

-1-

CANADA

DEPARTMENT OF MINES AND TECHNICAL SURVEYS
MINES BRANCH

OTTAWA

Fuels Division
Technical Memorandum - 83/58-RBS

A NEW METHOD OF HYDROCARBON STRUCTURAL GROUP ANALYSIS

by

D.S. Montgomery and M. L. Boyd

This paper is published with the permission of the Director,
Mines Branch, Department of Mines and Technical Surveys,
Ottawa, Canada.

Paper Prepared for Presentation Before the Division of Gas
and Fuel Chemistry. American Chemical Society
Chicago, Illinois Meeting, September 7 - 12, 1958.

(Copy No. ___)

June 1, 1958.

A NEW METHOD OF HYDROCARBON
STRUCTURAL GROUP ANALYSIS

by

D.S. Montgomery and M.L. Boyd

Fuels Division, Mines Branch,
Department of Mines and Technical Surveys.

ABSTRACT

A method of hydrocarbon structural group analysis has been developed for application to pure compounds in which three chemical and two physical properties have been expressed in terms of five structural groups in a form which may be simultaneously solved by modern high speed computing equipment. The chemical properties include the carbon and hydrogen content as well as the number of aromatic carbon atoms present per molecule. The physical properties required for the analysis are the molar volume and molar refraction. This method has been tested on a selected group of 114 hydrocarbons whose properties have been determined by A. P. I. Project 42. The results of the application of this structural analysis system are described in detail and the accuracies attained have been tabulated.

PRESENTED BEFORE THE DIVISION OF GAS AND FUEL CHEMISTRY
AMERICAN CHEMICAL SOCIETY
Chicago, Illinois, Meeting, September 7-12, 1958.

by

D. S. Montgomery and M. L. Boyd
Department of Mines and Technical Surveys, Ottawa, Ontario.

INTRODUCTION

The synthesis and systematic study of the physical properties of pure hydrocarbons has been in progress for a number of years to facilitate the determination of the molecular structure of hydrocarbons directly or by analogy, that is by comparing certain properties of compounds of unknown structure with those of compounds whose structures are known.

The value of studying physical properties of a series of compounds, as a means of predicting the properties of unknown compounds and affording a means of checking the accuracy of the physical constants of compounds, has also been repeatedly demonstrated. In general, the study of the physical properties of hydrocarbons has been undertaken by a number of independent investigators and usually has been confined to hydrocarbons containing a limited number of types of structural groups, or, alternatively, only one physical property of a large variety of compounds has been examined.

This paper describes a method of simultaneously analyzing certain physical and chemical properties of liquid hydrocarbons to secure structural information. Although the system described is confined to specific classes of hydrocarbons, it is more general than any system so far proposed and involves the simultaneous consideration of three chemical and two physical properties to yield quantitative information concerning five structural groups. The investigation was promoted by a desire to improve existing structural analysis systems for pure hydrocarbons and to facilitate the study of naturally occurring hydrocarbons. The new possibilities offered by modern high speed computing equipment provided an additional incentive to re-examine and extend the earlier work in the field of structural analysis.

The most extensively used structural analysis system applied to hydrocarbon mixtures has been developed by Waterman and his school (13) beginning with the classical Waterman Ring Analysis of 1932 (19, 20) and extending to the n-d-M method of 1947 (14). van Krevelen employed some of the concepts of the Waterman Ring Analysis to develop a system that was particularly suited to the study of highly condensed aromatic structures which were assumed to be the major constituents of coal (7, 8). The chief criticism of the various methods of struc-

tural group analysis (15) of oil and coal is that when applied to pure compounds poor results are obtained. These considerations made it desirable to attempt to formulate a structural analysis system based on the physical properties of known compounds, but of such a form that it could be used with reasonable confidence to analyze the structure of high molecular weight material.

The method presented here was evolved from a method of carbon type analysis published by van Krevelen (4) in 1952. In this method, van Krevelen divided the carbon atoms in a structure into four main types:-

$$C_1 = \frac{CH}{C}; C_2 = \frac{CH_2}{C}; C_3 = \frac{CH_{aromatic}}{C}; C_4 = \frac{C}{C} \text{ aromatic};$$

where C in the denominator of each fraction represented the total number of carbon atoms per molecule. He then set up the following four quantitative relationships:

- (a) $C_1 + C_2 + C_3 + C_4 = 1$ (carbon balance)
- (b) $2C_1 + C_2 + C_3 = H/C$ (hydrogen balance)
- (c) $C_2 + C_4 = 2R/C$ (ring balance)
- (d) $C_3 + C_4 = fa$ (aromatic carbon balance)

These equations were not independent, and hence could not be solved simultaneously. van Krevelen solved these equations by giving an equation for C_1 as a function of H/C which represented the statistical probability of the occurrence of a CH_2 group in the molecule. The use of this equation was open to considerable question and consequently the method was never widely applied. Equations (a), (b) and (d), above, are true by definition; however, equation (c), the ring balance equation, is only valid for high molecular weight hydrocarbons where the factor $2/C$ can be neglected and where two junctions are associated with the formation of every ring. There are structures where this relation is not valid, such as in spiro compounds, and in three-dimensional structures where three rings possess a common side. van Krevelen's system was devised to elucidate the structure of coal and coal-like products where the proportion of saturated carbon atoms was small or negligible. While this choice of carbon types was suitable for the study of coal, it was undesirable for the study of petroleum since it failed to differentiate between chain and cyclic CH_2 groups. Hence, a five-type carbon classification was chosen which differed from that of van Krevelen by dividing his C_1 into two types. The carbon linkages have therefore been divided into the following five types.

C_1 = number per molecule of CH_3 , CH_2 , CH , and C groups in linear and branched chains.

- C_2 = number per molecule of CH_2 groups in saturated rings, including the case where the hydrogen atoms may be replaced by branched or linear chains.
- C_3 = number per molecule of CH groups which are junctions in fused saturated rings, as well as similarly situated groups where the hydrogen is replaced by linear or branched chains.
- C_4 = number per molecule of CH groups in aromatic rings, including the case where the hydrogen may be replaced by branched or linear chains.
- C_5 = number per molecule C groups which are junctions in fused aromatic rings, as well as junctions between saturated and aromatic rings.

It is possible to rewrite the carbon balance, the hydrogen balance and the aromatic carbon balance equations in terms of this new classification of structural groups. It should be especially noted that in the van Krevelen system the molecular weight was unknown and hence the structure was described in terms of fractions of the total number of carbon atoms. However, the present system was designed for the case where the molecular weight (as well as % carbon, % hydrogen, density, refractive index, and aromatic carbon content) was either known or could be determined. The carbon classification was therefore expressed in terms of the actual numbers of the different carbon types $C_1 - C_5$ present in the molecule. The Ring Balance Equation could not be used when analyzing an unknown hydrocarbon, as there was no accurate method of estimating the number of rings in the molecule. The fundamental basis of this structural analysis system, therefore, consisted in finding two additional physical properties which could be accurately expressed in terms of the above structural groups to give five independent equations which could then be solved simultaneously.

Theoretically, any two physical properties would be suitable, provided it were possible to express them as independent equations in terms of the structural types already defined. The two physical properties chosen were the molar volume and the molar refraction (the Lorentz-Lorenz expression). Both of these physical properties of liquids can be approximately described in terms of a linear combination of the atomic contributions, and both quantities have been extensively used for the purposes of elucidating structure (7, 18, 5, 10). It was assumed that the same functional form of the equation would apply to both physical properties, and that both properties could be expressed in terms of the same groups of chemical types. Owing to the intimate relation between these two quantities, it was felt that the loss in accuracy associated with grouping together so many different chemical linkages might easily yield two expressions which were merely linear combinations of each other. It remained for this investigation to demonstrate that the differences between the expression derived for the molar volume and that for the molar refraction were of such a magnitude that reliable structural information could be secured by the simultaneous solution of these expressions.

