Multiple scattering in the Compton effect. IV. Operator formalism for nonstationary electrons

Elaine Braun-Keller and Irving R. Epstein*
Department of Chemistry, Brandeis University, Waltham, Massachusetts 02154
(Received 9 July 1976; revised manuscript received 22 November 1976)

The problem of correcting experimental Compton profiles for the effects of multiple scattering is shown to be equivalent to obtaining the solution to a nonlinear operator equation. Two methods of solving the equation are presented. One, analogous to the Newton-Raphson technique for algebraic equations, leads to a new solution to the problem. Another more efficient approach, based upon the idea of successive approximation, is seen to be equivalent to the method of Felsteiner, Pattison *et al.*, though it fails to converge if the fraction of multiple scattering is too great. The double-scattering operator is derived and examined in some detail. Extensions to higher-order scattering are indicated.

I. INTRODUCTION

Considerable progress has been made in the past several years toward the goal of understanding and being able to correct for the effects of multiple scattering in the measurement of Compton profiles.[1] Insights have been gained into the factors which determine the intensity,[2-5] angular distribution,[2,4-6] spectrum,[3,5-10] and relative importance of elastic and inelastic scattering[9-11] in multiple scattering, though these studies have generally been based upon scattering from stationary electrons.

In addition, two methods have been proposed, and applied with considerable success, to derive the pure single scattered profile from experimental measurements which contain contributions from multiply scattered photons. In the approach of Halonen, Williams, and Paakari (HWP)[10] a theoretical single scattered profile is used to generate the multiple scattering contribution, the intensity of which is obtained from a Monte Carlo calculation. This method has the obvious drawbacks that a reliable calculated profile is required as a starting point and that it may prove exceedingly difficult to distinguish between two or more reasonably accurate calculations.

The technique proposed by Felsteiner, Pattison, and Cooper (FPC)[6] uses the experimental profile as a first approximation to the single scattered profile and obtains the latter by an iterative procedure in which the spectrum and intensity of multiple scattering are calculated by a Monte Carlo method. Particularly with its recent modification to allow for the effects of electron binding energies,[12] the FPC technique appears to offer great promise as a practical procedure for correcting experimental profiles. However, no rigorous demonstration has yet been given either that the iteration should converge or that, if it does converge, the result obtained is the true single scattered profile.

In this paper, we show that the problem of correcting experimental profiles for the effects of multiple scattering is equivalent to solving a nonlinear operator equation. We discuss several approaches to obtaining this solution, and find that one of them corresponds to the FPC iteration scheme. Analysis shows that the scheme should converge to the desired profile if the fraction of multiple scattering is not too large. A new, more generally convergent, but less computationally attractive method is also derived and applied to a model experiment. In addition to our detailed treatment of double scattering, we outline the extension of the operator formalism to higher-order scattering. In paper V[13] of this series, we show how the formalism developed here may be employed to obtain bounds on the errors associated with single scattered profiles which have been obtained by correcting experimental profiles using the techniques described here or any alternative methods.

II. OPERATOR FORMALISM

Consider a Compton scattering experiment in which incident photons of wavelength λ_0 are scattered through an angle θ_p. Let $F^{\theta_p}_{\lambda_0}(\lambda)$ be the normalized single scattered Compton profile, i.e., the intensity of scattered radiation as a function of wavelength after the usual corrections for background, resolution, absorption, and detector efficiency,[14] if there were no multiple scattering. Determination of $F^{\theta_p}_{\lambda_0}(\lambda)$ is the object of conventional Compton scattering experiments, since it is this quantity which is directly related to the electronic momentum distribution of the scatterer.[15] We define an operator $\hat{O}$ which when applied to $F^{\theta_p}_{\lambda_0}(\lambda)$

generates the normalized multiple scattered profile $G^{\theta_p}_{\lambda_0}(\lambda)$, i.e.,

$$\hat{O}F^{\theta_p}_{\lambda_0}(\lambda) = G^{\theta_p}_{\lambda_0}(\lambda)\,, \tag{1a}$$

$$\int F^{\theta_p}_{\lambda_0}(\lambda)d\lambda = \int G^{\theta_p}_{\lambda_0}(\lambda)\,d\lambda = 1\,. \tag{1b}$$

For the moment we shall limit ourselves to the case in which triple and higher-order scatterings make negligible contributions, so that the multiple scattered profile $G^{\theta_p}_{\lambda_0}$ is simply the (normalized) double scattered profile $D^{\theta_p}_{\lambda_0}$.[16] In Sec. IV we show how higher-order scattering may be included within the present framework.

Let x be the fraction of double scattering in the experimental profile J_E. Thus

$$\begin{aligned} J_E(\lambda) &= (1-x)\hat{I}F^{\theta_p}_{\lambda_0}(\lambda) + xD^{\theta_p}_{\lambda_0}(\lambda) \\ &= (1-x)\hat{I}F^{\theta_p}_{\lambda_0}(\lambda) + x\hat{O}F^{\theta_p}_{\lambda_0}(\lambda)\,, \end{aligned} \tag{2}$$

where $\hat{I}$ is the identity operator. In general, x will depend upon the scattering angle, the sample dimensions and orientation, the beam size,[5] and the cross sections for the various photon-electron interaction channels. The value of x is generally obtained from a Monte Carlo calculation, though it may also be estimated from the observed relative intensities of the elastic and Compton lines.[11] Equation (3) may be formally inverted to yield the desired profile $F^{\theta_p}_{\lambda_0}(\lambda)$:

$$F^{\theta_p}_{\lambda_0}(\lambda) = [\hat{I} - x(\hat{I} - \hat{O})]^{-1}J_E(\lambda)\,. \tag{3}$$

Obtaining $F^{\theta_p}_{\lambda_0}(\lambda)$ has now been reduced to solving Eq. (2) for $F^{\theta_p}_{\lambda_0}$ given J_E or, equivalently, carrying out the inversion in Eq. (3).

