ASME Journal of Tribology
Online journal at:
https://asmedigitalcollection.asme.org/tribology

The Viscoelastic Coefficient of Restitution: Integrating the Wave Equation With Hereditary Physics

Itzhak Green
Green MechaniX,
Cumming, GA 30041
e-mail: itzhak.green@gmx.com

The viscoelastic coefficient of restitution (vCOR) is a functional of the entire constitutive law, including relaxation, retardation, memory kernels, impact velocity, and its duration. This work predicts the vCOR by embedding hereditary material behavior into a Zener-type wave model for sphere-plate impact. Impact generates both local deformation and stress waves, which attenuate rapidly in viscoelastic bodies due to intrinsic molecular dissipation. Various existing contact models typically relate force to deformation and deformation rate through forms of stiffness and damping parameters, but they neglect true hereditary viscoelasticity. In contrast, hereditary behavior requires that stresses depend on the full strain history, as described by Boltzmann's superposition principle and, for nonlinear intrinsics, the Volterra integral framework. Here, the Hertzian point contact force is reformulated to instill hereditary viscoelasticity into Zener's wave equation. A physically consistent separation condition is imposed to determine an effective coefficient of restitution, accounting also for the residual deformation present at separation in viscoelastic impacts. The associated hysteretic energy loss is quantified, and a parametric study is conducted. Two solution strategies are developed: a direct numerical integration of the nonlinear ordinary integro-differential wave equation, and an internal variable evolution formulation that converts the problem into a system of ordinary differential equations for rapid time integration. Both approaches yield independently identical numerical results. [DOI: 10.1115/1.4071256]

Keywords: viscoelastic coefficient of restitution, hereditary viscoelasticity, convoluted Hertzian contact, Zener wave model, Maxwell material model, fractional calculus model, contact mechanics, indentation, oscillations

1 Introduction

The coefficient of restitution (COR) is, by necessity, an ad hoc construct introduced to facilitate the analysis and solution of impact problems. In its basic form, it is defined as the ratio of the restitution impulse to the deformation impulse, which is equivalently expressed as the ratio between the relative velocity of separation after collision and the relative velocity of approach before collision. The material composition and its constitutive character play a central role in governing the dynamic response and the redistribution of energy during and after impact. A viscoelastic coefficient of restitution (vCOR) emerges in collisions when at least one contact body exhibits a contact force that includes a convolution with a viscoelastic kernel. Viscoelastic materials differ fundamentally from purely elastic ones in two aspects: (1) their dynamic responses are inherently hereditary, and (2) they dissipate energy intrinsically. Consequently, the vCOR generalizes the classical COR concept to materials exhibiting these behaviors, embedding the additional complexities required to capture hereditary and dissipative effects.

In addition to factors that affect the classical COR, a vCOR depends on impact velocity, relaxation spectrum of the hereditary kernels (Prony, fractional, etc.), the interplay between impact duration and relaxation times, and a delayed recovery (being a major effect in real viscoelastic solids). Real viscoelastic bodies separate before they fully recover their shape, which reduces COR values relative to elastic predictions. For low velocity impacts, the vCOR is shown to be governed primarily by the ratio of loss modulus to storage modulus, i.e., the phase lag between stress and strain. Because the impact duration competes with the relaxation times, viscoelastic materials can exhibit increasing or decreasing vCOR with velocity depending on whether the glassy or rubbery regime dominates. All that shall be studied and realized herein.

The utility of the COR is vast in solving problems of engineering, science, medicine, sports, and recreation, where, once interjected, it provides the missing information (i.e., an equation) to complement the conservation of momentum, to readily solve impact problems. While it is customary to determine the COR by running experiments that mimic actual conditions of the collision and offer empirical COR values, it is beneficial, however, to predict the COR values analytically.

Various models have been offered over the years tracing back to the historic paper by Zener [1] in 1941 who investigated the elastic impact between a sphere pinging a laterally wide plate of finite thickness, and to as recent as that by Green [2], who fuses the said Zener model with an elastoplastic Jackson, Green, and Marghitu (JGM) model [3] to account for more severe impacts. The fused analytical COR predictions in Ref. [2] match meticulously

Contributed by the Tribology Division of ASME for publication in the Journal of Tribology. Manuscript received January 9, 2026; final manuscript received February 13, 2026; published online March 17, 2026. Assoc. Editor: Guo Xiang.

experimental and FEA results obtained by Higgs et al. [4–6]. The works by Jackson et al. [3] and Higgs et al. [4–6] have each detailed substantial literature coverages, along with the extensive literature review on impact problems and COR models, as well as by Banerjee et al. [7]. These reviews are quite recent and relevant. There are other recent publications on impact calculations including wave propagation in both colliding bodies [8–11]. These usually require expensive computational effort, and they are alternative approaches to the classical approaches. Central to COR modeling is the nature of the contact force. Chen et al. [12] tabulate an additional 15 references where each offers a different contact force model consisting of a "contact stiffness" which multiplies odd functional forms and combinations of the deformation, its time derivative, and where the COR is *part* of the functional forms so that the COR is somehow "known a priori." None of these models is capable of or intended to handle the hereditary nature of viscoelastic materials. Henceforth, another massive literature coverage is therefore sidestepped in the interest of conciseness. This work will use the comprehensive coverage of viscoelasticity by Mainardi [13], Gutierrez-Lemini [14] (without prejudice, as there are numerous first-rate textbooks on the topic), the author's work [15–21], and others as needed within.

The derivation of Zener's [1] wave equation is relevant and fundamental to this work, so his model is briefly summarized; however, and most importantly, his solution is particular to elastic materials, which is extraneous for viscoelastic impact. So, only Zener's initial derivation of the wave equation is appropriate, but that equation requires modifications to capture the distinct way viscoelastic materials behave dynamically, and the rest of his analysis is abandoned.

This work is organized as follows: (i) the Zener acoustic model is briefly covered, casting a modified wave equation along with the initial conditions, and setting the stage to (ii) the viscoelastic formulation by expressing Hertzian contact theory in a convolution form to be used in Boltzmann's superposition integral, and obtain the governing nonlinear integro-differential equation, (iii) selecting first a standard linear solid (SLS, also known as a Zener material model) as the testbed material to establish two autonomous methods of solution: the direct formulation which incorporates the Volterra series to handle the nonlinearity and the internal variable evolution (IVE) formulation, (iv) then select a real material (BUTYL-B252) for the plate and re-apply the said methods of solution, (v) perform a parametric study to gauge the effects of the impact velocity, and lastly (vi) conclude.

2 The Zener Acoustic Model

The Kirchhoff–Love thin plate model provides the foundation for the wave equation derived by Zener [1] for calculating the COR, which this work further extends. That derivation is universal for this class of problems of sphere-plate impact. The model ushers a nonlinear governing equation to describe a wave that propagates in a laterally wide plate as instigated by an impacting sphere. As the wave rebounds, it ejects the sphere from the plate surface after a finite contact duration and at a reduced velocity, thereby giving rise to the COR. The present work adopts Zener's wave equation development; however, most aspects of Zener's original analysis are extraneous to the viscoelastic impact problem considered here. The development of the governing wave equation is presented next.

For consistency and clarity, Zener's original nomenclature for variables is adhered to. Subscripts a and b designate, respectively, the sphere and the plate. Starting with the equivalent elastic moduli of the contacting bodies:

$$E'(t)=\frac{E'_a(t)E'_b(t)}{E'_a(t)+E'_b(t)} \quad \text{and} \quad E'_i(t)=\frac{E_i(t)}{1-\nu_i^2}, \quad i=a,\,b \tag{1}$$

Poisson's ratio, ν, for most viscoelastic materials can be taken as 0.5, being typical for incompressible materials (e.g., elastomeric rubber), but other values are admissible to represent, say, foams that have Poisson ratios between 0.1 and 0.4. It is henceforth understood that at least one of the colliding bodies is viscoelastic and possesses a time-dependent relaxation modulus of elasticity such that $E_i=E_i(t)$. This alone distinguishes the present work substantially from Zener's original analysis, because the time-dependent moduli of elasticity alter subsequently other parameters and lead to a governing wave equation of a modified form.

Zener [1] approached the acoustics problem by coupling two equations of motion (which are not repeated here). When a sphere strikes a thin, laterally large Kirchhoff–Love plate, the deformation is governed by a dispersive bending wave launched from the contact region. The resulting equation of motion (EOM) is then coupled with the EOM of a pinging sphere by the contact force, $F_s(t)$, acting between them. Accounting for time-dependent $E_i=E_i(t)$, a modified wave equation is now introduced:

$$\frac{d^2s(t)}{dt^2}+\frac{F_s(t)}{m}+\frac{d^2}{dt^2}\left(\alpha(t)\int_0^t F_s(t)dt\right)=0, \quad F_s(t)\in\mathbb{R}_{\geq 0} \tag{2}$$

In this equation, t is the time, m is the sphere mass, $s(t)$ is the interference coordinate,[1] $s(t)=z(t)-U(t)$, being the difference between the sphere center degree-of-freedom, $z(t)$, and the plate mid-plane deformation under the point of contact, $U(t)$. Following the Kirchhoff–Love thin plate theory, Zener relates the deformation, $U(t)$, to be proportional to the impulse, which is given by the integral of the contact force, $F_s(t)$, over the time in contact, i.e., $U(t)=\alpha(t)\int_0^t F_s(t)dt$. The wave equation (2) above is distinguished from Zener's original equation (13) in Ref. [1], because in the present work, the proportionality parameter, $\alpha=\alpha(t)$, is allowed to be time-dependent:

$$\alpha(t)=\frac{\sqrt{3\rho_b/E'_b(t)}}{16\rho_b h^2} \tag{3}$$

The density ρ_b in Eq. (3) is that of the plate, having a thickness, 2 h. Notably, for a solid elastic plate, $E'_b=$ conts. and $\alpha=$ const. Then, Eq. (2) reverts back to Eq. (13) in Ref. [1]. The upcoming numerical procedures solve all cases robustly and seamlessly.

It is emphasized that a positive $F_s(t)$ is the contact force acting upon the sphere in a direction that is opposite to the positive definition of $s(t)$ (having the same sense as $z(t)$). In other words, $F_s(t)>0$ is a compressive force between the contacting sphere and the plate, which also sets the validity range of Eq. (2). That equation is subject to initial conditions:

$$\begin{aligned} s(0)&=0\\ \dot{s}(0)&=v_o \end{aligned} \tag{4}$$

Zener then relates the point contact force and the interference by the well-known Hertzian[2] nonlinear relationship, $F_s(t)=ks(t)^{3/2}$, where $k=(4/3)r^{1/2}E'$. Notably, k itself is not a stiffness (nor does it have units of stiffness), but it may be regarded as "firmness." It shall be denoted below as the *kernel*, and because $E'=E'(t)$ is a function of time, so is the kernel, $k=k(t)$.

