Abel Inversion using the Maximum Entropy Method

Danilo R. Neuber[*]
Wolfgang von der Linden[†]

3rd October 2003

Contents

[*]Inst. für Theoretische Physik, Tel.: +43/3 16/8 73-81 88; neuber@itp.tu-graz.ac.at
[†]Inst. für Theoretische Physik, Tel.: +43/3 16/8 73-81 70; wvl@itp.tu-graz.ac.at

1 Introduction

The density ρ of a plasma cannot be measured directly, but its absorbance A of a laser beam can be obtained experimentally. The absorbance along a straight line is proportional to the corresponding line integral over the density $\rho(\boldsymbol{r})$. Assuming a radial symmetric plasma column (see Fig. 1), absorbance and density are related via the Abel transformation:

$$A(y) = \int_y^R \frac{2r}{\sqrt{r^2 - y^2}}\,\rho(r)\,dr, \tag{1}$$

which can be rewritten as linear integral equation of the form:

$$A(y) = \int_0^R K_y(r)\rho(r)\,dr \qquad \text{with} \qquad K_y(r) = \Theta(r-y)\frac{2r}{\sqrt{r^2 - y^2}}. \tag{2}$$

Due to the singularity of the integral kernel $K_y(r)$ at $r = y$, the inversion of (2) turns out to be an ill-conditioned problem which we will tackle by applying the Maximum Entropy-Method.

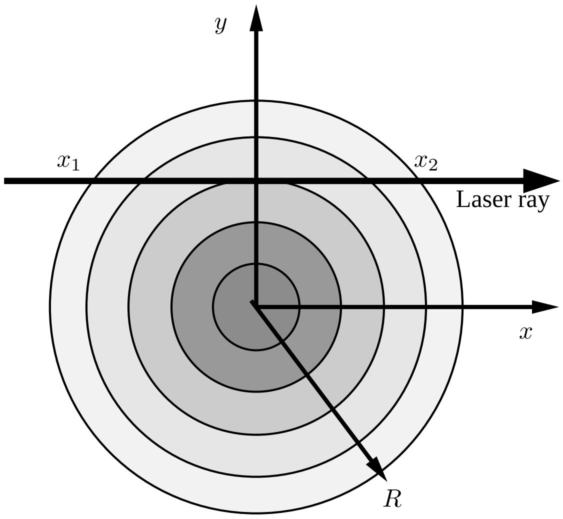

Figure 1: Experimental geometry for Abel inversion

2 Establishing the model matrix

For numeric calculations we approximate the density $\rho(r)$ by a piecewise constant function:

$$\rho(r) = \sum_{j=1}^{N_0} \rho_j \chi_j(r), \tag{3}$$

where $\chi_j(x)$ designates the characteristic function of the interval $(j-1)R/N_0 \leq x < jR/N_0$. Using this simple ansatz, we can substitute (3) into (2) and perform the integration analytically. Assuming that N_d measurements are made at positions

$$y_i = \left(i - \frac{1}{2}\right)\frac{R}{N_d}, \tag{4}$$

we obtain the matrix equation:

$$A_i \equiv A(y_i) = \sum_{j=1}^{N_0} \tilde{M}_{ij}\rho_j. \tag{5}$$

Setting

$$r_j = j\frac{R}{N_0}, \tag{6}$$

the model matrix $\tilde{M}$ is given by:

$$\tilde{M}_{ij} = \begin{cases} 2\left(\sqrt{r_j^2 - y_i^2} - \sqrt{r_{j-1}^2 - y_i^2}\right) & \text{if} \quad y_i \leq r_{j-1} \\ 2\left(\sqrt{r_j^2 - y_i^2}\right) & \text{if} \quad r_{j-1} < y_i \leq r_j \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

In order to ensure smoothness of the plasma density $\rho(r)$ we represent it by a natural cubic spline passing through N_k control points $(\hat{r}_k, \hat{\rho}_k)$ and calculate ρ_j by interpolation at r_j. Fixing $\hat{r}_k$ and r_j while varying only $\hat{\rho}_k$, we get ρ_j by a simple matrix multiplication: $\rho_j = \sum_k S_{jk}\hat{\rho}_k$. In order to obtain the spline matrix S_{jk} we calculate the cubic spline $S_k(r)$ passing through $(\hat{r}_l, \delta_{kl})$ and set $S_{jk} = S_k(r_j)$.

We finally arrive at a matrix equation relating $\hat{\rho}_k$ and the absorbance A_i:

$$A_i = \sum_{j,k} \tilde{M}_{ij} S_{jk} \hat{\rho}_k \equiv \sum_{k=1}^{N_k} M_{ik}\hat{\rho}_k \tag{8}$$

Note that the Maximum Entropy method ensures that the ordinates of the control points $\hat{\rho}_k$ will be always positive but interpolated values may be negative.

