

CHROM. 8831

## THE UTILIZATION OF GAS CHROMATOGRAPHIC RETENTION DATA IN COMBINATION WITH ELEMENTAL ANALYSIS FOR SUBSTANCE IDENTIFICATION

OTTO PIRINGER\*, MIHAI JALOBEANU and UNDINA STĂNESCU

*Institute of Stable Isotopes, C.P. 243, Cluj-Napoca (Romania)*

(Received September 4th, 1975)

---

### SUMMARY

A method is described in which combinations of retention increments, compatible with the empirical formula and the retention of an unknown substance, are used for its identification. A parameter,  $W$ , which relates the molecular weight,  $M$ , and the Kováts index,  $I$ , or the molecular retention index,  $Me$ , is defined as follows:

$$W = Me - M = 0.14 \cdot I - M + 2$$

Limits of the  $W$  values are given for several structural groups on the polar stationary phase Carbowax 20M, and examples of identifications are described. Possible combinations of increments can be determined with a programmable desk calculator. Some combinations are tabulated.

---

### INTRODUCTION

The retention parameters of a solute depend on its molecular structure, and gas chromatography (GC) is, in principle, appropriate for qualitative analysis. At present an extensive literature exists concerning the relations between retention behaviour and structure<sup>1-10</sup>. Refs. 1-10 are a few representative examples from this literature and contain other sources of information on this topic. Most of the publications deal with different kinds of differences in Kováts indices in the form of structural increments and describe rules for pre-calculation of retention data for solutes having known structures.

In spite of an enormous accumulation of retention data, today GC is of minor importance in structure elucidations when compared with ultraviolet, infrared, nuclear magnetic resonance and mass (MS) spectrometric methods. This situation has arisen because of the following reasons.

First, the existing retention data have been determined on too many different stationary phases<sup>11</sup>. However, no phase exists for which an exhaustive study has been

---

\* Present address: Physical Chemistry Department, Federal School of Technology, 1007 Lausanne, Switzerland.

made of the retention behaviour of all chemical classes that can be analyzed on that phase. The knowledge gained from such a study would be necessary for the identification of an unknown compound. The limits of variation of values of different structural increments, rather than constants, must be established, taking into account the exceptions in their behaviour and the many sources of errors.

Secondly, the retention behaviour of a substance is the sum of all of the incremental contributions in the molecule. It is necessary, therefore, to find a way of resolving this superposition of increments in the individual structural contributions.

Thirdly, in principle it is inconvenient to undertake the identification of substances with only retention parameters. Differences between indices determined on two stationary phases of different polarity are also the result of superpositions of increments and, consequently, are generally ambiguous. Moreover, in the case of a complex mixture, the unequivocal assignment of peaks from chromatograms obtained on two different stationary phases to the corresponding solutes is complicated by possible changes in the relative positions of the peaks and partial superpositions of the peaks. Finally, several forms of structural increments are known and there is a tendency to sub-divide them further with the aim of precise pre-calculation of retention data. This situation complicates rather than promotes the practical use of increments in substance identification.

## THEORY

A method of substance identification is now described, taking into account the above statements. A necessary condition for this method is the availability of the empirical formula of the unknown compound, in addition to its retention data. In organic synthesis or preparative-scale isolation (for example by GC) from a natural product, elemental analysis is a usual step. If multicomponent mixtures are investigated, the molecular weights can be determined by coupling GC and MS and the empirical formulae can be eventually obtained by high-resolution MS. In principle, it is possible to obtain all of the information on retention data and elemental composition by use of suitable gas chromatograph (for example, the Model 1200 elemental analyser and peak identifier from Chemical Data Systems).

### *Correlations between retention parameters*

The next step is the splitting of the total retention into combinations of structural increments, compatible with both the retention and the formula. A parameter is defined for this purpose which relates the molecular weight of a solute to its retention. Let  $Me$  be the molecular retention index<sup>12</sup> of a solute and  $M$  its molecular weight; the retention weight,  $W$ , is then defined as:

$$W = Me - M \quad (1)$$

The molecular retention index,  $Me$ , is given by:

$$Me = M_z + 14 \cdot \frac{\log r_{x/z}}{\log r_{(z+1)/z}} \quad (2)$$

where  $r_{x/z}$  and  $r_{(z+1)/z}$  are the relative retentions of the solute  $x$  and of the  $n$ -alkane

having  $(z + 1)$  carbon atoms, respectively, related to the  $n$ -alkane having  $z$  carbon atoms and molecular weight  $M_z$ . The retention of a solute is therefore the sum of its molecular weight and retention weight. The relation  $W_z = 0$  results from eqn. 2 for  $n$ -alkanes.  $W$  represents the interaction between a substance and the stationary phase, relative to a hypothetical  $n$ -alkane with the same molecular weight  $M$ , expressed in molecular-weight units. This is a practical and impressive expression for relative retentions, which is suitable for substance identification, for pre-calculation of retentions using increments and for the characterization of stationary phases.

Because retention parameters are usually given in the form of Kováts indices,  $I$ , the transformation to  $W$  is easily made:

$$W = 0.14 \cdot I - M + 2 \quad (3)$$

as in eqn. 3. If the adjusted retention time,  $t'_{R(x)}$ , of the solute  $x$  is related to that ( $t'_{R(s)}$ ) of a standard  $s$  having molecular weight  $M_s$  and retention weight  $W_s$ , the equation

$$W = M_s - M + W_s + \frac{14}{b} \cdot \log \frac{t'_{R(x)}}{t'_{R(s)}} \quad (4)$$

can be used for calculating the parameter  $W$  of the unknown solute, where  $b = \log r_{(z+1)/z}$ .

#### *Retention increments and empirical formulae*

All of the possible combinations of the structural increments which are compatible with the elemental analysis and the retention time can be calculated if the limits of their values are known. The essential assumption for this method of substance identification is that the variation limits of  $W$  on a stationary phase are known as exactly as possible for all of the classes of compounds that can be analyzed on that column. In order to extend the range of values for the various increments and thus reduce the ambiguity of interpretation, polar stationary phases should be used. Although there is no ideal stationary phase for this purpose and no exact study has been carried out, the usefulness of this method of identification can be illustrated with Carbowax 20M, a polar phase for which a considerable amount of retention data has been accumulated.

In Table I, the limits of  $W$  values (column 6) for different structural types (column 5) of compounds grouped in 56 classes (column 1) are arranged according to their elemental composition (column 2). The limits  $W' - W''$  were calculated from the data published in refs. 13 and 14, in the Chromatographic Data Tables in *J. Chromatogr.* and from our experimental measurements. Column 3 contains the number of double-bond equivalents,  $L$ , of the class of compounds under consideration, which may be calculated from the empirical formula  $C_rH_pO_qN_uX_v$  by use of

$$L = \frac{1}{2}(2r + 2 - p - v + u) \quad (5)$$

eqn. 5. For pentavalent nitrogen,  $3u$  should be used instead of  $u$  in eqn. 5. Column 4

TABLE I

VALUES OF  $L$ ,  $\Delta M$  AND  $W$  ON CARBOWAX 20M FOR SEVERAL CLASSES OF COMPOUNDSValues of  $W$  at 120°.