On the rebound, according to Zener [1], who assumes an elastic collision, separation occurs when the interference, $s(t)$, returns to zero. That condition is inappropriate for a viscoelastic collision where a different criterion is needed (which is already implied by the range of validity in Eq. (2)).

[1]Zener uses lowercase, s, for the interference, a symbol which is commonly reserved for the Laplace parameter. To distinguish between them, the interference in this work is always shown as a function of time, $s(t)$. The symbol s standing alone shall represent the Laplace parameter.

[2]Zener's model considered specifically a sphere impacting a flat plate. That should not be considered too restrictive, though. That model is straightforwardly extended to impact occurring between, say, two spheres. It is well-known that the Hertzian theory renders a composite radius, $r=r_1r_2/(r_1+r_2)$, where r_1 and r_2 are the two impacting spheres' radii. Any of the radii can be positive, infinity, or negative pertinent, respectively, to convex, flat, and concave geometries. Hence, the Hertzian theory leads to an equivalent sphere of a composite radius, r, pinging a flat surface having an equivalent elasticity modulus, $E'=E'(t)$, as given by Eq. (1).

3 A Viscoelastic Theoretical Background—A Formulation Preface

Sources by Mainardi [13] and Gutierres-Lemini [14] provide extensive coverage of linear viscoelasticity. Considerable treatment specific to time-dependent material properties, dynamics, and vibration particularly useful for this work are given in the author's work [15–21].

Viscoelastic materials are frequently called hereditary materials undergoing hereditary elasticity. The nonlinearity embedded in the Hertzian point contact model, $F_s(t) = k(t)(s(t))^{3/2}$, is caused by the nonlinear power of $s(t)$. That is a kinematical nonlinearity, not a material nonlinearity. Hence, Eq. (2) is nonlinear *even for linear viscoelastic materials*. Therefore, the elastic-viscoelastic correspondence principle *cannot* be used because it applies *only* to linear systems. That necessitates a numerical solution of Eq. (2). To further complicate things, it ought to be recognized that the proper definition of the force–displacement for viscoelastic materials must appear in the form of a convolution, denoted commonly by a "star" product:

$$F_s(t) = (k * u)(t) \quad \text{and} \quad u(t) = s(t)^{3/2} \tag{5}$$

Conveniently, a displacement function, $u(t) = s(t)^{3/2}$, is introduced to assist in additional mathematical operations. The appropriate condition of separation is when the contact force, $F_s(t)$, returns to zero succeeding the collision and upon the wave rebound. Unlike in Zener's work [1], contact is undone while $s(t)$ still "trails toward" zero[3] because of the delay embedded in the hereditary behavior of viscoelastic materials (mathematically it is embedded in the convolution above). And because of that delay, a residual deformation remains imprinted in the viscoelastic surface at the instant of separation.

Note that both the kernel, $k(t)$, and the response, $s(t)$, are functions of time. While $k(t)$ depends on the time-dependent material elasticity, $E'(t)$, the interference, $s(t)$, is obtained from the solution of the wave equation (2). Essentially, Eq. (5) expresses the Hertzian point contact model in a convolution form. It is well-known that a convolution can easily be calculated in the Laplace domain, followed by an inverse Laplace transform to extract the unknown, $s(t)$. While true for linear systems, it cannot be done herein because of the nonlinear term, $s(t)^{3/2}$, which thwarts the calculation of the convolution in the Laplace domain (that is because $\mathscr{L}(s(t)^{3/2}) \neq \mathscr{L}(s(t))^{3/2}$, where in fact $\mathscr{L}(s(t)^{3/2})$ is still a mathematical unknown). Hence, Eq. (2) must be evaluated in the time domain using a numerical procedure.

To establish the principles and to start off, an elementary form of Eq. (5) is that of Boltzmann's superposition integral for the stress–strain relationship and is given for simplicity, say, for uniaxial loading as

$$\sigma(t) = \begin{cases} E(t)\varepsilon(0) + \int_0^t \dot{\varepsilon}(\tau)E(t-\tau)d\tau \equiv E(t)\varepsilon(0) + \int_0^t \dot{\varepsilon}(\tau - t)E(\tau)d\tau & (\text{form } a) \\ E(0)\varepsilon(t) + \int_0^t \varepsilon(\tau)\dot{E}(t-\tau)d\tau \equiv E(0)\varepsilon(t) + \int_0^t \varepsilon(\tau - t)\dot{E}(\tau)d\tau & (\text{form } b) \end{cases} \tag{6}$$

The hereditary nature is immediately apparent—the time-dependent stress, $\sigma(t)$, at any time, t, is governed by a relaxation modulus, $E(t)$, convoluted with the strain, $\varepsilon(t)$, over all past time since $t = 0$. That is, the stress at any time, t, depends not just on the current state and rate of $E(t)$ and $\varepsilon(t)$, but on their values over all past history. Model functions that do not contain past history cannot represent hereditary viscoelasticity. It is pointed out that the two forms in Eq. (6) are mathematically equivalent rendering, of course, identical results for the stress either way. The choice of a form depends on either convenience or the availability of information regarding $E(t)$ or $\varepsilon(t)$. In that regard, one of the more advanced paradigms that captures relaxation robustly is through a fractional derivative (FC) model for the relaxation modulus of elasticity [13,15–22], using a Riemann–Liouville or a Caputo derivative (where both are actually defined by integral forms). But FC models commonly suffer from a singularity in the kernel rate at zero-time; hence, only *form a* (in Eq. (6)) is feasible for that material model.

The Laplace and Fourier transforms of Eq. (6) are

$$\begin{aligned} &\text{Laplace} \rightarrow \sigma(s) = sE(s)\varepsilon(s) \\ &\text{Frequency} \rightarrow \sigma(\omega) = (i\omega)E(\omega)\varepsilon(\omega) \end{aligned} \tag{7}$$

The terms $sE(s)$ and $E^*(\omega) = (i\omega)E(\omega)$ are known as the dynamic moduli. The real and imaginary parts of $E^*(\omega)$ are, respectively, the storage and loss moduli. The ratio $\tan\delta = \text{Im}(E^*(\omega))/\text{Re}(E^*(\omega))$ is the tangent loss, with a peak that lies somewhere within the transition between rubbery and glassy states.

For the uniaxial problem above, say a uniform rod, it is apparent that multiplying Eq. (6) by the rod cross-sectional area, A, and replacing $\varepsilon(t) = u(t)/L$, where $u(t)$ is a time-dependent displacement (typically, related to the unknown degree-of-freedom) and L is the rod length, then by way of $k(t) = E(t)A/L$, Eqs. (6) and (7) become:

$$F_s(t) = \begin{cases} k(t)u(0) + \int_0^t \dot{u}(\tau)k(t-\tau)d\tau \equiv k(t)u(0) + \int_0^t \dot{u}(t-\tau)k(\tau)d\tau & (\text{form } a) \\ k(0)u(t) + \int_0^t u(\tau)\dot{k}(t-\tau)d\tau \equiv k(0)u(t) + \int_0^t u(t-\tau)\dot{k}(\tau)d\tau & (\text{form } b) \end{cases} \tag{8}$$

and

$$\begin{aligned} &\text{Laplace} \rightarrow F_s(s) = sK(s)u(s) \\ &\text{Frequency} \rightarrow F_s(\omega) = (i\omega)K(\omega)u(\omega) \end{aligned} \tag{9}$$

The Volterra series [23–27] is a widely used analytical tool to capture the nonlinear hereditary behavior effects of viscoelastic materials over all time. Generally, it is expressed as

$$F_s(t) = h_0(t) + \int h_1(\tau)u(t-\tau)d\tau + \iint h_2(\tau_1, \tau_2)u(t-\tau_1)u(t-\tau_2)d\tau_1 d\tau_2 + \cdots \tag{10}$$

[3] In fact, as long as contact is in effect and while Eq. (2) governs, the interference, $s(t)$, may never return to or cross zero prior to ejection.

In Eq. (10), the force is $F_s(t)$, the kernels are $h_1(t) \equiv k(t)$ and $h_2(t_1, t_2) \equiv k(t_1 + t_2)$, and $u(t)$ belongs to a generic time signal, in our case $u(t) = s(t)^{3/2}$, and $du(t)/dt = (3/2)s(t)^{1/2} ds(t)/dt$.

Only two integral terms are shown in Eq. (10), where higher order terms (integrals) are implied. The leading term $h_0(t)$ in Eq. (10) has the same role of the "elastic" leading terms in Eq. (8). The first integral represents the Boltzmann's superposition principle; hence, the first-order kernel, $h_1(t)$, is identical to the kernel, $k(t)$, in Eq. (8). The successive integrals are intended to capture additional nonlinear effects, as the second-order kernel $h_2(t_1, t_2)$ is used to add correspondingly a quadratic nonlinear contribution, etc. For all the cases investigated in this work, it is found that the second nonlinear term is already superfluous because it adds an utterly insignificant value compared to the first integral term, i.e., the Boltzmann superposition principle of Eq. (8) is sufficient. From a practical point of view, the Volterra second term overburdens the computational efforts tremendously (just the second Volterra term alone consumes over 95% of the CPU time slowing computation considerably, possibly needlessly, where each higher order integral term would escalate that even more so). In summary, the Volterra series formulation is incorporated here to ensure that additional nonlinearities are not overlooked, but it is safe to bypass if chosen. In the forthcoming direct numerical integration procedure, the Volterra series is intrinsically built in (and can be activated simply by an input switch), but it will not be explicitly expressed any further because mathematical equations can become staggeringly involved. Of course, as a matter of caution, final results ought always to be judiciously verified.

The challenging part, though, is that regardless of whether Eqs. (8) and (10) are substituted into Eq. (2), that equation turns into a nonlinear ordinary *integro-differential* equation (OIDE). Choosing, for example, *form a* of Eq. (8) and applying the homogeneous initial conditions gives

$$\frac{d^2 s(t)}{dt^2} + \frac{3}{2m}\left(\int_0^t \sqrt{s(\tau)}\dot{s}(\tau)\, k(t-\tau) d\tau\right) + \frac{3}{2}\frac{d^2}{dt^2}\left[\alpha(t)\left(\int_0^t \sqrt{s(\tau)}\dot{s}(\tau)\, k(t-\tau) d\tau\right)\right] = 0 \tag{11}$$

Similar forms of OIDEs would appear to accommodate *form b* of Eq. (8); nevertheless, all forms need to be solved by a numerical time marching integration scheme.

