



*Харківський національний університет  
імені В. Н. Каразіна*

**“ІНФОРМАТИКА І ІНФОРМАЦІЙНІ ТЕХНОЛОГІЇ  
для хіміків”**

# **Лекція №7**

## **Теорія графів в хімії**

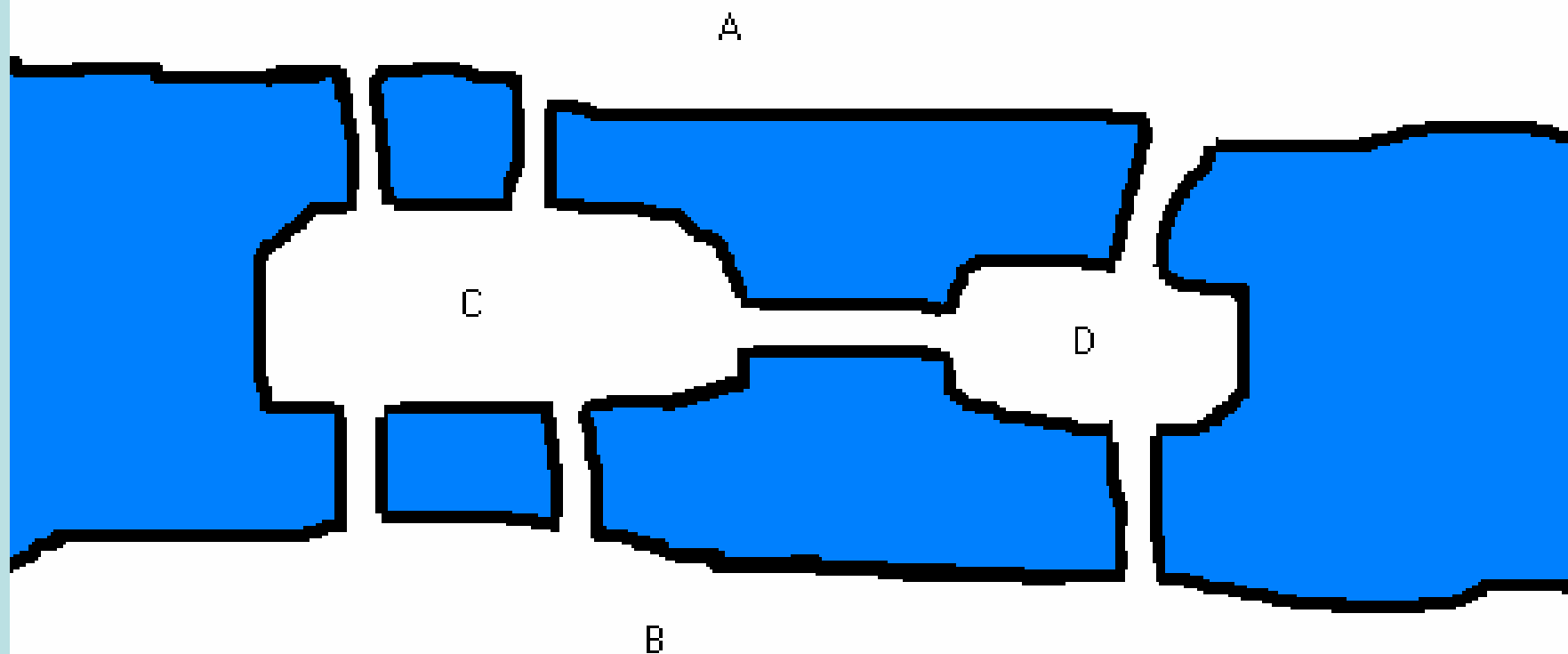
**В. В. Іванов**

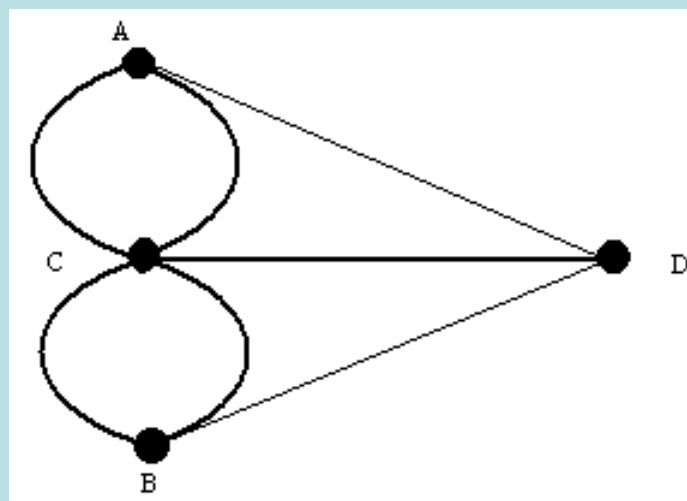
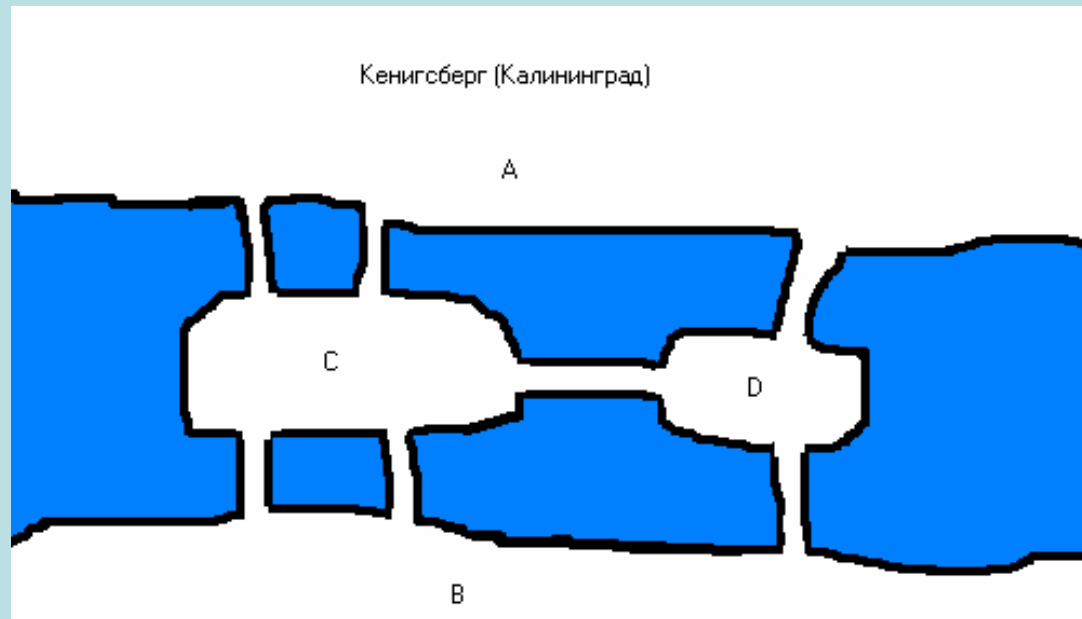
*Кафедра хімічного матеріалознавства*

# Леонард Эйлер (1707-1783)



Кенигсберг (Калининград)

