Journal of Chromatography, 542 (1991) 247–261
Elsevier Science Publishers B.V., Amsterdam

CHROM. 23 007

Computer spectrochromatography

Principles and practice of multi-channel chromatographic data processing

Yu. A. KALAMBET* and Yu. P. KOZMIN
Scientific and Technical Cooperative "Ampersand", P.O. Box 439, Moscow 123060 (U.S.S.R.)
and
M. P. PERELROYSEN
Novosibirsk Institute of Bioorganic Chemistry, Siberian Branch of the Academy of Sciences of the U.S.S.R., Pr. Lavrientieva 8, Novosibirsk 630090 (U.S.S.R.)

(First received June 6th, 1990; revised manuscript received November 23rd, 1990)

ABSTRACT

A new approach to multi-channel chromatographic data processing is introduced, based on the treatment of a chromatogram as a curve in multi-dimensional space. Each coordinate in this space is the signal-to-noise ratio for one detector. Every point in the space represents one spectrum-set of detector responses. Multi-dimensional space mathematics applied to the analysis of the chromatogram allows one to obtain individual substance detection profiles for overlapped peaks in the case of known spectra of each component, analyse peaks for homogeneity and determine the number of substances in overlapped peaks and their elution profiles without a prior knowledge of their spectra. All these tasks are solved with maximum accuracy owing to the principles of space coordinates construction. Theoretical considerations are illustrated by the results obtained for multi-wavelength chromatograms measured with a Milichrom chromatograph.

INTRODUCTION

The development of a wide range of ultraviolet (UV) spectral detection devices for chromatography probably started with the OB-4 chromatograph designed in the late 1970s in Novosibirsk [1]. This device is equipped with a rapid scanning UV detector with a forward optics double-beam measurement scheme. It has a 15-bit measurement accuracy and covers the absorbance range up to 12.8 (1.56 mm optical path cuvette). It was also the first microbore system with a column I.D. of 2 mm and a syringe pump volume of 2.5 ml.

We have been developing software for this device since 1983 and gained some experience in multi-wavelength data processing (*e.g.*, ref. 2) that may be useful when applied to other multi-wavelength or multi-channel chromatographic systems.

In particular, we succeeded in the analysis of unresolved peaks by spectral

0021-9673/91/$03.50 © 1991 Elsevier Science Publishers B.V.

criteria. Two types of such an analysis are distinguished. The first is applied to peak analysis in the case when the spectra of all components are known. In this instance a least-squares method allows individual elution profiles and good estimates of the amounts of each component to be obtained. The second type helps in analysis when none of the component spectra are known in advance. In this instance multiple regression and factor analysis methods allows estimates of substance spectra, elution profiles and amounts to be obtained, but the results may be less accurate than in the first type of analysis.

THEORY

We use the term "spectrochromatography" for all kinds of chromatographic systems where the result of measurement is a set of values instead of one detector response as in traditional chromatography. This may be UV–VIS spectrum measurement (SM) systems, mass spectrometer (MS) systems or just a set of different detectors measuring different characteristics of the single gas or liquid chromatographic flow. From the very beginning of problem formulation we stress that different detector outputs may be mixed, thus changing the view from uniform chromatographic systems where all of the channels have the same physical units of measurement.

Detector response space

Apart from the three-dimensional view of a chromatogram offered by many photodiode-array detection software systems, we use a different approach. We consider a spectrochromatogram to be a curve in multi-dimensional space, where each coordinate represents an individual detector response, *e.g.*, the response of one diode of an array (SM) or one mass (MS). One point in the detector response (DR) space represents one spectrum. In this DR space the chromatogram looks like a curve where successive measurements are connected by the line. In the case of a "spectrum" consisting of only two wavelengths, this will be a curve on the plane where one axis represents the absorbance for one wavelength and the other axis the absorbance at the other wavelength. If we subtract the baseline in such a chromatogram, the curve will become tightly grouped around zero with peaks looking like curve fragments starting from zero and returning back to zero. The analysis of curves in this space is not convenient, as the value of a signal given by each coordinate depends on, *e.g.*, the units of measurements.

Noise-normalized detector response space

The main reason why curve analysis in DR space is complicated is that many curve analysis techniques use the root-mean-square (RMS) approach. This approach minimizes the sum of squares of distances for each coordinate. Therefore, if we wish to use this approach, the expected errors should be comparable for all axes, otherwise the result will not be accurate. Even more, DR space coordinates may have different physical units of measurement and we may add square absorbance units to square millivolts.

There is one natural way to find unified space coordinates suitable for all kinds of curve analysis, *viz.*, we can use units of measurement equal to the noise (or the expected measurement error). In this instance each coordinate axis will represent the

signal-to-noise ratio (S/N) for one channel. Let us call this space a noise-normalized detector response (NNDR) space. NNDR and DR spaces may be transformed into one another by linear transformation:

$$D = RW \tag{1}$$

where D = normalized spectrum from NNDR space (detection vector), W = weight matrix with $W_{ij} = 0$ $(i \neq j)$ and $W_{ii} = 1/E_i$, E_i = expected error for the channel i, $R = (R_1, \ldots, R_N)$ = detector response vector and N = number of channels (detectors).

