Enumeration of the Staggered Conformers of Alkanes

S. J. Cyvin, J. Brunvoll, B. N. Cyvin, and W. Lüttke

Department of Physical Chemistry, The University of Trondheim, N-7034 Trondheim-NTH, Norway

Institut für Organische Chemie, Universität Göttingen, D-37077 Göttingen, Germany

Z. Naturforsch. 50a, 857–863 (1995); received March 8, 1995

Alkanoid systems are chemical graphs which can be embedded in the diamond lattice. They correspond to the staggered conformers of alkanes. Non-overlapping and overlapping systems are defined. The previously deduced algebraic solution for the numbers of unbranched alkanoids is revisited and reformulated in terms of generating functions. Computerized enumerations for non-overlapping alkanoids, both unbranched and branched, are reported. Some of the smallest overlapping unbranched alkanoids are generated by a method called combinatorial constructions.

Introduction

The traditional enumeration of alkanes (one methane, one ethane, one propane, two butanes, three pentanes, five hexanes, etc.) is a counting of constitutional (or structural) isomers. It started at least 120 years ago with the work of the celebrated mathematician Cayley [1, 2] and has since continued through modern times. So many authors have contributed to this research that only a small selection of references can be given here [1–12]. In sharp contrast to this situation, very little work has been done on the enumeration of conformers (or conformational isomers) of alkanes, although this topic seems clearly to be of prime interest in modern organic and structural chemistry. Whereas there is only one constitutional isomer of any normal alkane, e.g. n-pentane, there are four staggered conformers of n-pentane. These conformers can, at least in principle, be transformed into each other by internal rotation around the central bonds through energy barriers of different heights. Thus it is not surprising that the conformers of this kind possess different physical properties. In particular, they can, at least in theory, be distinguished by spectroscopical methods.

The first spectroscopic observations on rotational isomers are presumably due to Kohlrausch [13, 14], who studied the Raman spectra of monohaloalkanes. Kohlrausch and Köppl [15] included n-butane in their investigations. This molecule was reinvestigated by Szaasz et al. [16], who furnished the first detailed spectroscopical analysis of a simple hydrocarbon including its conformational isomerism. The book of Mizushima [17] is a classical monograph in the realm of internal rotation, while more recent reviews also are available [18, 19].

One single paper by Funck [20] stands as a pioneering work in the enumeration of alkane conformers. This author achieved a complete mathematical solution for the staggered conformers of unbranched alkanes. To wit, Funck's work was ahead of its time and has passed unnoticed among graph theoreticians and mathematical chemists, who later performed similar enumerations [21–24]. In fact, Balaban [23] rediscovered Funck's formulas independently after eighteen years. He formulated their derivation in the framework of graph theory, but his methods are essentially the same as those of Funck. In both cases the symmetry plays an important role. The numbers of alkanes under consideration [20, 23] do not take into account steric hindrances and therefore include many forms which are not expected to be realized chemically.

In the present work we have revisited Funck's formulas [20], cast them in a slightly modernized form, and deduced an alternative representation of the pertinent numbers in terms of generating functions. Furthermore, a computer program was designed in order to generate and enumerate a class of systems which is precisely defined in graph-theoretical terms and referred to as non-overlapping alkanoids. The members of this system represent staggered conformers of alkanes with strain-free carbon skeletons. Finally, the first computerized enumerations of staggered conformers of branched alkanes are reported. For these numbers, only some rudimentary results [23, 25] have been available so far.
Results and Discussion

Definitions and Basic Concepts

It is well known that any staggered alkane conformer can be embedded in a diamond lattice [23, 26]. In this lattice, also the "chair" configuration of cyclohexane and the configuration of adamantane are recognized.

For our purposes a graph may be defined simply as a set of vertices which are connected by edges. Two vertices which are directly connected by an edge, are said to be adjacent. Two edges which have a vertex in common, are said to be incident. The degree of a vertex is the number of other vertices which are adjacent to it.

We define an alkanoid system (or shortly an alkanoid) as a connected set of edges without any cycle, which can be embedded in a diamond lattice. Here an alkanoid is a chemical graph [27], which falls under the category of angle-restricted trees; the angle between any pair of incident edges is tetrahedral. An alkanoid is non-overlapping if no pair of vertices in it overlaps; otherwise it is overlapping.

