*Journal of Statistical Physics, Vol. 4, No. 2/3, 1972*

# On Interesting Walks in a Graph

Witold Brostow[1,3] and Andrzej Schinzel[2]

*Received January 27, 1971*

Notions of interesting walks and of their equivalence are introduced. A general formula for the number $\Gamma_l$ of equivalence classes of interesting walks of length $l$ in a given graph $G$ is derived and applied for $l \leqslant 5$ so as to express $\Gamma_l$ in terms of the adjacency matrix of $G$.

**KEY WORDS:** Graph theory; molecular structure; walks.

## 1. INTRODUCTION

A molecule of a chemical compound may be represented by a graph—a fact known to Cayley[1] more than a century ago, but little exploited since by chemists and physicists. In a liquid-state theory,[2] thermodynamic properties of substances and mixtures are expressed in terms of structural contributions from segments of various types contained in molecules and from interactions of such segments. Such contributions are determined by numbers of certain classes of walks on the graphs.

For the definition of a graph, a walk, and other basic graph-theoric notions not defined here, see, e.g., Harary[3]. To define a physically non-redundant walk in a graph, we introduce the following notation: for a given walk $w$, $E(w)$, $P(w)$, $L(w)$, and $l(w)$ are the set of endpoints, the set of points, the set of lines, and the length of $w$, respectively. We call two walks $u$ and $w$ equivalent, $u \sim w$, if

$$L(u) = L(w) \tag{1}$$

$$l(u) = l(w) \tag{2}$$

$$E(u) = E(w) \tag{3}$$

[1] Department of Chemical Engineering, University of Ottawa, Ottawa, Canada.
[2] Instytut Matematyczny Polskiej Akademii Nauk, Warsaw, Poland.
[3] Present address: Department of Chemistry, University of Montreal, Montreal, Canada.

A walk $w$ is considered to be redundant if there is a walk $u$ nonequivalent to $w$ such that (3) holds and also

$$L(u) \subset L(w) \tag{4}$$

$$P(u) = P(w) \tag{5}$$

$$l(u) \leqslant l(w) \tag{6}$$

Relevant for the applications mentioned above is the number $\Gamma_l(G)$ of equivalence classes of open nonredundant walks of length $l$ in a given graph $G$ (note that a walk equivalent to a nonredundant one is itself nonredundant). We call such walks interesting.

A problem originating in sociology—communication relationships in a group of people— treated by Ross and Harary[4] involves counting walks. In distinction to the present case, a graph representing their problem has all lines directed; consequently, the sociometric matrix $M$ is asymmetric. Moreover, Ross and Harary assume all walks which are not paths to be redundant. The difference may be seen by considering, e.g., a walk $w$ of shape $\gamma_2$, i.e., the walk number 2 for $l = 3$ in Fig. 1. Label the points in $w$ as $a$, $b$, and $c$, with $a$ at the left end. The walk is not a path, as the point $b$ occurs twice. However, the walk under consideration is interesting, as it is neither closed nor is there any walk $u$ of $l \leqslant 3$ fulfilling the conditions (3)–(5); from the point of view of the problem mentioned in the beginning of this section, the walk $w$ represents the effect of proximity of $c$ upon the $a \cdots b$ system.

In this paper, we give the general formula for $\Gamma_l(G)$ and apply it for $l \leqslant 5$ in order to express $\Gamma_l(G)$ in terms of the adjacency matrix of $G$. In principle, the same method may be used for $l > 5$; complications, however, grow quickly, as in the case of Harary and Manvel,[5] who count cycles of length $n \leqslant 5$ and find that "the situation gets rather out of hand for higher $n$."

## 2. GENERAL FORMULA

In order to find $\Gamma_l(G)$, we define the shape of a walk. We call two walks $u$ and $w$ isomorphic if there is a monomorphism $f$ on $P(u)$ such that $w = f(u)$, i.e., $u = u_1 u_2 \cdots u_n$, where $u_1, u_2, \ldots, u_n$ are points, and $w = f(u_1) f(u_2) \cdots f(u_n)$. The class of walks isomorphic to $w$ is the shape of $w$, denoted by $\sigma(w)$. Since $l(u) = l(w)$, we set $l(\sigma) = l(u)$ and call $l(\sigma)$ the length of $\sigma$. We prove the following lemma.

