

**Atmospheric retrievals with the
Tropospheric Emission Spectrometer
(TES)**

Industrial study group
Institute for Pure and Applied Mathematics

G. Bal, B. Borchers, K. Bowman, C. Estatico, M.
Jäckels, R. Ramlau, K. Ren, A. Tamasan

Los Angeles
2003

Introduction

The Tropospheric Emission Spectrometer (TES) on the EOS-Aura spacecraft will measure the global 3-dimensional distribution of ozone and other gases in the troposphere. It works in two different modes, the so called nadir and the limb mode. In the nadir-mode the measurements are given by a line-integral with respect to the altitude. This is an one-dimensional problem, the goal is the retrieval of the ozone concentration profile. In the limb mode the measurements are based on a two-dimensional integration over the altitude and the aperture angle. Here one wants to retrieve the ozone-distribution in two dimensions. This problem is more challenging than the retrieval problem in nadir mode. In our work, we focused only on the nadir mode.

The problem was illustrated to us by Kevin Bowman, who has worked on it for several years. The goal of the industrial study group was to develop new mathematical approaches and to get some theoretical and numerical results for these inverse problems.

Kevin and the group at JPL have already developed a statistical approach to this problem. Some members of our group reviewed this approach and came up with suggestions of ways in which it might be improved. These suggestions are discussed in Chapter 1. Other members of our group focused on a functional analytic approach, considering the nadir mode and developing results on the existence and uniqueness of solutions to a simplified version of the problem. These results are discussed in Chapter 2. Some members of the group also worked on a numerical method for the inversion. This approach is discussed in Chapter 3. An alternative Newton linearization approach is discussed in Chapter 4. Plans for future work are discussed in Chapter 5.

Chapter 1

Comments on the Existing Approach

Kevin and the group at JPL have developed a statistical approach for solving this inverse problem using nonlinear least squares minimization with ad-hoc regularization on discretized model. In this approach, the model is first discretized as $y = F(x) + n$, where x is a vector of concentrations at different altitudes, y is the resulting spectrum, contaminated by noise n . The solution is obtained by solving

$$\hat{x} = \min \|y - F(x)\|_{S_n^{-1}}^2 + \|x - x_c\|_{\Lambda^{-1}}^2$$

Here S_n represents the covariance matrix for the noise, and Λ and x_c determine the regularization.

In a pure Bayesian approach, if the prior distribution was multivariate normal (MVN), with prior mean x_c and prior covariance Λ , this would give the maximum a posteriori (MAP) solution. By linearizing around this solution, we could obtain an MVN posterior distribution. However, the JPL approach uses ad hoc regularization terms without an explicit prior. With this sort of regularization, the regularized solution is *biased*, with the magnitude and direction of the bias depending on both the true model and the averaging kernel.

After the inversion has been accomplished, it's possible to perform error analysis including the effects of bias introduced by the regularization. This error analysis is critical to atmospheric retrievals and data assimilation. However, this analysis requires us to make assumptions about the true model in much the same way that we would have to select a prior distribution in the Bayesian approach.

The approach that has been implemented has some advantages over the pure Bayesian approach. In the pure Bayesian approach, constructing priors from previous observations and simulation models can be a difficult process. Furthermore, the prior must be selected before we can begin to analyze the data. In the implemented approach, it's quite simple to construct an ad-hoc regularization scheme which stabilizes the inversion. By using the averaging kernel we can effectively postpone assumptions about the prior distribution.

Note that most Bayesians would consider the idea of selecting a prior after the data has been gathered and analyzed to be heretical. The problem here is that you run the risk of deluding yourself that results are good when in fact you over regularized in order to get a believable solution.

There were several suggestions for possible improvements to this basic approach.

1. The approach that has been implemented involves “regularization by discretization” in that the continuous problem is first reduced to the problem of finding the concentrations in 87 intervals, and then further reduced by interpolating down to 10 parameters for the actual minimization. This regularization by discretization could potentially be much stronger than the explicit regularization used. It would be good (and relatively simple) to perform some inversions with the 87 parameter model to determine the strength of the regularization introduced by the interpolation. It would also be worthwhile to further refine the discretization to see how this would effect the results.
2. Since the forward model is nonlinear, the uncertainty analysis is based on a linearization around the best model parameters. It isn’t clear whether this approximation is reasonable, particularly when the averaging kernel is used to estimate the bias for parameters which are far from the parameters used in the linearization. It should be relatively simple to compare the true nonlinear model and the linearized model for parameters of interest to see whether the approximation is reasonably accurate or not.
3. It appears that the inversion process might get stuck in a local minimum under some circumstances. Multi-start or more sophisticated randomized search strategies should be used to investigate both randomly selected cases and cases that have been identified as problematic. This could produce confidence that local minima are seldom a problem, or it could make it clear that some more sophisticated global optimization scheme is needed.
4. Markov Chain Monte Carlo (MCMC) methods might be used to sample from the posterior distribution. This would take care of problems caused by the nonlinearity of the model, and allow for a wide variety of prior distributions. This would also take care of situations in which the cost function is multimodal with significant local minima.
5. The forward model should be carefully checked to determine whether it is sufficiently numerically smooth.
6. The stopping criteria for the Levenberg–Marquardt method might be improved.
7. The group discussed the possibility of presmoothing the data before inversion. Another related idea would be use adaptive wavelet bases in constructing the ad-hoc regularization scheme.

Chapter 2

Uniqueness of Reconstruction in Simplified Setting

2.1 Setting of inverse problem

Consider a radiation source term at the surface $L(0, \nu)$, where ν is frequency, a volume source term of radiation $a(z, \nu)B(z, \nu)$, where z is altitude, $a(z, \nu)$ is the absorption profile in the atmosphere, and B represents black body radiation from the atmosphere at altitude z and frequency ν , which is assumed to follow Planck's law.

The measurements are the radiation intensity $L(Z, \nu)$ at the top of the atmosphere Z for all frequencies ν . The equation that L satisfies is

$\frac{\partial L(z, \nu)}{\partial z} + a(z, \nu)L(z, \nu) = a(z, \nu)B(z, \nu), \quad z \in (0, Z), \quad \nu \in (0, +\infty)$ $L(0, \nu) = L_0(\nu) \quad \text{known}$ $L(Z, \nu) \quad \text{measured .}$	(2.1)
---	-------

The first-order ODE is easily invertible and we get

$$L(Z, \nu) = e^{-\int_0^Z a(z, \nu) dz} L(0, \nu) + \int_0^Z a(z, \nu) B(z, \nu) e^{-\int_z^Z a(\zeta, \nu) d\zeta} dz. \quad (2.2)$$

Let us define the optical length

$$\tilde{\alpha}(z, \nu) = \int_0^z a(\zeta, \nu) d\zeta. \quad (2.3)$$

We recast the above integral as

$$\boxed{L(Z, \nu) = e^{-\tilde{\alpha}(0, \nu)} L(0, \nu) + \int_0^Z a(z, \nu) B(z, \nu) e^{-\tilde{\alpha}(z, \nu)} dz.} \quad (2.4)$$

An interesting inverse problem in atmosphere imaging is to assume that the radiation term $B(z, \mu)$ is known and to reconstruct as much as we can of $a(z, \nu)$ from $L(Z, \nu)$. The absorption map $a(z, \nu)$ depends on concentrations of particles in the atmosphere such as ozone, carbon monoxide/dioxide. These concentrations are what we are interested in. Let us assume that we have found a band of frequencies in which only one particle, such as *ozone*, contributes to absorption. Then we have as an approximate model that

$$\boxed{a(z, \nu) = c(z) \kappa(z, \nu),} \quad (2.5)$$

where $c(z)$ is the unknown concentration profile and $a(z, \nu)$ is a known signature of ozone. It looks like something of the form

$$\kappa_L(\nu, z) = \frac{S(\nu_0) \alpha_L(z)}{\pi(\nu - \nu_0)^2 + \alpha_L(z)^2}, \quad \alpha_L(z) \sim T(z)^{-1/2}, \quad (2.6)$$

where $T(z)$ is temperature and ν_0 is the center of the frequency band of interest.

