ON THE DETERMINATION OF THE EMPIRICAL FORMULA OF A HYDROCARBON [1]

By Edward W. Washburn

ABSTRACT

This paper discusses the precision aspects of the problem of determining the molecular weight and hydrogen content of a hydrocarbon and of combining the results so as to obtain the empirical formula. In certain cases the formula can be deduced from the molecular weight alone, in others from the combustion analysis alone. Where both are required, the accuracy necessary in one or both is, in many cases, adjustable within rather wide limits and is determinable in any case. A definite laboratory procedure is outlined for obtaining the desired result with the minimum of effort and inconvenience. By following this procedure it should be possible to determine the empirical formula of any pure hydrocarbon containing not more than 100 carbon atoms. A determination of the bromine- (or other-) addition number may, in some instances be substituted for the molecular weight determination or for the combustion analysis, or may be utilized to decrease the accuracy which would otherwise be required in either or both of these determinations. The requirements necessary for the determination of a reliable "average formula" of a mixture of hydrocarbons are formulated. The influence of impurities and of polymerization is discussed.

CONTENTS

[1] This discussion has been prepared in order to establish a procedure for use with the hydrocarbons which are in course of fractionation from petroleum under Project No. 6 of the American Petroleum Institute entitled "The Separation, Identification, and Determination of the Chemical Constituents of Commercial Petroleum Fractions." Financial assistance in this project has been received from a research fund of the American Petroleum Institute donated by John D. Rockefeller. This fund is being administered by the institute with the cooperation of the Central Petroleum Committee of the National Research Council.

I. INTRODUCTION

1. THE PROBLEM

The determination of the empirical formula of a hydrocarbon ordinarily involves (1) a combustion analysis in order to ascertain the hydrogen content; together with (2) a molecular weight determination. The purpose of this paper is to present a critical discussion of the precision aspects of the problems involved in measuring these two quantities and in combining them so as to obtain the empirical formula of the hydrocarbon.

The precision aspects present certain unusual and interesting features owing to the fact that the functions which connect the molecular weight and the hydrogen content with the values of n and x in the general formula, C_nH_{2n+x} are not continuous functions and are consequently not amenable to treatment by the methods ordinarily employed in precision-of-measurement discussions. This lack of continuity arises from the conditions (1) that n must be a whole number and (2) that x must be an even whole number. These two conditions, together with the character of the atomic weights of carbon and hydrogen, require further that the molecular weight, except when n is very large, must also be close to an even whole number. The whole-number relations thus involved call for the application of certain new principles of measurement and calculation. A general treatment of these principles has been given in a previous paper,[2] to which frequent reference will be made in the course of the application of these principles to the problems before us.

2. LABORATORY PROCEDURE

Given a sample of a pure hydrocarbon whose formula is desired, the first problem which presents itself is the laboratory procedure to be followed; that is, which of the two quantities, (*a*) molecular weight and (*b*) hydrogen content, should be determined first and how accurately should this determination be made?

The answer to this question in any specific instance is likely to depend upon attendant circumstances. For example, if the problem arose in a laboratory already equipped with, and operating, an accurate combustion apparatus, the most convenient procedure might be to make an accurate combustion analysis first. With this accurate value available it might be found that the molecular weight determination could be dispensed with or that a very rough determination would suffice.

[2] The principles of Measurement and of Calculation in their Application to the Determination of Diophantine Quantities. B. S. Jour. Research, 4 p. 221; 1930.

On the other hand, if the laboratory had no combustion apparatus in working condition but was instead equipped with suitable apparatus for exact molecular weight determinations, the investigator would perhaps prefer to determine the molecular weight accurately, since by so doing he might be able to dispense entirely with the combustion analysis or at worst would require only a rough value for the hydrogen content of his hydrocarbon.

A third case would be represented by a laboratory in which neither of the above outfits happened to be available so that, if required, both would have to be constructed (or assembled) and standardized. In the following treatment we shall assume a situation corresponding to this third case. The procedure appropriate to each of the other cases will, however, be brought out in the course of the discussion.

If now the investigator has at hand and in working order neither an accurate molecular weight apparatus nor an accurate combustion equipment (or indeed, if he has both of them), it will usually be advantageous to first make a rapid and approximate molecular weight determination.

As soon as the molecular weight is known approximately, one can determine (*a*) whether a combustion analysis is required and, if so, with what degree of accuracy and/or (*b*) whether a more accurate molecular weight determination may be needed or preferred, and if so, with what degree of accuracy.

II. SYMBOLS AND ABBREVIATIONS

$\left.\begin{matrix} n \\ x \end{matrix}\right\}$, defined by the formula C_nH_{2n+x}.

M the true molecular weight $= 14.0156n + 1.0078x$.

$r = \frac{n}{x}$.

$h = \frac{1.0078\ (2n+x)}{M} = \frac{\%H}{100}$.

max., maximum.

min., minimum.

h_a, any experimental value found for h.

$(\delta h)_{max.}$, maximum absolute error for the technic employed.[3]

M_a, any experimental value found for M.

$p_{max.} = \frac{(\delta M)_{max.}}{M}$, maximum fractional error for the technic employed.[3]

$\Delta = 0.0078\ (2n+x)$.

(I), one of the positive integers.

III. THE MOLECULAR WEIGHT

1. MATHEMATICAL RELATIONS

The laws of valency and of atomic proportions give us the following relations:

1. n is a positive integer.

[3] See p. 223 of reference 2.

2. x is an even whole number lying between $+2$ and $\frac{14-M}{6}$, the latter value being the one which makes the formula of the hydrocarbon C_nH_2.

3. $$M=14.0156\,n+1.0078\,x \tag{1}$$

For convenience it is sometimes desirable to express M as a whole number plus the quantity Δ introduced by the fact that the atomic weight of hydrogen is not exactly unity: that is,

$$M=14n+x+\Delta \tag{2}$$

where

$$\Delta=0.0078\ (2n+x) \tag{3}$$

Eliminating n from (2) and (3) and solving for Δ

$$\Delta=0.0011\,M+0.0067x \tag{4}$$

The maximum value of x is 2, hence the maximum value of Δ is

$$\Delta_{max.}=0.0011M+0.0133 \tag{5}$$

or

$$\frac{100\Delta_{max.}}{M}=\left(0.11+\frac{1.33}{M}\right)\text{ per cent} \tag{6}$$

The minimum value of Δ is obviously 0.0156, which is that for all hydrocarbons of the formula C_nH_2.

Stated in another way, the molecular weight of a hydrocarbon is greater than some even whole number by an amount which is never more than 0.19 per cent of the number and, as will appear later, is too small to be significant in connection with the problem of determining the empirical formula. Δ may therefore be neglected in practically all cases.

2. DEDUCTIONS FROM THE MOLECULAR WEIGHT

(a) EVALUATION OF M

Add to the observed value (M_a) of M the maximum error in making the determination (or divide it by $1-p_{max.}$ where $p_{max.}$ is the maximum fractional error) and take the nearest even whole number which is smaller. This is an upper limit for $M-\Delta$. Subtract from the observed value of M the maximum error (or divide it by $1+p_{max.}$) and take the nearest even whole number which is larger. This is a lower limit for $M-\Delta$. The true value of $M-\Delta$ will be an even whole number lying between these limits. Column 1 of the "M-table" (Table 1) shows all of the possible values of $M-\Delta$ up to 310. It will be noted that while above 60, all even integers are possible values of $M-\Delta$, below 60 only certain ones are possibilities.

