$$\pi_0 = \begin{pmatrix} 1 & 2 & \dots & i & \dots & m \\ j(1) & j(2) & \dots & j(i) & \dots & j(m) \end{pmatrix}$$

дает решение задачи об оптимальном назначении.

Доказательство. Доказательство носит конструктивный характер и дает алгоритм, позволяющий найти оптимальное решение за относительно малое количество шагов. Во-первых, заметим, что из условий (4.5) и (4.6) непосредственно следует для любой подстановки π неравенство

$$\sum_{i=1}^m a_{i, \pi(i)} \leq \sum_{i=1}^m (u_i + v_{\pi(i)}) = \sum_{i=1}^m u_i + \sum_{j=1}^m v_j, \quad (4.7)$$

из которого видно, что

$$\min \left(\sum_{i=1}^m u_i + \sum_{j=1}^m v_j \right) = M' \quad \text{существует, а для}$$

$$\max_{\pi \in S_m} \sum_{i=1}^m a_{i, \pi(i)} = M \quad \text{выполняется условие}$$

$$M = \max_{\pi \in S_m} \sum_{i=1}^m a_{i, \pi(i)} \leq \min \left(\sum_{i=1}^m u_i + \sum_{j=1}^m v_j \right) = M'. \quad (4.8)$$

Покажем, что $M = M'$. Для этого по заданному набору $\{u_i, v_j\}$, удовлетворяющему (4.6), определим множество упорядоченных пар индексов $i, j \in (1, 2, \dots, m)$

$$\{(i, j) / u_i + v_j = a_{ij}\} = P.$$

Всегда можно набор $\{u_i, v_j\}$ выбрать таким, что $P \neq \emptyset$. Например, полагая $u_i = \max_{1 \leq j \leq m} a_{ij}$, $v_j = 0$, получим, что все $(i, j) \in P$. Пусть P_i — множество всех таких индексов

j , что $u_i + v_j = a_{ij}$, тогда $P = \bigcup_{i=1}^m P_i$. Множеству P естественно сопоставить граф $G_P = (X, Y, \Gamma)$, в котором

$$X = \{x_i (i = 1, 2, \dots, m)\}, \quad Y = \{y_j (j = 1, 2, \dots, m)\},$$

$$\Gamma x_i = y_j (i = 1, 2, \dots, m), \quad \text{если} \quad j \in P_i (i = 1, 2, \dots, m).$$

Если граф G_P допускает максимальное паросочетание Π , $d(\Pi) = |X| = m$ и $(x_i, y_{j(i)})$ — дуги этого паросочетания, то подстановка $\pi(i) = j(i)$ ($i = 1, 2, \dots, m$) дает решение задачи об оптимальном назначении, поскольку в этом случае $M = M'$.

В случае, когда $d(\Pi) < |X| = m$, из теоремы 1.4 следует, что существует множество $A_0 \subset X$, $|\Gamma A_0| < |A_0|$. Определим тогда новые значения переменных u_i, v_j :

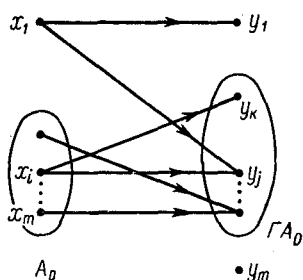


Рис. 97

$$u'_i = \begin{cases} u_i - 1 & (x_i \in A_0), \\ u_i & (x_i \notin A_0), \end{cases} \quad (4.9)$$

$$v'_j = \begin{cases} v_j + 1 & (y_j \in \Gamma A_0), \\ v_j & (y_j \notin \Gamma A_0). \end{cases}$$

Нетрудно проверить, что набор u'_i, v'_j ($i, j = 1, 2, \dots, m$) по-прежнему удовлетворяет условию двойственной задачи (рис. 97):

1) если $x_i \notin A_0$, то из формул (4.6), (4.9) следует неравенство $u'_i + v'_j \geq a_{ij}$;

2) если $x_i \in A_0, j \in \Gamma A_0$, то $v'_j + u'_i = (v_j + 1) + (u_i - 1) = u_i + v_j \geq a_{ij}$;

3) если $x_i \in A_0, j \notin \Gamma A_0$, то $(i, j) \notin P$ и $u_i + v_j > a_{ij}$ в силу целочисленности a_{ij} .

Для новых переменных сумма $\sum_{i=1}^m u'_i + \sum_{j=1}^m v'_j$ уменьшается по сравнению с суммой $\sum_{i=1}^m u_i + \sum_{j=1}^m v_j$ на величину

$|A_0| - |\Gamma A_0|$, в силу равенства

$$\begin{aligned}\sum_{i=1}^m u'_i + \sum_{j=1}^m v'_j &= \sum_{i=1}^m u_i - |A_0| + \sum_{j=1}^m v_j + |\Gamma A_0| = \\ &= \sum_{i=1}^m u_i + \sum_{j=1}^m v_j - (|A_0| - |\Gamma A_0|).\end{aligned}$$

По новым переменным u'_i, v'_j строим граф $G_{P'}$, где $P' = \{(i, j) | u'_i + v'_j = a_{ij}\}$ и проверяем условия (1.3) теоремы 1.4. Если они выполняются, то можно построить (используя один из рассмотренных выше алгоритмов) максимальное паросочетание, которое и дает оптимальный вариант задачи о назначении. В противном случае повторяем предыдущую процедуру и находим u''_i, v''_j , уменьшая на целое положительное число сумму $\sum_{i=1}^m u_i + \sum_{j=1}^m v_j$, сохра-

няя при этом условие $u''_i + v''_j \geq a_{ij}$ ($i, j = 1, 2, \dots, m$). В силу конечности множества индексов $i, j = 1, 2, \dots, m$ и целочисленности переменных $u_i, v_j \in \mathbb{Z}_+$, указанный процесс оборвется на конечном шаге и даст решение задачи об оптимальном назначении. ■

Задача 3. Показать, что и в случае нецелых значений a_{ij} ($i, j = 1, 2, \dots, n$) элементов матрицы эффективности A теорема 4.2 остается справедливой.

Задача 4. Доказать, что при оптимальном назначении (задача 1) хотя бы одно лицо обязательно будет выполнять ту работу, на которой его производительность максимальна

Работу алгоритма решения задачи об оптимальном назначении проиллюстрируем конкретным примером.

Пусть матрица эффективности $A = ((a_{ij}))$ имеет вид

$$A = \begin{vmatrix} 4 & 7 & 5 & 9 \\ 6 & 7 & 6 & \overline{2} \\ 8 & 9 & 13 & 10 \\ 5 & 7 & \overline{6} & 4 \end{vmatrix}$$

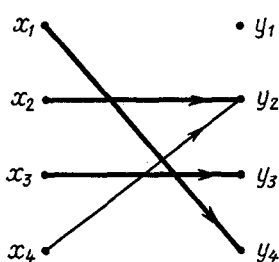
Полный перебор всех возможностей потребует расчета $m! = 24$ вариантов. Построим набор чисел u_i, v_j , положив $v_j = 0$ ($j = 1, 2, 3, 4$), а $u_i = \max_{j \in \{1, 2, 3, 4\}} a_{ij}$ ($i = 1, 2, 3, 4$).

Результат такого выбора переменных представим в табл. 2

ТАБЛИЦА 2

	$u_i \backslash v_j$	y_1	y_2	y_3	y_4
		0	$\uparrow 0$	0	$\uparrow 0$
x_1	9 $\downarrow$	4	7	5	<u>9</u>
x_2	7 $\downarrow$	6	<u>7</u>	6	2
x_3	13	8	9	<u>13</u>	10
x_4	7 $\downarrow$	5	<u>7</u>	6	4

Элементы матрицы A , равны $\max_{1 \leq j \leq 4} a_{ij} (i = 1, 2, 3, 4)$, под-



• y_1 черкнуты.

Строим граф G_P по множеству $P = \{(i, j), u_i + v_j = a_{ij}\}$. В данном случае $P = \{(1, 4), (2, 2), (3, 3), (4, 2)\}$, соответствующий граф G_P изображен на рис. 98; максимальное паросочетание графа $\Pi = \{x_1, y_4), (x_2, y_2), (x_3, y_3)\}$,

$$d(\Pi) = 3 < |X|.$$

Рис. 98

Находим множество $A_0 \subset X$ такое, что $|\Gamma A_0| < |A_0|$.

В данном случае $A_0 = \{(x_1, x_2, x_4)$, тогда $\Gamma A_0 = \{y_2, y_4\}$. Отмечаем в табл. 2 стрелками соответствующие переменные u_i и v_j и строим по ним новые переменные u'_i, v'_j по формуле (4.9), в результате чего получаем новую таблицу (табл. 3).

ТАБЛИЦА 3

	$u'_i \backslash v'_j$	y_1	y_2	y_3	y_4
		0	1	0	1
x_1	8	4	<u>7</u>	5	<u>9</u>
x_2	6	<u>6</u>	<u>7</u>	<u>6</u>	2
x_3	13	8	9	<u>13</u>	10
x_4	6	5	<u>7</u>	<u>6</u>	4

Строим граф $G_{P'}$, где $P' = \{(i, j), u_i' + v_j' = a_{ij}\}$, подчеркивая в таблице соответствующие элементы a_{ij} двойной чертой. На графе G_P (рис. 99) ищем максимальное паросочетание. Здесь удобно воспользоваться венгерским алгоритмом и заметить, что у графа G_P с паросочетанием Π , указанным выше, имеется тонкая чередующаяся

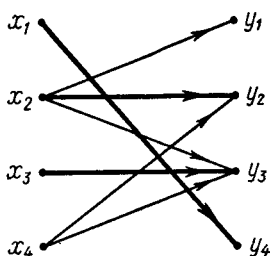


Рис. 99

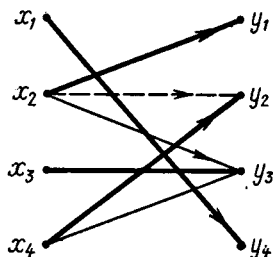


Рис. 100

цепь $\{(y_1, x_2), (x_2, y_2), (y_2, x_4)\}$. Перекрашивая нужные дуги, получаем новое паросочетание, очевидно, максимальное, так как $d(\Pi') = d(\Pi) + 1$, а $d(\Pi) = 3 = |X| - 1$ (рис. 100). Определяемая паросочетанием Π' подстановка

$$\pi = \begin{pmatrix} 1 & 2 & 3 & 4 \\ 4 & 1 & 3 & 2 \end{pmatrix}$$

дает оптимальное назначение $\pi(i) = j(i)$ ($i = 1, 2, 3, 4$).

З а м е ч а н и е. Теорема двойственности позволяет произвести косвенную проверку правильности полученного результата:

$$\sum_{i=1}^m a_{ij(i)} = 9 + 6 + 7 + 13 = 35,$$

$$\sum_{i=1}^m u_i + \sum_{j=1}^m v_{j(i)} = (8 + 6 + 13 + 6) + (1 + 1) = 35.$$

4. Задача об оптимальном назначении (задача II). Количество различных типов задач на оптимальное назначение определяется в конечном счете количеством различных оптимизационных параметров, представляющих практический интерес. Так, в предыдущем пункте в качестве

такого параметра была выбрана суммарная эффективность назначения

$$\sigma(\pi) = \sum_{i=1}^m a_{i, \pi(i)},$$

где π — некоторая подстановка множества индексов $X = \{1, 2, \dots, m\}$, $\pi \in S_m$, и задача I заключалась в нахождении оптимального значения параметра σ_0 , который равен $\max_{\pi} \sigma(\pi)$ по всем $\pi \in S_m$.

Для задачи II в качестве оптимизационного параметра выберем

$$\sigma(\pi) = \min_{i \in X} a_{i, \pi(i)} \quad (4.10)$$

и поставим вопрос о нахождении оптимального параметра σ_0 , равного $\max_{\pi \in S_m} \sigma(\pi)$. Иными словами, рассмотрим такое

назначение $\pi: i \rightarrow \pi(i)$ ($i = 1, 2, \dots, m$), при котором самое «узкое» место его будет наибольшим среди всевозможных назначений. Подобная задача возникает, например, при определении наибольшей скорости движения конвейера, которая определяется «узким» местом $\min_{1 \leq i \leq m} a_{i, \pi(i)}$ при усло-

вии, что a_{ij} выражает собой число единиц продукции, производимое в единицу времени рабочим i , если он выполняет работу j .

Решение поставленной задачи легко сводится к нахождению наибольшего паросочетания, причем здесь удобнее вести рассуждения в терминах теоремы 1.3.

Пусть π — произвольная подстановка, тогда сопоставим матрице $A = ((a_{ij}))_{m \times m}$ матрицу A_π порядка $m \times m$, состоящую из нулей и единиц так, что

$$a_{ij}^\pi = \begin{cases} 1, & \text{если } a_{ij} > \min_{1 \leq i \leq m} a_{i, \pi(i)}, \\ 0, & \text{если } a_{ij} \leq \min_{1 \leq i \leq m} a_{i, \pi(i)}. \end{cases} \quad (4.11)$$

Теорема 4.5. Если в матрице A_π существует m независимых единичных элементов, то подстановку π можно «улучшить» до подстановки π' такой, что

$$\min_{1 \leq i \leq m} a_{i, \pi'(i)} > \min_{1 \leq i \leq m} a_{i, \pi(i)}. \quad (4.12)$$

Подстановка π' будет отвечать одному из возможных максимальных паросочетаний Π ($d(\Pi) = m$) в графе G_π , струк-

тура множества дуг в котором определяется условиями (4.11).

Если максимальное число независимых единичных элементов меньше, чем m , $m = |X|$, то подстановка π оптимальна.

Доказательство. Первое утверждение вытекает из определения независимости элементов бинарной матрицы и равенства (4.11).

Для доказательства второго рассмотрим сначала частный случай, когда в матрице A_π существует строка i_0 , не содержащая ни одной единицы. Это означает по построению A_π , что для данной строки все элементы

$$a_{i_0 j} \leq \min_{1 \leq i \leq m} a_{i \pi(i)} \quad (4.13)$$

и, в частности, на этой строке $\min_{1 \leq i \leq m} a_{i \pi(i)} = a_{i_0 \pi(i_0)}$.

А тогда для любой другой подстановки π' выполнение условия (4.12) влечет за собой неравенство

$$a_{i_0 \pi'(i_0)} \geq \min_{1 \leq i \leq m} a_{i \pi'(i)} > a_{i_0 \pi(i_0)},$$

которое заведомо не выполняется в силу (4.13).

Для доказательства второго утверждения в общем случае остается заметить, что любая другая подстановка π' , для которой хотя бы один элемент $a_{ij}^{\pi'} = 0$ не может увеличить, как это видно из (4.11) и определения A_π , величины $\min_{1 \leq i \leq m} a_{i \pi'(i)}$. Подстановки π' , отличной от π и использующей только единичные элементы матрицы A_π , не существует в силу условия теоремы. ■

Из доказательства теоремы непосредственно видно, что алгоритм решения задачи в существенной своей части сводится к проверке условий теоремы 1.4 и нахождению максимального паросочетания.

Здесь уместно вновь вернуться к нахождению максимального паросочетания, причем интересно проследить, как алгоритм расчета максимального потока изменится с учетом специального характера графа и использованием лишь матрицы A_π в теореме 4.5.

Как нетрудно увидеть, задача о максимальном потоке для сети, состоящей из двудольного графа $G = (X, Y, \Gamma)$, дополненного входом и выходом, сводится к построению максимального множества независимых единичных элементов в матрице A_G порядка $m \times n$, где $m = |X|$, $n = |Y|$.

Рассмотрим таблицу T , в которой клетки с номерами (ij) ($i \in (1, 2, \dots, m)$, $j \in (1, 2, \dots, n)$), такими, что $a_{ij} = 1$ называются *допустимыми*. Клетки таблицы T , не лежащие на одной строке или в одном столбце, называются *независимыми*. Допустимые клетки таблицы T , образующие независимое множество, соответствуют множеству независимых единичных элементов матрицы A_G .

А л г о р и т м Ф о р д а — Ф а л к е р с о н а в
с л у ч а е м а т р и ч н о г о п р е д с т а в л е н и я
т р а н с п о р т н о й с е т и д л я п о с т р о е н и я
н а и б о л ь ш е г о п а р о с о ч е т а н и я
в д в у д о л ь н о м г р а ф е

1°. Построить в T произвольное множество независимых клеток, например, просматривая последовательно строки, начиная с первой и помещая единицу в любой допустимой клетке строки с последующим вычеркиванием строки и столбца, содержащих эту единицу. Пометить, например, знаком ($\times$) те строки, которые не содержат единиц.

2°. Просмотреть все помеченные строки, отыскать все допустимые клетки, не содержащие единицы, и пометить индексом соответствующей строки непомеченные столбцы, содержащие такие клетки. Если будет помечен столбец, не содержащий единиц, перейти к 4°. В противном случае — к 3°.

3°. Просмотреть помеченные столбцы и найти в них клетку с единицей. Строки, содержащие такие клетки и еще не помеченные, пометить номером соответствующего столбца. Если новых меток строкам присвоить нельзя, перейти к 5°. В противном случае — к 2°.

4°. Увеличить множество независимых клеток на единицу с помощью следующей операции, называемой «прорывом». В столбце, не содержащем единицы и только что помеченном, поставить единицу в клетке с меткой этого столбца; затем в строке, в которую поставлена единица и не помеченной меткой ($\times$), убрать единицу из клетки, соответствующей метке данной строки; и, наконец, в столбце, содержащем клетку, из которой убрана единица, поставить единицу в клетку, соответствующую метке данного столбца. Перейти к 2°.

5°. Закончить работу. Имеющееся размещение независимых клеток с единицами определяет в матри-

це A_G максимальное множество независимых элементов.

Обоснование алгоритма

Прежде всего заметим, что операция «прорыва» в п. 4⁰ по своему построению увеличивает имеющееся множество независимых клеток (с единицами) на одну клетку. Причем, применять эту операцию можно лишь конечное число раз, а именно, не более чем число строк, первоначально помеченных знаком ($\times$). В самом деле, при «прорыве» перемещение единицы по «вертикали» (без нарушения независимости клеток с единицами) возможно лишь в клетку, расположенную в строке с меткой ($\times$) (см. п. 1⁰ и определение независимых клеток).

В силу конечности таблицы и условия (п. 2⁰, 3⁰), обеспечивающего однократную маркировку строк и столбцов через конечное число шагов реализации алгоритма, появление новых меток прекратится. При этом, в силу условий п. 3⁰, все возможные случаи «прорыва», увеличивающего имеющиеся на данном этапе множество независимых клеток на единицу, были осуществлены, так что получившееся размещение независимых клеток не может быть увеличено и является максимальным.

Рассмотрим работу алгоритма на следующем примере. Пусть исходная транспортная сеть задана таблицей (рис. 101), в которой допустимые клетки не заштрихованы.

После применения первого правила алгоритма получаем множество независимых клеток, в которых расположены единицы. После применения второго и третьего правил алгоритма получаем соответствующую разметку столбцов и строк (рис. 102), причем пятый столбец, не содержащий единиц, помечается индексом первой строки ($\odot$). Теперь можно устроить «прорыв», перемещая 1 из клетки (1, 2) указанным выше способом; на рис. 103 изображено получившееся раз-

	1	2	3	4	5	6	7	
1		1						
2	1							
3			1					
4				1				
5								$\times$
6								$\times$
7						1		

Рис. 101

мещение независимых единиц в клетках (1,5), (2,1), (3,3) (4,4), (5,2) и (7,6). Просматривая вновь помеченные строки, находим столбец 7, не содержащий единиц, и поэтому до-

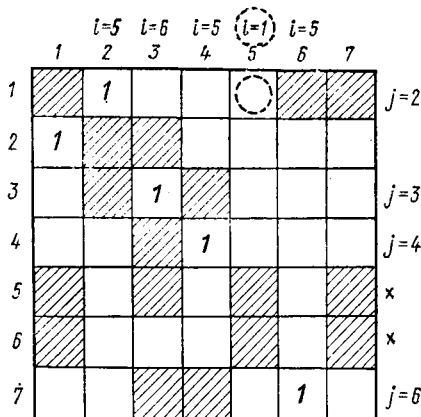


Рис. 102

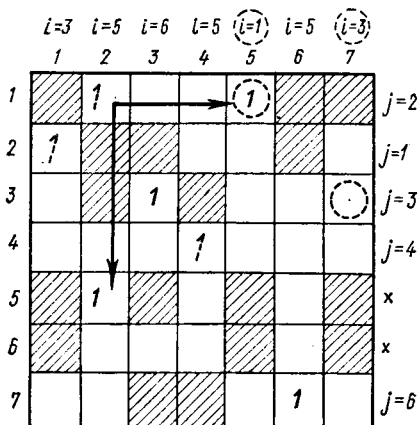


Рис. 103

пускающий прорыв (рис. 104). Полученное размещение единиц в клетках (1,5), (2,1), (3,7), (4,4), (5,2), (6,3), (7,6) максимально.

Задача 5. Построить матрицы смежности для графов G_1 и G_2 (см. рис. 94, 95) и указать максимально возможное размещение единиц, используя алгоритм Форда — Фалкерсона в матричной форме

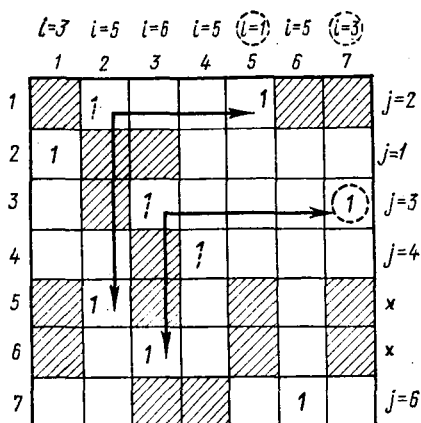


Рис. 104

Алгоритм решения задачи об оптимальном назначении (задача II)

1°. По исходной матрице A и произвольной подстановке $\pi \in S_n$ составить матрицу A_π , состоящую из нулей и единиц, причем так, что

$$a_{ij}^\pi = \begin{cases} 1, & \text{если } a_{ij} > \min_{1 \leq i \leq n} a_{i\pi(i)}, \\ 0, & \text{если } a_{ij} \leq \min_{1 \leq i \leq n} a_{i\pi(i)}. \end{cases}$$

2°. Найти в матрице A_π максимально возможное количество независимых единичных элементов. Если число таких элементов меньше, чем n , то перейти к 4°, если нет, то к 3°.

3°. По n независимым единичным элементам определить подстановку $\pi'(i) = j$ ($i = 1, 2, \dots, n$), j — номер столбца, содержащего единичный элемент строки i , и перейти к 1°.

4°. Закончить работу.

Подстановка, полученная в конце работы алгоритма и отвечающая набору независимых единичных элементов, является оптимальной.

Алгоритм решения задачи II оптимального назначения проиллюстрируем следующим примером. Пусть матрица

эффективности A представлена табл. 4, а исходная подстановка

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

ТАБЛИЦА 4

2	6	4	5	<u>9</u>	2	3
<u>5</u>	3	1	8	8	1	4
7	0	<u>5</u>	2	5	6	9
5	6	1	<u>7</u>	4	7	8
0	<u>8</u>	2	4	3	4	1
1	9	6	5	2	5	<u>3</u>
6	9	1	0	4	<u>6</u>	7

Числа, стоящие на пересечении строки и столбцов соответственно с номерами i , $\pi(i)$ ($i=1, 2, \dots, 7$), подчеркнуты. Для исходной матрицы A

$$\sigma(\pi) = \min_{1 \leq i \leq 7} a_{i, \pi(i)} = 3,$$

поэтому матрица A_π принимает вид, представленный таблицей на рис. 105, где незаштрихованные клетки соответствуют единичным элементам матрицы A_π , а заштрихованные — нулевым элементам.

В матрице A_π отыскиваем максимально возможное количество независимых клеток. Удобно при этом по возможности как можно больше клеток подстановки π . На рис. 105 такое назначение указано единицами. Поскольку матрица A_π совпадает с матрицей A_G , представленной таблицей на рис. 103, то, применяя описанный алгоритм Форда — Фалкерсона, можно указать максимальное множество из 7 клеток (см. рис. 104) и тем самым определить новую подстановку

$$\pi' = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 \\ 5 & 1 & 6 & 4 & 2 & 3 & 6 \end{pmatrix}.$$

такую, что

$$\sigma(\pi') = \min_{1 \leq i \leq 7} a_{i, \pi'(i)} = 6 > 3 = \min_{1 \leq i \leq 7} a_{i, \pi(i)} = \sigma(\pi).$$

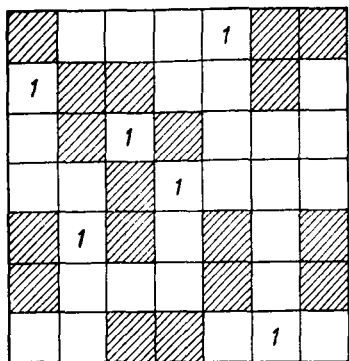


Рис. 105

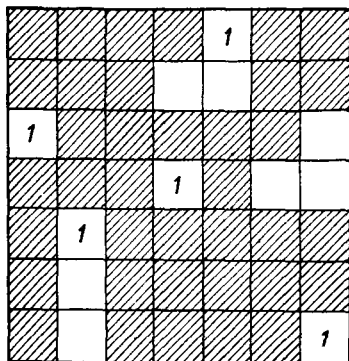


Рис. 106

Подстановка π' определяет матрицу $A_{\pi'}$ (рис. 106), для которой нетрудно определить, что максимальное независимое множество содержит $5 < 7$ клеток, и поэтому подстановка π' оптимальна.

Глава X

ГАМИЛЬТОНОВЫ МАРШРУТЫ И ЗАДАЧА О КОММИВОЯЖЕРЕ

В главе рассматривается одна из проблем теории графов, известная как «задача о коммивояжере». В теории графов эта проблема возникла под влиянием следующей практической задачи.

Район, который должен посетить бродячий торговец, содержит некоторое количество городов, расстояния между которыми известны. Требуется найти маршрут, проходящий через все пункты по одному разу и возвращающийся в исходный. Если таких маршрутов много, требуется найти кратчайший из них.

Сложность задачи проявляется, с одной стороны, в отсутствии достаточно общих теорем существования таких маршрутов, с другой стороны, в том, что нет пока эффективного алгоритма отыскания кратчайшего пути, который по объему вычислений существенно отличался бы от прямого перебора громадного числа вариантов (порядка $n!$, где n — число городов). В настоящее время для получения практических результатов при большом n используют не точные, а приближенные алгоритмы. В полной общности задача о коммивояжере формулируется в § 2.

§ 1. ГАМИЛЬТОНОВЫ ЦИКЛЫ И КОНТУРЫ

1. Существование гамильтоновых циклов в полностью неориентированном графе. Постановка задачи о коммивояжере предполагает, разумеется, наличие хотя бы одного маршрута, проходящего по всем городам, причем через каждый лишь по одному разу.

Определение 1.1. Пусть $G = (X, \bar{\Gamma})$ — полностью неориентированный граф. Цепь, проходящая через каждую вершину графа $G = (X, \bar{\Gamma})$ и притом только по одному разу, называется гамильтоновой. Цикл $\mu = \{(x_0, x_1),$

$(x_1, x_2), \dots, (x_{n-1}, x_n)\}$, проходящий через все вершины графа G и притом по одному разу через каждую (за исключением $x_0 = x_n$), называют гамильтоновым циклом.

На рис. 107 представлен один из гамильтоновых циклов

$$\mu_1 = \{1, 4), (4, 2), (2, 5), (5, 3), (3, 1)\}.$$

Те же ребра, расположенные в другом порядке, определяют цикл

$$\mu_2 = \{1, 3), (3, 5), (5, 2), (2, 4), (4, 1)\}.$$

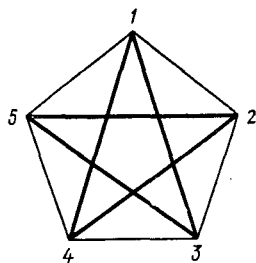


Рис. 107

Заметим, что в графе, изображенном на рис. 107, каждому гамильтонову циклу μ отвечает подстановка

$$\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ \pi(1) & \pi(2) & \dots & \pi(n) \end{pmatrix},$$

которая циклически переставляет элементы множества $X = \{1, 2, \dots, n\}$, т. е. для некоторого элемента $i \in X$ совокупность элементов $\{i, \pi(i), \dots, \pi^{n-1}(i), \pi^n(i) = i\}$ исчерпывает все X .

Задача 1. Доказать, что любая циклическая подстановка в множестве вершин $X = \{1, 2, \dots, n\}$ полного графа G определяет однозначно гамильтонов цикл.

Лемма 1.1. Если в графе $G = (X, \bar{G})$ ($|X| = n$) существует элементарная цепь $\mu = \{(x_0, x_1), \dots, (x_{m-1}, x_m)\} = \{x_0, x_1, \dots, x_m\}^*$ длины $m \leq n-1$ такая, что при некотором $i \in (1, 2, \dots, n)$ ребра $(x_0, x_i), (x_m, x_{i-1}) \in \bar{G}$ (рис. 108), то цепи μ можно сопоставить элементарный цикл $z_\mu = \{x_0, x_i, x_{i+1}, \dots, x_m, x_{i-1}, \dots, x_0\}$.

Доказательство. Очевидно, что z_μ — элементарный цикл. ■

Обозначим через G_μ подграф графа G , определяемый подмножеством вершин $\{x_0, x_1, \dots, x_m\} \subset X$. Тогда, по построению, цикл z_μ будет гамильтоновым в подграфе G_μ (см. рис. 108).

* В этом параграфе помимо обычного обозначения $\{(x_0, x_1), (x_1, x_2), \dots, (x_{n-1}, x_n)\}$ цепи μ будет использоваться и такое: $\{x_0, x_1, \dots, x_m\} = \mu$.

Определение 1.2. Элементарная цепь $\mu = \{x_0, x_1, \dots, x_m\}$ имеет тип цикла, если в подграфе G_μ есть гамильтонов цикл.

Нетрудно дать критерий того, что элементарная цепь имеет тип цикла, используя понятие степени вершин графа.

Определение 1.3. Элементарная цепь $\mu = \{x_0, x_1, \dots, x_m\}$ — полная, если ее нельзя продолжить добавлением новых ребер к какому-нибудь из ее концов, не нарушая свойства элементарности.

Лемма 1.2. Полная цепь $\mu = \{x_0, x_1, \dots, x_m\}$ длины m имеет тип цикла, если $\text{st } x_0 + \text{st } x_m \geq m + 1$.

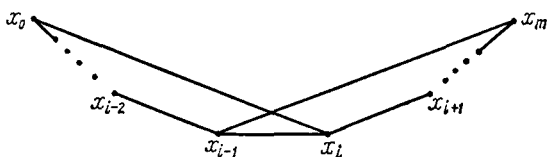


Рис. 108

Доказательство. Рассмотрим полную цепь $\mu = \{x_0, x_1, \dots, x_m\}$ и подграф G_μ (см. рис. 108). Поскольку цепь полная, то степени вершин $\text{st}_\mu x_0$ и $\text{st}_\mu x_m$ в подграфе G_μ совпадают со степенями $\text{st} x_0$ и $\text{st} x_m$ этих же вершин в графе G :

$$\text{st}_\mu x_0 = \text{st } x_0, \quad \text{st}_\mu x_m = \text{st } x_m, \quad (1.1)$$

иными словами, все ребра от вершин x_0 и x_m должны идти к вершинам графа $G_\mu = G(x_0, x_1, \dots, x_m)$.

Для подграфа G_μ условие

$$\text{st}_\mu x_m > m - \text{st}_\mu x_0 \quad (1.2)$$

очевидным образом обеспечивает для хотя бы одного ребра (x_0, x_i) существование ребра (x_{i-1}, x_m) . В силу леммы 1.1 цепь $\{x_0, x_i, \dots, x_m\}$ определяет гамильтонов цикл в графе G_μ . Поэтому из соотношений (1.1) и (1.2) следует утверждение леммы. ■

Лемма 1.3. Наидлиннейшая элементарная цепь в связном графе имеет тип цикла тогда и только тогда, когда граф имеет гамильтонов цикл.

Доказательство. Действительно, если $\mu = \{x_0, x_1, \dots, x_m\}$ — элементарная цепь наибольшей дли-

ны имеет тип цикла, то граф G имеет гамильтонов цикл. В противном случае нашлись бы такая вершина $x \in X$, что $x \notin G_\mu$ и, в силу связности графа G , ребро (x, x_i) для некоторого $i \in (0, 1, \dots, m)$, соединяющее вершину x с цепью μ . Поскольку в G_μ существует гамильтонов цикл z_μ и в нем ребер, по крайней мере, на одно больше, чем в μ , то элементарная цепь $\{x_1, x_i\} \cup (z_\mu \setminus \{x_i, x_{i+1}\})$ имеет длину, превосходящую длину цепи μ , что невозможно, так как μ , по условию, цепь наибольшей длины в графе G (рис. 109).

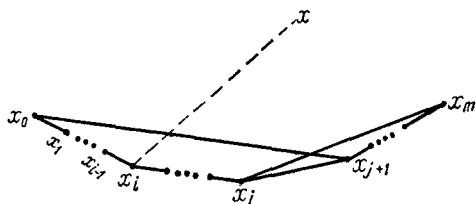


Рис. 109

Обратное утверждение очевидно, поскольку длина максимальной элементарной цепи в графе $G = (X, \Gamma)$ не превосходит числа $n - 1$ ($n = |X|$), а гамильтонов цикл графа содержит цепь длины $n - 1$. ■

Объединяя леммы 1.2 и 1.3, можно высказать следующую альтернативу.

Теорема 1.1. В связном графе а) либо имеется гамильтонов цикл, б) либо длина наибольших элементарных цепей графа удовлетворяет неравенству

$$m \geq \text{st } x_0 + \text{st } x_m, \quad (1.3)$$

где x_0 и x_m — начальная и конечная вершины элементарной цепи $\mu = \{x_0, x_1, \dots, x_m\}$ длины m .

Заменяя в неравенстве (1.3) сумму $\text{st } x_0 + \text{st } x_m$ на $\min (\text{st } x_0 + \text{st } x_m)$, где минимум берется по всем парам вершин $\{x_0, x_m\}$, $x_0 \neq x_m \in X$, из теоремы 1.1 получаем следующий результат.

Теорема 1.2. В графе без гамильтоновых циклов длина m его наибольших элементарных цепей удовлетворяет неравенству

$$m \geq \text{st } x_1 + \text{st } x_2, \quad (1.4)$$

где $\text{st } x_1 = \min_{x \in X} \text{st } x$, $\text{st } x_2 = \min_{\substack{x \neq x_1 \\ x \in X}} \text{st } x$.

Следствие 1. Если в графе $G = (X, \bar{\Gamma})$ ($|X| = n$) для любой пары вершин x_i и x_j выполняется соотношение

$$\text{st } x_i + \text{st } x_j \geq n - 1, \quad (1.4')$$

то граф G имеет гамильтонову цепь. Если же

$$\text{st } x_0 + \text{st } x_m \geq n, \quad (1.4'')$$

то граф G имеет гамильтонов цикл.

Следствие 2. Если степень каждой вершины графа G удовлетворяет соотношению

$$\text{st } x_i \geq \frac{1}{2} |X| = \frac{1}{2} n,$$

то граф имеет гамильтонов цикл.

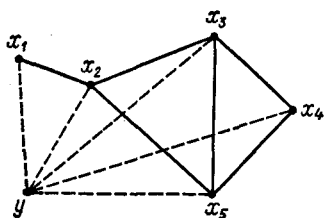


Рис. 110

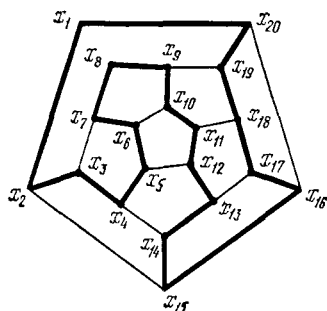


Рис. 111

Следствие 3. В полном графе с числом вершин $n \geq 2$ всегда существует гамильтонов цикл (см. рис. 107).

Как видно из условия (1.4'), наличие в графе вершин с большими степенями благоприятствует существованию в графе гамильтонова цикла. Это обстоятельство позволяет свести задачу отыскания гамильтоновой цепи в графе $G = (X, \bar{\Gamma})$ к отысканию гамильтонова цикла в графе $G' = (X', \bar{\Gamma}')$, который получается добавлением новой вершины y к множеству X и новых ребер, проведенных из y во все вершины $x \in X$.

Очевидно, что наличие цикла в графе G' равносильно существованию гамильтоновой цепи в графе G , и наоборот.

На рис. 110 гамильтоновой цепи $\mu = \{x_1, x_2, \dots, x_5\}$ графа $G = (\{x_1, x_2, \dots, x_5\}, \bar{\Gamma})$ отвечает гамильтонов цикл

$z_\mu = \{x_1, \dots, x_5, y, x_1\}$ в графе $G' = (\{x_1, x_2, \dots, x_5, y\}, \bar{\Gamma})$.

Приведем два классических примера гамильтоновых циклов. Правда, в этих примерах вопросы существования не охватываются рассмотренными результатами и построение гамильтоновых циклов требует для своего решения в каждом случае отдельного подхода.

1. Замкнутое кругосветное путешествие. Пусть даны 20 городов. Будем считать для простоты, что они располо-

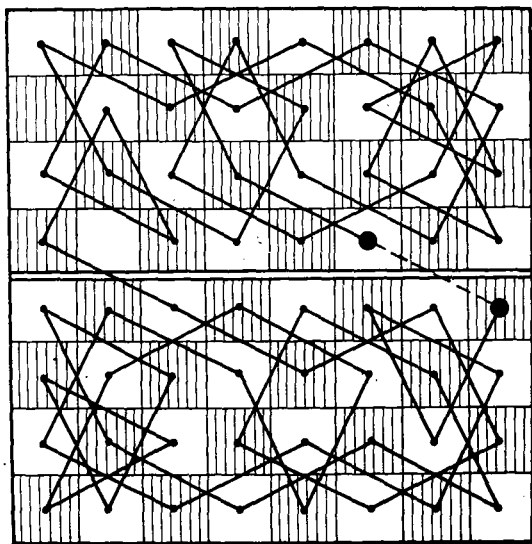


Рис. 112

жены в вершинах правильного додекаэдра, условно изображающего Землю. Тогда оказывается возможным осуществить «кругосветное путешествие», двигаясь по ребрам додекаэдра, побывав в каждом городе один только раз. Один из маршрутов приведен на рис. 111.

2. Ход шахматного коня. Можно обойти шахматную доску так, чтобы побывать в каждой клетке ровно по одному разу. Один из возможных циклов изображен на рис. 112.

Методы решения приведенных примеров, как правило, опираются на симметрию возникающих графов и не переносятся на более общие графы.

2. Гамильтоновы контуры. Для произвольных графов, как показывают простейшие примеры (рис. 113), еще более трудно, чем для полностью неориентированных, высказать какие-либо достаточные условия существования гамильтонова контура (пути). Так, графы изображенные на рис. 113, *а, б*, не имеют гамильтоновых контуров, в то время как соответствующие им полностью неориентированные графы удовлетворяют условию $\text{st } x_i \geq \frac{1}{2} |X|$, обесп-

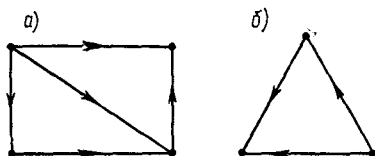


Рис. 113

печивающему существование гамильтоновых циклов. Тем не менее справедлива следующая теорема.

Теорема 1.3. В графе G , любая пара вершин которого соединена хотя бы в одном направлении*,

всегда существует гамильтонов путь.

Доказательство. Пусть $\mu = \{x_1, x_2, \dots, x_p\}$ — путь длины $p - 1$ ($p > 1$), все вершины которого различны, и пусть y — вершина графа, не принадлежащая μ . Тогда в силу свойств графа G можно составить путь вида $\mu_k = \{x_1, x_2, \dots, x_k, y, x_{k+1}, \dots, x_p\}$ при некотором $k \in \{0, 1, \dots, p\}$, содержащий вершину y (здесь $\mu_0 = \{y, x_1, \dots, x_p\}$, $\mu_p = \{x_1, \dots, x_p, y\}$).

Таким образом, шаг за шагом можно построить путь, содержащий все вершины графа по одному разу. ■

З а м е ч а н и е. Все результаты этого параграфа автоматически переносятся на полностью неориентированные графы. В частности, для полного полностью неориентированного графа всегда существует гамильтонов цикл.

Г а м и л ь т о н о в ы к о н т у р ы и ф а к т о р г р а ф а

Для выяснения существования гамильтоновых путей в графе можно использовать теорию транспортных сетей. Для этого понадобится определение.

Определение 1.4. Суграф $G' = (X, \Gamma')$ графа $G = (X, \Gamma)$ называется фактором, если каждая вершина его имеет полустепени захода и исхода, равные единице.

* Такой граф называется полным.

Очевидно, что каждый гамильтонов контур является фактором, но обратное не верно, ибо фактор может состоять из нескольких непересекающихся контуров.

Теорема 1.4. *Необходимое и достаточное условие существования факторов для графа $G = (X, \Gamma)$ состоит в том, что $|S| \leq |\Gamma S|$ для любого множества $S \subset X$.*

Доказательство. Рассмотрим двудольный граф $G' = (X, \bar{X}, \Gamma')$ в котором: 1) $|\bar{X}| = |X|$; 2) $\bar{x}_j \in \Gamma' x_i$ при условии, что $x_j \in \Gamma x_i$ ($i, j = 1, 2, \dots, n$). Тогда условие данной теоремы эквивалентно тому, что в графе $G' = (X, \bar{X}, \Gamma')$ существует паросочетание, отображающее множество X на $\bar{X}$ (теорема 1.4 гл. IX).

Отождествляя дуги паросочетания $(x_i, \bar{x}_j) \in G'$ с дугами $(x_i, x_j) \in G$, получаем искомый фактор в графе $G = (X, \Gamma)$. ■

Следствие. *Если в графе $G = (X, \Gamma)$ выполняется равенство $|\Gamma x| = |\Gamma^{-1}x| = t$ для любой вершины $x \in X$, то дуги этого графа можно распределить по t непересекающимся множествам $\Pi_1, \Pi_2, \dots, \Pi_t$, каждое из которых образует фактор.*

Доказательство. Действительно, в двудольном графе $G' = (X, \bar{X}, \Gamma')$ будем иметь равенство

$$|\Gamma'x| = t, \quad |\Gamma^{-1}x| = t,$$

и поэтому, в силу леммы 4.1 гл. IX, можно отобразить множество X на $\bar{X}$ и тем самым определить множество дуг Π_1 в графе $G = (X, \Gamma)$. В суграфе, остающемся после удаления дуг множества Π_1 , можно отобразить X на $\bar{X}$ и тем самым определить множество дуг Π_2 и т. д. ■

Рассмотрим примеры. 1. Для графа, изображенного на рис. 114, не существует гамильтонова контура, поскольку для вершины $d \in X$ нарушено условие теоремы. Действительно,

$$|\{d\}| - |\Gamma d| = 1 - 0 = 1.$$

2. Для графа $G = (X, \Gamma)$ (рис. 115) условия теоремы выполнены, и поэтому соответствующий ему граф $G' = (X, \bar{X}, \Gamma')$ (рис. 116) допускает паросочетания с числом дуг, равным $5 = |X|$, например,

$$\Pi_1 = \{(a, \bar{d}), (e, \bar{b}), (b, \bar{c}), (c, \bar{a}), (d, \bar{e})\}.$$

Паросочетание Π_1 определяет в графе $G = (X, \Gamma)$ фактор

$$\Phi_1 = \{(a, d), (d, e), (e, b), (b, c), (c, a)\},$$

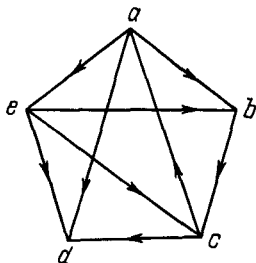


Рис. 114

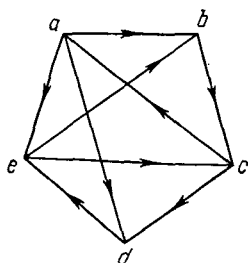


Рис. 115

который, как нетрудно видеть, является гамильтоновым контуром.

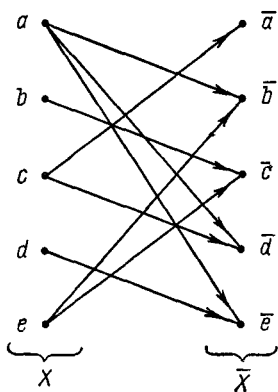


Рис. 116

Таким образом, задача отыскания фактора графа сводится к задаче о наибольшем паросочетании и допускает эффективное решение с помощью алгоритмов, рассмотренных в гл. IX. Для нахождения гамильтонова контура необходимо среди всех факторов рассмотреть лишь те, которые состоят из одного контура. Если таких факторов нет, то для построения гамильтонова контура используют «склейку» друг с другом различных факторов.

Гамильтоновы контуры и дефицит графа

Предыдущие результаты допускают обобщение и позволяют связать существование гамильтонова пути (контура) с дефицитом графа.

Определение 1.5. *Рассечением R графа $G = (X, \Gamma)$ называется такая совокупность элементарных путей или*

элементарных контуров, что 1) два пути рассечения не имеют общих вершин,

2) каждая вершина графа принадлежит одному из путей рассечения.

Среди различных путей рассечения будем различать элементарные контуры, которые обозначим через

$$\alpha_i = \{x_0^i, x_1^i, \dots, x_p^i = x_0^i\};$$

элементарные пути длины, не равной нулю, обозначаемые через

$$\beta_j = \{x_0^j, x_1^j, \dots, x_q^j\};$$

пути длины 0, т. е. совокупности отдельных вершин, которые обозначим

$$\gamma_k = \{x_k\}.$$

Обозначим через A^i множество вершин, принадлежащих контуру α_i , через B^j — множество вершин, принадлежащих пути β_j , через C^k — множество $\{x_k\}$; положим:

$$n(\alpha_i) = 0, \quad n(\beta_j) = 1, \quad n(\gamma_k) = 1.$$

Определение 1.6. Величина $\lambda(R)$ рассечения R ,

$$R = \{\alpha_1, \alpha_2, \dots, \alpha_p; \beta_1, \beta_2, \dots, \beta_q; \gamma_1, \gamma_2, \dots, \gamma_r\}$$

определяется формулой

$$\lambda(R) = \sum_{i=1}^p n(\alpha_i) + \sum_{j=1}^q n(\beta_j) + \sum_{k=1}^r n(\gamma_k).$$

Очевидно, что фактор графа является его рассечением величины $\lambda = 0$.

Теорема 1.5. Наименьшая величина рассечения графа $G = (X, \Gamma)$ равна дефициту этого графа, т. е.

$$\min_R \lambda(R) = \delta_0 = \max_{S \subset X} (|S| - |\Gamma S|).$$

Доказательство. Рассмотрим двудольный граф

$$G' = (X, \bar{X}, \Gamma'): \quad 1) |\bar{X}| = |X|, \quad 2) \bar{x}_j \in \Gamma' x_i, \text{ если}$$

$$x_j \in \Gamma x_i, \quad i, j \in (1, 2, \dots, n)$$

и рассечение

$$R = \{\alpha_1, \alpha_2, \dots, \alpha_p; \beta_1, \beta_2, \dots, \beta_q; \gamma_1, \gamma_2, \dots, \gamma_r\}.$$

Тогда контуру α_i в графе G' отвечает паросочетание, содержащее $|A^i|$ дуг; пути β_j отвечает паросочетание, содержащее $|B^j| - 1$ дуг. Объединяя все эти дуги, получим паросочетание Π с числом дуг

$$d(\Pi) = \sum_{i=1}^p |A^i| + \sum_{j=1}^q (|B^j| - 1).$$

Тогда справедливы следующие равенства:

$$\begin{aligned} |X| = n = \sum_{i=1}^p |A^i| + \sum_{j=1}^q |B^j| + \sum_{k=1}^r |C^k| &= d(\Pi) + \sum_{j=1}^q 1 + \\ &+ \sum_{k=1}^r |C^k| = d(\Pi) + \lambda(R), \end{aligned}$$

или

$$\lambda(R) = n - d(\Pi).$$

Отсюда, с учетом теоремы 1.5 гл. IX окончательно получаем:

$$\min_R \lambda(R) = n - \max_{\Pi} d(\Pi) = n - (n - \delta_0) = \delta_0. \blacksquare$$

Из теоремы 1.5 и определений 1.5, 1.6 непосредственно получаем следующие следствия.

Следствие 1. Если в графе $G = (X, \Gamma)$ существует гамильтонов контур, то его дефицит $\delta_0 = 0$.

Следствие 2. Если в графе G существует гамильтонов путь, то для дефицита δ_0 графа G справедливо неравенство $0 \leq \delta_0 \leq 1$.

Следствие 3. Если граф $G = (X, \Gamma)$ не содержит контуров, то его дефицит δ_0 равен наименьшему числу попарно непересекающихся элементарных путей (отдельная вершина при этом рассматривается как путь длины нуль).

Следствие 4. Если граф не содержит контуров, то необходимое и достаточное условие существования гамильтонова пути состоит в том, что $\delta_0 = 1$.

В качестве примера рассмотрим граф (см. рис. 114), не содержащий гамильтоновых контуров и с дефицитом $\delta_0 = 1$. В силу следствия 4 существует гамильтонов путь, который можно найти, рассмотрев наибольшее паросочетание

в графе $G' = (X, \bar{X}, \Gamma')$ (см. рис. 116), предварительно удалив из него дугу $(d, \bar{e})$. Искомый путь будет $\{e, b, c, a, d\}$.

Приведем теперь два примера практических задач, в которых возникает вопрос об отыскании гамильтонова пути.

1. Пусть $(x_1, x_2, \dots, x_n)$ — точки, изображающие некоторые действия. Точки x_i и x_j связываем дугой (x_i, x_j) , если только осуществление действия x_i перед x_j не влечет никаких потерь.

Оптимальный, т. е. с нулевыми потерями, порядок осуществления действий $(x_1, x_2, \dots, x_m)$ определяется одним из гамильтоновых путей графа.

2. Требуется указать такой порядок обработки n различных деталей, требующий двух последовательных технологических операций на двух станках, чтобы закончить работу как можно скорее. Известно время a_k , затрачиваемое для обработки k -й детали на первом станке и время b_k — на втором. Считается, что при оптимальном порядке i -я деталь обрабатывалась бы раньше, чем j -я в том и только в том случае, если

$$\min \{a_i, b_j\} \leq \min \{a_j, b_i\}.$$

При выполнении этого неравенства соединяются вершины x_i и x_j дугой (x_i, x_j) . В полученном графе находится гамильтонов путь. Если этот путь существует, то он укажет оптимальный порядок организации работы.

§ 2. ЗАДАЧА О КОММИВОЯЖЕРЕ

1. **Формулировка задачи.** Пусть A — квадратная матрица порядка $n \times n$ с неотрицательными коэффициентами

$A = ((a_{ij}))_{n \times n}$, $a_{ij} \geq 0$ ($i, j = 1, 2, \dots, n$); S_n — группа под-

становок порядка n , а $Z(n)$ — подмножество циклических подстановок ее. Элементы группы S_n будем обозначать

через $\pi = \begin{pmatrix} 1 & 2 & \dots & n \\ i_1 & i_2 & \dots & i_n \end{pmatrix}$. Каждой подстановке $\pi \in S_n$

сопоставим число

$$\sigma(\pi) = \sum_{i=1}^n a_{i, \pi(i)}. \quad (2.1)$$

Для заданной матрицы A , удовлетворяющей условиям

1) $a_{ij} \geq 0$ ($i, j = 1, 2, \dots, n$), 2) $a_{ii} = \infty$ ($i = 1, 2, \dots, n$),

найти циклическую подстановку $\pi_0 \in Z(n)$, для которой

$$\sigma(\pi_0) = \min_{\pi \in Z(n)} \sigma(\pi).$$

Будем обозначать иногда $\sigma(\pi_0)$ через $\sigma(A)$, подстановку π_0 называть оптимальной, матрицу A — матрицей «расстояний».

Геометрический смысл сформулированной задачи состоит в том, что элементы a_{ij} ($i \neq j$) матрицы A можно рассматривать как обобщенные длины дуг полного симметрического графа с числом вершин $|X| = n$, причем в отличие от матрицы смежности, когда равенство $a_{ij} = 0$ означает, что пара вершин $(x_i, x_j) \notin \Gamma$, в этой задаче равенство $a_{ij} = 0$ означает, что пара вершин x_i и x_j определяет дугу $(x_i, x_j) \in \Gamma$, имеющую нулевую обобщенную длину. Обобщенная длина дуг не является, вообще говоря, симметричной функцией вершин графа, что легко увидеть на конкретных примерах. Поскольку граф G — полный симметрический граф, то любая циклическая подстановка $\pi(i, \pi(i), \dots, \pi^{n-1}(i), \pi^n(i) = i)$ определяет в графе G гамильтонов контур $\mu_\pi = \{(i, \pi(i)), (\pi(i), \pi^2(i)), \dots, (\pi^{n-1}(i), i)\}$, а тогда задача о коммивояжере состоит в определении кратчайшего (оптимального по отношению к параметру σ) гамильтонова контура в полном графе.

Задача определения кратчайшего гамильтонова контура в произвольном графе $G = (X, \Gamma)$, длины дуг которого $b_{ij} \geq 0$, также легко сводится к задаче о коммивояжере. Для этого достаточно рассмотреть матрицу $A = ((a_{ij}))$ вида

$$a_{ij} = \begin{cases} b_{ij} & ((x_i, x_j) \in \Gamma), \\ \infty & ((x_i, x_j) \notin \Gamma). \end{cases} \quad i, j \in (1, 2, \dots, n) \quad (2.2)$$

Если при решении задачи получим $\sigma(\pi_0) = \infty$, то это означает, что в графе $G = (X, \Gamma)$ не существует ни одного гамильтонова контура.

Собственно задача о бродячем торговце, упомянутая во введении к этой главе, отвечает матрице A , определяемой соотношением (2.2) с дополнительными условиями:

1) A — симметричная матрица, так как $b_{ij} = b_{ji}$ — расстояния между городами;

2) коэффициенты матрицы A удовлетворяют неравенству «треугольника» $b_{ij} + b_{jm} \geq b_{im}$, $i, j, m \in (1, 2, \dots, n)$, $i \neq j \neq m$.

Как частный случай задача о коммивояжере содержит задачу об отыскании гамильтонова контура в произвольном графе $G = (X, \Gamma)$. Достаточно положить в матрице A

$$a_{ij} = \begin{cases} 0 & ((x_i, x_j) \in \Gamma), \\ \infty & ((x_i, x_j) \notin \Gamma). \end{cases} \quad i, j \in (1, 2, \dots, n)$$

Тогда гамильтонов контур в графе $G = (X, \Gamma)$ существует в том и только том случае, если оптимальное значение $\sigma(\pi_0) = 0$.

Отметим, наконец, связь задачи о коммивояжере с задачей 1 об оптимальном назначении, рассмотренной в гл. IX. Напомним, что в задаче о назначении параметр

$$\sigma'(\pi) = \sum_{i=1}^n a_{i, \pi(i)}$$

минимизировался по всем подстановкам из S_n , т. е. оптимальной была подстановка $\pi_1 \in S_n$ такая, что

$$\sigma(\pi_1) = \min_{\pi \in S_n} \sigma'(\pi).$$

В силу включения $Z(n) \subset S_n$ для любой подстановки $\pi \in Z(n)$ справедливо неравенство

$$\sigma(\pi) \geq \min_{\pi \in S_n} \sigma'(\pi).$$

Таким образом, решение задачи о назначении дает нижнюю оценку оптимальной длины гамильтонова контура π_0 , а именно:

$$\sigma(\pi_0) \geq \min_{\pi \in S_n} \sigma'(\pi).$$

2. Примеры практических задач, сводящихся к задаче о коммивояжере. Сформулированная в общем виде задача о коммивояжере часто встречается на практике там, где возникает вопрос о наилучшем упорядочении каких-либо действий, объектов, предметов и т. д.

Различные варианты задачи оптимального упорядочения объектов возникают при планировании рационального использования транспорта. Например, при определении маршрута развозки продуктов по торговым точкам города так, чтобы суммарная длина пробега машин была бы минимальной; при выборе минимальной по протяженности ав-

тобусной линии в городе, на ряде улиц которого возможно лишь одностороннее движение и фиксировано местоположение остановок автобуса.

Некоторые варианты этой задачи используются также при рациональной расстановке указателей в музее, когда известно местоположение каждого экспоната и расстояние между любыми двумя из них и требуется установить такой порядок осмотра, при котором затрачивается минимум времени на переходы между экспонатами.

Задача определения оптимального порядка обработки деталей в количестве n штук, если известно время t_{ij} переналадки станка для обработки i -й детали после обработки j -й детали, когда требуется установить такой порядок работы, при котором суммарное время на переналадку было бы минимальным, также основана на рассмотренной задаче о коммивояжере.

Ряд прикладных задач на нахождение оптимального гамильтонова цикла были рассмотрены в § 1 настоящей главы.

Нетрудно понять, что прямой способ перебора всех $(n - 1)!$ возможностей в задаче о коммивояжере высокой размерности n ($n \sim 10^2 - 10^3$) даже с использованием современных ЭВМ практически не применим.

Для частного вида матрицы A существуют достаточно эффективные алгоритмы. Так, например, для матрицы A , все элементы которой удовлетворяют неравенству треугольника, имеется удобный геометрический метод, основанный на элементарных аксиомах и теоремах геометрии (заметим, что если матрица A симметрична, то уже объем прямого перебора уменьшается в два раза, т. е. равен $\frac{(n-1)!}{2}$).

В произвольном случае алгоритмов, которые бы работали эффективно при $n \sim 100$, пока не существует. Наибольшей размерности, которой удалось достичь в расчете оптимального контура, является $n \sim 70$. В последнем случае был использован метод, основанный на модификации одного из наиболее эффективных методов решения задач о коммивояжере — методе «ветвей и границ».

§ 3. МЕТОД ВЕТВЕЙ И ГРАНИЦ

1. Общая схема метода. Среди методов нахождения оптимального решения экстремальных комбинаторных задач центральное место занимает в настоящее время метод

«ветвей и границ». Главная идея этого метода состоит в замене случайного перебора возможных решений целенаправленным. Существенным является наличие априорных критериев, с помощью которых в некоторых случаях удается в процессе перебора отбрасывать множества решений, в которых нет оптимального. Заметим при этом, что хотя при решении некоторых задач метод ветвей и границ не помогает избежать полного перебора, в большинстве случаев его применение к конкретным задачам значительно снижает объем вычислительной работы по сравнению со случайным полным перебором. Особенно это относится к случаю, когда достаточно построить приближенное решение с заданной степенью точности.

Перейдем теперь к точным формулировкам.

Экстремальной комбинаторной задачей называется задача о нахождении минимума функции

$$y = \sigma(z) \quad (z \in Z), \quad (3.1)$$

где Z — некоторое конечное множество.

В конкретных комбинаторных задачах Z представляет собой совокупность всех допустимых решений. Например, часто множество Z — это множество (возможно упорядоченных) выборок m элементов из n заданных. Так, в задаче о коммивояжере $m = n$ и множество Z является множеством циклических подстановок из n элементов $|Z| = (n-1)!$. Решение z^* , доставляющее минимум функции $\sigma(z)$ ($z \in Z$), называется оптимальным решением задачи.

Функцию $\sigma(z)$ ($z \in Z$) называют обычно *целевой функцией*.

Рассмотрим функцию $\gamma(A)$, определенную на всех подмножествах множества Z , за исключением пустого множества:

$$\gamma: 2^Z \setminus \{\emptyset\} \rightarrow \mathbb{R}$$

и удовлетворяющую условиям

$$а) \gamma(A) \leq \min_{z \in A} \sigma(z) \text{ для любого подмножества } A \subseteq Z;$$

$$б) \gamma(A_1) \geq \gamma(A_2), \text{ если } A_1 \subseteq A_2 \subseteq Z. \quad (3.2)$$

Определение 3.1. *Значение функции $\gamma(A)$ на множестве A называют нижней границей множества A .*

Выбор функции $\gamma(A)$ для решения конкретных задач в достаточной степени произволен. Имеет смысл выбирать

в качестве $\gamma(A)$ функцию, вычисление которой на данном множестве A требует значительно меньшего числа операций, чем нахождение минимума функции $\sigma(Z)$ на A .

Введение нижней границы множества является первым основным моментом метода ветвей и границ. Пусть дано некоторое разбиение множества Z на подмножества $Z_1^{(1)}, Z_2^{(1)}, \dots, Z_{n_1}^{(1)}$

$$Z = \bigcup_{i=1}^{n_1} Z_i^{(1)}, \quad Z_i^{(1)} \cap Z_j^{(1)} = \emptyset \quad (i \neq j), \quad (3.3)$$

тогда из соотношений 3.2 следует

$$\sigma(z^*) = \min_{z \in Z} \sigma(z) \geq \min_{1 \leq i \leq n_1} \gamma(Z_i^{(1)}) = \gamma(Z_{i_0}^{(1)}). \quad (3.4)$$

Отсюда можно оценить близость $\sigma(z) - \sigma(z^*)$ произвольного решения $z \in Z$ к оптимальному z^* следующим образом:

$$\sigma(z) - \sigma(z^*) \leq \sigma(z) - \gamma(Z_{i_0}^{(1)}) = \epsilon^1(z). \quad (3.5)$$

Если величина $\epsilon^1(z)$ лежит в пределах выбранной для конкретной задачи точности, то решение z можно принять за приближенное решение задачи. Если найдется такой элемент $z \in Z$, для которого

$$\sigma(z) = \gamma(Z_{i_0}^{(1)}), \text{ т. е. } \epsilon^1(z) = 0, \quad (3.6)$$

то из неравенства (3.5) следует, что z — оптимальное решение, т. е. $z = z^*$. Таким образом, в этом случае существует оптимальное решение z^* , принадлежащее $Z_{i_0}^{(1)}$. С этой точки зрения множество $Z_{i_0}^{(1)}$ является «наиболее перспективным» по сравнению с другими множествами $Z_i^{(1)}$ ($i \neq i_0$), входящими в разбиение множества Z . Напомним, что в заданном разбиении (3.3) множества Z , «наиболее перспективное» подмножество $Z_{i_0}^{(1)}$ имеет наименьшую границу

$$\gamma(Z_{i_0}^{(1)}) = \min_{1 \leq i \leq n_1} \gamma(Z_i^{(1)}). \quad (3.7)$$

Поскольку априори о точном решении z^* нельзя высказать каких-либо определенных суждений, то условие (3.7) принимают за решающее правило, по которому из множеств $Z_1^{(1)}, \dots, Z_{n_1}^{(1)}$ данного разбиения множества Z вы-

бирают множество $Z_{i_0}^{(1)}$ с целью сужения «области поиска оптимального решения».

Пусть теперь множество $Z_{i_0}^{(1)}$ представимо в виде

$$Z_{i_0}^{(1)} = \sum_{v=1}^{\nu(i_0)} Z_v^{(2)}, \quad Z_v^{(2)} \cap Z_{v'}^{(2)} = \emptyset \quad v \neq v'. \quad (3.8)$$

Тогда совокупность множеств

$$Z_1^{(1)}, Z_2^{(1)}, \dots, Z_{i_0-1}^{(1)}, \quad Z_{i_0+1}^{(1)}, \dots, Z_{n_1}^{(1)}, \\ Z_1^{(2)}, \dots, Z_{\nu}^{(2)}, \dots, Z_{\nu(i_0)}^{(2)}$$

образует новое разбиение множества Z , которое после переобозначений

$$Z_i^{(1)} \rightarrow Z_i^{(2)} \quad (1 \leq i \leq i_0 - 1), \quad Z_i^{(1)} \rightarrow Z_{i+\nu(i_0)-1}^{(2)} \quad (i_0 + 1 \leq \\ \leq i \leq n_1), \quad Z_i^{(2)} \rightarrow Z_{i_0+i-1}^{(2)} \quad (1 \leq i \leq \nu(i_0))$$

запишем в виде

$$Z = \sum_{i=1}^{n_2} Z_i^{(2)}. \quad (3.9)$$

Для последовательности множеств $\{Z_i^{(2)}\}$ ($i = 1, 2, \dots, n_2$) вычисляем их новые границы $\gamma(Z_i^{(2)})$ и на основании критерия (3.7) выбираем среди множеств, входящих в разбиение (3.9), множество $Z_{i_0}^{(2)}$ с наименьшей границей. Продолжая указанный процесс, на каждом k -м шаге получаем разбиение

множества Z вида $Z = \bigcup_{i=1}^{n_k} Z_i^{(k)}$, причем переход от разбиения множества Z с номером k к разбиению с номером $k+1$ осуществляется путем разбиения множества $Z_{i_0}^{(k)}$, удовлетворяющего условию

$$\gamma(Z_{i_0}^{(k)}) = \min_{1 \leq i \leq n_k} \gamma(Z_i^{(k)}). \quad (3.10)$$

Описанный процесс последовательного разбиения подмножеств решений $Z_i^{(k)}$ называется процессом «ветвления» исходного множества решений Z в соответствии с выбранным правилом ветвления (3.7). Процесс ветвления

является вторым основным моментом в методе ветвей и границ. Процесс ветвления исходного множества Z может быть схематически изображен в виде дерева вариантов, вершинами которого являются подмножества $Z_i^{(k)}$ (рис. 117), причем все вершины с фик-

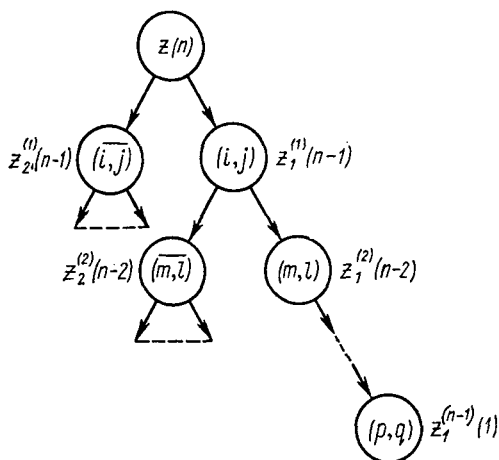


Рис. 117

сированным индексом k лежат на одном горизонтальном уровне.

Заметим, что в процессе ветвления происходит «измельчение» множеств, входящих в разбиение

$$\min_{1 \leq i \leq n_{k+1}} |Z_i^{(k+1)}| \leq \min_{1 \leq i \leq n_k} |Z_i^{(k)}|. \quad (3.11)$$

Отсюда при достаточно большом k ($k \rightarrow \infty$) можно получить разбиение множества Z на подмножества $Z_i^{(k)}$, состоящие из одного элемента. Это означает, что в процессе ветвления произведен полный перебор решений.

На каждом конечном шаге ветвления с использованием критерия (3.10) сужения области поиска оптимального решения может и не произойти. Действительно, хотя справедливо неравенство (3.11), для минимальных границ k -го и $(k+1)$ -го разбиения выполняется соотношение

$$\min_{1 \leq i \leq n_{k+1}} \gamma(Z_i^{(k+1)}) \geq \min_{1 \leq i \leq n_k} \gamma(Z_i^{(k)}).$$

Поэтому, вообще говоря, справедливо соотношение

$$|Z_{i_0}^{(k+1)}| \geq |Z_{i_0}^{(k)}|.$$

Рассмотрим случай, важный с точки зрения уменьшения общего объема вычислений, когда выполняется противоположное неравенство. Пусть при некотором k , $k < |Z|$, разбиение множества Z , $Z = \bigcup_{i=1}^{n_k} Z_i^{(k)}$ содержит подмножество, состоящее из одного элемента z . Пусть таким подмножеством будет $Z_1^{(k)}$ и пусть

$$\sigma(z) = \gamma(Z_1^{(k)}) > \min_{1 \leq i \leq n_k} \gamma(Z_i^{(k)}).$$

Рассмотрим все те подмножества $Z_i^{(k)}$ ($i = 1, 2, \dots, n_k$), для которых выполняется неравенство

$$\gamma(Z_i^{(k)}) > \gamma(Z_1^{(k)}).$$

Очевидно, что в процессе дальнейшего ветвления вершины, отвечающие таким подмножествам, не подлежат ветвлению, поскольку для всех таких $Z_i^{(k)}$ выполняется условие

$$\min_{z \in Z_i^{(k)}} \sigma(z) \geq \gamma(Z_i^{(k)}) > \gamma(Z_1^{(k)}).$$

Описанная ситуация будет реализована в рассматриваемом ниже варианте метода ветвей и границ при решении задачи о коммивояжере.

