Cleaning-up dirty isotachopherograms in time and frequency domain

Available online at www.sciencedirect.com

ScienceDirect

Procedia Chemistry 2 (2010) 67–75

Procedia Chemistry

www.elsevier.com/locate/procedia

5th Conference by Nordic Separation Science Society (NoSSS2009)

Cleaning-up dirty isotachopherograms in time and frequency domain

Jetse C. Reijenga[a*] and Andrus Seiman[a,b]

[a] *Department of Chemical Engineering and Chemistry, Eindhoven University of Technology, P.O.Box 513, 5600 MB Eindhoven, The Netherlands*
[b] *Department of Chemistry, Tallinn University of Technology, Akadeemia tee 15, 12618 Tallinn, Estonia*

Abstract

In chromatography and electrophoresis S/N is increased by time or frequency domain filtering, multiple injections or sample stacking (in case of CZE). Bottom line is noise reduction without peak distortion. Isotachophoresis (ITP) using conductivity detection is also relevant because of its entirely different information content and the importance in CZE stacking. Several time- and frequency domain filters (Fourier & Walsh transform) were compared. Results were:

- Time domain filters can be equally applied to either integral or differential signal;
- Fourier and Walsh transforms are likewise suitable, whereas Walsh better preserves sharp boundaries;
- Disturbance by detector transfer function can only be partly un-done with de-convolution; conditions are critical.

Keywords: Isotachophoresis; signal processing; noise filtering; Fourier transform; Walsh transform; de-convolution; low-pass filter

1. Introduction

Detection systems of most analytical separation equipment have a built-in time constant to reduce baseline noise. This applies equally well to Gas and Liquid Chromatography, Capillary Electrophoresis and Isotachophoresis (ITP). This feature is usually incorporated in the equipment for reason that the user is enabled to clean-up the base line as much as possible, preferably without affecting the information needed. It has to be variable, because broad peaks (wide zone boundaries in ITP) enable a higher degree of noise reduction than narrow peaks (sharp zone boundaries in ITP).

The range of time constants to choose from, and consequently the degree of noise reduction obtained in practice, are limited. Figure 1 illustrates an experimental baseline in a P/ACE 5500 UV detector at 214 nm with time constants between 0.1 and 2.0 s. One can also observe that what remains is always low-frequency noise. Total gain in noise reduction in this typical, but representative example is only a factor 5 at most.

* Corresponding author: J.C. Reijenga. Tel.: +31 40 2473096; fax: +31 40 2473815;
E-mail address: j.c.reijenga@tue.nl

1876-6196/09/$– See front matter
doi:10.1016/j.proche.2009.12.012

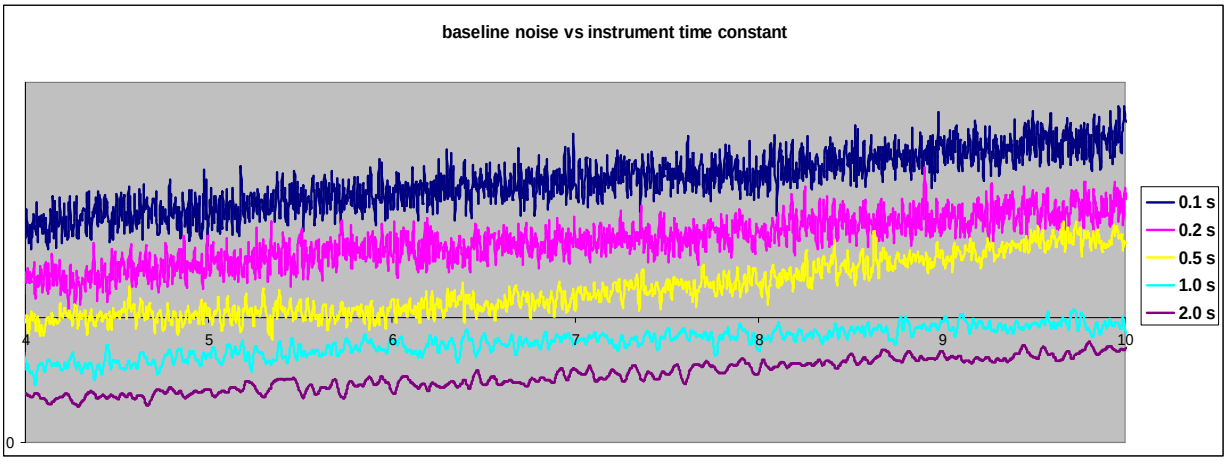

Fig. 1. Experimental baseline noise in P/ACE 5500 CE with UV detection at 214 nm. Horizontal axis in minutes, amplitude is relative

