ELSEVIER

Step by step filter based program for calculations of highly informative derivative curves

Vesselin Petrov [a,*], Liudmil Antonov [b,1], Hitoshi Ehara [c,2],
Nobuyuki Harada [c,2]

[a] *Institute of Organic Chemistry, Bulgarian Academy of Science, Sofia 1113, Bulgaria*
[b] *Department of Ecology, National Forestry University, Sofia 1756, Bulgaria*
[c] *Institute for Chemical Reaction Science, Tohoku University, Sendai 980-8577, Japan*

Abstract

We have developed a microcomputer program, based on a recently described approach called step by step filter (SBSF), for calculation of derivative curves directly from spectra recorded as a function of wavelength. This program avoids the long wavelength attenuation featured at conventional method for derivative curves calculation, and in this extent could be very helpful for daily spectroscopy practice. The features of the SBSF program include: easy treatment of data through a windowed environment, calculating of both conventional and step by step filter derivatives, possibilities for selection of the mathematical conditions for smoothing and differentiation simultaneous plotting of the original curve and its derivative and a mouse pointer. Several examples from different branches of the molecular spectroscopy (absorption UV-VIS, CD and fluorescence) are provided and discussed in the terms of advantages of SBSF. © 2000 Elsevier Science Ltd. All rights reserved.

Keywords: Numerical smoothing differentiation; Spectroscopy; Data processing; Chemometrics

1. Introduction

The importance of the derivative spectroscopy for interpretation of UV-VIS spectra (Talsky, 1994; Antonov and Stoyanov, 1993a) and for quantitative analysis (Ojeda et al., 1995; Bridge et al., 1987) is well known and documented. Its popularity results as from the good SNR in the absorption UV-VIS spectra as well as from the possibilities for easy obtainment of derivative curves in the commercial instruments and software,

where well studied procedures of Savitzky and Golay (1964) or Butler and Hopkins (1970) are implemented. On the other hand a problem appears when the spectrum is recorded as a function of wavelength (λ in nm) (Antonov and Stoyanov, 1993b; Antonov, 1997a), since a strong attenuation of the derivative curve at the long wavelengths is observed, leading to substantial loss of spectral information and causing troubles of the users. The problem is seriously enough since the wavelength linear spectra are generated from instruments containing grating monochromator, used practically in all commercial UV-VIS spectrophotometers available now. Recently these problems were solved by a new approach, called SBSF (Antonov, 1997a) and the aim of this smdy is to demonstrate the achievement of SBSF procedure as well as the original software developed.

---

* Corresponding author.

E-mail address: veso@orgchm.bas.bg (V. Petrov)