NNDR space is very useful, as with its help one can analyse in a unified way physically different signals, *e.g.*, conductivity and radioactivity. One point in the NNDR space represents the normalized spectrum.

Hence NNDR space is just such a space where one coordinate represents one channel and expected errors for all coordinates are equal.

Total S/N ratio profile

The NNDR space chromatogram presentation contains additional information that can be easily extracted and successfully used.

Before the baseline subtraction, we need to find the peak beginning and end. However, in a multi-detector system it is very difficult to use any one channel for this purpose. An effective and consistent way is to use a "summary report" of a chromatogram, *i.e.*, the profile of the total S/N for all detectors (see Fig. 1). The S/N profile gives a good visual representation of the chromatographic curve end eliminates the question of "which channel is better".

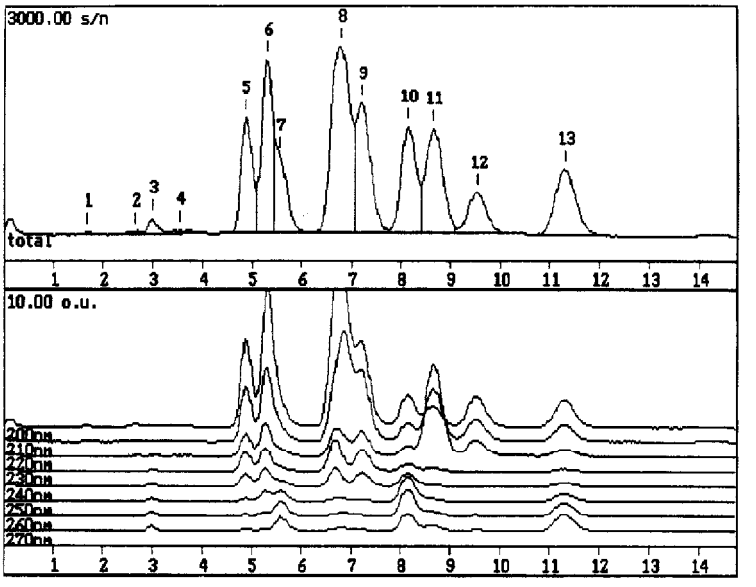

Fig. 1. Multi-channel chromatogram (bottom) and signal-to-noise ratio profile for this chromatogram (top). Abscissa in microlitres; ordinates are different and labelled at the top left corners. Output of Chrom&Spec software.

Peak shape in NNDR space

Hereafter, unless specially noted, we shall consider a chromatogram with a subtracted baseline, *i.e.*, each detector response will be substituted with the difference between the actual detector response value and the estimated baseline value at that point. The method of baseline estimation implemented in our program is described under Discussion.

In the case of a homogeneous peak and fulfillment of a detector response proportionality law (the detector response is proportional to substance concentration), all points belonging to the same peak should be positioned along the straight line with an average distance from the line close to 1:

$$D = cQ + e \tag{2}$$

where D = detection vector, c = concentration of substance, $Q = SW$ = substance spectrum S transformed into NNDR space and e = normalized error of measurement. Eqn. 2 is valid for all peak spectra with different concentration c and error vector e values.

A multi-component peak in NNDR space will be a linear combination of several detection vectors, one for each pure component:

$$D = c_1Q_1 + c_2Q_2 + \ldots + e \tag{3}$$

where c_k = concentration of the kth component and

$$Q_k = S_kW \tag{4}$$

is transformed to the NNDR space kth pure component spectrum S_k. As before, eqn. 3 can be written for every point within a peak.

Substance quantification

In the case when pure spectra for all substances within an overlapped peak are known in advance, one can use the RMS approach to decompose every measured spectrum of the peak into spectra of this pre-defined basis. In the case of known spectra, eqn. 3 represents a set of N equations with K unknowns, where K is the number of individual components and N, as before, is the number of channels. In the case when $K \leqslant N$ we can apply the RMS approach (see, *e.g.*, ref. 3) which will give solution of eqn. 3 in the form

$$C = DQ^T(QQ^T)^{-1} \tag{5}$$

where $C = (c_1, c_2, \ldots, c_K)$ is a concentration vector, D = detection vector and $Q = (NK)$ matrix composed of K detection vectors made from pure component spectra by eqn. 4.

Of course, the spectra of pure components should be linearly independent (not too similar), otherwise matrix inversion in eqn. 5 cannot be applied. If these conditions are fulfilled, decomposition is allowed and gives a unique concentration vector C.

