Viorel Barbu

SPRINGER MONOGRAPHS IN MATHEMATICS

Nonlinear Differential Equations of Monotone Types in Banach Spaces

Springer

S*pringer* **M***onographs in* **M***athematics*

Viorel Barbu

Nonlinear Differential Equations of Monotone Types in Banach Spaces

Springer

Viorel Barbu
Fac. Mathematics
Al. I. Cuza University
Blvd. Carol I 11
700506 Iasi
Romania
vb41@uaic.ro

ISSN 1439-7382
ISBN 978-1-4419-5541-8 e-ISBN 978-1-4419-5542-5
DOI 10.1007/978-1-4419-5542-5
Springer New York Dordrecht Heidelberg London

Library of Congress Control Number: 2009943993

Mathematics Subject Classification (2010): 34G20, 34G25, 35A16

Printed on acid-free paper

Springer is part of Springer Science+Business Media (www.springer.com)

Contents

Preface

In the last decades, functional methods played an increasing role in the qualitative theory of partial differential equations. The spectral methods and theory of C_0-semigroups of linear operators as well as Leray–Schauder degree theory, fixed point theorems, and theory of maximal monotone nonlinear operators are now essential functional tools for the treatment of linear and nonlinear boundary value problems associated with partial differential equations.

An important step was the extension in the early seventies of the nonlinear dynamics of accretive (dissipative) type of the Hille–Yosida theory of C_0-semigroups of linear continuous operators. The main achievement was that the Cauchy problem associated with nonlinear m-accretive operators in Banach spaces is well posed and the corresponding dynamic is expressed by the Peano exponential formula from finite-dimensional theory. This fundamental result is the corner stone of the whole existence theory of nonlinear infinite dynamics of dissipative type and its contribution to the development of the modern theory of nonlinear partial differential equations cannot be underestimated.

Previously, in mid-sixties, some spectacular properties of maximal monotone operators and their close relationship with convex analysis and m-accretivity were revealed. In fact, Minty's discovery that in Hilbert spaces nonlinear maximal monotone operators coincide with m-accretive operators was crucial for the development of the theory. Although with respect to the middle and end of the seventies, little new material on this subject has come to light, the field of applications grew up and still remains in actuality. In the meantime, some excellent monographs were published where these topics were treated exhaustively and with extensive bibliographical references. Here, we confine ourselves to the presentation of basic results of the theory of nonlinear operators of monotone type and the corresponding dynamics generated in Banach spaces. These subjects were also treated in the author's books *Nonlinear Semigroups and Differential Equations in Banach Spaces* (Noordhoff, 1976) and *Analysis and Control of Nonlinear Infinite Dimensional Systems* (Academic Press, 1993), but the present book is more oriented to applications. We refrain from an exhaustive treatment or details that would divert us from the principal aim of this book: the presentation of essential results of the theory and its illustration by significant problems of nonlinear partial differential equations. Our aim is to present functional tools for the study of a large class of nonlinear problems and open to the reader the way towards applications. This book can serve as a teaching text for graduate students and it is self-contained. One assumes, however, basic knowledge of real and functional analysis as well as of differential equations. The literature on this argument is so vast and accessible in print that I have dispensed with detailed references or bibliographical comments. I have confined myself to a selected bibliography arranged at the end of each chapter.

The present book is based on a graduate course given by the author at the University of Iaşi during the past twenty years as well as on one-semester graduate courses at the University of Virginia in 2005 and the University of Trento in 2008.

In the preparation of the present book, I have received valuable help from my colleagues, Ioan Vrabie and Cătălin Lefter (Al.I. Cuza University of Iaşi), Gabriela Marinoschi (Institute of Mathematical Statistics and Applied Mathematics of the Romanian Academy) and Luca Lorenzi from University of Parma, who read the preliminary draft of the book and made numerous comments and suggestions which have permitted me to improve the presentation and correct the errors. Elena Mocanu from the Institute of Mathematics in Iaşi has done a great job in preparing and processing this text for printing.

Iaşi, September 2009 *Viorel Barbu*

Acronyms

$\mathbf{R}$	the real line $(-\infty,\infty)$
$\mathbf{R}^N$	the N-dimensional Euclidean space
$\mathbf{R}^+ = (0,+\infty)$,	
$\mathbf{R}^- = (-\infty,0)$,	
$\overline{\mathbf{R}} = (-\infty,+\infty]$,	
$\mathbf{R}^N_+ = \{(x_1,...,x_N);x_N > 0\}$	
Ω	open subset of $\mathbf{R}^N$
$\partial\Omega$	the boundary of Ω
$Q = \Omega \times (0,T)$,	
$\Sigma = \partial\Omega \times (0,T)$,	where $0 < T < \infty$
$\|\cdot\|_X$	the norm of a linear normed space X
X^*	the dual of space X
$L(X,Y)$	the space of linear continuous operators from X to Y
∇f	the gradient of the map $f : X \to \mathbf{R}$
∂f	the subdifferential of $f : X \to \mathbf{R}$
B^*	the adjoint of the operator B
$\overline{C}$	the closure of the set C
$\operatorname{int} C$	the interior of C
$\operatorname{conv} C$	the convex hull of C
sign	the signum function on X : $\operatorname{sign} x = x/\|x\|_X$ if $x \neq 0$ $\operatorname{sign} 0 = \{x;\ \|x\| \leq 1\}$
$C^k(\Omega)$	the space of real-valued functions on Ω that are continuously differentiable up to order k, $0 \leq k \leq \infty$
$C^k_0(\Omega)$	the subspace of functions in $C^k(\Omega)$ with compact support in Ω
$\mathscr{D}(\Omega)$	the space $C^\infty_0(\Omega)$
$\frac{d^k u}{dt^k}$, $u^{(k)}$	the derivative of order k of $u : [a,b] \to X$
$\mathscr{D}'(\Omega)$	the dual of $\mathscr{D}(\Omega)$ (i.e., the space of distributions on Ω)
$C(\overline{\Omega})$	the space of continuous functions on $\overline{\Omega}$

$L^p(\Omega)$	the space of p-summable functions $u:\Omega\to\mathbf{R}$ endowed with the norm $\|u\|_p=(\int_\Omega \lvert u(x)\rvert^p dx)^{1/p}$, $1\le p<\infty$, $\|u\|_\infty=\operatorname{ess\,sup}_{x\in\Omega}\lvert u(x)\rvert$ for $p=\infty$
$L^p_m(\Omega)$	the space of p-summable functions $u:\Omega\to\mathbf{R}^m$
$W^{m,p}(\Omega)$	the Sobolev space $\{u\in L^p(\Omega); D^\alpha u\in L^p(\Omega), \lvert\alpha\rvert\le m, 1\le p\le\infty\}$
$W^{m,p}_0(\Omega)$	the closure of $C^\infty_0(\Omega)$ in the norm of $W^{m,p}(\Omega)$
$W^{-m,q}(\Omega)$	the dual of $W^{m,p}_0(\Omega)$; $(1/p)+(1/q)=1$, $p<\infty$, $q>1$
$H^k(\Omega), H^k_0(\Omega)$	the spaces $W^{k,2}(\Omega)$ and $W^{k,2}_0(\Omega)$, respectively
$L^p(a,b;X)$	the space of p-summable functions from (a,b) to X (Banach space) $1\le p\le\infty$, $-\infty\le a<b\le\infty$
$AC([a,b];X)$	the space of absolutely continuous functions from $[a,b]$ to X
$BV([a,b];X)$	the space of functions with bounded variation on $[a,b]$
$BV(\Omega)$	the space of functions with bounded variation on Ω
$W^{1,p}([a,b];X)$	the space $\{u\in AC([a,b];X);\ du/dt\in L^p([a,b];X)\}$

Chapter 1
Fundamental Functional Analysis

Abstract The aim of this chapter is to provide some standard basic results pertaining to geometric properties of normed spaces, convex functions, Sobolev spaces, and variational theory of linear elliptic boundary value problems. Most of these results, which can be easily found in textbooks or monographs, are given without proof or with a sketch of proof only.

1.1 Geometry of Banach Spaces

Throughout this section X is a real normed space and X^* denotes its dual. The value of a functional $x^* \in X^*$ at $x \in X$ is denoted by either (x^*,x) or $x^*(x)$, as is convenient. The norm of X is denoted by $\|\cdot\|$, and the norm of X^* is denoted by $\|\cdot\|_*$. If there is no danger of confusion we omit the asterisk from the notation $\|\cdot\|_*$ and denote both the norms of X and X^* by the symbol $\|\cdot\|$.

We use the symbol lim or $\to$ to indicate *strong convergence* in X and w-lim or $\rightharpoonup$ for *weak convergence* in X. By w^*-lim or $\rightharpoonup$ we indicate *weak-star* convergence in X^*. The space X^* endowed with the weak-star topology is denoted by X^*_w.

Define on X the mapping $J : X \to 2^{X^*}$:

$$J(x) = \{x^* \in X^*;\ (x^*,x) = \|x\|^2 = \|x^*\|^2\}, \qquad \forall x \in X. \tag{1.1}$$

By the Hahn–Banach theorem we know that for every $x_0 \in X$ there is some $x_0^* \in X^*$ such that $(x_0^*,x_0) = \|x_0\|$ and $\|x_0^*\| \le 1$.

Indeed, the linear functional $f : Y \to \mathbf{R}$ defined by $f(x) = \alpha\|x_0\|$ for $x = \alpha x_0$, where $Y = \{\alpha x_0;\ \alpha \in \mathbf{R}\}$, has a linear continuous extension $x_0^* \in X^*$ on X such that $|(x_0^*,x)| \le \|x\|\ \forall x \in X$. Hence, $(x_0^*,x_0) = \|x_0\|$ and $\|x_0^*\| \le 1$ (in fact, $\|x_0^*\| = 1$). Clearly, $x_0^*\|x_0\| \in J(x_0)$ and so $J(x_0) \ne \emptyset$ for every $x_0 \in X$.

The mapping $J : X \to X^*$ is called the *duality mapping* of the space X. In general, the duality mapping J is multivalued.

V. Barbu, *Nonlinear Differential Equations of Monotone Types in Banach Spaces*,
Springer Monographs in Mathematics, DOI 10.1007/978-1-4419-5542-5_1,

The inverse mapping $J^{-1} : X^* \to X$ defined by $J^{-1}(x^*) = \{x \in X;\ x^* \in J(x)\}$ also satisfies

$$J^{-1}(x^*) = \{x \in X;\ \|x\| = \|x^*\|,\ (x^*,x) = \|x\|^2 = \|x^*\|^2\}.$$

If the space X is reflexive (i.e., $X = X^{**}$), then clearly J^{-1} is just the duality mapping of X^* and so $D(J^{-1}) = X^*$. As a matter of fact, reflexivity plays an important role everywhere in the following and it should be recalled that a normed space is reflexive if and only if its dual X^* is reflexive (see, e.g., Yosida [16], p. 113).

It turns out that the properties of the duality mapping are closely related to the nature of the spaces X and X^*, more precisely to the convexity and smoothing properties of the closed balls in X and X^*.

Recall that the space X is called *strictly convex* if the unity ball B of X is strictly convex, that is the boundary ∂B contains no line segments.

The space X is said to be *uniformly convex* if for each $\varepsilon > 0$, $0 < \varepsilon < 2$, there is $\delta(\varepsilon) > 0$ such that if $\|x\| = 1$, $\|y\| = 1$, and $\|x-y\| \geq \varepsilon$, then $\|x+y\| \leq 2(1-\delta(\varepsilon))$.

Obviously, every uniformly convex space X is strictly convex. Hilbert spaces as well as the spaces $L^p(\Omega)$, $1 < p < \infty$, are uniformly convex spaces (see, e.g., Köthe [13]). Recall also that, by virtue of the Milman theorem (see, e.g., Yosida [16], p. 127), every uniformly convex Banach space X is reflexive. Conversely, it turns out that every reflexive Banach space X can be renormed such that X and X^* become strictly convex. More precisely, one has the following important result due to Asplund [4].

Theorem 1.1. *Let X be a reflexive Banach space with the norm $\|\cdot\|$. Then there is an equivalent norm $\|\cdot\|_0$ on X such that X is strictly convex in this norm and X^* is strictly convex in the dual norm $\|\cdot\|_0^*$.*

Regarding the properties of the duality mapping associated with strictly or uniformly convex Banach spaces, we have the following.

Theorem 1.2. *Let X be a Banach space. If the dual space X^* is strictly convex, then the duality mapping $J : X \to X^*$ is single-valued and demicontinuous (i.e., it is continuous from X to X_w^*). If the space X^* is uniformly convex, then J is uniformly continuous on every bounded subset of X.*

Proof. Clearly, for every $x \in X$, $J(x)$ is a closed convex subset of X^*. Because $J(x) \subset \partial B$, where B is the open ball of radius $\|x\|$ and center 0, we infer that if X^* is strictly convex, then $J(x)$ consists of a single point. Now, let $\{x_n\} \subset X$ be strongly convergent to x_0 and let x_0^* be any weak-star limit point of $\{J(x_n)\}$. (Because the unit ball of the dual space is w^*-compact (Yosida [16], p. 137) such an x_0^* exists.) We have $(x_0^*,x_0) = \|x_0\|^2 \geq \|x_0^*\|^2$ because the closed ball of radius $\|x_0\|$ in X^* is weak-star closed. Hence $\|x_0\|^2 = \|x_0^*\|^2 - (x_0^*,x_0)$. In other words, $x_0^* = J(x_0)$, and so

$$J(x_n) \rightharpoonup J(x_0),$$

as claimed. □

To prove the second part of the theorem, let us first establish the following lemma.

Lemma 1.1. *Let X be a uniformly convex Banach space. If $x_n \rightharpoonup x$ and $\limsup_{n\to\infty} \|x_n\| \le \|x\|$, then $x_n \to x$ as $n \to \infty$.*

Proof. One can assume of course that $x \ne 0$. By hypothesis, $(x^*, x_n) \to (x^*, x)$ for all $x \in X$, and so, by the weak lower semicontinuity of the norm in X,

$$\|x\| \le \liminf_{n\to\infty} \|x_n\| \le \|x\|.$$

Hence, $\lim_{n\to\infty} \|x_n\| = \|x\|$. Now, we set

$$y_n = \frac{x_n}{\|x_n\|}, \qquad y = \frac{x}{\|x\|}\,.$$

Clearly, $y_n \rightharpoonup y$ as $n \to \infty$. Let us assume that $y_n \not\to y$ and argue from this to a contradiction. Indeed, in this case we have a subsequence y_{n_k}, $\|y_{n_k} - y\| \ge \varepsilon$, and so there is $\delta > 0$ such that $\|y_{n_k} + y\| \le 2(1-\delta)$. Letting $n_k \to \infty$ and using once again the fact that the norm $y \to \|y\|$ is weakly lower semicontinuous, we infer that $\|y\| \le 1 - \delta$. The contradiction we have arrived at shows that the initial supposition is false. □

Proof of Theorem 1.2 (continued). Assume now that X^* is uniformly convex. We suppose that there exist subsequences $\{u_n\}, \{v_n\}$ in X such that $\|u_n\|, \|v_n\| \le M$, $\|u_n - v_n\| \to 0$ for $n \to \infty$, $\|J(u_n) - J(v_n)\| \ge \varepsilon > 0$ for all n, and argue from this to a contradiction. We set $x_n = u_n\|u_n\|^{-1}$, $y_n = v_n\|v_n\|^{-1}$. Clearly, we may assume without loss of generality that $\|u_n\| \ge \alpha > 0$ and that $\|v_n\| \ge \alpha > 0$ for all n. Then, as easily seen,

$$\|x_n - y_n\| \to 0 \quad \text{as } n \to \infty$$

and

$$(J(x_n) + J(y_n), x_n) = \|x_n\|^2 + \|y_n\|^2 + (x_n - y_n, J(y_n)) \ge 2 - \|x_n - y_n\|.$$

Hence

$$\frac{1}{2}\,\|J(x_n) + J(y_n)\| \ge 1 - \frac{1}{2}\,\|x_n - y_n\|, \qquad \forall n.$$

Inasmuch as $\|J(x_n)\| = \|J(y_n)\| = 1$ and the space X^* is uniformly convex, this implies that $\lim_{n\to\infty}(J(x_n) - J(y_n)) = 0$. On the other hand, we have

$$J(u_n) - J(v_n) = \|u_n\|(J(x_n) - J(y_n)) + (\|u_n\| - \|v_n\|)J(y_n),$$

so that $\lim_{n\to\infty}(J(u_n) - J(v_n)) = 0$ strongly in X^*. □

Now, let us give some examples of duality mappings.

1. $X = H$ is a Hilbert space identified with its own dual. Then $J = I$, the identity operator in H. If H is not identified with its dual H^*, then the duality mapping $J : H \to H^*$ is the canonical isomorphism Λ of H onto H^*. For instance, if $H = H_0^1(\Omega)$ and $H^* = H^{-1}(\Omega)$ and Ω is a bounded and open subset of $\mathbf{R}^N$, then $J = \Lambda$ is defined by

$$(\Lambda u, v) = \int_\Omega \nabla u \cdot \nabla v \, dx, \qquad \forall u, v \in H_0^1(\Omega). \tag{1.2}$$

In other words, $J = \Lambda$ is the Laplace operator $-\Delta$ under Dirichlet homogeneous boundary conditions in $\Omega \subset \mathbf{R}^N$. Here $H_0^1(\Omega)$ is the Sobolev space $\{u \in L^2(\Omega); \nabla u \in L^2(\Omega);\ u = 0 \text{ on } \partial\Omega\}$. (See Section 1.3 below.)

2. $X = L^p(\Omega)$, where $1 < p < \infty$ and Ω is a measurable subset of $\mathbf{R}^N$. Then, the duality mapping of X is given by

$$J(u)(x) = |u(x)|^{p-2} u(x) \|u\|_{L^p(\Omega)}^{2-p}, \quad \text{a.e. } x \in \Omega,\ \forall u \in L^p(\Omega). \tag{1.3}$$

Indeed, it is readily seen that if Φ_p is the mapping defined by the right-hand side of (1.3), we have

$$\int_\Omega \Phi_p(u) u \, dx = \left(\int_\Omega |u|^p dx \right)^{2/p} = \left(\int_\Omega |\Phi_p(u)|^q dx \right)^{2/q}, \quad \text{where } \frac{1}{p} + \frac{1}{q} = 1.$$

Because the duality mapping J of $L^p(\Omega)$ is single-valued (because L^p is uniformly convex for $p > 1$) and $\Phi_p(u) \in J(u)$, we conclude that $J = \Phi_p$, as claimed. If $X = L^1(\Omega)$, then as we show later (Corollary 2.7)

$$J(u) = \{v \in L^\infty(\Omega);\ v(x) \in \operatorname{sign} u(x) \cdot \|u\|_{L^1(\Omega)}, \text{ a.e. } x \in \Omega\}. \tag{1.4}$$

3. Let X be the Sobolev space $W_0^{1,p}(\Omega)$, where $1 < p < \infty$ and Ω is a bounded and open subset of $\mathbf{R}^N$. (See Section 1.3 below.) Then,

$$J(u) = -\sum_{i=1}^N \frac{\partial}{\partial x_i} \left(\left| \frac{\partial u}{\partial x_i} \right|^{p-2} \frac{\partial u}{\partial x_i} \right) \|u\|_{W_0^{1,p}(\Omega)}^{2-p}. \tag{1.5}$$

In other words, $J : W_0^{1,p}(\Omega) \to W^{-1,q}(\Omega)$, $(1/p) + (1/q) = 1$, is defined by

$$(J(u), v) = \sum_{i=1}^N \int_\Omega \left| \frac{\partial u}{\partial x_i} \right|^{p-2} \frac{\partial u}{\partial x_i} \frac{\partial v}{\partial x_i} \, dx \|u\|_{W_0^{1,p}(\Omega)}^{2-p}, \qquad \forall v \in W_0^{1,p}(\Omega). \tag{1.6}$$

We later show that the duality mapping J of the space X can be equivalently defined as the subdifferential (Gâteaux differential if X^* is strictly convex) of the function $x \to 1/2\,\|x\|^2$.

1.2 Convex Functions and Subdifferentials

Here we briefly present the basic results pertaining to convex analysis in infinite-dimensional spaces. For further results and complete treatment of the subject we refer the reader to Moreau [14], Rockafellar [15], Brezis [8], Barbu and Precupanu [6] and Zălinescu [17].

Let X be a real Banach space with dual X^*. A *proper convex function* on X is a function $\varphi : X \to (-\infty, +\infty] = \overline{\mathbf{R}}$ that is not identically $+\infty$ and that satisfies the inequality

$$\varphi((1-\lambda)x+\lambda y) \le (1-\lambda)\varphi(x)+\lambda\varphi(y) \tag{1.7}$$

for all $x, y \in X$ and all $\lambda \in [0,1]$.

The function $\varphi : X \to (-\infty, +\infty]$ is said to be *lower semicontinuous* (l.s.c.) on X if

$$\liminf_{u\to x} \varphi(u) \ge \varphi(x), \qquad \forall x \in X,$$

or, equivalently, every level subset $\{x \in X;\ \varphi(x) \le \lambda\}$ is closed.

The function $\varphi : X \to]-\infty, +\infty]$ is said to be *weakly lower semicontinuous* if it is lower semicontinuous on the space X endowed with weak topology.

Because every level set of a convex function is convex and every closed convex set is weakly closed (this is an immediate consequence of Mazur's theorem, Yosida [16], p. 109), we may therefore conclude that a proper convex function is lower semicontinuous if and only if it is weakly lower semicontinuous.

Given a lower semicontinuous convex function $\varphi : X \to (-\infty, +\infty] = \overline{\mathbf{R}}$, $\varphi \not\equiv \infty$, we use the following notations:

$$D(\varphi) = \{x \in X;\ \varphi(x) < \infty\} \quad \text{(the effective domain of } \varphi), \tag{1.8}$$

$$\operatorname{Epi}(\varphi) = \{(x,\lambda) \in X \times \mathbf{R};\ \varphi(x) \le \lambda\} \quad \text{(the epigraph of } \varphi). \tag{1.9}$$

It is readily seen that $\operatorname{Epi}(\varphi)$ is a closed convex subset of $X \times \mathbf{R}$, and as a matter of fact its properties are closely related to those of the function φ.

Now, let us briefly describe some elementary properties of l.s.c., convex functions.

Proposition 1.1. *Let $\varphi : X \to \overline{\mathbf{R}}$ be a proper, l.s.c., and convex function. Then φ is bounded from below by an affine function; that is there are $x_0^* \in X^*$ and $\beta \in \mathbf{R}$ such that*

$$\varphi(x) \ge (x_0^*, x) + \beta, \qquad \forall x \in X. \tag{1.10}$$

Proof. Let $E(\varphi) = \operatorname{Epi}(\varphi)$ and let $x_0 \in X$ and $r \in \mathbf{R}$ be such that $\varphi(x_0) > r$. By the classical separation theorem (see, e.g., Brezis [7]), there is a closed hyperplane $H = \{(x,\lambda) \in X \times \mathbf{R};\ -(x_0^*, x) + \lambda = \alpha\}$ that separates $E(\varphi)$ and (x_0, r). This means that

$$-(x_0^*, x) + \lambda \ge \alpha, \qquad \forall x \in E(\varphi) \quad \text{and} \quad -(x_0^*, x_0) + r < \alpha.$$

Hence, for $\lambda = \varphi(x)$, we have

$$-(x_0^*, x) + \varphi(x) \ge -(x_0^*, x_0) + r, \qquad \forall x \in X,$$

which implies (1.10). □

Proposition 1.2. *Let $\varphi : X \to \overline{\mathbf{R}}$ be a proper, convex, and l.s.c. function. Then φ is continuous on* $\operatorname{int} D(\varphi)$.

Proof. Let $x_0 \in \operatorname{int} D(\varphi)$. We prove that φ is continuous at x_0. Without loss of generality, we assume that $x_0 = 0$ and that $\varphi(0) = 0$. Because the set $\{x : \varphi(x) > -\varepsilon\}$ is open it suffices to show that $\{x : \varphi(x) < \varepsilon\}$ is a neighborhood of the origin. We set $C = \{x \in X;\ \varphi(x) \leq \varepsilon\} \cap \{x \in X;\ \varphi(-x) \leq \varepsilon\}$. Clearly, C is a closed balanced set of X (i.e., $\alpha x \in C$ for $|\alpha| \leq 1$ and $x \in C$). Moreover, C is absorbing; that is, for every $x \in X$ there exists $\alpha > 0$ such that $\alpha x \in C$ (because the function $t \to \varphi(tx)$ is convex and finite in a neighborhood of the origin and therefore it is continuous). Because X is a Banach space, the preceding properties of C imply that it is a neighborhood of the origin, as claimed. □

The function $\varphi^* : X^* \to \overline{\mathbf{R}}$ defined by

$$\varphi^*(p) = \sup\{(p,x) - \varphi(x);\ x \in X\} \tag{1.11}$$

is called the *conjugate* of φ.

Proposition 1.3. *Let $\varphi : X \to \overline{\mathbf{R}}$ be l.s.c., convex, and proper. Then φ^* is l.s.c., convex, and proper on the space X^*.*

Proof. As supremum of a set of affine functions, φ^* is convex and l.s.c. Moreover, by Proposition 1.2 we see that $\varphi^* \not\equiv \infty$. □

Proposition 1.4. *Let $\varphi : X \to (-\infty, +\infty]$ be a weakly lower semicontinuous function such that every level set $\{x \in X;\ \varphi(x) \leq \lambda\}$ is weakly compact. Then φ attains its infimum on X. In particular, if X is reflexive and φ is an l.s.c. proper convex function on X such that*

$$\lim_{\|x\| \to \infty} \varphi(x) = \infty, \tag{1.12}$$

then there exists $x_0 \in X$ such that $\varphi(x_0) = \inf\{\varphi(x);\ x \in X\}$.

Proof. Let $d = \inf\{\varphi(x);\ x \in X\}$ and let $\{x_n\} \subset X$ such that $d \leq \varphi(x_n) \leq d + (1/n)$. Then $\{x_n\}$ is weakly compact in X and, therefore, there is $\{x_{n_k}\} \subset \{x_n\}$ such that $x_{n_k} \rightharpoonup x$ as $n_k \to \infty$. Because φ is weakly semicontinuous, this implies that $\varphi(x) \leq d$. Hence $\varphi(x) = d$, as desired. If X is reflexive, then formula (1.12) implies that $\{x \in X;\ \varphi(x) \leq \lambda\}$ are weakly compact. As seen earlier, every convex and l.s.c. function is weakly lower semicontinuous, therefore we can apply the first part. □

Given a function f from a Banach space X to $\mathbf{R}$, the mapping $f' : X \times X \to \mathbf{R}$ defined by

$$f'(x,y) = \lim_{\lambda \downarrow 0} \frac{f(x + \lambda y) - f(x)}{\lambda}, \qquad x, y \in X, \tag{1.13}$$

(if it exists) is called the *directional derivative* of f at x in direction y.

The function $f : X \to \mathbf{R}$ is said to be *Gâteaux differentiable* at $x \in X$ if there exists $\nabla f(x) \in X^*$ (the *Gâteaux differential*) such that

$$f'(x,y) = (\nabla f(x), y), \qquad \forall y \in X. \tag{1.14}$$

If the convergence in (1.13) is uniform in y on bounded subsets, then f is said to be *Fréchet differentiable* and ∇f is called the *Fréchet differential* (derivative) of f.

Given an l.s.c., convex, proper function $\varphi : X \to \overline{\mathbf{R}}$, the mapping $\partial\varphi : X \to X^*$ defined by

$$\partial\varphi(x) = \{x^* \in X^*;\ \varphi(x) \le \varphi(y) + (x^*, x-y),\ \forall y \in X\} \tag{1.15}$$

is called the *subdifferential* of φ.

In general, $\partial\varphi$ is a multivalued operator from X to X^* not everywhere defined and can be seen as a subset of $X \times X^*$.

An element $x^* \in \partial\varphi(x)$ (if any) is called a *subgradient* of φ in x. We denote as usual by $D(\partial\varphi)$ the set of all $x \in X$ for which $\partial\varphi(x) \ne \emptyset$.

Let us pause briefly to give some simple examples.

1. $\varphi(x) = 1/2\,\|x\|^2$. Then, $\partial\varphi = J$ (the duality mapping of the space X). Indeed, if $x^* \in J(x)$, then

$$(x^*, x-y) = \|x\|^2 - (x^*, y) \ge \frac{1}{2}\,(\|x\|^2 - \|y\|^2), \qquad \forall x \in X.$$

Hence $x^* \in \partial\varphi(x)$. Now, let $x^* \in \partial\varphi(x)$; that is,

$$\frac{1}{2}(\|x\|^2 - \|y\|^2) \le (x^* - y, x), \qquad \forall y \in X. \tag{1.16}$$

We take $y = \lambda x,\ 0 < \lambda < 1$, in (1.16), getting

$$(x^*, x) \ge \frac{1}{2}\,\|x\|^2(1+\lambda).$$

Hence, $(x^*, x) \ge \|x\|^2$. If $y = \lambda x$, where $\lambda > 1$, we get that $(x^*, x) \le \|x\|^2$. Hence, $(x^*, x) = \|x\|^2$ and $\|x^*\| \ge \|x\|$. On the other hand, taking $y = x + \lambda u$ in (1.16), where $\lambda > 0$ and u is arbitrary in X, we get

$$\lambda(x^*, u) \le \frac{1}{2}\,(\|x+\lambda u\|^2 - \|x\|^2),$$

which yields

$$(x^*, u) \le \|x\|\,\|u\|.$$

Hence, $\|x^*\| \le \|x\|$. We have therefore proven that $(x^*, x) = \|x\|^2 = \|x^*\|^2$ as claimed.

2. Let K be a closed convex subset of X. The function $I_K : X \to \overline{\mathbf{R}}$ defined by

$$I_K(x) = \begin{cases} 0, & \text{if } x \in K, \\ +\infty, & \text{if } x \notin K, \end{cases} \tag{1.17}$$

is called the *indicator function* of K, and its dual function H,

$$H_K(p) = \sup\{(p,u);\ u \in K\}, \qquad \forall p \in X^*,$$

is called the *support function* of K. It is readily seen that $D(\partial I_K) = K$, $\partial I_K(x) = 0$ for $x \in \operatorname{int} K$ (if nonempty) and that

$$\partial I_K(x) = N_K(x) = \{x^* \in X^*;\ (x^*, x-u) \geq 0,\ \forall u \in K\}, \qquad \forall x \in K. \tag{1.18}$$

For every $x \in \partial K$ (the boundary of K), $N_K(x)$ is the *normal cone* at K in x.

3. Let φ be convex and Gâteaux differentiable at x. Then $\partial\varphi(x) = \nabla\varphi(x)$. Indeed, because φ is convex, we have

$$\varphi(x + \lambda(y-x)) \leq (1-\lambda)\varphi(x) + \lambda\varphi(y)$$

for all $x, y \in X$ and $0 \leq \lambda \leq 1$. Hence,

$$\frac{\varphi(x+\lambda(y-x)) - \varphi(x)}{\lambda} \leq \varphi(y) - \varphi(x),$$

and letting λ tend to zero, we see that $\nabla\varphi(x) \in \partial\varphi(x)$. Now, let w be an arbitrary element of $\partial\varphi(x)$. We have

$$\varphi(x) - \varphi(y) \leq (w, x-y), \qquad \forall y \in X.$$

Equivalently,

$$\frac{\varphi(x+\lambda y) - \varphi(x)}{\lambda} \geq (w,y), \qquad \forall \lambda > 0,\ y \in X,$$

and this implies that $(\nabla\varphi(x) - w, y) \geq 0$ for all $y \in X$. Hence, $w = \nabla\varphi(x)$.

By the definition of $\partial\varphi$ it is easily seen that $\varphi(x) = \inf\{\varphi(u);\ u \in X\}$ iff $0 \in \partial\varphi(x)$. There is a close relationship between $\partial\varphi$ and $\partial\varphi^*$. More precisely, we have the following.

Proposition 1.5. *Let X be a reflexive Banach space and let $\varphi : X \to \overline{\mathbf{R}}$ be an l.s.c., convex, proper function. Then the following conditions are equivalent.*

(i) $x^* \in \partial\varphi(x)$,
(ii) $\varphi(x) + \varphi^*(x^*) = (x^*, x)$,
(iii) $x \in \partial\varphi^*(x^*)$.

In particular, $\partial\varphi^* = (\partial\varphi)^{-1}$ and $(\varphi^*)^* = \varphi$.

Proof. By definition of φ^*, we see that

$$\varphi^*(x^*) \geq (x^*, x) - \varphi(x), \qquad \forall x \in X,$$

with equality if and only if $0 \in \partial_x(-(x^*, x) + \varphi(x))$. Hence, (i) and (ii) are equivalent. Now, if (ii) holds, then x^* is a minimum point for the function $\varphi^*(p) - (x, p)$ and so

$x \in \partial\varphi^*(x^*)$. Hence, (ii) $\Rightarrow$ (iii). Because conditions (i) and (ii) are equivalent for φ^*, we may equivalently express (iii) as $\varphi^*(x^*) + (\varphi^*)^*(x) = (x^*, x)$. Thus, to prove (ii) it suffices to show that $(\varphi^*)^* = \varphi$. It is readily seen that $(\varphi^*)^* = \varphi^{**} \le \varphi$. We suppose now that there exists $x_0 \in X$ such that $\varphi^{**}(x_0) < \varphi(x_0)$, and we argue from this to a contradiction. We have, therefore, $(x_0, \varphi^{**}(x_0)) \notin \mathrm{Epi}(\varphi)$ and so, by the separation theorem, it follows that there are $x_0^* \in X^*$ and $\alpha \in \mathbf{R}$ such that $(x_0^*, x_0) + \alpha\varphi^{**}(x_0) > \sup\{(x_0^*, x) + \alpha\lambda;\ (x, \lambda) \in \mathrm{Epi}(\varphi)\}$. After some calculation, it follows that $\alpha < 0$. Then, dividing this inequality by $-\alpha$, we get that

$$-\left(x_0^*, \frac{x_0}{\alpha}\right) - \varphi^{**}(x_0) > \sup\left\{\left(x_0^*, -\frac{x}{\alpha}\right) - \lambda;\ (x, \lambda) \in \mathrm{Epi}(\varphi)\right\}$$
$$= \sup\left\{\left(-\frac{x_0^*}{\alpha}, x\right) - \varphi(x);\ x \in D(\varphi)\right\} = \varphi^*\left(-\frac{x_0^*}{\alpha}\right),$$

which clearly contradicts the definition of φ^{**}. $\square$

We mention without proof the following density result. (See, e.g., [2].)

Proposition 1.6. *Let $\varphi : X \to \overline{\mathbf{R}}$ be an l.s.c., convex, and proper function. Then $D(\partial\varphi)$ is a dense subset of $D(\varphi)$.*

Proposition 1.7. *Let φ be an l.s.c., proper, convex function on X. Then* $\mathrm{int}\,D(\varphi) \subset D(\partial\varphi)$.

Proof. Let $x_0 \in \mathrm{int}\,D(\varphi)$ and let $V = B(x_0, r) = \{x;\ \|x - x_0\| < r\}$ be such that $V \subset D(\varphi)$. We know by Proposition 1.2 that φ is continuous on V and this implies that the set $C = \{(x, \lambda) \in V \times \mathbf{R};\ \varphi(x) < \lambda\}$ is an open convex set of $X \times \mathbf{R}$. Thus, there is a closed hyperplane, $H = \{(x, \lambda) \in X \times \mathbf{R};\ (x_0^*, x) + \lambda = \alpha\}$, that separates $(x_0, \varphi(x_0))$ from $\overline{C}$. Hence, $(x_0^*, x_0) + \varphi(x_0) < \alpha$ and

$$(x_0^*, x) + \lambda \ge \alpha, \qquad \forall (x, \lambda) \in \overline{C}.$$

This yields

$$\varphi(x_0) - \varphi(x) < -(x_0^*, x_0 - x), \qquad \forall x \in V.$$

But, for every $u \in X$, there exists $0 < \lambda < 1$ such that $x = \lambda x_0 + (1 - \lambda)u \in V$. Substituting this x in the preceding inequality and using the convexity of φ, we obtain that

$$\varphi(x_0) \le \varphi(u) + (x_0^*, x_0 - u), \qquad \forall u \in X.$$

Hence, $x_0 \in D(\partial\varphi)$ and $x_0^* \in \partial\varphi(x_0)$. $\square$

There is a close connection between the range of subdifferential $\partial\varphi$ of a lower semicontinuous convex function $\varphi : X \to \overline{\mathbf{R}}$ and its behavior for $\|x\| \to \infty$. Namely, one has

Proposition 1.8. *The following two conditions are equivalent.*

(j) $R(\partial\varphi) = X^*$, *and* $\partial\varphi^* = (\partial\varphi)^{-1}$ *is bounded on bounded subsets,*
(jj) $\lim_{\|x\| \to \infty} \varphi(x)/\|x\| = +\infty$.

Proof. (jj) $\Rightarrow$ (j). If (jj) holds, then by Proposition 1.4 it follows that for each $f \in X^*$ the equation $f \in \partial\varphi(x)$ or, equivalently, $0 \in \partial(\varphi(x) - f(x))$, has at least one solution $x \in D(\partial\varphi)$. Moreover, if $\{f\}$ remains in a bounded subset of X^*, the same is true of $(\partial\varphi)^{-1}f$.

(j) $\Rightarrow$ (jj). By Proposition 1.5 we have

$$\varphi(x) \geq (x^*, x) - \varphi^*(x^*), \qquad \forall x^* \in X^*,\ \forall x \in X.$$

This yields, for $x^* = \rho J(x)\|x\|^{-1}$,

$$\varphi(x) \geq \rho\|x\| - \varphi^*(\rho J(x)\|x\|^{-1}), \qquad \forall \rho > 0,\ \forall x \in X.$$

Because φ^* and $\partial\varphi^*$ are bounded on bounded subsets, the latter implies (jj). □

1.3 Sobolev Spaces and Linear Elliptic Boundary Value Problems

Throughout this section, until further notice, we assume that Ω is an open subset of $\mathbf{R}^N$. To begin with, let us briefly recall the notion of *distribution*. Let $f = f(x)$ be a complex-valued function defined on Ω. By the *support* of f, abbreviated $\operatorname{supp} f$, we mean the closure of the set $\{x \in \Omega;\ f(x) \neq 0\}$ or, equivalently, the smallest closed set of Ω outside of which f vanishes identically. We denote by $C^k(\Omega)$, $0 \leq k \leq \infty$, the set of all complex-valued functions defined in Ω that have continuous partial derivatives of order up to and including k (of any order $< \infty$ if $k = \infty$). Let $C_0^k(\Omega)$ denote the set of all functions $\varphi \in C^k(\Omega)$ with compact support in Ω.

It is readily seen that $C_0^\infty(\Omega)$ is a linear space. We may introduce in $C_0^\infty(\Omega)$ a convergence as follows. We say that the sequence $\{\varphi_k\} \subset C_0^\infty(\Omega)$ is convergent to φ, denoted $\varphi_k \Rightarrow \varphi$, if

(a) There is a compact $K \subset \Omega$ such that $\operatorname{supp}\varphi_k \subset K$ for all $k = 1, \ldots$.
(b) $\lim_{k\to\infty} D^\alpha\varphi_k = D^\alpha\varphi$ uniformly on K for all $\alpha = (\alpha_1, \ldots, \alpha_n)$.

Here $D^\alpha = D_{x_1}^{\alpha} \cdots D_{x_N}^{\alpha_n}$, $D_{x_i} = \partial/\partial x_i$, $i = 1, \ldots, n$. Equipped in this way, the space $C_0^\infty(\Omega)$ is denoted by $\mathscr{D}(\Omega)$. As a matter of fact, $\mathscr{D}(\Omega)$ can be redefined as a locally convex, linear topological space with a suitable chosen family of seminorms.

Definition 1.1. A linear continuous functional u on $\mathscr{D}(\Omega)$ is called a *distribution* on Ω.

In other words, a *distribution* is a linear functional u on $C_0^\infty(\Omega)$ having the property that $\lim_{k\to\infty} u(\varphi_k) = 0$ for every sequence $\{\varphi_k\} \subset C_0^\infty(\Omega)$ such that $\varphi_k \Rightarrow 0$.

The set of all distributions on Ω is a linear space, denoted by $\mathscr{D}'(\Omega)$.

The distribution is a natural extension of the notion of locally summable function on Ω for if $f \in L^1_{\mathrm{loc}}(\Omega)$, then the linear functional u_f on $C_0^\infty(\Omega)$ defined by

$$u_f(\varphi) = \int_\Omega f(x)\varphi(x)dx, \qquad \forall \varphi \in C_0^\infty(\Omega)$$

is a distribution on Ω; that is, $u_f \in \mathscr{D}'(\Omega)$. Moreover, the map $f \to u_f$ is injective from $L^1_{\text{loc}}(\Omega)$ to $\mathscr{D}'(\Omega)$.

Given $u \in \mathscr{D}'(\Omega)$, by definition, the derivative of order $\alpha = (\alpha_1, ..., \alpha_n)$, $D^\alpha u$, of u, is the distribution

$$(D^\alpha u)(\varphi) = (-1)^{|\alpha|} u(D^\alpha \varphi), \qquad \forall \varphi \in \mathscr{D}(\Omega), \quad \text{where } |\alpha| = \alpha_1 + \cdots + \alpha_n.$$

Let Ω be an open subset of $\mathbf{R}^N$ and let m be a positive integer. Denote by $H^m(\Omega)$ the set of all real valued functions $u \in L^2(\Omega)$ such that distributional derivatives $D^\alpha u$ of u of order $|\alpha| \le m$ all belong to $L^2(\Omega)$. In other words,

$$H^m(\Omega) = \{u \in L^2(\Omega);\ D^\alpha u \in L^2(\Omega),\ |\alpha| \le m\}. \tag{1.19}$$

This is the *Sobolev space* of order m on Ω. It is easily seen that $H^m(\Omega)$ is a linear space by $(u_1 + u_2)(x) = u_1(x) + u_2(x)$, $(\lambda u)(x) = \lambda u(x)$, $\forall \lambda \in \mathbf{R}$, a.e., $x \in \Omega$, under the convention that two $L^2(\Omega)$ functions u_1, u_2 represent the same element of $H^m(\Omega)$ if $u_1(x) = u_2(x)$, a.e., $x \in \Omega$. In other words, we do not distinguish two functions in $H^m(\Omega)$ that coincide almost everywhere. In this context we say that $u \in H^m(\Omega)$ is continuous, differentiable, or absolutely continuous if there exists a function $\bar{u} \in H^m(\Omega)$ which has these properties and coincides almost everywhere with u on Ω.

We present below a few basic properties of Sobolev spaces and refer to the books of Brezis [7], Adams [1] and Barbu [5] for proofs.

Proposition 1.9. $H^m(\Omega)$ *is a Hilbert space with the scalar product*

$$\langle u, v\rangle_m = \sum_{|\alpha| \le m} \int_\Omega D^\alpha u(x) D^\alpha v(x) dx, \qquad \forall u, v \in H^m(\Omega). \tag{1.20}$$

If $\Omega = (a,b)$, $-\infty < a < b < \infty$, then $H^1(\Omega)$ reduces to the subspace of absolutely continuous functions on the interval $[a,b]$ with derivative in $L^2(a,b)$.

Proposition 1.10. $H^1(a,b)$ *coincides with the space of absolutely continuous functions* $u : [a,b] \to \mathbf{R}$ *having the property that* $u' \in L^2(a,b)$. *Moreover, for each function* $u \in H^1(a,b)$ *the derivative* $D^1 u$ *in the sense of distributions coincides with the ordinary derivative* u' *that exists almost everywhere.*

More generally, for an integer $m \ge 1$ and $1 \le p \le \infty$, one defines the Sobolev space

$$W^{m,p}(\Omega) = \{u \in L^p(\Omega);\ D^\alpha u \in L^p(\Omega),\ |\alpha| \le m\} \tag{1.21}$$

with the norm

$$\|u\|_{m,p} = \left(\sum_{|\alpha| \le m} \int_\Omega |D^\alpha u(x)|^p dx \right)^{1/p}. \tag{1.22}$$

For $0 < m < 1$, the space $W^{m,p}(\Omega)$ is defined by (see Adams [1], p. 214)

$$W^{m,p}(\Omega) = \left\{ u \in L^p(\Omega);\ \frac{u(x)-u(y)}{|x-y|^{m+(N/p)}} \in L^p(\Omega \times \Omega) \right\}$$

with the natural norm. For $m > 1$, $m = s + a$, $s = [m]$, $0 < a < 1$, define

$$W^{m,p}(\Omega) = \{u \in W^{s,p}(\Omega);\ D^\alpha u \in W^{a,p}(\Omega);\ |\alpha| \le s\}.$$

It turns out that, if $u \in W^{1,p}(a,b)$, then there is an absolutely continuous function $\bar{u}$ with $\bar{u}' \in L^p(a,b)$ such that $\bar{u}(x) = u(x)$ and $\bar{u}'(x) = (D^1u)(x)$, a.e., $x \in (a,b)$. Conversely, any absolutely continuous function u with u' in $L^p(a,b)$ belongs to $W^{1,p}(a,b)$ and u' coincides, a.e. on (a,b), with the distributional derivative D^1u of u.

Proposition 1.10 and its counterpart in $W^{1,p}(\Omega)$ show that, in one dimension, the Sobolev spaces are just the classical spaces of absolutely continuous functions with derivatives in $L^p(\Omega)$.

It turns out, via regularization, that $C_0^\infty(\mathbf{R}^N)$ is dense in $H^1(\mathbf{R}^N)$.

We recall that an open subset Ω of $\mathbf{R}^N$ and its boundary $\partial\Omega$ are said to be of class C^1 if for each $x \in \partial\Omega$ there are a neighborhood U of x and a one-to-one mapping φ of $Q = \{x = (x', x_N) \in \mathbf{R}^N;\ \|x'\| < 1,\ |x_N| < 1\}$ onto U such that

$$\varphi \in C^1(\overline{Q}), \qquad \varphi^{-1} \in C^1(\overline{U}), \qquad \varphi(Q_+) = U \cap \Omega, \qquad \varphi(Q_0) = U \cap \partial\Omega,$$

where $Q_+ = \{(x', x_N) \in Q;\ x_N > 0\}$, $Q_0 = \{(x', 0);\ \|x'\| < 1\}$.

We are now ready to formulate the extension theorem for the elements of the space $H^1(\Omega)$, a result upon which most of the properties of this space are built.

Theorem 1.3. *Let Ω be an open subset of $\mathbf{R}^N$ that is of class C^1. Assume that either $\partial\Omega$ is compact or $\Omega = \mathbf{R}^N_+$. Then, there is a linear operator $P : H^1(\Omega) \to H^1(\mathbf{R}^N)$ and a positive constant C independent of u, such that*

$$(Pu)(x) = u(x), \quad \text{a.e. } x \in \Omega, \qquad \forall u \in H^1(\Omega), \tag{1.23}$$

$$\|Pu\|_{L^2(\mathbf{R}^N)} \le C\|u\|_{L^2(\Omega)}, \qquad \forall u \in H^1(\Omega), \tag{1.24}$$

$$\|Pu\|_{H^1(\mathbf{R}^N)} \le C\|u\|_{H^1(\Omega)}, \qquad \forall u \in H^1(\Omega). \tag{1.25}$$

Theorem 1.3 follows from the next extension result.

Let $u \in H^1(Q_+)$ and let $u^* : Q \to \mathbf{R}$ be the extension of u to Q

$$u^*(x', x_N) = \begin{cases} u(x', x_N) & \text{if } x_N \ge 0 \\ u(x', -x_N) & \text{if } x_N < 0. \end{cases}$$

Then $u^* \in H^1(Q)$ and $\|u^*\|_{L^2(Q)} \le 2\|u\|_{L^2(Q_+)}$, $\|u^*\|_{H^1(Q)} \le 2\|u\|_{H^1(Q_+)}$. The general result follows by a specific argument involving partition of unity (see, e.g., Brezis [7] or Barbu [5]).

Now, we mention without proof an important property of the space $H^1(\Omega)$ known as the *Sobolev embedding theorem.*

Theorem 1.4. *Let Ω be an open subset of $\mathbf{R}^N$ of class C^1 with compact boundary $\partial\Omega$, or $\Omega = \mathbf{R}^N_+$, or $\Omega = \mathbf{R}^N$. Then, if $N > 2$,*

$$H^1(\Omega) \subset L^{p^*}(\Omega) \quad \textit{for } \frac{1}{p^*} = \frac{1}{2} - \frac{1}{N}. \tag{1.26}$$

If $N = 2$, then $H^1(\Omega) \subset L^p(\Omega)$ for all $p \in [2, \infty[$.

The inclusion relation (1.26) should be considered of course in the algebraic and topological sense; that is,

$$\|u\|_{L^{p^*}(\Omega)} \le C\|u\|_{H^1(\Omega)} \tag{1.27}$$

for some positive constant C independent of u.

Theorem 1.4 has a natural extension to the Sobolev space $W^{m,p}(\Omega)$ for any $m > 0$. More precisely, we have (see Adams [1], p. 217)

Theorem 1.5. *Under the assumptions of Theorem 1.4, we have*

$$\begin{aligned} &W^{m,p}(\Omega) \subset L^{p^*}(\Omega) && \textit{if } 1 \le p < \frac{N}{m},\ \frac{1}{p^*} = \frac{1}{p} - \frac{m}{N}, \\ &W^{m,p}(\Omega) \subset L^{q}(\Omega) && \textit{for all } q \ge p \textit{ if } p = \frac{N}{m}, \\ &W^{m,p}(\Omega) \subset L^{\infty}(\Omega) && \textit{if } p > \frac{N}{m}. \end{aligned}$$

Remark 1.1. If Ω is a bounded and open subset of $\mathbf{R}^N$ of class C^1, then the following norm on $W^{1,p}(\Omega)$,

$$\|u\|_{1,p} = |\nabla u|_{L^p(\Omega)} + |u|_{L^q(\Omega)},$$

where $1 \le q \le p^*$ if $1 \le p < N$, $1 \le q < \infty$ if $p = N$ and $1 \le q \le \infty$ if $p > N$,

$$\frac{1}{p^*} = \frac{1}{p} - \frac{1}{N}$$

is equivalent with the norm (1.22) for $m = 1$ (see, e.g., Brezis [7], p. 170).

