Journal of Chromatography, 134 (1977) 249–261

CHROM. 9840

COMPUTER-ASSISTED PREDICTION OF GAS CHROMATOGRAPHIC SEPARATIONS

R. J. LAUB, J. H. PURNELL and P. S. WILLIAMS
Department of Chemistry, University College of Swansea, Swansea, Wales SA2 8PP (Great Britain)
(Received October 29th, 1976)

SUMMARY

A computer program is presented which allows the prediction of gas–liquid chromatographic retention behavior, and, thus, the separation of complex mixtures, with binary stationary phases. The method is based upon the law of diachoric solutions,

$$K_R = \varnothing_A K^0_{R(A)} + \varnothing_S K^0_{R(S)}$$

where K_R is the liquid–gas partition coefficient for a solute with a stationary phase composed of a mixture of A and S of volume fractions, $\varnothing_A$ and $\varnothing_S$, respectively; $K_{R(A)}$ and $K_{R(S)}$ are the solute partition coefficients for pure A and pure S. Computer-drawn plots of K_R *vs.* $\varnothing_A$ and α *vs.* $\varnothing_A$ are used to find the optimum A + S composition which will separate a given mixture, and, in addition, predict the α value for the two most difficult solute pairs at that point. This enables the column length necessary for the baseline separation of all components to be calculated. With certain limitations, relative retention times and weight fractions, rather than partition coefficients and volume fractions, can be employed. In the examples given, a 5-component steroid mixture and a 40-component hydrocarbon mixture are virtually completely separated by the computer-assisted application of the described procedure. The only limitations to the method appear to be the turn-around time of the computer installation, and the accuracy of the K_R data.

INTRODUCTION

In a recent series of papers[1–4], we have shown that a simple equation,

$$K_R = \varnothing_A K^0_{R(A)} + \varnothing_S K^0_{R(S)} \quad (1)$$

describes the gas–liquid chromatographic (GLC) retention behavior of a wide variety of solutes with an equally varied number of binary stationary phases; K_R is the liquid–gas solute partition coefficient for a stationary phase composed of A of volume fraction, $\varnothing_A$, and S, of volume fraction, $\varnothing_S$. $K^0_{R(A)}$ and $K^0_{R(S)}$ are the corresponding solute

partition coefficients with pure A and pure S, respectively. Eqn. 1 has more recently[5] been shown to apply to a ternary solvent mixture and so may well be more properly written as:

$$K_R = \sum_i \varnothing_i K^0_{R(i)} \tag{2}$$

We have chosen to call eqn. 2 the law of diachoric solutions.

Eqn. 2 is of considerable importance in analytical GLC, since, with it, the optimum composition of A and S which will separate any given mixture can be predicted[6]. Briefly, for a binary stationary phase, the relative volatility, $\alpha_{2/1}$ of solutes 1 and 2 is given by:

$$\alpha_{2/1} = K_{R_2}/K_{R_1} = \frac{\varnothing_A \Delta K^0_{R_2} + K^0_{R(S)_2}}{\varnothing_A \Delta K^0_{R_1} + K^0_{R(S)_1}} \tag{3}$$

where $\Delta K^0_R = K^0_{R(A)} - K^0_{R(S)}$. In practice, we have found it easier to employ the alternative relation,

$$\varnothing_A = \frac{K^0_{R(S)_2} - \alpha K^0_{R(S)_1}}{\alpha \Delta K^0_{R_1} - \Delta K^0_{R_2}} \tag{4}$$

and to calculate values of $\varnothing_A$ at discrete values of α within set limits, these being 1.00 (impossible separation) and 1.10 (4400 theoretical plates required for k', the capacity factor, larger than 10). α values larger than 1.10 are neglected since, in general, with mixtures of sufficient complexity to warrant use of our method, larger optimum values are rare. Further, we find no difficulty in constructing or using columns of the corresponding efficiency. This approach, of course, considerably reduces the required number of calculations.

We have also found that α *vs.* $\varnothing_A$ plots can be treated as triangles rather than curves, where points of inversion on the abscissa are defined at $\alpha = 1$ (when α falls below 1, the pairs are reversed, such that $\alpha \geqslant 1$ is always true). α *vs.* $\varnothing_A$ plots are thus composed of a set of partially overlapping inverted triangles; regions in which no overlap occurs are called windows. The largest minimum α value obtained for any A + S pair is found by comparison of the windows, *i.e.*, by inspection, during which the corresponding $\varnothing_A$ (optimum) is also found. The minimum number of plates required for baseline separation at the best $\varnothing_A$ value, $N_{\text{req.}}$, can then be calculated from the relation[7]

$$N_{\text{req.}} = 36 \left(\frac{\alpha}{\alpha - 1}\right)^2 \left(\frac{k' + 1}{k'}\right)^2 \tag{5}$$

Since the analyst generally knows the number of plates per foot of column he can obtain with a specified packing, liquid loading, etc., eqn. 5 immediately yields the length of column required. Hence, the window diagram procedure, in conjunction with eqn. 5, allows prediction of all parameters needed to separate a given mixture. The method can even be applied to mixtures of totally unknown composition and complexity[8].

