Arch. Rational Mech. Anal. 155 (2000) 299–334
Digital Object Identifier (DOI) 10.1007/s002050000115

Weak Solutions for a Class of Nonlinear Systems of Viscoelasticity

Sophia Demoulini

Communicated by C. M. Dafermos

Abstract

The principal focus of the article is the construction of classical weak solutions of the initial value problem for a class of systems of viscoelasticity in arbitrary spatial dimension. The class of systems studied is large enough to incorporate certain requirements dictated by frame indifference and also has a structure which allows for a variational treatment of the time-discretized problem. Weak solutions for this system are constructed under certain monotonicity hypotheses and are shown to satisfy various *a priori* estimates, in particular giving improved regularity for the time derivative. Also measure-valued solutions are obtained under a uniform dissipation condition, which is much weaker than monotonicity. A special case of the viscoelastic system is the gradient flow of a non-convex potential, for which measure-valued solutions are here obtained, a new result in the vectorial case. Furthermore, in this setting it is possible to show that these measure-valued solutions satisfy a certain property which ensures they coincide with the classical weak solution when this exists, as for example in the convex case where existence and uniqueness are well known.

1. Introduction and statement of results

1.1. The equations

This article concerns the system of equations of viscoelasticity

$$u_{tt} = \nabla \cdot \Sigma(\nabla u, \nabla u_t) \;=\; \nabla \cdot \sigma(\nabla u) + \nabla \cdot \tau(\nabla u, \nabla u_t) \tag{1.1}$$

as well as that of the gradient flow,

$$v_t = \nabla \cdot \tau(\nabla v), \tag{1.2}$$

which can be viewed as a special case of (1.1) when $\sigma = 0$ and $\tau(F, \dot{F}) = \tilde{\tau}(\dot{F})$. At fixed time t, $u, v : \Omega \subset \mathbb{R}^n \to \mathbb{R}^m$, where Ω is an open, bounded set with Lipschitz boundary and $m, n \in \mathbb{N}$. Both systems are supplemented with a Dirichlet boundary condition

$$u(t, x) = 0, \quad v(t, x) = 0 \quad \forall t \geqq 0, \; x \in \partial\Omega, \tag{1.3}$$

as well as initial conditions,

$$\begin{aligned} u(0, x) &= u_0(x), \\ u_t(0, x) &= v_0(x), \end{aligned} \tag{1.4}$$

where $u_0, \; v_0 \in H_0^1(\Omega)$ in the case of (1.1), and

$$v(0, x) = v_0(x), \tag{1.5}$$

where $v_0 \in H_0^1(\Omega)$ in the case of (1.2). The relation between the total tensor Σ, the elastic tensor σ and the dissipative tensor τ is given through the definitions

$$\sigma(F) \equiv \Sigma(F, 0), \quad \tau(F, \dot{F}) = \Sigma(F, \dot{F}) - \Sigma(F, 0)$$

so that τ vanishes with the strain rate, $\tau(F, 0) = 0$. For (1.1) it is also assumed that there exists an energy function $\mathcal{W}$, not necessarily convex, with $\mathcal{W}(F, \dot{F}) = W(F) + G(F, \dot{F})$ where $\nabla_F W(F) = \sigma(F)$ and $\nabla_{\dot{F}} G(F, \dot{F}) = \tau(F, \dot{F})$. In the case of (1.2), it is assumed that G is a potential, $\tau(F) = \nabla_F G$. Assumptions made on the energy and stress tensors are discussed below and are given precisely in Section 1.2. It will be seen that monotonicity properties of σ are irrelevant for the existence of a classical weak solution of (1.1), while monotonicity of τ in the $\dot{F}$ variables is important.

Although here $m, n \in \mathbb{N}$ are not restricted, physically interesting is (1.1) with $n = 3$, $m = 1, 3$. When $n, m = 3$ this system comprises the equations of viscoelasticity. In certain cases, as in shearing, the equations degenerate to a single dependent variable so that u is scalar-valued. In the case of a single equation (with $n = 3, m = 1$), (1.1) represents antiplane shear motion of a viscoelastic solid.

Informal description of results. The theorems summarized here are stated precisely in Section 1.3. Firstly, without any monotonicity assumptions on σ, τ other than the dissipation condition (1.17) or (1.17′), existence of a Young measure solution of (1.1) is shown with the method of time-discretization. From the proof, without the use of a dissipation condition, there follows also the existence of a Young measure solution of the gradient flow. Next, I turn to the question of when these measure-valued solutions are *classical weak solutions,* that is, when are the Young measures obtained delta functions. In Theorem A it is shown that under a monotonicity assumption on τ which does not amount to full convexity of G, the constructed Young measure solution is the *unique, classical weak solution* to (1.1).

For the gradient flow (1.2) the same monotonicity condition amounts to strict convexity of G (which here depends only on $\dot{F}$) and in this case a classical weak solution is recovered directly from the regularization used to obtain measure-valued solutions. In fact it is shown in Theorem B that, for G strictly convex, *any measure-valued solution satisfying a pseudomonotonicity-type condition* (*an independence property*), *is the unique classical solution of* (1.2). Of course, this conclusion also follows from the one made about (1.1) in Theorem A; however, it merits a separate proof which I give, in order to illustrate that this pseudomonotonicity condition reflects the differences in type between (1.1) and (1.2). (The presence of the σ term changes significantly the type as well as the convergence properties of the approximate solutions obtained under the same regularization).

In the remainder of the introduction I discuss the background to the equations, assumptions made and their physical significance.

The gradient flow. The role of the monotonicity of τ in the existence of classical weak solutions is exemplified here: under the hypothesis of strict monotonicity of τ, as for example in the case of a strictly convex potential $\nabla G = \tau$, it is well known that classical weak solutions to the system (1.2) exist, while when τ is non-monotone such solutions do not exist in general. Indeed, in regions of backwards monotonicity the equation with initial conditions on the hyperplane $t = 0$ ceases to be well posed and the forward–backward nature of the system forces oscillations in approximating sequences.

In the absence of any monotonicity conditions, the best that can be expected in general is the existence of very weak, typically Young measure, solutions. Combining variational methods and the Young measure theory developed by Tartar [28] to treat conservation laws, time-discretization was first implemented in the context of non-monotone evolutionary equations by Kinderlehrer & Pedregal [20] for the scalar case of (1.2) to obtain Young measure solutions. In the scalar case a Young measure solution is unique if the independence property mentioned above holds and coincides with a classical solution when one exists (cf. [13]). In the vectorial case, however, independence is not immediately satisfied by construction, hence uniqueness follows only for the measure-valued solutions which satisfy it. This difference is linked to the fact that in the scalar case the relaxation of the non-convex energy is convex, equivalent to a monotonicity condition for the gradient, while in the vectorial case the relaxation is quasiconvex which seems not to imply any such condition.

In this connection there are certain open questions, namely, how to solve the gradient flow (1.2) under weaker convexity assumptions on G, such as polyconvexity or quasiconvexity. Under these more general conditions (or even under uniform strict quasiconvexity introduced by Evans [17]), the existence of weak solutions of the gradient flow is apparently unknown. In the same vein is the result of a classical weak solution of (1.2) under the assumptions of *quasimonotonicity* in Zeidler [29], Hamburger [19] and Landes [22], extending the proof of Zhang [30] and the stronger assumption of *pseudomonotonicity* introduced by J.-L. Lions [23, Chapter 2.4]).

Weak solutions of elasticity. Before discussing viscoelasticity, it may be useful to make a comparison with the equation of elastodynamics,

$$u_{tt} = \nabla \cdot \sigma(\nabla u), \tag{1.6}$$

with $\sigma = \nabla_F W$ non-monotone. Existence of a Young measure solution was shown in [14] for the scalar equation $m = 1$ via the time-discretization method. The difference in type between the three nonlinear equations, (1.2), (1.6) and (1.1) is reflected in this method, for example, in the way compactness of the approximate solutions is obtained in order to yield a measure-valued solution as well as in the properties of such solutions. The question arises as to what extent this approach yields classical weak solutions when such exist.

In this connection the study of elastodynamics is undertaken in [15,16]: in the first stage we consider (1.6) with $n = m = 1$ with σ strictly monotone. It is shown in [15] that the measure-valued solutions constructed by time-discretization satisfy *entropy inequalities* and that DiPerna's theorem (1983) applies, so we can deduce that these are classical weak solutions.

The second stage of the analysis in [16] concerns (1.6) for $n = m = 3$ with W polyconvex and is motivated by recent results of Qin (1998) and DAFERMOS [12, Chapter 5] on symmetrizing these equations by embedding them into an enlarged system of conservations laws which includes the evolution of null Lagrangians. Approximations are produced by constrained minimization and it is shown that in the limit the elasticity equation is satisfied in a measure-valued sense while these constraints hold in the classical weak sense.

The equations of viscoelasticity. Much less has been established for dynamical viscoelasticity (1.1). When τ depends nonlinearly on $\dot{F}$, existence has been addressed mostly for $n = m = 1$. In this case, without any monotonicity assumptions on σ and strict monotonicity of τ only in the u_{tx} variable and dissipative τ, (1.17), DAFERMOS [11] showed the existence of a Holder solution and investigated its time-asymptotic properties (also as these pertain to boundary conditions). Incorporating the constraints of infinite energy for total compression and local invertibility of u discussed below, ANTMAN & SEIDMAN [4] showed the existence of classical weak solutions for $n = m = 1$ under (1.17). Under strict monotonicity of Σ in both u_x and u_{tx}, ANTMAN & KOCH [6] found classical time-periodic solutions of the one-dimensional nonlinear equation using Hopf bifurcation. In three dimensions with dissipative τ, POTIER-FERRY [26] showed the asymptotic stability of the equilibrium in $W^{2,p}$.

The special case of (1.1) with $\tau(F, \dot{F}) = \dot{F}$ gives rise to the semilinear case where the strictly monotone dependence on the highest derivative term gives the equation its (forward) parabolic character. This case has been the subject of significantly more literature, [1,2,25,27,18] among others, where the existence of weak solutions, asymptotic analysis and stability have been investigated. In one dimension this example is physically realistic; however, in many dimensions it conflicts with the requirement of *frame indifference*.

Frame indifference and other constraints when $n = m$. Physical considerations delimit the *constitutive equation*, or functional form, of the energies W, G and stress tensors σ, τ. Frame indifference is the requirement that when a rigid rotation, possibly time-dependent, is superimposed on the motion u, then scalar quantities like the energy remain *invariant*. In other words this axiom states that under the action of $SO(n)$ on invertible matrices $\mathrm{Mat}^+(n)$ given by $(Q, F) \mapsto QF$, the set of frame indifferent energies and stress tensors are those which are *fixed points* under the induced action of $SO(n)$ on the functions on $\mathrm{Mat}^+(n)$ and their derivatives (tensorial quantities like stress tensors vary covariantly). In the case of elasticity this means $W(QF) = W(F)$, equivalently, $\sigma(QF) = Q\sigma(F)$ for all rotations Q and invertible F.

For viscoelasticity this condition implies that for all Q, F as above, $\dot{F} \in \mathrm{Mat}(n)$ and for all $\dot{Q} \in T_Q SO(n)$, i.e., with $\dot{Q}Q^* + Q\dot{Q}^* = 0$, the stress tensor satisfies

$$\Sigma(QF, \dot{Q}F + Q\dot{F}) = Q\Sigma(F, \dot{F}) \tag{1.7}$$

(which is equivalent to $\mathcal{W}(QF, \dot{Q}F + Q\dot{F}) = \mathcal{W}(F, \dot{F})$ when there is an energy function). Another form of (1.7) is obtained (see, e.g., [3, Chapter 12]) using the polar decomposition $F = RU$ for $F \in \mathrm{Mat}^+(n)$ with $R \in SO(n)$ and $U^* = U$, $U^2 = F^*F$. Then (1.7) implies $\Sigma(F, \dot{Q}F + Q\dot{F}) = R\Sigma(U, \dot{U})$ for all $\dot{Q}$, $\dot{R}$ as above and $\dot{U}$ determined by $\dot{F} = \dot{R}U + R\dot{U}$.

Mathematical consequences of frame indifference are analysed in, for example, [3, Chapter 12], [9, Chapters 3, 4], [12, Chapter 2], [24, Chapter 2]. A similar invariance requirement is that of *isotropy*. Two further constraints desirable on physical grounds are $\nabla u \in \mathrm{Mat}^+(n)$ for which $\det \nabla u > 0$, restricting to orientation preserving deformations, and the condition $W(F) \to +\infty$ as $\det F \to 0+$, reflecting that infinite energy is required for total compression. These significantly constrain the constitutive equations for W and Σ (as well as for higher order tensors) and are hard to handle: for example, $\det \nabla u > 0$ is not a weakly closed condition in Sobolev spaces.

A theorem of Noll (1958) states that a tensor Σ is frame indifferent if and only if there is a symmetric tensor S such that $\sigma(F) = R\sigma(U) = FS(U)$ for all $F \in \mathrm{Mat}^+(n)$ with $F = RU$ and R, U as above. For viscoelasticity this condition states (cf. [3, Chapter 12]) that there exists a symmetric tensor S such that

$$\tau(F, \dot{F}) = FS(U, \dot{U}). \tag{1.8}$$

Based on this equivalent condition, Antman [5] showed that certain constitutive equations for τ with *affine* dependence on $\dot{F}$ (e.g., all forms $\tau(x, F, \dot{F}) = a(x)\dot{F}$ or $\tau(x, F, \dot{F}) = E_F(x, F)\dot{F}$) conflict with (1.8) and hence with frame indifference.

An important aspect is the implication of frame indifference for the *convexity* properties of W and G. Although convexity does not directly conflict with frame indifference, these two conditions together are physically too restricted: a theorem of Coleman and Noll (1959) for elastostatics (see [10, Section 8, Theorem 2]) implies that frame indifference and convexity of the elastic energy W preclude stability under compression stresses (in a precisely given sense), stability being a postulate of elasticity theory. Furthermore, convexity is incompatible with the property that W become singular on singular matrices, cf. [9, Theorem 4.8-1]. It follows that W or $\mathcal{W}$ cannot be convex.

Morrey's theorem (1952) links the weak lower semicontinuity and quasiconvexity under conditions on W disallowing singularities. In this regard Ball in [7] established that the condition of *polyconvexity* on W is sufficient for the conclusion of Morrey's theorem to hold under conditions sufficiently weak to allow singular energies and treat the constraint $\det \nabla u > 0$.

Turning now to properties of $G(F, \cdot)$, I have made two assumptions to obtain classical weak solutions, namely (1.11) and (1.18). To justify the former in the context of frame indifference, Dafermos provided an example which is presented in Appendix B. Ball recently observed that the strict convexity of $G(F, \cdot)$ implied by (1.18) conflicts with (1.8) (see Appendix B) and thus is too stringent for frame indifference (except when $m = n = 1$, and indeed strict monotonicity is assumed in [11] and [4]). Nevertheless, non-strict convexity of $G(F, \cdot)$ is compatible with (1.8) and this is also proved in Appendix B. Therefore, it would be desirable to replace the assumption (1.18) with the weaker assumption of convexity

$$((\tau(F, \dot{F}) - \tau(H, \dot{H})) \cdot (\dot{F} - \dot{H}) \geqq 0, \tag{1.9}$$

although I am not able to circumvent (1.18) in the present framework.

1.2. Assumptions

1.2.1. Assumptions for viscoelasticity. We consider the class of stress tensors for $\Sigma(F, \dot{F}) = \sigma(F) + \tau(F, \dot{F})$ such that there exist $W \in C^1(\mathrm{Mat}(m \times n))$, $G \in C^1(\mathrm{Mat}(m \times n), \mathrm{Mat}(m \times n))$ with

$$\sigma(F) = \partial_F W(F), \tag{1.10}$$

$$\tau(F, \dot{F}) = \partial_{\dot{F}} G(F, \dot{F}), \tag{1.11}$$

and $G(F, 0) = \tau(F, 0) = 0$ for all F. It is assumed that σ and $\tau(F, \cdot)$ are Lipschitz continuous with constants L_σ for σ, and L_τ, uniformly in F for τ. These two Lipschitz conditions are used to derive the estimates in Section 2 and are also essential in proving that a classical weak solution exists (Theorem 3.3).

For every $F, \dot{F} \in \mathrm{Mat}(m \times n)$ we assume the following growth conditions

$$k(|F|^2 - 1)^+ \leqq W(F) \leqq K(|F|^2 + 1), \tag{1.12}$$

$$c(|\dot{F}|^2 - 1)^+ \leqq G(F, \dot{F}) \leqq C(|\dot{F}|^2 + |F|^2 + 1), \tag{1.13}$$

$$\sigma(F) \leqq s|F|, \tag{1.14}$$

$$\begin{aligned} -\tilde{m}|\dot{F}| &\leqq \min\{G_F(F, \dot{F}), \tau(F, \dot{F}\} \\ &\leqq \max\{G_F(F, \dot{F}), \tau(F, \dot{F})\} \leqq m(|F| + |\dot{F}| + 1), \end{aligned} \tag{1.15}$$

where s and $c < C, k < K, \tilde{m} < m$ are positive constants independent of $F, \dot{F}$. It will be seen that if τ satisfies, instead of (1.15), the more restrictive condition

$$|\tau(F, \dot{F})| \leqq m|\dot{F}|, \tag{1.16}$$

then certain estimates on the derivatives of the solution become independent of time (cf. Lemma 2.6). Note that G^{qc}, the quasiconvex envelope of G *in the variables* $\dot{F}$, also satisfies (1.13).

