

Review

Mössbauer spectroscopy of metalloproteins and the use of Fourier transforms

W. R. DUNHAM¹, L. J. HARDING² and R. H. SANDS¹

¹ Biophysics Research Division, The University of Michigan, Ann Arbor, MI, USA

² Information Technology Division, The University of Michigan, Ann Arbor, MI, USA

(Received July 19, 1992/February 4, 1993) – EJB 92 1004/0

A method for obtaining accurate, quantitative ⁵⁷Fe Mössbauer spectra from biological samples is illustrated stepwise in a data reduction procedure. Exact criteria are presented for deciding when it is necessary to account for the effects of the Beer-Lambert law in the Mössbauer spectra from biological samples. This procedure makes extensive use of the fast Fourier transform and other computer techniques in its data reduction and its curve-fitting stages. A method for optimizing sample thickness is presented. The choice of truncation in Fourier space as a means to numerically stabilize the deconvolution procedure is defended. Several advantages for curve fitting in Fourier space are shown. Maximization of information content is discussed for Mössbauer spectral simulation techniques.

For over 15 years, this laboratory has published Mössbauer spectra from iron-containing proteins and other biological samples using a unique method of data acquisition and reduction [1–4]. During the course of these studies, this method has not been described in detail (except in theses) nor has the motivation for its use been presented. The following description is an attempt to correct these shortcomings.

The source lineshape convolution underlying Mössbauer spectra

In Fig. 1, the fundamental parts of a Mössbauer spectrometer are illustrated in a cartoon. The radioactive source (⁵⁷Co in a rhodium matrix) decays via electron capture [5] to emit gamma rays (principally at 122 and 14 keV energies) and X-rays (at 20 keV and 6 keV energies). This emitted beam (I_0) is collimated, passes through the sample and is counted as single photons by an argon gas proportional counter. The energy of the source radiation is modulated via the relativistic Doppler shift [6] to yield a count rate that varies with source velocity, $I(v)$. The entire process, including the nuclear resonance fluorescence (Mössbauer effect) and the non-resonant events, has been described by the following expression [2, 7]:

$$I(v) = I_0(1-f_s) + I_a + I_b + I_c + f_s I_0 \int_{-\infty}^{\infty} S(v-E) \cdot \exp[-\sigma(E)f_a c \ell \mathcal{L}/1000] dE \quad (1)$$

where $I(v)$ and I_0 are the detected 14 keV count rates (corrected for pileup and solid angle effects) in the presence and absence of a sample, respectively. f_s and f_a are the Debye-Waller factors for the source and absorber, $S(v-E)$ is the source lineshape, $\sigma(E)$ is the Mössbauer effect cross-section for the absorber, c is the absorber molarity, ℓ is the sample thickness and \mathcal{L} is Avogadro's number. I_a is the non-resonant 14 keV background count rate and will be assumed here to

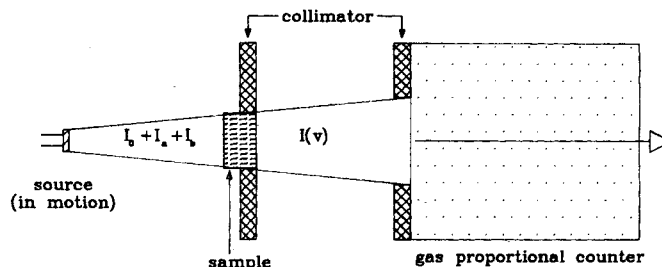


Fig. 1. The layout of a basic Mössbauer spectrometer.

arise entirely from 122 keV gamma rays and will be given the value measured by our argon counter within a month of the source manufacture date. I_b is the background rate due to the source but not attenuable by the sample. I_c is the cosmic background rate. The details of the measurements of these various functions and parameters is given elsewhere [4]. Typical values for a new source (35 mCi ⁵⁷Co on 6 μ m Rh foil) are: $I_0 = 20500$, $I_a = 750$, $I_b = 26$ and $I_c = 4$ counts/s.

It is common practice to present Mössbauer data as a plot of transmission (%), $100 I(v)/I(-\infty)$, as a function of source velocity. In 1971, Ure and Flinn [7] made a major contribution to this spectroscopic field by demonstrating a method to obtain the cross-section, $\sigma(v)$, from the data by a process of Fourier deconvolution, followed by taking the logarithm of the data as is usually the case in optical spectroscopy to account for the Beer-Lambert law. Their method was later modified [3] by a change to the apodization procedure and a detailed explanation of the laboratory and computer methods used to implement their Fourier deconvolution procedure. However, from a previous paper by Blume and Kistner [8], one can show that Eqn (1) is erroneous for the cases where magnetic splittings (nuclear Zeeman, and magnetic hyperfine interactions [9] are part of the Hamiltonian for the absorbing ⁵⁷Fe nucleus. These innovations were integrated into a data reduction procedure by Filter et al. [3, 4]. The purpose of this paper is to demonstrate the advantages of this procedure over that in common practice for biological samples.

Correspondence to W. R. Dunham, Biophysics Research Division, The University of Michigan, 2200 Bonisteel Blvd., Ann Arbor, MI 48109, USA

Optimum sample thickness

Because inelastic cross-sections for the relevant photon energies are proportional to Z^2 (Segre, see [10]), one can often model a biological sample as having the iron nuclei suspended in water when substituting values into Eqn (1). It is obviously important that high- Z elements, such as Cl, be controlled for this model to be valid, but for many samples with organic buffering agents the approximation is sufficient for our argument.

