TOWARDS AN OPTIMAL INVERSION METHOD FOR REMOTE ATMOSPHERIC SENSING

Jean I. F. King *
NAS/NRC Senior Resident Research Fellow
Air Force Cambridge Research Laboratories
Optical Physics Laboratory (CROI)
L. G. Hanscom Field
Bedford, Massachusetts 01730

ABSTRACT

The inference of atmospheric structure from satellite radiometric observations requires an inversion algorithm. The ideal inversion technique should be accurate, self-limiting, free from bias, stable against noise, felxible, and simple in application. A variety of techniques has been spawned to meet these demands. One class, the nonlinear inversion methods, copes with the serious problem of data noise in an unusual fashion. Unlike linear techniques which require *a priori* data smoothing, the nonlinear method can be applied directly to raw data. The algorithm discriminates the noise input by resolving the inferences into two types of solution, associating the real roots with atmospheric structure while ascribing the imaginary roots to noise.

The algorithm appears capable of further refinement with the possibility of inferring systematic noise structure, i.e. pinpointing a channel consistently in error. An example is given of this error-sensing ability of the nonlinear inversion technique.

* On leave from GCA Corporation, Bedford, Massachusetts 01730

1. INTRODUCTION

The advent of meteorological satellites has led to intensive investigation of the possibilities and limitations inherent in remote atmospheric sensing. The upwelling radiation intercepted by the orbiting spectrometer is an integral transform of the physical state of the atmosphere, symbolically expressible as

$$I[\varkappa(\nu)] = \Omega\left\{B[T(u)]\right\}, \tag{1}$$

with $\varkappa$ the monochromatic absorption coefficient at frequency ν and B the Planck intensity, an implicit function of the vertical absorber distribution u through the temperature. For nadir frequency scanning in the far infrared, the integral operator Ω follows from radiative transfer theory as a Laplace transform

$$I(0, 1/\varkappa) = \int_0^\infty B(u)e^{-\varkappa u} \varkappa du = \varkappa L[B(u)]. \tag{2}$$

Clearly the deduction of vertical thermal structure requires the solution of the inverse problem

$$B(u) = L^{-1}\left[\frac{I(0, 1/\varkappa)}{\varkappa}\right]. \tag{3}$$

It might appear at first sight as if we were making a problem where none existed. Certainly the inverse Laplace transform is known for a variety of functions and extensive tabulations are readily accessible. Two features, both related to the observation of real data, intrude into the pure functional space of the mathematician. These are the unavoidable noise residing in the data and the finite number of sensing channels. The fact of noise is particularly pernicious, imposing a sharp upper limit to the information deduction. Inversion, akin to differentiation, tends to amplify data error. The finite channel number forces one to infer a continuous vertical structure from a limited sampling of the radiation field. Thus the goal of inversion cannot be to infer <u>all</u> the temperature structure. This is inaccessible <u>in principle</u> because of the presence of noise. Our aim must be the more modest one of seeking the optimum inversion technique which yields all the valid inferences inherent in the observations. A more subtle corollary to this is that the inversion algorithm must be capable of discriminating noise from the data, that is a proper attribution of the signals either to radiating atmospheric sources or to extraneous noise.

Let us set forth the requirements of an optimum, ideal inversion procedure

1) Accuracy: The inversion should reconstruct all the temperature structure inherent in the data.

2) Self-limiting: The inversion method should not infer more structure than that permitted by the noise level.

3) Freedom from bias: The inference should not be weighted towards any predetermined structure, such as a climatological set. Neither should the functional representation chosen force an unnatural configuration on the thermal profile.

4) Noise stability: The algorithm should be capable of discriminating noise from signals arising from emitting atmospheric sources.

5) Flexibility: The technique should be applicable to any meteorological situation and to any choice of channel sensing input.

6) Simplicity: The inversion procedure should be compatible for computer programming and real time temperature data readout.

