ISSN 1070-4280, Russian Journal of Organic Chemistry, 2006, Vol. 42, No. 1, pp. 1–11. © Pleiades Publishing, Inc., 2006.
Original Russian Text © I.G. Zenkevich, 2006, published in Zhurnal Organicheskoi Khimii, 2006, Vol. 42, No. 1, pp. 9–20.

General Relations Holding in Variation of Physical Properties of Organic Compounds within Homologous Series

I. G. Zenkevich

Research Institute of Chemistry, St. Petersburg State University, Universitetskii pr. 26, St. Petersburg, 198504 Russia
e-mail: igor@IZ6246.spb.edu

Received December 27, 2004

Abstract—The patterns of variation of different parameters (*A*) of organic compounds within a homologous series (such as boiling point under atmospheric pressure, critical temperature, critical pressure, refractive index, relative density, viscosity, surface tension, saturated vapor pressure, dielectric constant, first adiabatic ionization energies, etc.) are identical, and they can be described in terms of a single linear recurrent equation, $A(n + 1) = aA(n) + b$. This equation relates a property of any member of a homologous series to the corresponding parameters of preceding homologs. The largest deviations from the proposed relation were revealed only in some homologous series for a few (1 or 2) simplest representatives which are characterized by deviations of all parameters.

DOI: 10.1134/S1070428006010015

Classical physical constants of organic compounds, such as boiling and melting points, refractive indices (n_D^{20}), densities d_4^{20}, etc., retain so far their significance for characterization and identification of individual substances [1]. It is also beyond doubt that maximally reliable determination of their thermodynamic parameters is important. Moreover, modern methods for investigation of organic compounds extend the number of parameters necessary for their successful application. For example, energies of ionization determine the potential of mass spectrometric and, in particular photoionization detection [2]. Distribution coefficients in heterophase solvent systems (e.g., in the standard octanol–water system) not only characterize lipophilic properties of substances but also ensure estimation of the order of their elution in reversed-phase high-performance liquid chromatography (HPLC) [3]. Such analytical parameters as gas chromatographic retention indices (*RI*), which can be used for identification of both individual substances and components of complex mixtures, are uniquely dependent on their boiling points.

However, not all organic compounds have been characterized by a complete set of different constants. The most thorough was the historically first reference database [4], whereas the number of known electron affinities is incommensurably lesser (reliable methods for determination of this parameter were developed only in 1980s). First of all, the lack of reference data is typical for most higher homologs in any homologous series (as a consequence of a large number of isomers), synthetically inaccessible compounds, or components of difficultly separable mixtures. In addition, a considerable number of organic substances were characterized by unreliable constants which were determined for samples with a poor purity (prior to chromatographic methods have been introduced into practice); therefore, such values are to be refined. Taking the above stated into account, calculation of physical constants is very important not only for already characterized compounds but also for control of reliability of the known values.

Methods for calculation of all constants should meet a number of requirements. First of all, they should be based on readily available rather than exotic data, whose accuracy should be no less that the required accuracy of the obtained estimates. For example, an error of no greater than ±1–2°C is admissible for boiling point; on subsequent recalculation to retention index, this error is equivalent to 5–10 units, i.e., it is comparable with the up-to-date interlaboratory experimental error on standard nonpolar stationary phases. For refractive indices n_D^{20} and densities d_4^{20}, the error should not exceed ±0.001, otherwise the significance of these parameters essentially decreases. From this viewpoint, we must recognize that the potential of ACD software (HyperCube Corp., Canada), despite the simplicity of its use, is far from being optimal: the confidence interval for boiling points is as large as ±(20–30°C) and more, and that for n_D^{20} and d_4^{20}, 0.002–0.01, depending on the structure.

The presently used methods for theoretical estimation of physical constants (A) can be divided into four main groups:

(1) Estimation of a parameter on the basis of other parameters of the same compound;

(2) Calculation of a parameter from analogous parameters of other structurally related compounds;

(3) Correlation of physical parameters with the position of a compound in the corresponding taxonomic group (i.e., homologous series, groups of congeners, etc.). In the simplest case, e.g., in a series of unbranched homologs, the position of a compound is characterized by the number of carbon atoms (n_C), i.e., $A = f(n_C)$;

(4) Calculation by additivity schemes.

The first group methods utilize various physicochemical relations; on the whole, their accuracy is poor [5, 6]. A typical example is estimation on the basis of molar refraction [$MR_D = f(M, n_D^{20}, d_4^{20})$]. Using MR_D values calculated by additivity schemes, refractive index of a compound may be estimated from its density and vice versa.

The second group methods are based on the structural analogy principle [7] or, in other words, on correlations between physical constants of compounds belonging to different taxonomic groups. These methods were used most widely to assess boiling points; they ensured the maximal accuracy in the estimation of that parameter. The boiling points of homologs occupying equivalent positions in the corresponding homologous series are related to each other through semilog dependences like (1) [8]:

$$\log A(2) = a \log A(1) + bY + c. \tag{1}$$

Here, Y is a parameter ensuring equivalent positions of the involved homologs in the corresponding series ($Y = M$, n_C, MR_D, etc); and a, b, and c are coefficients determined by the least-squares procedure (boiling points are given in K). A slight modification of Eq. (1) allows it to be extended to multidimensional taxonomic groups of organic compounds [9]. An analogous comparative approach has been applied to inorganic compounds. For example, the existence of a linear relation between boiling points of tetrahalosilanes (SiF_4, $SiCl_4$, $SiBr_4$, SiI_4) and the corresponding tetrahalomethanes makes it possible to calculate boiling point of any of the above compounds from the data for other analogs [10].

