QUASI-LINEAR PARABOLIC SYSTEMS IN DIVERGENCE FORM WITH WEAK MONOTONICITY

NORBERT HUNGERBÜHLER

Abstract

We consider the initial and boundary value problem for the quasi-linear parabolic system

$$\frac{\partial u}{\partial t} - \operatorname{div} \sigma\big(x, t, u(x,t), Du(x,t)\big) = f \qquad \text{on } \Omega \times (0,T),$$
$$u(x,t) = 0 \qquad \text{on } \partial\Omega \times (0,T),$$
$$u(x,0) = u_0(x) \quad \text{on } \Omega$$

for a function $u : \Omega \times [0,T) \to \mathbb{R}^m$ with $T > 0$. Here, $f \in L^{p'}(0,T; W^{-1,p'}(\Omega; \mathbb{R}^m))$ for some $p \in (2n/(n+2), \infty)$, and $u_0 \in L^2(\Omega; \mathbb{R}^m)$. We prove existence of a weak solution under classical regularity, growth, and coercivity conditions for σ but with only very mild monotonicity assumptions.

1. Introduction

On a bounded open domain $\Omega \subset \mathbb{R}^n$, we consider the initial and boundary value problem for the quasi-linear parabolic system

$$\frac{\partial u}{\partial t} - \operatorname{div} \sigma\big(x, t, u(x,t), Du(x,t)\big) = f \qquad \text{on } \Omega \times (0,T), \tag{1}$$
$$u(x,t) = 0 \qquad \text{on } \partial\Omega \times (0,T), \tag{2}$$
$$u(x,0) = u_0(x) \quad \text{on } \Omega \tag{3}$$

for a function $u : \Omega \times [0,T) \to \mathbb{R}^m$, $T > 0$. Here, $f \in L^{p'}(0,T; W^{-1,p'}(\Omega; \mathbb{R}^m))$ for some $p \in (2n/(n+2), \infty)$, $u_0 \in L^2(\Omega; \mathbb{R}^m)$, and σ satisfies the conditions (P0)–(P2) below. A feature of the Young measure technique we use is that we can treat a class of problems for which the classical monotone operator methods developed by M. Višik [23], G. Minty [21], F. Browder [5], H. Brézis [3], J.-L. Lions [20], and others do not apply. The reason for this is that σ does not need to satisfy the strict

DUKE MATHEMATICAL JOURNAL
Vol. 107, No. 3, © 2001
Received 22 December 1999. Revision received 10 July 2000.
2000 *Mathematics Subject Classification*. Primary 35K55.

monotonicity condition of a typical Leray-Lions operator. The tool we use in order to prove the needed compactness of approximating solutions is Young measures. The methods are inspired by [6] and [12].

To fix some notation, let $\mathbb{M}^{m\times n}$ denote the real vector space of $m \times n$ matrices equipped with the inner product $M : N = M_{ij}N_{ij}$ (with the usual summation convention).

The following notion of monotonicity plays a role in part of the exposition. Instead of assuming the usual pointwise monotonicity condition for σ, we also use a weaker, integrated version of monotonicity which is called quasi monotonicity (see [6]). The definition is phrased in terms of gradient Young measures. Note, however, that although quasi monotonicity is “monotonicity in integrated form,” the gradient $D\eta$ of a quasi-convex function η is not necessarily quasi-monotone.

Definition 1
A function $\eta : \mathbb{M}^{m\times n} \to \mathbb{M}^{m\times n}$ is said to be strictly p-quasi-monotone if

$$\int_{\mathbb{M}^{m\times n}} \big(\eta(\lambda) - \eta(\bar\lambda)\big) : \big(\lambda - \bar\lambda\big)\, d\nu(\lambda) > 0$$

for all homogeneous $W^{1,p}$-gradient Young measures ν with center of mass $\bar\lambda = \langle \nu, \mathrm{id}\rangle$ which are not a single Dirac mass.

A simple example is the following. Assume that η satisfies the growth condition

$$\big|\eta(F)\big| \le C|F|^{p-1}$$

with $p > 1$ and the structure condition

$$\int_\Omega \big(\eta(F + \nabla\phi) - \eta(F)\big) : \nabla\phi\, dx \ge c \int_\Omega |\nabla\phi|^p\, dx$$

for a constant $c > 0$ and for all $\phi \in C_0^\infty(\Omega)$ and all $F \in \mathbb{M}^{m\times n}$. Then η is strictly p-quasi-monotone. This follows easily from the definition if one uses that for every $W^{1,p}$-gradient Young measure ν there exists a sequence $\{Dv_k\}$ generating ν for which $\{|Dv_k|^p\}$ is equiintegrable (see [9], [14]).

Now, we state our main assumptions.

(P0) (Continuity) We assume that $\sigma : \Omega \times (0, T) \times \mathbb{R}^m \times \mathbb{M}^{m\times n} \to \mathbb{M}^{m\times n}$ is a Carathéodory function; that is, $(x, t) \mapsto \sigma(x, t, u, F)$ is measurable for every $(u, F) \in \mathbb{R}^m \times \mathbb{M}^{m\times n}$, and $(u, F) \mapsto \sigma(x, t, u, F)$ is continuous for almost every $(x, t) \in \Omega \times (0, T)$.

(P1) (Growth and coercivity) There exist $c_1 \ge 0$, $c_2 > 0$, $\lambda_1 \in L^{p'}(\Omega \times (0, T))$, $\lambda_2 \in L^1(\Omega \times (0, T))$, $\lambda_3 \in L^{(p/\alpha)'}(\Omega \times (0, T))$, $0 < \alpha < p$, such that

$$\left|\sigma(x,t,u,F)\right| \leq \lambda_1(x,t) + c_1\big(|u|^{p-1} + |F|^{p-1}\big),$$
$$\sigma(x,t,u,F) : F \geq -\lambda_2(x,t) - \lambda_3(x,t)|u|^{\alpha} + c_2|F|^p.$$

(P2) (Monotonicity) We assume that σ satisfies one of the following conditions:

(a) for all $(x,t) \in \Omega \times (0,T)$ and all $u \in \mathbb{R}^m$, the map $F \mapsto \sigma(x,t,u,F)$ is a C^1-function and is monotone; that is,

$$\big(\sigma(x,t,u,F) - \sigma(x,t,u,G)\big) : (F-G) \geq 0$$

for all $(x,t) \in \Omega \times (0,T)$, $u \in \mathbb{R}^m$, and $F, G \in \mathbb{M}^{m\times n}$;

(b) there exists a function $W : \Omega \times (0,T) \times \mathbb{R}^m \times \mathbb{M}^{m\times n} \to \mathbb{R}$ such that $\sigma(x,t,u,F) = (\partial W/\partial F)(x,t,u,F)$, and $F \mapsto W(x,t,u,F)$ is convex and a C^1-function for all $(x,t) \in \Omega \times (0,T)$ and all $u \in \mathbb{R}^m$;

(c) σ is strictly monotone; that is, σ is monotone, and $(\sigma(x,t,u,F) - \sigma(x,t,u,G)) : (F-G) = 0$ implies $F = G$;

(d) $\sigma(x,t,u,F)$ is strictly p-quasi-monotone in F.

The Carathéodory condition (P0) ensures that $\sigma(x,t,u(x,t),U(x,t))$ is measurable on $\Omega \times (0,T)$ for measurable functions $u : \Omega \times (0,T) \to \mathbb{R}^m$ and $U : \Omega \times (0,T) \to \mathbb{M}^{m\times n}$ (see, e.g., [25]). (P1) states standard growth and coercivity conditions. They are used in the construction of approximate solutions by a Galerkin method and when we pass to the limit. The strict monotonicity condition (c) in (P2) ensures existence of weak solutions of the corresponding parabolic systems by standard methods. However, the main point is that we do not require strict monotonicity or monotonicity in the variables (u, F) in (a), (b), or (d), as it is usually assumed in previous work (see, e.g., [2], [4], [16], [18], [17], [19], and the references therein).

We prove the following result.

THEOREM 2

If σ satisfies the conditions (P0)–(P2) for some $p \in (2n/(n+2), \infty)$, then the parabolic system (1)–(3) has a weak solution $u \in L^p(0,T; W_0^{1,p}(\Omega))$ for every $f \in L^{p'}(0,T;\, W^{-1,p}(\Omega))$ and every $u_0 \in L^2(\Omega)$.

Remark. The result for case (d) in (P2) answers, in particular, a question by J. Frehse [10].