The previous may be assumed to be common knowledge, but there are still manufacturing companies who proudly present a time constant of zero seconds, which is inherently untrue. Optimizing time constant is always time consuming, because one needs to do several runs at different values to check whether the peaks/zones are not affected by the filter. Post-processing of un-filtered data is always preferred, both for reasons of time gain, and for flexibility. If one assumes a time constant of 1.0 s is always a nice compromise, then one never knows whether hopefully sharp zone boundaries or nicely stacked peaks are unaffected. [1, 2, 3, 4]

2. Materials and Methods

CZE baseline experiments were carried out in P/ACE 5500 CE with UV detection at 214 nm, using a background electrolyte of Tris-Borate at pH 8.3, at 10 kV in a 75 µm capillary of 300/360 mm. ITP experiments were carried out by dr. Marak of Comenius University Bratislava in their home-made equipment with conductivity detection. They are traces from the separation of model mixtures in anionic mode with 0.01 M Chloride as leading at pH=3.5 with β-alanine as counter-ion). In addition to Microsoft Excel and Elsevier's CLEOPATRA programs, the time- and frequency domain filtering was carried out in a program written in PowerBasic (PowerBASIC Inc, Venice, Florida, USA), using algorithms as mentioned in the text (program available from the author).

3. ITP vs. CZE & Chromatography

Signal interpretation in ITP is different from CZE and Chromatography, see Fig.2. In ITP, identification relates to signal amplitude (step height), and amount relates to zone length (e.g. by using peak distance in the time differential). The challenge in improving detection limit in ITP is comparable to improving resolution in CZE (reduce peak overlap, in case of ITP in the differential signal). Likewise, the challenge of improving resolution in ITP is comparable with increasing S/N in CZE or chromatography.

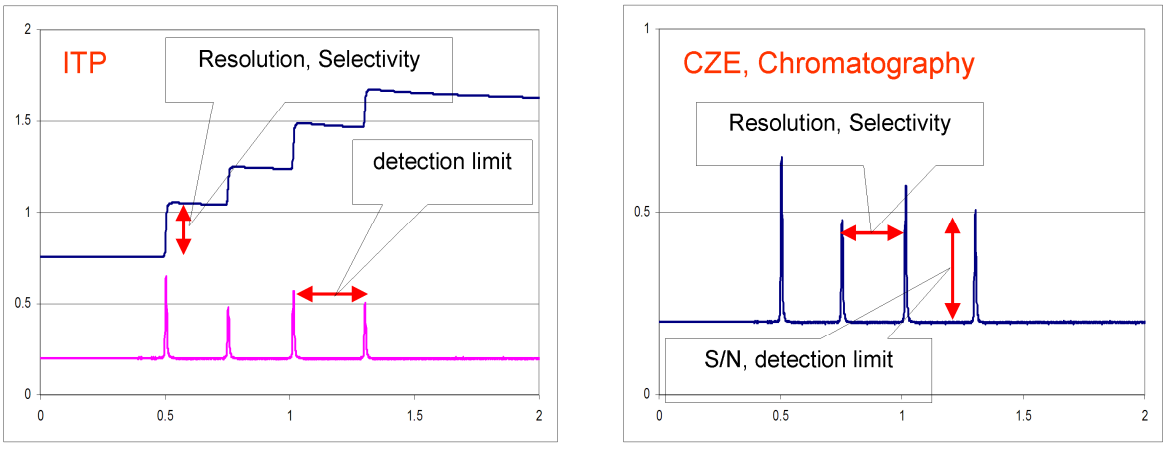

Fig. 2. Concepts of resolution, selectivity and detection limit in ITP detection signals (left) are entirely different from those in CZE and Chromatography (right). Horizontal axis in minutes, amplitude as relative step height (left) or relative absorbance (right).

Another point specific for ITP is the fact that some information is obtained from the stepwise signal, other information is preferably obtained from its time differential: zone lengths, as measured from the inflection points of the stepwise signal. Because differentiate & integrate are simple and fast operations, noise filtering can be done on either of these two as shown below.