By making a spectrum decomposition for every peak point, we can obtain

a concentration profile for each component. To present results it is usually more convenient to use an absorbance profile instead of a concentration profile. The absorbance profile may be reconstructed as

$$A_k(t) = c_k(t)S_k \tag{6}$$

where $A_k(t)$ = evaluated absorption (DR space!) vector of the kth component, $c_k(t)$ = concentration profile of the kth component obtained by eqn. 5 and S_k = kth component spectrum. Eqn. 6 should be evaluated at every peak point. To stress this fact, we introduced a time dependence into eqn. 6. Note that no peak shape considerations were used!

In the case when $K = N$, eqn. 5 will give an exact solution of N equations with N unknowns instead of the RMS approach, and this exact solution can be found for *any* N linearly independent basis spectra, no matter whether correct ones or taken as an arbitrary choice! To obtain reliable results the number of channels should be at least twice the number of detected substances, so that the residual error could help in the validation of results.

Use of NNDR space angle

To help in visual NNDR space chromatogram analysis, we can use angles between time-adjacent vectors in NNDR space, *i.e.*, we can associate additional characteristics with each peak point: an angle between NNDR vectors for that and the preceding point. These angles can be plotted as an angle profile of the peak. In the case of a homogeneous peak, NNDR space points are located along the line and angles inside the peak should be close to zero. At the very beginning and at the end of the peak the signal value is low, and the predicted angles are large in comparison with those inside the peak. Non-homogeneous peaks will show more complex structures (see Figs. 5–7).

In some instances we can also use the NNDR space inter-vector angle as a measure in comparisons of spectra. Comparisons of spectra should be performed taking into account the measurement error for each channel. Hence we can use the NNDR space angle as a measure of spectral similarity in the case of a spectral library search and other cases of spectral comparisons.

Many useful criteria may be introduced using angles. For example, in the case of spectrum decomposition for two known spectra, the angle between basis vectors should be much greater than the expected inter-vector angle caused by noise.

One argument for or against correct basic spectra selection is the residual error of spectrum decomposition. The problem is that the decomposed spectrum may differ from the sum of basis spectra with coefficients obtained by the RMS approach and an angle between initial and resynthesized vectors may be calculated. To obtain a numerical value of the angle we can use the tangent, *i.e.*, the ratio of the length of the residual vector to the length of initial vector (see Fig. 4, residual error value).

We shall mention one of the reasons why the angle is so attractive and why we consider it as an alternative to the correlation coefficient in spectra comparison: the angle is a natural measure, easily understandable by most people, and psychologically it is much easier to imagine an angle than a correlation coefficient.

Further chromatogram analysis

As a chromatogram is considered as a curve in a multi-dimensional space, we can try to analyse a curve without any *a priori* knowledge such as substance spectra. To do so we have to apply multi-dimensional space mathematics, *i.e.*, principal component and factor analysis. The theory and a detailed description of the algorithms for this type of analysis will be published elsewhere; here we shall only present an overview of the calculations made by our program. At this stage of curve analysis we apply the following.

(a) We obtain a subspace with a minimum number of dimensions that contain the entire peak curve within admissible error by principal component analysis. The number of subspace dimensions will represent the number of components in the peak and pure spectra should be linear combinations of subspace basis vectors (eigenvectors).

(b) We filter noise by placing the whole curve for the peak into subspace.

(c) We find the best candidates for the role of "pure spectrum" for each component among all of the vectors in the peak. That is, we assume that there exists some point within the peak where the first component is "pure" and assume the spectrum at this point to be a true pure first-component spectrum, and so on.

(d) We obtain elution profiles of each component within the peak and their amounts.

To apply this type of curve analysis, one should note the propositions made: (a) eluted substances have linearly independent (different) spectra; and (b) the elution profiles are shifted with respect to each other so that for each component there is a point within an analysed chromatogram region where only this component is present. The spectrum at this point will be taken as a pure spectrum of this component.

EXPERIMENTAL

Equipment and software

All measurements were performed using a Milichrom chromatograph (Nautchpribor, Orel, U.S.S.R.), which is an industrial version of the OB-4 chromatograph [1].

Two software systems were used. The first was developed at the Institute of Molecular Genetics of the Academy of Sciences of the U.S.S.R. for the ISKRA-226 computer (U.S.S.R.) with a special interface unit. The system was written in BASIC and Assembly languages. The second chromatographic data processor, Chrom&Spec, is a product of the "Ampersand" Cooperative (Moscow, U.S.S.R.) for IBM PC-compatible computers. With the Milichrom chromatograph data acquisition is performed by a special interface board, otherwise several analog-to-digital converting (ADC) devices are supported. Data import from LKB WaveScan primary data files and ASCII text files is also possible. The system is written in C and Assembly languages.