**Lemma.** A walk isomorphic to an interesting one is itself interesting.

*Proof.* Assume that $f$ is a monomorphism on $P(w)$, that $f(w)$ is interesting, and that $w$ is not. Clearly, $w$ is not closed, hence it is redundant. Therefore, there exists a walk $u$ nonequivalent to $w$ satisfying (3)–(6). We get

$$\begin{aligned} E(f(u)) &= fE(u) = fE(w) = E(f(w)) \\ P(f(u)) &= fP(u) = fP(w) = P(f(w)) \\ L(f(u)) &= fL(u) \subset fL(w) = L(f(w)) \\ l(f(u)) &= l(u) \leqslant l(w) = l(f(w)) \end{aligned}$$

Since $f(w)$ is interesting, we infer that $f(u) \sim f(w)$. But then

$$\begin{aligned} L(f(u)) &= L(f(w)) \\ l(f(u)) &= l(f(w)) \\ E(f(u)) &= E(f(w)) \end{aligned}$$

and since $f$ is a monomorphism,

$$\begin{aligned} L(u) &= L(w) \\ l(u) &= l(w) \\ E(u) &= E(w) \end{aligned}$$

and $u \sim w$, a contradiction.

In view of the lemma, we define an interesting shape as the shape of an interesting walk. Two shapes $\sigma_1$ and $\sigma_2$ are called equivalent if there are equivalent walks $w_1$ and $w_2$ such that $\sigma_i = \sigma(w_i)$, with $i = 1, 2$. The relation has the usual properties, but this fact is not necessary for the proof of the following theorem.

**Theorem.** The number $\Gamma_l(G)$ of equivalence classes of interesting walks of length $l$ in $G$ is given by the fomula

$$\Gamma_l(G) = \sum_{\sigma \in S} [w(\sigma, G)/e(\sigma)] \tag{7}$$

where $S$ is any maximal set of pairwise nonequivalent interesting shapes $\sigma$ of length $l$, $w(\sigma, G)$ is the number of walks of shape $\sigma$ in $G$, and $e(\sigma)$ is the maximal number of distinct equivalent walks of shape $\sigma$.

*Proof.* Let $w$ be any interesting walk. Since $S$ is maximal, $\sigma(w)$ is equivalent to a certain shape $\sigma_l \in S$. By definition of the equivalence of shapes, there exist walks $u$ and $t$ such that

$$u \sim t, \qquad \sigma(u) = \sigma(w), \qquad \sigma(t) = \sigma_l \in S.$$

By the definition of a shape, there exists a monomorphism $f$ on $P(u)$ such that $w = f(u)$; on the other hand,

$$L(u) = L(t), \qquad l(u) = l(t), \qquad E(u) = E(t)$$

It follows that

$$L(w) = L(f(u)) = fL(u) = fL(t) = L(f(t))$$
$$l(w) = l(f(u)) = l(u) = l(t) = l(f(t))$$
$$E(w) = E(f(u)) = fE(u) = fE(t) = E(f(t))$$

hence $w \sim f(t)$ and $\sigma(f(t)) = \sigma(t) \in S$. Therefore, any equivalence class $C$ is represented by a walk of one and only one shape $\sigma \in S$. We prove that it is represented by at least $e(\sigma)$ walks of shape $\sigma$. Indeed, let $w \in C$, $\sigma(w) = \sigma$. By the definition of $e(\sigma)$, there exist distinct equivalent walks $w_i$ $[i = 1, 2, \ldots, e(\sigma)]$ with $\sigma(w_i) = \sigma$. Since $\sigma(w_1) = \sigma(w)$, there is a monomorphism $g$ on $P(w)$ such that $w = g(w_1)$. It follows from

$$L(w_i) = L(w_1), \qquad l(w_i) = l(w_1), \qquad E(w_i) = E(w_1), \qquad 1 \leqslant i \leqslant e(\sigma)$$

that

$$L(g(w_i)) = L(w), \qquad l(g(w_i)) = l(w), \qquad E(g(w_i)) = E(w)$$

hence the walks $g(w_i)$ all of shape $\sigma$ are equivalent to $w$. Since $g$ is a monomorphism, they are also distinct. Since $e(\sigma)$ is the maximal number of distinct equivalent walks of shape $\sigma$, $C$ is represented exactly by $e(\sigma)$ walks of shape $\sigma$. Thus, shape $\sigma$ represents $w(\sigma, G)/e(\sigma)$ equivalence classes and the total number of the classes is given by formula (7).

