TRANSACTIONS OF THE
AMERICAN MATHEMATICAL SOCIETY
Volume 353, Number 12, Pages 5095–5120
S 0002-9947(01)02847-1
Article electronically published on June 21, 2001

ON THE WELLPOSEDNESS OF CONSTITUTIVE LAWS INVOLVING DISSIPATION POTENTIALS

WOLFGANG DESCH AND RONALD GRIMMER

Abstract. We consider a material with memory whose constitutive law is formulated in terms of internal state variables using convex potentials for the free energy and the dissipation. Given the stress at a material point depending on time, existence of a strain and a set of inner variables satisfying the constitutive law is proved. We require strong coercivity assumptions on the potentials, but none of the potentials need be quadratic.

As a technical tool we generalize the notion of an Orlicz space to a cone "normed" by a convex functional which is not necessarily balanced. Duality and reflexivity in such cones are investigated.

1. Introduction

1.1. **Survey.** The rheological behavior of a deformable material is described by the constitutive law. For a perfectly elastic material, the stress $\sigma(t)$ at time t is completely determined by the present strain $\epsilon(t)$. Energy dissipating phenomena like viscoelasticity, elastoplasticity and viscoplasticity, however, are characterized by the fact that $\sigma(t)$ depends on the whole history of the strain up to time t and vice versa:

$$\sigma(t) = \mathcal{F}(\epsilon(t-\cdot)) \quad \text{or} \quad \epsilon(t) = \mathcal{G}(\sigma(t-\cdot)). \tag{1.1}$$

The operators $\mathcal{F}$ and $\mathcal{G}$, mapping a tensor valued function into a tensor, may be, for instance, convolution operators in the case of viscoelasticity, or hysteresis operators in the case of plasticity.

Instead of stating the constitutive relation (1.1) explicitely, the following model assumes the existence of internal ("hidden") state variables V, whose change reflects the mechanism for energy dissipation.

$$\begin{aligned} &(\sigma(t), A(t)) \in \partial\psi(\epsilon(t), V(t)), \\ &\dot{V}(t) \in \partial\phi^*(-A(t)). \end{aligned} \tag{1.2}$$

Here $\epsilon(t) \in \mathbb{R}^M$ and $\sigma(t) \in \mathbb{R}^M$ describe the strain and stress at a material point at time t. Depending on the geometry of the problem, they may be scalars or symmetric 3×3-tensors, possibly with the restriction that the trace is 0 (in the

Received by the editors March 9, 2000 and, in revised form, February 16, 2001.

1991 *Mathematics Subject Classification.* Primary 73B05, Secondary 46E30.

Key words and phrases. Dissipation potential, viscoplastic material, constitutive equation, Orlicz space.

Supported by Spezialforschungsbereich F 003 "Optimierung und Kontrolle" at the Karl-Franzens-Universität Graz, grant GAUK 19/1997. W. D. acknowledges the kind hospitality of Southern Illinois University, Carbondale.

incompressible case), identified with vectors from $\mathbb{R}^6$ or $\mathbb{R}^5$. $V(t) \in \mathbb{R}^N$ is an internal state variable (some kind of hidden strain), while $A(t) \in \mathbb{R}^N$ is the corresponding thermodynamic force (some hidden stress). Notice that the internal state variables are usually purely hypothetical and not accessible to measurement. The potential ψ gives the density of free energy (per unit mass), while ϕ governs the law of energy dissipation by plasticity or viscoelastic damping. This model has been adapted from [8, Section 2.4], and many particular instances of this formal framework are found throughout the same monograph. We have restricted the consideration to the isothermic case and omitted the effects of temperature and heat flux.

To obtain the constitutive relationship in explicit form (1.1), one has to solve

Problem 1.1. Given initial states $\epsilon(0) = \epsilon_0 \in \mathbb{R}^M$, $V(0) = V_0 \in \mathbb{R}^N$, and the stress history $\sigma(\cdot) \in \mathbf{W}^{1,2}([0,T], \mathbb{R}^M)$, find $\epsilon(\cdot)$, $V(\cdot)$, and $A(\cdot)$ solving (1.2).

Our paper will give sufficient conditions such that this problem admits at least one solution.

In the case of linear elasticity and viscoelasticity this boils down to solving linear differential equations. However, as soon as nonlinear elements enter the scene, the explicit form of (1.1) is usually out of reach. Even worse, if ϕ and ψ are not smooth enough so that (1.2) reduces to an ordinary differential equation with Lipschitz continuous right hand side, it can be rather laborious to prove existence and uniqueness ("wellposedness") of solutions. This situation appears inevitably, if plasticity or viscoplasticity are considered. Examples of sophisticated ad hoc reasoning to prove the wellposedness of some seemingly simple elastoplastic constitutive models are found in [3]. On a more abstract level, this case has been investigated quite intensely in literature. We give a survey on existing results later. The key to these results is the convex structure of the model, suggesting that there should be some general theory of existence and uniqueness for (1.2) based on convexity rather than smoothness of the right hand side.

Our paper is meant to be a step on the way to a general theory of wellposedness for (1.2). We will give another set of sufficient conditions for the existence of solutions to Problem 1.1. We will not require smoothness beyond convexity and lower semi-continuity. Our conditions concern the growth and coercivity behavior of both potentials.

1.2. **Our conditions.** Our growth assumptions on ϕ are rather restrictive, ruling out rate independent phenomena of perfect plasticity and elastoplasticity, but well suited for viscoplasticity like e.g. Odqvist's law [8, p. 282]: (Here, σ_p is the inelastic stress, corresponding to the variable $-A$ in the abstract model. Its deviator is denoted by σ_p'.)

$$\phi^*(\sigma_p) = \frac{\lambda^*}{N^*+1}\left(\frac{\sigma_{eq}}{\lambda^*}\right)^{N^*+1},$$
$$\sigma_{eq}^2 = \frac{3}{2}\sigma_p' : \sigma_p'.$$

On the other hand, we allow for very general free energy potentials ψ. Nonlinear elastic potentials are well established in literature, take for instance the Neohookean potential and its generalizations for rubber-like materials [14, p. 242]. Global Lipschitz continuity of the derivative may fail for instance, if the elastic potential is infinite for some strains, like the following potential proposed to describe the elastic behavior of biological tissues in orthotropic experimental settings [15]:

$$\begin{aligned}\psi(\epsilon) &= -C\ln(1-Q),\\ Q &= \frac{1}{2}(a_{11}E_{11}^2 + a_{22}E_{22}^2 + 2a_{12}E_{11}E_{22}),\\ E_{ii} &= \frac{1}{2}(\epsilon_{ii}^2 + 2\epsilon_{ii}) \text{ Green-St. Venant strain in direction } i.\end{aligned}$$

Such potentials describe materials that cannot be strained beyond a certain margin. Our setting includes constitutive laws built from very general and possibly pathologic elastic elements combined with viscoplastic and viscoelastic elements.

From the viewpoint of particular constitutive laws, the combination of a nonsmooth elastic potential with viscoplasticity is certainly a special case of marginal interest only. The main message of our work is the comparison to existing results on the wellposedness of Problem 1.1: There seems to be a tradeoff between the coercivity properties of ϕ and ψ. Whenever coercivity assumptions on ϕ are imposed, assumptions on ψ can be relaxed. We expect, that a general theory of wellposedness will require hypotheses on the coercivity of $\phi + \psi$ and $\phi^* + \psi^*$.

1.3. **Literature.** To our knowledge, wellposedness of (1.2) has not been treated in the literature in the full generality of the framework outlined above. It should be anticipated that in many cases both ϕ and ψ are sufficiently smooth such that the constitutive equations reduce to ordinary differential equations with Lipschitz continuous right hand side. However, the references cited below, mostly motivated by problems of elastoplasticity, show that frequently the assumption of smooth potentials can be significantly relaxed if the convex structure of the equation is exploited.

The assumption that $\psi(\epsilon, V)$ is quadratic means that σ and A depend linearly on the state, i.e., all elastic components of the model are Hookean. In [1] this is called a constitutive equation of gradient type. There, existence and uniqueness of the solution to the dynamic problem (i.e., the partial differential equation governing motion of an elastoplastic body) is proved. The coercivity assumptions on ϕ are no longer needed, in fact, the relation between A and $\dot{V}$ need not even be given by a subdifferential, but by any maximal monotone operator. The approach in [1] consists in setting up the dynamic problem as a semigroup and choosing a suitable norm which makes the generator an m-dissipative operator. Reference [1] does not deal explicitely with (1.2), but the ideas in [1] can be easily modified to set up an inhomogeneous evolution equation for the state $(\sigma(t), \epsilon(t), V(t))$.

Reference [9] contains the discussion of a one-dimensional quasi-static problem with the Prandtl-Reuss constitutive law. Here the free energy is again quadratic, and ϕ^* is the indicator function of the yield characteristic. This means rate-independent dissipation, i.e., plasticity. Therefore, this reference provides also an instructive example for the difficulties implied by the non-coercivity of ϕ.

Quasi-static problems, in a somewhat different setting with a quadratic free energy potential, are also treated in [4, Chapter 3]. Here the irreversible phenomena are modelled by a Lipschitz continuous perturbation of the linear elastic equation. Existence of solutions for perfect plasticity is treated as a limiting case of Lipschitz continuous problems. A partial uniqueness result, namely uniqueness of the stress field, is inferred from the quadratic structure of the free energy.

More general free energy potentials are treated in [7]. Again, ϕ^* is the indicator function of a convex set. Some coercivity and smoothness assumptions on ψ are imposed. In particular, ψ is twice continuously differentiable with respect to V, and the Hessian is Lipschitz continuous. This, in turn, implies not only existence of solutions but also uniqueness and Lipschitz continuous dependence on initial data. In [7], the state space may be an infinite dimensional Hilbert space.

Our paper is almost complementary to the work cited above. The conditions on the free energy potential ψ are relaxed to the level that almost no conditions on ψ with respect to V are imposed. On the contrary, we require very strong coercivity conditions on ϕ and its Fenchel conjugate ϕ^*, so that rate-independent phenomena are ruled out. We hope to overcome this drawback by creating hybrids between the conditions stated in the literature and our work in forthcoming research.

1.4. **Orlicz spaces.** A significant part of this paper is devoted to adapt technical tools from the concept of Orlicz spaces. Nevertheless, the idea of the wellposedness proof can be understood without these tools if the Orlicz spaces $\mathbf{L}^\phi$ and $\mathbf{L}^{\phi^*}$ in this paper are replaced by $\mathbf{L}^p([0,T],\mathbb{R}^N)$ and $\mathbf{L}^q([0,T],\mathbb{R}^N)$ with $1 < p < \infty$, $p^{-1} + q^{-1} = 1$. This pertains exactly to the case of Odqvist's law of viscoplasticity, possibly complemented by a nonlinear elastic component.

For non-quadratic potentials, Orlicz spaces are the natural "energy spaces" similar to $\mathbf{L}^2$ in the case of linear elasticity and viscoelasticity. The technical key in our approach is an a priori estimate on the dissipated energy

$$\int_0^T [\phi(\dot V(t)) + \phi^*(-A(t))]\, dt,$$

which implies that $\dot V(\cdot)$ and $-A(\cdot)$ live in spaces constructed by the convex functions ϕ and ϕ^* very similarly as Orlicz spaces. It is some kind of reflexivity of these spaces that enables us to set up weak compactness and thus convergence of some approximating solutions. Therefore we need some tools concerning duality.