We note also the following compactness embedding result.

Theorem 1.6. *Let Ω be an open and bounded subset of $\mathbf{R}^N$ that is of class C^1. Then, the injection of the space $H^1(\Omega)$ into $L^2(\Omega)$ is compact.*

The "trace" to $\partial\Omega$ of a Function $u \in H^1(\Omega)$

If Ω is an open C^1 subset of $\mathbf{R}^N$ with the boundary $\partial\Omega$, then each $u \in C(\overline{\Omega})$ is well defined on $\partial\Omega$. We call the restriction of u to $\partial\Omega$ the *trace* of u to $\partial\Omega$ and it is denoted by $\gamma_0(u)$. If $u \in L^2(\Omega)$, then $\gamma_0(u)$ is no longer well defined. We have, however, the following.

Lemma 1.2. *Let Ω be an open subset of class C^1 with compact boundary $\partial\Omega$ or $\Omega = \mathbf{R}^N_+$. Then, there is $C > 0$ such that*

$$\|\gamma_0(u)\|_{L^2(\partial\Omega)} \leq C\|u\|_{H^1(\Omega)}, \qquad \forall u \in C_0^\infty(\mathbf{R}^N). \tag{1.28}$$

Taking into account that for domains Ω of class C^1 the space $\{u|_\Omega;\ u \in C_0^\infty(\mathbf{R}^N)\}$ is dense in $H^1(\Omega)$ (see, e.g., Adams [1], p. 54, or Brezis [7], p. 162), a natural way to define the *trace* of a function $u \in H^1(\Omega)$ is the following.

Definition 1.2. Let Ω be of class C^1 with compact boundary or $\Omega = \mathbf{R}^N_+$. Let $u \in H^1(\Omega)$. Then $\gamma_0(u) = \lim_{j\to\infty} \gamma_0(u_j)$ in $L^2(\partial\Omega)$, where $\{u_j\} \subset C_0^\infty(\mathbf{R}^N)$ is such that $u_j \to u$ in $H^1(\Omega)$.

It turns out that the definition is consistent; that is, $\gamma_0(u)$ is independent of $\{u_j\}$. Indeed, if $\{u_j\}$ and $\{\bar{u}_j\}$ are two sequences in $C_0^\infty(\mathbf{R}^N)$ convergent to u in $H^1(\Omega)$, then, by (1.28),

$$\|\gamma_0(u_j - \bar{u}_j)\|_{L^2(\partial\Omega)} \leq C\|u_j - \bar{u}_j\|_{H^1(\Omega)} \to 0 \quad \text{as } j \to \infty.$$

Moreover, it follows by Lemma 1.2 that the map $\gamma_0 : H^1(\Omega) \to L^2(\partial\Omega)$ is continuous. As a matter of fact, it turns out that the trace operator $u \to \gamma_0(u)$ is continuous from $H^1(\Omega)$ to $H^{1/2}(\partial\Omega)$ and so it is completely continuous from $H^1(\Omega)$ to $L^2(\partial\Omega)$.

In general (see Adams [1], p. 114), we have $W^{m,p}(\Omega) \subset L^q(\partial\Omega)$ if $mp < N$ and

$$p \leq q \leq \frac{(N-1)p}{(N-mp)}\,.$$

Definition 1.3. Let Ω be any open subset of $\mathbf{R}^N$. The space $H_0^1(\Omega)$ is the closure (the completion) of $C_0^1(\Omega)$ in the norm of $H^1(\Omega)$.

It follows that $H_0^1(\Omega)$ is a closed subspace of $H^1(\Omega)$ and in general it is a proper subspace of $H^1(\Omega)$. It is clear that $H_0^1(\Omega)$ is a Hilbert space with the scalar product

$$\langle u, v\rangle_1 = \sum_{i=1}^N \int_\Omega \frac{\partial u}{\partial x_i}\,\frac{\partial v}{\partial x_i}\,dx + \int_\Omega uv\,dx$$

with the corresponding norm

$$\|u\|_1 = \left(\int_\Omega (|\nabla u(x)|^2 + u^2(x))dx\right)^{1/2}.$$

Roughly speaking, $H_0^1(\Omega)$ is the subspace of functions $u \in H^1(\Omega)$ that are zero on $\partial\Omega$. More precisely, we have the following.

Proposition 1.11. *Let Ω be an open set of class C^1 and let $u \in H^1(\Omega)$. Then, the following conditions are equivalent.*

(i) $u \in H_0^1(\Omega)$.
(ii) $\gamma_0(u) \equiv 0$.

Proposition 1.12 below is called the *Poincaré inequality*.

Proposition 1.12. *Let Ω be an open and bounded subset of $\mathbf{R}^N$. Then there is $C > 0$ independent of u such that*

$$\|u\|_{L^2(\Omega)} \le C \|\nabla u\|_{L^2(\Omega)}, \qquad \forall u \in H_0^1(\Omega).$$

In particular, Proposition 1.12 shows that if Ω is bounded, then the scalar product

$$((u,v)) = \int_\Omega \nabla u(x) \cdot \nabla v(x) dx$$

and the corresponding norm

$$\|u\| = \left(\int_\Omega |\nabla u(x)|^2 dx \right)^{1/2}$$

define an equivalent Hilbertian structure on $H_0^1(\Omega)$.

We denote by $H^{-1}(\Omega)$ the dual space of $H_0^1(\Omega)$; that is, the space of all linear continuous functionals on $H_0^1(\Omega)$. Equivalently,

$$H^{-1}(\Omega) = \{u \in \mathscr{D}'(\Omega);\ |u(\varphi)| \le C_u \|\varphi\|_{H^1(\Omega)}, \qquad \forall \varphi \in C_0^\infty(\Omega)\}.$$

The space $H^{-1}(\Omega)$ is endowed with the dual norm

$$\|u\|_{-1} = \sup\{|u(\varphi)|;\ \|\varphi\| \le 1\}, \qquad \forall u \in H^{-1}(\Omega).$$

By Riesz's theorem, we know that $H^{-1}(\Omega)$ is isometric to $H_0^1(\Omega)$. Note also that

$$H_0^1(\Omega) \subset L^2(\Omega) \subset H^{-1}(\Omega)$$

in the algebraic and topological sense. In other words, the injections of $L^2(\Omega)$ into $H^{-1}(\Omega)$ and of $H_0^1(\Omega)$ into $L^2(\Omega)$ are continuous. Note also that the above injections are dense.

There is an equivalent definition of $H^{-1}(\Omega)$ given in Theorem 1.7 below.

Theorem 1.7. *The space $H^{-1}(\Omega)$ coincides with the set of all distributions $u \in \mathscr{D}'(\Omega)$ of the form*

$$u = f_0 + \sum_{i=1}^N \frac{\partial f_i}{\partial x_i} \quad \text{in } \mathscr{D}'(\Omega), \text{ where } f_i \in L^2(\Omega),\ i = 1, \dots, N.$$

The space $W_0^{1,p}(\Omega)$, $p \ge 1$, is similarly defined as the closure of $C_0^1(\Omega)$ into $W^{1,p}(\Omega)$ norm. The dual of $W_0^{1,p}(\Omega)$ is the space

$$W^{-1,q}(\Omega), \qquad \frac{1}{p}+\frac{1}{q}=1$$

defined as in Theorem 1.7 with $f_0, f_1, \ldots, f_N \in L^q$.

Variational Theory of Elliptic Boundary Value Problems

We begin by recalling an abstract existence result, the *Lax–Milgram lemma*, which is the foundation upon which all the results of this section are built. Before presenting it, we need to clarify certain concepts.

Let V be a real Hilbert space and let V^* be the topological dual space of V. For each $v^* \in V^*$ and $v \in V$ we denote by (v^*, v) the value $v^*(v)$ of functional v^* at v. The functional $a : V \times V \to \mathbf{R}$ is said to be *bilinear* if for each $u \in V$, $v \to a(u,v)$ is linear and for each $v \in V$, $u \to a(u,v)$ is linear on V. The functional a is said to be *continuous* if there exists $M > 0$ such that

$$|a(u,v)| \le M\|u\|_V \, \|v\|_V, \qquad \forall u, v \in V.$$

The functional a is said to be *coercive* if

$$a(u,u) \ge \omega\|u\|_V^2, \qquad \forall u \in V,$$

for some $\omega > 0$, and *symmetric* if

$$a(u,v) = a(v,u), \qquad \forall u, v \in V.$$

Lemma 1.3 (Lax–Milgram). *Let $a : V \times V \to \mathbf{R}$ be a bilinear, continuous, and coercive functional. Then, for each $f \in V^*$, there is a unique $u^* \in V$ such that*

$$a(u^*, v) = (f, v), \qquad \forall v \in V. \tag{1.29}$$

Moreover, the map $f \to u^*$ is Lipschitzian from V^* to V with Lipschitz constant $\le \omega^{-1}$. If a is symmetric, then u^* minimizes the function $u \to (1/2)a(u,u) - (f,u)$ on V; that is,

$$\frac{1}{2}a(u^*, u^*) - (f, u^*) = \min\left\{\frac{1}{2}a(u,u) - (f,u);\ u \in V\right\}. \tag{1.30}$$

If a is symmetric, then the Lax–Milgram lemma is a simple consequence of Riesz's representation theorem. Indeed, in this case $(u,v) \to a(u,v)$ is an equivalent scalar product on V and so, by the Riesz theorem, the functional $v \to (f,v)$ can be represented as (1.29) for some $u^* \in V$. In the general case we proceed as follows. For each $u \in V$, the functional $v \to a(u,v)$ is linear and continuous on V and we denote it by $Au \in V^*$. Then, the equation

$$a(u,v) = (f,v), \qquad \forall v \in V$$

can be rewritten as $Au = f$. Then, the conclusion follows because $R(A)$ is simultaneously closed and dense in V^*.

Weak Solutions to the Dirichlet Problem

Consider the Dirichlet problem

$$\begin{cases} -\Delta u + c(x)u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{1.31}$$

where Ω is an open set of $\mathbf{R}^N$, $c \in L^\infty(\Omega)$, and $f \in H^{-1}(\Omega)$ is given.

Definition 1.4. The function u is said to be a *weak* or *variational solution* to the Dirichlet problem (1.31) if $u \in H_0^1(\Omega)$ and

$$\int_\Omega \nabla u(x) \cdot \nabla \varphi(x) dx + \int_\Omega c(x)u(x)\varphi(x) dx = (f, \varphi) \tag{1.32}$$

for all $\varphi \in H_0^1(\Omega)$ (equivalently, for all $\varphi \in C_0^\infty(\Omega)$).

In (1.32), ∇u is taken in the sense of distributions and (f, φ) is the value of the functional $f \in H^{-1}(\Omega)$ into $\varphi \in H_0^1(\Omega)$. If $f \in L^2(\Omega) \subset H^{-1}(\Omega)$, then

$$(f, \varphi) = \int_\Omega f(x)\varphi(x) dx.$$

By the Lax–Milgram lemma, applied to the functional

$$a(u, v) = \int_\Omega (\nabla u(x) \cdot \nabla v(x) + c\,uv) dx, \qquad u, v \in V = H_0^1(\Omega),$$

we obtain the following.

Theorem 1.8. *Let Ω be a bounded open set of $\mathbf{R}^N$ and let $c \in L^\infty(\Omega)$ be such that $c(x) \geq 0$, a.e. $x \in \Omega$. Then, for each $f \in H^{-1}(\Omega)$ the Dirichlet problem (1.31) has a unique weak solution $u^* \in H_0^1(\Omega)$. Moreover, u^* minimizes on $H_0^1(\Omega)$ the functional*

$$\frac{1}{2} \int_\Omega (|\nabla u(x)|^2 + c(x)u^2(x)) dx - (f, u) \tag{1.33}$$

and the map $f \to u^*$ is Lipschitzian from $H^{-1}(\Omega)$ to $H_0^1(\Omega)$.

Weak Solutions to the Neumann Problem

Consider the boundary value problem

$$\begin{cases} -\Delta u + cu = f & \text{in } \Omega, \\ \dfrac{\partial u}{\partial \nu} = g & \text{on } \partial\Omega, \end{cases} \tag{1.34}$$

where $c \in L^\infty(\Omega),\ c(x) \geq \rho > 0$, and $f \in L^2(\Omega),\ g \in L^2(\partial\Omega)$.

Definition 1.5. The function $u \in H^1(\Omega)$ is said to be a *weak solution* to the Neumann problem (1.34) if

$$\int_\Omega \nabla u \cdot \nabla v\, dx + \int_\Omega cuv\, dx = \int_\Omega fv\, dx + \int_{\partial\Omega} gv\, d\sigma, \qquad \forall v \in H^1(\Omega). \tag{1.35}$$

Because for each $v \in H^1(\Omega)$ the trace $\gamma_0(v)$ is in $L^2(\partial\Omega)$, the integral $\int_{\partial\Omega} gv\, d\sigma$ is well defined and so (1.35) makes sense.

Theorem 1.9. *Let Ω be an open subset of $\mathbf{R}^N$. Then, for each $f \in L^2(\Omega)$ and $g \in L^2(\partial\Omega)$, problem* (1.34) *has a unique weak solution $u \in H^1(\Omega)$ that minimizes the functional*

$$u \to \frac{1}{2} \int_\Omega (|\nabla u(x)|^2 + c(x)u^2(x))dx - \int_\Omega f(x)u(x)dx - \int_{\partial\Omega} gu\, d\sigma \qquad \text{on } H^1(\Omega).$$

Proof. One applies the Lax–Milgram lemma on the space $V = H^1(\Omega)$ to the functional $a(u,v) = \int_\Omega (\nabla u \cdot \nabla v + cuv)dx,\ \forall u,v \in H^1(\Omega)$, and $(\widetilde{f}, v) = \int_\Omega fv\, dx + \int_{\partial\Omega} gv\, d\sigma$. □

Regularity of the Weak Solutions

We briefly recall here the regularity of the weak solutions to the Dirichlet problem

$$\begin{cases} -\Delta u = f & \text{in } \Omega, \\ u = 0 & \text{on } \partial\Omega. \end{cases} \tag{1.36}$$

By Theorem 1.8 we know that if Ω is a bounded and open subset of $\mathbf{R}^N$ and $f \in L^2(\Omega)$, then problem (1.36) has a unique solution $u \in H_0^1(\Omega)$. It turns out that if $\partial\Omega$ is smooth enough, then this solution is actually in $H^2(\Omega) \cap H_0^1(\Omega)$. More precisely, we have the following theorem.

Theorem 1.10. *Let Ω be a bounded and open subset of $\mathbf{R}^N$ of class C^2. Let $f \in L^2(\Omega)$ and let $u \in H_0^1(\Omega)$ be the weak solution to* (1.36). *Then, $u \in H^2(\Omega)$ and*

$$\|u\|_{H^2(\Omega)} \leq C\|f\|_{L^2(\Omega)}, \tag{1.37}$$

where C is independent of f.

To prove the theorem, one first shows that $u \in H^2(\Omega')$ for each open subset $\Omega' \subset \Omega$ compactly embedded in Ω (interior regularity). The most delicate part (boundary regularity) follows by the *method of tangential quotients* due to L. Nirenberg. In short, the idea is to reduce problem (1.36) to an elliptic Dirichlet problem on $\mathbf{R}^N_+$ and to estimate separately the tangential quotients $(\nabla u)_h$, $h=(h_1,...,h_{N-1},0)$ and the normal quotient $(\nabla u)_h$, $h=(0,...,0,h_N)$ in order to show that $v \in H^2(\mathbf{R}^N_+)$. For details we refer to Brezis' book [7]. (See also [5].)

In particular, Theorem 1.10 implies that if $A: H^1_0(\Omega) \to H^{-1}(\Omega)$ is the elliptic operator $A = -\Delta$ in $\mathscr{D}'(\Omega)$; that is,

$$(Au,\varphi) = \int_\Omega \nabla u \cdot \nabla \varphi \, dx, \qquad \forall \varphi \in H^1_0(\Omega),$$

then

$$\{u \in H^1_0(\Omega);\ Au \in L^2(\Omega)\} \subset H^2(\Omega)$$

and

$$\|u\|_{H^2(\Omega)} \le C\|Au\|_{L^2(\Omega)}, \qquad \forall u \in H^1_0(\Omega) \cap H^2(\Omega).$$

Theorem 1.10 remains true if Ω is an open, convex, and bounded subset of $\mathbf{R}^N$. For the proof which uses some specific geometrical properties of Ω we refer the reader to Grisvard [10]. More generally, we have the following.

Theorem 1.11. *If Ω is of class C^{m+2} and $f \in H^m(\Omega)$, then the weak solution u to problem* (1.36) *belongs to $H^{m+2}(\Omega)$ and*

$$\|u\|_{m+2} \le C\|f\|_m, \qquad \forall f \in H^m(\Omega).$$

If $m > N/2$, then $u \in C^2(\overline{\Omega})$. In particular, if Ω is of class C^∞ and $f \in C^\infty(\overline{\Omega})$, then $u \in C^\infty(\overline{\Omega})$.

We conclude this section with a regularity result for the weak solution $u \in H^1(\Omega)$ to Neumann's problem

$$\begin{cases} u - \Delta u = f & \text{in } \Omega, \\ \dfrac{\partial u}{\partial \nu} = 0 & \text{on } \partial\Omega. \end{cases} \tag{1.38}$$

Theorem 1.12. *Under the assumptions of Theorem* 1.10 *the weak solution $u \in H^1(\Omega)$ to problem* (1.38) *belongs to $H^2(\Omega)$ and*

$$\|u\|_{H^2(\Omega)} \le C\|f\|_{L^2(\Omega)}, \qquad \forall f \in L^2(\Omega). \tag{1.39}$$

Theorem 1.10 remains true in $L^p(\Omega)$ for $p > 1$. Namely, we have (Agmon, Douglis and Nirenberg [2])

Theorem 1.13. *Let Ω be a bounded open subset of $\mathbf{R}^N$ with smooth boundary $\partial\Omega$ and let $1 < p < \infty$. Then, for each $f \in L^p(\Omega)$, the boundary value problem*

$$-\Delta u = f \quad in\ \ \Omega, \qquad u = 0 \quad on\ \ \partial\Omega$$

has a unique weak solution $u \in W^{1,p}_0(\Omega) \cap W^{2,p}(\Omega)$. Moreover, one has

$$\|u\|_{W^{2,p}(\Omega)} \le C\|f\|_{L^p(\Omega)},$$

where C is independent of f.

The Space $BV(\Omega)$

Let Ω be an open subset of $\mathbf{R}^N$ with smooth boundary $\partial\Omega$.

A function $f \in L^1(\Omega)$ is said to be of bounded variation on Ω if its gradient Df in the sense of distributions is an $\mathbf{R}^N$-valued measure on Ω; that is,

$$\|Df\| := \sup\left\{\int_\Omega f \operatorname{div}\psi\, d\xi : \psi \in C_0^\infty(\Omega;\mathbf{R}^N),\ |\psi|_\infty \le 1\right\} < +\infty,$$

or, equivalently,

$$\|Df\| = \int_\Omega |Df(x)|dx,$$

where $|Df|$ is the total variation of measure Df.

The space of all functions of bounded variation on Ω is denoted by $BV(\Omega)$. It is a Banach space with the norm

$$\|f\|_{BV(\Omega)} = |f|_{L^1(\Omega)} + \|Df\|.$$

Let $f \in BV(\Omega)$. Then there is a Radon measure μ_f on $\overline{\Omega}$ and a μ_f-measurable function $\sigma_f : \Omega \to \mathbf{R}^N$ such that $|\sigma_f(x)| = 1$, μ_f, a.e., and

$$\int_\Omega f \operatorname{div}\psi\, d\xi = -\int_\Omega \psi\cdot\sigma_f\, d\mu_f, \qquad \forall\psi \in C_0^1(\Omega;\mathbf{R}^N). \tag{1.40}$$

For each $f \in BV(\Omega)$ there is the trace $\gamma(f)$ on $\partial\Omega$ (assumed sufficiently smooth) defined by

$$\int_\Omega f \operatorname{div}\psi\, d\xi = -\int_\Omega \psi\cdot\sigma_f\, d\mu_f + \int_{\partial\Omega} \gamma(f)\psi\cdot\nu\, dH^{N-1}, \quad \forall\psi \in C^1(\overline{\Omega};\mathbf{R}^N), \tag{1.41}$$

where ν is the outward normal and dH^{N-1} is the Hausdorff measure on $\partial\Omega$. We have that $|\gamma(f)|_N \in L^1(\partial\Omega; dH^{N-1})$.

We denote by $BV^0(\Omega)$ the space of all $BV(\Omega)$ functions with vanishing trace on $\partial\Omega$. By the Poincaré inequality it follows that, on $BV^0(\Omega)$, $\|Df\|$ is a norm equivalent with $\|f\|_{BV^0(\Omega)}$.

Theorem 1.14. *Let $1 \le p \le N/(N-1)$ and Ω be a bounded open subset. Then, we have $BV(\Omega) \subset L^p(\Omega)$ with continuous and compact embedding. Moreover, the function $u \to \|Du\|$ is lower semicontinuous in $L^p(\Omega)$.*

We refer the reader to Ambrosio, Fusco and Pallara [3] for proofs and other basic results on functions with bounded variations.

Weak compactness in $L^1(\Omega)$

Let Ω be a measurable subset of $\mathbf{R}^N$. Contrary to what happens in $L^p(\Omega)$ spaces with $1 < p < \infty$ that are reflexive, a bounded subset $\mathscr{M}$ of $L^1(\Omega)$ is not necessarily weakly compact. This happens, however, under some additional conditions on $\mathscr{M}$.

Theorem 1.15. (Dunford–Pettis) *Let $\mathscr{M}$ be a bounded subset of $L^1(\Omega)$ having the property that the family of integrals $\{\int_E u(x)dx;\ E \subset \Omega$ measurable, $u \in \mathscr{M}\}$ is uniformly absolutely continuous; that is, for every $\varepsilon > 0$ there is $\delta(\varepsilon) > 0$ independent of u, such that $\int_E |u(x)|dx \leq \varepsilon$ for $m(E) < \delta(\varepsilon)$ (m is the Lebesgue measure). Then the set $\mathscr{M}$ is weakly sequentially compact in $L^1(\Omega)$.*

For the proof, we refer the reader to Edwards [9], p. 270.

Theorem 1.15 remains true, of course, in $(L^1(\Omega))^m$, $m \in \mathbf{N}$.

1.4 Infinite-Dimensional Sobolev Spaces

Let X be a real (or complex) Banach space and let $[a,b]$ be a fixed interval on the real axis. A function $x : [a,b] \to X$ is said to be *finitely valued* if it is constant on each of a finite number of disjoint measurable sets $A_k \subset [a,b]$ and equal to zero on $[a,b] \setminus \cup_k A_k$. The function x is said to be *strongly measurable* on $[a,b]$ if there is a sequence $\{x_n\}$ of finite-valued functions that converges strongly in X and almost everywhere on $[a,b]$ to x. The function x is said to be *Bochner integrable* if there exists a sequence $\{x_n\}$ of finitely valued functions on $[a,b]$ to X that converges almost everywhere to x such that

$$\lim_{n\to\infty} \int_a^b \|x_n(t) - x(t)\| dt = 0.$$

A necessary and sufficient condition guaranteeing that $x : [a,b] \to X$ is Bochner integrable is that x is strongly measurable and that $\int_a^b \|x(t)\| dt < \infty$. The space of all Bochner integrable functions $x : [a,b] \to X$ is a Banach space with the norm

$$\|x\|_1 = \int_a^b \|x(t)\| dt,$$

and is denoted by $L^1(a,b;X)$.

More generally, the space of all (classes of) strongly measurable functions x from $[a,b]$ to X such that

$$\|x\|_p = \left(\int_a^b \|x(t)\|^p dt \right)^{1/p} < \infty$$

for $1 \leq p < \infty$ and $\|x\|_\infty = \operatorname{ess\,sup}_{t\in[a,b]} \|x(t)\| < \infty$, is denoted by $L^p(a,b;X)$. This is a Banach space in the norm $\|\cdot\|_p$.

If X is reflexive, then the dual of $L^p(a,b;X)$ is the space $L^q(a,b;X^*)$, where $p < \infty$, $1/p + 1/q = 1$ (see Edward [9]). Recall also that a function $x : [a,b] \to X$ is said to be *weakly measurable* if for any $x^* \in X^*$, the function $t \to (x^*, x(t))$ is measurable. According to the Pettis theorem, if X is separable then every weakly measurable function is strongly measurable, and so these two notions coincide.

An X-valued function x defined on $[a,b]$ is said to be *absolutely continuous* on $[a,b]$ if for each $\varepsilon > 0$ there exists $\delta(\varepsilon)$ such that $\sum_{n=1}^{N} \|x(t_n) - x(s_n)\| \le \varepsilon$, whenever $\sum_{n=1}^{N} |t_n - s_n| \le \delta(\varepsilon)$ and $(t_n, s_n) \cap (t_m, s_m) = \emptyset$ for $m \ne n$. Here, (t_n, s_n) is an arbitrary subinterval of (a,b).

A classical result in real analysis says that any real-valued absolutely continuous function is almost everywhere differentiable and it is expressed as the indefinite integral of its derivative. It should be mentioned that this result fails for X-valued absolutely continuous functions if X is a general Banach space.

However, if the space X is reflexive, we have (see, e.g., Komura [12]):

Theorem 1.16. *Let X be a reflexive Banach space. Then every X-valued absolutely continuous function x on $[a,b]$ is almost everywhere differentiable on $[a,b]$ and*

$$x(t) = x(a) + \int_a^t \frac{d}{ds} x(s) ds, \qquad \forall t \in [a,b], \tag{1.42}$$

where $(dx/dt) : [a,b] \to X$ is the derivative of x; that is,

$$\frac{d}{dt} x(t) = \lim_{\varepsilon \to 0} \frac{x(t+\varepsilon) - x(t)}{\varepsilon}.$$

Let us denote, as above, by $\mathscr{D}(a,b)$ the space of all infinitely differentiable real-valued functions on $[a,b]$ with compact support in (a,b), and by $\mathscr{D}'(a,b;X)$ the space of all continuous operators from $\mathscr{D}(a,b)$ to X. An element u of $\mathscr{D}'(a,b;X)$ is called an X-valued distribution on (a,b). If $u \in \mathscr{D}'(a,b;X)$ and j is a natural number, then

$$u^{(j)}(\varphi) = (-1)^j u(\varphi^{(j)}), \qquad \forall \varphi \in \mathscr{D}(a,b),$$

defines another distribution $u^{(j)}$, which is called the derivative of order j of u.

We note that every element $u \in L^1(a,b;X)$ defines uniquely the distribution (again denoted u)

$$u(\varphi) = \int_a^b u(t)\varphi(t) dt, \qquad \forall \varphi \in \mathscr{D}(a,b), \tag{1.43}$$

and so $L^1(a,b;X)$ can be regarded as a subspace of $\mathscr{D}'(a,b;X)$. In all that follows, we identify a function $u \in L^1(a,b;X)$ with the distribution u defined by (1.43).

Let k be a natural number and $1 \le p \le \infty$. We denote by $W^{k,p}([a,b];X)$ the space of all X-valued distributions $u \in \mathscr{D}'(a,b;X)$ such that

$$u^{(j)} \in L^p(a,b;X) \quad \text{for } j = 0, 1, \ldots, k. \tag{1.44}$$

Here, $u^{(j)}$ is the derivative of order j of u in the sense of distributions.

We denote by $A^{1,p}([a,b];X)$, $1 \le p \le \infty$, the space of all absolutely continuous functions u from $[a,b]$ to X having the property that they are a.e. differentiable on (a,b) and $(du/dt) \in L^p(a,b;X)$. If the space X is reflexive, it follows by Theorem 1.16 that $u \in A^{1,p}([a,b];X)$ if and only if u is absolutely continuous on $[a,b]$ and $(du/dt) \in L^p(a,b;X)$.

It turns out that the space $W^{1,p}$ can be identified with $A^{1,p}$. More precisely, we have (see Brezis [7]) the following theorem.

Theorem 1.17. *Let X be a Banach space and let $u \in L^p(a,b;X)$, $1 \le p \le \infty$. Then the following conditions are equivalent.*

(i) $u \in W^{1,p}([a,b];X)$.

(ii) *There is $u^0 \in A^{1,p}([a,b];X)$ such that $u(t) = u^0(t)$, a.e., $t \in (a,b)$. Moreover, $u' = du^0/dt$, a.e. in (a,b).*

Proof. For simplicity, we assume that $[a,b] = [0,T]$.

Let $u \in W^{1,p}([0,T];X)$; that is, $u \in L^p(0,T;X)$ and $u' \in L^p(0,T;X)$, and define the regularization u_n of u,

$$u_n(t) = n \int_0^T u(s)\rho((t-s)n)ds, \qquad \forall t \in [0,T], \tag{1.45}$$

where $\rho \in \mathscr{D}(\mathbf{R})$ is such that $\int \rho(s)ds = 1$, $\rho(t) = \rho(-t)$, $\operatorname{supp}\rho \subset [-1,1]$. It is well known that $u_n \to u$ in $L^p(0,T;X)$ for $n \to \infty$. Note also that u_n is infinitely differentiable. Let $\varphi \in \mathscr{D}(0,T)$ be arbitrary but fixed. Then, by (1.45), we see that

$$\begin{aligned}\int_0^T \frac{du_n}{dt}(t)\varphi(t)dt &= -\int_0^T u_n(t)\frac{d\varphi}{dt}(t)dt = -\int_0^T u(t)\frac{d\varphi_n}{dt}(t)dt\\ &= u'(\varphi_n) = \int_0^T u_n'\varphi\, dt \quad \text{if } \operatorname{supp}\varphi \subset \left(\frac{1}{n}, T-\frac{1}{n}\right).\end{aligned}$$

Hence,

$$\frac{du_n}{dt} = u_n', \quad \text{a.e. in } \left(\frac{1}{n}, T-\frac{1}{n}\right).$$

On the other hand, letting n tend to ∞ in the equation

$$u_n(t) - u_n(s) = \int_s^t \frac{du_n}{d\tau}(\tau)d\tau,$$

we get

$$u(t) - u(s) = \int_s^t u'(\tau)\,d\tau, \quad \text{a.e. } t,s \in (0,T),$$

because $(u')_n \to u'$ in $L^p(0,T;X)$. The latter equation implies that u admits an extension to an absolutely continuous function u^0 on $[0,T]$ that satisfies the equation

$$u^0(t) - u^0(0) = \int_0^t u'(\tau)d\tau, \qquad \forall t \in [0,T].$$

Hence, (i) $\Rightarrow$ (ii).

Conversely, assume now that $u \in A^{1,p}([0,T];X)$. Then,

$$u'(\varphi) = -\int_0^T u(t)\varphi'(t)dt = -\lim_{\varepsilon\to 0}\int_0^T u(t)\,\frac{\varphi(t)-\varphi(t-\varepsilon)}{\varepsilon}\,dt$$
$$= -\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_0^{T-\varepsilon}(u(t)-u(t+\varepsilon))\varphi(t)dt - \lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_{T-\varepsilon}^T u(t)\varphi(t)dt$$
$$+\lim_{\varepsilon\to 0}\frac{1}{\varepsilon}\int_0^{\varepsilon} u(t)\varphi(t-\varepsilon)dt, \qquad \forall\varphi\in\mathscr{D}(0,T).$$

Hence

$$u'(\varphi) = \int_0^T \frac{du}{dt}(t)\varphi(t), \qquad \forall\varphi\in\mathscr{D}(0,T).$$

This shows that $u' \in L^p(0,T;X)$ and $u' = du/dt$. □

Theorem 1.18. *Let X be a reflexive Banach space and let $u \in L^p(a,b;X)$, $1 < p \le \infty$. Then the following two conditions are equivalent.*

(i) $u \in W^{1,p}([a,b];X)$.
(ii) *There is $C > 0$ such that*

$$\int_a^{b-h} \|u(t+h)-u(t)\|^p dt \le C|h|^p, \qquad \forall h \in [0,b-a]$$

with the usual modification in the case $p=\infty$.

Proof. (i) $\Rightarrow$ (ii). By Theorem 1.17, we know that

$$u(t+h)-u(t) = \int_t^{t+h}\frac{du^0}{ds}(s)ds, \qquad \forall t,t+h\in[a,b],$$

where $u^0 \in A^{1,p}([a,b];X)$ that is, $(du^0/dt) \in L^p(a,b;X)$. This yields via the Hölder inequality and Fubini theorem

$$\int_a^{b-h}\|u(t+h)-u(t)\|^p dt \le |h|^{p-1}\int_a^{b-h}dt\int_t^{t+h}\left\|\frac{du^0}{ds}\right\|^p ds \le |h|^p\int_a^b\left\|\frac{du^0}{ds}\right\|^p ds$$

and this implies estimate (ii).

(ii) $\Rightarrow$ (i). Let u_n be the regularization of u. A simple straightforward computation involving formula (1.45) reveals that $\{u_n'\}$ is bounded in $L^p(a,b;X)$. Because $u_n \to u$ in $L^p(a,b;X)$, $u_n' \to u'$ in $\mathscr{D}'(a,b;X)$, and $\{u_n'\}$ is weakly compact in $L^p(a,b;X)$, which is reflexive, we infer that $u' \in L^p(a,b;X)$, as claimed. □

Remark 1.2. If $u \in W^{1,1}([a,b];X)$, then it follows as above that

$$\int_a^{b-h}\|u(t+h)-u(t)\|dt \le C|h|, \qquad \forall h\in[0,b-a].$$

However, this inequality does not characterize the functions u in $W^{1,1}([a,b];X)$, but the functions u with bounded variation on $[a,b]$.

Let V be a reflexive Banach space and H be a real Hilbert space such that $V \subset H \subset V'$ in the algebraic and topological senses. Here, V' is the dual space of V and H is identified with its own dual. Denote by $|\cdot|$ and $\|\cdot\|$ the norms of H and V, respectively, and by $(\cdot,\cdot)$ the duality between V and V'. If $v_1, v_2 \in H$, then (v_1, v_2) is the scalar product in H of v_1 and v_2.

Denote by $W_p([a,b];V)$, $1 < p < \infty$, the space

$$W_p([a,b];V) = \{u \in L^p(a,b;V);\ u' \in L^q(a,b;V')\}, \qquad \frac{1}{p} + \frac{1}{q} = 1, \tag{1.46}$$

where u' is the derivative of u in the sense of $\mathscr{D}'(a,b;V)$. By Theorem 1.17, we know that every $u \in W_p([a,b];V)$ can be identified with an absolutely continuous function $u^0 : [a,b] \to V'$. However, we have a more precise result.

Theorem 1.19. *Let $u \in W_p([a,b];V)$. Then there is a continuous function $u^0 : [a,b] \to H$ such that $u(t) = u^0(t)$, a.e., $t \in (a,b)$. Moreover, if $u, v \in W_p([a,b];V)$, then the function $t \to (u(t), v(t))$ is absolutely continuous on $[a,b]$ and*

$$\frac{d}{dt}(u(t),v(t)) = (u'(t),v(t)) + (u(t),v'(t)), \quad \textit{a.e. } t \in (a,b). \tag{1.47}$$

Proof. Let $u, v \in W_p([a,b];V)$ and $\psi(t) = (u(t), v(t))$. As we have seen in Theorem 1.17, we may assume that $u, v \in AC([a,b];V')$ and

$$\lim_{\varepsilon \downarrow 0} \int_a^{b-\varepsilon} \left\| \frac{u(t+\varepsilon) - u(t)}{\varepsilon} - u'(t) \right\|_{V'}^q dt = 0,$$

$$\lim_{\varepsilon \downarrow 0} \int_a^{b-\varepsilon} \left\| \frac{v(t+\varepsilon) - v(t)}{\varepsilon} - v'(t) \right\|_{V'}^q dt = 0.$$

Then, we have, by the Hölder inequality,

$$\lim_{\varepsilon \downarrow 0} \int_a^{b-\varepsilon} \left| \frac{\psi(t+\varepsilon) - \psi(t)}{\varepsilon} - (u'(t),v(t)) - (u(t),v'(t)) \right| dt = 0.$$

Hence, $\psi \in W^{1,1}([a,b];\mathbf{R})$ and $(d\psi/dt)(t) = (u'(t),v(t)) + (u(t),v'(t))$, a.e. $t \in (a,b)$, as claimed.

Now, in equation (1.47) we take $v = u$ and integrate from s to t. We get

$$\frac{1}{2}(|u(t)|^2 - |u(s)|^2) = \int_s^t (u'(\tau),u(\tau))d\tau.$$

Hence, the function $t \to |u(t)|$ is continuous. On the other hand, for every $v \in V$ the function $t \to (u(t), v)$ is continuous. Because $|u(t)|$ is bounded on $[a,b]$, this implies that for every $v \in H$ the function $t \to (u(t), v)$ is continuous; that is, $u(t)$ is H-weakly continuous. Then, from the obvious equation

$$|u(t)-u(s)|^2 = |u(t)|^2 + |u(s)|^2 - 2(u(t),u(s)), \qquad \forall t,s \in [a,b]$$

it follows that $\lim_{s\to t} |u(t)-u(s)| = 0$, as claimed. □

The spaces $W^{1,p}([a,b];X)$, as well as $W_p([a,b];V)$, play an important role in the theory of differential equations in infinite-dimensional spaces.

The following compactness result, which is a sharpening of the Arzelà–Ascoli theorem, is particularly useful in this context.

Theorem 1.20 (Aubin). *Let* X_0, X_1, X_2 *be Banach spaces such that* $X_0 \subset X_1 \subset X_2$, X_i *reflexive for* $i = 0,1,2$, *and the injection of* X_0 *into* X_1 *is compact. Let* $1 < p_i < \infty$, $i = 0,1$. *Then the space*

$$W = L^{p_0}(a,b;X_0) \cap W^{1,p_1}([a,b];X_2)$$

is compactly embedded in $L^{p_0}(a,b;X_1)$.

The proof relies on the following property of the spaces X_i (see Lions [11], p. 58). For every $\varepsilon > 0$ there exists $C_\varepsilon > 0$ such that

$$\|u\|_{X_1} \le \varepsilon \|u\|_{X_0} + C_\varepsilon \|u\|_{X_2}, \qquad \forall u \in X_0.$$

References

1. D. Adams, *Sobolev Spaces*, Academic Press, San Diego, 1975.
2. S. Agmon, A. Douglis, L. Nirenberg, Estimates near the boundary for solutions of elliptic partial differential equations satisfying general boundary conditions, *Comm. Pure Appl. Math.*, **12** (1959), pp. 623–727.
3. L. Ambrosio, N. Fusco, D. Pallara, *Functions of Bounded Variations and Free Discontinuous Processes*, Oxford University Press, Oxford, UK, 2000.
4. E. Asplund, Average norms, *Israel J. Math.,* **5** (1967), pp. 227–233.
5. V. Barbu, *Partial Differential Equations and Boundary Value Problems*, Kluwer, Dordrecht, 1998.
6. V. Barbu, T. Precupanu, *Convexity and Optimization in Banach Spaces*, D. Reidel, Dordrecht, 1986.
7. H. Brezis, *Analyse Fonctionnelle. Théorie et Applications*, Masson, Paris, 1983.
8. H. Brezis, *Opérateurs Maximaux Monotones et Semigroupes de Contractions dans un Espace de Hilbert*, North-Holland, Amsterdam, 1973.
9. R.E. Edwards, *Functional Analysis*, Holt, Rinehart and Winston, New York, 1965.
10. P. Grisvard, *Elliptic Problems in Non Smooth Domains*, Pitman Advanced Publishing Program, Boston, 1984.
11. J.L. Lions, *Quelques Méthodes de Resolution des Problèmes aux Limites Nonlinéaires*, Dunod-Gauthier Villars, Paris, 1969.
12. Y. Komura, Nonlinear semigroups in Hilbert spaces, *J. Math. Soc. Japan*, **19** (1967), pp. 508–520.
13. G. Köthe, *Topological Vector Spaces*, Springer-Verlag, Berlin, 1969.
14. J.J. Moreau, *Fonctionnelles Convexes*, Seminaire sur les équations aux dérivées partielles, Collège de France, Paris, 1966–1967.
15. R.T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton, NJ, 1969.
16. K. Yosida, *Functional Analysis,* Springer-Verlag, New York, 1980.
17. C. Zălinescu, *Convex Analysis in General Vector Spaces*, World Scientific, Singapore, 2002.

Chapter 2
Maximal Monotone Operators in Banach Spaces

Abstract In this chapter we present the basic theory of maximal monotone operators in reflexive Banach spaces along with its relationship and implications in convex analysis and existence theory of nonlinear elliptic boundary value problems. However, the latter field is not treated exhaustively but only from the perspective of its implications to nonlinear dynamics in Banach spaces.

2.1 Minty–Browder Theory of Maximal Monotone Operators

If X and Y are two linear spaces, we denote by $X \times Y$ their Cartesian product. The elements of $X \times Y$ are written as $[x,y]$, where $x \in X$ and $y \in Y$.

If A is a multivalued operator from X to Y, we may identify it with its graph in $X \times Y$:

$$\{[x,y] \in X \times Y;\ y \in Ax\}. \tag{2.1}$$

Conversely, if $A \subset X \times Y$, then we define

$$Ax = \{y \in Y;\ [x,y] \in A\}, \qquad D(A) = \{x \in X;\ Ax \neq \emptyset\}, \tag{2.2}$$

$$R(A) = \bigcup_{x \in D(A)} Ax, \qquad A^{-1} = \{[y,x];\ [x,y] \in A\}. \tag{2.3}$$

In this way, here and in the following we identify the operators from X to Y with their graphs in $X \times Y$ and so we equivalently speak of subsets of $X \times Y$ instead of operators from X to Y.

If $A, B \subset X \times Y$ and λ is a real number, we set:

$$\lambda A = \{[x, \lambda y];\ [x,y] \in A\}; \tag{2.4}$$

$$A + B = \{[x, y+z];\ [x,y] \in A,\ [x,z] \in B\}; \tag{2.5}$$

$$AB = \{[x,z];\ [x,y] \in B,\ [y,z] \in A \text{ for some } y \in Y\}. \tag{2.6}$$

V. Barbu, *Nonlinear Differential Equations of Monotone Types in Banach Spaces*, Springer Monographs in Mathematics, DOI 10.1007/978-1-4419-5542-5_2,

Throughout this chapter, X is a real Banach space with dual X^*. Notations for norms, convergence, and duality pairings are as introduced in Chapter 1, Section 1.1. In particular, the value of functional $x^* \in X^*$ at $x \in X$ is denoted by either (x, x^*) or (x^*, x). For the sake of simplicity, we denote by the same symbol $\|\cdot\|$ the norm of X and of X^*. If X is a Hilbert space unless otherwise stated we implicitly assume that it is identified with its own dual.

Definition 2.1. The set $A \subset X \times X^*$ (equivalently the operator $A : X \to X^*$) is said to be *monotone* if

$$(x_1 - x_2, y_1 - y_2) \geq 0, \qquad \forall [x_i, y_i] \in A,\ i = 1,2. \tag{2.7}$$

A monotone set $A \subset X \times X^*$ is said to be *maximal monotone* if it is not properly contained in any other monotone subset of $X \times X^*$.

Note that if A is a single-valued operator from X to X^*, then A is monotone if

$$(x_1 - x_2, Ax_1 - Ax_2) \geq 0, \qquad \forall x_1, x_2 \in D(A). \tag{2.8}$$

A simple example of a monotone subset of $X \times X^*$ is the duality mapping J of X. (See Section 1.1.) Indeed, by definition of J,

$$(x_1-x_2, y_1-y_2) = \|x_1\|^2 + \|x_2\|^2 - (x_1, y_2) - (x_2, y_1) \geq (\|x_1\| - \|x_2\|)^2, \quad \forall [x_i, y_i] \in J.$$

As a matter of fact, it turns out that J is maximal monotone in $X \times X^*$. Indeed, if $[u, v] \in X \times X^*$ is such that $(u - x, v - y) \geq 0$, $\forall [x, y] \in J$, then, because $J : X \to X^*$ is onto, there is $[x, y] \in J$ such that

$$2y = v + w, \qquad w \in J(u).$$

This yields

$$(u - x, w - y) \leq 0$$

and because $[u, w], [x, y] \in J$ we get

$$\|x\|^2 = \|y\|^2 = \|u\|^2 = \|w\|^2, \qquad (u, y) + (x, w) \geq 2\|x\|^2.$$

Hence,

$$(u, y) + (x, w) = 2\|x\|^2 = 2\|u\|^2$$

and this, clearly, implies that

$$(u, y) = (x, w) = (x, v) = \|u\|^2 = \|x\|^2.$$

Hence,

$$(u, v) \geq (x, v) + (u, y) - (x, y) = \|u\|^2 = \|v\|^2$$

and therefore $[u, v] \in J$, as claimed.

Definition 2.2. Let A be a single-valued operator from X to X^* with $D(A) = X$. The operator A is said to be *hemicontinuous* if, for all $x, y \in X$,

$$w\text{-}\lim_{\lambda \to 0} A(x + \lambda y) = Ax.$$

A is said to be *demicontinuous* if it is continuous from X to X^*_w; that is,

$$w\text{-}\lim_{x_n \to x} Ax_n = Ax.$$

A is said to be *coercive* if

$$\lim_{n \to \infty} (x_n - x^0, y_n) \|x_n\|^{-1} = \infty \tag{2.9}$$

for some $x^0 \in X$ and all $[x_n, y_n] \in A$ such that $\lim_{n \to \infty} \|x_n\| = \infty$.

A is said to be bounded if it is bounded on each bounded subset.

Proposition 2.1. *Let $A \subset X \times X^*$ be maximal monotone. Then:*

(i) *A is weakly–strongly closed in $X \times X^*$; that is, if $y_n = Ax_n$, $x_n \rightharpoonup x$ in X, and $y_n \to y$ in X^*, then $[x, y] \in A$,*
(ii) *A^{-1} is maximal monotone in $X^* \times X$,*
(iii) *For each $x \in D(A)$, Ax is a closed convex subset of X^*.*

Proof. (i) From the obvious inequality

$$(x_n - u, y_n - v) \geq 0, \qquad \forall [u, v] \in A,$$

we see that $(x - u, y - v) \geq 0$, $\forall [u, v] \in A$, and because A is maximal, this implies $[x, y] \in A$, as claimed.

(ii) This is obvious.

(iii) By (i) it is clear that Ax is a closed subset of X^* for each $x \in D(A)$. Now, let $y_0, y_1 \in Ax$ and let $y_\lambda = \lambda y_0 + (1 - \lambda) y_1$, where $0 < \lambda < 1$. From the inequalities

$$(x - u, y_0 - v) \geq 0, \qquad (x - u, y_1 - v) \geq 0, \qquad \forall [u, v] \in A,$$

we see that $(x - u, y_\lambda - v) \geq 0$, $\forall [u, v] \in A$, which implies that $[x, y_\lambda] \in A$ because A is maximal. The proof is complete. $\square$

It has been shown by G. Minty in the early 1960s that the coercive maximal monotone operators are surjective. This important result, which implies a characterization of a maximal monotone operator A in terms of the surjectivity of $A + J$ (J is the duality mapping) is a consequence of the following existence theorem.

Theorem 2.1. *Let X be a reflexive Banach space and let A and B be two monotone sets of $X \times X^*$ such that $0 \in D(A)$, B is single-valued, hemicontinuous, and coercive; that is,*

$$\lim_{\|x\|\to\infty} \frac{(x,Bx)}{\|x\|} = +\infty. \tag{2.10}$$

Then there exists $x \in K = \overline{\operatorname{conv} D(A)}$ *such that*

$$(u-x, Bx+v) \geq 0 \qquad \forall [u,v] \in A. \tag{2.11}$$

Here, $\overline{\operatorname{conv} D(A)}$ is the convex hull of the set $\overline{D(A)}$; that is, the set

$$\left\{ \sum_{i=1}^{m} \lambda_i x_i,\ x_i \in \overline{D(A)},\ 0 \leq \lambda_i \leq 1,\ \sum_{i=1}^{m} \lambda_i = 1,\ m \in \mathbf{N} \right\}.$$

In particular, if A is maximal monotone, it follows from (2.11) that $0 \in Ax + Bx$.

We first prove the following lemma.

Lemma 2.1. *Let X be a finite-dimensional Banach space and let B be a hemicontinuous monotone operator from X to X^*. Then B is continuous.*

Proof. Let us show first that B is bounded on bounded subsets. Indeed, otherwise there exists a sequence $\{x_n\} \subset X$ such that $\|Bx_n\| \to \infty$ and $x_n \to x_0$ as $n \to \infty$. We have

$$(x_n - x, Bx_n - Bx) \geq 0, \qquad \forall x \in X,$$

and therefore

$$\left(x_n - x,\ \frac{Bx_n}{\|Bx_n\|} - \frac{Bx}{\|Bx_n\|} \right) \geq 0,\ \ \forall x \in X.$$

Without loss of generality, we may assume that $Bx_n\|Bx_n\|^{-1} \to y_0$ as $n \to \infty$. This yields

$$(x_0 - x, y_0) \geq 0, \qquad \forall x \in X,$$

and therefore $y_0 = 0$. The contradiction can be eliminated only if B is bounded. Now, let $\{x_n\}$ be convergent to x_0 and let y_0 be a cluster point of $\{Bx_n\}$. Again by the monotonicity of B, we have

$$(x_0 - x, y_0 - Bx) \geq 0, \qquad \forall x \in X.$$

If in this inequality we take $x = tu + (1-t)x_0$, $0 \leq t \leq 1$, u arbitrary in X, we get

$$(x_0 - u, y_0 - B(tu + (1-t)x_0)) \geq 0, \qquad \forall t \in [0,1],\ u \in X.$$

Then, letting t tend to zero and using the hemicontinuity of B, we get

$$(x_0 - u, y_0 - Bx_0) \geq 0, \qquad \forall u \in X,$$

which clearly implies that $y_0 = Bx_0$, as claimed. □

The next step in the proof of Theorem 2.1 is the case where X is finite-dimensional.

