Theor Chem Acc (2007) 117:339–351
DOI 10.1007/s00214-006-0163-8

REGULAR ARTICLE

Graphs to chemical structures 5. Combinatorial enumeration of centroidal and bicentroidal three-dimensional trees as stereochemical models of alkanes

Shinsaku Fujita

Received: 27 April 2006 / Accepted: 18 July 2006 / Published online: 7 September 2006

Abstract Three-dimensional trees (3D-trees), which are defined as a 3D version of trees, are enumerated by Fujita's proligand method formulated in Part 1 to Part 3 of this series (Fujita in Theor Chem Acc 113:73–79, 113:80–86, 2005; 115:37–53, 2006). Such 3D-trees are classified into centroidal and bicentroidal 3D-trees, which correspond to respective promolecules having proligands as substituents. In order to enumerate such centroidal and bicentroidal 3D-trees, cycle indices with chirality fittingness (CI-CFs) are formulated as being composed of three kinds of sphericity indices, i.e., a_d for homospheric cycles, c_d for enantiospheric cycles, and b_d for hemispheric cycles. The CI–CFs are capable of giving itemized results with respect to chiral and achiral 3D-trees so that they are applied to derive functional equations $\left(a(x), c\left(x^2\right)\right.$, and $\left.b(x)\right)$. The generating functions of planted 3D-trees, which are formulated and calculated elsewhere, are introduced into such functional equations. Thereby, the numbers of 3D-trees or equivalently those of alkanes as stereoisomers are calculated and collected up to a carbon content of 20 in a tabular form. Now, the enumeration problem initiated by a mathematician Cayley (Philos Mag 47(4):444–446, 1874) has been solved in such a systematic and integrated manner as satisfying both mathematical and chemical requirements.

Keywords 3D-tree · Alkane · Promolecule · Cycle index · Sphericity index · Stereoisomer enumeration

S. Fujita (✉)
Department of Chemistry and Materials Technology,
Kyoto Institute of Technology, Matsugasaki, Sakyoku,
Kyoto 606-8585, Japan
e-mail: fajitas@chem.kit.ac.jp

1 Introduction

The enumeration of trees or alkanes as their chemical counterparts has been one of the classical problems of chemical combinatorics [1–8], since it was initiated by a mathematician Cayley [9,10]. Henze and Blair [11] obtained the number of alkane of a given carbon content where the alkane were regarded as graphs, not as three-dimensional objects (3D-objects). Pólya [12,13] applied his main theorem (Hauptsatz) to the evaluation of the number of trees as graphs by using alternating groups and symmetric groups. Although the alternating group of degree 3 (or 4) was able to count spatial isomers in terms of Pólya's treatment, it did not distinguish between achiral stereoisomers and two enantiomers of chiral ones. Moreover, the symmetric group of degree 3 (or 4) was not applicable to count spatial isomers, because it was incapable of determining such stereochemical relationships as enantiomeric and diastereomeric ones. Even when Pólya's theorem took account of asymmetric carbon centers (cf. Sect. 42 of [12,13]), it overlooked problems on pseudoasymmetry and *meso*-compounds, which required comprehensive understanding on the enantiomeric and diastereomeric relationships. Although Robinson et al. [14] has reported the enumeration of alkanes as stereoisomers by modifying Pólya's cycle indices (CIs), it is still desirable to develop a more systematic method for comprehending the problems on pseudoasymmetry and *meso*-compounds.

The enumeration of alkanes as stereoisomers requires thorough consideration on 3D configurations, where such stereoisomers are regarded as the 3D-objects having inner structure. For this purpose, we have developed the unit-subduced-cycle-index (USCI) approach by means of an algebraic derivation [15–18] as well

as by means of a diagrammatical formulation [19–21]. Fujita's USCI approach has emphasized the concept of *sphericities* of orbits governed by coset representations as well as the concepts of *proligands* and *promolecules* [22–24]. According to these concepts, we have proposed *the proligand method* for enumerating stereoisomers in Parts 1 to 3 of this series [25–27]. To accomplish the enumeration of alkanes, the concepts brought about by Fujita's USCI approach and the proligand method should be extended to be able to determine the nested nature appearing in the structures of alkanes.

In the present paper, the 3D-trees will be formulated by using the concepts of proligands and promolecules, which are also used to model stereoisomers having rotatable ligands. The non-rigid nature of the 3D-trees and that of the stereoisomers will be explained in the common theoretical framework. Then, Fujita's proligand method will be applied to the enumeration of such 3D-trees as models of alkanes. Such 3D-trees will be categorized into centroidal and bicentroidal 3D-trees, the numbers of which will be evaluated distinctly by using generating functions.

2 Centroidal and bicentroidal three-dimensional trees

After a tree is defined usually as a graph which has n vertices and e edges satisfying the relation $n = e + 1$, a three-dimensional tree (3D-tree) is defined as a 3D version of such a tree. A vertex of a given tree has incident branches, where the total number of vertices in such branches is equal to $n - 1$. Let m be the number of vertices contained in the largest branch among the branches attaching to the vertex. According to Jordan [28], trees are classified into two categories, i.e., trees having a centroid and trees having a bicentroid. Thus, there emerge two cases:

1. A given tree has an exceptional vertex (M) called a *centroid*, which satisfies the relation $m < \frac{1}{2}n$. The tree is called a *centroidal tree*.
2. A given tree has two adjacent vertices (M_1 and M_2), each of which satisfies the relation $m = \frac{1}{2}n$. The exceptional graph (M_1–M_2) composed of the two adjacent vertices and the relevant edge is called a *bicentroid*. The tree is called a *bicentroidal tree*.

All of the vertices of the tree other than the centroid or the bicentroid satisfy the relation $m > \frac{1}{2}n$. There are no cases in which a given tree has both a centroid and a bicentroid. Obviously, these properties of trees maintain in 3D trees, because the 3D-trees are regarded as a kind of trees.

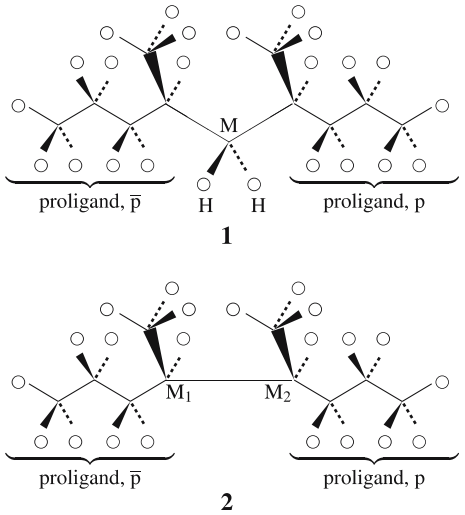


Fig. 1 Three-dimensional trees with a centroid (**1**) and with a bicentroid (**2**). Each joint of two or more edges represents a vertex. Each *open circle* represents an end vertex. The vertex M is referred to as a centroid. The vertices M_1 and M_2 construct a bicentroid

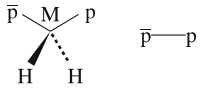
For example, the vertex (M) of **1** shown in Fig. 1 is a centroid, because the relation $m < \frac{1}{2}n$ is true when we place $n = 35$ (for the alkane $C_{11}H_{24}$) and $m = 16$ (for the alkyl ligand C_5H_{11}). Hence, **1** is a centroidal 3D-tree.

On the other hand, the vertices (M_1 and M_2) of **2** shown in Fig. 1 construct a bicentroid, because the relation $m = \frac{1}{2}n$ is true when we place $n = 32$ (for the alkane $C_{10}H_{22}$) and $m = 16$ (for the alkyl ligand C_5H_{11}). Hence, **2** is a bicentroidal 3D-tree. Obviously, any bicentroidal 3D-tree can be divided into two segments (branches) having the same number (m) of vertices, where the division occurs at the edge of the centroid in agreement of the relation $m = \frac{1}{2}n$.

3 Promolecules

3.1 Promolecules as models of 3D-trees

In contrast to trees as graphs, the 3D-trees as the 3D-objects suffer from multiple (or, in fact, infinite) conformational changes due to the free rotation around each edge, even though an appropriate fixed conformer is depicted, as shown in Fig. 1. Note that each carbon atom (non-terminal vertex) is designated by a joint of two or more bonds; each hydrogen atom (terminal vertex) is designated by an open circle. In agreement with chemical conventions, the configurations of carbon atoms in such a fixed conformer are designated by using wedges and hashed bonds attached to the carbon chain. Because every conformational change cannot be illustrated even



3 **4**

Fig. 2 Promolecule (**3** and **4**) as models of the 3D-trees (**1** and **2**). The *symbols p and* p̄ represent a pair of enantiomeric proligands. The *symbol H* represents an achiral proligand. Note that such a hydrogen atom (H) is also regarded as a proligand

for a single 3D-tree, the test of the congruence between two 3D-trees is not an easy task.