If $M-\Delta$ is to be definitely evaluated by the above procedure alone from any possible observed value of M, it is obvious that (with the single exception of methane) the molecular weight must be determined with a maximum error of less than one unit.[4]

[4] With an error considerably larger than this, evaluation may result for certain values of $M-\Delta$ less than 60, if the actual error is in a favorable direction. On this point see further pp. 231 and 232 of reference 2.

TABLE 1.—*The M-table*

[Formula of hydrocarbon, C_nH_{2n+x}. Molecular weight, $M=14n+x+\Delta$.[1]
[Percentage of hydrogen $=100h$. The table includes all hydrocarbons with molecular weight less than 312.]

$M-\Delta$[1]	n	h	$2n+x$	$M-\Delta$[1]	n	h	$2n+x$
16	1	0.2514	4	136	10	0.1184	16
					11	.0296	4
26	2	.0775	2	138	10	.1313	18
28	2	.1438	4		11	.0438	6
30	2	.2012	6	140	10	.1438	20
					11	.0576	8
38	3	.0530	2	142	10	.1559	22
40	3	.1007	4		11	.0709	10
42	3	.1438	6				
44	3	.1830	8	144	11	.0839	12
50	4	.0403	2	146	11	.0966	14
52	4	.0775	4		12	.0138	2
54	4	.1079	6	148	11	.1089	16
56	4	.1438	8		12	.0272	4
58	4	.1735	10	150	11	.1208	18
					12	.0403	6
62	5	.0325	2	152	11	.1325	20
64	5	.0630	4		12	.0530	8
66	5	.0915	6	154	11	.1438	22
68	5	.1184	8		12	.0654	10
70	5	.1438	10	156	11	.1548	24
72	5	.1677	12		12	.0775	12
74	6	.0272	2	158	12	.0892	14
					13	.0128	2
76	6	.0530	4	160	12	.1007	16
78	6	.0775	6		13	.0252	4
80	6	.1007	8	162	12	.1079	18
82	6	.1228	10		13	.0373	6
84	6	.1438	12	164	12	.1228	20
					13	.0491	8
86	6	.1638	14	166	12	.1334	22
	7	.0234	2		13	.0607	10
				168	12	.1438	24
88	7	.0458	4		13	.0719	12
90	7	.0672	6				
92	7	.0876	8	170	12	.1539	26
94	7	.1071	10		13	.0829	14
96	7	.1258	12		14	.0119	2
98	7	.1438	14	172	13	.0941	16
	8	.0206	2		14	.0234	4
100	7	.1610	16	174	13	.1041	18
	8	.0403	4		14	.0347	6
				176	13	.1144	20
102	8	.0593	6		14	.0458	8
104	8	.0775	8	178	13	.1244	22
106	8	.0950	10		14	.0566	10
108	8	.1079	12	180	13	.1343	24
					14	.0672	12
110	8	.1281	14				
	9	.0183	2	182	13	.1438	26
112	8	.1438	16		14	.0775	14
	9	.3598	4		15	.0111	2
114	8	.1589	18		13	.1532	28
	9	.0530	6	184	14	.0876	16
					15	.0219	4
116	9	.0695	8				
118	9	.0853	10	186	14	.0974	18
120	9	.1007	12		15	.0325	6
				188	14	.1071	20
122	9	.1155	14		15	.0429	8
	10	.0165	2	190	14	.1166	22
124	9	.1299	16		15	.0530	10
	10	.0325	4	192	14	.1258	24
126	9	.1438	18		15	.0630	12
	10	.0480	6				
128	9	.1573	20	194	14	.1349	26
	10	.0630	8		15	.0727	14
					16	.0104	2
130	10	.0775	10	196	14	.1438	28
132	10	.0915	12		15	.0822	16
					16	.0206	4
134	10	.1052	14	198	14	.1525	30
	11	.0150	2		15	.0915	18
					16	.0305	6

[1] $\Delta=0.0078\ (2n+x)$.

TABLE 1.—*The M-table*—Continued

$M-\Delta$	n	h	$2n+x$	$M-\Delta$	n	h	$2n+x$
200	15	0.1007	20	254	18	0.1506	38
	16	.0403	8		19	.1031	26
202	15	.1097	22		20	.0555	14
	16	.0499	10		21	.0079	2
204	15	.1184	24				
	16	.0593	12	256	19	.1101	28
					20	.0630	16
206	15	.1271	26		21	.0157	4
	16	.0685	14	258	19	.1171	30
	17	.0098	2		20	.0703	18
208	15	.1355	28		21	.0234	6
	16	.0775	16	260	19	.1239	32
	17	.0194	4		20	.0775	20
210	15	.1438	30		21	.0310	8
	16	.0863	18	262	19	.1306	34
	17	.0288	6		20	.0846	22
212	15	.1519	32		21	.0385	10
	16	.0950	20	264	19	.1373	36
	17	.0380	8		20	.0915	24
					21	.0458	12
214	16	.1035	22				
	17	.0471	10	266	19	.1438	38
216	16	.1079	24		20	.0984	26
	17	.0560	12		21	.0530	14
					22	.0076	2
218	16	.1200	26	268	19	.1502	40
	17	.0647	14		20	.1052	28
	18	.0092	2		21	.0601	16
220	16	.1281	28		22	.0150	4
	17	.0733	16	270	20	.1119	30
	18	.0183	4		21	.0672	18
222	16	.1360	30		22	.0224	6
	17	.0816	18	272	20	.1184	32
	18	.0272	6		21	.0741	20
224	16	.1438	32		22	.0296	8
	17	.0899	20	274	20	.1249	34
	18	.0598	8		21	.0809	22
226	16	.1514	34		22	.0368	10
	17	.0980	22	276	20	.1313	36
	18	.0446	10		21	.0876	24
228	17	.1060	24		22	.0438	12
	18	.0530	12	278	20	.1376	38
230	17	.1138	26		21	.0942	26
	18	.0613	14		22	.0507	14
	19	.0088	2		23	.0072	2
232	17	.1215	28	280	20	.1438	40
	18	.0695	16		21	.1007	28
	19	.0174	4		22	.0576	16
234	17	.1291	30		23	.0144	4
	18	.0775	18	282	20	.1499	42
	19	.0258	6		21	.1071	30
236	17	.1365	32		22	.0643	18
	18	.0853	20		23	.0214	6
	19	.0342	8	284	21	.1135	32
238	17	.1438	34		22	.0709	20
	18	.0931	22		23	.0284	8
	19	.0423	10	286	21	.1197	34
240	17	.1510	36		22	.0775	22
	18	.1007	24		23	.0352	10
	19	.0504	12	288	21	.1258	36
					22	.0839	24
242	18	.1082	26		23	.0420	12
	19	.0583	14	290	21	.1319	38
	20	.0083	2		22	.0902	26
244	18	.1155	28		23	.0486	14
	19	.0661	16		24	.0069	2
	20	.0165	4	292	21	.1072	40
246	18	.1228	30		22	.0966	28
	19	.0737	18		23	.0552	16
	20	.0246	6		24	.0133	4
248	18	.1298	32	294	21	.1438	42
	19	.0812	20		22	.1028	30
	20	.0325	8		23	.0617	18
250	18	.1369	34		24	.0206	6
	19	.0886	22	296	21	.1496	44
	20	.0403	10		22	.1089	32
252	18	.1438	36		23	.0681	20
	19	.0959	24		24	.0272	8
	20	.0480	12				

