The problem of error in deconvolution

1970 J. Phys. A: Gen. Phys. 3 462

(http://iopscience.iop.org/0022-3689/3/5/002)

J. Phys. A: Gen. Phys., 1970, Vol. 3. Printed in Great Britain

The problem of error in deconvolution

A. F. JONES† and D. L. MISELL‡§

† Department of Mechanics, The Johns Hopkins University, Baltimore, Maryland 21218, U.S.A.

‡ J. J. Thomson Physical Laboratory, University of Reading, Reading, Berkshire, England

MS. received 2nd April 1970

Abstract. The subject of this paper is the solution of the convolution integral, which relates the observed line profile to the true profile which is obtained in the absence of instrumental resolution effects. In the case of perfectly accurate data this problem has (by assumption) a well-defined solution. This paper considers the practical problem that results when the experimental curve is subject to error. Estimates are given for the importance of the error and a criterion is developed for optimizing the deconvolution procedure. This leads to a deconvolution function which when convoluted with the experimental curve gives the true solution. Finally the method of solution by iteration is examined and the effect of error upon it studied.

1. Introduction

Frequently in experimental physics, and elsewhere, the following integral, known as the convolution integral, occurs:

$$g(x) = \int_{-\infty}^{\infty} f(u)\, S(x-u)\, \mathrm{d}u. \tag{1}$$

The function $f(x)$ represents a spectrum of some kind which depends on a one-dimensional variable x and it is this spectrum that one is attempting to measure. True values for the spectrum are not found, however, because of the presence of distortion. This usually takes the form of resolution effects in the recording instrument (as in, for example, spectroscopy, optics and electron optics). Instead of measuring $f(x)$ one obtains an experimental curve $g(x)$, which is a weighted average of the true curve. Equation (1) defines this averaging mathematically, where $S(x)$ is the weighting function (or instrument broadening function) for the particular experiment being considered; $S(x)$ is usually a quantity that can be determined experimentally (see, for example, Stokes 1948).

With $S(x)$ known, the task that one very often faces is to deconvolute the experimental curve $g(x)$ and hence obtain the true function $f(x)$ that would be observed in the absence of instrument resolution effects. Because the convolution integral occurs so frequently, a very large number of papers have been published in the literature which, directly or indirectly, involve this problem (e.g. Burger and van Cittert 1932, 1933, Fellgett and Schmeidler 1952, Burr 1955, Skarsgard *et al.* 1961, Morrison 1963, Jones *et al.* 1967, Berreman 1968, Louër *et al.* 1969). However, the number of distinct methods of solution that they contain appears quite limited. We shall now briefly enumerate and describe these methods.

The first and commonest method of solution is by iteration (e.g. Burger and van Cittert 1932, 1933, Skarsgard *et al.* 1961, Rollett and Higgs 1962, Khidir and Decius

§ Present address: Department of Physics, Queen Elizabeth College, Campden Hill Road, London W8, England.