Lemma 2.2. *Let X be a finite-dimensional Banach space and let A and B be two monotone subsets of $X \times X^*$ such that $0 \in \underline{D(A)}$, and B is single-valued, continuous, and satisfies* (2.10). *Then there exists $x \in \overline{\operatorname{conv} D(A)}$ such that*

$$(u-x, Bx+v) \geq 0, \qquad \forall [u,v] \in A. \tag{2.12}$$

Proof. Redefining A if necessary, we may assume that the set $K = \overline{\operatorname{conv} D(A)}$ is bounded. Indeed, if Lemma 2.1 is true in this case, then replacing A by $A_n = \{[x,y] \in A;\ \|x\| \leq n\}$, we infer that for every n there exists $x_n \in K_n = K \cap \{x;\ \|x\| \leq n\}$ such that

$$(u-x_n, Bx_n+v) \geq 0, \qquad \forall [u,v] \in A_n. \tag{2.13}$$

This yields

$$(x_n, Bx_n)\|x_n\|^{-1} \leq \|\xi\|, \quad \text{for some } \xi \in A0,$$

and, by the coercivity condition (2.10), we see that there is $M > 0$ such that $\|x_n\| \leq M$ for all n. Now, on a subsequence, for simplicity again denoted n, we have $x_n \to x$. By (2.13) and the continuity of B, it is clear that x is a solution to (2.12), as claimed.

Let $T : K \to K$ be the multivalued operator defined by

$$Tx = \{y \in K;\ (u-y, Bx+v) \geq 0, \qquad \forall [u,v] \in A\}.$$

Let us show first that $Tx \neq \emptyset$, $\forall x \in K$. To this end, define the sets

$$K_{uv} = \{y \in K;\ (u-y, Bx+v) \geq 0\},$$

and notice that

$$Tx = \bigcap_{[u,v] \in A} K_{uv}.$$

Inasmuch as K_{uv} are closed subsets (if nonempty) of the compact set K, to show that $\bigcap_{[u,v]\in A} K_{uv} \neq \emptyset$ it suffices to prove that every finite collection $\{K_{u_i,v_i};\ i = 1,\dots,m\}$ has a nonempty intersection. Equivalently, it suffices to show that the system

$$(u_i - y, Bx + v_i) \geq 0, \qquad i = 1,\dots,m, \tag{2.14}$$

has a solution $y \in K$ for any set of pairs $[u_i, v_i] \in A$, $i = 1,\dots,m$.

Consider the function $H : U \times U \to \mathbf{R}$,

$$H(\lambda,\mu) = \sum_{i=1}^{m} \mu_i \left(\sum_{j=1}^{m} \lambda_j u_j - u_i, Bx + v_i \right), \qquad \forall \lambda, \mu \in U, \tag{2.15}$$

where

$$U = \left\{ \lambda \in \mathbf{R}^m;\ \lambda = (\lambda_1,\dots,\lambda_m),\ \lambda_i \geq 0,\ \sum_{i=1}^{m} \lambda_i = 1 \right\}.$$

The function H is continuous, convex in λ, and concave in μ. Then, according to the classical Von Neumann min–max theorem from game theory, it has a saddle point $(\lambda_0,\mu_0) \in U \times U$; that is,

$$H(\lambda_0,\mu) \le H(\lambda_0,\mu_0) \le H(\lambda,\mu_0), \qquad \forall \lambda,\mu \in U. \tag{2.16}$$

On the other hand, we have

$$\begin{aligned} H(\lambda,\lambda) &= \sum_{i=1}^{m} \lambda_i \left(\sum_{j=1}^{m} \lambda_j u_j - u_i, Bx + v_i \right) \\ &= \sum_{i=1}^{m}\sum_{j=1}^{m} \lambda_i\lambda_j (v_i, u_j - u_i) + \sum_{i=1}^{m}\sum_{j=1}^{m} \lambda_i\lambda_j (u_j - u_i, Bx) \le 0, \qquad \forall \lambda \in U, \end{aligned}$$

because, by monotonicity of B, $(v_i - v_j, u_i - u_j) \ge 0$ for all i,j.

Then, by (2.16) we see that

$$H(\lambda_0,\mu) \le 0, \ \ \forall \mu \in U;$$

that is,

$$\sum_{i=1}^{m} \mu_i \left(\sum_{j=1}^{m} (\lambda_0)_j u_j - u_i, Bx + v_i \right) \le 0, \qquad \forall \mu \in U.$$

In particular, it follows that

$$\left(\sum_{j=1}^{m} (\lambda_0)_j u_j - u_i, Bx + v_i \right) \le 0, \qquad \forall i = 1,\dots,m.$$

Hence, $y = \sum_{j=1}^{m} (\lambda_0)_j u_j \in K$ is a solution to (2.14). We have therefore proved that T is well defined on K and that $T(K) \subset K$. It is also clear that for every $x \in K$, Tx is a closed convex subset of X and T is upper semicontinuous on K. Indeed, because the range of T belongs to a compact set, to verify that T is upper-semicontinuous it suffices to show that T is closed in $K \times K$; that is, if $[x_n,y_n] \in T$, $x_n \to x$, and $y_n \to y$, then $y \in Tx$. But the last property is obvious if one takes into account the definition of T. Then, applying the classical Kakutani fixed point theorem (see, e.g., Deimling [11]), we conclude that there exists $x \in K$ such that $x \in Tx$, thereby completing the proof of Lemma 2.2. □

Proof of Theorem 2.1. The proof relies on standard finite-dimensional approximations of equations in Banach spaces (the Galerkin method). Let Λ be the family of all finite dimensional subspaces X_α of X ordered by the inclusion relation. For every $X_\alpha \in \Lambda$, denote by $j_\alpha : X_\alpha \to X$ the injection mapping of X_α into X and by $j_\alpha^* : X^* \to X_\alpha^*$ the dual mapping; that is, the projection of X^* onto X_α^*. The operators $A_\alpha = j_\alpha^* A j_\alpha$ and $B_\alpha = j_\alpha^* B j_\alpha$ map X_α into X_α^* and are monotone in $X_\alpha \times X_\alpha^*$. Because B is hemicontinuous from X to X^* and the j_α^* are continuous from X^* to X_α^* it follows by Lemma 2.1 that B_α is continuous from X_α to X_α^*.

We may therefore apply Lemma 2.2, where $X = X_\alpha$, $A = A_\alpha$, $B = B_\alpha$, and $K = K_\alpha = \overline{\mathrm{conv}\, D(A_\alpha)}$. Hence, for each $X_\alpha \in \Lambda$, there exists $x_\alpha \in K_\alpha$ such that

$$(u - x_\alpha, B_\alpha x_\alpha + v) \geq 0, \qquad \forall [u,v] \in A,$$

or, equivalently,

$$(u - x_\alpha, Bx_\alpha + v) \geq 0, \qquad \forall [u,v] \in A_\alpha. \tag{2.17}$$

By using the coercivity condition (2.10), we deduce from (2.17) that $\{x_\alpha\}$ remain in a bounded subset of X. The space X is reflexive, thus every bounded subset of X is sequentially weakly compact and so there exists a sequence $\{x_{\alpha_n}\} \subset \{x_\alpha\}$ such that

$$x_{\alpha_n} \rightharpoonup x \quad \text{in } X \text{ as } n \to \infty. \tag{2.18}$$

Moreover, because the operator B is bounded on bounded subsets, we may assume that

$$Bx_{\alpha_n} \rightharpoonup y \quad \text{in } X^* \text{ as } n \to \infty. \tag{2.19}$$

Because the closed convex subsets are weakly closed, we infer that $x \in K$. Moreover, by (2.17)–(2.19), we see that

$$\limsup_{n\to\infty} (x_{\alpha_n}, Bx_{\alpha_n}) \leq (u - x, v) + (u, y), \qquad \forall [u,v] \in A. \tag{2.20}$$

Without loss of generality, we may assume that A is maximal in the class of all monotone subsets $\widetilde{A} \subset X \times X^*$ such that $D(\widetilde{A}) \subset K = \overline{\mathrm{conv}\, D(A)}$. (If not, we may extend A by Zorn's lemma to a maximal element of this class.) To complete the proof, let us show first that

$$\limsup_{n\to\infty} (x_{\alpha_n} - x, Bx_{\alpha_n}) \leq 0. \tag{2.21}$$

Indeed, if this is not the case, it follows from (2.20) that

$$(u - x, v + y) \geq 0, \qquad \forall [u,x] \in A,$$

and because $x \in K$ and A is maximal in the class of all monotone operators $\widetilde{A}$ with domain in K, it follows that $[x, -y] \in A$. Then, putting $u = x$ in (2.20), we obtain (2.21), which contradicts the working hypothesis.

Now, for u arbitrary but fixed in $D(A)$ consider $u_\lambda = \lambda x + (1-\lambda)u$, $0 \leq \lambda \leq 1$, and notice that, by virtue of the monotonicity of B, we have

$$(x_{\alpha_n} - u_\lambda, Bx_{\alpha_n}) \geq (x_{\alpha_n} - u_\lambda, Bu_\lambda).$$

This yields

$$(1-\lambda)(x_{\alpha_n} - u, Bx_{\alpha_n}) + \lambda(x_{\alpha_n} - x, Bx_{\alpha_n}) \geq (1-\lambda)(x_{\alpha_n} - u, Bu_\lambda) + \lambda(x_{\alpha_n} - x, Bu_\lambda)$$

and so, by (2.20) and (2.21),

$$(x-u,Bu_\lambda) \le \limsup_{n\to\infty}(x_{\alpha_n}-u,Bx_{\alpha_n}) \le (u-x,v), \qquad \forall [u,v]\in A.$$

Inasmuch as B is hemicontinuous, the latter inequality yields for $\lambda \to 1$,

$$(u-x,v+Bx)\ge 0, \qquad \forall [u,v]\in A,$$

thereby completing the proof of Theorem 2.1.

We now use Theorem 2.1 to prove a fundamental result in the theory of maximal monotone operators due to G. Minty [19] and F. Browder [9] and which has opened the way to applications of the existence theory of nonlinear operatorial equations of monotone type.

Theorem 2.2. *Let X and X^* be reflexive and strictly convex. Let $A\subset X\times X^*$ be a monotone subset of $X\times X^*$ and let $J: X\to X^*$ be the duality mapping of X. Then A is maximal monotone if and only if, for any $\lambda>0$ (equivalently, for some $\lambda>0$), $R(A+\lambda J)=X^*$.*

Proof. "If" part. Assume that $R(A+\lambda J)=X^*$ for some $\lambda>0$. We suppose that A is not maximal monotone, and argue from this to a contradiction. If A is not maximal monotone, there exists $[x_0,y_0]\in X\times X^*$ such that $[x_0,y_0]\notin A$ and

$$(x-x_0,y-y_0)\ge 0, \qquad \forall [x,y]\in A. \tag{2.22}$$

On the other hand, by hypothesis, there exists $[x_1,y_1]\in A$ such that

$$\lambda J(x_1)+y_1=\lambda J(x_0)+y_0.$$

Substituting $[x_1,y_1]$ in place of $[x,y]$ in (2.22), this yields

$$(x_1-x_0,J(x_1)-J(x_0))\le 0.$$

Taking into account the definition of J, we get

$$\|x_1\|^2+\|x_0\|^2\le (x_1,J(x_0))+(x_0,J(x_1)),$$

and therefore

$$(x_1,J(x_0))=(x_0,J(x_1))=\|x_1\|^2=\|x_0\|^2.$$

Hence

$$J(x_0)=J(x_1),$$

and, because the duality mapping J^{-1} of X^* is single-valued (because X is strictly convex), we infer that $x_0=x_1$. Hence $[x_0,y_0]=[x_1,y_1]\in A$, which contradicts the hypothesis.

"Only if" part. The space X^* being strictly convex, J is single-valued and demicontinuous on X (Theorem 1.2). Let y_0 be an arbitrary element of X^* and let $\lambda>0$. Applying Theorem 2.1, where

$$Bu = \lambda J(u) - y_0, \qquad \forall u \in X,$$

we conclude that there is $x \in X$ such that

$$(u - x, \lambda J(x) - y_0 + v) \geq 0, \qquad \forall [u,v] \in A.$$

A is maximal monotone, therefore this implies that $[x, -\lambda J(x) + y_0] \in A$; that is, $y_0 \in \lambda J(x) + Ax$. Applying Theorem 2.1, we have implicitly assumed that $0 \in D(A)$. If not, we apply this theorem to $Bu = \lambda J(u + u_0) - y_0$ and $Au \overset{\text{def}}{=\!=} A(u + u_0)$, where $u_0 \in D(A)$. □

We later show that the assumption that X^* is strictly convex can be dropped in Theorem 2.2.

Let $\Phi_p(x) = J(x)\|x\|^{p-1}$, where $p > 0$. Theorem 2.2 extends to the case where J is replaced by Φ_p. We have the following theorem.

Theorem 2.3. *Let X and X^* be reflexive and strictly convex and let $A \subset X \times X^*$ be a monotone set. Then A is maximal monotone if and only if, for each $\lambda > 0$ and $p > 0$, $R(A + \lambda \Phi_p) = X^*$.*

Proof. The proof is exactly the same as that of Theorem 2.2, so it is only outlined.

If $R(A + \lambda \Phi_p) = X^*$ and if $[x_0, y_0] \in X \times X^*$ is such that

$$(x - x_0, y - y_0) \geq 0, \qquad \forall [x,y] \in A$$

then, choosing $[x_1, y_1] \in A$ such that

$$\lambda \Phi_p(x_1) + y_1 = \lambda \Phi_p(x_0) + y_0$$

and, substituting into the above inequality, we obtain

$$(x_1 - x_0, J(x_1)\|x_1\|^{p-1} - J(x_0)\|x_0\|^{p-1}) \leq 0$$

and this yields as above

$$(x_1, J(x_0)) = (x_0, J(x_1)) = \|x_0\|^2 = \|x_1\|^2;$$

that is, $J(x_0) = J(x_1)$ and $x_0 = x_1$. Hence

$$[x_0, y_0] = [x_1, y_1].$$

”The only if part” follows exactly as in the proof of Theorem 2.2. □

Now, we use Theorem 2.1 to derive a maximality criterion for the sum $A + B$.

Corollary 2.1. *Let X be reflexive and let B be a hemicontinuous monotone and bounded operator from X to X^*. Let $A \subset X \times X^*$ be maximal monotone. Then $A + B$ is maximal monotone.*

Proof. By Asplund's theorem (Theorem 1.1 in Chapter 1), we may take an equivalent norm in X such that X and X^* are strictly convex. It is clear that after this operation the monotonicity properties of $A, B, A+B$ as well as maximality do not change. Also, without loss of generality, we may assume that $0 \in D(A)$; otherwise, we replace A by $u \to A(u+u_0)$, where $u_0 \in D(A)$ and B by $u \to B(u+u_0)$. Let y_0 be arbitrary but fixed in X^*. Now, applying Theorem 2.1, where B is this time the operator $u \to Bu + J(u) - y_0$, we infer that there is an $x \in \overline{\mathrm{conv}\, D(A)}$ such that

$$(u-x, J(x)+Bx-y_0+v) \geq, \qquad \forall [u,v] \in A.$$

(Because $(u, Bu+J(u)-y_0) \geq (u,Bu)+\|u\|^2-\|y_0\|\,\|u\| \geq \|u\|^2-\|B0\|\,\|u\|-\|y_0\|\,\|u\|$, clearly condition (2.10) holds.) As A is maximal monotone, this yields

$$y_0 \in Ax + Bx + J(x),$$

as claimed. □

In particular, it follows by Corollary 2.1 that every monotone, hemicontinuous, and bounded operator from X to X^* is maximal monotone. We now prove that the boundedness assumption is redundant.

Theorem 2.4. *Let X be a reflexive Banach space and let $B : X \to X^*$ be a monotone hemicontinuous operator. Then B is maximal monotone in $X \times X^*$.*

Proof. Suppose that B is not maximal monotone. Then, there exists $[x_0, y_0] \in X \times X^*$ such that $y_0 \neq Bx_0$ and

$$(x_0 - u, y_0 - Bu) \geq 0, \qquad \forall u \in X. \tag{2.23}$$

For any $x \in X$, we set $u_\lambda = \lambda x_0 + (1-\lambda)x$, $0 \leq \lambda \leq 1$, and put $u = u_\lambda$ in (2.23). We get

$$(x_0 - x, y_0 - Bu_\lambda) \geq 0, \qquad \forall \lambda \in [0,1],\ u \in X,$$

and, letting λ tend to 1,

$$(x_0 - x, y_0 - Bx_0) \geq 0, \qquad \forall x \in X.$$

Hence $y_0 = Bx_0$, which contradicts the hypothesis. □

Corollary 2.2. *Let X be a reflexive Banach space and let A be a coercive maximal monotone subset of $X \times X^*$. Then A is surjective; that is, $R(A) = X^*$.*

Proof. Let $y_0 \in X^*$ be arbitrary but fixed. Without loss of generality, we may assume that X, X^* are strictly convex, so that by Theorem 2.2 for every $\lambda > 0$ the equation

$$\lambda J(x_\lambda) + Ax_\lambda \ni y_0 \tag{2.24}$$

has a (unique) solution $x_\lambda \in D(A)$. Now, we multiply (in the sense of the duality pairing $(\cdot,\cdot)$) equation (2.24) by $x_\lambda - x^0$, where x^0 is the element arising in the coercivity condition (2.9). We have

$$\lambda \|x_\lambda\|^2 + (x_\lambda - x^0, Ax_\lambda) = (x_\lambda - x^0, y_0) + \lambda (x_0, Jx_\lambda).$$

By (2.9), we deduce that $\{x_\lambda\}$ is bounded in X as $\lambda \to 0$ and so we may assume (taking a subsequence if necessary) that $\exists x_0 \in X$ such that

$$w\text{-}\lim_{\lambda \downarrow 0} x_\lambda = x_0.$$

Letting λ tend to zero in (2.24), we see that

$$\lim_{\lambda \downarrow 0} Ax = y_0.$$

Because, as seen earlier, maximal monotone operators are weakly–strongly closed in $X \times X^*$, we conclude that $y_0 \in Ax_0$. Hence $R(A) = X^*$, as claimed. □

In particular, the next corollary follows by Corollary 2.2 and Theorem 2.4.

Corollary 2.3. *A monotone, hemicontinuous, and coercive operator B from a reflexive Banach space X to its dual X^* is surjective.*

The Sum of Two Maximal Monotone Operators

A problem of great interest because of its implications for the existence theory for partial differential equations is to know whether the sum of two maximal monotone operators is again maximal monotone. Before answering this question, let us first establish some facts related to Yosida approximation of the maximal monotone operators.

Let us assume that X is a reflexive strictly convex Banach space with strictly convex dual X^*, and let A be maximal monotone in $X \times X^*$.

According to Corollaries 2.1 and 2.2, for every $x \in X$ the equation

$$0 \in J(x_\lambda - x) + \lambda Ax_\lambda \tag{2.25}$$

has a solution x_λ. Inasmuch as

$$(x - u, Jx - Ju) \geq (\|x\| - \|u\|)^2, \qquad \forall x, u \in X,$$

and J^{-1} is single-valued (because X is strictly convex), it is readily seen that x_λ is unique. Define

$$\begin{aligned} J_\lambda x &= x_\lambda, \\ A_\lambda x &= \lambda^{-1} J(x - x_\lambda), \end{aligned} \tag{2.26}$$

for any $x \in X$ and $\lambda > 0$.

The operator $A_\lambda : X \to X^*$ is called the *Yosida approximation* of A and plays an important role in the smooth approximation of A. We collect in Proposition 2.2 several basic properties of the operators A_λ and J_λ.

Proposition 2.2. *Let X and X^* be strictly convex and reflexive. Then:*

(i) A_λ *is single-valued, monotone, bounded, and demicontinuous from X to X^*.*
(ii) $\|A_\lambda x\| \le |Ax| = \inf\{\|y\|;\ y \in Ax\}$ *for every* $x \in D(A),\ \lambda > 0$.
(iii) $J_\lambda : X \to X$ *is bounded on bounded subsets and*

$$\lim_{\lambda\to 0} J_\lambda x = x, \qquad \forall x \in \overline{\operatorname{conv} D(A)}. \tag{2.27}$$

(iv) *If* $\lambda_n \to 0,\ x_n \to x,\ A_{\lambda_n} x_n \rightharpoonup y$ *and*

$$\limsup_{n,m\to\infty} (x_n - x_m, A_{\lambda_n} x_n - A_{\lambda_m} x_m) \le 0, \tag{2.28}$$

then $[x,y] \in A$ *and*

$$\lim_{m,n\to\infty} (x_n - x_m, A_{\lambda_n} x_n - A_{\lambda_m} x_m) = 0.$$

(v) *For* $\lambda \to 0,\ A_\lambda x \rightharpoonup A^0 x,\ \forall x \in D(A)$, *where $A^0 x$ is the element of minimum norm in Ax; that is,* $\|A^0 x\| = |Ax|$. *If X^* is uniformly convex, then* $A_\lambda x \to A^0 x$, $\forall x \in D(A)$.

The main ingredient of the proof is the following lemma which has an intrinsic interest.

Lemma 2.3. *Let X be a reflexive Banach space and let A be a maximal monotone subset of $X \times X^*$. Let $[u_n, v_n] \in A$ be such that $u_n \rightharpoonup u$, $v_n \rightharpoonup v$, and either*

$$\limsup_{n,m\to\infty} (u_n - u_m, v_n - v_m) \le 0 \tag{2.29}$$

or

$$\limsup_{n\to\infty} (u_n - u, v_n - v) \le 0. \tag*{(2.29)$'$}$$

Then $[u,v] \in A$ and $(u_n, v_n) \to (u,v)$ as $n \to \infty$.

Proof. Assume first that condition (2.29) holds. Because A is monotone, we have

$$\lim_{n,m\to\infty} (u_n - u_m, v_n - v_m) = 0.$$

Let $n_k \to \infty$ be such that $(u_{n_k}, v_{n_k}) \to \mu$. Then, clearly, we have $\mu \le (u,v)$. Hence

$$\limsup_{n\to\infty} (u_n, v_n) \le (u,v),$$

and by monotonicity of A we have

$$(u_n - x, v_n - y) \ge 0, \qquad \forall [x,y] \in A,$$

and therefore

$$(u-x, v-y) \geq 0, \qquad \forall [x,y] \in A,$$

which implies $[u,v] \in A$ because A is maximal monotone. The second part of the lemma follows by the same argument. □

Proof of Proposition 2.2. (i) We have

$$(x-y, A_\lambda x - A_\lambda y) = (J_\lambda x - J_\lambda y, A_\lambda x - A_\lambda y) + ((x - J_\lambda x) - (y - J_\lambda y), A_\lambda x - A_\lambda y),$$

and because $A_\lambda x \in AJ_\lambda x$, we infer that

$$(x-y, A_\lambda x - A_\lambda y) \geq 0$$

because A and J are monotone.

Let $[u,v] \in A$ be arbitrary but fixed. If we multiply equation (2.25) by $J_\lambda x - u$ and use the monotonicity of A, we get

$$(J_\lambda x - u, J(J_\lambda x - x)) \leq \lambda (u - J_\lambda x, v),$$

which yields

$$\|J_\lambda x - x\|^2 \leq \|x-u\| \, \|J_\lambda x - x\| + \lambda \|x-u\| \, \|v\| + \lambda \|v\| \, \|J_\lambda x - x\|.$$

This implies that J_λ and A_λ are bounded on bounded subsets.

Now, let $x_n \to x_0$ in X. We set $u_n = J_\lambda x_n$ and $v_n = A_\lambda x_n$. By the equation

$$J(u_n - x_n) + \lambda v_n = 0,$$

it follows that

$$\begin{aligned}((u_n - x_n) \, &-(u_m - x_m), J(u_n - x_n) - J(u_m - x_m)) + \lambda (u_n - u_m, v_n - v_m) \\ &+ \lambda (x_m - x_n, v_n - v_m) = 0.\end{aligned}$$

Because, as seen previously, J_λ is bounded, this yields

$$\lim_{n,m \to \infty} (u_n - u_m, v_n - v_m) \leq 0$$

and

$$\lim_{n,m \to \infty} ((u_n - x_n) - (u_m - x_m), J(u_n - x_n) - J(u_m - x_m)) = 0.$$

Now, let $n_k \to \infty$ be such that $u_{n_k} \rightharpoonup u$, $v_{n_k} \rightharpoonup v$, and $J(u_{n_k} - x_{n_k}) \rightharpoonup w$. By Lemma 2.3, it follows that $[u,v] \in A$, $[u - x_0, w] \in J$, and therefore

$$J(u - x_0) + \lambda v = 0.$$

We have therefore proven that $u = J_\lambda x_0$, $v = A_\lambda x_0$, and by the uniqueness of the limit we infer that $J_\lambda x_n \rightharpoonup J_\lambda x_0$ and $A_\lambda x_n \rightharpoonup A_\lambda x_0$, as claimed.

(ii) Let $[x,x^*] \in A$. Again, by the monotonicity of A, we have

$$0 \leq (x - J_\lambda x, x^* - A_\lambda x) \leq \|x^*\| \, \|x - x_\lambda\| - \lambda^{-1} \|x - x_\lambda\|^2.$$

Hence,

$$\lambda \|A_\lambda x\| = \|x - x_\lambda\| \leq \lambda \|x^*\|, \qquad \forall x^* \in Ax,$$

which implies (ii).

(iii) Let $x \in \overline{\operatorname{conv} D(A)}$ and $[u,u^*] \in A$. We have

$$(J_\lambda x - u, A_\lambda x - u^*) \geq 0,$$

and therefore

$$\|J_\lambda x - x\|^2 \leq \lambda (u - J_\lambda x, u^*) + (u - x, J(J_\lambda x - x)).$$

Let $\lambda_n \to 0$ be such that $J(J_{\lambda_n} x - x) \rightharpoonup y$ in X^*. This yields

$$\overline{\lim_{\lambda_n \to 0}} \, \|J_{\lambda_n} x - x\|^2 \leq (u - x, y).$$

Because u is arbitrary in $D(A)$, the preceding inequality extends to all $u \in \overline{\operatorname{conv} D(A)}$, and in particular we may take $u = x$ and infer that $J_{\lambda_n} x \to x$ for all such sequences $\{\lambda_n\}$. This implies (2.27).

(iv) We have

$$\begin{aligned}(x_n - x_m, A_{\lambda_n} x_n - A_{\lambda_m} x_m) A_{\lambda_m} &= (J_{\lambda_n} x_n - J_{\lambda_m} x_m, A_{\lambda_n} x_n - A_{\lambda_m} x_m) \\ &\quad + ((x_n - J_{\lambda_n} x_n) - (x_m - J_{\lambda_m} x_m), A_{\lambda_n} x_n - A_{\lambda_m} x_m) \\ &\geq ((x_n - J_{\lambda_n} x_n) - (x_m - J_{\lambda_m} x_m), A_{\lambda_n} x_n - A_{\lambda_m} x_m) \\ &= ((x_n - J_{\lambda_n} x_n) - (x_m - J_{\lambda_m} x_m), \lambda_n^{-1} J(x_n - J_{\lambda_n} x_n) \\ &\quad - \lambda_m^{-1} J(x_m - J_{\lambda_m} x_m)).\end{aligned}$$

(Here we have used the monotonicity of A and $A_\lambda x \in AJ_\lambda x$.)

$A_{\lambda_n} x_n = -\lambda_n^{-1}(J_{\lambda_n} x_n - x_n)$ and x_n remain in bounded subsets of X^* and X, respectively, therefore we infer that

$$\lim_{m,n \to \infty} (x_n - x_m, A_{\lambda_n} x_n - A_{\lambda_m} x_m) = 0$$

and

$$\lim_{m,n \to \infty} (J_{\lambda_n} x_n - J_{\lambda_m} x_m, A_{\lambda_n} x_n - A_{\lambda_m} x_m) = 0.$$

Then, by Lemma 2.3 we conclude that $[x,y] \in A$ because

$$\lim_{n \to \infty} (J_{\lambda_n} x_n - x_n) = - \lim_{n \to \infty} \lambda_n J^{-1}(A_{\lambda_n} x_n) = 0.$$

(v) Because Ax is a closed convex subset of X^*, and X^* is reflexive and strictly convex, the projection A^0x of 0 into Ax is well defined and unique.

Now, let $x \in D(A)$ and let $\lambda_n \to 0$ be such that $A_{\lambda_n}x \rightharpoonup y$ in X^*. As seen in the proof of (iv), $y \in Ax$, and because $\|A_{\lambda_n}x\| \le \|A^0x\|$, we infer that $y = A^0x$. Hence, $A_\lambda x \rightharpoonup A^0x$ for $\lambda \to 0$. If X^* is uniformly convex, then, by Lemma 1.1, we conclude that $A_\lambda x \to Ax$ (strongly) in X^* as $\lambda \to 0$.

In general, a maximal monotone operator $A : X \to X^*$ is not weakly–weakly closed, that is from $x_n \rightharpoonup u$ and $v_n \rightharpoonup v$ where $[u_n, v_n] \in A$ does not follow that $[u, v]$ belongs to A. However, by Lemma 2.3 we derive the following result.

Corollary 2.4. *Let X be a reflexive Banach space and let $A \subset X \times X^*$ be a maximal monotone subset. Let $[u_n, v_n] \in A$ be such that $u_n \rightharpoonup u$, $v_n \rightharpoonup v$, and*

$$\limsup_{n\to\infty}(u_n, v_n) \le (u, v).$$

Then, $[u, v] \in A$.

This simple property is, in particular, useful when one passes to the limit in approximating nonlinear equations involving maximal monotone operators.

We also note also the following consequence of Proposition 2.2.

Proposition 2.3. *If $X = H$ is a Hilbert space identified with its own dual, then:*

(i) $J_\lambda = (I + \lambda A)^{-1}$ *is nonexpansive in H (i.e., Lipschitz continuous with Lipschitz constant not greater than 1),*
(ii) $\|A_\lambda x - A_\lambda y\| \le \lambda^{-1}\|x - y\|, \ \forall x, y \in D(A), \ \lambda > 0,$
(iii) $\lim_{\lambda\to 0} A_\lambda x = A^0x, \ \forall x \in D(A).$

Proof. (i) We set $x_\lambda = (I + \lambda A)^{-1}x$, $y_\lambda = (I + \lambda A)^{-1}y$ (I is the unity operator in H). We have

$$x_\lambda - y_\lambda + \lambda(Ax_\lambda - Ay_\lambda) \ni x - y. \tag{2.30}$$

Multiplying by $x_\lambda - y_\lambda$ and using the monotonicity of A, we get

$$\|x_\lambda - y_\lambda\| \le \|x - y\|, \qquad \forall \lambda > 0.$$

Now, multiplying (scalarly in H) equation (2.30) by $Ax_\lambda - Ay_\lambda$, we get (ii).

Regarding (iii), it follows by Proposition 2.1(v). □

Corollary 2.5. *Let X be a reflexive Banach space and let A be maximal monotone in $X \times X^*$. Then both $\overline{D(A)}$ and $\overline{R(A)}$ are convex.*

Proof. Without any loss of generality, we may assume that X and X^* are strictly convex. Then, as seen in Proposition 2.1, $J_\lambda x \to x$ for every $x \in \overline{\text{conv}\, D(A)}$. Because $J_\lambda x \in D(A)$ for all $\lambda > 0$ and $x \in X$, we conclude that $\overline{\text{conv}\, D(A)} = \overline{D(A)}$, as claimed. Because $R(A) = D(A^{-1})$ and A^{-1} is maximal monotone in $X^* \times X$, we conclude that $\overline{R(A)}$ is also convex. □

We now establish an important property of monotone operators with nonempty interior domain.

Theorem 2.5. *Let A be a monotone subset of $X \times X^*$. Then A is locally bounded at any interior point of $D(A)$.*

Following an idea due to Fitzpatrick [13], we first prove the following technical lemma.

Lemma 2.4. *Let $\{x_n\} \subset X$ and $\{y_n\} \subset X^*$ be such that $x_n \to 0$ and $\|y_n\| \to \infty$ as $n \to \infty$. Let $B(0,r)$ be the closed ball $\{x;\ \|x\| \le r\}$. Then there exist $x^0 \in B(0,r)$ and $\{x_{n_k}\} \subset \{x_n\}$, $\{y_{n_k}\} \subset \{y_n\}$ such that*

$$\lim_{k\to\infty}(x_{n_k} - x^0, y_{n_k}) = -\infty. \tag{2.31}$$

Proof. Suppose that the lemma is false. Then there exists $r > 0$ such that for every $u \in B(0,r)$ there exists $C_u > -\infty$ such that

$$(x_n - u, y_n) \ge C_u, \qquad \forall n \in \mathbf{N}.$$

We may write $B(0,r) = \cup_k \{u \in B(0,r);\ (x_n - u, y_n) \ge -k,\ \forall n\}$. Then, by the Hausdorff–Baire theorem we infer that there is k_0 such that

$$\operatorname{int}\{u \in B(0,r);\ (x_n - u, y_n) > -k_0,\ \forall n\} \ne \emptyset.$$

In other words, there are $\varepsilon > 0$, $k_0 \in \mathbf{N}$, and $u_0 \in B(0,r)$ such that

$$\{u;\ \|u - u_0\| \le \varepsilon\} \subset \{u \in B(0,r);\ (x_n - u, y_n) > -k_0,\ \forall n\}.$$

Now, we have

$$(x_n - u, y_n) \ge -k_0 \quad \text{and} \quad (x_n - u_0, y_n) \ge C_{u_0}.$$

Summing up, we get

$$(2x_n + u_0 - u, y_n) \ge -k_0 + C, \qquad \forall u \in B(u_0, \varepsilon),$$

where $C = C_{u_0}$. Now, we take $u = u_0 + 2x_n + w$, where $\|w\| = \varepsilon/2$. For n sufficiently large, we therefore have

$$(w, y_n) \le -C + y_0, \qquad \forall w, \|w\| = \frac{\varepsilon}{2},$$

which clearly contradicts the fact that $\|y_n\| \to \infty$ as $n \to \infty$. $\square$

Proof of Theorem 2.5. The method of proof is due to Brezis, Crandall and Pazy [7]. Let $x_0 \in \operatorname{int} D(A)$ be arbitrary. Without loss of generality, we may assume that $x_0 = 0$. (This can be achieved by shifting the domain of A.) Let us assume that A is not locally bounded at 0. Then there exist sequences $\{x_n\} \subset X$, $\{y_n\} \subset X^*$ such that $[x_n, y_n] \in A$, $\|x_n\| \to 0$, and $\|y_n\| \to \infty$. According to Lemma 2.4, for every ball $B(0,r)$, there exists $x^0 \in B(0,r)$ and $\{x_{n_k}\} \subset \{x_n\}$, $\{y_{n_k}\} \subset \{y_n\}$ such that

$$\lim_{k\to\infty}(x_{n_k} - x^0, y_{n_k}) = -\infty.$$

Let r be sufficiently small so that $B(0,r) \subset D(A)$. Then, $x^0 \in D(A)$ and by the monotonicity of A it follows that

$$(x_{n_k} - x^0, y) \to -\infty \quad \text{as } k \to \infty,$$

for some $y \in Ax^0$. The contradiction we have arrived at completes the proof. □

In particular, Theorem 2.5 implies that every monotone operator A everywhere defined on X is locally bounded.

Now we are ready to prove the main result of this section, due to Rockafellar [24].

Theorem 2.6. *Let X be a reflexive Banach space and let A and B be maximal monotone subsets of $X \times X^*$ such that*

$$(\operatorname{int} D(A)) \cap D(B) \neq \emptyset. \tag{2.32}$$

Then $A + B$ is maximal monotone in $X \times X^*$.

Proof. As in the previous cases, we may assume without loss of generality that X and X^* are strictly convex. Moreover, shifting the domains and ranges of A and B, if necessary, we may assume that $0 \in (\operatorname{int} D(A)) \cap D(B)$, $0 \in A0$, $0 \in B0$. We prove that $R(J + A + B) = X^*$. To this aim, consider an arbitrary element y in X^*. Because the operator B_λ is demicontinuous, bounded, and monotone, and so is $J : X \to X^*$, by Corollaries 2.1 and 2.2, it follows that, for every $\lambda > 0$, the equation

$$Jx_\lambda + Ax_\lambda + B_\lambda x_\lambda \ni y \tag{2.33}$$

has a solution $x_\lambda \in D(A)$. (J and J^{-1} are single-valued and X, X^* are strictly convex, thus it follows by standard arguments involving the monotonicity of A and B that x_λ is unique.) Multiplying equation (2.33) by x_λ and using the obvious inequalities

$$(x_\lambda, Ax_\lambda) \geq 0, \qquad (x_\lambda, B_\lambda x_\lambda) \geq 0,$$

we infer that

$$\|x_\lambda\| \leq \|y\|, \qquad \forall \lambda > 0.$$

Moreover, because $0 \in \operatorname{int} D(A)$, it follows by Theorem 2.5 that there exist constants $\rho > 0$ and $M > 0$ such that

$$\|x^*\| \leq M, \ \forall x^* \in Ax, \qquad \|x\| \leq \rho. \tag{2.34}$$

Multiplying equation (2.33) by $x_\lambda - \rho w$ and using the monotonicity of A, we get

$$(x_\lambda - \rho w, Jx_\lambda + B_\lambda x_\lambda - y) + (x_\lambda - \rho w, A(\rho w)) \le 0, \qquad \forall \|w\| = 1.$$

By (2.34), we get

$$\|x_\lambda\|^2 - \rho(w, B_\lambda x_\lambda) \le M(\rho + \|x_\lambda\|) + \|x_\lambda\|(\rho + \|y\|).$$

Hence,

$$\|x_\lambda\|^2 + \rho\|B_\lambda x_\lambda\| \le \|x_\lambda\|(\rho + M + \|y\|) + M\rho, \qquad \forall \lambda > 0.$$

We may, therefore, conclude that $\{B_\lambda x_\lambda\}$ and $\{y_\lambda = y - Jx_\lambda - B_\lambda x_\lambda\}$ are bounded in X^* as $\lambda \to 0$. Inasmuch as X is reflexive, we may assume that on a subsequence, again denoted λ,

$$x_\lambda \rightharpoonup x_0, \qquad B_\lambda x_\lambda \rightharpoonup y_1, \qquad y_\lambda \in Ax_\lambda \rightharpoonup y_2, \qquad Jx_\lambda \rightharpoonup y_0.$$

Inasmuch as $A + J$ is monotone, we have

$$(x_\lambda - x_\mu, B_\lambda x_\lambda - B_\mu x_\mu) \le 0, \qquad \forall \lambda, \mu > 0.$$

Then, by Proposition 2.2(iv), we have

$$\lim_{\lambda,\mu \to 0} (x_\lambda - x_\mu, B_\lambda x_\lambda - B_\mu x_\mu) = 0$$

and $[x_0, y_1] \in B$. Then, by equation (2.33), we see that

$$\lim_{\lambda,\mu \to 0} (x_\lambda - x_\mu, Jx_\lambda + y_\lambda - Jx_\lambda - y_\mu) = 0, \qquad y_\lambda \in Ax_\lambda,\ y_\mu \in Ax_\mu,$$

and, because $A + J$ is maximal monotone, it follows by Lemma 2.3 (see Corollary 2.5) that $[x_0, y_0 + y_2] \in A + J$. Thus, letting λ tend to zero in (2.33), we see that

$$y \in J(x_0) + Ax_0 + Bx_0,$$

thereby completing the proof. □

In particular, Theorems 2.4 and 2.6 lead to the following.

Corollary 2.6. *Let X be a reflexive Banach space, $A \subset X \times X^*$ a maximal monotone operator, and let $B : X \to X^*$ be a demicontinuous monotone operator. Then $A + B$ is maximal monotone.*

More generally, it follows from Theorem 2.6 that if A, B are two maximal monotone sets of $X \times X^*$, and $D(B) = X$, then $A + B$ is maximal monotone.

We conclude this section with a result of the same type in Hilbert spaces.

Theorem 2.7. *Let $X = H$ be a Hilbert space identified with its own dual and let A, B be maximal monotone sets in $H \times H$ such that $D(A) \cap D(B) \neq \emptyset$ and*

$$(v, A_\lambda u) \ge -C(\|u\|^2 + \lambda\|A_\lambda u\|^2 + \|A_\lambda u\| + 1), \qquad \forall [u, v] \in B. \tag{2.35}$$

Then $A+B$ is maximal monotone.

Proof. We have denoted by $A_\lambda = \lambda^{-1}(I-(I+\lambda A)^{-1})$ the Yosida approximation of A. For any $y \in H$ and $\lambda > 0$, consider the equation

$$x_\lambda + Bx_\lambda + A_\lambda x_\lambda \ni y, \tag{2.36}$$

which, by Corollaries 2.5 and 2.6 has a solution (clearly unique) $x_\lambda \in D(B)$. Let $x_0 \in D(A) \cap D(B)$. Taking the scalar product of (2.36) with $x_\lambda - x_0$ and using the monotonicity of B and A_λ yields

$$(x_\lambda, x_\lambda - x_0) + (y_0, x_\lambda - x_0) + (A_\lambda x_0, x_\lambda - x_0) \le (y, x_\lambda - x_0).$$

Because, as seen in Proposition 2.2,

$$\|A_\lambda x_0\| \le |Ax_0|, \qquad \forall \lambda > 0,$$

this yields

$$\|x_\lambda\| \le M, \qquad \forall \lambda > 0.$$

Next, we multiply equation (2.36) by $A_\lambda x_\lambda$ and use inequality (2.35) to get, after some calculations,

$$\|A_\lambda x_\lambda\| \le C, \qquad \forall \lambda > 0.$$

Now, for a sequence $\lambda_n \to 0$, we have

$$x_{\lambda_n} \rightharpoonup x, \qquad A_{\lambda_n} x_{\lambda_n} \rightharpoonup y_1, \qquad y_{\lambda_n} \rightharpoonup y_2,$$

where $y_\lambda = y - x_\lambda - A_\lambda x_\lambda \in Bx_\lambda$.

Then, arguing as in the proof of Theorem 2.6, it follows by Proposition 2.2 that $[x, y_1] \in A$, $[x, y_2] \in B$, and this implies that $y \in x + Ax + Bx$, as claimed. □

Proposition 2.4. *Let X be the Euclidean space $\mathbf{R}^N$ and $A : \mathbf{R}^N \to \mathbf{R}^N$ be a monotone, everywhere defined, and upper-semicontinuous operator (multivalued) such that the set Ax is convex for each $x \in \mathbf{R}^N$. Then A is maximal monotone in $\mathbf{R}^N \times \mathbf{R}^N$.*

Proof. We recall that A is said to be upper-semicontinuous if its graph is closed in $\mathbf{R}^N \times \mathbf{R}^N$. One must prove that there is $\lambda > 0$ such that for each $f \in \mathbf{R}^N$ equation $\lambda x + Ax \ni f$ has solution. We rewrite this equation as

$$x \in \frac{1}{\lambda} f - \frac{1}{\lambda} Ax$$

and apply the Kakutani fixed point theorem to operator $x \xrightarrow{T} (1/\lambda)f - (1/\lambda)Ax$ on the closed ball $K_R = \{x \in \mathbf{R}^N;\ \|x\| \le R\}$. By Theorem 2.5 we know that $A(K_R)$ is bounded for each $R > 0$. Then, choosing λ sufficiently large, it follows that $T(K_R) \subset K_R$ and so T has a fixed point in K_R, as claimed. □

Consider now a monotone measurable function $\psi : \mathbf{R}^N \to \mathbf{R}^N$; that is,

$$(\psi(x)-\psi(y),x-y)_N \geq 0, \qquad \forall x,y \in \mathbf{R}^N.$$

(Here $(\cdot,\cdot)_N$ is the Euclidean scalar product.) We associate with ψ the following multivalued graph (the Filipov mapping)

$$\widetilde{\psi}(x)=\bigcap_{\delta>0}\ \bigcap_{m(E)=0} \overline{\text{conv}\ \psi\ (B_\delta(x)\setminus E)}, \qquad \forall x \in \mathbf{R}^N,$$

where $B_\delta(x)=\{y\in \mathbf{R}^N;\ \|y-x\|_N \leq \delta\}$ and $m(E)$ is the Lebesgue measure of the subset $E \subset \mathbf{R}^N$. In the special case where $N=1$, the Filipov mapping $\widetilde{\psi}$ is obtained by "filling the jumps" of ψ in discontinuity points; that is,

$$\widetilde{\psi}(x)=[\psi(x-0),\psi(x+0)], \qquad \forall x\in \mathbf{R}.$$

Proposition 2.5. *The operator $\widetilde{\psi}$ is maximal monotone in $\mathbf{R}^N \times \mathbf{R}^N$.*

Proof. The monotonicity of $\widetilde{\psi}$ follows immediately from that of ψ. It is also easily seen that $\widetilde{\psi}$ is upper semicontinuous and has convex values. Then the conclusion follows by Proposition 2.4. □

Monotone Operators in Complex Banach Spaces

Let $\widetilde{X}$ be a complex Banach space and let $\widetilde{X}^*$ be its dual.

A monotone subset $A\subset \widetilde{X}\times\widetilde{X}^*$ is called *monotone* if

$$\text{Re}(x-y,x^*-y^*)\geq 0 \quad \text{for all } [x,x^*],[y,y^*]\in A.$$

If we represent $\widetilde{X}$ as $X+iX$, where X is a real Banach space and $A_1,A_2\subset X\times X^*$ are defined by

$$A_1(x,\widetilde{x})+iA_2(x,\widetilde{x})=A(x+i\widetilde{x}), \qquad \forall x,\widetilde{x}\in X,$$

then the monotonicity condition reduces to

$$(x-y,A_1(x,\widetilde{x})-A_1(y,\widetilde{y}))+(\widetilde{x}-\widetilde{y},A_2(x,\widetilde{x})-A_2(y,\widetilde{y}))\geq 0.$$

Define the operator $\mathscr{A}: X\times X\to X^*\times X^*$ by

$$\mathscr{A}(x,\widetilde{x})=\{A_1(x,\widetilde{x}),A_2(x,\widetilde{x})\};$$

that is, $A_1=\text{Re}A$, $A_2=\text{Im}A$. Then A is monotone in $\widetilde{X}\times\widetilde{X}^*$ if and only if $\mathscr{A}$ is monotone in $(X\times X)\times(X^*\times X^*)$. Similarly, A is maximal monotone (i.e., it is maximal in the class of monotone operators) if and only if $\mathscr{A}$ is maximal monotone.

In this way, the whole theory of maximal monotone operators in real Banach spaces extends mutatis–mutandis to maximal monotone operators in complex Banach spaces.

2.2 Maximal Monotone Subpotential Operators

The subdifferential of a lower semicontinuous convex function is an important example of maximal monotone operator that closes the bridge between the theory of nonlinear maximal monotone operators and convex analysis. Such an operator is also called a *subpotential maximal monotone operator*.

Theorem 2.8. *Let X be a real Banach space and let $\varphi : X \to \overline{\mathbf{R}}$ be an l.s.c. proper convex function. Then $\partial\varphi$ is a maximal monotone subset of $X \times X^*$.*

Proof. It is readily seen that $\partial\varphi$ is monotone in $X \times X^*$. To prove that $\partial\varphi$ is maximal monotone, we assume for simplicity that X is reflexive and refer the reader to Rockafellar's work [26] for the proof in the general case.

Continuing, we fix $y \in X^*$ and consider the equation

$$Jx + \partial\varphi(x) \ni y. \tag{2.37}$$

Let $f : X \to \overline{\mathbf{R}}$ be the convex, l.s.c. function defined by

$$f(x) = \frac{1}{2}\,\|x\|^2 + \varphi(x) - (x,y).$$

By Proposition 1.1, we see that

$$\lim_{\|x\|\to\infty} f(x) = +\infty,$$

and so, by Proposition 1.4, we conclude that there exists $x_0 \in X$ such that

$$f(x_0) = \inf\{f(x);\ x \in X\}.$$

This yields

$$\frac{1}{2}\,\|x_0\|^2 + \varphi(x_0) - (x_0,y) \le \frac{1}{2}\,\|x\|^2 + \varphi(x) - (x,y), \qquad \forall x \in X;$$

that is,

$$\begin{aligned}\varphi(x_0) - \varphi(x) &\le (x_0 - x, y) + \tfrac{1}{2}\,(\|x\|^2 - \|x_0\|^2)\\ &\le (x_0 - x, y) + (x - x_0, Jx), \qquad \forall x \in X.\end{aligned}$$

In the latter inequality we take $x = tx_0 + (1-t)u$, $0 < t < 1$, where u is an arbitrary element of X. We get

$$\varphi(x_0) - \varphi(u) \le (x_0 - u, y) + (u - x_0, w_t),$$

where $w_t \in J(tx_0 + (1-t)u)$.

For $t \to 1$, $w_t \rightharpoonup w \in J(x_0)$ because, as seen earlier, J is strongly–weakly closed in $X \times X^*$. Hence,

$$\varphi(x_0) - \varphi(u) \le (x_0 - u, y - w), \qquad \forall u \in X,$$

and this inequality shows that $y - w \in \partial\varphi(x_0)$; that is, x_0 is a solution to equation (2.37). We have therefore proven that $R(J + \partial\varphi) = X^*$. □

In particular, this result leads to a simple proof of Proposition 1.6: *if* $\varphi : X \to \overline{\mathbf{R}}$ *is an l.s.c., convex, and proper function, then* $D(\partial\varphi)$ *is a dense subset of* $D(\varphi)$.