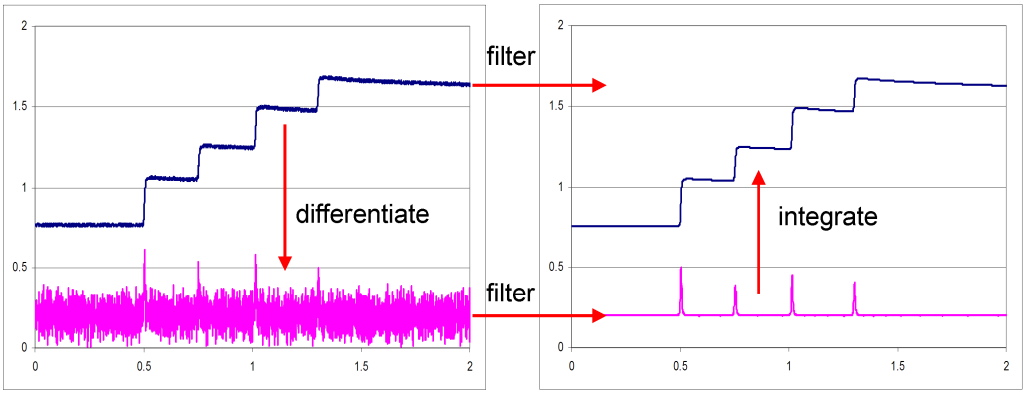

Figure 3. Filtering original signal or its first derivative (followed by integration) yield exactly same result. Axes as in Figure 2 left.

Results as shown in Figure 3 top right are identical. One might argue that the differential signal in this example looks awful and useless, but that is because the differential has been calculated over two adjacent points, so that all noise will amplify. Would we have differentiated from a polynomial fit, then the differential would have been much better, but only because filtering and differentiating would have been combined in one operation, still a valid alternative often used in the pre-computer age.

Another choice to be made is between time domain filters and frequency domain (spectra). The latter of course require transform, filtering and back transform.

4. Time domain filtering

4.1. Principle of operation

These are essentially moving average filters, where each data point is replaced by a number of adjacent points, each of them multiplied by a normalized filter function. By definition, time domain filters are low-pass (high frequencies are reduced) and this may lead to decrease of zone boundary sharpness (or broadening of differential peaks).

4.2. Comparison of filters

Three time domain filters are compared: Savitzky-Golay [2], Exponential and Gaussian [1], see Figure 4.

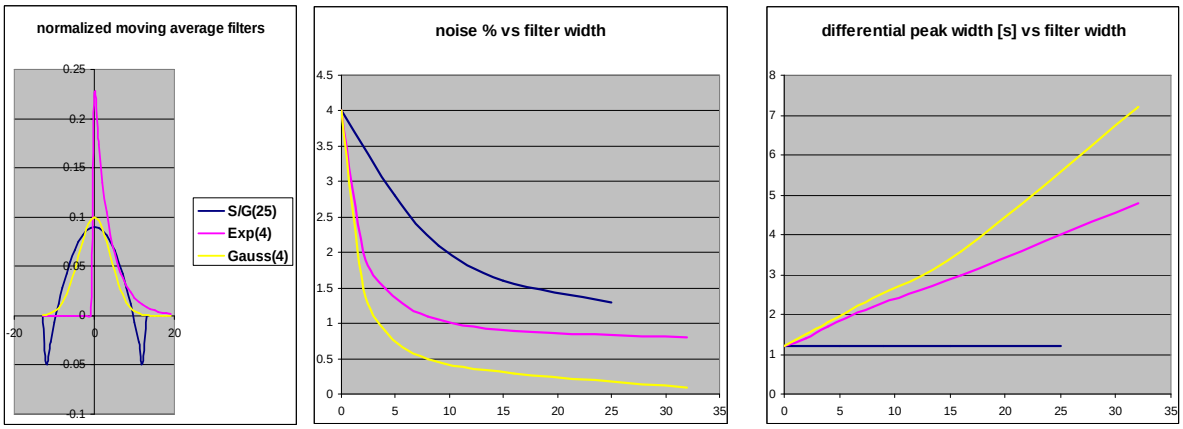

Figure 4. Three different time domain moving average filters Savitzky-Golay, Exponential and Gaussian (normalized function on the left) and their effect on noise level (middle) and differential peak width (right).

Noise levels decrease with increasing filter width, but differential peak width increases, meaning that apparent zone boundary sharpness deteriorates. What is most striking from the above is that S/G filters leave the peak width largely unaffected. Their only draw-back would seem that S/G filter widths exceeding 25 are not available: This problem can be well solved by either lowering the sampling frequency of signal digitization, or by bunching adjacent points (with inherent noise reduction as a free bonus). Just as in chromatography, 20 data points for the narrowest peak are enough; ITP requires also not more than 20 points for the steepest zone boundary. Cation-analyses may require 10-20 Hz because zone boundaries are inherently sharper [5], but for anion analyses in ITP, 5-10 Hz is usually enough.

