Variational principle for general diffusion problems

June 28, 2004

Luca Petrelli & Adrian Tudorascu
Department of Mathematical Sciences
Carnegie Mellon University, Pittsburgh, PA 15213
luca@andrew.cmu.edu, adriant@andrew.cmu.edu

Abstract

We employ the Monge-Kantorovich mass transfer theory to study existence of solutions for a large class of parabolic partial differential equations. We deal with non-homogeneous nonlinear diffusion problems (of Fokker-Planck type) with time dependent coefficients. This work greatly extends the applicability of known techniques based on constructing weak solutions by approximation with time-interpolants of minimizers arising from Wasserstein-type implicit schemes ([1], [11], [13], [14], [15]). It also generalizes the results from [18] and [20], where proofs of convergence in the case of a right hand side in the equation is given by these methods. To prove existence of weak solutions we establish an interesting maximum principle for such equations. This involves comparison with the solution for the corresponding homogeneous, time-independent equation.

Keywords: Anomalous diffusion, diffusion equations, optimal mass transportation, Wasserstein distance, discretized gradient flow, implicit schemes, nonhomogeneous, non-autonomous problem, weak solution.


1 Introduction

In the present work we study general diffusion problems with drift and diffusion coefficients which may be explicitly time dependent. The boundary conditions considered here are of Neumann type. In the absence of a forcing term, this causes the solution to preserve its total mass at all times. Conservation of total mass is clearly not achieved in the presence of a non-homogeneous term. Under certain conditions the sign of the initial data is preserved by the solution. These features (i.e. conservation of mass and sign preservation) are crucial to the variational approach by techniques based on time-discretized implicit schemes in the Wasserstein distance.

More precisely, we consider the following problem (where all gradients are spatial):

\[
\begin{cases}
    u_t - \nabla \cdot (u \nabla \Psi(x,t)) - \Delta f(t, u) = g(x, t, u) \quad \text{in } \Omega \times (0, T), \\
    (u \nabla \Psi + \nabla f(t, u)) \cdot \nu = 0 \quad \text{on } \partial \Omega \times (0, T), \\
    u(\cdot , 0) = u_0 \quad \text{in } \Omega,
\end{cases}
\]

(P)

where \(u_0\) is the initial datum and \(f : \mathbb{R}^2 \to \mathbb{R}, \ g : \mathbb{R}^{N+2} \to \mathbb{R} \) and \(\Psi : \mathbb{R}^{N+1} \to \mathbb{R}\) are functions with certain properties.

This general problem is connected to several issues of current interest in statistical physics, including anomalous diffusion in many systems. Interesting examples are mentioned in [16]: diffusion in plasmas, surface growth and transport of fluid in porous media, relative diffusion in turbulent media. In special cases even exact solutions can be computed ([6], [16], [17]). For example, [16] produces an exact solution for the equation:

\[
u_t = -\nabla \cdot [F(\cdot , t)u] + a(t)\Delta u^\nu,
\]
where \( \nu \geq 1 \) and \( \mathbf{F}(x, t) := k_1(t) + k_2(t)x \) falls into our admissible category of drift coefficients (indeed, \( \mathbf{F}(\cdot, t) \) is conservative for all \( t \)).

The homogeneous case (\( g \equiv 0 \)) comprises a large class of Fokker-Planck equations which describe the evolution in time of the probability density functions (p.d.f.’s) of stochastic processes that are solutions for certain stochastic differential equations (see [10]). For the linear, generic case see [11]. For the general case, we found [2] of great help as to its stochastic derivation and interpretation. It is shown that one is “[…] led to accept that the stochastic force depends […] on the probability distribution”. More precisely, \( u(\cdot, t) \) is the p.d.f. of a stochastic process \( \{X_t\} \) which is the solution for the Ito stochastic differential equation:

\[
dX_t = -\Psi(X_t, t)dt + H(X_t, t)dW_t, \tag{1.1}
\]

where \( \{W_t\} \) is the standard Wiener process and \( H \) is the stochastic force. Here

\[
H(x, t) = \left[ \frac{2f(t, u(x, t))}{u(x, t)} \right]^{1/2}.
\]

Techniques inspired by the Monge-Kantorovich optimal mass transfer theory are known to be applicable to solving certain parabolic partial differential equations. Jordan, Kinderlehrer and Otto show in [11] that the solution for the heat equation is the gradient flow of a free energy functional with respect to the Wasserstein distance. In [14] Otto gives a similar result for the porous medium equation and more recently Agueh, in [1], proves existence for more general parabolic equations:

\[
\frac{\partial u}{\partial t} = \text{div} \left\{ u \nabla c^* (\nabla (F'(u) + V)) \right\},
\]

where \( c \) is some appropriate cost function, \( c^* \) its Legendre transform, \( F \) the internal energy and \( V \) a convex potential. Agueh uses more general Monge-Kantorovich distances (induced by the cost function \( c \)) than the quadratic cost Wasserstein metric; he is, therefore, able to extend previous results, such as Otto’s [14]. Key to all these results is the construction of a time step approximation, which is then linearly interpolated and shown to converge to some notion of weak solution. This sequence in time is recovered from the variational principle:

\[
\mathbf{u}_k := \arg\min \left\{ \frac{1}{2} d(u, u_{k-1})^2 + hS(u) \right\}, \tag{1.2}
\]

where \( d(u, u_{k-1}) \) represents the Wasserstein distance between the measures \( u dx \) and \( u_{k-1} dx \), having the same total mass, and \( S(u) \) represents the free energy of the system. Conservation of mass allows for comparison between different time steps by means of the Wasserstein distance. In [12], Kinderlehrer and Walkington present a numerical algorithm, which uses the main idea of this variational principle, to solve certain parabolic partial differential equations with forcing terms. Mass is no longer conserved, therefore, in order to be able to compare different time steps, they suggest including a time-averaged correction in the discrete scheme:

\[
\mathbf{u}_k := \arg\min \left\{ \frac{1}{2} d(u, u_{k-1} + h\tilde{g})^2 + hS(u) \right\}, \tag{1.3}
\]

where now the minimum has to be taken over \( u \) with the same mass as \( u_{k-1} + h\tilde{g} \) and \( \tilde{g} \) is the time average over the \( k \)-th interval (i.e., \( [kh, (k + 1)h) \)) of the non-homogeneous data \( g \). In this paper we prove convergence for the cases studied in [12] and we extend the proof to (\( P \)).

Let us remark that the presence of \( g \) is not the only novelty of this work. In fact, through some nontrivial changes of standard methods we were able to include a direct time dependence for both the drift coefficient \( \Psi \) and the nonlinear diffusion term \( f \). While the concept in itself is quite simple, we now work with time averages for the above quantities, the proofs require some nontrivial induction. The authors have not been able to find an existence result in this generality in the literature. Proofs of convergence, using these methods, in less general cases than the ones considered here may be found in [18] and [20].

The main advantage of such a method over more traditional ones is evident in (1.2): no gradient of the approximating functions enters the variational principle. This allows constructing approximations from discontinuous functions, an appealing situation for problems which exhibit steep gradients. Also, with
this method we are able to minimize directly in the weak topology, naturally induced by the Wasserstein distance. Furthermore, our method has the remarkable property of constructing good approximations for the solution which presents interesting numerical advantages (e.g. reduction of oscillations, stability). Another interesting consequence of this new approach is that we do not require convexity of the potential. This makes our result available for problems where the potential might have more than one well (e.g. [5]).

The paper is organized as follows; in Section 2 we set up the format and we remind the reader of the main properties of the Wasserstein distance. We also present a maximum principle which will be used extensively in the sequel. We find this maximum principle a remarkable result. It is inspired by the pioneering work of Otto ([14]).

In Section 3 we recover the Euler equations of the variational principle and prove necessary inequalities that will be used to show convergence of the interpolation of the time-step approximations to the weak solution. We also prove the main result in the case of a linear diffusion coefficient where all that is needed is weak convergence. In order to prove the main result in the general nonlinear case we need strong convergence, which we recover through a compactness result for \( u^k \) in \( L^1(\Omega \times [0,T]) \). This is taken care of in Section 4, where we use the Riesz-Frechet-Kolmogorov criterion to establish the desired compactness (leading to the proof of existence).

We conclude in the last section with a partial uniqueness result. Long time behavior of solutions is to be addressed in future work.

2 The Weak Solution

Let \( u_0 \in L^1(\Omega) \) be the initial data and denote \( Q := \Omega \times (0,T) \).

**Definition 1.** A weak solution for \((P)\) is a function \( u(x,t) \) satisfying

(i) \( u \in L^\infty((0,T);L^1(\Omega)) \);

(ii) \( u\nabla \Psi + \nabla f(\cdot,u) \in L^1(Q) \) and \( g(\cdot,\cdot,u) \in L^1(Q) \);

(iii) For every \( \zeta \in C^\infty_c(\Omega \times [0,T]) \), i.e. smooth functions on \( \bar{\Omega} \times [0,T] \) that vanish near \( t = T \), we have

\[
\iint_Q \{ u\partial_t \zeta - (u\nabla \Psi + \nabla f(\cdot,u)) \cdot \nabla \zeta + g(\cdot,\cdot,u)\zeta \} \,dxdt = -\int_\Omega u_0 \zeta(\cdot,0) \,dx. \tag{2.1}
\]

By means of the Wasserstein metric Otto proves (in [14]) the existence of weak solutions for \((P)\) for the case \( \Omega = \mathbb{R}^N \), \( f(s) = s^p \), \( \Psi \equiv g \equiv 0 \), and nonnegative, bounded and compactly supported initial data (here \( T = \infty \)). The weak solution \( u \) is found to satisfy \( 0 \leq u \leq \|u_0\|_{L^\infty(\Omega)} \) \( \text{a.e. in } \Omega \times (0,\infty) \) and conserve \( \|u(\cdot,t)\|_{L^1(\Omega)} = \|u_0\|_{L^1(\Omega)} \) for all \( t > 0 \). For uniqueness we are referred to [21].

Denote by \((f1), (f2)\) the following properties of a function \( f : \mathbb{R} \times [0,\infty) \to \mathbb{R} \),

\[
(u - v)(f(t,u) - f(t,v)) \geq c|u - v|^{\omega} \text{ for all } u, v \geq 0, \tag{f1}
\]

for some \( c > 0 \) independent of \( t \geq 0, \omega \geq 2 \) and

\[
f(\cdot, s) \text{ is Lipschitz continuous, uniformly for } s \text{ in bounded sets.} \tag{f2}
\]

**Remark:** Note that all power functions \( f(s) = s^p \), \( p \geq 1 \) satisfy \((f1)\) (and, trivially, \((f2)\)). Indeed, notice that, if, say, \( x \geq y \geq 0 \), then \( x^p - y^p \geq (x-y)^p \). This leads to \( (x-y)(x^p - y^p) \geq |x-y|^{p+1} \) for all \( x, y \geq 0 \).

