COMMUNICATIONS ON
PURE AND APPLIED ANALYSIS
Volume **11**, Number **5**, September **2012**

10.3934/cpaa.2012.11.

pp. **2079–2123**

AN ABSTRACT EXISTENCE THEOREM FOR PARABOLIC SYSTEMS

Hans Wilhelm Alt

University Bonn, Institute for Applied Mathematics
Endenicher Allee 60, 53115 Bonn, Germany

(Communicated by Bernd Kawohl)

Abstract. In this paper we prove an abstract existence theorem which can be applied to solve parabolic problems in a wide range of applications. It also applies to parabolic variational inequalities. The abstract theorem is based on a Gelfand triple (V, H, V^*), where the standard realization for parabolic systems of second order is $(W^{1,2}(\Omega), L^2(\Omega), W^{1,2}(\Omega)^*)$. But also realizations to other problems are possible, for example, to fourth order systems. In all applications to boundary value problems the set $M \subset V$ is an affine subspace, whereas for variational inequalities the constraint M is a closed convex set.

The proof is purely abstract and new. The corresponding compactness theorem is based on [5]. The present paper is suitable for lectures, since it relays on the corresponding abstract elliptic theory.

1. Introduction.

In this paper we give an abstract existence proof for parabolic systems. The abstract theorem has been applied to many boundary value problems. Among other things it includes also cases in which the parabolic part is degenerated, therefore it contains elliptic-parabolic problems. It also includes the case of a convex constraint, therefore it contains variational inequalities. The proof for the combination of both effects is new, and has been presented by me in the lecture about partial differential equations in 2003.

There are two main reasons for this approach. One is mathematical, and consists of degenerate parabolic systems occurring in physical applications. The theory in this paper applies for example to boundary value problems of parabolic systems as shown in 11.1. Other applications one finds in [5], [6], [18], [4]. The proof is based on the estimates given in sections 7 and 8. The theory is more general than the parabolic existence theorems in [1], [12], [15], [19].

The other reason lies in theoretical physics and is the entropy principle as formulated in rational thermodynamics, see e.g. [20]. It implies that the estimate, which is the basic estimate of our approach, is equivalent with this entropy inequality. Thus the equations coming from physics are left for mathematical treatment in the original physical setting.

The method of this paper is worthwhile to make some comments. First of all it is a purely abstract formulation of the underlying variational inequality. In this formulation spaces with respect to the space variable are a Hilbert space H for the

2000 *Mathematics Subject Classification.* 35D30, 35K59, 35K65, 35K90, 35R37, 47F05, 47J30, 47J35, 47H05.

Key words and phrases. Parabolic equations, existence theorem, abstract theory.

parabolic term and a Banach space V for the elliptic term, such that (V, H, V^*) is a Gelfand triple, which in a sloppy formulation means that

$$V \hookrightarrow H \hookrightarrow V^*. \tag{1.1}$$

The standard case for $V \hookrightarrow H$ is, that H and V satisfy

$$V \subset H \text{ with a continuous mapping } \mathrm{Id}_V : V \to H. \tag{1.2}$$

The general case can be reduced to this special situation, see 13.2. The relation $H \hookrightarrow V^*$ is explained in 13.1. To generalize this paper to a Banach space H is an interesting question, see section 12.

This paper contains variational inequalities in an abstract setting. Essentially this is achieved by a choice of the set M. Let us explain this choice. If M is an affine subspace, we are dealing, for example, with a standard boundary value problem. That is, if Ω is a bounded Lipschitz domain,

$$\begin{aligned} &V = W^{1,2}(\Omega), \quad \Gamma \subset \partial\Omega \text{ closed},\\ &M = \{u \in V\,;\ u = u_1 \text{ on } \Gamma\}, \end{aligned}$$

we are looking for the solution u of the problem

$$\begin{aligned} &\partial_t u - \mathrm{div} a(u, \nabla u) = f \quad \text{in }]0, T[\times\Omega,\\ &u = u_1 \quad \text{in }]0, T[\times\Gamma,\\ &\nu \bullet a(u, \nabla u) = g \quad \text{in }]0, T[\times(\partial\Omega \setminus \Gamma),\\ &u = u_0 \quad \text{in } \{0\} \times \Omega. \end{aligned}$$

If $\partial_t u$ exists as a function in $L^2([0, T] \times \Omega)$, a weak version of this is

$$\begin{aligned} &u \in L^2([0, T]; M) \text{ and } u = u_0 \text{ in } \{0\} \times \Omega \text{ and}\\ &\int_0^T \int_\Omega (\zeta \cdot \partial_t u + \nabla\zeta \bullet a(u, \nabla u) - \zeta f)\, \mathrm{dL}^n\, \mathrm{dL}^1 - \int_0^T \int_{\partial\Omega} \zeta g\, \mathrm{dH}^{n-1}\, \mathrm{dL}^1 = 0\\ &\text{for } \zeta = u - v \text{ with } v \in L^2([0, T]; M). \end{aligned}$$

A different situation arises for a variational inequality. We assume that the inequality is given by the inequality $u \geq 0$. Then the strong version of the Dirichlet problem, with Dirichlet data $u_1 \geq 0$ and initial data $u_0 \geq 0$, reads

$$\begin{aligned} &u \geq 0 \quad \text{in }]0, T[\times\Omega,\\ &\partial_t u - \mathrm{div} a(u, \nabla u) = f \quad \text{in } (]0, T[\times\Omega) \cap \{u > 0\},\\ &\nu \bullet a(u, \nabla u) = 0 \quad \text{on } (]0, T[\times\Omega) \cap \partial\{u > 0\},\\ &u = u_1 \quad \text{on }]0, T[\times\partial\Omega,\\ &u = u_0 \quad \text{on } \{0\} \times \Omega. \end{aligned}$$

If $a_i(u, \nabla u) = \sum_j a_{ij}(u)\partial_j u$ and an elliptic matrix $(a_{ij})_{ij}$, it is easy to see, that the condition at the free surface $(]0, T[\times\Omega) \cap \partial\{u > 0\}$ is $\partial_\nu u = 0$. Let us write the solution in a weak form. If we define

$$M = \{u \in V\,;\ u \geq 0 \text{ almost everywhere}\},$$

this weak version can be written, if $\partial_t u$ exists as a function, as

$$
\begin{aligned}
&u \in L^2([0,T];M) \text{ and } u = u_0 \text{ in } \{0\} \times \Omega \text{ and} \\
&\int_0^T \int_\Omega ((u-v)\cdot\partial_t u + \nabla(u-v)\bullet a(u,\nabla u) - (u-v)f)\,\mathrm{dL}^n\,\mathrm{dL}^1 \le 0 \\
&\text{for all } v \in L^2([0,T];M).
\end{aligned}
$$

Now, in general one does not know that $\partial_t u$ is a function, it is only defined as a distribution. However, then the term $u\partial_t u$ in the equation is not defined, but this term formally can be written as $\frac{1}{2}\partial_t(u^2)$ and then can be integrated. One obtains the following weak version

$$
\begin{aligned}
&u \in L^2([0,T];M) \text{ and} \\
&\int_\Omega \left(\frac{1}{2}\left(|u(\bar t)|^2 - |u_0|^2\right) - v(\bar t)(u(\bar t) - u_0)\right)\mathrm{dL}^n \\
&+ \int_0^{\bar t}\int_\Omega (\partial_t v(u-u_0) + \nabla(u-v)a(u,\nabla u) - (u-v)f)\,\mathrm{dL}^n\,\mathrm{dL}^1 \le 0 \\
&\text{for almost all } \bar t \in]0,T[, \text{ and this for all } v \in C^\infty([0,T];M).
\end{aligned}
$$

This is of the form of the existence theorem, as we formulate it in 6.2, in the special case that $b(u) = u$.

Therefore in this paper the goal is to prove the general existence theorem 6.2 in an abstract setting for a general closed and convex set $M \subset V$. The usual version for parabolic equations, that is M is a subspace, is a consequence of this general theorem and formulated in 6.3.

There is a large class of problems to which this existence theorem can be applied. Some are presented in section 11. Realistic elliptic-parabolic boundary value problems, which fall under the theorem in this paper, one finds in [11]. We mention, that this paper also contains vector valued versions of such variational inequalities. However, the large class of problems where $u \mapsto b(u)$ has a jump, is not contained in this paper. However, the basic estimates in this paper generalize to such jump nonlinearities, so that the existence theorems can be used as an approximating step.

The constraint in this paper is a time independent set $M \subset V$. It often happens that the more general case of a time dependent constraint occurs. The proof is more involved, therefore not contained in this paper (see the argumentation in [13]).

There are different approaches to parabolic existence theory with a constraint. In particular the approach by [15], where a variational formulation is formulated and the elliptic part is of gradient structure. In this situation one can multiply by $\partial_t u$ and obtains an estimate on the time derivative.

Note: The main part of this paper has been presented during my lecture in 2003. I hope that this theory, which builds on the theory of corresponding stationary problems, can therefore be used in future lectures on functional analysis or on partial differential equations.

2. **Motivation.**

In the following we present some formal observations, which show that the type of system we consider is a consequence of general necessities. Consider a system of partial differential equations

$$\partial_t v_k + \mathrm{div} q_k = \tau_k \quad \text{for } k = 1,\dots,N, \tag{2.1}$$

where, in order to complete the system, we have to specify terms and determine the independent variables. Independent of this we want to derive an estimate. Therefore let us multiply the k-th equation by a function λ_k. Then summing over k we get

$$\sum_k \lambda_k \partial_t v_k + \sum_k \lambda_k \mathrm{div} q_k = \sum_k \lambda_k \tau_k. \tag{2.2}$$

This identity should behave like a parabolic one, that is, integrated over a domain $]t_0, t_1[\times\Omega$ it should be controlled by initial and boundary terms. Thus we require that an identity

$$\sum_k \lambda_k \partial_t v_k = \partial_t w \tag{2.3}$$

holds with a quantity w. If this is true, we obtain

$$\begin{aligned} 0 = &\sum_k \lambda_k \partial_t v_k + \sum_k \lambda_k \mathrm{div} q_k - \sum_k \lambda_k \tau_k \\ = &\ \partial_t w + \mathrm{div}\Big(\sum_k \lambda_k q_k\Big) - \sum_{kj} \partial_j \lambda_k \cdot q_{kj} - \sum_k \lambda_k \tau_k, \end{aligned}$$

where $q_k = (q_{kj})_{j=1,\dots,n}$. Writing $\tau_k = r_k + g_k$ this gives

$$\partial_t w + \mathrm{div}\Big(\sum_k \lambda_k q_k\Big) + \Big(\sum_{kj} \partial_j \lambda_k \cdot (-q_{kj}) + \sum_k \lambda_k \cdot (-r_k)\Big) = \sum_k \lambda_k g_k, \tag{2.4}$$

where usually the q_{kj}-term and the r_k-term are dissipative terms and g_k denotes an external term. Integrating this identity over $]t_0, t_1[\times\Omega$ we obtain

$$\begin{aligned} &\int_\Omega w(t_1, x)\,\mathrm{d}x \\ &+ \int_{t_0}^{t_1} \int_\Omega \Big(\sum_{kj} \partial_j \lambda_k(t,x) \cdot (-q_{kj}(t,x)) \quad + \sum_k \lambda_k(t,x) \cdot (-r_k(t,x))\Big)\,\mathrm{d}x\,\mathrm{d}t \\ &= \int_\Omega w(t_0, x)\,\mathrm{d}x - \int_{t_0}^{t_1} \int_{\partial\Omega} \sum_k \lambda_k(t,x) q_k(t,x) \bullet \nu(x)\,\mathrm{d}\mathrm{H}^{n-1}(x)\,\mathrm{d}t \\ &\quad + \int_{t_0}^{t_1} \int_\Omega \sum_k \lambda_k(t,x) g_k(t,x)\,\mathrm{d}x\,\mathrm{d}t\,. \end{aligned}$$

The first term on the right side contains initial conditions and the second term boundary conditions, whereas the last one is the external term. Therefore the two terms on the left side are the essential ones. The positivity of the second integrand

$$D := \sum_{kj} \partial_j \lambda_k \cdot (-q_{kj}) + \sum_k \lambda_k \cdot (-r_k),$$

if postulated, means that the dissipative term has a sign. Assuming that our N equations are linearly independent we introduce independent variables u_k, $k = 1, \dots, N$, and denote the vector $u = (u_k)_k$. Let us for a moment assume that further $q_k = -\sum_l a_{kl}(u)\nabla u$, and for simplicity $r_k = 0$. Then, if $\lambda_k = \lambda_k(u)$, the dissipative term reads

$$D = \sum_k \nabla \lambda_k \cdot (-q_k) = \sum_{ml} \Big(\sum_k \lambda_k{}'_{u_m}(u) a_{kl}(u)\Big) \nabla u_m \bullet \nabla u_l,$$

which has the correct sign $D \geq 0$, if the matrix

$$\left(\sum_k \lambda_{k\,'u_m}(u) a_{kl}(u) \right)_{ml}$$

is positive semidefinite. We have a freedom to choose the functions u_k, they are not determined by the equations, and properties determined by the equations are independent of the choice of u_k. One particular choice is $\lambda_k = u_k$, in which case (2.3) becomes

$$\sum_k u_k \partial_t v_k = \partial_t w \ . \tag{2.5}$$

Another possibility would be, to choose the functions v_k as independent variable, which would not change the procedure of this paper. However, what we do not assume, is that u_k and v_k are the same, that is a very special case.

Thus what remains is to study the first term on the left side of (2.2), that is (2.5). Let us introduce the vector notation $u = (u_k)_k$ as above and $v = (v_k)_k$. If $w = \varphi(v)$ in (2.5), thus introducing v as independent set of variables, equation (2.5) is equivalent to

$$\sum_k u_k \partial_t v_k = \partial_t \varphi(v) = \sum_k \varphi_{\,'v_k}(v) \partial_t v_k.$$

Therefore, if the derivatives $\partial_t v_k$ are independent from each other, we derive as necessary condition

$$u_k = \varphi_{\,'v_k}(v). \tag{2.6}$$

On the other hand, if $v = \beta(u)$, hence we introduce u as independent set of variables, equation (2.5) is equivalent to, if $w = \Phi(u)$ (it is $\Phi = \varphi \circ \beta$),

$$\sum_k u_k \partial_t \beta_k(u) = \partial_t(\Phi(u)),$$

which one can write as

$$\sum_{kl} u_k \beta_{k\,'u_l}(u) \partial_t u_l = \sum_l \Phi_{\,'u_l}(u) \partial_t u_l.$$

If the derivatives $\partial_t u_l$ are independent from each other, we derive as necessary condition

$$\sum_k u_k \beta_{k\,'u_l}(u) = \Phi_{\,'u_l}(u).$$

By taking the derivative of this equation with respect to u_m one obtains

$$\beta_{m\,'u_l}(u) + \sum_k u_k \beta_{k\,'u_l u_m}(u) = \Phi_{\,'u_l u_m}(u),$$

from which it follows that $(\beta_{m\,'u_l}(u))_{m,l}$ is a symmetric matrix, therefore

$$\beta_{k\,'u_l}(u) = \beta_{l\,'u_k}(u) \quad \text{for all } k, l. \tag{2.7}$$

If this is true for all u, one has for a certain function ψ

$$v_k = \beta_k(u) = \psi_{\,'u_k}(u) \quad \text{for all } k. \tag{2.8}$$

We have seen, that (2.5) implies conditions (2.6) and (2.8), that is

$$v_k = \psi_{\,'u_k}(u) \quad \text{and} \quad u_k = \varphi_{\,'v_k}(v) \quad \text{for all } k, \tag{2.9}$$

which means, that u and v are dual variables to each other. As a consequence

$$\begin{aligned}(\varphi(\beta(u)) + \psi(u))'_{u_l} &= \sum_k \varphi'_{v_k}(\beta(u))\beta_k{}'_{u_l}(u) + \psi'_{u_l}(u) \\ &= \sum_k u_k \beta_k{}'_{u_l}(u) + v_l = \sum_k (u_k\beta_k(u))'_{u_l},\end{aligned}$$

so that

$$\varphi(v) + \psi(u) = \sum_k u_k v_k + \text{const.} \quad \text{for } u \text{ and } v \text{ as in (2.9).} \tag{2.10}$$

3. **Conjugate function.** Let H be a Hilbert space and $\psi : H \to \mathbb{R} \cup \{+\infty\}$ be a mapping, which is not identical to $+\infty$, and which is convex and lower semicontinuous. The dual mapping ψ^*, defined by

$$\psi^*(z^*) := \sup_{z \in H}\Big((\, z\,,\, z^*\,)_H - \psi(z)\Big), \tag{3.1}$$

is a mapping with the same properties, that is, $\psi^* : H \to \mathbb{R} \cup \{+\infty\}$ is not identical $+\infty$, convex and lower semicontinuous. The map ψ^* is called the conjugate convex function of ψ, or Fenchel transformation of ψ. The main inequality reads

$$(\, z\,,\, z^*\,)_H \le \psi(z) + \psi^*(z^*) \quad \text{for all } z, z^* \in H, \tag{3.2}$$

which is Young's inequality. In this inequality, for given z^*, the equality holds for z, if the supremum of $\psi^*(z^*)$ in definition (3.1) is attained for this z. It also follows that $(\psi^*)^* = \psi$. Simple cases are

3.1 Examples. Let $H = \mathbb{R}$.

(1) For $p \in]1, \infty[$ the convex function is $\psi_p(x) := \frac{1}{p}|x|^p$. Then the dual mapping of ψ_p is $(\psi_p)^*(x^*) = \psi_{p^*}(x^*) = \frac{1}{p^*}|x^*|^{p^*}$, if p^* is the dual exponent of p, that is

$$\frac{1}{p} + \frac{1}{p^*} = 1.$$

Then for $\delta > 0$

$$ab = \delta a \cdot \frac{b}{\delta} \le \psi_p(\delta a) + \psi_{p^*}\left(\frac{b}{\delta}\right) = \frac{\delta^p}{p}a^p + \frac{1}{\delta^{p^*}p^*}b^{p^*}$$

is the corresponding Young inequality.

(2) If the convex function ψ is given by

$$\psi(x) := \frac{1}{2}\max(x, 0)^2, \quad \text{then} \quad \psi^*(x^*) = \begin{cases} +\infty \text{ for } x^* < 0, \\ \frac{1}{2}|x^*|^2 \text{ for } x^* \ge 0, \end{cases}$$

is the corresponding conjugate function ψ^*.

A subgradient $z^* \in H$ of ψ in z is defined by

$$\psi(\bar z) \ge \psi(z) + (\, \bar z - z\,,\, z^*\,)_H \quad \text{for all } \bar z \in H, \tag{3.3}$$

and the subdifferential

$$\partial\psi(z) := \{z^* \in H\,;\; z^* \text{ is subgradient of } \psi \text{ in } z\}. \tag{3.4}$$

3.2 Lemma. It is

$$z^* \in \partial\psi(z) \quad \Longleftrightarrow \quad z \in \partial\psi^*(z^*)$$

and one of these statements implies, that $\psi(z) < +\infty$ and $\psi^*(z^*) < +\infty$, and is equivalent to

$$\psi(z) + \psi^*(z^*) = (\, z \,,\, z^* \,)_H \,.$$

Proof. The statement $z^* \in \partial\psi(z)$ implies by (3.4), that (3.3) is satisfied, which implies $\psi(z) < +\infty$, since $\psi(\bar{z})$ is not infinite for all $\bar{z} \in H$, that is, is finite for some $\bar{z} \in H$. Then (3.3) is equivalent to

$$0 \geq \psi(z) + \psi^*(z^*) - (\, z \,,\, z^* \,)_H \,,$$

which is symmetric in (ψ, z) and (ψ^*, z^*). We mention, that this inequality is the inverse Young inequality (3.2) and therefore really must be an equality. □

The Weierstraß function E_ψ is nonnegative, if the function ψ is convex. Then the graph of ψ lies above a plain.

3.3 Weierstraß E-function. For $z_2^* \in \partial\psi(z_2)$ we define

$$\begin{aligned} E_\psi(z_1, z_2, z_2^*) := \quad & \psi(z_1) - \psi(z_2) - (\, z_1 - z_2 \,,\, z_2^* \,)_H \\ & \psi(z_1) - (\psi(z_2) + (\, z_1 - z_2 \,,\, z_2^* \,)_H) \geq 0, \end{aligned}$$

which is nonnegative since z_2^* lies in $\partial\psi(z_2)$.
Note: If ψ is differentiable in z_2 then $(z_1, z_2) \mapsto E_\psi(z_1, z_2, \nabla\psi(z_2))$ is the usual E-function depending on two variables.

