The use of prior information for extracting the post-edge background

K. V. Klementev

Moscow State Engineering Physics Institute, 115409 Kashirskoe sh. 31, Moscow, Russia. E-mail: klmn@htsc.mephi.ru

A new method for extracting the post-edge background μ_0 is proposed, the method of Bayesian smoothing. A further evolution of the smoothing spline method is considered as well. Both techniques are capable to take into account prior information about the peculiarities on the μ_0. In addition, since the Bayesian approach works in terms of the posterior probability density functions, it contains a natural way to determine the errors of the μ_0 construction, which has always been an unresolvable problem for any other method. Even with use of the prior information, which narrows the posterior probabilities, the errors of μ_0 are shown to be larger than the experimental noise.

Keywords: post-edge background; **Bayesian methods.**

1. Introduction

The most difficult procedure in extracting EXAFS from the measured absorption is the construction of μ_0, since one cannot definitely distinguish the environmental-born part of absorption from the atomic-like one. All methods for determination of the post-edge background are based on the assumption of its smoothness, and the only criterion for its validity is the absence of low-frequency structure in EXAFS $\chi(k) \cdot k^w$, i.e. the small absolute value of the Fourier transform (FT) at low r. In the present paper we consider a further evolution of the smoothing spline method and propose a new method of Bayesian smoothing.

2. Smoothing spline

Owing to the fast algorithm and easy program realization, the approximation of μ_0 by a smoothing spline has become widespread. Let $N+1$ experimental values of μ_i are defined on the mesh E_i (extrapolated pre-edge background already subtracted). The smoothing spline μ_0 minimizes the functional

$$J(\mu_0, \mu) = \int_{E_{\min}}^{E_{\max}} [\mu_0'']^2 \, dE + \frac{1}{\alpha} \sum_{i=0}^{N} [\mu_{0i} - \mu_i]^2. \qquad (1)$$

The smoothing parameter (or regularizer) α is a measure of compromise between the smoothness of μ_0 and its deviation from μ. The optimal regularizer should lead to μ_0 containing only low-frequency oscillations and, hence, to χ containing only structural oscillations. The formulation of a new criterion for the optimal α was considered by Klementev (2000*b*).

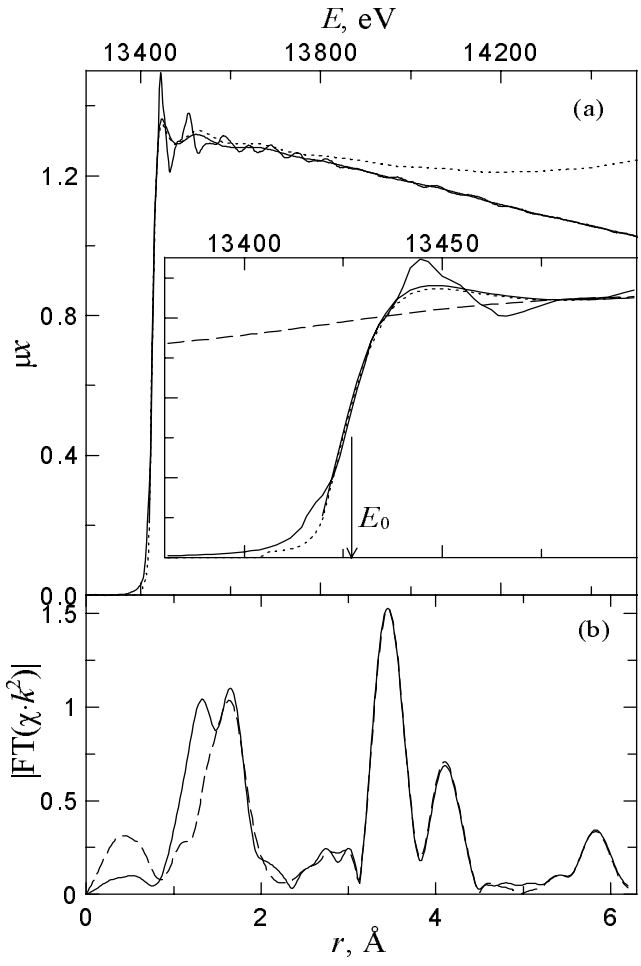

Figure 1
On the smoothing spline method. (a) Absorption coefficient μ, prior function $p(E)$, calculated by FEFF8 (dots) and various post-edge backgrounds, (b) Module of FT of EXAFS functions obtained with these backgrounds. Solid lines — $\mu_0(E)$ and FT obtained with use of the prior function; dashed lines — dittos without prior function. The regularizer α is the same for both cases.