RESULTS

Peak analysis in the case of known pure spectra

One of the applications of this possibility was the evaluation of the amount of protein (RNA polymerase) crosslinked to DNA [2]. In this case the chromatographic column was used for the partial separation of the DNA–protein complex from

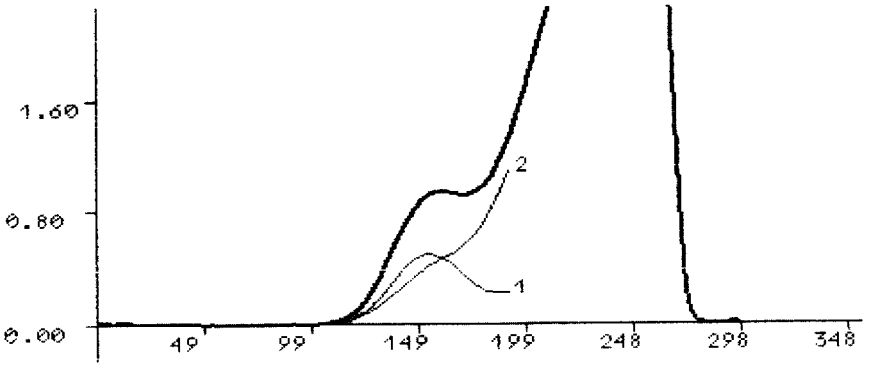

Fig. 2. Gel chromatogram of DNA–protein complex monitored at 230 nm and profiles of individual component elution: (1) DNA; (2) protein. Abscissa shows volume in microlitres, ordinate the absorption in o.u. Individual elution profiles were evaluated by eqn. 6. Output of ISKRA-226 software.

unbound protein (Fig. 2) by gel chromatography. We analysed the first part of the peak corresponding to the complex. The results of the spectrum decomposition are shown in Fig. 3, and associated numerical information in Fig. 4. Individual substance spectra were measured in advance and for the purpose of analysis they were loaded from disk files.

Factor analysis: what is visible to the user?

We have tried to provide a simple and convenient user interface in our software. We shall therefore try to illustrate how peak factor analysis is implemented to show that even an inexperienced user may extract some new information from the chromatogram.

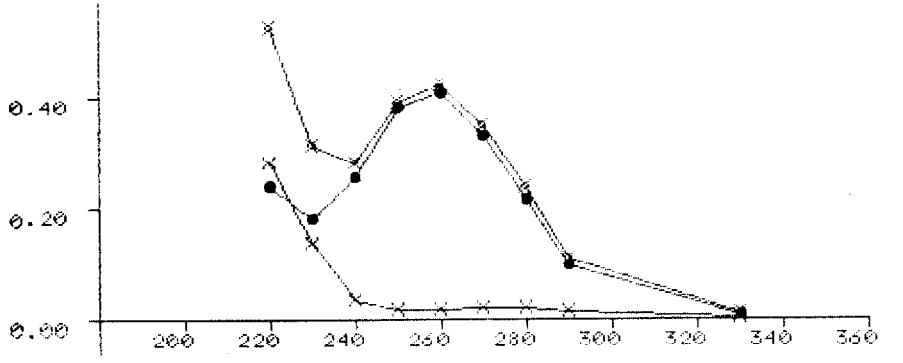

Fig. 3. Illustration of spectrum decomposition. Abscissa shows wavelength in nanometres, ordinate the absorption. The higher curve shows the chromatogram region spectrum and the lower curves (●) DNA and (×) protein library spectra. Associated textual program output is shown in Fig. 4. Output of ISKRA-226 software.

Chromatogram:
Date 18.05.85; 9 wavelengths; flow 10 mcl/min; User Burova;
Sample: pAO3 plasmid DNA: RNA polymerase complex;
Column: TSK-gel HW65, volume 300 mcl;
Eluent: 10mM TEA, 10 mM MgCl2, 100 mM KCl in water;
Wavelengths: 220 230 240 250 260 270 280 290 330 nm;

Section: volume 110-145 mcl.
Volume 35.8 mcl; Integral $8.29*10^{-4}$ o.u.*ml;
Mean absorption at 260 nm 0.41 o.u.

a) Spectrum:

Wavelength nm	Weight	Ratio
220	0.31	1.23
230	0.43	0.78
240	0.53	0.70
250	0.42	0.93
260	0.60	1.00
270	0.49	0.83
280	0.49	0.57
290	0.46	0.28
330	0.51	0.04

b) Spectral decomposition:

Substance name	Signal	Concentration mcg/ml	Concentration M/l	Amount mcg	Molecular mass Dalton
pAO3 DNA	79.1%	20.6	$1.96*10^{-8}$	0.74	1050000
RNA polymerase	19.4%	31.7	$7.05*10^{-8}$	1.13	450000

Residual RMS error 3.1%

Fig. 4. Spectral analysis of a chromatogram region. Weight column in the spectrum printout (a) stands for the reverse of the baseline noise. The weight equals 1 if the baseline noise is absent and hence the expected error equals the discretization error, 0.0004 o.u. in the case of the Milichrom device. Hence in the above example the baseline noise equals approximately 3 discretes for 220 nm (where discrete stands for the smallest voltage change that can be measured by the device) and is less than 2 discretes for 260 nm. The spectrum is normalized for the response at one wavelength (260 nm in this instance), called the reference wavelength. Spectrum decomposition is made by eqn. 5 with weights for DR-NNDR space transformation listed in Table (a). Molar ratio of DNA to protein amounts shows that 3–4 RNA polymerase molecules are bound to DNA, which was confirmed by electron microscopy [2]. Output of "ISKRA-226" software.