An alkanoid (system) is clearly a model of the carbon skeleton of a staggered alkane conformer and represents as such this \( \text{C}_n\text{H}_{2n+2} \) conformer. The vertices correspond to the carbon atoms, the edges to the carbon-carbon bonds. The vertices in an alkanoid may clearly have the degrees one, two, three and four, corresponding to primary, secondary, tertiary and quaternary carbons, respectively. An alkanoid (or the corresponding alkane) is said to be unbranched if all its vertices are of degree one or two; it is branched if it has at least one vertex of degree three or four.

The numbers of vertices and edges in an alkanoid are identified by the symbols \( n \) and \( m \), respectively. Then \( n=m+1 \), a general relation for trees in graph theory. Here it applies to both unbranched and branched alkanoids.

Two alkanoids are said to be isomorphic (equivalent, identical) if they can be superimposed on each other by a translation, occasionally combined with any symmetry operation. Here it should be noted that the symmetry of overlapping alkanoids is defined with reference to idealized objects where the overlapping does not cause any distortion. Thus, for instance, the coiled alkanoids which correspond to \( n \)-alkanes and are embedded in the (chair) cyclohexane lattice, belong either to the symmetry group \( C_2 \) or \( C_2h \), depending on whether \( n \) is odd or even, respectively.

Algebraic Solutions for Unbranched Systems

The above definitions assure that the number of non-isomorphic alkanoids is identical to the number of staggered alkane conformers. It is noted in particular that enantiomeric pairs are not counted double. When the unbranched systems are concerned, the non-overlapping and overlapping alkanoids taken together correspond exactly to the conformers considered by Funck [20] and which he called spectral isomers.

An enumeration method which essentially is the same as the one of Funck [20], and where the symmetry is exploited, is called "stupid sheep counting" [28]. Here "stupid" refers to the counting, not to the sheep. This name was coined under reference to Redelmeier [29], who says: "There is a well known way to count cattle in a herd: count the number of legs and divide by four."

In the application of stupid sheep counting to unbranched alkanoids, the "crude total" (corresponding to the number of sheep legs) is

\[
R_m = 3^{m-2} \quad (m > 1), \quad R_1 = 1. \hspace{1cm} (1)
\]

Here \( m \) (number of edges) is used as the leading parameter; \( m=1 \) pertains to ethane. The numbers (1) count systems belonging to different symmetry groups a certain numbers of times according to

\[
R_m = Z_m(C_{2h}/D_{2h}) + Z_m(C_{2v}) + 2Z_m(C_2) + 2Z_m(C_3) + 4Z_m(C_1). \hspace{1cm} (2)
\]

Here \( Z_m(G) \) is the number of systems belonging to the symmetry group \( G \); in particular, \( D_{2h} \) is unique for \( m=1 \) as the symmetry of the ethane skeleton. The total number of systems is simply

\[
Z_m = Z_m(C_{2h}/D_{2h}) + Z_m(C_{2v}) + Z_m(C_2) + Z_m(C_3) + Z_m(C_1). \hspace{1cm} (3)
\]

Now the relatively simple formulas for the symmetrical systems, which are consistent with those of Funck [20], are given below.

\[
Z_m(C_{2h}/D_{2h}) = \frac{1}{2} [1 - (-1)^m], \hspace{1cm} (4)
\]

\[
Z_m(C_{2v}) = \frac{1}{2} [1 + (-1)^m], \hspace{1cm} (5)
\]

\[
Z_m(C_2) = \frac{1}{2} 3^{(m-1)/2} - \frac{1}{2}, \hspace{1cm} (6)
\]
Here the "floor" function is employed; $\lceil x \rceil$ is the largest integer smaller than or equal to $x$. On eliminating $Z_m(C_i)$ from (2) and (3) one obtains

$$Z_m = \frac{1}{4} [R_m + 3 Z_m(C_{2h}/D_{\infty h}) + 3 Z_m(C_{2v}) + 2 Z_m(C_2) + 2 Z_m(C_i) + 2 Z_m(C_s)] ,$$

(8)

where the division by four appears as in the sheep counting. On inserting into (8) from (1) and (4)–(7) one arrives at the result

$$Z_m = \frac{1}{4} (3^{m-2} + 1) + 3^{(m-3)/2} (m = 3, 5, 7, \ldots),$$

(9)

when $m$ is odd, and

$$Z_m = \frac{1}{4} (3^{m-2} + 1)^2 (m = 2, 4, 6, \ldots)$$

(10)

when $m$ is even. A similar result for $Z_m(C_1)$ can also be obtained from the above equations.

Generating functions represent a powerful tool in different enumeration problems [6, 10, 22, 28, 30, 31]. The application to the present problem of unbranched alkanoids is outlined in the following.