TABLE 1.—*The M-table*—Continued

$M-\Delta$	n	h	$2n+x$	$M-\Delta$	n	h	$2n+x$
298	22	0.1149	34	306	22	0.1382	42
	23	.0744	22		23	.0987	30
	24	.0034	10		24	.0592	18
300	22	.1208	36		25	.0198	6
	23	.0806	24	308	22	.1438	44
	24	.0403	12		23	.1046	32
302	22	.1267	38		24	.0654	20
	23	.0867	26		25	.0262	8
	24	.0467	14	310	22	.1494	46
	25	.0067	2		23	.1104	34
304	22	.1325	40		24	.0715	22
	23	.0927	28		25	.0325	10
	24	.0530	16				
	25	.0134	4				

(b) EVALUATION OF n AND x

Equation 1 may be written—

$$n=\frac{M-1.0078x}{14.0156} \tag{7}$$

If limiting values of M and x are known, they may be put into the equation to determine limiting values for n. Thus $n_{\text{min.}}$ is obtained by substituting $M_{\text{min.}}$ and $x_{\text{max.}}$ and taking the nearest integer which is larger. $n_{\text{max.}}$ is obtained by substituting $M_{\text{max.}}$ and $x_{\text{min.}}$ and taking the nearest integer which is smaller. If nothing is known about the value of x, $x_{\text{max.}}$ should be taken as $+2$ and $x_{\text{min.}}$ as $\frac{14-M_{\text{max.}}}{6}$. The possible values of n are the integers lying between $n_{\text{min.}}$ and $n_{\text{max.}}$

For the complete evaluation of n by the above procedure when x is unknown, the following condition, which is both necessary and sufficient, must be fulfilled. (See the M-table.)

The values of $M-\Delta$ falling within the range $M_{\text{min.}}$ to $M_{\text{max.}}$, inclusive, must be wholly within one of the following inclusive ranges: 16 to 16, 26 to 30, 38 to 44, 50 to 58, 62 to 72, 74 to 84, 88 to 96, 102 to 108, 116 to 120, 130 to 132, or 144 to 144. These values of $M-\Delta$ are printed in bold-face type in Table 1.

From equation (1) x can obviously be computed, if M and n are known, hence, a definite evaluation of $M-\Delta$ as one of the numbers in bold-face type completely determines the empirical formula of the compound.

If the molecular weight can not be identified as one of those corresponding to a single formula, a combustion analysis (or substitute therefor) will always be required except in the following special cases:

In the M-table two hydrocarbons appear with the molecular weight 86. One of these is the saturated hydrocarbon, C_6H_{14}. The other has the formula C_7H_2. This very unsaturated hydrocarbon is not known, and perhaps does not exist. In any event its properties would readily distinguish it from C_6H_{14}. For all practical purposes, therefore, the value 86 might be added to the list of bold-face values in the table. For similar reasons the same statement can be made

with respect to the molecular weights 98, 110, 122, 134, 146, and 158, and, with somewhat less confidence, with respect to a number of other values. In other words, whenever the M-table is used, certain of the hydrocarbons there shown may be eliminated as possibilities in a given case on the grounds that they could not have the properties possessed by the given hydrocarbon.

(c) ILLUSTRATIVE EXAMPLES

Example 1

Given: $M_a = 91$, $p_{max.} = 0.03$.

The true value of $M - \Delta$ must be an even integer lying between the limits

$$M_{max.} - \Delta > 91/0.97 > 93.8 = 92$$
$$M_{min.} - \Delta < 91/1.03 < 88.4 = 90$$

Hence, it is 90 or 92.

From the M-table it is obvious that $n = 7$ and that the hydrocarbon is C_7H_6 or C_7H_8. If we repeat the molecular weight determination without increase of accuracy and find $M_a = 93$, the possible values of $M - \Delta$ are now 92 and 94. The true value must, therefore, be 92 and the hydrocarbon is C_7H_8.

If we prefer to make the evaluation with the aid of a combustion analysis, we note (Table 1) that the possible values for h are 0.0672 and 0.0876. Hence, a combustion analysis accurate to better than $\frac{1}{2}$ $(0.0876 - 0.0672) = 0.01$ unit will suffice to definitely evaluate h.

Example 2

Given: Two determinations of M, $M_1 = 91$, $M_2 = 87$; and $p_{max.} = 0.1$.

$$M_{max.} - \Delta > 87/0.9 > 96.5 = 96$$
$$M_{min.} - \Delta < 91/1.1 < 82.7 = 84$$

From the M-table we find that n is either 6 or 7 and, if the hydrocarbon C_7H_2 is ruled out, definite evaluation is possible, if we redetermine M with sufficient accuracy. If, however, we prefer to make the evaluation from combustion analysis we prepare the following table:

$M-\Delta$	86	88	90	92	94	96	84	86
h	0.0234	0.0458	0.0672	0.0876	0.1071	0.1258	0.1438	0.1638
Δh	0.0224	0.0214	0.0204	0.0195	0.0187	0.0180	0.0200	

The smallest value for Δh is 0.018. Hence, it will suffice, if $(\delta h)_{max.} < 0.009$.

Example 3

Given:

$M_a = 301$ and $p_{max.} = 0.02$

$M_{max.} - \Delta > 301/98 > 306 = 306$

$M_{min.} - \Delta < 301/1.02 < 295 = 296$

The possible values of n are evidently 21, 22, 23, 24, and 25.

Arranging the corresponding values of h in the M-table in ascending order and computing the differences, we find the smallest difference to be $\Delta h = 0.0008$ for the hydrocarbons $C_{25}H_6$ and $C_{24}H_6$. These are very improbable hydrocarbons and might perhaps be ruled out. The next smallest value for Δh is 0.0024 for the hydrocarbons $C_{24}H_{18}$ and $C_{23}H_{18}$. To distinguish between these the error in the combustion analysis should preferably be less than $(\delta h)_{max.} = \frac{0.0024}{2} = 0.0012$. In other words, a careful combustion analysis must be made. Suppose the result is $h_a = 0.149 \pm 0.001$. Obviously $h = 0.1496$ and the hydrocarbon must be $C_{21}H_{44}$.

For all hydrocarbons with molecular weights of the order of 300 or larger, a careful combustion analysis is usually unavoidable. Whenever, therefore, $M_a/(1-p_{max.}) > 300$, it is best to proceed immediately with the combustion analysis and to use the methods to be explained below in place of the M-table for deducing the formula of the hydrocarbon.

We shall now take up the consideration of the combustion analysis and the conclusions which can be derived therefrom.

IV. THE COMBUSTION ANALYSIS

1. GENERAL CONSIDERATIONS

The purpose of the combustion analysis is to determine what fraction of the compound, by weight, is hydrogen. This fraction is represented by $h = \frac{\%H.}{100}$

This can be obtained (1) from the percentage of hydrogen alone, (2) from the percentage of carbon alone, or (3) by combining both values.

Method (1), which requires only a determination of the percentage of hydrogen, is the best of the three methods, if the sample is a pure hydrocarbon. Under these circumstances the carbon determination is unnecessary and of no value.