To assure this type of congruence, the concepts of proligands and promolecules [15,22] are applied to this case. The branches selected in the process of determining a centroid or a bicentroid are regarded as proligands as follows:

1. When a given 3D-tree is a centroidal 3D-tree, we first select the centroid M and the relevant incident edges so as to construct a skeleton with substitution positions. When a given 3D-tree is a bicentroidal 3D-tree, the bicentroid (M_1–M_2) is regarded as a skeleton with substitution positions.
2. Next, we derive proligands by starting from the 3D-tree. When the part selected as the skeleton is deleted from the 3D-tree, the remaining part is partitioned into several branches, each of which is linked to the vertex chosen as each of the substitution positions in the selection of the skeleton. These branches can be regarded as proligands.
3. After the specification of the skeleton and the proligands, they are considered to construct a kind of promolecule.

By following this procedure, the centroidal 3D-tree (**1**) gives a promolecule (**3**) shown in Fig. 2, where a skeleton of $\mathbf{T}_d$-symmetry is selected by the first step of the above procedure. Then, the branches around the centroid, i.e., two open circles (H), an R-CH(CH_3) ($CH_2CH_2CH_3$) segment (p), and an S-CH(CH_3) ($CH_2CH_2CH_3$) segment (p̄), are regarded as proligands, where H is an achiral proligand, while p and p̄ are chiral proligands of an enantiomeric pair.

On the other hand, the bicentroidal 3D-tree (**2**) gives a promolecule (**4**) shown in Fig. 2, where the resulting skeleton belongs to $\mathbf{D}_{\infty d}$-symmetry.

As a result, Jordan's criteria for classifying trees into centroidal and bicentroidal trees are revised by considering promolecules as models of 3D-trees. The centroidal 3D-trees correspond to centroidal promolecules which are based on a skeleton belonging to a point group (e.g., a $\mathbf{T}_d$-skeleton for alkanes). On the other hand, the bicentroidal 3D-trees correspond to bicentroidal promolecules based on a two-nodal skeleton belonging to the $\mathbf{D}_{\infty h}$-point group. Hence, the enumeration of 3D-trees is found to be equivalent to that of promolecules classified into centroidal and bicentroidal ones.

3.2 Enumeration of promolecules

3.2.1 Promolecules for centroidal 3D-trees

Following Fujita's proligand method reported in Parts 1 to 3 of this series [25–27], Theorem 1 of Part 3 [27] is applied to the enumeration of promolecules under the action of a point group. The tetrahedral skeleton contained in the promolecule (**3**) belongs to $\mathbf{T}_d$-symmetry so that the four substitution positions are governed by the right coset representation (RCR) ($\mathbf{C}_{3v}\backslash\mathbf{T}_d$. The permutations of the RCR are collected as products of cycles in Table 1, which also contains the corresponding products of sphericity indices (PSIs) and products of dummy variables (PDVs). Thereby, the cycle index with chirality fittingness (CI-CF) for this case is calculated from the PSIs as follows:

$$\text{CI-CF}(\mathbf{T}_d, \$_d) = \frac{1}{24}\left(b_1^4 + 3b_2^2 + 8b_1b_3 + 6a_1^2c_2 + 6c_4\right), \tag{1}$$

which counts achiral promolecules and enantiomeric pairs of chiral promolecules. Note that each pair of enantiomers is counted just once in this enumeration.

Theorem 3 of [27] for the enumeration of ligands under the action of the maximum chiral subgroup can be applied to this case so as to derive the following CI-CF from the data for the **T** part of Table 1:

$$\text{CI-CF}(\mathbf{T}, b_d) = \frac{1}{12}\left(b_1^4 + 3b_2^2 + 8b_1b_3\right), \tag{2}$$

which counts achiral promolecules and chiral promolecules, where two enantiomers of each pair are counted separately.

The first proposition of Theorem 4 for the enumeration of achiral ligands [27] can be applied to this case so as to derive the following CI-CF_A:

$$\begin{aligned}\text{CI-CF}_A(\mathbf{T}_d, \$_d) &= 2\text{CI-CF}(\mathbf{T}_d, \$_d) - \text{CI-CF}(\mathbf{T}, b_d)\\ &= \frac{1}{2}(a_1^2c_2 + c_4),\end{aligned} \tag{3}$$

which counts achiral promolecules only.

The second proposition of Theorem 4 for the enumeration of chiral ligands [27] can be applied to obtain the following CI-CF_C:

Table 1 Right coset representation $(\mathbf{C}_{3v}\backslash)\mathbf{T}_d$ and products of sphericity indices

	$\mathbf{T}_d$		RCR[a]	PSI[b]	PDV[c]
T	I	∼	(1)(2)(3)(4)	b_1^4	s_1^4
	$C_{2(1)}$	∼	(1 2)(3 4)	b_2^2	s_2^2
	$C_{2(2)}$	∼	(1 4)(2 3)	b_2^2	s_2^2
	$C_{2(3)}$	∼	(1 3)(2 4)	b_2^2	s_2^2
	$C_{3(1)}$	∼	(1)(2 3 4)	b_1b_3	s_1s_3
	$C_{3(3)}$	∼	(1 2 4)(3)	b_1b_3	s_1s_3
	$C_{3(2)}$	∼	(1 4 3)(2)	b_1b_3	s_1s_3
	$C_{3(4)}$	∼	(1 3 2)(4)	b_1b_3	s_1s_3
	$C_{3(1)}^2$	∼	(1)(2 4 3)	b_1b_3	s_1s_3
	$C_{3(4)}^2$	∼	(1 2 3)(4)	b_1b_3	s_1s_3
	$C_{3(3)}^2$	∼	(1 4 2)(3)	b_1b_3	s_1s_3
	$C_{3(2)}^2$	∼	(1 3 4)(2)	b_1b_3	s_1s_3
$\mathbf{T}\sigma_{d(1)}$	$\sigma_{d(1)}$	∼	$\overline{(1)(2\,4)(3)}$	$a_1^2c_2$	$s_1^2s_2$
	$S_{4(3)}$	∼	$\overline{(1\,2\,3\,4)}$	c_4	s_4
	$S_{4(3)}^3$	∼	$\overline{(1\,4\,3\,2)}$	c_4	s_4
	$\sigma_{d(6)}$	∼	$\overline{(1\,3)(2)(4)}$	$a_1^2c_2$	$s_1^2s_2$
	$\sigma_{d(2)}$	∼	$\overline{(1)(2)(3\,4)}$	$a_1^2c_2$	$s_1^2s_2$
	$\sigma_{d(4)}$	∼	$\overline{(1\,2)(3)(4)}$	$a_1^2c_2$	$s_1^2s_2$
	$S_{4(1)}$	∼	$\overline{(1\,4\,2\,3)}$	c_4	s_4
	$S_{4(1)}^3$	∼	$\overline{(1\,3\,2\,4)}$	c_4	s_4
	$\sigma_{d(3)}$	∼	$\overline{(1)(2\,3)(4)}$	$a_1^2c_2$	$s_1^2s_2$
	$S_{4(2)}^3$	∼	$\overline{(1\,2\,4\,3)}$	c_4	s_4
	$\sigma_{d(5)}$	∼	$\overline{(1\,4)(2)(3)}$	$a_1^2c_2$	$s_1^2s_2$
	$S_{4(2)}$	∼	$\overline{(1\,3\,4\,2)}$	c_4	s_4

Each cycle appearing in an improper rotation is designated by an overbar, which represents the inversion of ligand chirality.

[a]Each permutation of the right coset representation $(\mathbf{C}_{3v}\backslash)\mathbf{T}_d$ is expressed as a product of cycles. The product of cycles is designated by an overbar, when the corresponding symmetry operation is an improper one ($\in \mathbf{T}\sigma_{d(1)}$)

[b]*PSI* Product of sphericity indices

[c]*PDV* Product of dummy variables

$$\begin{aligned}&\text{CI-CF}_C(\mathbf{T}_d, \$_d)\\ &\quad = \text{CI-CF}(\mathbf{T}, b_d) - \text{CI-CF}(\mathbf{T}_d, \$_d)\\ &\quad = \frac{1}{24}\left(b_1^4 + 3b_2^2 + 8b_1b_3 - 6a_1^2c_2 - 6c_4\right)\end{aligned} \tag{4}$$

which counts chiral promolecules only, where each pair of enantiomers is counted just once.

