

CHROM. 12,253

## RESOLUTION OF OVERLAPPING CHROMATOGRAPHIC PEAKS BY PARAMETER ESTIMATION OF THEIR MODEL

M. ROSENBAUM, V. HANČIL and R. KOMERS

*Institute of Chemical Process Fundamentals, Czechoslovak Academy of Sciences, Prague (Czechoslovakia)*

---

### SUMMARY

Overlapping chromatographic peaks were resolved by means of a Hewlett-Packard 9825 calculator using a maximum likelihood method. A functional model of the peaks to be resolved, had to be stated prior to the resolution. The model consisted of the structural parameters (position, width, height), which could be used for qualitative and quantitative evaluation, and of nuisance parameters which vary from case to case. The optimization procedure used was based on a Marquardt modification of the Newton Raphson method. This study is supplemented by examples of peak resolving of two hardly separable mixtures, *e.g.*, *m*- and *p*-cresols, and *m*- and *p*-xylenes.

---

### INTRODUCTION

Evergrowing significance of instrumental analysis in chemical sciences implies the introduction of integrated measurement systems consisting of micro-processors and sometimes calculators. Data are often gathered by the conversion of analog information into digital information. The data are then programmed by standard methods. However, such procedures cease to give acceptable answers when one works on the limit of the resolving power of the measuring system.

In our case the analyser was a gas chromatographic apparatus (see Experimental) with a linear recorder output. Chromatograms were evaluated by means of a Hewlett-Packard 9825A calculator which was equipped with a Hewlett-Packard 9872A plotter-digitalizer.

### MODEL OF SYSTEM

Direct evaluation of chromatograms assumes a defined background (baseline) and total resolution of peaks in the chromatogram. In such a case, geometrical properties of the records obtained can be used easily for analytical purposes. Physical properties of the chromatographed compounds, however, are often very similar. To achieve their desired resolution one should use rather expensive and sophisticated apparatus, which might not necessarily be available, or the problem

can be solved by the general method of resolving overlapping peaks. This method requires the assumption of functional form of the model of the response of the individual component in the analysed mixture. The chromatogram can be regarded as the superposition of responses ( $R$ ) of  $L$  components.

$$R(x) = \sum_{i=1}^L \Phi_i(x) \quad (1)$$

Of key significance to a successful resolution is the selection of a suitable mathematical description of single component response, *i.e.*, definition of its model. In the present study it is assumed that chromatographic behaviour of a component can be described as a function of position,  $x$ , and structural parameter,  $p$ , as follows:

$$\Phi_i(x) = \Phi_i(x, p_{1i}, p_{2i}, \dots, p_{ki}) \quad (2)$$

Background of measured response can be considered as a component of the system, the model of which is described by particular parameters,  $q$ . Background can be, *i.e.*, the linear function of the position,  $x$ , on record.

$$\Phi_b(x) = q_1(x) + q_2 \quad (3)$$

The  $q$ 's, so called nuisance parameters, on one hand, increase the number of unknown parameters and make the analysis of data more complex, but on the other hand, they enable estimation of structural parameters,  $p$ , which are freed from background influence.

In eqn. 2, the dimension of  $k$  equals 3 for the majority of the models (position, height, width), hence the number of parameters which are to be estimated per component is 3.

#### PARAMETER ESTIMATION

Parameters can be estimated by several methods<sup>1</sup>. However, most suitable seems to be the maximum likelihood method which, at reasonable simplifying assumptions<sup>1</sup>, leads to the least squares method. An objective function of this method is given by the equation

$$S = \sum_{j=1}^N \left[ M(x_j) - \sum_{i=1}^L \Phi_i(x_j, \bar{p}_i) - \Phi_b(x_j, \bar{q}) \right]^2 \quad (4)$$

where the quantity of  $M(x_j)$  is the normalized ordinate of the response at chosen  $x_j$ . The interval of point reading,  $M(x_j)$ , can be varied depending on chromatogram complexity in zones of interest. Maximum likelihood estimates of parameters  $p$  and  $q$  define the value of minimum objective function  $S(\bar{p}, \bar{q})$ , (eqn. 4), given data  $M(x_j)$ ,  $x_j$  for all  $j$ , *i.e.*,  $j = 1, 2, \dots, N$ .