[1] http://www.orgchm.bas.bg/ ~ lantonov
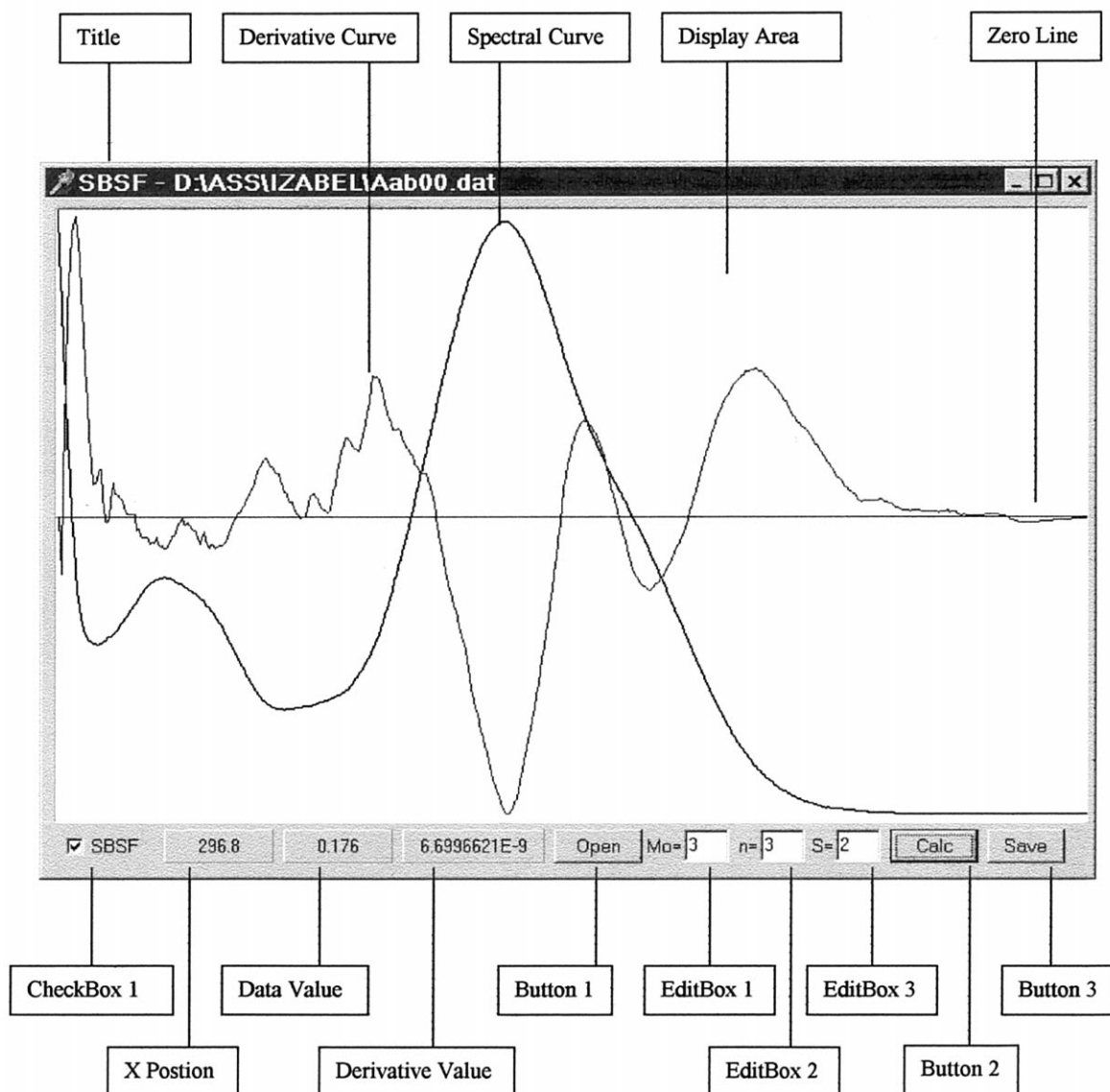
[2] http://www.chiral.icrs.tohoku.ac.jp

2. Description of the method

It is well known that any absorption spectrum presented as uniformly spaced points (with total number w) can be easy differentiated by using the classic Savitzky and Golay procedure, based of polynomial least squares method (Savitzky and Golay, 1964). In this case spectral points i to be differentiated is approximate through its neighboring points from $i - m$ to $i + m$ (the value $2m + 1$ is called filter width) by using n-order polynomial ($n < 2m + 1$). The value of m, selected by the researcher, is a constant determining the smoothing (differentiation) through the whole spectrum.

But if the spectrum is linear to the wavelength scale the classic Savitzky and Golay procedure leads to attenuating s_{th} derivative $d^s A(\lambda)/d\lambda^s$ (Fig. 1b and d) (Antonov, 1997a). The more informative derivative ($d^s A(\lambda)/d\tilde{v}^s$) (where $\tilde{v}$ represents wavenumbers in cm^{-1}) can be obtained only if m is not kept constant. It should depend on λ in order the value of the filter width to be constant in cm^{-1}, not in nm. In this case an initial value m_0 can be selected and definite filter width as a function of λ can be calculated:

$$m_i = \text{integer}\left[\frac{\lambda_i^2 m_0}{\lambda_{m_0+1}^2 + \Delta\lambda m_0(\lambda_i - \lambda_{m_0+1})}\right] \quad (1)$$

Scheme 1.