Proof. Since ψ is convex, the property $z_2^* \in \partial\psi(z_2)$ implies that E_ψ is nonnegative. □

For $z_0 \in H$ the translated function is

$$\psi_{z_0}(z) := \psi(z_0 + z). \tag{3.5}$$

Then the subdifferential $\partial\psi_{z_0}(z) = \partial\psi(z_0 + z)$ is just a shift and for the conjugate convex function $(\psi_{z_0})^*(z^*) = \psi^*(z^*) - (\, z_0 \,,\, z^* \,)_H$. For a convex C^1-function one obtains

3.4 Lemma. Let $\psi : H \to \mathbb{R} \cup \{+\infty\}$ as above and $D \subset H$ be open and convex. Assume that ψ is finite and continuously differentiable on D and $\psi = \infty$ in $H \setminus D$. Then

(1) $b := \nabla\psi : D \to H$ is monotone (increasing) in D.

(2) For $z \in D$ the subdifferential is $\partial\psi(z) = \{b(z)\}$.

(3) For $z \in D$ and $z^* = b(z)$ it is

$$(\, z \,,\, z^* \,)_H = \psi^*(z^*) + \psi(z).$$

With this the Weierstraß E-function is defined for $(z_1, z_2) \in D \times D$ by

$$\begin{aligned} (z_1, z_2) \mapsto E_\psi(z_1, z_2, \nabla\psi(z_2)) &= \psi(z_1) - \psi(z_2) - (\, z_1 - z_2 \,,\, \nabla\psi(z_2) \,)_H \\ &= \psi(z_1) - \psi(z_2) - (\, z_1 - z_2 \,,\, b(z_2) \,)_H \,. \end{aligned} \tag{3.6}$$

3.5 Lemma. Let ψ and D as in 3.4. Then

$$E_{\psi^*}(z_1^*, z_2^*, z_2) = E_\psi(z_2, z_1, z_1^*)$$

for $z_1, z_2 \in D$ and $z_1^* = b(z_1)$, $z_2^* = b(z_2)$.

Proof. By 3.4(3)

$$\begin{aligned} &E_{\psi^*}(z_1^*, z_2^*, z_2) = \psi^*(z_1^*) - \psi^*(z_2^*) - (\, z_1^* - z_2^* \,,\, z_2 \,)_H \\ &= (-\psi(z_1) + (\, z_1 \,,\, z_1^* \,)_H) + (\psi(z_2) - (\, z_2 \,,\, z_2^* \,)_H) - (\, z_1^* - z_2^* \,,\, z_2 \,)_H \\ &= -\psi(z_1) + \psi(z_2) + (\, z_1 - z_2 \,,\, z_1^* \,)_H = E_\psi(z_2, z_1, z_1^*), \end{aligned}$$

if we take the definition. □

Concerning b we will need for the next sections the following definition, which essentially is the conjugate function $z^* \mapsto \psi_{z_0}^*(z^*)$ for $z^* = b(z)$.

3.6 Definition. Let ψ and D as in 3.4 and b as in 3.4(1). For $z_0 \in D$ define a function $B_{z_0} : D \to \mathbb{R}$ by

$$\begin{aligned} B_{z_0}(z) :=\ & (\psi_{z_0})^*(b(z)) + \psi_{z_0}(0) \ \Big(= E_\psi(z_0, z, b(z))\Big) \\ =\ & (\, z - z_0 \,,\, b(z) \,)_H - \psi(z) + \psi(z_0) \\ =\ & \int_0^1 (\, z - z_0 \,,\, b(z) - b((1-s)z_0 + sz) \,)_H \, \mathrm{d}s. \end{aligned}$$

It is B_{z_0} nonnegative, that is $B_{z_0} \geq 0$.

Proof. Using 3.4(3) and 3.4(1) one obtains the identities for $B_{z_0}(z)$. By the convexity of ψ, or the monotonicity of b, or the nonnegativity of E_ψ, the terms in the definition are nonnegative. □

Then

3.7 Lemma. For $z_0, z_1, z_2 \in D$

$$\begin{aligned} (\, z_1 - z_0 \,,\, b(z_1) - b(z_2) \,)_H =\ & B_{z_0}(z_1) - B_{z_0}(z_2) + E_\psi(z_1, z_2, \nabla\psi(z_2)) \\ \geq\ & B_{z_0}(z_1) - B_{z_0}(z_2) \ . \end{aligned}$$

Proof. By the previous definition 3.6

$$\begin{aligned} & B_{z_0}(z_1) - B_{z_0}(z_2) \\ =\ & (\, z_1 - z_0 \,,\, b(z_1) \,)_H - \psi(z_1) - (\, z_2 - z_0 \,,\, b(z_2) \,)_H + \psi(z_2) \\ =\ & (\, z_1 - z_0 \,,\, b(z_1) - b(z_2) \,)_H + (\, z_1 - z_2 \,,\, b(z_2) \,)_H - \psi(z_1) + \psi(z_2) \\ =\ & (\, z_1 - z_0 \,,\, b(z_1) - b(z_2) \,)_H - E_\psi(z_1, z_2, \nabla\psi(z_2)) \\ \leq\ & (\, z_1 - z_0 \,,\, b(z_1) - b(z_2) \,)_H \end{aligned}$$

using the formula for E_ψ in (3.6). □

Whereas $\psi_{z_0}^*$ for the example in 3.1 with $p < \infty$ grows at infinity of order p^*, in general one has only the following lemma.

3.8 Superlinearity of $\psi_{z_0}^*$. Let $\psi : H \to H$ be convex and lower semicontinuous, and bounded on bounded subsets of H. Then for $\delta > 0$ and $z_0 \in H$ there exists a constant C_{δ, z_0} so that

$$\| z^* \|_H \leq \delta \psi_{z_0}^*(z^*) + C_{\delta, z_0} \quad \text{for all } z^* \in H.$$

Proof. We compute

$$\begin{aligned}
(\psi_{z_0})^*(z^*) &\geq \sup_{z\in\partial \mathrm{B}_{\frac{1}{\delta}}(0)\subset H}\Big(\langle z\,,\,z^*\rangle_H - \psi_{z_0}(z)\Big)\\
&\geq \sup_{z\in\partial \mathrm{B}_{\frac{1}{\delta}}(0)}\langle z\,,\,z^*\rangle_H - \sup_{z\in\partial \mathrm{B}_{\frac{1}{\delta}}(0)}\psi_{z_0}(z)\\
&= \frac{1}{\delta}\|z^*\|_H - \sup_{z\in\partial \mathrm{B}_{\frac{1}{\delta}}(0)}\psi_{z_0}(z).
\end{aligned}$$

□

This implies, if ψ and $D = H$ as in 3.4, and b as in 3.4(1) and B as in 3.6, that there is a constant $\widetilde{C}_{\delta,z_0} := C_{\delta,z_0} - \delta\psi_{z_0}(0)$ with

$$\|b(z)\|_H \leq \delta B_{z_0}(z) + \widetilde{C}_{\delta,z_0} \quad \text{for all } z\in H \text{ and } \delta > 0. \tag{3.7}$$

4. **Elliptic theorem.** The parabolic existence proof is based on the following elliptic theorem, which is formulated on a closed, convex set M of a Banach space V,

$$M\subset V \text{ nonempty, closed, and convex.} \tag{4.1}$$

Further a map

$$F : M \to V^* \tag{4.2}$$

is given, where V^* is the dual space of V. By $(w, w^*) \mapsto \langle w\,,\,w^*\rangle_V := w^*(w)$ for $w \in V$ and $w^* \in V^*$ we denote the dual product of V. The main assumption for the elliptic existence theorem 4.2 is the

4.1 Continuity condition. The following holds: Let $u_m, u \in V$ and $v^* \in V^*$ with

$$\left\{\begin{array}{c} u_m, u\in M \text{ and } u_m\to u \text{ weakly in } V \text{ for } m\to\infty,\\ F(u_m)\to v^* \text{ weakly}^* \text{ in } V^* \text{ for } m\to\infty, \text{ and}\\ \limsup_{m\to\infty}\langle u_m\,,\,F(u_m)\rangle_V \leq \langle u\,,\,v^*\rangle_V\,, \end{array}\right\}$$

then

$$\left\{\begin{array}{c} \langle u-v\,,\,F(u)-v^*\rangle_V \leq 0 \text{ for all } v\in M, \text{ and}\\ \limsup_{m\to\infty}\langle u_m\,,\,F(u_m)\rangle_V = \langle u\,,\,v^*\rangle_V\,. \end{array}\right\}$$

With this the following theorem is satisfied.

4.2 Theorem. Let V be a separable reflexive Banach space and $M \subset V$ as in (4.1), and let $F : M \to V^*$ with the following properties:

(1) ***Boundedness.*** The map F is bounded on bounded subsets of M.

(2) ***Continuity property.*** The map F satisfies condition 4.1.

(3) ***Coercivity.*** For some $\bar{u}\in M$

$$\frac{\langle u-\bar{u}\,,\,F(u)\rangle_V}{\|u-\bar{u}\|_V}\to\infty \quad \text{for } u\in M,\ \|u-\bar{u}\|_V\to\infty.$$

Under these assumptions there exists $u \in M$, so that

$$\langle u-v\,,\,F(u)\rangle_V \leq 0 \quad \text{for all } v\in M. \tag{4.3}$$

Note: The condition, that $\bar{u}$ belongs to M, in general is necessary for the theorem.

We call (4.3) the variational inequality for F with respect to M.

Proof. A proof can be found in [17, Chap. 3], [2], [22, Kap. 3]. A version of this theorem can be found in [8, Theorem (5.2.3)]. □

There are many examples, which fall under this theorem, the standard ones are monotone operators and compact perturbations of monotone operators. The condition 4.2(2), that is 4.1, is usually connected with the name "pseudomonotone operators", although the definition of pseudomonotone is a little bit different, but it is equivalent under the complete assumptions of 4.2.

5. **Time discrete problem.** We consider a Hilbert space H and a Banach space V as in the introduction, that is (1.2),

$$V \subset H, \quad \mathrm{Id}_V : V \to H \text{ continuous,} \tag{5.1}$$

is satisfied. For more general V we refer to 13.2.

The approximative problem is given for discrete times t^i with $t^i < t^{i+1}$. For simplicity we consider the case of a constant time step $h > 0$, that is,

$$t^i = ih \quad \text{ for } i \in \mathbb{N}. \tag{5.2}$$

The constraint is approximated by

$$M^i \subset V \quad \text{ nonempty closed convex,} \tag{5.3}$$

where here the set M^i may change in time. On M^i an "elliptic" operator is given by

$$A^i : M^i \to V^*, \tag{5.4}$$

where this map is defined recursively, that is, it may depend on the solution for smaller i. The "parabolic" part is given by a map

$$b = \nabla\psi \text{ with } \psi : H \to \mathbb{R} \text{ convex and continuously differentiable.} \tag{5.5}$$

We approximate the parabolic problem by a time discrete version, that is we replace the time derivative of $b(u)$ by time differences

$$t \mapsto \frac{1}{h}(b(u(t)) - b(u(t-h))).$$

Under assumptions (5.1)–(5.5) the problem is to find inductively in i a solution with given starting value $u_0 \in H$.

5.1 Time discrete problem. With $u^0 = u_0 \in H$ find inductively for $i \geq 1$ elements u^i with

$$\begin{gathered} u^i \in M^i \text{ and} \\ \left(u^i - v \,,\, \tfrac{1}{h}(b(u^i) - b(u^{i-1})) \right)_H + \left\langle u^i - v \,,\, A^i(u^i) \right\rangle_V \leq 0 \\ \text{for all } v \in M^i. \end{gathered} \tag{5.6}$$

Here u^{i-1}, for $i = 1$, is the initial value $u^0 := u_0 \in H$, and for $i \geq 2$, is the known vector from previous time step. Thus the solution $u^i \in M^i$ is constructed inductively in i, therefore the operator A^i may contain also information from the previous time steps, e.g. it may depend on u^{i-1}.

For the existence proof we deal with certain assumptions about the operator

$$u \mapsto F^i(u) = \lambda \mathrm{Id}_V^*(\mathrm{D}\psi(u) - \mathrm{D}\psi(u^{i-1})) + A^i(u), \tag{5.7}$$

among them the continuity condition 4.1. We will show that the "parabolic" part $b = \nabla\psi$ by assumption gives the necessary property, provided the embedding from V into H is compact. As consequence only the map $A^i : M^i \to V^*$ has to satisfy 4.1. The map F^i satisfies the following

5.2 Remark. It holds for $u \in M^i$ and $v \in V$

$$\begin{aligned}&\big\langle\, v \,,\, \lambda \mathrm{Id}_V^*(\mathrm{D}\psi(u) - \mathrm{D}\psi(u^{i-1})) + A^i(u) \,\big\rangle_V \\ &= \lambda \big(\, v \,,\, b(u) - b(u^{i-1}) \,\big)_H + \big\langle\, v \,,\, A^i(u) \,\big\rangle_V \ .\end{aligned}$$

Proof. $(\, v \,,\, b(u) \,)_H = (\, v \,,\, \nabla\psi(u) \,)_H = \mathrm{D}\psi(u)(v) = \langle\, v \,,\, \mathrm{Id}_V^* \mathrm{D}\psi(u) \,\rangle_V$. □

5.3 Lemma. Let the inclusion $\mathrm{Id}_V : V \to H$ be compact. If the map $A^i : M^i \to V^*$ satisfies 4.1, then for all $\lambda \in \mathbb{R}$ the map $\lambda \mathrm{Id}_V^*(\mathrm{D}\psi - \mathrm{D}\psi(u^{i-1})) + A^i$ satisfies condition 4.1 on M^i.

Proof. We assume that 4.1 is satisfied for A^i, and we have to show that condition 4.1 for the map F^i in (5.7) is true. Therefore let $u_m, u \in M^i$ with $u_m \to u$ weakly in V and $F^i(u_m) \to v^*$ weakly* in V^* for $m \to \infty$, and

$$\limsup_{m\to\infty} \big\langle\, u_m \,,\, F^i(u_m) \,\big\rangle_V \le \langle\, u \,,\, v^* \,\rangle_V \ .$$

Since $\mathrm{Id}_V : V \to H$ is compact, hence completely continuous, it follows that $u_m \to u$ strongly in H for $m \to \infty$. Since b is continuous, $b(u_m) \to b(u)$ strongly in H, and since for $v \in V$

$$\langle\, v \,,\, \mathrm{Id}_V^* \mathrm{D}\psi(u_m) \,\rangle_V = \langle\, v \,,\, \mathrm{D}\psi(u_m) \,\rangle_H = (\, v \,,\, b(u_m) \,)_H \,,$$

it follows that $\mathrm{Id}_V^* \mathrm{D}\psi(u_m) \to \mathrm{Id}_V^* \mathrm{D}\psi(u)$ strongly in V^* as $m \to \infty$. Defining

$$\widetilde{v}^* := v^* - \lambda \mathrm{Id}_V^*(\mathrm{D}\psi(u) - \mathrm{D}\psi(u^{i-1}))$$

we see that the properties of F^i imply

$$\begin{aligned}&A^i(u_m) \to \widetilde{v}^* \text{ weakly}^* \text{ in } V^* \text{ for } m \to \infty \ , \\ &\limsup_{m\to\infty} \big\langle\, u_m \,,\, A^i(u_m) \,\big\rangle_V \le \langle\, u \,,\, \widetilde{v}^* \,\rangle_V \ .\end{aligned}$$

Since it is assumed that 4.1 for the map A^i is satisfied, we conclude

$$\begin{aligned}&\big\langle\, u - v \,,\, A^i(u) - \widetilde{v}^* \,\big\rangle_V \le 0 \text{ for all } v \in M^i, \\ &\limsup_{m\to\infty} \big\langle\, u_m \,,\, A^i(u_m) \,\big\rangle_V = \langle\, u \,,\, \widetilde{v}^* \,\rangle_V \ .\end{aligned}$$

Inserting the definition of $\widetilde{v}^*$ one gets

$$\begin{aligned}&\big\langle\, u - v \,,\, F^i(u) - v^* \,\big\rangle_V \le 0 \text{ for all } v \in M^i, \\ &\limsup_{m\to\infty} \big\langle\, u_m \,,\, F^i(u_m) \,\big\rangle_V = \langle\, u \,,\, v^* \,\rangle_V \ .\end{aligned}$$

Therefore it has been shown that condition 4.1 for the map F^i is fulfilled. □

Similar one can show, that property 4.1 for F^i implies this property for A^i. It is also enough to assume the boundedness condition for A^i. This is because the following lemma holds.

5.4 Lemma. Let $\mathrm{Id}_V : V \to H$ be compact. If b is defined as in (5.5), then b is bounded on bounded subsets of V.

Proof. Assume this is not true. Then there is an $R > 0$ and $u_m \in V$, $m \in \mathbb{N}$, with $\|u_m\|_V \leq R$ and $\|b(u_m)\|_H \to \infty$ in H as $m \to \infty$. Since $\mathrm{Id}_V : V \to H$ is compact, hence completely continuous, u_m has in H a convergent subsequence, that is there is $u \in H$ with $u_m \to u$ in H for a subsequence $m \to \infty$. Since b is continuous, we conclude $b(u_m) \to b(u)$ for the subsequence, a contradiction. □

We can now formulate the theorem for time discrete solutions.

5.5 Theorem. Let (5.1)–(5.5) be satisfied and $\mathrm{Id}_V : V \to H$ be compact. Further assume

(1) ***Boundedness.*** The map A^i is bounded on bounded subsets of M^i.

(2) ***Continuity condition.*** The map $A^i : M^i \to V^*$ satisfies 4.1.

(3) ***Coercivity.*** There exists a $\lambda > 0$ independent of i, and a $\bar{u}^i \in M^i$, such that

$$\frac{\lambda\left(u - \bar{u}^i\, , b(u)\right)_H + \left\langle u - \bar{u}^i\, , A^i(u)\right\rangle_V}{\|u - \bar{u}^i\|_V} \to \infty \quad \text{for } u \in M^i \text{ with } \left\|u - \bar{u}^i\right\|_V \to \infty. \tag{5.8}$$

Under these assumptions it follows, that for given $u_0 \in H$ and for $h \leq \frac{1}{\lambda}$ the time discrete problem 5.1 has a solution.

Proof. Let u^0 be as in 5.1 and $i \geq 1$. Because of remark 5.2 the problem in 5.1 can be formulated as

$$\left\langle u - v\, , F_h^i(u)\right\rangle_V \leq 0 \quad \text{for all } v \in M^i \tag{5.9}$$

with $u \in M^i$ and with $F_h^i : M^i \to V^*$ given by

$$F_h^i(u) := \frac{1}{h}\mathrm{Id}_V^*(\mathrm{D}\psi(u) - \mathrm{D}\psi(u^{i-1})) + A^i(u).$$

Property (1) implies 4.2(1) for F_h^i by using 5.4. Since (2) is satisfied, the statement 5.3 shows that 4.2(2) is valid for the map F_h^i. Since ψ is convex, the first term on the right side of

$$u \mapsto F_h^i(u) = \left(\frac{1}{h} - \lambda\right)\mathrm{Id}_V^*(\mathrm{D}\psi(u) - \mathrm{D}\psi(u^{i-1})) + F_{\frac{1}{\lambda}}(u)$$

is monotone in u, if $h \leq \frac{1}{\lambda}$, where $\lambda > 0$ is the number for which (5.8) holds. Hence the coercivity 4.2(3) for F_h^i is satisfied, since (5.8) is satisfied for $F_{\frac{1}{\lambda}}$. Consequently there is a solution of the variational inequality (5.9). □

Alternatively, the time discrete solution in 5.1 can be formulated as in 5.6. For this we construct for each sequence $u^j \in H$ with $j \in \mathbb{N} \cup \{0\}$ a step function in time by

$$\begin{aligned} u_h(t) &:= \begin{cases} u^i & \text{for } (i-1)h < t \leq ih,\ i \in \mathbb{N}, \\ u^0 & \text{for } t \leq 0, \end{cases} \\ M_h(t) &:= M^i \quad \text{for } (i-1)h < t \leq ih,\ i \in \mathbb{N}. \end{aligned} \tag{5.10}$$

Similar one defines the elliptic operator

$$A_h(t,u) := \begin{cases} A^i(u) & \text{for } (i-1)h < t \le ih,\ i \in \mathbb{N}, \\ 0 & \text{for } t \le 0. \end{cases} \tag{5.11}$$

With these definitions equation 5.1 becomes

5.6 Time discrete problem. Find a step function u_h with

$$\begin{aligned} &u_h(t) = u_0 \text{ for } t < 0,\ u_h(t) \in M_h(t) \text{ for } t > 0, \\ &\big(\, u_h(t) - v\,,\, \tfrac{1}{h}(b(u_h(t)) - b(u_h(t-h)))\,\big)_H \\ &+ \langle\, u_h(t) - v\,,\, A_h(t, u_h(t))\,\rangle_V \le 0 \quad \text{for } t > 0 \text{ and for } v \in M_h(t). \end{aligned} \tag{5.12}$$

By a step function we mean a function as in (5.10).