No.	Elements	$L$	$\Delta M$	Class of compound	$W' - W''$
$T_0$					
0	C,H	0	0	<i>n</i> -Alkanes	0
1	C,H	0	0	Methyl- and ethyl-alkanes	-3 to -8
2	C,H	1	12	Alkenes	6- 14
				1-, 3- and 4-Alkenes	6- 9
				2-Alkenes	10- 14
3	C,H	1	12	Monocycloalkanes (C <sub>5</sub> and C <sub>6</sub> rings)	22- 26
4	C,H	2	10	Conjugated dienes	26- 35
5	C,H	2	10	Alkynes	28- 42
6	C,H	2	10	Bicycloalkanes (fused C <sub>5</sub> and C <sub>6</sub> rings)	32- 44
7	C,H	3	8	Tricycloalkanes (fused C <sub>5</sub> and C <sub>6</sub> rings)	38- 59
8	C,H	4	6	Alkylbenzenes	51- 71
				Alkylethylbenzenes	51- 60
				Monoalkylbenzenes	55- 64
				Alkylmethylbenzenes	60- 71
9	C,H	5	4	Conjugated alkenylbenzenes	77- 87
10	C,H	5	4	Tetralin	94-101
11	C,H	7	0	Naphthalene	123-138
$T_1$					
12	C,H,O <sub>1</sub>	0	2	Ethers	7- 12
				<i>n</i> -C <sub>z</sub> -O-R $z \geq 3$	7- 9
				$z = 2$	10- 12
13	C,H,O <sub>1</sub>	0	2	Alkyl methyl ethers	18- 20
14	C,H,O <sub>1</sub>	0	2	Tertiary alcohols	44- 53
15	C,H,O <sub>1</sub>	0	2	Secondary alcohols	60- 71
				<i>n</i> -C <sub>z</sub> -C(OH)-R $z \geq 3$	60- 64
				$z = 2$	62- 67
				$z = 1$	63- 71
16	C,H,O <sub>1</sub>	0	2	Primary alcohols	78- 92
17	C,H,O <sub>1</sub>	1	0	Aldehydes, ketones	47- 59
				<i>n</i> -C <sub>z</sub> -CO-R $z \geq 3$	47- 50
				$z = 2$	50- 54
				$z = 1$	54- 59
				$z = 0$	53- 57
18	C,H,O <sub>1</sub>	1	0	Cyclic ethers	48- 64
				Oxiranes	48- 56
				Oxetanes, oxolanes, oxanes	51- 64
19	C,H,O <sub>1</sub>	2	12	$\alpha,\beta$ -Unsaturated carbonyl compounds	67- 82
20	C,H,O <sub>1</sub>	2	12	Cycloketones	90- 95
21	C,H,O <sub>1</sub>	4	8	Phenols	182-193
$T_2$					
22	C,H,O <sub>2</sub>	1	2	Esters	27- 46
				<i>n</i> -C <sub>z</sub> -COOR $z \geq 3$	27- 30
				$z = 2$	31- 35
				$z = 1$	35- 39
				$z = 0$	41- 46
23	C,H,O <sub>2</sub>	1	2	Cycloethers	64- 71
				(Dioxolanes, dioxanes)	

TABLE I (continued)

No.	Elements	L	$\Delta M$	Class of compound	$W' - W''$
24	C,H,O <sub>2</sub>	1	2	$\alpha$ -tert.-Hydroxyketones	61- 72
25	C,H,O <sub>2</sub>	1	2	$\alpha$ -sec.-Hydroxyketones	87-100
26	C,H,O <sub>2</sub>	1	2	$\alpha$ -Hydroxyketones	105-121
27	C,H,O <sub>2</sub>	1	2	Acids	135-155
28	C,H,O <sub>2</sub>	2	0	$\alpha$ -Diketones	50- 66
29	C,H,O <sub>2</sub>	5	8	Alkyl benzoates	84- 89
30	C,H,O <sub>3</sub>	1	4	Cycloethers (trioxanes)	75- 78
31	C,H,N <sub>1</sub>	0	1	Amines	59- 69
32	C,H,N <sub>1</sub>	2	11	Nitriles	87-106
33	C,H,N <sub>1</sub>	4	7	Pyridines	87- 98
34	C,H,N <sub>1</sub>	4	7	Anilines	150-166
35	C,H,O <sub>2</sub> ,N <sub>1</sub>	2	3	Secondary nitroalkanes	70- 79
36	C,H,O <sub>2</sub> ,N <sub>1</sub>	2	3	Primary nitroalkanes	80-101
37	C,H,O <sub>2</sub> ,N <sub>1</sub>	6	9	Nitrobenzenes	128-142
38	C,H,Cl <sub>1</sub>	0	6	Secondary chloroalkanes	18- 22
39	C,H,Cl <sub>1</sub>	0	6	Primary chloroalkanes	27- 32
40	C,H,Cl <sub>1</sub>	4	12	Chlorobenzenes	62- 75
41	C,H,Cl <sub>2</sub>	0	13	gem-Dichloroalkanes	27- 32
42	C,H,Br <sub>1</sub>	0	9	Secondary bromoalkanes	-6 to -10
43	C,H,Br <sub>1</sub>	0	9	Primary bromoalkanes	2- 7
44	C,H,Br <sub>1</sub>	4	1	Bromobenzenes	35- 46
45	C,H,I <sub>1</sub>	0	0	Secondary iodoalkanes	-36 to -40
46	C,H,I <sub>1</sub>	0	0	Primary iodoalkanes	-23 to -29
47	C,H,I <sub>1</sub>	4	6	Iodobenzenes	-4 to 11
48	C,H,S <sub>1</sub>	0	4	Thioalcohols	59- 66
49	C,H,S <sub>1</sub>	3	12	Thiophene	61- 67
50	C,H,S <sub>2</sub>	0	8	Disulphides	68- 74
Aromatic substituents					
51	H,O <sub>1</sub>	0	2	-OH	120-130
52	H,N <sub>1</sub>	0	1	-NH <sub>2</sub>	93-104
53	H,O <sub>2</sub> ,N <sub>1</sub>	2	3	-NO <sub>2</sub>	70- 80
54	Cl <sub>1</sub>	0	6	-Cl	6- 16
55	Br <sub>1</sub>	0	9	-Br	-14 to -25
56	I <sub>1</sub>	0	0	-I	-48 to -61

represents the difference,  $\Delta M \geq 0$ , between the molecular weight of a member from the considered class and the molecular weight,  $M_z$ , of its nearest  $n$ -alkane

$$\Delta M = M - 16 - 14k \quad (6)$$

where  $k$  is the greatest integer for which  $\Delta M > 0$ .

The classes 51-56 are substituents in aromatic compounds and are applicable only together with one of the classes 8, 9, 10, 11, 21 and 29. They are used for aromatic compounds with more than one functional group, by adding the corresponding  $W$  values. Different classes can be further sub-divided with regard to their  $W$  values into sub-classes (1-alkenes, 2-alkenes, etc.) for more sophisticated investigations. New classes result from combinations of two neighbouring structural elements in a non-additive manner as a result of conjugation (classes 4 and 9) or of steric hindrance effects (classes 24, 28, etc.). Branched alkyl groups, iodine and secondary aliphatic bromine derivatives result in negative values of  $W$ .

### Observations on Table I

The limits,  $W' - W''$ , for nitrogen- and sulphur-containing compounds are less precise in comparison with the other classes of compound. Relatively few data are available for the determination of the variation limits for alkynes and iodine derivatives, and for the effects of conjugation and of steric hindrance of neighbouring functional groups. Little is known about the limits of different substituents and functional groups in saturated, aromatic and heterocyclic rings. The values of the limits for increments of high polarity can be severely affected by experimental errors.

The variation of  $W$  with temperature is not generally specific for a particular class of compounds, except for cycloalkanes and some of their derivatives for which high positive temperature gradients of between 1.2 and 2.5 units of  $W$  per  $10^\circ$  have been observed in the order tricyclo > bicyclo > monocyclo. The other classes of compounds have gradients of less than one unit of  $W$  per  $10^\circ$ . The accumulation of alkyl groups causes the temperature gradient to increase to about one unit of  $W$  per  $10^\circ$ .