Minimalization represents solution of a system of nonlinear equations which must be reached iteratively, solving linear approximations at a series of, with respect to objective function value (eqn. 4), ever better iteration values of the parameters. One of the best methods we have had at our disposal is the Marquardt

procedure<sup>2</sup>, a combination of the Newton Raphson method and the steepest descent method. It combines advantages of both methods, *i.e.*, fast convergence in the cases of well conditioned Hessian of the linear approximations, and reliability in other cases<sup>2</sup>. We have used a modified version<sup>3</sup> of FORTRAN subroutine BSOLVE which was translated into HPL language.

#### PROGRAM

The evaluation of the chromatograms is divided into three phases, I, II, and III, which can be realized separately, in sequence. One segment of the program corresponds to each phase. The first one provides digital representation of the chromatogram using a peripheral unit plotter-digitalizer HP 9872A. Values of the abscissas and the corresponding ordinates of a chromatographic record are collected into vector,  $X$ . Odd components are reserved for abscissa quantities, even ones are the ordinates (responses). The base of the ordinate digitalization is parallel to the record paper perforation. The scale of digitalization can be chosen by selection of two reference points,  $P_1$  and  $P_2$ , which are set by the user and to which reference values are assigned. In this step any conversion of data can also be realized. Conversion is optional and must be defined by the user. After this optional step, guesses of parameters  $p$  and  $q$  are introduced into vector  $B[J]$ . Restrictions to the parameters are introduced by means of the array  $M[2,J]$ , where  $M[1,J]$  is the minimum value and  $M[2,J]$  the maximum value of the parameters. Into vector  $V[J]$ , a code is introduced; if  $V[J]$  equals zero, the  $i$ th parameter is kept constant during the whole treatment; otherwise the parameter is optimized. In the next step, memory content, defined by  $B[*]$ ,  $M[*]$ ,  $V[*]$  and  $X[*]$ , is recorded on a digital cartridge (trc 0; ref 2). By repetition of the described digitalization procedure, better digital representation may be obtained. Having completed the previous digitalization the second segment is loaded into the internal calculator memory from the data cartridge memory with a minimalization program; see Appendix. This program which goes up to line 21 [goto "PROGRAM"], declares by specifying in a subscript declarator, the number of dimensions in the arrays and the size of each dimension. In this section constant  $F$ , for minimalization procedure control, is also evaluated from input data. The section between "BSOLVE" (line 22) and "ret" (line 165) is a memory-saving modification of FORTRAN subroutine BSOLVE by Ball described in ref. 3. This program can be easily modified by introduction of string memory-saving representation into all statements where  $P$  array variables are used. With this version, one can evaluate up to 6 overlapping bands, *i.e.*, 18 parameters simultaneously. However, this subroutine is rather difficult to interpret and therefore a simpler version is described. This simpler version is not capable of resolving more than two peaks, due to memory limitation.

The third segment of the program provides a print of total output information in comprehensive form together with a graph of calculations. Circles in this graph (see Figs. 1-3) are measured points which have been read off the record. The full line represents resulting model dependence as superposition of individual component models (dotted lines) and background (linear dependence). Any systematic deviations are quickly recognized by disagreement in the course of the full line and points. Correct guesses of lower and upper limits of parameters define a space in which minimum of  $S$ , eqn. 4 should be located. Thus computational time was shortened.

The ability of this method to resolve overlapping responses mathematically is made possible by properties of minimizing procedure as well as by the fact that additional nuisance parameters were used to describe the background. The nuisance parameters describe that part of the response which is not an intrinsic component of the studied system, but is given by actual experimentation. If we had not used these nuisance parameters, we would have obtained background dependent structural parameters, *i.e.*, parameters which are somehow dependent on experimentation. Such situations should be avoided in all cases.

## EXPERIMENTAL

### *Materials*

*m*-Xylene, *p*-xylene and *m*-cresol (all not less than 99.5% GC analysis) were obtained from the Institute of Chemical Technology (Prague, Czechoslovakia). *p*-Cresol (reagent grade, 99.5%) was supplied by Ferak (West Berlin, G.F.R.). Phosphoric acid (reagent grade) was supplied by Lachema (Brno, Czechoslovakia). Bentone 34 and dilauryl phthalate were obtained from Applied Science Labs. (State College, Pa., U.S.A.), dinonyl phthalate (for GLC) from BDH (Poole, Great Britain).

