Brief Reports

Brief Reports are accounts of completed research which, while meeting the usual **Physical Review** *standards of scientific quality, do not warrant regular articles. A Brief Report may be no longer than four printed pages and must be accompanied by an abstract. The same publication schedule as for regular articles is followed, and page proofs are sent to authors.*

Number of relevant independent points in x-ray-absorption fine-structure spectra

Edward A. Stern
Physics Department FM-15, University of Washington, Seattle, Washington 98195
(Received 5 April 1993; revised manuscript received 1 June 1993)

A discussion is given of the amount of information present in x-ray-absorption fine-structure spectra. It is shown that the usual formulations underestimate the degrees of freedom by at least one.

I. INTRODUCTION

The purpose of this paper is to determine the number of relevant independent points, N_1, that are present in an x-ray-absorption fine-structure (XAFS) spectrum and can be used to define parameters needed to fit the spectrum. In general, the number of parameters that can be determined from the XAFS is less than N_I, and, thus, N_I is a fundamental limitation to the amount of information that can be determined by XAFS.

There is some uncertainty in the literature on how to calculate N_I. The formula is usually given by[1]

$$N_I = 2\delta k \delta R / \pi \tag{1}$$

where δk is the k region of the XAFS spectrum analyzed and δR is the region in R phase over which the fit is made. As an example, the δk range is usually determined on its low-value end by where the background can reliably be separated from the total absorption and on the high-value end by where the noise becomes too large compared to the signal to tolerate. The δR range is the region to be fitted, which, when Fourier filtering is used, is the width of the window in r space. However, N_I has also been given by[2]

$$N_I = (2\delta k \delta R / \pi) + 1 \ . \tag{2}$$

What we show here is that the correct expression is neither of the above but is

$$N_I = (2\delta k \delta R / \pi) + 2 \ . \tag{3}$$

Though, in many cases N_I is large compared to one, and the differences among Eqs. (1)–(3) are not important, there are other cases where increasing N_I by one or two may make a significant difference, and, thus, there is a practical reason, in addition to the aesthetic one, to determine the correct expression for N_I. For example, the usable XAFS range for Pb metal near its melting point is quite limited because of the large amplitude of vibrations. In that case[3] $\delta k = 2$ Å^{-1} and $\delta R = 1.5$ Å giving $2\delta k \delta R / \pi = 2$. Thus, adding 2 to this value doubles the estimate of information content of the data. This doubling was essential to obtain useful information.

We give a straightforward derivation for N_I, which shows that Eq. (3) is the correct one. Although the derivation discusses the case of XAFS, the results are general and are applicable to any general function, such as the determination of correlation functions from diffused scattering measurements using x rays or neutrons. The concept of N_I is very closely related to ideas of information theory and how many parameters (or degrees of freedom) are required to define a given function measured over a time interval (corresponding to the k interval in XAFS). The usual formulation of this problem assumes that all frequencies (corresponding to $2r$, twice the radial distance, in XAFS) up to a maximum are present in the function and give Eq. (2) for the number of parameters. In the case of XAFS, the r range starts at a finite value, r_1, and this has to be accounted for in the derivation. As is shown below, it is because $r_1 \neq 0$ that a one has to be added to Eq. (2) to obtain Eq. (3).

II. DERIVATION

Consider the analysis of an XAFS spectrum $\chi(k)$, which is idealized to a situation where $\chi(k_i)$ is measured at discreet values of k_i with a random root-mean-square (rms) uncertainty $\sigma(k_i)$. For simplicity, $\sigma(k_i) = \sigma_k$ is assumed independent of k_i. It is assumed that the measurement and the background subtraction to isolate $\chi(k)$ introduce no systematic errors. The values of k_i are assumed to be on a uniform grid with a constant spacing Δk_0 so that

$$k_i = k_1 + n_i \Delta k_0 \ , \tag{4}$$

where n_i is an integer between zero and N_k. Thus the maximum value of k is

$$k_2 = k_1 + N_k \Delta k_0 \ . \tag{5}$$

In practice,[4] the spectrum $\chi(k)$ is analyzed only over a

finite interval of the conjugate variable, i.e., $r_1 - r_2$. The lower limit r_1 is determined by where the structural information differs significantly from any residual background signal not correctly subtracted from the total absorption, while the upper limit r_2 is set by either the strength of the signal or in the case considered here, the fitting will be done explicitly in r space to clearly define the region over which the spectrum is being analyzed. The function to be fit in r space is the Fourier transform of $\chi(k)$, namely,

$$f(r_j)=\sum_i \chi(k_i)e^{i2k_i r_j} . \tag{6}$$

Another way to define the region of the fit is to filter $f(r_j)$ with a window of $r_1 \rightarrow r_2$ and back transform to k space and then fit the filtered $\chi(k)$ in k space. This is equivalent to the direct fitting in r space, but fitting directly in r-space saves one transformation on the data and any attendant distortion introduced by the filtering, so it is a more direct way to define the fitting procedure.

