

Regular article

Graphs to chemical structures 2. Extended sphericity indices of cycles for stereochemical extension of Pólya's coronas

Shinsaku Fujita

Department of Chemistry and Materials Technology, Kyoto Institute of Technology, Matsugasaki, Sakyo-ku, Kyoto 606-8585, Japan

Received: 13 April 2004 / Accepted: 26 July 2004 / Published online: 26 November 2004
© Springer-Verlag 2004

Abstract In order to enumerate nonrigid isomers, we have proposed the proligand approach, where extended sphericity indices of k -cycles have been defined according to the enantiospheric, homospheric, or hemispheric nature of each k -cycle. Then, cycle indices with chirality fittingness have been defined so as to enumerate nonrigid stereoisomers with chiral and achiral ligands. Results of the proligand approach using tetramethylmethane as an example have been compared with those based on Pólya's corona. Thereby, Pólya's corona is concluded to be concerned with graphs, but not with chemical structures, where it is incapable of treating chiral ligands properly.

Keywords: Cycle index—Chirality fittingness—Sphericity index—Pólya's corona

1 Introduction

In a previous work [1], we pointed out that Pólya's theorem [2,3] is insufficient to deal with chemical structures (or stereoisomers) in which both achiral and chiral ligands are taken into consideration. In other words, Pólya's theorem is concerned with graphs, but not with chemical structures, where it treats atoms only (or at most achiral ligands) and incapable of treating chiral ligands properly. In fact, many graph-theoretical studies based on Pólya's theorem have been conducted within this limitation, where each object occupying a vertex (node) of a graph has been implicitly considered to be structureless, as implicated in books [4,5,6] and reviews [7,8,9].

The insufficiency of Pólya's theorem, in particular, has become obvious in its application to tetrahedral

molecules having the formula $CABp\bar{p}$, in which A and B are achiral ligands and p and \bar{p} represent a pair of enantiomeric chiral ligands [10]. The insufficiency, however, may be overlooked unconsciously or neglected as a trivial thing, unless the abstract meaning of A, B, p, and \bar{p} is formulated in terms of *proligands* and *promolecules*, where such proligands are regarded as being structureless but have achirality/chirality [11]. This situation has been critically discussed in terms of “conceptual revolution from graphs to chemical structures” [10].

After introducing the concepts of proligand and promolecule, we can see that the insufficiency of Pólya's theorem becomes clearer when such proligands are replaced by actual ligands so as to convert a promolecule into a nonrigid molecule. Let us consider $CABp\bar{p}$ in which each of A, B, p, and \bar{p} is derived from a methyl ligand (a methyl group). By placing $A = CX_2Y$, $B = CXY_2$, $p = R-CXYZ$, and $\bar{p} = S-CXYZ$, we obtain two achiral molecules (**1** and **2**), which are diastereomeric to each other, as shown in the top row of Fig. 1.

On the other hand, graph-theoretical approaches are incapable of discriminating between the two molecules (**1** and **2**) so as to recognize only one molecule (**3**) as a graph. In fact, Pólya's corona that is an extension of Pólya's theorem for treating such nonrigid isomers as **3** does not discriminate between $R-CXYZ$ and $S-CXYZ$ during its enumeration process [2,3]. This means that Pólya's corona lacks the concept of proligands (especially chiral proligands such as p and \bar{p}).

In this paper, we will first review Pólya's corona to show the previously described insufficiency by using an enumeration problem in which the four methyl ligands of tetramethylmethane [2,2-dimethylpropane, $C(CH_3)_4$] are replaced by a set of substituted methyl ligands. Then, we will propose a new approach based on the concepts of proligand and promolecule (called here the *proligand approach*), where both achiral and chiral (pro)ligands are taken into consideration. Thereby, we will show the superiority of our approach over Pólya's corona by using the same enumeration problem as an example.

Correspondence to: S. Fujita
e-mail: fujitas@chem.kit.ac.jp

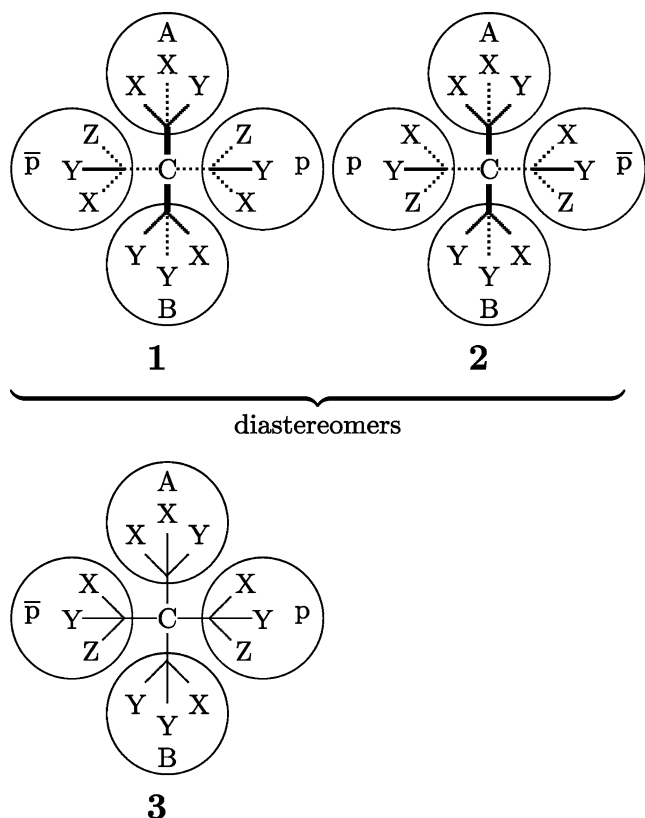


Fig. 1 Two molecules for chemical structure (1 and 2) versus one molecule for graph (3). The symbols X , Y , and Z represent appropriate atoms. Each encircled segment in **1** or **2** represents a substituted methyl ligand, which corresponds to an achiral ligand (A or B) or a chiral ligand (p or \bar{p}). Each encircled segment in **3** represents a substituted methyl ligand as a graph.