Let us assume that $\mu_0(E)$ is approximately known in advance. Denote this prior function as $p(E)$. Now we will tend the second derivative of the sought $\mu_0(E)$ not to zero (at the specified deviation of μ_0 from μ) but to the second derivative of $p(E)$. The sought $\mu_0(E)$ is now minimizes the functional

$$J^*(\mu_0, \mu) = \int_{E_{\min}}^{E_{\max}} [\mu_0'' - p'']^2 \, dE + \frac{1}{\alpha} \sum_{i=0}^{N+1} [\mu_{0i} - \mu_i]^2. \qquad (2)$$

As seen, in fact there is no need to know $p(E)$ itself, its second derivative is sufficient. Introducing $\tilde{\mu}_{0i} = \mu_{0i} - p_i$, one obtains:

$$\begin{aligned} J^*(\mu_0, \mu) &= \int_{E_{\min}}^{E_{\max}} [\tilde{\mu}_0'']^2 \, dE + \frac{1}{\alpha} \sum_{i=0}^{N+1} [\tilde{\mu}_{0i} - (\mu_i - p_i)]^2 \\ &= J(\tilde{\mu}_0, \mu_i - p_i). \end{aligned} \qquad (3)$$

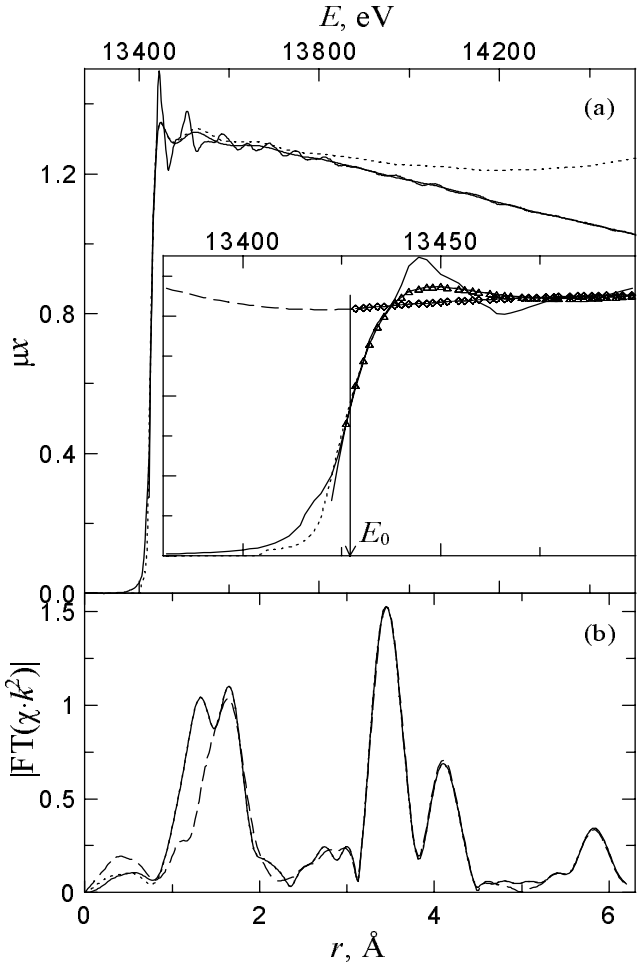

Figure 2
Extraction of EXAFS from the measured absorption using the Bayesian smoothing. Prior function $p(E)$ for the atomic-like absorption is drawn on (a) by dots. Solid lines — $\mu_0(E)$ and FT obtained with the use of the prior function; dashed lines — dittos without prior function. The dot line on (b), seen in the low-r region only, is obtained without additional requirement for $\mu_0(E)$ to pass through a point immediately before E_0. The regularizer α is the same for all cases and equals to the optimal one found for the smoothing spline.

Thus, the problem is reduced to the preceding one in which instead of initial data μ_i the difference $\mu_i - p_i$ is used. The sought μ_0 is found from the smoothed $\tilde{\mu}_0$ as $\mu_{0i} = \tilde{\mu}_{0i} + p_i$. Fig. 1 shows an example [1] of the atomic-like absorption approximation by the smoothing spline with and without the use of the prior function which was calculated by FEFF8 program (Ankudinov *et al.*, 1998) using self-consistent potential calculations with the ground state exchange correlation potential and the simple cubic perovskite structure. The calculated background was then multiplied by a constant factor to match the step height. Additional features due to possible multielectron excitation channels should be added to the $p(E)$, if known. However, we do not intend to discuss here the validity of the prior function, we just demonstrate how the prior knowledge can be used for EXAFS post-edge background removal. As seen, we cannot subtract the calculated background directly, but we can incorporate this function into the EXAFS analysis by the procedure described above. Also seen that the use of $p(E)$ has led to disappearance of the spurious peak at $r < 1$ Å and to considerable correction of the first shell signal.