6. **The main theorem.** In the following we describe the main theorem of this paper. Given a Hilbert space H and a Banach space V such that (V, H, V^*) is a Gelfand triple satisfying $V \hookrightarrow H \hookrightarrow V^*$. As pointed out in section 13, we can work with a special case and can assume that

$$V \subset H \text{ with a continuous mapping } \mathrm{Id}_V : V \to H. \tag{6.1}$$

Besides these spaces we consider a set

$$\begin{aligned} \mathcal{M} = \{u \in L^p([0,T];V)\,; \quad & B_0(u) \in L^\infty([0,T]), \\ & u(t) \in M \text{ for almost all } t\} \end{aligned} \tag{6.2}$$

with a time independent constraint

$$M \subset V \text{ nonempty, closed, and convex.} \tag{6.3}$$

We assume that the "parabolic part" of our problem is given by a map

$$\begin{aligned} b : H \to H \text{ monotone and continuous, in fact} \\ b = \nabla\psi, \quad \psi : H \to \mathbb{R} \text{ convex and continuously differentiable,} \end{aligned} \tag{6.4}$$

where the functional B_0 in the above definition is given by

$$B_0(z) = \psi^*(b(z)) + \psi(0),$$

see the definition in 3.6. On $\mathcal{M}$ we denote the "elliptic part" of the problem by

$$\begin{aligned} \mathcal{A} : \mathcal{M} \to L^{p^*}([0,T];V^*), \\ \mathcal{A}(u)(t) = A(t, u(t)) \quad \text{with} \quad A(t, \bullet) : M \to V^*. \end{aligned} \tag{6.5}$$

We shall present some illustrating examples in section 11, in particular 11.1.

According to the continuity condition in 4.1 we assume for the parabolic problem in this section, that a time version of this condition is satisfied.

6.1 Continuity condition. Let $u_m, u \in L^p([0,T];V)$ and $v^* \in L^{p^*}([0,T];V^*)$ with

$$\left\{\begin{array}{c} u_m(t), u(t) \in M \text{ for almost all } t \in]0,T[\text{ and} \\ u_m \to u \text{ weakly in } L^p([0,T];V) \text{ for } m \to \infty, \\ \{B(u_m)\,;\ m \in \mathbb{N}\} \text{ bounded in } L^\infty([0,T];\mathbb{R}) \text{ and} \\ b(u_m) \to b(u) \text{ strongly in } L^1([0,T];H) \text{ for } m \to \infty, \\ \mathcal{A}(u_m) \to v^* \text{ weakly in } L^{p^*}([0,T];V^*) \text{ for } m \to \infty, \text{ and} \\ \limsup_{m\to\infty} \int_0^T \langle\, u_m(t)\,,\, \mathcal{A}(u_m)(t)\,\rangle_V \,\mathrm{d}t \le \int_0^T \langle\, u(t)\,,\, v^*(t)\,\rangle_V \,\mathrm{d}t, \end{array}\right\}$$

then

$$\left\{\begin{array}{c} \int_0^T \langle\, u(t) - v(t)\,,\, \mathcal{A}(u)(t) - v^*(t)\,\rangle_V \,\mathrm{d}t \le 0 \\ \text{for all } v \in L^p([0,T];V) \text{ with } v(t) \in M \text{ for almost all } t\text{ , and} \\ \limsup_{m\to\infty} \int_0^T \langle\, u_m(t)\,,\, \mathcal{A}(u_m)(t)\,\rangle_V \,\mathrm{d}t = \int_0^T \langle\, u(t)\,,\, v^*(t)\,\rangle_V \,\mathrm{d}t. \end{array}\right\}$$

We mention, that by (3.7) the condition that $B(u_m) \in L^\infty([0,T])$ are bounded implies also that $b(u_m) \in L^\infty([0,T];H)$ are bounded.

With this assumption we can prove the following existence theorem, where the structure of the theorem is the same as in 5.5. We mention, that in concrete cases the proof that $\mathcal{A}$ maps into $L^{p^*}([0,T];V^*)$ usually immediately gives the boundedness condition 6.2(1).

6.2 Existence theorem. Let H be a Hilbert space and V a separable reflexive Banach space as in (6.1) and with a compact embedding $\mathrm{Id}_V : V \to H$. and let M as in (6.3). Moreover, let $\mathcal{A} : \mathcal{M} \to L^{p^*}([0,T];V^*)$ as in (6.5), and let $b : H \to H$ as in (6.4), with the following properties:

(1) ***Boundedness.*** $\mathcal{A}$ maps sets in $\mathcal{M}$, which are bounded in the $L^\infty([0,T])$-norm of B_0 and bounded in the $L^p([0,T];V)$-norm, into bounded sets of $L^{p^*}([0,T];V^*)$.

(2) ***Continuity condition.*** $\mathcal{A}$ satisfies the condition 6.1.

(3) ***Coercivity.*** For almost all $t \in]0,T[$

$$\langle\, u - \bar{u}\,,\, A(t,u)\,\rangle_V \ge c_0 \|u - \bar{u}\|_V^p - C_0 B_{\bar{u}}(u) - G_0(t)$$

for all $u \in M$. Here $\bar{u} \in M$ and $G_0 \in L^1([0,T])$, and $c_0 > 0$ and C_0 are constants.

Then there exist solutions of the "evolution problem", that is for given $u_0 \in \mathrm{clos}_H(M)$ there is a $u \in L^p([0,T];V)$ with

$$u(t) \in M \quad \text{for almost all } t,$$

$$\left\{\begin{array}{l} B_{\bar{u}}(u(\bar{t})) - B_{\bar{u}}(u_0) + (\,\bar{u} - v(\bar{t})\,,\, b(u(\bar{t})) - b(u_0)\,)_H \\ \quad + \int_0^{\bar{t}} \Big(-(\,\partial_t(\bar{u} - v)(t)\,,\, b(u(t)) - b(u_0)\,)_H \\ \qquad\qquad + \langle\, u(t) - v(t)\,,\, A(t,u(t))\,\rangle_V \Big) \,\mathrm{d}t \le 0 \\ \text{for almost all } \bar{t} \in]0,T[, \end{array}\right\} \tag{6.6}$$

and this for all $v \in C^\infty([0,T];V)$ with $v(t) \in M$ for almost all t.

Proof. The parabolic terms in the solution property in (6.6) we denote as

$$\begin{aligned}\Phi_{\bar u}(u,v)(\bar t) := \ & B_{\bar u}(u(\bar t)) - B_{\bar u}(u_0) + (\, \bar u - v(\bar t)\, ,\, b(u(\bar t)) - b(u_0)\,)_H \\ & - \int_0^{\bar t} (\, \partial_t(\bar u - v(t))\, ,\, b(u(t)) - b(u_0)\,)_H \, \mathrm{d}t,\end{aligned} \tag{6.7}$$

see definition 8.1. It is used in the proof the theorem in section 10. □

In the special case that the set $M \subset V$ is an affine space, that is the constraints are defined by equations only, the existence theorem has the following form.

6.3 Existence theorem. Let H be a Hilbert space and V a separable reflexive Banach space as in (6.1) and with a compact embedding $\mathrm{Id}_V : V \to H$. Further, let $M \subset V$ be a nonempty closed affine set, $\mathcal{M}$ and $\mathcal{A}$ as in (6.3) and (6.5), and $b : H \to H$ as in (6.4), with the assumptions 6.2(1)–6.2(3).

Then there exist solutions of the "evolution equation", that is if $u_0 \in \mathrm{clos}_H(M)$ there is a $u \in L^p([0,T];V)$ with

$$\begin{aligned}& u(t) \in M \quad \text{for almost all } t, \\ & \left\{ \begin{aligned} \int_0^T \Big(-(\, \partial_t \xi(t)\, ,\, b(u(t)) - b(u_0)\,)_H & \\ + \langle\, \xi(t)\, ,\, A(t,u(t))\, \rangle_V \Big)\, \mathrm{d}t = 0 & \end{aligned} \right\} \\ & \text{for all } \xi \in C_0^\infty([0,T[;V) \text{ with } \xi(t) \in M_1 \text{ for all } t.\end{aligned} \tag{6.8}$$

Here $M_1 \subset V$ is the subspace, such that $M = u_1 + M_1$ for every $u_1 \in M$.

The proof of this statement uses the general existence theorem.

Proof. In 6.2 we have proved the inequality, using the notation in (6.7),

$$\Phi_{\bar u}(u,v)(\bar t) + \int_0^{\bar t} \langle\, u(t) - v(t)\, ,\, A(t,u(t))\, \rangle_V \, \mathrm{d}t \le 0$$

for all $v \in C^\infty([0,T];V)$ which satisfies $v(t) \in M$ for all t. Here M now is an affine subspace contained in V. It follows that this inequality then also holds for all

$$v \in W^{1,p}(]0,T[;M) \subset W^{1,1}(]0,T[;H) \cap L^p(]0,T[;M).$$

Now $u_0 \in \mathrm{clos}_H(M)$, hence there are $u_{0\varepsilon} \in M$ so that $u_{0\varepsilon} \to u_0$ as $\varepsilon \to 0$ in H. Then define $u_{\delta\varepsilon} \in W^{1,p}(]0,T[;M)$ as in 8.3 and let

$$v := u_{\delta\varepsilon} - \xi \in W^{1,p}(]0,T[;M)$$

with $\xi \in W^{1,\infty}(]0,T[;V)$ and $\xi(t) \in M_1$. Since $t \mapsto A(t,u(t))$ is in $L^{p^*}(]0,T[;V^*)$ and since $u_{\delta\varepsilon} \to u$ as $\delta \to 0$ in $L^p(]0,T[;V)$, we get

$$\begin{aligned}&\int_0^{\bar t} \langle\, u(t) - v(t)\, ,\, A(t,u(t))\, \rangle_V \, \mathrm{d}t = \int_0^{\bar t} \langle\, u(t) - u_{\delta\varepsilon}(t) + \xi(t)\, ,\, A(t,u(t))\, \rangle_V \, \mathrm{d}t \\ &\longrightarrow \int_0^{\bar t} \langle\, \xi(t)\, ,\, A(t,u(t))\, \rangle_V \, \mathrm{d}t.\end{aligned}$$

We also get

$$\begin{aligned}\Phi_{\bar u}(u,v)(\bar t) =\ & \Phi_{\bar u}(u,u_{\delta\varepsilon}-\xi)(\bar t)\\ =\ & B_{\bar u}(u(\bar t)) - B_{\bar u}(u_0)\\ & + (\,\bar u - u_{\delta\varepsilon}(\bar t) + \xi(\bar t)\,,\, b(u(\bar t)) - b(u_0)\,)_H\\ & - \int_0^{\bar t} (\,\partial_t(\bar u - u_{\delta\varepsilon} + \xi)(t)\,,\, b(u(t)) - b(u_0)\,)_H \,\mathrm{d}t\\ =\ & \Phi_{\bar u}(u,u_{\delta\varepsilon})(\bar t)\\ & + (\,\xi(\bar t)\,,\, b(u(\bar t)) - b(u_0)\,)_H\\ & - \int_0^{\bar t} (\,\partial_t\xi(t)\,,\, b(u(t)) - b(u_0)\,)_H \,\mathrm{d}t.\end{aligned}$$

Since $\liminf_{\varepsilon\to 0}\liminf_{\delta\to 0}\Phi_{\bar u}(u,u_{\delta\varepsilon})(\bar t) \geq 0$ for almost all $\bar t > 0$ we obtain

$$\begin{aligned}&(\,\xi(\bar t)\,,\, b(u(\bar t)) - b(u_0)\,)_H\\ &- \int_0^{\bar t} (\,\partial_t\xi(t)\,,\, b(u(t)) - b(u_0)\,)_H \,\mathrm{d}t + \int_0^{\bar t} \langle\,\xi(t)\,,\, A(t,u(t))\,\rangle_V \,\mathrm{d}t \leq 0.\end{aligned}$$

This is obviously equivalent to the assertion. The left side is a linear form in ξ. Now we can replace ξ by $-\xi$ to obtain that the left side equals zero. If we now restrict $\xi \in C_0^\infty([0,T[;M_1)$, that is with compact support in $[0,T[$, then ξ vanishes in a neighbourhood of T, and therefore one chooses $\bar t$ close to T in order to get

$$\int_0^T (\,-\partial_t\xi(t)\,,\, b(u(t)) - b(u_0)\,)_H \,\mathrm{d}t + \int_0^T \langle\,\xi(t)\,,\, A(t,u(t))\,\rangle_V \,\mathrm{d}t = 0.$$

□

7. **Parabolic estimates.** We treat here the case of a constant constraint

$$M_h(t) = M^i = M \subset V \quad \text{nonempty, closed and convex.} \tag{7.1}$$

For sets, which are not constant in time, additional terms will occur in the following lemmata. To be precise, let us consider solutions

$$\begin{gathered} u_h(t) = u_0 \in H \text{ for } t < 0,\\ u_h(t) \in M,\ w_h^*(t) \in V^* \text{ for } t > 0, \text{ with}\\ \big(\,u_h(t) - v\,,\, \tfrac{1}{h}(b(u_h(t)) - b(u_h(t-h)))\,\big)_H + \langle\, u_h(t) - v\,,\, w_h^*(t)\,\rangle_V \leq 0\\ \text{for } t > 0 \text{ and for } v \in M. \end{gathered} \tag{7.2}$$

These elements can be given by different circumstances such as the time discrete solution with $w_h^*(t) = A_h(t,u_h(t))$. We assume that the quantities in (7.2) are step functions in time, that is in the following computations (5.10) is assumed, which means

$$\begin{aligned} u_h(t) &= \begin{cases} u^i & \text{for } (i-1)h < t \leq ih,\ i \in \mathbb{N},\\ u^0 & \text{for } -h < t \leq 0,\end{cases}\\ w_h^*(t) &= w^{*i} \quad \text{for } (i-1)h < t \leq ih,\ i \in \mathbb{N}.\end{aligned} \tag{7.3}$$

For the convergence of the time discrete solutions we have to show estimates which are independent of h. The first basic estimate is the

7.1 Energy estimate. Let u_h and w_h^* as in (7.2). Then we conclude, if $\bar{u}$ is a step function, and if $\bar{t}$ is a multiple of h,

$$B_{\bar{u}(\bar{t})}(u_h(\bar{t})) + \int_0^{\bar{t}} \langle\, u_h(t) - \bar{u}(t)\,,\, w_h^*(t)\,\rangle_V \,\mathrm{d}t$$
$$+ \int_0^{\bar{t}} \frac{1}{h} E_\psi(u_h(t), u_h(t-h), b(u_h(t-h)))\,\mathrm{d}t$$
$$\le B_{\bar{u}(0)}(u_0) + \int_0^{\bar{t}} \big(\, \partial_t^{-h}\bar{u}(t)\,,\, b(u_h(t))\,\big)_H\,,$$

provided $\bar{u}(t) \in M$ for $t > 0$ and $\bar{u}(0) \in H$.

If the function $\bar{u}$ is constant, that is $\bar{u} \in M$, and if we neglect the last part on the left side, $E_\psi \ge 0$, we obtain the standard version of the estimate

$$B_{\bar{u}}(u_h(\bar{t})) + \int_0^{\bar{t}} \langle\, u_h(t) - \bar{u}\,,\, w_h^*(t)\,\rangle_V \,\mathrm{d}t \le B_{\bar{u}}(u_0). \tag{7.4}$$

Proof. With (7.3) the inequality (7.2) reads

$$\big(\, u^i - v\,,\, b(u^i) - b(u^{i-1})\,\big)_H + h\,\langle\, u^i - v\,,\, w^{*i}\,\rangle_V \le 0$$

for $v \in M^i$ and for $i \in \mathbb{N}$. Setting $v = \bar{u}^i \in M^i$ and summing over $i = 1, \ldots, k$, one gets

$$\sum_{i=1}^k \big(\, u^i - \bar{u}^i\,,\, u^{*i} - u^{*i-1}\,\big)_H + \sum_{i=1}^k h\,\langle\, u^i - \bar{u}^i\,,\, w^{*i}\,\rangle_V \le 0.$$

Here we have used the notation $u^{*i} := b(u^i)$. It holds

$$\begin{aligned} \big(\, u^i\,,\, u^{*i}\,\big)_H &= \psi^*(u^{*i}) + \psi(u^i),\\ \big(\, u^i\,,\, u^{*i-1}\,\big)_H &\le \psi^*(u^{*i-1}) + \psi(u^i), \end{aligned} \tag{7.5}$$

where for the inequality we also can write

$$\big(\, u^i\,,\, u^{*i-1}\,\big)_H = \psi^*(u^{*i-1}) + \psi(u^i) - E_\psi(u^i, u^{i-1}, u^{*i-1})$$

by taking the identity for $\psi^*(u^{*i-1})$ into account (see section 3). Therefore one obtains

$$\begin{aligned}
&\sum_{i=1}^k \big(\, u^i - \bar{u}^i\,,\, u^{*i} - u^{*i-1}\,\big)_H\\
&= \sum_{i=1}^k \Big(\big(\, u^i\,,\, u^{*i}\,\big)_H - \big(\, u^i\,,\, u^{*i-1}\,\big)_H \Big) - \sum_{i=1}^k \big(\, \bar{u}^i\,,\, u^{*i} - u^{*i-1}\,\big)_H\\
&= \sum_{i=1}^k \big(\psi^*(u^{*i}) - \psi^*(u^{*i-1})\big) - \sum_{i=1}^k \big(\, \bar{u}^i\,,\, u^{*i} - u^{*i-1}\,\big)_H + \sum_{i=1}^k E_\psi(u^i, u^{i-1}, u^{*i-1})\\
&= \sum_{i=1}^k \Big(\big(\psi^*(u^{*i}) - \big(\, \bar{u}^i\,,\, u^{*i}\,\big)_H\big) - \big(\psi^*(u^{*i-1}) - \big(\, \bar{u}^{i-1}\,,\, u^{*i-1}\,\big)_H\big) \Big)\\
&\quad + \sum_{i=1}^k \big(\, \bar{u}^i - \bar{u}^{i-1}\,,\, u^{*i-1}\,\big)_H + \sum_{i=1}^k E_\psi(u^i, u^{i-1}, u^{*i-1}).
\end{aligned}$$

Using that the first term is a telescope sum, and using the definition in 3.6, one gets that this is

$$= B_{\bar{u}^k}(u^k) - B_{\bar{u}^0}(u^0) + \sum_{i=1}^{k} h \left(\frac{1}{h}(\bar{u}^i - \bar{u}^{i-1}) \, , \, u^{*i-1} \right)_H + \sum_{i=1}^{k} E_\psi(u^i, u^{i-1}, u^{*i-1}).$$

Here in the term $B_{\bar{u}^0}(u^0)$ and in the term $\frac{1}{h}(\bar{u}^i - \bar{u}^{i-1})$ for $i = 1$ the function $\bar{u}^0$ occurs. Now rewriting terms as step functions in time, one gets the result. □

The second estimate is the following

7.2 Compactness in time. Let (7.1) be satisfied, and let u_h and w_h^* as in (7.2) as well as $u_h^*(t) = b(u_h(t))$. Then for t an s being a multiple of h, $s = jh$, we infer

$$E_{\psi^*}(u_h^*(t+s), u_h^*(t), u_h(t)) \le s \cdot \frac{1}{j} \sum_{i=1}^{j} \langle \, u_h(t) - u_h(t+ih) \, , \, w_h^*(t+ih) \, \rangle_V \, .$$

Proof. It is assumed that t and s are multiple of h, say,

$$t = kh, \quad t + s = (k+j)h. \tag{7.6}$$

As in the previous proof we write for $t^i = ih$, $i \in \mathbb{N}$, problem (7.2) as

$$\left(u^i - v \, , \, u^{*i} - u^{*i-1} \right)_H \le h \left\langle v - u^i \, , \, w^{*i} \right\rangle_V \tag{7.7}$$

for $v \in M$, where again we use the notation $u^{*i} := b(u^i)$.

