Article

Two-Level Scheme for Identification of the Relaxation Time Spectrum Using Stress Relaxation Test Data with the Optimal Choice of the Time-Scale Factor

Anna Stankiewicz

Department of Technology Fundamentals, Faculty of Production Engineering,
University of Life Sciences in Lublin, 20-612 Lublin, Poland; anna.m.stankiewicz@gmail.com

Abstract: The viscoelastic relaxation spectrum is vital for constitutive models and for insight into the mechanical properties of materials, since, from the relaxation spectrum, other material functions used to describe rheological properties can be uniquely determined. The spectrum is not directly accessible via measurement and must be recovered from relaxation stress or oscillatory shear data. This paper deals with the problem of the recovery of the relaxation time spectrum of linear viscoelastic material from discrete-time noise-corrupted measurements of a relaxation modulus obtained in the stress relaxation test. A two-level identification scheme is proposed. In the lower level, the regularized least-square identification combined with generalized cross-validation is used to find the optimal model with an arbitrary time-scale factor. Next, in the upper level, the optimal time-scale factor is determined to provide the best fit of the relaxation modulus to experiment data. The relaxation time spectrum is approximated by a finite series of power–exponential basis functions. The related model of the relaxation modulus is proved to be given by compact analytical formulas as the products of power of time and the modified Bessel functions of the second kind. The proposed approach merges the technique of an expansion of a function into a series of independent basis functions with the least-squares regularized identification and the optimal choice of the time-scale factor. Optimality conditions, approximation error, convergence, noise robustness and model smoothness are studied analytically. Applicability ranges are numerically examined. These studies have proved that using a developed model and algorithm, it is possible to determine the relaxation spectrum model for a wide class of viscoelastic materials. The model is smoothed and noise robust; small model errors are obtained for the optimal time-scale factors. The complete scheme of the hierarchical computations is outlined, which can be easily implemented in available computing environments.

Keywords: viscoelasticity; relaxation time spectrum; linear relaxation modulus; hierarchical identification algorithm; regularized least-squares identification; time-scale factor optimal selection; modified Bessel functions of the second kind; singular value decomposition

1. Introduction

Numerous mathematical rheological models are used to describe the mechanical properties of viscoelastic materials [1–3]. The Maxwell and Kelvin–Voight models are, probably, the best-known rheological models. However, deeper insight into the complex behavior of viscoelastic materials is provided by the relaxation spectrum [2,4,5]. The relaxation spectrum is vital for constitutive models and for insight into the properties of a material, since, from the relaxation spectrum, other material functions used to describe rheological properties can be uniquely determined [4,6,7]. It is commonly used to describe, analyze, compare, and improve the mechanical properties of polymers [4,8,9], concrete [10], asphalt [11], rubber [12], wood [13], glass [14], dough [15], polymeric textile materials [16], geophysics [17] and biological materials [18]. The spectrum is not directly measurable; therefore, it must be recovered from oscillatory shear or stress relaxation data [2,5]. A number of different methods and algorithms have been proposed during the last five decades for the recovery

of the relaxation spectrum of a viscoelastic material from oscillatory shear data. These works, using different models, approaches and computational techniques, have included contributions by Baumgaertel and Winter [19], Honerkamp and Weese [20], Malkin [21], Malkin et al. [22], Stadler and Bailly [23], Davis and Goulding [24], Davis et al. [25] and Cho [26]. These studies, but also many others, have created new directions of research on discrete and continuous relaxation spectra identification based on dynamic moduli data, and have been conducted ever since [8,11,27].

However, a classical way of studying viscoelasticity is also the stress relaxation test, where time-dependent shear stress is studied for the step increase in strain [1–3,28]. For some materials, for example, highly hydrated biological plants, the stress relaxation test is much easier to perform and is more suitable than dynamic oscillatory shear tests [3,29]. There are only a few papers that deal with spectrum determination from time measurements of the relaxation modulus. Additionally, some of them only address specific materials. The first works came from the turn of the 1940s and 1950s. Alfrey and Doty [30] proposed a simple differential model, being, in fact, the first-order Post–Widder formula [31] for the inverse Laplace transform; for details, see the newer edition [32]. Ter Harr [33] approximated the spectrum of relaxation frequencies by the modulus multiplied by time inverse of the relaxation frequency, which can be thought as the Post–Widder inversion formula of the zero order. After many years, Bažant and Yunping [34] and Goangseup and Bažant [35] introduced a two-stage approach of approximating the stress (or retardation) data via multiple differentiable models of the relaxation modulus (creep compliance) and, next, by applying the Post–Widder formula to designate the related relaxation (retardation) spectrum model. The effectiveness of this approach depends, among other aspects, on the function applied to approximate the relaxation modulus. In [34], a logarithmic–exponential model of the relaxation modulus is proposed, for which the authors state the third-order Post–Widder approximation to be satisfactory.

The relaxation spectrum modeling based on the known pairs of Laplace transforms was initiated by Macey [36], who described the relaxation modulus of viscoelastic ceramic material by the modified Bessel function of the second kind and zero order, which corresponds to the exponential–hyperbolic model of the spectrum. To describe the mechanical properties of polyisobutylene, Sips [37] introduced a simple relaxation spectrum model given by the difference of two exponential functions and a related logarithmic model of the modulus. This model was augmented to consider a long-term modulus by Yamamoto [38] and applied to test the rheological properties of the plant cell wall.

Both the algorithms based on the Post–Widder formula and those using the pairs of Laplace transforms assumed rather narrow classes of models. Thus, the scope of their effective applicability is limited to similar, not much wider classes of 'real' relaxation characteristics. A wider range of applicability has been offered by the models based on the expansion of an unknown spectrum into a series of basis functions (or polynomials) forming a complete basis in a function space, for example, in the space of real-valued square-integrable functions. Stankiewicz [39] and Stankiewicz and Gołacki [18] derived algorithms of the optimal regularized identification of relaxation and retardation spectra in the classes of models generated by different special functions, where various rules were applied for the choice of regularization parameters.

However, articles [18,39] were based on such a definition of the relaxation spectrum, according to which, the modulus was directly given by the Laplace integral of the spectrum. This spectrum definition is not often used in the literature. Therefore, computationally efficient algorithms to determine the relaxation spectrum applied to time measurements of the relaxation modulus are still desirable. Recently, for the dominant definition of the relaxation spectrum in the literature, a class of algorithms for relaxation spectrum recovery, which combines the technique of an expansion of a function into a series in an orthonormal basis with the least-squares regularized identification, has been derived by Stankiewicz [40]. Legendre, Laguerre, Chebyshev and Hermite functions were used as the basis functions for the spectrum model. In [40], the problem of determining the spectrum of relaxation

frequencies was considered. The complementary problem of determining the spectrum of relaxation times based on the stress relaxation data is both analytically and numerically more difficult than the determination of the spectrum of relaxation frequencies, because the relaxation time occurs in the denominator of the exponential function in the spectrum definitional formula. This very task is the subject of this article. Previous studies [18,39,40] have suggested that by selecting an appropriate time-scale factor, a better fit of the model to the measurement data can be obtained.

The objective of the present paper was to develop a model and an identification algorithm for the determination of the continuous relaxation time spectrum based on discrete-time measurements of the relaxation modulus, which, taking into account the ill-posedness of the original problem of the spectrum recovery and the idea of the optimal selection of the time-scale factor, will provide: (a) good approximation of the relaxation spectrum and modulus also due to the best choice of the time-scale factor; (b) smoothness of the spectrum fluctuations, even for noise-corrupted measurements; (c) noise robustness; (d) applicability to a wide range of viscoelastic materials; (e) ease of the implementation of the model and identification algorithm in available computing packages. The idea of the optimal choice of time-scale factor is used here for the first time in the context of relaxation spectrum identification. The goal of this work was the synthesis of the respective model and identification scheme and the analysis of their properties. A further purpose was the numerical verification of the model and algorithm for double-mode Gauss-like distribution used to describe the viscoelastic properties of various materials, in particular different polymers, wood and glass.

A new hierarchical algorithm of noise robust approximation of the continuous spectrum of relaxation times by finite series of power–exponential basis functions is proposed. The components of the relaxation modulus model are given by compact analytical formulas described by the product of power of time and the modified Bessel function of the second kind. The main properties of the basis functions of relaxation spectrum and modulus models have been studied; positive definiteness, upper bounds, monotonicity and asymptotic properties have been examined. Ranges of applicability for different time-scale factors are determined. Since the problem of relaxation spectrum identification is an ill-posed inverse problem, the regularized least-squares identification technique is applied combined with generalized cross-validation to guarantee the stability of the scheme. The quadratic identification index refers to the measured relaxation modulus. The task of determining the best "regularized" model is solved at the lower level of the identification scheme. The appropriate choice of the time-scale factor on the upper level of the scheme ensures the best fit of the relaxation modulus to experimental data.

The optimality conditions are derived both for the identification problem solved at the lower level and the task of the optimal choice of the time-scale factor at the upper level of the scheme. It is proved that the smoothness of the vector of the optimal model parameters implies smoothness of the fluctuations of the relaxation spectrum model. A direct formula and upper and lower bounds for the square integral norm of the smoothed spectrum model are derived. The accuracy of the spectrum model for noisy measurements of the relaxation modulus is studied, and the linear convergence to the model that we would obtain for the noise-free measurements is proved.

To design a numerical algorithm based on the scheme, the computations should be arranged hierarchically in a two-level scheme, i.e., for each iteration of the minimization procedure at the upper level, the whole numerical procedure must be realized for the lower-level task. The complete computational procedure for determining the best model is described. The singular value decomposition technique is applied to simplify the algebraic computations. The identification scheme can be easily implemented in available computing environments. The numerical studies, especially the strong smoothing of the relaxation spectrum models in the example of a double-mode Gauss-like spectrum, with a simultaneous very good fit to the experimental data of the relaxation modulus models, suggests

the need to explore the applicability of the regularized weighted least-squares [41] in the lower-level identification task. This will be the subject of further work.

In Appendix A, the proofs and derivations of some mathematical formulas are given. Some tables have been moved to Appendix B to increase the clarity of the article.

2. Materials and Methods

2.1. Relaxation Time Spectrum

The uniaxial, non-aging and isothermal stress–strain equation for a linear viscoelastic material is represented by a Boltzmann superposition integral [2]:

$$\sigma(t) = \int_{-\infty}^{t} G(t-u)\dot{\varepsilon}(u)du, \tag{1}$$

where u is the past time variable in the range $-\infty$ to the present time t. In Equation (1), variables $\sigma(t)$ and $\varepsilon(t)$ denote, respectively, the stress and strain at time t, and $G(t)$ is the linear relaxation modulus. Modulus $G(t)$ is given by [2,5]:

$$G(t) = \int_{0}^{\infty} \frac{H(\tau)}{\tau} e^{-t/\tau} d\tau, \tag{2}$$

where $H(\tau)$ characterizes the distributions of relaxation times τ. The continuous relaxation spectrum $H(\tau)$ is a generalization of the discrete Maxwell spectrum [2,5] to a continuous function of the relaxation times τ.

The problem of relaxation spectrum identification is the problem of solving a system of Fredholm integral equations of the first kind (2) obtained for discrete-time measurement data. In this paper, time measurements of the relaxation modulus are considered. This problem is ill-posed in the Hadamard sense [42], i.e., small changes in measured relaxation modulus can lead to arbitrarily large changes in the determined relaxation spectrum. As a remedy, some reductions in the set of admissible solutions or the appropriate regularization of the original problem are used. Both the techniques are applied here.

2.2. Models

Assume that $H(\tau) \in L^2(0,\infty)$, where $L^2(0,\infty)$ is the space of real-valued square-integrable functions on the interval $(0,\infty)$. It is known that the set of the linearly independent functions $\{e^{-\alpha\tau}, \tau e^{-\alpha\tau}, \tau^2 e^{-\alpha\tau}, \ldots\}$ form a basis of the space $L^2(0,\infty)$ [43] (p. 125); here, α is a positive time-scaling factor. Since the maximum

$$\max_{\tau \geq 0} \overline{h}_k(\tau,\alpha) = \left(\frac{k}{\alpha}\right)^k e^{-k}$$

of the function $\overline{h}_k(\tau,\alpha) = \tau^k e^{-\alpha\tau}$ grows rapidly with k, it is convenient to expand the relaxation spectrum into a series of scaled basis functions

$$h_k(\tau,\alpha) = \left(\frac{\alpha\tau}{k}\right)^k e^{-\alpha\tau+k}, \; k = 1, 2, \ldots, \tag{3}$$

with the first function

$$h_0(\tau,\alpha) = e^{-\alpha\tau}, \tag{4}$$

as follows:

$$H(\tau) = \sum_{k=0}^{\infty} g_k h_k(\tau,\alpha), \tag{5}$$

where g_k are real model coefficients.

It is practical to replace the infinite summation in Equation (5) with a finite one of K first terms, i.e., to approximate the spectrum $H(\tau)$ by a model of the form

$$H_K(\tau,\alpha) = \sum_{k=0}^{K-1} g_k h_k(\tau,\alpha), \tag{6}$$

where the lower index is the number of model summands. Then, according to (2), the respective model of the relaxation modulus is described by:

$$G_K(t, \alpha) = \int_0^\infty \frac{H_K(\tau, \alpha)}{\tau} e^{-t/\tau} d\tau = \sum_{k=0}^{K-1} g_k \phi_k(t, \alpha), \tag{7}$$

where the functions

$$\phi_k(t, \alpha) = \int_0^\infty \frac{h_k(\tau, \alpha)}{\tau} e^{-t/\tau} d\tau. \tag{8}$$

For computational purposes, function $h_0(\tau, \alpha)$ is defined by (4). However, since most mathematicians agreed that $0^0 = 1$ [44], the general Formula (3) can also be applied in further analysis for $k = 0$.

The basis functions $\phi_k(t, \alpha)$ (8) of the modulus model (7) are given by a compact analytical formula specified by the following theorem proved in Appendix A.1.

Theorem 1. *Let $\alpha > 0$, $k \geq 0$ and $t > 0$. Then, the basis functions $\phi_k(t, \alpha)$ (8) are given by:*

$$\phi_k(t, \alpha) = 2e^k \left(\frac{\sqrt{\alpha t}}{k} \right)^k K_k \left(2\sqrt{\alpha t} \right), \tag{9}$$

where $K_k(x)$ is the modified Bessel function of the second kind [45,46] of integer order k.

The first Function (9) is as follows

$$\phi_0(t, \alpha) = 2K_0 \left(2\sqrt{\alpha t} \right). \tag{10}$$

The modified Bessel functions of the second kind, and especially $K_0(x)$ [47], have many applications in science and engineering, for example, in physics to describe the flow of magneto-hydrodynamic (MHD) viscous fluid in a Darcy-type porous medium [48], in engineering to derive a closed analytical form of the model of a axial-flux permanent magnet machine with segmented multipole-Halbach PM array [49] and to describe the per-unit-length internal impedance of two-layer cylindrical conductors [50]. The applications of the Bessel functions in the description of the dynamic response of a mono-pile foundation in homogeneous soil and varied layered soil–rock conditions under horizontal dynamic loads [51], to obtain a fully coupled poroelastic solution for spherical indentation into a half space with an impermeable surface when the indenter is subjected to step displacement loading [52] and to express a distribution of the traveling distance in heterogeneous populations [53] come from material science and ecology.

Five first basis functions $h_k(\tau, \alpha)$ (3) are shown in Figure 1 for two different values of the time-scaling factor α. Figure 2 shows the related functions $\phi_k(t, \alpha)$ (9). The basis functions $h_k(\tau, \alpha)$ and $\phi_k(t, \alpha)$ are dimensionless. It is seen from Figure 1 that the maximum of each scaled basis function $h_k(\tau, \alpha)$ (3) and (4) is equal one; however, the relaxation time t^{max} corresponding to the maximum, for a given parameter α, depends on the index k according to the formula $t^{max} = k/\alpha$; i.e., grows with k. This means that increasing the number of model components K will allow for good modeling of multimodal spectra, which is confirmed by the example presented in the final part of the paper. Reducing the time-scale factor α shifts the spectrum maxima towards larger relaxation times. From Figure 2, it is seen that the Debye decay monotonicity of basis functions for the relaxation modulus model is in good agreement with the courses of the relaxation modulus obtained in an experiment for real materials; for example: concrete [10] (Figure 13), rubber [12] (Figure 2), elastic polyacrylamide hydrogels [28] (Figures 2a,b, 4a, A5, A7 and A8a), sugar beet [18] (Figure 1) and several foods [3] (Figures 3–39).
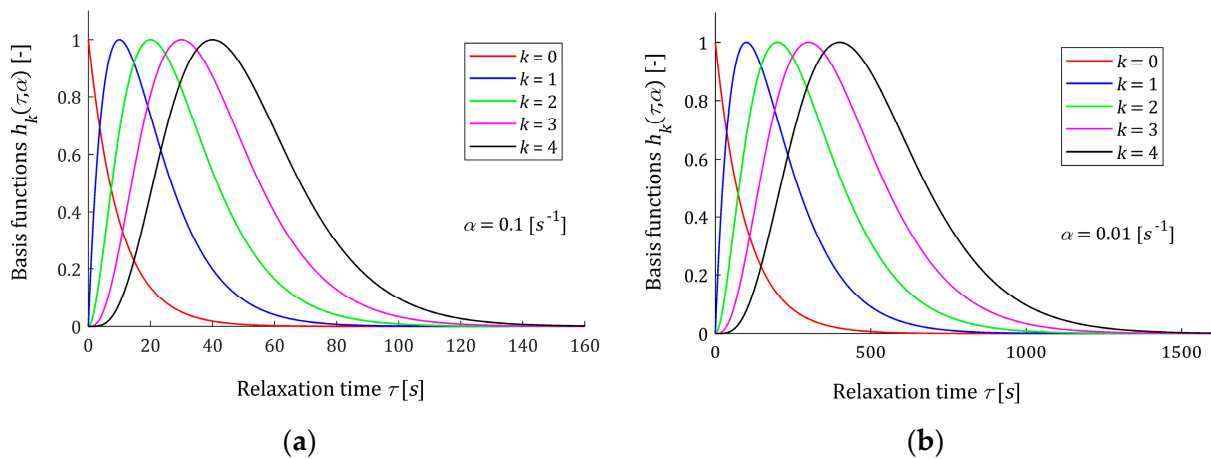
Figure 1. Basis functions $h_k(\tau, \alpha)$ (3) and (4) of the relaxation spectrum model $H_K(\tau, \alpha)$ (6) for time-scaling factors: (**a**) $\alpha = 0.1 \left[\text{s}^{-1} \right]$ and (**b**) $\alpha = 0.01 \left[\text{s}^{-1} \right]$, $k = 0, 1, 2, 3, 4$.
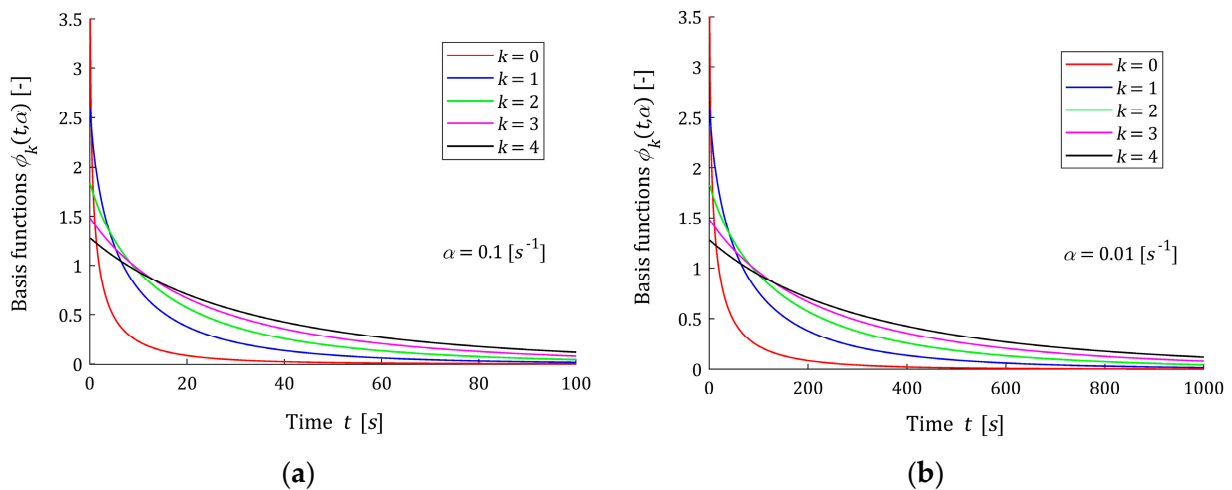
Figure 2. Basis functions $\phi_k(t, \alpha)$ (9) of the relaxation modulus model for time-scaling factors: (**a**) $\alpha = 0.1 \left[\text{s}^{-1} \right]$ and (**b**) $\alpha = 0.01 \left[\text{s}^{-1} \right]$, $k = 0, 1, 2, 3, 4$.