To obtain measure-valued solutions in Theorem A(i), τ is assumed to be *uniformly dissipative,*

$$\tau(F, \dot{F}) \cdot \dot{F} \geqq \gamma|\dot{F}|^2, \tag{1.17}$$

with $\gamma > 0$, (equivalently, $G(F, \cdot)$ is strictly convex in $\dot{F}$ at $(F, 0)$ uniformly in F). This is used to derive the energy estimate in Lemma 2.1. It will be seen in the proof that it suffices to impose (1.17) on $\theta = \nabla_{\dot{F}} G^{qc}$, where G^{qc} is the quasiconvex envelope of G *in the variables* $\dot{F}$. Moreover, it will be seen in the proof that uniform dissipation can be generalized to

$$\tau(F, \dot{F}) \cdot \dot{F} \geqq \gamma |\dot{F}|^2 - \delta |F|^2 \tag{1.17'}$$

with $\gamma, \delta > 0$, at the expense of the estimates in Lemma 2.4 being valid on finite but arbitrary time intervals (cfc Corollary 2.5).

The dissipation condition is strengthened to *uniform strict monotonicity* in the variables $\dot{F}$ in order to obtain classical weak solutions in Theorem A(ii),

$$((\tau(F, \dot{F}) - \tau(H, \dot{H})) \cdot (\dot{F} - \dot{H}) \geqq \kappa |\dot{F} - \dot{H}|^2 - l |F - H|^2 \tag{1.18}$$

with κ, l positive constants (depending only on τ). The compatibility of this condition with frame indifference is discussed in the introduction and Appendix B. It will be seen that the monotonicity of σ is not relevant when σ is Lipschitz.

1.2.2. Assumptions for the gradient flow. We assume there exists a potential $G \in C^1(\mathrm{Mat}(m \times n))$ with $\tau = \nabla_F G$ continuous (not necessarily Lipschitz) and such that

$$(k|F|^2 - \tilde{k})^+ \leqq G(F) \leqq K|F|^2 + \tilde{K} \tag{1.19}$$

$$\tau(F) \leqq m(|F| + 1). \tag{1.20}$$

With these conditions, the existence of a measure-valued solutions follows in Theorem B(i). A classical weak solution is recovered in Theorem B(ii) under the additional assumption

$$(\tau(A) - \tau(B)) \cdot (A - B) > 0 \quad \text{for all } A, B \in \mathrm{Mat}(m \times n) \text{ unless } A = B. \tag{1.21}$$

Remark 1.1. The use of the L^2 setting in either of these two problems is most likely not essential and it is expected that the present results remain valid (with the obvious modifications) if p-polynomial growth is assumed instead (with duality between $W^{1,p}$ and $W^{-1,p'}$). But for convenience I restrict to quadratic growth. Also notice that under a change of variable (1.1) is formally equivalent to the first order system

$$w_t = \nabla v,$$
$$v_t = \nabla \cdot \Sigma(w, \nabla v) = \nabla \cdot \Sigma(w, w_t).$$

The validity of such a change of variable hinges on the regularity of the functions involved but is not relevant in the analysis following and thus not pursued.

Notation. As is customary, H^p is the Sobolev space $W^{p,2}$. Spaces of the form $Y(I, X(\Omega))$, with I an interval or all of $\mathbb{R}^+$ and Y, X Banach spaces (usually Sobolev or L^q), are often abbreviated by $Y(X)$, e.g., $L^\infty(H_0^1(\Omega))$. Using $Q_\infty = \mathbb{R}^+ \times \Omega$, notation such as $H^1_{\mathrm{loc}}(Q_\infty) \cap H^1(L^2)$ is clear. The Banach space of continuous, p-growth functions on $\mathbb{R}^n$ or $\mathbb{R}^{m\times n}$, under the natural norm, is denoted by $\mathcal{E}^p$, and by $\mathcal{E}_0^p$ its separable subspace (under the same norm) of functions f such that $f(A) = o(1+|A|^p)$. Here the spaces $\mathcal{E}^2$ and $\mathcal{E}_0^2$ will be used. For $a,\ b \in \mathbb{R}^n$ the Euclidean dot product is written simply by ab and for $A,\ B \in \mathrm{Mat}(m \times n)$ the dot product is written $A \cdot B = A : B \equiv \mathrm{tr}(B^* A)$. By G^{qc} is denoted the quasiconvex envelope of $(F, \dot{F}) \mapsto G(F, \dot{F})$ in the $\dot{F}$ variables as well as that of $F \mapsto G(F)$ (in F). The operator ∇ is differentiation in the x variables (unless notation such as $\nabla_{t,x}$ is used). When not specified, $\|\cdot\|$ stands for the spatial norm in $L^2(\Omega)$.

On the space of maps $Q_\infty \to M$ a Young measure is a probability measure on the product space, $\boldsymbol{\nu} \in \mathcal{P}(Q_\infty \times M)$ with marginal on Q_∞ the Lebesgue measure, that is, $\boldsymbol{\nu}(S \times M) = \mathcal{L}^{1+n}(S)$ for all measurable sets $S \subset Q_\infty$. It can be also viewed as a parametrized family of probability measures $\boldsymbol{\nu} : (t, x) \mapsto \nu_{t,x}$ where for a.e. $(t, x) \in Q_\infty$ the measure $\nu_{t,x}$ is in $\mathcal{P}(M)$, the space of probability measures on the target space. For brevity $\mathbb{R}^{m\times n}$ is written in place of $\mathrm{Mat}(m \times n)$. In the present context $M = \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n}$ or $M = \mathbb{R}^{m\times n}$ and for maps f on M (real-valued or otherwise) the distributional notation $\langle \boldsymbol{\nu}, f\rangle$ represents the integral $\int_M f\, d\boldsymbol{\nu}$; $\langle \boldsymbol{\nu}, \alpha\rangle$ stands for the integral of the identity, also written as $\langle \boldsymbol{\nu}, \mathrm{id}\rangle$. Below appear essentially three types of Young measures, $\boldsymbol{\mu}$ generated by sequences of spatial gradients, $\boldsymbol{\nu}$ generated by their time derivatives and $\boldsymbol{\xi}$ generated by product sequences of such gradients, with marginals $\boldsymbol{\mu}$, $\boldsymbol{\nu}$. For clarity the integration variables $\alpha,\ \beta$ sometimes appear explicitly to distinguish the two arguments of $\boldsymbol{\xi}$, so that $\langle \boldsymbol{\xi}, f(\alpha)\rangle = \langle \boldsymbol{\mu}, f\rangle$ and $\langle \boldsymbol{\xi}, f(\beta)\rangle = \langle \boldsymbol{\nu}, f\rangle$. The theory of gradient Young measures was developed in [21].

1.3. Statement of the main theorems

Theorem A (The system of viscoelasticity).

(i) Existence of Young measure solutions. *Under the assumptions on W, G, σ, τ in Section 1.2.1, except* (1.18)*, and for initial data $u_0,\ v_0 \in H_0^1(\Omega)$ there exists $u \in H^2_{\mathrm{loc}}(L^2) \cap W^{1,\infty}_{\mathrm{loc}}(H_0^1) \cap L^\infty(H_0^1)$ with $\nabla u_t \in L^2(Q_\infty)$ and a Young measure $\boldsymbol{\xi} \in \mathcal{P}(Q_\infty \times \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n})$ with marginals $\boldsymbol{\mu},\ \boldsymbol{\nu} \in \mathcal{P}(Q_\infty \times \mathbb{R}^{m\times n})$ respectively on each space $\mathbb{R}^{m\times n}$ such that* (1.1) *is satisfied in the sense that for all $\zeta \in L^2_{\mathrm{loc}}(\mathbb{R}^+, H_0^1(\Omega))$,*

$$\int_0^T \int_\Omega (\ \langle \boldsymbol{\xi}, \Sigma\rangle \cdot \nabla\zeta + u_{tt}\zeta\ )\ dx\, dt = 0. \tag{1.22}$$

Equivalently, $\nabla \cdot \langle \boldsymbol{\xi}, \Sigma\rangle = u_{tt}$ in $L^2_{\mathrm{loc}}(\mathbb{R}^+, H^{-1}(\Omega))$ where $\langle \boldsymbol{\xi}, \Sigma\rangle = \langle \boldsymbol{\mu}, \sigma\rangle + \langle \boldsymbol{\xi}, \tau\rangle$. In addition,

$$\nabla u_t = \langle \boldsymbol{\nu}, \mathrm{id}\rangle = \partial_t \langle \boldsymbol{\mu}, \mathrm{id}\rangle\,, \quad (t, x) \in Q_\infty \textit{ a.e.} \tag{1.23}$$

$$\mathrm{supp}\ \xi_{t,x} \subseteq \{G(F, \cdot) = G^{qc}(F, \cdot)|_{F=\nabla u(t,x)}\}, \quad (t, x) \in Q_\infty \textit{ a.e.} \tag{1.24}$$

As $t \to 0^+$, $(u(t), u_t(t)) \longrightarrow (u_0, v_0)$ strongly in $L^2(\Omega)$, so that the initial data are attained in $C_{\mathrm{loc}}(\mathbb{R}^+, H_0^1(\Omega))$.

The t-uniform and L^2 bounds hold:

$$\|\nabla u(t)\|_{L^\infty(\mathbb{R}^+, L^2(\Omega))} + \|\nabla u_t\|_{L^2(Q_\infty)} \leqq C \tag{1.25}$$

and for every $T > 0$,

$$\sup_{t\in[0,T]} \|\nabla u_t(t)\|_{L^2(\Omega)} + \|u_{tt}\|_{L^2(Q_T)} \leqq C_T, \tag{1.26}$$

where the constants C *and* C_T *depend only on* $\|u_0\|_{H_0^1(\Omega)}$ *and* $\|v_0\|_{H_0^1(\Omega)}$ *and* C_T *depends also on* T. *Under the more restrictive condition* (1.16) *in place of* (1.15), (1.26) *holds with* C_T *independent of* T. *Similarly, if* (1.17′) *replaces* (1.17), (1.25) *holds on arbitrary intervals* $[0, T]$ *but with* C *depending* T.

(ii) Existence and uniqueness of a classical weak solution. *Assuming in addition* (1.18)*, there exists a unique classical weak solution* u *to* (1.1) *which satisfies*

$$\int_0^T \int_\Omega \left((\sigma(\nabla u) + \tau(\nabla u, \nabla u_t)) \cdot \nabla\zeta + u_{tt}\zeta \right) dx\, dt = 0$$

for all ζ *as above. This solution can be obtained by the approximation which is used to prove existence in Theorem A(i). In fact, given any measure-valued solution* $(u, \boldsymbol{\xi})$ *with* $u \in H^2_{\mathrm{loc}}(L^2) \cap H^1_{\mathrm{loc}}(H^1_0)$ *and* $\boldsymbol{\xi} \in \mathcal{P}(Q_\infty \times \mathbb{R}^{m\times n})$ *for which* (1.22)–(1.24) *hold and such that* $\boldsymbol{\xi}$ *satisfies*

$$\int_\Omega \langle \boldsymbol{\xi}, \tau(\alpha, \beta)\rangle \cdot \langle \boldsymbol{\nu}, \beta\rangle \, dx \geqq \int_\Omega \langle \boldsymbol{\xi}, \tau(\alpha, \beta) \cdot \beta\rangle \, dx, \tag{1.27}$$

then u *coincides with the unique classical weak solution.*

Furthermore, the initial value problem (1.1)–(1.4) *is well posed in the sense that if two solutions* $u, \tilde{u}$ *correspond to pairs of initial data* (u_0, v_0), $(\tilde{u}_0, \tilde{v}_0)$ *then for every* $T > 0$ *and* $0 \leqq t \leqq T$,

$$\|u_t - \tilde{u}_t\|^2(t) + \|\nabla u - \nabla\tilde{u}\|^2(t) \leqq k(T)\left(\|v_0 - \tilde{v}_0\|^2 + \|\nabla u_0 - \nabla\tilde{u}_0\|^2\right), \tag{1.28}$$

where $\|\cdot\|$ *is the* $L^2(\Omega)$ *norm.*

Theorem B (The gradient flow).

(i) Existence of Young measure solutions. *Under the assumptions on* G, τ *in Section 1.2.2, except* (1.21) *and for initial data* $v_0 \in H_0^1(\Omega)$ *there exists* $v \in H^1_{\mathrm{loc}}(Q_\infty) \cap H^1_{\mathrm{loc}}(\mathbb{R}^+, L^2(\Omega)) \cap L^\infty(\mathbb{R}^+, H^1_0(\Omega))$ *and a Young measure* $\boldsymbol{\nu} : (t, x) \in Q_\infty \mapsto \nu_{t,x} \in \mathcal{P}(\mathbb{R}^{m\times n})$ *such that*

$$\int_0^T \int_\Omega \left(\langle \boldsymbol{\nu}, \tau\rangle \cdot \nabla\zeta + v_t\zeta \right) dx\, dt = 0, \tag{1.29}$$

$$\nabla v = \langle \boldsymbol{\nu}, id\rangle \quad (t, x) \in Q_\infty \text{ a.e.}, \tag{1.30}$$

$$\operatorname{supp} \boldsymbol{\nu} \subseteq \{G = G^{qc}\}, \tag{1.31}$$

and such that the independence property holds:

$$\int_\Omega \langle \boldsymbol{\nu}, \tau\rangle \cdot \langle \boldsymbol{\nu}, \mathrm{id}\rangle \, dx = \int_\Omega \langle \boldsymbol{\nu}, \tau \cdot id\rangle \, dx \quad \forall t \geqq 0. \tag{1.32}$$

The initial data are attained in the sense that $v(t) \longrightarrow v_0$ *strongly in* $L^2(\Omega)$ *as* $t \to 0^+$. *Also the global* t*-uniform bounds hold:*

$$\|\nabla v(t)\|_{L^\infty(\mathbb{R}^+, L^2(\Omega))} + \|v_t\|_{L^2(Q_\infty)} \leqq C(\|\nabla v_0\|_{L^2(\Omega)}). \tag{1.33}$$

(ii) Recovery of classical weak solutions. *Assuming in addition* (1.21)*, the Young measure solution which is constructed in the proof of* (i)*, or any measure-valued solution* $(v, \boldsymbol{\nu})$ *with* $v \in H^1_{\mathrm{loc}}(Q_\infty)$ *and* $\boldsymbol{\nu} \in \mathcal{P}(Q_\infty \times \mathbb{R}^{m\times n})$ *satisfying* (1.29)–(1.31) *as well as the independence property* (1.32) *is the unique classical weak solution of* (1.2) *so that*

$$\int_0^T \int_\Omega (\ \tau(\nabla v) \cdot \nabla \zeta + v_t \zeta\ )\ dx\, dt = 0$$

for all ζ *as above.*

Remark 1.2. Property (1.32) asserts the *independence* of τ and the identity relative to $\boldsymbol{\nu}$ and is related to the property of *pseudomonotonicity* (cf. ZEIDLER [29]) mentioned in the introduction. Notice that any classical weak solution of the gradient flow satisfies it and thus coincides with the Young measure solution provided $\nabla u \subseteq \{G = G^{qc}\}$, (t, x) a.e. In the scalar case the measure-valued solution in Theorem B(i) is unique within the general class of measure-valued solutions satisfying (1.29)–(1.31) and (1.32) ([13, Theorem 4.2]). This uniqueness cannot be recovered (at least in this framework) in the vectorial case. Finally notice that (1.33) is global in t without restricting the growth condition of τ (1.20).

The proofs of Theorems A and B appear in Section 4. The statements of existence, uniqueness and other properties asserted in the theorems are proved separately in Sections 2 and 3.

1.4. A general lemma

The following compactness lemma states that when partial derivatives of Sobolev functions converge weakly, then some regularity can be gained in the corresponding variables. It will be used to show the attainment of the initial data continuously in time.

Lemma 1.3. *Assume* $U \subset \mathbb{R}^n$ *is an open set (not necessarily bounded) with Lipschitz boundary,* $p > 1$*, and that the sequence* $f^k \in W^{1,p}_{\mathrm{loc}}(\mathbb{R} \times U; \mathbb{R}^m)$ *converges* $f^k \rightharpoonup f$ *weakly in* $W^{1,p}_{\mathrm{loc}}(\mathbb{R} \times U)$ *and that for some* $t_0 \in \mathbb{R}$*,* $f^k(t_0) \longrightarrow f(t_0)$ *strongly in* $L^p(U)$*. Then* $f^k \longrightarrow f$ *strongly in* $C_{\mathrm{loc}}(\mathbb{R}, L^p(U))$.