We simplify the model as follows. We assume that

$$\kappa_L(\nu, z) = \frac{S(\nu_0) \alpha_L(z)}{\pi(\nu - \nu_0)^2} \equiv \mu(\nu) \alpha_L(z).$$

so that both formulas agree when $|\nu - \nu_0| \rightarrow \infty$. As an approximation, we thus obtain that μ varies between 0 and ∞ as ν varies in \mathbb{R} . We recast it as

$$\boxed{\kappa(\nu, z) = \mu(\nu) g(z)} \quad (2.7)$$

for some function g strictly positive. This is the slightly simplified model we analyze in the sequel.

Physically the B term is of the form

$$B(z, \nu) \sim \frac{\nu^3}{e^{-h\nu/kT(z)} - 1} \approx \nu^2 T(z) \quad \text{and} \quad L_0(\nu) \approx \nu^2 L_0, \quad (2.8)$$

where L_0 is a constant. Since both source terms are linear in ν^2 we can change variables $L(z, \nu)/\nu^2 \rightarrow L(z, \mu)$ with $\mu = (\nu - \nu_0)^2$. Let us define the rescaled optical length

$$\alpha(z) = \int_z^Z c(\zeta) g(\zeta) d\zeta. \quad (2.9)$$

Accounting for all the simplifications the inverse problem we wish to solve has the form

$$\boxed{D(\mu) = L_0 e^{-\mu \alpha(0)} + \int_0^Z \mu c(z) g(z) T(z) e^{-\mu \alpha(z)} dz} \quad (2.10)$$

Here $D(\mu)$ is our measurement of $\mu \in \mathbb{R}^+$. Important aspects are that g is a positive (and known) function and that $c(z)$ is a positive unknown concentration profile. We also assume that the temperature profile $T(z)$ is known.

2.2 Uniqueness and ill-posedness in a simple case

Since $\alpha'(z) < 0$ on $(0, Z)$ we can perform the change of variables $z \rightarrow \alpha(z)$, define the inverse map $\alpha \rightarrow z(\alpha)$ and recast the integral equation as

$$\begin{aligned} D(\mu) &= L_0 e^{-\mu\alpha(0)} + \int_0^{\alpha(0)} \mu c(z(\alpha)) g(z(\alpha)) T(z(\alpha)) \left| \frac{dz}{d\alpha} \right| e^{-\mu\alpha} d\alpha \\ &= L_0 e^{-\mu\alpha(0)} + \int_0^{\alpha(0)} \mu T(z(\alpha)) e^{-\mu\alpha} d\alpha. \end{aligned} \quad (2.11)$$

We have used that $\alpha(Z) = 0$. Upon integrating by parts in the above integral, we obtain that

$$D(\mu) = (L_0 - T(0)) e^{-\mu\alpha(0)} + T(Z) + \int_0^{\alpha(0)} T'(z(\alpha)) \frac{dz}{d\alpha} e^{-\mu\alpha} d\alpha. \quad (2.12)$$

The data $D(\mu) - T(Z)$ is thus the *Laplace transform* of the distribution

$$\boxed{T'(z(\alpha)) \frac{dz}{d\alpha} + (L_0 - T(0)) \delta(\alpha - \alpha(0)) \equiv h(\alpha)} \quad (2.13)$$

We can thus reconstruct $\alpha(0)$, L_0 , and $T'(z(\alpha))z'(\alpha)$ on $(0, \alpha(0))$ from the measurements $D(\mu)$. This is known to be a *severely ill-posed* problem.

The above problem is equivalent to

$$\frac{-T'(z)}{c(z)g(z)} = h\left(\int_z^Z c(\zeta)g(\zeta)d\zeta\right). \quad (2.14)$$

We thus obtain the Volterra type integral equation

$$\boxed{c(z) = \frac{-T'(z)}{h\left(\int_z^Z c(\zeta)g(\zeta)d\zeta\right)g(z)}} \quad (2.15)$$

It can be solved on $(0, Z)$ backwards starting from

$$c(Z) = \frac{-T'(Z)}{h(0)g(Z)}. \quad (2.16)$$

This shows uniqueness of the reconstruction and provides a constructive method to obtain $c(z)$ in $(0, Z)$.

2.3 Small inclusions

Solving the full problem for $c(z)$ is unlikely to provides us with a good estimate of the strong and localized variations that may occur in the concentration profile. Rather one may want to impose the presence of such variations and try to reconstruct them directly. Here is a caricature of a model.

Let us assume that $f \equiv 0$ and that a background profile $c_0(z)$ is more or less known. The true profile is

$$\boxed{c(z) = c_0(z) + \delta c(z)}. \quad (2.17)$$

The assumption on $\delta c(z)$ is *not* that it is small in L^∞ but rather that it is small in L^1 and of “small” support.

Let us assume that $\delta c(z)$ takes the value δc on an interval centered at $z = z_0$ and of size δz and takes the value 0 elsewhere:

$$\delta c(z) = \delta c \chi_{\{|z-z_0| \leq \delta z\}}(z).$$

We first recast the radiative transfer equation as an equation for $H(z, \mu) = L(z, \mu) - T(z)$ after rescaling and obtain that

$$H(Z, \mu) = e^{-\mu\alpha(0)} H(0) + \int_0^Z -\frac{\partial T}{\partial z} e^{-\mu\alpha(z)} dz. \quad (2.18)$$

Let us denote by $\alpha_0(z)$ the optical length for c_0 only. We then observe that

$$H[c+\delta c](Z, \mu) = e^{-\mu\alpha(0)} H(0) e^{-\mu\delta c\delta z g(z_0)} + \int_0^Z -\frac{\partial T}{\partial z} e^{-\mu\alpha_0(z)} e^{-\mu\delta c\delta z g(z_0)\chi_{[z, z_0]}(z_0)} dz + O(\delta z^2).$$

Assuming we know the background we thus have access to information of the form

$$\boxed{\mu \mapsto \mu\delta c\delta z g(z_0) \left[e^{-\mu\alpha_0} H(0) + \int_{z_0}^Z -\frac{\partial T}{\partial z} e^{-\mu\alpha_0(z)} dz \right]}.$$

Taking the ratio at two different values of μ gives a functional on z_0 out of which we should be able to reconstruct z_0 (may be with additional assumptions on where z_0 is, or may be using more than two values of μ). Once z_0 is known, we easily get $\delta c\delta z$ from the above expression. So if the background is known with sufficient accuracy we can deduce that

$$\boxed{z_0 \text{ and } \delta c\delta z \text{ are “easily” reconstructed even from limited measurements.}}$$

Notice that if the noise in the data is of the order of $O(\delta z^2)$ so that higher order terms cannot be reconstructed, the product $\delta c\delta z$ (i.e. the total amount of ozone variation in the layer) is all we can get. If the noise in the data is larger than $\delta c\delta z$, then even this information cannot be retrieved unless a more careful statistical model is considered.