Method (2) would require an absolute accuracy in the carbon determination equal to that required in method (1) for the hydrogen determination. This practically eliminates this method from consideration.

Method (3) has the following advantages: (*a*) It is not necessary to know the mass of the sample used; (*b*) the result is not affected by the presence of impurities in the sample, except such as give volatile products which are absorbed; (*c*) it is also not affected by a partial oxidation of the sample, provided all of the oxidation products are retained by the sample; (*d*) through an almost exact compensation of air-buoyancy effects, it is unnecessary to correct the weighings to vacuum, if NaOH (or "Ascarite") is used to absorb the carbon dioxide and $MgClO_4.3H_2O$ ("Dehydrite") followed by P_2O_5 to absorb the water.

A comparison of the values of h as given by the three methods in the case of combustion analyses of naphthalene and of a petroleum fraction respectively, is shown in Table 2.[5]

[5] The data in this table are taken from the combustion analyses made by Bruun, B. S. Jour. Research, **2**, p. 487; 1929.

TABLE 2.—*Illustrating the value of* $h=\frac{\%H}{100}$ *as obtained by three different methods of calculation*

1. NAPHTHALENE

Run	$h_1=\frac{\%H}{100}$	Deviation	$h_2=\frac{100-\%C}{100}$	Deviation	$h_3=\frac{\%H}{\%H+\%C}$	Deviation
1	0.0627	0.0002	0.0639	0.0002	0.0627_7	0.0002_6
2	.0621	.0004	.0634	.0003	$.0621_8$	$.0003_3$
3	.0626	.0001	.0639	.0002	$.0626_3$	$.0001_7$
Average	.0625	.0002	.0637	.0002	$.0625_1$	$.0002_8$
Theoretical	.0629		.0629		.0629	

2. GAS-OIL FRACTION

1	0.1366	0.0002	0.1390	0.0002	0.1369	0.0002
2	.1364	.0000	.1395	.0003	.1368	.0001
3	.1361	.0003	.1392	.0001	.1365	.0002
Average	.1364	$.0001_7$	.1392	.0002	.1367	$.0001_7$
h_1	.1364	±.0002				
h_3	.1367	±.0002				
Difference	.0003	±.0003				

From these data it would appear to be a conservative conclusion to state that the value of h can, if necessary, be determined within ± 0.0005 unit; also a reasonable and safe value of the "maximum error of the method," $(\delta h)_{max.}$, is ± 0.001. This is obviously a degree of accuracy attainable without much difficulty and it will be found ample for practically all cases.[6]

2. MATHEMATICAL RELATIONS

From the formula, C_nH_{2n+x}, molecular weight $=M$, together with the values 12.000 and 1.0078 for the atomic weights of carbon and hydrogen, respectively, a number of mathmetical relations connecting h, n, x, and M can be easily derived by purely algebraic processes.

Those which will be employed [7] in the following discussion are

$$h=\frac{\%\ \mathrm{H}}{100}\ \text{(definition)}$$

$$r=\frac{n}{x}=\frac{1/2\left(\frac{1}{h}-1\right)}{6.95535-1/h} \tag{8}$$

$$x=M(1.15893h-0.16666) \tag{9}$$

[6] On this point see further Sec. XI, p. 890.

[7] Other relations, such as $n=\frac{(1-h)M}{12}$ and $2n+x=\frac{Mh}{1.0078}$ might alternatively be employed, and in some instances would be shorter and more direct. The ones adopted, however, and the procedures based upon them are applicable to all situations and yield the maximum amount of information with little chance of going astray.

3. CLASSIFICATION INTO TYPE GROUPS

In discussing the calculation of n and x from the experimentally determined quantities h and M, the various cases which present themselves fall naturally into three groups which can be defined as follows:

Group I, r is a positive quantity, $x=+2$.

Group II, r is infinite, $x=0$.

Group III, r and x are negative quantities.

The characteristics of each group, together with illustrative examples will be discussed in order. In this discussion $(\delta h)_{\text{max.}}$ will be taken as equal to ± 0.001 for the reasons explained above. The discussion can, therefore, be generalized by substituting $(\delta h)_{\text{max.}}$ for ± 0.001 wherever it is used.

4. Group I. SATURATED HYDROCARBONS, C_nH_{2n+2}

(r is positive, $x=+2$)

While this group is characterized by a positive value for r, it is not necessary to calculate r in order to determine whether or not a given case belongs to the group. This can be determined directly from the value of h as follows: (1) It is obvious that $r\left(=\frac{n}{x}\right)$ can be a positive quantity only when $x=+2$; that is, only for a saturated hydrocarbon. (2) Every saturated hydrocarbon will have a larger value of h than any unsaturated hydrocarbon. (3) The highest value of h for an unsaturated hydrocarbon occurs in the hydrocarbon type, C_nH_{2n}, and is equal to $\frac{2\times 1.0078n}{14.0156n}=0.1438$.

Group I is therefore completely defined by the relation

$$h>0.1438$$

Whenever, therefore, the value of h is greater than 0.1438, the hydrocarbon must be a saturated one; that is, $x=+2$ and $n=2r$.

Example 1

Given, h (found) $=0.1559$, $(\delta h)_{\text{max.}}=0.001$.

Substituting in equation (8) gives

$$11>(2r=n)>9.3$$

The true value of n must, therefore, be 10 and the hydrocarbon is $C_{10}H_{22}$. No molecular weight determination is necessary.

Example 2

Given the following two experimental values for h, $h_1=0.1521$, $h_2=0.1508$, and $(\delta h)_{\text{max.}}=0.001$. Evidently we may write

$$h_{\text{max.}}=h_2+0.001=0.1518$$

$$h_{\text{min.}}=h_1-0.001=0.1511$$

From these two values and equation (8) we find

$$2r_{max.} = n_{max.} \gtrdot 16.16 = 16$$

$$2r_{min} = n_{min.} \lessdot 15.2 = 16$$

Hence, $n = 16$ and the hydrocarbon is $C_{16}H_{34}$.

If $(\delta h)_{max.} = 0.001$, no possible single value of h will lead to the evaluation of n for saturated hydrocarbons containing more than 10 carbon atoms, but if more than one determination of h is available such evaluation may result, if n does not exceed about 17.[8]

Example 3

Given

$$h \text{ (found)} = 0.1464, \ (\delta h)_{max.} = 0.001$$

Hence

$$75.0 > (2r = n) > 33.8$$

n must, therefore, be a whole number lying between 75 and 34, inclusive. To find its value a molecular weight determination is required.

The facts concerning saturated hydrocarbons may be summed up as follows:

1. If $(\delta h)_{max.}$ does not exceed 0.001, the formula of the hydrocarbon can always be derived from the combustion analysis alone, for all hydrocarbons up to and including C_7; and if a sufficiently favorable value of h is obtained, evaluation is possible up to and including C_{10}.

2. For saturated hydrocarbons between C_{10} and C_{17} the formula of the hydrocarbon may be derived from the combustion analysis alone, if two sufficiently favorable values of h are obtained. In general, however, a molecular weight determination will be desirable for all hydrocarbons above C_8 and will be required for all above C_{17}.

