

Application of Redfield–Pólya’s Theorem to the enumeration of the substitution isomers of linear polycyclic aromatic hydrocarbons

Enrico Benassi

Received: 21 March 2013 / Accepted: 18 July 2013 / Published online: 28 July 2013
© Springer Science+Business Media New York 2013

Abstract In this brief work, we shall obtain the general *formulae* for the enumeration of the linear polycyclic aromatic hydrocarbons isomers when the hydrogen atom or the CH group is substituted by one or more atoms or different groups (substitution isomers). Such *formulae* have been derived from Redfield–Pólya’s Theorem (Burnside in Theory of groups of finite order. Cambridge University Press, Cambridge, 1897; Redfield in Am J Math 49:433, 1927) application.

Keywords Redfield–Pólya’s Theorem · Polycyclic aromatic hydrocarbons (PAHs) · Substitution isomers

1 Introduction

Redfield–Pólya’s Theorem, also known as Pólya’s enumeration Theorem (PET), is a generalisation of Cauchy–Frobenius–Burnside’s lemma [1] on the number of orbits of a group action on a set. It was first published by Redfield [2], and 10 years later it was independently rediscovered by Pólya [3], who also significantly diffused the result by applying it to many counting problems, in particular to the enumeration of chemical compounds. Redfield–Pólya’s Theorem application to the molecular compounds isomers enumeration is based on the cycle index of a permutation group of a finite set, that of molecular substitutions sites, which are transformed by the symmetry operations of a molecular point group [3–8].

Products derived from coal or petroleum contain polycyclic aromatic hydrocarbons (PAHs) as major components. Average descriptive parameters are usually obtained for such materials; nevertheless, obtaining detailed compositional and structural informa-

E. Benassi (✉)
Centro S3, CNR Istituto di Nanoscienze, Via G. Campi 213/a, 41125 Modena, Italy
e-mail: enrico.benassi@unimore.it

tion can be of crucial importance. Many of the PAH are known to express a remarkable toxicity or mutagenic behaviour in various biological test systems. The toxic or carcinogenic activity can be related to specific structures and positions of ring substitution, as occurs in the cases of the isomers of methylphenanthrene, methylchrysene, benzo[*x*]pyrenes and benzo[*x*]chrysenes [9–13]. The importance of disposing of an univocal, correct and efficient identification of the isomers of PAHs is crucial also in the field of the environmental sciences and green chemistry [14], for which the detailed characterisation of such products is essential to obtain information about the complex structure of fossil fuels such as coal, oil shale, and petroleum. The problem of enumerating naphthalene and anthracene isomers is a classical problem, which has been originally discussed by Pólya himself [15]. In the early 80s, Dias [16] applied the graph theory concepts [17–19] for the enumeration of the PAH isomers obtained by different collections of regular hexagons arranged with adjacent sides. In this paper, we are interest into obtaining the general *formulae* for the enumeration of the substitution isomers of linear PAHs, by using the group theory derivations.

Let $\mathbf{G} = \{g_i\}_{i=1,2,\dots,\Gamma}$ be the molecular point group of order Γ , and let $D = \{d_j\}_{j=1,2,\dots,\Delta}$ the domain given by the set of the Δ substitutional sites of the molecule, where Δ is a sum of divisors of Γ . Moreover, let $P(g_k, D)$ be the set of all the distinct permutations, that forms a permutation group $P(\mathbf{G}, D) = \{P(g_k, D)\}_{k=1,2,\dots,\gamma}$ of order $\gamma \leq \Gamma$. Such an action is a homomorphism from \mathbf{G} to $P(\mathbf{G}, D)$. Since $P(\mathbf{G}, D) \subset S_\Delta$, from Lagrange’s Theorem it follows that γ is a divisor of $\Delta!$. Each permutation $p_k = P(g_k, D)$ may be expressed as a product of disjoint cycles $\prod_{j=1}^{\Delta} x_j^{\varepsilon_{j_p}}$, where ε_{j_p} indicates the number of cycles of degree j in the permutation $p \in P(\mathbf{G}, D)$, and thus the cycle index of the permutation group $P(\mathbf{G}, D)$ may be written as the following arithmetical mean:

$$Z(P(\mathbf{G}, D)) = \frac{1}{\gamma} \sum_{p \in P(\mathbf{G}, D)} \prod_{j=1}^{\Delta} x_j^{\varepsilon_{j_p}} = \frac{1}{\gamma} \sum_{l=1}^r h_l \prod_{j=1}^{\Delta} x_j^{\varepsilon_{j_l}}, \tag{1.1}$$

where the first sum runs over all the γ permutations, and the second sum runs over all the r permutation classes of $P(\mathbf{G}, D)$, h_l is the number of permutations of l th class ($\sum_{l=1}^r h_l = \gamma$), and the exponent ε_{j_l} is the number of cycles of degree j in the permutations of the l th class ($\sum_{j=1}^{\Delta} j \varepsilon_{j_l} = \Delta$). Exploiting some previous results [20], Baraldi et al. [8] obtained the *formula*:

$$Z(P(\mathbf{D}_n, D)) = \frac{x_1^{n(C_n)}}{2n} \sum_{k|n} \varphi(k) x_k^{(\Delta - n(C_n))/k} + \begin{cases} \frac{1}{2} x_1^{n(C_2)} x_2^{(\Delta - n(C_2))/2} & (\text{nodd}) \\ \frac{1}{4} \left[x_1^{n(C_2)} x_2^{(\Delta - n(C_2))/2} + x_1^{n(C'_2)} x_2^{\Delta - n(C'_2)/2} \right] & (\text{neven}) \end{cases}, \tag{1.2}$$

where $\varphi(k)$ is Euler’s function, which determines the number of positive integers $\leq k$ which are coprime to k .

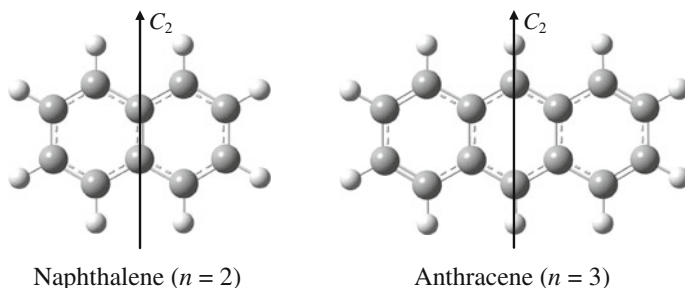


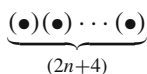
Fig. 1 Examples of even- (*left*) and odd- n (*right*) PAHs. One of the three C_2 axes is pictured

2 How to obtain the general formulae for the enumeration of the linear PAHs substitution isomers

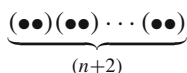
The general molecular formula for PAHs is $C_{4n+2}H_{2n+4}$, where n is the number of fused benzene rings. Their point group is $D_{2h} = \{E, 3C_2 i, 3\sigma_v\}$, but for the isomers enumeration considering the sub-group of the proper rotations $D_2 = \{E, 3C_2\}$ is enough [5]. The sites available for the substitution that lead to the isomers formation are the positions occupied by the hydrogen atoms in these hydrocarbons. The odd- n PAHs (e.g., the anthracene, $n = 3$) have two sites over a C_2 axis, while the even- n PAHs (e.g., naphthalene, $n = 2$) have no substitution site over the C_2 axes (Fig. 1). This origins a different cyclic structure of the $2n + 4$ sites permutations in the two PAHs classes.

To apply Pólya's Theorem, it is needed to identify the cyclic structure of the permutation classes corresponding to the permutations of the active sites set connected to the symmetry operation of the point group. To each permutation class, a term of the cycles index is associated. If we symbolise a m -length cycle with $\underbrace{(\bullet \bullet \bullet \dots \bullet)}_{(m)}$, for

the even- n PAHs, the permutation classes are two, and they have the following cyclic structure:



for the identity permutation, and:



for the permutations corresponding to the C_2 operations. Therefore, the associated cycles index is:

$$Z(D_2, x_1, x_2) = \frac{1}{4} (x_1^{2n+4} + 3x_2^{n+2}), \quad (\text{even } n) \quad (2.1)$$

where x_1 and x_2 indicate the fictitious (or undetermined) variable associated to the cycle (\bullet) and $(\bullet\bullet)$, respectively. (In general, x_m indicates the fictitious (or undetermined) variable associated to the m -length cycle.) The power of these variables coincides with the number of same kind cycles that appear in the permutation, and the number that multiply them is the number of permutations of that class. In the case of odd- n PAHs, the cyclic structure of the three permutations corresponding to the C_2 operations is not the same anymore. Two of them show the previously analysed shape, and one (that corresponding to the two sites over a C_2 axis) has the following structure:

$$(\bullet)(\bullet) \underbrace{(\bullet\bullet)(\bullet\bullet) \cdots (\bullet\bullet)}_{(n+1)},$$

that gives rise to an additional class. Therefore, the cycles index for the odd- n PAHs is:

$$Z(\mathbf{D}_2, x_1, x_2) = \frac{1}{4} \left(x_1^{2n+4} + x_1^2 x_2^{n+1} + 2x_2^{n+2} \right). \quad (\text{odd } n) \tag{2.2}$$

For the examples reported in Fig. 1, the formulae (2.1) and (2.2) becomes:

$$Z(\mathbf{D}_2, x_1, x_2) = \frac{1}{4} \left(x_1^8 + 3x_2^4 \right), \quad (\text{naphthalene, } n = 2) \tag{2.3}$$

$$Z(\mathbf{D}_2, x_1, x_2) = \frac{1}{4} \left(x_1^{10} + x_1^2 x_2^4 + 2x_2^5 \right), \quad (\text{anthracene, } n = 3) \tag{2.4}$$

that are known results [2].

Now, if we want to enumerate the isomers that are obtained when one site may be occupied by p atoms or groups different from each other, the following substitution is required:

$$x_i = \underbrace{r^i + s^i + t^i + \cdots + z^i}_{(p)}, \tag{2.5}$$

where the right member is called figures counting series. Inserting (2.5) into (2.1) and (2.2), we obtain the counting polynomial for the PAHs substitution isomers, i.e.:

$$\begin{aligned} & Z \left(\mathbf{D}_2, (r + s + t + \cdots + z), (r^2 + s^2 + t^2 + \cdots + z^2) \right) \\ &= \frac{1}{4} \left[(r + s + t + \cdots + z)^{2n+4} + 3(r^2 + s^2 + t^2 + \cdots + z^2)^{n+2} \right], \quad (\text{even } n) \\ Z^{(e)} &= \frac{1}{4} \left[\left(\sum_{j=1}^p \alpha_j \right)^{2n+4} + 3 \left(\sum_{j=1}^p \alpha_j^2 \right)^{n+2} \right] \tag{2.6} \\ & Z \left(\mathbf{D}_2, (r + s + t + \cdots + z), (r^2 + s^2 + t^2 + \cdots + z^2) \right) \end{aligned}$$

$$\begin{aligned}
&= \frac{1}{4} \left[(r+s+t+\dots+z)^{2n+4} + (r+s+t+\dots+z)^2 (r^2+s^2+t^2+\dots+z^2)^{n+1} \right. \\
&\quad \left. + 2(r^2+s^2+t^2+\dots+z^2)^{n+2} \right] \\
Z^{(o)} &= Z^{(e)} + \frac{1}{2} \left(\sum_{j=1}^p \alpha_j^2 \right)^{n+1} \sum_{j=1}^p \alpha_j \alpha_{k>j} \quad (\text{odd } n) \quad (2.7)
\end{aligned}$$

To count the total number of isomers, connected to the substitution (2.5), it is needed to put $x_i = p$ into (2.1) and (2.2). So we have:

$$Z(\mathbf{D}_2, p, p) = \frac{1}{4} (p^{2n+4} + 3p^{n+2}) = \frac{p^{n+2}}{4} (p^{n+2} + 3), \quad (\text{even } n) \quad (2.8)$$

$$Z(\mathbf{D}_2, p, p) = \frac{1}{4} (p^{2n+4} + p^{n+3} + 2p^{n+2}) = \frac{p^{n+2}}{4} (p^{n+2} + p + 2). \quad (\text{odd } n) \quad (2.9)$$