It is a pleasure to report that a nonlinear inversion method has been developed which gives promise not only of fulfilling these stringent demands, but appears capable of discriminating between systematic and random errors. That is to say the inversion algorithm will go beyond indicating the presence of noise, by pinpointing which channel is in error and by how much! To indicate how this can be done, let us review the history of the inversion problem in sufficient depth to place this nonlinear Fourier inversion method in perspective.

2. INVERSION THEORY AND CRITIQUE

Most inversion methods are basically linear techniques modified in various ways to cope with the noise problem. They proceed by expanding the Planck intensity in a suitably chosen orthogonal polynomial set

$$B(u) = \sum_{j=1}^{n} B_j P_j(u). \tag{4}$$

Substituting this expression into Eq. (2) leads to a linear simultaneous set of equations which must be satisfied by the n measured intensities

$$I(0,1/\varkappa_i) = \sum_{j=1}^{n} B_j p_j(\varkappa_i) + \epsilon(\varkappa_i) \quad i = 1, 2, \ldots, n, \tag{5}$$

where $p_j(\varkappa) = \varkappa L[P_j(u)$ and ϵ is the noise vector. Experience has shown that the best choice for representing the source is a set of empirical orthogonal functions based on climatological data. The solution consists of a matrix inversion to determine the n weights of the prechosen orthogonal members, with the data appropriately smoothed by the subtraction of the error vector ϵ.

Nonlinear inversion is a completely different approach to the problem. Rather than specifying in advance the functional representation, the weights <u>and members</u> of the set are inferred directly from the intensity data. The data points are used to generate a unique characteristic equation whose eigenfunction solutions form the members of the set.

In the first application to nonlinear inversion, the Planck intensity was approximated by spline functions, i.e., a series of slabs (step functions) or ramps. Since the nonlinear inversion method using spline functions is documented elsewhere (King 1964), we shall merely epitomize its merits and demerits relative to linear methods before proceeding with the new approach.

Three advantages come to mind. First is the fact that the intensity data determine directly the choice of members of the set. For spline functions this corresponds to the slab thicknesses or the distance between successive ramps. The second feature, the uniqueness of the inferred profile, follows as a corollary of the first. The thicknesses are given by the algorithm as the n roots, necessarily unique, of an nth degree polynomial. Stated explicitly, for any suitably chosen 2n intensities, there is one and only one array of n slabs which will fit the data. The third feature is the most subtle, most unexpected, and most important. This is the response of the algorithm to noise in the data. In this event one or more of the roots become negative. These inadmissible roots characteristically have small weights associated with them. The remaining valid roots are relatively unperturbed and preserve a high fidelity representation of the temperature profile. Thus the algorithm acts as a filter, discriminating between the rapidly varying noise components and the lower frequencies associated with temperature structure.

Balancing somewhat these three advantages of member choice, uniqueness, and noise discrimination vis-a-vis the linear methods are four restrictions associated with the nonlinear spline function inversion. First is the stipulation that the Planck intensity be approximated by a slab or ramp solution. For example, a constant tropospheric lapse-rate cannot be satisfactorily fit by a single ramp configuration. More serious, perhaps is the algorithm requirement that the channels be chosen at consecutive integral multiples of the absorption coefficient of the most transparent channel. The channel positions in practice are chosen out of engineering considerations, and it is highly unlikely that a choice on that basis would be optimal for the algorithm. A third condition is the need for an independent determination of the temperature at the top of the atmosphere [B(0)] in the slab algorithm. For the ramp solution, inputs for both B(0) and B'(0) are necessary. Finally, the Prony algorithm is applicable only for transmittances of an exponential function type.

Certainly the most interesting and potentially the most useful feature of the nonlinear approach lies in its treatment of noise in the data. A natural question is whether the nonlinear technique has an underlying physical basis in transfer theory or if the inadmissible roots are mere artifacts of a solution algorithm. Accordingly, an effort was made to establish the constraints necessary to yield slab or ramp solutions to the transfer equation. The search proved fruitless. Indeed it appeared that solutions of the spline function type were incompatible with classical transfer theory.