2.2.1. Positive Definiteness of the Basis Functions

The basis functions of the relaxation spectrum and modulus models are positive definite. For the functions $h_k(\tau, \alpha)$ (3) and (4), these properties are obvious. Since, according to the property A.1 in [54], the Bessel functions of the second kind $K_v(x)$ are positive for $x > 0$ and real v, the positive definiteness of the functions $\phi_k(t, \alpha)$ (9) and $\phi_0(t, \alpha)$ (10) directly result.

2.2.2. Asymptotic Properties of the Basis Functions

For $x \to 0$, the following asymptotic formula is found in the literature [45]:

$$K_k(x) \sim \frac{1}{2} \Gamma(k) \left(\frac{x}{2} \right)^{-k},$$ (11)

for modified Bessel functions of the order $k > 0$. Thus, for $t \to 0$, Formulas (11) and (9) imply

$$\phi_k(t, \alpha) \sim \Gamma(k) \left(\frac{e}{k} \right)^k,$$

which means that for t near zero, the values of basis functions $\phi_k(t, \alpha)$ decrease with increasing index k; see Figure 2.

For argument $x \to \infty$, the asymptotic exponential formula holds [54]

$$K_k(x) \sim \sqrt{\frac{\pi}{2x}} e^{-x} . \tag{12}$$

From (9) and (12), the next asymptotic formula follows for large t and $k \geq 1$

$$\phi_k(t, \alpha) \sim \frac{\sqrt{\pi}}{k^k} \left(\sqrt{\alpha t} \right)^{k - \frac{1}{4}} e^{-2\sqrt{\alpha t} + k} , \tag{13}$$

while for $t \to \infty$ and the first basis function, we have

$$\phi_0(t, \alpha) \sim \sqrt{\frac{\pi}{\sqrt{\alpha t}}} e^{-2\sqrt{\alpha t}} . \tag{14}$$

The multiplication by power function $\left(\sqrt{\alpha t} \right)^{k - \frac{1}{4}}$ in (13) causes that the greater k is, the slower the basis function $\phi_k(t, \alpha)$ decreases, which is seen in Figure 2. The first basis function $\phi_0(t, \alpha)$ decreases faster than the exponential function, which is also confirmed by the analysis of the course of the basis functions in Figure 2.

By the asymptotic Formulas (13) and (14), the basis functions tend to 0 as $t \to \infty$; i.e., for an arbitrary, $\alpha > 0$ and $k = 0, 1, 2, \ldots$

$$\lim_{t \to \infty} \phi_k(t, \alpha) = 0. \tag{15}$$

2.2.3. Upper Bounds for the Basis Functions

In [54], Corollary 3.4, the following inequality is proved:

$$K_0(x) < \sqrt{\frac{\pi}{2x}} e^{-x},$$

whence

$$\phi_0(t, \alpha) = 2K_0 \left(2\sqrt{\alpha t} \right) < \sqrt{\frac{\pi}{\sqrt{\alpha t}}} e^{-2\sqrt{\alpha t}} . \tag{16}$$

From Theorem 3.1 in [54], for $k \geq 1$ we have

$$K_k(x) < \frac{2^{k-1} \Gamma(k)}{x^{k+1}} .$$

This inequality applied into (9) gives the following upper bound

$$\phi_k(t, \alpha) < 2e^k \left(\frac{\sqrt{\alpha t}}{k} \right)^k \frac{2^{k-1} \Gamma(k)}{\left(2\sqrt{\alpha t} \right)^{k+1}} = \frac{e^k}{2k^k} \frac{\Gamma(k)}{\sqrt{\alpha t}} . \tag{17}$$

Note that having in mind the positive definiteness of the basis functions $\phi_k(t, \alpha)$, the limit (15) for the first function $\phi_k(t, \alpha)$ yields from (16), while, for any $k \geq 1$, can be derived from (17).

2.2.4. Monotonicity of the Basis Functions

As indicated above, the basis functions $h_k(\tau, \alpha)$ (3) for $k \geq 1$ have a global maximum equal to 1 for the relaxation time $\tau = k\alpha^{-1}$, while the first function $h_0(\tau, \alpha)$ (4) is monotonically decreasing.

Since the basis function $\phi_k(t, \alpha)$ (9) is the product of a monotonically decreasing Bessel function $K_k(x)$ [54] of increasing argument $2\sqrt{\alpha t}$ and a monotonically increasing power

function $\left(\sqrt{\alpha t}\right)^k$, its monotonicity is not obvious. By applying the differentiation formula, we obtain [54] (Equation (A.13)):

$$\frac{d}{dx}[x^v K_v(x)] = -x^v K_{v-1}(x) \tag{18}$$

which holds for any real-valued order v, to (9), gives for integer $k \geq 1$

$$\frac{d}{dt}[\phi_k(t,\alpha)] = -4\alpha \left(\frac{e}{2k}\right)^k \left(2\sqrt{\alpha t}\right)^{k-1} K_{k-1}\left(2\sqrt{\alpha t}\right),$$

whereas, having in mind the positive definiteness of the Bessel function $K_{k-1}\left(2\sqrt{\alpha t}\right)$ for all $t > 0$, we immediately conclude that the function $\phi_k(t,\alpha)$ is monotonically decreasing; see Figure 2. For the first basis function $\phi_0(t,\alpha)$ (10), the next differentiation formula is [54] (Equation (A.14)):

$$\frac{d}{dx}[K_v(x)] = -\frac{1}{2}[K_{v-1}(x) + K_{v+1}(x)], \tag{19}$$

which is satisfied for any real v, implies

$$\frac{d}{dt}[\phi_0(t,\alpha)] = -\sqrt{\frac{\alpha}{t}}\left[K_{-1}\left(2\sqrt{\alpha t}\right) + K_1\left(2\sqrt{\alpha t}\right)\right],$$

which shows that $\phi_0(t,\alpha)$ is also a monotonically decreasing function.

2.2.5. Ranges of Applicability

In the models, the parameter $\alpha > 0$ is a time-scaling factor. The following rule holds: the lower the parameter α, the greater the relaxation times. The above is illustrated by Figures 1 and 2. Through the optimal choice of the scaling factor, the best fit of the model to the experimental data is achieved, which will be the subject of study in Section 3.2.

Following [40], on the basis of the relaxation modulus course, the range of applicability is specified as the time t, for which the first K basis functions $\phi_k(t,\alpha)$ no longer permanently exceed; i.e., for any $\theta > t$, $\varepsilon = 0.5\%$ of its maximum value. Specifically,

$$t_{app}(\alpha) = \max_{0 \leq k \leq K-1} \min_{t > 0} \{t : |\phi_k(\theta,\alpha)| \leq 0.005 \cdot \phi_{kmax}(\alpha) \text{ for any } \theta \geq t\}, \tag{20}$$

where

$$\phi_{kmax}(\alpha) = \max_{t \geq 0} |\phi_k(t,\alpha)|.$$

Similarly, in [40], the range of applicability specified directly for the relaxation times τ was defined on the basis of the variability in the basis functions $h_k(\tau)$; i.e.,

$$\tau_{app}(\alpha) = \max_{0 \leq k \leq K-1} \min_{\tau > 0} \{\tau : |h_k(\vartheta,\alpha)| \leq 0.005 \cdot h_{kmax}(\alpha) \text{ for any } \vartheta \geq \tau\}, \tag{21}$$

with $h_{kmax}(\alpha)$ defined by

$$h_{kmax}(\alpha) = \max_{\tau \geq 0} |h_k(\tau,\alpha)|.$$

The times $t_{app}(\alpha)$ (20) and $\tau_{app}(\alpha)$ (21) for different values of α are summarized in Table 1 for $K = 5$ and $K = 12$. The same data for $K = 6, \ldots, 11$ are given in Table A1 in Appendix B.

Table 1. Ranges of the applicability of the model for various time-scale parameters for $K = 5$ and $K = 12$.

Time-Scale Factor α [s]	K = 5	K = 5	K = 12	K = 12
	Range [1] of Relaxation Times $\tau_{app}(\alpha)$ [s]	Range [1] of Times $t_{app}(\alpha)$ [s]	Range [1] of Relaxation Times $\tau_{app}(\alpha)$ [s]	Range [1] of Times $t_{app}(\alpha)$ [s]
0.0001	144,305.22	282,360.6	255,824.35	662,077.1
0.001	14,430.52	28,236.06	25,582.435	66,207.71
0.01	1443.05	2823.61	2558.243	6620.77
0.1	144.305	282.36	255.824	662.08
1	14.43	28.236	25.5824	66.208
10	1.4431	2.824	2.5582	6.621
100	0.1443	0.282	0.2558	0.662

[1] The upper bounds t_{app} (20) and $\tau_{app}(\alpha)$ (21) of the applicability intervals $[0, t_{app}(\alpha)]$ and $[0, \tau_{app}(\alpha)]$ are given.

A review of the data from these tables shows that for any fixed α, both $\tau_{app}(\alpha)$ and $t_{app}(\alpha)$ grow almost linearly with the number of model summands K.

2.3. Least-Squares Regularized Identification

Identification consists of the selection, within the chosen class of models given by (6) and (7), of such a model that ensures the best fit to the measurement results. Suppose a certain identification experiment (stress relaxation test [2,3]) resulted in a set of measurements of the relaxation modulus $\{\overline{G}(t_i) = G(t_i) + z(t_i)\}$ at the sampling instants $t_i \geq 0$, $i = 1, \ldots, N$, where $z(t_i)$ is the measurement noise. It is assumed that the number of measurements $N \geq K$. As a measure of the model (7) accuracy, the quadratic index is taken

$$Q_N(g_K, \alpha) = \sum_{i=1}^{N} \left[\overline{G}(t_i) - G_K(t_i, \alpha) \right]^2, \tag{22}$$

where $g_K = \begin{bmatrix} g_0 & \cdots & g_{K-1} \end{bmatrix}^T$ is an K-element vector of unknown coefficients of the models (6) and (7). The identification index (22) is rewritten in the compact form as

$$Q_N(g_K, \alpha) = \| \overline{G}_N - \Phi_{N,K}(\alpha) g_K \|_2^2, \tag{23}$$

where

$$\Phi_{N,K}(\alpha) = \begin{bmatrix} \phi_0(t_1, \alpha) & \cdots & \phi_{K-1}(t_1, \alpha) \\ \vdots & \ddots & \vdots \\ \phi_0(t_N, \alpha) & \cdots & \phi_{K-1}(t_N, \alpha) \end{bmatrix}, \quad \overline{G}_N = \begin{bmatrix} \overline{G}(t_1) \\ \vdots \\ \overline{G}(t_N) \end{bmatrix} \tag{24}$$

and $\|\cdot\|_2$ denotes the square norm in the real Euclidean space $\mathcal{R}^N$. Thus, the optimal identification of the relaxation spectrum in the class of models defined by (6) and (7) consists of determining the model parameter g_K through solving the following optimization task:

$$\min_{g_K \in \mathcal{R}^K} \| \overline{G}_N - \Phi_{N,K}(\alpha) g_K \|_2^2. \tag{25}$$

The matrix $\Phi_{N,K}(\alpha)$ is usually ill-conditioned. In consequence, the problem (25) is ill posed in the sense of Hadamard [42]. The parameter g_K minimizing identification index (23) is not unique, and even the normal (with the lowest Euclidean norm) solution of (25) is a non-continuous and unbounded function of the measurement vector $\overline{G}_N$. Therefore, when the data are noisy, even small changes in $\overline{G}_N$ would lead to arbitrarily large artefacts in optimal model parameters g_K. To deal with the ill-posedness, Tikhonov regularization [55], replacing the original ill-posed problem (25) by a nearby problem with a modified square functional of the form:

$$\min_{g_K \in \mathcal{R}^K} \| \overline{G}_N - \Phi_{N,K}(\alpha) g_K \|_2^2 + \lambda \| g_K \|_2^2, \tag{26}$$

where $\lambda > 0$ is a regularization parameter, can be used. The regularized task (26) is well-posed; that is, the solution always exists, is unique and continuously depends on both the matrix $\boldsymbol{\Phi}_{N,K}$ and on the measurement data $\overline{G}_N$. The parameter vector solving (26) is given by:

$$\overline{g}_K^\lambda(\alpha) = \left(\boldsymbol{\Phi}_{N,K}^T(\alpha)\boldsymbol{\Phi}_{N,K}(\alpha) + \lambda\mathbb{I}_{K,K} \right)^{-1} \boldsymbol{\Phi}_{N,K}^T(\alpha)\overline{G}_N, \tag{27}$$

where $\mathbb{I}_{K,K}$ is the K-dimensional identity matrix.

Following [40], the generalized cross-validation GCV [42,56] is applied, according to which, the best regularization parameter is [42,56]:

$$\lambda_{GCV}(\alpha) = min\left\{ \lambda : \ \lambda = arg \min_{\lambda \geq 0} V_{GCV}(\lambda, \alpha) \right\}, \tag{28}$$

where the GCV functional is defined by [56]

$$V_{GCV}(\lambda, \alpha) = \frac{\left\| \boldsymbol{\Xi}(\lambda, \alpha)\overline{G}_N \right\|_2^2}{tr[\boldsymbol{\Xi}(\lambda, \alpha)]^2}, \tag{29}$$

with the matrix

$$\boldsymbol{\Xi}(\lambda, \alpha) = \mathbb{I}_{N,N} - \boldsymbol{\Phi}_{N,K}(\alpha)\left(\boldsymbol{\Phi}_{N,K}^T(\alpha)\boldsymbol{\Phi}_{N,K}(\alpha) + \lambda\mathbb{I}_{K,K} \right)^{-1} \boldsymbol{\Phi}_{N,K}^T(\alpha). \tag{30}$$

Here, $\boldsymbol{\Xi}(\lambda, \alpha)\overline{G}_N$ is the residual vector for the regularized solution (27); $tr[\boldsymbol{\Xi}(\lambda, \alpha)]$ denotes the trace of the symmetric matrix $\boldsymbol{\Xi}(\lambda, \alpha)$. The optimization problem in (28) has a unique solution, and the resulting parameter $\overline{g}_K^{\lambda_{GCV}(\alpha)}(\alpha)$ differs the least from the normal solution of the problem (26) that we would obtain for the ideal (not noise-corrupted) measurements of the relaxation modulus [56].

3. Results and Discussion

In this section, the necessary optimality condition for the regularized identification task with GCV's choice of regularization parameter is given in the form of an algebraic nonlinear equation. Next, the problem of the choice of the optimal scale factor is stated, and the respective necessary optimality condition is derived. A hierarchical two-level identification scheme with the optimal choice of the scale factor is proposed. The numerical realization and the application of a singular value decomposition technique are discussed. A complete computational procedure is outlined. The analysis of the smoothing of the model and model accuracy for noisy measurements of the relaxation modulus is presented.

3.1. Necessary Optimality Condition

The GCV functional (29) is a differentiable function of the regularization parameter λ. The necessary optimality condition for the minimization task solved in (28) is derived in Appendix A.3.

Theorem 2. *The optimal regularization parameter $\lambda_{GCV}(\alpha)$ solves the following equation:*

$$\overline{G}_N^T(\boldsymbol{\Psi}(\alpha) + \lambda\mathbb{I}_{N,N})^{-3}\overline{G}_N tr\left[(\boldsymbol{\Psi}(\alpha) + \lambda\mathbb{I}_{N,N})^{-1}\right] = \overline{G}_N^T(\boldsymbol{\Psi}(\alpha) +$$
$$\lambda\mathbb{I}_{N,N})^{-2}\overline{G}_N \ tr\left[(\boldsymbol{\Psi}(\alpha) + \lambda\mathbb{I}_{N,N})^{-2}\right]. \tag{31}$$

where

$$\boldsymbol{\Psi}(\alpha) = \boldsymbol{\Phi}_{N,K}(\alpha)\boldsymbol{\Phi}_{N,K}^T(\alpha). \tag{32}$$

In Appendix A.4, the following formula:

$$\frac{\partial V_{GCV}(\lambda,\alpha)}{\partial \lambda} =$$
$$\frac{-2\overline{G}_N^T(\boldsymbol{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-3}\,\overline{G}_N tr\left[(\boldsymbol{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-1}\right]+2\overline{G}_N^T(\boldsymbol{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-2}\overline{G}_N tr\left[(\boldsymbol{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-2}\right]}{\left[tr\left[(\boldsymbol{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-1}\right]\right]^3} \tag{33}$$

is developed, describing the derivative of the GCV functional (29) minimized in (28) directly as a function of matrices $\overline{G}_N$ and $\boldsymbol{\Psi}(\alpha)$.

3.2. Choice of the Time-Scale Factor

The optimal regularization parameter $\lambda_{GCV}(\alpha)$ and matrix $\boldsymbol{\Phi}_{N,K}(\alpha)$ depend on the time-scale factor. Therefore, the optimal model parameter

$$\widehat{g}_K^{\lambda_{GCV}}(\alpha) = \overline{g}_K^\lambda(\alpha)\Big|_{\lambda=\lambda_{GCV}(\alpha)} \tag{34}$$

and the optimal identification index

$$Q_{Nopt}(\alpha) = Q_N\left(\widehat{g}_K^{\lambda_{GCV}}(\alpha),\alpha\right) = \|\overline{G}_N - \boldsymbol{\Phi}_{N,K}(\alpha)\widehat{g}_K^{\lambda_{GCV}}(\alpha)\|_2^2 \tag{35}$$

also depend on α. By the optimal choice of the scaling factor, the best fit of the model to the experimental data can be achieved. By (27), (30) and (34), and having (A9) in mind, the index $Q_{Nopt}(\alpha)$ (35) is expressed by the following formula:

$$Q_{Nopt}(\alpha) = \lambda_{GCV}^2(\alpha)\overline{G}_N^T(\boldsymbol{\Psi}(\alpha) + \lambda_{GCV}(\alpha)\mathbb{I}_{N,N})^{-2}\overline{G}_N. \tag{36}$$

The smaller the index $Q_{Nopt}(\alpha)$, the smaller the model error results. Thus, the problem of the choice of the best time-scale factor α_{opt} takes the form

$$\min_{\alpha>0} Q_{Nopt}(\alpha) = Q_{Nopt}\left(\alpha_{opt}\right). \tag{37}$$

Before we state the necessary optimality condition of the task (37), we prove the following result concerning the basis functions $\phi_k(t,\alpha)$ as the functions of the time-scale factor. The proof is given in Appendix A.5.