The elementary geometrical series

$$(1 - 3 x)^{ -1} = 1 + 3 x + 9 x^2 + 27 x^3 + \ldots$$

(11)

gives readily the generating function $R(x)$ for the crude total (1) as

$$R(x) = x + x^2 (1 - 3 x)^{ -1} = x (1 - 2 x) (1 - 3 x)^{ -1} .$$

(12)

The appropriate functions $Z(C_{2h}/D_{\infty h}; x)$ and $Z(C_{2v}; x)$ corresponding to (4) and (5), respectively, are also straightforward. From

$$(1 - x)^{ -1} = 1 + x + x^2 + x^3 + \ldots$$

(13)

one obtains

$$Z(C_{2h}/D_{\infty h}; x) + Z(C_{2v}; x) = x (1 - x^2)^{ -1} + x^2 (1 - x^2)^{ -1} = x (1 - x)^{ -1} = x + x^2 + x^3 + x^4 + \ldots$$

(14)

In order to deduce the generating functions which correspond to (5)–(7), start with the auxiliary function

$$S(x) = (1 - x)^{ -1} (1 - 3 x)^{ -1} = 1 + 4 x + 13 x^2 + 40 x^3 + \ldots,$$

(15)

where it is observed that the first differences of the coefficients form a geometrical series. Now,

$$Z(C_2; x) = x^3 (1 + x) S(x^2) = x^3 (1 - x)^{ -1} (1 - 3 x^2)^{ -1} = x^3 + x^4 + 4 x^5 + 4 x^6 + 13 x^7 + 13 x^8 + 40 x^9 + \ldots .$$

(16)

Furthermore,

$$Z(C_1; x) = x^4 S(x^2) = x^4 + 4 x^6 + 13 x^8 + 40 x^{10} + \ldots$$

(17)

and

$$Z(C_i; x) = x^5 S(x^2) = x^5 + 4 x^7 + 13 x^9 + 40 x^{11} + \ldots$$

(18)

A slight simplification is achieved when taking the sum of the two functions (17) and (18):

$$Z(C_i; x) + Z(C_1; x) = x^4 (1 - x^2)^{ -1} (1 - 3 x^2)^{ -1} + x^5 (1 - x^2)^{ -1} (1 - 3 x^2)^{ -1} = x^4 (1 - x)^{ -1} (1 - 3 x^2)^{ -1} = x^4 + x^5 + 4 x^6 + 4 x^7 + 13 x^8 + 13 x^9 + 40 x^{10} + \ldots$$

(19)

A relation analogous to (8) is valid for the appropriate generating functions. On inserting from (12), (14), (16) and (18) we finally arrived at the generating function $Z(x)$ for the $Z_m$ numbers of unbranched alkanoids:

$$Z(x) = \sum_{m=1}^{\infty} Z_m x^m = x (1 - 3 x - 2 x^2 + 8 x^3 - 3 x^4) (1 - x)^{ -1} (1 - 3 x)^{ -1} (1 - 3 x^2)^{ -1} = x + x^2 + 2 x^3 + 4 x^4 + 10 x^5 + 25 x^6 + 70 x^7 + 196 x^8 + 574 x^9 + 1681 x^{10} + \ldots$$

(20)

**Computer Programming for Unbranched Systems**

A computer program was based on a fundamental principle as formulated below.

**Proposition:** All alkanoids with $m$ edges, where $m > 1$, are generated by adding an edge in all possible positions, one at a time, to all alkanoids with $m - 1$ edges.

This principle is obviously valid for the non-overlapping alkanoids separately, as well as for the non-overlapping and overlapping alkanoids taken together. However, the principle is not valid for the overlapping alkanoids separately.

An analogous principle is valid for the benzenoids [32]. In fact, the present program is very much analogous to a program for the generation and enumeration of benzenoids [33, 34]. Yet it had to be more
complicated since the coordinates for positions in a three-dimensional lattice (viz. the diamond lattice) had to be introduced, in contrast to the hexagonal lattice for benzenoids, which is planar and therefore requires coordinates in two dimensions only. In both cases, for alkanoids as well as for benzenoids, the symmetry plays an important role. The symmetry group for a generated system was determined by inspecting the positions of the system after a set of geometrical transformations, which may or may not be symmetry operations for this particular system. The same transformations are also employed in order to assure that only the non-isomorphic systems are counted. A modification of the program was designed so as to generate and enumerate the non-overlapping unbranched alkanoids separately.

This is not the place to give more detailed descriptions of the program algorithms, so we give at once the computational results. Table 1 shows the numbers of non-overlapping unbranched alkanoid systems, corresponding to the staggered conformers of unbranched alkanes.