Proof. By compact imbedding, the sequence $\left(f^k\right)_{k\geq 1}$ converges to f in $L^p_{\mathrm{loc}}(\mathbb{R} \times U)$. Fix $\omega \subset\subset U$ (unless U is bounded), let $T > t_0$ and set $W_t \equiv [t_0, T] \times \omega$. Given $\varepsilon > 0$ there is $K(\varepsilon) > 0$ such that

$$\begin{aligned}
\sup_{t_0 \leqq t \leqq T} & \|f^k(t) - f^l(t)\|_{L^p(\omega)} - \|f^k(t_0) - f^l(t_0)\|_{L^p(\omega)} \\
&= p \sup_{t_0 \leqq t \leqq T} \int_{t_0}^t \int_\omega |f^k - f^l|^{p-2}(f^k - f^l)(f_t^k - f_t^l)\, dx\, ds \\
&\leqq p \sup_{t_0 \leqq t \leqq T} \|f^k - f^l\|_{L^p(W_t)}^{\frac{p}{p'}} \|f_t^k - f_t^l\|_{L^p(W_t)} \\
&\leqq p\, \|f^k - f^l\|_{L^p(W_T)}^{\frac{p}{p'}} \left(\|f_t^k\|_{L^p(W_T)} + \|f_t^l\|_{L^p(W_T)}\right) \\
&\leqq \varepsilon
\end{aligned}$$

for all $k, l \geqq K(\varepsilon)$, using the strong convergence of f^k locally in L^p and the boundedness of the derivatives f_t^k. (The equality is valid as $\|f^k - f^l\|_{L^p(\Omega)}^p(t)$ is weakly differentiable in t.) Thus $(f^k)_{k\geqq 1}$ is Cauchy in $C([t_0, T], L^p(U))$. □

2. Existence theory for measure-valued solutions

Step I: The discretization. We discretize (1.1) in the time variable, implicitly in the quantities u_{tt}, ∇u_t and explicitly in $\sigma(\nabla u)$: this choice is designed in order to obtain a good energy estimate in Lemma 2.1 and uniform bounds in Lemmas 2.4 and 2.6. Let $h > 0$ be the time step and define $u^{h,0} = u_0$, $v^{h,0} = v_0$ and $u^{h,-1} = u_0 - hv_0$, $v^{h,-1} = 0$. For $j = 1, 2, \ldots$ define $u^{h,j}$ to be the solutions of the following regularization:

$$\frac{v^{h,j} - v^{h,j-1}}{h} = \nabla \cdot \left(\sigma(\nabla u^{h,j-1}) + \tau(\nabla u^{h,j-1}, \nabla v^{h,j})\right), \tag{2.1}$$

where

$$v^{h,j} = \frac{u^{h,j} - u^{h,j-1}}{h} \quad \text{and} \quad \frac{v^{h,j} - v^{h,j-1}}{h} = \frac{u^{h,j} - 2u^{h,j-1} + u^{h,j-2}}{h^2}.$$

For fixed h, j, (2.1) is the Euler-Lagrange system corresponding to the functional $I_{h,j} : H_0^1(\Omega) \to \mathbb{R}$,

$$\begin{aligned}
I_{h,j}(z) = \int_\Omega \Bigg(& \nabla \cdot \sigma(\nabla u^{h,j-1}) z + hG\left(\nabla u^{h,j-1}, \frac{\nabla z - \nabla u^{h,j-1}}{h}\right) \\
& + \frac{(z - 2u^{h,j-1} + u^{h,j-2})^2}{2h^2} \Bigg)\, dx.
\end{aligned}$$

Consider the relaxed functional $I_{h,j}^{qc}$ obtained by replacing G in $I_{h,j}$ with its quasiconvex envelope in the $\dot{F}$ variables, G^{qc}. Then there exists $u^{h,j} \in H_0^1(\Omega)$ such that $\bar{I}_{h,j} \equiv \inf_{z \in H_0^1(\Omega)} I_{h,j}(z) = \inf_{z \in H_0^1(\Omega)} I_{h,j}^{qc}(z) = I_{h,j}^{qc}(u^{h,j})$. Let $(u^{h,j,k})_k$ be a minimizing sequence for $I_{h,j}$ and $I_{h,j}^{qc}$ so that $I_{h,j} = \lim_{k\to\infty} I_{h,j}(u^{h,j,k}) =$

$\lim_{k\to\infty} I^{qc}_{h,j}(u^{h,j,k})$. By the growth conditions (1.12) and (1.13), $(u^{h,j,k})_k$ converges subsequentially to $u^{h,j}$ weakly in $H^1_0(\Omega)$, strongly in $L^2(\Omega)$ and pointwise for a.e. x in Ω.

These sequences are bounded uniformly in h, j as is shown in Lemma 2.4. Since $\int_\Omega \tilde{G}(x, \nabla u^{h,j,k}(x))dx \longrightarrow \int_\Omega \tilde{G}(x, \nabla u^{h,j}(x))dx$ where $\tilde{G}(x, \nabla u^{h,j}(x)) = G^{qc}(\nabla u^{h,j-1}, \frac{\nabla u^{h,j}-\nabla u^{h,j-1}}{h})(x)$ and $\tilde{G}(x,\cdot) \in \mathcal{E}^2$ is quasiconvex, the sequence $G(x, \nabla u^{h,j,k}(x))$ is weakly precompact in L^1. Consider the $W^{1,2}$-gradient Young measures $\boldsymbol{\nu}^{h,j} = (\nu^{h,j}_x)_{x\in\Omega} \in \mathcal{P}(\mathbb{R}^{m\times n})$ generated by $(\nabla v^{h,j,k} \equiv h^{-1}(\nabla u^{h,j,k} - \nabla u^{h,j-1}))_{k\geqq 1}$ and set for all $\boldsymbol{\nu}^{h,0} \equiv \delta_{\nabla v_0}$ for all h. Then

$$f(\nabla v^{h,j,k}) \rightharpoonup \left\langle \boldsymbol{\nu}^{h,j}, f\right\rangle \equiv \int_{\mathbb{R}^n} f(\alpha)\, d\boldsymbol{\nu}^{h,j}$$

weakly in L^1 for all $f \in \mathcal{E}^2$ and weakly in L^2 for $f \in \mathcal{E}^1$. Then it follows that

$$\nabla v^{h,j} = \left\langle \boldsymbol{\nu}^{h,j}, \mathrm{id}\right\rangle = h^{-1}(\nabla u^{h,j} - \nabla u^{h,j-1}) \tag{2.2}$$

pointwise for a.e. $x \in \Omega$. By relaxation it follows that

$$\operatorname{supp} \boldsymbol{\nu}^{h,j} \subseteq \{G(F,\cdot) = G^{qc}(F,\cdot)|_{F=\nabla u^{h,j-1}}\}. \tag{2.3}$$

On the set $\{F : F = \nabla u^{h,j-1}\}$,

$$\nabla\cdot\left\langle \boldsymbol{\nu}^{h,j}, \tau(F,\cdot)\right\rangle = \nabla\cdot\left\langle \boldsymbol{\nu}^{h,j}, \theta(F,\cdot)\right\rangle = \nabla\cdot\theta(F, \nabla v^{h,j}), \tag{2.4}$$

where equality holds in $H^{-1}(\Omega)$.

To take the Gateaux derivative of $I^{qc}_{h,j}$ at the minimizer $u^{h,j}$ I apply a recent theorem of BALL, KIRCHHEIM & KRISTENSEN [8] which establishes that the differentiability properties of a function are inherited, and in fact improved, by its quasiconvex envelope. More precisely, in the present setting this theorem implies $G^{qc}(F,\cdot)$ is $\mathcal{C}^1$ and $\theta = \nabla_{\dot F} G^{qc}$ is locally Lipschitz in $\dot F$ uniformly in F (with the same constant L_τ). Thus the weak equation corresponding to the relaxed functional is

$$\begin{aligned} 0 &= \int_\Omega \left(\sigma(\nabla u^{h,j-1})\cdot\nabla\zeta + \theta(\nabla u^{h,j-1}, \nabla v^{h,j})\cdot\nabla\zeta + \frac{v^{h,j} - v^{h,j-1}}{h}\zeta\right) dx \\ &= \int_\Omega \left(\left(\sigma(\nabla u^{h,j-1}) + \left\langle \nu^{h,j}_x, \tau(\nabla u^{h,j-1},\cdot)\right\rangle\right)\cdot\nabla\zeta + \frac{v^{h,j} - v^{h,j-1}}{h}\zeta\right) dx \end{aligned}$$

for all $\zeta \in H^1_0(\Omega)$; it is in this sense that (2.1) holds in H^{-1}. (The latter equation is obtained by considering

$$\int_\Omega \left(\left(\sigma(\nabla u^{h,j-1}) + \tau(\nabla u^{h,j-1}, \nabla v^{h,j,k})\right)\cdot\nabla\zeta + \frac{v^{h,j,k} - v^{h,j-1}}{h}\zeta\right) dx.$$

and taking the limit in k.)

Step II: Estimates. The convergence of the approximate sequence relies on uniform estimates which imply the weak compactness of the sequence. The most significant estimate is the *discretized energy non-increase*. The classical analogue (A.1) which is shown in the appendix holds for smooth solutions of (1.1).

Lemma 2.1 (Energy norm estimates**).** *For each $h > 0$ and $j = 0, 1, \dots$ define*

$$E_{h,j} = \int_\Omega W(\nabla u^{h,j}) + \frac{(u^{h,j} - u^{h,j-1})^2}{2h^2}\, dx.$$

Then, under the assumptions of Theorem A with uniform dissipation (1.17), *for all* $0 \leqq h \leqq \frac{2\gamma}{L_\sigma} \equiv h_*$,

$$E_{h,j} - E_{h,j-1} \leqq 0, \tag{2.5}$$

$$\sup_{0<h\leqq h_*, j\in\mathbb{N}_0} \left\{ E_{h,j} + \sum_{j=0}^{\infty} h \int_\Omega \left(\left(\frac{\gamma - L_\sigma h}{2} \right) |\nabla v^{h,j}|^2 + \frac{(v^{h,j} - v^{h,j-1})^2}{2h} \right) dx \right\}$$
$$\leqq E_{h,0}. \tag{2.6}$$

If the dissipation condition (1.17′) *is assumed instead then there is $\alpha > 0$ such that*

$$E_{h,j} - (1 + \alpha h) E_{h,j-1} \leqq 0 \tag{2.7}$$

(and a version of (2.6) *holds with modifications resulting in Corollary 2.5).*

Remark 2.2 *(*Energy estimation in the presence of hyperbolic terms*).* In the hyperbolic case (1.6) in [14] the convergence of the time-discretized approximation hinged precisely on the observation that *time-discretization for second order dynamics in time leads to energy non-increase and hence uniform estimates exist*. This means that the discretization chosen is good if it is compatible with such an energy estimate. The same is true here for (1.1) owing to its hyperbolic term $\sigma(\nabla u)$ and is in contrast with the case of the gradient flow, where the estimates follow directly from the discretization.

Proof. The reason for the technical choice in (2.1) to discretize backwards (i.e., in $u^{h,j-1}$) in the σ term and forwards in the τ term will be clear in this proof. Temporarily suppress the dependence on h and reinstate it later. As above, let $v^{h,j} = \frac{1}{h}(u^{h,j} - u^{h,j-1})$. Then

$$\begin{aligned}
E_j - E_{j-1} &= \int_\Omega \left(W(\nabla u^j) - W(\nabla u^{j-1}) + \frac{(v^j)^2 - (v^{j-1})^2}{2} \right) dx \\
&= \int_\Omega \left(\int_0^1 \frac{d}{ds} W\left(s\nabla u^j + (1-s)\nabla u^{j-1} \right) ds \right. \\
&\quad \left. + v^j (v^j - v^{j-1}) - \frac{(v^j - v^{j-1})^2}{2} \right) dx \\
&= \int_\Omega \left(\int_0^1 \sigma(s\nabla u^j + (1-s)\nabla u^{j-1}) ds \right) \cdot h\nabla v^j \\
&\qquad - h\nabla v^j \cdot \left(\sigma(\nabla u^{j-1}) + \theta(\nabla u^{h,j-1}, \nabla v^j) \right) - \frac{1}{2}(v^j - v^{j-1})^2\, dx
\end{aligned}$$

$$= \int_\Omega \left(\int_0^1 \left(\sigma(s\nabla u^j + (1-s)\nabla u^{j-1}) - \sigma(\nabla u^{j-1}) \right) ds \right) \cdot h\nabla v^j$$

$$- h\theta(\nabla u^{h,j-1}, \nabla v^j) \cdot \nabla v^j - \frac{(v^j - v^{j-1})^2}{2}\, dx$$

$$\leqq \int_\Omega \left(\left(\int_0^1 s L_\sigma h |\nabla v^j| ds \right) \cdot h|\nabla v^j| - h\gamma |\nabla v^j|^2 - \frac{(v^j - v^{j-1})^2}{2} \right) dx$$

$$\leqq \int_\Omega \left(\left(\frac{L_\sigma h^2}{2} - h\gamma \right) |\nabla v^j|^2 - \frac{(v^j - v^{j-1})^2}{2} \right) dx, \tag{2.8}$$

where the Lipschitz continuity of σ was used in the first inequality, as well as (1.17) applied to θ together with the independence property (3.3) and (2.4):

$$\theta(a, \nabla v^j) \cdot \nabla v^j = \left\langle \boldsymbol{\nu}^{h,j}, \tau(a,\lambda) \cdot \lambda \right\rangle \geqq \gamma \left\langle \boldsymbol{\nu}^{h,j}, |\lambda|^2 \right\rangle \geqq \gamma |\nabla v^{h,j}|^2.$$

Hence for $h < \frac{2\gamma}{L_\sigma}$ (2.5), (2.6) follow from the lemma as well as the estimates in Remark 2.3 below.

If (1.17′) is assumed in place of (1.17), the first inequality in (2.8) is changed by the addition of the term $h \int_\Omega \delta |\nabla u^{h,j-1}|^2 dx$ on the right-hand side from which (2.7) follows. (A "local" version of (2.6) holds in this case where both the right-hand side as well as the telescoping terms are multiplied by a constant depending on hj. The useful consequence thereof is given in Corollary 2.5). □

Remark 2.3. From the proof transpires the following more precise estimate which gives the extent to which the estimate above diverges from the classical estimate. For all $0 \leqq h \leqq \frac{2\gamma}{L_\sigma}$,

$$0 \leqq \sum_{j=0}^{\infty} h \int_\Omega \left(\int_0^1 \left(\sigma(\nabla u^{h,j-1}) - \sigma(\nabla u^s) \right) ds \right.$$

$$\left. + \theta(\nabla u^{h,j-1}, \nabla v^{h,j}) \right) \cdot \nabla v^{h,j} dx$$

$$\leqq \sum_{j=0}^{\infty} h \int_\Omega \left(\frac{\gamma - L_\sigma h}{2} \right) |\nabla v^{h,j}|^2\, dx$$

$$\leqq \sum_{j=0}^{\infty} h \int_\Omega \left(\frac{\gamma - L_\sigma h}{2} \right) \left\langle \boldsymbol{\nu}^{h,j}, |\lambda|^2 \right\rangle dx \ < \infty,$$

where $\nabla u^s = s\nabla u^j + (1-s)\nabla u^{j-1}$ for $s \in [0,1]$. The latter inequality follows from the proof above by replacing $\theta(\nabla u^{h,j-1}, \nabla v^{h,j})) \cdot \nabla v^{h,j}$ with $\left\langle \boldsymbol{\nu}^{h,j}, \tau(\nabla u^{h,j-1}, \lambda) \cdot \lambda \right\rangle$, made possible by (2.4), (3.3) and by using the lower semicontinuity property $\|\nabla v^{h,j}\|_{L^2} \leqq \| \left\langle \boldsymbol{\nu}^{h,j}, |\lambda|^2 \right\rangle \|_{L^2}$.

The non-increase of the $E_{h,j}$ and the chosen discretization form the basis for the estimates below and thus of (1.25), (1.26).