We now seek to construct the operator $\hat{O}$. Photons which scatter twice through solid angles Ω_1 $[=(\theta_1, \varphi_1)]$ and Ω_2 have a spectral distribution given by

$$D^{(\Omega_1,\Omega_2)}_{\lambda_0}(\lambda) = \int F^{\theta_1}_{\lambda_0}(\lambda')F^{\theta_2}_{\lambda'}(\lambda)\,d\lambda'\,, \tag{4}$$

where the notation $F^{\theta}_{\lambda_1}(\lambda_2)$ signifies the normalized Compton profile which would be observed at a scattering angle θ and a wavelength λ_2 as a result of single scattering photons of incident wavelength λ_1. The integration in Eq. (4) is over all possible intermediate wavelengths λ'. We now wish to relate $F^{\theta}_{\lambda_1}(\lambda_2)$ to $F^{\theta_p}_{\lambda_0}(\lambda)$ for any λ_1, λ_2, and θ.

A photon of wavelength λ_1 which scatters through an angle θ from an electron with momentum component p_z along the scattering vector undergoes a wavelength shift given by

$$\Delta\lambda = \lambda_2 - \lambda_1 = \frac{2h}{mc}\sin^2\tfrac{1}{2}\theta + 2\frac{p_z}{mc}\lambda_1\sin\tfrac{1}{2}\theta\,. \tag{5}$$

We define

$$a(\theta) = \lambda_0 + (2h/mc)\sin^2\tfrac{1}{2}\theta\,, \tag{6a}$$

$$b(\theta) = (2\lambda_0/mc)\sin\tfrac{1}{2}\theta\,. \tag{6b}$$

Then, using Eq. (6), Eq. (5) may be written as

$$\lambda_2 = \lambda_1 - \lambda_0 + a(\theta) + (\lambda_1/\lambda_0)b(\theta)\,p_z \tag{7a}$$

or

$$p_z = \lambda_0[\lambda_2 - \lambda_1 + \lambda_0 - a(\theta)]/\lambda_1 b(\theta)\,. \tag{7b}$$

When the two profiles are both normalized, we must have

$$F^{\theta}_{\lambda_1}(\lambda_2) = F^{\theta_p}_{\lambda_0}(\lambda)$$

for wavelengths λ and λ_2 which correspond to the same electron momentum p_z. Thus, by setting $\lambda_2 = \lambda$, $\lambda_1 = \lambda_0$, and $\theta = \theta_p$ in Eq. (7a) and using Eq. (7b) for p_z, we obtain the general result that

$$\begin{aligned} F^{\theta}_{\lambda_1}(\lambda_2) = &\left[\frac{\lambda_0 b(\theta_p)}{\lambda_1 b(\theta)}\right] \\ &\times F^{\theta_p}_{\lambda_0}\left(a(\theta_p) + \frac{\lambda_0 b(\theta_p)}{\lambda_1 b(\theta)}\{\lambda_2 - \lambda_1 + \lambda_0 - a(\theta)\}\right), \end{aligned} \tag{8}$$

where the factor in square brackets is required to preserve the normalization of the profiles.

Equation (8) now allows us to write the double scattered profile D of Eq. (4) completely in terms of single scattered profiles at the experimental conditions λ_0, θ_p. We obtain

$$\begin{aligned} D^{(\Omega_1,\Omega_2)}_{\lambda_0}(\lambda) = \int &\frac{b(\theta_p)^2}{b(\theta_1)b(\theta_2)}\,\frac{\lambda_0}{\lambda'} \\ &\times F^{\theta_p}_{\lambda_0}\left(a(\theta_p) + \frac{b(\theta_p)}{b(\theta_1)}(\lambda' - a(\theta_1))\right) \\ &\times F^{\theta_p}_{\lambda_0}\left(a(\theta_p) + \frac{\lambda_0}{\lambda'}\frac{b(\theta_p)}{b(\theta_2)}\right. \\ &\left.\times[\lambda - \lambda' + \lambda_0 - a(\theta_2)]\right)d\lambda'\,. \end{aligned} \tag{9}$$

The spectral distribution given by Eq. (9) is *not*, as has sometimes been held,[6] a convolution of two single scattered profiles, since the shape of the second profile depends upon the intermediate wavelength λ'. In fact $F^{\theta_2}_{\lambda'}(\lambda)$ is not even integrable as a function of λ'.