The option that performs peak factor analysis is called "Analyse". Before its activation it is necessary to mark a region of interest with the help of a vertical bar

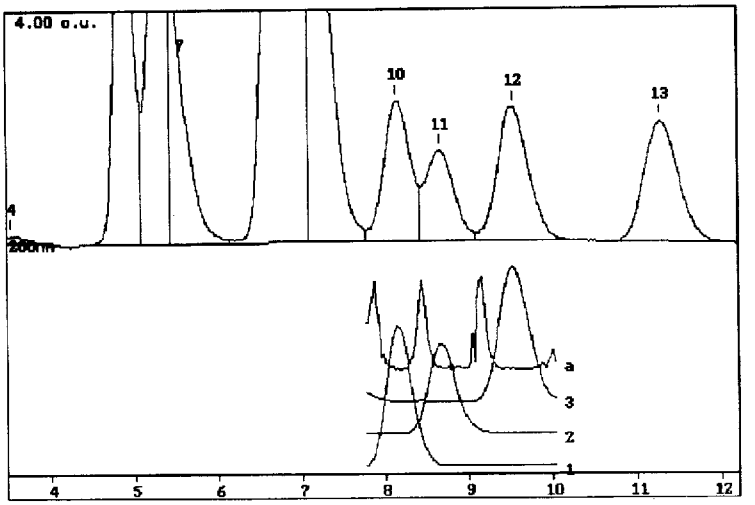

Fig. 5. Factor analysis of a chromatogram region. Here and in Figs. 6 and 7 the abscissa is in minutes, and the ordinate in absorbance units with full-scale labelled at the top left corner. Chromatogram at 200 nm and individual elution profiles for each component (curves 1, 2 and 3) are shown. Angle curve (a) is shown in arbitrary ordinate units. Output of Chrom&Spec software.

cursor. After the activation of the option the program performs some calculations and shows sets of numbers (NNDR space covariation matrix eigenvalues) to the user (an experienced user can use this set of numbers to select a number of substances in the region). Subsequently the program asks for the number of components and offers some value. The user can accept this value or enter his or her own. For the case of Milichrom data usually there is no need to change the value.

As the value is entered or confirmed, the program asks for the "pure spectrum"

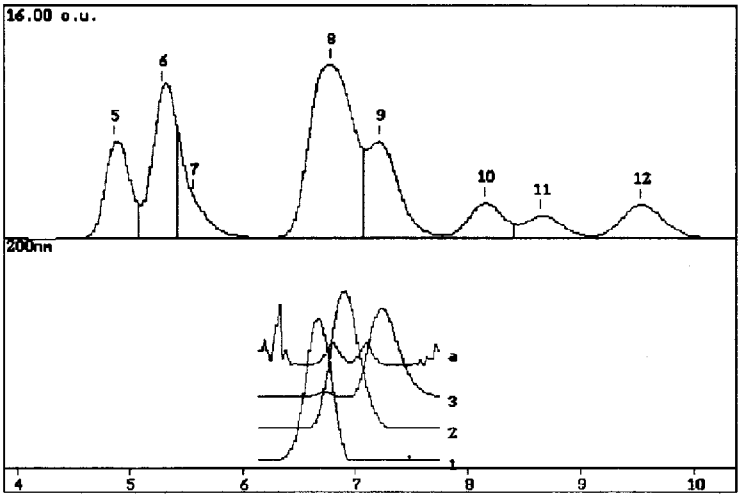

Fig. 6. Factor analysis of a chromatogram region. All notations as in Fig. 5.

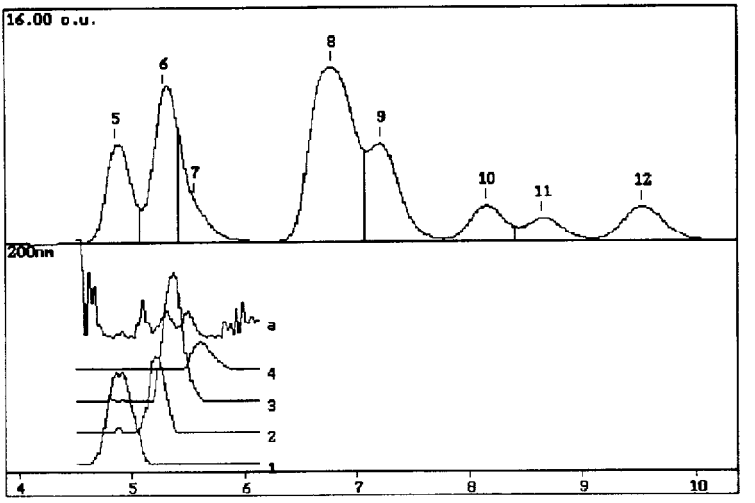

Fig. 7. Factor analysis of a chromatogram region. All notations as in Fig. 5.

selection mode (automatic or manual). In the case of the manual "pure spectrum" selection mode, one has to set the cursor to the position of the first "pure component" and press the Enter key and so on for every "pure component" position. In the case of the automatic mode the program selects these positions itself. Our experience shows that the spectrum selection algorithm works quite well and even if the user sees an angle function for the peak on the screen he or she usually cannot select better "pure spectrum" positions than those selected by the program.