Fig. 1. (a) Simulated spectrum in the wavenumber scale; (b) the same spectrum transformed into the wavelength scale; derivatives of (b): (c) Second derivative using SBSF method ($m_0 = 2$; $n = 3$; $s = 2$); (d) second derivative using conventional procedure (Gorry, 1990) ($m = 2$; $n = 3$; $s = 2$); (e) fourth derivative using SBSF method ($m_0 = 4$; $n = 5$; $s = 4$); (f) fourth derivative using the conventional procedure (Gorry, 1990) ($m = 4$; $n = 5$; $s = 4$).

```
                    ┌─────────┐
                   (   Begin   )
                    └─────────┘
                         │
                         ▼
           ┌──────────────────────────┐ ◄───────────────────┐
           │        Read Data          │                     │
           └──────────────────────────┘                     │
                         │                                    │
                         ▼                                    │
           ┌──────────────────────────┐                     │
           │    Read Begin Values      │                     │
           └──────────────────────────┘                     │
                         │                                    │
                         ▼          No    ┌──────────────────────────────────────┐
                       ╱   ╲ ───────────► │ Savitzky & Golay procedure using P. Gorry │
                      ╱ SBSF ?╲           │              [10] method                │
                       ╲     ╱            └──────────────────────────────────────┘
                        ╲   ╱                           │
                         │ Yes                          │
                         ▼                              │
           ┌──────────────────────────┐                │
           │    Calculation of m_w     │                │
           └──────────────────────────┘                │
                         │                              │
                         ▼                              │
           ┌──────────────────────────┐                │
           │  Points range 1÷w- m_w    │                │
           │ Calculation of the matrices│               │
           │  X,Y and B using eqs. 4 & 5│               │
           └──────────────────────────┘                │
                         │                              │
                         ▼                              │
           ┌──────────────────────────┐                │
           │ Points range w- m_w +1 ÷ w│               │
           │ Calculation of the matrices│               │
           │  X,Y and B using eqs. 7 & 8│               │
           └──────────────────────────┘                │
                         │ ◄────────────────────────────┘
                         ▼
           ┌──────────────────────────┐
           │       Show Result         │──────┐
           └──────────────────────────┘      │
                         │                    │
                         ▼                    │
           ┌──────────────────────────┐      │
           │          Save             │      │
           └──────────────────────────┘      │
                         │                    │
                         └────────────────────┤
                                              ▼
                                          ┌─────────┐
                                         (    End    )
                                          └─────────┘
```

Scheme 2.

Thus the derivative spectrum $(d^s A(\lambda_i)/d\tilde{v}^s)$ $(i = 1 \div w)$ can be calculated using the expression:

$$\frac{d^s A(\lambda_i)}{d\tilde{v}^s} = s![B_i]_s, \qquad (2)$$

where $[B_i]_s$ is the corresponding s_{th} element of the vector B_i obtained according the following matrix calculation:

$$X_i B_i = Y_i \qquad (3)$$

$$X_i = \begin{bmatrix} 2m_i + 1 & \sum\limits_{j=-m_i}^{m_i} x_{i+j} & \cdots & \sum\limits_{j=-m_i}^{m_i} x_{i+j}^{n} \\ \vdots & \vdots & \ddots & \vdots \\ \sum\limits_{j=-m_i}^{m_i} x_{i+j}^{n} & \sum\limits_{j=-m_i}^{m_i} x_{i+j}^{n+1} & \cdots & \sum\limits_{j=-m_i}^{m_i} x_{i+j}^{2n} \end{bmatrix}$$

(4)

$$Y_i = \begin{bmatrix} \sum\limits_{j=-m_i}^{m_i} A(\lambda_{i+j}) \\ \vdots \\ \sum\limits_{j=-m_i}^{m_i} A(\lambda_{i+j}) \cdot x_{i+j}^{n} \end{bmatrix}$$

(5)

$$x_{i+j} = a\left(\frac{1}{\lambda_i} - \frac{1}{\lambda_i + j\Delta\lambda}\right)$$

(6)

where a is the conversion factor between wavelength [in nm] and wavenumber [in cm^{-1}], equal to 1×10^7.

The essence of SBSF is to calculate $(\mathrm{d}^s A(\lambda_i)/\mathrm{d}\tilde{v}^s)$ directly form the spectrum linear in wavelength (Fig. 1c and e), keeping the filter window constant in cm^{-1}, according to Eq. (1). However, the constant (in cm^{-1}) filter width leads to change of the filter window as a function of λ and as a result the convolution coefficients B_i, cannot be tabulated. This fact requires their calculation at every $A(\lambda_i)$, but allows the derivative to be obtained without attenuation and loss of spectral information. The reliability of SBSF has already proved by careful analysis in the second and forth derivative spectra (Antonov, 1997a,b).