To demonstrate the theory, materials and geometry need to be chosen. As indicated above, subscripts a and b represent the sphere and the plate, respectively. Four common cases are possible for Eq. (1), noting that

$$E'_{a/b}(t) = \frac{E_{a/b}(t)}{1 - \nu^2_{a/b}} \tag{12}$$

Emphasizing that any modulus of elasticity might assume a time-dependent relationship, $E_{a/b}(t)$.

- Case 1—The sphere and the plate are both solids with time-unvarying E_a and E_b: instance (i) is restricted to pure elastic deformations as completed by Zener [1], instance (ii) includes elastoplastic deformation and completed by Green [2] who fused in the JGM [3] elastoplastic model. Case 1 is complete and is not pursued further herein.
- Case 2—Impact occurs between two similar (or identical) viscoelastic materials. Cases like that appear commonly in granular flows, biomedical cells (red blood cells) in plasma flows, and cartilage-on-cartilage contact (being further complicated by porosity and the synovial fluid, see Refs. [16–20]). For this case, $E_a(t) \approx E_b(t)$ and $\rho_a \approx \rho_b$, so that from Eq. (1) we have, $E'(t) = E'_{a/b}(t)/2 = E_{a/b}(t)/(2(1 - \nu^2_{a/b}))$.
- Case 3—A "heavy" elastic (say, metallic or glass) ball impacts a viscoelastic plate or substrate such as in mats or subfloors. Here, $E_a \gg E_b$, where $E_a = \text{const.}$, but the plate is viscoelastic, $E_b = E_b(t)$. Then, from Eq. (1), we have $E'(t) \approx E'_b(t) = E_b(t)/(1 - \nu_b^2)$.
- Case 4—This is a material reversal of case 3—the sphere is viscoelastic and the plate is a solid ("hard") surface having $E_b = \text{const.}$ For which $E_a \ll E_b$ and $E_a = E_a(t)$, so that $E'(t) \approx E'_a(t) = E_a(t)/(1 - \nu_a^2)$. This case is likely to represent impacts in sports such as tennis, racquetball, squash, etc. (of course, for hollow balls, a change in internal air pressure during impact may necessitate further thermodynamic considerations).

It is obvious that in principle the last three cases are all quite similar mathematically. If none of these represent a specific case of interest, the original definition of $E' = E(t)$ as given by Eq. (1) ought to be used. So, regardless of which case above is chosen, the forthcoming methodology is applicable to all cases (including case 1, as it will be seen). Case 3 is subjectively chosen to further the numerical procedures. Also, to reduce the number of parameters involved, it is further assumed that the sphere diameter equals half of the plate thickness, i.e., $2r \approx h$.

4 The Viscoelastic Material Characterization

Two numerical solution methods are presented below: (1) a direct ("brute force") method that solves straightforwardly the nonlinear OIDE, Eq. (11), and (2) the IVE method being an exceptionally efficient method that solves Eq. (2) using supplementary variables to bypass the need for the OIDE—the theory is extended further below. For both methods, it is important that the material representation has a time domain kernel representation. Depending on the formulation (as in Eq. (8)), it may also be required that the kernel has a nonsingular time derivative. Most viscoelastic materials have such representations, but the fractional derivative (FC) models (a model of which is used herein) have a singularity in the time derivative at $t = 0$, and it will be appropriately handled.

To set off the solution scheme, a rather basic SLS (known also as the Zener material model) is chosen (see Eq. (2.73) in Ref. [13] that gives only a Laplace representation). The time, Laplace, and frequency representations are

$$E(t) = E_g\left[\frac{\alpha}{\beta} + e^{-\beta t}\left(1 - \frac{\alpha}{\beta}\right)\right]$$
$$\underbrace{sE(s) = E_g \frac{\alpha + s}{\beta + s}|_{s=i\omega}}_{\text{Laplace}} = \underbrace{i\omega E(i\omega) \triangleq E^*(\omega) = E_g \frac{\alpha + i\omega}{\beta + i\omega}}_{\text{Frequency}} \tag{13}$$

Note that the time representation, $E(t)$, in Eq. (13) is backtracked from an inverse Laplace transform of the Laplace representation, $E(s)$, in Eq. (13). Here, E_g is the glassy modulus, while $E_r = E_g(\alpha/\beta)$ is the rubbery modulus. Maintaining a ratio, $\alpha/\beta = 1/4$ as in Ref. [13], the specific values here are $\alpha = 300$ 1/s and $\beta = 1200$ 1/s, leading to

$$E(t) = E_g\left[\frac{1}{4} + \left(\frac{3}{4}\right)e^{-1200t}\right] \tag{14}$$

A glassy modulus value of $E_g = 15$ MPa is picked as typical for an elastomer of a Shore A hardness of 90 durometer. Both Eqs. (13) and (14) will serve in the following (recall though that moduli still need to be corrected for Poisson's ratio by $(1 - \nu^2)^{-1}$ according to Eq. (1)). Equations (13) and (14) are shown in Fig. 1 for the parameters chosen (t_0 and $\omega_{\min}$ will be discussed later).

5 The Direct Formulation

Recall that case 3 above is the subject of this investigation. Recall also that the kernel is $k(t) = (4/3)r^{1/2}E'(t)$. The sphere radius is 10 mm, it is made of steel, $\rho_a = 7800$ kg/m^3, so that its mass is 32.67 g. The sphere is assumed to impact with $v_o = 10$ m/s (a parameter that will be varied later). The density of the viscoelastic

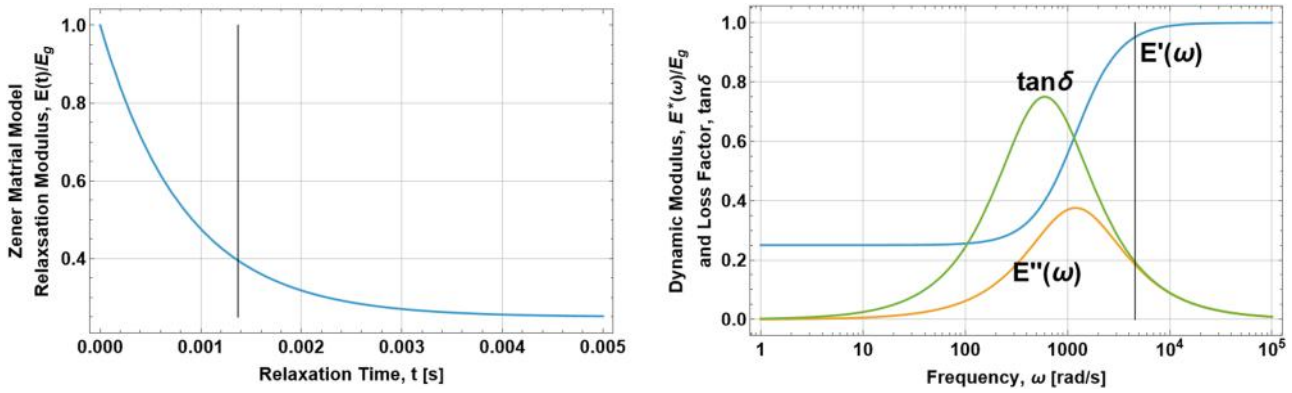

Fig. 1 Time and frequency domain representations for a standard linear solid (Zener material) with $\alpha = 300$ 1/s and $\beta = 1200$ 1/s; $t_0 = 1.34$ ms, and $\omega_{min} = 4684$ rad/s are shown by vertical lines, where the pertinent regions of relevance are $t \le t_0$ and $\omega \ge \omega_{min}$

material is taken as typical, $\rho_b = 1100$ kg/m^3. Note also that the direct formulation is coded in two programming languages: (1) Mathematica (being a modern and convenient platform that contains many useful built-in functions, e.g., ListConvolve[] and NDSolve[]) and (2) Fortran (a tried-and-true robust procedural language that when compiled is by leaps and bounds faster than interpreted languages such as Mathematica or MATLAB; however, many details and functions have to be hand-coded). The reason for using a compiled language such as Fortran is because of the nonlinear second Volterra double integral that is overly CPU-intensive, especially as time "marches forward." A Fortran compiled code executes in a matter of seconds (rarely minutes) even for the most demanding condition (10–60 s is typical, depending on the magnitude of the time increment). A Mathematica interpreted code for the same parameters may run for many minutes to hours even on a very fast and modern workstation (AMD Ryzen 9, 32GB of RAM, 32 logical processors, running at 4.5 GHz). If the Volterra series is bypassed, the compiled direct formulation code executes in a fraction of a second (even less than 0.1 s).

The direct formulation consists of the following steps: (1) setup data and equations for the geometry and material properties, (2) impose the initial conditions, (3) initiate a self-starting time integration loop by incrementing time from zero to some T_{max} (typically it is the time where the rubbery behavior levels; for the examples showcased here, T_{max} ranges in about 2–5 ms) with a small time marching increment δt (an order of 1 μs is typical, so there are approximately a few thousands time-steps), (4) calculate the convolutions (i.e., the viscoelastic force, $F_s(t)$) using summations in Fortran or the built-in function ListConvolve[] in Mathematica, (5) calculate the second time-derivatives of the impulse of $F_s(t)$ by a backward difference formula of order, $O(\delta t^5)$, (6) isolate the acceleration from the OIDE, Eq. (11), and integrate twice by explicit Adams-Bashforth time marching and Runge–Kutta self-starting formulations, both of order, $O(\delta t^5)$, (7) once the time marching loop is done, postprocess the result, with particular interest of the outcomes, $s(t)$, $\dot{s}(t)$, and $F_s(t)$.