2. Choice of the Galerkin base

Let $s \geq 1 + n(1/2 - 1/p)$. Then $W_0^{s,2}(\Omega) \subset W_0^{1,p}(\Omega)$. For $\zeta \in L^2(\Omega)$, we consider the linear bounded map

$$\phi : W_0^{s,2}(\Omega) \longrightarrow \mathbb{R}, \qquad v \longmapsto (\zeta, v)_{L^2},$$

where $(\cdot,\cdot)_{L^2}$ denotes the inner product of L^2. By the Riesz representation theorem there exists a unique $K\zeta \in W_0^{s,2}(\Omega)$ such that

$$\phi(v) = (\zeta, v)_{L^2} = (K\zeta, v)_{W^{s,2}} \quad \text{for all } v \in W_0^{s,2}(\Omega).$$

The map $L^2 \to L^2$, $\zeta \mapsto K\zeta$, is linear, symmetric, bounded, and (due to the compact embedding $W_0^{s,2}(\Omega) \subset L^2(\Omega)$) compact. Moreover, since

$$(\zeta, K\zeta)_{L^2} = (K\zeta, K\zeta)_{W^{s,2}} \geq 0,$$

the operator K is (strictly) positive. Hence, there exists an L^2-orthonormal base $W := \{w_1, w_2, \dots\}$ of eigenvectors of K and positive real eigenvalues λ_i with $Kw_i = \lambda_i w_i$. This, in particular, means that $w_i \in W_0^{s,2}(\Omega)$ for all i and that, for all $v \in W_0^{s,2}(\Omega)$,

$$\lambda_i (w_i, v)_{W^{s,2}} = (Kw_i, v)_{W^{s,2}} = (w_i, v)_{L^2}. \tag{4}$$

Notice that the functions w_i are therefore orthogonal also with respect to the inner product of $W^{s,2}(\Omega)$. In fact, for $i \neq j$, we get, by choosing $v = w_j$ in (4),

$$0 = \frac{1}{\lambda_i}\big(w_i, w_j\big)_{L^2} = \big(w_i, w_j\big)_{W^{s,2}}.$$

Notice also that, by choosing $v = w_i$ in (4),

$$1 = \|w_i\|_{L^2}^2 = \big(w_i, w_i\big)_{L^2} = \lambda_i \big(w_i, w_i\big)_{W^{s,2}} = \lambda_i \|w_i\|_{W^{s,2}}^2.$$

Thus, $\widetilde{W} = \{\tilde{w}_1, \tilde{w}_2, \dots\}$, with $\tilde{w}_i := \sqrt{\lambda_i}\, w_i$, is an orthonormal set for $W_0^{s,2}(\Omega)$. Actually, $\widetilde{W}$ is a basis for $W_0^{s,2}(\Omega)$. To see this, observe that, for arbitrary $v \in W_0^{s,2}(\Omega)$, the Fourier series

$$s_n(v) := \sum_{i=1}^{n} \big(\tilde{w}_i, v\big)_{W^{s,2}} \tilde{w}_i \longrightarrow \tilde{v} \quad \text{in } W_0^{s,2}(\Omega)$$

converges to some $\tilde{v}$. On the other hand, we have

$$s_n(v) = \sum_{i=1}^{n} (w_i, v)_{L^2} w_i \longrightarrow v \quad \text{in } L^2(\Omega)$$

and, by the uniqueness of the limit, $\tilde{v} = v$.

We need below the L^2-orthonormal projector $P_k : L^2 \to L^2$ onto $\operatorname{span}(w_1, w_2, \dots, w_k)$, $k \in \mathbb{N}$. Of course, the operator norm $\|P_k\|_{\mathscr{L}(L^2, L^2)} = 1$. But notice also that $\|P_k\|_{\mathscr{L}(W^{s,2}, W^{s,2})} = 1$ since, for $u \in W^{s,2}(\Omega)$,

$$P_k u = \sum_{i=1}^{k} (w_i, u)_{L^2} w_i = \sum_{i=1}^{k} \big(\tilde{w}_i, u\big)_{W^{s,2}} \tilde{w}_i.$$

3. Galerkin approximation

We make the following ansatz for approximating solutions of (1)–(3):

$$u_k(x,t) = \sum_{i=1}^{k} c_{ki}(t) w_i(x),$$

where $c_{ki} : [0,T) \to \mathbb{R}$ are supposed to be measurable-bounded functions. Each u_k satisfies the boundary condition (2) by construction in the sense that $u_k \in L^p(0,T; W_0^{1,p}(\Omega))$. We take care of the initial condition (3) by choosing the initial coefficients $c_{ki}(0) := (u_0, w_i)_{L^2}$ such that

$$u_k(\cdot,0) = \sum_{i=1}^{k} c_{ki}(0) w_i(\cdot) \longrightarrow u_0 \quad \text{in } L^2(\Omega) \text{ as } k \longrightarrow \infty. \tag{5}$$

We try to determine the coefficients $c_{ik}(t)$ in such a way that for all $k \in \mathbb{N}$ the system of ordinary differential equations

$$\big(\partial_t u_k, w_j\big)_{L^2} + \int_\Omega \sigma\big(x,t,u_k,Du_k\big) : Dw_j \, dx = \big\langle f(t), w_j \big\rangle \tag{6}$$

(with $j \in \{1,2,\ldots,k\}$) is satisfied in the sense of distributions. In (6), $\langle \cdot,\cdot \rangle$ denotes the dual pairing of $W^{-1,p'}(\Omega)$ and $W_0^{1,p}(\Omega)$. Now, we fix $k \in \mathbb{N}$ for the moment. Let $0 < \varepsilon < T$ and $J = [0,\varepsilon]$. Moreover, we choose $r > 0$ large enough such that the set $B_r(0) \subset \mathbb{R}^k$ contains the vector $(c_{1k}(0),\ldots,c_{kk}(0))$, and we set $K = \overline{B_r(0)}$. Observe that, by (P0), the function

$$F : J \times K \longrightarrow \mathbb{R}^k,$$

$$\big(t,c_1,\ldots,c_k\big) \longmapsto \left(\big\langle f(t), w_j \big\rangle - \int_\Omega \sigma\left(x,t,\sum_{i=1}^{k} c_i w_i, \sum_{i=1}^{k} c_i Dw_i \right) : Dw_j \, dx \right)_{j=1,\ldots,k}$$

is a Carathéodory function. Moreover, each component F_j may be estimated on $J \times K$ by

$$\begin{aligned} \big|F_j\big(t,c_1,\ldots,c_k\big)\big| \le{}& \|f(t)\|_{W^{-1,p'}} \|w_j\|_{W_0^{1,p}} \\ &+ \left(\int_\Omega \left| \sigma\left(x,t,\sum_{i=1}^{k} c_i w_i, \sum_{i=1}^{k} c_i Dw_i \right) \right|^{p'} dx \right)^{1/p'} \\ &\times \left(\int_\Omega \big|Dw_j\big|^p \, dx \right)^{1/p}. \end{aligned} \tag{7}$$

Using the growth condition in (P1), the right-hand side of (7) can be estimated in such a way that

$$\big|F_j\big(t,c_1,\ldots,c_k\big)\big| \le C(r,k) M(t) \tag{8}$$

uniformly on $J \times K$, where $C(r,k)$ is a constant that depends on r and k, and where $M(t) \in L^1(J)$ (independent of j, k, and r). Thus, the Carathéodory existence result on ordinary differential equations (see, e.g., [13]) applied to the system

$$c_j'(t) = F_j\big(t, c_1(t), \ldots, c_k(t)\big), \tag{9}$$

$$c_j(0) = c_{kj}(0) \tag{10}$$

(for $j \in \{1, \ldots, k\}$) ensures existence of a distributional, continuous solution c_j (depending on k) of (9)–(10) on a time interval $[0, \varepsilon')$, where $\epsilon' > 0$, a priori, may depend on k. Moreover, the corresponding integral equation

$$c_j(t) = c_j(0) + \int_0^t F_j\big(\tau, c_1(\tau), \ldots, c_k(\tau)\big)\, d\tau \tag{11}$$

holds on $[0, \epsilon')$. Then $u_k := \sum_{j=1}^k c_j(t) w_j$ is the desired (short-time) solution of (6) with initial condition (5).

Now, we want to show that the local solution constructed above can be extended to the whole interval $[0, T)$ independent of k. As a word of warning we should mention that the solution need not be unique.

The first thing we want to establish is a uniform bound on the coefficients $|c_{ki}(t)|$. Since (6) is linear in w_j, it is allowable to use u_k as a test function in equation (6) in place of w_j. This gives, for an arbitrary time τ in the existence interval,

$$\underbrace{\int_0^\tau \big(\partial_t u_k, u_k\big)_{L^2}\, dt}_{=:I} + \underbrace{\int_0^\tau \int_\Omega \sigma\big(x, t, u_k, Du_k\big) : Du_k\, dx\, dt}_{=:II} = \underbrace{\int_0^\tau \big\langle f(t), u_k\big\rangle}_{=:III}.$$

For the first term, we have

$$I = \frac{1}{2}\big\|u_k(\cdot, \tau)\big\|^2_{L^2(\Omega)} - \frac{1}{2}\big\|u_k(\cdot, 0)\big\|^2_{L^2(\Omega)}.$$

Using the coercivity in (P1) for the second term, we obtain

$$II \geq -\|\lambda_2\|_{L^1(\Omega\times(0,T))} - \|\lambda_3\|_{L^{(p/\alpha)'}(\Omega\times(0,T))} \|u_k\|^\alpha_{L^p(\Omega\times(0,\tau))} + c_2\|u_k\|^p_{L^p(0,\tau;W_0^{1,p}(\Omega))}.$$

For the third term, we finally get

$$III \leq \|f\|_{L^{p'}(0,T;W^{-1,p'}(\Omega))} \|u_k\|_{L^p(0,\tau;W_0^{1,p}(\Omega))}.$$

The combination of these three estimates gives

$$\big|\big(c_{ki}(\tau)\big)_{i=1,\ldots,k}\big|^2_{\mathbb{R}^k} = \big\|u_k(\cdot, \tau)\big\|^2_{L^2(\Omega)} \leq \bar{C}$$

for a constant $\bar{C}$ that is independent of τ (and of k).