The use of Eq. (1) unfortunately leads to loss of great number of spectral points in long wavelength area. This can be easy make off using asymmetrical filter width defined by the following equations:

a)

b)

c)

d)
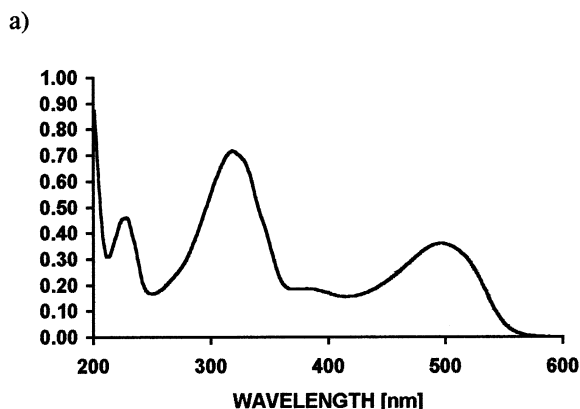

Fig. 2. (a) Spectrum of 4-aminoazobenzen; (b) Second derivative using conventional procedure (Gorry, 1990) ($m = 3$; $n = 3$; $s = 2$); (c) Second derivative using SBSF method ($m_0 = 3$; $n = 3$; $s = 2$); (d) the long wavelength part of the curve from (b) zoomed.
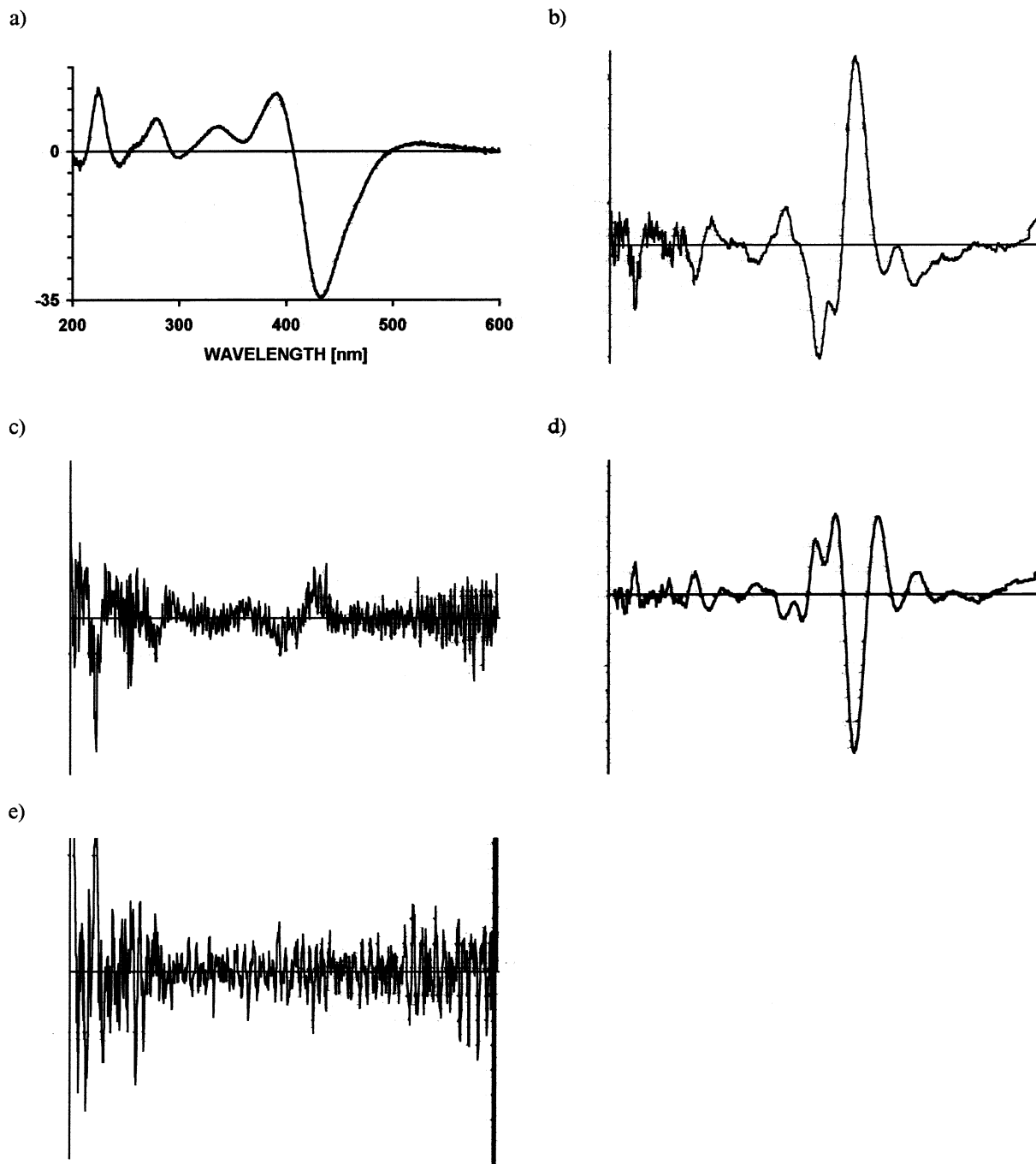
Fig. 3. (a) CD spectrum of a chiral azulene derivative; (b) Second derivative using SBSF method ($m_0 = 5$; $n = 3$; $s = 2$); (c) second derivative using conventional procedure (Gorry, 1990) ($m = 5$; $n = 3$; $s = 2$); (d) fourth derivative using SBSF method ($m_0 = 15$; $n = 5$; $s = 4$); (e) fourth derivative using conventional procedure ($m = 15$; $n = 5$; $s = 4$).

