A Least-Squares Method for Unfolding Convolution Products in X-ray Diffraction Line Profiles

Fumiyoshi Yokoyama*

(Received January 27, 1982)

Synopsis

A deconvolution method for the X-ray diffraction line profile is proposed, which is based on the conventional least-squares method. The true profile is assumed to be a functional form. The numerical values of parameters of the function assumed are determined so that the calculated profile, which is a convolution of the function and the instrumental profile, has a minimum deviation from the observed one. The method is illustrated by analysis of the X-ray powder diffraction profile of sodium chloride 222 reflexion under the assumption that the true profile is a Gaussian or a Cauchy function.

1. Introduction

The broadening of the X-ray diffraction line profile involves valuable information about the crystallite-size and the lattice distortion in a specimen. The observed broadening, however, has no value until the correction, i.e., deconvolution or unfolding, is made by the instrumental broadening inherent in the X-ray optical system. The deconvolution methods, therefore, have been proposed and developed by many investigators. Among them, the Fourier transform method proposed

* Department of Industrial Chemistry
by Stokes' [1] which is based on the mathematically strict relationship between the convolution and its Fourier transform has been widely employed. In this method, however, experimental errors often result in spurious sinusoidal components in the true profile, and the details of the shape become unreliable. Besides the Fourier transform method, the successive unfolding method [2] is useful, although this requires that data should have a high statistical accuracy and that the function desired is everywhere differential.

This communication deals with a deconvolution method which is an application of the conventional least-squares method to the convolution relation in the X-ray diffraction measurements. The analytical expressions and the practical procedures of the method will be described systematically and in detail.

2. Basis of the Analysis

In the X-ray diffraction, the observed profile $h$ is a convolution of the true profile $f$ and the instrumental one $g$, in the form

\[ h = f * g \]  

First, let $f$ be a function with a variable $x$ and the parameters $p_j$ ($j=1, 2, \ldots, n$) and have the bounded support

\[ f(x, p_1, p_2, \ldots, p_n) = 0; \quad x \leq a_1, \quad x \geq a_2 \]  

Let the height of $g$ at $x$ be denoted as $g(x)$ and normalized with the bounded support

\[ \int_{b_1}^{b_2} g(x) \, dx = 1 \]  

\[ g(x) = 0; \quad x \leq b_1, \quad x \geq b_2 \]  

Then, equation (1) can be rewritten in the integral form

\[ h(x) = \int_{a_1}^{a_2} f(y, p_1, p_2, \ldots, p_n) g(x-y) \, dy \]  

\[ h(x) = 0; \quad x \leq c_1, \quad x \geq c_2 \]  

where $p_1$ denotes the position of $f$ and the other parameters $p_j$ ($j \geq 2$) are assigned for the profile constants such as width, height and so on. Then, we have the relations about the limits

\[ c_1 = a_1 + b_1, \quad c_2 = a_2 + b_2 \]
Equation (4) is illustrated in Fig. 1-(a).

Secondary, if $f$ is fixed at a certain point, i.e., $p_i$ in $f = \text{constant}$, $f$ has the parameters $p_j(j \geq 2)$. Then, the shift of $g$ from its origin should be taken into account for the shift of $h$ along the abscissa. Thus, in this case, we denote $p_i$ as a parameter assigned for the shift of $g$, and have the expression

$$h(x) = \int_{a_1}^{a_2} f(y, p_2, p_3, \ldots, p_n) g(x-y, p_i) dy$$  \hspace{1cm} (6)

where $f, g$ and $h$ have the same corresponding definitions as in the first case. Equation (6) is illustrated in Fig. 1-(b).

Now, equation (1) is replaced by the form

$$h = g * f$$  \hspace{1cm} (7)

According to the above consideration, we can readily understand that this gives two integral equations

$$h(x) = \int_{b_1}^{b_2} g(y, p_i) f(x-y, p_2, p_3, \ldots, p_n) dy$$  \hspace{1cm} (8)

and

$$h(x) = \int_{b_1}^{b_2} g(y) f(x-y, p_2, p_3, \ldots, p_n) dy$$  \hspace{1cm} (9)

whose graphical images will correspond to Fig. 1-(a) and Fig. 1-(b), respectively, though the exchanges of symbols should be necessarily made.