4.3. Noise filtering and detection limit

Whereas in CZE and LC/GC detection limit follows from plate number and S/N, in ITP it follows from minimum detectable zone length and thus zone boundary sharpness. Tiny zones can of course be masked by both noise and lack of detector resolution (zone-boundary sharpness). Experimental isotachopherograms of a one-second zone were used, to which different levels of artificial white noise were added (Figure 5). At higher noise levels the zone can hardly be distinguished, or quantified. Filtering the worst signal (16% noise) is just possible with a 25 point S/G filter, as long as zone boundary profile is not further affected. The filtered signal can now even be quantified.
This result also illustrates how a definition of detection limit in ITP differs from CZE or Chromatography. In ITP with conductivity detection the most logical definition of (qualitative) limit of detection L_D would be: a zone length so short, that the two peaks in the differential signal can just be distinguished. Logically and as a consequence, the (quantitative) limit of determination L_Q has to be defined such that the differential peaks mentioned are sufficiently separated (e.g. with resolution 1), in order to enable to quantify the zone length from the distance of peak tops.

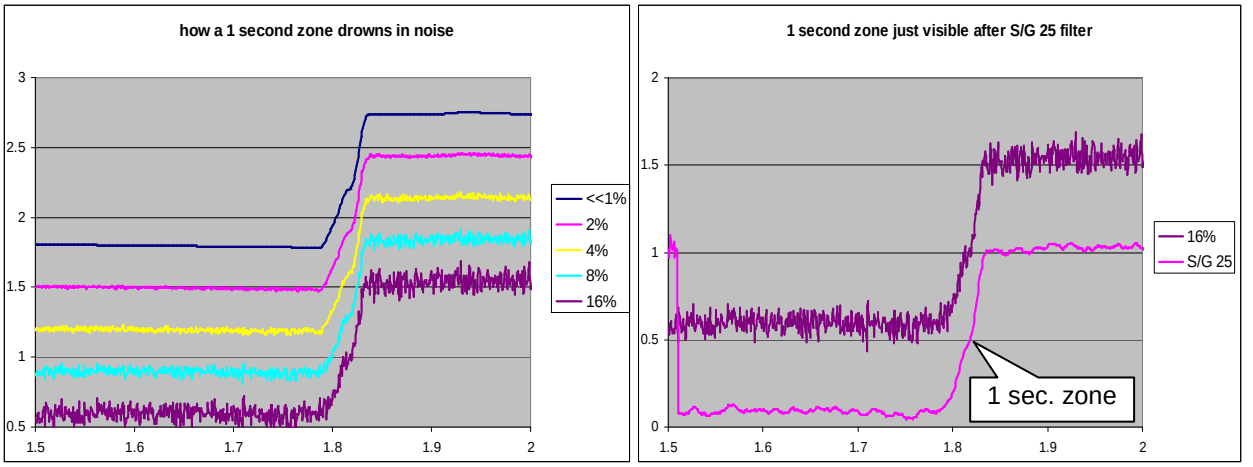

Figure 5. Experimental Isotachopherograms of a one-second zone with different levels of white noise added (left) and removed with a 25 point S/G filter (right). Horizontal axis in minutes, vertical as relative step height.

In the limiting case of hypothetically sharp zone boundaries (Figure 6 left), or in case the sampling frequency is too low, time domain filtering should be used with caution. The example here illustrates distortion by Exponential and even S/G filters (overshoot), a typical case of filter-overkill.

