Optimal control of the spatial motion of a viscoelastic rod

Thomas I. Seidman
Department of Mathematics and Statistics
University of Maryland Baltimore County
Baltimore, MD 21250,
e-mail: ⟨seidman@math.umbc.edu⟩
and
Stuart S. Antman
Department of Mathematics,
Institute for Physical Science and Technology, and
Institute for Systems Research
University of Maryland College Park
College Park, MD 20742
e-mail: ⟨ssa@math.umd.edu⟩

10/31/'02 [May 2, 2010]

Abstract

1. Introduction

We consider a model [1] of a nonlinearly viscoelastic rod moving in 3-dimensional space, taking into account not only longitudinal and transverse motions, but also shear and torsional motion: One may visualize this, discretized, as a chain of hard vertebrae connected by a viscous springy material. This paper is related to the forthcoming [3] much as [9] was related to [2]. The papers [2] and [9] treat the purely longitudinal motion of a straight rod. As in [9], we show that the attainment of optimality for certain control problems is intimately related to the considerations involved in showing the existence of solutions, in particular, to the requirement that a subsequential limit of solutions to some approximating problems should be solutions of a desired limit problem.

For the model we consider, the 'geometric state' at each point of the reference configuration (which we take parametrized by $s \in [0,1]$) consists of the position $\mathbf{r} = \mathbf{r}(t,s)$ in 3-dimensional space and the orientation of the vertebral cross-section. The latter may be specified by a pair of 'directors' — an orthonormal basis $\{\mathbf{d}_1, \mathbf{d}_2\}$ for the plane of the cross-section, which is then extended to a properly oriented basis for $\mathbb{R}^3$. This specification is equivalent to specification of a 3×3 rotation matrix $D = D(t,s) \in SO(3)$ (i.e., D is orthogonal with determinant $+1$) which transforms the fixed coordinate system to this one. Thus, the relevant space (pointwise) is $\mathcal{M} = \mathbb{R}^3 \times SO(3)$ and the actual geometric state is a function q, e.g., $q(t\cdot) \in C^1([0,1] \to \mathcal{M})$.

Our present model presents a significant new difficulty which does not arise for the restricted version of [2], [9]: this state space is here a manifold rather than a linear space. More precisely, since most of our analysis works with velocities (momenta) and local strains which lie in the tangent space, we note that $SO(3)$ is 3-dimensional, so the tangent space to $\mathcal{M}$ is pointwise isomorphic to $\mathbb{R}^6$ and the tangent space to $\mathcal{X}_0$ is isomorphic to a fixed linear space of $\mathbb{R}^6$-valued functions — but the relation to that fixed space is varying and derivatives of that relation complicate our analysis.

Of course, the analytic difficulties already ocurring in [2], [9] also continue to be relevant in the present more general setting:

- Preclusion of 'total compression': Locally, the rod material should be bounded away from passing through itself. This consideration constrains the domain of the constitutive function σ and ensures that it cannot possibly be uniformly Lipschitzian; at the same time we also avoid imposing on σ any growth rate for the response to large extensions (stretching).
- Nonlinearity of the viscous dissipation: We assume a strong monotonicity in the response to strain rate, but no such strong structural condition will be imposed regarding the dependence of this on the strain itself.

2. Formulation

It will be desirable to treat both the constitutive relations and the velocities in the coordinatization given pointwise by $D(t,s)$. Thus we introduce $v, p \in \mathbb{R}^3$ such that

$$\mathbf{r}_s = Dv \qquad \mathbf{r}_t = Dp. \tag{2.1}$$

Since $D = D(t,s) \in SO(3)$, both derivatives D_s, D_t involve skew-symmetric matrices and so can be represented on $\mathbb{R}^3$ by cross products: we define $v, w \in \mathbb{R}^3$ so

$$D_s = D[u\times] \qquad D_t = D[w\times] \tag{2.2}$$

where $[u\times], [w\times]$ are interpreted as 3×3 matrices. Then the vectors

$$\eta = \begin{pmatrix} u \\ v \end{pmatrix} \qquad \xi = \begin{pmatrix} w \\ p \end{pmatrix} \tag{2.3}$$

in $\mathbb{R}^6$ represent *velocity* and *strain*, respectively, in the pointwise coordinatization. With a little manipulation, equality of mixed partials gives

$$\eta_t = \xi_s + A\xi \qquad \text{with } A = A(\eta) := \begin{pmatrix} u\times & 0 \\ v\times & u\times \end{pmatrix}. \tag{2.4}$$

All the physics and the particularity of the situation then reside in specification of the inertia matrix M and the constitutive function

$$\hat{\sigma} : \mathbb{R}^6 \times \mathbb{R}^6 \longrightarrow \mathbb{R}^6 : y, z \mapsto \sigma = \hat{\sigma}(y, z). \tag{2.5}$$

It is our choice of the pointwise coordinatization which makes M a material property along the rod, which we assume homogeneous[1] for simplicity and ensures the appropriate frame indifference for the contact forces given by σ. Set

$$M = \begin{pmatrix} J & 0 \\ 0 & \rho I \end{pmatrix} \qquad B = B(\xi) := \begin{pmatrix} -[Jw]\times & 0 \\ 0 & \rho p\times \end{pmatrix}, \tag{2.6}$$

(where the scalar ρ is linear mass density and the3×3 matrix J is the density of moment of inertia). Standard continuum mechanics gives

$$M\xi_t = [\sigma]_s - A^*\sigma - B\xi + f \qquad \text{with } \sigma = \hat{\sigma}(\eta, \eta_t). \tag{2.7}$$

[Here f corresponds to any external body forces and, for present purposes, we assume $f = 0$.] We note that $\rho > 0$, J is positive definite, and B is skew; the dissipativity of the stress-strain relation given by (2.5) is indicated by (3.2), below. For future reference we also note that

$$[M\xi_t + B\xi] = D^*[DM\xi]_t \tag{2.8}$$

For a more detailed discussion of this derivation, see [1], [3], although the notation here is rather different.

