Inversion of bremsstrahlung spectra emitted by solar plasma

M. Piana
Dipartimento di Fisica dell'Universitá di Genova
and Istituto Nazionale di Fisica Nucleare,
via Dodecaneso 33, I-16146 Genova, Italy

Received 4 August 1993, accepted 18 January 1994

Abstract. Bremsstrahlung radiation coming from the solar atmosphere is linked to the distribution function of electrons in the solar plasma through a Volterra equation of the first kind. We have assumed the Bethe-Heitler approximation for the bremsstrahlung cross section and we have applied the SVD method to the integral equation with discrete data. Using Tikhonov's regularisation technique, reconstructions of the electronic distribution functions have been obtained, both for simulated and real data.

Key words: radiation mechanisms: thermal, non-thermal - methods: analytical - sun: X-rays

1 Introduction

The processes of X-ray emission from astrophysical sources are well known: X-ray radiation in the range from few keV to few MeV is produced by the interaction of electrons with the ambient medium; more precisely, the typical phenomena causing this kind of emission are (Johns & Lin 1992): collisional bremsstrahlung of electrons with the ions of a cosmic plasma; interaction between electrons and a magnetic field in a synchrotron process and inverse Compton interaction of electrons with photons.

It is important to notice that the analysis of the X-ray spectra provides information about the physical processes occuring in the source objects, particularly about the energy transport phenomena and the acceleration mechanisms

of particles (Thompson et al. 1992), such as in the case of the X-ray radiation coming from the plasma in the high solar chromosphere. In flare conditions, the acceleration of $\sim 10 - 100$ keV electrons occurs (Lin 1974; Dennis 1985) and so, for X-ray spectra over the $20 - 100$ keV energy range, the dominant emission mechanism is collisional bremsstrahlung of electrons with the ions of the plasma (Brown 1971). In fact Compton backscattering and synchrotron processes involve electrons of higher energies (Korchak 1971) and so they can be ignored as primary source of X-rays. However, Bay and Ramaty (1978) examined the effects on the observed X-ray spectrum produced by the photons Compton backscattered from the photosphere from a originally isotropic radiation source; more recently, Johns and Lin (1992) showed that this alteration of the photon spectrum have measurable small effects on the deconvolved electron spectrum. Anyway, in this paper Compton backscattering of photons emitted toward the solar photosphere will be neglected.

The bremsstrahlung of electrons with the ions of the plasma is a collisional process and so it is characterized by a cross section which, generally, depends on the energy of the photons and of the X-ray producing electrons. If this cross section is analytically known, the photon spectrum can be directly related to the electron distribution. Generally, the study of the emission process can be made under particular hypotheses about the physical conditions of the source. First of all, bremsstrahlung radiation is considered optically thin (Brown & Emslie 1988), that is absorption can be neglected; this implies that the observed X-ray spectra are quite similar to the ones emerging from the emitting region. Moreover the electron velocity distribution is characterized by isotropic conditions. Finally the plasma is assumed to be hydrogen dominated, so that the ions are almost completely protons. Under these hypotheses, the equation linking the distribution function of electrons with the photon spectrum can be written in the following way (Brown & Craig 1986):

$$g(\epsilon) = \int_V n(\mathbf{r}) \int_\epsilon^\infty F(E, \mathbf{r}) K(\epsilon, E) dE d\mathbf{r} \tag{1}$$

where V is the volume of the emitting region, $n(\mathbf{r})$ is the density of the ions in the plasma, E is the electron energy, ϵ is the photon energy, $F(E, \mathbf{r})$ is the electron distribution function, $K(\epsilon, E)$ is the bremsstrahlung cross section, $g(\epsilon)$ is the total rate of photon emission, in (photon/sec) per unit ϵ. It must be noted that g represents the flux at the sun; in practice, measurements regard fluxes at the earth, but here only the shape of the spectrum and the relative errors in it are considered, so that multiplicative constants are irrelevant.
Averaging the electron distribution function over the volume of the emitting region and weighting this function with the density of the ions, one obtains:

$$g(\epsilon) = \int_\epsilon^\infty \overline{F}(E) K(\epsilon, E) dE \tag{2}$$

$$\overline{F}(E) = \frac{1}{\overline{n}V} \int_V n(\mathbf{r}) F(E, \mathbf{r}) d\mathbf{r} \tag{3}$$

$$\overline{n} = \frac{1}{V} \int_V n(\mathbf{r}) d\mathbf{r} \tag{4}$$

In general, the bremsstrahlung cross section $K(\epsilon, E)$ can be written in the following form:

$$K(\epsilon, E) = \frac{1}{\epsilon E} k(\epsilon, E) \tag{5}$$

where $k(\epsilon, E)$ is a particular function which depends on the kind of approximation. In the Bethe-Heitler approximation (Koch & Motz 1959):

$$k(\epsilon, E) = K_\circ \log \frac{1 + \sqrt{1 - \frac{\epsilon}{E}}}{1 - \sqrt{1 - \frac{\epsilon}{E}}} \quad E \geq \epsilon \tag{6}$$

$$k(\epsilon, E) = 0 \qquad E < \epsilon \tag{7}$$

where, for hydrogen:

$$K_\circ = \frac{8}{3} \frac{r_\circ^2}{137} mc^2 \tag{8}$$

with $r_\circ$ the classical radius and m the rest mass of the electron respectively. The Bethe-Heitler formula holds in the following conditions (Koch & Motz 1959): $\frac{2\pi Z}{137\beta} \ll 1$ (Born approximation); $\frac{137}{\sqrt{Z}} \gg \frac{E_0 E}{k}$ (no screening effects); $\beta \ll 1$ (non relativistic limit) where Z is the atomic number of the element of the target (in this case $Z = 1$), β is the ratio between the speed of the electron and the speed of the light, E_0 and E are the initial and final energies of the electron and k is the energy of the emitted photon. For the electron energies considered here, the Bethe-Heitler formula is a good approximation (though, as usual for the Born approximation, it is smaller than the true cross section).