Now choose $k \in \mathbb{N}$ and set $v = u^k$, and sum over $i = k+1, \ldots, k+j$. The result is

$$\sum_{i=k+1}^{k+j} \left(u^i - u^k \, , \, u^{*i} - u^{*i-1} \right)_H \le \sum_{i=k+1}^{k+j} h \left\langle u^k - u^i \, , \, w^{*i} \right\rangle_V \, .$$

Now by the identity 3.4(3) and Young's inequality (3.2), see (7.5), we compute for the left side

$$\begin{aligned}
&\sum_{i=k+1}^{k+j} \left(u^i - u^k \, , \, u^{*i} - u^{*i-1} \right)_H \\
&= \sum_{i=k+1}^{k+j} \left(\left(u^i \, , \, u^{*i} \right)_H - \left(u^i \, , \, u^{*i-1} \right)_H \right) - \sum_{i=k+1}^{k+j} \left(u^k \, , \, u^{*i} - u^{*i-1} \right)_H \\
&\ge \sum_{i=k+1}^{k+j} \left(\psi^*(u^{*i}) - \psi^*(u^{*i-1}) \right) - \left(u^k \, , \, \sum_{i=k+1}^{k+j} (u^{*i} - u^{*i-1}) \right)_H \\
&= \psi^*(u^{*k+j}) - \psi^*(u^{*k}) - \left(u^k \, , \, u^{*k+j} - u^{*k} \right)_H \\
&= E_{\psi^*}(u^{*k+j}, u^{*k}, u^k)
\end{aligned}$$

with E_{ψ^*} defined in 3.3 and since $u^k \in \partial\psi^*(u^{*k})$, a consequence of $u^{*k} \in \partial\psi(u^k)$, see 3.2. Thus we have shown

$$\begin{aligned} E_{\psi^*}(u^{*k+j}, u^{*k}, u^k) &\leq \sum_{i=k+1}^{k+j} \left(u^i - u^k \, , \, u^{*i} - u^{*i-1} \right)_H \\ &\leq \sum_{i=k+1}^{k+j} h \left\langle u^k - u^i \, , \, w^{*i} \right\rangle_V \\ &= s \cdot \frac{1}{j} \sum_{i=1}^{j} \left\langle u^k - u^{k+i} \, , \, w^{*k+i} \right\rangle_V . \end{aligned}$$

We rewrite this as

$$\begin{aligned} & E_{\psi^*}(u_h^*(t+s), u_h^*(t), u_h(t)) = E_{\psi^*}(u^{*k+j}, u^{*k}, u^k) \\ & \leq s \cdot \frac{1}{j} \sum_{i=1}^{j} \left\langle u^k - u^{k+i} \, , \, w^{*k+i} \right\rangle_V \\ & = s \cdot \frac{1}{j} \sum_{i=1}^{j} \left\langle u_h(t) - u_h(t+ih) \, , \, w_h^*(t+ih) \right\rangle_V . \end{aligned}$$

□

This proof works for a general convex set M. If M is a subspace one obtains a slightly better estimate.

7.3 Lemma. Let M be an affine subspace, and u_h and w_h^* as in (7.2). Then if t and s are multiple of h and $s = jh$, we infer

$$\begin{aligned} & \left(u_h(t+s) - u_h(t) \, , \, b(u_h(t+s)) - b(u_h(t)) \right)_H \\ & = s \left\langle u_h(t) - u_h(t+s) \, , \, \frac{1}{j} \sum_{i=1}^{j} w_h^*(t+ih) \right\rangle_V . \end{aligned}$$

Proof. As in the previous proof we know that (7.7) is satisfied, but now for an affine subspace M, so that

$$\left(v \, , \, u_h^{*i} - u_h^{*i-1} \right)_H = -h \left\langle v \, , \, w_h^{*i} \right\rangle_V \tag{7.8}$$

for $v \in M_1$, a subspace for which $M = \bar{u}_1 + M_1$ with $\bar{u}_1 \in M$. Now again choose $k \in \mathbb{N}$, and sum over $i = k+1, \ldots, k+j$. The result is

$$\left(v \, , \, u_h^{*k+j} - u_h^{*k} \right)_H = -h \left\langle v \, , \, \sum_{i=k+1}^{k+j} w_h^{*i} \right\rangle_V .$$

Now set $v = u_h^{k+j} - u_h^k$, and write the result in terms of functions in time, that is $t = kh$, see (7.3). □

8. **Parabolic identity.** In all theories about parabolic problems there is one equation, which plays an exceptional role, and it has to be proved for the continuous limit problem. For some parabolic problems it is connected to an inequality, which is postulated for the formulation of a solution. In this paper it is connected to the following.

8.1 Definition. Let $u : [0,T] \to H$ be measurable, $\bar{u} \in H$, such that $B_{\bar{u}}(u) \in L^\infty([0,T])$. Assume initial data $u_0 \in H$. For $v \in W^{1,1}(]0,T[;H)$ define

$$\begin{aligned}\Phi_{\bar{u}}(u,v)(t) := \quad & B_{\bar{u}}(u(t)) - B_{\bar{u}}(u_0) + (\, \bar{u} - v(t)\, ,\, b(u(t)) - b(u_0)\,)_H \\ & - \int_0^t (\, \partial_t(\bar{u} - v)(s)\, ,\, b(u(s)) - b(u_0)\,)_H \,\mathrm{d}s,\end{aligned}$$

so that $\Phi_{\bar{u}}(u,v) \in L^\infty([0,T])$.

Proof. We have to show that the definition is well posed. Since $B_{\bar{u}}(u) \in L^\infty([0,T])$ statement (3.7) implies $b(u) \in L^\infty([0,T];H)$. Then the term under the integral, since $\partial_t v \in L^1(]0,T[;H)$, is integrable. □

The definition can also be written as

$$\begin{aligned}\Phi_{\bar{u}}(u,v)(t) = \quad & \big(B_{\bar{u}}(u(t)) + (\, \bar{u} - v(t)\, ,\, b(u(t))\,)_H\,\big) \\ & -\big(B_{\bar{u}}(u_0) + (\, \bar{u} - v(0)\, ,\, b(u_0)\,)_H\,\big) \\ & - \int_0^t (\, \partial_t(\bar{u} - v)(s)\, ,\, b(u(s))\,)_H \,\mathrm{d}s.\end{aligned}$$

We mention, that only $B_{\bar{u}}(u) \in L^\infty([0,T])$ is assumed. The fact that u_0 is a "initial value" for u, is only determined by this definition. The term $\Phi_{\bar{u}}(u,v)$ is the parabolic term in the differential inequality (6.6). That this coincides with the parabolic term $\partial_t b(u)$ in the differential equation, is shown formally in the following lemma. The argumentation is essentially the same as the proof in the time discrete case in 10.3.

8.2 Remark. Formally

$$\Phi_{\bar{u}}(u,v)(\bar{t}) = \int_0^{\bar{t}} (\, u(t) - v(t)\, ,\, \partial_t b(u(t))\,)_H \,\mathrm{d}t$$

for every function u which has initial data $u(0) = u_0$.

Proof. We write

$$\begin{aligned}&\int_0^{\bar{t}} (\, u(t) - v(t)\, ,\, \partial_t b(u(t))\,)_H \,\mathrm{d}t \\ &= \int_0^{\bar{t}} (\, u(t) - \bar{u}\, ,\, \partial_t b(u(t))\,)_H \,\mathrm{d}t + \int_0^{\bar{t}} (\, \bar{u} - v(t)\, ,\, \partial_t(b(u(t)) - b(u(0)))\,)_H \,\mathrm{d}t.\end{aligned}$$

Now from 3.7

$$\begin{aligned}&(\, u(t) - \bar{u}\, ,\, b(u(t)) - b(u(t-\delta))\,)_H \\ &= B_{\bar{u}}(u(t)) - B_{\bar{u}}(u(t-\delta)) + E_\psi(u(t), u(t-\delta), \nabla\psi(u(t-\delta)),\end{aligned}$$

and if one considers the limit $\delta \to 0$ after dividing by δ, one obtains formally

$$\partial_t B_{\bar{u}}(u) = (\, u - \bar{u}\, ,\, \partial_t b(u)\,)_H\, .$$

Therefore the first term in the above identity is

$$\int_0^{\bar{t}} (\, u(t) - \bar{u}\, ,\, \partial_t b(u(t))\,)_H \,\mathrm{d}t = \int_0^{\bar{t}} \partial_t B_{\bar{u}}(u(t))\,\mathrm{d}t = B_{\bar{u}}(u(\bar{t})) - B_{\bar{u}}(u(0)),$$

and for the second term one computes

$$\int_0^{\bar t} (\, \bar u - v(t)\, ,\, \partial_t(b(u(t)) - b(u(0)))\,)_H \, \mathrm{d}t$$
$$= (\, \bar u - v(\bar t)\, ,\, b(u(\bar t)) - b(u(0))\,)_H - \int_0^{\bar t} (\, \partial_t(\bar u - v(t))\, ,\, b(u(t)) - b(u(0))\,)_H \, \mathrm{d}t.$$

Add both terms in order to get $\Phi_{\bar u}(u, v)(\bar t)$. □

This statement indicates, that formally $\Phi_{\bar u}(u, u) = 0$. In a rigorous way this will be proved in the following lemma.

8.3 Lemma. Let $u : [0, T] \to H$ be measurable, $\bar u \in H$, such that $B_{\bar u}(u) \in L^\infty([0, T])$. Then

$$\liminf_{\varepsilon\to 0} \liminf_{\delta\to 0} \Phi_{\bar u}(u, u_{\delta\varepsilon})(\bar t) \geq 0$$

for almost all $\bar t > 0$, provided $u_{0\varepsilon} \to u_0$ in H as $\varepsilon \to 0$. Here $u_{\delta\varepsilon} \in W^{1,1}(]0, T[; H)$ is the function

$$u_{\delta\varepsilon}(t) := \frac{1}{\delta}\int_{t-\delta}^{t} \widetilde{u}_\varepsilon(s)\, \mathrm{d}s, \qquad \text{where} \qquad \widetilde{u}_\varepsilon(t) := \begin{cases} u(t) & \text{for } t > 0, \\ u_{0\varepsilon} & \text{for } t < 0. \end{cases}$$

The function $u_0 \in H$ is in definition 8.1 of $\Phi_{\bar u}$.
Hint: If $u \in L^p(]0, T[; V)$ and $u_{0\varepsilon} \in V$, then $u_{\delta\varepsilon} \in W^{1,p}(]0, T[; V)$. The function $u_0 \in H$ is the initial datum.

Proof. Obviously $\Phi_{\bar u}(u, u_{\delta\varepsilon})$ is defined, since $u_{\delta\varepsilon} \in W^{1,1}(]0, T[; H)$ with $\partial_t u_{\delta\varepsilon} = \partial_t^{-\delta}\widetilde{u}_\varepsilon$, where $\partial_t^{-\delta}$ is the backward differential quotient. Now by definition 8.1

$$\begin{aligned}\Phi_{\bar u}(u, u_{\delta\varepsilon})(\bar t) = \quad & B_{\bar u}(u(\bar t)) - B_{\bar u}(u_0) + (\, \bar u - u_{\delta\varepsilon}(\bar t)\, ,\, b(u(\bar t)) - b(u_0)\,)_H \\ & - \int_0^{\bar t} (\, \partial_t(\bar u - u_{\delta\varepsilon}(t))\, ,\, b(u(t)) - b(u_0)\,)_H \, \mathrm{d}t.\end{aligned}$$

We see that with discrete partial integration, defining $u(t) := u_0$ for $t < 0$ and therefore $b(u(t)) = b(u_0)$ for $t < 0$,

$$\begin{aligned}&\int_0^{\bar t} (\, \partial_t(\bar u - u_{\delta\varepsilon}(t))\, ,\, b(u(t)) - b(u_0)\,)_H \, \mathrm{d}t \\ &= \int_0^{\bar t} \big(\, \partial_t^{-\delta}(\bar u - \widetilde{u}_\varepsilon(t))\, ,\, b(u(t)) - b(u_0)\, \big)_H \, \mathrm{d}t \\ &= \frac{1}{\delta}\Big(\int_0^{\bar t} (\, \bar u - \widetilde{u}_\varepsilon(t)\, ,\, b(u(t)) - b(u_0)\,)_H \, \mathrm{d}t \\ &\qquad - \int_{-\delta}^{\bar t-\delta} (\, \bar u - \widetilde{u}_\varepsilon(t)\, ,\, b(u(t+\delta)) - b(u_0)\,)_H \, \mathrm{d}t \Big) \\ &= \frac{1}{\delta}\int_{\bar t-\delta}^{\bar t} (\, \bar u - \widetilde{u}_\varepsilon(t)\, ,\, b(u(t)) - b(u_0)\,)_H \, \mathrm{d}t - \int_{-\delta}^{\bar t-\delta} \big(\, \bar u - \widetilde{u}_\varepsilon(t)\, ,\, \partial_t^{+\delta} b(u(t))\, \big)_H \, \mathrm{d}t\end{aligned}$$

$$
\begin{aligned}
= \ & \frac{1}{\delta}\int_{\bar t-\delta}^{\bar t} (\,\bar u - u(t)\,,\, b(u(t)) - b(u_0)\,)_H \,\mathrm{d}t \\
& - \int_{-\delta}^{\bar t-\delta} \big(\,\bar u - u(t)\,,\, \partial_t^{+\delta} b(u(t))\,\big)_H \,\mathrm{d}t \\
& - \frac{1}{\delta}\int_{-\delta}^{0} (\,u_0 - u_{0\varepsilon}\,,\, b(u(t+\delta)) - b(u_0)\,)_H \,\mathrm{d}t.
\end{aligned}
$$

Here we have used that $\bar t - \delta > 0$, and that $\widetilde u_\varepsilon(t) = u(t)$ for $t > 0$. Now, let us have a look at each of the three terms in the above identity.

In the first term, since $t \mapsto b(u(t))$ is in $L^\infty([0,T];H)$ by assumption and (3.7), the integrand $t \mapsto (\,\bar u - u(t)\,,\, b(u(t)) - b(u_0)\,)_H$ is integrable. Therefore this term converges for almost all $\bar t$ to $(\,\bar u - u(\bar t)\,,\, b(u(\bar t)) - b(u_0)\,)_H$, a term which occurs also in the formula for $\Phi_{\bar u}(u, u_{\delta\varepsilon})(\bar t)$, hence this term cancels.

The last term is, again since $t \mapsto b(u(t))$ is in $L^\infty([0,T];H)$,

$$
\begin{aligned}
& \left| \frac{1}{\delta}\int_{-\delta}^{0} (\,u_0 - u_{0\varepsilon}\,,\, b(u(t+\delta)) - b(u_0)\,)_H \,\mathrm{d}t \right| \\
& \le \|u_0 - u_{0\varepsilon}\|_H \cdot \operatorname*{ess\,sup}_{t\in[0,T]} \|b(u(t)) - b(u_0)\|_H \to 0
\end{aligned}
$$

as $u_{0\varepsilon} \to u_0$ for $\varepsilon \to 0$.

Concerning the second term we use the inequality 3.7, that is for all t, s

$$
\begin{aligned}
& (\,u(t) - \bar u\,,\, b(u(t)) - b(u(s))\,)_H \\
& = B_{\bar u}(u(t)) - B_{\bar u}(u(s)) + E_\psi(u(t), u(s), \nabla\psi(u(s))) \\
& \ge B_{\bar u}(u(t)) - B_{\bar u}(u(s)),
\end{aligned}
$$

therefore with $s = t + \delta$

$$
\begin{aligned}
& - \big(\,\bar u - u(t)\,,\, \partial_t^{+\delta} b(u(t))\,\big)_H \\
& = -\frac{1}{\delta}\,(\,\bar u - u(t)\,,\, b(u(t+\delta)) - b(u(t))\,)_H \\
& \le \frac{1}{\delta}\,(B_{\bar u}(u(t+\delta)) - B_{\bar u}(u(t))) = \partial_t^{+\delta} B_{\bar u}(u(t)),
\end{aligned}
$$

and the second term becomes

$$
\begin{aligned}
& - \int_{-\delta}^{\bar t-\delta} \big(\,\bar u - u(t)\,,\, \partial_t^{+\delta} b(u(t))\,\big)_H \,\mathrm{d}t \le \int_{-\delta}^{\bar t-\delta} \partial_t^{+\delta} B_{\bar u}(u(t))\,\mathrm{d}t \\
& = \int_0^{\bar t} \partial_t^{-\delta} B_{\bar u}(u(t))\,\mathrm{d}t = \frac{1}{\delta}\int_{\bar t-\delta}^{\bar t} B_{\bar u}(u(t))\,\mathrm{d}t - B_{\bar u}(u_0).
\end{aligned}
$$

Since $t \mapsto B_{\bar u}(u(t))$ is integrable, this converges for a subsequence $\delta \to 0$ (a subsequence of an a-priori given sequence $\delta \to 0$, see the remark at the end of this proof) for almost all $\bar t$ to $B_{\bar u}(u(\bar t)) - B_{\bar u}(u_0)$, a term which occurs in the formula for $\Phi_{\bar u}(u, u_{\delta\varepsilon})(\bar t)$.

Remark: We mention that for the sequence $\delta \to 0$ one has to apply a certain trick. First one chooses a subsequence so that the limit with respect to this subsequence is the limes inferior in the assertion. With this subsequence one has to go into the above proof with the choice of a subsubsequence. □

9. **Compactness theorem.** The main statement of this section is the compactness result in 9.3, whose proof is based on the corresponding result in [5]. Before we show this we present a useful lemma.

9.1 Lemma. Let $\bar{t} = mh$ and assume u_h are step functions, see (5.10). If

$$\int_0^{\bar{t}-s} e(u_h(t+s), u_h(t)) \le C \cdot \omega(s)$$

for $s > 0$ which are multiple of h, then this inequality holds for any real $s > 0$. Here ω is a concave function and $e : H \times H \to \mathbb{R}$ continuous.

The lemma applies to our function $\omega(s) = s$. By the way, the property $\omega(0) = 0$ is assumed in 9.3.

Proof. Let s be arbitrary, $0 < s < \bar{t}$, and choose $j \in \mathbb{N}$ with

$$s = jh + \sigma, \quad 0 \le \sigma < h,$$

so that

$$\left(1 - \tfrac{\sigma}{h}\right) \cdot jh + \tfrac{\sigma}{h} \cdot (j+1)h = s. \tag{9.1}$$

Then if $t = (i-1)h + \tau$, $0 < \tau \le h$, we compute by (5.10) for step functions u_h

$$u_h(t+s) = \begin{cases} u^{i+j} & \text{if } \sigma + \tau \le h, \\ u^{i+j+1} & \text{if } \sigma + \tau > h, \end{cases}$$

and therefore

$$\begin{aligned}
&\int_0^{mh-s} e(u_h(t+s), u_h(t))\,\mathrm{d}t \\
&= \sum_{i=1}^{m-j} \int_{(i-1)h}^{ih-\sigma} e(u_h(t+s), u_h(t))\,\mathrm{d}t + \sum_{i=1}^{m-j-1} \int_{ih-\sigma}^{ih} e(u_h(t+s), u_h(t))\,\mathrm{d}t \\
&= \sum_{i=1}^{m-j} (h-\sigma) e(u^{i+j}, u^i) + \sum_{i=1}^{m-j-1} \sigma e(u^{i+j+1}, u^i)
\end{aligned}$$

$$\begin{aligned}
&= \left(1 - \frac{\sigma}{h}\right) \int_0^{mh-jh} e(u_h(t+jh), u_h(t))\,\mathrm{d}t \\
&\quad + \frac{\sigma}{h} \int_0^{mh-(j+1)h} e(u_h(t+(j+1)h), u_h(t))\,\mathrm{d}t \\
&\le C \cdot \left(\left(1 - \frac{\sigma}{h}\right) \cdot \omega(jh) + \frac{\sigma}{h} \cdot \omega((j+1)h) \right) \le C \cdot \omega(s)
\end{aligned}$$

by (9.1), since ω is concave. □

The following statement we will apply in the main proof.

9.2 Theorem. Let H be a Hilbert space and $V \subset H$ a Banach space with compact embedding $\mathrm{Id}_V : H \to V$. Further, let

$$b = \nabla\psi : H \to H$$

as in (6.4). Then for $R > 0$ there exist a continuous function $\omega_R : [0,\infty[\to [0,\infty[$ with $\omega_R(0) = 0$ so, that for all $\delta > 0$

$$\left\|u^1\right\|_V \le R, \quad \left\|u^2\right\|_V \le R, \quad E_{\psi^*}(b(u^1), b(u^2), u^2) \le \delta \tag{9.2}$$

implies

$$\left\|b(u^1) - b(u^2)\right\|_H \le \varepsilon \tag{9.3}$$

with $\varepsilon := \omega_R(\delta)$.