For definiteness, we consider the problem boundary conditions corresponding to having one end rigidly fixed:

$$q = (D, \mathbf{r})\Big|_{s=0} \equiv (D_*, \mathbf{r}_*) = \text{const.} \in \mathcal{M}$$

and with the contact load specified at the other end:

$$[D\sigma]\Big|_{s=1} = \nu = \nu(t).$$

Here we abused notation slightly in also using D for the 6×6 matrix $\begin{pmatrix} D & 0 \\ 0 & D \end{pmatrix}$.

It follows that we have for (2.7) the boundary conditions

$$\xi\Big|_{s=0} = 0, \qquad \hat{\sigma}(\eta, \eta_t)\Big|_{s=1} = D_1^*\nu \tag{2.9}$$

[1]This makes M constant, although our derivation would admit a possible (suppressed) dependence on s. We comment that C^1 dependence on s, for M and also for $\hat{\sigma}$, would be easy to handle and, with moderate care, one can even treat piecewise continuous material properties.

with $D_1(t) := D(t,1)$. [Note that, without retaining D as a separate variable, we can recover D_1 as needed from η by solving the $SO(3)$-valued ordinary differential equation in s:

$$D_s = D[v\times] \qquad D\Big|_{s=0} = D_* \tag{2.10}$$

for each fixed t and evaluating at $s = 1$.] With these boundary conditions the weak form of (2.7) with $f = 0$ becomes

$$\langle \zeta, M\xi_t\rangle + \langle \zeta_s + A\zeta, \varphi'(\eta) + \sigma^{\mathrm{D}}(\eta, \eta_t)\rangle = -\langle \zeta, B\xi\rangle + [D_1\zeta(1)]\cdot\nu \tag{2.11}$$

for all suitable $\mathbb{R}^6$-valued test functions ζ. [Note that $\langle\cdot,\cdot\rangle$ is a product pivoting on the usual L^2 inner product: $\langle f, g\rangle = \int_0^1 f\cdot g\,ds$; later we will also use $\langle f\rangle$ (without the comma) for $\int_0^1 f\,ds$.]

We adjoin to this the initial conditions

$$\xi\Big|_{t=0} = \overset{\circ}{\xi} \qquad \eta\Big|_{t=0} = \overset{\circ}{\eta} \tag{2.12}$$

and, assuming suitable regularity, note that $\overset{\circ}{\xi}, \overset{\circ}{\eta}$ can be obtained as in (2.1), (2.2), (2.3) from $q, q_t\Big|_{t=0}$, which are presumably given.

As formulated above we are considering, as a system for the unknown variables ξ, η (which are be taken in some appropriate spaces of $\mathbb{R}^6$-valued functions on $\mathcal{Q} = \mathcal{Q}_T = [0,T]\times[0,1]$):

$$\text{(2.4) and (2.7) with (2.9) and (2.12) [and (2.10)]} \tag{2.13}$$

with the functions $\overset{\circ}{\xi}, \overset{\circ}{\eta}, \nu$ (and D_*) as data. Using (2.10) we obtained D as part of the solution process for (2.13), so forces and velocities could be converted to the fixed (laboratory) coordinate system, if desired. Obviously, one can recover $\mathbf{r}$ also by using $\mathbf{r}_*$ and integrating Du in s for each t — or, using the original initial data, integrating Dp in t for each s.

3. Energy

Our fundamental structural hypotheses regarding the constitutive function $\hat\sigma$ are hyperelasticity of the equilibrium response and uniform monotonicity with respect to the second variable. Thus we first assume that the equilibrium response is given by a potential (stored energy function):

$$\hat\sigma(\eta, 0) = \frac{d\varphi}{dy} = \varphi'(y) \qquad \text{where } \varphi : \mathbb{R}^6 \to [0,\infty] \tag{3.1}$$

[Note that there is no suggestion that φ should have any convexity property.] We then introduce $\sigma^{\mathrm{D}}(y,z) := \hat\sigma(y,z) - \hat\sigma(y,0)$ and assume a uniform dissipativity condition:

$$[z_1 - z_2]\cdot[\hat\sigma(y,z_1) - \hat\sigma(y,z_2)] = [z_1 - z_2]\cdot[\sigma^{\mathrm{D}}(y,z_1) - \sigma^{\mathrm{D}}(y,z_2)] \geq \mu|z_1 - z_2|^2 \tag{3.2}$$

where $\mu > 0$ is a fixed constant; note that, if we assume — as we do, henceforth, for simplicity — that $\hat{\sigma}$ is at least of class C^1 where finite, then (3.2) is equivalent to having (pointwise in $y, z \in \mathbb{R}^6$)

$$\frac{\partial \hat{\sigma}}{\partial z} = \frac{\partial \sigma^{\mathrm{D}}}{\partial z} \geq \mu \tag{3.3}$$

with the inequality in the sense of quadratic forms so the 6×6 matrix σ_z^{D} is (uniformly) positive-definite. [We note that (3.2) gives, in particular, $z \cdot \sigma^{\mathrm{D}}(y,z) \geq \mu |z|^2$ since $\sigma^{\mathrm{D}}(y,0) = 0$.] This will permit us to get a fundamental estimate for

$$\mathcal{E} = \mathcal{E}(t) := \tfrac{1}{2}\langle \xi, M\xi \rangle + \langle \varphi(\eta) \rangle + \int_0^t \langle \eta_t, \sigma^{\mathrm{D}}(\eta, \eta_t) \rangle \tag{3.4}$$

— which we recognize as the sum of kinetic energy, potential energy, and also cumulative dissipative work.