The pursuit of this logic has led to the formulation of a new wave theory of radiative transfer which contrasts to the corpuscular approach of classical transfer theory. Although the preliminary outlines are clear, at this date (July 1968) the theory is not yet complete. The paper, "Remote Sensing and Inversion Techniques: State of Art, Kine (1967), indicates the progress and general direction of this research.

Our prime concern here is not transfer theory itself, but rather the insight it provides for inversion. In this vein it is a pleasure to report that the wave transfer theory has as a direct consequence a new inversion method based on nonlinear Fourier analysis. Let us back up a bit to see the problem and how the new inversion technique fulfills the need.

In the slab formulation the upwelling intensity is approximated by a set of exponentially weighted step functions of height B_j and thickness $u_{j+1} - u_j$.

$$I(0,1/\varkappa) - B(0) = \int_0^\infty e^{-\varkappa u} \, dB(u)$$

$$= \Sigma B_j e^{-\varkappa u_j} = \Sigma B_j x_j^{\varkappa} , \qquad (6)$$

where we have substituted $u_j = -\ln x_j$. By specifying integral values of we are led to a nonlinear simultaneous equation set which is mathematically equivalent to the moment problem in physics. This set

$$I(0,1/\varkappa_i) - B(0) = \alpha_i = \Sigma B_j x_j^{\varkappa_i} \quad \varkappa_i = 0, 1, \ldots, 2n-1 \qquad (7)$$

possesses a unique solution of n slabs which can be obtained using the Prony algorithm.

We are led to ask the three questions

1) Are there other nonlinear sets soluble by a similar algorithm?

2) Are these sets more flexible, i.e., less rigid than the slab or ramp configurations?

3) Can the intensity sampling points be arbitrarily chosen, free from the Prony requirement of equally spaced intervals?

The answer to these questions is affirmative, leading to the hope that we are approaching the goal described earlier of an optimum inversion method.

3. NONLINEAR FOURIER INVERSION

3.1 Application to Noise-Free Data

We begin by noting that the most general solution of the transfer equation involves waves, i.e., exponential functions of imaginary argument, as well as the attenuating exponentials of classical theory

$$I(u,1/\varkappa) = c \left[\sum_{j=1}^{\infty} \left(\frac{b_j e^{i\omega_j u}}{\varkappa - i\omega_j} + \frac{b_j e^{-i\omega_j u}}{\varkappa + i\omega_j} \right) + \frac{1}{\varkappa} + u + Q \right] . \qquad (8)$$

By specifying odd parity for the Planck intensity B(u), the upwelling intensity $I(0,1/\varkappa)$ and its inverse Laplace transform B(u) become

$$I(0,1/\varkappa) = c\left[\sum_{j=1}^{\infty} \frac{\varkappa b_j \omega_j}{\varkappa^2 + \omega_j^2} + \frac{1}{\varkappa}\right],$$

$$B(u) = L^{-1}\left[\frac{I(0,\ 1/\varkappa)}{\varkappa}\right] = c\left[\sum_{j=1}^{\infty} b_j \sin \omega_j u + u\right]. \tag{9}$$

The form of these equations suggests that by fitting the intensity at 2n channels we may obtain the unique set of n Fourier sine terms of amplitude b_j and frequency ω_j prescribed by the observations.

Let us check the method with a quadratic case, seeing to what extent we can reconstruct the familiar profile

$$B(u) = 1 - \exp(-u) \tag{10}$$

on the basis of the four intensity observations

$$I(0,1/\varkappa_i) - B(0) = \alpha_i = \int_0^\infty e^{-u} dB(u) = \frac{1}{\varkappa_i + 1} \qquad \varkappa_i = 1,\ 2,\ 3,\ 4. \tag{11}$$