Как можно понять из вышесказанного, эффективность метода ветвей и границ существенным образом зависит от способа разбиения множеств Z , выбора правила ветвления и способа вычисления границы множеств. В следующем пункте рассмотрим все эти вопросы в связи с решением задачи о коммивояжере.

2. Решение задачи о коммивояжере методом ветвей и границ (точный алгоритм). Пусть $Z(n)$ — множество всех гамильтоновых контуров π , отвечающих матрице порядка, $n \times n$. Зафиксируем некоторую дугу, например (x_i, x_j) , и разобьем множество $Z(n)$ на два непересекающихся подмножества: подмножество $Z_1^{(1)}(n-1)$ всех контуров, содержащих дугу (x_i, x_j) , и подмножество $Z_2^{(1)}(n-1)$, сос-

стоящее из контуров, не содержащих дуги (x_i, x_j) . Указанный процесс ветвления (по дуге (x_i, x_j)) схематически представлен на рис. 117.

Уже на первом шаге ветвления вместо одной задачи отыскания кратчайшего контура на множестве $Z(n)$ получаем две задачи для подмножеств $Z_1^{(1)}(n-1)$ и $Z_2^{(1)}(n-1)$, каждая с меньшим числом сравниваемых контуров. Так, например, для множества $Z_1^{(1)}(n-1)$ соответствующая матрица «расстояний» A_1 будет иметь порядок $n-1$, поскольку эта матрица получается из исходной матрицы A зачеркиванием i -й строки и j -го столбца. В самом деле, в дальнейших построениях гамильтоновых контуров с помощью ветвления узла $Z_1^{(1)}(n-1)$ элементы i -й строки a_{ik} ($k = 1, 2, \dots, n$) и j -го столбца a_{kj} ($k = 1, 2, \dots, n$) не могут быть использованы, так как гамильтонов контур является элементарным по определению. По этим же соображениям исключается в дальнейшем дуга (x_j, x_i) , т. е. в матрице A_1 полагают элемент $a_{ji} = \infty$.

Этот процесс преобразования матрицы A в матрицу A_1 по дуге (x_i, x_j) будем называть в дальнейшем *процессом стягивания дуги* (x_i, x_j) .

Продолжая процесс ветвления узлов $Z_1^{(1)}(n-1)$ и $Z_2^{(1)}(n-1)$, в конце концов придем к подмножествам, состоящим лишь из одного гамильтонова контура. Например, продвигаясь по ветви $Z_1^{(1)}(n-1) \rightarrow Z_1^{(2)}(n-2) \rightarrow \dots \rightarrow Z_1^{(k)}(n-k) \rightarrow \dots$, через $(n-1)$ шагов получим множество $Z_1^{(n-1)}(1)$, состоящее из контура $\{(i, j), (m, l), \dots, (p, q)\}$. Как уже отмечалось в п. 1, одно из таких множеств является оптимальным решением задачи. Конечно, проведение полного ветвления множества $Z(n)$ равносильно по объему работы прямому перебору всех $(n-1)!$ возможных контуров множества $Z(n)$, но, используя понятие границы множества Z_i , можно попутно отбрасывать «бесперспективные» узлы, из которых не могут выходить ветви, ведущие к решению.

С учетом сказанного сформулируем следующее правило завершения работы алгоритма. Процесс работы алгоритма, состоящий в последовательном ветвлении исходного множества контуров $Z(n)$ с одновременным вычислением и сравнением нижних границ получающихся подмножеств, проис-

ходит до тех пор, пока не будут выполнены следующие два условия: а) получено подмножество, состоящее из одного гамильтонова контура,

б) нижняя граница этого подмножества, т. е. длина этого контура меньше (или равна) нижних границ всех ранее построенных подмножеств (узлов ветвления).

Полученный таким путем гамильтонов контур является искомым решением.

Рассмотрим теперь подробнее, каким образом выбирают дугу (x_i, x_j) , по которой ведется стягивание, и как определяется нижняя граница подмножеств контуров $Z_1^{(k)}$.

Вычисление нижней границы для подмножества $Z_1^{(k)}$

Это вычисление производится путем операции приведения матрицы A_k . Рассмотрим, например, определение нижней границы исходного множества Z .

Определение 3.2. *Операция приведения над матрицей A состоит в том, что 1) из каждого элемента каждой строки вычитают минимальный элемент этой же строки (приведение по строкам), 2) из каждого элемента каждого столбца после приведения по строкам вычитают минимальный элемент этого же столбца (приведение по столбцам).*

Полученную в результате матрицу будем обозначать через $\tilde{A}$ и называть *приведенной*. Приведенная матрица $\tilde{A}$ содержит в каждой строке и столбце, по крайней мере, по одному нулевому элементу $\tilde{a}_{lm} = 0$. Обозначим через D множество пар индексов, отвечающих нулевым элементам:

$$D = \{(l, m) \mid \tilde{a}_{lm} = 0\}. \quad (3.12)$$

Очевидно, что длина оптимального контура $\sigma_0(\tilde{A})$ связана с длиной оптимального контура $\sigma_0(A)$ в исходной матрице A соотношением

$$\sigma_0(A) = \sigma_0(\tilde{A}) + \left(\sum_{k=1}^n \alpha_k + \sum_{l=1}^n \beta_l \right), \quad (3.13)$$

где α_k ($k = 1, 2, \dots, n$) — минимальный элемент в k -й строке матрицы A , в β_l ($l = 1, 2, \dots, n$) — минимальный элемент в l -м столбце матрицы A после приведения по строкам.

Величину $\sum_{k=1}^n \alpha_k + \sum_{l=1}^n \beta_l = h$ называют *константой*

приведения.

В приведенной матрице $\tilde{A}$, $\tilde{a}_{ij} \geq 0$, поэтому для любого контура $\pi \in Z(n)$ его длина $\sigma(\pi) \geq 0$ и $\sigma_0(\tilde{A}) \geq 0$. Из формулы (3.13) получаем следующую нижнюю границу длины контура для множества $Z(n)$:

$$\sigma_0(A) = \min_{\pi \in Z(n)} \sigma(\pi) \geq \sum_{k=1}^n \alpha_k + \sum_{l=1}^n \beta_l = h,$$

отсюда

$$\gamma(Z(n)) = h = \sum_{k=1}^n \alpha_k + \sum_{l=1}^n \beta_l. \quad (3.14)$$

Заметим, что из формулы (3.13) и определения множества D следует, что если множество D содержит хотя бы одну совокупность элементов, из которых можно образовать циклическую подстановку множества $X = \{1, 2, \dots, n\}$, то нижняя граница множества $Z(n)$ является *точно* нижней гранью $\sigma(\pi)$ на $Z(n)$. Действительно, в этом случае

$\sigma_0(\tilde{A}) = 0$, а величина $h = \sum_{k=1}^n \alpha_k + \sum_{l=1}^n \beta_l$ является длиной некоторого контура.

Аналогичную процедуру приведения и вычисления нижней границы можно провести для подмножеств $Z_1^{(k)}$ с матрицами меньшего порядка, чем n . Так, например, для множества $Z_1^{(1)}$ имеем $\gamma(Z_1^{(1)}) = a_{ij} + h_1$, где h_1 — константа приведения матрицы A_1 , полученной стягиванием матрицы A по дуге (i, j) .

В ы ч и с л е н и е н и ж н е й г р а н и ц ы д л я
п о д м н о ж е с т в $Z_2^{(k)}$

Вычисление производится по известной нижней границе «предыдущего» подмножества (так, например, для $Z_2^{(1)}$ таковым будем исходное множество $Z(n)$ (см. рис. 117) путем добавления к его границе числа α_{ii} , выражающего собой

минимальное увеличение длины любого контура, не содержащего по определению множества $Z_2^{(1)}$ дуги (i, j) .

Действительно, длина гамильтонова контура, не содержащего дуги (i, j) , а также (j, i) , возрастает по сравнению с длиной контуров, проходящих через дугу (i, j) (предполагается, что поскольку дуга (i, j) — дуга ветвления, то $\tilde{a}_{ij} = 0$ (см. ниже), не менее чем на величину α'_{ij} , где

$$\alpha'_{ij} = \min_{k \neq i} \tilde{a}_{ik} + \min_{k \neq j} \tilde{a}_{kj}. \quad (3.15)$$

Иллюстрацией к этому рассуждению может служить рис. 118.

Таким образом, для нижней границы множества $Z_2^{(1)}$ получаем формулу

$$\gamma(Z_2^{(1)}) = h + \alpha'_{ij}, \quad (3.16)$$

где α'_{ij} определено соотношением (3.15).

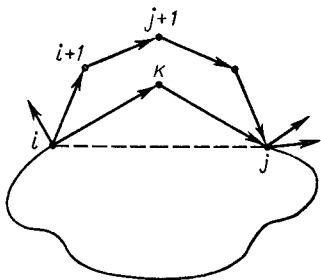


Рис. 118

Выбор дуги (i, j) , по которой производится ветвление

Выбор дуги осуществляется так, чтобы величина α'_{lm} , выражающая увеличение длины контура при невключении в него дуги (l, m) , была бы максимальной по всем дугам (l, m) , для которых в приведенной матрице $\tilde{A}$, $\tilde{a}_{lm} = 0$, т. е.

$$\alpha'_{ij} = \max_{(l, m) \in D} \alpha'_{lm}, \quad \text{где} \quad \alpha'_{lm} = \min_{k \neq m} \tilde{a}_{lk} + \min_{k \neq l} \tilde{a}_{km}, \quad (3.17)$$

а множество D определено соотношением (3.12).

При работе алгоритма может возникнуть необходимость произвести ветвление какого-либо множества $Z_2^{(k)}$. Рассмотрим, например, множество $Z_2^{(2)}$ ($n-2$) на рис. 117. В этом случае исходной матрицей для $Z_2^{(2)}$ ($n-2$) будет матрица A , в которой необходимо 1) провести стягивание по дуге (i, j) , так как элементы множества $Z_2^{(2)}$ ($n-2$) обязательно содержат дугу (i, j) , 2) заменить элемент a_{mi} на ∞ ,

так как $a_{ml} \notin Z_2^{(2)} (n - 2)$. В результате получим матрицу $\bar{A}_2$, после приведения которой по формуле (3.17) определим дугу для стягивания в множестве $Z_2^{(2)} (n - 2)$.

Задача 2. Показать, что граница множества $Z_2^{(k)} (n - k)$ может быть определена формулой

$$\gamma(Z_2^{(k)}) = \sum_{(i, j) \in M} a_{ij} + h', \quad (3.18)$$

где h' — константа приведения матрицы $\bar{A}_k$, а множество M — множество пар дуг, обязательно входящих в контуры из $Z_2^{(k)} (n - k)$.

Формулу (3.18) можно использовать для проверки правильности расчетов по формулам (3.16).

Полная схема работы алгоритма ветвей и границ будет выявлена при расчете следующего конкретного примера.

Пусть матрица «расстояний» — матрица A представлена табл. 5.

ТАБЛИЦА 5

$i \backslash j$	1	2	3	4	5	« α_k »
1	∞	30	40	15	6	6
2	10	∞	18	7	9	7
3	20	30	∞	0	10	0
4	25	10	35	∞	5	5
5	9	8	7	6	∞	6

1°. Приведение матрицы A и получение границы множества $Z(n)$ производится по формуле (3.14). В табл. 5 приведены в столбце « α_k » минимальные элементы строк; после приведения по строкам получим матрицу A' (табл. 6).

После приведения матрицы A' по столбцам матрица $\tilde{A}$ имеет вид, представленный табл. 7. Константа приведения

$$h = \sum_{k=1}^5 \alpha_k + \sum_{l=1}^5 \beta_l = 24 + 6 = 30. \text{ По формуле (3.14) гра-}$$

ница множества $Z(n) : \gamma(Z(n)) = 30$.

2°. Выбор дуги ветвления и оценка нижней границы мно-

$i \backslash j$	1	2	3	4	5
1	∞	24	34	9	0
2	3	∞	11	0	2
3	20	30	∞	0	10
4	20	5	30	∞	0
5	3	2	1	0	∞
« β_I »	3	2	1	0	0

жества $Z_2^{(1)}$. Из вида матрицы $\tilde{A}$ (табл. 7) определяем множество D ,

$$D = \{(1,5), (2,1), (2,4), (3,4), (4,5), (5,1), (5,2), (5,3), (5,4)\}$$

и вычисляем по формуле (3.15) α'_{lm} , $(l, m) \in D$,

$$\alpha'_{lm} = \{9, 0, 0, 10, 3, 0, 3, 10, 0\};$$

отсюда

$$\max_{(l, m) \in D} \alpha'_{lm} = 10 = \alpha'_{34}, \text{ т. е. дуга } (3,4) \text{ — дуга ветвления.}$$

Нижняя граница множества $Z_2^{(1)}$, элементы которого не содержат дугу (3,4), из формулы (3.16) равна

$$\gamma(Z_2^{(1)}) = h + \alpha'_{ij} = 30 + 10 = 40.$$

Ветвление по дуге $(i, j)' = (3, 4)$ изображено на рис. 119.

3°. Стягивание матрицы A по дуге $(i, j) = (3, 4)$ и приведение получившейся матрицы по методу первого шага алгоритма. В результате стягивания матрицы A получим

матрицу A_1 , отвечающую множеству $Z_1^{(1)}$ ($n-1$) и проделаем процедуры первого и второго шагов алгоритма до тех пор, пока множество Z_i , $i > 1$ не будет содержать только один контур (рис. 120).

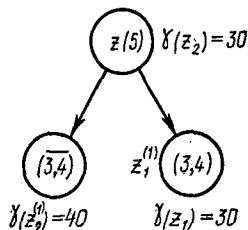


Рис. 119

Стягивая матрицу A по дуге $(3,4)$ после приведения получаем матрицу расстояний $\tilde{A}_1$ (табл. 8). Здесь элемент a_{43} заменен на ∞ в соответствии с замечанием, сделанным выше.

Поскольку для матрицы $\tilde{A}_1$ константа приведения $h_1 = 0$, то матрица $\tilde{A}_1$ совпадает со своей приведенной $\tilde{\tilde{A}}_1$.

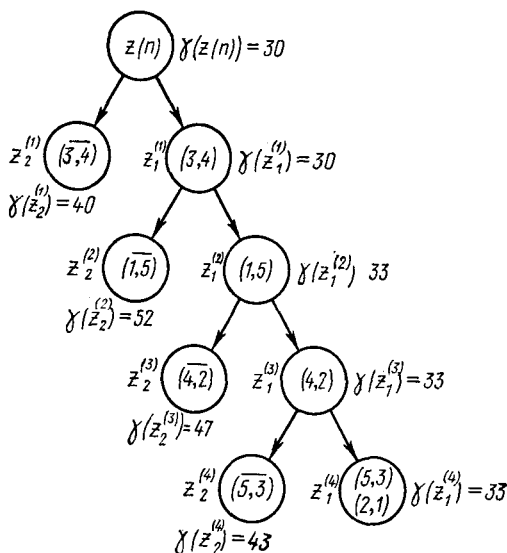


Рис. 120

и соответственно $\gamma(Z_1^{(1)}) = 30 = \gamma(Z(n))$. Выбираем в качестве дуги ветвления дугу $(1,5)$, получаем множества $Z_2^{(2)}$, $Z_1^{(2)}$ и повторяем процедуру до тех пор, пока матрица расстояний будет иметь порядок 2×2 (см. рис. 120).

Множество $Z_1^{(4)}$ состоит из единственного цикла, содержащего дуги $\{(3,4), (1,5), (4,2), (5,3), (2,1)\}$. Длина этого цикла $\sigma(\pi_1) = 33$.

4°. Ветвление множества $Z_2^{(k)}$, если $\gamma(Z_2^{(k)}) < \sigma(\pi_1)$.

В данном случае нижние границы множеств $Z_2^{(1)}$, $Z_2^{(2)}$, $Z_2^{(3)}$, $Z_2^{(4)}$ больше, чем длина получившегося контура π_1 . поэтому ветвление множеств $Z_2^{(k)}$ дает контуры, длина ко-

орых будет заведомо превосходить длину контура π_1 .
 Таким образом, контур π_1 — оптимальный.

В общей ситуации необходимо было бы произвести ветвление тех вершин $Z_2^{(k)}$, для которых $\gamma(Z_2^{(k)}) < \sigma(\pi_1)$, поучить контуры $\pi_2, \dots, \pi_m$, сравнить длины этих контуров между собой и получить тем самым оптимальный контур.

З а м е ч а н и е. Как видно из описания алгоритма, его работа будет особенно эффективна при реализации алгоритма на многопроцессорных ЭВМ, так как в этом случае возможно проводить одновременное ветвление нескольких вершин дерева вариантов.

ТАБЛИЦА 7

$i \backslash j$	1	2	3	4	5
1	∞	22	33	9	0
2	0	∞	10	0	2
3	17	28	∞	0	10
4	17	3	29	∞	0
5	0	0	0	0	∞

ТАБЛИЦА 8

$i \backslash j$	1	2	3	5
1	∞	22	33	0
2	0	∞	10	2
4	17	3	∞	0
5	0	0	0	∞

ТАБЛИЦА 9

$i \backslash j$	1	2	3	4	5
1	∞	30	46	19	2
2	10	∞	19	3	2
3	21	15	∞	35	0
4	26	22	30	∞	4
5	3	4	3	2	∞

Задача 3. Решить задачу о коммивояжере для матрицы «р стояний» A , приведенной в табл. 9.

Особенность рассмотренного алгоритма, основанного методе ветвей и границ, состоит в том, что оптимальн контур определяется последовательностью дуг, выбор торых на каждом шаге определяется условием: на дуге, которой ведется стягивание, величина α'_{lm} , $(l, m) \in D$, стигала бы максимума (см. (3.17)).

Изменяя принцип выбора дуги для стягивания, пол чаем другую реализацию метода ветвей и границ. Опиш кратко алгоритм, в котором выбор дуги для стягиван производится на основе минимального значения конту функции (см. [3]).

Определение 3.3. Пусть $Z(\alpha, \beta)$ — множество гамильн новых контуров, полного графа, проходящих через дугу (α, β) F^n

Функция $(\alpha, \beta) \rightarrow \mathbb{R}^+$, определяемая формулой

$$F_1^n[(\alpha, \beta)] = \sum_{\pi \in Z(\alpha, \beta)} \sigma(\pi), \quad (3.1)$$

называется контур-функцией первого порядка.

Правило выбора начальной дуги (i, j) контура π_0 опр деляется из условия минимальности контура-функции:

$$(i, j) \in \pi_0 \Rightarrow F_1^n[(i, j)] = \min_{(\alpha, \beta) \in U_A} F_1^n[(\alpha, \beta)], \quad (3.2)$$

где U_A — множество дуг графа G .

Последующая дуга контура определяется по матрице , получаемой из исходной стягиванием по дуге (i, j) , в сос ветствии с формулой (3.20).

Значение контур-функции для пары дуг (i, j) , (k, m) обозначим через $F_2^n[(i, j), (k, m)]$, т. е.

$$F_2^n[(i, j), (k, m)] = \sum_{\pi \in Z_1((i, j), (k, m))} \sigma(\pi). \quad (3.2)$$

Здесь $Z_1((i, j), (k, m))$ — множество гамильтоновых кои туров, содержащих дуги $((i, j), (k, m))$ в исходном граф

Аналогично, формула (3.21) определяет контур-функци k -го порядка $F_k^n[(i, j), (k, m), \dots, (p, r)]$.

Преимущество указанного способа выбора дуги для стягивания (см. (3.20)) состоит в том, что для вычисления значения контур-функции k -го порядка можно получить систему рекуррентных соотношений, по которым легко провести вычисления. При этом над матрицей A , помимо процесса стягивания, т. е. зачеркивания определенных строк и столбцов, никакие другие операции не производятся.

Приведем без доказательства формулу для контур-функции первого порядка

$$F_1^n[(k, l)] = (n-3)! [\theta^n - c_l^n - b_k^n + (n-1)a_{kl}^n - a_{lk}^n], \quad (3.22)$$

где

$$\begin{aligned} c_l^n &= \sum_{\substack{i=1 \\ i \neq l}}^n a_{il}^n, \quad b_k^n = \sum_{\substack{i=1 \\ i \neq k}}^n a_{ki}^n, \quad \theta^n = \sum_{i=1}^n c_i^n = \sum_{i=1}^n \sum_{j=1}^n a_{ij}^n = \\ &= \sum_{i \neq j=1}^n a_{ij}^n, \quad a_{ij}^n \text{—элементы исходной матрицы расстояний } A. \end{aligned}$$

Проводя конкретные расчеты по алгоритму «ветвей и границ», нетрудно заметить, что объем вычислительной работы существенно зависит от выбора вершин ветвления. Для матриц произвольного вида, вообще говоря, этот объем работы не намного меньше, чем при непосредственном переборе $(n-1)!$ возможных вариантов (n — размерность задачи). В практических расчетах весьма важным является применение приближенных методов решения задачи с небольшим объемом вычислительной работы.

Описанный выше в п. 1 алгоритм решения задачи о коммивояжере является точным в том смысле, что правило завершения работы алгоритма гласит: вычислительный процесс заканчивается тогда и только тогда, когда полученный контур оптимален.

Определение 3.4. Алгоритм решения задачи о коммивояжере называется приближенным, если правило завершения его работы таково, что вычислительный процесс заканчивается тогда (но не только тогда), когда найденный контур оптимален.

Например, можно совсем не проводить ветвления в методе «ветвей и границ» и заканчивать работу алгоритма как только в результате процесса последовательного стяги-

вания матрицы A одним из рассмотренных выше способов* получится гамильтонов контур, который принимается за оптимальный.

Отметим, наконец, возможность использования для решения задачи о коммивояжере приближенного метода статистических испытаний (метод Монте—Карло). Точ-

ность этого метода невелика, порядка $\frac{1}{\sqrt{n}}$, где n —размер-

ность матрицы A . Достоинством метода является его простота реализации на ЭВМ и использование небольшого объема памяти машины $\sim n^2 + 2n$.

* Приближенный алгоритм решения задачи методом «ветвей и границ» был реализован на ЭВМ М-220, и расчеты, проведенные Н. Г. Дудниковым (в дипломной работе) дали удовлетворительные результаты по точности и скорости счета. При этом выяснилось преимущество выбора дуги для стягивания с использованием контур-функции по сравнению со способом, основанным на формуле (3.17).

**ЭЛЕМЕНТЫ ТЕОРИИ КОНЕЧНЫХ АВТОМАТОВ.
МАТРИЦЫ СМЕЖНОСТИ ГРАФОВ
И ОПЕРАЦИИ НАД НИМИ**

Тесная связь теории конечных автоматов с теорией графов и удобств матричного описания графов в вопросах, связанных с конечными автоматами, обусловила объединение этих вопросов в одну главу.

Теория конечных автоматов применяется в настоящее время, например, при конструировании электронно-вычислительных машин. В связи с этим возникает так называемая задача декомпозиции автоматов, т. е. построения сложных автоматов из более простых блоков, при помощи различных способов их соединения. В данной главе рассматривается весьма частный случай этой задачи: вопрос о разложении данного автомата на два, соединенных параллельно.

В этой главе на частных примерах будет показано, каким образом методы теории графов могут быть применены к решению задач декомпозиции.

Кроме чисто практического применения, теория автоматов имеет и теоретическое значение. Понятие конечного автомата есть попытка строгого математического определения понятия алгоритма. Известны и другие попытки определить понятие алгоритма, а именно: машины Тьюринга и теория рекурсивных функций. При этом вопрос о полноте таким образом определенного понятия алгоритма выходит уже за рамки математических построений.

Отметим, что для простоты изложения в этой главе мы ограничились случаем простых графов. Большинство определений и теорем этой главы можно перенести и на случай мультиграфов.

§ 1. НЕКОТОРЫЕ ПОНЯТИЯ ТЕОРИИ КОНЕЧНЫХ АВТОМАТОВ

1. Конечные автоматы и нагруженные мультиграфы. Определим прежде всего понятие конечного автомата.

Определение 1. Конечным автоматом называется совокупность пяти объектов $Z = (A, B, Q, \varphi, \psi)$, где

1) A — конечное множество, называемое в дальнейшем входным алфавитом автомата;

2) B — конечное множество, называемое выходным алфавитом автомата,

3) Q — конечное множество (множество состояний автомата) с отмеченным элементом q_0 — начальным состоянием автомата; Q иногда будем называть алфавитом состояний автомата,

4) $\varphi : Q \times A \rightarrow Q$ — функция перехода,

5) $\psi : Q \times A \rightarrow B$ — функция выхода.

Опишем теперь работу конечного автомата. Будем считать, что автомат работает в дискретном времени ($t = 0, 1, 2, 3, \dots$). В начальный момент времени $t = 0$ автомат находится в отмеченном состоянии q_0 . В момент времени $t = 1$ на вход автомата подается некоторая буква a входного алфавита*. В результате этого автомат переходит в состояние $q = \varphi(q_0, a)$ и на его выходе появляется буква $b = \psi(q_0, a)$. В момент $t = 2$ на вход автомата может быть подана другая буква входного алфавита a' . Тогда в момент $t = 2$ автомат переходит в состояние $q' = \varphi(q, a')$, а на его выходе появляется буква $b' = \psi(q, a')$. Работа автомата продолжается таким образом, пока на его вход подаются буквы алфавита A .

Обозначим через $a(t)$ букву алфавита A , подаваемую на вход автомата в момент t , через $q(t)$ — состояние автомата и через $b(t)$ — букву алфавита B , появляющуюся на выходе в этот момент. Функция $a(t)$ определена для $t = 1, 2, \dots, t_{\max}$, при $t = t_{\max} + 1$ она не определена. Это означает, что в момент $t_{\max} + 1$ автомат прекращает свою работу. При этом $q(t)$ и $b(t)$ определены для тех же t . Очевидно, всегда $q(0) = q_0$.

Условимся называть словом над некоторым алфавитом всякую конечную последовательность его букв. Рассмотрим некоторое (произвольное!) слово $a_1 \dots a_n$ входного алфавита A . Будем последовательно подавать буквы этого слова на вход Z , т. е. положим

$$a(t) = a_t \quad (t = 1, 2, \dots, n).$$

В процессе работы автомата на выходе его появятся после-

* Буквой алфавита мы называем произвольный его элемент.

довательно буквы $b(1), b(2), \dots, b(n)$, а сам автомат пройдет через состояния $q(1), \dots, q(n)$. Обозначив $b_t = b(t)$, будем говорить, что на выходе автомата появилось слово $b_1 \dots b_n$. Сформулируем полученное утверждение в виде теоремы.

Теорема 1.1. *Любой конечный автомат Z задает функцию, определенную на множестве слов над алфавитом A со значениями в множестве слов над алфавитом B .*

Не следует думать, что любая функция, отображающая слова алфавита A в слова алфавита B , отвечает некоторому конечному автомату. Нетрудно видеть, что множество конечных автоматов с фиксированными A и B счетно (если не различать автоматы, отличающиеся друг от друга только обозначениями состояний), а множество функций описанного в теореме вида несчетно (докажите это!).

Ближайшая наша цель заключается в том, чтобы показать, что каждый автомат может быть эквивалентным образом описан с помощью некоторого **нагруженного мультиграфа**.

Определение 1.2. *Нагруженным мультиграфом с множеством индексов Y называется пара (G, f) , где $G = (X, \Gamma)$ — мультиграф, а $f: \Gamma \rightarrow Y$ — отображение, определенное на всем множестве Γ .*

Каждому конечному автомату $Z = (A, B, Q, \varphi, \psi)$ поставим в соответствие нагруженный мультиграф $G = (Q, \Gamma)$ с множеством индексов $A \times B$ и отображением f , которые строятся ниже.

Пусть $q_i, q_j \in Q$ — произвольная пара состояний автомата. Определим подмножество $A[q_i, q_j]$ множества A формулой

$$A[q_i, q_j] = \{a \in A \mid \varphi(q_i, a) = q_j\}.$$

Определим число $n(q_i, q_j)$ равенством $n(q_i, q_j) = |A[q_i, q_j]|$. Перенумеруем элементы множества $A[q_i, q_j]$ каким-либо образом. Будем обозначать через $a_k[q_i, q_j]$ — k -й элемент $A[q_i, q_j]$ ($k=1, 2, \dots, n(q_i, q_j)$) (если $n(q_i, q_j)=0$, то $A[q_i, q_j]$ пусто).

Теперь определим множество Γ формулой

$$\Gamma = \{(q_i, q_j, k) \in Q \times Q \times \mathbb{Z}_+ \mid k \leq n(q_i, q_j)\},$$

а функцию f — формулой

$$f((q_i, q_j, k)) = (a_k[q_i, q_j], \psi(q_i, a_k[q_i, q_j])).$$

Пусть (G, f) , где $G = (Q, \Gamma)$ — нагруженный мультиграф с множеством индексов $A \times B$, обладающий свойством:

А. Для любого элемента $a \in A$ из каждой вершины $q \in Q$ выходит одна и только одна дуга $f \in \Gamma$ такая, что

$$\text{Пр}_1 f(l) = a.$$

(Здесь $\text{Пр}_1 : A \times B \rightarrow A$ — проекция прямого произведения на первую компоненту.)

Определим теперь недостающие компоненты φ и ψ конечного автомата следующим образом. Пусть $l(q, a)$ — дуга G , выходящая из q , и такая, что

$$\text{Пр}_1 f(l(q, a)) = a.$$

В силу свойства A такая дуга определена однозначно для любых $a \in A$, $q \in Q$, тогда

$$\varphi(q, a) = \text{Пр}_2 l(q, a)$$

$\text{Пр}_2 : Q \times Q \times \mathbb{Z}_+ \rightarrow Q$ — проекция на вторую компоненту; напомним, что $l(q, a) \in \Gamma \subset Q \times Q \times \mathbb{Z}_+$, а

$$\psi(q, a) = \overline{\text{Пр}_2} f(l(q, a)),$$

где $\text{Пр}_2 : A \times B \rightarrow B$ — проекция на вторую компоненту.

Теорема 1.2. Построенные выше отображения устанавливают взаимно-однозначное соответствие между множеством конечных автоматов с заданными A , B , Q и множеством нагруженных мультиграфов описанного выше типа и обладающих свойством A , если считать неразличимыми все мультиграфы, которые отличаются друг от друга лишь перенумерацией дуг.

Доказательство вполне очевидно.

Пример 1. Пусть для автомата $A = \{a_1, a_2, a_3\}$, $Q = \{q_1, q_2, q_3, q_4\}$, $B = \{b_1, b_2\}$, а функции φ и ψ заданы табл. 10 и 11.

ТАБЛИЦА 10

$\begin{array}{c} \backslash \\ Q \\ A \end{array}$	q_1	q_2	q_3	q_4
a_1	q_3	q_1	q_4	q_3
a_2	q_4	q_3	q_1	q_2
a_3	q_2	q_2	q_2	q_1

ТАБЛИЦА II

$\begin{matrix} Q \\ A \end{matrix}$	q_1	q_2	q_3	q_4
a_1	b_1	b_2	b_1	b_1
a_2	b_1	b_1	b_2	b_1
a_3	b_2	b_2	b_1	b_2

Нагруженный мультиграф, отвечающий этому конечному автомату, изображен на рис. 121.

Если автомат Z обладает тем свойством, что из равенства

$$\varphi(q_j, a_{k_1}) = \varphi(q_j, a_{k_2})$$

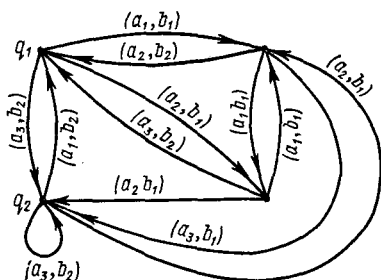


Рис. 121

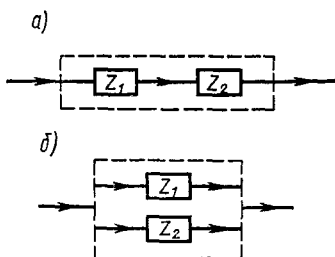


Рис. 122

следует $a_{k_1} = a_{k_2}$, то соответствующий этому автомату нагруженный мультиграф будет простым. Для облегчения изложения будем рассматривать только такие автоматы. Можно доказать, что каждый автомат, не обладающий этим свойством, может быть заменен на обладающий им без изменения функции, описанной в теореме 1.1.

2. Операции над конечными автоматами. Определим теперь операции последовательного и параллельного соединения двух конечных автоматов (рис. 122, а и б).

Определение 1.3. Автомат Z , отвечающий последовательной работе автоматов $Z_1 = (A_1, B_1, Q_1, \varphi_1, \psi_1)$ и $Z_2 = (A_2, B_2, Q_2, \varphi_2, \psi_2)$, считается определенным в том и только том случае, когда $B_1 = A_2$, и при этом

$Z = (A, B, Q, \varphi, \psi)$,
где $A = A_1$, $B = B_2$, $Q = Q_1 \times Q_2$, а φ и ψ определяются формулами

$$\begin{aligned}\varphi\left(\left(q_i^{(1)}, q_j^{(2)}\right), a_k^{(1)}\right) &= \left(\varphi_1\left(q_i^{(1)}, a_k^{(1)}\right), \right. \\ &\quad \left. \varphi_2\left(q_j^{(2)}, \psi_1\left(q_i^{(1)}, a_k^{(1)}\right)\right)\right), \\ \psi\left(\left(q_i^{(1)}, q_j^{(2)}\right), a_k^{(1)}\right) &= \psi_2\left(q_j^{(2)}, \psi_1\left(q_i^{(1)}, a_k^{(1)}\right)\right),\end{aligned}$$

т. е. автоматы Z_1 и Z_2 соединены таким образом, что выходной сигнал автомата Z_1 автоматически передается на вход автомата Z_2 .

Определение 1.4. Автомат Z , отвечающий параллельной* работе (см. рис. 122, б) автоматов Z_1 и Z_2 , определен всегда, и при этом

$Z = (A, B, Q, \varphi, \psi)$,
где $A = A_1 \times A_2$, $B = B_1 \times B_2$, $Q = Q_1 \times Q_2$, а φ и ψ определяются следующим образом:

$$\begin{aligned}\varphi\left(\left(q_i^{(1)}, q_j^{(2)}\right), \left(a_k^{(1)}, a_l^{(2)}\right)\right) &= \left(\varphi_1\left(q_i^{(1)}, a_k^{(1)}\right), \right. \\ &\quad \left. \varphi_2\left(q_j^{(2)}, a_l^{(2)}\right)\right), \\ \psi\left(\left(q_i^{(1)}, q_j^{(2)}\right), \left(a_k^{(1)}, a_l^{(2)}\right)\right) &= \left(\psi_1\left(q_i^{(1)}, a_k^{(1)}\right), \right. \\ &\quad \left. \psi_2\left(q_j^{(2)}, a_l^{(2)}\right)\right).\end{aligned}$$

Это определение, очевидно, отвечает способу соединения автоматов Z_1 и Z_2 , при котором на входы этих автоматов одновременно подаются буквы $a_k^{(1)}$ и $a_l^{(2)}$ соответствующих алфавитов, а выходной буквой считается пара букв $(b_k^{(1)}, b_l^{(2)})$, появившихся на выходах Z_1 и Z_2 в данный момент.

✓ Двумя описанными способами можно получать довольно большие структурные схемы автоматов (примеры даны на

* В литературе иногда рассматриваются два способа параллельного соединения автоматов, отвечающих параллельной одновременной и параллельной неодновременной работе автоматов. Наше определение описывает первый из этих способов.

рис. 123) Можно (мы здесь этого делать не будем) определить и другие способы построения сложных автоматов из более простых.

Предлагаем читателю самостоятельно определить условия, при которых определены автоматы со структурными схемами (см. рис. 123, а и б) и выписать выражения элемен-

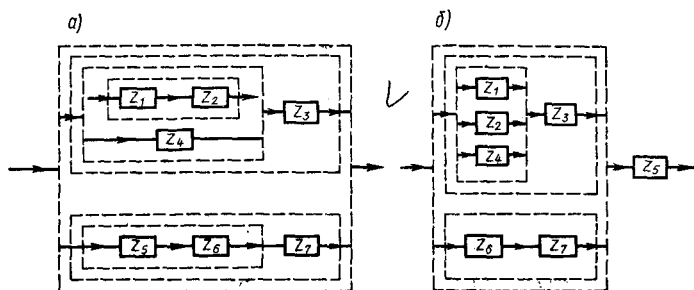


Рис. 123

тов результирующих автоматов через элементы $Z_1, \dots, Z_7$.

В связи с возможностью построения более сложных автоматов из более простых возникает обратная задача, так называемая задача декомпозиции автомата Z . Иными словами, нужно выяснить, при помощи соединения каких более простых автоматов можно построить данный автомат Z и вычертить структурную схему соединения таких автоматов. В общем случае эта задача имеет неоднозначное решение и поэтому ищется декомпозиция, оптимальная в том или ином смысле. Мы не будем здесь рассматривать эту задачу в полном объеме.

Рассмотрим еще понятие изоморфизма двух конечных автоматов Z_1 и Z_2 .

✓ **Определение 1.5.** Два автомата $Z = (A_1, B_1, Q_1, \Phi_1, \Psi_1)$ и $Z_2 = (A_2, B_2, Q_2, \Phi_2, \Psi_2)$ называются изоморфными, если существуют такие взаимно-однозначные отображения алфавитов $f: A_1 \rightarrow A_2$ и $g: B_1 \rightarrow B_2$, что слово $a_1^{(1)} \dots a_n^{(1)}$ перерабатывается автоматом Z_1 в слово $b_1^{(1)} \dots b_n^{(1)}$ тогда и только тогда, когда слово $f(a_1^{(1)}) \dots f(a_n^{(1)})$ перерабатывается автоматом Z_2 в слово $g(b_1^{(1)}) \dots g(b_n^{(1)})$, т. е. отображения слов этих двух автоматов отличаются лишь переобозначением алфавитов.

Очевидно, что *достаточным* условием изоморфизма автоматов будет изоморфизм их графоидов (G_1, f_1) и (G_2, f_2) , иначе говоря, такой изоморфизм графов G_1 и G_2 , который каждое ребро, отмеченное парой $(a^{(1)}, b^{(1)})$, переводит в ребро, отмеченное парой $(f(a^{(1)}), g(b^{(1)}))$. Устанавливать изоморфизм конечных автоматов можно с помощью алгоритма распознавания изоморфизма графов, который будет приведен в следующем параграфе.

§ 2. МАТРИЦЫ СМЕЖНОСТИ ГРАФОВ И ОПЕРАЦИИ НАД НИМИ

1. Матрицы смежности графа. Основная теорема об изоморфизме. Матрицы смежности являются одним из способов задания графов. При этом, как будет видно из дальнейшего, мультиграф $G = (X, \Gamma)$ определен своей матрицей смежности с точностью до изоморфизма.

Определение 2.1. Матрицей смежности мультиграфа $G = (X, \Gamma)$ называется матрица размера $n \times n$, где $n = |X|$, в которой элемент a_{ij} равен числу элементов Γ , имеющих вид (x_i, x_j, t) .

В графе без кратных дуг элементы a_{ij} могут быть равны единице или нулю в зависимости от того, входит ли дуга (x_i, x_j) в Γ или нет. Очевидно, что мультиграф определен с точностью до нумерации ребер.

Выясним теперь, каким образом по матрицам двух графов без кратных дуг можно определить, изоморфны ли эти графы.

Теорема 2.1. Графы $G = (X, \Gamma)$ и $G' = (X', \Gamma')$ с матрицами смежности $\|a_{ij}\|$ и $\|a'_{ij}\|$ соответственно изоморфны тогда и только тогда, когда

- 1) выполнено равенство $|X| = |X'| (= n)$ и
- 2) существует такая подстановка $\sigma \in S_n$, что $a_{ij} = a'_{\sigma(i)\sigma(j)}$, где S_n — группа подстановок множества $\{1, \dots, n\}$.

Доказательство. Пусть Φ — изоморфизм G и G' . Этот изоморфизм однозначно характеризуется взаимно-однозначным отображением $\varphi: X \rightarrow X'$. Тогда очевидно, что условие 1) выполнено. Кроме того, φ индуцирует подстановку S_n следующим образом: если $\varphi(x_i) = x'_j$, то $\sigma(i) = j$. Так как φ — взаимно-однозначное отображение, то σ — подстановка на множестве $\{1, \dots, n\}$. Далее, если $a_{ij} = 1$, то $(x_i, x_j) \in \Gamma$, $(\varphi(x_i), \varphi(x_j)) \in \Gamma'$ и $a'_{\sigma(i)\sigma(j)} = 1$.

Если $a_{ij} = 0$, то $(x_i, x_j) \notin \Gamma$, $(\varphi(x_i), \varphi(x_j)) \notin \Gamma'$ и $a'_{\sigma(i)\sigma(j)} = 0$.

Обратно, если $|X| = |X'|$, и существует подстановка $\sigma \in S_n$ такая, что $a_{ij} = a_{\sigma(i)\sigma(j)}$, то отображение φ , определенное равенством $\varphi(x_i) = x'_{\sigma(i)}$, индуцирует изоморфизм графов. Докажем это. Во-первых, очевидно, что φ взаимно-однозначно. Во-вторых, если $(x_i, x_j) \in \Gamma$, то $a_{ij} = 1$, $a'_{\sigma(i)\sigma(j)} = a_{ij} = 1$ и $(x'_{\sigma(i)}, x'_{\sigma(j)}) = (\varphi(x_i), \varphi(x_j)) \in \Gamma'$. Обратно, если $(x_i, x_j) \notin \Gamma$, то $a_{ij} = 0$, $a'_{\sigma(i)\sigma(j)} = a_{ij} = 0$ и $(x'_{\sigma(i)}, x'_{\sigma(j)}) = (\varphi(x_i), \varphi(x_j)) \notin \Gamma'$. ■

Доказательство теоремы 1 с незначительными усложнениями переносится на случай мультиграфов.

2. Модификация основной теоремы. Алгоритм распознавания изоморфизма графов. Заметим, что использовать непосредственно теорему 1 при практическом распознавании изоморфизма для графов с большим $n = |X|$ нецелесообразно, так как количество подстановок, подлежащих проверке, равно $n!$ и весьма быстро растет с ростом n . Используем метод, который для большинства графов позволит свести перебор к минимуму. Этот метод основан на построении графа соответствия для данных графов G и G' .

Очевидно, что если графы G и G' изоморфны, то у соответствующих вершин одинаковы полустепени исхода и захода. В связи с этим строим граф соответствия как граф, множество вершин которого есть $X \vee X'$, и вершина из X соединена с вершиной из X' в том и только в том случае, если у этих вершин в графах G и G' соответственно одинаковые полустепени исхода и захода. Вершины из X между собой не соединены, вершины из X' тоже.

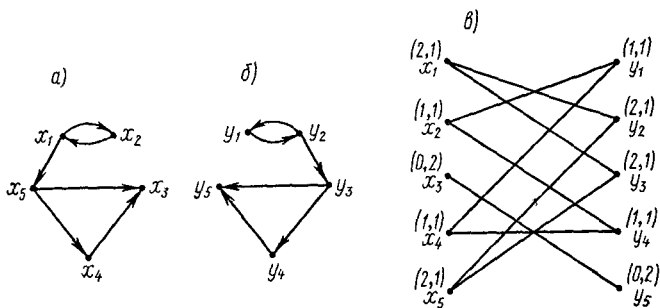


Рис. 124

Пример. 1. На рис. 124, а, б, в изображены соответственно графы G , G' и граф соответствия, построенный для графов G и G' . Из рассмотрения графа соответствия видно, что при наличии изоморфизма между G и G' вершины x_1 и x_5 могут соответствовать только вершинам y_2 и y_3 , вершины x_2 и x_4 — y_1 и y_4 , а x_3 — y_5 . Таким образом, вместо $5! = 120$ подстановок нужно проверить четыре:

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ y_2 & y_1 & y_5 & y_4 & y_3 \end{pmatrix}, \quad \begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ y_2 & y_4 & y_5 & y_1 & y_3 \end{pmatrix},$$

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ y_3 & y_1 & y_5 & y_4 & y_2 \end{pmatrix}, \quad \begin{pmatrix} x_2 & x_3 & x_4 & x_5 \\ y_4 & y_5 & y_1 & y_2 \end{pmatrix}.$$

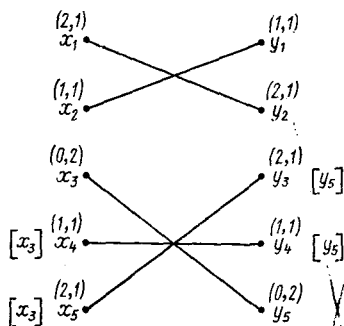


Рис. 125

С помощью модифицированного графа соответствия (см. рис. 125) можно еще уменьшить количество подстановок. Поясним способ построения такого графа на примере. Замечая, что при изоморфизме вершина x_3 обязательно переходит в вершину y_5 , выписываем около каждой вершины из G , соединенной с x_3 , символ x_3 , а у вершин G' , соединенных с y_5 , — символ y_5 . Теперь видно, что x_5 с полустепенями

(2,1) может переходить только в y_3 , но не в y_2 . Рассуждая далее так же, получим изображенный на рис. 125 модифицированный граф соответствия, который задает единственную подстановку

$$\begin{pmatrix} x_1 & x_2 & x_3 & x_4 & x_5 \\ y_2 & y_1 & y_5 & y_4 & y_3 \end{pmatrix},$$

которую только и нужно проверить по теореме 2.1. Такая проверка немедленно показывает изоморфизм графов.

В общем случае модифицированный граф соответствия строится следующим образом. Пусть $\bar{X}$ — такое подмножество X , что для каждой вершины $x_i \in \bar{X}$ не существует другой вершины $x_j \in X$ с теми же полустепенями исхода и захода. Пусть $\bar{X}'$ — подмножество в X' , строящееся анало-

гично. Соединим вершины из $\bar{X}$ с вершиной $\bar{X}'$ в том и только том случае, когда у них равны полустепени исхода и захода. Если полученный при этом граф не определяет взаимно-однозначного соответствия между $\bar{X}$ и $\bar{X}'$, то считаем модифицированный граф соответствия неопределенным. Если же такое соответствие установлено, то около каждой вершины x_i из $X \setminus \bar{X}$ выпишем те $x_j \in \bar{X}$, для которых $(x_i, x_j) \in \Gamma$, и те $x_j \in \bar{X}$, для которых $(x_j, x_i) \in \Gamma$ (объединяя эти вершины в две группы в соответствии с направлением стрелок). Прделаав то же самое с вершинами $X' \setminus \bar{X}'$, соединяем $x \in X \setminus \bar{X}$ с $x' \in X' \setminus \bar{X}'$ тогда и только тогда, когда у них одинаковы полустепени исхода и захода, и, кроме того, выписанные около x и x' вершины отвечают друг другу при взаимно-однозначном соответствии между $\bar{X}$ и $\bar{X}'$, определенном выше.

Очевидно, что если G изоморфен G' , то должно быть выполнено условие: *каждая вершина X соединена с какой-либо вершиной X' и наоборот. При этом в каждой компоненте связности модифицированного графа соответствия содержится одинаковое число вершин из X и X' .*

Если это условие выполнено, то модифицированный граф соответствия определяет одно или несколько взаимно-однозначных отображений X и X' . Сформулируем теперь алгоритм распознавания изоморфизма графов.

А л г о р и т м 2.1

1°. Проверить условие $|X| = |X'|$. Если оно выполнено — перейти к 2°, если нет — к 8°.

2°. Выписать вершины G' и G'' с их полустепенями исхода и захода (p, s) . Определить множества $\bar{X}$ и $\bar{X}'$. Если $\bar{X} = X$ или $\bar{X}' = X'$, то перейти к 4°.

3°. Выписать около каждой вершины совокупности смежных им вершин из $\bar{X}$, $\bar{X}'$. Построить модифицированный граф соответствия, если он определен — перейти к 5°, если нет — к 8°.

4°. Построить граф соответствия.

5°. Проверить условие В, если оно выполнено, перейти к 6°, если нет — к 8°.

6°. Проверить все подстановки, определяемые графом соответствия. Если есть подстановки, реализующие изоморфизм, — перейти к 7°, если нет — к 8°.

7°. Результат: графы изоморфны и изоморфизм определяют эти подстановки.

8°. Графы не изоморфны.

В большинстве практических задач этот алгоритм позволяет свести перебор к сравнительно малому количеству подстановок, реализуемый на ЭЦВМ даже при весьма больших n . Исключение представляет тот случай, когда все (или почти все) пары (p, s) равны между собой.

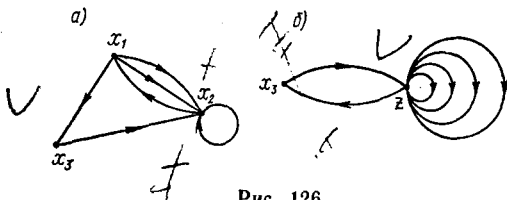


Рис. 126

✓ **3. Компоненты связности.** При помощи матриц смежности графов можно определить количество компонент связности. Для этого определим операцию элементарной склейки вершин на мультиграфе и выясним, как эта операция преобразует матрицу смежности.

Определение 2.2. Мультиграф $G' = (X', \Gamma')$ получен из мультиграфа $G = (X, \Gamma)$ при помощи операции элементарной склейки вершин x_i и x_j из X , если

1) $X' = (X \setminus (\{x_i\} \cup \{x_j\})) \cup \{z\}$, где $z \notin X$;

2) $(x_m, x_i, n) \in \Gamma'$ ($m, l \neq i, j$) тогда и только тогда, когда $(x_m, x_i, n) \in \Gamma$,

$(x_m, z, n) \in \Gamma'$ ($m \neq i, j$) тогда и только тогда, когда или $n = 2k$ и $(x_m, x_i, k) \in \Gamma$ или $n = 2k + 1$ и $(x_m, x_j, k) \in \Gamma$,

$(z, z, n) \in \Gamma'$ тогда и только тогда, когда или $n = 4k$ и $(x_i, x_i, k) \in \Gamma$, или $n = 4k + 1$ и $(x_j, x_j, k) \in \Gamma$, или $n = 4k + 2$ и $(x_i, x_j, k) \in \Gamma$, или $n = 4k + 3$ и $(x_j, x_i, k) \in \Gamma$,

$(z, x_m, n) \in \Gamma'$ тогда и только тогда, когда $n = 2k$ и $(x_i, x_m, k) \in \Gamma$ или $n = 2k + 1$ и $(x_j, x_m, k) \in \Gamma$.

Иными словами, при склейке вершин x_i и x_j склеиваются и концы дуг, совпадающие с x_i и x_j , а сами дуги сохраняются. Иллюстрацией может служить рис. 126, где изображен граф G' (рис. 126, б), полученный элементарной склейкой вершин x_1 и x_2 графа G , изображенного на рис. 126, а. Очевидно, склеивание двух вершин, лежащих в одной и той же компоненте связности, не изменяет количества компонент связности.

Обозначим $|X|$ через n . Перенумеруем множество вершин графа, полученного элементарной склейкой x_i и x_j следующим образом. Номера вершин, начиная с первого до $i-1$, сохраним. Номера вершин, начиная с $i+1$ до $j-1$, уменьшим на единицу, номера остальных вершин уменьшим на два, а вершине z присвоим номер $n-1$:

вершины $x_1 \dots x_{i-1} \quad x_{i+1} \dots x_{j-1} \quad x_{i+1} \dots x_n \quad z$;
 новые номера $1 \dots i-1 \quad i \dots j-2 \quad j-1 \dots n-2 \quad n-1$

Обозначим через $f(k)$ ($k = 1, 2, \dots, n-2$) старый номер вершины с новым номером k . Тогда матрица $\|a'_{lk}\|$ нового графа строится по матрице $\|a_{lk}\|$ первоначального графа по формулам:

$$a'_{lk} = a_{f(l)f(k)} \quad (l, k \leq n-2), \quad a'_{n-1 k} = a_{if(k)} + a_{jf(k)} \quad (k \leq n-2),$$

$$a'_{l n-1} = a_{f(l)i} + a_{f(l)j} \quad (l \leq n-2), \quad a'_{n-1 n-1} = a_{ii} + a_{jj}.$$

Например, матрицы смежности G и G' (см. рис. 126) есть соответственно

$$\begin{vmatrix} 0 & 2 & 1 \\ 1 & 1 & 0 \\ 0 & 1 & 0 \end{vmatrix} \quad \text{и} \quad \begin{vmatrix} 0 & 1 \\ 1 & 4 \end{vmatrix}.$$

Заметим еще, что если в матрице смежности мультиграфа отличны от нуля лишь элементы, стоящие на главной диагонали, то число компонент связности равно числу вершин мультиграфа, т. е. размеру матрицы, так как мультиграф содержит только петли. На основании этого сформулируем алгоритм подсчета числа компонент связности по матрице мультиграфа.

А л г о р и т м 2.2

1°. Найти ненулевой элемент матрицы a_{ij} , не стоящий на главной диагонали, если он существует, перейти к 2°, если нет — к 3°.

2°. Произвести над матрицей операцию, отвечающую склейке вершин x_i и x_j , перейти к 1°.

3°. Подсчитать количество p строк матрицы.

Результат: количество компонент связности мультиграфа равно p .

4. Операции над матрицами. Перейдем теперь к разбору операций над графами* и соответствующих им операций над матрицами. Из таких операций рассмотрим операции умножения и суперпозиции графов, отвечающих параллельной и последовательной работе конечных автоматов (см. § 1).

Начнем с параллельной работы. Множеством вершин графа $G = G_1 \times G_2$ будет множество $X_1 \times X_2$ (если $G_1 = (X_1, \Gamma_1)$, $G_2 = (X_2, \Gamma_2)$). Если графы G_1 и G_2 (с соответствующими отображениями f_1, f_2 в множества индексов) отвечают конечным автоматам Z_1 и Z_2 , то равенство $X = X_1 \times X_2$ очевидно. Далее, если возможны (при некоторых входных буквах на входе) переходы из $x_i^{(1)}$ в $x_j^{(1)}$ и из $x_k^{(2)}$ в $x_l^{(2)}$, то возможен и переход из $(x_i^{(1)}, x_k^{(2)})$ в $(x_j^{(1)}, x_l^{(2)})$ в автомате Z , полученном параллельным соединением Z_1 и Z_2 .

Определение 2.3. Граф $G = G_1 \times G_2 = (X, \Gamma)$ называется произведением графов G_1 и G_2 , если $X = X_1 \times X_2$ и $((x_i^{(1)}, x_k^{(2)}), (x_j^{(1)}, x_l^{(2)})) \in \Gamma$ тогда и только тогда, когда $(x_i^{(1)}, x_j^{(1)}) \in \Gamma_1$ и $(x_k^{(2)}, x_l^{(2)}) \in \Gamma_2$ (с точностью до порядка сомножителей $\Gamma = \Gamma_1 \times \Gamma_2$).

Для того чтобы рассматривать матрицу смежности графа G , нужно упорядочить элементы из $X_1 \times X_2$ (предполагая, что X_1 и X_2 уже как-либо упорядочены). Будем упорядочивать произведение в так называемом лексикографическом порядке, т. е. расставляем их так:

$$(x_1^{(1)}, x_1^{(2)}), \dots, (x_1^{(1)}, x_l^{(2)}), (x_2^{(1)}, x_1^{(2)}), \dots, (x_2^{(1)}, x_l^{(2)}), \dots, (x_m^{(1)}, x_1^{(2)}), \dots, (x_m^{(1)}, x_l^{(2)}),$$

где $m = |X_1|$, $l = |X_2|$. Иными словами, если i, j, k, l — целые числа, то считаем $(i, k) < (j, l)$, если $i < j$ (при любых k и l) или $i = j$ и $k < l$. Тогда, если $A_1 = \|a_{ij}^{(1)}\|$, $A_2 = \|a_{ij}^{(2)}\|$, то матрица A будет иметь вид $A = \|b_{(i,k),(j,l)}\|$, причем $b_{(i,k),(j,l)} = a_{ij}^{(1)} a_{kl}^{(2)}$.

В развернутом виде A является клеточной матрицей вида

* Начиная с этого момента, для простоты рассматриваются графы без кратных ребер.

$$\left\| \begin{array}{cccc} a_{11}^{(1)} A_2 & a_{12}^{(1)} A_2 & \dots & a_{1m}^{(1)} A_2 \\ a_{21}^{(1)} A_2 & a_{22}^{(1)} A_2 & \dots & a_{2m}^{(1)} A_2 \\ \dots & \dots & \dots & \dots \\ a_{l1}^{(1)} A_2 & a_{l2}^{(1)} A_2 & \dots & a_{lm}^{(1)} A_2 \end{array} \right\|,$$

т. е. она составлена указанным образом из матриц $a_{ij}^{(1)} A_2$, где $a_{ij}^{(1)} A_2 = A_2$ при $a_{ij}^{(1)} = 1$ и нулевой матрице при $a_{ij}^{(1)} = 0$. Всякая матрица, полученная указанным способом из некоторой матрицы меньшего размера, называется *правильной клеточной матрицей*.

Определение 2.4: Матрица A порядка $n = ml$ называется *правильной клеточной матрицей (ПКМ)*, если ее можно записать в виде

$$\left\| \begin{array}{ccc} A_{11} & \dots & A_{1m} \\ \dots & \dots & \dots \\ A_{m1} & \dots & A_{mm} \end{array} \right\|,$$

где A_{ik} — матрицы порядка l , причем все ненулевые A_{ik} ($i, k = 1, 2, \dots, m$) равны между собой.

Например, из двух матриц

$$\left\| \begin{array}{cc|cc} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array} \right\| \quad \text{и} \quad \left\| \begin{array}{cc|cc} 0 & 1 & 1 & 1 \\ 1 & 0 & 1 & 0 \\ \hline 0 & 0 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{array} \right\|$$

первая есть ПКМ, а вторая — нет.

Рассмотрим теперь преобразование над графами, соответствующее последовательной работе автоматов Z_1 и Z_2 . Пусть нагруженные графы (G_1, f_1) и (G_2, f_2) соответствуют автоматам Z_1 и Z_2 . Маркировка дуг графов G_1 и G_2 позволяет задать разбиения множества дуг, выходящих из данной вершины в каждом из этих графов, следующим образом. В графе G_1 для множества дуг $(x_{i_0}^{(1)}, x_i^{(1)})$ зададим разбиение на классы, относя к одному классу дуги, помеченные одной и той же буквой *выходного* алфавита автомата Z_1 , т. е. дуги $(x_{i_0}^{(1)}, x_i^{(1)})$ и $(x_{i_0}^{(1)}, x_j^{(1)})$ относятся к одному классу тогда и только тогда, когда

$$\text{Pr}_2 f_1((x_{i_0}^{(1)}, x_i^{(1)})) = \text{Pr}_2 f_1((x_{i_0}^{(1)}, x_j^{(1)})),$$

где $\text{Пр}_2 : A \times B \rightarrow B$ — проекция на вторую компоненту произведения (и далее $\text{Пр}_1 : A \times B \rightarrow A$ — проекция на первую компоненту).

В графе G_2 зададим разбиение дуг $(x_{k_0}^{(2)}, x_k^{(2)})$ по букве входного для Z_2 алфавита (отметим, что выходной алфавит Z_1 есть входной алфавит Z_2), т. е. дуги $(x_{k_0}^{(2)}, x_k^{(2)})$ и $(x_{k_0}^{(2)}, x_l^{(2)})$ лежат в одном классе, если

$$\text{Пр}_1 f_2(x_{k_0}^{(2)}, x_k^{(2)}) = \text{Пр}_1 f_2(x_{k_0}^{(2)}, x_l^{(2)}).$$

Заметим, что в силу однозначности работы Z_2 из этого условия вытекает равенство $x_k^{(2)} = x_l^{(2)}$, т. е. каждый класс содержит в точности один элемент.

Теперь заметим, что переход из вершины $(x_l^{(1)}, x_k^{(2)})$ в вершину $(x_j^{(1)}, x_l^{(2)})$ возможен тогда и только тогда, когда дуги $(x_l^{(1)}, x_j^{(1)})$ и $(x_k^{(2)}, x_l^{(2)})$ принадлежат классам, помеченным одной и той же буквой. Заметим, что число ребер, выходящих из $x_{k_0}^{(2)}$, одинаково при любых k_0 , если только работа автомата всегда определена.

Сейчас мы несколько обобщим эти рассуждения, дав определение суперпозиции графов. Предварительно удобно определить две теоретико-множественные операции над графами и матрицами — объединение и пересечение.

✓ **Определение 2.5.** Объединением графов $G_1 = (X_1, \Gamma_1)$ и $G_2 = (X_2, \Gamma_2)$ называется граф $G = (X, \Gamma)$, где $X = X_1 \cup X_2$, $\Gamma = \Gamma_1 \cup \Gamma_2$.

Определение 2.6. Пересечением графов $G = (X_1, \Gamma_1)$ и $G_2 = (X_2, \Gamma_2)$ называется граф $G = (X, \Gamma)$, где $X = X_1 \cap X_2$, $\Gamma = \Gamma_1 \cap \Gamma_2$.

В этих определениях Γ_1 и Γ_2 — подмножества

$$(X_1 \cup X_2) \times (X_1 \cup X_2). \quad \text{При этом } \Gamma_1 \cap \Gamma_2 \subset (X_1 \cap X_2) \times (X_1 \cap X_2),$$

и определение 4 корректно.

Соответствующие операции над матрицами изучим при $X_1 = X_2$. Заметим, что элементы матриц графов без кратных ребер могут быть равны только нулю или единице. Определим следующие операции. Если a, b — переменные, принимающие значения нуль или единица, то определим выражения $a \vee b$ и $a \wedge b$ равенствами:

$$a \vee b = a + b - ab, \quad a \wedge b = ab.$$

Очевидно, что $a \vee b$ и $a \wedge b$ снова принимают значения нуль или единица. Если графам G_1 и G_2 отвечают матрицы $A_1 = \|a_{ij}^{(1)}\|$ и $A_2 = \|a_{ij}^{(2)}\|$, то графам $G = G_1 \cup G_2$ и $G' = G_1 \cap G_2$, как легко видеть, отвечают матрицы

$$A_1 \cup A_2 = \|a_{ij}^{(1)} \vee a_{ij}^{(2)}\| \text{ и } A_1 \cap A_2 = \|a_{ij}^{(1)} \wedge a_{ij}^{(2)}\|$$

соответственно. Эти матрицы называются пересечением и объединением матриц A_1 и A_2 .

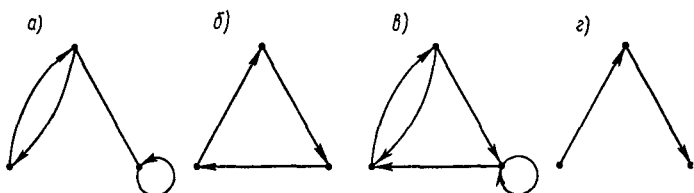


Рис. 127

Пример. На рис. 127, а, б, в и г изображены графы G_1 , G_2 , $G_1 \cup G_2$ и $G_1 \cap G_2$ соответственно, их матрицы

$$A_1 = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{vmatrix}, \quad A_2 = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{vmatrix}, \quad A_1 \cup A_2 = \begin{vmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 1 & 0 & 1 \end{vmatrix},$$

$$A_1 \cap A_2 = \begin{vmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{vmatrix}.$$

Теперь дадим определение суперпозиции графов*.

Определение 2.7. Пусть граф G_1 разбит на суграфы $G_{11}, \dots, G_{1p}$ так, что

$$G_1 = G_{11} \cup \dots \cup G_{1p}; \quad G_{1i} \cap G_{1j} = (X, \emptyset) \quad (i \neq j),$$

а граф G_2 — на суграфы $G_{21}, \dots, G_{2p}$, каждый из которых имеет не более одной выходящей дуги из каждой вершины, и при этом

$$G_2 = G_{21} \cup \dots \cup G_{2p}; \quad G_{2i} \cap G_{2j} = (X, \emptyset) \quad (i \neq j).$$

* Наше определение отличается от общепринятого; см., например, [8].

Суперпозицией графов по данным разбиениям называется граф G

$$G = G_1 \star G_2 = (G_{11} \times G_{21}) \cup \dots \cup (G_{1p} \times G_{2p}).$$

Отметим, что граф G_{1i} соответствует в теории конечных автоматов суграфу G_1 , в котором сохранены дуги, помеченные выходной буквой b_i , а G_{2i} — суграфу G_2 , в котором сохранены дуги, помеченные входной буквой b_i . Единственность дуги, выходящей из данной вершины в G_{2i} , отвечает однозначности работы автомата Z_2 .

Выясним, какой вид имеет матрица графа $G_1 \star G_2$. При этом нам удобно пользоваться упорядочением пар, обратным лексикографическому, т. е. $(i, k) < (l, m)$ тогда и только тогда, когда $k < m$ или $k = m$ и $i < l$.

Разбиения графов G_1 и G_2 порождают разбиение матриц:

$$A_1 = \bigcup_{i=1}^p A_{1i}, A_{1i} \cap A_{1j} = A^0 (i \neq j), A_2 = \bigcup_{i=1}^p A_{2i};$$

$$A_{2i} \cap A_{2j} = A^0 (i \neq j)$$

где A^0 — матрица, содержащая только нули, и в каждой строке каждой из матриц A_{2i} содержится не более одной единицы. Тогда матрица A , отвечающая $G_1 \star G_2$, равна

$$A = \bigcup_{i=1}^p A_{1i} \times A_{2i} = \left\| \begin{array}{cccc} A_{2i11} & A_{2i12} & \dots & A_{2i1m} \\ \cdot & \cdot & \cdot & \cdot \\ A_{2im1} & A_{2im2} & \dots & A_{2imm} \end{array} \right\|,$$

где в каждой клеточной строке содержится не более одной из матриц A_{2i} (некоторые клетки могут быть нулевыми). Матрицы такого вида называются *суперпозиционными клеточными матрицами*.

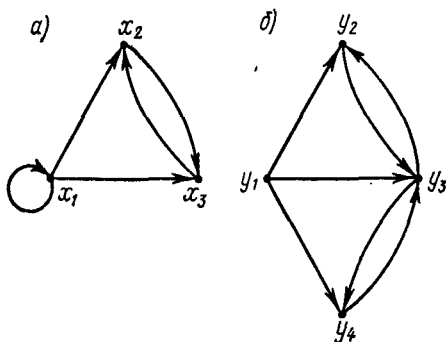


Рис. 128

✓ Пример 3. На рис. 128, а и б изображены графы G_1 и G_2 ; матрицы этих графов соответственно есть

$$A_1 = \begin{pmatrix} 1 & 1 & 1 \\ 0 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix} \text{ и } A_2 = \begin{pmatrix} 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 \end{pmatrix},$$

матрица графа $G_1 \times G_2$ есть

$$A_1 \times A_2 = \begin{pmatrix} 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 & 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 1 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix},$$

а сам граф изображен на рис. 129. Из этих рисунков видно, что разложение графов по операции $\times$ дает упрощение уже при небольшом числе вершин.