On this way we face a technical difficulty. The definition of an Orlicz space is based on the integral

$$\int_0^T \rho(|u(t)|)\, dt$$

where $\rho : [0,\infty) \to [0,\infty]$ is a convex, lower semicontinuous function with $\rho(0) = 0$. The only difficulty to generalize from $\rho(|u|)$ to a general convex function $\phi(u)$ lies in the possibility that $\phi(-u) \neq \phi(u)$. Physically speaking, this means that the material may behave quite differently under compression and extension, respectively. As a mathematical consequence, the generalized "Orlicz space" is rather a cone than a vector space. Since estimates holding for u are no longer automatically valid for $-u$ as well, all estimates are one-sided. This requires some changes in the notion of the dual space, and some technical adaptations. Due to the strong growth conditions on ϕ and ϕ^*, some properties of the spaces $\mathbf{L}^p$ with $1 < p < \infty$ can be carried over. In particular, we obtain some properties close to reflexivity.

The tools of convex analysis needed in this paper can be found, for instance, in the monographs [2, 13]. Moreover, [9] (mentioned already above) starts with a compact, but very informative survey of convex analysis applied to elasticity. The theory of Orlicz spaces is described in [5] and [10].

1.5. **Structure of the paper.** In the next section we adapt as much as we need from the theory of Orlicz spaces. The core of our paper is Section 3 containing the proof of existence for solutions of (1.2).

2. Orlicz space considerations

In this section we generalize the notion of an Orlicz space on an interval $[0, T]$ (see [5, 10]). The purpose of this section is mainly to provide a tool of weak compactness by an analog to the reflexivity of certain Orlicz spaces.

This section contains information about

2.1 Basic definitions,
2.2 Growth conditions,
2.3 Duality,
2.4 Notions of convergence,
2.5 Some technical lemmas.

2.1. **Basic definitions.** Throughout this section we assume

Hypothesis 2.1. Let $\phi : \mathbb{R}^N \to [0, \infty)$ be everywhere defined, convex, and continuous. Moreover, we assume that ϕ is coercive, i.e.

$$\lim_{|x|\to\infty} \frac{1}{|x|}\phi(x) = \infty.$$

By ϕ^* we denote, as usual, its Fenchel-Legendre transform

$$\phi^*(x) = \sup_{y\in\mathbb{R}^N} [\langle x, y\rangle - \phi(y)].$$

Remark 2.1. Hypothesis 2.1 implies that ϕ^* is also everywhere defined, continuous and coercive.

Proof. By [13, Theorems 2.35 and 11.8], for a convex, lower semicontinuous function $\phi : \mathbb{R}^N \to [0, \infty]$, the following are equivalent:

(1) ϕ is bounded on bounded subsets of $\mathbb{R}^N$.
(2) ϕ is continuous everywhere in $\mathbb{R}^N$.
(3) ϕ^* is coercive. □

Definition 2.1. Let ϕ satisfy Hypothesis 2.1. We define

$$\mathcal{L}^\phi([0, T], \mathbb{R}^N) = \big\{u \in \mathbf{L}^1([0, T], \mathbb{R}^N) \mid \exists\alpha > 0 : \int_0^T \phi(\alpha^{-1}u(t))\, dt \le 1\big\}.$$

By $\mathbf{L}^\phi([0, T], \mathbb{R}^N)$ we denote, as usual, the set of equivalence classes in $\mathcal{L}^\phi([0, T], \mathbb{R}^N)$ with respect to equality almost everywhere. For $u \in \mathbf{L}^\phi([0, T], \mathbb{R}^N)$, we define

$$N_\phi(u) = \inf\big\{\alpha > 0 : \int_0^T \phi(\alpha^{-1}u(t))\, dt \le 1\big\}.$$

In the classical theory, $N_\phi(u)$ is known as the Luxemburg norm of the function u.

Remark 2.2. $\mathbf{L}^\phi([0,T],\mathbb{R}^N)$ is a convex cone. Moreover, for $u,v \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$, $\lambda > 0$, the following properties hold:

$$\begin{aligned}
&N_\phi(u) \geq 0, \text{ and } N_\phi(u) = 0 \text{ implies } u = 0,\\
&N_\phi(u+v) \leq N_\phi(u) + N_\phi(v),\\
&N_\phi(\lambda u) = \lambda N_\phi(u),\\
&\int_0^T |u(t)|\,dt \leq M N_\phi(u) \text{ with some } M \text{ independent of } u.
\end{aligned}$$

2.2. **Growth conditions.** The following growth conditions and their relation to reflexivity of Orlicz spaces have already been investigated in [5, 10]:

Hypothesis 2.2.

(1) ϕ satisfies the Δ_2-growth condition: There exist some $K > 1$, $M > 0$, such that $\phi(2x) \leq K\phi(x)$ for all $x \in \mathbb{R}^N$ with $|x| \geq M$.
(2) ϕ satisfies the ∇_2-growth condition: There exists some $\ell > 1$, $M > 0$, such that $\phi(x) \leq \frac{1}{2\ell}\phi(\ell x)$ for all $x \in \mathbb{R}^N$ with $|x| \geq M$.

Remark 2.3. Hypothesis 2.2 implies that ϕ^* also satisfies Δ_2 and ∇_2.

Proof. For the case of even $\phi : \mathbb{R} \to [0,\infty)$, the equivalence between the Δ_2-condition for ϕ and the ∇_2-condition for ϕ^* can be found in [5, Theorem 4.2], see also [10, Section 2.3]. We give a direct proof not involving derivatives:

First assume that ϕ satisfies Δ_2 with a constant K, and let $\ell = K/2$. Given y, choose x such that

$$\phi(x) + \phi^*(y) = \langle x, y\rangle. \tag{2.1}$$

Since ϕ^* is coercive, x is large when y is also. Therefore we get the estimate

$$\langle x, y\rangle = \frac{1}{2\ell}\langle 2x, \ell y\rangle \leq \frac{1}{2\ell}(\phi(2x) + \phi^*(\ell y)) \leq \frac{K}{2\ell}\phi(x) + \frac{1}{2\ell}\phi^*(\ell y).$$

Subtracting (2.1), we obtain

$$\frac{1}{2\ell}\phi^*(\ell y) - \phi^*(y) \geq 0.$$

On the other hand, assume that ϕ^* satisfies ∇_2 with a constant ℓ and put $K = 2\ell$. Given x, choose y such that (2.1) holds. We have

$$\begin{aligned}
\phi^*(y) + \phi(x) = \langle y, x\rangle = 2\ell\langle \frac{1}{\ell}y, \frac{1}{2}x\rangle &\leq 2\ell\phi^*(\frac{1}{\ell}y) + 2\ell\phi(\frac{1}{2}x)\\
&\leq \phi^*(y) + 2\ell\phi(\frac{1}{2}x).
\end{aligned}$$

□

If Hypothesis 2.2 holds, then $\mathbf{L}^\phi([0,T],\mathbb{R}^N)$ is contained in some reflexive $\mathbf{L}^p([0,T],\mathbb{R}^N)$ and vice versa, as the following lemma shows.

Lemma 2.4.

(1) *If ϕ is bounded on bounded sets and satisfies the Δ_2-growth condition, then there exists some $p \in (1,\infty)$ such that $\mathbf{L}^p([0,T],\mathbb{R}^N) \subset \mathbf{L}^\phi([0,T],\mathbb{R}^N)$.*
(2) *If there exist constants $C_1, C_2 > 0$ such that $\phi(x) \geq C_2$ if $x \geq C_1$, and if ϕ satisfies the ∇_2-growth condition, then there exists some $p \in (1,\infty)$ such that $\mathbf{L}^\phi([0,T],\mathbb{R}^N) \subset \mathbf{L}^p([0,T],\mathbb{R}^N)$.*

Proof. 1) Let $C_1 \geq 1$ be sufficiently large such that $\phi(2x) \leq K\phi(x)$ for $|x| > C_1$, and $\phi(x) \leq C_2$ for $|x| \leq K$. Then, by induction, if $|y| \leq 2^n C_1$, we have $\phi(x) \leq K^n C_2$. From this we infer that

$$\phi(x) \leq \begin{cases} C_2 & \text{if } |x| \leq C_1, \\ C_2 K^{1+\log_2(x/C_1)} & \text{else.} \end{cases}$$

Therefore, in any case $\phi(x) \leq C_2(1 + KC_1^{-p}x^p)$ with $p = \log_2(K)$.

2) Take ℓ according to the ∇_2-growth condition. Immediately we conclude that $\phi(x) \geq (2\ell)^k C_2$ whenever $|x| \geq \ell^k C_1$ and k is a positive integer. This implies

$$\phi(x) \geq C_2(2\ell)^{\frac{\log(|x|/C_1\ell)}{\log(\ell)}} = C_2(2\ell)^{-\frac{\log(C_1\ell)}{\log(\ell)}}|x|^{\frac{\log(2\ell)}{\log(\ell)}}$$

whenever $|x| \geq C_1\ell$. This implies that $\mathbf{L}^\phi([0,T],\mathbb{R}^N) \subset \mathbf{L}^p([0,T],\mathbb{R}^N)$ with $p = 1 + \frac{\log 2}{\log(\ell)}$. □

Monotone convergence plays a major role as a technical tool. For shorthand we need the following definition:

Definition 2.2. Let $u_n, u \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$. We say that u_n converges monotonically to u if there exists a sequence $\alpha_n \in \mathbf{L}^\infty([0,T],\mathbb{R})$ such that $0 \leq \alpha_n(t) \leq \alpha_{n+1}(t) \leq 1$, $\alpha_n(t) \to 1$ almost everywhere, and $u_n(t) = \alpha_n(t)u(t)$.

We emphasize the fact that monotonicity here is not understood in the sense of an ordered vector space: The underlying ordering

$$u \preceq v \iff (\exists \alpha \in \mathbf{L}^\infty([0,T],[0,1])) : u = \alpha v$$

is not compatible with addition.

Lemma 2.5. *Let ϕ satisfy Hypotheses 2.1 and 2.2. Let u_n and u be in $\mathbf{L}^\phi([0,T],\mathbb{R}^N)$ such that u_n converges monotonically to u. Then $N_\phi(u - u_n) \to 0$.*

Proof. Given some $\epsilon = 2^{-m} > 0$, we want to show that for sufficiently large n we have

$$\int_0^T \phi(\frac{1}{\epsilon}(u(t) - u_n(t)))\,dt \leq 1.$$

We define the set

$$Q_n = \{t \in [0,T] : \phi(2^m(u(t) - u_n(t))) \geq \frac{1}{2T}\}.$$

Notice that the Lebesgue measure of Q_n converges to 0. Using the constant K from the Δ_2-condition Hypothesis 2.2, we obtain the estimate

$$\begin{aligned} &\int_0^T \phi(2^m(u(t) - u_n(t)))\,dt \\ &\quad \leq \int_{Q_n} \phi(2^m(u(t) - u_n(t)))\,dt + \int_{[0,T]\setminus Q_n} \phi(2^m(u(t) - u_n(t)))\,dt \\ &\quad \leq \int_{Q_n} K^m \phi(u(t))\,dt + \frac{1}{2}. \end{aligned}$$

Choosing n sufficiently large such that

$$K^m \int_{Q_n} \phi(u(t))\,dt < \frac{1}{2},$$

we obtain the desired estimate. □

We provide also a weaker version of this lemma which does not require Hypothesis 2.2:

Lemma 2.6. *Let ϕ be bounded on bounded sets. Let w_n be a bounded sequence in $\mathbf{L}^\infty([0,T],\mathbb{R}^N)$ such that $w_n(t) \to 0$ almost everywhere as $t \to \infty$. Then $N_\phi(w_n) \to 0$.*

Proof. Let $\|w_n\|_{\mathbf{L}^\infty} \le C_1$. Given $\epsilon > 0$, we put

$$C_2 = \sup_{|x|\le C_1} \phi(\frac{1}{\epsilon}x).$$

We define

$$Q_n = \{t \in [0,T] \mid \phi(\frac{1}{\epsilon}w_n(t)) > \frac{1}{2T}\}.$$

Since $w_n \to 0$ a.e., the measure of Q_n tends to 0 as $n \to \infty$. We choose n sufficiently large such that the measure of Q_n is bounded by $1/(2C_2)$, and estimate

$$\begin{aligned}\int_0^T \phi(\frac{1}{\epsilon}w_n(t))\,dt &= \int_{Q_n} \phi(\frac{1}{\epsilon}w_n(t))\,dt + \int_{[0,T]\setminus Q_n} \phi(\frac{1}{\epsilon}w_n(t))\,dt \\ &\le \frac{1}{2C_2}C_2 + T\frac{1}{2T} \le 1,\end{aligned}$$

thus $N_\phi(w_n) \le \epsilon$. □

2.3. **Duality.** We investigate the duality of the spaces $\mathbf{L}^\phi([0,T],\mathbb{R}^N)$ and $\mathbf{L}^{\phi^*}([0,T],\mathbb{R}^N)$, which is well established for Orlicz spaces [10, Chapters 4.1, 4.2]. The following definition mimics the definition of the dual space of an Orlicz space:

Definition 2.3. By $(\mathbf{L}^\phi)^*([0,T],\mathbb{R}^N)$ we denote the set of all functions

$$F : \mathbf{L}^\phi([0,T],\mathbb{R}^N) \to [-\infty,\infty)$$

with the following properties:

(1) $F(u+v) = F(u) + F(v)$ for all $u, v \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$.
(2) $F(\lambda u) = \lambda F(u)$ for all $u \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$ and all $\lambda \ge 0$.
(3) $F(0) = 0$.
(4) There exists a constant $M > 0$ such that $F(u) \le MN_\phi(u)$ for all $u \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$.