Now, let

$$\Lambda := \big\{t \in [0, T) : \text{ there exists a weak solution of (9)–(10) on } [0, t)\big\}.$$

Λ is nonempty since we proved local existence above.

Moreover, Λ is an open set. To see this, let $t \in \Lambda$, and let $0 < \tau_1 < \tau_2 \le t$. Then, by (11) and (8), we have

$$\left|c_{kj}(\tau_1) - c_{kj}(\tau_2)\right| \le \int_{\tau_1}^{\tau_2} \left|F_j\big(\tau, c_{k1}(\tau), \dots, c_{kk}(\tau)\big)\right| d\tau \le C\big(\bar{C}, k\big) \int_{\tau_1}^{\tau_2} \left|M(t)\right| d\tau.$$

Since $M \in L^1(0, T)$, this implies that $\tau \mapsto c_{kj}(\tau)$ is uniformly continuous. Thus, we can restart to solve (6) at time t with initial data $\lim_{\tau \nearrow t} u_k(\tau)$ and hence get a solution of (9)–(10) on $[0, t + \epsilon)$.

Finally, we prove that Λ is also closed. To see this, we consider a sequence $\tau_i \nearrow t$, $\tau_i \in \Lambda$. Let $c_{kj,i}$ denote the solution of (9)–(10) we constructed on $[0, \tau_i]$, and define

$$\tilde{c}_{kj,i}(\tau) := \begin{cases} c_{kj,i}(\tau) & \text{if } \tau \in [0, \tau_i], \\ c_{kj,i}(\tau_i) & \text{if } \tau \in (\tau_i, t). \end{cases}$$

The sequence $\{c_{kj,i}\}_i$ is bounded and equicontinuous on $[0, t)$, as seen above. Hence, by the Arzela-Ascoli theorem, a subsequence (again denoted by $\tilde{c}_{kj,i}(\tau)$) converges uniformly in τ on $[0, t)$ to a continuous function $c_{kj}(\tau)$. Using the Lebesgue convergence theorem in (11), it is now easy to see that $c_{kj}(\tau)$ solves (9) on $[0, t)$. Hence, $t \in \Lambda$, and thus Λ is indeed closed. And as claimed, it follows that $\Lambda = [0, T)$.

4. Compactness of the Galerkin approximation

By testing equation (6) by u_k in place of w_j, we obtain, as in Section 3, that the sequence $\{u_k\}_k$ is bounded in

$$L^\infty\big(0, T; L^2(\Omega)\big) \cap L^p\big(0, T; W_0^{1,p}(\Omega)\big).$$

Therefore, by extracting a suitable subsequence that is again denoted by u_k, we may assume

$$\begin{aligned} u_k &\overset{*}{\rightharpoonup} u \quad \text{in } L^\infty\big(0, T; L^2(\Omega)\big), \\ u_k &\rightharpoonup u \quad \text{in } L^p\big(0, T; W_0^{1,p}(\Omega)\big). \end{aligned}$$

At this point, the idea is to use J.-P. Aubin's lemma in order to prove compactness of the sequence $\{u_k\}$ in an appropriate space. Technically, this is achieved by the following lemma, which is slightly more flexible than, for example, the version in [20, Chapter 1, Section 5.2] or in [22].

LEMMA 3

Let B, B_0, and B_1 be Banach spaces, B_0 and B_1 reflexive. Let $i : B_0 \to B$ be a compact linear map, and let $j : B \to B_1$ be an injective-bounded linear operator. For T finite and $1 < p_i < \infty$, $i = 0, 1$,

$$W := \left\{ v \mid v \in L^{p_0}\big(0,T;B_0\big),\, \frac{d}{dt}(j \circ i \circ v) \in L^{p_1}\big(0,T;B_1\big) \right\}$$

is a Banach space under the norm $\|v\|_{L^{p_0}(0,T;B_0)} + \|j \circ i \circ v\|_{L^{p_1}(0,T;B_1)}$. *Then if* $V \subset W$ *is bounded, the set* $\{i \circ v \mid v \in V\}$ *is precompact in* $L^{p_0}(0,T;B)$.

The proof of Lemma 3 is given in Appendix A.

Now, we apply Lemma 3 to the following case: $B_0 := W_0^{1,p}(\Omega)$, $B := L^q(\Omega)$ (for some q with $2 < q < p^* := np/(n-p)$ if $p < n$ and $2 < p < \infty$ if $p \geq n$), and $B_1 := (W_0^{s,2}(\Omega))'$. Since we assume that $p \in (2n/(n+2), \infty)$, we have the following chain of continuous injections:

$$B_0 \overset{i}{\hookrightarrow} B \overset{i_0}{\hookrightarrow} L^2(\Omega) \overset{\gamma}{\cong} \big(L^2(\Omega)\big)' \overset{i_1}{\hookrightarrow} B_1. \tag{12}$$

Here, $L^2(\Omega) \cong (L^2(\Omega))'$ is the canonical isomorphism γ of the Hilbert space $L^2(\Omega)$ and its dual. For $i : B_0 \to B$ we take simply the injection mapping, and for $j : B \to B_1$ we take the concatenation of injections and the canonical isomorphism given by (12), that is, $j := i_1 \circ \gamma \circ i_0$.

Then, as stated at the beginning of this section, $\{u_k\}_k$ is a bounded sequence in $L^p(0,T;B_0)$. Observe that the time derivative $(d/dt)(j \circ i \circ u_k)$ is, according to (6), given by

$$\frac{d}{dt}\big(j \circ i \circ u_k\big) : [0,T) \longrightarrow B_1 = \big(W_0^{s,2}(\Omega)\big)',$$
$$t \longmapsto \left(\phi \longmapsto -\int_\Omega \sigma\big(x,t,u_k,Du_k\big) : D(P_k\phi)\,dx + \big\langle f(t), P_k\phi\big\rangle\right).$$

(We recall that the projection operators P_k are self-adjoint with respect to the L^2 inner product.) Now we claim that indeed $\{\partial_t j \circ i \circ u_k\}_k$ is a bounded sequence in $L^{p'}(0,T;(W_0^{s,2}(\Omega))')$. Namely, we have by the growth condition in (P1) that

$$\begin{aligned} &\left| -\int_0^T \int_\Omega \sigma\big(x,t,u_k,Du_k\big) : D(P_k\phi)\,dx\,dt + \big\langle f, P_k\phi\big\rangle \right| \\ &\quad \leq C\Big(\|\lambda_1\|_{L^{p'}((0,T)\times\Omega)} + \|u_k\|^{p-1}_{L^p(0,T;W_0^{1,p}(\Omega))} \\ &\qquad + \|f\|_{L^{p'}(0,T;W^{-1,p'}(\Omega))}\Big)\|P_k\phi\|_{L^p(0,T;W_0^{1,p}(\Omega))}, \end{aligned} \tag{13}$$

and the claim follows since

$$\|P_k\phi\|_{L^p(0,T;W_0^{1,p}(\Omega))} \leq \|P_k\phi\|_{L^p(0,T;W_0^{s,2}(\Omega))} \leq \|\phi\|_{L^p(0,T;W_0^{s,2}(\Omega))}.$$

In the last inequality we used the remark at the end of Section 2.

Hence, from Lemma 3, we may conclude that there exists a subsequence, which we still denote by u_k, having the property that

$$u_k \longrightarrow u \quad \text{in } L^p\big(0, T; L^q(\Omega)\big) \text{ for all } q < p^* \text{ and in measure on } \Omega \times (0, T).$$

Notice that, in order to have the strong convergence simultaneously for all $q < p^*$, the usual diagonal sequence procedure applies.

For further use, we note that from (13) we can conclude that $\partial_t u$ (or rather $\partial_t(j \circ i \circ u)$) is an element of the space $L^{p'}(0, T; W^{-1,p'}(\Omega))$. (This follows easily from the fact that the set $\{\phi \in L^p(0, T; W_0^{1,p}(\Omega)) : \exists k \in \mathbb{N} \text{ such that } P_k\phi = \phi\}$ is dense in $L^p(0, T; W_0^{1,p}(\Omega))$, as proved in Appendix B.) See also Appendix C.

Recall that the space

$$\Big\{u \in L^p\big(0, T; W_0^{1,p}(\Omega)\big) : \partial_t(j \circ i \circ u) \in L^{p'}\big(0, T; W^{-1,p'}(\Omega)\big)\Big\}$$

is continuously embedded in

$$C^0\big([0, T]; L^2(\Omega)\big).$$

Hence, we have that $u \in C^0([0, T]; L^2(\Omega))$ after possible modification of u on a Lebesgue zero-set of $[0, T]$. This gives $u(t, \cdot) \in L^2(\Omega)$ a pointwise interpretation for all $t \in [0, T]$ and allows in particular the statement that $u(t, \cdot)$ attains its initial value

$$u(\cdot, 0) = u_0 \tag{14}$$

continuously in $L^2(\Omega)$ (see Appendix C for a proof of (14)).

At this point we mention that, in the case when σ depends only on t and in a strictly quasi-monotonic way on Du, a quite simple proof gives the existence result. This is carried out in Appendix D. However, to obtain the general result stated in Theorem 2, some more work is needed in order to pass to the limit.