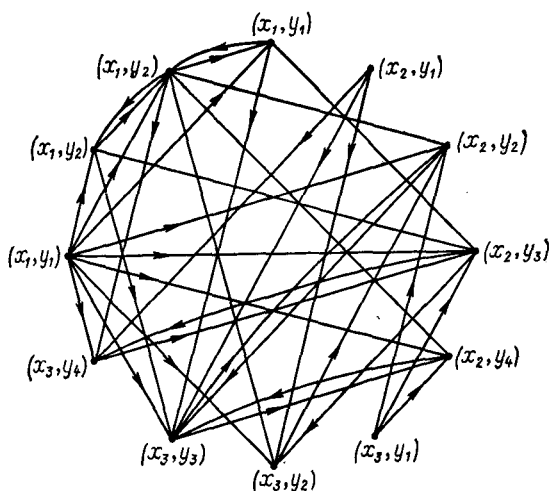


Рис. 129

✓ Пример 4. Пусть G_1, G_2 — те же графы, что и в предыдущем примере, разложение G_2 показано на рис. 130, а, а разложение G_1 — на рис. 130, б. Тогда матрицей графа $G_1 * G_2$ будет матрица

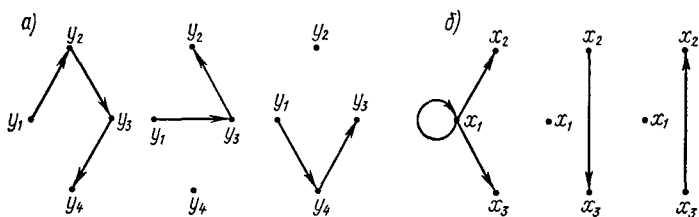


Рис. 130

$$A = \begin{vmatrix} 1 & 1 & 1 \\ 0 & 0 & 0 \\ 0 & 0 & 0 \end{vmatrix} \times \begin{vmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{vmatrix} \cup \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 0 \end{vmatrix} \times$$

$$\times \begin{vmatrix} 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 \end{vmatrix} \cup \begin{vmatrix} 0 & 0 & 0 \\ 0 & 0 & 0 \\ 0 & 1 & 0 \end{vmatrix} \times \begin{vmatrix} 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{vmatrix},$$

что при обратном лексикографическом упорядочении равно

$$A = \begin{vmatrix} 000 & 111 & 000 & 000 \\ 000 & 000 & 001 & 000 \\ 000 & 000 & 000 & 010 \\ 000 & 000 & 111 & 000 \\ 000 & 000 & 000 & 000 \\ 000 & 000 & 000 & 000 \\ 000 & 000 & 000 & 111 \\ 000 & 001 & 000 & 000 \\ 000 & 000 & 000 & 000 \\ 000 & 000 & 000 & 000 \\ 000 & 000 & 000 & 000 \\ 000 & 000 & 000 & 000 \\ 000 & 000 & 010 & 000 \end{vmatrix}.$$

Соответствующий граф изображен на рис. 131.

В заключение этого параграфа заметим, что число единиц, которые присутствуют в матрице графа $G_1 * G_2$, не может превышать числа m^2l , где m — число вершин G_1 , l — число вершин G_2 . Это непосредственно ясно из способа построения матрицы $G_1 * G_2$.

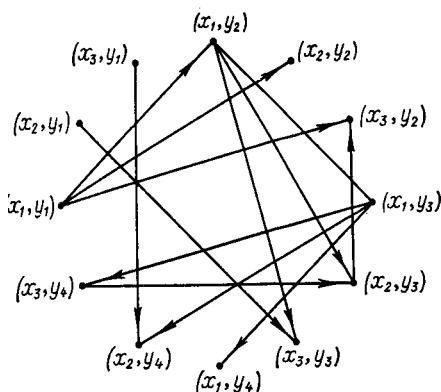


Рис. 131

§ 3. РАЗЛОЖЕНИЕ ГРАФОВ ПО ОПЕРАЦИЯМ ПРОИЗВЕДЕНИЯ И СУПЕРПОЗИЦИИ

1. Вводные замечания. В этом параграфе будут приведены алгоритмы разложения графа по операции произведения и операции суперпозиции. Число графов, непосредственно разлагающихся по этим операциям, весьма невелико, поэтому для данного графа G будем искать изоморфный ему граф G' , для которого справедливо одно из равенств

$$G' = G_1 \times G_2 \text{ или } G_1 \star G_2,$$

где G_1 и G_2 — более простые по сравнению с G графы.

Очевидно, если граф разлагается в этом смысле по одной из двух операций, то существует такая подстановка $\sigma \in S_n$ ($n = |X|$), которая переводит матрицу графа G в соответствии с теоремой 2.1 в правильную клеточную матрицу (ПКМ) или в суперпозиционную клеточную матрицу (СКМ), в зависимости от операции, по которой ведется разложение. Чтобы выяснить существование такой подстановки, нам нужно пересмотреть $n!$ перестановок, что в случае больших n весьма затруднительно. Ниже предлагается метод уменьшения числа проверяемых подстановок. Заметим, что если матрица A есть ПКМ или СКМ, то одновременная перестановка клеточных строк и столбцов или одновременная перестановка строк и столбцов каждой клетки с помощью одной и той же подстановки оставляет матрицу ПКМ или СКМ соответственно: кроме того, $n = ml$ ($m, l \neq n$).

Пусть $n = ml$, S_n — группа подстановок на n элементах, S_m и S_l — группы подстановок m и l элементов соответственно. Обозначим через $S_m S_l$ подгруппу S_n , элементы которой отвечают паре подстановок $\sigma \in S_m$, $\sigma' \in S_l$, и действуют так: подстановка $\sigma \in S_m$ представляет собой «блоки» по l элементов, а $\sigma' \in S_l$ — элементы каждого блока из числа m блоков. Тогда если σ переводит матрицу A в ПКМ (или СКМ), то подстановка $\sigma\sigma_1$ ($\sigma_1 \in S_m S_l$) тоже переводит A в ПКМ (соответственно СКМ), и наоборот.

Таким образом, из каждого класса эквивалентности по $S_m S_l$ нужно проверить по одной подстановке.

Каждой подстановке σ сопоставим теперь пару разбиений множества $\{1, \dots, n\}$ при $n = ml$ следующим образом. Рассмотрим числа $\sigma(1), \dots, \sigma(n)$. Классы первого из разбиений есть

$$\{\sigma(1), \dots, \sigma(l)\}, \{\sigma(l+1), \dots, \sigma(2l)\}, \dots, \{\sigma((m-1)l+1), \dots, \sigma(ml)\}$$

(число классов m). Обозначим эти множества $P_1, \dots, P_m$. Классы второго разбиения есть

$$\{\sigma(1), \sigma(l+1), \dots, \sigma((m-1)l+1)\}, \{\sigma(2), \sigma(l+2), \dots, \sigma((m-1)l+2)\}, \dots, \{\sigma(l), \sigma(2l), \dots, \sigma(ml)\}$$

Эти классы обозначим $Q_1, \dots, Q_l$. Так построена пара разбиений (P, Q) , обладающая свойством: для любых i, j , $P_i \cap Q_j$ содержит ровно один элемент. И обратно, пара (P, Q) с этим свойством определяет подстановку. Именно через $\sigma(1)$ обозначим элемент, содержащийся в $P_1 \cap Q_1$, $\sigma(2)$ — элемент, содержащийся в $P_1 \cap Q_2$, и так далее (пользуемся лексикографическим упорядочением множеств $P_i \cap Q_j$). Очевидно, что каждая подстановка из $S_m S_l$ только перенумеровывает элементы P и элементы Q (не смешивая их между собой), не изменяя самих разбиений. Наоборот, любая такая перенумеровка соответствует элементу из $S_m S_l$. Пары (P, Q) , отличающиеся только нумерацией, назовем *эквивалентными*. Очевидно, что нужно выбрать по одной паре (P, Q) из каждого класса эквивалентности и проверить подстановки, обратные тем, которые отвечают выбранным парам. Нетрудно видеть, что количество подстановок, подлежащих проверке, уменьшено до $\frac{n!}{k!l!}$.

2. Операция умножения. Если $s(x)$, $p(x)$ — полустепени исхода и захода вершины x графа $G = G_1 \times G_2$, то справедливы соотношения

$$s(x_i, y_j) = s(x_i) s(y_j); \quad p(x_i, y_j) = p(x_i) p(y_j).$$

В связи с этим из всех возможных разбиений чисел $s(x)$ и $p(x)$ графа G на множители

$$s_\alpha = s_{1\alpha} s_{2\alpha}; \quad p_\alpha = p_{1\alpha} p_{2\alpha} \quad (3.1)$$

выберем всевозможные разбиения, подчиняющиеся условиям:

$$\text{если } x_\alpha, x_\beta \in P_i, \text{ то } s_{1\alpha} = s_{1\beta} \text{ и } p_{1\alpha} = p_{1\beta}, \quad (3.2)$$

$$\text{если } x_\alpha, x_\beta \in Q_j, \text{ то } s_{2\alpha} = s_{2\beta} \text{ и } p_{2\alpha} = p_{2\beta}.$$

Теперь сформулируем алгоритм разложения графа по операции $\times$.

А л г о р и т м 3.1

1°. Подсчитать число вершин n графа G . Если $n = ml$ ($m, l \neq n$), перейти к 2°, если n простое — к 8°.

2°. Записать матрицу смежности G и проверить, не является ли она ПКМ, если да — перейти к 7°, если нет — к 3°.

3°. Для каждой вершины образовать всевозможные разложения чисел s и p на множители, аналогично равенству (3.1).

4°. Образовать пару (P, Q) , удовлетворяющую соотношениям (3.2). Перейти к 5°. Если такой пары нет, перейти к 8°.

5°. Найти подстановку, отвечающую паре (P, Q) . Обратную ей подстановку применить в матрице смежности A . Если получится ПКМ, перейти к 7°.

6°. Построить новую пару (P, Q) , неэквивалентную всем предыдущим и удовлетворяющую (3.2). Перейти к 5°. Если такой пары нет — перейти к 8°.

7°. По ПКМ построить матрицы смежности графов G_1 и G_2 . Результат: $G = G_1 \times G_2$.

8°. Граф G не разложим в произведение двух графов.

3. Операция суперпозиции. Приведем теперь алгоритм разложения G по операции суперпозиции.

А л г о р и т м 3.2

1°. Подсчитать число вершин n графа G . Если $n = ml$ ($m, l \neq n$) — перейти к 2°, если n простое — к 6°.

2°. Записать матрицу смежности G и проверить условие $\sum_{i,j} a_{ij} \leq m^2 l$. Если оно выполнено, перейти к 3°, если нет — к 6°.

3°. Проверить, не является ли полученная матрица СКМ, если да — перейти к 5°, если нет — к 4°.

4°. Образовать произвольную пару разбиений (P, Q) (неэквивалентную всем рассматривавшимся ранее). Перейти к 3°. Если такой пары не существует, перейти к 6°.

5°. По данной СКМ построить матрицы графов G_1 и G_2 . Результат: $G = G_1 * G_2$.

6°. Граф G не разложим в суперпозицию двух графов.

Рассмотрим примеры работы алгоритмов.

П р и м е р 1. Пусть матрица графа G есть

$$A = \left\| \begin{array}{cc|cc} 0001 & 0000 \\ 0001 & 0101 \\ 0000 & 1000 \\ 0100 & 1000 \\ \hline 1011 & 0111 \\ 0011 & 0000 \\ 0001 & 0001 \\ 0011 & 0011 \end{array} \right\|.$$

Применим алгоритм 3.1.

1°. Здесь $n = 8$. Пусть $m = 2, l = 4$.

2°. Проверка показывает, что A не есть ПКМ. Переходим к 3°. Выписываем

$$\begin{array}{ll} s_1 = 1 = 1 \cdot 1 & p_1 = 1 = 1 \cdot 1 \\ s_2 = 3 = 3 \cdot 1 \text{ или } 1 \cdot 3 & p_2 = 1 = 1 \cdot 1 \\ s_3 = 1 = 1 \cdot 1 & p_3 = 3 = 1 \cdot 3 \text{ или } 3 \cdot 1 \\ s_4 = 2 = 1 \cdot 2 \text{ или } 2 \cdot 1 & p_4 = 6 = 1 \cdot 6 \text{ или } 6 \cdot 1 \text{ или } 2 \cdot 3 \text{ или } 3 \cdot 2 \\ s_5 = 6 = 1 \cdot 6 \text{ или } 6 \cdot 1 \text{ или } 2 \cdot 3 \text{ или } 3 \cdot 2 & p_5 = 2 = 1 \cdot 2 \text{ или } 2 \cdot 1 \\ s_6 = 2 = 2 \cdot 1 \text{ или } 1 \cdot 2 & p_6 = 2 = 1 \cdot 2 \text{ или } 2 \cdot 1 \\ s_7 = 2 = 2 \cdot 1 \text{ или } 1 \cdot 2 & p_7 = 2 = 1 \cdot 2 \text{ или } 2 \cdot 1 \\ s_8 = 4 = 1 \cdot 4 \text{ или } 4 \cdot 1 \text{ или } 2 \cdot 2 & p_8 = 4 = 1 \cdot 4 \text{ или } 4 \cdot 1 \text{ или } 2 \cdot 2 \end{array}$$

5°. Легко проверить, что пара разбиений

$$P_1 = \{4, 5, 6, 8\}, \quad P_2 = \{1, 2, 3, 7\}$$

$$Q_1 = \{1, 6\}, \quad Q_2 = \{3, 4\}, \quad Q_3 = \{7, 8\}, \quad Q_4 = \{2, 5\}$$

удовлетворяет соотношениям (3.2) для выделенных способов разложения на множители. Соответствующая подстановка и обратная ей:

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 6 & 4 & 8 & 5 & 1 & 3 & 7 & 2 \end{pmatrix} \text{ и } \sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 5 & 8 & 6 & 2 & 4 & 1 & 7 & 3 \end{pmatrix};$$

подстановка σ^{-1} переводит A в матрицу A' :

$$A' = \left\| \begin{array}{cc|cc} 0100 & 0100 & & \\ 0001 & 0001 & & \\ 0110 & 0110 & & \\ 1110 & 1110 & & \\ \hline 0100 & 0000 & & \\ 0001 & 0000 & & \\ 0110 & 0000 & & \\ 1110 & 0000 & & \end{array} \right\|,$$

которая есть ПКМ. Переходим к 7⁰.

7⁰. Легко видеть, что матрицы A_1 и A_2 для G_1 и G_2 соответственно имеют вид:

$$A_1 = \left\| \begin{array}{cc} 1 & 1 \\ 1 & 0 \end{array} \right\|; \quad A_2 = \left\| \begin{array}{cccc} 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 1 \\ 0 & 1 & 1 & 0 \\ 1 & 1 & 1 & 0 \end{array} \right\|.$$

Результат: $G = G_1 \times G_2$.

Пример 2. Пусть матрица графа G есть

$$\left\| \begin{array}{cc|cc|cc|cc} 00 & 00 & 00 & 00 & & & & \\ 00 & 01 & 00 & 00 & & & & \\ \hline 00 & 00 & 00 & 00 & & & & \\ 00 & 10 & 00 & 00 & & & & \\ \hline 00 & 00 & 10 & 00 & & & & \\ 01 & 00 & 00 & 01 & & & & \\ \hline 00 & 00 & 00 & 10 & & & & \\ 00 & 00 & 00 & 01 & & & & \end{array} \right\|.$$

1⁰. $n = 8$. Представляем $n = 8 = 2 \cdot 4$.

2⁰. Условие $\sum_{i,j} a_{ij} = 7 < 2^2 \cdot 4 = 16$ выполнено.

3⁰. Проверим, является ли A СКМ. Для этого выпишем различные ненулевые 2×2 клетки

$$\left\| \begin{array}{cc} 0 & 0 \\ 1 & 0 \end{array} \right\|, \quad \left\| \begin{array}{cc} 1 & 0 \\ 0 & 0 \end{array} \right\|, \quad \left\| \begin{array}{cc} 0 & 0 \\ 0 & 1 \end{array} \right\|, \quad \left\| \begin{array}{cc} 1 & 0 \\ 0 & 1 \end{array} \right\|.$$

Видно, что они не могут образовывать нужное разбиение матрицы, так как, например, две последние клетки имеют непустое пересечение.

4°. Отметим, что единичной подстановке (рассмотренной в 3°) отвечает пара разбиений

$$\{1, 2\}, \{3, 4\}, \{5, 6\}, \{7, 8\}; \{1, 3, 5, 7\} \{2, 4, 6, 8\}.$$

Чтобы из этой пары получить неэквивалентную ей пару разбиений, нужно в одном из них (например, во втором) поменять местами два элемента (например, 1 и 2) разных классов, которые при первом разбиении принадлежат одному и тому же классу. Получим пару разбиений

$$\{1, 2\}, \{3, 4\}, \{5, 6\}, \{7, 8\}; \{2, 3, 5, 7\}, \{1, 4, 6, 8\},$$

отвечающую подстановке

$$\sigma = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 1 & 3 & 4 & 5 & 6 & 7 & 8 \end{pmatrix}.$$

Применяя обратную подстановку (которая в данном случае совпадает с σ , так как $\sigma^2 = e$) к матрице A , получим A' .

3°. Предоставляем читателю проверить, что A' не является СКМ. Продолжая проверку подстановок последовательным применением 4° и 3°*, придем к подстановке

$$\sigma_1 = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 8 & 1 & 5 & 3 & 7 & 2 & 6 & 4 \end{pmatrix}, \text{ для которой } \sigma^{-1} = \begin{pmatrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 \\ 2 & 6 & 4 & 8 & 3 & 7 & 5 & 1 \end{pmatrix}.$$

Применяя подстановку σ^{-1} к A , получим

$$A'' = \begin{vmatrix} 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{vmatrix},$$

различные 2×2 клетки которой есть

$$\begin{vmatrix} 0 & 1 \\ 0 & 0 \end{vmatrix}, \quad \begin{vmatrix} 1 & 0 \\ 0 & 0 \end{vmatrix}, \quad \begin{vmatrix} 0 & 0 \\ 1 & 0 \end{vmatrix},$$

причем в каждой клеточной строке каждая такая клетка встречается не более одного раза. Поэтому матрица A'' есть СКМ. Матрицы графов G_1 и G_2 вместе с разбиениями их на подматрицы выписаны ниже:

* Поскольку число подстановок, подлежащих проверке, равно $8! (2!4!)^{-1} = 840$, мы лишены возможности привести работу алгоритма полностью.

$$A_1 = \begin{Bmatrix} 1 & 1 \\ 1 & 0 \end{Bmatrix} = \begin{Bmatrix} 0 & 1 \\ 0 & 0 \end{Bmatrix} \cup \begin{Bmatrix} 1 & 0 \\ 1 & 0 \end{Bmatrix}; \quad A_2 = \begin{Bmatrix} 0100 \\ 0001 \\ 0110 \\ 1110 \end{Bmatrix} =$$

$$= \begin{Bmatrix} 0100 \\ 0000 \\ 0010 \\ 0100 \end{Bmatrix} \cup \begin{Bmatrix} 0000 \\ 0001 \\ 0000 \\ 1000 \end{Bmatrix} \cup \begin{Bmatrix} 0000 \\ 0000 \\ 0100 \\ 0010 \end{Bmatrix}.$$

Соответствующие графы с разбиением показаны на рис. 132.

Предоставляем читателю выписать графическую интерпретацию графа $G_1 * G_2$.

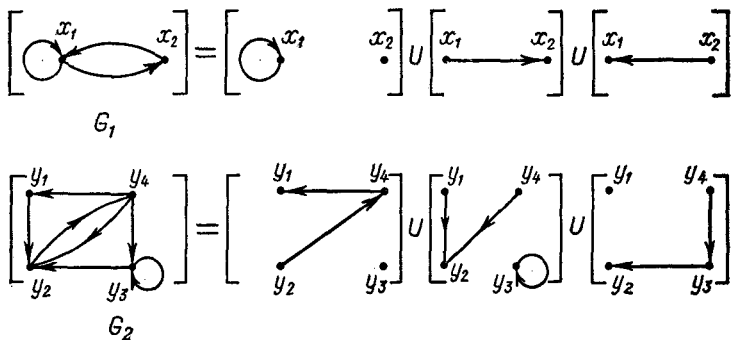


Рис. 132

Приведем формулировки некоторых общих задач теории автоматов.

1. Задача синтеза конечного автомата. По данному отображению множества слов над алфавитом A в множество слов над алфавитом B найти автомат Z , реализующий данное отображение.

2. Задача анализа конечного автомата. По данному графоиду конечного автомата Z найти отображение слов над его входным алфавитом в множество слов над его выходным алфавитом, которое реализует этот автомат.

3. Задача декомпозиции автоматов, частные случаи которой обсуждались выше.

При решении всех трех задач аппарат теории графов оказывается весьма полезным. Эти вопросы можно найти в специальной литературе (см. например, цитированную выше книгу А. Н. Мелихова и приведенную в ней обширную библиографию).

Глава XII

ПРОИЗВОДЯЩИЕ ФУНКЦИИ И ИХ ПРИМЕНЕНИЕ В ПЕРЕЧИСЛИТЕЛЬНЫХ ЗАДАЧАХ ТЕОРИИ ГРАФОВ

Предварительное понятие о задачах, которыми занимается теория перечисления графов, можно получить, попытавшись ответить, например, на такие вопросы. Сколько существует неизоморфных графов, состоящих из n вершин и m дуг, или, в более общем виде, — сколько существует неизоморфных графов из n вершин, обладающих одним из следующих свойств: или состоящих из m компонент связности, или связных и не имеющих циклов, или плоских, или с четными степенями вершин и т. п.? Эти задачи относятся, в сущности, к комбинаторике. Вместе с тем приемы решения перечислительных задач теории графов имеют свою специфику. В этой главе будет рассмотрено несколько комбинаторных задач, с тем чтобы способом, предложенным Пойа [10], ввести мощный инструмент комбинаторики — производящие функции.

§ 1. РИСУНОЧНЫЕ РЯДЫ И ПРОИЗВОДЯЩИЕ ФУНКЦИИ

1. Сочетания. Начнем с достаточно простой задачи о сочетаниях.

Определение 1.1. Пусть имеется n каких-то отличных друг от друга объектов $a_1, a_2, \dots, a_n$. Назовем m — сочетанием из n элементов $a_1, a_2, \dots, a_n$ выборку m из них без учета порядка.

Найдем C_n^m — число m — сочетаний из n элементов. Решение этой задачи начнем со случая $n = 3$. Рассмотрим рис. 133. В первом столбце квадратиком обозначено отсутствие элемента a_j ($j = 1, 2, 3$); 2-выборка a_2a_3 , например, получается, если взять левую фигуру первой строки (квадратик) и правые фигуры из второй и третьей строк. Вообще, каждая m -выборка ($m \leq 3$), получается, если взять по одной фигуре из каждой строки, причем, m из них из

второго столбца. Таким образом, рис. 133 содержит явно альтернативы для каждого предмета — войти или не войти в выборку, и неявно — все возможные выборки из трех предметов.

Представим выборку a_2a_3 в виде символического произведения, изображенного на рис. 134, в котором сомножителями являются левая фигура первой строки и правые фигуры второй и третьей строки рис. 133. Каждая m -вы-

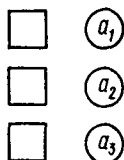


Рис. 133

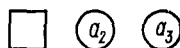


Рис. 134

борка также может быть записана в виде символического произведения, а все возможные выборки представлены различными фигурами, изображенными на рис. 135.

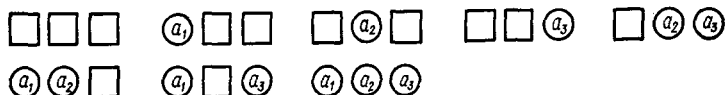


Рис. 135

Следующий важный шаг состоит в том, чтобы считать каждую строку рис. 133 и весь рис. 135 символической суммой своих фигур и ввести символическое произведение сумм рис. 133; получится ряд, изображенный на рис. 136.

Возникает вопрос, в каком смысле складываются и перемножаются фигуры. Вообще говоря, можно было бы дать строгое математическое определение этим действиям. Однако мы не будем этого делать, исходя из тех соображений, что рис. 136 дает излишнюю информацию по сравнению с

$$\begin{aligned}
 & (\square + \textcircled{a_1}) (\square + \textcircled{a_2}) (\square + \textcircled{a_3}) = \\
 & = \square \square \square + \textcircled{a_1} \square \square + \square \textcircled{a_2} \square + \square \square \textcircled{a_3} + \square \textcircled{a_2} \textcircled{a_3} + \\
 & + \textcircled{a_1} \square \textcircled{a_3} + \textcircled{a_1} \textcircled{a_2} \square + \textcircled{a_1} \textcircled{a_2} \textcircled{a_3}
 \end{aligned}$$

Рис. 136

той, которая требуется постановкой задачи. Нас интересует лишь число m -сочетаний, а рис. 136 содержит все m -сочетания явно.

Сопоставим каждой фигуре рис. 133, 135, 136 числовую переменную x в степени, равной числу содержащихся в рисунке предметов, квадратику сопоставим $x^0 = 1$ (рис. 137).

$$\square \rightarrow x^0, \textcircled{a_1} \rightarrow x^1, \dots, \square \textcircled{a_1} \textcircled{a_2} \rightarrow x^2, \dots, \textcircled{a_1} \textcircled{a_2} \textcircled{a_3} \rightarrow x^3$$

Рис. 137

При таком соответствии нижние две строки рис. 136 переходят в сумму

$$P(x) = 1 + C_3^1 x + C_3^2 x^2 + C_3^3 x^3, \quad (1.1)$$

так как рисунки, переходящие в x^m , соответствуют m -выборкам из трех предметов. Весь ряд (рис. 136) переходит в равенство

$$(1 + x)^3 = 1 + C_3^1 x + C_3^2 x^2 + C_3^3 x^3. \quad (1.2)$$

Производя фактическое возведение в степень в левой части равенства (1.2) и сравнивая коэффициенты при одинаковых степенях в левой и правой частях, получаем, что

$$C_3^1 = 3, \quad C_3^2 = 3, \quad C_3^3 = 1.$$

Возвращаясь к случаю произвольного числа n предметов и рассматривая рисунок, аналогичный рис. 133, но содержащий n строк, легко получить равенство

$$(1 + x)^n = 1 + C_n^1 x + \dots + C_n^m x^m + \dots + C_n^n x^n. \quad (1.3)$$

Это равенство дает повод назвать числа C_n^m биномиальными коэффициентами.

2. Производящая функция и перечисляющий ряд. Принципиальное значение равенства (1.3) состоит в том, что оно сводит решение комбинаторной задачи к алгебраической.

Определение 1.2. Функция $(1 + x)^n$ называется производящей для чисел C_n^m в том смысле, что C_n^m — коэффициент при x^m в разложении этой функции в ряд Тейлора, а сумма (конечный ряд) в правой части равенства (1.3) называется перечисляющим рядом.

Из формулы (1.3) вытекает, что $C_n^n = 1$. Хотя число C_n^0 не имеет комбинаторного смысла, удобно положить $C_n^0 = 1$ и переписать равенство (1.3) в виде

$$(1+x)^n = \sum_{j=0}^n C_n^j x^j. \quad (1.3')$$

Из (1.3) или (1.3') получаем цепочку равенств

$$C_n^m = \frac{1}{m!} \frac{d^m}{dx^m} (1+x)^n \Big|_{x=0} = \frac{n(n-1) \dots (n-m)}{m!} = \frac{n!}{(n-m)! m!}. \quad (1.4)$$

Производящая функция $(1+x)^n$ удобна при нахождении различных, полезных в ряде задач, соотношений между числами C_n^m . Например, из тождества

$$(1+x)^{n+1} = (1+x)^n (1+x)$$

немедленно следует рекуррентная по n формула

$$C_{n+1}^m = C_n^m + C_n^{m-1}. \quad (1.5)$$

Учитывая, что $C_n^0 = C_n^n = 1$, легко получаем с помощью (1.5) следующую таблицу:

ТАБЛИЦА 12

$m \backslash n$	0	1	2	3	4	5
0	1					
1	1	1				
2	1	2	1			
3	1	3	3	1		
4	1	4	6	4	1	
5	1	5	10	10	5	1

Из равенств (1.4) вытекает соотношение

$$C_n^m = C_n^{n-m}. \quad (1.6)$$

Полагая в тождестве (1.3) $x = 1$, получаем равенство для суммы биномиальных коэффициентов

$$\sum_{j=0}^n C_n^j = 2^n. \quad (1.7)$$

С помощью производящей функции $(1+x)^n$ можно вывести ряд других интересных соотношений между числами C_n^m .

3. Выборки с различными видами предметов. Положим теперь, что имеется не n различных предметов, а n различных видов предметов. Требуется найти число m -выборок, состоящих из объектов указанных n видов, причем в каждую выборку могут входить по несколько предметов одного вида. Порядок предметов несуществен. Обозначим искомое число через $f(n, m)$.

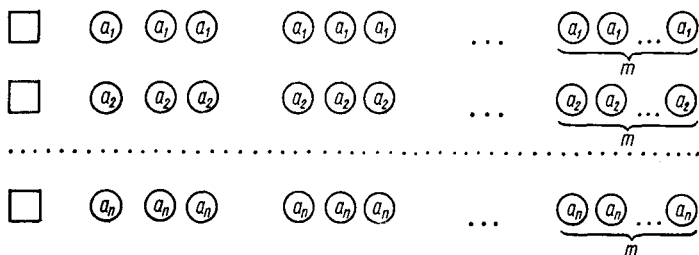


Рис. 138

Обратимся опять к вспомогательному рисунку (рис. 138). Каждый ряд рис. 138 состоит из $m+1$ фигуры, каждая фигура отвечает вхождению 0, 1, 2, ..., m предметов данного вида в m -выборку.

Рассуждения, аналогичные проведенным выше, приводят к следующему результату: производящей функцией для чисел $f(n, m)$ является функция $(1+x+x^2+\dots+x^m)^n$. Таким образом, справедливо равенство

$$f(n, m) = \frac{1}{m!} \frac{d^m}{dx^m} (1+x+x^2+\dots+x^m)^n \Big|_{x=0}. \quad (1.8).$$

Можно написать и другие производящие функции для чисел $f(n, m)$. Например, легко понять, что добавление к сумме $1+x+x^2+\dots+x^m$ степеней x^k при $k > m$ не влияет на значение правой части формулы (1.8). Следовательно, производящей функцией для чисел $f(n, m)$ будет также функция $(1-x)^{-n}$, так как

$$(1-x)^{-n} = (1+x+x^2+\dots+x^m+\dots)^n. \quad (1.9)$$

Из равенств (1.8) и (1.9) вытекает формула

$$f(n, m) = \frac{1}{m!} \frac{d^m}{dx^m} (1-x)^{-n} \Big|_{x=0}. \quad (1.8')$$

Задача 1. Пусть предметы k -го ($k \leq n$) вида входят в m -выборку не более $j(k)$ раз ($j(k) \leq m$). Найти производящую функцию для чисел $F(n, m)$ m -выборок в этом случае.

В следующем параграфе мы рассмотрим две задачи, иллюстрирующие применение изложенных здесь идей.

§ 2. ЗАДАЧИ О РАЗМЕНЕ МОНЕТ И РАССЕЧЕНИИ МНОГОУГОЛЬНИКОВ

Задача 1. Сколькими способами можно разменять монету достоинством 20 копеек на медные монеты?

Решение. Имеются четыре типа медных монет: достоинством в одну, две, три и пять копеек. Обозначим через P_n — число способов заплатить сумму в n копеек этими четырьмя типами монет. Очевидно, что $P_1 = 1$, $P_2 = 2$, $P_3 = 3$, $P_4 = 4$, $P_5 = 6$. Нужно найти P_{20} . Положим для удобства $P_0 = 1$.

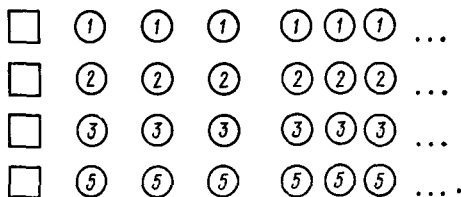


Рис. 139

Для дальнейших рассуждений полезно иметь перед глазами рисунок (рис. 139). Один из способов размена монеты в 20 копеек можно представить в виде символического произведения, изображенного на рис. 140. Все возможные спо-



Рис. 140

собы размена будут содержаться в произведении, изображенном на рис. 141.

$$\begin{aligned}
 & \left(\square + \textcircled{1} + \textcircled{1} \textcircled{1} + \textcircled{1} \textcircled{1} \textcircled{1} + \dots \right) \cdot \\
 & \left(\square + \textcircled{2} + \textcircled{2} \textcircled{2} + \textcircled{2} \textcircled{2} \textcircled{2} + \dots \right) \cdot \\
 & \left(\square + \textcircled{3} + \textcircled{3} \textcircled{3} + \textcircled{3} \textcircled{3} \textcircled{3} + \dots \right) \cdot \\
 & \left(\square + \textcircled{5} + \textcircled{5} \textcircled{5} + \textcircled{5} \textcircled{5} \textcircled{5} + \dots \right) = \\
 & = \square \square \square \square + \dots + \textcircled{1} \cdot \textcircled{2} \textcircled{2} \cdot \square \cdot \textcircled{5} \textcircled{5} \textcircled{5} + \dots
 \end{aligned}$$

Рис. 141

Для того чтобы извлечь из последней строки рис. 141 те и только те рисунки, которые отвечают способам размена монеты в 20 копеек, сопоставим каждой фигуре переменную x в степени, равной общей стоимости монет фигуры (рис. 142). При таком соответствии последняя строка рис. 141 переходит в перечисляющий ряд

$$\begin{aligned}
 \square & \rightarrow x^0, \textcircled{1} \rightarrow x^1, \textcircled{2} \rightarrow x^2, \textcircled{3} \rightarrow x^3, \textcircled{5} \rightarrow x^5, \textcircled{1} \textcircled{1} \rightarrow x^2, \\
 \textcircled{1} \cdot \square \cdot \textcircled{3} \textcircled{3} \textcircled{3} \cdot \square & \rightarrow x^9, \textcircled{1} \cdot \textcircled{2} \textcircled{2} \cdot \square \cdot \textcircled{5} \textcircled{5} \textcircled{5} \rightarrow x^{20}, \dots
 \end{aligned}$$

Рис. 142

$$P(x) = P_0 + P_1 x + P_2 x^2 + \dots + P_n x^n + \dots, \quad (2.1)$$

а весь рис. 141 соответствует равенству

$$\begin{aligned}
 (1 - x)^{-1} (1 - x^2)^{-1} (1 - x^3)^{-1} (1 - x^5)^{-1} &= P_0 + P_1 x + \\
 &+ P_2 x^2 + \dots + P_n x^n + \dots, \quad (2.2)
 \end{aligned}$$

т. е.

$$F(x) = (1 - x)^{-1} (1 - x^2)^{-1} (1 - x^3)^{-1} (1 - x^5)^{-1} \text{ — произво-}$$

водящая функция.

Опишем рекурсивную процедуру подсчета коэффициента P_{20} . Имеем разложения

$$\begin{aligned}
 (1 - x)^{-1} (1 - x^2)^{-1} (1 - x^3)^{-1} (1 - x^5)^{-1} &= \\
 &= P_0 + P_1 x + \dots + P_n x^n + \dots,
 \end{aligned}$$

$$\begin{aligned}
(1-x)^{-1} (1-x^2)^{-1} (1-x^3)^{-1} &= \\
&= a_0 + a_1 x + \dots + a_n x^n + \dots, \\
(1-x)^{-1} (1-x^2)^{-1} &= b_0 + b_1 x + \dots + b_n x^n + \dots, \\
(1-x)^{-1} &= 1 + x + \dots + x^n + \dots,
\end{aligned}
\tag{2.3}$$

причем $b_0 = b_1 = a_0 = a_1 = 1$, $a_2 = 2$, $P_5 = 6$. Из первых двух разложений (2.3) вытекает равенство

$$\begin{aligned}
(1-x^5) (P_0 + P_1 x + \dots + P_n x^n + \dots) &= \\
&= a_0 + a_1 x + \dots + a_n x^n + \dots,
\end{aligned}$$

откуда следуют соотношения между коэффициентами a_j и P_j

$$a_n = P_n - P_{n-5}$$

или

$$P_n = P_{n-5} + a_n. \tag{2.4}$$

Точно так же получаем соотношения:

$$a_n = a_{n-3} + b_n, \tag{2.5}$$

$$b_n = b_{n-2} + 1. \tag{2.6}$$

Из соотношения (2.6) и того факта, что $b_0 = b_1 = 1$, имеем равенства

$$b_{2m} = 1 + m, \quad b_{2m+1} = 1 + m. \tag{2.7}$$

Применяя формулу (2.4) несколько раз, выражаем коэффициент P_{20} через коэффициенты a_j

$$\begin{aligned}
P_{20} = P_{15} + a_{20} = P_{10} + a_{15} + a_{20} = P_5 + a_{10} + a_{15} + \\
+ a_{20} = 6 + a_{10} + a_{15} + a_{20}.
\end{aligned}
\tag{2.8}$$

Воспользовавшись формулой (2.5), получаем равенства:

$$\begin{aligned}
a_{20} = a_{17} + b_{20} = \dots = a_2 + b_5 + b_8 + b_{11} + b_{14} + b_{17} + \\
+ b_{20} = 2 + 3 + 5 + 6 + 8 + 9 + 11 = 44, \\
a_{15} = a_{12} + b_{15} = \dots = a_0 + b_3 + b_6 + b_9 + b_{12} + b_{15} = \\
= 1 + 2 + 4 + 5 + 7 + 8 = 27,
\end{aligned}$$

$$a_{10} = a_7 + b_{10} = \dots = a_1 + b_4 + b_7 + b_{10} = 1 + 3 + 4 + 6 = 14.$$

Таким образом, $P_{20} = 6 + 14 + 27 + 44 = 91$. Итак, монету достоинством в 20 копеек можно разменять на медные 91 способом.



Рис. 143

Задача 2. Сколькими способами можно расsects выпуклый n -угольник на $n - 2$ треугольника $n - 3$ непересекающимися диагоналями? Многоугольник считается жестко закрепленным на плоскости.

Решение. Обозначим через D_n искомое число способов. Легко видеть, что $D_3 = 1$, $D_4 = 2$, $D_5 = 5$ (рис. 143).

Основная идея решения этой задачи заключается в следующем. Будем строить расsections многоугольника, не являющегося треугольником, из расsections других многоугольников с меньшим числом сторон. С этой целью выделяем одну из сторон многоугольника и называем ее базой. Тот треугольник, который опирается на базу, назовем базовым и обозначим значком Δ . В рассматриваемом n -угольнике есть два расsectionных на треугольники многоугольника с меньшим, чем n , числом сторон, располагающиеся слева и справа от базового треугольника (рис. 144).

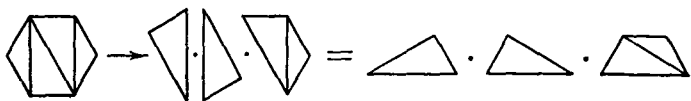


Рис. 144

Конечно, один или оба таких многоугольника могут отсутствовать. Будем считать, что в этом случае имеются вырожденные многоугольники и рисовать вместо них отрезок (рис. 145). Рис. 144 и 145 можно рассматривать как представление n -угольника в виде символического произведения базового треугольника на многоугольники, располагающиеся слева и справа от него.

Введем рисуночный ряд, состоящий из всех n -угольников, рассеченных всеми возможными способами (рис. 146). Во второй строке рис. 146 воспроизведен один из способов

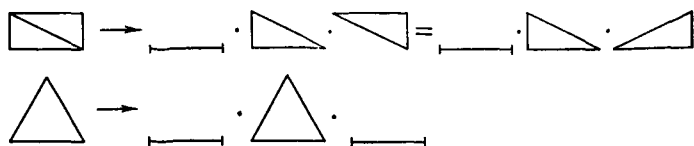


Рис. 145

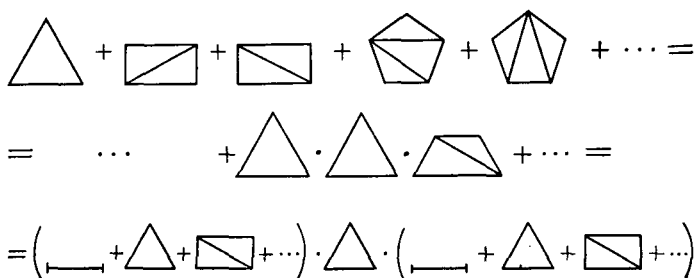


Рис. 146

рассечения шестиугольника на треугольники, показанный на рис. 144. Так как каждая фигура первой строки рис. 146 может быть представлена в виде символического произведения базового треугольника на многоугольники с меньшим числом сторон, то фигуры первой строки рис. 146 находятся во взаимно-однозначном соответствии с разложением произведения трех сомножителей третьей строки рис. 146. Средний сомножитель этого произведения — базовый треугольник, два других — рисуночные ряды верхней строки рис. 146 плюс отрезок.

На рис. 147 показано соответствие фигур степеням переменной x , с помощью которого от рисуночного ряда пере-

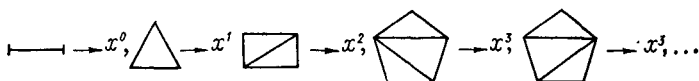


Рис. 147

ходим к перечисляющему. Обозначим через $D(x)$ перечисляющий ряд:

$$D(x) = D_3x + D_4x^2 + \dots + D_nx^{n-2} + \dots$$

Тогда соотношение между первой и третьей строками рис. 146 аналитически запишется в виде

$$D(x) = [1 + D(x)]x[1 + D(x)] = x[1 + D(x)]^2. \quad (2.9)$$

Выбирая тот корень квадратного относительно $D(x)$ уравнения (2.9), который конечен при $x = 0$, получаем формулу

$$D(x) = \frac{1 - 2x - \sqrt{1 - 4x}}{2x}, \quad (2.10)$$

правая часть которой есть производящая функция. Учитывая разложение

$$\sqrt{1 - 4x} = 1 - 2x - \dots - \frac{2^n 1 \cdot 3 \cdot 5 \dots (2n - 3)}{n!} x^n - \dots,$$

из формулы (2.10) получаем формулу для искомого комбинаторного числа

$$D_n = \frac{2 \cdot 6 \cdot 10 \cdot \dots \cdot (4n - 10)}{(n - 1)!}.$$

§ 3. ПЕРЕЧИСЛЕНИЕ КОРНЕВЫХ ДЕРЕВЬЕВ

Рассмотрим граф, являющийся деревом, одна вершина которого отмечена и названа *корнем*. Два корневых дерева называются *изоморфными*, если существует взаимно-однозначное отображение вершин одного дерева в вершины другого, переводящее корень дерева в корень и сохраняющее отношение смежности вершин. Пусть требуется найти T_n — число различных неизоморфных корневых деревьев с n вершинами. Легко видеть, что $T_1 = 1$, $T_2 = 1$, $T_3 = 2$, $T_4 = 4$.

На рис. 148 корень дерева отмечен стрелкой, направленной вниз. Основная идея решения этой задачи аналогич-

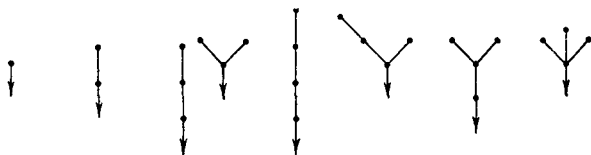


Рис. 148

на идее решения задачи о рассеении многоугольника. Мы будем строить дерево с n вершинами из более мелких. Для этого выделим «главные ветви» дерева, содержащие ребра, инцидентные корню, и будем рассматривать любое дерево как символическое произведение корня на фигуры, состоящие из изоморфных ветвей — одной, двух и т. д. (рис. 149).

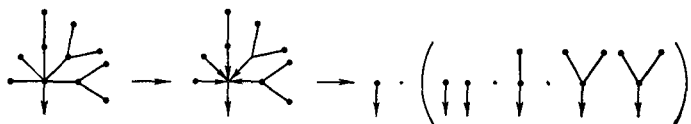


Рис. 149

После этого замечания переходим к рисуночному ряду: бесконечной сумме всех неизоморфных корневых деревьев (рис. 150). На рис. 150 бесконечное число строк. Первая строка — рисуночный ряд. Вторая строка — ряд, относящийся к всевозможным наборам корневых деревьев из одной вершины. Третья строка — ряд из фигур, составленных из корневых деревьев с двумя вершинами. Четвертая и пятая строка — ряды из фигур, соответствующих двум неизоморфным корневым деревьям из трех вершин.

Установим соответствие фигур степеням переменной x (рис. 151). Тогда рис. 150 приводит к соотношению Кэли

$$T(x) = T_1 x + T_2 x^2 + \dots + T_n x^n + \dots = \\ = x(1-x)^{-T_1} (1-x^2)^{-T_2} \dots (1-x^n)^{-T_n} \dots$$

Более компактно это соотношение можно записать в виде

$$T(x) = x \prod_{n=1}^{\infty} (1-x^n)^{-T_n}. \quad (3.1)$$

Формула (3.1) отличается от формул (2.2) и (2.10) тем, что в производящую функцию вошли искомые числа T_n . Тем не менее, сравнивая коэффициенты при одинаковых степенях переменной в левой и правой частях формулы (3.1), можно последовательно вычислять T_n . Выпишем несколько первых членов перечисляющего ряда $T(x)$:

Глава XIII

ТЕОРЕМА ПОЙА

Решение большого количества задач перечисления графов опирается на изящную теорему, принадлежащую Пойа. Подробный разбор этой теоремы оправдан тем, что математические понятия, необходимые для формулировки и доказательства этой теоремы, имеют принципиальное значение для всех комбинаторных задач. Трудно дать определение комбинаторики; к этому разделу математики относят задачи, в которых речь идет о подсчете возможностей, о числе способов выбора элементов некоторого множества, о числе множеств данного типа и т. п. Основная сложность заключается в том, что часто приходится считать некоторые возможности или способы выбора не отличающимися друг от друга в некотором, иногда довольно трудно поддающемся формальному описанию смысле. Например, утверждение о том, что существует всего пять правильных многогранников, подразумевает, что их расположение в пространстве не имеет значения. В современной математике существует понятие, соответствующее этой ситуации — понятие отношения эквивалентности. Можно сказать, что в комбинаторных задачах подсчитывается число классов эквивалентности. Как правило, отношение эквивалентности порождается некоторой группой подстановок, и изучение подобных отношений лежит в основе теоремы Пойа. Кроме того, эта теорема дает решение комбинаторных задач с помощью перечня конфигураций, являющегося обобщением понятий производящей функции и перечисляющего ряда.

§ 1. ЦИКЛОВОЙ ИНДЕКС ГРУППЫ ПОДСТАНОВОК МНОЖЕСТВА

1. Группа подстановок множества. Рассмотрим конечное множество T , $|T| = N$.

Определение 1.1. Пусть $T = \{t_1, t_2, \dots, t_N\}$. Подстановкой g множества T называется взаимно-однозначное отображение T на себя:

$$g: T \rightarrow T.$$

Образ элемента $t \in T$ при отображении g будем обозначать $g(t)$.

Определение 1.2. Композицией $g_1 \circ g_2$ двух подстановок g_1 и g_2 называется такое отображение множества T в себя, для которого

$$g_1 \circ g_2(t) = g_1(g_2(t)).$$

Приведем без доказательства следующую лемму.

Лемма 1.1. Композиция $g_1 \circ g_2$ подстановок g_1, g_2 есть подстановка множества T . Композиция есть ассоциативная операция:

$$g_1 \circ (g_2 \circ g_3) = (g_1 \circ g_2) \circ g_3.$$

В силу свойства взаимной однозначности подстановок как отображений для каждой подстановки g существует обратная подстановка g^{-1} :

$$g^{-1}(t_j) = t_i, \text{ если } t_j = g(t_i).$$

Подстановки g и g^{-1} удовлетворяют равенствам

$$g \circ g^{-1}(t) = g^{-1} \circ g(t) = t, \quad \forall t \in T.$$

Обозначим через e тождественную подстановку множества T :

$$e(t) = t, \quad \forall t \in T.$$

Тогда можно записать, что $g^{-1} \circ g = g \circ g^{-1} = e$. Из проведенных рассуждений и леммы 1 вытекает, что совокупность всех подстановок множества T образует группу.

Определение 1.3. Группа всех подстановок множества T , $|T| = N$ называется симметрической группой степени N и обозначается S_N .

Группа S_N содержит $N!$ подстановок, где N — число элементов множества T . Все рассматриваемые в дальнейшем группы G подстановок множества T будут являться, следовательно, подгруппами группы S_N .

2. Эквивалентность элементов множества T относительно подстановки. Фиксируем некоторую группу G .

Определение 1.4. Пусть g подстановка множества T , $g \in G$. Элемент $t_i \in T$ будем считать эквивалентным элементу $t_j \in T$ относительно подстановки g , если существует целое m такое, что $t_i = g^m(t_j)$, где $g^0 = e$, $g^m = \underbrace{g \circ g \circ \dots \circ g}_m$

при m -натуральном, $g^n = (g^{-1})^{|n|}$ при n целом, отрицательном.

Если t_i эквивалентен t_j , то пишут $t_i \stackrel{g}{\sim} t_j$.

Лемма 1.2. Выполняются следующие свойства отношения $t_i \stackrel{g}{\sim} t_j$:

- 1) $t \sim t, \forall t \in T$ (рефлексивность),
- 2) если $t_i \stackrel{g}{\sim} t_j$, то $t_j \stackrel{g}{\sim} t_i$ (симметричность),
- 3) из $t_i \stackrel{g}{\sim} t_j, t_j \stackrel{g}{\sim} t_k$ следует, что $t_i \stackrel{g}{\sim} t_k$ (транзитивность).

Доказательство. Свойство 1) вытекает из равенства $t = g^0(t)$. Свойство 2) справедливо ввиду того, что если $t_i = g^m(t_j)$, то $t_j = g^{-m}(t_i)$. Наконец, если справедливы равенства $t_i = g^m(t_j), t_j = g^n(t_k)$, то $t_i = g^{m+n}(t_k)$. ■

Определение 1.5. Обозначим через $\tau_g(t)$ класс элементов эквивалентных элементу t относительно подстановки g . Класс эквивалентности $\tau_g(t)$ называется циклом.

Основанием к этому определению служит следующая лемма.

Лемма 1.3. Класс эквивалентности $\tau_g(t)$ состоит из элементов $t, g(t), g^2(t), \dots, g^k(t)$, где k — натуральное число, зависящее, вообще говоря, от g и t , т. е. $k = k(g, t)$, такое, что справедливо равенство $g^{k+1}(t) = t$.

Таким образом, можно сказать, что элементы множества $\tau_g(t)$ циклически переставляются подстановкой g .

Доказательство леммы 1.3. Пусть $g \neq e$, ибо в случае $g = e$ доказательство тривиально ($k = 0$). Рассмотрим последовательность

$$t, g(t), g^2(t), \dots, g^j(t), \dots \quad (1.1)$$

Так как множество T конечно, последовательность (1.1) не может состоять из различных элементов. Пусть отрезок этой последовательности от элемента t до $g^{(k)}(t)$ состоит

из несовпадающих элементов, а $g^{k+1}(t)$ равен одному из предшествующих

$$g^{k+1}(t) = g^p(t) \quad (0 \leq p \leq k).$$

Пусть $p \neq 0$, тогда имеем равенства

$$g^p(t) = g(g^{p-1}(t)) \quad \text{и} \quad g^p(t) = g(g^k(t)).$$

В силу взаимной однозначности отображения g отсюда следует равенство $g^k(t) = g^{p-1}(t)$. Это противоречит сделанному предположению. Следовательно, для каждого $t \in T$ существует такое натуральное число $k = k(g, t)$, что $g^{k+1}(t) = t$.

Ясно, что если q — натуральное число и $q > k$, то справедливо равенство $g^q(t) = g^j(t)$, где $j = q - \left[\frac{q}{k+1} \right] (k+1)$ — остаток от деления числа q на $k+1$; очевидно, $0 \leq j \leq k$.

Если q — целое отрицательное число, то, выбрав натуральное n такое, что $n(k+1) - |q| > 0$ получим

$$g^q(t) = g^{-|q|}(t) = g^{-|q|}(g^{n(k+1)}(t)) = g^{n(k+1)-|q|}(t). \quad \blacksquare$$

Пусть $C_1(g), C_2(g), \dots, C_n(g)$ — различные циклы множества T . Непосредственным следствием теоремы 1.1 гл. I является следующая теорема.

Теорема 1.1. Циклы $C_j(g)$ ($j = 1, 2, \dots, n$) образуют разбиение множества T :

$$T = \bigcup_{j=1}^n C_j(g), \quad C_p(g) \cap C_q(g) = \emptyset \quad (p \neq q).$$

3. Цикловой индекс групп подстановок. Пусть для данной подстановки g множество T разбивается на j_1 циклов длины 1 ($k=1$), j_2 циклов длины 2 ($k=2$) и т. д.

Определение 1.6. Типом подстановки g называется последовательность $(j_1, j_2, \dots, j_N)$, где N — число элементов множества T .

Из определения чисел j_i и теоремы 1.1 вытекает равенство

$$j_1 + 2j_2 + 3j_3 + \dots + Nj_N = N.$$

Рассмотрим некоторую группу G подстановок множества T . Каждой подстановке $g \in G$ сопоставим произведение

$$x_1^{j_1} x_2^{j_2} \dots x_N^{j_N},$$

где $(j_1, j_2, \dots, j_N)$ — тип g , $x_k \in R^1$ ($k = 1, 2, \dots, N$).

Определение 1.7. Цикловым индексом группы подстановок G называется многочлен от $N = |T|$ переменных

$$P_G(x_1, x_2, \dots, x_N) = \frac{1}{|G|} \sum_{g \in G} x_1^{j_1(g)} x_2^{j_2(g)} \dots x_N^{j_N(g)},$$

где $|G|$ — число элементов группы G .

Например, если G — тривиальная подгруппа группы S_N , т. е. подгруппа, содержащая лишь один элемент — тождественную подстановку e , то справедливо равенство $P_G(x_1, x_2, \dots, x_N) = x_1^N$.

Пример 1. Пусть дано множество $T = \{t_1, t_2, t_3\}$. Требуется найти цикловой индекс группы S_3 всех подстановок множества из трех элементов.

Каждому из 6 элементов группы G сопоставим граф, вершинами которого являются элементы множества T , а дуги соответствуют отображению g (табл. 13). Из таблицы заключаем, что

$$P_{S_3}(x_1, x_2, x_3) = \frac{1}{6} (x_1^3 + 3x_1 x_2 + 2x_3).$$

ТАБЛИЦА 13

Элемент $G = S_3$	Граф	Слагаемое в P_{S_3}
e		x_1^3
g_1		$x_1^1 x_2^1$
g_2		$x_1^1 x_2^1$
g_3		$x_1^1 x_2^1$
g_4		x_3^1
g_5		x_3^1

Пример 2. Пусть T — множество вершин тетраэдра (рис. 152), а G — группа подстановок T , которые могут быть осуществлены вращениями тетраэдра как целого. Подсчитать цикловой индекс группы G .

Таких подстановок $4 \times 3 = 12$, тогда как всех подстановок четырех элементов $4! = 24$. Подстановки $g \in G$ можно разбить на классы:

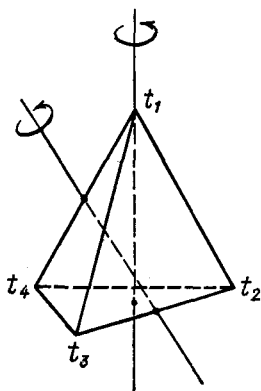


Рис. 152

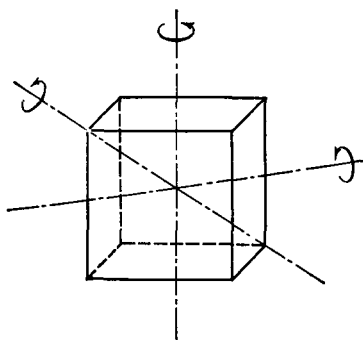


Рис. 153

а) тождественная подстановка,

б) 8 поворотов на $\pm 120^\circ$ вокруг четырех осей, проходящих через вершины T и центры противоположных граней,

с) три поворота на 180° вокруг трех осей, проходящих через центры противоположных ребер.

Геометрически очевидно, что для подстановок из класса а) имеется четыре цикла по одному элементу; для подстановок из класса б) один цикл длины один и один цикл длины три; для подстановок из класса с) — два цикла длины два. Поэтому справедлива формула

$$P_G(x_1, x_2, x_3, x_4) = \frac{1}{12} (x_1^4 + 8x_1x_3 + 3x_2^2).$$

Пример 3. Пусть T — множество вершин куба, G — группа подстановок T , осуществимых с помощью вращений куба. Найти цикловой индекс группы G .

Всего таких подстановок $6 \cdot 4 = 24$. Их можно разбить на классы (рис. 153):

а) тождественная подстановка,

б) шесть поворотов на $\pm 90^\circ$ вокруг осей, проходящих через центры противоположных граней,

с) три поворота на 180° вокруг осей, проходящих через центр противоположных граней,

д) шесть поворотов на 180° вокруг осей, проходящих через центры противоположных ребер,

е) восемь поворотов на $\pm 120^\circ$ вокруг осей, проходящих через противоположные вершины куба.

Вклад подстановки класса а) в циклический индекс равен x_1^8 , подстановка класса б) дает вклад x_4^2 , из класса с) — вклад x_2^4 , из класса д) — вклад x_4^2 , из класса е) — вклад $x_1^2 x_3^2$. Отсюда получим следующее выражение для циклового индекса группы G :

$$P_G(x_1, x_2, \dots, x_8) = \frac{1}{24} (x_1^8 + 9x_2^4 + 6x_4^2 + 8x_1^2 x_3^2).$$

Пример 4. Пусть T — множество граней куба, G — подстановки T , осуществимые с помощью вращений. Найти цикловой индекс группы G .

Тождественная подстановка дает в цикловой индекс вклад x_1^6 , подстановка класса б) — вклад $x_1^2 x_4$, подстановка из с) — вклад $x_1^2 x_2^2$, из д) — вклад x_2^3 , из е) — вклад x_3^2 . Следовательно, справедливо равенство

$$P_G(x_1, x_2, \dots, x_6) = \frac{1}{24} (x_1^6 + 6x_1^2 x_4 + 3x_1^2 x_2 + 6x_2^3 + 8x_3^2).$$

§ 2. ТРАНЗИТИВНЫЕ МНОЖЕСТВА ГРУППЫ ПОДСТАНОВОК

1. Эквивалентность элементов множества, порожденная группой подстановок. В § 1 рассмотрено понятие эквивалентности двух элементов множества T относительно подстановки g . Введем сейчас понятие эквивалентности относительно группы подстановок G .

Определение 2.1. Элемент $t_i \in T$ называется эквивалентным элементу $t_j \in T$ относительно группы G подстановок множества T , если существует такая подстановка $g \in G$, что $t_i = g(t_j)$.

Лемма 2.1. Введенное отношение эквивалентности рефлексивно, симметрично, транзитивно.

Доказательство. В силу равенства $t = e(t)$, $e \in G$ отношение рефлексивно. Из равенства $t_i = g(t_j)$ вытекает равенство $t_j = g^{-1}(t_i)$. Так как $g^{-1} \in G$, то отношение симметрично. Наконец, если справедливо равенство $t_i = g_1(t_j)$, $t_j = g_2(t_k)$, где $g_1, g_2 \in G$, то выполняется равенство $t_i = g_2 \circ g_1(t_k)$, $g_2 \circ g_1 \in G$, и отношение транзитивно. ■

Обозначим $\tau_G(t)$ — класс эквивалентности элемента t относительно группы G .

Определение 2.2. Классы эквивалентности $\tau_G(t)$ называются транзитивными множествами группы G .

Для группы G , состоящей из тождественной подстановки, транзитивные множества содержат по одному элементу. Для симметрической группы S_N , напротив, существует лишь одно транзитивное множество — все T .

2. Стационарная подгруппа элемента множества T . Пусть g — некоторая подстановка. Рассмотрим множество

$$T_g = \{t \in T \mid g(t) = t\}.$$

Множество T_g состоит из неподвижных точек отображения g . Для произвольного элемента $t \in T$ определим подмножество $G_t \subset G$, положив $G_t = \{g \in G \mid g(t) = t\}$.

Лемма 2.2. *Множество G_t есть подгруппа группы G , так называемая стационарная подгруппа элемента $t \in T$.*

Доказательство. Пусть подстановки g_1, g_2 принадлежат G_t , т. е. $g_1(t) = t, g_2(t) = t$. Тогда справедливы равенства $g_1 \circ g_2(t) = g_1(g_2(t)) = g_1(t) = t$, следовательно, $g_1 \circ g_2 \in G_t$. Для тождественной подстановки выполняется равенство $e(t) = t$, поэтому $e \in G_t$.

Пусть $g \in G_t$, тогда получим цепочку равенств $g^{-1}(t) = g^{-1}(g(t)) = g^{-1} \circ g(t) = e(t) = t$, следовательно, $g^{-1} \in G_t$. ■

Лемма 2.3. *Пусть $t_1 \in \tau_G(t)$; тогда справедливы равенства $|G_{t_1}| = |G_t| = \frac{|G|}{|\tau_G(t)|}$.*

Доказательство. В силу условия леммы существует элемент $h \in G$, такой, что $t_1 = h(t), t = h^{-1}(t_1)$. Пусть подстановка $g \in G_t$; рассмотрим подстановку $g' = h \circ g \circ h^{-1}$. Имеем

$$g'(t_1) = h \circ g \circ h^{-1}(t_1) = h \circ g(t) = h(t) = t_1,$$

следовательно, $g' \in G_{t_1}$. Для каждой подстановки $g' \in G_{t_1}$ существует лишь единственная подстановка $g \in G$ такая, что $g' = h \circ g \circ h^{-1}$, так как из равенства $g' = h \circ g \circ h^{-1} = h \circ g_1 \circ h^{-1}$, $g_1 \in G_t$ и взаимной однозначности отображения h следует $g = g_1 = h^{-1} \circ g' \circ h$.

Итак, установлено взаимно-однозначное соответствие между G_{t_1} и G_t , а следовательно, справедливо равенство $|G_{t_1}| = |G_t|$.

Для доказательства равенства $|G_t| = |G|/|\tau_G(t)|$ введем отношение эквивалентности подстановок группы G относительно подгруппы G_t . А именно, подстановку $h \in G$ будем считать эквивалентной подстановке $f \in G$, если существует подстановка $g \in G_t$ такая, что $h = f \circ g$. В силу

того что G_t — подгруппа, введенное отношение рефлексивно, симметрично и транзитивно. Классы эквивалентности $\tau_t(h)$ образуют разбиение группы G как множества: $G = \bigcup_{j=1}^p \tau_t(h_j)$. Из определения эквивалентности вытекает равенство $|\tau_t(h_j)| = |G_t|$.

Каждый класс $\tau_t(h_j)$ находится во взаимно-однозначном соответствии с элементами из $\tau_G(t)$. В самом деле, если $f_1 f_2 \in \tau_t(h_j)$, то справедливы равенства

$$f_1(t) = h_j \circ g_1(t) = h_j(t); \quad f_2(t) = h_j \circ g_2(t) = h_j(t),$$

так как $g_1, g_2 \in G_t$. Поэтому все подстановки класса $\tau_t(h_j)$ переводят элемент $t \in T$ в элемент $h_j(t) \in \tau_G(t)$. Обратно, если $t' \in \tau_G(t)$ и $t' = h_1(t)$, $t' = h_2(t)$, то подстановка $h_1^{-1} \circ h_2 \in G_t$, т. е. подстановки h_1 и h_2 — эквивалентны. Следовательно, число p различных классов эквивалентности $\tau_t(h_j)$ равно $|\tau_G(t)|$.

Высказанные утверждения позволяют установить равенства

$$|G| = \sum_{j=1}^p |\tau_t(h_j)| = p |G_t| = |\tau_G(t)| |G_t|. \blacksquare$$

3. Теорема о числе транзитивных множеств. Основным результатом настоящего параграфа является следующая теорема.

Теорема 2.1. Число транзитивных множеств $N(G)$ группы G равно

$$N(G) = \frac{1}{|G|} \sum_{g \in G} |T_g|.$$

Доказательство. Рассмотрим все пары (g, t) , где $g \in G$, $t \in T$, для которых $g(t) = t$. Число n таких пар можно подсчитать двумя способами. С одной стороны находим, что при фиксированной подстановке $g \in G$ число таких пар равно $|T_g|$. Суммируя по всем элементам группы G , получаем равенство

$$n = \sum_{g \in G} |T_g|. \quad (2.1)$$

С другой стороны, при фиксированном элементе $t \in T$ число пар равно $|G_t|$, следовательно,

$$n = \sum_{t \in T} |G_t|.$$

Преобразуем сумму $\sum_{t \in T} |G_t|$. Существует набор $t_j \in T$

($j = 1, 2, \dots, k$), $k = N(G)$, такой, что $\bigcup_{j=1}^k \tau_G(t_j) = T$, поэтому справедливы равенства

$$\begin{aligned} n = \sum_{t \in T} |G_t| &= \sum_{j=1}^k \sum_{t \in \tau_G(t_j)} |G_t| = \sum_{j=1}^k |\tau_G(t_j)| |G_{t_j}| = \\ &= |G| N(G). \end{aligned} \quad (2.2)$$

Последний знак равенства в (2.2) установлен в силу леммы 2.3. Из равенств (2.1) и (2.2) следует утверждение теоремы. ■

§ 3. КЛАССЫ ЭКВИВАЛЕНТНЫХ ОТОБРАЖЕНИЙ

1. Эквивалентность отображений конечных множеств. Рассмотрим отображения $f: T \rightarrow R$ множества T в некоторое конечное множество R . Так как T и R — конечные множества, то совокупность всех отображений $f: T \rightarrow R$ также содержит конечное число элементов. Это число равно $R^{|T|}$. Будем обозначать множество отображений T в R через R^T .

Группа G подстановок множества T порождает эквивалентность отображений в соответствии со следующим определением.

Определение 3.1. Отображение $f_i \in R^T$ эквивалентно отображению $f_j \in R^T$ относительно группы G , если существует такая подстановка $g \in G$, что справедливо равенство

$$f_i = f_j \circ g^{-1}. \quad (3.1)$$

Лемма 3.1. Отношение эквивалентности (3.1) рефлексивно, симметрично, транзитивно.

Доказательство леммы очевидным образом вытекает из групповых свойств подстановок и проводится аналогично доказательству леммы 2.1.

Отношение эквивалентности (3.1) порождает классы $\tau_G(f)$ эквивалентных отображений. Эти классы $\tau_G(f_j)$ ($j =$

$= 1, 2, \dots, p)$ образуют разбиение множества отображений R^T .

Пример 5. Пусть $T = \{t_1, t_2\}$, $G = S_2$, R — числовое множество, содержащее два элемента 0 и 1. Найти классы эквивалентных отображений R^T .

Существует четыре отображения множества T в множество R и два элемента группы S_2 (табл. 14, рис. 154). Очевидны равенства:

$$f_j(g_1^{-1}(t)) = f_j(e(t)) = f_j(t) \quad (j = 1, 2, 3, 4);$$

кроме того,

$$f_1(g_2^{-1}(t)) = f_1(t), \quad f_2(g_2^{-1}(t)) = f_2(t).$$

ТАБЛИЦА 14

R^T	f_1	f_2	f_3	f_4
t_1	0	1	0	1
t_2	0	1	1	0

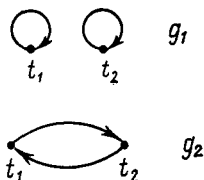


Рис. 154

Следовательно, имеем

$$\tau_G(f_1) = \{f_1\}, \quad \tau_G(f_2) = \{f_2\}.$$

Из равенств

$$f_3(g_2^{-1}(t)) = f_4(t), \quad f_4(g_2^{-1}(t)) = f_3(t)$$

вытекает

$$\tau_G(f_3) = \tau_G(f_4) = \{f_3, f_4\}.$$

Таким образом, имеется три класса эквивалентных отображений.

Пример 6. Пусть T — множество граней куба, G — группа подстановок граней куба, осуществимых с помощью вращений. Множество R пусть состоит из двух элементов: $R = \{c, k\}$ (можно считать, что эти два элемента есть синяя и красная краски). Найти классы эквивалентных отображений R^T .

Отображение $f: T \rightarrow R$ можно отождествить с раскраской граней куба в красный и синий цвета. Всех способов раскраски фиксированного в пространстве куба 2^6 . Если два раскрашенных различно куба расположены параллельно в пространстве, то может случиться, что один из них можно повернуть так, что кубы перестанут казаться раскрашенными различно. Очевидно, эти две раскраски эквивалентны относительно группы подстановок граней куба, осуществимых с помощью вращений.