Theorem 3. *Let $\alpha > 0$ and $t > 0$. Then, for $k \geq 1$, derivatives of the basis functions $\phi_k(t,\alpha)$ (9) with respect to the parameter α are given by:*

$$\frac{d}{d\alpha}[\phi_k(t,\alpha)] = -e\left(\frac{t}{k}\right)\left(\frac{k-1}{k}\right)^{k-1}\phi_{k-1}(t,\alpha). \tag{38}$$

For $k = 0$, the following formula holds:

$$\frac{d}{d\alpha}[\phi_0(t,\alpha)] = -\frac{1}{e\alpha}\phi_1(t,\alpha). \tag{39}$$

Thus, positive definite functions $\phi_k(t,\alpha)$ are monotonically decreasing functions of α for any fixed $t > 0$.

The following necessary optimality condition is derived in Appendix A.6:

Theorem 4. *The optimal time-scale factor α_{opt} satisfies the following equation:*

$$\lambda_{GCV}(\alpha)\overline{G}_N^T Y(\alpha)^{-1}\boldsymbol{\Omega}_{N,K}(\alpha)Y(\alpha)^{-2}\overline{G}_N = \frac{d\lambda_{GCV}(\alpha)}{d\alpha}\overline{G}_N^T Y(\alpha)^{-2}\overline{G}_N -$$
$$\lambda_{GCV}(\alpha)\frac{d\lambda_{GCV}(\alpha)}{d\alpha}\overline{G}_N^T Y(\alpha)^{-3}\overline{G}_N, \tag{40}$$

where derivative $\frac{d\lambda_{GCV}(\alpha)}{d\alpha}$ is given by the equation

$$
\begin{aligned}
\frac{d\lambda_{GCV}(\alpha)}{d\alpha} \Big\{ & 3\overline{\boldsymbol{G}}_N^T \boldsymbol{Y}(\alpha)^{-4}\overline{\boldsymbol{G}}_N tr\Big[\boldsymbol{Y}(\alpha)^{-1}\Big] - \overline{\boldsymbol{G}}_N^T \boldsymbol{Y}(\alpha)^{-3}\overline{\boldsymbol{G}}_N tr\Big[\boldsymbol{Y}(\alpha)^{-2}\Big] \\
& -2\overline{\boldsymbol{G}}_N^T \boldsymbol{Y}(\alpha)^{-2}\overline{\boldsymbol{G}}_N \ tr\Big[\boldsymbol{Y}(\alpha)^{-3}\Big]\Big\} = \\
& 2\overline{\boldsymbol{G}}_N^T \boldsymbol{Y}(\alpha)^{-2}\boldsymbol{\Omega}_{N,K}(\alpha)\boldsymbol{Y}(\alpha)^{-1}\overline{\boldsymbol{G}}_N \ tr\Big[\boldsymbol{Y}(\alpha)^{-2}\Big] + \\
& 2\overline{\boldsymbol{G}}_N^T \boldsymbol{Y}(\alpha)^{-2}\overline{\boldsymbol{G}}_N \ tr\Big[\boldsymbol{\Omega}_{N,K}(\alpha)\boldsymbol{Y}(\alpha)^{-3}\Big] - \\
& 2\overline{\boldsymbol{G}}_N^T \boldsymbol{Y}(\alpha)^{-3}\boldsymbol{\Omega}_{N,K}(\alpha)\boldsymbol{Y}(\alpha)^{-1}\overline{\boldsymbol{G}}_N tr\Big[\boldsymbol{Y}(\alpha)^{-1}\Big] - \\
\overline{\boldsymbol{G}}_N^T \boldsymbol{Y}(\alpha)^{-2}\boldsymbol{\Omega}_{N,K}(\alpha)&\boldsymbol{Y}(\alpha)^{-2}\overline{\boldsymbol{G}}_N tr\Big[\boldsymbol{Y}(\alpha)^{-1}\Big] - \overline{\boldsymbol{G}}_N^T \boldsymbol{Y}(\alpha)^{-3}\overline{\boldsymbol{G}}_N tr\Big[\boldsymbol{\Omega}_{N,K}(\alpha)\boldsymbol{Y}(\alpha)^{-2}\Big],
\end{aligned}
\tag{41}
$$

with symmetric matrices

$$
\boldsymbol{Y}(\alpha) = \boldsymbol{\Psi}(\alpha) + \lambda_{GCV}(\alpha)\mathbb{I}_{N,N},
\tag{42}
$$

$$
\boldsymbol{\Omega}_{N,K}(\alpha) = \frac{d}{d\alpha}\boldsymbol{\Psi}(\alpha) = \boldsymbol{\Theta}_{N,K}(\alpha)\boldsymbol{\Phi}_{N,K}^T(\alpha) + \boldsymbol{\Phi}_{N,K}(\alpha)\boldsymbol{\Theta}_{N,K}^T(\alpha).
\tag{43}
$$

where $\boldsymbol{\Psi}(\alpha)$ is defined by (32) and

$$
\boldsymbol{\Theta}_{N,K}(\alpha) = \frac{d}{d\alpha}\boldsymbol{\Phi}_{N,K}(\alpha) =
$$
$$
-\begin{bmatrix}
\frac{1}{e\alpha}\phi_1(t_1,\alpha) & et_1\phi_0(t_1,\alpha) & \cdots & e\left(\frac{t_1}{K-1}\right)\left(\frac{K-2}{K-1}\right)^{K-2}\phi_{K-2}(t_1,\alpha) \\
\vdots & \vdots & \ddots & \vdots \\
\frac{1}{e\alpha}\phi_1(t_N,\alpha) & et_N\phi_0(t_N,\alpha) & \cdots & e\left(\frac{t_N}{K-1}\right)\left(\frac{K-2}{K-1}\right)^{K-2}\phi_{K-2}(t_N,\alpha)
\end{bmatrix}.
\tag{44}
$$

3.3. Two-Level Identification Scheme

To find the optimal time-scale factor α_{opt} and the optimal model of the relaxation time spectrum, the following two-level scheme is applied:

3.3.1. Lower Level

Given time-scale factor $\alpha > 0$, find regularization parameter $\lambda_{GCV}(\alpha)$ solving the GCV minimization task (28).

3.3.2. Upper Level

Find the time-scale factor $\alpha_{opt} > 0$ minimizing identification index $Q_{Nopt}(\alpha)$ (36), i.e., solving optimization task (37). Take

$$
\widehat{\boldsymbol{g}}_K = \widehat{\boldsymbol{g}}_K^{\lambda_{GCV}}(\alpha_{opt}),
\tag{45}
$$

where $\widehat{\boldsymbol{g}}_K^{\lambda_{GCV}}(\alpha_{opt})$ is defined by (34) as a parameter of the best model of the relaxation time spectrum.

Having the optimal model parameter $\widehat{\boldsymbol{g}}_K$, the optimal model of the relaxation spectrum is determined according to the formula resulting directly from (6):

$$
H_{Kopt}(\tau) = H_K(\tau, \alpha_{opt}) = \sum_{k=0}^{K-1} \widehat{g}_k h_k(\tau, \alpha_{opt}),
\tag{46}
$$

where $\widehat{g}_k$ are elements of the vector $\widehat{\boldsymbol{g}}_K$.

To design numerical realization of the scheme, we need:

1. A numerical procedure for solving the lower-level GCV minimization task (28);
2. An iterative scheme for solving the upper-level problem (37) of choosing the best time-scale factor.

For any given parameter α, the GCV function $V_{GCV}(\lambda, \alpha)$ (29) is differentiable with respect to λ. Partial derivative $\frac{\partial V_{GCV}(\lambda, \alpha)}{\partial \lambda}$ is given by (33) as a function of the experiment data $\overline{G}_N$ and the matrix $\boldsymbol{\Psi}(\alpha)$ (32), which depends on the time instants t_i used in the relaxation experiment and on the time-scaling factor. An arbitrary gradient optimization method can be implemented to solve the GCV minimization task (28). Additionally, an arbitrary gradient method can be used to solve optimization problem (37); the derivative of the index $Q_{Nopt}(\alpha)$ (36) is described by Formula (A24) derived in Appendix A.6. However, the optimal parameter α_{opt} can be also found by solving the necessary optimality condition from Theorem 4, i.e., the two scalar algebraic Equations (40) and (41), in fact.

3.4. Algebraic Background of the Identification Scheme

Formulas (27)–(30), fundamental for lower-level optimization task (28), are elegant but generally unsuitable for computational purposes. Following [40], the singular value decomposition (SVD) technique [57] will be applied. Let SVD of the $N \times K$ dimensional matrix $\boldsymbol{\Phi}_{N,K}(\alpha)$ take the form [57]:

$$\boldsymbol{\Phi}_{N,K}(\alpha) = \boldsymbol{U}(\alpha)\boldsymbol{\Sigma}(\alpha)\,\boldsymbol{V}(\alpha)^T, \tag{47}$$

where $\boldsymbol{\Sigma}(\alpha) = diag\left(\sigma_1(\alpha), \ldots, \sigma_{r(\alpha)}(\alpha), 0, \ldots, 0\right) \epsilon \mathcal{R}^{N,K}$ is a diagonal matrix containing the non-zero singular values $\sigma_1(\alpha), \ldots, \sigma_{r(\alpha)}(\alpha)$ of the matrix $\boldsymbol{\Phi}_{N,K}(\alpha)$, matrices $\boldsymbol{V}(\alpha) \in \mathcal{R}^{K,K}$ and $\boldsymbol{U}(\alpha) \in \mathcal{R}^{N,N}$ are orthogonal and $r(\alpha) = rank[\boldsymbol{\Phi}_{N,K}(\alpha)] < N$. Taking advantage of the diagonal structure of $\boldsymbol{\Sigma}(\alpha)$ and orthogonality of the matrices $\boldsymbol{V}(\alpha)$ and $\boldsymbol{U}(\alpha)$, it may be simply proved that the parameter $\overline{g}_K^\lambda(\alpha)$ (27) is given by

$$\overline{g}_K^\lambda(\alpha) = \boldsymbol{V}(\alpha)\boldsymbol{\Lambda}_\lambda(\alpha)\,\boldsymbol{U}(\alpha)^T\,\overline{G}_N, \tag{48}$$

where $K \times N$ diagonal matrix $\boldsymbol{\Lambda}_\lambda(\alpha)$ is as follows:

$$\boldsymbol{\Lambda}_\lambda(\alpha) = diag\left(\frac{\sigma_1(\alpha)}{[\sigma_1(\alpha)]^2 + \lambda}, \ldots, \frac{\sigma_{r(\alpha)}(\alpha)}{\left[\sigma_{r(\alpha)}(\alpha)\right]^2 + \lambda}, 0, \ldots, 0\right). \tag{49}$$

Using SVD (47) and introducing N dimensional vector $\boldsymbol{Y}(\alpha) = \boldsymbol{U}(\alpha)^T\,\overline{G}_N$, the GCV function (29) is also expressed by a convenient analytical formula:

$$V_{GCV}(\lambda, \alpha) = \frac{\left[\sum_{i=1}^{r(\alpha)} \frac{\lambda^2 [y_i(\alpha)]^2}{\left([\sigma_i(\alpha)]^2 + \lambda\right)^2} + \sum_{i=r(\alpha)+1}^{N} [y_i(\alpha)]^2\right]}{\left[N - r(\alpha) + \sum_{i=1}^{r(\alpha)} \frac{\lambda}{\left([\sigma_i(\alpha)]^2 + \lambda\right)}\right]^2}, \tag{50}$$

as a function of the singular values $\sigma_i(\alpha)$ and elements $y_i(\alpha)$ of the vector $\boldsymbol{Y}(\alpha)$.

Similarly, the identification index $Q_{Nopt}(\alpha)$ (36) minimized in the upper level of the scheme is expressed using the SVD (47). Since, by virtue of (47) and (32) we have

$$\left(\boldsymbol{\Psi}(\alpha) + \lambda_{GCV}(\alpha)\mathbb{I}_{N,N}\right)^{-2} = \boldsymbol{U}(\alpha)\left[\boldsymbol{\Sigma}(\alpha)\,\boldsymbol{\Sigma}(\alpha)^T + \lambda_{GCV}(\alpha)\mathbb{I}_{N,N}\right]^{-2}\boldsymbol{U}(\alpha)^T,$$

the index $\widehat{Q}_N(\alpha)$ is expressed as

$$Q_{Nopt}(\alpha) = \lambda_{GCV}^2(\alpha)\left[\sum_{i=1}^{r(\alpha)} \frac{[y_i(\alpha)]^2}{\left([\sigma_i(\alpha)]^2 + \lambda_{GCV}(\alpha)\right)^2} + \sum_{i=r(\alpha)+1}^{N} [y_i(\alpha)]^2\right]. \tag{51}$$

3.5. Computational Algorithm for Model Identification

To design a numerical algorithm of the scheme, the communication between the levels should also be resolved. The computations must be arranged hierarchically in a two-level structure, i.e., for each iteration of the minimization procedure at the upper level, the whole numerical procedure must be realized for the lower-level GCV task (28). The complete computational procedure for determining the optimal model is given below.

Step 0: Perform the experiment—stress relaxation test [1–3,28]—and record the measurements $\overline{G}(t_i)$, $i = 1, \ldots, N$, of the relaxation modulus at times $t_i \geq 0$.

Step 1: Determine the optimal regularization parameter $\lambda_{GCV}(\alpha_{opt})$ in the following two-level computations.

> **Step 1.0**: Choose the initial point α^0 for the numerical procedure applied to solve the upper-level task (37).
>
> **Step 1.1**: Let α^m be the m-th iterate in the numerical procedure chosen to solve the upper-level task (37). For $\alpha = \alpha^m$, solve the lower-level minimization task (28) according to the chosen numerical optimization procedure and determine the regularization parameter $\lambda_{GCV}(\alpha^m)$. The algebraic formula $V_{GCV}(\lambda, \alpha)$ (50) is applied.
>
> **Step 1.2**: Using $\lambda_{GCV}(\alpha^m)$, compute, according to the numerical procedure selected to solve the upper-level task (37), with the index $Q_{Nopt}(\alpha)$ described by (51), the new parameter α^{m+1}, which is the next approximation of α_{opt}. If for α^{m+1} the stopping rule of the chosen numerical procedure is satisfied, i.e.,
>
> $$\left| \alpha^{m+1} - \alpha^m \right| \leq \varepsilon_1$$
>
> or
>
> $$\left| Q_{Nopt}\left(\alpha^{m+1}\right) - Q_{Nopt}(\alpha^m) \right| \leq \varepsilon_2,$$
>
> where ε_1 and ε_2 are preselected small positives, put $\alpha_{opt} = \alpha^m$ as the optimal time-scale factor, $\lambda_{GCV}(\alpha^m)$ as $\lambda_{GCV}(\alpha_{opt})$, and go to Step 2. Otherwise, return to Step 1.1 and continue the computations for $\alpha = \alpha^{m+1}$.

Step 2: Compute the vector of the optimal model parameters $\widehat{g}_K$ according to (45) and the best model of the relaxation spectrum $H_{Kopt}(\tau)$ given by (46).

Remark 1. *The appealing feature of the scheme is that only the values of $\lambda_{GCV}(\alpha^m)$, not the related parameters vector $\widehat{g}_K^{\lambda_{GCV}}(\alpha^m)$ (34), are used for α^m in successive iterations of the numerical procedure solving the upper-level task (37).*

Remark 2. *The regularization parameter $\lambda_{GCV}(\alpha^m)$ resulting from the lower-level minimization task (28) in each iteration of the upper level is the solution of the GCV problem (28) for current α^m. Thus, the respective vector $\widehat{g}_K^{\lambda_{GCV}}(\alpha^m)$ (34) can be treated as an approximate solution of the overall identification problem.*

Remark 3. *The selection of the initial time-scaling factor α^0 in Step 1.0 may be based on the data concerning model applicability summarized in Tables 1 and A1 or α^0 can be selected by comparison, for different values of α, the first basis function $\phi_k(t, \alpha)$ (9), with the experiment results $\{\overline{G}(t_i)\}$.*

Remark 4. *The SVD (47) of the matrix $\boldsymbol{\Phi}_{N,K}(\alpha)$, of computational complexity $\mathcal{O}(NK^2)$ [57], must be computed only once for successive α^m generated in Step 1.1. SVD is accessible in the form of optimized numerical procedures in most commonly used computational packets.*

The above procedure and communication between the levels are illustrated in Figure 3.
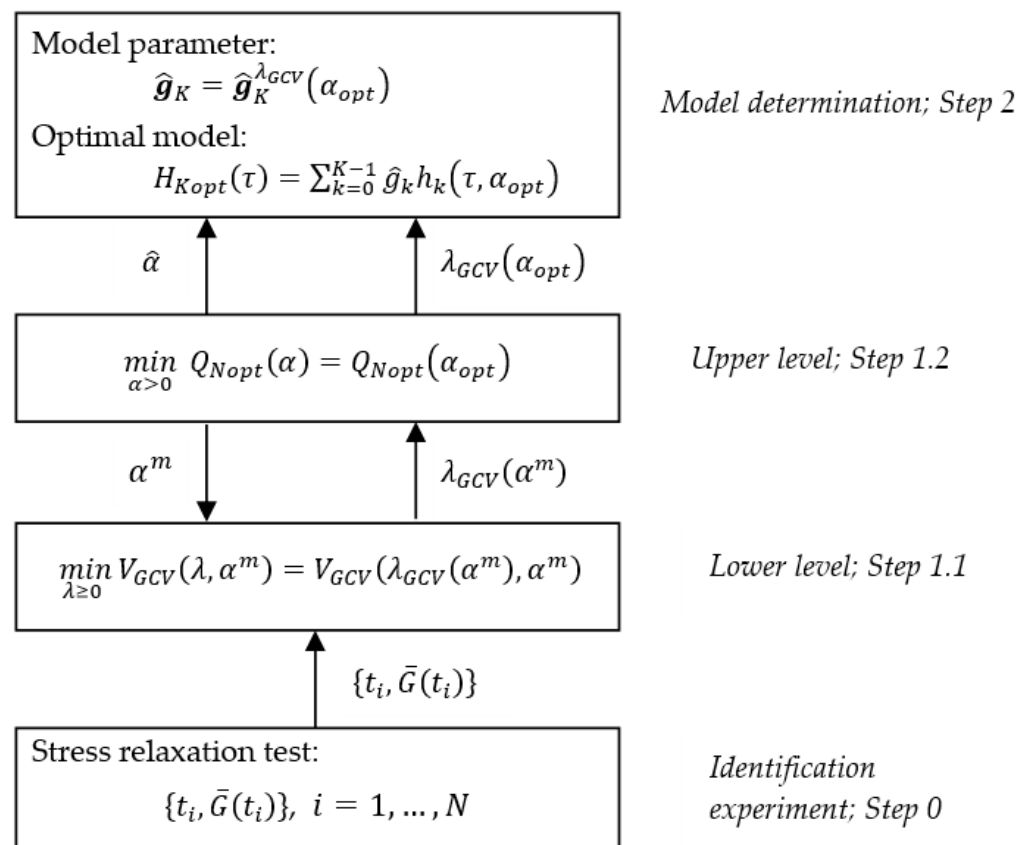
Figure 3. Hierarchical procedure of the two-level scheme for determination of the optimal spectrum relaxation model.