In order to avoid the technical difficulties raised by working with the Wasserstein metric on unbounded domains we assume \( \Omega \) to be bounded. Also, for convenience, we assume momentarily that \( 0 < T < \infty \). Let \( u_1, u_2 : \Omega \to \mathbb{R} \) be two nonnegative and integrable functions that satisfy \( \int_\Omega u_1 \,dx = \int_\Omega u_2 \,dx > 0 \). Denote:

\[
P(u_1, u_2) := \left\{ \text{nonnegative Borel measure } \mu \text{ on } \Omega \times \Omega \mid \begin{align*}
\iint_{\Omega \times \Omega} \xi(x)d\mu(x,y) &= \int_\Omega \xi(x)u_1(x) \,dx \\
\iint_{\Omega \times \Omega} \xi(y)d\mu(x,y) &= \int_\Omega \xi(y)u_2(y) \,dy \text{ for all } \xi \in C(\Omega) \end{align*} \right\}
\]

Then we have:
Definition 2. The $p^{th}$ power of the Wasserstein distance of order $p$ is defined as

$$d_p(u_1, u_2)^p := \inf_{\mu \in P(u_1, u_2)} \int_{\Omega \times \Omega} |x - y|^p d\mu(x, y).$$

Assume $u_0 \geq 0$ is a nonzero element of $L^1(\Omega)$, $\Psi \equiv g \equiv 0$ and $f \equiv f(u)$. Formally, a solution of $(P)$ is a gradient flow of the functional $S(u) := \int_\Omega \phi(u) dx$ with respect to the Wasserstein metric of order 2 (simply denoted by $d$), i.e., a differentiable function $(0, \infty) \ni t \rightarrow u(t) \in \mathcal{M}$ such that $u(0) = u_0$ that satisfies

$$m_{u(t)}(\partial_t u(t), v) + \langle \nabla S(u(t)), v \rangle = 0 \text{ for all } v \in T_{u(t)} \mathcal{M}_{u_0},$$

where

$$\mathcal{M}_{u_0} := \left\{ u \mid u \text{ is measurable, nonnegative and } \int_\Omega ud\mathcal{X} = \int_\Omega u_0 d\mathcal{X} \right\},$$

$T_{u(t)} \mathcal{M}_{u_0}$ is the tangent space to $\mathcal{M}_{u_0}$ at $u(t)$ and

$$\phi$$

is a solution of $s\phi'(s) - \phi(s) = f(s)$ in $[0, \infty)$ (2.2)

(see, for example, [15]). It is well-known that any gradient flow on the manifold $(\mathcal{M}, m)$ ($m$ is a metric tensor on $T \mathcal{M}$ that induces the Wasserstein metric on $\mathcal{M}$; see [15], [14]) admits a natural time-discretized variational scheme of the form $(d_m, m)$ being the metric induced by $m$ on $\mathcal{M}$:

Fix $h > 0$ and let $\rho_0 \geq 0$ be integrable of positive mass. For every integer $k \geq 1$ find $\rho_k \in \mathcal{M}_{\rho_0}$ that minimizes

$$\frac{1}{2h}d_m(\rho, \rho_{k-1})^2 + S(\rho)$$

among all $\rho \in \mathcal{M}_{\rho_0}$. This motivates our approach of proving that the scheme (2.5) has a unique solution in our case and we will show how to eventually construct a weak solution for $(P)$ by standard interpolation techniques. In order to do this we need several results. Some of them are not going to prove since they are simply direct generalizations of similar results proved in [11], [15], [14] and [13].

When we drop the assumption $\Psi \equiv 0$ and we let $f = f(t, u)$ we need consider time averages for the quantities $\Psi$ and $f$. We will still be able to define a sequence $u_k$ which will be used later to construct the time interpolation. Let:

$$\Psi_k^h := \frac{1}{h} \int_{k \cdot h}^{(k+1)h} \Psi(t, \cdot) dt \quad \text{and} \quad f_k^h := \frac{1}{h} \int_{k \cdot h}^{(k+1)h} f(t, \cdot) dt.$$ (2.3)

The functional $S$ will be replaced by a sequence of functionals $\{S_k^h\}_k$ given by

$$S_k^h(u) := \frac{1}{h} \int_{k \cdot h}^{(k+1)h} \int_\Omega \left\{ \phi(t, u(x)) + u(x)\Psi(x, t) \right\} dxdt,$$

where $\phi(t, \cdot)$ is the continuous solution for (2.2) with $f = f(t, \cdot)$, i.e.

$$\phi(t, \cdot) \text{ is a solution of } s\phi'(s) - \phi(s) = f(t, s) \text{ in } (0, \infty) \times [0, \infty).$$ (2.4)

Remark: Note that, whenever $f(t, \cdot)$ is continuous and has the right growth, then $\phi(t, \cdot)$ is $C^1$ and it satisfies $\phi^*(t, \cdot) \circ \phi'(t, \cdot) = f(t, \cdot)$, where $\phi^*$ is the Legendre transform in space of $\phi$ and $\phi'$ is its spatial derivative.

Assume

$$\Psi : \mathbb{R}^N \times (0, \infty) \rightarrow \mathbb{R}$$

is differentiable and locally Lipschitz in $x \in \mathbb{R}^N$. (Ψ1)

Lemma 1. Let $u^* \in L^1(\Omega)$ be nonnegative and $h > 0$ be given. If (f1) holds, then, for each integer $k \geq 1$, the functional

$$I_k^h[u^*](u) := \frac{1}{2} d(u, u^*)^2 + \int_{k \cdot h}^{(k+1)h} \int_\Omega \left\{ \phi(t, u(x)) + u(x)\Psi(x, t) \right\} dxdt$$ (2.5)

admits a unique minimizer in $\mathcal{M}_{u^*}$.
Notice that, due to (2.3), we have
\[ S_k^h(u) = \int_{\Omega} \{ \phi_k^h(u) + u \Psi_k^h \} \, dx \]
and
\[ I_k^h[u^*](u) = \frac{1}{2}d(u, u^*)^2 + hS_k^h(u). \]
The proof is already quite standard (see, for example, [11]). We just point out that, due to the boundedness of \( \Omega \), we can omit all the technical details used there to obtain bounds on the second moments of the measure densities involved. Also as in [11], we can avoid using the strict convexity of the square of the Wasserstein metric (see [14]) by noting that the uniqueness of the minimizer is due to the convexity of \( d(\cdot, u^*)^2 \) (easy to prove) and the strict convexity of the \( S_k^h \) (due to (f1)). The minimizer is obtained as the weak \( L^1 \) limit of a minimizing sequence by applying the superlinear growth \( L^1 \)-weak convergence criterion to \( \phi_k^h \), which, also as a consequence of (f1), satisfies for all \( k \) and all \( h > 0 \),
\[ \frac{\phi_k^h(s)}{s} = \int_1^s \frac{I_k^h(\sigma)}{\sigma^2} \, d\sigma + C \geq c \log s + f_k^h(0) \left( 1 - \frac{1}{s} \right) + C \text{ for all } s > 1 \]
\( (\Psi_k^h \) has no adverse effect on the superlinear growth of \( I_k^h[u^*] \) since \( \int_{\Omega} u \Psi_k^h \, dx \geq -\|\Psi\|_{L^\infty(Q)} \|u\|_{L^1(\Omega)} \).
A trivial iterative application of this lemma yields:

**Corollary 1.** Given \( u_0 \geq 0 \) of finite, positive mass, the following iterative scheme:

For any integer \( k \geq 1 \), “ \( u_k \) minimizes \( I_k^h[u_{k-1}] \) in \( \mathcal{M}_{u_0} \)” has a unique solution \( \{u_k\}_{k \geq 0} \). \hspace{1cm} (2.6)

**Remark:** The property (f1) is sufficient for giving the right growth of \( \phi \) in order to infer existence of the minimizer. However, this is not the reason why (f1) is used throughout the paper (it is needed to prove compactness in section 4). In fact, weaker assumptions on \( \phi \) (see [1]) than the ones implied by (f1) are sufficient.

These \( u_k \)'s turn out to be exactly what we need in the case \( g \equiv 0 \) to construct the solution of \( P \) by interpolating time. In [12], Kinderlehrer and Walkington suggest modifying (2.6) as follows:

**Corollary 2.** Given \( u_0 \geq 0 \) of finite, positive mass, the following iterative scheme:

For any integer \( k \geq 1 \), “ \( u_k \) minimizes \( I_k^h[u_{k-1}] \) in \( \mathcal{M}_{u_{k-1}} \)” has a unique solution \( \{u_k\}_{k \geq 0} \). \hspace{1cm} (2.7)

At each step \( k \geq 0 \) the function \( v_k : \Omega \to \mathbb{R} \) is defined as follows:
\[ v_k := u_k + \int_{kh}^{(k+1)h} g(\cdot, \tau, u_k(\cdot)) \, d\tau. \hspace{1cm} (2.8) \]

We make the following assumptions on \( g \):
\[ g(x, \cdot, \cdot) \text{ is nonnegative in } [0, \infty) \times [0, \infty) \text{ for all } x \in \mathbb{R}^N; \hspace{1cm} (g1) \]
\[ g(x, t, u) \leq C(1 + u) \text{ for } u \geq 0 \text{ locally uniformly with respect to } (x, t), \ t \geq 0; \hspace{1cm} (g2) \]
\[ g(x, t, \cdot) \text{ is continuous on } [0, \infty); \hspace{1cm} (g3) \]
\[ \{g(x, t, u)\}_{(x,u)} \text{ is equicontinuous on } [0, \infty) \text{ with respect to } (x,u). \hspace{1cm} (g4) \]
The fact that all the \( v_k \)'s and the \( u_k \)'s from Corollary 2 are nonnegative and have finite, positive mass (nondecreasing in \( k \)) follows trivially from (g1)-(g3).