Table 1 collects the total number and the counting polynomials of the isomers, $p = 2$ for and $n = 2-7$. The counting polynomial of the isomers may be interpreted as follows. In case one site may be occupied by a H atom (r) or another X atom (s) in naphthalene, we have a single isomer $C_{10}H_8(r^8)$ and $C_{10}X_8(s^8)$, 2 isomers $C_{10}H^7X(r^7s)$ and $C_{10}HX_7(r^7s^7)$, 10 isomers $C_{10}H_6X_2(r^6s^2)$ and $C_{10}H_2X_6(r^2s^6)$, 14 isomers $C_{10}H_5X_3(r^5s^3)$ and $C_{10}H_3X_5(r^3s^5)$, and 22 isomers $C_{10}H_4X_4(r^4s^4)$. The other counting polynomials may be analogously interpreted. The total number of isomers coincides with the sum of the coefficients of the counting polynomial terms. The results concerning the cases of $n = 2$ (naphthalene) and 3 (anthracene) are equivalent to those reported by Pólya [15] although he does not provide the number of isomers when the sites are occupied by more than $4 \times$ atoms in the case of anthracene.

As an example of triple substitution ($x_i = r^i + s^i + t^i$), we shall consider naphthalene. The Eq. (2.6) reduces to the following polynomial:

$$\begin{aligned}
&Z(\mathbf{D}_2, (r+s+t), (r^2+s^2+t^2)) \\
&= (r^8 + s^8 + t^8) + 2(r^7s + r^7t + st^7 + rs^7 + s^7t + rt^7) \\
&\quad + 10(r^6s^2 + r^2s^6 + r^6t^2 + s^6t^2 + r^2s^6 + s^2t^6) \\
&\quad + 14(r^5s^3 + r^3s^5 + r^5t^3 + s^5t^3 + r^3s^5 + s^3t^5) \\
&\quad + 22(r^4s^4 + r^4t^4 + s^4t^4) + 14(r^6st + rs^6t + rst^6) \\
&\quad + 42(r^5s^2t + r^2s^5t + r^5st^2 + rs^5t^2 + r^2st^5 + rs^2t^5) \\
&\quad + 70(r^4s^3t + r^3s^4t + r^4st^3 + rs^4t^3 + r^3st^4 + rs^3t^4) \\
&\quad + 114(r^4s^2t^2 + r^2s^4t^2 + r^2s^2t^4) \\
&\quad + 140(r^3s^3t^2 + r^3s^2t^3 + r^2s^3t^3). \quad (2.10)
\end{aligned}$$

The total number of the isomer is 1,701, coinciding with the sum of the coefficients of (2.10). Also in this case, our results generalise those found by Pólya [15].

Table 1 Total number of isomers and counting polynomial of the isomers for the first six linear PAHs, when the hydrogen atoms substituted with another atom or group

Molecules ^a	Total number of isomers	Counting polynomial of the isomers
Naphthalene ($n = 2$)	76	$(r^8 + s^8) + 2(r^7s + rs^7) + 10(r^6s^2 + r^2s^6) + 14(r^5s^3 + r^3s^5) + 22r^4s^4$
Anthracene ($n = 3$)	288	$(r^{10} + s^{10}) + 3(r^9s + rs^9) + 15(r^8s^2 + r^2s^8) + 32(r^7s^3 + r^3s^7) + 60(r^6s^4 + r^4s^6) + 66r^5s^5$
Tetracene ($n = 4$)	1,072	$(r^{12} + s^{12}) + 3(r^{11}s + rs^{11}) + 21(r^{10}s^2 + r^2s^{10}) + 55(r^9s^3 + r^3s^9) + 135(r^8s^4 + r^4s^8) + 198(r^7s^5 + r^5s^7) + 246r^6s^6$
Pentacene ($n = 5$)	4,224	$(r^{14} + s^{14}) + 4(r^{13}s + rs^{13}) + 28(r^{12}s^2 + r^2s^{12}) + 94(r^{11}s^3 + r^3s^{11}) + 266(r^{10}s^4 + r^4s^{10}) + 508(r^9s^5 + r^5s^9) + 777(r^8s^6 + r^6s^8) + 868r^7s^7$
Hexacene ($n = 6$)	16,576	$(r^{16} + s^{16}) + 4(r^{15}s + rs^{15}) + 36(r^{14}s^2 + r^2s^{14}) + 140(r^{13}s^3 + r^3s^{13}) + 476(r^{12}s^4 + r^4s^{12}) + 1,092(r^{11}s^5 + r^5s^{11}) + 2,044(r^{10}s^6 + r^6s^{10}) + 2,860(r^9s^7 + r^7s^9) + 3,270r^8s^8$
Heptacene ($n = 7$)	66,048	$(r^{18} + s^{18}) + 5(r^{17}s + rs^{17}) + 45(r^{16}s^2 + r^2s^{16}) + 208(r^{15}s^3 + r^3s^{15}) + 792(r^{14}s^4 + r^4s^{14}) + 2,156(r^{13}s^5 + r^5s^{13}) + 4,074(r^{12}s^6 + r^6s^{12}) + 7,984(r^{11}s^7 + r^7s^{11}) + 11,034(r^{10}s^8 + r^8s^{10}) + 12,190r^9s^9$