### *Apparatus*

The gas chromatographic measurements were carried out with a PYE Model 64 heated dual flame-ionization detector programmed chromatograph (Series 104, Pye Unicam, Cambridge, Great Britain). It was equipped with a 0-1 mV recorder (Honeywell, Elektronik 194). The columns employed were 170 × 0.4 and 170 × 0.3 cm I.D. glass tubes for the separation of the xylene and cresol mixtures, respectively. The column temperature of the former was 105° and for the latter 136°. Nitrogen was used as the carrier gas at a flow-rate of approximately 30 ml/min. The samples were introduced with a 10- $\mu$ l Hamilton microsyringe. The injection port and detector were maintained at 200 and 150°, respectively.

### *Column packings*

The following mixed stationary phases were used: Bentone 34 and dinonyl phthalate<sup>4</sup> (2:1) for the separation of *m*- and *p*-xylenes, and dilauryl phthalate and phosphoric acid (9:1) for the separation of *m*- and *p*-cresols. Gas-Chrom Q of the particle size 80-100 mesh (Applied Science Labs.) served as the support in both cases. The packings were prepared in the usual manner by dissolving the individual components of each pair in chloroform and slurring with the support. The former phase contained 5% and the latter 10% of the corresponding mixed phases. The packed columns were preconditioned at 120° for 12 h with the carrier gas flowing.

### *Calibration mixtures*

For the determination of *p*-xylene (PX) and *m*-xylene (MX), and *p*-cresol (PC) and *m*-cresol (MC), the calibration mixtures were prepared in the ratios of approximately 1:9, 1:1 and 9:1, respectively of the mass concentration ratio of PX/MX (Ia, Ib and Ic) or PC/MC (IIa, IIb and IIc). Therefore, the following amounts of individual compounds were precisely weighed and mixed: 0.02927 g of

PX and 0.26005 g of MX (Ia), 0.05002 g of PX and 0.05428 g of MX (Ib), 0.14595 g of PX and 0.01785 g of MX (Ic); 0.02635 g of PC and 0.26979 g of MC (IIa), 0.06353 g of PC and 0.06423 g of MC (IIb), 0.12420 g of PC and 0.02295 g of MC (IIc).

## RESULTS

Six mixtures were analyzed, their chromatograph records digitalized and resolved. All records were treated according to Gaussian models of peaks as follows:

$$\Phi = \Phi_{max} \cdot \exp [(x - x_{max})/\sigma]^2 \quad (5)$$

Values of parameters  $\Phi_{max}$ ,  $x_{max}$ ,  $\sigma$  which are optimum with respect to the objective function  $S$ , (eqn. 4) are given in Table I. Results of xylene mixtures treatments are also shown in Figs. 1-3.

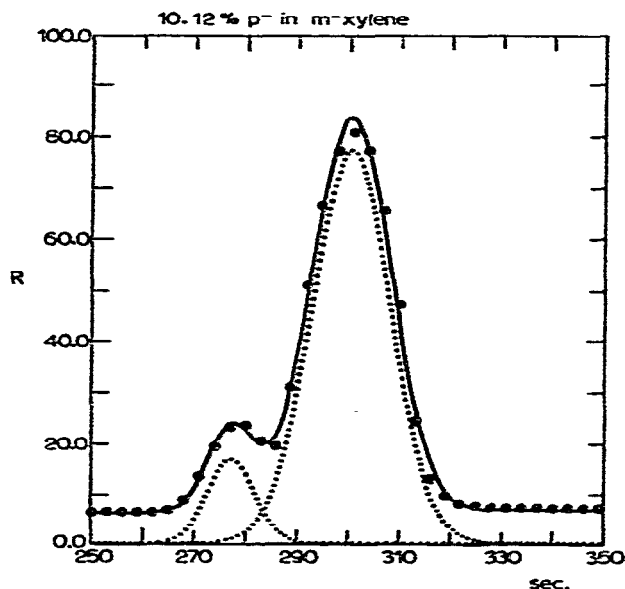


Fig. 1. Chromatogram of mixture Ia. Lines: ●●●● = experimental data; ···· = resolved component responses; — = sum of background and resolved responses.