The following equations provide complementary kernel details for convolution calculations (including those in the Volterra series), along with the signals $u(t)$ and $\dot{u}(t)$, as they are applied in the two forms of Eq. (8) or (10). Starting with *form a*, note that by Eq. (4), $s(0) = 0$:

$$\begin{aligned} \dot{u}(t) &= \frac{3}{2} s(t)^{1/2} \times \dot{s}(t); \quad u(0) = 0; \quad \dot{u}(0) = 0 \\ h_1(t) &= k(t) = 2.667 \times 10^6 \left(\frac{1}{4} + \frac{3e^{-1200t}}{4} \right) \\ h_2(t_1, t_2) &= k(t_1 + t_2) = 2.667 \times 10^6 \left(\frac{1}{4} + \frac{3}{4} e^{-1200(t_1+t_2)} \right) \\ h_0(t) &= k(t) \cdot u(0) = 0 \end{aligned} \tag{15}$$

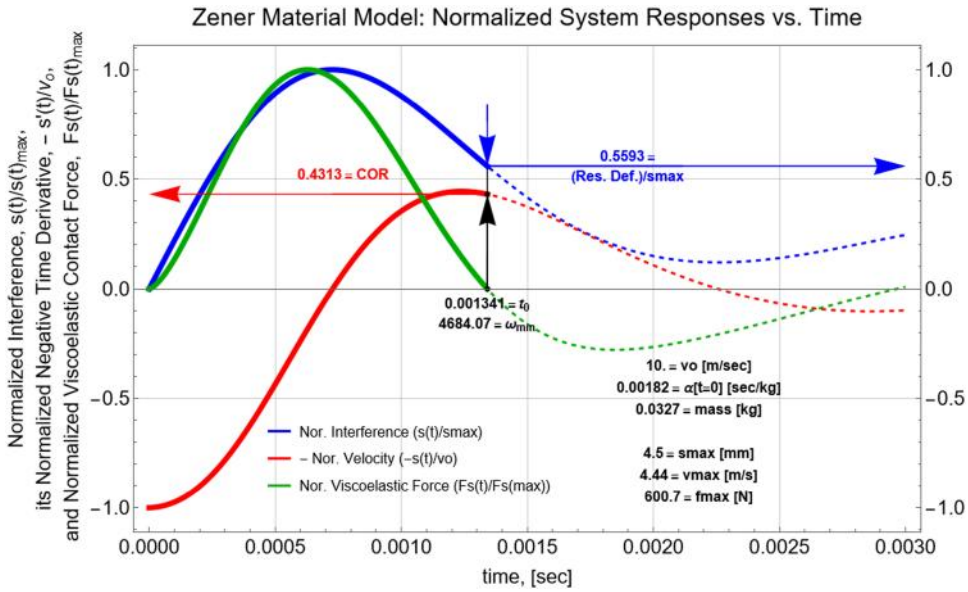

Fig. 2 Normalized system responses versus time obtained by the direct time integration, and the IVE formulations for the Zener material (SLS) model

and then for *form b*:

$$\begin{aligned} &u(t) = s(t)^{3/2}; \quad u(0) = 0 \\ &h_1(t) = \dot{k}(t) = -2.4 \times 10^9 \times e^{-1200t} \\ &h_2(t_1, t_2) = \dot{k}(t_1 + t_2) = -2.4 \times 10^9 \times e^{-1200(t_1+t_2)} \\ &h_0(t) = k(0) \cdot u(t) = 2.667 \times 10^6 \times u(t) \end{aligned} \tag{16}$$

It is indeed verified that it does not matter which form is used for the Zener material; both forms produce identical results regardless of the value of δt chosen (as long as it is sufficiently small). It is also verified that the direct ("brute-force") integration method produces identical results for the final outcomes, whether programed in Mathematica or in Fortran (while using their distinct intrinsic functions). Another convergence test is done by superficially "transforming" the Zener viscoelastic material (as specified by Eqs. (15) and (16)) into an elastic material by merely nullifying the decaying parameter (i.e., plummeting the exponent value of 1200 1/s to zero), executing the direct method codes, and comparing the results with a known solution for elastic collision. The values for COR = 0.79, and the instant of ejection $t_0 = 1.4$ ms, duly match those predicted by Eqs. (13) and (13*b*) in Ref. [2] for an "inelasticity" parameter, $\lambda = 0.138$. Lastly, the most decisive convergence benchmark is that the results of the direct method match perfectly those of the *upcoming* solution method of the IVE—the two numerical procedures are completely distinct, they do not share anything other than, of course, the same physics.

For the parameters stated above, the results of the simulation are shown in Fig. 2. The sequence of arrows points to values of the final results. Starting with the instant when the contact force $F_s(t)$ returns to zero—that is designated by t_0 to signify the instant of ejection. Then, at that instant, the values of the normalized velocity and the normalized residual deformation are pinpointed.

Figure 2 is now discussed in detail. For all three responses to be shown on the same plot, they are normalized as follows: $s(t)/s(t)_{\max}$, $-\dot{s}(t)/v_o$, and $F_s(t)/F_s(t)_{\max}$. First note that $\dot{s}(t)/v_o$ is shown in a negative sense because the rebound takes place in the opposite direction of the initial impact. By definition, the relative speed ratio at separation onset yields the COR. Initially, the force, $F_s(t)$, is quite in phase with the interference, $s(t)$, where both increase from zero initial conditions to their maxima. But then, as time progresses further, the plate inertial effects and the hysteretic energy loss (shown in Fig. 3) prompt the viscous contact force, $F_s(t)$, to drop rather quickly until it reaches the value of zero that signifies separation—at that instant $t_0 =$ 1.34 ms, separation or ejection occurs, rendering thus the COR value, $|\dot{s}(t)/v_o| \equiv \text{COR} = 0.4313$. Numerically, the instant t_0 is determined by a root-finding routine that acts upon a Lagrange interpolated function of the force values of $F_s(t)$ at the first occurrence of a negative value. At that instant of t_0, it is apparent that there is a residual depression left in the viscoelastic surface, having a magnitude of $0.5593 \times 4.5 = 2.517$ mm (which will dimmish, of course, in due time after ejection). So, the pertinent regions of relevance are $t \le t_0$ and $\omega \ge \omega_{\min}$, where $t_0 = 1.34$ ms and $\omega_{\min} = 4684$ rad/s are also shown by the vertical lines in Fig. 1.

Past the instant of separation at t_0, "pseudo" or "ghost" responses still appear in Fig. 2 in dashed lines, but they are irrelevant to the stated problem of restitution. That is because a negative force means that the sphere "pulls" the plate surface in a direction opposite to the impact. Because no retention mechanisms (e.g., adhesion) are present, the sphere had already departed from the plate, where Eq. (2) or (11) cease to govern the problem at the instant of departure, t_0. The novelty in the "ghost" lines is that they would describe a different case where a "strong" adhesive or glue is spread on the plate surface prior to impact, and once the sphere hits, it adheres to the plate being unable to separate form it (at least for some time); then, some sort of an oscillatory motion ensues. Perhaps classical adhesion models (e.g., the Lennard–Jones potential) ought to be considered to supplement Eq. (5). While that might be another interesting topic to pursue, it is beyond the objectives and scope of this work.

A parametric plot of the viscoelastic force, $F_s(t)$, versus the deformation, $s(t)$, reveals the hysteretic loop as shown in Fig. 3. The area captured inside the loop equals the hysteretic energy loss, and it is calculated by

$$W_h = \int F_s(t) ds(t) = \int_0^{t_0} F_s(t) \frac{ds(t)}{dt} dt$$

Because the outcomes $F_s(t)$ and $ds(t)/dt$ (e.g., as shown in Fig. 2) are available in equidistantly time-spaced arrays, the second quadrature form is more convenient and is easily calculated by Simpson's rule to be 0.793 Joule for this case of the SLS-Zener material.

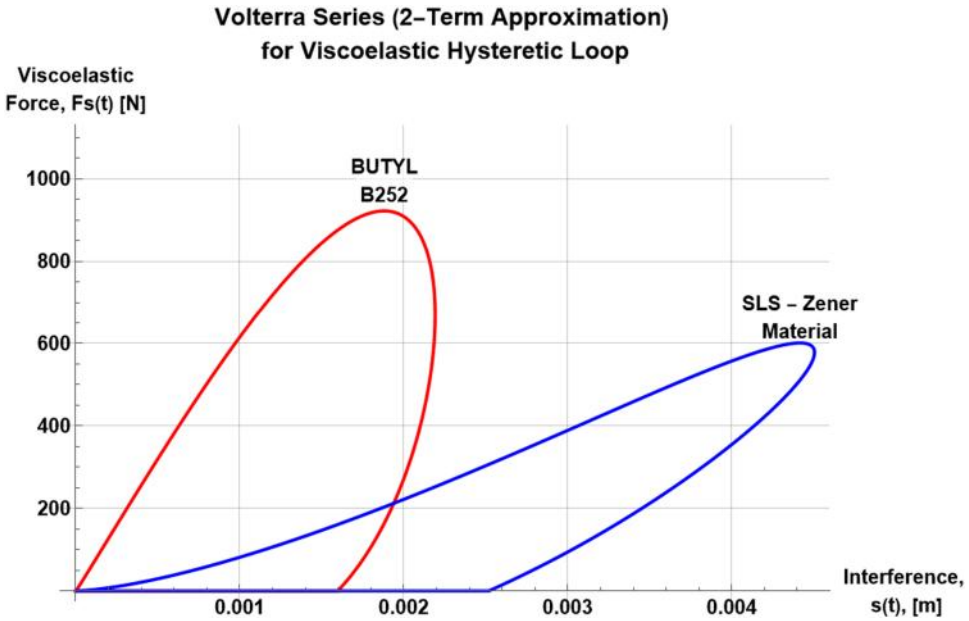

Fig. 3 Hysteretic loops for Zener material ($F_s(t)_{\max} = 600.7$ N, $s(t)_{\max} = 4.5$ mm, and energy loss = 0.793 J) and BUTYL-B252 ($F_s(t)_{\max} = 921.7$ N, $s(t)_{\max} = 2.19$ mm, and energy loss = 1.172 J)

6 The Internal Variable Evolution Formulation

Offering yet another formulation serves a few purposes: (1) the forthcoming formulation is mathematically robust—there are no approximations of any kind, and (2) it circumvents the need for dealing with the aforementioned ordinary integro-differential wave equation (OIDE) as in Eq. (11). The formulation is recasts into *a set* of first-order ordinary differential equations (ODEs) that are rapidly solved by the internal Mathematica function NDSolve[] with ease, (3) execution is practically instantaneous, and is most efficient to conduct a parametric investigation, and (4) as a minimum, it offers an additional verification to the outcomes of the direct method that solves the OIDE as discussed above. The formulation is general and it can be amended to handle a variety of constitutive models in addition to the Prony-based Maxwell models. The mathematical development below, however, targets the generalized Maxwell model because of its simplicity (that model is perhaps the most commonly used). It ought to be recognized that the aforementioned Zener material model is in essence an elementary or a special case of the generalized Maxwell model, which is expressed as a Prony series:

$$k(t) = G_\infty + \sum_{k=1}^{N} G_k e^{-t/\tau_k} \tag{17}$$

The parameter G is used here generically (it is not necessarily a "shear modulus"). So, correspondingly, Zener's material parameters are $N=1$, $G_\infty = 0.25k(0)$, $G_1 = 0.75k(0)$, and $k(0) = G_\infty + G_1 = 2.667 \times 10^6$ is the glassy value (complying with Eq. (15), and $\tau_1 = 1/\beta = 1/1200$ s is the relaxation time). Returning now to any N (a value of $N=3$ will be used later), the state of each Maxwell element, k ($k=1, \ldots, N$), is described by an internal variable, $Q_k(t)$, and is governed by a first-order differential equation (see the equivalence in Ref. [13], Eq. (2.19*a*)):

$$\begin{aligned} \dot{Q}_k + \frac{Q_k}{\tau_k} &= G_k \frac{du(t)}{dt}, \qquad k = 1, \ldots, N \\ &= \frac{3}{2} G_k s(t)^{1/2} \dot{s}(t) \end{aligned} \tag{18}$$