3 Quantified MaxEnt

Taking into account noise η which is always present in expermental data, we have to rewrite (8) as

$$A_i = \sum_{j=1}^{N} M_{ij}\hat{\rho}_j + \eta_i \tag{9}$$

In order to infer $\hat{\rho}$, we apply Bayesian probability theory for the calculation of the posterior probability $p(\hat{\boldsymbol{\rho}}|\boldsymbol{A},\mathcal{I})$ of a density distribution $\hat{\boldsymbol{\rho}}$ given the absorbance data $\boldsymbol{A}$. Using Bayes' Theorem we obtain:

$$p(\hat{\boldsymbol{\rho}}|\boldsymbol{A},\mathcal{I}) = \frac{p(\boldsymbol{A}|\hat{\boldsymbol{\rho}},\mathcal{I})p(\hat{\boldsymbol{\rho}}|\mathcal{I})}{p(\boldsymbol{A}|\mathcal{I})}, \tag{10}$$

Assuming uncorrelated Gaussian noise with mean zero and variance σ_i^2 the likelihood is given by:

$$p(\boldsymbol{A}|\hat{\boldsymbol{\rho}},\mathcal{I}) = \prod_{i=1}^{N_d} \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{1}{2}\sum_{i=1}^{N_d}\left[\frac{A_i - \sum_j M_{ij}\hat{\rho}_j}{\sigma_i}\right]^2\right) \tag{11}$$

As prior we choose the MaxEnt prior (see e.g. [Siv96] for further discussion):

$$p(\hat{\boldsymbol{\rho}}|\alpha,\boldsymbol{m},I) = \frac{(2\pi)^{-N/2}\alpha^{N/2}}{\sqrt{\hat{\rho}_1\cdots\hat{\rho}_N}} \exp(\alpha S), \tag{12}$$

with hyperparameter α, default model $\boldsymbol{m}$ and entropy S defined by:

$$S = \sum_{j=1}^{N} \hat{\rho}_j - m_j - \hat{\rho}_j \log\frac{\hat{\rho}_j}{m_j} \tag{13}$$

The default model $\boldsymbol{m}$ is the MaxEnt-solution in case of strong regularization ($\alpha \to \infty$).

In order to obtain the maximum posterior or MAP solution for fixed hyperparameter α, we have to maximize

$$\exp\left(-\frac{\chi^2}{2} + \alpha S\right)$$

denoting the misfit by $\chi^2 = \sum_i (A_i - \sum_j M_{ij}\hat{\rho}_j)^2/\sigma_i^2$ and ignoring the denominator in (12) since we assume that it does not vary much compared to the exponential. The maximum is calculated numerically by Newton's method, for the details we refer to [Lin01].

Finally we have to specify the optimal hyperparameter α, which is fixed by the relation:

$$\chi^2 = N_d, \tag{14}$$

since we expect each data point A_i to deviate by σ_i from its true value on average. Starting with large α (in order to ensure convergence of the Newton iterations), the hyperparameter is determined by interval bisection.

4 Exercises

Using the Program

The mock data for the absorbance $\boldsymbol{A}$ are generated by an Abel transformation of the density function:

$$\rho(r) = r^2(1 - r^2) + \frac{1}{5} \exp\left\{-250\left(r - \frac{1}{4}\right)^2\right\}, \tag{15}$$

which is evaluated at $N_0 =$ `N_grid` equally spaced points and multiplied by the model matrix $\tilde{M}$ defined by (7). Then, Gaussian noise with zero mean and standard deviation `Sigma noise` is added; `seed` may be changed to obtain different sequencies of normally distributed random numbers.

Before starting the reconstruction by pressing the button `Compute`, the standard deviation of the data points corresponding to σ_i in (11), the value of the default model $\boldsymbol{m}$ which is chosen flat and the initial value of the hyperparameter α have to be specified. Moreover, the number of data points `N_data` and the number of control points `N_knots` for spline interpolation may be set.

Suggested exercises

- Increase the initial value for `alpha` to `1.e6` and watch how the reconstruction converges, starting from the default model.

- Set the default model to `1.e-1`: The reconstructed density no longer vanished at $r = 0$. A look at Fig. 1 reveals that there is only one measurement taking into account the density $\rho(0)$ at the center, while all measurement include information on $\rho(R)$. Therefore, the density near the center is mainly determined by our choice for the default model — in case of insufficient data we will get what we used as input!

- Increase both `Noise` and `Standard deviation` to `0.03` and `0.003`; in the latter case you may need to increase the number of knot points for the spline interpolation, for otherwise there are not enough degrees of freedom to reduce the misfit $\chi^2 \leq N_d$.

  Note that you will probably not achive convergence if you choose the noise level greater that the standard deviation for the reconstruction.

- Restore default values (press `Reset`) and try different numbers of knot points `N_knot`. For too few points the criterion $\chi^2 = N_d$ cannot be satisfied, if too many points are used, ripples appear in the reconstructed density — the data are overfitted.

References

[Lin01] W. von der Linden, A. P.: *Skriptum Wahrscheinlichkeitstheorie, Statistik und Datenanalyse*. 1. 2001 4

[Siv96] Sivia, D.S.: *Data Analysis – A Bayesian Tutorial*. 1. Oxford University Press, 1996 4