The sequel will use the notation \( u_k^n \) instead of \( u_k \) when we want to emphasize the dependence of each such function on \( h \), dependence that will become increasingly important. For every positive integer \( n \), following [11], [14] etc., we define
\[ u^h : \Omega \times [0, T) \to [0, \infty), \ u^h(x, t) := u_k^n(x) \text{ if } kh \leq t < (k + 1)h, \hspace{1cm} (2.9) \]
for all \( k \in \mathbb{N}, \ k < n \) where \( h := T/n \).
Denote by $(\bar{\Omega})$. Here $\mu$ and consider the interpolant $u^h$ instead of $u$.

2.1 A discrete comparison principle

It is obvious that, for a given $h > 0$, we only need $\{u^h\}_{0 \leq k \leq n-1}$ to completely define $u^h$.

Denote by $(f3)$ the property

$$f(t, \cdot) \text{ is differentiable and } \frac{\partial f}{\partial s} \text{ is positive and monotone in time.} \quad (f3)$$

We are now ready for the following:

**Lemma 2.** If $0 \leq u_0 \leq M_0 < \infty$ a.e. in $\Omega$ for large enough $M_0$, then there exists $0 < M = M(M_0) < \infty$ such that $0 \leq u^h \leq M$ a.e. in $\Omega$ for all $h > 0$ if one of the following is true

(i) $f$ satisfies $(f3)$, $\lim_{t \to \infty} \phi(t, s) = \infty$ uniformly with respect to $t > 0$ and $g \equiv 0$;

(ii) $f$ satisfies $(f3)$, $\lim_{s \to \infty} \phi(t, s) = \infty$ uniformly in $t > 0$ and for $s > 0$ large enough we have

$$\eta s \frac{\partial f}{\partial s}(t, \eta s + \eta - 1) - (\eta s + \eta - 1) \frac{\partial f}{\partial s}(t, s) \text{ does not change sign for all } t > 0, \eta > 1, \quad (2.10)$$

being nonnegative if $\frac{\partial f}{\partial s}(\cdot, s)$ is increasing and nonpositive otherwise.

Remark: Note that all functions $f(t, s) = a(t)s^p$ for $p \geq 1$ and $\eta$ a monotone and bounded from below away from zero, are covered by the above statement even if $g$ is not identically zero. For $p = 1$ the expression in $(2.10)$ is nonnegative while for $p = 1$ it is nonpositive. Also, $(f1)$ implies $\lim_{s \to \infty} \phi'(s) = \infty$.

Remark: Note that it follows from some easy calculations that we have properties similar to $(f3)$, $(2.10)(i)$ and $(2.10)(ii)$ for $\phi_k$ and $f_k$. Also, in the first part of the following proof we drop the $k$'s to make it more readable.

Proof: Observe that in all cases $\phi'$ is strictly increasing and continuous on $(0, \infty)$ and $\phi'((0, \infty)) = (\phi'(0), \infty)$. Set, for convenience, $\theta := \phi'$ and $\vartheta := (\phi')^{-1}$. We are going to adjust the proof in [20] to fit our purpose. First, we claim that:

If $u \leq U^\phi_{\infty}$ a.e. in $\Omega$, then $u \leq U^\phi_{\infty}$ a.e. in $\Omega$, \quad (2.11)

where $U^\phi_{\infty} := \theta \circ (C - \Psi)$ ($C$ is a constant function) and $u$ is the minimizer from Lemma 1. Take $\mu \in P(u^*, u)$ to be the optimal transfer plan and define $v_0$ and $v_1$ by

$$\int_{\Omega} v_0 \xi dx = \int_{\{u(y) > U(y)\}} \xi(x) d\mu(x, y) \text{ for all } \xi \in C(\bar{\Omega}),$$

$$\int_{\Omega} v_1 \xi dx = \int_{\{u(y) > U(y)\}} \xi(y) d\mu(x, y) = \int_{\{u(y) > U(y)\}} u \xi dy$$

for all $\xi \in C(\bar{\Omega})$. Here $U := U^\phi_{\infty}$. The equality (valid for all $\xi \in C(\bar{\Omega} \times \Omega)$)

$$\int_{\Omega \times \Omega} \xi(x, y) d\mu(x, y) = \int_{\{u(y) \leq U(y)\}} \xi(x, y) d\mu(x, y) +$$

$$+(1 - \lambda) \int_{\{u(y) > U(y)\}} \xi(x, y) d\mu(x, y) + \lambda \int_{\{u(y) > U(y)\}} \xi(x, x) d\mu(x, y)$$

is valid for $\lambda \in [0, 1]$. The equality $(2.11)$ is valid for all $\xi \in C(\bar{\Omega} \times \Omega)$.
defines for every $\lambda \ll 1$ a plan $\mu_\lambda \in P(u^*, u_\lambda)$ with $u_\lambda := u - \lambda(v_1 - v_0) \in \mathcal{M}_{u^*}$. Then
\[
\frac{1}{2h} d(u^*, u_\lambda)^2 + \int_\Omega \{\phi(u_\lambda) + \Psi u_\lambda\} dx \leq I[u^*](u) + \int_\Omega \{\phi(u_\lambda) - \phi(u) + \Psi u_\lambda - \Psi u\} dx. \tag{2.12}
\]
due to the definition of $d$ and $\mu_\lambda$. Let
\[
\beta(\lambda) := \int_\Omega [\phi(u_\lambda) - \phi(u) + \Psi(u_\lambda - u)] dx = \int_\Omega [\phi(u - \lambda v) - \phi(u) - \lambda v \Psi] dx,
\]
where $v := v_1 - v_0 \in T_u\mathcal{M}_{u^*}$. We have
\[
\beta'(0) = -\int_\Omega v(\theta(u) + \Psi) dx.
\]
Since $\theta$ is strictly increasing and $v_1 = u\chi_{\{u > U\}} > U \geq v_0$ on $\{u > U\}$ the splitting (take into account that $v_1 = 0$ on $\{u \leq U\}$)
\[
\beta'(0) = -\int_{\{u > U\}} v(\theta(u) - \theta(U)) dx + \int_{\{u \leq U\}} v_0(\theta(u) - \theta(U)) dx - \int_\Omega v(\Psi + \theta(U)) dx,
\]
shows that $\beta'(0) < 0$. Indeed, $\Psi + \theta(U) \equiv C = \text{const}$ and $v$ has null average so the last term of the inequality above vanishes. Therefore, $\beta(0) = 0 > \beta'(0)$ and so $\beta(\lambda)$ is negative for sufficiently small $\lambda > 0$. This, however, contradicts the minimality of $\lambda > 0$. Hence, we wish to show $\mu_\lambda$.

Note that, if we choose $m_k$ large enough, we have no trouble defining all of the above. Let also, for $c_k$, constants to be defined later, $M_k := \theta_k(m_k) + c_k$, for $1 \leq k \leq n$. Next we claim:
\[
u^*_k(x) \leq \|u^*\|_{L^\infty(\Omega)} \leq \theta(\|u^*\|_{L^\infty(\Omega)}) + \|\Psi\|_{L^\infty(\Omega)} - \Psi(x),
\]
for a.e. $x \in \Omega$.

Once we reconsider the right notations in the picture we have just proved the following adjusted claim:

If $u^* \leq U_{\mathcal{M}_k}^\theta \Phi_k$ a.e. in $\Omega$, then $u \leq U_{\mathcal{M}_k}^\theta \Phi_k$ a.e. in $\Omega$, \tag{2.13}

where the constants $M_k$ will be chosen below. Let $b$ a constant, which we choose later, and define the sequence $\{m_k\}_{1 \leq k \leq n}$ by
\[
m_k := \theta_k(a_k) - b \quad \text{where} \quad a_k := (1 + Ch)^k(M_0 + b) + (1 + Ch)^k - 1. \tag{2.14}
\]

Note that, if we choose $M_0$ large enough, we have no trouble defining all of the above. Let also, for $c_k$, constants to be defined later, $M_k := \theta_k(m_k) + c_k$, for $1 \leq k \leq n$. Next we claim:
\[
u^*_k \leq U_{\mathcal{M}_k}^\theta \Phi_k \text{ a.e. in } \Omega \text{ for all } 1 \leq k \leq n. \tag{2.15}
\]

This would bring our proof to its conclusion due to the boundedness of the right hand side uniformly with respect to $h$ (proven later). We prove (2.15) by induction. Note that, due to (2.13), it suffices to show that $\nu^*_{k-1} \leq U_{\mathcal{M}_{k-1}}^\theta \Phi_{k-1}$ a.e. in $\Omega$. The induction hypothesis reads $\nu^*_{k-1} \leq U_{\mathcal{M}_{k-1}}^\theta \Phi_{k-1}$ a.e. in $\Omega$. Then (2.8) and (g2) imply
\[
u_{k-1} \leq (1 + Ch)\nu_{k-1} + Ch \leq (1 + Ch)U_{\mathcal{M}_{k-1}}^\theta \Phi_{k-1} + Ch.
\]

Consequently, we wish to show
\[
(1 + Ch)U_{\mathcal{M}_{k-1}}^\theta \Phi_{k-1} + Ch \leq U_{\mathcal{M}_k}^\theta \Phi_k.
\]

This is equivalent to
\[
(1 + Ch)\theta_{k-1}(m_{k-1}) + c_{k-1} - \Psi_{k-1} + Ch \leq \theta_k(m_k) + c_k - \Psi_k. \tag{2.16}
\]
The relations in (2.14) give
\[ \theta_k(m_k) = \theta_k((1 + Ch)a_{k-1} + Ch) - b = \theta_k((1 + Ch)\vartheta_{k-1}(m_{k-1}) + b) + Ch - b. \]

Therefore, as all the \( \theta_k, \vartheta_k \) are strictly increasing, (2.16) is equivalent to
\[ \theta_k((1 + Ch)\vartheta_{k-1}(m_{k-1}) + e^{k-1} - \Psi_{k-1}) + Ch) - (e^k - \Psi_k) \leq \]
\[ \leq \theta_k((1 + Ch)\vartheta_{k-1}(m_{k-1}) + b) + Ch) - b. \]

We now choose \( b = LT + ||\Psi_0||_{L^\infty(Q)} + ||\Psi||_{L^\infty(Q)}, \) where \( L \) is the Lipschitz constant in time for \( \Psi, \) i.e. we have:
\[ ||\Psi(x, t + h) - \Psi(x, t)|| \leq LH \quad \text{for all } x \in \Omega \text{ and } t \in [0, T] \]  
and we define the \( e^k \) recursively as \( e^0 := ||\Psi_0||_{L^\infty(Q)} \) and \( e^k := e^{k-1} + LH. \)

Therefore, we obtain \( 0 \leq e^{k-1} - \Psi_{k-1} \leq e^k - \Psi_k \leq b \) and to prove (2.16) it is enough to show that
\[ q(s) := \theta_k((1 + Ch)\vartheta_{k-1}(m_{k-1}) + s) + Ch) - s \text{ is increasing on } (0, \infty). \]

This result follows from (2.10) with the positive sign. Indeed,
\[ q'(s) = (1 + Ch)\theta'_k((1 + Ch)\vartheta_{k-1}(m_{k-1}) + s) + Ch) \frac{d}{ds} [\vartheta_{k-1}(m_{k-1}) + s] - 1 \]
\[ = \eta \theta'_k(\eta \vartheta_{k-1}(m_{k-1}) + s) + \eta - 1)\vartheta'_{k-1}(m_{k-1}) + s) - 1, \]
where \( \eta := 1 + Ch. \)

Taking into account that \( \vartheta'_{k-1}(x) = 1/[\theta'_{k-1}(\theta(x))] \), and setting \( y = \vartheta_{k-1}(m_{k-1} + s) \)
we obtain that \( q \) is nondecreasing if
\[ \frac{\eta f'_k(\eta y + \eta - 1)}{\eta y + \eta - 1} \geq \frac{f'_{k-1}(y)}{y} \]
which follows from (f3) (which implies \( f'_k \geq (f_{k-1})' \)) and (2.10) with \( \frac{\partial f}{\partial y} \) nondecreasing.