1962, Jones *et al.* 1967, Ergun 1968). The experimental curve $g(x)$ is taken as an initial approximation to $f(x)$ and a scheme is set up to continuously improve the approximation; details of this method are discussed in § 4. Normally only three or four iterations are recommended since it is observed empirically that the solution begins to diverge if more are attempted. The second method is by the use of a Fourier transform (e.g. Smith 1934, Stokes 1948, Rollett and Higgs 1962); this changes equation (1) into an algebraic equation which is easily solved for $\bar{f}(t)$, the Fourier transform of $f(x)$. The final solution for $f(x)$ is then obtained by an inverse Fourier transform. Whilst this solution is ideally exact, it has practical drawbacks, namely (i) one only has values for $g(x)$ over a finite (not an infinite) range of x, (ii) transforms are laborious to calculate numerically, (iii) noise on the experimental profile $g(x)$ produces high-frequency oscillations on the Fourier transform $\bar{g}(t)$ and, (iv) $\bar{f}(t)$ is all too often given by an expression which has zeros in its divisor for large values of t. An extension of this second method attempts to avoid some of these difficulties; the ideal Fourier transform solution is inverted to obtain an expression which involves the original functions $f(x)$ and $S(x)$ more directly (e.g. Hardy and Young 1949, Allen *et al.* 1964, Sauder 1966, Jones and Misell 1967, Berreman 1968). A third method is one which recognizes that it is only a finite range over which $g(x)$ is determined. A Fourier series expansion thus replaces the Fourier transform. Truncation of the series leads to a set of algebraic equations which can be solved to yield a solution (e.g. Moore 1968). Another method in quite common use is the representation of $g(x)$ and $S(x)$ by suitable analytic functions; the convolution integral is then solved explicitly, by Fourier inversion, to give $f(x)$ (e.g. Shull 1946, Schoening *et al.* 1952, Ruland 1965, 1968, Saksena *et al.* 1968). This method of solution for $f(x)$ is rather limited in that the line shapes for both $g(x)$ and $S(x)$ must be explicitly given; many line profiles defy curve fitting by simple analytic expressions. An alternative way of writing down the convolution integral, when $g(x)$ is known at a series of points, is as a set of linear equations and this set of equations is solved by matrix inversion, for $f(x)$ (e.g. Paterson 1950, Louër *et al.* 1969). The use of Hermite polynomials, as an alternative to Fourier series, for $g(x)$ and $S(x)$ has been used to give an explicit solution for $f(x)$ in terms of the coefficients of the Hermite polynomials (e.g. Berry 1947, Hossfeld 1968). However, it should be pointed out that the representation of $g(x)$ as a sum of Hermite polynomials is not of general application; the actual line shape for $g(x)$ is a critical factor in the success of this method. Finally, the method of moments in which $S(x)$ is represented as a polynomial expansion gives a solution for $f(x)$ in terms of the moments of $g(x)$ (e.g. Flynn and Seymour 1960, Eastabrook and Wilson 1952, Young *et al.* 1967, Sauder 1966); a source of error in the solution for $f(x)$ is due to the unpredictable behaviour of the higher-order moments of $g(x)$, when $g(x)$ is subject to errors.

All these methods have one feature in common; some form of approximation is always made. In fact, a simple argument quickly demonstrates that an approximation is a necessary part of a method of solution and that in general equation (1) need not have any solution at all. The argument is as follows: The effect of the averaging described by the convolution integral is to smooth out any irregularities, or roughness, in the original curve $f(x)$. Only if some very large irregularities exist in $f(x)$ would any roughness still be visible in the experimental curve $g(x)$. Unfortunately, $g(x)$ always contains some small rough variations, which do not actually originate from $f(x)$ but are the effects of random background noise. However, unless one can distinguish between the true experimental curve and the noise, the irregularities can only be

interpreted as very large (but spurious) variations in the original distribution $f(x)$. Furthermore, if $g(x)$ is sufficiently rough, the corresponding variations in $f(x)$ are indefinitely large, and indeed, examples can be constructed where it can be shown that there is no solution at all of the convolution integral, e.g. if $S(x) = \sin ax/x$ its Fourier transform $\bar{S}(t)$ is zero for $|t| > a$, which yields a contradiction if $g(x)$ is allowed to be general function. So we may conclude that equation (1) cannot necessarily be inverted at all.

Such arguments as the above, whilst correct, represent extremes. For example, suppose we make the not unrealistic assumption that $f(x)$ is piecewise analytic whilst $S(x)$ is a Gaussian distribution. It then follows that $g(x)$ is an analytic function. Moreover, one of the properties of an analytic function is that its value, over the complete range of the variable x, can be determined from knowledge of its value over an indefinitely small range of x (the property of analytic continuation). Thus we are led to conclude that it should be possible to determine $f(x)$ from a knowledge of $g(x)$ over any indefinitely small range of x. The fallacy here is clear: we are applying the mathematical formula (1) to real situations without any regard to the practical limitations that one is subject to in real life. In particular, the amount of information that can be gleaned from an empirical curve is severely limited and the above argument implicitly assumes that the values of the curve $g(x)$ can be read to any specified degree of accuracy at all points of x.

Similarly we cannot conclude that there exists no practical deconvolution method by which equation (1) can be solved to give $f(x)$. Although a strict mathematical solution appears forbidden, the correlation between the mathematics and reality has always to be kept in mind; and so we have the problem of optimal-deconvolution, i.e. the determination of the maximum amount of information that can be extracted from a realistic curve $g(x)$ about the true curve $f(x)$ and some knowledge concerning the accuracy of this information. These problems, although of importance, have not been given much attention previously in the literature. Even in this paper we shall not be examining the full problem because of all the difficulties involved. Instead, a partly idealized situation is considered, namely, we assume that the experimental curve, which includes the true convoluted curve plus a small amount of random noise, is known, in principle, exactly over the complete range of x, $-\infty < x < \infty$. We then try to ascertain how much of this information is relevant to the determination of $f(x)$ (§ 2).