Branching of aliphatic chains and the presence of alkyl groups in the vicinity of functional groups result in negative retention effects. For alkanes, alcohols, halogen derivatives and for alkyl groups *ortho* to hydroxyl and alkoxy groups in aromatic compounds, the decreases of  $W$  lie between  $-3$  and  $-8$  for one alkyl group, except for iodine derivatives. For iodine derivatives smaller decreases of  $W$  from  $-1$  to  $-2$  have been found for an alkyl group in the vicinity of an iodine atom. An ethyl (Et) or methyl (Me) group in the vicinity of an oxygen atom in ketones, aldehydes, esters, alicyclic ethers, and *ortho* to bulky functional groups (e.g., nitro) in aromatic compounds, causes the corresponding values of  $W$  to decrease by  $-8$  to  $-12$  units; this effect is even more pronounced in cyclic ethers where decreases of  $-13$  to  $-16$  units are observed. This effect decreases rapidly with distance; e.g., decreases of 0 to  $-4$  units occur for ethyl or methyl groups *meta* to functional groups. In alkanes, the negative retention effects of branching decrease in the order 4-Et > 2-Et, 2-Me > 3-Et, 3-Me. The effects are generally additive; e.g., two methyl groups on the same carbon atom are equivalent to two separate methyl groups, but two vicinal methyl groups are only equivalent to one separate methyl group. In aromatic compounds, two *o*-methyl groups produce a positive effect of ca. 8 units and only longer groups show negative values (0 to  $-4$ ).

Alkyl substituents of carbon atoms of double bonds produce only very small effects (0 to  $-2$ ). The  $W$  values for double bonds in alkenes at positions 1 and 3 are a little higher than for the more inner positions. *cis*-Isomers give higher retentions than the corresponding *trans*-isomers.

Conjugation of double bonds produces additional positive effects, but vicinal carbonyl and hydroxyl groups decrease the  $W$  values. Two hydroxy groups in benzene lower the retention by ca.  $-85$ ,  $-60$  and  $-55$  units for *ortho*, *para* and *meta* substitution, respectively.

The first member of each homologous series generally exhibits anomalous values of  $W$ . Methyl esters have values of 5–10 units greater than those of the higher alkyl esters. Vinyl esters and esters of cyclic alcohols exhibit positive values of 3 and 14–16 units, respectively. The values of  $W$  for cyclic alcohols are obtained by summing the maximum values for the corresponding cycloalkane and the secondary alcohol.

The increase in retention produced by halogen substituents of the carbon atoms of double bonds is not completely established. The effect of two *gem*-chlorine atoms is similar to that of only one chlorine atom. In aromatic compounds the retentions of dichloro derivatives decreases in the order *ortho* > *para* > *meta*. The retention weights of such derivatives are lower (by *ca.* 8 units) than expected from a simple addition of the *W* values for monochlorobenzenes and for the chloro group itself (classes 40 and 54, respectively).

#### The identification procedure

In the case of an unknown chemical compound, the experimentally measured values for *W*, *L* and  $\Delta M$  result from combinations of the *n* structural increments contained in the compound and can be obtained by a summation of these incremental

$$W = \sum_{i=1}^n W_i, L = \sum_{i=1}^n L_i, \Delta M = \sum_{i=1}^n \Delta M_i \quad (7)$$

values as in eqn. 7. All of the considered combinations of the structural increments should be compatible with eqn. 7 and with the empirical formula of the solute. It is convenient to begin the identification by eliminating nitrogen-, halogen- and sulphur-containing increments, *i.e.*, classes 31–50 and 52–56, respectively. For nitrogen, sulphur and the halogens the variety of possible classes is much smaller than in the case of compounds containing carbon, hydrogen and oxygen. After subtracting all of the possible combinations of nitrogen-, halogen- and sulphur-containing increments and their corresponding *M* and  $\Delta M$  values, the remainder of the molecule is investigated with regard to the possible combinations of the carbon-, hydrogen- and oxygen-containing classes.

In Table II, the limits of *W* values for pairs of increments (obtained from Table I) are given. The Table is divided into seven parts according to the  $\Delta M$  values, and in each part the possible combinations of classes (column 4) are given in order of increasing values of the lower limit of *W* (column 3). Column 1 gives the number of oxygen atoms in the molecule, and column 5 gives the minimum mass,  $M_{\min.}$ , that a substance in the respective class combination must have.  $M_{\min.}$  provides additional information which is helpful in the selection of structures from the compatible set.

If sets of class combinations are selected from Table II, it is necessary eventually to take into account branching of alkyl groups which results in lower *W* values as discussed in the previous section. Combinations of more than two classes can be resolved in a reasonable time only with a programmable desk calculator. In order to simplify the program, an algorithm will now be described, in which only the first 29 classes in Table I are considered. These classes contain the structural types which appear most frequently in compounds of carbon, hydrogen and oxygen.

The information collected from Table I is organized into a file for processing by a computer. The table is divided into three parts,  $T_0$ ,  $T_1$  and  $T_2$ , according to the number of oxygen atoms present, 0, 1 or 2. In sub-table  $T_i$  the lines of data are ordered in increasing values of  $L_i$ , decreasing values of  $\Delta M_i$  and increasing values of  $W_i - W_i'$ . The first terms of the parts  $T_0$ ,  $T_1$  and  $T_2$  are denoted by  $N_0$ ,  $N_1$  and  $N_2$ , respectively.

TABLE II

VALUES OF  $W$  FOR PAIRS OF INCREMENTS OBTAINED FROM TABLE I BY COMBINATION OF CLASSES CONTAINING CARBON, HYDROGEN AND OXYGEN

No. of oxygen atoms	$L$	$W' - W''$	Combination	$M_{\text{mix}}$
$\Delta M = 0$				
0	0	0	0	16
0	0	-3 to -8	1	58
1	1	13-26	2,12	72
1	1	24-34	2,13	58
1	1	29-38	3,12	114
2	2	33-60	2,22	86
1	1	40-46	3,13	100
1	1	47-59	17	30
1	1	48-64	18	44
2	2	49-72	3,22	128
2	2	50-66	28	58
1	1	50-67	2,14	86
1	1	66-79	3,14	100
1	1	66-85	2,15	72
2	2	70-85	2,23	72
1	1	82-97	3,15	86
1	1	84-106	2,16	58
2	2	86-97	3,23	142
0	7	89-130	7, 8	240
2	2	94-118	17,17	72
2	2	95-123	17,18	72
2	2	96-128	18,18	86
2	2	97-107	12,20	128
1	1	100-118	3,16	100
0	7	103-122	4, 9	156
0	7	105-129	5, 9	128
2	2	108-115	13,20	114
0	7	109-131	6, 9	226
3	3	117-141	20,22	142
0	7	120-136	4,10	184
0	7	122-143	5,10	156
0	7	123-138	11	128
0	7	126-145	6,10	254
2	2	134-148	14,20	114
2	9	135-160	8,29	212
2	2	141-169	2,27	72
2	2	150-166	15,20	100
3	3	154-166	20,23	156
2	2	157-181	3,27	114
2	2	168-187	16,20	114
1	8	170-197	11,17	156
1	8	171-202	11,18	170
3	3	225-250	20,27	128
1	8	233-264	8,21	170
$\Delta M = 2$				
1	0	7-12	12	60
1	0	18-20	13	46
2	1	27-46	22	60
1	0	44-53	14	74



TABLE II (continued)

No. of oxygen atoms	L	W' - W''	Combination	M <sub>int.</sub>
2	1	54-71	12,17	74
2	1	55-76	12,18	60
1	0	60-71	15	60
2	1	61-72	24	88
2	1	64-71	23	74
2	1	65-79	13,17	60
2	1	66-84	13,18	46
3	2	74-105	17,22	88
3	2	75-110	18,22	102
0	6	76-118	7, 7	326
0	6	77-106	4, 8	130
1	0	78-92	16	32
0	6	79-113	5, 8	102
0	6	83-101	2, 9	130
0	6	83-115	6, 8	200
2	1	87-100	25	74
2	1	91-112	14,17	102
2	1	92-117	14,18	102
0	6	99-113	3, 9	172
0	6	100-115	2,10	130
2	1	105-121	26	60
2	1	107-130	15,17	88
2	1	108-135	15,18	88
3	2	111-130	17,23	102
3	2	112-135	18,23	116
0	6	116-127	3,10	200
2	8	122-148	7,29	298
2	1	125-151	16,17	74
2	1	126-156	16,18	74
1	7	130-150	11,12	172
2	1	135-155	27	46
1	7	141-158	11,13	158
2	8	150-184	11,22	186
1	7	167-182	9,20	186
1	7	167-191	11,14	186
4	10	168-178	29,29	270
3	2	182-214	17,27	74
1	7	183-209	11,15	172
3	2	183-219	18,27	88
1	7	184-196	10,20	214
2	8	187-209	11,23	200
1	7	201-230	11,16	158
1	7	246-264	51,11	144
$\Delta M = 4$				
2	0	14-24	12,12	118
2	0	25-32	12,13	104
3	1	34-58	12,22	104
2	0	36-40	13,13	90
3	1	45-66	13,22	88
2	0	51-65	12,14	118
4	2	54-92	22,22	118