Else, we set \( b = 0 \) and \( c^k \) as above and show that \( q(s) \) is decreasing on \( (0, \infty). \)

As \( e^k - \Psi_k \geq 0, \) it follows that \( q(0) \geq q(c^k - \Psi_k) \) which gives us the bound in this case.

Of course, if \( g \) is null, then \( C = 0 \) and \( q \) is constant. Thus, (i) is proved if we can show that the \( U^\phi_k \Psi_k \) are bounded: recalling the various definitions involved we have:
\[ U^\phi_k \Psi_k = \vartheta_k \circ (M_k - \Psi_k) = \vartheta_k \circ (\theta_k(a_k) - b + c_k - \Psi_k) \leq \vartheta_k \circ \theta_k(a_k) = a_k \]
when we have (2.10) with the positive sign. Now it follows easily from their definition that the terms \( a_k \)
are bounded by \( c^{C+1} (\Theta_0 + b + 1) - 1. \)

In case we have (2.10) with the negative sign then the following is true
\[ U^\phi_k \Psi_k \leq \vartheta_k \circ (\vartheta_k(c^{C+1} (\Theta_0 + 1) + C T + ||\Psi_0||_{L^\infty(Q)} + ||\Psi||_{L^\infty(Q)}) \leq \vartheta_k \circ \theta_k(C') = C' \]
for such a \( C' \) which is bound to exist thanks to our hypothesis on the uniform growth of \( \phi_k. \) \]

\[ \square \]

Remark: If \( \Psi \equiv 0 \) and \( f = f(u) \), then we can replace \( U^\phi_k \Psi \) in (2.11) by a constant and the proof remains valid. Also, in this case \( f \) need only be increasing (none of (f1), (f2), (f3) or (2.10) are needed).

Remark: Existence of a uniform lower bound for all \( u^n \) provided that \( u_0 \) be essentially bounded from below away from zero does not seem to follow once we have the uniform upper bound (as it did in [1], [13], [20]). That is, unless we impose unrealistic conditions on the the potential \( \Psi \) and the lower bound of \( u_0 \). The existence of such uniform lower bound would make possible proving compactness in Section 4 without using (f1) but weaker assumptions. We thank the referee for this observation.
2.2 The approximate Euler equation

Denote by \( \mathbb{N} \) the set of all nonnegative integers.

**Proposition 1.** Let \( u_0 \in L^\infty(\Omega) \) be nonnegative of positive mass and \( \{u_k\}_{k \in \mathbb{N}} \) be the solution of (2.7). Then \( u_k \in L^\infty(\Omega) \) and \( f_k(u_k) \in W^{1,\infty}(\Omega) \) for all \( k \geq 1 \). Furthermore

\[
\left| \int_{\Omega} \left\{ \frac{1}{h}(u_k - u_{k-1})\zeta + u_k \nabla \Psi_k \cdot \nabla \zeta + \nabla f_k(u_k) \cdot \nabla \zeta - \left( \frac{1}{h} \int_{(k-1)h}^{kh} g(\cdot, \tau, u_{k-1}) d\tau \right) \zeta \right\} dx \right| \leq \frac{1}{2h} \|\nabla^2 \zeta\|_\infty d(v_{k-1}, u_k)^2,
\]

(2.18)

for every \( \zeta \in C^\infty(\bar{\Omega}) \).

We will next give a sketch of the proof following mainly [20].

**Proof:** Consider a smooth vector field \( \bm{\theta} \in C^\infty_0(\Omega; \mathbb{R}^N) \) and the corresponding flux, i.e. a one-parameter family \( \{\vartheta(\cdot; \tau)\}_{\tau \in \mathbb{R}} \) of diffeomorphisms, given by

\[
\vartheta(\cdot; \tau) = \xi \circ \vartheta(\cdot; \tau) \quad \text{for all} \quad \tau \in \mathbb{R}, \quad \vartheta(\cdot; 0) = id_\Omega.
\]

Then, as mentioned in [11], \( \vartheta(\cdot; \tau) \) is invertible for all \( \tau \in \mathbb{R} \) and \( \vartheta(\Omega; \tau) = \Omega \) for \( |\tau| \) sufficiently small. For such \( \tau \) we define

\[
u_\tau := \frac{u_k \circ \vartheta^{-1}(\cdot; \tau)}{\det J \vartheta(\cdot; \tau) \circ \vartheta^{-1}(\cdot; \tau)},
\]

which can be shown to lie in \( M_{v_{k-1}} \). This means that \( u_\tau \) is admissible in the variational principle (2.7) and we have

\[
\frac{1}{2} d(v_{k-1}, u_k)^2 + h \int_{\Omega} \{\phi_k(u_k) + u_k \Psi_k\} dx \leq \frac{1}{2} d(v_{k-1}, u_\tau)^2 + h \int_{\Omega} \{\phi_k(u_\tau) + u_\tau \Psi_k\} dx.
\]

(2.19)

Note that \( u_\tau \) is the so-called push-forward density of \( u_k \) under the diffeomorphism \( \vartheta(\cdot; \tau) \), i.e.

\[
\int_{\Omega} u_\tau(x) \zeta(x) dx = \int_{\Omega} u_k(x) \zeta(\vartheta(x; \tau)) dx \quad \text{for all} \quad \zeta \in C(\bar{\Omega}).
\]

Therefore, we can use \( \Psi_k \) instead of \( \zeta \) and, as \( \frac{1}{2} (\Psi_k(\vartheta(\cdot; \tau)) - \Psi_k(\cdot)) \) converges as \( \tau \downarrow 0 \) uniformly in \( \Omega \) to \( \nabla \Psi_k \cdot \xi \), we have

\[
\lim_{\tau \to 0} \frac{1}{\tau} \left( \int_{\Omega} (u_\tau - u_k) \Psi_k dx \right) = \int_{\Omega} \nabla \Psi_k \cdot \xi u_k dx.
\]

(2.20)

Now \( \frac{d}{d\tau} [\det J \vartheta(\cdot; \tau)] |_{\tau = 0} = \nabla \cdot \xi \) easily gives

\[
\lim_{\tau \to 0} \frac{1}{\tau} \left( \int_{\Omega} \phi_k(u_\tau) dx - \int_{\Omega} \phi_k(u_k) dx \right) = - \int_{\Omega} f_k(u_k) \nabla \cdot \xi dx.
\]

(2.21)

Let \( \mu \in P(u_{k-1}, u_k) \) be optimal in the definition of \( d(u_{k-1}, u_k) \). Then

\[
\int_{\Omega \times \Omega} \zeta(x, y) d\mu_\tau(x, y) = \int_{\Omega \times \Omega} \zeta(x, \vartheta(y; \tau)) d\mu(x, y) \quad \text{for all} \quad \zeta \in C(\bar{\Omega} \times \bar{\Omega})
\]

defines a \( \mu_\tau \in P(u_{k-1}, u_\tau) \) and thus we have

\[
\frac{1}{\tau} \left( \frac{1}{2} d(v_{k-1}, u_\tau)^2 - \frac{1}{2} d(v_{k-1}, u_k)^2 \right) \leq \int_{\Omega \times \Omega} \frac{1}{2} \left( \frac{1}{2} |\vartheta(y; \tau) - x|^2 - \frac{1}{2} |x - y|^2 \right) d\mu(x, y),
\]

which, together with (2.19), (2.20) and (2.21), yields (upon replacing \( \xi \) also by \( -\xi \))

\[
\int_{\Omega \times \Omega} (y - x) \cdot \xi(y) d\mu(x, y) + h \int_{\Omega} (\nabla \Psi_k \cdot \xi u_k - f_k(u_k) \nabla \cdot \xi) dx = 0 \quad \text{for all} \quad \xi \in C^\infty_c(\Omega; \mathbb{R}^N).
\]

(2.22)
Now consider $\zeta \in C^\infty_c(\Omega)$ which gives $\nabla \zeta := \xi \in C^\infty_c(\Omega; \mathbb{R}^N)$. As the following holds

$$
\zeta(y) - \zeta(x) = (y - x) \cdot \nabla \zeta(y) + (1/2)(y - x)^T (\nabla^2 \zeta)(z)(y - x),
$$

for some $z \in \Omega$ (here we think of $\zeta$ as being extended by zero outside $\Omega$), we infer

$$
\left| \int_\Omega (u_k - v_{k-1}) \zeta dx - \int_{\Omega \times \Omega} (y - x) \cdot \nabla \zeta(y) d\mu(x,y) \right| 
$$

(2.23)

$$
= \int_{\Omega \times \Omega} |\zeta(y) - \zeta(x) - (y - x) \cdot \nabla \zeta(y)| d\mu(x,y) 
\leq \frac{1}{2} \sup_\Omega |\nabla^2 \zeta| \int_{\Omega \times \Omega} |y - x|^2 d\mu(x,y) 
= \frac{1}{2} \|\nabla^2 \zeta\|_{L^\infty} d(v_{k-1}, u_k)^2 
$$

for all $\zeta \in C^\infty_c(\Omega)$ with $\nabla \zeta \in C^\infty_c(\Omega; \mathbb{R}^N)$.