![Fig. 1. Schematic representations of convolution relations. (a): equation (4); (b): equation (6).](image-url)
Equations (4), (6), (8) and (9) are equivalent of one another; there is only a difference that the parameter \( P \) for the shift of \( h \) is set in \( f \) or in \( g \). Any of them will give the same shape of \( f \), and the same shape and position of \( h \), when the curve fitting of the calculated profile \( h \) to the observed one is performed by the least-squares method presented here. An incorrect position of \( f \) in (6) and (8), or that of \( g \) in (4) and (9), however, misleads the final position of \( g \) or \( f \), respectively.

In the least-squares method, we know that the partial derivatives of the calculated profile \( h \) with respect to the parameters \( P \) are required for the construction of the normal equations. If both \( f \) and \( g \) are represented by the functional forms, the derivative forms of \( h \) will be obtained. Then, any of above expressions about \( h \) will become available for the purpose. In this study, \( f \) is assumed to be a functional form, while \( g \) is an observed profile, which is asymmetrical around the peak and will not be exactly replaced by a single function (see Fig. 3), though the function of an asymmetrical form for the X-ray diffraction line profiles has been proposed [3]. From such considerations, it is concluded that (4) and (9) are more suitable for practical use than (6) and (8); either \( g \) in the formers doesn’t have the parameter \( P \) and its derivative with respect to \( P \) is not required. Here, let us determine to apply only equation (4) to the true profile analysis. From (4), we have the formula of the calculated profile \( h \) which is convenient for the performance of the integration

\[
h(x_i)_{\text{calc.}} = \sum_k f(x_i, p_1, p_2, \ldots, p_n) g(x_i - y_k) \Delta y \tag{10}
\]

where index \( i \) refers to the data point of the observed profile. Then the partial derivative of (10) becomes

\[
\frac{\partial h(x_i)_{\text{calc.}}}{\partial p_j} = \sum_k \frac{\partial f(x_i, p_1, p_2, \ldots, p_n)}{\partial p_j} g(x_i - y_k) \Delta y \tag{11}
\]

The least-squares procedure involves the minimization of the function

\[
S = \sum_i W_i \left[ h(x_i)_{\text{obs.}} - h(x_i)_{\text{calc.}} \right]^2 \tag{12}
\]

from which we can derive finally the normal equations, in which (10) and (11) are used (see Appendix A). \( W_i \) in (12) is a weight of the \( i \) th observed data \( h(x_i)_{\text{obs.}} \). By the iterative solvings of the normal equations, we can obtain a set of the optimum parameters \( P \) of the true profile \( f \). The reliability factors used are
A Least-Squares Method for Unfolding Convolution Products in X-ray Diffraction Line Profiles

\[ R_1 = \frac{\sum_{i} |h(x_i)_{\text{obs}} - h(x_i)_{\text{calc}}|}{\sum_{i} h(x_i)_{\text{obs}}} \]  
(13)

and

\[ R_2 = m\sqrt[2]{\frac{\sum_{i} [h(x_i)_{\text{obs}} - h(x_i)_{\text{calc}}]^2}{m}} / \sum_{i} h(x_i)_{\text{obs}} \]  
(14)

3. Practice of the Analysis

The practical procedures needed in the analysis of the powder diffraction profile of sodium chloride 222 reflexion by the present least-squares method are described.

3.1 Preparation of the Profiles \( h \) and \( g \)

As a standard material to give the instrumental broadening, quartz was selected. We adopted the internal-standard method for the measurements, as this guarantees that both materials are subjected to the same geometrical conditions. In this experiment, the homogenous powder mixture of sodium chloride (purity 99.99%, as milled to fine powder) and quartz (smaller than 400 mesh, but removed of smallest fraction by suspension in methanol [4]) in the weight ratio 1:1 was prepared, giving about the same order of intensities of both reflexions mentioned later. The measurements were performed by Ni-filtered Cu K\( \alpha \) radiation and the usual optical arrangement under the symmetrical reflexion method.