The use of $A = f(n_C)$ functions (usually nonlinear) in the third group methods is convenient, but the required dependences are difficult to reveal (they are often deduced empirically). For instance, Kreglewky's equation (2) [11] remains so far the most accurate for the estimation of boiling points of normal C_1–C_{100} alkanes, but it cannot be applied to all other series of organic compounds.

$$\log(a - T_b) = b - cn_C^k. \tag{2}$$

Power functions like $A = an_C^k + b$, where $k < 1$, were proposed [12] to approximate refractive indices n_D^{20} and relative densities d_4^{20}. From the constraint $k < 1$ it follows that there are limiting values of these parameters at $n_C \to \infty$, i.e., $\lim A(n_C \to \infty) = b$, which were estimated at 1.475±0.003 and 0.857±0.008 for n_D^{20} and d_4^{20}, respectively.

The choice of optimal approximating functions is facilitated by the procedure involving selection of simpler relations for the second finite differences, followed by their double integration. Just in this way, Eq. (3) [which is analogous to Eq. (1)] was derived for estimation of boiling points [13]:

$$\log T_b = a \log n_C + bn_C + c. \tag{3}$$

Likewise, more general relation (4) applicable to both $X = T_b$ and n_D^{20} and d_4^{20} was obtained [14]:

$$X = A(n_C - B)^{(C+2)}/[(C+1)(C+2)] + Dn_C + E. \tag{4}$$

Here, A, B, C, D, and E are constants.

Optimal choice of approximating functions ensures a fairly high accuracy of the corresponding estimates.

An example of additivity schemes is the use of so-called boiling point numbers N [15, 16]. Values of N are estimated from increments of different atoms and structural fragments of organic molecules $N = \Sigma N_i$ (CH_2, 3.15; CH_3, 3.65; H, 1.20; F, 0.65; Cl, 3.40; Br, 5.88; I, 8.35; Si, 4.20; etc.) using the empirical equation $T_b \approx 230N^{1/3} - 270$. Naturally, the presence in this equation of coefficients characterized by 1 to 2 significant digits implies that the number of reliable significant digits in the result will not be greater. An analogous algorithm underlies ACD software. The accuracy of estimates obtained by additivity schemes is comparable to that inherent to the first group methods but is lower than the accuracy of calculations by semilog equations.

Diversity of approaches and methods for estimation of physical parameters of organic compounds, some of which were briefly discussed above, makes it possible to select optimal procedures in each particular case, depending on the available data. However, different algorithms can also be proposed for the same purpose. For example, no attempts have been made so far to estimate different parameters (A) of homologs from the corresponding values of preceding homologs in the same series. In fact, this means that recurrent equations like (5) must be used:

$$A(n + 1) = f[A(n)]. \tag{5}$$

Just that approach is the subject of the present work.

General characteristic of the potential of recurrent relations (with atmospheric-pressure boiling points as an example). It seems to be unreasonable to perform detailed analysis of mathematical equations used in chemistry without using particular examples; therefore, such parameter as boiling point will be considered first. The possibility of using recurrent relations like (5) for estimation of various physical constants originates from general principles of comparing structures of compounds belonging to different homologous series (structural analogy principle), which are reflected, e.g., in Eq. (1). In fact, parameters of series (2) may be estimated from the data both for a different series of compounds [series (1)] and for the same series with some “shift” along the number of carbon atoms. In other words, the data for $A(n – k)$ ($k \geq 1$) may be used for the calculation of $A(n)$. From the practical viewpoint, the most important is the use of data for the nearest preceding homolog, i.e., when $k = 1$. For example, boiling points of 2-methylalkanes may be estimated from the correlation with boiling points of the preceding homologs. Let us assume the data for 2-methylheptane to be arbitrarily unknown.

Compound	n_C	T_b, °C	Structural analog	$n_C - 1$	T_b, °C
2-Methylbutane	5	27.9	2-Methylpropane	4	–11.7
2-Methylpentane	6	60.3	2-Methylbutane	5	27.9
2-Methylhexane	7	90.1	2-Methylpentane	6	60.3
2-Methylheptane	8	$X = ?$	2-Methylhexane	7	90.1
2-Methyloctane	9	143.3	2-Methylheptane	8	–
2-Methylnonane	10	167	2-Methyloctane	9	143.3

For $N = 4$, the parameters of Eq. (1) are as follows: $Y = n_C$, $a = 0.59811$, $b = 1.0452 \times 10^{-2}$, $c = 2.2753$; the calculated boiling point is $T_b = 116.3$°C, and the reference value is 117.6°C. Checking of this calculation procedure showed that the use of simple linear equations like (6) ensures higher accuracy (at least for boiling points), as compared to general semilog relations like (1).

$$T_b(n + 1) = aT_b(n) + b. \tag{6}$$

Here, the coefficients a and b are calculated by the least-squares procedure. Moreover, analogous linear relations are typical of other physical constants of homologs (see below); therefore, Eq. (6) in the general form looks as follows:

$$A(n + 1) = aA(n) + b. \tag{7}$$

No such simple recurrent relations were considered or even noted in any known manual on the calculation and approximation of physical constants (see, e.g., [5–7, 17]). Nonlinear variation of physical constants, primarily of boiling points, is discussed even in textbooks on organic chemistry but exclusively on a qualitative level. For example, according to [18], the plot of boiling points of n-alkanes versus their molecular weights initially ascends steeply, for the gain in molecular weight by 14 units per CH_2 group is a relatively large relative variation while the molecular weight is small; the boiling points of C_1 and C_2 homologs differ by 72°C, whereas the corresponding difference for C_9 and C_{10} alkanes is only 25°C.