We have indicated elsewhere[5,6] that the above described procedure can be computerized. The logic is not as simple as one might suppose, however, and so we herein describe the algorithms we are currently using to produce window diagrams and, hence, predict chromatographic separations. Two examples are given, namely, a 40-component hydrocarbon mixture, and a 5-component steroid mixture. Relative retention times and weight percents, rather than partition coefficients and volume fractions, are used for the latter for reasons which are discussed.

PROGRAM DESCRIPTION

A flow chart of the program is given in Fig. 1. A brief description of the routines is now presented.

Arrays and variables

N: the number of solutes.

CKR(1,1), CKR(1,2): correspond to $K^0_{R(S)}$ and $K^0_{R(A)}$, respectively, for each of the solutes.

XLIM: the maximum α which need be considered (any plot of α which has $\alpha >$ XLIM over the entire volume fraction range is ignored). XLIM must be greater than the highest window. It is also used to specify the vertical axis of the window diagram to the graph plotter.

IJPAIR(L,1), IJPAIR(L,2): pairs of solutes which have relevant plots of α.

MAXL: the number of relevant plots of α.

ALPHA(101,3): ALPHA(101,1): the minimum α at 101 equally-spaced points along the volume fraction axis; ALPHA(101,2), APHA(101,3): the two solutes which comprise each point in ALPHA(101,1).

AINV(M,3): AINV(M,1): the values of $\varnothing_A$ at which inversion of the α plots occur; AINV(M,2), AINV(M,3): the corresponding solutes.

PEAK(M+1,2,3): used to hold descriptions of windows obtained from subroutine PEAKDET (corresponds to M+1 sets of A1, P1, A2, P2, A3, P3).

Main routine

N, CKR(1,1), CKR(1,2), and XLIM are read in. All pairs of solutes corresponding to the relevant plots of α are stored in IJPAIR(L,1), IJPAIR(L,2). K_R is plotted *vs.* $\varnothing_A$ using only CKR(1,1) and CKR(1,2).

The minimum values of α at 101 equally-spaced points along the $\varnothing_A$ axis are calculated and stored in array ALPHA.

The pairs of solutes in IJPAIR are searched for all points of inversion. The values are ordered and stored in array AINV.

Subroutine PEAKDET is called for each window. Subroutine PLOT is called to order the points for graph plotting, and the window diagram is then plotted.

Subroutines

PEAKDET(M,A1,P1,A2,P2,A3,P3): this subroutine produces a description of the window specified by M. On exit, A2 will indicate the highest α value of the window, and P2 the corresponding $\varnothing_A$. A1, P1, A3, and P3 are used to give any further information on the shape of the window. If, on exit, they are zero, the window is