Although the $\chi(k_i)$ is known only over the range $\delta k = k_2 - k_1$, it is common practice[3] to extend the transform and make it periodic over a longer range $K > \delta k$. The function $\chi(k_i)$ is then defined to be zero in the range outside δk. This is done so that the points r_j can be more closely spaced, since they are given by

$$r_j = \pi n_j / K , \tag{7}$$

where n_j are integers.

Because of the random noise in $\chi(k_i)$, a corresponding mean-squared fluctuation in $f(r_j)$ occurs, which can be calculated from Eq. (6) to be

$$\begin{aligned}\langle |\Delta f(r_j)|^2 \rangle &\equiv \sigma_r^2 \\ &= \left\langle \sum_{il} \Delta\chi(k_i)e^{ik_i r_j}\Delta\chi(k_l)e^{-ik_l r_j} \right\rangle \\ &= (N_k+1)\sigma_k^2 ,\end{aligned} \tag{8}$$

where the relation $\langle \Delta\chi(k_i)\Delta\chi(k_l)\rangle = \sigma_{k_i}\delta_{il}$ has been used, since the measurement points are independent and $\langle \chi(k_i)\rangle = 0$. However, there is generally a correlation between the points r_j and $r_l = r_j + \Delta r$, which can be calculated to be

$$\langle \Delta f(r_j)\Delta f(r_j+\Delta r)\rangle = \sigma_k^2 \sum_i e^{i2k_i\Delta r} . \tag{9}$$

The sum in (9) can be well approximated by an integral, assuming Δk_0 is small compared to the variations in $\chi(k_i)$, as should be the case for good data collection. In that case

$$\langle \Delta f(r_1)\Delta f(r_1+\Delta r)\rangle = \sigma_k^2 (e^{i2\bar{k}\Delta r}\sin\delta k\,\Delta r/\Delta k_0 \Delta r) \tag{10}$$

where $\bar{k}$ is the midpoint of the k range [$\bar{k} = (k_1+k_2)/2$].

Equation (10) indicates that points separated by Δr are not independent but are correlated, since, if the points were independent, the average in Eq. (10) would be the product of the average of each term. The average of each term is zero, and thus a *necessary* condition for the points to be independent is for the average in Eq. (10) to be zero. This occurs for

$$\Delta r = m\pi/\delta k , \tag{11}$$

where m is an integer. Thus, the independent points in r space are spaced a distance apart by the amount

$$\Delta r_0 = \pi/\delta k . \tag{12}$$

These points are *sufficient* to completely define the function $\chi(k_i)$ in the range $k_1 \rightarrow k_1 + \delta k$ by the Fourier series

$$\chi(k_i) = \sum_{i=0}^{\delta r/\Delta r_0} f(r_1 + n_i\Delta r_0)e^{i2k_j(r_1+n_i\Delta r_0)} . \tag{13}$$

As proven by the Fourier theorem, this function is periodic with a periodicity δk, but it fully reproduces the data in the required range.

Note that though the extension of the periodicity of the $\chi(k_i)$ from the measured range δk to K by adding zeros increased the density of points in $f(r_j)$, this extension, not surprisingly, did not increase the information in the data, since the spacing of independent points is given by the measured range δk only. The function $f(r_j)$ of Eq. (6) is, in general, complex and is fitted in the finite range $r_1 \le r_j \le r_2$. Such a complex function can be defined to be a real function over the range $r_1 \le r_i \le r_2$ and $-r_1 \ge r_i \ge -r_2$ by the relation

$$f(-r_j) = f^*(r_j) , \tag{14}$$

which follows from Eq. (6), since $\chi(k_i)$ is real.