DEVELOPMENT OF THE FIVE EQUATIONS

It should be emphasized at the outset that this method as presently constituted is applicable only to certain specific classes of compounds. In the division of the carbon groups into the various types, care was taken that the hydrogen balance equation should always be satisfied. It was recognized that this equation was not strictly true as defined for the normal and branched paraffins. This case will subsequently be discussed. The present system will deal with monocyclic compounds and fused ring compounds, both saturated aromatic, but will not include polycyclic non-fused structures. Spiro compounds and three-dimensional ring systems are also excluded. It can be shown that the contributions to the molar volume and molar refraction of junction atoms in polycyclic non-fused compounds are not identical with any of the five types dealt with here, but in fact, represent a 6th (saturated) and a 7th (aromatic) type of linkage. Another class of compounds excluded are those containing olefinic or acetylenic bonds. In this work only double bonds existing in aromatic rings have been considered.

The method depends upon the ability to measure the density and refractive index of hydrocarbons in the liquid state at 20° C. at one atmosphere pressure, or on the capacity to correct to this standard state measurements made under other conditions. Except where otherwise specified, the coefficients of the molar volume and molar refraction equations have been determined from the properties of the hydrocarbons prepared by A. P. I. Project 42 (16).

The Carbon, Hydrogen and Aromatic Carbon Balance Equations The carbon balance, hydrogen balance, and aromatic carbon balance can be written as follows, by definition:

$$C_1 + C_2 + C_3 + C_4 + C_5 = \epsilon C \quad (1)$$

$$2C_1 + 2C_2 + C_3 + C_4 = \epsilon H \quad (2)$$

$$C_4 + C_5 = \epsilon C_a \quad (3)$$

The total number of carbon atoms per molecule ϵC was calculated from the carbon analysis and the molecular weight. The total number of hydrogen atoms per molecule ϵH , was similarly calculated from the hydrogen analysis and the molecular weight. The total number of aromatic carbon atoms in the molecule, ϵC_a , was not quite so readily available, although it could be determined by spectroscopic means (1). For the purposes of the present investigation (testing the method on known compounds), direct measurement of this quantity was not required.

The Molar Volume Equation The following molar volume equation was developed for the purpose of this structural analysis system:

$$\begin{aligned} \text{M. V.} &= C_1(16.38 + \frac{30.61}{\xi C}) + C_2(13.20 + \frac{28.48}{\xi C}) + & (4) \\ \text{at} & & \\ 20^\circ\text{C.} & & \\ 1 \text{ atm.} & C_3(10.981 + \frac{20.679}{\xi C}) + C_4(12.406 + \frac{14.042}{\xi C} - 1.96 \frac{C_1}{\xi C} + 10.13 \frac{C_2}{\xi C}) \\ \text{press.} & + C_5(5.124 - \frac{5.238}{\xi C}) \end{aligned}$$

A detailed account of this equation is in preparation. However to clarify the basis of this method of structural analysis, a brief description of the manner in which this equation was developed will now be given.

Initially, on the basis of Kurtz and Sankin's (9) work, the following general form of the molar volume equation was assumed:

$$\text{M. V.} = v_1 C_1 + v_2 C_2 + v_3 C_3 + v_4 C_4 + v_5 C_5 + K \quad (5)$$

A preliminary investigation to evaluate the coefficients in this expression indicated several difficulties with this functional form. No system of constant coefficients v_1 to v_5 could be found for the five structural types covered by the present classification. This work also revealed that k was not constant.

To overcome these difficulties, the following approach was adopted. From the definition of ξC , in Equation 1, the following expression was derived by multiplication by k :

$$\frac{C_1 K}{\xi C} + \frac{C_2 K}{\xi C} + \frac{C_3 K}{\xi C} + \frac{C_4 K}{\xi C} + \frac{C_5 K}{\xi C} = K \quad (6)$$

When K in equation 5 was replaced by the value in equation 6, the following equation was obtained:

$$\begin{aligned} \text{M. V.} &= C_1(v_1 + \frac{K}{\xi C}) + C_2(v_2 + \frac{K}{\xi C}) + C_3(v_3 + \frac{K}{\xi C}) & (7) \\ &+ C_4(v_4 + \frac{K}{\xi C}) + C_5(v_5 + \frac{K}{\xi C}) \end{aligned}$$

To determine the coefficients of C_1 to C_5 , use was made of the fact that for a molecule containing only one species of carbon atom (C_i), Equation 7, assumed the following form, as C_i was equal to ξC :

$$\text{M. V.} = v_i C_i + K_i \quad (8)$$

When the molar volume was plotted against C_i , the resultant straight line had a slope of v_i and an intercept of k_i . Consequently, v_i represented the contribution to the molar volume of a C_i group in the presence of an infinite number of C_i groups.

It will subsequently be shown that the K's in each bracket are not identical, these have hence been designated k_1 to k_5 respectively. The quantity K in the above equation expresses the increase in the molar volume that takes place on reducing the molecular weight of a particular species.

Determination of v_1 and k_1

The molar volume for the 13 normal paraffins on the list of A. P. I. Project 42 (16) was plotted against the number of carbon atoms in the molecule. The equation of the resulting straight line was determined by the method of least squares and found to be as follows:

$$M. V. = 16.38C_1 + 30.61 \quad (9)$$

Determination of v_2 and k_2

These quantities were calculated in a manner similar to that used for the calculation of v_1 and k_1 , using the physical properties given by Ward and Kurtz (21) for cyclopentane, cyclohexane, cycloheptane and cyclooctane. The molar volume of these compounds could be expressed by the following equation:

$$M. V. = 13.20C_2 + 28.48 \quad (10)$$

Determination of v_4 and k_4

Some difficulties were encountered in obtaining suitable data for compounds containing only C_4 's. The following compounds were used: benzene (properties taken from Egloff (3), cyclooctatetraene (properties by Eccleston (2) et al), and cyclopentadiene (Ward and Kurtz (21). A word of explanation about the use of the cyclopentadiene is required. Since it contained one C_2 group its use here is, strictly speaking, not justified in view of the presence of two different types of groups, C_2 and C_4 . Due to the scarcity of data, the observed molar volume of the cyclopentadiene was taken and the contribution of one CH_2 group was deducted. This amounted to considering a hypothetical compound C_4H_4 having four C_4 groups and a molar volume of $\frac{M. V. (\text{cyclopentadiene}) - 1(13.20 + \frac{28.48}{5})}{5}$. The resulting least squares equation of the straight line through these three compounds was found to be:

$$M. V. = 12.406C_4 + 14.042 \quad (11)$$

Determination of v_3 and k_3

A slightly different method had to be used to determine v_3 and k_3 , since no compounds existed containing C_3 groups only. Compounds containing C_2 and C_3 groups had to be used, and for this reason the following procedure was

adopted. $\left[(M.V. \text{ observed} - C_2(13.20 + \frac{28.48}{\xi C})) / C_3 \right]$ was plotted against $1/\xi C$. The result of this plot was a straight line whose slope was k_3 and intercept v_3 . Using the following compounds from the A.P.I. List (16): bicyclooctane 543, decahydronaphthalene 569 and 570, perhydrofluorene 561, perhydropyrene 578 and perhydrochrysene 575, v_3 was determined to be 10.981 and k_3 to be 20.679. The coefficient of C_3 was undoubtedly a function of the ring size, but the values found for five- and six-membered rings were almost identical.