Proof. Let x be any element of $D(\varphi)$ and let $x_\lambda = J_\lambda x$ be the solution to the equation (see (2.25))

$$J(x_\lambda - x) + \lambda\partial\varphi(x_\lambda) \ni 0.$$

Multiplying this equation by $x_\lambda - x$, we get

$$\|x_\lambda - x\|^2 + \lambda(\varphi(x_\lambda) - \varphi(x)) \leq 0, \qquad \forall \lambda > 0.$$

Because, by Proposition 1.1, φ is bounded from below by an affine function and $\varphi(x) < \infty$, this yields

$$\lim_{\lambda \to 0} x_\lambda = x.$$

As $x_\lambda \in D(\partial\varphi)$ and x is arbitrary in $D(\varphi)$, we conclude that

$$\overline{D(\varphi)} = \overline{D(\partial\varphi)},$$

as claimed. □

For every $\lambda > 0$, define the function

$$\varphi_\lambda(x) = \inf\left\{\frac{\|x-u\|^2}{2\lambda} + \varphi(u);\ u \in X\right\}, \qquad \forall x \in X, \tag{2.38}$$

where $\varphi : X \to \overline{\mathbf{R}}$ is an l.s.c. proper convex function. By Propositions 1.1 and 1.4 it follows that $\varphi_\lambda(x)$ is well defined for all $x \in X$ and the infimum defining it is attained (if the space X is reflexive). This implies by a straightforward argument that φ_λ is convex and l.s.c. on X. (Because φ_λ is everywhere defined, we conclude by Proposition 1.2, that φ_λ is continuous.)

The function φ_λ is called the *Moreau regularization* of φ (see [21]), for reasons that become clear in the following theorem.

Theorem 2.9. *Let X be a reflexive and strictly convex Banach space with strictly convex dual. Let $\varphi : X \to \overline{\mathbf{R}}$ be an l.s.c. convex, proper function and let $A = \partial\varphi \subset X \times X^*$. Then the function φ_λ is convex, continuous, Gâteaux differentiable, and $\nabla\varphi_\lambda = A_\lambda$ for all $\lambda > 0$. Moreover:*

$$\varphi_\lambda(x) = \frac{\|x - J_\lambda x\|^2}{2\lambda} + \varphi(J_\lambda x), \qquad \forall \lambda > 0,\ x \in X; \tag{2.39}$$

$$\lim_{\lambda \to 0} \varphi_\lambda(x) = \varphi(x), \qquad \forall x \in X; \tag{2.40}$$

$$\varphi(J_\lambda x) \leq \varphi_\lambda(x) \leq \varphi(x), \qquad \forall \lambda > 0,\ x \in X. \tag{2.41}$$

If X is a Hilbert space (not necessarily identified with its dual), then φ_λ is Fréchet differentiable on X.

Proof. We observe that the subdifferential of the function

$$u \to \frac{\|x-u\|^2}{2\lambda} + \varphi(u)$$

is just the operator $u \to \lambda^{-1}J(u-x) + \partial\varphi(u)$ (see Theorem 2.10 below). This implies that every solution x_λ of the equation

$$\lambda^{-1}J(u-x) + \partial\varphi(u) \ni 0$$

is a minimum point of the function

$$u \to \frac{1}{2\lambda}\|x-u\|^2 + \varphi(u).$$

Recalling that $x_\lambda = J_\lambda x$, we obtain (2.39). Regarding inequality (2.41), it is an immediate consequence of (2.38). To prove (2.40), assume first that $x \in D(\varphi)$. Then, as seen in Proposition 2.3, $\lim_{\lambda\to 0} J_\lambda x = x$, and by (2.41) and the lower semicontinuity of φ, we infer that

$$\varphi(x) \le \liminf_{\lambda\to 0} \varphi(J_\lambda x) \le \liminf_{\lambda\to 0} \varphi_\lambda(x) \le \varphi(x).$$

If $x \notin D(\varphi)$ (i.e., $\varphi(x) = +\infty$), then $\lim_{\lambda\to 0} \varphi_\lambda(x) = +\infty$ because otherwise there would exist $\{\lambda_n\} \to 0$ and $C > 0$ such that

$$\varphi_{\lambda_n}(x) \le C, \qquad \forall n.$$

Then, by (2.39), we see that $\lim_{n\to\infty} J_{\lambda_n} x = x$, and again by (2.41) and the lower semicontinuity of φ, we conclude that $\varphi(x) \le C$, which is absurd.

To conclude the proof, it remains to show that φ_λ is Gâteaux differentiable and $\nabla\varphi_\lambda = A_\lambda$. By (2.39), it follows that

$$\begin{aligned}\varphi_\lambda(y) - \varphi_\lambda(x) &\le (J_\lambda(y) - J_\lambda(x), A_\lambda y) + \frac{1}{2\lambda}(\|y-J_\lambda(y)\|^2 - \|x-J_\lambda(x)\|^2)\\ &= (y-x, A_\lambda y) + (J_\lambda(y) - y, A_\lambda y) + (x - J_\lambda(x), A_\lambda y)\\ &\quad + \frac{1}{2\lambda}(\|y-J_\lambda(y)\|^2 - \|x-J_\lambda(x)\|^2) \le (y-x, A_\lambda y).\end{aligned}$$

Hence,

$$\varphi_\lambda(y) - \varphi_\lambda(x) - (y-x, A_\lambda x) \le (y-x, A_\lambda y - A_\lambda x) \tag{2.42}$$

for all $x, y \in X$ and $\lambda > 0$. The latter inequality clearly implies that

$$\lim_{t\downarrow 0} \frac{\varphi_\lambda(x+tu) - \varphi_\lambda(x)}{t} \le (u, A_\lambda x), \qquad \forall u, x \in X,$$

because, as seen earlier, A_λ is demicontinuous. Hence, φ_λ is Gâteaux differentiable and $\nabla\varphi_\lambda = (\partial\varphi)_\lambda = A_\lambda$.

Now, assume that X is a Hilbert space. Then, as seen earlier in Proposition 2.3, $A_\lambda : X \to X$ is Lipschitz continuous with the Lipschitz constant not greater than $2/\lambda$. Then, by inequality (2.42), we see that

$$|\varphi_\lambda(x) - \varphi_\lambda(y) - (x-y, A_\lambda x)| \le \frac{2}{\lambda}\,\|x-y\|^2, \qquad \forall x,y \in X,$$

and this shows that φ_λ is Fréchet differentiable. □

Let us consider the particular case where $\varphi = I_K$ (see (1.17)), K is a closed convex subset of X, and X is a Hilbert space Then

$$(I_K)_\lambda(x) = \frac{\|x - P_K x\|^2}{2\lambda}, \qquad \forall x \in X,\ \lambda > 0, \tag{2.43}$$

where $P_K x$ is the projection of x on K. (Because K is closed and convex, $P_K x$ is uniquely defined.) Moreover, as previously seen, we have

$$P_K = J_\lambda = (I + \lambda A)^{-1}, \qquad \forall \lambda > 0. \tag{2.44}$$

It should be said that (2.38) is a convenient way to regularize the convex l.s.c. functions φ in infinite dimensions and, in particular, in Hilbert spaces, the main advantage being that the regularization φ_λ remains convex and is C^1 with Lipschitz differential $\nabla\varphi_\lambda$.

A problem of great interest in convex optimization as well as for calculus with convex functions is to determine whether given two l.s.c., convex, proper functions f and g on X, $\partial(f+g) = \partial f + \partial g$. The following theorem due to Rockafellar [25] gives a general answer to this question.

Theorem 2.10. *Let X be a Banach space and let $f : X \to \overline{\mathbf{R}}$ and $g : X \to \overline{\mathbf{R}}$ be two l.s.c., convex, proper functions such that $D(f) \cap \operatorname{int} D(g) \ne \emptyset$. Then*

$$\partial(f+g) = \partial f + \partial g. \tag{2.45}$$

Proof. If the space X is reflexive, (2.45) is an immediate consequence of Theorem 2.6. Indeed, as seen in Proposition 1.7, $\operatorname{int} D(\partial g) = \operatorname{int} D(g)$ and so $D(\partial f) \cap \operatorname{int} D(\partial g) \ne \emptyset$. Then, by Theorem 2.6, $\partial f + \partial g$ is maximal monotone in $X \times X^*$. On the other hand, it is readily seen that $\partial f + \partial g \subset \partial(f+g)$. Hence, $\partial f + \partial g = \partial(f+g)$.

In the general case, Theorem 2.10 follows by a separation argument we present subsequently.

Because the relation $\partial f + \partial g \subset \partial(f+g)$ is obvious, let us prove that

$$\partial(f+g) \subset \partial f + \partial g.$$

To this end, consider $x_0 \in D(\partial f) \cap D(\partial g)$ and $w \in \partial(f+g)(x_0)$, arbitrary but fixed. We prove that $w = w_1 + w_2$, where $w_1 \in \partial f(x_0)$ and $w_1 \in \partial g(x_0)$. Replacing the functions f and g by $x \to f(x+x_0) - f(x_0) - (x, z_1)$ and $x \to g(x+x_0) - g(x_0) - (x, z_2)$, respectively, where $w = z_1 + z_2$, we may assume that $x_0 = 0$, $w = 0$, and $f(0) = g(0) = 0$. Hence, we should prove that $0 \in \partial f(0) + \partial g(0)$. Consider the sets E_i, $i = 1, 2$, defined by

$$\begin{aligned} E_1 &= \{(x, \lambda) \in X \times \mathbf{R};\ f(x) \le \lambda\}, \\ E_2 &= \{(x, \lambda) \in X \times \mathbf{R};\ g(x) \le -\lambda\}. \end{aligned}$$

Inasmuch as $0 \in \partial(f+g)(0)$, we have

$$0 = (f+g)(0) = \inf\{(f+g)(x);\ x \in X\},$$

and therefore $E_1 \cap \operatorname{int} E_2 = \emptyset$. Then, by the separation theorem there exists a closed hyperplane that separates the sets E_1 and E_2. In other words, there are $w \in X^*$ and $\alpha \in \mathbf{R}$ such that

$$\begin{aligned} (w, x) + \alpha\lambda &\le 0, \qquad \forall (x, \lambda) \in E_1, \\ (w, x) + \alpha\lambda &\ge 0, \qquad \forall (x, \lambda) \in E_2. \end{aligned} \tag{2.46}$$

Let us observe that the hyperplane is not vertical; that is, $\alpha \ne 0$. Indeed, if $\alpha = 0$, then this would imply that the hyperplane $(w, x) = 0$ separates the sets $D(f)$ and $D(g)$ in the space X, which is not possible because $D(f) \cap \operatorname{int} D(g) \ne \emptyset$. Hence, $\alpha \ne 0$, and to be more specific we assume that $\alpha > 0$. Then, by (2.46), we see that

$$g(x) \le -\lambda \le (w, x) \le -\alpha f(x), \qquad \forall x \in X,$$

and, therefore, $(1/\alpha) w \in \partial f(0)$, $-(1/\alpha) w \in \partial g(0)$ (i.e., $0 \in \partial f(0) + \partial g(0)$), as claimed. □

Theorem 2.11. *Let $X = H$ be a real Hilbert space (identified with its own dual) and let A be a maximal monotone subset of $H \times H$. Let $\varphi : H \to \overline{\mathbf{R}}$ be an l.s.c., convex, proper function such that $D(A) \cap D(\partial g) \ne \emptyset$ and, for some $h \in H$,*

$$\varphi((I + \lambda A)^{-1}(x + \lambda h)) \le \varphi(x) + C\lambda(1 + \varphi(x)), \qquad \forall x \in D(\varphi),\ \lambda > 0. \tag{2.47}$$

Then $A + \partial\varphi$ is maximal monotone and $\overline{D(A + \partial\varphi)} = \overline{D(A)} \cap \overline{D(\varphi)}$.

Proof. We proceed as in the proof of Theorem 2.7. Let y be arbitrary but fixed in H. Then, for every $\lambda > 0$, the equation

$$x_\lambda + A_\lambda x_\lambda + \partial\varphi(x_\lambda) \ni y$$

has a unique solution $x_\lambda \in D(\partial\varphi)$. We multiply the preceding equation by $x - J_\lambda(x_\lambda + \lambda h)$ and use condition (2.47). This yields

$$\|A_\lambda x_\lambda\|^2 + (A_\lambda x_\lambda, J_\lambda(x_\lambda) - J_\lambda(x_\lambda + \lambda h)) \le C\lambda(\|y\| + \|h\| + \|x_\lambda\| + \varphi(x_\lambda) + 1),$$

where $J_\lambda = (I + \lambda A)^{-1}$. We get

$$\|A_\lambda x_\lambda\|^2 \le C(\|y\| + \|h\| + \|x_\lambda\| + \varphi(x_\lambda) + 1).$$

On the other hand, multiplying the latter equation by $x_\lambda - x_0$, where $x_0 \in D(A) \cap D(\partial\varphi)$, we get

$$\|x_\lambda\|^2 + \varphi(x_\lambda) \le C(\|A_\lambda x_0\|^2 + \varphi(x_0) + 1).$$

Hence, $\{A_\lambda x_\lambda\}$ and $\{x_\lambda\}$ are bounded in H. Then, as seen in the proofs of Theorems 2.6 and 2.7, this implies that $x_\lambda \rightharpoonup x$, where x is the solution to the equation

$$x + \partial\varphi(x) + Ax \ni y.$$

Now, let us prove that

$$\overline{D(A)} \cap \overline{D(\varphi)} \subset \overline{D(A) \cap D(\varphi)} \subset \overline{D(A) \cap D(\partial\varphi)}.$$

Let $u \in \overline{D(A)} \cap \overline{D(\varphi)}$ be arbitrary but fixed and let h be as in condition (2.47). Clearly, there is a sequence $\{u_\lambda\} \subset D(\varphi)$ such that $u_\lambda + \lambda h \in D(\varphi)$ and $u_\lambda \to u$ as $\lambda \to 0$. Let $v_\lambda = J_\lambda(u_\lambda + \lambda h) \in D(A) \cap D(\varphi)$ (by condition (2.47)). We have

$$\|v_\lambda - u\| \le \|J_\lambda(u_\lambda + \lambda h) - J_\lambda u\| + \|u - J_\lambda u\| \to 0 \quad \text{as } \lambda \to 0,$$

because $u \in \overline{D(A)}$ (see Proposition 2.2). Hence,

$$\overline{D(A)} \cap \overline{D(\varphi)} \subset \overline{D(A) \cap D(\varphi)}.$$

Now, let u be arbitrary in $D(A) \cap D(\varphi)$ and let $x_\lambda \in D(A) \cap D(\partial\varphi)$ be the solution to

$$x_\lambda + \lambda(Ax_\lambda + \partial\varphi(x_\lambda)) \ni u.$$

By the definition of $\partial\varphi$, we have

$$\lambda(\varphi(x_\lambda) - \varphi(u)) \le (u - x_\lambda - \lambda Ax_\lambda, x_\lambda - u) \le -\|u - x_\lambda\|^2 + \lambda\|A^0 u\|\ \|u - x_\lambda\|, \quad \forall \lambda > 0.$$

Hence, $x_\lambda \to u$ for $\lambda \to 0$, and so $D(A) \cap D(\varphi) \subset \overline{D(A) \cap D(\partial\varphi)}$, as claimed. □

Remark 2.1. In particular, condition (2.45) holds if

$$(A_\lambda(x + \lambda h), y) \ge -C(1 + \varphi(x)), \qquad \forall \lambda > 0,$$

for some $h \in H$, and all $[x, y] \in \partial\varphi$.

In fact, condition (2.47) can be seen as an abstract substitute for the maximum principle because in some specific situations (for instance, if A is an elliptic operator) it can be checked via maximum principle arguments.

We conclude this section with an explicit formula for $\partial\varphi$ in term of the directional derivative, φ'.

Proposition 2.6. *Let X be a Banach space and let $\varphi : X \to \overline{\mathbf{R}}$ be an l.s.c., convex, proper function on X. Then, for all $x_0 \in D(\partial\varphi)$,*

$$\partial\varphi(x_0) = \{x_0^* \in X^*;\ \varphi'(x_0,u) \geq (u,x_0^*), \qquad \forall u \in X\}. \tag{2.48}$$

Proof. Let $x_0^* \in \partial\varphi(x_0)$. Then, by the definition of $\partial\varphi$,

$$\varphi(x_0) - \varphi(x_0 + tu) \leq -t(u,x_0^*), \qquad \forall u \in X,\ t > 0,$$

which yields

$$\varphi'(x_0,u) \geq (u,x_0^*), \qquad \forall u \in X.$$

Assume now that $(u,x_0^*) \leq \varphi'(x_0,u),\ \forall u \in X$. Because φ is convex, the function $t \to (\varphi(x_0+tu) - \varphi(x_0)/t)$ is monotonically increasing and so we have

$$(u,x_0^*) \leq t^{-1}(\varphi(x_0+tu) - \varphi(x_0)), \qquad \forall u \in X,\ t > 0.$$

Hence $x_0^* \in \partial\varphi(x_0)$, and the proof is complete. □

Formula (2.48) can be taken as an equivalent definition of the subdifferential $\partial\varphi$, and it may be used to define the generalized gradients of nonconvex functions.

It turns out that, if φ is continuous at x, then

$$\varphi'(x_0,u) = \sup\{(u,x_0^*);\ x_0^* \in \partial\varphi(x_0)\}, \qquad u \in X. \tag{2.49}$$

Examples of Subpotential Operators

There is a general characterization of maximal monotone operators that are subdifferential of l.s.c. convex functions due to Rockafellar [23]. A set $A \subset X \times X^*$ is said to be *cyclically monotone* if

$$(x_0 - x_1, x_0^*) + \cdots + (x_{n-1} - x_n, x_{n-1}^*) + (x_n - x_0, x_n^*) \geq 0, \tag{2.50}$$

for all $[x_i,x_i^*] \in A,\ i = 0,1,\dots,n$. A is said to be *maximal cyclically monotone* if it is cyclically monotone and has no cyclically monotone extensions in $X \times X^*$. It turns out that the class of subdifferential mappings coincides with that of maximal cyclically monotone operators. More precisely, one has the following.

Theorem 2.12. *Let X be a real Banach space and let $A \subset X \times X^*$. The set A is the subdifferential of an l.s.c., convex, proper function from X to $\mathbf{R}$ if and only if A is maximal cyclically monotone.*

We leave to the reader the proof of this theorem and we concentrate on some significant examples of subdifferential mappings.

1. *Maximal monotone sets (graphs) in* $\mathbf{R} \times \mathbf{R}$. Every maximal monotone set (graph) of $\mathbf{R} \times \mathbf{R}$ is the subdifferential of an l.s.c., convex, proper function on $\mathbf{R}$.

Indeed, let β be a maximal monotone set in $\mathbf{R}\times\mathbf{R}$ and let $\beta^0:\mathbf{R}\to\mathbf{R}$ be the function defined by

$$\beta^0(r)=\{y\in\beta(r);\ |y|=\inf\{|z|;\ z\in\beta(r)\}\},\qquad \forall r\in\mathbf{R}.$$

We know that $\overline{D(\beta)}=[a,b]$, where $-\infty\le a\le b\le\infty$. The function β^0 is monotonically increasing and so the integral

$$j(r)=\int_{r_0}^{r}\beta^0(u)du,\qquad \forall r\in\mathbf{R}, \tag{2.51}$$

where $r_0\in D(\beta)$, is well defined (unambiguously a real number or $+\infty$). Clearly, the function j is continuous on (a,b) and convex on $\mathbf{R}$. Moreover,

$$\liminf_{r\to b} j(r)\ge j(b)\quad\text{and}\quad \liminf_{r\to a} j(r)\ge j(a).$$

Finally,

$$j(r)-j(t)=\int_t^r\beta^0(u)du\le v(r-t),\qquad \forall[r,v]\in\beta,\ t\in\mathbf{R}.$$

Hence $\beta=\partial j$, where j is the l.s.c. convex function defined by (2.51).

It is easily seen that if $\beta:\mathbf{R}\to\mathbf{R}$ is a continuous and monotonically increasing function, then β is a maximal monotone graph in $\mathbf{R}\times\mathbf{R}$ in the sense of general definition; that is, the range of $u\to u+\beta(u)$ is all of $\mathbf{R}$. (By a monotonically increasing function we mean, here and everywhere in the following, a monotonically nondecreasing function.) If β is a monotonically increasing function discontinuous in $\{r_j\}_{j=1}^{\infty}$, then as seen earlier one gets from β a maximal monotone graph $\widetilde{\beta}\subset\mathbf{R}\times\mathbf{R}$ by "filling" the jumps of β in r_j; that is,

$$\widetilde{\beta}(r)=\begin{cases}\beta(r), & \text{for } r\ne r_j,\\ [\beta(r_j-0),\beta(r_j+0)], & \text{for } r=r_j.\end{cases}$$

(See Proposition 2.4.)

2. *Self-adjoint operators.* Let H be a real Hilbert space (identified with its own dual) with scalar product $(\cdot,\cdot)$ and norm $|\cdot|$, and let A be a linear self-adjoint positive operator on H. Then, $A=\partial\varphi$, where

$$\varphi(x)=\begin{cases}\dfrac{1}{2}\,|A^{1/2}x|^2, & x\in D(A^{1/2}),\\ +\infty, & \text{otherwise.}\end{cases} \tag{2.52}$$

(Here, $A^{1/2}$ is the square root of the operator A.)

Conversely, any linear, densely defined operator that is the subdifferential of an l.s.c. convex function on H is self-adjoint.

To prove these assertions, we note first that any self-adjoint positive operator A in a Hilbert space is maximal monotone. Indeed, it is readily seen that the range of the operator $I+A$ is simultaneously closed and dense in H. On the other hand, if $\varphi : H \to \overline{\mathbf{R}}$ is the function defined by (2.52), then clearly it is convex, l.s.c., and

$$\varphi(x) - \varphi(u) = \frac{1}{2}\,(|A^{1/2}x|^2 - |A^{1/2}u|^2) \leq (Ax, x-u), \qquad \forall x \in D(A),\ u \in D(A^{1/2}).$$

Hence $A \subset \partial\varphi$, and, because A is maximal monotone, we conclude that $A = \partial\varphi$.

Now, let A be a linear, densely defined operator on H of the form $A = \partial\psi$, where $\psi : H \to \overline{\mathbf{R}}$ is an l.s.c. convex function. By Theorem 2.9, we know that $A_\lambda = \nabla\psi_\lambda$, where $A_\lambda = \lambda^{-1}(I - \lambda A)^{-1}$. This yields

$$\frac{d}{dt}\,\psi_\lambda(tu) = t(A_\lambda u, u), \qquad \forall u \in H,\ t \in [0,1],$$

and therefore $\psi_\lambda(u) = (A_\lambda u, u)/2$ for all $u \in H$ and $\lambda > 0$. Calculating the Fréchet derivative of ψ_λ, we see that

$$\nabla\psi_\lambda = A_\lambda = \frac{1}{2}\,(A_\lambda + A_\lambda^*).$$

Hence $A_\lambda = A_\lambda^*$, and letting $\lambda \to 0$, this implies that $A = A^*$, as claimed.

More generally, if A is a linear continuous, symmetric operator from a Hilbert space V to its dual V^* (not identified with V), then $A = \partial\varphi$, where $\varphi : V \to \mathbf{R}$ is the function

$$\varphi(u) = \frac{1}{2}\,(Au, u), \qquad \forall u \in V.$$

Conversely, every linear continuous operator $A : V \to V'$ of the form $\partial\varphi$ is symmetric.

In particular, in virtue of Theorem 1.10, if Ω is a bounded and open domain of $\mathbf{R}^N$ with sufficiently smooth boundary (of class C^2, for instance), then the operator $A : D(A) \subset L^2(\Omega) \to L^2(\Omega)$ defined by

$$Ay = -\Delta y, \qquad \forall y \in D(A), \quad D(A) = H_0^1(\Omega) \cap H^2(\Omega),$$

is self-adjoint and $A = \partial\varphi$, where $\varphi : L^2(\Omega) \to \overline{\mathbf{R}}$, is given by

$$\varphi(y) = \begin{cases} \dfrac{1}{2}\displaystyle\int_\Omega |\nabla y|^2 dx & \text{if } y \in H_0^1(\Omega), \\ +\infty & \text{otherwise.} \end{cases}$$

This result remains true for a nonsmooth bounded open domain if it is convex.

3. *Convex integrands.* Let Ω be a measurable subset of the Euclidean space $\mathbf{R}^N$ and let $L^p(\Omega)$, $1 \le p < \infty$, be the space of all p summable functions on Ω. We set $L^p_m(\Omega) = (L^p(\Omega))^m$.

The function $g : \Omega \times \mathbf{R}^m \to \overline{\mathbf{R}}$ is said to be a *normal convex integrand* if the following conditions hold.

(i) For almost all $x \in \Omega$, the function $g(x,\cdot) : \mathbf{R}^m \to \overline{\mathbf{R}}$ is convex, l.s.c., and not identically $+\infty$.

(ii) g is $\mathscr{L} \times \mathscr{B}$ measurable on $\Omega \times \mathbf{R}^m$; that is, it is measurable with respect to the σ-algebra of subsets of $\Omega \times \mathbf{R}^m$ generated by products of Lebesgue measurable subsets of Ω and Borel subsets of $\mathbf{R}^m$.

We note that if g is convex in y and $\operatorname{int} D(g(x,\cdot)) \neq \emptyset$ for every $x \in \Omega$, then condition (ii) holds if and only if $g = g(x,y)$ is measurable in x for every $y \in \mathbf{R}^m$ (see Rockafellar [27]).

A special case of an $\mathscr{L} \times \mathscr{B}$ measurable integrand is the *Carathéodory integrand*. Namely, one has the following.

Lemma 2.5. *Let $g = g(x,y) : \Omega \times \mathbf{R}^m \to \mathbf{R}$ be continuous in y for every $x \in \Omega$ and measurable in x for every y. Then g is $\mathscr{L} \times \mathscr{B}$ measurable.*

Proof. Let $\{z_i^n\}_{i=1}^\infty$ be a dense subset of $\mathbf{R}^m$ and let $\lambda \in \mathbf{R}$ arbitrary but fixed. Inasmuch as g is continuous in y, it is clear that $g(x,y) \le \lambda$ if and only if for every n there exists z_i^n such that $\|z_i^n - y\| \le (1/n)$ and $g(x,z_i^n) \le \lambda + (1/n)$. Denote by Ω_{in} the set $\{x \in \Omega;\ g(x,z_i^n) \le \lambda + (1/n)\}$ and put $Y_{in} = \{y \in \mathbf{R}^m;\ \|y - z_i^n\| \le 1/n\}$. Inasmuch as

$$\{(x,y) \in \Omega \times \mathbf{R}^m;\ g(x,y) \le \lambda\} = \bigcap_{n=1}^\infty \bigcup_{i=1}^\infty \Omega_{in} \times Y_{in},$$

we infer that g is $\mathscr{L} \times \mathscr{B}$ measurable, as desired. $\square$

Let us assume, in addition to conditions (i) and (ii), the following.

(iii) There are $\alpha \in L^q_m(\Omega)$, $1/p + 1/q = 1$, and $\beta \in L^1(\Omega)$ such that

$$g(x,y) \ge (\alpha(x),y) + \beta(x), \quad \text{a.e. } x \in \Omega,\ y \in \mathbf{R}^m, \tag{2.53}$$

where $(\cdot,\cdot)$ is the usual scalar product in $\mathbf{R}^m$.

(iv) There is $y_0 \in L^p_m$ such that $g(x,y_0) \in L^1(\Omega)$.

Let us remark that if g is independent of x, then conditions (iii) and (iv) automatically hold by virtue of Proposition 1.1.

Define on the space $X = L^p_m(\Omega)$ the function $I_g : X \to \overline{\mathbf{R}}$,

$$I_g(y) = \begin{cases} \displaystyle\int_\Omega g(x,y(x))dx & \text{if } g(x,y) \in L^1(\Omega), \\ +\infty & \text{otherwise.} \end{cases} \tag{2.54}$$

Proposition 2.7. *Let g satisfy assumptions* (i)–(iv). *Then the function I_g is convex, lower semicontinuous, and proper. Moreover,*

$$\partial I_g(y) = \{w \in L^q_m(\Omega);\ w(x) \in \partial g(x, y(x)),\ a.e.\ x \in \Omega\}. \tag{2.55}$$

Here, ∂g is the subdifferential of the function $y \to g(x,y)$.

Proof. Let us show that I_g is well defined (unambiguously a real number or $+\infty$) for every $y \in L^q_m(\Omega)$. Note first that for every Lebesgue measurable function $y : \Omega \to \mathbf{R}^m$ the function $x \to g(x, y(x))$ is Lebesgue measurable on Ω. For a fixed $\lambda \in \mathbf{R}$, we set

$$E = \{(x,y) \in \Omega \times \mathbf{R}^m;\ g(x,y) \le \lambda\}.$$

Let us denote by $\mathscr{S}$ the class of all sets $S \subset \Omega \times \mathbf{R}^m$ having the property that the set $\{x \in \Omega;\ (x, y(x)) \in S\}$ is Lebesgue measurable. Obviously, S contains every set of the form $T \times D$, where T is a measurable subset of Ω and D is an open subset of $\mathbf{R}^m$. Because $\mathscr{S}$ is a σ-algebra, it follows that it contains the σ-algebra generated by the products of Lebesgue measurable subsets of Ω and Borel subsets of $\mathbf{R}^m$. Hence, $E \in \mathscr{S}$, and therefore $g(x, y(x))$ is Lebesgue measurable; that is, I_g is well defined. By assumption (i), it follows that I_g is convex, whereas by (iv) we see that $I_g \not\equiv +\infty$. Let $\{y_n\} \subset L^p_m(\Omega)$ be strongly convergent to y. Then there is $\{y_{n_k}\} \subset \{y_n\}$ such that

$$y_{n_k}(x) \to y(x), \quad \text{a.e. } x \in \Omega \text{ for } n_k \to \infty.$$

Then, by assumption (iii) and by Fatou's lemma, it follows that

$$\begin{aligned}\liminf_{n_k\to\infty} &\int_\Omega (g(x, y_{n_k}(x)) - (\alpha(x), y_{n_k}(x)) - \beta(x))dx \\ &\ge \int_\Omega (g(x, y(x)) - (\alpha(x), y(x)) - \beta(x))dx,\end{aligned}$$

and therefore

$$\liminf_{n_k\to\infty} I_g(y_{n_k}) \ge I_g(y).$$

Clearly, this implies that $\liminf_{n\to\infty} I_g(y_n) \ge I_g(y)$; that is, I_g is l.s.c. on X.

Let us now prove (2.55). It is easily seen that every $w \in L^q_m(\Omega)$ such that $w(x) \in \partial g(x, y(x))$ belongs to $\partial I_g(y)$. Now, let $w \in \partial I_g$; that is,

$$\int_\Omega (g(x, y(x)) - g(x, u(x)))dx \le \int_\Omega (w(x), y(x) - u(x))dx, \qquad \forall u \in L^p_m(\Omega).$$

Let D be an arbitrary measurable subset of Ω and let $u \in L^p_m(\Omega)$ be defined by

$$u(x) = \begin{cases} y_0 & \text{for } x \in D, \\ y(x) & \text{for } x \in \Omega \setminus \overline{D}, \end{cases}$$

where y_0 is arbitrary in $\mathbf{R}^m$. Substituting in the previous inequality, we get

$$\int_D (g(x,y(x)) - g(x,y_0) - (w(x), y(x) - y_0))dx \le 0.$$

D is arbitrary, therefore this implies, a.e. $x \in \Omega$,

$$g(x,y(x)) \le g(x,y_0) + (w(x), y(x) - y_0), \qquad \forall y_0 \in \mathbf{R}^m.$$

Hence, $w(x) \in \partial g(x,y(x))$, a.e. $x \in \Omega$, as claimed. □

The case $p = \infty$ is more subtle, because the elements of $\partial I_g(y) \subset (L_m^\infty(\Omega))^*$ are no longer Lebesgue integrable functions on Ω. It turns out, however, that in this case $\partial I_g(y)$ is of the form $\mu_a + \mu_s$, where $\mu_a \in L_m^1(\Omega)$, $\mu_a(x) \in \partial g(y(x))$, a.e., $x \in \Omega$, and μ_s is a singular element of $(L_m^\infty(\Omega)))^*$. We refer the reader to Rockafellar [28] for the complete description of ∂I_g in this case.

Now, let us consider the special case where

$$g(x,y) = I_K(y) = \begin{cases} 0 & \text{if } y \in K, \\ +\infty & \text{if } y \notin K, \end{cases}$$

K being a closed convex subset of $\mathbf{R}^m$. Then, I_g is the indicator function of the closed convex subset $\mathscr{K}$ of $L_m^p(\Omega)$ defined by

$$\mathscr{K} = \{y \in L_m^p(\Omega);\ y(x) \in K,\ \text{ a.e. } x \in \Omega\},$$

and so by formula (2.55) we see that the normal cone $N_{\mathscr{K}} \subset L_m^q(\Omega)$ to $\mathscr{K}$ is defined by

$$N_{\mathscr{K}}(y) = \{w \in L_m^q(\Omega);\ w(x) \in N_K(y(x)),\ \text{ a.e. } x \in \Omega\}, \tag{2.56}$$

where $N_K(y) = \{z \in \mathbf{R}^m;\ (z, y-u) \ge 0,\ \forall u \in K\}$ is the normal cone at K in $y \in K$.

In particular, if $m = 1$ and $K = [a,b]$, then

$$\begin{aligned} &N_{\mathscr{K}}(y) = \{w \in L^q(\Omega);\ w(x) = 0,\ \text{a.e. in } [x \in \Omega;\ \ a < y(x) < b], \\ &w(x) \ge 0,\ \text{a.e. in } [x \in \Omega;\ y(x) = b],\ w(x) \le 0,\ \text{a.e. in } [x \in \Omega;\ y(x) = a]\}. \end{aligned} \tag{2.57}$$

Let us take now $K = \{y \in \mathbf{R}^m;\ \|y\| \le \rho\}$. Then,

$$N_K(y) = \begin{cases} 0 & \text{if } \|y\| < \rho, \\ \bigcup_{\lambda > 0} \lambda y & \text{if } \|y\| = \rho, \end{cases}$$

and so $N_{\mathscr{K}}$ is given by

$$\begin{aligned} &N_{\mathscr{K}}(y) = \{w \in L_m^q(\Omega);\ w(x) = 0,\ \text{a.e. in } [x \in \Omega; \|y(x)\| < \rho], w(x) = \lambda(x)y(x), \\ &\text{a.e. in } [x \in \Omega;\ \|y(x)\| = \rho],\ \text{where } \lambda \in L_m^q(\Omega),\ \lambda(x) \ge 0,\ \text{a.e. } x \in \Omega\}. \end{aligned}$$

Elliptic nonlinear operators on bounded open domains of $\mathbf{R}^N$ with appropriate boundary value conditions represent another source of maximal monotone operators and, in particular, of subpotential operators. We give a few examples here.

Corollary 2.7. *The mapping* $\phi_1 : L^1(\Omega) \to L^\infty(\Omega)$ *defined by*

$$\phi_1(u) = \{\|u\|_{L^1(\Omega)} w;\ w(x) \in L^\infty(\Omega),\ w(x) \in \operatorname{sign} u(x) \text{ a.e. } x \in \Omega\}$$

is the duality mapping J *of the space* $X = L^1(\Omega)$.

Proof. It is easily seen that $\phi_1(u) \in J(u)$, $\forall u \in L^1(\Omega)$. On the other hand, by Proposition 2.7 we have

$$\partial \|u\|_{L^1(\Omega)} = \{w \in L^\infty(\Omega);\ w(x) \in \operatorname{sign} u(x),\ \text{ a.e. } x \in \Omega\}.$$

This implies that

$$\partial \left(\frac{1}{2}\, \|u\|^2_{L^1(\Omega)} \right) = \phi_1(u), \qquad \forall u \in L^1(\Omega)$$

and, because by Theorem 2.8 the mapping $\partial \left(\frac{1}{2}\, \|u\|^2_{L^1(\Omega)} \right)$ is maximal monotone in $L^1(\Omega) \times L^\infty(\Omega)$, we conclude that so is ϕ_1 and, because $\phi_1 \subset J$, we have $\phi_1 = J$ as claimed. □

4. *Semilinear elliptic operators in* $L^2(\Omega)$. Let Ω be an open bounded subset of $\mathbf{R}^N$, and let $g : \mathbf{R} \to \overline{\mathbf{R}}$ be a lower semicontinuous, convex, proper function such that $0 \in D(\partial g)$.

Define the function $\varphi : L^2(\Omega) \to \overline{\mathbf{R}}$ by

$$\varphi(y) = \begin{cases} \displaystyle\int_\Omega \left(\frac{1}{2}\, |\nabla y|^2 + g(y) \right) dx & \text{if } y \in H^1_0(\Omega) \text{ and } g(y) \in L^1(\Omega), \\ +\infty & \text{otherwise.} \end{cases} \tag{2.58}$$

Proposition 2.8. *The function* φ *is convex, l.s.c., and* $\not\equiv +\infty$. *Moreover, if the boundary* $\partial\Omega$ *is sufficiently smooth* (*for instance, of class* C^2) *or if* Ω *is convex, then* $\partial\varphi \subset L^2(\Omega) \times L^2(\Omega)$ *is given by*

$$\begin{aligned} \partial\varphi = \{\ [y,w];\ w \in L^2(\Omega);\ y \in H^1_0(\Omega) \cap H^2(\Omega), \\ w(x) + \Delta y(x) \in \partial g(y(x)),\ \text{a.e. } x \in \Omega\}. \end{aligned} \tag{2.59}$$

Proof. It is readily seen that φ is convex and $\not\equiv +\infty$. Let $\{y_n\} \subset L^2(\Omega)$ be strongly convergent to y as $n \to \infty$. As seen earlier,

$$\liminf_{n\to\infty} \int_\Omega g(y_n)dx \geq \int_\Omega g(y)dx,$$

and it is also clear, by weak lower semicontinuity of the $L^2(\Omega)$-norm, that

$$\liminf_{n\to\infty}\int_\Omega |\nabla y_n|^2dx \ge \int_\Omega |\nabla y|^2dx.$$

Hence, $\liminf_{n\to\infty}\varphi(y_n)\ge\varphi(y)$.

Let us denote by $\Gamma\subset L^2(\Omega)\times L^2(\Omega)$ the operator defined by the second part of (2.59); that is,

$$\Gamma=\{\ [y,w]\in(H_0^1(\Omega)\cap H^2(\Omega))\times L^2(\Omega);\\ w(x)\in-\Delta y(x)+\partial g(y(x)),\ \text{a.e. } x\in\Omega\}.$$

The inclusion $\Gamma\subset\partial\varphi$ is obvious, thus it suffices to show that Γ is maximal monotone in $L^2(\Omega)$. To this end, observe that $\Gamma=A_2+B$, where $A_2y=-\Delta y$, $\forall y\in D(A_2)=H_0^1(\Omega)\cap H^2(\Omega)$, and $By=\{v\in L^2(\Omega);\ v(x)\in\partial g(y(x)),\text{ a.e. } x\in\Omega\}$. As seen earlier, the operators A_2 and B are maximal monotone in $L^2(\Omega)\times L^2(\Omega)$. Replacing B by $y\to By-y_0$, where $y_0\in B(0)$, we may assume without loss of generality that $0\in B(0)$. On the other hand, it is readily seen that $(B_\lambda u)(x)=\beta_\lambda(u(x))$, a.e. $x\in\Omega$ for all $u\in L^2(\Omega)$, where $\beta=\partial g$, and $\beta_\lambda=\lambda^{-1}(1-(1+\lambda\beta)^{-1})$ is the Yosida approximation of β. We have

$$(A_2u,B_\lambda u)=-\int_\Omega \Delta u\beta_\lambda(u)dx\ge 0,\qquad \forall u\in H_0^1(\Omega)\cap H^2(\Omega),$$

or, equivalently,

$$\int_\Omega g(1+\lambda A_2)^{-1}y(x)dx\le\int_\Omega g(y(x))dx,\qquad \forall y\in L^2(\Omega),$$

which results from the following simple argument. We set $z=(I+\lambda A_2)^{-1}y$:

$$z-\lambda\Delta z=y\quad\text{in }\ \Omega;\ z\in H_0^1(\Omega)\cap H^2(\Omega).$$

If we multiply the latter by $\beta_\mu(z)=(1/\mu)(z-(1+\mu\beta)^{-1}z)$, $\mu>0$, and integrate on Ω, we obtain that

$$\int_\Omega \beta_\mu(z)(z-y)\le 0,\qquad \forall\mu>0,$$

because (inasmuch as $\beta'_\mu\ge 0$) we have

$$\int_\Omega \Delta z\beta_\mu(z)dx=-\int_\Omega \beta'_\mu(z)|\nabla z|^2dx\le 0,\qquad \forall\mu>0.$$

This yields

$$\int_\Omega g_\mu(z)dx\le\int_\Omega g_\mu(y)dx,\qquad \forall\mu>0,$$

where $g_\mu=\beta_\mu$. Then, letting $\mu\to 0$, and recalling Theorem 2.9, we get the desired inequality. (As a matter of fact, this calculation works if $\beta_\lambda\in C^1(\mathbf{R})$ but, in a general situation, we replace β_λ by a C^1 mollifier regularization $(\beta_\lambda)_\varepsilon$ and let ε tend to zero.)

Then, applying Theorem 2.7 (or Theorem 2.11), we may conclude that $\Gamma = A_2 + B$ is maximal monotone. □

Remark 2.2. Because $A_2 + B$ is coercive, it follows from Corollary 2.2 that $R(A_2+B) = L^2(\Omega)$. Hence, for every $f \in L^2(\Omega)$, the Dirichlet problem

$$\begin{cases} -\Delta y + \beta(y) \ni f, & \text{a.e. in } \Omega, \\ y = 0, & \text{on } \partial\Omega, \end{cases} \tag{2.60}$$

has a unique solution $y \in H_0^1(\Omega) \cap H^2(\Omega)$.

In the special case, where $\beta \subset \mathbf{R} \times \mathbf{R}$ is given by

$$\beta(r) = \begin{cases} 0 & \text{if } r > 0, \\ \mathbf{R}^- & \text{if } r = 0, \end{cases}$$

problem (2.60) reduces to the celebrated *obstacle problem*

$$\begin{cases} -\Delta y = f, & \text{a.e. in } [y > 0], \\ -\Delta y \geq f,\ y \geq 0, & \text{a.e. in } \Omega, \\ y = 0, & \text{on } \partial\Omega. \end{cases} \tag{2.61}$$

This is an *elliptic variational inequality* describing a free boundary problem, which is discussed in some detail later.

We also note that the solution y to (2.60) is the limit in $H_0^1(\Omega)$ of the solutions y_ε to the approximating problem

$$\begin{cases} -\Delta y + \beta_\varepsilon(y) = f, & \text{in } \Omega, \\ y = 0, & \text{on } \partial\Omega, \end{cases} \tag{2.62}$$

where β_ε is the Yosida approximation of β. Indeed, multiplying (2.62) by y_ε, we get

$$\|y_\varepsilon\|^2_{H_0^1(\Omega)} + \|\Delta y_\varepsilon\|^2_{L^2(\Omega)} \leq C, \qquad \forall \varepsilon > 0,$$

and therefore $\{y_\varepsilon\}$ is bounded in $H_0^1(\Omega) \cap H^2(\Omega)$. This yields

$$\int_\Omega |\nabla(y_\varepsilon - y_\lambda)|^2 dx + \int_\Omega (\beta_\varepsilon(y_\varepsilon) - \beta_\lambda(y_\lambda))(y_\varepsilon - y_\lambda) dx = 0,$$

and, therefore,

$$\int_\Omega |\nabla(y_\varepsilon - y_\lambda)|^2 dx + \int_\Omega (\beta_\varepsilon(y_\varepsilon) - \beta_\lambda(y_\lambda))(\varepsilon\beta_\varepsilon(y_\varepsilon) - \lambda\beta_\lambda(y_\lambda)) dx \leq 0,$$

because $\beta_\varepsilon(y) \in \beta((1+\varepsilon\beta)^{-1}y)$ and β is monotone. Hence, $\{y_\varepsilon\}$ is Cauchy in $H_0^1(\Omega)$, and so $y = \lim_{\varepsilon\to 0} y_\varepsilon$ exists in $H_0^1(\Omega)$. This clearly also implies that

$$\begin{aligned}
\Delta y_\varepsilon \to \Delta y \qquad & \text{weakly in } L^2(\Omega),\\
y_\varepsilon \to y \qquad & \text{weakly in } H^2(\Omega),\\
\beta_\varepsilon(y_\varepsilon) \to g \qquad & \text{weakly in } L^2(\Omega).
\end{aligned}$$

Now, by Proposition 2.2(iv), we see that $g(x) \in \beta(y(x))$, a.e. $x \in \Omega$, and so y is the solution to problem (2.60).

5. *Nonlinear boundary Neumann conditions.* Let Ω be a bounded and open subset of $\mathbf{R}^N$ with the boundary $\partial\Omega$ of class C^2. Let $j : \mathbf{R} \to \overline{\mathbf{R}}$ be an l.s.c., proper, convex function and let $\beta = \partial j$. Define the function $\varphi : L^2(\Omega) \to \overline{\mathbf{R}}$ by

$$\varphi(u) = \begin{cases} \dfrac{1}{2}\displaystyle\int_\Omega |\nabla u|^2 dx + \int_{\partial\Omega} j(u)dx & \text{if } u \in H^1(\Omega),\ j(u) \in L^1(\partial\Omega),\\ +\infty & \text{otherwise.} \end{cases} \tag{2.63}$$

Because for every $u \in H^1(\Omega)$ the trace of u on $\partial\Omega$ is well defined and belongs to $L^2(\partial\Omega)$ (see Definition 1.2), formula (2.63) makes sense. Moreover, arguing as in the previous example, it follows that φ is convex and l.s.c. on $L^2(\Omega)$. Regarding its subdifferential $\partial\varphi \subset L^2(\Omega) \times L^2(\Omega)$, it is completely described in Proposition 2.9, due to Brezis [3].

Proposition 2.9. *We have*

$$\partial\varphi(u) = -\Delta u, \qquad \forall u \in D(\partial\varphi), \tag{2.64}$$

where

$$D(\partial\varphi) = \left\{ u \in H^2(\Omega);\ -\frac{\partial u}{\partial \nu} \in \beta(u),\ \text{a.e. on } \partial\Omega \right\}$$

and $\partial/\partial\nu$ is the conormal derivative to $\partial\Omega$. Moreover, there are some positive constants C_1, C_2 such that

$$\|u\|_{H^2(\Omega)} \le C_1\|u - \Delta u\|_{L^2(\Omega)} + C_2, \qquad \forall u \in D(\partial\varphi). \tag{2.65}$$

Proof. Let $A : L^2(\Omega) \to L^2(\Omega)$ be the operator defined by

$$\begin{aligned}
Au &= -\Delta u,\ \ u \in D(A),\\
D(A) &= \left\{ u \in H^2(\Omega);\ -\frac{\partial u}{\partial \nu} \in \beta(u),\ \text{a.e. on } \partial\Omega \right\}.
\end{aligned}$$

Note that A is well defined because, for every $u \in H^2(\Omega)$, $(\partial u/\partial\nu) \in H^{1/2}(\partial\Omega)$. It is easily seen that $A \subset \partial\varphi$. Indeed, by Green's formula,

$$\begin{aligned}
\int_\Omega Au(u-v)dx &= \int_\Omega \nabla u(\nabla u - \nabla v)dx + \int_{\partial\Omega} \beta(u)(u-v)dx\\
&\ge \frac{1}{2}\int_\Omega |\nabla u|^2 dx + \int_{\partial\Omega} j(u)dx - \frac{1}{2}\int_\Omega |\nabla v|^2 dx - \int_{\partial\Omega} j(v)dx
\end{aligned}$$

for all $u \in D(A)$ and $v \in H^1(\Omega)$. Hence,

$$(Au, u-v) \geq \varphi(u) - \varphi(v), \qquad \forall u \in D(A),\ v \in L^2(\Omega).$$

(Here, $(\cdot,\cdot)$ is the usual scalar product in $L^2(\Omega)$.) Thus, to show that $A = \partial\varphi$, it suffices to prove that A is maximal monotone in $L^2(\Omega) \times L^2(\Omega)$; that is, $R(I+A) = L^2(\Omega)$. Toward this aim, we fix $f \in L^2(\Omega)$ and consider the equation $u + Au = f$:

$$\begin{aligned} u - \Delta u &= f \quad \text{in } \Omega, \\ \frac{\partial u}{\partial \nu} + \beta(u) &\ni 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{2.66}$$

We approximate (2.66) by

$$\begin{cases} u - \Delta u = f & \text{in } \Omega, \\ \dfrac{\partial u}{\partial \nu} + \beta_\lambda(u) = 0 & \text{on } \partial\Omega, \end{cases} \tag{2.66$'$}$$

where $\beta_\lambda = \lambda^{-1}(1-(1+\lambda\beta)^{-1})$, $\lambda > 0$. Recall that β_λ is Lipschitz continuous with Lipschitz constant $1/\lambda$ and $\beta_\lambda(u) \to \beta^0(u)$, $\forall u \in D(\beta)$, for $\lambda \to 0$.