Suppose that a set of proligands selected from the following proligand warehouse:

$$\mathbf{X} = \{\mathrm{H}, \mathrm{X}, \mathrm{Y}, \mathrm{Z}; \mathrm{p}, \overline{\mathrm{p}}; \mathrm{q}, \overline{\mathrm{q}}; \mathrm{r}, \overline{\mathrm{r}}, \mathrm{s}, \overline{\mathrm{s}}\}, \tag{5}$$

where the symbols H, X, Y, and Z represent achiral proligands, while the pairs, p/$\overline{\mathrm{p}}$, q/$\overline{\mathrm{q}}$, and r/$\overline{\mathrm{r}}$, denote enantiomeric pairs of chiral proligands. Then, the following ligand inventories are calculated:

$$a_d = \mathrm{H}^d + \mathrm{X}^d + \mathrm{Y}^d + \mathrm{Z}^d \tag{6}$$

$$\begin{aligned}c_d = {} & \mathrm{H}^d + \mathrm{X}^d + \mathrm{Y}^d + \mathrm{Z}^d\\ & + 2\mathrm{p}^{d/2}\overline{\mathrm{p}}^{d/2} + 2\mathrm{q}^{d/2}\overline{\mathrm{q}}^{d/2} + 2\mathrm{r}^{d/2}\overline{\mathrm{r}}^{d/2} + 2\mathrm{s}^{d/2}\overline{\mathrm{s}}^{d/2}\end{aligned} \tag{7}$$

$$\begin{aligned}b_d = {} & \mathrm{H}^d + \mathrm{X}^d + \mathrm{Y}^d + \mathrm{Z}^d\\ & + \mathrm{p}^d + \overline{\mathrm{p}}^d + \mathrm{q}^d + \overline{\mathrm{q}}^d + \mathrm{r}^d + \overline{\mathrm{r}}^d + \mathrm{s}^d + \overline{\mathrm{s}}^d.\end{aligned} \tag{8}$$

These ligand inventories are introduced to Eqs. 1–4 to give generating functions for counting the respective numbers of objects. Enumeration of such promolecules itemized with respect to their subsymmetries have been discussed in a previous report [22] and in Fujita's book [15]. For the sake of convenience for further discussions, the results are cited from Chap. 21 of Fujita's book [15], as shown in Fig. 3. Note that each promolecule in Fig. 3 is selected as a representative of promolecules of the same type. For examples, the promolecule **5** having X^4 is a representative of promolecules having X^4, Y^4, Z^4, H^4, and so on. The promolecule (**3**) corresponds to the type **15** shown in Fig. 3.

3.2.2 Promolecules for bicentroidal 3D-trees

Promolecules for bicentroidal 3D-trees contain a two-nodal skeleton of $\mathbf{D}_{\infty h}$-symmetry. The two positions of the skeleton are governed by the right coset representation $(\mathbf{C}_{\infty v}\backslash)\mathbf{D}_{\infty h}$. According to the treatment reported previously [23], we consider the factor group:

$$\mathbf{K} = \mathbf{D}_{\infty h}/\mathbf{C}_\infty = \{\mathbf{C}_\infty I, \mathbf{C}_\infty C_2, \mathbf{C}_\infty \sigma_v, \mathbf{C}_\infty \sigma_h\} \tag{9}$$

and its subgroups:

$$\mathbf{K}_1 = \mathbf{C}_\infty/\mathbf{C}_\infty = \{\mathbf{C}_\infty I\} \tag{10}$$

$$\mathbf{K}_2 = \mathbf{D}_\infty/\mathbf{C}_\infty = \{\mathbf{C}_\infty I, \mathbf{C}_\infty C_2\} \tag{11}$$

$$\mathbf{K}_3 = \mathbf{C}_{\infty v}/\mathbf{C}_\infty = \{\mathbf{C}_\infty I, \mathbf{C}_\infty \sigma_v\} \tag{12}$$

$$\mathbf{K}_4 = \mathbf{C}_{\infty h}/\mathbf{C}_\infty = \{\mathbf{C}_\infty I, \mathbf{C}_\infty \sigma_h\} \tag{13}$$

$$\mathbf{K}_5 = \mathbf{K} = \mathbf{D}_{\infty h}/\mathbf{C}_\infty \tag{14}$$

Thereby, the right coset representation $(\mathbf{C}_{\infty v}\backslash)\mathbf{D}_{\infty h}$ can be replaced by the right coset representation of the factor group, i.e., $(\mathbf{K}_3\backslash)\mathbf{K}$. Note that the factor group $\mathbf{K}$ is isomorphic to the point group $\mathbf{C}_{2v} = \{I, C_2, \sigma_v, \sigma_h\}$, while the subgroup $\mathbf{K}_3$ is isomorphic to the point group $\mathbf{C}_s = \{I, \sigma_v\}$. This means that the sphericity indices for the RCR $(\mathbf{C}_{\infty v}\backslash)\mathbf{D}_{\infty h}$ having infinite symmetry operations can be discussed in terms of the RCR $(\mathbf{K}_3\backslash)\mathbf{K}$ having finite operations.

When we sequentially number the two substitution positions of the two-nodal skeleton for the bicentroidal 3D-trees, we can obtain the following right coset representation:

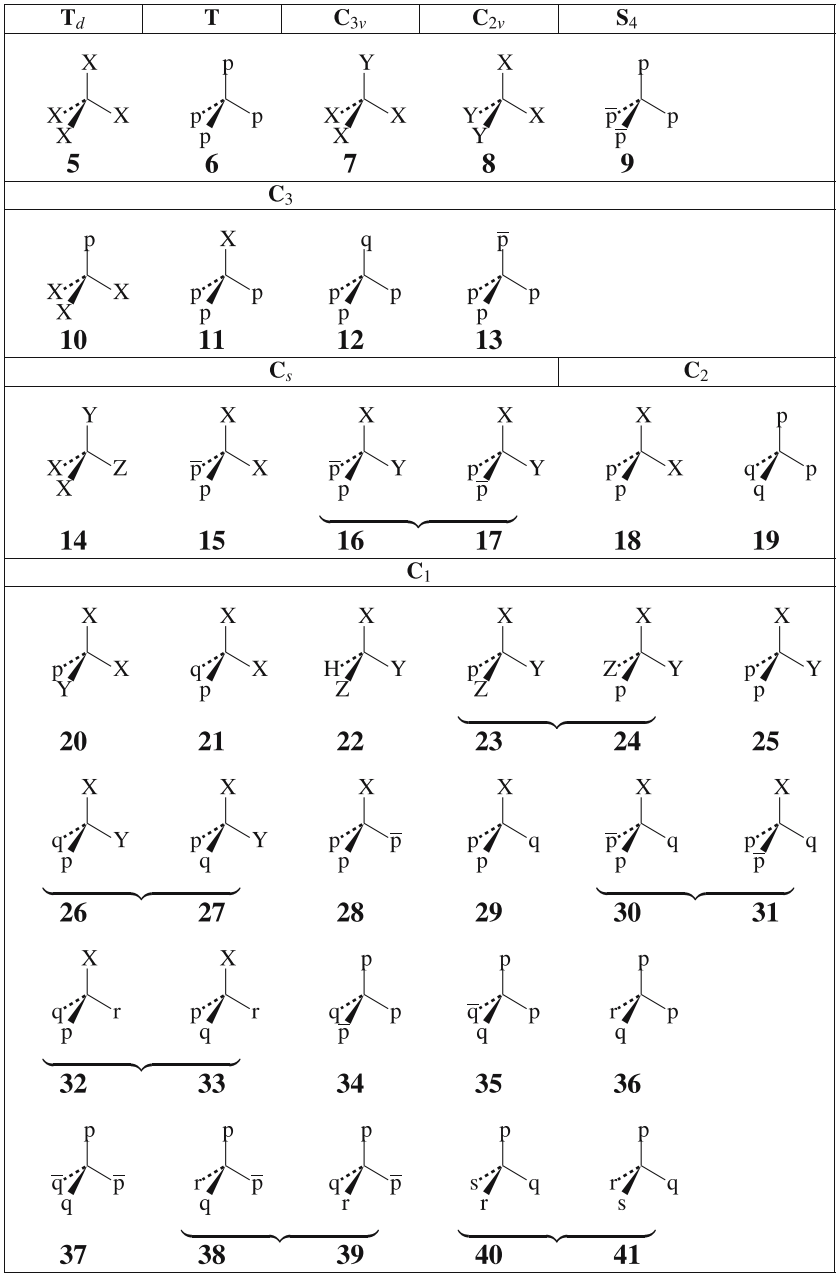


Fig. 3 Promolecules for centroidal 3D-trees, which are derived from a tetrahedral skeleton of $\mathbf{T}_d$-symmetry [22,15]. The *symbols X, Y, Z, and H* represent achiral proligands, while the *symbols* $p/\overline{\mathrm{p}}$, $q/\overline{\mathrm{q}}$, $r/\overline{\mathrm{q}}$, *and* $s/\overline{\mathrm{s}}$ represent enantiomeric pairs of chiral proligands. Either one selected from a enantiomeric pair of promolecules is depicted as a representative

$$(\mathbf{K}_3\backslash)\mathbf{K} = \{(1)(2), (1\ 2), \overline{(1)(2)}, \overline{(1\ 2)}\}, \tag{15}$$

which permutes the two positions, where each overbar represents the alternation of the chirality of a proligand occupying the position. The products of sphericity indices b_1^2 and b_2 are assigned to the cycles (1 2) and (1)(2), while a_1^2 and c_2 are assigned to $\overline{(1)(2)}$ and $\overline{(1\ 2)}$. Following Theorem 1 of Part 3 [27], the CI-CF for this case is obtained as follows:

$$\text{CI-CF}(\mathbf{K}; \$_d) = \frac{1}{4}\left(b_1^2 + b_2 + a_1^2 + c_2\right), \tag{16}$$

where the symbol $\$_d$ represents a sphericity index, a_d, b_d, or c_d, according to the respective sphericity. An equation equivalent to Eq. 16 has been reported previously on the basis of Fujita's USCI approach [23].