Lemma 2.4 (Integral and uniform estimates I**).** *Under the hypotheses of Lemma 2.4 with uniform dissipation there exists a positive constant M such that for all $h \leqq h_* = \frac{2\gamma}{L_\sigma}$ the following integral estimates hold, where $\|\cdot\|$ represents the $L^2(\Omega)$ norm:*

$$\sum_{j=0}^{\infty}\left(h\|\nabla v^{h,j}\|^2 + \|v^{h,j} - v^{h,j-1}\|^2\right) \leqq M. \tag{2.9}$$

Also the uniform estimates hold:

$$\sup_{h\leqq h_*, j\in\mathbb{N}_0} \left\{\|\nabla u^{h,j}\| + \|v^{h,j} - v^{h,j-1}\|\right\} \leqq M. \tag{2.10}$$

Consequently, for all $h \leqq h_*$ the minimizing sequences can be taken to satisfy $\sup_{h,j,k}\left\{\|\nabla u^{h,j,k}\|\right\} \leqq M$. The minimal values of the functionals $I_{h,j} = I^{qc}_{h,j}$ are uniformly bounded. The Young measures $\boldsymbol{\nu}^{h,j}$ satisfy,

$$\sum_{j=0}^{\infty} h \int_\Omega \left(\frac{\gamma - L_\sigma h}{2}\right)\left\langle \boldsymbol{\nu}^{h,j}, |\lambda|^2\right\rangle dx < \infty. \tag{2.11}$$

When (1.17′) *is assumed, M depends on hj .*

Proof. The restriction on $h \leqq h_*$ applies only when (2.6) is used. Equation (2.9) follows immediately from (2.8) by summation, and $(2.10)^i$ is true since $E_{h,j} \leqq E_0$. (Notation such as $(2.10)^i$ is used to refer to the first term of (2.10) and so on.) From $(2.9)^{ii}$ also follows $(2.10)^{ii}$. Recall that $I_{h,j}(u^{h,j,k}) \longrightarrow I^{qc}_{h,j}(u^{h,j})$ and hence by (1.13) the minimizing sequences can be always assumed to be bounded in $H^1_0(\Omega)$ independently of h, j, k by the estimate

$$\begin{aligned}
\bar{I}_{h,j} &= I^{qc}_{h,j}(u^{h,j}) \leqq I^{qc}_{h,j}(u^{h,j-1}) \\
&= \int_\Omega \left(\sigma(\nabla u^{h,j-1})\cdot\nabla u^{h,j-1} + \frac{(u^{h,j-1} - u^{h,j-2})^2}{2h^2}\right) dx \\
&\leqq L_\sigma\|\nabla u^{h,j-1}\|^2_{L^2(\Omega)} + \sup_{h\leqq h_*, j} E^{h,j-1},
\end{aligned}$$

thus concluding $\sup_{h\leqq h^*,j} I_{h,j} \leqq M$ by $(2.10)^i$ and Lemma 2.1. The inequality (2.11) was explained in the proof of Lemma 2.1. □

Step III: Interpolation and convergence to a limiting solution. Introduce time-dependent approximating solutions of (1.1) by interpolating in t the discrete solutions obtained above, both piecewise constantly and continuously piecewise linearly in time. Let $\chi^{h,j}(t)$ be the characteristic function of the interval $[hj, h(j+1))$ and define for $t \geqq 0$ and x a.e. in Ω

$$\begin{aligned}
u^h(t) &= \sum_j \chi^{h,j}(t) u^{h,j+1}, \\
U^h(t) &= \sum_j \chi^{h,j}(t) \left(u^{h,j} + (t-hj) v^{h,j+1} \right), \\
v^h(t) &= \sum_j \chi^{h,j}(t) v^{h,j+1} \equiv U^h_t, \\
V^h(t) &= \sum_j \chi^{h,j}(t) \left(v^{h,j} + (t-hj) \frac{v^{h,j+1} - v^{h,j}}{h} \right), \\
z^h(t) &= \sum_j \chi^{h,j}(t)\, \frac{v^{h,j+1} - v^{h,j}}{h} \equiv V^h_t, \\
\boldsymbol{\nu}^h_{(t,x)} &= \sum_j \chi^{h,j}(t) \nu^{h,j}_x, \\
\boldsymbol{\mu}^h_{(t,x)} &= \delta_{\nabla U^h(t,x)}, \\
\boldsymbol{\xi}^h_{(t,x)} &= \boldsymbol{\mu}^h_{(t,x)} \times \boldsymbol{\nu}^h_{(t,x)}.
\end{aligned} \tag{2.12}$$

Also define $u^{h,k} = \sum_j \chi^{h,j} u^{h,j,k}$ and similarly $v^{h,k}$. By (2.3), supp $\boldsymbol{\nu}^h \subseteq \{G(F,\cdot) = G^{qc}(F,\cdot)|_{F=\nabla u^h}\}$ for a.e. (t,x), and by (2.2),

$$\left\langle \boldsymbol{\nu}^h, \mathrm{id} \right\rangle = \partial_t \left\langle \boldsymbol{\mu}^h, \mathrm{id} \right\rangle. \tag{2.13}$$

For $\zeta \in H^1_{\mathrm{loc}}(Q_\infty)$ such that $\zeta(t,\cdot) \in H^1_0(\Omega)$ for a.e. t the interpolates (2.12) satisfy the weak equation

$$\nabla \cdot \left(\sigma(\nabla u^h(t-h)) + \left\langle \boldsymbol{\nu}^h, \tau(\nabla u^h(t-h), \cdot) \right\rangle \right) = z^h \quad \text{in } H^{-1}_{\mathrm{loc}}(Q_\infty).$$

Equivalently for such ζ,

$$\int_0^T \int_\Omega \left(\left(\sigma(\nabla u^h(t-h)) + \left\langle \boldsymbol{\nu}^h, \tau(\nabla u^h(t-h), \cdot) \right\rangle \right) \cdot \nabla \zeta + z^h \zeta \right) dx\, dt = 0. \tag{2.14}$$

By relaxation it was shown above (cf. (2.4)) that $\nabla \cdot \theta(\nabla u^h(t-h), \nabla v^h) = \nabla \cdot \left\langle \boldsymbol{\nu}^h, \tau(\nabla u^h(t-h), \cdot) \right\rangle = \nabla \cdot \left\langle \boldsymbol{\nu}^h, \theta(\nabla u^h(t-h), \cdot) \right\rangle$ in H^{-1}_{loc} so that (2.14) can also be written in these terms. By Lebesgue differentiation on $(t-\varepsilon, t+\varepsilon)$, the time integrand in (2.14) vanishes for a.e. $t \geqq 0$ so that the equation is also satisfied pointwise a.e. in time.

Corollary 2.5 (Integral and uniform estimates I**).** *Under* (1.17) *Lemma 2.4 implies the uniform estimates* $\|\nabla u^h\|_{L^\infty(L^2(\Omega))} + \|\nabla U^h_t\|_{L^2(Q_\infty)} \leqq M$. *Under* (1.17′) *these depend on time, and for each* $T > 0$, $\|\nabla u^h\|_{L^\infty([0,T],L^2(\Omega))} + \|\nabla U^h_t\|_{L^2([0,T],L^2(\Omega))} \leqq M_T$.

The final estimates in the following lemma concern bounds for V^h, v^h, z^h. Their classical counterpart is (A.3) in the appendix.

Lemma 2.6 (Integral and uniform estimates II**).** *For each $T > 0$ there exists a constant M_T such that for all $h \leqq h_*$ the regularizing solutions satisfy*

$$\sup_{t\in[0,T]} \int_\Omega |\nabla v^h|^2(t,x)\,dx = \sup_{t\in[0,T]} \int_\Omega |\nabla V^h|^2(t,x)\,dx) < M_T, \quad (2.15)$$

$$\int_0^T \int_\Omega |z^h|^2(t,x)\,dx\,dt < M_T. \quad (2.16)$$

Consequently, the minimising sequences can be taken to satisfy $\sup_{t\in[0,T]} \|\nabla v^{h,k}\|(t) < M_T$ independently of h, k. The Young measure sequence $(\boldsymbol{\nu}^h)_{h>0}$ generated by the ∇v^h is bounded in the space $L^1_{\rm loc}(\mathcal{E}^{2'}) \cap L^2_{\rm loc}(\mathcal{E}^{1'}) \cap L^\infty(\Omega, \mathcal{P}(\mathbb{R}^{m\times n}))$, (in particular, $\|\boldsymbol{\nu}^h\|_{L^\infty(\mathcal{P}(\mathbb{R}^{m\times n}))} = 1$). Under the growth condition (1.16) *in place of* (1.15)*, M_T above is independent of T.*

Proof. A technical obstruction to the estimation here arises from the fact that G^{qc} may not have partial derivatives in the direction F (cf. [8, Proposition 5.4]) while a direct analogue of the classical case relies on the existence of such derivatives (see (A.3)). Thus the present argument is inevitably more technical and circumvents this difficulty by enabling use of the regularity of G instead; here this is achieved by considering the infimum of a functional related to $I_{h,j}$ in terms of the Young measure $\boldsymbol{\nu}^{h,j}$.

It follows from [8] that $\theta = G^{qc}_{\dot F}$ exists, satisfies the same growth as τ in (1.15) and $\theta(F,\cdot)$ is Lipschitz continuous with constant L_θ uniformly in F. Consider the functional for $z \in H^1_0(\Omega)$,

$$J_{h,j}(z) = \int_\Omega \left(\nabla\cdot\sigma(\nabla u^{h,j-1})z + G(\nabla u^{h,j-1}, \nabla z) + \frac{(z - v^{h,j-1})^2}{2h} \right) dx$$

(corresponding to the Euler-Lagrange system given in the remark in Section 1.2). Then,

$$I_{h,j}(hz + u^{h,j-1}) = h\ J_{h,j}(z) + \int_\Omega \nabla\cdot\sigma(\nabla u^{h,j-1})u^{h,j-1}\,dx.$$

Thus the relaxed functional $J^{qc}_{h,j}$ (defined similarly as $I^{qc}_{h,j}$) attains its infimum at $v^{h,j}$, which by Step I of Section 2 can be represented in terms of the gradient Young measure $\boldsymbol{\nu}^{h,j}$:

$$\begin{aligned} J^{qc}_{h,j}(v^{h,j}) &= \inf_{z\in H^1_0(\Omega)} J_{h,j}(z) \\ &= \int_\Omega \Bigg(\nabla\cdot\sigma(\nabla u^{h,j-1})v^{h,j} \\ &\qquad + \left\langle \boldsymbol{\nu}^{h,j}, G(\nabla u^{h,j-1},\cdot)\right\rangle + \frac{(v^{h,j} - v^{h,j-1})^2}{2h} \Bigg) dx \end{aligned}$$

$$= h^{-1}\left(I^{qc}_{h,j}(u^{h,j}) + \int_\Omega \sigma(\nabla u^{h,j-1}) \cdot \nabla u^{h,j-1}\, dx\right).$$

Since $v^{h,j}$ minimizes $J^{qc}_{h,j}$,

$$\begin{aligned}
J^{qc}_{h,j}(v^{h,j}) &\leqq J^{qc}_{h,j}(v^{h,j-1})\\
&= \int_\Omega \left(\nabla \cdot \sigma(\nabla u^{h,j-1}) v^{h,j-1} + G^{qc}(\nabla u^{h,j-1}, \nabla v^{h,j-1})\right) dx
\end{aligned}$$

(and by lower semicontinuity along $(v^{h,j-1,k})_k$)

$$\leqq \liminf_{k\to\infty} \int_\Omega \left(\nabla \cdot \sigma(\nabla u^{h,j-1}) v^{h,j-1,k} + G^{qc}(\nabla u^{h,j-1}, \nabla v^{h,j,k-1})\right) dx$$

(and since the limit infimum is the limit and $G^{qc}(a,\cdot) \leqq G(a,\cdot)$)

$$\leqq \int_\Omega \left(\nabla \cdot \sigma(\nabla u^{h,j-1}) v^{h,j-1} + \left\langle \boldsymbol{\nu}^{h,j-1}, G(\nabla u^{h,j-1}, \cdot)\right\rangle\right) dx$$

(and if the integral on the right-hand side is momentarily called $\Psi^{h,j}$)

$$\begin{aligned}
&\leqq \Psi^{h,j} + \int_\Omega \frac{(v^{h,j-1} - v^{h,j-2})^2}{2h}\, dx\\
&= J^{qc}_{h,j-1}(v^{h,j-1}) + \int_\Omega \left(\sigma(\nabla u^{h,j-1}) - \sigma(\nabla u^{h,j-2})\right) \cdot \nabla v^{h,j-1}\, dx\\
&\qquad + \int_\Omega \left(\left\langle \boldsymbol{\nu}^{h,j-1}, G(\nabla u^{h,j-1}, \lambda)\right\rangle - \left\langle \boldsymbol{\nu}^{h,j-1}, G(\nabla u^{h,j-2}, \lambda)\right\rangle\right) dx\\
&\equiv J^{qc}_{h,j-1}(v^{h,j-1}) + S_1^{h,j} + S_2^{h,j}. \qquad (2.17)
\end{aligned}$$

Summing (2.17) on both sides in $1 \leqq j \leqq j_h$ (where j_h will be chosen soon), gives a relationship with telescopes as

$$\begin{aligned}
&\sup_{j\leqq j_h} \int_\Omega G^{qc}(\nabla u^{h,j-1}, \nabla v^{h,j}) + \sum_{j=1}^{j_h} \int_\Omega \frac{(v^{h,j-1} - v^{h,j-2})^2}{2h} dx\\
&\qquad \leqq J_0 + \sum_{j=1}^{j_h} (S_1^{h,j} + S_2^{h,j}) + \left|\sup_{j\leqq j_h} \int_\Omega \nabla \cdot \sigma(\nabla u^{h,j-1}) v^{h,j}\, dx\right|.
\end{aligned}$$

Using (1.14) the last integral is bounded above by $\frac{1}{2\varepsilon}\|\nabla u^{h,j-1}\|^2 + \frac{\varepsilon}{2}\|\nabla v^{h,j}\|^2$ for some $\varepsilon > 0$ small, while by (1.13) the first integral on the left-hand side is bounded below by $c(\|\nabla v^{h,j}\|^2 - 1)^+$. Therefore,

$$\begin{aligned}
&\left(c - \frac{\varepsilon}{2}\right) \sup_{j\leqq j_h} \|\nabla v^{h,j}\|^2 + \sum_{j=1}^{j_h} \frac{\|v^{h,j-1} - v^{h,j-2}\|^2}{2h}\\
&\qquad \leqq J_0 + \sum_{j=1}^{j_h} (S_1^{h,j} + S_2^{h,j}) + \sup_{j\leqq j_h} \frac{1}{2\varepsilon} \|\nabla u^{h,j-1}\|^2. \qquad (2.18)
\end{aligned}$$

Now, given $T > 0$, fix $h > 0$ and set $j_h = [\frac{T}{h}]$, where $j_h \to \infty$ as $h \to 0$. It remains to estimate the sums of $S_1^{h,j}$, $S_2^{h,j}$ in terms of T. First,

$$S_1^h = \sum_{j=1}^{j_h} \int_\Omega \Big(\sigma(\nabla u^{h,j-1}) - \sigma(\nabla u^{h,j-2})\Big) \cdot \nabla v^{h,j-1}\, dx \leqq L_\sigma \sum_j^\infty h\|\nabla v^{h,j}\|^2 < M$$

for some constant M independent of h (and T), by $(2.9)^i$. Next, using (1.15)) on G_F,

$$S_2^h = \sum_{j=1}^{j_h} \int_\Omega \Big(\Big\langle \boldsymbol{\nu}^{h,j-1}, G(\nabla u^{h,j-1}, \lambda) - G(\nabla u^{h,j-2}, \lambda) \Big\rangle \Big)\, dx$$

$$= \sum_{j=1}^{j_h} \int_\Omega \Big\langle \boldsymbol{\nu}^{h,j-1},\ \int_0^1 G_F(s\nabla u^{h,j-1} + (1-s)\nabla u^{h,j-2},\ \lambda) \cdot h\nabla v^{h,j-1}\, ds \Big\rangle\, dx$$

$$\leqq \sum_{j=1}^{j_h} \int_\Omega \Big\langle \boldsymbol{\nu}^{h,j-1},\ \int_0^1 m(|\nabla u^{h,j-2} + sh\nabla v^{h,j-1}| + |\lambda| + 1)h|\nabla v^{h,j-1}|\, ds \Big\rangle\, dx$$

$$\leqq \sum_{j=1}^{j_h} \int_\Omega \Big(mh^{\frac{1}{2}}|\nabla u^{h,j-2}|\, h^{\frac{1}{2}}|\nabla v^{h,j-1}| + mh^2|\nabla v^{h,j-1}|^2$$

$$+ m\Big\langle \boldsymbol{\nu}^{h,j-1}, |\lambda|\Big\rangle h|\nabla v^{h,j-1}| + mh|\nabla v^{h,j-1}| \Big)\, dx$$

$$\leqq \frac{m}{2} \sum_{j=1}^{j_h} h\|\nabla u^{h,j-2}\|^2 + m\sum_{j=1}^{j_h} (\frac{3h}{2} + h^2)\|\nabla v^{h,j-1}\|^2$$

$$+ \frac{m}{2}\sum_{j=1}^{j_h} h \int_\Omega \Big\langle \boldsymbol{\nu}^{h,j-1}, |\lambda|^2\Big\rangle\, dx + \frac{m}{2}\sum_{j=1}^{j_h} h\mathcal{L}^n(\Omega)^2$$

$$\leqq \frac{m}{2} T \sup_{h,j} \|\nabla u^{h,j}\|^2 + m(1+h)\sum_{j=1}^{\infty} h\|\nabla v^{h,j}\|^2$$

$$+ m\sum_{j=1}^{\infty} h \int_\Omega \Big\langle \boldsymbol{\nu}^{h,j-1}, |\lambda|^2\Big\rangle\, dx + \frac{m}{2} T\mathcal{L}^n(\Omega)^2$$

$$\equiv M_T$$

and M_T depends only on T, using $(2.9)^i$, $(2.10)^i$ and (2.11). Clearly this estimate is independent of T if $|\tau(F, \dot{F})| \leqq m|\dot{F}|$ (in agreement with a similar assertion in Lemma A.1). Therefore (2.18) implies that

$$\sup_{\substack{1 \leqq j \leqq [\frac{T}{h}] \\ h>0}} \int_\Omega |\nabla v^{h,j}|^2\, dx + \sum_{j=1}^{[\frac{T}{h}]} h \int_\Omega \frac{(v^{h,j} - v^{h,j-1})^2}{h^2}\, dx \leqq M_T$$

which is precisely (2.15), (2.16).

