INTERNATIONAL JOURNAL FOR NUMERICAL METHODS IN ENGINEERING
Int. J. Numer. Meth. Engng 2005; **63**:428–454
Published online 3 March 2005 in Wiley InterScience (www.interscience.wiley.com). DOI: 10.1002/nme.1161

Comparison of numerical methods for identification of viscoelastic line spectra from static test data

S. Gerlach and A. Matzenmiller*, †

Department of Mechanical Engineering, Institute of Mechanics, University of Kassel, Moenchebergstr. 7, 34109 Kassel, Germany

SUMMARY

Viscoelastic line spectra are identified from creep or relaxation data of static experiments with different numerical methods, which may or may not depend on additional informations, to be provided by the user, about the unknown parameters. If the least square method is applied, a non-linear optimization problem with non-negative constraints on the parameters has to be solved. Its solution can be achieved directly by using a gradient-based optimization algorithm like the projected Newton method of Bertsekas. However, appropriate starting values for the unknown parameters must be chosen. The problem can be alleviated by dividing the identification task into three successive steps, based on the Tschebyscheff approximation and the quadratic optimization method by Wolfe.

Alternatively, the identification task can be reduced to a quadratic optimization problem, if the user provides additional informations about the distribution of the respondance times of the spectra. The windowing-method of Emri and Tschoegl is based on this assumption. If the line spectrum is assumed to have equally distributed spectrum lines on the logarithmic axis, the identification problem can also be solved by standard regularization techniques, like the truncated singular value decomposition or the Tikhonov regularization.

The choice of qualified respondance times as additional information requires some experience with the identification task at hand. Its results may be improved after several reruns of the algorithms. Various applications of the methods to test and experimental data are given and a comparison of their performance is discussed. Copyright © 2005 John Wiley & Sons, Ltd.

KEY WORDS: linear viscoelasticity; line spectra; inverse problem; parameter identification; non-negative constraints; regularization technique; non-linear optimization

*Correspondence to: A. Matzenmiller, Department of Mechanical Engineering, Institute of Mechanics, University of Kassel, Moenchebergstr. 7, 34109 Kassel, Germany.
†E-mail: amat@ifm.maschinenbau.uni-kassel.de

Contract/grant sponsor: Publishing Arts Research Council; contract/grant number: 98 1846389
Contract/grant sponsor: DFG

Received 20 October 2003
Revised 12 March 2004
Accepted 7 May 2004

1. INTRODUCTION

The time-dependent mechanical behaviour of materials shows up in different phenomena. Stress relaxation under constant loading, decay of vibrations or strain and stress rate-dependence are some examples of viscoelastic material behaviour.

Thus, the constitutive equations, linking the stresses to the strains in the material, are functions of time. For infinitesimal strains the material behaviour of polymers may be described sufficiently accurate by the theory of the linear viscoelasticity in many cases. Based on the Boltzmann superposition principle, the constitutive equations for linear viscoelastic materials are functionals of the entire preceding stress or strain history

$$\varepsilon(t) = \int_{-\infty}^{t} J(t-\tau)\,\frac{\mathrm{d}\sigma(\tau)}{\mathrm{d}\tau}\,\mathrm{d}\tau, \quad \sigma(t) = \int_{-\infty}^{t} G(t-\tau)\,\frac{\mathrm{d}\varepsilon(\tau)}{\mathrm{d}\tau}\,\mathrm{d}\tau \tag{1}$$

where σ' and ε' are the time derivatives of the stress and strain history. $J(t)$ and $G(t)$ are the viscoelastic material functions, describing the behaviour of the material after the application of a constant stress or strain jump. $J(t) = \varepsilon(t)/\sigma_0$ is the creep function and $G(t) = \sigma(t)/\varepsilon_0$ the relaxation function. Both kernels, the retardation function J as well as the relaxation function G, must be determined from experimental data. Especially, the relaxation function is needed as the input data for the displacement formulation of the FE-method, if a boundary value problem with a viscoelastic material model must be analysed. Due to the fact that frequently only creep data are available from static experiments, an identification scheme is needed to determine the parameters for the retardation function from the given creep data. Afterwards the retardation function must be interconverted into the relaxation function.

The scope of this paper is the fitting of an appropriate retardation function to the discrete creep data. A finite Dirichlet–Prony series is chosen as the characteristic material function, representing well the linear viscoelastic behaviour in most cases. The retardation function

$$J_{\mathrm{N}}(t) = J_0 + \sum_{k=1}^{N} J_k(1 - \mathrm{e}^{-t/\tau_k}) \tag{2}$$

corresponds to the generalized rheological model of Kelvin. The unknown parameters J_k and τ_k represent the discrete viscoelastic spectrum and completely determine the time-dependent behaviour. The multipliers J_k in front of the exponentials define the spectrum strengths, whereas τ_k in the exponent of the kernel function describes the respondance time. For physically realistic materials these parameters must be positive. Therefore, the fitting algorithm has to take into account the constraints

$$\begin{aligned} &J_0 \geqslant 0 \\ &J_k \geqslant 0, \quad \tau_k \geqslant 0, \quad k = 1, 2, \ldots, N \end{aligned} \tag{3}$$

The emphasis of this paper is not the discussion on the advantages of various spectra functions, but on a comparison of different numerical methods for the identification of discrete creep spectra.

2. CLASSIFICATION OF IDENTIFICATION METHODS

In the following four different numerical methods are discussed, which generate the discrete retardation spectrum from the discrete response function of a static creep experiment. These methods can be divided into two different groups in a way whether the user has to make additional assumptions about the distribution of the respondance times or not.

2.1. Identification algorithms without additional informations

The identification of the retardation spectrum without any additional information leads to a non-linear optimization problem with non-negative constraints, which must be taken into account by the method applied. The discrete least square approach leads to a set of non-linear normal equations. The iterative solution of the resulting system of non-linear equations involves the difficulty that the convergence to the minimum error is not guaranteed for an arbitrary starting vector. In Section 3 the projected Newton method of Bertsekas is shortly outlined, which iteratively solves the non-linear optimization problem for a given set of appropriate starting values for the unknown parameters.

Alternatively, a new identification algorithm is presented in Section 4, where the identification task is solved in three steps [1]. At first a qualified tendency function is determined to approximate the test data sufficiently well. Such tendency functions are often already available for many materials. By applying the Tschebyscheff-approximation to the tendency function, the respondance times and the spectrum strengths are calculated. In the third step the respondance times from the Tschebyscheff approach are kept and the spectrums strengths are redetermined by approximating the discrete creep data with the Dirichlet–Prony series. With the given respondance times the determination of the new spectrum strengths leads to a quadratic optimization problem, which may be solved with Wolfe's version of the simplex method in view of the non-negative constraints on the strength parameters.

This class of identification algorithms has the advantage that no assumptions are required about the respondance times of the model function in advance.

2.2. Identification algorithms based on additional informations

The identification task can be reduced to a quadratic optimization problem, if the user provides additional informations about the distribution of the respondance times of the spectrum. The specification of a set of suitable respondance times is a meaningful strategy, as long as enough respondance times are used. In this case the experimental data can be approximated with sufficient accuracy, although the optimal respondance times for the least error are, in general, not among the chosen ones. Usually, more respondance times are used as necessary to approximate the experimental data. A large spectrum with many specified respondance times has the advantage of successfully describing the experimental data. However, with respect to the finite element analysis of rate-dependent boundary value problems, a small number of spectrum lines is desirable, since each spectrum line requires the numerical time integration of an ordinary differential equation. For instance not more than 10 spectrum lines can be defined in many commercial finite element programs.

Based on a given set of respondance times, different numerical methods have been developed to determine the corresponding spectrum lines: e.g. the collocation method of Schapery [2],

the smoothness constraint approach of Clauser and Knauss [3], the multidata method of Cost and Becker [4] or the enhanced multidata method of Bradshaw and Brinson [5].

In this paper, the windowing-method of Emri and Tschoegl [6–8], shortly recalled in Section 5, as well as the regularization techniques, based on the singular value decomposition (SVD)—see Section 6—are used to identify the viscoelastic line spectrum. For both methods the respondance times must be carefully chosen in advance.

3. PROJECTED NEWTON METHOD OF BERTSEKAS

The identification of the viscoelastic line spectrum from given experimental data requires the solution of an inverse problem, which has no unique solution in general, because it is ill-posed according to the definitions of Hadamard. Therefore, the objective of the parameter identification is not to determine the exact solution, but to find the optimal solution, which minimizes the least square functional f_{LS} of the model response $\mathbf{M}(\mathbf{x})$ and the experimental data $\mathbf{d}$.

$$f_{\mathsf{LS}}(\mathbf{x}) = \tfrac{1}{2}||\mathbf{M}(\mathbf{x}) - \mathbf{d}||_2^2 \to \text{Min.}, \quad \mathbf{M}(\mathbf{x}) \in \mathbb{R}^M, \quad \mathbf{x} \in \mathbb{R}^{(2N+1)}, \quad \mathbf{d} \in \mathbb{R}^M \tag{4}$$

The components of the vector $\mathbf{x}$ represent the unknown material parameters

$$\mathbf{x}^{\mathrm{T}} = [\, J_0 \;\; J_1 \;\; \ldots \;\; J_N \;\; \tau_1 \;\; \ldots \;\; \tau_N \,] \tag{5}$$

which are subjected to the constraints

$$x_i \geqslant 0, \quad i = 1, 2, \ldots, 2N+1 \tag{6}$$

The vector $\mathbf{d}$ contains the values of the experimental creep data at different points t_j in time.

$$\mathbf{d}^{\mathrm{T}} = [\, d(t_1) \;\; d(t_2) \;\; d(t_3) \;\; \cdots \;\; d(t_M) \,] \tag{7}$$

The model response $\mathbf{M}(\mathbf{x})$ can be given as a function of the unknown parameters $\mathbf{x}$ and the points t_j in time of the experiment as follows:

$$\mathbf{M}(\mathbf{x}) = \begin{bmatrix} J_0 + & J_1(1-\mathrm{e}^{-t_1/\tau_1}) + & \cdots + & J_N(1-\mathrm{e}^{-t_1/\tau_N}) \\ J_0 + & J_1(1-\mathrm{e}^{-t_2/\tau_1}) + & \cdots + & J_N(1-\mathrm{e}^{-t_2/\tau_N}) \\ \vdots & \vdots & \vdots & \vdots \\ J_0 + & J_1(1-\mathrm{e}^{-t_M/\tau_1}) + & \cdots + & J_N(1-\mathrm{e}^{-t_M/\tau_N}) \end{bmatrix} \tag{8}$$

The optimization problem (4) is solved by the gradient-based iteration method of Bertsekas [9].

$$\mathbf{x}^{(k+1)} = \mathscr{P}\{\mathbf{x}^{(k)} - \alpha^{(k)}\mathbf{s}^{(k)}\}, \quad k = 0, 1, 2, \ldots \tag{9}$$

$$\mathscr{P}\{\mathbf{x}\}_i := \max(0, x_i^{(k)}), \quad i = 0, 1, \ldots 2N+1 \tag{10}$$

The operator $\mathscr{P}$ projects those parameters, which violate constraint (6), onto the positive range of real numbers. The vector $\mathbf{x}^{(k)}$ represents the parameters of the preceding step or the starting

vector $\mathbf{x}^{(0)}$. The parameter $\alpha^{(k)}$ denotes a step-size, chosen by the Goldstein–Armijo descent test. The descent direction $\mathbf{s}^{(k)}$ must be determined in each step by an iteration matrix $\mathbf{B}$ and the gradient $\nabla f(\mathbf{x})$ of Equation (4).

$$\mathbf{s}^{(k)} = \mathbf{B}\nabla f(\mathbf{x}) \tag{11}$$

This matrix $\mathbf{B}$ should be positive definite, which can be achieved by the following choices:

$$\begin{aligned} \textit{Gradient method}&: \mathbf{B} = \mathbf{I} \\ \textit{Newton method}&: \mathbf{B} = \nabla^2[f(\mathbf{x}^{(k)})]^{-1} \\ \textit{Gauss–Newton method}&: \mathbf{B} = [\mathbf{J}^{\mathrm{T}}\mathbf{J}]^{-1} \end{aligned} \tag{12}$$

with the Jacobi-matrix

$$\mathbf{J} = \frac{\mathrm{d}(\mathbf{M}(\mathbf{x}))^{\mathrm{T}}}{\mathrm{d}\mathbf{x}} \tag{13}$$

In order to ensure a descent within each iteration step of the algorithm, a additional condition of the Bertsekas algorithm is, that the matrix has to be 'diagonalized'. For further details on this terminology see Reference [9].