3. Bayesian smooth curve

The method of Bayesian smoothing is ideologically similar to the smoothing spline method, but has one global advantage: Within no other existent approach one can define and determine the errors of the μ_0 construction. Since the Bayesian smoothing method finds the posterior distributions for all μ_{0i}, one can find not only average values but also standard deviations and any desirable moments. In addition, within the framework of the method it is possible (before the post-edge background construction) to deconvolve μ with the monochromator resolution curve. The weakness of the method is its low speed. On a modern PC the curve drawn through $N \sim 500$ points is smoothed for a few minutes.

Detailed formalism of the Bayesian smoothing, though quite simple, would take several pages, so we cannot give explicit expressions here, but refer to the e-print (Klementev, 2000*b*). To determine $\overline{\mu}_{0i}$ and $\delta\mu_{0i}$ one should find eigenvalues and eigenvectors of a special five-diagonal square $N \times N$ matrix. In Fig. 2 the Bayesian smoothing was done on the mesh of $N = 536$ experimental points above absorption edge, with and without the prior function described in the preceding section. In addition, we demanded from the Bayesian curve in Fig. 2 to pass through a point nearest at left to E_0 (for this, the five-diagonal matrix should be slightly changed, again see e-print (Klementev, 2000*b*)); this requirement practically does not affect μ_0, but affects $\delta\mu_{0i}$ in the low-k region (see below). The regularizer was equalized to the optimal value found for the method of smoothing spline. As seen, we have obtained completely the same EXAFS function as that given by the previous method. So, the only thing that warrants such a slow method is that it gives the errors of the μ_0 construction.

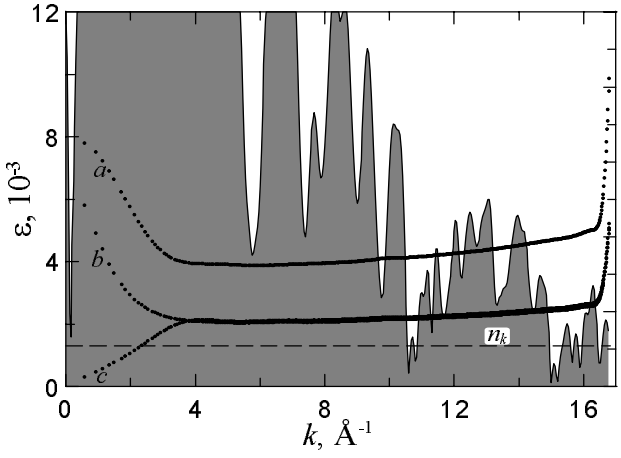

Figure 3
Errors of $\chi(k)$ extraction by the method of Bayesian smoothing without (*a*) and with (*b* and *c*) prior information specifying the second derivative. The curve *c* uses the additional information that $\mu_0(E)$ passes through a point immediately before E_0. Solid line with the filling — the envelope of $\chi(k)$ (not weighted). Dashed line — the noise level estimated from FT.

[1] Here for examples is used the spectrum at the Bi L_3 absorption edge for $Ba_{0.6}K_{0.4}BiO_3$ at 50 K recorded in transmission mode at D-21 line (XAS-13) of DCI (LURE, Orsay, France) using the Si(311) monochromator. Energy step $\sim$ 1 eV, counting time 1 s.

4. Errors of μ_0 construction and noise

For the correct subsequent fitting of the EXAFS signal, one should determine, ideally, all sources of errors. While some of them are quite transparent, others even cannot be defined, for instance, the errors of μ_0 construction. One may argue that the errors of μ_0 are not as important since are of low frequency and presumably are outside (in r-space) the region of EXAFS analysis. This may be true only if one performs fitting of the $\chi(r)$ or filtered $\chi(k)$. But many researchers try to fit the entire $\chi(k)$ function, together with possibly undetermined slow varying background. This is the first reason why we have to know the errors of μ_0. The second, and the main, reason is that the μ_0 drawn is not the atomic background itself but an artificial representation of it. Therefore, even if our μ_0 is of zero frequency, considerable spectral range (especially the signal from the first shell) may be distorted by this representation. Speaking of the errors of μ_0 construction, we mean the deviation of our artificial μ_0 from the real background. The real background must be *nearly* a middle line of the oscillating part of μ with the second derivative which is either *nearly* small or *nearly* close to the second derivative of the prior function. The Bayesian analysis gives errors $\delta\mu_{0i}$ within which lie *all* such functions, and, among them, the real post-edge background. Thus, the errors $\delta\mu_{0i}$ are very important because they account for possible deviation from the true atomic background.