2 Pólya's corona

Pólya's corona was introduced in his famous article [2,3] to solve enumeration problems of nonrigid molecules due to rotations around bonds. This is essentially equivalent to the concept of wreath products, which was later used to count stereoisomers and positional isomers in a generalized wreath product method [12]. These methods and related topics have been reviewed extensively in books [3,5] and reviews [7,9,13,14]. The present section is devoted to discussion on the scope and limitations of Pólya's corona by using derivatives of tetramethylmethane as an example.

First we find the extended cycle index (CI) of a methyl ligand. Since the three hydrogens of a methyl ligand are controlled by a permutation group represented by

$$\{(1)(2)(3), (1\ 3\ 2), (1\ 2\ 3), (1\ 2)(3), (1\ 3)(2), (1)(2\ 3)\}, \quad (1)$$

we obtain the corresponding CI:

$$\psi_1 = \frac{1}{6}s_1^3 + \frac{1}{3}s_3 + \frac{1}{2}s_1s_2, \quad (2)$$

To extend this CI to the case of tetramethylmethane as a corona, we introduce the following corona ψ_k on the basis of Eq. (2):

$$\psi_k = \frac{1}{6}s_k^3 + \frac{1}{3}s_{3k} + \frac{1}{2}s_k s_{2k}, \quad (3)$$

where k represents a multiplicity factor (1–4) for calculating the CI $\varphi(\psi)$ of tetramethylmethane:

$$\varphi(\psi_k) = \frac{1}{24}\psi_1^4 + \frac{1}{8}\psi_2^2 + \frac{1}{4}\psi_1^2\psi_2 + \frac{1}{3}\psi_1\psi_3 + \frac{1}{4}\psi_4. \quad (4)$$

Thus, four methyls, which are the same or different according to the values of k in Eq. (3), are substituted for the four positions of methane. Note that the four positions of methane are controlled by the symmetric group of degree 4 so as to give Eq. (4). When a set of ligand is selected from $\{X, Y, Z\}$, the corresponding ligand inventory is obtained as follows:

$$s_d = X^d + Y^d + Z^d. \quad (5)$$

After introducing Eq (3) into Eq. (4), the ligands inventory (Eq. 5) is introduced into the resulting polynomial concerning s_d so as to give a generating function (f), as shown in Eq. (6).

$$\begin{aligned} f = & (X^{12} + Y^{12} + Z^{12}) \\ & + (X^{11}Y + XY^{11} + X^{11}Z + \dots) \\ & + 2(X^{10}Y^2 + X^2Y^{10} + X^{10}Z^2 + \dots) \\ & + 3(X^9Y^3 + X^3Y^9 + X^3Z^9 + \dots) \\ & + 4(X^8Y^4 + X^4Y^8 + X^8Z^4 + \dots) \\ & + 4(X^7Y^5 + X^5Y^7 + X^7Z^5 + \dots) \\ & + 5(X^6Y^6 + X^6Z^6 + Y^6Z^6) \\ & + 2(X^{10}YZ + XYZ^{10} + XY^{10}Z) \\ & + 4(X^9Y^2Z + X^2Y^9Z + XY^9Z^2 + \dots) \\ & + 6(X^8Y^3Z + X^3Y^8Z + X^8YZ^3 + \dots) \\ & + 8(X^8Y^2Z^2 + X^2Y^8Z^2 + X^2Y^2Z^8) \\ & + 8(X^7Y^4Z + X^4Y^7Z + XY^7Z^4 + \dots) \\ & + 12(X^7Y^3Z^2 + X^7Y^2Z^3 + X^3Y^7Z^2 + \dots) \\ & + 9(X^6Y^5Z + X^5Y^6Z + XY^6Z^5 + \dots) \\ & + 16(X^6Y^4Z^2 + X^4Y^6Z^2 + X^6Y^2Z^4 + \dots) \\ & + 18(X^6Y^3Z^3 + X^3Y^6Z^3 + X^3Y^3Z^6) \\ & + 16(X^5Y^5Z^2 + X^2Y^5Z^5 + X^5Y^2Z^5) \\ & + 21(X^5Y^4Z^3 + X^4Y^5Z^3 + X^5Y^3Z^4 + \dots) \\ & + 25X^4Y^4Z^4 \end{aligned} \quad (6)$$

The coefficient of each term $X^xY^yZ^z$ ($x + y + z = 12$) in Eq. (6) represents the number of isomers (as graphs) with x of X , y of Y , and z of Z . It should be noted that the coefficients collected in Eq. (6) were obtained by using Pólya's corona.

To show the meaning of the intermediate equation (Eq. 2), the ligand inventory (Eq. 5) is introduced into Eq. (2) so as to produce

$$\begin{aligned} \psi_1 = & X^3 + X^2Y + XY^2 + Y^3 + X^2Z + XYZ \\ & + Y^2Z + XZ^2 + YZ^2 + Z^3. \end{aligned} \quad (7)$$

Each term on the right-hand side of Eq. (7) shows the type of ligand to be considered. Moreover, its coefficient represents its number to be calculated, although each coefficient is equal to 1 in this case. In particular, the term XYZ in Eq. (7) corresponds to a chiral ligand $CXYZ$, where the value 1 of the coefficient of the term XYZ shows that $R-CXYZ$ and $S-CXYZ$ are equalized as one graph.

3 Proligand approach

The present approach is based on the concepts of proligand and promolecule [11], where the replacement of proligands by usual ligands converts a promolecule into a molecule. For the sake of convenience, the present approach is called the proligand approach.