Let us show first that equation (2.66)$'$ has a unique solution $u_\lambda \in H^2(\Omega)$. Indeed, consider the operator $u \xrightarrow{T} v\big|_{\partial\Omega}$ from $L^2(\partial\Omega)$ to $L^2(\partial\Omega)$, where $v \in H^1(\Omega)$ is the solution to the linear boundary value problem

$$v - \Delta v = f \quad \text{in } \Omega, \qquad v + \lambda\,\frac{\partial v}{\partial \nu} = (1+\lambda\beta)^{-1}u \quad \text{on } \partial\Omega. \tag{2.67}$$

(The existence of v is an immediate consequence of the Lax–Milgram lemma.) Moreover, by Green's formula we see that

$$\begin{aligned} \|v-\bar v\|^2_{L^2(\Omega)} &+ \int_\Omega |\nabla(v-\bar v)|^2 dx + \frac{1}{\lambda}\int_{\partial\Omega}(v-\bar v)^2 dx \\ &\leq \frac{1}{\lambda}\int_{\partial\Omega}((1+\lambda\beta)^{-1}u - (1+\lambda\beta)^{-1}\bar u)(v-\bar v)dx, \end{aligned}$$

where $\{v,u\}$ and $\{\bar v,\bar u\}$ satisfy (2.67). Because

$$|(1+\lambda\beta)^{-1}x - (1+\lambda\beta)^{-1}y| \leq |x-y|, \qquad \forall x,y \in \mathbf{R},\ \lambda > 0,$$

we infer that

$$\|v-\bar v\|^2_{H^1(\Omega)} + \frac{1}{2\lambda}\,\|Tu - T\bar u\|^2_{L^2(\partial\Omega)} \leq \frac{1}{\lambda}\,\|u-\bar u\|^2_{L^2(\partial\Omega)}.$$

Because, by the trace theorem, the map $v \to v\big|_{\partial\Omega}$ is continuous from $H^1(\Omega)$ into $H^{1/2}(\partial\Omega) \subset L^2(\partial\Omega)$, we have

$$\|v-\bar v\|_{H^1(\Omega)} \geq C\|Tu - T\bar u\|_{L^2(\partial\Omega)},$$

and so the map T is a contraction of $L^2(\partial\Omega)$. Applying the Banach fixed point theorem, we therefore conclude that there exists $u \in L^2(\partial\Omega)$ such that $Tu = u$, and so problem (2.66)′ has a unique solution $u_\lambda \in H^1(\Omega)$. We have

$$\begin{cases} u_\lambda - \Delta u_\lambda = f & \text{in } \Omega, \\ \dfrac{\partial u_\lambda}{\partial \nu} = -\beta_\lambda(u_\lambda) & \text{on } \partial\Omega. \end{cases} \tag{2.68}$$

We note that $\beta_\lambda(u_\lambda) \in H^1(\Omega)$ (because β_λ is Lipschitz) and so its trace to $\partial\Omega$ belongs to $H^{1/2}(\partial\Omega)$, we conclude by the classical regularity theory for the linear Neumann problem (see Theorem 1.12) that $u_\lambda \in H^2(\Omega)$.

Let us postpone for the time being the proof of the following estimate,

$$\|u_\lambda\|_{H^2(\Omega)} \le C(1 + \|f\|_{L^2(\Omega)}), \qquad \forall \lambda > 0, \tag{2.69}$$

where C is independent of λ and f.

Now, to obtain existence in problem (2.66), we pass to limit $\lambda \to 0$ in (2.68). Inasmuch as the mapping

$$u \to \left(u\big|_{\partial\Omega}, \frac{\partial u}{\partial \nu}\big|_{\partial\Omega} \right)$$

is continuous from $H^2(\Omega)$ to $H^{3/2}(\partial\Omega) \times H^{1/2}(\partial\Omega)$ and the injection of $H^2(\Omega)$ into $H^1(\Omega) \subset L^2(\Omega)$ is compact, we may assume, selecting a subsequence if necessary, that, for $\lambda \to 0$,

$$\begin{aligned} u_\lambda &\rightharpoonup u && \text{in } H^2(\Omega), \\ u_\lambda &\to u && \text{in } H^1(\Omega), \\ u_\lambda\big|_{\partial\Omega} &\to u\big|_{\partial\Omega} && \text{in } H^{3/2}(\partial\Omega) \subset L^2(\partial\Omega), \\ \frac{\partial u_\lambda}{\partial \nu} &\to \frac{\partial u}{\partial \nu} && \text{in } H^{1/2}(\partial\Omega) \subset L^2(\partial\Omega). \end{aligned} \tag{2.70}$$

Moreover, because by (2.69) $\{\beta_\lambda(u_\lambda)\}$ is bounded in $L^2(\partial\Omega)$, we may assume that, for $\lambda \to 0$,

$$\beta_\lambda(u_\lambda) \rightharpoonup g \quad \text{in } L^2(\partial\Omega). \tag{2.71}$$

It is clear by (2.68), (2.70), and (2.71) that

$$\begin{cases} u - \Delta u = f & \text{in } \Omega, \\ \dfrac{\partial u}{\partial \nu} + g = 0, & \text{a.e. on } \partial\Omega. \end{cases}$$

Let us show that $g(x) \in \beta(u(x))$, a.e. $x \in \Omega$. Indeed, the operator $\widetilde{\beta} \subset L^2(\partial\Omega) \times L^2(\partial\Omega)$ defined by

$$\widetilde{\beta} = \{[u,v] \in L^2(\partial\Omega) \times L^2(\partial\Omega);\ v(x) \in \beta(u(x)) \text{ a.e. } x \in \partial\Omega\}$$

is obviously maximal monotone, and

$$\widetilde{\beta}_\lambda(u)(x) = \beta_\lambda(u(x)), \qquad ((I+\lambda\widetilde{\beta})^{-1}u)(x) = (1+\lambda\beta)^{-1}u(x), \quad \text{a.e. } x \in \partial\Omega.$$

By (2.71), $\widetilde{\beta}_\lambda(u_\lambda) \rightharpoonup g$, $(I+\lambda\widetilde{\beta})^{-1}u_\lambda \to u$, and $\widetilde{\beta}_\lambda(u_\lambda) \in \widetilde{\beta}((I+\lambda\widetilde{\beta})^{-1}u_\lambda)$, therefore we conclude that $g \in \widetilde{\beta}(u)$ (because $\widetilde{\beta}$ is strongly–weak closed). We have therefore proved that u is a solution to equation (2.66), and because f is arbitrary in $L^2(\Omega)$, we infer that $A = \partial\varphi$. Finally, letting λ tend to zero in the estimate (2.69), we obtain (2.65), as claimed. □

Proof of estimate (2.69). Multiplying equation (2.68) by $u_\lambda - u_0$, where $u_0 \in D(\beta)$ is a constant, we get after some calculation involving Green's lemma that

$$\int_\Omega (u_\lambda^2 + |\nabla u_\lambda|^2)dx \le C\left(\int_\Omega f^2 dx + 1\right).$$

(We denote by C several positive constants independent of λ and f.) Hence,

$$\|u_\lambda\|_{H^1(\Omega)} \le C(\|f\|_{L^2(\Omega)} + 1), \qquad \forall \lambda > 0. \tag{2.72}$$

If Ω' is an open subset of Ω such that $\overline{\Omega}' \subset \Omega$, then we choose $\rho \in C_0^\infty(\Omega)$ such that $\rho = 1$ in $\overline{\Omega}'$. We set $v = \rho u_\lambda$ and note that

$$v - \Delta v = \rho f - u_\lambda \Delta\rho - 2\nabla\rho \cdot \nabla u_\lambda \quad \text{in } \Omega. \tag{2.73}$$

Because v has compact support in Ω, we may assume that $v \in H^2(\mathbf{R}^N)$, and equation (2.73) extends to all of $\mathbf{R}^N$. Then, taking the Fourier transform and using Parseval's formula, we get

$$\|v\|_{H^2(\mathbf{R}^N)} \le C(\|f\|_{L^2(\Omega)} + \|u_\lambda\|_{H^1(\Omega)}),$$

and, therefore, by (2.72) we get the internal estimate

$$\|u_\lambda\|_{H^2(\Omega')} \le C(\|f\|_{L^2(\Omega)} + 1), \qquad \forall \lambda > 0, \tag{2.74}$$

where C is dependent of $\Omega' \subset\subset \Omega$.

To obtain H^2-estimates near the boundary $\partial\Omega$, we use the classical method of tangential quotients. Namely, let $x_0 \in \partial\Omega$, U be a neighborhood of x_0, and $\varphi : U \to Q$ be such that $\varphi \in C^2(U)$, $\varphi^{-1} \in C^2(Q)$, $\varphi^{-1}(Q_+) = \Omega \cap U$, and $\varphi^{-1}(Q_0) = \partial\Omega \cap U$, where $Q = \{y \in \mathbf{R}^N;\ \|y'\| < 1,\ |y_N| < 1\}$, $Q_+ = \{y \in Q;\ 0 < y_N < 1\}$, $Q_0 = \{y \in Q;\ y_N = 0\}$, and $y = (y', y_N) \in \mathbf{R}^N$. (Because $\partial\Omega$ is of class C^2, such a pair (U,φ) always exists.) Now, we "transport" equation (2.73) from $U \cap \Omega$ on Q, using the local coordinate φ. We set

$$w(y) = u_\lambda(\psi(y)), \qquad \forall y \in Q_+,\ \psi = \varphi^{-1},$$

and notice that w satisfies on Q_+ the boundary value problem

$$\begin{cases} w - \displaystyle\sum_{k,j=1}^{N} \frac{\partial}{\partial y_j}\left(a_{kj}(y)\,\frac{\partial w}{\partial y_k}\right) + \sum_{j=1}^{N} b_j(y)\,\frac{\partial w}{\partial y_j} + c(y)w = g(y) & \text{in } Q_+, \\ \dfrac{\partial w}{\partial \nu} + \beta_\lambda(w) = 0 & \text{on } Q_0, \end{cases} \tag{2.75}$$

where $g(y) = f(\psi(y))$,

$$a_{kj}(y) = \sum_{1=1}^{N} \frac{\partial \varphi_k}{\partial x_\ell}\,\frac{\partial \varphi_j}{\partial x_\ell}, \qquad \forall y \in Q_+,\ \varphi = (\varphi_1,\dots,\varphi_N),$$

and

$$\frac{\partial w}{\partial \nu} = \sum_{i,j=1}^{N} \frac{\partial w}{\partial y_j}\,\frac{\partial \varphi_j}{\partial x_i}\,\cos(\nu, x_i)$$

(ν is the conormal derivative to $\partial\Omega$). Because $\varphi_N(x) = 0$ is the equation of the surface $\partial\Omega \cap U$, we may assume that $\partial\varphi_N/\partial x_j = \cos(\nu, x_j)$, and so

$$\frac{\partial w}{\partial \nu} = -\sum_{j=1}^{N} \frac{\partial w}{\partial y_j}\,a_{jN} \quad \text{on } Q_0.$$

Assuming for a while that $f \in C^1(\overline{\Omega})$, we see that $z = \partial w/\partial y_i$, $1 \le i \le N-1$, satisfies the equation

$$\begin{cases} z - \displaystyle\sum_{k,j=1}^{N} \frac{\partial}{\partial y_j}\left(a_{kj}\,\frac{\partial z}{\partial y_k}\right) + \sum_{j=1}^{N}\left(\widetilde{b}_j\,\frac{\partial z}{\partial y_j} + \widetilde{c}_j\,\frac{\partial w}{\partial y_j}\right) & \\ \qquad\qquad + c(y)z + c'(y)w = \dfrac{\partial}{\partial y_i}\,g(y) & \text{in } Q_+, \\ \dfrac{\partial z}{\partial \nu} = -\beta_\lambda'(u_\lambda)z + \displaystyle\sum_{j=1}^{N} \frac{\partial w}{\partial y_j}\,\frac{\partial a_{JN}}{\partial y_i} & \text{on } Q_0. \end{cases} \tag{2.76}$$

Now, let $\rho \in C_0^\infty(\overline{Q}_+)$ be such that $\rho(y)=0$ for $\|y'\| \ge \frac{2}{3}$, $\frac{2}{3} < y_i < 1$, and $\rho(y)=1$ for $\|y'\| < \frac{1}{2}$ and $0 \le y_N \le \frac{1}{2}$. Multiplying (2.76) by $\rho^2 z$ and integrating on Q_+, we get

$$\begin{aligned} \int_{Q_+} \rho^2 z^2 dy &+ \sum_{k,j=1}^{N} \int_{Q_+} a_{kj}(y)\,\frac{\partial z}{\partial y_k}\,\frac{\partial}{\partial y_j}(\rho^2 z)dy + \int_{Q_0} \rho^2 \beta_\lambda'(u_\lambda)z^2 dy \\ &= \sum_{j=1}^{N} \int_{Q_0} \rho^2\,\frac{\partial w}{\partial y_j}\,\frac{\partial a_{jN}}{\partial y_i}\,dy + \int_{Q_+} \frac{\partial}{\partial y_i}\,g(y)z(y)dy \\ &\quad - \sum_{j=1}^{N} \int_{Q_+} \rho^2\left(\widetilde{b}_j\,\frac{\partial z}{\partial y_j} + \widetilde{c}_j\,\frac{\partial w}{\partial y_j}\right) z\,dy + \int_{Q_+} (cz + c'w)z\rho^2 dy. \end{aligned}$$

Taking into account that

$$\sum_{k,j=1}^{N} a_{kj}(y)\xi_i\xi_j \geq \omega\|\xi\|^2, \qquad \forall y \in Q_+,\ \xi \in \mathbf{R}^N,$$

we find after some calculations that

$$\sum_{j=1}^{N} \int_{Q_+} \rho^2(y) \left(\frac{\partial z}{\partial y_j}\right)^2 dy \leq C(\|g\|^2_{L^2(Q_+)} + \|w\|^2_{H^2(Q_+)} + 1).$$

Hence,

$$\left\|\rho\ \frac{\partial^2 w}{\partial y_i \partial y_j}\right\|_{L^2(Q_+)} \leq C(\|f\|_{L^2(\Omega)} + \|u_\lambda\|_{H^1(\Omega)} + 1)$$

for $i = 1, 2, ..., N-1,\ j = 1, ..., N$.

Because $a_{NN}(y) \geq w_0 > 0$ for all $y \in Q_+$, by equation (2.75) and the last estimate, we see that

$$\left\|\frac{\partial^2 w}{\partial y_N^2}\right\|_{L^2(Q_+)} \leq C(\|f\|_{L^2(\Omega)} + \|u_\lambda\|_{H^1(\Omega)} + 1).$$

Hence,

$$\|\rho w\|_{H^2(Q_+)} \leq C(\|f\|_{L^2(\Omega)} + 1).$$

Equivalently,

$$\|(\rho \cdot \varphi)u_\lambda\|_{H^2(U\cap\Omega)} \leq C(\|f\|_{L^2(\Omega)} + 1), \qquad \forall \lambda > 0.$$

Hence, there is a neighborhood $U' \subset U$ such that

$$\|u_\lambda\|_{H^2(U'\cap\Omega)} \leq C(\|f\|_{L^2(\Omega)} + 1), \qquad \forall \lambda > 0. \tag{2.77}$$

Now, taking a finite partition of unity subordinated to such a cover $\{U\}$ of $\partial\Omega$ and using the local estimates (2.74) and (2.77), we get (2.69). This completes the proof of Proposition 2.9. □

We have incidentally proved that, for every $f \in L^2(\Omega)$, the boundary value problem (2.66) has a unique solution $u \in H^2(\Omega)$. If $\beta \subset \mathbf{R} \times \mathbf{R}$ is the graph

$$\beta(0) = \mathbf{R}, \qquad \beta(r) = \emptyset \quad \text{for } r \neq 0,$$

then (2.66) reduces to the classical Dirichlet problem. If

$$\beta(r) = \begin{cases} 0 & \text{if } r > 0, \\ (-\infty, 0] & \text{if } r = 0, \end{cases} \tag{2.78}$$

then problem (2.66) can be equivalently written as

$$\begin{cases} y - \Delta y = f & \text{in } \Omega, \\ y\dfrac{\partial y}{\partial \nu} = 0, \quad y \geq 0, \quad \dfrac{\partial y}{\partial \nu} \geq 0 & \text{on } \partial\Omega. \end{cases} \tag{2.79}$$

This is the celebrated *Signorini's problem*, which arises in elasticity in connection with the mathematical description of friction problems. This is a problem of unilateral type and the subset Γ_0 that separates $\{x \in \partial\Omega;\ y > 0\}$ from $\{x \in \partial\Omega;\ (\partial y/\partial \nu) > 0\}$ is a *free boundary* and it is one of the unknowns of the problem.

For other unilateral problems of physical significance that can be written in the form (2.66), we refer to the book of Duvaut and Lions [12].

Remark 2.3. As mentioned earlier, Proposition 2.9 and its corollaries remain valid if Ω is an *open, bounded, and convex subset* of $\mathbf{R}^N$. The idea is to approximate such a domain Ω by smooth domain Ω_ε, to use the estimate (2.69) (which is valid on every Ω_ε with a constant C independent of ε), and to pass to the limit. It is useful to note that the constant C in estimate (2.69) is independent of β.

6. *The nonlinear diffusion operator.* Let Ω be a bounded and open subset of $\mathbf{R}^N$ with a sufficiently smooth boundary $\partial\Omega$. Denote as usual by $H^1_0(\Omega)$ the Sobolev space of all $u \in H^1(\Omega)$ having null trace on $\partial\Omega$ and by $H^{-1}(\Omega)$ the dual of $H^1_0(\Omega)$. Note that $H^{-1}(\Omega)$ is a Hilbert space with the scalar product

$$\langle u,v\rangle = (J^{-1}u,v) \qquad \forall u,v \in H^{-1}(\Omega),$$

where $J = -\Delta$ is the canonical isomorphism (duality mapping) of $H^1_0(\Omega)$ onto $H^{-1}(\Omega)$ and $(\cdot,\cdot)$ is the pairing between $H^1_0(\Omega)$ and $H^{-1}(\Omega)$.

Let $j : \mathbf{R} \to \overline{\mathbf{R}}$ be an l.s.c., convex, proper function and let $\beta = \partial j$. Define the function $\varphi : H^{-1}(\Omega) \to \overline{\mathbf{R}}$ by

$$\varphi(u) = \begin{cases} \displaystyle\int_\Omega j(u(x))dx & \text{if } u \in L^1(\Omega) \text{ and } j(u) \in L^1(\Omega), \\ +\infty & \text{otherwise.} \end{cases} \tag{2.80}$$

It turns out (see Proposition 2.10 below) that the subdifferential $\partial\varphi : H^{-1}(\Omega) \to H^{-1}(\Omega)$ of φ is just the operator $u \to -\Delta\beta(u)$ with appropriate boundary conditions.

The equation $\lambda u - \Delta\beta(u) = f$ is known in the literature as the *nonlinear diffusion equation* or the *porous media equation*.

Proposition 2.10. *Let us assume that*

$$\lim_{|r|\to\infty} \frac{j(r)}{|r|} = +\infty. \tag{2.81}$$

Then the function φ is convex and lower semicontinuous on $H^{-1}(\Omega)$. Moreover, $\partial\varphi \subset H^{-1}(\Omega) \times H^{-1}(\Omega)$ is given by

$$\begin{aligned} \partial\varphi = \{[u,w] \in (H^{-1}(\Omega) \cap L^1(\Omega)) \times H^{-1}(\Omega);\ w = -\Delta v, \\ v \in H^1_0(\Omega),\ v(x) \in \beta(u(x)),\ a.e.\ x \in \Omega\}. \end{aligned} \tag{2.82}$$

Proof. Obviously, φ is convex. To prove that φ is l.s.c., consider a sequence $\{u_\lambda\} \subset H^{-1}(\Omega) \cap L^1(\Omega)$ such that $u_n \to u$ in $H^{-1}(\Omega)$ and $\varphi(u_n) \le \lambda$; that is,

$\int_\Omega j(u_n)dx \le \lambda$, $\forall n$. We must prove that $\int_\Omega j(u)dx \le \lambda$. We have already seen in the proof of Proposition 2.7 that the function $u \to \int_\Omega j(u)dx$ is lower semicontinuous on $L^1(\Omega)$. Because this function is convex, it is weakly lower semicontinuous in $L^1(\Omega)$ and so it suffices to show that $\{u_n\}$ is weakly compact in $L^1(\Omega)$. According to the Dunford–Pettis criterion (see Theorem 1.15), we must prove that the integrals $\int |u_n|dx$ are uniformly absolutely continuous; that is, for every $\varepsilon > 0$ there is $\delta(\varepsilon)$ such that $\int_E |u_n(x)|dx \le \varepsilon$ if $m(E) \le \delta(\varepsilon)$ (E is a measurable set of Ω) and m is the Lebesgue measure. By condition (2.81), for every $p > 0$ there exists $R(p) > 0$ such that $j(r) \ge p|r|$ if $|r| \ge R(p)$. This clearly implies that $\int_\Omega |u_n(x)|dx \le C$.

Moreover, for every measurable subset E of Ω, we have

$$\begin{aligned}\int_E |u_n(x)|dx &\le \int_{E\cap\{|u_n|\ge R(p)\}} |u_n(x)|dx + \int_{E\cap\{|u_n|<R(p)\}} |u_n(x)|dx \\ &\le \frac{1}{p}\int_\Omega |u_n(x)|dx + R(p)m(E) \le \varepsilon,\end{aligned}$$

if we choose $p > (2\varepsilon)^{-1}\sup\int_\Omega |u_n(x)|dx$ and $m(E) \le (\varepsilon/(2R(p)))$. Hence, $\{u_n\}$ is weakly compact in $L^1(\Omega)$.

To prove (2.82), consider the operator $A \subset H^{-1}(\Omega) \times H^{-1}(\Omega)$ defined by

$$Au = \{-\Delta v;\ v \in H_0^1(\Omega),\ v(x) \in \beta(u(x)),\ \text{a.e. } x \in \Omega\},$$

where $D(A) = \{u \in H^{-1}(\Omega) \cap L^1(\Omega);\ \exists v \in H_0^1(\Omega),\ v(x) \in \beta(u(x)), \text{a.e. } x \in \Omega\}$. To prove that $A = \partial\varphi$, proceeding as in the previous case, we show separately that $A \subset \partial\varphi$ and that A is maximal monotone. Let us show first that $R(I+A) = H^{-1}(\Omega)$. Let f be arbitrary but fixed in $H^{-1}(\Omega)$. We must show that there exist $u \in H^{-1}(\Omega) \cap L^1(\Omega)$ and $v \in H_0^1(\Omega)$ such that

$$u - \Delta v = f \quad \text{in } \Omega, \qquad v(x) \in \beta(u(x)), \quad \text{a.e. } x \in \Omega;$$

or equivalently,

$$\begin{aligned} u - \Delta v = f \quad &\text{in } \Omega,\ u(x) \in \gamma(v(x)), \quad \text{a.e. } x \in \Omega, \\ &u \in H^{-1}(\Omega) \cap L^1(\Omega),\ v \in H_0^1(\Omega), \end{aligned} \tag{2.83}$$

where $\gamma = \beta^{-1}$.

Consider the approximating equation

$$\gamma_\lambda(v) - \Delta v = f \quad \text{in } \Omega, \qquad v = 0 \quad \text{on } \partial\Omega, \tag{2.84}$$

where $\gamma_\lambda = \lambda^{-1}(1 - \lambda\gamma)^{-1}$, $\lambda > 0$. It is readily seen that (2.84) has a unique solution $v_\lambda \in H_0^1(\Omega)$. Indeed, because $-\Delta$ is maximal monotone from $H_0^1(\Omega)$ to $H^{-1}(\Omega)$ and $v \to \gamma_\lambda(v)$ is monotone and continuous from $H_0^1(\Omega)$ to $H^{-1}(\Omega)$ (in fact, from $L^2(\Omega)$ to itself), we infer by Corollary 2.1 that $v \to \gamma_\lambda(v) - \Delta v$ is maximal monotone in $H_0^1(\Omega) \times H^{-1}(\Omega)$, and by Corollary 2.2 that it is surjective. Let $v_0 \in D(\gamma)$. Multiplying equation (2.84) by $v_\lambda - v_0$, we get

$$\int_\Omega |\nabla v_\lambda|^2 dx + \int_\Omega \gamma(v_0)(v_\lambda - v_0)dx \le (v_\lambda - v_0, f).$$

Hence, $\{v_\lambda\}$ is bounded in $H_0^1(\Omega)$. Then, on a subsequence, again denoted by λ, we have

$$v_\lambda \rightharpoonup v \quad \text{in } H_0^1(\Omega), \qquad v_\lambda \to v \quad \text{in } L^2(\Omega).$$

Thus, extracting further subsequences, we may assume that

$$\begin{aligned} v_\lambda(x) &\to v(x), \quad \text{a.e. } x \in \Omega, \\ (1+\lambda\gamma)^{-1} v_\lambda(x) &\to v(x), \quad \text{a.e. } x \in \Omega, \end{aligned} \tag{2.85}$$

because, by condition (2.81) and Proposition 1.7, it follows that $D(\gamma) = R(\beta) = \mathbf{R}$ (β is coercive) and so $\lim_{\lambda\to 0}(1+\lambda\gamma)^{-1}r = r$ for all $r \in \mathbf{R}$ (Proposition 2.2).

We get $g_\lambda = \gamma_\lambda(v_\lambda)$. Then, letting λ tend to zero in (2.84), we see that $g_\lambda \to u$ in $H^{-1}(\Omega)$ and

$$u - \Delta v = f \quad \text{in } \Omega, \qquad v \in H_0^1(\Omega).$$

It remains to be shown that $u \in L^1(\Omega)$ and $u(x) \in \gamma(v(x))$, a.e. $x \in \Omega$.

Multiplying equation (2.84) by v_λ, we see that

$$\int_\Omega g_\lambda v_\lambda dx \le C, \qquad \forall \lambda > 0.$$

On the other hand, for some $u_0 \in D(j)$ we have $j(g_\lambda(x)) \le j(u_0) + (g_\lambda(x) - u_0)v$, $\forall v \in \beta(g_\lambda(x))$. This yields

$$\int_\Omega j(g_\lambda(x))dx \le C, \qquad \forall \lambda > 0,$$

because $(1+\lambda\gamma)^{-1}v_\lambda \in \beta(g_\lambda)$.

As seen before, this implies that $\{g_\lambda\}$ is weakly compact in $L^1(\Omega)$. Hence, $u \in L^1(\Omega)$ and

$$g_\lambda \rightharpoonup u \quad \text{in } L^1(\Omega) \text{ for } \lambda \to 0. \tag{2.86}$$

On the other hand, by (2.85) it follows by virtue of the Egorov theorem that for every $\varepsilon > 0$ there exists a measurable subset $E_\varepsilon \subset \Omega$ such that $m(\Omega \setminus E_\varepsilon) \le \varepsilon$, $\{(1+\lambda\gamma)^{-1}v_\lambda\}$ is bounded in $L^\infty(E_\varepsilon)$, and

$$(1+\lambda\gamma)^{-1}v_\lambda \to v \quad \text{uniformly in } E_\varepsilon \text{ as } \lambda \to 0. \tag{2.87}$$

Recalling that $g_\lambda(x) \in \gamma((1+\lambda\gamma)^{-1}v_\lambda(x))$ and that the operator

$$\widetilde{\gamma} = \{[u,v] \in L^1(E_\varepsilon) \times L^\infty(E_\varepsilon);\ u(x) \in \gamma(v(x)), \text{ a.e. } x \in E_\varepsilon\},$$

is maximal monotone in $L^1(E_\varepsilon)\times L^\infty(E_\varepsilon)$, we infer, by (2.86) and (2.87), that $[u,v]\in\widetilde{\gamma}$; that is, $v(x) \in \beta(u(x))$, a.e. $x \in E_\varepsilon$. Because ε is arbitrary, we infer that $v(x) \in \beta(u(x))$, a.e. $x \in \Omega$, as desired. □

To prove that $A \subset \partial\varphi$, we must use the definition of A. However, in order to avoid a formal calculus with symbol (w,u), we need the following lemma, which is a special case of a general result due to Brezis and Browder [8].

Lemma 2.6. *Let Ω be an open subset of $\mathbf{R}^N$. If $w \in H^{-1}(\Omega) \cap L^1(\Omega)$ and $u \in H_0^1(\Omega)$ are such that*

$$w(x)u(x) \geq -|h(x)|, \quad \text{a.e. } x \in \Omega, \tag{2.88}$$

for some $h \in L^1(\Omega)$, then $wu \in L^1(\Omega)$ and

$$w(u) = \int_\Omega w(x)u(x)dx. \tag{2.89}$$

(Here, $w(u)$ is the value of functional $w \in H^{-1}(\Omega)$ at $u \in H_0^1(\Omega)$.)

Proof. The exact meaning of Lemma 2.6 is that, for u in $H_0^1(\Omega)$, the distribution $w \in H^{-1}(\Omega)$ computed at u is represented by the integral (2.89). This is of course obvious if $u \in C_0^\infty(\Omega)$ or $u \in C_0^1(\Omega)$ but less obvious if $u \in H_0^1(\Omega)$. The proof relies on an approximation result for the functions of $H_0^1(\Omega)$ due to Hedberg [15].

Let $u \in H_0^1(\Omega)$. Then there exists a sequence $\{u_n\} \subset C_0^1(\Omega)$ such that $u_n \to u$ in $H_0^1(\Omega)$ and

$$|u_n(x)| \leq \inf(n, |u(x)|), \qquad u_n(x)u(x) \geq 0, \quad \text{a.e. } x \in \Omega. \tag{2.90}$$

(Such a sequence can be chosen by mollifying the function u.) Then, $w(u_n)$ can be represented as

$$w(u_n) = \int_\Omega w(x)u_n(x)dx, \qquad \forall n. \tag{2.91}$$

On the other hand, by (2.88) we have

$$wu_n + |h|\frac{u_n}{u} = (wu + |hy|)\frac{u_n}{u} \geq 0, \quad \text{a.e. in } \Omega,$$

and so, by the Fatou lemma, $wu + |h| \in L^1(\Omega)$ and

$$\liminf_{n\to\infty} \int_\Omega \left(wu + |h|\frac{u_n}{u}\right)dx \geq \int_\Omega (wu + |h|)dx$$

because, on a subsequence, $u_n(x) \to u(x)$, a.e. $x \in \Omega$.

We have, therefore, proved that $wu \in L^1(\Omega)$ and

$$\liminf_{n\to\infty} \int_\Omega wu_n dx \geq \int_\Omega wu\,dx.$$

On the other hand, $wu_n \to wu$, a.e. in Ω, and, by (2.90), $|wu_n| \leq |wu|$, a.e. in Ω. Then, by the Lebesgue dominated convergence theorem, we infer that $wu_n \to wu$ in $L^1(\Omega)$, and letting $n \to \infty$ in (2.91) we get (2.89), as desired. □

Now, to conclude the proof of Proposition 2.10, consider an arbitrary element $[u,-\Delta u] \in A$; that is, $u \in H^{-1}(\Omega) \cap L^1(\Omega)$, $v \in H_0^1(\Omega)$, $v(x) \in \beta(u(x))$, a.e. $x \in \Omega$. We have

$$\langle Au, u-\bar{u}\rangle = (v, u-\bar{u}), \qquad \forall \bar{u} \in H^{-1}(\Omega) \cap L^1(\Omega).$$

Because $v(x)(u(x)-\bar{u}(x)) \geq j(u(x)) - j(\bar{u}(x))$, a.e., $x \in \Omega$, it follows by Lemma 2.6 that

$$\begin{aligned}\langle Au, u-\bar{u}\rangle &= (v, u-\bar{u}) = \int_\Omega v(x)(u(x)-\bar{u}(x))dx \\ &\geq \int_\Omega j(u(x))dx - \int_\Omega j(\bar{u}(x))dx, \qquad \forall \bar{u} \in D(\varphi).\end{aligned}$$

Hence,

$$\langle Au, u-\bar{u}\rangle \geq \varphi(u) - \varphi(\bar{u}), \qquad \forall \bar{u} \in H^{-1}(\Omega),$$

thereby completing the proof.

Remark 2.4. As seen in Proposition 1.7, condition (2.81) is equivalent to $R(\beta)=\mathbf{R}$ and β^{-1} is bounded on bounded sets.

2.3 Elliptic Variational Inequalities

Let X be a reflexive Banach space with the dual X^* and let $A : X \to X^*$ be a monotone operator (linear or nonlinear). Let $\varphi : X \to \overline{\mathbf{R}}$ be a lower semicontinuous convex function on X, $\varphi \not\equiv +\infty$. If f is a given element of X, consider the following problem.

Find $y \in X$ such that

$$(y-z, Ay) + \varphi(y) - \varphi(z) \leq (y-z, f), \qquad \forall z \in X. \tag{2.92}$$

This is an *abstract elliptic variational* inequality associated with the operator A and the convex function φ, and it can be equivalently expressed as

$$Ay + \partial\varphi(y) \ni f, \tag{2.93}$$

where $\partial\varphi \subset X \times X^*$ is the subdifferential of φ. In the special case where $\varphi = I_K$ is the indicator function of a closed convex

$$I_K(x) = \begin{cases} 0 & \text{if } x \in K, \\ +\infty & \text{otherwise,} \end{cases}$$

problem (2.92) becomes:

Find $y \in K$ such that

$$(y-z, Ay) \le (y-z, f), \qquad \forall z \in K. \tag{2.94}$$

It is useful to notice that if the operator A is itself a subdifferential $\partial\psi$ of a continuous convex function $\psi : X \to \mathbf{R}$, then the variational inequality (2.92) is equivalent to the minimization problem (the Dirichlet principle)

$$\min\{\psi(z) + \varphi(z) - (z, f);\ z \in X\} \tag{2.95}$$

or, in the case of problem (2.94),

$$\min\{\psi(z) - (z, f);\ z \in K\}. \tag{2.96}$$

As far as existence in problem (2.92) is concerned, we note first the following result.

Theorem 2.13. *Let $A : X \to X^*$ be a monotone demicontinuous operator and let $\varphi : X \to \overline{\mathbf{R}}$ be a lower semicontinuous, proper, convex function. Assume that there exists $y_0 \in D(\varphi)$ such that*

$$\lim_{\|y\|\to\infty} \frac{(y-y_0, Ay) + \varphi(y)}{\|y\|} = +\infty. \tag{2.97}$$

Then, problem (2.92) has at least one solution. Moreover, the set of solutions is bounded, convex, and closed in X and if the operator A is strictly monotone (i.e., $(Au - Av, u - v) = 0 \Longleftrightarrow u = v$), then the solution is unique.

Proof. By Theorem 2.4, the operator $A + \partial\varphi$ is maximal monotone in $X \times X^*$. By condition (2.97) it is also coercive, therefore we conclude (see Corollary 2.2) that it is surjective. Hence, equation (2.93) (equivalently, (2.92)) has at least one solution.

The set of all solutions y to (2.92) is $(A + \partial\varphi)^{-1}(f)$, thus we infer that this set is closed and convex (see Proposition 2.1). By the coercivity condition (2.97), it is also bounded. Finally, if A (or, more generally, if $A + \partial\varphi$) is strictly monotone, then $(A + \partial\varphi)^{-1} f$ consists of a single element. □

In the special case $\varphi = I_K$, we have the following.

Corollary 2.8. *Let $A : X \to X^*$ be a monotone demicontinuous operator and let K be a closed convex subset of X. Assume either that there is $y_0 \in K$ such that*

$$\lim_{\|y\|\to\infty} \frac{(y-y_0, Ay)}{\|y\|} = +\infty, \tag{2.98}$$

or that K is bounded. Then problem (2.92) has at least one solution. The set of all solutions is bounded, convex, and closed. If A is strictly monotone, then the solution to (2.92) is unique.

To be more specific, we assume in the following that $X = V$ is a Hilbert space, $X^* = V'$, and

$$V \subset H \subset V' \tag{2.99}$$

algebraically and topologically, where H is a real Hilbert space identified with its own dual. The norms of V and H are denoted by $\|\cdot\|$ and $|\cdot|$, respectively. For $v \in V$ and $v' \in V'$ we denote by (v,v') the value of v' at v; if $v,v' \in H$, this is the scalar product in H of v and v'. The norm in V' is denoted by $\|\cdot\|_*$.

Let $A \in L(V,V')$ be a linear continuous operator from V to V such that, for some $\omega > 0$,

$$(v,Av) \geq \omega\|v\|^2, \qquad \forall v \in V.$$

Very often, the operator A is defined by the equation

$$(u,Av) = a(u,v), \qquad \forall u,v \in V, \tag{2.100}$$

where $a : V \times V \to \mathbf{R}$ is a bilinear continuous functional on $V \times V$ such that

$$a(v,v) \geq \omega\|v\|^2, \qquad \forall v \in V. \tag{2.101}$$

In terms of a, the variational inequality (2.92) on V becomes

$$a(y,y-z) + \varphi(y) - \varphi(z) \leq (y-z,f), \qquad \forall z \in V, \tag{2.102}$$

and (2.94) reduces to

$$y \in K,\ \ a(y,y-z) \leq (y-z,f), \qquad \forall z \in K. \tag{2.103}$$

As we show later in the application, V is usually a Sobolev space on an open subset Ω of $\mathbf{R}^N$, $H = L^2(\Omega)$ and A is an elliptic differential operator on Ω with appropriate homogeneous boundary value conditions. The set K incorporates various unilateral conditions on the domain Ω or on its boundary $\partial\Omega$.

By Theorem 2.8, we have the following existence result for problem (2.102).

Corollary 2.9. *Let $a : V \times V \to \mathbf{R}$ be a bilinear continuous functional satisfying condition* (2.101) *and let $\varphi : V \to \overline{\mathbf{R}}$ be an l.s.c., convex, proper function. Then, for every $f \in V'$, problem* (2.102) *has a unique solution $y \in V$. The map $f \to y$ is Lipschitz from V' to V.*

Similarly for problem (2.103).

Corollary 2.10. *Let $a : V \times V \to \mathbf{R}$ be a bilinear continuous functional satisfying condition* (2.101) *and let K be a closed convex subset of V. Then, for every $f \in V'$, problem* (2.103) *has a unique solution y. The map $f \to y$ is Lipschitz continuous from V' to V.*

A problem of great interest when studying equation (2.102) is whether $Ay \in H$. To answer this problem, we define the operator $A_H : H \to H$,

$$A_H y = Ay \quad \text{for } y \in D(A_H) = \{u \in V;\ Au \in H\}. \tag{2.104}$$

The operator A_H is positive definite on H and $R(I + A_H) = H$ (I is the unit operator in H). (Indeed, by Lemma 1.3, the operator $I + A$ is surjective from V to V'.) Hence, A_H is maximal monotone in $H \times H$.

Theorem 2.14. *Under the assumptions of Corollary* 2.8*, suppose in addition that there exist* $h \in H$ *and* $C \in \mathbf{R}$ *such that*

$$\varphi(I+\lambda A_H)^{-1}(y+\lambda h) \le \varphi(y) + C\lambda, \qquad \forall \lambda > 0,\ y \in V. \tag{2.105}$$

Then, if $f \in H$*, the solution* y *to* (2.102) *belongs to* $D(A_H)$ *and*

$$|Ay| \le C(I + |f|). \tag{2.106}$$

Proof. Let $A_\lambda \in L(H,H)$ be the Yosida approximation of A_H; that is,

$$A_\lambda = \lambda^{-1}(I - (I+\lambda A_H)^{-1}), \qquad \lambda > 0.$$

Let $y \in V$ be the solution to (2.102). If in (2.102) we set $z = (I+\lambda A_H)^{-1}(y+\lambda h)$ and use condition (2.105), we get

$$(Ay, A_\lambda y) - (Ay, (I+\lambda A_H)^{-1}h) \le (A_\lambda y, f) - ((I+\lambda A_H)^{-1}h, f).$$

Because $(Ay, A_\lambda y) \ge |A_\lambda y|^2$ for all $\lambda > 0$ and $y \in V$, we get

$$|A_\lambda y|^2 \le |A_\lambda y|\,|h| + |A_\lambda y|\,|f| + |f|\,|h|, \qquad \forall \lambda > 0.$$

(Here, we have assumed that A is symmetric; the general case follows by Theorem 2.11.) We get the estimate

$$|A_\lambda y| \le C(1+|f|), \qquad \forall \lambda > 0,$$

where C is independent of λ and f. This implies that $y \in D(A_H)$ and estimate (2.106) holds. □

Corollary 2.11. *In Corollary* 2.10*, assume in addition that* $f \in H$ *and*

$$(I+\lambda A_H)^{-1}(y+\lambda h) \in K \quad \textit{for some } h \in H \textit{ and all } \lambda > 0. \tag{2.107}$$

Then, the solution y *to variational inequality* (2.94) *belongs to* $D(A_H)$*, and the following estimate holds,*

$$|Ay| \le C(1+|f|), \qquad \forall f \in H. \tag{2.108}$$

The Obstacle Problem

Throughout this section, Ω is an open and bounded subset of the Euclidean space $\mathbf{R}^N$ with a smooth boundary $\partial\Omega$. In fact, we assume that $\partial\Omega$ is of class C^2. However, if Ω is convex, this regularity condition on $\partial\Omega$ is no longer necessary.

Let $V = H^1(\Omega)$, $H = L^2(\Omega)$, and $A : V \to V'$ be defined by

$$(z,Ay) = a(y,z) = \sum_{i=1}^{N} \int_{\Omega} a_{ij}(x) y_{x_i}(x) z_{x_j}(x)dx + \int_{\Omega} a_0(x)y(x)z(x)dx \\ + \frac{\alpha_1}{\alpha_2} \int_{\partial\Omega} y(x)z(x)d\sigma_x, \qquad \forall y,z \in V, \tag{2.109}$$

where α_1, α_2 are two nonnegative constants such that $\alpha_1 + \alpha_2 > 0$. If $\alpha_2 = 0$, we take $V = H_0^1(\Omega)$ and $A : H_0^1(\Omega) \to H^{-1}(\Omega)$ is defined by

$$(z,Ay) = a(y,z) = \sum_{i=1}^{N} \int_{\Omega} a_{ij}(x) y_{x_i} z_{x_j}(x)dx \\ + \int_{\Omega} a_0(x)y(x)z(x)dx, \qquad \forall y,z \in H_0^1(\Omega). \tag{2.110}$$

Here, $a_0, a_{ij} \in L^\infty(\Omega)$ for all $i,j = 1,...,N$, $a_{ij} = a_{ji}$, and

$$a_0(x) \geq 0, \qquad \sum_{i,j=1}^{N} a_{ij}(x)\xi_i\xi_j \geq \omega\|\xi\|_N^2, \qquad \forall \xi \in \mathbf{R}^N,\ x \in \Omega, \tag{2.111}$$

where ω is some positive constant and $\|\cdot\|_N$ is the Euclidean norm in $\mathbf{R}^N$.

If $\alpha_1 = 0$, we assume that $a_0(x) \geq \rho > 0$, a.e. $x \in \Omega$.

The reader will recognize, of course, in the operator defined by (2.109) the second order elliptic operator

$$A_0 y = -\sum_{i,j=1}^{N} (a_{ij} y_{x_i})_{x_j} + a_0 y \tag{2.112}$$

with the boundary value conditions

$$\alpha_1 y + \alpha_2 \frac{\partial y}{\partial \nu} = 0 \quad \text{on } \partial\Omega, \tag{2.113}$$

where $\partial/\partial\nu$ is the conormal derivative,

$$\frac{\partial y}{\partial \nu} = \sum_{i,j=1}^{N} a_{ij} y_{x_j} \cos(\nu, e_i). \tag{2.114}$$

Similarly, the operator A defined by (2.110) is the differential operator (2.112) with the Dirichlet homogeneous conditions: $y = 0$ on $\partial\Omega$.

Let $\psi \in H^2(\Omega)$ be a given function and let K be the closed convex subset of $V = H^1(\Omega)$ defined by

$$K = \{y \in V;\ y(x) \geq \psi(x),\ \text{a.e. } x \in \Omega\}. \tag{2.115}$$

Note that $K \neq \emptyset$ because $\psi^+ = \max(\psi, 0) \in K$. If $V = H_0^1(\Omega)$, we assume that $\psi(x) \leq 0$, a.e. $x \in \partial\Omega$, which implies as before that $K \neq \emptyset$.

Let $f \in V'$. Then, by Corollary 2.10, the variational inequality

$$a(y, y-z) \le (y-z, f), \qquad \forall z \in K \tag{2.116}$$

has a unique solution $y \in K$.

Formally, y is the solution to the following boundary value problem known in the literature as the *obstacle problem*,

$$\begin{cases} A_0 y = f \quad \text{in } \Omega^+ = \{x \in \Omega;\ y(x) > \psi(x)\}, \\ A_0 y \ge f, \qquad y \ge \psi \quad \text{in } \Omega, \\ y = \psi \quad \text{in } \Omega \setminus \Omega^+, \qquad \dfrac{\partial y}{\partial \mu} = \dfrac{\partial \psi}{\partial \mu} \quad \text{on } \partial\Omega^+ = S, \end{cases} \tag{2.117}$$

$$\alpha_1 y + \alpha_2 \frac{\partial y}{\partial \nu} = 0 \quad \text{on } \partial\Omega, \tag{2.118}$$

where μ is the conormal to $\partial\Omega^+$.

Indeed, if $\psi \in C(\overline{\Omega})$ and y is a sufficiently smooth solution, then Ω^+ is an open subset of Ω and so, for every $\varphi \in C_0^\infty(\Omega^+)$ there is $\rho > 0$ such that $y \pm \rho\varphi \ge \psi$ on Ω (i.e., $y \pm \rho\varphi \in K$). Then, if we take $z = y \pm \rho\varphi$ in (2.116), we see that

$$\sum_{i,j=1}^N \int_\Omega a_{ij} y_{x_i} \varphi_{x_j} dx + \int_\Omega a_0 y \varphi \, dx = (f, \varphi), \qquad \forall \varphi \in C_0^\infty(\Omega^+).$$

Hence, $A_0 y = f$ in $\mathscr{D}'(\Omega^+)$.

Now, if we take $z = y + \varphi$, where $\varphi \in H^1(\Omega)$ and $\varphi \ge 0$, we get

$$\sum_{i,j=1}^N \int_\Omega a_{ij} y_{x_i} \varphi_{x_j} dx + \int_\Omega a_0 y \varphi \, dx \ge (f, \varphi),$$

and, therefore, $A_0 y \ge f$ in $\mathscr{D}'(\Omega)$.

The boundary conditions (2.118) are obviously incorporated into the definition of the operator A if $\alpha_2 = 0$. If $\alpha_2 > 0$, then the boundary conditions (2.118) follow from the inequality (2.116) if $\alpha_1 + \alpha_2(\partial\psi/\partial\nu) \le 0$, a.e. on $\partial\Omega$ (see Theorem 2.13 following). As for the equation

$$\frac{\partial y}{\partial \mu} = \frac{\partial \psi}{\partial \mu} \quad \text{on } \partial\Omega^+,$$

this is a transmission property that is implied by the conditions $y \ge \psi$ in Ω and $y = \psi$ in $\partial\Omega^+$, if y is smooth enough.

In the problem (2.117) and(2.118), the surface $\partial\Omega^+ = S$ that separates the domains Ω^+ and $\Omega \setminus \overline{\Omega}^+$ is not known a priori and is called the *free boundary*. In classical terms, this problem can be reformulated as follows. Find the free boundary S and the function y that satisfy the system

$$\begin{cases} A_0 y = f & \text{in } \Omega^+, \\ y = \psi & \text{in } \Omega \setminus \Omega^+, \\ \dfrac{\partial y}{\partial \mu} = \dfrac{\partial \psi}{\partial \mu} & \text{on } S, \\ \alpha_1 + \alpha_2 \dfrac{\partial u}{\partial \nu} = 0 & \text{on } \partial\Omega. \end{cases}$$

In the variational formulation (2.116), the free boundary S does not appear explicitly but the unknown function y satisfies a nonlinear equation. Once y is known, the free boundary S can be found as the boundary of the coincidence set $\{x \in \Omega;\ y(x) = \psi(x)\}$.

There exists an extensive literature on the regularity properties of the solution to the obstacle problem and of the free boundary. We mention in this context the earlier work of Brezis and Stampacchia [6], Brezis [3], and the books of Kinderlehrer and Stampacchia [17] and Friedman [14], which contain complete references on the subject. Here, we present only a partial result.

Proposition 2.11. *Assume that $a_{ij} \in C^1(\overline{\Omega})$, $a_0 \in L^\infty(\Omega)$, and that conditions* (2.111) *hold. Furthermore, assume that $\psi \in H^2(\Omega)$ and*

$$\alpha_1 \psi + \alpha_2 \frac{\partial \psi}{\partial \nu} \le 0, \quad \text{a.e. on } \partial\Omega. \tag{2.119}$$

Then, for every $f \in L^2(\Omega)$, the solution y to variational inequality (2.116) *belongs to $H^2(\Omega)$ and satisfies the complementary system*

$$\begin{cases} (A_0 y(x) - f(x))(y(x) - \psi(x)) = 0, & \text{a.e. } x \in \Omega,\ y(x) \ge \psi(x), \\ A_0 y(x) \ge f(x), & \text{a.e. } x \in \Omega, \end{cases} \tag{2.120}$$

along with the boundary value conditions

$$\alpha_1 y + \alpha_2 \frac{\partial y}{\partial \nu}(x) = 0, \quad \text{a.e. } x \in \partial\Omega. \tag{2.121}$$

Moreover, there exists a positive constant C independent of f such that

$$\|y\|_{H^2(\Omega)} \le C(\|f\|_{L^2(\Omega)} + 1). \tag{2.122}$$

Proof. We apply Corollary 2.11, where $H = L^2(\Omega)$, $V = H^1(\Omega)$ (respectively, $V = H_0^1(\Omega)$ if $\alpha_2 = 0$), A is defined by (2.109) (respectively, (2.110)), and K is given by (2.115).

Clearly, the operator $A_H : L^2(\Omega) \to L^2(\Omega)$ is defined in this case by

$$\begin{cases} (A_H y)(x) = (A_0 y)(x), \text{ a.e. } x \in \Omega,\ y \in D(A_H), \\ D(A_H) = \left\{ y \in H^2(\Omega);\ \alpha_1 y + \alpha_2 \dfrac{\partial y}{\partial \nu} = 0, \text{ a.e. on } \partial\Omega \right\}. \end{cases}$$

We shall verify condition (2.107) with $h = A_0\psi$. To this end, consider for $\lambda > 0$ the boundary value problem

$$\begin{cases} w + \lambda A_0 w = y + \lambda A_0 \psi & \text{in } \Omega, \\ \alpha_1 w + \alpha_2 \dfrac{\partial w}{\partial \nu} = 0 & \text{on } \partial\Omega, \end{cases}$$

which has a unique solution $w \in D(A_H)$. (See Theorems 1.10 and 1.12.)