Through the substitution of expression (6),(7) into Eq. (2), one obtains:

$$g(\epsilon) = \frac{1}{\epsilon} \int_\epsilon^\infty f(E) \log \frac{1 + \sqrt{1 - \frac{\epsilon}{E}}}{1 - \sqrt{1 - \frac{\epsilon}{E}}} \frac{dE}{E} \tag{9}$$

where

$$f(E) = K_\circ \overline{F}(E) \tag{10}$$

Equation (9) is a Volterra equation of first kind (Riesz & Nagy 1953). By a change of variables (Brown 1971), it can be reduced to the Abel equation (Courant & Hilbert 1989) and therefore it is possible to give its exact solution (Brown & Craig 1986):

$$f(E) = \frac{E^{\frac{1}{2}}}{\pi} \frac{d}{dE} \int_E^\infty \left(-\frac{d(\epsilon g(\epsilon))}{d\epsilon} \frac{1}{\sqrt{\epsilon - E}} \right) d\epsilon \tag{11}$$

This analytic solution of the Volterra equation cannot be used in practical cases: in fact, in order to obtain the electron distribution from Eq. (11), it should be necessary to know $g(\epsilon)$ on the whole energy interval (up to infinite photon energy) and with such an accuracy to be able to evaluate its second derivative. On the contrary, in the reality, detectors provide a data vector $\mathbf{g}$, whose generic component, affected by noise, is the integral of $g(\epsilon)$ on a given range of energy.

2 The singular system of the integral operator

First of all, Eq. (9) is considered in the case where the data function $g(\epsilon)$ is given in a finite set of points ϵ_n, i.e. the integration effect of the detectors is neglected. In such a case we have:

$$g_n = \frac{1}{\epsilon_n} \int_{\epsilon_n}^{\infty} f(E) \log \frac{1+\sqrt{1-\frac{\epsilon_n}{E}}}{1-\sqrt{1-\frac{\epsilon_n}{E}}} \frac{dE}{E} \qquad n = 1, \ldots, N \tag{12}$$

One can define the functions ϕ_n so that:

$$\phi_n(E) = 0 \qquad E \leq \epsilon_n \tag{13}$$

$$\phi_n(E) = \frac{1}{\epsilon_n} \frac{1}{E} \log \frac{1+\sqrt{1-\frac{\epsilon_n}{E}}}{1-\sqrt{1-\frac{\epsilon_n}{E}}} \qquad E > \epsilon_n \tag{14}$$

In this way, if the distribution function of electrons belongs to $L^2(0,\infty)$ (that is to the space of functions f such that $\int_0^\infty |f(x)|^2 dx < \infty$) the expression (f, ϕ_n) can be introduced in order to indicate the scalar product between the functions f and ϕ_n in the Hilbert space $L^2(0,\infty)$, i.e.:

$$\frac{1}{\epsilon_n} \int_{\epsilon_n}^{\infty} f(E) \log \frac{1+\sqrt{1-\frac{\epsilon_n}{E}}}{1-\sqrt{1-\frac{\epsilon_n}{E}}} \frac{dE}{E} \equiv (f, \phi_n) \qquad n = 1, \ldots, N \tag{15}$$

At this point, it is natural to define the operator $L : L^2(0,\infty) \to Y$ which transforms f into the vector whose n-th component is:

$$(Lf)_n = (f, \phi_n) \qquad n = 1, \ldots, N \tag{16}$$

Then, the linear inverse problem with discrete data (12) assumes the operator form:

$$\mathbf{g} = Lf \tag{17}$$

In general (Bertero et al. 1985), the data space Y has an euclidean structure, defined by the scalar product:

$$(\mathbf{g}, \mathbf{h}) = \sum_{n=1}^{N} w_n g_n h_n^* \qquad . \tag{18}$$

In this paper the weights are chosen in such a way that the scalar product (18) approximates the L^2-scalar product. To this aim, in the case of uniform sampling, i.e.:

$$\epsilon_n = \epsilon_1 + d(n-1) \qquad n = 1, \ldots, N \tag{19}$$

it is (Bertero et al 1985):

$$w_n = d \quad , \qquad n = 1, \ldots, N \tag{20}$$

while in the case of geometrical sampling, i.e.:

$$\epsilon_n = \epsilon_1 \Delta^{n-1} \qquad n = 1, \ldots, N \tag{21}$$

a good choice for the weights w_n is (Bertero et al. 1984):

$$w_n = (\log \Delta)\epsilon_n \quad n = 1, \ldots, N \quad . \tag{22}$$

In this particular problem geometrical sampling provides better results and therefore it will be preferred. The reasons for this greater efficiency reside, first of all, in the fact that geometrical sampling is more appropriate for equations of convolution with respect to the dilation group (i.e. those integral equations whose kernel depends on the ratio between the variable of the data function and the one of the solution). Moreover, from the data analysis point of view, data are summed in bins of width which increases logarithmically with photon energy.

As L is a linear finite-rank operator, it is always possible to introduce its singular system (Bertero et al. 1985), that is the set of triples $\{\sigma_k; u_k, \mathbf{v}_k\}_{k=1}^N$ such that:

$$Lu_k = \sigma_k \mathbf{v}_k \qquad k = 1, \ldots, N \tag{23}$$

$$L^* \mathbf{v}_k = \sigma_k u_k \qquad k = 1, \ldots, N \tag{24}$$

with L^* the adjoint of L.