In the presence of magnetic splittings, circular polarization of the gamma-ray beam requires that Eqn (1) be revised to account for the density matrix for the impinging photons [8]. The effects of polarization are very complicated [4], but the underlying principle can be simply demonstrated by answering the question: How much of I_0 can be removed by an absorption process that requires a photon that is right circularly polarized? The answer is 'one-half' at a maximum because normally the source radiation is not circularly polarized; it is composed of half right- and half left-circularly polarized light. Because the absorption process requires right-circularly polarized light, it cannot affect photons with the 'wrong' polarization. As mentioned above, this effect requires that Eqn (1) be rewritten to account for circular polarization of I_0 . In the following argument, we will assume that $\sigma(\nu)$ contains only a quadrupole pair describing each particular iron environment. Under this assumption, Eqn (1) is valid because there is no polarization in the absorption lines from a randomly oriented sample where its $\sigma(\nu)$ is the sum of quadrupole pairs [4]. In the laboratory, this assumption is equivalent to taking the spectrum at zero applied magnetic field and at a high temperature so that the high internal magnetic field of the iron atom is averaged to zero (for paramagnetic samples).

With this understanding, we can substitute numbers into Eqn (1) that model a protein sample as an iron atom dissolved in water. For an absorption displaying the smallest possible linewidth (calculated from the nuclear lifetime and decay branching ratios as full width at half maximum [FWHM] = 0.0975 mm/s), the absorption maximum is given by the expression for resonance fluorescence [6]:

$$\sigma_{\max} = \sigma_0 = 2.566 \times 10^{-18} \text{ cm}^2. \quad (2)$$

The Mössbauer spectra of many biological samples are quadrupole pairs with linewidths ≈ 0.16 mm/s (FWHM). f_a is assumed to be 0.8. With these assumptions,

$$f_s c \ell \mathcal{L}\sigma(\nu)/1000 = 386 c \ell \frac{\Gamma^2}{\Gamma^2 + \nu^2} \text{ cm}^{-1} \quad (3)$$

for one of the two- lines. Using mass attenuation coefficients for water that are available in the literature [11], we can rewrite Eqn (1) to model our spectrometer's count rate when running a biological sample at maximum resonance as:

$$I(0) \approx 20500e^{-2.0\ell} \left\{ (1-f_s) + f_s \int_{-\infty}^{\infty} S(-E) \cdot \exp \left[-386c\ell \frac{\Gamma^2}{\Gamma^2 + E^2} \right] dE \right\} + 750e^{-0.16\ell} + 30. \quad (4)$$

If $x = 0.05$, then $e^{-x} = 0.951$. Thus, the approximation that $e^{-x} \approx 1 - x$ is certainly valid for the cases when the argument of the exponential is less than 0.05. In these situations the convolution performed by the integration can be approximated by a doubling of the linewidth. Accordingly, the depth

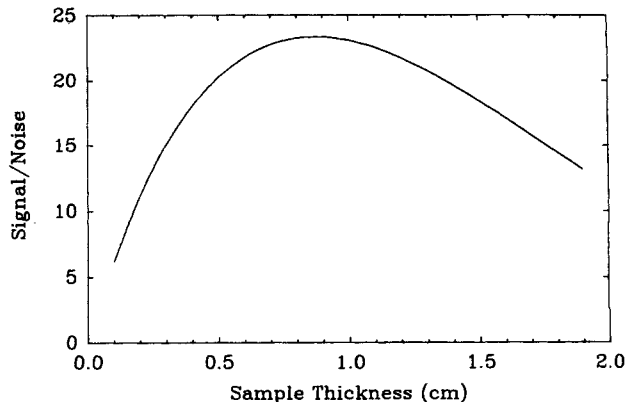


Fig. 2. The signal/noise ratio of a 10 μM , single-iron environment (quadrupole pair) after a 28-h collection time as a function of sample thickness. The open beam 14.4 keV rate is 20500 counts/s; the rate due to 122-keV gamma rays is 750 counts/s and the background rate is 30 counts/s. The non-resonant absorption is modeled as that of water.

of resonance is divided by a factor of two, leading to the approximation:

$$I(0) \approx 20500e^{-2.0\ell} \{ (1-f_s) + f_s(1-193c\ell) \} + 750e^{-0.16\ell} + 30. \quad (5)$$

The number of counts represented by the resonance is $[I(\infty) - I(0)]t$, where t is the duration time (s) of the experiment. Because radioactive decay is described by Poisson statistics (see [12]), the 'noise' in the experiment can be approximated as the square root of $I(\infty)t$. Thus, a trivial definition of the signal/noise ratio (S/N) is:

$$S/N = [I(\infty) - I(0)]t / [I(\infty)t]^{1/2}. \quad (6)$$

Arbitrarily, we let $c = 10^{-5}$ (10 μM) and $t = 10^5$ (28 h). We can then graph Eqn (6) as a function of ℓ (Fig. 2). This graph shows that sample thicknesses between 0.3–2 cm give reasonable signal/noise ratios with the maximum occurring around 0.9 cm. With the above assumptions, the graph is only rigorously valid for weak absorbers ($c\ell < 1.3 \times 10^{-4}$); however, if the absorber is strong, one is usually less concerned about optimizing signal/noise ratios.

Thus, there is a large latitude allowed for choice of sample size. The optimization of signal/noise may not coincide with the point of view of the biochemist whose task it is to supply the sample. Our spectrometer has a cylindrical sample volume of 200 μl with a 0.69-cm diameter; therefore the sample thickness is 0.54 cm.

When is it necessary to account for the Beer-Lambert law?