Subsiitution of these values into Eq. (9) yields the following four relations which must be satisfied

$$\frac{1}{2} = \frac{b_1\omega_1}{1+\omega_1^2} + \frac{b_2\omega_2}{1+\omega_2^2}$$

$$\frac{1}{3} = \frac{2b_1\omega_1}{4+\omega_1^2} + \frac{2b_2\omega_2}{4+\omega_2^2}$$

$$\frac{1}{4} = \frac{3b_1\omega_1}{9+\omega_1^2} + \frac{3b_2\omega_2}{9+\omega_2^2} \qquad (12)$$

$$\frac{1}{5} = \frac{4b_1\omega_1}{16+\omega_1^2} + \frac{4b_2\omega_2}{16+\omega_2^2}$$

After clearing of fractions and eliminating b_1 and b_2 from the equations, we obtain the following two equations which must be satisfied

$$c_1 - c_2\,(\omega_1^2+\omega_2^2) + c_3\omega_1^2\omega_2^2 = 0$$

$$c_4 - c_5\,(\omega_1^2+\omega_2^2) + c_6\omega_1^2\omega_2^2 = 0 \qquad (13)$$

with

$$c_1 = \frac{27\alpha_3}{5} - \frac{64\alpha_2}{5} + \frac{\alpha_1}{3} = \frac{17}{180} \qquad (14)$$

$$c_2 = \frac{3\alpha_3}{5} + \frac{16\alpha_2}{15} - \frac{\alpha_1}{3} = \frac{7}{180}$$

$$c_3 = \frac{\alpha_3}{15} - \frac{4\alpha_2}{15} + \frac{\alpha_1}{3} = \frac{17}{180}$$

$$c_4 = \frac{64\alpha_4}{7} - \frac{324\alpha_3}{35} + \frac{8\alpha_2}{5} = \frac{1}{21}$$

$$c_5 = \frac{-4\alpha_4}{7} + \frac{36\alpha_3}{35} \;\frac{-2\alpha_2}{5} = \frac{1}{105}$$

$$c_6 = \frac{\alpha_4}{28} - \frac{4\alpha_3}{35} + \frac{\alpha_2}{10} = \frac{1}{84}$$

The characteristic equation must be a binomial with roots at ω_1^2 and ω_2^2

$$(x - \omega_1^2)(x - \omega_2^2) = x^2 - (\omega_1^2 + \omega_2^2)\, x + \omega_1^2 \omega_2^2 = 0. \qquad (15)$$

This equation is compatilbe with the Eqs. (13) if and only if the following determinant vanishes

$$\begin{vmatrix} x^2 & x & 1 \\ c_1 & c_2 & c_3 \\ c_4 & c_5 & c_6 \end{vmatrix} = 0, \qquad (16)$$

yielding as the characteristic equation

$$x^2 - \frac{c_1c_6 - c_3c_4}{c_2c_6 - c_3c_5}\, x + \frac{c_1c_5 - c_2c_4}{c_2c_6 - c_3c_5} =$$

$$x^2 - \frac{85}{11}\, x + \frac{120}{55} = 0. \qquad (17)$$

This equation possesses the roots

$$\omega_1^2 = 0.29350082\,, \omega_1 = \pm\, 0.54175716$$

$$\omega_2^2 = 7.43377190\,, \omega_2 = \pm\, 2.72649443. \qquad (18)$$

We determine b_1 and b_2 now by back substitution of these roots into any pair of Eqs. (12)

$$b_1 = 1.10607838, \quad b_2 = 0.11364969 \qquad (19)$$

Thus we have inferred from the four intensity measurements the following two Fourier sine terms

$$\begin{aligned} B(u) &= b_1 \sin \omega_1 u + b_2 \sin \omega_2 u \\ &= 1.10607838 \sin 0.54175716u \\ &+ 0.11364969 \sin 2.72649443u. \end{aligned} \tag{20}$$