3.6. Analysis

Recently, a class of robust algorithms of the approximation of the continuous spectrum of relaxation frequencies by finite series of orthonormal functions has been developed and analyzed in detail in [40]. However, these results were derived using the orthogonality of the basis functions, so they are not applicable here. Therefore, both the analysis of the smoothing of the model and model accuracy for noisy measurements of the relaxation modulus must be carried out anew here.

3.6.1. Smoothness

The purpose of regularization relies on the stabilization of the resulting vector $\bar{g}_K^\lambda(\alpha)$ (27). Since the basis functions $h_k(\tau, \alpha)$ (3) and (4) are such that $h_k(\tau, \alpha) \leq 1$ for any arguments, the following inequality

$$\max_{\tau \geq 0} |H_K(\tau, \alpha)| \leq \sum_{k=0}^{K-1} |g_k|$$

holds for an arbitrary time-scale factor, which means that the smoothing of the vector of model parameters results in the limitation of the respective relaxation spectrum. The mechanism of the vector $\bar{g}_K^\lambda(\alpha)$ (27) stabilization via Tikhonov regularization is explained in many papers, for example, [20,40,55]. The following rule holds: the greater the regularization parameter λ is, the more highly bounded the fluctuations of the vector $\bar{g}_K^\lambda(\alpha)$ are; see [40].

The norm $\|H_K(\tau, \alpha)\|_2$ is a measure of smoothing of the relaxation spectrum model, where $\|\cdot\|_2$ also means the square norm in $L^2(0, \infty)$. In Appendix A.7, the following proposition is proved:

Proposition 1. *For an arbitrary time-scale factor α and arbitrary vector of model parameters g_K, we have*

$$\|H_K(\tau,\alpha)\|_2^2 = g_K^T \Gamma(\alpha) g_K = \frac{1}{2\alpha} g_K^T \Gamma_1 g_K, \tag{52}$$

where $K \times K$ symmetric real matrices, $\Gamma(\alpha)$ described by (A39) and Γ_1 given by

$$
\Gamma_1 = \begin{bmatrix}
1 & \frac{e}{2} & \cdots & \left(\frac{e}{2}\right)^j \frac{j!}{j^j} & \cdots & \left(\frac{e}{2}\right)^{K-1} \frac{(K-1)!}{(K-1)^{K-1}} \\
\frac{e}{2} & 2\left(\frac{e}{2}\right)^2 & \cdots & \left(\frac{e}{2}\right)^{1+j} \frac{(1+j)!}{j^j} & \cdots & \left(\frac{e}{2}\right)^K \frac{(K)!}{(K-1)^{K-1}} \\
\vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\
\left(\frac{e}{2}\right)^k \frac{k!}{k^k} & \left(\frac{e}{2}\right)^{k+1} \frac{(k+1)!}{k^k} & \cdots & \left(\frac{e}{2}\right)^{k+j} \frac{(k+j)!}{k^k \, j^j} & \cdots & \left(\frac{e}{2}\right)^{k+K-1} \frac{(k+K-1)!}{k^k \, (K-1)^{K-1}} \\
\vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\
\left(\frac{e}{2}\right)^{K-1} \frac{(K-1)!}{(K-1)^{K-1}} & \left(\frac{e}{2}\right)^{K-1+j} \frac{(K)!}{(K-1)^{K-1}} & \cdots & \left(\frac{e}{2}\right)^{K-1+j} \frac{(K-1+j)!}{(K-1)^{K-1} \, j^j} & \cdots & \left(\frac{e}{2}\right)^{2K-2} \frac{(2K-2)!}{(K-1)^{2(K-1)}}
\end{bmatrix} \tag{53}
$$

are a positive definite; matrix $\Gamma(\alpha)$ for an arbitrary time-scale factor $\alpha > 0$.

The matrix Γ_1 is independent of the time-scale factor; only the multiplier $\frac{1}{2\alpha}$ in the last expression of (52) depends on α. Since for any symmetric non-negative definite matrix, the eigenvalues and singular values are identical, by virtue of the Rayleigh–Ritz inequalities [58] (Lemma I):

$$\lambda_{min}(X) x^T x \leq x^T X x \leq \lambda_{max}(X) x^T x, \tag{54}$$

which hold for any $x \epsilon \mathcal{R}^m$ and any symmetric matrix $X = X^T \epsilon \mathcal{R}^{m,m}$, where $\lambda_{min}(X)$ and $\lambda_{max}(X)$ are minimal and maximal eigenvalues of the matrix X, Equation (52) implies the following estimations:

$$\frac{1}{2\alpha} \sigma_{min}(\Gamma_1) \|g_K\|_2^2 \leq \|H_K(\tau,\alpha)\|_2^2 \leq \frac{1}{2\alpha} \sigma_1(\Gamma_1) \|g_K\|_2^2,$$

where $\sigma_1(\Gamma_1)$ and $\sigma_{min}(\Gamma_1)$ denote the largest and the minimal singular values of matrix Γ_1 (53). Thus, the next result is derived.

Proposition 2. *For an arbitrary time-scale factor α and arbitrary vector of model parameters g_K, the following inequalities hold:*

$$\frac{1}{\sqrt{2\alpha}} \sqrt{\sigma_{min}(\Gamma_1)} \|g_K\|_2 \leq \|H_K(\tau,\alpha)\|_2 \leq \frac{1}{\sqrt{2\alpha}} \sqrt{\sigma_1(\Gamma_1)} \|g_K\|_2. \tag{55}$$

The square roots of the singular values $\sigma_1(\Gamma_1)$ and $\sigma_{min}(\Gamma_1)$ for $K = 5, 6, \ldots 12$ are summarized in Table 2. However, the lower bound of this norm is useful only for small K and small time-scale factors.

Table 2. The square roots of the largest $\sigma_1(\Gamma_1)$ and minimal $\sigma_{min}(\Gamma_1)$ singular values of the matrix Γ_1 (53) for $K = 4, 5, \ldots 12$ model summands.

K	4	5	6	7	8	9	10	11	12
$\sqrt{\sigma_1(\Gamma_1)}$	3.666396	4.325185	4.927061	5.485792	6.010247	6.506534	6.979079	7.431223	7.865567
$\sqrt{\sigma_{min}(\Gamma_1)}$	0.206481	0.087523	0.034642	0.013243	0.004953	0.001824	6.6358×10^{-4}	2.3925×10^{-4}	8.5615×10^{-5}

Since $\sqrt{\sigma_1(\Gamma_1)}$ grows with K, the greater the number of model summands there are, the greater the time-scaling factor should be to achieve pre-assumed multiplier $\frac{1}{\sqrt{2\alpha}} \sqrt{\sigma_1(\Gamma_1)}$ in the estimation (55). In [59], a decreasing sequence of upper bounds for the largest singular

value of a non-negative definite square matrix is constructed, given by Equation (19) in [59]. This result applied to the $K \times K$ matrix Γ_1 means that

$$\psi_n(\Gamma_1) = \frac{tr(\Gamma_1)}{K} + \left[\frac{(K-1)^{2n-1}}{(K-1)^{2n-1}+1} tr \left[\left(\Gamma_1 - \frac{tr(\Gamma_1)}{K} \mathbb{I}_{K,K} \right)^{2n} \right] \right]^{\frac{1}{2n}}, \tag{56}$$

is a decreasing sequence of upper bounds for $\sigma_1(\Gamma_1)$. The right inequality in (55) can be weakened to the following:

Proposition 3. *For an arbitrary time-scale factor α and arbitrary vector of model parameters g_K, the sequence of inequalities hold:*

$$\|H_K(\tau,\alpha)\|_2 \leq \frac{1}{\sqrt{2\alpha}} \sqrt{\psi_n(\Gamma_1)} \|g_K\|_2 \tag{57}$$

for $n = 1, 2, \ldots$, with the coefficients $\psi_n(\Gamma_1)$ (56).

In Table A2 in the Appendix B, the sequences $\sqrt{\psi_n(\Gamma_1)}$ are summarized for $n = 1, 2, \ldots, 10$ and $K = 4, 5, \ldots 12$ model summands; for comparison, in the last row, $\sqrt{\sigma_1(\Gamma_1)}$ are given. The sequence $\sqrt{\psi_n(\Gamma_1)}$ quickly decreases to $\sqrt{\sigma_1(\Gamma_1)}$; already, the sixth to eighth estimates equal $\sqrt{\sigma_1(\Gamma_1)}$. However, it should be remembered that the right inequality in (55) and (57) only give the upper bounds of the norm $\|H_K(\tau,\alpha)\|_2$.

To summarize, the smoothness of the optimal solution $\bar{g}_K^\lambda(\alpha)$ (27) of discrete regularized problem (26) guarantees that the fluctuations in the respective spectrum of relaxation, in particular the resulting spectrum of relaxation $\widehat{H}_K(\tau)$ (46), are also bounded. The time-scale factor α also affects the smoothness of the spectrum model.

3.6.2. Convergence and Noise Robustness

The relaxation spectrum recovery from experimental data is an inverse ill-posed problem, in which the identification index refers to the measured relaxation modulus, but not directly to the unknown relaxation spectrum $H(\tau)$. Therefore, we cannot estimate the model error $\|H(\tau) - H_K(\tau,\alpha)\|_2$ directly. As a reference point for the model $H_K(\tau,\alpha)$ (6) and (27), we will consider the model of the spectrum that we would obtain for the same time-scale factor α and regularization parameter λ on the basis of ideal (undisturbed) measurements of the relaxation modulus:

$$\widetilde{H}_K(\tau,\alpha) = \sum_{k=0}^{K-1} g_K^\lambda(\alpha) h_k(\tau,\alpha), \tag{58}$$

where $g_K^\lambda(\alpha)$ is the vector of the regularized solution of (26) given by (compare (27))

$$g_K^\lambda(\alpha) = \left(\Phi_{N,K}^T(\alpha) \Phi_{N,K}(\alpha) + \lambda \mathbb{I}_{K,K} \right)^{-1} \Phi_{N,K}^T(\alpha) G_N \tag{59}$$

for noise-free measurements of relaxation modulus $G_N = \begin{bmatrix} G(t_1) & \cdots & G(t_N) \end{bmatrix}^T$; c.f., Equation (24). In Appendix A.8, the following estimations are derived:

Proposition 4. *For an arbitrary time-scale factor α and arbitrary regularization parameter λ, the error between the relaxation time spectrum models $H_K(\tau,\alpha)$ (6) and (27) and $\widetilde{H}_K(\tau,\alpha)$ (58) and (59) is estimated using the following inequality:*

$$\|H_K(\tau,\alpha) - \widetilde{H}_K(\tau,\alpha)\|_2 \leq \frac{1}{\sqrt{2\alpha}} \sqrt{\sigma_1(\Gamma_1)} \left[\sum_{i=1}^{r(\alpha)} \frac{[\sigma_i(\alpha)]^4}{\left\{ [\sigma_i(\alpha)]^2 + \lambda \right\}^4} \right]^{\frac{1}{4}} \|z_N\|_2 \tag{60}$$

where $z_N = \begin{bmatrix} z(t_1) & \cdots & z(t_N) \end{bmatrix}^T$ is the vector of measurement noises.

According to inequality (60), the accuracy of the spectrum approximation depends both on the measurement noises and regularization parameter and on the singular values $\sigma_1(\alpha), \ldots, \sigma_{r(\alpha)}(\alpha)$ of the matrix $\boldsymbol{\Phi}_{N,K}(\alpha)$ (47) depending on the time-scale factor. By (60), the spectrum $H_K(\tau, \alpha)$ tends to the noise-free spectrum $\widetilde{H}_K(\tau, \alpha)$ in each relaxation time τ, at which they are both continuous, linearly with respect to the norm $\|z_N\|_2$, as $\|z_N\|_2 \to 0$.

3.7. Example

Consider a viscoelastic material of the relaxation spectrum described by the double-mode Gauss-like distribution:

$$H(\tau) = \left[\beta_1 e^{-\left(\frac{1}{\tau} - m_1\right)^2 / q_1} + \beta_2 e^{-\left(\frac{1}{\tau} - m_2\right)^2 / q_2} \right] / \tau, \tag{61}$$

where the parameters are as follows: $\beta_1 = 467$ Pa·s, $m_1 = 0.0037$ s^{-1}, $q_1 = 1.124261 \times 10^{-6}$ s^{-2}, $\beta_2 = 39$ Pa·s, $m_2 = 0.045$ s^{-1} and $q_2 = 1.173 \times 10^{-3}$ s^{-2}. Spectra of this type are tested at the stage of developing new identification methods, for example, in [23] (Figure 2), [24] (Figures 9, 11 and 17) and [25] (Figures 2, 3, 6, 7–11 and 14), because they describe the viscoelastic properties of various materials, in particular: polymers [60] (Figures 4b and 8b), polyacrylamide gels [28] (Figure A4), glass [61] (Figure 2), wood [13] and sugar beet [18] (Figure 2). It is shown in Appendix A.9 that the corresponding 'real' relaxation modulus is

$$\begin{aligned}
G(t) = \frac{\sqrt{\pi}}{2} \Bigg[&\beta_1 \sqrt{q_1} \, e^{\frac{1}{4}t^2 q_1 - m_1 t} erfc\left(\frac{\frac{1}{2}tq_1 - m_1}{\sqrt{q_1}} \right) + \\
&\beta_2 \sqrt{q_2} \, e^{\frac{1}{4}t^2 q_2 - m_2 t} erfc\left(\frac{\frac{1}{2}tq_2 - m_2}{\sqrt{q_2}} \right) \Bigg],
\end{aligned} \tag{62}$$

where the complementary error function $erfc(x)$ is defined by [62]:

$$erfc(x) = \frac{2}{\sqrt{\pi}} \int_x^{\infty} e^{-z^2} dz. \tag{63}$$

In the experiment, $N = 5000$ sampling instants t_i were generated with the constant period in the time interval $\mathcal{T} = [0, 1550]$ seconds selected in view of the course of the modulus $G(t)$ (62). Additive measurement noises $z(t_i)$ were selected independently by random choice with uniform distribution on the interval $[-0.005, 0.005]$ Pa. The 'real' spectrum (61), modulus (62) and the basis functions $h_k(\tau, \alpha)$, $\phi_k(t, \alpha)$ were simulated in Matlab R2022a using the special functions *besselk* and *erfc*. For the singular value decomposition procedure, *svd* was applied. New calculation algorithms of the modified Bessel function of the second kind are constantly being developed; recently, an algorithm for parallel calculation was proposed [63].

3.7.1. Optimal Models

For $K = 3, 4, \ldots, 12$, the optimal time-scaling factors α_{opt} were determined via the proposed two-level identification scheme and are given in Table 3 together with the related regularization parameters $\lambda_{GCV}(\alpha_{opt})$. Next, the vectors of optimal model parameters $\widehat{g}_K = \widehat{g}_K^{\lambda_{GCV}}(\alpha_{opt})$ (45) were computed and are given in Table A3 in Appendix B; the elements of these vectors are both negative and positive. The square norms $\|\widehat{g}_K\|_2$ and $\|H_{Kopt}(\tau)\|_2$ are also enclosed in Table 3 as the measures of the solution smoothness. The norm $\|H_{Kopt}(\tau)\|_2$ of the optimal model $H_{Kopt}(\tau)$ (46) was determined through (52) based on $\widehat{g}_K$. For the 'real' spectrum $H(\tau)$ (61), the norm $\|H(\tau)\|_2 = 19.2562$ Pa. The approximation error between $H(\tau)$ (61) and their models $H_{Kopt}(\tau)$ (46) was estimated via relative integral error $ER1(K)$, defined by:

$$ER1(K)^2 = \frac{\|H(\tau) - H_{Kopt}(\tau)\|_2^2}{\|H(\tau)\|_2^2} = \frac{\int_0^{\infty} \left[H(\tau) - H_{Kopt}(\tau) \right]^2 d\tau}{\int_0^{\infty} H(\tau)^2 d\tau}, \tag{64}$$

which is expressed in a percentage in the penultimate column of Table 3. For the bimodal spectrum, the values of this error of the order of 33% are not surprising in the context of the ill-conditioned inverse problem. To compare the approximations $H_{Kopt}(\tau)$ (46) obtained for successive integer K, the square index $ER2(K)$ defined by the distance:

$$ER2(K) = \|H_{(K+1)opt}(\tau) - H_{Kopt}(\tau)\|_2^2 = \int_0^\infty \left[H_{(K+1)opt}(\tau) - H_{Kopt}(\tau) \right]^2 d\tau, \quad (65)$$

is applied; see the last column of Table 3. The optimal indices $Q_{Nopt}(\alpha_{opt})$ (37) are given in Table 3, too. The optimal models $H_{Kopt}(\tau)$ (46) and the 'real' spectrum $H(\tau)$ (61) are plotted in Figure 4 for $K = 3, 4, \dots 12$. The analysis of the data from Table 3 and, in particular, the analysis of Figure 4b–d show that increasing the number of model components above $K = 8$ does not significantly affect the course of the model $H_{Kopt}(\tau)$ and its accuracy. This is emphasized by the values of the $ER1(K)$ (64) and $ER2(K)$ (65) indices; in particular, the integral square error $ER2(K)$ between successive $H_{Kopt}(\tau)$ and $H_{(K+1)opt}(\tau)$ spectrum approximations, which relates to the square norm $\|H(\tau)\|_2^2$ for $K \geq 8$ does not exceed 0.046%, and for $K \geq 10$, it is of the order of 0.009%.

Table 3. The parameters of the optimal models in the example: optimal time-scale factors α_{opt}, upper α_{min} and lower α_{max} bounds of the interval $[\alpha_{min}, \alpha_{max}]$ of suboptimal time-scale factors defined by inequality (66) for $\varepsilon_\alpha = 0.1$, regularization parameters $\lambda_{GCV}(\alpha_{opt})$, optimal identification indices $Q_{Nopt}(\alpha_{opt})$ defined in (37), square norms: $\|\hat{g}_K\|_2$ of the vector of optimal model parameter $\hat{g}_K$ (45) and $\|H_{Kopt}(\tau)\|_2$ of the optimal relaxation spectrum model $H_{Kopt}(\tau)$ (46), relaxation spectrum approximation error $ER1(K)$ defined in (64) and the distance between successive approximations $H_{Kopt}(\tau)$ measured by $ER2(K)$ (65).

K	α_{opt} [s^{-1}]	α_{min} [s^{-1}]	α_{max} [s^{-1}]	$\lambda_{GCV}(\alpha_{opt})$ [−]	$Q_{Nopt}(\alpha_{opt})$ [Pa^2]	$\|\hat{g}_K\|_2$ [Pa]	$\|H_{Kopt}(\tau)\|_2$ [Pa]	$ER1(K)$ [%]	$ER2(K)$ [Pa$^2 \cdot$s]
3	0.00520	0.00205	0.0080	0.0445	8.63505×10^{-4}	0.7055	13.0248	90.459	12.876
4	0.01675	0.0164	0.01745	0.0063	3.43945×10^{-5}	7.3485	16.1607	39.160	3.1049
5	0.02025	0.0192	0.0213	0.0071	2.71552×10^{-5}	4.6724	17.1465	35.638	2.4592
6	0.02375	0.0220	0.0255	0.0078	2.48511×10^{-5}	3.6493	18.0443	33.986	0.7563
7	0.02655	0.0234	0.0290	0.0083	2.48256×10^{-5}	2.8846	18.5088	33.364	0.1432
8	0.02865	0.0234	0.0318	0.0089	2.51617×10^{-5}	2.3555	18.5679	32.824	0.1692
9	0.03005	0.0241	0.03425	0.0099	2.52412×10^{-5}	2.0639	18.3756	32.701	0.0111
10	0.03215	0.0255	0.0367	0.0109	2.51143×10^{-5}	1.9058	18.4224	32.631	0.0336
11	0.03390	0.0276	0.03915	0.0122	2.48521×10^{-5}	1.8020	18.3498	32.879	0.0327
12	0.03670	0.02935	0.04195	0.0127	2.44452×10^{-5}	1.7198	18.4432	32.919	0.0328

In Figure 5, the optimal models of the relaxation modulus $G_K(t, \alpha_{opt})$ computed for $\hat{g}_K$ (45) according to (7) are plotted, where the measurements $\overline{G}(t_i)$ of the 'real' modulus $G(t)$ (62) are also marked. The optimal models $G_K(t, \alpha_{opt})$ have been well fitted to the experimental data, as indicated by the mean-square model errors $Q_{Nopt}(\alpha_{opt})/N$, which for $8 \leq K \leq 12$ vary in the range from $4.889 \cdot 10^{-9}$ Pa^2 to $5.0482 \cdot 10^{-9}$ Pa^2. Thus, models $G_K(t, \alpha_{opt})$ for different K practically coincide with the measurement points and with each other; see Figure 5.

