Alkanes and Alkyls: their Structural Isomers, Clanes and Cycle Indices

M.Yu.Kornilov, Kiev National Taras Shevchenko University

kornilovster@gmail.com

Alkanes are investigated in every detail, that any original news about them seems more and more sophisticated and complicated [1]. Newertheless this area of organic chemistry remains quite mysterious and attractive. Some bases of the enumeration of the isomer numbers have been included into the course books for students of Kiev National Taras Shevchenko University [2 – 3]. We are suggesting here two new aspects of the alkanes structure and topology, discussing them together with alkyls.

I. The **first aspect** is considering alkane isomers C_nH_{2n+2} from the point of the content of primary A, secondary B, tertiary C and quaternary carbon atoms D in their molecules, designating their common formulae as $A_aB_bC_cD_d$, where $a + b + c + d = n$; $3a + 2b + c = 2n + 2$ and calling them as **clanes** of the corresponding alkane homologues (Table 1). Similar data could be made for alkyls (Table 2).

Table 1. Alkane homologue clanes

a	*c*	*d*	Number of carbon atoms *n* in alkane clanes $A_aB_bC_cD_d$								
			2	3	4	5	6	7	8	9	10
2	0	0	A_2	A_2B	A_2B_2	A_2B_3	A_2B_4	A_2B_5	A_2B_6	A_2B_7	A_2B_8
3	1	0			A_3C	A_3BC	A_3B_2C	A_3B_3C	A_3B_4C	A_3B_5C	A_3B_6C
4	2	0					A_4C_2	A_4BC_2	$A_4B_2C_2$	$A_4B_3C_2$	$A_4B_4C_2$
5	3	0							A_5C_3	A_5BC_3	$A_5B_2C_3$
6	4	0									A_6C_4
4	0	1				A_4D	A_4BD	A_4B_2D	A_4B_3D	A_4B_4D	A_4B_5D
5	1	1						A_5CD	A_5BCD	A_5B_2CD	A_5B_3CD
6	2	1								A_6C_2D	A_6BC_2D
6	0	2							A_6D_2	A_6BD_2	$A_6B_2D_2$
7	1	2									A_7CD_2
Clanes total			1	1	2	3	4	5	7	8	10
Isomers total			1	1	2	3	5	9	18	35	75

Table 2. Alkyl homologue clanes

a	*c*	*d*	Number of carbon atoms *n* in alkyl clanes $A_aB_bC_cD_d$									
			1	2	3	4	5	6	7	8	9	10
1	0	0	A	AB	AB_2	AB_3	AB_4	AB_5	AB_6	AB_7	AB_8	AB_9
2	1	0			A_2C	A_2BC	A_2B_2C	A_2B_3C	A_2B_4C	A_2B_5C	A_2B_6C	A_2B_7C
3	0	1				A_3D	A_3BD	A_3B_2D	A_3B_3D	A_3B_4D	A_3B_5D	A_3B_6D
3	2	0					A_3C_2	A_3BC_2	$A_3B_2C_2$	$A_3B_3C_2$	$A_3B_4C_2$	$A_3B_5C_2$
4	1	1						A_4CD	A_4BCD	A_4B_2CD	A_4B_3CD	A_4B_4CD
4	3	0							A_4C_3	A_4BC_3	$A_4B_2C_3$	$A_4B_3C_3$
5	0	2							A_5D_2	A_5BD_2	$A_5B_2D_2$	$A_5B_3D_2$
5	2	1								A_5C_2D	A_5BC_2D	$A_5B_2C_2D$
5	4	0									A_5C_4	A_5BC_4
6	1	2									A_6CD_2	A_6BCD_2
6	3	1										A_6C_3D
6	5	0										A_6C_5
7	0	3										A_7D_3
Clanes total			1	1	2	3	4	5	7	8	10	13

Isomers total	1	1	2	4	8	17	39	89	211	507

Note: The only variable is the *b* (the number of B carbon atoms type) in each row of Tables 1 and 2.

The figures in Table 1 are associated with the distributions of figures *n* into *k* integer positive summands *x*, *y*, *z*,... Alkanes correspond to the case, when $k = 3$ and $a + b + c + d = x + y + z + 2 = n$, e.g. number 8 could be presented as 10 such distributions. Decanes $C_{10}H_{22}$ with 10 ($n = 8 + 2$) carbon atoms in their molecules could be distributed in the same number of clanes: A_2B_8, A_3B_6C, $A_4B_4C_2$, $A_5B_2C_3$, A_6C_4, A_4B_5D, A_5B_3CD, A_6BC_2D, $A_6B_2D_2$, A_7CD_2 (Table 1). Each alkyl could be represented as a subtraction of one H from the alkane producing a lower unit: A → B, B → C, C → D. For example, the alkane A_2B_8 minus A is the alkyl AB_9, whereas the alkane A_2B_8 minus B is the alkyl A_2B_7C, etc. (Table 2).