Determination of v_5 and k_5

These quantities were determined in exactly the same manner as that used for the determination of v_3 and k_3 . Some difficulties were experienced in obtaining suitable liquid state molar volume data at 20°C for fused ring aromatic compounds that contained only C_4 and C_5 . Liquid molar volume data at 20°C have been determined by Ubbelohde (11, 12) for several fused ring aromatic compounds. This author also determined the molar volume for the solid state at 20°C. By plotting $(M.V._1 - M.V._s)$ versus C_5/C_4 , it was possible to express this shrinkage in volume as a function of the degree of condensation. By means of this relationship, the solid state molar volume data of van Krevelen (17) for fused ring aromatic compounds was "converted" to the liquid state at 20°C. Using the data obtained in this manner for anthracene, chrysene, phenanthrene (Ubbelohde), and for dibenzanthracene, chrysene, pyrene and coronene (van Krevelen's data converted to the liquid state) v_5 was determined to be 5.124 and k_5 to be -5.238.

The molar volume equation thus obtained was used to calculate the molar volume for the appropriate classes of compounds on the A.P.I. List (16) and to compare the results with the experimental values. The results of this comparison showed that a study of the interactions between various types of structural groups had to be made. The most significant interactions were found when C_1 and C_4 occurred in the same molecule and when C_2 and C_4 occurred together. In the former case, data from Ward and Kurtz (21) were used to evaluate the magnitude of the interaction, and in the latter case A.P.I. data (16) were employed. In brief, the method of determining the functional form and magnitude of the interaction terms consisted of obtaining the difference between observed and calculated molar volume per C_4 group and plotting this difference against $C_1/\xi C$ in the first case and $C_2/\xi C$ in the second. Two straight lines were obtained, from which were derived the two terms in $C_1/\xi C$ and $C_2/\xi C$ which were added to the C_4 term in Equation 4. The accuracy with which this "corrected" equation predicted the molar volume of the A.P.I. (16) hydrocarbons is given in Table I.

Table I - Accuracy of the Authors' Molar Volume Equation Applied to A. P. I. 42 Hydrocarbons

Class of Compound	No. in Class	Average Value of $\frac{MV_{calc} - MV_{obs}}{MV_{obs}} \times 100$	Standard Deviation
n-paraffins	13	-0.02	0.07
br. -paraffins	38	+0.31	0.80
Monocyclic saturates	27	+0.10	0.33
Fused ring saturates	18	+0.66	0.80
Monocyclic aromatics	21	+0.03	0.19
Fused ring aromatics	13	- .46	0.89

The Molar Refraction Equation The following molar refraction equation was developed for the purpose of this analysis:

$$\begin{aligned}
 \text{M. R. at } 20^{\circ}\text{C. 1 atm. press.} &= C_1 \left(4.63 + \frac{2.314}{\epsilon C} \right) + C_2 \left(4.468 + \frac{0.868}{\epsilon C} - \frac{0.245C}{\epsilon C^2} \right) + \\
 &C_3 \left(3.693 + \frac{0.3395}{\epsilon C} \right) + \\
 &C_4 \left(4.5445 - \frac{1.021}{\epsilon C} - 0.396 \frac{C}{\epsilon C^2} - 5.701 \frac{C^2}{\epsilon C^3} \right) + \\
 &C_5 \left(5.734 - \frac{14.333}{\epsilon C} \right) \dots \dots \dots (12)
 \end{aligned}$$

In this equation, M. R. referred to the Lorentz-Lorenz expression for the molar refraction $\left(\frac{n^2 - 1}{n^2 + 2} \frac{M}{d} \right)$, where n was the refractive index for the sodium D line at 20°C. The coefficients in this equation were determined in exactly the same manner, and using the same compounds as in the molar volume equation. Attention is drawn to the fact that, in the determination of the C₅ term, the molar refraction data for the fused ring aromatic compounds of van Krevelen (17) refer to measurements made of the compounds in benzene solution and have been referred to by van Krevelen (6) as "hypothetical liquid state data". The accuracy with which this equation predicts the molar refraction of the A. P. I. 42 (16) hydrocarbons is indicated in Table II.

METHODS OF SOLUTION OF THE SET OF FIVE EQUATIONS

Equations 1, 2, 3, 4 and 12 constitute the three chemical and two physical property equations that, when solved simultaneously for C_1 , C_2 , C_3 , C_4 and C_5 , form the proposed system of analysis. The first three of these equations are linear but equations 4 and 12 are quadratic in C_1 and C_2 . This set of linear and quadratic equations was initially solved in the following manner. The solution was obtained by reducing the system to a linear form by substituting an initial value $C_1 = C_2 = 0$ in the non-linear terms. The resultant set of linear equations was then solved by the standard methods of matrix algebra.

Table II - Accuracy of the Authors' Molar Refraction Equation, Applied to A. P. I. 42 Hydrocarbons

Class of Compound	No. in Class	Average Value of $\frac{MR_{calc} - MR_{obs}}{MR_{obs}} \times 100$	Standard Deviation
n-paraffins	9	+0.012	0.056
br. -paraffins	37	+0.19	0.26
Monocyclic saturates	26	-0.01	0.20
Fused ring saturates	19	-0.008	0.26
Monocyclic aromatics	19	+0.16	0.51
Fused ring aromatics	8	+0.17	0.64

The new values of C_1 and C_2 from this solution were then substituted in the non-linear terms and another solution was obtained. This iterative procedure was repeated until two consecutive solutions were equal. The critical aspect of the solution was whether or not the iterative procedure would converge. From the practical point of view the rate of convergence was important. The number of iterations varied from three for the paraffins to thirty for some of the fused ring aromatic compounds.

Since the molar volume and the molar refraction equations were quadratic in C_1 and C_2 , in general there would be four roots. Since it was clear that the iterative procedure yielded only one root, it was desirable to obtain a method of solution which would give all the roots. For this reason the set of five equations was solved by a second method. It was possible, by simple algebraic re-arrangement of the three linear equations, to express C_3 , C_4 and C_5 in terms of C_1 and C_2 . These values for C_3 , C_4 and C_5 were then substituted

in the molar volume and molar refraction equations and the following two quadratic equations in C_1 and C_2 resulted:

$$\frac{1.96}{\xi C} C_1^2 + \frac{(0.077 \xi C - 1.96 \xi C_a - 8.17 C_2 - 1.96 \xi H - 9.349)}{\xi C} C_1 + (13) \\ (-15.193 C_2 - 11.479 \frac{C_2}{\xi C} + 10.13 \frac{C_2 \xi C_a}{\xi C} + 10.13 \frac{C_2 \xi H}{\xi C} - 10.13 \frac{C_2^2}{\xi C^2} + \\ 3.699 \xi C + 1.425 \xi C_a + 7.282 \xi H - 6.637 \frac{\xi C_a}{\xi C} + 19.280 \frac{\xi H}{\xi C} + \\ 1.399 - M.V.) = 0$$

and

$$\frac{0.396}{\xi C} C_1^2 + \frac{(2.516 \xi C + 5.852 C_2 - 0.396 \xi C_a - 0.396 \xi H - 11.338)}{\xi C} C_1 + (14) \\ + (7.666 C_2 - 12.784 \frac{C_2}{\xi C} - 5.70 \frac{C_2 \xi C_a}{\xi C} - 5.701 \frac{C_2 \xi H}{\xi C} + 5.701 \frac{C_2^2}{\xi C^2} \\ + 0.852 \xi C_a - 1.360 \frac{\xi C_a}{\xi C} - 1.190 \xi H + 13.312 \frac{\xi H}{\xi C} + 4.883 \xi C \\ - 12.973 - M.R.) = 0$$

These two quadratics were solved graphically by taking arbitrary values of C_2 and solving each equation for C_1 . These values of C_1 were plotted against C_2 and the intersection of the molar volume and molar refraction curves represented the roots of the system.

RESULTS

The iterative method of solution has been tested by application to 121 known compounds, of which 114 are from the list of properties of A. P. I. Project 42 (16) and seven are fused ring aromatic compounds whose properties were determined by van Krevelen (17). All calculations were made by the International Business Machines Company on a Type 650 Magnetic Drum Processing Machine. The results are shown in Table III. In this table, D_1 to D_5 refer to the differences between calculated and observed values of C_1 to C_5 (i. e. $[C_1(\text{calc.}) - C_1(\text{observed})]$).