(Continued on p. 432)

TABLE II (continued)

No. of oxygen atoms	<i>L</i>	<i>W'</i> - <i>W''</i>	Combination	<i>M</i> <sub>min.</sub>
0	5	57-85	2, 8	104
2	0	62-73	13,14	104
0	5	64-94	4, 7	160
0	5	66-101	5, 7	188
2	0	67-83	12,15	104
0	5	70-103	6, 7	286
3	1	71-83	12,23	118
3	1	71-99	14,22	118
0	5	73-97	3, 8	146
3	1	75-78	30	90
0	5	77-87	9	104
2	0	78-91	13,15	90
3	1	82-91	13,23	104
2	0	85-104	12,16	90
3	1	87-117	15,22	104
2	0	88-106	14,14	118
4	2	91-117	22,23	132
0	5	94-102	10	132
2	0	96-112	13,16	76
2	0	104-124	14,15	104
3	1	105-138	16,22	90
3	1	108-124	14,23	104
2	7	110-124	4,29	188
2	7	112-131	5,29	160
2	7	116-133	6,29	258
2	0	120-142	15,15	90
2	0	122-145	14,16	90
3	1	124-142	15,23	90
1	6	124-146	9,17	132
1	6	125-151	9,18	146
4	2	128-142	23,23	146
2	0	138-163	15,16	76
0	6	141-160	10,17	160
1	6	141-166	8,20	160
3	3	142-163	16,23	104
1	6	142-165	10,18	174
3	1	142-167	12,27	90
3	1	153-175	13,27	74
2	0	156-184	16,16	62
4	2	162-201	22,27	104
3	1	179-208	14,27	104
3	1	195-226	15,27	90
4	2	199-226	23,27	118
1	6	208-228	4,21	146
1	6	210-235	5,21	118
3	1	213-247	16,27	76
4	2	270-310	27,27	90
$\Delta M = 6$				
0	4	44-73	2, 7	162
0	4	51-71	8	78
0	4	52-70	4, 4	106
0	4	54-77	4, 5	78

TABLE II (continued)

No. of oxygen atoms	L	W' - W''	Combination	M <sub>max.</sub>
0	4	56-84	5, 5	50
0	4	58-79	4, 6	120
0	4	60-85	3, 7	232
0	4	60-86	5, 6	148
0	4	64-88	6, 6	246
1	5	84-99	9,12	148
2	6	90-103	2,29	162
1	5	95-107	9,13	134
1	5	98-130	8,17	106
1	5	99-135	8,18	120
1	5	101-113	10,12	176
2	6	104-133	9,22	162
2	6	106-115	3,29	204
1	5	112-121	10,13	162
1	5	121-140	9,14	162
2	6	121-147	10,22	190
1	5	128-154	7,20	246
1	5	137-158	9,15	148
1	5	138-154	10,14	162
2	6	141-158	9,23	176
1	5	154-172	10,15	148
1	5	155-179	9,16	134
2	6	158-172	10,23	204
1	5	172-193	10,16	162
3	7	174-184	20,29	218
0	11	174-209	8,11	204
1	5	188-207	2,21	120
1	5	200-213	51, 9	120
1	5	204-219	3,21	162
2	6	212-242	9,27	148
1	5	217-227	51,10	148
2	6	229-256	10,27	176
$\Delta M = 8$				
0	3	32-49	2, 4	80
0	3	34-56	2, 5	52
0	3	38-58	2, 6	122
0	3	38-59	7	164
0	3	48-61	3, 4	66
0	3	50-68	3, 5	94
0	3	54-70	3, 6	192
1	4	58-83	8,12	122
1	4	69-91	8,13	108
2	5	78-117	8,22	136
2	5	84-89	29	136
1	4	85-118	7,17	192
1	4	86-123	7,18	206
1	4	95-124	8,14	136
1	4	111-142	8,15	122
2	5	115-142	8,23	150
1	4	116-130	4,20	136
1	4	118-137	5,20	108

(Continued on p. 434)

TABLE II (continued)

No. of oxygen atoms	L	W' - W''	Combination	M <sub>min.</sub>
1	4	122-139	6,20	206
1	4	129-162	8,16	108
3	6	131-148	17,29	164
3	6	132-153	18,29	178
0	10	154-174	9, 9	206
0	10	171-188	9,10	234
1	4	174-197	51, 8	94
1	4	182-193	21	94
2	5	186-226	8,27	122
2	5	229-252	17,21	122
2	5	230-257	18,21	136
$\Delta M = 10$				
0	2	12- 28	2, 2	54
0	2	26- 35	4	54
0	2	28- 40	2, 3	68
0	2	28- 42	5	26
0	2	32- 44	6	124
0	2	44- 52	3, 3	138
1	3	45- 71	7,12	208
1	3	56- 79	7,13	194
2	4	65-105	7,22	222
1	3	73- 94	4,17	82
1	3	74- 99	4,18	96
1	3	75-101	5,17	54
1	3	76-106	5,18	68
1	3	79-103	6,17	152
1	3	80-108	6,18	166
1	3	82-112	7,14	180
3	5	91-101	12,29	180
1	3	96-109	2,20	82
1	3	98-130	7,15	180
3	5	102-109	13,29	166
2	4	102-130	7,23	236
4	6	111-135	22,29	194
1	3	112-121	3,20	152
1	3	116-151	7,16	194
3	5	128-142	14,29	194
0	9	128-158	8, 9	180
3	5	144-160	15,29	180
0	9	145-172	8,10	208
4	6	148-160	23,29	208
0	9	149-173	4,11	180
0	9	151-180	5,11	152
0	9	155-182	6,11	250
3	5	162-181	16,29	166
2	4	173-214	7,27	208
2	4	180-190	20,20	166
2	4	189-205	12,21	138
2	4	200-213	13,21	124
3	5	207-215	51,29	152
3	5	209-239	21,22	152
4	6	219-244	27,29	180

TABLE II (continued)

No. of oxygen atoms	<i>L</i>	<i>W'</i> - <i>W''</i>	Combination	<i>M</i> <sub>min.</sub>
2	4	226-246	14,21	152
2	4	242-264	15,21	138
3	5	246-264	21,23	166
2	4	260-285	16,21	123
2	4	305-319	51,21	110
$\Delta M = 12$				
0	1	6-14	2	28
0	1	22-26	3	70
1	2	33-47	4,12	98
1	2	35-54	5,12	70
1	2	39-56	6,12	168
1	2	44-55	4,13	84
1	2	46-62	5,13	56
1	2	50-64	6,13	154
1	2	53-73	2,17	56
2	3	53-81	4,22	112
1	2	54-78	2,18	70
2	3	55-88	5,22	84
2	3	59-90	6,22	182
1	2	69-85	3,17	98
1	2	70-88	4,14	112
1	2	70-90	3,18	112
1	2	72-95	5,14	84
1	2	76-97	6,14	140
1	2	86-106	4,15	98
1	2	88-113	5,15	70
1	2	90-95	20	84
2	3	90-106	4,23	126
2	3	92-113	5,23	98
1	2	92-115	6,15	140
2	3	96-115	6,23	196
0	8	102-142	8, 8	154
1	2	104-127	4,16	84
1	2	106-134	5,16	56
1	2	110-136	6,16	154
0	8	115-146	7, 9	266
0	8	129-152	2,11	154
2	3	137-154	17,20	112
2	3	138-159	18,20	126
0	8	145-164	3,11	196
2	10	161-176	9,29	238
2	3	161-190	4,27	98
2	3	163-197	5,27	70
2	3	167-199	6,27	168
1	9	213-233	11,20	210