As usual, we take $\zeta = \xi$ in (2.11) to obtain the desired estimate. Using (2.4), noting that $\eta_t \cdot \varphi'(\eta) = d\varphi(\eta)/dt$, and computing $d\mathcal{E}/dt$ from (3.4), we see that

$$\mathcal{E}(t) = \mathcal{E}(0) + \int_0^t D_1\xi \cdot \nu(\tau)\, d\tau. \tag{3.5}$$

We are able to apply the Gronwall Inequality to obtain a bound for $\mathcal{E}(t)$ once we estimate $|D_1\xi(1)| = |\xi(1)| \leq \|\xi\|_\infty$ in terms of $\mathcal{E}$. Without full details, this proceeds as:

$$\begin{aligned} \|\xi\|_\infty &\leq \|\xi_s\|_1 = \|\eta_t - A\xi\|_1 \\ &\leq C\left[\|\eta_t\| + \|\eta\|\|\xi\|\right] \leq \cdots \end{aligned}$$

where $\|\cdot\|_p$ is the $L^p(0,1)$-norm and $\|\cdot\| = \|\cdot\|_2$. Since M is positive definite and $\eta_t \cdot \sigma$ dominates $|\eta_t|^2$, our estimate for $\mathcal{E}(t)$ gives:

$$\begin{array}{l} \xi \text{ is bounded in } L^\infty(\to L^2), \\ \langle \varphi(\eta) \rangle \text{ is bounded pointwise in } t \in [0,T], \\ \eta_t \cdot \sigma \text{ is bounded in } L^1(\mathcal{Q}) \text{ so } \eta_t \text{ in } L^2(\mathcal{Q}), \\ \text{so } \eta \text{ is bounded in } L^\infty(\to L^2), \\ \text{and } \xi \text{ is bounded in } L^2(\to H^1) \text{ and in } L^2(\to L^\infty). \end{array} \tag{3.6}$$

in terms of the $L^2(0,T)$-norm of $\nu(\cdot)$)

4. Total compression

So far we have not said much about the potential $\varphi(\cdot)$, other than non-negativity, but we now note on geometric grounds that its domain $\mathcal{A}$ cannot be all of $\mathbb{R}^6$. This easiest to see when the rod is straight and it is obvious that, if the longitudinal component of $\mathbf{r}$ were not increasing in s (making the corresponding component of v positive), then the rod material would be interpenetrating itself. For a physical rod with some (very small) cross-sectional diameter h, one similarly sees that the vertebrae of our discretized visualization should not interpenetrate by having $|D_s|$ too large when the longitudinal component of v is (positive but) small. We control this by introducing an auxiliary function $\psi : \mathbb{R}^6 \longrightarrow [0,\infty]$ which blows up as its argument approaches the boundary of $\mathcal{A}$ so $\mathcal{A} = \bigcup_c \mathcal{A}_c$ where $\mathcal{A}_c := \{y \in \mathbb{R}^6 : \psi(y) \le c\}$.

One natural extension of our considerations in [2], taking advantage of the formulation there in terms of an inequality, is to impose as a possible hypothesis the assumption (perhaps restricting the choice of the representing ψ and its relation to the constitutive function $\hat{\sigma}$) that:

There is some γ in $\mathbb{R}^6$ of the form $(0,\overline{\gamma})$ with $\overline{\gamma} \in \mathbb{R}^3$ and some $\beta \in \mathbb{R}$ such that

$$\psi'(y) \cdot z \le \gamma \cdot [\hat{\sigma}(y,z)] + \beta \tag{4.1}$$

for all z and for all $y \in \mathcal{A} \setminus \mathcal{A}_{\overline{c}}$ where $\overline{c}$ is large enough that $\mathring{\eta}\ (s) \in \mathcal{A}_{\overline{c}}$ for each s.

We show that this hypothesis, together with (3.6), uniformly bounds the solution pointwise against total compression:

$$\eta(t,s) \in \mathcal{A}_c \ \text{ for some } c, \tag{4.2}$$

i.e., $\psi(\eta(\overline{t},\overline{s})) \le c$ for $0 \le \overline{s} \le 1$, $0 \le \overline{t} \le T$. It would be convenient if we could have (4.1) with ψ coercive, so this would simultaneously show a bound on $|\eta|$, but we do not require this and do not expect it.

