Wavelet transform for the evaluation of peak intensities in flow-injection analysis

M. Bos and E. Hoogendam

Department of Chemical Technology, University of Twente, P O Box 217, 7500 AE Enschede (Netherlands)

(Received 18th September 1991, revised manuscript received 13th April 1992)

Abstract

The application of the wavelet transform in the determination of peak intensities in flow-injection analysis was studied with regard to its properties of minimizing the effects of noise and baseline drift. The results indicate that for white noise and a favourable peak shape a signal-to-noise ratio of 2 can be tolerated at the 5% error level, which means that a significant reduction in the detection limit can be obtained in comparison with the classical signal-processing methods. With regard to the influence of a changing baseline it was found that its d c level has a negligible effect, but a linear or exponentially rising baseline introduces an error that depends on the chosen frequency of the wavelet that is used to determine the peak intensity. The optimum choice of this frequency, in turn, depends on the shape of the peak that is studied. In this respect significant differences were observed for pure Gaussian and exponentially modified Gaussian peaks.

Keywords Flow injection, Signal processing methods, Peak intensities, Wavelet transform

The main limitations in the determination of peak intensities in flow-injection analysis (FIA) are noise and a drifting baseline. Peak overlap generally is not a problem because it can easily be avoided by adjusting the sample rate. Several methods have been described [1–4] that address the automatic evaluation of the peak-shaped FIA signals. In general, the procedure can be fairly straightforward if sufficient attention is paid to the design of the FIA system in order to obtain a baseline free from drift and a well formed peak. The aforementioned problems only come into play when a FIA system is operated near its detection limits. Here automatic signal processing encounters practical difficulties in locating the peak and finding the right baseline correction method.

The detection of weak signals embedded in stochastic noise is a more general problem for which correlation of the observed signal with a replica of the desired signal is in use. Correlation is most effective if the shape and the time of occurrence of the relevant signal is known. Instead of correlation a matched filter can be used. Also in this case the shape of the signal or at least the band width should be known.

This paper shows the advantages of the use of a signal-processing technique called wavelet transformation to tackle these problems. The main characteristic of this technique is that it transforms the information contained in the signal into a two-dimensional time–frequency form. In this transform time and frequency information of the signal is retained. This permits the choice of a coefficient in this two-dimensional representation that reflects the peak intensity optimally filtered from noise. This choice can be automated by searching for the maximum coefficient in the
transform of a well defined peak obtained for a sample with a relatively high concentration. The position of this maximum on the horizontal time axis of the transform conveys the positional information of the peak whereas its position on the vertical frequency axis determines the noise-filtering characteristic that can be obtained.

Automatic full compensation of baseline d.c. level drift and, to a lesser extent, of a sloping baseline of various mathematical forms can be realised by proper choice of the form of the analysing wavelet.

**THEORY**

The purpose of the wavelet transform procedure is to decompose a signal into localized contributions that are characterized by a so-called scale parameter. The mathematical techniques to perform this decomposition have been described in a number of papers [5–9]. Here only a short review of the main equations is given.

Just like other types of transforms, e.g., the Fourier transform, the wavelet transform realizes a correspondence between physical space \( t \) and spectral space \( \omega \), by convoluting the signal \( s(t) \) to be analysed with a given analysing function \( \psi(t) \) defined in the two-dimensional space \([\omega, t]\). Figure 1 shows four examples of analysing functions \( \psi(t) \). It shows how the time-scale methods windowed Fourier transform and wavelet transform achieve an optimum compromise between time resolution \( \Delta t \) and frequency resolution \( \Delta \omega \): \( \Delta t \Delta \omega = \text{constant} \). Owing to their ability to encompass time variations of spectral properties,
time-scale representations are particularly adapted to the analysis of non-periodic, non-stationary signals.

**Analysing wavelets**

If the analysing wavelet is denoted as the function \( g(t) \) and its Fourier transform by \( \hat{g}(\omega) \), a number of “self-similar” elementary wavelets \( g_{a,b}(t) \) concentrated in time and frequency can be obtained from some basic wavelength \( q_{\text{basic}}(t) \) by shifts in the time variable and by dilations which act both on the time and the frequency variable

\[
\begin{align*}
g_{a,b}(t) &= \frac{1}{\sqrt{a}} \, q_{\text{basic}} \left( \frac{t - b}{a} \right) \\
\end{align*}
\]

where \( 1/\sqrt{a} \) is a normalizing constant that ensures that all analysing functions have “unit energy.”