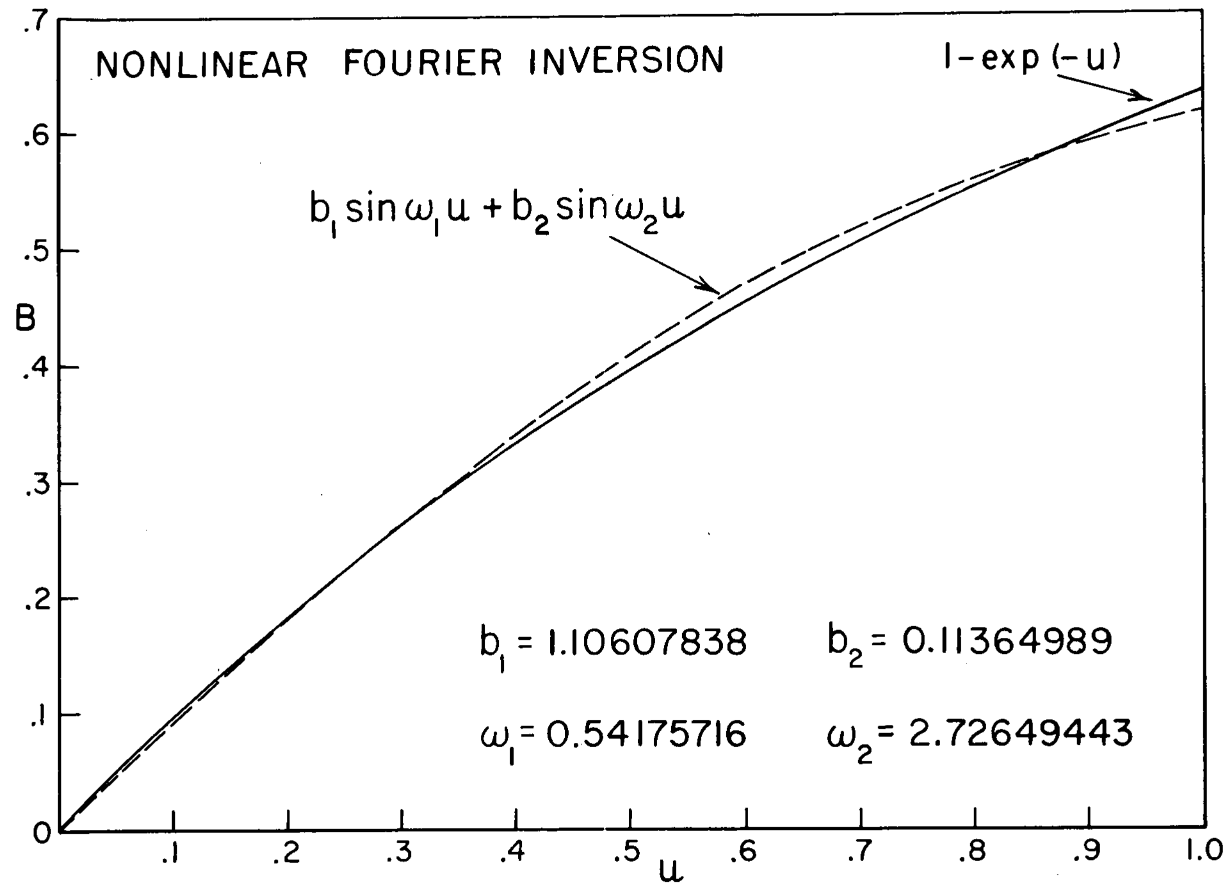

Figure 1. Nonlinear Fourier Inversion (inferred values versus actual profile.)

Figure 1 displays the inferred compared to the actual profile. The agreement is highly gratifying. Particularly encouraging is the fact that the lower frequency component has some eleven times the weight of the higher term, an indication that the method converges rapidly. Although the intensities were chosen at integral values of the absorption coefficient, this need not have been the case, since the inversion algorithm is independent of channel choice. Separation is desirable, however, to avoid near singular matrices.

Although the algebra becomes tedious as more Fourier terms are inferred it should be emphasized that once the channel sites are fixed, the routine leading to the coefficient of the characteristic equation can be fixed once and for all. A program for direct time computer readout is certainly accessible.

The algorithm will surely respond to noise in the data by having one or more of the roots of the characteristic equation becoming negative. In fact one can expect that the algorithm not only will discriminate against noise but infer which one (or more) of the channels is in error. This could be considered the ultimate in inversion, a method which would not only infer all the valid information implicit in the observations but also which channels are in error and by how much.

