Journal of Physics: Conference Series **490** (2014) 012149 doi:10.1088/1742-6596/490/1/012149

Numerical aspects of grazing incidence XRF

Eduardo Miqueles[1], Carlos Pérez[1] and Vanessa I.T. Suaréz[2]

[1] Brazilian Synchrotron Light Source/CNPEM - Campinas, SP/Brazil
[2] Universidade Federal de Viçosa/Dept. of Physics - Viçosa, MG/Brazil

E-mail: edu.miqueles@gmail.com

Abstract. In this article, we give a different mathematical approach for background aspects of grazing incidence x-ray fluorescence, GIXRF for short. Our contribution comes from an applied point of view, in order to have a computer program to simulate the fluorescence intensity from a stacking of thin layer films. A typical ill-posed inverse problem is formulated. Our aim is to reconstruct the fluorescence intensity for a variety of grazing angle measurements. We rederive some classical equations pointing out the numerical aspects of the inversion procedure and giving new directions for direct in inverse algorithms.

1. Introduction

Grazing incidence x-ray fluorescence - GIXRF - experiments are used to study the propagation of electromagnetic waves in a stratified media. The fluorescence yield by wave penetration at glancing incidence above the critical angle, provides information about the upper layers. The theory for GIXRF is beyond the scope of this article. For a complete description of the main equations and physical properties we refer to [1, 2, 3, 4]. A typical layered media, representing a thin-film with 5 layers, is shown in figure 1. Interfaces separating each layer are marked at the left, starting from 1 to 6. The eletromagnetic source is placed at vacuum, and is represented by the asterisk in figures 1 and 2. In this short example, the incoming wave strikes the first interface with a grazing angle θ. Assuming θ above the critical angle, the electromagnetic wave penetrates in the medium, were reflected and transmitted waves propagates according to Snell's law. If θ varies within a short interval, containing the critical angle, we have a family of waves striking the first interface, propagating through layers, and coming back to the surface, were the fluorescence intensity is finally measured, as depicted in figure 2. In general, we denote Σ_j as the jth interface separating layer j and $j+1$. Each layer has a complex refraction index n_j and the pair $(a_j, b_j) \in \mathbb{C}$ at each interface denote the amplitude of the incident and reflected wave at Σ_j.

Each layer is a composition of many chemical elements, say copper, cobalt, zinc, gold, among others. According to the scattering theory [1, 3], the total fluorescence intensity for a given element m is given by

$$F_m(\theta, \boldsymbol{u}) = \sum_{j=1}^{N} c_{mj}(\boldsymbol{u}) \int_0^{h_j} f_j(z, \theta, \boldsymbol{u}) e^{-d_{mj}(\boldsymbol{u})z} \mathrm{d}z, \quad f_j(z, \theta, \boldsymbol{u}) = |a_j(\theta, \boldsymbol{u}) e^{ikzp_j(\theta)} + b_j(\theta, \boldsymbol{u}) e^{-ikzp_j(\theta)}|^2, \qquad (1)$$

with $c_{mj}(\boldsymbol{u})$ and $d_{mj}(\boldsymbol{u})$ appropriate physical constants, dependent on the parameter $\boldsymbol{u} \in \mathbb{R}^N$. Each entry of the vector $\boldsymbol{f} = (f_j) \in \mathbb{R}^N$ represents the incident fluorescent intensity at the layer