Some explanation is required concerning the application of this method to the paraffins. For this class of compounds, Equation 2, the hydrogen balance, is not strictly true as given but should be:

$$2C_1 + 2C_2 + C_3 + C_4 = \xi H - 2 \quad (15)$$

TABLE III. ANALYSIS OF PURE HYDROCARBONS TO SHOW THE ACCURACY OF THE FIVE STRUCTURAL GROUP ANALYSIS SYSTEM

API Compd. No.	C ₁		D ₁		C ₂		D ₂		C ₃		D ₃		C ₄		D ₄		C ₅		D ₅		
	Observed	Calc.	Observed	Calc.	Observed	Calc.	Observed	Calc.	Observed	Calc.	Observed	Calc.	Observed	Calc.	Observed	Calc.	Observed	Calc.	Observed	Calc.	
Group I - n-paraffins																					
528	12	12.07	+0.07	0	-0.094	-0.094	0	0	0.025	+0.025	0	0	0.025	+0.025	0	0	-0.025	-0.025	0	0	
529	13	13.04	+0.04	0	-0.070	-0.070	0	0	0.030	+0.030	0	0	0.030	+0.030	0	0	-0.030	-0.030	0	0	
531	14	13.99	-0.01	0	-0.016	-0.016	0	0	0.026	+0.026	0	0	0.026	+0.026	0	0	-0.026	-0.026	0	0	
532	15	15.04	+0.04	0	-0.030	-0.030	0	0	-0.015	-0.015	0	0	-0.015	-0.015	0	0	0.015	+0.015	0	0	
534	16	15.99	-0.02	0	0.024	+0.024	0	0	-0.016	-0.016	0	0	-0.016	-0.016	0	0	0.016	+0.016	0	0	
535	17	17.00	0.0	0	-0.0084	-0.0084	0	0	0.0049	+0.0049	0	0	0.0049	+0.0049	0	0	-0.0049	-0.0049	0	0	
537	18	17.91	-0.09	0	0.10	+0.10	0	0	-0.012	-0.012	0	0	-0.012	-0.012	0	0	0.012	+0.012	0	0	
540	20	19.59	-0.01	0	0.031	+0.031	0	0	-0.021	-0.021	0	0	-0.021	-0.021	0	0	0.021	+0.021	0	0	
106	26	25.90	-0.10	0	0.12	+0.12	0	0	-0.020	-0.020	0	0	-0.020	-0.020	0	0	0.020	+0.020	0	0	
Group II - branched paraffins																					
1	26	25.68	-0.32	0	0.22	+0.22	0	0	0.099	+0.099	0	0	0.099	+0.099	0	0	-0.099	-0.099	0	0	
107	34	33.97	-0.03	0	0.039	+0.039	0	0	-0.0098	-0.0098	0	0	-0.0098	-0.0098	0	0	0.0098	+0.0098	0	0	
109	26	25.40	-0.60	0	0.69	+0.69	0	0	-0.088	-0.088	0	0	-0.089	-0.089	0	0	0.089	+0.089	0	0	
133	36	36.22	+0.22	0	-0.13	-0.13	0	0	-0.092	-0.092	0	0	-0.092	-0.092	0	0	0.092	+0.092	0	0	
163	23	22.90	-0.10	0	0.11	+0.11	0	0	-0.0030	-0.0030	0	0	-0.0030	-0.0030	0	0	0.0030	+0.0030	0	0	
164	34	34.19	+0.19	0	-0.12	-0.12	0	0	-0.070	-0.070	0	0	-0.070	-0.070	0	0	0.070	+0.070	0	0	
184	28	28.34	+0.34	0	-0.47	-0.47	0	0	0.14	+0.14	0	0	0.14	+0.14	0	0	-0.14	-0.14	0	0	
191	32	32.10	+0.10	0	-0.044	-0.044	0	0	-0.061	-0.061	0	0	-0.061	-0.061	0	0	0.061	+0.061	0	0	
2	26	25.78	-0.22	0	0.16	+0.16	0	0	0.067	+0.067	0	0	0.067	+0.067	0	0	-0.067	-0.067	0	0	
22	26	25.27	-0.73	0	0.63	+0.63	0	0	0.10	+0.10	0	0	0.10	+0.10	0	0	-0.10	-0.10	0	0	
23	26	24.62	-1.38	0	1.33	+1.33	0	0	0.053	+0.053	0	0	0.053	+0.053	0	0	-0.052	-0.052	0	0	
25	25	24.85	-0.15	0	0.072	+0.072	0	0	0.075	+0.075	0	0	0.075	+0.075	0	0	-0.075	-0.075	0	0	
27	26	25.82	-0.18	0	0.097	+0.097	0	0	0.080	+0.080	0	0	0.080	+0.080	0	0	-0.080	-0.080	0	0	
3	26	25.80	-0.20	0	0.13	+0.13	0	0	0.073	+0.073	0	0	0.073	+0.073	0	0	-0.073	-0.073	0	0	
4	26	25.46	-0.54	0	0.44	+0.44	0	0	0.098	+0.098	0	0	0.098	+0.098	0	0	-0.098	-0.098	0	0	
5	28	27.72	-0.28	0	0.21	+0.21	0	0	0.077	+0.077	0	0	0.077	+0.077	0	0	-0.077	-0.077	0	0	
500	19	18.68	-0.32	0	0.31	+0.31	0	0	0.0016	+0.0016	0	0	0.0016	+0.0016	0	0	-0.0016	-0.0016	0	0	
51	26	25.22	-0.18	0	0.097	+0.097	0	0	0.080	+0.080	0	0	0.080	+0.080	0	0	-0.080	-0.080	0	0	
510	20	19.34	-0.66	0	0.62	+0.62	0	0	0.040	+0.040	0	0	0.040	+0.040	0	0	-0.040	-0.040	0	0	
511	20	19.75	-0.25	0	0.27	+0.27	0	0	-0.027	-0.027	0	0	-0.027	-0.027	0	0	0.027	+0.027	0	0	
512	14	13.97	-0.03	0	0.012	+0.012	0	0	0.018	+0.018	0	0	0.018	+0.018	0	0	-0.018	-0.018	0	0	
53	26	24.91	-1.09	0	1.02	+1.02	0	0	0.068	+0.068	0	0	0.068	+0.068	0	0	-0.068	-0.068	0	0	
545	16	15.68	-0.32	0	0.30	+0.30	0	0	0.018	+0.018	0	0	0.018	+0.018	0	0	-0.018	-0.018	0	0	
546	13	12.66	-0.34	0	0.34	+0.34	0	0	-0.0030	-0.0030	0	0	-0.0030	-0.0030	0	0	0.0030	+0.0030	0	0	
549	10	9.66	-0.34	0	0.29	+0.29	0	0	0.052	+0.052	0	0	0.052	+0.052	0	0	-0.052	-0.052	0	0	
55	26	25.17	-0.83	0	0.75	+0.75	0	0	0.079	+0.079	0	0	0.079	+0.079	0	0	-0.079	-0.079	0	0	
554	21	20.86	-0.14	0	0.15	+0.15	0	0	-0.011	-0.011	0	0	-0.011	-0.011	0	0	0.011	+0.011	0	0	
556	14	10.79	-3.21	0	3.40	+3.40	0	0	-0.20	-0.20	0	0	-0.20	-0.20	0	0	0.20	+0.20	0	0	
557	18	17.29	-0.71	0	0.74	+0.74	0	0	-0.023	-0.023	0	0	-0.023	-0.023	0	0	0.023	+0.023	0	0	
6	30	29.70	-0.30	0	0.25	+0.25	0	0	0.042	+0.042	0	0	0.042	+0.042	0	0	-0.042	-0.042	0	0	
63	28	27.79	-0.21	0	0.12	+0.12	0	0	0.096	+0.096	0	0	0.096	+0.096	0	0	-0.096	-0.096	0	0	
69	26	25.94	-0.06	0	-0.055	-0.055	0	0	0.12	+0.12	0	0	0.12	+0.12	0	0	-0.12	-0.12	0	0	
7	32	31.74	-0.26	0	0.18	+0.18	0	0	0.077	+0.077	0	0	0.077	+0.077	0	0	-0.077	-0.077	0	0	
8	11	10.92	-0.08	0	0.032	+0.032	0	0	0.052	+0.052	0	0	0.052	+0.052	0	0	-0.052	-0.052	0	0	