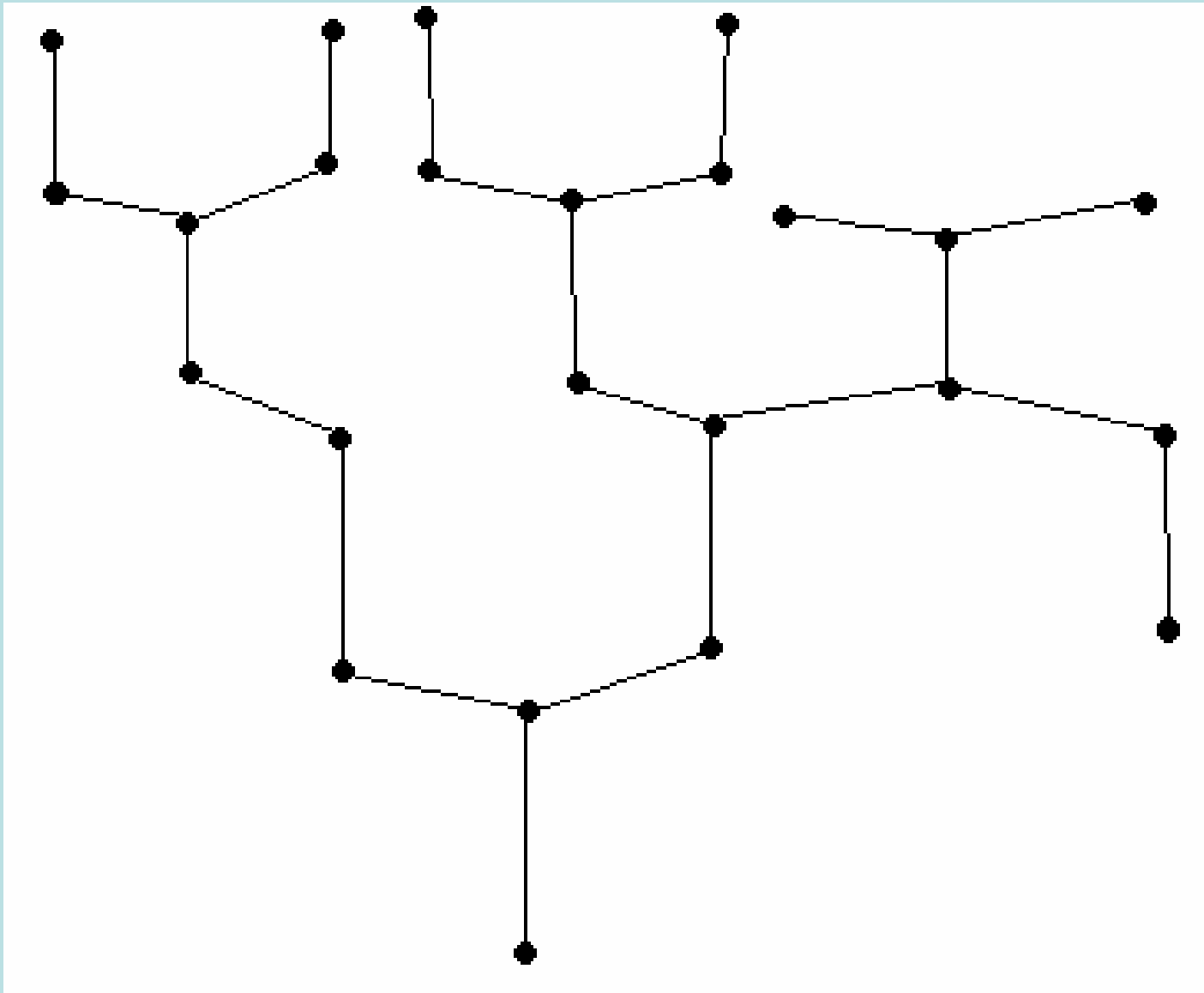




(вершина, вертекс, vertex), а мост – ребро (edge):

## Немного терминологии

- **Граф** ( $G$ )
- **вершина**, вертекс, (vertex)
- **ребро** (edge)
- **Маршрут**.
- Маршрут называется **цепью** если все его ребра различны.
- Цепь в которой первая и последняя вершины совпадают называется **циклом**.
- Граф называется **связным** если любая пара его вершин соединяется цепью.
- Связный граф, в котором нет циклов называется **деревом**.



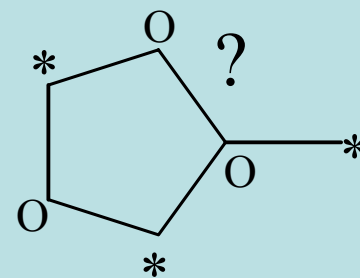
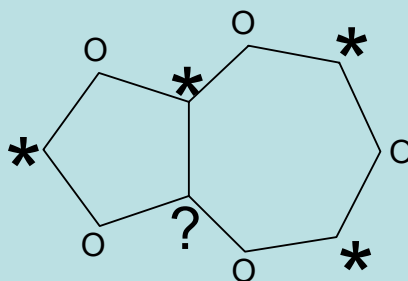
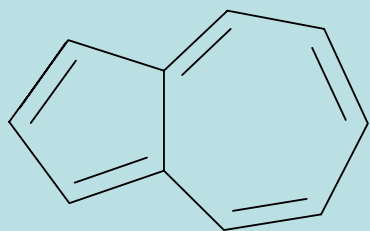
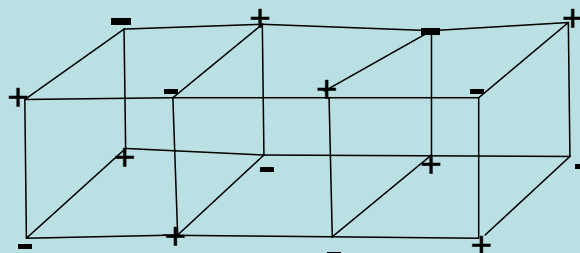
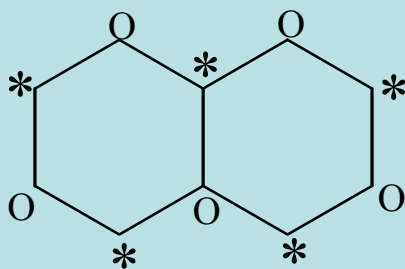
Свойства вещества зависят не только от его состава (молекулярной формулы), но и от того, в каком порядке связаны между собой атомы в молекуле

Александр Михайлович Бутлеров (1828-1886)

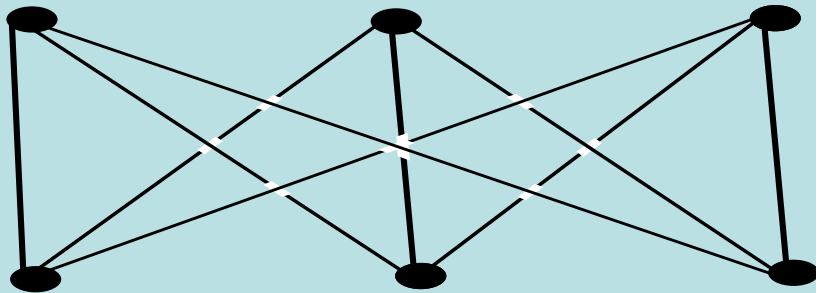
$\text{CH}_3\text{-CH}_2\text{-CH}_2\text{-CH}_3$  бутан

$\text{CH}_3\text{-CH}(\text{CH}_3)\text{-CH}_3$  изобутан

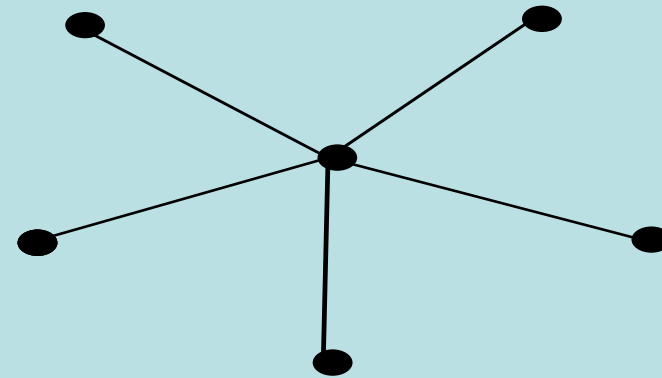
- Граф содержащий цикл включающий **все** ребра называется *эйлеровым*.
- *Степень вершины графа  $v_i$*  (кол-во ребер соединенных с данной вершиной ( $i$ )).
- *двудольный* Граф. Графы **альтернанных углеводородов** являются двудольными.



• Двудольный граф называется *полным двудольным* графом, если каждая вершина одного подмножества ( $n$  элементов  $V_1$ ) соединена со всеми вершинами (каждой вершиной) другого подмножества ( $m$  элементов  $V_2$ ).



$U_{33}$



$U_{15}$

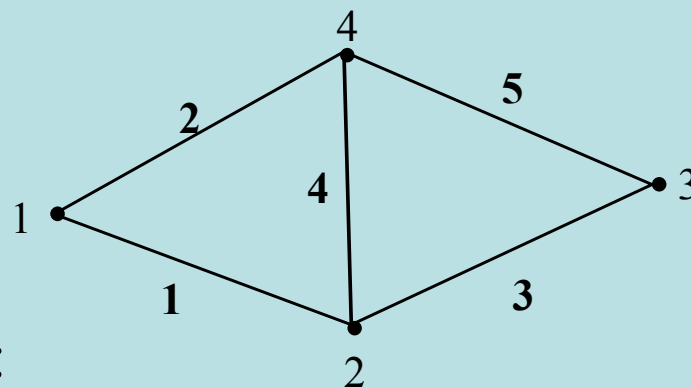
### Решение задачи о мостах (*Теорема Эйлера*)

Для того, чтобы граф был эйлеровым необходимо и достаточно, чтобы степени всех его вершин были четными (*укр.* парными).

*Необходимость:* на каждую вершину нужно войти столько раз, сколько и выйти. Т.е. степень вершины должна быть четной.



# Матричное представление Графа



Ребра этого графа:

1. {1-2},
2. {1-4},
3. {2-3},
4. {2-4},
5. {3-4}.