Equation (9) gives the double scattered profile that would be observed if there were only a single pair of angles (Ω_1, Ω_2) which allowed double scattered photons to emerge at θ_p. Such an example will be considered later. However, in any real experiment, there are many ways in which two scatterings can combine to yield the observed scattering angle, i.e., there is a distribution of (Ω_1, Ω_2). We denote the probability that a photon

will scatter successively into solid angles between Ω_1 and $\Omega_1 + d\Omega_1$, then Ω_2 and $\Omega_2 + d\Omega_2$ and finally emerge with total scattering angle θ_p as $P^{\theta_p}(\Omega_1, \Omega_2)\, d\Omega_1\, d\Omega_2$. This probability distribution will be determined by the differential and total scattering cross sections and by the sample geometry[4] as well as by the binding energies of the electrons.[10,12,17] It is perhaps most conveniently obtained by a Monte Carlo procedure,[10] although under certain circumstances simplifications can be made. For instance, in the infinite radius limit,[4] $P^{\theta_p}(\Omega_1, \Omega_2)$ reduces to a simple analytic expression in closed form. Also, if we assume perfect collimation of the incident beam to zero diameter, the problem of calculating the probability distribution may be reduced to a three-dimensional integration, thereby significantly increasing the efficiency of the Monte Carlo computation relative to existing procedures.[6]

Given the form of P^{θ_p}, we may write the actual double scattered profile as

$$D^{\theta_p}_{\lambda_0}(\lambda) = \int D^{(\Omega_1,\Omega_2)}_{\lambda_0}(\lambda) P^{\theta_p}(\Omega_1, \Omega_2)\, d\Omega_1\, d\Omega_2 . \tag{10}$$

Combining Eqs. (1), (9), and (10), we are now able to identify the operator $\hat{O}$. We may view $\hat{O}$ as a bilinear operator[18] which maps ordered pairs of functions (single scattered profiles) into a single function (a double scattered profile). We write $\hat{O}$ as the composition of three operators:

$$\hat{O} = \hat{A}\hat{R}\hat{C} . \tag{11}$$

$\hat{C}$ is the Compton scaling and translation operator of Eq. (8). If f and g are any two functions, then

$$\hat{C}(f(\lambda), g(\lambda)) = \left(\frac{b(\theta_p)}{b(\theta_1)} f\left[a(\theta_p) + \frac{b(\theta_p)}{b(\theta_1)} [\lambda' - a(\theta_1)] \right] , \frac{b(\theta_p)}{b(\theta_2)} \frac{\lambda_0}{\lambda'} g\left[a(\theta_p) + \frac{\lambda_0}{\lambda'} \frac{b(\theta_p)}{b(\theta_2)} [\lambda - \lambda' + \lambda_0 + a(\theta_2)] \right] \right) . \tag{12}$$

$\hat{C}$ is a linear operator on ordered pairs of functions. The operator $\hat{R}$ corresponds to the double scattering integration of Eq. (9):

$$\hat{R}(h_1(\lambda'), h_2(\lambda')) = \int h_1(\lambda') h_2(\lambda')\, d\lambda' . \tag{13}$$

Note that $\hat{R}$ is bilinear. Finally, $\hat{A}$ is the angular integration, Eq. (10), which is a linear map:

$$\hat{A} f(\lambda) = \int f^{(\Omega_1,\Omega_2)}(\lambda) P^{\theta_p}(\Omega_1, \Omega_2)\, d\Omega_1\, d\Omega_2 . \tag{14}$$

Comparing Eqs. (8)–(10) with Eqs. (11)–(14), we see that

$$D^{\theta_p}_{\lambda_0} = \hat{O} F^{\theta_p}_{\lambda_0} = \hat{A}\hat{R}\hat{C}(F^{\theta_p}_{\lambda_0}, F^{\theta_p}_{\lambda_0}) \equiv \hat{B}(F^{\theta_p}_{\lambda_0}, F^{\theta_p}_{\lambda_0}) , \tag{15}$$

and thus $\hat{B}$ is a bilinear operator. Having defined $\hat{O}$, we now return to the problem of solving Eq. (2).

III. SOLUTION OF THE EQUATION

Since the operator $\hat{O}$ is nonlinear, the simple expansion of Eq. (4) as either

$$[\hat{I} - x(\hat{I} - \hat{O})]^{-1} = \hat{I} + x(\hat{I} - \hat{O}) + x^2(\hat{I} - \hat{O})^2 + \cdots \tag{16a}$$

or as

$$[\hat{I} - x(\hat{I} - \hat{O})]^{-1} = \frac{1}{1-x}\hat{I} - \frac{x}{(1-x)^2}\hat{O} + \frac{x^2}{(1-x)^3}\hat{O}^2 + \cdots \tag{16b}$$

may not be valid. Nonlinear operator equations like Eq. (2) are difficult to solve, since direct expansion of the inverse operator as in Eqs. (16) may not be possible, even if the inverse of $\hat{O}$ were known explicitly. Several iterative techniques are available for dealing with such equations, however, and we present two such approaches here.

A. Newton-Raphson method

One of the most powerful ways of solving nonlinear algebraic equations of the form $f(x) = 0$ is the Newton-Raphson iterative procedure. Provided that the derivative $f'(x)$ exists and is nonzero for all x, a root of the equation, if one exists, is obtained as the limit of the sequence $\{x_n\}$, where x_{n+1} is given by

$$x_{n+1} = x_n - f(x_n)/f'(x_n) = x_n - [f'(x_n)]^{-1} f(x_n) \tag{17}$$

and the initial approximation x_0 may be chosen arbitrarily.