It follows now that $(v^{h,k})_k$ can be assumed to be bounded in $H_0^1(\Omega)$ independently of h, j, k as

$$\int_\Omega G(\nabla u^{h,j-1}, \nabla v^{h,j,k})dx \overset{k\to\infty}{\longrightarrow} \int_\Omega G^{qc}(\nabla u^{h,j-1}, \nabla v^{h,j})dx$$

so that it may be always assumed that

$$\left|\int_\Omega G(\nabla u^{h,j-1}, \nabla v^{h,j,k})dx\right| \leqq \left|\int_\Omega G^{qc}(\nabla u^{h,j-1}, \nabla v^{h,j})dx + 1\right|.$$

The bounds for the Young measures also follow,

$$\begin{aligned}\|\boldsymbol{\nu}^{h,j}\|_{L^1(\Omega,\mathcal{E}^{2'})} &= \int_\Omega \left| \sup_{\|f\|_{\mathcal{E}^2}=1} \int_{\mathbb{R}^n} f(\alpha)d\boldsymbol{\nu}^{h,j}\right| dx \leqq \int_\Omega \int_{\mathbb{R}^n} |\alpha|^2 d\boldsymbol{\nu}^{h,j}\ dx + 1\\ &= \lim_{k\to\infty} \int_\Omega |\nabla v^{h,j,k}|^2\, dx + 1\end{aligned}$$

and by (2.15) the bound asserted in $L^1_{\text{loc}}(\mathcal{E}^{2'})$ follows. Boundedness in $L^2(\Omega, \mathcal{E}^{1'})$ follows similarly and boundedness in $L^\infty(\Omega, \mathcal{P}(\mathbb{R}^{m\times n}))$ uses the fact that $\boldsymbol{\nu}^{h,j}$ are probability measures. □

We are now able to let h tend to zero. Firstly, observe two relations between the iterates: for each $t \in [hj, h(j+1))$,

$$U^h(t) - u^h(t-h) = (t-hj)v^{j+1} \equiv (t-hj)v^h(t) \equiv (t-hj)U_t^h(t), \quad (2.19)$$

$$V^h(t) - v^h(t) = (t-hj-h)z^h(t) \equiv (t-hj-h)V_t^h(t). \quad (2.20)$$

Thus for any Lipschitz function f with constant L_f,

$$|f(U^h(t)) - f(u^h(t-h))| \leqq L_f h|v^h(t)| = L_f h|U_t^h(t)|, \quad (2.21)$$

$$|f(V^h(t)) - f(v^h(t-h))| \leqq L_f h|z^h(t)| = L_f h|V_t^h(t)|. \quad (2.22)$$

Lemma 2.7 (Convergence**).** *There exists a pair $(u, \boldsymbol{\xi})$ with*

$$u \in W^{1,\infty}(\mathbb{R}^+,\ H_0^1(\Omega)) \cap H^2_{\text{loc}}(\mathbb{R}^+,\ L^2(\Omega)) \cap H^1_{\text{loc}}(\mathbb{R}^+,\ L^2(\Omega))$$

and $u_{tt} \in L^2_{\text{loc}}(Q_\infty)$, $\nabla u_t \in L^2(Q_\infty)$ and $\boldsymbol{\xi} = (\xi_{t,x})_{(t,x)\in Q_\infty}$ a Young measure such that $u, \boldsymbol{\xi}$ are weak limit points along a subsequence in $h \to 0$

$$(u^h, U^h, \nabla u^h, \nabla U^h, v^h, V^h, \nabla v^h, \nabla V^h) \ \rightharpoonup\ (u, u, \nabla u, \nabla u, u_t, u_t, \nabla u_t, \nabla u_t)$$

weakly∗ in $L^\infty(L^2)$ and weakly in $L^2_{\text{loc}}(Q_\infty)$. In particular,

$$(U^h, V^h) \ \rightharpoonup\ (u, u_t) \qquad \textit{weakly in } H^1_{\text{loc}}(Q_\infty).$$

Strong convergence then follows, namely,

$$\begin{aligned} U^h &\longrightarrow u && \text{strongly in } L^p_{\text{loc}}(L^2)\ \forall p \geqq 1, \\ V^h &\longrightarrow u_t && \text{strongly in } L^2_{\text{loc}}(Q_\infty). \end{aligned} \tag{2.23}$$

Time derivatives converge also as

$$\begin{aligned} (\nabla v^h, \nabla V^h) &\rightharpoonup (\nabla u_t, \nabla u_t) && \text{weakly in } L^2(Q_\infty), \\ z^h &\rightharpoonup u_{tt} && \text{weakly in } L^2_{\text{loc}}(Q_\infty) \end{aligned}$$

and the convergence of $(z^h)_{h>0}$ *is weakly in* $L^2(Q_\infty)$ *if the growth of* τ *is* $|\tau(F, \dot F)| \leqq m|\dot F|$ *in* (1.15). *Also* $(\boldsymbol{\xi}^h, \boldsymbol{\mu}^h) \rightharpoonup (\boldsymbol{\xi}, \boldsymbol{\mu})$ *weakly∗ in* $L^\infty(\mathcal{P})$ *and* $\boldsymbol{\nu}^h \rightharpoonup \boldsymbol{\nu}$ *weakly∗ in* $L^1_{\text{loc}}(\mathcal{E}^{2'}) \cap L^2_{\text{loc}}(\mathcal{E}^{1'}) \cap L^\infty(\mathcal{P})$. *Moreover,* $\boldsymbol{\mu}, \boldsymbol{\nu}$ *are the projections of* $\boldsymbol{\xi}$ *on each component* $\mathbb{R}^{m\times n}$ *and are Young measures generated by spatial gradients in the sense of* [21].

Proof. The above assertions follow from the uniform in h bounds on the iterates (2.12) provided by Lemmas 2.4 and 2.6. Firstly, $(U^h)_h$ is bounded in $W^{1,\infty}(H^1_0)$ by the uniform bounds $(2.10)^i$, (2.15) and by compact embedding it is precompact in $L^p(L^2)$ for all $p \geqq 1$. Also it follows that $(U^h)_h$ is bounded in $H^1_{\text{loc}}(Q_\infty)$. These bounds give the asserted weak∗ convergence in $L^\infty(L^2)$. In particular $(V^h)_h$ is bounded in $L^\infty(H^1_0)$. By $(2.9)^i$ $(\nabla V^h)_h$ is bounded in $L^2(Q_\infty)$ and, by (2.16), $(z^h = V^h_t)_h$ is bounded in $L^2(Q_T)$ for all T. Therefore $(V^h)_h$ is bounded in $H^1_{\text{loc}}(Q_\infty)$ and is precompact in $L^2_{\text{loc}}(Q_\infty)$. These implications use the fact that the limits of piecewise constant and continuous piecewise linear interpolates are the same, as was shown in [20, Lemma 6.3]. (Thus U^h, u^h have the same weak limit u above and similarly V^h, v^h converge to v and, since $v^h = \partial_t U^h$, then $v = u_t$ and similarly $z^h = \partial_t V^h$ converge to u_{tt}.)

Considering the Young measures, note that $\partial_t \langle \boldsymbol{\mu}^h, \alpha\rangle = \langle \boldsymbol{\nu}^h, \alpha\rangle (t, x)$ a.e. where $\boldsymbol{\mu}^h$ and $\boldsymbol{\nu}^h$ are the projections of $\boldsymbol{\xi}^h$ on $\mathbb{R}^{m\times n}$. By Lemma 2.6, $\boldsymbol{\nu}^h$ has a weak limit point $\boldsymbol{\nu}$ in the topology indicated while $(\boldsymbol{\xi}^h, \boldsymbol{\mu}^h)$ are bounded in the space $L^\infty(\mathcal{P})$ so we may extract weak∗ limit points $\boldsymbol{\mu}, \boldsymbol{\nu}, \boldsymbol{\xi}$ respectively and $\boldsymbol{\mu}, \boldsymbol{\nu}$ are the projections of $\boldsymbol{\xi}$ (not necessarily a product measure), and, moreover, these are Young measures. For example, $\boldsymbol{\xi}$ is a Young measure generated by a diagonal subsequence of $(\nabla u^h, \nabla v^h)$; (this construction was shown explicitly in [13, Lemma 2.4] for the single equation (1.2) and applies here). This fact is of analytical interest but not essential in obtaining a measure-valued solution here. A separate fact proved in Lemma A.3 is that the piecewise constant and the continuous piecewise linear interpolations generate the same Young measures. □

Letting $h \to 0$ in (2.14) we obtain the weak equation (1.22) stated in Theorem A,

$$\int_0^T \int_\Omega (\langle \boldsymbol{\xi}, \sigma + \tau\rangle \cdot \nabla\zeta + u_{tt}\zeta)\, dx\, dt = 0,$$

where $\langle \boldsymbol{\xi}, \sigma + \tau\rangle = \langle \boldsymbol{\mu}, \sigma\rangle + \langle \boldsymbol{\xi}, \tau\rangle$. The relations on $\boldsymbol{\xi}, \boldsymbol{\mu}, \boldsymbol{\nu}$ in (1.23), (1.24) follow directly from (2.2), (2.3). It remains to discuss the initial data.

Corollary 2.8 (Initial data)**.** *As* $h \to 0^+$, $U^h \longrightarrow u$ *and* $V^h \longrightarrow u_t$ *strongly in* $C_{\text{loc}}(\mathbb{R}^+; L^2(\Omega))$. *Thus* u *attains its initial data* (u_0, v_0) *strongly in* $L^2(\Omega) \times L^2(\Omega)$.

Proof. This follows directly from Lemma 2.7 using Lemma 1.3. As $(U^h, V^h)_h$ converges in $H^1_{\mathrm{loc}}(Q_\infty)$ and $(U^h, V^h)(0, \cdot) = (u_0, v_0)$ for all h, thus a subsequence converges strongly in $C_{\mathrm{loc}}(\mathbb{R}^+, L^2(\Omega))$ and the initial data are obtained as asserted. □

3. Existence and uniqueness of a classical weak solution

This section is divided into two parts: in Section 3.1 the solution obtained by time-discretization in Section 2 is shown to be the unique classical weak solution (1.1) or (1.2) if the monotonicity conditions (1.18) and (1.21) are assumed. In Section 3.2 I give sufficient conditions for any measure-valued solution to coincide with this classical weak solution.

3.1. Classical weak from measure-valued solutions

First, without assuming (1.18) or (1.21) I show that the Young measures $\boldsymbol{\nu}^h$ obtained in Section 2 satisfy a property related to monotonicity which in the case of the gradient flow also holds for the limiting measure $\boldsymbol{\nu}$. Then, imposing (1.18) and (1.21), I show that the convergence of $(U^h_t)_{h>0}$ can be improved to strong convergence in $H^1_{\mathrm{loc}}(Q_\infty)$: this is related to a partial smoothing property of the solution operator.

Lemma 3.1 (Independence**).** *The identity function and τ are independent with respect to the Young measure $\boldsymbol{\nu}^h$ obtained in Section 2, that is,*

$$\left\langle \boldsymbol{\nu}^h, \tau(\nabla u^h(t-h), \cdot) \cdot \mathrm{id} \right\rangle = \left\langle \boldsymbol{\nu}^h, \tau(\nabla u^h(t-h), \cdot) \right\rangle \cdot \left\langle \boldsymbol{\nu}^h, \mathrm{id} \right\rangle \tag{3.1}$$

for a.e. $(t, x) \in Q_\infty$. *In the case of the gradient flow, passing to the limit in h,* (1.32) *holds, and in fact,*

$$\langle \boldsymbol{\nu}, \tau \cdot \mathrm{id} \rangle = \langle \boldsymbol{\nu}, \tau \rangle \cdot \langle \boldsymbol{\nu}, \mathrm{id} \rangle, \qquad (t, x) \text{ a.e.} \tag{3.2}$$

Proof. Suppose $u^{h,j,k} \to u^{h,j}$ is a minimizing sequence and let $\boldsymbol{\mu}^{h,j}, \boldsymbol{\nu}^{h,j}$ be generated as in Step I, Section 2. First we see that

$$\left\langle \boldsymbol{\nu}^{h,j}, \tau \cdot \mathrm{id} \right\rangle = \left\langle \boldsymbol{\nu}^{h,j}, \tau \right\rangle \cdot \left\langle \boldsymbol{\nu}^{h,j}, \mathrm{id} \right\rangle \quad \text{for a.e. } x \in \Omega. \tag{3.3}$$

We know that for all $\zeta \in H^1_0(\Omega)$

$$\int_\Omega \left(\left(\sigma(\nabla u^{h,j-1}) + \tau(\nabla u^{h,j-1}, \nabla v^{h,j,k}) \right) \cdot \nabla\zeta + \frac{v^{h,j,k} - v^{h,j-1}}{h} \zeta \right) dx$$
$$\xrightarrow{k\to\infty} \int_\Omega \left(\left(\sigma(\nabla u^{h,j-1}) + \left\langle \boldsymbol{\nu}^{h,j}, \tau(\nabla u^{h,j-1}, \cdot) \right\rangle \right) \cdot \nabla\zeta \right.$$
$$\left. + \frac{v^{h,j} - v^{h,j-1}}{h} \zeta \right) dx = 0$$

from which it is deduced that $\left(\nabla \cdot \tau(\nabla u^{h,j-1}, \nabla v^{h,j,k})\right)_{k\geqq 0}$ is strongly convergent in $H^{-1}(\Omega)$ to $\nabla \cdot \langle \boldsymbol{\nu}^{h,j}, \tau\rangle$, since $v^{h,j,k} \to v^{h,j}$ strongly in $L^2(\Omega)$ and weakly in $H_0^1(\Omega)$. Therefore, (suppressing the explicit dependence on $\nabla u^{h,j-1}$),

$$\lim_{k\to\infty} \int_\Omega \nabla \cdot \tau(\nabla v^{h,j,k}) \cdot v^{h,j,k} \phi\, dx = -\int_\Omega \langle \boldsymbol{\nu}^{h,j}, \tau\rangle \cdot \nabla v^{h,j} \phi\, dx$$
$$= -\int_\Omega \langle \boldsymbol{\nu}^{h,j}, \tau\rangle \cdot \langle \boldsymbol{\nu}^{h,j}, \mathrm{id}\rangle \phi\, dx$$

for any $\phi \in C_0^\infty(\Omega)$, that is,

$$\tau(\nabla v^{h,j,k}) \cdot \nabla v^{h,j,k} \longrightarrow \langle \boldsymbol{\nu}^{h,j}, \tau\rangle \cdot \langle \boldsymbol{\nu}^{h,j}, \mathrm{id}\rangle$$

in the sense of distributions. But since $\tau \cdot \mathrm{id} \in \mathcal{E}^2$ it is also true that the L^1-weak limit of the same sequence is given by $\langle \boldsymbol{\nu}^{h,j}, \tau \cdot \mathrm{id}\rangle$ and by equating the two expressions we obtain (3.1).