$\mathbf{P}^{\phi^*}([0,T],\mathbb{R}^N)$ denotes the set of all $F \in (\mathbf{L}^\phi)^*([0,T],\mathbb{R}^N)$ satisfying the following "monotone convergence property"

> If $u_n \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$ converge monotonically to u (see Definition 2.2), then $F(u) = \lim_{n\to\infty} F(u_n)$ (the limit being finite or negative infinite).

For $F \in (\mathbf{L}^\phi)^*([0,T],\mathbb{R}^N)$ we define

$$O_{\phi^*}(F) = \inf\{M > 0 \mid \forall u \in \mathbf{L}^\phi([0,T],\mathbb{R}^N) : F(u) \le MN_\phi(u)\}.$$

If ϕ is even, then O_{ϕ^*} is the Orlicz norm $\|\cdot\|_{\phi^*}$. However, we avoid the notation by a norm symbol since in the general case O_{ϕ^*} is no longer a norm. We also emphasize that, according to the definition above, $-\infty$ can be taken as a value of $F(u)$.

Proposition 2.7. *Let $v \in \mathbf{L}^{\phi^*}([0,T],\mathbb{R}^N)$. Then v can be identified with the following functional $F_v \in \mathbf{P}^{\phi^*}([0,T],\mathbb{R}^N)$:*

$$F_v(u) = \int_0^T \langle v(t), u(t)\rangle \, dt \in [-\infty, \infty) \text{ for all } u \in \mathbf{L}^{\phi}([0,T],\mathbb{R}^N).$$

Moreover, $O_{\phi^*}(F_v) \leq 2N_{\phi^*}(v)$.

Proof. Let $u \in \mathbf{L}^{\phi}([0,T],\mathbb{R}^N)$. Without loss of generality, let $N_\phi(u) \leq 1$ and $N_{\phi^*}(v) \leq 1$, so that

$$\int_0^T \phi(u(t))\, dt \leq 1 \text{ and } \int_0^T \phi^*(v(t))\, dt \leq 1.$$

Care must be taken handling the notion of the integral which may be negative infinite. Let

$$\begin{aligned} M_+ &= \{t \in [0,T] \mid \langle v(t), u(t)\rangle \geq 0\}, \\ M_- &= \{t \in [0,T] \mid \langle v(t), u(t)\rangle < 0\}. \end{aligned}$$

Notice that

$$\int_{M_+} \langle v(t), u(t)\rangle \, dt \leq \int_{M_+} (\phi^*(v(t)) + \phi(u(t)))\, dt \leq 2.$$

The integral of a negative measurable function, finite or not, can always be defined:

$$\int_{M_-} \langle v(t), u(t)\rangle \, dt \in [-\infty, 0].$$

Since one of the two integrals is always finite, the sum can be defined and satisfies

$$F_v(u) = \int_0^T \langle v(t), u(t)\rangle \, dt \leq 2.$$

It is easy to see that F_v is additive and $F_v(\lambda u) = \lambda F_v(u)$ for positive λ. In order to show that $F_v \in \mathbf{P}^{\phi^*}([0,T],\mathbb{R}^N)$, we have to check the monotone convergence property. Let u_n converge monotonically to u. Thus we have measurable functions $\alpha_i : [0,T] \to [0,1]$ with $u_i(t) = \alpha_i(t)u(t)$, $0 \leq \alpha_1(t) \leq \alpha_2(t) \leq \cdots \to 1$. Constructing M_+ and M_- as above, we infer from the principle of monotone convergence that

$$\int_{M_\pm} \langle v(t), u_i(t)\rangle \, dt = \int_{M_\pm} \alpha_i(t)\langle v(t), u(t)\rangle \, dt \to \int_{M_\pm} \langle v(t), u(t)\rangle \, dt.$$

Again, we notice that one of the limits is finite, so that we can take their sum and have

$$\int_0^T \langle v(t), u_i(t)\rangle \, dt \to \int_0^T \langle v(t), u(t)\rangle \, dt.$$

□

There is also a converse estimate:

Proposition 2.8. *If $v \in \mathbf{L}^1([0,T],\mathbb{R}^N)$ such that for each piecewise constant function $u \in \mathbf{L}^{\phi}([0,T],\mathbb{R}^N)$ we have*

$$\int_0^T \langle v(t), u(t)\rangle \, dt \leq M N_\phi(u),$$

then $v \in \mathbf{L}^{\phi^*}([0,T],\mathbb{R}^N)$ and $N_{\phi^*}(v) \leq M$.

Proof. Without loss of generality we may assume that $M = 1$. In this case we have to show that $\int_0^T \phi^*(v(t))\,dt \le 1$. We take the averaging approximation

$$v_n^i = \frac{n}{T}\int_{(i-1)T/n}^{iT/n} v(t)\,dt \text{ for } n \in \mathbb{N},\ i = 1\cdots n,$$
$$v_n(t) = v_n^i \text{ for } t \in [\frac{(i-1)T}{n}, \frac{iT}{n}).$$

Let u be any step function with $N_\phi(u) \le 1$. We define the averaging approximations u_n of u in the same way. A direct computation yields

$$\int_0^T \langle v_n(t), u(t)\rangle\,dt = \int_0^T \langle v(t), u_n(t)\rangle\,dt.$$

Using Jensen's inequality, we see that for $n > T$

$$\int_0^T \phi(u_n(t))\,dt = \sum_{i=1}^n \frac{T}{n}\phi(\frac{n}{T}\int_{(i-1)T/n}^{iT/n} u(t)\,dt)$$
$$\le \sum_{i=1}^n \int_{(i-1)T/n}^{iT/n} \phi(u(t))\,dt = \int_0^T \phi(u(t))\,dt \le 1.$$

Therefore, for such u we have $\int_0^T \langle v_n(t), u(t)\rangle\,dt \le 1$.

Now choose $z_n^i \in \mathbb{R}^N$ such that

$$\phi^*(v_n^i) + \phi(z_n^i) = \langle v_n^i, z_n^i\rangle.$$

Suppose first that $\frac{T}{n}\sum_{i=1}^n \phi(z_n^i) > 1$. Then there exists a factor $\beta < 1$ such that $\frac{T}{n}\sum_{i=1}^n \phi(\beta z_n^i) = 1$. Putting

$$u(t) = \beta z_n^i \text{ for } t \in [\frac{(i-1)T}{n}, \frac{iT}{n}),$$

we obtain that $\int_0^T \phi(u(t))\,dt \le 1$. Therefore

$$\frac{T}{n}\sum_{i=1}^n \phi^*(v_n^i) = \frac{T}{n}\sum_{i=1}^n \langle v_n^i, z_n^i\rangle - \frac{T}{n}\sum_{i=1}^n \phi(z_n^i)$$
$$\le \frac{1}{\beta}\int_0^T \langle v_n(t), u(t)\rangle\,dt - \frac{1}{\beta}\frac{T}{n}\sum_{i=1}^n \phi(\beta z_n^i) \le \frac{1}{\beta} - \frac{1}{\beta} = 0.$$

Now assume that $\frac{T}{n}\sum_{i=1}^n \phi(z_n^i) \le 1$. We repeat the same computation with $\beta = 1$ and obtain

$$\frac{T}{n}\sum_{i=1}^n \phi^*(v_n^i) \le 1 - 0 = 1.$$

In either case we have proved that

$$\int_0^T \phi^* v_n(t)\,dt \le 1.$$

Now, since $v_n(t) \to v(t)$ almost everywhere, we infer that

$$\phi^*(v(t)) \le \liminf_{n\to\infty} \phi^*(v_n(t)).$$

By Fatou's Lemma we conclude that

$$\int_0^T \phi^*(v(t))\,dt \leq 1.$$

□

Theorem 2.9. *Suppose that ϕ satisfies Hypothesis 2.1. Then, using the identification from Proposition 2.7, we have $\mathbf{P}^{\phi^*}([0,T],\mathbb{R}^N) = \mathbf{L}^{\phi^*}([0,T],\mathbb{R}^N)$.*

Proof. With respect to Propositions 2.7 and 2.8, all we have to show is that for each $F \in \mathbf{P}^{\phi^*}([0,T],\mathbb{R}^N)$ there exists a unique function $v \in \mathbf{L}^{\Phi^*}([0,T],\mathbb{R}^N)$ such that $F(u) = \int_0^T \langle v(t), u(t)\rangle\,dt$ for all $u \in \mathbf{L}^{\phi}([0,T],\mathbb{R}^N)$.

The uniqueness is easily proved. Suppose both v_1 and v_2 represent the same functional F. Since ϕ is bounded on bounded sets, $\mathbf{L}^{\infty}([0,T],\mathbb{R}^N)$ is contained in $\mathbf{L}^{\phi}([0,T],\mathbb{R}^N)$, so that

$$\int_0^T \langle v_1(t), u(t)\rangle = \int_0^T \langle v_2(t), u(t)\rangle \text{ for all } u \in \mathbf{L}^{\infty}([0,T],\mathbb{R}^N).$$

This implies that $v_1 = v_2$ almost everywhere.

