

Group Vector Space Method for Estimating Enthalpy of Vaporization of Organic Compounds at the Normal Boiling Point

Wei Wenying,* Han Jinyu, and Xu Wen

Key Laboratory for Green Chemical Technology of State Education Ministry, School of Chemical Engineering & Technology, Tianjin University, Tianjin 300072, P.R. China

Received March 31, 2004

The specific position of a group in the molecule has been considered, and a group vector space method for estimating enthalpy of vaporization at the normal boiling point of organic compounds has been developed. Expression for enthalpy of vaporization $\Delta_{\text{vap}}H(T_b)$ has been established and numerical values of relative group parameters obtained. The average percent deviation of estimation of $\Delta_{\text{vap}}H(T_b)$ is 1.16, which show that the present method demonstrates significant improvement in applicability to predict the enthalpy of vaporization at the normal boiling point, compared the conventional group methods.

INTRODUCTION

To physical properties of organic compounds, enthalpy of vaporization at the normal boiling point is required in many calculations or process simulations and product designs. It is not always possible, however, to find reliable experimental values of these properties for the compounds of interest in the literature, nor is it practical to measure the properties as the need arises. So, estimation methods are profusely employed.

Numerous techniques for predicting the enthalpy of vaporization $\Delta_{\text{vap}}H(T_b)$ at the normal boiling point have been proposed. The famous Trouton's rule expressed it as a constant value. Chen¹ presented a correlation, which was expressed by the critical pressure, the critical temperature, and the normal boiling point. Another estimation method was group contribution. Two representatives of the group contribution methods developed in recent years are the methods by Hoshino et al.² and by Ma and Zhao.³ Estimations by these two methods all need data of the normal boiling point, which sometimes are not acquirable experimentally. The accuracies of estimating enthalpies of vaporization with these two methods are all not too high, and they require the data of the normal boiling point.

In this work, the specific position of a group in the molecule was considered. A group vector space method will be proposed to predict the enthalpy of vaporization $\Delta_{\text{vap}}H(T_b)$ at the normal boiling point for a wide variety of organic compounds with higher accuracy. This method provides the important advantage of quick estimation without requiring the data of normal boiling points.

GROUP VECTOR SPACE FOR ORGANIC COMPOUNDS

In this work, we select 38 simple groups to describe organic compounds. The molecule is considered to be in a given space, and every group in the molecule is only a point in the space. To write conveniently, graphs with a different number of points are all expressed as graphs with five points.

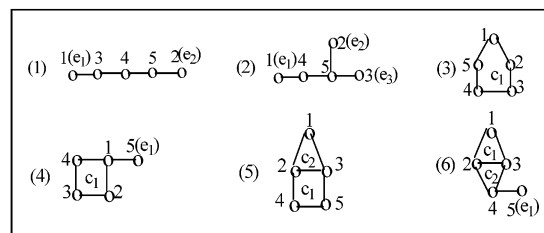


Figure 1. Six topologic graphs expressing organic compounds.

Thus, the organic molecules can be expressed as six topologic graphs in Figure 1.

The topologic structure of a molecule can be described by the distance matrix of the molecule. The distance matrix is a square symmetric matrix. It is required to input $n \times n$ elements for the n -order square matrix. The input quantity for the large molecule is too large, and, furthermore, the distance matrix is unable to describe the molecular structure characteristics such as chains, rings, and branches sufficiently. To overcome the above limitation, the following group vector space method has been proposed.

Considering the chain graph first, the dimension number of the space is equal to the number of end points (e_i) on the chain, and one end point has determined a dimension of the space. The coordinate of an end point in the dimension determined by it is zero, while the coordinate of another point in this dimension is the distance from that point to the end point (the distance between two adjacent groups is defined as 1). For one example, the no. 1 molecule in Figure 1 has three end points, so the dimension of its space is 3, namely e_1 , e_2 , and e_3 . The distances of the group 1 to every dimension e_1 , e_2 , and e_3 are 0, 3, and 3, respectively, and then the coordinate of this group in the space is (0, 3, 3). For the cyclic graph, one ring (c_i) represents a dimension. In that dimension the coordinate of the ring point equals the number of points on the ring, and the coordinate of nonring point equals the sum of the distance from the point to the ring and the number of points on the ring. If the route from the ring point to the end point is nonunique, the shortest route should be selected. So the dimension number m of the space

* Corresponding author e-mail: wwy7324@eyou.com.