## CONCLUSION

The described procedure provides a possibility for the treatment of overlapping peaks on chromatographs which are otherwise difficult to handle. The use of nuisance parameters,  $q$ , enabled us to obtain structural parameters without incidental effect of particular experimentations. This procedure can be used not only in gas chromatographic peaks resolution, but also in spectroscopy, liquid chromatography and various other techniques. Connection of the on-line calculator is also possible.

TABLE I  
CALCULATED CHARACTERISTICS OF THE TREATED CHROMATOGRAMS

Sample	Composition % of <i>p</i> -isomer	Retention related to max. (sec)		Height of peak		Standard deviation		Ratio of rel. content of <i>p</i> -isomer to rel. area*	Figure
		<i>p</i> -	<i>m</i> -	<i>p</i> -	<i>m</i> -	<i>p</i> -	<i>m</i> -		
Ia	10.12	277.3	300.6	17.01	77.49	6.47	10.86	0.82	1
Ib	47.96	279.9	302.4	60.82	73.45	9.30	10.31	0.89	2
Ic	89.10	281.5	303.4	89.39	26.32	12.36	6.75	0.86	3
IIa	8.9	6221	6691	23.83	81.30	3.12	8.82	0.95	
IIb	49.7	6250	6716	62.78	55.94	5.26	6.33	0.97	
IIc	84.4	6230	6692	72.30	28.3	7.01	3.1	0.98	

\* Area is proportional to the product of the height of the peak and the standard deviation.

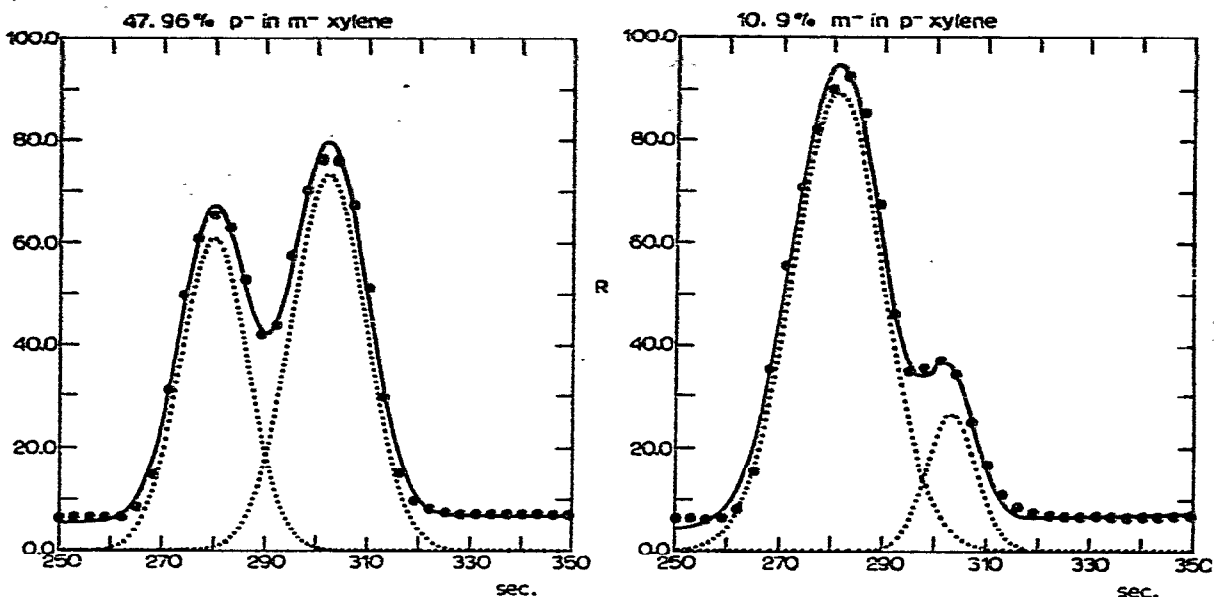


Fig. 2. Chromatogram of mixture Ib. Line symbols same as in Fig. 1.

Fig. 3. Chromatogram of mixture Ic. Line symbols same as in Fig. 1.