This is a nonhomogeneous linear ODE of the general form, $\dot{y} + ay = b(t)$, which is routinely solved by an integrating factor, $\exp\left(\int a dt\right) = \exp\left(\int \frac{1}{\tau_k} dt\right) = \exp(t/\tau_k)$. Assuming a pristine past, i.e., $Q_k(t) = 0$ and $du(t)/dt = 0$ for any $t \le 0$, the solution of Eq. (18) for $Q_k(t)$ is straightforwardly obtained from the integrating factor by

$$Q_k(t) = \int_0^t G_k \frac{du(t)}{dt} e^{-(t-\xi)/\tau_k} d\xi, \quad k = 1, \ldots, N \tag{19}$$

This equation already has the mathematical form of a convolution. The viscoelastic force is then

$$F_s(t) = G_\infty u(t) + \sum_{k=1}^{N} Q_k(t) \tag{20}$$

Substituting Eq. (19) into (20) gives

$$F_s(t) = G_\infty u(t) + \sum_{k=1}^{N} \int_0^t (G_k e^{-(t-\xi)/\tau_k}) \frac{du(t)}{dt} d\xi \tag{21}$$

Equation (21) is the standard Boltzmann superposition principle, where the term in parentheses is part of the time-dependent relaxation kernel (i.e., the "memory" or hereditary function in Eq. (18)). As the time-response signal is $u(t) = s(t)^{3/2}$, the integral and summation collect all past "history" for $F_s(t)$.

In summary, the governing ODE, Eq. (2), is now coupled with N ODEs of Eq. (18), all solved for the objective unknown, $s(t)$, along with the auxiliary internal variables $Q_k(t)$; hence, combining the set of $N+1$ ODEs gives

$$\begin{aligned} &\frac{d^2 s(t)}{dt^2} + \frac{F_s(t)}{m} + \frac{d^2\left(\alpha(t)\int_0^t F_s(t)dt\right)}{dt^2} = 0 \\ &\dot{Q}_k + \frac{Q_k}{\tau_k} = \frac{3}{2} G_k s(t)^{1/2} \dot{s}(t) \quad \text{for } k = 1, \ldots, N \end{aligned} \tag{22a}$$

Now, Eq. (20) is substituted in the governing EOM (i.e., the first equation in Eq. (22)), along with $u(t) = s(t)^{3/2}$ to give finally

$$\begin{aligned} &\frac{d^2 s(t)}{dt^2} + \frac{G_\infty s(t)^{3/2} + \sum_{k=1}^{N} Q_k(t)}{m} \\ &\quad + \frac{d^2\left[\alpha(t)\int_0^t \left(G_\infty s(t)^{3/2} + \sum_{k=1}^{N} Q_k(t)\right) dt\right]}{dt^2} = 0 \\ &\dot{Q}_k + \frac{Q_k}{\tau_k} = \frac{3}{2} G_k s(t)^{1/2} \dot{s}(t) \quad \text{for } k = 1, \ldots, N \end{aligned} \tag{22b}$$

The set of $N+1$ equations is solved simultaneously, very efficiently, and remarkably fast (in a *fraction* of a second) by Mathematica's function, NDSolve[]. While it is now appropriate to show the results of this formulation, it is essentially unnecessary because the outcomes shown in Figs. 2 and 3 accurately represent the current formulation just as well.

The IVE formulation is now used for the upcoming parametric investigation. There is, of course, a great number of parameters in this problem along with various combinations that can be investigated in a parametric study. The most compelling parameter is the impact velocity, v_o, which is now being varied in the range 1–30 m/s in steps of 1 m/s. Figure 4 shows that the COR increases with v_o and decreases with a faster decay β. While the latter is expected, the former is quite unusual (compared to, say, an elasto-plastic COR case as it was investigated by Jackson et al. [3] and Green [2]). The reason for that behavior is embedded in Figs. 1 and 2, noting that $t_o = 1.34$ ms and $\omega_{\min} = 2\pi/t_0 = 4684$ rad/s. Only higher frequencies are relevant for the problem (because only shorter times than t_0 are involved). As it is evident from Fig. 1, $\omega_{\min}$ lies far away past the transition from rubbery to glassy states, where $\omega_{\min}$ itself tends to also increase with the impact velocity. So, noting that the storage modulus (i.e., "elasticity") is nearly at its maximum and still increasing with frequency past $\omega_{\min}$, the loss modulus and $\tan\delta$ (where both imply "damping") are on a steep decline. The outcome is a COR that increases with v_o to a level, as shown in Fig. 4. With the current parameters of the SLS (Zener) material, the problem clearly lies in the "glassy" state.

7 A Real Material—Butyl-B252

It ought to be recognized that an SLS expressed by a single Prony term can rarely describe real viscoelastic material dynamics because usually it is unable to robustly capture relaxation and frequency intrinsic that span over frequency decades of specific interest. The SLS model is used above for its simplicity, and mainly to establish the principles needed to meet the objective of the stated problem in this work, i.e., the COR computation, but there would be little practical use for such a material model. So, a real rubber material is now chosen, the BUTYL-B252 as documented by Bagley and Torvik [28]. They used a fractional calculus (FC) model with a fractional derivative of order $\alpha = 1/2$ that fits the actual data exceptionally well in three frequency decades of their interest, 10^2–10^5 Hz:

$$\begin{aligned} &\sigma(t) = E_0 \varepsilon(t) + E_1 D^\alpha[\varepsilon(t)] \quad \text{and} \quad \alpha = 1/2 \\ &E_0 = 3 \times (7.6 \times 10^5) \text{ Pa}; \; E_1 = 3 \times (2.95 \times 10^5) \text{ Pa} \end{aligned} \tag{23a}$$

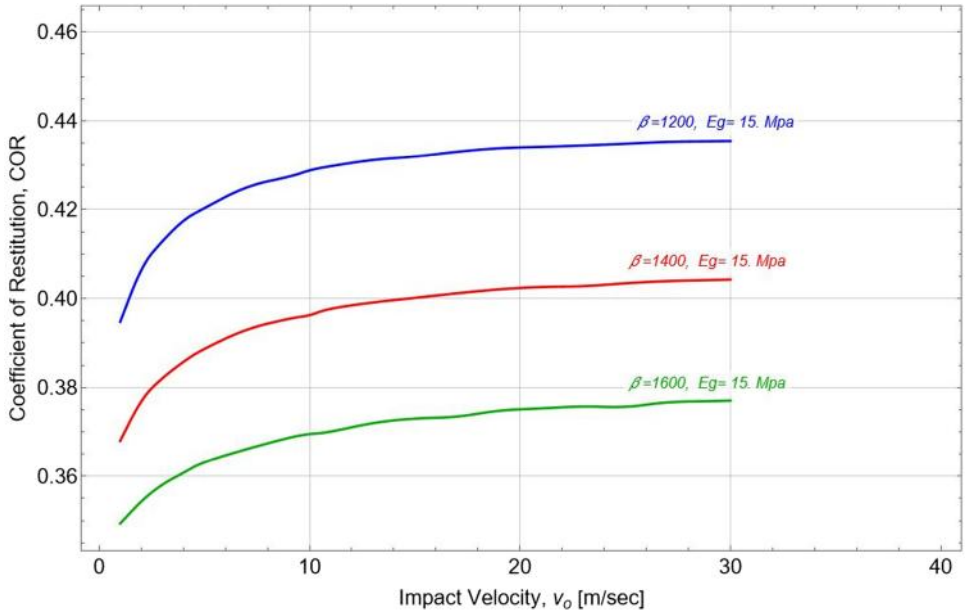

Fig. 4 The COR as a function of impact velocity, with β as a parameter

Because the fractional half-derivative is a convolution of a singular kernel, $E_{1/2}(t) = 1/\sqrt{\pi t}$, with the strain, $\varepsilon(t)$, a matching power-law in time representation is achieved for the elasticity modulus (see Refs. [15,29]):

$$E(t) = 3 \times 10^5 \left(7.6 + \frac{2.95}{\sqrt{\pi t}}\right) \tag{23b}$$

Now, the BUTYL-B252 actual data are wholly delegated to this equation. Modifying the denominator to $\sqrt{\pi \cdot \text{Max}(\delta t, t)}$, where δt is incrementally small, removes the said singularity and allows the Boltzmann superposition convolution integral to numerically trace back to $t = 0$ or $t - \tau = 0$. While $E(t)$ forms all the kernels in Eqs. (15) and (16), they all multiply homogeneous initial conditions in Eq. (8); hence, that denominator modification has utterly nought consequences. Conveniently, δt is selected here to equal the time marching increment used in the direct integration method of the OIDE. The forthcoming ejection time, t_0, is indeed shown to reside well within the time range of 10^{-5}–10^{-2} s that corresponds to the aforementioned frequency range.

Another effective bypass of the singularity is to use the Mittag–Leffler function, $E_{\alpha=1/2}(-x) = \exp(x^2)\,\text{erfc}(x)$. This functional form naturally appears as the solution to fractional differential equations of order $\alpha = 1/2$, which is specifically used by Bagley and Torvik [28] to describe their data (see Fig. 5). A nonlinear curve-fit to Eq. (23*b*) lends the following parameters:

$$\begin{gathered} E(t) = E_0 + E_1 e^{\beta_1^2 t}\text{erfc}\left(\beta_1\sqrt{t}\right) \triangleq E_0 + E_1 \cdot \text{CERF}\left(\beta_1\sqrt{t}\right) \\ E_0 = 0.442\,\text{MPa}; \quad E_1 = 0.6765\,\text{GPa}; \quad \beta_1 = 700\,1/\text{s} \end{gathered} \tag{24}$$

Letting $x = \beta_1\sqrt{t}$, the term in Eq. (24), $e^{\beta_1^2 t}\text{erfc}\left(\beta_1\sqrt{t}\right)$, is tagged as the $\text{CERF}\left(\beta_1\sqrt{t}\right)$ function following Refs. [15,29,30]. CERF is devised to avert numerical instabilities in computation. An updated version is reformulated in the Appendix and used throughout this work. Note also that only one CERF term appears in Eq. (24), and that is because FC models are generally robust, where a single term is often sufficient to represent $E(t)$ over several decades of frequency. Similar to Eqs. (13) and (14), the Laplace and Fourier representations for Eq. (24) are (see Refs. [15–21,29])

$$sE(s) = E_0 + \left.\frac{E_1\sqrt{s}}{\sqrt{s} + \beta_1}\right|_{s=i\omega} \triangleq E^*(\omega) = E_0 + \frac{E_1\sqrt{i\omega}}{\sqrt{i\omega} + \beta_1} \tag{25}$$

The glassy modulus is obtained by letting s approach infinity in Eq. (25), giving $Eg = 676.5$ MPa.