Consider the limit as $h \to 0$. By the $L^2(Q_\infty)$ weak convergence (and thus by strong $H^{-1}(Q_\infty)$ compactness) of $(u^h_{tt})_{h>0}$ we deduce the strong convergence of $\nabla \cdot A^h \equiv \nabla \cdot (\sigma(\nabla u^h) + \langle \boldsymbol{\nu}^h, \tau\rangle)$ in H^{-1}_{loc}. Let $\nabla \cdot A^h \to \nabla \cdot A$ strongly in $H^{-1}_{\mathrm{loc}}(Q_\infty)$ (the limit of operators in divergence form is also in divergence form). Also A coincides with the weak L^2 limit of A^h: since $\sigma(\nabla u^h) \rightharpoonup \langle \boldsymbol{\mu}, \sigma\rangle$ and $\langle \boldsymbol{\nu}^h, \tau\rangle \rightharpoonup \langle \boldsymbol{\nu}, \tau\rangle$ weakly in $L^2_{\mathrm{loc}}(Q_\infty)$ then $A = \langle \boldsymbol{\mu}, \sigma\rangle + \langle \boldsymbol{\xi}, \tau\rangle = \langle \boldsymbol{\xi}, \sigma + \tau\rangle$. Apply the weak equation (2.14) against $U^h_t = v^h$ recalling $v^h \rightharpoonup u_t$ weakly in $H^1_{\mathrm{loc}}(Q_\infty)$ (cf. (Lemma 2.7)). Thus for any $\phi \in C_0^\infty(Q_\infty)$,

$$\int_0^\infty \int_\Omega \left(\langle \boldsymbol{\mu}^h, \sigma\rangle + \langle \boldsymbol{\nu}^h, \tau\rangle\right) \cdot \langle \boldsymbol{\nu}^h, \mathrm{id}\rangle \phi\, dx\, dt$$
$$\longrightarrow \int_0^\infty \int_\Omega \left(\langle \boldsymbol{\mu}, \sigma\rangle + \langle \boldsymbol{\xi}, \tau\rangle\right) \cdot \langle \boldsymbol{\nu}, \mathrm{id}\rangle\, \phi\, dx\, dt.$$

Now restrict to the gradient flow, $\sigma = 0$ and $\tau = \tau(\dot{F})$ and thus $\boldsymbol{\xi} = \boldsymbol{\nu}$. As before, $\langle \boldsymbol{\nu}^h, \tau \cdot \mathrm{id}\rangle$ converges to $\langle \boldsymbol{\nu}, \tau \cdot \mathrm{id}\rangle$ weakly in $L^1_{\mathrm{loc}}(Q_\infty)$. But it also converges to $\langle \boldsymbol{\nu}, \tau\rangle \cdot \langle \boldsymbol{\nu}, \mathrm{id}\rangle$ in the same topology by the div-curl lemma using that $(z^h)_h$ in (2.12) are compact in $H^{-1}_{\mathrm{loc}}(Q_\infty)$. Equating the two limits we obtain the stronger version of (1.32) as claimed. □

Remark 3.2 (Independence property for viscoelasticity). In the case of (1.1) the limit of (3.1) as $h \to 0$ is related to a chain rule property involving $\boldsymbol{\mu}$, $\boldsymbol{\nu}$ and the limit of (2.13). This is described in Lemma A.2. The analysis shows that these properties are relevant to the question of uniqueness of measure-valued solutions.

Theorem 3.3 (Existence of a classical weak solution**).** *Assume that* (1.18) *holds. Then the solution obtained in Section 2 is a classical weak solution and satisfies for all $T > 0$*

$$\int_0^T \int_\Omega \left((\sigma(\nabla u) + \tau(\nabla u, \nabla u_t)) \cdot \nabla \zeta + u_{tt} \zeta \right) dx\, dt = 0 \qquad (3.4)$$

for all $\zeta \in H^1_{\mathrm{loc}}(Q_\infty)$ with zero trace on $\partial\Omega$.

Proof. The use of Gronwall's inequality is natural in this context. I show that the sequence $(\nabla U_t^h)_{h>0}$ obtained in the existence proof is Cauchy in $L^2(Q_\infty)$ from which it can be deduced that $\boldsymbol{\xi} = \delta_{(\nabla u, \nabla u_t)}$. Apply the weak equation (2.14) on $v^h = U_t^h$ and subtract at indices h and h', recalling that ∇v^h generates $\boldsymbol{\nu}$. Below $e(h, h'; T)$ represents the error terms arising from exchanging the piecewise constant with the piecewise differentiable interpolates as shown. Note that under (1.18) $\boldsymbol{\nu}^h j = \delta_{\nabla v^{h,j}}$ and thus the $\boldsymbol{\nu}^h, \boldsymbol{\nu}^{h'}$ below are sums of delta functions. For given $T > 0$ have (specifying the dependence in time only when it is $t - h$),

$$
\begin{aligned}
&\int_{Q_T} \Big((z^h - z^{h'}) \cdot (V^h - V^{h'}) + \kappa |\nabla v^h - \nabla v^{h'}|^2 \Big)\, dx\, dt \\
&\quad = \int_{Q_T} \left(\frac{1}{2} \partial_t |V^h - V^{h'}|^2 + \kappa \langle \boldsymbol{\nu}^h \times \boldsymbol{\nu}^{h'}, |\alpha - \beta|^2 \rangle \right) dx\, dt \\
&\quad \leqq \int_{Q_T} \ \frac{1}{2} \partial_t |V^h - V^{h'}|^2 + l |\nabla u^h - \nabla u^{h'}|^2 \\
&\qquad\qquad + \langle \boldsymbol{\nu}^h \times \boldsymbol{\nu}^{h'}, (\tau(\nabla u^h(t-h), \alpha) \\
&\qquad\qquad - \tau(\nabla u^{h'}(t - h'), \beta)) \cdot (\alpha - \beta) \rangle \ dx\, dt \\
&\quad = -\int_{Q_T} \bigg(\Big(\sigma(\nabla U^h) - \sigma(\nabla U^{h'}) \Big) \cdot (\nabla v^h - \nabla v^{h'}) \\
&\qquad\qquad + l |\nabla U^h - \nabla U^{h'}|^2 \bigg) dx\, dt + e(h, h'; T) \qquad\qquad (3.5) \\
&\quad \leqq \left(\frac{L_\sigma^2}{2\varepsilon} + l \right) \int_{Q_T} |\nabla U^h - \nabla U^{h'}|^2 \, dx\, dt \\
&\qquad\qquad + \frac{\varepsilon}{2} \int_{Q_T} |\nabla v^h - \nabla v^{h'}|^2 \, dx\, dt + e(h, h'; T) \\
&\quad \leqq \left(\frac{L_\sigma^2}{2\varepsilon} + l \right) T \int_{Q_T} \int_0^t |\nabla U_t^h - \nabla U_t^{h'}|^2 ds\, dx\, dt \\
&\qquad\qquad + \frac{\varepsilon}{2} \int_{Q_T} |\nabla U_t^h - \nabla U_t^{h'}|^2 \, dx\, dt + e(h, h'; T) \\
&\quad \leqq \left(\frac{L_\sigma^2}{2\varepsilon} + l \right) T \int_0^t \|\nabla U_t^h - \nabla U_t^{h'}\|_{L^2(Q_T)}^2 (s)\, ds \\
&\qquad\qquad + \frac{\varepsilon}{2} \|\nabla U_t^h - \nabla U_t^{h'}\|_{L^2(Q_T)}^2 + e(h, h'; T)
\end{aligned}
$$

where the first inequality holds by the lower semicontinuity of the norm, the second uses (1.18), the equality is by (3.1) and (2.14) and the following inequality uses the Lipschitz condition on σ (and $\nabla U^h - \nabla U^{h'} = \int_0^t (\nabla U_t^h - \nabla U_t^{h'}) ds$). The error term is estimated using (2.20),

$$\begin{aligned} e(h,h';T) &\leqq \int_{Q_T} \Big(\big|\sigma(\nabla u^h)(t-h) - \sigma(\nabla U^h)(t)\big| \\ &\quad + \big|\sigma(\nabla u^{h'})(t-h') - \sigma(\nabla U^{h'})(t)\big| \Big) \ |\nabla U_t^h - \nabla U_t^{h'}|\, dx\, dt \\ &\quad + l \int_\Omega (|\nabla u^h - \nabla u^{h'}|^2 - |\nabla U^h - \nabla U^{h'}|^2)\, dx\, dt \\ &\quad + \int_\Omega |z^h - z^{h'}||v^h - V^h + v^{h'} - V^{h'}|\, dx\, dt \\ &\leqq (h+h')C \left(\|\nabla v^h\|^2 + \|\nabla v^{h'}\|^2 + \|z^h\|^2 + \|z^{h'}\|^2 \right) \end{aligned}$$

with $C = C(T, L_\sigma, \sup_{t,h} \|\nabla u^h\|)$ a positive constant independent of h, h' and $e(h,h';T) \to 0$ as $h, h' \to 0$ by $(2.10)^i$, (2.15) and (2.16). Choosing $\varepsilon = \frac{\kappa}{2}$ (from (1.18)) the first and last line in (3.5) give,

$$\begin{aligned} &\frac{1}{2}\|V^h - V^{h'}\|^2(T) + \frac{\kappa}{2}\int_0^T \|\nabla v^h - \nabla v^{h'}\|^2(t)\, dt \\ &\qquad \leqq \left(\frac{L_\sigma^2}{\kappa} + l\right) T \int_0^T \int_0^t \|\nabla v^h - \nabla v^{h'}\|^2(s)\, ds\, dt + e(h,h';T). \qquad (3.6) \end{aligned}$$

Fix $T_* > 0$ and note that $e(h,h;T) \leqq e(h,h;T_*)$, and by the Growall inequality we have for all $T \leqq T_*$ (ignoring the first term in (3.6)),

$$\begin{aligned} &\int_0^T \left(\|U_t^h - U_t^{h'}\|^2_{L^2(Q_t)} + \|\nabla U_t^h - \nabla U_t^{h'}\|^2_{L^2(Q_t)} \right) dt \\ &\qquad\qquad \leqq e(h,h';T_*)e^{cT} \overset{h,h'\to 0}{\longrightarrow} 0. \end{aligned}$$

We conclude that

$$(U_t^h, \nabla U_t^h) \longrightarrow (u_t, \nabla u_t) \quad \text{strongly in } L^2_{\mathrm{loc}}(Q_\infty),$$

that is, $U^h \longrightarrow u$ strongly in $H^1_{\mathrm{loc}}(Q_\infty)$ (similarly for v^h). Therefore, $\boldsymbol{\xi} = \delta_{(\nabla u, \nabla u_t)}$ and (3.4) is proved. □

Well-posedness and uniqueness for the solution immediately follow: if u^h, v^h are constructed corresponding to varying initial data (u_0^h, v_0^h) and $(\tilde{u}_0^h, \tilde{v}_0^h)$, then

$$\|U^h - \tilde{U}^h\|_{H^1(Q_\infty)} \leqq C\left(\|u_0^h - \tilde{u}_0^h\|_{H_0^1(\Omega)} + \|v_0^h - \tilde{v}_0^h\|_{H_0^1(\Omega)} \right).$$

For this the above proof can be repeated, now including the difference in L^2 of the initial data on the right-hand side of (3.6). This is summarized as follows:

Corollary 3.4 (Uniqueness and well-posedness I**).** *Assume* (1.18) *holds. The solution obtained with time discretization in Section 2 is the unique, classical weak solution and is well posed with respect to initial data* (1.4).

Restrict again to the gradient flow and assume (1.21). I close this section by showing how the independence property (1.32) can be used to recover this solution from a measure-valued solution (in particular the one obtained by the time-discretization).

Lemma 3.5 (Recovery of classical weak solutions of the gradient flow**).** *Under the monotonicity assumption* (1.21) *the Young measure solution constructed in Section 2, or any measure-valued solution* $(v, \boldsymbol{\nu})$ *with* $v \in H^1_{\mathrm{loc}}(Q_\infty)$, $\boldsymbol{\nu} \in \mathcal{P}(Q_\infty \times \mathbb{R}^{m\times n})$ *which satisfies* (1.29)–(1.31) *and* (1.32)*, is the unique classical weak solution of* (1.2).

Proof. Let $(v, \boldsymbol{\nu})$, be a measure-valued solution satisfying the assumptions. By the weak equation (1.29) applied to $\langle \boldsymbol{\nu}, \mathrm{id}\rangle - \nabla v$,

$$
\begin{aligned}
0 &= \int_0^t \int_\Omega \langle \boldsymbol{\nu}, \tau\rangle \cdot (\langle \boldsymbol{\nu}, id\rangle - \nabla v)\, dx\, ds \\
&= \int_0^t \int_\Omega \langle \boldsymbol{\nu}, (\tau(\alpha) - \tau(\nabla v)) \cdot (\alpha - \nabla v)\rangle\; dx\; ds \\
&= \int_0^t \int_\Omega \langle \boldsymbol{\nu}, (\tau(\alpha) - \tau(\nabla v)) \cdot (\alpha - \nabla v)\rangle\; dx\; ds.
\end{aligned}
$$

By assumption the last term is non-negative, thus the support of the measure lies where the integrand vanishes. This forces $\langle \boldsymbol{\nu}, \alpha\rangle = \nabla v$ for (t, x) a.e. Uniqueness follows from Theorem 3.3. □

3.2. Sufficient conditions for classical weak solutions

Consider the class $\mathcal{S}$ of probability measures $\boldsymbol{\xi} \in \mathcal{P}(Q_\infty \times \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n})$ with marginal $\boldsymbol{\nu} \in \mathcal{P}(Q_\infty \times \mathbb{R}^{m\times n})$ in the second factor $\mathbb{R}^{m\times n}$ such that for a.e. $t \in \mathbb{R}^+$ (1.27) holds,

$$
\int_\Omega \langle \boldsymbol{\xi}, \tau(\alpha, \beta)\rangle \cdot \langle \boldsymbol{\nu}, \beta\rangle\; dx \geqq \int_\Omega \langle \boldsymbol{\xi}, \tau(\alpha, \beta) \cdot \beta\rangle\; dx.
$$

In this section it is shown that under (1.18) the unique classical weak solution of (1.1) obtained in the above sections is unique also within the class of measure-valued solutions in $\mathcal{S}$. (In other words, a measure-valued solution $(u, \boldsymbol{\xi})$ with $\boldsymbol{\xi} \in \mathcal{S}$ is not genuinely measure-valued but coincides with the classical weak solution obtained in the previous theorem). First I show another relation between the measures $\boldsymbol{\mu}, \boldsymbol{\nu}$.

Lemma 3.6 (Kato-type inequalities**).** *Consider a pair of Young measures* $\boldsymbol{\mu}, \boldsymbol{\nu} \in \mathcal{P}(Q_\infty \times \mathbb{R}^{m\times n})$*, each parametrized by* $(t, x) \in Q_\infty$ *and generated by two sequences in* $L^2_{\mathrm{loc}}(Q_\infty)$, $(f^h)_{h>0}$ *and time derivatives* $(f^h_t)_{h>0}$ *respectively with* $f^h(t, \cdot) = 0$*. Assume that the Young measure representation holds weakly in* $L^1_{\mathrm{loc}}(Q_\infty)$ *for functions in* $\mathcal{E}^2$*. Then the Kato-type inequalities hold,*

$$
|\partial_t \langle \boldsymbol{\mu}, |\alpha|\rangle| \leqq \langle \boldsymbol{\nu}, |\beta|\rangle, \quad (t, x) \text{ a.e.} \tag{3.7}
$$

and more generally, for $1 \leqq p \leqq 2$,

$$|\partial_t \langle \boldsymbol{\mu}, |\alpha|^p \rangle| \leqq p \left(\left\langle \boldsymbol{\mu}, |\alpha|^2 \right\rangle\right)^{\frac{p-1}{2}} \left(\left\langle \boldsymbol{\nu}, |\beta|^{\frac{2}{3-p}} \right\rangle\right)^{\frac{3-p}{2}}, \quad (t, x) \text{ a.e.} \tag{3.8}$$

For any $T > 0$ *and* $1 \leqq p \leqq 2$ *the integral form holds,*

$$\int_{Q_T} \langle \boldsymbol{\mu}, |\alpha|^p \rangle \, dx\, dt \leqq T^{p-1} \int_{Q_T} \int_0^t \langle \boldsymbol{\nu}, |\beta|^p \rangle \, ds\, dx\, dt. \tag{3.9}$$

Proof. It can be shown by approximation by positive smooth functions bounded away from 0 that the distributional derivative of $|f^h|^p$ is in $L^{\frac{2}{p}}_{\text{loc}}(Q_\infty)$ and $\partial_t |f^h|^p = p|f^h|^{p-2}(f^h \cdot f^h_t)$ for a.e. (t, x) on the set $\{f^h \neq 0\}$ and 0 a.e. on the complement. (To see this, approximate each f^h by $A^{k,\varepsilon}_{(h)} = (|f^k_{(h)}|^2 + \varepsilon^2)^{\frac{p}{2}}$ where $f^k_{(h)}$ are smooth and converge pointwise a.e. to $f_{(h)}$, and compute $\lim_{k\to\infty} \lim_{\varepsilon\to 0} \partial_t A^{k,\varepsilon}_{(h)}$).) Then,

$$|\partial_t |f^h|^p| \leqq p|f^h|^{p-1}|f^h_t|, \quad (t, x) \text{ a.e.} \tag{3.10}$$

so that for any measurable E compactly supported in Q_∞,

$$\int_E |\partial_t |f^h|^p| \, dx\, dt \leqq p \left(\int_E |f^h|^2 \, dx\, dt\right)^{\frac{p-1}{2}} \left(\int_E |f^h_t|^{\frac{2}{3-p}} \, dx\, dt\right)^{\frac{3-p}{2}} < \infty$$

and in fact $\partial_t |f^h|^p$ is bounded in $L^{\frac{2}{p}}$. Since $\partial_t |f^h|^p \rightharpoonup \partial_t \langle \boldsymbol{\mu}, |a|^p \rangle$ in the sense of distributions and hence weakly in $L^{\frac{2}{p}}$,

$$\int_E \partial_t |f^h|^p \, dx\, dt \longrightarrow \int_E \partial_t \langle \boldsymbol{\mu}, |\alpha|^p \rangle \, dx\, dt$$

by the Young measure representation for $|f^h|^p$. Also,

$$\int_E |f^h_t|^{\frac{2}{3-p}} \, dx\, dt \longrightarrow \int_E \left\langle \boldsymbol{\nu}, |\beta|^{\frac{2}{3-p}} \right\rangle dx\, dt$$

(the representation holds since $\frac{2}{3-p} \leqq 2$). Then (3.7) and (3.8) follow by Lebesgue differentiation. The integral form (3.9) follows more readily: for $T > 0$ and $0 \leqq t \leqq T$,

$$\begin{aligned} \int_0^t \int_\Omega \partial_t |f^h|^p \, dx\, dt &= \int_\Omega |f^h|^p(t)\, dx \\ &= \int_\Omega \left| \int_0^t f^h_t \, ds \right|^p dx \leqq \int_\Omega \left(\int_0^t |f^h_t|^p \, ds \right) t^{p-1} \end{aligned}$$

and so by integrating

$$\int_0^T \int_\Omega |f^h|^p \, dx\, dt \leqq T^{p-1} \int_0^T \int_\Omega \int_0^t |f^h_t|^p \, ds\, dx\, dt$$

and then (3.9) follows by taking the limit in h. □

Remark 3.7. The case $p = 1$ is the only one in which the bound in (3.7) is independent of the L^2 norm of the sequence generating the measures. (The usual Kato inequality refers to this case.) For $p = 2$ (and in general for $p \geqq 2$ changing the topology of convergence accordingly) it is easier to justify both taking the weak derivative of $\partial_t |f^h|^p$ and the chain rule formula.