It is relevant to point out that this model makes several important and interrelated approximations. The most unrealistic of these is that $g(x)$ is known at all points x. In practice, deconvolution is always attempted from a much more limited amount of data (typically 30–70 readings for a one-dimensional curve). We are assuming that we have far more information about the curve than is actually the case. Further, it is not possible to take readings from the curve $g(x)$ with 100% accuracy as we are assuming. This last point is not too important, since we can always attribute any error in the readings to additional background noise, although this interpretation partly obscures a problem that does exist, namely, the question of at what intervals of x should readings of the experimental curve be taken. Clearly they should not be too far apart or information is lost but it is not clear what can be gained by taking readings indefinitely close together. Lastly it is over a finite, not an infinite, range of x that $g(x)$ is determined. Theoretically this may seem unimportant (see the argument on analytic continuation earlier) but in practice, owing to the restriction of describing $g(x)$ by a finite number of points, it means that less information is known about $f(x)$

near the end points of the range than elsewhere. This fact is always reflected in the results of 'laboratory tests' on the methods of solution described earlier. When a curve $g(x)$, for which the true spectrum $f(x)$ is known for comparison purposes, is deconvoluted, the answer always has its maximum variance from the true answer near the end points of the range. No theoretical estimate for the magnitude of this variance has been proposed, however.

§ 3 of this paper describes a deconvolution function, which replaces $S(x)$ in the convolution integral, equation (1). This function $M(x)$ allows for the effects of errors in $g(x)$ and it is proposed that $M(x)$ gives an optimum solution for $f(x)$. We also discuss the optimization of the solution for $f(x)$ for two specific methods of deconvolution, namely Fourier transforms (§ 3) and iterative procedures (§ 4). Whilst for illustrative purposes we have used Gaussian profiles for $S(x)$ and $g(x)$, the equations derived, relating to the optimization of $f(x)$, are of a general applicability.

2. True information contained by an experimental curve

The arguments that follow rest largely on the physical interpretation of the Fourier transform and so we present here a brief review of the relevant points we shall be using. The Fourier transform for a function $F(x)$ is defined by

$$\bar{F}(t) = \int_{-\infty}^{\infty} F(x) \exp(-itx)\, dx \tag{2}$$

and its inverse transform by

$$F(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} \bar{F}(t) \exp(itx)\, dt. \tag{3}$$

The Fourier transform can be regarded as the limit of a Fourier series when the range over which it is applied becomes infinite. With this interpretation, the integral (2) defines a *coefficient* $\bar{F}(t)$ while the integral (3) defines a *summation* of harmonics $\exp(it_0x)$ for various values of t_0. These harmonics are weighted by the factor $\bar{F}(t_0)/2\pi$ and then equation (3) indicates that the sum of all these harmonics, weighted in this fashion, is equal to the original function $F(x)$. However, since the range of x is infinite, there is a continuous range of values for t_0 instead of a discrete set and then equation (3) is written not as a summation but as its limiting form, an integral.

Taking the Fourier transform of equation (1) gives

$$\bar{g}(t) = \bar{f}(t)\, \bar{S}.(t) \tag{4}$$

since, as is well known, the convolution integral becomes a product in transform space (see e.g. Stokes 1948). Consider now the interpretation of this equation; $\bar{g}(t)$ represents the components of the harmonics which comprise the experimental curve $g(x)$, while $\bar{f}(t)$ represents the components of the harmonics which comprise the theoretical curve $f(x)$. $\bar{g}(t)$ and $\bar{f}(t)$ are related by a modification factor $\bar{S}(t)$, which multiplies $\bar{f}(t)$ to make it equal to $\bar{g}(t)$. Thus $g(x)$ contains the same harmonics as $f(x)$ but in an altered ratio, altered in fact in proportion to $\bar{S}(t)$.