3. If it is known that the hydrocarbon is saturated, a combustion analysis is unnecessary, since the formula can be calculated from the molecular weight determination alone. (See equation (7).) Or stated in another way, the combustion analysis need only be accurate enough to show that h is definitely greater than 0.1438. A greater degree of accuracy is ordinarily of no real value (unless it is desired to avoid, where possible, the necessity of a molecular weight determination) since (*a*) for saturated hydrocarbons of low molecular weight, only a moderate degree of accuracy in the molecular weight is required in order to determine n by means of equation (7) alone; and since (*b*) for higher molecular weights the accuracy required in the molecular weight determination is not materially diminished by a more accurate knowledge of h.

Procedure for Group I

Compute $r_{min.}$ and $r_{max.}$ with the aid of equation (8). Then find

$$n_{min.} \lessdot 2r_{min.} = (I)_{min.} \tag{10}$$

and

$$n_{max.} \gtrdot 2r_{max.} = (I)_{max.} \tag{11}$$

[8] On this point see further the discussion on pp. 233 and 234 of reference 2.

If $(I)_{min.} = (I)_{max.}$, n is completely evaluated.

If $(I)_{max.} > (I)_{min.}$, and a molecular weight determination is required, the accuracy in M should be $(\delta M)_{max.} < 7$ units, in order to be certain of evaluating n from a single determination of M.

5. Group II. HYDROCARBONS OF THE TYPE C_nH_{2n}

$$(x = 0,\ r = \infty)$$

This group includes only, but all, hydrocarbons of the type C_nH_{2n}. For all members of the group $h = 0.1438$. If, therefore, the value 0.1438 is included within the limits

$$h\ (\text{found}) \pm [(\delta h)_{max.} = 0.001]$$

the hydrocarbon in all probability has the formula C_nH_{2n}. The only possible alternative is a hydrocarbon containing more than 62 carbon atoms.[9]

For all hydrocarbons belonging to this group the formula must be determined from the molecular weight. The combustion analysis is of no further help.

Procedure for Group II

Determine M and compute n from the relation

$$\frac{M_{max.}}{14.0156} > (n = I) > \frac{M_{min.}}{14.0156} \qquad (12)$$

In order to be certain of evaluating n from a single determination of M, the accuracy in M should be $(\delta M)_{max.} < 7$ units. If n is found to be greater than 62, the combustion analysis must be repeated with an accuracy sufficient to identify the group type with certainty.

6. Group III. HYDROCARBONS OF THE TYPE C_nH_{2n+x}

(x is negative, r is negative)

This group includes all hydrocarbons having negative values of x. The group is completely defined also by the relation

$$h < 0.1438$$

In order to determine n and x for members of this group both M and h are determined and utilized in the calculation. The details of the calculation are discussed in 8 below.

7. EVALUATION OF n AND x FROM CUMBUSTION ANALYSIS ALONE

(a) EVALUATION OF x

The complete evaluation of x from combustion analysis alone is, in general, possible only when $h \geq 0.1438$. As we have shown above (p. 877), when $h = 0.1438$, $x = 0$ and when $h > 0.1438$, $x = +2$. In particular cases it is possible also for other values of h, provided some

[9] For $(\delta h)_{max.} = 0.0005$ the only alternative is a hydrocarbon with more than 123 carbon atoms.

limitation can be placed upon n. Since, however, the knowledge of such a limitation is practically equivalent to an approximate knowledge of the molecular weight, or at all events of an upper limit for it, such cases are best treated in connection with molecular weight data.

(b) EVALUATION OF n

Complete evaluation of n from combustion analysis alone is possible only when $h>0.1438$ and when n is less than a fixed value determined by and calculable from $(\delta h)_{max.}$ as in the examples already discussed (p. 878, *supra*).

(c) COMBINATION VALUES OF n AND x

If no limitation is placed upon the molecular weight, values of $h<0.1438$ give an infinite number of possible values for n and x. If, however, we agree to limit our field to hydrocarbons for which n is not greater than some fixed value, say 100, then for each value of h, or in practice, for each value of $h\pm(\delta h)_{max.}$, there is only a finite number of combination values possible for n and x and these are all calculable.[10] The calculation is made by computing $r_{max.}$ and $r_{min.}$ with the aid of equation (8) and then determining the possible combination values of n and x by Diophantine analysis. Since, however, this set of possible combination values is usually rather large, if h only is known, it is of practical interest only when M can not be determined. The more restricted set which is limited by an approximate knowledge of M is readily calculable by the methods which will now be described.

8. PROCEDURE FOR GROUP III

1. Determine $M_{max.}$ and $M_{min.}$ as directed in Section III, 2 (a).
2. Determine $x_{max.}$ and $x_{min.}$ from relation (9) which gives

$$-x_{max.}\geqslant M_{max.}\,[0.16666-1.1589\,(h-(\delta h)_{max.})]=2\,(I) \quad (13)$$

$$-x_{min.}\leqslant M_{min.}\,[0.16666-1.1589\,(h+(\delta h)_{max.})]=2\,(I) \quad (14)$$

3. Determine $n_{max.}$ and $n_{min.}$ from equation (7) which gives with sufficient accuracy

$$n_{max.}\geqslant \frac{M_{max.}-x_{min.}}{14}=(I)_{max.} \quad (15)$$

$$n_{min.}\leqslant \frac{M_{min.}-x_{max.}}{14}=(I)_{min.} \quad (16)$$

4. If n and x are not evaluated at this point, determine $-r_{max.}$ and $-r_{min.}$ from relation (8).
5. Using in turn each possible value of x as obtained from relations (13) and (14) above, determine the integers lying between $x\times r_{max.}$ and $x\times r_{min.}$ These integers, together with their corresponding values of x, constitute possible combination values of n and x.
6. From the set of combination values thus obtained strike out any which are inconsistent with relations (15) and (16) and tabulate the hydrocarbon formulas of the remainder, together with their molecular weights and values of h.

[10] See p. 241 of reference 2.

7. Determine by inspection of this table the next step in the procedure.

This next step is rather difficult to set forth in explicit terms, since it varies so greatly with the nature of the table obtained in (6) and with the desires of the investigator. It can best be presented by means of concrete examples.

If more than one experimental value of M and/or of h is available, the procedure just outlined may be modified accordingly. This can also best be presented by means of concrete examples.

Example 1

Given:

$M_a = 513$; $p_{max} = 0.1$

$h_a = 0.1398$; $(\delta h)_{max.} = 0.001$

$M_{min.} < 513/1.1 < 466 = 466 + \Delta$

$M_{max.} > 513/0.9 > 570 = 570 + \Delta$

The small quantity Δ may be neglected.

$$-x_{max.} > 570\ [0.1666 - 1.1589 \times 0.1388] > 3.2 = 2$$

$$-x_{min.} < 466\ [0.1666 - 1.1589 \times 0.1408] < 0.98 = 2$$

Hence $x = -2$

$$n_{min.} < \frac{466+2}{14} < 33.4 = 34$$

$$n_{max.} > \frac{570+2}{14} > 40.8 = 40$$

$$-r_{min.} = -1/2\left(\frac{1}{0.1388} - 1\right) \Big/ \left(6.955 - \frac{1}{0.1388}\right) = 12.4$$

$$-r_{max.} = -1/2\left(\frac{1}{0.1408} - 1\right) \Big/ \left(6.955 - \frac{1}{0.1408}\right) = 20.6$$

For $x = -2$ these give $n_{max.} = 41$ and $n_{min.} = 25$, which are wider limits than the above. The hydrocarbon, therefore, belongs to the type C_nH_{2n-2} and n must be between 34 and 40, inclusive. For this type the interval, ΔM, is constant and equal to 14.0156.