Table 1. Values of Group Parameters^a

no.	group	$\Delta_{vap}H_{0i}$	$\Delta_{vap}H_{Li}$	$\Delta_{vap}H_{pi}$	no.	group	$\Delta_{vap}H_{0i}$	$\Delta_{vap}H_{Li}$	$\Delta_{vap}H_{pi}$
1	-CH ₃	0.179	3.565	-2.707	20	>C=O	2.015	0.000	20.690
2	>CH ₂	-0.291	1.362	3.818	21	(>C=O) _R	25.920	0.000	-33.377
3	>CH-	-0.170	1.535	1.719	22	O=CH-	9.875	0.000	0.702
4	>C<	0.515	4.842	-17.014	23	-COOH	73.274	0.000	-95.424
5	=CH ₂	0.964	3.375	-4.714	24	-COO-	-13.130	19.577	9.827
6	=CH-	-0.139	1.263	4.549	25	-SH	0.000	6.471	3.910
7	=C<	-0.460	11.131	-25.220	26	-S-	1.832	0.715	15.410
8	=C=	3.745	0.000	0.449	27	(-S-) _R	0.000	3.178	12.226
9	≡CH	3.567	0.000	-2.991	28	-NH ₂	4.494	-4.218	16.180
10	=CH-	0.004	3.763	1.578	29	>NH	0.771	-9.885	43.203
11	(>CH ₂) _R	3.629	3.387	-2.922	30	(>NH) _R	23.079	-32.887	53.376
12	(>CH-) _R	-0.427	-0.215	3.520	31	>N-	0.000	-4.855	27.731
13	(>C<) _R	-3.013	2.467	-1.959	32	(-N=) _R	3.525	0.479	9.449
14	(=CH-) _R	-2.406	4.672	-3.293	33	-CN	25.731	-30.832	32.425
15	(=C<) _R	-1.244	0.359	8.550	34	-NO ₂	0.004	0.004	26.445
16	-OH	0.000	-0.155	36.663	35	-F	0.189	0.112	1.259
17	(-OH) _{AC}	-10.517	0.000	64.086	36	-Cl	1.913	2.021	5.989
18	-O-	-0.638	-0.818	13.646	37	-Br	2.488	7.477	-3.900
19	(-O-) _R	-1.615	-1.227	15.335	38	-I	-0.001	10.471	0.369

^a Subscript R represents a ring group; AC represents a connection to the aromatic ring.

for a graph is equal to the sum of the number k_e of end points (e_i) and the number k_c of rings (c_i) in the graph. Another example is the no. 6 molecule in Figure 1; it has two rings and one end point, so it has three dimensions c_1 , c_2 , and e_1 . The distance of the group 5 to e_1 dimension decided by it is zero; the sum (is 5) of the distance (is 2) to ring 1 and the number (is 3) of points on the ring 1 is its coordinate at c_2 dimension; similarly its coordinate at c_1 dimension is 4. So, the coordinate of the group 5 in the space is (5, 4, 0). Every point in the graph has m coordinates in the m -dimension space. The graph may be described by a space matrix, the number of rows in the matrix equals the number of points in the graph, and the number of columns equals the dimension number of the space. These space matrices of the above six topologic graphs are as follows:

$$\begin{aligned}
 & \begin{matrix} 1 & \begin{bmatrix} e_1 & e_2 \\ 0 & 4 \end{bmatrix} \\ 2 & \begin{bmatrix} 4 & 0 \end{bmatrix} \\ 3 & \begin{bmatrix} 1 & 3 \end{bmatrix} \\ 4 & \begin{bmatrix} 3 & 1 \end{bmatrix} \\ 5 & \begin{bmatrix} 2 & 2 \end{bmatrix} \end{matrix}; \quad \begin{matrix} 1 & \begin{bmatrix} e_1 & e_2 & e_3 \\ 0 & 3 & 3 \end{bmatrix} \\ 2 & \begin{bmatrix} 3 & 0 & 2 \end{bmatrix} \\ 3 & \begin{bmatrix} 3 & 2 & 0 \end{bmatrix} \\ 4 & \begin{bmatrix} 1 & 2 & 2 \end{bmatrix} \\ 5 & \begin{bmatrix} 2 & 1 & 1 \end{bmatrix} \end{matrix}; \quad \begin{matrix} 1 & \begin{bmatrix} c_1 \\ 5 \end{bmatrix} \\ 2 & \begin{bmatrix} 5 \end{bmatrix} \\ 3 & \begin{bmatrix} 5 \end{bmatrix} \\ 4 & \begin{bmatrix} 5 \end{bmatrix} \\ 5 & \begin{bmatrix} 5 \end{bmatrix} \end{matrix}; \\
 & \begin{matrix} 1 & \begin{bmatrix} c_1 & e_1 \\ 4 & 1 \end{bmatrix} \\ 2 & \begin{bmatrix} 4 & 2 \end{bmatrix} \\ 3 & \begin{bmatrix} 4 & 3 \end{bmatrix} \\ 4 & \begin{bmatrix} 4 & 2 \end{bmatrix} \\ 5 & \begin{bmatrix} 5 & 0 \end{bmatrix} \end{matrix}; \quad \begin{matrix} 1 & \begin{bmatrix} c_1 & c_2 \\ 3 & 5 \end{bmatrix} \\ 2 & \begin{bmatrix} 3 & 4 \end{bmatrix} \\ 3 & \begin{bmatrix} 3 & 4 \end{bmatrix} \\ 4 & \begin{bmatrix} 4 & 4 \end{bmatrix} \\ 5 & \begin{bmatrix} 4 & 4 \end{bmatrix} \end{matrix}; \quad \begin{matrix} 1 & \begin{bmatrix} c_1 & c_2 & e_1 \\ 3 & 4 & 3 \end{bmatrix} \\ 2 & \begin{bmatrix} 3 & 3 & 2 \end{bmatrix} \\ 3 & \begin{bmatrix} 3 & 3 & 2 \end{bmatrix} \\ 4 & \begin{bmatrix} 4 & 2 & 1 \end{bmatrix} \\ 5 & \begin{bmatrix} 5 & 4 & 0 \end{bmatrix} \end{matrix}
 \end{aligned}$$

The numerals in the left of the matrix represent the numbers of points in the space, the symbols above the matrix are the dimensions of the space, and the element of the matrix b_{ij} is the distance from i -group to j -dimension. That is, the values of the row vector in the matrix are the distance from a group to all dimensions, and the values of the column vector are the distance from all groups to a dimension. The matrixes show that the space position of point i in the graph can be represented by the m -dimension vector ($b_{i1}, b_{i2}, \dots, b_{im}$). Then the module α_i of point i vector is

$$\alpha_i = (\sum_{j=1}^m b_{ij}^2)^{1/2} \quad (i = 1, 2, 3, 4, 5)$$

Table 2. Comparison of Accuracy between the Widely Used Methods and the One Proposed for $\Delta_{vap}H(T_b)$ Estimation

method	Hoshino method ²	Ma method ³	the proposed method
no. of data	411	483	524
APE ^a	1.5	1.45	1.16

^a APE = average percent error.

The average square root of the module of some point i in the graph is defined as the module index ν_i of this point vector. That is

$$\nu_i = \alpha_i / (\sum_{j=1}^5 \alpha_j^2)^{1/2} \quad (i = 1, 2, 3, 4, 5)$$

The quantity ν_i is used to describe the point i position in the space. On the analogy of this, the module index ν_i of group i in the molecule is taken to characterize the position of that group in the molecular space. Thus, every simple group, except halogen groups, has its own independent module index. For the four halogen groups, their module indexes were determined to be the same as those of the hydrocarbon groups with which they were connected.