In order to demonstrate the significance of this, suppose $S(x)$ is a normalized Gaussian profile, i.e.

$$S(x) = \frac{\alpha}{\pi^{1/2}} \exp(-\alpha^2 x^2) \tag{5}$$

so that

$$\bar{S}(t) = \exp\left(\frac{-t^2}{4\alpha^2}\right) \tag{6}$$

and further let the percentage noise in the experimental curve $g(x)$ be $\nu\%$ (by which we mean that the root-mean-square of the noise, or instrument error, is $\nu\%$ of the maximum value of the true experimental curve). Then if we make the assumption that the percentage noise in the transformed experimental curve $\bar{g}(t)$ is also $\nu\%$ (see below), it can be seen that there is a limit to the amount of useful information contained in $\bar{g}(t)$ about the desired transform curve $\bar{f}(t)$. For imagine that $g(x)$ has been measured in some experiment and its Fourier transform obtained, then all the values of $\bar{g}(t)$ for t greater than some t_* would be less than the value that we are allowing for the noise. This is because $\bar{g}(t)$ is the value of the theoretical curve $\bar{f}(t)$ after multiplication by the exponential factor of equation (6) and, if t_* is selected to be sufficiently large, this product can be made indefinitely small. Thus for high harmonics the true information contained in $\bar{g}(t)$ (and hence the original $g(x)$) is less than the misinformation contained about the noise, and values of $\bar{g}(t)$ for $|t| > t_*$ cannot be depended on.

Looking at this in another way, the inversion of equation (4) yields

$$\bar{f}(t) = \bar{g}(t)\, T(t) \tag{7}$$

where

$$T(t) = \frac{1}{\bar{S}(t)}.$$

Thus if we insist on attaching importance to the high harmonics, i.e. assume the values of $\bar{g}(t)$ are correct for large t, then we obtain an erroneous end result, since we are obliged to multiply $\bar{g}(t)$ by an exponentially large factor, which raises the noise to an intolerable level. (This, incidentally, is the reason that all methods of deconvolution require some form of approximation, or smoothing of the original data, to obtain the solution. An exact solution, if it exists, would be dominated by the noise and so noise has to be reduced in importance by some procedure.) We conclude that there exists a cut-off frequency t_* beyond which $g(x)$ contains no reliable information about the harmonic components of $f(x)$. This is not meant to imply that the data for $|t| < t_*$ are perfect; the noise is still present and causes some uncertainty in this region. The point is that for $|t| > t_*$, the uncertainty approaches 100% and, if not corrected for, this uncertainty leads to large, spurious variations in the correct solution.

The value of t_* will depend on the particular experiment and the exact forms of $S(x)$ and $g(x)$. Three possible cases can be distinguished: (i) $f(x)$ varies slowly in comparison with $S(x)$, (ii) $f(x)$ varies over a length scale comparable with the length scale of $S(x)$, and (iii) $f(x)$ varies quickly in comparison with $S(x)$. In the first of these cases, the transform of $f(x)$ will decrease sharply to zero in comparison with the transform of $S(x)$ and so the noise level (i.e. the fraction $\nu/100$) is reached primarily as a result of the decrease in the value of $\bar{f}(t)$ and not because $\bar{S}(t)$ has become small. In this case there is no essential loss of information owing to convolution. (There is always *some* loss of information about $f(x)$ because of the presence of noise in $g(x)$; the distinction here is whether additional information has been lost because of the convolution effect or whether the same amount of information would be lost if the experiment could be reproduced with the same noise present but without any convolution.) Case (iii) is the converse of this situation. $\bar{S}(t)$ is a much sharper function

than $\tilde{f}(t)$ and as a result the convolution of $f(x)$ with $S(x)$ to form $g(x)$ causes substantial loss of information about the higher harmonics of $f(x)$. Since in this case $g(x)$ does not contain the information to regenerate $f(x)$ accurately, there is no point in trying to deconvolute $g(x)$; one can only attempt to improve the experimental equipment. Case (ii) is the transition state between these two extremes. There is some loss of information about $f(x)$, since the product $\tilde{f}(t) \cdot \tilde{S}(t)$ reaches the noise level more quickly than $\tilde{f}(t)$ would do alone. On the other hand, one can agree to accept this loss and derive an answer which is an approximation to $f(x)$. An important point to note, however, is that cases (ii) and (iii) cannot be distinguished from each other by experiment; in case (iii), $g(x)$ will contain no high harmonics since they are eliminated by convolution of $f(x)$ with $S(x)$ and thus $g(x)$ will vary on a length scale comparable with the length scale of $S(x)$. However, this is also the characteristic which defines case (ii). Thus a deconvolution, when the length scales of $g(x)$ and $S(x)$ are similar, should be attempted only if one can be sure by other means that one is dealing with case (ii) and not case (iii). Even then, one risks the loss of the finer details in $f(x)$.