Assume that r_1 and r_2 are chosen so that they each coincide with an independent point and are given by

$$\begin{aligned} r_1 &= m_1\Delta r_0 , \\ r_2 &= m_2\Delta r_0 . \end{aligned} \tag{15}$$

In this case the number of independent points in the range

$$\delta r = r_2 - r_1 \tag{16}$$

is

$$n_i = \delta r/\Delta r_0 + 1 , \tag{17}$$

the one coming from the fact that there is one more point than there are spaces. The total number of independent points N_I is twice n_i because of contribution from the negative-r region so that

$$N = 2n_i = 2\delta r\delta k/\pi + 2 . \tag{18}$$

III. DISCUSSION

Formula (18) differs from the standard one in Eq. (2) by an additional one, and it is important to understand this difference. The standard formula assumes that $r_1 = 0$. In such a case the total r-space range is $2\delta r$ covering the re-

gion $-r_2 \leq r_i \leq r_2$ and the number of independent points is the number of spaces $2\delta r/\Delta r_0$ plus one or

$$N_I = 2\delta r \delta k/\pi + 1 , \tag{19}$$

agreeing with the standard formula.

Another way to count the number of parameters needed to define $f(r_i)$ is to consider only the positive range from r_1 to r_2 and for each independent point, except for $r=0$, two numbers are required, namely, the Re and Im parts of $f(r_i)$. For the point $r=0$, Eq. (14) shows that $f(0)$ is real, and only one number is required. Thus, for the range $0-r_2$ the number of points with $r \neq 0$ is $\delta r/\Delta r_0$ and the number of parameters required to define $f(r_i)$ is

$$N_I' = 2\delta r/\Delta r + 1 ,$$

where the 1 comes from the point $r=0$.

However, if $r_1 \neq 0$, then the number of independent points contained in the range r_1-r_2 is $\delta r/\Delta r+1$, and each point requires two parameters to define $f(r_i)$, giving the total number of parameters to define $f(r_i)$ as

$$N_I = 2(\delta r/\Delta r + 1) ,$$

in agreement with Eq. (18). Again, the difference between Eqs. (17) and (18) comes about because $f(r_i)$ in the latter does not contain the point $r=0$. For the XAFS case the latter is always the case, and, thus, the number of parameters necessary to define the function $f(r_i)$ is given by Eq. (18).

N_I is a measure of the total amount of information contained in an XAFS spectrum. It is the maximum number of parameters that can be determined by a fit to the spectrum. This occurs, in spite of the fact that the N_k+1 points in k space are all independent measurements. The derivation also points out that the number of the parameters is uniformly spaced in r space, and the parameters that are used to fit the data must be associated with the r range being fit and their number must be less than or equal to the value of N_I associated with that r range. Thus, when fitting several shells of atoms, it may occur that the number of parameters required to fit the outer shells is larger than the number of independent points enclosed in the r-space range of the outer shell, though less than the total number of points enclosed in the r-space range, which include the first shell. This can occur because the spacing between the first shell is larger than the spacing between outer shells. In this case, one cannot borrow the independent points associated with the first shell and use them to determine the parameters of the outer shells, since the first shell-independent points do not contain such information.

IV. SUMMARY

It has been shown that the number of relevant independent points, or degrees of freedom, of an XAFS spectrum is given by Eq. (3), which is one more than the standard formula from information theory [Eq. (2)] because XAFS does not include the vicinity of $r=0$ in its analysis.

It should be noted that extending the k range of the data over which a Fourier transform into r space is performed by adding zeros makes the r space transform more continuous, and its appearance more pleasing, but no new information is added by this addition of more points in r space. The limitation of the amount of information present in the XAFS spectrum, or any related spectrum such as the diffuse elastic scattering of x rays or neutrons, restricts the number of structural parameters that can be determined to be below N_I.

Also, adding more points k_i to the measurement of $\chi(k_i)$ by increasing their density but not increasing the range δk does not increase the information content, i.e., the degrees of freedom, of the measurement. If $\sigma(k_i)$ for each point remains the same, increasing the number of points would reduce the random noise in the data, but this could be done just as well by spending more time on measuring the original points to add up to the same total measurement time as that used with the increased number of points.

ACKNOWLEDGMENTS

Stimulating discussions with Matthew Newville, Professor John J. Rehr, and Professor Yizhak Yacoby are gratefully acknowledged. Research was supported by DOE Grant No. DE-FG-06-90ER454 25.

[1]P. A. Lee, P. H. Citrin, P. Eisenberger, and B. M. Kincaid, Rev. Mod. Phys. **53**, 769 (1981).

[2]S.-L. Lin, E. A. Stern, A. J. Kalb (Gilboa), and Y. Zhang, Biochem. **30**, 2323 (1991); see also L. Brillouin, *Science and Information Theory* (Academic, New York, 1967), p. 94.

[3]E. A. Stern, P. Livins, and Z. Zhang, Phys. Rev. B **43**, 8850 (1991).

[4]See, e.g., *X-Ray Absorption, Principles, Applications, Techniques of SEXAFS, and XANES,* edited by D. C. Konigsberger and R. Prins (Wiley, New York, 1988).