Chapter 3

New approaches for numerical inversions in the nadir mode

For fixed frequency ν the monochromatic radiance $L(\nu, z)$ emerging from an atmospheric path can be written as

$$L(\nu, z) = L(\nu, 0) \tau(\nu, 0, z) + \int_0^z J(\xi) \frac{d}{d\xi} \tau(\nu, \xi, z) d\xi . \quad (3.1)$$

In (3.1) $\tau(\nu, \xi, z)$ is the transmittance from ξ to z and can be expressed in terms of the absorption coefficients $\kappa_i(\nu, z)$ of the different absorbers:

$$\tau(\nu, \xi, z) = e^{-\int_{\xi}^z \sum_i \kappa_i(\nu, s) \rho_i(s) ds} . \quad (3.2)$$

For simplicity we consider only the case, that scattering albedo can be neglected and that there's only one absorber, ozone, which distribution $\rho(z)$ we want to recover. Because measurements are always taken at a constant height, which is given by the satellite, we set $z \equiv 1$. Further we introduce the function $B(\nu, z)$, so that we can rewrite (3.1) as

$$L(\nu, 1) = L(\nu, 0) \tau(\nu, 1) + \int_0^1 B(\nu, \xi) \frac{d}{d\xi} \tau(\nu, \xi, 1) d\xi . \quad (3.3)$$

This equation will be the starting point of our considerations.

3.1 Approach 1

From (3.3) we get easily via integration by parts

$$L(\nu, 1) = (L(\nu, 0) - B(\nu, 0)) \tau(\nu, 0, 1) + B(\nu, 1) - \int_0^1 B_\xi(\nu, \xi) \tau(\nu, \xi, 1) d\xi \quad (3.4)$$

with

$$B(\nu, T) = \frac{\nu^3}{e^{\frac{-h\nu}{k_B T}} - 1},$$

$$\tau(\nu, \xi, 1) = e^{-\int_\xi^1 \kappa(\nu, s) \rho(s) ds}.$$

Because we want to retrieve the unknown distribution $\rho(\xi)$, (3.4) seems to be easier to handle since there the derivative of B is taken, which can be calculated analytically.

The idea is to solve the nonlinear equation (3.4) iteratively with an initial guess $\varrho_1(\xi)$ for the unknown distribution $\varrho(\xi)$. In every iteration step we define pseudo-data $g_k(\varrho_k)$ by

$$g_k(\varrho_k) := (L(\nu, 0) - B(\nu, 0)) e^{-\int_0^1 \kappa(\nu, s) \rho_k(s) ds} + B(\nu, 1) - L(\nu, 1) \quad (3.5)$$

and obtain

$$F(\varrho_{k+1}) := \int_0^1 B_\xi(\nu, \xi) e^{-\int_0^1 \kappa(\nu, s) \rho_{k+1}(s) ds} d\xi = g_k(\varrho_k). \quad (3.6)$$

We can solve this nonlinear operator equation by several iteration methods. For convenience we consider the Landweber method

$$\varrho_{k+1} = \varrho_k + \beta F'^* (g_k - F\varrho_k), \quad k \in \mathbb{N}, \quad (3.7)$$

with F'^* the adjoint of the Fréchet derivative F' . Obvious the Fréchet derivative of F in respect to ϱ is given by

$$F'(\varrho) h = \int_0^1 B_\xi(\nu, \xi) \kappa(\nu, \xi) h(\xi) e^{-\int_0^1 \kappa(\nu, s) \rho_{k+1}(s) ds} d\xi, \quad (3.8)$$

so that the calculation of the adjoint is straight forward and we obtain

$$\begin{aligned}
\langle F'(\varrho) h, \tilde{h} \rangle_{L_2[\nu_{min}, \nu_{max}]} &= \int_{\nu_{min}}^{\nu_{max}} \int_0^1 B_\xi(\nu, \xi) \kappa(\nu, \xi) h(\xi) e^{-\int_\xi^1 \kappa(\nu, s) \varrho(s) ds} d\xi \tilde{h}(\nu) d\nu \\
&= \int_0^1 \int_{\nu_{min}}^{\nu_{max}} B_\xi(\nu, \xi) \kappa(\nu, \xi) \tilde{h}(\nu) e^{-\int_\xi^1 \kappa(\nu, s) \varrho(s) ds} d\nu h(\xi) d\xi \\
&= \langle h, F'^*(\varrho) \tilde{h} \rangle_{L_2[0,1]}, \tag{3.9}
\end{aligned}$$

where the adjoint F'^* is defined by

$$F'^*(\varrho) \tilde{h} := \int_{\nu_{min}}^{\nu_{max}} B_\xi(\nu, \xi) \kappa(\nu, \xi) \tilde{h}(\nu) e^{-\int_\xi^1 \kappa(\nu, s) \varrho(s) ds} d\nu. \tag{3.10}$$

3.2 Approach 2

$$L(\nu, 1) = L(\nu, 0) \tau(\nu, 0, 1) + \int_0^1 B(\nu, \xi) \frac{\partial \tau(\nu, \xi, 1)}{\partial \xi} d\xi, \tag{3.11}$$

with

$$\begin{aligned}
B(\nu, \xi) &- \text{Planck's law} \\
\tau(\nu, \xi, 1) &- \text{Transmittance function} \\
\tau(\nu, \xi, 1) &= \exp \left\{ - \int_\xi^1 \sum_{i=1}^n \kappa_i(\nu, s) \varrho_i(s) ds \right\} \\
\varrho_i &- \text{searched for distribution}
\end{aligned}$$

In the following, we might consider only the special case where the spectrum only contains information on O_3 . Thus, the transmittance function can be written as

$$\tau(\nu, \xi, 1) = \exp \left\{ - \int_\xi^1 \kappa(\nu, s) \varrho(s) ds \right\}. \tag{3.12}$$

There are several possibilities for a stable inversion of equation (3.11). As could be seen in the previous section, a difficulty arises from the fact that $\tau(\nu, 0, 1)$ is unknown. We have

$$\int_0^1 \frac{\partial \tau(\nu, \xi, 1)}{\partial \xi} d\xi = \tau(\nu, 1, 1) - \tau(\nu, 0, 1),$$

and with

$$\tau(\nu, 1, 1) = \exp \left\{ - \int_1^1 \kappa(\nu, s) \varrho(s) ds \right\} = 1$$

follows

$$L(\nu, 1) - L(\nu, 0) = \int_0^1 (B(\nu, \xi) - L(\nu, 0)) \frac{\partial \tau(\nu, \xi, 1)}{\partial \xi} d\xi . \quad (3.13)$$

