RESOLUTION ENHANCEMENT OF SPECTRA USING DIFFERENTIATION

MARKUS HEGLAND

ABSTRACT. Finding the positions and heights of spectral lines becomes an extremely difficult problem when the lines are close. The most popular methods currently available for this task use nonlinear fitting. For such methods, a particular challenge is the determination of good starting values for the line positions. An alternative method is based on the reformulation of the discrete nonlinear fitting problem as a continuous linear one which does not require any starting values. However, this linear problem turns out to be severely ill-posed, and, in addition, the solution does not have the properties required for the application of the standard convergence theory.

The situation is different if one only solves the linear problem approximately. While this “approximate problem” is still ill-posed, it can be analysed by applying the standard convergence theory. This approximation procedure is called “enhancement”. To a certain extent, enhancement can remedy some of the problems occurring in non-linear fitting by preconditioning the data such that the spectral lines are narrower, and thus better separated. It is demonstrated that this enhancement can be done in a stable way using numerical differentiation. A new technique is proposed which is found to have superior enhancement properties compared to that achieved with the original Taylor-series based method.

1. Introduction

Many spectra in chemistry and physics are well approximated by superpositions of similar “peaks” of roughly known shapes [AGG64]. In the simplest situation, and after scaling the independent variable $x$, they can be described by

$$f(x) = \sum_{j=1}^{m} c_j b(x - x_j)$$

where the shape functions $b(x)$ are the “peaks” and include Gaussians $b(x) = e^{-x^2}$ and Lorentzians $b(x) = 1/(1 + x^2)$ and convolutions of these two. A basic problem of spectroscopy is the determination of the heights $c_i$ and the locations $x_i$ of the peaks from measurements.

Estimates $y^{(i)}$ of $f$ are obtained from measurements, performed at the points $x^{(i)} = ih$ which are assumed to be equidistant. The measurements contain (unknown) errors $\epsilon^{(i)}$ and are modelled as

$$y^{(i)} = f(x^{(i)}) + \epsilon^{(i)}, \quad i = 1, \ldots, n.$$
Note that \(x_j\) denotes the locations of the peaks whereas \(x^{(i)}\) denotes the location of the measurement points. While a variety of types of errors can occur in practice, it will be assumed here that the \(e^{(i)}\) are independent normally distributed random variables.

A well established technique used in spectroscopy uses nonlinear least squares [Mad80, Bjö96] to determine the positions and heights of the peaks. For the case of Lorentzian peaks, the determination of the heights and locations can be done in the Fourier domain using Prony’s method [KMOS92]. If the peaks are well separated, this approach yields reliable estimates. If the peaks are very close, however, it is very difficult to get good starting values, and, in many cases, it is impossible to see, even from visual inspection, how many peaks are present. An example of this situation will be presented in the last section.

From the formula for \(f\), it follows that the spectrum can be represented as the convolution of a kernel with linear combinations of shifted delta distributions \(\delta(x - x_j)\). More specifically, one introduces the distribution

\[
u(x) = \sum_{j=1}^{m} c_j \delta(x - x_j).
\]

This \(\nu\) is an element of the dual \(H^{-1}(\mathbb{R})\) of the Sobolev space \(H^1(\mathbb{R})\). Loosely speaking, the Sobolev space \(H^1(\mathbb{R})\) is the space of functions for which

\[
\|f\|_1 := \sqrt{\int_{\mathbb{R}} f(x)^2 dx + \int_{\mathbb{R}} f'(x)^2 dx}
\]

exists and \(\| \cdot \|_1\) is the Sobolev norm [Ada75]. It is a consequence of the Sobolev embedding theorem that the delta distributions, and, thus, \(\nu\) are elements of \(H^{-1}(\mathbb{R})\). In general, the number \(m\) of peaks is unknown.