The most interesting cases arise when methylene groups, i.e. their notations B, are absent in the structure (e.g. A_6C_4, A_7D_3). They correspond to the most overloaded molecules of alkanes and alkyls that in some cases cannot exist owing to the abnormally long lengths of C–C bonds, e.g. in $C[(CH_3)_3C]_4$.

II. The **second aspect** concerning alkanes and alkyls is being discussed here for the first timey. It is the ability of their hydrogen atoms to be substituted by any monovalent atom, e.g. Cl, Na etc., or any monovalent group, e.g. COOH, OCH_3 etc., except for alkyl. Let us assume that there is no limitation of the number of such substituents or their mutual arrangements. We can also allow for free rotation around all of the C–C bonds after a substitution.

Any exhaustive substitution of H-atoms in alkanes can be described by the so called cycle indices which take into account all automorphisms which could be done by symmetry operations of the basic structure and a cycle index expansion into a counting series. Thus, the basic structure of the methyl group has 6 symmetry operations that correspond to one E (identity operation) (term f_1^3), three C_2 axes (term $3f_1^2f_2^1$) and two C_3 axes (term $2f_3^1$). This fact can be described by the corresponding cycle index $Z_{CH_3}(1)$ with the denominator which equals to the sum of the coefficients in the numerator (the meaning of f_m^n is given in [2]):

$$Z_{CH_3} = \frac{f_1^3 + 3f_1^2f_2^1 + 2f_3^1}{6} = T \qquad (1)$$

Then, the basic structure of methylene group has one E (term f_1^2) and one C_2 axis (term f_2^1). This can be presented by the cycle index (2).

$$Z_{CH_2} = \frac{f_1^2 + f_2^1}{2} = D \qquad (2)$$

The basic structure CH group has one E symmetry operation, that can be shown by the cycle index (3).

$$Z_{CH} = f_1^1 = P \qquad (3)$$

The basic structure with tetragonal C-atom, having 24 symmetry operations, refers to the cycle index (4):

$$Z_C = \frac{f_1^4 + 6f_1^2f_2^1 + 3f_2^2 + 8f_1^1f_3^1 + 6f_4^1}{24} = Q \qquad (4)$$

In some cases design formulae looks like recurrent dependence, e.g.

$$Z_{alkanol} = \frac{C^3(x) + 3C(x)C(x^2) + 2C(x^3)}{6}x = C(x) \qquad (5)$$

$$Z_{alkane} = \frac{C^4(x) + 6C^2(x)C(x^2) + 3C^2(x^2) + 8C(x)C(x^3) + 6C(x^4)}{24}x + C(x) - \frac{C^2(x) - C(x^2)}{2} \qquad (6)$$

(equation (6) is taken from R.C.Read [3]).

Cycle indices of any alkane derivatives can be described as a combination of basic expressions (1) – (4). To enumerate alkanes and their derivatives their cycle indices are transfered into the counting series (7):

$$1 + a_1x + a_2x^2 + a_3x^3 + ...+ a_ix^i +... \qquad (7)$$

by the simplest substitution $(1 + x^m)^n$ instead of f_m^n, where a_i is equal to the corresponding substituted basic structure isomer numbers. For alkanes it looks like

$$1 + x + x^2 + x^3 + 2x^4 + 3x^5 + 5x^6 + 9x^7 + 18x^8 + 35x^9 + 75x^{10} + ... \qquad (8)$$

Alkanes C_nH_{2n+2}										
n	1	2	3	4	5	6	7	8	9	10
Alkane	Methane	Ethane	Propane	Butane	Pentane	Hexane	Heptane	Octanes	Nonanes	Decanes
Numbers of alkane isomers										
C_nH_{2n+2}	1	1	1	2	3	5	9	18	35	75

The coefficients a_i correspond to isomer numbers with any equal substituents (F, Cl, COOH, OCH_3 etc., except of alkyls). For two or more types of substituents the substitution $(1 + x^m + y^m +...)^n$ has to be used, giving rise the series (9):

$$1 + a_1x + a_2x^2 + a_3x^3 + ... + a_1y + a_2y^2 + a_3y^3 + ...+ b_1xy + b_2xy^2 + b_2x^2y + ... \qquad (9)$$

If alkane fragments are connected in a continuous flexible chain, its cycle index equals to the product of cycle indices of methyl and methylene groups, e.g. ethyl $CH_3–CH_2–$ has the cycle index (10):

$$Z_{CH_3CH_2} = \left(\frac{f_1^3 + 3f_1^2f_2^1 + 2f_3^1}{6}\right) \times \left(\frac{f_1^2 + f_2^1}{2}\right) = T \times D \qquad (10)$$