5. The Young measure generated by the Galerkin approximation

The sequence (or at least a subsequence) of the gradients Du_k generates a Young measure $\nu_{(x,t)}$, and since u_k converges in measure to u on $\Omega \times (0, T)$, the sequence (u_k, Du_k) generates the Young measure $\delta_{u(x,t)} \otimes \nu_{(x,t)}$ (see, e.g., [11]). Now, we collect some facts about the Young measure ν in the following proposition.

PROPOSITION 4

The Young measure $\nu_{(x,t)}$ generated by the sequence $\{Du_k\}_k$ has the following properties:

(i) $\nu_{(x,t)}$ *is a probability measure on* $\mathbb{M}^{m\times n}$ *for almost all* $(x, t) \in \Omega \times (0, T)$*;*

(ii) $\nu_{(x,t)}$ *satisfies* $Du(x, t) = \langle \nu_{(x,t)}, \mathrm{id}\rangle$ *for almost every* $(x, t) \in \Omega \times (0, T)$*;*

(iii) $\nu_{(x,t)}$ *has finite pth moment for almost all* $(x, t) \in \Omega \times (0, T)$*;*

(iv) $\nu_{(x,t)}$ *is a homogeneous* $W^{1,p}$*-gradient Young measure for almost all* $(x, t) \in \Omega \times (0, T)$.

Proof
(i) The first observation is simple. To see that $\nu_{(x,t)}$ is a probability measure on $\mathbb{M}^{m\times n}$ for almost all $(x,t)\in\Omega\times(0,T)$, it suffices to recall the fact that Du_k is a bounded sequence in $L^1(\Omega\times(0,T))$ and to use the fundamental theorem in [1].

(ii) As we stated at the beginning of Section 4, $\{Du_k\}_k$ is bounded in $L^p(0,T;L^p(\Omega))$, and we may assume that

$$Du_k \rightharpoonup Du \quad \text{in } L^p\big(0,T;L^p(\Omega)\big).$$

On the other hand, it follows that the sequence $\{Du_k\}_k$ is equiintegrable on $\Omega\times(0,T)$, and hence, by the Dunford-Pettis theorem (see, e.g., [7]), the sequence is sequentially weakly precompact in $L^1(\Omega\times(0,T))$, which implies that

$$Du_k \rightharpoonup \langle\nu_{(x,t)},\mathrm{id}\rangle \quad \text{in } L^1\big(0,T;L^1(\Omega)\big).$$

Hence, we have $Du(x,t)=\langle\nu_{(x,t)},\mathrm{id}\rangle$ for almost every $(x,t)\in\Omega\times(0,T)$.

(iii) The next thing we have to check is that $\nu_{(x,t)}$ has finite pth moment for almost all $(x,t)\in\Omega\times(0,T)$. To see this, we choose a cutoff function $\eta\in C_0^\infty(B_{2\alpha}(0);\mathbb{R}^m)$ with $\eta=\mathrm{id}$ on $B_\alpha(0)$ for some $\alpha>0$. Then the sequence

$$D(\eta\circ u_k)=(D\eta)(u_k)Du_k$$

generates a probability Young measure $\nu^\eta_{(x,t)}$ on $\Omega\times(0,T)$ with finite pth moment; that is,

$$\int_{\mathbb{M}^{m\times n}}|\lambda|^p\,d\nu^\eta_{(x,t)}(\lambda)<\infty$$

for almost all $(x,t)\in\Omega\times(0,T)$. Now, for $\phi\in C_0^\infty(\mathbb{M}^{m\times n})$ we have

$$\phi\big(D(\eta\circ u_k)\big)\rightharpoonup\big\langle\nu^\eta_{(x,t)},\phi\big\rangle=\int_{\mathbb{M}^{m\times n}}\phi(\lambda)\,d\nu^\eta_{(x,t)}(\lambda)$$

weakly in $L^1(\Omega\times(0,T))$. Rewriting the left-hand side, we have also (see, e.g., [11])

$$\phi\big((D\eta)(u_k)Du_k\big)\rightharpoonup\int_{\mathbb{M}^{m\times n}}\phi\big(D\eta(u(x,t))\lambda\big)\,d\nu_{(x,t)}(\lambda).$$

Hence,

$$\nu^\eta_{(x,t)}=\nu_{(x,t)} \quad \text{if } |u(x,t)|<\alpha.$$

Since α was arbitrary, it follows that indeed $\nu_{(x,t)}$ has finite pth moment for almost all $(x,t)\in\Omega\times(0,T)$.

(iv) Finally, we have to show that $\{\nu_{(x,t)}\}_{x\in\Omega}$ is for almost all $t\in(0,T)$ a $W^{1,p}$-gradient Young measure. To see this, we take a quasi-convex function q on $\mathbb{M}^{m\times n}$ with $q(F)/|F|\to1$ as $F\to\infty$. Then we fix $x\in\Omega$, $\delta\in(0,1)$ and use inequality (1.21) from [15, Lemma 1.6] with u replaced by $u_k(x,t)$, with $a := u(x,t) - Du(x,t)x$, and with $X := Du(x,t)$. Furthermore, we choose $r > 0$ such that $B_r(x) \subset \Omega$. Observe that the singular part of the distributional gradient vanishes for u_k and that, after integrating the inequality over the time interval $[t_0-\epsilon, t_0+\epsilon] \subset (0,T)$, we get

$$\begin{aligned}\int_{t_0-\epsilon}^{t_0+\epsilon} &\int_{B_r(x)} q\big(Du_k(y,t)\big)\,dy\,dt \\ &+ \frac{1}{(1-\delta)r}\int_{t_0-\epsilon}^{t_0+\epsilon}\int_{B_r(x)\setminus B_{\delta r}(x)} \big|u_k(y,t) - u(x,t) - Du(x,t)(y-x)\big|\,dy\,dt \\ &\qquad \geq \big|B_{\delta r}(x)\big| \int_{t_0-\epsilon}^{t_0+\epsilon} q\big(Du(x,t)\big)\,dt.\end{aligned}$$

Letting k tend to infinity in the inequality above, we obtain

$$\begin{aligned}\int_{t_0-\epsilon}^{t_0+\epsilon} &\int_{B_r(x)} \int_{\mathbb{M}^{m\times n}} q(\lambda)\,d\nu_{(y,t)}(\lambda)\,dy\,dt \\ &+ \frac{1}{(1-\delta)r}\int_{t_0-\epsilon}^{t_0+\epsilon}\int_{B_r(x)\setminus B_{\delta r}(x)} \big|u(y,t) - u(x,t) + Du(x,t)(y-x)\big|\,dy\,dt \\ &\qquad \geq \big|B_{\delta r}(x)\big| \int_{t_0-\epsilon}^{t_0+\epsilon} q\big(Du(x,t)\big)\,dt.\end{aligned}$$

Now, we let $\epsilon \to 0$ and $r \to 0$ and use the differentiability properties of Sobolev functions (see, e.g., [8]) and obtain that, for almost all $(x, t_0) \in \Omega \times (0,T)$,

$$\int_{\mathbb{M}^{m\times n}} q(\lambda)\,d\nu_{(x,t_0)}(\lambda) \geq \frac{\big|B_{\delta r}(x)\big|}{\big|B_r(x)\big|} q\big(Du(x,t_0)\big).$$

Since $\delta \in (0,1)$ was arbitrary, we conclude that Jensen's inequality holds true for q and the measure $\nu_{(x,t)}$ for almost all $(x,t) \in \Omega \times (0,T)$. Using the characterization of $W^{1,p}$-gradient Young measures of [14] (e.g., in the form of [15, Theorem 8.1]), we conclude that in fact $\{\nu_{x,t}\}_{x\in\Omega}$ is a $W^{1,p}$-gradient Young measure on Ω for almost all $t \in (0,T)$. By the localization principle for gradient Young measures, we conclude then that $\nu_{(x,t)}$ is a homogeneous $W^{1,p}$-gradient Young measure for almost all $(x,t) \in \Omega \times (0,T)$. □

6. A parabolic div-curl inequality

In this section, we prove a parabolic version of a "div-curl lemma" (see also [6, Lemma 11]), which is the key ingredient for passing to the limit in the approximating equations and for proving that the weak limit u of the Galerkin approximations u_k is indeed a solution of (1)–(3).

Let us consider the sequence

$$I_k := \Big(\sigma\big(x,t,u_k,Du_k\big) - \sigma\big(x,t,u,Du\big)\Big) : \big(Du_k - Du\big),$$

and let us prove that its negative part I_k^- is equiintegrable on $\Omega \times (0,T)$. To do this, we write I_k^- in the form

$$\begin{aligned} I_k &= \sigma\big(x,t,u_k,Du_k\big) : Du_k - \sigma\big(x,t,u_k,Du_k\big) : Du \\ &\quad - \sigma\big(x,t,u,Du\big) : Du_k + \sigma\big(x,t,u,Du\big) : Du =: II_k + III_k + IV_k + V_k. \end{aligned}$$

The sequences II_k^- and V_k^- are easily seen to be equiintegrable by the coercivity condition in (P1). Then, to see equiintegrability of the sequence III_k, we take a measurable subset $S \subset \Omega \times (0,T)$ and write

$$\begin{aligned} &\int_S \big|\sigma\big(x,t,u_k,Du_k\big) : Du\big|\,dx\,dt \\ &\quad\le \left(\int_S \big|\sigma\big(x,t,u_k,Du_k\big)\big|^{p'}\,dx\,dt\right)^{1/p'} \left(\int_S |Du|^p\,dx\,dt\right)^{1/p} \\ &\quad\le C\left(\int_S \Big(\big|\lambda_1(x,t)\big|^{p'} + |u_k|^p + \big|Du_k\big|^p\Big)\,dx\,dt\right)^{1/p'} \left(\int_S |Du|^p\,dx\,dt\right)^{1/p}. \end{aligned}$$

The first integral is uniformly bounded in k (see Section 4). The second integral is arbitrarily small if the measure of S is chosen small enough. A similar argument gives the equiintegrability of the sequence IV_k.