API Compd. No.	C ₁		D ₁		C ₂		D ₂		C ₃		D ₃		C ₄		D ₄		C ₅		D ₅		
	Observed	Calc.	Observed	Calc.	Observed	Calc.	Observed	Calc.	Observed	Calc.	Observed	Calc.	Observed	Calc.	Observed	Calc.	Observed	Calc.	Observed	Calc.	
Group III - monochloric naphthalene																					
539	8	8.52	+0.52	6	5.52	-0.49	0	-0.029	-0.029	0	-0.029	-0.029	0	-0.029	-0.029	0	0.029	+0.029	0	+0.029	
528	8	7.68	-0.32	6	6.36	+0.36	0	-0.039	-0.039	0	-0.039	-0.039	0	-0.039	-0.039	0	0.039	+0.039	0	+0.039	
573	10	9.93	-0.07	5	5.07	-0.07	0	0.0039	+0.0039	0	0.0039	+0.0039	0	0.0039	+0.0039	0	-0.0039	-0.0039	0	-0.0039	
572	10	10.50	+0.50	6	5.54	-0.46	0	-0.032	-0.032	0	-0.032	-0.032	0	-0.032	-0.032	0	0.032	+0.032	0	+0.032	
582	16	13.53	-0.47	5	5.46	+0.46	0	0.011	+0.011	0	0.011	+0.011	0	0.011	+0.011	0	-0.011	-0.011	0	-0.011	
506	13	12.55	-0.45	6	6.48	+0.48	0	-0.025	-0.025	0	-0.025	-0.025	0	-0.025	-0.025	0	0.025	+0.025	0	+0.025	
509	17	16.80	-0.20	6	6.24	+0.24	0	-0.038	-0.038	0	-0.038	-0.038	0	-0.038	-0.038	0	0.038	+0.038	0	+0.038	
110	20	19.73	-0.27	5	5.26	+0.26	0	0.017	+0.017	0	0.017	+0.017	0	0.017	+0.017	0	-0.017	-0.017	0	-0.017	
88	19	19.15	+0.15	6	5.94	-0.16	0	0.016	+0.016	0	0.016	+0.016	0	0.016	+0.016	0	-0.016	-0.016	0	-0.016	
117	21	21.00	0.0	5	5.05	+0.05	0	-0.048	-0.048	0	-0.048	-0.048	0	-0.048	-0.048	0	0.048	+0.048	0	+0.048	
64	21	20.18	-0.82	5	5.76	+0.76	0	0.063	+0.063	0	0.063	+0.063	0	0.063	+0.063	0	-0.063	-0.063	0	-0.063	
100	20	20.60	+0.60	6	5.49	-0.51	0	-0.097	-0.097	0	-0.097	-0.097	0	-0.097	-0.097	0	0.097	+0.097	0	+0.097	
102	20	19.77	-0.23	6	6.29	+0.29	0	-0.055	-0.055	0	-0.055	-0.055	0	-0.055	-0.055	0	0.055	+0.055	0	+0.055	
75	20	19.40	-0.60	6	6.59	+0.59	0	0.052	+0.052	0	0.052	+0.052	0	0.052	+0.052	0	-0.052	-0.052	0	-0.052	
106	20	19.56	-0.44	6	6.48	+0.48	0	-0.044	-0.044	0	-0.044	-0.044	0	-0.044	-0.044	0	0.044	+0.044	0	+0.044	
76	20	19.58	-0.42	6	6.40	+0.40	0	0.025	+0.025	0	0.025	+0.025	0	0.025	+0.025	0	-0.025	-0.025	0	-0.025	
77	20	19.63	-0.37	6	6.34	+0.34	0	0.031	+0.031	0	0.031	+0.031	0	0.031	+0.031	0	-0.031	-0.031	0	-0.031	
78	20	19.60	-0.40	6	6.37	+0.37	0	0.028	+0.028	0	0.028	+0.028	0	0.028	+0.028	0	-0.028	-0.028	0	-0.028	
162	20	19.94	-0.06	6	6.10	+0.10	0	-0.035	-0.035	0	-0.035	-0.035	0	-0.035	-0.035	0	0.035	+0.035	0	+0.035	
74	22	21.73	-0.27	5	5.27	+0.27	0	-0.0008	-0.0008	0	-0.0008	-0.0008	0	-0.0008	-0.0008	0	0.0008	+0.0008	0	+0.0008	
60	21	20.62	-0.38	6	6.35	+0.35	0	0.037	+0.037	0	0.037	+0.037	0	0.037	+0.037	0	-0.037	-0.037	0	-0.037	
91	22	22.38	+0.38	6	5.70	-0.30	0	-0.074	-0.074	0	-0.074	-0.074	0	-0.074	-0.074	0	0.074	+0.074	0	+0.074	
160	24	24.81	+0.81	5	4.10	-0.90	0	0.099	+0.099	0	0.099	+0.099	0	0.099	+0.099	0	-0.099	-0.099	0	-0.099	
69	25	24.76	-0.24	6	6.23	+0.23	0	0.016	+0.016	0	0.016	+0.016	0	0.016	+0.016	0	-0.016	-0.016	0	-0.016	
136	29	29.02	+0.02	6	6.07	+0.07	0	-0.099	-0.099	0	-0.099	-0.099	0	-0.099	-0.099	0	0.099	+0.099	0	+0.099	
138	33	33.04	+0.04	6	6.07	+0.07	0	-0.10	-0.10	0	-0.10	-0.10	0	-0.10	-0.10	0	0.10	+0.10	0	+0.10	
Group IV - monochloric naphthalene																					
538	8	8.00	0	0	0.11	+0.11	0	-0.12	-0.12	6	5.88	-0.12	0	0.12	+0.12	0	0.12	+0.12	0	0.12	
513	8	8.06	+0.06	0	0.22	+0.22	0	-0.27	-0.27	6	5.73	-0.27	0	0.27	+0.27	0	0.27	+0.27	0	0.27	
571	10	9.96	-0.04	0	0.084	+0.084	0	-0.042	-0.042	6	5.96	-0.04	0	0.042	+0.042	0	0.042	+0.042	0	0.042	
521	13	18.86	-0.14	0	-0.061	-0.061	0	0.20	+0.20	6	6.20	+0.20	0	-0.20	-0.20	0	-0.20	-0.20	0	-0.20	
87	19	18.64	-0.36	0	0.64	+0.64	0	-0.28	-0.28	6	5.72	-0.28	0	0.28	+0.28	0	0.28	+0.28	0	0.28	
10	19	20.28	+1.28	0	-0.69	-0.69	0	-0.59	-0.59	6	5.61	-0.59	0	0.59	+0.59	0	0.59	+0.59	0	0.59	
99	20	19.92	-0.08	0	0.045	+0.045	0	0.037	+0.037	6	6.04	+0.04	0	-0.037	-0.037	0	-0.037	-0.037	0	-0.037	
101	20	19.71	-0.29	0	0.37	+0.37	0	-0.076	-0.076	6	5.92	-0.08	0	0.076	+0.076	0	0.076	+0.076	0	0.076	
79	20	19.66	-0.34	0	0.35	+0.35	0	-0.093	-0.093	6	5.99	-0.01	0	0.093	+0.093	0	0.093	+0.093	0	0.093	
103	20	19.72	-0.28	0	0.30	+0.30	0	-0.025	-0.025	6	5.98	-0.02	0	0.025	+0.025	0	0.025	+0.025	0	0.025	
80	20	19.65	-0.35	0	0.41	+0.41	0	-0.061	-0.061	6	5.94	-0.06	0	0.061	+0.061	0	0.061	+0.061	0	0.061	
81	20	19.78	-0.22	0	-0.099	-0.099	0	0.32	+0.32	6	6.32	+0.32	0	-0.32	-0.32	0	-0.32	-0.32	0	-0.32	
82	20	19.79	-0.21	0	-0.25	-0.25	0	0.46	+0.46	6	6.46	+0.46	0	-0.46	-0.46	0	-0.46	-0.46	0	-0.46	
152	20	20.31	+0.31	0	-0.65	-0.65	0	0.35	+0.35	6	6.34	+0.34	0	-0.35	-0.35	0	-0.35	-0.35	0	-0.35	
161	20	20.13	+0.13	0	-0.52	-0.52	0	0.39	+0.39	6	6.39	+0.39	0	-0.39	-0.39	0	-0.39	-0.39	0	-0.39	
9	21	20.79	-0.21	0	-0.30	-0.30	0	0.51	+0.51	6	6.51	+0.51	0	-0.51	-0.51	0	-0.51	-0.51	0	-0.51	
54	22	21.54	-0.46	0	1.45	+1.45	0	-0.99	-0.99	6	5.00	-1.00	0	0.99	+0.99	0	0.99	+0.99	0	0.99	
146	23	22.79	-0.21	0	0.94	+0.94	0	-0.73	-0.73	6	5.27	-0.73	0	0.73	+0.73	0	0.73	+0.73	0	0.73	
68	25	24.73	-0.27	0	0.095	+0.095	0	0.37	+0.37	6	6.37	+0.37	0	-0.37	-0.37	0	-0.37	-0.37	0	-0.37	