Cycle indices of normal (nonbranched) alkyls $CH_3–(CH_2)_n–$ could be presented as (11):

$$Z_{CH_3(CH_2)_n} = \left(\frac{f_1^3 + 3f_1^2f_2^1 + 2f_3^1}{6}\right) \times \left(\frac{f_1^2 + f_2^1}{2}\right)^n = T \times D^n \qquad (11)$$

The situation with *branched* alkyls changes completely because of (1) more number of methyl groups present, and (2) C and D types of carbon atoms appear, and (3) their symmetry group changes as well. If the latter is C_2, C_3 (also T_h), then the cycle index is transfered into so called the Kranz (*germ*. for crown) product of the components V and W, that we notify as $V\{W\}$. For example, 1-methylethyl (isopropyl) $(CH_3)_2CH–$ and 2,2-dimethylethyl (*tert*-butyl) $(CH_3)_3C–$ have (12) and (13) cycle indices, correspondingly.

$$Z_{1\text{-methylethyl}} = \frac{\left[\frac{f_1^3+3f_1^1f_2^1+2f_3^1}{6}\right]^2 + \left[\frac{f_2^3+3f_2^1f_4^1+2f_6^1}{6}\right]}{2} f_1^1 = D\{T\} \times P \qquad (12)$$

$$Z_{2,2\text{-dimethylethyl}} = \frac{\left[\frac{f_1^3+3f_1^1f_2^1+2f_3^1}{6}\right]^3 + 3\left[\frac{f_1^3+3f_1^1f_2^1+2f_3^1}{6}\right]\left[\frac{f_2^3+3f_2^1f_4^1+2f_6^1}{6}\right] + 2\left[\frac{f_3^3+3f_3^1f_6^1+2f_9^1}{6}\right]}{6} \qquad (13)$$

$$= T\{T\}$$

The cycle indices and counting series of the simplest alkanes and alkyls are given in the Tables 3 and 4.

Table 3. The simplest alkanes with their cycle indices and counting series

IUPAC name	Cycle index	Counting series
methane	Q	$1 + x + x^2 + x^3 + x^4$

ethane	D{T}	$1 + x + 2x^2 + 2x^3 + 2x^4 + x^5 + x^6$
propane	D{T} × D	$1 + 2x + 4x^2 + 5x^3 + 6x^4 + 5x^5 + 4x^6 + 2x^5 + x^6$
butane	D{T × D}	$1 + 2x + 6x^2 + 9x^3 + 14x^4 + 14x^5 + 14x^6 + 9x^7 + 6x^8 + 2x^9 + x^{10}$
2-methylpropane	T{T} × P	$1 + 2x + 3x^2 + 5x^3 + 6x^4 + 6x^5 + 6x^6 + 5x^7 + 3x^8 + 2x^9 + x^{10}$
pentane	D {T × D} × D	$1 + 3x + 9x^2 + 17x^3 + 29x^4 + 37x^5 + 42x^6 + 37x^7 + 29x^8 + 17x^9 + 9x^{10} + 3x^{11} + x^{12}$
2-methylbutane	D {T} × T × D × P	$1 + 4x + 10x^2 + 19x^3 + 29x^4 + 37x^5 + 40x^6 + 37x^7 + 29x^8 + 19x^9 + 10x^{10} + 4x^{11} + x^{12}$
2,2-dimethylpropane	Q {T}	$1 + x + 2x^2 + 3x^3 + 4x^4 + 4x^5 + 5x^6 + 4x^7 + 4x^8 + 3x^9 + 2x^{10} + x^{11} + x^{12}$