The Newton method is used here with its inverse of the Hesse matrix determined analytically. Its numerical evaluation in each iteration step is indeed very expensive. If the Hesse matrix $\mathbf{B}$ is not positive definite, it is replaced by the identity matrix $\mathbf{I}$ in the iteration step at hand. However, this simplification slows down the convergence of the gradient method. For future work on this topic the author recommend to use the Gauss–Newton-matrix in those case where the Hesse matrix in not positive definite.

4. TSCHEBYSCHEFF-APPROXIMATION OF A TENDENCY FUNCTION AND SIMPLEX-METHOD BY WOLFE

The algorithm proposed next solves the identification task in three steps by approximating a qualified tendency function through the creep data by means of the Tschebyscheff method in order to find the relaxation times. Afterwards the spectrum strengths are redetermined by solving again the least square problem with the quadratic optimization method of Wolfe [1] but keeping the retardation times from the previous step of the Tschebyscheff approximation. The three steps are outlined in the following.

4.1. Approximation of the creep data by a qualified tendency function J_{T}

The creep data are approximated by a qualified tendency function J_{T}, which are already available for many materials. Here, a power law ($J_{\mathsf{T}} = \gamma t^{\alpha}$) or a hyperbolic tangents function ($J_{\mathsf{T}} = b\ \tanh(ct)^{d}$) are introduced among others as demonstration examples for the tendency function. The advantage of the appropriately chosen tendency functions consists in the fact, that a very few numbers of parameters in the tendency function already represent the test data sufficiently well. Hence, the determination of the free parameters of the tendency function is

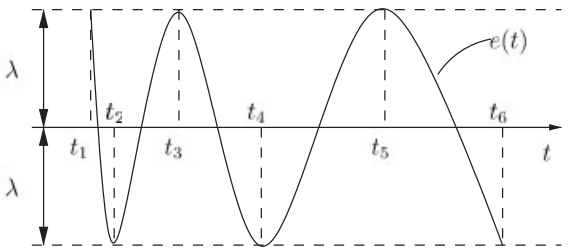

Figure 1. Schematic representation of the alternating error function $\varepsilon(t)$ with $N = 2$.

much easier than the solution of the original identification problem. In the case of a power law the linear regression method provides the only two parameters, namely the coefficient γ and the exponent α in double logarithmic scale of the strain versus time diagram.

4.2. Tschebyscheff-approximation of tendency function J_T by model function J_N

The use of the maximum norm $\| \ \|_\infty$ in the theory of function approximations leads to the class of Tschebyscheff-methods. The advantage of this class of schemes is that they converge to the minimal solution for the non-linear approximation with exponentials—see Reference [10]. By rearranging the objective function f_{TA} of the Tschebyscheff approach,

$$f_{\mathsf{TA}}(\mathbf{x}) = \| J_\mathsf{N}(\mathbf{x}) - J_\mathsf{T} \|_\infty \to \text{Min.}, \quad J_\mathsf{N} : \mathbb{R} \to \mathbb{R}, \quad \mathbf{x} \in \mathbb{R}^{(2N+1)}, J_\mathsf{T} : \mathbb{R} \to \mathbb{R} \tag{14}$$

a set of non-linear equations is obtained

$$J_\mathsf{N}(t_\mu) + (-1)^\mu \lambda = J_\mathsf{T}(t_\mu) \quad \text{for } \mu = 1, 2, \ldots, 2N + 2 \tag{15}$$

with J_N according to Equation (2) and unknowns J_0, J_i, τ_i, and λ for $i = 1, 2, \ldots, N$ and given N. The solution of Equation (15) leads to the error function $e = J_T - J_N$, alternating between two extrema with the absolute value λ at the points $t_1, t_2, t_3, \ldots, t_\mu, \ldots, t_{2N+2}$ in time as shown in Figure 1. The error function is determined by solving iteratively the system of $(2N + 2)$ non-linear equations, given by (15). Starting values for the material parameters $(J_0, J_1, J_2, \ldots, J_N; \tau_1, \tau_2, \ldots, \tau_N)$ and for a set of points $(t_1, t_2, \ldots, t_\mu, \ldots, t_{2N+2})$ in time must be chosen for the first alternating error function.

4.3. Redetermination of spectrum strengths by Wolfe's method

The approximation error form the tendency function obviously influences the determination of th retardation times. However, the spectrum strength are best fitted to the original data in the next step. By taking the calculated retardation times from step 2, the retardation strengths are computed from the creep data with the least square functional,

$$f_{\mathsf{LS}}(\mathbf{x}) = \tfrac{1}{2} \, ||\mathbf{H}\mathbf{x} - \mathbf{d}||_2^2 \to \text{Min.}, \quad \mathbf{H} \in \mathbb{R}^{M \times (N+1)}, \quad \mathbf{x} \in \mathbb{R}^{(N+1)}, \quad \mathbf{d} \in \mathbb{R}^M \tag{16}$$

under the condition that constraints (6) hold. This minimum problem is a special case of problem (4) since the respondance times are known. The coefficient matrix $\mathbf{H}$ is stated

as follows:

$$\mathbf{H} = \begin{bmatrix} 1 & (1-\mathrm{e}^{-t_1/\tau_1}) & (1-\mathrm{e}^{-t_1/\tau_2}) & \cdots & (1-\mathrm{e}^{-t_1/\tau_N}) \\ 1 & (1-\mathrm{e}^{-t_2/\tau_1}) & (1-\mathrm{e}^{-t_2/\tau_2}) & \cdots & (1-\mathrm{e}^{-t_2/\tau_N}) \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & (1-\mathrm{e}^{-t_M/\tau_1}) & (1-\mathrm{e}^{-t_M/\tau_2}) & \cdots & (1-\mathrm{e}^{-t_M/\tau_N}) \end{bmatrix} \tag{17}$$

The column vector $\mathbf{d}$ collects the experimental data as in (7), the reduced column vector $\mathbf{x}$ comprises the elastic compliance J_0 and the unknown spectrum strengths J_k

$$\mathbf{x}^{\mathrm{T}} = [\, J_0 \;\; J_1 \;\; \ldots \;\; J_N \,] \tag{18}$$

By introducing the slack variables q and r_k, the inequality constraints may be replaced by equations of the form

$$-J_0 + q^2 = 0, \quad -J_k + r_k^2 = 0, \;\; \text{for } k = 1, \ldots, N \tag{19}$$

By means of the Lagrange multiplier method the constraint equations can be built into the least square functional. The corresponding Lagrange function h is quadratic in the unknown strength parameters

$$h = \hat{h}(J_o, J_k, q, r_k, u_0, u_k) = f_{\mathsf{LS}}(J_0, J_k) + u_0(q^2 - J_0) + \sum_{k=1}^{N} u_k(r_k^2 - J_k) \tag{20}$$

with the Lagrange multipliers u_0 and u_k. The necessary conditions for the minimal solution

$$\partial h/\partial J_0 = 0 \quad \text{and} \quad \partial h/\partial J_k = 0 \quad \text{for } k = 1, 2, \ldots, N$$

lead to a set of $(N+1)$ linear equations:

$$\mathbf{H}^{\mathrm{T}}\mathbf{d} = \mathbf{H}^{\mathrm{T}}\mathbf{H}\,\mathbf{x} - \tfrac{1}{2}\mathbf{u} \tag{21}$$

The column vector $\mathbf{u}$ contains the unknown Lagrange multipliers

$$\mathbf{u}^{\mathrm{T}} = [u_0 \;\; u_1 \;\; u_2 \;\; \cdots \;\; u_N] \tag{22}$$

The partial derivatives $\partial h/\partial u_0 = 0$ and $\partial h/\partial u_k = 0$ provide the modified constraints (19). The derivatives $\partial h/\partial q = 0$ and $\partial h/\partial r_v = 0$ lead to simple non-linear algebraic equations.

$$u_0 q = 0 \tag{23}$$

$$u_k r_k = 0 \quad \text{for } k = 1, 2, \ldots, N \tag{24}$$

which are equivalent to

$$u_k x_k = 0 \quad \text{for } k = 0, 1, \ldots, N \tag{25}$$

This system of linear (21) and non-linear (25) equations is solved by the modified Simplex method of Wolfe [11], who introduced non-negative artifical variables, inserted into the column vector **v**

$$\mathbf{v}^{\mathrm{T}} = [v_0 \;\; v_1 \;\; v_2 \;\; \cdots \;\; v_N] \tag{26}$$

The modified Simplex method of Wolfe is a special solution technique for quadratic optimization problems with constraints, in contrast to the Bertsekas algorithm, which is a general solution method for all kinds of constrained non-linear optimization problems. Wolfe's method always converges to the solution point in a number of finite steps and the starting vector has not to be chosen by the user, as it is the case in Bertsekas algorithm. According to Wolfe's approach the system of linear equations is expanded by means of the artifical variables **v**

$$\mathbf{H}^{\mathrm{T}}\mathbf{d} = \mathbf{H}^{\mathrm{T}}\mathbf{H}\,\mathbf{x} - \tfrac{1}{2}\,\mathbf{u} + \mathbf{v} \tag{27}$$

The necessary objective function for the Simplex method is formulated as

$$z = \sum_{k=0}^{N} v_k \rightarrow \text{Min!} \tag{28}$$

By inserting (27) into (28), the following equation is obtained:

$$z = c - c_0 J_0 - \sum_{k=1}^{N} c_k J_k + \frac{1}{2}\,u_0 + \sum_{k=1}^{N} \frac{1}{2}\,u_k \tag{29}$$

with the variables

$$c = \sum_{i=1}^{M} \left\{ d(t_i) \left[1 + \sum_{k=1}^{N} (1 - \mathrm{e}^{-t_i/\tau_k}) \right] \right\} \tag{30}$$

$$c_0 = \sum_{i=1}^{M} \left[1 + \sum_{k=1}^{N} (1 - \mathrm{e}^{-t_i/\tau_k}) \right] \tag{31}$$

$$c_k = \sum_{i=1}^{M} \left\{ (1 - \mathrm{e}^{-t_i/\tau_k}) \left[1 + \sum_{k=1}^{N} (1 - \mathrm{e}^{-t_i/\tau_k}) \right] \right\} \tag{32}$$

Equation (28) represents a linear optimization problem, subjected to linear (27) and non-linear (25) constraints. The objective function (28) is minimized, if all artificial variables are equal to zero and, therefore, the linear constraints (27) are equivalent to (21).