To the unknown substance we assign a quadruplet in terms of the quantities (*q*, *L*,  $\Delta M$  and *W*) and suppose that this substance results from a combination of some of the constituents in Table I. The problem is to determine the numbers  $j_1, j_2, \dots, j_p$  (corresponding to the numbers of the constituents in column 1 of Table I) such that eqns. 8-11 are satisfied.

$$q = O_{j_1} + O_{j_2} + \dots + O_{j_p}, \text{ when } O_{j_1} \geq O_{j_2} \geq \dots \geq O_{j_p} \quad (8)$$

$$L = L_{j_1} + L_{j_2} + \dots + L_{j_p} \quad (9)$$

$$\Delta M = \Delta M_{j_1} + \Delta M_{j_2} + \dots + \Delta M_{j_p} \quad (10)$$

$$W'_{j_1} + W'_{j_2} + \dots + W'_{j_p} \leq W \leq W''_{j_1} + W''_{j_2} + \dots + W''_{j_p} \quad (11)$$

The algorithm provides all combinations of  $p$  components  $j_1$  to  $j_p$  for which the conditions 8–11 hold. From the first condition, 8, we obtain the sub-table  $T_k$  where we look for the component  $j_i$ . Here,  $N_3$  is the end of the file,

$$\begin{aligned} &\text{if } O_{j_i} = 0, 0 \leq j_i < N_1 \\ &\text{if } O_{j_i} = 1, N_1 \leq j_i < N_2 \\ &\text{if } O_{j_i} = 2, N_2 \leq j_i < N_3 \end{aligned}$$

For example, if the number of oxygen atoms,  $q$ , in the investigated molecule is 3, then:

$$O_{j_1} = O_{j_2} = O_{j_3} = 1, O_{j_4} = \dots = O_{j_p} = 0$$

$$\text{or } O_{j_1} = 2, O_{j_2} = 1, O_{j_3} = \dots = O_{j_p} = 0$$

No other possibility may be used to represent  $q$  as a sum of  $p$  components providing condition 8 is satisfied. We note that some of the numbers  $j_1$  to  $j_p$  can be zero, corresponding to the fact that the unknown substance results from a combination of less than  $p$  constituents. This case has been taken into account in Table I where all of the characteristics of the first constituent are equal to zero.

When  $p = 3$ ,  $q$  can be represented as sum of three numbers in the following manner:

$$\begin{aligned} 0 &= 0 + 0 + 0 \\ 1 &= 1 + 0 + 0 \\ 2 &= 2 + 0 + 0 \quad \text{or} \quad 1 + 1 + 0 \\ 3 &= 2 + 1 + 0 \quad \text{or} \quad 1 + 1 + 1 \\ 4 &= 2 + 2 + 0 \quad \text{or} \quad 2 + 1 + 1 \\ 5 &= 2 + 2 + 1 \\ 6 &= 2 + 2 + 2 \end{aligned}$$

Hence, if the unknown substance is a combination of three constituents, it contains at most six oxygen atoms.

The algorithm is programmed according to the flow diagrams in Fig. 1a–e. The  $t'_{R(s)}$ ,  $M_s$ ,  $W_s$  and  $b$  data corresponding to a chromatogram are inserted first, followed by the data referring to the unknown substance,  $t'_{R(x)}$ ,  $r$ ,  $p$  and  $q$ . Values of  $M$ ,  $\Delta M$ ,  $L$  and  $W$  are then computed. The value of  $q$  is analyzed by means of a program,  $P$ ,

and a sub-table,  $S$ , is chosen which contains the components of of the unknown substance:

If  $q = 0$ , then  $j_1, j_2, j_3 \in T_0$

If  $q = 1$ , then  $j_1 \in T_1$  and  $j_2, j_3 \in T_0$

If  $q = 2$ , then  $j_1 \in T_2$  and  $j_2, j_3 \in T_0$  or  $j_1, j_2 \in T_1$  and  $j_3 \in T_0$

If  $q = 3$ , then  $j_1 \in T_2, j_2 \in T_1, j_3 \in T_0$  or  $j_1, j_2, j_3 \in T_1$

If  $q = 4$ , then  $j_1, j_2 \in T_2$  and  $j_3 \in T_0$  or  $j_1 \in T_2$  and  $j_2, j_3 \in T_1$

If  $q = 5$ , then  $j_1, j_2 \in T_2$  and  $j_3 \in T_1$

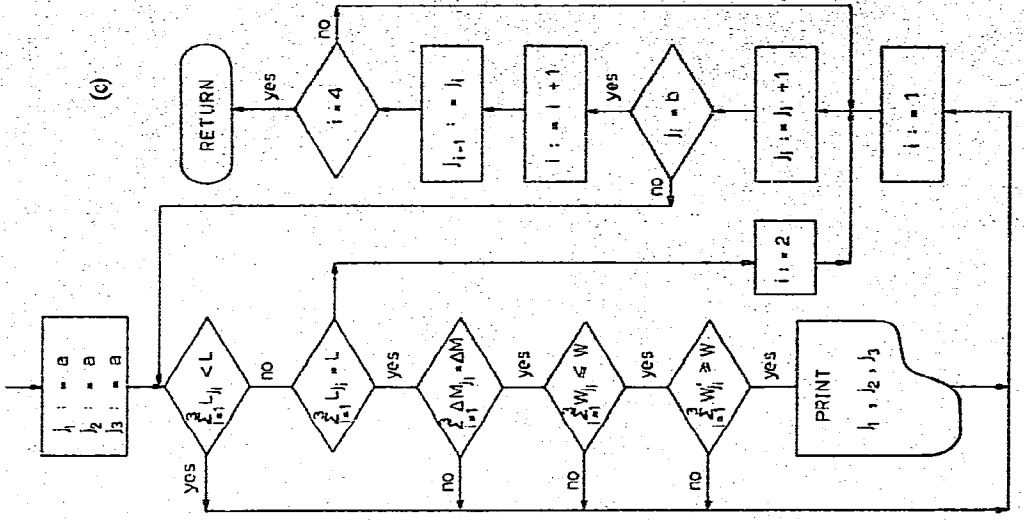
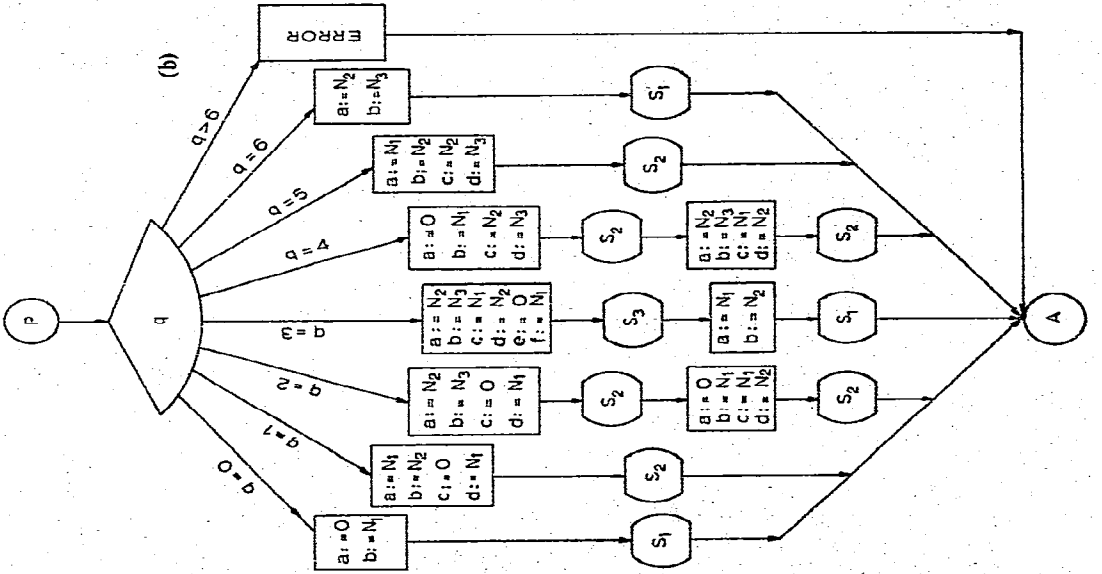
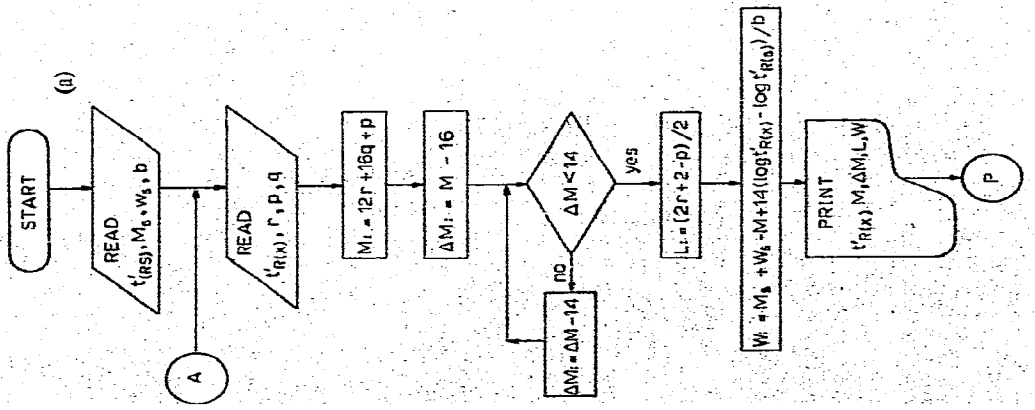
If  $q = 6$ , then  $j_1, j_2, j_3 \in T_2$