Theorem 3 of [27] for the enumeration of ligands under the action of the maximum chiral subgroup can be applied to this case so as to derive the following CI-CF:

$$\text{CI-CF}(\mathbf{K}_2; b_d) = \frac{1}{2}\left(b_1^2 + b_2\right), \tag{17}$$

which counts achiral promolecules and chiral promolecules, where two enantiomers of each pair are counted separately.

The first proposition of Theorem 4 for the enumeration of achiral ligands [27] can be applied to this case so as to derive the following CI-CF$_A$:

$$\text{CI-CF}_A(\mathbf{K}, \$_d) = 2\text{CI-CF}(\mathbf{K}, \$_d) - \text{CI-CF}(\mathbf{K}_2, b_d)$$
$$= \frac{1}{2}\left(a_1^2 + c_2\right), \tag{18}$$

which counts achiral promolecules only, where each pair of enantiomers is counted just once.

The second proposition of Theorem 4 for the enumeration of chiral ligands [27] can be applied to obtain the following CI-CF$_C$:

$$\text{CI-CF}_C(\mathbf{K}, \$_d) = \text{CI-CF}(\mathbf{K}_2, b_d) - \text{CI-CF}(\mathbf{K}, \$_d)$$
$$= \frac{1}{4}\left(b_1^2 + b_2 - a_1^2 - c_2\right) \tag{19}$$

which counts chiral promolecules only.

Suppose that a set of proligands selected from the following proligand warehouse:

$$\mathbf{X}' = \{\mathrm{X}, \mathrm{Y}; \mathrm{p}, \overline{\mathrm{p}}; \mathrm{q}, \overline{\mathrm{q}}\}, \tag{20}$$

In terms of Theorem 1 of Part 3 [27], we use the following ligand inventories:

$$a_d = \mathrm{X}^d + \mathrm{Y}^d \tag{21}$$

$$c_d = \mathrm{X}^d + \mathrm{Y}^d + 2\mathrm{p}^{d/2}\overline{\mathrm{p}}^{d/2} + 2\mathrm{q}^{d/2}\overline{\mathrm{q}}^{d/2} \tag{22}$$

$$b_d = \mathrm{X}^d + \mathrm{Y}^d + \mathrm{p}^d + \overline{\mathrm{p}}^d + \mathrm{q}^d + \overline{\mathrm{q}}^d \tag{23}$$

These inventories (Eqs. 21–23) are introduced into Eq. 16 and the resulting equation is expanded to give the following generating function:

$$F = \left[\mathrm{X}^2 + \mathrm{Y}^2\right] + \mathrm{XY}$$
$$+ \frac{1}{2}\left[(\mathrm{Xp} + \mathrm{X}\overline{\mathrm{p}}) + (\mathrm{Xq} + \mathrm{X}\overline{\mathrm{q}}) + (\mathrm{Yp} + \mathrm{Y}\overline{\mathrm{p}})\right.$$
$$\left. + (\mathrm{Yq} + \mathrm{Y}\overline{\mathrm{q}})\right] + \frac{1}{2}\left[(\mathrm{p}^2 + \overline{\mathrm{p}}^2) + (\mathrm{q}^2 + \overline{\mathrm{q}}^2)\right]$$
$$+ \frac{1}{2}\left[(\mathrm{pq} + \overline{\mathrm{pq}}) + (\mathrm{p}\overline{\mathrm{q}} + \overline{\mathrm{p}}\mathrm{q})\right] + [\mathrm{p}\overline{\mathrm{p}} + \mathrm{q}\overline{\mathrm{q}}], \tag{24}$$

where the coefficient of each term represents the number of stereoisomers. Note that this enumeration counts a pair of enantiomers just once. In other words, the number of achiral promolecules (bicentroidal 3D-trees) plus pairs of enantiomeric promolecules (bicentroidal 3D-trees) is obtained by Eq. 24. Hence, the term $(1/2)(\mathrm{Xp} + \mathrm{X}\overline{\mathrm{p}})$, for example, represents one pair of enantiomers (Xp and $\mathrm{X}\overline{\mathrm{p}}$).

The inventories (Eqs. 21–23) are introduced into Eq. 17 and the resulting equation is expanded to give the following generating function:

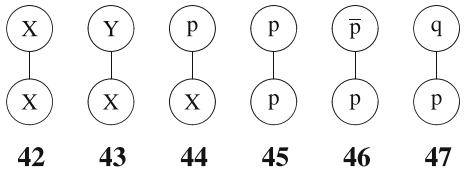


Fig. 4 Promolecules for bicentroidal 3D-trees [23]. The promolecules X–X, X–Y, and p–$\overline{\text{p}}$ are achiral. The other promolecules A–p, p–p, and p–q are chiral, where an appropriate enantiomer is depicted for each pair of enantiomers

$$F' = [\mathrm{X}^2 + \mathrm{Y}^2] + \mathrm{XY}$$
$$+ \left[(\mathrm{Xp} + \mathrm{X}\overline{\mathrm{p}}) + (\mathrm{Xq} + \mathrm{X}\overline{\mathrm{q}}) + (\mathrm{Yp} + \mathrm{Y}\overline{\mathrm{p}})\right.$$
$$\left. + (\mathrm{Yq} + \mathrm{Y}\overline{\mathrm{q}})\right] + \left[(\mathrm{p}^2 + \overline{\mathrm{p}}^2) + (\mathrm{q}^2 + \overline{\mathrm{q}}^2)\right]$$
$$+ \left[(\mathrm{pq} + \overline{\mathrm{pq}}) + (\mathrm{p}\overline{\mathrm{q}} + \overline{\mathrm{p}}\mathrm{q})\right] + [\mathrm{p}\overline{\mathrm{p}} + \mathrm{q}\overline{\mathrm{q}}], \tag{25}$$

where achiral promolecules plus two enantiomeric promolecules of every pair are counted separately as models of bicentroidal 3D-trees.

The inventories (Eqs. 21–23) are introduced into Eq. 18 and the resulting equation is expanded to give the following generating function:

$$F_A = \left[\mathrm{X}^2 + \mathrm{Y}^2\right] + \mathrm{XY} + \left[\mathrm{p}\overline{\mathrm{p}} + \mathrm{q}\overline{\mathrm{q}}\right], \tag{26}$$

where achiral promolecules are only counted as models of the corresponding bicentroidal 3D-trees.

Finally, the inventories (Eqs. 21–23) are introduced into Eq. 19 and the resulting equation is expanded to give the following generating function:

$$F_C = \frac{1}{2}\left[(\mathrm{Xp} + \mathrm{X}\overline{\mathrm{p}}) + (\mathrm{Xq} + \mathrm{X}\overline{\mathrm{q}}) + (\mathrm{Yp} + \mathrm{Y}\overline{\mathrm{p}})\right.$$
$$\left. + (\mathrm{Yq} + \mathrm{Y}\overline{\mathrm{q}})\right] + \frac{1}{2}\left[(\mathrm{p}^2 + \overline{\mathrm{p}}^2) + (\mathrm{q}^2 + \overline{\mathrm{q}}^2)\right]$$
$$+ \frac{1}{2}\left[(\mathrm{pq} + \overline{\mathrm{pq}}) + (\mathrm{p}\overline{\mathrm{q}} + \overline{\mathrm{p}}\mathrm{q})\right], \tag{27}$$

where each pair of enantiomeric promolecules is counted just once as models of the corresponding bicentroidal 3D-trees.

Obviously, the results shown in Eqs. 24–27 satisfy $F_A = 2F - F'$, $F_C = F' - F$, and $F = F_A + F_C$.

The promolecules for bicentroidal 3D-trees are illustrated in Fig. 4, where an arbitrary one of enantiomers is depicted as a representative of this chiral problem (e.g., X–p, p–p, and p–q). The results essentially equivalent to but more detailed than Fig. 4 have been obtained by employing two methods of Fujita's USCI approach [23,29].

4 Enumeration of 3D-trees

Any 3D-tree (or any promolecule) can be characterized by the number n of non-terminal nodes (chemically, carbon nodes of an alkane) so as to be represented by the term x^n. For example, we assign x^{11} to **1** and x^{10} to **2**. Hereafter, we consider non-terminal nodes (as carbon nodes) only, so that the power n of the term x^n is called *carbon content* for the sake of convenience. In this treatment, however, an endpoint (i.e., hydrogen) is taken into consideration in terms of $x^0(=1)$ to treat stereoisomers properly. Although such hydrogen content does not appear in the term x^n, the multiple participations of the term x^0 should be explicitly considered in the process of proving relevant propositions which are intended to characterize the 3D nature of the 3D-tree.