The key to the argument is the form of γ, with the observation that, although our model does *not* give $D[\sigma_s - A^*\sigma] = [D\sigma]_s$, the form of A is such that this *does* hold as an identity for the "lower components." Thus,

$$\begin{aligned} \gamma \cdot \sigma\Big|_{\overline{s}} &= (D\gamma)\Big|_{\overline{s}} \cdot \left[\nu - \textstyle\int_{\overline{s}}^1 (\sigma_s - A^*\sigma)\right] \\ &= (D\gamma)\Big|_{\overline{s}} \cdot \left[\nu - \textstyle\int_{\overline{s}}^1 (DM\xi)_t\right]. \end{aligned} \tag{4.3}$$

The argument now proceeds as in [2]: if we ever were to have $\psi(\eta)\Big|_{(\overline{t},\overline{s})} > \overline{c}$ we could find $0 < \tau < \overline{t} \le T$ with

$$\psi(\eta)\Big|_{(\tau,\overline{s})} = \overline{c} \text{ and } \psi(\eta)\Big|_{(t,\overline{s})} > \overline{c} \text{ for } \tau \le t \le \overline{t}.$$

Integrating $[\psi(\eta)]_t = \psi' \cdot \eta_t$ and using (4.1) gives

$$\psi\Big|_{(\overline{t},\overline{s})} \le \overline{c} + \int_\tau^{\overline{t}} \left(\beta + (D\gamma)\Big|_{\overline{s}} \cdot \left[\nu - \int_{\overline{s}}^1 (DM\xi)_t\right]\right)$$

and we note that the right side of this is uniformly bounded in terms of the bounds in (3.6) once we note, e.g., that $|(D\gamma)_t| \le \|\xi\|_\infty |\overline{\gamma}|$.

5. Two optimal control problems

At this point we see how state compactness can be provided either directly by a constraint or indirectly through a term in the objective function to facilitate the argument for existence of an optimal control. The two examples we consider here each refer to boundary control — considering the system (2.13) with fixed initial data and fixed $\mathbf{r}_*, D_*$, but with the endpoint contact force $\nu(\cdot)$ taken as a control. Each requires that the state reach a target:

$$(q, q_t)\Big|_{t=T} \in \mathcal{S} \tag{5.1}$$

where $\mathcal{S}$ is a suitable closed set in the appropriate space.

EXAMPLE 5.1: Minimize the time T required to control the state to the given target set as in (5.1), subject to a control constraint: $|\nu(t)| \le 1$ and a state constraint:

$$\eta(t) = \eta(t, \cdot) \in \mathcal{K} \tag{5.2}$$

for each $t \in [0, T]$.

EXAMPLE 5.2: With T fixed, find a control $\nu(\cdot)$ minimizing the cost

$$\mathcal{J} = \int_0^T \left[\|\eta(t, \cdot)\|_{\mathcal{V}}^2 + |\nu(t)|^2 \right] dt \tag{5.3}$$

subject to (5.1).

For Example 5.1 we assume that the constraint set $\mathcal{K}$ lies in $L^\infty([0,1] \to \mathbb{R}^6)$, uniformly avoids total compression: $\eta(t,s) \in \mathcal{A}_c$ for some c, and is compact with respect to the topology of pointwise ae convergence,

For Example 5.2 we assume that the space $\mathcal{V}$ is embedded into $L^\infty([0,1] \to \mathbb{R}^6)$ and compactly embedded into L^2 and also assume that $\hat{\sigma}$ satisfies the condition (4.1).

For each of the examples we assume that the choice of target set $\mathcal{S}$ is consistent with the problem: there is at least one admissible control (and T for Example 5.1) such that (5.1) is satisfied. We will also assume a linear growth rate for $\hat{\sigma}$

$$|\sigma(y, z)| \le C[|y| + |z|] \quad \text{for all } z \in \mathbb{R}^6,\ y \in \mathcal{A}_c. \tag{5.4}$$

It then follows, for each of these examples, that there will be a minimizing sequence $\{\nu^k\}$ with a corresponding sequence of solutions $[\eta^k, \xi^k]$ of (2.13). Without loss of generality, we may assume that $\nu^k \rightharpoonup \nu^\infty$ in $L^2(0,T)$ and then must show that there is a (possibly subsequential) limit $[\eta^\infty, \xi^\infty]$ of the solutions in a sense which permits us to conclude that $[\eta^\infty, \xi^\infty]$ satisfies (2.13) and gives the terminal condition (5.1) at T. In each case, the key will be to show the convergence: $\eta_t^k \to \eta_t^\infty$ in $L^2(\mathcal{Q})$.

Note that the $L^2(0,T)$-bound on $\{\nu^k\}$ makes (3.6) applicable and ensures (4.2) for Example 5.2 as well. With (5.4) we then also have a bound in $L^2(\mathcal{Q})$ for $\{\sigma^k := \hat{\sigma}(\eta^k, \eta_t^k)\}$. This bounds $[\sigma^k]_s$ in $L^2(\to H^{-1})$ and $(A^k)^*\sigma^k, B^k\xi^k$ are bounded in $L^2(\to L^1)$ so ξ_t^k is bounded in $L^2(\to$ [some space]). Given the bound on ξ^k in $L^2(\to H^1)$, we may apply the Aubin Compactness Theorem [4] to see that $\{\xi^k\}$ lies in a compact subset of $L^2(\to H^r[0,1])$ for any $r < 1$: we may assume $\xi^k \to \xi^\infty$ there for some ξ^∞. We also have $B^k \to B^\infty$; since $\{\xi^k\}$ is uniformly bounded in $L^\infty(\to L^2)$, the same bound holds for ξ^∞, etc.