CORRELATION AND GROUP PARAMETERS

The expression of the physical property f for a simple group method is

$$f = a + \sum_i n_i \Delta f_i \quad (1)$$

where the subscript i represents the group type, Δf_i is the contribution value of the i -type group, n_i is the number of i -type groups in the molecule, and a is the correlation constant.

The group has an individual contribution because a different group has a different effect on the physical property; for two identical groups, whose positions in the molecule are different, they have different contributions, then the position contribution is induced to express the different position of groups in the molecule. So, in this study, the group contribution was divided into two parts: the indepen-

Table 3. Correlation Accuracy and the Error Distribution of $\Delta_{vap}H(T_b)$ for Six Groups of Organic Compounds

compound	$\Delta_{vap}H(T_b)/\text{kJ}\cdot\text{mol}^{-1}$			compound	$\Delta_{vap}H(T_b)/\text{kJ}\cdot\text{mol}^{-1}$		
	no. of data	absolute deviation	percent deviation (%)		no. of data	absolute deviation	percent deviation (%)
aliphatic hydrocarbon	124	0.202	0.902	O, S compounds	188	0.343	1.331
naphthenic hydrocarbon	40	0.215	0.952	N compounds	67	0.309	0.975
aromatic hydrocarbon	18	0.301	1.253	halogenated compounds	87	0.390	1.401
total					524	0.302	1.16

Table 4. Estimation Results of T_b and $\Delta_{vap}H(T_b)$ for 10 Kinds of Compounds Using This Proposed Method

$\Delta_{vap}H(T_b)/\text{kJ}\cdot\text{mol}^{-1}$						$\Delta_{vap}H(T_b)/\text{kJ}\cdot\text{mol}^{-1}$					
compound	exp data	proposed method		Ma method ³		compound	exp data	proposed method		Ma method ³	
		est data	% error	est data	% error			est data	% error	est data	% error
2,2-dimethylpentane	29.23	29.46	0.77	29.33	0.34	cyclooctane	35.90	35.87	−0.09	35.39	−1.42
2,2,4-trimethylpentane	30.79	30.92	0.42	30.87	0.27	methyl isopropanone	32.35	32.25	−0.03	32.08	−0.83
3-methylnonane	38.26	37.86	−1.05	37.49	−2.02	1,4-diethylbenzene	39.40	40.64	3.15	32.59	−17.3
2-methyl-1-butene	25.50	25.26	−0.91	25.60	−0.41	1-propanol	41.44	41.91	1.13	39.47	−4.75
propyl cyclopentane	34.70	34.62	−0.22	34.06	−1.84	isobutyl formate	33.60	34.52	2.75	34.30	2.07
average									1.05		3.13

dent contribution of the group and the position contribution of the group. To improve the estimation accuracy, the constant of the group is added to the correlation equation. The physical property f is expressed as follows

$$f = a + \sum_i (\Delta f_{0i} + n_i \Delta f_{li} \sum_{j=1}^{n_i} \nu_j \Delta f_{pi}) \quad (2)$$

where Δf_{0i} is the constant of the i -type group, Δf_{li} is the independent contribution of the i -type group, Δf_{pi} is the position contribution of the i -type group, $\sum_{j=1}^{n_i} \nu_j$ is the module index sum of the i -type groups.

Accordingly, the expression of the enthalpy of vaporization at the normal boiling point is written as follows

$$\Delta_{vap}H(T_b)/\text{kJ}\cdot\text{mol}^{-1} = 13.41 + \sum_i (\Delta_{vap}H_{0i} + n_i \Delta_{vap}H_{Li} + \sum_{j=1}^{n_i} \nu_j \Delta_{vap}H_{pi}) \quad (3)$$

where $\delta_{vap}h(T_b)$ is the enthalpy of vaporization at the normal boiling point, and the corresponding units employed are in $\text{kJ}\cdot\text{mol}^{-1}$.

A great deal of experimental data in the literature^{4,5} has been used to optimize the values of group parameters in eq 3. The simple method of least-squares data regression is used for optimization with the following objective function

$$OF = \sum_j \left| \frac{Q_j^{exp} - Q_j^{cal}}{Q_j^{exp}} \right| \quad (4)$$

where Q_j^{exp} and Q_j^{cal} are the experimental value and the calculation value of the enthalpy of vaporization at the normal boiling point of organic compounds. n is the number of compounds used in the regression. The group parameters obtained by correlating data are listed in Table 1.