After the end of the spectrum selection dialogue the program shows elution profiles for each component for the first wavelength (Figs. 5–7). Then it asks for the next wavelength to show. The sum of elution profiles is approximately equal to the peak profile for the channel shown on the screen (with subtracted baseline, of course). At this moment the scale factor can be changed so that peaks with different signal levels can be investigated and more details examined.

If the user presses the Escape key to quite the elution profile display loop, the program lists (in numerical form) each substance spectrum normalized for the reference wavelength (selected previously by the user) and the area for the peak normalized for the total area of the selected region (Fig. 8).

The final question is whether it is necessary to adjust peaks boundaries in accordance with the results of factor analysis. If the user replies positively, peak boundaries are shifted so that the number of peaks in the region equals the number of pure components and the resulting area of each peak at the reference wavelength is the same as that calculated by factor analysis. After this procedure, area estimates obtained by the "Analyse" option are permanently stored in the peak pattern.

```
Chromatogram: test1-i.ODS
Data: 31/07/1989 13:05:10
Operator: Nagaev I.
File: 97311306.CHM
Flow rate: 50.0 mcl/min
Column: LiChrospher C18, 5 um
Eluent: 80% MeOH: 20% water
Sample: mixture of methyl esters of benzoic acid, benzene,
        toluene, etc.
8 channel(s)
Reference channel: 200nm
```

No	200nm	210nm	220nm	230nm	240nm	250nm	260nm	270nm	Q-ty	Conc.%
1	1.000	0.560	0.347	0.250	0.417	0.794	0.870	0.564	26.03	34.83
2	1.000	2.527	4.568	0.162	0.078	0.149	0.239	0.285	18.03	24.12
3	1.000	0.720	0.544	0.015	0.004	0.014	0.037	0.046	30.68	41.05

Fig. 8. Table produced as a result of peak factor analysis. It includes estimated spectrum for each component (ratio of absorption at a given wavelength to absorption at the reference wavelength), amount (amount units are the same as quantity (area in arbitrary units) and not shown here) and relative amount. The relative amount is calculated as an area percentage with respect to the total area of the region for the reference wavelength.

Results of peak deconvolution

In some instances, as for peaks 10, 11 and 12, a regional structure is evident and the user gains only a better quantitative estimate of each peak area, as can be seen from Figs. 5 and 8. Curves 1, 2 and 3 show elution profiles for each component at a wavelength of 200 nm, curve a–angle profile. The angle profile shows three distinct minima that correspond to fairly pure component regions where one component is changed by the other and the spectra at adjacent points differ significantly or at the peak boundaries, where S/N is low.

In other instances, as with peaks 8 and 9, a regional structure is not so evident and only an experienced user can "find" more than one peak in peak 8. Evidence for the complex peak structure is the slightly shifted position of the peak 8 maximum at different wavelengths and the peak shape at a wavelength of 210 nm (Fig. 1). In fact, when we start to analyse this region we find three separate peaks as shown at Fig. 6.

The most impressive example of peak analysis is the group of peaks 5–7. Three peaks can be distinguished in this region and operator reflected this when marking the peak boundaries in the manual mode. Besides, factor analysis of the region detects four substances with different spectra. Their estimated elution profiles are shown in Fig. 7, curves 1–4. In fact, if we consider the angle curve [see Fig. 7, curve (a)], two distinct local minima within peak 6 are found. These minima show that the peak structure is complex, and at least two components are present. Of course, as these minima do not reach zero, there is no point where an appropriate substance is pure and the spectrum at the minimum of the angle function reflects the real substance spectrum with a significant error, but nevertheless it can be used to prove the presence of extra

components and to estimate the elution profile and the amount of the component by the order of magnitude.

DISCUSSION

None of the methods described under Theory is new and they have been applied to multi-channel chromatographic data previously [4–10]. What is new in our approach is that we introduced normalization of data for the expected error prior to data analysis (NNDR space), which allowed us (a) to perform all calculations with maximum accuracy, (b) to introduce and use the total signal-to-noise profile and (c) to introduce and use the angle profile of the chromatogram.

Although the methods used have long been known, they have not received the attention they deserve, and we hope that in the future their contribution will increase. We shall now discuss some of the advantages the user may gain from a wider application of these methods.

