



PAPER

The art of model fitting to experimental results

To cite this article: Pedro J Sebastião 2014 *Eur. J. Phys.* **35** 015017

View the [article online](#) for updates and enhancements.

Related content

- [Study of large-angle anharmonic oscillations of a physical pendulum using an acceleration sensor](#)
João C Fernandes, Pedro J Sebastião, Luís N Gonçalves et al.
- [Detecting anharmonicity at a glance](#)
M Gilberti, M Stellato, S Barbieri et al.
- [The influence of spring length on the physical parameters of simple harmonic motion](#)
C A Triana and F Fajardo

Recent citations

- [¹H NMR Relaxometry and Diffusometry Study of Magnetic and Nonmagnetic Ionic Liquid-Based Solutions: Cosolvent and Temperature Effects](#)
Maria Beira *et al*
- [Molecular dynamics of 1-ethyl-3-methylimidazolium triflate ionic liquid studied by ¹H and ¹⁹F nuclear magnetic resonances](#)
Magdalena Wencka *et al*
- [Hydration of MgO-Based Cement: Water Dynamics by ¹H Fast Field-Cycling NMR Relaxometry](#)
Francesca Martini *et al*

The art of model fitting to experimental results

Pedro J Sebastião

Departamento de Física, Instituto Superior Técnico-IST, Universidade de Lisboa-UL,
Av. Rovisco Pais, 1049-001 Lisboa, Portugal
Condensed Matter Physics Center (IST site), Universidade de Lisboa-UL, Complexo
Interdisciplinar, Av. Rovisco Pais, 1049-001 Lisboa, Portugal

E-mail: pedro.jose.sebastiao@tecnico.ulisboa.pt

Received 1 September 2013, revised 15 October 2013

Accepted for publication 31 October 2013

Published 13 December 2013

Abstract

Model fitting to experimental results is presented within the context of graduate physics laboratories and generalized to other graduate and post-graduate levels, and diverse research fields. In most cases the analysis of experimental results, in terms of mathematical models available to describe the obtained results, extends beyond the numerical minimization of statistical estimators, like the chi-square, in the model's parameter space. Dedicated fitting procedures, not easily or directly available in common data analysis software's packages, are required to obtain the best fitting set of parameters that present a consistent physical meaning. A simple but powerful web-based solution is presented, and its relative advantage in comparison with known commercial and open source solutions is discussed.

Keywords: model fitting, data analysis

 Online supplementary data available from stacks.iop.org/EJP/35/015017/mmedia

(Some figures may appear in colour only in the online journal)

1. Introduction

The model fitting of experimental results is a teaching subject considered in physical sciences and engineering graduate courses. The most basic and common example concerns the case of the fit of a straight line to a data set where the dependent observable depends linearly on an independent variable. This case is often presented together with the statistical concept of one-dimensional linear regression and further extended to multiple dimensions [1]. The numerical procedure most often used to estimate the slope and intercept in a linear fit is the least-squares minimization [1]. This simple model fit can be made using a large variety of software packages, some of them well known to both teachers and students. In a recent work Peterlin [2] presented a comparison between spreadsheet versus

statistics suite solutions for data analysis and graphing in physics laboratories. Simple linear regressions and polynomial regressions can be performed using both spreadsheet and dedicated software packages. However, in the case of more complex nonlinear and/or multidimensional model fits, the spreadsheet approach often becomes impractical and other more dedicated statistical suites (either commercial or open source) have to be considered [2].

In introductory physics laboratories, it is often required to analyse experimental results with physical models that have nonlinear dependences on the physical parameters. An electrical oscillatory circuit might be one such case [3]. A serial circuit with an ac power supply, a resistor, R , a capacitor, C , and a coil, L , might operate in different regimes, depending on the type of voltage being applied to the circuit and the values of R , C and L . Another example is the classical mechanical system formed by a mass suspended on a spring. Within the elastic limit the length of the spring depends linearly on the value of the mass in a static experiment or presents a motion with a damped amplitude (oscillatory or not) after an initial spring's elongation/compression. When subjected to an external oscillatory force, the spring–mass system might present a resonance when the frequency of the external force matches the resonance frequency of the system that depends on the elastic constant of the spring, the value of the suspended mass and the friction's coefficient, respectively.

In advanced physics laboratories students and researchers need to analyse experimental results obtained with different experimental techniques such as: x-ray diffraction, nuclear magnetic resonance (NMR) and polarizing optical microscopy, to name a few. These experimental results are in most cases analysed in terms of nonlinear physical models. In professional research work, experimental results obtained for a studied system often include different observables and depend on one or more experimental variables. In those cases, the best results are obtained when the data analysis is performed with a global best-fitting target. This goal adds a degree of complexity to the model-fitting procedures that cannot be easily addressed using spreadsheet software and in some cases requires considerable expertise in the use of some of the dedicated statistical suites available, like those referred to by Peterlin [2]. The approach to obtain the best global model fit to a set of experimental results, where different observables and independent variables have to be considered, requires a set of skills that extends the basic technical competence in numerical analysis and least-squares minimization.

Another important issue concerning data analysis in physics laboratories is the possibility of collaborating with other students in the same problem (team work). The possibility of sharing data space, models, plots and solutions, in addition to the possibility of continuing the work away from the laboratories, can also be an important point when choosing the computer solution to use for the data analysis. In advanced research work, the collaboration might include researchers across the world. The existence of web services as repositories of files that can be shared among different users help to overcome part of this problem. However, the use of commercial software might in some cases be a serious limitation to off-campus work and in some cases prevents the sharing of software solutions due to license limitations that might restrict the use of conventional software packages.

In this work, a web-based solution is presented (available at <http://fitteia.org>) for professional model fitting, graphing, data analysis and report writing that can also be used for teaching at both graduate and post-graduate levels¹. This solution spares users the overhead technical work associated with the least-squares minimization, data plotting and report writing, necessary when using conventional software packages, thus allowing users to focus on the

¹ Users can register, or reset their login password, using a valid email address. The system assigns a password that users can change afterwards. The user can be assigned one of two privilege levels: normal and privileged. The user has three running modes available: basic, advanced and expert (<http://fitteia.org>).

'art' of model fitting and parameter's estimation. The system allows for a considerable increase of the user's flexibility to perform his work and independence regarding his choice of both hardware and operating system. Personal computers (laptop or desktop), main-frame computers, tablets and smartphones are devices that can be used to fully access the system without losing performance.