API Compd. No.	C ₁		D ₁	C ₂		D ₂	C ₃		D ₃	C ₄		D ₄	C ₅		D ₅
	Observed	Calc.		Observed	Calc.		Observed	Calc.		Observed	Calc.		Observed	Calc.	
Group V - fused ring naphthenes															
527	4	2.93	-1.07	10	11.11	+1.11	4	3.96	-0.04	0	-0.037	-0.037	0	0.037	+0.037
125	12	12.19	+0.19	10	9.97	-0.03	4	3.84	-0.16	0	-0.16	-0.16	0	0.16	+0.16
178	18	17.29	-0.71	6	6.75	+0.75	2	1.96	-0.04	0	-0.035	-0.035	0	0.035	+0.035
141	12	11.43	-0.57	10	10.69	+0.69	4	3.88	-0.12	0	-0.12	-0.12	0	0.12	+0.12
544	11	10.83	-0.17	8	8.25	+0.25	2	1.91	-0.09	0	-0.087	-0.087	0	0.087	+0.087
175	15	15.03	+0.03	8	8.09	+0.09	2	1.88	-0.12	0	-0.12	-0.12	0	0.12	+0.12
166	8	8.02	+0.02	12	12.08	+0.08	6	5.89	-0.11	0	-0.11	-0.11	0	0.11	+0.11
62	21	20.08	-0.92	8	8.97	+0.97	2	1.95	-0.05	0	-0.053	-0.053	0	0.053	+0.053
177	17	15.21	-1.79	8	9.87	+1.87	4	3.92	-0.08	0	-0.078	-0.078	0	0.078	+0.078
193	15	13.63	-1.37	8	9.48	+1.48	4	3.89	-0.11	0	-0.11	-0.11	0	0.11	+0.11
106	16	15.89	-0.11	7	7.15	+0.15	2	1.95	-0.05	0	-0.046	-0.046	0	0.046	+0.046
196	8	7.74	-0.26	11	10.45	-0.55	6	6.81	+0.81	0	-0.19	-0.19	0	0.19	+0.19
575	0	0.41	+0.41	12	11.60	-0.40	6	5.99	-0.01	0	-0.011	-0.011	0	0.011	+0.011
578	0	-0.80	-0.80	10	10.80	+0.80	6	6.00	0	0	0.00077	+0.00077	0	-0.00077	-0.00077
577	0	-0.44	-0.44	10	10.41	+0.41	6	6.02	+0.02	0	0.025	+0.025	0	-0.025	-0.025
561	0	-0.40	-0.40	9	9.37	+0.37	4	4.03	+0.03	0	0.034	+0.034	0	-0.034	-0.034
543	0	-0.344	-0.34	6	6.34	+0.34	2	2.01	+0.01	0	0.0064	+0.0064	0	-0.0064	-0.0064
Group VI - fused ring aromatics															
567	1	0.28	-0.72	0	-0.085	-0.085	0	0.81	+0.81	8	8.81	+0.81	2	1.19	-0.81
574	0	0.55	+0.55	10	9.56	-0.44	2	1.90	-0.10	2	1.90	-0.10	4	4.10	+0.10
559	11	10.07	-0.93	0	-0.47	-0.47	0	1.41	+1.41	8	9.41	+1.41	2	0.59	-1.41
16	16	15.93	-0.07	3	2.84	-0.16	0	0.23	+0.23	4	4.23	+0.23	2	1.77	-0.23
188	15	15.60	+0.60	4	3.48	-0.52	0	-0.073	-0.073	4	3.93	-0.073	2	2.07	+0.073
174	15	14.03	-0.97	0	-1.28	-1.28	0	2.24	+2.24	8	10.24	+2.24	2	-0.24	-2.24
165	8	7.67	-0.33	10	10.21	+0.21	2	2.12	+0.12	2	2.12	+0.12	4	3.88	-0.12
179	8	11.11	+3.11	4	1.22	-2.78	0	-0.33	-0.33	8	7.67	-0.33	6	6.33	-0.33
173	21	20.53	-0.47	4	4.71	+0.71	0	-0.24	-0.24	4	3.76	-0.24	2	2.24	+0.24
Group VII - fused ring aromatics (data from Van Krevelen(17))															
chrysenes	0	1.52	+1.52	0	1.18	+1.18	0	-2.70	-2.70	12	9.30	-2.70	6	8.70	+2.70
pyrene	0	0.11	+0.11	0	-0.014	-0.014	0	-0.095	-0.095	10	9.90	-0.10	6	6.10	+0.10
coronene	0	-2.76	-2.76	0	4.62	+4.62	0	-1.86	-1.86	12	10.14	-1.86	12	13.86	+1.86
fluoranthene	0	0.72	+0.72	1	2.43	+1.43	0	-2.16	-2.16	8	5.84	-2.16	4	6.16	+2.16
B-truxene	0	-4.06	-4.06	3	9.68	+6.68	0	-2.64	-2.64	12	9.36	-2.64	12	14.64	+2.64
fluoranthene	0	-0.54	-0.54	0	0.74	+0.74	0	-0.20	-0.20	10	9.80	-0.20	6	6.20	+0.20
dodecahydrotriphenylene	0	-5.71	-5.71	12	16.39	+4.39	0	1.32	+1.32	0	1.32	+1.32	6	4.68	-1.32

At the outset, the effect of this comparatively small change was not known and the original hydrogen balance equation was used. The results on the n-paraffins are shown in Table IV. An examination of these results indicated that, although in error, they followed a very recognizable pattern. Since results for all the normal and branched paraffins on the A.P.I. 42 (16) list gave this same pattern, it was found possible to have the program for the Type 650 computer written in such a way that when this pattern occurred the revised hydrogen balance equation 15, would be substituted in place of the original.