Having established the equiintegrability of I_k^-, it follows by the Fatou lemma [6, Lemma 6] that

$$\begin{aligned} X &:= \liminf_{k\to\infty} \int_0^T \int_\Omega I_k\,dx\,dt \\ &\ge \int_0^T \int_\Omega \int_{\mathbb{M}^{m\times n}} \sigma\big(x,t,u,\lambda\big) : (\lambda - Du)\,d\nu_{(x,t)}(\lambda)\,dx\,dt. \end{aligned} \tag{15}$$

On the other hand, we now see that $X \le 0$. According to Mazur's theorem (see, e.g., [24, Theorem 2, p. 120]), there exists a sequence v_k in $L^p(0,T;W_0^{1,p}(\Omega))$ where each v_k is a convex linear combination of $\{u_1,\dots,u_k\}$ such that $v_k \to u$ in $L^p(0,T;W_0^{1,p}(\Omega))$. In particular, $v_k(t,\cdot) \in \operatorname{span}(w_1,w_2,\dots,w_k)$ for all $t \in [0,T]$.

Now, we have

$$\begin{aligned} X &= \liminf_{k\to\infty} \int_0^T \int_\Omega \sigma\big(x,t,u_k,Du_k\big) : \big(Du_k - Du\big)\,dx\,dt \\ &= \liminf_{k\to\infty} \Bigg(\int_0^T \int_\Omega \sigma\big(x,t,u_k,Du_k\big) : \big(Du_k - Dv_k\big)\,dx\,dt \end{aligned}$$

$$
\begin{aligned}
&\qquad + \int_0^T \int_\Omega \sigma\big(x,t,u_k,Du_k\big) : \big(Dv_k - Du\big)\,dx\,dt\Bigg) \\
&\le \liminf_{k\to\infty} \Bigg(\bigg(\int_0^T \int_\Omega \big|\sigma\big(x,t,u_k,Du_k\big)\big|^{p'}\,dx\,dt\bigg)^{1/p'} \|v_k - u\|_{L^p(0,T;W^{1,p}(\Omega))} \\
&\qquad + \big\langle f, u_k - v_k\big\rangle - \int_0^T \int_\Omega (u_k - v_k)\partial_t u_k\,dx\,dt\Bigg). \qquad (16)
\end{aligned}
$$

Observe that $u_k - v_k \in \operatorname{span}(w_1, w_2, \ldots, w_k)$, which allows us to use (6) in the inequality above. The first factor in the first term in (16),

$$
\bigg(\int_0^T \int_\Omega \big|\sigma\big(x,t,u_k,Du_k\big)\big|^{p'}\,dx\,dt\bigg)^{1/p'},
$$

is uniformly bounded in k by the growth condition in (P1) and the bound for u_k in $L^p(0,T;W^{1,p}(\Omega))$ (see Section 4). The second factor,

$$
\|v_k - u\|_{L^p(0,T;W^{1,p}(\Omega))},
$$

converges to zero for $k \to \infty$ by construction of the sequence v_k. Hence, the first term in (16) vanishes in the limit.

The second term in (16),

$$
\big\langle f, u_k - v_k\big\rangle,
$$

converges to zero since $u_k - v_k \rightharpoonup 0$ in $L^p(0,T;W^{1,p}(\Omega))$.

Finally, for the last term in (16), we have

$$
\begin{aligned}
&-\int_0^T \int_\Omega \big(u_k - v_k\big)\partial_t u_k\,dx\,dt \\
&\qquad = -\int_0^T \int_\Omega \frac{1}{2}\partial_t u_k^2\,dx\,dt + \int_0^T \int_\Omega v_k \partial_t u_k\,dx\,dt \qquad (17) \\
&\qquad = -\frac{1}{2}\big\|u_k(\cdot,T)\big\|_{L^2(\Omega)}^2 + \frac{1}{2}\big\|u_k(\cdot,0)\big\|_{L^2(\Omega)}^2 + \int_0^T \int_\Omega v_k \partial_t u_k\,dx\,dt.
\end{aligned}
$$

Concerning the last term in (17), we claim that for $k \to \infty$ we have

$$
\int_0^T \int_\Omega v_k \partial_t u_k\,dx\,dt \longrightarrow \int_0^T \int_\Omega u\,\partial_t u\,dx\,dt = \frac{1}{2}\big\|u(\cdot,T)\big\|_{L^2(\Omega)}^2 - \frac{1}{2}\|u_0\|_{L^2(\Omega)}^2. \qquad (18)
$$

To see this, let $\epsilon > 0$ be given. Then there exists M such that for all $l \ge m \ge M$ we have

(i) $|\int_0^T \int_\Omega (u - v_m)\partial_t u\,dx\,dt| \le \epsilon$. This is possible since $\partial_t(j \circ i \circ u) \in L^{p'}(0,T;W^{-1,p'}(\Omega))$ and $v_m \to u$ in $L^p(0,T;W_0^{1,p}(\Omega))$.

(ii) $|\int_0^T \int_\Omega (v_l - v_m)\partial_t u_l \, dx \, dt| \leq \epsilon$. This is possible by (13) since $v_l - v_m \in$ span $(w_1, \ldots, w_l)$ for all fixed $t \in (0, T)$.

Now, we fix $m \geq M$ and choose $m_0 \geq m$ such that, for all $l \geq m_0$,

$$\left| \int_0^T \int_\Omega v_m \big(\partial_t u - \partial_t u_l\big) \, dx \, dt \right| \leq \epsilon.$$

This is possible since $\partial_t u_l \overset{*}{\rightharpoonup} \partial_t u$ in $L^{p'}(0, T; (W_0^{s,2}(\Omega))')$. Combination yields, for all $l = l(\epsilon) \geq m_0(\epsilon)$,

$$\begin{aligned}
&\left| \int_0^T \int_\Omega v_l \partial_t u_l \, dx \, dt - \int_0^T \int_\Omega u \partial_t u \, dx \, dt \right| \\
&\qquad \leq \left| \int_0^T \int_\Omega \big(v_l - v_m\big)\partial_t u_l \, dx \, dt \right| + \left| \int_0^T \int_\Omega v_m \big(\partial_t u_l - \partial_t u\big) \, dx \, dt \right| \\
&\qquad\quad + \left| \int_0^T \int_\Omega \big(v_m - u\big)\partial_t u \, dx \, dt \right| \leq 3\epsilon.
\end{aligned}$$

This establishes (18). On the other hand, since $\{u_k\}_k$ is bounded in $L^\infty(0, T; L^2(\Omega))$, we have (after extraction of a further subsequence if necessary) that $u_k(\cdot, T) \rightharpoonup u(\cdot, T)$ in $L^2(\Omega)$ (see Appendix C for a proof). Hence,

$$\liminf_{k\to\infty} \big\| u_k(\cdot, T) \big\|_{L^2(\Omega)} \geq \big\| u(\cdot, T) \big\|_{L^2(\Omega)}. \tag{19}$$

By construction of u_k, we also have

$$\lim_{k\to\infty} \big\| u_k(\cdot, 0) \big\|_{L^2(\Omega)} = \|u_0\|_{L^2(\Omega)}. \tag{20}$$

Using (19), (20), and (18) in (17), we conclude

$$\liminf_{k\to\infty} - \int_0^T \int_\Omega (u_k - v_k)\partial_t u_k \, dx \, dt \leq 0.$$

This establishes $X \leq 0$, and we infer from (15) that the following "div-curl inequality" holds.

LEMMA 5

The Young measure $\nu_{(x,t)}$ generated by the gradients Du_k of the Galerkin approximations u_k has the property that

$$\int_0^T \int_\Omega \int_{\mathbb{M}^{m\times n}} \big(\sigma\big(x, t, u, \lambda\big) - \sigma\big(x, t, u, Du\big)\big) : \big(\lambda - Du\big) \, d\nu_{(x,t)}(\lambda) \, dx \, dt \leq 0. \tag{21}$$

7. Passage to the limit

We start with the easiest case.