Next, the derivative of $\tau(\nu, \xi, 1)$ with respect to ξ is computed. Setting

$$f(\nu, s) = \kappa(\nu, s) \varrho(s)$$

we get

$$\frac{\partial \tau(\nu, \xi, 1)}{\partial \xi} = f(\nu, \xi) \exp \left\{ - \int_{\xi}^1 f(\nu, s) ds \right\} .$$

Setting

$$F(\varrho)(\nu) := \int_0^1 (B(\nu, \xi) - L(\nu, 0)) \kappa(\nu, \xi) \varrho(\xi) \exp \left\{ - \int_{\xi}^1 \kappa(\nu, s) \varrho(s) ds \right\} d\xi . \quad (3.14)$$

we have to solve the nonlinear operator equation

$$L(\nu, 1) - L(\nu, 0) = F(\varrho)(\nu) . \quad (3.15)$$

We wish to remark that the structure of the operator F is close to the attenuated Radon transform. Almost all regularization methods make use of the Fréchet derivative $F'(\varrho)$ of F and its adjoint $F'(\varrho)^*$. In many cases, the choice of the proper function spaces is crucial for the existence of the derivative. If we assume that the kernel $\kappa(s, \nu)$ is bounded,

$$|\kappa(\nu, s)| \leq C \forall (\nu, s) \in [\nu_{min}, \nu_{max}] \times [0, 1] , \quad (3.16)$$

we can prove that F is Fréchet differentiable. First, we need to prove some preliminary results:

Proposition 3.2.1 *Let K be an integral operator defined by*

$$K\varrho(\nu) = \int_{\xi}^1 \kappa(\nu, s) \varrho(s) ds ,$$

where κ fulfills (3.16). Then

$$\left| \int_{\xi}^1 \kappa(\nu, s) \varrho(s) ds \right| \leq C \|\varrho\| \quad (3.17)$$

Proof:

This is easily seen by Hölder inequality:

$$\begin{aligned} \left| \int_{\xi}^1 \kappa(\nu, s) \varrho(s) ds \right| &\leq \left(\int_{\xi}^1 |\kappa(\nu, s)|^2 ds \right)^{1/2} \left(\int_{\xi}^1 |\varrho(s)|^2 ds \right)^{1/2} \\ &\leq C \|\varrho\| \end{aligned}$$

■

Proposition 3.2.2 *Under the assumptions of Proposition 3.2.1, the estimate*

$$\begin{aligned} \exp \left\{ - \int_{\xi}^1 \kappa(\nu, s) (\varrho(s) + h(s)) ds \right\} &= \exp \left\{ - \int_{\xi}^1 \kappa(\nu, s) (\varrho(s)) ds \right\} \\ &\quad - \int_{\xi}^1 \kappa(\nu, s) (\varrho(s)) ds \cdot \exp \left\{ - \int_{\xi}^1 \kappa(\nu, s) (\varrho(s)) ds \right\} + O(\|h\|^2) \end{aligned} \quad (3.18)$$

holds for $\|h\| \rightarrow 0$.

Proof:

By Taylor's expansion, we have

$$e^{-(x+t)} = e^{-x} - te^{-x} + \frac{t^2}{2} \int_0^1 e^{-(x+\theta t)} d\theta \quad (3.19)$$

Setting $x = \int_0^1 \kappa(\nu, s) \varrho(s) ds$ and $t = \int_0^1 \kappa(\nu, s) h(s) ds$. Using (3.17) we can estimate the last term in (3.19) by

$$\begin{aligned} \left| \frac{t^2}{2} \int_0^1 e^{-(x+\theta t)} d\theta \right| &\leq \frac{C}{2} \|h\|^2 \int_0^1 e^{C1(\|\varrho\| + \theta \|h\|)} d\theta \\ &\leq \frac{C}{2} \|h\|^2 e^{C(2\|\varrho\|)} = O(\|h\|^2) \end{aligned}$$

for $\|h\| \leq \|\varrho\|$. Thus, (3.18) holds for $\|h\| \rightarrow 0$.

■

Theorem 3.2.3 *The operator $F : L_2(0, 1) \rightarrow L_2(\nu_{min}, \nu_{max})$ defined in (3.14) is Fréchet-differentiable. Its derivative is given by*

$$F'(\varrho)h(\nu) = \int_0^1 \kappa(\nu, \xi) \left(h(\xi) - \varrho(\xi) \int_{\xi}^1 \kappa(\nu, s) h(s) ds \right) e^{-\int_{\xi}^1 \kappa(\nu, s) \varrho(s) ds} d\xi \quad (3.20)$$

Proof:

We have

$$\begin{aligned} F(\varrho + h)(\nu) &= \int_0^1 \kappa(\nu, 1) (h(\xi) + \varrho(\xi)) e^{\int_{\xi}^1 \kappa(\nu, s) (\varrho(s) + h(s)) ds} \\ &\stackrel{(3.18)}{=} \int_0^1 \kappa(\nu, \xi) (\varrho(\xi) + h(\xi)) \left(e^{\int_{\xi}^1 \kappa(\nu, s) \varrho(s) ds} - \int_{\xi}^1 \kappa(\nu, s) h(s) ds \cdot e^{-\int_{\xi}^1 \kappa(\nu, s) \varrho(s) ds} + O(\|h\|^2) \right) d\xi \\ &= F(\varrho)(\nu) + \underbrace{\int_0^1 \kappa(\nu, \xi) h(\xi) e^{\int_{\xi}^1 \kappa(\nu, s) \varrho(s) ds} d\xi - \int_0^1 \kappa(\nu, \xi) \varrho(\xi) \int_{\xi}^1 \kappa(\nu, s) h(s) ds \cdot e^{\int_{\xi}^1 \kappa(\nu, s) \varrho(s) ds} d\xi}_{=: F'(\varrho)h(\nu)} \\ &\quad + \int_0^1 \kappa(\nu, \xi) h(\xi) \int_{\xi}^1 \kappa(\nu, s) h(s) ds \cdot e^{\int_{\xi}^1 \kappa(\nu, s) \varrho(s) ds} d\xi + O(\|h\|^2) \int_0^1 \kappa(\nu, \xi) (\varrho(\xi) + h(\xi)) d\xi . \end{aligned}$$

It is easily seen that the last two terms behave like $O(\|h\|^2)$, and we have finally shown

$$F(\varrho + h) = F(\varrho) + F'(\varrho)h + O(\|h\|^2) ,$$

which proves the Fréchet-differentiability of F . ■

Theorem 3.2.4 *The adjoint operator $F'(\varrho)^*$ of $F'(\varrho)$ (in the L_2 sense) is given by*

$$F'(\varrho)^*g(\xi) = \int_{\nu_{min}}^{\nu_{max}} \kappa(\nu, \xi) e^{-\int_{\xi}^1 \kappa(\nu, s) \varrho(s) ds} g(\nu) d\nu - \int_{\nu_{min}}^{\nu_{max}} \int_0^{\xi} \kappa(\nu, s) \varrho(s) \cdot e^{-\int_{\xi}^1 \kappa(\nu, s) \varrho(s) ds} ds g(\nu) d\nu \quad (3.21)$$

Proof:

To simplify the notation we set $F'(\varrho) = F'_1(\varrho) - F'_2(\varrho)$ with

$$\begin{aligned} F'_1(\varrho)h(\nu) &= \int_0^1 \kappa(\nu, \xi)h(\xi) \cdot e^{-\int_{\xi}^1 \kappa(\nu, s)\varrho(s) ds} d\xi \\ F'_2(\varrho)h(\nu) &= \int_0^1 \kappa(\nu, \xi)\varrho(\xi) \int_{\xi}^1 \kappa(\nu, s)h(s) ds \cdot e^{-\int_{\xi}^1 \kappa(\nu, s)\varrho(s) ds} d\xi . \end{aligned}$$

We obtain

$$\begin{aligned} \langle F'_1(\varrho)h, g \rangle &= \int_{\nu_{min}}^{\nu_{max}} \int_0^1 \kappa(\nu, \xi)h(\xi) e^{-\int_{\xi}^1 \kappa(\nu, s)\varrho(s) ds} d\xi g(\nu) d\nu \\ &= \int_0^1 h(\xi) \int_{\nu_{min}}^{\nu_{max}} \kappa(\nu, \xi) e^{-\int_{\xi}^1 \kappa(\nu, s)\varrho(s) ds} g(\nu) d\nu d\xi , \end{aligned}$$

and thus

$$F'_1(\varrho)^* g(\xi) = \int_{\nu_{min}}^{\nu_{max}} \kappa(\nu, \xi) e^{-\int_{\xi}^1 \kappa(\nu, s)\varrho(s) ds} g(\nu) d\nu . \quad (3.22)$$

The adjoint of F'_2 is computed by

$$\begin{aligned} \langle F'_2(\varrho)h, g \rangle &= \int_{\nu_{min}}^{\nu_{max}} \int_0^1 \kappa(\nu, \xi)\varrho(\xi) \int_{\xi}^1 \kappa(\nu, s)\varrho(s) ds \cdot e^{-\int_{\xi}^1 \kappa(\nu, s)\varrho(s) ds} g(\nu) d\nu \\ &= \int_{\nu_{min}}^{\nu_{max}} \int_0^1 \kappa(\nu, \xi)\varrho(\xi) \int_0^1 \chi_{[\xi, 1]}(s)\kappa(\nu, s)h(s) ds \cdot e^{-\int_{\xi}^1 \kappa(\nu, s)\varrho(s) ds} d\xi g(\nu) d\nu \\ &= \int_0^Z h(s) \int_{\nu_{min}}^{\nu_{max}} \int_0^1 \kappa(\nu, \xi)\varrho(\xi)\chi_{[\xi, 1]}(s) \kappa(\nu, s) \cdot e^{-\int_{\xi}^1 \kappa(\nu, s')\varrho(s') ds'} d\xi g(\nu) d\nu ds . \end{aligned}$$

Now, we have

$$\begin{aligned} \chi_{[\xi, 1]}(s) &= \begin{cases} 1 & \text{for } \xi \leq s \leq 1 \\ 0 & \text{otherwise} \end{cases} \\ &= \chi_{[0, s]}(\xi) , \end{aligned}$$

and thus

$$F'_2(\varrho)^* g(s) = \int_{\nu_{min}}^{\nu_{max}} \int_0^s \kappa(\nu, \xi)\varrho(\xi)\kappa(\nu, s) \cdot e^{-\int_{\xi}^1 \kappa(\nu, s')\varrho(s') ds'} dz' g(\nu) d\nu . \quad (3.23)$$

The assumption follows from $F'(\varrho)^* = F'_1(\varrho)^* - F'_2(\varrho)^*$. ■

Chapter 4

A Newton linearization approach for solving the atmospheric retrieval problem

We want to solve for $a \in L^2[0, 1]$, the following equation

$$G(a) = f$$

where

- $f \in L^2[\nu_{min}, \nu_{max}]$ is the known data, and
- $G : L^2[0, 1] \longrightarrow L^2[\nu_{min}, \nu_{max}]$ is the forward operator defined by

$$G(a)(\nu) = E(a)(0, \nu) + \int_0^1 E(a)(s, \nu) ds$$

with

$$E(a)(s, \nu) = \exp \left\{ - \int_s^1 a(t)k(t, \nu) dt \right\},$$

for $k : [0, 1] \times [\nu_{min}, \nu_{max}] \longrightarrow \mathbb{R}$ a known function.

4.1 Linearization by Newton method

Let ϵ estimate data error and let $G'(a) : L^2[0, 1] \longrightarrow L^2[\nu_{min}, \nu_{max}]$ denote the Fréchet derivative of the operator G at the point a . Let $j = 0$.

Start with initial guess $a_0(t)$ and solve for $h \in L^2[0, 1]$ the linearized problem

$$G'(a_j)h = f - G(a_j),$$

then update $a_{j+1} = a_j + h$ and repeat until $\|f - G(a_j)\| \leq \epsilon$.

4.2 The Fréchet derivative for the linearization

By considering the Taylor's expansion $e^{-(l+m)} = e^{-l}(1 - m) + O(m^2)$, with $l = \int_s^1 a(t)k(t, \nu) dt$ and $m = \int_s^1 h(t)k(t, \nu) dt$, we can write

$$E(a+h)(s, \nu) = E(a)(s, \nu) \left(1 - \int_s^1 h(t)k(t, \nu) dt\right) + O(\|h\|^2).$$

We have

$$\begin{aligned} G(a+h)(\nu) &= E(a+h)(0, \nu) + \int_0^1 E(a+h)(s, \nu) ds \\ &= E(a+h)(0, \nu) + \int_0^1 \left\{ E(a)(s, \nu) \left(1 - \int_s^1 h(t)k(t, \nu) dt\right) + O(\|h\|^2) \right\} ds \\ &= E(a+h)(0, \nu) + \int_0^1 E(a)(s, \nu) ds - \int_0^1 E(a)(s, \nu) \int_s^1 h(t)k(t, \nu) dt ds + O(\|h\|^2) \\ &= E(a)(0, \nu) - E(a)(0, \nu) \int_0^1 h(t)k(t, \nu) dt \\ &\quad + \int_0^1 E(a)(s, \nu) ds - \int_0^1 E(a)(s, \nu) \int_s^1 h(t)k(t, \nu) dt ds + O(\|h\|^2) \\ &= E(a)(0, \nu) + \int_0^1 E(a)(s, \nu) ds \\ &\quad - E(a)(0, \nu) \int_0^1 h(t)k(t, \nu) dt - \int_0^1 E(a)(s, \nu) \int_s^1 h(t)k(t, \nu) dt ds + O(\|h\|^2) \\ &= G(a)(\nu) + [G'(a)h](\nu) + O(\|h\|^2). \end{aligned}$$

Hence, the Fréchet derivative $G'(a) : L^2[0, 1] \longrightarrow L^2[\nu_{min}, \nu_{max}]$ is given by

$$[G'(a)h](\nu) = -E(a)(0, \nu) \int_0^1 h(t)k(t, \nu) dt - \int_0^1 E(a)(s, \nu) \int_s^1 h(t)k(t, \nu) dt ds.$$

4.3 Landweber approach for the linearized problem

At each step of the Newton method, we have to solve the linearized problem:

$$G'(a_j)h = f - G(a_j).$$

In order to regularize the resolution, we solve the latter equation by using the Landweber iterative method as follows. Let $h_0(t) = 0$ be the initial guess and

let η be a fixed value. Compute, for a certain $\beta > 0$,

$$h_{k+1} = h_k - \beta[G'^*(a_j)G'(a_j)h_k - G'^*(a_j)(f - G(a_j))]$$

until $\|G'^*(a_j)G'(a_j)h_k - G'^*(a_j)(f - G(a_j))\| \leq \eta$.