Table IV - Analysis of n-paraffins to Illustrate the Use of Uncorrected Hydrogen Balance Equation

A. P. I. 42 Compd. #	C ₁		C ₂		C ₃		C ₄		C ₅	
	obs.	calc.								
528	12	10.14	0	3.75	0	-1.89	0	0.11	0	-0.11
529	13	11.13	0	3.75	0	-1.88	0	0.12	0	-0.12
531	14	12.09	0	3.78	0	-1.87	0	0.12	0	-0.12
532	15	13.15	0	3.73	0	-1.89	0	0.11	0	-0.11
534	16	14.12	0	3.77	0	-1.89	0	0.11	0	-0.11
535	17	15.15	0	3.72	0	-1.87	0	0.13	0	-0.13
537	18	16.05	0	3.82	0	-1.88	0	0.12	0	-0.12
540	20	18.17	0	3.71	0	-1.88	0	0.12	0	-0.12
106	26	24.12	0	3.74	0	-1.86	0	0.13	0	-0.13

For simplicity and also to illustrate that the use of the revised hydrogen balance equation led to accurate results, the revised hydrogen balance (equation 15) was used in analyzing groups I and II (Paraffins) and the original hydrogen balance (equation 2) for the remainder. A summary of the accuracy of this analysis is given in Table V.

The graphical method of solution which has been described was applied to 17 representative compounds taken from the five major classes. The relatively small number of compounds examined by this method was due to the fact that this method was much more time-consuming than the iterative procedure. The nature of the roots was essentially the same in all cases, they all possessed only one real root less than $\sum C$. There was one exception to this rule, A.P.I. compound #133 possessed two real roots but one of these lay in the second quadrant, which has no physical meaning. One representative example of the graphical method of solution showing the intersection between the molar volume and molar refraction curves which yields the single root less than $\sum C$ is given in Figure 1.

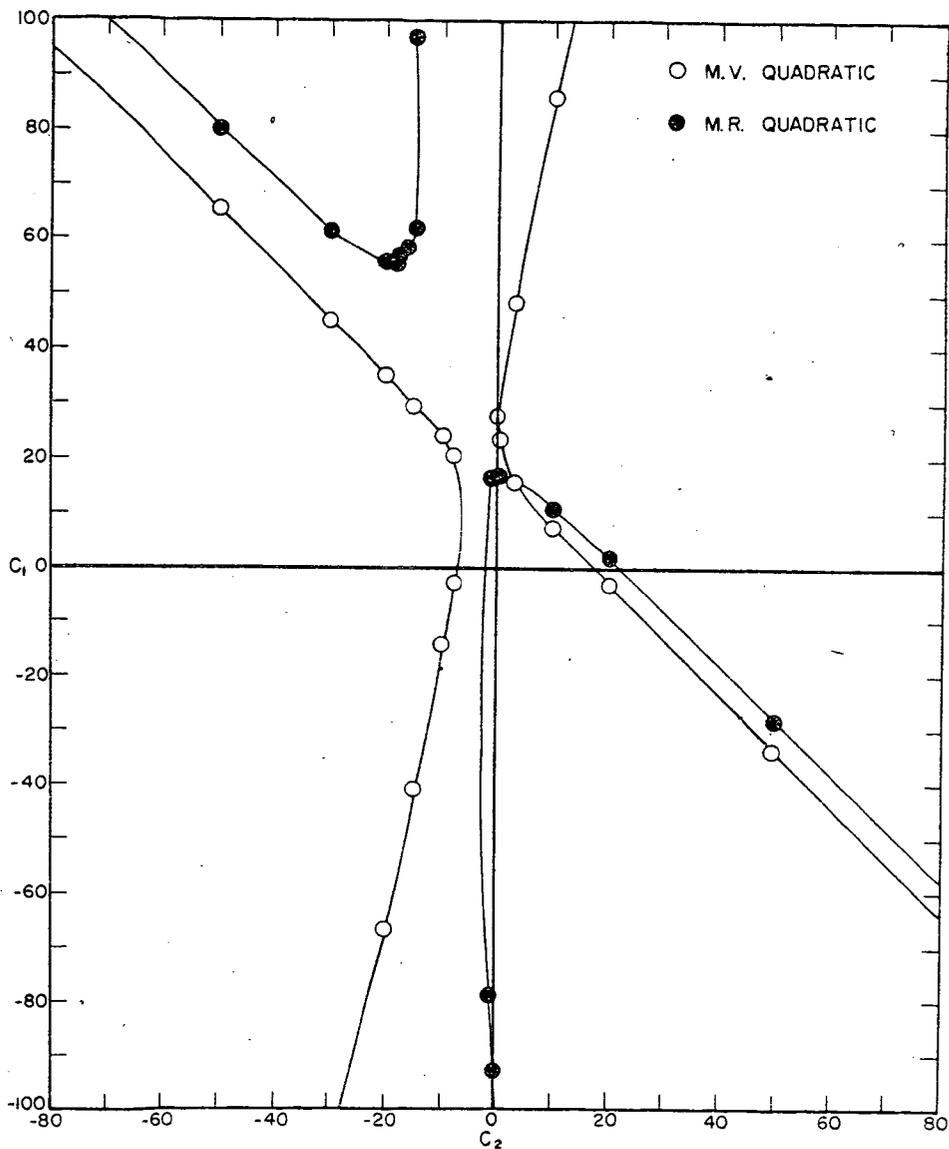


Figure I - Graphical Solution for A.P.I. Compound #16 to Illustrate the Existence of Only One Real Root.

DISCUSSION

Some additional information about the structure can be obtained by a detailed examination of the "pattern" of the results in Table III, within each of the seven groups of compounds into which this table is divided. This information can be used as a valuable aid when applying the method to unknown compounds. Before discussing the significance of the errors, it should be pointed out that for a given composition in terms of the five characteristic groups, there is a fixed relationship between the errors. If the error in C_1 is taken as x and the error in C_3 as y , then the error in C_2 will be $-(x + y)$, that in C_4 will be y and that in C_5 will be $-y$.

Considering, first of all, the branched paraffins, Group II, it will be noted that whereas C_2 should be zero, it is in fact positive and sometimes as high as +3. There appeared to be a direct relationship between the number of branches and the magnitude of the error. On the average, the error in C_2 per branch was +0.42 C_2 group. If, in the analysis of an unknown, C_2 , C_3 , C_4 and C_5 were less than 0.1, the compound could be predicted to be a normal paraffin with a high degree of certainty. However, if C_2 were between +0.1 and +0.4, then the compound would probably have one branch. On the other hand, if C_2 were greater than +0.4, the compound would probably have more than one branch.

The accuracy of the analysis of Group III, monocyclic naphthenes, was unusually high, due, no doubt, to the uniformity of the structures within the group. Little difficulty should be experienced in identifying this group in the analysis of unknown structures, since C_2 here must be greater than 3 (minimum ring size). A few structural effects were noted in the errors in C_1 and C_2 . The size of the ring and the number of branches on the ring would appear to be the effects having the most influence, whereas branching on a single side chain did not appreciably influence the results.

The accuracy of the results for Group IV, the monocyclic aromatics, was practically the same as for Group III, and for essentially the same reasons. The accuracy for compounds with a normal side chain was very high, as would be expected, since this was the type of compound upon which the interaction effect of C_1 and C_4 was based. Where one branched side chain was attached to the ring, the error in C_1 was about -0.3 and in C_2 about +0.3. The largest error for the entire group occurred for cases where there were more than one side chain attached to the ring. As with Group III, there should be no difficulty in recognizing this class of compound when analyzing unknown compounds, since C_4 should always be six in this class. At this level the errors were sufficiently small that no doubt should exist as to the structure.