We consider now the convolution problem, where the value of the cut-off is determined only by $g(x)$. For simplicity, we first consider a $g(x)$ which is a Gaussian distribution centred on $x = 0$, i.e.

$$g(x) = A \exp(-\beta^2 x^2) \tag{8}$$

and

$$\tilde{g}(t) = \frac{A\pi^{1/2}}{\beta} \exp\left(\frac{-t^2}{4\beta^2}\right). \tag{9}$$

We can define the cut-off point as the value of t when $\tilde{g}(t)$ decreases to the noise level. Our assumption that the error in $\tilde{g}(t)$ is the same as that in $g(x)$, i.e. $\nu/100$, then yields $t_* = 2\beta\{\ln(100/\nu)\}^{1/2}$. However, since the definition of t_* is somewhat arbitrary and also since reasonable changes in ν do not alter the value of t_* greatly, it is convenient to assume $\nu \simeq 5$ (say) and so obtain the simpler equation $t_* \simeq 4\beta$. In this case, t_* is directly related to the halfwidth of the experimental function H_g by $t_* = 8(\ln 2)^{1/2}/H_g$ or

$$t_* \simeq \frac{6.5}{H_g}. \tag{10}$$

While this relation has been derived for a particular functional form (equation (8)), it can be seen that a similar cut-off criterion can be obtained in the general case for $g(x)$ that are likely to be encountered in practice. The argument then rests on order-of-magnitude considerations; when taking the Fourier transform of $g(x)$, the positive and negative parts of the integral will largely cancel if $\exp(-itx)$ oscillates sufficiently quickly compared with $g(x)$. If we let the dividing line be the point when $\exp(-itx)$ oscillates through a complete period over the halfwidth of $g(x)$, we again obtain a similar result to equation (10), approximating $2\pi \simeq 6{\cdot}3$. It should be emphasized, however, that this is only a convenient estimate for t_*. In practice the value of t_* depends on the individual experiment and the numerical factor in (10) should be altered to suitably reflect the accuracy of the experiment and the particular form of $g(x)$.

In order to obtain an estimate of the information lost because of the convolution effect, we again take equations (5) and (8) as typical functions. Then

$$\tilde{f}(t) = \frac{A\pi^{1/2}}{\beta} \exp\left\{-\frac{t^2}{4}\left(\frac{1}{\beta^2} - \frac{1}{\alpha^2}\right)\right\} \tag{11}$$

and it can be seen that the useful range in $\tilde{g}(t)$ is less than that of $\tilde{f}(t)$ by a factor of $\beta\{(1/\beta^2)-(1/\alpha^2)\}^{1/2}$ or, if $\beta \ll \alpha$,

fractional decrease in the useful range of t $$\simeq 1-\tfrac{1}{2}\left(\frac{\beta}{\alpha}\right)^2 = 1-\tfrac{1}{2}\left(\frac{H_S}{H_g}\right)^2 \qquad (12)$$

where H_S = halfwidth of $S(x)$.

3. A deconvolution function

In order to counter the problem of high-harmonic noise in $g(x)$, $T(t)$ in the deconvolution expression (7) should be replaced by another function $R(t)$, which is equal to $T(t)$ for $|t| < t_*$ but is defined in the remaining range in such a way as to reduce the importance of the noise. Hence we define

$$R(t) = \begin{cases} T(t) & |t| < t_* \\ 0 & |t| \geqslant t_*. \end{cases} \qquad (13)$$

Then

$$f(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} \tilde{g}(t)\, R(t) \exp(itx)\, dt \qquad (14)$$

or replacing $\tilde{g}(t)$ from its definition equation (2) as a Fourier transform and reversing the order of integration, we obtain

$$f(x) = \int_{-\infty}^{\infty} g(u)\, du \frac{1}{2\pi}\int_{-\infty}^{\infty} R(t) \exp\{it(x-u)\}\, dt. \qquad (15)$$