Table 4. The simplest alkyls with their cycle indices and counting series

IUPAC name	Cycle index	Counting series
methyl	T	$1 + x + x^2 + x^3$
ethyl	T × D	$1 + 2x + 3x^2 + 3x^3 + 2x^4 + x^5$
propyl	T × D^2	$1 + 3x + 6x^2 + 8x^3 + 8x^4 + 6x^5 + 3x^6 + x^7$
1-methylethyl	D {T} × P	$1 + 2x + 3x^2 + 4x^3 + 4x^4 + 3x^5 + 2x^6 + x^7$
butyl	T × D^3	$1 + 4x + 10x^2 + 17x^3 + 22x^4 + 22x^5 + 17x^6 + 10x^7 + 4x^8 + x^9$
1-methylpropyl	T^2 × D × P	$1 + 4x + 9x^2 + 15x^3 + 19x^4 + 19x^5 + 15x^6 + 9x^7 + 4x^8 + x^9$
2-methylpropyl	D {T} × D × P	$1 + 3x + 6x^2 + 9x^3 + 11x^4 + 11x^5 + 9x^6 + 6x^7 + 3x^8 + x^9$
2,2-dimethylethyl	T {T}	$1 + x + 2x^2 + 3x^3 + 3x^4 + 3x^5 + 3x^6 + 2x^7 + x^8 + x^9$
pentyl	T × D^4	$1 + 5x + 15x^2 + 31x^3 + 49x^4 + 61x^5 + 61x^6 + 49x^7 + 31x^8 + 15x^9 + 5x^{10} + x^{11}$
1-methylbutyl	T^2 × D^2 × P	$1 + 5x + 14x^2 + 28x^3 + 43x^4 + 53x^5 + 53x^6 + 43x^7 + 28x^8 + 14x^9 + 5x^{10} + x^{11}$
2-methylbutyl	T^2 × D^2 × P	$1 + 5x + 14x^2 + 28x^3 + 43x^4 + 53x^5 + 53x^6 + 43x^7 + 28x^8 + 14x^9 + 5x^{10} + x^{11}$
3-methylbutyl	D {T} × D^2 × P	$1 + 4x + 10x^2 + 18x^3 + 26x^4 + 31x^5 + 31x^6 + 26x^7 + 18x^8 + 10x^9 + 4x^{10} + x^{11}$
1,1-dimethylpropyl	D {T} × T × D	$1 + 3x + 7x^2 + 12x^3 + 17x^4 + 20x^5 + 20x^6 + 17x^7 + 12x^8 + 7x^9 + 3x^{10} + x^{11}$
1,2-dimethylpropyl	D {T} × T × P^2	$1 + 4x + 9x^2 + 16x^3 + 23x^4 + 27x^5 + 27x^6 + 23x^7 + 16x^8 + 9x^9 + 4x^{10} + x^{11}$
2,2-dimethylpropyl	T {T} × D	$1 + 2x + 4x^2 + 6x^3 + 8x^4 + 9x^5 + 9x^6 + 8x^7 + 6x^8 + 4x^9 + 2x^{10} + x^{11}$
1-ethylpropyl	D {T × D} × P	$1 + 3x + 8x^2 + 15x^3 + 23x^4 + 28x^5 + 28x^6 + 23x^7 + 15x^8 + 8x^9 + 3x^{10} + x^{11}$

The full list of alkanes from methane up to undecanes and alkyls from methyl up to octyls with their IUPAC names, cycle indices, clane compositions and coefficients of the counting series is presented in the Atlas [4].

Conclusions.

A discussion of structural and topological aspects of alkanes have been presented together with corresponding alkyls by the use of clanes symbolism and cycle indices of individual isomers. Clanes permit the concentration of a wide variety of structural isomers to the limited numbers of types that undergo the same conformity as a partition of any integer number into four positive summands. This approach permits us to easily pick out and outline the isomers with the most branched molecules. The cycle index approach provides the opportunity to enumerate all cases of hydrogen atoms substitution by any other atoms and their groups that exist and can exist in organic chemistry.

Acknowledgement.

The author is grateful to D.Sci. S.S.Tratch (Chemistry Department of MSU, Organic Chemistry Chair) for fruitful discussion of the main ideas of the paper.

References.

1. *L.Bytautas, D.J.Klein*. Chemical Combinatorics for Alkane-Isomer Enumeration and More. – J. Chem. Inf. Comput. Sci. 1998, V. 38, pp. 1063 – 1078; *L.Bytautas, D.J.Klein*. Alkane Isomer Combinatorics: Stereostructure Enumeration and Graph-Invariant and Molecular-Property Distributions. ibid. 1999, V. 39, pp. 803 – 818; *L.Bytautas, D.J.Klein*. Mean Wiener Numbers and Other Mean Extensions for Alkane Trees. ibid. 2000, V. 40, pp. 471 – 481.

2. *М.Ю.Корнілов*. Числа ізомерів органічних сполук. Навчальний посібник для магістрів-хіміків. – К., ВПЦ "Київський університет", 2011, 79 с. (M.Yu.Kornilov. Isomer numbers of organic compounds. The textbook for magisters-chemists. – Kyiv. VPC "Kyiv University", 2011, 79 pp.).

3. *R.C.Read*. The Enumeration of Acyclic Chemical Compounds. In: Chemical Applications of Graph Theory, Ed. A.T.Balaban. Acad. Press, London etc., 1976, pp. 25 – 61.

4. *М.Ю.Корнілов*. Атлас алканів – К., ВПЦ "Київський університет", 2013, 50 с. (M.Yu.Kornilov. Atlas of alkanes – Kyiv. VPC "Kyiv University", 2013, 50 pp.).