A feasible starting vector **v**, which initiates the Simplex method for the minimization of Equation (28), is given by Equation (27) for $\mathbf{v} = \mathbf{H}^{\mathrm{T}}\mathbf{d}$, when J_k and u_k are assumed to be equal to zero. For a clear outline of the method as well as for the development of the numerical optimization scheme the necessary variables are inserted into the Simplex tableau. The first row defines the basis variables, set equal to zero ($= 0$), whereas the first column corresponds to the non-basis variables, which are non-zero ($\neq 0$). The last row contains the value z of the objective function according to Equation (29). The initial tableau is shown in Table I, where J_k and u_k are basis variables. The minimum of the objective function is computed with the standard Simplex algorithm for the case of linear equations by successively exchanging the

Table I. Initial tableau of the Simplex method.

Basis	Value	J_0	J_1	J_2	$\cdots$	J_N	u_0	u_1	u_2	$\cdots$	u_N	v_0	v_1	$\cdots$	v_N
v_0	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
v_1	$\cdots$	$\cdots$				$\cdots$	$\cdots$				$\cdots$	$\cdots$			
v_2	$\mathbf{H}^{\mathrm{T}}\mathbf{d}$	$\cdots$		$\mathbf{H}^{\mathrm{T}}\mathbf{H}$		$\cdots$	$\cdots$		$-\frac{1}{2}\mathbf{I}$		$\cdots$	$\cdots$		$\mathbf{I}$	$\cdots$
$\cdots$	$\cdots$	$\cdots$				$\cdots$	$\cdots$				$\cdots$	$\cdots$			$\cdots$
v_N	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$	$\cdots$
z	c	c_0	c_1	c_2	$\ldots$	c_N	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	$\frac{1}{2}$	0	0	0	0

basis and non-basis variables in the Simplex tableau, which needs to be constructed after each iteration step. For the linear constraint equations the exchange rules of the Simplex method control, which variable enters and which leaves the basis at each step. In order to satisfy also the non-linear equations (25), the usual rules of the Simplex method must be extended. Thus, at least one variable in all the pairs $(u_0, J_0), (u_1, J_1), \ldots, (u_N, J_N)$ must be non-basic, since at most only one component of each doublet may be nonzero. A detailed outline of Wolfe's method with examples is given in Reference [11].

5. WINDOWING-METHOD OF EMRI AND TSCHOEGL

Emri and Tschoegl proposed a recursive identification scheme by using the characteristic property of the exponentials in the kernel. The method is based on the fact that a single Dirichlet–Prony term e^{-t/τ_k} is significantly time-dependent only in the range of two time decades, called the modelling window. By means of the modelling window for each kernel function an exactly defined subset of experimental data can be chosen to be used for the determination of each spectrum line. A set of appropriate respondance times τ_k must be preselected to successfully apply this method to the identification of the discrete line spectrum. The respondance times are selected in the time interval, where the creep curve on logarithmic scale significantly increases. In each time decade within this interval about four respondance times are assumed with an equidistant spacing. For further details see References [6–8]. In the following the key equations of the scheme are summarized for completeness. At first the creep data are normalized

$$\hat{d}(t_j) = \frac{\max d(t_j) - d(t_j)}{\max d(t_j) - \min d(t_j)} \tag{33}$$

as well as the spectrum strengths

$$x_k = J_k \Big/ \sum_{i=1}^{N} J_k \tag{34}$$

The absolute error Δ_j between the data points and the retardation function is introduced and the least square norm $f^{\Delta}(x_k)$ is formed with M_{W} function values in the modelling

window.

$$f^{\Delta}(x_k) = \sum_{j=1}^{M_W} \Delta_j = \sum_{j=1}^{M_W} \left[\sum_{i=m}^{k-1} x_i \mathrm{e}^{-t_j/\tau_i} + x_k \mathrm{e}^{-t_j/\tau_k} + \sum_{i=k+1}^{N} x_i \mathrm{e}^{-t_j/\tau_i} - \hat{d}(t_j) \right]^2 \tag{35}$$

The parameter m depends on the number of respondance times preselected per decade. The minimization of the error function supplies the necessary conditions (36) for the determination of the normalized spectrum strengths x_k to the associated respondance time τ_k

$$\frac{\partial f^{\Delta}(x_k)}{\partial x_k} = 0 \rightarrow x_k = \sum_{j=1}^{M_W} \frac{\left[\hat{d}(t_j) - \sum_{i=m}^{k-1} x_i \mathrm{e}^{-t_j/\tau_i} - \sum_{i=k+1}^{N} x_i \mathrm{e}^{-t_j/\tau_i} \right]}{\mathrm{e}^{-2t_j/\tau_k}} \mathrm{e}^{-t_j/\tau_k} \tag{36}$$

For the application to non-smoothed test data the procedure is modified according to Reference [7] by introducing the relative error $\delta_j = \Delta_j / \hat{d}(t_j)$. The minimization of the relative error function leads to

$$\frac{\partial f^{\delta}(x_k)}{\partial x_k} = 0 \rightarrow \sum_{j=1}^{M_W} \frac{\left[\hat{d}(t_j) - \left(\sum_{i=m}^{k-1} x_i \mathrm{e}^{-t_j/\tau_i} + \sum_{i=k+1}^{N} x_i \mathrm{e}^{-t_j/\tau_i} + x_k \mathrm{e}^{-t_j/\tau_k} \right) \right]}{\left(\sum_{i=m}^{k-1} x_i \mathrm{e}^{-t_j/\tau_i} + \sum_{i=k+1}^{N} x_i \mathrm{e}^{-t_j/\tau_i} + x_k \mathrm{e}^{-t_j/\tau_k} \right)^3} \hat{d}(t_j) \mathrm{e}^{-t_j/\tau_k} = 0 \tag{37}$$

The set of non-linear equations for the spectrum strengths x_k is iteratively solved by the following recursion formula, which has some similarity to the Gauss–Seidel solution method:

$$x_k^{(n+1)} = \frac{A(x_k^{(n)})}{B(x_k^{(n)})} \tag{38}$$

$$A(x_k^{(n)}) = \sum_{j=1}^{M_W} \frac{\left[\hat{d}(t_j) - \left(\sum_{i=m}^{k-1} x_i^{(n)} \mathrm{e}^{-t_j/\tau_i} + \sum_{i=k+1}^{N} x_i^{(n)} \mathrm{e}^{-t_j/\tau_i} \right) \right]}{\left(\sum_{i=m}^{k-1} x_i^{(n)} \mathrm{e}^{-t_j/\tau_i} + \sum_{i=k+1}^{N} x_i^{(n)} \mathrm{e}^{-t_j/\tau_i} + x_k^{(n)} \mathrm{e}^{-t_j/\tau_k} \right)^3} \hat{d}(t_j) \mathrm{e}^{-t_j/\tau_k} \tag{39}$$

$$B(x_k^{(n)}) = \sum_{j=1}^{M_W} \frac{\left[\mathrm{e}^{-t_j/\tau_k} \right]^2}{\left(\sum_{i=m}^{k-1} x_i^{(n)} \mathrm{e}^{-t_j/\tau_i} + \sum_{i=k+1}^{N} x_i^{(n)} \mathrm{e}^{-t_j/\tau_i} + x_k^{(n)} \mathrm{e}^{-t_j/\tau_k} \right)^3} \hat{d}(t_j) \tag{40}$$

The initial values ($x_k^{(1)}$ for $k = 1, \ldots, N$) are found by recursively evaluating them from (36), starting with the relative strength for the largest respondance time τ_N. In the first sweep all spectrum strengths x_k are set to zero. In the following sweeps all positive strength values will be saved and negative ones are set to zero. Emri and Tschoegl suggest to repeat this step up to three times to speed up the iteration algorithm in (38) based on the minimization of the relative error. The application of (38) is terminated according to Emri and Tschoegl as soon as the norm of changes of the parameters $||\mathbf{x}^{(n+1)} - \mathbf{x}^{(n)}||$ is smaller than a given criterion. The convergence criterion of this iteration scheme works successfully. However, it is in contrast to

other methods, which use a small value of the residual norm to terminate the recursion. In the last step the unknown spectrum strengths J_k are computed from the normalized values x_k

$$J_k = x_k[\max d(t_j) - \min d(t_j)] \tag{41}$$

The instantaneous modulus J_0 can be determined as follows:

$$J_0 = \max d(t_j) - \sum_{i=1}^{N} J_k \tag{42}$$

The identified spectrum may be improved by the repeated application of the algorithm, after an enhanced set of respondance times is selected on the basis the preceding result. The number of respondance times is increased in the time range, where two or more closely spaced spectrum strengths are determined clearly greater than zero. It is also possible that closely spaced spectrum lines may be replaced by a few or even by a single spectrum line without a significant loss of accuracy for the resulting exponential series approximation. In comparison to the initially identified line spectrum this strategy usually reduces the discrete spectrum. Emri and Tschoegl explain in detail—see Reference [6]—how the selection of the respondance times can be improved in order to reduce the number of spectrum lines.