Convolutions in \(L_2(\mathbb{R})\) are represented by integral operators of the form

\[
(A\nu)(x) := \int_{\mathbb{R}} b(x - t)\nu(t)dt.
\]

For convenience, the same notation will be used when \(\nu \in H^{-1}(\mathbb{R})\). The relationship between the distribution \(\nu\) and the model function \(f\) is then recast as

\[
Au = f.
\]

It is assumed in the following that this equation has a unique solution. If one is able to compute the functional \(\nu\) from the data, one can then determine the locations \(x_j\) and heights \(c_j\) from scalar products \((\nu, w) = \int w(t)\nu(t)dt\), with appropriate elements \(w\) of \(H^1(\mathbb{R})\).

The basic problem is thereby reduced the reconstruction of \(\nu\) from measured data. Assume that the data is interpolated to obtain an \(f_\epsilon\) such that

(2) \[ f_\epsilon(x^{(i)}) = y^{(i)}, \quad i = 1, \ldots, n, \]

and

(3) \[ \|f_\epsilon - Au\|_1 \leq \epsilon. \]

It can be seen that \(\epsilon\) can be chosen such that \(e^{(i)} = O(\epsilon), \quad i = 1, \ldots, m\). The reconstruction problem can then be reformulated as the problem of finding \(\nu\) satisfying these constraints from a given \(f_\epsilon\). Note that this interpolation approach does
not lose any information, in fact, the measurements \( y^{(i)} \) are recovered by sampling \( f_\epsilon(x) \) at \( x = x^{(i)} \).

The advantage of this interpolation reinterpretation is that the problem of the reconstruction of \( u \) from \( f_\epsilon \) has been widely studied. In particular, a well-known result by Micchelli and Rivlin [MR77] gives bounds for the error of an “optimal reconstruction” of \( u \) from the data \( f_\epsilon \). In the sense of Micchelli and Rivlin, an optimal reconstruction algorithm is defined by a (not necessarily linear) operator \( R_\alpha : H^{-1}(\mathbb{R}) \to H^{-1}(\mathbb{R}) \) such that the worst case error

\[
E_R(\rho, \epsilon) = \sup_{f_\epsilon, v \in H^{-1}(\mathbb{R})} \{ ||v - Rf_\epsilon|| \mid ||Av - f_\epsilon|| \leq \epsilon, \|v\| \leq \rho \}
\]

is minimised for \( R = R_\alpha \). Then, \( u_\epsilon = R_\alpha(y_\epsilon) \) is called an optimal reconstruction. The main theorem of Micchelli/Rivlin provides the following error bound for an optimal reconstruction

\[
e_\epsilon \leq ||u_\epsilon - u||_{-1} \leq 2e_\epsilon
\]

where

\[
e_\epsilon = \sup \{ ||v||_{-1} \mid ||Av|| \leq \epsilon, \|v\|_{-1} \leq \rho \}
\]

and \( \rho = ||u||_{-1} \). For simplicity, \( e_\epsilon \) will be referred to as the reconstruction error. It follows that, if \( e_\epsilon \) is small, any optimal reconstruction is close to the actual solution \( u \). If \( e_\epsilon \) is large, an optimal reconstruction is not guaranteed to be a good approximation. The worst-case optimality of \( R_\alpha \) implies that for any reconstruction \( \tilde{R} \) there are \( g_\epsilon, w \in H^{-1}(\mathbb{R}) \) with \( ||Aw - g_\epsilon|| \leq \epsilon \) and \( ||w|| \leq \rho \) such that \( e_\epsilon \leq ||w - R(g_\epsilon)|| \). If \( e_\epsilon \) is large it follows that \( R(g_\epsilon) \) is not a good approximation of \( u \).

It \( A^{-1} \) is unbounded, then there exists a sequence \( z_i \) with \( ||z_i||_{-1} = 1 \) such that \( \mu_i = ||A^{-1}z_i||_{-1} \) is unbounded. Thus, with \( v_i = \rho/\mu_i A^{-1}z_i \), one obtains a sequence with \( ||v_i||_{-1} = 1 \) and \( ||Av_i||_{-1} = \rho/\mu_i \) such that, for large enough \( i \),

\[
v_i \in \{ \|Av\| \leq \epsilon, \|v\|_{-1} \leq \rho \}
\]

and, as a direct result, \( e_\epsilon \geq ||v_i|| = \rho \). From this it follows that, among other possibilities, the operator \( R_\alpha \) with \( R_\alpha(f_\epsilon) = 0 \) is an optimal reconstruction. Thus, from a practical point of view, the reconstruction problem defined by the conditions (2) may not be tractable.