4.1 Enumeration of centroidal 3D-trees

4.1.1 Formulation

According to this type of expressions, e.g., x^n for 3D-trees of carbon content n, we first discuss the enumeration of centroidal 3D-trees. Let $\widehat{B}(x)$ be a generating function for enumerating achiral centroidal promolecules plus enantiomeric pairs of chiral centroidal promolecules up to carbon content n, i.e.

$$\widehat{B}(x) = \sum_{k=0}^{n} \widehat{B}_k x^k \tag{28}$$

Our target is to calculate $\widehat{B}_n$ which satisfies the criterion of centroidal 3D-trees. This can be accomplished by using Theorem 6 of Part 3 [27], where we employ the CI-CF($\mathbf{T}_d, \$_d$) shown in Eq. 1. The SIs a_d, c_d, and b_d involved in Eq. 1 are replaced by the terms $a(x^d)$, $c(x^d)$, and $b(x^d)$, respectively, where we emphasize that such terms as $a(x^d)$, $c(x^d)$, and $b(x^d)$ are calculated as generating functions before they are substituted for the SIs. Because Eq. 1 for CI-CF($\mathbf{T}_d, \$_d$) ignores the centroid of the parent promolecule tentatively (cf. **3**), the following functional equation is obtained by multiplying by x.

$$\widehat{B}(x) = \frac{x}{24}\big(b(x)^4 + 3b(x^2)^2 + 8b(x)b(x^3) + 6a(x)^2c(x^2) + 6c(x^4)\big). \tag{29}$$

Note that, in contrast to the cases of planted 3D-trees, the addition of 1 for the initial (trivial) promolecule is inessential in this case, because the further recursive use of Eq. 29 is not intended for centroidal 3D-trees.

The criterion for centroidal 3D-trees means that the maximum number (m), which is the number of non-terminal vertices in the largest proligand, is restricted to satisfy the following condition:

$$\frac{1}{2}n - 1 \le m < \frac{1}{2}n \tag{30}$$

or equivalently $2m < n \le 2m+2$. Hence, the warehouse of proligands are restricted within the range of carbon content equal to or less than m as follows:

$$\mathbf{X}_m = \{1, x^1, x^2, \ldots, x^m\}, \tag{31}$$

where x^k $(k = 0, 1, 2, \ldots, m)$ is presumed to be selected multiply so as to represent isomeric proligands.

By examining the ligand inventories corresponding to Eqs. 6–8, the respective functions appearing in the right-hand side of Eq. 29 are found to indicate that $a(x)$ is a generating function for enumerating achiral proligands (cf. Eq. 6); $c(x)$ is a generating function for enumerating ordered enantiomeric pairs called "diploids" (cf. Eq. 7); and $b(x)$ is a generating function for enumerating achiral proligands and chiral proligands, where two enantiomers of each pair counted separately (cf. Eq. 6). Hence, we presume the following generating functions:

$$a(x) = \sum_{k=0}^{m} \alpha_k x^k \tag{32}$$

$$c(x) = \sum_{k=0}^{m} \gamma_k x^k \tag{33}$$

$$b(x) = \sum_{k=0}^{m} \beta_k x^k, \tag{34}$$

where we place $\alpha_0 = 1$, $\gamma_0 = 1$, $\beta_0 = 1$ for the trivial cases; and α_k, γ_k, and β_k represent the numbers of proligands at issue. Once the generating functions for proligands (Eqs. 32–34) are evaluated, they are introduced into Eq. 29 to obtain the target number as the coefficient $\widehat{B}_n$ of the term x^n appearing in the resulting generating function $\widehat{B}(x)$. It should be noted that the $\widehat{B}_n$ is effective only for the value n derived from a given value m by using Eq. 30.

Let $\widehat{E}(x)$ be a generating function for enumerating achiral centroidal promolecules plus chiral centroidal promolecules up to carbon content n, where two enantiomers of each pair are counted separately.

$$\widehat{E}(x) = \sum_{k=0}^{n} \widehat{E}_k x^k. \tag{35}$$

Then, the target at issue is to obtain $\widehat{E}_n$ which satisfies the criterion of centroidal 3D-trees. This can be accomplished by using Theorem 6 of Part 3 [27] after an adequate CI-CF is selected for the purpose of calculating $\widehat{E}_n$, i.e., CI-CF($\mathbf{T}, b_d$) shown in Eq. 2. Because Eq. 2 for

CI-CF($\mathbf{T}, b_d$) ignores the centroid of the parent promolecule tentatively (cf. **3**), the following functional equation is obtained by multiplying by x:

$$\widehat{E}(x) = \frac{x}{12}\left(b(x)^4 + 3b\left(x^2\right)^2 + 8b(x)b\left(x^3\right)\right). \tag{36}$$

The same generating functions for proligands as shown in Eqs. 32–34 are effective to this case. Once they are evaluated, they are introduced into Eq. 36 so as to obtain the target number as the coefficient $\widehat{E}_n$ of the term x^n appearing in the resulting generating function $\widehat{E}(x)$, where n satisfies Eq. 30.

Let $\widehat{A}(x)$ be a generating function for enumerating achiral centroidal promolecules only up to carbon content n, i.e.,

$$\widehat{A}(x) = \sum_{k=0}^{n} \widehat{A}_k x^k. \tag{37}$$

Our next target is to calculate $\widehat{A}_n$ which satisfies the criterion of centroidal 3D-trees. This can be accomplished by using Theorem 6 of Part 3 [27] after the selection of the CI-CF to be adopted. Because Eq. 3 for CI-CF$_A(\mathbf{T}_d, \$_d)$ ignores the centroid of the parent promolecule tentatively (cf. **3**), the following functional equation is obtained by multiplying by x:

$$\widehat{A}(x) = \frac{x}{2}\left(a(x)^2 c\left(x^2\right) + c\left(x^4\right)\right). \tag{38}$$

The same generating functions for proligands as shown in Eqs. 32–34 are effective to this case. After the evaluation of them, they are introduced into Eq. 38 to obtain the target number as the coefficient $\widehat{A}_n$ appearing on the term x^n of the resulting generating function $\widehat{A}(x)$ (cf. Eq. 30 for n).

Let $\widehat{C}(x)$ be a generating function for enumerating chiral centroidal promolecules, where each pair of two enantiomers is counted just once:

$$\widehat{C}(x) = \sum_{k=0}^{n} \widehat{C}_k x^k. \tag{39}$$

We then calculate $\widehat{C}_n$ which satisfies the criterion of centroidal 3D-trees, where we use Theorem 6 of Part 3 [27] after the selection of the CI-CF to be adopted here. Because Eq. 1 for CI-CF$_C(\mathbf{T}_d, \$_d)$ ignores the centroid of the parent promolecule tentatively (cf. **3**), the following functional equation is obtained by multiplying by x:

$$\widehat{C}(x) = \frac{x}{24}\left(b(x)^4 + 3b\left(x^2\right)^2 + 8b(x)b\left(x^3\right) - 6a(x)^2 c\left(x^2\right) - 6c\left(x^4\right)\right). \tag{40}$$

The same generating functions for proligands as shown in Eqs. 32–34 are effective to this case. Once they are evaluated, they are introduced into Eq. 40 to obtain the target number as the coefficient $\widehat{C}_n$ of the term x^n appearing in the resulting generating function $\widehat{C}(x)$ (cf. Eq. 30 for n).

4.1.2 Calculation

As shown in the preceding formulation, the same generating functions for proligands (Eqs. 32–34) are effective to evaluate $\widehat{B}(x)$ (Eq. 29), $\widehat{E}(x)$ (Eq. 36), $\widehat{A}(x)$ (Eq. 38), and $\widehat{C}(x)$ (Eq. 39) for centroidal 3D-trees. This means that the enumeration of such centroidal 3D-trees (or alkanes) is ascribed to the enumeration of such proligands. After regarding proligands as planted promolecules, the enumeration of proligands are conducted on the basis of Fujita's proligand method described in Parts 1–3 of this series [25–27]. Although the enumeration of proligands will be mentioned elsewhere in this series, several results are shown as follows in order to enumerate the 3D-trees:

$$\begin{aligned} a(x) = {} & 1 + x + x^2 + 2x^3 + 3x^4 \\ & + 5x^5 + 8x^6 + 14x^7 + 23x^8 \\ & + 41x^9 + 69x^{10} + 122x^{11} + 208x^{12} + 370x^{13} \\ & + 636x^{14} + 1134x^{15} + \cdots \end{aligned} \tag{41}$$

$$\begin{aligned} c\left(x^2\right) = {} & 1 + x^2 + x^4 + 2x^6 + 5x^8 + 11x^{10} + 28x^{12} \\ & + 74x^{14} + 199x^{16} + \cdots \end{aligned} \tag{42}$$

$$\begin{aligned} b(x) = {} & 1 + x + x^2 + 2x^3 + 5x^4 + 11x^5 + 28x^6 + 74x^7 \\ & + 199x^8 + 551x^9 + 1553x^{10} \\ & + 4436x^{11} + 12832x^{12} \\ & + 37496x^{13} + 110500x^{14} + 328092x^{15} + \cdots, \end{aligned} \tag{43}$$

where each coefficient represents α_k, γ_k, or β_k for the generating function shown in Eqs. 32–34.