Proof. The function ω_R can be constructed, if for every $\varepsilon > 0$ there exist $\delta > 0$ such that (9.2) implies (9.3). Thus we have to show

$$\forall\ R > 0\ \forall\ \varepsilon > 0\ \exists\ \delta > 0 :\ (\ (9.2) \text{ implies } (9.3)\ )\ .$$

Assume this is not true. This is equivalent to

$$\exists\ R > 0\ \exists\ \varepsilon > 0\ \forall\ \delta > 0 :\ (\ (9.2) \text{ and not } (9.3)\ )\ .$$

Let such numbers $R > 0$ and $\varepsilon > 0$ be given. Hence for small $\delta > 0$ there are $u_\delta^1, u_\delta^2 \in V$ with $\left\|u_\delta^1\right\|_V \le R$, $\left\|u_\delta^2\right\|_V \le R$, and $E_{\psi^*}(b(u_\delta^1), b(u_\delta^2), u_\delta^2) \le \delta$, but

$$\left\|b(u_\delta^1) - b(u_\delta^2)\right\|_H > \varepsilon. \tag{9.4}$$

The boundedness in V and the compactness of the embedding $V \hookrightarrow H$ imply that there are $u^1, u^2 \in H$ with $u_\delta^1 \to u^1$ and $u_\delta^2 \to u^2$ in H for a subsequence $\delta \to 0$. Since b is continuous it follows for this subsequence that $b(u_\delta^1) \to b(u^1)$ and $b(u_\delta^2) \to b(u^2)$ in H. Similarly since ψ is continuous we obtain $\psi(u_\delta^1) \to \psi(u^1)$ and $\psi(u_\delta^2) \to \psi(u^2)$ in $\mathbb{R}$. Hence by 3.5

$$\begin{aligned} \delta \ge\ & E_{\psi^*}(b(u_\delta^1), b(u_\delta^2), u_\delta^2) = E_\psi(u_\delta^2, u_\delta^1, b(u_\delta^1)) \\ =\ & \psi(u_\delta^2) - \psi(u_\delta^1) + \big(\, u_\delta^1 - u_\delta^2\,,\, b(u_\delta^1)\,\big)_H \\ \to\ & \psi(u^2) - \psi(u^1) + \big(\, u^1 - u^2\,,\, b(u^1)\,\big)_H = E_\psi(u^2, u^1, b(u^1)) \ge 0 \end{aligned}$$

as $\delta \to 0$. It follows that

$$0 = E_\psi(u^2, u^1, b(u^1)) = \psi(u^2) - \psi(u^1) + \big(\, u^1 - u^2\,,\, b(u^1)\,\big)_H\ . \tag{9.5}$$

Besides this we compute for every $v \in H$

$$\begin{aligned} 0 \le\ & E_\psi(u^2 + v, u^1, b(u^1)) \\ =\ & \psi(u^2 + v) - \psi(u^1) + \big(\, u^1 - u^2 - v\,,\, b(u^1)\,\big)_H \\ =\ & \psi(u^2 + v) - \psi(u^2) - \big(\, v\,,\, b(u^1)\,\big)_H + E_\psi(u^2, u^1, b(u^1)) \\ =\ & \psi(u^2 + v) - \psi(u^2) - \big(\, v\,,\, b(u^1)\,\big)_H \end{aligned}$$

by inserting the identity (9.5). We get

$$0 \le \psi(u^2 + v) - \psi(u^2) - \big(\, v\,,\, b(u^1)\,\big)_H$$

for all $v \in H$, that is

$$b(u^1) \in \partial\psi(u^2) = \{b(u^2)\}\ ,$$

since ψ is differentiable. We conclude $b(u^1) = b(u^2)$ and therefore by (9.4)

$$\varepsilon < \left\|b(u_\delta^1) - b(u_\delta^2)\right\|_H \to \left\|b(u^1) - b(u^2)\right\|_H = 0$$

as $\delta \to 0$, a contradiction. ∎

With this we are able to prove the main result.

9.3 Compactness result. Let V be a separable reflexive Banach space and H a Hilbert space with compact embedding $\mathrm{Id}_V : H \to V$. Further, let $p \in]1,\infty[$, $t_0 < t_1$, and

$$b = \nabla\psi : H \to H$$

as in (6.4) (see also (5.5)). For $C > 0$ let

$$\begin{aligned}\mathcal{K}_C := \Big\{ \quad & u \in L^p([t_0,t_1];V)\ ;\ \|u\|^p_{L^p([t_0,t_1];V)} \le C, \\ & \operatorname*{ess\,sup}_{t\in]t_0,t_1[} B_0(u(t)) \le C, \\ & \forall s \in [0, t_1-t_0] : \int_{t_0}^{t_1-s} E_{\psi^*}(b(u(t+s)), b(u(t)), u(t))\,\mathrm{d}t \le C\cdot\omega(s) \ \Big\}.\end{aligned}$$

Here ω is continuous with $\omega(0) = 0$. Then

$$\{b(u)\,;\ u \in \mathcal{K}_C\} \text{ is precompact in } L^1([t_0,t_1];H).$$

Proof part 1. We prove that

$$\sup_{u\in\mathcal{K}_C} \int_{t_0}^{t_1-s} \|b(u(t+s)) - b(u(t))\|_H\,\mathrm{d}t \to 0 \quad \text{ as } s \to 0.$$

For this define

$$\begin{aligned}\mathcal{T}^s_R(u) := \Big\{ \quad & t \in [t_0, t_1-s]\ ;\ \|u(t)\|^p_V \le R^p \text{ and } \|u(t+s)\|^p_V \le R^p \\ & \text{and } \frac{1}{\omega(s)} E_{\psi^*}(b(u(t+s)), b(u(t)), u(t)) \le R^p \ \Big\}\ ,\end{aligned} \tag{9.6}$$

where we assume that $\omega(s) \neq 0$. Then for $u \in \mathcal{K}_C$

$$\int_{t_0}^{t_1-s} \left(\|u(t)\|^p_V + \|u(t+s)\|^p_V + \frac{1}{\omega(s)} E_{\psi^*}(b(u(t+s)), b(u(t)), u(t)) \right) \mathrm{d}t \le 3C.$$

Considering the integrand on $[t_0,t_1-s] \setminus \mathcal{T}^s_R(u)$ we see that

$$R^p \cdot \mathrm{L}^1([t_0,t_1-s] \setminus \mathcal{T}^s_R(u)) \le 3\cdot C\ , \tag{9.7}$$

therefore the Lebesgue measure of the set $[t_0,t_1-s] \setminus \mathcal{T}^s_R(u)$ is estimated by a constant depending on R alone. By (3.7) we compute for $u \in \mathcal{K}_C$

$$\|b(u(t))\|_H \le C_1 := C + \widetilde{C}_{1,0} \quad \text{ for almost all } t \in [t_0,t_1]. \tag{9.8}$$

Then by 9.2 and (9.8)

$$\|b(u(t+s)) - b(u(t))\|_H \le \begin{cases} \omega_R(\omega(s)\cdot R^p) & \text{for } t \in \mathcal{T}^s_R(u), \\ 2\cdot C_1 & \text{for } t \in [t_0,t_1-s] \setminus \mathcal{T}^s_R(u), \end{cases}$$

and integrating this gives

$$\begin{aligned}&\int_{t_0}^{t_1-s} \|b(u(t+s)) - b(u(t))\|_H\,\mathrm{d}t \\ &\le (t_1-t_0)\cdot\omega_R(\omega(s)\cdot R^p) + 2\cdot C_1\cdot\frac{3\cdot C}{R^p}\ .\end{aligned}$$

The right side is independent of u. First we choose R large enough, so that the second term on the right becomes small, and then s small, so that the first term is small. It follows that the integral is small uniformly in $u \in \mathcal{K}_C$. □

Proof part 2. We prove that

$$\sum_{i\le k}\frac{1}{\delta}\int_{t_0+(i-1)\delta}^{t_0+i\delta}\int_{t_0+(i-1)\delta}^{t_0+i\delta}\|b(u(t))-b(u(s))\|_H\,\mathrm{d}t\,\mathrm{d}s$$
$$\le\frac{2}{\delta}\int_0^\delta\left(\int_{t_0}^{t_1-s}\|b(u(t))-b(u(t+s))\|_H\,\mathrm{d}t\right)\mathrm{d}s\le 2\cdot\varepsilon_R^\delta$$

with

$$\varepsilon_R^\delta:=(t_1-t_0)\cdot\omega_R(\omega(\delta)\cdot R^p)+2\cdot C_1\cdot\frac{3\cdot C}{R^p},$$

if both ω_R and ω are assumed to be concave and monotone, and of course $\omega_R(0)=0$ and $\omega(0)=0$. The estimate is true for every time interval $[0,\delta]$ with

$$\delta=\frac{t_1-t_0}{k}\quad\text{ and } k\in\mathbb{N}. \tag{9.9}$$

Indeed

$$\sum_{i\le k}\frac{1}{\delta}\int_{t_0+(i-1)\delta}^{t_0+i\delta}\int_{t_0+(i-1)\delta}^{t_0+i\delta}\|b(u(t))-b(u(s))\|_H\,\mathrm{d}t\,\mathrm{d}s$$
$$=\sum_{i\le k}\frac{1}{\delta}\int_0^\delta\int_0^\delta\|b(u(t_0+(i-1)\delta+s_1))-b(u(t_0+(i-1)\delta+s_2))\|_H\,\mathrm{d}s_2\,\mathrm{d}s_1$$
$$=2\cdot\sum_{i\le k}\frac{1}{\delta}\int_0^\delta\int_{s_1}^\delta\|b(u(t_0+(i-1)\delta+s_1))-b(u(t_0+(i-1)\delta+s_2))\|_H\,\mathrm{d}s_2\,\mathrm{d}s_1$$
$$=2\cdot\sum_{i\le k}\frac{1}{\delta}\int_0^\delta\int_0^{\delta-s_1}\|b(u(t_0+(i-1)\delta+s_1))-b(u(t_0+(i-1)\delta+s_1+s))\|_H\cdot\mathrm{d}s\,\mathrm{d}s_1$$
$$\le\frac{2}{\delta}\int_{t_0}^{t_1}\int_0^{\min(\delta,t_1-t)}\|b(u(t))-b(u(t+s))\|_H\,\mathrm{d}s\,\mathrm{d}t$$
$$\le\frac{2}{\delta}\int_0^\delta\int_{t_0}^{t_1-s}\|b(u(t))-b(u(t+s))\|_H\,\mathrm{d}t\,\mathrm{d}s\,.$$

This gives the result. □

Proof main part. We choose a time step $\delta>0$ as in (9.9) and approximate each function by a step function as follows. We define

$$\beta(s,v)(t):=\sum_{i=1}^k b(v(t_0+(i-1)\delta+s))\mathcal{X}_{]t_0+(i-1)\delta,t_0+i\delta]}(t)$$

for $v\in L^p([t_0,t_1];H)$ and $s\in[0,t_0-t_1]$ and we approximate

$$b(u)\quad\text{ by }\quad\beta(s,u_R^\delta)$$

for $u\in\mathcal{K}_C$ and a suitable s which we choose later. Here

$$u_R^\delta(t):=\begin{cases}u(t)\text{ if }t\in\mathcal{T}_R^\delta(u),\\ \bar{u}\text{ elsewhere},\end{cases}$$

where it is assumed that $\bar{u} \in \mathcal{K}_C$. We compute

$$\begin{aligned}
&\int_{t_0}^{t_1} \left\| b(u(t)) - \beta(s, u_R^\delta)(t) \right\|_H \mathrm{d}t \\
&= \int_{t_0}^{t_1} \left\| b(u(t)) - b(u_R^\delta(t)) \right\|_H \mathrm{d}t + \int_{t_0}^{t_1} \left\| b(u_R^\delta(t)) - \beta(s, u_R^\delta)(t) \right\|_H \mathrm{d}t
\end{aligned}$$

and by (9.8) and (9.7)

$$\begin{aligned}
&\int_{t_0}^{t_1} \left\| b(u(t)) - b(u_R^\delta(t)) \right\|_H \mathrm{d}t = \int_{[t_0,t_1]\setminus \mathcal{T}_R^\delta(u)} \| b(u(t)) - b(\bar{u}) \|_H \mathrm{d}t \\
&\leq 2 \cdot C_1 \cdot \mathrm{L}^1([t_0, t_1] \setminus \mathcal{T}_R^\delta(u)) = 2 \cdot C_1 \cdot \frac{3 \cdot C}{R^p} := \varepsilon_R \ ,
\end{aligned}$$

hence

$$\begin{aligned}
&\int_{t_0}^{t_1} \left\| b(u(t)) - \beta(s, u_R^\delta)(t) \right\|_H \mathrm{d}t \\
&= \varepsilon_R + \int_{t_0}^{t_1} \left\| b(u_R^\delta(t)) - \beta(s, u_R^\delta)(t) \right\|_H \mathrm{d}t \ .
\end{aligned}$$

Integrating over the offset s we obtain

$$\begin{aligned}
&\frac{1}{\delta} \int_0^\delta \int_{t_0}^{t_1} \left\| b(u(t)) - \beta(s, u_R^\delta)(t) \right\|_H \mathrm{d}t \, \mathrm{d}s \\
&= \varepsilon_R + \frac{1}{\delta} \int_0^\delta \int_{t_0}^{t_1} \left\| b(u_R^\delta(t)) - \beta(s, u_R^\delta)(t) \right\|_H \mathrm{d}t \, \mathrm{d}s \\
&\leq \varepsilon_R + \frac{1}{\delta} \int_0^\delta \sum_{i=1}^k \int_{t_0+(i-1)\delta}^{t_0+i\delta} \left\| b(u_R^\delta(t)) - b(u_R^\delta(t_0 + (i-1)\delta + s)) \right\|_H \mathrm{d}t \, \mathrm{d}s \\
&= \varepsilon_R + \frac{1}{\delta} \sum_{i=1}^k \int_{t_0+(i-1)\delta}^{t_0+i\delta} \int_{t_0+(i-1)\delta}^{t_0+i\delta} \left\| b(u_R^\delta(t)) - b(u_R^\delta(s)) \right\|_H \mathrm{d}t \, \mathrm{d}s.
\end{aligned}$$

By setting $\mathcal{T}_i := [t_0 + (i-1)\delta, t_0 + i\delta] \cap \mathcal{T}_R^\delta(u)$ this is

$$\begin{aligned}
&= \ \varepsilon_R + \frac{2}{\delta} \sum_{i=1}^k \int_{[t_0+(i-1)\delta, t_0+i\delta]\setminus\mathcal{T}_i} \int_{\mathcal{T}_i} \| b(u(t)) - b(\bar{u}) \|_H \mathrm{d}t \, \mathrm{d}s \\
&\quad + \frac{1}{\delta} \sum_{i=1}^k \int_{\mathcal{T}_i} \int_{\mathcal{T}_i} \| b(u(t)) - b(u(s)) \|_H \mathrm{d}t \, \mathrm{d}s \\
&= \ \varepsilon_R + 2 \sum_{i=1}^k \int_{[t_0+(i-1)\delta, t_0+i\delta]\setminus\mathcal{T}_i} \left(\frac{1}{\delta} \int_{\mathcal{T}_i} \| b(u(t)) - b(\bar{u}) \|_H \mathrm{d}t \right) \mathrm{d}s \\
&\quad + \frac{1}{\delta} \sum_{i=1}^k \int_{\mathcal{T}_i} \int_{\mathcal{T}_i} \| b(u(t)) - b(u(s)) \|_H \mathrm{d}t \, \mathrm{d}s \\
&\leq \ \varepsilon_R + 2 \cdot \mathrm{L}^1([t_0, t_1] \setminus \mathcal{T}_R^\delta(u)) \cdot 2 \cdot C_1 \\
&\quad + \frac{1}{\delta} \sum_{i=1}^k \int_{t_0+(i-1)\delta}^{t_0+i\delta} \int_{t_0+(i-1)\delta}^{t_0+i\delta} \| b(u(t)) - b(u(s)) \|_H \mathrm{d}t \, \mathrm{d}s \\
&\leq 3 \cdot \varepsilon_R + 2 \cdot \varepsilon_R^\delta
\end{aligned}$$

by the above estimate. Hence

$$\sup_{u\in\mathcal{K}_C}\frac{1}{\delta}\int_0^\delta\int_{t_0}^{t_1}\big\|b(u(t))-\beta(s,u_R^\delta)(t)\big\|_H\,\mathrm{d}t\le 3\cdot\varepsilon_R+2\cdot\varepsilon_R^\delta\le\varepsilon, \tag{9.10}$$

if R is large and then δ is small. We then can choose an $s=s_u$ such that for $u\in\mathcal{K}_C$

$$\int_{t_0}^{t_1}\big\|b(u(t))-\beta(s_u,u_R^\delta)(t)\big\|_H\,\mathrm{d}t\le 2\cdot\varepsilon. \tag{9.11}$$

It follows that the functions $b(u)$ lie in an 2ε-neighbourhood of the step functions $\beta(s_u,u_R^\delta)$,

$$\begin{gathered}\{b(u)\in L^1([t_0,t_1];H)\,;\ u\in\mathcal{K}_C\}\\ \subset \mathrm{B}_{2\varepsilon}\big(\{\beta(s_u,u_R^\delta)\in L^1([t_0,t_1];H)\,;\ u\in\mathcal{K}_C,\ R\text{ large},\ \delta\text{ small}\}\big)\end{gathered}$$

in the topology with respect to $L^1([t_0,t_1];H)$. □

Proof last part. From the previous proof it follows, that the precompactness of

$$\{b(u)\,;\ u\in\mathcal{K}_C\}\subset L^1([t_0,t_1];H)$$

follows from the precompactness of

$$\{\beta(s_u,u_R^\delta)\,;\ u\in\mathcal{K}_C,\ R\text{ large},\ \delta\text{ small}\}\subset L^1([t_0,t_1];H),$$

since the first set is contained in an 2ε-neigbourhood of the second set, ε an arbitrary small number. The second set depends on ε, which is allowed. Indeed this is true, since R was chosen large enough and δ small enough, both depending on ε.

Therefore the precompactness of the second set in $L^1([t_0,t_1];H)$ has to be shown. Since these are step functions, we have to show the precompactness of the steps in H, that is the precompactness of

$$\{b(u_R^\delta(t_0+(i-1)\delta+s_u))\,;\ u\in\mathcal{K}_C,\ i\le k,\ R\text{ large},\ \delta\text{ small}\}\subset H.$$

We show instead the precompactness of the larger set

$$\{b(u)\,;\ u\in V,\ \|u\|_V\le R\}\subset H$$

for large R. But this follows from the compactness of the embedding $V\hookrightarrow H$. Then bounded sets in V are precompact in H, and the continuous function b transforms this to a precompact set in H. □

The compactness of the functions $b(u_h)$ in $L^1([0,T];H)$ implies, that for a sequence $h\to 0$ these functions have a strong limit b^* in $L^1([0,T];H)$. Then, if the functions u_h already have a weak limit u, one can apply the following lemma, whose proof is classical. It shows that the limits satisfy $b^*=b(u)$. Note, that this is true, even if b is not strictly increasing, however it must be monotone.

9.4 Lemma. If $u_m,u\in L^p([0,T];V)$ and $b^*\in L^1([0,T];H)$ with

$$\begin{gathered}u_m\to u\text{ weakly in }L^p([0,T];V)\text{ as }m\to\infty,\\ b(u_m)\to b^*\text{ strongly in }L^1([0,T];H)\text{ as }m\to\infty,\end{gathered}$$

then

$$b^*=b(u).$$

Proof. Define $R : L^p([0,T];H) \times L^1([0,T];H) \to L^\infty([0,T];H)$ by

$$R(v,\beta)(t) := \frac{b(v(t)) - \beta(t)}{1 + \|b(v(t))\|_H + \|\beta(t)\|_H}.$$

We want to show, that $R(u,b^*) = 0$, that is $b(u) = b^*$. Now it follows from the assumption, that for a subsequence $b(u_m(t)) \to b^*(t)$ strongly in H for almost all t as $m \to \infty$. Therefore (for this subsequence)

$$R(v,b(u_m))(t) \to R(v,b^*)(t) \text{ strongly in } H \text{ as } m \to \infty. \tag{9.12}$$

To continue, we have to use the standard monotonicity argument for b, which implies

$$0 \le (\, v(t) - u_m(t)\,,\, b(v(t)) - b(u_m(t))\,)_H\,,$$

and therefore $0 \le (\, v(t) - u_m(t)\,,\, R(v,b(u_m))(t)\,)_H$ for almost all t. It follows that

$$\begin{aligned} 0 \le \quad & \int_0^T (\, v(t) - u_m(t)\,,\, R(v,b(u_m))(t)\,)_H \,\mathrm{d}t \\ = \quad & \int_0^T (\, v(t)\,,\, R(v,b(u_m))(t)\,)_H \,\mathrm{d}t - \int_0^T (\, u_m(t)\,,\, R(v,b(u_m))(t)\,)_H \,\mathrm{d}t. \end{aligned}$$

Since $R(v,b(u_m))$ is bounded in $L^\infty([0,T];H)$, it follows from (9.12) for every $q < \infty$, that $R(v,b(u_m)) \to R(v,b^*)$ strongly in $L^q([0,T];H)$ for $m \to \infty$. Setting $q = p^*$ we see that

$$\int_0^T (\, v(t)\,,\, R(v,b(u_m))(t)\,)_H \,\mathrm{d}t \to \int_0^T (\, v(t)\,,\, R(v,b^*)(t)\,)_H \,\mathrm{d}t \quad \text{ as } m \to \infty$$

for $v \in L^p([0,T];H)$. Now to the convergence of the second term. Since $u_m \to u$ weakly in $L^p([0,T];V)$, which is continuously embedded into $L^p([0,T];H)$, and since $R(v,b(u_m)) \to R(v,b^*)$ strongly in $L^{p^*}([0,T];H)$ for $m \to \infty$, it follows

$$\int_0^T (\, u_m(t)\,,\, R(v,b(u_m))(t)\,)_H \,\mathrm{d}t \to \int_0^T (\, u(t)\,,\, R(v,b^*)(t)\,)_H \,\mathrm{d}t \quad \text{ as } m \to \infty.$$

Altogether we conclude

$$0 \le \int_0^T (\, v(t) - u(t)\,,\, R(v,b^*)(t)\,)_H \,\mathrm{d}t$$

for all $v \in L^p([0,T];H)$. We apply now a Minty type argument, that is we replace v by $u + \varepsilon(v-u)$ and letting $\varepsilon \searrow 0$, to obtain

$$0 \le \int_0^T (\, v(t) - u(t)\,,\, R(u,b^*)(t)\,)_H \,\mathrm{d}t$$

for all $v \in L^p([0,T];H)$. Since $\widetilde{v} = v - u$ is an arbitrary element of $L^p([0,T];H)$ this gives

$$0 \le \int_0^T (\, \widetilde{v}(t)\,,\, R(u,b^*)(t)\,)_H \,\mathrm{d}t$$

for all $\widetilde{v} \in L^p([0,T];H)$, and therefore also

$$0 = \int_0^T (\, \widetilde{v}(t)\,,\, R(u,b^*)(t)\,)_H \,\mathrm{d}t,$$

which finally implies $R(u,b^*) = 0$ almost everywhere. $\square$

Having proved that $b^* = b(u)$ it follows that $u \in \mathcal{M}$, provided $u_m \in \mathcal{M}$ and $t \mapsto B_0(u_m(t))$ is bounded in $L^\infty([0,T])$ uniformly in m. This sequentially closedness is also true under some more general convergence conditions as shown below.