2. Three theoretical models

2.1. Damped harmonic oscillator

One physics experiment that can be used to illustrate some aspects of advanced model fitting to experimental results concerns the description of motions in a spring–mass system consisting of a mass, m , suspended on a spring with elastic constant K . In static equilibrium, the length of the spring, $\Delta\ell$, depends on the equilibrium between the weight and on the spring's elastic force. Therefore,

$$\Delta\ell = \frac{g}{K}m, \quad (1)$$

where g is gravitational acceleration. The mass remains at its vertical position of equilibrium, z_0 .

In general terms, the equation of motion that describes the position of the mass with respect to its equilibrium position $Z = z - z_0$ is the well-known second degree differential equation

$$\frac{d^2Z(t)}{dt^2} + \frac{b}{m} \frac{dZ(t)}{dt} + \frac{K}{m}Z(t) = \frac{F_{\text{external}}(t)}{m}, \quad (2)$$

where b is the friction coefficient and $F_{\text{external}}(t)$ is a user-defined external force acting on the system. Equation (2) has several solutions. In the case of $F_{\text{external}}(t) = 0$, the equation of motion $Z(t)$ depends only on the system's parameters m , b and K . $\omega_0 = \sqrt{K/m}$ is the natural frequency of the system and $\lambda = b/(2m)$ is the damping coefficient. The particular solution of equation (2) when $\lambda < \omega_0$ is

$$Z(t) = Z_M e^{-\lambda t} \cos(\omega t + \phi), \quad (3)$$

where $\omega = \sqrt{\omega_0^2 - \lambda^2}$. ϕ and Z_M depend on the initial conditions of the motion. Starting from its initial position $Z(0)$ the mass m oscillates around the equilibrium position z_0 with period $T = 2\pi/\omega$ and decreasing amplitude.

If $\lambda \geq \omega_0$, the mass m has a damped aperiodic motion towards its final position z_0 starting from $Z(0)$.

In the particular case $F_{\text{external}}(t) = F_0 \cos(\omega_a t)$ and $\lambda < \omega_0$, it is easy to demonstrate that

$$Z(t) = Z_M e^{-\lambda t} \cos(\omega t + \phi) + A(\omega_a) \cos(\omega_a t + \alpha), \quad (4)$$

where

$$A(\omega_a) = \frac{F_0}{m\sqrt{(\omega_0^2 - \omega_a^2)^2 + 4\lambda^2\omega_a^2}} \quad (5)$$

and

$$\tan \alpha = \frac{2\lambda\omega_a}{\omega_0^2 - \omega_a^2}. \quad (6)$$

Obviously, other external forces will produce different permanent regimes after the initial natural motion of the mass vanishes.

2.2. X-ray diffraction

In advanced physics laboratories, students often have contact with x-ray diffraction experiments where it is possible to obtain structural information about condensed matter systems. In the case of soft matter systems that do not present long-range structural order, large-angle diffraction profiles can reveal details of the local molecular organization in the range from 10^{-10} to 5×10^{-9} m [4].

X-ray counters used in the experiments often produce profiles with a large number of points. The x-ray profiles where the scattered intensity is plotted as a function of scattered vector amplitude can, in some cases, be described by a sum of the Gaussian profiles

$$I(q) = I_0 + \sum_{k=0}^N a_k e^{-(q-q_k)^2/b_k^2}, \quad (7)$$

where I_0 is the background contribution, q_k , a_k and b_k are the wave vectors, amplitudes and line widths of each characteristic structural distance numbered with index k , respectively.

2.3. NMR relaxation

Condensed matter physics students, at the PhD level, occasionally have to obtain molecular dynamics information using NMR relaxometry (NMRD) [5, 6]. The proton spin–lattice relaxation rate, T_1^{-1} , as a function of frequency and temperature might present features that can be interpreted in terms of relaxation models that are associated with different types of molecular motions [5–8]. Molecular rotations and reorientations (R), translational self-diffusion (SD) and slow collective motions are common examples considered when analysing NMRD results obtained for isotropic liquids and liquid crystals [6]. In the case of an isotropic phase of a liquid crystal of ellipsoid molecules T_1^{-1} can be approximated by

$$T_1^{-1} = (T_1^{-1})_R + (T_1^{-1})_{SD} + (T_1^{-1})_{OPF} \quad (8)$$

with

$$(T_1^{-1})_R = A_R \left[\frac{\tau_S}{1 + \omega^2 \tau_S^2} + \frac{4\tau_S}{1 + 4\omega^2 \tau_S^2} + \frac{\tau_L}{1 + \omega^2 \tau_L^2} + \frac{4\tau_L}{1 + 4\omega^2 \tau_L^2} \right], \quad (9)$$

$$(T_1^{-1})_{SD} = \frac{9}{8} \gamma^4 \hbar^2 \left(\frac{\mu_0}{4\pi} \right)^2 \frac{n\tau_D}{d^3} [T(\alpha, \omega\tau_D) + 4T(\alpha, 2\omega\tau_D)] \quad (10)$$

and

$$(T_1^{-1})_{OPF} = \frac{A_{OPF}}{\omega^{1/2}} \int_{\omega_{Cl}/\omega}^{\omega_{Ch}/\omega} \frac{\sqrt{x}}{1 + (x + \omega_0/\omega)^2} dx. \quad (11)$$

In this relaxation model (equation (8)), τ_L and τ_S are correlation times associated with molecular rotations/reorientations along the long molecular axis and around a perpendicular axis, and τ_D is a correlation time associated with the translational diffusion displacements. n is the density of spins and d is the distance of the closest lateral approach between molecules [9]. ω_{Cl} and ω_0 are low cut-off frequencies, and ω_{Ch} is a high cut-off frequency associated with fluctuations of order (OPF) observed close to the isotropic–nematic transition. A_R and A_{OPF} are pre-factors that depend on additional physical parameters of the liquid crystal compound (see the electronic supplementary information (ESI), available from stacks.iop.org/EJP/35/015017/mmedia). τ_S , τ_L and τ_D are temperature dependent and are usually described by Arrhenius laws $\tau_S = \tau_{S_{ref}} \exp(E_S(T^{-1} - T_{ref}^{-1})/8.31)$, $\tau_L = \tau_{L_{ref}} \exp(E_L(T^{-1} - T_{ref}^{-1})/8.31)$ and $\tau_D = \tau_{D_{ref}} \exp(E_D(T^{-1} - T_{ref}^{-1})/8.31)$, respectively. T_{ref} is a reference temperature and E_S , E_L and E_D are activation energies.