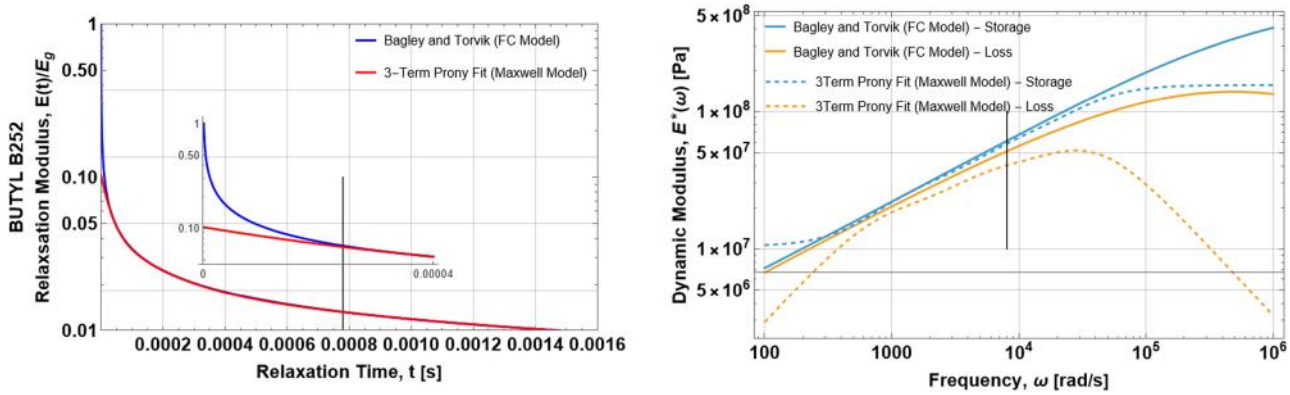

Fig. 5 Time and frequency domain moduli for BUTYL-B252; $t_0 = 0.78$ ms and $\omega_{\text{min}} = 8102$ rad/s are shown by vertical lines to indicate limits on regions of present relevance: $t \le t_0$ and $\omega \ge \omega_{\text{min}}$

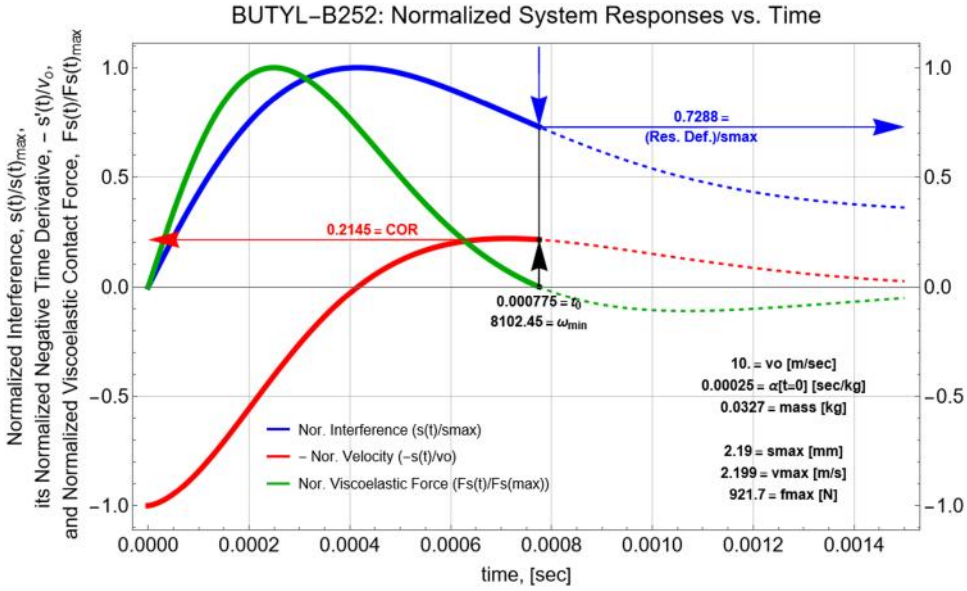

Fig. 6 System responses versus time for BUTYL-B252 obtained by the direct time integration (OIDE) for the actual data representation of Eqs. (23b) and (24)

Another widespread alternative to represent the Bagley and Torvik data is to fit a Prony series Maxwell model. It will be shown, however, that representation is inferior to the CERF model; yet, it is still used here because it is the most commonly used material model, and the COR problem can be solved expeditiously by the IVE method developed above. However, some important physical differences between the material models will surface. So, the three-term Prony series as fitted by Szumski [29] is used (see Eq. (113) there):

$$E(t) = E_0 + \sum_{i=1}^{N=3} E_i e^{-\beta_i t}$$
$$\{E_0, E_1, E_2, E_3\} = \{10.421, 20.181, 38.204, 87.344\}\ \text{MPa}$$
$$\{\beta_1, \beta_2, \beta_3\} = \{972, 6426, 34221\}\ \text{s}^{-1} \tag{26}$$

The corresponding Laplace and Fourier transform representations are (see Refs. [15–21,29])

$$sE(s) = E_0 + s\sum_{i=1}^{N=3} \frac{E_i}{s+\beta_i}$$
$$E^*(\omega) = E_0 + i\omega \sum_{i=1}^{N=3} \frac{E_i}{i\omega+\beta_i} \tag{27}$$

Repeating the estimation of the glassy modulus for the three-term Prony gives $Eg = 156.15$ MPa. The relaxation and dynamic moduli are depicted in Fig. 5 for both models, FC-CERF and Prony series. The following is observed:

(1) While at "long" times the three-term Prony series relaxation modulus follows the Bagley and Torvik data closely [22] (referred to as FC-CERF), it simply cannot match it well

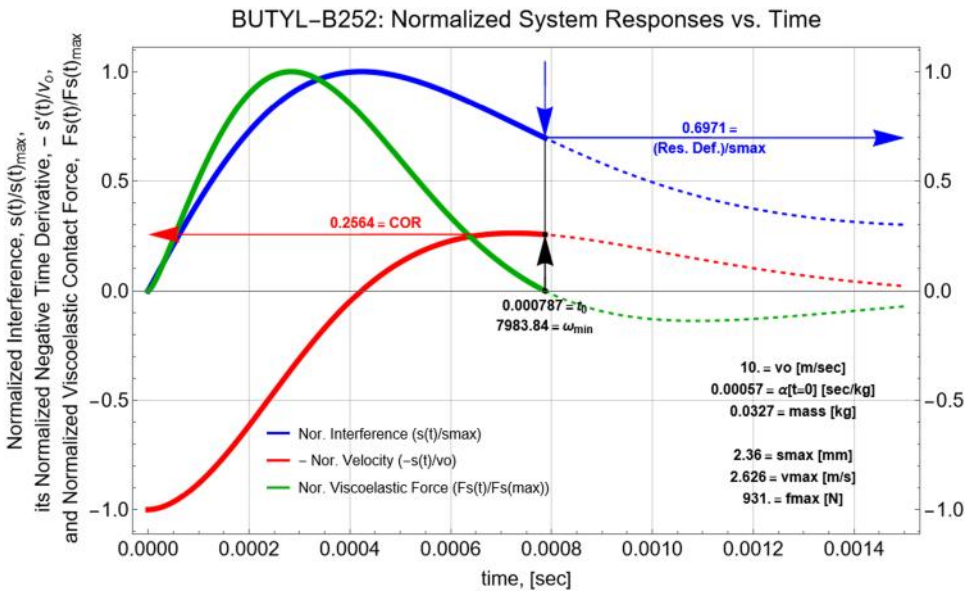

Fig. 7 System responses versus time for BUTYL-B252 obtained by the direct time integration (OIDE) for the Maxwell three-term Prony model, Eq. (26)

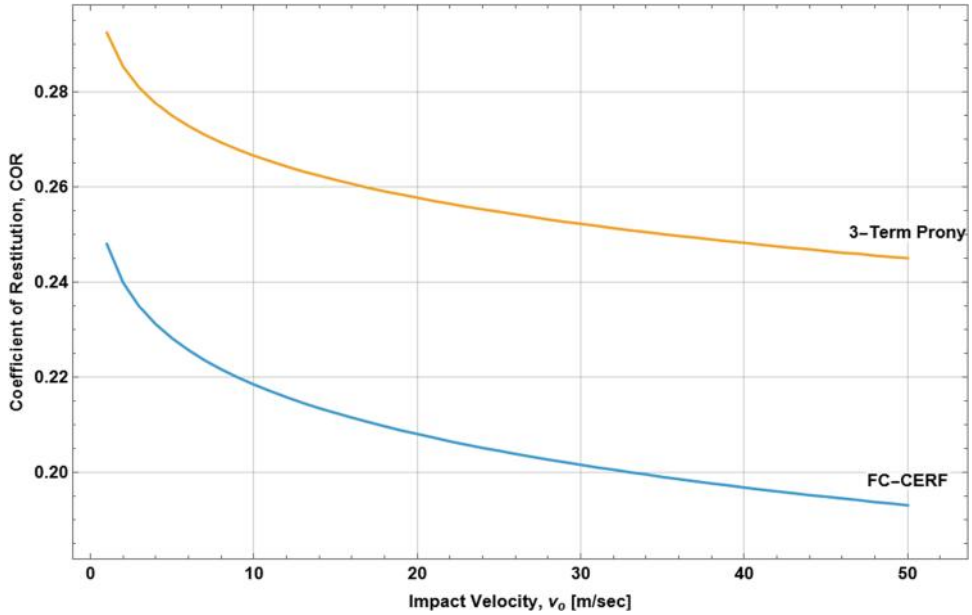

Fig. 8 Comparison of the two material models for BUTYL-B252—the COR values as a function of impact velocity

at "smaller" times (the inset in Fig. 5 reveals that in a time zoom-in, $t \leq 0.4$ ms). The mismatch between the glassy moduli of the Prony and the FC-CERF models is apparent. Such a mismatch is common when Maxwell Prony series models are attempted to match FC models. In which case, more Prony terms would be required in an arduous numerical procedure (if at all successful) that is likely to confront ill-conditioned matrices during the curve-fitting process.

