Comparison of Methods for Resolving Overlapping Curves. Application to an Infrared Study of Rotational Isomerism in 2-Oxo-furans and -thiophens

By **Derek J. Chadwick,** * **John Chambers,** and **Roger L. Snowden,** Dyson Perrins Laboratory, Oxford University, South Parks Road, Oxford OX1 3QY

Graphical, analogue computer, and digital computer methods for curve resolution are compared by applying them to profiles generated by computer combination of individual i.r. curves. Precision of resolution is sensitive to method, degree of curve overlap, and relative curve half-widths. The methods are used to determine thermodynamic parameters for the rotameric equilibria of five heterocyclic compounds.

A PROBLEM common to many branches of spectroscopy is the decomposition of a profile consisting of overlapped symmetric curves into its components or into quantities proportional to the areas of the components. In this study, part of a continuing interest in rotational isomerism and i.r. spectroscopy,[1,2] graphical, analogue computer using the Du Pont 310 curve resolver, and digital computer methods of curve separation are compared in an attempt to establish a reliable approach to the determination of thermodynamic parameters for rotameric equilibria. The conclusions are, however, applicable to other branches of spectroscopy.

The graphical methods[3] may broadly be classified as the wing method, the methods of the minimum and midpoint, and the method of outer halves. The wing method is the least precise and involves a subjective estimate of the contribution made by a small band to the wing of a large band situated some distance away. The methods of the minimum and mid-point involve the dropping of a perpendicular either from the minimum between, or from a point mid-way between, the maxima of two overlapped curves, and integrating the areas so obtained. The method of outer halves involves taking the leading half of the higher band and the trailing half of the lower as being proportional to the respective band areas. Peak heights may also be used as quantities proportional to band areas.

In the analogue computer method, bands consisting of 26 straight-line segments (the slopes and lengths of 24 of which were variable) were summed on an oscilloscope. A semi-silvered mirror system allowed superimposition of theoretical and experimental contours. The shapes, heights, and positions of the bands were individually adjusted until the fit of the synthesised to the experimental contour was optimised; the areas of the component bands relative to an arbitrary standard were read off from an integrator meter. In the digital computer method the equations and programs described by Pitha and Jones[4] were used, the problem being to apply least-squares analysis to non-linear functions using iteration to achieve acceptable fits. For i.r. spectro-

† Listings and cards of all of the programs referred to in the text are available on application to the authors.

[1] D. J. Chadwick and G. D. Meakins, *Chem. Comm.*, 1970, 637; D. J. Chadwick, J. Chambers, G. D. Meakins, and R. L. Snowden, *ibid.*, 1971, 624, 625; *J.C.S. Chem. Comm.*, 1972, 742.

[2] D. J. Chadwick, J. Chambers, G. D. Meakins, and R. L. Snowden, *J.C.S. Perkin II*, 1972, 1959.

scopy, the optimisation program uses the Voigt convolution (i)[5] in transmittance units where $(T/T_0)_\nu$ is the computed ordinate at wavenumber ν in transmittance units, x_1 is the peak height in absorbance units, x_2 is the

$$(T/T_0)_\nu = \exp{-2 \cdot 30258}$$
$$\left\{ \alpha + \sum_{P=1}^{M} x_{1(P)}[1 + x_3{}^2{}_{(P)}(\nu - x_{2(P)})^2]^{-1} \right.$$
$$\left. \exp\left[-x_4{}^2{}_{(P)}(\nu - x_2{}^2{}_{(P)})\right] \right\} \quad \text{(i)}$$

position in wavenumbers of the band centre, x_3 and x_4 jointly define the half band width and Lorentzian : Gaussian ratio, α is a baseline displacement in absorbance units, and M is the number of component bands. The values of $(x_1-x_4)_P$ and α were varied to minimise $\Sigma[(T/T_0)_{\nu,\text{obs}} - (T/T_0)_{\nu,\text{calc}}]^2$.

The methods of curve resolution are (1) analogue, (2) digital, (3) outer halves, (4) minimum, (5) mid-point, and (6) peak heights, and were compared first by overlapping (using an I.C.L. 1906A computer †) two i.r. bands (the carbonyl absorptions of methyl *m*-chlorobenzoate and methyl 3-methyl-2-furyl ketone at 30°) known, by measurement of half-widths, to be symmetric. [The bands were recorded separately on dilute solutions with carbon tetrachloride as solvent using a Perkin-Elmer 521 grating spectrophotometer (purged continuously with dry air) fitted with a Stogate data logger: this provided digitised spectra in binary-coded decimal, on eight-track paper tape, in units of transmittance $\times 10,000$; the input voltage to the logger was derived from a read-out potentiometer clamped on to the transmitting potentiometer of the spectrophotometer.] Computer-generated drawings of profiles from seven different degrees of overlap are in Figure 1. A paper tape record of the profiles suitable for subsequent digital curve resolution was produced simultaneously. Artificial profiles were thus generated of which the true curve resolution was known (area of right hand band A_r/area of left hand band $A_l = 0.8$). The profiles were then decomposed using the six methods (Table 1). Integrations were by Simpson's rule (using a computer), the error term in Simpson's equation being insignificant for the increments used here.