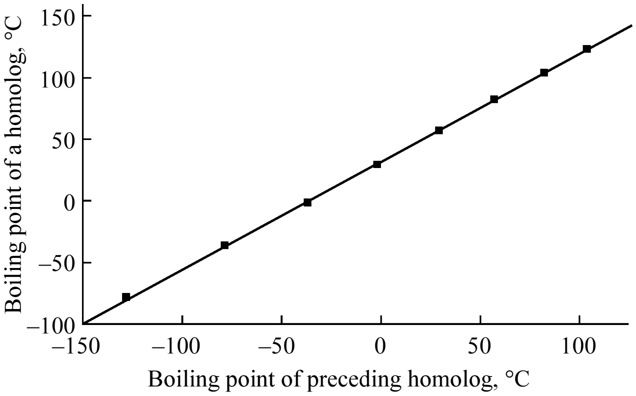
The revealed relation [Eq. (7)] may be illustrated by boiling points of 12 n-alkyl-substituted benzenes (from benzene to undecylbenzene) and 9 unbranched perfluorinated alkanes (CF_4 to C_9F_{20}); their boiling points are given in Table 1. Hereinafter, the statement “data for N contiguous homologs are used” means that the properties of $N - 1$ members of the series 2, ...N correlate with those of ($N - 1$) preceding homologs 1, ...$N - 1$, so that the overall number of couples is equal to $N - 1$. For each group of homologs, coefficients a and b in equations like (6), correlation coefficients r, and general dispersions S_0 are given. The latter may be used to characterize the accuracy of approximation. Taking into account extremely high correlation coefficients ($r = 0.9999$), it is not surprising

Table 1. Boiling points of *n*-alkylbenzenes and perfluoro-*n*-alkanes and their approximation by Eq. (6)

Alkylbenzenes		Perfluoroalkanes	
homolog	T_b, °C	homolog	T_b, °C
Benzene	80.1	CF_4	–128.0
Toluene	110.6	C_2F_6	–78.2
Ethylbenzene	136.2	C_3F_8	–36.6
Propylbenzene	159.2	C_4F_{10}	–1.7
Butylbenzene	183.3	C_5F_{12}	29.4
Pentylbenzene	205.5	C_6F_{14}	57.2
Hexylbenzene	226.1	C_7F_{16}	82.4
Heptylbenzene	246	C_8F_{18}	103.8
Octylbenzene	264.4	C_9F_{20}	123
Nonylbenzene	282		
Decylbenzene	298		
Undecylbenzene	313.2		
Parameters of Eq. (6)			
$a = 0.939 \pm 0.005$, $b = 33.4 \pm 1.0$, $r = 0.9999$, $S_0 = 1.0$ ($N = 11$)		$a = 0.872 \pm 0.005$, $b = 31.8 \pm 0.4$, $r = 0.9999$, $S_0 = 1.1$ ($N = 8$)	

that the plot illustrating Eq. (6) for perfluoroalkanes (see figure) is an almost ideal straight line. Here, the correlation character does not depend on the nature of homologous difference (CH_2 or CF_2).

Insofar as the greatest anomalies of almost all parameters of organic compounds are observed most frequently for one or two simplest homologs, the correlation found for alkylbenzenes may be improved by excluding the data for benzene. As a result, the coefficients in Eq. (6) change insignificantly, and the standard deviation and S_0 decrease: $a = 0.946 \pm 0.003$, $b = 31.7 \pm 0.8$, $r = 0.9999$, $S_0 = 0.7$ ($N = 10$). Exclusion of the data for the simplest homologs in the above and other homologous series (if necessary) is not a serious limitation of the proposed approach, for various constants of these substances are usually known and there is no need of calculating them.



Plot of the linear relation $T_b(n + 1) = aT_b(n) + b$ for C_1–C_9 unbranched perfluoroalkanes.

Correlations of boiling points like (6) are observed for all series of organic compounds, some of which (the simplest) are listed in Table 2. In all cases, the correlation coefficients are no less than 0.999, and S_0 range from 0.6 to 1.4. The plots of such dependences ($r > 0.999$) look like that shown in figure.

Apart from Eq. (7), one more recurrent procedure may be proposed for the calculation of physical constants of higher homologs from the data for three preceding contiguous members of any homologous series. In fact, the equation

$$A(n) = aA(n-1) + b$$

may be written as

$$A(n-1) = aA(n-2) + b.$$

Then,

$$a = [A(n) - A(n-1)]/[A(n-1) - A(n-2)].$$

$$b = [A(n-1)^2 - A(n) \times A(n-2)]/[A(n-1) - A(n-2)].$$

Finally, Eq. (8) is obtained:

$$A(n+1) = \{A(n-1)^2 + A(n)[A(n) - A(n-1) - A(n-2)]\}/[A(n-1) - A(n-2)]. \quad (8)$$

It should be emphasized that the coefficients a are the ratio of the differences in the boiling points of two contiguous homologs, i.e., $a = \Delta T(n, n-1)/\Delta T(n-1, n-2) < 1$; therefore, the boiling points $T_b = A$ may be expressed in any temperature scale (°C, K, °F, °R) without necessary reduction to the Kelvin scale, as with the use of Eq. (1). In the present work all boiling points are given in the Celsius scale (°C). Naturally, other physical constants can be given in any units of measurement.

Turning back to aromatic hydrocarbons, some comments should be given to the calculations of boiling points of homologs having 18 and more carbon atoms. It is clear that N-fold ($N > 1$) successive application of a recurrent equation leads to increased extrapolation error. However, the given data indicate that the error does not exceed 1°C upon "ascend" by even five

Homolog	Calculated T_b, °C		Reference value, °C	Error, °C
	Eq. (6)	Eq. (8)		
Dodecylbenzene	327.5	327.6	327.6	–0.1; 0
Tridecylbenzene	341.0	341.4	341.2	–0.2; +0.2
Tetradecylbenzene	353.8	354.4	354	–0.2; +0.4
Pentadecylbenzene	365.8	366.8	366	–0.2; +0.8
Hexadecylbenzene	377.1	378.5	378	–0.9; +0.5

carbon atoms within a homologous series. Errors in the estimation of boiling points of polar organic compounds having active hydrogen atoms are larger, but they are comparable with the corresponding experimental errors in the determination of boiling points of higher homologs.