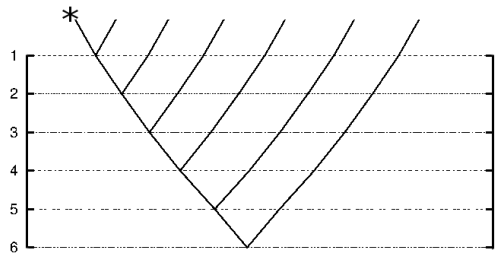

Figure 1. Single x-ray incidence

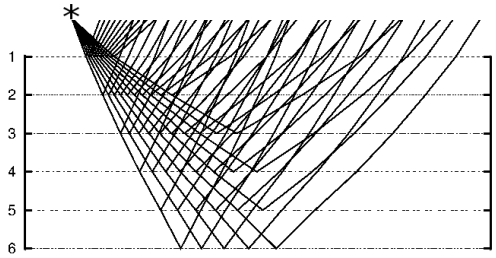

Figure 2. Total x-rays incidence

j, whereas $\boldsymbol{h} = (h_j) \in \mathbb{R}^N$ is the thickness of the jth layer and $|\cdot|$ stands for the complex modulus. Function $\theta \in \Theta \mapsto F_m(\theta, \boldsymbol{u}) \in \mathbb{R}$ is a real-valued function of the angle θ. In pratice, a measure of total fuorescence F_m is obtained, say $\{q_1, q_2, \ldots, q_v\}$ satisfying $F_m(\theta_k) = q_k$ (at least hypothetically), $\theta_k \in \Theta$. The mathematical challenge can be stated as the following non-linear fitting problem:

Problem [♠]: Find $\boldsymbol{u} \in \mathbb{R}^N$ such that $q_k - F_m(\theta_k, \boldsymbol{u})$ is minimized.

There are many optimization techniques to solve the above non-linear least squares problem, for instance the Gauss-Newton or the Levenberg-Marquardt method. Like most of the iterative approaches, we find a sequence $\{\boldsymbol{u}^k\}$ converging to the approximated solution, and the computation of each iterate strongly depends on the value $F_m(\theta, \boldsymbol{u}^k)$, for a given θ. The parameter $\boldsymbol{u}$ could be the vector of layer thickness $\boldsymbol{u} = \boldsymbol{h}$, or the concentration of a given element within each layer.

2. The direct problem

We aim at calculating the function value $F_m(\theta, \boldsymbol{u})$, for a fixed pair $(\theta, \boldsymbol{u})$, through (1). Let us denote the inner integral in (1) by $\ell_{jm}(\theta, \boldsymbol{u})$. It is easy to note that ℓ_{jm} resembles a Laplace transform of function f_j and it is easily calculated using a quadrature formula. Before that, we have to evaluate function f_j, which is dependent on the amplitude values a_j and b_j. From the matrix formalism [1, 3, 4, 5], these amplitudes are given by the Fresnel equations and the principle of reversibility:

$$\left\{ \begin{array}{l} \nu_j a_j + \nu_j r_j b_j - a_{j-1} t_j = 0 \\ \frac{r_j}{\nu_j} a_j + \frac{1}{\nu_j} b_j - b_{j-1} t_j = 0 \end{array} \right. \Leftrightarrow \left[\begin{array}{c} a_{j-1} \\ b_{j-1} \end{array} \right] = \frac{1}{t_j} \left[\begin{array}{cc} \nu_j & 0 \\ 0 & \frac{1}{\nu_j} \end{array} \right] \left[\begin{array}{cc} 1 & r_j \\ r_j & 1 \end{array} \right] \left[\begin{array}{c} a_j \\ b_j \end{array} \right], \quad \nu_j \equiv \exp\left(i \frac{2\pi}{\lambda} p_j h_j \right) \tag{2}$$

where t_j, r_j are the transmission and reflection coefficients[1] at Σ_j, respectively. As stablished in the literature, the matrix relation in (2) is used to find (a_j, b_j) in terms of (a_{j-1}, b_{j-1}) in a recursive way, see [3, 5, 6]. Now, we take a different approach. Let us write down the Fresnel equations in (2), with index j varying from 1 to $N + 1$. To make easier the equations, we suppose $N = 2$; the generalization will go straightforward. In the absence of reflected waves at the substrate, i.e., $b_{N+1} = 0$, we will have:

$$\left[\begin{array}{cccccc} \nu_1 & 0 & 0 & 0 & r_1\nu_1 & 0 \\ -t_2 & \nu_2 & 0 & 0 & 0 & r_2\nu_2 \\ 0 & -t_3 & \nu_3 & 0 & 0 & 0 \\ \frac{r_1}{\nu_1} & 0 & 0 & -t_1 & \frac{1}{\nu_1} & 0 \\ 0 & \frac{r_2}{\nu_2} & 0 & 0 & -t_2 & \frac{1}{\nu_2} \\ 0 & 0 & \frac{r_3}{\nu_3} & 0 & 0 & -t_3 \end{array} \right] \left[\begin{array}{c} a_1 \\ a_2 \\ a_3 \\ b_0 \\ b_1 \\ b_2 \end{array} \right] = \left[\begin{array}{c} a_0 t_1 \\ 0 \\ 0 \\ -b_3 \frac{1}{\nu_3} \\ 0 \\ 0 \end{array} \right] \Leftrightarrow \underbrace{\left[\begin{array}{cc} \mathbf{H}_U & \mathbf{Q} \\ \mathbf{D} & \mathbf{H}_L \end{array} \right]}_{\text{matrix } \mathbf{L}} \underbrace{\left[\begin{array}{c} \mathbf{a} \\ \mathbf{b} \end{array} \right]}_{\text{unknown}} = \left[\begin{array}{c} a_0 t_1 \boldsymbol{\rho}_1 \\ \mathbf{0} \end{array} \right] \tag{3}$$