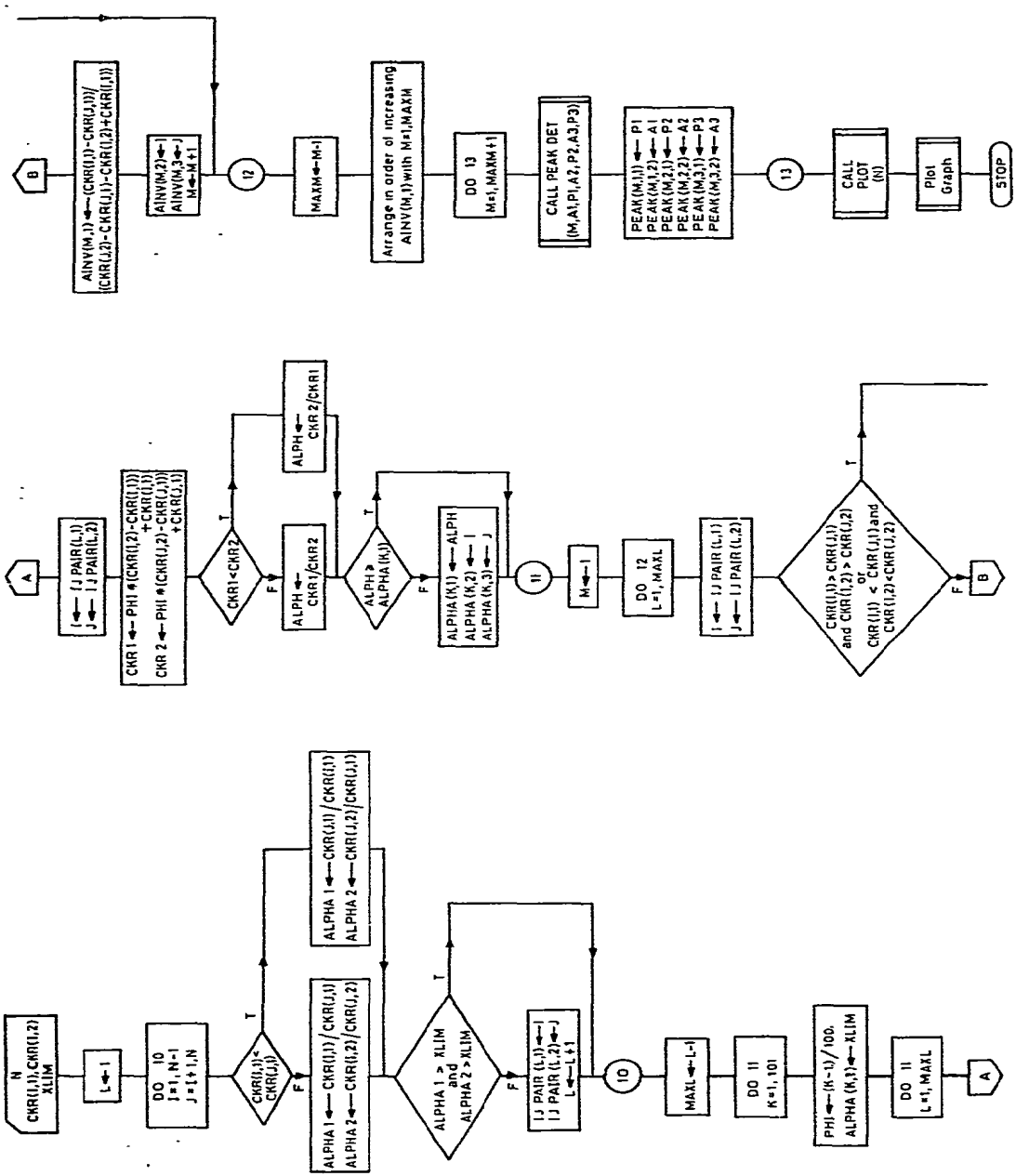

Fig. 1. Flow chart of the program used to predict the optimum column composition for the separation of an *n*-component mixture.

indicated to have a simple "triangular" shape, bounded by the points of inversion adjacent to P2. The actual values (if other than zero) may be inferred from the following brief description of PEAKDET. A1 to P3 initialized to zero.

Special cases tested: if M = 1 and AINV(1,1) = 0.0, return; if M = MAXM and AINV(MAXM,1) = 1.0, return; if AINV(M−1,1) = (AINV(M,1), return. These would all be superfluous windows (*i.e.*, do not exist).

All the points of ALPHA contained in the window which is bounded by the points of inversion are now examined (note that some care must be taken in selecting these if either or both points of inversion are exact at two decimal places). There may of course be no points of ALPHA within the window (special case, see later).

The point of ALPHA within the window with the highest α value is taken, and the slope of the plot which passes through it is calculated. If the slope is positive, the point of ALPHA immediately to the right of this point [or AINV(M,1) if it is numerically smaller] is considered, and the intersection of the two plots corresponding to the two points is calculated using subroutine INTERCEPT. Similarly, for a negative slope, the point to the left is examined (for the case of zero slope, *i.e.*, a horizontally topped window, see later). The values of A2 and P2 are now known.

If the plot which passes through AINV(M−1,1) (with positive slope) does not pass through (P2,A2), then one or more intersections of the plots must have occurred between these two points. Points of ALPHA to the left of P2 are examined successively until a different plot is encountered, and intersection is again calculated using INTERCEPT (which now has values for P1 and A1). Similarly, P3 and A3 will describe any intersection immediately to the right of P2.

For a horizontally topped window, a similar method to the above is used; P1/A1 and P2/A2 are used to store each end of the horizontal part.

The course of action is slightly different for the windows at each end of the diagram, as they are not necessarily bound on each side by points of inversion.

If no member of ALPHA is contained in a window, the intersection of the plots (with correct slopes: positive for AINV(M − 1,1) and negative for AINV(M,1)) through the points of inversion is calculated and described by A2 and P2.

INTERCEPT(I,J,PHI,I2,J2,PHI2,PEAK,PHIP): this subroutine takes in the two pairs of solutes specifying the two plots (I/J and I2/J2) and also the two limits within which the intersection occurs (PHI and PHI2). The mean of the limits is calculated. The slopes of the plots are known at this point, so that the correct form of α may be chosen (*i.e.*, such that $\alpha \geqslant 1$). The intersection is then calculated using simple algebra. The best α and $\varnothing_A$ values are given by PEAK and PHIP, respectively, on exit.

CHECK(PEAK,PHIP): after each call to INTERCEPT, a call to CHECK is made. This subroutine takes in a value of α and a value of $\varnothing_A$, and works through all the relevant plots; if any have a lower value of α at the given value of $\varnothing_A$, the value of α (PEAK) is uprated. This procedure will decrease any error in a case such as that illustrated in Fig. 2, which will be most significant when the slopes of the plots are steep. It will be noted that the points of ALPHA are 0.01 of the volume fraction axis apart: errors in the drawing of the windows will therefore be negligible.