We may also assume $\eta_t^k \rightharpoonup \eta_t^\infty$ for some $\eta_t^\infty \in L^2(\mathcal{Q})$ so also $\eta^k \rightharpoonup \eta^\infty$ with $(\eta^\infty)_t = \eta_t^\infty$. We now set $\sigma^\infty := \hat{\sigma}(\eta^\infty, \eta_t^\infty)$, noting that, while we can assume weak convergence for σ^k, it is not clear at this point that the limit would be σ^∞.

We next need $\eta^k \to \eta^\infty$, e.g., strongly in $L^2(\mathcal{Q})$. For Example 5.1 we adapt an argument from [8]. Let $\mathcal{K}_*$ be the set $\mathcal{K}$, topologized in $L^2(0,1)$. Since the $\mathcal{K}$-topology gives pointwise ae convergence and $\mathcal{K}_*$ is bounded in $L^\infty(0,1)$, use of the Dominated Convergence Theorem gives continuity of the identity: $\mathcal{K} \to \mathcal{K}_*$. Thus, $\mathcal{K}_*$ is also compact. Boundedness of $\{\eta_t^k\}$ shows $\{\eta^k\}$ equicontinuous from $[0,T]$ to $\mathcal{K}_*$ so the Arzelá-Ascoli Theorem gives precompactness in $C([0,T] \to \mathcal{K}_*)$ and we therefore have $\eta^k \to \eta^\infty$ there and *a fortiori* in $L^2(\mathcal{Q})$. For Example 5.2 the assumed compact embedding: $V \hookrightarrow L^2(0,1)$ and boundedness of η_t^k makes the Aubin Theorem [4] applicable and we again get strong $L^2(\mathcal{Q})$-convergence $\eta^k \to \eta^\infty$ and $A^k = A(\eta^k) \to A^\infty = A(\eta^\infty)$. With (5.4), we now note that Krasnosel'skiĭ's Theorem on the continuity of Nemitskiĭ operators (cf., e.g., [7]) then gives strong $L^2(\mathcal{Q})$-convergence:

$$\hat{\sigma}(\eta^k, \eta_t^\infty) \to \hat{\sigma}(\eta^\infty, \eta_t^\infty), \tag{5.5}$$

with the second argument held fixed as η_t^∞.

Finally, we wish to show that we have convergence to the correct limit:

$$\sigma^k = \hat{\sigma}(\eta^k, \eta_t^k) \to \sigma^\infty = \hat{\sigma}(\eta^\infty, \eta_t^\infty), \tag{5.6}$$

which, with our results above, would show that ξ^∞ is the solution of (2.13) corresponding to the (weak limit) control $\nu^\infty(\cdot)$. Fairly standard arguments then show that this control is admissible with respect to the constraints and is then the desired optimal control.

To see (5.6) we use (2.11) for each k with $\zeta = \zeta^k = \xi^k - \xi^\infty$ so $\zeta^k \to 0$. Note that $\zeta_s + A^k\zeta$ is then $[\eta_t^k - \eta_t^\infty] + [(A^k - A^\infty)\xi^\infty]$. Thus, integrating (2.11) and noting that $\xi^k = \xi^\infty$ at $t = 0$, we obtain

$$\begin{aligned} \tfrac{1}{2}\langle \zeta^k, M\zeta^k\rangle\Big|_T \quad & + \langle \eta_t^k - \eta_t^\infty, \hat{\sigma}((\eta^k, \eta_t^k) - \hat{\sigma}(\eta^k, \eta_t^\infty)\rangle_{\mathcal{Q}} \\ & = \int_0^T D_1^k\zeta^k(1)\cdot\nu^k - \langle \zeta^k, B^k\xi^k\rangle_{\mathcal{Q}} \\ & \quad - \langle (A^k - A^\infty)\xi^\infty, \sigma^k - \sigma^\infty\rangle_{\mathcal{Q}} - \langle \zeta^k, M\xi_t^\infty\rangle_{\mathcal{Q}} \\ & \quad - \langle \eta_t^k - \eta_t^\infty, \sigma^\infty\rangle_{\mathcal{Q}} - \langle \eta_t^k - \eta_t^\infty, \hat{\sigma}(\eta^k, \eta_t^\infty) - \hat{\sigma}(\eta^\infty, \eta_t^\infty)\rangle_{\mathcal{Q}}. \end{aligned} \tag{5.7}$$

From (3.2) we have

$$\mu\|\eta_t^k - \eta_t^\infty\|_{\mathcal{Q}}^2 \le \langle \eta_t^k - \eta_t^\infty, \hat\sigma(\eta^k,\eta_t^k) - \hat\sigma(\eta^k,\eta_t^\infty)\rangle_{\mathcal{Q}}$$

and we wish to show that the right-hand side goes to 0 as $k \to \infty$. Since we have $\zeta^k \to 0$ in suitable spaces and $\eta_t^k - \eta_t^\infty \rightharpoonup 0$ and (3.6) and (5.6), each term on the right of (5.7) goes to 0 so $\eta_t^k \to \eta_t^\infty$ in $L^2(\mathcal{Q})$. Now, however, we have $[\eta^k, \eta_t^k] \to [\eta^\infty, \eta_t^\infty]$ in $L^2(\mathcal{Q} \to \mathbb{R}^1 2)$ and can again use Krasnosel'skiĭ's Theorem, in view of (5.4), to conclude (5.6).