Even if the sample preparation as indicated above is made, a 2\( \theta \) dependence of the instrumental broadening can not be avoided because of a complex broadening effect by many factors: flat specimen, transparency, misalignment and so on [5]. Preceding to the analysis, the integral breadths and the half-widths of the reflexions from quartz in the mixture, in the different 2\( \theta \) ranges were evaluated after the \( \Delta_0, \Delta_0 \) doublet separation and subsequent subtraction of the background scatterings by the conventional straight line method (Fig. 2). It is found that there exists a 2\( \theta \) dependence of the instrumental profile. Consequently, in this study, both quartz 11·2 (2\( \theta \) = 50.1°) and 12·1 (2\( \theta \) = 60.0°) (Fig. 3) were adopted as the instrumental profiles so that sodium chloride 222 (2\( \theta \) = 56.5°) was situated close to and at the
Fig. 2. Integral breadth (---) and half-width (-----) of powder reflexions of quartz versus 2θ.

Fig. 3. Powder diffraction line profiles of quartz 11·2 and 12·1.

middle of them.

The observed profile of sodium chloride 222, hereafter denoted as \( h_{\text{obs}} \) or \( h \), was prepared by the same procedures as applied in quartz \( g \). The boundaries of these profiles were defined in accordance with (3a) and (4a). Then the data of \( h_{\text{obs}} \) consisted of \( m \) = 63 readings at equal interval 0.0125°. Those of \( g \) 11·2 and 12·1 consisted of 99 and 101 readings at the half of the former, 0.00625°, respectively. Both profiles of quartz were normalized to be unity and their Bragg angles
A Least-Squares Method for Unfolding Convolution Products in X-ray Diffraction Line Profiles

with respect to $\alpha_1$ radiation were turned to be origins of the diffraction geometry (Fig. 3).

3.2 Estimation of the Initial Values

In this section, let us show how the initial values of parameters $\beta_j (j=1,2,3)$ of the true profile $f$ are estimated when that is assumed to be a Gaussian

$$f = P_3 \exp\left[-\frac{(x)}{(P_3)^2}\right]$$

or a Cauchy

$$f = P_3 \left[1 + \frac{1}{(P_2)^2(P - P_3)^2}\right]$$

where commonly in both functions, $x$, $P_1$: peak position; $P_2$: integral breadth; $P_3$: peak height. Denoting these initial values as $P_1^o$, $P_2^o$ and $P_3^o$, they are estimated as follows.

(a) Peak Position

Let the angular position where the observed profile $h$ has a maximum be $\chi_{h,\max}$ and that for $g$ be $\chi_{g,\max}$, the initial position will be approximated well by

$$P_1^o = \chi_{h,\max} - \chi_{g,\max}$$

In this analysis, $\chi_{g,\max}$ is neglected, as that is very small (Fig. 3).

(b) Peak Height and Integral Breadth

These two parameters are correlated with each other by the relation

$$S_f = P_3^o P_2^o$$

which is an area under the entire curve (15) or (16) and should be equal to that of the observed profile $h$, $S_h$, as $g$ is normalized to be unity, i.e., $S_g = 1$ in equation (B3) (see Appendix B). Then, equation (18) is rewrittten by

$$S_h = P_3^o P_2^o$$

Here, let us consider the use of the initial estimates of the integral breadth $P_2^o$ derived from the three relations below

$$P_2^o(g) = \sqrt{(P_{2,h})^2 - (P_{2,g})^2}$$

$$P_2^o(c) = P_{2,h} - P_{2,g}$$
where both $P_{z,h}$ and $P_{z,g}$ are the observed integral breadths of $h$ and $g$, and equations (20) and (21) are based on the supposition that the related profiles would be a Gaussian and a Cauchy, respectively. By substituting the estimates obtained from the above three relations in (19), three pairs of $P^*$ and $P^*$ are determined. In the least-squares calculations, the integral breadth and the peak height are treated as an independent variable. In section 4, the initial estimates obtained by this method will be tabulated.

3.3 Notes on the Calculation

Equations (10) and (11) require the height of $g$ at a given point $(X'-Y_k)$. This problem was solved by the Legendre's extrapolation method, in which a set of four reading data $(X, g(x))$ adjacent to each $(X_k-Y_k)$ was used to give $g(X'_k-Y_k)$. In order to reduce as much the error resulting from the definite width $A Y$ in the calculation of (10) as possible, the Simpson's integration method was used, though this method was applied also in the normalization of $g$, since the profiles $f$ indicated later and $g$ had the relatively narrow lines and therefore their height near the peaks varied abruptly with an increment $A Y$ and the product of $f$ with $g$ in (10) was expected to have the similar tendency. But, this method was not applied in the calculation of (11), as the accuracy of numerical value of (11) in the normal equations is not so influential factor on the final results of the analysis as that of (10).