3.7.2. Optimal and Sub-Optimal Time-Scale Factors

The identification index $Q_{Nopt}(\alpha)$ (35) minimized in the upper-level task (37) as a function of time-scale factor α is plotted in Figure 6a–d for $K = 6, 8, 10, 12$. The optimal parameters α_{opt} are marked. The analysis of these plots suggests that there is such a neighborhood of α_{opt}, namely a closed interval $[\alpha_{min}, \alpha_{max}]$, that $\alpha_{min} \leq \alpha_{opt} \leq \alpha_{max}$ and for each $\alpha \in [\alpha_{min}, \alpha_{max}]$, the identification index $Q_{Nopt}(\alpha)$ differs from the minimal $Q_{Nopt}(\alpha_{opt})$ not more than $\varepsilon_\alpha \cdot Q_{Nopt}(\alpha_{opt})$, i.e.,

$$Q_{Nopt}(\alpha) - Q_{Nopt}(\alpha_{opt}) \leq \varepsilon_\alpha \cdot Q_{Nopt}(\alpha_{opt}), \quad (66)$$

where ε_α is a small positive number. Inequality (66) means the deterioration of the model error is not greater than ε_α percent of the optimal $Q_{Nopt}(\alpha_{opt})$. Any parameter α from the

interval $[\alpha_{min}, \alpha_{max}]$ is a suboptimal time-scale factor. In Figure 6, $\varepsilon_\alpha = 0.1$, which means a 10% level of sub-optimality, is assumed, and the values of α_{min} and α_{max} are marked on the small subfigures. They are also given in Table 3. With the increase in the number of model components, the optimal parameter α_{opt} increases and the range of sub-optimal time-scale factors shifts. For an exemplary number of model components, $K = 9$, the optimal model $H_{Kopt}(\tau)$ (46) and models $H_K(\tau, \alpha)$ (6), optimal in the sense of lower-level task (28) for suboptimal parameters α_{min} and α_{max}, are plotted in Figure 7a; the 'real' spectrum $H(\tau)$ (61) is presented, too. The corresponding optimal regularization parameters are as follows: $\lambda_{GCV}(\alpha_{min}) = 0.0167$ and $\lambda_{GCV}(\alpha_{max}) = 0.0094$. For respective vectors of optimal model parameters, we have $\|\widehat{g}_K^{\lambda_{GCV}}(\alpha_{min})\|_2 = 2.5114$ Pa and $\|\widehat{g}_K^{\lambda_{GCV}}(\alpha_{max})\|_2 = 2.5119$ Pa. The norms of the relaxation spectra are $\|H_K(\tau, \alpha_{min})\|_2 = 17.4472$ Pa and $\|H_K(\tau, \alpha_{max})\|_2 = 20.2069$ Pa. In Figure 7b, the 'real' modulus $G(t)$ (62) and the models: $G_K(t, \alpha_{opt})$ (7) and (45) and $G_K(t, \alpha_{min})$, $G_K(t, \alpha_{max})$ computed for $\widehat{g}_K^{\lambda_{GCV}}(\alpha_{min})$, $\widehat{g}_K^{\lambda_{GCV}}(\alpha_{max})$ according to (7) are plotted. However, the relaxation moduli are almost identical (Figure 7b), the spectrum models differ (Figure 7a), which emphasizes the importance of the time-scale factor optimal selection. Too strong smoothing of the relaxation spectrum models, with a simultaneous very good fit to the experimental data of the relaxation modulus models, indicates the need to modify the quality index in the regularized task (26). For example, the square term $\|g_K\|_2^2$ in the second component of the objective function can be replaced by the quadratic form $g_K^T W g_K$, and a positive definite weight matrix W or regularized weighted least-squares approach [41] can be applied. This will be the subject of further work.
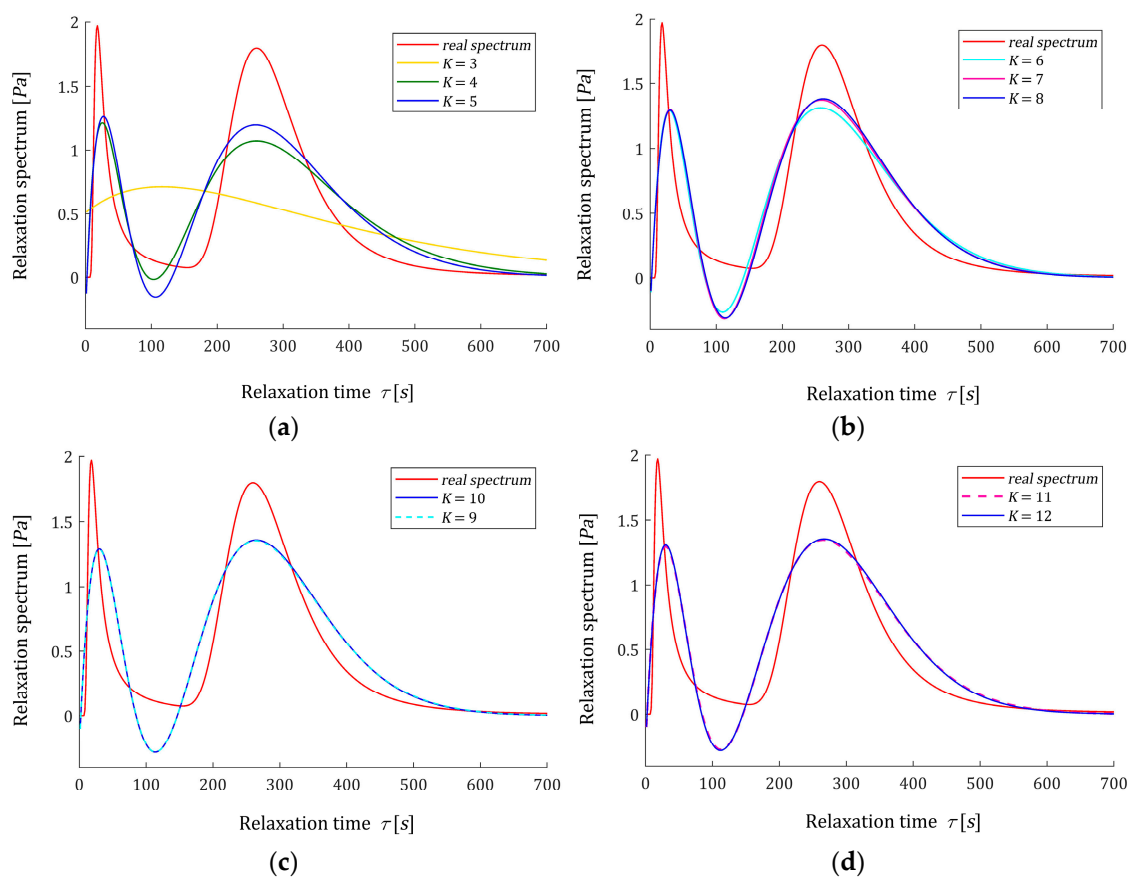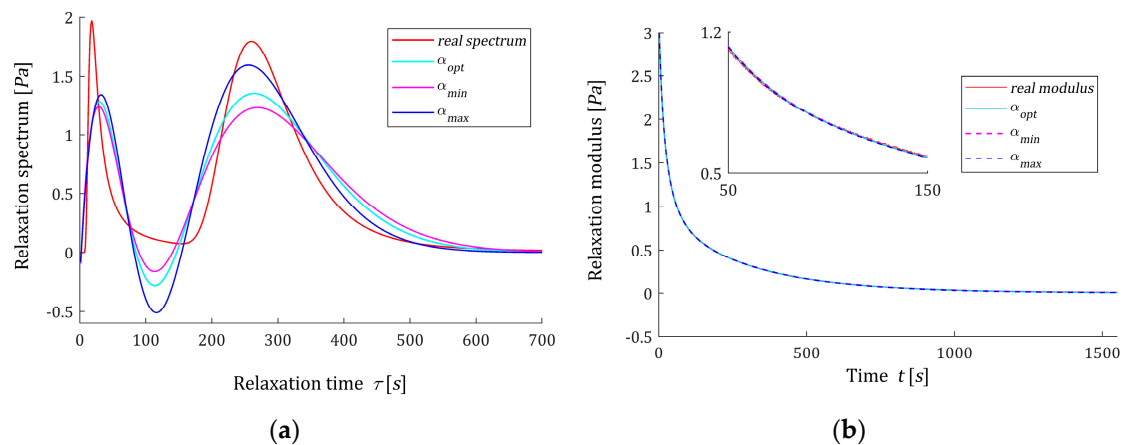

Figure 4. Relaxation time spectrum $H(\tau)$ (61) (solid red line) from the example and the corresponding models $H_{Kopt}(\tau)$ (46) for K summands of the model: (**a**) $K = 3, 4, 5$; (**b**) $K = 6, 7, 8$; (**c**) $K = 9, 10$; (**d**) $K = 11, 12$.
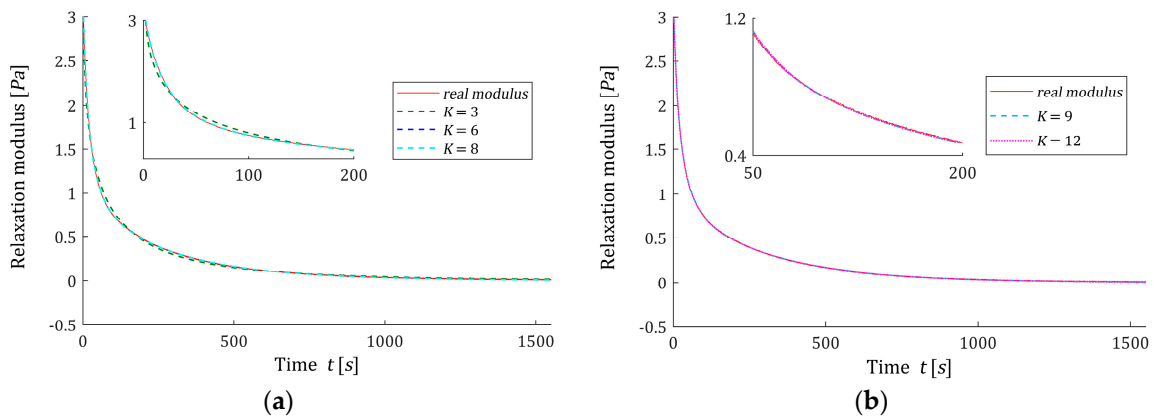
Figure 5. The measurements $\overline{G}(t_i)$ of the 'real' relaxation modulus $G(t)$ (62) (red points) from the example and the optimal approximated models $G_K(t, \alpha_{opt})$ (7) computed for $\widehat{g}_K$ (45) for K summands of the model: (**a**) $K = 3, 6, 8$; (**b**) $K = 9, 12$.
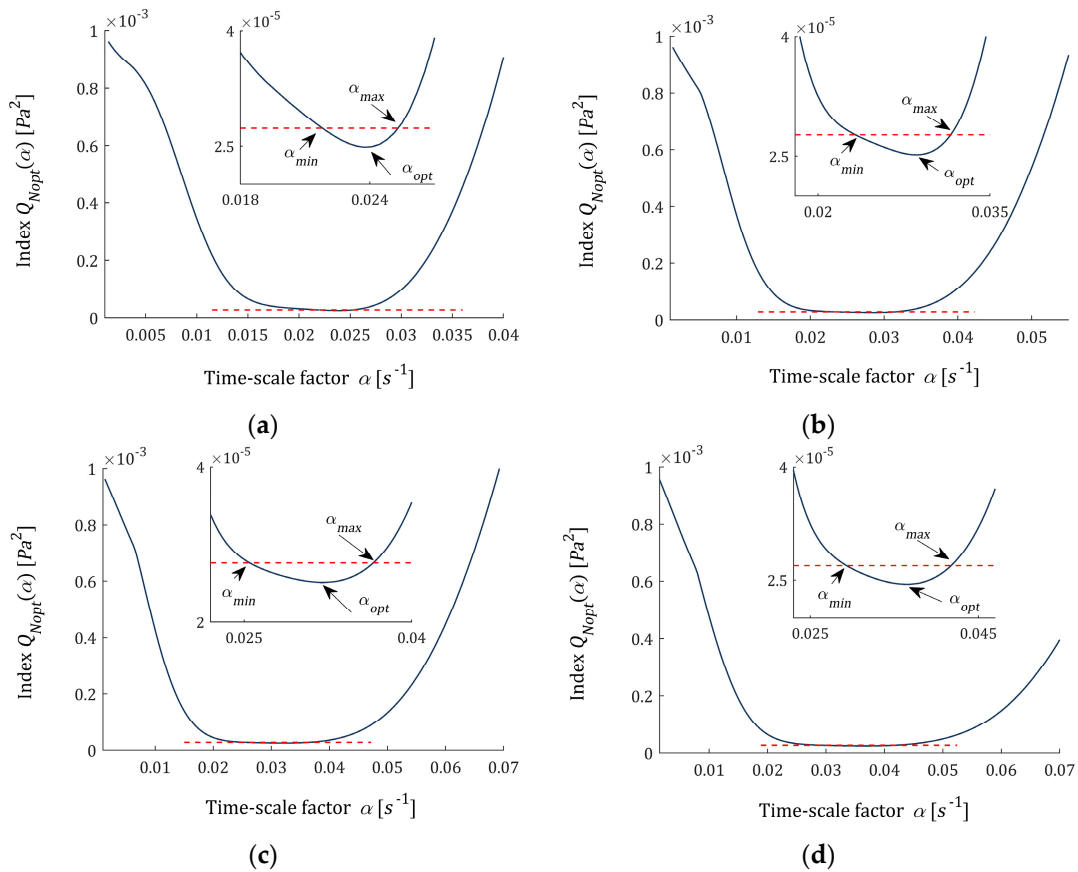
Figure 6. The identification index $Q_{Nopt}(\alpha)$ (35) (solid navy blue line) minimized in the upper level task (37) from the example for: (**a**) $K = 6$, (**b**) $K = 8$, (**c**) $K = 10$, (**d**) $K = 12$ summands of the relaxation spectrum model (6); on small subfigures, dashed red line of the value $(1 + \varepsilon_\alpha)Q_{Nopt}(\alpha_{opt})$ determines the interval $[\alpha_{min}, \alpha_{max}]$ of suboptimal time-scale factors defined by (66); α_{opt} is the optimal time-scale factor solving problem (37) for the sub-optimality factor $\varepsilon_\alpha = 10\%$.

Figure 7. Relaxation time spectra and moduli from the example and the corresponding models for $K = 9$ summands of the model: (**a**) 'real' spectrum $H(\tau)$ (61) (solid red line) and the models: optimal $H_{Kopt}(\tau)$ (46) and suboptimal $H_K(\tau, \alpha_{min})$ and $H_K(\tau, \alpha_{max})$; (**b**) 'real' modulus $G(t)$ (62) (red points), the optimal approximated model $G_K(t, \alpha_{opt})$ (7) and (45), the suboptimal models $G_K(t, \alpha_{min})$ and $G_K(t, \alpha_{max})$ computed for $\widehat{g}_K^{\lambda_{GCV}}(\alpha_{min})$ and $\widehat{g}_K^{\lambda_{GCV}}(\alpha_{max})$ according to (7).

4. Conclusions

In this paper, a new robust hierarchical algorithm for the identification of the relaxation time spectrum, which combines the technique of an expansion of a function into a series of basis functions with the least-squares regularized identification and the optimal choice of the time-scale factor, has been derived. The task of determining the best 'regularized' model was solved at the lower level, while the optimal time-scale factor was selected on the upper level of the identification scheme. The continuous spectrum of relaxation times was approximated by finite series of power–exponential basis functions, while the components of the relaxation modulus model were proven to be described by the product of power of time and the modified Bessel function of the second kind. In the present paper, the problem of the optimal choice of the time-scale factor to ensure the best fit of the model to experimental data has been formulated and solved for the first time in the context of the relaxation spectrum identification.

The necessary optimality conditions both for the optimal regularized least-squares identification task and the problem of the optimal selection of the time-scale factor were derived in the form of nonlinear algebraic equations. The main properties of the basis functions of the relaxation spectrum and modulus models, their positive definiteness, convenient upper bounds, monotonicity, asymptotic properties and wide range of applicability for different time-scale factors indicated the possibility of using the proposed model and identification algorithm to determine the spectrum of a wide class of viscoelastic materials.

The overly strong smoothing of the relaxation spectrum models in the example, with a very good fit to the experimental data of the relaxation modulus models, indicates the need to modify the quality index in the lower level identification task. An introduction of a weight matrix in the second component of the objective function or a direct application of regularized weighted least-squares should be investigated. Another solution is the selection of such a spectrum model which guarantees the assumed level of smoothing and optimal adjustment to the relaxation modulus measurement data. These approaches will be the subject of further work.

The presented scheme of the relaxation spectrum identification can be easily modified for retardation spectrum recovery from creep compliance measurements obtained in the standard creep test.

Institutional Review Board Statement: Not applicable.

Appendix A

Appendix A.1. Proof of Theorem 1

To prove the theorem, we first derive the following integral formula:

$$\mathcal{I} = \int_0^\infty \tau^{v-1} e^{-\alpha\tau} e^{-t/\tau} d\tau = 2\left(\sqrt{\frac{t}{\alpha}}\right)^v K_v\left(2\sqrt{\alpha t}\right), \tag{A1}$$

where $v \in \mathcal{R}$.

By applying in the integral of the left-hand side of (A1) the substitution $\sqrt{\alpha}\,\tau = \sqrt{t} e^y$ we have:

$$\mathcal{I} = \int_{-\infty}^\infty \left(\sqrt{\frac{t}{\alpha}}\right)^v e^{y(v-1)} e^{-\sqrt{\alpha t}\,e^y} e^{-\sqrt{\alpha t}\,e^{-y}} e^y dy,$$

which can be expressed as

$$\mathcal{I} = \left(\sqrt{\frac{t}{\alpha}}\right)^v \int_{-\infty}^\infty e^{vy} e^{-2\sqrt{\alpha t}\,\cosh(y)} dy. \tag{A2}$$

Since, bearing in mind that $\cosh(y)$ is an even function, the integral in (A2) is rewritten as

$$\mathcal{I} = \left(\sqrt{\frac{t}{\alpha}}\right)^v \left[\int_0^\infty e^{-vy}\,e^{-2\sqrt{\alpha t}\,\cosh(y)} dy + \int_0^\infty e^{vy}\,e^{-2\sqrt{\alpha t}\,\cosh(y)} dy\right],$$

or equivalently

$$\mathcal{I} = 2\left(\sqrt{\frac{t}{\alpha}}\right)^v \int_0^\infty e^{-2\sqrt{\alpha t}\,\cosh(y)} \cosh(vy) dy,$$

whence, by the following integral representation formula for the Bessel functions [45,46] (Equation (5) on p. 181 in [46]):

$$K_v(x) = \int_0^\infty e^{-x\,\cosh(y)} \cosh(vy) dy,$$

where $x > 0$ and $v \in \mathcal{R}$, Formula (A1) directly follows.