Классы эквивалентности можно пересчитать и описать следующим образом: а) все грани красные — одна раскраска в классе,

б) пять граней красных, одна синяя — 6 раскрасок в классе, в) две противоположные грани синие, остальные — красные — три раскраски в классе, г) две смежные грани синие, четыре — красные — 12 раскрасок, д) три грани, прилегающие к одной вершине, синие — 8 раскрасок, е) две противоположных грани и одна оставшаяся синие — 12 раскрасок.

Классы ж), з), и), к) получаются из классов г), в), б), а) заменой краски «синяя» на «красную».

Здесь перечислено 10 классов раскраски, содержащих в сумме

$$1 + 6 + 3 + 12 + 8 + 12 + 12 + 3 + 6 + 1 = 2^3$$

раскрасок. Следовательно, перечислены все возможные классы.

2. Группа подстановок множества R^T , порожденная группой G . Полезно обсудить другой подход к рассмотрению классов эквивалентных отображений, по форме более близкий подходу § 2. Для этого рассмотрим множество $\tilde{G}$ отображений множества R^T в множество R^T , которое порождено группой G в следующем смысле. По каждой подстановке $g \in G$ построим отображение $\pi_g \in \tilde{G}$ по формуле

$$\pi_g(f) = f \circ g^{-1}, \quad (3.2)$$

где $f \in R^T$. Очевидно, что $\pi_g(f) \in R^T$.

Лемма 3.2. *Отображение $\pi_g \in \tilde{G}$ есть подстановка множества R^T .*

Доказательство. Пусть выполнено равенство $\pi_g(f_1) = \pi_g(f_2)$, эквивалентное в силу (3.2) равенству

$$f_1 \circ g^{-1} = f_2 \circ g^{-1}. \quad (3.3)$$

В силу свойства взаимной однозначности подстановки g равенство (3.3) означает, что значения отображений $f_1, f_2 \in R^T$ совпадают на всех элементах $t \in T$. ■

Композиция двух отображений π_{g_1} и π_{g_2} определяется как обычно:

$$\pi_{g_1} \circ \pi_{g_2}(f) = \pi_{g_1}(\pi_{g_2}(f)). \quad (3.4)$$

Из равенств (3.2) и (3.4) вытекает равенство

$$\pi_{g_1} \circ \pi_{g_2}(f) = f \circ g_2^{-1} \circ g_1^{-1}. \quad (3.5)$$

Лемма 3.3. *Множество $\tilde{G}$ есть группа относительно операции композиции (3.4).*

Доказательство. Операция композиции (3.4) —

ассоциативна. Единицей группы $\tilde{G}$ является подстановка π_e , $e \in G$, так как

$$\pi_e(f) = f \circ e^{-1} = f \circ e = f, \quad \forall f \in R^T.$$

Элементом $(\pi_g)^{-1}$, обратным элементу π_g , является подстановка $\pi_{g^{-1}}$, так как

$$\pi_g \circ \pi_{g^{-1}}(f) = f \circ g \circ g^{-1} = f \circ e = f$$

в силу равенства (3.5). ■

Лемма 3.4. Соответствие $\pi: G \rightarrow \tilde{G}$, заданное равенством (3.2), есть гомоморфизм групп G и $\tilde{G}$.

Доказательство. Из формулы 3.5 вытекают равенства

$$\pi_{g_1 \circ g_2}(f) = f \circ (g_1 \circ g_2)^{-1} = f \circ g_2^{-1} \circ g_1^{-1} = \pi_{g_1} \circ \pi_{g_2}(f). \quad \blacksquare$$

У группы $\tilde{G}$ подстановок множества отображений R^T есть транзитивные классы. Из формулы (3.2) и определения транзитивных классов, данного в § 2, вытекает следующая лемма.

Лемма 3.5. Отображения f_i и $f_j \in R^T$ эквивалентны тогда и только тогда, когда они принадлежат одному и тому же транзитивному классу группы $\tilde{G}$.

Определение 3.2. Множество элементов $g \in G$ таких, что $\pi_g = \pi_e \in \tilde{G}$, называется ядром отображения π и обозначается $\ker \pi$.

Пример. Пусть множество R состоит из одного элемента $R = \{r\}$. Тогда множество R^T для любого T состоит также из одного элемента. Естественно, что и группа всех подстановок множества R^T содержит лишь единичный элемент. Ясно, что какова бы ни была группа G подстановок множества T , группа $\tilde{G} = \{e\}$. Поэтому ядром отображения π будет являться вся группа G .

Лемма 3.6. Пусть $\{g\}$ — множество подстановок из G , переходящих в π_{g_0} при отображении π . Справедливы равенства

$$|\{g\}_{\pi_{g_0}}| = |\ker \pi|, \quad |G| = |\tilde{G}| |\ker \pi|.$$

Лемма есть непосредственное следствие леммы 3.4.

§ 4. ВЕС И ПЕРЕЧЕНЬ ФИГУР И КОНФИГУРАЦИЙ

Имея в виду дальнейшее приложение теории в комбинаторике и перечислении графов, введем следующие термины.

Определение 4.1. *Фигурами называются элементы множества R . Конфигурациями называются отображения $f \in R^T$.*

Существенным моментом теории перечисления Пойа является приписывание фигурам и конфигурациям весов. В простейшем случае весами могут служить вещественные числа. Отметим, что в этом случае веса можно складывать и перемножать. Введем общее понятие веса.

Определение 4.2. *Пусть на множестве W введены две ассоциативные и коммутативные операции: сложение «+» и умножение «·». Пусть, кроме того, определено произведение элементов $w \in W$ на рациональные числа. Если выполнены следующие условия:*

- 1) *относительно операции сложения множество W является группой,*
- 2) *справедливы равенства*

$$(p + q)w = pw + qw, \quad p(qw) = (pq)w,$$

$$(pw_1) \cdot (qw_2) = (pq)w_1 \cdot w_2,$$

где p, q — рациональные числа $w_1, w_2 \in W$, то говорят, что множество W есть коммутативная алгебра над полем рациональных чисел.

Например, множество полиномов от n переменных $x_1, x_2, \dots, x_n$ с рациональными коэффициентами и введенным стандартным образом операциями сложения и умножения является коммутативной алгеброй над полем рациональных чисел.

Определение 4.3. *Пусть имеется произвольное отображение $\alpha: R \rightarrow W$. Весом $w_\alpha(r)$ фигуры $r \in R$ называется образ фигуры r при отображении α .*

З а м е ч а н и е. В дальнейшем никогда не придется рассматривать одновременно несколько отображений R в W , поэтому значок α у веса $w_\alpha(r)$ мы будем опускать.

Определение 4.4. *Перечнем $L(R)$ множества фигур R называется сумма весов фигур $r \in R$:*

$$L(R) = \sum_{r \in R} w(r).$$

Если в качестве множества W взять множество вещественных чисел и придать вес, равный 1 всем фигурам, то перечень фигур — это просто число элементов множества R .

Определение 4.5. Пусть определены веса $w(r)$ фигур множества R . Весом конфигурации $f \in R^T$ называется произведение

$$\tilde{w}(f) = \prod_{t \in T} w(f(t)), \quad \tilde{w}(f) \in W.$$

Перечнем конфигураций называется перечень множества R^T :

$$L(R^T) = \sum_{f \in R^T} \tilde{w}(f).$$

В дальнейшем будет показано, что перечень конфигураций есть обобщение понятия перечисляющего ряда.

Лемма 4.1. Если конфигурации $f_1, f_2 \in R^T$ эквивалентны относительно группы G , то их веса равны.

Доказательство. По условию леммы выполнено соотношение $f_1(t) = f_2 \circ g^{-1}(t)$, $g \in G$. Следовательно, справедлива следующая цепочка равенств:

$$\begin{aligned} \tilde{w}(f_1) &= \prod_{t \in T} w(f_1(t)) = \prod_{t \in T} w(f_2(g^{-1}(t))) = \\ &= \prod_{t \in T} w(f_2(t)) = \tilde{w}(f_2), \end{aligned}$$

поскольку в силу взаимной однозначности подстановки g^{-1} второе и третье произведения имеют одни и те же сомножители, расставленные, может быть, в другом порядке. Порядок несуществен в силу коммутативности операции умножения.

Лемма 4.1 позволяет однозначным образом ввести вес класса эквивалентных конфигураций.

Определение 4.6. Весом класса эквивалентности $\bar{w}[\tau_G(f)] \in W$ называется значение веса любой входящей в него конфигурации. Перечень классов эквивалентности — это сумма весов классов эквивалентности.

Пример 7. Рассмотрим раскрашивание граней куба из примера 6. Пусть множеством весов является множество полиномов от двух переменных s и k с рациональными коэффициентами. Найти перечень классов эквивалентных раскрасок.

Положим вес элемента «синий» из R равным c , элемента «красный», равным k . Тогда классы эквивалентных раскрасок получают веса

$$k^6, k^5c, k^4c^2, k^4c^2, k^3c^3, k^3c^3, k^2c^4, k^2c^4, kc^5, c^6$$

соответственно. Мы видим, что разные классы могут иметь одинаковые веса. Перечень L классов эквивалентных раскрасок равен

$$L = k^6 + k^5c + 2k^4c^2 + 2k^3c^3 + 2k^2c^4 + kc^5 + c^6.$$

§ 5. ТЕОРЕМА ПОЙА

1. Соотношение между перечнем фигур и перечнем конфигураций. Справедливо следующее полезное утверждение.

Лемма 5.1. *Перечень конфигураций $f \in R^T$ связан с перечнем фигур $r \in R$ соотношением*

$$L(R^T) = (L(R))^{|\mathcal{T}|}. \quad (5.1)$$

Доказательство. В силу определения перечня фигур правая часть формулы (5.1) может быть записана в виде

$$(L(R))^{|\mathcal{T}|} = \left(\sum_{r \in R} w(r) \right)^{|\mathcal{T}|} = (w(r_1) + w(r_2) + \dots + w(r_p))$$

$$(w(r_1) + w(r_2) + \dots + w(r_p)) \dots (w(r_1) + w(r_2) + \dots + w(r_p)), \quad (5.2)$$

где $p = |R|$ и число сомножителей в произведении равно $|\mathcal{T}|$. Если раскрыть скобки в правой части формулы (5.2), то получим сумму

$$(L(R))^{|\mathcal{T}|} = \sum_{i_1, i_2, \dots, i_p} w(r_{i_1}) w(r_{i_2}) \dots w(r_{i_p}), \quad (5.3)$$

содержащую $|R|^{|\mathcal{T}|}$ членов. Общий член этой суммы имеет вид

$$w(r_{i_1}) w(r_{i_2}) \dots w(r_{i_p}) \quad (5.4)$$

и получается выбором i_1 -го слагаемого в первом сомножителе, i_2 -го слагаемого — во втором, i_p -го — в последнем. Конечно, часть сомножителей произведения (5.4) может совпадать между собой.

Установим взаимно-однозначное соответствие между элементами множества $\mathcal{T}$ и сомножителями произведения в формуле (5.2). Пусть элементу t_1 соответствует первый сом-

ножитель, элементу t_2 — второй, элементу t_N , где $N = |T|$ — последний сомножитель:

$$\overbrace{(\omega(r_1) + \dots + \omega(r_p))}^{t_1} \overbrace{(\omega(r_1) + \dots + \omega(r_p))}^{t_2} \dots \overbrace{(\omega(r_1) + \dots + \omega(r_p))}^{t_N}.$$

Тогда каждое слагаемое формулы (5.3) будет отвечать некоторому отображению $f: T \rightarrow R$ такому, что справедливы равенства

$$f(t_1) = r_{i_1}, f(t_2) = r_{i_2}, \dots, f(t_N) = r_{i_p}.$$

Кроме того, это слагаемое будет равно $\tilde{\omega}(f)$ — весу конфигурации f . Так как число конфигураций $f \in R^T$, равное $|R|^{|T|}$, совпадает с числом слагаемых суммы (5.3), то правая часть формулы равна перечню конфигураций

$$L(R^T) = \sum_{f \in R^T} \tilde{\omega}(f). \blacksquare$$

2. Перечень кусочно постоянных конфигураций. Пусть множество T разбито на k непересекающихся подмножеств

$$T = \bigcup_{j=1}^k T_j, \quad T_i \cap T_j = \emptyset \quad (i \neq j).$$

Рассмотрим подмножество $\Phi_k \subseteq R^T$ конфигураций f постоянных на каждом подмножестве T_j :

$$f(t) = r_j, \quad \forall t \in T_j \quad (j = 1, 2, \dots, k).$$

Лемма 5.2. Для перечня подмножества Φ_k справедливо равенство

$$L(\Phi_k) = \prod_{i=1}^k \sum_{r \in R} (\omega(r))^{|T_j|}. \quad (5.5)$$

Д о к а з а т е л ь с т в о. Рассмотрим разложение произведения правой части равенства (5.5):

$$\prod_{j=1}^k \sum_{r \in R} (\omega(r))^{|T_j|} = ((\omega(r_1))^{|T_1|} + (\omega(r_2))^{|T_1|} + \dots + (\omega(r_p))^{|T_1|}) \times ((\omega(r_1))^{|T_2|} + \dots + (\omega(r_2))^{|T_2|}) \dots ((\omega(r_1))^{|T_k|} + \dots + (\omega(r_p))^{|T_k|}). \quad (5.6)$$

Общий член суммы, получаемый при раскрытии скобок в (5.6), имеет вид

$$(\omega(r_{i_1}))^{|T_1|} (\omega(r_{i_2}))^{|T_2|} \dots (\omega(r_{i_p}))^{|T_k|} \quad (5.7)$$

и соответствует одной и только одной конфигурации, для которой выполняются равенства:

$$f(t) = r_{i_1}, \quad \forall t \in T_1,$$

$$f(t) = r_{i_2}, \quad \forall t \in T_2,$$

$$\dots \dots \dots$$

$$f(t) = r_{i_p}, \quad \forall t \in T_k$$

(среди элементов r_{i_j} могут встречаться одинаковые).

Кроме того, произведение (5.7) равно весу конфигурации. Следовательно, вся сумма, возникающая при раскрытии скобок в (5.6), равна перечню $L(\Phi_k)$. ■

Пример 8. Пусть имеется n шаров и три ящика A_1, A_2, A_3 . Сколькими способами P_n можно разместить шары по ящикам так, чтобы в первых двух ящиках было по одинаковому количеству шаров?

В качестве множества T выберем множество, состоящее из трех ящиков $\{A_1, A_2, A_3\}$. Множество R — совокупность неотрицательных чисел, не превосходящих n : $R = \{0, 1, 2, 3, \dots, n\}$. Конфигурация f — способ разместить шары по ящикам. В данном случае должны выполняться равенства

$$f(A_1) = f(A_2); \quad f(A_1) + f(A_2) + f(A_3) = n.$$

Положим $T_1 = \{A_1, A_2\}$, $T_2 = \{A_3\}$. Придадим вес фигурам — элементам из R — по правилу $\omega(j) = x^j$, $j \in R$. Тогда вес конфигураций, удовлетворяющих условиям задачи, равен x^n :

$$\begin{aligned} \omega(f) &= \omega(f(A_1)) \omega(f(A_2)) \omega(f(A_3)) = x^{f(A_1)} x^{f(A_2)} x^{f(A_3)} = \\ &= x^{f(A_1)+f(A_2)+f(A_3)} = x^n. \end{aligned}$$

По лемме 5.2 перечень L конфигураций, постоянных на T_1 и T_2 (но не обязательно имеющих вес x^n),

$$L = (1 + x^2 + x^4 + \dots + x^{2n}) (1 + x + x^2 + \dots + x^n). \quad (5.8)$$

Искомое число способов — коэффициент при x^n в разложении правой части (5.8) по степеням x . Легко сообразить, что тот же коэффициент остается при x^n , если рассмотреть вместо (5.8) выражение

$$L_1 = (1 + x^2 + x^4 + \dots + x^{2n} + x^{2n+2} + \dots) (1 + x + x^2 + \dots + x^n + x^{n+1} + \dots) = (1 - x^2)^{-1} (1 - x)^{-1}.$$

Функции L и L_1 являются производящими функциями комбинаторных чисел P_n . Имеем разложение

$$L_1 = \frac{1}{4} (1 + x)^{-1} + \frac{1}{2} (1 - x)^{-2} + \frac{1}{4} (1 - x)^{-1},$$

откуда находим следующую формулу:

$$P_n = \frac{1}{2} (n + 1) + \frac{1}{4} (1 + (-1)^n).$$

3. Теорема Пойа. Пусть на множестве T действует группа подстановок G . Обозначим через F класс эквивалентных относительно G конфигураций, через $\omega(F)$ — вес класса F . Для перечня классов эквивалентных конфигураций $L = \sum_F \bar{\omega}(F)$ справедливо равенство

$$L = P_G \left(\sum_{r \in R} \omega(r), \sum_{r \in R} (\omega(r))^2, \dots, \sum_{r \in R} (\omega(r))^N \right), \quad (5.9)$$

где N — число элементов в T , $P_G(x_1, x_2, \dots, x_N)$ — цикловой индекс группы G .

Доказательство. Фиксируем некоторый вес $\omega \in W$. Пусть D_ω — совокупность конфигураций, имеющих вес ω :

$$D_\omega = \{f \in R^T \mid \tilde{\omega}(f) = \omega\}.$$

По лемме 4.1 из $f \in D_\omega$ вытекает, что $\varphi \in D_\omega$, если $\varphi \in \tau_G(f)$.

Следовательно, группа $\tilde{G}$ подстановок множества R^T , введенная в § 3 этой главы, переводит множество D_ω в себя. Поэтому для подсчета n_{D_ω} — числа классов эквивалентных конфигураций, обладающих весом ω , можно применить теорему 2.1 и получить равенство

$$n_{D_w} = \frac{1}{|\tilde{G}|} \sum_{\pi_g \in \tilde{G}} \Psi_w(\pi_g), \quad (5.10)$$

где $\Psi_w(\pi_g)$ — число неподвижных точек подстановки π_g , т. е. конфигураций $f \in D_w$ таких, что $\pi_g(f) = f$. В силу леммы 3.6 формулу (5.10) можно преобразовать к виду

$$n_{D_w} = \frac{1}{|G|} \sum_{g \in G} \Psi_w(g), \quad (5.11)$$

где $\Psi_w(g)$ — число конфигураций $f \in D_w$ таких, что $f \circ g^{-1} = f$. Умножая обе части равенства (5.11) на w и суммируя по всем весам $w \in W$, получим равенство

$$\sum_{w \in W} w n_{D_w} = \frac{1}{|G|} \sum_{w \in W} w \sum_{g \in G} \Psi_w(g). \quad (5.12)$$

Левая часть есть перечень классов эквивалентности

$$\sum_{w \in W} w n_{D_w} = \sum_F \bar{w}(F) = L. \quad (5.13)$$

Изменим в правой части равенства (5.12) порядок суммирования. Получим следующее выражение для перечня классов эквивалентности:

$$L = \frac{1}{|G|} \sum_{g \in G} \sum_{w \in W} w \Psi_w(g). \quad (5.14)$$

Внутренняя сумма в правой части равенства (5.14) есть перечень конфигураций f таких, что $f \circ g^{-1} = f$. Так как g^{-1} — подстановка множества T , то T разбивается на циклы T_j ($j = 1, 2, \dots, k$) подстановки g^{-1} (теорема 1.1). Условие $f = f \circ g^{-1}$ означает, что

$$f(t) = f(g^{-1}(t)) = f(g^{-2}(t)) = \dots,$$

поэтому конфигурация f постоянна на каждом цикле T_j ($j = 1, 2, \dots, k$).

По лемме 5.2 справедливо равенство

$$\sum_{w \in W} w \Psi_w(g) = \prod_{j=1}^k \sum_{r \in R} (w(r))^{|T_j|}.$$

Пусть $(j_1, j_2, \dots, j_N)$ тип подстановки g^{-1} . Следовательно, среди чисел $|T_j|$ число 1 встречается j_1 раз, число 2 встречается j_2 раз и т. д. Отсюда вытекает, что

$$\sum_{w \in W} w \Psi_w(g) = \left\{ \sum_{r \in R} w(r) \right\}^{i_1} \times \\ \times \left\{ \sum_{r \in R} (w(r))^2 \right\}^{i_2} \dots \left\{ \sum_{r \in R} (w(r))^N \right\}^{i_N}. \quad (5.15)$$

Выражение (5.15) может быть получено, если подставить выражения

$$x_1 = \sum_{r \in R} w(r), \quad x_2 = \sum_{r \in R} (w(r))^2, \quad \dots, \quad x_N = \sum_{r \in R} (w(r))^N$$

в произведение $x_1^{i_1} x_2^{i_2} \dots x_N^{i_N}$, являющееся слагаемым в цикловом индексе $P_G(x_1, x_2, \dots, x_N)$ группы G , отвечающим подстановке g^{-1} . Из этого замечания и равенств (5.14), (5.15) вытекает утверждение теоремы. ■

Пример 9. Вновь рассмотрим задачу об окрашивании граней куба в две краски: синюю и красную. Требуется, пользуясь теоремой Пойа, найти перечень классов эквивалентных раскрасок.

Цикловой индекс группы подстановок граней, осуществимых с помощью вращений, найден в примере 4:

$$P_G(x_1, x_2, \dots, x_6) = \frac{1}{24} (x_1^6 + 6x_1^2 x_4 + 3x_1^2 x_2^2 + 6x_2^3 + 8x_3^2).$$

Пусть, как и в примере 6, множество весов есть множество полиномов с рациональными коэффициентами от двух переменных c и k . Положим вес элемента «синий» равным c , «красный» — k . В силу равенства (5.9) перечень классов эквивалентных раскрасок есть

$$\frac{1}{24} [(c+k)^6 + 6(c+k)^2(c^4+k^4) + 3(c+k)^2(c^2+k^2) + 6(c^2+k^2)^3 + 8(c^3+k^3)^2] = k^6 + k^5c + 2k^4c^2 + 2k^3c^3 + 2k^2c^4 + kc^5 + c^6,$$

что находится в полном согласии с результатами примеров 6 и 7.

§ 6. ЦИКЛОВЫЕ ИНДЕКСЫ ДВУХ ГРУПП, ЧАСТО ВСТРЕЧАЮЩИХСЯ В ТЕОРИИ ПЕРЕЧИСЛЕНИЙ ГРАФОВ

1. Цикловой индекс симметрической группы S_p . Пусть подстановка $g \in S_p$ имеет тип $(j_1, j_2, \dots, j_p)$. Определим число подстановок, имеющих тот же тип. Пусть сначала тип

g есть $(0, 0, \dots, 0, 1)$. Из всех $p!$ перестановок элементов множества T подстановку g порождают ровно p перестановок вида $t_{i_1}, t_{i_2}, \dots, t_{i_p}$, где $t_{i_{k+1}} = g(t_{i_k})$ ($k = 0, 1, \dots, p-1$), $t_{i_1} = g(t_{i_p})$. Таким образом, имеется $(p-1)!$ подстановок рассмотренного типа. Пусть теперь тип g есть $(j_1, j_2, \dots, j_p)$. Очевидно существует $p! [j_1! j_2! \dots j_p!]^{-1}$ различных способов объединить p элементов множества T в упорядоченные группы, причем j_1 групп содержит по одному элементу, j_2 групп — по два элемента и т. д. Эти упорядоченные группы отвечают циклам подстановок длины n ($n = 1, 2, \dots, p$) (сами циклы есть просто подмножества T). Так как ровно n перестановок из n элементов ($n = 1, 2, \dots, p$) соответствуют действию подстановки g на цикле длины n , то искомое число подстановок есть $p! [j_1! j_2! \dots j_p! 1^{j_1} 2^{j_2} \dots p^{j_p}]^{-1}$. Следовательно, справедлива формула

$$P_{S_p}(y_1, y_2, \dots, y_p) = \frac{1}{p!} \sum_{(j)} \frac{p!}{j_1! 2^{j_2} j_2! \dots p^{j_p} j_p!} y_1^{j_1} y_2^{j_2} \dots y_p^{j_p}, \quad (6.1)$$

где суммирование проводится по всем таким наборам целых неотрицательных чисел $(j_1, j_2, \dots, j_p)$, для которых $j_1 + 2j_2 + \dots + pj_p = p$.

2. Цикловой индекс группы-кранца. Пусть Q и R — конечные множества, G и H — группы подстановок множества Q и R соответственно. Рассмотрим множество $Q \times R$ всех пар (q, r) точек $q \in Q, r \in R$.

Определим подстановку γ множества $Q \times R$ следующим образом. Пусть $f: Q \rightarrow H$ — некоторое отображение на всю группу H , рассматриваемую просто как множество. Пусть g — некоторая подстановка из группы G . Тогда действие подстановки γ на пару (q, r) определяется формулой

$$\gamma[g, f]((q, r)) = (g(q), f(q)(r)). \quad (6.2)$$

Количество подстановок $\gamma[g, f]$, очевидно, равно $|G| |H|^{|Q|}$. Нетрудно проверить, что они образуют группу. Эта группа называется *группой-кранца* и обозначается GlH .

Теорема 6.1. Пусть $P_G(y_1, y_2, \dots)$ и $P_H(x_1, x_2, \dots)$ — цикловые индексы групп подстановок G и H . Справедливо равенство

$$P_{G(H)}(x_1 x_2, \dots) = P_G[P_H(x_1, x_2, x_3, \dots),$$

$$P_H(x_2, x_4, x_6, \dots) \dots], \quad (6.3)$$

где правая часть получена подстановкой $y_k = P_H(x_k, x_{2k}, x_{3k}, \dots)$ в цикловой индекс группы G .

Доказательство. Рассмотрим подстановку $\gamma[g, f] \in G[H]$. Пусть $g \in G$ имеет тип $(j_1, j_2, \dots, j_p)$, где $p = |Q|$. Определим тип подстановки γ . Фиксируем некоторый цикл $(q_1, q_2, \dots, q_k)$ подстановки g и рассмотрим множество T всех пар (q_i, r) , где $i = 1, 2, \dots, k$, а r пробегает все множество R . Очевидно, что это множество переходит в себя при преобразовании γ . Найдем, на какие циклы оно распадается под действием подстановки γ .

Пусть $h_j = f(q_j)$ ($j = 1, 2, \dots, k$). Оказывается, что эти циклы зависят только от произведения $h = h_k, h_{k-1} \dots h_1$. Положим, что подстановка h имеет тип $(b_1, b_2, b_3, \dots)$, и рассмотрим действие подстановки $\gamma[g, f]$ и ее степеней на элемент $[q_1, r]$. Последовательные применения γ дают следующий результат:

$$(q_1, r) \rightarrow (q_2, h_1(r)) \rightarrow (q_3, h_2 h_1(r)) \rightarrow \dots$$

$$\dots \rightarrow (q_k, h_{k-1} \dots h_1(r)) \rightarrow (q_1, h_k h_{k-1} \dots h_1(r)).$$

Следовательно, если элемент r принадлежит циклу подстановки h длины l , то элемент (q_1, r) порождает цикл длины kl . Таким образом, множество T распадается на b_1 циклов длины k , b_2 циклов длины $2k$, b_3 циклов длины $3k$ и т. д.

Для каждого множества T рассмотрим выражения $|H|^{-k} x_k^{b_1} x_{2k}^{b_2} \dots$. Если просуммировать эти произведения при фиксированных $q \in (q_1, q_2, \dots, q_k)$ по всем $f \in H^Q$, то получим $P_H(x_k, x_{2k}, x_{3k}, \dots)$, так как произведение $h = h_k h_{k-1} \dots h_1$ пробегает группу H k раз, когда h_i ($i = 1, 2, \dots, k$) пробегает группу H один раз.

Так как тип подстановки g есть $(j_1, j_2, \dots)$, то эта подстановка дает вклад

$$[P_H(x_1, x_2, x_3, \dots)]^{j_1} [P_H(x_2, x_4, x_6, \dots)]^{j_2} [P_H(x_3, x_6, x_9, \dots)]^{j_3} \dots$$

в цикловой индекс группы $G[H]$. После суммирования по всем $g \in G$ этих вкладов получаем утверждение теоремы. ■

Глава XIV

НЕКОТОРЫЕ МЕТОДЫ ПЕРЕЧИСЛЕНИЯ ГРАФОВ

Теория перечисления графов имеет многочисленные применения. Перечисление корневых деревьев, предпринятое Кэли (см. § 3 гл. XII), имело целью подсчет числа изомеров насыщенных углеводородов $C_n H_{2n+2}$. Структурную формулу таких веществ можно рассматривать как дерево, в котором вершины имеют степень 1 или 4. Результаты других задач перечисления находят приложения в химии, теоретической и статистической физике, биологии, теории планирования эксперимента и т. д.

Изучение методов перечисления графов полезно не только с целью получения результата — искомого комбинаторного числа, оно позволяет также глубже уяснить смысл основных понятий теории графов. К их числу следует отнести изоморфизм графов, группу автоморфизмов графа, корневые графы и графы с отмеченными ребрами. В первом параграфе дается новое определение графа, полнее раскрывающее комбинаторную природу этого понятия.

Заметим, что если бы удалось найти число всех плоских графов и всех 4-хроматических, то это бы решило проблему 4 красок. К сожалению, эти задачи перечисления до сих пор не решены.

§ 1. ПЕРЕЧИСЛЕНИЕ ГРАФОВ С p ВЕРШИНАМИ И k ДУГАМИ

1. Граф как конфигурация. Пусть T — некоторое множество, $|T| = p$. Рассмотрим полный граф без петель $K_p(T, T \times T \setminus E)$, где $E \subset T \times T$ — диагональ, $E = \{(t_i, t_i) \mid (i = 1, 2, \dots, p)\}$. Граф K_p имеет $p(p-1)$ дуг. Произвольный граф $G = (X, \Gamma)$ с p вершинами и k дугами — это суграф графа K_p . Обозначим через R множество из двух элементов $R = \{0, 1\}$ с весами $\omega(0) = 1$, $\omega(1) = x$. Очевидно, что графы $G = (X, \Gamma)$ с p вершинами и k дугами ($k \leq$

$\leq p(p-1)$) находятся во взаимно-однозначном соответствии с конфигурациями $f: K_p \rightarrow R$, причем

$$f((t_i, t_j)) = \begin{cases} 1, & \text{если } (t_i, t_j) \in \Gamma, \\ 0 & \text{в противном случае.} \end{cases}$$

Вес конфигурации $\tilde{w}(f) = x^k$.

Обозначим через Q_p группу автоморфизмов графа K_p .

Теорема 1.1. Совокупность изоморфных графов с p вершинами и k дугами есть класс эквивалентных относительно группы Q_p конфигураций $f: K_p \rightarrow R$.

Доказательство теоремы вытекает непосредственно из определения автоморфизмов графа.

Применим теорему Пойа для подсчета числа неизоморфных графов без петель с p вершинами и k дугами. Обозначим это число g_{pk} . Следствием теоремы 1.1 и теоремы Пойа является следующая теорема.

Теорема 1.2. Перечисляющий полином $g_p(x)$ удовлетворяет равенству

$$g_p(x) = P_{Q_p}(1+x, 1+x^2, \dots, 1+x^{p(p-1)}), \quad (1.1)$$

где $P_{Q_p}(x_1, x_2, \dots, x_{p(p-1)})$ — цикловой индекс группы Q_p , рассматриваемой как группа подстановок дуг графа K_p .

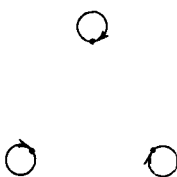
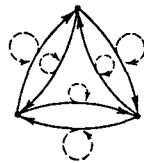
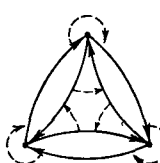
Таким образом, задача свелась к нахождению циклового индекса группы Q_p .

2. Цикловой индекс группы Q_p . Группу Q_p можно определить как группу подстановок упорядоченных пар точек. Для получения ее циклового индекса воспользуемся цикловым индексом группы S_p . Рассмотрим сначала случай $p=3$. Тогда из формулы (6.1) гл. XIII находим следующее выражение для циклового индекса группы S_3 :

$$P_{S_3}(y_1, y_2, y_3) = \frac{1}{6}(y_1^3 + 3y_1y_2 + 2y_3).$$

В табл. 15 показано, какую подстановку $\tilde{g} \in Q_3$ порождает подстановка $g \in S_3$. Суммируя по всем подстановкам из S_3 , получим равенство

$$P_{Q_3}(x_1, x_2, \dots, x_6) = \frac{1}{6}(x_1^6 + 3x_2^3 + 2x_3^2).$$

Член PS_3	Подстановка $g \in S_3$	Подстановка $\tilde{g} \in Q_3$	Член PQ_3
y_1^3			x_1^6
$y_1 y_2$			x_2^3
y_3			x_3^2

В силу теоремы 1.2 для перечисляющего полинома $g_3(x)$ справедлива формула

$$\begin{aligned}
 g_3(x) &= \frac{1}{6}(1+x)^6 + \frac{1}{2}(1+x^2)^3 + \frac{1}{3}(1+x^3)^2 = \\
 &= 1 + x + 4x^2 + 4x^3 + 4x^4 + x^5 + x^6.
 \end{aligned}$$

На рис. 155 показаны все графы с тремя вершинами.

В случае произвольного p член циклового индекса группы Q_p , соответствующий члену $y_1^{j_1} y_2^{j_2} \dots y_p^{j_p}$ в цикловом индексе группы S_p , можно вычислить следующим образом. Рассмотрим сначала те и только те упорядоченные пары, которые состоят из точек, лежащих на циклах подстановки $\underline{g} \in S_p$ одинаковой длины n . При действии подстановки $\underline{g}$ пар точек, порожденной подстановкой $\underline{g}$, такие пары образуют $2nC_{1n}^2 + (n-1)j_n$ циклов длины n (первое слагаемое отвечает парам точек, лежащих на разных циклах, второе — на одном и том же; $C_{1n}^2 = 0$ при

$j_n < 2$). Упорядоченные пары (t_{i_1}, t_{i_2}) такие, что точка t_{i_1} лежит на цикле подстановки $g \in S_p$ длины q , а точка t_{i_2} — длины r , сами образуют циклы подстановки $\tilde{g}$

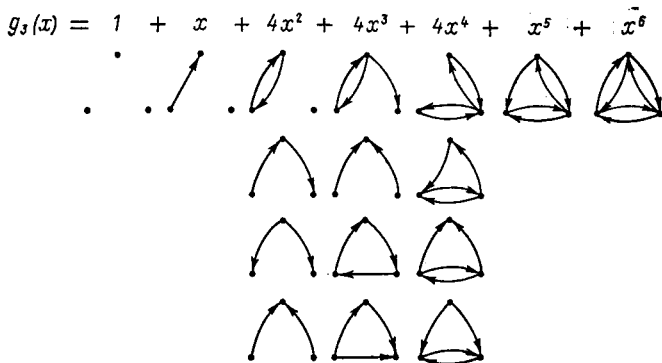


Рис. 155

длины $m(q, r)$ — наименьшего общего кратного чисел q и r . Таких циклов будет $2j_q j_r d(q, r)$, где $d(q, r)$ — наибольший общий делитель чисел q и r . Следовательно, член $y_1^{j_1} y_2^{j_2} \dots y_p^{j_p}$ отвечает члену

$$\prod_{n=1}^p x_n^{2nC_n^2 + (n-1)j_n} \prod_{1 \leq q < r \leq p} x_{m(q,r)}^{2j_q j_r d(q,r)}$$

в цикловом индексе группы Q_q . Отсюда вытекает равенство

$$\begin{aligned} P_{Q_p}(x_1, x_2, \dots, x_{p(p-1)}) &= \\ &= \frac{1}{p!} \sum_{(j)} \frac{p!}{1^{j_1} j_1! 2^{j_2} j_2! \dots p^{j_p} j_p!} \prod_{n=1}^p x_n^{2nC_n^2 + (n-1)j_n} \times \\ &\times \prod_{1 \leq q < r \leq p} x_{m(q,r)}^{2j_q j_r d(q,r)}, \quad (1.2) \end{aligned}$$

где суммирование проводится по всем наборам натуральных чисел $(j_1, j_2, \dots, j_p)$ таких, что $j_1 + 2j_2 + \dots + pj_p = p$.

Пример 1. Требуется подсчитать число графов с четырьмя вершинами.

Имеем $p = 4$, следовательно, суммирование в формуле (1.2) проводится по следующим наборам чисел (j_1, j_2, j_3, j_4) : $(4, 0, 0, 0)$; $(2, 1, 0, 0)$; $(1, 0, 1, 0)$; $(0, 2, 0, 0)$; $(0, 0, 0, 1)$. Каждому набору отвечают следующие слагаемые:

$$\frac{4!}{4} x_1^{2C_4^2} = x_1^{12} \text{ (первое),}$$

$$\frac{4!}{2 \cdot 2} x_1^2 x_2^1 x_2^4 = 6x_1^2 x_2^5 \text{ (второе),}$$

$$\frac{4!}{3} x_3^2 x_3^2 = 8x_3^4 \text{ (третье),}$$

$$\frac{4!}{4 \cdot 2} x_2^6 = 3x_2^6 \text{ (четвертое),}$$

$$\frac{4!}{4} x_4^3 = 6x_4^3 \text{ (пятое).}$$

Суммируя эти слагаемые, получаем равенство

$$P_{Q_4}(x_1, x_2, \dots, x_{12}) = \frac{1}{24} (x_1^{12} + 6x_1^2 x_2^5 + 8x_3^4 + 3x_2^6 + 6x_4^3). \quad (1.4)$$

Из формул (1.4) и (1.1) при $p = 4$ получаем следующее выражение для $g_4(x)$:

$$g_4(x) = 1 + x + 5x^2 + 13x^3 + 27x^4 + 38x^5 + 48x^6 + 38x^7 + \\ + 27x^8 + 13x^9 + 5x^{10} + x^{11} + x^{12}.$$

Симметрия коэффициентов вытекает из того факта, что взятие дополнения графа G относительно полного графа K_p однозначно определяет сам граф G .

§ 2. ПЕРЕЧИСЛЕНИЕ ПОЛНОСТЬЮ НЕОРИЕНТИРОВАННЫХ ГРАФОВ С p ВЕРШИНАМИ И k РЕБРАМИ

1. Конфигурация, отвечающая полностью неориентированному графу. Рассмотрим полный полностью неориентированный граф K_p с p вершинами. Обозначим через R_p группу автоморфизмов графа K_p (R_p — группа перестановок неупорядоченных пар точек). Тогда различные неизоморфные симметрические графы с p вершинами и $k \leq \frac{p(p-1)}{2}$ ребрами можно отождествить с классами эквивалентных относительно группы R_p конфигураций $f: K_p \rightarrow R$, где $R = \{0, 1\}$. Пусть веса фигур удовлетворя-

ют равенствам $\omega(0) = 1$, $\omega(1) = x$, тогда вес конфигураций $\tilde{\omega}(f)$ равен x^k . Обозначим через $\bar{g}_p(x)$ перечисляющий полином

$$\bar{g}_p(x) = \sum_{k=0}^{\frac{p(p-1)}{2}} \bar{g}_{pk} x^k$$

($\bar{g}_{pk}$ — число неизоморфных графов с p вершинами и k ребрами). Аналогом теоремы 1.2 будет следующая теорема.

Теорема 2.1. *Перечисляющий полином $\bar{g}_p(x)$ удовлетворяет равенству*

$$\bar{g}_p(x) = P_{R_p} \left(1 + x, 1 + x^2, \dots, 1 + x^{\frac{p(p-1)}{2}} \right), \quad (2.1)$$

где $P_{R_p}(x_1, x_2, \dots, x_{p(p-1)/2})$ — цикловой индекс группы R_p .

2. Подсчет циклового индекса группы подстановок R_p . Для нахождения циклового индекса группы R_p воспользуемся, как и в § 1, цикловым индексом группы S_p . Пусть сначала $p = 4$. Цикловой индекс группы S_4 дается выражением

$$P_{S_4}(y_1, y_2, y_3, y_4) = \frac{1}{24} (y_1^4 + 6y_1^2 y_2 + 8y_1 y_3 + 3y_2^2 + 6y_4).$$

Рассмотрим табл. 16.

Из табл. 16 получаем формулу

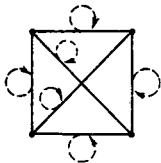
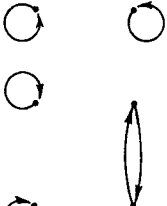
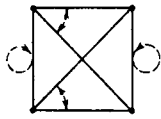
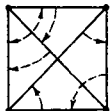
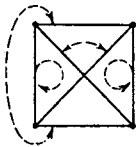
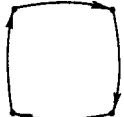
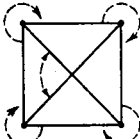
$$P_{R_4}(x_1, x_2, \dots, x_6) = \frac{1}{24} (x_1^6 + 9x_1^2 x_2^2 + 8x_3^2 + 6x_2 x_4). \quad (2.2)$$

Из равенств (2.2) и (2.1) следует, что справедливо следующее выражение для перечисляющего полинома $\bar{g}_4(x)$:

$$\bar{g}_4(x) = 1 + x + 2x^2 + 3x^3 + 2x^4 + x^5 + x^6.$$

На рис. 156 изображены все графы с четырьмя вершинами.

В общем случае вклад от члена $y_1^{i_1} y_2^{i_2} \dots y_p^{i_p} \in P_{S_p}$ в цикловой индекс P_{R_p} может быть разбит на две независимые части, которые затем следует перемножить. Первая часть порождается парами, точки которых лежат на циклах подстановки $g \in S_p$ одинаковой длины n , вторая часть — остальными парами точек.

Член P_{S_4}	Подстановка $g \in S_4$	Подстановка $\tilde{g} \in R_4$	Член P_{R_4}
y_1^4			x_1^6
$y_1^2 y_2$			$x_1^2 x_2^2$
$y_1 y_3$			x_3^2
y_2^2			$x_1^2 x_2^2$
y_4			$x_2 x_4$

Первую часть удобно разделить на вклады от циклов нечетной и четной длин. Для нечетных циклов справедливо соответствие

$$y_{2n+1}^{j_{2n+1}} \rightarrow x_{2n+1}^{(2n+1) C_{j_{2n+1}}^2 + n j_{2n+1}},$$

для четных циклов — соответствие

$$y_{2n}^{j_{2n}} \rightarrow (x_n x_{2n}^{n-1})^{j_{2n}} x_{2n}^{2n C_{j_{2n}}^2}.$$

$$q_4(x) = 1 + x + 2x^2 + 3x^3 + 2x^4 + x^5 + x^6$$

Рис. 156

Вторая часть вклада порождается соответствием

$$y_q^{j_q} y_r^{j_r} \rightarrow x_{m(q,r)}^{j_q j_r d(q,r)},$$

где $m(q, r)$ — наименьшее общее кратное, $d(q, r)$ — наибольший общий делитель чисел q и r . Таким образом, установлено равенство

$$P_{R_p}(x_1, x_2, \dots, x_{\frac{p(p-1)}{2}}) = \frac{1}{p!} \sum_{(j)} \frac{p!}{\prod_{i=1}^p j_i!} \times$$

$$\times \prod_{n=0}^{\left[\frac{p-1}{2}\right]} x_{2n+1}^{(2n+1) C_{j_{2n+1}}^2 + n j_{2n+1}} \prod_{n=1}^{\left[\frac{p}{2}\right]} (x_n x_{2n}^{n-1})^{j_{2n}} x_{2n}^{2n C_{j_{2n}}^2} \times$$

$$\times \prod_{1 \leq q < r \leq p} x_{m(q,r)}^{j_q j_r d(q,r)}. \quad (2.3)$$

Подстановка (2.3) в (2.1) решает поставленную задачу перечисления. Например, для $p = 5$ имеем следующее вы-

ражение для перечисляющего полинома:

$$g_5(x) = 1 + x + 2x^2 + 4x^3 + 6x^4 + 6x^5 + 6x^6 + 4x^7 + \\ + 2x^8 + x^9 + x^{10}.$$

3. Перечисление мультиграфов. В теореме 2.1. было установлено выражение для перечисляющего полинома $\underline{g}_p(x)$ полностью неориентированных графов без кратных ребер. Очевидно, что если в качестве множества фигур R взять множество $\{0, 1, 2\}$, то конфигурацию $f: K_p \rightarrow R$ — можно отождествить с мультиграфом, в котором существует не более двух ребер, соединяющих пары вершин. Поэтому, заменяя перечень $L(R) = 1 + x$ в равенстве (2.1) на перечень $L(R) = 1 + x + x^2$, получаем перечисляющий полином для графов, кратность ребер в которых не превосходит двух.

Следствие 1. Перечисляющий полином $\bar{g}'_p(x)$ для полностью неориентированных мультиграфов с кратностью ребер, не большей двух, равен

$$\bar{g}'_p(x) = P_{R_p} \left(1 + x + x^2, 1 + x^2 + x^4, \dots, \right. \\ \left. 1 + x^{\frac{p(p-1)}{2}} + x^{p(p-1)} \right). \quad (2.4)$$

Если не накладывать ограничений на кратность ребер, то взяв в качестве множества фигур R множество $Z_+ = \{0, 1, 2, \dots\}$ с весами $w(n) = x^n$, приходим к следующему утверждению.

Следствие 2. Перечисляющий полином $\bar{g}''_p(x)$ для мультиграфов с p вершинами имеет вид

$$\bar{g}''_p(x) = P_{R_p} \left((1-x)^{-1}, (1-x^2)^{-1}, \dots, \left(1-x^{\frac{p(p-1)}{2}} \right)^{-1} \right). \quad (2.5)$$

Аналогичные следствия вытекают и из теоремы 1.2.

§ 3. ПЕРЕЧИСЛЕНИЕ МАРКИРОВАННЫХ ГРАФОВ

1. Маркировка. Приемы перечисления графов, основанные на теореме Пойа, далеко не исчерпывают всех способов перечисления. Разработаны и другие методы. Мы по-

знакоимся с двумя из них. Начнем с перечисления маркированных графов.

Вершинами графа могут быть элементы самых различных по своей природе множеств. Часто бывает удобно в качестве множества вершин рассматривать какое-либо конкретное, например, числовое множество.

Определение 3.1. Пусть X и Q — два конечных множества. Взаимно-однозначное отображение $f: Q \rightarrow X$ называется маркировкой множества X множеством Q . Множество Q есть множество меток, X — маркированное множество.

Маркированное множество X может являться множеством вершин некоторого графа $G = (X, \Gamma)$. В этом случае естественно ввести следующее определение.

Определение 3.2. Пара (G, f) , где $G = (X, \Gamma)$ граф, а f — маркировка множества X , называется маркированным графом.

С помощью маркировки f можно определить граф $G(f) = (Q, \Gamma_f)$. Будем считать, что дуга $(q_\alpha, q_\beta, t) \in \Gamma_f$, где $q_\alpha, q_\beta \in Q$, в том и только в том случае, если дуга $(f(q_\alpha), f(q_\beta), t) \in \Gamma$.

Определение 3.3. Два маркированных графа (G, f_1) , (G, f_2) называются изоморфными, если графы G_{f_1} и G_{f_2} совпадают. Маркировки f_1 и f_2 изоморфных маркированных графов называются эквивалентными.

2. Маркировка и группа автоморфизмов графа. Рассмотрим граф $G = (X, \Gamma)$ без петель и кратных дуг. Очевидно, что для любых двух маркировок f_1 и f_2 графа G существует такой элемент g симметрической группы S_n подстановок множества Q ($n = |Q|$), что справедливо равенство

$$f_2 = f_1 \circ g. \quad (3.1)$$

Теорема 3.1. Пусть f — некоторая маркировка графа G . Совокупность H подстановок $g \in S_n$ таких, что маркировки $f \circ g$ эквивалентны маркировке f , изоморфна группе автоморфизмов $\text{Aut}(G)$ графа G .

Доказательство. Докажем, что H — подгруппа группы S_n . Очевидно, что тождественная подстановка $e \in H$. Пусть подстановка $g \in H$. Это означает, что графы G_f и $G_{f \circ g}$ совпадают. Отсюда следует, что дуги (q_α, q_β) и $(g(q_\alpha), g(q_\beta))$ одновременно либо принадлежат, либо не принадлежат множеству Γ_f . Из этого утверждения вытекает, что $g^{-1} \in H$. Аналогично доказывается, что композиция двух подстановок из H также принадлежит H .

По заданной подстановке $g \in H$ построим отображение

$$\varphi = f \circ g \circ f^{-1} \quad (3.2)$$

множества вершин X графа в себя. Из формулы (3.2) вытекает, что φ — взаимно-однозначное отображение, так как отображения f и g взаимно однозначны.

Пусть x_1 и x_2 — смежные вершины, т. е. $(x_1, x_2) \in \Gamma$. Тогда $(\varphi(x_1), \varphi(x_2)) \in \Gamma$, так как дуга $(g \circ f^{-1}(x_1), g \circ f^{-1}(x_2)) \in \Gamma_f$, следовательно, $\varphi \in \text{Aut}(G)$. Обратно, если $\varphi \in \text{Aut}(G)$, то $g = f^{-1} \circ \varphi \circ f \in H$. В силу взаимной однозначности маркировки f формула (3.2) устанавливает изоморфизм групп H и $\text{Aut}(G)$. ■

Следствие. *Данный граф $G = (X, \Gamma)$ можно маркировать $n! / |\text{Aut}(G)|$ способами, где $n = |X|$.*

3. Число маркированных графов с n вершинами. Для того чтобы подсчитать, сколько существует маркированных графов с n вершинами, достаточно рассмотреть полный граф $K_n (Q, Q \times Q \setminus E)$, который содержит $n(n-1)$ дуг. Тогда любой граф G_f , порожденный графом G с n вершинами, есть суграф K_n . Следовательно, справедлива следующая теорема.

Теорема 3.2. *Число неизоморфных маркированных графов с n вершинами равно $2^{n(n-1)}$.*

Для полностью неориентированных маркированных графов с вершинами справедлив аналогичный результат.

Теорема 3.3. *Число неизоморфных маркированных полностью неориентированных графов с n вершинами равно*

$$2^{\frac{n(n-1)}{2}}.$$

На рис. 157 показаны все маркированные полностью неориентированные графы с тремя вершинами.

4. Перечисление маркированных деревьев. Концевыми вершинами дерева T называются вершины x , инцидентные одному ребру ($\text{st } x = 1$).

Лемма 3.1. *Любое дерево T с числом вершин более одной имеет, по крайней мере, две концевые вершины.*

Доказательство. Применим метод индукции по числу ребер. Для дерева из одного ребра утверждение леммы, очевидно, справедливо. Пусть оно справедливо для всех деревьев из n ребер. Удалим из дерева T , содержащего $n+1$ ребро, какое-либо ребро. Так как в дереве любое ребро — перешеек, то дерево T распадается на две компоненты связности, также являющиеся деревьями. Если каждая из них содержит более двух вершин, то в силу предпо-

ложения индукции, в каждой компоненте содержится, по крайней мере, по две концевые вершины, а всего концевых вершин не менее четырех. После восстановления удаленного ребра увеличивается степень лишь у двух концевых вершин. Следовательно, по крайней мере, две концевые вершины

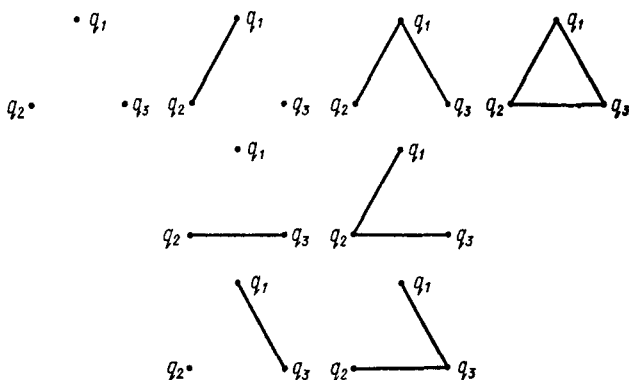


Рис. 157

остаются. Если одна из компонент связности содержит лишь одну вершину, то при восстановлении удаленного ребра эта вершина становится концевой. Остается концевой, по крайней мере, еще одна концевая вершина у второй компоненты связности. ■

Теорема 3.4. Число неизоморфных маркированных деревьев из n вершин равно n^{n-2} .

Доказательство. Подсчитаем число различных, быть может изоморфных, деревьев, которые можно построить на множестве меток Q . Обозначим через

$$q_1, q_2, \dots, q_n \quad (3.3)$$

элементы множества Q . Пусть T — некоторое дерево с множеством вершин Q . Пусть q_{i_1} — первая концевая вершина в последовательности (3.3), а $u_1 = (q_{i_1}, q_{j_1})$ — инцидентное ей ребро. Удалим из дерева T ребро u_1 вместе с вершиной q_{i_1} . Получим новое дерево T_1 . Пусть q_{i_2} — первая в последовательности (3.3) концевая вершина дерева T_1 , а $u_2 = (q_{i_2}, q_{j_2})$ — инцидентное ей ребро. Повторяем этот процесс до тех пор, пока после удаления ребра $u_{n-2} = (q_{i_{n-2}}, q_{j_{n-2}})$ не остается единственное ребро $u_{n-1} =$

$= (q_{i_{n-1}}, q_{j_{n-1}})$, соединяющее две оставшиеся вершины. Последовательность

$$q_{i_1}, q_{j_2}, \dots, q_{j_{n-2}} \quad (3.4)$$

однозначно определяется деревом T . Заметим, что любая вершина $q_i \in Q$ может войти в последовательность (3.4) более одного раза.

По последовательности (3.4) можно однозначно восстановить дерево T . Для этого находим первую вершину q_{i_1} в последовательности (3.3), которая не содержится в последовательности (3.4). Она определяет ребро $u_1 = (q_{i_1}, q_{j_1})$. Вычеркиваем один элемент q_{j_1} из последовательности (3.4) и вершину q_{i_1} — из последовательности (3.3) и продолжаем построение для оставшихся вершин. Легко доказать, что получающийся граф является деревом. В последовательности (3.4) каждый элемент q_{j_m} может совпадать с любой из n вершин множества Q . Различные последовательности вида (3.4) определяют по построению различные деревья. ■

§ 4. ПЕРЕЧИСЛЕНИЕ ДЕРЕВЬЕВ

1. Теорема о характеристике неподобия для деревьев. В этом параграфе производящая функция числа неизоморфных деревьев будет выражена через производящую функцию корневых деревьев.

Определение 4.1. Пусть T — дерево, G_T — группа автоморфизмов дерева T . Две вершины x_i, x_j дерева T называются подобными, если существует автоморфизм $\sigma \in G_T$ такой, что $\sigma(x_i) = x_j$.

Аналогично определяется подобие ребер.

Определение 4.2. Числом неподобных вершин дерева T (числом неподобных ребер) называется число классов эквивалентности вершин (ребер) относительно группы G_T .

Теорема 4.1. Пусть p' и q' — числа неподобных вершин и ребер дерева, q'' — число ребер, соединяющих подобные вершины. Тогда имеет место соотношение

$$p' - (q' - q'') = 1. \quad (4.1)$$

Доказательству теоремы предположим несколько лемм. Пусть T_α — поддереву T (T_α — подграф T , являющийся деревом).

Определение 4.3. Поддереву T_α называется характеристическим, если любые две различные вершины T_α не подобны.

Определение 4.4. Характеристическое поддерево T_α , обладающее тем свойством, что любая вершина $y \in T$, смежная с T_α , имеет в T_α , по крайней мере, одну вершину, подобную себе, называется максимальным характеристическим поддеревом.

Лемма 4.1. Любое дерево T обладает, по крайней мере, одним максимальным характеристическим поддеревом.

Доказательство. Пусть T_1 — поддерево, состоящее из вершины x_1 . Очевидно, что T_1 — характеристическое поддерево. Если все вершины $y \in T$, смежные с x_1 , подобны x_1 , то T_1 — максимальное характеристическое поддерево. Пусть вершина x_2 смежна с x_1 и не подобна x_1 . Тогда, присоединяя к T_1 вершину x_2 и ребро (x_1, x_2) , получим характеристическое поддерево T_2 . Если все вершины $y \in T$, смежные с x_1 или x_2 , подобны либо x_1 , либо x_2 , то T_2 — максимальное. Если вершина x_3 смежна либо с x_1 , либо с x_2 и не подобна ни x_1 , ни x_2 , то, присоединяя к T_2 вершину x_3 и либо ребро (x_1, x_3) , либо ребро (x_2, x_3) , получим характеристическое поддерево T_3 и т. д. После конечного числа шагов получим максимальное характеристическое поддерево T_n . ■

Лемма 4.2. Пусть T_0 — максимальное характеристическое поддерево дерева T . Тогда T_0 содержит по одному элементу из каждого класса эквивалентности подобных вершин дерева T .

Доказательство. В силу свойства максимальнойности поддерева T_0 любая вершина $x \in T$, смежная с T_0 , имеет в T_0 подобную себе. Пусть вершина x_1 , не смежная с T_0 и ни одна вершина, подобная x_1 , не содержится в T_0 . Рассмотрим цепь μ , соединяющую вершину x_1 с какой-то вершиной $y \in T_0$. Пусть $(x_1, x_2, \dots, x_m, y)$ — совокупность вершин цепи μ . Тогда ни одна из вершин, подобных x_2 , не содержится в T_0 . Действительно, в противном случае существовал бы автоморфизм $\sigma \in G_T$ такой, что $\sigma(x_2) \in T_0$. Следовательно, вершина $\sigma(x_1)$ была бы смежна с $\sigma(x_2) \in T_0$ и в силу максимальнойности T_0 принадлежала бы T_0 . Аналогичным свойством обладают вершины $x_3, x_4, \dots, x_m$. Значит, вершина x_m , будучи смежной с вершиной $y \in T_0$, не имеет в T_0 подобных. Это противоречит максимальнойности T_0 . ■

Определение 4.5. Ребро симметрии дерева T называется ребро (x, y) такое, что вершины x и y подобны.

Заметим, что число ребер симметрии q'' входит в формулировку теоремы 4.1.

Лемма 4.3. Каждое дерево T имеет не более одного ребра симметрии.

Доказательство. Пусть $l = (x, y)$ — ребро симметрии. Удалим это ребро из дерева. Тогда дерево T разобьется на два несвязанных между собой дерева T_x и T_y , содержащих соответственно вершины x и y . Если $\sigma \in G_T$ такой, что $\sigma(x) = y$, то автоморфизм σ порождает изоморфизм деревьев T_x и T_y . В самом деле, если вершина $z \in T_x$, то $\sigma(z) \in T_y$, так как z связана цепью с вершиной x , а следовательно, $\sigma(z)$ должна быть связана цепью с вершиной $y = \sigma(x)$ (автоморфизм сохраняет отношение смежности). Из изоморфизма деревьев T_x и T_y следует, что $|T_x| = |T_y|$. Пусть $l' = (x', y')$ — ребро симметрии дерева T , отличное от l . Пусть, для определенности, оно находится в дереве T_x . Удалим ребро l' из дерева, сохранив в нем ребро l . Тогда для соответствующих деревьев $T_{x'}$ и $T_{y'}$ будем иметь равенство $|T_{x'}| = |T_{y'}|$. В то же время это равенство не может иметь места, так как либо $T_{x'}$, либо $T_{y'}$ есть часть дерева T_x . ■

Лемма 4.4. Пусть T_0 — максимальное характеристическое поддереву дерева T . Тогда T_0 содержит по одному элементу из каждого класса эквивалентности ребер, за исключением ребра симметрии, если таковое существует.

Доказательство. Лемма 4.4 есть непосредственное следствие леммы 4.2 и определения поддерева T_0 .

Доказательство теоремы 4.1. Пусть s_0 и s_1 — число вершин и ребер поддерева T_0 . В силу теоремы 1.3 гл. IV $s_0 - s_1 = 1$, в силу лемм 4.2 и 4.4, $s_0 = p'$, $s_1 = q' - q''$. Следовательно, справедливо равенство (4.1). ■

2. Производящая функция чисел неизоморфных деревьев. Заметим, что для данного дерева T число p' есть число неизоморфных корневых деревьев, которые можно получить из T . Следовательно, если равенство (4.1) написать столько раз, сколько существует неизоморфных деревьев с n вершинами, и просуммировать эти равенства, то в правой части получим число t_n неизоморфных деревьев с n вершинами, сумма чисел p' дает число T_n корневых деревьев:

$$T_n - \sum (q' - q'') = t_n. \quad (4.2)$$

Сумму чисел $q' - q''$ можно найти следующим способом. Аналогично корневому дереву рассмотрим дерево с отмеченным ребром. Тогда сумма чисел $q' - q''$ по всем деревьям с n вершинами будет равна числу неизоморфных деревьев с отмеченными ребрами, не являющимися ребрами симметрии. Изоморфизм деревьев с отмеченными ребрами — это такой изоморфизм деревьев, при котором отмеченное ребро соответствует отмеченному. Рассмотрим два корневых дерева: дерево A с i вершинами и дерево B с j вершинами, причем $i + j = n$. Соединим корень дерева A с корнем дерева B ребром. Будем считать это ребро отмеченным. Мы получим дерево с n вершинами и отмеченным ребром.

Лемма 4.5. Пусть A — дерево, A_i ($i = 1, 2, \dots, t$) — неизоморфные корневые деревья, полученные из дерева A . Пусть B — корневое дерево, неизоморфное деревьям A_i . Тогда деревья D_j ($i = 1, 2, \dots, t$) с отмеченными ребрами, полученные соединением корней деревьев A_i с корнем дерева B отмеченным ребром, неизоморфны.

Доказательство. Предположим, что дерево D_i изоморфно дереву D_j , $f(D_i) = D_j$, где f — взаимно-однозначное отображение деревьев, сохраняющее отношение смежности и переводящее отмеченное ребро в отмеченное. Предположим, что корень дерева B (в составе дерева D_i) переходит в корень дерева B (в составе дерева D_j), а корень дерева A_i переходит в корень дерева A_j . Отсюда немедленно вытекает, что корневое дерево A_i изоморфно корневому дереву A_j , что противоречит условию теоремы. Альтернативное предположение о переходе при отображении f корня дерева A_i в корень дерева B и корня дерева B в корень дерева A_j привело бы к заключению о существовании изоморфизма корневых деревьев A_i , A_j и B . ■

Заметим, что ребро симметрии может появиться только при соединении корней изоморфных корневых деревьев. Отсюда следует, что число q_n деревьев с n вершинами и отмеченными ребрами, не являющимися ребрами симметрии, равно

$$q_n = \sum_{\substack{i+j=n \\ i < j}} T_i T_j + C_{T_{n/2}}, \quad (4.3)$$

причем второй член равенства (4.3) соответствует соединению ребрами различных корней корневых деревьев, полученных из одного и того же дерева. Формулу (4.3) можно привести к виду

$$q_n = \frac{1}{2} \sum_{\substack{i+j=n \\ i \neq j}} T_i T_j + \frac{1}{2} T_{n/2} T_{n/2} - \frac{1}{2} T_{n/2} =$$

$$= \frac{1}{2} \sum_{i+j=n} T_i T_j - \frac{1}{2} T_{n/2}. \quad (4.4)$$

Подставляя (4.4) в (4.2), получаем формулу

$$t_n = T_n + \frac{1}{2} \sum_{i+j=n} T_i T_j - \frac{1}{2} T_{n/2}, \quad (4.5)$$

откуда — следующее выражение для производящей функции деревьев:

$$t(x) = T(x) + \frac{1}{2} [T^2(x) - T(x^2)]. \quad (4.6)$$

Глава XV

ЭЛЕКТРИЧЕСКИЕ СЕТИ

Источником теоретических концепций теории графов, изложенных в этой главе, явилось исследование электрических сетей Кирхгофом, предпринятое им в прошлом веке.

§ 1. КОЦЕПИ В ГРАФЕ

Введем понятие, двойственное понятию 1-цепей графа, рассмотренному ранее в гл. IV.

1. Линейные пространства коцепей графа. В этом параграфе будут рассматриваться только связные графы. В гл. IV вершинам и ребрам графа G были сопоставлены алгебраические объекты — группы 0- и 1-цепей. Целесообразно провести подобное сопоставление и в отношении функций, заданных на множествах вершин или ребер графа.

Определение 1.1. Пусть c^* есть гомоморфизм группы $C_0(G)$ в группу вещественных чисел $\mathbb{R}$

$$c^* : C_0(G) \rightarrow \mathbb{R}; \quad c^*(mc_1 + nc_2) = mc^*(c_1) + nc^*(c_2), \quad m, n \in \mathbb{Z},$$

где $c_1, c_2 \in C_0(G)$. Гомоморфизм c^* называется нульмерной коцепью, или 0-коцепью.

Определение 1.2. Суммой коцепей c_1^* и c_2^* называется такая коцепь c^* , что

$$c^*(c) = c_1^*(c) + c_2^*(c), \quad \forall c \in C_0(G).$$

Произведением вещественного числа $\alpha \in \mathbb{R}$ на коцепь c^* называется коцепь αc^* , обладающая свойством

$$(\alpha c^*)(c) = \alpha c^*(c), \quad \forall c \in C_0(G).$$

Множество коцепей с введенными операциями сложения и умножения на вещественные числа образует линейное

пространство, называемое *пространством нульмерных коцепей* $C_0^*(G)$. Если вместо группы нульмерных цепей $C_0(G)$ рассматривается группа 1-цепей, то получается линейное пространство одномерных коцепей $C_1^*(G)$ в полной аналогии с определениями 1.1 и 1.2.

Рассмотрим коцепь $\kappa_j^* \in C_0^*(G)$ такую, что

$$x_j^*(x_i) = \delta_{ji} = \begin{cases} 1, & \text{если } i = j, \\ 0 & \text{в противном случае} \end{cases} \quad (1.1)$$

и коцепь $u_k^* \in C_1^*(G)$ такую, что

$$u_k^*(u_e) = \delta_{ke}. \quad (1.2)$$

Справедлива следующая, приводимая без доказательства теорема.

Теорема 1.1. *Любая коцепь $c \in C_0^*(G)$ однозначно представима в виде*

$$c = \sum_{j=1}^{s_0} \alpha_j \kappa_j^*. \quad (1.3)$$

Любая коцепь $c \in C_1^*(G)$ однозначно представима в виде

$$c = \sum_{k=1}^{s_1} \beta_k u_k^*. \quad (1.4)$$

Равенства (1.3) и (1.4) означают, что размерность линейного пространства $C_0^*(G)$ ($C_1^*(G)$) равна числу вершин (ребер) графа, а коцепи κ_j^* (u_k^*) образуют базис в $C_0^*(G)$ ($C_1^*(G)$).

2. Кограничный оператор и его свойства. Введем гомоморфизм, сопряженный граничному оператору $d: C_1(G) \rightarrow C_0(G)$.

Определение 1.3. *Дифференциалом, или кограничным оператором ∂ в пространстве $C_0^*(G)$ называется отображение*

$$\partial: C_0^*(G) \rightarrow C_1^*(G),$$

определенное следующим образом:

$$1) (\partial c^*)(c) = c^*(dc), \quad \forall c \in C_1(G), \quad \forall c^* \in C_0^*(G),$$

$$2) \partial(\alpha_1 c_1^* + \alpha_2 c_2^*) = \alpha_1 \partial c_1^* + \alpha_2 \partial c_2^*, \quad c_1^*, c_2^* \in C_0^*(G).$$

Найдем дифференциал коцепи x_j^* . Для его вычисления достаточно знать значения ∂x_j^* на любой 1-цепи u_k . В соответствии с (1.3) получаем равенства

$$\begin{aligned} (\partial x_j^*)(u_k) &= x_j^*(du_k) = x_j^*(x_{k_1} - x_{k_2}) = x_j^*(x_{k_1}) - x_j^*(x_{k_2}) = \\ &= \delta_{jk_1} - \delta_{jk_2}, \end{aligned} \quad (1.5)$$

где x_{k_1} — конец, x_{k_2} — начало ребра u_k при канонической ориентации. Следовательно, если ребро u_k не инцидентно вершине κ_j , значение ∂x_j^* на u_k равно нулю, если же ребро u_k инцидентно x_j , то значение ∂x_j^* на u_k равно плюс или минус единице в зависимости от того, является ли вершина x_j концом или началом ребра u_k .

Определение 1.4. Коцепь $b \in C_1^*(G)$, представляемая в виде $b = \partial c^*$, где $c^* \in C_0(G)$, называется кограницей.