Group V, the fused ring naphthenes, presented a much more difficult case to interpret. Since there appeared to be many more structural factors which could influence the analysis, and the examples available were very limited, no regular pattern between structure and the errors in the analysis could be determined. However, several important aspects of the errors associated with this

Table V - Summary of the Accuracy of the Hydrocarbon Analysis

Class of Comps.	No. in Class	\bar{D}_1	C_1	\bar{D}_2	C_2	\bar{D}_3	C_3	\bar{D}_4	C_4	\bar{D}_5	C_5
n-paraffins	9	-0.008	0.05	+0.006	0.07	+0.0002	0.02	+0.0002	0.02	-0.0002	0.02
br. -paraffins	34	-0.39	0.60	+0.36	0.63	+0.03	0.07	+0.03	0.07	-0.03	0.07
Monocyclic saturates	26	-0.11	0.40	+0.13	0.38	-0.01	0.04	-0.01	0.04	+0.01	0.04
Monocyclic aromatics	19	-0.09	0.37	+0.13	0.51	-0.03	0.40	-0.03	0.40	+0.03	0.40
Fused ring saturates	17	-0.49	0.56	+0.49	0.61	-0.006	0.21	-0.06	0.06	+0.06	0.06
A.P.I. fused ring aromatics	9	+0.09	1.20	-0.54	0.95	+0.45	0.82	+0.45	0.82	-0.45	0.82
Fused ring aromatics (van Krevelen)	7	-1.53	2.49	+2.72	2.25	-1.19	1.43	-1.19	1.43	+1.19	1.43

\bar{D} = arithmetic mean of C_i (calc) - C_i (obs)

σ = standard deviation of errors from the mean

$$\sigma = \sqrt{\frac{\sum (D_i - \bar{D}_i)^2}{N}}$$

group should be mentioned. The mean error and standard deviation were very small for C_3 , so that there should be no difficulty in detecting the presence of a relatively small number of C_3 groups in a saturated compound. An examination of the analysis of compounds 543, 561, 577 and 578 indicated that in cases where C_1 was actually zero the analysis yielded small negative values for C_1 ; when this occurred, if the value of C_1 was added to the value obtained for C_2 , the true value of C_2 was obtained. This suggested that in the analyses of unknown compounds, if a negative value of C_1 occurred, it should be replaced by zero and the negative value of C_1 should be added to C_2 to give the correct amount of C_2 present (#575 was an exception to this rule). Side chains and carbon atoms common to three rings appeared to have some influence on the accuracy of the analysis.

Group VI consisted of the A.P.I. fused ring aromatics. The error here was, of course, higher than for the preceding groups, but still small enough to enable the structural analysis system to be used. In this class, where the largest number of structural groups was present in a single compound, the largest number of interaction effects could be expected. The total number of compounds available in this group was so small that it was not possible to make any generalizations about the relationship between the errors in the predicted analysis and the structure. Attention is drawn to the analysis of compound #179, which exhibited the highest error in C_1 and C_2 . This result was not unexpected, as during the evaluation of the interaction effect of C_2 on C_4 it was observed that #179 did not fit the straightline relation on which this interaction factor was based.

Group VII consisted of fused ring aromatic compounds whose properties had been determined by van Krevelen (17). The errors in the predicted values of the structural groups in this class are considerably greater than in the other classes. It should be noted that the results for coronene and dodecahydrotriphenylene were not unexpected. When the original coefficients for C_5 were calculated, the data for coronene were not consistent with the other compounds and were therefore omitted. Consequently, it was expected that the accuracy of the analysis of this compound would not be very high. Dodecahydrotriphenylene represented a type of structure which the present system was not designed to treat. The unusual characteristic of this structure was the presence of C_2 and C_5 in the absence of C_4 . It will be recalled that fused ring aromatics containing C_2 groups were not successfully dealt with by determining the interaction effect of C_2 on C_4 . Clearly, this method of structural analysis was not applicable to the case where $C_4 = 0$ and C_2 and $C_5 = 0$. As in Group V, use could be made of the fact that negative values of C_1 usually indicated a true value for C_1 of zero, and that if this negative value of C_1 were added to the value obtained for C_2 , the resulting revised C_2 would be considerably closer to the true value. The rather high error for this group as a whole was considered to be due to the general unreliability of the physical property data, which has already been discussed. The accuracy of the molar volume and the molar refraction equations for this group was not as great as for all other classes. However, it was considered necessary to include this group, because of the scarcity of data for fused ring aromatic compounds.

The average accuracy for all A. P. I. hydrocarbons analyzed was considerably less than one carbon group for each type. This was considered adequate for determining the average number of the various structural groups.

It was realized, initially, that the molar volume and the molar refraction were very intimately related properties and that the equations developed in terms of the five structural groups might not be independent. However, this investigation revealed that these two equations were in fact independent, and that the differences between them were of such a magnitude that the solutions obtained were in substantial agreement with the known values of the groups present. Consequently, it was established that it was possible to deduce a considerable amount of structural information from these two closely related properties.

The results of the application of this method indicated the merits and some of the deficiencies of this type of structural analysis. As physical properties become available on "new" types of structures, the scope of this method may be extended. This approach could be applied to types of structures not already covered, by developing analogous equations for other physical properties and also by considering additional interaction effects. One of the merits of this system lay in the manner in which the coefficients in the molar volume and molar refraction equations were determined. This gave confidence in extrapolating beyond the molecular weight range of the known compounds that were used to establish the system.

- (1) Burdett, R. A., Taylor, L. W., Jones, L. C. Jr., "Molecular Spectroscopy". Report of a Conference Organized by the Spectroscopic Panel of the Hydrocarbon Research Group of the Institute of Petroleum, London, Oct. 28-29, 1954, p. 36.
- (2) Eccleston, B. H., Coleman, H. J., Adams, N. G.; J. Am. Chem. Soc. 72. 3866-70 (1950).
- (3) Egloff, G., Physical Constants of Hydrocarbons, Vol. III, 25, A. C. S. Monograph, Reinhold, New York, 1946.
- (4) Krevelen, D. W. van, Brennstoff-Chemie, 33, 260-8 (1952).
- (5) Ibid., 34, 167-82 (1953).
- (6) Krevelen, D. W. van, Blom, L., and Cherrmin, H. A. G., Nature 171. 1075-6 (1953).
- (7) Krevelen, D. W. van, and Cherrmin, H. A. G., Fuel, 33, 79-87 (1954).
- (8) Krevelen, D. W. van, and Schuyer, J., "Coal Science, Aspects of Coal Constitution", Chapters VI and VII, Elsevier, Amsterdam, 1957.
- (9) Kurtz, S. S. Jr., and Sankin, A., Ind. Eng. Chem. 46, 2186-91 (1954).
- (10) Lorentz, H. A., Ann Physik. 9. 641 (1880).
- (11) Mahdi, A. A. K. Al., and Ubbelohde, A. R., Proc. Roy. Soc. (London), 220A. 143-56 (1953).
- (12) Mahdi, A. A. K. Al., and Ubbelohde, A. R. Changements de Phases, Comptes-rendus de la deuxieme Reunion Annuelle, Soc. Chim. Phys., Paris, June 1952, pp. 360-65.
- (13) Nes, K. van, and Westen, H. A. van, "Aspects of the Constitution of Mineral Oils", pp. 299-314, Elsevier, Amsterdam, 1951.
- (14) Ibid., pp. 318-49 and pp. 445-53.
- (15) Ibid., Chapter IV.
- (16) Schiessler, R. W., and Whitmore, F. C., Ind. Eng. Chem. 47, 1660-5 (1956); Am. Doc. Inst., Doc. 4597.
- (17) Schuyer, J., Blom, L., and Krevelen, D. W. van. Trans. Far. Soc. 49, 1391-1401 (1953).

- (18) Schuyer, J., and Krevelen, D.W. van, *Fuel*, 33, 176-83 (1954).
- (19) Vlugter, J.C., Waterman, H.I., and Westen, H.A. van, *J. Inst. Petroleum Technol.* 18, 735-50 (1932).
- (20) Vlugter, J.C., Waterman, H.I., and Westen, H.A. van, *J. Inst. Petroleum Technol.* 21, 661-76 (1935).
- (21) Ward, A.L., and Kurtz, S.S. Jr., *Ind. Eng. Chem. Anal. Ed.* 10, 559-76 (1938).