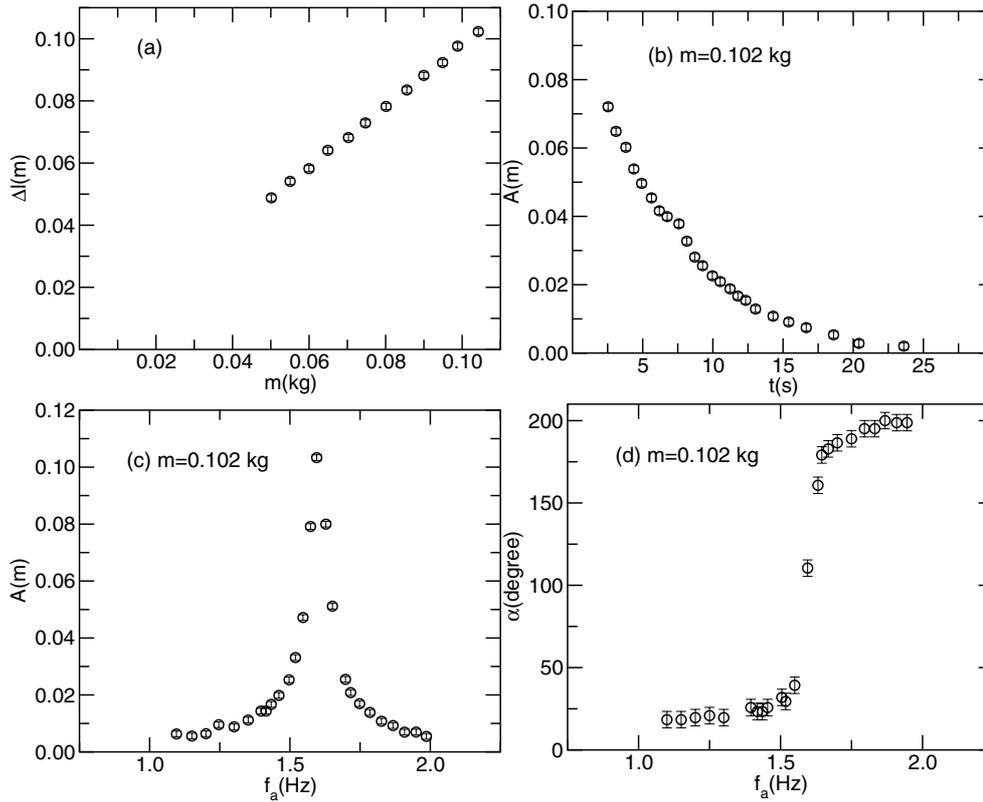


Figure 1. Experimental results obtained for a spring–mass system [10], as explained in the text. (a) The spring’s length as a function of the suspended mass; (b) amplitude of oscillation measured at time nT , $n = 0, 1, 2, 3, \dots$, where T is the oscillation period; (c) amplitude of oscillation measured as a function of the frequency of the harmonic applied force; (d) phase difference between the mass motion’s amplitude and the applied force as a function of the oscillation frequency of the force.

3. Experimental results

3.1. Damped harmonic oscillator

Using a damped harmonic oscillator experimental setup [10], which includes a spring with unknown elastic constant K , a mass $m = 0.102$ kg, and a unknown friction coefficient b , it is possible to obtain four sets of experimental results corresponding to the motions of the harmonic oscillator, described by the model equations (1), (3), (5) and (6). These experimental results are presented in figures 1(a), (b), (c) and (d), respectively. The position of the mass was determined with an uncertainty of $\pm 10^{-3}$ m. The frequency of the oscillation of the external force acting on the system was controlled with a resolution of ± 0.01 Hz. The phase difference between oscillations of the mass and the external force was measured with an uncertainty of 5° . The results of figure 1(b) were obtained regularly with a sampling time equal to the period of oscillation.

3.2. X-ray diffraction

In figure 2 is presented the x-ray diffraction profile obtained for a liquid crystal compound using $\text{Cu } K_\alpha$ radiation with a wave length $\lambda = 1.54 \text{ \AA}$.

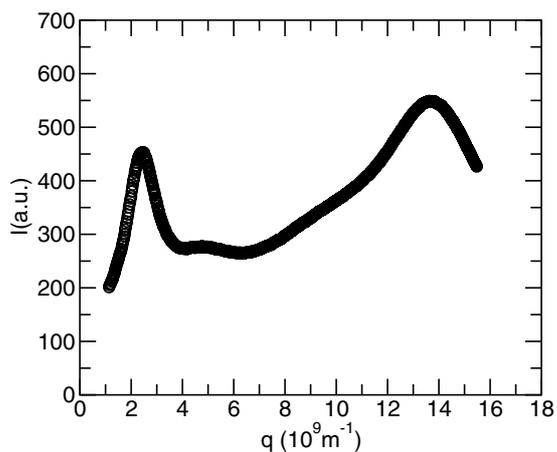


Figure 2. X-ray diffraction profile obtained for a liquid crystal compound.

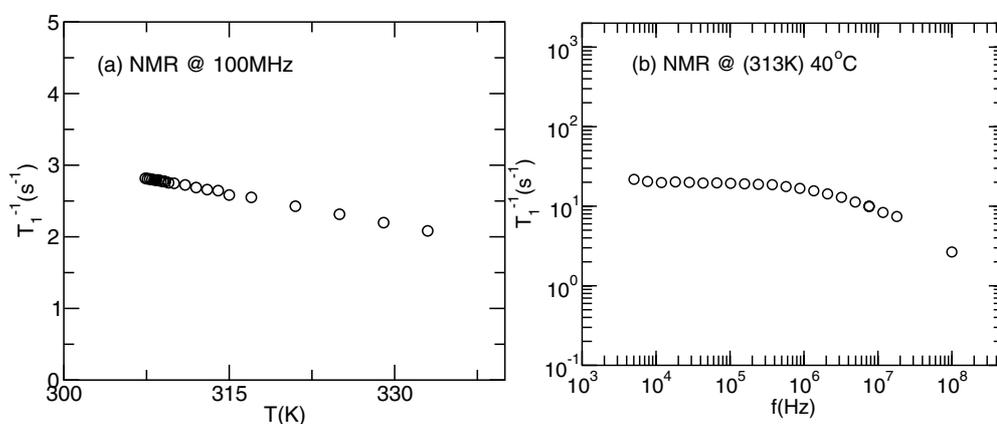


Figure 3. Experimental ¹H spin–lattice relaxation rates: (a) results obtained at a fixed frequency for a Larmor resonance frequency of 100 MHz; (b) results obtained as a function of the Larmor frequency for a fixed temperature $T = 40$ °C.

3.3. NMR relaxation

In figure 3 are presented the hydrogen proton spin–lattice relaxation rate experimental results as a function of temperature for a fixed Larmor frequency of 100 MHz, and as a function of frequency for a fixed temperature. The results for frequencies below 20 MHz were obtained using fast field cycling NMR relaxometry [11–13]. For the measurements at 100 MHz, standard NMR relaxometry techniques were used [7].