The analysing wavelet function has to fulfil a number of conditions

\[
\int |g(t)|^2 \, dt < \infty
\]

which ensures finite energy and

\[
\int |\hat{g}(\omega)|^2 \frac{d\omega}{|\omega|} < \infty
\]

which gives a restriction on the behaviour of \( \hat{g}(\omega) \) around the zero frequency. It ensures a short-wave like behaviour and generally implies that \( g_{\text{basic}}(t) \) does not have a d.c. component.

A third restriction imposed on the function representing the analysing wavelet is

\[
\hat{g}_{\text{basic}}(\omega) = 0 \quad \text{for} \quad \omega < 0
\]

The wavelet transform is now defined as the function \( S(b,a) \) on the open \((b,a)\) half-plane \((b\ \text{arbitrary}, \ a > 0)\)

\[
S(b,a) = \int_{-\infty}^{\infty} [g_{b,a}(t)]^* s(t) \, dt
\]

where the asterisk denotes that the function \( g \) is complex. This equation can be rewritten as

\[
S(b,a) = \frac{1}{\sqrt{a}} \int_{-\infty}^{\infty} g_{\text{basic}}^* \left( \frac{t - b}{a} \right) s(t) \, dt
\]

The conditions imposed on the wavelet function imply that \( \hat{g}(\omega) \) is negligible above a certain frequency \( \omega_{\text{max}} \), making \( S(b,a) \) insensitive to the higher Fourier components of the signal \( s(\omega) \). This eliminates the influence of small-scale features. Further, these conditions make \( g \) negligible outside an interval \([t_{\text{min}}, t_{\text{max}}]\) on the \( t \)-axis, causing \( S(b,a) \) to be insensitive to the values of \( s(t) \) such that \( t - b \) lies outside of \([t_{\text{min}}, t_{\text{max}}]\). The latter is the origin of the desired localization in time.

A commonly used analysing wavelet that fulfils these conditions is the Morlet wavelet

\[
g_{\text{basic,Morlet}}(t) = e^{i\omega_0 t} e^{-t^2/2} + \text{small corrections}
\]

These small corrections are numerically negligible when \( \omega_0 > 5 \) but have to be added since \( \exp[-(\omega - \omega_0)^2] \) does not vanish for \( \omega < 0 \).

Figure 2 shows a number of these wavelets with different values for the scale parameter \( a \).

**Discretization of the wavelet transform**

In practice, the signals to be analysed will be acquired by computer at discrete time intervals. This necessitates the use of a discrete form of Eqn 6

\[
S(iT_s,a) = T_s \frac{1}{\sqrt{a}} \sum_{n}^{} g_{\text{basic}}^* \left[ \frac{(n-i)T_s}{a} \right] s(nT_s)
\]

where \( 1/T_s \) is the sampling frequency. Let \( g_{a}^*(iT_s) = g^*(iT_s/a) \). Then, for each value of \( a \), the analysing wavelet is sampled, yielding the sequence \( g_{a}^*(iT_s) \). The convolution product between \( s(nT_s) \) and \( g_{a}^*(iT_s) \) is then computed. As \( g(t) \) has finite support, the number of sampling points of \( g_{a}^*(t) \) is finite and grows linearly with \( a \).

In this work the wavelet of Eqn 7 is used with \( \omega_0 = 5 \).
All $g^*(t)$ wavelets are derived from $g^*_{\text{basic}}(t)$ through the dilation operator by sampling with a narrower interval.

More details and a more efficient algorithm that is suitable for larger applications can be found in [10].

**EXPERIMENTAL**

**Chemicals**

Imidazole samples were prepared from a stock solution of 0.00563 M imidazole (Merck, analytical-reagent grade) by dispensing from a piston burette and diluting with carbon dioxide-free deionized water that had been passed through a Millipore Q1 filter. With these samples a reagent stream of 0.1992 M acetic acid diluted from pure acetic acid (Merck, analytical-reagent grade) was used. This was standardized titrimetrically.

18-Crown-6 (Merck, zur Synthese) samples were prepared by dissolving the required amount of substance in deionized water obtained from a Millipore Q2 installation. The reagent stream was 0.2 M barium chloride (Merck, analytical-reagent grade, standardized titrimetrically).