According to Lemma 2, all $u_k$ are uniformly bounded in $L^\infty$ for $1 \leq k \leq n - 1$ independently of $h$. Obviously, so are $f_k(u_k)$ for $1 \leq k \leq n - 1$. We will actually prove that $f_k(u_k) \in W^{1,\infty}(\Omega)$. For this consider:

$$
\left| \int_{\Omega \times \Omega} (y - x) \cdot \xi(y) d\mu(x,y) \right| \leq \text{diam}(\Omega) \int_{\Omega \times \Omega} |\xi(y)| d\mu(x,y) 
$$

(2.24)

$$
= \text{diam}(\Omega) \int_\Omega |\xi(y)| u_k(y) dy \leq \text{diam}(\Omega) \|u_k\|_{L^p(\Omega)} \|\xi\|_{L^p(\Omega; \mathbb{R}^N)},
$$

and so, by (2.22)

$$
\int_\Omega f_k(u_k) \nabla \cdot \xi dx \leq \left( \frac{1}{h} \text{diam}(\Omega) \|u_k\|_{L^p(\Omega)} + \|\Psi\|_{L^\infty(\Omega)} \|u_k\|_{L^p(\Omega)} \right) \|\xi\|_{L^p(\Omega; \mathbb{R}^N)},
$$

for all $\xi \in C^\infty_c(\Omega; \mathbb{R}^N)$. This implies

$$
\nabla f_k(u_k) \in L^p(\Omega) \text{ for all } 1 \leq p \leq \infty.
$$

Therefore, $\{f_k(u_k)\}_{k \in \mathbb{N}} \subset W^{1,\infty}(\Omega)$ and, as $\Omega$ is bounded, we conclude that $\{f_k(u_k)\}_{k \in \mathbb{N}} \subset W^{1,p}(\Omega)$ for any $1 \leq p \leq \infty$.

It is important to realize that (2.22) can be more refined for $p = 2$ (we will make use of it later). Indeed

$$
\left| \int_{\Omega \times \Omega} (y - x) \cdot \xi(y) d\mu(x,y) \right| \leq \left( \int_{\Omega \times \Omega} |x - y|^2 d\mu(x,y) \right)^{1/2} \times \left( \int_\Omega u_k(y) |\xi(y)|^2 dy \right)^{1/2}
$$

(2.25)

$$
\leq \left( \|u_k\|_{L^\infty} \right)^{1/2} \|\xi\|_{L^2(\Omega)},
$$

Note that (2.22) now entails

$$
\int_{\Omega \times \Omega} (y - x) \cdot \xi(y) d\mu(x,y) + h \int_\Omega (u_k \nabla \Psi_k + \nabla f_k(u_k)) \cdot \xi dx = 0,
$$

(2.26)

for all $\xi \in C^\infty_c(\Omega; \mathbb{R}^N)$. In view of (2.25) and (2.26) we obtain

$$
\|\nabla f_k(u_k)\|_{L^2(\Omega)}^2 \leq \frac{C}{h^2} d(v_{k-1}, u_k)^2 + C \|\nabla \Psi\|_{L^\infty(\Omega)}^2,
$$

(2.27)

for $C > 0$ independent of $k, h$. The only thing that remains to be proved is that (2.18) holds for all $\zeta \in C^\infty_c(\Omega)$ (because (ii) is an obvious consequence of (i)). We refer the reader to [4], [8], [9] for the existence of a convex function $\Phi : \Omega \to \mathbb{R}$ which a.e. satisfies $\nabla \Phi \in \partial\zeta$ and

$$
\int_{\Omega \times \Omega} \varphi(x,y) d\mu(x,y) = \int_\Omega u_k(y) \varphi(\nabla \Phi(y), y) dy \text{ for all } \varphi \in C(\bar{\Omega} \times \bar{\Omega}).
$$

Applying this to $\varphi(x,y) := x \cdot \xi(y)$ gives, together with (2.26),

$$
\frac{1}{h} \left( \nabla \Phi - \text{id}_\Omega \right) u_k - u_k \nabla \Psi_k - \nabla f_k(u_k) = 0 \text{ a.e. in } \Omega.
$$

(2.28)
Next we integrate back against any \( \xi \in C^\infty(\bar{\Omega}; \mathbb{R}^N) \) and, according to (2.8), (2.23) and (2.26), the proof is complete once we note that (2.23) remains valid for any \( \xi \in C^\infty(\bar{\Omega}) \) (not necessarily compactly supported in \( \Omega \)).

Remark: We have seen that \( f_k(u_k) \in W^{1,\infty}(\Omega) \). Therefore, they are Lipschitz functions. Due to (f1) we can, in fact, deduce that the minimizers themselves are regular. More specifically, the \( u_k \)'s themselves are Lipschitz functions if \( \omega = 2 \) or Hölder continuous functions of uniform Hölder exponent if \( \omega > 2 \).

3 Convergence to the weak solution

We group in this section a collection of results needed in the rest of the proof. Key to the following is the inequality in the first Lemma, which is an extension of a known result, see for example [11], [7] or [12], to take into consideration the non-homogeneity of the equations we consider.

Lemma 3. There exists a constant \( C > 0 \) independent of \( h \) such that

\[
\sum_{k=1}^{n-1} d(v_{k-1}^h, u_k^h)^2 \leq Ch,
\]

where \( n \geq 1 \) is integer and \( nh = T \).

Proof: Corollary 2 yields:

\[
\frac{1}{2h}d(v_{k-1}^h, u_k^h)^2 \leq \int_{\Omega} [\phi_k(v_{k-1}^h) - \phi_k(u_k^h)] dx + \int_{\Omega} (v_{k-1}^h - u_k^h) \Psi_k dx.
\]

Since the \( \phi_k \) are convex we have, due to (2.8):

\[
\phi_k(v_{k-1}^h(x)) \leq \phi_k(u_{k-1}^h(x)) + (\phi_k)'(u_{k-1}^h(x)) \int_{(k-1)h}^{kh} g(x, \tau, u_{k-1}^h(x)) d\tau.
\]

This implies

\[
\sum_{k=1}^{n-1} \int_{\Omega} [\phi_k(v_{k-1}^h) - \phi_k(u_k^h)] dx \leq \sum_{k=1}^{n-1} \int_{\Omega} [\phi_k(u_{k-1}^h) - \phi_k(u_k^h)] dx + \phi_k(M) \sum_{k=1}^{n-1} \int_{(k-1)h}^{kh} g(\cdot, \cdot, u_{k-1}^h) dx d\tau \leq \sum_{k=1}^{n-1} \int_{\Omega} [\phi_k(u_{k-1}^h) - \phi_k(u_k^h)] dx + \phi_k(M) \int_{0}^{T-h} g(\cdot, \cdot, u^h) dx d\tau,
\]

where \( M > 0 \) is a common upper bound for all \( u_k^h \)'s and \( v_k^h \)'s. We also have

\[
\sum_{k=1}^{n-1} \int_{\Omega} [\phi_k(u_{k-1}^h) - \phi_k(u_k^h)] dx = \sum_{k=1}^{n-1} \int_{\Omega} [\phi_k(u_{k-1}^h) - \phi_{k-1}(u_{k-1}^h)] dx + \sum_{k=1}^{n-1} \int_{\Omega} [\phi_{k-1}(u_{k-1}^h) - \phi_k(u_k^h)] dx \leq \sum_{k=1}^{n-1} \int_{\Omega} \frac{1}{k} \int_{(k-1)h}^{kh} [\phi(\tau, u_{k-1}^h) - \phi(\tau + h, u_{k-1}^h)] d\tau dx + \int_{\Omega} [\phi_0(u_0^h) - \phi_N(u_N^h)] dx \leq \sum_{k=1}^{n-1} \int_{\Omega} \frac{1}{h} \int_{(k-1)h}^{kh} Ch d\tau dx + C \leq C'.
\]
This and the previous result give us a bound on the first term. Now we estimate
\[ \sum_{k=1}^{n-1} \int_{\Omega} (u_{k-1}^h - u_k^h) \Psi_k \, dx = \sum_{k=1}^{n-1} \int_{\Omega} (u_{k-1}^h - u_k^h) \Psi_k \, dx + \sum_{k=1}^{n-1} \int_{(k-1)h}^{kh} g(\cdot, t, u_{k-1}^h) \Psi_k \, dt \, dx \leq \]
\[ \leq \sum_{k=1}^{n-1} \int_{\Omega} (\Psi_k - \Psi_{k-1}) u_k^h \, dx + \sum_{k=1}^{n-1} \int_{\Omega} (\Psi_{k-1} u_{k-1}^h - \Psi_k u_k^h) \, dx + \sum_{k=1}^{n-1} \int_{\Omega} Ch(1 + M) \|\Psi\|_\infty \, dx \leq \]
\[ \leq MLT|\Omega| + \int_{\Omega} (\Psi_0 u_0^h - \Psi_n u_n^h) \, dx + CT(1 + M) \|\Psi\|_\infty |\Omega|. \]
The last two equations, together with (g1), give by summation
\[ \frac{1}{2h} \sum_{k=1}^{n-1} d(u_{k-1}^h, u_k^h)^2 \leq C, \]
for $C > 0$ independent of $h$. \hfill \Box

Since $\{u^h\}$ is bounded in $L^\infty(Q)$ we also deduce that
\[ u^h \rightharpoonup u \text{ weakly }^{*} \text{ in } L^\infty(Q) \text{ as } h \downarrow 0. \] (3.2)

Assume now that $f$ and $g$ are affine in $u$. More precisely, consider the problem:
\[ \begin{cases}
    u_t - \nabla \cdot (u \nabla \Psi) - \beta \Delta u = a(x,t)u + b(x,t) \text{ in } \Omega \times (0,T), \\
    \frac{\partial u}{\partial \nu} + \beta u = 0 \quad \text{on } \partial \Omega \times (0,T),
\end{cases} \] (3.3)

We still need all the conditions on $f$ and $g$ to remain imposed so we assume $\beta > 0$, $0 \leq a(x,t)$, $b(x,t) \leq M < \infty$ for all $(x,t) \in Q$. We next develop some useful tools to prove:

**Proposition 2.** The problem (3.3) admits a weak solution for nonnegative and bounded initial data.

Let us start by defining $f^h : (0,\infty) \times \mathbb{R} \rightarrow \mathbb{R}$ by
\[ f^h(t, s) := f_k^h(s) \text{ whenever } kh \leq t < (k+1)h. \] (3.4)

Now, according to (2.27), (3.1) and Proposition 1, we have
\[ \int_h^T \int_{\Omega} |\nabla f^h(t, u^h(x,t))|^2 \, dx \, dt = \sum_{k=1}^{n-1} \int_{kh}^{(k+1)h} \left( \int_{\Omega} |\nabla f_k^h(u_k^h)|^2 \, dx \right) \, dt = \]
\[ = h \sum_{k=1}^{n-1} \int_{\Omega} |\nabla f_k^h(u_k^h)|^2 \, dx \leq \frac{C}{h} \sum_{k=1}^{n-1} d(u_{k-1}^h, u_k^h)^2 + C(n-1)h \|\Psi\|^2_\infty \leq C', \]
(as $(n-1)h = T - h < T$ for some $C' > 0$ independent of $h$).