From Eqn (4) one can see that the response of the spectrometer will become non-linear with respect to concentration when the argument of the exponential in the transmission integral [7] becomes too large. We showed that this error is approximately 0.001 if the exponential argument is 0.05. This kind of error is negligible unless the number of counts in the resonant peak of the data is greater than a million. (From Poisson statistics, we know that the noise is equal to the square root of the counts.) In most biological experiments, there are other unknowns (concentration and sample volume and geometry) that lessen the utility of such an accurate quantitation even if it were accessible. Furthermore, the difficulties in obtaining absolute quantitations are

usually seen as beside the point because: (a) in most laboratories, quantitation is achieved by running standards, and (b) most biochemists are chiefly interested in relative intensities when using Mössbauer spectroscopy.

In the following treatment, the effect of the Beer-Lambert law (Eqn 4) on relative intensity measurements will be emphasized. The arguments will be appropriate to the case where two lines are compared in one spectrum or in two different spectra, as in a quantitative calibration experiment. To make the example concrete, assume that the sample thickness is 1 cm (from the maximum in Fig. 2) and that the protein concentration is 100 μM , where there is one iron atom in one environment and three iron atoms in another environment. This example is also appropriate to comparing data from a single-iron 100 μM protein to that of a single-iron 300 μM protein.

Substituting into Eqn (4), the exponential argument at resonance ($\nu = 0$) is 0.03088 for the one-iron environment and 0.09264 for the three-iron environment. The error due to ignoring the Beer-Lambert law is 1.5% and 4.5%, respectively. The intensity ratio is 1 : 2.909 for an error of 3%. In this example, it was assumed that the protein molecule has four iron atoms; a 3% error does not lead to an incorrect determination of the quantity of iron in the protein, for example. Furthermore, most laboratories use absorption areas instead of peak intensities, so that the error initially appears to be negligible in this situation.

Suppose, however, that the protein molecule (100 μM) contains 16 iron atoms so that the above exponential arguments are multiplied by four relative to the preceding paragraph. In this case the intensity errors are 5.9% and 16.4% for the two lines; the intensity ratio is 1 : 2.665 for an error of 11.2% or 1.8 of the 16 irons in the protein. Ignoring the Beer-Lambert law is a disaster in this case. Furthermore, it has been common practice in other laboratories to use broadened linewidths (three-parameter lineshapes) in an attempt to correct for the effects of these non-linearities and thereby to use areas (not peak intensities) for quantitation. Not only is this procedure open to large errors in interpretation and to errors due to noise when (as in this example) there are 16, possibly different, iron atoms in the protein, but the resolution of the spectroscopy is sacrificed unnecessarily. Therefore, the argument that one should look at areas to lessen the effects of exponentiation need not apply when the spectrum is so complicated that interpretation is not suitable to curve-fitting methods. A good example of the relevance of the Beer-Lambert law to protein Mössbauer spectra can be found with component 1 of nitrogenase [13], a protein with over 30 iron atoms/molecule.

However, all of the preceding arguments have assumed that one can use Eqn (4) to compare spectra or the lines in the spectra. To use Eqn (4), one must determine f_a , the absorber Debye-Waller factor. The determination of this factor for a protein sample [4] requires very precise knowledge of the spectral intensity on an absolute scale. On this scale, a 5% intensity error is very large because the temperature range over which we are often constrained for biological samples is very small, usually 125–175 K. Only in this range is the spectrum composed of pure quadrupole pairs without any complication by magnetic interaction or the onset of the liquid state and its effects on recoil-free fraction.

Therefore, the question of when it is necessary to incorporate the Beer-Lambert law in a Mossbauer data reduction scheme depends greatly on what is to be done with the data. However, when a single-iron environment reaches 300 μM

with a 1-cm sample thickness, accuracy in quantitative spectroscopic studies requires that one choose some method to account for the effects of this law. Thus, in a Mössbauer laboratory, work on biological compounds is continuously threatened by a lack of standard procedures to handle the Beer-Lambert law. In the following, reasons are presented for preferring one of these methods.

Why should one deconvolve the transmission integral?

By solving Eqn (1) for the integral, one arrives at an expression for the 'transmission integral' [7], $T(\nu)$:

$$T(\nu) = [I(\nu) - I_0(1 - f_s) - I_a - I_b - I_c] / f_s I_0. \quad (7)$$

This function describes the interaction of the sample with those particular gamma rays that could have been absorbed by the Mössbauer effect; i.e. the recoil-free fraction of the 14.4 keV gamma rays. Forming the data into this array is analogous to correcting for 'dark current' and 'stray light' in an optical spectrometer. The mathematics of Eqn (7) are straightforward, but every variable in the expression must be measured for each spectrum. Although this task is simplified by routine, this laboratory has found that there are many subtleties masked in some of the parameters. For example, self-absorption in the source [14] (by the ^{57}Fe left by the decay of ^{57}Co) results in a time dependence in the source Debye-Waller factor [4].

The major importance of the function, $T(\nu)$, is that its evaluation is a required step toward reaching the desired quantity, $\sigma(\nu)$. From Eqn (1) one can see that

$$T(\nu) = \int_{-\infty}^{\infty} S(\nu - E) \cdot \exp[-f_a n \sigma(E) / A] dE \quad (8)$$

where we have rewritten Eqn (1) in terms of the number of ^{57}Fe atoms, n , and the cross-sectional area of the sample, A . The integral defines a convolution of the source lineshape, S , with the absorption function of the sample lineshape, σ . There are three obvious ways to deal with this convolution: (a) one can deconvolve the integral to obtain $\sigma(\nu)$; (b) one can fit $T(\nu)$ with computer-simulated spectra that have been convolved explicitly; or (c) one can ignore the problem and fit the raw data directly. Reasons have been given for rejecting this latter option when attempting quantitative work on biological samples. The reasons for preferring deconvolution of the data over convolution of the simulated spectra involve computer-simulation techniques and signal/noise considerations. They also depend on whether the absorber is strong or weak.