[1] For a grazing incidence angle θ at the surface, the coefficients (t_j, r_j) at Σ_j - see [3] - are given by $t_j = \frac{2p_{j-1}}{p_{j-1}+p_j}$ and $r_j = \frac{p_{j-1}-p_j}{p_{j-1}+p_j}$, with $p_j = \sqrt{\theta^2 - 2\delta_j - 2i\beta_j}$

with $\mathbf{a} = (a_1, a_2, \ldots, a_{N+1})^T$ and $\mathbf{b} = (b_0, b_1, \ldots, b_N)^T$. In the multidimensional case (right hand side of (3)), $\mathbf{D} = \text{diag}[(r_j \nu_j^{-1})] \in \mathbb{C}^{(N+1)\times(N+1)}$ is a diagonal matrix while $\mathbf{H}_U, \mathbf{H}_L \in \mathbb{C}^{(N+1)\times(N+1)}$ are the so called Hessenberg matrices, upper and lower respectively. Also, $\boldsymbol{\rho}_j \in \mathbb{C}^{N+1}$ is the jth canonical vector. A similar approach, not exactly in the same matrix context, was also stablished in [2]. It should be noted that (3) gives us the following results:

$$\left\{ \begin{array}{l} \mathbf{H}_U \mathbf{a} + \mathbf{Q} b = a_0 t_1 \boldsymbol{\rho}_1 \\ \mathbf{D}\mathbf{a} + \mathbf{H}_L \mathbf{b} = \mathbf{0} \end{array} \right. \quad \Rightarrow \quad \left\{ \begin{array}{l} \mathbf{a} = -\mathbf{D}^{-1}\mathbf{H}_L \mathbf{b} \\ \left[\mathbf{Q} - \mathbf{H}_U \mathbf{D}^{-1} \mathbf{H}_L\right] \mathbf{b} = a_0 t_1 \boldsymbol{\rho}_1 \end{array} \right. \tag{4}$$

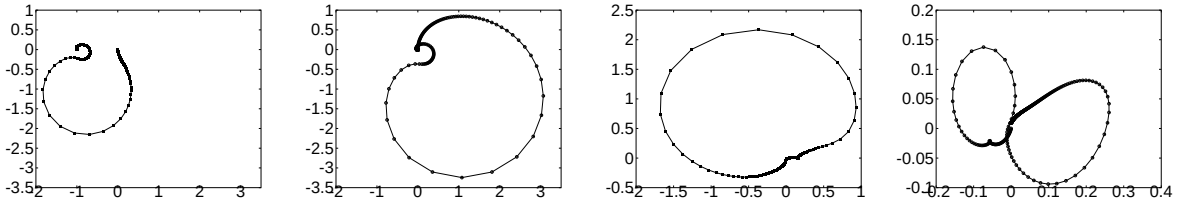

Figure 3. From left to right: Plot of $\{a_1(\theta), a_2(\theta), b_0(\theta), b_1(\theta)\}$ in the complex plane, with $0 \leqslant \theta \leqslant 1^\circ$.