The computation of the singular system of L is based on the observation that the singular values are the square root of the eigenvalues of the operator LL^* and on the fact that this is represented by the matrix:

$$LL^* = G^T W \tag{25}$$

where W is the matrix defined by:

$$W_{nm} = w_n \delta_{nm} \qquad n, m = 1, \ldots, N \tag{26}$$

and G is the Gram matrix (Bertero et al. 1985):

$$G_{nm} = (\phi_n, \phi_m) \tag{27}$$

With the help of MAXIMA, a software package for symbolic calculation (Computer Aided Mathematics Group of Symbolics 1988), we have found closed expressions for the matrix elements of the Gram matrix. The result is:
for $m > n$:

$$G_{mn} = \frac{4}{(\epsilon_n \epsilon_m)^{\frac{3}{2}}} \log \frac{1+\sqrt{\frac{\epsilon_n}{\epsilon_m}}}{1-\sqrt{\frac{\epsilon_n}{\epsilon_m}}} - \frac{2}{\epsilon_n (\epsilon_m)^2} \log \frac{\epsilon_n}{\epsilon_m} + \frac{2}{\epsilon_n (\epsilon_m)^2} \left(1+\frac{\epsilon_m}{\epsilon_n}\right) \log \left(1-\frac{\epsilon_n}{\epsilon_m}\right) \tag{28}$$

for $m = n$:

$$G_{mm} = \frac{1}{(\epsilon_m)^3} 8log2 \tag{29}$$

For $m < n$:

$$G_{mn} = G_{nm} \tag{30}$$

By standard routines one can compute the eigenvalues of $G^T W$ and therefore the singular values of L. Table 1 contains some singular values for $N = 20$, $N = 40$, $N = 80$ sampled points. The energy range is the same, that is: $\epsilon_1 = 10$ keV, $\epsilon_N = 100$ keV.

k	$N = 20$	$N = 40$	$N = 80$
1	0.17119	0.16604	0.16359
2	$3.58509 \text{x} 10^{-2}$	$3.47064 \text{x} 10^{-2}$	$3.41459 \text{x} 10^{-2}$
3	$1.51269 \text{x} 10^{-2}$	$1.45881 \text{x} 10^{-2}$	$1.43213 \text{x} 10^{-2}$
4	$8.67455 \text{x} 10^{-3}$	$8.34862 \text{x} 10^{-3}$	$8.18690 \text{x} 10^{-3}$
5	$5.69975 \text{x} 10^{-3}$	$5.47583 \text{x} 10^{-3}$	$5.36655 \text{x} 10^{-3}$
6	$4.10162 \text{x} 10^{-3}$	$3.93174 \text{x} 10^{-3}$	$3.85213 \text{x} 10^{-3}$
7	$3.12954 \text{x} 10^{-3}$	$2.98723 \text{x} 10^{-3}$	$2.92571 \text{x} 10^{-3}$
8	$2.49864 \text{x} 10^{-3}$	$2.36838 \text{x} 10^{-3}$	$2.31858 \text{x} 10^{-3}$
9	$2.06407 \text{x} 10^{-3}$	$1.93530 \text{x} 10^{-3}$	$1.89325 \text{x} 10^{-3}$
10	$1.74934 \text{x} 10^{-3}$	$1.62079 \text{x} 10^{-3}$	$1.58396 \text{x} 10^{-3}$
11	$1.50107 \text{x} 10^{-3}$	$1.38321 \text{x} 10^{-3}$	$1.34974 \text{x} 10^{-3}$
12	$1.28968 \text{x} 10^{-3}$	$1.19970 \text{x} 10^{-3}$	$1.16830 \text{x} 10^{-3}$
13	$1.10402 \text{x} 10^{-3}$	$1.05431 \text{x} 10^{-3}$	$1.02383 \text{x} 10^{-3}$

Table.1. Comparison of singular values corresponding to different numbers of geometrically sampled points, in the energy range $\epsilon_1 = 10$ keV, $\epsilon_N = 100$ keV.

The singular functions can be computed through the formula (Bertero et al. 1985):

$$u_k = \frac{1}{\sigma_k} \sum_{n=1}^{N} w_n (\mathbf{v}_k)_n \phi_n \tag{31}$$

In Fig. 1 the first six singular functions are represented, for $N = 40$, $\epsilon_1 = 10$ keV and $\epsilon_N = 100$ keV. It is interesting to note that every singular function

is zero at $E = \epsilon_1$ (this is due to the fact that the functions ϕ_n are zero at $E = \epsilon_1$) and that the k-th singular function has $k-1$ zeroes in the interval (ϵ_1, ∞). As the differences of the singular functions corresponding to $N = 20$ and $N = 80$ with respect to those corresponding to $N = 40$ are negligible, they are not represented here.

Finally, it can be shown (Bertero et al. 1985) that there is a unique solution of the system:

$$\|Lf - \mathbf{g}\| = minimum \tag{32}$$

$$\|f\| = minimum \tag{33}$$

This function is called the generalized solution or normal pseudosolution and it can be represented in the form:

$$f^\dagger = \sum_{k=1}^{N} \frac{1}{\sigma_k}(\mathbf{g}, \mathbf{v}_k)u_k \tag{34}$$

As a consequence, it is natural to introduce the generalized inverse operator $L^\dagger$ such that

$$L^\dagger : Y \to L^2(0, \infty)$$

and

$$L^\dagger \mathbf{g} = f^\dagger \tag{35}$$

The following relationship occurs (Bertero et al. 1985):

$$\frac{\|\delta f^\dagger\|}{\|f^\dagger\|} \leq C(L)\frac{\|\delta \mathbf{g}\|}{\|\mathbf{g}\|} \tag{36}$$

where $C(L) = \|L\|\|L^\dagger\| = \frac{\sigma_1}{\sigma_N}$ is the so called condition number. If $C(L) \gg 1$, the problem of computing the generalized solution is ill-conditioned and this ill-conditioning is related to the fact that the original problem (2) is ill-posed. As one can see in Table 1, the first singular values slowly decrease for increasing N: in fact they converge to the singular values of the integral operator when $N \to \infty$. On the other hand $C(L)$ is rapidly increasing. We have $C(L) \simeq 630$ for $N = 20$, $C(L) \simeq 2000$ for $N = 40$ and $C(L) \simeq 6100$ for $N = 80$. As usual, the ill-conditioning of the problem increases for increasing values of N, since the singular values of the integral operator tend to zero.