## 3. COUNTING WALKS FOR $l \leqslant 5$

Denote the walk reverse to $w$ by $w'$. It is easy to see that if $\sigma(u) = \sigma(w) = \sigma$, then $\sigma(u') = \sigma(w')$. This permits us to call $\sigma' = \sigma(u')$ the reverse shape of $\sigma(u)$. In Fig. 1 we have adopted the convention of drawing shapes beginning from the left end, so that the respective reverse shapes begin from the right. The figure shows all possible shapes of walks for $l \leqslant 5$. If $\sigma = \sigma'$, then $e(\sigma) \geqslant 2$, since clearly $w \sim w'$.

Let $G$ have $p$ points labeled $v_1, v_2, \ldots, v_p$ and let $A$ be the adjacency matrix of $G$. Set

$$A^{(l)} = [a_{ij}^{(l)}]$$

$$a_i^{(l)} = \sum_{j=1}^{p} a_{ij}^{(l)}$$

$$a^{(l)} = \sum_{i=1}^{p} a_i^{(l)}$$

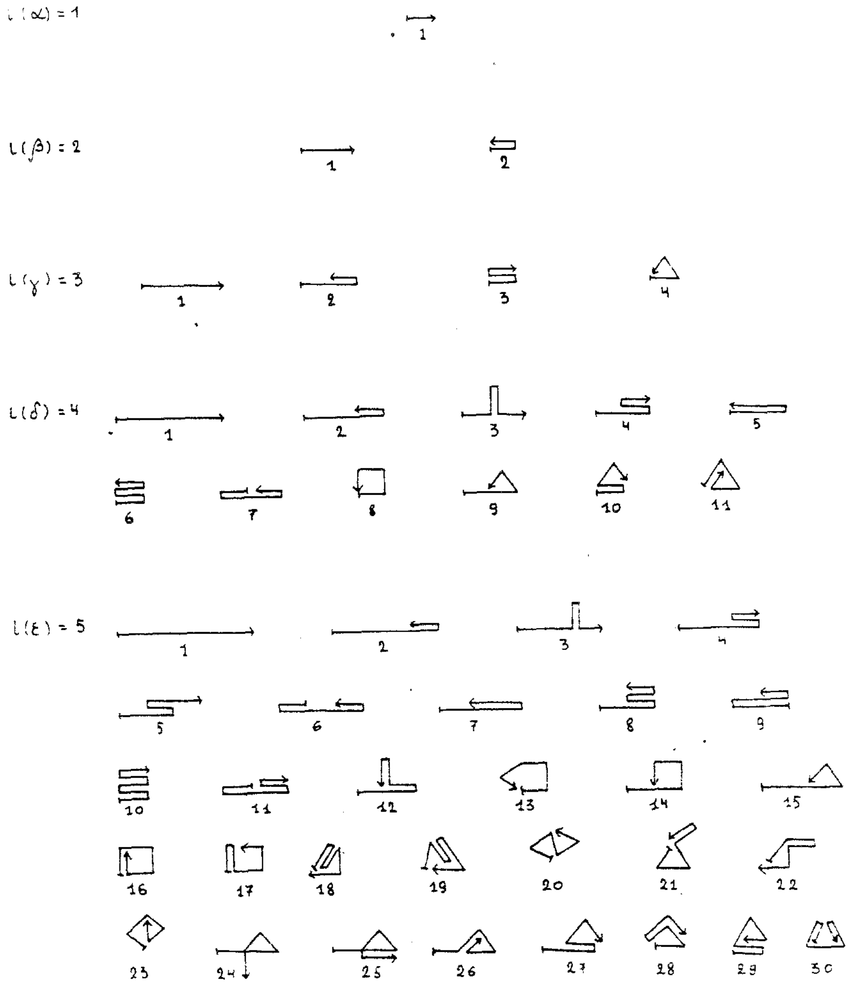


Fig. 1

It is known (cf., e.g., Theorem 13.1 in Ref. 3) that $a_{ij}^{(l)}$ is the number of walks of length $l$ between points $v_i$ and $v_j$; hence, the total number of open walks of length $l$ in $G$ is

$$w_l = w_l(G) = a^{(l)} - \operatorname{tr} A^l$$

We proceed to compute $\Gamma_l(G) = \Gamma_l$ in terms of $A$. We abbreviate $w(G, \sigma)$ to $w(\sigma)$; note that $w(\sigma) = w(\sigma')$. We substitute for $\sigma$ the letters $\alpha, \ldots, \epsilon$ with the subscripts as used in Fig. 1.