User gain

The user of a multi-channel chromatographic system gains the following from spectral detection: first, an estimate of peak purity; if the peak fits along a straight line within admissible error, then the peak can be considered pure; second, peak deconvolution for the known spectra; this task may be performed manually or in an automated mode depending on user needs; and third, analysis of a peak with unknown composition. These attributes lead to simpler method development, faster solvent optimization procedures, quantification of chromatographically unresolved peaks, additional information on the number of substances in the unknown peak, etc.

Another application direction is in the routine determination of amounts of components in repetitive chromatographic analyses of the same mixture. In this instance a simpler chromatographic system may be used, as the spectra are known in advance and spectrum decomposition partially replaces chromatographic separation. Individual elution profiles may also be useful in preparative chromatography. Many other areas of application are possible.

Sources of error

Usually no benefit can be gained without drawbacks. In this instance, the drawback is a potential error caused by improper use of the software. We shall outline here some of the problem areas.

An estimate of peak purity may be wrong because (a) spectral analysis cannot distinguish substances with identical spectra and (b) different errors of spectrum evaluation discussed below may be treated as spectral impurities.

Spectrum decomposition for the known spectra may be wrong because of (a) wrong basis spectra, (b) identical or almost identical basis spectra, (c) too many substances in a region and (d) errors in spectra evaluation.

Analysis of an unknown chromatographic peak may be wrong because (a) spectral analysis cannot distinguish substances with identical or very similar spectra, (b) there are too many substances in a region, (c) chromatographic separation of peaks is poor and (d) errors in spectra evaluation.

Estimate of expected error

There are several potential error sources in spectral measurement. Strictly we do not measure the spectrum of the substance that elutes within a peak. We try to evaluate this spectrum based on measured chromatographic data. This evaluation produces errors, and here we try to list and discuss some of them which are important from our point of view. We deal with evaluation and measurement errors only, ignoring other error sources, such as substance spectrum changes in different solvents.

To obtain NNDR space spectrum coordinates, one should know the expected error for each channel. At present we use high-frequency baseline noise instead of expected error throughout the program. Errors from other sources in many instances seem to be proportional to the baseline noise, so we consider normalization for the baseline noise to be the first greatest step in the direction of equal measurement errors by each channel. Further steps are possible, but they will not provide such an increase in accuracy as the first step did.

Spectrum correction for measurement timing

This is one type of error that in some instances can be precisely corrected. Its source is that measurements at different wavelengths of the same spectrum are made at different moments in time. The first wavelength response is measured first, and by the time of measurement of the last wavelength the absorption of the first may have changed significantly. This error source is important for PDA detectors but with the Milichrom device it is especially important because of the long minimum measurement time per wavelength (0.15 s). We solve this problem by data interpolation. That is, we substitute measured values for new ones, calculated for the time that corresponds to the middle of the spectrum measurement cycle. This is done once when the chromatogram finishes, as we use three-point quadratic interpolation and need previous and subsequent measurements at the same wavelength. We found that this type of interpolation gives adequate results and more complex interpolation methods are not required.

With multiple detectors in series, when measurements are separated not only by time but also in space, this error may be the greatest error source, because in addition to time adjustment difficulties the peaks change their shape from detector to detector.

Baseline subtraction

Correct baseline detection is one of the most important conditions of precise spectrum measurement. There is no sense in measuring an absolute absorbance value, as different solvent and gradient absorbance profiles may change the spectrum considerably. Therefore, absorbances should be measured with respect to the baseline and the spectrum should be constructed from these difference absorbances. In our system there are baseline points on the chromatogram and the baseline itself is linear under the peaks. Baseline points are considered to belong to the baseline on all channels/wavelengths. Hence the baseline for each channel is determined in the traditional way for chromatography, but the time positions of baseline points are common for all channels. To provide the best choice of baseline points we use for peak detection purposes our synthetic "total" channel (Fig. 1) by default.

The baseline subtraction error is theoretically always greater than or equal to the baseline noise error. To decrease its value, visual investigation of peak detection results

can be recommended and in some instances baseline noise filtration by a noise filtration algorithm may help. In all instances non-linear baseline changes cannot be accounted for and we consider this error as one of the greatest error sources.

Measurement error

From the very beginning we estimate the minimum measurement error. This error is equal to one discrete of signal conversion or baseline noise for the channel, whichever is greater. The expected error of the measured detector response at any point within a chromatogram cannot be less than this value. Also, with high absorbances another error source is important: the small amount of light passed through the cuvette increases the expected error owing to statistical fluctuations in the number of light quanta. The error increases to infinity in the case of signal overflow.

Because of different error sources, pure peak fits along the straight line not within baseline noise accuracy, but slightly worse. Most of the errors mentioned increase as the baseline noise increases and therefore the use of baseline noise for calculations at least makes the expected errors for each channel comparable and allows the RMS approach to be used adequately.