Theorem 3.8 (Uniqueness and well-posedness II**).** *Suppose τ satisfsfies* (1.18) *and $\tilde{u} \in H^2_{\text{loc}}(L^2) \cap H^1_{\text{loc}}(H^1_0)$ is such that there exists a measure $\tilde{\boldsymbol{\xi}} \in \mathcal{P}(Q_\infty \times \mathbb{R}^{m\times n} \times \mathbb{R}^{m\times n})$ with marginals $\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\nu}}$ for which* (1.22)–(1.24) *and* (1.27) *hold with initial data* (1.4) *and the Young measure representation holds in L^1 for functions in $\mathcal{E}^2$. Then $\tilde{u}$ coincides with the unique classical solution u of Theorem 3.3, i.e., $u = \tilde{u}$ (t, x) a.e. in Q_∞.*

Proof. Suppose that $(u, \boldsymbol{\xi})$ is the solution of Theorem 3.4 , where of course $\boldsymbol{\xi} = \delta_{(\nabla u, \nabla u_t)}$, and $(\tilde{u}, \tilde{\boldsymbol{\xi}})$ a measure-valued solution with $\tilde{\boldsymbol{\xi}} \in \mathcal{S}$ and $\tilde{\boldsymbol{\mu}}, \tilde{\boldsymbol{\nu}}$ respectively the projections of $\tilde{\boldsymbol{\xi}}$ on each argument. We will fix $T > 0$, apply the (3.4) and (1.22) with $\zeta = \nabla u - \nabla\tilde{u}$ and subtract. Then

$$\int_{Q_T} \left(\partial_t \frac{|u_t - \tilde{u}_t|^2}{2} + \kappa \langle \tilde{\boldsymbol{\nu}}, |\nabla u_t - \tilde{\beta}|^2 \rangle \right) dx\, dt$$

(using (1.18), (1.27))

$$\begin{aligned}
&\leqq \int_{Q_T} \Big((u_{tt} - \tilde{u}_{tt}) \cdot (u_t - \tilde{u}_t) + \langle \tilde{\boldsymbol{\xi}}, \tau(\nabla u, \nabla u_t) \\
&\quad - \tau(\tilde{\alpha}, \tilde{\beta}) \rangle \cdot \langle \tilde{\boldsymbol{\nu}}, \nabla u_t - \tilde{\beta} \rangle + l \langle \tilde{\boldsymbol{\mu}}, |\nabla u - \tilde{\alpha}|^2 \rangle \Big)\, dx\, dt \\
&= - \int_{Q_T} \big(- \langle \tilde{\boldsymbol{\mu}},\ \sigma(\nabla u) - \sigma(\tilde{\alpha}) \rangle \cdot (\nabla u_t - \nabla \tilde{u}_t) \\
&\quad + l \langle \tilde{\boldsymbol{\mu}}, |\nabla u - \tilde{\alpha}|^2 \rangle \big)\, dx\, dt \\
&\leqq L_\sigma \int_{Q_T} \big(\langle \tilde{\boldsymbol{\mu}}, |\nabla u - \tilde{\alpha}| \rangle \cdot \langle \tilde{\boldsymbol{\nu}}, |\nabla u_t - \tilde{\beta}| \rangle \\
&\quad + l \left\langle \tilde{\boldsymbol{\mu}}, |\nabla u - \tilde{\alpha}|^2 \right\rangle \big)\, dx\, dt
\end{aligned}$$

(now using (3.7) on the first term and (3.9) with $p = 2$ on the second term)

$$\begin{aligned}
&\leqq L_\sigma \int_{Q_T} \left(\int_0^t \langle \tilde{\boldsymbol{\nu}}, |\nabla u_t - \tilde{\beta}| \rangle ds \right) \cdot \langle \tilde{\boldsymbol{\nu}}, |\nabla u_t - \tilde{\beta}| \rangle \\
&\quad + lT \int_{Q_T} \int_0^t \langle \tilde{\boldsymbol{\nu}}, |\nabla u_t - \tilde{\beta}|^2 \rangle ds\, dx\, dt
\end{aligned}$$

(and by Jensen's inequality)

$$\begin{aligned}
&\leqq \left(\frac{L_\sigma^2}{2\varepsilon} + l \right) T \int_{Q_T} \int_0^t \left| \langle \tilde{\boldsymbol{\nu}}, |\nabla u_t - \tilde{\beta}|^2 \rangle \right| dx\, dt \\
&\quad + \frac{\varepsilon}{2} \int_{Q_T} \langle \tilde{\boldsymbol{\nu}}, |\nabla u_t - \tilde{\beta}|^2 \rangle\, dx\, dt
\end{aligned}$$

$$\leqq \left(\frac{L_\sigma^2}{2\varepsilon} + l\right) T \int_0^T \left\| \langle \tilde{\boldsymbol{\nu}}, |\nabla u_t - \tilde{\beta}|^2 \rangle \right\|_{L^1(Q_t)} (t)\, dt + \frac{\varepsilon}{2} \left\| \langle \tilde{\boldsymbol{\nu}}, |\nabla u_t - \tilde{\beta}|^2 \rangle \right\|_{L^1(Q_T)}. \tag{3.11}$$

Choose $\varepsilon = \frac{\kappa}{2}$ and apply Gronwall's inequality on

$$F(T) = \|u_t - \tilde{u}_t\|^2_{L^2(Q_T)} + \int_0^T \|\langle \tilde{\boldsymbol{\nu}}, |\nabla u - \tilde{\beta}|^2 \rangle\|_{L^1(Q_t)}\, dt.$$

Then by (3.11) there is a constant $c > 0$ such that $F'(t) \leqq cF(t)$ so that $F(t) \leqq F(0)\mathrm{e}^{ct} = 0$. Assuming the same initial data for $u, \tilde{u}$, this shows that $u = \tilde{u}$ a.e. in Q_T for every $T \geqq 0$ but also as in Theorem 3.3 this proof also implies well-posedness (1.28). □

Remark 3.9. The measures $\boldsymbol{\xi}, \boldsymbol{\mu}, \boldsymbol{\nu}$ are not expected to be unique in general (only their first moments): there simply is not enough information (or constraints) imposed on these measures, either by the assumptions of the theorem or the method of construction of the solution, to guarantee such uniqueness. It would be desirable have a further criterion to select a unique measure, especially in the case, when the solution is not classical weak but genuinely measure-valued.

4. Proof of the main theorems, Theorems A and B

The proof, in three steps, of the existence of Young measure solutions of (1.1) (Theorem A(i)) is the content of Section 2. The approximate solutions constructed in Steps I, II are shown in Lemma 2.7 to converge to the Young measure solution satisfying (1.22)–(1.24). The estimates in (1.25) and (1.26) are true by Lemmas 2.4 and 2.6 and Corollary 2.5. The initial data are attained as claimed by Corollary 2.8.1. The measure-valued solution obtained above is in fact a classical weak solution under (1.18) by Theorem 3.3 and in Theorem 3.8 it is shown to be unique within a general class of measure-valued solutions.

The existence of Young measure solutions for (1.2) follows from that of (1.1), with the simplification that the measure $\boldsymbol{\nu}$ suffices to represent the solution as τ depends only on ∇u. Under strict convexity of G it follows from Theorem 3.3 that it is a unique classical weak solution and a proof particular to the gradient flow case appears in Lemma 3.5.

5. Appendix

A. Energy dissipation for classical weak and measure-valued solutions

Viscoelasticity. The uniform estimates in (1.25) are obtained by taking limits of approximate solutions constructed by time-discretization. This is in analogy with the classical energy estimates which a weak solution satisfies given in the next lemma.

Lemma A.1 (Classical energy estimates**).** *Assume* $u \in H^1_{\text{loc}}(Q_\infty) \cap L^2(H^1_0(\Omega))$ *is a strong distributional solution (or classical solution) of* (1.1),(1.4) *given in Section 1.2. Then for* $0 \leqq s \leqq t$*,*

$$E(t) + \int_s^t \int_\Omega \gamma |\nabla u_t|^2 \, dx \, dt \leqq E(t) + \int_s^t \int_\Omega \tau(\nabla u, \nabla u_t) \cdot \nabla u_t \, dx \, dt = E(s), \quad \text{(A.1)}$$

where

$$E(t) = \int_\Omega \left(W(\nabla u)(t,x) + \frac{{u_t}^2}{2}(t,x) \right) dx, \quad \text{(A.2)}$$

and for all $T > 0$*,*

$$\int_0^T \int_\Omega u_{tt}^2 \, dx \, dt + \sup_{0 \leqq t \leqq T} \int_\Omega |\nabla u_t|^2 \, dx \leqq C_T, \quad \text{(A.3)}$$

where the constant C_T *depends on* T *and the initial data. If the growth of* G_F *is given by* $G_F(F, \dot{F}) \leqq m|\dot{F}|$ *then the estimate in* (A.1) *is independent of* T*.*

Proof. Standard multiplication of the equation (by u_t for (A.1) and by u_{tt} for (A.3)) and integration provides these estimates. I sketch the proof of (A.3) assuming (A.1). Below K_0 is a generic positive constant depending only on the initial data and C_T is a generic positive constant depending only on T. Boundary terms on $\partial\Omega$ vanish.

$$\int_0^T \int_\Omega u_{tt}^2 \, dx \, dt = -\int_0^T \int_\Omega (\, \sigma(\nabla u) \cdot \nabla u_{tt} + \tau(\nabla u, \nabla u_t) \cdot \nabla u_{tt} \,) \, dx \, dt \stackrel{\text{def}}{\equiv} -\text{I} - \text{II}. \quad \text{(A.4)}$$

Since σ is Lipschitz,

$$|\text{I}| \leqq \left| \int_\Omega \sigma(\nabla u) \cdot \nabla u_t(T) \, dx - K_0 + \int_0^T \int_\Omega \langle \sigma'(\nabla u); (\nabla u_t, \cdot \nabla u_t) \rangle \, dx \, dt \right|$$
$$\leqq \frac{s^2}{2\varepsilon} \|\nabla u\|^2_{L^\infty(L^2)} + \frac{\varepsilon}{2} \|\nabla u_t(T)\|^2_{L^2(\Omega)} + K_0 + L_\sigma \|\nabla u_t\|^2_{L^2(Q_\infty)}.$$

For II we have,

$$\text{II} = \int_\Omega G(\nabla u, \nabla u_t)(T,x) \, dx - K_0 - \int_0^T \int_\Omega G_F(\nabla u, \nabla u_t) \cdot \nabla u_t \, dx \, dt$$

and by the growth of G, G_F

$$c \int_\Omega |\nabla u_t|^2(T) \, dx \leqq \int_\Omega G(\nabla u, \nabla u_t)(T) \, dx$$

and using again (A.1)

$$\int_0^T \int_\Omega G_F(\nabla u, \nabla u_t) \cdot \nabla u_t \, dx \, dt \leqq \frac{mT}{2} \|\nabla u\|_{L^\infty(L^2)} + 2m \|\nabla u_t\|^2_{L^2(Q_\infty)} + \frac{mT|\Omega|}{2} \equiv C_T$$

where clearly the estimate is independent of T if τ satisfies the restricted growth condition above. Then by choosing ε sufficiently small in term I (A.4) implies (A.3). □

For a Young measure solution, the classical result generalizes in the following sense.

Lemma A.2 (Dissipation of energy, independence and a chain rule property**).** *Let* $(u, \boldsymbol{\mu}, \boldsymbol{\xi}, \boldsymbol{\nu})$ *be a solution of* (1.1)*),* (1.3), (1.4)*), found in Section 2. Define (the energy)* $E(t) = \int_\Omega \langle \boldsymbol{\mu}, W \rangle (t, x) \, dx + \frac{1}{2} \|u_t\|^2_{L^2(\Omega)}(t)$*. Then,* $t \mapsto E(t)$ *is an absolutely continuous, decreasing function of time. Moreover,* $\int_\Omega W(\nabla U^h)(t, x) \, dx$ *converges to* $\int_\Omega \langle \mu_{t,x}, W \rangle \, dx$ *weakly in* $W^{1,1}_{\text{loc}}(\mathbb{R}^+)$ *and*

$$\partial_t \int_\Omega \langle \mu_{t,x}, W \rangle \, dx = \int_\Omega (\langle \mu_{t,x}, \sigma \rangle \cdot \nabla u_t(t) - D_{t,x}) \, dx, \tag{A.5}$$

where

$$D_{t,x} = \int_\Omega (\langle \nu_{t,x}, \tau(\alpha) \cdot \alpha \rangle - \langle \nu_{t,x}, \tau \rangle \cdot \langle \nu_{t,x}, id \rangle) \, dx. \tag{A.6}$$

The case where $\int_\Omega D_{t,x} \, dx$ *vanishes is of interest and is discussed in the next section. In that case the chain rule generalizes to the Young measure solution in a natural way,*

$$\partial_t \int_\Omega \langle \mu_{t,x}, W \rangle \, dx = \int_\Omega \langle \mu_{t,x}, \sigma \rangle \cdot \partial_t \langle \mu_{t,x}, id \rangle \, dx$$

(recalling (1.23)*). Also* (A.1) *generalizes naturally to*

$$E(t) + \int_s^t \int_\Omega \gamma \left\langle \boldsymbol{\nu}, \alpha^2 \right\rangle dx \, dt \leqq E(t) + \int_s^t \int_\Omega \langle \boldsymbol{\nu}, \tau \cdot \alpha \rangle \, dx \, dt = E(s) \tag{A.7}$$

for all $0 \leqq s \leqq t$*. When* $\int_0^t \int_\Omega D_{t,x} \, dx = 0$ *we get the better analogue of* (A.1)*, namely,*

$$E(t) + \int_s^t \int_\Omega \langle \boldsymbol{\nu}, \tau \rangle \cdot \langle \boldsymbol{\nu}, id \rangle \, dx \, dt \leqq E(s).$$

Proof. As W is Lipschitz and $\nabla U^h(\cdot, x) \in H^1_{\text{loc}}(\mathbb{R}^+)$ for a.e. $x \in \Omega$, then $W(\nabla U^h)(\cdot, x) \in W^{1,1}_{\text{loc}}(\mathbb{R}^+)$ for such x and $\partial_t W(\nabla U^h) = \sigma(\nabla U^h) \cdot \nabla U^h_t$. Furthermore, $\int_\Omega W(\nabla U^h) \, dx \in W^{1,1}_{\text{loc}}(\mathbb{R}^+)$ (in fact bounded there) and the relation $\partial_t \int_\Omega W(\nabla U^h) \, dx = \int_\Omega \sigma(\nabla U^h) \cdot \nabla U^h_t \, dx$ is justified by dominated convergence. □

I compute now the weak limits of these quantities in terms of the Young measure $\boldsymbol{\mu}$ and using (2.14) and Lemma A.3.

Claim. *The term* $\int_0^t\int_\Omega W(\nabla U^h)(t,x)\,dxds$ *converges to* $\int_0^t\int_\Omega \langle \boldsymbol{\mu}, W\rangle\,dxds$ *weakly in* $W^{1,1}_{\mathrm{loc}}(\mathbb{R}^+)$.