More work is required to prove existence. Without loss of generality we assume that $F(u) \leq N_\phi(u)$ for all $u \in \mathbf{L}^{\phi}([0,T],\mathbb{R}^N)$. For any measurable $Q \subset [0,T]$ we define

$$\chi_Q(t) = \begin{cases} 1 & \text{if } t \in Q, \\ 0 & \text{if } t \notin Q. \end{cases}$$

Let $x \in \mathbb{R}^N$ be such that $\phi(\pm x) \leq 1/T$. Then the functions $\pm\chi_Q x$ are in $\mathbf{L}^{\phi}([0,T],\mathbb{R}^N)$, with $N_\phi(\pm\chi_Q x) \leq 1$. Therefore we can define $\mu_x(Q) = F(\chi_Q x)$, mapping the Lebesgue σ-algebra into $[-1,1]$. Since F is additive and has the monotone convergence property (recall Definition 2.3), the set function μ_x is σ-additive, i.e., it is a signed measure. If the Lebesgue measure of Q is zero, then $\chi_Q x = 0$ almost everywhere, so that $\mu_x(Q) = 0$. Therefore, μ_x is absolutely continuous. By the Radon-Nikodym theorem there exists $v_x \in \mathbf{L}^1([0,T],\mathbb{R})$ such that for all Lebesgue sets Q

$$\mu_x(Q) = \int_Q v_x(t)\,dt.$$

Since ϕ is continuous, we can choose a basis $x_1, \cdots, x_N$ of $\mathbb{R}^N$ such that $\phi(\pm x_i) \leq 1/T$. Let y_i be the dual basis $\langle y_i, x_j\rangle = \delta_{i,j}$ and define

$$v(t) = \sum_{i=1}^N v_{x_i}(t) y_i.$$

Then the additivity and homogeneity of F imply that for all step functions $u(t) = \sum_{j=1}^n \chi_{Q_j}(t) u_j$ with pairwise disjoint measurable Q_j and vectors $u_j \in \mathbb{R}^N$ the following identity holds:

$$F(u) = \sum_{j=1}^N \sum_{i=1}^N F(\chi_{Q_j} x_i)\langle y_i, u_j\rangle = \int_0^T \langle v(t), u(t)\rangle\,dt.$$

Before we extend this identity to general u, we infer from Proposition 2.8 that $v \in \mathbf{L}^{\phi^*}([0,T],\mathbb{R}^N)$ with $N_{\phi^*}(v) \leq 1$.

Now we assume first that u is bounded. As a consequence, both u and $-u$ are contained in $\mathbf{L}^\phi([0,T],\mathbb{R}^N)$. Without loss of generality we may assume that $N_\phi(u) \leq 1$ and $N_\phi(-u) \leq 1$. We take the averaging approximation

$$u_n(t) = \frac{n}{T}\int_{(i-1)T/n}^{iT/n} u(s)\,ds \text{ for } t \in [\frac{(i-1)T}{n}, \frac{iT}{n}).$$

Clearly, $u_n(t) \to u(t)$ almost everywhere. From the boundedness of u we infer that u_n are uniformly bounded. The latter fact implies that

$$\int_0^T \langle v(t), u_n(t)\rangle \to \int_0^T \langle v(t), u(t)\rangle \text{ as } n \to \infty.$$

The uniform boundedness of u_n implies that $\pm u_n$ and $\pm(u-u_n)$ are also contained in $\mathbf{L}^\phi([0,T],\mathbb{R}^N)$, so that in particular $F(\pm u_n) = \pm F(u_n)$ is defined. Moreover, from Lemma 2.6 we see that $N_\phi(\pm(u-u_n)) \to 0$ as $n \to \infty$. Since

$$\pm F(u) = F(\pm u_n) + F(\pm(u-u_n)) \leq \pm F(u_n) + N_\phi(\pm(u-u_n))$$

we infer that

$$\pm F(u) \leq \limsup_{n\to\infty} \pm F(u_n),$$

i.e.

$$F(u) = \lim_{n\to\infty} F(u_n) = \lim_{n\to\infty}\int_0^T \langle v(t), u_n(t)\rangle\,dt = \int_0^T \langle v(t), u(t)\rangle\,dt.$$

Finally, we extend this result to general $u \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$. For this purpose we define

$$u_n(t) = \begin{cases} u(t) & \text{if } |u(t)| \leq n, \\ 0 & \text{else.} \end{cases}$$

Notice that each u_n is bounded and converges monotonically to u. Therefore, by the monotone convergence property

$$F(u) = \lim_{n\to\infty} F(u_n) = \lim_{n\to\infty}\int_0^T \langle v(t), u_n(t)\rangle\,dt = \int_0^T \langle v(t), u(t)\rangle\,dt.$$

□

The theorem above implies reflexivity whenever all $F \in (\mathbf{L}^\phi)^*([0,T],\mathbb{R}^N)$ have the monotone convergence property. This is the case in classical Orlicz spaces when ϕ satisfies Hypothesis 2.2 (compare the reflexivity theorem [10, Corollary 4.1.9]). If $-\mathbf{L}^\phi([0,T],\mathbb{R}^N) \neq \mathbf{L}^\phi([0,T],\mathbb{R}^N)$, we will not have reflexivity in general, but we can decompose F into a part $f \in \mathbf{L}^{\phi^*}([0,T],\mathbb{R}^N)$ and a "singular" part:

Theorem 2.10. *Let ϕ satisfy Hypothesis 2.1 and Hypothesis 2.2. Then for each $F \in (\mathbf{L}^\phi)^*([0,T],\mathbb{R}^N)$, there exists a unique $F_1 \in \mathbf{P}^{\phi^*}([0,T],\mathbb{R}^N)$ such that $F_1(u) = F(u)$ whenever $\pm u \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$. Moreover we have*

1. $F_1(u) = \lim F(u_n)$ *if u_n converge monotonically to u and $\pm u_n \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$.*
2. $F_1(u) \leq O_{\phi^*}(F)N_\phi(u)$.
3. $F(u) \leq F_1(u)$ *for all $u \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$.*

Proof. Let $F \in (\mathbf{L}^\phi)^*([0,T],\mathbb{R}^N)$. Without loss of generality we may assume that $O_{\phi^*}(F) \leq 1$. Suppose first that $F_1 \in \mathbf{P}^{\phi^*}([0,T],\mathbb{R}^N)$ exists such that $F_1(u) = F(u)$

if $\pm u \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$. Let u_n converge monotonically to $u \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$ such that $\pm u_n \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$. Then

$$F_1(u) = \lim_{n\to\infty} F_1(u_n) = \lim_{n\to\infty} F(u_n).$$

This implies assertion (1), assertion (2), and uniqueness. From Lemma 2.5 we infer that $N_\phi(u-u_n)$ converges to 0 which implies assertion (3).

To prove the existence of such a decomposition, notice first that if u_n is any sequence converging monotonically to u, then $\lim_{n\to\infty} F(u_n)$ exists (finite or negative infinite). In fact, by Lemma 2.5 we have that $F(u_{n+k})-F(u_n) \le N_{\phi^*}(u-u_n) \to 0$, which implies easily that $\limsup_{n\to\infty} F(u_n) \le \liminf_{n\to\infty} F(u_n)$. Next we show that the limit does not depend on the sequence u_n if $\pm u_n \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$. For this purpose, let $u_n = \alpha_n u$ and $v_n = \beta_n u$ be two sequences converging monotonically to u such that $\pm u_n \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$. For now, no restriction is imposed on $-v_n$. We start with the following estimates:

$$\begin{aligned} F(\alpha_m u) &= -F(-\alpha_m u) \ge F(\alpha_m\beta_n u) - N_\phi((1-\beta_n)(-\alpha_m u)),\\ F(\beta_n u) &\le F(\alpha_m\beta_n u) + N_\phi((1-\alpha_m)\beta_n u)\\ &\le F(\alpha_m\beta_n u) + N_\phi((1-\alpha_m)u). \end{aligned}$$

Using Lemma 2.5, we can first choose m_k such that $N_\phi((1-\alpha_{m_k})u) \le \frac{1}{k}$. Then we use that $-u_{m_k} \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$ and choose n_k such that $N_\phi((1-\beta_{n_k})(-\alpha_{m_k}u)) \le \frac{1}{k}$. The estimates above yield that $F(u_{m_k}) \ge F(v_{n_k}) - \frac{2}{k}$. Therefore

$$\lim_{m\to\infty} F(u_m) \ge \lim_{n\to\infty} F(v_n). \tag{2.2}$$

In particular, if also $\pm v_n \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$, then by symmetry $F(u_m)$ and $F(v_n)$ have the same limit.

We may now define $F_1 u = \lim_{n\to\infty} F(u_n)$ where u_n is any sequence converging monotonically to u with $\pm u_n \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$. It is easily seen that $F_1 \in (\mathbf{L}^\phi)^*([0,T],\mathbb{R}^N)$ with $O_{\phi^*}(F_1) \le 1$. If $\pm u \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$, then we may put $u_n = u$ and obtain $F(u) = F_1(u)$. To show that $F_1 \in \mathbf{P}^{\phi^*}([0,T],\mathbb{R}^N)$, let $u_m = \alpha_m u$ be a sequence converging monotonically to u with $\pm u_m \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$ and $v_n = \beta_n u$ a sequence converging monotonically to u (without the requirement that $-v_n \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$). From Lemma 2.5 we obtain immediately that $F_1(u) \le \lim_{n\to\infty} F_1(v_n)$. On the other hand, applying (2.2) to F_1 instead of F, we see that $F_1(u) = \lim_{n\to\infty} F_1(u_m) \ge \lim_{n\to\infty} F_1(v_n)$. □

By Theorem 2.10 and Theorem 2.9, we may define

Definition 2.4. Given $F \in (\mathbf{L}^\phi)^*([0,T],\mathbb{R}^N)$, let $f \in \mathbf{L}^{\phi^*}([0,T],\mathbb{R}^N)$ be such that

$$\int_0^T f(t)u(t)\,dt = F(u) \text{ for all } u \text{ such that } \pm u \in \mathbf{L}^\phi([0,T],\mathbb{R}^N).$$

We call f the function representative of F.

Example 2.1. Let $N = 1$, let

$$\phi(x) = \begin{cases} x^2 & \text{if } x \le 0,\\ x^4 & \text{else.} \end{cases}$$

Then

$$\mathbf{L}^\phi([0,T],\mathbb{R}) = \{f_+ - f_- \mid f_+ \in \mathbf{L}^4([0,T],[0,\infty)), f_- \in \mathbf{L}^2([0,T],[0,\infty))\}.$$

Notice that ϕ satisfies Hypothesis 2.2. Let

$$F(u) = \begin{cases} \int_0^T g(t)u(t)\,dt & \text{if } u \in \mathbf{L}^4([0,T],\mathbb{R}), \\ -\infty & \text{else,} \end{cases}$$

with $g \in \mathbf{L}^2([0,T],\mathbb{R})$. Then the function representative of F is

$$F_1(u) = \int_0^T g(t)u(t)\,dt.$$

2.4. **Notions of convergence.** Since the difference of two functions in $\mathbf{L}^\phi([0,T],\mathbb{R}^N)$ need not be contained in $\mathbf{L}^\phi([0,T],\mathbb{R}^N)$, the usual definition of a metric by the Luxemburg norm must fail. We still may introduce a notion of weak* convergence. We have to modify the metric on the reals in order to include the value $-\infty$, which is a possible value of the functionals in $(\mathbf{L}^\phi)^*([0,T],\mathbb{R}^N)$. Strong convergence can be defined by weak* convergence which is uniform on sets with bounded N_ϕ.

Definition 2.5. Let $(F_\lambda)_{\lambda\in\Lambda}$ be a net in $(\mathbf{L}^\phi)^*([0,T],\mathbb{R}^N)$ and let

$$F \in (\mathbf{L}^\phi)^*([0,T],\mathbb{R}^N).$$

We topologize the semi-closed interval $[-\infty,\infty)$ by the metric $d(x,y) = |e^x - e^y|$.

(1) We say that F_λ converges to F weakly* if $F_\lambda(u) \to F(u)$ in $[-\infty,\infty)$ for all $u \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$.
(2) We say that F_λ converges to F strongly if $F_\lambda(u) \to F(u)$ in $[-\infty,\infty)$ uniformly for all $u \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$ with $N_\phi(u) \le 1$.

In the case of an Orlicz vector space, these notions of convergence are just weak* convergence and convergence with respect to the Orlicz norm. In Banach spaces, the following lemma is the well-known weak* compactness of the closed unit ball.