Case (d)
Suppose that $\nu_{(x,t)}$ is not a Dirac mass on a set $(x,t) \in M \subset \Omega \times (0,T)$ of positive Lebesgue measure $|M| > 0$. Then, by the strict p–quasi monotonicity of $\sigma(x,t,u,\cdot)$ and by the fact that $\nu_{(x,t)}$ is a homogeneous $W^{1,p}$-gradient Young measure (see Section 5) for almost all $(x,t) \in \Omega \times (0,T)$, we have, for a.e. $(x,t) \in M$,

$$\int_{\mathbb{M}^{m\times n}} \sigma\big(x,t,u,\lambda\big) : \lambda \, d\nu_{(x,t)}(\lambda) > \int_{\mathbb{M}^{m\times n}} \sigma\big(x,t,u,\lambda\big)\, d\nu_{(x,t)}(\lambda) : \underbrace{\int_{\mathbb{M}^{m\times n}} \lambda \, d\nu_{(x,t)}(\lambda)}_{=\,Du(x,t)}.$$

Hence, by integrating over $\Omega \times (0,T)$, we get, together with Lemma 5,

$$\begin{aligned}\int_0^T \int_\Omega \int_{\mathbb{M}^{m\times n}} & \sigma\big(x,t,u,\lambda\big)\, d\nu_{(x,t)}(\lambda) : Du(x,t)\, dx\, dt \\ &\geq \int_0^T \int_\Omega \int_{\mathbb{M}^{m\times n}} \sigma\big(x,t,u,\lambda\big) : \lambda \, d\nu_{(x,t)}(\lambda)\, dx\, dt \\ &> \int_0^T \int_\Omega \int_{\mathbb{M}^{m\times n}} \sigma\big(x,t,u,\lambda\big)\, d\nu_{(x,t)}(\lambda) : Du(x,t)\, dx\, dt,\end{aligned}$$

which is a contradiction. Hence, we have $\nu_{(x,t)} = \delta_{Du(x,t)}$ for almost every $(x,t) \in \Omega \times (0,T)$. From this it follows that $Du_k \to Du$ on $\Omega \times (0,T)$ in measure for $k \to \infty$ (see, e.g., [11]), and thus $\sigma(x,t,u_k,Du_k) \to \sigma(x,t,u,Du)$ almost everywhere on $\Omega \times (0,T)$ (up to extraction of a further subsequence). Since, by the growth condition in (P1), $\sigma(x,t,u_k,Du_k)$ is equiintegrable, it follows that $\sigma(x,t,u_k,Du_k) \to \sigma(x,t,u,Du)$ in $L^1(\Omega \times (0,T))$ by the Vitali convergence theorem. Now, we take a test function $w \in \cup_{i\in\mathbb{N}} \operatorname{span}(w_1,\dots,w_i)$ and $\phi \in C_0^\infty([0,T])$ in (6) and integrate over the interval $(0,T)$ and pass to the limit $k \to \infty$. The resulting equation is

$$\int_0^T\int_\Omega \partial_t u(x)\phi(t)w(x)\, dx\, dt + \int_0^T\int_\Omega \sigma\big(x,t,u,Du\big) : Dw(x)\phi(t)\, dx\, dt = \big\langle f, \phi w\big\rangle$$

for arbitrary $w \in \cup_{i\in\mathbb{N}} \operatorname{span}(w_1,\dots,w_i)$ and $\phi \in C^\infty([0,T])$. By density of the linear span of these functions in $L^p(0,T;W^{1,p}(\Omega))$, this proves that u is in fact a weak solution. Hence, the theorem follows in case (d).

Now, we prepare the proof of Theorem 2 in the remaining cases as follows. Observe that the integrand in (21) is nonnegative by monotonicity. Thus, it follows from Lemma 5 that the integrand must vanish almost everywhere with respect to the product measure $d\nu_{(x,t)} \otimes dx \otimes dt$. Hence, we have that, for almost all $(x,t) \in \Omega \times (0,T)$,

$$\big(\sigma\big(x,t,u,\lambda\big) - \sigma\big(x,t,u,Du\big)\big) : \big(\lambda - Du\big) = 0 \quad \text{on } \operatorname{spt} \nu_{(x,t)} \tag{22}$$

and thus

$$\operatorname{spt}\nu_{(x,t)} \subset \big\{\lambda \mid \big(\sigma\big(x,t,u,\lambda\big)-\sigma\big(x,t,u,Du\big)\big):\big(\lambda-Du\big)=0\big\}. \tag{23}$$

Now, we proceed with the proof in the single cases (a), (b), and (c) of (P2). We start with the simplest case (c).

Case (c)

By strict monotonicity, it follows from (22) or (23) that $\nu_{(x,t)}=\delta_{Du(x,t)}$ for almost all $(x,t)\in\Omega\times(0,T)$, and hence $Du_k\to Du$ in measure on $\Omega\times(0,T)$. The rest of the proof is identical to the proof for case (d).

Case (b)

We start by showing that, for almost all $(x,t)\in\Omega\times(0,T)$, the support of $\nu_{(x,t)}$ is contained in the set where W agrees with the supporting hyperplane $L:=\{(\lambda, W(x,t,u,Du)+\sigma(x,t,u,Du)(\lambda-Du))\}$ in $Du(x,t)$; that is, we want to show that

$$\begin{aligned}\operatorname{spt}\nu_{(x,t)} &\subset K_{(x,t)}\\ &=\Big\{\lambda\in\mathbb{M}^{m\times n}: W\big(x,t.u,\lambda\big)=W\big(x,t,u,Du\big)+\sigma\big(x,t,u,Du\big):\big(\lambda-Du\big)\Big\}.\end{aligned}$$

If $\lambda\in\operatorname{spt}\nu_{(x,t)}$, then, by (23),

$$(1-\tau)\big(\sigma\big(x,t,u,Du\big)-\sigma\big(x,t,u,\lambda\big)\big):\big(Du-\lambda\big)=0\quad\text{for all }\tau\in[0,1]. \tag{24}$$

On the other hand, by monotonicity, we have for $\tau\in[0,1]$ that

$$0\le(1-\tau)\Big(\sigma\big(x,t,u,Du+\tau\big(\lambda-Du\big)\big)-\sigma\big(x,t,u,\lambda\big)\Big):\big(Du-\lambda\big). \tag{25}$$

Subtracting (24) from (25), we get

$$0\le(1-\tau)\Big(\sigma\big(x,t,u,Du+\tau\big(\lambda-Du\big)\big)-\sigma\big(x,t,u,Du\big)\Big):\big(Du-\lambda\big) \tag{26}$$

for all $\tau\in[0,1]$. But, by monotonicity, in (26) also the reverse inequality holds, and we may conclude that

$$\Big(\sigma\big(x,t,u,Du+\tau\big(\lambda-Du\big)\big)-\sigma\big(x,t,u,Du\big)\Big):\big(\lambda-Du\big)=0 \tag{27}$$

for all $\tau\in[0,1]$, whenever $\lambda\in\operatorname{spt}\nu_{(x,t)}$. Now, it follows from (27) that

$$\begin{aligned}W\big(x,t,u,\lambda\big)&=W\big(x,t,u,Du\big)+\int_0^1\sigma\big(x,t,u,Du+\tau\big(\lambda-Du\big)\big):\big(\lambda-Du\big)\,d\tau\\ &=W\big(x,t,u,Du\big)+\sigma\big(x,t,u,Du\big):\big(\lambda-Du\big),\end{aligned}$$

as claimed.

By the convexity of W we have $W(x,t,u,\lambda) \geq W(x,t,u,Du)+\sigma(x,t,u,Du) : (\lambda - Du)$ for all $\lambda \in \mathbb{M}^{m\times n}$, and thus L is a supporting hyperplane for all $\lambda \in K_{(x,t)}$. Since the mapping $\lambda \mapsto W(x,t,u,\lambda)$ is by assumption continuously differentiable, we obtain

$$\sigma\big(x,t,u,\lambda\big) = \sigma\big(x,t,u,Du\big) \quad \text{for all } \lambda \in K_{(x,t)} \supset \operatorname{spt} \nu_{(x,t)}, \tag{28}$$

and thus

$$\bar{\sigma} := \int_{\mathbb{M}^{m\times n}} \sigma\big(x,t,u,\lambda\big)\, d\nu_{(x,t)}(\lambda) = \sigma\big(x,t,u,Du\big). \tag{29}$$

Now consider the Carathéodory function

$$g\big(x,t,u,p\big) = \big|\sigma\big(x,t,u,p\big) - \bar{\sigma}(x,t)\big|.$$

The sequence $g_k(x,t) = g(x,t,u_k(x,t), Du_k(x,t))$ is equiintegrable, and thus

$$g_k \rightharpoonup \bar{g} \quad \text{weakly in } L^1\big(\Omega \times (0,T)\big)$$

and the weak limit $\bar{g}$ is given by

$$\begin{aligned}\bar{g}(x,t) &= \int_{\mathbb{R}^m \times \mathbb{M}^{m\times n}} \big|\sigma\big(x,t,\eta,\lambda\big) - \bar{\sigma}(x,t)\big|\, d\delta_{u(x,t)}(\eta) \otimes d\nu_{(x,t)}(\lambda)\\ &= \int_{\operatorname{spt} \nu_{(x,t)}} \big|\sigma\big(x,t,u(x,t),\lambda\big) - \bar{\sigma}(x,t)\big|\, d\nu_{(x,t)}(\lambda) = 0\end{aligned}$$

by (28) and (29). Since $g_k \geq 0$, it follows that

$$g_k \longrightarrow 0 \quad \text{strongly in } L^1\big(\Omega \times (0,T)\big).$$

This again suffices to pass to the limit in the equation, and the proof of case (b) is finished.