Fourier deconvolution allows one to increase signal/noise on weak samples

In Fig. 3 is an example of how one can first separate the signal from the noise (by fast Fourier transformation), perform the deconvolution (by dividing by the transform of the source lineshape), then reject that part of the Fourier space where the signal/noise is less than one, thus significantly lowering the noise component in the final spectrum (compare Fig. 3A and 3K). Although Fig. 3 shows each step of the deconvolution process pictorially, two of the details can be reviewed to advantage. The mathematical basis of the process is the deconvolution theorem of Fourier transforms, which states that the Fourier transform of a convolution of two functions is equal to the product of the Fourier transforms of the two functions. Therefore, a deconvolution is

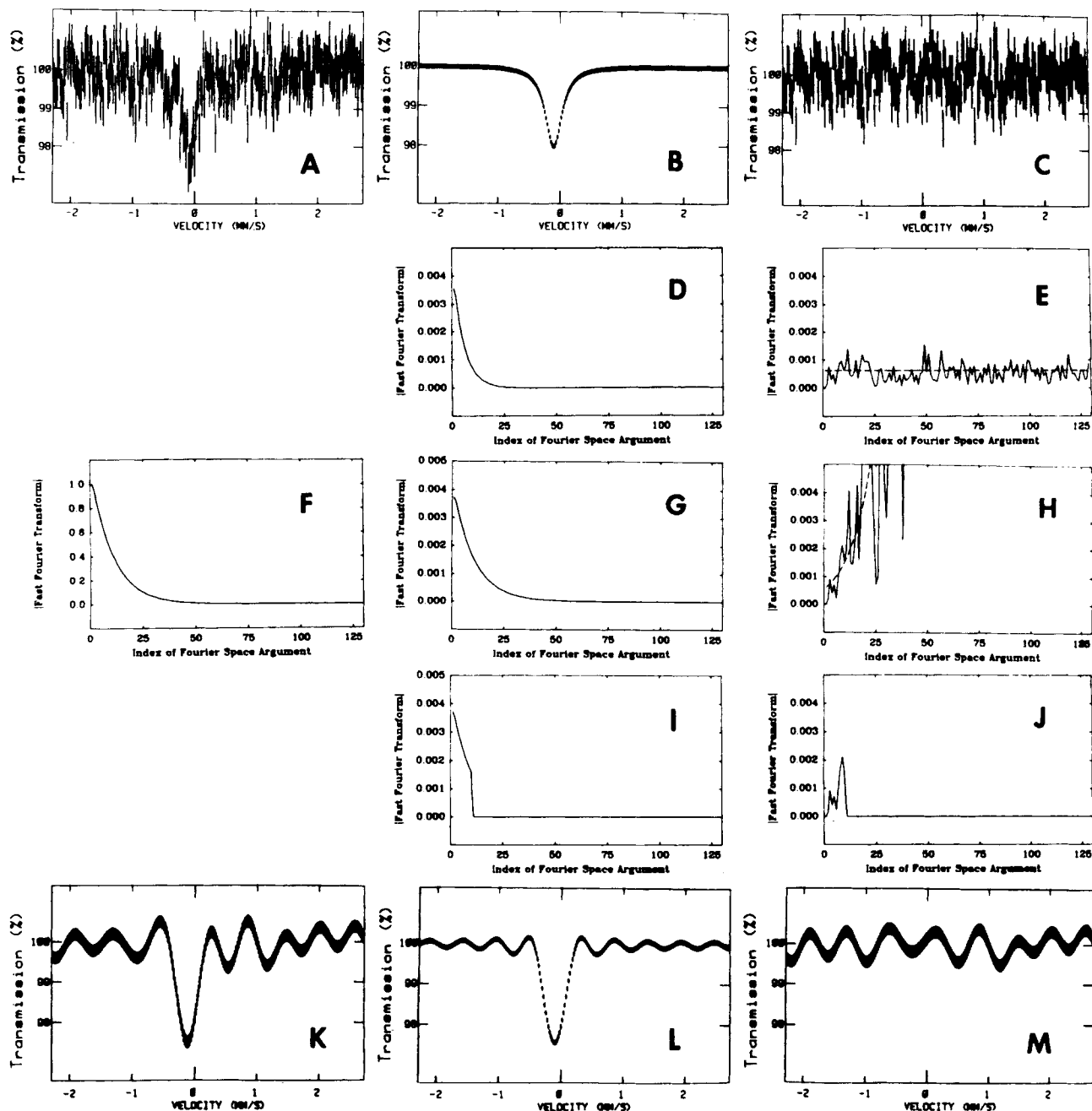


Fig. 3. The individual steps of the Fourier deconvolution procedure shown separately for the signal and the noise of a spectrum. The spectrum (A) is the sum of two parts: a signal (B) with FWHM = 0.32 mm/s and amplitude 0.02, and a noise component (C) whose root-mean-square amplitude is 0.005. The absolute magnitudes of the fast Fourier transforms of the two components are shown in (D) and (E). These two transforms are divided by the fast Fourier transform of the source (F), with FWHM = 0.16 mm/s, to form the deconvolved transforms (G) and (H). These are truncated at the 11-th coefficient; (I) and (J) and back-transformed to form the deconvolved spectra (K) and (L), which are summed to form the resultant spectrum (M). The dashed lines in (E) and (H) are the root-mean-square noise levels, before and after deconvolution.

equivalent to a division of two functions in Fourier space. Also, the truncation point in the Fourier spaces is situated where the signal/noise equals one, when the line in Fig. 3G has the same ordinate as the dashed line in Fig. 3H.