Now, when Gaussian and Cauchy functions are applied to $f$, we have to notice on the limits $a_1$ and $a_2$ of $Y_k$, because tails persist extensively in these functions, the more in the latter. Practically, definitions for $a_1$ and $a_2$ will be rather below than those in (5)

$$Q_1 = C_{1,\text{obs.}} - b_{2,\text{obs.}}, \quad a_2 = C_{2,\text{obs.}} - b_{1,\text{obs.}}$$

The first restriction of the interval of $Y_k$ in (10) and (11) was determined from (23)

$$a_1 \leq Y_k \leq a_2$$

In addition to this, the following point should be considered. As both the functions have a symmetrical form around the peak, we have the
relation
\[ p_1^I - \frac{\Delta}{2} \leq \chi_k \leq p_1^I + \frac{\Delta}{2} \] (25)

where \( p_1^I \) is a refined value of the peak position of \( f \) at the \( I \)th iteration in the least-squares calculation, and \( \Delta \) is an absolute value of the full interval of \( \chi_k \). In the present functions, we are forced to truncate the tails at a angular position, where it is necessary for the height of \( f \) to be sufficiently small compared with the peak height. The second restriction of \( \chi_k \) was thus determined as
\[
p_1^I - \frac{\Delta}{2} (f = p_1^{3/1000}) \leq \chi_k \leq p_1^I + \frac{\Delta}{2} \] (26)
in which \( \frac{\Delta}{2} (f = p_1^{3/1000}) \) means a position where the height of \( f \) is lowered down to one thousandth of the new peak height \( p_1^I \) at \( I \)th iteration.

In the computer program, both condition equations (24) and (26) were given in parallel.

4. Results and Discussion

The profile analysis of sodium chloride 222 was carried out on the four combinations of the instrumental profile \( \mathcal{G} \): quartz 11·2 or 12·1 and the true profile \( f \): Gaussian or Cauchy (Fig. 4 (a)-(d) and Table 1). All the weights \( w_k \) in equation (12) were set to be unity in any case. It was after four to seven iterations that a perfect convergence was accomplished. In the true profile Cauchy ((b) and (d)), the convergence is assured even when a integral breadth (or a peak height) is appreciably larger (or smaller) than its final value. On the other hand, in the true profile Gaussian ((a) and (c)), a relatively better approximated pair of the initial estimates is required than in the former. This may be ascribed to the character of the normal equations constructed in each case; the equations are approximated ones originated from linear equations of the truncated Taylor series expansion.

In any case of (a) to (d), the peak position is corrected nearly by the angle \( = -\chi_{g, \text{max}} \) (see also Fig. 3) to shift to a slightly higher angle and becomes closer to the standard Bragg angle of sodium chloride 222 (\( 2\theta (\chi) = 56.493^\circ \) at room temp. 12°C). Although the optical center or origin of diffraction geometry should be situated at the standard Bragg angle of quartz, it is not always necessary to give the precise position for the origin if we want only the shape of \( f \) and don't want
Fig. 4. Profile analyses by the least-squares method. (---): true profile $f$; (---): calculated profile $k_{calc}$; (•••): observed profile $k_{obs}$. Intensity is a relative one to the maximum of observed profile.
Table 1. Initial estimates of parameters of the true profile and analytical results. (a): quartz 11·2 - Gaussian; (b): 11·2 - Cauchy; (c): 12·1 - Gaussian; (d): 12·1 - Cauchy. If a initial set is successful in convergence, it is denoted by (a)-(d) at the right side in the row of $f_{\text{initial}}$.