To illustrate the procedure of calculation, let us consider the case of $m = 3$ for the enumeration of centroidal 3D-trees. Then, the terms up to x^3 are selected from Eqs. 41–43 as follows:

$$a(x) = 1 + x + x^2 + 2x^3 \tag{44}$$

$$c\left(x^2\right) = 1 + x^2 + x^4 + 2x^6 \tag{45}$$

$$b(x) = 1 + x + x^2 + 2x^3 \tag{46}$$

These generating functions are introduced into the functional equations shown in Eqs. 29, 36, 38, and 40. The

resulting equations are expanded to give the following generating functions:

$$\widehat{B}(x) = x + x^2 + 2x^3 + 4x^4 + 5x^5 + 6x^6 + 9x^7 + 8x^8 \cdots \quad (47)$$

$$\widehat{E}(x) = x + x^2 + 2x^3 + 4x^4 + 5x^5 + 6x^6 + 11x^7 + 9x^8 \cdots \quad (48)$$

$$\widehat{A}(x) = x + x^2 + 2x^3 + 4x^4 + 5x^5 + 6x^6 + 7x^7 + 7x^8 \cdots \quad (49)$$

$$\widehat{C}(x) = 2x^7 + x^8 + \cdots \quad (50)$$

Among the coefficients in the right-hand sides, those of the terms x^7 and ^{8} are effective, because the maximum number $m = 3$ is available in the cases of $n = 7$ and $n = 8$. Hence, we obtain $\widehat{B}_7 = 9$ and $\widehat{B}_8 = 8$ from Eq. 47; $\widehat{E}_7 = 11$ and $\widehat{E}_8 = 9$ from Eq. 48; $\widehat{A}_7 = 7$ and $\widehat{A}_8 = 8$ from Eq. 49; as well as $\widehat{C}_7 = 2$ and $\widehat{C}_8 = 1$ from Eq. 50. They are listed in the $n = 7$ and $n = 8$ rows of Table 2. This procedure is repeated by moving m from 0 to 9. Thereby, we obtain the results summarized in Table 2.

4.2 Enumeration of bicentroidal 3D-trees

4.2.1 Formulation

In this subsection, we shall discuss the enumeration of bicentroidal 3D-trees. Let $\widetilde{B}(x)$ be a generating function for enumerating achiral bicentroidal promolecules plus enantiomeric pairs of chiral bicentroidal promolecules up to carbon content n, i.e.

$$\widetilde{B}(x) = \sum_{k=0}^{n} \widetilde{B}_k x^k \quad (51)$$

Our target is to calculate $\widetilde{B}_n$ which satisfies the criterion of bicentroidal 3D-trees. This can be accomplished by using Theorem 6 of Part 3 [27] after the selection of the CI-CF to be adopted. Because Eq. 16 for CI-CF($\mathbf{K}, \$_d$) ignores the central edge only (cf. **4**), the following functional equation is obtained by changing the SIs (a_d, c_d, and b_d) of Eq. 16 into the generating functions ($a(x^d)$, $c(x^d)$, and $b(x^d)$):

$$\widetilde{B}(x) = \frac{1}{4}\left(b(x)^2 + b(x^2) + a(x)^2 + c(x^2)\right). \quad (52)$$

Let $\widetilde{E}(x)$ be a generating function for enumerating achiral bicentroidal promolecules plus chiral bicentroidal promolecules up to carbon content n, where two enantiomers of each pair are counted separately:

$$\widetilde{E}(x) = \sum_{k=0}^{n} \widetilde{E}_k x^k \quad (53)$$

By employing CI-CF($\mathbf{K}_2, b_d$) of Eq. 17 in Theorem 6 of Part 3 [27], the following functional equation is obtained:

$$\widetilde{E}(x) = \frac{1}{2}\left(b(x)^2 + b(x^2)\right). \quad (54)$$

Let $\widetilde{A}(x)$ be a generating function for enumerating achiral bicentroidal promolecules up to carbon content n, i.e.,

$$\widetilde{A}(x) = \sum_{k=0}^{n} \widetilde{A}_k x^k \quad (55)$$

By applying Theorem 6 of Part 3 [27] to CI-CF$_A$($\mathbf{K}, \$_d$) of Eq. 18, the following functional equation is obtained:

$$\widetilde{A}(x) = \frac{1}{2}\left(a(x)^2 + c(x^2)\right). \quad (56)$$

Let $\widetilde{C}(x)$ be a generating function for enumerating chiral bicentroidal promolecules up to a carbon content n, where each pair of enantiomers is counted just once:

$$\widetilde{C}(x) = \sum_{k=0}^{n} \widetilde{A}_k x^k \quad (57)$$

Theorem 6 of Part 3 [27] is applied to CI-CF$_C$($\mathbf{K}, \$_d$) of Eq. 19 so as to give the following functional equation:

$$\widetilde{C}(x) = \frac{1}{4}\left(b(x)^2 + b(x^2) - a(x)^2 - c(x^2)\right). \quad (58)$$

By examining the ligand inventories corresponding to Eqs. 21–23, the functions appearing in the right-hand side of Eq. 52, 54, 56, or 58 are found to indicate that $a(x)$ is a generating function for enumerating achiral proligands (cf. Eq. 21); $c(x)$ is a generating function for enumerating ordered enantiomeric pairs called "diploids" (cf. Eq. 22); and $b(x)$ is a generating function for enumerating achiral proligands and chiral proligands, where two enantiomers of each pair are counted separately (cf. eq. 23). Because the criterion of deciding bicentroidal 3D-trees requires the relation $m = \frac{1}{2}n$, we presume the following generating functions each composed of a single term:

$$a(x) = \alpha_{n/2} x^{n/2} \quad (59)$$

$$c(x) = \gamma_n x^{n/2} \quad (60)$$

$$b(x) = \beta_{n/2} x^{n/2}, \quad (61)$$

where the subscript n of γ_n is selected to be twice of $\alpha_{n/2}$ and $\beta_{n/2}$. Note that these ligand inventories are used on condition that n is tentatively fixed. Because these generating functions are composed of a single term, they

Table 2 The numbers of centroidal 3D-trees or centroidal alkanes

n	m	$\widehat{B}_n$	$\widehat{E}_n$	$\widehat{A}_n$	$\widehat{C}_n$
1	0	1	1	1	0
2	0	0	0	0	0
3	1	1	1	1	0
4	1	1	1	1	0
5	2	3	3	3	0
6	2	2	2	2	0
7	3	9	11	7	2
8	3	8	9	7	1
9	4	38	55	21	17
10	4	46	70	22	24
11	5	203	345	61	142
12	5	283	494	72	211
13	6	1,299	2,412	186	1,113
14	6	2,004	3,788	220	1,784
15	7	9,347	18,127	567	8,780
16	7	15,758	30,799	717	15,041
17	8	72,505	143,255	1,755	70,750
18	8	129,281	256,353	2,209	127,072
19	9	589,612	1,173,770	5,454	584,158
20	9	1,098,656	2,190,163	7,149	1,091,507

The numbers of centroidal 3D-trees are obtained under several conditions, i.e., $\widehat{B}_n$: achiral and chiral centroidal 3D-trees, where a pair of enantiomers is counted just once; $\widehat{E}_n$: achiral and chiral centroidal 3D-trees, where two enantiomers of each pair are counted separately; $\widehat{A}_n$: achiral centroidal 3D-trees; $\widehat{C}_n$: chiral centroidal 3D-trees, where each pair of enantiomers is counted just once

are introduced into Eqs. 52, 54, 56, and 58 to give the following equations:

$$\widetilde{B}(x) = \widetilde{B}_n x^n = \frac{1}{4}\left(\beta_{n/2}^2 + \beta_{n/2} + \alpha_{n/2}^2 + \gamma_n\right)x^n \tag{62}$$

$$\widetilde{E}(x) = \widetilde{E}_n x^n = \frac{1}{2}\left(\beta_{n/2}^2 + \beta_{n/2}\right)x^n \tag{63}$$

$$\widetilde{A}(x) = \widetilde{A}_n x^n = \frac{1}{2}\left(\alpha_{n/2}^2 + \gamma_n\right)x^n \tag{64}$$

$$\widetilde{C}(x) = \widetilde{C}_n x^n = \frac{1}{4}\left(\beta_{n/2}^2 + \beta_{n/2} - \alpha_{n/2}^2 - \gamma_n\right)x^n, \tag{65}$$

where the last term x^n of Eqs. 52, 54, 56, or 58 remains to give each coefficient, i.e., $\widetilde{B}_n$, $\widetilde{E}_n$, $\widetilde{A}_n$, or $\widetilde{C}_n$. Note that these generating functions should be used on condition that n is tentatively fixed as the n in the ligand inventories (Eqs. 59–61) is fixed. Once the generating functions for proligands (Eqs. 59–61) are evaluated, Eqs. 62–65 are easily obtained.

4.2.2 Calculation

As shown in the preceding formulation, the same generating functions for proligands (Eqs. 59–61) are effective to evaluate $\widetilde{B}(x)$ (Eq. 62) $\widetilde{E}(x)$ (Eq. 63), $\widetilde{A}(x)$ (Eq. 64), and $\widetilde{C}(x)$ (Eq. 65) for centroidal 3D-trees. Thus, the coefficients derived by using Eqs. 59–61 are introduced into Eqs. 62–62 to give $\widetilde{B}_n$, $\widetilde{E}_n$, $\widetilde{A}_n$, and $\widetilde{C}_n$.