3.1 Enumeration of chiral and achiral Ligands

A ligand for substituting for a proligand can be derived from an appropriate ligand skeleton which has n substitution positions. Let \mathbf{H} be the point group of a ligand skeleton. Then the substitution positions are governed by a sum of coset representations, i.e., $\sum_i \beta_i \mathbf{H}/(\mathbf{H}_i)$, where \mathbf{H}_i represents a subgroup of \mathbf{H} and β_i represents the multiplicity of each coset representation. The sum can be regarded as a permutation group $\hat{\mathbf{H}}$ of degree n . In this paper, we will mainly consider such faithful cases. But the discussions in this paper hold true for unfaithful cases in which the order of $\hat{\mathbf{H}}$ is not always equal to the order of \mathbf{H} .

An element \hat{H} of $\hat{\mathbf{H}}$ is called a *proper element* if it corresponds to a proper rotation of \mathbf{H} , and is called an *improper element* if it corresponds to an improper rotation of \mathbf{H} . Such an improper element is represented by a product of cycles with overlines, which emphasize the change of the chirality of each ligand to the opposite chirality.

Let the cycle decomposition of an improper element \hat{H} of $\hat{\mathbf{H}}$ be represented by the number $v_k(\hat{H})$ of k -cycles, where we place $k = 1, 2, \dots, n$. If k is even, a k -cycle for an improper element is called an *enantiospheric k -cycle* and is ascribed to a sphericity index c_k . If k is odd, a k -cycle for an improper element is called a *homospheric k -cycle* and is ascribed to a sphericity index a_k . Whether k is even or odd, a k -cycle for a proper element is called a *hemispheric k -cycle* and is ascribed to a sphericity index b_k . Thereby, the sphericity indices a_k , c_k , and b_k control modes of ligand packing, which are equivalent to those due to chirality fittingness proposed (CF) previously [1,15]. Thereby, we define a CI with (CI-CF) by using the sphericity indices a_k , c_k , and b_k .

Let $\hat{\mathbf{H}}$ be a sum of coset representations of a point group \mathbf{H} , as described earlier. An element \hat{H} of $\hat{\mathbf{H}}$ is a permutation of degree n , which is represented by a cycle decomposition involving the number $v_k(\hat{H})$ of k -cycles [$\sum_{k=1}^n kv_k(\hat{H}) = n$]. Each of the k -cycles corresponds to a sphericity index $\$k$, where $\$k$ is a_k if \hat{H} is an improper

element and k is odd, $\$k$ is c_k if \hat{H} is an improper element and k is even, and $\$k$ is b_k if \hat{H} is a proper element. Hence, the element \hat{H} corresponds to a product of sphericity indices $\$1^{v_1(\hat{H})} \$2^{v_2(\hat{H})} \dots \$n^{v_n(\hat{H})}$. Thereby, a CI-CF for the present case is defined as follows:

$$\text{CI-CF}(\hat{H}; \$d) = \frac{1}{|\hat{\mathbf{H}}|} \sum_{\hat{H} \in \hat{\mathbf{H}}} \$1^{v_1(\hat{H})} \$2^{v_2(\hat{H})} \dots \$n^{v_n(\hat{H})}, \quad (8)$$

where $\$d$ is a_d if \hat{H} is an improper element and d is odd, $\$d$ is c_d if \hat{H} is an improper element and d is even, and $\$d$ is b_d if \hat{H} is a proper element.

Let $\hat{\mathbf{H}}$ be a sum of coset representations of the maximum chiral subgroup $\hat{\mathbf{H}}$ of \mathbf{H} . An element \hat{H} of $\hat{\mathbf{H}}$ is a permutation of degree n , which is also an element of $\hat{\mathbf{H}}$.

$$\text{CI-CF}(\hat{\mathbf{H}}'; b_d) = \frac{1}{|\hat{\mathbf{H}}'|} \sum_{\hat{H} \in \hat{\mathbf{H}}'} b_1^{v_1(\hat{H})} b_2^{v_2(\hat{H})} \dots b_n^{v_n(\hat{H})}, \quad (9)$$

The following theorem for enumerating ligands has essentially the same meaning as described for the enumeration of isomers [10].

Theorem 1. Let $\hat{\mathbf{H}}$ be a sum of coset representations of a point group, which governs a set Δ of n positions. Suppose that the cycle structure of $\hat{H}(\in \hat{\mathbf{H}})$ is represented by

$$(1^{v_1(\hat{H})} 2^{v_2(\hat{H})} \dots n^{v_n(\hat{H})}),$$

where $\sum_{k=1}^n kv_k(\hat{H}) = n$. Each position of Δ is occupied by an achiral or a chiral ligand selected from a set of ligands,

$$\mathbf{X} = \{x_1, x_2, \dots, x_m; p_1, p_2, \dots, p_{m'}; \bar{p}_1, \bar{p}_2, \dots, \bar{p}_{m'}\},$$

where each x_j represents an achiral ligand and each pair of p_j and \bar{p}_j represents an enantiomeric pair of chiral ligands. Consider isomers having θ_1 of x_1 , θ_2 of x_2 , ..., θ_m of x_m ; θ'_1 of p_1 , θ'_2 of p_2 , ..., $\theta'_{m'}$ of $p_{m'}$; and θ''_1 of \bar{p}_1 , θ''_2 of \bar{p}_2 , ..., $\theta''_{m'}$ of $\bar{p}_{m'}$, where

$$\begin{aligned} [\theta] &: \theta_1 + \theta_2 + \dots + \theta_m \\ &+ \theta'_1 + \theta'_2 + \dots + \theta'_{m'} \\ &+ \theta''_1 + \theta''_2 + \dots + \theta''_{m'} = n. \end{aligned} \quad (10)$$