The spectra in Fig. 3L and 3M also illustrate another property of the Fourier deconvolution process: the spectra have 'ripples', which result from the truncation in Fourier space. Some scientists have found these ripples to be horrifying artifacts that cannot be understood or trusted. In fact, they are simply what is left after the undesirable, high 'fre-

quency' part of the Fourier space is discarded. By comparing Fig. 3C with 3M, one can see that the ripples were in Fig. 3C, but were not easily seen under the 'noise'. However, it is important to realize that the signal/noise ratio in Fourier space is not changed by Fourier deconvolution because both signal and noise are divided by the same numbers, the source function (Fig. 3F). The enhancement of signal/noise ratio is made possible by the separation of signal from noise by the fast Fourier transformation. Thus, Fourier transformation is a valuable tool for Mössbauer data reduction, due mostly

to the fact that the Mössbauer absorption lineshape decays exponentially in Fourier space while the noise function is flat (Fig. 3D and 3E).

In this data reduction method, ripples are totally irrelevant because all of the data-fitting procedures take place in Fourier space where there are no ripples. In a standard data reduction procedure, the spectrum analogous to Fig. 3K is transformed to a plot of $n\sigma(\nu)/A$ by taking the negative logarithm according to Eqn (8) and dividing by f_a . Next, the spectrum is routinely shifted back to Fourier space for computer simulation. In Fourier space the trial vectors are very much shorter and simpler to generate (see below) so that the simulator programs run at least tenfold faster in Fourier space and with a greater precision. In a later section, other advantages to curve fitting in Fourier space will be presented. The important point here is that Fourier deconvolution is not as much trouble as it seems because one should use Fourier transforms anyway for other reasons.

In high-signal/noise spectra, Fourier deconvolution allows an accurate accounting of the Beer-Lambert law

If the signal had been much stronger in Fig. 3 and the noise had been much less, then the truncation point in Figs 3I and 3J would have been at a much larger argument in the Fourier space. Accordingly, the lineshape in Fig. 3K would have a much smaller artifact from truncation and would therefore have resembled the real lineshape to a much greater extent. This is fortunate because the subsequent logarithmic step is very sensitive to amplitude when the amplitudes are greater than 0.1, so that the rippling artifact would work against attaining high precision in the data reduction procedure. Thus, Fourier deconvolution benefits Mössbauer spectra from both strong and weak absorbers, but for different reasons. For weak absorbers, the reason is signal/noise enhancement; for strong absorbers, the reason is that the Beer-Lambert law can be accounted for in this way.

Why choose truncation as the apodization procedure?

The operator of any modern NMR spectrometer is well aware of the many possibilities that exist for manipulating the free induction decay signals. It is common for the software to allow many choices for apodization procedures. Our truncation procedure is only one of the choices in this analogy (the free induction decay from a pulsed NMR machine is also back-transformed to form the spectrum in the energy domain). We emphasize that multiplication by any function in Fourier space is an explicit convolution (see above). Because of this complication, the safest procedure for Mössbauer spectroscopists is to leave the elements of the Fourier space unchanged. On the other hand, it is clear from comparing Fig. 3G and 3H that the Fourier deconvolution procedure must be stabilized numerically by some apodization procedure. Otherwise, the back-transformed spectra will be dominated by noise from the high arguments of the Fourier space. Our only disagreement with the treatment of Ure and Flinn [7] is with their choice to stabilize the Fourier deconvolution by multiplying the Fourier space with a Gaussian lineshape. This multiplication in Fourier space is an explicit convolution that carries with it a lowering of the peak intensities in the energy space. Thus regardless of the signal/noise ratio, the Beer-Lambert law cannot be accurately treated by this process. It is not possible to prove the superiority of truncation over Gaussian apodization for every poss-

ible intensity and shape of $n\sigma(\nu)/A$ but, for the extreme case where there is no noise in the spectrum, the correct solution is to leave the Fourier space unchanged. Truncation allows the smoothest approach to this asymptote. For NMR spectroscopists, this choice of apodization procedure is immediately acceptable because it is understood to be the only procedure that can maintain peak intensity in the energy space.

Data simulation in Fourier space

The preceding description of the Fourier deconvolution method has shown how it succeeds in solving problems for Mössbauer spectroscopists, but no reasons for preferring this method over convolution of the spectral simulations have as yet been presented. To understand our bias for deconvolution, it is necessary to inspect the process of *curve fitting* in more detail. Our curve-fitting procedures involve the minimization of variance of a trial vector, Θ , and an experimental vector, x . For example, one possible goodness-of-fit criterion is:

$$S^2 = N^{-1} \sum_{i=1}^N (x_i - a\Theta_i - b)^2 \quad (9)$$

where a is a scaling parameter and b is the 'baseline' for the trial vector, Θ_i . Although a and b are seemingly trivial parameters, they represent a time disaster for non-linear minimization procedures. A change in any of the independent variables in Θ can make an intensity change that in turn presents a problem for choosing the best values of a and b for the particular choice of independent variables. In general, the baseline is difficult to determine from an experimental Mössbauer spectrum, so that the determination of the best values of a and b is in itself a non-linear minimization problem that must be solved for every choice of parameters for the trial vector. Because a and b are coupled in their effects on intensities, the solution of this problem greatly slows the rate of convergence for the fitting procedure.

From elementary least-squares approximation theory, if $b = 0$, then the best choice for a is

$$a = x \cdot \Theta / \Theta \cdot \Theta. \quad (10)$$

Therefore, if the value of b is chosen, then the value of a can be derived by evaluating two dot-products. Although the use of Eqn (10) greatly speeds convergence, the difficulty of determining the baseline remains as a major drag on the minimization process.