Теорема 1.2. Значение кограницы $b \in C_1^*(G)$ на любом цикле $z \in C_1(G)$ равно нулю.

Доказательство этой теоремы непосредственно вытекает из определений 1.3 и 1.4.

Теорема 1.3. Размерность подпространства кограниц $B \subset C_1^*(G)$ равна $s_0 - 1$.

Доказательство. В силу теоремы 1.1 любая 0-коцепь $c^* \in C_0^*(G)$ представима в виде (1.1). Из равенства (1.1) и определения 1.3 вытекает, что любая кограница есть линейная комбинация кограниц ∂x_j^* ($j = 1, \dots, s_0$). Докажем, что совокупность $\{\partial x_j^*\}$ линейно зависима. Для этого рассмотрим кограницу

$$b_1 = \partial x_1^* + \partial x_2^* + \dots + \partial x_{s_0}^*. \quad (1.6)$$

Значение кограницы b_1 на любой цепи вида u_k ($k = 1, 2, \dots, s_1$) равно нулю. В самом деле, пусть κ_{k_1} — начало, x_{k_2} — конец ребра u_k . Тогда $\partial x_{k_1}^*(u_k) = -1$, $\partial x_{k_2}^*(u_k) = +1$, $\partial x_j^*(u_k) = 0$ при $j \neq k_1, k_2$ и, следовательно, $b_1(u_k) = 0$.

Докажем, что из равенства

$$b_2 \equiv \sum_{j=1}^{s_0-1} \alpha_j x_j^* = 0$$

вытекает, что все $\alpha_j = 0$. Для этого рассмотрим значение коцепи b_2 на ребре u_{k_1} , инцидентном вершине x_{s_0} . Пусть $x_{m_1} (m_1 < s_0)$ — другой конец ребра u_{k_1} , тогда $b_2(u_{k_1}) = \pm \alpha_{m_1} = 0$. Пусть u_{k_2} — ребро инцидентное вершине x_{m_1} , $x_{m_2} \neq x_{s_0}$ — конец ребра u_{k_2} , тогда $b_2(u_{k_2}) = \pm \alpha_{m_2} = 0$, потому что $\sigma_{m_1} = 0$. В силу связности графа этот процесс приведет к доказательству сформулированного утверждения. Следовательно, любые $s_0 - 1$ кограниц вида dx_j^* линейно независимы. ■

Пусть $z_1, z_2, \dots, z_\nu \in C_1(G)$ — базис независимых циклов графа G и пусть

$$z_j = \sum_{k=1}^{s_1} \alpha_k^{(j)} u_k. \quad (1.7)$$

Рассмотрим коцепи $z_1^*, z_2^*, \dots, z_\nu^* \in C_1^*(G)$ такие, что

$$z_j^* = \sum_{k=1}^{s_1} \alpha_k^{(j)} u_k^*, \quad (1.8)$$

где коэффициенты $\alpha_k^{(j)}$ те же, что и в равенствах (1.7).

Теорема 1.4. Коцепи $z_1^*, z_2^*, \dots, z_\nu^*$, где $\nu = \nu(G)$, линейно независимы и не лежат в B .

Доказательство. Вычислим значение коцепи z_j^* на цикле z_j :

$$\begin{aligned} z_j^*(z_j) &= \left(\sum_{m=1}^{s_1} \alpha_m^{(j)} u_m^* \right) \left(\sum_{k=1}^{s_1} \alpha_k^{(j)} u_k \right) = \\ &= \sum_{m=1}^{s_1} \sum_{k=1}^{s_1} \alpha_m^{(j)} \alpha_k^{(j)} u_m^*(u_k) = \sum_{k=1}^{s_1} \alpha_k^{(j)^2}. \end{aligned} \quad (1.9)$$

Равенство (1.9) означает, что любая коцепь $z_j^* (j=1, 2, \dots, \nu)$ не является кограницей. Совокупность коцепей $z_1^*, z_2^*, \dots, z_\nu^*$ линейно независима, так как из равенства

$$\sum_{j=1}^{\nu} \beta_j z_j^* = 0 \quad (1.10)$$

вытекало бы, что все $\beta_j = 0$. Для этого достаточно вычислить значения левой части равенства (1.10) на таких ребрах u_k ($k = 1, 2, \dots, v$), которые принадлежат только одному из циклов $z_1, z_2, \dots, z_v$ и не принадлежат остальным. ■

Обозначим через P линейную оболочку коцепей $z_1^*, \dots, z_v^*$.

По теореме 1.2 кограницы $b \in B$ обращаются в ноль на циклах. Выясним, на каких цепях $c \in C_1(G)$ принимают нулевые значения коцепи подпространства P . Для этого рассмотрим множество R_{x_j} ребер, инцидентных вершине x_j в графе. Пусть это будут ребра $u_{j_1}, \dots, u_{j_n}$. Рассмотрим цепь

$$r_{x_j} = \pm u_{j_1} \pm u_{j_2} \pm \dots \pm u_{j_n}, \quad (1.11)$$

причем перед ребром u_{j_k} ставится знак «+», если x_j — конец u_{j_k} при канонической ориентации, а знак «—» в противном случае.

Теорема 1.5. *Справедливы равенства*

$$z_k^*(r_{x_j}) = 0 \quad (k = 1, 2, \dots, v; j = 1, 2, \dots, s_0). \quad (1.12)$$

$$dx_j^*(r_{x_j}) = m, \quad (1.12')$$

где $m = \text{st}x_j$.

Доказательство. Если вершина x_j не инцидентна ни одному ребру, входящему в цикл z_k , то очевидно, что равенство (1.12) справедливо. В противном случае вершина x_j инцидентна сразу двум ребрам цикла z_k , причем для одного из них вершина x_j является началом, для другого — концом при индуцированной ориентации ребер. Отсюда вытекает утверждение теоремы. ■

Линейное пространство $C_1^*(G)$ может быть сделано эвклидовым пространством, если в нем ввести скалярное произведение коцепей.

Определение 1.5. *Скалярное произведение $\langle c_1^*, c_2^* \rangle$ коцепей задается формулой*

$$\langle c_1^*, c_2^* \rangle = c_1^*(c_2), \quad (1.13)$$

где c_2 — такая цепь, что $c_2 = \sum_{j=1}^{s_1} \alpha_j u_j$, если $c_2^* = \sum_{j=1}^{s_1} \alpha_j u_j^*$.

Теорема 1.6. Пусть 1-коцепи $c_1^* = \sum_{j=1}^{s_1} \alpha_j u_j^*$, $c_2^* = \sum_{i=1}^{s_1} \beta_i u_i^*$, тогда справедливы равенства

$$\langle c_1^*, c_2^* \rangle = \sum_{j=1}^{s_1} \alpha_j \beta_j = \sum_{j=1}^{s_1} c_1^*(u_j) c_2^*(u_j).$$

Доказательство. Равенство (1.14) вытекает непосредственно из свойства линейности коцепей (определение 1.2) и соотношений (1.2). ■

Теорема 1.7. Для любой кограницы $b \in B \subset C_1^*(G)$ и любого элемента $c \in P$ выполняется соотношение

$$\langle b, c \rangle = 0. \quad (1.14)$$

Доказательство. В силу теоремы 1.4 любой элемент $c \in P$ есть линейная комбинация коцепей $z_j^* (j = 1, 2, \dots, v)$

$$c = \sum_{i=1}^v \alpha_i z_i^*. \quad (1.15)$$

Вычислим $\langle b, c \rangle$. Используя (1.14), получаем

$$\langle b, c \rangle = b \left(\sum_{i=1}^v \alpha_i z_i \right) = \sum_{i=1}^v \alpha_i b(z_i) = 0.$$

Последнее равенство выполняется в силу теоремы 1.2. ■

§ 2. ПРИМЕНЕНИЕ ГРАФОВ В ТЕОРИИ ЭЛЕКТРИЧЕСКИХ СЕТЕЙ

1. Пассивные и активные элементы электрических сетей. С точки зрения теории графов электрическая сеть — это полностью неориентированный граф G , ребрами которого являются электрические элементы — сопротивления, емкости и индуктивности. Ребром может также служить произвольная комбинация этих элементов. Сопротивления, емкости и индуктивности являются пассивными элементами. Активными элементами электрических сетей служат генераторы тока и напряжения.

2. Сети с пассивными элементами. Для сопротивления, индуктивности и емкости справедливо одно из трех соот-

ношений между силой тока, протекающего через элемент, и разностью потенциалов $v(t)$ на концах элемента:

$$v(t) = Ri(t), v(t) = L \frac{di(t)}{dt}, \frac{dv(t)}{dt} = \frac{1}{C} i(t), \quad (2.1)$$

где R, L, C — величины сопротивления, индуктивности и емкости, считающиеся не зависящими от силы тока (линейные сети) и положительными. Если $i(t) > 0$, то ток течет в направлении, отвечающем канонической ориентации ребра.

После преобразования Лапласа из (2.1) получаем равенства

$$v(p) = Ri(p), v(p) = Lp i(p), v(p) = \frac{1}{Cp} i(p) \quad (2.2)$$

или в общем виде $v(p) = Y(p)i(p)$, где $Y(p)$ — адмитанс ребра (начальные значения тока $i(0)$ и напряжения $v(0)$ полагаем равными нулю).

Равенства (2.2) есть лишь соотношения между током и напряжением на каждом элементе. Для того чтобы узнать, каковы на самом деле эти величины в случае, когда элементы объединены в сеть, применяют законы Кирхгофа: 1) сумма значений тока на всех ребрах, инцидентных каждой вершине графа, равна нулю; 2) сумма значений напряжений на каждом цикле графа равна нулю.

Будем рассматривать ток $I(p)$ и напряжение $V(p)$ как элементы пространства $C_1^*(G)$. Тогда из законов Кирхгофа и результатов § 1 вытекает, что $I(p) \in P$, $V(p) \in B$, т. е.

$$I(p) = \sum_{k=1}^v I_k(p) z_k^*, \quad (2.3)$$

$$V(p) = \sum_{m=1}^{s_0-1} V_m(p) \partial x_m^*. \quad (2.4)$$

Величины $I_k(p)$ называются *токами в циклах*. Зная

$$I_k(p) (k = 1, \dots, v) \text{ и } V_m(p) (m = 1, \dots, s_0 - 1),$$

можно найти ток $i(p)$ и напряжение $v(p)$ на каждом ребре, вычислив значение $I(p)$ или $V(p)$ на этом ребре.

В правых частях равенств (2.3) и (2.4) всего $v + s_0 - 1 = s_1$ неизвестных коэффициентов. Воспользуемся соотно-

шениями (2.2), чтобы получить алгебраическую систему уравнений для $I_k(p)$ и $V_m(p)$.

Подставим значения $v_j(p) = V(p)(u_j)$ и $i_j(p) = I(p)(u_j)$ в систему (2.2). Получим соотношения

$$V(p)(u_j) = Y_j(p) I(p)(u_j), \quad (2.5)$$

которые приводят к следующей линейной алгебраической системе уравнений:

$$\sum_{m=1}^{s_0-1} V_m(p) \partial x_m^*(u_j) - Y_j(p) \sum_{k=1}^v I_k(p) z_k^*(u_j) = 0 \quad (2.6)$$

$$(j = 1, 2, \dots, s_1).$$

Теорема 2.1. Система (2.6) имеет только нулевые решения.

Доказательство. В самом деле, в силу теоремы 1.7 коцепи $I(p)$ и $V(p)$ ортогональны:

$$\langle I(p), V(p) \rangle = \sum_{k=1}^{s_1} I(p)(u_k) V(p)(u_k) = 0. \quad (2.7)$$

Подставляя соотношения (2.5) в (2.7), получим равенство

$$\sum_{k=1}^{s_1} Y_k(p) (I(p)(u_k))^2 = 0. \quad (2.8)$$

В силу того, что $Y_k(p) > 0$, из (2.8) вытекает равенство нулю величин $I(p)(u_k)$ ($k = 1, \dots, s_1$). Линейная независимость коцепей $z_1^*, \dots, z_v^*$ влечет тогда равенства $I_k(p) = 0$. Из соотношений (2.5) и линейной независимости кограниц ∂x_j^* вытекают равенства $V_m(p) = 0$ ($m = 1, 2, \dots, s_0 - 1$). ■

Утверждение теоремы 2.1 физически означает, что в электрических сетях, состоящих только из пассивных элементов, при нулевых начальных условиях для токов и напряжений электрический ток и напряжения всегда равны нулю.

3. Сети с активными элементами. В силу того что генераторы тока и напряжения имеют вполне определенные значения тока и напряжения, их включение в качестве ребер в граф электрической сети накладывало бы дополнительные и иногда противоречивые ограничения на потенциалы в вершинах, к которым подсоединяются генераторы напряжения, и на токи в циклах, в которые встроены генераторы тока. Принимается, что генераторы напряжения вклю-

чаются последовательно с пассивными элементами, а генераторы тока параллельно пассивным элементам, и вся эта совокупность является ребром графа. Следовательно, если $i(t)$, $v(t)$ — ток и напряжение на пассивном элементе, то на ребре ток есть $i(t) - j(t)$, где $j(t)$ — заданный ток генератора тока, а напряжение есть $v(t) - e(t)$, где $e(t)$ — э.д.с. генератора напряжения. Законы Кирхгофа, сформулированные для сети в предыдущем пункте, по-прежнему остаются в силе. Остаются справедливыми также соотношения

$$v_k(p) = Y_k(p) i_k(p) \quad (2.9)$$

между током и напряжением на каждом пассивном элементе. Из сказанного вытекает, что токи и напряжения в цепях с активными элементами представимы в виде (2.3) и (2.4), а уравнения для коэффициентов $I_k(p)$, $V_m(p)$ получаются из следующей системы соотношений:

$$\begin{aligned} v_k(p) &= V(p)(u_k) + e_k(p) = Y_k(p) i_k(p) = \\ &= Y_k(p)(I(p)(u_k) + j_k(p)). \end{aligned} \quad (2.10)$$

Из (2.10) получаем равенство

$$\begin{aligned} \sum_{m=1}^{s_0-1} V_m(p) \partial x_m^*(u_k) - Y_k(p) \sum_{j=1}^s I_j(p) z_j^*(u_k) = \\ = -e_k(p) + Y_k(p) j_k(p). \end{aligned} \quad (2.11)$$

В силу теоремы 2.1 неоднородная система уравнений (2.11) имеет единственное решение.

Аналогичная система уравнений может быть получена для сетей из пассивных элементов при произвольных начальных условиях для значений тока и напряжения на ребрах.

§ 3. МАТРИЦЫ, АССОЦИИРОВАННЫЕ С ГРУППАМИ ОДНОМЕРНЫХ И НУЛЬМЕРНЫХ ЦЕПЕЙ ГРАФА

Матрицы инцидентий и матрица циклов графа естественным образом возникают при «координатном» рассмотрении граничного оператора $d: C_1(G) \rightarrow C_0(G)$ относительно выбранных в $C_0(G)$ и $C_1(G)$ соответствующих «базисов» — системы образующих $\{x_i \mid (i = 1, 2, \dots, s_0)\}$ в группе $C_0(G)$ и системы образующих $\{u_j \mid (j = 1, 2, \dots, s_1)\}$ в группе $C_1(G)$.

1. Матрица инцидентий вершин и ребер графа, не имеющего петель. Дадим определение матрицы инцидентий.

Определение 3.1. Матрицей инцидентий D графа G называется матрица гомоморфизма $d: C_1(G) \rightarrow C_0(G)$ относительно образующих элементов $\{x_i \ (i = 1, 2, \dots, s_0)\}$ в $C_0(G)$ и $\{u_j \ (j = 1, 2, \dots, s_1)\}$ в $C_1(G)$:

$$d(u_j) = \sum_{i=1}^{s_0} d_{ij} x_i \quad (j = 1, 2, \dots, s_1). \quad (3.1)$$

Применим к обеим частям равенства (3.1) гомоморфизм x_i^* , тогда

$$d_{ij} = x_i^* d(u_j) = x_i^* (x_{j_2} - x_{j_1}) \quad (i = 1, 2, \dots, s_0, \ j = 1, 2, \dots, s_1), \quad (3.2)$$

где x_{j_2} — конец ребра u_j , а x_{j_1} — начало ребра u_j при канонической ориентации (см. гл. IV).

Из равенства (3.2) окончательно получим, что

$$d_{ij} = \begin{cases} 0, & \text{если вершина } x_i \text{ не инцидентна дуге } u_j, \\ +1, & \text{если вершина } x_i \text{ является концом дуги } u_j, \\ -1, & \text{если вершина } x_i \text{ является началом дуги } u_j. \end{cases} \quad (3.3)$$

Таким образом, можно дать следующее определение матрицы инцидентий, очевидным образом равносильное определению 3.1.

Определение 3.2. Матрица инцидентий графа G — это матрица D порядка $(S_0 \times S_1)$, элементы которой задаются соотношением (3.3).

Задача 1. Показать, что справедливы равенства

$$u_j^* (r_{x_i}) = \partial x_i^* (u_j) = d_{ij} \quad (i = 1, 2, \dots, S_0, \ j = 1, 2, \dots, S_1), \quad (3.4)$$

где 1-цепь r_{x_i} определена формулой (1.11) § 1.

Теорема 3.1. Ранг матрицы D равен $S_0 - 1$.

Доказательство. Из теоремы 2.1 гл. IV следует, что ядро гомоморфизма d $\text{Ker } d = Z(G)$ имеет размерность, равную цикломатическому числу $v(G) = S_1 - S_0 + 1$ графа G^* . Отсюда получаем, что

* Здесь и всюду ниже рассматривается связный граф G .

$$\text{rang } D = S_1 - \text{rang } Z(G) = S_1 - (S_1 - S_0 + 1) = S_0 - 1. \blacksquare$$

З а м е ч а н и е 1. Тот факт, что $\text{rang } D \leq S_0 - 1$, непосредственно следует из определения матричных элементов d_{ij} . Действительно, каждое ребро графа без петель инцидентно двум и только двум вершинам графа, поэтому любой столбец матрицы D содержит только два ненулевых элемента $(+1)$ и (-1) . Если сложить все строки матрицы D , то получим равенства

$$\sum_{i=1}^{S_0} d_{ij} = 0 \quad (j = 1, 2, \dots, S_1), \quad (3.5)$$

которые означают линейную зависимость всех строк матрицы инцидентности.

Если 1-коцепь $c^1 \in P(G)$, то тогда справедлива система равенств

$$\sum_{j=1}^{S_1} d_{ij} c_j = 0 \quad (i = 1, 2, \dots, S_0), \quad (3.6)$$

$$c_j = c^1(u_j) \quad (j = 1, 2, \dots, S_1).$$

Действительно, из условия $c^1 \in P(G)$ и теоремы 1.5 следует, что значение коцепи c^1 на 1-цепи вида r_{x_i} при каждом $i = 1, 2, \dots, S_0$ равно нулю, т. е.

$$c^1(r_{x_i}) = \sum_{j=1}^{S_1} c_j \mu_j^*(r_{x_i}) = 0 \quad (i = 1, 2, \dots, S_0).$$

Подставляя сюда значение коцепи $u_j^*(r_{x_i})$ из формулы (3.4), приходим к системе (3.6).

Таким образом, матрица инцидентий D является матрицей линейной системы уравнений для нахождения коэффициентов разложения по базису $\{u_j^*\}$ коцепи $c^1 \in C_1^*(G)$, удовлетворяющей первому закону Кирхгофа в вершинах графа G .

З а м е ч а н и е 2. Обозначим через D_T произвольный невырожденный минор порядка $(S_0 - 1) \times (S_0 - 1)$ матрицы D . Тогда множество ребер, отвечающих столбцам матрицы D_T , и множество вершин графа X образуют суграф $T(G)$ графа G , являющийся частичным деревом графа G .

Задача 2. Доказать, что $|\det D_T| = |\det D_T^{-1}| = 1$.

2. Матрица циклов графа. Рассмотрим определение матрицы циклов графа.

Определение 3.3. Матрицей циклов Z графа G , порожденной частичным деревом $T(G)$, называется матрица разложения базисных 1-циклов $z_1, z_2, \dots, z_{S_1}$ по образующим $\{u_1, u_2, \dots, u_{S_1}\}$ группы $C_1(G)$, т. е.

$$z_i = \sum_{j=1}^{S_1} \alpha_{ij} u_j \quad (i = 1, 2, \dots, S_1). \quad (3.7)$$

Теорема 3.2. Ранг матрицы Z равен цикломатическому числу $S_1 - S_0 + 1 = \nu(G)$ графа G .

Доказательство следует из линейной независимости 1-циклов $z_1, z_2, \dots, z_{S_1}$, образующих базу 1-циклов графа G .

Применим к обеим частям равенства (3.7) функционалы u_j^* ($j = 1, 2, \dots, S_1$), тогда получим:

$$\alpha_{ij} = u_j^*(z_i) = \begin{cases} 0, & \text{если ребро } u_j \text{ не принадлежит циклу } z_i, \\ +1, & \text{если ориентация ребра } u_j \text{ в цикле } z_i \text{ совпадает с канонической ориентацией } u_j \\ -1, & \text{если ориентация ребра } u_j \text{ в цикле } z_i \text{ противоположна канонической ориентации.} \end{cases} \quad (3.8),$$

Используя соотношения (3.8), можно дать следующее определение матрицы циклов графа G .

Определение 3.5. Матрицей циклов графа G , порожденной деревом $T(G)$, называется матрица Z порядка $\nu \times S_1$, где ν — цикломатическое число графа G , элементы которой определяются из соотношения (3.8).

Если $\{u_j \mid (j = 1, 2, \dots, S_1)\}$ — образующие элементы группы $C_1(G)$, то через c будем обозначать S_1 — мерный вектор с целочисленными компонентами $c = (c_1, c_2, \dots, c_{S_1})$, где $c_1, c_2, \dots, c_{S_1}$ — коэффициенты цепи $c \in C_1(G)$ по образующим группы $C_1(G)$,

$$c = \sum_{j=1}^{S_1} c_j u_j \quad (c_j \in \mathbb{Z}).$$

Если цепь c — 1-цикл, то соответствующий вектор c называется *циклическим*. Пусть 1-цепь $c \in Z(G)$, т. е. $dc = 0$,

тогда в базисе $\{u_j \ (j = 1, 2, \dots, S_1)\}$ условие равенства нулю дифференциала d на цепи c означает, что

$$Dc = 0. \quad (3.9)$$

Цикл c раскладывается по образующим $\{z_1, z_2, \dots, z_v\}$ подгруппы $Z(G)$,

$$c = \sum_{k=1}^v \gamma_k z_k.$$

Соответственно для вектора c справедливо разложение

$$c = \sum_{k=1}^v \gamma_k z_k. \quad (3.10)$$

Если z_k — соответствующий циклу z_k вектор, то из определения 3.4 следует, что $z_k = (a_{k_1}, a_{k_2}, \dots, a_{k_{s_1}})$ — строка цикломатической матрицы Z . Тогда из (3.9) и (3.10) следует, что

$$Dz_k^{tr} = 0 \ (k = 1, 2, \dots, v), \quad (3.11)$$

где z_k^{tr} — (S_1) -мерный вектор-столбец.

Условие (3.11) можно переписать в виде условия ортогональности матрицы D и транспонированной матрицы Z^{tr}

$$DZ^{tr} = 0. \quad (3.12)$$

Полученный результат о связи матриц D и Z сформулируем в виде теоремы.

Теорема 3.3. Матрица циклов и матрица инцидентий графа удовлетворяют соотношению

$$DZ^{tr} = 0. \quad (3.13)$$

З а м е ч а н и е 3. Пусть $T(G)$ — частичное дерево графа G . Перенумеруем ребра графа таким образом, чтобы дуги $\{u_{v+1}, u_{v+2}, \dots, u_{s_1}\}$ принадлежали дереву $T(G)$. Множество дуг дерева $T(G)$ обозначим через $T(T = \{u_{v+1}, u_{v+2}, \dots, u_{s_1}\})$. Через T_* обозначим дополнение $T_* = \bar{G}/T$. Очевидно, что при указанной нумерации дуг графа матрица циклов Z имеет вид

$$\overbrace{\begin{pmatrix} 1 & 0 \\ 1 & \\ \vdots & \\ 0 & \cdot 1 \end{pmatrix}}^{T^*} \overbrace{\begin{pmatrix} 0 \\ Z_T \end{pmatrix}}^T \quad (3.14)$$

3. Решение неоднородной линейной системы первого закона Кирхгофа. Пусть дана неоднородная система уравнений относительно переменных $q_1, q_2, \dots, q_{s_1}$, $q_j \in \mathbb{R}$ ($j = 1, 2, \dots, S_1$) с матрицей D

$$\sum_{j=1}^{s_1} d_{ij} q_j = p_i, \quad (i = 1, 2, \dots, S_0). \quad (3.15)$$

При $p_i = 0$ ($i = 1, 2, \dots, S_0$) однородная система (3.15), как уже отмечалось выше, является записью первого закона Кирхгофа для цепи $c_1^* = \sum_{j=1}^{s_1} q_j u_j^*$, при $p_i \neq 0$ неоднородная система (3.15) отвечает первому закону Кирхгофа для электрической сети с внешними генераторами тока силы p_i ($i = 1, 2, \dots, S_0$).

Теорема 3.4. 1) Система (3.15) совместна, тогда и только тогда, когда выполняются условия

$$\sum_{i=1}^{S_0} p_i = 0;$$

2) если система (3.15) совместна, то решение системы — вектор-столбец $q = (q_1, q_2, \dots, q_{s_1})^{tr}$ представим в виде

$$q = Z^{tr} \gamma + (q_T, D_T^{-1} p)^{tr}, \quad (3.16)$$

где γ — произвольный v -мерный вектор Z^{tr} — транспонированная матрица циклов графа G относительно дерева T , q_T — произвольный v -мерный вектор, D_T — невырожденная подматрица матрицы D , столбцы которой занумерованы множеством индексов T , $|T| = S_0 - 1$.

Доказательство. Условия совместности системы (3.14) следуют из свойства (3.5) матрицы D . Действительно, если

$$\sum_{i=1}^{S_0} d_{ij} = 0 \quad (j = 1, 2, \dots, S_1),$$

то

$$\sum_{i=1}^{S_0} p_i = \sum_{i=1}^{S_0} \sum_{j=1}^{S_1} d_{ij} q_j = \sum_{j=1}^{S_1} q_j \left(\sum_{i=1}^{S_0} d_{ij} \right) = 0.$$

Решение неоднородной системы (3.15) представим в виде

$$q = q_0 + q_1.$$

где q_0 — общее решение однородной системы, q_1 — частное решение неоднородной системы

$$Dq_0 = 0, \quad (3.17a)$$

$$Dq_1 = p, \quad p = (p_1, \dots, p_{S_0}). \quad (3.17b)$$

В силу условия (3.10) решение однородного уравнения (3.17a) можно записать в виде

$$q_0 = \sum_{k=1}^v \gamma_k z_k, \quad (3.18)$$

где вектор $Z_k = (z_{k_1}, z_{k_2}, \dots, z_{k_{S_1}})$ — k -я строка цикломатической матрицы относительно дерева T , а γ_k — произвольные действительные числа.

Определим вектор-столбец $\gamma = (\gamma_1, \gamma_2, \dots, \gamma_v)^T$, тогда равенство (3.17a) означает, что

$$q_0 = Z^T \gamma, \quad q_0 = (q_1^0, \dots, q_{S_0}^0) \quad (3.19)$$

или, что

$$q_j^0 = \sum_{k=1}^v \gamma_k z_{kj} = \sum_{k=1}^v u_j^*(z_k) = \sum_{k=1}^v \gamma_k z_k^*(u_j), \quad j = 1, 2, \dots, S_1.$$

Для нахождения частного решения системы q_1 заметим, что $\text{rang } D = S_0 - 1$, поэтому достаточно рассмотреть, в (3.17b), например, первые $(S_0 - 1)$ уравнений

$$\sum_{i=1}^{S_1} d_{ij} q_i^{(1)} = p_i \quad (i = 1, 2, \dots, S_0 - 1). \quad (3.20)$$

Пусть D_T — невырожденная подматрица матрицы D порядка $(S_0 - 1) \times (S_0 - 1)$, отвечающая дереву $T(G)$.

Из системы (3.20) величины $q_j^{(1)}$ ($j = \nu + 1, \dots, S_1$) однозначно определяются через величины p_i ($i = 1, 2, \dots, S_0 - 1$) и произвольные значения $q_j^{(1)}$ ($j = 1, 2, \dots, \nu$).

Таким образом, частное решение системы q_1 запишется в виде

$$q_1 = (q_{T_*}, D_T^{-1} p)^{tr},$$

где $q_{T_*} = (q_1, q_2, \dots, q_\nu)^{tr}$, $D_T^{-1} p = (q_{\nu+1}^{(1)}, \dots, q_{S_1}^{(1)})$ — решение невырожденной неоднородной системы (3.17 б) с матрицей D_T .

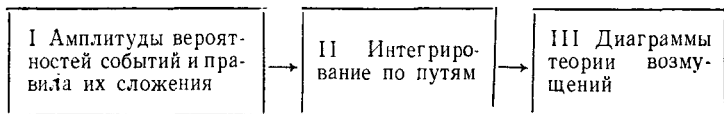
ЛИТЕРАТУРА

1. Б а с а к е р Р., С а а т и Г. Конечные графы и сети. М., «Наука», 1974.
2. Б е р ж К. Теория графов и ее применение. М., ИЛ, 1962.
3. В а с и л ь е в П. Н. Сб. тр. ЦЭМИ АН СССР, вып. 1 (1970), 148—176.
4. М е л и х о в. Ориентированные графы и конечные автоматы. М., «Наука», 1971.
5. О р е О. Теория графов. М., «Наука», 1968.
6. Прикладная комбинаторная математика. Под. ред. Э. Б е к к е н - б а х а. М., «Мир», 1968.
7. Ф о р д Л. Р., Ф а л к е р с е н Д. Р. Потоки в сетях. М., «Мир», 1966.
8. Х а р а р и Ф. Теория графов. М., «Мир», 1973.
9. Х о л л М. Комбинаторика. М., «Мир», 1970.
10. P o l y a G. On picture-writing, Amer. math monthly, 63 (1956), 689—697.

ТЕОРИЯ ВОЗМУЩЕНИЙ КОМПЛЕКСНЫХ МАРКОВСКИХ ЦЕПЕЙ

В работах Р. Фейнмана в конце 40-х годов, посвященных новой формулировке нерелятивистской квантовой механики и созданию квантовой электродинамики, можно выделить, по крайней мере, три важные идеи, послужившие в дальнейшем источником многих приложений в различных разделах теоретической физики. Это, во-первых, представление о бесконечномерном интегрировании по траекториям, во-вторых, введение так называемых диаграмм теории возмущений — графических схем для изображения и интерпретации членов ряда теории возмущений и, наконец, понятие о комплексной амплитуде вероятностей событий в квантовой механике и правилах их сложения.

Если следовать хронологической последовательности появления статей в 1948—1949 гг., то генетическую связь указанных трех идей можно изобразить так:



Отсюда видно, что вероятностная концепция нерелятивистской квантовой механики послужила отправной точкой для формулировки метода функционального интегрирования (интегрирования по траекториям).

Если связь $I \rightarrow II$ явно представлена в работах Фейнмана (см., например [10],) то переход $II \rightarrow III$ в деталях нигде явно не проделан, хотя логическая связь между интегрированием по траекториям и диаграммами теории возмущений у Фейнмана была, и авторы дополнения попытаются ее здесь восстановить. Последовательно и математически строго удастся это проделать, отправляясь от дискретной

модели нерелятивистской квантовой механики — так называемой комплексной марковской цепи (КМ-цепи), введенной в работе [5].

Глава I

КОМПЛЕКСНЫЕ МАРКОВСКИЕ ЦЕПИ И СУММИРОВАНИЕ ПО ПУТЯМ

Определение дискретной комплексной марковской цепи (КМ-цепи) естественным образом связано с теорией графов и теорией матриц.

В § 1 рассматривается связь теории графов с представлением матриц и операций над ними на графах. Здесь же вводятся понятия графа матрицы, суммы и упорядоченного произведения графов матриц. Эти понятия составляют теоретико-графовую основу приложения КМ-цепи к физическим идеям Фейнмана.

В § 2 определяется однородная КМ-цепь и вводится понятие амплитуды Грина КМ-цепи, которая выступает в приложениях как комплексная амплитуда вероятности перехода квантовой системы из одного состояния в другое. Представление амплитуды Грина КМ-цепи на графах матриц реализует правила сложения амплитуд вероятностей событий по Фейнману. Представление амплитуды Грина КМ-цепи на графе матрицы в виде суммы по путям этого графа приводит естественным образом к дискретному аналогу так называемого континуального интеграла Фейнмана.

Такая интерпретация амплитуды Грина КМ-цепи позволяет получить обобщение «континуального интеграла» Фейнмана, рассматривая последний как предел решения разностной схемы, отвечающей однородной КМ-цепи. Соответствующий пример рассматривается в § 3. Для вычисления амплитуды Грина в этом примере эффективно используются методы, развитые в учебном руководстве [6].

§ 1. АЛГЕБРА МАТРИЦ И ГРАФЫ

I. Граф матрицы. Представление произведения матриц на сумме графов перемножаемых матриц. Предварительно рассмотрим определение.

Определение 1.1. Суммой $G_1 + G_2$ двух мультиграфов $G_1 = (X, \Gamma_{G_1})$ и $G_2 = (X, \Gamma_{G_2})$ с одним и тем же множест-

вом вершин X называется мультиграф $G' = (X, \Gamma_{G_1+G_2})$, множество дуг $\Gamma_{G_1+G_2}$ которого определяется равенством

$$\Gamma_{G_1+G_2} = \Gamma_{G_1} \vee \Gamma_{G_2} \quad (1.1)$$

(см. § 1 гл. I). Наглядно равенство (1.1) означает, что множество дуг мультиграфа $G_1 + G_2$ получается объединением дуг Γ_{G_1} и Γ_{G_2} (рис. 158).

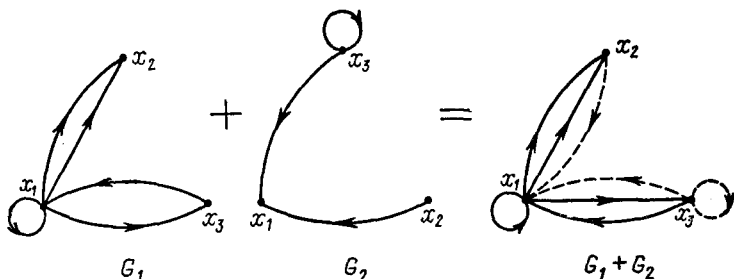


Рис. 158

Из определения 1.1 следует, что множество Γ_{G_1} дуг графа $G_1 + G_2$ совместно с множеством вершин X образует суграф* графа $G_1 + G_2$, который будем обозначать по-прежнему через G_1 . Аналогичное обозначение используется для суграфа (X, Γ_{G_2}) графа $G_1 + G_2$.

На рисунках удобно различить дуги графа $G_1 + G_2$, отвечающие суграфу G_1 от дуг суграфа G_2 так, как это изображено на рис. 158.

Определение 1.2. Пусть $A = ((a_{ij}))$ — произвольная матрица порядка $N \times N$ с элементами $a_{ij} \in \mathbb{C}$. Граф $G_A = (X_A, \Gamma_A)$ без кратных дуг, множество вершин X_A и множество дуг Γ_A которого определяется из условий:

$$1) X_A = \{x_1, x_2, \dots, x_N\}, \quad (1.2a)$$

$$2) \Gamma_A = \{(x_i, x_j) \mid a_{ij} \neq 0\}, \quad (1.2б)$$

назовем графом матрицы A . Матричный элемент $a_{ij} \neq 0$ будем называть весом дуги (x_i, x_j) и обозначать через $f((x_i, x_j))$.

Из условия (1.2 б)) следует, что если элемент $a_{ij} = 0$, то дуги (x_i, x_j) в графе G_A нет.

* См. § 1 гл. I.

Таким образом, произвольной матрице A однозначным образом можно сопоставить:

1) граф без кратных дуг $G_A = (X_A, \Gamma_A)$, определяемый условиями (1.2);

2) весовую функцию f , определенную на множестве Γ_A дуг графа $f: \Gamma_A \rightarrow \mathbb{C} \setminus \{0\}$ формулой

$$f((x_i, x_j)) = a_{ij} \quad (i, j = 1, 2, \dots, N). \quad (1.3)$$

При описанном соответствии матриц и графов без кратных дуг следует иметь в виду, что различным двум матрицам может соответствовать один и тот же граф, точнее — два графа, изоморфные друг другу в смысле определения 2.1 гл. III.

Действительно, рассмотрим преобразование матрицы A , состоящее в одновременной перестановке любых двух ее строк и соответствующих им столбцов с теми же номерами. Обозначим через $\sigma_{i_k j_k}$ операцию перестановки двух строк и столбцов матрицы A с номерами i_k и j_k и рассмотрим композицию из N перестановок строк и столбцов матрицы A

$$\sigma = \sigma_{i_{n j_n} i_{n-1 j_{n-1}}} \dots \sigma_{i_1 j_1}. \quad (1.4)$$

Пусть A' — матрица, получившаяся из матрицы A указанной перестановкой строк и столбцов,

$$A' = \sigma(A), \quad a'_{\sigma(i) \sigma(j)} = a_{ij} \quad (i, j = 1, 2, \dots, N). \quad (1.5)$$

Всякая перестановка $\sigma_{i_k j_k}$ строк и столбцов матрицы определяет на множестве $\{1, 2, \dots, N\}$ подстановку, которую обозначим через $\tilde{\sigma}_{i_k j_k}$,

$$\tilde{\sigma}_{i_k j_k} = (1, 2, \dots, i_k, \dots, j_k, \dots, n) \cdot (1, 2, \dots, j_k, \dots, i_k, \dots, n).$$

Соответственно, композиции (1.4) отвечает подстановка $\sigma \in S_N$, где S_N — группа всех подстановок множества $\{1, 2, \dots, N\}$.

Таким образом, преобразованию σ матрицы A в матрицу A' отвечает изоморфизм соответствующих им графов $\sigma(G_A) = G_{A'}$, который задается отображением множества

вершин графа G_A в множество вершин графа G_A , по формуле*

$$\sigma(x_i) = x'_{\sigma(i)}, \quad x_i \in X_A, \quad x'_{\sigma(i)} \in X_{A'}.$$

Определение 1.3. Скажем, что матрица A приводится к матрице A' , если существует такая перестановка σ строк и столбцов матрицы, A , что $A' = \sigma(A)$.

Как отмечено выше, приводимым матрицам соответствуют изоморфные графы.

Если матрица A приводится к блочно-диагональной матрице с числом блоков, равным p , то это очевидно, означает, что граф G_A изоморфен графу, состоящему из p компонент связности.

Определение 1.4. Матрицу A , не приводящуюся к блочно-диагональной, будем называть неприводимой матрицей.

Перейдем теперь к представлению произведения матриц A_1 и A_2 на графе $G_{A_1} + G_{A_2}$. Рассмотрим на графе $G_{A_1} + G_{A_2}$ упорядоченную тройку вершин (x_i, x_k, x_j) такую, что

$$x_k \in \Gamma_{G_{A_1}}(x_i), \quad x_j \in \Gamma_{G_{A_2}}(x_k). \quad (1.6)$$

Последовательность вершин (x_i, x_k, x_j) , удовлетворяющая условию (1.6), определяет последовательность дуг l_{ij}^2

$$l_{ij}^2 = \{(x_i, x_k) \in \Gamma_{G_{A_1}}(x_k, x_j) \in \Gamma_{G_{A_2}}\} = \{u_1, u_2\}, \quad (1.7)$$

которая на графе $G_{A_1} + G_{A_2}$ представляет собой путь длины два такой, что первая дуга пути принадлежит суграфу G_{A_1} , вторая — суграфу G_{A_2} . Такой путь l на графе $G_1 + G_2$ назовем *чередующимся путем*, отвечающем упорядоченной цепочке (A_1, A_2) матриц A_1 и A_2 .

Обозначим через $l_{ij}^{(2)}(A_1, A_2)$ множество всех чередующихся путей длины два на графе $G_1 + G_2$ из вершины x_i в вершину x_j . Из определения 1.2 графа матрицы следует, что каждому чередующемуся пути $l \in l_{ij}^{(2)}(A_1, A_2)$ отвечает отличное от нуля произведение $a_{ik}^{(1)} a_{kj}^{(2)}$ элементов матриц A_1 и A_2 .

Справедливо и обратное утверждение. Каждый ненулевой

* См. § 2 гл. XI.

член суммы, выражающий собой матричный элемент $[A_1 A_2]_{ij}$ произведения матриц A_1 и A_2

$$[A_1 A_2]_{ij} = \sum_{k=1}^N a_{ik}^{(1)} \cdot a_{kj}^{(2)}, \quad (1.8)$$

определяет на графе $G_{A_1} + G_{A_2}$, в точности один путь длины два.

Таким образом, суммирование ненулевых членов разложения (1.8) связано с геометрической структурой графа $G_{A_1} + G_{A_2}$, точнее со структурой множества всех чередующихся путей длины два на графе $G_{A_1} + G_{A_2}$, отвечающих упорядоченной цепочке (A_1, A_2) .

Если величину $a_{ik}^{(1)} a_{kj}^{(2)}$ рассматривать как вклад $f(l)$ соответствующего чередующегося пути длины два на графе $G_{A_1} + G_{A_2}$, то из равенства (1.8) получаем следующее правило.

Правило вычисления матричных элементов $[A_1 A_2]_{ij}$ произведения матриц $A_1 A_2$ на графе $G_{A_1} + G_{A_2}$

Элемент $[A_1 A_2]_{ij}$ произведения матриц A_1 и A_2 равен сумме вкладов всех чередующихся путей длины два на графе $G_{A_1} + G_{A_2}$, ведущих из вершины x_i в вершину x_j ,

$$[A_1 A_2]_{ij} = \sum_{k=1}^N a_{ik}^{(1)} a_{kj}^{(2)} = \sum_{l \in l_{ij}^2(A_1 A_2)} f(l).$$

З а м е ч а н и е 1. На графе $G_{A_1} + G_{A_2}$ можно рассматривать пути l произвольной длины $n (n > 2)$ с произвольным чередованием дуг суграфов G_{A_1} и G_{A_2} . Такие пути будем называть *произвольно чередующимися* путями на графе $G_{A_1} + G_{A_2}$. Множество всех произвольно чередующихся путей длины n на графе $G_{A_1} + G_{A_2}$ обозначим через $\widehat{l}_{ij}^{(n)}(A_1 A_2)$.

Обобщим предыдущие рассуждения на произвольное число матриц $A_1, A_2, \dots, A_n$.

Определение 1.5. *Путь l длины n из вершины x_i в вершину x_j на графе $\sum_{i=1}^n G_{A_i}$, отвечающем упорядоченной цепочке матриц $(A_1, A_2, \dots, A_n)$, определяется набором $(n +$*

$+1$)-й вершины $(x_i, x_{i_1}, \dots, x_{i_{n-1}}, x_j)$ таким, что если $l = \{(x_i, x_{i_1}), (x_{i_1}, x_{i_2}), \dots, (x_{i_{n-1}}, x_j)\}$, то

$$(x_i, x_{i_1}) \in \Gamma_{A_1}, (x_{i_1}, x_{i_2}) \in \Gamma_{A_2}, \dots, (x_{i_{n-1}}, x_j) \in \Gamma_{A_n}.$$

Такой путь l длины n на графе $\sum_{i=1}^n G_{A_i}$ назовем чередующимся путем упорядоченной цепочки матриц $(A_1, A_2, \dots, A_n)$.

Если через $l_{ij}^{(n)}(A_1, A_2, \dots, A_n)$ обозначить множество всех чередующихся путей длины n цепочки $(A_1, A_2, \dots, A_n)$, соединяющих вершин x_i и x_j , то множества $L_{ij}^{(n)}(A_1, A_2, \dots, A_n)$ образуют разбиение множества $L^n(A_1, A_2, \dots, A_n)$ — множества всех чередующихся путей длины n упорядоченной цепочки матриц $(A_1, A_2, \dots, A_n)$ на графе $\sum_{i=1}^n G_{A_i}$.

$$L^n(A_1, A_2, \dots, A_n) = \bigcup_{i,j=1}^N l_{ij}^{(n)}(A_1, A_2, \dots, A_n),$$

$$l_{ij}^{(n)} \cap l_{i_1 j_1}^{(n)} = \delta_{ii_1} \delta_{jj_1}. \quad (1.9)$$

Вкладом чередующегося пути l цепочки $(A_1, A_2, \dots, A_n)$ назовем величину

$$f(l) = \prod_{k=0}^{n-1} f((x_{i_k}, x_{i_{k+1}})), \quad (1.10)$$

где

$$f((x_{i_k}, x_{i_{k+1}})) = a_{i_k i_{k+1}}^{(k+1)}, a_{i_k i_{k+1}}^{(k+1)} \in A_{k+1} \quad (k = 0, 1, \dots, n-1; i_0 = i; i_n = j).$$

Правило вычисления элементов произведения матриц $A_1 A_2 \dots A_n$ на графе $\sum_{i=1}^n G_{A_i}$

Матричный элемент $[A_1 A_2 \dots A_n]_{ij}$ произведения матриц $A_1, A_2, \dots, A_n$ равен сумме вкладов всех чередующихся

путей n длины на графе $\sum_{i=1}^n G_{A_i}$ из вершины x_i в вершину x_j ,

$$[A_1 A_2 \dots A_n]_{ij} = \sum_{e \in I_{ij}^n(A_1 A_2 \dots A_n)} f(e), \quad (1.11)$$

где вклад $f(l)$ определяется из формулы (1.10)

Положим $A_1 = A_2 = \dots = A_n = A$. Тогда $G_{A_i} = G_A$ ($i = 1, 2, \dots, N$) и поэтому любой чередующийся путь на графе $\sum_{i=1}^n G_{A_i}$ представляет собой путь длины n на графе G_A . Соответственно множество $L^n(A_1, A_2, \dots, A_n)$ всех чередующихся путей длины n на графе $\sum_{i=1}^n G_{A_i}$ в этом случае совпадает с множеством $L^n(A)$ всех путей длины n на графе G_A . Отсюда и из формулы (1.11) для матрицы $(A)^n$ получаем следующее правило.

Правило вычисления элементов матрицы $(A)^n$ на графе G_A

Матричный элемент $[A^n]_{ij}$ равен сумме вкладов всех путей длины n на графе G_A из вершины x_i в вершину x_j

$$[A^n]_{ij} = \sum_{e \in I_{ij}^{(n)}(A)} f(e), \quad (1.12)$$

где

$$f(l) = \prod_{k=0}^{n-1} a_{i_k i_{k+1}} \quad (i_0 = i, i_n = j).$$

В качестве примера рассмотрим матрицу A вида

$$A = \begin{pmatrix} 0 & 5 & 5 & 2 & 1 \\ 0 & 0 & 0 & 1 & 0 \\ 2 & 4 & 0 & 0 & 0 \\ 0 & 0 & 0 & 3 & 0 \\ 0 & 0 & 0 & 8 & 0 \end{pmatrix}$$

Если изобразить граф G_A матрицы A и применить правило (1.12), то нетрудно получить, что

$$[A^n]_{13} = \begin{cases} 0, & n = 2k \ (k = 1, 2, \dots), \\ 5 \cdot 10^n, & n = 2k + 1 \ (k = 0, 1, \dots), \end{cases} \quad (1.13)$$

$$[A^n]_{14} = \sum_{k=0}^{\left[\frac{n-2}{2}\right]} 13 \cdot 10^k 3^{n-2-2k} + \sum_{k=0}^{\left[\frac{n-1}{2}\right]} 2 \cdot 10^k 3^{n-1-2k} + \\ + \sum_{k=0}^{\left[\frac{n-3}{2}\right]} 20 \cdot 10^k 3^{n-3-2k}.$$

Здесь $\left[\frac{n-2}{2}\right]$ означает целую часть числа $\frac{n-2}{2}$.

З а м е ч а н и е 2. В случае совпадения двух матриц в упорядоченной цепочке $(A_1, A_2, \dots, A_n)$, например, матрицы A_{i_1} и A_{i_2} ($i_1, i_2 = 1, 2, \dots, n$), множество $L^n(A_1, A_2, \dots, A_n)$ всех чередующихся путей длины n на графе $\sum_{k=1}^n G_{A_k}$ можно рассматривать

на графе $\sum_{\substack{k=1 \\ k \neq i_2}}^n G_{A_k}$, но уже как множество чередующихся путей длины n с отмеченным чередованием дуг суграфа $G_{A_{i_1}} \equiv G_{A_{i_2}}$,

а именно, i_1 -я и i_2 -я дуги пути l на графе $\sum_{\substack{k=1 \\ k \neq i_2}}^n G_{A_k}$ должны принадлежать суграфу $G_{A_{i_1}} \equiv G_{A_{i_2}}$.

Для построения дискретного аналога континуального интеграла Фейнмана и его обобщения полезным будет следующее замечание.

З а м е ч а н и е 3. На множестве всех путей $L^n(A)$ длины n в графе G_A определим функционал

$$f : L^n(A) \rightarrow \mathbb{C}$$

по формуле

$$f(l) = \prod_{k=0}^{n-1} f((x_{i_k} x_{i_{k+1}})),$$

где

$$f((x_{i_k}, x_{i_{k+1}})) = a_{i_k i_{k+1}} \quad l = \{(x_{i_0}, x_{i_1}, \dots, x_{i_n}) \in L^n(A).$$

По функционалу f определим функционал Φ на любом подмножестве путей вида $l_{ij}^{(n)}(A)$, образующих разбиение множества $L^n(A)$ аддитивным образом:

$$\Phi(Ul) = \sum_{l \in l_{ij}^{(n)}(A)} f(l).$$

Из правила вычисления элементов $[A^n]_{ij}$ матрицы A^n следует, что значение функционала Φ на подмножестве $l_{ij}^{(n)}(A)$ равно матричному элементу матрицы A^n ,

$$\Phi(Ul) = \sum_{l \in l_{ij}^{(n)}(A)} f(l) = [A^n]_{ij}. \quad (1.14)$$

Рассмотрим подробнее частный, но важный для приложений случай вычисления n -й степени матрицы A , когда $A = B + C$ и коммутатор $[B, C] \neq 0$.

П р а в и л о в ы ч и с л е н и я э л е м е н т о в м а т р и ц ы $(B + C)^n$

Матричный элемент $[(B + C)^n]_{ij}$ равен сумме вкладов всех произвольно чередующихся путей l длины n на графе $G_B + G_C$ из вершин κ_i в вершину κ_j :

$$[(B + C)^n]_{ij} = \sum_{l \in \hat{l}_{ij}^{(n)}(B, C)} f(l), \quad (1.15)$$

здесь вклад $f(l)$ пути $l = (u_1, u_2, \dots, u_n)$ определяется формулой

$$f(l) = \prod_{k=0}^{n-1} f(u_k), \quad u_k = (x_{i_k}, x_{i_{k+1}}), \quad i_0 = i, \quad i_n = j, \quad (1.16)$$

где

$$f(u_k) = b_{i_k i_{k+1}}, \quad \text{если } u_k \in G_B,$$

$$f(u_k) = c_{i_k i_{k+1}}, \quad \text{если } u_k \in G_C.$$

Для обоснования правила (1.15), (1.16) представим матрицу $(B + C)^n$ в виде суммы 2^n слагаемых, каждое из кото-

рых является произведением, составленным из матриц B и C так, что общее число сомножителей в каждом слагаемом равно n :

$$\begin{aligned}(B + C)^n &= B^n + BC^{n-1} + \dots + B^{n-1}C + BC B^{n-2} + \\ &+ BC^2 B^{n-3} + \dots + BC^{n-2}B + \dots + C^n = \\ &= \sum_{k=1}^{2^n} P_k(B, C). \quad (1.17)\end{aligned}$$

Поскольку общий член $P_k(B, C)$ разложения (1.17) представляет собой произведение n матриц, то в силу правила вычисления элементов произведения (1.11) матричный элемент $[P_k(B, C)]_{ij}$ равен сумме вкладов всех чередующихся путей $l_{ij}^{(n)}(B, C, k)$ длины n на графе $\sum_{i=1}^{n_1(k)} G_B + \sum_{i=1}^{n_2(k)} G_C$, где $n_1(k)$ и $n_2(k)$ — число сомножителей в $P_k(B, C)$, равных матрице B и матрице C соответственно. Очевидно, что выполняется равенство

$$n_1(k) + n_2(k) = n \quad (k = 1, 2, \dots, 2^n).$$

В силу замечания 2 множество чередующихся путей длины n на графе $\sum_{i=1}^{n_1(k)} G_B + \sum_{i=1}^{n_2(k)} G_C$ совпадает с множеством чередующихся путей длины n на графе $G_B + G_C$ с отмеченным чередованием дуг суграфов G_B и G_C . Поэтому матричный элемент $[P_k(B, C)]_{ij}$ при любом $k=1, 2, \dots, 2^n$ определяется вкладами $f(l)$ путей длины n на графе $G_B + G_C$, причем различным k ($k=1, 2, \dots, 2^n$) отвечают пути длины n с различным чередованием дуг суграфов G_B и G_C , так как чередование дуг суграфов G_B и G_C в пути $l \in l_{ij}^{(n)}(BC)$ определяется расположением матриц B и C в произведении $P_k(B, C)$.

Из разложения (1.17) и правила вычисления элементов произведения матриц получаем требуемое равенство.

З а м е ч а н и е 4. Правило вычисления элементов матрицы $(B + C)^n$ на графе $G_B + G_C$ является той теоретико-графовой основой, на которой удастся построить диаграммную интерпретацию членов ряда теории возмущений КМ-цепи (см. ниже гл. 2).

Граф $G_B + G_C$ будем называть в дальнейшем *аддитивным графом суммы двух матриц* B и C и обозначать через $\widehat{G}_{B+C}$, т. е.

$$\widehat{G}_{B+C} = G_B + G_C.$$

Внимательный читатель заметит, что в общем случае аддитивный граф суммы двух матриц не совпадает с графом суммы этих матриц, т. е. что $\widehat{G}_{B+C} \neq G_{B+C}$.

2. Упорядоченное произведение графов. Представление произведения матриц на упорядоченном произведении графов перемножаемых матриц. Пути длины n , дающие вклады в (1.11) при достаточно большом n , вообще говоря, не являются простыми на графе G_A . Например, если в вершине x_i в графе G_A имеется петля (x_i, x_i) и $a_{ij} \neq 0$ при некотором $j \neq i$ ($i = 1, 2, \dots, N$), то при любом n существует путь длины n .

Поэтому определим теперь граф $G_A^{(n)}$, называемый *n -й степенью графа* G_A такой, что матричный элемент матрицы A^n определяется вкладом только простых путей на графе $G_A^{(n)}$.

Вначале отметим, что любому графу без кратных дуг $G = (X, \Gamma)$ ($\Gamma \neq \emptyset$) можно сопоставить двудольный граф $\tilde{G} = (\tilde{X}, \tilde{\Gamma})$, для которого множество 1) вершин $\tilde{X} = {}^1X \vee {}^2X$, 2) дуг $\tilde{\Gamma}$ определено отображением $\Gamma : X \rightarrow 2^X$.

Примером может служить граф G_1 и отвечающий ему двудольный граф $\tilde{G}_1$, изображенные на рис. 159.

Граф G является гомоморфным образом графа $\tilde{G}$ при

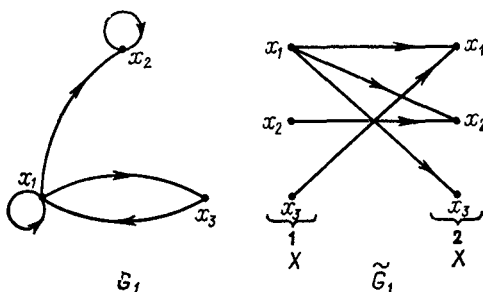


Рис. 159

гомоморфизме $\Phi^2, \Phi^2: \tilde{G} \rightarrow G$, который задается следующими двумя отображениями вершин и дуг графа:

$$\varphi: \tilde{x} \rightarrow x, \quad \varphi(x_i \in \overset{1}{X}) = x_i, \quad \varphi(x_i \in \overset{2}{X}) = x_i \quad (i=1, 2, \dots, N), \quad (1.18)$$

$$\psi: \tilde{\Gamma} \rightarrow \Gamma, \quad \psi((x_i, x_j)) = (\varphi(x_i), \varphi(x_j)) \quad (i, j=1, 2, \dots, N),$$

где $x_i \in \overset{1}{X}$, $x_j \in \overset{2}{X}$.

Геометрически гомоморфизм Φ^2 означает склейку при каждом i вершины $x_i \in \overset{1}{X}$ и вершины $x_i \in \overset{2}{X}$ в одну вершину с сохранением дуг $\tilde{\Gamma}$.

Определение 1.6. Упорядоченным произведением $G_{A_1 A_2 \dots A_n}$ графов $G_{A_1}, G_{A_2}, \dots, G_{A_n}$ матриц $A_1, A_2, \dots, A_n$ соответственно назовем граф $(\tilde{X}, \tilde{\Gamma})$, множество вершин $\tilde{X}$ и множество дуг $\tilde{\Gamma}$ которого определяется из условий (см. с. 6):

$$1) \tilde{X} = \bigvee_{i=1}^{n+1} \overset{i}{X}, \quad (\overset{i}{X} = \{i\} \times X),$$

$$2) \tilde{\Gamma} = \Gamma_{A_1} U \Gamma_{A_2} U \dots U \Gamma_{A_n}, \quad \Gamma_{A_i}: \overset{i}{X} \rightarrow 2^{\overset{i+1}{X}} \quad (i=1, 2, \dots, n).$$

Если $A_1 = A_2 = \dots = A_n = A$, то граф $G_{A_1, A_2, \dots, A_n}$ называется n -й степенью графа G_A и обозначается через $G_A^{(n)}$.

Граф $G_{A_1 A_2 A_3}$, являющийся упорядоченным произведением графов $G_1 \equiv G_{A_1}$ (см. рис. 159), $G_{A_2} \equiv G_1$ и графа $G_2 = G_{A_2}$ (рис. 160), изображен на рис. 161

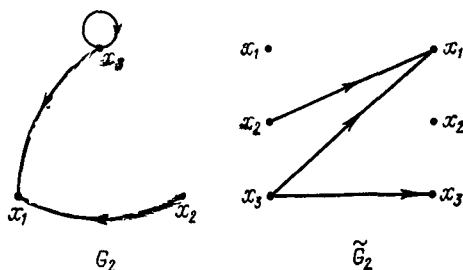


Рис. 160

Граф $\sum_{k=1}^n G_{A_k}$ является гомоморфным образом графа $G_{A_1 A_2 \dots A_n}$ при гомоморфизме Φ^n , который определяется аналогично гомоморфизму Φ^2 в (1.18), и геометрически состоит в склейке при каждом i ($i = 1, 2, \dots, N$) всех вершин $x_i \in X$,

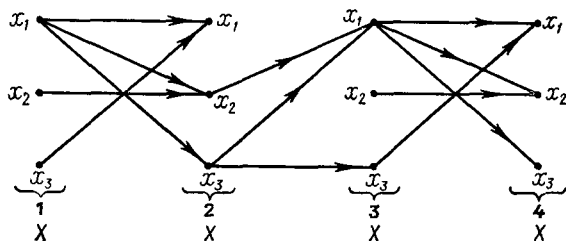


Рис. 161

$\dots, x_i \in X^{n+1}$ в одну вершину x_i графа $\sum_{k=1}^n G_{A_k}$ так, что каждой дуге $(x_i, x_j) \in \Gamma_{A_k}$ однозначно соответствует дуга (x_i, x_j) суграфа G_{A_k} в графе $\sum_{k=1}^n G_{A_k}$.

Из определения 1.6 и определения графа G_{A_k} матрицы A_k следует, что каждое отличное от нуля слагаемое суммы

$$[A_1 A_2 \dots A_n]_{ij} = \sum_{1 \leq i_1, i_2, \dots, i_{n-1} \leq N} a_{ii_1}^{(1)} a_{i_1 i_2}^{(2)} \dots a_{i_{n-1} j}^{(n)} \quad (i, j = 1, 2, \dots, N) \quad (1.19)$$

представляет собой вклад пути $l, l = \{x_i, x_{i_1}, \dots, x_{i_{n-1}}, x_j \mid x_{i_k} \in X\}$ длины n на графе $G_{A_1 A_2 \dots A_n}$.

В случае, когда $A_1 = A_2 = \dots = A_n = A$, из формулы (1.19) получаем следующее правило.

Правило вычисления элементов матрицы $(A)^n$ на графе $G_A^{(n)}$

Матричный элемент $[A^n]_{ij}$ матрицы $(A)^n$ равен сумме вкладов всех простых путей длины n из вершины x_i в вер-

шину x_j на графе $G_A^{(n)}$, являющемся n -й степенью графа G_A , т. е.

$$[A^n]_{ij} = \sum_{l \in l_{ij}^{(n)}(A^n)} f(l), \quad (1.20)$$

где $l_{ij}^{(n)}(A^n)$ обозначает множество всех путей длины n графа $G_A^{(n)}$ из вершины x_i в вершину x_j .

З а м е ч а н и е 5. Понятие графа, являющегося упорядоченным произведением графов перемножаемых матриц, и представление на нем матричных элементов произведения матриц является теоретико-графовой основой связи между КМ-цепью и дискретным аналогом континуального интеграла (интеграла по траекториям) Фейнмана.

Задача 1. Пусть σ_1 , σ_2 и σ_3 — двухрядные матрицы Паули:

$$\sigma_1 = \begin{pmatrix} 0 & 1 \\ 1 & 0 \end{pmatrix}, \quad \sigma_2 = \begin{pmatrix} 0 & -i \\ i & 0 \end{pmatrix}, \quad \sigma_3 = \begin{pmatrix} 1 & 0 \\ 0 & -1 \end{pmatrix}.$$

Вычислить, используя правило (1.20), произведения матриц

$$1) \sigma_1, \sigma_2, \sigma_3, \quad 2) (\sigma_i)^n \quad (i = 1, 2, 3), \quad \lim_{n \rightarrow \infty} \sum_{k=0}^n \frac{\sigma_i^k}{k!} \quad (i = 1, 2, 3).$$

§ 2. СУПЕРПОЗИЦИЯ АМПЛИТУД ВЕРОЯТНОСТЕЙ СОБЫТИЙ ПО ФЕЙНМАНУ И ОПРЕДЕЛЕНИЕ КОМПЛЕКСНОЙ МАРКОВСКОЙ ЦЕПИ

В этом параграфе излагаются основные положения фейнмановского подхода к вероятности в квантовой механике и способам ее вычисления, следуя работам [10], [13], [14].

Затем определяется комплексная марковская цепь (КМ-цепь), которая, в частности, может служить математической моделью вероятностного подхода Фейнмана в квантовой механике.

Определение КМ-цепи тесным образом связано с понятием графа матрицы и множеством путей на графе. Целесообразность использования понятия графа при рассмотрении дискретного аналога теории Фейнмана связана с тем, что графы матриц являются простейшей и в то же время естественной моделью, на которой реализуются правила

Фейнмана относительно сложения амплитуд вероятностей и вычисления вероятностей событий.

1. Амплитуды вероятностей и правила их сложения. В основе формализма фейнмановского подхода лежит представление об амплитуде вероятности, связанной с каким-либо движением системы в пространстве — времени. Сам Фейнман пишет об этом: «...новый, выдвигаемый квантовой механикой способ изображать мир ...состоит в том, чтобы задавать амплитуду любого события, которое может случиться. Если событие состоит в регистрации частицы, то можно задать амплитуду обнаружения этой частицы в тех или иных местах и в то или иное время. Вероятность обнаружить частицу тогда будет пропорциональна квадрату абсолютной величины амплитуды» ([13], стр. 219).

Итак, первое положение вероятностной концепции квантовой механики гласит следующее.

Вероятность $P(A)$ события* A дается квадратом абсолютной величины комплексного числа $Z(A)$, называемого *амплитудой вероятности данного события A* :

$$P(A) = |Z(A)|^2. \quad (2.1)$$

Второе положение определяет правило сложения амплитуд вероятностей.

Если событие A может произойти несколькими взаимно исключающими способами $B_1, B_2, \dots, B_k$, то амплитуда вероятности $Z(A)$ такого события является суммой амплитуд вероятностей каждого отдельного способа $Z(B_i)$ ($i = 1, 2, \dots, k$)

$$Z(A) = \sum_{i=1}^k Z(B_i). \quad (2.2)$$

Из формул (2.1) и (2.2) следует, что вероятность суммы попарно несовместимых событий не равна сумме вероятностей отдельных слагаемых, т. е.

$$\begin{aligned} P\left(\sum_{i=1}^k B_i\right) &= P(A) = |Z(A)|^2 = \left|\sum_{i=1}^k Z(B_i)\right|^2 \neq \\ &\neq \sum_{i=1}^k |Z(B_i)|^2 = \sum_{i=1}^k P(B_i). \end{aligned} \quad (2.3)$$

* Событие здесь рассматривается как совокупность начальных и конечных условий опыта. Эти условия полностью определяются путем измерения в начальный и конечный моменты времени.

В таком случае говорят, что между амплитудами вероятности отдельных событий происходит интерференция.

Следует отметить, что правило сложения амплитуд вероятности (2.2) в физике было известно задолго до появления квантовой теории. А именно, в волновой оптике явление интерференции (дифракции) света количественно описывается с помощью соотношений (2.1) и (2.3). Если имеется например, два когерентных точечных источника света B_1 и B_2 , излучающих волны с амплитудами $Z(B_1)$ и $Z(B_2)$, то интенсивность $I(B_1, B_2)$ света, создаваемая этими источниками в точке x , равна

$$I(B_1, B_2) = |Z(B_1) + Z(B_2)|^2 = I(B_1) + I(B_2) + 2\sqrt{I(B_1)I(B_2)}\cos\Delta(x),$$

где $I(B_1) = |Z(B_1)|^2$, $I(B_2) = |Z(B_2)|^2$, $\Delta(x)$ — разность фаз между $Z(B_1)$ и $Z(B_2)$. Член $2\sqrt{I(B_1)I(B_2)}\cos\Delta(x)$ обуславливает интерференцию.

Интенсивность света $I = |Z(B)|^2$ также допускает вероятностную трактовку, если обратиться к корпускулярной природе света. Тогда величина $I(x) = |Z(x)|^2$ выражает собой вероятность попадания фотона в точку x .

Вне зависимости от физического истолкования амплитуд вероятностей и правила их сложения соотношениям (2.1)—(2.3) в дискретном случае отвечает следующая абстрактная схема построения вероятностного пространства событий [14].

Пусть $\bar{\Omega} = \{A_1, A_2, \dots, A_k\}$ — произвольное конечное множество, элементы которого назовем *элементарными истинными событиями* (или просто элементарными событиями) и пусть каждое A_i само состоит из $\gamma(i)$ элементов $b_j^{(i)}$ ($j = 1, 2, \dots, \gamma(i)$). Каждый элемент $b_j^{(i)}$ назовем *элементарным виртуальным событием*, а всю совокупность их при $i = 1, 2, \dots, k$ — множеством элементарных виртуальных событий Ω , $\Omega = \bigcup_{i=1}^k \bigcup_{j=1}^{\gamma(i)} \{b_j^{(i)}\}$.

Изобразим схематически множества $\bar{\Omega}$ и Ω в виде двудольного графа $G = (\Omega \cup \bar{\Omega}, \Gamma)(\Gamma: \Omega \rightarrow 2^{\bar{\Omega}})$ (рис. 162) и припишем каждой дуге $(b_j^{(i)}, A_i)$ ($i = 1, 2, \dots, k$, $j = 1, 2, \dots, \gamma(i)$) произвольное комплексное число c_{ij} , кото-

рое назовем *амплитудой элементарного виртуального события* $b_j^{(i)}$.