4. Model fits

The analysis of experimental results presented in figures 1–3 with respect to the physics models previously introduced requires the fitting of the models' equations to the experimental

results in order to obtain the physical parameters relevant for each system. One common way to obtain the best fit is by the minimization of the least squares

$$\chi^2(p_1, \dots, p_m) = \sum_{k=1}^N \frac{(y_k^e - y_k^f(p_1, \dots, p_m, x_k^e))^2}{\sigma_k^{e2}}, \quad (12)$$

where $y_k^f(p_1, \dots, p_m, x_k^e)$ is the value of the model equation calculated for the value x_k^e , and the model parameters p_k with $k = 1, \dots, M$, $(x_k^e, y_k^e \pm \sigma_k^e)$ are the experimental points $1, \dots, N$, with σ_k^e being the experimental uncertainty of y^e . In most experiments, the uncertainty of the independent variable is controlled in order to reduce its influence on the measured observable. The case where the uncertainties of both x and y are important will be addressed below.

A rule of thumb, in normal practice, is that a ‘moderately’ good fit is obtained when the value of $\chi^2(p_1, \dots, p_m) \simeq N - M$ ($N - M$ being the number of degrees of freedom) is obtained after the minimization [14].

The χ^2 minimization considering the model equations (1), (3)–(8) presents different types of challenges. Technically, the fit of equation (1) to the data of figure 1(a) is the most simple one and can be considered a trivial case. The fits of the other model equations require more sophisticated numerical methods to perform the least-squares minimization and some user skills in the use of the available software packages.

On *fitteia*’s website (see footnote 1), users can perform both simple and advanced model fits to experimental data with different degrees of complexity using a common interface that drastically reduces the user’s overhead work required to perform identical data analysis using other software packages. The *fitter* module available presents three distinct sections: (1) data, (2) plot parameters, and (3) function and parameters. In the data section the experimental points are introduced as a table with three columns where the values of ‘ x^e ’, ‘ y^e ’ and ‘ σ^e ’ (one point $(x_k^e, y_k^e \pm \sigma_k^e)$ per line) are separated by tab characters or white spaces. Copy and paste actions from spread sheet tables or text files can help filling up the data table. In the plot parameters section, the plots can be set according to the data being analysed. The model equation $y = f(p_1, \dots, p_m, x_k^e)$ is introduced in the one-line text box in the function and parameters section, using the syntax of language C. Both dependent and independent variable names (y and x) and the parameter names (p_1, \dots, p_M) have to be declared (additional information is available in the ESI, available from stacks.iop.org/EJP/35/015017/mmedia).

The *fitteia*’s *fitter* module core is a powerful minimization routine from the CERN library called MINUIT [15]. MINUIT contains three different minimization methods SCAn, SIMplex and Mlgrad. A sequential use of these methods and a call to MINUIT MINOS method to compute the uncertainties of the fitting parameters is a powerful strategy to find the best absolute minimum of equation (12). As with all minimization routines the user should provide an ‘educated guess’ of the initial values of the fitting parameters to minimize the computation time and to avoid stopping the minimization on a local minimum in the parameters’ space. Nevertheless, *fitteia* has proven to be quite robust with respect to this issue and all parameters are initialized with a finite value to prevent some unnecessary mistakes.

4.1. Damped harmonic oscillator

The first step in the analysis of the experimental results obtained for the damped harmonic oscillator is to obtain the best numerical model fit to each data set of the experimental results separately.

Table 1. Model parameters obtained from the best model fits to the experimental results explained in the text. $m = 0.102$ kg.

Data Figure	$\Delta\ell(m)$ 1(a)	$A(t)$ 1(b)	$A(f_a)$ 1(c)	$\alpha(f_a)$ 1(d)	All 4
K (N m^{-1})	9.9 ± 0.2	–	10.359 ± 0.006	10.27 ± 0.03	10.4 ± 0.1
d_0 (10^{-3}m)	-0.6 ± 1.3	–	–	–	2.7 ± 0.3
λ (s^{-1})	–	0.139 ± 0.004	0.163 ± 0.003	0.12 ± 0.02	0.142 ± 0.001
Z_0 (m)	–	0.103 ± 0.002	–	–	0.105 ± 0.001
Z_∞ (m)	–	0.003 ± 0.001	–	–	0.002 ± 0.001
F_0 (N)	–	–	0.037 ± 0.001	–	0.033 ± 0.001
a_0	–	–	–	0.32 ± 0.03	0.34 ± 0.02

In the case of the spring's elongation as a function of the suspended mass (figure 1(a)) the model used was $\Delta\ell = (g/K)m + d_0$. d_0 was introduced in equation (1) to take into account possible offsets in the experiment, when measuring $\Delta\ell$.

In the case of the data set corresponding to the free damped oscillations, equation (3) can be simplified when the experimental measurements are made regularly at times $t = t_0 + nT$, where T is the period of oscillation, $n = 0, \dots, N$, $Z(t_0 + nT) = (Z_0 - Z_\infty)e^{-\lambda nT} + Z_\infty$. Z_∞ should be null, but it can be introduced to take into account possible offset deviations from the initial equilibrium value.

In the case of forced oscillations, the amplitude of oscillation $A(\omega_a)$ given by equation (5) can be written in terms of $f_a = \omega_a/(2\pi)$, $A(f_a) = (F_0/m)[(K/m - 4\pi^2 f_a^2)^2 + 16\pi^2 \lambda^2 f_a^2]^{-1/2}$. As for the phase difference between the external force oscillations and the mass oscillations α given by equation (6) some attention is required since $\arctan(x)$ changes by a value of π when the argument changes sign. Therefore, using the (? :) conditional operator equation (6) can be written as $\alpha = (f_a < \sqrt{K/(4\pi^2 m)}) ? \arctan(4\pi \lambda f_a / (K/m - 4\pi^2 f_a^2)) + \alpha_0 : \arctan(4\pi \lambda f_a / (K/m - 4\pi^2 f_a^2)) + \alpha_0 + \pi$, α_0 is an offset angle.

The fitting parameters obtained from each best fit to the experimental results are presented in table 1. The fitting PDF reports produced with *fitteia's report* module can be found in the ESI (available from stacks.iop.org/EJP/35/015017/mmedia).

When analysing the fitting parameters in table 1, it becomes obvious that the independent fits of $\Delta\ell(m)$, $A(t)$, $A(f_a)$ and $\alpha(f_a)$ give different results for the physical parameters of the system K and λ . The differences depend on the influence of the experimental conditions and experimental uncertainties of the experimental results. A discussion is required in order to conclude about the values of K and λ for the system.