^a The last members are very reactive. The superior acenes are not stable

3 Conclusions

In this work, we have obtained the general and simple *formulae* for the enumeration of the linear PAHs isomers when the Hydrogen atom or the CH group is substituted by one or more atoms or different groups (substitution isomers). Such *formulae* have been obtained through Redfield–Pólya's Theorem application, considering the molecular symmetry. The general molecular *formula* for PAHs is $C_{4n+2}H_{2n+4}$, where n is the number of fused benzene rings, and their point group is $D_{2h} = \{E, 3C_2, i, 3\sigma_v\}$, but for the isomers enumeration considering the sub-group of the proper rotations $D_2 = \{E, 3C_2\}$ suffices. Since the odd- n PAHs have two sites over a C_2 axis, while the even- n PAHs have no substitution site over the C_2 axes, a different cyclic structure of the $2n + 4$ sites permutations in the two PAHs classes is originated. The total number of non-equivalent substitution isomers is 76, 288, 1,072, 4,224, 16,576, and 6,648 for naphthalene, anthracene, tetracene, pentacene, hexacene, and heptacene respectively. The results obtained for naphthalene and anthracene generalise those achieved by Pólya [15].

References

1. W. Burnside, *Theory of Groups of Finite Order* (Cambridge University Press, Cambridge, 1897)
2. J.H. Redfield, *Am. J. Math.* **49**, 433 (1927)
3. G. Pólya, *Acta Math.* **68**, 145 (1937)
4. G. Pólya, R.C. Read, *Combinatorial Enumeration of Groups, Graphs, and Chemical Compounds* (Springer, Berlin, 1987)
5. A.T. Balaban, in *Chemical Group Theory. Introduction and Fundamentals*, chap. 5, ed. by D. Bonchev, D.H. Rouvray (Gordon and Breach Science, Yverdon, 1994)
6. K. Balasubramanian, in *Chemical Group Theory. Techniques and Applications*, chap. 2, ed. by D. Bonchev, D.H. Rouvray (Gordon and Breach Science, Amsterdam, 1995)
7. S. Fujita, *Symmetry and Combinatorial Enumeration in Chemistry* (Springer, Berlin, 1991)
8. I. Baraldi, C. Fiori, D. Vanossi, *J. Math. Chem.* **25**, 23 (1999)
9. W.C. Hueper, W.D. Conway, *Chemical Carcinogenesis and Cancers* (Charles C. Thomas, Springfield, 1964)
10. S.S. Hecht, W.E. Bondinell, D.J. Hoffmann, *Natl. Cancer Inst.* **52**, 1121 (1974)
11. E.J. Lavoie, L. Tulley-Freiler, V. Bedenko, D. Hoffmann, *Cancer Res.* **41**, 3441 (1981)
12. M.L. Lee, M. Novotny, K.D. Bartle, *Analytical Chemistry of Polycyclic Aromatic Compounds* (Academic, New York, 1981)
13. J.G. Speight, *The Chemistry and Technology of Coal* (Marcel Dekker, New York, 1983)
14. M. Nishioka, *Jpn. Patent 58-49788*, Tokkaisho (1983)
15. G. Pólya, *Helv. Chim. Acta* **19**, 22 (1936)
16. J.R. Dias, *J. Chem. Inf. Comput. Sci.* **22**, 15 (1982)
17. A.T. Balaban, *Rev. Roum. Chim.* **11**, 1097 (1966)
18. A.T. Balaban, F. Harary, *Tetrahedron* **24**, 2505 (1968)
19. A.T. Balaban, *Chemical Applications of Graph Theory*, chap. 5 (Academic Press, London, 1976)
20. N.L. Biggs, *Discrete Mathematics* (Oxford Science, Oxford, 1989)