[3] A. W. Baker and M. D. Yeaman, *Spectrochim Acta*, 1966, 22, 1773.

[4] J. Pitha and R. N. Jones, *Canad. J. Chem.*, 1966, 44, 3031.

[5] R. N. Jones and K. Seshadri, *Spectrochim. Acta*, 1963, 19, 1013.

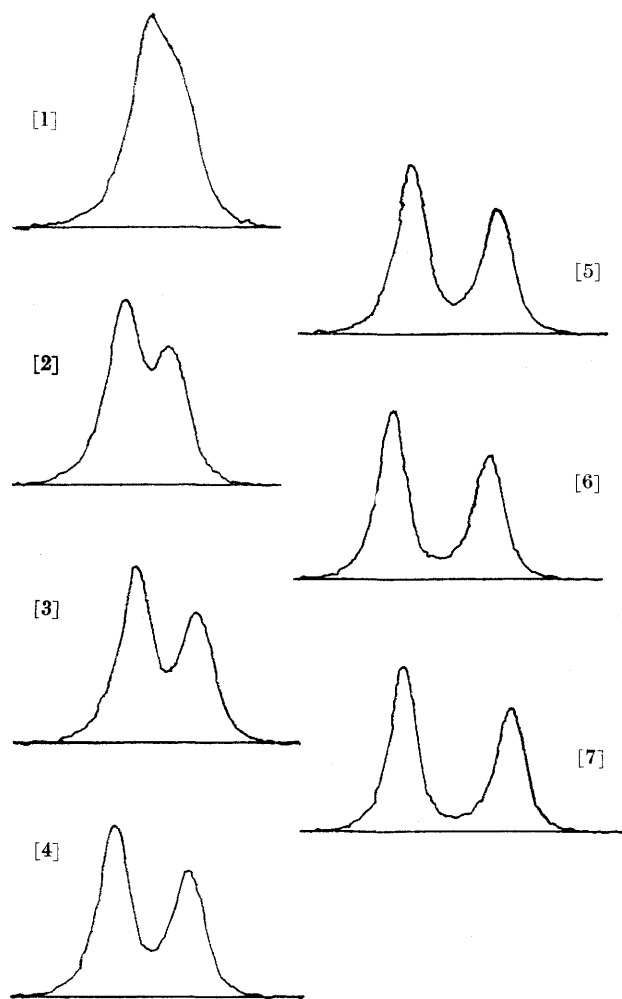At minimal overlap, all methods give similar results. However, as the degree of overlap increases, the accuracy

FIGURE 1 Profiles from seven different degrees of overlap with $A_r/A_1 = 0.8$

of the methods varies considerably. For overlap number $\leqslant$[3], the digital method (2) leads to considerable errors, whereas the analogue method (1) leads to errors even at overlap [4]. Of the graphical methods, the method of outer halves (3) yields variable results and the

TABLE 1

Application of methods (1)—(6) to decomposition of an artificial profile with $A_r/A_1 = 0.8$

	Method (A_r/A_1)					
Overlap	(1)	(2)	(3)	(4)	(5)	(6)
[1]	0·57	0·40	Methods not applied			
[2]	0·59	0·56	0·72	0·62	0·76	0·75
[3]	0·60	0·65	0·64	0·74	0·73	0·74
[4]	0·66	0·70	0·65	0·71	0·74	0·74
[5]	0·75	0·72	0·77	0·76	0·76	0·74
[6]	0·72	0·74	0·66	0·75	0·76	0·74
[7]	0·74	0·75	0·76	0·79	0·77	0·75

method of the minimum (4) deteriorates at overlap [2]. The methods of mid-point (5) and peak heights (6) give

consistently good estimates; however, in most methods there is a tendency to underestimate the true value.