Now, we prove Formula (9). Since, by (3) and (8) the basis functions

$$\phi_k(t, \alpha) = \left(\frac{\alpha}{k}\right)^k e^k \int_0^\infty \tau^{k-1} e^{-\alpha\tau} e^{-t/\tau} d\tau$$

from (A1), we have

$$\phi_k(t, \alpha) = 2\left(\frac{\alpha}{k}\right)^k e^k \left(\sqrt{\frac{t}{\alpha}}\right)^k K_k\left(2\sqrt{\alpha t}\right),$$

whence, Formula (9) follows. $\square$

Appendix A.2. Algebraic Matrix Properties

In this appendix, a new differential matrix property, used to obtain the optimality conditions and to simplify numerical computations, is derived. First, for convenience, some matrix identities and inequalities known in the literature are also provided.

Appendix A.2.1. Matrix Identities and Inequalities

The following identities hold for arbitrary matrices A, B, C and D [57] (Equations (6.4.2)), (2.1.1) and (2.1.4), in order:

$$tr[AB] = tr[BA], \tag{A3}$$

$$(AB)^{-1} = B^{-1}A^{-1}, \tag{A4}$$

$$(BC + A)^{-1} = A^{-1} - A^{-1}B\left(I + CA^{-1}B\right)^{-1}CA^{-1}, \tag{A5}$$

provided that respective matrices are invertible.

The following differential properties hold for arbitrary differentiable matrix functions $A(x)$ and $B(x)$ [57] (Equations (P2.1.2a) and (P2.1.2b), respectively):

$$\frac{\partial}{\partial x}[A(x)B(x)] = \frac{\partial A(x)}{\partial x}B(x) + A(x)\frac{\partial B(x)}{\partial x}, \tag{A6}$$

$$\frac{\partial}{\partial x}\left[A(x)^{-1}\right] = -A(x)^{-1}\frac{\partial A(x)}{\partial x}A(x)^{-1}, \tag{A7}$$

assuming that matrix $A(x)$ is invertible.

Appendix A.2.2. New Matrix Property

Let us now prove the following result.

Proposition A1. *Let $A(x)$ be a symmetric positive definite $m \times m$ matrix function differentiable with respect to the real argument x. Then*

$$\frac{\partial tr\left[A(x)^{-1}\right]}{\partial x} = -tr\left[A(x)^{-1}\frac{\partial A(x)}{\partial x}A(x)^{-1}\right]. \tag{A8}$$

Proof of Proposition A1. Since the trace of the matrix is a linear function, bearing in mind definition of the matrix derivative with respect to the scalar argument x [57] (Section 2.1.8), we have

$$\frac{\partial tr\left[A(x)^{-1}\right]}{\partial x} = tr\left[\frac{\partial}{\partial x}\left[A(x)^{-1}\right]\right],$$

which, in view of (A7), directly yields (A8); the proposition is proved. $\square$

Appendix A.3. Proof of Theorem 2

First, we express $\Xi(\lambda, \alpha)$ (30) in a more useful equivalent form. By the matrix identity (A5) for $\lambda > 0$, we have

$$\lambda\left(\boldsymbol{\Phi}_{N,K}(\alpha)\boldsymbol{\Phi}_{N,K}^{T}(\alpha) + \lambda\mathbb{I}_{N,N}\right)^{-1} = \mathbb{I}_{N,N} - \boldsymbol{\Phi}_{N,K}(\alpha)\left(\lambda\mathbb{I}_{K,K} + \boldsymbol{\Phi}_{N,K}^{T}(\alpha)\boldsymbol{\Phi}_{N,K}(\alpha)\right)^{-1}\boldsymbol{\Phi}_{N,K}^{T}(\alpha),$$

whereas, by (30) and bearing in mind the notation (32), we obtain

$$\Xi(\lambda, \alpha) = \lambda\left(\boldsymbol{\Psi}(\alpha) + \lambda\mathbb{I}_{N,N}\right)^{-1}. \tag{A9}$$

Since, by (30), the GCV functional $V_{GCV}(\lambda, \alpha)$ (29) is expressed as

$$V_{GCV}(\lambda, \alpha) = \frac{\overline{\boldsymbol{G}}_N^T \boldsymbol{\Xi}(\lambda, \alpha) \boldsymbol{\Xi}(\lambda, \alpha) \overline{\boldsymbol{G}}_N}{tr[\boldsymbol{\Xi}(\lambda, \alpha)]^2},$$

its derivative with respect to λ is given by

$$\frac{\partial V_{GCV}(\lambda, \alpha)}{\partial \lambda} = \frac{\frac{\partial}{\partial \lambda}\left[\overline{\boldsymbol{G}}_N^T \boldsymbol{\Xi}(\lambda, \alpha) \boldsymbol{\Xi}(\lambda, \alpha) \overline{\boldsymbol{G}}_N\right] tr[\boldsymbol{\Xi}(\lambda, \alpha)] - 2\overline{\boldsymbol{G}}_N^T \boldsymbol{\Xi}(\lambda, \alpha) \boldsymbol{\Xi}(\lambda, \alpha) \overline{\boldsymbol{G}}_N \frac{\partial tr[\boldsymbol{\Xi}(\lambda, \alpha)]}{\partial \lambda}}{tr[\boldsymbol{\Xi}(\lambda, \alpha)]^3}. \tag{A10}$$

Now, we determine the partial derivatives which appear in the nominator of the right-hand side of (A10). By Formula (A6), we have

$$\frac{\partial}{\partial \lambda}\left[\overline{\boldsymbol{G}}_N^T \boldsymbol{\Xi}(\lambda, \alpha) \boldsymbol{\Xi}(\lambda, \alpha) \overline{\boldsymbol{G}}_N\right] = 2\overline{\boldsymbol{G}}_N^T \boldsymbol{\Xi}(\lambda, \alpha) \frac{\partial \boldsymbol{\Xi}(\lambda, \alpha)}{\partial \lambda} \overline{\boldsymbol{G}}_N. \tag{A11}$$

Equation (A9) yields

$$\frac{\partial \boldsymbol{\Xi}(\lambda, \alpha)}{\partial \lambda} = (\boldsymbol{\Psi}(\alpha) + \lambda \mathbb{I}_{N,N})^{-1} + \lambda \frac{\partial}{\partial \lambda}\left[(\boldsymbol{\Psi}(\alpha) + \lambda \mathbb{I}_{N,N})^{-1}\right], \tag{A12}$$

where the derivative of the inverse matrix can be found due to property (A7), which, when applied to (A12), results in

$$\frac{\partial \boldsymbol{\Xi}(\lambda, \alpha)}{\partial \lambda} = (\boldsymbol{\Psi}(\alpha) + \lambda \mathbb{I}_{N,N})^{-1} - \lambda(\boldsymbol{\Psi}(\alpha) + \lambda \mathbb{I}_{N,N})^{-1}(\boldsymbol{\Psi}(\alpha) + \lambda \mathbb{I}_{N,N})^{-1},$$

whereas, after algebraic manipulations, we obtain

$$\frac{\partial \boldsymbol{\Xi}(\lambda, \alpha)}{\partial \lambda} = (\boldsymbol{\Psi}(\alpha) + \lambda \mathbb{I}_{N,N})^{-2} \boldsymbol{\Psi}(\alpha). \tag{A13}$$

Combining (A11), (A13) and (A9) yields

$$\frac{\partial}{\partial \lambda}\left[\overline{\boldsymbol{G}}_N^T \boldsymbol{\Xi}(\lambda, \alpha) \boldsymbol{\Xi}(\lambda, \alpha) \overline{\boldsymbol{G}}_N\right] = 2\lambda \overline{\boldsymbol{G}}_N^T (\boldsymbol{\Psi}(\alpha) + \lambda \mathbb{I}_{N,N})^{-3} \boldsymbol{\Psi}(\alpha) \overline{\boldsymbol{G}}_N. \tag{A14}$$

To find $\frac{\partial tr[\boldsymbol{\Xi}(\lambda, \alpha)]}{\partial \lambda}$, expression (A9) is used, which implies

$$tr[\boldsymbol{\Xi}(\lambda, \alpha)] = \lambda \, tr\left[(\boldsymbol{\Psi}(\alpha) + \lambda \mathbb{I}_{N,N})^{-1}\right], \tag{A15}$$

whence

$$\frac{\partial tr[\boldsymbol{\Xi}(\lambda, \alpha)]}{\partial \lambda} = tr\left[(\boldsymbol{\Psi}(\alpha) + \lambda \mathbb{I}_{N,N})^{-1}\right] + \lambda \frac{\partial tr\left[(\boldsymbol{\Psi}(\alpha) + \lambda \mathbb{I}_{N,N})^{-1}\right]}{\partial \lambda}. \tag{A16}$$

By Proposition A1 and Equation (A8), we have

$$\frac{\partial tr\left[(\boldsymbol{\Psi}(\alpha) + \lambda \mathbb{I}_{N,N})^{-1}\right]}{\partial \lambda} = -tr\left[(\boldsymbol{\Psi}(\alpha) + \lambda \mathbb{I}_{N,N})^{-2}\right];$$

therefore, derivative $\frac{\partial tr[\boldsymbol{\Xi}(\lambda, \alpha)]}{\partial \lambda}$ (A16) takes the form

$$\frac{\partial tr[\boldsymbol{\Xi}(\lambda, \alpha)]}{\partial \lambda} = tr\left[(\boldsymbol{\Psi}(\alpha) + \lambda \mathbb{I}_{N,N})^{-1}\right] - \lambda \, tr\left[(\boldsymbol{\Psi}(\alpha) + \lambda \mathbb{I}_{N,N})^{-2}\right]. \tag{A17}$$

Now, we can derive the necessary optimality condition. Since, in view of (A9), matrix $\boldsymbol{\Xi}(\lambda, \alpha)$ is positive definite for any $\lambda > 0$, derivative $\frac{\partial V_{GCV}(\lambda, \alpha)}{\partial \lambda}$ given by the fraction from

the right-hand side of (A10) is equal to zero, if and only if its nominator is equal zero to, i.e.,

$$\frac{\partial}{\partial \lambda}\left[\overline{G}_N^T \Xi(\lambda,\alpha)\Xi(\lambda,\alpha)\overline{G}_N\right]tr[\Xi(\lambda,\alpha)] = 2\overline{G}_N^T \Xi(\lambda,\alpha)\Xi(\lambda,\alpha)\overline{G}_N\frac{\partial tr[\Xi(\lambda,\alpha)]}{\partial \lambda}. \qquad (A18)$$

By (A18), (A14), (A9), and (A17), we have

$$2\lambda^2\overline{G}_N^T(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-3}\mathbf{\Psi}(\alpha)\overline{G}_N tr\left[(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-1}\right]=$$
$$2\lambda^2\overline{G}_N^T(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-2}\overline{G}_N\left\{tr\left[(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-1}\right]-\lambda\ tr\left[(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-2}\right]\right\},$$

which is equivalent to

$$\overline{G}_N^T(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-3}\mathbf{\Psi}(\alpha)\overline{G}_N tr\left[(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-1}\right]=\overline{G}_N^T(\mathbf{\Psi}(\alpha)+$$
$$\lambda\mathbb{I}_{N,N})^{-2}\overline{G}_N\left\{tr\left[(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-1}\right]-\lambda\ tr\left[(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-2}\right]\right\}, \qquad (A19)$$

whence, by standard algebraic manipulations, Equation (31) follows directly. The theorem is proved. $\square$

Appendix A.4. Derivation of the Formula (33)

To derive the Formula (33) describing the derivative $\frac{\partial V_{GCV}(\lambda,\alpha)}{\partial \lambda}$ directly as a function of the matrices $\mathbf{\Psi}(\alpha)$ and $\overline{G}_N$, let us consider the nominator of the right-hand side of (A10), i.e., the function

$$M(\lambda,\alpha) = \frac{\partial}{\partial \lambda}\left[\overline{G}_N^T\Xi(\lambda,\alpha)\Xi(\lambda,\alpha)\overline{G}_N\right]tr[\Xi(\lambda,\alpha)]-$$
$$2\overline{G}_N^T\Xi(\lambda,\alpha)\Xi(\lambda,\alpha)\overline{G}_N\frac{\partial tr[\Xi(\lambda,\alpha)]}{\partial \lambda},$$

which, by (A14), (A15), (A9) and (A17), is expressed as

$$M(\lambda,\alpha) = 2\lambda^2\overline{G}_N^T(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-3}\mathbf{\Psi}(\alpha)\overline{G}_N\ tr\left[(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-1}\right]-$$
$$2\lambda^2\overline{G}_N^T(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-2}\overline{G}_N\left\{tr\left[(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-1}\right]-\lambda\ tr\left[(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-2}\right]\right\}.$$

Through tedious algebraic manipulations, the above expression is rewritten as

$$M(\lambda,\alpha) = -2\lambda^3\overline{G}_N^T(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-3}\ \overline{G}_N tr\left[(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-1}\right]+$$
$$2\lambda^3\overline{G}_N^T(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-2}\overline{G}_N tr\left[(\mathbf{\Psi}(\alpha)+\lambda\mathbb{I}_{N,N})^{-2}\right]. \qquad (A20)$$

Now, (A10) combined with (A20) and (A15) yields Formula (33).

Appendix A.5. Proof of Theorem 3

We first prove Equation (39) for $k = 0$; next, Formula (38) is derived.
Based on (10), property (19) and bearing in mind that $K_1(x) = K_{-1}(x)$ [54], we have

$$\frac{d}{d\alpha}[\phi_0(t,\alpha)] = -\frac{\sqrt{t}}{\sqrt{\alpha}}\left[K_{-1}\left(2\sqrt{\alpha t}\right)+K_1\left(2\sqrt{\alpha t}\right)\right] = -2\frac{\sqrt{t}}{\sqrt{\alpha}}K_1\left(2\sqrt{\alpha t}\right), \qquad (A21)$$

whence, by (9) and the last equation in (A21), we immediately obtain Equation (39).
Since, for an arbitrary $k \geq 1$, the basis function $\phi_k(t,\alpha)$ (9) is expressed as

$$\phi_k(t,\alpha) = 2\left(\frac{e}{2k}\right)^k\left(2\sqrt{\alpha t}\right)^k K_k\left(2\sqrt{\alpha t}\right),$$

by Formula (18), we immediately have

$$\frac{d}{d\alpha}[\phi_k(t,\alpha)] = -2\frac{\sqrt{t}}{\sqrt{\alpha}}e^k\left(\frac{1}{2k}\right)^k\left(2\sqrt{\alpha t}\right)^k K_{k-1}\left(2\sqrt{\alpha t}\right),$$

which is rewritten as

$$\frac{d}{d\alpha}[\phi_k(t,\alpha)] = -2e\left(\frac{t}{k}\right)\left(\frac{k-1}{k}\right)^{k-1}e^{k-1}\left(\frac{\sqrt{\alpha t}}{k-1}\right)^{k-1}K_{k-1}\left(2\sqrt{\alpha t}\right),$$

whence, in view of (9), Formula (38) follows directly, which completes the proof. □

Appendix A.6. Proof of Theorem 4

Let us find the derivative of the function $Q_{Nopt}(\alpha)$. Based on (36) and property (A6), and bearing in mind definition (42), we have

$$\frac{dQ_{Nopt}(\alpha)}{d\alpha} = 2\lambda_{GCV}(\alpha)\frac{d\lambda_{GCV}(\alpha)}{d\alpha}\overline{\boldsymbol{G}}_N^T\boldsymbol{Y}(\alpha)^{-2}\overline{\boldsymbol{G}}_N+$$
$$2\lambda_{GCV}^2(\alpha)\overline{\boldsymbol{G}}_N^T\left[\frac{d}{d\alpha}\boldsymbol{Y}(\alpha)^{-1}\right]\boldsymbol{Y}(\alpha)^{-1}\overline{\boldsymbol{G}}_N,$$

whereas by property (A7), the next formula follows

$$\frac{dQ_{Nopt}(\alpha)}{d\alpha} = 2\left[\lambda_{GCV}(\alpha)\frac{d\lambda_{GCV}(\alpha)}{d\alpha}\overline{\boldsymbol{G}}_N^T - \lambda_{GCV}^2(\alpha)\overline{\boldsymbol{G}}_N^T\boldsymbol{Y}(\alpha)^{-1}\left[\frac{d}{d\alpha}\boldsymbol{Y}(\alpha)\right]\right]\boldsymbol{Y}(\alpha)^{-2}, \quad \text{(A22)}$$

where, by (42):

$$\frac{d}{d\alpha}\boldsymbol{Y}(\alpha) = \frac{d}{d\alpha}\boldsymbol{\Psi}(\alpha) + \frac{d\lambda_{GCV}(\alpha)}{d\alpha}\mathbb{I}_{N,N} = \boldsymbol{\Omega}_{N,K}(\alpha) + \frac{d\lambda_{GCV}(\alpha)}{d\alpha}\mathbb{I}_{N,N}. \quad \text{(A23)}$$

Substituting (A23) into (A22), we obtain

$$\frac{dQ_{Nopt}(\alpha)}{d\alpha} = 2\lambda_{GCV}(\alpha)\frac{d\lambda_{GCV}(\alpha)}{d\alpha}\overline{\boldsymbol{G}}_N^T\boldsymbol{Y}(\alpha)^{-2}\overline{\boldsymbol{G}}_N-$$
$$2\lambda_{GCV}^2(\alpha)\frac{d\lambda_{GCV}(\alpha)}{d\alpha}\overline{\boldsymbol{G}}_N^T\boldsymbol{Y}(\alpha)^{-3}\overline{\boldsymbol{G}}_N - 2\lambda_{GCV}^2(\alpha)\overline{\boldsymbol{G}}_N^T\boldsymbol{Y}(\alpha)^{-1}\boldsymbol{\Omega}_{N,K}(\alpha)\boldsymbol{Y}(\alpha)^{-2}\overline{\boldsymbol{G}}_N. \quad \text{(A24)}$$

Whence, if the parameter α_{opt} solves the minimization task (37), then $\lambda_{GCV}(\alpha_{opt})$ is a solution of the algebraic Equation (40), resulting directly from equating the derivative $\frac{dQ_{Nopt}(\alpha)}{d\alpha}$ (A24) to zero. The inspection of (40) shows that two derivatives, $\boldsymbol{\Omega}_{N,K}(\alpha) = \frac{d}{d\alpha}\boldsymbol{\Psi}(\alpha)$ introduced in (A23) and $\frac{d\lambda_{GCV}(\alpha)}{d\alpha}$, are necessary to define (40) well.