One of the sub-routines  $S_1$ ,  $S_2$  and  $S_3$  is then run:  $S_1$ , when all of the components are in the same sub-table;  $S_2$ , when one component is in one sub-table and the other two components are in another sub-table; and  $S_3$ , when each component is in a different sub-table. The parameters of the sub-routines are the limits of the sub-tables:  $a$  and  $b$  for  $S_1$ ;  $a, b, c$  and  $d$  for  $S_2$ ; and  $a, b, c, d, e$  and  $f$  for  $S_3$ . The program provides values of  $j_1, j_2$  and  $j_3$  for verifying the conditions in eqns. 9-11. It is to be noted that if  $\sum_i \Delta M_{j_i} > 14$ , then  $\sum_i \Delta M_{j_i} = \sum_i \Delta M_{j_i} - 14$ . From the values of  $j_1, j_2$  and  $j_3$ , those which have no chemical meaning are rejected. The allowed triplets  $j_1, j_2$  and  $j_3$  are then printed and the program awaits input data from a new unknown substance, analyzed on the same chromatogram.

This program has been used on a Hewlett-Packard 9100 desk calculator with the file in the 9101A extended memory and the results were printed on a teletype. The program has four parts: input of data; computation of  $L, \Delta M, M$  and  $W$ ; choice of the parameters  $a, b, c, d, e$  and  $f$ ; and provision of the order numbers  $j_1, j_2$  and  $j_3$  of the constituents that can be present in the sample (carried out by the program,  $P$ ). The values of  $M, t'_{R(x)}, L, \Delta M$  and  $W$  are first printed by the teletype, followed by the solutions. If all of the allowed possible combinations of classes have been collected from Table II or calculated by means of the computer, structural formulae can be proposed for the investigated molecule. The retention weights corresponding to these formulae are then estimated using Table I and the observations on Table I, and these values are compared with the experimental value of  $W$ . The use of a second stationary phase, which is selective for the supposed compounds and for which retention data are available, is recommended for further investigations, together with information from spectroscopic methods and knowledge of the method of synthesis, or origin, of the sample.

## EXPERIMENTAL

A glass column (1.8 m  $\times$  3.0 mm I.D.), A, was packed with 5% Carbowax 20M on Chromosorb G, AW-DMCS (Applied Sciences Labs., State College, Pa., U.S.A.) (100-120 mesh) and connected to a Perkin-Elmer F-11 gas chromatograph equipped with a flame-ionization detector (FID). Argon was used as the carrier gas at a flow-rate of 25 ml/min. A stainless-steel open capillary column (50 m  $\times$  0.25 mm I.D.), B, coated with Carbowax 20M was also used. It was connected to a Perkin-Elmer Model 990 gas chromatograph equipped with a FID. The carrier gas was helium at a flow-rate of 1 ml/min.