However, this problem is greatly modified when viewed in Fourier space. If the vectors in Eqn (9) are re-defined as vectors in the transform space, then $b = 0$ because the baseline is now contained in the first element of Fourier space. One can simply ignore this element when determining the value of S^2 by beginning the summation in Eqn (9) at $i = 2$ and the baseline problem disappears. The value of a is determined quickly at each iteration by using Eqn (10) (this equation is also valid when a , x_i and Θ_i are complex numbers). The value of b is determined after minimization by comparing x_1 with Θ_1 .

There is another computational advantage to fitting in Fourier space. If the spectrum to simulate is written as a sum of m lines of various intensities but the same linewidth, then, in Fourier space, the spectrum can be expressed [15] as

$$\Theta_k = \exp[-|2\pi\Gamma_a(k-1)/N|] \sum_{j=1}^m X_j \cdot \exp[-i2\pi\delta_j(k-1)/N] \quad (11)$$

where δ_j is the isomer shift and X_j is the intensity of the j th resonance, Γ_a is the halfwidth at half maximum of the absorber's Lorentzian lineshape. Either of the exponential functions in Eqn (11) can be generated by iteration formulae. For example, if the lineshape vector is denoted as Θ_i , then one can relate sequential elements of this vector as

$$\Psi_{i+1} = \Psi_i \exp[-2\pi\Gamma_a/N]. \quad (12)$$

Therefore, the entire vector, Ψ_i , can be generated by a single call to the exponential function in the computer program. The complex exponential functions and trigonometric functions can also be generated in a similar manner. In addition, these vectors are relatively short for most spectra because the simulation vectors need extend only to the truncation point.

Although the above may seem esoteric to most biochemists, these details represent a major time saving in computer-simulation methods. None of these are possible without working in Fourier space. The 'baseline problem' is a major time-waster if one convolves the simulated lineshapes because the use of Eqn (10) is not possible in this case. What is not clear to us is the extent to which energy-domain spectral fitting techniques can be optimized. Once we began to fit our Mossbauer spectra (and EPR spectra also) in Fourier space we never returned to our former methods.

Goodness-of-fit criteria

In our first paper on Fourier deconvolution [2], it was shown that the use of the χ^2 -distribution to describe the statistical properties of spectral fitting errors was inappropriate. This approach does not differentiate between parts of the spectrum that contain mostly signal and those that contain mostly noise. The idea of using a weighted χ^2 -distribution was promoted instead where the weighting coefficients were proportional to the size of the signal. However, when one is fitting in Fourier space, it soon becomes obvious that in this domain most of the undesirable noise has already been rejected (by the truncation). Therefore, a non-weighted, root-mean-square error is more appropriate to spectral fitting in the Fourier space than to fitting in the energy space. Although spectral weighting is also possible in Fourier space (the signal/noise ratio is highest at low arguments; see Fig. 3D and 3E), there is an advantage (in addition to speed) to omitting the weighting of the root-mean-square error.

This advantage lies in speeding the convergence of the non-linear minimization procedures. As a minimization course approaches a 'good fit', the criteria for improving the fit fall increasingly on the small details of the spectrum. These small details are contained mostly in the higher arguments of the Fourier space. The advantage of not decreasing the importance of the higher arguments of Fourier space is that these arguments can thereby improve their guidance of the convergence of the minimization process when the changes are small. The period during the final refinements turns out to be where most of the time is spent in minimization procedures.

Note that if the quantity in Eqn (9), $X_i - a\Theta_i$ (the fit residual), is envisioned to be a vector, then the length of this vector is proportional to the root-mean-square error. The length of a vector is invariant to unitary transformation; the fast Fourier transformation is a unitary transformation. Therefore, the root-mean-square error in energy space is being simultaneously minimized by our Fourier space procedures. However, in this new method the energy space has

been changed relative to that from the 'old days' (see Fig. 3A and 3K).

We realize that our use of a flat weighting function in Fourier space is in conflict with one of the maxims of information theory. Translating to the language used here, information theory claims that to maximize signal/noise ratio, the Fourier transform of the data treatment should match the Fourier transform of the signal in the data. Accordingly one should use a weighting function in Fourier space. However, because spectroscopists would find it difficult to accept data filtered in this way and because the goodness-of-fit criterion is only used to gauge the success of our minimization efforts, the use of Fourier space weighting is not recommended.

Does Fourier deconvolution hide minor resonances in a spectrum?

The focal point for criticism of the Fourier deconvolution method for Mossbauer spectra is the appearance of the back-transformed/deconvolved data and derives from the lack of 'high-frequency noise' in the spectra. The back-transformed spectra contain 'ripples', which can be mistaken as artifacts of the truncation in Fourier space and obscurers of the 'true nature' of the spectra. Statements such as 'the fast Fourier transform is a linear transformation' have sometimes failed to convince skeptics that signal/noise ratios or relative amplitudes are not influenced by the process of division in the Fourier space. Likewise skeptics are not always convinced by the explanation that most of the 'ripples' are from the deconvolution of the noise component of the spectrum. We have also heard the claim that when the spectra contain resonance lines of widely varying amplitudes, such as when magnetic interactions dominate, then the 'ripples' will obscure the weaker lines in the spectra.

To illustrate the counter arguments to these claims, the following examples have been constructed. In these examples (Fig. 4) an 'experimental noise spectrum' taken with no absorber present has been added to a theoretical spectrum consisting of two lines with the intensity ratio of 10:1. The calculated spectra with a linewidth (FWHM) of 0.30 mm/s are deconvolved by a source function of identical lineshape, but one-half the linewidth (FWHM = 0.15 mm/s). The deconvolved spectra are back-transformed, divided by two, and plotted over the 'raw data' in Fig. 4A–C. The associated Fourier spaces are plotted in Fig. 4D–F. The three examples (Fig. 4) differ in their signal/noise ratios from the case with the poor signal/noise ratio to the case where only the noise due to digitizing the signal is represented. The twofold reduction in the back-transformed data facilitates comparison of the convolved and deconvolved data. (Because deconvolution conserves area, the deconvolved data would be twice the amplitude of the convolved data without this division.) Specifically, a 0.30-mm/s-wide line has one-half the amplitude of a 0.15-mm/s-wide line with equal area. In these examples, the Beer-Lambert law and its complications are ignored. Therefore, only the effects of Fourier deconvolution are visible in Fig. 4.