6. REGULARIZATION OF SINGULAR VALUES

Honerkamp [12] and Honerkamp and Weese [13] used the regularization techniques to identify the discrete viscoelastic spectrum from experimental data of dynamical tests. Similar to their approach the respondance times τ_k are chosen here as equidistantly distributed on the logarithmic time axis. On the assumption that no large differences are between the strength values of neighbouring spectrum lines, the regularization techniques, based on the SVD, can be used to determine the strength values J_k. The elastic creep compliance J_0 can be computed from the experimental data at the point $t = t_{\min}$ in time. With given values for the respondance times τ_k the least square problem

$$f_{\mathrm{LS}}(\mathbf{x}) = \tfrac{1}{2}\,||\mathbf{A}\mathbf{x} - \mathbf{d}||_2^2 \to \text{Min.}, \quad \mathbf{A} \in \mathbb{R}^{M\times N}, \ \ \mathbf{x} \in \mathbb{R}^N, \ \ \mathbf{d} \in \mathbb{R}^M \tag{43}$$

is quadratic in the vector $\mathbf{x}$ of the unknown spectrum strengths.

$$\mathbf{x}^{\mathrm{T}} = [\, J_1 \ \ J_2 \ \ \ldots \ \ J_N \,] \tag{44}$$

The vector $\mathbf{d}$ contains the values of the experimental data as in (7). The components of the matrix $\mathbf{A}$ are exponential functions of the chosen respondance times τ_k and the fixed points t_j in time of the experimental data

$$\mathbf{A} = \begin{bmatrix} (1-\mathrm{e}^{-t_1/\tau_1}) & (1-\mathrm{e}^{-t_1/\tau_2}) & \cdots & (1-\mathrm{e}^{-t_1/\tau_N}) \\ (1-\mathrm{e}^{-t_2/\tau_1}) & (1-\mathrm{e}^{-t_2/\tau_2}) & \cdots & (1-\mathrm{e}^{-t_2/\tau_N}) \\ \vdots & \vdots & \ddots & \vdots \\ (1-\mathrm{e}^{-t_M/\tau_1}) & (1-\mathrm{e}^{-t_M/\tau_2}) & \cdots & (1-\mathrm{e}^{-t_M/\tau_N}) \end{bmatrix} \tag{45}$$

Obviously this least square problem can be solved with Bertsekas algorithm also. Instead, additional insight is given by the singular values, defined below. Therefore, the least square problem is solved be means of the SVD of the matrix $\mathbf{A} \in \mathbb{R}^{M\times N}$, resulting in the matrix product:

$$\mathbf{A} = \mathbf{V}\,\mathbf{\Sigma}\,\mathbf{U}^{\mathrm{T}} = \sum_{i=1}^{N} \mathbf{u}_i \sigma_i \mathbf{v}_i^{\mathrm{T}} \tag{46}$$

with the orthonormal matrices of the singular vectors $\mathbf{U} = (\mathbf{u}_1, \ldots, \mathbf{u}_N) \in \mathbb{R}^{N\times N}$ and $\mathbf{V} = (\mathbf{v}_1, \ldots, \mathbf{v}_M) \in \mathbb{R}^{M\times M}$. The diagonal matrix $\mathbf{\Sigma} = \mathrm{diag}(\sigma_1, \ldots, \sigma_N)$ comprises the singular values. With (46) the solution $\mathbf{x}_{\mathsf{LS}}$ of the least square problem is given by

$$\mathbf{x}_{\mathsf{LS}} = \mathbf{A}^{\dagger}\mathbf{y} = \sum_{i=1}^{r} \frac{\mathbf{v}_i^{\mathrm{T}}\mathbf{d}}{\sigma_i}\,\mathbf{u}_i \tag{47}$$

with r as the rank of the coefficient matrix $\mathbf{A}$: $r = \mathrm{rank}(\mathbf{A})$. If the matrix $\mathbf{A}$ is invertible ($r = N$), then $\mathbf{A}^{\dagger}$ corresponds to the inverse $\mathbf{A}^{-1}$. Otherwise, the matrix $\mathbf{A}$ is rank deficient ($r < N$), i.e. the number of independent equations is smaller than the number of unknowns. In this case the matrix $\mathbf{A}^{\dagger}$ represents its pseudo-inverse.

Two characteristic features of the SVD are of interest with respect to the inverse problem [14]: First, the singular values σ_i may be ordered such that they gradually decrease to zero

$$\sigma_1 \geqslant \cdots \geqslant \sigma_N \geqslant 0 \tag{48}$$

If the matrix $\mathbf{A}$ is rank deficient ($r < N$), all singular values right of σ_r are zero,

$$\sigma_1 \geqslant \cdots \geqslant \sigma_r \geqslant 0, \quad \sigma_{r+1} = \cdots = \sigma_N = 0 \tag{49}$$

and the numerical schemes based on straight forward inversion of $\mathbf{A}$ usually fail. If one or more columns of the matrix $\mathbf{A}$ are nearly a linear combination of the other columns, the matrix is numerically rank deficient. Therefore, particular large differences between successive singular values may occur.

As a second feature of the SVD it is found that the number of sign changes in the components ('waviness') of the singular vectors $\mathbf{u}_i$ and $\mathbf{v}_i$ increases as the index i is enlarged. The series representation of the solution vector $\mathbf{x}_{\mathsf{LS}}$ in (47) shows that the solution is dominated by those singular vectors, which have many sign changes in their components, since their corresponding singular values are small and appear in the denominator. The uncertainty of the experimental data in the right-hand side vector $\mathbf{d}$ is amplified, leading to instabilities of the solution: small changes in the experimental data cause large changes in the solution. This observation is a consequence of ill-posed problems.

The regularization technique approaches this undesired phenomenon by introducing filter factors $f(\sigma_i, \alpha)$ as functions of the singular values σ_i and the regularization parameter α. The filter factors shall reduce the influence of those components of the SVD, which are amplified by the reciprocal of tiny singular values. The regularized least square solution is given by

$$\mathbf{x}_{\mathsf{LS}}^{\alpha} = \sum_{i=1}^{N} f(\sigma_i, \alpha)\frac{\mathbf{v}_i^{\mathrm{T}}\mathbf{d}}{\sigma_i}\,\mathbf{u}_i \tag{50}$$

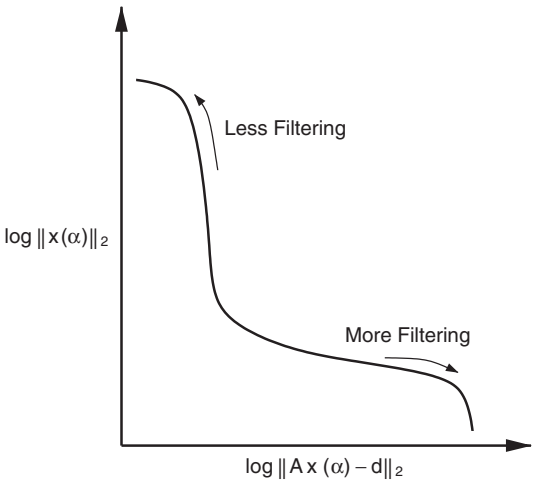

Figure 2. Plot of solution norm versus residual norm for L-curve-criterion of Hansen [14].

The most common regularization technique is the truncated SVD, where those members of the series representation in (50) are neglected, whose singular values σ_i are smaller than the regularization parameter α. Hence, an obvious filter function is

$$f(\alpha, \sigma_i) = \begin{cases} 1 & \text{for } \sigma_i \geqslant \alpha \\ 0 & \text{for } \sigma_i < \alpha \end{cases} \tag{51}$$

According to the idea of the Tikhonov regularization the least square problem in (43) is replaced by its regularized counterpart

$$\|\mathbf{Ax} - \mathbf{d}\|_2 + \alpha^2 \|\mathbf{x}\|_2 \rightarrow \text{Min} \tag{52}$$

which enforces the solution norm $\|\mathbf{x}\|_2$ to be small in addition to minimizing the residual norm $\|\mathbf{Ax} - \mathbf{d}\|_2$. The regularization parameter α controls the size of the solution norm in the weighted combination of the residual norm $\|\mathbf{Ax} - \mathbf{d}\|_2$ and the solution norm $\|\mathbf{x}\|_2$. A large α favors a small solution norm at the cost of a large residual norm. The resulting filter factor of the Tikhonov regularization becomes

$$f(\alpha, \sigma_i) = \frac{\sigma_i^2}{\sigma_i^2 + \alpha^2} \tag{53}$$

The choice of the optimal value for the regularization parameter α is the main difficulty in the applications. The L-curve criterion of Hansen [14, 15] is used here and illustrated in Figure 2. Both, the logarithm of the norm of the solution vector $\|\mathbf{x}(\alpha)\|$ of (52) and the logarithm of the defect norm $\|\mathbf{Ax}(\alpha) - \mathbf{y}^{\varepsilon}\|$ are displayed as functions of the regularization parameter α along the axes of the diagram. For discrete ill-posed problems the plot on the log–log-scale of the solution norm versus the residual norm is L-shaped in many cases. The optimal regularization parameter marks the point, where both, the residual and the solution norm, are small.

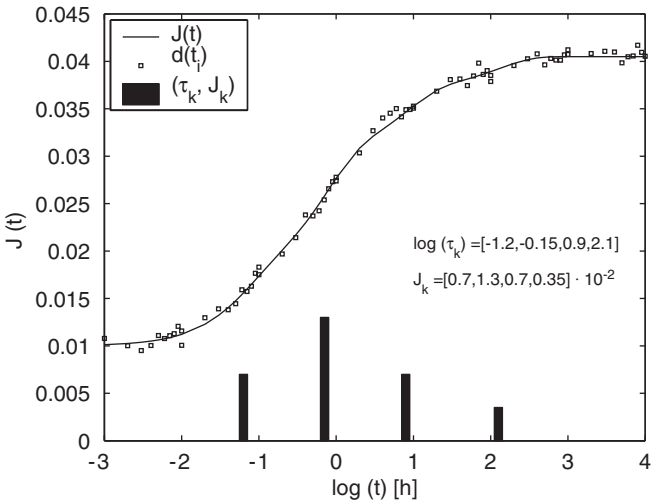

Figure 3. Given retardation spectrum (τ_k, J_k), resulting creep curve $J(t)$ and perturbed test data $d(t_i)$.

7. NUMERICAL EXAMPLES

The performance of the identification methods presented are studied via a set of artificially generated test data as well as through experimental data from the creep response of a unidirectional composite as determined by Sarabi [16].

7.1. Generated test data

In order to generate the test data, a discrete creep spectrum was assumed, as it is shown in Figure 3 by means of the four spectrum lines. Then roughly a 100 time points t_i are selected and the corresponding function values of the creep response $J(t_i)$ are chosen as the set of preliminary test data. These test data are perturbed with a random error of less than 5%. That means the function values are randomly modified, by adding or subtracting of up to 5% of their original value. The corresponding creep function and an example for the set of perturbed test data are shown in Figure 3. The results for the different identification methods are summarized and discussed in the following subsections.

7.1.1. Results of the projected Newton method. The main difficulty of the non-linear optimization method of Bertsekas is to provide appropriate starting values for the iterative solution of the unknown parameters. For an increasing number of Dirichlet–Prony terms it becomes more and more laborious to find appropriate starting values, for which the iterative algorithm converges. In the present investigation the spectrum parameters are successfully identified with up to four spectrum lines, involving not more than nine unknown material parameters. The starting values for the spectrum strengths are chosen as follows:

$$J_0 = J(t_{\min}), \quad J_k = [J(t_{\max}) - J(t_{\min})]/N \ \text{ for } k = 1, 2, \ldots, N$$

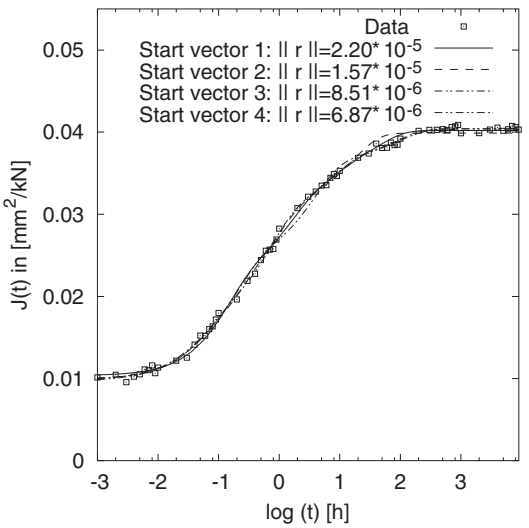

Figure 4. Fit of creep data for four different starting vectors $\mathbf{x}^{(i)}$.

with N as the number of Maxwell chains. Four different sets of starting values for the respondance times are examined

$$\log(\tau_k^{(1)}) = [-2, -1, 0, 1], \quad \log(\tau_k^{(2)}) = [-1, 0, 1, 2]$$

$$\log(\tau_k^{(3)}) = [-2, -0.6, 0.7, 2], \quad \log(\tau_k^{(4)}) = [-2, 0, 1, 3]$$

The data fitting in Figure 4 shows that the four different starting vectors supply four slightly differing approximations, although all the four curves fit the data points well. The associated line spectra are depicted in Figure 5(a), where the spectrum strength and relaxation times can be read off. The figure clearly illustrates that the identified spectra also differ from each other. Therefore, the unique solution for the material parameters is difficult to identify. A feasible approach for the determination of an acceptable result may be the following strategy: If different starting vectors lead to almost the same set of parameters, the solution likely represents a stable, local minimum of the least square function (4) and may indicate a global minimum.