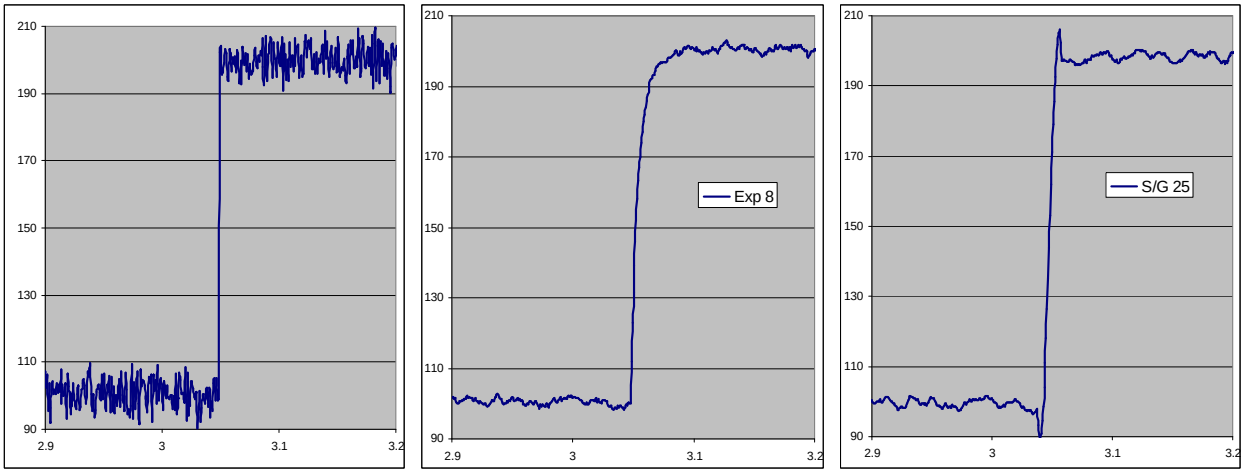

Fig. 6. Infinitely sharp noisy ITP boundary (left) filtered by 8-point Exponential (middle) or 25 point Savitzky-Golay filter (right). Axes as in Figure 5.

5. Frequency domain filtering

5.1. Principle of operation [3]

Different transforms are compared: Fourier (FFT) and Walsh (FWT). FFT consist of series of coefficients of sine functions with increasing frequency, FWT transforms are likewise coefficients of periodic functions that are either 0 or 1 (also known as binary FFT) [4]. After transform, a normalized low-pass filter in frequency domain is applied. This filter function is 1 at low frequencies and 0 at high frequencies and can have different shapes. Both FFT and FWT functions have additive properties: transform of sum of signals equals sum of transform of signals. This is illustrated in Figure 7, a simulation of ITP detection signal with addition of noise.