Let the symbol B_0 denote the number of isomers of such isomers as having $[\theta]$ under the action of $\hat{\mathbf{H}}$, where achiral isomers and enantiomeric pairs are enumerated combinatorially. Let the symbol B'_0 denote the number of isomers of such isomers as having $[\theta]$ under the action of $\hat{\mathbf{H}}$, where achiral isomers and chiral isomers (both of each enantiomer) are enumerated combinatorially. Generating functions for calculating B_0 and B'_0 are represented by

$$\begin{aligned} \sum_{[\theta]} B_0 x_1^{\theta_1} x_2^{\theta_2} \dots x_m^{\theta_m} p_1^{\theta'_1} p_2^{\theta'_2} \dots p_{m'}^{\theta'_{m'}} \bar{p}_1^{\theta''_1} \bar{p}_2^{\theta''_2} \dots \bar{p}_{m'}^{\theta''_{m'}} \\ = \text{CI-CF}(\hat{\mathbf{H}}; \$d), \end{aligned} \quad (11)$$

$$\begin{aligned} \sum_{[\theta]} B'_0 x_1^{\theta_1} x_2^{\theta_2} \dots x_m^{\theta_m} p_1^{\theta'_1} p_2^{\theta'_2} \dots p_{m'}^{\theta'_{m'}} \bar{p}_1^{\theta''_1} \bar{p}_2^{\theta''_2} \dots \bar{p}_{m'}^{\theta''_{m'}} \\ = \text{CI-CF}(\hat{\mathbf{H}}'; b_d), \end{aligned} \quad (12)$$

where the summation is concerned with all of the partitions ($[\theta]$) shown in Eq. (10). The sphericity indices $\$_d$ in the CI-CF are ligand inventories replaced by

$$a_d = x_1^d + x_2^d + \cdots + x_n^d \quad (13)$$

$$c_d = x_1^d + x_2^d + \cdots + x_n^d + 2p_1^{d/2} \bar{p}_1^{d/2} + 2p_2^{d/2} \bar{p}_2^{d/2} + \cdots + 2p_n^{d/2} \bar{p}_n^{d/2}, \quad (14)$$

$$b_d = x_1^d + x_2^d + \cdots + x_n^d + p_1^d + p_2^d + \cdots + p_n^d + \bar{p}_1^d + \bar{p}_2^d + \cdots + \bar{p}_n^d. \quad (15)$$

For the present purpose of ligand enumeration, we should place $p_i = \bar{p}_i$ ($i = 1, 2, \dots, n$), because chiral and achiral ligands should be differentiated but Eqs. (8) and (9) are considered on the common basis. Thereby, Eqs. (13), (14) and (15) are simplified as follows:

$$a_d = x_1^d + x_2^d + \cdots + x_n^d, \quad (16)$$

$$c_d = x_1^d + x_2^d + \cdots + x_n^d + 2(p_1^d + p_2^d + \cdots + p_n^d), \quad (17)$$

$$b_d = x_1^d + x_2^d + \cdots + x_n^d + 2(p_1^d + p_2^d + \cdots + p_n^d). \quad (18)$$

Since Eq. (11) gives the number of achiral ligands (A) plus enantiomeric pairs (C), we can place $A + C = \text{CI-CF}(\hat{\mathbf{H}}; \$_d)$. Since Eq. (12) gives the number of achiral ligands (A) plus chiral ligands (C) plus their enantiomers (C), we can place $A + 2C = \text{CI-CF}(\hat{\mathbf{H}}'; b_d)$ under the assumption given by Eqs. (17) and (18). Then we have $C = \text{CI-CF}(\hat{\mathbf{H}}'; \$_d) - \text{CI-CF}(\hat{\mathbf{H}}; b_d)$ and $A = 2\text{CI-CF}(\hat{\mathbf{H}}; b_d) - \text{CI-CF}(\hat{\mathbf{H}}'; b_d)$. When a ligand selected from the set of such chiral ligands and such achiral ligands is placed on a position of a given skeleton of \mathbf{G} -symmetry, we should place $\psi_{(a)1} = A$, $\psi_{(c)1} = A + 2C$, and $\psi_{(b)1} = A + 2C$. Hence, we obtain the following equations:

$$\psi_{(a)1} = 2\text{CI-CF}(\hat{\mathbf{H}}; \$_d) - \text{CI-CF}(\hat{\mathbf{H}}'; b_d), \quad (19)$$

$$\psi_{(c)1} = \text{CI-CF}(\hat{\mathbf{H}}'; b_d), \quad (20)$$

$$\psi_{(b)1} = \text{CI-CF}(\hat{\mathbf{H}}'; b_d), \quad (21)$$

where $\psi_{(a)1}$ corresponds to homosphericity, $\psi_{(c)1}$ corresponds to enantiosphericity, and $\psi_{(b)1}$ corresponds to hemisphericity.

It should be noted that $\psi_{(a)1}$, $\psi_{(c)1}$, and $\psi_{(b)1}$ produce respective generating functions for generating ligand inventories by introducing Eqs. (16), (17) and (18). As an example, let us again examine derivatives of tetramethylmethane.

$$\hat{\mathbf{H}} = \mathbf{C}_{3v}/(\mathbf{C}_s) = \{(1)(2)(3), (1\ 3\ 2), (1\ 2\ 3), \overline{(1\ 2)(3)}, \overline{(1\ 3)(2)}, \overline{(1)(2\ 3)}\},$$

$$\hat{\mathbf{H}}' = \mathbf{C}_3/(\mathbf{C}_1) = \{(1)(2)(3), (1\ 3\ 2), (1\ 2\ 3)\},$$

where an overbar represents an improper element. According to Eqs. (8) and (9), we have the corresponding CIs-CF:

$$\text{CI-CF}(\hat{\mathbf{H}}; \$_d) = \frac{1}{6}b_1^3 + \frac{1}{3}b_3 + \frac{1}{2}a_1c_2, \quad (22)$$

$$\text{CI-CF}(\hat{\mathbf{H}}'; b_d) = \frac{1}{3}b_1^3 + \frac{2}{3}b_3. \quad (23)$$