PLOT(N): this subroutine examines the arrays, PEAK and AINV, and orders all relevant points (omitting the zeros produced by PEAKDET). It also outputs the total number of relevant points (N) which are subsequently used by the graph plotting

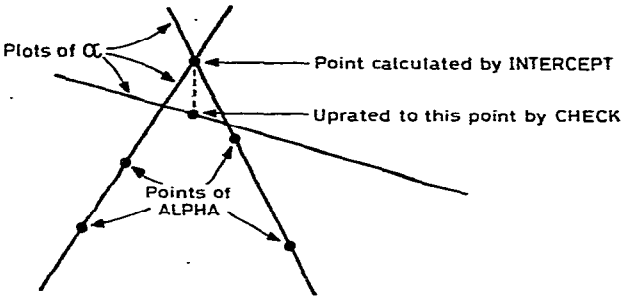

Fig. 2. Illustration of the subroutine CHECK (*cf.* text).

routines. The window diagram is then simply plotted using only these points. Note that each window is described by a maximum of 5 points (3 from PEAKDET and the two points of inversion). This was found to be perfectly acceptable.

EXPERIMENTAL

All computer studies were carried out on an ICL 1904S with a CalComp graph plotter; the program was written in Fortran IV. A Perkin-Elmer Model F-33 gas chromatograph equipped with a flame ionization detector was employed for the GLC work. The column temperature was monitored with a Hewlett-Packard Model 2802A precision Pt resistance system, and was found to be constant to within $\pm 0.05°$. Chromosorb G AW DMCS (120–140 mesh) solid support was used throughout. Columns were constructed from 1/8-in. O.D. stainless-steel tubing for the hydrocarbon mixture, and glass-lined stainless-steel tubing (SGE, North Melbourne, Australia) for the steroid mixture. Liquid phases were obtained from BDH (Poole, Great Britain) [squalane and di-*n*-nonyl phthalate (DNNP)] or Applied Science (State College, Pa., U.S.A.) (L-45 and QF-1), and were used as received. Column packings [1–5% (w/w) for squalane and for DNNP, and 5% for L-45 and for QF-1] were made by dissolving the required amount of stationary phase in a suitable solvent, adding the support, and removing the solvent by rotary evaporation. Binary stationary phases comprised mechanical mixtures of the appropriate amounts of (support + A) and (support + S). Columns were packed by pressurizing (500 p.s.i.g.) a stainless-steel reservoir containing the coated support, and tapping. Squalane–DNNP columns constructed via this technique reproducibly gave in excess of 750 effective plates per foot length for the later eluting solutes of the 40-component mixture. The first four solutes gave in excess of 1300 plates per foot, and, in one case, 1400 plates per foot. The number of plates for a given column packing was found to be additive in the number of plates per foot up to at least 75 feet.

RESULTS AND DISCUSSION

The volume fraction of one component of a binary mixture, $\varnothing_A$, is related to the weight fraction, m_A, by:

$$\varnothing_A = \frac{m_A \varrho_S^0}{\varrho_A^0 + m_A (\varrho_S^0 - \varrho_A^0)} \tag{6}$$

where ϱ_i^0 is the density of the pure ith medium. Thus, if $\varrho_S^0 \approx \varrho_A^0$, $\varnothing_A \approx m_A$; this situation broadly applies to high-molecular-weight polymeric silicone oils or gums and so, weight fraction, rather than volume fraction, can be used in dealing with their mixtures. This obviates the need to measure densities, a difficult procedure at high temperature.

Mechanically mixed phases, of course, can always be treated as completely ideal (or, completely immiscible) liquids. In the special situation that a particular solute has partition coefficients which are identical with both A and S (*i.e.*, K_R *vs.* $\varnothing_A$ is a horizontal line), corrected (for dead volume) retention times of mixture components relative to that solute (α) only, rather than K_R values; need be determined. These two approximations, namely, the use of relative retention times and weight fractions, considerably simplifies the practical application of the optimization procedure. If, as will be usual for any given stationary phase pair, it is found that $K^0_{R(S)}$ and $K^0_{R(A)}$ for the "standard" solute are not identical, the partition coefficients (rather than α values) of mixture components must be determined.

In the first example of computer-assisted GLC optimization, we make use of the data of Touchstone and co-workers[9], who chromatographed (underivatized) steroids on QF-1, L-45, and various intimate mixtures of these, at 240°. They found that estrone had the same retention time with both pure phases as well as with mixtures of the two, and so they reported retention times of other steroids relative to those for this solute. They also found that these α values varied linearly with the weight percent of L-45 in QF-1 over the entire range, 0–100%. Table I gives the α values of five of the steroids with the pure phases, and Fig. 3 shows the computer-drawn plot of α *vs.* m_A. Only the two end-points of each line were used to construct this figure. We see that there are two solute pairs which would be virtually impossible to separate with pure QF-1, and at least one pair would also be difficult to separate with pure L-45. Fig. 4 shows the computer window diagram for the five solutes, and we find that there are several column compositions at which separation can be achieved. The highest minimum α value according to the computer print-out is 1.124, which occurs at $m_A = 0.617$. Since an α value of 1.124 requires 3000 plates, and since we obtained 250 plates per foot with these phases, a column 12 ft. long containing a mechanical mixture of L-45 and QF-1 packings such that $m_A = 0.617$ is needed. Fig. 5 shows achievement of the predicted chromatogram of the five steroids; resolution is virtually complete. Also given are the predicted positions of the peaks relative to estrone (not shown); the error on No. 4 is 3.0% and on No. 5 is 2.3%.