(2) Also in the frequency domain, the three-term Prony series fits the FC-CERF model reasonably well within a limited frequency range (800–8000 rad/s), but it falls short in the

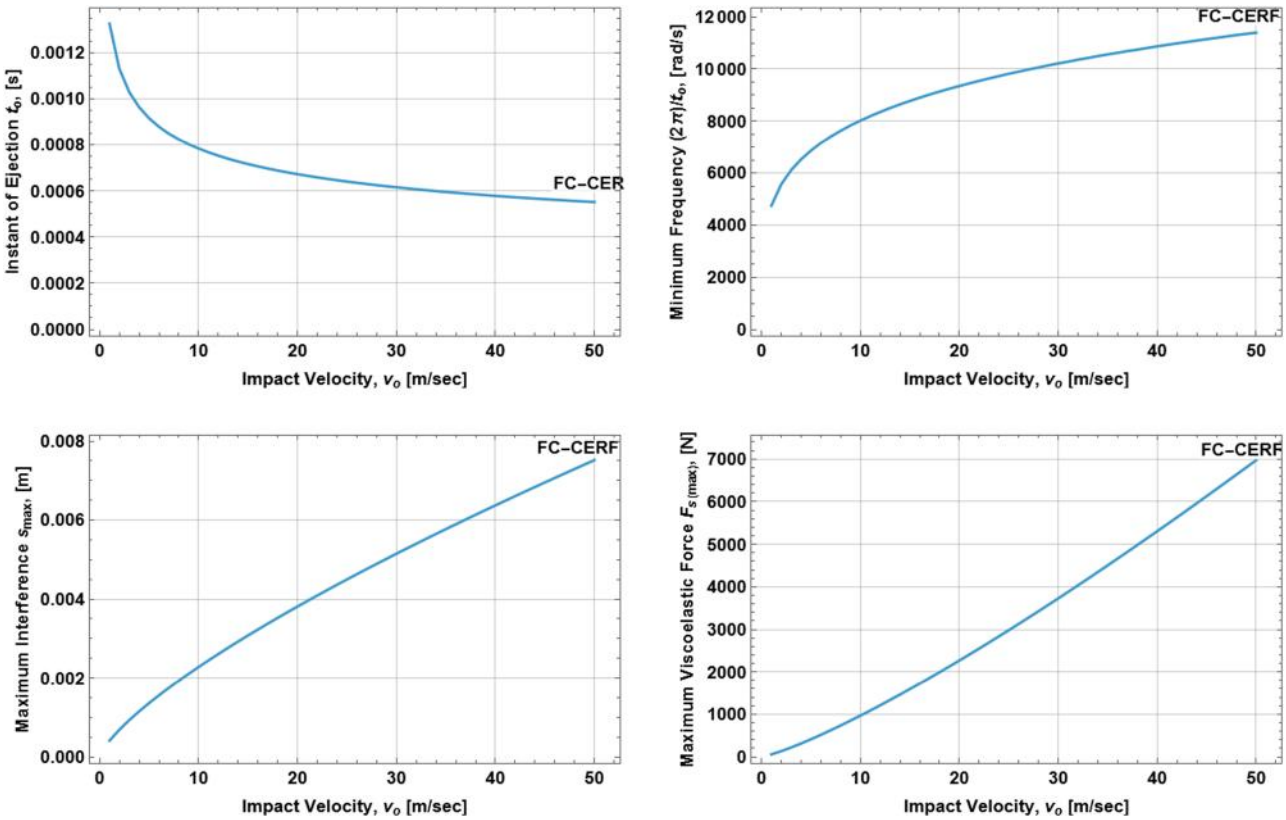

Fig. 9 Effects of the impact velocity on the time to ejection, the minimum frequency of interest, the maximum interference (indentation), and the maximum contact force

frequency range of our interest, i.e., $\omega \geq \omega_{\min} = 8102$ rad/s. Noticing that while the FC-CERF data and model have just entered the transition from rubbery toward glassy, the three-term Prony model is already past its own transition with a leveling storage modulus and with a sharply declining loss modulus. Recall that the Boltzmann convolution integral accounts for all time, i.e., tracing all the way back to $t = 0$. Hence, these Maxwell model shortcomings are bound to affect the COR values, and of course they do (as further discussed below);

(3) While the Maxwell model can use both *forms a* and *b* in Eq. (8), for the FC-based models only *form a* is feasible because the time derivative of $e^{\beta_1^2 t}\text{erfc}\left(\beta_1\sqrt{t}\right)$ is singular at $t = 0$.

The aforementioned direct formulation is now applied to the nonlinear OIDE using the data-faithful of Eq. (23*b*), then to the numerically fitted FC-CERF model of Eq. (24), and lastly to the three-term Prony series model approximation of Eq. (26). The results are shown in Figs. 6 and 7, still for an impact velocity of $v_o = 10$ m/s. Figure 6 shows outcomes for the first two material models of Eqs. (23*b*) and (24) (but because the results are numerically identical, only one plot is shown).

Figure 7 represents the solution obtained for the three-term Prony series model of Eq. (26), using the direct formulation on the nonlinear OIDE, but it is also a true representation of the solution obtained from the IVE formulation (as both formulations produce again numerically identical results).

While the dynamic responses trend quite similarly between Figs. 6 and 7, the values for COR are visibly off where the three-term Prony model produces a COR value that is about 24% higher than that for the FC-CERF model. That is caused by the dramatic drop of the loss modulus for the three-term Prony model (see Fig. 5) where the FC-CERF model still maintains a much higher value of the loss modulus, resulting in a corresponding lower COR. Some of the other outcomes are in better agreement (e.g., the maximum contact forces and t_0 are somewhat closer, while the maximum deformations are only about 11% apart). The hysteretic energy loss is shown in Fig. 3 and is about 1.172 J with only a small difference between the material models.

Next, the versatility of the direct time integration code is now used to perform a parametric investigation to gauge the impact velocity effects on the results (the compiled code takes about 6 s to complete at once the simulation for the entire impact velocity range of $v_o = \{1, 50\}$ m/s, in steps of 1 m/s). The first comparison is given in Fig. 8, showing how the COR is affected by the two material models. As explained above, the three-term Prony model enters the "glassy" state sooner leading to more restoring "elasticity," while the FC-CERF model is still within a "rubbery" state, i.e., energy loss is manifested as "damping," and hence the lower COR values. Yet, the COR values for both material models decline with the increase in impact velocity (in contrast to the trends shown in Fig. 4 for the previous SLS material example).

Lastly, because the FC-CERF model provides a close representation of the data in Ref. [22], only that material model is now used for a parametric investigation to gauge how other parameters are affected by the impact velocity, and that is shown in Fig. 9. With an increasing impact velocity, v_o, the duration to ejection, t_0, decreases (hence, $\omega_{\min}$ increases), the maximum interference (deformation) and the maximum viscoelastic force also increase. Evidently, the impact velocity affects these parameters quite significantly. The highest frequency value of $\omega_{\min}$ is nearly 12,000 rad/s, which is still well situated in the material representation, shown in Fig. 5. The hysteretic energy loss also increases monotonically with v_o, where at $v_o = 50$ m/s, it reaches the value of 30.7 J.

8 Conclusions

This work computes the COR for collisions involving at least one viscoelastic material using fully numerical procedures. Starting from the modified governing wave equation, the Hertzian point contact law is reformulated into a convolution representation suitable for direct insertion into the governing equation of motion. Two numerical approaches are developed:

(1) *A direct time marching scheme* that operates on the full nonlinear integro-differential governing equation (OIDE).
(2) *An internal variable evolution for*mulation that replaces the more laborious direct time marching scheme by introducing evolution equations yielding a system of first-order ODEs that can be integrated rapidly.
(3) Both approaches produce independently identical numerical results for the same material and conditions. The direct method codes can accommodate any arbitrary time-dependent kernels; the internal variable evolution method is constructed specifically for a Maxwell material model. The direct method codes also have the Volterra series built in to capture additional nonlinear effects (which are found to be insignificant in the cases investigated herein).
(4) Time histories are generated for the key quantities of interest: interference (deformation), relative velocity, and contact force. Defining separation as the instant when the contact force returns to zero pinpoints the corresponding ejection relative velocity and consequently the vCOR value. Because the material response is hereditary, a residual indentation remains at the surface at the instant of ejection.
(5) Two materials are examined:
   (i) a Zener material model, equivalent to a single Prony Maxwell representation, and
   (ii) a real material, BUTYL-B252, modeled accurately using a fractional derivative constitutive law.
(6) The results show that the COR may either increase or decrease with impact velocity, depending on whether the material response is dominated by rubbery or glassy behavior. Although the impact duration decreases with increasing velocity, the deformation amplitude, contact force, and hysteretic energy loss all increase, often substantially.
(7) The central contribution of this work is the incorporation of hereditary material behavior directly into a modified wave equation, together with numerical solution methods that compute vCOR without relying on ad hoc combinations of stiffness and damping parameters, which cannot capture hereditary effects. The compiled direct numerical formulation ought to be the method of choice because it is the most robust, it is speedy, it can handle any kernel form, and it incorporates the Volterra series whenever it is necessary.

Conflict of Interest

There are no conflicts of interest.

Data Availability Statement

The authors attest that all data for this study are included in the article.