Матрица смежности

$$A(G) = \begin{vmatrix} 0 & 1 & 0 & 1 \\ 1 & 0 & 1 & 1 \\ 0 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix},$$

Матрица инцидентности

$$B(G) = \begin{vmatrix} 1 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 1 & 0 \\ 0 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 1 \end{vmatrix}$$

**Теорема 1** (Л.Эйлер, 1736) Сумма степеней вершин графа равна удвоенному кол-ву ребер этого графа.

$$\sum_{i=1}^N v_i = 2\beta$$

**Теорема 2** Пусть граф  $G$  содержит  $m$  ребер и  $n$  вершин. Матрица инцидентности равна  $B = B(G)$ . Тогда матрица смежности этого графа равна

$$A(G) = B B^+ - \text{diag}\{d_1, d_2, \dots, d_n\},$$

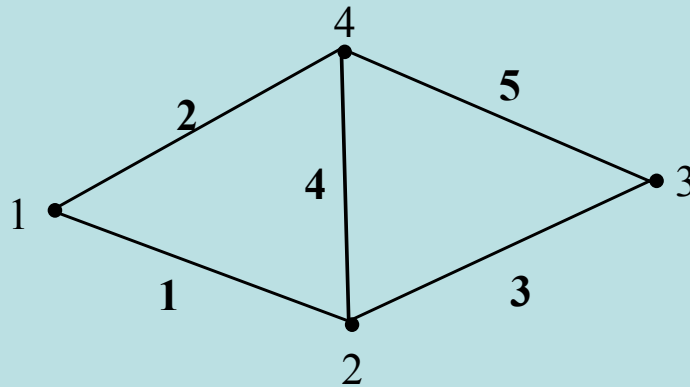
где  $\text{diag}\{d_1, d_2, \dots, d_n\}$  – диагональная матрица в которой  $d_i$  – степень вершины  $i$

**Теорема 3** Пусть граф  $G$  содержит  $m$  ребер и  $n$  вершин. Матрица инцидентности равна  $V=B(G)$ . Тогда

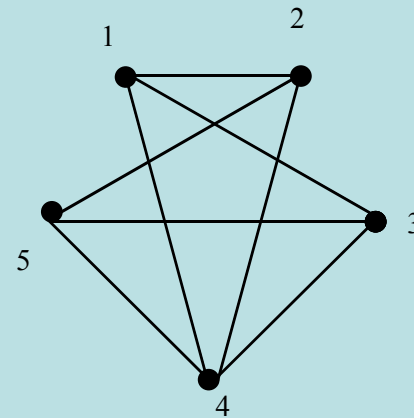
$$A_E(G) = B^+B - 2I$$

Где  $A_E(G)$  – матрица смежности реберного графа (то есть графа вершины которого соответствуют ребрам  $G$ ).

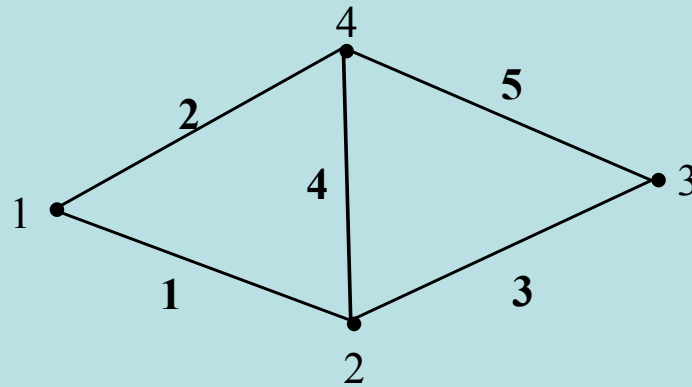
Для приведенного выше графа матрица смежности реберного графа имеет вид:



$$A_E(G) = \begin{vmatrix} 0 & 1 & 1 & 1 & 0 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 0 & 0 & 1 & 1 \\ 1 & 1 & 1 & 0 & 1 \\ 0 & 1 & 1 & 1 & 0 \end{vmatrix}$$



**матрица расстояний  $D(G)$ .**



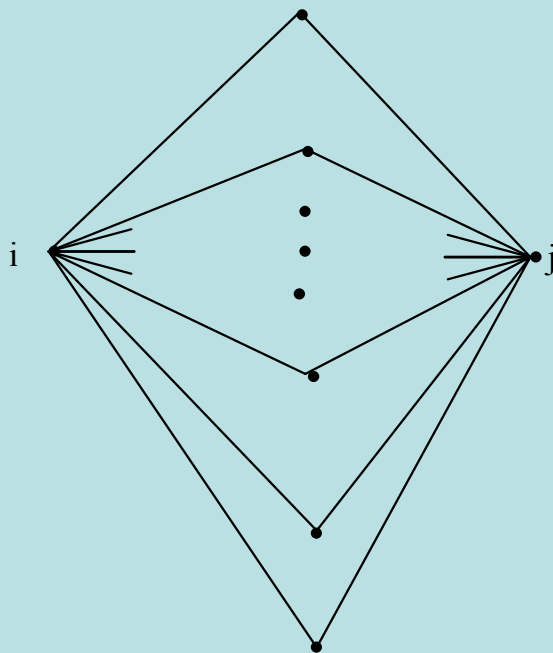
$$D(G) = \begin{vmatrix} 0 & 1 & 2 & 1 \\ 1 & 0 & 1 & 1 \\ 2 & 1 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{vmatrix}$$

**Теорема 4** Пусть  $G$  помеченный граф с вершинами  $\{1,2,3,\dots, N\}$

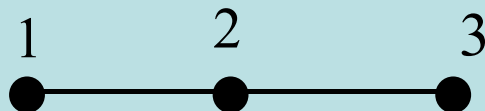
и матрицей смежности  $A(G)=A=[a_{ij}]$ .

Тогда матричный элемент  $(A^n)_{ij}$  равен числу маршрутов длины  $n$ , соединяющих вершину  $i$  с вершиной  $j$  в графе  $G$

$$n=2. \quad a_{ik} = a_{kj} = 1 \quad c_{ij} = \sum_k a_{ik} a_{kj}$$



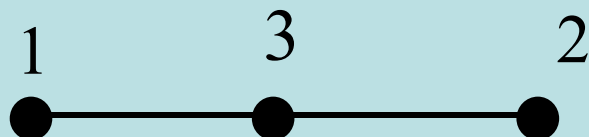
## Топологические инварианты графа для графа



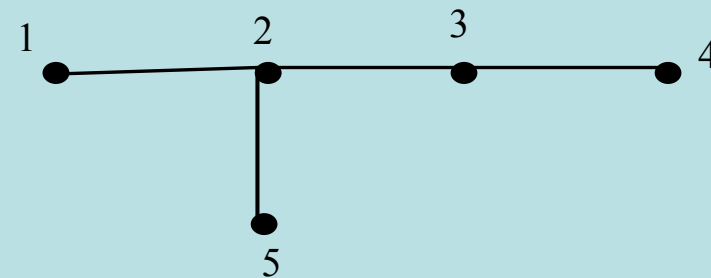
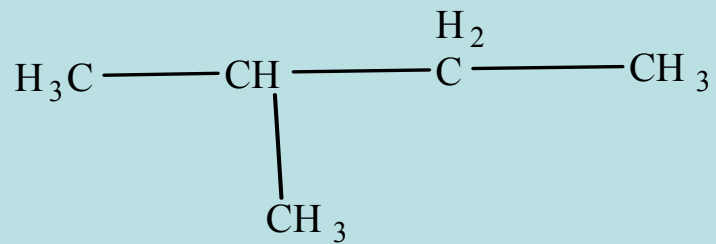
топологическая матрица имеет вид:

$$\begin{pmatrix} 0 & 1 & 0 \\ 1 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}$$

:  
а в случае такой нумерации:



$$\begin{pmatrix} 0 & 0 & 1 \\ 0 & 0 & 1 \\ 1 & 1 & 0 \end{pmatrix}$$



$$A(G) = \begin{vmatrix} 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 1 \\ 0 & 1 & 0 & 1 & 0 \\ 0 & 0 & 1 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 \end{vmatrix}$$

$$D(G) = \begin{vmatrix} 0 & 1 & 2 & 3 & 2 \\ 1 & 0 & 1 & 2 & 1 \\ 2 & 1 & 0 & 1 & 2 \\ 3 & 2 & 1 & 0 & 3 \\ 2 & 1 & 2 & 3 & 0 \end{vmatrix}$$

# Задача о числе изомеров

А.Кэли (1821-1895)



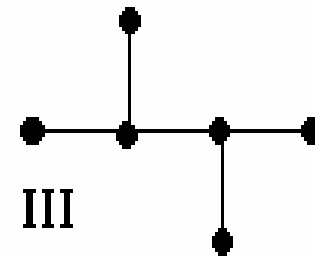
k	1	2	3	4	5	6	7	8	9	10	11	12	13
f(k)	1	1	1	2	3	5	9	18	35	75	159	357	799