Let now $G^h$ be the extension by zero of $\nabla f^h(\cdot, u^h)$ (which is defined on $\Omega \times (h,T)$) to $Q$. Then $\{G^h\}_{h>0}$ is bounded in $L^2(Q)$, which implies there exists $G \in L^2(Q; \mathbb{R}^N)$ such that, up to a subsequence (not relabelled)
\[ G^h \rightharpoonup G \text{ weakly } \text{ in } L^2(Q; \mathbb{R}^N). \] (3.5)

The following lemma is also crucial to the proof of the main result of the paper (see Theorem 1; next section!)

**Lemma 4.** Let $\varepsilon > 0$ be given. Then, for sufficiently small $h > 0$, the following holds:
\[ \left| - \int_h^{T-h} \int_{\Omega} u^h(x,t) \frac{1}{h} \left( \zeta(x,t+h) - \zeta(x,t) \right) dx \, dt - \frac{1}{h} \int_0^h \int_{\Omega} u_0(x) \zeta(x,t+h) dx \, dt \right| + \]
\[ + \int_h^T \int_{\Omega} \{ \nabla f^h(t, u^h) + u^h \nabla \Psi \} \cdot \nabla \zeta \, dx \, dt + \frac{1}{h} \int_{T-h}^T \int_{\Omega} u^h \zeta \, dx \, dt - \]
\[ \int_h^T \int_{\Omega} u^h \zeta \, dx \, dt = \]
\[ = \int_h^T \int_{\Omega} u^h \zeta \, dx \, dt - \int_h^{T-h} \int_{\Omega} u^h \zeta \, dx \, dt - \frac{1}{h} \int_0^h \int_{\Omega} u_0(x) \zeta(x,t+h) dx \, dt + \]
\[ + \int_h^T \int_{\Omega} \{ \nabla f^h(t, u^h) + u^h \nabla \Psi \} \cdot \nabla \zeta \, dx \, dt + \frac{1}{h} \int_{T-h}^T \int_{\Omega} u^h \zeta \, dx \, dt - \]
\[ \int_h^T \int_{\Omega} u^h \zeta \, dx \, dt. \]
for all \( \zeta \in C_0^\infty(\widehat{\Omega} \times [0, T]) \).

Proof: Following [1], [11], [14], we consider \( \zeta = \zeta(x, t) \in C_0^\infty(\widehat{\Omega} \times [0, T]) \) in (2.18), integrate over \([kh, (k + 1)h]\) and sum for \( k = 1, ..., n - 1 \). However, as shown in [20], these operations do not lead directly to (3.6) and it takes a couple of steps to get there. We basically need to show how \( O(h) \) and \( \varepsilon \) come into play. For this, note that by doing the operations suggested to obtain (3.6), the terms containing \( g \) add up to:

\[
-\sum_{k=1}^{n-1} \int_{kh}^{(k+1)h} \frac{1}{h} \int_{(k-1)h}^{kh} g(x, t, \tau, u_{i,k-1}(x)) \zeta(x, t) \, d\tau \, dt.
\]

Exactly as in [20], let us now write each term as:

\[
\frac{1}{h} \int_{(k-1)h}^{kh} g(x, t, \tau, u_{i,k-1}(x)) \zeta(x, t) \, d\tau = T_{1,k} + T_{2,k},
\]

where

\[
T_{1,k} := \frac{1}{h} \int_{(k-1)h}^{kh} \{g(x, \tau, u_{i,k-1}(x)) - g(x, t - h, u_{i,k-1}(x))\} \zeta(x, t) \, d\tau,
\]

and

\[
T_{2,k} := g(x, t - h, u_{i,k-1}(x)) \zeta(x, t),
\]

If \( h \) is sufficiently small, we have, due to (4.4),

\[
|g(x, t - h, u_{i,k-1}(x))| \leq \varepsilon / T,
\]

as \( t \in [kh, (k + 1)h] \) implies \( |\tau - t + h| \leq h \). Finally, integrating the \( T_{1,k} \)'s over \([kh, (k + 1)h]\) and adding we obtain a quantity whose absolute value is less than \((n - 1)\varepsilon / T < \varepsilon \). As for the \( T_{2,k} \)'s, note that for all \( t \in [kh, (k + 1)h] \), we have \( g(x, t - h, u_{i,k-1}(x)) = g(x, t - h, u^h(x, t - h)) \). Therefore, by integrating over \([kh, (k + 1)h]\) and adding we obtain, upon changing the variable \( t \rightarrow t - h \):

\[
\sum_{k=1}^{n-1} \int_{kh}^{(k+1)h} T_{2,k} \, dt = \int_0^{T-h} g(x, t, u^h(x, t)) \zeta(x, t + h) \, dt.
\]

Note that this is close to what we want, i.e. \( \int_0^{T-h} g(x, t, u^h(x, t)) \zeta(x, t) \, dt \), within \( O(h) \) distance. Also, note that the terms containing \( \Psi \) do not simply add up to what is displayed in (3.6). However, the argument for the latter terms is similar to the one above. Thus, (3.6) is proved.

Proof of Proposition 2: Fix \( t \in (0, T) \) and notice that \( u^h(\cdot, t) = u^h_k \) for \( k := \left\lfloor \frac{t}{h} \right\rfloor \). We know that all \( u^h_k \)'s (since \( f(u^h_k) \) is in \( H^1 \) and \( f \) is affine) lie in \( H^1(\Omega) \) for \( k \geq 1 \) and so, if we pick any \( \xi \in C_0^\infty(Q) \), we can write (we have equality for the spatial integrals and then integrate in \( t \)):

\[
\int_0^T \int_{\Omega} u^h_k \xi \, dx \, dt = -\int_0^T \int_{\Omega} u^h \xi \, dx \, dt.
\]

It follows, in view of (3.2) and (3.5), that

\[
\int_Q (\beta u \xi_i + G_i \xi) \, dx \, dt = 0, \quad i = 1, ..., N,
\]

where \( G_i \) are the components of \( G \) defined in (3.5). This, along with the arbitrariness of \( \xi \in C_0^\infty(Q) \) imply that \( G \) is the distributional spatial gradient of \( \beta u \) and so (3.5) becomes

\[
C^h \rightarrow \beta \nabla u \text{ weakly in } L^2(Q; \mathbb{R}^N).
\]

We now have all the ingredients for passing to the limit in the left hand side of (3.6) if both \( f \) and \( g \) are affine in \( u \). Therefore \( u \) is a weak solution. \( \square \)

What happens if at least one of \( f \) and \( g \) is not affine? The next section deals with this issue.
4 The main existence result

The weak ∗ convergence result (3.2) that follows from the uniform bound of $u^h$ in $L^\infty(Q)$ is not enough to show convergence for (3.6) as $h \downarrow 0$. The nonlinearity of the problem requires stronger convergence; for that purpose we show, in Proposition 3, that $\{u^h\}_{h \downarrow 0}$ is relatively compact in $L^1(Q)$.

**Proposition 3.** Assume (f1)-(f3), and (g1)-(g3) hold. Then $\{u^h\}_{h \downarrow 0}$ is relatively compact in $L^1(Q)$.

**Proof:** Let $U^h : \mathbb{R}^{N+1} \to [0, \infty)$ be the extension by 0 of $u^h$ outside $Q$. Let $w^h(x,t) := f^h(t,u^h(x,t))$. We shall apply the Riesz–Fréchet–Kolmogorov criterion (see, e.g. [3]) to prove that $\{U^h\}_{h \downarrow 0} |_Q = \{u^h\}_{h \downarrow 0}$ is relatively compact in the strong topology of $L^1(Q)$. We have

$$\int_Q |U^h(x,e,t + \tau) - U^h(x,t)| dx dt \leq$$

(4.1)

for that purpose we show, in Proposition 3, that $\{U^h\}_{h \downarrow 0}$ is relatively compact in $L^1(Q)$.

The idea is to show that the left hand side of (4.1) is bounded by some quantity that tends to vanish independently of $h$ as $\tau$ and $|e|$ tend to vanish. As we can see, there are two different issues that need be addressed; precompactness in space and precompactness in time. We will deal with them separately.

**Lemma 5.** (Precompactness in space) There exists a function $l = l(e)$ such that $\lim_{|e| \downarrow 0} l(e) = 0$ and

$$\int_\Omega |U^h(x,e,t) - U^h(x,t)| dx dt \leq l(e)$$

for all $\tau$, $h$.