Compound	Boiling point, °C					ΔT_b, °C
	three preceding members			calculated by Eq. (8)	reference value	
Octane	36.1	68.6	98.4	125.5	125.7	–0.2
1-Decene	93.6	121.3	146.9	170.6	170.6	0.0
1-Chlorooctane	107.8	134.5	159.9	184.1	184.6	–0.5
1-Octanol	138.1	157.5	176.4	194.8	195.3	–0.7
Hexanoic acid	163.2	186.4	205.4	221	223	–2

In practice, high accuracy in the estimation of boiling points is not always necessary. Taking into account that variation of the coefficients *a* and *b* in Eq. (6) in going from one homologous series to another is not very large, the boiling point of any organic compound may be estimated in the simplest way using the data for only one preceding homolog and average values of the above coefficients. Naturally, these average values will be refined as the number of examined homologous series increases; on the basis of the data for several tens series, the average values of *a* and *b* may now be assumed equal to 0.926±0.014 and 34.2±2.5, respectively. In this case, the average accuracy in the estimation of boiling points, $\Delta T_\mathrm{b} \approx [(\Delta a T_\mathrm{b})^2 + (\Delta b)^2]^{1/2}$, is lower than in the estimation by Eqs. (6) and (8) and is about ±2.9°C for $T_\mathrm{b} \approx 100$°C and ±3.8°C for $T_\mathrm{b} \approx 200$°C. Examples of the application of this calculation version are given in Table 3 for 12 organic compounds belonging to different classes in the molecular weight range from 96 to 178. The average errors ΔT_b are about ±1°C, i.e., the results are comparable with those calculated using ACD software, but the calculation procedure is much simpler.

The proposed approach is important for those organic compounds which so far have not been characterized by reliable T_b values, primarily due to difficulties in their synthesis and isolation. For example, a boiling point of 72.9–77.8°C was reported for perfluorocyclohexane (*cyclo*-C_6F_{12}); these data seem to be clearly overestimated, for the boiling points of the preceding homologs are –33 (*cyclo*-C_3F_6), –6.4 (*cyclo*-C_4F_8), and +22.0°C (*cyclo*-C_5F_{10}). Such an "intuitive" statement must be confirmed by specific estimations. ACD calculation gives a value of T_b = 55.4±20°C, which is inadmissible because of very large confidence interval. The use of Eq. (6) for two couples of the corresponding homologs leads to T_b = 52.3°C. Calcula-

Table 2. Parameters of Eq. (6) for approximation of boiling points of organic compounds

Homologous series	Excluded simplest homologs	Parameters of Eq. (6)			
		a	*b*	*r*	S_0
Alkanes	1	0.905±0.003	37.1±0.4	0.9999	0.9
1-Alkenes	1	0.913±0.003	36.3±0.3	0.9999	0.6
Alkylbenzenes	–	0.939±0.005	33.4±1.0	0.9999	1.0
1-Chloroalkanes	–	0.924±0.005	35.1±0.6	0.9999	1.2
1-Bromoalkanes	–	0.928±0.004	35.2±0.6	0.9992	0.9
1-Alkanols	1, 2	0.956±0.001	25.5±1.4	0.9998	1.0
Alkyl acetates	1	0.944±0.009	29.8±1.4	0.9998	0.9
Methyl alkanoates[a]	–	0.978±0.009	25.3±1.2	0.9997	1.4
Alkanoic acids[a]	1, 2	0.930±0.011	32.6±2.4	0.9996	1.3
Alkanenitriles	1, 2	0.950±0.007	29.5±1.4	0.9998	0.9
Perfluoroalkanes[b]	–	0.872±0.005	31.8±0.4	0.9999	1.1

[a] Some boiling points were determined with a low accuracy (within 2-4°C); average values were used in the calculations.

[b] Homologous difference CF_2.

Table 3. Approximation of boiling points of various organic compounds from the boiling points of preceding homologs using Eq. (6) ($a = 0.926 \pm 0.014$, $b = 34.2 \pm 2.5$) and by ACD software

Compound	M	T_b, °C	Boling point of preceding homolog, °C	Calculated T_b, °C			
				Eq. (6)	ΔT_b, °C	ACD	ΔT_b, °C
2-Ethylfuran[a]	96	94	62.8	92.4	–1.6	94.5±9	+0.5
1-Fluorohexane	104	93.15	62.8	92.4	–0.8	91.6±3	–1.6
2-Heptanol	116	160.3	137.4	161.4	+1.1	159.4±3	–0.9
2-Propylthiophene	126	157.5–159.5	135	159.2	+1.7; –0.3	158.4±3	+0.9; –1.1
Hexyl nitrite	131	129–130	104	130.5	+0.5	129.6±3	–0.4
N-Propylaniline	135	222	204.7	223.7	+1.7	221.1±9	–0.9
4-Butylpyridine	135	207–209	189	209.2	+0.2	207.0±9	0
4-Butylphenol	136	248	232.8	249.8	+1.8	246.2±9	–1.8
Undecanenitrile	167	260.8	244.1	260.2	–0.6	253[b]	–[b]
1-Propylnaphthalene[a]	170	272.8; 274.5	258.7	273.8	+1.0; –0.7	275.3±3	+2.5; +0.8
Decanoic acid	172	269	254	269.4	+0.4	269.6±3	+0.6
Butyl benzoate	178	250	231.2	248.3	–1.7	250.4±8	+0.4
Average ΔT_b, °C				±1.0		±1.0	

[a] Compounds with the minimal (94°C) and maximal boiling points (273–274°C) define the range of variation of T_b estimate errors by Eq. (6) with average coefficients (±2.8 to ±4.4°C).

[b] Invalid experimental boiling point in the ACD database; it was not taken into account while estimating ΔT_b.

tion by the same equation with the coefficients a and b for perfluoroalkanes (as closest analogs of perfluorocycloalkanes; Table 2) gives $T_b = 50.9°C$. The availability of data for the three preceding homologs makes it possible to use recurrent Eq. (8) to obtain $T_b = 52.3°C$. Finally, the known structural analogy procedure based on Eq. (1) gives $T_b = 51.8°C$ by comparing the data for perfluorocycloalkanes and perfluoroalkanes ($a = 0.18743$, $b = 8.0049 \times 10^{-3}$, $s = 4.2163$, $A = n_C$). Thus, calculations by four different methods led to almost coinciding boiling points of perfluorocyclohexane which finally may be characterized by an average value of 51.8±0.7°C.