Thus, we have shown:

THEOREM 5.1: *Under the hypotheses noted above — compactness and compression avoidance for $\mathcal{K}$, consistency of the target set $\mathcal{S}$, and the growth condition* (5.4) *on the constitutive function — and with suitably regular initial data, the optimization of Example 5.1 attains its minimum: There exists a minimum-time control.*

THEOREM 5.2: *Under the hypotheses noted above and with suitably regular initial data, the optimization of Example 5.2 attains its minimum: There exists an optimal control.*

6. A further condition and estimate

To show the existence of an optimal control without externally imposing compactness on η as in the first two examples, we must obtain compactness otherwise. This will force us to impose a further condition on the constitutive function $\hat\sigma(\cdot,\cdot)$ and also to require more regularity of the control.

We will impose the same auxiliary constitutive condition as is to be used in [3] for obtaining existence of solutions — a generalization of the condition used in [2], there weakening a condition on $\partial\hat\sigma/\partial y$ used similarly in [5]. We now control the η-dependence in terms of the viscous dissipativity σ^{D} by requiring:

$$|\,[\sigma_z^{\mathrm{D}}(y,z)]^{-1/2}\hat\sigma_y(y,z)\,| \le \lambda[1 + z\cdot\sigma^{\mathrm{D}}(y,z) + \varphi(y)] \tag{6.1}$$

for some constant λ. The condition (6.1) is to hold for all $z \in \mathbb{R}^6$ and for all $y \in \mathcal{A}_c$; we may permit λ to depend on c here. [Note that (3.3) already ensures existence of the positive definite matrix $[\sigma_z^{\mathrm{D}}(y,z)]^{-1/2}$ appearing here.]

The significance of (6.1) appears in obtaining an estimate for an artificial pseudo-energy $\mathcal{H}$, imitating (3.4):

$$\mathcal{H} = \mathcal{H}(t) := \tfrac{1}{2}\langle \xi_t, M\xi_t\rangle + \int_0^t \langle \eta_{tt}, \sigma_z^{\mathrm{D}}(\eta,\eta_t)\eta_{tt}\rangle. \tag{6.2}$$

To this end, we differentiate (2.11) with respect to t, getting

$$\langle \zeta, M\xi_{tt}\rangle + \langle \zeta_s + A\zeta, [\hat\sigma(\eta,\eta_t)]_t\rangle = -\langle A_t\zeta, \sigma\rangle - \langle \zeta, [B\xi]_t\rangle + \zeta(1)\cdot[D_1^*\nu]_t\,, \tag{6.3}$$

and take $\zeta = \xi_t$. Note that then $\zeta_s + A\zeta = \eta_{tt} - A_t\xi$ and $[\hat{\sigma}(\eta, \eta_t)]_t = \hat{\sigma}_y\,\eta_t + \sigma_z^{\mathrm{D}}\,\eta_{tt}$ so

$$\begin{aligned} H'(t) &= \langle \zeta, M\zeta_t \rangle + \langle \eta_{tt}, \sigma_z^{\mathrm{D}}\eta_{tt} \rangle \\ &= [\zeta, M\xi_{tt}\rangle + \langle \eta_{tt} - A_t\xi, \sigma_t \rangle] + \langle A_t\xi, \sigma_t \rangle - \langle \eta_{tt}, \hat{\sigma}_y\eta_t \rangle \\ &= -\langle \eta_{tt}, \hat{\sigma}_y\eta_t \rangle + \langle A_t\xi, \sigma_t \rangle \\ &\qquad - \langle A_t\zeta, \sigma \rangle - \langle \zeta, [B\xi]_t \rangle + \zeta(1)] \cdot [D_1^*\nu]_t\,. \end{aligned} \tag{6.4}$$

We first wish to control the term $\langle \eta_{tt}, \hat{\sigma}_y\,\eta_t \rangle$ and of course the hypothesis (6.1) is precisely designed for this estimation, since we can take

$$\eta_{tt} \cdot \hat{\sigma}_y\,\eta_t = \Big([\sigma_z^{\mathrm{D}}]^{1/2}\eta_{tt} \Big) \cdot \Big([\sigma_z^{\mathrm{D}}]^{1/2}\hat{\sigma}_y\,\eta_t \Big)\,.$$

With this in hand, it is rather messy, but not especially difficult, to estimate the remaining terms in terms of $\mathcal{H}$ and (3.6), noting that the boundary term now involves the t-derivative ν'. These estimates enable us to apply the Gronwall Inequality to (6.4) and so provide the desired bound on $\mathcal{H}(t)$. We may conclude that

$$\begin{array}{l} \zeta = \xi_t \text{ is bounded in } L^\infty(\to L^2), \\ \eta_{tt} \text{ is bounded in } L^2(\mathcal{Q}) \text{ so } \eta_t \text{ is bounded in } L^\infty(\to L^2), \\ \text{whence } \xi_s \text{ is also bounded in } L^\infty(\to L^2), \\ \text{so } \xi \text{ is bounded in } L^\infty(\to H^1) \hookrightarrow L^\infty(\mathcal{Q}), \\ \text{and } \xi_t \text{ is bounded in } L^2(\to H^1) \text{ and in } L^2(\to L^\infty). \end{array} \tag{6.5}$$

in terms of the $H^1(0,T)$-norm of $\nu(\cdot)$ and, of course, the initial data and the parameters of our previous estimates. For more detail, we again refer to [3].