The identified line spectra for five different sets of perturbed test data reveal in Figure 5(b) that small changes in the data cause only small deviations of the identified parameters.

7.1.2. Results with Tschebyscheff/Wolfe identification. Figures 6(a) and 7(a) show the approximation of the test data by two different tendency functions.

An obvious choice is the hyperbolic tangent-function:

$$J_T(t) = J_0 + b \tanh(ct)^d$$

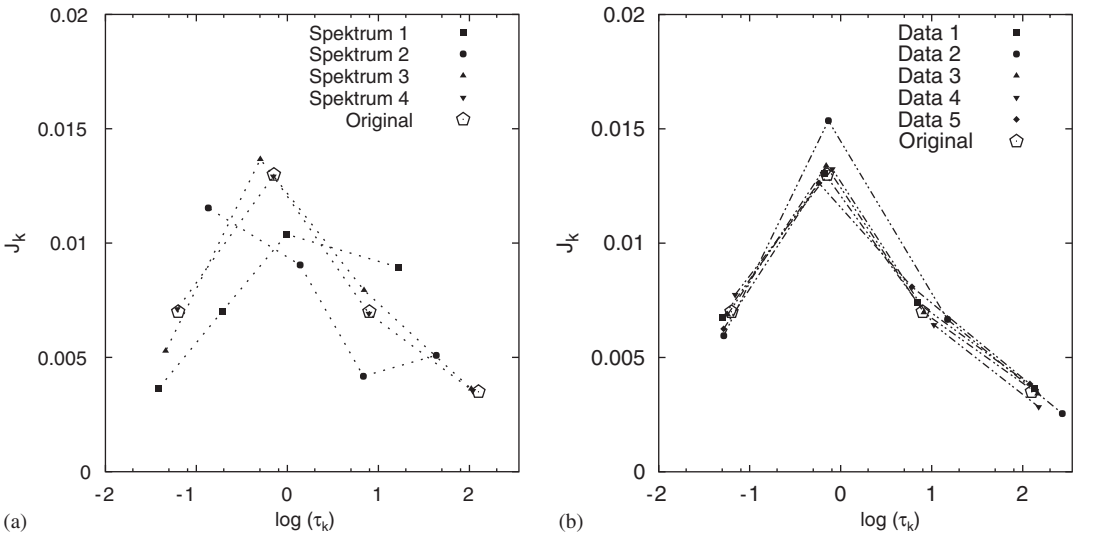

Figure 5. (a) Identified line spectra; and (b) line spectra for five different sets of perturbed test data found with starting vector $\mathbf{x}^{(4)}$.

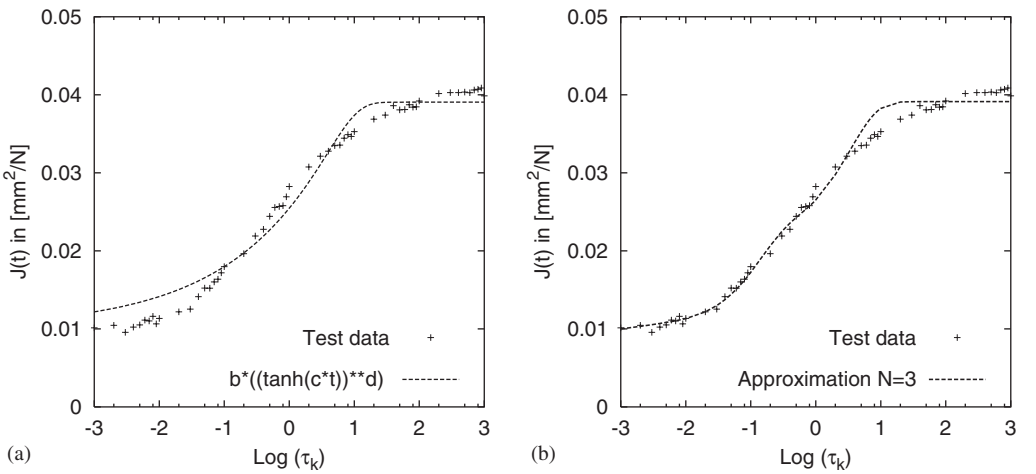

Figure 6. (a) Approximation with tendency function $J_T(t) = b\ \tanh(ct)^d$; and (b) approximation with three exponentials.

with

$$J_0 = 0.010129, \quad b = 0.028926, \quad c = 0.113593 \text{ and } d = 0.291198$$

and after repeatedly plotting the discrete data points in log–log-diagrams it follows:

$$J_T(t) = J_0 + 10^{(-10^a \cdot t^b + u)}$$

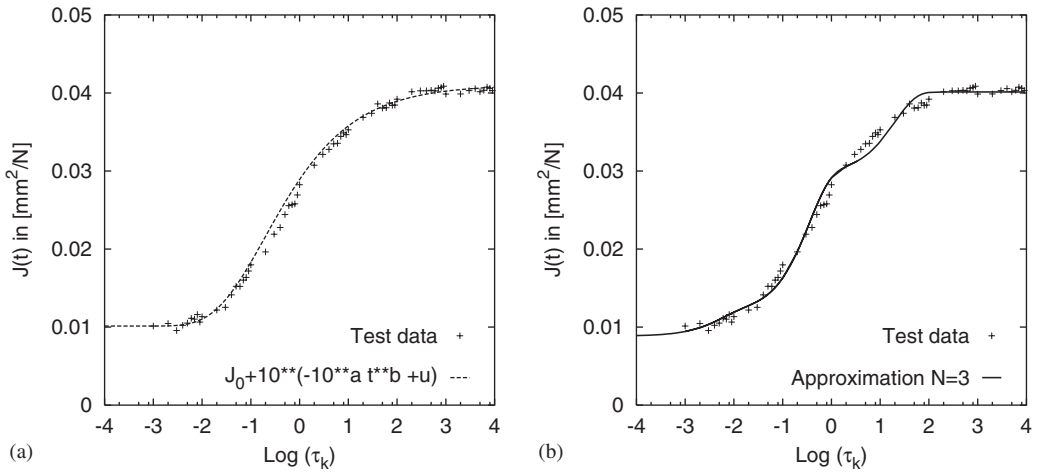

Figure 7. (a) Approximation with tendency function $J_T(t) = J_0 + 10^{(-10^a t^b + u)}$; and (b) approximation with three exponentials.

with

$$J_0 = 0.010129, \quad a = -0.665401, \quad b = -0.426788 \text{ and } u = -1.51$$

The results of the Tschebyscheff-approximation and the successive enhancement with the Simplex-method by Wolfe are given in Figures 6(b) and 7(b) for a finite Dirichlet–Prony series with three exponentials. For the Tschebyscheff approximation with four exponential terms no successful starting vector has been found to converge the algorithm. The corresponding line spectra are shown in Figures 3(a) and (b). The approximation of the test data is not very good in the first case—see Figure 6(b), but better in the second one—see Figure 7(b), because the hyperbolic tangent as tendency function obviously differs from the data around time $t = 10$. An improvement of the approximation for the exponential function in Figure 7(b) may be achieved by increasing the number of Dirichlet–Prony terms. However, it becomes increasingly more difficult to find good starting values for the convergence of the Tschebyscheff iterative method with a growing number of terms in the Dirichlet–Prony series.

For the successful convergence of the Tschebyscheff iteration scheme it is necessary to provide starting values for the respondance times τ_k, which are sufficiently close to the unknown solution. Furthermore, it becomes less likely to find an alternating error function with $2N + 2$ extrema, if the number of Dirichlet–Prony terms is large, because the intervals, in which the individual exponentials e^{-t/τ_k} are clearly time-dependent functions, may be spaced to narrowly.

Note, that the simple power function γt^β cannot fit the test data well, since for more than three decades the data points follow the course of exponentials, which are bounded from above, because their argument is negative. Therefore, the curve fitting with a Dirichlet–Prony series is not successful, if the respondance times are estimated with the Tschebyscheff approximation of the power function in this case (Figures 8(a) and (b)).

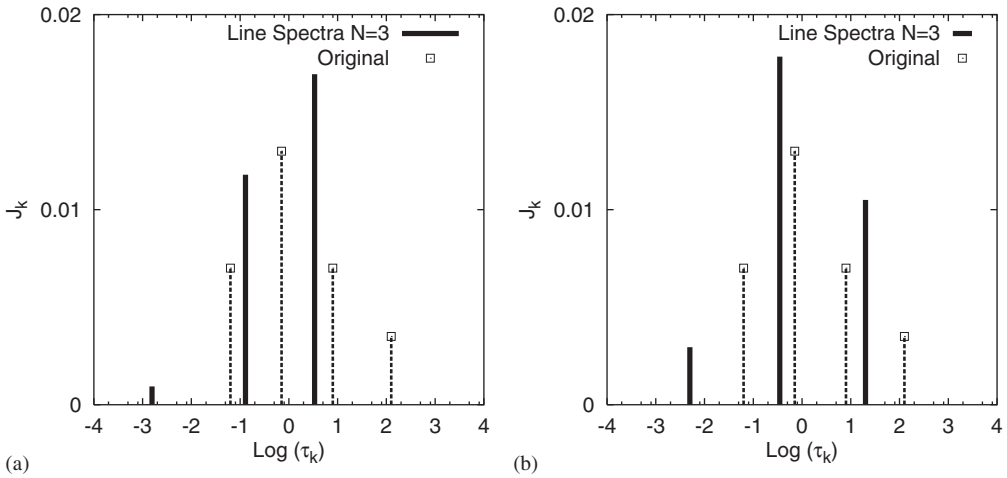

Figure 8. Original and identified line spectrum: (a) based on tendency function $J_T(t) = b\ \tanh(ct)^d$; and (b) based on tendency function $J_T(t) = J_0 + 10^{(-10^a t^b + u)}$.