Extensions of recurrent relations to estimation of physical constants. As shown above, the proposed approach implies the use of both general recurrent equations like (7) and particular relations like (8) which require the data for three preceding homologs. In addition, preliminary estimates can be obtained from the constants of only one preceding homolog and average values of the coefficients a and b.

General recurrent Eq. (7) may be treated as the mathematical equation $x(n) = A(n) = ax(n-1) + b$, which has the following solution:

$$A(n) = ca^n + b(a^n - 1)/(a - 1), \qquad (9)$$

where a, b, and c are constants. Equation (9) is a combination of power functions with the exponents differing by unity since $(a^n - 1)/(a - 1) = a^{n-1} + a^{n-2} + ... + a + 1$; it would be quite difficult to deduce such a relation empirically. Therefore, while speaking about general classification of methods for calculating physical parameters (see above), recurrent relations should undoubtedly be treated as a particular group; however, from the mathematical viewpoint, they are not standalone and are a version of correlations like $A = f(n_C)$. Here, the general form of function (9), as well as of initial recurrent relation (7), was previously unknown.

A number of important extensions follow from Eq. (9). First, depending on the coefficient a, the values of $A(n)$ can either increase unlimitedly ($a \geq 1$) or tend to a certain limit, $\lim A = b/(1 - a)$ ($a < 1$) as the number of carbon atoms infinitely increases (hypothetically, $n_C \to \infty$). In this connection, properties of organic compounds may be divided into two groups: those having no limit upon hypothetical infinite increase of the number of carbon atoms in molecule (e.g., boiling point [12, 14], critical temperature, viscosity, partition coefficient in heterophase solvent systems, etc.) and those having such a limit (refractive index, relative density, surface tension, critical pressure, ionization energy, dielectric constant, dipole moment, etc). Calculation of coefficients a in equations like (7) provides a rigorous mathematical support to the conclusion implying unlimited variation of a physical parameter: if $a > 1$, the value of that parameter should increase with no limit as the number of carbon atoms rises. The opposite statement is not always valid. For instance, boiling point T_b and critical temperature T_{crit} (see below) may increase unlimitedly (theoretically), but the corresponding coefficient a is less than zero. This means that Eq. (7) is not absolutely rigorous for description of variation of these constants within a homologous series; nevertheless, it is fulfilled with a high accuracy. Estimates of such “false” limits from the data in Table 2 fall into the range from 400 to 600°C, i.e., they exceed a reasonable upper limit of measurement of boiling points under atmospheric pressure, which is quite admissible.

Probably, estimation of physical constants of homologs from the data for contiguous members (differing by ±1 carbon atom) is important from the practical viewpoint. From two equations like (7), $A(n) = a \times A(n-1) + b$ and $A(n+1) = aA(n) + b$, we obtain

$$A(n) \approx [aA(n-1) + A(n+1)]/(1 + a). \qquad (10)$$

For preliminary estimation of $A(n)$, average values of coefficients a may be used, which characterize variation of the same parameter within different homologous series, e.g., $a = 0.926$ for T_b. Estimates obtained by such a simple procedure are often better than arithmetic or geometric means of constants of neighboring homologs, especially for the first members of a homologous series. This may be illustrated using ethylbenzene as an example. Its reference boiling point is 136.2°C, and the boiling points of its closest homologs, toluene and propylbenzene, are 110.6 and 159.2°C, respectively.

Method	T_b, °C
Arithmetic mean	134.9
Geometric mean	132.7
Geometric mean of absolute boiling temperatures	135.8
Eq. (10), $a \approx 0.926$	135.8

Having considered general properties of recurrent relations like (7), let us turn back to their application to the calculation of various physical constants of organic compounds.

Critical temperature T_{crit} and critical pressure p_{crit}. Insofar as the physical nature and mainly the pattern of variation of T_{crit} within a homologous series are identical to those typical of T_b, general recurrent Eq. (7) should be applicable to critical temperature as well. This may be verified using available reference data [19]. For example, T_{crit} (°C) and p_{crit} (atm) for the most thoroughly characterized normal C_2–C_{19} alkanes (methane is excluded due to anomalous values of these parameters) are as follows:

Alkane	T_{crit}	P_{crit}	Alkane	T_{crit}	P_{crit}	Alkane	T_{crit}
C_2H_6	32.3	48.2	C_8H_{18}	296.7	24.6	$C_{14}H_{30}$	422
C_3H_8	96.8	42.0	C_9H_{20}	322	22.5	$C_{15}H_{32}$	437
C_4H_{10}	152.0	37.5	$C_{10}H_{22}$	346	20.8	$C_{16}H_{34}$	452
C_5H_{12}	196.6	33.3	$C_{11}H_{24}$	367	19.2	$C_{17}H_{36}$	462
C_6H_{14}	234.7	29.9	$C_{12}H_{26}$	386	17.9	$C_{18}H_{38}$	477
C_7H_{16}	267.0	27.0	$C_{13}H_{28}$	404	–[a]	$C_{19}H_{40}$	487

[a] Starting from tridecane, P_{crit} values were determined with a poor accuracy.

Correlation (7) has the following parameters: $a = 0.885 \pm 0.005$, $b = 63.1 \pm 1.9$, $r = 0.9997$, $S_0 = 2.9$ ($N = 17$). As with boiling points, the coefficient a is less than unity, indicating the existence of a “false” limit at about 550°C. However, this does not preclude the use of Eq. (7) in the temperature range up to 500°C.

The values of p_{crit} decrease as the number of carbon atoms rises; therefore, p_{crit} should tend to a limiting value. The parameters of the equation $p_{crit}(n+1) = ap_{crit}(n) + b$ for C_2–C_{12} alkanes are as follows: $a = 0.84 \pm 0.06$, $b = 1.9 \pm 0.2$, $r = 0.9998$, $S_0 = 0.19$ ($N = 10$). The limit $\lim p_{crit}(n_C \to \infty) = b/(1-a)$ was thus estimated at 11.6 atm; this value is likely to be only the best asymptotic approximation for the given limited set of data.