METHOD COMPARISON AND ESTIMATION RESULTS

Extensive comparisons are made between eq 3 and the Hoshino method² and the Ma³ method. The result is listed in Table 2.

The enthalpy of vaporization $\Delta_{vap}H(T_b)$ at the normal boiling point is 524 compounds, which are divided into 6

groups; the average deviations between the calculated value by this method and the experimental data are listed in Table 3.

To test the applicability and reliability of the new methods, estimations of the new method for 10 compounds not used in the parameter regression have made. These results compare favorably with the lower deviation Ma³ method in Table 4.

The proposed method gives, as shown in Tables 2–4, more accurate predictions for $\Delta_{vap}H(T_b)$ than the compared methods.

CONCLUSION

The specific position of a group in the molecule has been considered, and a group vector space method for estimating the enthalpy of vaporization of organic compounds developed. Expression for enthalpy of vaporization $\Delta_{vap}H(T_b)$ has been proposed, with the numerical values of relative group parameters presented. The average percent deviation of estimation of the above property is 1.16, which shows that the proposed method demonstrates a significant improvement in applicability to predict $\Delta_{vap}H(T_b)$, compared to the conventional group methods, and the estimation does not need data of normal boiling points.

APPENDIX: ILLUSTRATIVE EXAMPLE OF ESTIMATION

The procedure for estimating the physical property of an organic compound can be readily performed. First, the structural formula was drawn and the numbering of groups made. The number of every group in the molecule and the signs of dimensions are random, but the numbers of a type of groups are continuous to make the calculation convenient. For every molecule, possibly there are several different numberings, which have no effect on the result of estimation. According to the structural formula, the vector space matrix for the molecule was drawn. Then, module α_i of group i in the molecule and corresponding module index ν_i may be calculated by the equation $\alpha_i = (\sum_{j=1}^m b_{ij}^2)^{1/2}$ and equation $\nu_i = \alpha_i / (\sum_{j=1}^m \alpha_j^2)^{1/2}$. Finally, substituting each $\sum_{j=1}^{n_i} \nu_j$ along with the values of the corresponding parameters found in Table 1 into a relative correlation, the estimation value of

Table 5. Results of $\Delta_{\text{vap}}H(T_b)$ Estimation for 1 Methyl Propyl Thioether

no.	α_i	ν_i	no.	α_i	ν_i	group	n_i	$\sum_{j=1}^{n_i} \nu_j$	$\Delta_{\text{vap}}H_{0i}$	$\Delta_{\text{vap}}H_{Li}$	$\Delta_{\text{vap}}H_{pi}$
1	4.000	0.516	4	2.828	0.365	-CH ₃	2	5.164	0.179	3.565	-2.707
2	4.000	0.516	5	3.162	0.408	>CH ₂	2	3.867	-0.291	1.362	3.818
3	3.162	0.408				-S-	1	2.041	1.832	0.715	15.410

Table 6. Results of $\Delta_{\text{vap}}H(T_b)$ Estimation for 4-Fluoro-1-methylbenzene

no.	α_i	ν_i	no.	α_i	ν_i	group	n_i	$\sum_{j=1}^{n_i} \nu_j$	$\Delta_{\text{vap}}H_{0i}$	$\Delta_{\text{vap}}H_{Li}$	$\Delta_{\text{vap}}H_{pi}$
1	7.000	0.369	5	6.708	0.354	-CH ₃	1	0.369	0.179	3.565	-2.707
2	6.083	0.321	6	6.325	0.333	(=C<) _R	2	0.701	-1.244	0.359	8.550
3	7.211	0.380	7	6.708	0.354	(=CH-) _R	4	1.374	-2.406	4.672	-3.293
4	6.325	0.333	8	7.211	0.380	-F-	1	0.380	0.189	0.112	1.259

the physical property can be obtained. Two examples are shown as follows:

Example 1. Estimate $\Delta_{\text{vap}}H(T_b)$ of methyl propyl thioether, its experimental value of $\Delta_{\text{vap}}H(T_b)$ is 32.08 kJ·mol⁻¹.