The above method may be generalized to nonlinear operator equations.[19] Given an operator $\hat{O}$, if the solution to the equation $\hat{O}f = 0$ exists, it is the limit of the sequence of functions $\{f_n\}$, where

$$f_{n+1} = \hat{I} f_n - (\hat{O}' f_n)^{-1} \hat{O} f_n \tag{18}$$

provided that $\hat{O}'$, the Frechet derivative of the operator $\hat{O}$ (see Appendix A) exists and its inverse $\hat{O}'^{-1}$ exists and is bounded.

We wish to solve by Newton's method the equation

$$\hat{T} J_S = (1-x)\hat{I} J_S + x\hat{B}(J_S, J_S) - J_E = 0 , \tag{19}$$

where the bilinear operator $\hat{B}$ is defined in Eqs. (12)–(15) and J_S is the single scattered profile

$F^{\theta_p}_{\lambda_0}(\lambda)$. We have by Eq. (18)

$$J_S^{(n+1)} = [\hat{I} - \hat{T}'^{-1}(J_S^{(n)}, \cdot)]\hat{T}J_S^{(n)} . \tag{20}$$

The operator $\hat{T}$ at J_S has the Frechet derivative

$$\hat{T}'(J_S, f) = (1-x)f + x\hat{B}'(J_S, f) , \tag{21}$$

where $\hat{B}'(J_S, f) = \hat{B}(J_S, f) + \hat{B}(f, J_S)$ or, in the present case,

$$\hat{T}'(J_S^{(n)}, \hat{T}J_S^{(n)}) = (1-x)\hat{T}J_S^{(n)} + x\hat{B}'(J_S^{(n)}, \hat{T}J_S^{(n)}) . \tag{22}$$

Since $\hat{T}'$ is linear in its second argument, its inverse has the series expansion [starting from Eq. (21)]

$$\hat{T}'^{-1} = [(1-x)\hat{I} + x\hat{B}'(J_S, \cdot)]^{-1} f = \frac{1}{1-x}\left[\hat{I} - \frac{x}{1-x}\hat{B}'(J_S, \cdot) + \frac{x^2}{(1-x)^2}\hat{B}'^2(J_S, \cdot) - \cdots\right] f . \tag{23}$$

Inserting this expansion in Eq. (20), we obtain

$$J_S^{(n+1)} = J_S^{(n)} - \left\{\hat{I} - \frac{1}{1-x}\left[\hat{I} - \frac{x}{1-x}\hat{B}'(J_S^{(n)}, \cdot) + \frac{x^2}{(1-x)^2}\hat{B}'^2(J_S^{(n)}, \cdot) - \cdots\right]\right\}[(1-x)\hat{I} + x\hat{B}(\cdot, \cdot) - J_E]J_S^{(n)} . \tag{24}$$

Using the linearity of $B'(J_S^{(n)}, \cdot)$, Eq. (24) affords the following iteration scheme:

$$J_S^{(n+1)} = J_S^{(n)} - \left\{\frac{1}{1-x}(1-x)J_S^{(n)} + x\hat{B}(J_S^{(n)}, J_S^{(n)}) - J_E + \sum_{i=1}^{\infty}(-)^i\left[x^i\hat{B}'^{(i)}(J_S^{(n)}, J_S^{(n)}) + \frac{x^{2i}}{(1-x)^i}\hat{B}'^{(i)}(J_S^{(n)}, \hat{B}(J_S^{(n)}, J_S^{(n)})) + (-)^{i+1}\left(\frac{x}{1-x}\right)^i\hat{B}'^{(i)}(J_S^{(n)}, J_E)\right]\right\} . \tag{25}$$

The sequence $J_S^{(n)}$ will converge to the correct single scattering profile if $x < 0.5$. The series of Eq. (25) diverges if $x \geq 0.5$. However if we replace the series expansion [Eq. (23)] of $\hat{T}'^{-1}$ by the series

$$\begin{aligned}\hat{T}'^{-1} &= [(1-x)\hat{I} + x\hat{B}'(J_S, \cdot)]^{-1} \\ &= [\hat{I} - x(\hat{I} - \hat{B}'(J_S, \cdot))]^{-1} \\ &= \hat{I} + x(\hat{I} - \hat{B}'(J_S, \cdot)) + x^2(\hat{I} - \hat{B}'(J_S, \cdot))^2 + \cdots ,\end{aligned} \tag{26}$$

then the iteration will converge to J_S for all $x < 1$, though convergence will be somewhat slower for $x < 0.5$ than if Eq. (23) had been employed.

While the calculations in Eq. (25) appear formidable, the series in i converges after 3 or 4 terms, and for small x, the method should be practicable. As a test, we have applied Eq. (25) to a very simple model of a Compton scattering experiment. Figure 1 shows a single electron surrounded by a planar ring of electrons. Photons are incident along the radius of the ring, and a detector with perfect collimation is located at a scattering angle θ_p. Clearly, all single scattered photons detected must have $\theta_1 = \theta_p$, while double scattering requires $\theta_1 = 180° - \theta_p$, $\theta_2 = 180°$ for detection. The distribution function $P^{\theta_p}(\Omega_1, \Omega_2)$ has been reduced to a δ function, removing the angular integrations.