These equations are introduced into Eqs. (19), (20), and (21) so as to give

$$\psi_{(a)1} = a_1c_2, \quad (24)$$

$$\psi_{(c)1} = \frac{1}{3}b_1^3 + \frac{2}{3}b_3, \quad (25)$$

and

$$\psi_{(b)1} = \frac{1}{3}b_1^3 + \frac{2}{3}b_3. \quad (26)$$

By considering achiral ligands only (X , Y , and Z) Eqs. (16), (17), and (18) generate the same ligand inventories:

$$a_d = c_d = b_d = X^d + Y^d + Z^d \quad (27)$$

By introducing Eq. (27) into Eqs. (24), (25), and (26), we obtain

$$\psi_{(a)1} = X^3 + X^2Y + XY^2 + Y^3 + X^2Z + Y^2Z + XZ^2 + YZ^2 + Z^3 \quad (28)$$

$$\psi_{(c)1} = X^3 + X^2Y + XY^2 + Y^3 + X^2Z + 2XYZ + Y^2Z + XZ^2 + YZ^2 + Z^3, \quad (29)$$

$$\psi_{(b)1} = X^3 + X^2Y + XY^2 + Y^3 + X^2Z + 2XYZ + Y^2Z + XZ^2 + YZ^2 + Z^3 \quad (30)$$

Compare these equations derived by the present proligand approach with Eq. (7) derived by Pólya's corona. In particular, the term $2XYZ$ appearing in the right-hand side of Eq. (29) stems from two modes of compensated pairwise packing of $R\text{-CXYZ}$ and $S\text{-CXYZ}$ for an enantiospheric orbit [1]. On the other hand, the term $2XYZ$ appearing in the right-hand side of Eq. (30) represents a respective occupation by $R\text{-CXYZ}$ or $S\text{-CXYZ}$ for a hemispheric orbit, although the same term XYZ is used to represent $R\text{-CXYZ}$ as well as $S\text{-CXYZ}$.

3.2 Molecules derived from promolecules

Let $\hat{\mathbf{G}}$ be a sum of coset representations of a point group \mathbf{G} , which governs a set $\hat{\Delta}$ of \hat{n} positions of a skeleton. Suppose that the cycle structure of $\hat{G}(\in \hat{\mathbf{G}})$ is represented by

$$(1^{\mu_1(\hat{G})} 2^{\mu_2(\hat{G})} \dots \hat{n}^{\mu_{\hat{n}}(\hat{G})}),$$

where $\sum_{k=1}^{\hat{n}} k\mu_k(\hat{G}) = \hat{n}$. Each position of $\hat{\Delta}$ is occupied by an achiral or a chiral ligand selected from a set of ligands enumerated by Eqs. (19), (20), and (21). When k of ligands enumerated by Eq. (19) occupy k equivalent positions of the skeleton (corresponding to $k^{\mu_k(\hat{G})}$), the corresponding k of $\psi_{(a)1}$ become transitive to give $\psi_{(a)k}$. This holds true for $\psi_{(c)1}$ and $\psi_{(b)1}$. As a result, we obtain the following inventories:

$$\psi_{(a)k} = 2\text{CI-CF}(\hat{\mathbf{H}}; \mathcal{S}_{kd}) - \text{CI-CF}(\hat{\mathbf{H}}'; b_{kd}), \quad (31)$$

$$\psi_{(c)k} = \text{CI-CF}(\hat{\mathbf{H}}'; b_{kd}), \quad (32)$$

$$\psi_{(b)k} = \text{CI-CF}(\hat{\mathbf{H}}'; b_{kd}), \quad (33)$$

where k represents a multiplicity factor, as described previously. Since $\psi_{(a)k}$, $\psi_{(c)k}$, and $\psi_{(b)k}$ are a kind of extension of sphericity indices (a_k , c_k , and b_k), they are here called *extended sphericity indices of k -cycles*. Note again that the equality between Eq. (32), and Eq. (33) stems from the assumption due to Eqs. (17) and (18).

The present treatment is based on the simplification that ligands enumerated by Eqs. (11), and (12) are regarded as proligands, where their substitution is controlled by the sphericity of an orbit contained in a promolecule. The sphericity of an orbit is taken into consideration in the form of extended sphericity indices, i.e., $\psi_{(a)k}$ for a homospheric k -cycle, $\psi_{(c)k}$ for an enantiospheric k -cycle, and $\psi_{(b)k}$ for a hemispheric k -cycle. Thereby, we define the CI-CF for a (pro)molecule of $\hat{\mathbf{G}}$ -symmetry as follows:

$$\text{CI-CF}(\hat{\mathbf{G}}; \psi_{(\mathcal{S})k}) = \frac{1}{|\hat{\mathbf{G}}|} \sum_{\hat{G} \in \hat{\mathbf{G}}} \psi_{(\mathcal{S})1}^{\mu_1(\hat{G})} \psi_{(\mathcal{S})2}^{\mu_2(\hat{G})} \cdots \psi_{(\mathcal{S})n}^{\mu_n(\hat{G})}, \quad (34)$$

where $\psi_{(\mathcal{S})k}$ is $\psi_{(a)k}$ (Eq. 31) if \hat{G} is an improper element and k is odd, $\psi_{(\mathcal{S})k}$ is $\psi_{(c)k}$ (Eq. 32) if \hat{G} is an improper element and k is even, and $\psi_{(\mathcal{S})k}$ is $\psi_{(b)k}$ (Eq. 33) if \hat{G} is a proper element.