In order to determine whether these conclusions were more generally applicable, the overlap experiment was repeated but with the ' left hand ' band (that of methyl 3-methyl-2-furyl ketone) halved in width. Computer-generated drawings of the profiles from six different degrees of overlap are in Figure 2. The profiles were decomposed as before (Table 2), the true curve resolution being $A_1/A_r = 0.63$. The situation at minimal overlap [6] is now less satisfactory, with analogue (1) and mid-point (5) methods yielding unacceptably high values: because of the differing half-widths of the component curves, the method of peak heights (6) is highly inaccurate. Of the graphical methods, the method of the minimum (4) (as in Table 1) seems the most satisfactory,
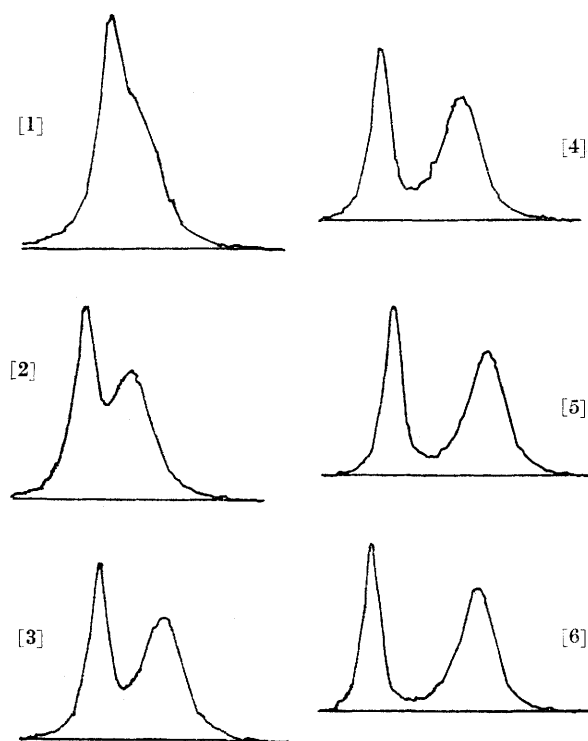

FIGURE 2 Profiles from six different degrees of overlap with $A_1/A_r = 0.63$

TABLE 2

Application of methods (1)—(6) to the decomposition of a profile with $A_1/A_r = 0.63$

	Method (A_1/A_r)					
Overlap	(1)	(2)	(3)	(4)	(5)	(6)
[1]	1·67	1·92	Methods not applied			
[2]	0·94	0·96	1·00	1·03	1·08	1·50
[3]	0·85	0·81	0·86	0·84	0·93	1·44
[4]	0·82	0·75	0·80	0·75	0·83	1·39
[5]	0·77	0·70	0·77	0·71	0·77	1·37
[6]	0·74	0·67	0·71	0·66	0·73	1·35

although overall the digital computer method is the most accurate. However, for overlap numbers $\leqslant$[4], gross errors result from all methods.

A further comparison of the methods was made by overlapping, to varying extents, two symmetric Voigt functions derived from a digital computer curve resolution of the i.r. carbonyl bands of 17-oxo-5α-androstan-3β-yl thiophen-2-carboxylate. Computer-generated drawings of profiles from five different degrees of overlap are in Figure 3. The profiles were decomposed as before (Table 3), the true resolution being $A_r/A_1 = 0.76$. Most methods give similar results at minimal overlap [5] although, because the component curves again have very different half-widths, the method of peak heights (6) is

TABLE 3

Application of methods (1)—(6) to decomposition of a profile derived from overlap of Voigt functions fitted to the i.r. carbonyl profile of 17-oxo-5α-androstan-3β-yl thiophen-2-carboxylate with $A_r/A_1 = 0.76$

Overlap	(1)	(2)	(3)	(4)	(5)	(6)
[1]	0·68	0·76	Methods not applied			
[2]	0·71	0·76	0·77	0·90	0·92	1·36
[3]	0·71	0·76	0·76	0·80	0·85	1·31
[4]	0·69	0·76	0·76	0·77	0·80	1·29
[5]	0·74	0·76	0·75	0·76	0·78	1·29

inaccurate. As the degree of overlap increases, the graphical methods again deteriorate; the computer methods do not, however, show a marked deterioration
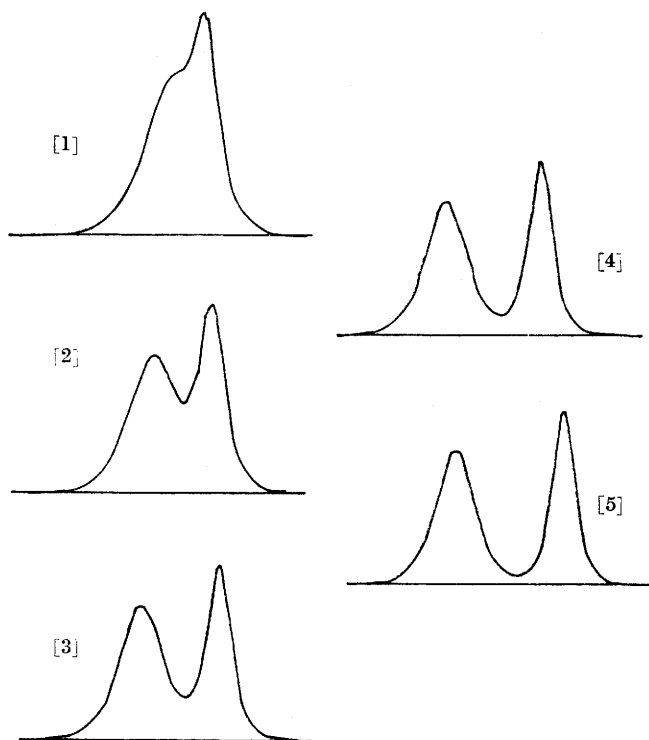

FIGURE 3 Profiles from five different degrees of overlap $A_r/A_1 = 0.76$

and even at gross overlap [1] the digital method yields an accurate resolution: thus, the decomposition of artificially overlapped functions is not a rigorous test

for computer-based curve resolution methods (cf. columns 1 and 2 in Tables 1 and 3).