Определим на совокупности всех подмножеств* $N(\Omega)$ множества Ω комплексно-значную функцию $Z: N(\Omega) \rightarrow \mathbb{C}$, принимающую на подмножестве $\{b_j^{(i)}\}$ значение

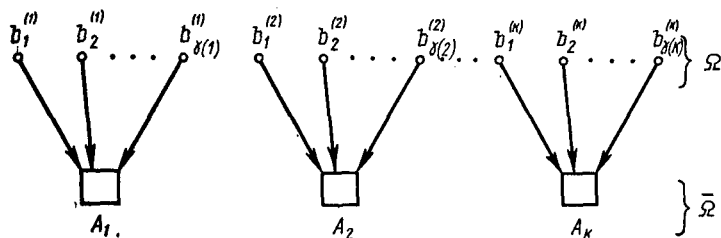


Рис. 162

$$Z(\{b_j^{(i)}\}) = c_{ij} \quad (i = 1, 2, \dots, k, \quad j = 1, 2, \dots, \gamma(i)) \quad (2.4)$$

и являющуюся аддитивной функцией на каждом из множеств A_i :

$$Z\left(\left\{\bigcup_{j=1}^{\gamma(i)} b_j^{(i)}\right\}\right) = \sum_{j=1}^{\gamma(i)} Z(\{b_j^{(i)}\}) = \sum_{j=1}^{\gamma(i)} c_{ij}. \quad (2.5)$$

Элементы $B \in N(\Omega)$ назовем *виртуальными событиями*, а значение функции $Z(B)$ — *амплитудой виртуального события* B .

Функцию $Z: N(\Omega) \rightarrow \mathbb{C}$ нормируем по отношению к множествам $A_1, A_2, \dots, A_k$, $A_i \in N(\Omega)$ так, чтобы

$$\sum_{i=1}^k |Z(A_i)|^2 = 1. \quad (2.6)$$

Тройку $(\Omega, N(\Omega), Z)$, где Ω — произвольное множество, $N(\Omega)$ — алгебра подмножества множества Ω , $Z(B)$ — амплитуда, определенная на $N(\Omega)$ и удовлетворяющая условиям

* Совокупность всех подмножеств $N(\Omega)$ является алгеброй множеств. Последнее означает, что само $\Omega \in N(\Omega)$ объединение, пересечение и разность двух множеств из $N(\Omega)$ опять принадлежит Ω .

(2.5) и (2.6), назовем *пространством виртуальных событий с комплексной амплитудой* Z .

Рассмотрим совокупность всех подмножеств $M(\bar{\Omega})$ множества элементарных истинных событий $\bar{\Omega} = \{A_1, A_2, \dots, A_k\}$. Элементы $A \in M(\bar{\Omega})$ называются *истинными событиями*.

Поскольку элементарные истинные события A_i сами состоят из элементарных виртуальных событий $\{b_j^{(i)}\}$, $A_i = \bigcup_{j=1}^{r(i)} b_j^{(i)}$, то $A_i \in N(\Omega)$, и поэтому любое множество

$A \in M(\bar{\Omega})$ можно рассматривать как элемент из алгебры $N(\Omega)$. Тогда комплексная амплитуда Z , определенная на алгебре виртуальных событий $N(\Omega)$, индуцирует на алгебре истинных событий $N(\Omega)$ функцию

$$P(A \in M(\bar{\Omega})) = |Z(A \in N(\Omega))|^2. \quad (2.7)$$

Значение $P(A)$ будем называть *вероятностью истинного события* A , а величину $Z(A)$ — *амплитудой вероятности события* A .

Как следует из формул (2.4)—(2.7), функция $Z : N(\Omega) \rightarrow \mathbb{C}$ удовлетворяет правилам Фейнмана (2.1) и (2.2) относительно сложения амплитуд вероятностей и вычисления вероятностей истинных событий A_i ($i = 1, 2, \dots, k$).

Описанная выше абстрактная схема правил сложения и вычисления амплитуд вероятностей событий может быть естественным образом реализована на графах, являющихся графами матриц.

Действительно, пусть $Z = ((z_{ij}))_{N \times N}$ — неприводимая унитарная матрица с комплексными коэффициентами и пусть G_Z — граф матрицы Z . Граф матрицы G_Z связный граф без кратных дуг. Рассмотрим на графе G_Z множество $L^n(Z)$ всех путей длины n ,

$$L^n(Z) = \bigcup_{1 \leq n, j_0 j_n \leq N} l_{j_0 j_n}^{(n)}(Z), \quad (2.8)$$

где $l_{j_0 j_n}^{(n)}(Z)$ — множество путей длины n из вершины x_{j_0} , в вершину x_{j_n} , или множество путей длины n с «закрепленными концами» x_{j_0} и x_{j_n} . Напомним, что через $[Z^n]_{j_0 j_n}$ обозначаются элементы матрицы Z^n .

В соответствии с абстрактной схемой сложения амплитуд вероятности проведем следующее сопоставление:

- 1) элементарное виртуальное событие $\{b_j^{(i)}\}$ $\longleftrightarrow$ путь l длины n на графе G_Z с закрепленными концами x_{j_0} и x_{j_n}
- 2) множество Ω элементарных виртуальных событий $\longleftrightarrow$ множество всех путей $L^n(Z)$ на графе G_Z
- 3) элементарное истинное событие $A_i = \bigcup_{j=1, 2, \dots, k}^{j^{(i)}} \{b_j^{(i)}\}$ $\longleftrightarrow$ множество $l_{j_0 j_n}^{(n)}(Z)$ путей длины n с закрепленными концами x_{j_0} и x_{j_n}
- 4) амплитуда элементарного виртуального события $Z(\{b_j^{(i)}\})$ $\longleftrightarrow$ вклад $f(l)$ пути $l \in l_{j_0 j_n}^{(n)}(Z)$, умноженный на величину c , зависящую от j_0 и N , $c = c(j_0, N)$
- 5) амплитуда вероятности истинного события $Z(A_i)$ $\longleftrightarrow$ матричный элемент $[Z^n]_{j_0 j_n}$ матрицы Z^n , умноженный на $c = c(j_0, N)$

Величина $c = c(j_0, N)$ определяется из условия нормировки (2.6). Так как матрица Z унитарна, то

$$\sum_{j_n=1}^N |[Z^n]_{j_0 j_n}|^2 = 1 \quad (j_0 = 1, 2, \dots, N). \quad (2.9)$$

Отсюда следует, что для выполнения условия нормировки (2.6) необходимо выбрать числа $c(j_0, N)$ ($j_0 = 1, 2, \dots, N$) таким образом, чтобы выполнялось условие

$$\sum_{i=1}^N |c(j_0, N)|^2 = 1. \quad (2.10)$$

В силу замечания 3 в § 1 любая функция $Z: B(L^n(Z)) \rightarrow \mathbb{C}$, значение которой на элементах $b_j^{(i)} \longleftrightarrow l \in l_{j_0 j_n}^{(n)}(Z)$, и $A_i = \bigcup_{j=1}^{j^{(i)}} \{b_j^{(i)}\} \longleftrightarrow l_{j_0 j_n}^{(n)}(Z)$ определено равенствами

$$Z(\{b_j^{(i)}\}) \leftrightarrow l = c(j_0, N) f(l), \quad l \in l_{j_0 j_n}^{(n)}(Z), \quad (2.11 \text{ а})$$

$$Z\left(\bigcup_j \{b_j^{(i)}\}\right) = Z(A_i) = c(j_0, N) [Z^n]_{j_0 j_n}, \quad (2.11 \text{ б})$$

удовлетворяет условию аддитивности (2.5) на подмножествах A_i ($i = 1, 2, \dots, k$), являющихся элементарными истинными событиями.

Таким образом, с точностью до набора чисел $\{c(j_0, N), j_0 = 1, 2, \dots, N\}$ из (2.10) унитарная матрица Z , граф матрицы G_Z и множество путей $L^n(Z)$ на графе G по формулам (2.11 а) и (2.11 б) определяют на алгебре $N(L^n(Z))$ всех путей длины n функцию $Z : N(\Omega) \rightarrow c$, удовлетворяющую правилам вычисления амплитуд вероятностей (2.1) и (2.2).

Само понятие вероятности в квантовой механике не изменяется по сравнению с обычной теорией вероятностей. Радикально изменяется лишь способ вычисления вероятности события и понятие самого события, вероятность которого вычисляется.

Как уже отмечалось (см. сноску на стр. 310), понятие «событие» в квантовой механике в конечном итоге неразрывно связано с процессом измерения в начальный и конечный моменты времени. Поэтому если мы говорим о вероятности какого-либо события в квантовой механике, например, о том, что квантовая система занимает в данный момент определенное положение, то неявно предполагается проведение измерений положения системы в этот же момент времени.

В формальной схеме, рассмотренной выше, такие события («снабженные» измерением) назывались истинными событиями. Вероятность истинных событий вычисляется по формуле (2.1).

Если измерения над системой в данный момент времени не производятся, то вопрос о вероятности «события» найти систему в некотором положении в этот момент времени является не корректным. Подобные «события» были определены выше как виртуальные события. Для виртуальных событий существует лишь комплексная амплитуда события $Z(B)$, которая «становится» амплитудой вероятности события и определяет вероятность этого события по правилу (2.1), если производится измерение.

Подробнее о связи измерения в квантовой механике и вычислением вероятностей событий см. в [13].

Заметим, что правила сложения амплитуд вероятностей допускают большой произвол в выборе функции

$$Z : N(L^n(Z)) \rightarrow c.$$

Как уже отмечалось выше, любая комплексно-значная функция $Z(B)$, удовлетворяющая соотношениям (2.11),

обеспечивает выполнение условий (2.1), (2.2). Поэтому для проведения последовательного анализа на пространстве виртуальных событий $(L^n(Z), N(L^n(Z)))$ и получения содержательных результатов потребуем, чтобы функция $Z: N(L^n(Z)) \rightarrow \mathbb{C}$ была аддитивной функцией на любом подмножестве $\Theta_{i_\alpha i_\beta} \in N(L^n(Z))$ ($\alpha, \beta = 0, 1, 2, \dots, n$) множества путей $L^n(Z)$ с двумя «закрепленными вершинами» x_{i_α} и x_{i_β} .

Точный смысл этого условия состоит в том, что

$$Z\left(\bigcup_{l \in \Theta_{i_\alpha i_\beta}} l\right) = \sum_{l \in \Theta_{i_\alpha i_\beta}} Z(l), \quad (2.12)$$

где l — произвольный путь длины n , проходящей через вершины графа G_Z x_{i_α} и x_{i_β} «в моменты времени t_α и t_β ». Последнее означает, что если записать путь l длины n двумя способами, указывая один раз последовательность дуг, составляющих путь,

$$l = \{u_1, u_2, \dots, u_\alpha, \dots, u_\beta, \dots, u_n\},$$

а другой раз последовательность проходимых вершин вдоль пути l ,

$$l = \{x_{i_0}, x_{i_1}, \dots, x_{i_\alpha}, \dots, x_{i_\beta}, \dots, x_{i_n}\},$$

то тогда вершина x_{i_α} является концом дуги u_α , а вершина x_{i_β} концом дуги u_β . Отметим, что множество путей $\Theta_{i_\alpha i_\beta}$ при $\alpha = 0, \beta = n$ ранее обозначалось через $l_{i_0 i_n}^{(n)}(Z)$.

Определим на множестве $N(\Omega)$, где $\Omega = L^n(Z)$, условную амплитуду $Z(B_1/B_2)$, $B_1, B_2 \in N(\Omega)$.

Определение 2.1. Условная амплитуда $Z(B_1/B_2)$ события B_1 при наступлении события B_2 определяется как отношение

$$Z(B_1/B_2) = \frac{Z(B_1 \cdot B_2)}{Z(B_2)}, \quad (2.13)$$

где $B_1 \cdot B_2$ — пересечение событий B_1 и B_2

2. Определение комплексной марковской цепи. Предварительно кратко рассмотрим вещественный случай [15]. Вещественная дискретная цепь Маркова (ВМ-цепь) с конечным числом состояний определяется как последовательность случайных величин $\{\xi_k, k = 0, 1, \dots, n\}$ такая, что распределение значений случайной величины ξ_k полностью

определяется распределением значений случайной величины ξ_{k-1} . Уточним это определение.

Обозначим через $x_1, x_2, \dots, x_N$ ($x_i \in \mathbb{R}^1$) возможные значения случайной величины ξ_k при каждом $k = 0, 1, \dots, n$.

Определение 2.2. Последовательность случайных величин $\{\xi_k, k = 0, 1, \dots, n\}$ является вещественной цепью Маркова, если условное распределение величины ξ_k при фиксированных значениях величин $\xi_{k-1} = x_{i_{k-1}}, \dots, \xi_1 = x_{i_1}, \xi_0 = x_{i_0}$ совпадает с условным распределением ξ_k при условии, что $\xi_{k-1} = x_{i_{k-1}}$, т. е.

$$P(\xi_k = x_{i_k} / \xi_{k-1} = x_{i_{k-1}}, \dots, \xi_0 = x_{i_0}) = \\ = P(\xi_k = x_{i_k} / \xi_{k-1} = x_{i_{k-1}}). \quad (2.14)$$

Определение 2.3. Матрица $R(k-1, k) = ((P(\xi_k = x_{i_k} / \xi_{k-1} = x_{i_{k-1}})))$ называется матрицей перехода цепи Маркова на k -м шаге. Цепь Маркова называется однородной, если матрица перехода $R(k-1, k)$ не зависит от k^* .

Матричные элементы матрицы перехода $R = ((p_{i_{k-1} i_k}))$ удовлетворяют условиям $(p_{i_{k-1} i_k} = P(\xi_k = x_{i_k} / \xi_{k-1} = x_{i_{k-1}}))$

$$0 \leq p_{i_{k-1} i_k} \leq 1 \quad (i_k, i_{k-1} = 1, 2, \dots, N), \quad (2.15)$$

$$\sum_{i_k=1}^N p_{i_{k-1} i_k} = 1 \quad (i_{k-1} = 1, 2, \dots, N).$$

Марковские цепи являются адекватной моделью для описания следующей часто встречающейся ситуации. Предположим, что некоторая физическая система S , которая в любой момент времени t может находиться в одном из состояний, число которых конечно, переходит из одного состояния в другое в дискретные моменты времени t_k ($k = 0, 1, \dots, n$). Обозначим возможные состояния системы S через $x_1, x_2, \dots, x_N$. Определим случайные величины $\{\xi_k, k = 0, 1, \dots, n\}$ так, что событие $\{\xi_k = x_{i_k} (i_k = 1, 2, \dots, N, k = 0, 1, \dots, n), x_{i_k} \in \mathbb{R}^1\}$ означает нахождение системы S в момент t_k в состоянии x_{i_k} . Событие $\{\xi_k = x_{i_k} x_{i_k} \in \mathbb{R}^1\}$ для удобства обозначения будем записывать в виде (t_k, x_{i_k}) . Тогда элементы матрицы R рассматриваются как вероят-

* Всякую ниже, если не оговорено особо, рассматриваются однородные марковские цепи.

ности перехода системы S за время $\tau = t_k - t_{k-1}$ из одного состояния в другое, причем $p_{i_{k-1} i_k}$ — вероятность перехода системы S из состояния $(t_{k-1}, x_{i_{k-1}})$ в состояние (t_k, x_{i_k}) .

При исследовании процессов, описываемых марковскими цепями, рассматривается следующая задача. По заданному в начальный момент $t = 0$ распределению случайной величины ξ_0 (или по заданным начальным вероятностям $p_{i_0}^{(0)}$ ($i_0 = 1, 2, \dots, N$)) обнаружить систему S в состоянии x_{i_0} требуется найти вероятность $P(t_n, x_{i_n})$ того, что система в момент времени t_n находится в состоянии x_{i_n} . Из формулы полной вероятности следует, что

$$P(t_n, x_{i_n}) = \sum_{i_0=1}^N p_{i_0}^{(0)} P(t_n, x_{i_n} / t_0, x_{i_0}). \quad (2.16)$$

Условную вероятность $P(t_n, x_{i_n} / t_0, x_{i_0})$ называют *вероятностью перехода системы S из состояния x_{i_0} в момент t_0 в состояние x_{i_n} в момент t_n* . Из условия несовместимости состояний системы S и однородности цепи Маркова следует, что матрица перехода за n шагов записывается в виде

$$((P(t_n, x_{i_n} / t_0, x_{i_0}))) = (R)^n, \quad (2.17)$$

где $R = ((p_{i_{k-1} i_k}))$ — матрица перехода за один шаг.

Рассмотренная конструкция вещественной марковской цепи при переходе к пределу при $n \rightarrow \infty$ и увеличении числа возможных состояний $N \rightarrow \infty$ (если такой предел существует) дает описание непрерывных систем классического типа, например, описание движения броуновских частиц.

Для волновых систем, подобных системе нерелятивистской квантовой механики, может быть получен [5] соответствующий предельный переход, если использовать обобщение вещественных марковских цепей — комплексные марковские цепи.

По существу КМ-цепь является вероятностным дискретным аналогом процессов квантовой механики, в которых с одной стороны дальнейшее развитие процесса определяется состоянием системы в настоящий момент времени и не зависит от способа, которым это состояние достигнуто, а с другой — если над системой не производится измерений

в данный момент времени, то положение системы не определено, и она может находиться одновременно в нескольких из возможных состояний.

Указанные обстоятельства совместно с понятием амплитуд вероятностей событий и правилами их сложения находят точное математическое выражение в конструкции КМ-цели к определению которой и перейдем.

Пусть S — некоторая система, которая в любой момент времени $t (0 \leq t \leq T)$ может находиться в любом из своих возможных состояний $x_1, x_2, \dots, x_N$. Возможные состояния системы S отождествим с вершинами $x_1, x_2, \dots, x_N$ графа G_Z — графа унитарной неприводимой матрицы Z . Положим $t_k = k\tau$, где $\tau = \frac{T}{n} (k = 0, 1, \dots, n)$. Напомним, что путь l длины n проходит в момент t_k через вершину x_{j_k} , если k -я дуга пути l имеет своим концом вершину x_{j_k} .

Определение 2.4. *Виртуальным состоянием системы S в момент времени t_k назовем множество всех путей длины n на графе G_Z , проходящих в момент t_k через вершину x_{j_k} и будем обозначать через*

$$x_{j_k} (k = 0, 1, \dots, n; \quad j_k = 1, 2, \dots, N).$$

По аналогии с вещественным случаем определим условную амплитуду перехода системы S из одного виртуального состояния в другое за конечное число шагов $m (m \leq n)$.

Определение 2.5. *Условная амплитуда перехода $Z(x_{j_{k+m}} / x_{j_k})$ системы S из виртуального состояния $B_2 = x_{j_k}$ в виртуальное состояние $B_1 = x_{j_{k+m}}$ определяется формулой*

$$Z(x_{j_{k+m}} / x_{j_k}) = [Z^m]_{j_k \ j_{k+m}}. \quad (2.18)$$

Теорема 2.1. *Условная амплитуда того, что система находится в виртуальном состоянии x_{j_k} при условии, что в моменты времени $t_0, t_1, \dots, t_{k-1}$ она находилась в виртуальных состояниях $x_{j_0}, x_{j_1}, \dots, x_{j_{k-1}}$ равна условной амплитуде того, что система S находится в момент t_k в состоянии x_{j_k} при условии, что в момент времени t_{k-1} она находилась в состоянии $x_{j_{k-1}}$, т. е.*

$$Z(x_{j_k} / x_{j_{k-1}}, \dots, x_{j_0}) = Z(x_{j_k} / x_{j_{k-1}})^*. \quad (2.19)$$

* Левая часть равенства определена в силу свойства амплитуды (2.12) и определения условной амплитуды (2.13).

Доказательство. Для простоты возьмем $k = 2$. Для случая $k > 2$ доказательство производится аналогичным образом.

Из определения 2.1 имеем:

$$Z(x_{j_2} / x_{j_1} x_{j_0}) = \frac{Z(x_{j_2} x_{j_1} x_{j_0})}{Z(x_{j_1} x_{j_0})} = \frac{Z(x_{j_2} x_{j_1})}{Z(x_{j_2} x_{j_1})} \times \\ \times \frac{Z(x_{j_2} x_{j_1} x_{j_0})}{Z(x_{j_1} x_{j_0})} = \frac{Z(x_{j_2} x_{j_1})}{Z(x_{j_1})} \frac{Z(x_{j_2} x_{j_1} x_{j_0})}{Z(x_{j_2} x_{j_1})} \frac{Z(x_{j_1})}{Z(x_{j_1} x_{j_0})}.$$

Так как при $k = 2$ правая часть формулы (2.19) имеет вид

$$Z(x_{j_2} / x_{j_1}) = \frac{Z(x_{j_2} x_{j_1})}{Z(x_{j_1})},$$

то для доказательства утверждения теоремы при $k = 2$ достаточно показать, что выполняется равенство

$$\frac{Z(x_{j_2} x_{j_1} x_{j_0})}{Z(x_{j_2} x_{j_1})} \frac{Z(x_{j_1})}{Z(x_{j_1} x_{j_0})} = 1. \quad (2.20)$$

Рассмотрим событие $x_{j_1} x_{j_0} \in B(L^n(Z))$, являющееся объединением элементарных виртуальных событий — путей длины n с закрепленными двумя вершинами x_{j_0} и x_{j_1} ,

$$x_{j_1} \cdot x_{j_0} = \bigcup_{1 \leq j_2, j_3, \dots, j_n \leq N} x_{j_0} \cdot x_{j_1} \cdot x_{j_2} \dots x_{j_n}. \quad (2.21)$$

Из свойства (2.12) аддитивности амплитуды $Z: B(L^n(Z)) \rightarrow \mathbb{C}$ на подмножествах путей указанного вида и формулы (2.11a) из равенства (2.21) следует, что

$$Z(x_{j_1} \cdot x_{j_0}) = \sum_{1 \leq j_2, j_3, \dots, j_n \leq N} Z(x_{j_0} x_{j_1} x_{j_2} \dots x_{j_n}) = \\ = \sum_{1 \leq j_2, j_3, \dots, j_n \leq N} c(j_0, N) Z_{j_0 j_1} Z_{j_1 j_2} \dots Z_{j_{n-1} j_n} = \\ = c(j_0, N) Z_{j_0 j_1} \sum_{1 \leq j_2, j_3, \dots, j_n \leq N} Z_{j_1 j_2} \dots Z_{j_{n-1} j_n}.$$

Аналогичным образом вычисляем безусловные амплитуды событий

$$x_{j_2} x_{j_1}, x_{j_2} x_{j_1} x_{j_0}, x_{j_1}:$$

$$Z(x_{j_2} x_{j_1}) = Z_{j_1 j_2} \sum_{1 \leq j_0 \leq N} c(j_0, N) Z_{j_0 j_1} \times \\ \times \left\{ \sum_{1 \leq j_3, \dots, j_n \leq N} Z_{j_2 j_3} \dots Z_{j_{n-1} j_n} \right\},$$

$$Z(x_{j_2} x_{j_1} x_{j_0}) = \sum_{1 \leq j_3, \dots, j_n \leq N} c(j_0, N) Z_{j_0 j_1} Z_{j_1 j_2} \dots Z_{j_{n-1} j_n} = \\ = c(j_0, N) Z_{j_0 j_1} Z_{j_1 j_2} \left\{ \sum_{1 \leq j_3, \dots, j_n \leq N} Z_{j_2 j_3} \dots Z_{j_{n-1} j_n} \right\},$$

$$Z(x_{j_1}) = \sum_{1 \leq j_0, j_2, \dots, j_n \leq N} c(j_0, N) Z_{j_0 j_1} Z_{j_1 j_2} \dots Z_{j_{n-1} j_n} = \\ = \sum_{1 \leq j_0 \leq N} c(j_0, N) Z_{j_0 j_1} \left\{ \sum_{1 \leq j_2, \dots, j_n \leq N} Z_{j_1 j_2} \dots Z_{j_{n-1} j_n} \right\}.$$

Подставляя вычисленные безусловные амплитуды событий в левую часть (2.20), получаем требуемое равенство. ■

Теорема 2.1 выражает условие, аналогичное условию марковости вещественной цепи случайных величин.

Предыдущие рассуждения приводят к следующему определению комплексной марковской цепи.

Определение 2.4. Пусть Z — неприводимая унитарная матрица с комплексными элементами, G_Z — граф матрицы Z , $L^n(Z)$ — множество путей длины n на графе G_Z ; пусть случайная величина принимает значения $1, 2, \dots, N$ при каждом $k = 0, 1, \dots, n$. Значению случайной величины $\{v_k = j_k, k = 0, 1, \dots, n, j_k = 1, \dots, N\}$ на пространстве виртуальных событий $(L^n(Z), B(L^n(Z)), Z)$ с комплексной амплитудой Z сопоставим виртуальное состояние системы x_{j_k} .

Последовательность случайных величин $\{v_k, k = 0, 1, \dots, n\}$ называется однородной комплексной марковской цепью, отвечающей матрице Z порядка $N \times N$, если выполнено условие:

матрица условной амплитуды перехода за m шагов определяется формулой

$$((Z(x_{j_{k+m}} / x_{j_k}))) = (Z)^m_{j_k j_{k+m}}. \quad (2.22)$$

Таким образом, КМ-цепь однозначно определена, если задана совокупность объектов: унитарная матрица Z и граф матрицы G_Z . В этом случае будем обозначать КМ-цепь через (Z, G_Z) .

Заметим, что понятие виртуального состояния системы как множества всех путей на графе G_Z , проходящих через фиксированную вершину графа в заданный «момент времени», является принципиальным моментом в приведенной выше конструкции КМ-цепи и соответствует физическому смыслу понятия состояния квантовой системы, над которой не производят измерения.

Рассмотрим виртуальные состояния x_{j_k} системы в момент времени t_k ($k = 1, \dots, n-1$). Аналогично вещественному случаю виртуальные состояния x_{j_k} ($j_k = 1, \dots, N$) образуют разложения вероятностного пространства Ω на попарно несовместимые виртуальные события

$$\Omega = x_1 + x_2 + \dots + x_N, \quad (2.23)$$

$$x_{j_k} x'_{j'_k} = \phi, \quad j_k \neq j'_k.$$

При $t_k = t_0 = 0$ виртуальное состояние x_{j_0} системы S — множество путей длины n , выходящих из вершины графа x_{j_0} , отождествим с состоянием x_{j_0} системы S в начальный момент $t_{j_0} = 0$. Тогда в любой другой момент t_k ($k \neq 0$), система S , находившаяся в момент $t_0 = 0$ в состоянии x_{j_0} , представляется «пучком» путей x_{j_0} на граф G_Z , и поскольку пересечение множеств путей x_{j_0} и x_{j_k} ($j_k = 1, 2, \dots, N$) не пусто, то это означает, что система S в момент времени t_k находится сразу в нескольких попарно несовместимых виртуальных состояниях x_{j_k} ($j_k = 1, 2, \dots, N$).

В момент времени $t=t_n$ пересечение $x_{j_0} x_{j_n}$ виртуальных состояний системы S представляет собой множество путей длины n с закрепленными концами x_{j_0} и x_{j_n} и поэтому в силу соответствия (см. п. 1) является истинным элементарным событием A_i :

$$A_i = x_{j_0} x_{j_n}. \quad (2.24)$$

Учитывая определение 2.1, из (2.24) следует, что безусловная амплитуда события A_i равна

$$Z(A_i) = Z(x_{j_0} x_{j_n}) = Z(x_{j_n} / x_{j_0}) Z(x_{j_0}).$$

Отсюда, с учетом определения (2.4) условной амплитуды, получаем, что

$$Z(A_i) = [Z^n]_{j_0 j_n} Z(x_{j_0}). \quad (2.25)$$

Сравнивая формулу (2.25) и (2.116) приходим к следующей интерпретации величин $c(j_0, N)$, фигурирующих в схеме сложения амплитуд вероятности. Величина $c(j_0, N)$ ($j_0 = 1, 2, \dots, N$) представляет собой амплитуду события—обнаружить систему S в виртуальном состоянии x_{j_0} или в силу проведенного выше отождествления — в состоянии x_{i_0} в момент $t = t_0 = 0$. Иными словами, величины $(c(j_0, N), j_0 = 1, 2, \dots, N)$ в совокупности задают начальное распределение амплитуды вероятности

$$Z(x_{j_0}) = c(j_0, N) \quad (j_0 = 1, 2, \dots, N),$$

причем вероятность $p_{j_0}^{(0)}$ системе S находиться в момент $t_0 = 0$ в состоянии x_{j_0} , равна

$$p_{j_0}^{(0)} = |Z(x_{j_0})|^2 = |c(j_0, N)|^2.$$

В силу условия (2.10) для начальных вероятностей $p_{i_0}^{(0)}$ выполняется равенство

$$\sum_{j_0=1}^N p_{j_0}^{(0)} = 1,$$

аналогичное условию нормировки начального распределения вероятностей для ВМ-цепи.

Выберем величину $c(i_0, N)$ из условия, что

$$c(j_0, N) = \delta_{i_0 j_0} = \begin{cases} 1, & i_0 = j_0, \\ 0, & i_0 \neq j_0. \end{cases} \quad i_0 = 1, 2 \dots N \quad (2.26)$$

Тогда с вероятностью $p_{i_0}^{(0)} = 1$ система S в момент $t_0 = 0$ будет находиться в состоянии x_{i_0} и в силу равенства (2.25) амплитуда истинного события A_i совпадает с условной амплитудой перехода системы из начального состояния системы x_{i_0} в конечное x_{j_n} .

Определение 2.5. Условную амплитуду $Z(x_{j_n} / x_{j_0})$ будем называть амплитудой Грина КМ-цепи и обозначать через $K_{j_0 j_n}^{(n)}$.

Из определения 2.4 следует, что

$$K_{j_0 j_n}^{(n)} = [Z^n]_{j_0 j_n} \quad (j_0, j_n = 1, 2, \dots, N). \quad (2.27)$$

Амплитуда Грина $K_{j_0 j_n}^{(n)}$ является амплитудой вероятности перехода системы S из состояния x_{j_0} в x_{j_n} за n шагов. В соответствии с правилами (2.1—2.2) вероятность $p_{j_0 j_n}$ такого перехода определяется формулой

$$p_{j_0 j_n} = |K_{j_0 j_n}^{(n)}|^2.$$

Отсюда и из условия унитарности матрицы Z получаем классические условия на вероятности перехода:

$$\sum_{j_n=1}^N p_{j_0 j_n} = 1 \quad (j_0 = 1, 2, \dots, N). \quad (2.28)$$

3. Связь амплитуды Грина КМ-цепи с дискретным аналогом интеграла Фейнмана по траекториям. Рассмотрим дискретную квантовую систему S , состояния $x_i (i=1, 2, \dots, N)$ которой представляют собой координаты квантового объекта, например, электрона на прямой, отстоящие друг от друга на расстоянии h , $h = \frac{x}{N}$, $x_i = ih (i=1, 2, \dots, N)$.

Движение системы S в дискретном времени $t_k (k=1, 2, \dots, n)$ описывается комплексной марковской цепью (Z, G_Z) .

Если в момент $t=0$ электрон находился в точке x_{j_0} , то событие — обнаружить его в точке x_{j_n} в момент $t=t_n$ представляется множеством путей длины n на графе G_Z из вершины x_{j_0} в вершину x_{j_n} . Амплитуда такого события — амплитуда Грина $K_{j_0 j_n}^{(n)}$ КМ-цепи в силу формулы (2.27) и правила вычисления матричных элементов матрицы Z^n на графе G_Z (см. 1.12) равна сумме вкладов всевозможных способов осуществления события

$$K_{j_0 j_n}^{(n)} = [Z^n]_{j_0 j_n} = \sum_{l \in l_{j_0 j_n}^{(n)}(Z)} f(l), \quad (2.29)$$

где l — путь длины n на графе G_Z из вершины x_{j_0} в вершину x_{j_n} .

Используя результаты § 1, перейдем к представлению n -й степени матрицы Z (амплитуды Грина КМ-цепи) на графе $G_Z^{(n)}$, являющемся n -й степенью графа G_Z (см. определение 1.6).

Граф $G_Z^{(n)}$ представляет собой «развертку» в дискретном времени $t_k = k\tau$ ($k = 0, 1, \dots, n$) графа G_Z так, что любой

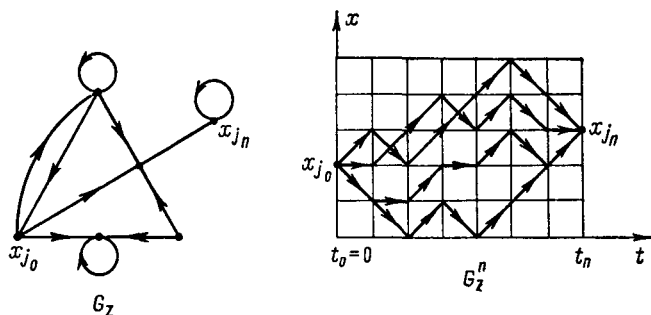


Рис. 163

путь l графа G_Z изображается в виде простого пути l на графе $G_Z^{(n)}$, причем движение вдоль пути l происходит в направлении возрастания времени t_k ($k = 0, 1, \dots, n$).

На рис. 163 изображен граф G_Z , $G_Z^{(n)}$ ($n = 7$) и ряд путей длины $n = 7$ из вершины x_{j_0} в вершину x_{j_n} . Граф $G_Z^{(n)}$ естественно назвать *дискретным пространством-временем* (x, t) , в котором происходит движение системы S , определяемое КМ-цепью (Z, G_Z) . Тогда пути l на графе $G_Z^{(n)}$ — это дискретные траектории движения системы S в пространстве-времени, а множество $l_{j_0 j_n}^{(n)}(Z^n)$ — множество всевозможных траекторий, выходящих из точки x_{j_0} в момент $t_0 = 0$ и приходящих в точку x_{j_n} в момент $t = t_n$. В силу (1.20) для амплитуды Грина КМ-цепи $K_{j_0 j_n}^{(n)} = [Z^n]_{j_0 j_n}$ справедливо равенство

$$K_{j_0 j_n}^{(n)} = \sum_{l \in l_{j_0 j_n}^{(n)}(Z^n)} f(l), \quad (2.30)$$

где $\bar{l}$ — траектория движения системы из x_{j_0} в x_{j_n} . Отсюда следует, что амплитуда Грина $K_{j_0 j_n}^{(n)}$ КМ-цепи (Z, G_Z) — это функция, которая зависит от координат начального x_{j_0} и конечного x_{j_n} положений системы, причем значение этой функции определяется вкладом всех траекторий, допустимых для данной системы матрицей Z — вкладом всех «интерферирующих альтернатив движения электрона» из точки x_{j_0} в точку x_{j_n} по Р. Фейнману [14].

Представление (2.30) в виде суммы $\sum_{l \in l_{ij}^n(Z^n)} f(l)$ по траек-

ториям на графе $G_Z^{(n)}$ можно считать точным дискретным аналогом представления решения уравнений квантовой механики в виде так называемого *континуального интеграла*, или «интеграла по траекториям» [14].

При $\hbar \rightarrow 0$ и $\tau = \tau(\hbar) \rightarrow 0$ сумма по траекториям, стоящая в правой части равенства (2.30), превращается в «интеграл» по всем непрерывным траекториям $x(t)$, $x(0) = x_{j_0}$, $x(t = t_n) = x_{j_n}$, от вклада траектории $f(l)$ [4]. Поэтому если существует предел $K(x, t)$ функции Грина $K_{j_0 j_n}^{(n)}(\tau, \hbar)$ КМ-цепи (Z, G_Z) при $\hbar \rightarrow 0$ и $\tau = \tau(\hbar) \rightarrow 0$ ($\hbar j_n \rightarrow x$, $t_n \rightarrow t$), то этот предел $K(x, t)$ можно принять за определение «интеграла по траекториям».

Пример существования интеграла по траекториям в указанном смысле для уравнения

$$\frac{\partial u}{\partial t} = c \frac{\partial u}{\partial x},$$

где c — скорость света, рассматривается в следующем параграфе.

§ 3. КМ-ЦЕПЬ ФОТОНА

Рассмотрим простейшее модельное уравнение*, описывающее динамику фотона,

* Авторы сознательно рассматривают тривиальный пример линейного дифференциального уравнения первого порядка с постоянными коэффициентами и соответствующую этому уравнению разностную схему, применяя достаточно сложную технику операторного исчисления, построенного в [6]. Тем самым читателю предоставляется возможность на примере, допускающем точное решение, познакомиться с общими актуальными методами решения

$$\frac{\partial u}{\partial t} = c \frac{\partial u}{\partial x} \quad (-\infty < x < +\infty, \quad 0 \leq t < +\infty, \\ c = \text{const}) \quad (3.1)$$

и построим КМ-цепь (Z, G_Z) , рассматривая разностную аппроксимацию этого уравнения на равномерной сетке $\Omega_{\tau, h}$ с шагами τ и h по координатам x и t соответственно:

$$\Omega_{\tau, h} = \{x_j, t_k \mid x_j = jh, \quad j = 0, 1, \dots, N; \quad t_k = k\tau, \\ k = 0, 1, \dots, n\}.$$

С помощью операторных методов [6] найдем матрицу Z^n , определяющую амплитуду Грина $K_{j_0 j_n}^{(n)}(\tau, h)$ ($j_0, j_n = 1, 2, \dots, N$) КМ-цепи (Z, G_Z) и докажем сходимость амплитуды $K_{j_0 j_n}^{(n)}(\tau, h)$ к функции Грина уравнения (3.1) при $h \rightarrow 0$ и $\tau(h) \rightarrow 0$.

Решение уравнения (3.1) будем искать в классе 2π -периодических по $x \in \mathbb{R}^1$ функций $u(x, t)$

$$u(x + 2\pi, t) = u(x, t), \quad (3.2)$$

задавая при $t = 0$ начальное условие

$$u(x, 0) = u_0(x) \quad (0 \leq x \leq 2\pi). \quad (3.3)$$

Легко проверить, что функция

$$u(x, t) = \tilde{u}_0(x + ct),$$

где $\tilde{u}_0(x)$ — периодическое продолжение функции $u_0(x)$ на всю прямую $\mathbb{R}^1$ является точным решением задачи (3.1) — (3.3).

Найдем приближенное решение поставленной задачи Коши (3.1), (3.3) с краевыми условиями (3.2) на отрезке $0 \leq t \leq T$:

$$\begin{cases} \frac{\partial u}{\partial t} = c \frac{\partial u}{\partial x}, \\ u(0, t) = u(2\pi, t) \quad (0 \leq t \leq T), \\ u(x, 0) = u_0(x) \quad (0 \leq x \leq 2\pi), \end{cases} \quad (3.4)$$

задач математической физики, предназначенными в первую очередь для построения асимптотических решений уравнений с переменными коэффициентами и нелинейных уравнений, а также материал для самостоятельной работы и возможных обобщений рассмотренного примера.

в соответствии с чем сетку $\Omega_{\tau, h}$ выберем, полагая

$$N = \frac{2\pi}{h}, \quad n = \frac{T}{\tau}. \quad (3.5)$$

Для построения функции, приближающей решение уравнения (3.1) в точках сетки $\Omega_{\tau, h}$, рассмотрим явную разностную схему (максимального порядка точности [9]) вида

$$\frac{u_j^{k+1} - u_j^k}{\tau} = c\varepsilon \left\{ \frac{u_{j+1}^k - u_j^k}{h} - \left(\frac{h}{2} - \frac{\tau}{2} \right) \times \right. \\ \left. \times \frac{u_{j+1}^{k+1} - u_j^{k+1} - u_{j+1}^k + u_j^k}{\tau h} \right\}, \quad (3.1')$$

где u_j^k — значение сеточной функции $u_{\tau, h}(x, t)$, определенной в узлах сетки: $x_j = jh$, $t_k = k\tau$ ($j = 0, 1, \dots, N$; $k = 0, 1, \dots, n$).

Разностное уравнение (3.1') рассматривается во внутренних узлах сетки: x_j , t_k ($j = 1, 2, \dots, N-1$, $k = 1, 2, \dots, n-1$). В граничных узлах сетки x_j , t_k ($j = 0, N$, $k = 0$) рассматриваются начальные и граничные условия для уравнения (3.1'), аппроксимирующие соответствующие условия (3.2) и (3.3) для уравнения (3.1).

В данном случае эти условия для уравнения (3.1') имеют вид

$$u_0^k = u_N^k \quad (k = 0, 1, \dots, n), \quad (3.2')$$

$$u_j^0 = \varphi_j \quad (j = 0, 1, \dots, N), \quad (3.3')$$

где $\{\varphi_j, j = 0, 1, \dots, N\}$ — произвольный набор чисел с условием $\varphi_0 = \varphi_N$, аппроксимирующий начальную функцию $u_0(x)$. Если $u_0(x)$ — гладкая функция, то можно положить

$$\varphi_j = u_0(x = jh) \quad (j = 0, 1, \dots, N). \quad (3.6)$$

Если разностное уравнение (3.1') и краевые условия (3.2') позволяют выразить значение сеточной функции на временном слое $t_{k+1} = (k+1)\tau$: $(u_0^{k+1}, \dots, u_N^{k+1})$ через значение сеточной функции на слое $t_k = k\tau$: $(u_0^k, \dots, u_N^k)$, то можно шаг за шагом вычислить значения решения разностного уравнения в узлах сетки на слоях $t = \tau$, $t =$

$= 2\tau, \dots, t = n\tau$, поскольку значения u_j^0 ($j = 0, 1, \dots, N$) известны и определяются равенствами (3.6).

Введем отношение шагов сетки $\alpha = \frac{\tau}{h}$ и выберем систему единиц, в которой $c = 1$, тогда систему (3.1') можно преобразовать к виду

$$(1 + \varepsilon\alpha) u_j^{k+1} + (1 - \varepsilon\alpha) u_{j+1}^{k+1} = (1 - \varepsilon\alpha) u_j^k + (1 + \varepsilon\alpha) u_{j+1}^k. \quad (3.1'')$$

С учетом краевого условия (3.2') при каждом $k = 0, 1, \dots, n$ сеточная функция $u_{\tau, h}(x, t)$ задается выбором N чисел $(u_0^k, u_1^k, \dots, u_{N-1}^k)$. Обозначив через u^k N -мерный вектор $u^k = (u_0^k, u_1^k, \dots, u_{N-1}^k)$ систему (3.1), (3.2) запишем в матричной форме

$$A_1 u^{k+1} = A_0 u^k \quad (k = 0, 1, \dots, n-1), \quad (3.7)$$

где матрицы A_1 и A_0 имеют вид

$$A_1 = \left\| \begin{array}{cccccc} (1 + \varepsilon\alpha) & (1 - \varepsilon\alpha) & 0 & \dots & 0 \\ 0 & (1 + \varepsilon\alpha) & (1 - \varepsilon\alpha) & \dots & 0 \\ 0 & 0 & (1 + \varepsilon\alpha) & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ 0 & 0 & 0 & \dots & (1 + \varepsilon\alpha) & (1 - \varepsilon\alpha) \\ (1 - \varepsilon\alpha) & 0 & 0 & \dots & (1 + \varepsilon\alpha) \end{array} \right\|_{N \times N} \quad (3.8)$$

$$A_0 = \left\| \begin{array}{cccccc} (1 - \varepsilon\alpha) & (1 + \varepsilon\alpha) & 0 & \dots & 0 \\ 0 & (1 - \varepsilon\alpha) & (1 + \varepsilon\alpha) & \dots & 0 \\ 0 & 0 & (1 - \varepsilon\alpha) & \dots & 0 \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ \cdot & \cdot & \cdot & \cdot & \\ 0 & 0 & 0 & \dots & (1 - \varepsilon\alpha) & (1 + \varepsilon\alpha) \\ (1 + \varepsilon\alpha) & 0 & 0 & \dots & (1 - \varepsilon\alpha) \end{array} \right\|_{N \times N}$$

Нетрудно проверить, что определитель $\det B$ матрицы B вида

$$B = \begin{vmatrix} b_1 & b_2 & 0 & \dots & 0 \\ 0 & b_1 & b_2 & \dots & 0 \\ 0 & 0 & b_1 & \dots & 0 \\ \cdot & & & \cdot & \cdot \\ \cdot & & & \cdot & \cdot \\ \cdot & & & \cdot & \cdot \\ 0 & 0 & 0 & \dots & b_1 & b_2 \\ b_2 & 0 & 0 & \dots & b_1 & b_2 \end{vmatrix}_{N \times N}$$

равен

$$\det B = b_1^N + (-1)^N b_2^N \quad (N \geq 3). \quad (3.9)$$

Отсюда следует, что матрица A_1 при условии $\varepsilon\alpha \neq 0$ для любого $N = 3, 4, \dots$ невырождена, так как

$$\det A_1 = (1 + \varepsilon\alpha)^N + (-1)^{N+1} (1 - \varepsilon\alpha)^N \neq 0.$$

Используя невырожденность матрицы A_1 , систему (3.7) разрешаем относительно вектора u^{k+1} :

$$u^{k+1} = A_1^{-1} A_0 u^k \quad (k = 0, 1, \dots, n-1). \quad (3.10)$$

Покажем, что матрица перехода $Z = A_1^{-1} A_0$ от слоя t_k к слою t_{k+1} — унитарная неприводимая матрица и тем самым определяет некоторую КМ-цепь (Z, G_Z) . Для этого найдем собственные значения матрицы $Z = A_1^{-1} A_0$ из уравнения

$$\det(\lambda I - A_1^{-1} A_0) = \det A_1^{-1} \det(\lambda A_1 - A_0) = 0. \quad (3.11)$$

Используя равенство (3.9) и явный вид матриц A_1 и A_0 из (3.8), уравнение (3.11) можно переписать в виде

$$[(1 + \varepsilon\alpha)\lambda - (1 - \varepsilon\alpha)]^N + (-1)^{N+1} [(1 - \varepsilon\alpha)\lambda - (1 + \varepsilon\alpha)]^N = 0 \quad (3.12)$$

Отсюда для собственных значений λ_j имеем равенства

$$\lambda_j = \frac{1 - \varepsilon\alpha + (1 + \varepsilon\alpha) e^{i\gamma_j}}{1 + \varepsilon\alpha + (1 - \varepsilon\alpha) e^{i\gamma_j}} \quad (j = 0, 1, \dots, N-1), \quad (3.12a)$$

где

$$\gamma_j = \frac{2\pi j}{N} + \frac{\pi(N+1)}{N} \quad (j = 0, 1, \dots, N-1). \quad (3.12b)$$

В силу формул (3.12) все собственные числа λ_j ($j = 0, 1, \dots, N-1$) попарно различны и по модулю равны единице:

$$|\lambda_j| = \left| \frac{\beta + e^{i\gamma_j}}{1 + \beta e^{i\gamma_j}} \right| = 1,$$

где $\beta = \frac{1 - \varepsilon\alpha}{1 + \varepsilon\alpha}$. Отсюда следует, что матрица $Z = A_1^{-1} A_0$ унитарная, а ее неприводимость будет доказана в п. 4 § 1 гл. II.

Таким образом, разностная схема (3.1') для уравнения (3.1) определяет однородную комплексную марковскую цепь (Z, G_Z) с матрицей $Z = A_1^{-1} A_0$, где A_1 и A_0 определены равенствами (3.8).

Решение разностной схемы (3.10) на n -м временном слое можно выразить через начальные данные в виде

$$u^n = (Z)^n u^0. \quad (3.13)$$

где $u^0 = (\varphi_0, \dots, \varphi_{N-1})$.

Матрица Z^n —это матрица условной амплитуды перехода КМ-цепи (Z, G_Z) за n шагов на графе G_Z (см. § 2).

Определение матрицы Z^n эквивалентно решению разностной задачи (3.1') — (3.3').

Решение задачи (3.1) — (3.3) можно получить, если разностную схему (3.1) записать в виде псевдодифференциального уравнения. Для этого продолжим сеточную функцию u_j^k специальным образом до гладкой по x и t функции $u(t, x)$ ($u_j^k = u(t_k, x_j)$) так, чтобы при каждом $t = t_k$ для функции $u(t, x)$ выполнялись краевые условия (3.2). С этой целью рассмотрим векторное пространство M_N функций, представляющее собой линейную оболочку векторов

$$e^{imx}, \quad m = 0, \pm 1, \pm 2, \dots, \pm M, \quad M = \frac{N-1}{2}, \\ N - \text{нечетно.}$$

Очевидно, что для заданного набора чисел $(u_0^k, u_1^k, \dots, u_N^k) = u^k$ с условием $u_0^k = u_N^k$ ($k = 0, 1, \dots, n$) существует функция $u^k(x) \in M_N$ такая, что

$$u^k(x = jh) = u_j^k \quad (j = 0, 1, \dots, N),$$

причем в силу определения пространства M_N краевое условие (3.2') удовлетворяется автоматически:

$$u^k(0) = u^k(Nh = 2\pi) \quad (k = 0, 1, \dots, n).$$

Таким образом, начальные и краевые условия (3.2') — (3.3') в узлах сетки $\Omega_{\tau, h}$ выполнены, если решение разностного уравнения (3.1) при $t = t_k$ рассматривать как сужение на сетку гладких функций из пространства M_N .

Далее, используя соотношения

$$e^{h \frac{\partial}{\partial x}} u(t, x) = e^{i(-ih \frac{\partial}{\partial x})} u(t, x) = u(t, x + h)$$

$$e^{\tau \frac{\partial}{\partial t}} u(t, x) = e^{i(-i\tau \frac{\partial}{\partial t})} u(t, x) = u(t + \tau, x)$$

и учитывая равенство $\tau = ah$, перепишем задачу (3.1') — (3.3') в виде псевдодифференциального уравнения с начальными данными, принадлежащими пространству функций M_N :

$$Lu = \left[(1 + \varepsilon\alpha) e^{ah \frac{\partial}{\partial t}} + (1 + \varepsilon\alpha) e^{ah \frac{\partial}{\partial t}} e^{h \frac{\partial}{\partial x}} - (1 + \varepsilon\alpha) e^{h \frac{\partial}{\partial x}} - (1 - \varepsilon\alpha) \right] u(x, t) = 0, \quad (3.14)$$

$$u(x, 0) = u_0(x) \in M_N. \quad (3.14a)$$

Обозначим через A оператор $A = -i \frac{\partial}{\partial x}$ и положим

$\omega = hA$, $h = \frac{2\pi}{N}$. Решение задачи (3.14) будем искать [6] в виде

$$u(x, t) = e^{iAS\left(\begin{smallmatrix} 2 \\ x, t, \omega \end{smallmatrix}\right)} \varphi\left(\begin{smallmatrix} 2 \\ x, t, \omega \end{smallmatrix}\right) u_0(x). \quad (3.15)$$

Здесь $S(x, t, \lambda)$, $\varphi(x, t, \lambda)$ — вещественные бесконечно дифференцируемые по всем своим аргументам и 2π -периодические по $x \in \mathbb{R}^1$ функции. Выражение вида $F\left(\begin{smallmatrix} 2 \\ x, \omega \end{smallmatrix}\right)$ представляет собой оператор, являющийся функцией от упорядоченной пары некоммутирующих друг с другом операторов ω и x . Индексы 1 и 2 над операторами ω и x в выражении $F\left(\begin{smallmatrix} 2 \\ x, \omega \end{smallmatrix}\right)$ показывают, что первым действует оператор

$\omega = -ih \frac{\partial}{\partial x}$, вторым — оператор умножения на x . Точный смысл этого утверждения состоит в следующем определении действия оператора $F \begin{pmatrix} 2 & 1 \\ x & \omega \end{pmatrix}$ на функции $v(x) \in M_N$.

$$F \begin{pmatrix} 2 & 1 \\ x & \omega \end{pmatrix} v(x) = \frac{1}{2\pi} \sum_{-M}^{+M} e^{imx} F(x, mh) \int_0^{2\pi} e^{-lmy} v(y) dy. \quad (3.16)$$

Подставим выражение (3.15) в уравнение (3.14) и, воспользовавшись теоремой о коммутации с экспонентой [6], получим

$$\begin{aligned} e^{iAS} \left[(1 + \varepsilon\alpha) e^{i\alpha\omega \frac{\partial S}{\partial t}} + (1 - \varepsilon\alpha) \left[e^{i\alpha\omega \frac{\partial S}{\partial t}} \right] \left[e^{\left[h \frac{\partial}{\partial x} + i\omega \frac{\partial S}{\partial x} \right]} \right] \right. \\ \left. - (1 + \varepsilon\alpha) e^{\left[h \frac{\partial}{\partial x} + i\omega \frac{\partial S}{\partial x} \right]} - (1 - \varepsilon\alpha) \right] \varphi u_0(x) = \\ = e^{iAS} L \varphi u_0 = 0, \quad (3.17) \end{aligned}$$

где у функций S , φ опущены аргументы (x, t, ω) .

Для того чтобы функция $u(x, t)$ из (3.15) являлась решением уравнения (3.14), достаточно, чтобы выполнялось равенство

$$L \varphi u_0 = 0, \quad (3.18)$$

где через L обозначается операторное выражение, стоящее в квадратных скобках в равенстве (3.17).

Положим функцию $\varphi \equiv 1$, а функцию $S(x, t, \omega)$ определим как решение уравнения

$$\begin{aligned} (1 + \varepsilon\alpha) e^{i\alpha\omega \frac{\partial S}{\partial t}} + (1 - \varepsilon\alpha) e^{i\alpha\omega \frac{\partial S}{\partial t}} e^{i\omega \left(\frac{\partial S}{\partial x} + 1 \right)} - \\ - (1 + \varepsilon\alpha) e^{i\omega \left(\frac{\partial S}{\partial x} \right) + 1} - (1 - \varepsilon\alpha) = 0. \quad (3.19) \end{aligned}$$

При таком выборе функции $S(x, t, \omega)$ прямой подстановкой в уравнение (3.14а) можно показать, что функция

$$u(x, t) = e^{iAS} \begin{pmatrix} 2 & 1 \\ x & t, \omega \end{pmatrix} u_0(x), \quad (3.20)$$

где S удовлетворяет уравнение (3.19), является точным решением уравнения (3.14а).

Для того чтобы решение из (3.20) удовлетворяло начальному условию (3.14б), необходимо положить

$$S(x, t, \omega)|_{t=0} = 0. \quad (3.21)$$

Таким образом, для рассматриваемого случая операторный метод позволяет полностью свести решение разностной задачи (3.14а) — (3.14б) к решению нелинейного уравнения первого порядка (3.19) относительно функции $S(x, t, \omega)$, зависящей от параметра $\omega \in [-\pi, +\pi]$, с начальным условием (3.21). Уравнение (3.19) относительно функции $S(x, t, \omega)$ называется *уравнением Гамильтона—Якоби*. Решение уравнения Гамильтона—Якоби (3.19) с начальным условием (3.21) получим методом бихарактеристик [6].

Предварительно разрешим уравнение (3.19) относительно производной $\frac{\partial S}{\partial t}$. Полагая $\beta = \frac{1 - \varepsilon\alpha}{1 + \varepsilon\alpha}$, из (3.19) получаем равенство

$$\begin{aligned} \frac{\partial S}{\partial t} = \frac{1}{i\alpha\omega} \ln \left| \frac{\beta + e^{i\omega\left(\frac{\partial S}{\partial x} + 1\right)}}{1 + \beta e^{i\omega\left(\frac{\partial S}{\partial x} + 1\right)}} \right| + \\ + \frac{1}{\alpha\omega} \arg \left(\frac{\beta + e^{i\omega\left(\frac{\partial S}{\partial x} + 1\right)}}{1 + \beta e^{i\omega\left(\frac{\partial S}{\partial x} + 1\right)}} \right). \end{aligned}$$

Поскольку $\left| \frac{\beta + e^{i\omega\left(\frac{\partial S}{\partial x} + 1\right)}}{1 + \beta e^{i\omega\left(\frac{\partial S}{\partial x} + 1\right)}} \right| = 1$, то для определения

функции $S(x, t, \omega)$ окончательно получаем задачу (задачу Коши)

$$\frac{\partial S}{\partial t} = \frac{1}{\alpha\omega} \arg \left(\frac{\beta + e^{i\omega\left(\frac{\partial S}{\partial x} + 1\right)}}{1 + \beta e^{i\omega\left(\frac{\partial S}{\partial x} + 1\right)}} \right), \quad (3.22)$$

$$S|_{t=0} = 0$$

с вещественным гладким гамильтонианом $H(x, p) =$

$$= \frac{1}{\alpha \omega} \arg \left(\frac{\beta + e^{i\omega(p+1)}}{1 + \beta e^{i\omega(p+1)}} \right). \text{ Метод бихарактеристик дает}$$

для задачи (3.22) решение $S(x, t, \omega)$ в виде

$$S(x, t, \omega) = \frac{t}{\alpha \omega} \left[\operatorname{arctg} \left(\frac{\sin \omega}{\beta + \cos \omega} \right) - \operatorname{arctg} \left(\frac{\beta \sin \omega}{1 + \beta \cos \omega} \right) \right]. \quad (3.23)$$

Отсюда и из формулы (3.20) точное решение разностной задачи (3.14а) — (3.14б) представимо в виде

$$u(x, t) = e^{\frac{it}{ah} \left[\operatorname{arctg} \left(\frac{\sin \omega}{\beta + \cos \omega} \right) - \operatorname{arctg} \left(\frac{\beta \sin \omega}{1 + \beta \cos \omega} \right) \right]} u_0(x), \quad (3.24)$$

где, напомним, оператор $F(\omega, t)$ на функциях $u_0(x) \in M_N$ определяется формулой

$$u(x, t) = F(\omega, t) u_0(x) = \frac{1}{2\pi} \sum_{m=-M}^{+M} e^{imx} F(mh, t) \times \\ \times \int_0^{2\pi} e^{-imy} u_0(y) dy. \quad (3.25)$$

Полагая в формуле (3.25)

$$u_0(x) = \delta_h(x - j_0 h) = \\ = \sum_{m=-M}^{+M} e^{im(x-j_0 h)} \quad (0 \leq j_0 \leq N-1), \quad (3.26)$$

получим следующее выражение для разностной функции Грина уравнения (3.14а)

$$K_h(x, t, j_0) = \\ = \sum_{m=-M}^{+M} e^{im(x-j_0 h)} e^{\frac{it}{ah} \left[\operatorname{arctg} \frac{\sin mh}{\beta + \cos mh} - \operatorname{arctg} \frac{\beta \sin mh}{1 + \beta \cos mh} \right]}, \quad (3.27)$$

где $\beta = \frac{1 - \varepsilon \alpha}{1 + \varepsilon \alpha}$. Отсюда следует окончательная формула для амплитуды Грина $K_{j_0 j_n}^{(n)}(\tau, h)$ КМ-цепи (Z, G_Z) , если рас-

смотреть ограничение полученной в (3.27) функции $K_h(x, t, j_0)$ на слой $t_n = n\tau = T$ сетки $\Omega_{\tau, h}$ в точке $x = j_n h$ и поменять местами индексы j_0 и j_n^* :

$$K_{j_0 j_n}^{(n)}(\tau, h) = \sum_{m=-M}^{+M} e^{-im(j_n - j_0)h} e^{in \left[\arctg \frac{\sin mh}{\beta + \cos mh} - \arctg \frac{\beta \sin mh}{1 + \beta \cos mh} \right]}. \quad (3.28)$$

Относительно сходимости при $h \rightarrow 0$ построенного решения $K_h(j_n, h, n\tau, j_0) \equiv K_{j_0 j_n}^{(n)}(\tau, h)$ разностной задачи (3.1') — (3.3') с начальным условием

$$K_h(j_n h, 0, j_0) = \sum_{m=-M}^{+M} e^{-im(j_n - j_0)h} \equiv \frac{1}{h} \delta_{j_n j_0} = \begin{cases} \frac{1}{h}, & j_n = j_0 \quad (j_n = 0, 1, \dots, N) \\ 0, & j_n \neq j_0 \quad (0 \leq j_0 \leq N-1) \end{cases} \quad (3.29)$$

к решению дифференциальной задачи (3.1) — (3.3) с начальным условием $u_0(x) = \delta(x - x_0)$ заметим следующее. Функция $K_h(x, t, j_0)$ является гладким продолжением на отрезок $[0, 2\pi]$ сеточной функции $K_{j_0}^{(n)}(\tau, h)$, определенной в узлах сетки $\Omega_{\tau, h}$. Поэтому под сходимостью решения разностной задачи $K_{j_0}^{(n)}(\tau, h)$ к решению $u(x, t)$ дифференциальной задачи будем понимать сходимость при $h \rightarrow 0$ последовательности гладких функций $K_h(x, t, j_0) \in M_N$ к функции $u(x, t)$. В силу формул (3.15), (3.23) и (3.26) функция $K_h(x, t, j_0)$ представляется в виде

$$K_h(x, t, j_0) = e^{iAS(t, h, A)} \delta_h(x - j_0 h), \quad (3.30)$$

* Как следует из формулы (3.13), решение $u_{j_n}^{(n)}(j_0)$ ($j_n = 0, 1, \dots, N$) разностной задачи Коши с начальным условием

$$u_i^{(0)} = \delta_{ij_0} = \begin{cases} 1, & i = j_0 \\ 0, & i \neq j_0 \end{cases} \quad (i = 0, 1, \dots, N-1) \text{ при фиксирован-}$$

ном j_0 ($j_0 = 0, 1, \dots, N-1$) в узле $(x_j = jh, t_n = n\tau)$ сетки $\Omega_{\tau, h}$ является матричным элементом $[Z^n]_{j_n j_0}$ транспонированной матрицы $(([Z^n]_{j_0 j_n}))' = ((K_{j_0 j_n}^{(n)}))'$.

где $A = -i \frac{\partial}{\partial x}$.

Хорошо известно [2], что последовательность функций $\delta_h(x - j_0 h)$ при $h \rightarrow 0$ ($j_0 h \rightarrow x_0$) сходится к обобщенной функции $\delta(x - x_0)$ по норме пространства $H_{-1}(S^1)$, где $H_{-1}(S^1)$ — пространство обобщенных функций первого порядка на окружности S^1 . Напомним, что норма функции $f(x) \in H_{-1}(S^1)$ определяется следующим образом:

$$\|f\|_{H_{-1}(S^1)} = \int_0^{2\pi} f(x) \left(-\frac{d^2}{dx^2} + 1 \right)^{-1} f(x) dx.$$

Нетрудно показать, что оператор $e^{iAS(t, h, A)}$ из пространства $H_s(S^1)$ в $H_s(S^1)$ (s — целое) равномерно по h ограничен и при $h \rightarrow 0$ сходится к оператору сдвига $e^{iAS(t, 0)} = e^{+ct \frac{\partial}{\partial x}}$. Отсюда следует, что при каждом $t \in [0, T]$

$$\lim_{h \rightarrow 0} K_h(x, t, j_0) = u(x, t) = \delta(x + ct), \quad (3.31)$$

(по норме пространства $H_{-1}(S^1)$). Полагая, как обычно, норму $\|\cdot\|_B$ в пространстве B функций от x и t равной

$$\|f\|_B = \max_{t \in [0, T]} \|f\|_{H_{-1}(S^1)}, \quad (3.32)$$

из (3.31) получаем сходимость последовательности $K_h(x, t, j_0)$ к точному решению $u(x, t)$ в пространстве B .

Глава II

ДИАГРАММЫ ФЕЙНМАНА В ТЕОРИИ ВОЗМУЩЕНИЙ КМ-ЦЕПИ И КВАНТОВОЙ ТЕОРИИ

Графические схемы, впервые введенные Р. Фейнманом [11], [12] для представления членов ряда теории возмущений в квантовой электродинамике, — диаграммы Фейнмана, находят в настоящее время широкое применение в различных разделах теоретической физики. В этой главе на примере нерелятивистской квантовой механики (§ 2) предпринята попытка понять диаграммы Фейнмана в теории возмущений непрерывной квантовой системы с точки

зрения диаграмм дискретной модели этой системы — КМ-цепи, определенной в гл. I. Диаграммы теории возмущений рассматриваются здесь как классы эквивалентных траекторий, имеющих вклады одного и того же порядка малости относительно малого параметра, по которому ведется разложение в ряды теории возмущений (в приложениях таким параметром обычно является константа взаимодействия системы с внешним полем).

Реализация идеи разложения всей совокупности траекторий, определяющей полную амплитуду вероятности перехода системы из одного состояния в другое, на классы эквивалентности основана на возможности представления амплитуды Грина КМ-цепи на графе, представляющем собой сумму графов матриц. На этом пути удастся получить именно тот вид диаграмм, который обычно используется в литературе для интерпретации членов ряда теории возмущений в нерелятивистской квантовой механике.

В § 1 приводится формула (1.21) для вычисления вклада диаграмм теории возмущений КМ-цепи в полную амплитуду Грина КМ-цепи, которая является дискретным аналогом известных правил вычисления вкладов диаграмм теории возмущений в квантовой механике в импульсном представлении. При указанном подходе выясняется целесообразность и естественность рассмотрения рядов теории возмущений в импульсном представлении, в котором волновая функция непрерывной системы выражается в координатах: 1) импульс $p \in \mathbb{R}^3$, время $t \in \mathbb{R}^1$ или 2) импульс $p \in \mathbb{R}^3$, энергия $E \in \mathbb{R}^1$. Переход от координат (p, t) к координатам (p, E) в непрерывном случае осуществляется с помощью преобразования Фурье по переменным $t \rightarrow E$.

Аналогичные понятия, а именно, « p, t » и « p, E »-представления определяются для КМ-цепи соответственно в § 1 и 2. При этом преобразованию Фурье, осуществляющему переход от переменных (p, t) к переменным (p, E) в непрерывной системе, сопоставляется (§ 2) преобразование диаграммы теории возмущений дискретной КМ-цепи, которое уменьшает в переменных (p, E) число топологически неэквивалентных диаграмм.

Диаграмма Фейнмана в теории возмущений нерелятивистской квантовой механики определяется как совокупность объектов, состоящая из графа диаграммы и коцепи, определенной на дугах графа со значениями в группе коэффициентов, выбор которой зависит от выбора представления КМ-цепи.

В « p , E »-представлении КМ-цепи соответствующая кощепь принимает значение в четырехмерном векторном пространстве импульсов $\mathfrak{R}^4$ и удовлетворяет первому закону Кирхгофа (гл. XV, § 1) в каждой внутренней вершине графа диаграммы. Это свойство диаграмм квантовой механики при переходе к квантовой теории поля (§ 3) интерпретируется как закон сохранения 4-вектора «энергии-импульса» взаимодействующих частиц при столкновениях*.

§ 1. РЕГУЛЯРНАЯ ТЕОРИЯ ВОЗМУЩЕНИЙ КМ-ЦЕПИ

1. Связь разложения в ряд теории возмущений матриц и соответствующих им графов. Требование унитарности матрицы Z в определении КМ-цепи (Z, G_Z) , важное для приложений, в частности, в квантовой механике (см. § 2 гл. I), несущественно для построения теории возмущений КМ-цепи. Поэтому в этом параграфе будет рассматриваться КМ-цепь с произвольной неприводимой матрицей* порядка N , обозначаемой через A , $A = ((a_{ij}))$, $a_{ij} \in \mathbb{C}$.

Рассмотрим задачу о разложении амплитуды Грина $K^{(n)}(\epsilon)$ однородной КМ-цепи $(A(\epsilon), G_{A(\epsilon)})$, аналитически зависящей от малого параметра ϵ ($|\epsilon| \ll 1$), в ряд Тейлора

$$K^{(n)}(\epsilon) = K_0^{(n)} + \epsilon^1 K_1^{(n)} + \dots + \epsilon^m K_m^{(n)} + \dots = \sum_{r=0}^{\infty} \epsilon^r K_r^{(n)}, \quad (1.1)$$

называемый *рядом теории возмущений КМ-цепи*. Аналитическая зависимость КМ-цепи от параметра ϵ здесь означает, что элементы матрицы $A(\epsilon)$ — аналитические функции по ϵ на интервале $(-1, +1)$.

Используя интерпретацию суммы и произведения матриц на графах, каждому члену $K_r^{(n)}$ порядка r в разложении (1.1) удастся сопоставить семейство графов достаточно простой структуры, называемых *графами теории возмущений КМ-цепи*, и сформулировать правила вычисления матричных элементов каждого члена разложения (1.1) по этим графам. Идея подобного сопоставления чрезвычайно проста и состоит в следующем.

* Рассмотрение неприводимых матриц не ограничивает общности. Если матрица Z приводится к квазидиагональной с блоками $Z_1, Z_2, \dots, Z_p$, то граф G_Z матрицы Z состоит из p компонент связности $G_1, G_2, \dots, G_p$ с «матрицами смежности» $Z_1, Z_2, \dots, Z_p$ соответственно и, поэтому дальнейшие построения можно проводить для каждой из компонент связности G_i независимо от других.