However, if one has additional information, such as smoothness information about the solution, one might be able to compute a good approximation. Unfortunately, in the case of spectra, one cannot guarantee the required smoothness. What one does know, however, is that \( u(x) \) is a linear combination of a finite number of delta distributions. It will now be shown, that if these delta distributions are not too close (in a sense to be defined), this information can be utilised to produce a convergent reconstruction.

First, one observes that one can stably reconstruct a smooth “approximation” of \( u \). This smooth approximation, termed resolution enhancement, is then used to obtain initial values for a nonlinear fitting procedure. The overall reconstruction procedure will be further developed in future work. In the remaining sections, the enhancement procedure is developed.

The basic resolution enhancement idea is developed in Section 2. In Section 3, the implementation of resolution enhancement using numerical differentiation is presented, and a new, “optimal” method is introduced. Section 3 also provides some
experimental evidence that the method is capable of separating several close peaks. The last section summarises the findings and suggests further work.

It should be noted that other important practical problems of spectroscopy are not discussed here. Such problems include “background subtraction” [PS90] or the determination of the width of the peaks. However, in the case of overlapping peaks, it is virtually impossible to determine the peak width or even the number of peaks [Per68]. Additional domain knowledge is required in order to resolve such difficulties.

2. Resolution Enhancement

Linear, translational-invariant approximations of \( u(x) \) are of the form

\[
\tilde{u}(x) = \sum_{j=1}^{m} c_{j} \tilde{b}(x - x_{j}).
\]

Such approximations allow the determination of \( x_{j} \) and \( c_{j} \) from sampled data if the function \( \tilde{b}(x) \) is “high and narrow” compared to the distances \( |x_{i} - x_{j}| \) and the sampling rate. Typically, regularisation [Tik63] based reconstruction methods for \( u(x) \) produce such approximations. The width of the peaks of regularisation-based methods depends on the size of the regularisation parameter. While this might suggest that one could get a practical reconstruction technique using regularisation, the discussion in the previous section revealed that this approach too is bound to fail in practice for some examples which is mainly due to the fact that one does not know how close the peaks are.

Enhancement techniques capitalise on the simple observation that a finite set of distinct points \( x_{i} \) has a minimal positive distance \( \min_{i \neq j} \{ |x_{i} - x_{j}| \} \). Any approximation of the above form for which the width of \( \tilde{b}(x) \) is much less than \( \min_{i \neq j} \{ |x_{i} - x_{j}| \} \) allows a relatively precise determination of the \( x_{j} \) from appropriately sampled data. In practice, however, a fixed resolution is chosen and some peaks may not be separated.

In order to do quantitative comparisons one needs to define the width of a function \( b(x) \). Here it is suggested to use the second moment to define the (squared) width as

\[
\Delta x^{2} = \frac{\int_{\mathbb{R}} x^{2} b(x)^{2} dx}{\int_{\mathbb{R}} b(x)^{2} dx}.
\]

In Figure 1 four different peaks all with maximal height and width 1 are displayed. It can be argued that maybe this definition puts too much weight on the “tails” as one can easily see that two Lorentzian peaks of width 1 can be better separated than two Gaussian peaks due to the fact that the Lorentzian peaks are more narrow around zero. However, one often compares peaks \( b(x) \) with similar behaviour at infinity, i.e., similar tails, so that this is not a real concern.

Now one can study the broadening effect of any convolution operator \( E \). If \( b_{E} = Eb \) with width \( \Delta_{E}x^{2} \) we introduce the enhancement factor \( \alpha_{E} \) of the operator \( E \) as the square root of the ratio of the squared widths:

\[
\alpha_{E} = \sqrt{\frac{\Delta_{E}x^{2}}{\Delta x^{2}}}
\]
Figure 1. Gaussian, Lorentzian, rectangular and triangular peaks all with maximal height $b(0) = 1$ and squared width $\Delta x^2 = 1$.