9.5 Lemma. If $u_m, u \in L^p([0,T];V)$ and

$$u_m \to u \text{ weakly in } L^p([0,T];V) \text{ as } m \to \infty,$$
$$b(u_m) \to b(u) \text{ weakly in } L^2([0,T];H) \text{ as } m \to \infty,$$
$$B_0(u_m) \text{ bounded in } L^\infty([0,T]),$$

then

$$u_m \in \mathcal{M} \text{ for } m \in \mathbb{N} \quad \text{implies} \quad u \in \mathcal{M}.$$

Proof. By the Lemma of Mazur u is the strong limit of elements in

$$\operatorname{conv}\{u_m \in V\,;\ m \in \mathbb{N}\}.$$

This implies that the weak convergence in $L^p([0,T];V)$ lets the set

$$\{u \in L^p([0,T];V)\,;\ u(t) \in M \text{ for almost all } t\},$$

which is closed and convex, invariant. On the other hand, by the same reason $w := b(u)$ is the strong limit in $L^2([0,T];H)$ of finite sums

$$\sum_{m\in\mathbb{N}} \alpha_m w_m \to w, \qquad \sum_{m\in\mathbb{N}} \alpha_m = 1, \quad \alpha_m \geq 0$$

with $w_m := b(u_m)$. Since ψ^* is lower semicontinuous, one concludes

$$\psi^*(w(t)) = \quad \psi^*\left(\lim_{m\to\infty} \sum_{m\in\mathbb{N}} \alpha_m w_m(t)\right) \leq \lim_{m\to\infty} \psi^*\left(\sum_{m\in\mathbb{N}} \alpha_m w_m(t)\right)$$

and the convexity of ψ^* implies, that

$$\psi^*\left(\sum_{m\in\mathbb{N}} \alpha_m w_m(t)\right) \leq \sum_{m\in\mathbb{N}} \alpha_m \psi^*(w_m(t)) \leq C - \psi(0),$$

since $\psi^*(w_m(t)) = B_0(u_m(t)) - \psi(0)$ and $B_0(u_m(t)) \leq C$ for some C. Since $\psi^*(w(t)) = B_0(u(t)) - \psi(0)$ we conclude $B_0(u) \in L^\infty([0,T])$. □

10. **Convergence proof.** In the following we give a convergence proof of the main theorem of this paper, formulated in 6.2. We consider the case of a time independent constraint $M \subset V$, and the boundedness condition 6.2(1) and the continuity condition in 6.1 are satisfied. Besides this we assume the coerciveness in 6.2(3).

First we show that we have approximative solutions of the time discrete problem. We define the time discrete operator to the map $(s,z) \mapsto A(s,z)$ from (6.5) by

$$A_h(t,z) := \frac{1}{h}\int_{(i-1)h}^{ih} A(s,z)\,\mathrm{d}s \quad \text{ for } (i-1)h < t \leq ih. \tag{10.1}$$

The initial data are $u_0 \in H$ and b is given by (6.4). With this we show

10.1. There are solutions u_h, which are step function with $u_h(t) = u_0$ for $t < 0$ and $u_h(t) \in M$ for $t > 0$, of the time discrete problem

$$\left(u_h(t) - v\,,\, \frac{1}{h}(b(u_h(t)) - b(u_h(t-h))) \right)_H + \langle\, u_h(t) - v\,,\, A_h(t, u_h(t))\,\rangle_V \leq 0$$

for $v \in M$. This is true for $t > 0$ which are multiple of h, and then also for all $t > 0$.

With $A^i(u) := A_h(t^i, u)$, $t^i = ih$, and $M^i = M$ for $i \in \mathbb{N}$ we have to show the assumptions in 5.5.

We want to prove 5.5(1). Let S be a bounded set in M^i. Define

$$S_i := \{u_i \in L^p([0,T];V)\,;\ u \in S\} \quad \text{with} \quad u_i(t) := \left\{ \begin{array}{l} u \text{ if } (i-1)h < t \le ih \\ \bar{u} \text{ elsewhere} \end{array} \right\}$$

and $\bar{u}$ an element in M. Then S_i is a bounded set in $L^p([0,T];M)$. Since the embedding $V \hookrightarrow H$ is compact, the values of u and, by 5.4, the values of $b(u)$ both are bounded in H for $u \in S$. It follows, by the definition in 3.6, that $B_0(u)$ are bounded, say $B_0(u) \le R$. We obtain

$$B_0(u_i(t)) = \left\{ \begin{array}{l} B_0(u) \text{ if } (i-1)h < t \le ih, \\ B_0(\bar{u}) \text{ elsewhere }, \end{array} \right.$$

that is

$$\| B_0(u_i) \|_{L^\infty([0,T])} \le \max(R, B_0(\bar{u})).$$

Hence on S_i also the $L^\infty([0,T])$-norm of B_0 is bounded. It follows from 6.2(1), that $\mathcal{A}$ on S_i is bounded in $L^{p^*}([0,T];V^*)$, say,

$$\begin{aligned} R^* &\ge \| \mathcal{A}(u_i) \|_{L^{p^*}([0,T];V^*)} = \left(\int_0^T \| A(t, u_i(t)) \|_{V^*}^{p^*} \, \mathrm{d}t \right)^{\frac{1}{p^*}} \\ &\ge \left(\int_{(i-1)h}^{ih} \| A(t,u) \|_{V^*}^{p^*} \, \mathrm{d}t \right)^{\frac{1}{p^*}} \ge h^{-\frac{1}{p}} \left\| \int_{(i-1)h}^{ih} A(t,u) \, \mathrm{d}t \right\|_{V^*} \\ &= h^{1-\frac{1}{p}} \left\| \frac{1}{h} \int_{(i-1)h}^{ih} A(t,u) \, \mathrm{d}t \right\|_{V^*} = h^{\frac{p-1}{p}} \| A^i(u) \|_{V^*}. \end{aligned}$$

Therefore A^i is bounded on S, which shows 5.5(1).

We want to prove 5.5(2) and we know 6.2(2), that is the continuity condition 6.1. Let a sequence be given for A^i as in 5.5(2), which is stated in 4.1, that is $u_m, u \in M^i = M$ with $u_m \to u$ weakly in V and $A^i(u_m) \to v^*$ weakly* in V^* for $m \to \infty$ and such that

$$\limsup_{m \to \infty} \langle\, u_m \,,\, A^i(u_m) \,\rangle_V \le \langle\, u \,,\, v^* \,\rangle_V \,.$$

We have to show, that the conclusions in 4.1 are true. Define

$$u_{im}(t) := \left\{ \begin{array}{l} u_m \text{ if } (i-1)h < t \le ih \\ \bar{u} \text{ elsewhere} \end{array} \right\}, \quad v_i^*(t) := \left\{ \begin{array}{l} v^* \text{ if } (i-1)h < t \le ih \\ A(t,\bar{u}) \text{ elsewhere} \end{array} \right\},$$

and u_i to u as u_{im} to u_m. Then for $m \to \infty$ it converges $u_{im} \to u_i$ weakly in $L^p([0,T];V)$, and $\mathcal{A}(u_{im}) \to v_i^*$ weakly* in $L^{p^*}([0,T];V^*)$, and

$$\limsup_{m \to \infty} \int_0^T \langle\, u_{im}(t) \,,\, A(t, u_{im}(t)) \,\rangle_V \, \mathrm{d}t \le \int_0^T \langle\, u_i(t) \,,\, v_i^*(t) \,\rangle_V \, \mathrm{d}t.$$

Since the embedding $V \hookrightarrow H$ is compact, u_m converges to u strongly in H. Since b is continuous also $b(u_m) \to b(u)$ strongly in H, and by the definition 3.6 it follows that $B(u_m) \to B(u)$ in $\mathbb{R}$. This implies that $\{B(u_{im})\,;\ m \in \mathbb{N}\}$ is bounded in

$L^\infty([0,T];\mathbb{R})$ and that $b(u_{im}) \to b(u_i)$ strongly in $L^2([0,T];H)$. Thus the assumptions in the continuity condition 6.1 are fulfilled, hence 6.1 can be applied, and we conclude

$$\int_0^T \langle\, u_i(t) - v(t)\,,\, A(t,u_i(t)) - v_i^*(t)\,\rangle_V \,\mathrm{d}t \le 0 \tag{10.2}$$

for all $v \in L^p([0,T];V)$ with $v(t) \in M$ for almost all t and

$$\limsup_{m\to\infty} \int_0^T \langle\, u_{im}(t)\,,\, A(t,u_{im}(t))\,\rangle_V \,\mathrm{d}t = \int_0^T \langle\, u_i(t)\,,\, v_i^*(t)\,\rangle_V \,\mathrm{d}t.$$

Plugging in the definitions for u_{im} and u_i, the last identity becomes

$$\limsup_{m\to\infty} \left\langle\, u_m\,,\, A^i(u_m)\,\right\rangle_V = \langle\, u\,,\, v^*\,\rangle_V\,.$$

Setting with given $\widetilde{v} \in M$

$$v(t) = \left\{\begin{array}{l} \widetilde{v} \text{ if } (i-1)h < t \le ih \\ \bar{u} \text{ elsewhere} \end{array}\right\}$$

one obtains similarly from (10.2)

$$\left\langle\, u - \widetilde{v}\,,\, A^i(u) - v^*\,\right\rangle_V \le 0.$$

Thus the conclusions of the continuity condition 5.5(2), see 4.1, is satisfied. Therefore the continuity condition 5.5(2) is fulfilled.

It remains to show 5.5(3). From the coerciveness 6.2(3)

$$\langle\, u - \bar{u}\,,\, A(t,u)\,\rangle_V \ge c_0\|u-\bar{u}\|_V^2 - C_0 B_{\bar{u}}(u) - G_0(t) \tag{10.3}$$

for all $t > 0$ and $u \in M$ one gets, since $\bar{u}$ does not depend on time, the same estimate for the operator A_h,

$$\langle\, u - \bar{u}\,,\, A_h(t,u)\,\rangle_V \ge c_0\|u-\bar{u}\|_V^2 - C_0 B_{\bar{u}}(u) - G_{0h}(t), \tag{10.4}$$

where

$$G_{0h}(t) := \frac{1}{h}\int_{(i-1)h}^{ih} G_0(s)\,\mathrm{d}s \qquad \text{for } (i-1)h < t \le ih$$

in other notation

$$\left\langle\, u - \bar{u}\,,\, A^i(u)\,\right\rangle_V \ge c_0\|u-\bar{u}\|_V^2 - C_0 B_{\bar{u}}(u) - G_{0h}(t^i). \tag{10.5}$$

Since by 3.7

$$\Big(\, u - \bar{u}\,,\, b(u) - b(u^{(i-1)})\,\Big)_H \ge B_{\bar{u}}(u) - B_{\bar{u}}(u^{(i-1)})$$

we obtain

$$\begin{aligned} &\Big(\, u - \bar{u}\,,\, \lambda(b(u) - b(u^{(i-1)}))\,\Big)_H + \left\langle\, u - \bar{u}\,,\, A^i(u)\,\right\rangle_V \\ &\ge c_0\|u-\bar{u}\|_V^p + \lambda(B_{\bar{u}}(u) - B_{\bar{u}}(u^{(i-1)})) - C_0 B_{\bar{u}}(u) - G_{0h}(t^i) \\ &\ge c_0\|u-\bar{u}\|_V^p - \lambda B_{\bar{u}}(u^{(i-1)}) - G_{0h}(t^i) \end{aligned}$$

if $\lambda \ge C_0$, therefore, since $p > 1$,

$$\frac{\big(\, u - \bar{u}\,,\, \lambda(b(u) - b(u^{(i-1)}))\,\big)_H + \left\langle\, u - \bar{u}\,,\, A^i(u)\,\right\rangle_V}{\|u-\bar{u}\|_V} \to \infty$$

as $\|u\|_V \to \infty$. This shows coercivity 5.5(3) with $\bar{u}^i = \bar{u}$. Since all assumptions are fulfilled 5.5 is applicable.

10.2. We prove: There is $u \in L^p([0,T];V)$ with $u(t) \in M$ for almost all t, such that for a subsequence $h \to 0$ the following convergence holds:

$$\begin{aligned}
&u_h \to u \text{ weakly in } L^p([0,T];V),\\
&\mathcal{A}(u_h) = A(\bullet, u_h) \to u^* \text{ weakly}^* \text{ in } L^{p*}([0,T];V^*),\\
&\{B_{\bar{u}}(u_h)\,;\ 0 < h < h_0\} \text{ bounded in } L^\infty([0,T]),\\
&b(u_h) \to b(u) \text{ strongly in } L^1([0,T];H).
\end{aligned}$$

Here the approximative functions u_h are defined only in $[0,T_h]$, where T_h is the largest multiple of h less or equal T. It is irrelevant, how u_h is defined in $]T_h,T]$, the main thing is that it stays bounded, for example we define it by $\bar{u}$. Then the statements hold on the interval $[0,T]$.

We take a solution of 10.1, set $w_h^*(t) := A_h(t, u_h(t))$ in (7.2), and use the a-priori estimate 7.1. We obtain

$$B_{\bar{u}}(u_h(\bar{t})) + \int_0^{\bar{t}} \langle\, u_h(t) - \bar{u}\,,\, w_h^*(t)\,\rangle_V \;\mathrm{d}t \le B_{\bar{u}}(u_0).$$

The coerciveness (10.4) leads to

$$\begin{aligned}
&\int_0^{\bar{t}} \langle\, u_h(t) - \bar{u}\,,\, w_h^*(t)\,\rangle_V \;\mathrm{d}t\\
&\ge c_0 \int_0^{\bar{t}} \|u_h(t) - \bar{u}\|_V^p \,\mathrm{d}t - C_0 \int_0^{\bar{t}} B_{\bar{u}}(u_h(t))\,\mathrm{d}t - \int_0^{\bar{t}} G_{0h}(t)\,\mathrm{d}t.
\end{aligned}$$

This implies

$$\begin{aligned}
&B_{\bar{u}}(u_h(\bar{t})) + c_0 \int_0^{\bar{t}} \|u_h(t) - \bar{u}\|_V^p \,\mathrm{d}t\\
&\le B_{\bar{u}}(u_0) + C_0 \int_0^{\bar{t}} B_{\bar{u}}(u_h(t))\,\mathrm{d}t + \int_0^{\bar{t}} G_{0h}(t)\,\mathrm{d}t.
\end{aligned} \tag{10.6}$$

Since $G_0 \in L^1([0,T])$ the last term on the right is bounded uniformly in h. Then a Gronwall argumentation on the inequality (10.6) gives the boundedness of the sets

$$\{B_{\bar{u}}(u_h)\,;\ 0 < h < h_0\} \text{ in } L^\infty([0,T_h]).$$

Using this one gets from (10.6) the "parabolic" estimate

$$\operatorname*{ess\,sup}_{t\in[0,T_h]} B_{\bar{u}}(u_h(t)) + \int_0^{T_h} \|u_h\|_V^p \le C, \tag{10.7}$$

where C depends only on u_0, $\bar{u}$, G_0, $B_{\bar{u}}$, c_0, T, and obvious quantities like p, n, V, H. In particular, C is independent of h. Hence for a subsequence $h \to 0$ there exists the weak limit $u_h \to u$ in $L^p([0,T];V)$. Since $u_h(t) \in M$ for almost all t, it follows that $u(t) \in M$ for almost all t.

Equation (10.7) says that (define for example $u_h(t) := \bar{u}$ for $T_h < t < T$) the set $\{u_h\,;\ 0 < h < h_0\}$ satisfies the required boundedness assumption in 6.2(1). It follows by the boundedness condition 6.2(1) that the set $\{\mathcal{A}(u_h)\,;\ 0 < h < h_0\}$ is bounded in $L^{p^*}([0,T];V^*)$. Therefore there is a subsequence $h \to 0$ so (all subsequent subsequences have to be chosen as subsequence of the previous subsequence), that the weak* limit $\mathcal{A}(u_h) \to u^*$ exists in $L^{p^*}([0,T];V^*)$.

What is missing is the strong convergence of $b(u_h)$ as $h \to 0$. To derive this we go into the second estimate in section 7, that is 7.2,

$$E_{\psi^*}(b(u_h(t+s)), b(u_h(t)), u_h(t)) \leq s \cdot \frac{1}{j} \sum_{i=1}^{j} \langle\, u_h(t) - u_h(t+ih)\,,\, w_h^*(t+ih)\,\rangle_V\,.$$

This gives

$$\begin{aligned}
&\int_0^{T-s} E_{\psi^*}(b(u_h(t+s)), b(u_h(t)), u_h(t))\,\mathrm{d}t \\
&\leq s \cdot \frac{1}{j} \sum_{i=1}^{j} \int_0^{T-s} \langle\, u_h(t) - u_h(t+ih)\,,\, w_h^*(t+ih)\,\rangle_V\, \mathrm{d}t \\
&\leq 2s \cdot \|u_h\|_{L^p([0,T];V)} \|w_h^*\|_{L^{p^*}([0,T];V^*)}.
\end{aligned}$$

By the estimates proved so far, that is the estimates for u_h and $w_h^* = \mathcal{A}(u_h)$ the right side is bounded by a constant times s. Thus we obtain

$$\int_0^{T-s} E_{\psi^*}(b(u_h(t+s)), b(u_h(t)), u_h(t))\,\mathrm{d}t \leq C \cdot s.$$

This estimate is fulfilled for all $s > 0$, not only for multiple of h, see lemma 9.1. Then the compactness theorem 9.3 implies that $\{b(u_h)\,;\ 0 < h < h_0\}$ is precompact in $L^1([0,T];H)$. Hence there is a subsequence $h \to 0$ so (all subsequent subsequences have to be chosen as subsequence of the previous subsequence), that $b(u_h)$ strongly in $L^1([0,T];H)$ to a limit b^*. But then by 9.4 it is $b(u) = b^*$, so that $b(u_h) \to b(u)$ strongly in $L^1([0,T];H)$.