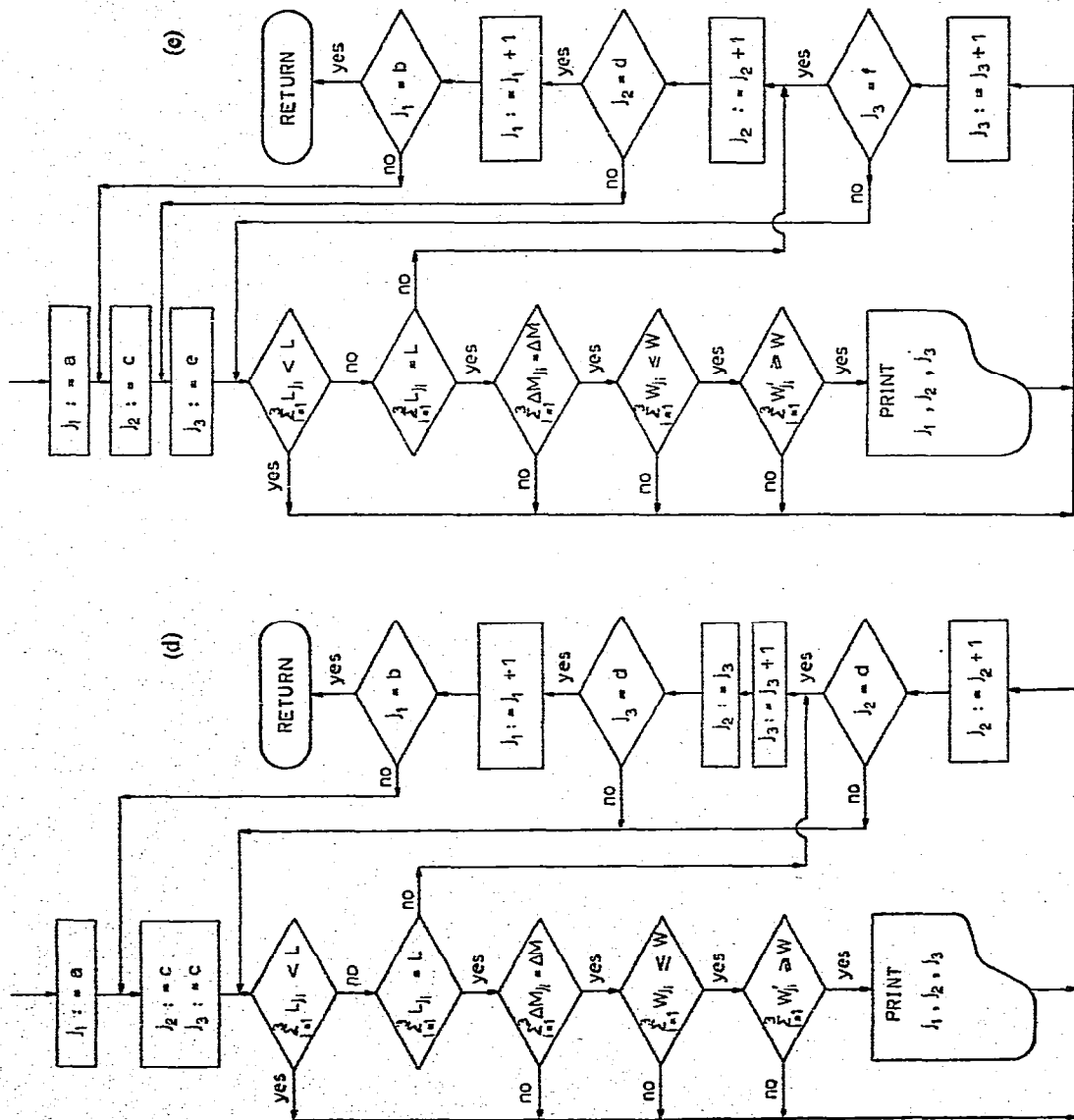


Fig. 1. Flow diagram for the computer program. (a) Diagram of the data input; (b) diagram of the selection of sub-tables; (c) sub-routine  $S_1$ ; (d) sub-routine  $S_2$ ; (e) sub-routine  $S_3$ .

TABLE III  
VALUES OF  $b$  FOR CARBOWAX 20M AT SEVERAL TEMPERATURES

Temperature (°C)	$b$	Temperature (°C)	$b$	Temperature (°C)	$b$
100	0.248	140	0.201	180	0.162
110	0.235	150	0.190	190	0.153
120	0.223	160	0.180	200	0.145
130	0.212	170	0.171	230	0.121

For MS analysis (example 5), the Model 990 gas chromatograph was coupled directly to the ion source of a Varian MAT-311 mass spectrometer.

All of the retention times were measured with a Perkin-Elmer SIP-1 integrator.

For exact temperature control, a calibrated thermocouple was mounted in the oven. In Table III, values of  $b$  are given at different temperatures. In order to determine the temperature dependence of  $W$ , the investigated sample must be run at two or three different temperatures. This is also necessary for samples of low volatility, in order to extrapolate the values to 120°. For preparative-scale separations, a Perkin-Elmer F-21 gas chromatograph was used. This was equipped with a column (9 m × 8 mm I.D.) packed with 5% Carbowax 20M on Chromosorb G AW-DMCS (60–80 mesh). A Hewlett-Packard Model 185 analyzer was used for the micro-analysis of carbon, hydrogen and nitrogen.

#### APPLICATIONS

In order to test the selectivity of the method with regard to the possibilities of combination of increments and to illustrate its advantages and limitations, the following six examples are given. The first three examples serve to fix the procedure and the algorithm. The other examples represent real cases investigated in our laboratory and emphasize the various possibilities of application of the method itself or in combination with other identification techniques. A reduction in selectivity resulted from the limits of the  $W$  values for the individual classes being too large. This was inevitable as the knowledge of the retention behaviour on this phase is incomplete, and the errors in the published data were different.

##### Example 1

First, the retention times for naphthalene,  $C_{10}H_8$  (1), 1,5-pentanediol,  $C_5H_{12}O_2$  (2) and 3-(4-hydroxycyclohexyl)-1-propene,  $C_9H_{16}O$  (3) were estimated for column A at 120° and an argon flow-rate of ca. 25 ml/min. Under these experimental conditions, a retention time,  $t'_{R(3)}$ , of 19.16 min was found for *n*-hexadecane ( $M_s = 226$ ,  $W_s = 0$ ) as reference solute. From Table I, values of  $W$  were then estimated, using average values of the corresponding intervals  $W' - W''$  for the first two compounds: for 1,  $W = 130$ ; for 2,  $W = 2 \cdot 85 = 170$ . For compound 3, a value  $W = 7 + 26 + 71 = 104$  was obtained taking into account the observation that for cyclic alcohols the upper limits for classes 3 and 15 must be used. From eqn. 4 and with  $b = 0.223$  (Table III), we obtain the adjusted retention times:  $t'_{R(1)} = 62.12$ ,  $t'_{R(2)} = 110.0$  and  $t'_{R(3)} = 37.14$  min.

TABLE IV  
DATA PRINTED OUT BY THE TELETYPE FOR EXAMPLE 1

Substance, $x$	$M$	$t'_{R(x)}$ (min)	$L$	$\Delta M$	$W$	Allowed triplets of classes ( $j_1, j_2, j_3$ )
1. Naphthalene	128	62.12	7	0	130	11,0,0
2. 1,5-Pentanediol	104	110.0	0	4	169.5	16,16,1 16,16,0
3. 3-(4-Hydroxycyclohexyl)-1-propene	140	37.14	2	12	104	3,3,14 5,0,16,1 2,2,16 5,1,15 6,1,16 6,0,15 4,0,16 5,0,15 3,3,15 5,1,16 3,2,15 4,1,16 4,0,15 6,0,16,1 3,2,16,1* 6,1,15

\* An additional combination of the allowed triplets with class 1 is considered by the computer program.

These values of the retention times were then used as input data for the desk calculator in order to verify the program and to test how many combinations of classes are compatible with these data. The set of  $t'_{R(x)}$ ,  $M_s$ ,  $W_s$  and  $b$  and the three sets of  $t'_{R(x)}$ ,  $r$ ,  $p$  and  $q$  for the three compounds were introduced into the computer (Fig. 1a). The values represented in Table IV appeared on the teletype. For 1, only class 11 results as a possible alternative. For 2, no selection can be made between 1,5-pentanediol and the isomers of methyl-1,4-butanediol. For 3, 16 combinations of classes are possible, but in all of them an aliphatic hydroxyl group is combined with an alkene or a cycloalkane. Additional information must be considered in order to obtain further selection. First, the observation on Table I referring to the high temperature gradient of cycloalkanes can be used. The italicized combinations are selected on this basis. Taking into account the minimum molecular-weight criteria (column 5 in Table II), only the combinations 3,2,15 and 3,2,16,1 are compatible. For a final decision between these combinations, information from other sources must be available.

#### Example 2

A compound having the empirical formula  $C_9H_9O_3N$  and a value of  $W = 192$  was considered. For pentavalent nitrogen, the values  $L = 7$  and  $\Delta M = 9$ , respectively, may be obtained from eqns. 5 and 6. Tests of the combinations of classes were then started with the nitrogen-containing increments. Only classes 37 ( $\Delta M = 9$ ), 35 and 36 ( $\Delta M = 3$ ) are compatible with eqn. 7. On subtracting from the experimental values of  $L$ ,  $\Delta M$  and  $W$  and from the empirical formula the corresponding values for nitrobenzene (class 37), a residue having one oxygen atom ( $L = 1$ ,  $\Delta M = 0$  and  $W = 50-64$ ) results. From Table II it can be seen that the classes 17 and 18 and the combination 2,14 are compatible with this residue. The combination 2,14 requires a value of  $M_{min.} = 86$  and can be eliminated.

Following the same procedure with classes 35 and 36, a C, H, O, residue having  $L = 5$ ,  $\Delta M = 6$  and  $W = 91-122$  results, and from Table II the following combinations of classes are compatible: 9,12, 9,13, 8,17, 8,18, 10,12, 10,13, 9,14 and 10,22. On applying the minimum-mass criteria, only combinations 8,17 and 8,18 remain. The combinations of 8,18,35 and 8,18,36 also exceed the molecular weight of the

compound; hence the combinations of 8,17,35, 8,17,36, 17,37 and 18,37 remain. In order to make a choice between the remaining combinations, more information from other sources is necessary.

For trivalent nitrogen, the combination 17,22,33 is obtained in a similar manner.