So if we can evaluate

$$M(x) = \frac{1}{2\pi}\int_{-\infty}^{\infty} R(t) \exp(itx)\, dt \qquad (16)$$

then equation (15) is an expression for $f(x)$ as a convolution, i.e.

$$f(x) = \int_{-\infty}^{\infty} g(u)\, M(x-u)\, du. \qquad (17)$$

This function can be computed and once obtained can be used as a deconvolution function for the particular weighting function $S(x)$ under consideration. For example, if $S(x)$ is the Gaussian equation (5),

$$\begin{aligned} M(x) &= \frac{1}{2\pi}\int_{-t_*}^{+t_*} \exp\left(\frac{t^2}{4\alpha^2}\right) \exp(itx)\, dt \\ &= \frac{1}{2\pi} t_* \int_{-1}^{+1} \exp\left(\frac{t_*^2}{4\alpha^2} T^2\right) \exp(it_* xT)\, dT. \end{aligned} \qquad (18)$$

However, because of the discontinuity in the function $R(t)$, $M(x)$ will behave as $\sin x/x$ for large x and so the convergence of equation (17) will be slow. Thus it is better to smooth out this discontinuity and define

$$R(t) = \begin{cases} T(t) & |t| < t_* \\ U(t) & |t| > t_* \end{cases} \qquad (19)$$

where $U(t)$ is defined so as to make $R(t)$ continuous at $t = t_*$ with its first n derivatives, and such that $U(t) \to 0$ as $t \to \infty$. $M(x)$ will then behave as $\sin x/x^{n+2}$ for large x

and convergence will be improved. For the Gaussian equation (5), we may choose $n = 1$ for example and define (matching the value and first differential at $t = t_*$)

$$U(t) = \exp\left[\frac{1}{4\alpha^2}\{-(|t|-2t_*)^2+2t_*^2\}\right]. \tag{20}$$

Note that since we have no reliable information about harmonics beyond $t = t_*$, it is as correct to choose $U(t) = 0$ as it is to select $U(t)$ as any other bounded function. The noise problem we are attempting to eliminate results from allowing $U(t)$ to be unbounded, which magnifies the noise beyond proportion.

These arguments hold for any weighting function $S(x)$ for which the Fourier transform can be obtained. The function $M(x)$ defined by equations (16) and (18) can always be derived as a set of numerical values and then used in equation (17) as a universal deconvolution function. It is obviously preferable if the Fourier transform of $S(x)$ can be found analytically but, even if $S(x)$ is only known numerically, this method of solution would be preferable to the original method of solving the convolution integral with $S(x)$; $R(t)$ has only to be calculated once and then deconvolution involves only a single set of integrations (instead of two sets or more by solving the original convolution integral (1) by Fourier transforms), while the importance of the high-harmonic noise is automatically reduced.

4. Analysis of the iterative method

Since a very common method of solution of the convolution integral is by iteration, we present an analysis of the method. Solution by iteration is normally attempted by defining

$$f_1(x) = g(x) \tag{21}$$

and

$$f_{n+1}(x) = f_n(x)+g(x)-f_n(x)*S(x) \tag{22}$$

where the * indicates the convolution of $f_n(x)$ with $S(x)$.

We shall determine the criterion for this iterative scheme to converge and show that if it does, the solution is unique and thus $f_n(x) \to f(x)$.

By induction, one can show

$$f_n(x) = g(x)+g(x)\sum_{r=1}^{r=n-1}\{1-*S(x)\}^r \tag{23}$$

or taking Fourier transforms of both sides of equation (23),

$$\bar{f}_n(t) = \bar{g}(t)\sum_{r=0}^{r=n-1}\{1-\bar{S}(t)\}^r. \tag{24}$$

Summing the series on the right-hand side of equation (24) gives

$$\bar{S}(t)\bar{f}_n(t) = \bar{g}(t)-\bar{g}(t)\{1-\bar{S}(t)\}^n. \tag{25}$$

If now one assumes that solution to the original convolution problem exists, and furthermore that the solution is unique, then $f_n(x) \to f(t)$, i.e. $f_n(x) \to f(x)$ (except perhaps at a few discrete points, which we assume is irrelevant for practical problems) as $n \to \infty$ if and only if

$$\{1-\bar{S}(t)\}^n \to 0 \qquad \text{as} \qquad n \to \infty \tag{26}$$

i.e.