The numbers in Table 1 are seen to coincide with those of Funck [20] (and of Balaban [23]) up to \(m = 5\) (\(C_6H_{14}\)). Up to this \(m\) value there are no overlapping alkanoids. On subtracting the numbers of Table 1 from the formula numbers (cf. Section on algebraic solutions above) one obtains the numbers of overlapping unbranched alkanoids, which are entered in Table 2.

**Combinatorial Constructions for Unbranched Systems**

Both Funck [20] and Balaban [23] depicted some of the smallest alkanes which they enumerated. This can be done more or less systematically and is of course practicable only up to a limited number of systems. We shall not repeat such general depictions here but concentrate upon the overlapping unbranched alkanoids. It appears that a systematic method, which has been referred to as combinatorial constructions and applied to the generation of helicenes [35, 36], can straightforwardly be adapted for the present purpose. The below definition is analogous to the definition of irreducible helicenes.

**Definition:** An overlapping alkanoid is *irreducible* if no smaller overlapping alkanoid can be produced from it by deleting one edge. Otherwise it is *reducible*.

It is clear that an irreducible overlapping alkanoid must be unbranched and possess exactly one pair of overlapping vertices. The system is associated with a cycle which is broken but would be restored if the two overlapping vertices coalesce. Cycles on the diamond lattice have been described and depicted by Saunders [37]. Notice that one cycle often may be associated with more than one irreducible overlapping alkanoid. All reducible overlapping alkanoids can clearly be generated by adding edges to the irreducible systems.

The smallest overlapping alkanoid is a unique irreducible system with \(m = 6\), which is associated with the cyclohexane cycle; see Figure 1. Let it be denoted by \(U_1\). The overlapping unbranched alkanoids with \(m = 7\) are generated by adding one edge to one of the (symmetrically equivalent) ends of \(U_1\). The three possibilities are shown in Fig. 2, where the coiled system (at the bottom) is designated \(U_2\). The pertinent symmetry groups are indicated in the figure. It is still practicable to display in full all the fourteen overlapping unbranched alkanoids with \(m = 8\) as in Fig. 3, where we at this stage disregard the numerals. Here the first twelve systems are generated by additions of one edge at a time to the three systems of Figure 2. The coiled \(m = 8\) system is identified by the symbol

---

**Table 1. Numbers of non-overlapping unbranched alkanoids (staggered conformers of \(n\)-alkanes).**

<table>
<thead>
<tr>
<th>Formula</th>
<th>(C_{2h}/D_{xh})</th>
<th>(C_{2v})</th>
<th>(C_s)</th>
<th>(C_1)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(C_7)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(C_8)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(C_9)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(C_{10})</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(C_{11})</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

**Table 2. Numbers of overlapping unbranched alkanoids (staggered conformers of \(n\)-alkanes).**

<table>
<thead>
<tr>
<th>Formula</th>
<th>(C_2)</th>
<th>(C_s)</th>
<th>(C_1)</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>(C_6)</td>
<td>1</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(C_7)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(C_8)</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(C_{10})</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(C_{11})</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(C_{12})</td>
<td>0</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

* \(D_{xh}\) only for \(m = 1\).*
In addition, one has two irreducible systems ($U_4$ and $U_5$), both associated with the eight-membered cycle on the diamond lattice. This cycle is the perimeter of two cyclohexane rings which share two edges, and it is a fragment of adamantane. A real demonstration of the method of combinatorial constructions starts below for the $m = 9$ systems.

**Case A:** Additions to $U_1$.

**Subcase A1:** There should be no $U_2$ as a subgraph. This subcase splits further into Subcases A1.1 and A1.2 where three edges are added to the same end of $U_1$, or two and one edge are added to the two ends of $U_1$, respectively. The $6 \times 3$ possibilities under Subcase A1.1 are mapped in the two top rows of Figure 3; here the numerals indicate the numbers of possibilities for adding one edge at a time to the given $m = 8$ systems. The possibilities under Subcase A1.2 are mapped in the third row from top in Figure 3. Symmetry is taken into account for additions to the two $m = 8$ systems of symmetry $C_s$.

**Subcase A2:** Additions to $U_2$ (which itself is an addition to $U_1$), while there should be no $U_3$. Again two further subcases are distinguished: Subcase A2.1, where two edges are added to the same end of $U_2$, and Subcase A2.2, where one edge is added to each of the two ends of $U_2$. The $2 \times 3$ possibilities in the former case and $2 + 1$ in the latter are mapped in Figure 3.