A good enhancement $E$ has an enhancement factor $\alpha_E$ close to zero. It can be seen that an enhancement $E$ with a small enhancement factor does increase the resolving power introduced in [AD98] for Lorentzian peaks. (One can also define the inverse of $A$ to have a resolution factor $\alpha_{A^{-1}} = 0$.)

In the best case, the enhancement preserves the shape of the peaks $b(x)$ and only decreases the width. This will be called ideal enhancement $E_\alpha$ defined by

$$(E_\alpha b)(x) = b\left(\frac{x}{\alpha}\right)$$

with $0 < \alpha \leq 1$. As $E_\alpha$ should treat all the “translated peaks” $b(x - x_j)$ in the same way, $E_\alpha$ is linear and translation-invariant. This uniquely determines the operator $E_\alpha$. Moreover, if $A_\alpha : H^{-1}(\mathbb{R}) \to H^{-1}(\mathbb{R})$ is the integral operator defined by

$$(A_\alpha f)(x) = \int_{\mathbb{R}} b\left(\frac{x - t}{\alpha}\right) f(t) dt$$

one gets

$$E_\alpha = A_\alpha A^{-1}.$$ 

The squared width of an ideally enhanced peak is

$$\Delta_\alpha x^2 = \frac{\int_{\mathbb{R}} x^2 b(x/\alpha)^2 dx}{\int_{\mathbb{R}} b(x/\alpha)^2} = \alpha^2 \Delta x^2$$

and from this it follows that

$$\alpha_{E_\alpha} = \alpha.$$ 

Thus in the case of ideal enhancement $E_\alpha$ the enhancement factor equals the dilation imposed on the peaks.

The determination of an ideally enhanced spectrum $E_\alpha f$ from data $f_\varepsilon$ can now be reformulated as a reconstruction problem where one wants to reconstruct the solution of the equation

$$AA_\alpha^{-1}u = f$$

from $f_\varepsilon$ where $\|f - f_\varepsilon\|_1 \leq \varepsilon$. Alternatively, as convolution operators commute, this amounts the computation of $A_\alpha u$ where $u = A^{-1}f$. The operator $A_\alpha$ acts like a mollifier in this case and regularises the problem. In contrast to standard regularisation theory, however, one does not attempt to reconstruct $u$ at all as the $\alpha$ is chosen independently of the measurement error. It will now be seen that even
If it is infeasible to reconstruct \( u \) in many important cases one can still reconstruct \( A_\alpha u \), i.e., the reconstruction error

\[
e = \sup \{ \| A_\alpha v \| : \| Av \| \leq \epsilon, \| f \| \leq \rho \}
\]

approaches zero whenever the data error \( \epsilon \) does because \( \alpha \) is fixed, independent of \( \epsilon \). The following lemma uses standard Fourier transform methods.

**Lemma 1.** If the kernel \( b \) of \( A \) is Gaussian, i.e., \( b(x) = e^{-x^2} \) then

\[
A_\alpha = \alpha 2^{\frac{\alpha - 1}{2}} A^{\alpha^2}.
\]

If \( b \) is Lorentzian, i.e., \( b(x) = \frac{1}{1 + x^2} \) then

\[
A_\alpha = \alpha (2\pi)^{\frac{\alpha - 1}{2}} A^{\alpha^2}.
\]

**Proof.** The Fourier transform of the enhanced kernel is

\[
\hat{b}_\alpha(k) = \int_{\mathbb{R}} b(x/\alpha) e^{-ikx} \, dx = \alpha \hat{b}(\alpha k)
\]

if \( \hat{b}(k) \) is the Fourier transform of the kernel \( b(x) \).