For the consecutive values of $l$, we obtain as follows.

$l = 1$. $S = \{\alpha_1\}$, $e(\alpha_1) = 2$. Hence,

$$\Gamma_1 = \tfrac{1}{2}\, w(\alpha_1) = \tfrac{1}{2} w_1 = \tfrac{1}{2} a^{(1)} = \tfrac{1}{2}\,\mathrm{tr}\, A^2$$

$l = 2$. $S = \{\beta_1\}$, $e(\beta_1) = 2$. Hence,

$$\Gamma_2 = \tfrac{1}{2} w(\beta_1) = \tfrac{1}{2} w_2 = \tfrac{1}{2}(a^{(2)} - \mathrm{tr}\, A^2) = \tfrac{1}{2}(a^{(2)} - a^{(1)})$$

$l = 3$. $S = \{\gamma_1, \gamma_2\}$, $e(\gamma_1) = 2$, $e(\gamma_2) = 1$. Hence,

$$\begin{aligned}\Gamma_3 &= \tfrac{1}{2} w(\gamma_1) + w(\gamma_2) = \tfrac{1}{2}[w(\gamma_1) + w(\gamma_2) + w(\gamma_2)] \\ &= \tfrac{1}{2}[w_3 - w(\gamma_3)] = \tfrac{1}{2}[w_3 - w_1] = \tfrac{1}{2}(a^{(3)} - a^{(1)} - \mathrm{tr}\, A^3]\end{aligned}$$

$l = 4$. $S = \{\delta_1, \delta_2, \delta_3, \delta_9\}$, $e(\delta_1) = e(\delta_3) = 2$, $e(\delta_2) = 1$, $e(\delta_9) = 2$.

To see the last equality, note that the walks (12342) and (12432) are both of shape $\delta_9$. Hence,

$$\begin{aligned}\Gamma_4 &= \tfrac{1}{2} w(\delta_1) + w(\delta_2) + \tfrac{1}{2} w(\delta_3) + \tfrac{1}{2} w(\delta_9) \\ &= \tfrac{1}{2}[w_4 - 2w(\delta_4) - w(\delta_9) - 2w(\delta_{10}) - w(\delta_{11})]\end{aligned}$$

Now,

$$w(\delta_4) = w(\beta_1) = w_2 = a^{(2)} - a^{(1)}$$

$$w(\delta_9) = \sum_{i=1}^{p} a_{ii}^{(3)}(a_i^{(1)} - 2) = \sum a_{ii}^{(3)} a_i^{(1)} - 2\,\mathrm{tr}\, A^3$$

$$w(\delta_{10}) = w(\delta_{11}) = w(\gamma_4) = \mathrm{tr}\, A^3$$

Hence,

$$\Gamma_4 = \tfrac{1}{2}\left(a^{(4)} - 2a^{(2)} - 2a^{(1)} - \mathrm{tr}\, A^4 - \mathrm{tr}\, A^3 - \sum_{i=1}^{p} a_{ii}^{(3)} a_i^{(1)}\right)$$

$l = 5$. Here, $S = \{\epsilon_1, \epsilon_2, \epsilon_3, \epsilon_6, \epsilon_7, \epsilon_{12}, \epsilon_{14}, \epsilon_{15}, \epsilon_{20}, \epsilon_{24}\}$. Note that $\epsilon_{20}$ and $\epsilon_{23}$ are equivalent since the walks (123143) and (123413) are equivalent. Further,

$$e(\epsilon_1) = e(\epsilon_6) = e(\epsilon_{20}) = 2, \qquad e(\epsilon_2) = e(\epsilon_3) = e(\epsilon_9) = 1$$
$$e(\epsilon_{12}) = e(\epsilon_{14}) = e(\epsilon_{15}) = 2, \qquad e(\epsilon_{24}) = 4.$$

To see the last two equalities, note that (123242) and (124232) are both of shape $\epsilon_{12}$; (123452) and (125432) are both of shape $\epsilon_{14}$; (123453) and (123543) are both of shape $\epsilon_{15}$; (123425), (124325), (524321), and (523421) are all of shape $\epsilon_{24}$.