### Example 3

A substance having the empirical formula  $C_5H_9Br$  and a value of  $W = -1$  was investigated. As in the previous example, we obtained the values of  $L = 1$  and  $\Delta M = 7$  by calculation. Values corresponding to classes 42 and 43, respectively, were then subtracted, resulting in two sets of values for the residue:

$$\begin{array}{llll} (1) \text{ C,H} & L = 1, & \Delta M = 12, & W' - W'' = 5-9 \\ (2) \text{ C,H} & L = 1, & \Delta M = 12, & W' - W'' = -9 \text{ to } -14 \end{array}$$

In the first case, a combination of classes 2 and 42 is possible, with two isomers, 4-bromo-1-pentene and 3-bromo-2-methyl-1-butene. In the second case no combination is possible, because, the negative values of  $W$  for the residue require two alkyl groups and the resulting 3-bromo-2,2-dimethyl-1-propene is not compatible with the observation on alkyl groups at carbon double bonds.

### Example 4

In this example the problem is to identify the major impurity in a sample of ethyl acrylate. Column A was used at  $120^\circ$  ( $b = 0.223$ ). Using ethyl acrylate as the standard solute ( $M_s = 100$ ,  $W = 40$ ), retention times  $t'_{R(s)} = 0.80$  and  $t'_{R(x)} = 4.92$  min were measured, respectively. The empirical formula of the unknown compound was found to be  $C_7H_{14}O_3$  by preparative GC separation and elemental analysis. With these input data, the value of  $W = 43.4$  and the combinations of classes 12,22, 12,22,1 and 13,22,1 were obtained by use of the desk calculator. Ethyl- $\beta$ -ethoxy propionate (12,22) was proposed as the unknown compound and is related to the skeleton of ethyl acrylate. This conclusion was supported by MS analysis, although the molecular ion could not be detected.

### Example 5

Here, the problem is to identify the impurities in a sample of styrene (Table V). Column B was used at  $130^\circ$  together with the GC-MS technique. The mass spectra of all of the peaks had pronounced molecular ions and were characteristic of aromatic compounds. Although the empirical formulae of the impurities were not known, from the expected aromatic character of the solutes the presence of one of the classes 8, 9, 10, 11, 21 and 29 was assumed in all of the considered combinations of increments. Peaks 6-8 having  $\Delta M = 4$  and  $W < 87$  (the upper limit for class 9 in Table I) are due to alkylstyrenes, with reduced  $W$  values in accordance with the observations on Table I. From Table II, tetralin is assigned to peak 9. Peaks 1, 2, 4 and 5 having  $W < 71$  are alkylbenzenes. For peak 10, the combinations of classes 8,17 and 8,18 are compatible with the molecular weight,  $M = 120$ . Acetophenone could be identified by direct sampling. In the same manner, from the combinations 8,15 and 8,16 for peak 12 ( $\Delta M = 8$ ), methylphenylcarbinol was identified. For peak 11, the combina-

TABLE V

VALUES OF  $M$  AND  $W$  AND THE IDENTIFIED STRUCTURES OF THE SOLUTES IN EXAMPLE 5

Peak	$M$	$W$	$\Delta M$	Compound
1	120	58	6	Isopropylbenzene
2	134	54	6	1,3-Diethylbenzene
3	104	85	4	Styrene
4	148	48	6	3-Ethyl-1-isopropylbenzene
5	162	41	6	1,3-Diisopropylbenzene
6	132	85	4	3-Ethylstyrene
7	160	59	4	2,6-Diethylstyrene
8	160	65	4	2,4-Diethylstyrene
9	132	100	4	Tetralin
10	120	120	6	Acetophenone
11	128	125	0	Naphthalene
12	122	136	8	Methylphenylcarbinol

tion 5,9 and the class 11 are compatible with the molecular weight. Naphthalene was found by direct sampling. The alkylbenzenes and alkylstyrenes proposed in Table V were supported by "table matching".

#### Example 6

In this example, MS examination of the deuteration product of phenylacetylene showed two major ions with  $m/e$  105 and 123 and a small peak at  $m/e$  102, the molecular weight of undeuterated phenylacetylene. The peak at  $m/e$  123 is not compatible with the expected product.

Gas chromatography on column A at 150° gave two peaks with the adjusted retention times  $t'_{R(s)} = 2.25$  min (phenylacetylene,  $M_s = 102$  and  $W_s = 60 + 34 = 94$ ) and  $t'_{R(x)} = 9.13$  min ( $b = 0.188$ ). From eqn. 4, we obtained a molecular retention index,  $Me = 241$ . The peak having  $M = 123$  is due to a deuterium derivative of the unknown compound, because nitrogen was absent from the sample and an odd value of  $m/e$  is not possible for compounds containing carbon, hydrogen and oxygen. We therefore started with the values of  $M = 122$ ,  $\Delta M = 8$  and  $W = 119$  for a possible undeuterated compound. No combination from Table II is compatible with the molecular weight, except 8,15 of methylphenylcarbinol. For this substance a higher  $W$  value is expected (example 5) and so we repeated the procedure by subtracting three mass units from the molecular weight of the unknown compound. For  $M = 120$ ,  $\Delta M = 6$  and  $W = 121$ , the combinations of 8,17 and 8,18 (Table II) are compatible with the molecular weight. Acetophenone is in accordance with both the retention (example 5) and the mass spectra (the peak at  $m/e$  105 is characteristic of the benzoyl radical). Thus the deuteration product is a mixture of [ $^2\text{H}_3$ ]methyl phenyl ketone and undeuterated phenylacetylene.

#### CONCLUSIONS

A method for substance identification has been described, using combinations of structural increments compatible with retention data and empirical formulae.

A retention parameter, appropriate for this purpose was defined and an algorithm of calculus was given.

The possibilities and limitations of the method have been illustrated on few examples, using Carbowax 20M as stationary phase.

Further refinement of the described method requires an extensive study of a selection of a few universally applicable stationary phases, in order to establish exact limits of values for all of the classes of compounds that can be analyzed on these phases and to increase the selectivity. The application of the method will be facilitated by the use of appropriate combinations of detectors, columns and reaction tubes, which enable the determination of the retention data together with the elemental composition. It must not be forgotten that the most powerful procedure for structure elucidation is a combined application of various instrumental methods. Amongst these methods, GC occupies a place of importance comparable with that of the spectroscopic methods.

#### ACKNOWLEDGEMENT

We thank Dr. N. Palibroda for the mass spectra.

#### REFERENCES

- 1 A. Wehrli and E. Kováts, *Helv. Chim. Acta*, 42 (1959) 2709.
- 2 R. Kaiser, *Chromatographie in der Gasphase*, Vol. 3, Part 2, Bibliographisches Institut, Mannheim, Zürich, 2nd ed., 1969, p. 182.
- 3 D. A. Leathard and B. C. Shurlock, *Identification Techniques in Gas Chromatography*, Wiley-Interscience, London, 1970.
- 4 J. K. Haken, *J. Chromatogr. Sci.*, 11 (1973) 144.
- 5 G. Schomburg and G. Dielmann, *J. Chromatogr. Sci.*, 11 (1973) 151.
- 6 J. M. Takács, *J. Chromatogr. Sci.*, 11 (1973) 210.
- 7 L. Soják, J. Hrivňák, I. Ostrovský and J. Janák, *J. Chromatogr.*, 91 (1974) 613.
- 8 L. S. Etre, *Chromatographia*, 7 (1974) 39.
- 9 A. W. Ladon, *J. Chromatogr.*, 99 (1974) 203.
- 10 S. D. West and R. C. Hall, *J. Chromatogr. Sci.*, 13 (1975) 5.
- 11 J. R. Mann and S. T. Preston, *J. Chromatogr. Sci.*, 11 (1973) 216.
- 12 M. B. Evans, *Chromatographia*, 2 (1969) 397.
- 13 O. E. Schupp and J. S. Lewis (Editors), *Gas Chromatographic Data Compilation*, American Society for Testing and Materials, Philadelphia, 1967.
- 14 O. E. Schupp and J. S. Lewis (Editors), *Gas Chromatographic Data Compilation*, Supplement 1, American Society for Testing and Materials, Philadelphia, 1971.