TABLE I

CORRECTED RETENTION TIMES[9] RELATIVE TO ESTRONE AT 240°

No.	*Solute*	*α (estrone = 1)*	
		QF-1	*L-45*
3	Cholesterol	1.08	3.64
1	3α,20α-Dihydroxypregnane	0.80	1.31
2	Androsterone	0.82	0.83
5	17β-Hydroxyandrostan-3-one	1.06	0.95
4	Androstan-3,17-dione	1.79	0.92

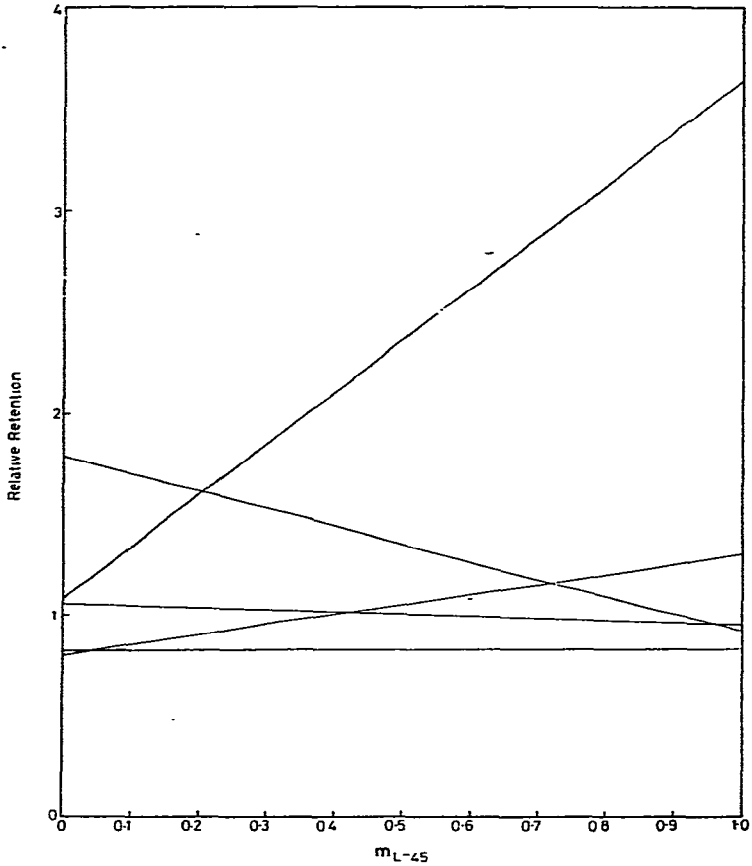

Fig. 3. Computer-drawn plot of relative retention time (estrone = 1.00) *vs.* weight fraction of L-45 in QF-1 for the five steroids of Table I.

A computer would not normally be used for the simple example given above, since the calculations may be done by hand in a few minutes. However, when the number of components in a mixture exceeds about 10, the required arithmetic may take several hours. For example, Table II lists $K^0_{R(S)}$ and $K^0_{R(A)}$ values for 40 hydrocarbons with squalane and DNNP at 100°. This mixture represents a very difficult separation, and, furthermore, an exceedingly tedious task if the window diagram procedure were to be applied by hand, Now, however, the limiting factor in the application of our procedure is the turn-around time of the computer installation, which is generally between a few minutes and two or three hours.

Fig. 6 shows the computer plot of K_R *vs.* $\varnothing_A$ for the 40 hydrocarbons, and Fig.7 the computer window diagram. The highest minimum α value is a plateau of 1.032 at $\varnothing_A = 0.10$–0.17 which would require about 38,000 plates if $k' > 10$. Since we found an average of 850 plates per foot for these packings, a column 50 ft. long containing a mechanical mixture such that $\varnothing_A = 0.17$ is dictated. Fig. 8 shows the chromatogram obtained with such a column, where all components have been resolved. Several of the early peaks were not quite baseline-separated, which we attribute to k' values being