Multiplying this equation by $(w-\psi)^- \in H^1(\Omega)$ and integrating on Ω, we get, via Green's formula,

$$\begin{aligned} &\int_\Omega |(w-\psi)^-|^2 dx + \lambda a((w-\psi)^-, (w-\psi)^-) \\ &- \frac{\lambda}{\alpha_2} \int_{\partial\Omega} \left(\alpha_1 \psi + \alpha_2 \frac{\partial \psi}{\partial \nu} \right) (w-\psi)^- d\sigma \\ &= \int_\Omega (y-\psi)(w-\psi)^- dx. \end{aligned}$$

Hence, in virtue of (2.119), $(w-\psi)^- = 0$, a.e. in Ω and so $w \in K$, as claimed. Then, by Corollary 2.11, we infer that $y \in D(A_H)$ and

$$\|A_H y\|_{L^2(\Omega)} \le C(\|f\|_{L^2(\Omega)} + 1),$$

and, because $\partial\Omega$ is sufficiently smooth (or Ω convex), this implies (2.122).

Now, if $y \in D(A_H)$, we have

$$a(y,z) = \int_\Omega A_0 y(x) z(x) dx, \qquad \forall z \in H^1(\Omega),$$

and so, by (2.116), we see that

$$\int_\Omega (A_0 y(x) - f(x))(y(x) - z(x)) dx \le 0, \qquad \forall z \in K. \tag{2.123}$$

The last inequality clearly can be extended by density to all $z \in K_0$, where

$$K_0 = \{u \in L^2(\Omega);\ u(x) \ge \psi(x), \quad \text{a.e. } x \in \Omega\}. \tag{2.124}$$

If in (2.123) we take $z = \psi + \alpha$, where α is any positive $L^2(\Omega)$ function, we get

$$(A_0 y)(x) - f(x) \ge 0, \quad \text{a.e. } x \in \Omega.$$

Then, for $z = \psi$, (2.123) yields

$$(y(x) - \psi(x))(A_0 y)(x) - f(x) = 0, \quad \text{a.e. } x \in \Omega,$$

which completes the proof. □

We note that under the assumptions of Theorem 2.14 the obstacle problem can be equivalently written as

$$A_H y + \partial I_{K_0}(y) \ni f, \tag{2.125}$$

where

$$\partial I_{K_0}(y) = \left\{ v \in L^2(\Omega);\ \int_\Omega v(x)(y(x) - z(x))dx \ge 0,\ \forall z \in K_0 \right\}$$

or, equivalently,

$$\partial I_{K_0}(y) = \{v \in L^2(\Omega);\ v(x) \in \beta(y(x) - \psi(x)),\ \text{a.e. } x \in \Omega\},$$

where $\beta : \mathbf{R} \to 2^{\mathbf{R}}$ is the maximal monotone graph,

$$\beta(r) = \begin{cases} 0 & \text{if } r > 0, \\ \mathbf{R}^- & \text{if } r = 0, \\ \emptyset & \text{if } r < 0. \end{cases} \tag{2.126}$$

Hence, under the conditions of Theorem 2.14, we may equivalently write the variational inequality (2.116) as

$$\begin{cases} (A_0 y)(x) + \beta(y(x) - \psi(x)) \ni f(x), & \text{a.e. } x \in \Omega, \\ \alpha_1 y + \alpha_2 \dfrac{\partial y}{\partial \nu} = 0, & \text{a.e. on } \partial\Omega, \end{cases} \tag{2.127}$$

and it is equivalent to the minimization problem

$$\min \left\{ \frac{1}{2} a(y,y) + \int_\Omega j(y(x) - \psi(x))dx - \int_\Omega f(x)y(x)dx;\ y \in L^2(\Omega) \right\},$$

where $j : \mathbf{R} \to \overline{\mathbf{R}}$ is defined by

$$j(r) = \begin{cases} 0 & \text{if } r \ge 0, \\ +\infty & \text{otherwise}. \end{cases} \tag{2.128}$$

A simple physical model for the obstacle problem is that of an elastic membrane that occupies a plane domain Ω and is limited from below by a rigid obstacle ψ while it is under the pressure of a vertical force field of density f. (See, e.g., Barbu [1].) The mathematical model of the water flow through an isotropic homogeneous rectangular dam can be described (by a device due to C. Baiocchi) as an obstacle problem of the above type. We mention in the same context the elastic–plastic problem (Brezis and Stampacchia [6]) or the mathematical model of oxygen diffusion in tissue.

2.4 Nonlinear Elliptic Problems of Divergence Type

We study here the boundary value problem

$$\lambda y - \mathrm{div}_x \beta(\nabla y(x)) \ni f(x), \qquad x \in \Omega, \tag{2.129}$$

$$y = 0 \qquad \text{on } \partial\Omega, \tag{2.130}$$

where Ω is a bounded and open domain of $\mathbf{R}^N$ with smooth boundary $\partial\Omega$, the function f is in $L^2(\Omega)$, and λ is a nonnegative constant. Here, $\beta : \mathbf{R}^N \to 2^{\mathbf{R}^N}$ is a maximal monotone graph in $\mathbf{R}^N \times \mathbf{R}^N$ such that $0 \in \beta(0)$.

Equation (2.129) describes the equilibrium state of diffusion-like processes where the diffusion flux $\mathbf{q}$ is a nonlinear function of the gradient ∇y of local density y. In the special case, where β is a potential function (i.e., $\beta = \nabla j$, $j : \mathbf{R} \to \mathbf{R}$), then the functional $\phi(y) = \int_\Omega j(\nabla y)dx + (\lambda/2)\int_\Omega y^2 dx$ can be viewed as the energy of the system and equation (2.129) describes the critical points of ϕ. The elliptic character of equation (2.129) is given by monotonicity assumption on β.

It should be said that equation (2.129) with boundary condition (2.130) might be highly nonlinear and so the best one can expect from the existence point of view is a weak solution.

Definition 2.3. The function $y \in L^1(\Omega)$ is said to be a weak solution to the Dirichlet problem (2.129) and (2.130) if $y \in W_0^{1,1}(\Omega)$ and there is $\eta \in (L^q(\Omega))^N$, $1 < p < \infty$, such that

$$\eta(x) \in \beta(\nabla y(x)), \quad \text{a.e. } x \in \Omega, \tag{2.131}$$

$$\lambda \int_\Omega y\psi\, dx + \int_\Omega \eta(x) \cdot \nabla\psi(x) dx = \int_\Omega f(x)\psi(x) dx, \qquad \forall \psi \in W_0^{1,p}(\Omega),\ \frac{1}{p} + \frac{1}{q} = 1. \tag{2.132}$$

Similarly, the function y is said to be a weak solution to equation (2.129) with the Neumann boundary value condition

$$\beta(\nabla y(x)) \cdot \nu(x) = 0 \quad \text{on } \partial\Omega \tag{2.133}$$

if $y \in W^{1,1}(\Omega)$ and there is $\eta \in (L^p(\Omega))^N$ which satisfies (2.131), and (2.132) holds for all $\psi \in W^{1,q}(\Omega)$. (Here ν is the normal to $\partial\Omega$.)

The first existence result for problem (2.129) and (2.130) concerns the case where β is single-valued.

Theorem 2.15. *Assume that $\beta : \mathbf{R}^N \to \mathbf{R}^N$ is continuous, monotonically increasing, and*

$$|\beta(r)| \le C_1(1 + |r|^{p-1}), \qquad \forall r \in \mathbf{R}^N, \tag{2.134}$$

$$\beta(r) \cdot r \ge \omega |r|^p - C_2, \qquad \forall r \in \mathbf{R}^N, \tag{2.135}$$

where $\omega > 0,\ p > 1,\ 2N/(N+2) \le p$. *Then, for each* $f \in W^{-1,q}(\Omega)$ *and* $\lambda > 0$, *there is a unique weak solution* $y \in W_0^{1,p}(\Omega)$ *to problem* (2.129) and (2.130).

Proof. We apply Corollary 2.3 to the operator $T : X \to X^*,\ X = W_0^{1,p}(\Omega)$, $X^* = W^{-1,q}(\Omega)$, defined by

$$(v,Tu) = \int_\Omega \beta(\nabla u(x)) \cdot \nabla v(x)dx + \lambda \int_\Omega u(x)v(x)\,dx, \qquad \forall u,v \in X = W_0^{1,p}(\Omega). \tag{2.136}$$

It is easily seen that T is monotone and demicontinuous. Indeed, if $u_j \to u$ strongly in $X = W_0^{1,p}(\Omega)$, then $\nabla u_j \to \nabla u$ strongly in $L^p(\Omega)$ and, by continuity of β, we have on a subsequence $\beta(\nabla u_j) \to \beta(\nabla u)$, a.e. on Ω. On the other hand, by (2.134) we have that $\{\beta(\nabla u_j)\}$ is bounded in $L^q(\Omega)$ and therefore it is weakly sequentially compact in $L^q(\Omega)$. Hence, we also have (eventually, on a subsequence)

$$\beta(\nabla u_j) \rightharpoonup \beta(\nabla u) \quad \text{in } \ (L^q(\Omega))^N.$$

Then, we infer that

$$\lim_{j\to\infty} \int_\Omega \beta(\nabla u_j) \cdot \nabla v\,dx = \int_\Omega \beta(\nabla u) \cdot \nabla v\,dx, \qquad \forall v \in X$$

and also

$$\lim_{j\to\infty} \int_\Omega u_j v\,dx = \int_\Omega uv\,dx,$$

because $W^{1,p}(\Omega) \subset L^2(\Omega)$ by Theorem 1.5. Hence,

$$Tu_j \rightharpoonup Tu \quad \text{in } \ X^* = W^{-1,q}(\Omega).$$

It is also clear by (2.135) that T is coercive; that is,

$$(u,Tu) \ge \omega \int_\Omega |\nabla u|^p dx - C_2, \qquad \forall u \in X.$$

This completes the proof. □

If $\lambda = 0$, we still have a solution $y \in W_0^{1,p}(\Omega)$, but in general it is not unique.

A similar existence result follows for problem (2.129) and (2.133), namely, the following.

Theorem 2.16. *Under the assumptions of Theorem* 2.15, *for each* $f \in (W^{1,p}(\Omega))^*$ *and* $\lambda > 0$ *there is a unique weak solution* $y \in W^{1,p}(\Omega)$ *to problem* (2.129) and (2.133).

Proof. One applies Corollary 2.3 to the operator $T : W^{1,p}(\Omega) \to (W^{1,p}(\Omega))^*$ defined by (2.136) for all $v \in W^{1,p}(\Omega)$.

It follows as in the previous case that T is monotone and demicontinuous. As regards the coercivity, we note that by (2.135) and (2.136) we have

$$(u,Tu) \geq \omega \int_\Omega |\nabla u|^p dx + \lambda \int_\Omega u^2 dx. \tag{2.137}$$

Recalling that (see Remark 1.1)

$$\|u\|_{W^{1,p}(\Omega)} \leq C(\|\nabla u\|_{L^p(\Omega)} + \|u\|_{L^q(\Omega)}), \qquad \forall u \in W^{1,p}(\Omega),$$

for $1 \leq q \leq Np/(N-p)$, $N > p$ and $q \geq 1$ for $N \geq p$, we see, by (2.137), that

$$(u,Tu) \geq \omega \|u\|^\alpha_{W^{1,p}(\Omega)}, \qquad \forall u \in W^{1,p}(\Omega),$$

where $\alpha = \max\{p,2\}$ and therefore T is coercive, as desired. Then Theorem 2.16 follows by Corollary 2.3. □

The above existence results extend to general maximal monotone (multivalued) graphs $\beta \subset \mathbf{R}^N \times \mathbf{R}^N$ satisfying assumptions (2.134) and (2.135); that is,

$$\sup\{|w|;\ w \in \beta(r)\} \leq C_1(1+|r|^{p-1}), \qquad \forall r \in \mathbf{R}^N, \tag{2.138}$$

$$w \cdot r \geq \omega |r|^p - C_2, \qquad \forall (w,r) \in \beta. \tag{2.139}$$

(Here, and everywhere in the following, we denote by $|r|$ the Euclidean norm of $r \in \mathbf{R}^N$.)

Theorem 2.17. *Let β be a maximal monotone graph in $\mathbf{R}^N \times \mathbf{R}^N$ satisfying conditions* (2.138) *and* (2.139) *for $\omega > 0$, and $p > 1$. Then, for each $f \in L^2(\Omega)$ and $\lambda > 0$ there is a unique weak solution $y \in W_0^{1,p}(\Omega)$ to problem* (2.129) *and* (2.130) (*respectively a unique weak solution $y \in W^{1,p}(\Omega)$ to problem* (2.129) *and* (2.133)) *in the following sense*

$$\lambda \int_\Omega y\psi + \int_\Omega \eta \cdot \nabla \psi\, dx = \int_\Omega f\psi\, dx,\ \forall \psi \in W_0^{1,p}(\Omega) \cap L^2(\Omega) \\ (\textit{respectively, } \forall \psi \in W^{1,p}(\Omega) \cap L^2(\Omega)), \tag{2.140}$$

where $\eta \in \beta(\nabla y)$, a.e. in Ω.

Of course, if p is such that $W^{1,p}(\Omega) \subset L^2(\Omega)$ (for instance if $p \geq (2N/(N+2))$, then (2.140) coincides with (2.132).

Proof. We prove the existence theorem in the case of problem (2.129) and (2.130) only, the other case (i.e., the Neumann boundary condition (2.133)) being completely similar. We first assume that $f \in W^{-1,q}(\Omega) \cap L^2(\Omega)$. We introduce the Yosida approximation of β

$$\beta_\varepsilon(r) = \frac{1}{\varepsilon}(r - ((1+\varepsilon\beta)^{-1}r) \in \beta((1+\varepsilon\beta)^{-1})), \qquad \forall r \in \mathbf{R}^N,\ \varepsilon > 0, \tag{2.141}$$

and consider the approximating problem

$$\begin{cases} \lambda y_\varepsilon - \operatorname{div}(\beta_\varepsilon(\nabla y_\varepsilon) + \varepsilon \nabla y_\varepsilon) = f & \text{in } \Omega, \\ \beta_\varepsilon(\nabla y_\varepsilon) = 0 & \text{on } \partial\Omega, \end{cases} \tag{2.142}$$

which, by Theorem 2.15 has a unique solution $y_\varepsilon \in H_0^1(\Omega)$. Indeed, β_ε is Lipschitz and it is readily seen that conditions (2.138) and (2.139) hold with $p = 2$ (with constants C independent of ε). On the other hand, by (2.138)and (2.139), we see that

$$\begin{aligned} |\beta_\varepsilon(r)| &< \sup\{|w|;\ w \in \beta((1+\varepsilon\beta)^{-1}r)\} \le C_1(|(1+\varepsilon\beta)^{-1}r|^{p-1}+1) \\ &\le C_3(|r|^{p-1}+1), \qquad \forall r \in \mathbf{R}^N,\ \forall \varepsilon > 0, \end{aligned} \tag{2.143}$$

and

$$\begin{aligned} \beta_\varepsilon(r)\cdot r &= \beta_\varepsilon(r)\cdot(1+\varepsilon\beta)^{-1}r + \varepsilon|\beta_\varepsilon(r)|^2 \\ &\ge \varepsilon|(1+\varepsilon\beta)^{-1}r|^p + C_4(\varepsilon|r|^{p-1}+1) \\ &\ge \omega|r|^p + C_5\varepsilon|r|^p + C_6, \qquad \forall r \in \mathbf{R}^N,\ \varepsilon > 0. \end{aligned} \tag{2.144}$$

(The constants C_i arising in (2.143) and (2.144) are independent of ε.)

We have therefore

$$\lambda \int_\Omega y_\varepsilon \psi\, dx + \int_\Omega (\beta_\varepsilon(\nabla y_\varepsilon) + \varepsilon\nabla y_\varepsilon)\cdot\nabla\psi\, dx = \int_\Omega f\psi\, dx, \qquad \forall \psi \in H_0^1(\Omega), \tag{2.145}$$

and so, for $\psi = y_\varepsilon$, we obtain that

$$\begin{aligned} \lambda \int_\Omega y_\varepsilon^2 dx + \varepsilon \int_\Omega |\nabla y_\varepsilon|^2 dx + \int_\Omega \beta_\varepsilon(\nabla y_\varepsilon)\cdot(1+\varepsilon\beta)^{-1}\nabla y_\varepsilon dx \\ + \varepsilon \int_\Omega |\beta_\varepsilon(\nabla y_\varepsilon)|^2 dx = \int_\Omega f y_\varepsilon dx. \end{aligned} \tag{2.146}$$

Taking into account that $\beta_\varepsilon(\nabla y_\varepsilon) \in \beta((1+\varepsilon\beta)^{-1}\nabla y_\varepsilon)$, it follows by (2.144) and (2.146) that

$$\begin{aligned} \lambda \int_\Omega y_\varepsilon^2 dx + \varepsilon \int_\Omega |\nabla y_\varepsilon|^2 dx + \omega \int_\Omega |(1+\varepsilon\beta)^{-1}\nabla y_\varepsilon|^p dx \\ + \varepsilon \int_\Omega |\beta_\varepsilon(\nabla y_\varepsilon)|^2 dx \le C \int_\Omega |f|^2 dx, \qquad \forall \varepsilon > 0. \end{aligned} \tag{2.147}$$

(Here and everywhere in the sequel, C is a positive constant independent of ε.) In particular, it follows by (2.147) that

$$\int_\Omega |(1+\varepsilon\beta)^{-1}\nabla y_\varepsilon - \nabla y_\varepsilon|^2 dx \to 0 \quad \text{as } \varepsilon \to 0, \tag{2.148}$$

because $\varepsilon^2|\beta_\varepsilon(r)|^2 = |(1+\varepsilon\beta)^{-1}r - r|^2$, $\forall r \in \mathbf{R}^N$.

Moreover, by (2.141), (2.143) we see that

$$\|\beta_\varepsilon(\nabla y_\varepsilon)\|_{L^q(\Omega)} \le C(\|(1+\varepsilon\beta)^{-1}\nabla y_\varepsilon\|^p_{L^p(\Omega)} + 1).$$

Then, on a subsequence again denoted ε, we have by (2.147) and (2.148),

$$y_\varepsilon \to y \qquad \text{weakly in } L^2(\Omega)\cap W^{1,p}(\Omega), \tag{2.149}$$

$$(1+\varepsilon\beta)^{-1}\nabla y_\varepsilon \to \nabla y \quad \text{weakly in } (L^p(\Omega))^N, \tag{2.150}$$

$$\beta_\varepsilon(\nabla y_\varepsilon) \to \eta \qquad \text{weakly in } (L^q(\Omega))^N, \tag{2.151}$$

as $\varepsilon \to 0$. Taking into account (2.145) and (2.150), (2.151), we obtain by the weak semicontinuity of the L^p-norm that

$$\begin{aligned} &\lambda \int_\Omega |\nabla y|^2 dx + \int_\Omega |\nabla y|^p dx \le C \int_\Omega f^2 dx, \quad \text{and} \\ &\lambda \int_\Omega y\psi\, dx + \int_\Omega \eta\cdot\nabla\psi\, dx = \int_\Omega f\psi\, dx, \qquad \forall \psi \in W_0^{1,p}(\Omega)\cap L^2(\Omega). \end{aligned} \tag{2.152}$$

Because $f \in W^{-1,q}(\Omega)\cap L^2(\Omega)$, the latter extends to all of $\psi \in W_0^{1,p}(\Omega)$. To complete the proof, it suffices to show that

$$\eta(x) \in \beta(\nabla y(x)), \quad \text{a.e. } x \in \Omega. \tag{2.153}$$

To this end, we start with the obvious inequality

$$\int_\Omega (\beta_\varepsilon(\nabla y_\varepsilon) - \zeta)\cdot((1+\varepsilon\beta)^{-1}\nabla y_\varepsilon - u)dx \ge 0, \tag{2.154}$$

for all $u \in L^p(\Omega)$ and $\zeta \in (L^q(\Omega))^N$ such that $\zeta(x) \in \beta(u(x))$, a.e. $x \in \Omega$. (This is an immediate consequence of monotonicity of β because, by (2.141), $\beta_\varepsilon(y) \in \beta((1+\varepsilon\beta)^{-1}y)$, $\forall y \in \mathbf{R}^N$, $\forall \varepsilon > 0$.)

Letting ε tend to zero in (2.154), we obtain that

$$\int_\Omega (\eta - \zeta)\cdot(y-u)dx \ge 0.$$

Now, choosing $u = (1+\beta)^{-1}(\eta+y)$ and $\zeta = \eta - y + u \in \beta(u)$, we obtain that

$$\int_\Omega (y-u)^2 dx = 0.$$

Hence, $y = u$ and $\eta = \zeta \in \beta(u)$, a.e. in Ω. This completes the proof of existence for $f \in W^{-1,q}(\Omega)\cap L^2(\Omega)$.

If $f \in L^2(\Omega)$, consider a sequence $\{f_n\} \subset W^{-1,q}(\Omega)\cap L^2(\Omega)$ strongly convergent to f in $L^2(\Omega)$. If y_n are corresponding solutions to problem (2.140), we obtain, by monotonicity of β,

$$\lambda \int_\Omega |y_n - y_m|^2 dx \le \|f_n - f_m\|_{W^{-1,q}(\Omega)} \|y_n - y_m\|_{W_0^{1,p}(\Omega)},$$

whereas, by estimate (2.152), we see that $\{y_n\}$ is bounded in $W_0^{1,p}(\Omega)$. Hence, on a subsequence, we have

$$\begin{aligned} y_n &\to y &&\text{strongly in } L^2(\Omega) \text{ and weakly in } W_0^{1,p}(\Omega)\\ \eta_n \in \beta(\nabla y_n) &\to \eta &&\text{weakly in } (L^q(\Omega))^N. \end{aligned}$$

Clearly, (y,η) verify (2.140) and arguing as above it follows also $\eta \in \beta(\nabla y)$, a.e. in Ω. This completes the proof of existence. The uniqueness is immediate by the monotonicity of β. □

We have chosen β multivalued not only for the sake of generality, but because this case arises naturally in specific problems. For instance, if β is the subdifferential ∂j of a lower semicontinuous convex function that is not differentiable, then β is necessarily multivalued and this situation occurs, for instance, in the description of stationary (equilibrium) states of systems with nondifferentiable energy.

Define the operator $A : D(A) \subset L^2(\Omega) \to L^2(\Omega)$,

$$\begin{cases} D(A) = \{y \in W_0^{1,p}(\Omega);\ \exists \eta \in (L^q(\Omega))^N;\ \eta(x) \in \beta(\nabla y(x)), \\ \qquad\qquad\qquad\qquad \text{a.e. } x \in \Omega,\ \operatorname{div}\eta \in L^2(\Omega)\}, \\ Ay = \{-\operatorname{div}\eta\}, \qquad \forall y \in D(A). \end{cases} \tag{2.155}$$

If β is single-valued, then A can be simply represented as

$$\begin{cases} Ay = -\operatorname{div}\beta(\nabla y), \qquad \forall y \in D(A) \\ D(A) = \{y \in W_0^{1,p}(\Omega);\ \operatorname{div}\beta(\nabla y) \in L^2(\Omega)\}. \end{cases} \tag{2.156}$$

We have the following theorem.

Theorem 2.18. *The operator A is maximal monotone in $L^2(\Omega) \times L^2(\Omega)$. Moreover, if $\beta = \partial j$, where $j : \mathbf{R}^N \to \mathbf{R}$ is a continuous convex function, then $A = \partial\varphi$, and $\varphi : L^2(\Omega) \to \overline{\mathbf{R}}$ (the energy function), is given by*

$$\varphi(y) = \begin{cases} \displaystyle\int_\Omega j(\nabla y)dx & \text{if } y \in W_0^{1,p}(\Omega) \text{ and } j(\nabla y) \in L^1(\Omega) \\ +\infty & \text{otherwise.} \end{cases} \tag{2.157}$$

Proof. Because (2.156) is taken in the sense of distributions on Ω, we have

$$(Ay,\psi) = \int_\Omega \beta(\nabla y)\cdot\nabla\psi\, dx, \qquad \forall \psi \in L^2(\Omega) \cap W_0^{1,p}(\Omega). \tag{2.158}$$

(Here $(\cdot,\cdot)$ is the duality defined by the scalar product of $L^2(\Omega)$.) This yields, of course,

$$(Ay - Az, y - z) \ge 0, \qquad \forall y,z \in W_0^{1,p}(\Omega) \cap L^2(\Omega)$$

and, by density, the latter extends to all $y,z \in D(A)$. Hence A is monotone.

To prove the maximal monotonicity, consider the equation

$$\lambda y + Ay \ni f, \tag{2.159}$$

where $\lambda > 0$ and $f \in L^2(\Omega)$. Taking into account (2.158), we rewrite (2.159) as

$$\lambda \int_\Omega y\psi + \int_\Omega \eta \cdot \nabla\psi\, dx = \int_\Omega f\psi\, dx, \qquad \forall \psi \in W_0^{1,p}(\Omega) \cap L^2(\Omega), \tag{2.160}$$

where $\eta \in (L^q(\Omega))^N,\ \eta(x) \in \beta(\nabla y(x))$, a.e. $x \in \Omega$.

On the other hand, by Theorem 2.17, there is a solution y to (2.140) and therefore to (2.159), because by (2.158) it also follows that

$$\operatorname{div} \eta(\psi) = -\int_\Omega f\psi + \lambda \int_\Omega fy \le C\|\psi\|_{L^2(\Omega)}, \qquad \forall \psi \in W_0^{1,p}(\Omega) \cap L^2(\Omega)$$

and, therefore, $\operatorname{div}\eta \in L^2(\Omega)$. Hence A is maximal monotone.

Now, if β is a subgradient maximal monotone graph of the form ∂j, it is easily seen that $A \subset \partial\varphi$; that is,

$$\varphi(y) - \varphi(z) \le \int_\Omega \eta(y-z)dx, \qquad \forall \eta \in Ay,\ y,z \in L^2(\Omega).$$

Because A is maximal in the class of monotone operators, we have therefore $A = \partial\varphi$, as claimed. □

It turns out that in the special case, where $\beta = \partial j$, assumptions (2.138) and (2.139) can be weakened to

(i) j is convex, continuous, $\inf j = j(0) = 0$.

$$\lim_{|r|\to\infty} \frac{j(r)}{|r|} = \lim_{|p|\to\infty} \frac{j^*(p)}{|p|} = +\infty. \tag{2.161}$$

$$\lim_{|r|\to\infty} \frac{j(-r)}{j(r)} < \infty. \tag{2.162}$$

Here j^* is the conjugate of j; that is, $j^*(p) = \sup\{(p\cdot u) - j(u);\ u \in \mathbf{R}^N\}$. By $|\cdot|$ we denote here the Euclidean norm in $\mathbf{R}^N$.

We come back to boundary value problem (2.129) and (2.133) in the more general context (2.161) and (2.162) which assume minimal growth conditions on β or j.

Theorem 2.19. *Under assumptions* (2.161) *and* (2.162)*, problem* (2.129) *and* (2.133) *has, for each $\lambda > 0$ and $f \in L^2(\Omega)$, a unique weak solution $y^* \in W^{1,1}(\Omega)$ in the following sense*

$$\begin{cases} \int_\Omega (\lambda yv + \eta\cdot\nabla v)dx = \int_\Omega fv\,dx, & \forall v \in C^1(\overline{\Omega}) \\ \eta \in (L^1(\Omega))^N,\ \eta(x) \in \beta(\nabla y(x)), & a.e.\ x \in \Omega \\ j^*(\eta) \in L^1(\Omega),\ j(\nabla y) \in L^1(\Omega). \end{cases} \tag{2.163}$$

Moreover, y^* is the unique minimizer of problem

$$\min\left\{\frac{\lambda}{2}\int_\Omega |y(x)-\frac{1}{\lambda}f(x)|^2dx+\int_\Omega j(\nabla y(x))dx;\ y\in W^{1,1}(\Omega)\right\}. \tag{2.164}$$

Proof. We assume for simplicity $\lambda = 1$. The existence of a unique minimizer u^* for problem (2.164) is an immediate consequence of Proposition 1.4 and of the fact that, under the first of conditions (2.161), the convex function

$$\varphi : L^2(\Omega) \to \overline{\mathbf{R}} = (-\infty,+\infty],\ \varphi(u) = \int_\Omega j(\nabla u(x))dx + \frac{1}{2}\int_\Omega (u-f)^2 dx$$

is weakly lower semicontinuous in the space $L^2(\Omega)$. Indeed, by the same argument as that used in the proof of Proposition 2.11, it follows by (2.161) that the set $\mathscr{M} = \{y \in W^{1,1}(\Omega);\ \varphi(y) \le \lambda\}$ is bounded in $W^{1,1}(\Omega)$; that is,

$$|\nabla y|_{(L^1(\Omega))^N} \le C \qquad \forall y \in \mathscr{M}$$

and

$$\left\{\int_E |\nabla y(x)|dx;\ E \subset \Omega,\ u \in \mathscr{M}\right\}$$

is uniformly absolutely continuous and so, by the Dunford–Pettis theorem (Theorem 1.15) $\mathscr{M}$ is weakly compact in $W^{1,1}(\Omega)$. Hence, if $\{y_n\} \subset \mathscr{M}$ is weakly convergent to y in $L^2(\Omega)$, it follows that $\nabla y_n \to \nabla y$ weakly in $(L^1(\Omega))^N$ and because the convex integrand $v \to \int_\Omega j(v)$ is weakly lower semicontinuous in $(L^1(\Omega))^n$ (because by Proposition 2.10 it is lower semicontinuous in $(L^1(\Omega))^n$), we infer that $y \in \mathscr{M}$. Hence $\mathscr{M}$ is closed in $L^2(\Omega)$ as claimed.

In order to prove that the minimizer y^* is a solution to (2.163), we start with the approximating equation

$$\text{Min}\left\{\int_\Omega \left(j_\varepsilon(\nabla y) + \frac{\varepsilon}{2}|\nabla y(x)|^2 + \frac{1}{2}|y-f|^2\right)dx;\ y \in H^1(\Omega)\right\}, \tag{2.165}$$

where $j_\varepsilon \in C^1(\mathbf{R}^N)$ is the function (see (2.38)),

$$j_\varepsilon(p) = \inf\left\{\frac{1}{2\varepsilon}|v-p|^2 + j(v);\ v \in \mathbf{R}^N\right\}.$$

Problem (2.165) has a unique solution $y_\varepsilon \in H^1(\Omega)$ which, as easily seen, satisfies the elliptic boundary value problem

$$\begin{aligned} y_\varepsilon - \varepsilon\Delta y_\varepsilon - \mathrm{div}_x(\partial j_\varepsilon(\nabla y_\varepsilon)) &= f \quad \text{in } \Omega,\\ (\varepsilon\nabla y_\varepsilon + \partial j_\varepsilon(\nabla y_\varepsilon))\cdot \nu &= 0 \quad \text{on } \partial\Omega. \end{aligned} \tag{2.166}$$

Equivalently,

$$\int_\Omega ((\varepsilon \nabla y_\varepsilon + \partial j_\varepsilon(\nabla y_\varepsilon)) \cdot \nabla v + y_\varepsilon v) dx = \int_\Omega f v\, dx, \qquad \forall v \in H^1(\Omega). \tag{2.167}$$

(The Gâteaux differential of the function arising in (2.165) is just the operator from the left-hand side of (2.166) or (2.167).) We recall that (see Theorem 2.9)

$$\begin{aligned} \partial j_\varepsilon(p) &= \frac{1}{\varepsilon}(p - (1+\varepsilon\beta)^{-1}p) \in \beta((1+\varepsilon\beta)^{-1}p), \qquad \forall p \in \mathbf{R}^N, \\ j_\varepsilon(p) &= \frac{1}{2\varepsilon}|p - (1+\varepsilon\beta)^{-1}p|^2 + j((1+\varepsilon\beta)^{-1}p). \end{aligned}$$

Then, it is readily seen by (2.165) that on a subsequence, again denoted $\{\varepsilon\} \to 0$, we have

$$\begin{aligned} y_\varepsilon &\to y^* && \text{weakly in } L^2(\Omega), \\ ((1+\varepsilon\beta)^{-1}\nabla y_\varepsilon - \nabla y_\varepsilon) &\to 0 && \text{strongly in } L^2(\Omega;\mathbf{R}^N), \\ (1+\varepsilon\beta)^{-1}\nabla y_\varepsilon &\to \nabla y^* && \text{weakly in } L^1(\Omega;\mathbf{R}^N). \end{aligned} \tag{2.168}$$

The latter follows by the obvious inequality

$$\begin{aligned} &\int_\Omega \left(j((1+\varepsilon\beta)^{-1}\nabla y_\varepsilon) + \frac{1}{2\varepsilon}|\nabla y_\varepsilon - (1+\varepsilon\beta)^{-1}\nabla y_\varepsilon|^2 + \frac{\varepsilon}{2}|\nabla y_\varepsilon|^2 + \frac{1}{2}|y_\varepsilon - f|^2 \right) dx \\ &\le \int_\Omega \left(j(\nabla y_\varepsilon) + \frac{\varepsilon}{2}|\nabla y_\varepsilon|^2 + \frac{1}{2}|y_\varepsilon - f|^2 \right) dx \\ &\le \int_\Omega \left(j_\varepsilon(\nabla v) + \frac{\varepsilon}{2}|\nabla v|^2 + \frac{1}{2}|v - f|^2 \right) dx, \qquad \forall v \in H^1(\Omega). \end{aligned} \tag{2.169}$$

On the other hand, by (2.169) and the first condition in (2.161), it follows via the Dunford–Pettis theorem (Theorem 1.15) that $\{(1+\varepsilon\beta)^{-1}\nabla y_\varepsilon\}$ is weakly compact in $L^1(\Omega;\mathbf{R}^N) = (L^1(\Omega))^N$ and so (2.168) follows. Then, taking into account the weak lower semicontinuity of functional φ in $L^1(\Omega;\mathbf{R}^N)$, we see that

$$\int_\Omega \left(j(\nabla y^*) + \frac{1}{2}|y^* - f|^2 \right) dx \le \int_\Omega \left(j(\nabla v) + \frac{1}{2}|v - f|^2 \right) dx, \qquad \forall v \in W^{1,1}(\Omega);$$

that is, y^* is optimal in problem (2.164).

Now, we recall the conjugacy inequality (see Proposition 1.5)

$$j(v) + j^*(p) \ge v \cdot p, \qquad \forall v, p \in \mathbf{R}^N$$

with equality if and only if $p \in \beta(v) = \partial j(v)$. This yields

$$\begin{aligned} &\int_\Omega (j((1+\varepsilon\beta)^{-1}\nabla y_\varepsilon) + j^*(\partial j(\nabla y_\varepsilon)))dx \ge \int_\Omega (1+\varepsilon\beta)^{-1}\nabla y_\varepsilon \cdot \partial j(\nabla y_\varepsilon) dx \\ &= \int_\Omega \nabla y_\varepsilon \cdot \partial j(\nabla y_\varepsilon) dx - \frac{1}{\varepsilon}\int_\Omega |\partial j_\varepsilon(\nabla y_\varepsilon)|^2 dx. \end{aligned}$$

(Here, $\partial j(\nabla y_\varepsilon)$ is any section of $\partial j(\nabla y_\varepsilon)$.)

Then, by (2.169), we see that $\{\int_\Omega j^*(\partial j(\nabla y_\varepsilon))dx\}$ is bounded and so, again by the second condition in (2.161) and by the Dunford–Pettis theorem, we infer that $\{\int_E \partial j(y_\varepsilon);\ E \subset \Omega\}$ is uniformly absolutely continuous and therefore $\{\partial j(\nabla y_\varepsilon)\}$ is weakly compact in $(L^1(\Omega))^N$. (Here, one uses the same argument as in the proof of Proposition 2.10; that is, write for each measurable set $E \subset \Omega$,

$$\begin{aligned}\int_E |\partial j(\nabla y_\varepsilon)|dx \le &\int_{E\cap[|\partial j(\nabla y_\varepsilon)|\ge R]} |\partial j(\nabla y_\varepsilon)|dx\\ &+\int_{E\cap[|\partial j(\nabla y_\varepsilon)\le R]} |\partial j(\nabla y_\varepsilon)|dx \le \eta,\end{aligned}$$

for $m(E) \le \delta(\eta)$.) Hence, we may assume that for $\varepsilon \to 0$,

$$\partial j(\nabla y_\varepsilon) \to \eta \quad \text{weakly in } (L^1(\Omega))^N,$$

where η satisfies

$$\int_\Omega (y^* v + \nabla v \cdot \eta)dx = \int_\Omega f v\, dx, \qquad \forall v \in C^1(\overline{\Omega}). \tag{2.170}$$

To conclude the proof, it remains to be shown that

$$\eta(x) \in \beta(\nabla y^*(x)), \quad \text{a.e. } x \in \Omega. \tag{2.171}$$

To this aim, we notice that, in virtue of (2.168) and the conjugacy equality, it follows by the weak lower semicontinuity of the convex integrand in $L^1(\Omega)$,

$$\begin{aligned}\int_\Omega (j(\nabla y^*) + j(\eta))dx &\le \liminf_{\varepsilon\to 0} \int_\Omega (1+\varepsilon\beta)^{-1}\nabla y_\varepsilon \cdot \partial j(\nabla y_\varepsilon)dx\\ &\le \liminf_{\varepsilon\to 0} \int_\Omega \nabla y_\varepsilon \cdot \partial j(\nabla y_\varepsilon)dx.\end{aligned} \tag{2.172}$$

On the other hand, by (2.167) and (2.169), we see that

$$\lim_{\varepsilon\to 0} \int_\Omega \nabla y_\varepsilon \cdot \partial j(\nabla y_\varepsilon)dx = -\int_\Omega (y^* - f)y^* dx. \tag{2.173}$$

We have also that

$$\begin{aligned}\nabla y^* \cdot \eta &\le j(\nabla y^*) + j^*(\eta), && \text{a.e. in } \Omega\\ -\nabla y^* \cdot \eta &\le j(-\nabla y^*) + j^*(\eta) \le Cj(\nabla y^*) + j(\eta), && \text{a.e. in } \Omega.\end{aligned}$$

(The second inequality follows by the convexity of j^*.) Hence, $\nabla y^* \cdot \eta \in L^1(\Omega)$ and so, by (2.170), (2.172) and (2.173), we see that

$$\int_\Omega (j(\nabla y^*) + j^*(\eta) - \nabla y^* \cdot \eta)dx \le 0,$$

because (2.170) extends by density to all $v \in W^{1,1}(\Omega)$ such that $\nabla v \cdot \eta \in L^1(\Omega)$. Recalling that $j^*(\nabla y^*) + j(\eta) - \nabla y^* \cdot \eta \geq 0$, a.e. in Ω, we infer that

$$j(\nabla y^*(x)) + j^*(\eta(x)) = \nabla y^*(x) \cdot \eta(x), \quad \text{a.e. } x \in \Omega,$$

which implies (2.171), as claimed. Hence, y^* is a weak solution in sense of (2.163).

Conversely, any weak solution y^* to (2.163) minimizes the functional φ. Indeed, we have

$$\begin{aligned}\varphi(y^*) - \varphi(v) &\leq \int_\Omega \left(j(\nabla y^*) - j(v) + \frac{1}{2}(|y^* - f|^2 - |v - f|^2) \right) dx \\ &\leq \int_\Omega (\eta \cdot (\nabla y^* - \nabla v) + (y^* - f)(u^* - v)) dx = 0, \qquad \forall v \in C^1(\overline{\Omega}).\end{aligned}$$

The latter inequality extends to all $v \in D\varphi \in \{z \in L^2(\Omega);\ \varphi(z) < \infty\}$. □

Remark 2.5. In particular, it follows by Theorem 2.19 that the operator A, defined by (2.155) in sense of (2.163), is maximal monotone in $L^2(\Omega) \times L^2(\Omega)$.

Remark 2.6. Theorem 2.15 extends to nonlinear elliptic boundary value problems of the form

$$\begin{aligned} \sum_{|\alpha| \leq m} D^\alpha A_\alpha(x, y, D^\beta y) &= f(x), \qquad x \in \Omega,\ |\beta| \leq m, \\ D^\alpha y &= 0 \qquad \text{on } \partial\Omega,\ |\alpha| < m, \end{aligned} \tag{2.174}$$

where $A_\alpha : \Omega \times \mathbf{R}^{mN} \to \mathbf{R}^{mN}$ are measurable functions in x continuous in other variables and satisfy the following conditions.

(i) $\sum_{|\alpha| \leq m} (A_\alpha(x,\xi) - A_\alpha(x,\eta)) \cdot (\xi - \eta) \geq 0,\ \forall \xi, \eta \in \mathbf{R}^{mN}$.

(ii) $\sum_{|\alpha| \leq m} A_\alpha(x,\xi) \cdot \xi \geq \omega \|\xi\|^p - C,\ \forall \xi \in \mathbf{R}^{mN}$, where $\omega > 0,\ p > 1$ and $\|\cdot\|$ is the norm in $\mathbf{R}^{mN}$.

(iii) $\|A_\alpha(x,\xi)\| \leq C_1 \|\xi\|^{p-1} + C_2,\ \forall \xi \in \mathbf{R}^{mN},\ x \in \Omega$.

(Here β is the multi-index $\{D_{x_j}^{\beta_j},\ j = 1, \ldots, N,\ \beta_j \leq m\}$.)

Indeed, under these assumptions the operator $A : X \to X'$, $X = W_0^{m,p}(\Omega)$, $X' = W^{-m,q}(\Omega)$, defined by

$$(Ay, z) = \sum_{|\alpha| \leq m} \int_\Omega A_\alpha(x, y(x), D^\beta y(x)) \cdot D^\alpha z(x) dx, \qquad \forall y, z \in W_0^{m,p}(\Omega)$$

is monotone, demicontinuous, and coercive. Then, the existence of a generalized solution $y \in W_0^{m,p}(\Omega)$ to problem (2.174) for $f \in L^2(\Omega)$ follows by Corollary 2.3. The details are left to the reader.

The nonlinear diffusion techniques and PDE-based variational models are very popular in image denoising and restoring (see, e.g., Rudin, Osher and Fatemi [29]). A gray value image is defined by a function f from a given domain Ω of $\mathbf{R}^d$, $d = 2,3$, to $\mathbf{R}$. In each point $x \in \Omega$, $f(x)$ is the light intensity of the corrupted image located in x. Then, a restored (denoised) image $u : \Omega \to \mathbf{R}$ is computed from the minimization problem (2.165); that is,

$$\text{Minimize}\left\{\frac{1}{2}\int_\Omega (u(x)-f(x))^2dx + \int_\Omega j(\nabla u(x))dx,\ u \in X(\Omega)\right\}, \tag{2.175}$$

where $j : \mathbf{R}^d \to \mathbf{R}$ is a given function and $X(\Omega)$ is a space of functions on Ω. The term $j(\nabla u)$ arising here is taken in order to smooth (mollify) the observation u. In order for the minimization problem to be well posed, one must assume that j is convex and lower semicontinuous and $X(\Omega)$ must be taken, in general, as a distribution space on Ω, for instance, the Sobolev space $W^{1,p}(\Omega)$, where $p \geq 1$. In this case, problem (2.175) is equivalent to the nonlinear diffusion equation

$$\begin{cases} -\mathrm{div}_x(\beta(\nabla y(x))) + y = f & \text{in } \Omega, \\ \beta(\nabla y(x)) \cdot \nu(x) = 0 & \text{on } \partial\Omega, \end{cases}$$

where $\beta : \mathbf{R}^d \to \mathbf{R}^d$ is the subdifferential of j and $\nu = \nu(x)$ is the normal to $\partial\Omega$ at x. The latter equation describes the filtering process of the original corrupted image f.

In the first image processing models, j was taken quadratic and most of the subsequent models have considered functions j of the form

$$j(\nabla y) \equiv |\nabla y|^p, \qquad p > 1,$$

and $X(\Omega)$ was necessarily taken as $W^{1,p}(\Omega)$. As mentioned above, the term $j(\nabla y)$ in the above minimization problem has a smoothing effect in restoring the degraded image f while preserving edges. For the second objective, $p = 1$ (i.e., $j(\nabla y) \equiv |\nabla y|$ and $X(\Omega) = W^{1,1}(\Omega)$) might be apparently the best choice. However, the functional arising in (2.175) is not lower semicontinuous in this latter case in $L^2(\Omega)$ because the functional $y \to \int_\Omega |\nabla y|dx$ is not lower semicontinuous in $L^2(\Omega)$. Thus $W^{1,1}(\Omega)$ must be replaced by the space $BV(\Omega)$ of functions u with bounded variations, and instead of the Sobolev norm $\int_\Omega |\nabla y|dx$ we should take the total variation functional of y. (This functional framework is briefly discussed below.) The case treated in Theorem 2.19 is an intermediate one between $L^p(\Omega)$ with $p > 1$ and $BV(\Omega)$.

The *BV* Approach to the Nonlinear Equations with Singular Diffusivity

As mentioned earlier, the existence theory for equation (2.129) developed above fails for $p = 1$, the best example being, perhaps, in the case where $\beta = \partial j$, $j(u) = |u|$. In this case, equation (2.171) reduces to the singular diffusion equation

$$y - \mathrm{div}_x(\mathrm{sign}(\nabla y)) \ni f \quad \text{in } \Omega, \tag{2.176}$$

with boundary value conditions

$$y = 0 \quad \text{on } \partial\Omega, \tag{2.177}$$

or

$$\mathrm{sign}\,(\nabla y) \cdot \nu = 0 \quad \text{on } \partial\Omega. \tag{2.178}$$

This equation comes formally from variational problems with nondifferentiable energy and it is our aim here to give a rigorous meaning to it. As noticed earlier, this equation is relevant in image restoration as well as in mathematical modeling of faceted crystal growth (see Kobayashi and Giga [18]). Formally, (2.176) is equivalent with the minimization problem (for Dirichlet null boundary condition)

$$\min\left\{\frac{1}{2}\int_\Omega |y-f|^2 dx + \int_\Omega |\nabla y| dx;\ y \in W_0^{1,1}(\Omega)\right\} \tag{2.179}$$

or

$$\min\left\{\frac{1}{2}\int_\Omega |y-f|^2 dx + \int_\Omega |\nabla y| dx;\ y \in W^{1,1}(\Omega)\right\} \tag{2.180}$$

in the case of Neumann boundary conditions. However, as mentioned earlier, problems (2.179) or (2.180) are not well posed in the $W^{1,1}(\Omega)$-setting, the main reason being that the energy functional

$$y \to \int_\Omega |\nabla y(x)| dx$$

is not lower semicontinuous and coercive in an appropriate space of functions on Ω (for instance in $L^p(\Omega),\ p \geq 1$). This fact suggests replacing the space $W^{1,1}(\Omega)$ by a larger space and more precisely by the space $BV(\Omega)$ defined in Section 1.3.

Consider the function $\varphi : L^p(\Omega) \to (-\infty, +\infty]$, $p \geq 1$, defined by

$$\varphi(y) = \begin{cases} \|Dy\| & \text{if } y \in L^p(\Omega) \cap BV^0(\Omega) \\ +\infty & \text{otherwise,} \end{cases} \tag{2.181}$$

respectively,

$$\psi(y) = \begin{cases} \|Dy\| & \text{if } y \in BV(\Omega) \\ +\infty & \text{otherwise.} \end{cases} \tag{2.182}$$

By Theorem 1.14, we know that functions φ and ψ are lower semicontinuous and convex in $L^p(\Omega)$ and, in particular, in $L^2(\Omega)$.

Then the minimization problems

$$\min \frac{1}{2} \int_{\Omega} |y-f|^2 dx + \|Dy\|; \qquad y \in BV^0(\Omega), \tag{2.183}$$

$$\min \frac{1}{2} \int_{\Omega} |y-f|^2 dx + \|Dy\|; \qquad y \in BV(\Omega), \tag{2.184}$$

which replace (2.179) and (2.180), respectively, have unique solutions $y \in BV^0(\Omega)$ and $v \in BV(\Omega)$, respectively. If we denote by $\partial\varphi, \partial\psi : L^2(\Omega) \to L^2(\Omega)$ the subdifferentials of functions φ and ψ; that is,

$$\partial\varphi(y) = \left\{ \eta \in L^2(\Omega);\ \varphi(y) - \varphi(z) \le \int_{\Omega} \eta(y-z)dx,\ \forall y,z \in D(\varphi) \right\}, \tag{2.185}$$

respectively,

$$\partial\psi(y) = \left\{ \xi \in L^2(\Omega);\ \psi(y) - \psi(z) \le \int_{\Omega} \xi(y-z)dx,\ \forall y,z \in D(\psi) \right\}, \tag{2.186}$$

we may write equivalently (2.183) and (2.184) as

$$y + \partial\varphi(y) \ni f \tag{2.187}$$

$$v + \partial\psi(v) \ni f, \tag{2.188}$$

respectively. In variational form, equation (2.187) can be rewritten as

$$\int_{\Omega} y(x)(y(x)-z(x))dx + \|Dy\| \le \|Dz\| + \int_{\Omega} f(x)(y(x)-z(x))dx, \qquad \forall y,z \in BV^0(\Omega),$$

with the obvious modification for (2.188). It is also useful to recall that this equation can be approximated by (see (2.166))

$$\begin{aligned} y_\varepsilon - \varepsilon\Delta y_\varepsilon - \operatorname{div}_x \beta_\varepsilon(\nabla y_\varepsilon) &= f \quad \text{in } \Omega, \\ y_\lambda &= 0 \quad \text{on } \partial\Omega, \end{aligned}$$

where β_ε is the Yosida approximation of $\beta(r) = \operatorname{sign} r$. The solutions y and v to equations (2.187) (respectively, (2.188)) are to be viewed as variational (generalized) solutions to (2.177) and (2.178) and, respectively, (2.187) and (2.188). Taking into account that for $y \in W^{1,1}(\Omega) \subset BV(\Omega)$, we have $\|Dy\| = |\nabla y|_{L^1(\Omega)}$, it follows that, if $y \in W_0^{1,1}(\Omega)$ and $\eta = -\operatorname{div}(\nabla y/|\nabla y|) \in L^2(\Omega)$, then $\eta \in \partial\varphi(y)$. Similarly, if $y \in W^{1,1}(\Omega)$, $\operatorname{sign}(\nabla y) \cdot \nu = 0$ on $\partial\Omega$ and $\xi = -\operatorname{div}(\nabla y/|\nabla y|) \in L^2(\Omega)$, then $\xi \in \partial\psi(y)$. Of course, in general, one might not expect that $y \in W^{1,1}(\Omega)$ and so, the above calculation remains formal. We may conclude, however, that in this generalized sense these equations have unique solutions $u \in BV^0(\Omega)$, respectively, $v \in BV(\Omega)$.