For the Gaussian kernel one gets

\[
\hat{b}_\alpha(k) = \frac{\alpha}{\sqrt{2\pi}} e^{-\alpha^2 k^2/4} = \alpha 2^{\frac{\alpha - 1}{2}} \left( \hat{b}(k) \right)^{\alpha^2}.
\]

As functions of convolution operators are defined by applying the functions to the Fourier transform of the kernel one gets

\[
A_\alpha = \alpha 2^{\frac{\alpha - 1}{2}} A^{\alpha^2}.
\]

In the case of the Lorentzian kernel one has

\[
\hat{b}(k) = \frac{1}{\sqrt{2\pi}} e^{-|k|}. \]

With similar reasoning as before one gets

\[
A_\alpha = \alpha (2\pi)^{\frac{\alpha - 1}{2}} A^{\alpha^2}.
\]

These representations lead to an estimate for the Micchelli/Rivlin reconstruction error \( e \):

**Proposition 1.** In the case of Gaussian peaks one has

\[
e = \sup \{ \| A_\alpha v \| : \| Av \| \leq \delta, \| v \| \leq \rho \} = \alpha 2^{\frac{\alpha - 1}{2}} \delta^{\alpha^2} \rho^1 - \alpha^2.
\]

In the case of Lorentzian peaks one has

\[
e = \sup \{ \| A_\alpha v \| : \| Av \| \leq \delta, \| v \| \leq \rho \} = \alpha (2\pi)^{\frac{\alpha - 1}{2}} \delta^{\alpha} \rho^1 - \alpha.
\]

**Proof.** This is a consequence of the previous lemma and the Hilbert scales interpolation theory [KP66] which states that

\[
\| A^\beta f \| \leq \| Af \|^{\beta} \cdot \| f \|^{1 - \beta}
\]

and that this bound is sharp in the sense that there is a sequence of elements \( f \) for which equality holds in the limit.

In the case of Gaussians one chooses \( \beta = \alpha^2 \) and in the case of Lorentzian peaks one sets \( \beta = \alpha \).
It was mentioned earlier that for a given width Lorentzians are easier to separate as they have “locally sharper peaks”. In addition, a consequence of the previous proposition is that the Lorentzians can also be better enhanced than the Gaussians as \( e \) goes faster to zero with \( \alpha \). In principle, however, one can see that enhancement factors \( \alpha < 0.5 \) are very hard to obtain if small but noticeable errors are present. While the above result only holds for ideal enhancements it seems plausible that similar observations could be made for other classes of enhancements, however, this has not been proven yet.

Often peaks look similar like Gaussians or Lorentzians and one would expect that the ideal enhancement would also behave in a similar way. The following proposition generalises the previous one and provides a tool to deal with other \( b(x) \).

**Proposition 2.** *If there exists a function \( \psi \) with convex inverse such that \( A^*A \leq \psi(A^*A) \) which means \( \alpha |\hat{b}(\alpha k)|^2 \leq \psi(|\hat{b}(k)|^2) \), for \( k \in \mathbb{R} \), then the enhancement error is bounded by \( e = \sup \{ \| A_\alpha v \| \mid \| Av \| \leq \delta, \| v \| \leq \rho \} \leq \rho \sqrt{\psi(\delta^2/\rho^2)} \).

**Proof.** This is a consequence of the variable Hilbert scales inequality, see [Heg92, Heg95] as from the convexity of \( \psi^{-1} \) one gets for any \( f \):
\[
\| A_\alpha f \|^2 \leq \| f \|^2 \psi(\| Af \|^2/\| f \|^2).
\]

This inequality applied to Lorentzian and Gaussian peaks again confirms the findings of Proposition 1, but, in addition, one can derive similar results for other peaks like convolutions of Lorentzians and Gaussians.

### 3. The Application of Numerical Differentiation

Ideal enhancement was shown to be a feasible method in the previous section. However, it still involves the solution of a usually severely ill-posed problem and implementations of ideal enhancement requires the application of a substantial amount of regularisation in order to prevent the “blowing up” of the measurement errors.