To find the solution $\{\boldsymbol{a}, \boldsymbol{b}\}$, we first solve a linear system to find the vector of reflected amplitudes $\mathbf{b}$, followed by a matrix product to compute the transmitted amplitudes $\mathbf{a}$, as given in (4). It is easy to note that the matrix $\mathbf{M} \equiv \mathbf{Q} - \mathbf{H}_U \mathbf{D}^{-1} \mathbf{H}_L$ is in the tridiagonal form. Therefore, solving $\mathbf{M}\mathbf{b} = a_0 t_1 \mathbf{p}_1$ requires about $O(N)$ operations using a fast algorithm. This is in contrast with the $O(N^3)$ operations requested by the complete Gaussian elimination, see [8, 7]. Another non-optimal approach, involves computing the analytical inverse of matrix $\mathbf{L}$ using the Aitken-block diagonalization (Banachiewicz inversion formula, see [9]). Such an inversion, although explicit, involves at least three tridiagonal inversion; being expensive than solving the system with matrix $\mathbf{M}$. We remark that our strategy has the same computational cost as the one depicted in [3, 5], where N two-dimensional linear systems must be solved to find $\mathbf{a}$ and $\mathbf{b}$.

Example: In order to verify the matrix formalism in (4), we consider an example extracted from [1]. Let us consider a 2-layered thin film, deposited on a thick Silicium substrate; first layer being Cobalt and second Gold. Layers have thickness around $h_1 = 3.4 \times 10^{-9}\mathsf{nm}$ and $h_2 = 2 \times 10^{-}9\mathsf{nm}$ respectively. Complex refractive indexes are given by $n_j = 1 - \delta_j - i\beta_j$ with $\boldsymbol{\delta} = (5.6 \times 10^{-6}, 10.5 \times 10^{-6}, 1.6 \times 10^{-6})$ and $\boldsymbol{\beta} = (19.8 \times 10^{-8}, 129.5 \times 10^{-8}, 0.84 \times 10^{-8})$. Figure 3 shows the complex solution $\boldsymbol{a} = (a_1, a_2) \in \mathbb{C}^2$ and $\boldsymbol{b} = (b_0, b_1) \in \mathbb{C}^2$ of equation (4), for a variety of glancing angles. The intensity function $f_0(z, \theta) = |e^{-ikzp_0} + b_0(\theta) e^{ikzp_0}|^2$ is shown in figures 4 and 5 for fixed values of z and θ. We considered the initial amplitude $a_0 = 1$. The results obtained here are slightly different from [1] since we are not considering mixtures within each layer. ■

3. The inverse problem

As previously stated, we want to solve a fitting problem [♠]. Let the measurement data be given by vector $\boldsymbol{q} \in \mathbb{R}^v$. Functional $\boldsymbol{u} \in \mathbb{R}^N \mapsto r(\boldsymbol{u}) \in \mathbb{R}^v$ (with $v \geqslant N$) has to be minimized using a standard optimization algorithm. Many differential techniques can be used in this problem since $\boldsymbol{r}$ is a differentiable function of $\boldsymbol{u}$. In fact,

$$\boldsymbol{r}(\boldsymbol{u}) = \boldsymbol{q} - \boldsymbol{F}_m(\boldsymbol{u}), \quad \boldsymbol{F}_m(\boldsymbol{u}) = (F_m(\theta_s, \boldsymbol{u})) \in \mathbb{R}^v, \quad F_m(\theta, \boldsymbol{u}) = \boldsymbol{c}_m^T(\boldsymbol{u}) \boldsymbol{\ell}_m(\theta, \boldsymbol{u}) \tag{5}$$

with m standing for a given element under investigation. Here, $\boldsymbol{c}_m^T$ is the mth row of the matrix $\boldsymbol{C} = (c_{mj})$ whereas $\boldsymbol{\ell}_m$ is the mth column of a matrix $\boldsymbol{L} = (\ell_{jm})$. To implement an optimization

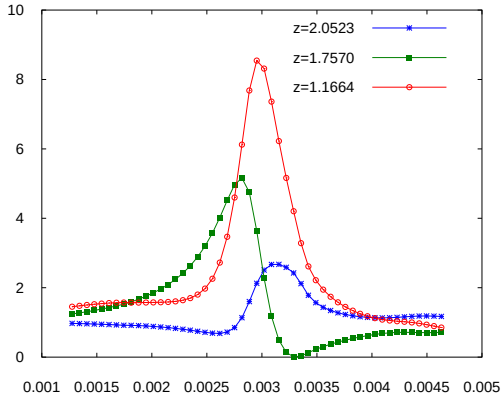

Figure 4. Intensity curve $f_0(z,\cdot)$ at depth $z \in \{2.0523, 1.7570, 1.1664\}$ (unity in nanometer range)