We now seek to extract the necessary compactness from this. Note, first that the last line of (6.5) can be improved, using a result from [10], [8] somewhat as for Example 5.1: we actually have ξ in a compact subset — uniformly fixed, again in terms of the $H^1(0,T)$-norm of $\nu(\cdot)$, etc. — of $C(\overline{\mathcal{Q}})$. Next, we consider (2.7) as an ordinary differential equation in s (at each t) for $\sigma = \sigma(t,\cdot) = \hat{\sigma}(\eta(t,\cdot), \eta_t(t,\cdot))$:

$$\sigma_s - A^*\sigma = B\xi + M\zeta \qquad \text{with } \sigma\Big|_{s=1} = \nu(t). \tag{6.6}$$

In view of (6.5), this ensures that $\sigma(t,s)$ is bounded uniformly in $(t,s) \in \overline{\mathcal{Q}}$; indeed, the set of functions $\{\sigma(t,\cdot) : t \in [0,T]\}$ is equicontinuous and so lies in a fixed compact set in $C([0,1])$.

Next we note that

$$\eta_{ts} = \xi_{ss} + [A\xi]_s \qquad [\hat{\sigma}(\eta, \eta_t)]_s = \hat{\sigma}_y\,\eta_s + \sigma_z^{\mathrm{D}}\,\eta_{ts}$$

so, multiplying (2.7) by $-\xi_{ss}$ and integrating, we obtain

$$-\langle \xi_{ss}, M\xi_t \rangle + \langle \eta_{ts}, \sigma_z^{\mathrm{D}}\eta_{ts} \rangle = -\langle \eta_{ts}, \hat{\sigma}_y\eta_s \rangle + \text{“other terms”}.$$

Integrating the first term by parts gives

$$\begin{aligned}\tfrac{1}{2}\langle \xi_s, M\xi_s\rangle_t + \langle \eta_{ts}, \sigma_z^{\mathrm{D}}\eta_{ts}\rangle \\ = [\eta_t - A\xi]\cdot M\xi_t\Big|_{s=1} - \langle \eta_{ts}, \hat{\sigma}_y \eta_s\rangle \qquad + \text{“other terms”}.\end{aligned} \tag{6.7}$$

To estimate $\eta_t(t,1)$ (and $\eta(t,1)$, which we need to bound $A(\eta)$), we proceed as follows: first,

$$\varphi(\eta)\Big|_0^T + \mu\|\eta_t\|^2_{L^2(0,T)} \le \int_0^T \sigma(\eta,\eta_t)\cdot\eta_t\,dt \le K\sqrt{T}\,\|\eta_t\|_{L^2(0,T)}$$

(as $\varphi \ge 0$ and $|\sigma| \le K$) so $\eta_t(\cdot,1)$ is bounded in $L^2(0,T)$ whence $\eta(\cdot,1)$ is bounded. Since $\eta \in \mathcal{A}_c$, we have a bound on φ' as well as on φ at $s=1$ so we can consider the structural description of $\hat{\sigma}$ as giving an ordinary differential equation in t for η along $s=1$:

$$\sigma^{\mathrm{D}}(\eta,\eta_t) = \sigma(t) - \varphi'(\eta) \qquad \eta\Big|_{t=0} = \overset{\circ}{\eta}(1). \tag{6.8}$$

Using (3.3) we can solve (6.8) to express this as an ordinary differential equation in standard form: $\eta_t = \Gamma(\eta, \sigma - \varphi')$ with Γ well-behaved. Thus, with η, σ, φ' bounded, we have η_t bounded pointwise along $s=1$. This permits estimation of the first term on the right of (6.7); the second term is estimated by again using (6.1) as earlier; the "other terms" also can be estimated as desired. Applying the Gronwall Inequality now shows that

$$\begin{aligned}&\xi_s \text{ is bounded in } L^\infty(\to L^2),\\ &\eta_{ts} \text{ is bounded in } L^2(\mathcal{Q}),\\ &\text{so } \eta_t \text{ is bounded in } L^2(\to H^1),\\ &\text{and } \eta_s \text{ is bounded in } L^\infty(\to L^2),\\ &\text{so } \eta \text{ is bounded in } L^\infty(\to H^1) \hookrightarrow L^\infty(\to C[0,1])\end{aligned} \tag{6.9}$$

in terms of the $H^1(0,T)$-norm of $\nu(\cdot)$, etc. As earlier, the results of [10], [8] then give η in a fixed compact subset of $C(\overline{\mathcal{Q}})$. Since (6.5) bounded η_{tt} in $L^2(\mathcal{Q})$ and we have here bounded η_t in $L^2(\to H^1)$, application of the Aubin Theorem also gives η_t in a fixed compact subset of $L^2(\to C[0,1])$; recall that we have an $L^\infty(\mathcal{Q})$ bound for η_t without compactness. Thus we have the desired compactness results:

$$\begin{aligned}&\xi, \eta \text{ are in fixed compact subsets of } C(\overline{\mathcal{Q}}),\\ &\eta_t \text{ is in a fixed compact subset of } L^2(\to C[0,1]).\end{aligned} \tag{6.10}$$

REMARK 6.1: We should note that in applying the Gronwall Inequality for the estimation above we implicitly needed $\mathcal{H}(0) < \infty$, i.e., $\zeta = \xi_t\Big|_{t=0}$ must be in $L^2([0,1] \to \mathbb{R}^6)$. To get this initial datum, we consider (2.7) at $t=0$, so we are using the initial data $\overset{\circ}{\xi}, \overset{\circ}{\eta}, \overset{\circ}{\eta}_t$ with $\overset{\circ}{\eta}_t$ obtained from considering (2.4) at $t = 0$. This L^2 regularity of $\zeta(0, \cdot)$ is thus really a restriction on the admissible initial data $\overset{\circ}{\xi}, \overset{\circ}{\eta}$.