Theorem 2: *Let us consider $\hat{\mathbf{G}}$ for representing a skeleton with \hat{n} positions, each of which is occupied by a ligand generated from a ligand skeleton of $\hat{\mathbf{H}}$ having n positions (the maximum chiral subgroup: $\hat{\mathbf{H}}'$). There are in total $\hat{n}\hat{n}$ positions. Each position is occupied by an achiral or a chiral ligand selected from a set of ligands,*

$$\mathbf{X} = \{x_1, x_2, \dots, x_m; p_1, p_2, \dots, p_{m'}; \bar{p}_1, \bar{p}_2, \dots, \bar{p}_{m'}\},$$

where each x_j represents an achiral ligand and each pair of p_j and \bar{p}_j represents an enantiomeric pair of chiral ligands. Consider isomers having θ_1 of x_1 , θ_2 of x_2, \dots, θ_m of x_m ; θ'_1 of p_1, θ'_2 of $p_2, \dots, \theta'_{m'}$ of $p_{m'}$; and θ''_1 of \bar{p}_1, θ''_2 of $\bar{p}_2, \dots, \theta''_{m'}$ of $\bar{p}_{m'}$, where

$$\begin{aligned} [\theta] &: \theta_1 + \theta_2 + \cdots + \theta_m \\ &+ \theta'_1 + \theta'_2 + \cdots + \theta'_{m'} \\ &+ \theta''_1 + \theta''_2 + \cdots + \theta''_{m'} = n\hat{n}. \end{aligned} \quad (35)$$

Let the symbol A_0 denote the number of isomers of such isomers as having $[\theta]$ under the action of $\hat{\mathbf{G}}$, where achiral isomers and enantiomeric pairs are enumerated A generating function for calculating A_0 is represented by

$$\begin{aligned} \sum_{[\theta]} A_0 x_1^{\theta_1} x_2^{\theta_2} \cdots x_m^{\theta_m} p_1^{\theta'_1} p_2^{\theta'_2} \cdots p_{m'}^{\theta'_{m'}} \bar{p}_1^{\theta''_1} \bar{p}_2^{\theta''_2} \cdots \bar{p}_{m'}^{\theta''_{m'}} \\ = \text{CI-CF}(\hat{\mathbf{G}}; \psi_{(\mathcal{S})k}), \end{aligned}$$

where the right-hand side is represented by Eq. (34), into which Eqs. (31), (32) and (33) are introduced under the

presumption of $p_i = \bar{p}_i$. The ligand inventories (Eqs.16, 17, and 18) are finally introduced into the resulting polynomial.

To show the versatility of theorem 2, let us continue the example of tetramethylmethane. According to Eqs. (31), (32), and (33), Eqs. (24), (25), and (26) are converted into extended sphericity indices as a set of inventories:

$$\psi_{(a)k} = a_k c_{2k}, \quad (37)$$

$$\psi_{(c)k} = \frac{1}{3} b_k^3 + \frac{2}{3} b_{3k}, \quad (38)$$

$$\psi_{(b)k} = \frac{1}{3} b_k^3 + \frac{2}{3} b_{3k}. \quad (39)$$

According to Eq. (34), the CI-CF for this case is obtained as follows:

$$\begin{aligned} \text{CI} - \text{CF}(\hat{\mathbf{G}}; \psi_{(\mathcal{S})k}) &= \frac{1}{24} \psi_{(b)1}^4 + \frac{1}{8} \psi_{(b)2}^2 + \frac{1}{4} \psi_{(a)1}^2 \psi_{(c)2} \\ &+ \frac{1}{3} \psi_{(b)1} \psi_{(b)3} + \frac{1}{4} \psi_{(c)4}. \end{aligned} \quad (40)$$

Compare this equation with Eq. (4) derived by Pólya's corona.

After introducing the extended sphericity indices (Eqs. 37, 38, and 39) into Eq. (40), the ligand inventories described in Eq. (27) are introduced into the resulting polynomial, since we consider X , Y , and Z as achiral ligands. Thereby, we obtain the following generating function:

$$\begin{aligned} F = & (X^{12} + Y^{12} + Z^{12}) \\ & + (X^{11}Y + XY^{11} + X^{11}Z + \cdots) \\ & + 2(X^{10}Y^2 + X^2Y^{10} + X^{10}Z^2 \cdots) \\ & + 3(X^9Y^3 + X^3Y^9 + X^9Z^3 + \cdots) \\ & + 4(X^8Y^4 + X^4Y^8 + X^8Z^4 + \cdots) \\ & + 4(X^7Y^5 + X^5Y^7 + X^7Z^5 + \cdots) \\ & + 5(X^6Y^6 + X^6Z^6 + Y^6Z^6) \\ & + 2(X^{10}YZ + XYZ^{10} + XY^{10}Z) \\ & + 4(X^9Y^2Z + X^2Y^9Z + X^9YZ^2 + \cdots) \\ & + 6(X^8Y^3Z + XY^8Z^3 + X^3YZ^8 + \cdots) \\ & + 10(X^8Y^2Z^2 + X^2Y^8Z^2 + X^2Y^2Z^8) \\ & + 9(X^7Y^4Z + X^4Y^7Z + X^7YZ^4 + \cdots) \\ & + 15(X^3Y^7Z^2 + X^7Y^3Z^2 + X^7Y^2Z^3 + \cdots) \\ & + 10(X^6Y^5Z + X^5Y^6Z + X^6YZ^5 + \cdots) \\ & + 22(X^6Y^4Z^2 + X^4Y^6Z^2 + X^6Y^2Z^4 + \cdots) \\ & + 24(X^6Y^3Z^3 + X^3Y^6Z^3 + X^3Y^3Z^6) \\ & + 22(X^5Y^5Z^2 + X^5Y^2Z^5 + X^2Y^5Z^5) \\ & + 30(X^5Y^4Z^3 + X^4Y^5Z^3 + X^5Y^3Z^4 + \cdots) \\ & + 39X^4Y^4Z^4. \end{aligned} \quad (41)$$

The coefficient of each term $X^x Y^y Z^z$ ($x + y + z = 12$) in Eq. (41) represents the number of isomers with x of X , y

of Y , and z of Z . It should be noted that the numbers collected in Eq. (41) were obtained by using the present proligand approach.