X_i

$$= \begin{pmatrix} 2m_i+1 & \sum\limits_{j=m_i-\alpha}^{m_i-\alpha} x_{i+j} & \cdots & \sum\limits_{j=m_i-\alpha}^{m_i-\alpha} x_{i+j} \\ \vdots & \vdots & \ddots & \vdots \\ \sum\limits_{j=m_i-\alpha}^{m_i-\alpha} x_{i+j}^n & \sum\limits_{j=m_i-\alpha}^{m_i-\alpha} x_{i+j}^{n+1} & \cdots & \sum\limits_{j=-m_i-\alpha}^{m_i-\alpha} x_{i+j}^{2n} \end{pmatrix} \quad (7)$$

$$Y_i = \begin{pmatrix} \sum\limits_{j=m_i-\alpha}^{m_i-\alpha} A(\lambda_{i+j}) \\ \vdots \\ \sum\limits_{j=m_i-\alpha}^{m_i-\alpha} A(\lambda_{i+j}) \cdot x_{i+j}^n \end{pmatrix} \quad (8)$$

where α is defined as

- $\alpha = 0$ when $i + m_i \leq w$
- $\alpha = i + m_I - w$ when $i + m_i > w$

Almost the same approach can be used for the short wavelength area, which we do not recommend because in this area only a few points are lost and the errors caused by the calculations are very high.