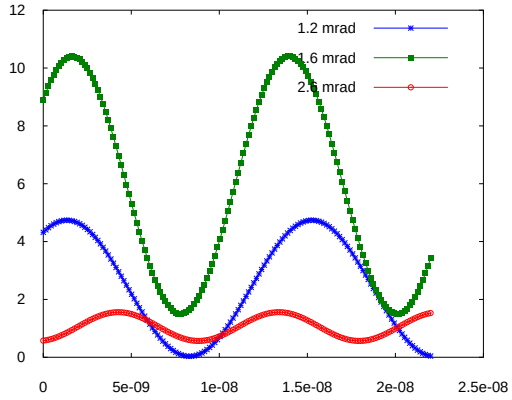

Figure 5. Intensity curve $f_0(\cdot,\theta)$ at glancing angles $\theta \in \{0.0012752, 0.0016107, 0.0026174\}$.

method, one must have the Jacobian of $\boldsymbol{r}$, which is written as $\boldsymbol{J_r}(\boldsymbol{u})^T = [\nabla \boldsymbol{r}_1(\boldsymbol{u}) \dots \nabla \boldsymbol{r}_v(\boldsymbol{u})]$ with $\boldsymbol{r}_s(\boldsymbol{u}) = \boldsymbol{q}_s - F_m(\theta_s, \boldsymbol{u})$. The gradient $\nabla \boldsymbol{r}_s$ depends on the derivatives $\partial_l F_m(\theta_s, \boldsymbol{u}) = \partial_l(\boldsymbol{c}_m)^T \boldsymbol{\ell}_m + \boldsymbol{c}_m^T \partial_l(\boldsymbol{\ell}_m)$, with ∂_l denoting $\equiv \frac{\partial}{\partial \boldsymbol{u}_l}$. Hence, to obtain $\partial_l F_m$, first we need alternatives to obtain the partial derivatives of $\boldsymbol{\ell}_m$, and therefore, the derivatives of $\boldsymbol{a}(\theta, \boldsymbol{u})$ and $\boldsymbol{b}(\theta, \boldsymbol{u})$. This is our new problem:

Problem [★]: Find $\partial_l \boldsymbol{a}(\theta, \boldsymbol{u}) \in \mathbb{C}^{N+1}$ and $\partial_l \boldsymbol{b}(\theta, \boldsymbol{u}) \in \mathbb{C}^{N+1}$ for $l = 1, \dots, v$

Ommiting the angle dependence for simplicity, equation (4) give us a direct formula to compute $\boldsymbol{a}$ and $\boldsymbol{b}$ through simple matrix product, where each matrix also depends on the parameter u. We use this fact to obtain the derivatives. In fact, using $\alpha \equiv a_0 t_1$ and $\boldsymbol{b}(\boldsymbol{u}) = \alpha \mathbf{M}(\boldsymbol{u})^{-1} \boldsymbol{\rho}_1$; and from trivial matrix calculus $\partial_l[\mathbf{M}^{-1}] = -\mathbf{M}^{-1}(\partial_l \mathbf{M})\mathbf{M}^{-1}$:

$$(\partial_l \boldsymbol{b}) = -\mathbf{M}^{-1}(\partial_l \mathbf{M})\mathbf{M}^{-1}(\alpha \boldsymbol{\rho}_1) \quad \Rightarrow \quad \begin{cases} \mathbf{M}(\partial_l \boldsymbol{b}) = -(\partial_l \mathbf{M})\boldsymbol{b} & \text{(a)} \\ \mathbf{M}\boldsymbol{b} = \alpha \boldsymbol{\rho}_1 & \text{(b)} \end{cases} \tag{6}$$