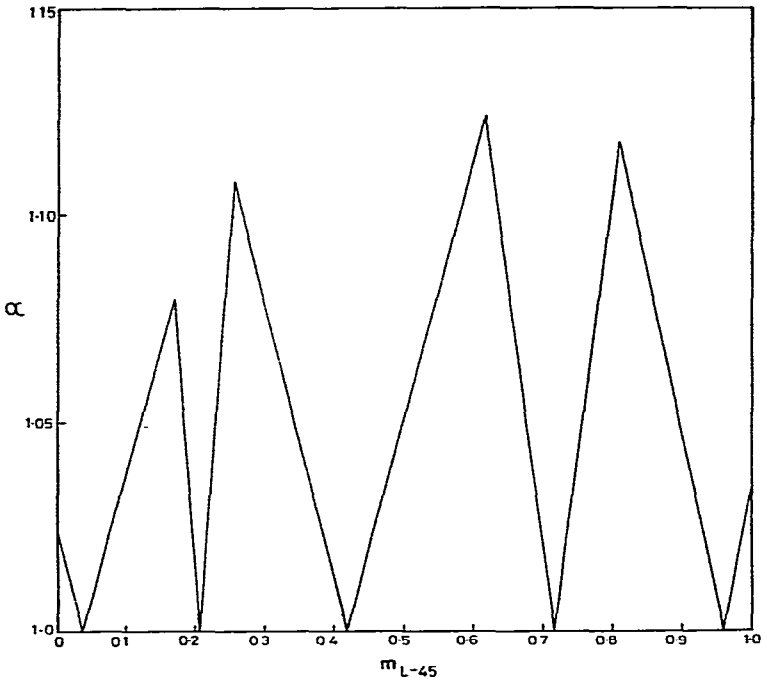

Fig. 4. Computer-drawn window diagram for the solutes of Fig. 3. Best α value is predicted to be 1.124 at $m_A = 0.617$.

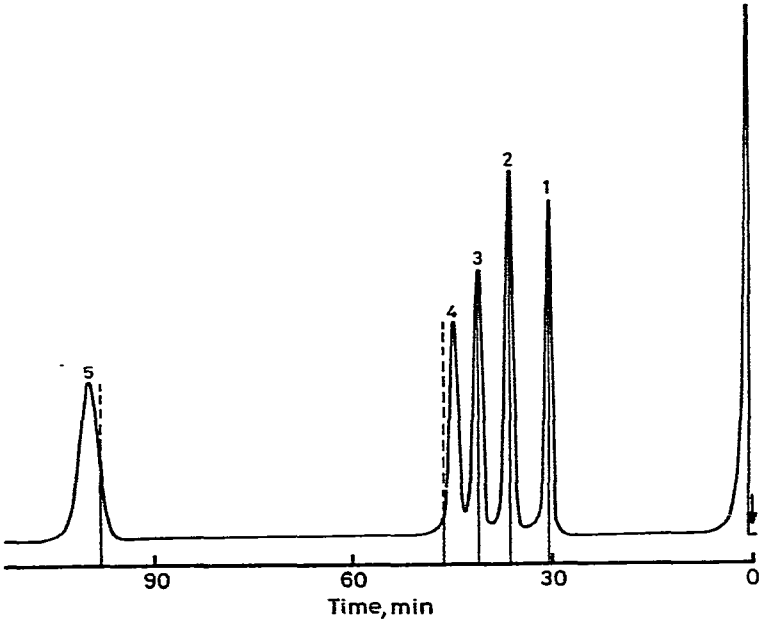

Fig. 5. Chromatogram of the solutes of Fig. 3 at $m_A = 0.617$. Column, 12 ft. × 1/8 in. O.D. glass-lined stainless steel; packing: 5% (w/w) mechanically mixed solvents on 120–140-mesh Chromosorb G AW DMCS; temperature, 240°. Vertical lines indicate computer-predicted positions of the peaks. Peak numbers, see Table I.

TABLE II

PARTITION COEFFICIENTS FOR NAMED SOLUTES WITH SQUALANE (S) AND Di-*n*-NONYL PHTHALATE (A) AT 100°

No.	*Name*	$K^{\circ}_{R(S)}$	$K^{\circ}_{R(A)}$
1	3-Methyl-1-butene	10.4	7.54
2	1-Pentene	13.4	9.81
3	1,3-Pentadiene (*cis* and *trans*)	18.1	16.6
4	4-Methyl-1-pentene	23.5	16.5
5	2,3-Dimethylbutane	27.3	16.5
6	1-Hexene	30.5	21.6
7	3-Methyl-2-pentene (*trans*)	38.5	28.0
8	2-Methyl-1,3-pentadiene	43.5	39.8
9	Benzene	50.8	58.1
10	Cyclohexane	60.7	39.8
11	1-Heptene	65.7	46.4
12	*n*-Heptane	74.6	46.4
13	2,4,4-Trimethyl-2-pentene	85.3	58.1
14	Methylcyclohexane	99.4	63.1
15	2,3,4-Trimethylpentane	116	72.2
16	Toluene	116	128
17	1-Octene	140	97.6
18	1-Octyne	143	149
19	2,6-Dimethyl-3-heptene (*cis* and *trans*)	163	107
20	4-Octyne	176	155
21	3-Octyne	185	164
22	4-Vinyl-1-cyclohexene	201	164
23	2-Octyne	222	211
24	Ethylbenzene	229	248
25	*p*-Xylene	254	271
26	1-Nonene	300	200
27	*o*-Xylene	300	336
28	*n*-Nonane	339	200
29	Isopropylbenzene	357	376
30	*n*-Propylbenzene	443	463
31	Ethoxybenzene	501	747
32	1,3,5-Trimethylbenzene	564	596
33	1-Decene	626	416
34	*n*-Decane	712	413
35	*p*-Cymene	773	788
36	1,2,3-Trimethylbenzene	800	884
37	1,3-Diethylbenzene	885	913
38	*n*-Butylbenzene	931	956
39	1,4-Diethylbenzene	960	988
40	Acetophenone	897	1890

less than 10. This was partially offset, however, by the larger (than 850) number of plates per foot for the solutes, and, thus, at least 98% (6σ) resolution was achieved.