With this simplification, we apply Eq. (25) to the problem of extracting the single scattered profile from an "experimental" profile. J_E is constructed by choosing a form for $F^{\theta_p}_{\lambda_0}$, computing $D^{\theta_p}_{\lambda_0}$ by the prescription of Eq. (9), and substituting in Eq. (2) using an assumed value of x. Table I gives the results obtained when the single scattered profile is taken as a Gaussian of half-width 0.833. All of the integrations required to compute the $\hat{B}$ operators were carried out numerically using a 20-point Gaussian quadrature.[20] Tails of all functions were set to zero for $|q| > 5$ a.u. The experimental profile J_E was used as the starting approximation $J_S^{(0)}$.

We see that even after a single iteration the ani-

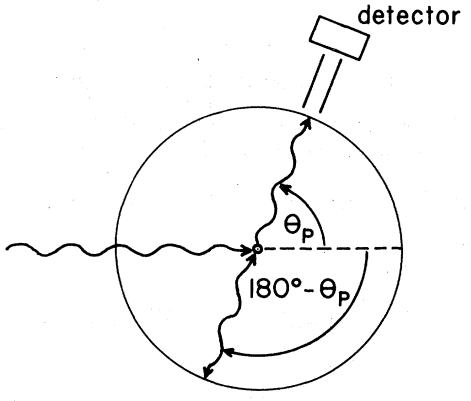

FIG. 1. A model Compton scattering experiment. Note that in order to be detected, single or double scattered photons must scatter first from the central electron.

TABLE I. Correction of a model experimental profile for double scattering using Eq. (25).[a]

λ (Å)	Exptl.[b] profile	1st iteration	2nd iteration	Exact single scat.
0.715	0.5009	0.5565	0.5565	0.5565
0.720	1.261	1.402	1.402	1.402
0.725	2.846	3.137	3.121	3.119
0.730	5.614	6.157	6.136	6.131
0.735	9.793	10.70	10.66	10.65
0.740	15.12	16.45	16.36	16.34
0.742	17.38	18.87	18.76	18.73
0.744	19.59	21.22	21.09	21.08
0.746	21.67	23.40	23.24	23.19
0.748	23.50	25.29	25.10	25.05
0.750	25.00	26.81	26.59	26.53
0.751	25.60	27.40	27.17	27.10
0.752	26.09	27.87	27.61	27.55
0.753	26.46	28.20	27.93	27.86
0.754	26.71	28.40	28.11	28.04
0.755	26.84	28.46	28.16	28.08
0.756	26.83	28.38	28.06	27.98
0.757	26.70	28.16	27.83	27.74
0.758	26.45	27.80	27.46	27.37
0.760	25.58	26.71	26.35	26.25
0.762	24.28	25.15	24.79	24.68
0.764	22.63	23.23	22.86	22.75
0.766	20.70	21.04	20.67	20.56
0.768	18.60	18.68	18.32	18.22
0.770	16.41	16.26	15.92	15.82
0.775	11.11	10.54	10.29	10.20
0.780	6.767	6.03	5.87	5.81
0.790	1.864	1.32	1.32	1.30
0.795	0.8534	0.485	0.514	0.511

[a] All profiles normalized to unity.
[b] Calculated from Eqs. (2) and (10) with $\theta_p = 150°$, $x = 0.1$.

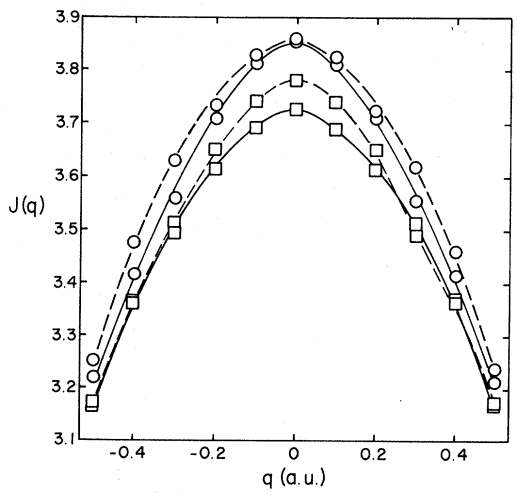

FIG. 2. Experimental (□) and corrected (○) Compton profiles for 2.0 cm (--) and 0.3 cm (——) samples of water.

sotropy has been greatly reduced, and accuracy has been improved by nearly an order of magnitude. The major drawback of this technique is that it appears to be computationally inefficient. Even in this simplified case where only one pair of scattering angles is allowed, several minutes of computer time are required to correct the model J_E for double scattering.

The correction procedure was also applied to two x-ray Compton profiles of water, with sample thicknesses 0.3 and 2.0 cm. The experiments were done by Halonen as part of the Compton profile standardization project.[21] The Monte Carlo routine gave the double scattered contributions as 10% in the thick sample, 4.5% in the thin. The results of the corrections are shown in Fig. 2. The peak heights are brought into agreement, although this may be fortuitous, since higher-order (more than double) scattering was neglected. The corrected profiles differ somewhat at low momenta away from $p = 0$. We believe that this is due to the symmetrizing of the experimental profiles. As we see from the model calculation, there is a discernible anisotropy in a Compton profile with 10% double scattering. If this is averaged out by symmetrizing the profile, the corrected profile will be asymmetric. Although this effect is somewhat more pronounced in the model, it may be a source of error in real systems as well.