**Equipment for FIA with enthalpymetric detection**

A schematic diagram of the FIA equipment is given in Fig. 3. The sampler and the injection valve (Skalar Model 1000 and 1550, respectively) are computer controlled (Apple IIE). The pump is a Gilson Minipuls 2. The reagent and carrier streams are immersed in a 5-l stirred water-bath that is thermally isolated from its surroundings by 1 cm of Styrofoam. The temperature of this water-bath is controlled to 0.01°C by a Tamson thermostat via a glass heat exchanger. The length of the PTFE coil (tubing 1 dm 0.5 mm o d 1.5 mm) to equilibrate the reagent thermally is 1 m. For the sample a 20-cm stainless-steel coil (tubing 1 dm 0.5 mm o d 1.5 mm) is used for this purpose. The enthalpymetric detector consists of a measuring and a reference thermistor (Philips, NTC626-12223) connected in a Wheatstone bridge circuit with fixed resistors of 21.5 kΩ. The supply voltage for this circuit is 5 V. The bridge signal is measured every 200 ms with a Hewlett-Packard Model 3478A multimeter connected to the computer by its HP-IB interface. The sample loop in the injection valve has a volume of 0.200 ml. The flow-rate of the carrier and the reagent streams is 2.5 ml min⁻¹.

**RESULTS AND DISCUSSION**

The suitability of the wavelet transform for the evaluation of FIA peaks and its optimum parameter settings were studied on synthetic and real signals. With the use of simulated FIA signals a systematic survey of the influence of noise and various forms of baseline drift could be obtained. Whereas the experiments in which an enthalpymetric detector was used show what can be gained with this technique in practice for noisy signals on a drifting baseline.

**Results for simulated signals**

FIA peaks can have a Gaussian shape, but more generally one finds a so-called exponen-
trially modified Gaussian peak (EMG) [11–14]. The simulations were carried out with both signal types. For the Gaussian peak, the following equation was used in the computer program:

$$s(t) = \frac{A}{\sqrt{2\pi}} \exp\left[-\frac{(p-i)^2}{2\sigma^2}\right]$$

(9)

where $A$ = peak height, $\sigma$ determines the width of the peak and $p$ is its position on the discretized $t$-axis.

The EGM peaks were constructed using the equations given by Foley [11]:

$$s(t) = \frac{A}{\tau} \exp\left(\frac{\sigma_g^2}{2\tau} - \frac{t-t_g}{\tau}\right)$$

$$\int_{-\infty}^{\infty} \frac{1}{\sqrt{\pi}} \exp\left(-\frac{x^2}{2}\right) dx$$

(10)

with

$$Z = \frac{t-t_g}{\sigma_g} - \frac{\sigma_g}{\tau}, \text{ time } t = 0 \text{ and } j = 1$$

201

where $A$ is the area of the peak, $\tau$ is the time constant of the exponential decay, $t_g$ determines the position of the peak on the time axis and the ratio $\tau/\sigma_g$ is a measure of its asymmetry.

Random white noise of uniform distribution was generated with the use of a pseudo-random number generator provided in the CSL software library [15].

---

**Fig 4** Schematic representation of a grid on the plane $(a,b)$ allowing arbitrary precise reconstruction of signals

---

**Fig 5** Calibration graphs for peak intensities determined by the wavelet transform for (○) EMG and (+) Gaussian peaks.

The transforms are carried out with a grid of $a$ and $b$ values that is depicted qualitatively in Fig 4 and mathematically described by

$$a_i = 2^{-\alpha i}, \quad b_{i,j} = \beta a_i, \quad i,j \in N$$

(11)

where $\alpha$ and $\beta$ are sampling constants. In the calculations, the values used were $\alpha = 0.25$ and $\beta = 0.1$. Using these values, the maximum wavelet transform coefficients were found at $a = 0.35$, $b = 0.28$ (Gaussian peak) and $a = 0.500$, $b = 0.25$ (EMG).

**Linearity and independence of baseline d c level** Figure 5 shows the perfect linear relationship between peak height and the maximum coefficient of the wavelet transform of the signal. Table 1 shows that the d c level has a negligible influence on the value of the transform.

**Baseline drift** The error produced by a linear drifting baseline

$$\text{baseline}[t] = \text{slope}(t - \text{offset})$$

(12)

is shown in Fig 6.