The possible hydrocarbons are therefore the following:

F	M	h
$C_{34}H_{66}$	474.508	0.1402
$C_{35}H_{68}$	488.523	.1404
$C_{36}H_{70}$	502.539	.1405
$C_{37}H_{72}$	516.554	.1405
$C_{38}H_{74}$	530.570	.1405
$C_{39}H_{76}$	544.585	.1406
$C_{40}H_{78}$	558.600	.1406

There is evidently nothing to be gained by repeating the combustion analysis. To be certain of identifying the hydrocarbon from one additional M determination it is evident that $(\delta M)_{max.}$ must be < 7 units, or, for the most unfavorable case, $p_{max.}$ must be less than 1.27 per cent. This is obtained from equation (32), page 235 of the preceding paper,[11] which for this case yields

[11] See footnote 2, p. 868.

$$p_{\max.}=\frac{\Delta M}{2M_{\max.}-14}=\frac{1}{2n_{\max.}-\frac{9}{7}}=\frac{1}{2\times 30-\frac{9}{7}}=1.27 \text{ per cent}$$

Example 2

Given:

$M_a=300$, $p_{\max.}=0.1$
$h_a=0.0770$ $(\delta h)_{\max.}=0.001$
$M_{\min.}\lessdot 300/1.1\lessdot 272.6=274+\Delta$
$M_{\max.}\gtrdot 300/0.9\gtrdot 333.3=332+\Delta$
$-x_{\max.}\gtrdot 332\ [0.1666-1.1589\times 0.0760]\gtrdot 26.1=26$
$-x_{\min.}\lessdot 274\ [0.1666-1.1589\times 0.0780]\lessdot 20.9=22$

$$n_{\min.}\lessdot\frac{274+22}{14}\lessdot 21.1=22$$

$$n_{\max.}\gtrdot\frac{332+26}{14}\gtrdot 25.6=24$$

$-r_{\min}=0.98$; $-r_{\max}=1.01$

Hence, $1.01>\frac{n}{-x}>0.98$

For

$x=$	$n=$
-22	22
-24	24
-26	26

The hydrocarbon must, therefore, belong to the type C_nH_n, for which $\Delta M=26$. The possibilities are

F	$-x$	M	h
$C_{22}H_{22}$	22	286.2	0.0775
$C_{24}H_{24}$	24	312.2	

To be certain of identifying the hydrocarbon by one additional M determination, it is evident [12] that

$$p_{\max.} \text{ must be} < \frac{26}{286+312} < 4.3 \text{ per cent}$$

Example 3

Given:

$M_a=302$, $p_{\max.}=0.1$
$h_a=0.0530$, $(\delta h)_{\max.}=0.001$
$M_{\min.}\lessdot 302/1.1\lessdot 274.5=276$
$M_{\max.}\gtrdot 302/0.9\gtrdot 335.5=334$
$-x_{\max.}\gtrdot 334\ [0.1666-1.1589\times 0.0520]\gtrdot 35.5=34$
$-x_{\min.}\lessdot 274\ [0.1666-1.1589\times 0.0540]\lessdot 28.7=30$
$n_{\min.}\lessdot \frac{276+30}{14}\lessdot 21.8=22$

[12] See p. 235 of reference 2.

$$n_{max.} > \frac{334+34}{14} > 26.3 = 26$$

$$-r_{min.} = 0.742;\ -r_{max.} = 0.757$$

$$0.757 > \frac{n}{-x} > 0.742$$

For

$x = -30$	$n =$ no possible integer
-32	$= 24$
-34	$=$ no possible integer

Hence, $n = 24$, $x = -32$ and the hydrocarbon is $C_{24}H_{16}$.

With the same values of h_a and $(\delta h)_{max.}$, the same result would have been obtained even though the maximum error in the molecular weight determination had been larger, as long as $M_{max.}$ was found to be <342 and $M_{min.}>266$. Similarly, with the same values of M_a and $p_{max.}$ the same result would have been obtained as long as $h \pm (\delta h)_{max.}$ was included within the limits 0.053 ± 0.002. Furthermore, by increasing the accuracy in the M determination the accuracy required in the h determination could be further materially lessened.

Example 4

In order to avoid the possibility of "mistakes,"[13] the investigator will usually run at least two combustion analyses, and, since M is being measured by a rapid method, at least two determinations of M might just as well be made. The following example illustrates a method which may be followed when more than one experimental value of h and/or of M is available.

Given:

$$M_1 = 542;\ M_2 = 539;\ p_{max.} = 0.1$$

$$h_1 = 0.0908;\ h_2 = 0.0916;\ (\delta h)_{max.} = 0.001.$$

$$M_{min.} < 542/1.1 < 492.27 = 492$$

$$M_{max.} > 539/0.9 > 599 = 598$$

$$-x_{max.} > 598\ [0.1666 - 1.1589 \times 0.0906] > 36.8 = 36$$

$$-x_{min.} < 492\ [0.1666 - 1.1589 \times 0.0918] < 29.6 = 30$$

$$n_{min.} < \frac{492+30}{14} < 37.3 = 38 \quad (a)$$

$$n_{max.} > \frac{598+36}{14} > 45.3 = 45$$

$$-r_{max.} = 1.256;\ -r_{min.} = 1.227$$

Hence $1.256 > \frac{n}{-x} > 1.227$

For

$x = -30$	$n = 37$
-32	$= 40$
-34	$= 42$
-36	$= 45$

[13] See p. 225, 237 of reference 2.

The first set is ruled out by (*a*) above, leaving as possibilities:

F	$-x$	M	h
$C_{40}H_{48}$	32	528.34	0.0918
$C_{42}H_{50}$	34	554.38	.0909
$C_{45}H_{54}$	36	594.41	.0915

In order to identify the hydrocarbon it is obviously necessary to make additional measurements. What shall they be and how accurately must they be made? No definite single answer can be given to this question. For example, suppose the hydrocarbon were $C_{42}H_{50}$. With $(\delta h)_{max.}=0.001$ any value of h between 0.0899 and 0.0919 is an experimental possibility. If, therefore, the combustion analysis were repeated and any value not greater than 0.0904 were obtained, the hydrocarbon would be identified. If, however, the correct value, 0.0909, were obtained, identification would fail. Similarly, if the hydrocarbon were $C_{40}H_{48}$ and the M determination were repeated without increase of accuracy, identification would result, if the value obtained in the measurement were less than $0.9\times 554.4=499$. In other words, if the investigator is fortunate in the errors which he makes, he will obtain the desired answer.[14]

While a definite answer can not be given to the question as formulated above, the following formulation, which is that used in the preceding examples, will yield such an answer: How accurately must M or h be measured in order that a single measurement of either will be certain to lead to identification?

For M we use equation (32) of the preceding paper.[15] This gives us

$$p_1=\frac{554.38-528.34}{554.38+528.34}=2.4 \text{ per cent.}$$

$$p_2=\frac{594.41-554.38}{594.41+554.38}=3.5 \text{ per cent.}$$

The answer is therefore $p_{max.}$ must be <2.4 per cent.