## LIST OF SYMBOLS

### Symbols

- $L$  = number of components of chromatogram including background;
- $M$  = measured response;
- $N$  = number of digitalized points on measured response;
- $p$  = component of vector of structural parameters  $\bar{p}$ ;
- $q$  = component of vector of nuisance parameters  $\bar{q}$ ;
- $R$  = response;
- $S$  = objective function (eqn. 4);
- $x$  = value of coordinate;
- $\Phi$  = part of response caused by one component or background;
- $\sigma$  = standard deviation, parameter of Gaussian model defined by eqn. 5.

### Indices

- $b$  = value related to background;
- $i$  = index of component;
- $j$  = index of digitalization;
- max = value related to maximum of  $\Phi$ .

## APPENDIX

*Listing of minimization program*

```

0: "CHROMATOGRAM RESOLUTION FOR 2 PEAKS":
1: dse "RESOLUTION FOR 2 PEAKS";stp
2: trk 0;flt 4
3: ent "number of peaks";K;3K+2+K;jmp KK15
4: ent "number of points";N;3mp NK200
5: 2+M;sfz 14
6: K(N+2)+N+W
7: dim F[5],B[K],M[2,K],V[K],X[MN]
8: dim A[K,K+2],C[K,K+2],P[W],Y[N]
9: dim K[6],I[2]
10: ent "EPS?";E
11: ent "FLA?";F[2]
12: ent "SUMMIN";F[4]
13: prt "SUMMA DIFf2"
14: ldf 2,B[1],M[1],V[1],X[1]
15: 0+I[2]
16: 1+K+G;1+G+H;KN+K[3];K[3]+K+K[4];K[4]+K+K[5]
17: 10+F[1];1+R
18: if E<=0;.001+E
19: if F[2]K=0;.01+F[2]
20: .001+F[3]
21: eto "PROGRAM"
22: "BSOLVE":0+K[6]
23: for I=1 to K
24: if V[I]#0;1+K[6]+K[6]
25: next I
26: eto "A";if K[6]K=3;-3+K[1];eto "Y"
27: "A":eto "B";if N<K[6];-2+K[1];eto "Y"
28: "B":1+I[1]
29: "C":if I[2]>0;eto "U"
30: for I=1 to K
31: K[3]+I+J;B[I]+P[J];K[4]+I+J
32: 1.0|abs(B[I])+.01+P[J]
33: next I
34: eto "Q"
35: "D":if F[4]>F[5] and I[2]>1;eto "I"
36: for L=1 to K
37: (L-1)N+C
38: if V[L]=0;eto "H"
39: for J=1 to K
40: K[3]+J+A;B[J]+P[A]
41: next J
42: L+K[3]+A;L+K[4]+B
43: P[B]+D;if abs(P[A])>F[3];abs(P[A])+D
44: .001D+D

```



```

45: P[A]+D→B
46: if B<=M[L] or B>=M[L];sto "E"
47: P[A]-D→P[A];-D→D;sto "F"
48: "E":P[A]+D→P[A]
49: "F":C→B;K[3]→A
50: for I=1 to N
51: call 'func'(I,A,B)
52: next I
53: for J=1 to N
54: J+C→A
55: (P[A]-Y[J])/D→P[A]
56: next J
57: "H":next L
58: "I":for I=1 to K
59: (I-1)N→C;0→A[I,G];if V[I]=0;sto "J"
60: for J=1 to N
61: C+J→B
62: A[I,G]-P[B]*Y[J]→A[I,G]
63: next J
64: for J=1 to K
65: 0→A[I,J];(J-1)N→H
66: for L=1 to N
67: C+L→B;M+L→A
68: A[I,J]+P[B]*P[A]→A[I,J]
69: next L
70: next J
71: if A[I,I]>1e-20;sto "K"
72: "J":for J=1 to G
73: 0→A[I,J]
74: next J
75: 1→A[I,I]
76: "K":next I
77: 0→F
78: for I=1 to K
79: F+A[I,G]↑2→F
80: next I
81: for I=1 to K
82: A[I,I]→A[I,H]
83: next I
84: for I=1 to K
85: A[I,G]/A[I,H]→A[I,G]
86: for J=1 to K
87: A[I,J]/(A[I,H]*A[J,H])→A[I,J]
88: next J
89: next I
90: F[2]/F[1]→R;sto +2
91: "L":F[1]*R→R
92: for I=1 to K
93: for J=1 to G
94: A[I,J]→C[I,J]
95: next J