The analysis of the experimental results in figure 1 can be improved if a detailed analysis of the experimental uncertainties is made and if the analysis of the experimental results is made by performing a fit that takes into account all data and a global least-squares minimum target.

In the case of the forced oscillations, it should be mentioned that the experimental uncertainty of both the amplitude of oscillation $A(f_a)$ and the phase difference α should be estimated taking into account the stability of the external force. Indeed, if the external frequency varies within the range $f_a \pm \sigma_{f_a}$, then

$$\sigma_k^{e2} \simeq \sigma_k^{y2} + \left(\frac{dy}{df_a}\right)_k^2 \sigma_{f_a}^2. \quad (13)$$

From equation (13), it is clear that $\sigma_k^e \simeq \sigma_k^y$ unless $\sigma_k^y \sim \sigma_{f_a} |dy/df_a|_k$, which means that the uncertainty of f_a is relevant, particularly when $|dy/df_a|_k$ is large.

The derivative dy/df_a must be calculated numerically [16]; otherwise, the use of the fitting model to compute the propagated uncertainty would introduce a bias in the minimization procedure. The *fitter* module of *fitteia* includes a calculator text box in the data section where users can introduce a sequence of operations involving the columns of the data table. Removing, adding and reordering of columns become quite easy, in addition to calculations involving columns. Therefore, starting from the columns in the data table (e.g. corresponding to ' f_a ', ' A ' and ' σ^A '), the calculations associated with equation (13) can be done by executing the following *fitteia* instructions:

```
c1 c2 c3 0.01
c1 c2 c3 dc2/dc1 c4
c1 c2 sqrt(c3*c3+c4*c4*c5*c5)
```

Each line of the above sequence transforms the data table into another. The first line transforms the three-column table into a four-column table keeping columns 1–3 and adding a fourth column with a constant value '0.01' corresponding to σ_{f_a} . The second line in the sequence transforms the four-column table by moving column four to the fifth position and introducing in the fourth position the numerical derivative of column two with respect to column one. Finally, a three-column table is obtained with columns one and two equal to the original ones and putting in third position a column with the values of σ_k^e calculated according to equation (13). The effect of the correction on the experimental error bars can be observed in figures 4(c) and (d).

The least-squares minimization, with a global minimum target that takes into account all data and the different model equations, is fairly simple using *fitteia*. This type of problem is equivalent to a multidimensional fit. *fitteia*'s 'Expert' user mode presents an interface where a second independent variable can be introduced and used as a flag in nested conditional tests that help to apply the proper physical model to the corresponding experimental data. Naming the second independent variable 'flag', we can use the statement 'flag = 1' for $\Delta\ell(m)$, 'flag = 2' for $A(t)$, 'flag = 3' for $A(f_a)$ and 'flag = 4' for $\alpha(f_a)$. In the data text box, the experimental results are introduced in blocks separated by reserved comment lines '# DATA flag = 1', etc.

Defining x and y as the generic independent and dependent variables, respectively, the generic model equation can be written using the (? :) conditional operator in a nested sequence.

```
y=(flag==1) ? g/K*x+d0 :
(flag==2) ? (Zt0-Zinf)*exp(-1*x)+Zinf :
(flag==3) ? F0/m/sqrt(pow(K/m-4*pi*pi*x*x, 2.0)+16*pi*pi*1*1*x*x) :
(x>sqrt(K/m)/(2*pi)) ? (atan(2*1*2*pi*x/(K/m-4*pi*pi*x*x))+pi+a0)*180/pi :
(atan(2*1*2*pi*x/(K/m-4*pi*pi*x*x))+a0)*180/pi
```

In the case of 'flag = 4' an additional test is required to choose the proper expression on the calculus of $\alpha(f_a)$ according to the condition $\omega > \omega_0$. $\exp(x) \equiv e^x$, $\text{pow}(x, p) \equiv x^p$, $\text{sqrt}(x) \equiv \sqrt{x}$ and $\text{atan}(x) \equiv \arctan(x)$.

In this way, the minimization of equation (12) takes into account all experimental results and selectively considers each model equation according to the corresponding data set.

$$\chi^2(K, \lambda, d_0, Z_{t_0}, Z_{t_\infty}, F_0, a_0) = \sum_{k=1}^{N_1} \frac{(\Delta\ell_k^e - \Delta\ell_k^t(K, d_0, m_k))^2}{\sigma_k^{e2}} + \sum_{k=1}^{N_2} \frac{(A_k^e - A_k^t(\lambda, Z_{t_0}, Z_{t_\infty}, t_k))^2}{\sigma_k^{e2}} + \sum_{k=1}^{N_3} \frac{(A_k^e - A_k^t(K, \lambda, F_0, f_{a_k}))^2}{\sigma_k^{e2}} + \sum_{k=1}^{N_4} \frac{(\alpha_k^e - \alpha_k^t(K, \lambda, F_0, f_{a_k}, a_0))^2}{\sigma_k^{e2}}. \quad (14)$$