Inspecting the overplots in Fig. 4A and 4B shows that the baseline ripples in the deconvolved spectra (solid lines) are actually filtered noise in the convolved spectra. The solid lines in these overplots approximately trace excursions that are visible in the raw data (crosses), except where a resonance occurs. At the resonance positions, the additional resolution associated with deconvolution is obvious. The solid lines in Fig. 4C have one-half the linewidth and the same

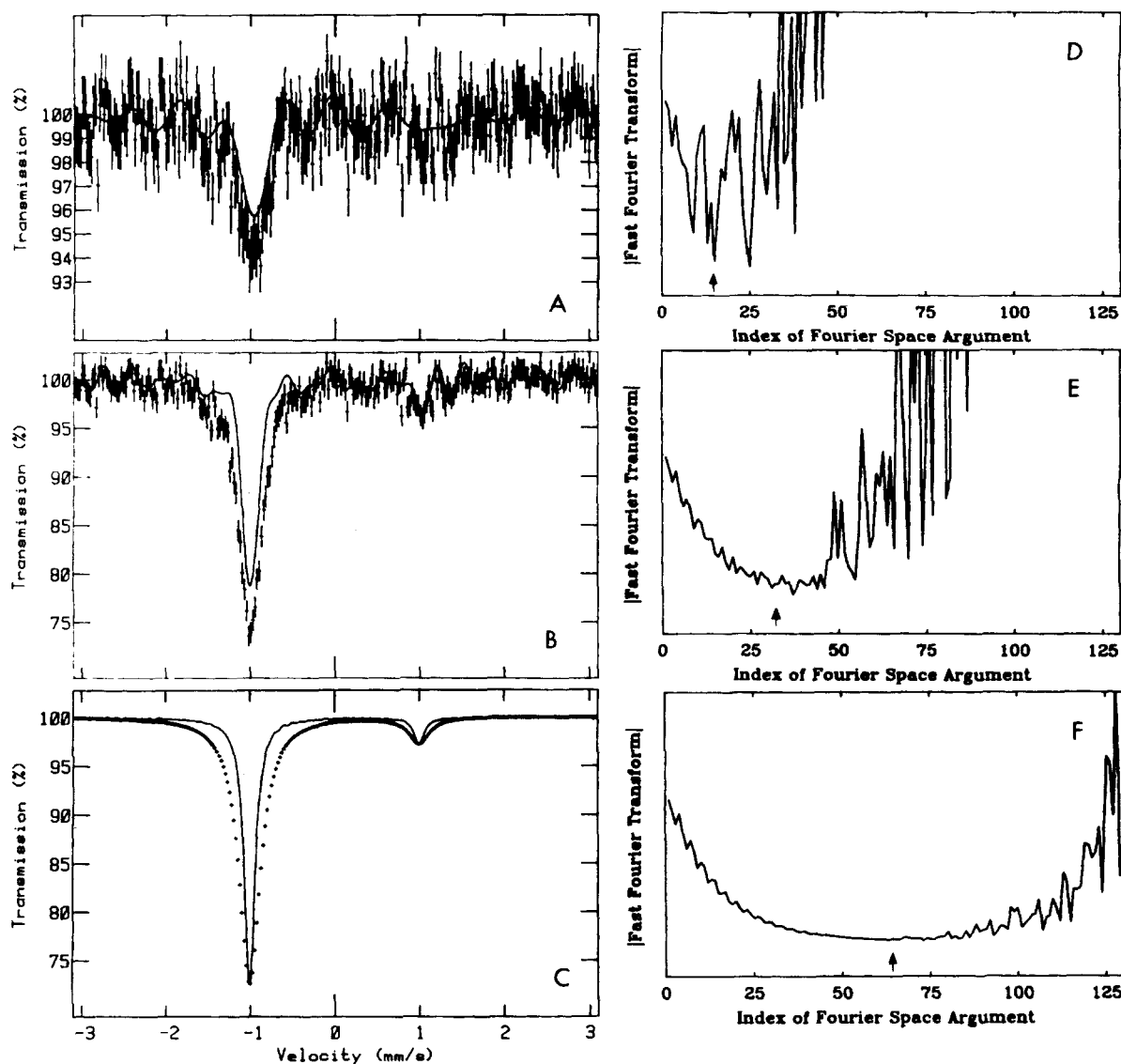


Fig. 4. An experimental noise spectrum. A signal consisting of two lines, with intensity ratio of 10:1 and a linewidth of 0.30 mm/s, has been added to an experimental noise spectrum at various signal/noise ratios. The associated Fourier spaces are also included (D–F) with the truncation points denoted by arrows. The spectral sums, denoted by crosses (A–C), are Fourier-transformed, divided by the transform of a 0.15-mm/s lineshape, truncated, back-transformed, divided by 2 and shown as solid lines in (A–C). The signal/noise ratio is 16 times larger in B than in A. In C only the noise due to digitizing the signal is present.

intensity as their convolved counterparts. In the Fourier domain all three spectra (Fig. 4D–F) have easily distinguishable signal and noise components. The signal is an exponentially decreasing function whereas the noise is an exponentially increasing function with increasing argument of the Fourier space. At the point where the noise is larger than the signal, the Fourier space is truncated (marked by arrows). This truncation point moves to higher arguments as the relative signal amplitude increases. These principles are easily seen in Fig. 4D–F. In the Fourier space, therefore, deconvolution amounts to a division of the signal and the noise by the same numbers; therefore the signal/noise ratio is unchanged for these elements of the Fourier space. The elements of the Fourier space that have a signal/noise ratio less than one are dropped from the analysis by the truncation. Therefore, Fourier deconvolution followed by truncation *must increase the signal/noise ratio* of the spectrum. It is true that the back-transformed spectrum has a modified lineshape as can be seen in Fig. 4A–C; however, this modification

is not present in the Fourier space where the curve-fitting procedures take place. Therefore, these data reduction methods are artifact-free and there can be no loss in sensitivity to weak resonances due to the Fourier deconvolution procedure.