Case (a)
We claim that in this case for almost all $(x,t) \in \Omega \times (0,T)$ the following identity holds for all $\mu \in \mathbb{M}^{m\times n}$ on the support of $\nu_{(x,t)}$:

$$\sigma\big(x,t,u,\lambda\big) : \mu = \sigma\big(x,t,u,Du\big) : \mu + \big(\nabla\sigma\big(x,t,u,Du\big)\mu\big) : \big(Du - \lambda\big), \tag{30}$$

where ∇ is the derivative with respect to the third variable of σ. Indeed, by the monotonicity of σ we have, for all $\tau \in \mathbb{R}$,

$$\big(\sigma\big(x,t,u,\lambda\big) - \sigma\big(x,t,u,Du + \tau\mu\big)\big) : \big(\lambda - Du - \tau\mu\big) \geq 0,$$

whence, by (22),

$$
\begin{aligned}
&-\sigma\big(x,t,u,\lambda\big):(\tau\mu)\\
&\qquad\geq -\sigma\big(x,t,u,Du\big):\big(\lambda-Du\big)+\sigma\big(x,t,u,Du+\tau\mu\big):\big(\lambda-Du-\tau\mu\big)\\
&\qquad=\tau\Big(\big(\nabla\sigma\big(x,t,u,Du\big)\mu\big)\big(\lambda-Du\big)-\sigma\big(x,t,u,Du\big):\mu\Big)+o(\tau).
\end{aligned}
$$

The claim follows from this inequality since the sign of τ is arbitrary. Since the sequence $\sigma(x,t,u_k,Du_k)$ is equiintegrable, its weak L^1-limit $\bar{\sigma}$ is given by

$$
\begin{aligned}
\bar{\sigma}&=\int_{\operatorname{spt}\nu_{(x,t)}}\sigma\big(x,t,u,\lambda\big)\,d\nu_{(x,t)}(\lambda)\\
&=\int_{\operatorname{spt}\nu_{(x,t)}}\sigma\big(x,t,u,Du\big)\,d\nu_{(x,t)}(\lambda)\\
&\qquad+\big(\nabla\sigma\big(x,t,u,Du\big)\big)^t\int_{\operatorname{spt}\nu_{(x,t)}}\big(Du-\lambda\big)\,d\nu_{(x,t)}(\lambda)\\
&=\sigma\big(x,t,u,Du\big),
\end{aligned}
$$

where we used (30) in this calculation. This finishes the proof of case (a) and hence of Theorem 2.

Remark. Notice that in case (a) we have $\sigma(x,t,u_k,Du_k)\rightharpoonup\sigma(x,t,u,Du)$, in case (b) we have $\sigma(x,t,u_k,Du_k)\to\sigma(x,t,u,Du)$ in $L^1(\Omega\times(0,T))$, and in case (c) we even have $Du_k\to Du$ in measure on $\Omega\times(0,T)$ as $k\to\infty$.

Appendices

Appendix A

Here we give the proof of the modified lemma of Aubin, that is, Lemma 3.

Proof of Lemma 3

Let $\tilde{B}_0:=j(i(B_0))\subset B_1$ be the Banach space equipped with the norm

$$
\|\tilde{x}\|_{\tilde{B}_0}:=\inf_{\substack{x\in B_0\\ j\circ i(x)=\tilde{x}}}\|x\|_{B_0},
$$

and let $\tilde{B}:=j(B)\subset B_1$ be the Banach space equipped with the norm

$$
\|x\|_{\tilde{B}}:=\big\|j^{-1}(x)\big\|_B.
$$

(We recall that j is supposed to be injective.) Now, we consider a bounded sequence $\{v_n\}_n$ in W. Let $\tilde{v}_n:=j\circ i\circ v_n$. Then $\{\tilde{v}_n\}_n$ is bounded in

$$
\tilde{W}:=\left\{\tilde{v}\;\middle|\;\tilde{v}\in L^{p_0}\big(0,T;\tilde{B}_0\big),\,\frac{d\tilde{v}}{dt}\in L^{p_1}\big(0,T;\tilde{B}_1\big)\right\},
$$

and by the usual Aubin lemma (see [20, Chapter 1, Section 5.2]) it follows that there exists a subsequence $\tilde{v}_\nu$ that converges strongly in $L^{p_0}(0, T; \tilde{B})$. By isometry of B and $\tilde{B}$, the claim follows. □

Appendix B

Let u be an arbitrary function in $L^p(0, T; W_0^{1,p}(\Omega))$. We want to construct a sequence $v_k \in L^p(0, T; W_0^{1,p}(\Omega))$ which has the following properties:

(i) $v_k \to u$ in $L^p(0, T; W_0^{1,p}(\Omega))$;

(ii) $v_k(t) \in \operatorname{span}(w_1, w_2, \ldots, w_k)$ for $0 \le t \le T$.

To construct the sequence $\{v_k\}_k$, we take $\epsilon > 0$ (with the intention to let $\epsilon \to 0$) and a standard mollifier δ_η in space-time. The function u is extended by zero outside $\Omega \times [0, T] \subset \mathbb{R}^{n+1}$. Choosing $\eta > 0$ small enough, we may achieve that

$$\|u * \delta_\eta - u\|_{L^p(0,T;W^{1,p}(\Omega))} < \epsilon.$$

Now, for a smooth function $\phi \in C^\infty(\bar{\Omega} \times [0, T])$ and $j \in \mathbb{N}$, let

$$Q_j(\phi)(x, t) := \phi\left(x, i\frac{T}{j}\right) \quad \text{if } t \in \left[i\frac{T}{j}, (i+1)\frac{T}{j}\right)$$

denote the step function approximation of ϕ in time. We fix $j \in \mathbb{N}$ large enough such that we have

$$\left\|u * \delta_\eta - Q_j(u * \delta_\eta)\right\|_{L^p(0,T;W^{1,p}(\Omega))} < \epsilon.$$

Finally, we choose k large enough such that

$$\left\|Q_j(u * \delta_\eta) - P_k \circ Q_j(u * \delta_\eta)\right\|_{L^p(0,T;W^{1,p}(\Omega))} < \epsilon,$$

where (as before) P_k denotes the $W^{s,2}(\Omega)$-projection onto $\operatorname{span}(w_1, w_2, \ldots, w_k)$. (Notice that this is possible since $t \mapsto Q_j(u * \delta_\eta)$ takes only finitely many values on $[0, T]$.)

Combination yields

$$\left\|u - P_k \circ Q_j(u * \delta_\eta)\right\|_{L^p(0,T;W^{1,p}(\Omega))} < 3\epsilon,$$

and hence the sequence $v_k = P_{k(\epsilon)} \circ Q_{j(\epsilon)}(u * \delta_{\eta(\epsilon)})$ for $\epsilon \to 0$ is a sequence with the properties (i)–(ii).

Appendix C

Here, we want to prove that

$$u_k(\cdot, T) \rightharpoonup u(\cdot, T) \quad \text{weakly in } L^2(\Omega)$$

and that

$$u(\cdot, 0) = u_0.$$

Since $\{u_k\}_k$ is bounded in $L^\infty(0, T; L^2(\Omega))$, it is clear that, for a (not relabeled) subsequence,

$$u_k(\cdot, T) \rightharpoonup z \quad \text{weakly in } L^2(\Omega),$$

and we have to show $z = u(\cdot, T)$. To shorten the notation, we write from now on $u(T)$ instead of $u(\cdot, T)$, and so on.

In order to prove the claim, note that (again, after a possible choice of a further subsequence)

$$-\operatorname{div} \sigma\big(x, t, u_k, Du_k\big) \rightharpoonup \chi \quad \text{weakly in } L^{p'}\big(0, T; W^{-1,p'}(\Omega)\big).$$

Now, we claim that, for arbitrary $\psi \in C^\infty([0, T])$ and $v \in W_0^{1,p}(\Omega)$,

$$\int_\Omega z\psi(T)v\,dx - \int_\Omega u_0\psi(0)v\,dx = \big\langle f - \chi, \psi v\big\rangle + \int_0^T \int_\Omega \psi' v u\,dx\,dt. \tag{31}$$

Since $\cup_{n\in\mathbb{N}} \operatorname{span}(w_1, \dots, w_n)$ is dense in $W_0^{1,p}(\Omega)$, it suffices to verify (31) for $v \in \operatorname{span}(w_1, \dots, w_n)$. Then, by testing (6) by $v\psi$, we have, for $m \geq n$,

$$\underbrace{\int_0^T\int_\Omega \partial_t u_m v\psi\,dx\,dt}_{= \int_\Omega u_m(T)\psi(T)v\,dx - \int_\Omega u_m(0)\psi(0)v\,dx - \int_0^T\int_\Omega u_m v\psi'\,dx\,dt} + \int_0^T\int_\Omega \sigma\big(x, t, u_m, Du_m\big) : Dv\psi\,dx\,dt = \langle f, v\psi\rangle.$$

Then (31) follows by letting m tend to infinity. By choosing $\psi(0) = \psi(T) = 0$ in (31), we have, in particular,

$$\big\langle f - \chi, \psi v\big\rangle = -\int_0^T \int_\Omega \psi' v u\,dx\,dt = \int_0^T \int_\Omega \psi v u'\,dx\,dt,$$

and hence

$$u' + \chi = f.$$

Using this and (31) we have, on the other hand,

$$\begin{aligned}
&\int_\Omega z\psi(T)v\,dx - \int_\Omega u_0\psi(0)v\,dx \\
&\qquad = \big\langle u', \psi v\big\rangle + \int_0^T \int_\Omega \psi' v u\,dx\,dt \\
&\qquad = \int_\Omega u\psi v\,dx\Big|_0^T \\
&\qquad = \int_\Omega u(T)\psi(T)v\,dx - \int_\Omega u(0)\psi(0)v\,dx.
\end{aligned} \tag{32}$$

Choosing $\psi(T) = 1$, $\psi(0) = 0$ in (32), we obtain that $u(0) = u_0$, and choosing $\psi(T) = 0$, $\psi(0) = 1$, we get $u(T) = z$, as claimed.