(a) The structural formula and the corresponding vector space matrix are drawn. For the two groups -CH₃, the numbers are continuous which are 1 and 2; for the two groups -CH₂-, the numbers are continuous which are 3 and 4. The vector of group 1 is (0,4) and that of group 2 is (4,0). In turn, the vectors of all groups in the molecule are written, which constitute in a matrix.

$$\begin{array}{cccccc}
 & & & & & \begin{matrix} e_1 & e_2 \end{matrix} \\
 & & & & & \begin{bmatrix} 1 & 0 & 4 \\ 2 & 4 & 0 \\ 3 & 3 & 1 \\ 4 & 2 & 2 \\ 5 & 1 & 3 \end{bmatrix} \\
 \begin{matrix} 2(e_2) & 3 & 4 & 5 & 1(e_1) \\ \text{CH}_3-\text{CH}_2-\text{CH}_2-\text{S}-\text{CH}_3 \end{matrix} & & & & &
 \end{array}$$

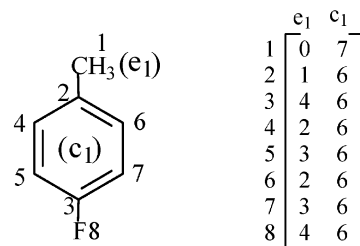
(b) The values of group parameters and the corresponding $\sum_{j=1}^{n_i} \nu_j$ are listed in Table 5.

(c) When the values of group parameters and the corresponding $\sum_{j=1}^{n_i} \nu_j$ in Table 5 are substituted into eq 3, the $\Delta_{\text{vap}}H(T_b)$ value estimated is 32.14 kJ·mol⁻¹. The relative deviation is 0.19%.

Example 2. Estimate $\Delta_{\text{vap}}H(T_b)$ of 4-fluoro-1-methylbenzene, its experimental value of $\Delta_{\text{vap}}H(T_b)$ is 33.90 kJ·mol⁻¹.

(a) The structural formula and the corresponding vector space matrix are drawn. According to example 1, the numbering and the vectors are written. Especially, the halogen group -F has its vector, which is the same as that of group 3. For the cyclic group 5, there are two routes to the e_1 dimension, but the shortest route should be selected, so its distance to the e_1 dimension is 3; the distance from

group 5 to the c_1 dimension is 6 which equals the number of points on the ring, so the coordinate of group 5 in the space is (3,6).



(b) The values of group parameters and corresponding $\sum_{j=1}^{n_i} \nu_j$ are listed in Table 6.

(c) When the values of group parameters and corresponding $\sum_{j=1}^{n_i} \nu_j$ in Table 6 are substituted into eq 3, the $\Delta_{\text{vap}}H(T_b)$ value estimated is 34.15 kJ·mol⁻¹. The relative deviation is 0.77%.

Supporting Information Available: Groups and their contributions of the Ma method and the data of enthalpy of vaporization used in this work. This material is available free of charge via the Internet at <http://pubs.acs.org>.

REFERENCES AND NOTES

- (1) Chen, N. C. Generalized correlation for Latent Heat of Vaporization. *J. Chem. Eng. Data* **1965**, *10*, 207-210.
- (2) Hoshino, D.; Nagahama, K.; Hirata, M. Prediction of the Entropy of Vaporization at the Normal Boiling Point by the Group Method. *Ind. Eng. Chem. Fundam.* **1983**, *22*, 430-433.
- (3) Ma, P.; Zhao, X. Modified Group Contribution Method for Predicting the Entropy of Vaporization at the Normal Boiling Point. *Ind. Eng. Chem. Res.* **1993**, *32*, 3180-3183.
- (4) Majer, V.; Svoboda, V. *Enthalpies of Vaporization of Organic Compounds*; Blackwell Scientific Publication: Oxford, 1985.
- (5) Dean, J. A. *Lange's Handbook of Chemistry*, 15th ed.; McGraw-Hill Book Co.: New York, 1999.

CI049888R