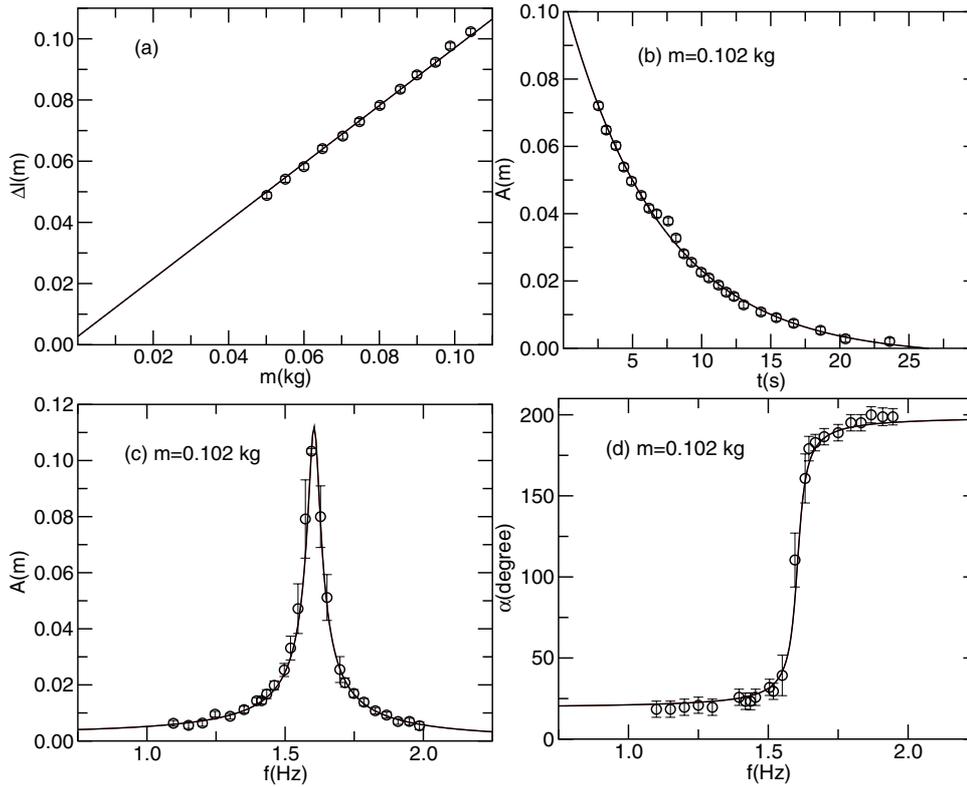


Figure 4. Experimental results obtained for a spring–mass system [10], taking into account the uncertainties of both x and y in the case of plots (c) and (d). (a) The spring’s length as a function of the suspended mass; (b) amplitude of oscillation measured at time nT , $n = 0, 1, 2, 3, \dots$, where T is the oscillation period; (c) amplitude of oscillation measured as a function of frequency of the harmonic applied force; (d) phase difference between the mass motion’s amplitude and the applied force as a function of oscillation frequency of the force.

As it happened that K and λ are present in different model equations, the global minimization searches for the values for these parameters that minimize χ^2 and consequently minimize all terms in equation (14) in a consistent way.

fitteia presents a simple interface that helps users to prepare the figures with the plots of the fitting curves according to the details of each data set.

In figure 4 are shown the best fits obtained using the above described global minimization procedure (additional information can be found in the ESI, available from stacks.iop.org/EJP/35/015017/mmedia). As can be observed in figures 4(c) and (d), the consequence of propagating the uncertainty in the frequency f_a to variables $A(f_a)$ and $\phi(f_a)$ is to obtain larger error bars for the experimental points where dy/dx is larger.

The values of the fitting parameters for this fit are presented in column ‘all’ in table 1. A discussion about the advantages of both minimization approaches can be made but the important point here is to present the two strategies.

4.2. X-ray diffraction

The data analysis of the experimental results of the x-ray diffraction experimental presented in figure 2 presents a different type of challenge with respect to the previous example. First, the

Table 2. Fitting parameters (as presented in the ESI fit report generated by *fitteia*, available from stacks.iop.org/EJP/35/015017/mmedia) obtained from the fit of model equation (7) to the experimental results of figure 2. $\chi^2 = 8.78204$.

$a = 153.21 \pm 5.4086$	$a2 = 200 \pm 29.723$
$a0 = 252.52 \pm 15.052$	$b2 = 4 \pm 0.12931$
$b0 = 1 \pm 0.014206$	$q2 = 10.419 \pm 0.12791$
$q0 = 2.4294 \pm 0.040112$	$a3 = 300 \pm 9.6683$
$a1 = 93.836 \pm 7.8405$	$b3 = 2.3904 \pm 0.14258$
$b1 = 1.7183 \pm 0.29467$	$q3 = 14.164 \pm 0.080498$
$q1 = 4.7695 \pm 0.14836$	

data set has 685 points, more than eight times the number of points of the previous example; second, we know in advance that the model fit is composed of a sum of several Gaussian curves and besides being necessary to obtain the curve that represents the sum there are clear advantages in being able to see each contribution separately.

The number of experimental points is not a limitation as far as the computation time remains acceptable, but even for a moderate number of points the fit of a nonlinear function with ten or more independent fitting parameters can be very CPU demanding. Therefore, a good strategy to analyse a large experimental data set is to choose a subset of the experimental points larger than the set of fitting parameters in order to obtain a first fit and then use the obtained fitting parameters as starting guess values for the final fit with all data. *fitteia*'s fitter module helps to achieve this by providing two conditional pre-processing commands, '# fitif (condition), (sampling step)' and '# plotif (condition), (sampling step)', that can be introduced in the DATA text box as reserved comment lines. The first command is considered when fitting, the second command is considered only when plotting the data and the fitting curves using the model parameter values (e.g. simulating the dependent variable as a function of the independent variable using the model equation). '# fitif 1, 10' can be used to select a subset of the experimental data where one point every ten points is considered for the fit. The condition here is always '1' so the test is applied to all data points.

fitteia's fitter module also helps users to generate curves that can be plotted together with the curve that corresponds to the model equation. If the q and y are the independent and dependent variables the scattered x-ray intensity can be modelled by equation (7) with four Gaussian contributions and an offset constant.

$$y = a + a0*\exp(-\text{pow}((q-q0)/b0,2.0)) + a1*\exp(-\text{pow}((q-q1)/b1,2.0)) + a2*\exp(-\text{pow}((q-q2)/b2,2.0)) + a3*\exp(-\text{pow}((q-q3)/b3,2.0))$$

If the plus symbol '+' is replaced by '\+', in the above expression, then, after compilation, the *fitteia*'s fitter interface is modified and allows for the selection of the line type, color and label of the curves associated with the terms added with each '\+'. In the following case

$$y = a \ \backslash+ \ a0*\exp(-\text{pow}((q-q0)/b0,2.0)) \ \backslash+ \ a1*\exp(-\text{pow}((q-q1)/b1,2.0)) \ \backslash+ \ a2*\exp(-\text{pow}((q-q2)/b2,2.0)) \ \backslash+ \ a3*\exp(-\text{pow}((q-q3)/b3,2.0))$$

each term will be represented by its corresponding plot line according to the values of a_k , b_k and q_k .

In figure 5 is presented the best fit obtained using both the pre-processing command '# fitif 1, 10' and the contributions corresponding to all the terms in the model equation.

There is a clear advantage in observing separately the contributions to total scattered intensity since, in addition to the best-fitting curves, simulation plots can put into evidence the effects of the model parameters separately. In table 2 are presented the values of the model

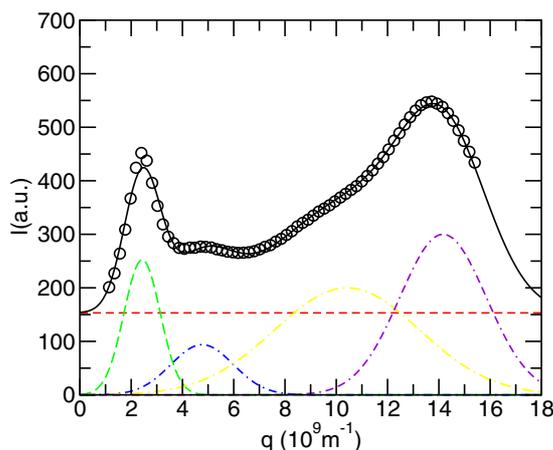


Figure 5. X-ray diffraction results fitted with model equation (7). The fit was done considering a subset of the experimental data (as an example of *fitteia*'s features), as explained in the text.

parameters corresponding to the best fit. They are presented as they are obtained from the *fitteia* report without taking into account the coherent correspondence between the significant digits in both values and uncertainties.