The relationships between this (outer-inner cycles) Newton-Landweber method and the following classical Landweber iteration scheme for non-linear operator, are quite evident

$$a_{j+1} = a_j - \tilde{\beta}G'^*(a_j)(G(a_j) - f).$$

In the first case we compute a better solution of the linearized Newton step, while in the second one we compute a more precise approximation of the derivative G' and G'^* at each new point.

For the Landweber iteration, we need the adjoint operator $G'^*(a) : L^2[\nu_{min}, \nu_{max}] \rightarrow L^2[0, 1]$ of the Fréchet derivative $G'(a)$, that is, by definition, we have to compute the operator $G'^*(a)$ such that $\langle G'(a)h, \phi \rangle_{L^2[\nu_{min}, \nu_{max}]} = \langle h, G'^*(a)\phi \rangle_{L^2[0, 1]}$. We obtain

$$\begin{aligned} \langle G'(a)h, \phi \rangle_{L^2[\nu_{min}, \nu_{max}]} &= \int_{\nu_{min}}^{\nu_{max}} [G'(a)h](\nu)\phi(\nu) d\nu \\ &= \int_{\nu_{min}}^{\nu_{max}} \left\{ -E(a)(0, \nu) \int_0^1 h(t)k(t, \nu) dt \right\} \phi(\nu) d\nu \\ &\quad + \int_{\nu_{min}}^{\nu_{max}} \left\{ -\int_0^1 E(a)(s, \nu) \int_s^1 h(t)k(t, \nu) dt ds \right\} \phi(\nu) d\nu. \end{aligned}$$

For the first integral, we have

$$\begin{aligned} &\int_{\nu_{min}}^{\nu_{max}} \left\{ -E(a)(0, \nu) \int_0^1 h(t)k(t, \nu) dt \right\} \phi(\nu) d\nu \\ &= \int_0^1 \left\{ \int_{\nu_{min}}^{\nu_{max}} \left(-E(a)(0, \nu)k(t, \nu)\phi(\nu) \right) d\nu \right\} h(t) dt, \end{aligned}$$

and for the second one:

$$\begin{aligned} &\int_{\nu_{min}}^{\nu_{max}} \left\{ -\int_0^1 E(a)(s, \nu) \int_s^1 h(t)k(t, \nu) dt ds \right\} \phi(\nu) d\nu \\ &= -\int_0^1 ds \int_s^1 dt \int_{\nu_{min}}^{\nu_{max}} E(a)(s, \nu)h(t)k(t, \nu) \phi(\nu) d\nu \\ &= -\int_0^1 dt \int_0^t ds \int_{\nu_{min}}^{\nu_{max}} E(a)(s, \nu)h(t)k(t, \nu) \phi(\nu) d\nu \\ &= \int_0^1 \left\{ \int_{\nu_{min}}^{\nu_{max}} \int_0^t \left(-E(a)(s, \nu)k(t, \nu)\phi(\nu) \right) ds d\nu \right\} h(t) dt. \end{aligned}$$

The adjoint operator $G'^*(a)$ is then defined as follows

$$[G'^*(a)\phi](t) = \int_{\nu_{min}}^{\nu_{max}} \left\{ -E(a)(0, \nu)k(t, \nu) + \int_0^t \left(-E(a)(s, \nu)k(t, \nu) \right) ds \right\} \phi(\nu) d\nu$$

We notice that both the linearized operator $G'(a)$ and its adjoint $G'^*(a)$ are the following convolution operators

$$G'(a) : L^2[0, 1] \longrightarrow L^2[\nu_{min}, \nu_{max}] \quad \text{such that} \quad [G'(a)h](\nu) = \int_0^1 g_a(t, \nu)h(t) dt,$$

and

$$G'^*(a) : L^2[\nu_{min}, \nu_{max}] \longrightarrow L^2[0, 1] \quad \text{such that} \quad [G'^*(a)\phi](t) = \int_{\nu_{min}}^{\nu_{max}} g_a(t, \nu)\phi(\nu)d\nu,$$

where

$$\begin{aligned} g_a(t, \nu) &= -k(t, \nu) \left(E(a)(0, \nu) + \int_0^t E(a)(s, \nu) \right) \\ &= -k(t, \nu) \left\{ \exp \left(- \int_0^1 a(\tau)k(\tau, \nu)d\tau \right) + \int_0^t \exp \left(- \int_s^1 a(\tau)k(\tau, \nu)d\tau \right) ds \right\} \end{aligned}$$

Indeed we have

$$\begin{aligned} [G'(a)h](\nu) &= \int_0^1 \left(- E(a)(0, \nu)h(t)k(t, \nu) \right) dt - \int_0^1 ds \int_s^1 dt E(a)(s, \nu)h(t)k(t, \nu) \\ &= \int_0^1 \left(- E(a)(0, \nu)h(t)k(t, \nu) \right) dt - \int_0^1 dt \int_0^t ds E(a)(s, \nu)h(t)k(t, \nu) \\ &= \int_0^1 \left(- E(a)(0, \nu)k(t, \nu) - \int_0^t E(a)(s, \nu)k(t, \nu) ds \right) h(t) dt \end{aligned}$$

4.4 Computational matters

For a given a , we need to compute

$$Y(s, \nu) := \exp \left(- \int_s^1 a(\tau)k(\tau, \nu)d\tau \right).$$

We could do this by solving the initial value problem

$$\begin{cases} \frac{\delta Y}{\delta s}(s, \nu) = a(s)k(s, \nu)Y(s, \nu) \\ Y(1, \nu) = 1 \end{cases}$$

by marching backwards with $N + 1$ equidistant nodes from $s_N = 1$ to $s_0 = 0$ and using a centered difference scheme:

$$\begin{aligned} Y(s_N, \nu) &= 1 \\ Y(s_{N-1}, \nu) &= Y(s_N, \nu) - \Delta s a(s_N)k(s_N, \nu)Y(s_N, \nu) \\ Y(s_j, \nu) &= Y(s_{j+2}, \nu) - 2\Delta s a(s_{j+1})k(s_{j+1}, \nu)Y(s_{j+1}, \nu), \end{aligned}$$

where j goes from $N - 2$ to 0 .

By a numerical point of view, the latter scheme is highly unstable. So we use the normal trapezoidal rule:

$$Y(s_j, \nu) = \exp \left\{ -\frac{\Delta s}{2} \left(a(s_j)k(s_j, \nu) + 2 \sum_{k=j+1}^{N-1} a(s_k)k(s_k, \nu) + a(s_N)k(s_N, \nu) \right) \right\},$$

where $\Delta s = 1/2N$.