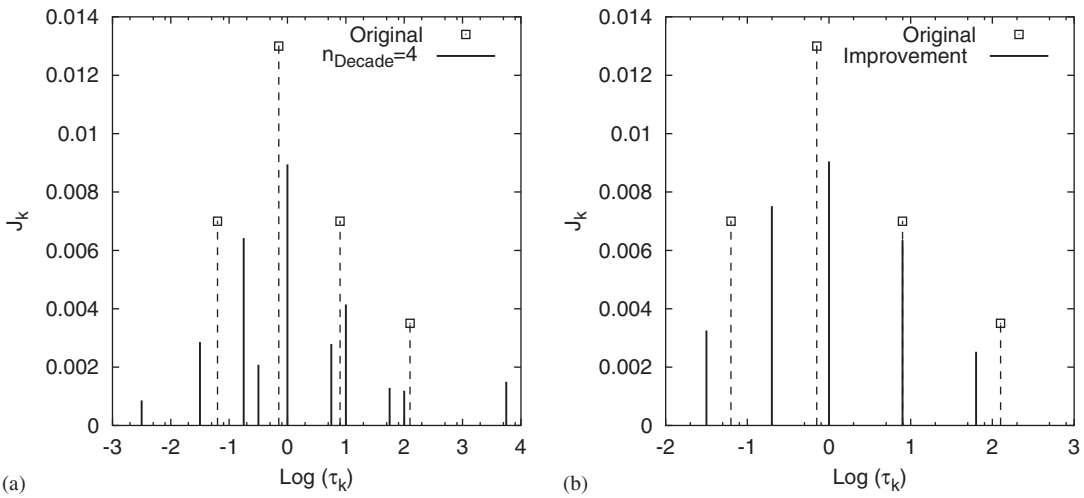

Figure 9. (a) Original and computed line spectrum found with four respondance times per time decade; and (b) line spectrum with reduced set of respondance times.

7.1.3. Results of windowing method. Starting with four preselected respondance times per time decade, equally spaced between $-3.0 \leqslant \log \tau \leqslant 4.0$, the windowing method uses in total $N = 29$ retardation time in the beginning and, finally, provides the discrete spectrum with 10 significant lines from the perturbed test data as given in Figure 9(a). The other strength values turn out to be zero or close to zero. The spectrum shows irregular jumps between neighbouring spectrum lines in comparison to the ordinal one. In order to 'smooth' the spectrum, a new set

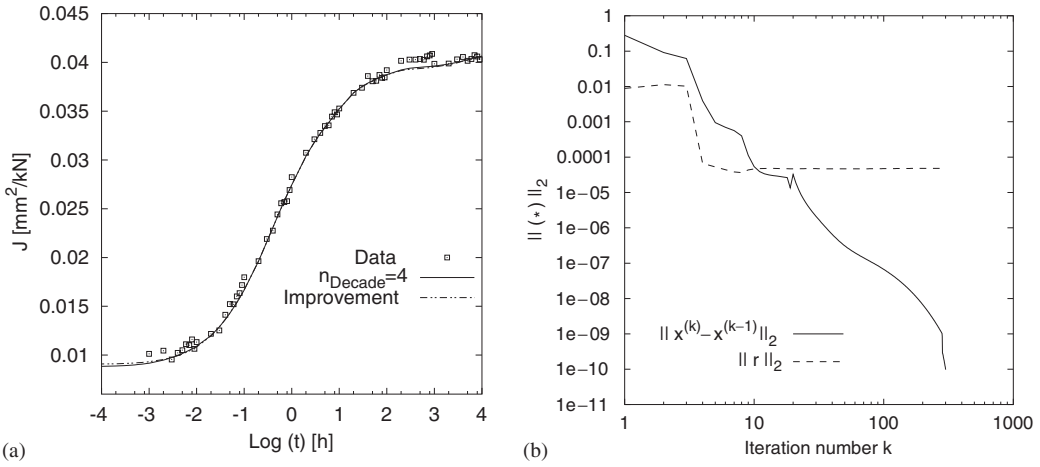

Figure 10. (a) Approximated creep curves; and (b) norm of parameter change and residual norm.

of respondance times is selected for the repeated run of the algorithm. Based on the previous result, the number of preselected respondance times is increased in the time range, where densely spaced spectrum lines appear with strengths clearly greater than zero. Vice versa, in the time range, where the spectrum strengths are nearly zero, the associated respondance times are neglected in the new set for the next run. By means of this strategy the number of spectrum lines can obviously be reduced as shown by the improved solution in Figure 9(b). The creep curves, corresponding to the spectra in Figures 9(a) and (b), nearly coincide and approximate the test data well—see Figure 10. Figure 10(b) shows the Euclidean norm of the change of the computed parameters $||\mathbf{x}^{(n)} - \mathbf{x}^{(n-1)}||_2$ between successive iteration steps versus the number of iterations. As soon as this norm is smaller than a given tolerance, the iteration is terminated and the current spectrum is taken as the final solution. In Figure 10(b) also the residual norm $\|\mathbf{Ax} - \mathbf{d}\|_2$, also denoted as the error norm, is plotted as a function of the iteration counter.

7.1.4. Results with regularization. For the approximation of the test data 10 respondance times are preselected, which are equidistantly distributed on the logarithmic time scale.

$$\log \tau_k = [-1.5,\ -1.0,\ -0.5,\ 0.0,\ 0.5,\ 1.0,\ 1.5,\ 2.0,\ 2.5,\ 3.0] \tag{54}$$

In Figure 11 the computed retardation strengths are presented together with the results for the truncated SVD—see Figure 11(a)—and for the Tikhonov regularization—see Figure 11(b). In contrast to the non-regularized solution with 10 non-zero singular values ($n_\sigma = 10$) all strengths computed from the truncated SVD method satisfy the constraints of non-negative strength values in (6), if not more than 8 non-zero singular values ($n_\sigma \leqslant 8$) are taken in this case. The same holds for the strength values calculated from the Tikhonov regularization with a parameter value of $\alpha = 0.2373$ according to the L-curve criterion. In consequence of the regularization the identified line spectra show no longer large irregular jumps between neighbouring spectrum lines, as it is the case for the un-regularized solution—see Figure 11(a) in the case of $n_\sigma = 10$ or Figure 11(b) in the case of $\alpha = 0.0$.

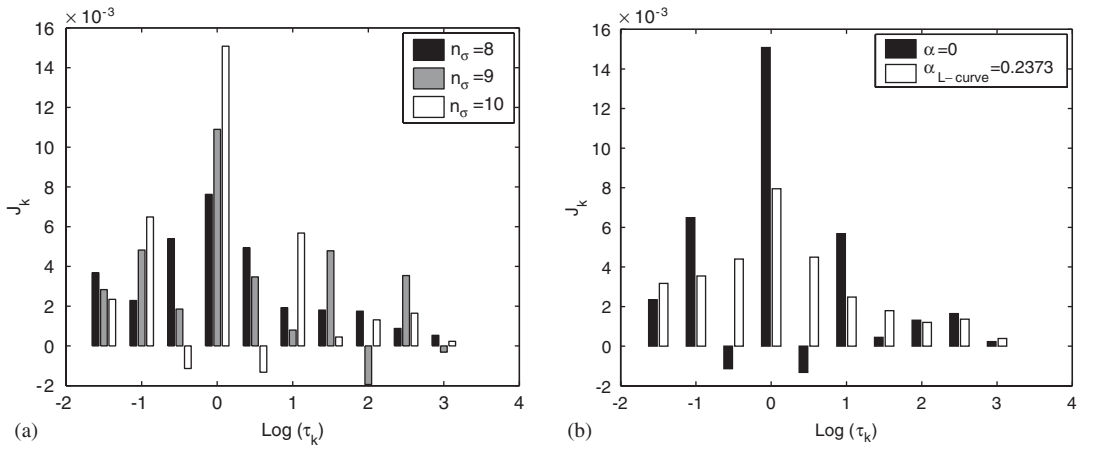

Figure 11. Identified retardation spectra with: (a) truncated SVD; and (b) Tikhonov regularization.

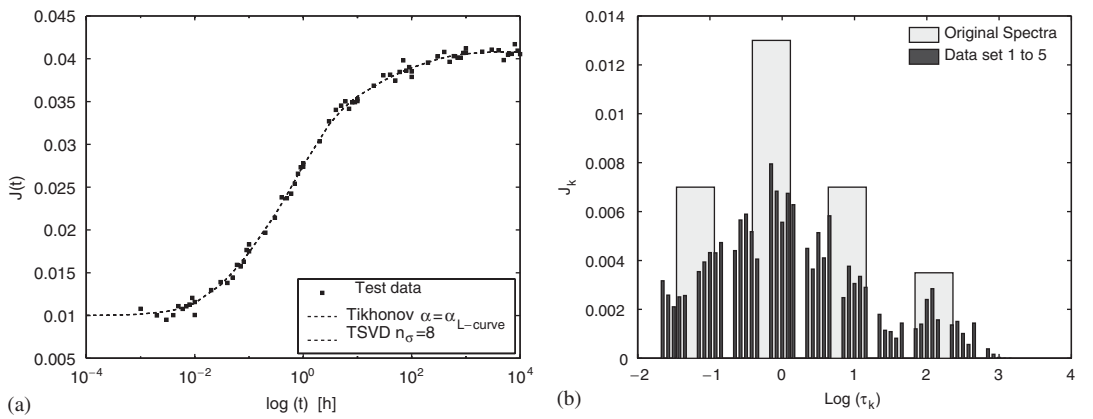

Figure 12. (a) Curve fit of test data; and (b) original and identified retardation spectra for five different sets of perturbed test data.

The associated creep functions are plotted in Figure 12(a), which shows an excellent fitting of the test data. In Figure 12(b) the identified spectra for five different perturbations of the creep data are compared. All spectra, regularized with Tikhonov's method, are very similar to each other for the different perturbations. It demonstrates the stability of the regularization method with respect to perturbations of the data. The distribution of lines of the identified spectra is comparable to the one of the original spectrum. The solutions of the inverse problem with five different perturbations may be viewed as approximations to the original spectrum, where the four spectrum lines are replaced by ten lines, however, with smaller values for the strengths.

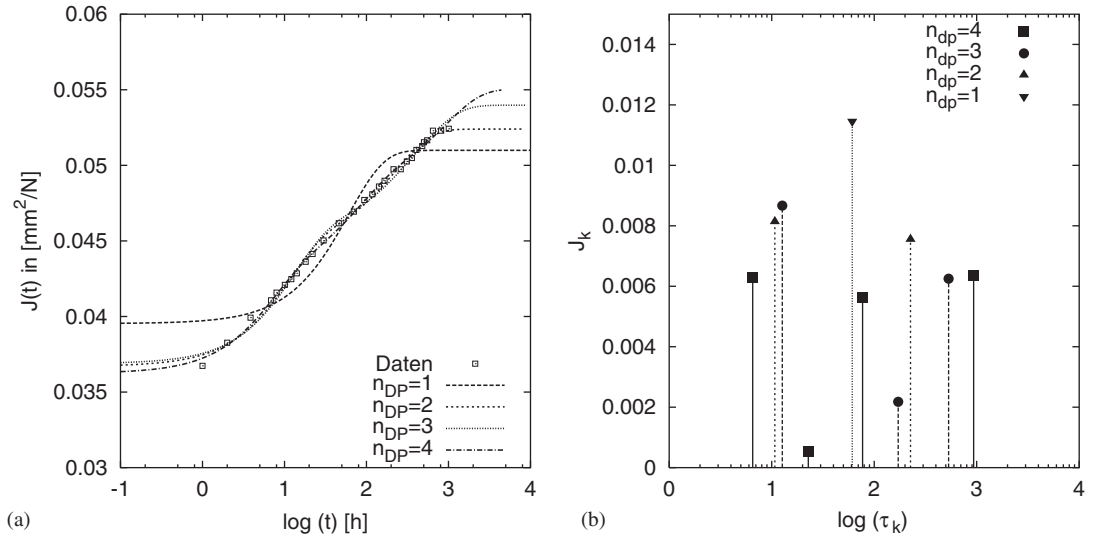

Figure 13. (a) Curve fit of experimental data; and (b) identified spectra with one to four spectrum lines.