This means that $\partial_l \boldsymbol{b}$ is obtained solving the tridiagonal linear system (6.a), while (6.b) was previously solved. To find $\partial_l \boldsymbol{b}$ we use (4) again. Indeed, from $\mathbf{D}\boldsymbol{a} + \mathbf{H}_L \boldsymbol{b} = 0$ and matrix calculus we obtain $-\mathbf{D}(\partial_l \boldsymbol{a}) = (\partial_l \mathbf{D})\boldsymbol{a} + (\partial_l \mathbf{H}_L)\boldsymbol{b} + \mathbf{H}_L(\partial_l \boldsymbol{b})$. It only remais to find matrices $\{\partial_l \mathbf{D}, \partial_l \mathbf{H}_L, \partial_l \mathbf{H}_U\}$. After some observation, it is easy to realize from (3) that

$$\mathbf{H}_L = -\mathrm{diag}(\boldsymbol{t}) + \sum_{k=1}^{N} \frac{\boldsymbol{\rho}_k \boldsymbol{\rho}_{k+1}^T}{\nu_k}, \quad \mathbf{H}_U = \sum_{k=1}^{N+1} \nu_k \boldsymbol{\rho}_k \boldsymbol{\rho}_k^T - t_{k+1} \boldsymbol{\rho}_{k+1} \boldsymbol{\rho}_k, \quad \mathbf{Q} = \sum_{k=1}^{N} r_k \nu_k \boldsymbol{\rho}_k \boldsymbol{\rho}_{k+1}^T, \quad \mathbf{D} = \sum_{k=1}^{N} \frac{r_k}{\nu_k} \boldsymbol{\rho}_k \boldsymbol{\rho}_k^T \tag{7}$$

where now follows immediately $-\partial_l \mathbf{M} = (\partial_l \mathbf{H}_U)\mathbf{D}^{-1}\mathbf{H}_L + \mathbf{H}_U \partial_l(\mathbf{D}^{-1})\mathbf{H}_L + \mathbf{H}_U \mathbf{D}^{-1} \partial_l(\mathbf{H}_L)$ with

$$\partial_l \mathbf{H}_L = \partial_l \left(\frac{1}{\nu_l} \right) \boldsymbol{\rho}_l \boldsymbol{\rho}_{l+1}^T, \quad \partial_l \mathbf{H}_U = \partial_l(\nu_l) \boldsymbol{\rho}_l \boldsymbol{\rho}_l^T, \quad \partial_l \mathbf{Q} = \partial_l(\nu_l) r_l \boldsymbol{\rho}_l \boldsymbol{\rho}_{l+1}^T, \quad \partial_l \mathbf{D} = \partial_l \left(\frac{r_l}{\nu_l} \right). \tag{8}$$

4. Future work

Our matrix formalism is new and very different from the one pointed out in [3, 5, 1]. A fast computation of the amplitude coefficients $\boldsymbol{a}, \boldsymbol{b} \in \mathbb{R}^{N+1}$ and their partial derivatives $\partial_l \boldsymbol{a}, \partial_l \boldsymbol{b} \in \mathbb{R}^{N+1}$ was presented. A complete algorithmic description, using standard differential optimization techniques, depends on the calculation of these vectors.

References

[1] de Boer D K G, 1991 Glancing-incidence x-ray fluorescence of layered materials, *Physical Review B*, **42(2)**, pp. 498.
[2] Pérez R D, Sánchez H J, Rubio M, Pérez C A, 1999 Mathematical model for evaluation of surface analysis data by total reflection xrf, *X-Ray Spectrometry*, **28**, pp.342.
[3] Sánchez H J, Pérez C A, Pérez R D, Rubio M, 1996 Surface analysis by total-reflection x-ray fluorescence, *Radiat.Phys.Chem.* **48**, pp.325.
[4] Parrat L G, 1994 Surface studies of solids by total reflection of x-rays, *Physical Review*, **95(2)**, pp 359.
[5] Król A, Sher C J and Kao Y H, 1988 X-ray fluorescence of layered synthetic materials with interfacial roughness, *Phyical Review B*, **38(13)**, 8579-92.
[6] Gerrard A and Burch J M, 1975, Introduction to Matrix Methods in Optics, *Dover*, John Wiley & Sons.
[7] Lewis J W, 1982 Inversion of tridiagonal matrices, *Numerische Mathematik*, **38(3)**, pp.333.
[8] Usmani R A, 1994 Inversion of jacobi's tridiagonal matrix, *Computer Math. Applic.*, **27(8)**, pp.59.
[9] Tian Y, Takane Y, 2009 The inverse of any two-by-two nonsingular partitioned matrix and three matrix inverse completion problems, *Computers & Mathematics with Applications*, **57(8)**, pp. 12941304.