10.3. Proof of 6.2.

For the sequence u_h we have the following time discrete inequality

$$\Phi_{\bar{u}}^h(u_h, v)(\widetilde{t}) + \int_0^{\widetilde{t}} \langle\, u_h(t) - v(t)\,,\, A_h(t, u_h(t)\,\rangle_V\, \mathrm{d}t \leq 0$$

for all $\widetilde{t} > 0$ and all $t \mapsto v(t) \in M$ with $v \in L^1([0,T];H)$, where

$$\Phi_{\bar{u}}^h(u_h, v)(\widetilde{t}) := \int_0^{\widetilde{t}} \big(\, u_h(t) - v(t)\,,\, \partial_t^{-h} b(u_h(t))\,\big)_H\, \mathrm{d}t$$

and $u_h(t) := u_0$ for $t < 0$. We now define for a given $\bar{t}$ a $\widetilde{t} = \bar{t}_h$ as a multiple of h with

$$\bar{t}_h - h < \bar{t} \leq \bar{t}_h.$$

We use this $\bar{t}_h$ and obtain

$$\Phi_{\bar{u}}^h(u_h, v)(\bar{t}_h) + \int_0^{\bar{t}_h} \langle\, u_h(t) - v(t)\,,\, A_h(t, u_h(t)\,\rangle_V\, \mathrm{d}t \leq 0.$$

Since u_h is a step function, i.e. $u_h(t) = u_h(ih)$ for $(i-1)h < t \leq ih$, see (5.10), we compute for $\bar{t}_h = \bar{i}h$

$$\begin{aligned}
&\int_0^{\bar{t}_h} \langle\, u_h(t)\,,\, A_h(t, u_h(t)\,\rangle_V\, \mathrm{d}t = \sum_{i=1}^{\bar{i}} h\, \langle\, u_h(ih)\,,\, A_h(ih, u_h(ih)\,\rangle_V \\
&= \sum_{i=1}^{\bar{i}} \int_{(i-1)h}^{ih} \langle\, u_h(ih)\,,\, A(t, u_h(ih)\,\rangle_V\, \mathrm{d}t = \int_0^{\bar{t}_h} \langle\, u_h(t)\,,\, A(t, u_h(t)\,\rangle_V\, \mathrm{d}t,
\end{aligned}$$

that is

$$\begin{aligned}&\Phi^h_{\bar u}(u_h,v)(\bar t_h)+\int_0^{\bar t_h}\langle\, u_h(t)\,,\,A(t,u_h(t)\,\rangle_V\ \mathrm{d}t\\&\le\int_0^{\bar t_h}\langle\, v(t)\,,\,A_h(t,u_h(t)\,\rangle_V\ \mathrm{d}t.\end{aligned}\tag{10.8}$$

The parabolic part we compute (this looks similar to 8.2)

$$\begin{aligned}\Phi^h_{\bar u}(u_h,v)(\bar t_h)\ &=\int_0^{\bar t_h}\big(\,u_h(t)-v(t)\,,\,\partial_t^{-h}b(u_h(t))\,\big)_H\ \mathrm{d}t\\&=\int_0^{\bar t_h}\big(\,u_h(t)-\bar u\,,\,\partial_t^{-h}b(u_h(t))\,\big)_H\ \mathrm{d}t\\&\quad+\int_0^{\bar t_h}\big(\,\bar u-v(t)\,,\,\partial_t^{-h}(b(u_h(t))-b(u_0))\,\big)_H\ \mathrm{d}t.\end{aligned}$$

With usage of lemma 3.7 this is

$$\begin{aligned}&\ge\int_0^{\bar t_h}\partial_t^{-h}B_{\bar u}(u_h(t))\,\mathrm{d}t\ +\frac1h\int_0^{\bar t_h}(\,\bar u-v(t)\,,\,b(u_h(t))-b(u_0)\,)_H\ \mathrm{d}t\\&\qquad-\frac1h\int_{-h}^{\bar t_h-h}(\,\bar u-v(t+h)\,,\,b(u_h(t))-b(u_0)\,)_H\ \mathrm{d}t\\&=\int_0^{\bar t_h}\partial_t^{-h}B_{\bar u}(u_h(t))\,\mathrm{d}t\ +\frac1h\int_{-\infty}^{\bar t_h}(\,\bar u-v(t)\,,\,b(u_h(t))-b(u_0)\,)_H\ \mathrm{d}t\\&\qquad-\frac1h\int_{-\infty}^{\bar t_h-h}(\,\bar u-v(t+h)\,,\,b(u_h(t))-b(u_0)\,)_H\ \mathrm{d}t\\&=\ \frac1h\int_{\bar t_h-h}^{\bar t_h}B_{\bar u}(u_h(t))\,\mathrm{d}t-B_{\bar u}(u_0)+\frac1h\int_{\bar t_h-h}^{\bar t_h}(\,\bar u-v(t)\,,\,b(u_h(t))-b(u_0)\,)_H\ \mathrm{d}t\\&\quad-\int_0^{\bar t_h-h}\big(\,\partial_t^{+h}(\bar u-v(t))\,,\,b(u_h(t))-b(u_0)\,\big)_H\ \mathrm{d}t\\&=\ B_{\bar u}(u_h(\bar t_h))-B_{\bar u}(u_0)+(\,\bar u-v_h(\bar t_h)\,,\,b(u_h(\bar t_h))-b(u_0)\,)_H\\&\quad-\int_0^{\bar t_h-h}\big(\,\partial_t^{+h}(\bar u-v(t))\,,\,b(u_h(t))-b(u_0)\,\big)_H\ \mathrm{d}t.\end{aligned}$$

Here we have used the choice of $\bar t_h$ above, and

$$v_h(t):=\frac1h\int_{(i-1)h}^{ih}v(s)\,\mathrm{d}s\qquad\text{for }(i-1)h<t\le ih.$$

Altogether we derived

$$\begin{aligned}\Phi^h_{\bar u}(u_h,v)(\bar t_h)\ &\ge\ B_{\bar u}(u_h(\bar t_h))-B_{\bar u}(u_0)+(\,\bar u-v_h(\bar t_h)\,,\,b(u_h(\bar t_h))-b(u_0)\,)_H\\&\quad-\int_0^{\bar t_h-h}\big(\,\partial_t^{+h}(\bar u-v(t))\,,\,b(u_h(t))-b(u_0)\,\big)_H\ \mathrm{d}t\\&=\ B_{\bar u}(u_h(\bar t))-B_{\bar u}(u_0)+(\,\bar u-v_h(\bar t)\,,\,b(u_h(\bar t))-b(u_0)\,)_H\\&\quad-\int_0^{\bar t_h-h}\big(\,\partial_t^{+h}(\bar u-v(t))\,,\,b(u_h(t))-b(u_0)\,\big)_H\ \mathrm{d}t.\end{aligned}$$

Now we know $b(u_h)\to b(u)$ strongly in $L^1([0,T];H)$, and therefore for a subsequence that $b(u_h)\to b(u)$ in H almost everywhere. Since $B_{\bar u}(u_h(\bar t))=\psi^*_{\bar u}(b(u_h(\bar t)))+\psi_{\bar u}(0)$, see definition 3.6, and if we now assume that $\partial_t v \in L^1([0,T];H)$, we conclude that for $h \to 0$ for almost all $\bar t$

$$\begin{aligned}
&B_{\bar u}(u_h(\bar t)) - B_{\bar u}(u_0) + (\, \bar u - v_h(\bar t)\,,\, b(u_h(\bar t)) - b(u_0)\,)_H \\
&- \int_0^{\bar t_h - h} \big(\, \partial_t^{+h}(\bar u - v(t))\,,\, b(u_h(t)) - b(u_0)\,\big)_H \,\mathrm{d}t \\
\longrightarrow\ & B_{\bar u}(u(\bar t)) - B_{\bar u}(u_0) + (\, \bar u - v(\bar t)\,,\, b(u(\bar t)) - b(u_0)\,)_H \\
&\quad - \int_0^{\bar t} (\, \partial_t(\bar u - v(t))\,,\, b(u(t)) - b(u_0)\,)_H \,\mathrm{d}t \\
=\ & \Phi_{\bar u}(u,v)(\bar t),
\end{aligned}$$

see definition 8.1. Therefore we have proved that for almost all $\bar t$

$$\liminf_{h\to 0} \Phi_{\bar u}^h(u_h,v)(\bar t_h) \geq \Phi_{\bar u}(u,v)(\bar t). \tag{10.9}$$

Since equation (10.8) reads

$$\begin{aligned}
&\Phi_{\bar u}^h(u_h,v)(\bar t_h) + \int_0^{\bar t_h} \langle\, u_h(t)\,,\, A(t,u_h(t)\,\rangle_V \,\mathrm{d}t \\
&\leq \int_0^{\bar t_h} \langle\, v(t)\,,\, A_h(t,u_h(t)\,\rangle_V \,\mathrm{d}t = \int_0^{\bar t_h} \langle\, v_h(t)\,,\, A(t,u_h(t)\,\rangle_V \,\mathrm{d}t \\
&\longrightarrow \int_0^{\bar t} \langle\, v(t)\,,\, u^*(t)\,\rangle_V \,\mathrm{d}t
\end{aligned}$$

for $h \to 0$ (the sequence has to be chosen as the above subsequence), we obtain

$$\begin{aligned}
&\liminf_{h\to 0} \Phi_{\bar u}^h(u_h,v)(\bar t_h) + \limsup_{h\to 0} \int_0^{\bar t_h} \langle\, u_h(t)\,,\, A(t,u_h(t)\,\rangle_V \,\mathrm{d}t \\
&\leq \int_0^{\bar t} \langle\, v(t)\,,\, u^*(t)\,\rangle_V \,\mathrm{d}t,
\end{aligned}$$

that is

$$\Phi_{\bar u}(u,v)(\bar t) + \limsup_{h\to 0} \int_0^{\bar t_h} \langle\, u_h(t)\,,\, A(t,u_h(t)\,\rangle_V \,\mathrm{d}t \leq \int_0^{\bar t} \langle\, v(t)\,,\, u^*(t)\,\rangle_V \,\mathrm{d}t.$$

Since u_h and $\mathcal{A}(u_h)$ are bounded, the contribution

$$\int_{\bar t}^{\bar t_h} \langle\, u_h(t)\,,\, A(t,u_h(t)\,\rangle_V \,\mathrm{d}t \to 0$$

as $h \to 0$, and therefore

$$\begin{aligned}
&\Phi_{\bar u}(u,v)(\bar t) + \limsup_{h\to 0} \int_0^{\bar t} \langle\, u_h(t)\,,\, A(t,u_h(t)\,\rangle_V \,\mathrm{d}t \\
&\leq \int_0^{\bar t} \langle\, v(t)\,,\, u^*(t)\,\rangle_V \,\mathrm{d}t.
\end{aligned} \tag{10.10}$$

Now we come to the missing term for the sequence. We set $v = u_\delta$ with u_δ from 8.3 (it is $\partial_t v \in L^1([0,T];H)$) and obtain

$$\begin{aligned}&\Phi_{\bar u}(u,u_\delta)(\bar t) + \limsup_{h\to 0}\int_0^{\bar t} \langle\, u_h(t)\,,\, A(t,u_h(t)\,\rangle_V \,\mathrm{d}t \\ &\le \int_0^{\bar t} \langle\, u_\delta(t)\,,\, u^*(t)\,\rangle_V \,\mathrm{d}t \longrightarrow \int_0^{\bar t} \langle\, u(t)\,,\, u^*(t)\,\rangle_V \,\mathrm{d}t\end{aligned}$$

as $\delta \to 0$. Since $\Phi_{\bar u}(u,u_\delta)(\bar t)$ in the limit $\delta \to 0$ is nonnegative as shown in 8.3, we arrive at

$$\limsup_{h\to 0}\int_0^{\bar t} \langle\, u_h(t)\,,\, A(t,u_h(t)\,\rangle_V \,\mathrm{d}t \le \int_0^{\bar t} \langle\, u(t)\,,\, u^*(t)\,\rangle_V \,\mathrm{d}t$$

This is the last property in the assumption of 6.2(2). (In reality one has to use the time interval $[0,T]$ instead of $[0,\bar t]$, that is one has to use for example the functions

$$\widetilde u_h(t) := \left\{\begin{array}{l} u_h(t) \text{ for } t \le \bar t \\ \bar u \text{ for } t > \bar t \end{array}\right\}, \quad \widetilde u^*(t) := \left\{\begin{array}{l} u^*(t) \text{ for } t \le \bar t \\ A(t,\bar u) \text{ for } t > \bar t \end{array}\right\},$$

and $\widetilde u$ defined with respect to u as $\widetilde u_h$ with respect to u_h. Note, that then

$$\limsup_{h\to 0}\int_0^{T} \langle\, \widetilde u_h(t)\,,\, A(t,\widetilde u_h(t)\,\rangle_V \,\mathrm{d}t \le \int_0^{T} \langle\, \widetilde u(t)\,,\, \widetilde u^*(t)\,\rangle_V \,\mathrm{d}t$$

holds.) Therefore we can use the conclusions of 6.2(2), that is for all $v \in L^p([0,T];V)$ we infer

$$\begin{aligned}\int_0^{\bar t} \langle\, u(t)-v(t)\,,\, A(t,u(t))-u^*(t)\,\rangle_V \,\mathrm{d}t \le 0, \quad \text{and} \\ \limsup_{h\to 0}\int_0^{\bar t} \langle\, u_h(t)\,,\, A(t,u_h(t))\,\rangle_V \,\mathrm{d}t = \int_0^{\bar t} \langle\, u(t)\,,\, u^*(t)\,\rangle_V \,\mathrm{d}t.\end{aligned} \tag{10.11}$$

Plugging the identity of (10.11) in the above equation (10.10), one gets

$$\Phi_{\bar u}(u,v)(\bar t) + \int_0^{\bar t} \langle\, u(t)-v(t)\,,\, u^*(t)\,\rangle_V \,\mathrm{d}t \le 0$$

and therefore, using the inequality in (10.11), one obtains

$$\Phi_{\bar u}(u,v)(\bar t) + \int_0^{\bar t} \langle\, u(t)-v(t)\,,\, A(t,u(t)\,\rangle_V \,\mathrm{d}t \le 0.$$

This is the assertion.

11. **Examples.** In the following we present some concrete examples. First there are second order boundary value problems, where $\Omega \subset \mathbb{R}^n$ is a bounded Lipschitz domain and

$$H = L^2(\Omega;\mathbb{R}^N) \quad \text{and} \quad V = W^{1,2}(\Omega;\mathbb{R}^N).$$

Here the usual isomorphism between $L^2(]t_0,t_1[\times\Omega;\mathbb{R}^N)$ and $L^2(]t_0,t_1[;H)$ is used. Therefore we identify $u(t,x) = u(t)(x)$. For simplicity we take $N = 1$, and we let $L^2(\Omega) = L^2(\Omega;\mathbb{R})$ and $W^{1,2}(\Omega) = W^{1,2}(\Omega;\mathbb{R})$. Under these assumptions the following standard example is true, where we do not choose the most general form.

11.1 Second order problem. We let $\Omega \subset \mathbb{R}^n$ as above and take a closed (may be empty) set $\Gamma \subset \partial\Omega$. On the time interval $[0,T]$ we consider the elliptic-parabolic boundary value problem

$$\begin{aligned}
&\partial_t \beta(x,u(t,x)) - \mathrm{div} a(x,u(t,x),\nabla u(t,x)) = f(t,x) \quad \text{for } (t,x) \in]0,T[\times\Omega,\\
&u(t,x) = u_1(x) \quad \text{for } (t,x) \in]0,T[\times\Gamma,\\
&a(u(t,x),\nabla u(t,x)) \bullet \nu(x) = 0 \quad \text{for } (t,x) \in]0,T[\times(\partial\Omega\setminus\Gamma),\\
&\beta(x,u(0,x)) = \beta(x,u_0(x)) \quad \text{for } x \in \Omega.
\end{aligned}$$

Here functions $u_1 \in W^{1,2}(\Omega)$ and $u_0 \in L^2(\Omega)$ and a right side $f \in L^2([0,T]\times\Omega)$ are given. Moreover $\beta : \Omega\times\mathbb{R} \to \mathbb{R}$ and $a : \Omega\times\mathbb{R}\times\mathbb{R}^n \to \mathbb{R}^n$ are Carathéodory functions, that is measurable in the first argument and continuous in the other arguments. We assume the following monotonicity and growth conditions

$$\begin{aligned}
&(\beta(x,z_1) - \beta(x,z_2))\cdot(z_1 - z_2) \geq 0,\\
&(a(x,z,p_1) - a(x,z,p_2)) \bullet (p_1 - p_2) \geq c\cdot|p_1 - p_2|^2,\\
&|\,\beta(x,z)\,| \leq C\cdot(1+|\,z\,|), \quad |\,a(x,z,p)\,| \leq C\cdot(1+|\,p\,|),
\end{aligned}$$

(1) Then, with a correct choice of $\mathcal{M}$, $\mathcal{A}$, and b, this example is of the general type. The condition 6.2(1) of theorem 6.3 is satisfied.

(2) If in addition $a(x,z_1,p) = a(x,z_2,p)$ whenever $\beta(x,z_1) = \beta(x,z_2)$, then the continuity condition 6.2(2) is satisfied.

(3) If in addition $\mathrm{H}^{n-1}(\Gamma) > 0$ or $|\beta(x,z) - \beta(x,0)| \geq c_1|z|$ for $z \in \mathbb{R}$, then the coercivity 6.2(3) is satisfied.

Proof (1). Since we have the boundary condition $u(t,\bullet) = u_1$ in the formulation of the problem, we set a time independent constraint

$$M := \{u \in V\,;\ u = u_1 \text{ almost everywhere on } \Gamma\}$$

with $V := W^{1,2}(\Omega)$. The map $b : H \to H$, $H = L^2(\Omega)$, is given by

$$b(u)(x) := \beta(x,u(x)) \quad \text{for } x \in \Omega \text{ and } u \in H.$$

If $\varphi : \Omega\times\mathbb{R} \to \mathbb{R}$ is a convex map in the second variable with $\frac{\mathrm{d}}{\mathrm{d}z}\varphi(x,z) = \beta(x,z)$ and $\varphi(x,0) = 0$, then $|\,\varphi(x,z)\,| \leq C(1+|\,z\,|^2)$ by the growth condition on β, and we are able to define $\psi : H \to \mathbb{R}$ by

$$\psi(u) := \int_\Omega \varphi(x,u(x))\,\mathrm{d}x \quad \text{for } u \in H.$$

It follows that $b = \nabla\psi$, that is

$$\begin{aligned}
&(\,\nabla\psi(u)\,,\,v\,)_H = \mathrm{D}\psi(u)(v) = \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\psi(u+\varepsilon v)\Big|_{\varepsilon=0}\\
&= \int_\Omega \frac{\mathrm{d}}{\mathrm{d}\varepsilon}\varphi(x,u(x)+\varepsilon v(x))\Big|_{\varepsilon=0}\,\mathrm{d}x = \int_\Omega \frac{\mathrm{d}}{\mathrm{d}z}\varphi(x,u(x))v(x)\,\mathrm{d}x = (\,b(u)\,,\,v\,)_H\,.
\end{aligned}$$

For the elliptic part we set $A(t,\bullet) : M \to V^*$ by

$$\langle\, v\,,\,A(t,u)\,\rangle_V := \int_\Omega \Big(\nabla v(x) \bullet a(x,u(x),\nabla u(x)) - v(x)f(t,x)\Big)\,\mathrm{d}x \quad \text{for } u,v \in V,$$

where $a(\bullet,u,\nabla u) \in L^2(\Omega)$ by the growth condition. The map

$$\mathcal{A} : L^2([0,T];V) \to L^2([0,T];V^*)$$

is given by $\mathcal{A}(u)(t) := A(t,u)$, the set $\mathcal{M}$ given in (6.2) is a subset of $L^2([0,T];V)$.