Appendix D

In this section, we want to assume that σ does not depend on x and u, and we want to replace condition (P2) by the following more classical quasi-monotonicity condition:

($P2'$) For all fixed $t \in [0, T)$, the map $\sigma(t, F)$ is strictly quasi-monotone in the variable F.

Here, by strictly *quasi-monotone*, we mean the following.

Definition 6

A function $\eta : \mathbb{M}^{m\times n} \to \mathbb{M}^{m\times n}$ is said to be strictly quasi-monotone if there exist constants $c > 0$ and $r > 0$ such that

$$\int_\Omega \big(\eta(Du) - \eta(Dv)\big) : (Du - Dv)\,dx \geq c \int_\Omega \big|Du - Dv\big|^r\,dx$$

for all $u, v \in W_0^{1,p}(\Omega)$.

We want to prove the following theorem.

THEOREM 7

If $\sigma(t, Du)$ satisfies conditions (P0), (P1), and ($P2'$) for some $p \in (2n/(n+2), \infty)$, then the parabolic system (1)–(3) has a weak solution $u \in L^p(0, T; W_0^{1,p}(\Omega))$ for every $f \in L^{p'}(0, T; W^{-1,p}(\Omega))$ and every $u_0 \in L^2(\Omega)$.

Since in this case we do not have to deal with x- and u-dependence of σ, the following simple proof is possible.

Proof

Let u_k and v_k be constructed as in the proof of Theorem 2. Then, by using $u_k - v_k$ as a test function in (6), we obtain

$$\begin{aligned}
\big\langle f, u_k - v_k\big\rangle - \int_0^T \int_\Omega \big(u_k - v_k\big)\partial_t u_k\,dx\,dt \\
= \int_0^T \int_\Omega \sigma\big(t, Du_k\big) : \big(Du_k - Dv_k\big)\,dx\,dt \\
= \int_0^T \int_\Omega \Big(\sigma\big(t, Du_k\big) - \sigma\big(t, Dv_k\big)\Big) : \big(Du_k - Dv_k\big)\,dx\,dt \\
+ \int_0^T \int_\Omega \sigma\big(t, Dv_k\big) : \big(Du_k - Dv_k\big)\,dx\,dt.
\end{aligned} \tag{33}$$

The first term on the left of (33), $\langle f, u_k - v_k\rangle$, converges to zero as $k \to \infty$ since $u_k - v_k \rightharpoonup 0$ in $L^p(0, T; W_0^{1,p}(\Omega))$. For the second term on the left of (33), we have seen in Section 6 that

$$\liminf_{k\to\infty} -\int_0^T \int_\Omega \big(u_k - v_k\big)\partial_t u_k \, dx \, dt \le 0$$

for $k \to \infty$. The last term on the right of (33) converges to zero for $k \to \infty$ since $\sigma(t, Dv_k) \to \sigma(t, Du)$ in $L^{p'}(0, T; L^{p'}(\Omega))$ (at least for a subsequence) and $Du_k - Dv_k \rightharpoonup 0$ in $L^p(0, T; L^p(\Omega))$. We conclude that

$$\begin{aligned} o(1) &= \int_0^T \int_\Omega \big(\sigma\big(t, Du_k\big) - \sigma\big(t, Dv_k\big)\big) : \big(Du_k - Dv_k\big) \, dx \, dt \\ &\ge c \int_0^T \int_\Omega \big|Du_k - Dv_k\big|^r \, dx \, dt. \end{aligned}$$

This implies $Du_k \to Du$ in measure for a suitable subsequence. The rest of the proof is as in case (d) in Section 7. □

Acknowledgments. This work was carried out while the author was at the Max Planck Institute for Mathematics in the Sciences in Leipzig. I am most grateful to Professors Stefan Müller and Eberhard Zeidler for their support and for providing and maintaining such a stimulating place.

I would like to thank Professors Müller and Michael Struwe for their continuing interest in this work. I am especially grateful to Professor Jens Frehse for initiating my interest in parabolic systems with weak monotonicity assumptions. Finally, I would like to thank Jan Kristensen and Mikhail Sytchev for inspiring discussions and many useful hints.

References

[1] J. M. BALL, "A version of the fundamental theorem for Young measures" in *PDEs and Continuum Models of Phase Transitions (Nice, 1988)*, Lecture Notes in Phys. **344**, Springer, Berlin, 1989, 207–215.

[2] L. BOCCARDO and F. MURAT, *Almost everywhere convergence of the gradients of solutions to elliptic and parabolic equations*, Nonlinear Anal. **19** (1992), 581–597.

[3] H. BRÉZIS, O*pérateurs maximaux monotones et semi-groupes de contractions dans les espaces de Hilbert*, North-Holland Math. Stud. **5**, North-Holland, Amsterdam, Notas Mat. **50**, American Elsevier, New York, 1973.

[4] H. BRÉZIS and F. E. BROWDER, *Strongly nonlinear parabolic initial-boundary value problems*, Proc. Nat. Acad. Sci. U.S.A. **76** (1979), 38–40.

[5] F. E. BROWDER, "Existence theorems for nonlinear partial differential equations" in *Global Analysis (Berkeley, Calif., 1968)*, Proc. Sympos. Pure Math. **16**, Amer. Math. Soc., Providence, 1970, 1–60.

[6] G. DOLZMANN, N. HUNGERBÜHLER, and S. MÜLLER, *Non-linear elliptic systems with measure-valued right hand side*, Math. Z. **226** (1997), 545–574.

[7] N. DUNFORD and J. T. SCHWARTZ, *Linear Operators, Parts I–III*, Wiley Classics Lib., Wiley, New York, 1988.

[8] L. C. EVANS and R. F. GARIEPY, *Measure theory and fine properties of functions*, Stud. Adv. Math., CRC Press, Boca Raton, Fla., 1992.

[9] I. FONSECA, S. MÜLLER, and P. PEDREGAL, *Analysis of concentration and oscillation effects generated by gradients*, SIAM J. Math. Anal. **29** (1998), 736–756, http://epubs.siam.org/sam-bin/dbq/toclist/SIMA

[10] J. FREHSE, private communication, 1999.

[11] N. HUNGERBÜHLER, *A refinement of Ball's theorem on Young measures*, New York J. Math. **3** (1997), 48–53, http://nyjm.albany.edu:8000/nyjm.html

[12] ———, *Quasilinear elliptic systems in divergence form with weak monotonicity*, New York J. Math. **5** (1999), 83–90, http://nyjm.albany.edu:8000/nyjm.html

[13] E. KAMKE, *Das Lebesgue-Stieltjes-Integral*, Teubner, Leipzig, 1960.

[14] D. KINDERLEHRER and P. PEDREGAL, *Gradient Young measures generated by sequences in Sobolev spaces*, J. Geom. Anal. **4** (1994), 59–90.

[15] J. KRISTENSEN, *Lower semicontinuity in spaces of weakly differentiable functions*, Math. Ann. **313** (1999), 653–710.

[16] R. LANDES, *On weak solutions of quasilinear parabolic equations*, Nonlinear Anal. **9** (1985), 887–904.

[17] ———, *A note on strongly nonlinear parabolic equations of higher order*, Differential Integral Equations **3** (1990), 851–862.

[18] R. LANDES and V. MUSTONEN, *A strongly nonlinear parabolic initial-boundary value problem*, Ark. Mat. **25** (1987), 29–40.

[19] ———, *On parabolic initial-boundary value problems with critical growth for the gradient*, Ann. Inst. H. Poincaré Anal. Non Linéaire **11** (1994), 135–158.

[20] J.-L. LIONS, *Quelques méthodes de résolution des problèmes aux limites non linéaires*, Dunod, Gauthier-Villars, Paris, 1969.

[21] G. J. MINTY, *Monotone (nonlinear) operators in Hilbert space*, Duke Math. J. **29** (1962), 341–346.

[22] J. SIMON, *Compact sets in the space* $L^p(0, T; B)$, Ann. Mat. Pura Appl. (4) **146** (1987), 65–96.

[23] M. I. VIŠIK, *Solubility of boundary-value problems for quasi-linear parabolic equations of higher orders* (in Russian), Mat. Sb. (N.S.) **59 (101)** (1962) suppl., 289–325.

[24] K. YOSIDA, *Functional Analysis*, 6th ed., Grundlehren Math. Wiss. **123**, Springer, Berlin, 1980.

[25] E. ZEIDLER, *Nonlinear Functional Analysis and Its Applications, II/B: Nonlinear Monotone Operators*, Springer, New York, 1990.

Department of Mathematics, University of Alabama at Birmingham, 452 Campbell Hall, 1300 University Boulevard, Birmingham, Alabama 35294-1170, USA; buhler@math.uab.edu;
current: Institut de Mathématiques, Université de Fribourg, Pérolles, 1700 Fribourg, Switzerland