Proof of claim. Notice first that the weak convergence asserted above cannot be deduced by boundedness alone. It is clear that $\int_\Omega W(\nabla U^h)(t,x)\,dx$ converges to $\int_\Omega \langle \boldsymbol{\mu}, W\rangle\, dx$ weakly in L^1_{loc} using the Young measure representation and the fact that u^h, U^h generate the same Young measure, see lemma A.3. Furthermore, $Q^h(t) \overset{\mathrm{def}}{\equiv} \partial_t \int_\Omega W(\nabla U^h)\,dx = \int_\Omega \sigma(\nabla U^h)\cdot\nabla U^h_t\,dx$. By the growth of σ and the integrability of ∇U^h, Q^h is bounded in $L^1_{\mathrm{loc}}(\mathbb{R}^+)$ and thus has a biting weak limit which must coincide with $\partial_t \int_\Omega \langle \boldsymbol{\mu}, W\rangle\, dx$ (as it coincides with it on sets exhausting Ω). This proves the second claim and it remains now to identify this time derivative: the desirable estimate of course is that it is given by $\int_\Omega \langle \boldsymbol{\mu}, \sigma\rangle\cdot\nabla u_t\,dx$; however, there is a correction term, $D_{t,x}$, which will now be identified.

We can compute the same limit also by the approximating weak equation (2.14) applied against ∇U^h_t. Let $W_0 = \int_\Omega W(\nabla u_0)\,dx$. For each $t \geqq 0$,

$$\begin{aligned}
\int_0^t \partial_t \int_\Omega W(\nabla U^h)(t)\,dx\,ds &= \int_\Omega W(\nabla U^h)(t)\,dx - W_0 \\
&= \int_0^t\int_\Omega \sigma(\nabla U^h)\cdot\nabla U^h_t\,dx\,ds \\
&= \int_0^t\int_\Omega \Big(\sigma(\nabla u^h(t-h))\cdot\nabla v^h \\
&\qquad + \Big(\sigma(\nabla U^h) - \sigma(\nabla u^h(t-h))\Big)\cdot\nabla v^h\Big)\,dx\,ds \\
&= -\int_0^t\int_\Omega \Big(\Big\langle \boldsymbol{\nu}^h, \tau\Big\rangle\cdot\nabla v^h + z^h v^h + I^h\Big)\,dx\,ds \\
&\overset{h\to 0}{\longrightarrow} -\int_0^t\int_\Omega (\langle \boldsymbol{\nu}, \tau\cdot id\rangle + u_{tt}u_t)\,dx\,ds \\
&= -\int_0^t\int_\Omega \big(\langle \boldsymbol{\nu}, \tau\rangle\cdot\langle \boldsymbol{\nu}, id\rangle + u_{tt}u_t + D_{t,x}\big)\,dx\,ds \\
&= \int_0^t\int_\Omega \Big(\langle \boldsymbol{\mu}, \sigma\rangle\cdot\nabla u_t\,dx\,ds + \int_0^t\int_\Omega D_{t,x}\Big)\,dx\,ds \\
&\quad \Big(= \int_0^t\int_\Omega \langle \boldsymbol{\mu}, W\rangle\,dx\,ds\Big),
\end{aligned}$$

where the error term $\int_0^t\int_\Omega I^h\,dx\,ds \to 0$ as $h \to 0$ by (2.21) (as σ is Lipschitz) and (2.9).

Now compute the weak limits of the time derivatives:

$$\begin{aligned}
\partial_t \int_\Omega W(\nabla U^h)\,dx &= \int_\Omega \sigma(\nabla U^h)\cdot \nabla U^h_t\,dx \\
&= -\int_\Omega \left(\left\langle \boldsymbol{\nu}^h, \tau\right\rangle \cdot \nabla v^h + z^h v^h + I^h\right)\,dx \\
&\overset{h\to 0}{\rightharpoonup} -\int_\Omega \left(\langle \boldsymbol{\nu}, \tau\rangle \cdot \langle \boldsymbol{\nu}, id\rangle + u_{tt}u_t + D_{t,x}\right)\,dx \\
&\qquad \text{weakly in } L^1_{\text{loc}}(\mathbb{R}^+) \\
&= \int_\Omega \left(\langle \boldsymbol{\mu}, \sigma\rangle \cdot \nabla u_t + D_{t,x}\right)\,dx,
\end{aligned}$$

where the error term $\int_\Omega \mathrm{I}^h\,dx \rightharpoonup 0$ in L^1 weakly by (2.21) and (2.15). This shows the chain rule (A.6) property as asserted.

Notice that as a consequence $\int_0^t\int_\Omega \langle \boldsymbol{\mu}, W\rangle\,dxds \in W^{1,1}_{\text{loc}}(\mathbb{R}^+)$ as a weak limit of a sequence in this space (*a priori* this term belongs only to L^1_{loc}); and hence this is also true for $\int_0^t\int_\Omega \left(\langle \boldsymbol{\mu}, \sigma\rangle \cdot \nabla u_t + D_{t,x}\right)\,dxds$. From this then follows that the map $t \mapsto E(t) \in W^{1,1}_{\text{loc}}(\mathbb{R}^+)$ since also $t \mapsto \|u_t\|_{L^2(\Omega)}(t) \in W^{1,1}_{\text{loc}}(\mathbb{R}^+)$.

It follows now that $E(t)$ is decreasing in t:

$$\begin{aligned}
\frac{d}{dt}E(t) &= \frac{d}{dt}\int_\Omega \left(\langle \boldsymbol{\mu}, \sigma\rangle \cdot \langle \boldsymbol{\nu}, \mathrm{id}\rangle - D_{t,x} + u_{tt}u_t\right)\,dx \\
&= -\int_\Omega \langle \boldsymbol{\nu}, \tau\cdot \mathrm{id}\rangle\,dx \leqq -\gamma\int_\Omega \left\langle \boldsymbol{\nu}, \alpha^2\right\rangle \leqq 0
\end{aligned}$$

and from this together with (1.17) follows (A.7). □

Above, the following has been used:

Lemma A.3. *The sequences $(\nabla U^h)_h$ and $(\nabla u^h)_h$ generate the same Young measure $\boldsymbol{\mu}$. Similarly, $(\nabla V^h)_h$ and $(\nabla v^h)_h$ generate $\boldsymbol{\nu}$.*

Proof. It is easy to see that the L^2 weak limits of $\phi(\nabla U^h)$ and $\phi(\nabla u^h)$ are represented by $\langle \boldsymbol{\mu}, \phi\rangle$ for all ϕ Lipschitz continuous by (2.21) and (2.19). The extension to all functions in $\mathcal{E}^1_0$ is immediate by density; and the extension to functions in $\mathcal{E}^p_0$ with $p > 1$ (which can be at most locally Lipschitz with a constant growing like the $p-1$ power of the size of the local domain) is straightforward by the available estimates. I give two arguments for $p > 1$ (as remarked in Section 1.2 the restriction $p = 2$ in the text is most likely not essential): first observe that the set of locally Lipschitz functions ϕ such that $\sup_{x,y\in B_\rho} |\phi(x) - \phi(y)| \leqq c_\rho |x-y|$ with $c_\rho = \rho^{p-1}$ is dense in $\mathcal{E}^p_0$ so it suffices to show the claim for such ϕ. Indeed, by (2.19) and (2.21), $\phi(\nabla U^h) - \phi(\nabla u^h)(t-h) \leqq ch|\nabla v^h|(t)$ and $c = |\nabla U^h - \nabla u^h(t-h)|^{p-1} \leqq (h|\nabla v^h|)^{p-1}$; integrate over Q_T to show that in the limit the difference is bounded by $ch^p\|\nabla v^h\|^p_{L^p(Q_T)}$ which vanishes by (2.9) or (2.10) and thus the Young measure representation of the two integrals in the limit is the same.

A direct argument for $\phi \in \mathcal{E}_0^p$ is provided by showing that the claim is true for the p-norm of the generating sequences. For simplicity we show this with $p = 2$

$$\int_0^T \int_\Omega |\nabla U^h|^2 \, dx \, dt = \int_0^T \int_\Omega |\nabla u^h|^2 + |\nabla U^h - \nabla u^h|^2 + \nabla u^h \cdot (\nabla U^h - \nabla u^h) \, dx \, dt$$

and the first term converges to $\int_0^T \int_\Omega \left\langle \boldsymbol{\nu}, (\mathrm{id})^2 \right\rangle \, dx \, dt$ while the other two terms are $O(h)$ by (2.19) and the uniform bounds (2.10). □

The gradient flow. As in viscoelasticity, the uniform estimates are provided by energy norm estimates. Define the *energy*

$$E(t) = \int_\Omega \langle \boldsymbol{\nu}, G \rangle \, dx.$$

Then $E \in L^1(\mathbb{R}^+)$, it is decreasing in t and $\lim_{t \downarrow \infty} E(t) = 0$. This is the counterpart of the classical energy estimate with

$$E(t) = \int_\Omega G(\nabla u)(t, x) \, dx$$

which satisfies the same properties.

B. Frame indifference and monotonicity in $\dot{F}$: an example

Assume $n = m = 3$. The following is an example proposed by Dafermos of a frame indifferent tensor τ which satisfies (1.11),

$$\tau(F, \dot{F}) = (\dot{F}F^{-1} + (\dot{F}F^{-1})^*)F^{-1*}, \tag{A.8}$$

then $\tau(F, \dot{F}) = \partial G_{\dot{F}}(F, \dot{F})$ where G is the positive semi-definite quadratic form

$$G(F, \dot{F}) = (det \ F)^{-1} tr \left(\dot{F}F^{-1} + (\dot{F}F^{-1})^* \right)^2. \tag{A.9}$$

This example is motivated by the known general form of frame indifferent tensors τ, namely polynomial in D with coefficients depending only on the principal invariants of D, where D is the symmetric part of $\dot{F}F^{-1}$ (cf. [12,9]). We may use this example to check whether (1.18) is satisfied. The potential G is convex but not strictly convex, in fact it vanishes where $\dot{F}F^{-1}$ is skew symmetric and therefore for such F, $\dot{F}$ the convexity of $G(F, \cdot)$ and monotonicity of τ degenerate: indeed, taking $F = H$ and $\dot{F} - \dot{H} = JF$ with $J + J^* = 0$, the left-hand side of (1.18) is 0.

Ball observed that (1.18) is incompatible with the necessary form of τ provided by Antman in (1.8): choose $F_1(t) = R(t)$, $F_2(t) = Q(t)$ two smooth curves in $SO(3)$ with $R(0) = Q(0) = 1$ (and so $(U, \dot{U}) = (1, 0)$ for F_1, F_2). Then (1.18) implies $(RS - QS) \cdot (\dot{R} - \dot{Q}) \geqq \kappa |\dot{R} - \dot{Q}|^2 - l|R - Q|^2$, impossible when evaluated at $(1, 0)$ since $|R - Q|(0) = 0$ and so the left-hand side vanishes.

This implies that strict convexity of $G(F, \cdot)$ is incompatible with frame indifference. This fact can be seen also directly from (1.7): taking $F = H = Q = 1$, $\dot{F} = -\dot{F}^*$, $\dot{Q} = -\dot{Q}^*$ (1.7) implies that the left-hand side of (1.18) vanishes, $(\tau(1, \dot{F}) - \tau(1, 0)) \cdot \dot{F} = 0$. Non-strict convexity, however, is compatible with frame indifference, as is evident by example (A.8).

Acknowledgements. I thank STUART ANTMAN, JOHN BALL, CONSTANTINE DAFERMOS and DAVID STUART for very helpful discussions. The problem was proposed by S. ANTMAN who also made available early versions of [5, 6] cited below. J. Ball suggested that frame indifference be incorporated which led to the study of the more general problem (1.1). In this respect it was helpful to know the extent to which the imposed assumptions (1.11), (1.18) are compatible with frame indifference, and this question was answered by the example furnished by C. DAFERMOS and the observation of J. BALL described in the text. I thank also the authors of [8] for making available a preliminary version of their preprint and in particular I thank JAN KRISTENSEN for many related discussions. This work was supported by the European Commission TMR Grant ERBFMBICT972343 held at the Mathematical Institute, University of Oxford.

References

1. G. ANDREWS, On the existence of solutions to the equation $u_{tt} = u_{xxt} + \sigma(u_x)_x$, *J. Diff. Eqs.* **35**, 200–231, 1980.
2. G. ANDREWS & J. BALL, Asymptotic behaviour and changes of phase in one-dimensional nonlinear viscoelasticity, *J. Diff. Eqs.* **44**, 306–341, 1982.
3. S. ANTMAN, *Nonlinear problems in elasticity*, Applied Mathematical Sciences, Vol. 107, Springer Verlag, New York 1995.
4. S. ANTMAN & T. SEIDMAN, Quasilinear Hyperbolic-Parabolic Equations of One-Dimensional Viscoelasticity, *J. Diff. Eqs.* **124**, 132–184, 1996.
5. S. ANTMAN, Physically unacceptable viscous stresses, *Z. Angew. Math. Phys.* **49**, 980–988, 1998.
6. S. ANTMAN & H. KOCH, Self-sustained oscillations of nonlinearly viscoelastic layers. To appear.
7. J. BALL, Convexity conditions and existence theorems in nonlinear elasticity, *Arch. Rational Mech. Anal.* **63**, 337–403, 1977.
8. J. BALL, B. KIRCHHEIM & J. KRISTENSEN, Regularity of quasiconvex envelopes. Preprint.
9. P. CIARLET, *Mathematical Elasticity, Vol. 1: Three-Dimensional Elasticity*, Studies in Mathematics and Its Appliations, Vol. 20, North-Holland, Amsterdam 1994.
10. B. COLEMAN & W. NOLL, On the Thermostatics of Continuous Media, *Arch. Rational Mech. Anal.* **4**, 97–128, 1959.
11. C. DAFERMOS, The mixed initial-boundary value problem for the equations of one-dimensional nonlinear viscoelasticity, *J. Diff. Eqs.* **6**, 71–86, 1969.
12. C. DAFERMOS, *Hyperbolic Conservation Laws in Continuum Physics*, Springer Verlag, New York 2000.
13. S. DEMOULINI, Young measure solution for a nonlinear parabolic equation of forward-backward type, *SIAM J. Math. Anal.* **27**, 376–403, 1996.
14. S. DEMOULINI, Young measure solutions for nonlinear evolutionary systems of mixed type, *Ann. Inst. Henri Poincaré, Analyse non linéaire*, **14**, 143–162, 1997.
15. S. DEMOULINI, D. STUART & A. TZAVARAS, Construction of entropy solutions for one-dimensional elastodynamics via time discretisation, *Ann. Inst. Henri Poincaré, Analyse non linéaire* **17**, 711–731, 2000
16. S. DEMOULINI, D. STUART & A. TZAVARAS, A variational approximation scheme for three dimensional elastodynamics with polyconvex energy. Arch RAtional Mechn. Anal., to appear
17. L. EVANS, *Weak convergence methods for nonlinear partial differential equations*, CBMS Vol 74, American Mathematical Society, Providence 1990.
18. G. FRIESECKE & G. DOLZMANN, Implicit time discretization and global existence for a quasilinear evolution equation with nonconvex energy, *SIAM J. Math. Anal.* **28**, 363–380, 1997.
19. C. HAMBURGER, Quasimonotonicty, regularity and duality for nonlinear sytems of partial differential equations, *Annali Mat. pura ed appl.* **69**, 321–354, 1995.

20. D. KINDERLEHRER & P. PEDREGAL, Weak convergence of integrands and the Young measure representation, *SIAM J. Math. Anal.* **23**, 1–19, 1992.
21. D. KINDERLEHRER & P. PEDREGAL, Gradient Young measures generated by sequences in Sobolev spaces, *J. Geom. Anal.* **4**, 59–90 1994.
22. R. LANDES, Quasimonotone versus pseudomonotone, *Proc. Roy. Soc. Edin.* A **126**, 705–717, 1996.
23. J. L. LIONS, *Quelques Méthodes de résolution des problemes aux limites non linéaires*, Études Mathématiques, Dunod, Paris 1969.
24. J. MARSDEN & T. HUGHES, *Mathematical Foundations of Elasticity*, (Prentice-Hall, 1983) Dover, New York 1993.
25. R. PEGO, Phase transitions in one-dimensional nonlinear viscoelasticity, *Arch. Rational Mech. Anal.* **97**, 353–394, 1987.
26. M. POTIER-FERRY, On the Mathematical Foundations of Elastic Stability Theory. I., *Arch. Rational Mech. Anal.* **78**, 55–72, 1982.
27. P. RYBKA, Dynamical modeling of phase transitions by means of viscoelasticity in many dimensions, *Proc. Roy. Soc. Edin.* A **121**, 101–138, 1992.
28. L. TARTAR, Compensated compactness and partial differential equations. In *Nonlinear Analysis and Mechanics*, KNOPS, ed., Vol. IV, Pitman Research Notes in Mathematics, 1979, pp. 136–212.
29. E. ZEIDLER, *Nonlinear Functional Analysis and Its Applications* II A, Springer Verlag, New York 1990.
30. K.-W. ZHANG, *On the Dirichlet problem for a class of quasilinear elliptic systems of partial differential equations in divergence form*, Chern, ed., Vol. 1306, Springer Lecture Notes in Mathematics 1988, pp. 262–277.

Mathematical Institute
University of Oxford
currently at:
CMS / DPMMS
University of Cambridge
Wilberforce Road
Cambridge CB3 OWB
UK
e-mail: S.Demoulini@dpmms.cam.ac.uk

(*Accepted July 1, 2000*)
Published online December 6, 2000 –