Next we compute for $N + 1$ equidistant nodes $t_0 = 0$ to $t_N = 1$,

$$W(t_p, \nu) = \int_0^{t_p} \exp \left(- \int_s^1 a(\tau)k(\tau, \nu) d\tau \right) ds$$

via the trapezoidal rule:

$$W(t_p, \nu) = \frac{\Delta s}{2} \left(Y(s_0, \nu) + 2 \sum_{j=1}^{p-1} Y(s_j, \nu) + Y(s_p, \nu) \right).$$

Then we obtain

$$G(a)(\nu) = Y(s_0, \nu) + W(t_N, \nu)$$

and

$$g_a(t_p, \nu) = -k(t_p, \nu) \left\{ Y(s_0, \nu) + W(t_p, \nu) \right\}.$$

The linearized operators can be computed now by the following formulas

$$[G'(a)h](\nu) = \frac{\Delta s}{2} \left\{ g_a(t_0, \nu)h(t_0) + 2 \sum_{j=1}^{N-1} g_a(t_j, \nu)h(t_j) + g_a(t_N, \nu)h(t_N) \right\}$$

respectively

$$[G'^*(a)\phi](t) = \frac{\Delta \nu}{2} \left\{ g_a(t, \nu_0)\phi(\nu_0) + 2 \sum_{j=1}^{M-1} g_a(t, \nu_j)\phi(\nu_j) + g_a(t, \nu_M)\phi(\nu_M) \right\}.$$

The numerical results of same sample problem have shown a low convergence speed of the Newton method, since the shift $G'^*(a_j)(G(a_j) - f)$ is very small after the first iterations.

4.5 Monotone considerations

Throughout we consider the pointwise monotony relation, i.e. for a and b functions on $[u, v]$ we say $a \leq b$ if $a(t) \leq b(t)$ for all $t \in [u, v]$.

An operator B is *monotonically decreasing* if $a < b$ implies $B(a) > B(b)$. An operator B is called *concave* if

$$B(\lambda a + (1 - \lambda)b) \geq \lambda B(a) + (1 - \lambda)B(b), \text{ for all } \lambda \in [0, 1]$$

and (?)

$$B(a + h) - B(a) \leq B'(a)h.$$

The following holds.

Proposition 4.5.1 *The forward operator G is monotonically decreasing and concave.*

Proof. Since $e^{-l} > e^{-m}$, if $l < m$, then, by considering $l = \int_s^1 a(t)k(t, \nu) dt$ and $m = \int_s^1 b(t)k(t, \nu) dt$, with $0 \leq a < b$, we can write

$$E(a)(s, \nu) = \exp\left(-\int_s^1 a(t)k(t, \nu) dt\right) > \exp\left(-\int_s^1 b(t)k(t, \nu) dt\right) = E(b)(s, \nu).$$

Then G is monotonically decreasing, as follows

$$G(a)(\nu) = E(a)(0, \nu) + \int_0^1 E(a)(s, \nu) ds > E(b)(0, \nu) + \int_0^1 E(b)(s, \nu) ds = G(b)(\nu).$$

Since $e^{-\lambda l + (1-\lambda)m} \leq \lambda e^{-l} + (1-\lambda)e^{-m}$, we can write

$$\begin{aligned} E(\lambda a + (1 - \lambda)b)(s, \nu) &= \exp\left(-\int_s^1 (\lambda a(t) + (1 - \lambda)b(t))k(t, \nu) dt\right) \\ &\leq \lambda \exp\left(-\int_s^1 a(t)k(t, \nu) dt\right) + (1 - \lambda) \exp\left(-\int_s^1 b(t)k(t, \nu) dt\right) \\ &= \lambda E(a)(s, \nu) + (1 - \lambda)E(b)(s, \nu). \end{aligned}$$

We have

$$\begin{aligned} G(\lambda a + (1 - \lambda)b)(\nu) &= E(\lambda a + (1 - \lambda)b)(0, \nu) + \int_0^1 E(\lambda a + (1 - \lambda)b)(s, \nu) ds \\ &\leq \lambda E(a)(0, \nu) + (1 - \lambda)E(b)(0, \nu) + \int_0^1 (\lambda E(a)(s, \nu) + (1 - \lambda)E(b)(s, \nu)) ds \\ &= \lambda E(a)(0, \nu) + (1 - \lambda)E(b)(0, \nu) \\ &\quad + \lambda \int_0^1 E(a)(s, \nu) ds + (1 - \lambda) \int_0^1 E(b)(s, \nu) ds \\ &= \lambda G(a)(\nu) + (1 - \lambda)G(b)(\nu). \end{aligned}$$

■

This allow us the analogy with one dimensional case. We call a a *super-solution* (*sub-solution*) if $G(a) \geq f$ ($G(a) \leq f$).

Lemma 4.5.2 *If a is a sub-solution, then the Newton method gives a larger sub-solution.*

Proof. Let a be a sub-solution and let h be the solution of $G'(a)h = f - G(a)$. We have that $G(a + h) \leq G(a) + G'(a)h = f$, which states that $a + h$ is a sub-solution too.

Furthermore, $G'(a)h \geq 0$ since $f - G(a) \geq 0$, which gives

$$G(a + h) \leq G(a) + G'(a)h \leq G(a).$$

■

Our question:

Since the function is (pointwise) concave, can we extend in same “**heuristic way**” the simple scheme Newton-Secant?

We know Newton-Secant works properly in the simple 1-D case, and it gives an useful bound for the distance between the iterative solution and the true (unknown) solution. Obviously it cannot be extended directly to multidimensional setting. We recall that then the secant method gives the new approximation $c = a + [g(a)/(g(b) - g(a))](b - a)$, which lies inside (a, b) if $g(a)g(b) < 0$.

The convergence of the Landweber method is often quite slow. In our first test examples, Landweber “converges” to a function that is far from the true solution (see Fig. 4.2, A), while the residual is very small (see Fig. 4.2, B). We try to include monotone ideas to pass by these Landweber trapping regions.

As first example, we test the following scheme.

Let a be the result coming from the Landweber method. Since the operator is concave, this solution is usually “sub-solution-close”, which means that $\int(G(a) - f) d\nu < 0$.

Let b be a super-solution (the simplest one is $b \equiv 0$, related to the upper bound $G(b) \equiv 2$).

Now we consider subsequent corrections of the Newton-Landweber solution, as follows.

Let $j = 1, 2, \dots$

- If j is even

$$c = \text{Landweber}(a + \gamma(b - a))$$

with few iterations, where $\text{Landweber}(h)$ denotes the Landweber method with h as initial guess.

- otherwise

$$c = a + \gamma(b - a)$$

Here γ has to be a shift value in the direction of the expected better solution, such as either $\gamma = \int G(a) d\nu / \int (G(b) - G(a)) d\nu$ (secant method) or $\gamma = 1/2$ (middle point).

If $\|G(c) - f\| < \epsilon$, where ϵ is the estimate data noise, then c is the solution and stop.

Otherwise,
 if $\int (G(c) - f) d\nu > 0$, let $b = c$, and
 if $\int (G(c) - f) d\nu < 0$, let $a = c$.

4.6 Numerical results

We test the simplified model with a discretization of 25 equidistant nodes in $[0, 1]$ for the altitude, and 50 equidistant nodes in $[\nu_{\min}, \nu_{\max}]$ for the frequencies, where $\nu_{\min} = 40$; $\nu_{\max} = 89$.