We may also remark that, although we have not pursued this here, the techniques used in [9] can also be adapted to this more general setting to weaken the regularity imposed above on the control $\nu(\cdot)$: we here required that ν' should be in $L^2(0,T)$ and a more careful analysis shows that it is sufficient for our purposes to require only that ν' be in $L^{4/3}(0,T)$ — and even that could be further weakened to permit jumps in ν at a finite number of specified times.

7. Optimal control

Finally, we consider (2.13) without the external imposition of compactness as in Examples 5.1 and 5.2, e.g., we consider

EXAMPLE 7.1: With T fixed in (2.13), find a control $\nu(\cdot)$ which minimizes the cost

$$\mathcal{J} = \int_0^T \left[|\nu(t)|^2 + |\nu'(t)|^2 \right] \, dt \tag{7.1}$$

subject to (5.1).

and wish to prove

THEOREM 7.1: *Under the hypotheses* (3.2), (4.1), (6.1), *and the consistency of the target set $\mathcal{S}$ (and with suitably regular initial data), the cost functional $\mathcal{J}$ in* (7.1) *attains its minimum: There exists an optimal control for Example 7.1.*

As for the previous examples, there is a minimizing sequence $\{\nu^k\}$ with a corresponding sequence of solutions $[\eta^k, \xi^k]$ of (2.13). We necessarily have $\{\nu^k\}$ bounded in $H^1(0,T)$ so, without loss of generality, we may now assume that $\nu^k \rightharpoonup \nu^\infty$ in $H^1(0,T)$ and then must show that there is a (possibly subsequential) limit $[\eta^\infty, \xi^\infty]$ of the solutions in a sense which permits us to conclude that $[\eta^\infty, \xi^\infty]$ satisfies (2.13) and gives the terminal condition (5.1) at T.

From our results in Sections 3, 4, 6, the hypotheses (3.2), (4.1), (6.1) ensure that $\{[\eta^k, \eta_t^k, \xi^k]\}$ uniformly avoid total compression, i.e., $\eta^k(t,s) \in \mathcal{A}_c$ with c independent of t, s, k, and will lie in a suitable fixed compact set so, possibly extracting a subsequence, we may assume that $\eta^k \to \eta^\infty$, $\eta_t^k \to \eta_t^\infty$, $\xi^k \to \xi^\infty$ as in (6.10). Since we have η^k, η_t^k pointwise L^∞-bounded, any growth condition on $\hat{\sigma}$ would be moot, so we can again apply Krasnosel'skiĭ's Theorem [7] to see that $\sigma^k := \hat{\sigma}(\eta^k, \eta_t^k) \to \hat{\sigma}(\eta^\infty, \eta_t^\infty) =: \sigma^\infty$ in, e.g., $L^2(\mathcal{Q})$. As in Section 5, this ensures that $[\eta^\infty, \xi^\infty]$ satisfies (2.13) and we have (5.1); in view of the weak lower semicontinuity of $\mathcal{J}$, this shows that ν^∞ is an optimal control.

Acknowledgments: We thank the organizers of the May, 2002, conference at Washington State University for the opportunity to participate in that conference and for this opportunity to contribute a paper dedicated to John

Cannon on the occasion of his 60th birthday and in honor of his own contributions to the areas of Partial Differential Equations, Distributed Parameter Control, and Inverse Problems.

The work of S.S. Antman was partially supported by grants from the NSF and the ARO.

References

[1] S.S. Antman, *Nonlinear Problems of Elasticity,* Springer-Verlag, New York (1995).

[2] S.S. Antman and T.I. Seidman, *Quasilinear hyperbolic-parabolic equations of one-dimensional viscoelasticity*, J. Diff. Eqns. **124**, pp. 132–185 (1996).

[3] S.S. Antman and T.I. Seidman, *The spatial motion of nonlinearly viscoelastic rods*, in preparation.

[4] J. P. Aubin, *Un théorème de compacité*, Comptes Rendus Acad. Sci. Paris **265**, pp. 5042–5045 (1963).

[5] C.M. Dafermos, *The mixed initial-boundary value problem for the equations of nonlinear one-dimensional viscoelasticity,* J. Diff. Eqs. **6**, pp. 71–86 (1969).

[6] D. French, S. Jensen, and T.I. Seidman, *A space-time finite element method for a class of nonlinear hyperbolic-parabolic equations*, Appl. Numer. Math. **31**, pp. 429–450 (1999).

[7] M.A. Krasnosel'skiĭ, *Topological Methods in the Theory of Nonlinear Integral Equations.* Pergamon, Oxford (1964).

[8] T.I. Seidman, *The transient semiconductor problem with generation terms,II,* in *Nonlinear Semigroups, PDE, and Attractors* (LNM #1394; T.E. Gill, W.W. Zachary, eds.), pp. 185–198, Springer-Verlag, New York (1989).

[9] T.I. Seidman and S.S. Antman, *Optimal control of a nonlinearly viscoelastic rod,* in *Control of Nonlinear Distributed Parameter Systems*, (G. Chen, I. Lasiecka, J. Zhou, eds.) pp. 273–283, Marcel Dekker, New York, (2001).

[10] J. Simon, *Compact sets in the space* $L^p(0,T;B)$*,* Ann. Mat. Pura Appl. **146** pp. 65–96 (1987).