I



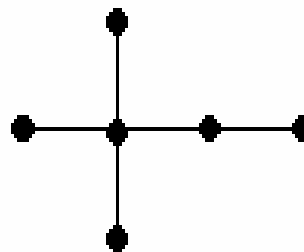
II



III



IV



V



- *индекс Винера*  $W$  
$$W = \frac{1}{2} \sum_{i,j}^N d_{ij}$$

- *индекс Рандича*  $\chi^{(1)}$  – характеристика молекулярной СВЯЗНОСТИ:

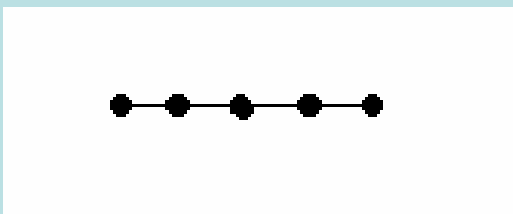
$$\chi^{(1)} = \sum_{(i,j)} (v_i v_j)^{-1/2}$$

$$\chi^{(k)} = \sum_{(i,\dots,\ell,\dots,j)} (v_i \dots v_\ell \dots v_j)^{-1/2}$$

- *Индексы загребской группы:*

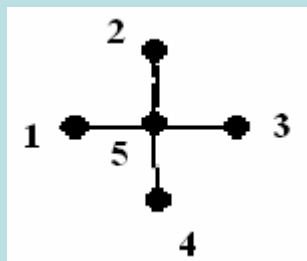
$$M_1(G) = \sum_i v_i^2$$

$$M_2(G) = \sum_{(i,j)} (v_i v_j)$$

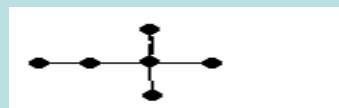


$$\chi^{(1)} = \sum_{(i,j)} (v_i v_j)^{-1/2}$$

$$\chi^{(1)} = \frac{1}{\sqrt{v_1 v_2}} + \frac{1}{\sqrt{v_2 v_3}} + \frac{1}{\sqrt{v_3 v_4}} + \frac{1}{\sqrt{v_4 v_5}} = \frac{1}{\sqrt{2}} + \frac{1}{2} + \frac{1}{2} + \frac{1}{\sqrt{2}} = 1 + \sqrt{2} = 2.414$$



$$\chi^{(1)} = \frac{1}{\sqrt{v_1 v_5}} + \frac{1}{\sqrt{v_2 v_5}} + \frac{1}{\sqrt{v_3 v_5}} + \frac{1}{\sqrt{v_4 v_5}} = \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{4}} + \frac{1}{\sqrt{4}} = \frac{4}{\sqrt{4}} = 2$$



Индекс	(н-гексан)	(2,2-диметилбутан)	(2-метилпентан)
$N_C$	6	6	6
$P_1$	5	5	5
$P_2$	4	7	5
$P_3$	3	3	3
$\chi^{(1)}$	2.914	2.561	2.770
$\chi^{(2)}$	1.707	2.914	2.183
$\chi^{(3)}$	0.957	1.061	0.866
$W$	35	28	32
$M_1(G)$	18	24	20
$M_2(G)$	16	22	18
$D^2$	7	4	5.6

Теплота атомизации алканов (ккал/моль):

$$\Delta H = 115.72 + 283.33 N - 6.321 \chi^{(1)} \quad r = 0.999, \quad \sigma = 0.960$$

**Индекс Рандича и температура кипения алканов**

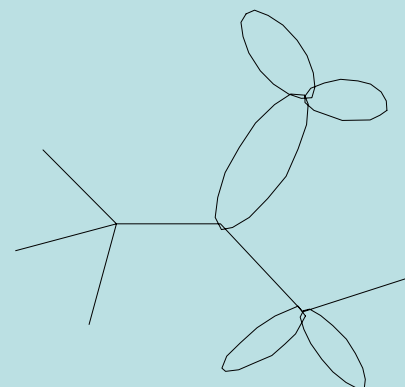
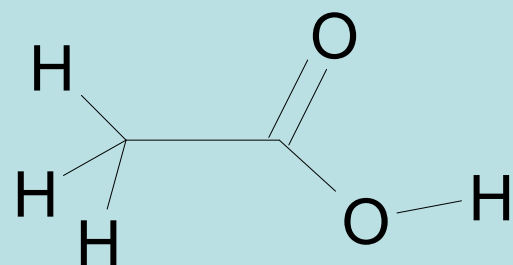
$$T(^{\circ}\text{C}) = 57.85 \chi^{(1)} - 97.9$$

	2-метил пентан	2,2-диметил бутан	2,3-диметил бутан	н-гексан
$\chi^{(1)}$	2.77	2.56	2.64	2.91
T(°C)	62.3	50.2	55.0	70.7
T(°C)exp	60.3	49.7	58.0	68.7

Минимальная блокирующая концентрация анестетиков

$$\lg \text{MBC} = -0.762 \chi^{(1)} + 3.55 \quad (n = 36, r = 0.98, \sigma = 0.39)_{20}$$

## Обобщение на системы с кратными связями и гетероатомы



$$d_{ii} = 1 - \frac{6}{Z_i}$$

$$d_{ij} = \sum_{i,j} k_{ij}$$

$$k_{ij} = \frac{1}{b_{ij}} \frac{36}{Z_i Z_j}$$

b = 1  $\sigma$ -bond

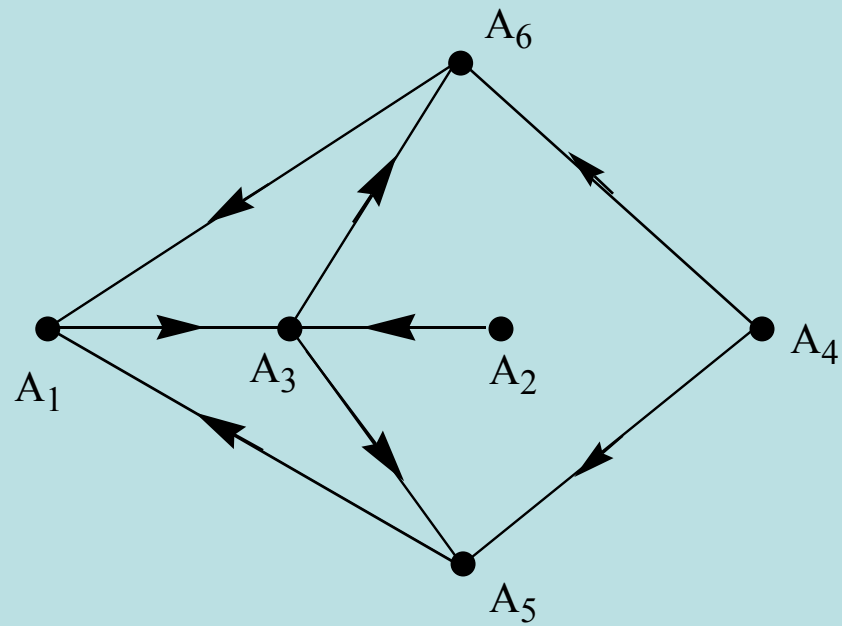
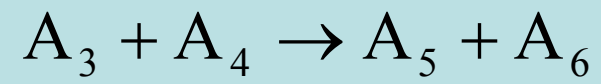
b = 2 double

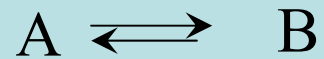
b = 3 triple

C	0
N	0.143
O	0.25
S	0.625

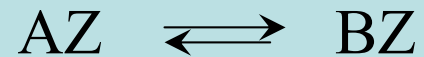
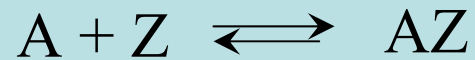
C-C	1
C=C	0.5
C $\equiv$ C	0.25

## Графы химических процессов

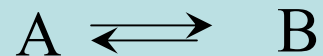
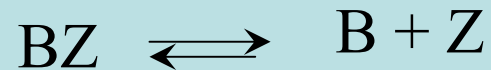




Z – катализатор



$$[Z] + [AZ] + [BZ] = 1$$



$$P = S - I + 1$$

P-число маршрутов,

I – число независимых промежуточных соединений,

S – число стадий.

$$P = 3 - 3 + 1 = 1$$

## Синтез винилхлорида



Промежуточные соединения: Z, ZC<sub>2</sub>H<sub>2</sub>, ZHCl

$$P = S - I + 1$$

Сколько маршрутов в 4-х стадийном процессе ?