**Subcase A3:** Additions to $U_3$. There are three possibilities as indicated in Figure 3. They include the coiled $m = 9$ system on the cyclohexane lattice.

**Case B:** Additions to $U_4$ or $U_5$. The $2 + 3$ possibilities emerge as indicated (Fig. 3) when the symmetry is properly taken into account. These possibilities include the coiled systems on the eight-membered ring lattice.

There are no more cases for $m = 9$; new irreducible overlapping alkanoids do not occur before $m = 10$. Therefore the total number of overlapping unbranched alkanoids is obtained on adding the numerals of Figure 3.

**Conclusion:** By the combinatorial constructions we arrived at the numbers 1, 3, 14 and 47 for overlapping unbranched alkanoids. These results are in perfect agreement with the corresponding numbers of Table 2. The pertinent symmetry distributions in this table were also confirmed by inspection of the symmetries of the constructed systems. This test of internal consistency between numbers from the algebraic solution, computerized enumeration and a paper-and-pencil method (combinatorial constructions) increase our confidence in the correctness of the present analyses.

**Branched Systems**

The present computer program in its original form (cf. the section on computer programming above) generates and enumerates all non-overlapping alkanoids (branched and unbranched). The branched systems were classified according to the $q$ number of vertices of degree four (corresponding to the quaternary carbon atoms). Numerical values are shown in Table 3.

We shall not go into further details here, especially because there is much more to be done in this area. Some of these aspects are outlined in the below conclusion.
Fig. 3. The 14 overlapping unbranched alkanoids with $m = 8$. The numerals indicate generations of $m = 9$ systems as explained in the text.

Table 3. Numbers of non-overlapping alkanoids (staggered conformers of alkanes).

<table>
<thead>
<tr>
<th>$m$</th>
<th>Formula</th>
<th>Unbranched</th>
<th>Branched (br.)</th>
<th>Grand total</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td></td>
<td>$q=0$</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>1</td>
<td>$C_2H_6$</td>
<td>1</td>
<td></td>
<td></td>
</tr>
<tr>
<td>2</td>
<td>$C_3H_8$</td>
<td>$1^a$</td>
<td>$1^b,c$</td>
<td>$1^b,c$</td>
</tr>
<tr>
<td></td>
<td>$C_4H_{10}$</td>
<td>$2^a$</td>
<td></td>
<td>$1^b,c$</td>
</tr>
<tr>
<td>3</td>
<td>$C_5H_{12}$</td>
<td>$4^a$</td>
<td>$2^b,c$</td>
<td>$1^b,c$</td>
</tr>
<tr>
<td>4</td>
<td>$C_6H_{14}$</td>
<td>$10^a$</td>
<td>$13^b,c$</td>
<td>$1^b,c$</td>
</tr>
<tr>
<td>5</td>
<td>$C_7H_{16}$</td>
<td>24</td>
<td>$61^b,c$</td>
<td>$7^c$</td>
</tr>
<tr>
<td>6</td>
<td>$C_8H_{18}$</td>
<td>67</td>
<td>321</td>
<td>36</td>
</tr>
<tr>
<td>7</td>
<td>$C_9H_{20}$</td>
<td>182</td>
<td>1612</td>
<td>231</td>
</tr>
<tr>
<td>8</td>
<td></td>
<td></td>
<td></td>
<td>3</td>
</tr>
</tbody>
</table>

Conclusion

In the above section on algebraic solutions the remarkable formulas of Funck [20] are revisited. They represent a complete algebraic solution for the number of unbranched staggered alkane conformers or unbranched alkanoids. Balaban [23, 25] pointed out an interesting one-to-one correspondence between these systems and the unbranched catafusenes [21], which represent a certain class of polycyclic aromatic hydrocarbons with exclusively six-membered rings. The two enumeration problems are analogous or basically identical.

A complete mathematical solution for the numbers of all catafusenes (branched and unbranched taken together) has been deduced by Harary and Read [30] and presented as a generating function. This function is considerably more complicated than $Z(x)$ of (20). Now it is tempting to investigate whether the same methods could be used to deduce an algebraic solution for all alkanoids (branched and unbranched). It seems that this might be possible, but it will not be easy, as we have learnt from some preliminary investigations. For example, while the branched catafusenes are distributed among six symmetry groups ($D_{3h}, C_{3h}, D_{2h}, C_{2v}, C_2, C_1$), we have ascertained that sixteen symmetry groups are possible for the branched alkanoids, viz.: $T_d, T, D_{3d}, D_3, S_6, C_{3v}, C_3, D_{2d}, D_2, S_4, C_{2h}, C_{2v}, C_2, C_1, C_1$. 