4.3. NMR relaxation

The analysis of the NMR data of figure 3 using model equation (8) with contributions (10) and (11) presents in addition to the technical problems referred to previously an additional technical challenge. In fact, the contribution of equation (11) must be obtained numerically and the integral has to be calculated every time the function is called. In addition, some of the functions used are too cumbersome to be introduced using a simple line of text. One solution for these problems is to ask the *fitteia* system's manager to include the necessary functions in the user functions library. Another solution is to use *fitteia* with an increased level of privilege. In the latter, users are allowed to program their own functions library and freely program the necessary code for their needs.

In the case of equation (8), all functions are already available in the system. The equation to be fitted to the spin–lattice relaxation data, T_1^{-1} , taking into account both its dependence on the temperature, T , and on the Larmor frequency, f , can be written taking advantage of the previous examples. In *fitteia*'s expert mode, it is possible to define a model that considers different independent variables (in the case of the spring–mass system three independent variables were considered). For the NMR data we can consider $z \equiv T_1^{-1}(T, f)$ and define x and y as independent variables. The meaning of x and y is either frequency or temperature depending on the context. In any case x refers to the first column and y takes the value assigned in the data section using the statement format '#DATA y=...'. The experimental data is introduced in two DATA blocks using the following arrangement:

```
# DATA y=100e6
333      2.08333      0.104167
329      2.1978      0.10989
.        .          .
.        .          .
```

```
# DATA y=313
      1e+08      2.65957      0.132978
      1.79991e+07      7.42497      0.371249
      :          :          :
      :          :          :
```

The frequency is always larger than the temperature and is never lower than 1 kHz. It is clear that when $y > 500$, $y \equiv \text{temperature}$ and x is the frequency and when $y < 500$, $y \equiv \text{frequency}$ and x is the temperature. In this way additional DATA can be introduced for different temperatures and different frequency dependences (see Sebastião *et al* [17]). Assuming that $(T_1^{-1})_R$ is the sum of two ‘BPP’ contributions ($(T_1^{-1})_R \equiv \text{BPP}(\tau_S) + \text{BPP}(\tau_L)$), $(T_1^{-1})_{SD}$ is calculated using ‘Torrey1’ function ($(T_1^{-1})_{SD} \equiv \text{Torrey1}$), and $(T_1^{-1})_{OPF} \equiv \text{OPF}$ (additional information is available in the ESI, available from stacks.iop.org/EJP/35/015017/mmedia)

```
z=(
  (y>500.0) ? BPP(y,Arot,tauL*exp(EL/8.31*(1.0/x-1.0/Tref))) :
              BPP(x,Arot,tauL*exp(EL/8.31*(1.0/y-1.0/Tref)))
) + (
  (y>500.0) ? BPP(y,Arot,tauS*exp(ES/8.31*(1.0/x-1.0/Tref))) :
              BPP(x,Arot,tauS*exp(ES/8.31*(1.0/y-1.0/Tref)))
)
\+
(
  (y>500.0) ?
  Torrey1(y,d,r,n,r*r*1e-20/(6*Dref*exp(-ED/8.31*(1.0/x-1.0/Tref)))) :
  Torrey1(x,d,r,n,r*r*1e-20/(6*Dref*exp(-ED/8.31*(1.0/y-1.0/Tref))))
)
\+
(
  (y>500.0) ? OPF(y,Aopf,f0,fcM,fcM,np) :
              OPF(x,Aopf,f0,fcM,fcM,np)
)
)
```

A close analysis of the above expression shows that, according to the number of ‘\+’ symbols, three contributing curves will be plotted separately. The first curve will correspond to the sum of two ‘BPP’ functions, the second will be ‘Torrey1’ and ‘OPF’ the third. This is a second example of the independent plot of the model contributions presented previously for the x-ray fit. In addition, it can be observed that each contribution includes a conditional test ‘ $y > 500$ ’ that changes the meaning of x and y . ‘ $\text{BPP}(f, A, \tau)$ ’ is a function of frequency f , a pre-factor A , and a characteristic time $\tau(T)$. Since $\tau(T)$ depends on the temperature T , ‘BPP’ depends on a frequency and a temperature. According to the data coding used x and y alternate meanings and if $y > 500$ it is for sure a frequency, as the experimental results were obtained for temperatures below 500 K.

The model parameters obtained for the best fit are presented in table 3 and the model-fitting curves are presented in figure 6.

The plots presented in figures 1–6 were produced using the *fitter* and *plotter* interface modules available on <http://fitteia.org> [18]. In the ESI (available from stacks.iop.org/EJP/35/015017/mmedia) additional information can be found concerning the use of *fitteia*’s web service.

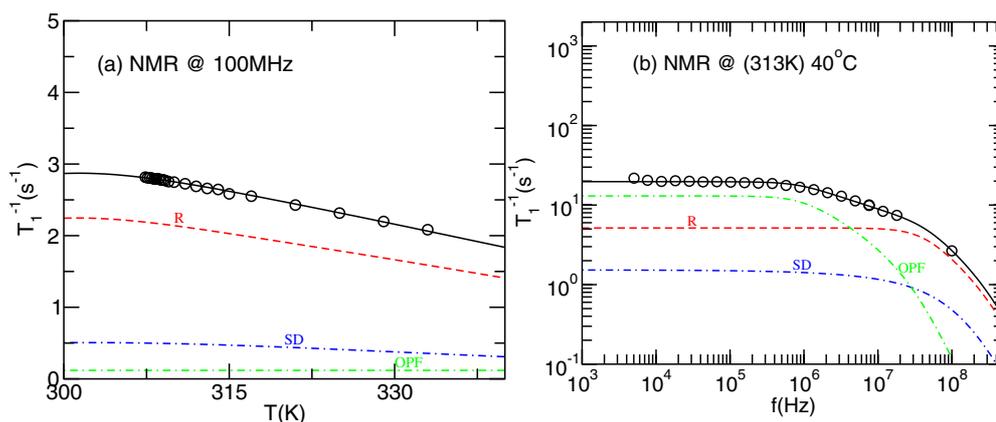


Figure 6. Model fit to the experimental ^1H spin-lattice relaxation rates: (a) fitting curves plotted as a function of temperature, for the fixed Larmor resonance frequency of 100 MHz; (b) fitting curves plotted as a function of the Larmor frequency, for the fixed temperature $T = 40^\circ\text{C}$ as explained in the text.