The data of Table II were determined relative to toluene, and we note that there are, in some cases, discrepancies between these and the $K^{0}_{R(S)}$ values of Miyake *et al.*[10] which we have cited elsewhere[3,6]. The $K^{0}_{R(S)}$ (and $K^{0}_{R(A)}$) value for toluene we employed for this work was taken from measurements we have recently made with a precision apparatus[11], which are in agreement with those reported by Martire and co-workers[12,13].

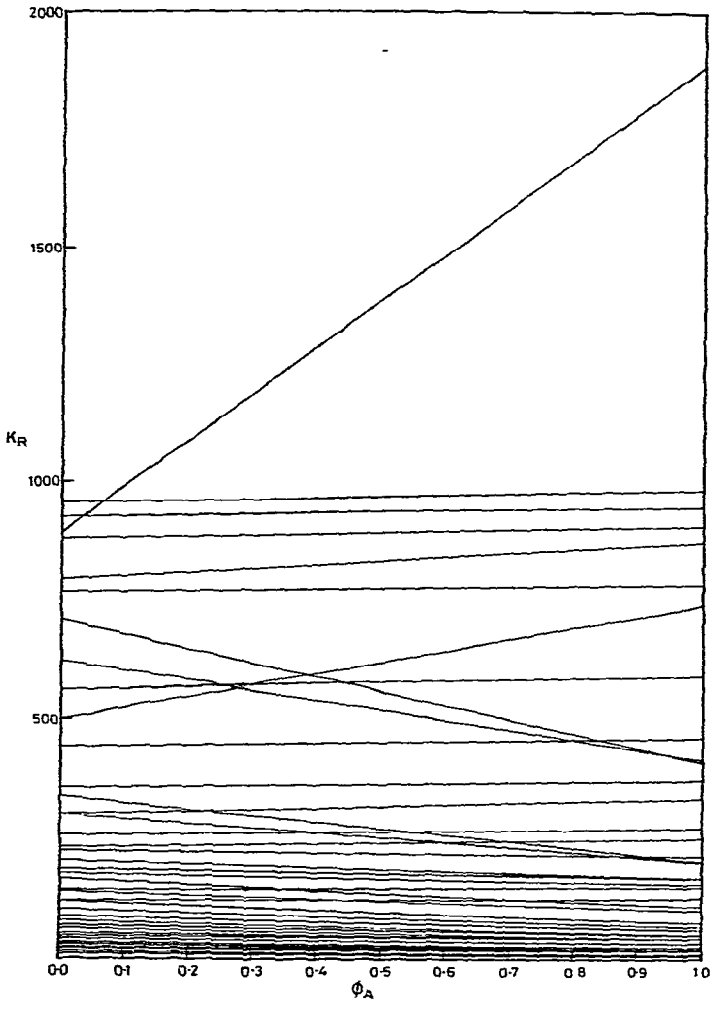

Fig. 6. Computer-drawn plot of K_R *vs.* $\varnothing_A$ for the forty solutes of Table II.

Virial effects[14] may result in differences in predicted *vs.* actual retention behavior; these will be approximately self-cancelling if K^0_R values are determined with the same carrier gas at the same column pressures as used for the analytical work. The effects are negligible when helium or hydrogen carriers are employed.

Solid support adsorption effects may cause errors in predicting retention behavior when low liquid loadings are used. The data in Table II were determined (relative to toluene) with columns containing from 1 to 5% (w/w) solvent for both squalane and DNNP, and agreement was within 3% for all solutes except acetophenone. Eqn. 2 should be obeyed for solutes which exhibit surface adsorption effects when mechanically mixed packings are employed as long as the amount injected is identical for each column, and the same liquid loadings used to determine K^0_R data comprise the column employed for the separation.