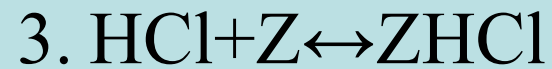
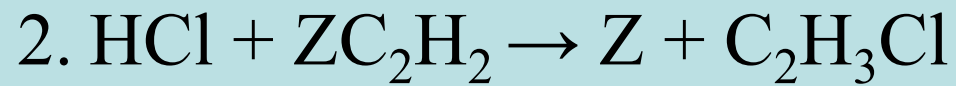
$$P = S - 3 + 1 = 4 - 3 + 1 = 2$$

Сколько стадий в 1 маршрутном процессе

$$S = P + I - 1 = 1 + 3 - 1 = 3$$

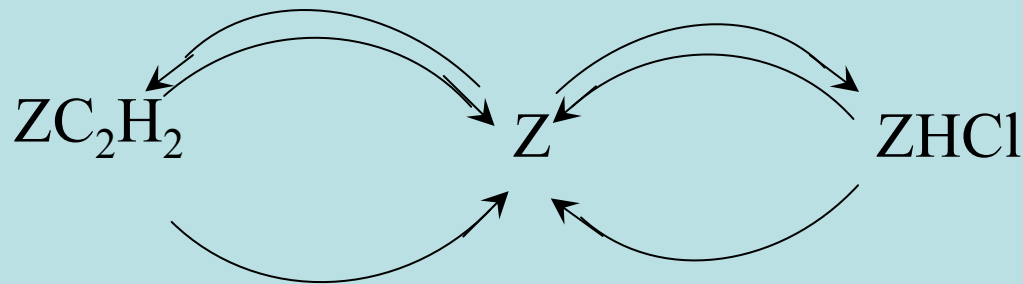


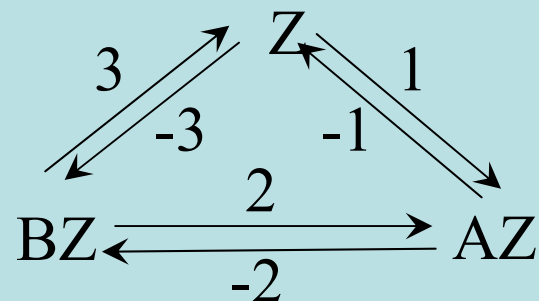
## 4-х стадийный процесс



Маршрут №1

Маршрут №2





Реакция линейна т.е. скорость (на каждой стадии) зависит от концентрации одного промежуточного соединения

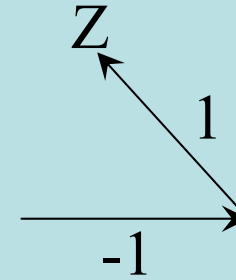
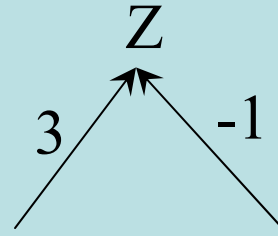
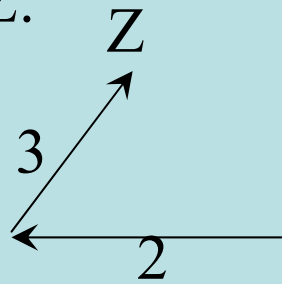
**Метод Хоурти-Темкина:** Рассматриваются только линейные комбинации  $K$  стадий схемы, в которых стехиометрические числа промежуточных соединений обращаются в нуль.

**В реакции с одним циклом ( $n$ -вершин)  $n^2$  каркасов.**

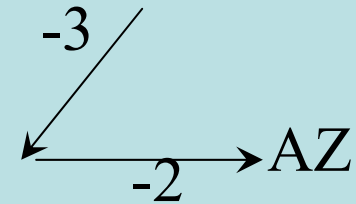
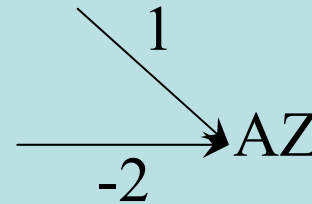
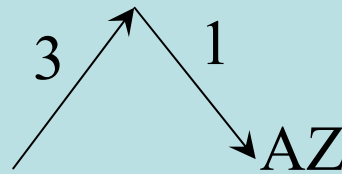
$$3^2 = 9$$

# КАРКАСЫ

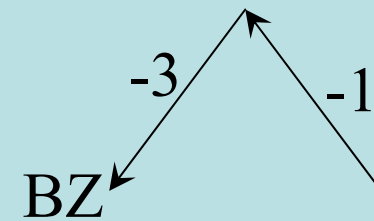
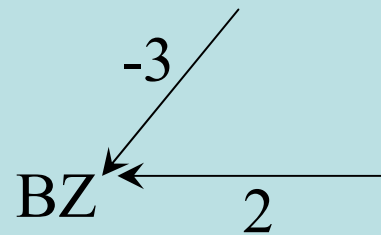
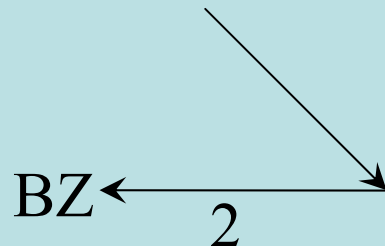
Вершина Z:



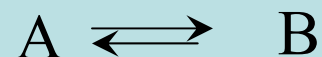
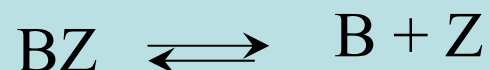
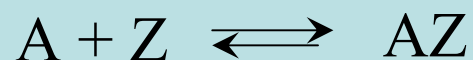
Вершина AZ:



Вершина BZ:



## Веса «дуг» прямой (+) и обратной (-) реакций



$$b_s^+ = w_s^+ / [x_i^+]$$

$$b_s^- = w_s^- / [x_i^-]$$

$$w_1^+ = k_1^+ [A][Z]$$

$$w_2^+ = k_2^+ [AZ]$$

$$w_3^+ = k_3^+ [BZ]$$

$$w_1^- = k_1^- [AZ]$$

$$w_2^- = k_2^- [BZ]$$

$$w_3^- = k_3^- [B][Z]$$

## Весы дуг

$$b_1^+ = k_1^+[A]$$

$$b_2^+ = k_2^+$$

$$b_3^+ = k_3^+$$

$$b_1^- = k_1^-$$

$$b_2^- = k_2^-$$

$$b_3^- = k_3^-[B]$$

## Весы каркасов

Каркас Z:  $B_{Z,1} = b_2^+ b_3^+$      $B_{Z,2} = b_3^+ b_1^-$      $B_{Z,3} = b_2^- b_1^-$

Каркас AZ:  $B_{AZ,1} = b_1^+ b_3^+$      $B_{AZ,2} = b_1^+ b_2^-$      $B_{AZ,3} = b_3^- b_2^-$

Каркас BZ:  $B_{BZ,1} = b_1^+ b_2^+$      $B_{BZ,2} = b_2^+ b_3^-$      $B_{BZ,3} = b_1^- b_3^-$

## Суммарный вес каркасов вершин

$$\text{Каркас } Z: \quad B_Z = B_{Z,1} + B_{Z,2} + B_{Z,3}$$

$$\text{Каркас } AZ: \quad B_{AZ} = B_{AZ,1} + B_{AZ,2} + B_{AZ,3}$$

$$\text{Каркас } BZ: \quad B_{BZ} = B_{BZ,1} + B_{BZ,2} + B_{BZ,3}$$

## Суммарный вес каркасов графа

$$B = B_Z + B_{AZ} + B_{BZ}$$

$$[x] = B_X / B$$

$$W = k_2^+[AZ] - k_2^-[BZ]$$

$$W = k_2^+[AZ] - k_2^-[BZ]$$

$$W = \frac{k_1^+k_2^+k_3^+[A] - k_1^-k_2^-k_3^-[B]}{k_1^+[A](k_2^+ + k_3^+ + k_2^-) + k_3^-[B](k_1^- + k_2^- + k_2^+) + k_2^+k_3^+ + k_1^-k_2^- + k_3^-k_1^-}$$