$$|1-\bar{S}(t)| < 1. \tag{27}$$

As is known, however, for a practical solution of the convolution problem, high-harmonic noise must be eliminated. Thus if equation (25) has *heuristic* convergence, then we must have that

$$\{1-\bar{S}(t)\}^n = \begin{cases} 0 & |t| < t_* \\ 1 & \text{for sufficiently large } t. \end{cases}$$

Although this criterion for convergence is not widely known, circumstances are such that most authors, who have used this method of deconvolution, in the past have satisfied it. This is because it is conventional to use a normalized weighting function when deconvoluting, i.e.

$$\int_{-\infty}^{\infty} S(x)\,\mathrm{d}x = 1.$$

Thus it follows that $\bar{S}(0) = 1$ and consequently in the immediate vicinity of $t = 0$, $\{1-\bar{S}(t)\}$ is always small. Thus $\{1-\bar{S}(t)\}^n \to 0$ for $|t| < t_*$ if t_* is sufficiently small. The other requirement that $\{1-\bar{S}(t)\}^n \to 1$ for large t is a consequence of the fact that $\bar{S}(t)$ is by nature small for large t. Thus if n is not selected to be too large (i.e. $n \ll 1/\bar{S}(t)$), we have

$$\{1-\bar{S}(t)\}^n \sim 1-n\bar{S}(t)$$

for large t, and so equation (25) reduces to

$$\bar{f}_n(t) = n\bar{g}_n(t)$$

for that region of t. Thus $f_n(x)$ contains the erroneous high-harmonics of $g_n(x)$ multiplied by the factor n. This is not necessarily disastrous (unlike exact deconvolution when they are amplified many times—see § 3), but nevertheless it is an undesirable feature of the method and one should attempt to restrict the number of iterations $i(=n-1)$ to a minimum. Indeed, that the method diverges if too many iterations are attempted is well known from empirical work (Khidir and Decius 1962, Jones *et al.* 1967, Ergun 1968, Jones and Misell—unpublished).

The criterion that establishes the minimum number of iterations that are necessary to generate the solution again can be derived from equation (25). It is that $\{1-\bar{S}(t)\}^n$ must be less than the experimental error, in the range of useful information in $g(x)$, i.e.

$$\{1-\bar{S}(t)\}^n < \frac{\nu}{100} \qquad \text{for all} \qquad |t| < t_*$$

or if we assume that the imaginary part of $\bar{S}(t)$ is negligible while the real part is monotonic decreasing either side of $t = 0$ in the relevant range (which is true for a symmetric $S(x)$), then

$$\bar{S}(t_*) > 1-\left(\frac{\nu}{100}\right)^{1/n}. \tag{28}$$

Since t_* is inversely proportional to the halfwidth H_g of $g(x)$ and since $\bar{S}(t)$ varies on a scale inversely proportional to the halfwidth H_S of $S(x)$, equation (28) relates the minimum number of iterations necessary to the ratio of the halfwidths of $g(x)$ and $S(x)$ when they are Gaussians. Unfortunately, the numerical values in the relationship depend on the form of $\bar{S}(t)$ and so it is difficult to generalize further. However,

for any individual experiment, $\bar{S}(t_*)$ can actually be computed and then used to determine the minimum value for the number of iterations. For the Gaussian distributions equations (5) and (8), one obtains the results shown in table 1.

Table 1. The minimum number of iterations i required to solve the convolution integral for various ratios between the halfwidths of the experimental curve $g(x)$ and the broadening function $S(x)$ when they are both Gaussian profiles and the experimental error $\nu = 2$.

Minimum number of iterations, i	Ratio of the halfwidths of $g(x)$ and $S(x)$
$i = 0$	$H_g/H_S > 15$
$i = 1$	$H_g/H_S > 5$
$i = 2$	$H_g/H_S > 4$
$i = 3$	$H_g/H_S > 3$

5. Conclusions

The intentions of this paper are:

(i) To draw attention to the fact that limits exist to the application of deconvolution in the resolution of spectra, and to try and indicate under what conditions one might reasonably hope to obtain a solution and under what conditions one should concentrate on improvement of the experimental equipment. There are, in fact, methods of deconvolution described in the literature which imply that details can be resolved in the spectrum on a length scale much less than that of the instrument function. This is misleading, however, since this can only be achieved in laboratory tests which use theoretical curves whose values are indefinitely accurate.