4 Comparison between two approaches

In comparison with Pólya's corona that uses a CI (such as Eq. 41) and a single inventory (such as Eq. 3), the present proligand approach uses a CI-CF (such as Eq. 41) and three inventories (such as Eqs. 37, 38, and 39). This difference results in the difference of isomer numbers, which is found by comparing the coefficient of each term in Eq. (6) and the corresponding one in Eq. (41).

To compare the two approaches, the factorization of each term into four factors is informative, as shown for the term $X^5Y^5Z^2$ in Table 1. This term can be factorized into 16 modes, each of which corresponds to a graph enumerated in Eq. (6) of Pólya's corona ($16X^5Y^5Z^2$).

On the other hand, the present proligand approach gives the coefficient 22 of $X^5Y^5Z^2$ (Eq. 41), which is ascribed to the categorization of isomers collected in the last column of Table 1.

To clarify the difference between the present proligand approach and Pólya's corona in detail, let us examine the case of $XY^2 \times X^2Y \times (XYZ)^2$, which corresponds to promolecules such as $ABp\bar{p}$, $ABpp$, and $AB\bar{p}\bar{p}$. Note that XYZ corresponds to a chiral proligand p or \bar{p} . Among the three isomers, two isomers have been described in Fig. 1 (1 and 2). The remaining one is a pair of enantiomers (4 and 5), as depicted in Fig. 2. Note that each pair of enantiomers is counted once in the present proligand approach.

This exhibits a sharp contrast to the fact that Pólya's corona regards $ABp\bar{p}$ (3 in Fig. 1), $ABpp$, and $AB\bar{p}\bar{p}$ (3 in Fig. 2) as the same graph so as to be counted once, as shown in Table 1.

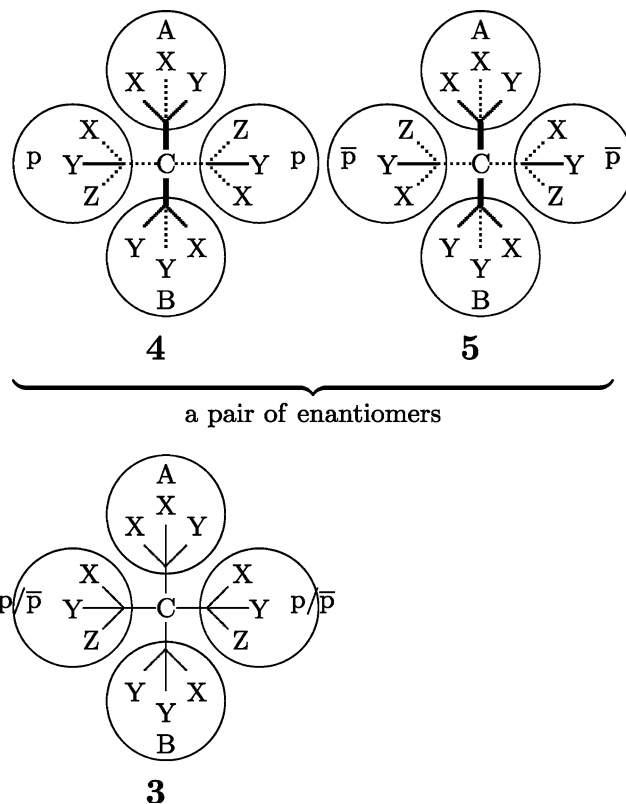


Fig. 2. One pair of enantiomers for chemical structure (4 and 5) versus one molecule for graph (the same as 3). For the symbols used, see Fig. 1.

Another difference between the present proligand approach and Pólya's corona appears in the isomer numbers of the factorization $X^2Z \times X^2Y \times Y^3 \times XYZ$ (1 versus 2 in Table 1), which corresponds to promolecules such as $ABCp$ (or $ABC\bar{p}$). According to the present approach, the two isomers (6 and 7) are depicted in Fig. 3, where an arbitrary enantiomer is selected from each pair of enantiomers, since both 6 and 7 are chiral.

Table 1 Factorization of the term $X^5Y^5Z^2$ for tetramethylmethane

Factors	Promolecule	Isomer number due to Pólya's corona	Isomer number due to the proligand approach
$XZ^2 \times XY^2 \times X^3 \times Y^3$	ABCD	1	1
$YZ^2 \times X^2Y \times X^3 \times Y^3$	ABCD	1	1
$X^2Z \times Y^2Z \times X^3 \times Y^3$	ABCD	1	1
$X^2Z \times Y^2Z \times XY^2 \times XY^2$	ABCD	1	1
$(X^2Z)^2 \times XY^2 \times Y^3$	A ² BC	1	1
$(Y^2Z)^2 \times X^2Y \times X^3$	A ² BC	1	1
$(X^2Y)^2 \times XY^2 \times Y^2Z$	A ² BC	1	1
$(XY^2)^2 \times X^2Y \times X^2Z$	A ² BC	1	1
$(X^2Y)^2 \times XZ^2 \times Y^3$	A ² BC	1	1
$(XY^2)^2 \times YZ^2 \times X^3$	A ² BC	1	1
$(X^2Y)^2 \times Y^2Z \times XYZ$	A ² Bp	1	1
$(XY^2)^2 \times X^2Z \times XYZ$	A ² Bp	1	1
$X^2Z \times X^2Y \times Y^3 \times XYZ$	ABCp	1	2
$Y^2Z \times XY^2 \times X^3 \times XYZ$	ABCp	1	2
$XY^2 \times X^2Y \times (XYZ)^2$	ABp \bar{p} , ABp ²	1	3
$X^3 \times Y^3 \times (XYZ)^2$	ABp \bar{p} , ABp ²	1	3
Total		16	22