Bibliographical Remarks

The main results of the theory of nonlinear maximal monotone operators in Banach spaces are essentially due to Minty [19, 20] and Browder [9, 10]). Other important contributions are due to Brezis [3]–[5], Lions [16] and Rockafellar [23]–[25], Moreau [21, 22], mainly in connection with the theory of subdifferential type operators. The first applications of the theory of maximal monotone operators to nonlinear elliptic equations of divergence type are due to Browder [9, 10]. The theory of elliptic variational inequalities and its treatment in framework of nonlinear analysis was initiated by Stampacchia and Lions (see [17] and [16] for complete references on the subject) and developed later in a large number of works mostly in connection with its applications to problems with free boundary.

References

1. V. Barbu, *Optimal Control of Variational Inequalities*, Pitman, London, 1983.
2. V. Barbu, *Analysis and Control of Nonlinear Infinite Dimensional Systems*, Academic Press, San Diego, 1993.
3. H. Brezis, Problemes unilatéraux, *J. Math. Pures Appl.,* **51** (1972), pp. 1–168.
4. H. Brezis, Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations, *Contributions to Nonlinear Functional Analysis*, E. Zarantonello (Ed.), Academic Press, New York, 1971.
5. H. Brezis, Equations et inéquations nonlinéaires dans les espaces vectorielles en dualité, *Ann. Institute Fourier*, 18 (1968), pp. 115–175.
6. H. Brezis, G. Stampacchia, Sur la régularité de la solution d'inéquations elliptiques, *Bull. Soc. Math. France,* 95 (1968), p. 153.
7. H. Brezis, M.G. Crandall, A. Pazy, Perturbations of nonlinear maximal monotone sets, *Comm. Pure Appl. Math.*, **13** (1970), pp. 123–141.
8. H. Brezis, F. Browder, Some properties of higher order Sobolev spaces, *J. Math. Pures Appl.*, **61** (1982), pp. 245–259.
9. F. Browder, *Problèmes Nonlinéaires*, Les Presses de l'Université de Montréal, 1966.
10. F. Browder, *Nonlinear Operators and Nonlinear Equations of Evolution in Banach Spaces*, Nonlinear Functional Analysis, Symposia in Pure Math., vol. **18**, Part 2, F. Browder (Ed.), American Mathematical Society, Providence, RI, 1970.
11. K. Deimling, *Nonlinear Functional Analysis*, Springer-Verlag, Berlin, 1985.
12. G. Duvaut, J.L. Lions, *Inequalities in Mechanics and Physics*, Springer-Verlag, Berlin, 1976.
13. P.M. Fitzpatrick, Surjectivity results for nonlinear mappings from a Banach space to its dual, *Math. Ann.*, **204** (1973), pp. 177–188.
14. A. Friedman, *Variational Principles and Free Boundary Problems*, John Wiley and Sons, New York, 1982.
15. L.I. Hedberg, Two approximation problems in function spaces, *Ark. Mat.*, **16** (1978), pp. 51–81.
16. J.L. Lions, *Quelques Méthodes de Résolution de Problèmes Nonlinéaires*, Dunod-Gauthier Villars, Paris, 1969.
17. D. Kinderlehrer, G. Stampacchia, *An Introduction to Variational Inequalities and Applications*, Academic Press, New York, 1980.
18. R. Kobayashi, Y. Giga, Equations with singular diffusivity, *J. Statistical Physics*, **95** (1999), pp. 1187–1220.
19. G. Minty, Monotone (nonlinear) operators in Hilbert spaces, *Duke Math. J.*, **29** (1962), pp. 341–346.

20. G. Minty, On the generalization of a direct method of the calculus of variations, *Bull. Amer. Math.Soc.*, **73** (1967), pp. 315–321.
21. J.J. Moreau, Proximité et dualité dans un espace hilbertien, *Bull. Soc. Math. France*, **93** (1965), pp. 273–299.
22. J.J. Moreau, *Fonctionnelle Convexes*, Seminaire sur les équations aux dérivées partielles, Collège de France, Paris 1966–1967.
23. R.T. Rockafellar, *Convex Analysis,* Princeton University Press, Princeton, NJ, 1969.
24. R.T. Rockafellar, On the maximal monotonicity of subdifferential mappings, *Pacific J. Math.*, **33** (1970), pp. 209–216.
25. R.T. Rockafellar, Local boundedness of nonlinear monotone operators, *Michigan Math. J.*, **16** (1969), pp. 397–407.
26. R.T. Rockafellar, On the maximality of sums of nonlinear operatorsw, *Trans. Amer. Math. Soc.*, **149** (1970), pp. 75–88.
27. R.T. Rockafellar, Integrals which are convex functional II, *Pacific J. Math.*, **39** (1971), pp. 439–469.
28. R.T. Rockafellar, Integral functionals, normal integrands and measurable selections, *Nonlinear Operators and the Calculus of Variations*, J.P. Gossez, E. Dozo, J. Mawhin, L. Waelbroeck (Eds.), Lecture Notes in Mathematics, Springer-Verlag, New York, 1976, pp. 157–205.
29. L.I. Rudin, S. Osher, E. Fatemi, Nonlinear total variation based noise removal algorithms, *Physica*, D, **60** (1992), pp. 259–260.

Chapter 3
Accretive Nonlinear Operators in Banach Spaces

Abstract This chapter is concerned with the general theory of nonlinear quasi-m-accretive operators in Banach spaces with applications to the existence theory of nonlinear elliptic boundary value problems in L^p-spaces and first-order quasilinear equations. While the monotone operators are defined in a duality pair (X, X^*) and, therefore, in a variational framework, the accretive operators are intrinsically related to geometric properties of the space X and are more suitable for nonvariational and nonHilbertian existence theory of nonlinear problems. The presentation is confined, however, to the essential results of this theory necessary to the construction of accretive dynamics in the next chapter.

3.1 Definition and General Theory

Throughout this chapter, X is a real Banach space with the norm $\|\cdot\|$, X^* is its dual space, and $(\cdot,\cdot)$ the pairing between X and X^*. We denote as usual by $J : X \to X^*$ the duality mapping of the space X.

Definition 3.1. A subset A of $X \times X$ (equivalently, a multivalued operator from X to X) is called *accretive* if for every pair $[x_1, y_1], [x_2, y_2] \in A$, there is $w \in J(x_1 - x_2)$ such that

$$(y_1 - y_2, w) \geq 0. \tag{3.1}$$

An accretive set is said to be *maximal accretive* if it is not properly contained in any accretive subset of $X \times X$.

An accretive set A is said to be *m-accretive* if

$$R(I + A) = X. \tag{3.2}$$

Here we have denoted I the unity operator in X, but when there is no danger of confusion, we simply write 1 instead of I.

V. Barbu, *Nonlinear Differential Equations of Monotone Types in Banach Spaces*, Springer Monographs in Mathematics, DOI 10.1007/978-1-4419-5542-5_3,

We denote by $D(A) = \{x \in X;\ Ax \neq \emptyset\}$ the domain of A and by $R(A) = \{y \in Ax;\ [x,y] \in A\}$ the range of A. As in the case of operators from X to X^*, we identify an operator (eventually multivalued) $A : D(A) \subset X \to X$ with its graph $\{[x,y];\ y \in Ax\}$ and so view A as a subset of $X \times X$.

A subset A is called *dissipative* (respectively, *maximal dissipative*, *m-dissipative*) if $-A$ is accretive (respectively, maximal, m-accretive).

Finally, A is said to be *ω-accretive* (*ω-m-accretive*), where $\omega \in \mathbf{R}$, if $A + \omega I$ is accretive (respectively, m-accretive). A subset $A \subset X \times X$ that is ω-accretive or ω-m-accretive for some $\omega \in \mathbf{R}$ is called *quasi-accretive*, respectively, *quasi-m-accretive*.

As we show below, the accretiveness of A is, in fact, a metric geometric property that can be equivalently expressed as

$$\|x_1 - x_2\| \leq \|x_1 - x_2 + \lambda(y_1 - y_2)\|, \qquad \forall \lambda > 0,\ [x_i, y_i] \in A,\ i = 1,2, \tag{3.3}$$

using the following lemma (Kato's lemma).

Lemma 3.1. *Let $x, y \in X$. Then there exists $w \in J(x)$ such that $(y, w) \geq 0$ if and only if*

$$\|x\| \leq \|x + \lambda y\|, \qquad \forall \lambda > 0 \tag{3.4}$$

holds.

Proof. Let x and y in X be such that $(y,w) \geq 0$ for some $w \in J(x)$. Then, by definition of J, we have

$$\|x\|^2 = (x,w) \leq (x + \lambda y, w) \leq \|x + \lambda y\| \cdot \|w\| = \|x + \lambda y\| \cdot \|x\|, \qquad \forall \lambda > 0,$$

and (3.4) follows.

Suppose now that (3.4) holds. For $\lambda > 0$, let w_λ be an arbitrary element of $J(x + \lambda y)$. Without loss of generality, we may assume that $x \neq 0$. Then, $w_\lambda \neq 0$ for λ small. We set $f_\lambda = w_\lambda \|w_\lambda\|^{-1}$. Because $\{f_\lambda\}_{\lambda > 0}$ is weak-star compact in X^*, there exists a generalized sequence, again denoted λ, such that $f_\lambda \rightharpoonup f$ in X^*. On the other hand, from the inequality

$$\|x\| \leq \|x + \lambda y\| = (x + \lambda y, f_\lambda) \leq \|x\| + \lambda(y, f_\lambda)$$

it follows that

$$(y, f_\lambda) \geq 0, \qquad \forall \lambda > 0.$$

Hence, $(y,f) \geq 0$ and $\|x\| \leq (x,f)$. Because $\|f\| \leq 1$, this implies that $\|x\| = (x,f)$, $\|f\| = 1$, and therefore $w = f\|x\| \in J(x)$, $(y,w) \geq 0$, as claimed. □

Proposition 3.1. *A subset A of $X \times X$ is accretive if and only if inequality* (3.3) *holds for all $\lambda > 0$* (*equivalently, for some $\lambda > 0$*) *and all* $[x_i, y_i] \in A$, $i = 1,2$.

Proposition 3.1 is an immediate consequence of Lemma 3.1. In particular, it follows that A is ω-accretive iff

$$\|x_1 - x_2 + \lambda(y_1 - y_2)\| \geq (1 - \lambda\omega)\|x_1 - x_2\|$$
$$\text{for } 0 < \lambda < \frac{1}{\omega} \quad \text{and} \quad [x_i, y_i] \in A,\ i = 1, 2. \tag{3.5}$$

Hence, if A is accretive, then the operator $(I + \lambda A)^{-1}$ is single-valued and nonexpansive on $R(I + \lambda A)$; that is,

$$\|(I + \lambda A)^{-1}x - (I + \lambda A)^{-1}y\| \leq \|x - y\|, \qquad \forall \lambda > 0,\ x, y \in R(I + \lambda A).$$

If A is ω-accretive, then it follows by (3.5) that $(I + \lambda A)^{-1}$ is single-valued and Lipschitzian with Lipschitz constant not greater than $1/(1 - \lambda\omega)$ on $R(I + \lambda A)$, $0 < \lambda < 1/\omega$.

Let us define the operators J_λ and A_λ:

$$J_\lambda x = (I + \lambda A)^{-1}x, \qquad x \in R(I + \lambda A); \tag{3.6}$$

$$A_\lambda x = \lambda^{-1}(x - J_\lambda x), \qquad x \in R(I + \lambda A). \tag{3.7}$$

As in the case of maximal monotone operators in $X \times X^*$ (see (2.26)), the operator A_λ is called the *Yosida approximation of* A, and, in the special case when $X = H$ is a Hilbert space, it is just the operator studied in Proposition 2.2.

In Proposition 3.2 below, we collect some elementary properties of J_λ and A_λ.

Proposition 3.2. *Let A be ω-accretive in $X \times X$. Then:*

(a) $\|J_\lambda x - J_\lambda y\| \leq (1 - \lambda\omega)^{-1}\|x - y\|$, $\forall \lambda \in (0, 1/\omega)$, $\forall x, y \in R(I + \lambda A)$.

(b) *A_λ is ω-accretive and Lipschitz continuous with Lipschitz constant not greater than $2/(1 - \lambda\omega)$ in $R(I + \lambda A)$, $0 < \lambda < 1/\omega$.*

(c) $A_\lambda x \in AJ_\lambda x$, $\forall x \in R(I + \lambda A)$, $0 < \lambda < 1/\omega$.

(d) $(1 - \lambda\omega)\|A_\lambda x\| \leq |Ax| = \inf\{\|y\|;\ y \in Ax\}$;

(e) $\lim_{\lambda \to 0} J_\lambda x = x$, $\forall x \in \overline{D(A)} \cap_{0<\lambda<1/\omega} R(I + \lambda A)$.

Proof. (a) and (b) are immediate consequences of inequality (3.5).

(c) Let $x \in R(I + \lambda A)$. Then, $A_\lambda x \in \lambda^{-1}((I + \lambda A)J_\lambda x - J_\lambda x) \in AJ_\lambda x$.

(d) For $x \in D(A) \cap R(I + \lambda A)$, we have $A_\lambda x = \lambda^{-1}(J_\lambda(I + \lambda A)x - J_\lambda x)$ and, therefore, $\|A_\lambda x\| \leq |Ax|(1 - \lambda\omega)^{-1}$, $\forall x \in D(A)$.

(e) For every $x \in D(A) \cap R(I + \lambda A)$, we have

$$\|J_\lambda x - x\| = \lambda\|A_\lambda x\| \leq \frac{\lambda}{1 - \lambda\omega}|Ax|, \qquad \forall \lambda \in \left(0, \frac{1}{\omega}\right).$$

Hence, $\lim_{\lambda \to 0} J_\lambda x = x$. Clearly, this extends to all of $\overline{D(A)} \bigcap_{0<\lambda<1/\omega} R(I + \lambda A)$, as claimed. □

In the following we confine ourselves to the study of accretive subsets, the extensions to the quasi-accretive sets being immediate.

Proposition 3.3. *An accretive set $A \subset X \times X$ is m-accretive if and only if $R(I + \lambda A) = X$ for all (equivalently, for some) $\lambda > 0$.*

Proof. Let A be m-accretive and let $y \in X$, $\lambda > 0$, be arbitrary but fixed. Then, the equation

$$x + \lambda Ax \ni y \tag{3.8}$$

may be written as

$$x = J_1\left(\frac{y}{\lambda} + \left(1 - \frac{1}{\lambda}\right)x\right).$$

Then, by the contraction principle, we infer that the equation has a solution for $1/2 < \lambda < +\infty$.

Now, fix $\lambda_0 > 1/2$ and write the preceding equation as

$$x = (I + \lambda_0 A)^{-1}\left(\left(1 - \frac{\lambda_0}{\lambda}\right)x + \frac{\lambda_0}{\lambda}\,y\right). \tag{3.9}$$

Because $J_{\lambda_0} = (I + \lambda_0 A)^{-1}$ is nonexpansive, this equation has a solution for $\lambda \in (\lambda_0/2, \infty)$. Repeating this argument, we conclude that $R(I + \lambda A) = X$ for all $\lambda > 0$. Assume now that $R(I + \lambda_0 A) = X$ for some $\lambda_0 > 0$. Then, if we set equation (3.8) into the form (3.9), we conclude as before that $R(I + \lambda A) = X$ for all $\lambda \in (\lambda_0/2, \infty)$ and so $R(I + \lambda A) = X$ for all $\lambda > 0$, as claimed. □

Combining Propositions 3.2 and 3.3, we conclude that $A \subset X \times X$ is m-accretive if and only if for all $\lambda > 0$ the operator $(I + \lambda A)^{-1}$ is nonexpansive on all of X.

Similarly, A is ω-m-accretive if and only if, for all $0 < \lambda < 1/\omega$,

$$\|(I + \lambda A)^{-1}x - (I + \lambda A)^{-1}y\| \le \frac{1}{1 - \lambda\omega}\,\|x - y\|, \qquad \forall x, y \in X. \tag{3.10}$$

By Theorem 2.2, if $X = H$ is a Hilbert space, then A is m-accretive if and only if it is maximal accretive.

A subset $A \subset X \times X$ is said to be *demiclosed* if it is closed in $X \times X_w$; that is, if $x_n \to x$, $y_n \rightharpoonup y$, and $[x_n, y_n] \in A$, then $[x, y] \in A$ (recall that $\rightharpoonup$ denotes weak convergence). A is said to be *closed* if $x_n \to x$, $y_n \to y$, and $[x_n, y_n] \in A$ for all $n \in \mathbf{N}$ imply that $[x, y] \in A$.

Proposition 3.4. *Let A be an m-accretive set of $X \times X$. Then A is closed and if $\lambda_n \in \mathbf{R}$, $x_n \in X$ are such that $\lambda_n \to 0$ and*

$$x_n \to x,\ A_{\lambda_n} x_n \to y \quad \textit{for } n \to \infty, \tag{3.11}$$

then $[x, y] \in A$. If X^* is uniformly convex, then A is demiclosed, and if

$$x_n \to x,\ A_{\lambda_n} x \rightharpoonup y \quad \textit{for } n \to \infty, \tag{3.12}$$

then $[x, y] \in A$.

Proof. Let $x_n \to x$, $y_n \to y$, $[x_n, y_n] \in A$. Because A is accretive, we have

$$\|x_n - u\| \le \|x_n + \lambda y_n - (u + \lambda v)\|, \qquad \forall [u, v] \in A,\ \lambda > 0.$$

Hence,

$$\|x-u\| \le \|x+\lambda y-(u+\lambda v)\|, \qquad \forall [u,v] \in A,\ \lambda > 0.$$

Now, A being m-accretive, there is $[u,v] \in A$ such that $u+\lambda v = x+\lambda y$. Substituting in the latter inequality, we see that $x = u$ and $y = v \in Ax$, as claimed.

Now, if λ_n, x_n satisfy condition (3.11), then $\{A_{\lambda_n} x_n\}$ is bounded and so $J_{\lambda_n} x_n - x_n \to 0$. Because $A_{\lambda_n} x_n \in AJ_{\lambda_n} x_n$, $J_{\lambda_n} x_n \to x$, and A is closed, we have that $[x,y] \in A$. We assume now that X^* is uniformly convex. Let x_n, y_n be such that $x_n \to x$, $y_n \rightharpoonup y$, $[x_n, y_n] \in A$. Inasmuch as A is accretive, we have

$$(y_n - v, J(x_n - u)) \ge 0, \qquad \forall [u,v] \in A,\ n \in \mathbf{N}^*.$$

On the other hand, recalling that J is continuous on X (Theorem 1.2), we may pass to the limit $n \to \infty$ to obtain

$$(y - v, J(x-u)) \ge 0, \qquad \forall [u,v] \in A.$$

Now, if we take $[u,v] \in A$ such that $u+v = x+y$, we see that $y = v$ and $x = u$. Hence, $[x,y] \in A$, and so A is demiclosed. The final part of Proposition 3.4 is an immediate consequence of this property, remembering that $A_{\lambda_n} x_n \in AJ_{\lambda_n} x_n$. □

Remark 3.1. Note that an m-accretive set of $X \times X$ is maximal accretive. Indeed, if $[x,y] \in X \times X$ is such that

$$\|x-u\| \le \|x+\lambda y-(u+\lambda v)\|, \qquad \forall [u,v] \in A,\ \lambda > 0,$$

then, choosing $[u,v] \in A$ such that $u+\lambda v = x+\lambda y$, we see that $x = u$ and so $v = y \in Ax$. These two properties are equivalent, however, in Hilbert spaces.

If X^* is uniformly convex, then it follows that, for every $x \in D(A)$, we have the following algebraic description of Ax

$$Ax = \{y \in X;\ (y-v, J(x-u)) \ge 0, \qquad \forall [u,v] \in A\}.$$

In particular, it follows that Ax is a closed convex subset of X. Denote by $A^0 x$ the element of minimum norm on Ax (i.e., the projection of the origin into Ax). Because the space X is reflexive, by Proposition 1.4 it follows that $A^0 x \neq \emptyset$ for every $x \in D(A)$. The set $A^0 \subset A$ is called the *minimal section* of A. If the space X is strictly convex, then, as easily seen, A^0 is single-valued.

Proposition 3.5. *Let X and X^* be uniformly convex and let A be an m-accretive set of $X \times X$. Then:*

(i) $A_\lambda x \to A^0 x$, $\forall x \in D(A)$ *for* $\lambda \to 0$.
(ii) $\overline{D(A)}$ *is a convex set of* X.

Proof. (i) Let $x \in D(A)$. As seen in Proposition 3.2, $\|A_\lambda x\| \le |Ax| = \|A^0 x\|$, $\forall \lambda > 0$. Now, let $\lambda_n \to 0$ be such that $A_{\lambda_n} x \rightharpoonup y$. By Proposition 3.1, we know that $y \in Ax$, and thus

$$\lim_{n\to\infty} \|A_{\lambda_n} x\| = \|y\| = \|A^0 x\|.$$

The space X is uniformly convex; therefore this implies that $A_{\lambda_n} x \to y = A^0 x$ (Lemma 1.1). Hence, $A_\lambda x \to A^0 x$ for $\lambda \to 0$.

(ii) Let $x_1, x_2 \in D(A)$, and $0 \le \alpha \le 1$. We set $x_\alpha = \alpha x_1 + (1-\alpha)x_2$. Then, as it is readily verified,

$$\|J_\lambda(x_\alpha) - x_1\| \le \|x_\alpha - x_1\| + \lambda |Ax_1|, \qquad \forall \lambda > 0,$$
$$\|J_\lambda(x_\alpha) - x_2\| \le \|x_\alpha - x_2\| + \lambda |Ax_2|, \qquad \forall \lambda > 0,$$

and, because the space X is uniformly convex, these estimates imply, by a standard geometrical device we omit here, that

$$\|J_\lambda(x_\alpha) - x_\alpha\| \le \delta(\lambda), \qquad \forall \lambda > 0,$$

where $\lim_{\lambda\to 0} \delta(\lambda) = 0$. Hence, $x_\alpha \in \overline{D(A)}$. □

Regarding the single-valued linear m-accretive (equivalently, m-dissipative) operators, it is useful to note the following density result.

Proposition 3.6. *Let X be a Banach space. Then any m-accretive linear operator $A : X \to X$ is densely defined (i.e., $\overline{D(A)} = X$).*

Proof. Let $y \in X$ be arbitrary but fixed. For every $\lambda > 0$, the equation $x_\lambda + \lambda A x_\lambda = y$ has a unique solution $x_\lambda \in D(A)$. We know that $\|x_\lambda\| \le \|y\|$ for all $\lambda > 0$ and so, on a subsequence $\lambda_n \to 0$,

$$x_{\lambda_n} \rightharpoonup x, \ \lambda_n A x_{\lambda_n} \rightharpoonup y - x \quad \text{in } X.$$

Because A is closed, its graph in $X \times X$ is weakly closed (it is a linear subspace of $X \times X$) and so $\lambda_n x_{\lambda_n} \to 0$, $A(\lambda_n x_{\lambda_n}) \rightharpoonup y - x$ imply that $y - x = 0$. Hence,

$$(1 + \lambda_n A)^{-1} y \rightharpoonup y.$$

We have, therefore, proven that $y \in \overline{D(A)}$ (recall that the weak closure of $D(A)$ coincides with the strong closure). □

We conclude this section by introducing another convenient way to define the accretiveness. Toward this aim, denote by $[\cdot,\cdot]_s$ the directional derivative of the function $x \to \|x\|$; that is (see (1.13)),

$$[x,y]_s = \lim_{\lambda \downarrow 0} \frac{\|x + \lambda y\| - \|x\|}{\lambda}, \qquad x, y \in X. \tag{3.13}$$

The function $\lambda \to \|x + \lambda y\|$ is convex, thus we may define, equivalently, $[\cdot,\cdot]_s$ as

$$[x,y]_s = \inf_{\lambda > 0} \frac{\|x + \lambda y\| - \|x\|}{\lambda}, \qquad \forall x, y \in X. \tag{3.14}$$

Roughly speaking, $[\cdot,\cdot]_s$ can be viewed as a "scalar product" on $X \times X$.

Let us now briefly list some properties of the bracket $[\cdot,\cdot]_s$.

Proposition 3.7. *Let X be a Banach space. We have the following.*

(i) $[\cdot,\cdot]_s : X \times X \to \mathbf{R}$ *is upper semicontinuous.*
(ii) $[\alpha x, \beta y]_s = \beta [x,y]_s$, *for all* $\beta \geq 0,\ \alpha \in \mathbf{R},\ x,y \in X$.
(iii) $[x, \alpha x + y]_s = \alpha \|x\| + [x,y]_s$ *if* $\alpha \in \mathbf{R}^+,\ x \in X$.
(iv) $|[x,y]_s| \leq \|y\|,\ [x,y+z]_s \leq [x,y]_s + [x,z]_s,\ \forall x,y,z \in X$.
(v) $[x,y]_s = \max\{(y,x^*);\ x^* \in \Phi(x)\},\ \forall x,y \in X$, *where*

$$\Phi(x) = \{x^* \in X^*;\ (x,x^*) = \|x\|,\ \|x^*\| = 1\},\ \textit{if}\ x \neq 0,$$
$$\Phi(0) = \{x^* \in X^*;\ \|x^*\| \leq 1\}.$$

Proof. (i) Let $x_n \to x$ and $y_n \to y$ as $n \to \infty$. For every n there exist $h_n \in X$ and $\lambda_n \in (0,1)$ such that $\|h_n\| + \lambda_n \leq 1/n$ and

$$[x_n, y_n]_s \leq (\|x_n + h_n + \lambda_n y_n\| - \|x_n + y_n\|)\lambda_n^{-1} + (1/n).$$

This yields

$$\limsup_{n\to\infty} [x_n, y_n]_s \leq [x,y]_s,$$

as claimed.

Note that (ii)–(iv) are immediate consequences of the definition. To prove (v), we note first that

$$\Phi(x) = \partial(\|x\|), \qquad \forall x \in X,$$

and apply Proposition 2.6. □

Now, coming back to the definition of accretiveness, we see that, in virtue of part (v) of Proposition 3.7, condition (3.3) can be equivalently written as

$$[x_1 - x_2, y_1 - y_2]_s \geq 0, \qquad \forall [x_i, y_i] \in A,\ i = 1,2. \tag{3.15}$$

Similarly, condition (3.5) is equivalent to

$$[x_1 - x_2, y_1 - y_2]_s \geq -\omega \|x_1 - x_2\|, \qquad \forall [x_i, y_i] \in A,\ i = 1,2. \tag{3.16}$$

Summarizing, we may see that a subset A of $X \times X$ is ω-accretive if one of the following equivalent conditions holds.

(i) If $[x_1,y_1],[x_2,y_2] \in A$, then there is $w \in J(x_1 - x_2)$ such that

$$(y_1 - y_2, w) \geq -\omega \|x_1 - x_2\|.$$

(ii) $\|x_1 - x_2 + \lambda(y_1 - y_2)\| \geq (1 - \lambda\omega)\|x_1 - x_2\|$ for $0 < \lambda < 1/\omega$ and all $[x_i, y_i] \in A$, $i = 1,2$.
(iii) $[x_1 - x_2, y_1 - y_2]_s \geq -\omega \|x_1 - x_2\|,\ \forall [x_i, y_i] \in A,\ i = 1,2$.

In applications, however, it is more convenient to use condition (i) to verify the ω-accretiveness.

We know that, if X is a Hilbert space, then a continuous accretive operator is m-accretive (see Lemma 1.3). This result was extended by R. Martin [11] to general Banach spaces. More generally, we have the following result established by the author in [1]. (See also [2].)

Theorem 3.1. *Let X be a real Banach space, A be an m-accretive set of $X \times X$, and let $B : X \to X$ be a continuous, m-accretive operator with $D(B) = X$. Then $A + B$ is m-accretive.*

This result (which can be compared with Corollary 2.6) is, in particular, useful to treat continuous nonlinear accretive perturbations of equations involving m-accretive operators.

Other m-accretive criteria for the sum $A+B$ of two m-accretive operators $A,B \in X \times X$ can be obtained approximating the equation $x + Ax + Bx \ni y$ by

$$x + Ax + B_\lambda x \ni y,$$

where B_λ is the Yosida approximation of B.

We illustrate the method on the following example.

Proposition 3.8. *Let X be a Banach space with uniformly convex dual X^* and let A and B be two m-accretive sets in $X \times X$ such that $D(A) \cap D(B) \neq \emptyset$ and*

$$(Au, J(B_\lambda u)) \geq 0, \qquad \forall \lambda > 0,\ u \in D(A). \tag{3.17}$$

Then $A + B$ is m-accretive.

Proof. Let $f \in X$ and $\lambda > 0$ be arbitrary but fixed. We approximate the equation

$$u + Au + Bu \ni f \tag{3.18}$$

by

$$u + Au + B_\lambda u \ni f, \qquad \lambda > 0, \tag{3.19}$$

where B_λ is the Yosida approximation B; that is,

$$B_\lambda = \lambda^{-1}(I - (I + \lambda B)^{-1}).$$

We may write equation (3.19) as

$$u = \left(1 + \frac{\lambda}{1+\lambda} A\right)^{-1} \left(\frac{\lambda f}{1+\lambda} + \frac{(I+\lambda B)^{-1} u}{1+\lambda}\right),$$

which, by the Banach fixed point theorem, has a unique solution $u_\lambda \in D(A)$ (because $(I+\lambda B)^{-1}$ and $(I+\lambda A)^{-1}$ are nonexpansive). Now, we multiply the equation

$$u_\lambda + Au_\lambda + B_\lambda u_\lambda \ni f \tag{3.20}$$

by $J(B_\lambda u_\lambda)$ and use condition (3.17) to get that

$$\|B_\lambda u_\lambda\| \le \|f\| + \|u_\lambda\|, \qquad \forall \lambda > 0.$$

On the other hand, multiplying (3.20) by $J(u_\lambda - u_0)$, where $u_0 \in D(A) \cap D(B)$, we get

$$\|u_\lambda - u_0\| \le \|u_0\| + \|f\| + \|\xi_0\| + \|B_\lambda u_0\| \le \|u_0\| + \|f\| + \|\xi_0\| + |Bu_0|, \qquad \forall \lambda > 0,$$

where $\xi_0 \in Au_0$. Hence,

$$\|u_\lambda\| + \|B_\lambda u_\lambda\| \le C, \qquad \forall \lambda > 0. \tag{3.21}$$

Now, multiplying the equation (in the sense of the duality between X and X^*)

$$u_\lambda - u_\mu + Au_\lambda - Au_\mu + B_\lambda u_\lambda - B_\mu u_\mu \ni 0$$

by $J(u_\lambda - u_\mu)$. Because A is accretive, we have

$$\|u_\lambda - u_\mu\|^2 + (B_\lambda u_\lambda - B_\mu u_\mu, J(u_\lambda - u_\mu)) \le 0, \qquad \forall \lambda, \mu > 0. \tag{3.22}$$

On the other hand,

$$\begin{aligned}(B_\lambda u_\lambda\ &- B_\mu u_\mu, J(u_\lambda - u_\mu)) \\ &\ge \left(B_\lambda u_\lambda - B_\mu u_\mu, J(u_\lambda - u_\mu) - J\left((I+\lambda B)^{-1}u_\lambda - (I+\mu B)^{-1}u_\mu\right)\right)\end{aligned}$$

because B is accretive and $B_\lambda u \in B((I+\lambda B)^{-1}u)$. Because J is uniformly continuous on bounded subsets (Theorem 1.2) and by (3.21) we have

$$\|u_\lambda - (I+\lambda B)^{-1}u_\lambda\| + \|u_\mu - (I+\mu B)^{-1}u_\mu\| \le C(\lambda + \mu),$$

this implies that $\{u_\lambda\}$ is a Cauchy sequence and so $u = \lim_{\lambda\to 0} u_\lambda$ exists. Extracting further subsequences, we may assume that

$$B_\lambda u_\lambda \rightharpoonup y, \qquad f - B_\lambda u_\lambda - u_\lambda \rightharpoonup z.$$

Then, by Proposition 3.4, we see that $y \in Bu$, $z \in Au$, and so u is a solution (obviously unique) to equation (3.18). □

If X is a Hilbert space and $A = \partial\varphi$, then Proposition 3.8 reduces to Theorem 2.11. We also note the following perturbation result.

Proposition 3.9. *Let X be a Banach space with a uniformly convex dual and let A, B be two m-accretive sets in $X \times X$ such that, for each $r > 0$,*

$$\|B^0x\| \le \alpha\|A^0x\| + Cr \quad \textit{for} \ \ \|x\| \le r,\ \forall x \in D(A), \tag{3.23}$$

where $0 < \alpha < 1$. Then $A + B$ is m-accretive. Here A^0 is the minimal section of A.

Proof. For $f \in X$ we approximate, as above, equation (3.18) by (3.19) and denote by $u_\lambda \in D(A)$ the solution to (3.19). We have, of course, that $\{u_\lambda\}$ is bounded in X (i.e., $\|u_\lambda\| \le r$, $\forall \lambda > 0$), and by Proposition 3.2, part (d), and by assumption (3.23) it follows that

$$\|B_\lambda u_\lambda\| \le \|B^0 u_\lambda\| \le \alpha \|A^0 u_\lambda\| + Cr \le \alpha(\|B_\lambda u_\lambda\| + \|f\| + r), \qquad \forall \lambda > 0.$$

This yields $\|B_\lambda u_\lambda\| \le C$, $\forall \lambda > 0$, and, arguing as in the proof of Proposition 3.8, we infer that, for $\lambda \to 0$,

$$\begin{aligned} u_\lambda &\to u \quad \text{in } X, \\ B_\lambda u_\lambda &\rightharpoonup \eta \quad \text{in } X, \\ w_\lambda = f - u_\lambda - B_\lambda u_\lambda &\rightharpoonup \xi \quad \text{in } X, \end{aligned}$$

where $\eta \in Bu$ and $\xi \in Ax$. Hence, by Proposition 3.4 we have $f \in R(I + A + B)$, as claimed. □

Remark 3.2. The accretivity property of an operator A defined in a Banach space X should not be mixed up with that of monotonicity. The first is defined for operators A from X to itself and is a metric geometric property, whereas the second is defined for operators A from X to dual space X^* and is a variational property. Of course, as mentioned earlier, these two concepts coincide if the space X is Hilbert and is identified with its own dual.

3.2 Nonlinear Elliptic Boundary Value Problem in L^p

In most situations, the m-accretive operators arise as partial differential operators on a domain Ω with appropriate boundary value conditions. These boundary value problems do not have an appropriate formulation in a variational functional setting (as in the case with elliptic boundary value problems in $L^p(\Omega)$ spaces or that of nonlinear elliptic problems of divergence type treated in Section 2.4) but have, however, an adequate treatment in the framework of m-accretive operator theory. We treat a few significant examples below. Throughout this section, Ω is a bounded and open subset of $\mathbf{R}^N$ with a smooth boundary, denoted $\partial\Omega$.

Semilinear Elliptic Operators in $L^p(\Omega)$

Let β be a maximal monotone graph in $\mathbf{R} \times \mathbf{R}$ such that $0 \in D(\beta)$.

Let $\widetilde{\beta} \subset L^p(\Omega) \times L^p(\Omega)$, $1 \le p < \infty$, be the operator defined by

$$\begin{aligned} \widetilde{\beta}(u(x)) &= \{v \in L^p(\Omega);\ v(x) \in \beta(u(x))),\ \text{a.e. } x \in \Omega\}, \\ D(\widetilde{\beta}) &= \{u \in L^p(\Omega);\ \exists v \in L^p(\Omega) \text{ so that } v(x) \in \beta(u(x)),\ \text{a.e. } x \in \Omega\}. \end{aligned} \tag{3.24}$$

It is easily seen that $\widetilde{\beta}$ is m-accretive in $L^p(\Omega) \times L^p(\Omega)$ and

$$((I+\lambda\widetilde{\beta})^{-1}u) = (1+\lambda\beta)^{-1}u(x), \quad \text{a.e. } x \in \Omega,\ \lambda > 0,$$
$$(\widetilde{\beta}_\lambda u)(x) = \beta_\lambda(u(x)), \quad \text{a.e. } x \in \Omega,\ \lambda > 0,\ u \in L^p(\Omega).$$

Very often, this operator $\widetilde{\beta}$ is called the *realization* of the graph $\beta \subset \mathbf{R} \times \mathbf{R}$ in the space $L^p(\Omega) \times L^p(\Omega)$.

Theorem 3.2. *Let $A : L^p(\Omega) \to L^p(\Omega)$ be the operator defined by*

$$\begin{aligned} Au &= -\Delta u + \widetilde{\beta}(u), && \forall u \in D(A),\\ D(A) &= W_0^{1,p}(\Omega) \cap W^{2,p}(\Omega) \cap D(\widetilde{\beta}) && \textit{if } 1 < p < \infty,\\ D(A) &= \{u \in W_0^{1,1}(\Omega);\ \Delta u \in L^1(\Omega)\} \cap D(\widetilde{\beta}) && \textit{if } p = 1. \end{aligned} \tag{3.25}$$

Then A is m-accretive and surjective in $L^p(\Omega)$.

We note that, for $p = 2$, this result has been proven in Proposition 2.8.

Proof. Let us show first that A is accretive. If $u_1, u_2 \in D(A)$ and $v_1 \in Au_1$, $v_2 \in Au_2$, $1 < p < \infty$, we have, by Green's formula,

$$\|u_1 - u_2\|_{L^p(\Omega)}^{p-2}(v_1 - v_2, J(u_1 - u_2)) = -\int_\Omega \Delta(u_1 - u_2)|u_1 - u_2|^{p-2}(u_1 - u_2)dx$$
$$+ \int_\Omega (\beta(u_1) - \beta(u_2))(u_1 - u_2)|u_1 - u_2|^{p-2}dx \geq 0$$

because β is monotone (recall that $J(u)(x) = |u(x)|^{p-2}u(x)\|u\|_{L^p(\Omega)}^{2-p}$ is the duality mapping of the space $L^p(\Omega)$). (In the previous formula and everywhere in the sequel, by $\beta(u_i)$, $i = 1, 2$, we mean single-valued sections of $\beta(u_i)$ which arise in the definition of Au_i.) In the case $p = 1$, consider the function $\gamma_\varepsilon : \mathbf{R} \to \mathbf{R}$ defined by

$$\gamma_\varepsilon(r) = \begin{cases} 1 & \text{for } r > \varepsilon,\\ \theta_\varepsilon(r) & \text{for } -\varepsilon \leq r \leq \varepsilon,\\ -1 & \text{for } r < -\varepsilon. \end{cases} \tag{3.26}$$

where $\theta_\varepsilon \in C^2[-\varepsilon, \varepsilon]$, $\theta_\varepsilon' > 0$ on $(-\varepsilon, \varepsilon)$, $\theta_\varepsilon(0) = 0$, $\theta_\varepsilon(\varepsilon) = 1$, $\theta_\varepsilon(-\varepsilon) = -1$, and $\theta_\varepsilon'(\varepsilon) = 0$, $\theta_\varepsilon'(-\varepsilon) = 0$. The function γ_ε is a smooth monotonically increasing approximation of the signum multivalued function,

$$\operatorname{sign} r = \begin{cases} 1 & \text{for } r > 0,\\ [-1, 1] & \text{for } r = 0,\\ -1 & \text{for } r < 0, \end{cases}$$

we invoke frequently in the following.

If $[u_i, v_i] \in A,\ i = 1,2$, then we have, via Greens' formula,

$$\int_\Omega (v_1 - v_2)\gamma_\varepsilon(u_1 - u_2)dx = \int_\Omega |\nabla(u_1 - u_2)|^2 \gamma'_\varepsilon(u_1 - u_2)dx + \int_\Omega (\beta(u_1) - \beta(u_2))\gamma_\varepsilon(u_1 - u_2)dx \geq 0, \qquad \forall \varepsilon > 0.$$

For $\varepsilon \to 0,\ \gamma_\varepsilon(u_1 - u_2) \to g$ in $L^\infty(\Omega)$, where $g \in J(u)\|u\|^{-1}_{L^1(\Omega)}$, $u = u_1 - u_2$; that is, $g(x) \in \operatorname{sign} u(x)$, a.e. $x \in \Omega$. Hence, A is accretive.

We prove that A is m-accretive, considering separately the cases $1 < p < \infty$ and $p = 1$.

Case 1. $1 < p < \infty$. Let us denote for $1 < p < \infty$ by A_p the operator $-\Delta$ with the domain $D(A_p) = W_0^{1,p}(\Omega) \cap W^{2,p}(\Omega)$. We have already seen that A_p is accretive in $L^p(\Omega)$. Moreover, by Theorem 1.14, we have that $R(I + A_p) = L^p(\Omega)$ and

$$\|u\|_{W^{2,p}(\Omega)\cap W_0^{1,p}(\Omega)} \leq C\|A_p u\|_{L^p(\Omega)}, \qquad \forall u \in D(A_p). \tag{3.27}$$

Hence, A_p is m-accretive $L^p(\Omega)$. Let us prove now that $R(I + A_p + \widetilde{\beta}) = L^p(\Omega)$. Replacing, if necessary, the graph β by $u \to \beta(u) - v_0$, where $v_0 \in \beta(0)$, we may assume that $0 \in \widetilde{\beta}(0)$ and so $\widetilde{\beta}_\lambda(0) = 0$. Then, by Green's formula, for all $\lambda > 0$,

$$\begin{aligned}(A_p u, J(\widetilde{\beta}_\lambda u)) &= -\|\widetilde{\beta}_\lambda(u)\|^{2-p}_{L^p(\Omega)} \int_\Omega \Delta u |\beta_\lambda(u)|^{p-2}\beta_\lambda(u)dx \\ &= \|\widetilde{\beta}(u)\|^{2-p}_{L^p(\Omega)} \int_\Omega |\nabla u|^2 \frac{d}{du}|\beta_\lambda(u)|^{p-2}\beta_\lambda(u)dx \geq 0,\end{aligned} \tag{3.28}$$

and so, by Proposition 3.8, we conclude that $R(I + A_p + \widetilde{\beta}) = L^p(\Omega)$, as claimed.

To prove the surjectivity of $A_p + \widetilde{\beta}$, consider the equation

$$\varepsilon u + A_p u + \widetilde{\beta}(u) \ni f, \qquad \varepsilon > 0,\ f \in L^p(\Omega), \tag{3.29}$$

which, as seen before, has a unique solution u_ε, and $u_\varepsilon = \lim_{\lambda\to 0} u^\varepsilon_\lambda$ in $L^p(\Omega)$, where u^ε_λ is the solution to the approximating equation $\varepsilon u + A_p u + \widetilde{\beta}_\lambda(u) \ni f$. By (3.28), it follows that $\|A_p u^\varepsilon_\lambda\|_{L^p(\Omega)} \leq C$, where C is independent of ε and λ. Hence, letting $\lambda \to 0$, we get $\|A_p u_\varepsilon\|_{L^p(\Omega)} \leq C$, $\forall \varepsilon > 0$, which, by estimate (3.27) implies that $\{u_\varepsilon\}$ is bounded in $W^{1,p}(\Omega) \cap W^{2,p}(\Omega)$. Selecting a subsequence, for simplicity again denoted u_ε, we may assume that

$$\begin{aligned} u_\varepsilon &\to u && \text{weakly in } W^{2,p}(\Omega), \text{ strongly in } L^p(\Omega), \\ A_p u_\varepsilon &\to A_p u && \text{weakly in } L^p(\Omega), \\ \widetilde{\beta}_\varepsilon(u_\varepsilon) &\to g && \text{weakly in } L^p(\Omega). \end{aligned}$$

By Proposition 3.4 we know that $g \in \widetilde{\beta}(u)$, therefore we infer that u is the solution to the equation $A_p u + \widetilde{\beta}(u) \ni f$; that is, $u \in W^{2,p}(\Omega)$ and

$$\begin{cases} -\Delta u + \beta(u) \ni f, & \text{a.e. in } \Omega, \\ u = 0 & \text{in } \partial\Omega. \end{cases} \tag{3.30}$$

Case 2. $p = 1$. We prove directly that $R(A_1 + \widetilde{\beta}) = L^1(\Omega)$ that is, for $f \in L^1(\Omega)$, equation (3.30) has a solution $u \in D(A_1) = \{u \in W_0^{1,1}(\Omega);\ \Delta u \in L^1(\Omega)\}$. (Here, $A_1 = -\Delta$ with the domain $D(A_1)$.)

We fix f in $L^1(\Omega)$ and consider $\{f_n\} \subset L^2(\Omega)$ such that $f_n \to f$ in $L^1(\Omega)$. As seen before, the problem

$$\begin{cases} -\Delta u_n + \beta(u_n) \ni f_n & \text{in } \Omega, \\ u_n = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.31}$$

has a unique solution $u_n \in H_0^1(\Omega) \cap H^2(\Omega)$. Let $v_n(x) = f_n(x) + \Delta u_n(x) \in \beta(u_n(x))$, a.e. $x \in \Omega$. By (3.31) we see that

$$\int_\Omega |v_n(x) - v_m(x)|dx \le \int_\Omega |f_n(x) - f_m(x)|dx,$$

because β is monotone and $-\Delta$ is accretive in $L^1(\Omega)$; that is, $\int_\Omega \Delta u \theta\, dx \le 0$, $\forall u \in D(A_1)$, for some $\theta \in L^\infty(\Omega)$ such that $\theta(x) \in \operatorname{sign} u(x)$, a.e. $x \in \Omega$. (It suffices to check the latter for $\theta = \gamma_\varepsilon(u)$ where γ_ε is given by (3.26) because, by density, it extends to all of $D(A_1)$.) Hence,

$$\begin{aligned} v_n &\to v && \text{strongly in } L^1(\Omega), \\ \Delta u_n &\to \xi && \text{strongly in } L^1(\Omega). \end{aligned} \tag{3.32}$$

Now, let $h_i \in L^p(\Omega),\ i = 0, 1, \dots, N,\ p > N$. Then, by a well-known result due to G. Stampacchia [12] (see also Dautray and Lions [9], p. 462), the boundary value problem

$$\begin{cases} -\Delta \varphi = h_0 + \displaystyle\sum_{i=1}^N \frac{\partial h_i}{\partial x_i} & \text{in } \Omega, \\ \varphi = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.33}$$

has a unique weak solution $\varphi \in H_0^1(\Omega) \cap L^\infty(\Omega)$ and

$$\|\varphi\|_{L^\infty(\Omega)} \le C \sum_{i=0}^N \|h_i\|_{L^p(\Omega)}, \qquad h_i \in L^p(\Omega). \tag{3.34}$$

This means that

$$\int_\Omega \nabla\varphi\cdot\nabla\psi\,dx = \int_\Omega h_0\psi - \sum_{i=1}^N \int_\Omega h_i \frac{\partial\psi}{\partial x_i}\,dx, \qquad \forall\psi\in H_0^1(\Omega). \tag{3.35}$$

Substituting $\psi = u_n$ in (3.35), we get, via Green's formula,

$$-\int_\Omega \varphi\Delta u_n dx = \int_\Omega \nabla\varphi\cdot\nabla u_n dx = \int_\Omega h_0 u_n dx - \sum_{i=1}^N \int_\Omega h_i \frac{\partial u_n}{\partial x_i}\,dx,$$

and, therefore, by (3.34),

$$\left| \int_\Omega h_0 u_n dx - \sum_{i=1}^N h_i \frac{\partial u_n}{\partial x_i}\,dx \right| \le C\|\Delta u_n\|_{L^1(\Omega)} \sum_{i=0}^N \|h_i\|_{L^p(\Omega)}.$$

Because $\{h_i\}_{i=0}^N \subset (L^p(\Omega))^{N+1}$ are arbitrary, we conclude that the sequence

$$\left\{ \left(u_n, \frac{\partial u_n}{\partial x_1}, \dots, \frac{\partial u_n}{\partial x_N} \right) \right\}_{n=1}^\infty$$

is bounded in $(L^q(\Omega))^{N+1}$, $1/p + 1/q = 1$. Hence,

$$\|u_n\|_{W^{1,q}(\Omega)} \le C\|\Delta u_n\|_{L^1(\Omega)}, \quad \text{where } 1 < q = \frac{p}{p-1} < \frac{N}{N-1}. \tag{3.36}$$

Therefore, $\{u_n\}$ is bounded in $W^{1,q}(\Omega)$ and, consequently, compact in $L^1(\Omega)$. Then, extracting a further subsequence if necessary, we may assume that

$$u_n \to u \quad \text{weakly in } W_0^{1,q}(\Omega) \text{ and strongly in } L^1(\Omega). \tag{3.37}$$

Then, by (3.32), it follows that $\xi = \Delta u$, and because the operator $\widetilde{\beta}$ is closed in $L^1(\Omega)\times L^1(\Omega)$, we see by (3.32) and (3.37) that $v(x)\in\beta(u(x))$, a.e. $x\in\Omega$, and $u\in W_0^{1,q}(\Omega)$. Hence $R(A) = L^1(\Omega)$ and, in particular, A is m-accretive. □

We have proved, therefore, the following existence result for the semilinear elliptic boundary value problem in $L^1(\Omega)$.

Corollary 3.1. *For every $f\in L^p(\Omega)$, $1<p<\infty$, the boundary value problem*

$$\begin{cases} -\Delta u + \beta(u) \ni f, & \text{a.e. in } \Omega, \\ u = 0 & \text{on } \partial\Omega, \end{cases} \tag{3.38}$$

has a unique solution $u\in W_0^{1,p}(\Omega)\cap W^{2,p}(\Omega)$. If $L^1(\Omega)$, then $u\in W_0^{1,q}(\Omega)$ with $\Delta u\in L^1(\Omega)$, where $1\le q<N/(N-1)$. Moreover, the following estimate holds:

$$\|u\|_{W_0^{1,q}(\Omega)} \le C\|f\|_{L^1(\Omega)}, \qquad \forall f\in L^1(\Omega). \tag{3.39}$$

In particular, A_1 is m-accretive in $L^1(\Omega)$, $D(A_1)\subset W_0^{1,q}(\Omega)$, and

$$\|u\|_{W_0^{1,q}(\Omega)} \leq C\|\Delta u\|_{L^1(\Omega)}, \qquad \forall u \in D(A_1).$$

Remark 3.3. It is clear from the previous proof that Theorem 3.2 and Corollary 3.1 remain true for more general linear second-order elliptic operators A_p on Ω.