**Если все стадии необратимы,**

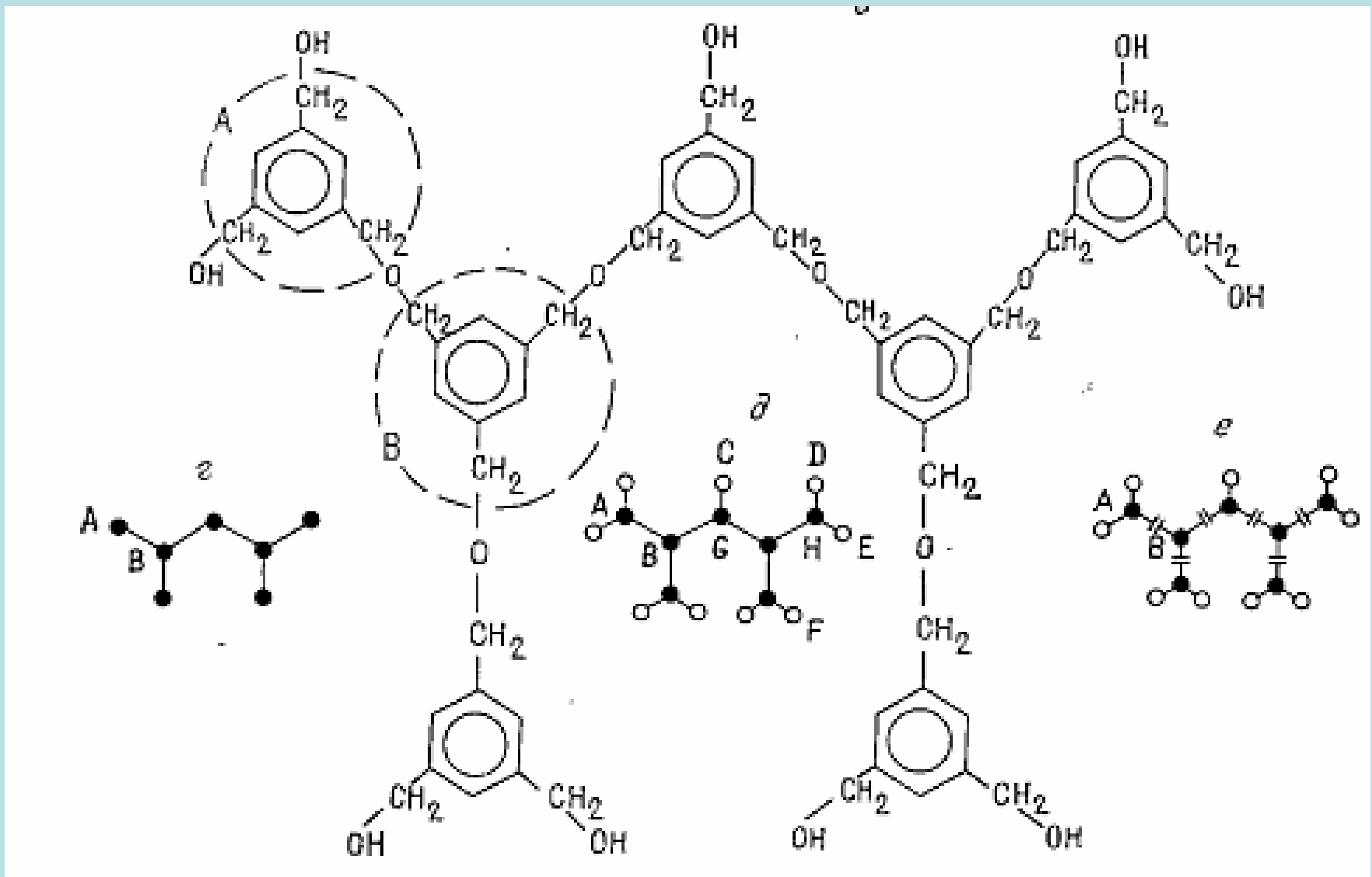
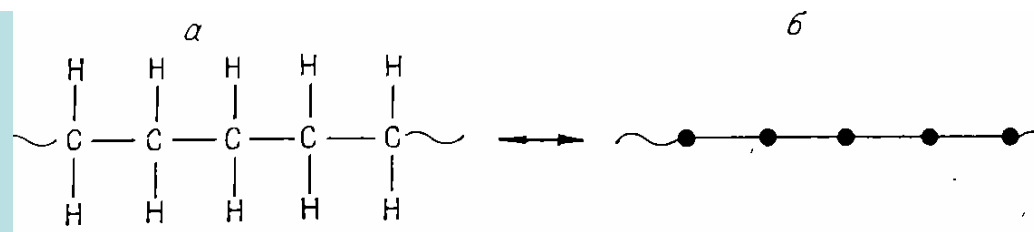
$$W = \frac{k_1^+k_2^+k_3^+[A]}{k_1^+[A](k_2^+ + k_3^+) + k_2^+k_3^+}$$

## **Теория графов в описании физико-химических свойств полимеров**

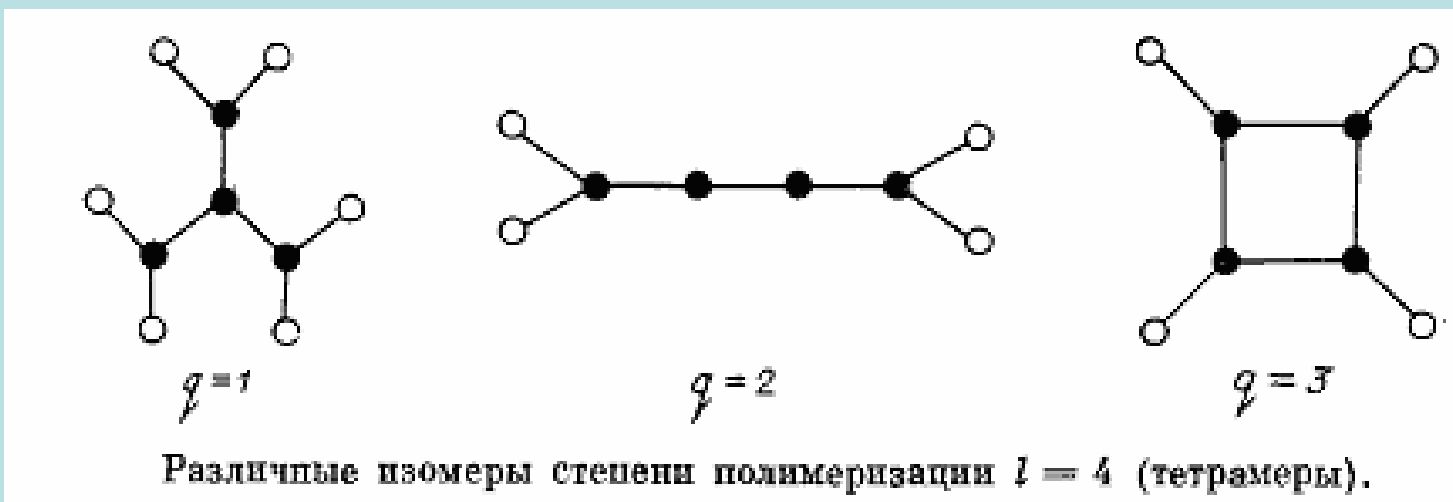
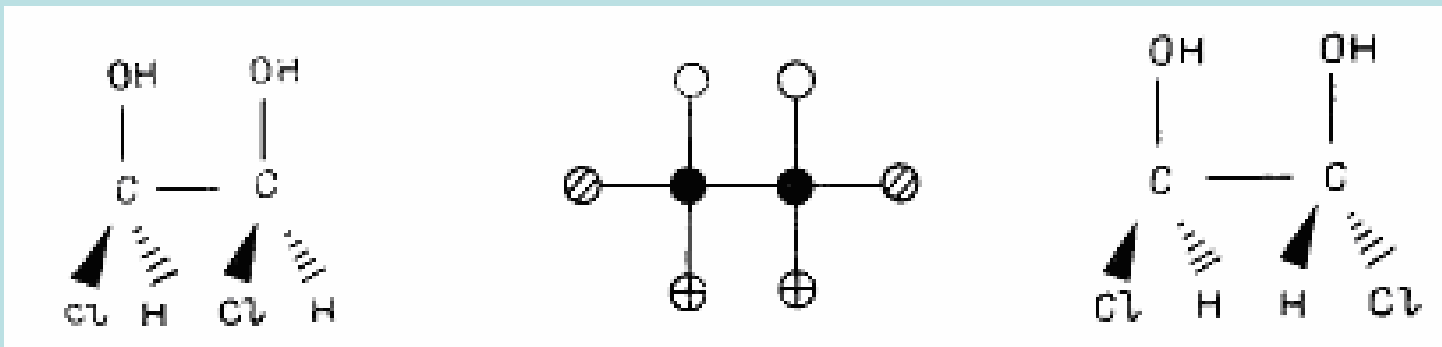
В химической физике полимеров решение многих задач значительно упрощается, если их удастся сформулировать в терминах теории графов.

По аналогии с углеводородами многие «структурно - аддитивные» свойства полимеров могут быть рассчитаны, исходя из средних чисел различных фрагментов малого размера в макромолекулах.