<table>
<thead>
<tr>
<th>Peak or Peak</th>
<th>Integral breadth</th>
<th>Peak or Peak</th>
<th>Integral breadth</th>
<th>Peak or Peak</th>
<th>Integral breadth</th>
</tr>
</thead>
<tbody>
<tr>
<td>Position [degree]</td>
<td>[degree]</td>
<td>[degree]</td>
<td>[degree]</td>
<td>[degree]</td>
<td>[degree]</td>
</tr>
<tr>
<td>(a) (b) 56.47</td>
<td>(G) - 0.104</td>
<td>-</td>
<td>1.468</td>
<td>(a) (b)</td>
<td>2.096</td>
</tr>
<tr>
<td>(c) (d) 56.47</td>
<td>(G) - 0.100</td>
<td>-</td>
<td>1.526</td>
<td>(c) (d)</td>
<td>2.212</td>
</tr>
<tr>
<td>(a) 56.485</td>
<td>0.087</td>
<td>-</td>
<td>1.666</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) 56.485</td>
<td>0.064</td>
<td>-</td>
<td>2.510</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) 56.500</td>
<td>0.080</td>
<td>-</td>
<td>1.803</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) 56.500</td>
<td>0.055</td>
<td>-</td>
<td>2.889</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a) 56.47</td>
<td>0.150</td>
<td>-</td>
<td>0.967</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(b) 56.47</td>
<td>0.156</td>
<td>-</td>
<td>0.992</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c) 56.47</td>
<td>0.148</td>
<td>-</td>
<td>0.975</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(d) 56.47</td>
<td>0.154</td>
<td>-</td>
<td>0.992</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$h_{\text{calc}}$</td>
<td>(56.47)</td>
<td>(0.153)</td>
<td>(1.000)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>(a): 0.08-0.12; (b): 0.07-0.08</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(c): 0.08-0.11; (d): 0.05-0.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

its correct position.

Any integral breadth of $f_{\text{final}}$ is the value of $\rho_i$ itself of the function and is necessarily concerned with its entire curve, while that of $h_{\text{calc}}$ is one derived by dividing the area within $c_{\text{min}} \leq x \leq c_{\text{max}}$ by the height. In either of (a) and (b), the area (= product of integral breadth and peak height) of $f_{\text{final}}$ and that of $h_{\text{calc}}$ are found to be identical. Computer outputs indicated that the truncation of tails in the Gaussian function was made by equation (26), which was the preceding condition to equation (24). It was confirmed that the effect of truncation by (26) on the analytical results was negligible. In contrast with this, in either of (b) and (d), the area of $h_{\text{calc}}$ has a value a little smaller than that of $f_{\text{final}}$. The difference corresponded to the area cut off by (24) and resulted from the persistent tails in the Cauchy function.
Fig. 5. Profile analyses by the Stokes' method. (---): true profile \( f \); (---): calculated profile \( h_{\text{calc}} \); (•••): observed profile \( h_{\text{obs}} \). Intensity is a relative one to the maximum of observed profile.

Regardless of whether quartz 11·2 or 12·1 is used as the instrumental profile, the reliability factors are a little better in cases of the true profile Cauchy than in cases of the true profile Gaussian. We can not conclude, however, that there is essentially a difference of the curve fitting between both functions, if it is taken into account that in any case, the fitting near the peak is good; the tail portions away from the peak to the higher angle in which the deviation is large in Gaussian cases have an error introduced during the doubling elimination; the background corrections have an uncertainty to some extent.

For the milled powders of sodium chloride, it can be expected that crystallite-size and strain broadening coexist. In the investigations on cold-worked metals and alloys, it has been indicated that these broadenings are very close to a Cauchy and a Gaussian distribution,
respectively [7]. It may be worthwhile to attempt the analysis by using a Voigtian which is a convolution of one or more Cauchy and Gaussian functions [8]. It should be noted here that in general the effect of TDS (thermal diffuse scattering) in the analysis of the X-ray powder diffraction profiles can be neglected [9].