Spectral slit width and accuracy

The spectral slit width is a very important system characteristic: the wider the spectral slit, the lower is the noise. On the other hand, a change in spectral slit width will cause a change in the spectral ratio for most wavelengths and will require recalculation or measurement of another copy of a library spectrum. Further, a wide spectral slit gives undesirable effects. Thus, it is possible to show that with a wide rectangular spectral slit and a narrow absorption band the measured spectral ratio on the upslope and downslope of the absorption band will depend on the substance concentration, *i.e.*, the Lambert–Beer law will be violated. This effect is due to the fact that the physically measured value is not the absorbance, but the amount of light passed through the cuvette, and absorbance is obtained after non-linear conversion of that value. The error in absorption measurement caused by the spectral slit width in the case of a constant spectral intensity of the light source and a rectangular spectral slit may be evaluated as

$$A_{app} - A \approx (dA/dL)^2 w^2 l/12 \tag{7}$$

where A_{app} = measured (apparent) absorption, A = real absorption, l = optical path length, dA/dL = first derivative of absorption for the wavelength and w = spectral slit width. More discussion of this equation will be published elsewhere. Note that eqn. 7 shows that the measured spectrum depends on substance concentration.

The effect of the spectral slit width is more pronounced with visible light detection than with UV detection because of narrow bands and hence higher values of dA/dL (unpublished data). Thus a wider spectral slit may cause a change in the library spectrum and an additional risk of treatment of homogeneous peaks as non-homogeneous because of violation of the Lambert–Beer law.

To minimize the effect of the spectral slit width it is possible to select wavelengths for detection and calculations near the absorption band maxima and minima. Another

possibility is to use a narrow spectral slit width detector and a short optical path cuvette.

Rapid scanning vs. diode-array detectors

There are some additional considerations on different measurement schemes implemented in different SM detector types. A comparison of forward and reverse optics schemes shows that the high total S/N level obtained with PDA reverse optics, in addition to its advantages, also has some drawbacks: (a) lower S/N characteristics per channel and (b) very high light intensity inside the cuvette. The former effect makes our type of data analysis less informative but may be partially compensated for by a greater number of channels. The latter may play a critical role in some instances, producing, *e.g.*, depurinization of DNA or some other kind of degradation of analyte substances. It is also necessary to understand that 512 signals from an array in the case of a UV spectrum are not independent and *ca.* 20 channels are more than sufficient to reconstruct the full UV spectrum of a substance with high accuracy. In many instances the number of channels needed for spectrum reconstruction may be significantly smaller.

On the other hand, the PDA detector may provide an optimum selection of wavelengths for correct data analysis.

The benefits of rapid scanning detectors such as the Milikhrom are that they provide a better S/N per channel, they provide almost all the information that diode arrays do, they produce less substance degradation and they save disk storage space.

Use of different hardware

Our software allows different hardware configurations to be used for data acquisition owing to the support of several ADC types. We are only starting to gain experience in the field of multi-detector systems. Two of the most obvious problems are adjustment of data from different detectors and peak shape changes.

ACKNOWLEDGEMENTS

The authors are grateful to Dr. A. A. Alexandrov for support with the first stages of investigations. They also thank Dr. I. Nagaev for his excellent unresolved chromatogram, which he indicated was much easier to obtain than a well resolved chromatogram.

REFERENCES

1 G. I. Baram, M. A. Grachev, N. I. Komarova, M. P. Perelroysen, Yu. A. Bolvanov, F. V. Kuzmin, V. V. Kargaltsev and E. A. Kuper, *J. Chromatogr.*, 264 (1983) 69–90.
2 Yu. A. Kalambet, E. I. Burova, A. A. Zhuchkov, V. L. Knorre and A. A. Alexandrov, in H. Kalasz and L. S. Ettre (Editors), *Chromatography '87*, Akadémiai Kiadó, Budapest, 1988, pp. 251–268.
3 G. E. Forsythe, M. A. Malcolm and C. B. Moler, *Computer Methods for Mathematical Computations*, Prentice Hall, Englewood Cliffs, NJ, 1977.
4 J. C. Nicolson, J. J. Meister, D. R. Patil and L. R. Field, *Anal. Chem.*, 56 (1984) 2447–2451.
5 E. R. Malinowski and D. G. Howery, *Factor Analysis in Chemistry*, Wiley, New York, 1980.
6 F. J. Knorr and J. H. Futrell, *Anal. Chem.*, 51 (1979) 1236–1241.
7 M. A. Sharaf and B. R. Kowalski, *Anal. Chem.*, 54 (1982) 1291–1296.
8 M. F. Delaney, *Anal. Chem.*, 56 (1984) 261R–277R.
9 S. D. Brown, T. Q. Barker, R. J. Larivee, S. L. Monfre and H. R. Wilk, *Anal. Chem.*, 60 (1988) 252R–273R.
10 S. D. Brown, *Anal. Chem.*, 62 (1990) 84R–101R.