In 1964, almost at the same time as Tikhonov [Tik63] published his ground-breaking work on regularisation, Allen, Gladney and Glarum [AGG64] suggested to use differentiation for enhancement. They developed their method in the Fourier domain where differentiation corresponds to multiplication with the independent variable. As the operator \( A \) is a convolution, it is a multiplication operator in the Fourier domain and the Fourier transform of the solution of \( Au = f \) is
\[
\hat{u}(k) = \frac{\hat{f}(k)}{\hat{b}(k)}
\]
where the \( \hat{\cdot} \) denotes the Fourier transform of any distribution \( v \). If the peak \( b(x) \) is symmetric and has a smooth Fourier transform \( \hat{b}(k) \) Allen et al [AGG64] suggest to
approximate $\hat{b}(k)^{-1}$ with a polynomial in $k$:

$$\frac{1}{\hat{b}(k)} \approx \sum_{j=0}^{q} \beta_j k^{2j}.$$ 

The conditions hold for a Gaussian, the Lorentzian for which $\hat{b}$ has a corner at 0 is treated slightly differently. From this one directly obtains an approximation for $u = A^{-1}f$:

$$u(x) \approx \sum_{j=0}^{q} \beta_j \frac{d^{2j}f(x)}{dx^{2j}}.$$ 

An advantage of this approach is that the severely ill-posed problem of inverting $A$ has now been replaced by the “less severely” ill-posed problem of computing derivatives. As the data $f(x)$ can be assumed to be smooth, the computation of these derivatives from measured data is feasible using, and can be stably computed using, e.g., Tikhonov regularisation. However, this method cannot be guaranteed to provide a good approximation of the spectrum $u(x)$ in the $H^{-1}(\mathbb{R})$ sense as follows from the discussion in the previous section.

But this is not what is of interest here, as all one attempts to compute is a function which has narrow peaks in the same locations as $f(x)$. Indeed, one can see in Figure 2 how the application of the method by Allen/Gladney and Glarum does reduce the width of a Gaussian peak. Notice in this figure that the peak narrowing comes at the cost of negative “side-lobes”, especially when higher derivatives are
Figure 3. Enhancement factors for the Allen/Gladney/Glarum method based on Gaussian peaks applied both to Gaussian and Lorentzian peaks.

used. This is maybe the largest change in the shape of the peak and may be considered counter-intuitive as “spectra are positive”. For many practical applications, however, one deals with second derivatives of the original spectrum which can have negative values. Also, one should keep in mind, that what one attempts to do here is not the reconstruction of the spectrum (preserving positivity) but mapping the spectrum to a function from which the locations $x_j$ and heights $c_j$ of the peaks can be determined. This function could in principle have any form whatsoever as long as it provides information about the locations and heights of the peaks. Especially when using higher order derivatives one should remember, however, that the resolution of this method is limited and the narrowing procedure does produce artifacts. However, as this method cannot separate peaks which are close together in the first place this is not a major concern in practice.

The effect of “approximating” the solution $A^{-1}$ by a linear combination of derivatives is very similar to the ideal enhancement. This enhancement interpretation shall now be further investigated. Consider the case of Gaussian peaks with $b(x) = e^{-x^2}$. The coefficients of the Allen/Gladney/Glarum method are then $\beta_j = (-1)^j \frac{1}{j!}$. Using these coefficients one gets for the peaks of the enhanced spectrum:

$$b_E(x) = \sum_{s=0}^{q} (-1)^s \frac{1}{4^s j!} b^{(2j)}(x).$$

Note that this particular enhancement formula is well-defined whenever $b(x)$ is $2q$ times differentiable and is not limited to Gaussian peaks. In particular, the same formula has been applied to Lorentzian peaks. The resulting enhancement both for Gaussian and Lorentzian peaks is displayed in Figure 3. It is seen that higher than 10th order derivatives are required in order to obtain enhancements of about 0.5 for Gaussian peaks. For Lorentzian peaks one actually gets enhancement of around $\alpha_E = 0.2$ using 10th order derivatives.

One can now ask if this limited enhancement power is due to the choice of the coefficients $\beta_j$ or if it is inherent in the enhancement method using derivatives. For
general derivative-based enhancement one has

$$b_E(x) = \sum_{j=0}^{q} \beta_j b^{(2j)}(x)$$

and the enhancement factor in the general case is

$$\alpha_E = \frac{\sum_{i,j=0}^{q} \beta_i \beta_j \int_{\mathbb{R}} x^{2i} b^{(2i)}(x) b^{(2j)}(x) dx}{\sum_{i,j=0}^{q} \beta_i \beta_j \int_{\mathbb{R}} b^{(2i)}(x) b^{(2j)}(x) dx} = \frac{\beta^T B_X \beta}{\beta^T B \beta}.$$  