Theorem 2.11. *Let $(F_\lambda)_{\lambda\in\Lambda}$ be a net in $(\mathbf{L}^\phi)^*([0,T],\mathbb{R}^N)$, bounded in the sense that $O_{\phi^*}(F_\lambda) \le C$ for all λ and some fixed constant C. Then there exists a subnet $\tilde\Lambda$ of λ and a functional $F \in (\mathbf{L}^\phi)^*([0,T],\mathbb{R}^N)$ such that $(F_\lambda)_{\lambda\in\tilde\Lambda}$ converges to F in the weak* sense.*

Proof. Let $B = \{u \in \mathbf{L}^\phi([0,T],\mathbb{R}^N) \mid N_\phi(u) \le 1\}$. By assumption, each F_λ maps B into the compact set $[-\infty, C]$. Due to Tikhonov's theorem there exists a subnet $\tilde\Lambda$ and a function $F : B \to [-\infty, C]$ such that $(F_\lambda(u))_{\lambda\in\tilde\Lambda} \to F(u)$ for all $u \in B$. It is now easy to extend F from B to the whole cone $\mathbf{L}^\phi([0,T],\mathbb{R}^N)$ by $F(u) = \mu F(\frac{1}{\mu}u)$ and to show that F is in fact additive and positive homogeneous. □

Lemma 2.12. *Let ϕ satisfy Hypotheses 2.1 and 2.2. Let $(f_\lambda)_{\lambda\in\Lambda}$ be a net in $\mathbf{L}^{\phi*}([0,T],\mathbb{R}^N)$ and $F \in (\mathbf{L}^\phi)^*([0,T],\mathbb{R}^N)$ be such that f_λ converges weakly* to F. Moreover, let f be the function representative of F. Then*

(1) *f_λ converges to f weakly in a suitable space $\mathbf{L}^q([0,T],\mathbb{R}^N)$ with $q \in (1,\infty)$.*
(2)

$$\int_0^T \phi^*(f(t))\,dt \le \liminf_{\lambda\in\Lambda} \int_0^T \phi^*(f_\lambda(t))\,dt.$$

Proof. By Lemma 2.4, we have $\mathbf{L}^p([0,T],\mathbb{R}^N) \subset \mathbf{L}^\phi([0,T],\mathbb{R}^N)$ for some $p \in (1,\infty)$. Evidently, for $u \in \mathbf{L}^p([0,T],\mathbb{R}^N)$ we have $\pm u \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$, thus

$$\int_0^T f_\lambda(t)u(t)\,dt \to F(u(t)) = \int_0^T f(t)u(t)\,dt.$$

Consequently, $f_\lambda \to f$ weakly in $\mathbf{L}^q$ with $\frac{1}{p} + \frac{1}{q} = 1$. Now part (2) is proved easily using the weak closedness of convex closed sets in $\mathbf{L}^q([0,T],\mathbb{R}^N)$ and the lower semicontinuity of ϕ^*. $\square$

Lemma 2.13. *Let $(u_\lambda)_{\lambda\in\Lambda}$ be a net in $\mathbf{L}^\phi([0,T],\mathbb{R}^N)$ converging in the weak* sense to $U \in (\mathbf{L}^\phi)^*([0,T],\mathbb{R}^N)$. Let u be the function representative of U. Let $v_\lambda(t) = \int_0^t u_\lambda\,dt$ and $v(t) = \int_0^t u(t)\,dt$. Then $v_\lambda \to v$ uniformly on $[0,T]$. In particular, v_λ converges to v strongly in $\mathbf{L}^{\phi^*}([0,T],\mathbb{R}^N)$.*

Proof. By Lemma 2.12, u_λ converges to u weakly in $\mathbf{L}^q([0,T],\mathbb{R}^N)$ for some $q \in (1,\infty)$. It is well known that this implies uniform convergence of the antiderivatives. Finally,

$$|\int_0^T \langle v_\lambda(t), f(t)\rangle\,dt - \int_0^T \langle v(t), f(t)\rangle\,dt| \le \|v_\lambda - v\|_\infty \int_0^T |f(t)|\,dt.$$

The latter integral is bounded whenever $N_\phi(f)$ is bounded. $\square$

2.5. **Some technical lemmas.** We close this section with some technicalities.

Lemma 2.14. *Let u_n be a sequence in $\mathbf{L}^1([0,T],\mathbb{R})$ converging to a function $u(t)$ almost everywhere. Suppose that there exists a sequence g_n converging to a function g almost everywhere and in the sense of $\mathbf{L}^1([0,T],\mathbb{R})$, such that $|u_n(s)| \le g_n(s)$ almost everywhere. Then*

$$\int_0^T |u_n(s) - u(s)|\,ds \to 0.$$

Proof. Put

$$w_n(s) = \max[-g(s), \min(g(s), u_n(s))].$$

Then $|w_n(s)| \le g(s)$, and

$$\int_0^T |w_n(s) - u_n(s)|\,ds \le \int_0^T |g_n(s) - g(s)|\,ds.$$

Since the latter integral converges to 0, the sequence u_n converges in $\mathbf{L}^1([0,T],\mathbb{R})$ if and only if the sequence w_n converges. Moreover, w_n converges to the same limit as u_n almost everywhere. Therefore, the lemma follows from the dominated convergence principle. $\square$

Lemma 2.15. *Let $v \in \mathbf{W}^{1,1}([0,T],\mathbb{R}^N)$ be such that $\dot{v} \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$ and $f \in \mathbf{L}^{\phi^*}([0,T],\mathbb{R}^N)$. Let $t \in (0,T)$. Then*

$$\limsup_{h\to 0+} \frac{1}{h}\int_0^t \langle v(s+h) - v(s), f(s)\rangle\,ds \le \int_0^t \langle \dot{v}(s), f(s)\rangle\,ds.$$

Proof. Without loss of generality we may assume that $N_\phi(\dot v) \le 1$ and $N_{\phi^*}(f) \le 1$. For $h > 0$, put $u_h(s) = h^{-1}\langle v(s+h) - v(s), f(s)\rangle$ and $u_0(s) = \langle \dot v(s), f(s)\rangle$. Then $u_h(s) \to u_0(s)$ almost everywhere as $h \to 0$. Moreover, by Jensen's inequality,

$$u_h(s) \le \phi(\frac{1}{h}\int_s^{s+h} \dot v(\tau)\,d\tau) + \phi^*(f(s)) \le \frac{1}{h}\int_s^{s+h} \phi(\dot v(\tau))\,d\tau + \phi^*(f(s)).$$

For $h \ge 0$ we decompose u_h in its positive and negative part: $u_h(s) = u_h^+(s) - u_h^-(s)$ with $u_h^+(s) = \max(u_h(s), 0)$ and $u_h^-(s) = \max(-u_h(s), 0)$. Put

$$g_h(s) = \frac{1}{h}\int_s^{s+h} \phi(\dot v(\tau))\,d\tau + \phi^*(f(s)).$$

Since $\phi(\dot v)$ and $\phi^*(f)$ are integrable, we have that

$$g_h \to \phi(\dot v) + \phi^*(f)$$

in $\mathbf{L}^1([0,T],\mathbb{R})$. For the positive part we obtain

$$|u_h^+(s)| = u_h^+(s) \le g_h(s),$$

so that Lemma 2.14 implies that u_h^+ converges to u_0^+ in $\mathbf{L}^1([0,T],\mathbb{R})$. For the negative part we utilize Fatou's Lemma and see that

$$\int_0^t u_0^-(s)\,ds \le \liminf_{h\to 0}\int_0^t u_h^-(s)\,ds.$$

Summing up we have that

$$\int_0^t u_0(s)\,ds \ge \limsup_{h\to 0}\int_0^t u_h(s)\,ds$$

which is the desired result. ☐

We recall also the following special case of [12, Corollary 1B]:

Lemma 2.16. *Let $\phi : \mathbb{R}^N \to [0,\infty]$ be convex and lower semicontinuous. Let $x \in \mathbf{L}^\infty([0,T],\mathbb{R}^N)$ and $y \in \mathbf{L}^1([0,T],\mathbb{R}^N)$ be such that for all $u \in \mathbf{L}^\infty([0,T],\mathbb{R}^N)$ the following holds:*

$$\int_0^T \phi(x(t)+u(t))\,dt - \int_0^T \phi(x(t))\,dt \ge \int_0^T \langle y(t), u(t)\rangle\,dt.$$

Then $y(t) \in \partial\phi(x(t))$ almost everywhere.

3. The existence theorem

3.1. **Statement of the result.** We rewrite problem (1.2) in the form

$$\begin{aligned}(\sigma(t), A(t)) &\in \partial\psi(\epsilon(t), V(t)),\\ \dot V(t) &\in \partial\phi^*(-A(t))\end{aligned} \tag{3.1}$$

with initial conditions

$$\epsilon(0) = \epsilon_0,\ V(0) = V_0. \tag{3.2}$$

Throughout this section, we assume that $\phi : \mathbb{R}^N \to [0,\infty)$ is a convex function satisfying Hypotheses 2.1 and 2.2. The function $\psi : \mathbb{R}^M \times \mathbb{R}^N \to [0,\infty]$ is convex and lower semicontinuous, with $\psi(0,0) = 0$. Moreover, we assume

Hypothesis 3.1.

(1) For arbitrary small $C_1 > 0$ there exist constants $C_2, C_3 > 0$ such that for all $\xi \in \mathbb{R}^M$, $v \in \mathbb{R}^N$
$$|\xi| \le C_1\psi(\xi, v) + C_2|v| + C_3.$$
(2) If $\infty > \psi(\frac{\xi_1+\xi_2}{2}, v) \ge \frac{1}{2}\psi(\xi_1, v) + \frac{1}{2}\psi(\xi_2, v)$ for some $\xi_1, \xi_2 \in \mathbb{R}^M$ and some $v \in \mathbb{R}^N$, then $\xi_1 = \xi_2$.
(3) For each $L_1 > 0$ there exist constants $L_2, L_3 > 0$ such that for each $v \in \mathbb{R}^N$ with $|v| \le L_1$ there exists $\xi \in \mathbb{R}^M$ with $|\xi| \le L_2$ and $\psi(\xi, v) \le L_3$.

The data σ, ϵ_0, and V_0 satisfy the following hypothesis. Notice that we require fairly strong smoothness assumptions on σ, combined with a compatibility assumption on the initial data.

Hypothesis 3.2. $\sigma \in \mathbf{W}^{1,1}([0,T], \mathbb{R}^M)$, $\epsilon_0 \in \mathbb{R}^M$, $V_0 \in \mathbb{R}^N$ such that $\psi(\epsilon_0, V_0) < \infty$ and there exists some $A_0 \in \mathbb{R}^N$ with $(\sigma(0), A_0) \in \partial\psi(\epsilon_0, V_0)$.

The aim of this section is to prove the following theorem:

Theorem 3.1. *If Hypotheses 2.1, 2.2, 3.1, and 3.2 are satisfied, then there exists at least one set of functions*
$$\epsilon \in \mathcal{C}([0,T], \mathbb{R}^M), \quad V \in \mathbf{W}^{1,1}([0,T], \mathbb{R}^N), \quad -A \in \mathbf{L}^{\phi^*}([0,T], \mathbb{R}^N)$$
such that $\dot{V} \in \mathbf{L}^{\phi}([0,T], \mathbb{R}^N)$, the equation (3.1) *is satisfied almost everywhere, and the initial conditions* (3.2) *hold.*

The proof will be given by an approximation procedure and takes the remainder of this section. This is a short outline:

(1) For $\lambda > 0$ we replace the subdifferentials $\partial\phi^*$ and $\partial\psi$ by their Yosida approximations. Consequently, the approximating problem yields an ordinary differential equation with Lipschitz continuous right hand side. We will investigate the limit as $\lambda \to 0$.
(2) We derive a priori bounds on the approximating state $\dot{V}_\lambda$ and costate A_λ in suitable Orlicz spaces. Moreover we derive $\mathbf{L}^\infty$ a priori bounds for the approximating state ϵ_λ.
(3) We take the function representatives A and $\dot{V}$ of weak* cluster points of the approximating $\dot{V}_\lambda$ and A_λ.
(4) We prove that ϵ_λ converges to a continuous function ϵ.
(5) We prove that ϵ, A, and V solve (3.1).