Alternatively, the following equations are derived from Eqs. 41–43:

$$a(x)^2 = 1 + x^2 + x^4 + 2^2x^6 + 3^2x^8 + 5^2x^{10} + 8^2x^{12} + 14^2x^{14} + 23^2x^{16} + 41^2x^{18} + 69^2x^{20} + \cdots \tag{66}$$

$$c(x^2) = 1 + x^2 + x^4 + 2x^6 + 5x^8 + 11x^{10} + 28x^{12} + 74x^{14} + 199x^{16} + 551x^{18} + 1553x^{20} + \cdots \tag{67}$$

$$b(x)^2 = 1 + x^2 + x^4 + 2^2x^6 + 5^2x^8 + 11^2x^{10} + 28^2x^{12} + 74^2x^{14} + 199^2x^{16} + 551^2x^{18} + 1553^2x^{20} + \cdots \tag{68}$$

$$b(x^2) = 1 + x^2 + x^4 + 2x^6 + 5x^8 + 11x^{10} + 28x^{12} + 74x^{14} + 199x^{16} + 551x^{18} + 1553x^{20} + \cdots \tag{69}$$

which are introduced into Eqs. 52, 54, 56, and 58. Thereby, the coefficient of each x^k $(k = 1, 2, \ldots, n, \ldots)$ represents a respective number to be obtained, i.e., $\widetilde{B}_k$, $\widetilde{E}_k$, $\widetilde{A}_k$, or $\widetilde{C}_k$. The values obtained up to carbon content 20 are collected in Table 3.

4.3 Total features of enumeration of 3D-trees

From the mathematical point of view, the classification into centroidal and bicentroidal 3D-trees is essential to the present combinatorial enumeration. The classification, however, is not easily ascribed to a distinct chemical meaning so that the total number of 3D-trees of each carbon content n is convenient for further discussions from the chemical point of view, as shown in Table 4. The partial numerical values obtained by Robinson et al. [14] by using an alternative method agree with those of Table 4 except several values.

Table 3 The numbers of bicentroidal 3D-trees or bicentroidal alkanes

n	$\widetilde{B}_n$	$\widetilde{E}_n$	$\widetilde{A}_n$	$\widetilde{C}_n$
2	1	1	1	0
4	1	1	1	0
6	3	3	3	0
8	11	15	7	4
10	42	66	18	24
12	226	406	46	180
14	1,455	2,775	135	1,320
16	10,132	19,900	364	9,768
18	76,596	152,076	1,116	75,480
20	604,919	1,206,681	3,157	601,762

The numbers of bicentroidal 3D-trees are obtained under several conditions, i.e., $\widetilde{B}_n$: achiral and chiral bicentroidal 3D-trees, where a pair of enantiomers is counted just once; $\widetilde{E}_n$: achiral and chiral bicentroidal 3D-trees, where two enantiomers of each pair are counted separately; $\widetilde{A}_n$: achiral bicentroidal 3D-trees; $\widetilde{C}_n$: chiral bicentroidal 3D-trees, where each pair of enantiomers is counted just once

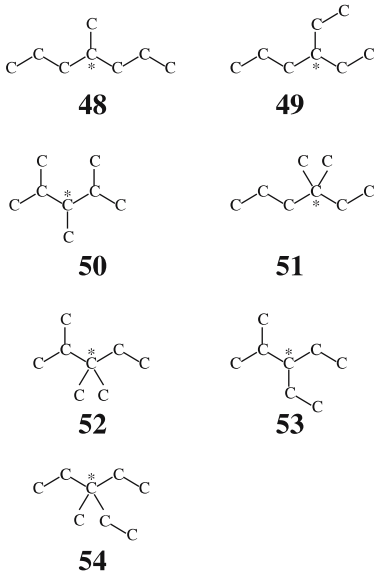
Table 4 The numbers of 3D-trees or alkanes

n	$\widehat{B}_n + \widetilde{B}_n$	$\widehat{E}_n + \widetilde{E}_n$	$\widehat{A}_n + \widetilde{A}_n$	$\widehat{C}_n + \widetilde{C}_n$
1	1	1	1	0
2	1	1	1	0
3	1	1	1	0
4	2	2	2	0
5	3	3	3	0
6	5	5	5	0
7	9	11	7	2
8	19	24	14	5
9	38	55	21	17
10	88	136	40	48
11	203	345	61	142
12	509	900	118	391
13	1,299	2,412	186	1,113
14	3,459	6,563	355	3,104
15	9,347	18,127	567	8,780
16	25,890	50,699	1,081	24,809
17	72,505	143,255	1,755	70,750
18	205,877	408,429	3,325	202,552
19	589,612	1,173,770	5,454	584,158
20	1,703,575	3,396,844	10,306	1,693,269

The numbers of 3D-trees are obtained under several conditions. See Tables 2 and 3. $\widehat{B}_n + \widetilde{B}_n$: achiral and chiral 3D-trees, where a pair of enantiomers is counted just once; $\widehat{E}_n + \widetilde{E}_n$: achiral and chiral 3D-trees, where two enantiomers of each pair are counted separately; $\widehat{A}_n + \widetilde{A}_n$: achiral 3D-trees; $\widehat{C}_n + \widetilde{C}_n$: chiral 3D-trees, where each pair of enantiomers is counted just once

To illustrate the difference between centroidal 3D-trees and bicentroidal ones, let us examine the $n = 8$ rows of Tables 2, 3, 4. Among 19 of 3D-trees ($\widehat{B}_8 + \widetilde{B}_8 = 19$ in Table 4), 8 centroidal 3D-trees ($\widehat{B}_8 = 8$ in Table 2) are depicted in Fig. 5, where they are categorized into 7 achiral 3D-trees (**48**–**54**) and 1 chiral one (**55**) according to the values $\widehat{A}_8 = 7$ and $\widehat{C}_8 = 1$ in Table 2. The 3D-trees are depicted as alkanes with

(achiral 3D-trees or alkanes of centroidal types)



(chiral 3D-trees or alkanes of centroidal types)

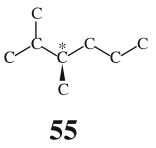
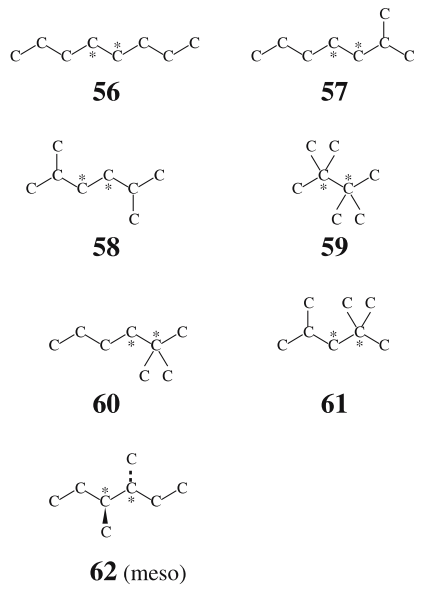


Fig. 5 Stereoisomers for centroidal 3D-trees or alkanes of carbon content 8. A *wedged edge* is used to represent the necessity of showing the configuration of the carbon node, as found in the chiral 3D-tree or alkane (**55**). The *stereoisomer with S-configuration* (**55**) is depicted as a representative. The *other stereoisomers without wedges* are achiral, i.e., **48**–**54**. Each *carbon with an asterisk* is a centroid

carbon content 8 (i.e., C_8H_{18}), where hydrogens at every endpoint are omitted for the sake of simplicity. Following chemical conventions, the configuration of the chiral 3D-tree (**55**) is expressed by a wedge. Each pair of enantiomeric stereoisomers (3D-trees) is counted just once in the present enumeration, so that an appropriate enantiomer (e.g., **55**) for the pair of enantiomeric stereoisomers is depicted as a representative. An asterisk is attached to the centroid of each 3D-tree.

In the present enumeration, the centroidal 3D-trees are ascribed to promolecules of respective types shown in Fig. 3. For example, the achiral 3D-trees **48**–**54** are ascribed to promolecules of the type **14** shown in Fig. 3. The promolecule **14** belongs to $\mathbf{C}_s$-symmetry, where four proligands X_2YZ are selected from achiral alkyl ligands (or hydrogens), e.g., X = –$CH_2CH_2CH_3$, Y = –CH_3, and Z = H for **48**. On the other hand, the chiral 3D-tree **55** is ascribed to a promolecule of the type **22**, which belongs to $\mathbf{C}_1$-symmetry (i.e., asymmetric configuration). The four proligands XYZH of **22** with respect to

(achiral 3D-trees or alkanes of bicentroidal types)



(chiral 3D-trees or alkanes of bicentroidal types)

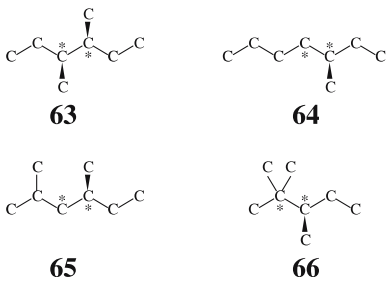


Fig. 6 Stereoisomers for bicentroidal 3D-trees or alkanes of carbon content 8. The seven stereoisomers in the *upper part* are achiral, while the four stereoisomers in the *lower part* are chiral. *Wedged edges and hashed edges* are used to represent the necessity of showing the configuration of the carbon node, whether the 3D-tree or alkane at issue is achiral or chiral. An appropriate enantiomer of each enantiomeric pair is depicted as a representative. *Each pair of adjacent asterisks* represents a bicentroid

55 are selected from achiral alkyl ligands (or hydrogens), i.e., X = –$CH_2CH_2CH_3$, X = –$CH(CH_3)_2$, Y = –CH_3, and Z = H.