**Proof:** We have

$$\int_\Omega |U^h(x,e,t) - U^h(x,t)| dx dt = \int_\Omega |u^h(x,e,t) - u^h(x,t)| dx dt + \int_\Omega u^h(x,t) dx dt.$$

From Lemma 2, $u^h$ is uniformly essentially bounded. Since $\text{meas}(\Omega \setminus (\Omega - e)) \to 0$ as $|e| \to 0$ and we infer that the second term of the inequality above tends to 0 as $|e| \downarrow 0$ uniformly in $\tau$. We turn now to the first term. According to (3.5) we have

$$\int_\Omega |\nabla u^h|^2 dx dt \leq C^2 \frac{1}{h} \sum_{k=1}^{n-1} d(v^h_{k-1}, u^h_k)^2 + nh \|\Psi\|_\infty \leq C',$$

for some $C' > 0$ independent of $h$. If $\tau \geq h$, we have

$$\int_\Omega |w^h(x,e,t) - w^h(x,t)| dx dt \leq |e| \int_\Omega |\nabla w^h(x,e,t)| dx dt ds \leq |e| \int_\Omega |\nabla w^h(x,e,t)| dx dt ds,$$

where $\Omega_e := ([\Omega + se] \cap (\Omega - (1-s)e)) \subset \Omega$ for sufficiently small $|e|$ and for $s \in [0,1]$ (since $\Omega$ is convex). Since $Q$ is bounded (thus the $L^1$ norm is bounded from above by the $L^2$ norm) the previous inequality chains and (3.5) implies

$$\int_\Omega |w^h(x,e,t) - w^h(x,t)| dx dt \leq C|e|,$$

(4.2)
for some constant \( C > 0 \). If \( \tau < h \) all we need to observe is
\[
\int_\tau^T \int_{\Omega \cap (\Omega+e)} |u^h(x+e,t) - w^h(x,t)|dxdt =
\]
\[
= \int_h^T \int_{\Omega \cap (\Omega+e)} |w^h(x+e,t) - w^h(x,t)|dxdt + (h - \tau) \int_{\Omega \cap (\Omega+e)} |w_0(x+e) - w_0(x)|dx.
\]
(Note \( w_0^h = w_0 = f(u_0) \) does not depend on \( h \).) While we treat the first integral as before and obtain something similar to (4.2), we know that the second integral tends to zero as \(|e| \to 0\). We end the proof by noting that, if \( kh < t \leq (k+1)h \), we have:
\[
|w^h(x+e,t) - w^h(x,t)| = |f_k(u_k^h(x+e)) - f_k(u_k^h(x))| =
\]
\[
= \frac{1}{h} \int_{kh}^{(k+1)h} \left| f(t, u_k^h(x+e)) - f(t, u_k^h(x)) \right| dt =
\]
\[
= |f(t_k, u_k^h(x+e)) - f(t_k, u_k^h(x))| \geq c|u_k^h(x+e) - u_k^h(x)|^{\alpha-1},
\]
where \( t_k \) is some \( t \) in \((kh, (k+1)h)\).
\[\square\]

We now prove

**Lemma 6.** (Precompress in time) There exists a function \( q = q(\tau) \) such that \( \lim_{\tau \to 0} q(\tau) = 0 \) and
\[
\int_0^{T-\tau} \int_{\Omega} |u^h(x, t+\tau) - u^h(x,t)|dxdt \leq q(\tau).
\]

**Proof:** In order to do this let us fix an \( h > 0 \) and let \( \tau = jh + \gamma \) with \( 0 \leq \gamma < h \) so that:
\[
T - \tau = nh - jh - \gamma = (n-j)h - \gamma < (n-j)h
\]
and estimate
\[
\int_0^{T-\tau} \int_{\Omega} |u^h(x, t+\tau) - u^h(x,t)|dxdt.
\]

Now, thanks to (f1) and the definitions of \( u^h \) and \( f^h \), in \( kh \leq t < (k+1)h \), we have
\[
\int_0^{T-\tau} \int_{\Omega} |u^h(x, t+\tau) - u^h(x,t)|dxdt \leq \sum_{k=0}^{n-j-1} \int_{kh}^{(k+1)h} \left| u^h(x, t+\tau) - u^h(x,t) \right| dxdt \leq
\]
\[
\leq C \sum_{k=0}^{n-j-1} \int_{kh}^{(k+1)h} \left( |u^h(x, t+\tau) - u^h(x,t)| f(t, u^h(x, t+\tau)) - f(t, u^h(x,t)) \right) dx =
\]
\[
= C \sum_{k=0}^{n-j-1} \left\{ \langle h - \gamma \rangle \int_{u_k^{h, j+1} - u_k^{h, j}} (f_k(u_k^{h, j}) - f_k(u_k^{h,j+1})) dx + \gamma \int_{u_k^{h, j+1} - u_k^{h, j}} (f_k(u_k^{h, j+1}) - f_k(u_k^{h,j})) dx \right\}
\]
(4.4)

Also
\[
(u_k^{h,j+1} - u_k^{h,j})(f_k(u_k^{h,j}) - f_k(u_k^{h,j+1})) = (u_k^{h,j+1} - u_k^{h})(f_k(u_k^{h,j}) - f_k(u_k^{h,j+1})) + (u_k^{h,j} - u_k^{h})(f_k(u_k^{h,j+1}) - f_k(u_k^{h,j})) =
\]
\[
= (u_k^{h,j+1} - u_k^{h})(f_k(u_k^{h,j}) - f_k(u_k^{h})) + \frac{u_k^{h,j+1} - u_k^{h,j}}{h} \int_{kh}^{(k+1)h} \{ f(t, u_k^{h,j}) - f(t + jh, u_k^{h,j+1}) \} dt \leq
\]
\[
\leq (u_k^{h,j+1} - u_k^{h})(f_k(u_k^{h,j}) - f_k(u_k^{h,j+1})) + 2M jhL(f; I),
\]
where \( I := [0, M] \) and \( L(f; I) \) is (according to (f2)) the uniform Lipschitz constant of \( f(\cdot, s) \) for \( s \in I \) (\( M \) being the uniform upper bound for all \( u_k^{h} \)). Next we claim the following:
\[
\sum_{k=0}^{n-j-1} \int_{\Omega} (u_k^{h,j+1} - u_k^{h})(f_k(u_k^{h,j}) - f_k(u_k^{h})) dx \leq C j + C \sqrt{\frac{\gamma}{h}}
\]
(4.5)
for all $j \geq 1$ and a constant $C > 0$ (independent of $k$, $q$ and $h$). To prove (4.5), we know that, for any $i \geq 1$, there exists a convex potential $\Phi$ with $\nabla \Phi: \Omega \rightarrow \Omega$ onto that satisfies
\[
d(\bar{v}_{k+i-1}^h, \bar{u}_{k+i}^h)^2 = \int_{\Omega} \bar{u}_{k+i}^h(y)|y - \nabla \Phi(y)|^2 dy,
\]
(4.6)
\[
\int_{\Omega} \bar{v}_{k+i-1}^h(x)\zeta(x)dx = \int_{\Omega} \bar{u}_{k+i}^h(y)(\nabla \Phi(y))dy \text{ for all } \zeta \in C(\bar{\Omega}).
\]
(4.7)
As in [14], let $\zeta := f_{k+j}(u_{k+j}^h) - f_k(u_k^h)$. Due to (4.7) we have
\[
\int_{\Omega} (\bar{u}_{k+i}^h - \bar{v}_{k+i-1}^h)(f_{k+j}(u_{k+j}^h) - f_k(u_k^h))dx = \int_{\Omega} \bar{u}_{k+i}^h(y)(\zeta(y) - (\nabla \Phi(y)))dy \leq
\]
\[
\leq \int_0^1 \left( \int_{\Omega} \bar{u}_{k+i}^h(y)|\nabla \zeta((1-t)y + t\nabla \Phi(y))|^2 dy \right)^{1/2} dt \left( \int_{\Omega} \bar{u}_{k+i}^h(y)|y - \nabla \Phi(y)|^2 dy \right)^{1/2}.
\]
This is where Otto’s lemma 3 ([14]; Appendix) applies to $|\nabla \zeta|$ (recall that the $u_k^h$’s are Lipschitz continuous). Thus, using (2.27) and (4.6), we discover
\[
\int_{\Omega} (\bar{u}_{k+i}^h - \bar{v}_{k+i-1}^h)(f_{k+j}(u_{k+j}^h) - f_k(u_k^h))dx \leq
\]
\[
\leq \frac{C}{h} \left[ d(\bar{u}_{k+i}^h, \bar{v}_{k+i-1}^h) + d(\bar{u}_{k+j}^h, u_k^h) + 2h\|\nabla \Phi\|_{\infty} \right]
\]
for some $C > 0$ independent of $k$, $i$, $h$. Next we claim there exists $C > 0$ independent of $k$, $j$, $h$ such that
\[
\int_{\Omega} (u_{k+j}^h - \bar{u}_{k+i}^h)(f_{k+j}(u_{k+j}^h) - f_k(u_k^h))dx \leq
\]
\[
\leq \frac{C}{h} \left[ d(\bar{u}_{k+j}^h, u_k^h) + d(\bar{u}_{k+j}^h, \bar{v}_{k+j-1}^h) + 2h\|\nabla \Phi\|_{\infty} \right] \sum_{i=1}^{j} d(\bar{u}_{k+i}^h, \bar{v}_{k+i-1}^h) + Cjh.
\]
(4.9)
To prove this claim, recall that $\zeta = u_{k+j}^h - \bar{u}_{k+i}^h$ and note that
\[
(u_{k+j}^h - \bar{u}_{k+i}^h)\zeta = \sum_{i=1}^{j} (u_{k+i}^h - \bar{v}_{k+i-1}^h)\zeta + \sum_{i=1}^{j} (u_{k+i-1}^h - \bar{u}_{k+i-1}^h)\zeta.
\]
We know, by (2.8), (g1), (g2) and the uniform boundedness of $u_k^h$ that $\bar{v}_{k+i-1}^h - u_{k+i-1}^h = O(h)$. We also take into account that $\{\bar{u}_{k+i}^h\}$, is bounded in $L^\infty$. Therefore,
\[
\int_{\Omega} (\bar{v}_{k+i-1}^h - u_{k+i-1}^h)(\bar{u}_{k+i}^h - \bar{u}_k^h)dx \leq C h.
\]
Summing up (4.8) and the inequalities above for $i = 1,..j$ we prove (4.9). Now write (4.9) for $k = 1,..n-j-1$ and add to obtain
\[
\sum_{k=1}^{n-j-1} \int_{\Omega} (u_{k+j}^h - \bar{u}_{k+i}^h)(u_{k+j}^h - \bar{u}_k^h)dx \leq
\]
\[
\leq \sum_{k=1}^{n-j-1} \frac{C}{h} \left[ d(u_k^h, \bar{v}_{k-1}^h) + d(u_{k+j}^h, \bar{v}_{k+j-1}^h) + 2h\|\nabla \Phi\|_{\infty} \right] \sum_{i=1}^{j} d(u_{k+i}^h, \bar{v}_{k+i-1}^h) + Cjh(n - j - 1).
\]
As $h(n - j - 1) \leq T$, all that remains to be proved is
\[
\sum_{k=1}^{n-j-1} \left[ d(u_k^h, \bar{v}_{k-1}^h) + d(u_{k+j}^h, \bar{v}_{k+j-1}^h) \right] \sum_{i=1}^{j} d(u_{k+i}^h, \bar{v}_{k+i-1}^h) \leq Cjh
\]
and
\[
2\|\nabla \Phi\|_{\infty} \sum_{i=1}^{j} d(u_{k+i}^h, \bar{v}_{k+i-1}^h) \leq C\sqrt{h}.
\]
In order to prove the first, rewrite the left hand side as
\[
\sum_{i=1}^j \sum_{k=1}^{n-j-1} [d(u_{k+i}, v_{k+i-1})d(u_{k+1}, v_{k-1}) + d(u_{k+i}, v_{k+i-1})d(u_{k+i+j}, v_{k+j-1})] \leq \frac{1}{2} \sum_{i=1}^j \sum_{k=1}^{n-j-1} [2d(u_{k+i}, v_{k+i-1})^2 + d(u_{k+i+j}, v_{k+j-1})^2 + d(u_{k+i}, v_{k+i-1})^2].
\]

It is easy to see that (3.1) gives the second inequality. Applying (4.5), (for \(j\) and \(j + 1\)) in (4.4) concludes our proof.