(ii) When deconvolution is permissible, to try to assess the accuracy with which it can be achieved.

(iii) To present an optimized method of deconvolution (which takes account of experimental error) in the ideal case where the experimental curve is known for all values of x.

(iv) To stimulate, we hope, further research in this area. The principal problem now seems to be how the knowledge of the experimental curve over only a finite instead of infinite range affects the solution.

Acknowledgments

One of the authors (D.L.M.) is grateful to the University of Reading Computing Unit for providing excellent facilities, and the other (A.F.J.) wishes to acknowledge the receipt of a grant under contract number ONR 4010(02).

References

ALLEN, L. C., GLADNEY, H. M., and GLARUM, S. H., 1964, *J. chem. Phys.*, **40**, 3135–41.
BERREMAN, D. W., 1968, *Appl. Opt.*, **7**, 1447–53.
BERRY, C. R., 1947, *Phys. Rev.*, **72**, 942–7.
BURGER, H. C., and VAN CITTERT, P. H., 1932, *Z. Phys.*, **79**, 722–30.
—— 1933, *Z. Phys.*, **81**, 428–34.
BURR, E. J., 1955, *Aust. J. Phys.*, **8**, 30–53.
EASTABROOK, J. N., and WILSON, A. J. C., 1952, *Proc. Phys. Soc.* B, **65**, 67–75.
ERGUN, S., 1968, *J. appl. Crystallogr.*, **1**, 19–23.
FELLGETT, P. B., and SCHMEIDLER, F. B., 1952, *Mon. Not. R. Astr. Soc.*, **112**, 445–51.

FLYNN, C. P., and SEYMOUR, E. F. W., 1960, *Proc. Phys. Soc.*, **75,** 337–44.
HARDY, A. C., and YOUNG, F. M., 1949, *J. Opt. Soc. Am.*, **39,** 265–70.
HOSSFELD, F., 1968, *Acta crystallogr.*, **A24,** 643–50.
JONES, A. F., and MISELL, D. L., 1967, *Br. J. appl. Phys.*, **18,** 1479–83.
JONES, R. N., VENKATARAGHAVAN, R., and HOPKINS, J. W., 1967, *Spectrochim. Acta*, **23A,** 925–39.
KHIDIR, A. L., and DECIUS, J. C., 1962, *Spectrochim. Acta,* **18,** 1629–39.
LOUËR, D., WEIGEL, D., and LOUBOUTIN, R., 1969, *Acta crystallogr.*, **A25,** 335–8.
MOORE, L., 1968, *J. Phys.* D: *Appl. Phys.*, **1,** 237–45.
MORRISON, J. D., 1963, *J. chem. Phys.*, **39,** 200–7.
PATERSON, M. S., 1950, *Proc. Phys. Soc.*, **A63,** 477–82.
ROLLETT, J. S., and HIGGS, L. A., 1962, *Proc. Phys. Soc.*, **79,** 87–93.
RULAND, W., 1965, *Acta crystallogr.*, **18,** 581.
—— 1968, *J. appl. Crystallogr.*, **1,** 90–101.
SAKSENA, B. D., AGARWAL, K. C., PAHWA, D. R., and PRADHAM, M. M., 1968, *Spectrochim. Acta*, **24A,** 1981–98.
SAUDER, W. C., 1966, *J. appl. Phys.*, **37,** 1495–508.
SCHOENING, F. R. L., VAN NIEKERK, J. N., and HAUL, R. A. W., 1952, *Proc. Phys. Soc.*, **B65,** 528–35.
SHULL, C. G., 1946, *Phys. Rev.*, **70,** 679–84.
SKARSGARD, L. D., JOHNS, H. E., and GREEN, L. E. S., 1961, *Radiat. Res.*, **14,** 261–80.
SMITH, L. P., 1934, *Phys. Rev.*, **46,** 343–51.
STOKES, A. R., 1948, *Proc. Phys. Soc.*, **61,** 382–91.
YOUNG, R. A., GERDES, R. J., and WILSON, A. J. C., 1967, *Acta crystallogr.*, **22,** 155–62.