This is just the Rayleigh quotient of $B_X$ with respect to the scalar product defined by $B$. Thus the best enhancement possible is just the square root of the smallest eigenvalue of the generalised eigenvalue problem

$$B_X \beta = \alpha^2 B \beta.$$  

The coefficient vector $\beta$ is the eigenvector corresponding to the smallest eigenvalue. With this one gets best possible enhancement for the method using even derivatives up to order $2q$. The computation of the optimal enhancements and the coefficients of the optimal method can be seen to relate closely to Hermite polynomials in the case of Gaussian peaks.

In Figure 4 the enhancement factors of both this “optimal” enhancement and the Allen/Gladney/Glarum method are displayed. One sees that at most about a 30 percent improvement can be achieved. Even this is of interest, in particular as the curve is very flat and one can achieve a fixed enhancement, say, of $\alpha_E = 1/2$ with lower order derivatives. This is of interest as lower order derivatives are more precisely reconstructed from measured data.

The reconstruction error of all the derivative-based reconstructions mainly originates in the reconstruction error of the highest derivative which is

$$e = \sup\{\|D^m Af\| \mid \|Af\| \leq \epsilon, \|f\| \leq \rho\}.$$  

In many cases this error can again be computed using variable Hilbert scales. In particular, for Gaussian and Lorentzian peaks one has
Proposition 3. Let \( \psi(s) > 0 \) be an increasing function for \( s > 0 \) with convex inverse where

1. For Gaussian peaks \( \psi(s) \geq |2\log(2s)|^m s \) for \( s > 0 \) or
2. for Lorentzian peaks \( \psi(s) \geq (\log(\sqrt{2\pi s}))^{2m} s \).

Then the reconstruction error of the derivative \( D^m A \) from inexact data \( f_\epsilon \) is bounded by

\[
e \leq \rho \sqrt{\psi(\epsilon^2/\rho^2)}.
\]

Proof. Only the case of Gaussian peaks is considered here, Lorentzian peaks can be treated in a similar way. For Gaussian peaks the condition on \( \psi \) is

\[
|2\log(2s)|^m s \leq \psi(s).
\]

Now substitute \( s := |\hat{b}(k)|^2 = \frac{1}{s}e^{-k^2/2} \) to get

\[
k^{2m} |\hat{b}(k)|^2 \leq \psi(|\hat{b}(k)|^2),
\]

which, by applying the inverse Fourier transform yields the operator inequality

\[
(D^m A)^*(D^m A) \leq \psi(A^* A),
\]

meaning, in the terms of variable Sobolev spaces

\[
\|D^m A u\|^2 \leq (u, \psi(A^* A) u) = \|u\|^2_{\psi}
\]

which is well-defined for all \( u \in H_\psi \), the variable Hilbert scale defined in [Heg95].

From this it follows that

\[
e \leq \sup\{\|u\|_{\psi} \ | \ |Au| \leq \epsilon, \|u\| \leq \rho\}
\]

and the application of the variable Hilbert scales theory provides the desired bound. \( \square \)

For convergence the behaviour close to \( s = 0 \) is important. It follows that the derivatives of both Gaussian and Lorentzian spectra can be reconstructed with an error “almost” proportional to the measurement error. As the Gaussian is “smoother” than the Lorentzian it comes as no surprise that the derivatives can be computed with less error (asymptotically) for Gaussian peaks.

The numerical implementation requires a stabilised method to numerically compute derivatives. Such a method has been suggested by Anderssen/de Hoog/Hegland [AdHH96] and is based on sampling and averaging. Essentially, it is a finite difference approximation. Other approaches to numerical differentiation are reviewed in [WJ95]. The sampled central difference for second derivatives used by Anderssen et al is defined as

\[
(D_{sh} y)_i = \frac{(y_{i-s} - 2y_i + y_{i+s})}{s^2h^2}
\]

with the averaging defined as

\[
(M_s y)_i = \frac{1}{s} \sum_{j=-(s-1)/2}^{(s-1)/2} y_{i+j}.
\]