3 Reconstruction algorithms (simulated data)

In order to reduce the ill-conditioning of the problem, an extremely direct method consists of using only $M < N$ singular values, singular vectors and

singular functions in the expansion (34). In other words, an estimate of the generalized solution is represented by:

$$f_M = \sum_{k=1}^{M} \frac{1}{\sigma_k}(\mathbf{g}, \mathbf{v}_k)u_k \tag{37}$$

where α_M is the last singular values greater than the signal-to- noise ratio (Bertero et al. 1984). Following this approach, which is called the truncated SVD method, the condition number clearly decreases though, in the present case, the results are not good; the reason could be due to the fact that this problem is not extremely ill-conditioned and so also, by truncating the expansion (34), a significative loss of information occurs.

An alternative, more sophisticated method is represented by regularisation theory. By definition (Bertero et al. 1987), a regularisation algorithm is a one parameter family of operators $\{R_\lambda\}_{\lambda>0}$:

$$R_\lambda : Y \to L^2(0, \infty)$$

characterized by precise properties and such that for a given choice of the regularisation parameter, the regularised solution

$$f_\lambda = R_\lambda \mathbf{g} \tag{38}$$

provides an approximation of the generalized solution.

It is useful to note (Bertero et al. 1987) that a noisy data vector $\mathbf{g}$ can be represented in the form

$$\mathbf{g} = Lf^\dagger + \mathbf{h} \tag{39}$$

where $Lf^\dagger$ is the noise free signal and $\mathbf{h}$ is a vector representing the noise (note that $\mathbf{h}$ can be signal dependent). So one obtains:

$$f_\lambda - f^\dagger = (R_\lambda L f^\dagger - f^\dagger) + R_\lambda \mathbf{h} \tag{40}$$

In the rhs of Eq. (40), the first term represents the approximation error induced by the introduction of R_λ instead of the generalized inverse $L^\dagger$; as $lim_{\lambda\to 0} R_\lambda = L^\dagger$ (Bertero et al. 1987), then this term tends to zero when $\lambda \to 0$. On the contrary, the second term represents the error induced by the noise and it grows to very large values when $\lambda \to 0$. So, one of the crucial points of the regularisation theory is the choice of the regularisation parameter which allows a compromise between the approximation error and the error propagation from the data to the solution.

A particular regularisation method is represented by Tikhonov's technique (Tikhonov 1963); it was first introduced in the case of Fredholm integral equations and then it was extended by Miller (Miller 1970) to the more general case of ill-posed problems in Hilbert spaces. In the case of inverse problems with

discrete data, this algorithm can be obtained (Bertero et al. 1987) minimising the functional

$$\Phi_\lambda[f] = \|Lf - \mathbf{g}\|^2 + \lambda\|f\|^2 \tag{41}$$

The result represents the regularised solution, whose explicit form is:

$$f_\lambda = \sum_{k=1}^{N} \frac{\sigma_k}{\sigma_k^2 + \lambda}(\mathbf{g}, \mathbf{v}_k)u_k \tag{42}$$

As this equation shows, a great advantage of Tikhonov's algorithm resides in the fact that it can be easily computed if one knows the singular system of the integral operator L. In the present case, the Gram matrix elements are analytically known and so Tikhonov's regularisation seems to be particularly suitable. It must be noted that this kind of approach avoids the discretisation of the solution and the problems of instability that every kind of numerical approximation implies in an ill- conditioned problem. In order to verify the fitness of Tikhonov's regularisation to the present case, this method has been firstly applied to simulated data affected by noise; secondly, the algorithm was also used to invert a set of real data.

The simulated data have been obtained by applying the operator L to a given distribution function of the electrons $f(E)$ and then by affecting the vector Lf with gaussian noise with zero mean and standard deviation equal to $2\eta(Lf)_n$. Here η represents the relative error on the data. As regards the functions $f(E)$, a typical form is:

$$f(E) = AE^{-\alpha} \tag{43}$$

For $A = 1$ and $\alpha = 3$, we have:

$$(Lf)_n = \frac{16}{45}\epsilon_n^{-4}$$

The simulated data obtained by affecting the vector Lf with gaussian noise is represented in Fig. 2(a), where it has been assumed again: $\epsilon_1 = 10$ keV, $\epsilon_N = 100$ keV, $N = 40$; the relative error is $\eta = 0.02$. By inserting this data vector into Tikhonov's formula (42) one obtains the regularised reconstruction of the electron distribution function.

As regards the estimate of the error on the solution, it must be noted that f_λ represents an approximation of the unknown source function, but, as the original problem (2) is ill-posed, the uncertainty with which the solution can be known may be infinite. Therefore, what can be done is to evaluate the so called confidence limit of the regularised solution; that is, for a given choice of the regularisation algorithm and of the regularisation parameter, one can determine the error propagation from the data to the regularised solution. To this aim, one can perform different regularised reconstructions (with the same value of the regularisation parameter), corresponding to different simulated vectors obtained by modifying the real data vector with random components having zero mean

and variance equal to the one used in the simulation. In this way, one obtains a 'strip' of regularised solutions whose upper and lower borders are the confidence limit of the reconstructed function, as every function in this 'strip' has the property of fitting the simulated data.

As regards the estimate of the resolution, we observe that the singular functions which significatively contribute to the reconstruction of the source function (through Eq. (42)) are certainly less than N; for example, in this case, the coefficients $\frac{\sigma_k}{\sigma_k^2+\lambda}|(\mathbf{g},\mathbf{v}_k)|$ become very small for $k > 29$. On the other hand, as already noted in Sect. 2, the singular function of order k is characterized by $k-1$ zeroes in (ϵ_1,∞) which approximately form a geometric progression of ratio δ:

$$E_n = E_1\delta^{n-1} \qquad n = 1,\ldots,k \tag{44}$$

Therefore, the reconstructed solution does not contain details in the intervals whose amplitudes coincide with the distances between adjacent zeroes of the last significative singular functions. This implies a uniform resolution, given by $log\delta$, in the variable $logE$. So, taking the average value of the ratios between two adjacent zeroes of the singular function of order $M = 29$, δ can be estimated equal to 1.09 and $E_1 = 10$ keV.