Refractive index n_D^{20} and relative density d_4^{20}. With rise in the number of carbon atoms, these dimensionless parameters of organic compounds asymptotically tend to limiting values which were estimated previously at 1.475±0.003 (n_D^{20}) and 0.857±0.008 (d_4^{20}) [12]. Therefore, just these constants may be expected to be approximated most accurately by correlations like (7). Table 4 contains the relative densities and refractive indices of unbranched C_6–C_{13} alkanes and parameters of the corresponding correlations, which fully support the above conclusion. It is interesting that the limiting value of n_D^{20}, $\lim n_D^{20}(n_C \to \infty)$, calculated from these data is $0.270/(1 - 0.813) = 1.475$, i.e., it coincides with the known estimate [12]. In the case of d_4^{20}, the coincidence is slightly worse.

Table 4. Experimental relative densities d_4^{20} and refractive indices n_D^{20} of C_6–C_{13} alkanes and parameters of the corresponding correlations like (7)

Number of carbon atoms	d_4^{20}	n_D^{20}
6	0.660	1.3751
7	0.684	1.3878
8	0.703	1.3974
9	0.718	1.4054
10	0.730	1.4120
11	0.740	1.4172
12	0.749	1.4216
13	0.756	1.4256
Parameters of Eq. (7)		
a	0.804±0.008	0.813±0.008
b	0.154±0.006	0.270±0.011
r	0.9997	0.9998
S_0	0.0007	0.0003

Unique potential of the proposed approach follows, e.g., from the fact that the erroneous value (probably, misprint) of n_D^{20} for decane, 1.4102 [20], was revealed by the anomalous position of the corresponding point on the plot of type (6) dependence. The correct value is n_D^{20} = 1.4120 [4], i.e., the error is only two third decimal points.

The application of recurrent relations to the estimation of d_4^{20} may be illustrated by numerous examples, the two simplest of which are given below.

Relative density d_4^{20}

Compound	three preceding homologs			calculated by Eq. (8)	reference value	Δd_4^{20}
Decane	0.684	0.703	0.718	0.730	0.730	0
1-Heptanol	0.810	0.816	0.819	0.820	0.822	–0.002

Likewise, a large number of examples illustrating estimation of n_D^{20} may be given, but it seems to be reasonable to present only two.

Refractive index n_D^{20}

Compound	three preceding homologs			calculated by Eq. (8)	reference value	Δn_D^{20}
Tetradecane	1.4172	1.4216	1.4256	1.4292	1.4290	+0.0002
1-Hexanol	1.386	1.399	1.410	1.419	1.418	+0.001

Dynamic viscosity η. Like many other physical constants, the applicability of Eq. (7) to estimation of dynamic viscosity of organic compounds within a homologous series may be illustrated most simply using normal alkanes. The data for C_5–C_{17} homologs (η, cP, 20°C) are given below.

Alkane	η	Alkane	η	Alkane	η
C_5H_{12}	0.240	$C_{10}H_{22}$	0.907	$C_{15}H_{32}$	2.841
C_6H_{14}	0.307	$C_{11}H_{24}$	1.182	$C_{16}H_{34}$	3.451
C_7H_{16}	0.417	$C_{12}H_{26}$	1.492	$C_{17}H_{36}$	4.209
C_8H_{18}	0.546	$C_{13}H_{28}$	1.878		
C_9H_{20}	0.714	$C_{14}H_{30}$	2.322		

The corresponding correlation has the following parameters: a = 1.204±0.006, b = 0.05±0.01, r = 0.9999, S_0 = 0.02 (N = 12). In keeping with the above stated [see comments to Eq. (9)], the constraint $a > 1$ implies unlimited increase of η as the number of carbon atoms increases. This does not contradict physical nature of dynamic viscosity. It should be noted that all quantities derived from η (e.g., kinematic viscosity which is the ratio of dynamic viscosity η and density d of a liquid) should also fit Eq. (7).

Surface tension σ. Here, only the values for nonpolar and weakly polar compounds fit Eq. (7). The data for normal C_6–C_{18} alkanes (σ, dyn/cm, 20°C) are given below [a = 0.848±0.008, b = 4.55±0.19, r = 0.9996, S_0 = 0.08 (N = 12)].

Alkane	σ	Alkane	σ	Alkane	σ
C_6H_{14}	18.41	$C_{11}H_{24}$	24.78	$C_{16}H_{34}$	27.64
C_7H_{16}	20.29	$C_{12}H_{26}$	25.48	$C_{17}H_{36}$	28.06
C_8H_{18}	21.78	$C_{13}H_{28}$	26.13	$C_{18}H_{38}$	28.42
C_9H_{20}	22.96	$C_{14}H_{30}$	26.69		
$C_{10}H_{22}$	23.89	$C_{15}H_{32}$	27.17		

A limiting value of σ ($n_C \to \infty$) may be anticipated; from the data for normal alkanes, it was estimated at 4.55/(1 – 0.848) ≈ 29.9. As concerns polar organic compounds which have labile hydrogen atoms and are strongly associated in the liquid phase, the corresponding σ–n_C dependences do not conform to Eq. (7) (cf. the data for alkanoic acids [19]).

Saturated vapor pressure (*p*) and derived physical constants. Insofar as correlation (7) is applicable to critical pressure (see above), it is reasonable to presume that Eq. (7) should also be appropriate to description of saturated vapor pressure of homologs at a constant temperature. However, the necessary experimental data are difficultly accessible; therefore, saturated vapor pressures p (mm) of C_7–C_{14} alkylbenzenes calculated by ACD software were used to illustrate the relation $p(n + 1) = ap(n) + b$ [$a = 0.322\pm0.003$, $b = 0.02\pm0.02$, $r = 0.9998$, $S_0 = 0.05$ ($N = 7$)].

n_C(R) in RPh	p(20°C)	n_C(R) in RPh	p(20°C)
C_1	21.4	C_5	0.25
C_2	6.9	C_6	0.083
C_3	2.3	C_7	0.028
C_4	0.84	C_8	0.010

Obviously, the theoretical limit of p at $n_C \to \infty$ is $p = 0$. In fact, taking into account that the standard deviation of b is equal to b, this coefficient should be assumed equal to zero; then, $\lim p(n_C \to \infty) = b/(1 - a) = 0$.