B. Successive approximation

An alternative, though less generally applicable technique for solving nonlinear equations is successive approximation. We rewrite Eq. (2) as a mapping:

$$\hat{F}\colon f \to \frac{1}{1-x} J_E - \frac{x}{1-x}\hat{O}f \,. \tag{27}$$

A solution to Eq. (3) is thus a *fixed point* of this mapping, that is, a function f_0 such that

$$\hat{F}\colon f_0 \to f_0 \,.$$

The f_0 which satisfies this condition is the exact single scattered profile J_S, since by assumption

$$J_S = \frac{1}{1-x} J_E - \frac{x}{1-x}\hat{O}J_S \,. \tag{28}$$

To generate a successive approximation scheme, we simply use the nth approximation to J_S on the right-hand side of Eq. (28) to yield the $(n+1)$-th approximation on the left-hand side:

$$J_S^{(n+1)} = \frac{1}{1-x} J_E - \frac{x}{1-x}\hat{O}J_S^{(n)} \,. \tag{29}$$

With the choice $J_S^{(0)} = J_E$, Eq. (29) is identical to the FPC iteration scheme.[6] *If* the operator $\hat{O}$

were linear, repeated application of Eq. (29) would yield a series solution of the form obtained in Eq. (16b) by expanding the inverse operator. Since $\hat{O}$ is not linear, one must take care not to attempt such a series solution but must follow the iterative prescription of Eq. (29). The work of Felsteiner *et al.*[6,11,12] suggests that their method converges rapidly in most cases. It should be more computationally efficient than the Newton-Raphson scheme.

From a mathematical point of view, Eq. (29) will converge to the correct J_S if $\hat{F}$ in eq. (27) is a *contraction mapping*.[22] While we have been unable to determine that this is the case in general, we do show in Appendix B that $\hat{F}$ *is not* a contraction mapping for $x \geq \frac{1}{2}$. Further, if we assume that the major contribution to the integral for the double scattered profile [Eq. (9)] is from $\lambda_0/\lambda_1 \sim 1$ (this turns out to be an excellent approximation), the operator becomes a convolution. Then it can be shown (see Appendix B) that for $\theta_1 = \theta_2$ the convergence is further limited to $x < \frac{1}{3}$. This condition on θ_1 and θ_2 would hold, for instance, in the experiment proposed by Felsteiner and Opher,[23] which would measure the spin dependence of the Compton profile. Thus the FPC self-consistent iteration procedure [Eq. (29)], if the correct operator is employed, should be the most effective way of correcting for multiple scattering so long as the fraction of multiple scattering is not too large.

IV. HIGHER-ORDER SCATTERING

The techniques developed in the previous sections have assumed that only single and double scattering contribute significantly to the observed profile. While this is probably a good approximation in most cases, higher-order scattering can be important, particularly if thick samples are used or if very high accuracy is required. The operator formalism is easily extended to deal with higher-order scattering.

We have thus far approximated the operator $\hat{O}$ of Eq. (1), which generates $G^{\theta_p}_{\lambda_0}$ from $F^{\theta_p}_{\lambda_0}$ by the double scattering operator $\hat{O}_2$, which generates $D^{\theta_p}_{\lambda_0}$ from $F^{\theta_p}_{\lambda_0}$. The true $\hat{O}$ is a sum over all orders of scattering

$$\hat{O} = \sum x_i \hat{O}_i \tag{30}$$

where the ith-order scattering operator $\hat{O}_i$ is defined by

$$\hat{O}_i F^{\theta_p}_{\lambda_0}(\lambda) = \int P^{\theta_p}(\Omega_1, \Omega_2, \ldots, \Omega_1) \times F^{\theta_1}_{\lambda_0}(\lambda_1) F^{\theta_2}_{\lambda_1}(\lambda_2) \ldots F^{\theta_i}_{\lambda_{i-1}}(\lambda_i) \times d\Omega_1\, d\Omega_2 \ldots d\Omega_i\, d\lambda_1\, d\lambda_2 \ldots d\lambda_{i-1}. \tag{31}$$

The distribution function P^{θ_p} is again best found by a Monte Carlo procedure. Defining

$$x_T = \sum_i x_i ,$$

Eq. (3) generalizes to

$$J_E = (1 - x_T)\hat{I} J_S + \sum_i x_i \hat{O}_i J_S . \tag{32}$$

Equation (32) may be solved by the same procedures as Eq. (2). The condition for the convergence of Eq. (25) becomes $x_T < \frac{1}{2}$, i.e., the amount of single scattering must exceed the sum of all the higher-order scatterings.

Since the operators $\hat{O}_i$ are i-linear, the Frechet derivatives $\hat{O}'_i$ are easy to compute. The Frechet derivative of an i-linear operator is a linear operator, and the required derivatives for our Newton-Raphson procedure are of the form

$$\hat{O}'_i(a, \ldots, a)(x, \ldots, x) = \sum_{j=1}^{i} P_j[\hat{O}(x, a, \ldots, a)] , \tag{33}$$

where P_j permutes the n-tuple $(x, a, \ldots, a)$ so as to place x in the jth position. Clearly, use of Eqs. (30) and (33) in Eq. (25) will render the Newton-Raphson procedure even more unwieldy. The complications introduced in the successive approximation approach should be less serious, though if higher-order scattering is non-negligible, x_T may be large enough to make convergence a problem.