The regularised reconstruction of the electron distribution function (43) is represented in Fig. 2(b), in some discrete values of the electron energy, for $\lambda = 8\text{x}10^{-6}$; bars corresponding to the estimated resolution and to the confidence limit have been plotted too. It must be noted that the value of the regularisation parameter has been chosen by comparing the regularised solution with the theoretical distribution function (43). Moreover, it is important to note that the regularisation method gives a good reconstruction of the distribution function only in the range (ϵ_1,ϵ_N), that is the range where we have information about the photon spectrum.

Another interesting form of the distribution function of electrons is:

$$f(E) = AE^{-\delta_1} \qquad E \le E_1 \tag{45}$$

$$f(E) = BE^{-\delta_2} \qquad E > E_1 \tag{46}$$

It can also be assumed that this function is continuous at $E = E_1$. In the case $A = 1$, $\delta_1 = 2$, $\delta_2 = 4$ the vector Lf has the form:
for $E_1 > \epsilon_n$:

$$(Lf)_n = \frac{1}{\epsilon_n}\left[-\frac{E_1^{-2}}{4}\log\frac{1+\sqrt{1-\frac{\epsilon_n}{E}}}{1-\sqrt{1-\frac{\epsilon_n}{E}}} + \epsilon_n^{-2}\left(\sqrt{1-\frac{\epsilon_n}{E_1}} - \frac{1}{3}\left(\sqrt{1-\frac{\epsilon_n}{E_1}}\right)^3\right) + \right.$$

$$\left. + \frac{E_1^2}{2}\epsilon_n^{-4}\left(\frac{16}{45} - \sqrt{1-\frac{\epsilon_n}{E_1}} + \left(\sqrt{1-\frac{\epsilon_n}{E_1}}\right)^3 - \frac{3}{5}\left(\sqrt{1-\frac{\epsilon_n}{E_1}}\right)^5 + \frac{1}{7}\left(\sqrt{1-\frac{\epsilon_n}{E_1}}\right)^7\right)\right]$$

for $E_1 \le \epsilon_n$:

$$(Lf)_n = \frac{72}{35}\epsilon_n^{-5}$$

Figure 3(a) represents the simulated data, obtained by the method previously described. The physical parameters are: $\epsilon_1 = 1$ keV, $\epsilon_N = 10$ keV, $E_1 = 3$ keV, $N = 40$ and the relative error $\eta = 0.01$. The regularised solution, corresponding to $\lambda = 1.1\text{x}10^{-3}$ is represented in Fig. 3(b).

A third type of function $f(E)$ is a power law with a smooth low energy turn over, whose analytic form is:

$$f(E) = AE^{-\delta} exp\left(-\frac{E_1}{E}\right) \tag{47}$$

The vector Lf is obtained by computing numerically the N integrals defined in (12). The simulated data used in the inversion are represented in Fig. 4(a), while Fig. 4(b) shows the corresponding regularised solution. Here $E_1 = 40$ keV, $\delta = 3$, $A = 1$; the photon energy range is $\epsilon_1 = 1$ keV, $\epsilon_N = 100$ keV; the number of sampled points is $N = 40$, while the value of the relative error η is 0.01 for $\epsilon < 10$ keV and 0.02 otherwise; the value of the regularisation parameter is $\lambda = 2\text{x}10^{-4}$.

4 The case of binned data

If one takes into account the fact that the detectors provide integrals of the spectrum over finite ranges of photon energy, the equation with discrete data takes the form:

$$g_n = \int_{\epsilon_n}^{\epsilon_{n+1}} d\epsilon \frac{1}{\epsilon} \left(\int_{\epsilon}^{\infty} f(E) \log \frac{1+\sqrt{1-\frac{\epsilon}{E}}}{1-\sqrt{1-\frac{\epsilon}{E}}} \frac{dE}{E} \right) \quad n = 1, \ldots, N \tag{48}$$

The problem can be written again in the form (17), with L defined as in Eq. (16), where, now, the functions ϕ_n are given by:
for $E < \epsilon_n$:

$$\phi_n(E) = 0 \tag{49}$$

for $\epsilon_n \leq E < \epsilon_{n+1}$

$$\phi_n(E) = \frac{1}{E} \int_{\epsilon_n}^{E} \log \frac{1+\sqrt{1-\frac{\epsilon}{E}}}{1-\sqrt{1-\frac{\epsilon}{E}}} \frac{d\epsilon}{\epsilon} \tag{50}$$

for $E \geq \epsilon_{n+1}$:

$$\phi_n(E) = \frac{1}{E} \int_{\epsilon_n}^{\epsilon_{n+1}} \log \frac{1+\sqrt{1-\frac{\epsilon}{E}}}{1-\sqrt{1-\frac{\epsilon}{E}}} \frac{d\epsilon}{\epsilon} \tag{51}$$

By defining the functions:

$$\phi_n^{(1)}(E) = \frac{1}{E} \int_{\epsilon_n}^{E} \log \frac{1+\sqrt{1-\frac{\epsilon}{E}}}{1-\sqrt{1-\frac{\epsilon}{E}}} \frac{d\epsilon}{\epsilon} \tag{52}$$

$$\phi_n^{(2)}(E) = \frac{1}{E}\int_{\epsilon_n}^{\epsilon_{n+1}} \log\frac{1+\sqrt{1-\frac{\epsilon}{E}}}{1-\sqrt{1-\frac{\epsilon}{E}}}\frac{d\epsilon}{\epsilon} \tag{53}$$

and using again MAXIMA, we have found the following closed expressions:

$$\phi_n^{(1)}(E) = \frac{1}{E}\left[-2li_2\left(\frac{1}{2}\right) - \log(\epsilon_n)\log\frac{1+\sqrt{1-\frac{\epsilon_n}{E}}}{1-\sqrt{1-\frac{\epsilon}{E}}} + \right.$$