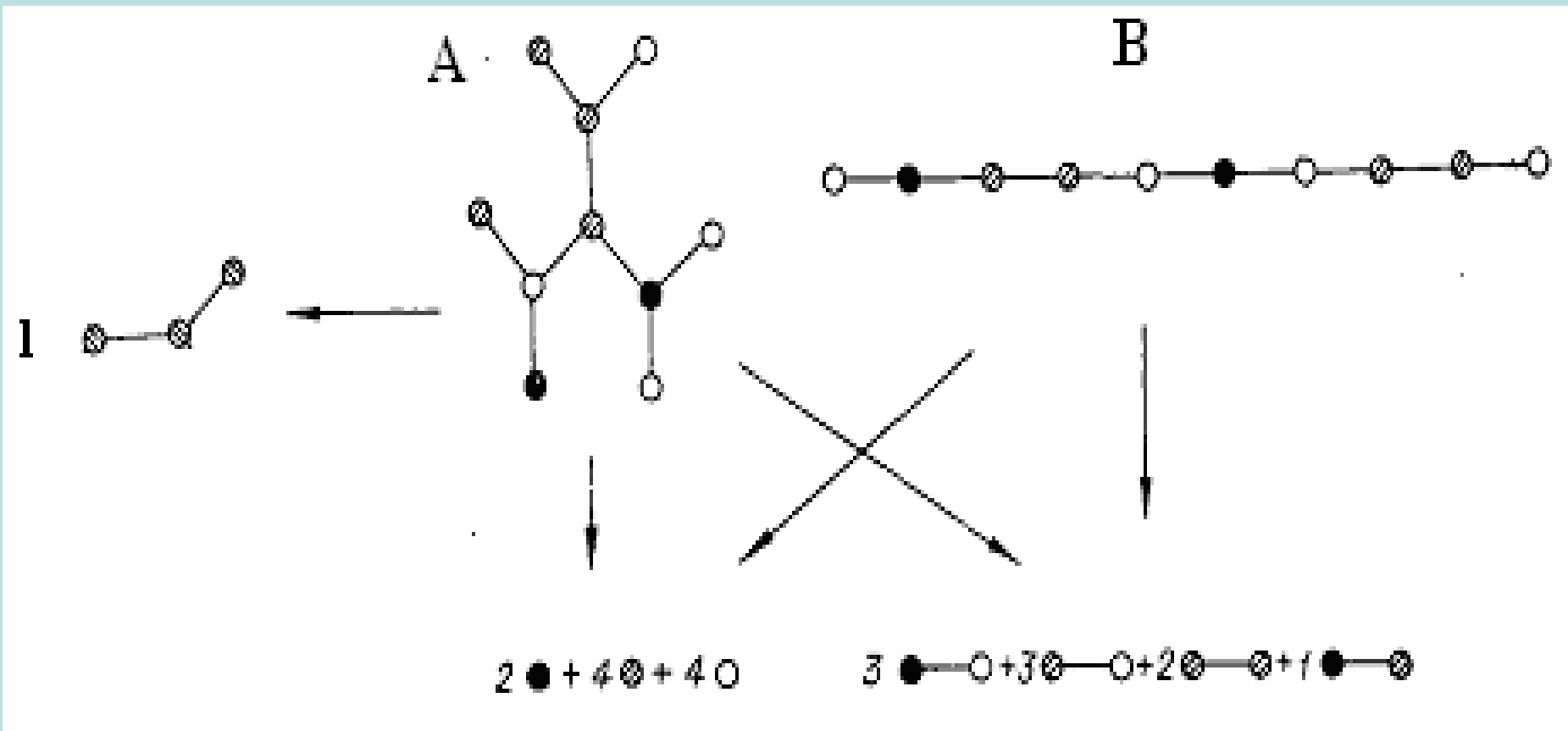
Молекулярный граф не всегда однозначно определяет индивидуальное химическое соединение.



Структура полимерной цепи определяется парой чисел  $(l, q)$

# Для полимеров топологическая матрица может быть очень большой !

локальную структуру можно задать, указав число вершин  
разного цвета в графе, различных пар смежных вершин, троек  
и т.



**Структурно-аддитивные свойства полимера** зависят от числа содержащихся в молекулярном графе подграфов того или иного вида с небольшим числом вершин.

В простейшем случае такое свойство, как **энтальпия образования молекулы**, задается только числом в ней разнотипных химических связей. Поэтому для ее расчета достаточно определить в раскрашенном графе числа подграфов, состоящих из пар вершин всевозможных комбинаций цветов

$$\Delta H_A = 3 (\bullet - \circ) + 3 (\otimes - \circ) + 2 (\otimes - \otimes) + 1 (\bullet - \otimes)$$

$$\Delta H_B = 3 (\bullet - \circ) + 3 (\otimes - \circ) + 2 (\otimes - \otimes) + 1 (\bullet - \otimes)$$

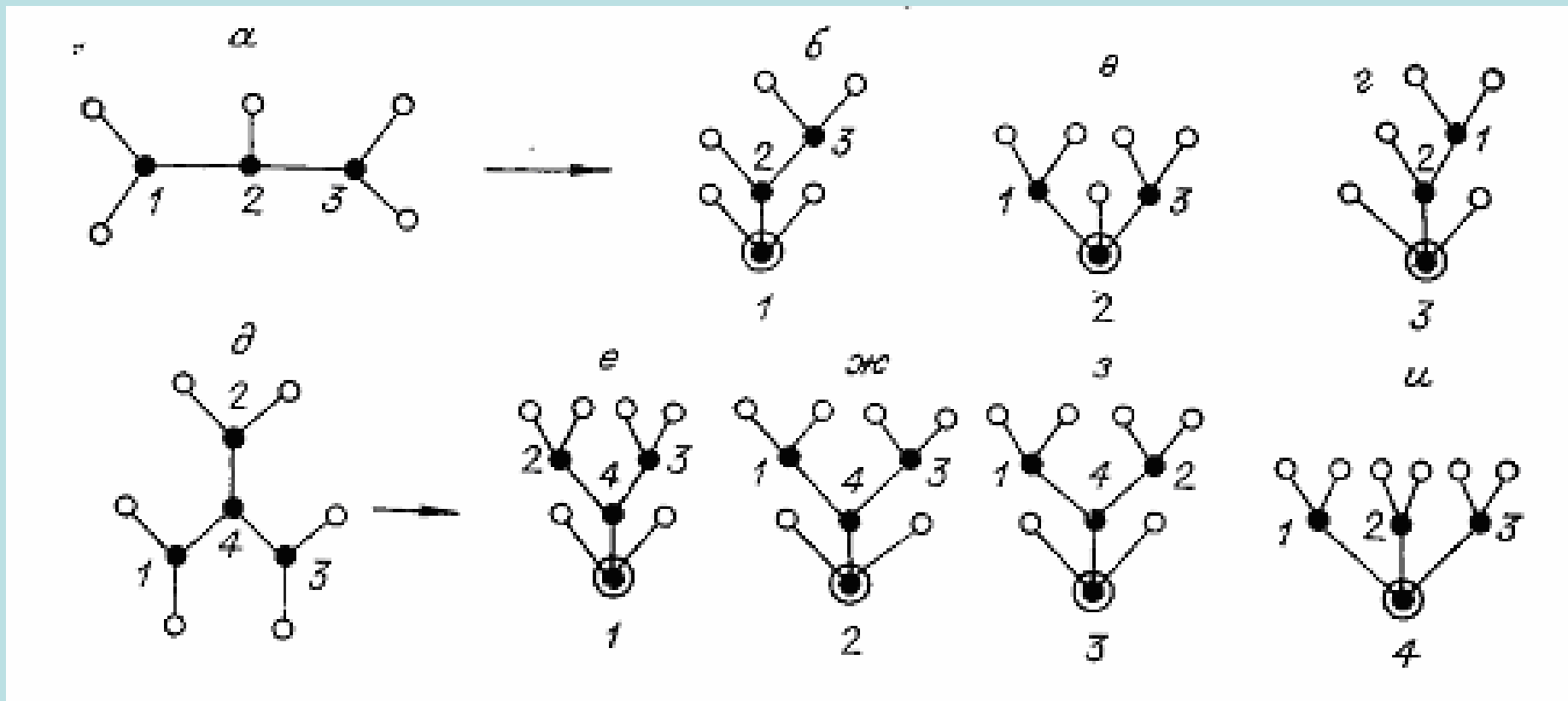
$$\Delta H_A = \Delta H_B + 1 (\otimes - \otimes - \otimes)$$

Среди свойств молекул, не являющихся структурно-аддитивными, имеются такие, которые определяются не локальной, а **глобальной структурой графа**. Такие свойства зависят от набора конформаций молекулы. Каждая из конформаций некоторого изомера характеризуется взаимным расположением его фрагментов в пространстве.

## Ансамбли полимерных молекул и случайные графы

На множестве всех возможных молекулярных графов задается **вероятностная мера**, так что вероятность определенного молекулярного графа пропорциональна концентрации изображаемых им молекул в полимерной системе.

### Корневые графы



Изучение **индивидуальных химических соединений**, состоящих из **тождественных молекул**, измеряемую характеристику можно считать **относящейся к любой конкретной молекуле**.