Table 3. Fitting parameters (as presented in the ESI fit report generated by *fitteia*, available from stacks.iop.org/EJP/35/015017/mmedia) obtained from the fit of model equation (8) to the experimental results as explained in the text. $\chi^2[313\text{ K}] = 8.819\,02$, $\chi^2[100\text{ MHz}] = 0.184\,23$, $\chi_r^2 = 9.003\,25$.

$A_{rot} = 5.7688 \times 10^{+08} \pm 8.45 \times 10^{+06}$	$D_{ref} = 5.44 \times 10^{-11}$ (fixed)
$\tau_{auS} = 1.45 \times 10^{-09}$ (fixed)	$ED = 32\,800$ (fixed)
$ES = 294\,12$ (fixed)	$T_{ref} = 313$ (fixed)
$\tau_{auL} = 3.36 \times 10^{-10}$ (fixed)	$A_{opf} = 17438 \pm 1017.9$
$EL = 468\,28 \pm 6860.6$	$f_0 = 1 \times 10^{+05}$ (fixed)
$d = 5$ (fixed)	$fcM = 4.25 \times 10^{+07}$ (fixed)
$n = 4.59 \times 10^{+22}$ (fixed)	$fcM = 6.8381 \times 10^{+05} \pm 998\,54$
$r = 4$ (fixed)	$np = 10$ (fixed)

5. Conclusions

It is shown that when analysing experimental results in terms of theoretical models the concept of best fit is not only related to the numerical quality of the model fit but also connected with the physical coherence between the model parameters obtained. This is particularly true in the case of experimental observables that depend on different experimental variables known prior to the experiment and that can be used to better describe the studied system. Multidimensional fits require not only additional computation skills but also a method to easily understand the details of the minimization process. A web-based interface to a software package is presented that allows users to perform online model fits, data plots and report writing using models with different degrees of complexity. The details of the methods used to fit models to experimental data with one, two, \dots , n dimensions were presented here. The methods used can be implemented in a relatively simple way using the interface but they can also be used to develop solutions using other numerical software packages. One important difference between the presented web-solution for the model fitting with respect to other solutions is the possibility of sharing working folders with other users involved in the project, no matter where they are physically or what type of computer system they use. The efficiency thus achieved is worth

noticing. Both graduate and post-graduate students find this solution quite useful during the learning stage, where they obtain the basic skills to independently perform model analysis and fitting to experimental data. Later, former students find the system also useful to perform more complex data analysis during their MSc and/or PhD research work.

Acknowledgments

The author wishes to thank *fitteia*'s users for the challenge of further improving the system and for the inclusion of additional features that simplify the process of model fitting to experimental results. Thanks to Instituto Superior Técnico and Universidade de Lisboa for hosting the two computers that presently run the *fitteia* server. Thanks also to Alberto Ferreira and João Godinho, former MEFT IST-ULisboa students and *fitteia* users, for their commitment to the development of a user friendly translator to C, included in the *fitteia* fitter interface since 2011.

References

- [1] Montgomery D C, Peck E A and Vining G 2001 *Introduction to Linear Regression Analysis (Wiley Series in Probability and Statistics)* 4th edn (New York: Wiley)
- [2] Peterlin P 2010 Data analysis and graphing in an introductory physics laboratory: spreadsheet versus statistics suite *Eur. J. Phys.* **31** 919–31
- [3] Fernandes J C, Ferraz A and Rogalski M S 2010 Computer-assisted experiments with oscillatory circuits *Eur. J. Phys.* **31** 299–306
- [4] Filip D, Cruz C, Sebastião P J, Ribeiro A C, Vilfan M, Meyer T, Kouwer P H J and Mehl G H 2007 Structure and molecular dynamics of the mesophases exhibited by an organosiloxane tetrapode with strong polar terminal groups *Phys. Rev. E* **75** 11704
- [5] Dong R Y 1997 *Nuclear Magnetic Resonance of Liquid Crystals* (New York: Springer)
- [6] Sebastião P J, Cruz C and Ribeiro A C 2009 Nuclear magnetic resonance spectroscopy of liquid crystals *Advances in Proton NMR Relaxometry in Thermotropic Liquid Crystals* ed R Y Dong (Singapore: World Scientific) pp 129–67 chapter 5
- [7] Farrar T C and Becker E D 1971 *Pulse and Fourier Transform NMR* (New York: Academic)
- [8] Abragam A 1961 *The Principles of Nuclear Magnetism* (Oxford: Clarendon)
- [9] Torrey H C 1953 Nuclear spin relaxation by translational diffusion *Phys. Rev.* **92** 962–9
- [10] PASCO Scientific Pasco model 9201a damped harmonic oscillator
- [11] Noack F 1986 NMR field-cycling spectroscopy—principles and applications *Prog. Nucl. Magn. Reson. Spectrosc.* **18** 171–276
- [12] Kimmich R and Ansaldo E 2004 Field-cycling NMR relaxometry *Prog. Nucl. Magn. Reson. Spectrosc.* **44** 257–320
- [13] Sousa D M, Marques G D, Cascais J M and Sebastião P J 2010 Desktop fast-field cycling nuclear magnetic resonance relaxometer *Solid State Nucl. Magn. Reson.* **38** 36–43
- [14] Press W H, Teukolsky S A, Vetterling W T and Flannery B P 1992 *Numerical Recipes in C: The Art of Scientific Computing* 2nd edn (New York: Cambridge University Press)
- [15] CERN 2000 Minuit <http://wwwasdoc.web.cern.ch/wwwasdoc/minuit/minmain.html>
- [16] Holoborodko P 2008 Smooth noise robust differentiators www.holoborodko.com/pavel/numerical-methods/numerical-derivative/smooth-low-noise-differentiators/
- [17] Sebastião P J, Gradišek A, Pinto L F V, Apih T, Godinho M H and Vilfan M 2011 Fast field-cycling NMR relaxometry study of chiral and nonchiral nematic liquid crystals *J. Phys. Chem. B* **115** 14348–58
- [18] Sebastião P 2009 <http://fitteia.org> All users can have access to the plots and fits presented in this work in *fitteia* folder Tutorial 1 under names that start with prefixes 'Basic-EJP', 'Expert-EJP' and 'Plot-EJP'