$$+2\log\left(1-\sqrt{1-\frac{\epsilon_n}{E}}\right)\log\frac{1+\sqrt{1-\frac{\epsilon_n}{E}}}{2} - \frac{1}{2}\left(\log\left(1+\sqrt{1-\frac{\epsilon_n}{E}}\right)\right)^2 -$$

$$-\log\left(1+\sqrt{1-\frac{\epsilon_n}{E}}\right)\left(\log\left(1-\sqrt{1-\frac{\epsilon_n}{E}}\right)-\log\epsilon_n\right)+\frac{1}{2}\left(\log\left(1-\sqrt{1-\frac{\epsilon_n}{E}}\right)\right)^2 -$$

$$\left.-\log\epsilon_n\log\left(1-\sqrt{1-\frac{\epsilon_n}{E}}\right)+2li_2\left(\frac{1-\sqrt{1-\frac{\epsilon_n}{E}}}{2}\right)\right]$$

and:

$$\phi_n^{(2)}(E) = \frac{1}{E}\left[\log(\epsilon_{n+1})\log\frac{1+\sqrt{1-\frac{\epsilon_{n+1}}{E}}}{1-\sqrt{1-\frac{\epsilon_{n+1}}{E}}} + \right.$$

$$-2\log\left(1-\sqrt{1-\frac{\epsilon_{n+1}}{E}}\right)\log\frac{1+\sqrt{1-\frac{\epsilon_{n+1}}{E}}}{2} + \frac{1}{2}\left(\log\left(1+\sqrt{1-\frac{\epsilon_{n+1}}{E}}\right)\right)^2 +$$

$$+\log\left(1+\sqrt{1-\frac{\epsilon_{n+1}}{E}}\right)\left(\log\left(1-\sqrt{1-\frac{\epsilon_{n+1}}{E}}\right)-\log\epsilon_{n+1}\right)-\frac{1}{2}\left(\log\left(1-\sqrt{1-\frac{\epsilon_{n+1}}{E}}\right)\right)^2 +$$

$$+\log\epsilon_{n+1}\log\left(1-\sqrt{1-\frac{\epsilon_{n+1}}{E}}\right) - 2li_2\left(\frac{1-\sqrt{1-\frac{\epsilon_{n+1}}{E}}}{2}\right) -$$

$$-\log(\epsilon_n)\log\frac{1+\sqrt{1-\frac{\epsilon_n}{E}}}{1-\sqrt{1-\frac{\epsilon}{E}}} +$$

$$+2\log\left(1-\sqrt{1-\frac{\epsilon_n}{E}}\right)\log\frac{1+\sqrt{1-\frac{\epsilon_n}{E}}}{2} - \frac{1}{2}\left(\log\left(1+\sqrt{1-\frac{\epsilon_n}{E}}\right)\right)^2 -$$

$$-\log\left(1+\sqrt{1-\frac{\epsilon_n}{E}}\right)\left(\log\left(1-\sqrt{1-\frac{\epsilon_n}{E}}\right)-\log\epsilon_n\right)+\frac{1}{2}\left(\log\left(1-\sqrt{1-\frac{\epsilon_n}{E}}\right)\right)^2 -$$

$$\left.-\log\epsilon_n\log\left(1-\sqrt{1-\frac{\epsilon_n}{E}}\right)+2li_2\left(\frac{1-\sqrt{1-\frac{\epsilon_n}{E}}}{2}\right)\right]$$

where li_2 stands for the dilogarithm function defined as:

$$li_2(x) = -\int_0^\infty \frac{log|1-t|}{t} dt$$

In terms of these functions the matrix elements of the Gram matrix are given by:
if $m > n$

$$G_{mn} = \int_{\epsilon_m}^{\epsilon_{m+1}} \phi_n^{(2)}(E)\phi_m^{(1)}(E)dE + \int_{\epsilon_{m+1}}^{\infty} \phi_n^{(2)}(E)\phi_m^{(2)}(E)dE \qquad (54)$$

if $m = n$

$$G_{mm} = \int_{\epsilon_m}^{\epsilon_{m+1}} \phi_m^{(1)}(E)\phi_m^{(1)}(E)dE + \int_{\epsilon_{m+1}}^{\infty} \phi_m^{(2)}(E)\phi_m^{(2)}(E)dE \qquad (55)$$

if $m < n$

$$G_{mn} = G_{nm} \qquad (56)$$

Finally, the integrals (54) and (55) can be computed numerically.

k	point case	binned case
1	0.16604	0.16078
2	3.47064×10^{-2}	3.35749×10^{-2}
3	1.45881×10^{-2}	1.41163×10^{-2}
4	8.34862×10^{-3}	8.06932×10^{-3}
5	5.47583×10^{-3}	5.28043×10^{-3}
6	3.93174×10^{-3}	3.77813×10^{-3}
7	2.98723×10^{-3}	2.85728×10^{-3}
8	2.36838×10^{-3}	2.25234×10^{-3}
9	1.93530×10^{-3}	1.82781×10^{-3}
10	1.62079×10^{-3}	1.51840×10^{-3}
11	1.38321×10^{-3}	1.28374×10^{-3}
12	1.19970×10^{-3}	1.10156×10^{-3}
13	1.05431×10^{-3}	9.56291×10^{-4}

Table 2. Comparison of singular values in the case of point data and in the case of binned data ($N = 40$ geometrically sampled points in the range $\epsilon_1 = 10$ keV, $\epsilon_N = 100$ keV).