Поскольку амплитуда Грина $K^{(n)}(\epsilon)$ представляет собой значение функционала Φ на множестве путей $L^n(G_{A(\epsilon)})$ длины n графа $G_{A(\epsilon)}$, $K^{(n)}(\epsilon) = \Phi(L^n(G_{A(\epsilon)}))$ (см. § 1 гл. I), то естественно ожидать, что разложению (1.1) отвечает соответствующее «разложение», или «расслоение» множества путей $L^n(G_{A(\epsilon)})$. Это «расслоение» множества путей $L^n(G_{A(\epsilon)})$, очевидно, должно индуцироваться разложением матрицы $A(\epsilon)$ в ряд Тейлора

$$A(\epsilon) = A_0 + \epsilon^1 A_1 + \dots + \epsilon^{m_1} A_{m_1} + O(\epsilon^{m_1+1}), \quad (1.2)$$

так как с алгебраической точки зрения для получения r -го приближения $K_r^{(n)}$ амплитуды Грина необходимо сгруппировать члены при ϵ^r в разложении (1.3):

$$\begin{aligned} K^{(n)}(\epsilon) &\equiv (A(\epsilon))^n = (A_0 + \epsilon A_1 + \dots + \epsilon^{m_1} A_{m_1} + \\ &+ O(\epsilon^{m_1+1}))^n = A_0^n + \epsilon A_0^{n-1} A_1 + \dots + \epsilon^2 A_0^{n-1} A_2 + \dots + \\ &+ \epsilon^{m_1} A_0^{n-1} A_{m_1} + \dots + O(\epsilon^m), \quad m = m_1 n. \end{aligned} \quad (1.3)$$

Однако в общем случае для произвольной матрицы $A(\epsilon)$ непосредственная реализация этой идеи разложения множества путей графа $G_{A(\epsilon)}$ затруднительна ввиду «устойчивости» графа G_A матрицы A относительно малых возмущений матрицы A , а следовательно, и «устойчивости» геометрической структуры множества путей $L^n(G_A)$ на графе G_A .

Устойчивость графа матрицы относительно малых возмущений матрицы означает, что если G_{A_1} — граф матрицы A_1 , а G_{A_2} — граф матрицы A_2 , причем,

$$A_2 = A_1 + \epsilon \Delta = A_2(\epsilon), \quad \epsilon \rightarrow 0, \quad (1.4)$$

$$A_1 = ((a_{ij}^{(1)})), \quad a_{ij}^{(1)} \neq 0 \quad (i, j = 1, 2, \dots, N),$$

Δ — произвольная матрицы ($\Delta \neq 0$), то (в силу определения 1.4) гл. I) граф G_{A_1} совпадает с графом G_{A_2} :

$$G_{A_1} \equiv G_{A_2}^*. \quad (1.5)$$

* Для выполнения равенства (1.5) достаточно, например, по-

$$\text{ложить в (1.4) } \epsilon = \frac{\min_{1 \leq i, j \leq N} |a_{ij}^{(1)}|}{\max_{1 \leq i, j \leq N} |\Delta_{ij}|}.$$

Устойчивость графа матрицы относительно малых возмущений применительно к матрице $A(\varepsilon)$ из (1.2) приводит к тому, что в формуле, выражающей амплитуду $K^{(n)}(\varepsilon)$ через вклады путей на графе $G_{A(\varepsilon)}$,

$$[K^{(n)}(\varepsilon)]_{ij} = \sum_{l \in l_{ij}^{(n)}(G_{A(\varepsilon)})} f(l), \quad (1.6)$$

где

$$f(l) = \prod_{k=1}^n f(u_k), \quad f(u_k) = a_{i_k i_{k+1}}(\varepsilon) = a_{i_k i_{k+1}}^{(0)} + \\ + \varepsilon a_{i_k i_{k+1}}^{(1)} + \dots + \varepsilon^{m_1} a_{i_k i_{k+1}}^{(m_1)},$$

если $u_k = (x_{i_k}, x_{i_{k+1}})$, ($i_1 = i$, $i_{n+1} = j$), можно при $\varepsilon \rightarrow 0$ заменить граф $G_{A(\varepsilon)}$ на G_{A_0} :

$$[K^{(n)}(\varepsilon)]_{ij} = \sum_{l \in l_{ij}^{(n)}(G_{A_0})} f(l) \quad (1.6a)$$

вне зависимости от структуры матриц $A_1, A_2, \dots, A_n$. При этом возникает «рассогласование» между возможностью разложить левую часть равенства (1.6a) в ряд (1.3) по степеням ε и независимостью структуры множества путей $l_{ij}^{(n)}(G_{A_0})$, по которым производится суммирование в правой части равенства (1.6a), от параметра ε . В то время как множество путей $l_{ij}^{(n)}(G_{A_0})$ однозначно определено матрицей A_0 , вклад $f(l)$ пути $l \in l_{ij}^{(n)}(G_{A_0})$ зависит от всех матриц A_i , $i = 0, 1, \dots, n$ в разложении (1.2) матрицы $A(\varepsilon)$ и параметра ε в силу формулы (1.6).

Последнее обстоятельство обеспечивает выполнение равенства (1.6a) при каждом $\varepsilon \in (-1, +1)$, но при разложении вклада пути $f(l) = \prod_{k=0}^{n-1} a_{i_k i_{k+1}}$ по степеням ε получающиеся коэффициенты разложения при ε^r в правой части равенства (1.6a) уже не допускают наглядной интерпретации в виде вкладов путей на графе. Чтобы обойти это затруднение, необходимо перейти от представления КМ-цепи на графе $G_{A(\varepsilon)}$ к представлению КМ-цепи на графе

$\hat{G}_{A(\varepsilon)}$ — аддитивном графе суммы матриц (1.2). Напомним (см. с. 306), что по определению

$$\hat{G}_{A(\varepsilon)} = G_{A_0} + G_{A_1} + \dots + G_{A_m} + G_{0(\varepsilon)^{m+1}}. \quad (1.7)$$

Обобщая очевидным образом полученное в § 1 гл. I правило вычисления матричных элементов n -й степени суммы двух матриц, для произвольного числа матриц $M_1, M_2, \dots, M_m$ получаем следующее правило.

П р а в и л о вычисления n -й степени суммы матриц $M_1 + M_2 + \dots + M_m$.

Матричный элемент $[(M_1 + M_2 + \dots + M_m)^n]_{ij}$ равен сумме вкладов всех произвольно чередующихся путей длины n на аддитивном графе $\hat{G}_m^{\sum_{i=1}^m M_i}$ суммы матриц

$$[(M_1 + M_2 + \dots + M_m)^n]_{ij} = \sum_{l \in \widehat{l}_{ij}^{(n)}(M_1, M_2, \dots, M_m)} f(l), \quad (1.8)$$

где

$$\hat{G}_m^{\sum_{i=1}^m M_i} = \sum_{i=1}^m G_{M_i}, \quad (1.9)$$

a вклад $f(l)$ пути l является произведением матричных элементов матриц $M_1, M_2, \dots, M_m$, причем число матричных элементов матрицы M_i ($i = 1, 2, \dots, m$) в произведении $f(l)$ равно числу дуг пути l , принадлежащих суграфу G_{M_i} графа $\hat{G}_m^{\sum_{i=1}^m M_i}$.

Воспользовавшись этим правилом для амплитуды Грина $K^{(n)}(\varepsilon)$ КМ-цепи $(A(\varepsilon), G_{A(\varepsilon)})$, где $A(\varepsilon)$ удовлетворяет (1.2), получим представление

$$K^{(n)}(\varepsilon) = (A(\varepsilon))^n = \sum_{l \in \widehat{L}^n(\widehat{G}_{A(\varepsilon)})} f(l), \quad (1.10)$$

где

$$\hat{G}_{A(\varepsilon)} = \sum_{i=1}^{m_1} G_{A_i} + G_{O(\varepsilon m_1 + 1)}, \quad (1.11)$$

а через $L^n(\hat{G}_{A(\varepsilon)})$ обозначается множество всех произвольно чередующихся путей длины n на графе $\hat{G}_{A(\varepsilon)}$.

КМ-цепь $(A(\varepsilon), \hat{G}_{A(\varepsilon)})$ эквивалентна КМ-цепи $(A(\varepsilon), \hat{G}_{A(\varepsilon)})$ в том смысле, что полные амплитуды Грина той и другой КМ-цепи совпадают. Но для приближенного вычисления амплитуды Грина $K^{(n)}(\varepsilon)$, как увидим в дальнейшем, более естественным и удобным является представление амплитуды $K^{(n)}(\varepsilon)$ на графе $\hat{G}_{A(\varepsilon)}$. Поэтому всюду ниже, если не будет оговорено особо, теория возмущений КМ-цепи $(A(\varepsilon), \hat{G}_{A(\varepsilon)})$ рассматривается на графе $\hat{G}_{A(\varepsilon)}$, являющемся аддитивным графом суммы матриц ряда Тейлора (1.2) для матрицы $A(\varepsilon)$.

Из формулы (1.10) в силу мультипликативности вклада пути $f(l)$ относительно дуг, входящих в путь l , следует, что для получения разложения амплитуды $K^{(n)}(\varepsilon)$ в r -м порядке теории возмущений ($r \leq m$) достаточно взять множество путей графа $\hat{G}_r = G_{A_0} + G_{A_1} + \dots + G_{A_r}$. Рассмотрим подробнее частный, но важный случай, когда в разложении (1.2) $A_n \equiv 0$, $n > 1$, $O(\varepsilon^{n+1}) \equiv 0$. Тогда амплитуда $K^{(n)}(\varepsilon)$ в любом порядке теории возмущений будет определяться вкладами путей длины n на графе $G_{A_0} + G_{A_1}$. Но любой путь l длины n на графе $G_{A_0} + G_{A_1}$ состоит из дуг суграфа G_{A_0} и суграфа G_{A_1} , причем вклад дуги суграфа G_{A_0} по порядку ε есть const , а вклад дуги суграфа $G_{A_1} = O(\varepsilon)$. Отсюда в силу мультипликативности вклада $f(l)$ пути l относительно дуг, входящих в него, и формулы (1.8) тогда следует, что для вычисления амплитуды Грина $K_r^{(n)}$ порядка $O(\varepsilon^r)$ из всего множества путей $L^n(G_{A_0} + G_{A_1})$ на графе $G_{A_0} + G_{A_1}$ необходимо рассмотреть только те произвольно чередующиеся пути длины n , которые содержат дуги суграфа G_{A_1} равно r раз ($r \leq n$). Обозначим множество таких путей через $l_{ij}^{(n, r)}(A(\varepsilon))$.

Таким образом, разложению амплитуды $K^{(n)}(\varepsilon)$ по степеням ε^r ($r = 0, 1, 2, \dots, m$) однозначно соответствует раз-

биение множества всех путей длины n на графе $\hat{G}_{A(\epsilon)}$ на подмножества путей вида $l_{ij}^{(n,r)}$. Множеству путей $l_{ij}^{(n,r)}(A(\epsilon))$ можно сопоставить семейство подграфов $\{G^{(r)}(A(\epsilon))\}$ порядка r графа $\hat{G}_{A(\epsilon)}$.

Оказывается, что достаточно просто можно описать структуру множества всех путей $l_{ij}^{(n,r)}(A(\epsilon))$ и соответствующих им подграфов, если рассмотреть так называемое « r, t »-представление КМ-цепи.

2. Определение « r, t »-представления КМ-цепи. Предположим, что матрица A_0 в разложении (1.2) матрицы $A(\epsilon)$ КМ-цепи $A(\epsilon)$, $G_{A(\epsilon)}$ имеет простую структуру, т. е. существует $N(N - \text{порядок матрицы } A_0)$ линейно независимых собственных векторов e_i матрицы A_0 ,

$$A_0 e_i = \lambda_i e_i \quad (i = 1, 2, \dots, N).$$

Определение 1.1. Представление матрицы $A(\epsilon) = A_0 + \epsilon A_1(\epsilon)$ в базисе собственных векторов матрицы A_0 назовем « r »-представлением* матрицы $A(\epsilon)$.

В « r »-представлении матрица $A(\epsilon)$ имеет вид

$$\tilde{A}(\epsilon) = U^{-1} A(\epsilon) U = [\lambda_1, \lambda_2, \dots, \lambda_N] + \epsilon \tilde{A}_1(\epsilon), \quad (1.12)$$

где U — матрица порядка $(N \times N)$, составленная из собственных векторов e_i матрицы A_0 . Через $[\lambda_1, \lambda_2, \dots, \lambda_N]$ обозначается диагональная матрица с элементами $\lambda_i (i = 1, 2, \dots, N)$ на диагонали, $\tilde{A}_1(\epsilon)$ — матрица, эквивалентная матрице $A_1(\epsilon)$, $\tilde{A}_1(\epsilon) = U^{-1} A_1(\epsilon) U$.

Определение 1.2. КМ-цепь вида $(\tilde{A}(\epsilon), \hat{G}_{\tilde{A}(\epsilon)})$, где матрица $\tilde{A}(\epsilon)$ — « r »-представление матрицы $A(\epsilon)$ КМ-цепи $(A(\epsilon), G_{A(\epsilon)})$, а граф $\hat{G}_{A(\epsilon)}$ определяется равенством

$$\hat{G}_{\tilde{A}(\epsilon)} = G_{\tilde{A}_0} + G_{\tilde{A}_1(\epsilon)}, \quad (1.13)$$

называется « r, t »-представлением КМ-цепи $(A(\epsilon), G_{A(\epsilon)})$.

* Использование термина « r »-представления связано с физическим смыслом матрицы A_0 для КМ-цепи нерелятивистской квантовой механики (см. ниже § 2).

Граф $\hat{G}_{\tilde{A}(\varepsilon)}$ называется графом матрицы $A(\varepsilon)$ в « p, t »-представлении.

Матрица $\tilde{K}^{(n)}(\varepsilon) = \tilde{A}(\varepsilon)^n$ называется « p, t »-представлением амплитуды Грина $K^{(n)}(\varepsilon)$ КМ-цепи $(A(\varepsilon), G_{A(\varepsilon)})$.

Из определения (1.2) следует, что граф $\hat{G}_{\tilde{A}(\varepsilon)}$ матрицы $A(\varepsilon)$ в « p, t »-представлении является суммой двух графов

$$\hat{G}_{\tilde{A}(\varepsilon)} = G_{[\lambda_1, \lambda_2, \dots, \lambda_n]} + G_{\tilde{A}_1(\varepsilon)},$$

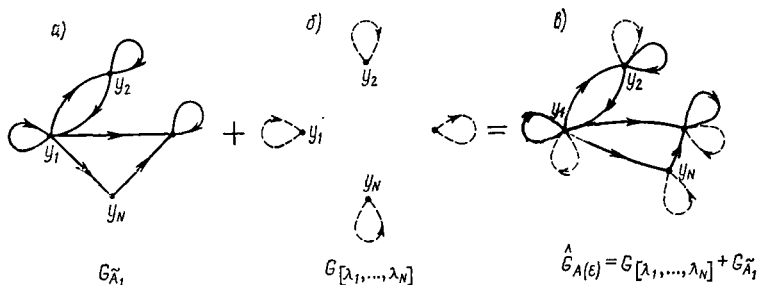


Рис. 164

где $G_{[\lambda_1, \lambda_2, \dots, \lambda_N]}$ — граф диагональной матрицы $[\lambda_1, \lambda_2, \dots, \lambda_N]$, состоящий из изолированных вершин, с одной петлей в каждой вершине, $G_{\tilde{A}_1(\varepsilon)}$ — граф матрицы $\tilde{A}_1(\varepsilon)$.

В качестве примера рассмотрим граф $G_{\tilde{A}_1(\varepsilon)}$ матрицы $\tilde{A}_1(\varepsilon)$, представленный на рис. 164, а, и граф $G_{[\lambda_1, \lambda_2, \dots, \lambda_N]}$ — на рис. 164, б (всюду ниже дуги графа $G_{[\lambda_1, \lambda_2, \dots, \lambda_N]}$ изображаются пунктирными линиями); граф $G_{[\lambda_1, \lambda_2, \dots, \lambda_N]} + G_{\tilde{A}_1(\varepsilon)} = \hat{G}_{\tilde{A}(\varepsilon)}$ изображен на рис. 164, в.

З а м е ч а н и е. Имея в виду приложение рассматриваемых понятий к теории возмущений в квантовой механике, подчеркнем, что вершины графа матрицы $\tilde{A}(\varepsilon)$ в « p, t »-представлении никак не связаны с возможными дискретными состояниями — вершинами $(x_1, x_2, \dots, x_N)$ графа $G_{A(\varepsilon)}$ для КМ-цепи, отвечающей матрице $Z^- =$

$= A(\varepsilon)$. В связи с этим вершины графа $\hat{G}_{A(\varepsilon)}$ обозначаются через $(y_1, y_2, \dots, y_N)$ в отличие от обозначений $(x_1, x_2, \dots, x_N)$ вершин графа $G_{A(\varepsilon)}$.

3. Диаграммы теории возмущений КМ-цепи в « p, t »-представлении. Рассмотрим разложение в ряд по степеням ε амплитуды Грина КМ-цепи в « p, t »-представлении

$$\tilde{K}^{(n)}(\varepsilon) = \tilde{K}_0^{(n)} + \varepsilon \tilde{K}_1^{(n)} + \dots + \varepsilon^m \tilde{K}_m^{(n)} = \sum_{r=0}^m \varepsilon^r \tilde{K}_r^{(n)}. \quad (1.14)$$

Из определения амплитуды Грина КМ-цепи и определения « p, t »-представления следует, что коэффициенты разложения (1.1) связаны с коэффициентами разложения (1.14) соотношением

$$K_r^{(n)} = U^1 \tilde{K}_r^{(n)} U^{-1} \quad (r = 0, 1, \dots, m).$$

Поэтому достаточно вычислить члены разложения КМ-цепи в « p, t »-представлении. Как следует из правила вычисления элементов n -й степени матрицы (см. формулу (1.12) гл. I), матричный элемент амплитуды

$$[\tilde{K}^{(n)}(\varepsilon)]_{ij} = [(\tilde{A}(\varepsilon))^n]_{ij} \quad (i, j = 1, 2, \dots, N)$$

определяется формулой

$$[\tilde{K}^{(n)}(\varepsilon)]_{ij} = \sum_{l \in l_{ij}^{(n)}(\tilde{A}(\varepsilon))} f(l), \quad (1.15)$$

где $f(l)$ — произведение весов $f(u_k)$ ($k = 1, 2, \dots, n$) дуг u_k , составляющих путь l длины n на графе $\hat{G}_{A(\varepsilon)}$.

Для простоты рассуждений положим вначале, что матрица $\tilde{A}_1(\varepsilon) = ((\tilde{a}_{ij}))_{N \times N}$ не зависит от ε , т. е. что в « p, t »-представлении матрица $A(\varepsilon)$ КМ-цепи $(A(\varepsilon), G_{A(\varepsilon)})$ имеет вид

$$\tilde{A}(\varepsilon) = [\lambda_1, \lambda_2, \dots, \lambda_N] + \varepsilon \tilde{A}_1. \quad (1.16)$$

Из соотношения (1.16) и из правила сложения графов следует, что вклад дуги u графа $\hat{G}_{A(\varepsilon)}$, принадлежащей суграфу $G_{[\lambda_1, \lambda_2, \dots, \lambda_N]}$, в величину $f(l)$ имеет поря-

док const, а вклад любой дуги u суграфа $G_{\tilde{A}_1}$ — поряд-
док $O(\varepsilon^1)$. Отсюда и из формулы (1.15) следует, что для
получения члена разложения амплитуды $[\tilde{K}^{(n)}(\varepsilon)]_{ij}$ по-
рядка ε^r ($r = 0, 1, \dots, m; m \leq n$) необходимо вычислить
вклад тех и только тех путей длины n на графе $G_{[\lambda_1, \lambda_2, \dots, \lambda_N]} +$
 $+ G_{\tilde{A}_1}$, которые содержат дуги суграфа $G_{\tilde{A}_1}$ ровно r
раз ($r \leq n$).

Обозначим множество всех таких путей длины n через
 $l_{ij}^{(n, r)}(\tilde{A}(\varepsilon))$ ($r = 0, 1, \dots, m$). Очевидно, что справедливо сле-
дующее разбиение множества $\hat{l}_{ij}^{(n)}(\tilde{A})$ всех путей дли-
ны n из вершины x_i в вершину x_j :

$$\hat{l}_{ij}^{(n)}(\tilde{A}(\varepsilon)) = \bigcup_{r=0}^m l_{ij}^{(n, r)}(\tilde{A}(\varepsilon)), \quad l_{ij}^{(n, r)} \cap l_{ij}^{(n, r')} = \emptyset, \quad r \neq r'. \quad (1.17')$$

Разбиение множества $\hat{l}_{ij}^{(n)}(\tilde{A})$ всех путей (1.17') в силу фор-
мулы (1.15) дает следующее представление амплитуды Гри-
на:

$$[\tilde{K}^{(n)}(\varepsilon)]_{ij} = \sum_{l \in \hat{l}_{ij}^{(n)}(\tilde{A}(\varepsilon))} f(l) = \sum_{r=0}^m \sum_{l \in l_{ij}^{(n, r)}(\tilde{A}(\varepsilon))} f(l). \quad (1.17)$$

Сравнивая разложение (1.17) и разложение (1.14) и учиты-
вая, что вклад $f(l) = O(\varepsilon^r)$, если $l \in l_{ij}^{(n, r)}(\tilde{A}(\varepsilon))$, для чле-
нов ряда теории возмущений амплитуды Грина КМ-цепи в
« r, t »-представлении окончательно получаем формулу

$$[\tilde{K}_r^{(n)}]_{ij} = \sum_{l \in l_{ij}^{(n, r)}(\tilde{A}(\varepsilon))} f(l). \quad (1.18)$$

Рассмотрим подробнее подмножество путей $l_{ij}^{(n, r)}(\tilde{A}(\varepsilon))$.
Пусть $r = 1$ и индекс $i \neq j$ ($i = 1, 2, \dots, N$), тогда мат-
ричный элемент $[\tilde{K}_1^{(n)}]_{ij} \neq 0$, если только $\tilde{a}_{ij}^{(1)} \neq 0$. При этом
любой путь $l \in l_{ij}^{(n, 1)}(\tilde{A}(\varepsilon))$ длины n можно представить на под-
графе $G'(\varepsilon^1)$ графа $\hat{G}(\tilde{A}(\varepsilon))$, изображенном на рис. 165, а,
если начальной вершине y_i пути l сопоставить вершину
 ξ_1 графа $G'(\varepsilon^1)$, конечной вершине пути y_j — вершину ξ_2
и ребро (ξ_1, ξ_2) ориентировать от вершины ξ_1 к вершине ξ_2 .

Любой путь $l \in l_{ij}^{(n, 1)}(\tilde{A}(\varepsilon))$ будет тогда однозначно определен, если задать два целых неотрицательных числа: число n_1 — кратность дуги-петли (y_i, y_i) , число n_2 — кратность дуги-петли (y_j, y_j) такие, что вместе с числом 1 — кратностью дуги (y_i, y_j) они определяют разбиение числа n ,

$$n_1 + n_2 + 1 = n. \quad (1.19)$$

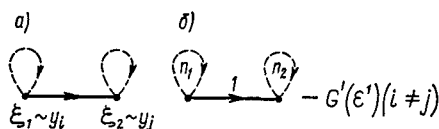


Рис. 165

Для пары чисел n_1 и n_2 , удовлетворяющих условию (1.19), примем обозначение $\langle n_1, n_2 \rangle$ (рис. 165, б). При $i = j$ множество путей $l_{ij}^{(n, 1)}(\tilde{A})$ можно представить на графе $G''(\varepsilon^1)$, изображенном на рис. 166.

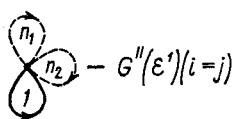


Рис. 166

Рассмотрение графа $G''(\varepsilon^1)$ с двумя пунктирными петлями при вершине ξ_i продиктовано желанием сохранить равенство (1.19) и для случая $i = j$. Кроме этого, как выяснится в дальнейшем, именно такое изображение графа является необходимым для определения « p, E »-представления КМ-цепи (см. § 2). Отметим, что вычисление матричных элементов $[K_r^{(n)}]_{ii}$ по графу $G''(\varepsilon^1)$ не усложняется при наличии двух петель вместо одной.

Целые числа n_1, n_2 и 1, удовлетворяющие условию (1.19) при заданном значении n , будем рассматривать как значения некоторой коцепи C_r^* , $r = 1$, определенной на дугах графа $G'(\varepsilon^1)$. Причем, имея в виду равенство (1.19) для значений коцепи $C_{r=1}^*$, будем говорить, что коцепь $C_{r=1}^*$ удовлетворяет «первому закону Кирхгофа». Последнему соглашению можно придать буквальный смысл (закона Кирхгофа) как «равенства нулю суммы токов в вершине графа». Это будет сделано ниже в § 2 в связи с рассмотрением диаграмм Фейнмана в квантовой механике.

Граф $G'(\varepsilon^1)$ вместе с коцепью $C_{r=1}^*$, удовлетворяющей первому закону Кирхгофа (1.19), назовем *диаграммой те-*

ории возмущения порядка $r = 1$ для КМ-цепи в « p, t »-представлении и будем изображать так, как это показано на рисунке 165,б, отмечая значение коцепи на соответствующих дугах графа.

Обобщим рассмотренное выше определение на диаграммы теории возмущений любого порядка $r (1 \leq r \leq m)$. Предварительно рассмотрим следующее определение.

Определение 1.3. Связный мультиграф G назовем эйлеровым путем относительно двух вершин графа « a » и « b », если в графе G существует эйлеров путь из вершины « a » в вершину « b », и соответственно эйлеровым контуром относительно вершины « c », если существует эйлеров контур, для которого вершина « c » является началом и концом.

Вершину « a », являющуюся началом эйлерова пути, обозначим через ξ_{out} , вершину « b » — через ξ_{in} . В случае эйлерова контура вершину « c » обозначим через $\xi_{out} = \xi_{in}$.

Определение 1.4. Диаграммой $\mathcal{L}_{p,t}(\varepsilon^r)$ теории возмущений порядка r для КМ-цепи в « p, t »-представлении назовем пару объектов $(G_{p,t}(\varepsilon^r), C_{r,p,t}^*)$, где 1) граф $G_{p,t}(\varepsilon^r) = (\Gamma, X = \{\xi_1, \xi_2, \dots, \xi_{s_0}\})$ ($1 \leq s_0 \leq N$) является или эйлеровым контуром относительно одной отмеченной вершины $\xi_{out} = \xi_{in}$, или эйлеровым путем относительно двух отмеченных вершин ξ_{out} и ξ_{in} , множество дуг Γ которого представляет собой объединение двух множеств Γ_{int} и Γ_{ext} .

Множество Γ_{int} состоит из r дуг, среди которых могут быть и петли, называемые внутренними дугами графа;

множество Γ_{ext} состоит из петель, называемых внешними петлями графа, причем число внешних петель $\omega(\xi_i)$ при каждой вершине ξ_i ($i = 1, 2, \dots, s_0$) вычисляется по формулам

$$\omega(\xi_i) = \begin{cases} p_1^{int} = q_1^{int} + 1, & i = 1, \\ p_i, & i = 2, 3, \dots, s_0 - 1, \text{ если } G_{p,t}(\varepsilon^r) \text{ — эйлеров путь} \\ q_{s_0}^{int} = p_{s_0}^{int} + 1, & i = s_0, \end{cases}$$

$$\omega(\xi_i) = \begin{cases} p_1^{int} + 1, & i = 1 \\ p_i^{int} = q_i^{int}, & i = 2, 3, \dots, s_0, \end{cases} \text{ если } G_{p,t}(\varepsilon^r) \text{ — эйлеров контур,}$$

где через p_i^{int} обозначается полустепень исхода вершины ξ_i в подграфе $G^{\text{int}} = (\Gamma_{\text{int}}, X)$ графа $G_{p,t}(\varepsilon^r)$, а через q_i^{int} — полустепень захода вершины ξ_i^* .

Вершины графа $G_{\xi_{\text{in}}}$ и ξ_{out} будем называть внешними вершинами, остальные вершины — внутренними.

2) $C_{r,p,t}^*$ — коцепь со значениями в $\mathbb{Z}_+$, принимающая на дугах графа значения

$$C_r^*(u_i^{\text{int}}) = 1 \quad (i = 1, 2, \dots, r) \quad C_r^*(u_i^{\text{ext}}) = n_i \quad (i = 1, 2, \dots, r+1),$$

которые удовлетворяют условию

$$\sum_{i=1}^{r+1} n_i + r = n. \quad (1.20)$$

Пару $(u, C_r^*(u))$, где u — дуга графа, а $C_r^*(u)$ — значение коцепи C_r^* на этой дуге, будем называть линией диаграммы $\mathcal{L}_{p,t}(\varepsilon^r)$. Набор целых чисел $n_1, n_2, \dots, n_{r+1}$, удовлетворяющих условию (1.20), обозначим через $\langle n_1 | n_2 | \dots | n_{r+1} \rangle$.

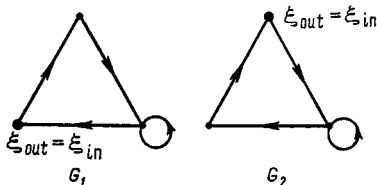


Рис. 167

Определение 1.5. Диаграммы, графы которых различаются только нумерацией внутренних вершин, назовем топологически эквивалентными.

Например, графы G_1 и G_2 (рис. 167) изоморфны в смысле обычного понятия изоморфизма двух графов (см., например, § 2 гл. III), но они не являются топологически эквивалентными, как графы с отмеченными вершинами.

Множество всех топологически неэквивалентных диаграмм порядка $r \{ \mathcal{L}_{p,t}^{(\alpha)}(\varepsilon^r), \alpha = 1, 2, \dots, t(r) \}$ будем обозначать через $\mathcal{L}(r)$.

На рис. 168** приведены в качестве примера все возмож-

* Напомним, что полустепень исхода $p(\xi_i)$ вершины ξ_i — это число дуг, выходящих из вершины ξ_i ; полустепень захода $q(\xi_i)$ — число дуг, заходящих в вершину ξ_i . В формулах для вычисления $\omega(\xi_i)$ использована связь между полустепенями исхода $p(\xi_i)$ и захода $q(\xi_i)$, которая следует для графов, имеющих эйлеров путь (контур). (См.: К. Б е р ж. Теория графов и ее применение. М., ИЛ, 1962).

** Здесь и далее внутренние дуги диаграмм изображаются жирными линиями.

ные топологически неэквивалентные диаграммы второго порядка ($r = 2$), на рис. 169 — ряд диаграмм третьего порядка ($r = 3$).

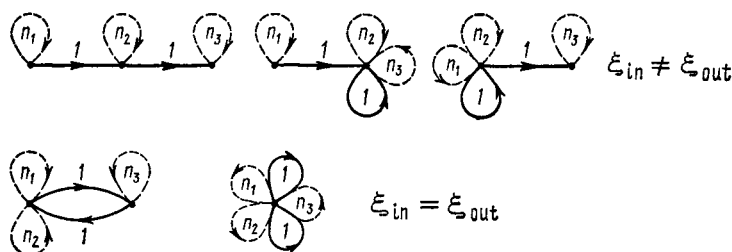


Рис. 168

Задача 1. Найти число $t(r)$ топологически неэквивалентных диаграмм порядка r .

4. Вычисление вкладов диаграмм теории возмущений в « p , t »-представлении. Перейдем теперь к формулировке правил вычисления вкладов диаграмм $\mathcal{L}_{p,t}^{(\alpha)}(\epsilon^r)$ ($\alpha = 1, 2, \dots, t(r)$) порядка r ($r = 0, 1, 2, \dots, m$) в амплитуду КМ-цепи $K_r^{(n)}$ порядка r . Предварительно рассмотрим пример

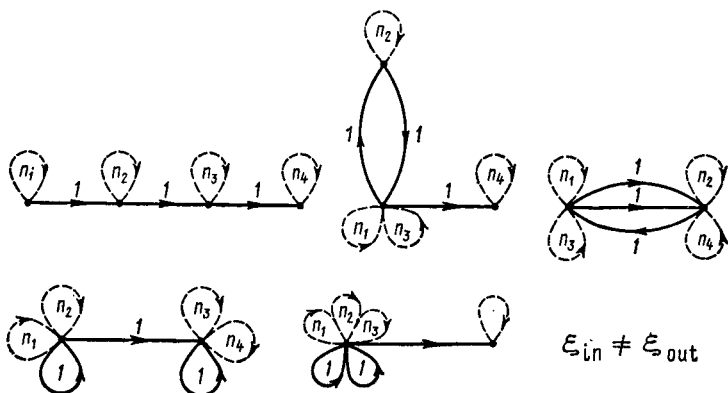


Рис. 169

вычисления вклада диаграммы порядка $r = 3$, изображенной на рис. 170, в матричный элемент $[\tilde{K}_3^{(n)}]_{ij}$ при $i=1, j=2$. Пусть граф $G_{p,t}(\epsilon^r)$ является подграфом графа $G_{\tilde{A}(\epsilon)}$

КМ-цепи $(\tilde{A}(\varepsilon), G_{\tilde{A}(\varepsilon)})$. На графе $G_{p,t}(\varepsilon^3)$ (см. рис. 170) существует эйлеров путь* длины 3 из вершины ξ_{out} в ξ_{in} , дуги которого выделены жирными линиями. Этот эйлеров путь длины 3, состоящий из внутренних дуг графа

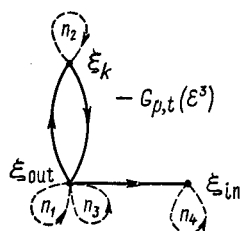


Рис. 170

$G_{p,t}(\varepsilon^r)$, дополняется до пути l длины n из вершины $x_{i=1} = \xi_{\text{out}}$ в вершину $x_{j=2} = \xi_{\text{in}}$ путем присоединения к внутренним трем дугам внешних петель при вершинах ξ_{out} , ξ_k , ξ_{in} в количестве $\tilde{n}_1 = n_1 + n_3$ при вершине ξ_{out} , n_2 при вершине ξ_k , n_4 при вершине $\xi_{\text{in}} = \xi_j$ так, что

$$\tilde{n}_1 + n_2 + n_4 + 3 = n. \quad (1.20a)$$

Вклад каждой внешней петли при вершине равен соответствующему этой вершине собственному значению матрицы $[\lambda_1, \lambda_2, \dots, \lambda_N]$ из (1.16), т. е. при вершине ξ_{out} этот вклад равен λ_i , при ξ_k — λ_k , при ξ_j — λ_j .

Вклад каждой внутренней дуги (ξ_{i_k}, ξ_{j_k}) равен соответствующему матричному элементу матрицы $\tilde{A}_1$ из (1.16). Тогда вклад $f(l)$ отдельного пути l длины n при фиксированных целых числах $\tilde{n}_1, n_2, n_4$, удовлетворяющих условию (1.20a), определяется равенством

$$f(l) = \lambda_i^{\tilde{n}_1} \lambda_k^{n_2} \lambda_j^{n_4} \tilde{a}_{ik} \tilde{a}_{ki} \tilde{a}_{ij} \quad (k \neq i, j). \quad (1.20b)$$

Суммируя выражения (1.20b) по $k = 1, 2, \dots, N$, $k \neq i, j$ и всем целым числам $\tilde{n}_1, n_2, n_4 \in \mathbb{Z}_+$, удовлетворяющим условию (1.20a), получаем для вклада диаграммы $\mathcal{L}$ выражение

$$f(\mathcal{L}_{p,t}(\varepsilon^3)) = \tilde{a}_{ij} \sum_{\substack{1 \leq k \leq N \\ k \neq i, j}} \tilde{a}_{ik} \tilde{a}_{ki} \sum_{\langle \tilde{n}_1 | n_2 | n_4 \rangle} \lambda_i^{\tilde{n}_1} \lambda_k^{n_2} \lambda_j^{n_4}.$$

* Указанный путь, строго говоря, является эйлеровым в суграфе $G_{p,t}(\varepsilon^3)$, состоящем из внутренних дуг диаграммы. Если дополнить этот путь внешними петлями при вершинах, то получим эйлеров путь в графе $G_{p,t}(\varepsilon^3)$.

Учитывая, что $\tilde{n}_1 = n_1 + n_3$, $\lambda_1 = \lambda_3 = \lambda_i$, последнюю формулу запишем в следующем виде, используя « δ -функцию» целочисленного аргумента

$$\delta_{im} = \begin{cases} 1, & i = m \\ 0, & i \neq m \quad (i, m = 1, 2, \dots, N); \end{cases}$$

$$f(\mathcal{L}) = \sum_{\substack{1 \leq k, m \leq N \\ k, m \neq i, j}} \tilde{a}_{ik} \tilde{a}_{km} \tilde{a}_{mj} \delta_{im} \sum_{\langle n_1 | n_2 | n_3 | n_4 \rangle} \lambda_i^{n_1} \lambda_k^{n_2} \lambda_m^{n_3} \lambda_j^{n_4} \quad (1.20в)$$

Формула (1.20в) обобщается на диаграммы любого порядка r ($r = 0, 1, 2, \dots, m$) следующим образом.

П р а в и л о в ы ч и с л е н и я в к л а д о в
д и а г р а м м т е о р и и в о з м у щ е н и й
в а м п л и т у д у Г р и н а К М - ц е п и
в « p, t » - п р е д с т а в л е н и и

1. Если граф диаграммы $\mathcal{L}_{p,t}(\varepsilon')$ — эйлеров путь, то полагаем вклад диаграммы $f(\mathcal{L}_{p,t}(\varepsilon'))$ в амплитуду $[\tilde{K}_r^{(n)}]_{ii}$ равным нулю при $i = 1, 2, \dots, N$ (соответственно полагаем, что $f(\mathcal{L}_{p,t}(\varepsilon')) = 0$ для амплитуды $[\tilde{K}_r^{(n)}]_{ij}$, $i \neq j$, если граф диаграммы $\mathcal{L}_{p,t}(\varepsilon')$ — эйлеров контур).

2. Если граф диаграммы $\mathcal{L}_{p,t}(\varepsilon')$ — эйлеров путь и $i \neq j$, то для того чтобы определить вклад диаграммы $f(\mathcal{L}_{p,t}(\varepsilon'))$ в амплитуду $[\tilde{K}_r^{(n)}]_{ij}$, $i \neq j$, необходимо:

а) присвоить вершине ξ_{out} индекс i ($\xi_{\text{out}} = \xi_i$), а вершине ξ_{in} — индекс j ($\xi_{\text{in}} = \xi_j$) и провести эйлеров путь l длины r из вершины ξ_{out} в вершину ξ_{in} в подграфе $G^{\text{int}} = (X, \Gamma_{\text{int}})$ графа $G_{p,t}(\varepsilon')$; обозначить через $\{\xi_i, \xi_{i_2}, \xi_{i_3}, \dots, \xi_{i_r}, \xi_j\} = \{\xi_{i_k}, i_k = 1, 2, \dots, N\}$ последовательность вершин, определяющих путь l ;

б) сопоставить индексу i_k вершины ξ_{i_k} ($k = 1, 2, \dots, r$, $r + 1$; $i_1 = i$, $i_{r+1} = j$) число $\lambda_{i_k}^{n_k}$:

$$i_k \rightarrow \lambda_{i_k}^{n_k} \quad (k = 1, 2, \dots, r, r + 1),$$

где λ_{i_k} — собственное значение матрицы A_0 , n_k — значение коцепи C_r^* на внешней петле в вершине ξ_{i_k} ;

в) сопоставить паре индексов (i_k, i_{k+1}) двух последовательных вершин ξ_{i_k} и $\xi_{i_{k+1}}$ ($i_1 = i, i_{r+1} = j, k = 1, 2, \dots, r+1$) элемент $\tilde{a}_{i_k i_{k+1}}$ матрицы $\tilde{A}_1(\epsilon)$, если $\xi_{i_k} \neq \xi_{i_{k+1}}$, и элемент $\tilde{a}_{i_k i_{k+1}}$, умноженный на « δ -функцию» целочисленного аргумента

$$\delta_{i_k i_{k+1}} = \begin{cases} 1, & i_k = i_{k+1}, \text{ если } \xi_{i_k} = \xi_{i_{k+1}}; \\ 0, & i_k \neq i_{k+1} \end{cases}$$

$$(i_k, i_{k+1}) \rightarrow \tilde{a}_{i_k i_{k+1}}, \text{ если } \xi_{i_k} \neq \xi_{i_{k+1}},$$

$$(i_k, i_{k+1}) \rightarrow \tilde{a}_{i_k i_{k+1}} \delta_{i_k i_{k+1}}, \text{ если } \xi_{i_k} = \xi_{i_{k+1}}.$$

Тогда вклад $f(\mathcal{L}_{p,t}(\epsilon^r))$ диаграммы $\mathcal{L}_{p,t}(\epsilon^r)$ определяется формулой

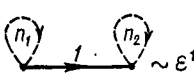
$$f(\mathcal{L}_{p,t}(\epsilon^r)) = \overbrace{\sum_{1 < i_2, i_3, \dots, i_r < N} \dots \sum}^{r-1} \tilde{a}_{ii_2} \tilde{a}_{i_2 i_3} \dots \tilde{a}_{i_r j} \left(\prod \delta_{i_k i_{k+1}} \right) \times \\ \times \sum_{\langle n_1 | n_2 | \dots | n_r \rangle} \lambda_i^{n_1} \lambda_{i_2}^{n_2} \dots \lambda_{i_r}^{n_r} \lambda_j^{n_{r+1}}, \quad (1.21)$$

где через $\prod \delta_{i_k i_{k+1}}$ обозначено произведение « δ -функций» по всевозможным парам индексов, совпадающих вершин в пути l , N — порядок матрицы $A(\epsilon)$ в КМ-цепи $(A(\epsilon), G_A(\epsilon))$.

3. Если граф диаграммы $\mathcal{L}$ — эйлеров контур, то вычисление вклада $f(G)$ производится по формуле (1.21) при $i = j$.

З а м е ч а н и е. Из формулы (1.21) следует, что если r_0 — число внутренних петель графа $G_{p,t}(\epsilon^2)$ ($r_0 \leq r$), то фактическая кратность суммы по индексам $i_2, i_3, \dots, i_r$ равна $r - r_0 - 1$.

В качестве примера вычислим вклады ряда диаграмм порядка $r = 2, 3$ ($\xi_{\text{out}} \neq \xi_{\text{in}}$). Выпишем результаты этих вычислений:



$$f(\mathcal{L}) = \tilde{a}_{ij} \sum_{\langle n_1 | n_2 \rangle} \lambda_i^{n_1} \lambda_j^{n_2},$$

$$\begin{array}{c} \textcircled{n_1} \xrightarrow{1} \textcircled{n_2} \xrightarrow{1} \textcircled{n_3} \sim \varepsilon^2 \end{array}$$

$$f(\mathcal{L}) = \sum_{\substack{1 \leq i_2 \leq N \\ i_2 \neq i, j}} \tilde{a}_{ii_2} \tilde{a}_{i_2 j} \left(\sum_{\langle n_1 | n_2 | n_3 \rangle} \lambda_i^{n_1} \lambda_{i_2}^{n_2} \lambda_j^{n_3} \right),$$

$$\begin{array}{c} \textcircled{n_1} \textcircled{n_2} \xrightarrow{1} \textcircled{n_3} \sim \varepsilon^2 \\ \textcircled{1} \end{array} \quad f(\mathcal{L}) = \tilde{a}_{ii} \tilde{a}_{ij} \sum_{\langle n_1 | n_2 | n_3 \rangle} \lambda_i^{n_1} \lambda_{i_2}^{n_2} \lambda_j^{n_3} =$$

$$= \tilde{a}_{ii} \tilde{a}_{ij} \sum_{\langle m_1 | n_3 \rangle} \lambda_i^{m_1} \lambda_j^{n_3},$$

$$(m_1 = n_1 + n_2)$$

$$\begin{array}{c} \textcircled{n_1} \xrightarrow{1} \textcircled{n_2} \textcircled{n_3} \sim \varepsilon^2 \\ \textcircled{1} \end{array} \quad f(\mathcal{L}) = \tilde{a}_{ij} \tilde{a}_{jj} \sum_{\langle n_1 | m_2 \rangle} \lambda_i^{n_1} \lambda_j^{m_2},$$

$$(m_2 = n_2 + n_3),$$

$$\begin{array}{c} \textcircled{n_1} \xrightarrow{1} \textcircled{n_2} \xrightarrow{1} \textcircled{n_3} \xrightarrow{1} \textcircled{n_4} \sim \varepsilon^3 \end{array}$$

$$f(\mathcal{L}) = \sum_{\substack{1 \leq i_2, i_3 \leq N \\ i_2, i_3 \neq i, j}} \tilde{a}_{ii_2} \tilde{a}_{i_2 i_3} \tilde{a}_{i_3 j} \left(\sum_{\langle n_1 | n_2 | n_3 | n_4 \rangle} \lambda_i^{n_1} \lambda_{i_2}^{n_2} \lambda_{i_3}^{n_3} \lambda_j^{n_4} \right),$$

$$\begin{array}{c} \textcircled{n_1} \textcircled{n_2} \\ \textcircled{n_3} \textcircled{n_4} \end{array} \xrightarrow{1} \xrightarrow{1} \xrightarrow{1} \sim \varepsilon^3$$

$$f(\mathcal{L}) = a_{ij}^2 \tilde{a}_{ji} \sum_{\langle m_1 | m_2 \rangle} \lambda_i^{m_1} \lambda_j^{m_2},$$

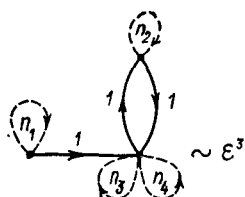
$$(m_1 = n_1 + n_3, m_2 = n_2 + n_4),$$

$$\begin{array}{c} \textcircled{n_2} \\ \textcircled{n_1} \textcircled{n_3} \end{array} \xrightarrow{1} \xrightarrow{1} \xrightarrow{1} \textcircled{n_4} \sim \varepsilon^3$$

$$f(\mathcal{L}) = \tilde{a}_{ij} \sum_{1 \leq k \neq i, j \leq N} \tilde{a}_{ik} \tilde{a}_{ki} \times$$

$$\times \left(\sum_{\langle m_1 | n_2 | n_4 \rangle} \lambda_i^{m_1} \lambda_k^{n_2} \lambda_j^{n_4} \right),$$

$$(m_1 = n_1 + n_3)$$



$$f(\mathcal{L}) = \tilde{a}_{ij} \sum_{1 \leq k \neq i, j \leq N} \tilde{a}_{jk} \tilde{a}_{kj} \times \\ \times \left(\sum_{\langle n_1 | n_2 | m_3 \rangle} \lambda_i^{n_1} \lambda_k^{n_2} \lambda_j^{n_3} \right) \\ (m_3 = n_3 + n_4).$$

Введение диаграмм $\mathcal{L}_{p, i}(\varepsilon^2)$ теории возмущений позволяет при вычислении r -го приближения $K_r^{(n)}$ амплитуды Грина КМ-цепи заменить суммирование вкладов всех путей $l \in l_{ij}^{(n, r)}$ на суммирование вкладов всех топологически неэквивалентных диаграмм порядка r , т. е. для амплитуды $K_r^{(n)}$ справедлива формула

$$|\tilde{K}_r^{(n)}|_{ij} \equiv \sum_{l \in l_{ij}^{(n, r)}(\tilde{A})} f(l) = \sum_{\alpha=1}^{t(r)} f(\mathcal{L}_{p, i}^{(\alpha)}(\varepsilon^r)) \quad (r = 0, 1, \dots, m), \quad (1.22)$$

где $t(r)$ — число неэквивалентных диаграмм порядка r .

Отсюда и из формулы (1.3) окончательно приходим к выводу, что полная амплитуда Грина КМ-цепи в « p, t »-представлении является суммой вкладов всех диаграмм во всех порядках r ($r = 0, 1, 2, \dots, m$) теории возмущений:

$$|\tilde{K}^{(n)}|_{ij} = \sum_{r=0}^m \sum_{\alpha=1}^{t(r)} f(\mathcal{L}_{p, i}^{(\alpha)}(\varepsilon^r)) = \sum_{r=0}^m f(\mathcal{L}(r)). \quad (1.23)$$

Формулу (1.23) совместно с определением диаграмм $\mathcal{L}_{p, i}(\varepsilon^r)$ и правилом вычисления их вкладов можно принять за определение полной амплитуды Грина КМ-цепи в « p, t »-представлении. Это новое определение амплитуды КМ-цепи, разумеется, эквивалентно определению амплитуды $\tilde{K}(\varepsilon)$ как n -й степени матрицы $\tilde{A}(\varepsilon)$, $\tilde{K}(\varepsilon) = (\tilde{A}(\varepsilon))^n$.

Но важно отметить следующее обстоятельство при переходе к пределу от дискретной модели к непрерывной. Например, в квантовой теории поля определение полной амплитуды перехода системы из одного состояния в другое (так называемой S -матрицы) связано с определением этой амплитуды в виде ряда теории возмущений по константе взаимодействия. Каждый член этого ряда представляет собой сумму вкладов всех топологически неэкви-

валентных диаграмм соответствующего порядка r ($r = 0, 1, 2, \dots, \infty$), допускаемых в квантовой теории поля (диаграммы квантовой теории поля будут определены ниже в § 3). При этом дополнительно возникает серьезная проблема суммирования ряда теории возмущений и так называемая проблема расходимостей некоторых членов этого ряда [1]. По-видимому, возможное обобщение понятия КМ-цепи на случай квантовой теории поля при обосновании соответствующего предельного перехода к непрерывным величинам позволило бы корректным образом вычислить амплитуду перехода.

В рассматриваемой в этом параграфе дискретной ситуации представление амплитуды Грина КМ-цепи в виде конечного ряда теории возмущений (1.23) оказывается полезным представлением для вычисления матричных элементов $[K^{(n)}]_{ij}$, если заметить следующее обстоятельство.

Графы диаграмм теории возмущений КМ-цепи могут быть отождествлены с подграфами графа $\hat{G}_{\sim A(\epsilon)}$ КМ-цепи в « p, t »-представлении. В самом деле, каждый граф $G_{p, t}(\epsilon')$ диаграммы порядка r представляет собой некоторый путь l длины n на графе $\hat{G}_{\sim A(\epsilon)}$, содержащий ровно r дуг графа $G_{\sim A_1(\epsilon)}$. Если этот путь простой, то множество вершин и дуг, входящих в него, является подграфом графа $\hat{G}_{\sim A(\epsilon)}$, который и отождествляется с графом $G_{p, t}(\epsilon')$. Если этот путь не простой, то он содержит в точности один простой путь, подграф которого отождествляется с графом диаграммы. При таком отождествлении двум графам диаграмм разного порядка может отвечать один и тот же подграф графа $G_{\sim A(\epsilon)}$.

Например, на рис. 171 приведены две диаграммы $G_{p, t}(\epsilon^2)$ и $G_{p, t}(\epsilon^4)$ второго и четвертого порядков соответственно, которым на графе $\hat{G}_{\sim A(\epsilon)}$, изображенном на рис. 172, отвечает один и тот же подграф, выделенный жирной линией.

Это соответствие графов диаграмм теории возмущений и подграфов графа КМ-цепи $G_{\sim A(\epsilon)}$ позволяет при вычислении амплитуды $K^{(n)}$ по формуле (1.12) не рассматривать вклады тех диаграмм, графы $G_{p, t}(\epsilon')$ которых не содержат

ся среди возможных подграфов графа $\hat{G}_{\tilde{A}(\epsilon)}$. Ради удобства терминологии в этом случае будем говорить, что вклады диаграмм, графы которых не являются подграфами графа $\hat{G}_{\tilde{A}(\epsilon)}$, равны нулю.

Заметим, что вклад диаграмм $\mathcal{L}_{p, i}(\epsilon')$ может быть равен нулю за счет интерференции членов в формуле (1.21), хотя

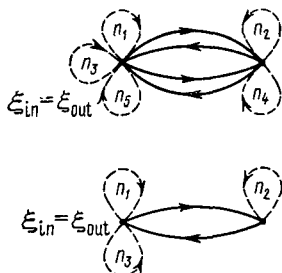


Рис. 171

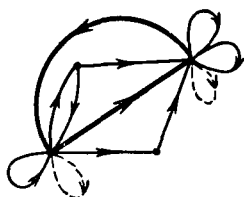


Рис. 172

граф диаграммы $\mathcal{L}_{p, i}(\epsilon')$ является подграфом графа $\hat{G}_{\tilde{A}(\epsilon)}$.

Таким образом, в конкретных задачах число диаграмм, вклады которых необходимо учитывать при вычислении амплитуды $[\tilde{K}_r^{(n)}]_{ij}$ по формуле (1.12), может быть весьма ограниченным. Это число зависит от конкретной структуры матрицы $\tilde{A}(\epsilon) = [\lambda_1, \dots, \lambda_N] + \epsilon \tilde{A}_1(\epsilon)$ — « p »-представления матрицы $A(\epsilon)$ КМ-цепи $(A(\epsilon), G_{A(\epsilon)})$.

В качестве примера приведем следующие два утверждения.

1. Если связный граф $G_{\tilde{A}_1}$ матрицы $\tilde{A}_1$ удовлетворяет условию

$$p(x_i) \cdot q(x_i) = 0 \quad (i = 1, 2, \dots, N)^*, \quad (1.24)$$

* Здесь $p(x_i)$ и $q(x_i)$ — полустепени исхода и захода вершины x_i в графе $G_{\tilde{A}_1}$.

то в разложении (1.1) матричный элемент

$$[K_r^{(n)}]_{ij} = 0 \quad (i, j = 1, 2, \dots, N) \quad (r = 2, 3, \dots). \quad (1.25)$$

Действительно, из условия (1.24) следует отсутствие в графе $G_{\tilde{A}_1}$ путей длины два. Отсюда и из определения $[K_r^{(n)}]_{ij}$ в формуле (1.18) следует равенство (1.25).

2. Если матричные диагональные элементы матрицы $\tilde{A}_1$

$$\tilde{a}_{ii} = 0, \quad (i = 1, 2, \dots, N),$$

то вклады всех диаграмм, для которых число внутренних петель $r_0 \neq 0$, равны нулю.

Доказательство очевидно в силу определения графа $\hat{G}_{A(\epsilon)}^{\sim}$, мультипликативности вклада $f(l)$, $l \in l_{ij}^{(n,r)}(\tilde{A})$ и следованного выше замечания о связи подграфов графа $\hat{G}_{A(\epsilon)}^{\sim}$ и графов диаграмм теории возмущений.

Приведенные утверждения сформулируем в терминах свойств матрицы $A = A_0 + \epsilon A_1$.

Теорема 1.1.* Пусть A_0 и A_1 — матрицы с вещественными коэффициентами и пусть A_0 — симметрическая в $\mathbb{R}^N$ матрица, а матрица A_1 удовлетворяет условиям:

а) $A_1^2 \equiv 0$, б) коммутатор $[A_0, A_1] = 0$

Тогда и только тогда в разложении (1.3) матрица

$$K_r^{(n)} \equiv 0 \quad (r = 2, 3, \dots).$$

Теорема 1.2. Если A_0 — эрмитова, а A_1 — косоэрмитова матрицы, то вклады диаграмм теории возмущений, имеющих хотя бы одну внутреннюю петлю при вершине, равны нулю.

Задача. Доказать справедливость утверждений теорем 1.1 и 1.2, используя диаграммную технику.

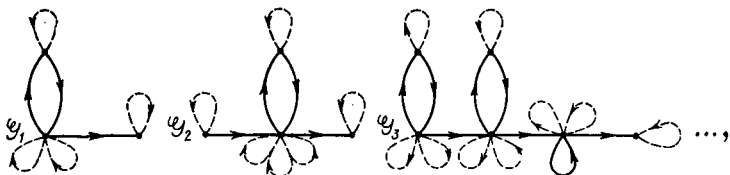
Отправляясь от свойств матрицы A_0 и A_1 , задающих КМ-цепь $(A(\epsilon), G_{A(\epsilon)})$, можно дать ряд достаточных условий, когда вклад диаграмм определенного типа равен нулю.

Примером может служить следующая теорема.

Теорема 1.3. Пусть A_0 — эрмитова матрица порядка

* Более общий случай см. в [3].

$(N \times N)$, $[A_0, A_1] = 0$, матрица A_1 удовлетворяет условию $A_1^{\text{tr}} = I A_1$, $I^2 = -E$ (E — единичная матрица).



Тогда вклад диаграмм вида G_1, G_2, G_3 в амплитуду $[\tilde{K}^{(n)}]_{ij}$ равен нулю при любых $i, j = 1, 2, \dots, N$ ($i \neq j$).

Доказательство теоремы относительно вклада диаграммы $\mathcal{L}_1$ немедленно следует из формулы (1.21) и свойств матрицы A_1 .

Доказательство теоремы для диаграмм $\mathcal{L}_2, \mathcal{L}_3, \dots$, и т. д. вытекает из следующего факта.

Если граф $G_\mu(\epsilon^r)$ диаграмм $\mathcal{L}_{p,t}(\epsilon^r)$ содержит в качестве подграфа граф $G_\nu(\epsilon^{r'})$ диаграммы $\mathcal{L}_{p,t}(\epsilon^{r'})$ ($r' \leq r$), вклад которой равен нулю, то и вклад диаграммы $\mathcal{L}_{p,t}(\epsilon^r)$ тоже равен нулю. Последнее утверждение непосредственно следует из формулы (1.21). ■

Используем полученные результаты для разложения по малому параметру ϵ ($\epsilon \ll 1$) КМ-цепи фотона, определенной в § 3 гл. I.

5. Теория возмущений КМ-цепи фотона. Матрица $Z(\epsilon)$ КМ-цепи фотона ($Z(\epsilon), G_{Z(\epsilon)}$) определяется равенством

$$Z(\epsilon) = A_1^{-1}(\epsilon) A_0(\epsilon), \quad (1.26)$$

где матрицы $A_1(\epsilon)$ и $A_0(\epsilon)$ имеют вид (3.8).

Для получения разложения амплитуды Грина $K^{(n)}(\epsilon) = (Z(\epsilon))^n$ в ряд теории возмущений (1.1) необходимо разложить в ряд по ϵ матрицу $Z(\epsilon)$,

$$Z(\epsilon) = Z_0 + \epsilon Z_1 + \dots + \epsilon^m Z_m + O(\epsilon^{m+1}) \quad (m \leq N).$$

Используя явный вид матриц $A_1(\epsilon)$ и $A_0(\epsilon)$ из (3.8), представим их в виде:

$$\begin{aligned} A_0(\epsilon) &= B_1 - \epsilon \alpha B_2, \\ A_1(\epsilon) &= B_1 + \epsilon \alpha B_2, \end{aligned} \quad (1.27)$$

где матрицы B_1 и B_2 уже не зависят от ε и их элементы определяются равенствами

$$B_1 = ((b_{ij}))_{N \times N}, \quad b_{ij} = \begin{cases} +1, & i = j-1, j \\ 0, & i \neq j-1, j \end{cases} \quad (i, j = 1, 2, \dots, N), \quad (1.28)$$

$$b_{01} = b_{N1} = 1$$

$$B_2 = ((c_{ij}))_{N \times N},$$

$$c_{ij} = \begin{cases} -1, & i = j-1, \\ -1, & i = j, \\ 0, & i \neq j, j-1 \end{cases} \quad (i, j = 1, 2, \dots, N).$$

$$c_{01} = c_{N1} = -1$$

Из формул (1.27) и (3.9) при $\varepsilon = 0$ следует, что

$$\det B_1 = \det A_1(0) = 2 \quad (N - \text{нечетное}). \quad (1.28')$$

Воспользуемся известной [6] формулой конечномерной теории возмущений

$$(B_1 + \varepsilon \alpha B_2)^{-1} = B_1^{-1} (I + \varepsilon \alpha B_2 B_1^{-1}) =$$

$$= B_1^{-1} \sum_{k=0}^{\infty} (-1)^k (\varepsilon \alpha)^k (B_2 B_1^{-1})^k. \quad (1.29)$$

Тогда для матрицы $Z(\varepsilon)$ с учетом (1.27) и (1.29) получаем разложение

$$Z(\varepsilon) = I - 2\varepsilon \alpha B_1^{-1} B_2 + O(\varepsilon^2). \quad (1.30)$$

Поскольку матрица $Z_0 = 1$ — диагональна, то матрица $Z(\varepsilon)$ из (2.35) является матрицей КМ-цепи в « p, t »-представлении ($Z(\varepsilon) \equiv \tilde{Z}(\varepsilon)$) и соответственно КМ-цепь $(Z(\varepsilon), G_{Z(\varepsilon)})$ в « p, t »-представлении имеет вид $(Z(\varepsilon), \hat{G}_{Z(\varepsilon)})$, где

$$\hat{G}_{Z(\varepsilon)} = G_{[1, 1, \dots, 1]} + G_{Z_1(\varepsilon)},$$

$$Z_1(\varepsilon) = -2\varepsilon \alpha B_1^{-1} B_2 + O(\varepsilon^2).$$

При $\varepsilon \rightarrow 0$ граф $G_{Z_1(\varepsilon)}$ матрицы $Z_1(\varepsilon)$ совпадает с графом $G_{B_1^{-1}B_2}$ матрицы $B_1^{-1}B_2$ так, что в первом порядке по ε амплитуда Грина $K_1^{(n)}$ полностью определяется вкладом диаграмм первого порядка, являющимися подграфами графа $G_{[1,1,\dots,1]} + G_{B_1^{-1}B_2}$, причем величина вклада дуги суграфа $G_{B_1^{-1}B_2}$ равна соответствующему матричному элементу матрицы $B_1^{-1}B_2$.

Вычислим матричные элементы матрицы $B_1^{-1}B_2 = Z_1$. Из (1.28') и (1.28'') получаем, что матричный элемент $[Z_1]_{ij}$ определяется соотношением

$$[Z_1]_{ij} = \sum_{k=1}^N b_{ik}^{-1} c_{kj} = b_{ij}^{-1} - b_{ij-1}^{-1} = (-1)^{i+j} - \delta_{ij} =$$

$$= \begin{cases} 0, & i = j \\ \pm 1, & i \neq j \end{cases} \quad (i, j = 1, 2, \dots, N), \quad (1.31)$$

где b_{ij}^{-1} ($i, j = 1, 2, \dots, N$) — элементы матрицы B_1^{-1} . При выводе формулы (1.31) использована следующая связь между элементами матрицы B_1^{-1} :

$$b_{ij}^{-1} + b_{ij-1}^{-1} = \delta_{ij} \quad (i, j = 1, 2, \dots, N, \quad b_{01} = b_{N1} = 1).$$

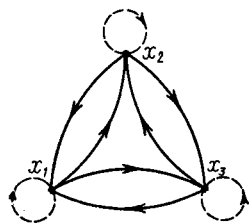


Рис. 173

Из формулы (1.31) и определения графа матрицы следует, что граф G_{Z_1} матрицы $Z_1 = B_1^{-1}B_2$ — полный симметрический граф без петель, а $G_{Z(\varepsilon)} = G_{[1,1,\dots,1]} + G_{B_1^{-1}B_2}$ — граф матрицы

$Z(\varepsilon)$ в «р, t»-представлении — это полный симметрический граф с петлями при каждой вершине вида, изображенного на рис. 173 (при $N = 3$). Так как $G_{Z(\varepsilon)}$ — полный граф, то он содержит в качестве своих подграфов

графы всех топологически неэквивалентных диаграмм $\mathcal{L}_{p,t}^{(\alpha)}(\varepsilon^r)$ ($\alpha = 1, 2, \dots, t(r)$ при $r \leq N$). Поэтому каждому члену разложения $K_r^{(n)}$ ($r = 0, 1, \dots, m, m \leq N$) амплитуды Грина $K^{(n)}(\varepsilon) = (Z(\varepsilon))^n$ в ряд (1.1) в данном случае отвечают диаграммы всех топологических типов при каждом r ($r = 0, 1, \dots, m, m \leq N$).

Вклад этих диаграмм вычисляется по формуле (1.21), в которой нужно положить $\lambda_i = 1$ ($i = 1, 2, \dots, N$), $\tilde{a}_{i_k i_{k+1}} = [Z_1]_{i_k i_{k+1}}$ ($i_k, i_{k+1} = 1, 2, \dots, N$). При $r = 0, 1$ вклады всех топологически неэквивалентных диаграмм полностью определяют по формуле (1.23) члены разложения (1.1) $K_0^{(n)}$ и $K_1^{(n)}$:

$$|K_0^{(n)}|_{ii} \sim \text{diagram 1}$$

$$|K_1^{(n)}|_{ij} \sim \text{diagram 2}$$

Diagram 1: A single loop with a vertical line through its center, labeled $i=j$.

Diagram 2: Two loops connected by a horizontal line. The left loop is labeled $i=j$ and the right loop is labeled $i \neq j$.

§ 2. ДИАГРАММЫ ФЕЙНМАНА В НЕРЕЛЯТИВИСТСКОЙ КВАНТОВОЙ МЕХАНИКЕ КАК ДИАГРАММЫ ТЕОРИИ ВОЗМУЩЕНИЙ КМ-ЦЕПИ В « p, E »-ПРЕДСТАВЛЕНИИ

1. Связь нестационарной теории возмущений в квантовой механике с диаграммами КМ-цепи в « p, t »-представлении. Пусть матрица Z_0^* такова, что амплитуда Грина $K_0^{(n)}$ (τ, h) КМ-цепи (Z_0, G_{Z_0}) при $h \rightarrow 0$ и $\tau = \tau(h) \rightarrow 0$ сходится к функции $K_0(x, t)$, удовлетворяющей уравнению Шредингера для свободного движения квантовой системы

$$i\alpha \frac{\partial K_0}{\partial t} = -\alpha^2 \Delta K_0 \equiv H_0 K_0 \quad (2.1)$$

при некоторой постоянной $\alpha > 0$. (В дальнейшем рассматривается система единиц, в которой $\alpha = 1$.) Существование такой КМ-цепи следует из результатов работы [5].

Сходимость амплитуды Грина $K_0^{(n)}$ (τ, h) к функции $K_0(x, t)$ определяется аналогично тому, как было рассмотрено в § 3 гл. I для модельного уравнения фотона $\frac{\partial u}{\partial t} = c \frac{\partial u}{\partial x}$.

Рассмотрим теперь регулярное возмущение КМ-цепи (Z_0, G_{Z_0}) , т. е. такую КМ-цепь $(Z(\epsilon), G_{Z(\epsilon)})$, аналитически

* В этом параграфе рассматриваются унитарные матрицы.

зависящую от малого параметра ε , представляющего собой константу взаимодействия частицы с внешним полем, что справедливо разложение

$$Z(\varepsilon) = Z_0 + \sum_{k=1}^n \varepsilon^k Z_k + O(\varepsilon^{n+1}). \quad (2.2)$$

Для простоты будет рассмотрен случай, когда все $Z_k \equiv 0$ при $k \geq 2$ и $O(\varepsilon^{n+1}) \equiv 0$.

Предположим, что существует предел в указанном выше смысле КМ-цепи $(Z(\varepsilon), G_{Z(\varepsilon)})$ при $\hbar \rightarrow 0$ и $\tau = \tau(\hbar) \rightarrow 0$ и пусть коммутатор

$$[Z_0, Z_1] \neq 0^*. \quad (2.3)$$

Тогда очевидно, что предел амплитуды Грина КМ-цепи $(Z(\varepsilon), G_{Z(\varepsilon)})$ — функция $K(x, t, \varepsilon)$ будет удовлетворять уравнению Шредингера с потенциалом $\varepsilon v(x, t, \varepsilon) \neq 0^{**}$, где $v(x, t, \varepsilon)$ — аналитическая функция по ε , гладким образом зависящим от $x \in \mathbb{R}^3$, $t \in \mathbb{R}^1$:

$$i \frac{\partial K}{\partial t} = -\Delta K + \varepsilon v(x, t, \varepsilon) K. \quad (2.4)$$

Пусть при $t = t_0$ функция $K(x, t, \varepsilon)$ удовлетворяет условию

$$K(x, t, \varepsilon)|_{t=t_0} = \delta(x - \xi). \quad (2.5)$$

Решение уравнения (2.4) при $t > t_0$, удовлетворяющее условию (2.5), называется *функцией Грина задачи Коши*

* Условие (2.3) исключает из рассмотрения тривиальный случай диаграмм теории возмущений КМ-цепи в $\langle p, t \rangle$ -представлении, когда матрицы Z_0 и Z_1 одновременно приводятся к диагональному виду так, что граф G_{Z_1} матрицы $\tilde{Z}_1 U^{-1} Z_1$ ($Z_0 U = [\lambda_1 \dots \lambda_N] U$) состоит из изолированных вершин с петлями при каждой.