Correlations like (7) are also fulfilled for some quantities derived from saturated vapor pressure. For example, the coefficients A, B, and C in the Antoine equation $\log p = A - B/(T + C)$ which describes temperature dependence of saturated vapor pressure are known for a large number of compounds in a wide temperature range [19]. All these coefficients fit Eq. (7), as illustrated by the values of B for normal C_5–C_{20} alkanes: $B(n + 1) = aB(n) + b$; $a = 0.915\pm0.009$, $b = 192\pm14$, $r = 0.9994$, $S_0 = 8.9$ ($N = 15$).

n_C	B	n_C	B	n_C	B	n_C	B
5	1075.82	9	1428.81	13	1677.43	17	1847.82*
6	1171.53	10	1501.17	14	1725.46	18	1888.73
7	1266.87	11	1572.48	15	1768.82	19	1917.0
8	1355.23	12	1625.93	16	1831.32	20	1948.7

The plot of this dependence shows distinct anomalies for two pairs of the $B(n + 1) - B(n)$ values, which clearly result from unreliable value of B for heptadecane (marked with an asterisk).

Dielectric constant ε. Historically, dielectric constants were measured under different conditions; therefore, their values are sometimes poorly reproducible, and comparison of dielectric constants for different homologs is difficult. Moreover, characterization of large series of homologs with dielectric constants is often unnecessary. However, even the available limited data demonstrate applicability of general relations like (7). For example, dielectric constants of C_1–C_7 primary alcohols at 25°C [21] may be used [$a = 0.73\pm0.04$, $b = 2.1\pm0.8$, $r = 0.995$, $S_0 = 0.6$ ($N = 6$)].

Alcohol	ε	Alcohol	ε
Methanol	32.65	1-Pentanol	15.0[b]
Ethanol	25.2	1-Hexanol	12.5
1-Propanol	21.3[a]	1-Heptanol	11.1
1-Butanol	17.7		

[a] Clearly deviating value 19.7 was replaced by the estimate obtained from the data for contiguous homologs using Eq. (10).

[b] No data were available; the value calculated as in note [a] is given.

The obtained results suggest the existence of some nonzero limiting value of ε at $n_C \to \infty$; this value was estimated at $2.1/(1–0.73) \approx 7.7$ which seems to be quite reasonable, taking into account the known dielectric constant of 1-decanol (ε = 8.1).

First adiabatic ionization energy *I*. After consideration of the above physical parameters, it is not surprising that general correlation (7) is also fulfilled for ionization energies which are determined by energies of the highest occupied molecular orbitals, $I(n + 1) = aI(n) + b$. The first adiabatic ionization energies of normal alkanes [21] regularly decrease with rise in n_C, approaching a certain limiting value: $a = 0.66\pm0.03$, $b = 3.3\pm0.3$, $r = 0.996$, $S_0 = 0.06$ ($N = 6$).

Alkane	I, eV	Alkane	I, eV
CH_4	12.6	C_5H_{12}	10.35
C_2H_6	11.6	C_6H_{14}	10.18
C_3H_8	11.07	C_7H_{16}	10.0
C_4H_{10}	10.6		

The up-to-date accuracy in the determination of ionization energies is not high (about ±0.1 eV), and the range of variation of this parameter is relatively narrow; therefore, the correlation coefficient is lower than in the above examples. The limiting value of I for alkanes is about 9.85 eV. The data for other homologous series, for which ionization energies of several contiguous members are known (1-alkenes, alkyl alkanoates, etc.), conform to the general pattern.

Distribution coefficients in heterophase organic solvent systems (K_d, log*P*). The most widely used heterophase system for determination of distribution coefficients of organic compounds is 1-octanol–water which simulates properties of cell membranes [22]; from the analytical viewpoint, combinations of low-boiling solvents, e.g., hexane–acetonitrile, or condensed phase–gas systems (so-called vapor phase analysis) are preferred. In all cases, K_d or $\log P$ values for homologs fit general relations like (7). As the simplest example, log*P* values of C_1–C_7-alkyl-substituted benzenes are given.

n_C(R) in RPh	log*P*	n_C(R) in RPh	log*P*
C_1	2.19	C_5	4.26
C_2	2.74	C_6	4.90
C_3	3.17	C_7	5.52
C_4	3.66		

The corresponding correlation has the following parameters: $a = 1.06\pm0.02$, $b = 0.34\pm0.08$, $r = 0.9991$, $S_0 = 0.05$ ($N = 6$). Insofar as $a > 1$, log*P* should have no limit at $n_C \to \infty$, which does not contradict physical sense of that constant. Therefore, the general recurrent relation should also be fulfilled for properties proportional to the number of carbon atoms in molecule [particular cases of the general equation at $a \approx 1$, $b \approx$ *const* or $A(n) = a'n + b'$, e.g., thermodynamic phase transition functions, distribution coefficients in heterophase solvent systems, etc.] or those represented in specially selected linearized scales (e.g., chromatographic retention indices).

To complete listing of properties of organic compounds, whose variation within a homologous series may be described in terms of the general relation $A(n + 1) = a\,A(n) + b$, those parameters must be noted, for which analogous relation still has not been verified. The main reasons, as with some parameters considered above (ε, I), are different conditions for their determination and clearly unsatisfactory characterization of homologous series. Such properties include dipole moments μ, specific heat capacities C_p, and heat conductivities λ. However, even scanty known data suggest the existence of the corresponding limiting values, which indicates the possibility for their approximation by correlations like (7). It is important that the list of physical parameters to which the proposed approach may be applied is not restricted to those considered in the present work. For example, an analogous correlation may be expected (and it really exists) for such a parameter as flash temperature, which is undoubtedly interesting from the technical viewpoint.