**Proof of Proposition 3 (conclusion):** We finally estimate (see (4.1)) the term \(\int_{T-\epsilon}^T \int_Q u^h(x, t) dx dt\). Trivially
\[
\int_{T-\epsilon}^T \int_Q u^h(x, t) dx dt \leq M \tau \text{ meas}(\Omega).
\]

This last result, together with Lemmas 5 and 6, is enough to show that the Riesz-Fréchet-Kolmogorov criterion (see, e.g. [3]) applies to give the desired \(L^1\) precompactness. Thus, Proposition 3 is proved.

Therefore, \(\{u^h\}\) is precompact in \(L^1(Q)\) and, up to a subsequence, we can pass to the limit in (3.6) to infer existence for (\(P\)). We may thus prove:

**Theorem 1.** Assume (f1) - (f3), (g1) - (g4) and (2.10). Then the problem (\(P\)) admits a nonnegative essentially bounded weak solution provided that \(\Omega\) is bounded and convex and the initial \(u_0\) is nonnegative and essentially bounded.

**Proof:** In order to be able to pass to the limit in (3.6) we need to show \(\nabla f(\cdot, u) \in L^2(Q; \mathbb{R}^N)\) and
\[
G^h \rightharpoonup \nabla f(\cdot, u) \text{ weakly in } L^2(Q; \mathbb{R}^N),
\]
i.e. \(G = \nabla f(\cdot, u)\) (see (3.5)). Note that this would follow from \(f^h(\cdot, u^h) \rightharpoonup f(\cdot, u)\) strongly in \(L^2(Q)\) exactly as in the proof of Proposition 2. Trivially
\[
||f^h(t, u^h) - f(t, u^h)||_{L^2(Q)} \leq ||f^h(t, u^h) - f(t, u^h)||_{L^2(Q)} + ||f(t, u^h) - f(t, u)||_{L^2(Q)}.
\]

As the second term is bounded and tends pointwise to zero (up to a subsequence not relabelled), all we need to show is
\[
\lim_{h \to 0} \int_Q |f^h(\cdot, u^h) - f(\cdot, u^h)|^2 dx dt = 0.
\]

Due to the boundedness of the integrand it suffices to show that it tends to zero in \(L^1(Q)\). We have
\[
\int_Q |f^h(t, u^h) - f(t, u^h)| dx dt \leq \frac{1}{h} \int_0^1 \sum_{k=0}^{n-1} \int_{tk}^{(k+1)h} \left( \int_{tk}^{(k+1)h} f(\tau, u^h(x)) - f(t, u^h(x)) d\tau \right) dt dx.
\]

As \(f(\cdot, s)\) is Lipschitz continuous in \(t\) with respect to \(s\) in bounded domains we infer:
\[
\int_Q |f^h(t, u^h) - f(t, u^h)| dx dt \leq C h,
\]

which brings the proof to its conclusion.

**Remark:** It is interesting to see what happens if we drop the “strong monotonicity” requirement (f1). Following the proof of Proposition 3 we discover that (f1) was only used to bound from above quantities of the type \(|u^h(t + \tau, x + e) - u^h(x, t)|^{\omega - 1}\) by \(|f^h(t + \tau, u^h(t + \tau, x + e)) - f^h(t, u^h(x, t))|\). In fact, the rest of the proof shows, without employing (f1) in any form, that \(\{f^h(\cdot, u^h)\}_h\) is relatively compact in \(L^1(Q)\). It follows that, up to a subsequence (not relabelled),
\[
f^h(\cdot, u^h) \rightarrow V \text{ strongly in } L^1(Q)
\]
for some nonnegative $V \in L^1(Q)$. We have seen in the proof of Theorem 1 that this along with (f2) imply $V^h := f(\cdot, u^h) \to V$ in $L^1(Q)$. Due to (f3), the operator $T : L^2(Q) \to L^2(Q)$ given by $(Tv)(x,t) := f(t,v(x,t))$ is maximal monotone. The strong convergence of $V^h$ to $V$ together with (3.2) imply $\langle h^h, V^h \rangle_{L^2(Q)} \to \langle u, V \rangle_{L^2(Q)}$ as $h \to 0$. Therefore, $V = f(\cdot, u)$ (see [19]).

This observation allows us to generalize the existence result in [20] to the problem

$$
\begin{cases}
  u_t - \nabla \cdot (u \nabla \Psi(x,t)) - \Delta \theta = a(x,t)u + b(x,t,\theta) & \text{in } \Omega \times (0,T), \\
  u(x,t) \in \mathcal{H}(t,\theta(x,t)) & \text{for a.e. } (x,t) \in Q, \\
  (u \nabla \Psi + \nabla \theta) \cdot \nu = 0 & \text{on } \partial \Omega \times (0,T), \\
  \theta(\cdot,0) = \theta_0 & \text{in } \Omega.
\end{cases}
$$

where $\theta_0$ is the initial datum, $\mathcal{H}$ is a maximal graph with an inverse $f$ satisfying (f2), (f3). Also, $a : \mathbb{R}^{N+1} \to \mathbb{R}$, $b : \mathbb{R}^{N+2} \to \mathbb{R}$ are such that (g1)-(g4) are satisfied by $g(x,t,u) := a(x,t)u + b(x,t, f(t,u))$.

5 Remarks on uniqueness

Uniqueness can be relatively easily proved under some restrictions. For the general case, see [1].

**Proposition 4.** Assume now $g$ is Lipschitz in $u$. Within the hypotheses of Theorem 1 there exists at most one solution for (P) which satisfies $\partial u/\partial t \in L^1(Q)$ provided that $\partial f / \partial s$ is bounded from below away from zero.

**Proof:** Define $\zeta$ as

$$
\zeta_\delta(x,t) := \phi_\delta(f(t,u_1(x,t)) - f(t,u_2(x,t))),
$$

where

$$
\phi_\delta(s) := \begin{cases} 
0 & \text{if } s \leq 0 \\
\frac{s}{\delta} & \text{if } 0 < s \leq \delta \\
1 & \text{if } s \geq \delta
\end{cases}
$$

We can then use, for the difference $u_1 - u_2$ of the solutions in (2.1), a smooth approximation of $\zeta_\delta$ so that we can write

$$
\int_Q \partial_t (u_1 - u_2) \zeta_\delta \, dx \, dt = - \int_Q \left\{ (u_1 - u_2) \nabla \Psi \cdot \nabla \zeta_\delta + \nabla (f(\cdot,u_1) - f(\cdot,u_2)) \cdot \nabla \zeta_\delta (g(\cdot,\cdot,u_1) - g(\cdot,\cdot,u_2)) \right\} \zeta_\delta \, dx \, dt
$$

Call now $Q_\delta := Q \cap \{ 0 < f(\cdot,u_1) - f(\cdot,u_2) < \delta \}$ then, recalling the definition of $\zeta_\delta$, we have

$$
\int_{Q_\delta} \partial_t (u_1 - u_2) \zeta_\delta \, dx \, dt = - \frac{1}{\delta} \int_{Q_\delta} \left\{ (u_1 - u_2) \nabla \Psi \cdot \nabla (f(\cdot,u_1) - f(\cdot,u_2)) + |\nabla (f(\cdot,u_1) - f(\cdot,u_2))|^2 \right\} \, dx \, dt
$$

$$
+ \frac{1}{\delta} \int_{Q_\delta} (g(\cdot,\cdot,u_1) - g(\cdot,\cdot,u_2))(f(\cdot,u_1) - f(\cdot,u_2)) \, dx \, dt + \int_{Q_\delta} \leq \frac{1}{\delta} \left( \int_{Q_\delta} (u_1 - u_2)^2 |\nabla \Psi|^2 \, dx \, dt \right)^{1/2} \| \nabla (f(\cdot,u_1) - f(\cdot,u_2)) \|_{L^2(Q_\delta)}
$$

$$
+ C(1 + M) \int_{Q_\delta} f(\cdot,u_1) - f(\cdot,u_2) \, dx \, dt + L_{\text{lip}}(g) \int_Q (u_1 - u_2)^+ \, dx \, dt,
$$

where $R_\delta := \{ f(\cdot,u_1) - f(\cdot,u_2) \geq \delta \}$. Here we use (g1), (g2) and ($\Psi$1). We recall that solutions of (P) are bounded, $0 \leq u \leq M$, and that in $Q_\delta$ we have $f(\cdot,u_1) - f(\cdot,u_2) < \delta$, therefore the second term in the last inequality is bounded and for $\delta \to 0$ the integral over $Q_\delta$ tends to 0, as the measure of the set $Q_\delta \setminus \{ u_1 = u_2 \}$ tends to 0. Since $\partial f / \partial s > c > 0$, it follows

$$
c|u_1 - u_2| \leq |f(t,u_1) - f(t,u_2)| \leq \delta.
$$

Now noticing that $\{ f(\cdot,u_1) - f(\cdot,u_2) \geq 0 \} = \{ u_1 - u_2 \geq 0 \}$ and letting $\delta \to 0$ gives

$$
\int_Q \partial_t (u_1 - u_2)^+ \, dx \, dt \leq L_{\text{lip}}(g) \int_Q (u_1 - u_2)^+ \, dx \, dt.
$$

18
Therefore, for any $0 < t < T$ we have
\[
\int_{\Omega} (u_1(t) - u_2(t))^+ \, dx \leq \int_{\Omega} (u_1(0) - u_2(0))^+ \, dx + \text{Lip}_u(g) \int_0^t \int_{\Omega} (u_1(\tau) - u_2(\tau))^+ \, dx \, d\tau.
\]
As $u_1(0) = u_2(0)$ in $\Omega$, we apply Gronwall’s Lemma to infer $(u_1 - u_2)^+ = 0$ a.e. in $Q$. Inverting the roles of $u_1$ and $u_2$ concludes the proof. □

Acknowledgements

The authors would like to thank their PhD advisor, David Kinderlehrer, for his valuable suggestions and continual support. This research was partially supported by NSF grant DMS 0305794.

References