V. CONCLUSIONS

The operator formalism described here provides a useful and flexible approach to the problem of multiple Compton scattering. It has enabled us to derive a new technique for correcting profiles and to understand better the advantages and limitations of an existing method. The operator approach can also be adapted to a number of related problems. For example, by employing a translation operator, which is linear, in place of the nonlinear multiple scattering operator, one may apply an expansion similar to that of Eq. (16) to derive a method for carrying out the $\alpha_1 - \alpha_2$ separation in x-ray Compton experiments. The method turns out to be equivalent to the commonly applied Rachinger technique.[24]

Perhaps the most important consequence of the operator view point is that it enables us to derive upper and lower bounds to the error in profiles which have been corrected for multiple scattering effects. These results will be presented in paper V of this series.

ACKNOWLEDGMENTS

We thank Clayton Keller for illuminating mathematical discussions. We are grateful to Dr. Anthony Tanner for a number of important insights into this problem. Acknowledgment is made to the donor of the Petroleum Research Fund, administered by the American Chemical Society, for support of this research.

APPENDIX A: FRECHET DERIVATIVES

We summarize some results from the theory of nonlinear operators.[18]

(1) Given spaces X and Y and an operator $\hat{O}$: $X \to Y$, the derivative of $\hat{O}$, called its *Frechet derivative*, is defined as follows:

If, at a point $x \in X$,

$$\hat{O}(x+h) - \hat{O}x = \hat{U}(x,h) + \hat{W}(x,h)\,, \tag{A1}$$

where $\hat{U}(x,h)$ is a bounded linear operator in $h \in X$ at x and

$$\lim_{\|h\| \to 0} \frac{\|\hat{W}(x,h)\|}{\|h\|} = 0\,, \tag{A2}$$

then $\hat{U}(x,h)$ is the Frechet differential of $\hat{O}$ with remainder $\hat{W}(x,h)$. The operator itself, $\hat{U}(x,\cdot)$ [also denoted $\hat{O}'(x,\cdot)$ or $d\hat{O}(x,\cdot)$] is called the *Frechet derivative of* $\hat{O}$ *at* x. Equivalently $\hat{U}$ is a linear transformation in h satisfying

$$\lim_{\|h\| \to 0} \frac{\|\hat{O}(x+h) - \hat{O}x - \hat{U}(x,h)\|}{\|h\|} = 0\,. \tag{A3}$$

(2) The Frechet derivative of a linear operator is the operator itself; i.e., it is clear that if $\hat{O}$ is linear, $\hat{U}(x,h) = \hat{O}h$ satisfies Eq. (A3).

(3) The Frechet derivative of a bilinear operator $\hat{B}$ at (x_1, x_2) in the direction (y_1, y_2) is the linear operator

$$\begin{aligned} d\hat{B}(x_1,x_2)(y_1,y_2) &= \lim_{t\to 0} \frac{\hat{B}(x_1+ty_1, x_2+ty_2) - \hat{B}(x_1,x_2)}{t} \\ &= \lim_{t\to 0} \frac{\hat{B}(x_1,x_2) + \hat{B}(x_1,ty_2) + \hat{B}(ty_1,x_2) + \hat{B}(ty_1,ty_2) - \hat{B}(x_1,x_2)}{t} \\ &= \lim_{t\to 0} \frac{t\hat{B}(x_1,y_2) + t\hat{B}(y_1,x_2) + t^2\hat{B}(y_1,y_2)}{t} \\ &= \hat{B}(x_1,y_2) + \hat{B}(y_1,x_2)\,. \end{aligned} \tag{A4}$$

In our application, we shall need only the derivative at (x,x) in the direction (y,y), which from Eq. (22) becomes

$$\hat{B}'(x,x)(y,y) = \hat{B}(x,y) + \hat{B}(y,x) \equiv \hat{B}(x,y) + \hat{B}^*(x,y) \tag{A5}$$

where $\hat{B}^*(x,y) = \hat{B}(y,x)$ is called the permutation of $\hat{B}$. To simplify the notation a bit, we shall write derivatives of the form (A5) as

$$B'(x,y) \equiv B'(x,x)(y,y)\,. \tag{A6}$$

APPENDIX B: CONVERGENCE OF THE SUCCESSIVE APPROXIMATION METHOD

The fixed-point theorem[22] states that an iteration procedure like Eq. (29), based upon the mapping $\hat{F}$ of Eq. (27), will converge to a root of Eq. (28) if $\hat{F}$ is a contraction mapping, i.e., if for all $f, g \in L^1$, there exists a real number $k < 1$ such that

$$\|\hat{F}f - \hat{F}g\| \le k\|f - g\| \tag{B1}$$

where the norm $\|\cdots\|$ is defined by

$$\|f\| \equiv \int_{-\infty}^{\infty} |f(y)|\,dy\,.$$

We first show by a counterexample that $\hat{F}$ in Eq. (27) cannot be a contraction mapping for $x \ge \frac{1}{2}$. From Eq. (27), we have

$$\|\hat{F}f - \hat{F}g\| = [x/(1-x)]\,\|\hat{O}f - \hat{O}g\|$$

for any f and g in L^1. Now choose f to be a normalized Compton profile and g to be identically zero. Then

$$\|\hat{F}f - \hat{F}g\| = \frac{x}{1-x}\,\|\hat{O}f\| = \frac{x}{1-x}$$

since $\hat{O}f$ is normalized to 1 by the definition of $\hat{O}$. Since $\|f-g\| = \|f\| = 1$, we have

$$\|\hat{F}f - \hat{F}g\| > k\|f - g\|$$

unless

$$x/(1-x) \le k < 1\,,$$

i.e., unless $x < 0.5$.