3.2. **The approximating problem.**

Definition 3.1. For $\lambda > 0$ we define the following approximate potentials:
$$\begin{aligned}
\phi_\lambda(x) &= \phi(x) + \frac{\lambda}{2}|x|^2, \\
\phi^*_\lambda(x) &= \inf_{y \in \mathbb{R}^N}(\phi^*(y) + \frac{1}{2\lambda}|y - x|^2), \\
\psi_\lambda(\xi, x) &= \inf_{\eta \in \mathbb{R}^M,\, y \in \mathbb{R}^N}(\psi(\eta, y) + \frac{1}{2\lambda}|\eta - \xi|^2 + \frac{1}{2\lambda}|y - x|^2).
\end{aligned}$$

With these potentials, we consider the approximating problem

$$\begin{aligned}(\sigma_\lambda(t), A_\lambda(t)) &\in \partial\psi_\lambda(\epsilon_\lambda(t), V_\lambda(t)),\\ \dot{V}_\lambda(t) &\in \partial\phi_\lambda^*(-A_\lambda(t)),\\ \dot{\epsilon}_\lambda(t) &= \frac{1}{\lambda}(\sigma(t) - \sigma_\lambda(t)),\end{aligned} \tag{3.3}$$

with initial conditions

$$\epsilon_\lambda(0) = \epsilon_0,\ V_\lambda(0) = V_0. \tag{3.4}$$

Remark 3.2. The function ϕ_λ^* as defined above is in fact the Fenchel-Legendre conjugate of ϕ_λ.

Proof. This follows from standard rules to calculate Fenchel transforms, see for instance [13, 11(3), Example 11.11, Theorem 11.23]. □

Lemma 3.3. *The subdifferentials $\partial\phi_\lambda^*$ and $\partial\psi_\lambda$ are Lipschitz continuous, in fact $\partial\phi_\lambda^*(x) = \frac{1}{\lambda}(x - J_\lambda(x))$ with $J_\lambda = (1 - \lambda\partial\phi_\lambda^*)^{-1}$. (A similar result holds for $\partial\psi_\lambda$.)*

Proof. See [2, Theorem 5.2]. □

Lemma 3.4. *For each $\lambda > 0$, problem* (3.3) *admits a unique solution.*

Proof. Because of Lemma 3.3, this is an ordinary differential equation with a Lipschitz continuous right hand side in a Banach space. □

Definition 3.2. Throughout the rest of this paper, let

$$\begin{aligned}&\epsilon_\lambda \in \mathcal{C}^1([0,T], \mathbb{R}^M), \quad \sigma_\lambda \in \mathcal{C}([0,T], \mathbb{R}^M),\\ &V_\lambda \in \mathcal{C}^1([0,T], \mathbb{R}^N), \quad A_\lambda \in \mathcal{C}([0,T], \mathbb{R}^N)\end{aligned}$$

be a solution of (3.3), (3.4).

3.3. **A priori estimates.** We estimate $\psi_\lambda(\epsilon(t), V(t))$, $N_\phi(\dot{V}_\lambda(\cdot))$ and $N_{\phi^*}(-A_\lambda(\cdot))$ in terms of the initial conditions and the energy fed into the system by the forcing stress, i.e., $\int_0^t \langle \sigma(s), \dot{\epsilon}_\lambda(s)\rangle\, ds$. Subsequently, we will derive a priori bounds for the energy input.

Lemma 3.5. *The following estimate holds for all $t \in [0,T]$:*

$$\begin{aligned}&\psi_\lambda(\epsilon_\lambda(t), V_\lambda(t)) + \int_0^t \phi_\lambda(\dot{V}_\lambda(s))\, ds + \int_0^t \phi_\lambda^*(-A_\lambda(s))\, ds\\ &\qquad = \psi_\lambda(\epsilon_0, V_0) + \int_0^t \langle \sigma_\lambda(s), \dot{\epsilon}_\lambda(s)\rangle\, ds.\end{aligned} \tag{3.5}$$

Proof. Since $\dot{V}_\lambda \in \partial\phi_\lambda^*(-A_\lambda)$ we have the identity

$$\langle -A_\lambda(s), \dot{V}_\lambda(s)\rangle = \phi_\lambda(\dot{V}_\lambda(s)) + \phi_\lambda^*(-A_\lambda(s)).$$

Notice that ψ_λ is continuously differentiable and its gradient is given by the pair $(\sigma_\lambda, A_\lambda)$. The chain rule implies

$$\begin{aligned}
&\psi_\lambda(\epsilon_\lambda(t), V_\lambda(t))\\
&= \psi_\lambda(\epsilon_0, V_0) + \int_0^t \langle \sigma_\lambda(s), \dot\epsilon_\lambda(s)\rangle\, ds + \int_0^t \langle A_\lambda(s), \dot V_\lambda(s)\rangle\, ds\\
&= \psi_\lambda(\epsilon_0, V_0) + \int_0^t \langle \sigma_\lambda(s), \dot\epsilon_\lambda(s)\rangle\, ds\\
&\quad - \int_0^t \phi_\lambda(\dot V_\lambda(s))\, ds - \int_0^t \phi_\lambda^*(-A_\lambda(s))\, ds.
\end{aligned}$$

□

Lemma 3.6. *Choose $B_\lambda(t) \in \mathbb{R}^N$ such that*

$$\phi_\lambda^*(-A_\lambda(t)) = \phi^*(-B_\lambda(t)) + \frac{1}{2\lambda}|A_\lambda(t) - B_\lambda(t)|^2. \tag{3.6}$$

Then $\dot V_\lambda(t) \in \partial\phi^*(-B_\lambda(t))$. Moreover, the following estimate holds

$$\begin{aligned}
&\psi_\lambda(\epsilon_\lambda(t), V_\lambda(t)) + \int_0^t \phi(\dot V_\lambda(s))\, ds + \int_0^t \phi^*(-B_\lambda(s))\, ds\\
&\quad + \int_0^t \frac{\lambda}{2}|\dot V_\lambda(s)|^2\, ds + \int_0^t \frac{1}{\lambda}|\sigma_\lambda(s) - \sigma(s))|^2\, ds\\
&\quad + \int_0^t \frac{1}{2\lambda}|B_\lambda(s) - A_\lambda(s)|^2\, ds\\
&= \psi(\epsilon_\lambda(t) - \lambda\sigma_\lambda(t), V_\lambda(t) - \lambda A_\lambda(t))\\
&\quad + \int_0^t \phi(\dot V_\lambda(s))\, ds + \int_0^t \phi^*(-B_\lambda(s))\, ds\\
&\quad + \frac{\lambda}{2}|\sigma_\lambda(s)|^2\, ds + \frac{\lambda}{2}|A_\lambda(s)|^2\, ds + \int_0^t \frac{\lambda}{2}|\dot V_\lambda(s)|^2\, ds\\
&\quad + \int_0^t \frac{1}{\lambda}|\sigma_\lambda(s) - \sigma(s)|^2\, ds + \int_0^t \frac{1}{2\lambda}|B_\lambda(s) - A_\lambda(s)|^2\, ds\\
&\le \psi(\epsilon_0, V_0) + \int_0^t \langle \sigma(s), \dot\epsilon_\lambda(s)\rangle\, ds.
\end{aligned} \tag{3.7}$$

Proof. Let $X \in \mathbb{R}^N$. Using the definition of ϕ_λ^* and the fact that $\dot V_\lambda(t) \in \partial\phi_\lambda^*(A_\lambda(t))$ we have the following estimate:

$$\begin{aligned}
&\phi^*(-B_\lambda(t) + X) - \phi^*(-B_\lambda(t))\\
&= \phi^*(-A_\lambda(t) + X - (-A_\lambda(t) + B_\lambda(t)))\\
&\quad + \frac{1}{2\lambda}|B_\lambda(t) - A_\lambda(t)|^2 - \phi_\lambda^*(-A_\lambda(t))\\
&\ge \phi_\lambda^*(-A_\lambda(t) + X) - \phi_\lambda^*(-A_\lambda(t)) \ge \langle \dot V_\lambda(t), X\rangle.
\end{aligned}$$

This shows that $\dot V_\lambda(t) \in \partial\phi^*(-B_\lambda(t))$. Now notice that $\psi_\lambda(\epsilon_0, V_0) \le \psi(\epsilon_0, V_0)$ and

$$\begin{aligned}
&\langle \sigma_\lambda(s), \dot\epsilon_\lambda(s)\rangle = \langle \sigma(s), \dot\epsilon_\lambda(s)\rangle - \langle \lambda\dot\epsilon_\lambda(s), \dot\epsilon_\lambda(s)\rangle\\
&= \langle \sigma(s), \dot\epsilon_\lambda(s)\rangle - \frac{1}{\lambda}|\sigma_\lambda(s) - \sigma(s)|^2.
\end{aligned}$$

Using the definitions of ϕ_λ and ϕ^*_λ in estimate (3.5), we obtain

$$\begin{aligned}&\psi_\lambda(\epsilon_\lambda(t), V_\lambda(t)) + \int_0^t \phi(\dot V_\lambda(s))\,ds + \frac{\lambda}{2}|\dot V_\lambda(s)|^2 \\ &+ \int_0^t \phi^*(-B_\lambda(s))\,ds + \frac{1}{2\lambda}\int_0^t |A_\lambda(s) - B_\lambda(s)|^2\,ds \\ \le\ & \psi(\epsilon_0, V_0) + \int_0^t \langle \sigma(s), \dot\epsilon_\lambda(s)\rangle\,ds - \frac{1}{\lambda}\int_0^t |\sigma_\lambda(s) - \sigma(s)|^2\,ds.\end{aligned}$$

Notice also that from $(\sigma_\lambda(t), A_\lambda(t)) \in \partial\psi_\lambda(\epsilon_\lambda(t), V_\lambda(t))$ we infer that

$$\psi_\lambda(\epsilon_\lambda(t), V_\lambda(t)) = \psi(\epsilon_\lambda(t) - \lambda\sigma_\lambda(t), V_\lambda(t) - \lambda A_\lambda(t)) + \frac{\lambda}{2}|\sigma_\lambda(t)|^2 + \frac{\lambda}{2}|A_\lambda(t)|^2.$$

□

All we need therefore is an estimate on the input energy. It is here the coercivity hypothesis 3.1 comes in.