Hence,

$$\begin{aligned}\Gamma_5 &= \tfrac{1}{2}w(\epsilon_1) + w(\epsilon_2) + w(\epsilon_3) + \tfrac{1}{2}w(\epsilon_6) + w(\epsilon_7) + \tfrac{1}{2}w(\epsilon_{12}) \\ &\quad + \tfrac{1}{2}w(\epsilon_{14}) + \tfrac{1}{2}w(\epsilon_{15}) + \tfrac{1}{2}w(\epsilon_{23}) + \tfrac{1}{4}w(\epsilon_{24}) \\ &= \tfrac{1}{2}[w_5 - 2w(\epsilon_4) - w(\epsilon_5) - 2w(\epsilon_8) - 2w(\epsilon_9) - w(\epsilon_{10}) - 2w(\epsilon_{11}) \\ &\quad - w(\epsilon_{12}) - w(\epsilon_{14}) - w(\epsilon_{15}) - w(\epsilon_{16}) - w(\epsilon_{17}) - 2w(\epsilon_{23}) - \tfrac{1}{2}w(\epsilon_{24}) \\ &\quad - 2w(\epsilon_{25}) - 2w(\epsilon_{26}) - 2w(\epsilon_{27}) - 2w(\epsilon_{28}) - w(\epsilon_{30})]\end{aligned}$$

Now,

$$w(\epsilon_4) = w(\epsilon_5) = w(\gamma_1) = w_3 - 2w_2 - w_1 = a^{(3)} - 2a^{(2)} + a^{(1)}$$

$$w(\epsilon_8) = w(\epsilon_9) = w(\epsilon_{11}) = w(\beta_1) = w_2 = a^{(2)} - a^{(1)}$$

$$w(\epsilon_{10}) = w(\alpha_1) = a^{(1)}$$

$$w(\epsilon_{12}) = 6 \sum_{i=1}^{p} \binom{a_i^{(1)}}{3}$$

$$w(\epsilon_{14}) = 2 \sum_{i \neq j} (a_i^{(1)} - 2)\binom{a_{ij}^{(2)}}{2} - w(\epsilon_{23})$$

$$w(\epsilon_{15}) = \sum_{i=1}^{p} a_{ii}^{(3)}(a_i^{(2)} - 2) - w(\epsilon_{23}) - w(\epsilon_{27})$$

$$w(\epsilon_{16}) = w(\epsilon_{17}) = w(\delta_8) = \operatorname{tr} A^4 - a^{(1)} - 2(a^{(2)} - a^{(1)}) = \operatorname{tr} A^4 - 2a^{(2)} + a^{(1)}$$

$$w(\epsilon_{24}) = 2 \sum_{i=1}^{p} a_{ii}^{(3)}\binom{a_i^{(1)} - 2}{2}$$

$$w(\epsilon_{25}) = w(\epsilon_{26}) = w(\epsilon_{27}) = w(\delta_9) = \sum_{i=1}^{p} a_{ii}^{(3)}(a_i^{(1)} - 2)$$

Hence,

$$\begin{aligned}\Gamma_5 = \tfrac{1}{2}\Big\{ & a^{(5)} - 3a^{(3)} + 6a^{(2)} - a^{(1)} - \operatorname{tr} A^5 - 3 \operatorname{tr} A^4 \\ & - 6 \sum_{i=1}^{p} \binom{a_i^{(1)}}{3} - \sum_{i=1}^{p} a_{ii}^{(3)} \left[a_i^{(2)} + \binom{a_i^{(1)} - 2}{2} + 5a_i^{(1)} - 9\right] \\ & - 2 \sum_{i \neq j} (a_i^{(1)} - 2)\binom{a_{ij}^{(2)}}{2}\Big\}\end{aligned}$$

## REFERENCES

1. A. Cayley, *Phil. Mag.* **13**:19 (1857).
2. W. Brostow, *Phys. Chem. Liq.*, to be published.
3. F. Harary, *Graph Theory*, Addison-Wesley, Reading, Mass. (1969).
4. I. C. Ross and F. Harary, *Psychometrika* **17**:195 (1952).
5. F. Harary and B. Manvel, preprint from the University of Michigan.