7.2. Experimental creep data of Sarabi [*16*]

The identification algorithms presented above are applied to the experimental creep data of Sarabi [16]—see Figure 13(a). The plots show that on the logarithmic time scale all data points are close to a straight line.

7.2.1. Results of projected Newton method. The performance of the non-linear optimization algorithm is studied via the 'experimental' creep data, which are approximated with one to four Dirichlet–Prony terms. Besides the starting values for the parameters only the number of exponentials must be specified. According to Figure 13(a) the fitting of the experimental data can be substantially improved by increasing the number of Maxwell chains in the series as expected. Figure 13(b) shows the calculated line spectra with one to four exponential terms, which were determined by several repeated runs of the algorithm with different starting values. The approximations may be expected to be the minimizer of the least square functional.

7.2.2. Results with Tschebyscheff/Wolfe identification. Figure 14(a) shows the crude approximation of the experimental data by the power law

$$J_{\mathrm{T}}(t) = \gamma t^{\alpha}$$

with

$$\gamma = 0.00492666 \text{ and } \alpha = 0.29972351$$

However, the approximation is adequate for the purpose of estimating the respondance times by a tendency function. Other elementary functions or a combination of them may turn into a better tendency function. One to four respondance times are estimated with the Tschebyscheff

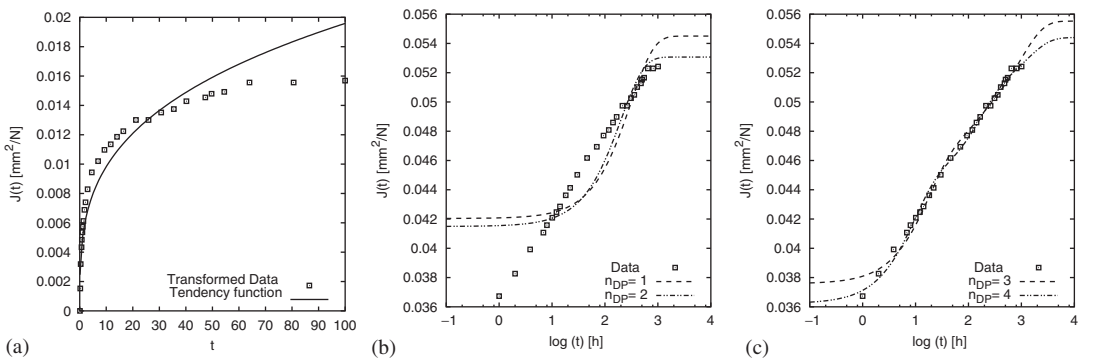

Figure 14. (a) Tendency function and creep data; (b) approximation with one or two spectrum lines; and (c) approximation with three or four spectrum lines.

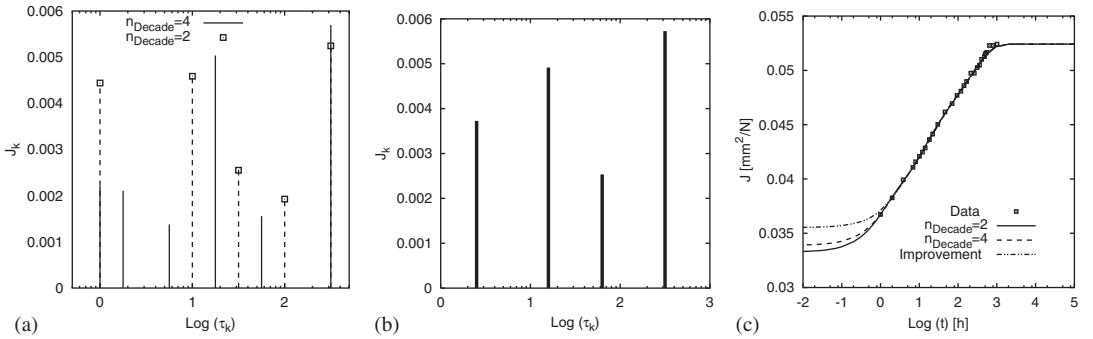

Figure 15. (a) Final line spectra found from two and four preselected respondance times per decade; (b) simplified line spectrum with improved respondance times; and (c) time approximation of creep data.

approximation of the power function. The strengths, belonging to the creep times, are found with Wolfe's method. The diagrams in Figures 14(b) and (c) show that an increasing number of exponential terms obviously improves the quality of the creep data approximation. The discontenting results in Figure 14(b) are based on two spectrum lines, whereas three and four exponential terms prove satisfactory as shown in Figure 14(c).

7.2.3. Results of windowing method. In the first step the creep data of Sarabi are approximated by selecting equally distributed respondance times with either two or four values per decade. This approach leads to the two line spectra in Figure 15(a) with satisfying approximations as shown in Figure 15(c).

The line spectrum, identified with four preselected respondance times per decade, contains altogether six spectrum lines, where three are close together between $\log \tau_k = 0.0$ and $\log \tau_k = 0.75$. This phenomenon indicates that an improved distribution of the spectrum strengths as well as a reduction of the number of identified spectrum lines may be possible. Therefore,

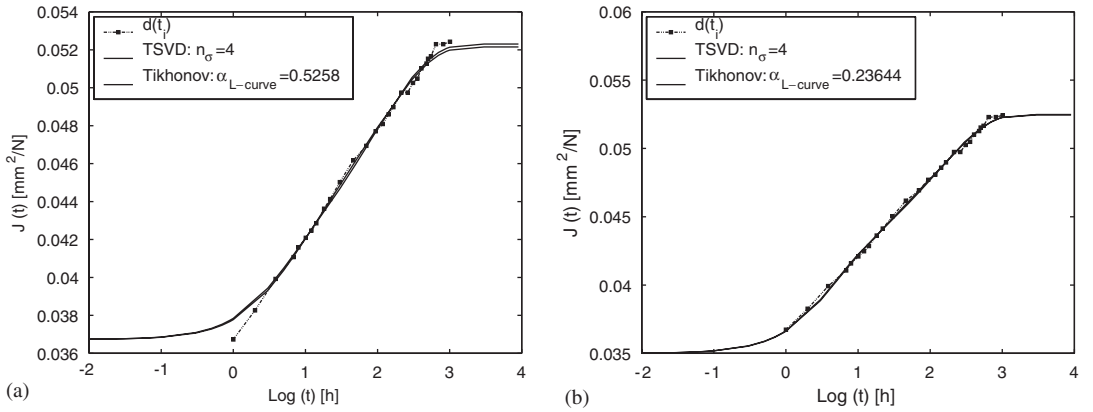

Figure 16. Approximation of creep data with different assumptions for the elastic compliance: (a) $J_0 = J(t_{\min})$; and (b) $J_0 = 0.035$ [mm^2/N].

additional respondance times are assumed in the range, where several spectrum strengths are clearly greater than zero. By repeating the windowing method, the final spectrum can be found, consisting now of only four lines—see Figure 15(b). The enhanced approximation of the data with this simplified spectrum is shown in Figure 15(c) together with the creep curves of the previous spectra in Figure 15(a). The algorithm of Emri/Tschoegl is based on the assumption that the last data point in time corresponds to the equilibrium stage with vanishing creep rates and the elastic compliance is determined from the following difference: $J_0 = J(t_{\max}) - \sum_{i=1}^{N} J_i$. As a consequence of this setting the constant term J_0 of the Dirichlet–Prony series is smaller than the creep value of the first datum point in the time interval.

7.2.4. Results with regularization. The estimate of appropriate respondance times forms the basis for the identification of the unknown material parameters with regularization of the inverse. For a suitable choice of creep times no negative values for the spectrum strength may arise necessarily. The following procedure proved to be useful.

First, the respondance times are selected such that they are equidistantly distributed within the experimental time interval on the logarithmic axis. Not too many respondance times per time decade may be chosen, i.e. a maximum of four per decade turns out to be sufficient, in order to avoid almost linear dependence of the rows in the matrix **A**. Otherwise, the condition number of the coefficient matrix becomes to large. If the preselected respondance times on the left-hand side of the time interval of the spectrum ($\min.\tau_k$) are too small, the SVD-method may yield negative values for the corresponding spectrum strengths.

In case the chosen creep times at the right-hand side of the spectrum ($\max.\tau_k$) are too large, the resulting creep curve keeps still growing after the last experimental data point is passed. Hence, the last experimental datum point in time is not consistent with the function value of the asymptote for the exponential series. On the basis of the aforementioned criteria the experimental data were approximated by five exponential sum terms with the following creep times: $\log(\tau_k) = [0.5,\ 1.0,\ 1.5,\ 2.0,\ 2.5]$ for $k = 1, 2, \ldots, 5$.

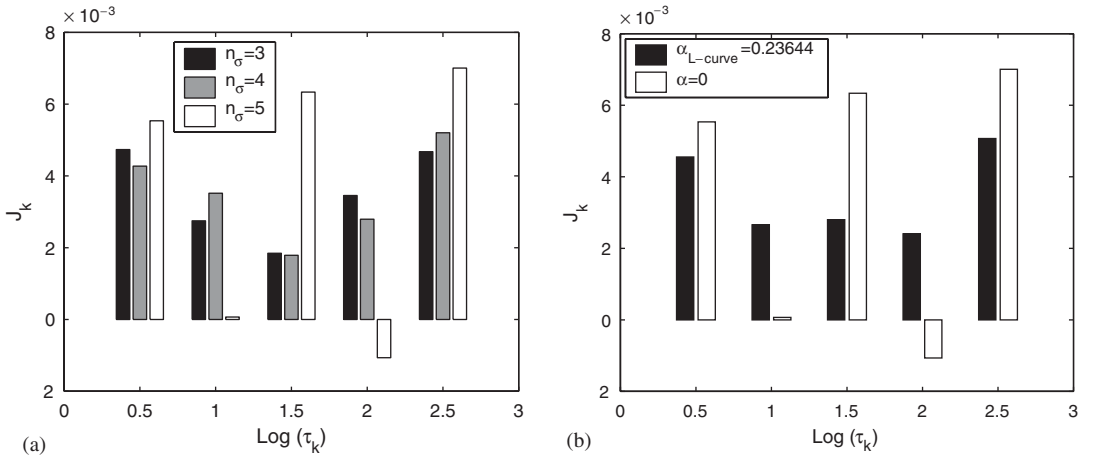

Figure 17. Regularized spectra for $J_0 = 0.035$ [mm^2/N] with (a) Truncated SVD; and (b) Tikhonov method.