Lemma 3.7. *If ψ satisfies Hypothesis 3.1, then for each $C_1 > 0$ the constants C_2 and C_3 can be chosen such that for all $\lambda \in (0, 1]$, $\xi \in \mathbb{R}^M$, $v \in \mathbb{R}^N$*

$$|\xi| \le C_3 + C_2|v| + C_1\psi_\lambda(\xi, v).$$

Proof. First choose C_3, C_2 according to Hypothesis 3.1. Put

$$\tilde C_3 = C_3 + \frac{1 + C_2^2}{2C_1}.$$

For any $\xi, \eta \in \mathbb{R}^M$, $v, w \in \mathbb{R}^N$, $\lambda \in (0, 1]$, we obtain the estimate

$$\begin{aligned}&\tilde C_3 + C_2|v| + C_1(\psi(\eta, w) + \frac{1}{2\lambda}|\eta - \xi|^2 + \frac{1}{2\lambda}|v - w|^2) \\ \ge\ & C_3 + \frac{1}{2}(C_1|\xi - \eta|^2 - 2|\xi - \eta| + \frac{1}{C_1}) \\ &+ \frac{1}{2}(C_1|v - w|^2 - 2C_2|v - w| + \frac{C_2^2}{C_1}) \\ &+ C_1\psi(\eta, w) + |\xi - \eta| + C_2|v - w| + C_2|v| \\ \ge\ & C_3 + C_2|w| + C_1\psi(\eta, w) - |\eta| + |\xi| \\ \ge\ & |\xi|.\end{aligned}$$

Given ξ and v, we take the infimum for all η, w and obtain

$$\tilde C_3 + C_2|v| + C_1\psi_\lambda(\xi, v) \ge |\xi|.$$

□

Lemma 3.8. *For all constants $C_2 > 0$, $C_1 > 0$ there exists some constant $C_4 > 0$ (depending on V_0 and T) such that for all functions $V \in \mathbf{W}^{1,1}([0, T], \mathbb{R}^N)$ with $V(0) = V_0$ and all $t \in (0, T]$*

$$(C_2 + 1)|V(t)| \le C_1\int_0^t \phi(\dot V(s))\,ds + C_4.$$

Proof. Since ϕ is coercive, there exists some constant $L > 0$ such that $C_1\phi(\dot V(s)) \geq (C_2+1)|\dot V(s)|$ whenever $|\dot V(s)| \geq L$. Put

$$Q = \{t \in [0,T] \mid |\dot V(t)| \geq L\}.$$

Then

$$\begin{aligned}
&(C_2+1)|V(t)| \leq (C_2+1)|V_0| + (C_2+1)\int_0^t |\dot V(s)|\,ds \\
&\leq (C_2+1)|V_0| + (C_2+1)\int_{[0,t]\setminus Q} |\dot V(s)|\,ds + (C_2+1)\int_{[0,t]\cap Q} |\dot V(s)|\,ds \\
&\leq (C_2+1)|V_0| + (C_2+1)TL + C_1\int_{[0,t]\cap Q} \phi(\dot V(s))\,ds.
\end{aligned}$$

The desired estimate follows with $C_4 = (C_2+1)(|V_0| + TL)$. □

Lemma 3.9. *Given $V_0 \in \mathbb{R}^N$, $\epsilon_0 \in \mathbb{R}^M$ and $\sigma \in \mathbf{W}^{1,1}([0,T],\mathbb{R}^M)$, there exist uniform bounds independent of $\lambda \in (0,1]$ for the following terms:*

$$\begin{aligned}
&\sup_{t\in[0,T]} \psi_\lambda(\epsilon_\lambda(t), V_\lambda(t)),\ \sup_{t\in[0,T]} |\epsilon_\lambda(t)|,\ \sup_{t\in[0,T]} |V_\lambda(t)|,\ \sup_{t\in[0,T]} \lambda^{1/2}|\sigma_\lambda(t)|,\\
&\frac{1}{2\lambda}\int_0^T |B_\lambda(t) - A_\lambda(t)|^2\,dt,\ \frac{1}{\lambda}\int_0^T |\sigma_\lambda(t) - \sigma(t)|^2\,dt,\\
&\int_0^T \phi(\dot V_\lambda(t))\,dt,\ \int_0^T \phi^*(-B_\lambda(t))\,dt.
\end{aligned}$$

Proof. Because of Lemma 3.6, all estimates follow if we can find a priori bounds for

$$\int_0^t \langle \sigma(s), \dot\epsilon_\lambda(s)\rangle\,ds = \langle \sigma(t), \epsilon_\lambda(t)\rangle - \langle \sigma(0), \epsilon_0\rangle - \int_0^t \langle \dot\sigma(s), \epsilon_\lambda(s)\rangle\,ds.$$

Since $\sigma \in \mathbf{W}^{1,1}([0,T],\mathbb{R}^M)$ is fixed, all we need is a uniform bound for $\epsilon_\lambda(t)$. Take $t \in (0,T]$ and choose $C_1, \cdots, C_4$ according to the Lemmas 3.7 and 3.8, such that $C_5 < 1$, where C_5 is defined by

$$C_5 = C_1\Big(\int_0^T |\dot\sigma(s)|\,ds + \sup_{s\in[0,T]}(|\sigma(s)|)\Big).$$

We use again Lemmas 3.6, 3.7, and 3.8:

$$\begin{aligned}
&\epsilon_\lambda(t) + |V_\lambda(t)|\\
&\leq C_3 + (C_2+1)|V_\lambda(t)| + C_1\psi_\lambda(\epsilon_\lambda(t), V_\lambda(t))\\
&\leq C_3 + C_4 + C_1\int_0^t \phi(\dot V_\lambda(s))\,ds + C_1\psi_\lambda(\epsilon_\lambda(t), V_\lambda(t))\\
&\leq C_3 + C_4\\
&\quad + C_1[\psi(\epsilon_0, V_0) + \langle \sigma(t), \epsilon_\lambda(t)\rangle - \langle \sigma(0), \epsilon_0\rangle - \int_0^t \langle \dot\sigma(s), \epsilon_\lambda(s)\rangle\,ds]\\
&\leq C_3 + C_4 + C_1\psi(\epsilon_0, V_0) + C_1|\sigma(0)|\,|\epsilon_0| + C_5 \sup_{s\in[0,t]}(|\epsilon_\lambda(s)|)\\
&= C_6 + C_5 \sup_{s\in[0,t]}(|\epsilon_\lambda(s)| + |V_\lambda(s)|)
\end{aligned}$$

with some positive constant C_6. Since $C_5 < 1$, we infer

$$|\epsilon_\lambda(t)| + |V_\lambda(t)| \le \frac{C_6}{1 - C_5}.$$

□

3.4. **Limits of approximating solutions.** Our a priori bounds imply that we may extract weakly*-convergent nets. Since we have no theorem on separability, we need to be somewhat careful about using sequences or nets.

Lemma 3.10. *There exists a net $\Lambda \subset (0,1)$, $\Lambda \to 0$, such that B_λ and $\dot V_\lambda$ converge to some $\tilde A \in (\mathbf{L}^\phi)^*([0,T],\mathbb{R}^N)$ and $\tilde W \in (\mathbf{L}^{\phi^*})^*([0,T],\mathbb{R}^N)$, respectively, in the weak * sense, and σ_λ converges to σ almost everywhere.*

Proof. Since $\sigma_\lambda \to \sigma$ in $\mathbf{L}^2([0,T],\mathbb{R}^M)$ as $\lambda \to 0$, there exists a sequence λ_n such that $\sigma_{\lambda_n}(t) \to \sigma(t)$ for all t not contained in a null-set N. From this sequence, we extract a subnet Λ according to Theorem 2.11 such that $\dot V_\lambda$ and B_λ converge in the weak* sense. □

Definition 3.3. Throughout the rest of this paper we consider a net $\Lambda \to 0$ according to Lemma 3.10. Moreover, let $A \in \mathbf{L}^{\phi^*}([0,T],\mathbb{R}^N)$ and $\dot V \in \mathbf{L}^\phi([0,T],\mathbb{R}^N)$ be the function representatives of $\tilde A$ and $\tilde W$, resp. We define $V \in \mathbf{W}^{1,1}([0,T],\mathbb{R}^N)$ by $V(t) = V_0 + \int_0^t \dot V(s)\,ds$.

Lemma 3.11. *V_λ converges uniformly to V along the net Λ.*

Proof. See Lemma 2.13. □

3.5. **Convergence of ϵ_λ.**

Lemma 3.12. *Let ψ satisfy Hypothesis 3.1. Let $v_\lambda, a_\lambda \in \mathbb{R}^N$, $\xi_\lambda, \tau_\lambda \in \mathbb{R}^M$, $\mu_\lambda \in [0,1]$ be nets such that $\tau_\lambda \to \tau$, $v_\lambda \to v$, $\mu_\lambda \to 0$, and $(\tau_\lambda, a_\lambda) \in \partial\psi(\xi_\lambda, v_\lambda - \mu_\lambda a_\lambda)$.*

Then the net ξ_λ converges to some $\xi \in \mathbb{R}^M$, and there is some $a \in \mathbb{R}^N$ such that $(\tau, a) \in \partial\psi(\xi, v)$.

Moreover, ξ is uniquely determined by v, τ, and the property that there exists some a with $(\tau, a) \in \partial\psi(\xi, v)$.

Proof. We start proving the uniqueness of ξ. Suppose $(\tau, a_i) \in \partial\psi(\xi_i, v)$ for $i = 1,2$. Then

$$\psi(\frac{\xi_1+\xi_2}{2}, v) \ge \psi(\xi_1, v) + \langle \tau, \frac{\xi_2-\xi_1}{2}\rangle,$$
$$\psi(\frac{\xi_1+\xi_2}{2}, v) \ge \psi(\xi_2, v) + \langle \tau, \frac{\xi_1-\xi_2}{2}\rangle.$$

Summing up the two inequalities, we obtain

$$\psi(\frac{\xi_1+\xi_2}{2}, v) \ge \frac{1}{2}\psi(\xi_1, v) + \frac{1}{2}\psi(\xi_2, v).$$

By Hypothesis 3.1 (2) we infer that $\xi_1 = \xi_2$.

Now we prove convergence: Let $b_\lambda = \frac{1}{|a_\lambda|} a_\lambda$. Then the net $v_\lambda + b_\lambda$ is bounded, and by Hypothesis 3.1 (3) we infer that there exists a net η_λ such that $\xi_\lambda + \eta_\lambda$ and $\psi(\xi_\lambda + \eta_\lambda, v_\lambda + b_\lambda)$ are bounded. Now

$$\begin{aligned} &\psi(\xi_\lambda + \eta_\lambda, v_\lambda + b_\lambda) - \langle \tau_\lambda, \eta_\lambda\rangle \\ &\ge \psi(\xi_\lambda, v_\lambda - \mu_\lambda a_\lambda) + \langle a_\lambda, b_\lambda + \mu_\lambda a_\lambda\rangle \\ &\ge \psi(\xi_\lambda, v_\lambda - \mu_\lambda a_\lambda) + |a_\lambda|. \end{aligned}$$

This implies that the nets a_λ and $\psi(\xi_\lambda, v_\lambda - \mu_\lambda a_\lambda)$ are bounded. From Hypothesis 3.1 (1) we infer immediately that the net ξ_λ is bounded. By compactness, there are subnets ξ_ν and a_ν converging to some ξ and a, respectively. Since a subdifferential is a closed operator, it is easily seen that $(\tau, a) \in \partial\psi(\xi, v)$. However, this relation determines ξ uniquely in terms of v and τ, so that any subnet of ξ_λ converges to the same ξ. $\square$

Lemma 3.13. *There exists a continuous function $\epsilon \in \mathcal{C}([0,T], \mathbb{R}^M)$ such that $\epsilon(0) = \epsilon_0$, and ϵ_λ converges to ϵ along Λ almost everywhere on $[0,T]$. Moreover, for almost all t there exists some $\hat{A}(t)$ such that $(\sigma(t), \hat{A}(t)) \in \partial\psi(\epsilon(t), V(t))$.*

Proof. We use the fact that $(\sigma_\lambda, A_\lambda) \in \partial\psi(\epsilon_\lambda - \lambda\sigma_\lambda, V_\lambda - \lambda A_\lambda)$. Whenever $\sigma_\lambda(t) \to \sigma(t)$, i.e., almost everywhere, Lemma 3.12 (with $\mu_\lambda = \lambda$) implies that $\epsilon_\lambda(t)$ converges to some $\epsilon(t)$ and the limit satisfies $(\sigma(t), \hat{A}(t)) \in \partial\psi(\epsilon(t), V(t))$ for a suitable $\hat{A}(t)$. Once again, Lemma 3.12 (with $\mu_\lambda = 0$) may be used to show that $\epsilon(t)$ depends continuously on $\sigma(t)$ and $V(t)$, thus ϵ is a continuous function. Since $(\sigma(0), A_0) \in \partial\psi(\epsilon_0, V_0)$, and $\epsilon(0)$ is determined uniquely by this property, we have $\epsilon(0) = \epsilon_0$. $\square$