As examples of bicentroidal 3D-trees, 11 bicentroidal 3D-trees ($\widetilde{B}_8 = 11$ in Table 3) are depicted in Fig. 6 among 19 of 3D-trees ($\widehat{B}_8 + \widetilde{B}_8 = 19$ in Table 4). They are categorized into 7 achiral 3D-trees (**56–62**) and 4 enantiomeric pairs of chiral ones (**63–66**) in agreement with the values $\widetilde{A}_8 = 7$ and $\widetilde{C}_8 = 4$ in Table 3. Following chemical conventions, the configuration of a carbon point to be specified is expressed by a wedge and/or a hashed line. A pair of asterisks is attached to the centroid of each 3D-tree.

Bicentroidal 3D-trees are ascribed to promolecules of respective types shown in Fig. 4. The achiral 3D-trees **56**, **58**, and **59** are ascribed to promolecules of the type **42** shown in Fig. 4. The promolecule **42** belongs to **K**-symmetry ($\mathbf{K} = \mathbf{D}_{\infty h}/\mathbf{C}_{\infty}$), where two proligands X_2 are selected from achiral alkyl ligands (or hydrogens), e.g., X = –$CH_2CH_2CH_3$ for **56**.

The achiral 3D-trees **57**, **60**, and **61** are ascribed to a promolecule of the type **43**, where we place two achiral ligands of different types, e.g., X = –$CH_2CH_2CH_2CH_3$ and Y = –$CH_2CH(CH_3)_2$ for **57**, according to the $\mathbf{K}_3$-symmetry ($\mathbf{K}_3 = \mathbf{C}_{\infty v}/\mathbf{C}_{\infty}$).

The achiral 3D-tree **62** is ascribed to a promolecule of the type **46**, in which the two proligands are represented as p = –*S*-$CH(CH_3)CH_2CH_3$ and $\overline{\text{p}}$ = –*R*-$CH(CH_3)CH_2CH_3$ so as to exhibit *meso*-character.

The chiral 3D-tree **63** is ascribed to a promolecule of the type **45**, in which the two proligands are represented as p = –*S*-$CH(CH_3)CH_2CH_3$ (or $\overline{\text{p}}$ = –*R*-$CH(CH_3)CH_2CH_3$) so as to exhibit $\mathbf{K}_2$-symmetry ($\mathbf{K}_2 = \mathbf{D}_{\infty}/\mathbf{C}_{\infty}$).

The chiral 3D-trees **64–66** are ascribed to a promolecule of the type **44**, where we place X = –$CH_2CH_2CH_2CH_3$ and Y = –$CH(CH_3)CH_2CH_3$ for **64**, and so on in agreement with the $\mathbf{K}_1$-symmetry ($\mathbf{K}_1 = \mathbf{C}_{\infty}/\mathbf{C}_{\infty}$).

5 Comparison with earlier accomplishments

The present paper is devoted to enumerate alkanes as 3D-structures, not as graphs. Henze and Blair [11] obtained the number of alkanes of a given carbon content where the alkane were regarded as graphs, not as 3D-objects. Pólya [12,13] applied his main theorem (Hauptsatz) to the evaluation of the number of trees as graphs, which were mathematical counterparts of alkanes.

According to Pólya (Eq. 2.50 of Sect. 51 [12,13]), the number of centroidal trees or alkanes as graphs was evaluated by the following functional equation:

$$r(x) = \frac{x}{24}\left(s(x)^4 + 3s\left(x^2\right)^2 + 8s(x)s\left(x^3\right) + 6s(x)^2 s\left(x^2\right) + 6s\left(x^4\right)\right), \tag{70}$$

which can be derived by using the products of dummy variables (PDVs) collected in Table 1. Obviously, Eq. 70 is a special case of Eq. 29 in which we place $s\left(x^d\right) = a\left(x^d\right) = b\left(x^d\right) = c\left(x^d\right)$ in the right-hand side. The results based on Eq. 70 were identical with those of Henze and Blair [11,30].

Robinson et al. [14] reported the enumeration of 3D-trees by using the dissimilarity characteristic equation of Otter [31]. Their treatment, however, did not take explicit account of the concept of sphericity. For example, they used the following equation:

$$\frac{x}{24}\left(s(x)^4 + 3s\left(x^2\right)^2 + 8s(x)s\left(x^3\right) + 6a(x)^2 s\left(x^2\right) + 6s\left(x^4\right)\right) \quad (71)$$

even in the case where Eq. 29 would be required if the concept of sphericity was taken into consideration. Obviously, Eq. 71 is a special case of Eq. 29 in which we place $s\left(x^d\right) = b\left(x^d\right) = c\left(x^d\right)$, so that $b\left(x^d\right)$ and $c\left(x^d\right)$ are mixed up. The mixing-up of $b\left(x^d\right)$ with $c\left(x^d\right)$ indicates that the two modes of transitivity ascribed to $c\left(x^d\right)$ (e.g., p/$\overline{\text{p}}$ and $\overline{\text{p}}$/p for $d = 2$) were replaced by other two modes of transitivity ascribed to $b\left(x^d\right)$ (e.g., p/p and $\overline{\text{p}}$/$\overline{\text{p}}$ for $d = 2$). The usage of Eq. 71 fortunately resulted in the same results as the present ones using Eq. 29, because the molecular formulas of p and $\overline{\text{p}}$ are equal by considering their carbon contents only.

6 Conclusions

Three-dimensional (3D) trees, which are defined as a 3D extension of trees, are enumerated by Fujita's proligand method formulated in Part 1 to Part 3 of this series [25–27]. Such 3D-trees are classified into centroidal and bicentroidal 3D-trees, which correspond to respective promolecules having proligands as substituents. Cycle indices with chirality fittingness (CI-CFs) are formulated to enumerate centroidal and bicentroidal 3D-trees, where they are composed of three kinds of sphericity indices (SIs), i.e., a_d for homospheric cycles, c_d for enantiospheric cycles, and b_d for hemispheric cycles, in order to obtain itemized results with respect to chiral and achiral 3D-trees. The CI-CFs are applied to derive functional equations $\left(a(x), c\left(x^2\right), \text{and } b(x)\right)$, into which the generating functions of planted 3D-trees are introduced. Thereby, the numbers of 3D-trees or equivalently those of alkanes as stereoisomers are calculated and collected up to carbon content 20 in a tabular form.

References

1. Harary F (1969) Graph theory. Addison-Wesley, Reading
2. Hosoya H (1972) Kagaku no Ryoiki 26:989–1001
3. Rouvray DH (1974) Chem Soc Rev 3:355–372
4. Polansky OE (1975) MATCH Commun Math Comput Chem 1:11–31
5. Balaban AT (ed) (1976) Chemical applications of graph theory. Academic, London
6. Biggs NL, Lloyd EK, Wilson RJ (1976) Graph theory 1736–1936. Oxford University Press, Oxford
7. Pólya G, Tarjan RE, Woods DR (1983) Notes on introductory combinatorics. Birkhäuser, Boston
8. Balasubramanian K (1985) Chem Rev 85:599–618
9. Cayley A (1874) Philos Mag 47(4):444–446
10. Cayley A (1875) Rep Br Assoc Adv Sci 45:257–305
11. Henze HR, Blair CM (1931) J Am Chem Soc 53:3077–3085
12. Pólya G (1937) Acta Math 68:145–254
13. Pólya G, Read RC (1987) Combinatorial enumeration of groups, graphs, and chemical compounds. Springer, Berlin Heidelberg New York
14. Robinson RW, Harary F, Balaban AT (1976) Tetrahedron 32:355–361
15. Fujita S (1991) Symmetry and combinatorial enumeration in chemistry. Springer, Berlin Heidelberg New York
16. Fujita S (1989) Theor Chim Acta 76:247–268
17. Fujita S (1990) J Math Chem 5:121–156
18. Fujita S (1990) Bull Chem Soc Jpn 63:203–215
19. Fujita S (2005) MATCH Commun Math Comput Chem 54:251–300
20. Fujita S (2006) MATCH Commun Math Comput Chem 55:5–38
21. Fujita S (2006) MATCH Commun Math Comput Chem 55:237–270
22. Fujita S (1991) Tetrahedron 47:31–46
23. Fujita S (1992) J Chem Inf Comput Sci 32:354–363
24. Fujita S (1993) Polyhedron 12:95–110
25. Fujita S (2005) Theor Chem Acc 113:73–79
26. Fujita S (2005) Theor Chem Acc 113:80–86
27. Fujita S (2006) Theor Chem Acc 115:37–53
28. Jordan C (1869) J Reine Angew Math 70:185–190
29. Fujita S (2000) J Chem Inf Comput Sci 40:426–437
30. Henze HR, Blair CM (1931) J Am Chem Soc 53:3042–3046
31. Otter R (1948) Ann Math 49:583–599