3. Structure of the program

The structure of the program (Scheme 1) can be described as follows:
1. Input of data by Button 1.
2. Setting of the mathematical conditions of smoothing (differentiation) by Edit Box 1, 2 and 3.
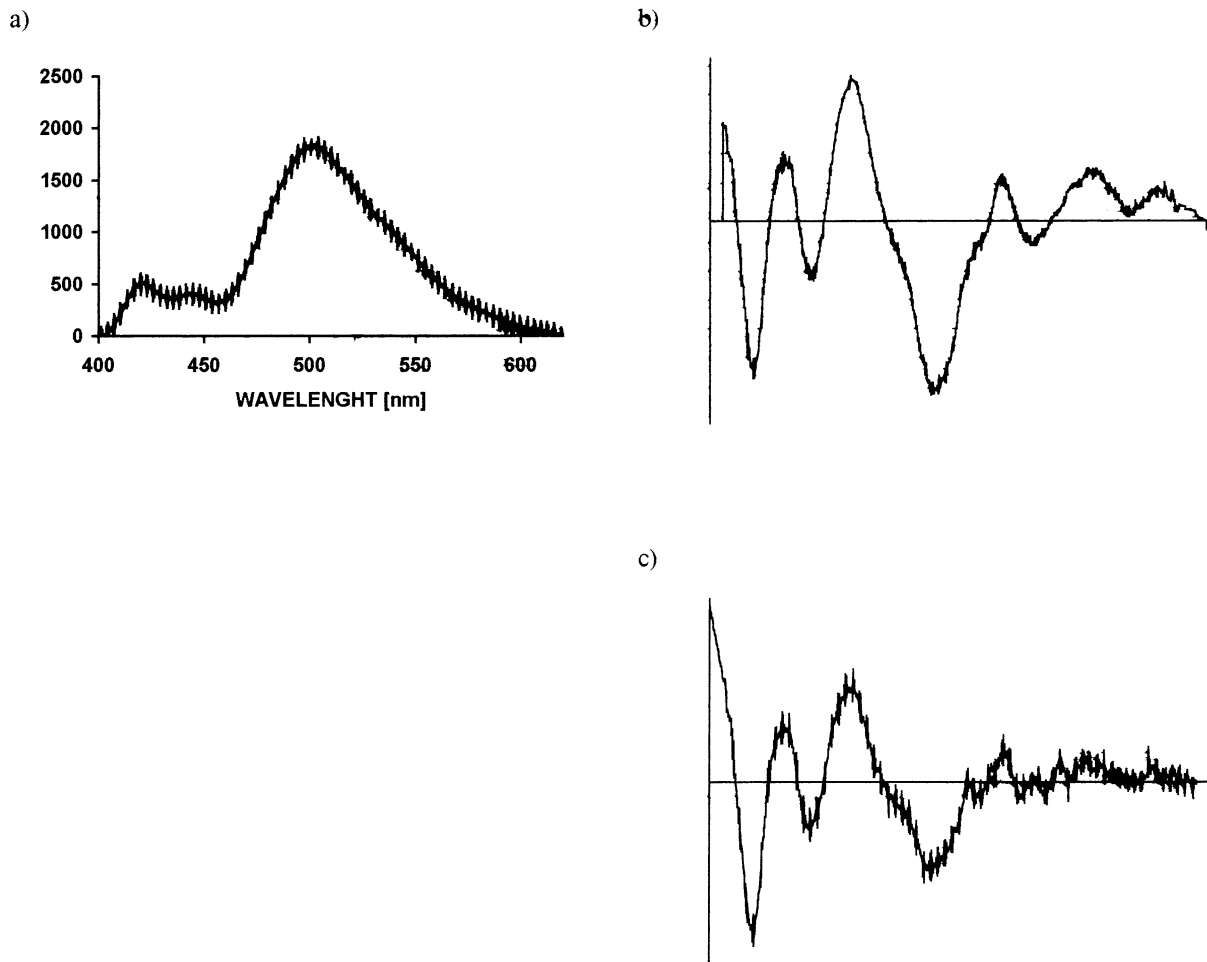
a)

b)

c)

Fig. 4. (a) Spectrum of BiAcridine; (b) second derivative using SBSF method ($m_0 = 30$; $n = 3$; $s = 2$); (c) second derivative using conventional procedure (Gorry, 1990) ($m = 30$; $n = 3$; $s = 2$).

3. Setting the mathematical model by CheckBox 1. If the 'SBSF' is checked the SBSF procedure is applied, otherwise the conventional Golay–Savitzky method is used.
4. Calculations of the derivative (smoothed) spectrum by Button 2. In the case of SBSF the resulting curve is given in blue, while in the conventional Golay–Savitzky procedure red color is used.
5. Saving results by Button 3

The block diagram of the program is shown on the Scheme 2.

The program is based on the described above step by step filter method, which is applied without changes. It is written on program language Delphi version 3.0 because it gives a possibility for easy calculation and visualization.

4. Experimental part

The spectra used below are recorded as follows:
1. Absorption spectra – on PE λ 17 W-VIS spectrophotometer.
2. Fluorescent spectra – on PE 440 MBF spectrophotometer.
3. CD absorption spectra – on Jasco J 550.

All of the spectra are recorded at optimal instrumental conditions in order the noise level to be lower as possible, but without using additional mathematical smoothing.

5. Results and discussion

In order the advantages of SBSF to be proved without any suspicious the developed program was tested as on simulated as well as on wide range of real molecular spectra.

5.1. Simulated spectrum

Three bands with the same half-width was generated (Fig. 1a) in coordinates $A(\tilde{v})$. After being transformed in coordinates $A(\lambda)$ these three bands (Fig. 1b) look as recorded on grating instrument. The transformation leads to visual asymmetry of the bands (although they are actually the same) caused by a compressing in the short wavelength area, and broadening in the long wavelength area. The visual broadening in the long wavelength region results in attenuation of the corresponding derivative curves (Fig. 1d and f) becoming higher with increase of the order of the derivative. This attenuation can not be avoided using the conventional Golay–Savitzky procedure. But as shown (Fig. 1c and d) the situation changes after using SBSF, where loss of spectral information at the long wavelengths even in higher order derivatives cannot be observed.