---

**TABLE 1**

<table>
<thead>
<tr>
<th>Baseline value</th>
<th>Intensity</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>EMG peak</td>
</tr>
<tr>
<td>0 0</td>
<td>55 5737</td>
</tr>
<tr>
<td>1 0</td>
<td>55 5745</td>
</tr>
<tr>
<td>5 0</td>
<td>55 5795</td>
</tr>
<tr>
<td>10 0</td>
<td>55 5856</td>
</tr>
<tr>
<td>100 0</td>
<td>55 6961</td>
</tr>
</tbody>
</table>
Figure 7 shows the errors produced by an exponential rise of the baseline. The equation used to simulate the exponentially rising baseline was

$$\text{baseline}[t] = \exp[(t - 1500) \times \text{factor}]$$  \hspace{1cm} (13)

*Noise* The influence of noise was tested by superimposing uniformly distributed random signals on the signals generated by Eqns 9 and 10 using various amplitudes for the noise. The summary of the results given in Fig 8 shows that this influence is less than 5% for a signal-to-noise ratio of less than 2 for the EMG peak and less than 10 for the Gaussian peak.

**Experimental results**

Table 2 shows the peak intensities found by the wavelet transform procedure for different samples of imidazole. The reagent used in this determination was 0.1992 M acetic acid. The linearity with concentration is within 5% down to the lowest concentration measured, i.e., 5.63 × 10⁻⁴ M Figure 9 shows the measured signal for this concentration and clearly demonstrates the robustness of the wavelet transform technique against noise and baseline fluctuations. Table 3 shows similar results for the reaction between 0.2

<table>
<thead>
<tr>
<th>Concentration (10⁻³ M)</th>
<th>Calculated intensity</th>
<th>Intensity/concentration</th>
<th>Deviation from mean (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>5.63</td>
<td>4.241</td>
<td>0.753</td>
<td>12</td>
</tr>
<tr>
<td>5.63</td>
<td>4.256</td>
<td>0.755</td>
<td>14</td>
</tr>
<tr>
<td>5.06</td>
<td>3.720</td>
<td>0.735</td>
<td>12</td>
</tr>
<tr>
<td>5.06</td>
<td>3.710</td>
<td>0.733</td>
<td>14</td>
</tr>
<tr>
<td>4.53</td>
<td>3.481</td>
<td>0.768</td>
<td>3.2</td>
</tr>
<tr>
<td>4.53</td>
<td>3.459</td>
<td>0.763</td>
<td>2.5</td>
</tr>
<tr>
<td>3.94</td>
<td>2.826</td>
<td>0.717</td>
<td>3.6</td>
</tr>
<tr>
<td>3.94</td>
<td>2.886</td>
<td>0.732</td>
<td>1.6</td>
</tr>
<tr>
<td>3.38</td>
<td>2.541</td>
<td>0.751</td>
<td>0.09</td>
</tr>
<tr>
<td>3.38</td>
<td>2.431</td>
<td>0.719</td>
<td>3.3</td>
</tr>
<tr>
<td>2.81</td>
<td>2.067</td>
<td>0.735</td>
<td>12</td>
</tr>
<tr>
<td>2.81</td>
<td>1.917</td>
<td>0.682</td>
<td>8.3</td>
</tr>
<tr>
<td>2.25</td>
<td>1.468</td>
<td>0.652</td>
<td>11.3</td>
</tr>
<tr>
<td>2.25</td>
<td>1.552</td>
<td>0.689</td>
<td>7.3</td>
</tr>
<tr>
<td>1.69</td>
<td>1.315</td>
<td>0.778</td>
<td>4.5</td>
</tr>
<tr>
<td>1.69</td>
<td>1.344</td>
<td>0.795</td>
<td>6.8</td>
</tr>
<tr>
<td>1.13</td>
<td>0.863</td>
<td>0.763</td>
<td>2.5</td>
</tr>
<tr>
<td>1.13</td>
<td>0.928</td>
<td>0.821</td>
<td>10.3</td>
</tr>
<tr>
<td>0.56</td>
<td>0.428</td>
<td>0.761</td>
<td>2.2</td>
</tr>
<tr>
<td>0.56</td>
<td>0.440</td>
<td>0.782</td>
<td>5.1</td>
</tr>
</tbody>
</table>
During the experiments it was found that the position of the maximum wavelet transform coefficient is stable within a 1-week period. This means that in practice one does not have to perform the complete wavelet transform of the FIA signal. A convolution of the signal with a Morlet wavelet with the right parameters $a$ and $b$ is then sufficient.