3.2 Algorithm Response to Noise

Let us examine now the reaction of the nonlinear Fourier inversion to data noise. As a test case we have retained the intensities in three of the directions associated with the sample profile [Eq. (10)], perturbing the third channel alone. The intensities now read

$$\begin{aligned} \alpha_1 &= 1/2 \\ \alpha_2 &= 1/3 \\ \alpha_3 &= 1/3 \\ \alpha_4 &= 1/5. \end{aligned} \tag{21}$$

The straightforward application of the algorithm as before yields the perturbed characteristic equation [cf. Eq.(17)]

$$x^2 + \frac{2785}{361} x - \frac{1656}{361} = 0, \tag{22}$$

with the roots, one positive and one negative,

$$\begin{aligned} \omega_1^{\ 2} &= +0.55472626 \\ \omega_2^{\ 2} &= -8.26940770. \end{aligned} \tag{23}$$

We determine as well by back substitution the weights

$$\begin{aligned} b_1\omega_1 &= 0.78193720 \\ b_2\omega_2 &= 0.02138690. \end{aligned} \tag{24}$$

Our algorithm has therefore inferred from the perturbed intensity measurements, the following interpolation formula for the intensity

$$
\begin{aligned}
I(0,1/\varkappa) &= \sum_{j=1}^{2} \frac{\varkappa b_j \omega_j}{\varkappa^2 + \omega_j^2} \\
&= \frac{0.78193720\varkappa}{\varkappa^2 + 0.55472626} + \frac{0.02138690\varkappa}{\varkappa^2 - 8.26940770},
\end{aligned} \tag{25}
$$

and the following temperature profile

$$
\begin{aligned}
B(u) &= \sum_{j=1}^{2} b_j \sin \omega_j u \\
&= 1.04986 \sin (0.7448u) + 0.007437i \sin (2.8757iu).
\end{aligned} \tag{26}
$$

These results deserve the closest attention and interpretation. We see from the intensity formula [Eq. (25)] that the algorithm has inferred a pole at $1/\varkappa = 1/\omega_2 = 0.34774$ to fit the erroneous measurement at $1/\varkappa = 1/3$. The contribution of the negative root at $1/\varkappa = 1/3$ is

$$
I_2(0,1/3) = \frac{0.02138690/3}{1-(8.26940770/9)} = .08781. \tag{27}
$$

The actual displacement is $(1/3)-(1/4) = (//12) = .08333$, which is within 5% of the true value. Moreover we see that the negative root term affects only slightly the three valid channels. This high fitting specificity follows from the concentration of the amplitude excursion near the pole.

Turning now to the inferred temperature profile, [Eq. (26)] we see that there is only one valid term. The second term is deemed inadmissible because of the imaginary character of the amplitude and frequency. We note, however, that the contribution of the second term is down some two orders of magnitude from the first. The valid term that remains preserves with commendable fidelity the character of the actual profile.

The inadmissible root response of the inversion algorithm to data error is reminiscent of the spline function inversion. Physically this is evidence that there does not exist any Fourier sine pair of arbitrary amplitudes and frequencies which will yield the observed intensities [Eq. (21)]. Hence we must ascribe the negative root to some effect, such as data error, extrinsic to the atmospheric emitting sources.

These results underlie the optimism we feel that the ideal inversion procedure is within grasp and the expectation with which we view its application to real data.

ACKNOWLEDGEMENTS

Research partially supported by NASA/GSFC under contract NAS5-3352

REFERENCES

King, Jean I. F., 1964: Inversion by Slabs of Varying Thickness. J. Atmos. Sci., 21, 324-326.

King, Jean I. F., 1967: Remote Sensing and Inversion Techniques: State of the Art, pp. 253-265 in Proceedings of the Specialist Conference on Molecular Radiation and Its Application to Diagnostic Techniquws, Marshall Space Flight Center, 5-6 October 1967. Document TMX-53711, ed. R. Goulard, NASA/MSFC, 489 pp.

C, 9