In the first approach the elastic compliance J_0 is assumed to be equal to the function value of the experimental data pair at the first point $t_{\min}$ in time. Hence: $J_0 = J(t_{\min})$. The resulting creep curves, regularized by either the TSVD ($n_\sigma = 4$) or the Tikhonov method ($\alpha = 0.5258$), exhibit insufficient approximations of the data points with small time values at the beginning—see Figure 16(a). The approximation of the creep process at the early stage may be enhanced by manually adjusting the elastic compliance to $J_0 = 0.035$[mm^2/N], which yields the creep curves in Figure 16(b). They differ only with respect to the type of regularization, which are the truncated singular value decomposition (TSVD) with four non-zero singular values ($n_\sigma = 4$) and the Tikhonov method with $\alpha = 0.2364$. The corresponding discrete spectra, belonging to the creep curves in Figure 16(b), are drawn in Figure 17(a) and (b) for the two different regularization methods. The spectrum for the SVD, truncated after the third singular value ($n_\sigma = 3$), is added to Figure 17(a). The regularized spectra may be compared to the non-regularized solution, which violates the constraint in (6) for the spectrum line at $\log(\tau_4) = 2.0$.

8. COMPARISON

The various identification methods require individual interactions with the user, provision of additional information and different numerical effort. Figure 18(a) shows that all identification schemes provide excellent approximations of the creep data in the time range of the experiment. The time range of the plots in Figure 18(b) is extended by an additional decade beyond the range of the measured data in the logarithmic representation. Thus, the creep curve should be plotted for each solution of the spectrum and compared to those derived from the other identification schemes. Special attention must be given to the last data points, indicating whether the creep process tends to its final equilibrium stage, defined by vanishing creep rates.

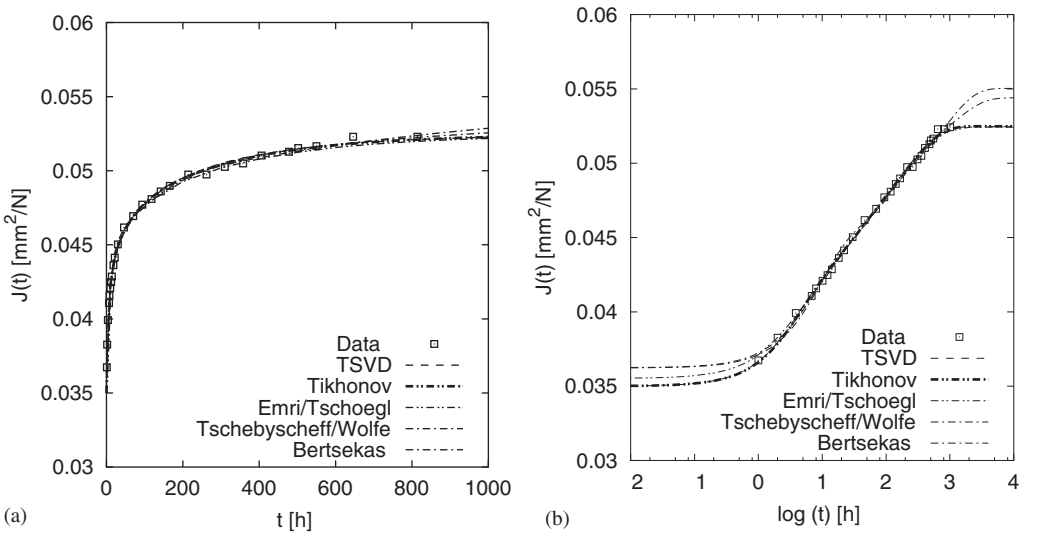

Figure 18. Various approximations of creep data from [16]: (a) linear time scale; and (b) logarithmic scale.

No additional information about the parameters prior to the identification is required for the straight forward solution of the non-quadratic least square functional with the optimization method of Bertsekas and the Tschebyscheff approximation of an appropriate tendency function with redetermination of the spectrum strengths. Both procedures treat the curve fitting problem of the data with the chosen model function as purely mathematical one. A disadvantage of these methods is that the convergence of the iterative procedures becomes more difficult with an increasing number of Dirichlet–Prony terms. The algorithms converge to stable minima of the objective functions for up to four exponential terms. If four or more sum terms are chosen, the starting values must lie in the vicinity of the objective function in order to achieve convergence.

The solution of the objective function with the projected Newton method of Bertsekas often depends on the starting values, since the sequence of iteratively computed parameters may converge to a local minimum. In practice the following point of view may be taken: if different starting vectors lead to nearly the same set of solution parameters, the final result represents at least a stable local minimum, which likely is also the global one.

Additional information about the unknown parameters is required for the windowing method as well as for the regularization schemes, as the distribution of the respondance times have to be assumed in advance. The set of spectrum lines can be reduced in the case of a qualified estimate of the respondance times. These methods have the advantage that they supply stable results, although the identified parameters may not represent the mathematically best approximation. But from an engineering point of view the identified variables represent the 'better' material parameters, since the discrete viscoelastic line spectrum can be interpreted as an approximation of a continuous spectrum generally.

Figure 19 shows the difference between the line spectra for Sarabi's experimental creep data [16] in connection with the various identification methods. Although the respondance times are within the same time range from $0 < \log \tau < 3$, the distributions of the spectrum lines

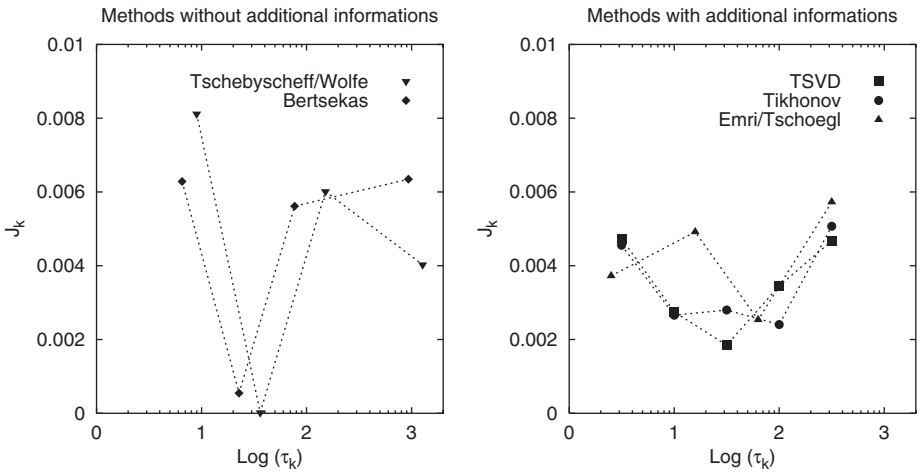

Figure 19. Identified line spectra of the different identification methods.

are quite different. The spectra from the methods without additional information show large differences between neighbouring spectrum lines. However, the other methods require additional information, but provide a smoother distribution of the line spectrum.

With the exception of some solutions from the regularization methods, all other schemes do not yield negative parameters, which are physically unrealistic. Negative spectrum strengths can be avoided also in the context of the regularization methods through a qualified choice of the respondance times as explained above.

9. CONCLUSIONS

In this paper, four different numerical methods were investigated for the identification of the discrete viscoelastic line spectra from theoretically generated data and experimental creep data. The non-negative constraints to the unknown material parameters are taken into account as an important aspect of the identification task. The methods presented may be classified into two main groups. The identification algorithms of the first class, the projected Newton method of Bertsekas and a combined approach, based on the Tschebyscheff approximation of a tendency function, require no estimates of the unknown creep times. These strategies have the advantage that the user needs no experience with the estimation of the creep times for the non-linear identification task.

The windowing method of Emri and Tschoegl as well as the regularization techniques, based on the SVD, are assigned to the second class of identification algorithms. These latter methods require an assumption about the distribution of the respondance times. Then they lead to a quadratic optimization problem.

All four method were successfully applied to test data as well as to experimental data and the results presented were compared to each other.

ACKNOWLEDGEMENTS

The authors gratefully acknowledge helpful discussions with Dr M. Brede and Prof. Dr W. Sippel and the financial support of the Deutsche Forschungsgemeinschaft (DFG) through the Graduiertenkolleg 'Identifikation von Material- und Systemeigenschaften'.

REFERENCES

1. Gerlach S, Matzenmiller A, Sippel W. Identification of retardation spectra by approximation theory. *Zeitschrift fur Angewandte Mathematik and Mechanik* 2001; **81**:S293–S294.
2. Schapery RA. A simple collocation method for fitting viscoelastic models to experimental data. *Technical Report 61-23 A*, California Institute of Technology, Pasadena, CA, 1961.
3. Clauser JF, Knauss WG. On the numerical determination of relaxation and retardation spectra for linearly viscoelastic materials. *Transactions of the Society of Rheology* 1968; **12**(1):143–153.
4. Cost TL, Becker EB. A multidata method of approximate laplace transform inversion. *International Journal for Numerical Methods in Engineering* 1970; **2**:207–219.
5. Bradshaw RD, Brinson LC. A sign control method for fitting and interconverting material functions for linear viscoelastic solids. *Mechanics of Time-Dependent Materials* 1997; **1**:85–108.
6. Tschoegl NW, Emri I. Generating line spectra from experimental response. Part I: relaxation modulus and creep compliance. *Rheologica Acta* 1993; **32**:311–321.
7. Emri I, Tschoegl NW. Generating line spectra from experimental response. Part IV: application to experimental data. *Rheologica Acta* 1994; **33**:60–70.
8. Emri I, Tschoegl NW. Determination of mechanical spectra from experimental response. *International Journal of Solids and Structures* 1995; **32**(6/7):817–826.
9. Bertsekas DP. Projected newton methods for optimization problems with simple constraints. *SIAM Journal on Control and Optimization* 1982; **20**:221–246.
10. Meinardus G. *Approximation von Funktionen und ihre numerische Behandlung*. Springer: Berlin, 1964.
11. Bunday BD. *Optimisation Methods in PASCAL*. Edward Arnold: London, 1987.
12. Honerkamp J. Ill–posed problems in rheology. *Rheologica Acta* 1989; **28**:363–371.
13. Honerkamp J, Weese J. Tikhonovs regularization method for ill-posed problems. *Continuum Mechanics and Thermodynamics* 1990; **2**:17–30.
14. Hansen PC. *Rank-deficient and Discrete Ill-Posed Problems*. SIAM: Philadelphia, 1997.
15. Hansen PC. Regularization tools: a matlab package for analysis and solution of discrete ill-posed problems. *Numerical Algorithms* 1994; **6**:1–35.
16. Sarabi B. Das Anstrengungsverhalten von Polymerwerkstoffen infolge ein- und zweiachsigen Kriechens: Ermittlung von Langzeitbemessungskennwerten. *Ph.D. Thesis*, University Kassel, 1984.
17. Baumgaertel M, Winter HH. Determination of discrete relaxation and retardation time spectra from dynamic mechanical data. *Rheologica Acta* 1989; **28**:511–519.
18. Tschoegl NW. *The Phenomenological Theory of Linear Viscoelastic Material Behaviour*. Springer: Berlin, 1. edition, 1989.