For h we note that $\Delta h_{min.}=0.0003$. Hence $(\delta h)_{max.}$ would have to be <0.00015.

It is evident that our most certain procedure is to repeat the M determination, with an accuracy better than 2.4 per cent, if practicable.

V. POSSIBLE SUBSTITUTES FOR THE COMBUSTION ANALYSIS OR THE MOLECULAR WEIGHT DETERMINATION

1. GENERAL CONSIDERATIONS

From the foregoing discussion it is evident that the purpose of the molecular weight determination and the combustion analysis is to provide us with two independent mathematical relations involving n and x, each relation being given with a known accuracy. These two relations, together with the Diophantine characters of n and

[14] See p. 231 of reference 2.

[15] See p. 235 of reference 2.

lead to the complete identification of both n and x, if the accuracy is sufficient.

Now it is obvious that any mathematical relation involving either or both n and x should be similarly utilizable, either as additional information or in place of the M or the h functions. Furthermore, any clean-cut chemical reaction, series of reactions or set of reactions in which the hydrocarbon is involved should, in principle, be capable of furnishing a mathematical relation of this character. The expression "clean cut" in this connection means simply that all molecules of the hydrocarbon react stoichiometrically alike. The desired relation is obtained by determining any stoichiometric quantity associated with the process. For example, in the combustion analysis itself the stoichiometric quantity determined is the number of mols of water produced per gram of hydrocarbon burned.

It is hardly worth while to discuss in detail the various chemical reactions involving hydrocarbons which might conceivably be used to supply the desired type of information. It will suffice to discuss one such case as an illustrative example.

Let us first assume that we have made the customary determinations of molecular weight and combustion analysis with the following results:

$$M_a = 129 \qquad p_{max.} = 0.2$$
$$h_a = 0.0766 \qquad (\delta h)_{max.} = 0.001$$

This is a Group III hydrocarbon. Hence, we have

$$M_{min.} \lessdot 129/1.2 \lessdot 107.5 = 108$$
$$M_{max.} \gtrdot 129/0.8 \gtrdot 161.2 = 160$$
$$-x_{max.} \gtrdot 160\ [0.1666 - 1.1589 \times 0.0756] \gtrdot 12.6 = 12$$
$$-x_{min.} \lessdot 108\ [0.1666 - 1.1589 \times 0.0776] \lessdot 8.3 = 10$$
$$n_{min.} \lessdot \frac{108+10}{14} \lessdot 8.42 = 9$$
$$n_{max.} \gtrdot \frac{160+12}{14} \gtrdot 12.3 = 12$$
$$-r_{min.} = 0.976; \quad -r_{max.} = 1.00$$
$$1.00 > \frac{n}{-x} > 0.976$$

For

$$x = -10 \qquad n = 10$$
$$-12 \qquad 12$$

The hydrocarbon is therefore of the type C_nH_n and is either $C_{10}H_{10}$ or $C_{12}H_{12}$.

2. UTILIZATION OF THE BROMINE-ADDITION NUMBER

Now let us assume that instead of making a combustion analysis we have determined the bromine- (or other-) addition number of the hydrocarbon. The following discussion is applicable to any addition reaction. Our data will be, let us say

$$M_a = 129 \qquad p_{max.} = 0.2$$
$$u_a = 0.0307 \qquad (\delta u)_{max.} = 0.0003$$

where u is the number of equivalents of bromine stoichiometrically added by 1 gram of the hydrocarbon.

If we let Z represent the number of equivalents of unsaturation [16] which remain in the molecule after the bromine addition, it follows from the laws of valency that

$$x=2-Mu-Z$$

where Mu and Z are even integers.

As in the preceding example

$$M_{max.}=160 \text{ and } M_{min.}=108$$

Hence

$$(Mu)_{max.} \gtrdot 160\times 0.031 \gtrdot 4.96=4$$
$$(Mu)_{min.} \lessdot 108\times 0.0304 \lessdot 3.31=4$$

And

$$Mu=4$$

Also

$$M_{max.} \gtrdot \frac{Mu}{u_{min.}} \gtrdot \frac{4}{0.0304} \gtrdot 131.5=130+\Delta$$

$$M_{min.} \lessdot \frac{4}{0.031} \lessdot 129=130+\Delta$$

Hence, $M-\Delta=130$ and the hydrocarbon must be $C_{10}H_{10}$.

Complete identification has resulted because we have assumed a sufficiently small value for $(\delta u)_{max.}$ If we had assumed a larger value, say $(\delta u)_{max.}=0.001$, we would have found 6 possibilities, namely, (C_9H_{20}), $C_{10}H_8$, $C_{10}H_{10}$, $C_{10}H_{12}$, $C_{10}H_{14}$, and $C_{11}H_{12}$. C_9H_{20} is ruled out because it is a saturated hydrocarbon. It will be noticed that $C_{12}H_{12}$ is not included among the possibilities.

Let us now assume that we have made only the combustion analysis and the bromine-addition determination. These will yield the following information, taking the same numerical data as in the preceding examples.

$$1.00>\frac{n}{-x}>0.976$$

$$0.031>u>0.0304$$

$$x=(2-Mu-Z)>\frac{14-M}{6}$$

$$M=14.0156n+1.0078x$$

n is an integer and x, Mu and Z are even integers.

When solved for n and x the above relations, together with equation (9), give

$$x=\frac{(Z-2)\ (0.1666-1.159h)}{u-0.1666+1.159h} \tag{17}$$

$$n=-rx \tag{18}$$

[16] Saturation is here defined by the condition, $x=+2$.

Three cases are possible as follows:

1. If Z is actually zero, equation (17) will be found to yield a definite value for x.

2. If Z is actually 2, equation (17) will be found to become indeterminate. In this case the number of combination values possible for n and x is the same as though the bromine determination had not been made; that is, it is the number corresponding to the limiting values of h. The only utility of the bromine determination in these circumstances is to eliminate as possibilities certain structural formulas.

3. If Z is actually greater than 2, equation (17) will lead to a smaller number of possibilities than correspond to value of h alone.

In the present example we find

$$-x=\frac{(Z-2)\ [0.1666-1.159\ (0.0766\pm 0.001)]}{0.166-1.159\ (0.0766\pm 0.001)-(0.0207\pm 0.0003)}$$

The denominator is a positive quantity. Hence, Z must be >2. If we assume that $n\not>50$, a Diophantine analysis yields the following as the only possible formulas for the hydrocarbon: $C_{10}H_{10}$, $C_{20}H_{20}$, $C_{30}H_{30}$, $C_{40}H_{40}$, $C_{45}H_{44}$, and $C_{50}H_{50}$. For the same condition the value of h by itself leads to 30 possibilities.

3. CONCLUSIONS

The results obtained in the examples just discussed suggest that the bromine (or other) addition number may be a valuable aid in deducing the formula of a hydrocarbon. From a purely mathematical standpoint it has a material advantage over the combustion analysis because in many cases it enables us to deal with a small even integer instead of a large one, with a consequent gain in the precision required.

The writer has hesitated to include it definitely as a possible substitute for the combustion analysis, however, because he has been unable to satisfy himself that the present state of our knowledge of the reactions of any of the halogens with the hydrocarbons justifies the assumption that the reaction can be so controlled as to be always stoichiometric in character.[17] However, it should be valuable as additional evidence and should always be determined, if only for the purpose of accumulating additional evidence as to its stoichiometric reliability.