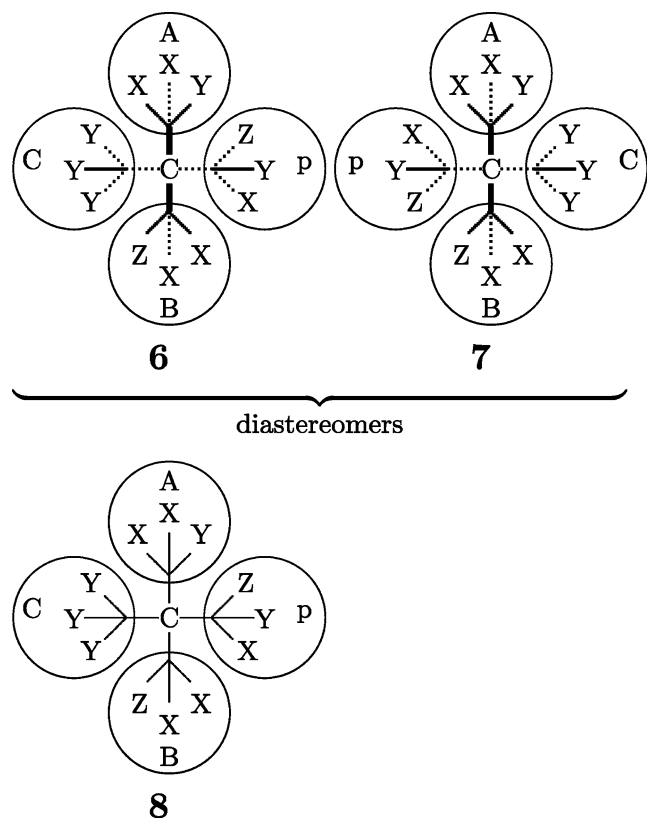


Fig. 3 Two molecules for chemical structure (6 and 7) versus one molecule for graph (8). For the *symbols* used, see Fig. 1. Since both 6 and 7 are chiral, an appropriate enantiomer is depicted as a representative of each pair of enantiomers.

Note again that a pair of enantiomers is counted once in the present proligand approach. The two isomers (6 and 7) are in diastereomeric relationship.

In contrast, Pólya's corona equalizes the two isomers (6 and 7) so as to regard them as one graph (8), as depicted in the bottom of Fig. 3.

5 Comparison with other approaches

The unit-subduced-CI (USCI) approach developed by us [1] has been used to count derivatives of tetramethylmethane [16], where the results have been itemized with respect to molecular formulas as well as to point-group symmetries. The total number of each molecular formula ($X^xY^yZ^z$) collected in Table 4 of [16] is identical with the result collected in Eq. (41). The USCI approach requires a mark table based on a group-subgroup relationship, which is not always a trivial thing, although it gives more detailed classification of isomers.

The partial-CI method [1], which is an application of the USCI approach, has been applied to the enumeration of tetramethylmethane derivatives [17]. It has given results equivalent to the USCI approach where it also requires a mark table. Moreover, the paper systematic enumeration of nonrigid isomers with given ligand symmetries has been discussed [17].

The characteristic-monomial method [18,19,20], has been applied to the enumeration of tetramethylmethane derivatives [21]. This method is based on linear representations and provides a simple method for enumeration of nonrigid isomers, where ligand inventories equivalent to Eqs. (28), (29), and (30) have been alternatively obtained.

As compared with these methods, one of the merits of the present proligand approach is that it requires knowledge on the permutation of each element. This merit is more remarkable when a detailed enumeration with respect to point-group symmetries is not required.

6 Conclusions

In order to enumerate nonrigid isomers, we have proposed the proligand approach, in which a promolecule is regarded as a skeleton having proligands and a molecule is produced by replacing each proligand by ligands. This treatment has provided a new formulation for a stereochemical extension of Pólya's coronas, where extended sphericity indices of k -cycles have been defined according to the enantiospheric, homospheric, or hemispheric nature of each k -cycle. Thereby, CIs-CF have been defined so as to enumerate nonrigid stereoisomers with chiral and achiral ligands. We can conclude that Pólya's corona is concerned with graphs, but not with chemical structures, where it is incapable of treating chiral ligands properly.

References

1. Fujita S (1991) Symmetry and combinatorial enumeration in chemistry. Springer, Berlin Heidelberg New York
2. Pólya G (1937) 9 Acta Math. 68:145-254
3. Pólya G, Read RC (1987) Combinatorial enumeration of groups, graphs, and chemical compounds. Springer, Berlin - Heidelberg New York
4. Harary F, Palmer EM (1973) Graphical enumeration. Academic, New York
5. Balaban AT (ed) (1976) Chemical applications of graph theory. Academic, London
6. Pólya G, Tarjan RE, Woods DR (1983) Notes on introductory combinatorics. Birkhäuser, Boston
7. Balasubramanian K (1985) Chem Rev 85:599-618
8. Lloyd EK (1987) In: King RB, Rouvray DH (eds) studies in physical and theoretical chemistry. Graph theory and topology in chemistry, vol 51. Elsevier, Amsterdam, pp 537-543
9. Balaban AT (1994) In: Bonchev D, Rouvray DH (eds) Chemical group theory. Introduction and fundamentals. Gordon & Breach, Switzerland, pp 159-208
10. Fujita S (2004) Theor Chem Acc 113: (in press)
11. Fujita S (1991) Tetrahedron 47:31-46
12. Balasubramanian K (1979) Theor Chim Acta 51:37
13. Rouvray DH (1974) Chem Soc Rev 3:355-372
14. Harary F, Palmer EM, Robinson RW, Read RC (1976) In: Balaban AT (ed) Chemical applications of graph theory. Academic, London, pp 11-24
15. Fujita S (1990) J Am Chem Soc 112:3390-3397
16. Fujita S (1990) Bull Chem Soc Jpn 63:2033-2043
17. Fujita S (2000) J Chem Inf Comput Sci 40:135-146
18. Fujita S (1998) Theor Chem Acc 99:224-230
19. Fujita S (1998) Chem Soc Jpn 71:1587-1596
20. Fujita S (1998) Theor Chem Acc 99:404-410
21. Fujita S (1999) Bull Chem Soc Jpn 72:2403-2407