Then the finite difference approximation enhancement operator is

\[
y_E = M_s \sum_{j=0}^{q} \beta_j D^j_{sh} y.
\]
In Figure 5 one sees how the Allen/Gladney/Glarum enhancement using finite difference approximations for the derivatives enhances a simulated spectrum. The averaging was chosen with $s = 33$ and the highest derivative in the enhancement used is 10. Note that the data consists of only one “large peak” $f(x)$ with around 1 percent measurement error. The dashed lines are the 5 underlying peaks $c_j b(x - x_j)$ which summed up give the “large peak”. This is a simulation and the data has been chosen such that it is impossible even to guess how many peaks are hidden in $f(x)$. After enhancement one notices that the line has dropped a lot and the 5 peaks are starting to become visible.

One could now attempt to use the optimal coefficients $\beta_j$ defined earlier. Experiments showed that this gives some improvement, however, the averaging and sampling procedure does interfere with the improved enhancement and so the improvement is less than one might expect. A better approach results if the coefficients $\beta_j$ are chosen such that the finite difference approximation including the averaging and sampling optimises the enhancement factor.

In order to define this new method one requires some notation. First let $c^j = (c^j_1, \ldots, c^j_n)^T$ be the vector of the approximate $k$-th derivative $b^{(k)}(x)$ at the measurement points:

$$c^j := M_s D^j_{sh} b$$

where $b := (b(x(1)), \ldots, b(x(n)))$ denotes the vector of values of $b(x)$ at the measurement points. The enhancement factor is then approximated by the term

$$\alpha_{E,h} = h^2 \frac{\sum_{i,j=0}^q \beta_i \beta_j (c^i)^T K c^j}{\sum_{i,j=0}^q \beta_i \beta_j (c^i)^T c^j} \quad = \quad \frac{\beta^T \tilde{B} X \beta}{\beta^T \beta}$$

where the matrix $K$ is diagonal with elements 1, 4, 9, \ldots, $n^2$ on the diagonal.

The smallest enhancement factor is again the square root of the smallest eigenvalue of a generalise eigenvalue problem, this time of

$$\tilde{B} X \beta = \alpha_{E,h}^2 \tilde{B} \beta$$

where the elements of $\tilde{B}$ are $(c^i)^T c^j$ and the elements of $\tilde{B} X$ are $(c^i)^T K c^j$. The coefficients $\beta_j$ of the procedure leading to this optimal enhancement are the components
of the eigenvector $\beta$ to the smallest eigenvalue. This method has been implemented in MATLAB. The results of this enhancement are displayed in Figure 6 for the same data as was used for the Allen/Gladney/Glarum method. Note that now the five peaks become clearly visible.

4. Conclusion

It was seen that the reconstruction of spectra in the sense that the errors of the reconstruction tend to zero if the data errors do is infeasible. However, this is not necessary and enhancements with fixed errors can be reconstructed. In particular, methods based on differentiation have been further studied. It is seen that high order derivatives are required if any substantial enhancement is to be achieved. The method by Allen/Gladney/Glarum gives good results and even better performance is obtained with a new method where the coefficients are chosen such that best possible enhancement is obtained. The numerical differentiation is done with a method based on sampling and averaging suggested by Anderssen, deHoog and Hegland in [AdHH96].

The aim of this work was to show that enhancement is a feasible alternative to reconstructing for the analysis of spectra. A new method is proposed which improves the enhancement capability of earlier enhancement techniques based on numerical differentiation. Future work will investigate further properties of this “optimal method”, in particular, aspects relating to the best possible enhancement and the effect of numerical approximation errors and data errors on enhancement are to be discussed. Further aspects to be analysed is the choice of the width of the peaks.

A MATLAB demonstrator code for the enhancement procedure can be obtained from the author.

Acknowledgements

I would like to thank R.S. Anderssen, CSIRO, for his ongoing support and numerous fruitful discussions and Ryan Bettis and Andrew Torda, ANU, for providing me with background information on spectroscopy. Finally, I thank Berwin Turlach,
ANU, for providing me with the reference to the unpublished report [Rit97] which contains a discussion of Baysian analysis of spectra.

References