Nomenclature

h = half plate thickness, Eq. (3)
m = sphere mass
r = sphere radius
s = Laplace complex variable
t = time
v = velocity
G_i = Maxwell model *i*th material constant, Eq. (17)
D^α = fractional derivative (FC) of order α, Eq. (23)
$h_i(t)$ = Volterra series *i*th kernel, Eq. (10)
$k(t)$ = kernel, Eq. (5)

$s(t)$ = relative displacement (interference), $z(t) - U(t)$
$u(t)$ = displacement function, $s(t)^{3/2}$, Eq. (5)
$z(t)$ = sphere center degree-of-freedom
$E(t)$ = time-dependent elastic modulus
$E_i(t)$ = elastic modulus for material $i = a$ or b
$F_s(t)$ = contact force
$Q_i(t)$ = Maxwell model ith internal variable, Eq. (18)
$U(t)$ = plate mid-plane displacement
α = SLS material property, Eq. (13)
$\alpha(t)$ = Zener's plate parameter, Eq. (3)
β_i = SLS material property, Eq. (13), ith decay parameter, Eq. (17)
δt = time marching increment in the numerical direct formulation
ε = strain
λ = Zener's "inelasticity" parameter
ν = Poisson's ratio
ρ = density
σ = stress
τ = dummy time variable
τ_i = relaxation time, $1/\beta_i$, Eq. (17)
ω = frequency

Subscripts

0 = instant designating when $F_s(t) = 0$
a = sphere
b = plate
o = impact velocity

Superscripts

α = fractional derivative order, Eq. (23)
$'$ = equivalent modulus, Eq. (1)
$*$ = complex dynamic modulus, Eq. (7)

Appendix: The Exponentially Scaled Complementary Error Function

The exponentially scaled complimentary error function is $e^{x^2}\text{erfc}(x)$. Its direct computation would use the intrinsic functions $\exp(x^2)$ and $\text{erfc}(x)$. These two monotonic functions trend oppositely where the former hastily increases while the latter hastily diminishes. Direct computation runs into over- and underflow problems in double-precision computation at about $x \approx 26.6$, to be specific. To allow computation for x values larger than that (indeed, the case herein of BUTYL-B252 requires x values of 40 and above), adjustments are needed. An early implementation [15,29,30] of the function CERF(x) uses the rational approximation (7.1.26) of $\text{erfc}(x)$ given in Ref. [31]. While that early CERF(x) implementation is valid for any $x \geq 0$, it is of single-precision significance producing a 5% error already at about $x = 20$, with an error that escalates quite rapidly with larger x (it still trends correctly above $x = 20$). In that implementation, the leading exponential $\exp(x^2)$ simply vanishes by cancelation leaving a rational expression that is numerically stable. The following development uses a similar cancelation concept to achieve a higher accuracy valid for any value of x. Starting with the known [31,32] asymptotic expansion of $\text{erfc}(x)$ for "large" x,

$$\text{erfc}(x) \approx \frac{e^{-x^2}}{x\sqrt{\pi}} \sum_{n=0}^{\infty} (-1)^n \frac{(2n-1)!!}{(2x^2)^n} \tag{A1}$$

where (!!) is the double-factorial. It is immediately apparent that

$$e^{x^2}\text{erfc}(x) \approx \frac{1}{x\sqrt{\pi}} \sum_{n=0}^{\infty} (-1)^n \frac{(2n-1)!!}{(2x^2)^n} \tag{A2}$$

leaving a numerically stable expansion on the right-hand side of Eq. (A2). Note that for "large" x, CERF(x) trends as $(x\sqrt{\pi})^{-1}$. So, now CERF(x) can be defined for any value of x, "small" or "large,"

$$\text{CERF}(x) = \begin{cases} e^{x^2}\text{erfc}(x) & \forall \quad x \leq x_s \\ \dfrac{1}{x\sqrt{\pi}}\left[1 - \dfrac{1}{2x^2} + \dfrac{3}{(2x^2)^2} - \dfrac{15}{(2x^2)^3} + \dfrac{105}{(2x^2)^4} - \dfrac{945}{(2x^2)^5} + \ldots\right] & \forall \quad x > x_s \end{cases} \tag{A3}$$

For double-precision accuracy, $x_s \approx 25$ is a reasonable switch value where the intrinsic functions can safely be used for $x \leq x_s$; however, the expansion is to be used for $x > x_s$. For practical double-precision computation, Eq. (A3) with a six-term expansion renders an error relative to the bracketed term in the order of 10^{-16} for any x. Equation (A3) is easy to program and is transportable to any legacy computer language. However, some modern and recently updated languages contain the CERF(x) function intrinsically: For example, PYTHON has the function erfcx(x) in its library, while Fortran 2008 and above has intrinsically the built-in function erfc_*scaled*(x). Mathematica does not have a dedicated function but CERF(x) can straightforwardly be programed via the *HermiteH*[n,x] polynomial using this single line of code,

$$\text{Cerf}[x\_] := 2/\text{Sqrt}[Pi]\text{Hermite}H[-1, N[x]] \tag{A4}$$

In summary, the aforementioned functions are all synonyms as they achieve a common objective: to provide high-precision and numerical stability in the computation of the scaled complimentary error function for any x value.

References

[1] Zener, C., 1941, "The Intrinsic Inelasticity of Large Plates," Phys. Rev., **59**(8), pp. 669–673.
[2] Green, I., 2022, "The Prediction of the Coefficient of Restitution Between Impacting Spheres and Finite Thickness Plates Undergoing Elastoplastic Deformations and Wave Propagation," Nonlinear Dyn., **109**(4), pp. 2443–2458.
[3] Jackson, R. L., Green, I., and Marghitu, D. B., 2010, "Predicting the Coefficient of Restitution of Impacting Elastic-Perfectly Plastic Spheres," Nonlinear Dyn., **60**(3), pp. 217–229.
[4] Marinack, M. C., Musgrave, R. E., and Higgs, C. F., 2013, "Experimental Investigations on the Coefficient of Restitution of Single Particles," Tribol. Trans., **56**(4), pp. 572–580.
[5] Patil, D., and Fred Higgs, C., 2017, "Critical Plate Thickness for Energy Dissipation During Sphere–Plate Elastoplastic Impact Involving Flexural Vibrations," ASME J. Tribol., **139**(4), p. 041104.
[6] Patil, D., and Fred Higgs, C., 2018, "Experimental Investigations on the Coefficient of Restitution for Sphere–Thin Plate Elastoplastic Impact," ASME J. Tribol., **140**(1), p. 011406.
[7] Banerjee, A., Chanda, A., and Das, R., 2017, "Historical Origin and Recent Development on Normal Directional Impact Models for Rigid Body Contact Simulation: A Critical Review," Arch. Comput. Methods Eng., **24**(2), pp. 397–422.
[8] Zhu, R., Liu, X. N., and Huang, G. L., 2015, "Study of Anomalous Wave Propagation and Reflection in Semi-Infinite Elastic Metamaterials," Wave Motion, **55**, pp. 73–83.
[9] Khan, M. H., Li, B., and Tan, K. T., 2018, "Impact Load Wave Transmission in Elastic Metamaterials," Int. J. Impact Eng., **118**, pp. 50–59.
[10] Wen, K., and Chen, X., 2020, "Analysis of the Stress Wave and Rarefaction Wave Produced by Hypervelocity Impact of Sphere Onto Thin Plate," Def. Technol., **16**(5), pp. 969–979.
[11] Moon, S.-I., Kang, T., Seo, J.-S., Lee, J.-H., Han, S.-W., and Park, J.-H., 2018, "Plate Bending Wave Propagation Behavior Under Metal Sphere Impact Loading," J. Mech. Sci. Technol., **32**(3), pp. 1117–1124.
[12] Chen, W., Wang, X., He, Y., Liu, Y., and Cai, Y., 2025, "Study on Modeling and Validation of an Improved Contact Force in Multibody System With Clearance Joint," J. Mech. Sci. Technol., **39**(9), pp. 4917–4939.
[13] Mainardi, F., 2022, *Fractional Calculus and Waves in Linear Viscoelasticity: An Introduction to Mathematical Models*, World Scientific (Connect), Singapore.
[14] Gutierrez-Lemini, D., 2014, *Engineering Viscoelasticity*, Springer US, Boston, MA.
[15] Szumski, R., and Green, I., 1991, "Constitutive Laws in Time and Frequency Domains for Linear Viscoelastic Materials.," J. Acoust. Soc. Am., **90**(4), pp. 2292–2292.
[16] Smyth, P., Green, I., Jackson, R., and Hanson, R. R., 2012, "Stress Relaxation of Articular Cartilage in Unconfined Compression," ASME/STLE 2012

International Joint Tribology Conference, American Society of Mechanical Engineers, pp. 35–37.
[17] Smyth, P. A., Green, I., Jackson, R. L., and Hanson, R. R., 2014, "Biomimetic Model of Articular Cartilage Based on In Vitro Experiments," J. Biomimetics Biomater. Biomed. Eng., **21**, pp. 75–91.
[18] Smyth, P. A., Varney, P. A., and Green, I., 2016, "A Fractional Calculus Model of Viscoelastic Stator Supports Coupled With Elastic Rotor–Stator Rub," ASME J. Tribol., **138**(4), p. 041101.
[19] Smyth, P. A., and Green, I., 2017, "Analysis of Coupled Poroviscoelasticity and Hydrodynamic Lubrication," Tribol. Lett., **65**(1), p. 1.
[20] Smyth, P. A., and Green, I., 2018, "Storage and Loss Characteristics of Coupled Poroviscoelastic and Hydrodynamic Systems for Biomimetic Applications," ASME J. Tribol., **140**(4), p. 041703.
[21] Varney, P., and Green, I., 2014, "Rotordynamic Analysis Using the Complex Transfer Matrix: An Application to Elastomer Supports Using the Viscoelastic Correspondence Principle," J. Sound Vib., **333**(23), pp. 6258–6272.
[22] Bagley, R. L., and Torvik, P. J., 1986, "On the Fractional Calculus Model of Viscoelastic Behavior," J. Rheol., **30**(1), pp. 133–155.
[23] Kong, F., Yu, J., Xu, Y., and Fang, J., 2025, "A Volterra-Series-Based Semi-Analytical Method for Fractionally Damped Nonlinear Systems Under Stochastic Excitation," Mech. Syst. Signal Process., **241**, p. 113579.
[24] Moch, K., Gainaru, C., and Böhmer, R., 2024, "Nonlinear Susceptibilities and Higher-Order Responses Related to Physical Aging: Wiener–Volterra Approach and Extended Tool–Narayanaswamy–Moynihan Models," J. Chem. Phys., **161**(1), p. 014502.
[25] Patlashenko, I., Givoli, D., and Barbone, P., 2001, "Time-Stepping Schemes for Systems of Volterra Integro-Differential Equations," Comput. Methods Appl. Mech. Eng., **190**(43–44), pp. 5691–5718.
[26] Shaw, S., and Whiteman, J. R., 1998, "Some Partial Differential Volterra Equation Problems Arising in Viscoelasticity," The Proceedings of the EQUADIFF 9 Conference.
[27] ScienceDirect, "Volterra Series—An Overview|ScienceDirect Topics," https://www.sciencedirect.com/topics/computer-science/volterra-series
[28] Bagley, R. L., and Torvik, P. J., "Fractional Calculus in the Transient Analysis of Viscoelastically Damped Structures," AIAA J., **23**(6), p. 918.
[29] Szumski, R., 1993, "A Finite Element Formulation for the Time Domain Vibration Analysis of an Elastic-Viscoelastic Structure," Ph.D. Thesis, Georgia Institute of Technology, Atlanta, GA.
[30] Smyth, P. A., and Green, I., 2015, "Fractional Calculus Model of Articular Cartilage Based on Experimental Stress-Relaxation," Mech. Time-Depend. Mater., **19**(2), pp. 209–228.
[31] Abramowitz, M., and Stegun, I. A., eds., 2013, *Handbook of Mathematical Functions: With Formulas, Graphs, and Mathematical Tables*, Dover Publications, New York.
[32] Olver, F., 1997, *Asymptotics and Special Functions*, A K Peters/CRC Press, New York.