5. Analysis by the Stokes' method

In Fig. 5 are shown the analytical results obtained by the Stokes' method [1]. It is a logical conclusion that the calculated profiles, which are the convolution of \( f \) and \( g \), show an almost perfect curve fitting to the observed one, as both the deconvoluted \( f \) are solutions which satisfy the convolution relation \( h_{\text{obs}} = f \ast g \). The erroneous sinusoidal components in tails of \( f \) are originated from the experimental errors in \( g \) or in \( F \) which is a Fourier transform of \( f \). In this experiment, the broadening of observed 222 reflexion of sodium chloride is not so much large compared with that of quartz, i.e., the integral breadth of the former is 0.15° and those of 11.2 and 12.1 in the latter are nearly the same 0.11°. Such a condition makes the final shape of \( f \) uncertain because errors in \( F(t)(t: \text{a variable in the Fourier space}) \) for large values of \( t \) increase anomalously with increasing \( t \) and consequently \( f \) is ill synthesized. It so happens that the large scale picture of \( f \) in (a) resembles the Gaussian curves in Fig. 4, while that of \( f \) in (b) resembles a curve with an intermediate shape of the Gaussian and the Cauchy ones, sensitively affected by the small difference of shape between the two instrumental profiles. In order to confirm the reliability of the results, it will be preferable that \( f \) is estimated so as to lie within the band with a probability [10].

6. Consideration of the Instrumental Profile

In this analysis, we were forced to adopt the two instrumental profiles which are considered to be non-ideal, because the 2\( \theta \) dependence of the instrumental line shape was apparent (Fig. 2 and 3). It is assumed in such a treatment that the convolution relation stands approximately in each combination shown above; an expected result is regarded to be intermediate of two corresponding results.
deconvoluted by these two instrumental profiles.

The determination of the instrumental profile becomes the more important problem particularly when the X-ray absorption coefficient of a specimen is low and largely different from that of a standard material, because the X-ray transparency with a large sample thickness is one of the effective broadening factors and causes a distortion of diffraction lines [5]. In the preliminary experiment of this study, even in the present combination of the relatively high absorption coefficient materials compared with organic materials, a line shape of quartz from quartz only and that from their mixture exhibited different broadenings. From this result, separate measurements of these two materials reasonably expected to bring different instrumental broadenings on their profiles. Then, an application of the convolution relation to them can not be justified. On the other hand, the internal-standard method allows the application, as that guarantees that both materials are subjected to the same instrumental broadening effects including the X-ray transparency. The methods of the absorption corrections [11] [12] on the distorted line shapes have a room for improvement. These are the reasons why the internal-standard method was adoptd.

Now, for example, in analysis of the X-ray line broadenings in an organic polymer substance, isotactic polystyrene, the use of an organic instrumental standard, a well-crystallized hexamethylene-tetramine, whose X-ray transparency is similar to that of the former, was attempted [13]. This may become a solution to remove the above kind of problem to a degree.

Appendix A

Setting $\delta^{\theta} = 0$, we have finally the normal equations

$$M \delta = b$$  \hspace{1cm} (A1)

where $M$, $\delta$ and $b$ are expressed as follows.

is a square matrix of order $n$ ($n$: the number of parameters), which can be written as

$$M = \tilde{A} WA$$  \hspace{1cm} (A2)

where $A$ is an asymmetric matrix of order $m \times n$ ($m$: the number of data), whose $ij$ th element is given by (11); $\tilde{A}$ is its transposed matrix;
$W$ is a diagonal matrix of order $m \times m$ with respect to weights of the observed data. $b$ is a column matrix of order $n$, represented by

$$b = \tilde{A}W\Delta$$

where $\Delta$ is a column matrix of order $m$, whose $i$th element is given by

$$(\Delta)_i = h(x_i)_{\text{obs}} - h(x_i)_{\text{calc}}.$$  

From (A1), we have a set of corrections in the form of a column matrix of order $n$

$$\delta = M^T b$$

which are added to the initial estimates $p_0$, and a refined set of estimates for the next iteration is provided. This procedure is repeated until the corrections $\delta$ become negligibly small.

Appendix B

The areas under the integrable curves $f(x)$ and $g(x)$ are expressed, respectively in the forms

$$S_f = \int f(x) \, dx$$

and

$$S_g = \int g(x) \, dx$$

Let $h(x)$ be a convolution of $f(x)$ and $g(x)$, the area of $h(x)$ is given by

$$S_h = \int \left[ \int f(y)g(x-y) \, dy \right] \, dx$$

$$= \int \left[ \int f(y)g(x-y) \, dx \right] \, dy$$

$$= \int f(y)S_g \, dy = S_fS_g$$

References