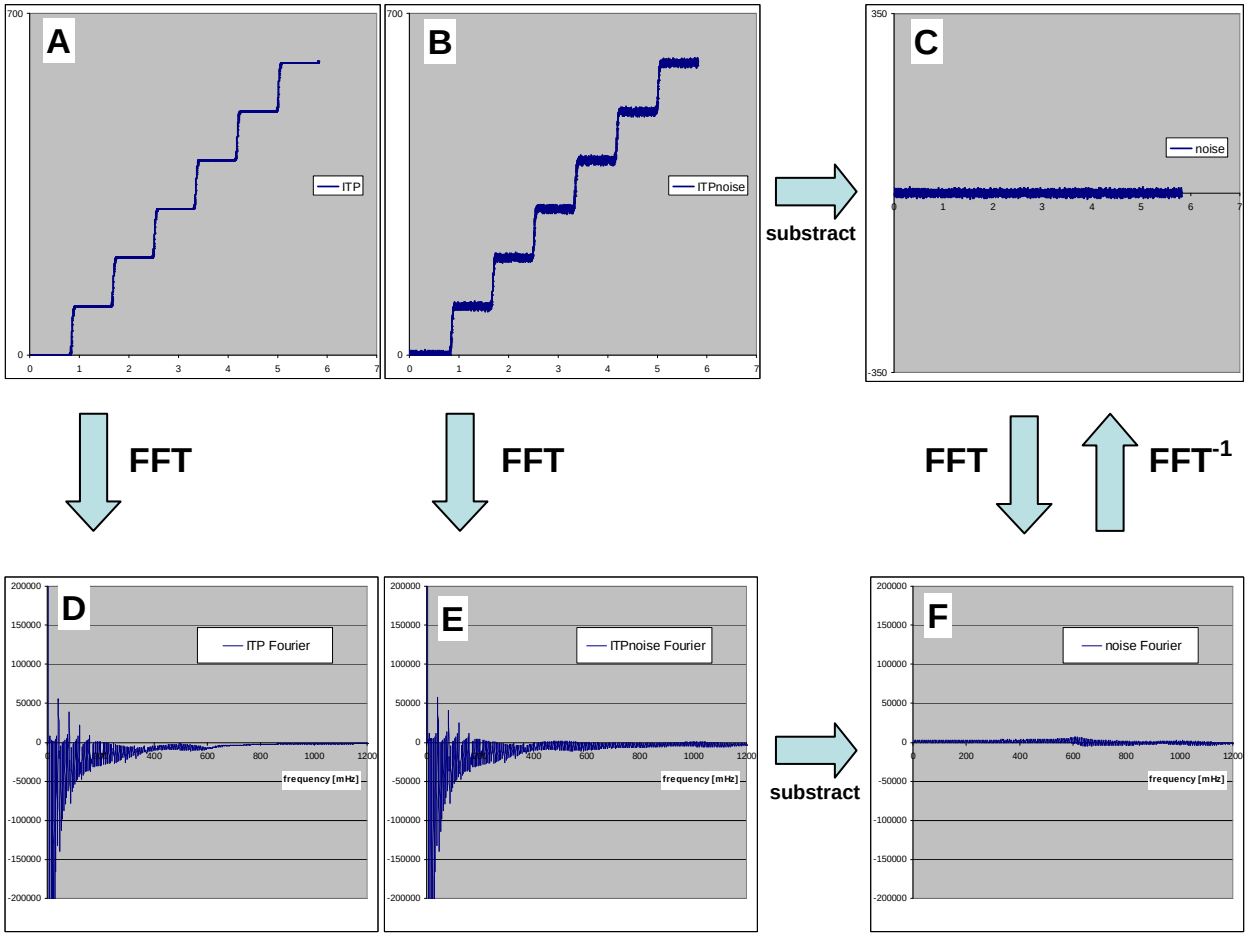

Fig. 7. Additive properties of Fourier Tansform, for the example of noise-less step signal and white noise. A axes of A, B, C as in Figure 5, for D, E, F the frequency axes are in mHz, spectral amplitude is dimensionless.

Noise is added to signal A and B is obtained. A and B are transformed into spectra D and E respectively. Subtracting B - A gives C, the noise level which was added. Subtracting E – D gives F, the spectrum of the noise. Obviously C and F can be converted from one another by the FFT algorithm. In the example given, FFT spectra (D, E) are very similar but subtracting those reveals that noise is evenly spread out over all frequencies (F), whereas peak info concentrates below 600 mHz. Low pass filtering at 600 mHz and inverse FFT would therefore be the method of choice for enhancing S/N in the example given.

In case zone transitions are instantaneous (infinitely sharp), information content of the noisy signal in the frequency domain covers a wider frequency range than just 0-600 mHz. Optimum Fourier filtering as a consequence is less straightforward. This is illustrated in Figure 8, showing (A) an infinitely sharp and noisy zone transition. Fourier filtering with 600 mHz cut-off leads to overshoot-oscillation in time domain (Figure 8 B). The frequency of this oscillation, not surprisingly, corresponds to 600 mHz and the slope at the inflection point is decreased. Fast Walsh transform on the other hand, hardly affects this slope, nor does it introduce oscillations (Figure 8 C). Walsh transfer filtering is thus preferred over FFT filtering for step functions, at least in case of white noise.

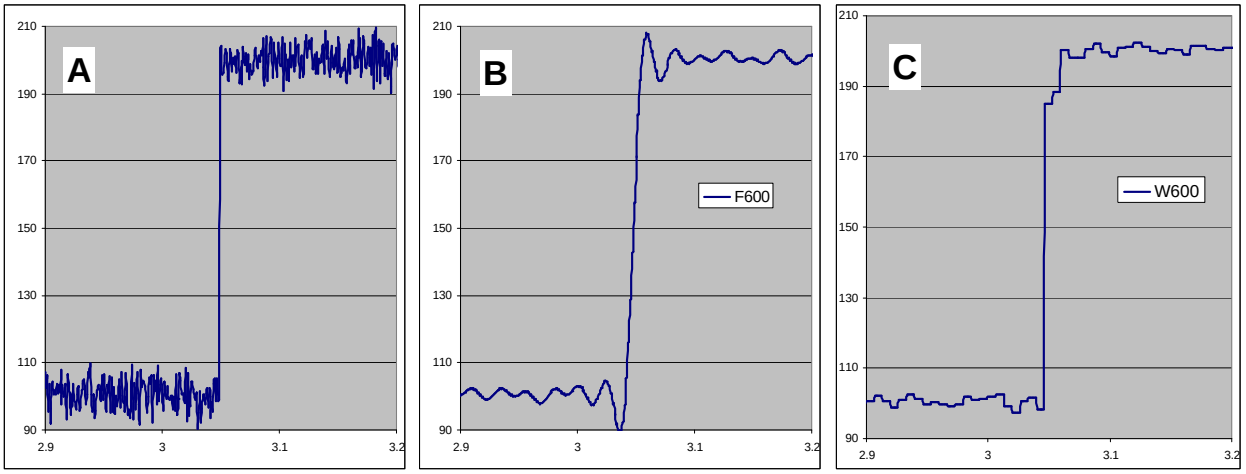

Figure 8. Infinitely sharp noise zone boundary (A) after 600 mHz low-pass filtering in Fourier domain (B) and Walsh domain (C). Axes as in Figure 5.

6. De-convolution

6.1. Principle of operation

Detector signals can be distorted by the detector itself (cell geometry, zone boundary profile, lamp noise etc). Some of these can be modelled as a matrix multiplication in the frequency domain (but not noise, this is additive). The time equivalent of this frequency matrix is called the transfer function or disturb function. For example, in the limiting case of no distortion, the transfer function is a Dirac pulse, which has the value 1 for t=0 and the value 0 at t>0. The frequency spectrum consists of the value 1 for all frequencies.
In case the transfer function is known (or can be estimated), the disturb spectrum is calculated. Then a matrix division of raw detector spectrum and disturb spectrum (and back transform) may improve signal “sharpness”, albeit at the expense of S/N.