To obtain the abstract equation (6.8) we have to identify $u(t,x) = u(t)(x)$ with $u(t) \in V$. The test functions ζ have compact support in $[0,T[\times(\overline{\Omega}\setminus\Gamma)$, which is equivalent to $\zeta \in C^\infty([0,T]\times\overline{\Omega})$ with $\zeta = 0$ in a neighbourhood of $\{T\}\times\overline{\Omega}$ and $[0,T]\times\Gamma$. Therefore in (6.8) it is assumed that $\zeta(t,x) = \xi(t)(x)$, and now $\xi \in C_0^\infty([0,T[;M_1)$, where $M_1 := M - u_1$ is a subspace and $\xi(t) = 0$ for t close to the final time T. □

Proof (2). Let $u_m, u \in L^2([0,T];W^{1,2}(\Omega))$ and $v^* \in L^2([0,T];W^{1,2}(\Omega)^*)$ as in the continuity condition 6.1. By the ellipticity condition on a

$$\begin{aligned}0 \le\ & \int_0^T \int_\Omega (\nabla u_m - \nabla v)\bullet\big(a(x,u_m,\nabla u_m) - a(x,u_m,\nabla v)\big)\,\mathrm{d}x\,\mathrm{d}t\\ =\ & \int_0^T \langle\, u_m\,,\, A(t,u_m)\,\rangle_{W^{1,2}(\Omega)}\,\mathrm{d}t - \int_0^T \langle\, v\,,\, A(t,u_m)\,\rangle_{W^{1,2}(\Omega)}\,\mathrm{d}t\\ & - \int_0^T\int_\Omega (\nabla u_m - \nabla v)\bullet\big(a(x,u_m,\nabla v) - f(t,x)\big)\,\mathrm{d}x\,\mathrm{d}t.\end{aligned}$$

Since by assumption we have a dependence $a(x,z,p) = \widetilde{a}(x,\beta(x,z),p)$, where also $\widetilde{a}$ is a Carathéodory function, and since $b(u_m) \to b(u)$ strongly in $L^1([0,T];H)$, we conclude for a subsequence $m \to \infty$ that $\beta(x,u_m(t,x)) \to \beta(x,u(t,x))$ for almost all (t,x). Hence for a subsequence

$$a(x,u_m,\nabla v) \to a(x,u,\nabla v) \quad \text{strongly in } L^2([0,T];L^2(\Omega)).$$

Therefore we get

$$\begin{aligned}0 \le \limsup_{m\to\infty} & \int_0^T \langle\, u_m\,,\, A(t,u_m)\,\rangle_{W^{1,2}(\Omega)}\,\mathrm{d}t - \int_0^T \langle\, v\,,\, v^*\,\rangle_{W^{1,2}(\Omega)}\,\mathrm{d}t\\ - \int_0^T\int_\Omega & (\nabla u - \nabla v)\bullet\big(a(x,u,\nabla v) - f(t,x)\big)\,\mathrm{d}x\,\mathrm{d}t.\end{aligned}$$

For $v = u$ one obtains the second conclusion of continuity condition 6.1. Then

$$\begin{aligned}0 \le & \int_0^T \langle\, u - v\,,\, v^*\,\rangle_{W^{1,2}(\Omega)}\,\mathrm{d}t\\ - & \int_0^T\int_\Omega (\nabla u - \nabla v)\bullet\big(a(x,u,\nabla v) - f(t,x)\big)\,\mathrm{d}x\,\mathrm{d}t.\end{aligned}$$

Now replace v by $v_\varepsilon := u + \varepsilon(v-u)$ and obtain for $\varepsilon \to 0$

$$\begin{aligned}0 \le & \int_0^T \langle\, u - v\,,\, v^*\,\rangle_{W^{1,2}(\Omega)}\,\mathrm{d}t\\ - & \int_0^T\int_\Omega (\nabla u - \nabla v)\bullet\big(a(x,u,\nabla u) - f(t,x)\big)\,\mathrm{d}x\,\mathrm{d}t\\ = & \int_0^T \langle\, u - v\,,\, v^* - A(t,u)\,\rangle_{W^{1,2}(\Omega)}\,\mathrm{d}t,\end{aligned}$$

which is the first conclusion of the continuity condition 6.1. □

Proof (3). To prove 6.2(3) we compute with $\bar{u} = u_1$

$$\begin{aligned}
&\langle\, u - \bar{u}\,,\; A(t,u) - A(t,\bar{u})\,\rangle_{W^{1,2}(\Omega)} \\
&= \int_\Omega (\nabla u - \nabla\bar{u})\bullet\big(a(\bullet,u,\nabla u) - a(\bullet,\bar{u},\nabla\bar{u})\big)\,\mathrm{dL}^n \\
&= \int_\Omega (\nabla u - \nabla\bar{u})\bullet\big(a(\bullet,u,\nabla u) - a(\bullet,u,\nabla\bar{u})\big)\,\mathrm{dL}^n \\
&\quad + \int_\Omega (\nabla u - \nabla\bar{u})\bullet\big(a(\bullet,u,\nabla\bar{u}) - a(\bullet,\bar{u},\nabla\bar{u})\big)\,\mathrm{dL}^n \\
&\geq c\|\nabla(u-\bar{u})\|^2_{L^2(\Omega)} - 2C\|\nabla(u-\bar{u})\|_{L^2(\Omega)}\cdot\|1+|\nabla\bar{u}|\|_{L^2(\Omega)},
\end{aligned}$$

hence

$$\begin{aligned}
&\langle\, u - \bar{u}\,,\; A(t,u)\,\rangle_{W^{1,2}(\Omega)} \\
\geq\;& c\|\nabla(u-\bar{u})\|^2_{L^2(\Omega)} - \|\nabla(u-\bar{u})\|_{W^{1,2}(\Omega)}\cdot\|A(t,\bar{u})\|_{W^{1,2}(\Omega)^*} \\
&-2C\|\nabla(u-\bar{u})\|_{L^2(\Omega)}\cdot\|1+|\nabla\bar{u}|\|_{L^2(\Omega)}.
\end{aligned}$$

This gives the desired estimate, since in case $\mathrm{H}^{n-1}(\Gamma) > 0$ the Poincaré inequality can be used, and otherwise $\|u-\bar{u}\|^2_{L^2(\Omega)} \leq C_1(1+B_{\bar{u}}(u))$ by assumption on β. □

The fact, that one proves that $\mathcal{A}$ maps into $L^2([0,T];V^*)$, usually gives the boundedness condition 6.2(1) as a byproduct. We mention, that also the case, that $a(x,u(t,x),\nabla u(t,x))$ has a controlled unbounded term in $u(t,x)$, can be treated. Similar arguments apply to a right side $f(t,x,u(t,x))$.

Also the case of systems of elliptic equations is covered, that is the case $N > 1$. Then different components of this system may satisfy different boundary conditions.

In general, there is not, as usual in parabolic equations, a uniqueness theorem as consequence of the existence theory. Uniqueness theorems are available only under additional assumptions (see e.g. [5]). The reason for nonuniqueness is, that the theory works for elliptic-parabolic problems, and it is well known that for elliptic problems in general the solution is not unique.

11.2 Non-uniqueness. Let $\beta : \mathbb{R} \to \mathbb{R}$ be given and consider the problem

$$\begin{aligned}
&\partial_t\beta(u) - \operatorname{div}(a\nabla u) = f(u) \quad \text{ in }]0,\infty[\times\Omega, \\
&\partial_\nu u = 0 \quad \text{ on }]0,\infty[\times\partial\Omega, \\
&u(0,x) = u_0(x) \quad \text{ for } x \in \Omega,
\end{aligned}$$

with a strictly positive function $a \in L^\infty(]0,\infty[\times\Omega)$ and a continuous sublinear function $f : \mathbb{R} \to \mathbb{R}$. We assume that β is continuous, sublinear, and weakly monotone non-decreasing. Under these assumptions there is a weak solution for each initial datum $u_0 \in L^\infty(\Omega)$. Let us take the example

$$\beta(u) = \begin{cases} \beta_- + c_-(u-u_-) & \text{for } u \leq u_-, \\ \beta_+ + c_+(u-u_+) & \text{for } u \geq u_+, \\ \text{monotone and continuous for } u_- \leq u \leq u_+, \end{cases}$$

where $c_- \geq 0$, $c_+ \geq 0$, $u_- < u_+$, and $\beta_- \leq \beta_+$. Then the following is true:

(1) If $\beta_- = \beta_+$ and $f = 0$, then for $u_- \leq u_0 \leq u_+$ each function $u = \text{const.}$ is a solution, if this constant lies in $]u_-,u_+[$.

(2) If $u_- \le u_0 \le u_+$, then each function can be obtained as limit of unique solutions of problems with strictly monotone β.

Proof (2). Let $\bar{u}_0 = \text{const.}$ with $u_- \le \bar{u}_0 \le u_+$. For $\delta > 0$ and $\varepsilon > 0$ let us consider the problem

$$\begin{aligned}
&\partial_t \beta_\delta(u) - \operatorname{div}(a\nabla u) = -\varepsilon(u - \bar{u}_0) \quad \text{ in }]0,\infty[\times\Omega,\\
&\partial_\nu u = 0 \quad \text{ on }]0,\infty[\times\partial\Omega,\\
&u(0,x) = u_0(x) \quad \text{ for } x \in \Omega,
\end{aligned}$$

where

$$\beta_\delta(u) := \beta(u) + \delta u\,.$$

Let $u = u_{\delta\varepsilon}$ be the unique weak solution of this approximate problem. By the maximum principle $u_- \le u \le u_+$, so that the differential equation is

$$\delta\partial_t u - \operatorname{div}(a\nabla u) + \varepsilon(u - \bar{u}_0) = 0,$$

and from there the energy estimate

$$\begin{aligned}
&\frac{\delta}{2}\int_\Omega (u(t,x) - \bar{u}_0)^2\,\mathrm{d}x + \int_0^t \int_\Omega \Big(a(t,x)|\nabla(u(t,x) - \bar{u}_0)|^2 + \varepsilon(u(t,x) - \bar{u}_0)\Big)\,\mathrm{d}x\,\mathrm{d}t\\
&\le \frac{\delta}{2}\int_\Omega (u_0(x) - \bar{u}_0)^2\,\mathrm{d}x
\end{aligned}$$

holds. It follows immediately, if one chooses $\delta = \delta_\varepsilon$ small enough, that $u_{\delta_\varepsilon\varepsilon} \to \bar{u}_0$ when $\varepsilon \to 0$. □

The problem where $u_- = u_+$ and $\beta_- < \beta_+$ is not contained in this theorem, although the theory is capable to treat this case as a limit $u_+ - u_- \searrow 0$, but we do not discuss this here.

The standard case of parabolic equations is, that the solutions are continuous in time. This is not the case here, and it has again to do with the elliptic-parabolic character of the problem, see [5, Introduction].

11.3 Non-continuity. Consider the problem

$$\begin{aligned}
&\partial_t \beta(u) - \Delta u = 0 \quad \text{ in }]0,\infty[\times\Omega,\\
&u(t,x) = u_1(t) \quad \text{ for } t > 0,\ x \in \partial\Omega,\\
&u(0,x) = u_0(x) \quad \text{ for } x \in \Omega,
\end{aligned}$$

where $\Omega = \mathrm{B}_R(0)$ is a ball, $u_1 < 0$ continuous, and

$$\beta(u) = \max(u, 0).$$

Then the following holds:

(1) If the initial data are positive somewhere, there is a solution u, which has a jump in time.

(2) The solution in (1) is the limit of the unique solutions with $\beta = \beta_\varepsilon$,

$$\beta_\varepsilon(u) = \begin{cases} u & \text{for } u \ge 0,\\ \varepsilon u & \text{for } u \le 0, \end{cases}$$

as $\varepsilon \to 0$.

Proof (1). In the case $n = 1$ for suitable initial data and boundary data, a weak solution is

$$u(t,x) = \begin{cases} s(t) \cdot \eta\left(\dfrac{x}{s(t)}\xi\right) & \text{for } |x| \le s(t), \\ \xi\eta'(\xi) \cdot (|x| - s(t)) & \text{for } |x| \ge s(t), \end{cases}$$

which is a C^1-function until the jump happening at time $t = t_{crit}$, where

$$s(t) = 2\xi\sqrt{t_{crit} - t},$$

and a function constant in x

$$u(t,x) = u_1(t)$$

after the jump. Here the boundary data, we assume $R > s(0)$, are

$$u_1(t) = \begin{cases} \xi\eta'(\xi) \cdot (R - s(t)) \text{ before the jump } t \le t_{crit}, \\ \text{less than zero after the jump } t \ge t_{crit}. \end{cases}$$

Further

$$\eta(y) = 1 - \sum_{i=1}^{\infty} \frac{y^{2i}}{i!(2i-1)},$$

or equivalently,

$$\eta''(y) - 2y\eta'(y) + 2\eta(y) = 0,$$
$$\eta(0) = 1,\ \eta'(0) = 0,\ \xi \approx 0.92414 \text{ first positive zero of } \eta\ .$$

The boundary data u_1 can be continuous for all t. For $n > 1$ the procedure is similar. □

One sees in both examples 11.2 and 11.3 that the method in this paper is closed among all problems with a weakly monotone β. We mention that uniqueness theorems as well as regularity theorems one can find in different literature. One can also think about differential equations of fourth order. Depending on the boundary condtion, say, if they are given as Dirichlet conditions of the values and first derivatives, one uses $H = L^2(\Omega)$ and $V = W^{2,2}(\Omega)$ as spaces. Also other spaces have applications.

12. **Generalizations.** A generalization is the situation that H is a Banach space, e.g. as in [9]. We do not consider this situation here, but it is one of the first things to do. If $H = L^q(\Omega)$ with $q > 1$, more general monotone functions β,

$$|\,\beta(x,z)\,| \le C(1 + |\,z\,|^s) \quad \text{with } s > 1,$$

would be allowed. Another class of problems have even stronger growth like

$$\beta(u) = \mathrm{e}^u,$$

or more general an arbitrary growth. This class also belongs to the closure of problems of this method, but is not considered here.

An important example comes from diffusion in a chemical system. Although this seems to be an application of this theory, it is not clear, how the standard assumptions fit our theorem, see e.g. [24]. The source term, which is there already in the ODE version, is monotone, but does not give a contribution to the standard diffusion term. That is, if one leaves the ODE version unchanged, one has to manipulate the elliptic part of the problem to make the theorem work. This is done in [24].

As mentioned in the paper, a generalization to a time dependent constraint is possible, but not contained in this paper. This is important, as examples in [15], [16] and [13] show. It is also possible to show the existence locally in time, as presented in [3]. For this a weaker coercivity condition is enough.

Of big interest it would be, to generalize the set V to a locally convex topological vector space. The 3D incompressible Navier-Stokes equation does not satisfy the assumption in the present version. Leray's weak solution of Navier-Stokes equation uses $D(\Omega) = C_0^\infty(\Omega)$ and its dual space $D'(\Omega) = (D(\Omega))^*$ for the formulation. The generalization would go into this direction. The incompressible 2D Navier-Stokes equation is covered by the existence theorem of this paper because

$$\|u\|^2_{L^4([t_0,t_1]\times\Omega)} \le C \cdot \|u\|_{L^2([t_0,t_1];W^{1,2}(\Omega))} \|u\|_{L^\infty([t_0,t_1];L^2(\Omega))} \qquad \text{for } n = 2\,.$$

13. **Appendix.** In the following we deal with statements about Gelfand triples $(\widetilde{V}, H, \widetilde{V}^*)$, that is, about a Hilbert space H and a Banach space $\widetilde{V}$ satisfying

$$\widetilde{V} \hookrightarrow H \hookrightarrow \widetilde{V}^*, \tag{13.1}$$

by which one means, that the following line of mappings

$$\widetilde{V} \underset{I}{\to} H \underset{J_H}{\cong} H^* \underset{I^*}{\to} \widetilde{V}^* \tag{13.2}$$

exist, where $I : \widetilde{V} \to H$ is a continuous linear map. Here $J_H : H \to H^*$ is the isomorphism of the Riesz representation theorem and $I^* : H^* \to \widetilde{V}^*$ the adjoint map of $I : \widetilde{V} \to H$. We can also write

$$\widetilde{V} \underset{I}{\to} H \underset{J := I^* \circ J_H}{\to} \widetilde{V}^*. \tag{13.3}$$

In literature the notion of a Gelfand triple includes the injectivity of I and J, or equivalently, the injectivity of I and I^*. We mention the following

13.1 Proposition. Let H be a Hilbert space and $\widetilde{V}$ a Banach space, such that $(\widetilde{V}, H, \widetilde{V}^*)$ satisfies (13.2). Then

(1) $\|I(v)\|_H \le C\|v\|_{\widetilde{V}}$ for all $v \in \widetilde{V}$.

(2) $J := I^* \circ J_H : H \to \widetilde{V}^*$ has the representation

$$\langle\, v\,,\, J(u)\,\rangle_{\widetilde{V}} = (\, I(v)\,,\, u\,)_H \qquad \text{for } u \in H,\ v \in \widetilde{V}.$$

(3) The map $\widetilde{J} := I^* \circ J_H \circ I : \widetilde{V} \to \widetilde{V}^*$ satisfies

$$\Big\langle\, v_1\,,\, \widetilde{J}(v_2)\,\Big\rangle_{\widetilde{V}} = (\, I(v_1)\,,\, I(v_2)\,)_H \qquad \text{for } v_1, v_2 \in \widetilde{V}.$$

(4) J injective $\iff$ $I(\widetilde{V})$ dense in H.

The following has to be applied, if a general space $\widetilde{V}$ is given.

13.2 Lemma. The general case in 13.1, with the assumption that I is injective, can be reduced to the special case

$$V \subset H \text{ with a continuous mapping } \mathrm{Id} : V \to H$$

via

$$V := I(\widetilde{V}) \subset H, \quad \|v\|_V := \left\|I^{-1}v\right\|_{\widetilde{V}}.$$

Remark: If $\widetilde{V}$ is a Hilbert space, also V becomes a Hilbert space with $(\, v_1 \,,\, v_2 \,)_V := \big(\, I^{-1}v_1 \,,\, I^{-1}v_2 \,\big)_{\widetilde{V}}$.

The following theorem is functional analysis (a proof is contained in [7]). In this paper it is unspoken used when bounded sets in $L^{p^*}([t_0,t_1];V^*)$ occurred and when it was used that such sets are weakly sequentially compact.

13.3 Lemma. If V is a separable reflexive Banach space and $1 < p < \infty$. Then

$$L^{p^*}([t_0,t_1];V^*) \cong (L^p([t_0,t_1];V))^*.$$

The Isomorphism I_p going from the left space to the right one is given by

$$\langle\, v \,,\, I_p(f^*) \,\rangle_{L^p([t_0,t_1];V)} := \int_{t_0}^{t_1} \langle\, v(t) \,,\, f^*(t) \,\rangle_V \;\mathrm{d}t$$

for $f^* \in L^{p^*}([t_0,t_1];V^*)$ and for $v \in L^p([t_0,t_1];V)$.

REFERENCES

[1] H. Amann, "Linear and Quasilinear Parabolic Problems, Volume I: Abstract Linear Theory," Monograph in Mathematics, Birkhäuser Basel, 1995.

[2] H. W. Alt, "Elliptische Probleme mit freiem Rand," Lecture Notes **21** SFB 256, Bonn, 1991.

[3] H. W. Alt, *Partielle Differentialgleichungen III*, Vorlesung Wintersemester 2003/04, Universität Bonn, unpublished manuscript.

[4] H. W. Alt and E. DiBenedetto, *Nonsteady flow of water and oil through inhomogeneous porous media*, Ann. Scuola Norm. Sup. Pisa, Cl. Sci., **12** (1985), 335–392.

[5] H. W. Alt and S. Luckhaus, *Quasilinear elliptic-parabolic differential equations,* Math. Z., **183** (1983), 311–341.

[6] H. W. Alt, S. Luckhaus and A. Visintin, *On nonstationary flow through porous media*, Ann. Mat. Pura Appl., **136** (1984), 303–316.

[7] A.-K. Becher, "Ein abstrakter Existenzsatz für elliptisch-parabolische Systeme," Diplomarbeit 2005, Universität Bonn.

[8] M. S. Berger, "Nonlinearity and Functional Analysis," Lectures on Nonlinear Problems in Mathematical Analysis, Academic Press, 1977.

[9] F. Bernis, *Existence results for doubly nonlinear higher order parabolic equations on unbounded domains*, Math. Ann., **279** (1988), 373–394.

[10] E. DiBenedetto, "Degenerate Parabolic Equations," Universitext, Springer, 1993.

[11] G. Duvaut and J. L. Lions, "Inequalities in Mechanics and Physics," Grundlehren der mathematischen Wissenschaften 219, Springer-Verlag, 1976.

[12] A. Friedman, "Partial Differential Equations," Holt, Rinehart and Winston New York, 1969.

[13] U. Fermum, "Nichtlineare elliptisch-parabolische Gleichungen mit zeitabhängigen Hindernissen," Diplomarbeit 2005, Universität Bonn.

[14] W. Jäger and J. Kačur, *Solution of doubly nonlinear and degenerate parabolic problems by relaxation schemes*, Math. Modelling Numer. Anal., **29** (1995), 605–627.

[15] N. Kenmochi, *Solvability of nonlinear evolution equations with time-dependent constraints and application*, Bull. Fac. Education Chiba Univ., **30** (1981), 1–87.

[16] N. Kenmochi and I. Pawlow, *A class of nonlinear elliptic-parabolic equations with time dependent constraints*, Nonlinear Analysis, **10** (1986), 1181–1202.

[17] D. Kinderlehrer and G. Stampacchia, "An Introduction to Variational Inequalities and their Application," Academic Press, 1980.

[18] D. Kröner and S. Luckhaus, *Flow of oil and water in a porous medium*, J. Differential Equations, **55** (1984), 276–288.

[19] O. A. Ladyženskaja, V. A. Solonnikov and N. N. Ural'ceva, "Linear and Quasilinear Equations of Parabolic Type," Translations of Mathematical Monographs, Vol. 23, 1968.

[20] I. Müller, "Thermodynamics," Interaction of mechanics and mathematics series, Pitman Boston London Melbourne, 1985.

[21] P. A. Raviart, *Sur la résolution de certaines equations paraboliques non linéaires*, J. Functional Analysis, **5** (1970), 299–328.
[22] M. Růžička, "Nichtlineare Funktionalanalysis. Eine Einführung," See also the version in `aam.mathematik.uni-freiburg.de/IAM/homepages/rose/springer.html`, Springer, 2004.
[23] J. Simon, *Compact sets in the space* $L^p(0,T;B)$, Ann. Mat. Pura Appl., **146** (1987).
[24] M. Schmidt, "Ein Existenzsatz für parabolische Systeme zur Beschreibung von chemischen Reaktionen," Diplomarbeit 2006, Universität Bonn.
[25] A. Visintin, "Models of Phase Transition," Birkhäuser, 1996.

Received January 2011; revised May 2011.

Remark: In contrast to the originally published paper this version of the paper contains the corrections of some minor mistakes.

E-mail address: `alt@ma.tum.de`