By virtue of (32) and (A6), we obtain

$$\frac{d}{d\alpha}\boldsymbol{\Psi}(\alpha) = \frac{d}{d\alpha}\boldsymbol{\Phi}_{N,K}(\alpha)\boldsymbol{\Phi}_{N,K}^T(\alpha) + \boldsymbol{\Phi}_{N,K}(\alpha)\frac{d}{d\alpha}\boldsymbol{\Phi}_{N,K}^T(\alpha). \quad \text{(A25)}$$

Derivative $\frac{d}{d\alpha}\boldsymbol{\Phi}_{N,K}(\alpha)$ of the matrix $\boldsymbol{\Phi}_{N,K}(\alpha)$ follows from Theorem 3 and the structure of $\boldsymbol{\Phi}_{N,K}(\alpha)$ (24). By (24), (38) and (39), bearing in mind that $0^0 = 1$, matrix $\boldsymbol{\Theta}_{N,K}(\alpha) = \frac{d}{d\alpha}\boldsymbol{\Phi}_{N,K}(\alpha)$ is given by (44). Thus, $\frac{d}{d\alpha}\boldsymbol{\Psi}(\alpha) = \boldsymbol{\Omega}_{N,K}(\alpha)$ (A25) is expressed by (43).

The regularization parameter $\lambda_{GCV}(\alpha)$, which is defined by the minimization task (28), satisfies the necessary condition (31) from Theorem 2; thus, to evaluate derivative $\frac{d\lambda_{GCV}(\alpha)}{d\alpha}$, Equation (31) is used. By (31), bearing in mind the notation (42), we have

$$\overline{\boldsymbol{G}}_N^T\boldsymbol{Y}(\alpha)^{-3}\overline{\boldsymbol{G}}_N tr\left[\boldsymbol{Y}(\alpha)^{-1}\right] = \overline{\boldsymbol{G}}_N^T\boldsymbol{Y}(\alpha)^{-2}\overline{\boldsymbol{G}}_N\ tr\left[\boldsymbol{Y}(\alpha)^{-2}\right].$$

Differentiating both sides of the above equation with respect to α, and bearing in mind the properties (A7) and (A8), we obtain

$$\overline{G}_N^T Y(\alpha)^{-3}\left[\tfrac{d}{d\alpha}Y(\alpha)^3\right]Y(\alpha)^{-3}\overline{G}_N tr\left[Y(\alpha)^{-1}\right] + \overline{G}_N^T Y(\alpha)^{-3}\overline{G}_N tr\left[Y(\alpha)^{-1}\tfrac{d}{d\alpha}Y(\alpha)Y(\alpha)^{-1}\right] =$$
$$\overline{G}_N^T Y(\alpha)^{-2}\left[\tfrac{d}{d\alpha}Y(\alpha)^2\right]Y(\alpha)^{-2}\overline{G}_N\ tr\left[Y(\alpha)^{-2}\right] + \overline{G}_N^T Y(\alpha)^{-2}\overline{G}_N\ tr\left[Y(\alpha)^{-2}\tfrac{d}{d\alpha}Y(\alpha)^2 Y(\alpha)^{-2}\right], \tag{A26}$$

where derivative $\tfrac{d}{d\alpha}Y(\alpha)$ is given by (A23). By (42) and property (A6), we have

$$\frac{d}{d\alpha}Y(\alpha)^2 = \left[\frac{d}{d\alpha}Y(\alpha)\right]Y(\alpha) + Y(\alpha)\frac{d}{d\alpha}Y(\alpha), \tag{A27}$$

and

$$\frac{d}{d\alpha}Y(\alpha)^3 = \left[\frac{d}{d\alpha}Y(\alpha)\right]Y(\alpha)^2 + Y(\alpha)\frac{d}{d\alpha}Y(\alpha)^2,$$

whence

$$\frac{d}{d\alpha}Y(\alpha)^3 = \left[\frac{d}{d\alpha}Y(\alpha)\right]Y(\alpha)^2 + Y(\alpha)\left[\frac{d}{d\alpha}Y(\alpha)\right]Y(\alpha) + Y(\alpha)^2\frac{d}{d\alpha}Y(\alpha). \tag{A28}$$

By (A28), bearing in mind the that matrices $Y(\alpha)$ and $\tfrac{d}{d\alpha}Y(\alpha)$ are symmetric, we have

$$\overline{G}_N^T Y(\alpha)^{-3}\left[\frac{d}{d\alpha}Y(\alpha)^3\right]Y(\alpha)^{-3}\overline{G}_N = 2\overline{G}_N^T Y(\alpha)^{-3}\left[\frac{d}{d\alpha}Y(\alpha)\right]Y(\alpha)^{-1}\overline{G}_N + \overline{G}_N^T Y(\alpha)^{-2}\left[\frac{d}{d\alpha}Y(\alpha)\right]Y(\alpha)^{-2}\overline{G}_N. \tag{A29}$$

Similarly, by (A27), we obtain

$$\overline{G}_N^T Y(\alpha)^{-2}\left[\frac{d}{d\alpha}Y(\alpha)^2\right]Y(\alpha)^{-2}\overline{G}_N = 2\overline{G}_N^T Y(\alpha)^{-2}\left[\frac{d}{d\alpha}Y(\alpha)\right]Y(\alpha)^{-1}\overline{G}_N. \tag{A30}$$

Including (A29), (A30) and (A27), Equation (A26) is, after algebraic manipulations, rewritten as

$$2\overline{G}_N^T Y(\alpha)^{-3}\left[\tfrac{d}{d\alpha}Y(\alpha)\right]Y(\alpha)^{-1}\overline{G}_N tr\left[Y(\alpha)^{-1}\right] + \overline{G}_N^T Y(\alpha)^{-2}\left[\tfrac{d}{d\alpha}Y(\alpha)\right]Y(\alpha)^{-2}\overline{G}_N tr\left[Y(\alpha)^{-1}\right] +$$
$$\overline{G}_N^T Y(\alpha)^{-3}\overline{G}_N tr\left[Y(\alpha)^{-1}\tfrac{d}{d\alpha}Y(\alpha)Y(\alpha)^{-1}\right] = 2\overline{G}_N^T Y(\alpha)^{-2}\left[\tfrac{d}{d\alpha}Y(\alpha)\right]Y(\alpha)^{-1}\overline{G}_N\ tr\left[Y(\alpha)^{-2}\right] +$$
$$\overline{G}_N^T Y(\alpha)^{-2}\overline{G}_N\ tr\left[\left[Y(\alpha)^{-2}\left[\tfrac{d}{d\alpha}Y(\alpha)\right]Y(\alpha)^{-1} + Y(\alpha)^{-1}\left[\tfrac{d}{d\alpha}Y(\alpha)\right]Y(\alpha)^{-2}\right]\right].$$

By (A23), the last equation can be expressed as follows:

$$2\overline{G}_N^T Y(\alpha)^{-3}\left[\Omega_{N,K}(\alpha) + \tfrac{d\lambda_{GCV}(\alpha)}{d\alpha}\mathbb{I}_{N,N}\right]Y(\alpha)^{-1}\overline{G}_N tr\left[Y(\alpha)^{-1}\right] + \overline{G}_N^T Y(\alpha)^{-2}\left[\Omega_{N,K}(\alpha)+\right.$$
$$\tfrac{d\lambda_{GCV}(\alpha)}{d\alpha}\mathbb{I}_{N,N}\right]Y(\alpha)^{-2}\overline{G}_N tr\left[Y(\alpha)^{-1}\right] + \overline{G}_N^T Y(\alpha)^{-3}\overline{G}_N tr\left[Y(\alpha)^{-1}\left[\Omega_{N,K}(\alpha) + \tfrac{d\lambda_{GCV}(\alpha)}{d\alpha}\mathbb{I}_{N,N}\right]Y(\alpha)^{-1}\right] =$$
$$2\overline{G}_N^T Y(\alpha)^{-2}\left[\Omega_{N,K}(\alpha) + \tfrac{d\lambda_{GCV}(\alpha)}{d\alpha}\mathbb{I}_{N,N}\right]Y(\alpha)^{-1}\overline{G}_N\ tr\left[Y(\alpha)^{-2}\right] + \overline{G}_N^T Y(\alpha)^{-2}\overline{G}_N\ tr\left[\left[Y(\alpha)^{-2}[\Omega_{N,K}(\alpha)+\right.\right.$$
$$\tfrac{d\lambda_{GCV}(\alpha)}{d\alpha}\mathbb{I}_{N,N}\right]Y(\alpha)^{-1} + Y(\alpha)^{-1}\left[\Omega_{N,K}(\alpha) + \tfrac{d\lambda_{GCV}(\alpha)}{d\alpha}\mathbb{I}_{N,N}\right]Y(\alpha)^{-2}\right]\right],$$

which, after painstaking algebraic manipulations, yields Equation (41), which completes the proof. $\square$

Appendix A.7. Proof of Proposition 1

By (6), for any time-scale factor α and any vector of model parameters g_K, we have

$$\|H_K(\tau,\alpha)\|_2^2 = \int_0^\infty \left[\sum_{k=0}^{K-1} g_k h_k(\tau,\alpha)\right]^2 d\tau, \tag{A31}$$

which is rewritten as

$$\|H_K(\tau,\alpha)\|_2^2 = \sum_{k=0}^{K-1}\sum_{j=0}^{K-1} g_k g_j \gamma_{kj}(\alpha). \tag{A32}$$

where the functions

$$\gamma_{kj}(\alpha) = \int_0^\infty h_k(\tau,\alpha)h_j(\tau,\alpha)d\tau, \tag{A33}$$

for $k, j = 0, 1, \ldots K - 1$.

By virtue of (A33), we have $\gamma_{kj}(\alpha) = \gamma_{jk}(\alpha)$ for any $k, j = 0, 1, \ldots K - 1$. The proof is based on the following integral formula [64] (Equation (3.351.3)):

$$\int_0^\infty \tau^n e^{-2\alpha\tau}d\tau = \frac{n!}{(2\alpha)^{n+1}}. \tag{A34}$$

For $k = j = 0$, definition (A33) and (4) immediately imply

$$\gamma_{00}(\alpha) = \int_0^\infty e^{-2\alpha\tau}d\tau = \frac{1}{2\alpha}. \tag{A35}$$

For $k = 0$ and $j = 1, \ldots K - 1$, by virtue of (A33), (3), (4) and (A34), we have

$$\gamma_{0j}(\alpha) = \left(\frac{e\alpha}{j}\right)^j \int_0^\infty \tau^j e^{-2\alpha\tau}d\tau = \left(\frac{e\alpha}{j}\right)^j \frac{j!}{(2\alpha)^{j+1}} = \frac{1}{2\alpha}\left(\frac{e}{2}\right)^j \frac{j!}{j^j}. \tag{A36}$$

For $k, j = 1, \ldots K - 1$, through (A33), (3) and (A34), we obtain

$$\gamma_{kj}(\alpha) = e^{k+j}\left(\frac{\alpha}{k}\right)^k\left(\frac{\alpha}{j}\right)^j \int_0^\infty \tau^{k+j} e^{-2\alpha\tau}d\tau = e^{k+j}\left(\frac{\alpha}{k}\right)^k\left(\frac{\alpha}{j}\right)^j \frac{(k+j)!}{(2\alpha)^{k+j+1}}.$$

The last equation is expressed as

$$\gamma_{kj}(\alpha) = \frac{1}{2\alpha}\left(\frac{e}{2}\right)^{k+j}\frac{(k+j)!}{k^k\, j^j}. \tag{A37}$$

Bearing in mind (A33), Equation (A32) is rewritten as the quadratic form

$$\|H_K(\tau,\alpha)\|_2^2 = g_K^T \Gamma(\alpha)g_K, \tag{A38}$$

where symmetric $K \times K$ matrix $\Gamma(\alpha) = \left[\gamma_{kj}(\alpha)\right]$, composed by the functions $\gamma_{kj}(\alpha)$ defined by (A33), through (A35)–(A37), is as follows

$$\Gamma(\alpha) = \frac{1}{2\alpha}\begin{bmatrix} 1 & \frac{e}{2} & \cdots & \left(\frac{e}{2}\right)^j\frac{j!}{j^j} & \cdots & \left(\frac{e}{2}\right)^{K-1}\frac{(K-1)!}{(K-1)^{K-1}} \\ \frac{e}{2} & 2\left(\frac{e}{2}\right)^2 & \cdots & \left(\frac{e}{2}\right)^{1+j}\frac{(1+j)!}{j^j} & \cdots & \left(\frac{e}{2}\right)^K\frac{(K)!}{(K-1)^{K-1}} \\ \vdots & \vdots & \ddots & \vdots & \ddots & \vdots \\ \left(\frac{e}{2}\right)^k\frac{k!}{k^k} & \left(\frac{e}{2}\right)^{k+1}\frac{(k+1)!}{k^k} & \cdots & \left(\frac{e}{2}\right)^{k+j}\frac{(k+j)!}{k^k\, j^j} & \cdots & \left(\frac{e}{2}\right)^{k+K-1}\frac{(k+K-1)!}{k^k\,(K-1)^{K-1}} \\ \vdots & \vdots & \vdots & \vdots & \ddots & \vdots \\ \left(\frac{e}{2}\right)^{K-1}\frac{(K-1)!}{(K-1)^{K-1}} & \left(\frac{e}{2}\right)^{K-1+j}\frac{(K)!}{(K-1)^{K-1}} & \cdots & \left(\frac{e}{2}\right)^{K-1+j}\frac{(K-1+j)!}{(K-1)^{K-1}\, j^j} & \cdots & \left(\frac{e}{2}\right)^{2K-2}\frac{(2K-2)!}{(K-1)^{2(K-1)}} \end{bmatrix}. \tag{A39}$$

Since matrix $\Gamma(\alpha)$ can be expressed as $\Gamma(\alpha) = \frac{1}{2\alpha}\Gamma_1$, where Γ_1 is given by (53), Equation (52) is derived.

According to (A38) and (A31), the quadratic form $g_K^T \Gamma(\alpha)g_K$ is expressed as

$$g_K^T \Gamma(\alpha)g_K = \int_0^\infty \left[\sum_{k=0}^{K-1} g_k h_k(\tau,\alpha)\right]^2 d\tau.$$

Thus, $g_K^T \Gamma(\alpha)g_K \geq 0$ for an arbitrary vector g_K, and $g_K^T \Gamma(\alpha)g_K = 0$, if and only if $\sum_{k=0}^{K-1} g_k h_k(\tau,\alpha) = 0$ for almost all $\tau > 0$. Since the basis functions $h_k(\tau,\alpha)$ are independent, the last equality holds, if and only if $g_k = 0$ for all $k = 0, 1, \ldots, K - 1$, i.e., only if the vector

$g_K = 0$, which yields the positive definiteness of $\Gamma(\alpha)$. Whence, $\Gamma_1 = 2\alpha\Gamma(\alpha)$. is positive definite, too; this finishes the proof. □

Appendix A.8. Proof of Proposition 4

Since, by (6) and (58), bearing in mind (A32), (A33), (A38) and $\Gamma(\alpha) = \frac{1}{2\alpha}\Gamma_1$, we have

$$\|H_K(\tau,\alpha) - \widetilde{H}_K(\tau,\alpha)\|_2^2 = \frac{1}{2\alpha}\left[\overline{g}_K^\lambda(\alpha) - g_K^\lambda(\alpha)\right]^T \Gamma_1 \left[\overline{g}_K^\lambda(\alpha) - g_K^\lambda(\alpha)\right],$$

in view of the right inequality in (54), the following estimation holds:

$$\|H_K(\tau,\alpha) - \widetilde{H}_K(\tau,\alpha)\|_2^2 \le \frac{1}{2\alpha}\sigma_1(\Gamma_1)\|\overline{g}_K^\lambda(\alpha) - g_K^\lambda(\alpha)\|_2^2. \tag{A40}$$

By virtue of (27) and (59), and bearing in mind (48), we obtain

$$\overline{g}_K^\lambda(\alpha) - g_K^\lambda(\alpha) = V(\alpha)\Lambda_\lambda(\alpha)\, U(\alpha)^T z_N,$$

where the orthogonal matrices $V(\alpha)$ and $U(\alpha)$ are defined by SVD (47), and the diagonal matrix $\Lambda_\lambda(\alpha)$ is given by (49). Thus,

$$\|\overline{g}_K^\lambda(\alpha) - g_K^\lambda(\alpha)\|_2^2 = tr\left[U(\alpha)\Lambda_\lambda^T(\alpha)\Lambda_\lambda(\alpha)\, U(\alpha)^T z_N z_N^T\right],$$

which, using the Schwarz inequality for matrices [57] and the orthogonality of $U(\alpha)$, is estimated as follows:

$$\|\overline{g}_K^\lambda(\alpha) - g_K^\lambda(\alpha)\|_2^2 \le tr\left[\Lambda_\lambda^T(\alpha)\Lambda_\lambda(\alpha)\Lambda_\lambda^T(\alpha)\Lambda_\lambda(\alpha)\right]^{\frac{1}{2}} tr\left[z_N z_N^T z_N z_N^T\right]^{\frac{1}{2}}. \tag{A41}$$

Since

$$tr\left[z_N z_N^T z_N z_N^T\right] = z_N^T z_N tr\left[z_N z_N^T\right] = \left[z_N^T z_N\right]^2 = \|z_N\|_2^4,$$

and by the diagonal structure of $\Lambda_\lambda(\alpha)$ (49), we have

$$tr\left[\Lambda_\lambda^T(\alpha)\Lambda_\lambda(\alpha)\Lambda_\lambda^T(\alpha)\Lambda_\lambda(\alpha)\right] = \sum_{i=1}^{r(\alpha)} \frac{[\sigma_i(\alpha)]^4}{\left\{[\sigma_i(\alpha)]^2 + \lambda\right\}^4},$$

inequality (A41) is equivalently rewritten as

$$\|\overline{g}_K^\lambda(\alpha) - g_K^\lambda(\alpha)\|_2^2 \le \left[\sum_{i=1}^{r(\alpha)} \frac{[\sigma_i(\alpha)]^4}{\left\{[\sigma_i(\alpha)]^2 + \lambda\right\}^4}\right]^{\frac{1}{2}} \|z_N\|_2^2. \tag{A42}$$

Now, estimations (A40) and (A42) yield inequality (60), and the proposition is proved. □

Appendix A.9. Derivation of the Formula (62)

To derive Equation (62), note that due to (2), the relaxation modulus induced by the spectrum

$$H(\tau) = \left[e^{-\left(\frac{1}{\tau}-m\right)^2/q}\right]/\tau \tag{A43}$$

is described by

$$G(t) = \int_0^\infty e^{-\left(\frac{1}{\tau}-m\right)^2/q} \frac{1}{\tau^2} e^{-t/\tau} d\tau.$$

Through (2) and (A43), applying the substitution $\frac{1}{\tau} = x$, the above integral is rewritten as

$$G(t) = -\int_{\infty}^{0} e^{-(x-m)^2/q} e^{-tx} dx = \int_{0}^{\infty} e^{-(x-m)^2/q} e^{-tx} dx$$

and next, as follows:

$$G(t) = e^{\frac{1}{4}t^2 q - mt} \int_{0}^{\infty} e^{\frac{-(x+\frac{1}{2}tq-m)^2}{q}} dx. \tag{A44}$$

Applying in the integral (A44) the substitution $\frac{\left(x+\frac{1}{2}tq-m\right)}{\sqrt{q}} = z$, the formula describing relaxation modulus is expressed as

$$G(t) = \sqrt{q}\, e^{\frac{1}{4}t^2 q - mt} \int_{(\frac{1}{2}tq-m)/\sqrt{q}}^{\infty} e^{-z^2} dz = \frac{\sqrt{q\pi}}{2}\, e^{\frac{1}{4}t^2 q - mt} erfc\left(\frac{\frac{1}{2}tq - m}{\sqrt{q}}\right), \tag{A45}$$

where the function $erfc(x)$ is defined by (63). Formula (62) follows directly from the last expression of the right-hand side of (A45), by comparison of (A43) and (61).