5.2. Absorption spectrum of 4-aminoazobenzen

The absorption spectrum of 4-aminoazobenzen in EtOH/water (50%/50%) is presented on Fig. 2a. It is a feature of such kind of aminoazodyes to exhibit color change under protonation (Stoyanova et al., 1996) determining their wide use as pH indicators and resulting in appearance of new bands as shown in Fig. 2a. It is very important for the practice the structure of these batochromicaly shifted bands to be studied in the term of evaluation of the number and positions of the corresponding sub-bands.

The instrumentally obtained second derivative of the protonated spectrum is presented on Fig. 2b. It is clear that even after smoothing the noise is not removed and due to above cited attenuation any information about the structure of the bands and there positions can not be gained directly. In Fig. 2c the second derivative obtained by SBSF at the same polynomial order and filter width is presented for comparison. This curve is even visually less noisy (in comparison with Fig. 2d) and very easy for interpretation. The position and structure of the band at ≈ 500 nm due to protonation can be easy estimated using the mouse pointer.

5.3. CD spectrum of a chiral azulene derivative

It is well known that CD spectroscopy is very powerful for determining the absolute stereochemistry of natural products (Harada and Nakanishi, 1983). In general, CD spectra are measured at extremely low SNR, and therefore sometimes requiring additional smoothing. Fig. 3a shows the CD spectrum of a chiral azulene derivative, which was selected as a typical CD spectrum exhibiting medium intensity but low SNR especially in the long and short wavelength regions. The second derivative spectra of the compound were calcu-lated using the same polynomial order and filter width by the old procedure of Golay–Savitzky (Fig. 3c) and also by the SBSF method (Fig. 3b). A substantial difference in the quality of the derivative curves is evident. Without suspicion the increase of the filter width could lead to some improvement of the Golay–Savitzky derivative, but it will cause curve distortion. The derivatives obtained by the SBSF method are optimal concerning these opposite trends, noise removal and curve deformation, especially for the higher order derivatives as shown in Fig. 3d and e.

5.4. Fluorescent spectrum of bi-acridine

Fluorescent spectra consist also overlapping bands and sometimes it is important to estimate their positions by derivative spectroscopy. It is not easy since the SNR is lower comparing with the conventional W-VIS spectroscopy. On Fig. 4 the fluorescence spectrum of

Bi-Acridine with unhomogenious noise, as well as its second derivatives are shown. Again SBSF shows better possibility for noise removal at the same mathematical conditions.

6. Conclusions

The newly developed SBSF, avoiding the loss of spectral information in the derivative curves is reported. Achievement of the method is described allowing the differentiation to be performed without loss of the end spectral points as it is in the conventional methods. The advantages of SBSF are clearly shown in real spectra from various branches of molecular spectroscopy. A SBSF microcomputer program is developed allowing easy treatment of spectra recorded as a function of wavelength.

7. Program availability

The program is freeware, but the users are kindly requested to cite both reference Antonov (1997a) and this paper in the case of use. The SBSF program can be obtained by writing to the authors. It is also available with two working examples at http://www.orgchm. bas.bg/ ∼ lantonov or via ftp://ftp.orgchm.bas.bg/pub/ Chemistry_Software/software/MS-WINDOWS95/veso/ SBSF.zip.

References

Talsky, G., 1994. Derivative Spectroscopy. VCH, Weinheim.

Antonov, L., Stoyanov, S., 1993a. Appl. Spectr. 47, 1030.

Ojeda, C.B., Rojas, F.S., Sano Pavon, J.M., 1995. Talanta 42, 1195.

Bridge, T.P., Fell, A.F., Wardman, R.H., 1987. J. Soc. Dyers Colour 103, 17.

Savitzky, A., Golay, M.J.E., 1964. Anal. Chem. 36, 1627.

Butler, W.L., Hopkins, D.W., 1970. Photochem. Photobiol. 12, 439.

Antonov, L., Stoyanov, S., 1993b. Appl. Spectr. 47, 1712.

Antonov, L., 1997a. Trends Anal. Chem. 16, 536.

Antonov, L., 1997b. Anal. Chim. Acta 349, 295.

Gorry, P.A., 1990. Anal. Chem. 62, 570.

Stoyanova, T., Stoyanov, S., Antonov, L., Petrova, V., 1996. Dyes and Pigments 31, 1.

Harada, N., Nakanishi, K., 1983. Circular Dichroic Spectroscopy-Exciton Coupling in Organic Stereochemistry. University Science Books.