The function k is $k(s, \nu) = (\sqrt{10}(N + M))^{-1} * \sqrt{s + \nu}$.

The input (true) solution is the function $a(t)_* = \exp\{-(t - 0.5)^2/0.2^2\}$.

The initial guess for the Landweber method is
 $a(t)_0 = a(t)_* + 0.3 * \exp\{-t(-0.3)^2/0.1^2\} + 0.5 * \exp\{-(t - 0.5)^2/0.2^2\} + 0.3 * \exp\{-(t - 0.7)^2/0.1^2\}$.

Fig. 4.1, shows true solution, initial guess and reconstructions by Landweber method after 50 iterations; the restoration errors and the residuals are shown in Fig. 4.2-A and Fig. 4.2-B (diamond = true solution; plus sign = Landweber initial guess; asterisk = Landweber last iteration).

After the first 30 iterations, the Landweber method does not improve the solutions. Note that at the latter iteration the residual is small, while the errors are still high.

In order to get away from the region where the Landweber method is static, we apply the monotone trick of the previous section.

Fig. 4.3, show the first 5 iterations of the method. Fig. 4.4, show the next, up to the 20-th, iterations of the method (diamond = true solution; plus sign = Landweber initial guess; asterisk = Landweber last iteration; dashed line = Landweber iteration in the monotone scheme; dotted line = iteration with linear combination in the monotone scheme; points = last iteration of the monotone scheme).

The restoration errors and the residuals are shown in Fig. 4.5 and an improvement of the solution happens.

Unfortunately, the algorithm does not give monotonic decreasing of the

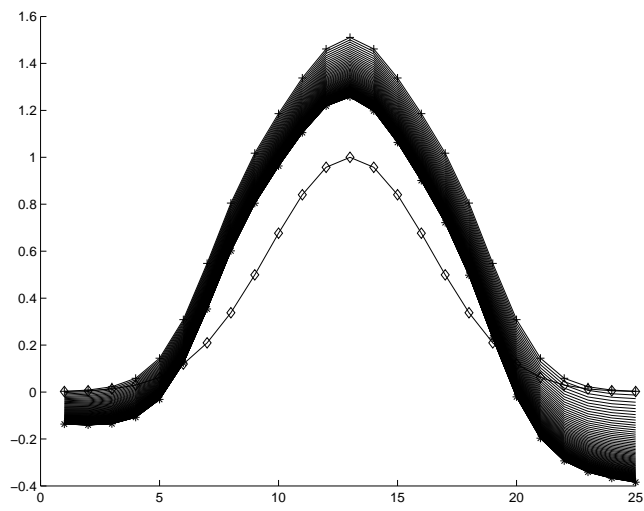
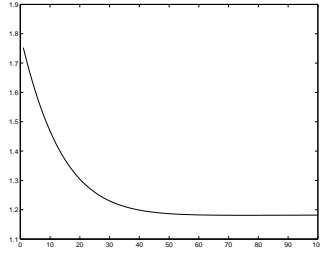
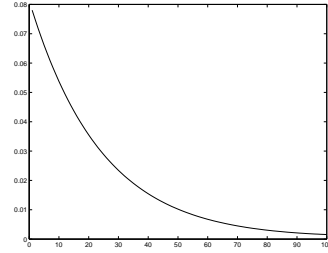


Figure 4.1: Iterations of the Landweber method



A : Reconstruction errors



B : Residuals

Figure 4.2: Reconstruction errors and Residuals of the Landweber Method

restoration error.

As we expect, differing from the $1-d$ case, a subsequent iterations could be worst than the previous. The key point is that the concept of "sub-solution-close" a , and its measure $\int (G(a) - f) d\nu$, are only heuristic considerations.

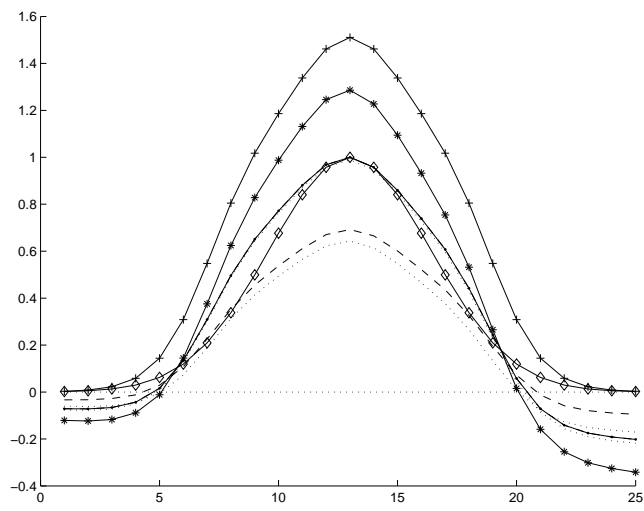


Figure 4.3: First 5 iterations of the monotonic method

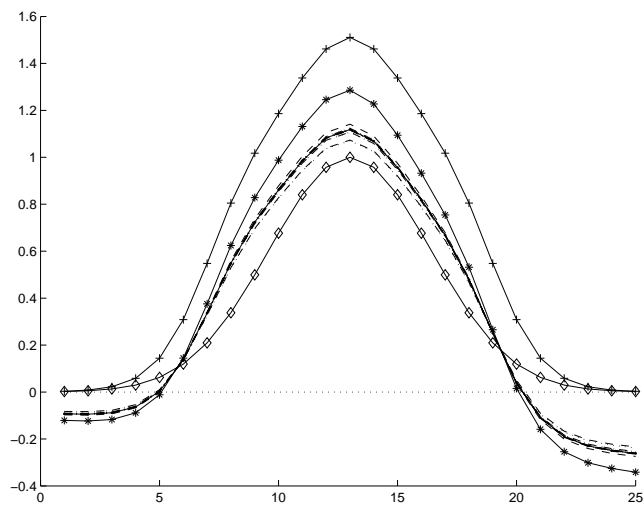
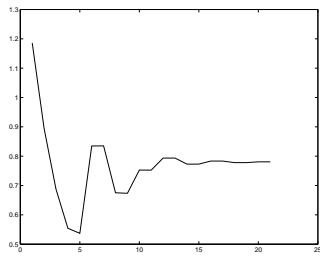
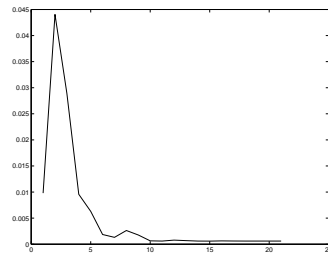


Figure 4.4: Next iterations of the monotonic method



A : Reconstruction errors



B : Residuals

Figure 4.5: Reconstruction errors and Residuals of the monotonic method

Chapter 5

Future Work

All of the members of the group have expressed an interest in continuing to work on this problem. Brian Borchers is interested in following up on the suggestions for improving and validating the implementation of the existing statistical approach discussed in Chapter 1. Guillaume Bal and Kui Ren would like to extend the work on small inclusions discussed in Chapter 2. Ronny Ramlau and Markus Jäckels are continuing to work on the numerical scheme discussed in Chapter 3. Alexandru Tamasan and Claudio Estatico would like to continue working on the implementation of the Newton linearization approach discussed in chapter 4.