```

```

96: R+CEI, I]→CEI, I]
97: next I
98: for I=1 to K
99: I+I→C
100: for J=C to G
101: CEI, J]÷CEI, I]→CEI, J]
102: next J
103: for J=1 to K
104: if I=J]sto "M"
105: for L=C to G
106: CEJ, L]-CEI, L]÷CEJ, I]→CEJ, L]
107: next L
108: "M":next J]next I
109: Ø→S→T
110: for I=1 to K
111: CEI, G]÷ACI, H]→CEI, H]
112: KI3]÷I→A
113: BEI]÷CEI, H]→B; if B>ME2, I]; ME2, I]→B
114: ME1, I]→PEA]]; if PEA]KB; B→PEA]
115: T+CEI, H]÷ACI, G]÷ACI, H]→T
116: S+CEI, H]÷CEI, H]→S; PEA]-SEI]→CEI, H]
117: next I
118: T/√(SF)→T; Ø→S; if T<Ø; 2→S; -T→T
119: if T>1; I→T
120: Ø→KI2]; rad; if T<-1; -I→T
121: if T>=-1 and T<Ø; I→KI2]
122: if T<Ø; abs(T)→T
123: if T=Ø; sto "N"
124: atn(√((1-T+2)/T))→T; if KI2]=1; π-T→T
125: sto "P"
126: "N":π/2→T
127: "P":180T/π→T; if S>Ø; 180-T→T
128: "Q":KI3]÷A; KI5]÷B
129: for I=1 to N
130: cll 'func'(I, A, B)
131: next I
132: Ø→U
133: for I=1 to N
134: I+KI5]→A
135: U+PEA]↑2→U
136: next I
137: if U<1e-10; sto "T"
138: if I[2]>Ø; sto "R"
139: K→KI1]; sto "X"
140: "R": if U>=FI5]; sto "U"
141: Ø→KI1]
142: for I=1 to K
143: I+KI3]→A
144: if abs(CEI, H])÷(FI3]+abs(PEA]))>E; I+KI1]→KI1]

```

```

145: next I
146: if K[I]=0; goto "S"
147: if R>1 and T>90;-1+K[I]
148: goto "W"
149: "S":if R>1 and T<=45;-4+K[I]
150: goto "W"
151: "T":0+K[I]; goto "W"
152: "U":if I[I]>2; goto "V"
153: 1+I[I]->I[I]
154: goto "C";if I[I]>1; goto "D";if I[I]>2; goto "L"
155: "V":if R<1e5; goto "L"
156: -5+K[I]
157: "W":R+P[I]
158: for I=1 to K
159: K[I]+I+R;P[A]->B[I]
160: next I
161: "X":for I=1 to N
162: I+K[I]->R;P[A]->Y[I]
163: next I
164: U+P[I];1+I[I]->I[I]
165: "Y":ret
166: "func":p2+1+p3
167: P[p2+1]*exp(-((P[p2+2]-X[2p1-1])/P[p2+3])^2)+r1
168: P[p2+4]*exp(-((P[p2+5]-X[2p1-1])/P[p2+6])^2)+r1+r1
169: P[p2+7]*X[2p1-1]+P[p2+8]+r1-X[3p1]+P[p3]
170: ret
171: "PROGRAM":c11 'BSOLVE'
172: prt P[5];if K[1]>0; goto "PROGRAM"
173: wrt 6,"parameters B"
174: fmt 1,4f10.4
175: wrt 6.1,B[1],B[2],B[3],B[4]
176: wrt 6.1,B[5],B[6],B[7],B[8]
177: wrt 6,"SUMA",P[5]," ICON",K[1]
178: rew;trk 1;fdf 1;mk 1,290;rof 11;P[*],D[*]
179: end
*15939

```

## REFERENCES

- 1 M. G. Kendall and A. Stuart, in *The Advanced Theory of Statistics*, Vol. 2, Hafner, New York, 1961.
- 2 D. W. Marquardt, *J. Soc. Appl. Math.*, 11 (1963) 2, 431.
- 3 E. J. Henley and E. M. Rosen, *Material and Energy Balances Computations*, Wiley, New York, 1969.
- 4 P. Strnad, *Collect. Czech. Chem. Commun.*, 30 (1965) 2132.