Comparison with classical signal-processing procedures

To show the merits of the wavelet signal-processing method, the experimental data were also subjected to classical signal evaluation procedures in which the peak heights and peak areas are determined. The calibration data for wavelet intensity, peak height and peak area versus concentration were used in the procedure given by Funk et al. [16] to determine the detection limits of the various calibrations. The performance of the peak-area determination method was such that higher concentration ranges had to be used to fulfill the requirement that the lowest concentration used in the calibration series should be twice the test value $x_p$ given by Funk et al.'s procedure.

The detection limits obtained are given in Table 4, which shows that the performance of the calibration with peak area is much poorer than that with peak height, whereas the improvement due to the application of the wavelet transform is more pronounced for the 18-crown-6-determination than for imidazole. This can be explained by the fact that the imidazole peaks are "narrower" and is in agreement with the simulation results where higher noise levels could be tolerated for EMG than for Gaussian peaks.

The authors express their thanks to H F Bulthuis for sharing his experience in wavelet theory with them.

### TABLE 3
Peak intensities versus concentration for 18-crown-6 samples and barium chloride reagent

<table>
<thead>
<tr>
<th>Concentration ($10^{-3}$ M)</th>
<th>Calculated intensity</th>
<th>Intensity/concentration</th>
<th>Deviation from mean (%)</th>
</tr>
</thead>
<tbody>
<tr>
<td>4.50</td>
<td>3.293</td>
<td>0.731</td>
<td>0.1</td>
</tr>
<tr>
<td>4.50</td>
<td>3.208</td>
<td>0.712</td>
<td>2.4</td>
</tr>
<tr>
<td>4.05</td>
<td>2.823</td>
<td>0.697</td>
<td>4.5</td>
</tr>
<tr>
<td>4.05</td>
<td>2.809</td>
<td>0.693</td>
<td>5.0</td>
</tr>
<tr>
<td>3.60</td>
<td>2.583</td>
<td>0.717</td>
<td>1.7</td>
</tr>
<tr>
<td>3.60</td>
<td>2.515</td>
<td>0.698</td>
<td>4.3</td>
</tr>
<tr>
<td>3.15</td>
<td>2.279</td>
<td>0.725</td>
<td>0.9</td>
</tr>
<tr>
<td>3.15</td>
<td>2.136</td>
<td>0.678</td>
<td>7.1</td>
</tr>
<tr>
<td>2.70</td>
<td>1.916</td>
<td>0.709</td>
<td>2.8</td>
</tr>
<tr>
<td>2.70</td>
<td>2.028</td>
<td>0.731</td>
<td>5.8</td>
</tr>
<tr>
<td>2.25</td>
<td>1.615</td>
<td>0.717</td>
<td>1.8</td>
</tr>
<tr>
<td>2.25</td>
<td>1.727</td>
<td>0.767</td>
<td>6.0</td>
</tr>
<tr>
<td>1.80</td>
<td>1.468</td>
<td>0.815</td>
<td>11.6</td>
</tr>
<tr>
<td>1.80</td>
<td>1.326</td>
<td>0.737</td>
<td>1.9</td>
</tr>
<tr>
<td>1.35</td>
<td>0.965</td>
<td>0.714</td>
<td>2.1</td>
</tr>
<tr>
<td>1.35</td>
<td>1.102</td>
<td>0.816</td>
<td>11.7</td>
</tr>
</tbody>
</table>

### TABLE 4
Detection limits (M) for different signal-processing techniques

<table>
<thead>
<tr>
<th>Compound</th>
<th>Wavelet</th>
<th>Peak height</th>
<th>Peak area</th>
</tr>
</thead>
<tbody>
<tr>
<td>Imidazole</td>
<td>3.8 $\times 10^{-4}$</td>
<td>5.0 $\times 10^{-4}$</td>
<td>3.2 $\times 10^{-3}$</td>
</tr>
<tr>
<td>18-Crown-6</td>
<td>4.0 $\times 10^{-4}$</td>
<td>8.5 $\times 10^{-4}$</td>
<td>6.3 $\times 10^{-3}$</td>
</tr>
</tbody>
</table>

### REFERENCES