6.2. Simulation example [6]

The Cleopatra program [6] used to be commercially available for educational purposes from Elsevier (Amsterdam, The Netherlands) in the 1980’s. It was used to illustrate the principle with a simulation example. Suppose a detected isotachopherogram lacks resolution, in part due to detector transfer function (Figure 9 left), and the suspected transfer function for example is a Gaussian (Figure 9 right), we now apply FFT to both.

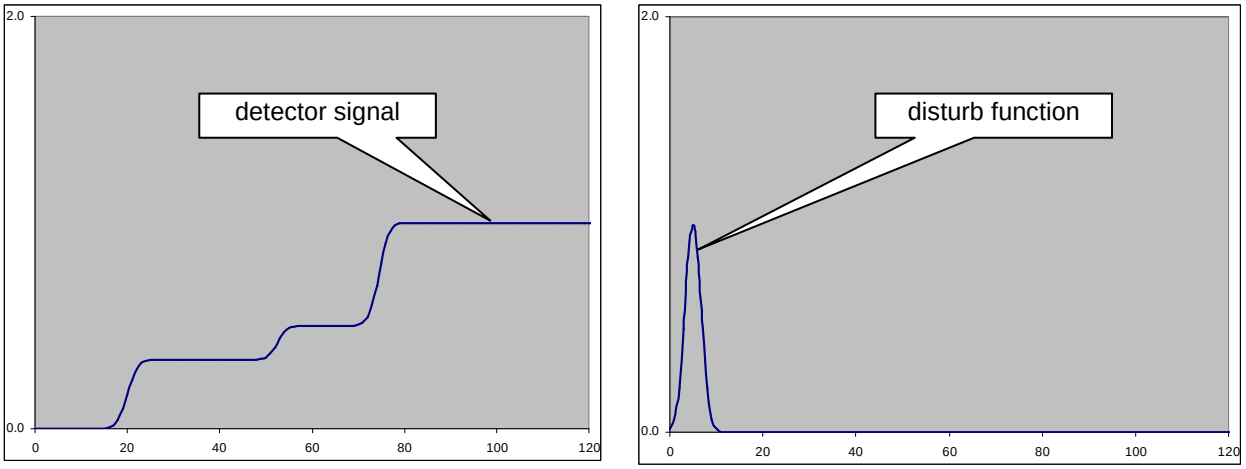

Fig. 9. An isotachopherogram of conductivity detection (left) for which a Gaussian disturb function (right) is assumed. Axes have dimensionless data points.

Resulting spectra are shown in figure 10. Theoretically, the non-disturbed spectrum can be obtained by matrix division.

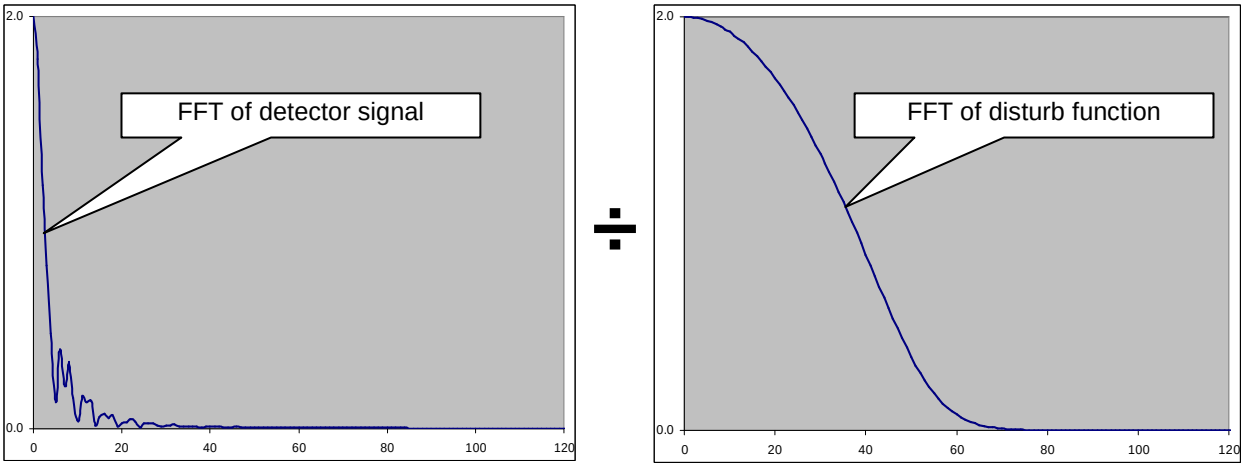

Fig. 10 Fourier spectrum of detector signal (left) and of assumed disturb function (right). Axes have dimensionless data points