3.6. **The limiting functions provide a solution.**

Lemma 3.14. *For all $t \in [0,T]$ we have*

$$\liminf_{\lambda \in \Lambda} \psi_\lambda(\epsilon_\lambda(t), V_\lambda(t)) \geq \psi(\epsilon(t), V(t)).$$

Proof. Choose η_λ and W_λ such that

$$\psi_\lambda(\epsilon_\lambda(t), V_\lambda(t)) = \psi(\eta_\lambda(t), W_\lambda(t)) + \frac{1}{2\lambda}[|\eta_\lambda(t) - \epsilon_\lambda(t)|^2 + |W_\lambda(t) - V_\lambda(t)|^2].$$

By Lemma 3.9, $\psi_\lambda(\epsilon_\lambda(t), V_\lambda(t))$ remains bounded as $\lambda \to 0$, so that $|\eta_\lambda(t) - \epsilon_\lambda(t)|^2 + |W_\lambda(t) - V_\lambda(t)|^2 \to 0$. Consequently $\eta_\lambda(t) \to \epsilon(t)$ and $W_\lambda(t) \to V(t)$. Since ψ is lower semicontinuous, we infer that

$$\liminf_{\lambda \to 0} \psi_\lambda(\epsilon_\lambda(t), V_\lambda(t)) \geq \liminf_{\lambda \to 0} \psi(\eta_\lambda(t), W_\lambda(t)) \geq \psi(\epsilon(t), V(t)).$$

$\square$

Lemma 3.15. *For any $\eta \in \mathbf{L}^\infty([0,T], \mathbb{R}^M)$, $W \in \mathbf{L}^\infty([0,T], \mathbb{R}^N)$ we have*

$$\begin{aligned} &\int_0^t \psi(\epsilon(s) + \eta(s), V(s) + W(s))\, ds - \int_0^t \psi(\epsilon(s), V(s))\, ds \\ \geq\ &\int_0^t \langle \sigma(s), \eta(s) \rangle\, ds + \int_0^t \langle A(s), W(s) \rangle\, ds. \end{aligned}$$

In particular, $(\sigma(t), A(t)) \in \partial\psi(\epsilon(t), V(t))$ for almost all t.

Proof.

$$\begin{aligned}
&\int_0^t \psi(\epsilon(s)+\eta(s), V(s)+W(s))\,ds - \int_0^t \psi(\epsilon(s), V(s))\,ds \\
\geq\ &\int_0^t [\psi_\lambda(\epsilon(s)+\eta(s), V(s)+W(s)) - \psi_\lambda(\epsilon_\lambda(s), V_\lambda(s))]\,ds \\
&+ \int_0^t [\psi_\lambda(\epsilon_\lambda(s), V_\lambda(s)) - \psi(\epsilon(s), V(s))]\,ds \\
\geq\ &\int_0^t [\langle \sigma_\lambda(s), \epsilon(s)+\eta(s)-\epsilon_\lambda(s)\rangle + \langle A_\lambda(s), V(s)+W(s)-V_\lambda(s)\rangle]\,ds \\
&+ \int_0^t [\psi_\lambda(\epsilon_\lambda(s), V_\lambda(s)) - \psi(\epsilon(s), V(s))]\,ds.
\end{aligned}$$

Now take limits for $\lambda \to 0$. From Lemma 2.13 we infer that

$$\int_0^t \langle A_\lambda(s), V(s) - V_\lambda(s)\rangle\,dt \to 0.$$

Using Lemma 3.14 we get the desired inequality. Finally Lemma 2.16 implies that $(\sigma(t), A(t)) \in \partial\psi(\epsilon(t), V(t))$ almost everywhere. □

Lemma 3.16. *Let t be a Lebesgue point of the function $\psi(\epsilon(\cdot), V(\cdot))$, i.e.,*

$$\lim_{h\to 0} \frac{1}{h}\int_t^{t+h} \psi(\epsilon(s), V(s))\,ds = \psi(\epsilon(t), V(t)).$$

Then

$$\begin{aligned}
&\psi(\epsilon(t), V(t)) - \psi(\epsilon(0), V(0)) \\
\geq\ &\langle \sigma(t), \epsilon(t)\rangle - \langle \sigma(0), \epsilon(0)\rangle - \int_0^t \langle \dot\sigma(s), \epsilon(s)\rangle\,ds + \int_0^t \langle A(s), \dot V(s)\rangle\,ds.
\end{aligned}$$

Proof. Let $h > 0$. We apply Lemma 3.15 with $\eta(s) = \epsilon(s+h) - \epsilon(s)$ and $W(s) = V(s+h) - V(s)$.

$$\begin{aligned}
&\frac{1}{h}\int_t^{t+h} \psi(\epsilon(s), V(s))\,ds - \frac{1}{h}\int_0^h \psi(\epsilon(s), V(s))\,ds \\
=\ &\frac{1}{h}\int_0^t [\psi(\epsilon(s+h), V(s+h)) - \psi(\epsilon(s), V(s))]\,ds \\
\geq\ &\frac{1}{h}\int_0^t \langle \epsilon(s+h) - \epsilon(s), \sigma(s)\rangle\,ds + \frac{1}{h}\int_0^t \langle V(s+h) - V(s), A(s)\rangle\,ds \\
=\ &\frac{1}{h}\int_t^{t+h} \langle \sigma(s-h), \epsilon(s)\rangle\,ds - \frac{1}{h}\int_0^h \langle \sigma(s), \epsilon(s)\rangle\,ds \\
&+ \int_h^t \langle \frac{1}{h}(\sigma(s-h) - \sigma(s)), \epsilon(s)\rangle\,ds \\
&- \int_0^t \langle \frac{1}{h}(V(s+h) - V(s)), -A(s)\rangle\,ds.
\end{aligned}$$

We take limits for $h \to 0$ and use Lemma 2.15. Notice also that $\psi(\epsilon_0, V_0) \leq \liminf_{s\to 0} \psi(\epsilon(s), V(s))$. Therefore

$$\begin{aligned}&\psi(\epsilon(t), V(t)) - \psi(\epsilon_0, V_0)\\ &\geq \langle \sigma(t), \epsilon(t)\rangle - \langle \sigma(0), \epsilon_0\rangle - \int_0^t \langle \dot{\sigma}(s), \epsilon(s)\rangle\, ds - \int_0^t \langle \dot{V}(s), -A(s)\rangle\, ds.\end{aligned}$$

□

Lemma 3.17. *For almost all $t \in [0, T]$, we have $\dot{V}(t) \in \partial\phi^*(-A(t))$.*

Proof. Integration by parts in Lemma 3.6 yields

$$\begin{aligned}&\psi(\epsilon_\lambda(t) - \lambda\sigma_\lambda(t), V_\lambda(t) - \lambda A_\lambda(t))\\ &\quad + \int_0^t \phi(\dot{V}_\lambda(s))\, ds + \int_0^t \phi^*(-B_\lambda(s))\, ds\\ &\leq \psi(\epsilon_0, V_0) + \langle \epsilon_\lambda(t), \sigma(t)\rangle - \langle \epsilon_0, \sigma(0)\rangle - \int_0^t \langle \dot{\sigma}(s), \epsilon_\lambda(s)\rangle\, ds.\end{aligned}$$

Using the lower semicontinuity of the convex functions involved, we obtain at any point t where $\epsilon_\lambda(t) \to \epsilon(t)$

$$\begin{aligned}&\psi(\epsilon(t), V(t)) + \int_0^t \phi(\dot{V}(s))\, ds + \int_0^t \phi^*(-A(s))\, ds\\ &\leq \psi(\epsilon_0, V_0) + \langle \epsilon(t), \sigma(t)\rangle - \langle \epsilon_0, \sigma(0)\rangle - \int_0^t \langle \dot{\sigma}(s), \epsilon(s)\rangle\, ds.\end{aligned}$$

If t is a Lebesgue point of $\psi(\epsilon(t), V(t))$, we may compare with Lemma 3.16 and obtain

$$\int_0^t \phi(\dot{V}(s))\, ds + \int_0^t \phi^*(-A(s))\, ds \leq -\int_0^t \langle \dot{V}(s), A(s)\rangle\, ds.$$

Since $\phi(\dot{V}(s)) + \phi^*(-A(s)) \geq \langle \dot{V}(s), -A(s)\rangle$ everywhere, we infer that $\phi(\dot{V}(s)) + \phi^*(-A(s)) = \langle \dot{V}(s), -A(s)\rangle$ almost everywhere, which says that $\dot{V}(s) \in \partial\phi^*(-A(s))$ almost everywhere. □

References

[1] H.-D. Alber, *Materials with Memory*, Lecture Notes in Mathematics 1682, Springer, Berlin, Heidelberg, New York 1998. MR **99i:**73040

[2] J.-P. Aubin, *Optima and Equilibria, An Introduction to Nonlinear Analysis*, 2nd ed., Graduate Texts in Mathematics 140, Springer, Berlin, Heidelberg, New York 1998. CMP 2000:06

[3] M. Brokate and P. Krejčí, Wellposedness of kinematic hardening models in elastoplasticity, Christian-Albrechts-Universität Kiel, Berichtsreihe des Mathematischen Seminars Kiel, Bericht 96–4, Februar 1996.

[4] J. R. Ionescu and M. Sofonea, *Functional and Numerical Methods in Viscoplasticity*, Oxford University Press, Oxford, New York, Tokyo 1993. MR **95e:**73025

[5] M. A. Krasnosel'skii and Y. B. Rutickii, *Convex Functions and Orlicz Spaces*, P. Noordhoff, Groningen 1961. MR **23:**A4016

[6] P. Krejčí, *Hysteresis, Convexity, and Dissipation in Hyperbolic Equations*, Gakkotosho, Tokyo, 1996.

[7] P. Laborde and Q. S. Nguyen, Étude de l'équation d'évolution des systémes dissipatifs standards, Mathematical Modelling and Numerical Analysis 24 (1990), 67–84. MR **90m:**58025

[8] J. Lemaitre and J. L. Chaboche, *Mechanics of Solid Materials*, Cambridge University Press, Cambridge 1994.

[9] J. J. Moreau, Application of convex analysis to the treatment of elastoplastic systems, in *Applications of Methods of Functional Analysis to Problems in Mechanics*, P. Germain and B. Nayroles, eds., Lecture Notes in Mathematics 503, Springer, Belin, Heidelberg, New York 1975, pp. 57–89. MR **58:**25196

[10] M. M. Rao and Z. D. Ren, *Theory of Orlicz Spaces*, Pure and Applied Mathematics 146, Marcel Dekker, New York 1991. MR **92e:**46059

[11] R. T. Rockafellar, Integrals which are convex functionals, Pacific J. Math. 24 (1968), 525–539. MR **38:**4984

[12] R. T. Rockafellar, Integrals which are convex functionals II, Pacific J. Math. 39 (1971), 439–469. MR **46:**9710

[13] R. T. Rockafellar and R. J.-B. Wets, *Variational Analysis*, Grundlehren der Mathematischen Wissenschaften 317, Springer, Berlin, Heidelberg, New York 1998. MR **98m:**49001

[14] E. S. Şuhubi, Thermoelastic solids, in *Continuum Physics II: Continuum Mechanics of Single Substance Bodies*, C. Eringen, ed., Academic Press, New York, San Francisco, London 1975.

[15] K. Takamizawa and K. Hayashi, Strain energy density function and uniform strain hypothesis for arterial mechanics, J. Biomechanics 20 (1987), 7–17.

Institut für Mathematik, Universität Graz, Heinrichstrasse 36, A-8010 Graz, Austria

E-mail address: `georg.desch@kfunigraz.ac.at`

Department of Mathematics, Southern Illinois University, Carbondale, Illinois 62901