It is possible to argue that the truncation points chosen in Fig. 4D–F were based on the signal amplitude of the major line and were consequently not the optimal choice for the minor line in the spectrum. The Fourier spaces should be truncated at lower arguments because the minor line's signal intensity is smaller than that of the major line. These are valid arguments and can give rise to procedures with added sensitivity to the minor lines in the spectrum because they amount to rejecting more of the Fourier space where the signal/noise ratio is less than one for the minor lines. However, this choice of truncation point would result in a further distortion of the major resonance lineshape and is not recommended. The signal/noise ratio in a single element of the Fourier space is unaffected by division by any number. The

advantages to be gained by over-truncating are marginal and depend on the unpredictable character of the noise in the elements of Fourier space between the two truncation points.

On the other hand, the principle that one should shorten the Fourier space to increase the sensitivity to 'weak' resonances is established in the preceding paragraph. Applying this principle to all spectra implies that truncation in Fourier space enhances rather than decreases the sensitivity of the spectroscopy to 'weak' resonances. However, this additional enhancement is not very large as explained previously (see also Fig. 4A–C).

When a strong magnetic interaction is present at the ^{57}Fe nucleus, the Mössbauer spectra can be very complex with many lines of differing intensity. In this situation, the Beer-Lambert law becomes complex as well [3, 4, 8] so that it is standard procedure in this laboratory to perform the Fourier deconvolution on the data as shown in Fig. 4, but not to take the logarithm of the data as implied in Eqn (1). Instead, the computer-simulated spectra from the left- and right-handed polarizations of the gamma-ray beam are exponentiated, then summed to form a trial simulation for the deconvolved spectra [4]. The curve fitting takes place in Fourier space as usual. The two-line spectra in Fig. 4 model the application to magnetic spectra by demonstrating that the Fourier deconvolution process is appropriate for spectra with highly differing intensities. Mathematically, all the information presented above concerning the application of Fourier deconvolution to multi-line spectra are implied by the statement that the fast Fourier transform is a linear, unitary transformation: vector lengths are unchanged and summation is preserved by fast Fourier transformation.

Conclusion

The preceding comments are an attempt to review and justify the methods of this laboratory for data reduction of Mössbauer spectra. It is shown that they are often beneficial and sometimes necessary for the proper treatment of the data from biological samples. On the other hand, these methods are not always necessary and are difficult to implement. The

laboratory took ten years to develop the enabling machinery and software. Nevertheless, there are many advantages, some of them seem crucial to us, to this method of handling Mössbauer data. The method was designed to deal routinely with the effects of the Beer-Lambert law on Mössbauer spectra. It performs this function well and, as a side benefit, also provides enhanced resolution and signal/noise ratios, even in the presence of magnetic field splittings. The problem is simply seen from a 'transformed' point of view.

REFERENCES

1. Wu, C.T. (1975) PhD thesis (Physics), University of Michigan, University Microfilms 7609549.
2. Dunham, W.R., Wu, C.T., Polichar, R.M., Sands, R.H. & Harding, L.J. (1977) *Nuclear Instrum. Methods* 145, 537–553.
3. Filter, W.F., Dunham, W.R., Polichar, R.M., Sands, R.H. & Harding, L.J. (1978) in *Frontiers of biological energetics*, vol. I (Dutton, P.L., Leigh J.S. & Scarpa, A., eds) pp. 603–616, Academic Press, New York.
4. Filter, W.F. (1983) PhD thesis (Physics) University of Michigan, University Microfilms 8324177.
5. Greenwood, N.N. & Gibb, T.C. (1971) *Mössbauer spectroscopy*, Chapman and Hall, London.
6. Frauenfelder, H. (1962) *The Mössbauer effect*, W.A. Benjamin, New York.
7. Ure, M.D.C. & Flinn, P.A. (1971) *Mössbauer Effect Methods* 7, 245–262.
8. Blume, M. & Kistner, O.C. (1968) *Phys. Rev.* 171, 417–425.
9. DeBenedetti, S., Barros, F. deS. & Hoy, G.R. (1966) *Annu. Rev. Nuclear Sci.* 16, 31–88.
10. Evans, R.D. (1955) *The atomic nucleus*, p. 711, McGraw-Hill, New York.
11. Grodstein, G.W. (1957) *National Bureau of Standards Circular* 583, US Government Printing Office, Washington DC.
12. Evans, R.D. (1955) *The atomic nucleus*, p. 785, McGraw-Hill, New York.
13. Dunham, W.R., Hagen, W.R., Braaksma, A., Grande, H.J. & Haaker, H. (1985) *Eur. J. Biochem.* 146, 497–502.
14. Housley, R.M. (1965) *Nuclear Instrum. Methods* 35, 77–82.
15. Dunham, W.R., Fee, J.A., Harding, L.J. & Grande, H.J. (1980) *J. Magn. Reson.* 40, 351–359.