Appendix B

Table A1. Ranges of the applicability of the model for various time-scale parameters for $K = 6, \ldots 11$.

Time-Scale Factor α [s]	Range [1] of Relaxation Times $\tau_{app}(\alpha)$ [s]	Range [1] of Times $t_{app}(\alpha)$ [s]	Range [1] of Relaxation Times $\tau_{app}(\alpha)$ [s]	Range [1] of Times $t_{app}(\alpha)$ [s]
	$K = 6$	$K = 6$	$K = 7$	$K = 7$
0.0001	161,655.35	337,647.7	178,346.044	392,372
0.001	16,165.53	33,764.77	17,834.606	39,238.5
0.01	1616.55	3376.48	1783.462	3923.85
0.1	161.65	337.65	178.348	392.5
1	16.165	33.765	17.836	39.26
10	1.616	3.376	1.784	3.95
100	0.1616	0.337	0.1783	0.392
α [s]	$K = 8$	$K = 8$	$K = 9$	$K = 9$
0.0001	194,529.30	446,719.4	210,306.21	500,804.6
0.001	19,452.93	44,671.94	21,030.621	50,080.46
0.01	1945.29	4467.19	2103.062	5008.046
0.1	194.53	446.72	210.306	500.81
1	19.453	44.672	21.0306	50.08
10	1.945	4.467	2.1031	5.008
100	0.1945	0.447	0.2103	0.501
α [s]	$K = 10$	$K = 10$	$K = 11$	$K = 11$
0.0001	225,748.07	554,697	240,907.43	608,443.6
0.001	22,574.807	55,469.7	24,090.743	60,844.38
0.01	2257.481	5546.97	2409.074	6084.45
0.1	225.748	554.69	240.907	608.46
1	22.5748	55.469	24.0907	60.86
10	2.2575	5.547	2.4091	6.1
100	0.2257	0.555	0.2409	0.608

[1] The upper bounds $t_{app}(\alpha)$ (20) and $\tau_{app}(\alpha)$ (21) of the applicability intervals $[0, t_{app}(\alpha)]$ and $[0, \tau_{app}(\alpha)]$ are given.

Table A2. The square roots of the upper bounds $\psi_n(\Gamma_1)$ (56) and the square roots of the largest $\sigma_1(\Gamma_1)$ and minimal $\sigma_{min}(\Gamma_1)$ singular values of the matrix Γ_1 (53) for $K = 4, 5, \ldots 12$ model summands.

K	4	5	6	7	8	9	10	11	12
n					$\sqrt{\psi_n(\Gamma_1)}$				
1	3.680199	4.353830	4.974923	5.556319	6.106199	6.630173	7.132299	7.615642	8.082583
2	3.671711	4.330308	4.931579	5.489724	6.013827	6.510086	6.982965	7.435811	7.871217
3	3.667383	4.325690	4.927344	5.485966	6.010365	6.506625	6.979169	7.431342	7.865747
4	3.666555	4.325232	4.927078	5.485799	6.010251	6.506536	6.979081	7.431227	7.865575
5	3.666421	4.325190	4.927062	5.485792	6.010247	6.506534	6.979079	7.431224	7.865568
6	3.666400	4.325186	4.927061	5.485792	6.010247	6.506534	6.979079	7.431223	7.865567
7	3.666397	4.325186	4.927061	5.485792	6.010247	6.506534	6.979079	7.431223	7.865567
8	3.666396	4.325186	4.927061	5.485792	6.010247	6.506534	6.979079	7.431223	7.865567
9	3.666396	4.325186	4.927061	5.485792	6.010247	6.506534	6.979079	7.431223	7.865567
10	3.666396	4.325186	4.927061	5.485792	6.010247	6.506534	6.979079	7.431223	7.865567
$\sqrt{\sigma_1(\Gamma_1)}$	3.666396	4.325185	4.927061	5.485792	6.010247	6.506534	6.979079	7.431223	7.865567

Table A3. Optimal parameters $\widehat{g}_K = \widehat{g}_K^{\lambda_{GCV}}(\hat{\alpha})$ (45) of the relaxation spectrum models from Example 1; the elements of the vectors $\widehat{g}_K$ are expressed in [Pa].

					$\widehat{g}_K$				
$K = 3$	$K = 4$	$K = 5$	$K = 6$	$K = 7$	$K = 8$	$K = 9$	$K = 10$	$K = 11$	$K = 12$
0.50819	−0.32787	−0.31484	−0.28238	−0.26421	−0.26104	−0.26344	−0.25958	−0.25722	−0.24893
0.48916	3.25406	2.44705	1.86341	1.57535	1.44701	1.38846	1.26949	1.18396	1.042852
−0.01177	−5.37283	−1.74791	−0.21483	0.21735	0.33298	0.38151	0.53024	0.62627	0.77959
	3.79955	−1.90531	−1.81625	−1.22415	−0.94252	−0.85336	−0.72742	−0.63340	−0.49513
		3.00974	−0.42686	−0.80274	−0.65867	−0.57137	−0.58067	−0.58715	−0.58709
			2.49772	0.35990	0.00038	0.018141	−0.06714	−0.11806	−0.19722
				1.85736	0.61157	0.39343	0.27044	0.22125	0.12879
					1.25669	0.58048	0.40004	0.34372	0.26284
						0.75498	0.46114	0.34725	0.26954
							0.59592	0.35547	0.25643
								0.46177	0.31367
									0.49927

References

1. Malkin, A.I.A.; Malkin, A.Y.; Isayev, A.I. *Rheology: Concepts, Methods and Applications*; ChemTec: Deerfield Beach, FL, USA, 2006. Available online: https://books.google.pl/books?id=8rGafjhgz-UC (accessed on 3 December 2022).
2. Ferry, J.D. *Viscoelastic Properties of Polymers*, 3rd ed.; John Wiley & Sons: New York, NY, USA, 1980.
3. Rao, M.A. *Rheology of Fluid, Semisolid, and Solid Foods: Principles and Applications*; Springer: New York, NY, USA, 2013. Available online: https://books.google.pl/books?id=9-23BAAAQBAJ (accessed on 9 March 2023).
4. Ankiewicz, S.; Orbey, N.; Watanabe, H.; Lentzakis, H.; Dealy, J. On the use of continuous relaxation spectra to characterize model polymers. *J. Rheol.* **2016**, *60*, 1115–1120. [CrossRef]
5. Anderssen, R.S.; Davies, A.R.; de Hoog, F.R.; Loy, R.J. Derivative based algorithms for continuous relaxation spectrum recovery. *J. Non-Newton. Fluid Mech.* **2015**, *222*, 132–140. [CrossRef]
6. Mead, D.W. Numerical interconversion of linear viscoelastic material functions. *J. Rheol.* **1994**, *38*, 1769–1795. [CrossRef]
7. Hajikarimi, P.; Moghadas Nejad, F. Chapter 6-Interconversion of constitutive viscoelastic functions. In *Applications of Viscoelasticity*; Hajikarimi, P., Nejad, F.M., Eds.; Elsevier: Amsterdam, The Netherlands, 2021; pp. 107–139. [CrossRef]
8. Chang, F.-L.; Bin, H.; Huang, W.-T.; Chen, L.; Yin, X.-C.; Cao, X.-W.; He, G.-J. Improvement of rheology and mechanical properties of PLA/PBS blends by in-situ UV-induced reactive extrusion. *Polymer* **2022**, *259*, 125336. [CrossRef]
9. Pogreb, R.; Loew, R.; Bormashenko, E.; Whyman, G.; Multanen, V.; Shulzinger, E.; Abramovich, A.; Rozban, A.; Shulzinger, A.; Zussman, E.; et al. Relaxation spectra of polymers and phenomena of electrical and hydrophobic recovery: Interplay between bulk and surface properties of polymers. *J. Polym. Sci. Part B Polym. Phys.* **2017**, *55*, 198–205. [CrossRef]
10. Sun, Y.; Huang, B.; Chen, J. A unified procedure for rapidly determining asphalt concrete discrete relaxation and retardation spectra. *Constr. Build. Mater.* **2015**, *93*, 35–48. [CrossRef]
11. Luo, L.; Xi, R.; Ma, Q.; Tu, C.; Ibrahim Shah, Y. An improved method to establish continuous relaxation spectrum of asphalt materials. *Constr. Build. Mater.* **2022**, *354*, 129182. [CrossRef]

12. Martinez, J.M.; Boukamel, A.; Méo, S.; Lejeunes, S. Statistical approach for a hyper-visco-plastic model for filled rubber: Experimental characterization and numerical modeling. *Eur. J. Mech.-A/Solids* **2011**, *30*, 1028–1039. [CrossRef]
13. Bardet, S.; Gril, J. Modelling the transverse viscoelasticity of green wood using a combination of two parabolic elements. *C. R. Mécanique* **2002**, *330*, 549–556. [CrossRef]
14. Zhang, N.Z.; Bian, X.L.; Ren, C.; Geng, C.; Mu, Y.K.; Ma, X.D.; Jia, Y.D.; Wang, Q.; Wang, G. Manipulation of relaxation processes in a metallic glass through cryogenic treatment. *J. Alloys Compd.* **2022**, *894*, 162407. [CrossRef]
15. Yazar, G.; Demirkesen, I. Linear and Non-Linear Rheological Properties of Gluten-Free Dough Systems Probed by Fundamental Methods. *Food Eng. Rev.* **2023**, *15*, 56–85. [CrossRef]
16. Demidov, A.V.; Makarov, A.G. Spectral Simulation of Performance Processes of Polymeric Textile Materals. *Fibre Chem.* **2023**, *54*, 222–226. [CrossRef]
17. Lau, H.C.P.; Holtzman, B.K.; Havlin, C. Toward a Self-Consistent Characterization of Lithospheric Plates Using Full-Spectrum Viscoelasticity. *AGU Adv.* **2020**, *1*, e2020AV000205. [CrossRef]
18. Stankiewicz, A.; Golacki, K. Approximation of the continuous relaxation spectrum of plant viscoelastic materials using Laguerre functions. *Electron. J. Pol. Agric. Univ. Ser. Agric. Eng.* **2008**, *11*. Available online: http://www.ejpau.media.pl/articles/volume11/issue1/art-20.pdf (accessed on 7 March 2023).
19. Baumgaertel, M.; Winter, H.H. Determination of discrete relaxation and retardation time spectra from dynamic mechanical data. *Rheol. Acta* **1989**, *28*, 511–519. [CrossRef]
20. Honerkamp, J.; Weese, J. Determination of the relaxation spectrum by a regularization method. *Macromolecules* **1989**, *22*, 4372–4377. [CrossRef]
21. Malkin, A.Y. The use of a continuous relaxation spectrum for describing the viscoelastic properties of polymers. *Polym. Sci. Ser. A* **2006**, *48*, 39–45. [CrossRef]
22. Malkin, A.Y.; Vasilyev, G.B.; Andrianov, A.V. On continuous relaxation spectrum. Method of calculation. *Polym. Sci. Ser. A* **2010**, *52*, 1137–1141. [CrossRef]
23. Stadler, F.J.; Bailly, C. A new method for the calculation of continuous relaxation spectra from dynamic-mechanical data. *Rheol. Acta* **2009**, *48*, 33–49. [CrossRef]
24. Davies, A.R.; Goulding, N.J. Wavelet regularization and the continuous relaxation spectrum. *J. Non-Newton. Fluid Mech.* **2012**, *189–190*, 19–30. [CrossRef]
25. Davies, A.R.; Anderssen, R.S.; de Hoog, F.R.; Goulding, N.J. Derivative spectroscopy and the continuous relaxation spectrum. *J. Non-Newton. Fluid Mech.* **2016**, *233*, 107–118. [CrossRef]
26. Cho, K.S. Power series approximations of dynamic moduli and relaxation spectrum. *J. Rheol.* **2013**, *57*, 679–697. [CrossRef]
27. Takeh, A.; Shanbhag, S. A Computer Program to Extract the Continuous and Discrete Relaxation Spectra from Dynamic Viscoelastic Measurements. *Appl. Rheol.* **2013**, *23*, 24628. [CrossRef]
28. Pérez-Calixto, D.; Amat-Shapiro, S.; Zamarrón-Hernández, D.; Vázquez-Victorio, G.; Puech, P.-H.; Hautefeuille, M. Determination by Relaxation Tests of the Mechanical Properties of Soft Polyacrylamide Gels Made for Mechanobiology Studies. *Polymers* **2021**, *13*, 629. [CrossRef] [PubMed]
29. Joyner, H.S. *Rheology of Semisolid Foods*; Springer: Berlin/Heidelberg, Germany, 2019. Available online: https://books.google.pl/books?id=siy-DwAAQBAJ (accessed on 9 March 2023).
30. Alfrey, T.; Doty, P. The Methods of Specifying the Properties of Viscoelastic Materials. *J. Appl. Phys.* **1945**, *16*, 700–713. [CrossRef]
31. Widder, D.V. *An Introduction to Transformation Theory*; Academic Press: Cambridge, MA, USA, 1971. Available online: https://books.google.pl/books?id=wERoPQAACAAJ (accessed on 9 March 2023).
32. Alfrey, T. *Mechanical Behavior of High Polymers*; Interscience Publishers: Geneva, Switzerland, 1965.
33. ter Haar, D. An easy approximate method of determining the relaxation spectrum of a viscoelastic materials. *J. Polym. Sci.* **1951**, *6*, 247–250. [CrossRef]
34. Bažant, Z.P.; Yunping, X. Continuous Retardation Spectrum for Solidification Theory of Concrete Creep. *J. Eng. Mech.* **1995**, *121*, 281–288. [CrossRef]
35. Goangseup, Z.; Bažant, Z.P. Continuous Relaxation Spectrum for Concrete Creep and its Incorporation into Microplane Model M4. *J. Eng. Mech.* **2002**, *128*, 1331–1336. [CrossRef]
36. Macey, H.H. On the Application of Laplace Pairs to the Analysis of Relaxation Curves. *J. Sci. Instrum.* **1948**, *25*, 251. [CrossRef]
37. Sips, R. Mechanical behavior of viscoelastic substances. *J. Polym. Sci.* **1950**, *5*, 69–89. [CrossRef]
38. Yamamoto, R. Stress relaxation property of the cell wall and auxin-induced cell elongation. *J. Plant Res.* **1996**, *109*, 75–84. [CrossRef]
39. Stankiewicz, A. A scheme for identification of continuous relaxation time spectrum of biological viscoelastic materials. *Acta Sci. Pol. Ser. Tech. Agrar.* **2003**, *2*, 77–91.
40. Stankiewicz, A. A Class of Algorithms for Recovery of Continuous Relaxation Spectrum from Stress Relaxation Test Data Using Orthonormal Functions. *Polymers* **2023**, *15*, 958. [CrossRef]
41. Zhang, B.; Makram-Ebeid, S.; Prevost, R.; Pizaine, G. Fast solver for some computational imaging problems: A regularized weighted least-squares approach. *Digit. Signal Process.* **2014**, *27*, 107–118. [CrossRef]
42. Hansen, P.C. *Rank-Deficient and Discrete Ill-Posed Problems*; Society for Industrial and Applied Mathematics: Philadelphia, PA, USA, 1998. [CrossRef]

43. Szabatin, J. *Podstawy Teorii Sygnałów*; Wydawnictwa Komunikacji i Łączności: Warszawa, Poland, 1982. (In Polish)

44. Knuth, D.E. Two Notes on Notation. *Am. Math. Mon.* **1992**, *99*, 403–422. [CrossRef]

45. Yang, Z.-H.; Chu, Y.-M. On approximating the modified Bessel function of the second kind. *J. Inequal. Appl.* **2017**, *2017*, 41. [CrossRef]

46. Watson, G.N. *A Treatise on the Theory of Bessel Functions*, 3rd ed.; Cambridge University Press: Cambridge, UK, 1995. Available online: https://books.google.nl/books?id=Mlk3FrNoEVoC (accessed on 19 December 2022).

47. Martin, P.; Maass, F. Accurate analytic approximation to the Modified Bessel function of Second Kind K0(x). *Results Phys.* **2022**, *35*, 105283. [CrossRef]

48. Rehman, N.U.; Hussain, S.T.; Aziz, A. Pulsatile Darcy flow of water-based thermally radiative carbon nanotubes between two concentric cylinders. *Numer. Methods Partial. Differ. Equ.* **2023**, *39*, 213–230. [CrossRef]

49. Okita, T.; Harada, H. 3-D Analytical Model of Axial-Flux Permanent Magnet Machine with Segmented Multipole-Halbach Array. *IEEE Access* **2023**, *11*, 2078–2091. [CrossRef]

50. Lovrić, D.; Vujević, S. Accurate Computation of Internal Impedance of Two-Layer Cylindrical Conductors for Arguments of Arbitrary Magnitude. *IEEE Trans. Electromagn. Compat.* **2018**, *60*, 347–353. [CrossRef]

51. Das, R.; Manna, B.; Banerjee, A. Spectral element formulation for rock-socketed mono-pile under horizontal dynamic loads. *Soil Dyn. Earthq. Eng.* **2023**, *169*, 107863. [CrossRef]

52. Liu, M.; Huang, H. Poroelastic response of spherical indentation into a half space with an impermeable surface via step displacement. *J. Mech. Phys. Solids* **2021**, *155*, 104546. [CrossRef]

53. Yamamura, K. Dispersal distance of heterogeneous populations. *Popul. Ecol.* **2002**, *44*, 93–101. [CrossRef]

54. Gaunt, R.E. Inequalities for modified Bessel functions and their integrals. *J. Math. Anal. Appl.* **2014**, *420*, 373–386. [CrossRef]

55. Tikhonov, A.N.; Arsenin, V.Y. *Solutions of Ill-Posed Problems*; John Wiley & Sons: New York, NY, USA, 1977.

56. Wahba, G. Practical Approximate Solutions to Linear Operator Equations When the Data are Noisy. *SIAM J. Numer. Anal.* **1977**, *14*, 651–667. [CrossRef]

57. Golub, G.H.; Van Loan, C.F. *Matrix Computations*; Johns Hopkins University Press: Baltimore, MD, USA, 2013.

58. Lee, C.-H. Upper and lower matrix bounds of the solution for the discrete Lyapunov equation. *IEEE Trans. Autom. Control.* **1996**, *41*, 1338–1341. [CrossRef]

59. Rojo, O.; Soto, R.; Rojo, H. Bounds for the spectral radius and the largest singular value. *Comput. Math. Appl.* **1998**, *36*, 41–50. [CrossRef]

60. Kwakye-Nimo, S.; Inn, Y.; Yu, Y.; Wood-Adams, P.M. Linear viscoelastic behavior of bimodal polyethylene. *Rheol. Acta* **2022**, *61*, 373–386. [CrossRef]

61. Wang, J.; Wang, X.; Ruan, H. On the mechanical β relaxation in glass and its relation to the double-peak phenomenon in impulse excited vibration at high temperatures. *J. Non-Cryst. Solids* **2020**, *533*, 119939. [CrossRef]

62. Deaño, A.; Temme, N.M. Analytical and numerical aspects of a generalization of the complementary error function. *Appl. Math. Comput.* **2010**, *216*, 3680–3693. [CrossRef]

63. Takekawa, T. Fast parallel calculation of modified Bessel function of the second kind and its derivatives. *SoftwareX* **2022**, *17*, 100923. [CrossRef]

64. Gradshteyn, I.S.; Ryzhik, I.M. *Table of Integrals, Series, and Products*; Academic Press: Boston, MA, USA, 2007. [CrossRef]