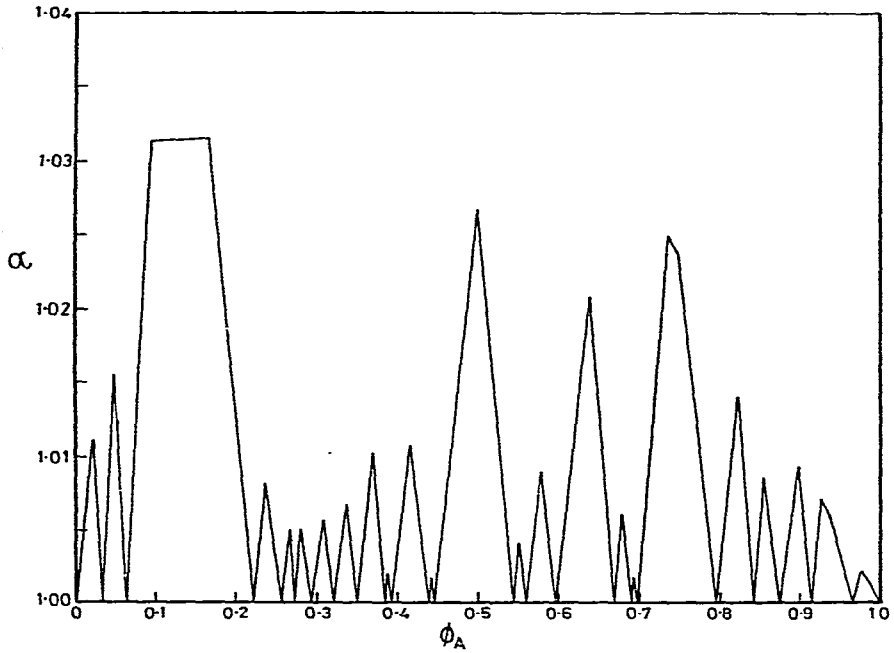

Fig. 7. Computer-drawn window diagram for the solutes of Fig. 6. The best α value is predicted to be a plateau of 1.032 at $\varnothing_A = 0.10$–0.17.

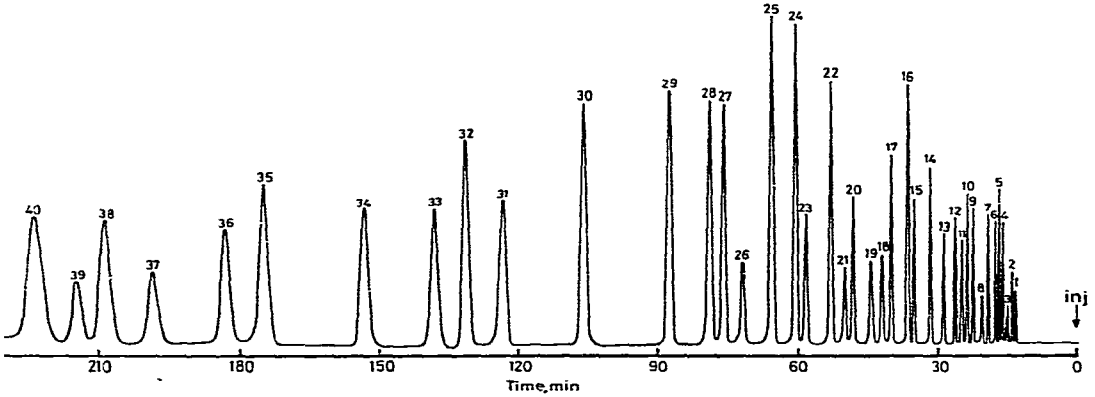

Fig. 8. Chromatogram of the solutes of Fig. 6 at $\varnothing_A = 0.168$. Column, 50 ft. × 1/8 in. O.D. stainless steel; packing, 1.5% (w/w) mechanically mixed solvents on 120–140-mesh Chromosorb G AW DMCS; temperature, 100°. Peak numbers, see Table II.

As can be seen, however, very complex mixtures require accurate data and careful packing preparation, since, generally, the more useful windows may have steep sides. These, we feel, are the only limitations to the application of our procedure.

ACKNOWLEDGEMENTS

We gratefully acknowledge financial support from the Science Research Council, and the Foxboro Company, Foxboro, Mass., U.S.A.

REFERENCES

1 J. H. Purnell and J. M. Vargas de Andrade, *J. Amer. Chem. Soc.*, 97 (1975) 3585.
2 J. H. Purnell and J. M. Vargas de Andrade, *J. Amer. Chem. Soc.*, 97 (1975) 3590.
3 R. J. Laub and J. H. Purnell, *J. Amer. Chem. Soc.*, 98 (1976) 30.
4 R. J. Laub and J. H. Purnell, *J. Amer. Chem. Soc.*, 98 (1976) 35.
5 R. J. Laub and J. H. Purnell, *Anal. Chem.*, 48 (1976) 799.
6 R. J. Laub and J. H. Purnell, *J. Chromatogr.*, 112 (1975) 71.
7 J. H. Purnell, *J. Chem. Soc.*, 1268 (1960).
8 R. J. Laub and J. H. Purnell, *Anal. Chem.*, 48 (1976) 1720.
9 J. C. Touchstone, C.-H. Wu, A. Nikolski and T. Murawec, *J. Chromatogr.*, 29 (1967) 235.
10 H. Miyake, M. Mitooka and T. Matsumoto, *Bull. Chem. Soc. Jap.*, 38 (1965) 1062.
11 R. M. Franklin, R. J. Laub and J. H. Purnell, in preparation.
12 J. P. Sheridan, D. E. Martire and Y. B. Tewari, *J. Amer. Chem. Soc.*, 94 (1972) 3294.
13 J. M. Janini, J. W. King and D. E. Martire, *J. Amer. Chem. Soc.*, 96 (1974) 5368.
14 R. J. Laub and R. L. Pecsok, *J. Chromatogr.*, 98 (1974) 511.