Of course, when figure 10 left is divided by the right, higher frequencies (noise!) are amplified, because spectrum values higher than 70 are close to zero. Resulting spectrum is called a de-correlated spectrum, depicted in figure 11, left. Low-pass filtering is therefore necessary prior to back-transform. Filter settings are very critical but some resolution gain can be obtained, as seen in figure 11, right. Zone boundaries are sharper (inflection points are steeper), but the danger of oscillation is still obvious. Obviously the oscillation frequency in this case can be estimated as approximately 70 on the frequency axis of figure, 11 left.

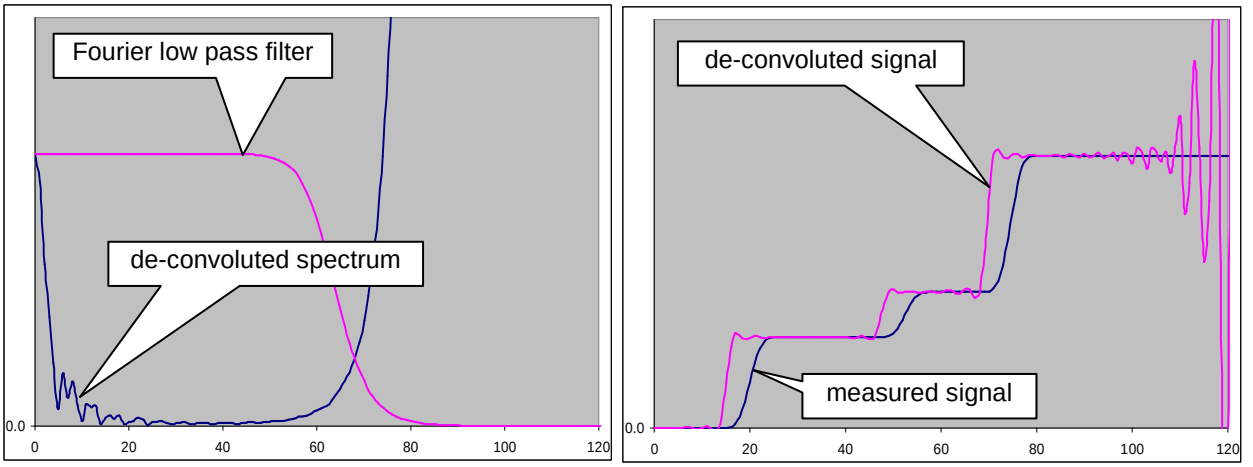

Figure 11. Unfiltered de-convolution spectrum and potential filter function (left) and measured vs improved signal (right). Axes have dimensionless data points

In choosing optimum FFT filter, one hypothetically needs a-priori information about the maximum frequency present in the original signal, excluding noise, and of course estimates of the transfer function. Among possible other signal distortions are curved zone boundaries, detector overshoots, step height drift, even electrode reactions when using contact conductivity etc.

7. Conclusions

Many options for post-run noise filtering are available; all can potentially improve signal to noise ratio in isotachophoresis. Care should however be taken that S/N is not increased at the expense of resolution. The Savitzky-Golay filter is better than other time domain filters. In the frequency domain Walsh is better than Fourier Transform for very sharp zones.
Improving signal quality using de-convolution techniques with presumed detector disturb functions however is tricky, as it always increases noise level. Frequency domain filtering is crucial and filter settings always critical.

Acknowledgement
Clean experimental isotachopherograms were kindly provided by Dr. Jozef Marak of Comenius University, Bratislava.

References

1. D.L. Massart et al, Handbook of Chemometrics and Qualimetrics, Elsevier, Amsterdam,1997.
2. A. Savitzky and M.J.E. Golay, Analytical Chemistry 36 (1964) 1627.
3. W.H. Press et al, Numerical Recipes, 3rd edition Source Code CD-ROM, Cambridge University Press, 1992.
4. E.E. Wallingford and J.C. Culpepper, IEEE 3rd International Computer Software and Applications Conference Proceedings 1979, 824.
5. J.C. Reijenga, Detection and Signal Evaluation in Capillary Isotachophoresis, *thesis*, Eindhoven University of Technology, 1984.
6. G. Kateman et al, CLEOPATRA software, Elsevier, Amsterdam, 1985.