** Для того чтобы рассмотреть случай, когда потенциал взаимодействия v зависит от времени t , предполагается, что КМ-цепь не является однородной, т. е. условная амплитуда перехода $[Z]_{ij}$ из состояния (x_i, t_{k_0}) в состояние (x_j, t_{k_0+1}) зависит от k_0 ($k_0 = 0, 1, \dots, n-1$). Все результаты, полученные для однородной КМ-цепи, остаются в силе; при этом следует лишь учесть, что амплитуда Грина неоднородной КМ-цепи определяется не степенью матрицы $Z(\varepsilon)$, а произведением матриц $\prod_{k=0}^{n-1} Z(t_k(\varepsilon))$.

для уравнения Шредингера и обозначается через $K(x, \xi; t, t_0; \varepsilon)$.

Для решения задачи (2.4), (2.5) справедливо следующее разложение в ряд теории возмущений по малому параметру ε :

$$\begin{aligned} K(x, \xi; t, t_0; \varepsilon) = & K_0(x, \xi; t, t_0, 0) + \\ & + \left(\frac{\varepsilon}{i}\right) \int_{t_0}^t e^{-iH_0(t-t_1)} v(x, t_1, 0) e^{-iH_0(t_1-t_0)} \delta(x - \xi) dt_1 + \\ & + \left(\frac{\varepsilon}{i}\right)^2 \int_{t_0}^t e^{-iH_0(t-t_2)} v(x, t_2, 0) \int_{t_0}^{t_2} e^{-iH_0(t_2-t_1)} v(x, t_1, 0) \times \\ & \times e^{-iH_0(t_1-t_0)} \delta(x - \xi) dt_1 dt_2 + \frac{\varepsilon^2}{i} \int_{t_0}^t e^{-iH_0(t-t_1)} \frac{\partial v}{\partial \varepsilon}(x, t, 0) \times \\ & \times e^{-iH_0(t_1-t_0)} \delta(x - \xi) dt_1 + O(\varepsilon^3), \quad (2.6) \end{aligned}$$

где выражение $e^{-iH_0(t-t_0)} f(x) = u(x, t)$ представляет собой запись решения задачи Коши

$$i \frac{\partial u}{\partial t} = H_0 u, \quad u(x, t)|_{t=t_0} = f(x).$$

Как увидим, разложение (2.6) допускает естественную аналогию в дискретной ситуации, в связи с чем можно будет дать однозначное графическое представление каждого члена ряда теории возмущений. А именно, разложению (2.6) решения $K(x, \xi; t, t_0; \varepsilon)$ уравнения Шредингера от-

вечает разложение амплитуды Грина $K^{(n)}(\tau, h, \varepsilon) = \prod_{k_0=0}^{n-1} Z(t_{k_0}, \varepsilon)$ соответствующей КМ-цепи в ряд по степеням ε .

Для получения нулевого и первого членов разложения амплитуды $K^{(n)}(\tau, h, \varepsilon) \equiv K^{(n)}(\varepsilon)$, очевидно, достаточно рассмотреть разложение в ряд по ε произведения матриц $\prod_{k_0=0}^{n-1} (Z_0 + \varepsilon Z_1(t_{k_0}))$. Используя результаты § 1, в частности, очевидное обобщение формулы (1.21) на случай неоднородной КМ-цепи, для амплитуды Грина $[\tilde{K}^{(n)}(\varepsilon)]_{ij}$ в « p ,

t -представлении в первом порядке теории возмущений по ε получаем

$$[\tilde{K}_1^{(n)}]_{ij} = \sum_{\langle n_1 | n_2 \rangle} \lambda_j^{n_2} [\tilde{Z}_1(t_{n_1})]_{ij} \lambda_i^{n_1}, \quad (2.7)$$

где $\tilde{Z}_1(t_{n_1})$ — матрица $Z_1(t_{n_1})$ в « p »-представлении,

$$\tilde{Z}_1(t_{n_1}) = U^{-1} Z_1(t_{n_1}) U, \quad Z_0 U = U [\lambda_1, \dots, \lambda_N] \quad (\text{см. § 1}).$$

Формула (2.7) для амплитуды Грина КМ-цепи является дискретным аналогом первого (по ε) члена разложения (2.6). Для того чтобы в этом убедиться, необходимо в формуле (2.6) перейти к « p , t »-представлению $\tilde{K}(p, p'; t, t_0; \varepsilon)$ функции Грина $K(x, \xi; t, t_0; \varepsilon)$, используя преобразование Фурье F по переменным $x \in \mathbb{R}^3$ и $\xi \in \mathbb{R}^3$,

$$\begin{aligned} \tilde{K}(p, p'; t, t_0; \varepsilon) &= F_{\xi \rightarrow p'} F_{x \rightarrow p} K(x, \xi; t, t_0; \varepsilon) = \\ &= \frac{1}{(2\pi)^3} \int \int_{\mathbb{R}^3 \mathbb{R}^3} e^{-ipx} e^{ip'\xi} K(x, \xi; t, t_0; \varepsilon) dx d\xi. \end{aligned}$$

Для функции Грина в « p , t »-представлении из (2.6) тогда получаем разложение

$$\begin{aligned} \tilde{K}(p, p'; t, t_0; \varepsilon) &= e^{-ip^2(t-t_0)} \delta(p - p') + \\ &+ \left(\frac{\varepsilon}{i}\right) \int_{t_0}^t e^{-ip^2(t-t_1)} \tilde{v}(p - p', t_1, 0) e^{-ip'^2(t_1-t_0)} dt_1 + \\ &+ \left(\frac{\varepsilon}{i}\right)^2 \int_{t_0}^t e^{-ip^2(t-t_1)} \int_{\mathbb{R}^1} \tilde{v}(p - q, t_2, 0) \int_{t_0}^{t_2} e^{-iq^2(t_2-t_1)} \times \\ &\times \tilde{v}(q - p', t_1, 0) e^{-ip'^2(t_1-t_0)} dt_1 dq dt_2 + \\ &+ \frac{\varepsilon^2}{i} \int_{t_0}^t e^{-ip^2(t-t_1)} \frac{\partial \tilde{v}}{\partial t}(p - p', t_1, 0) e^{-ip'^2(t_1-t_0)} dt_1 + O(\varepsilon^3), \end{aligned} \quad (2.8)$$

где функция $\tilde{v}(p, t, 0)$ — преобразование по Фурье потенциала взаимодействия $v(x, t, 0)$. Формула (2.8) справедлива при $t - t_0 > 0$; при $t - t_0 < 0$ получаем, что амплитуда

$\tilde{K}(p, p'; t, t_0; \varepsilon) \equiv 0$. Первый (по ε) член разложения (2.8) имеет вид

$$\int_{t_0}^t e^{-ip^2(t-t_1)} \tilde{v}(p-p', t_1, 0) e^{-ip'^2(t_1-t_0)} dt_1 \quad (2.8a)$$

и соответствует первому члену разложения (2.7), который можно переписать в виде

$$[\tilde{K}_1^{(n)}]_{ij} = \frac{1}{\tau} \sum_{n_1=0}^{n-1} \lambda_j^{n-n_1-1} [\tilde{Z}_1(t_{n_1})]_{ij} \lambda_i^{n_1} \Delta t_{n_1}, \quad (2.8b)$$

где $n = \frac{t}{\tau}$, $\Delta t_{n_1} = \tau$,

Из сравнения подынтегрального выражения в (2.8a) и общего члена суммы в (2.8b) можно сопоставить:

1) волновой функции $e^{ip'^2(t_1-t_0)}$ свободно движущейся частицы с импульсом $p' \in \mathbb{R}^3$ в момент времени t_1 сомножитель $\lambda_i^{n_1} \equiv \lambda_i(t_1)$ при $t_1 = n_1\tau$ в (2.8b):

$$e^{-ip'^2(t_1-t_0)} \leftrightarrow \lambda_i^{n_1},$$

соответственно,

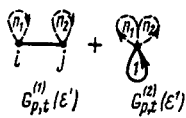
$$e^{-ip^2(t-t_1)} \leftrightarrow \lambda_j^{n-n_1-1};$$

2) «потенциалу» $\tilde{v}(p-p', t_1, 0)$ в (2.8b) — матричный элемент:

$$\tilde{v}(p-p', t_1, 0) \leftrightarrow [\tilde{Z}_1(t_{n_1})]_{ij}.$$

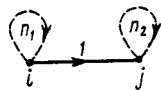
Установленное соответствие между первым членом ряда теории возмущений в (2.8a) и первым членом разложения амплитуды Грина КМ-цепи, с одной стороны, и соответствие между разложениями амплитуды Грина КМ-цепи и диаграммами соответствующего порядка (см. § 1) в « p, t »-представлении — с другой, позволяют сопоставить первому члену ряда теории возмущений (2.8a) диаграммы

$$\begin{aligned} & \int_{t_0}^t e^{-ip^2(t-t_1)} \tilde{v}(p-p', t_1, 0) e^{-ip'^2(t_1-t_0)} dt_1 \sim \\ & \sim \sum_{n_1=0}^{n-1} \lambda_j^{n-n_1-1} [\tilde{Z}_1(t_{n_1})]_{ij} \lambda_i^{n_1} \sim \end{aligned}$$



Как диаграмма $G_{p,t}^{(1)}(\varepsilon^1)$, так и диаграмма $G_{p,t}^{(2)}(\varepsilon^1)$ допускает следующую физическую интерпретацию. Начиная с момента времени $t = t_0$ свободная частица движется с импульсом p' до момента $t_{n_1} = n_1\tau$. В момент $t_{n_1} = n_1\tau$ частица под действием «потенциала» $\tilde{v}(p, t, 0) \leftrightarrow [\tilde{Z}_1(t_{n_1})]_{ij}$ переходит в другое состояние с импульсом $p \neq p'$ и далее движется свободно. В этом случае говорят, что частица «рассеялась» на потенциале один раз.

Однократное «рассеяние» на потенциале $v(x, t, 0)$ может произойти в любой момент времени $t_{n_1} = n_1\tau$ ($0 \leq n_1 \leq n-1$), поэтому полная амплитуда перехода $[\tilde{K}^{(n)}]_{ij}$ в первом порядке по ε получается суммированием вкладов всех возможных переходов при $n_1 = 0, 1, \dots, n-1$. При $i \neq j$ амплитуда $[\tilde{K}_1^{(n)}]_{ij}$ определяется вкладом диаграммы $\mathcal{L}_{p,t}^{(1)}(\varepsilon^1)$

$$\mathcal{L}_{p,t}^{(1)}(\varepsilon^1) - \text{diagram} \sim f(\mathcal{L}_{p,t}^{(1)}(\varepsilon^1)) = [\tilde{K}_1^{(n)}]_{ij},$$


при $i = j$ — вкладом диаграммы $\mathcal{L}_{p,t}^{(2)}(\varepsilon^1)$

$$\mathcal{L}_{p,t}^{(2)}(\varepsilon^1) - \text{diagram} \sim f(\mathcal{L}_{p,t}^{(2)}(\varepsilon^1)) = [\tilde{K}_1^{(n)}]_{ii}.$$


Аналогичным образом можно сопоставить каждому члену ряда теории возмущений (1.7) для функции Грина $\tilde{K}(p, p'; t, t_0; \varepsilon)$ в « p, t »-представлении диаграммы $\mathcal{L}_{p,t}(\varepsilon^r)$ теории возмущений соответствующей КМ-цепи в « p, t »-представлении.

Например, во втором порядке теории возмущений амплитуде

$$\begin{aligned} \tilde{K}_2(p, p'; t, t_0) = & \int_{t_0}^t e^{-ip'(t-t_2)} \int_{\mathbb{R}^1} \tilde{v}(p-q, t_2, 0) \times \\ & \times \int_{t_0}^{t_2} e^{-iq^2(t_2-t_1)} \tilde{v}(q-p', t_1, 0) e^{-ip'^2(t_1-t_0)} dt_1 dq dt_2 \end{aligned}$$

соответствуют диаграммы $\{\mathcal{L}_{p,t}^{(\alpha)}(\varepsilon^2), \alpha = 1, \dots, t(2)\}$ КМ-цепи, изображенные на рис. 174.

2. Определение « p, E »-представления диаграмм теории возмущений. С ростом числа r порядка теории возмущений число топологически не эквивалентных диаграмм, определяющих r -й член разложения функции Грина в « p, t »-представлении, быстро растет. Оказывается, что существует

$$= \tilde{\kappa}_2(p_1, p_2, p_3; t, t_0)$$

Рис. 174

такое представление КМ-цепи, в котором число неэквивалентных диаграмм порядка r значительно меньше, чем в случае « p, t »-представления КМ-цепи. Это новое представление КМ-цепи, которое мы будем называть « p, E »-представлением, естественным образом возникает в связи с « p, E »-представлением в квантовой механике. « p, E »-представление в квантовой механике — это представление волновой функции системы $\tilde{\Psi}(p, p'; t, t_0)$ в координатах: энергия E , $E \in \mathbb{R}^1$ и импульс $p, p \in \mathbb{R}^3$.

Волновая функция в « p, t »-представлении $\tilde{\Psi}(p, t)$ связана с волновой функцией в « p, E »-представлении преобразованием Фурье

$$\tilde{\Psi}(p, t) = \frac{1}{(2\pi)^{1/2}} \int_{-\infty}^{+\infty} e^{-iEt} \tilde{\tilde{\Psi}}(p, E) dE.$$

По отношению к « x, t »-представлению волновая функция $\tilde{\tilde{\Psi}}(p, E)$ представляет собой четырехмерное преобразование Фурье по переменным $x \rightarrow p, t \rightarrow E$

$$\Psi(x, t) = \frac{1}{(2\pi)^2} \int_{\mathbb{R}^1} \int_{\mathbb{R}^3} e^{-iEt} e^{ipx} \tilde{\tilde{\Psi}}(p, E) dE dp. \quad (2.9)$$

Рассмотрим преобразование диаграмм « p, t »-представления $\{\mathcal{L}_{p,t}^{(\alpha)}(e^r)\}$ в диаграммы « p, E »-представления $\{\mathcal{L}_{p,E}^{(\alpha)}(e^r)\}$, индуцированное переходом от координат (p, t) к координатам (p, E) . Это преобразование обозначим через F_L . Преобразование F_L диаграмм « p, t »-представления — это

пара двух преобразований F_G и F_{C^*} , $F_L = (F_G, F_{C^*})$, где F_G — преобразование графа $G_{p, t}(\epsilon^r)$ диаграммы $\mathcal{L}_{p, t}(\epsilon^r)$, а F_{C^*} — преобразование соответствующей коцепи.

Определение 2.1. Преобразованием $F_G(G_{p, t}(\epsilon^r))$ графа $G_{p, t}(\epsilon^r) = (\Gamma_{\text{int}} \cup \Gamma_{\text{ext}}, X)$ порядка r называется граф (X, Γ) , множество вершин X которого состоит из $(r + 1)$ вершин $(x_1, x_2, \dots, x_r, y)$, а множество дуг Γ — образ отображения

$$F_G : \Gamma_{\text{int}} \cup \Gamma_{\text{ext}} \rightarrow X \times X,$$

которое определяется условиями: 1) каждый внутренней дуге u_i^{int} ($i = 1, 2, \dots, r$) ставится в соответствие дуга $(y, x_i) \in \Gamma$ ($i = 1, 2, \dots, r$), 2) каждой внешней петле при внутренней вершине u_i^{ext} ($i = 1, 2, \dots, r$) ставится в соответствие дуга $(x_{\alpha_i}, x_{\beta_i})$ так, что если $u_i^{\text{ext}} \rightarrow (x_{\alpha_i}, x_{\beta_i})$, $u_{i+1}^{\text{ext}} \rightarrow (x_{\alpha_{i+1}}, x_{\beta_{i+1}})$, то $x_{\beta_i} = x_{\alpha_{i+1}}$, и тройка вершин $(x_{\alpha_i}, x_{\beta_i}, x_{\beta_{i+1}})$ определяет путь на графе (X, Γ) , 3) внешней петле u_1^{ext} и u_{r+1}^{ext} ставится в соответствие дуга (y, x_1) и (x_r, y) соответственно.

Дуги графа (X, Γ) , инцидентные выделенной вершине $y \in X$ назовем внешними, не инцидентные вершине y — внутренними.

Определение 2.2. Граф (X, Γ) , полученный преобразованием F_G графа $G_{p, t}(\epsilon^r)$, называется графом порядка r диаграммы $\mathcal{L}_{p, E}(\epsilon^r)$ КМ-цепи в « p, E »-представлении и обозначается через $G_{p, E}(\epsilon^r)$, т. е.

$$G_{p, E}(\epsilon^r) = F_G(G_{p, t}(\epsilon^r)).$$

Пример преобразования F_G для графов второго порядка изображен на рис. 175.

Как следует из определения 2.1, преобразование F_G переводит множество всех топологически неэквивалентных графов порядка r в « p, t »-представлении в один граф порядка r в « p, E »-представлении (рис. 176).

Рассмотрим теперь преобразование F_{C^*} целочисленных коцепей C_r^* , определенных на графах $G_{p, t}^{(a)}(\epsilon^2)$ и удовлетво-

ряющих при заданных целых числах n и r условию
 $n_1 + n_2 + \dots + n_r + n_{r+1} + r = n, \forall \alpha = 1, 2, \dots, t(r). \quad (2.10)$

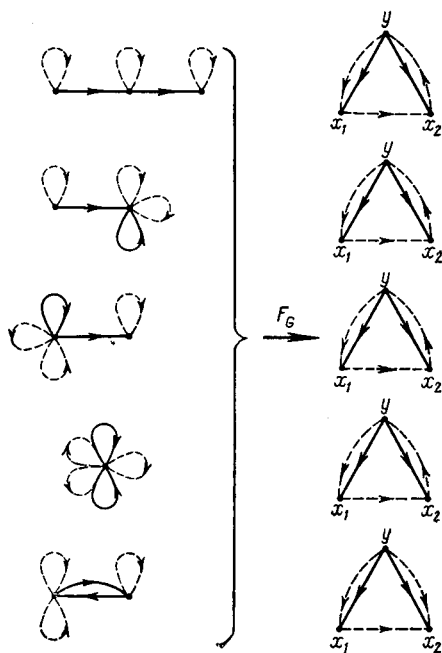


Рис. 175

Определение 2.3. Преобразованием $F_{C^*}(C_r^*)$ целочисленной коцепи $C_{r, p, t}^*$ называется коцепь $C_{r, p, E}^*$ со значением в $\mathbb{R}^1$, удовлетворяющая первому закону Кирхгофа в каждой вершине графа $G_{p, E}(e^r)$.

Если обозначить значение коцепи $C_{r, p, E}^*$ на дугах графа через

$$q_i = C_r^*((y, x_i)) \quad (i = 1, 2, \dots, r), \quad p_i = C_r^*((x_i, x_{i+1})) \\ (q_i, p_i \in \mathbb{R}^1) \quad (i = 1, 2, \dots, r)$$

$$p_{in} = C_r^*(y, x_1), \quad p_{out} = C_r^*(x_r, y), \quad (p_{in}, p_{out} \in \mathbb{R}^1),$$

то выполняются условия

$$q_{i+1} + p_i = p_{i+1} \quad (i = 1, 2, \dots, r-1), \quad (2.11)$$

$$\sum_{i=1}^r q_i + p_{\text{in}} = p_{\text{out}}. \quad (2.11a)$$

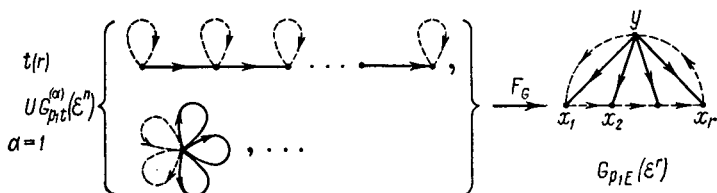


Рис. 176

Условие (2.11a) является следствием равенства (2.10) для соответствующей коцепи C_r^* в « p , t »-представлении.

Определение 2.4. Пара объектов $(G_{p,E}(\epsilon^r), C_{r,p,E}^*)$, где 1) $G_{p,E}(\epsilon^r)$ — граф, полученный преобразованием F_G графа $G_{p,t}(\epsilon^r)$ диаграммы в « p , t »-представлении, 2) $C_{r,p,E}^*$ — коцепь со значением в $\mathbb{R}^4$, являющаяся преобразованием F_C коцепи $C_{r,p,t}^*$ называется диаграммой теории возмущений $\mathcal{L}_{p,E}(\epsilon^r)$ КМ-цепи в « p , E »-представлении.

Рассмотренное преобразование F_L диаграмм « p , t »-представления в диаграммы « p , E » — представления согласовано с преобразованием Фурье по переменным время $t \rightarrow E$ (энергия) волновой функции в « p , t »-представлении в том смысле, что каждая диаграмма порядка r в « p , E »-представлении и только она одна определяет полностью r -й член разложения функции Грина в « p , E »-представлении. Для того чтобы в этом убедиться, в разложении (2.8) применим преобразование Фурье к обеим частям равенства по переменным $t \rightarrow E$, $t_0 \rightarrow E_0$. Для простоты рассмотрим, например, Фурье-образ первого члена разложения (2.9). Вводя функцию

$$\Theta(\lambda) = \begin{cases} 1, & \lambda > 0 \\ 0, & \lambda < 0, \end{cases}$$

запишем первый член разложения (2.8) в виде

$$\tilde{K}_1(p, p'; t, t_0) = \int_{\mathbb{R}^1} e^{-ip^2(t-t_1)} \Theta(t-t_1) \tilde{v}(p - p', t_1, 0) e^{-ip'^2(t_1-t_0)} \Theta(t_1-t_0) dt_1. \quad (2.12)$$

Далее заметим, что функция $e^{-ip^2(t-t_1)} \Theta(t-t_1)$ удовлетворяет уравнению Шредингера в « p, t »-представлении с правой частью

$$\left(i \frac{\partial}{\partial t} - p^2\right) e^{-ip^2(t-t_1)} \Theta(t-t_1) = \delta(t-t_1). \quad (2.13)$$

Переходя к Фурье-образам по переменной E , из уравнения (2.13) получаем, что

$$F_{t \rightarrow E} (e^{-ip^2(t-t_1)} \Theta(t-t_1)) = \lim_{v \rightarrow 0+0} \frac{1}{E - p^2 + iv} e^{+iEt_1} * \quad (2.13a)$$

Аналогично, для преобразования Фурье $t_0 \rightarrow E_0$ имеем:

$$F_{t_0 \rightarrow E_0} (e^{-ip'^2(t_1-t_0)} \Theta(t_1-t_0)) = \lim_{v \rightarrow 0+0} \frac{1}{E_0 - p'^2 + iv} e^{-iE_0 t_1}. \quad (2.13b)$$

З а м е ч а н и е. Преобразование Фурье $F_{t_0 \rightarrow E_0}$ записывается в виде

$$F_{t_0 \rightarrow E_0} \varphi(t_0) = \frac{1}{2\pi^{1/2}} \int_{\mathbb{R}^1} e^{-iE_0 t_0} \varphi(t_0) dt_0.$$

Из формулы (2.12) с учетом (2.13b) окончательно получаем, что

$$F_{t_0 \rightarrow E_0} F_{t \rightarrow E} \tilde{K}_1(p, p'; t, t_0) = \int_{\mathbb{R}^1} \frac{e^{+iEt_1}}{E - p^2 + iv} \tilde{v}(p - p', t_1, 0) \frac{e^{-E_0 t_1}}{E_0 - p'^2 + iv} dt_1 = \frac{1}{E - p^2 + iv} \tilde{v}(p - p', E - E_0, 0) \frac{1}{E_0 - p'^2 + iv}, \quad (2.14)$$

* Это обычная регуляризация с выходом в комплексную плоскость широко используется в теории обобщенных функций (см., например, [2]).

где $\tilde{v}(p, E, 0)$ — Фурье-образ потенциала $v(x, t, 0)$ при четырехмерном преобразовании Фурье от переменных (x, t) к переменным (p, E) .

Функция $\frac{1}{E - p^2 + i\nu}$ называется *свободным пропагатором*, или функцией распространения для нерелятивистской свободно движущейся бесспиновой частицы.

Введем четырехмерные векторы «энергии-импульса»

$$p_2 = (E, \mathbf{p}), \quad \mathbf{p} \in \mathbb{R}^3 \quad \text{и} \quad p_1 = (E_0, \mathbf{p}'), \quad \mathbf{p}' \in \mathbb{R}^3 {}^*.$$

Тогда в первом порядке теории возмущений для функции Грина в « p, E »-представлении получаем формулу

$$\begin{aligned} \tilde{K}_1(p_2, p_1) &= \frac{1}{E - p^2 + i\nu} \\ \tilde{v}(p_2 - p_1, 0) &\frac{1}{E_0 - p'^2 + i\nu}, \end{aligned} \quad (2.15)$$

которую можно представить в виде

$$\begin{aligned} \tilde{K}_1(p_2, p_1) &= \frac{1}{(2\pi)^2} \int_{\mathbb{R}^1} \frac{1}{E - p^2 + i\nu} \times \\ &\times \tilde{v}(q, 0) \frac{1}{E_0 - p'^2 + i\nu} \delta(q - p_2 + p_1) dq. \end{aligned} \quad (2.16)$$

Для полного ряда амплитуды $K(p_2, p_1, \varepsilon)$ с точностью до ε^2 включительно из (2.8) следует разложение

$$\begin{aligned} \tilde{K}(p_2, p_1) &= \frac{1}{E - p^2 + i\nu} \delta(E - E_0) \delta(\mathbf{p} - \mathbf{p}') + \\ &+ \left(\frac{\varepsilon}{i}\right) \int_{\mathbb{R}^4} \frac{1}{E - p^2 + i\nu} \tilde{v}(q, 0) \frac{1}{E_0 - p'^2 + i\nu} \delta(q - p_2 + \\ &+ p_1) dq + \left(\frac{\varepsilon}{i}\right)^2 \int_{\mathbb{R}^1} \int_{\mathbb{R}^3} \int_{\mathbb{R}^4} \frac{1}{E - p^2 + i\nu} \tilde{v}(q_2, 0) \frac{1}{E_q - q_2^2 + i\nu} \times \\ &\times \tilde{v}(q_1, 0) \frac{1}{E_0 - p'^2 + i\nu} \delta(q_2 - p + q) \delta(q_1 - q + \end{aligned}$$

* Здесь и далее трехмерные векторы обозначаются жирным шрифтом.

$$+ p_0) dq dq_1 dq_2 + \frac{\varepsilon^2}{i} \int_{\mathbb{R}^4} \frac{1}{E - p_2 + i\nu} \frac{\partial \tilde{v}}{\partial \varepsilon}(q, 0) \times \\ \times \frac{1}{E_0 - p'^2 + i\nu} \delta(q - p_2 + p_1) dq + O(\varepsilon^3). \quad (2.17)$$

Сопоставим каждому члену ряда $\tilde{K}_r(p_2, p_1)$ теории возмущений порядка r в разложении (2.17) диаграмму $\mathcal{L}_{p, E}(\varepsilon')$ порядка r в « p, E »-представлении. Предварительно произведем преобразование Γ_Φ над графом $G_{p, E}(\varepsilon')$. Это преобразование состоит в удалении отмеченной вершины y у графа



Рис. 177

$G_{p, E}(\varepsilon')$. При этом дуги, инцидентные вершине y , сохраняются и будут изображаться стрелками с одной вершиной — концом x_i для внутренних дуг вида (y, x_i) и началом x_r — для дуги (x_r, y) . Преобразование Γ_Φ над графом $G_{p, E}(\varepsilon')$ изображено на рис. 177, а, б.

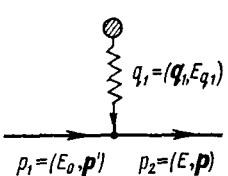
Объект, возникающий из графа $G_{p, E}(\varepsilon')$ в результате преобразования Γ_Φ , будем по-прежнему называть «графом» и обозначать через $G_\Phi(\varepsilon')$. Графы $G_\Phi(\varepsilon')$ ($r = 0, 1, \dots$) — это графы Фейнмана теории возмущений порядка r в нерелятивистской квантовой механике.

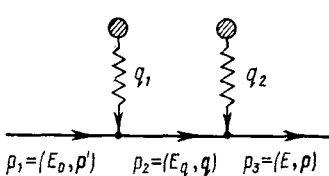
Граф $G_\Phi(\varepsilon')$ можно изобразить и так, как это показано на рис. 177, в, отмечая волнистой линией взаимодействие частицы с рассеивающим центром, который символически изображается заштрихованным кружком, а сплошной прямой стрелкой — свободное движение частицы. Такую форму изображения графов теории возмущений обычно используют в физической литературе [1].

Граф Фейнмана $G_\Phi(\varepsilon')$ порядка r представляет собой «дерево» с числом вершин, равным r . Соответствующая графу $G_\Phi(\varepsilon')$ диаграмма $\mathcal{L}_\Phi = (G_\Phi, C_\Phi^*)$ называется *диаграммой Фейнмана* порядка r .

З а м е ч а н и е. Граф Фейнмана $G_{\Phi}(\varepsilon^r)$ порядка r можно рассматривать как геометрическую реализацию графа $G_{p, E}(\varepsilon^r)$ порядка r на плоскости R^2 . Эта реализация получается при стереографической проекции сферы S^2 из полюса $N \in S^2$ правильной геометрической реализации плоского графа $G_{p, E}(\varepsilon^r)$ на сфере S^2 при условии, что полюсу N отвечает выделенная вершина y графа $G_{p, E}(\varepsilon^r)$.

Приведем диаграммы, отвечающие первым трем членам ряда теории возмущений (2.17):

$$\begin{aligned} & \frac{1}{E - p^2 + i\nu} \delta(E - E_0) \delta(p - p') \sim \xrightarrow{(E, p)} \\ & \int_{R^1} \frac{1}{E - p^2 + i\nu} \tilde{v}(q_1, 0) \times \\ & \times \frac{1}{E_0 - p'^2 + i\nu} \delta(q_1 - p_2 + p_1) dq_1 \sim \end{aligned}$$


$$\begin{aligned} & \iint_{R^1 R^1} \frac{1}{E - p^2 + i\nu} \tilde{v}(q_2, 0) \times \\ & \times \frac{1}{E_q - q^2 + i\nu} \tilde{v}(q_1, 0) \times \\ & \times \frac{1}{E_0 - p'^2 + i\nu} \delta(q_2 - p_3 + p_2) \delta(q_1 - p_2 + p_3) dq_1 dq_2 dp_2 \end{aligned}$$


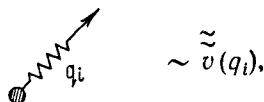
Сравнивая теперь члены разложения порядка r функции Грина в « p, E »-представлении из (2.17) с графами Фейнмана G_{Φ} порядка r ($r = 0, 1, \dots$) приходим к следующему правилу вычисления амплитуды $K_r(p_{r+1}, p_1)$ в каждом порядке r теории возмущений по заданной диаграмме Фейнмана $\mathcal{L}_{\Phi}(\varepsilon^r)$.

Для вычисления функции Грина $K_r(p_{r+1}, p_1)$ в « p, E »-представлении, в r -м порядке теории возмущений необходимо:

1) сопоставить каждой линии $(u, C_{r, p, E}^*(u) = p_i = (E, p_i))$ диаграммы $\mathcal{L}_{\Phi}(\varepsilon^r)$ пропагатор свободного движения частицы $\frac{1}{E_i - p_i^2 + i\nu}$

$$p_i \nearrow \sim \frac{1}{E_i - p_i^2 + i\nu},$$

каждой линии вида $(u, C_{r, p, E}^*(u) = q_i)$ Фурье компоненту потенциала взаимодействия $\tilde{v}(q_i)$, $q_i \in \mathbb{R}^4$:



$$\sim \tilde{v}(q_i),$$

а каждой вершине x_i — « δ -функцию» $\delta(q_i - p_{i+1} + p_i)$, выражающую закон сохранения импульса (1-й закон Кирхгофа) в каждой вершине графа G .

$$x_i \bullet \sim \delta(q_i - p_{i+1} + p_i);$$

$$2) \text{ перемножить величины } \frac{1}{E_i - p_i^2 + i\nu} \text{ и } \tilde{v}(q_i, 0)$$

в порядке движения по диаграмме $\mathcal{L}_{p, E}(e')$ от вершины x_1 к вершине x_r и умножить полученное произведение

$$\frac{1}{E_{r+1} - p_{r+1}^2 + i\nu} \tilde{v}(q_r) \frac{1}{E_r - p_r^2 + i\nu} \dots \tilde{v}(q_1) \frac{1}{E_1 - p_1^2 + i\nu}$$

на произведение « δ -функций»

$$\prod_{i=1}^{S_0} \delta(q_i - p_{i+1} + p_i);$$

3) полученное выражение проинтегрировать по «импульсам» потенциала $\tilde{v}(q_i)$, т. е. по переменным $q_1, q_2, \dots, q_r$, $q_i \in \mathbb{R}^4$ и по внутренним импульсам частицы $p_1, p_2, \dots, p_r$, $p_i \in \mathbb{R}^4$ ($p_i = (E_i, \mathbf{p}_i)$).

З а м е ч а н и е. Выбор « p, E »-представления в нерелятивистской квантовой механике является скорее вопросом удобства графической интерпретации членов ряда теории возмущений по константе взаимодействия z в этом представлении.

В релятивистской квантовой теории поля выбор « p, E »-представления по существу связан с релятивистской инвариантностью теории и поэтому в следующем параграфе диаграммы Фейнмана теории возмущений рассматриваются только в « p, E »-представлении.

§ 3. ДИАГРАММЫ ФЕЙНМАНА В КВАНТОВОЙ ТЕОРИИ ПОЛЯ

Диаграммы Фейнмана в квантовой теории поля являются естественным обобщением диаграмм нерелятивистской квантовой механики и представляют собой удобную графическую интерпретацию членов ряда теории возмущений по малой константе взаимодействия ϵ ($\epsilon \ll 1$) для полной амплитуды перехода квантовой системы из одного состояния в другое. Эта амплитуда является матричным элементом оператора, который в физической литературе носит название S -матрицы [1].

По сравнению с диаграммами квантовой механики диаграммы Фейнмана в квантовой теории поля выглядят намного сложнее. Это связано с различием физической картины тех явлений, которые рассматриваются в квантовой механике и в квантовой теории поля. Последняя описывает уже такие взаимодействия элементарных частиц с внешним полем, при которых возможно взаимное превращение частиц друг в друга. На языке теории графов различие между квантовой механикой и квантовой теорией поля это различие между графом, являющимся деревом (именно таков граф диаграмм Фейнмана в любом порядке теории возмущений по ϵ в квантовой механике), и произвольным связным графом с тем же числом вершин (такие графы возникают в квантовой теории поля).

Определение диаграмм Фейнмана в квантовой теории поля, так же как и диаграмм квантовой механики, тесным образом связано с первым законом Кирхгофа для линейных цепей, рассмотренном в § 2 гл. XV.

1. Определение диаграмм Фейнмана. Обозначим через $G_\Phi = (X_\Phi, \Gamma_\Phi)$ связный граф без кратных дуг и без петель с одной отмеченной вершиной $y \in X_\Phi$. Пусть число всех неотмеченных вершин равно S_0 , так что общее число всех вершин графа G_Φ равно $S_0 + 1$.

Неотмеченные вершины графа $G_\Phi \{x_1, x_2, \dots, x_{S_0}\}$ назовем *внутренними вершинами графа*, отмеченную вершину y — *внешней*. Таким образом, множество всех вершин графа G_Φ — это множество

$$X_\Phi = \{x_1, x_2, \dots, x_{S_0}, y\} = X \cup \{y\}. \quad (3.1)$$

Множество Γ_Φ всех дуг графа разобьем на два подмножества Γ_{int} и Γ_{ext} ; Γ_{ext} — множество дуг графа G_Φ , инцидент-

ных внешней вершине y , дуги множества Γ_{ext} называются *внешними дугами (линиями)*; Γ_{int} — множество дуг графа G_{Φ} , инцидентных только внутренним вершинам графа, дуги множества Γ_{int} называются *внутренними дугами (линиями)*.

Число внутренних дуг графа G_{Φ} обозначим через S_1 .

Рассмотрим на графе G_{Φ} 1-коцепь C_{Φ}^* со значениями в векторном пространстве $\mathbb{R}^4$. Будем различать значения коцепи C_{Φ}^* на внутренних и внешних дугах графа G_{Φ} . Примем следующие обозначения:

$$C_{\Phi}^*(u) = p, \quad p = (p_0, p_1, p_2, p_3) \in \mathbb{R}^4, \text{ если } u \in \Gamma_{\text{ext}}, \quad (3.2)$$

$$C_{\Phi}^*(u) = q, \quad q = (q_0, q_1, q_2, q_3) \in \mathbb{R}^4, \text{ если } u \in \Gamma_{\text{int}}.$$

Потребуем, чтобы в каждой вершине графа G_{Φ} выполнялся первый закон Кирхгофа, т. е. чтобы коцепь $C_{\Phi}^* \in P(G_{\Phi})$ (см. гл. XV § 1).

Определение 3.1.* Назовем диаграммой Фейнмана $\mathcal{L}_{\Phi}$ пару объектов (G_{Φ}, C_{Φ}^*) , где 1) G_{Φ} — связный граф без кратных дуг и петель с одной отмеченной вершиной, который называется графом Фейнмана, 2) C_{Φ}^* — 1-коцепь со значениями в $\mathbb{R}^4$, принадлежащая подпространству коразрезов $P(G_{\Phi})$ графа G_{Φ} (§ 3 гл. XV).

Число внутренних вершин S_0 графа G_{Φ} называют порядком диаграммы.

Пусть D — подматрица матрицы инцидентий графа G_{Φ} , не содержащая строки выделенной вершины y и столбцов, отвечающих дугам множества Γ_{ext} . Матрица D — матрица порядка $S_0 \times S_1$, элементы которой обозначаются через d_{ij} ($i = 1, 2, \dots, S_0$; $j = 1, 2, \dots, S_1$). Используя формулу (3.5) гл. XV и обозначения (3.2), запишем условие того, что 1-коцепь $C_{\Phi}^* \in P(G_{\Phi})$, в матричной форме следующим образом:

$$\sum_{j=1}^{S_1} d_{ij} q_j \pm p_i = 0, \quad (i = 1, 2, \dots, S_0) \quad (3.3a)$$

* Ср. с определением 2.4 § 2.

$$\sum_{i=1}^{S_0} \pm p_i = 0, \quad (3.36)$$

где знак «+» означает, что внешняя дуга $u_i \in \Gamma_{\text{ext}}$, $C_{\Phi}^*(u_i) = p_i$ направлена от вершины y к вершине x_i , знак «—» — наоборот, от вершины x_i к вершине y . Как следует из теоремы 3.4. гл. XV, условие (3.36) является условием совместности системы (3.3а) относительно неизвестных величин q_j ($j = 1, 2, \dots, S_1$) при заданных величинах p_i ($i = 1, 2, \dots, S_0$).

Приведем теперь физическую интерпретацию графа G_{Φ} и коцепи C_{Φ}^* , обычно используемую в квантовой теории поля [1]. Каждая дуга u_j графа G_{Φ} представляет собой свободное распространение релятивистской элементарной частицы массы $m_j \geq 0$, а каждая внутренняя вершина x_i графа G_{Φ} представляет взаимодействие элементарных частиц в малой области пространства-времени с внешним полем. В результате взаимодействия элементарных частиц возможны их взаимные превращения друг в друга. В зависимости от выбора взаимодействия и сорта частиц определяется число частиц, возникающих при одном акте взаимодействия в малой области x пространства-времени, что в свою очередь определяет степень внутренних вершин x_i ($i = 1, 2, \dots, S_0$) графа G_{Φ} . Так, например, в квантовой электродинамике [8] степень каждой вершины графа G_{Φ} равна трем, что связано с возможным рождением (уничтожением) электронно-позитронной пары частиц и фотона. Значение коцепи C_{Φ}^* на дугах графа представляет собой четырехмерный вектор энергии-импульса релятивистской частицы. Например, если $u_j \in \Gamma_{\text{ext}}$, то $C_{\Phi}^*(u_j) = p_j = (p_j^0, \mathbf{p}_j)$, где p_j^0 — энергия, а $\mathbf{p}_j \in \mathbb{R}^3$ — импульс частицы. Условие (3.36) тогда означает выполнение закона сохранения энергии-импульса в каждом процессе взаимодействия в области x_i ($i = 1, 2, \dots, S_0$).

Для частиц массы m_j , связанных с внешними дугами графа предполагается выполненным релятивистское условие связи энергии-импульса p с массой частицы m_j :

$$p_j^2 = (p_j^0)^2 - \sum_{i=1}^3 (p_j^i)^2 = m_j^2. \quad (3.4)$$

Если выполнено дополнительное условие

$$p_j^0 > 0, \quad (3.5)$$

то говорят, что вектор $p_j \in \mathbb{R}^4$ принадлежит массовой оболочке — положительной полё гиперboloида, уравнением которого в псевдоевклидовом пространстве $\mathbb{R}^4$ является соотношение (3.4).

Четырехмерные импульсы внутренних дуг графа, вообще говоря, не удовлетворяют условию (3.4) и соответствующие им элементарные частицы называются *виртуальными*.

В экспериментах по столкновению (рассеянию) частиц требуется обычно определить по заданным до столкновения 4-импульсам сталкивающихся частиц энергии-импульсы частиц, возникших в результате взаимодействия в некоторой области пространства-времени.

На графе G_Φ свободным частицам до взаимодействия отвечают внешние дуги, идущие от выделенной вершины y к вершинам x в которых происходит взаимодействие частиц. Соответственно, частицам, обнаруженным после рассеяния, отвечают внешние дуги графа G_Φ , направленные к вершине y .

Таким образом, вершине y графа G_Φ в физическом пространстве-времени отвечает «бесконечно удаленная точка» — начало движения системы невзаимодействующих свободных частиц до рассеяния и конец движения свободных частиц после рассеяния. Поэтому при геометрической реализации графа G_Φ вершина y изображаться не будет, а соответствующие этой вершине внешние дуги имеют только одну вершину (входа или выхода). Типичный пример подобной реализации графа Фейнмана представлен на рис. 178.

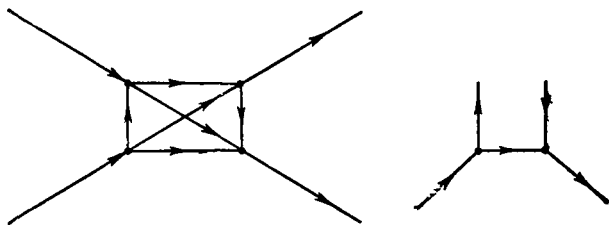


Рис. 178

З а м е ч а н и е. Указанную выше геометрическую реализацию графа Фейнмана можно получить путем стереографической проекции из полюса $N \in S^2$ геометрической реализации* графа G_Φ на сфере S^2 при условии, что внешней вершине y графа G_Φ соответствует точка N .

Отсутствие петель на графе G_Φ отвечает причинной упорядоченности двух последовательных взаимодействий в пространстве-времени (x, t) .

Данное выше определение диаграмм Фейнмана соответствует в физической литературе [1] диаграммам Фейнмана в импульсном представлении. Это связано с тем, что состояние свободной квантовой частицы помимо внутренних квантовых чисел определяется сохраняющимся между двумя последовательными взаимодействиями 4-импульсом частицы $p_j = (p_j^0, \mathbf{p}_j)$.

Аналогично тому, как это было проделано для диаграмм нерелятивистской квантовой механики (§ 1), можно рассмотреть координатное (x, t) -представление диаграмм Фейнмана. Поскольку переход от координат (x, t) к (p, E) с помощью четырехмерного преобразования Фурье — унитарное преобразование, то величины, подлежащие измерению, в процессах столкновения частиц не зависят от выбора представления. Поэтому всюду ниже будет рассматриваться только импульсное представление для описания взаимодействия частиц. В то время как число внешних дуг графа G_Φ определяется числом реальных частиц, т. е. частиц, обнаруженных в результате измерения до и после столкновения, число внутренних дуг S_1 при заданном порядке S_0 диаграммы может быть произвольным (при условии, что $S_0 - 1 \leq S_1 \leq 2C_{S_0}^2$, где $C_{S_0}^2$ — число сочетаний из S_0 по 2) и в конечном счете зависит от выбора физической модели, описывающей процесс рассеяния.

Порядок диаграммы S_0 однозначно соответствует r -му члену ряда теории возмущений (при $r = S_0$) в разложении полной амплитуды перехода системы (так называемой S -матрицы) из начального состояния при $t = -\infty$ в конечное при $t = +\infty$ по малой константе взаимодействия ϵ ($\epsilon \ll 1$).

* Здесь геометрическая реализация графа G_Φ не обязана быть правильной (сравни с § 2 гл. V), так как графа G_Φ в квантовой теории поля не обязательно является плоским, как это было в квантовой механике (см. § 2).

Определение полной амплитуды перехода (S -матрицы) является основной и крайне сложной задачей квантовой теории поля. Поэтому представляет интерес приближенное вычисление полной амплитуды перехода, приводящее к важным физическим следствиям уже при разложении с точностью до ϵ^2 или ϵ^3 [1]. Вычисление амплитуды перехода с точностью до ϵ' связано с рассмотрением графов Фейнмана с числом вершин $S_0 = r$.

Ниже будет рассмотрен подробнее случай взаимодействия релятивистских частиц с нулевым спином и массой $m > 0$, волновая функция $\Psi(x)$ которых удовлетворяет уравнению Кляйна—Гордона—Фока

$$\left(\Delta - \frac{\partial^2}{\partial t^2}\right)\Psi(x) \equiv \square \Psi(x) = m^2 \Psi(x), \quad x \in \mathbb{R}^4.$$

Здесь и всюду ниже используется система единиц, в которой скорость света c и константа Планка $\hbar$ равны единице.

2. Вклады диаграмм Фейнмана в амплитуду перехода в импульсном представлении [7]. Пусть $\mathcal{L}_\Phi$ —диаграмма Фейнмана порядка S_0 с числом внутренних дуг S_1 .

Определение 3.5. «Интегралом» Фейнмана, отвечающим диаграмме $\mathcal{L}_\Phi$ (вкладом диаграммы $\mathcal{L}_\Phi$ или амплитудой Фейнмана), называется формальное выражение вида

$$\begin{aligned} \mathcal{I}_\Phi(p_1, p_2, \dots, p_{S_0}) = & (i)^{S_1} \int_{\mathbb{R}^{4S_1}} \dots \int dq_1 dq_2 \dots dq_{S_1} \times \\ & \times \frac{\prod_{i=1}^{S_0} \delta\left(\sum_{j=1}^{S_1} d_{ij} q_j + p_i\right)}{\prod_{j=1}^{S_1} (m_j^2 - q_j^2 - i\nu)} P(q_1, q_2, \dots, q_{S_1}), \end{aligned} \quad (3.6)$$

где $q_j \in \mathbb{R}^4$ ($j = 1, 2, \dots, S_1$) — четырехмерный импульс внутренней дуги u_j графа G_Φ диаграммы $\mathcal{L}_\Phi$, ν — малый параметр ($\nu \rightarrow 0 + 0$), $P(q_1, q_2, \dots, q_{S_1})$ — спиновый полином порядка k относительно переменных $q_1, q_2, \dots, q_{S_1}$, d_{ij} — элементы матрицы инцидентий графа G_Φ .

Спиновый полином $P(q_1, q_2, \dots, q_{S_1})$ представляет собой многочлен относительно $q_1, q_2, \dots, q_{S_1}$ с матричными коэффициентами порядка 4×4 , которые в свою очередь выражаются через четырехрядные матрицы Дирака [8].

Функция $\frac{i}{m_j^2 - q_j^2 - i\nu}$ называется *пропагатором*, или

функцией распространения свободной релятивистской бесспиновой частицы массы $m_j > 0$ и является фундаментальным решением (функцией Грина) для уравнения Кляйна—Гордона — Фока в импульсном представлении. Последнее означает, что функция

$$K_j^+(x, t) = \lim_{\nu \rightarrow 0+0} \int_{\mathbb{R}^4} e^{i\langle q, \tilde{x} \rangle} \frac{i}{m_j^2 - q^2 - i\nu} dq,$$

где $\langle q, \tilde{x} \rangle = q_0 t - \sum_{i=1}^3 q_i x_i$, $q \in \mathbb{R}^4$, $x \in \mathbb{R}^3$, удовлетворяет уравнению

$$\square K_j^+(x, t) = +m_j^2 K_j^+(x, t) + \delta(x)\delta(t), \quad x \in \mathbb{R}^3, \quad t \in \mathbb{R}^1.$$

Нетрудно видеть, сравнивая выражения (3.6) и (2.16—2.17), что определение «интеграла» Фейнмана порядка $r = S_0$ в квантовой теории поля является прямым обобщением амплитуды вероятности Фейнмана в нерелятивистской квантовой механике в импульсном представлении, отвечающей r -му (при $r = S_0$) приближению по константе взаимодействия e . Формальное выражение (3.6) при определенных условиях на структуру графа G_Φ диаграммы $\mathcal{L}_\Phi$ при заданном порядке k полинома $P(q_1, q_2, \dots, q_{S_1})$ определяет сходящийся при $|q_j| \rightarrow +\infty$ несобственный интеграл.

Ниже будет показано, что этот интеграл существует, если выполнено неравенство

$$2S_1 - 4\nu(G_\Phi) - k > 0, \quad (3.7)$$

где $\nu(G_\Phi)$ — цикломатическое число графа G_Φ , k — порядок спинового полинома $P(q_1, q_2, \dots, q_{S_1})$.

Если G_Φ — связанный граф, то используя определение (гл. IV) цикломатического числа, неравенство (3.7) можно представить в виде

$$4S_0 - 2S_1 - (k + 4) > 0. \quad (3.8)$$

Из неравенства (3.8) следует, что при заданном порядке r ($r = S_0$) теории возмущений в разложении « S »-матрицы

в импульсном представлении диаграммы Фейнмана, описывающие процессы с большим числом (S_1) виртуальных частиц, дают неограниченно большой вклад. Так что использование «интеграла» Фейнмана в этом случае приводит к бессмысленному результату, и возникает задача о регуляризации (устранении расходимостей) выражения (3.6). В общем случае эта задача подробно рассматривается в [1].

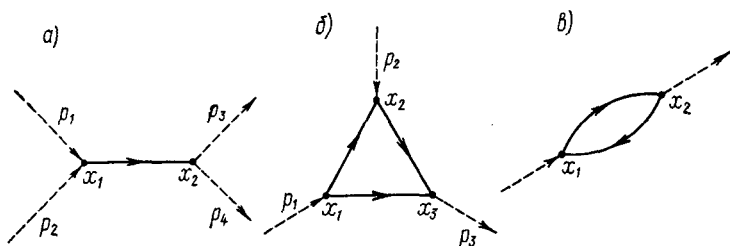


Рис. 179

При заданном порядке k полинома $P(q_1, q_2, \dots, q_{S_1})$ нетрудно привести примеры диаграмм, для которых выражение (3.26) определено корректным образом. Для $k = 0$ такие диаграммы порядка $r = 2$ и $r = 3$ изображены на рис. 179, а, б соответственно. При $k = 1$ или $k = 2$ «интеграл» Фейнмана для диаграммы на рис. 179, б является уже расходящимся при $|q_j| \rightarrow \infty$ интегралом, точно так же как и для диаграммы ($r = 2$), изображенной на рис. 179, в.

Всюду ниже будут рассматриваться только такие диаграммы, для которых при заданном порядке k полинома $P(q_1, q_2, \dots, q_{S_1})$ выполняется условие (3.7) так, что корректным является в этом случае употребление термина «интеграл Фейнмана» в собственном значении понятия интеграла.

Приступим теперь к выводу условия (3.7). В силу того что в каждой вершине x_i графа G_Φ выполняется закон сохранения 4-импульса, число независимых переменных импульсов, по которым ведется интегрирование в (3.6), на самом деле меньше, чем S_1 . Преобразуем подынтегральное выражение в (3.6), используя результаты § 3 гл. XV о матричном представлении первого закона Кирхгофа в электрических сетях.

Во-первых, воспользуемся тем, что $\text{rang } D = S_0 - 1$, и поэтому число независимых строк матрицы D равно

$S_0 - 1$. Пусть строка матрицы D с номером S_0 выражается через предыдущие $(S_0 - 1)$ строки в виде

$$d_{S_0} \equiv (d_{S_0 1}, d_{S_0 2}, \dots, d_{S_0 S_1}) = - \sum_{i=1}^{S_0-1} d_i, \\ d_i = (d_{i1}, d_{i2}, \dots, d_{iS_1}). \quad (3.9)$$

Равенство (3.9) выполняется в силу свойства матрицы инцидентий

$$\sum_{i=1}^{S_0} d_{ij} = 0 \quad (j = 1, 2, \dots, S_1).$$

Рассмотрим произведение

$$\prod_{i=1}^{S_0} \delta \left(\sum_{j=1}^{S_1} d_{ij} q_j + p_i \right) = \delta \left(\sum_{j=1}^{S_1} d_{S_0 j} q_j + p_{S_0} \right) \times \\ \times \prod_{i=1}^{S_0-1} \delta \left(\sum_{j=1}^{S_1} d_{ij} q_j + p_i \right). \quad (3.10)$$

Заменяя в аргументе первого сомножителя в (3.10) сумму

$\sum_{j=1}^{S_1} d_{S_0 j} q_j$ на выражение

$$\sum_{j=1}^{S_1} d_{S_0 j} q_j = \sum_{j=1}^{S_1} \left(- \sum_{i=1}^{S_0-1} d_{ij} \right) q_j = - \sum_{i=1}^{S_0-1} \left(\sum_{j=1}^{S_1} d_{ij} q_j \right) = \\ = - \sum_{i=1}^{S_0-1} (-p_i) = \sum_{i=1}^{S_0-1} p_i,$$

преобразуем произведение (3.10) к виду

$$\prod_{i=1}^{S_0} \delta \left(\sum_{j=1}^{S_1} d_{ij} q_j + p_i \right) = \delta \left(\sum_{i=1}^{S_0} p_i \right) \times \\ \times \prod_{i=1}^{S_0-1} \delta \left(\sum_{j=1}^{S_1} d_{ij} q_j + p_i \right). \quad (3.11)$$

С учетом (3.11) перепишем интеграл $\mathcal{J}_\Phi$ в виде

$$\mathcal{J}_\Phi(p_1, p_2, \dots, p_{S_0}) = (i)^{S_1} \delta\left(\sum_{i=1}^{S_0} p_i\right) \int_{\mathbb{R}^{S_1}} dq_1 dq_2 \dots dq_{S_1} \times \\ \times \frac{\prod_{i=1}^{S_0-1} \delta\left(\sum_{j=1}^{S_1} d_{ij} q_j + p_i\right)}{\prod_{j=1}^{S_1} (m_j^2 - q_j^2 - i\nu)} P(q_1, q_2, \dots, q_{S_1}). \quad (3.12)$$

В интеграле (3.12) сделаем замену переменных:

$$\sum_{i=1}^{S_1} d_{ij} q_j + p_i = \xi_i \quad (i = 1, 2, \dots, S_0 - 1). \quad (3.13)$$

Пусть D_T — минор матрицы D , отвечающий дереву $T(G)$ графа G , и пусть $\{u_{\nu+1}, u_{\nu+2}, \dots, u_{S_1}\}$ — дуги дерева $T(G)$, тогда в силу теоремы (3.4) § 3 гл. XV для системы (3.13) решение $q = (q_1, q_2, \dots, q_{S_1})^T$, $q_j \in \mathbb{R}^1$ представимо в виде

$$q = Z^T \gamma + (q_T, D_T^{-1}(\xi - p))^T,$$

или

$$q_i = (Z^T \gamma)_i + q_i^* \quad (i = 1, 2, \dots, \nu), \\ q_j = (Z^T \gamma)_j + [D_T^{-1}(\xi - p)]_j \quad (j = \nu + 1, \dots, S_1). \quad (3.13a)$$

Здесь произвольные переменные $\gamma_i \in \mathbb{R}^1$ ($i = 1, 2, \dots, \nu$) — это величины, которые в § 3 гл. XV в цепях Кирхгофа назывались токами в циклах.

Произвольные переменные q_i^* положим равными нулю, тогда якобиан матрицы перехода от переменных $(q_1, q_2, \dots, q_{S_1})$ к переменным $(\gamma_1, \gamma_2, \dots, \gamma_\nu, \xi_1, \dots, \xi_{S_0-1})$ равен (см. § 3 гл. XV)

$$\left| \frac{D(q_1, q_2, \dots, q_\nu, q_{\nu+1}, \dots, q_{S_1})}{D(\gamma_1, \gamma_2, \dots, \gamma_\nu, \xi_1, \dots, \xi_{S_0-1})} \right| = \det \begin{vmatrix} 1 & Z_T^T \\ 0 & D_T^{-1} \end{vmatrix} = 1.$$

Здесь Z_T^{tr} — невырожденная подматрица транспортной матрицы циклов Z , определитель которой равен единице.

В переменных $\gamma_1, \gamma_2, \dots, \gamma_\nu, \xi_1, \dots, \xi_{S_0-1}$ интеграл (3.12) примет вид

$$\mathcal{J}_\Phi(p_1, p_2, \dots, p_{S_0}) = (i)^{S_1} \delta \left(\sum_{i=1}^{S_0} p_i \right) \int_{\mathbb{R}^{4\nu} \times \mathbb{R}^{4(S_0-1)}} d\gamma_1 \dots$$

$$\dots d\gamma_\nu d\xi_1 \dots d\xi_{S_0-1} \frac{\prod_{i=1}^{S_0-1} \delta(\xi_i) P(q_1, \dots, q_{S_1})}{\prod_{j=1}^{S_1} (m_j^2 - q_j^2 - i\nu)} \cdot \quad (3.14)$$

Используя характеристическое свойство « δ -функций», проинтегрируем (3.14) по переменным $\xi_1, \xi_2, \dots, \xi_{S_0-1}$. В результате интегрирования получим следующее представление для $\mathcal{J}_\Phi(p_1, \dots, p_{S_0})$:

$$\mathcal{J}_\Phi(p_1, \dots, p_{S_0}) = (i)^{S_1} \delta \left(\sum_{i=1}^{S_0} p_i \right) \int_{\mathbb{R}^{4\nu}} d\gamma_1 \dots d\gamma_\nu \times$$

$$\times \frac{P(q_1(\gamma) \dots q_{S_1}(\gamma))}{\prod_{j=1}^{S_1} (m_j^2 - q_j^2(\gamma) - i\nu)}, \quad (3.15)$$

где векторы $q_j(\gamma)$ ($j = 1, 2, \dots, S_1$) определяются из уравнений (3.13), если в них положить все ξ_i ($i = 1, 2, \dots, S_0 - 1$) равными нулю, т. е.

$$q_j = (Z^{tr} \gamma)_j \quad (j = 1, 2, \dots, \nu), \quad (3.16)$$

$$q_j = (Z^{tr} \gamma)_j - [D_T^{-1}(p)]_j \quad (j = \nu + 1, \dots, S_1).$$

Если теперь в интеграле (3.15) перейти к сферическим координатам в пространстве $\mathbb{R}^{4\nu}$: (ρ, ω) , где $\omega = (\varphi_1, \dots, \varphi_{4\nu-1})$ — угловые координаты на сфере $S^{4\nu-1}$ размерности $(4\nu-1)$, то подынтегральное выражение при $\rho \rightarrow \infty$ (равномерно по $\omega \in S^{4\nu-1}$) ведет себя как функция вида $(\rho^{2S_1 - (4\nu-1) - k})^{-1}$.

Отсюда следует, что для сходимости интеграла (3.15) при $p \rightarrow \infty$ необходимо потребовать выполнения неравенства

$$2S_1 - (4\nu - 1) - k > 1, \quad (3.17)$$

которое совпадает с неравенством (3.7).

Дальнейшее исследование интеграла Фейнмана, в частности, исследование аналитических свойств выражения (3.6) как функции внешних импульсов частиц $p_1, p_2, \dots, p_{S_1}$, связано с параметрическим представлением, или так называемым « α »-представлением интеграла Фейнмана [1].

Вывод « α »-представления и многочисленные применения этого представления к исследованию свойств интеграла Фейнмана подробно рассматриваются в книге [7].

ЛИТЕРАТУРА К ДОПОЛНЕНИЮ

1. Боголюбов Н. Н., Ширков Д. В. Введение в теорию квантованных полей. М., «Наука», 1973.
2. Владимиров В. С. Уравнения математической физики. М., «Наука», 1972.
3. Карасев М. В. Матем. заметки, 18, № 2 (1975), 267—276.
4. Крылов В. Ю. Дополнение V к книге Годунова С. К. и Рябенского В. С. «Введение в теорию разностных схем». М., Физматгиз, 1962.
5. Маслов В. П. ДАН. 1970., т. 192, № 2.
6. Маслов В. П. Операторные методы. М., «Наука», 1973.
7. Nakamishi N. Graph Theory and Feynman integral. New York, 1971.
8. Новейшее развитие квантовой электродинамики. Сборник статей, ИЛ, 1954.
9. Самарский А. А. Введение в теорию разностных схем. М., «Наука», 1971.
10. Фейнман (Feynman R. P.) Rev. Mod. Phys. 20.367. 1948 (имеется перевод в сборнике «Вопросы причинности в квантовой механике», М., ИЛ, 1955, стр. 167).
11. Фейнман (Feynman R. P.) Phys. Rev. 76. 749. 1949 (имеется перевод в [7]).
12. Фейнман (Feynman R. P.) Phys. Rev. 76. 769. 1949 (имеется перевод в [7]).
13. Фейнман Р., Лейтон Р., Сэндс М. Фейнмановские лекции по физике. М., «Мир», 1965, т. 3, гл. 37—38.
14. Фейнман Р., Хибс А. Квантовая механика и интегралы по траектории. М., «Мир», 1968.
15. Феллер В. Введение в теорию вероятностей и ее приложения. М., «Мир», 1967, т. 2.

ОГЛАВЛЕНИЕ

Предисловие	3
Г л а в а I. Основные определения. Первые задачи теории графов	5
§ 1. Предварительные сведения из теории множеств	5
§ 2. Основные определения теории графов	9
§ 3. Задачи, послужившие основой теории графов	14
Г л а в а II. Алгоритмические задачи	18
§ 1. Задачи о кратчайших путях	18
§ 2. Алгоритм построения эйлерова цикла	22
§ 3. Лабиринты	24
§ 4. Алгоритмы теории графов	25
Г л а в а III. Гомоморфизмы мультиграфов. Группа автоморфизмов	28
§ 1. Вспомогательные сведения из теории групп	28
§ 2. Гомоморфизмы, изоморфизмы и автоморфизмы графов	33
Г л а в а IV. Циклы в графах. Цикломатическое число графов	43
§ 1. Цикломатическое число графа. Деревья	43
§ 2. База независимых циклов	46
Г л а в а V. Топологические инварианты двумерных поверхностей	59
§ 1. Эйлерова характеристика	60
§ 2. Гомологии двумерных поверхностей	75
Г л а в а VI. Раскраски графов. Хроматическое число	84
§ 1. Вычисление хроматического числа заданного графа	84
§ 2. Оценка хроматического числа плоского графа	95
Г л а в а VII. Ядро графа. Применение графов в теории игр	100
§ 1. Ядро графа и его свойства	100
§ 2. Графы в теории игр	111
Г л а в а VIII. Транспортные сети	114
§ 1. Потоки на транспортных сетях	114
§ 2. Некоторые прикладные задачи	119
§ 3. Поток, насыщающий выходные дуги	122

Г л а в а IX. Теорема о системе различных представителей и ее приложение в теории графов.	128
§ 1. Теорема о системе различных представителей и другие эквивалентные ей утверждения	128
§ 2. Эквивалентность теорем 1.1 — 1.6	134
§ 3. Венгерский алгоритм	144
§ 4. Прикладные задачи на паросочетания	150 ✓
Г л а в а X. Гамильтоновы маршруты и задача о коммивояжере.	168 ✓
§ 1. Гамильтоновы циклы и контуры	168
§ 2. Задача о коммивояжере	179
§ 3. Метод ветвей и границ	182
Г л а в а XI. Элементы теории конечных автоматов. Матрицы смежности графов и операции над ними	199
§ 1. Некоторые понятия теории конечных автоматов	199
§ 2. Матрицы смежности графов и операции над ними	206
§ 3. Разложение графов по операциям произведения и суперпозиции	219
Г л а в а XII. Производящие функции и их применение в перечислительных задачах теории графов	226
§ 1. Рисуночные ряды и производящие функции	226
§ 2. Задачи о размене монет и рассечении многоугольников	231
§ 3. Перечисление корневых деревьев	236
Г л а в а XIII. Теорема Пойа	239
§ 1. Цикловой индекс группы подстановок множества	240
§ 2. Транзитивные множества группы подстановок	245
§ 3. Классы эквивалентных отображений	248
§ 4. Вес и перечень фигур и конфигураций	252
§ 5. Теорема Пойа	254
§ 6. Цикловые индексы двух групп, часто встречающихся в теории перечислений графов	259
Г л а в а XIV. Некоторые методы перечисления графов	262
§ 1. Перечисление графов с p вершинами и k дугами	262
§ 2. Перечисление полностью неориентированных графов с p вершинами и k ребрами	266
§ 3. Перечисление маркированных графов	270
§ 4. Перечисление деревьев	274
	391

Глава XV. Электрические сети	279
§ 1. Коцепи в графе	279
§ 2. Применение графов в теории электрических сетей	284
§ 3. Матрицы, ассоциированные с группами одномерных и нульмерных цепей графа	287
Литература	294
Дополнение В. В. Белов, В. П. Маслов Теория возмущений комплексных марковских цепей	295
Глава I. Комплексные марковские цепи и суммирование по путям	296
§ 1. Алгебра матриц и графы	296
§ 2. Суперпозиция амплитуд вероятностей событий по Фейнману и определение комплексной марковской цепи	309
§ 3. КМ-цепь фотона	326
Глава II. Диаграммы Фейнмана в теории возмущений КМ-цепи и квантовой теории	337
§ 1. Регулярная теория возмущений КМ-цепи	339
§ 2. Диаграммы Фейнмана в нерелятивистской квантовой механике как диаграммы теории возмущений КМ-цепи в « p , E »-представлении	363
§ 3. Диаграммы Фейнмана в квантовой теории поля	378
Литература к дополнению	389

Владимир Владимирович БЕЛОВ,
Евгений Михайлович ВОРОБЬЕВ,
Виктор Евгеньевич ШАТАЛОВ

ТЕОРИЯ ГРАФОВ

Редактор А. И. Селиверстова. Переплет художника В. Л. Ильичева.
Художественный редактор В. И. Пономаренко. Технический редактор Н. В. Яшукова. Корректор В. В. Кожуткина.

Сдано в набор 11/IX-75 г. Подп. к печати 16/IV-76 г. Формат 84×108/32. Бум. тип. № 1. Объем 12,25 печ. л. Усл. п. л. 20,58. Уч.-изд. л. 18,62. Изд. № ФМ-556. Тираж 20 000 экз. Цена 81 коп. Зак. 637.

План выпуска литературы издательства
«Высшая школа» (вузы и техникумы) на 1975 г. Позиция № 79.

Москва, К-51, Неглинная ул., д. 29/14,
издательство «Высшая школа»

Ярославский полиграфкомбинат Союзполиграфпрома при Государственном комитете Совета Министров СССР по делам издательств, полиграфии и книжной торговли.
150014, Ярославль, ул. Свободы, 97.