Presumably, recurrent approach cannot be applied to such parameters which are characterized by strong alteration for homologs with even and odd numbers of carbon atoms. Among these, the most common property is melting point.

Up to now, the classical concepts of physical chemistry implied that different physical constants of organic compounds required development of radically different methods for their theoretical estimation. The existence of a simple common recurrent relation [$A(n + 1) = aA(n) + b$] between the value of any parameter of any member of a homologous series and the value of the same parameter of preceding homologs indicates that these concepts should be revised. The reason for identical patterns of variation of different properties is not their origin but the same chemical nature of compounds possessing these properties and differing from each other by similar units, homologous (CH_2) or other (e.g., CF_2).

Thus just similar patterns in variation of the molecular structure of organic compounds are responsible for mathematically identical patterns of variation of their properties within a homologous series. It is important that these patterns are identical for both nonpolar compounds and those containing labile hydrogen atoms and hence strongly associated in condensed phase (alcohols, carboxylic acids, etc.)

To conclude, it should be noted that recurrent relations like (7) may be regarded as a general law of variation of properties of organic compounds in homologous series (more exactly, within the series including unbranched homologs): most known properties of organic compounds are linearly related to the corresponding properties of preceding homologs.

EXPERIMENTAL

The physical constants of organic compounds were taken mostly from reference books [4, 19] and other available sources [20, 23]. The first adiabatic ionization energies of the simplest organic compounds were taken from [21]. If several values of T_b, n_D^{20}, d_4^{20}, etc. were available, either those encountered repeatedly or values containing the maximal number of significant digits were taken; boiling points were rounded to one decimal point, and n_D^{20} and d_4^{20}, to four decimal points. For compounds characterized by a temperature range, the average temperature was used.

Parameters of linear correlations were calculated using Origin 4.1 software. Calculations by recurrent equations were performed using QBasic program (including that adapted for a Casio PB 100 programmable calculator). ACD software was used to obtain comparative estimates of T_b, n_D^{20}, d_4^{20}, and $p(20°C)$.

The author thanks Dr. A.N. Marinicheva (Research Institute of Chemistry, St. Petersburg State University) for the analytical solution of the principal recurrent relation [Eq. (9)].

REFERENCES

1. *Organikum. Organisch-chemisches Grundpraktikum*, Berlin: Wissenschaften, 1976, 15th ed.
2. Vulikh, P.P., Polukhin, D.Yu., Revelsky, I.A., Yashin, Yu.S., and Napalkova, O.V., *Abstracts of 20th Int. Symp. on Capillary Chromatography*, Riva del Garda, Italy, 1998.
3. Smith, R.M., *Retention and Selectivity in Liquid Chromatography*, Amsterdam: Elsevier, 1995.
4. *Beilsteins Handbuch der organischen Chemie*, 1918, vols. 1–31.
5. Reid, R.C. and Sherwood, T.K., *The Properties of Gases and Liquids*, New York: McGraw–Hill, 1977, 3rd ed.
6. Stolyarov, E.A. and Orlova, N.G., *Raschet fiziko-khimicheskikh svoistv zhidkostei* (Calculation of Physicochemical Parameters of Liquids), Leningrad: Khimiya, 1976.
7. Karapet'yants, M.Kh., *Metody sravnitel'nogo rascheta fiziko-khimicheskikh svoistv* (Methods for Comparative Analysis of Physicochemical Properties), Moscow: Nauka, 1965.
8. Zenkevich, I.G., *Zh. Org. Khim.*, 1994, vol. 30, p. 1441.
9. Zenkevich, I.G., *Russ. J. Org. Chem.*, 1996, vol. 32, p. 1408.
10. Morachevskii, A.G. and Sladkov, I.B., *Fiziko-khimicheskie svoistva molekulyarnykh neorganicheskikh soedinenii. Eksperimental'nye dannye i metody rascheta* (Physicochemical Properties of Moleclar Inorganic Compounds. Experimental Data and Methods of Calculation), St. Petersburg: Khimiya, 1996, p. 93.
11. Kreglewsky, A. and Zwolinsky, B.J., *J. Phys. Chem.*, 1961, vol. 65, p. 1050.
12. Ioffe, B.V. and Zenkevich, I.G., *Zh. Fiz. Khim.*, 2000, vol. 74, p. 2101.
13. Zenkevich, I.G., *Russ. J. Org. Chem.*, 1998, vol. 34, p. 1399.
14. Zenkevich, I.G. and Kranics, B., *Chemometr. Intel. Lab. Systems*, 2003, vol. 67, p. 51.
15. Kinney, C.R., *J. Am. Chem. Soc.*, 1938, vol. 60, p. 3032.
16. Sladkov, I.B., *Zh. Prikl. Khim.*, 1999, vol. 72, p. 1266.
17. Tatevskii, V.M., Benderskii, V.A., and Yarovoi, S.S., *Rules and Methods for Calculating the Physico-Chemical Properties of Paraffinic Hydrocarbons*, New York: Pergamon, 1961.
18. Fieser, L.F. and Fieser, M., *Advanced Organic Chemistry*, New York: Reinhold, 1962.
19. *Spravochnik khimika* (Chemist's Handbook), Nikol'skii, B.P., Ed., Leningrad: Khimiya, 1971, vol. 1.
20. *Svoistva organicheskikh soedinenii. Spravochnik* (Properties of Organic Compounds. Reference Book), Potekhin, A.A., Ed., Leningrad: Khimiya, 1984.
21. Vilesov, F.I., *Uspekhi fotoniki* (Advances in Photonics), Leningrad: Leningr. Gos. Univ., 1969, vol. 1, p. 5.
22. Leo, A., Hansch, C., and Eilins, D., *Chem. Rev.*, 1971, vol. 71, p. 525.
23. *Handbook of Chemistry & Physics*, Weast, R.C., Ed., Boca Raton: CRC, 1982–1983, 63rd ed., p. 689.