We now simplify Eq. (9) by letting $\theta_1 = \theta_2$, and under the assumption $\lambda_0/\lambda' \sim 1$ we obtain: J_{double}

$=J_{\text{single}}*J_{\text{single}}$, where $*$ signifies convolution. Thus the mapping becomes

$$\hat{G}\colon f \to \frac{1}{1-x} J_E - \frac{x}{1-x} f*f \tag{B2}$$

where $f(q)\geq 0$ for all q, and $||f||=||f*f||=1$. We wish to show that G is a contraction mapping if and only if $x<\frac{1}{3}$.

Consider

$$\begin{aligned} f*f-g*g &= f*f-g*g+f*g-f*g \\ &= f*(f-g)+g*(f-g) \\ &= (f-g)*(f+g). \end{aligned}$$

We have

$$\begin{aligned} ||f*f-g*g|| &= ||(f-g)*(f+g)|| \leq ||f-g||\,||f+g|| \\ &= ||f-g||(||f||+||g||). \end{aligned}$$

The last equality holds because $f,g\geq 0$. Since $||f||=||g||=1$, we have

$$||f*f-g*g|| \leq 2||f-g||. \tag{B3}$$

Using Eqs. (B2) and (B3) we obtain

$$||\hat{G}f-\hat{G}g|| = \frac{x}{1-x}||f*f-g*g|| \leq \frac{2x}{1-x}||f-g||.$$

Thus in order for (B1) to hold and $\hat{G}$ to be a contraction map, we must have $2x/(1-x)<1$ or $x<\frac{1}{3}$, proving our assertion.

---

*Work supported in part by the Camille and Henry Dreyfus Foundation.

[1]For a review, see V. Halonen, I. R. Epstein, A. C. Tanner, and B. G. Williams, in *Compton Scattering: The Investigation of Electron Momentum Distributions*, edited by B. G. Williams (McGraw-Hill, New York, to be published).

[2]B. G. Williams and V. Halonen, Phys. Fenn. 10, 5 (1975).

[3]J. Felsteiner and P. Pattison, Nucl. Inst. Meth. 124, 449 (1975).

[4]Paper I: A. C. Tanner and I. R. Epstein, Phys. Rev. A 13, 335 (1976).

[5]Paper III: A. C. Tanner and I. R. Epstein, Phys. Rev. A 14, 328 (1976).

[6]J. Felsteiner, P. Pattison, and M. J. Cooper, Phil. Mag. 30, 537 (1974).

[7]B. G. Williams, P. Pattison, and M. J. Cooper, Phil. Mag. 30, 307 (1974).

[8]W. R. McIntire, Phys. Stat. Sol. 23, 359 (1974).

[9]Paper II: A. C. Tanner and I. R. Epstein, Phys. Rev. A 14, 313 (1976).

[10]V. Halonen, B. G. Williams, and T. Paakari, Phys. Fenn. 10, 107 (1975).

[11]P. Pattison, S. Mannien, J. Felsteiner, and M. J. Cooper, Phil. Mag. 30, 973 (1974).

[12]J. Felsteiner and P. Pattison, Phys. Rev. B 13, 2702 (1976).

[13]Paper V: E. Braun-Keller and I. R. Epstein, following paper, Phys. Rev. A 16, 1154 (1977).

[14]B. G. Williams, Acta. Cryst. A 32, 513 (1976).

[15]It is important to keep in mind that the profile F (as well as the multiply scattered profiles G and D) depends upon λ_0 and θ_p, i.e., on the experimental conditions, though the momentum distribution, which uniquely determines F, depends only upon the nature of the sample.

[16]For simplicity, we also assume that the electron momentum distribution is isotropic.

[17]The simplest way to allow for binding energy effects is to include a factor $f_1 f_2$ in $P^{\theta}P(\Omega_1,\Omega_2)$, where f_i is the fraction of electrons in the scatterer which have binding energies smaller than the Compton energy shift resulting from scattering by the angle θ_i.

[18]A bilinear operator has the properties: $\hat{O}(af,g)=\hat{O}(f,ag)=a\hat{O}(f,g)$; $\hat{O}(f_1,g)+\hat{O}(f_2,g)=\hat{O}(f_1+f_2,g)$; $\hat{O}(f,g_1)+\hat{O}(f,g_2)=\hat{O}(f,g_1+g_2)$.

[19]T. L. Saaty, *Modern Non-Linear Equations* (McGraw-Hill, New York, 1967).

[20]While the choice of a Gaussian J_S makes it possible to do some of the integrations analytically, this was not done, since part of the purpose of the calculation was to assess the timing and numerical accuracy of the computer program.

[21]B. G. Williams, Acta Cryst. A 32, 513 (1976).

[22]M. Rosenlicht, *Introduction to Analysis* (Scott Foresman, Glenview, Ill., 1968), pp. 170–173.

[23]J. Felsteiner and R. Opher (unpublished).

[24]W. A. Rachinger, J. Sci. Inst. 25, 254 (1948).