As regards the choice of the weights in the scalar product of the space of the binned data, some considerations, based on the use of Schwarz's inequality, suggest, for uniform sampling (Bertero & Pike 1991):

$$w_n = \frac{1}{d} \qquad n = 1, \ldots, N \qquad (57)$$

and, for geometrical sampling:

$$w_n = \frac{1}{\epsilon_1} \frac{1}{\Delta^n - \Delta^{n-1}} \qquad n = 1, \ldots, N \tag{58}$$

Table 2 shows that in the case $N = 40$, $\epsilon_1 = 10$ keV, $\epsilon_N = 100$ keV, there is not a great difference between singular values in the ideal situation of a spectrum which is known at the sampled photon energy and those corresponding to binned data (this is because the numerical $\Delta\epsilon$ is already small compared with the kernel resolution). More precisely, in the last case there are 12 singular values $> 10^{-3}$ instead of 13; this means that the integration between sampled points implies a small loss of information. It must be also noted that the differences between the singular functions in the two cases are completely negligible.

The construction of simulated data can be performed in a way similar to the one described in Sect. 3. Tikhonov regularisation algorithm provides reconstructions of $f(E)$ which are very similar to the ones contained in Fig. 2(b), Fig. 3(b), Fig. 4(b).

5 The case of real data

The application to real data refers to a spectrum emitted by the June 27 1980 solar flare and observed by Germanium detectors placed on a balloon (Lin & Schwartz 1987). These data correspond to a time interval around 16:16 UT, approximately the same of the results presented by Johns & Lin (1992) and by Thompson et al. (1992) (Fig. 5 of the Johns & Lin paper, Fig. 4 of the Thompson et al. paper). The data vector is represented in Fig. 5(a): on the X-axis we give the lowest energies of each channel, measured in keV, while, on the Y-axis we give the integral of the photon flux between the lowest energy and the highest energy of a channel, which coincides with the lowest energy of the following channel; the units on the Y-axis are $\mathrm{cm}^{-2}\mathrm{s}^{-1}\mathrm{keV}^{-1}$. We have $N = 37$ channels, in the range $\epsilon_1 = 14.09$ keV, $\epsilon_N = 286.99$ keV.

In real data application, the problem of the choice of the regularisation parameter becomes crucial. In this paper two criteria have been followed. The first is that known as Generalized Cross Validation (GCV) (Golub et al. 1979; Craven & Wahba 1979): the value of λ is given by the minimiser of the function (Bertero et al. 1987):

$$V(\lambda) = (N^{-1} Tr[I - A(\lambda)])^{-2} (N^{-1} \|[I - A(\lambda)]\mathbf{g}\|^2) \tag{59}$$

where

$$A(\lambda) = [LL^*][LL^* + \lambda I]^{-1} \tag{60}$$

In the present case, GCV provides $\lambda = 2.4\mathrm{x}10^{-6}$ and the regularised solution corresponding to this value of the regularisation parameter is represented in Fig.

5(b) (solid line); as it was already observed (Thompson et al. 1992), the reconstruction obtained with the GCV criterion is characterized by undersmoothed properties.

The other criterion, the Morozov's discrepancy principle (Bertero et al. 1987), provides a larger value of λ, that is $\lambda = 1.13\text{x}10^{-5}$; this number represents the solution of the equation

$$\|Lf_\lambda - \mathbf{g}\|^2 = \|\delta\mathbf{g}\|^2 \tag{61}$$

where the norm is that introduced in the data space. The regularised solution corresponding to the discrepancy principle value of the regularisation parameter is represented again in Fig. 5(b) (dashed line).

The error on the regularised solution has been estimated using the procedure described in Sect. 3; the 'strips' of regularised solutions, representing the confidence limit of the reconstructed function, are plotted in Fig. 6(a) for λ(GCV) and in Fig. 6(b) for λ(discrepancy principle). In these figures, the central part of the regularised solutions has been put in evidence, because the reconstructions are more stable for electron energies between 20 keV and 200 keV; for energies below 20 keV and over 200 keV the error on the solution sensibly grows and this should be due to the fact that the photon spectrum has large relative errors at these energies.

As regards the resolution, in Table 3 the coefficients $\frac{\sigma_k}{\sigma_k^2+\lambda}|(\mathbf{g}, \mathbf{v}_k)|$ are represented with λ given by the discrepancy principle; as one can see, these values indicate that the contribution provided by the last thirteen singular functions is negligible and so the sum (42) can be truncated at $M = 24$ without any loss of information.

k	$\frac{\sigma_k}{\sigma_k^2+\lambda}\|(\mathbf{g},\mathbf{v}_k)\|$	k	$\frac{\sigma_k}{\sigma_k^2+\lambda}\|(\mathbf{g},\mathbf{v}_k)\|$
1	2086.6847	20	151.3343
2	6185.9608	21	47.3225
3	8360.7937	22	11.9774
4	8098.9336	23	8.1065
5	5287.6286	24	3.0088
6	5287.6286	25	0.9804
7	3861.5898	26	0.5284
8	2880.9123	27	$6.8371\text{x}10^{-2}$
9	2302.8453	28	0.1294
10	1956.4356	29	0.1433
11	1704.5650	30	$6.1861\text{x}10^{-2}$
12	1443.9929	31	$8.9522\text{x}10^{-2}$
13	1319.3554	32	$9.4724\text{x}10^{-2}$
14	1069.2384	33	$6.1310\text{x}10^{-2}$
15	911.9663	34	$3.7625\text{x}10^{-3}$
16	758.7604	35	$2.5381\text{x}10^{-2}$
17	452.6179	36	$2.7757\text{x}10^{-2}$
18	431.3417	37	$5.3007\text{x}10^{-3}$
19	170.1266		

Table 3. Values of the coefficients $\frac{\sigma_k}{\sigma_k^2+\lambda}|(\mathbf{g},\mathbf{v}_k)|$ in the case of the real data (λ according to discrepancy principle).

Taking the average value of the ratios between two adjacent zeroes of the singular function of order $M = 24$, the resolution ratio δ can be estimated equal to 1.12 and $E_1 = 14.09$ keV. Note that the resolution on $f(E)$ is worst than the resolution on $g(\epsilon)$, since for the photon spectrum $\delta \simeq 1.08$, so that data are probably oversampled. Figure 7 represents the regularised solution in some discrete values of the electron energy for λ given by the discrepancy principle; in particular, a spectral slope can be noted at $E \simeq 30$ keV. Using the same method in the case of λ(GCV) one obtains the reconstruction represented in Fig. 8, where a further spectral slope is visible at $E \simeq 80$ kev (this second slope is not visible in Fig. 7, as the discrepancy principle probably provides an oversmoothed reconstruction of the electron energy distribution function).