Whether or not it is stoichiometric in a given instance could in principle be determined by carrying out the bromination in steps, in such a way that the amount of bromine added by the hydrocarbon in each step is controlled by the known activity of bromine in some second nonmiscible phase (gas or liquid) containing it. The graph of the amount added against the activity of the bromine in the nonmiscible phase should exhibit a flat corresponding to each type of stoichiometrically added bromine.[18]

---

[17] This does not refer to the possibility that addition may be accompanied by some substitution, because the latter can, of course, be determined by an acid titration and corrected for.

[18] For recent applications of the principle of this method to another situation, see Bancroft and Barnett, Proc. Nat. Acad. Sci., 16, pp. 118, 135; 1930.

VI. RÉSUMÉ AND GENERAL PROCEDURE

1. Make an approximate determination of the molecular weight.

(*a*) If M_a is less than 300, follow the procedure of Section III, 2, page 870.

(*b*) If M_a is of the order of 300 or greater, proceed to 2.

2. Make a combustion analysis.

(*a*) If $h-(\delta h)_{max.} > 0.1438$, follow the procedure of Section IV, 4, page 878.

(*b*) If 0.1438 is included within $h \pm (\delta h)_{max.}$, follow the procedure of Section IV, 5, page 879.

(*c*) If $h+(\delta h)_{max.}, < 0.1438$, follow the procedure of Section IV, 8, page 880.

3. Determine the halogen- (hydrogen-, acid-, or other-) addition number and, if the result is not zero, check the deductions of 1 and 2 by the procedure explained in Section V, 2, page 885.

4. If there is reason to suppose that the hydrocarbon may be an equilibrium mixture of polymers, it should be further investigated as described in Section IX below.

VII. EFFECTS OF IMPURITIES

The procedure outlined in the foregoing pages assumes that the hydrocarbon is pure; that is, that it contains a single molecular species. In practice, however, the requirement in this respect is that, if impurities are present, they must be of such natures and magnitudes as not to alter the measured values of h and M by such amounts as will lead to erroneous deductions. Thus if the impurities are all isomers of the principal constituent, they are without influence. Likewise, an impurity having the same hydrogen content as the principal constituent would not lead to an erroneous result, if the quantity present did not affect the measured molecular weight by a significant amount. Since, however, in general the natures of the impurities present will not be known, it is necessary to establish the purity of the sample before proceeding to determine its formula.[19]

VIII. THE "AVERAGE FORMULA" OF A MIXTURE OF HYDROCARBONS

If the procedure of the preceding pages be applied to a mixture of hydrocarbons, it may lead to a definite formula. The hydrocarbon corresponding to this formula may not, however, be present in the mixture and the result is of no value. If it is desired to find the so-called average formula of a mixture, the procedure here outlined may be used, but with the omission of those features of it which result from the Diophantine characters postulated for n and x. The average formula thus obtained will be $C_{a\pm b}H_{c\pm d}$, the subscripts being evaluated numerically. In this way the data of example 1, page 881, would yield the formula

$$C_{37.1\pm 3.7}H_{72.1\pm 6.3}$$

In determining the "average molecular weight" of a mixture by any of the methods involving the use of a solvent, the determination

[19] Suitable tests for purity have been discussed elsewhere. See Ind. Eng. Chem., 23, p. 985; 1930. It is obvious that the sample used for combustion must be carefully freed from all moisture.

should be made for at least two concentrations and preferably with at least two solvents in order to avoid the unnecessary and possibly erroneous assumption that the hydrocarbon mixture is free from the substance employed as the solvent.

IX. EFFECTS OF POLYMERIZATION

If the sample of the hydrocarbon is a mixture of polymers, certain precautions are necessary. If the polymers present are not in equilibrium with one another, the sample is a mixture. It may, in principle, be separated into its constituent hyrdocarbons by suitable methods of fractionation.

If, however, the sample is a mixture of polymers in equilibrium with one another, it will behave toward the phase rule like a pure substance, and may consequently meet the tests for purity as ordinarily applied.

If now the procedure of the preceding pages be applied to such a "pure substance," an erroneous result may be obtained. If it is desired to eliminate this possibility, it is necessary to make accurate molecular weight determinations and to demonstrate that the molecular weight is independent of concentration and/or temperature.

If the molecular weight varies appreciably with concentration and temperature, the hydrocarbon must be an equilibrium mixture of polymers of the general formula $(C_nH_{2n+x})_y$ where y is an integer. In such a case the information desired is the values of n and x, with, perhaps, the average value of y under some stated conditions.

To obtain the values of n and x, the combustion analysis should be made as accurately as possible and the molecular weight should be determined under conditions where the degree of polymerization is as small as possible. In this way it will be possible in many cases to determine n and x. If this proves not to be possible, recourse must be had to the information which can be obtained by converting the hydrocarbon into one or more of its chemical derivatives, a problem which will be specific for each case and can not be discussed in general terms.

X. OTHER CHEMICAL COMPOUNDS

The principles, which in the present paper have been developed in their application to the problem of determining the formula of a hydrocarbon, should be utilized in connection with the determination of the formula of any chemical compound. The widespread practice of reporting the results of a chemical analysis of a compound and comparing these results with the values calculated from an assumed formula should be abandoned. Instead, the investigator should deduce from the results of his analysis and their estimated accuracy, the set of chemical formulas consistent therewith. If then he can eliminate certain members of this set on the basis of auxiliary evidence, this evidence should be stated. Only when all but one member of the set can be thus eliminated can the formula be considered as established.

XI. CONCLUSIONS

By following the procedures described in the preceding pages it should be possible to determine the empirical formula of any molecularly pure hydrocarbon containing not more than, say, 100 carbon atoms.

The accuracy required in the molecular weight determination would never exceed that necessary to distinguish $C_{99}H_{198}$ from $C_{100}H_{200}$ and for this an accuracy of 0.5 to 1 per cent would be sufficient.

If the molecular weight can be determined with the required accuracy (which in no case need be better than ± 0.5 per cent), then the accuracy necessary in the combustion analysis would in no case be greater than that required to distinguish between $C_{100}H_{200}$ and $C_{100}H_{202}$. An accuracy of about ± 0.0008 unit in h is ample for this purpose. This degree of accuracy should be attainable in any instance.[20]

As regards the molecular weight determination, the required accuracy can probably be obtained in most cases where a molecular weight determination is possible. Hydrocarbons may exist, however, which are so nonvolatile and so insoluble in all solvents that it is not possible to determine their molecular weights. In such cases, however, it would be equally impossible to obtain them in the pure condition and/or to establish their purity. They are, therefore, not likely to be met with in practice. The special problems arising in connection with an equilibrium mixture of polymers are discussed in Section IX.

The writer desires to acknowledge the valued assistance of R. T. Leslie and S. T. Schicktanz for the computation of the M-table and for checking the computations throughout the manuscript.

Washington, June 24, 1930.

[20] The above statements are valid only if the sample is a molecularly pure hydrocarbon. In practic the possibility of definitely evaluating the formula of a hydrocarbon of high molecular weight will in many cases be determined, not by the accuracy attainable in the molecular weight and combustion determinations, but by the practicability of obtaining the hydrocarbon in the required degree of purity and of demonstrating the purity.