Thus, before attempting a curve resolution, the spectroscopist should test the applicability of the chosen method by an overlap experiment, since the accuracy of the methods depends critically upon the degree of overlap and upon the relative half-widths of the overlapping curves.

In order to demonstrate how the choice of curve resolution method may affect results derived therefrom, the six methods were applied to each of five compounds
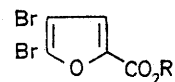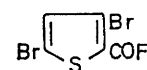

(I) R=Me
(II) R=Et
(III) R=Bu^t
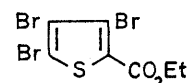
(IV)

(V)

[methyl (I), ethyl (II), and t-butyl (III) 4,5-dibromofuran-2-carboxylates, ethyl 3,4,5-tribromothiophen-2-carboxylate (IV), and 3,5-dibromothiophen-2-carbonyl

TABLE 4 [a]

Thermodynamic quantities derived for equilibrium (ii)

Compound	Method	$\Delta H^0/$ J mol⁻¹	$\Delta G^0_{303}/$ J mol⁻¹	$\Delta S^0/$ J mol⁻¹ K⁻¹
(I)	(1)	−1400	−750	−2·0
	(2)	−1400	−1550	0·0
	(3)	700	−550	4·5
	(4)	−900	−500	−1·0
	(5)	−100	−450	1·5
	(6)	−100	−250	0·5
(II)	(1)	−1500	−1150	−1·0
	(2)	−200	−2050	6·0
	(3)	−1100	−950	−0·5
	(4)	−900	−800	−0·5
	(5)	−600	−700	0·5
	(6)	−700	−250	−1·5
(III)	(1)	−3100	−1000	−7·0
	(2)	−2600	−1750	−3·0
	(3)	−2800	−900	−6·5
	(4)	−2600	−500	−7·0
	(5)	−2300	−600	−5·5
	(6)	−2700	−400	−8·0
(IV)	(1)	1500	1450	0·5
	(2)	400	1300	−4·0
	(3)	800	1550	−3·0
	(4)	900	1350	−2·0
	(5)	1000	1350	−1·5
	(6)	1000	1700	−2·5
(V)	(1)	1900	1450	−2·0
	(2)	2800	2700	1·0
	(3)	800	2100	−4·0
	(4)	2500	2200	1·0
	(5)	2100	2000	1·0
	(6)	2500	1700	3·0

[a] Estimated errors are ± 950 J mol⁻¹ for ΔH^0, ± 250 J mol⁻¹ for ΔG^0, and ± 4.0 J mol⁻¹ K⁻¹ for ΔS^0, assuming that the errors in the equilibrium constants are random and Gaussian.

fluoride (V) *] all of which exhibit doublet i.r. carbonyl absorptions (overlapped to a degree of [3] or [4], Table 1) resulting from an equilibrium between *syn-* and *anti-rotational* isomers. The variable temperature i.r. data were processed using the method and equations described previously,[2] the ratio of integrated absorption coefficients for *syn-* and *anti-*forms being taken as unity. Thermodynamic parameters so derived for the equilibrium (ii) are in Table 4. Although the absolute spread

form with higher carbonyl frequency $\rightleftharpoons$
form with lower carbonyl frequency   (ii)
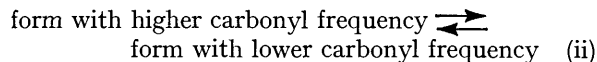
of values is large, when the limits of error are included, the spread of ΔH^0 values is just satisfactory. The

* Syntheses of compounds (I)—(III) have been reported,[6] those of compounds (IV) and (V) will be shortly.

$\Delta G_{303}{}^0$ values are less satisfactory, possibly because they are derived from individual area ratios rather than *via* a least squares fit to a range of area ratios as is the case with the ΔH^0 values.

Clearly, in a spectroscopic technique involving the decomposition of profiles arising from overlaps more serious than the rather favourable cases discussed here, it must be demonstrated that the quantities derived from such decompositions have any significance.

[6] D. J. Chadwick, J. Chambers, G. D. Meakins, and R. L. Snowden, *J.C.S. Perkin I*, 1973, 1766.