Having determined $\bar{\mu}_{0i}$ and $\delta\mu_{0i}$ (this is done simultaneously via $N \times N$ matrix inversion), we express the errors of $\chi(k)$ extraction as $\varepsilon_i = \delta\mu_{0i}/\bar{\mu}_{0i}$. For our example spectrum they are shown in Fig. 3 by dots. As seen, the introduction of the prior information has significantly diminished the errors ε_i. This is quite natural, since any decrease of our ignorance about μ_0 should narrow the posterior distribution of μ_{0i} for all i. Of course, this concerns the experimental information as well: ε_i are the smaller (among equal-length spectra) the more measured points N the spectrum has.

Near the hanging ends of the spectrum, the errors of μ_0 construction are, as expected, significantly larger than in the inner area. The ends should be deleted after $\chi(k)$ extraction or treated in a different way: we can take into account another information that the atomic-like absorption must coincide with the total absorption (minus pre-edge background) at energies $E < E_0$. As seen (curve c in Fig. 3, in this case the left spectrum end is saved for further analysis.

It is quite reasonable to demand that the errors of μ_0 construction were less than the EXAFS signal (the envelope of $\chi(k)$ in Fig. 3). For the Bayesian curve a this range is $0 \leq k <\sim 14\,\text{Å}^{-1}$, for the Bayesian curves b and c this range is wider: $0 \leq k <\sim 16\,\text{Å}^{-1}$.

Another factor that limits the spectrum length is the presence of noise. To determine the noise is a straightforward task for r-space, where EXAFS signals at high r have clearly noise character. Assuming noise in k-space and r-space, n_k and n_r, to be constant, via Parseval's identity one obtains (Newville *et al.*, 1999):

$$n_k^2 = n_r^2 \frac{\pi}{dk} \frac{2w+1}{k_{\max}^{2w+1} - k_{\min}^{2w+1}}. \quad (4)$$

where as n_r the mean value of $|\text{FT}[\chi(k) \cdot k^w]|$ serves, usually over the range $15 < r < 25\,\text{Å}$. This noise is smaller than the signal for all the spectrum (see Fig. 3).

Practically all programs for EXAFS spectra processing (IXS, 2000) to estimate the noise use the Fourier analysis. But then it is the noise that they use as uncertainties ε_i of $\chi(k)$ determination in the definition of χ^2-statistics:

$$\chi^2 = \frac{N_{\text{ind}}}{M} \sum_{i=1}^{M} \frac{[(\chi_{\text{exp}})_i - (\chi_{\text{mod}})_i]^2}{\varepsilon_i^2}. \quad (5)$$

It would be more correct to consider as ε_i the *larger* from the two: the noise and the errors derived from the μ_0 construction. In our case (and as a rule) the latter are noticeably greater.

It is well-known that understated ε_i lead to uncertainties in fitted parameters that are much too small. When the fitted parameters happen to conflict with the data of other experimental techniques, often vague 'systematic errors' are mentioned. We believe that the errors of μ_0, always neglected, contribute essentially to systematic errors, and here we have shown how this contribution can be determined.

Note finally, that all the stages of EXAFS-function extraction and its uncertainty estimations are realized in the freeware program VIPER (Klementev, 2000*a*).

This work was supported by RFBR (99-02-17343) and Programm "Superconductivity" (99010). I thank the Program Committee for the financial support of my participation in the XAFS-XI conference.

References

Ankudinov, A. L., Ravel, B., Rehr, J. J. & Conradson, S. D. (1998). *Phys. Rev. B*, **58**, 7565–7576.

IXS, (2000). Catalog of XAFS Analysis Programs, `http://ixs.csrri.iit.edu/catalog/XAFS_Programs`.

Klementev, K. V., (2000*a*). *VIPER for Windows (Visual Processing in EXAFS Researches)*, freeware, `www.crosswinds.net/~klmn/viper.html` .

Klementev, K. V. (2000*b*). E-arXiv: physics/0003086 (xxx.lanl.gov).

Newville, M., Boyanov, B. I. & Sayers, D. E. (1999). *J. Synchrotron Rad.* **6**, 264–265. (Proc. of Int. Conf. XAFS X).