**Физико-химические свойства полимерных систем** определяются **вкладами** всех **составляющие** ее  $(l, q)$ -изомеров,

$$f_N(l) = \frac{c(l)}{\sum_l c(l)} \quad f_w(l) = \frac{lc(l)}{\sum_l lc(l)} \quad f_z(l) = \frac{l^2c(l)}{\sum_l l^2c(l)}$$

молекулярно-структурные распределения:

$$f_N(l, q), f_w(l, q), f_z(l, q),$$

различают молекулы в соответствии с их молекулярными графами

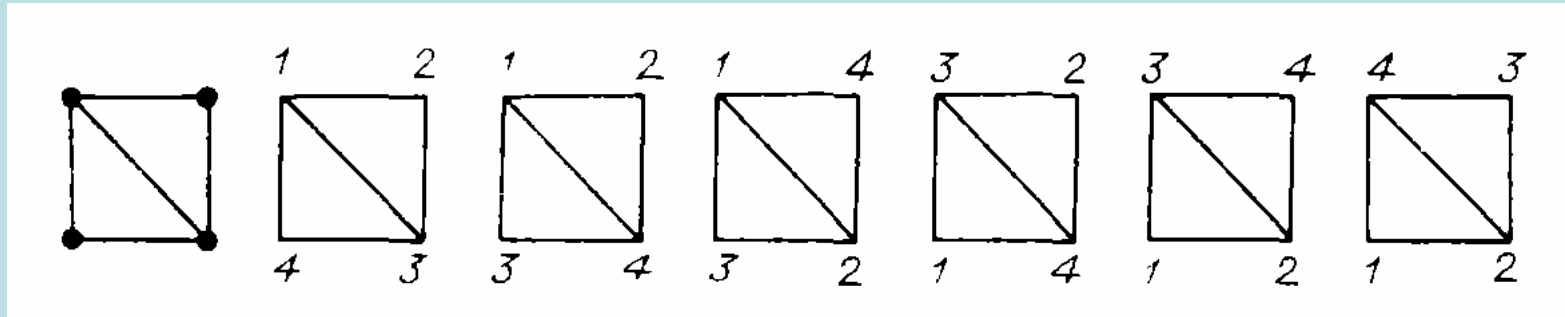
рассмотрим помеченные (пронумерованные) графы, в которых каждой из  $n$  вершин приписывается отличное от других целое число от 1 до  $n$  ( $n!$ ).

но при этом некоторые из них могут совпадать. Так, из  $4! = 24$

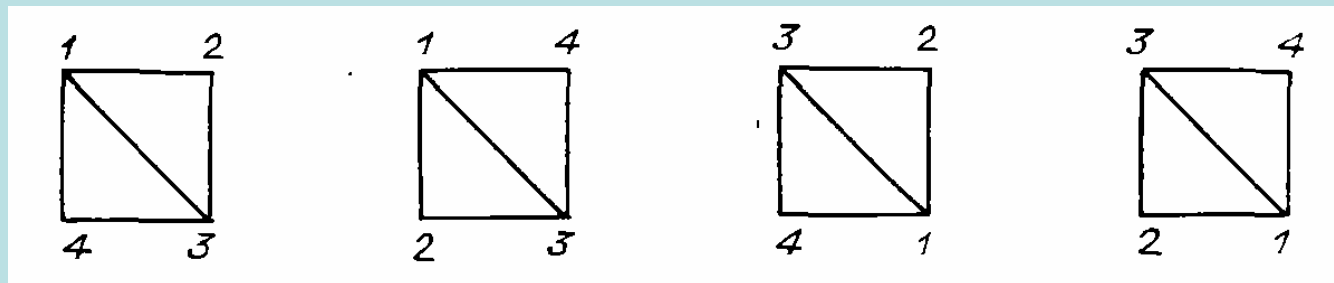
Два (помеченных) графа считаются одинаковыми и называются изоморфными, одинаковыми и называются изоморфными, если существует взаимнооднозначное отображение множества вершин одного графа на множество вершин другого, при котором сохраняется смежность вершин (и распределение пометок на них).



Непомеченный граф и полный набор различных его нумераций.



Полный набор изоморфов (S) первого среди помеченных



Число различных (неизоморфных) нумераций

$$W = n! / S$$

Вероятность состояния определяется лишь кратностью его вырождения, т. е. числом различных способов объединения  $N$  мономеров в заданную конфигурацию. При переходе из состояния 1 в состояние 2 изменение комбинаторной энтропии связано с **порядком группы автоморфизма**

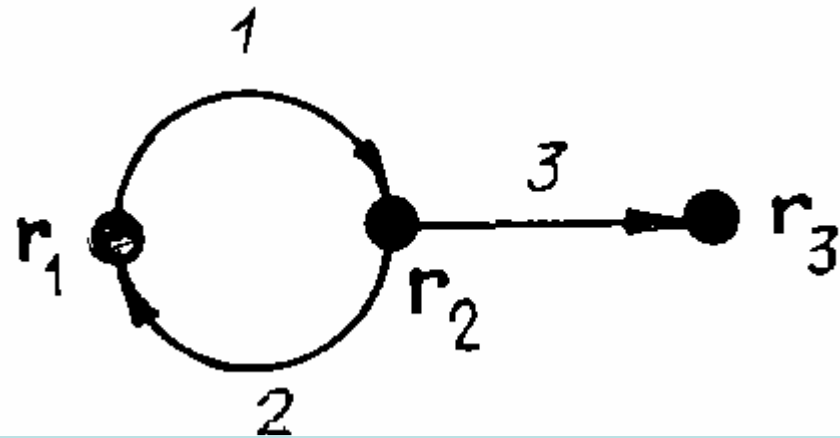
**Автоморфизм графа** есть отображение множества вершин на себя, сохраняющее смежность

$$\Delta S = S_2 - S_1 = R \ln \left( \frac{N(2)}{N(1)} \right)$$

$N(i)$  - число способов перевести в себя молекулярный граф мономера.

Кратность вырождения состояния  $i$ :

$$W(i) = \exp(\Delta S / R)$$



$$\mathbf{A} = \begin{pmatrix} 0 & 2 & 0 \\ 2 & 0 & 1 \\ 0 & 1 & 0 \end{pmatrix}, \quad \mathbf{B} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix},$$

матрица Кирхгофа,

$$\mathbf{K} = \mathbf{B}\mathbf{B}^T.$$

$$\mathbf{K} = \begin{pmatrix} 2 & -2 & 0 \\ -2 & 3 & -1 \\ 0 & -1 & 1 \end{pmatrix} = \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 0 \\ 0 & 1 & -1 \end{pmatrix} = \mathbf{B}\mathbf{B}^T.$$

Если соседние мономерные звенья соединены достаточно длинной полтмерной цепью, Конформация макромолекулы в такой модели характеризуется координатами  $r_i$  ее звеньев, а энергия имеет вид:

$$U = kT \sum_{i,j} \gamma_{ij} (r_i - r_j)^2$$

$$\begin{pmatrix} \mathbf{r}_{11} - \mathbf{r}_{12} & \mathbf{r}_{12} - \mathbf{r}_{11} & \mathbf{r}_{12} - \mathbf{r}_{13} \\ \mathbf{r}_{21} - \mathbf{r}_{22} & \mathbf{r}_{22} - \mathbf{r}_{21} & \mathbf{r}_{22} - \mathbf{r}_{23} \\ \mathbf{r}_{31} - \mathbf{r}_{32} & \mathbf{r}_{32} - \mathbf{r}_{31} & \mathbf{r}_{32} - \mathbf{r}_{33} \end{pmatrix} =$$

$$= \begin{pmatrix} \mathbf{r}_{11} & \mathbf{r}_{12} & \mathbf{r}_{13} \\ \mathbf{r}_{21} & \mathbf{r}_{22} & \mathbf{r}_{23} \\ \mathbf{r}_{31} & \mathbf{r}_{32} & \mathbf{r}_{33} \end{pmatrix} \begin{pmatrix} 1 & -1 & 0 \\ -1 & 1 & 1 \\ 0 & 0 & -1 \end{pmatrix}.$$

Поэтому энергия молекулы просто выражается через матрицу Кирхгофа ее графа:

$$\sum_{i-j} |\mathbf{r}_i - \mathbf{r}_j|^2 = \text{Tr}(\mathbf{RB}(\mathbf{RB})^T) = \text{Tr}(\mathbf{RBB}^T\mathbf{R}^T) = \text{Tr}(\mathbf{RKR}^T),$$

## Литература

1. Скоробогатов В.А. Алгоритмический анализ графов, 1988
2. Применение теории графов в химии. (под редакцией Н.С.Зефирова), Новосибирск «Наука», 1988, 306 с.

«В любой науке столько истины, сколько в  
ней математики»

*Иммануил Кант (1724-1804)*