6 Conclusions

The problem of the inversion of Eq. (2) has been already carried out by Johns & Lin (1992) and by Thompson et al. (1992). In both papers a discretisation of the integral equation has been performed. Johns and Lin achieve a regularisation by reducing the resolution of the data by widening the energy interval corresponding to each sampled value of the solution, so that no regularisation method is used in the inversion. A regularisation approach is considered

in the Thompson et al. paper, where the ill-posed nature of the problem is taken into account. In the present work, thanks to the closed expression of the Gram matrix we have obtained, we did not discretise the integral equation. Moreover, it has been possible to compute the singular system of the problem both in the case of point data (Eq. (12)) and in the case of integrated data (Eq. (48)); so we have shown that the results for the two cases are not significatively different. The knowledge of the singular system has provided information about the numerical instability and the resolution achievable in the presence of a given noise. Finally, Tikhonov's algorithm has been used to perform regularised reconstructions of the solution both from simulated and real data.

Acknowledgement

It is a pleasure to thank prof. J. C. Brown for interesting discussions and suggestions and for forwarding the experimental data provided to him by courtesy of R.P. Lin of U.C. Berkeley.

This work has been partly supported by NATO Grant No. 930384.

References

Bai T., Ramaty R., 1978, ApJ 219, 705
Bertero M., Boccacci P., Pike E.R., 1984, Proc. R. Soc. Lond. A 393, 51
Bertero M., De Mol C., Pike E.R., 1985, Inverse Problems 1, 301
Bertero M., De Mol C., Pike E.R., 1987, Inverse problems 4, 573
Bertero M., Pike E.R., 1991, Inverse Problems 7, 1
Brown J.C., 1971, Solar Phys. 18, 489
Brown J.C., Emslie A.G., 1988, ApJ 331, 554
Computer Aided Mathematics Group of Symbolics, 1988, Macsyma Reference Manual. Symbolics

Courant R., Hilbert D., 1989, Methods of Mathematical Physics. Interscience Publishers, New York, p. 158
Craig I.J.D., Brown J.C., 1986, Inverse Problems in Astronomy. Adam Hilger, Bristol, p. 35
Craven P., Wahba G., 1979, Num. Math. 31, 377
Dennis B.R., 1985, Solar Phys. 100, 465
Golub G.H., Heath M., Wahba G., 1979, Technometrics 21, 215
Johns C.M., Lin R.P., 1992, Solar Phys. 137, 121
Koch H.W., Motz J.W., 1959, Rev. Mod. Phys. 31, 920
Korchak A.A., 1971, Solar Phys. 18, 284
Lin R.P., 1974, Space Sci. Rev. 16, 189
Lin R.P., Schwartz R.A., 1987, ApJ 312, 462
Miller K., 1970, SIAM J. Math. Anal. 1
Riesz F.,Nagy B., 1953, Lecons d'analyse functionelle. Akademiai Kiado, Budapest, p. 143
Thompson A.M., Brown J.C., Craig I.J.D., Fulber C., 1992, A&A 265, 278
Tikhonov A.N., 1963, Sov. Math. Dokl. 4, 1035

Figure captions

Figure 1. Plot of the first six singular functions $u_k(E)$ for $N = 40$ sampled points in the photon energy range $\epsilon_1 = 10$ keV, $\epsilon_N = 100$ keV; (a) u_1 (solid line) and u_2 (dashed line); (b) u_3 (solid line) and u_4 (dashed line); (c) u_5 (solid line) and u_6 (dashed line).
Figure 2. Reconstruction of function (43) with $\alpha = 3$ and $A = 1$, in the case of $N = 40$ geometrically sampled points and $\epsilon_1 = 10$ keV, $\epsilon_N = 100$ keV. (a) Simulated data vector with $\eta = 0.02$; (b) true solution and regularised solution (dashed line) with $\lambda = 8\text{x}10^{-6}$.
Figure 3. Reconstruction of function (44),(45) with $A = 1$, $E_1 = 3$ kev, $\delta_1 = 2$, $\delta_2 = 4$, in the case of $N = 40$ geometrically sampled points and $\epsilon_1 = 1$ keV, $\epsilon_N = 10$ keV. (a) Simulated data vector with $\eta = 0.01$; (b) true solution (solid line) and regularised solution with $\lambda = 1.1\text{x}10^{-3}$.
Figure 4. Reconstruction of function (46) with $E_1 = 40$ keV, $\delta = 3$, $A = 1$, in the case of $N = 40$ geometrically sampled points and $\epsilon_1 = 1$ keV, $\epsilon_N = 100$ keV. (a) Simulated data vector with $\eta = 0.01$ for $\epsilon < 10$ keV and $\eta = 0.02$ otherwise; (b) true solution (solid line) and regularised solution with $\lambda = 2\text{x}10^{-4}$.
Figure 5. (a) Real data vector of $N = 37$ points in the energy range $\epsilon_1 = 14.09$ keV, $\epsilon_N = 286.99$ keV, from the HIREX instrument for the June 27 solar flare; this data corresponds to a time interval around 16:16 UT. (b) Regularised solutions corresponding to the real data vector; the values of the regularisation parameter are chosen according to GCV (solid line) and discrepancy principle (dashed line).
Figure 6. 'Strip' of regularised solutions which puts in evidence the confidence limit of the reconstruction, that is the error propagation from the data to the regularised solution:(a) λ chosen according to GCV;(b) λ chosen according to the discrepancy principle.
Figure 7. Regularised solution in correspondence with discrete values of the electron energy (λ given by discrepancy principle).
Figure 8. Regularised solution in correspondence with discrete values of the electron energy (λ given by GCV).