The Semilinear Elliptic Operator in $L^1(\mathbf{R}^N)$

The previous results partially extend to unbounded domains Ω. Below we treat the case $\Omega = \mathbf{R}^N$.

Let β be a maximal monotone graph in $\mathbf{R} \times \mathbf{R}$ such that $0 \in \beta(0)$ and let $A : L^1(\mathbf{R}^N) \to L^1(\mathbf{R}^N)$ be the operator

$$Au = -\Delta u + \widetilde{\beta}(u), \qquad \forall u \in D(A), \quad \text{in } \mathscr{D}'(\mathbf{R}^N), \tag{3.40}$$

where

$$\begin{aligned} D(A) &= \{u \in L^1(\mathbf{R}^N),\ \Delta u \in L^1(\mathbf{R}^N);\ u \in D(\widetilde{\beta})\},\\ D(\widetilde{\beta}) &= \{u \in L^1(\mathbf{R}^N);\ \exists \eta \in L^1(\mathbf{R}^N),\ \eta(x) \in \beta(u(x)),\ \text{a.e. } x \in \mathbf{R}^N\},\\ \widetilde{\beta}(u) &= \{\eta \in L^1(\mathbf{R}^N);\ \eta(x) \in \beta(u(x)),\ \text{a.e. } x \in \mathbf{R}^N\}. \end{aligned} \tag{3.41}$$

Here Δu is taken in the sense of distributions on $\mathbf{R}^N$; that is,

$$\Delta u(\varphi) = \int_{\mathbf{R}^n} u\Delta\varphi\, dx, \qquad \forall \varphi \in C_0^\infty(\mathbf{R}^N),$$

and the equation $Au = f$ is taken in the following distributional sense

$$\int_{\mathbf{R}^N} (-u\Delta\varphi + \eta\varphi)dx = \int_{\mathbf{R}^n} f\varphi\, dx, \qquad \forall \varphi \in C_0^\infty(\mathbf{R}^N),$$

where $\eta \in L^1(\mathbf{R}^N)$ is such that $\eta(x) \in \beta(u(x))$ a.e. $x \in \mathbf{R}^N$.

Theorem 3.3. *The operator A defined by equations* (3.40) *and* (3.41) *is m-accretive in* $L^1(\mathbf{R}^N) \times L^1(\mathbf{R}^N)$.

Proof. We fix $f \in L^1(\mathbf{R}^N)$ and consider the equation $\lambda u + Au \ni f$; that is,

$$\lambda u - \Delta u + \beta(u) \ni f \quad \text{in } \mathbf{R}^N, \tag{3.42}$$

which is taken in the above distributional sense. We prove that for each $\lambda > 0$ there is a unique solution $u = u(f)$ and that

$$\|u(f) - u(g)\|_{L^1(\mathbf{R}^n)} \leq \frac{1}{\lambda}\|f - g\|_{L^1(\mathbf{R}^N)}, \qquad \forall f, g \in L^1(\mathbf{R}^N). \tag{3.43}$$

To this end we consider the approximating equation

$$\lambda u_\varepsilon - \Delta u_\varepsilon + \beta_\varepsilon(u_\varepsilon) = f \quad \text{in } \mathscr{D}'(\mathbf{R}^N), \tag{3.44}$$

where $\beta_\varepsilon = (1-(1+\varepsilon\beta)^{-1})/\varepsilon$, $\forall \varepsilon > 0$.

We rewrite (3.44) as

$$\lambda u_\varepsilon - \Delta u_\varepsilon + \frac{1}{\varepsilon} u_\varepsilon = f + \frac{1}{\varepsilon}(1+\varepsilon\beta)^{-1} u_\varepsilon.$$

Equivalently,

$$u_\varepsilon - \frac{\varepsilon}{1+\varepsilon\lambda}\Delta u_\varepsilon = \frac{\varepsilon}{1+\varepsilon\lambda} f + \frac{1}{1+\varepsilon\lambda}(1+\varepsilon\beta)^{-1} u_\varepsilon. \tag{3.45}$$

On the other hand, it is well known that, for each $g \in L^1(\mathbf{R}^N)$ and constant $\mu > 0$, the equation

$$v - \mu\Delta v = g \quad \text{in } \mathscr{D}'(\mathbf{R}^N)$$

has a unique solution $v \in L^1(\mathbf{R}^N)$ and

$$\|v\|_{L^1(\mathbf{R}^N)} \le \|g\|_{L^1(\mathbf{R}^N)}.$$

(This means that the operator $A_1 = -\Delta$ is m-accretive in $L^1(\mathbf{R}^N)$.) If we set $v = T_\mu g$, we may rewrite (3.45) as

$$u_\varepsilon = T_{\varepsilon/(1+\varepsilon\lambda)}\left(\frac{\varepsilon}{1+\varepsilon\lambda} f + \frac{1}{1+\varepsilon\lambda}(1+\varepsilon\beta)^{-1} u_\varepsilon\right),$$

and so, by the Banach fixed point theorem, it follows the existence of a unique solution $u_\varepsilon \in L^1(\mathbf{R}^N)$ to (3.45). Moreover, as easily seen, we have

$$\|u_\varepsilon\|_{L^1(\mathbf{R}^N)} \le \frac{1}{\lambda}\|f\|_{L^1(\mathbf{R}^N)}. \tag{3.46}$$

We have also

$$\|\beta_\varepsilon(u_\varepsilon)\|_{L^1(\mathbf{R}^N)} \le \|f\|_{L^1(\mathbf{R}^N)}, \qquad \forall \varepsilon > 0. \tag{3.47}$$

Formally, (3.47) follows by multiplying (3.44) by sign $\beta_\varepsilon(u_\varepsilon)$ and integrating on $\mathbf{R}^N$. However, in order to prove it rigorously, we assume first that $f \in L^1(\mathbf{R}^N) \cap L^2(\mathbf{R}^N)$ and get the desired inequality by density argument. Indeed, in this case the solution u_ε to (3.44) belongs to $H^2(B_R) \cap H^1(\mathbf{R}^N)$ on each ball $B_R \subset \mathbf{R}^N$ of radius R and center 0 (see Theorem 1.10). Let $\rho \in C_0^\infty(R)$ be such that $\rho > 0$, $\rho(r) = 1$ for $0 \le r \le 1$ and $\rho(r) = 0$ for $r \ge 2$ and let $\varphi_R(x) = \rho\left(|x|^2/R^2\right)$. Finally, let $\chi = \gamma_\varepsilon$ be the function (3.26). Then, multiplying equation (3.44) by $\varphi_R\chi(\beta_\varepsilon(u_\varepsilon))$ and integrating on $\mathbf{R}^N$ (in fact on B_{2R}) we see that

$$\begin{aligned} &\lambda\int_{B_{2R}} u_\varepsilon\chi(\beta_\varepsilon(u_\varepsilon))\varphi_R dx + \int_{B_{2R}} \nabla u_\varepsilon\cdot\nabla(\varphi_R\chi(\beta_\varepsilon(u_\varepsilon))dx \\ &+ \int_{B_{2R}} \varphi_R\beta_\varepsilon(u_\varepsilon)\chi(\beta_\varepsilon(u_\varepsilon))dx = \int_{B_{2R}} f\varphi_R\, dx,\ \forall R. \end{aligned} \tag{3.48}$$

Keeping in mind that $\int_{B_R} \nabla u_\varepsilon \cdot \nabla(\chi(\beta_\varepsilon(u_\varepsilon))\varphi_R dx \geq 0$ and that $\varphi_R = 1$ on $[|x| < R]$, we see by (3.48)

$$\begin{aligned} &\lambda \int_{B_{2R}} u_\varepsilon \chi(\beta_\varepsilon(u_\varepsilon))\varphi_R dx + \int_{B_{2R}\setminus B_R} (\nabla u_\varepsilon \cdot \nabla \varphi_R)\chi(\beta_\varepsilon(u_\varepsilon))dx \\ &+ \int_{B_{2R}} \varphi_R \beta_\varepsilon(u_\varepsilon)\chi(\beta_\varepsilon(u_\varepsilon))dx \leq \int_{B_{2R}} f\varphi_R\, dx, \ \ \forall R. \end{aligned} \tag{3.49}$$

On the other hand, multiplying (3.44) by u_ε and integrating on $\mathbf{R}^N$, we see that

$$\lambda \int_{\mathbf{R}^N} |u_\varepsilon|^2 dx + 2\int_{\mathbf{R}^N} |\nabla u_\varepsilon|^2 dx \leq \int_{\mathbf{R}^N} |f|^2 dx. \tag{3.50}$$

Then, letting $R \to \infty$ and $\chi \to$ sign into (3.49), we obtain (3.47), as claimed.

Note also that assuming $f \in L^1(\mathbf{R}^N) \cap L^2(\mathbf{R}^N)$ besides (3.50) we have the estimate

$$2\lambda \int_{\mathbf{R}^N} |u_\varepsilon|^2 dx + \int_{\mathbf{R}^N} |\beta_\varepsilon(u_\varepsilon)|^2 dx \leq \int_{\mathbf{R}^N} |f|^2 dx. \tag{3.51}$$

(The latter follows as above multiplying equation (3.44) by $\varphi_N \beta_\varepsilon(u_\varepsilon)$ and integrating on $\mathbf{R}^N$.)

Moreover, we have by (3.44) for all $\varepsilon, \varepsilon' > 0$,

$$\lambda(u_\varepsilon - u_{\varepsilon'}) - \Delta(u_\varepsilon - u_{\varepsilon'}) + \beta_\varepsilon(u_\varepsilon) - \beta_{\varepsilon'}(u_{\varepsilon'}) = 0 \quad \text{in } \ \mathbf{R}^N$$

and we get, as above, that

$$\begin{aligned} &\lambda \int_{\mathbf{R}^N} |u_\varepsilon - u_{\varepsilon'}|^2 dx \\ &\leq \int_{\mathbf{R}^N} (\varepsilon|\beta_\varepsilon(u_\varepsilon)| + \varepsilon'|\beta_{\varepsilon'}(u_{\varepsilon'})|)(|\beta_\varepsilon(u_\varepsilon)| + |\beta_{\varepsilon'}(u_{\varepsilon'})|)dx, \qquad \forall \varepsilon, \varepsilon' > 0 \end{aligned}$$

because $(\beta_\varepsilon(u_\varepsilon) - \beta_{\varepsilon'}(u_{\varepsilon'}))(u_\varepsilon - u_{\varepsilon'}) \geq (\beta_\varepsilon(u_\varepsilon) - \beta_{\varepsilon'}(u_{\varepsilon'}))(\varepsilon\beta_\varepsilon(u_\varepsilon) - \varepsilon'\beta_{\varepsilon'}(u_{\varepsilon'}))$, $\forall \varepsilon, \varepsilon' \geq 0$.

By virtue of (3.51), this yields

$$\lambda \int_{\mathbf{R}^N} |u_\varepsilon - u_{\varepsilon'}|^2 dx \leq C(\varepsilon + \varepsilon'), \qquad \forall \varepsilon, \varepsilon' > 0. \tag{3.52}$$

Hence, on a subsequence, again denoted $\{\varepsilon\} \to 0$, we have

$$\begin{aligned} u_\varepsilon &\to u \qquad \text{strongly in } L^2(\mathbf{R}^N) \\ \beta_\varepsilon(u_\varepsilon) &\rightharpoonup \eta \qquad \text{in } L^2(\mathbf{R}^N) \\ \Delta u_\varepsilon &\rightharpoonup \Delta u \qquad \text{in } L^2(\mathbf{R}^N). \end{aligned} \tag{3.53}$$

Because β is maximal monotone, so is its realization $\widetilde{\beta} \subset L^2(\mathbf{R}^N) \times L^2(\mathbf{R}^N)$; that is,

$$\widetilde{\beta} = \{[u,v] \in L^2(\mathbf{R}^N) \times L^2(\mathbf{R}^N),\ v(x) \in \beta(u(x)), \quad \text{a.e. } x \in \mathbf{R}^N\}.$$

Then, by (3.53) it follows that $\eta(x) \in \beta(u(x))$, a.e. $x \in \mathbf{R}^N$. Moreover, by (3.49) and (3.50) we infer that $u, \eta \in L^1(\mathbf{R}^N)$ and $\Delta u = f - \eta \in L^1(\mathbf{R}^N)$. Hence (u, η) is a solution to (3.42).

If $f, g \in L^1(\mathbf{R}^N) \cap L^2(\mathbf{R}^N)$ and $u_\varepsilon(f), u_\varepsilon(g)$ are corresponding solutions to (3.44) we have

$$\begin{aligned}&\lambda \int_{\mathbf{R}^N} (u_\varepsilon(f) - u_\varepsilon(g))\varphi_R \chi(u_\varepsilon(f) - u_\varepsilon(g))\\&+ \int_{\mathbf{R}^N} \nabla(u_\varepsilon(f) - u_\varepsilon(g)) \cdot \nabla(\varphi_R \chi(u_\varepsilon(f) - u_\varepsilon(g)))dx\\&+ \int_{\mathbf{R}^N} (\beta_\varepsilon(u_\varepsilon(f)) - \beta_\varepsilon(u_\varepsilon(g)))\varphi_R \chi(u_\varepsilon(f) - u_\varepsilon(g))dx\\&+ \int_{\mathbf{R}^N} (f-g)\varphi_R \chi(u_\varepsilon(f) - u_\varepsilon(g))dx,\end{aligned}$$

where χ and φ_R are defined as above.

Letting $R \to \infty$ and $\chi \to$ sign we obtain that

$$\lambda \int_{\mathbf{R}^N} |u_\varepsilon(f) - u_\varepsilon(g)|dx \le \int_{\mathbf{R}^N} |f-g|dx$$

and for $\varepsilon \to 0$ we get (3.43); that is, $\|u(f) - u(g)\|_{L^1(\mathbf{R}^N)} \le (1/\lambda)\|f-g\|_{L^1(\mathbf{R}^N)}$. This implies by density that $u = u(f)$ extends as a solution to equation (3.41) for all $f \in L^1(\mathbf{R}^N)$.

It remains to prove the uniqueness. If u_1, u_2 are two solutions to (3.42), we have

$$\lambda(u_1 - u_2) - \Delta(u_1 - u_2) + \eta_1 - \eta_2 = 0 \quad \text{in } \mathscr{D}'(\mathbf{R}^N), \tag{3.54}$$

where $u_i, \eta_i \in L^1(\mathbf{R}^N)$ and $\eta_i \in \beta(u_i)$, a.e. in $\mathbf{R}^N$ for $i = 1, 2$.

We set $u = u_1 - u_2$ and take $u_\delta = u * \rho_\delta$ where ρ_δ is a C_0^∞ mollifier and $*$ stands for convolution product. We have

$$\lambda u_\delta - \Delta u_\delta + (\eta_1 - \eta_2) * \rho_\delta = 0 \quad \text{in } \mathbf{R}^N. \tag{3.55}$$

It follows, of course, that $u_\delta, (\eta_1 - \eta_2) * \rho_\delta \in L^1(\mathbf{R}^N)$ and $u_\delta \in H^1(\mathbf{R}^N)$ because, as easily seen,

$$\begin{aligned}\|u_\delta\|_{L^2(\mathbf{R}^N)} &\le \|\rho_\delta\|_{L^2(\mathbf{R}^N)} \|u\|_{L^1(\mathbf{R}^N)},\\ \|\nabla u_\delta\|_{L^2(\mathbf{R}^N)} &\le \|\nabla \rho_\delta\|_{L^2(\mathbf{R}^N)} \|u\|_{L^1(\mathbf{R}^N)}.\end{aligned} \tag{3.56}$$

Then, multiplying (3.55) by $\zeta(u_\delta)$, where $\zeta = \gamma_\varepsilon$ as above is a smooth approximation of the signum function (see (3.26)), we obtain

$$\lambda \int_{\mathbf{R}^N} u_\delta \zeta(u_\delta)dx + \int_{\mathbf{R}^N} ((\eta_1 - \eta_2) * \rho_\delta)\zeta(u_\delta)dx \le 0$$

and, letting $\zeta \to$ sign, we get

$$\lambda \int_{\mathbf{R}^N} |u_\delta(x)|dx + \liminf_{\delta \to 0} \int_{\mathbf{R}^N} ((\eta_1 - \eta_2) * \rho_\delta)\operatorname{sgn} u_\delta dx \le 0, \qquad \forall \delta > 0.$$

Taking into account that by the monotonicity of β, we have that $\operatorname{sgn}(\eta_1 - \eta_2) = \operatorname{sgn} u$, a.e. in $\mathbf{R}^N$, this yields

$$\liminf_{\delta \to 0} \int_{\mathbf{R}^N} ((\eta_1 - \eta_2) * \rho_\delta)(x) \operatorname{sgn} u_\delta(x) dx \geq 0.$$

Hence, $u_\delta \to 0$ as $\delta \to 0$ and this implies $u_1 = u_2$, as claimed. This completes the proof of Theorem 3.3. □

One might expect that for $\lambda \to 0$ the solution $u = y_\lambda$ to equation (3.42) is convergent (in an appropriate space) to a solution $y \in L^1_{\text{loc}}(\mathbf{R}^N)$ to equation

$$-\Delta y + \beta(y) \ni f \quad \text{in } \mathscr{D}'(\mathbf{R}^N). \tag{3.57}$$

It turns out that this is indeed the case and that equation (3.57) has a unique solution. More precisely, one has the following existence result due to Bénilan, Brezis and Crandall [3].

Theorem 3.4. *Assume that $f \in L^1(\mathbf{R}^N)$. Then,*

(i) *If $N = 1$ and $0 \in \operatorname{int} R(\beta)$, then equation (3.57) has a unique solution $y \in W^{1,\infty}(\mathbf{R})$ with $\Delta y \in L^1(\mathbf{R})$.*
(ii) *If $N = 2$ and $0 \in \operatorname{int} R(\beta)$, then there is a unique solution $y \in L^1_{\text{loc}}(\mathbf{R}^2) \cap W^{1,1}_{\text{loc}}(\mathbf{R}^2)$ with $\Delta y \in L^1(\mathbf{R}^2)$ and $\nabla y \in M^2(\mathbf{R}^2)$.*
(iii) *If $N \geq 3$, then there is a unique solution $y \in M^{N/(N-2)}(\mathbf{R}^N) \cap L^1_{\text{loc}}(\mathbf{R}^N)$ with $\Delta y \in L^1(\mathbf{R}^N)$.*

$R(\beta)$ is the range of β and $M^p(\mathbf{R}^N)$, $p \geq 1$, is the Marcinkiewicz class of order p; that is,

$$M^p(\mathbf{R}^N) = \Big\{ u : \mathbf{R}^N \to \mathbf{R} \text{ measurable}, \\ \min_{E \subset \mathbf{R}^N} \Big\{ \alpha \in \mathbf{R}^+ ; \int_E |u(x)| dx \leq \alpha (\operatorname{meas} E)^{1/q} \Big\} = \|u\|_M < \infty \Big\}, \\ \text{where } \frac{1}{p} + \frac{1}{q} = 1.$$

Proof. (Sketch) We are going to pass to the limit $\lambda \to 0$ in equation (3.42); that is,

$$\lambda y_\lambda - \Delta y_\lambda + \beta(y_\lambda) \ni f. \tag{3.42$'$}$$

The main problem is, however, the boundedness of $\{y_\lambda\}$ in $L^1(\mathbf{R}^N)$ or in $L^1_{\text{loc}}(\mathbf{R}^N)$.

We set $w_\lambda = \beta(y_\lambda)$ (or the section of it arising in $(3.42)'$ if β is multivalued).

We see that

$$\lambda \int_{\mathbf{R}^N} |y_\lambda(x+h) - y_\lambda(x)| dx + \int_{\mathbf{R}^N} |w_\lambda(x+h) - w_\lambda(x)| dx \leq \int_{\mathbf{R}_N} |f(x+h) - f(x)| dx, \ \forall h,$$

and

$$\int_{\mathbf{R}^N} |w_\lambda(x)|dx \le \int_{\mathbf{R}^N} |f(x)|dx. \tag{3.58}$$

Hence, by the Kolmogorov compactness theorem, $\{w_\lambda\}$ is compact in $L^1_{\text{loc}}(\mathbf{R}^N)$and so, there is $w \in L^1_{\text{loc}}(\mathbf{R}^N)$ such that, as $\lambda \to 0$,

$$w_\lambda \to w \quad \text{in } L^1_{\text{loc}}(\mathbf{R}^N). \tag{3.59}$$

On the other hand, by (3.58) and by Fatou's lemma, it follows that $w \in L^1(\mathbf{R}^N)$.

This implies that $\Delta y_\lambda = \lambda y_\lambda + w_\lambda - f$ is bounded in $L^1(\mathbf{R}^N)$ and so, if $N \ge 3$, we have (see [3])

$$\|y_\lambda\|_{M^{N/(N-2)}(\mathbf{R}^N)} + \|\nabla y_\lambda\|_{M^{N/(N-1)}(\mathbf{R}^N)} \le C, \qquad \forall \lambda > 0.$$

In particular, it follows that $\{y_\lambda\}$ is bounded in $W^{1,1}_{\text{loc}}(\mathbf{R}^N)$ and so $\{y_\lambda\}$ is compact in $L^1_{\text{loc}}(\mathbf{R}^N)$. Then, on a subsequence, $y_\lambda \to y$ in $L^1_{\text{loc}}(\mathbf{R}^N)$ and by (3.59), we infer that $w(x) = \beta(y(x))$, a.e. $x \in \mathbf{R}^N$. Clearly, y is a solution to (3.57) because $\Delta y_\lambda \to \Delta y$ in $\mathscr{D}'(\mathbf{R}^N)$ as $\lambda \to 0$.

We now consider the following.

The case $N = 2$. In this case, in order to get the boundedness of $\{y_\lambda\}$, one must assume further that $0 \in \operatorname{int} R(\beta)$. If we denote by $j : \mathbf{R} \to \overline{\mathbf{R}}$ the potential of β (i.e., $\beta = \partial j$), we have that $j(r) \ge c|r|$, for some $c > 0$ and $|r| \ge R_1$.

Indeed, as seen earlier (Proposition 1.5), $\operatorname{int} R(\beta) = \operatorname{int} D(\beta^{-1}) = \operatorname{int} D(j^*)$, where j^* is the conjugate of j:

$$j(r) = \sup\{rp - j^*(p),\ \forall p \in \mathbf{R}\}.$$

We have therefore $|j^*(p)| \le C$ for all $p \in \mathbf{R}$, $|p| \le r^*$, where $r^* > 0$ is suitably chosen. This yields

$$j(r) \ge \rho|r| - j^*\left(\rho\,\frac{r}{|r|}\right) \ge \frac{r^*}{2}\,|r| \quad \text{for } |r| \ge 1.$$

Now, we come back to equation (3.42) and notice that multiplying by $\operatorname{sign} y_\lambda$ we get as above

$$\int_{[|y_\lambda|>1]} \frac{\beta(y_\lambda)y_\lambda}{|y_\lambda|}\,dx \le \int_\Omega |f|dx$$

and taking into account that $\beta(y_\lambda)y_\lambda \ge j(y_\lambda) \ge c|y_\lambda|$ on $[|y_\lambda| \ge 1]$ we get

$$\int_{\mathbf{R}^N} |y_\lambda(x)|dx \le c, \qquad \forall \lambda > 0$$

and therefore $\{y_\lambda\}$ is bounded in $L^1(\mathbf{R}^N)$. Then, by the equation $\Delta y_\lambda = \lambda y_\lambda + w_\lambda - f$ and, by Lemma A.14 in [3], we infer that $\{\nabla y_\lambda\}$ is bounded in $M^2(\mathbf{R}^2)$. This implies that $y = \lim_{\lambda \downarrow 0} y_\lambda$ exists (on a subsequence) in $L^1_{\text{loc}}(\mathbf{R}^2)$ and also that $\nabla y \in M^2(\mathbf{R}^2)$. Then, by (3.59), we see that $w(x) \in \beta(y(x))$, a.e. $x \in \Omega$, and so y is the desired solution.

The case $N = 1$. It follows as above that $\{y_\lambda\}$ and $\{\beta_\lambda(y_\lambda)\}$ are bounded in $L^1(\mathbf{R}^N)$ and, because $\{y''_\lambda\}$ is bounded in $L^1(\mathbf{R})$, we also get that $\{y'_\lambda\}$ is bounded in $L^\infty(\mathbf{R})$. In fact, because $\{y'_\lambda\}$ is bounded in $L^1(\mathbf{R})$, then there is at least one $x_0 \in \mathbf{R}$ such that $\{y'_\lambda(x_0)\}$ is bounded and this, clearly, implies that $\{y'_\lambda\}$ is bounded in $L^\infty(\mathbf{R})$. Then we infer, as in the previous cases, that $y = \lim_{\lambda\downarrow 0} y_\lambda$ is the solution to (3.57) and satisfies the required conditions. The details are omitted. □

The Porous Media Equation in $L^1(\Omega)$

We have already studied this equation in the $H^{-1}(\Omega)$ space framework in Section 2.2. Here, we consider this equation in the L^1 space framework.

In the space $X = L^1(\Omega)$ define the operator

$$\begin{cases} Au = -\Delta\beta(u), & \forall u \in D(A), \\ D(A) = \{u \in L^1(\Omega);\ \beta(u) \in W_0^{1,1}(\Omega),\ \Delta\beta(u) \in L^1(\Omega)\}, \end{cases} \tag{3.60}$$

where β is a maximal monotone graph in $\mathbf{R}\times\mathbf{R}$ such that $0 \in \beta(0)$ and Ω is an open bounded subset of $\mathbf{R}^N$ with smooth boundary. More precisely, $A \subset L^1(\Omega)\times L^1(\Omega)$ is defined by

$$A = \{[u, -\Delta\eta], u \in L^1(\Omega),\ \eta \in W_0^{1,1}(\Omega), \Delta\eta \in L^1(\Omega),\ \eta(x) \in \beta(u(x)), \\ \text{a.e. } x \in \Omega\}. \tag{3.61}$$

We have the following.

Theorem 3.5. *The operator A is m-accretive in* $L^1(\Omega)\times L^1(\Omega)$.

Proof. Let $u, v \in D(A)$ and let γ be a smooth monotone approximation of the sign of the form considered earlier. (See (3.26).) Then, we have

$$\int_\Omega (Au - Av)\gamma(\beta(u) - \beta(v))dx = \int_\Omega |\nabla(\beta(u) - \beta(v))|^2 \gamma'(\beta(u) - \beta(v))dx \ge 0.$$

Letting $\gamma \to$ sign, we get

$$\int_\Omega (Au - Av)\xi\, dx \ge 0,$$

where $\xi(x) \in \text{sign}(\beta(u(x)) - \beta(v(x))) = \text{sign}(u(x) - v(x))$, a.e. $x \in \Omega$. Hence, A is accretive.

Let us prove now that $R(I + A) = L^1(\Omega)$. For $f \in L^1(\Omega)$, the equation

$$u + Au = f$$

can be equivalently written as

$$\beta^{-1}(v) - \Delta v = f \quad \text{in } \Omega,\ \ v \in W_0^{1,1}(\Omega),\ \Delta v \in L^1(\Omega). \tag{3.62}$$

But, according to Corollary 3.1, equation (3.62) has a solution $v \in W_0^{1,q}(\Omega)$, $\Delta v \in L^1(\Omega)$, $1 < q < N/(N-1)$. □

The Porous Media Equation in $\mathbf{R}^N$

Consider the equation

$$\lambda y(x) - \Delta\beta(y(x)) \ni f(x) \quad \text{in } \mathbf{R}^N, \tag{3.63}$$

where $\lambda > 0$, and β is a maximal monotone graph in $\mathbf{R}\times\mathbf{R}$ such that $0 \in \beta(0)$. By solution y to (3.63) we mean a function $y \in L^1(\mathbf{R}^N)$ such that $\exists \eta \in L^1_{\rm loc}(\mathbf{R}^N)$, $\eta(x) \in \beta(y(x))$, a.e. $x \in \mathbf{R}^N$, and

$$\lambda y - \Delta\eta = f \quad \text{in } \mathscr{D}'(\mathbf{R}^N). \tag{3.64}$$

Theorem 3.6. *Assume that $f \in L^1(\mathbf{R}^N)$. Then,*

(i) *If $N = 1$ and $0 \in \operatorname{int} D(\beta)$, then there is a unique solution $y \in L^1(\mathbf{R}^N)$ with $\eta \in L^1_{\rm loc}(\mathbf{R}) \cap W^{1,\infty}(\mathbf{R})$.*
(ii) *If $N = 2$ and $0 \in \operatorname{int} D(\beta)$, then there is a unique solution $y \in L^1(\mathbf{R}^N)$ with $\eta \in W^{1,1}_{\rm loc}(\mathbf{R}^2)$, $|\nabla\eta| \in M^2(\mathbf{R}^2)$.*
(iii) *If $N \geq 3$, then there is a unique solution $y \in L^1(\mathbf{R}^N)$, with $\eta \in M^{N/(N-2)}(\mathbf{R}^N)$.*

Proof. By substitution, $\beta(y) \to u$, equation (3.63) reduces to equation (3.57) with β^{-1} in the place of β and so, one can apply Theorem 3.4 to derive (i) $\sim$ (iii).

In the space $L^1(\mathbf{R}^N)$ consider the operator

$$Ay = -\Delta\beta(y), \qquad \forall y \in D(A) \tag{3.65}$$

defined by

$$\begin{aligned} D(A) = \{y \in L^1(\mathbf{R}^N);\ \exists \eta \in L^1_{\rm loc}(\mathbf{R}^N),\\ \eta(x) \in \beta(y(x)),\ \text{a.e. } x \in \Omega,\ \Delta\eta \in L^1(\mathbf{R}^N)\} \end{aligned} \tag{3.66}$$

$$Ay = \{-\Delta\eta \in L^1(\mathbf{R}^N);\ \eta \in \beta(y),\ \text{a.e. in } \mathbf{R}^N,\ \eta \in L^1_{\rm loc}(\mathbf{R}^N),\ y \in L^1(\mathbf{R}^N)\}. \tag{3.67}$$

□

We have the following.

Theorem 3.7. *Assume that β is a maximal monotone graph satisfying the conditions of Theorem 3.6. Then the operator A defined by (3.66) and (3.67) is m-accretive in $L^1(\mathbf{R}^N) \times L^1(\mathbf{R}^N)$.*

Proof. There is nothing left to do, except to apply Theorem 3.6 and to notice that by Theorem 3.3 we have also the accretivity inequality

$$\|u-v\|_{L^1(\mathbf{R}^N)} \le \frac{1}{\lambda}\,\|f-g\|_{L^1(\mathbf{R}^N)}$$

if u, v are solutions to (3.63) for f and g, respectively. □

3.3 Quasilinear Partial Differential Operators of First Order

Here, we study the first-order partial differential operator

$$(Au)(x) = \sum_{i=1}^{N} \frac{\partial}{\partial x_i} a_i(u(x)), \qquad x \in \mathbf{R}^N, \tag{3.68}$$

in the space $X = L^1(\mathbf{R}^N)$. We use the notations $a = (a_1, a_2, \dots, a_N)$, $\varphi_x = (\varphi_{x_1}, \dots, \varphi_{x_N})$, $a(u)_x = \sum_{i=1}^N (\partial/\partial x_i) a_i(u(x)) = \operatorname{div} a(u)$.

The function $a : \mathbf{R} \to \mathbf{R}^N$ is assumed to be continuous.

We define the operator A in $L^1(\mathbf{R}^N) \times L^1(\mathbf{R}^N)$ as the closure of the operator $A_0 \subset L^1(\Omega) \times L^1(\Omega)$ defined in the following way.

Definition 3.2. $A_0 = \{[u,v] \in L^1(\mathbf{R}^N) \times L^1(\mathbf{R}^N);\ a(u) \in (L^1(\mathbf{R}^N))^N\}$ and

$$\int_{\mathbf{R}^N} \operatorname{sign}_0(u(x)-k)((a(u(x))-a(k)) \cdot \varphi_x(x) + v(x)\varphi(x))dx \ge 0, \tag{3.69}$$

for all $\varphi \in C_0(\mathbf{R}^N)$ such that $\varphi \ge 0$, and all $k \in \mathbf{R}$. Here, $\operatorname{sign}_0 r = r/|r|$ for $r \ne 0$, $\operatorname{sign}_0 0 = 0$.

It is readily seen that, if $a \in C^1(\mathbf{R})$ and $u \in C_0^1(\mathbf{R}^N)$, then $u \in D(A_0)$ and $A_0 u = a(u)_x$. Indeed, if ρ is a smooth approximation of $r \to |r|$ of the form considered above, then we have

$$\begin{aligned}\int_{\mathbf{R}^N} \rho'(u(x)-k)a(u)_x \varphi\, dx &= \int_{\mathbf{R}^N} dx \left(\int_k^{u(x)} \rho'(s-k)a'(s)ds\right)_x \varphi(x)dx \\ &= -\int_{\mathbf{R}^N} dx \left(\left(\int_k^{u(x)} \rho'(s-k)a'(s)dx\right)\right) \cdot \varphi_x(s),\end{aligned}$$

where $a' = (a_1', a_2', \dots, a_N')$ is the derivative of a. Now, letting ρ' tend to sign_0, we get

$$\int_{\mathbf{R}^N} \operatorname{sign}_0(u(x)-k)(a(u(x)-a(k)) \cdot \varphi_x(x) + a(u(x))_x \varphi(x))dx = 0$$

for all $\varphi \in C_0(\mathbf{R}^N)$. Hence, $u \in D(A_0)$ and $A_0 u = (a(u))_x$.

Conversely, if $u \in D(A_0) \cap L^\infty(\mathbf{R}^N)$ and $v \in A_0 u$, then using the inequality (3.69) with $k = \|u\|_{L^\infty(\mathbf{R}^N)} + 1$ and $k = -(\|u\|_{L^\infty(\mathbf{R}^N)} + 1)$, we get

$$\int_{\mathbf{R}^N} ((a(u(x)) - a(k)) \cdot \varphi_x(x) + v(x)\varphi(x))dx \le 0, \qquad \forall \varphi \in C_0^\infty(\mathbf{R}^N),\ \varphi \ge 0,$$

respectively,

$$\int_{\mathbf{R}^N} ((a(u(x)) - a(k)) \cdot \varphi_x(x) + v(x)\varphi(x))dx \geq 0, \qquad \forall \varphi \in C_0^\infty(\mathbf{R}^N),\ \varphi \geq 0.$$

Hence, $-(a(u))_x + v = 0$ in $\mathscr{D}'(\mathbf{R}^N)$.

Let A be the closure of A_0 in $L^1(\mathbf{R}^N) \times L^1(\mathbf{R}^N)$; that is, $A = \{[u,v] \in L^1(\mathbf{R}^N) \times L^1(\mathbf{R}^N);\ \exists [u_n, v_n] \in A_0,\ u_n \to u,\ v_n \to v \text{ in } L^1(\mathbf{R}^N)\}$.

Theorem 3.8. *Let $a : \mathbf{R} \to \mathbf{R}^N$ be continuous and $\limsup_{r\to 0}(\|a(r)\|/|r|) < \infty$. Then A is m-accretive.*

We prove Theorem 3.8 in several steps but, before proceeding with its proof, we must emphasize that a function u satisfying (3.69) is not a simple distributional solution to equation $(a(u))_x = v$. Its precise meaning becomes clear in the context of the so-called entropy solution to the conservation law equation $u_t + (a(u))_x = v$ which is discussed later on in Chapter 5. We shall first prove the following.

Lemma 3.2. *A is accretive in $L^1(\mathbf{R}^N) \times L^1(\mathbf{R}^N)$.*

Proof. Let $[u,v]$ and $[\bar{u},\bar{v}]$ be two arbitrary elements of A_0. By Definition 3.2, we have, for $k = \bar{u}(y)$, $\varphi(x) = \psi(x,y)$ ($\psi \in C_0^\infty(\mathbf{R}^N \times \mathbf{R}^N)$, $\psi \geq 0$),

$$\begin{aligned}\int_{\mathbf{R}^N\times\mathbf{R}^N} \operatorname{sign}_0(u(x) - \bar{u}(y))(a(u(x)) - a(\bar{u}(y)) \cdot \psi_x(x,y)\\ + v(x)\psi(x,y))dx\,dy \geq 0.\end{aligned} \tag{3.70}$$

Now, it is clear that we can interchange u and $\bar{u}$, v and $\bar{v}$, x and y to obtain, by adding to (3.70) the resulting inequality,

$$\begin{aligned}\int_{\mathbf{R}^N\times\mathbf{R}^N} \operatorname{sign}_0(u(x) - \bar{u}(y))((a(u(x)) - a(\bar{u}(y)) \cdot (\psi_x(x,y))\\ + \psi_y(x,y)) + (v(x) - \bar{v}(y))\psi(x,y))dx\,dy \geq 0,\end{aligned} \tag{3.71}$$

for all $\psi \in C_0^\infty(\mathbf{R}^N \times \mathbf{R}^N)$, $\psi \geq 0$. Now, we take

$$\psi(x,y) = \frac{1}{\varepsilon^n}\,\varphi(x+y)\rho\left(\frac{x-y}{\varepsilon}\right),$$

where $\varphi \in C_0^\infty(\mathbf{R}^N)$, $\varphi \geq 0$, and $\rho \in C_0(\mathbf{R}^N)$ is such that $\operatorname{supp}\rho \subset \{y;\ \|y\| \leq 1\}$, $\int \rho(y)dy = 1$, $\rho(y) = \rho(-y)$, $\forall y \in \mathbf{R}^N$.

Substituting in (3.71), we get after some calculation that

$$\begin{aligned}\int_{\mathbf{R}^N\times\mathbf{R}^N} \operatorname{sign}_0(u(y+\varepsilon z) - \bar{u}(y))(2(a(u(y+\varepsilon z))\\ -a(\bar{u}(y))) \cdot \nabla\varphi(y+\varepsilon z)) + (v(y+\varepsilon z) - \bar{v}(y))\varphi(y+\varepsilon z))\rho(z)dy\,dz \geq 0.\end{aligned} \tag{3.72}$$

Now, letting ε tend to zero in (3.72), we get

$$\int_{\mathbf{R}^N} \theta(y)(v(y) - \bar{v}(y))\varphi(y)dy + 2\int_{\mathbf{R}^N} \theta(y)(a(u(y)) - a(\bar{u}(y))\cdot\nabla\varphi(y))dy \geq 0, \tag{3.73}$$

for all $\theta(y) \in \operatorname{sign}(u(y) - \bar{u}(y))$, a.e. $y \in \mathbf{R}^N$. Hence, for every $\varphi \in C_0^\infty(\mathbf{R}^N)$, $\varphi \geq 0$, there exists $\theta \in J(u - \bar{u})$ such that (3.73) holds, where J is the duality mapping of the space $L^1(\Omega)$ (see (1.4)). If in (3.73) we take $\varphi = \alpha(\varepsilon\|y\|^2)$, where $\alpha \in C_0^\infty(\mathbf{R})$, $\alpha \geq 0$, and $\alpha(r) = 1$ for $|r| \leq 1$, and let $\varepsilon \to 0$, we get

$$\int_{\mathbf{R}^N} \theta(y)(v(y) - \bar{v}(y))dy \geq 0$$

for some $\theta \in J(u - \bar{u})$. Hence, A_0 is accretive in $L^1(\mathbf{R}^N)$ and hence so is its closure A. □

In order to prove that A is m-accretive, taking into account that A_0 is accretive, it suffices to show that the range of $I + A_0$ is dense in $L^1(\mathbf{R}^N)$; that is, that the equation $u + a(u)_x = f$ has a solution (in the generalized sense) for a sufficiently large class of functions f. This means, adopting a terminology used in linear theory, that A_0 is essentially m-accretive. To this end, we approximate this equation by the following family of elliptic equations

$$u + a(u)_x - \varepsilon\Delta u = f \quad \text{in } \mathbf{R}^N. \tag{3.74}$$

Lemma 3.3. *Let $a \in C^1$, a' bounded, and let $\varepsilon > 0$. Then, for each $f \in L^2(\mathbf{R}^N)$, equation* (3.74) *has a solution $u \in H^2(\mathbf{R}^N)$.*

Proof. Denote by Λ the operator defined in $L^2(\mathbf{R}^N)$ by

$$\Lambda = -\Delta, \qquad D(\Lambda) = H^2(\mathbf{R}^N)$$

and let $Bu = -a(u)_x$, $\forall u \in D(B) = H^1(\mathbf{R}^N)$. The operator $T = (I + \varepsilon\Lambda)^{-1}B$ is continuous and bounded from $H^1(\mathbf{R}^N)$ to $H^2(\mathbf{R}^N)$, and therefore it is compact in $H^1(\mathbf{R}^N)$. For a given $f \in L^2(\mathbf{R}^N)$, equation (3.74) is equivalent to

$$u = Tu + (I + \varepsilon\Lambda)^{-1}f. \tag{3.75}$$

Let $D = \{u \in H^1(\mathbf{R}^N);\ \|u\|^2_{L^2(\mathbf{R}^N)} + \varepsilon\|\nabla u\|^2_{L^2(\mathbf{R}^N)} < R^2\}$, where $R = \|f\|_{L^2(\mathbf{R}^N)} + 1$. We note that

$$(I + \varepsilon\Lambda)^{-1}f \notin (I - tT)(\partial D), \qquad 0 \leq t \leq 1. \tag{3.76}$$

Indeed, otherwise there is $u \in \partial D$ and $t \in [0,1]$ such that

$$u - \varepsilon\Delta u + ta(u)_x = f \quad \text{in } \mathbf{R}^N,$$

and we argue from this to a contradiction. Multiplying the last equation by u and integrating on $\mathbf{R}^N$, we get

$$\|u\|^2_{L^2(\mathbf{R}^N)} + \varepsilon\|\nabla u\|^2_{L^2(\mathbf{R}^N)} + t\int_{\mathbf{R}^N} a(u)_x u\,dx = \int_{\mathbf{R}^N} fu\,dx.$$

On the other hand, we have

$$\int_{\mathbf{R}^N} a(u)_x u\,dx = -\int_{\mathbf{R}^N} a(u)\cdot u_x\,dx = -\int_{\mathbf{R}^N} \operatorname{div} b(u)dx = 0,$$

where $b(u) = \int_0^u a(s)ds$. Hence,

$$\|u\|^2_{L^2(\mathbf{R}^N)} + \varepsilon\|\nabla u\|^2_{L^2(\mathbf{R}^N)} \le \|f\|_{L^2(\mathbf{R}^N)}\|u\|_{L^2(\mathbf{R}^N)} \le (R-1)R < R^2,$$

and so $u \notin \partial D$.

Let us denote by $d(I-tT, D, (I+\varepsilon\Lambda)^{-1}f)$ the topological degree of the map $I-tT$ relative to D at the point $(I+\varepsilon\Lambda)^{-1}f$. By (3.76) and the invariance property of topological degree, it follows that (see [8] for the definition and basic properties of topological degree in Banach spaces)

$$d(I-tT, D, (I+\varepsilon\Lambda)^{-1}f) = d(I, D, (I+\varepsilon\Lambda)^{-1}f)$$

for all $0 \le t \le 1$. Hence,

$$d(I-T, D, (I+\varepsilon\Lambda)^{-1}f) = d(I, D, (I+\varepsilon\Lambda)^{-1}f) = 1$$

because $(I+\varepsilon\Lambda)^{-1}f \in D$. Hence, equation (3.75) has at least one solution $u \in D(\Lambda) = H^2(\mathbf{R}^N)$ and so the proof of Lemma 3.3 is complete. □

Lemma 3.4. *Under the assumptions of Lemma* 3.3*, if* $f \in L^p(\mathbf{R}^N) \cap L^2(\mathbf{R}^N)$, $1 \le p \le \infty$*, then* $u \in L^p(\mathbf{R}^N)$ *and*

$$\|u\|_{L^p(\mathbf{R}^N)} \le \|f\|_{L^p(\mathbf{R}^N)}. \tag{3.77}$$

Proof. We first treat the case $1 < p < \infty$. Let $\alpha_n : \mathbf{R} \to \mathbf{R}$ be defined by

$$\alpha_n(r) = \begin{cases} |r|^{p-2}r & \text{if } |r| \le n, \\ n^{p-2}r & \text{if } r > n, \\ n^{p-2}r & \text{if } r < -n. \end{cases}$$

If we multiply equation (3.74) by $\alpha_n(u) \in L^2(\mathbf{R}^N)$ and integrate on $\mathbf{R}^N$, we get

$$\int_{\mathbf{R}^N} \alpha_n(u)u\,dx \le \int_{\mathbf{R}^N} f\alpha_n(u)dx \tag{3.78}$$

because, as previously seen,

$$\int_{\mathbf{R}^N} a(u)_x \alpha_n(u)dx = \int_{\mathbf{R}^N} dx\left(\int_0^{u(x)} a'(s)\alpha_n(s)ds\right)_x dx = 0,$$

and

$$-\int_{\mathbf{R}^N} \Delta u \alpha_n(u)dx = \int_{\mathbf{R}^N} \alpha_n'(u)|\nabla u|^2 dx \geq 0,$$

because α_n is monotonically increasing. Note also the inequality

$$\alpha_n(r)r \geq |\alpha_n(r)|^q, \qquad \forall r \in \mathbf{R},\ \frac{1}{p}+\frac{1}{q}=1.$$

Then, using the Hölder inequality in (3.78), we get

$$\int_{\mathbf{R}^N} |\alpha_n(u)|^q dx \leq \left(\int_{\mathbf{R}^N} |f|^p dx\right)^{1/p} \left(\int_{\mathbf{R}^N} |\alpha_n(u)|^q dx\right)^{1/q},$$

whence

$$\int_{[|u(x)|\leq n]} |u(x)|^p dx \leq \|f\|^p_{L^p(\mathbf{R}^N)},$$

which clearly implies that $u \in L^p(\mathbf{R}^N)$ and that (3.77) holds. In the case $p=1$, we multiply equation (3.74) by $\delta_n(u)$, where

$$\delta_n(r) = \begin{cases} nr & \text{if } |r| \leq n^{-1}, \\ 1 & \text{if } r > n^{-1}, \\ -1 & \text{if } r < -n^{-1}. \end{cases}$$

Note that $\delta_n(u) \in L^2(\mathbf{R}^N)$ because $m\{x \in \mathbf{R}^N;\ |u(x)| > n^{-1}\} \leq n^2\|u\|^2_{L^2(\mathbf{R}^N)}$. Then, arguing as before, we get

$$\begin{aligned}\int_{[|u(x)|\geq n^{-1}]} |u(x)|dx &\leq \int_{\mathbf{R}^N} \delta_n(u)dx \leq \int_{\mathbf{R}^N} |f|\delta_n(u)|dx \\ &\leq n\int_{[|u|\leq n^{-1}]} |f|\,|u|dx + \int_{[|u|>n^{-1}]} |f|dx \leq \|f\|_{L^1(\mathbf{R}^N)}.\end{aligned}$$

Then, letting $n \to \infty$, we get (3.77), as desired.

Finally, in the case $p=\infty$, we set $M = \|f\|_{L^\infty(\mathbf{R}^N)}$. Then, we have

$$u - M + a(u)_x - \varepsilon\Delta(u-M) = f - M \leq 0, \quad \text{a.e. in } \mathbf{R}^N.$$

Multiplying this by $(u-M))^+$ (which, as is well known, belongs to $H^1(\mathbf{R}^N)$), we get $\int_{\mathbf{R}^N}((u-M)^+)^2dx \leq 0$ because

$$\int_{\mathbf{R}^N} a(u)_x(u-M)^+dx = 0,$$

$$-\int_{\mathbf{R}^N} \Delta(u-M)(u-M)^+dx = \int_{\mathbf{R}^N} |\nabla(u-M)^+|^2dx \geq 0.$$

Hence, $u(x) \leq M$, a.e. $x \in \mathbf{R}^N$. Now, we multiply the equation

$$u + M + (a(u))_x - \varepsilon\Delta(u+M) = f + M \geq 0$$

by $(u+M)^-$ and get as before that $(u+M)^- = 0$, a.e. in $\mathbf{R}^N$. Hence, $u \in L^\infty(\mathbf{R}^N)$ and

$$|u(x)| \le \|f\|_{L^\infty(\mathbf{R}^N)}, \quad \text{a.e. } x \in \mathbf{R}^N,$$

as desired. □

Lemma 3.5. *Under the assumptions of Lemma* 3.3, *let* $f, g \in L^2(\mathbf{R}^N) \cap L^1(\mathbf{R}^N)$ *and let* $u, v \in H^2(\mathbf{R}^N) \cap L^1(\mathbf{R}^N)$ *be the corresponding solutions to equation* (3.74)*. Then we have*

$$\|(u-v)^+\|_{L^1(\mathbf{R}^N)} \le \|(f-g)^+\|_{L^1(\mathbf{R}^N)}, \tag{3.79}$$

$$\|u-v\|_{L^1(\mathbf{R}^N)} \le \|(f-g)\|_{L^1(\mathbf{R}^N)}. \tag{3.80}$$

Proof. Because (3.80) is an immediate consequence of (3.79) we confine ourselves to the latter estimate. If we multiply the equation

$$u - v + (a(u) - a(v))_x - \varepsilon\Delta(u-v) = f - g$$

by $\xi \in L^\infty(\mathbf{R}^N)$, $\xi(x) \in \operatorname{sign}(u-v)^+$ (or, more precisely, by $\zeta(u-v)$, where ζ is given by (3.26)) and integrate on $\mathbf{R}^N$, we get

$$\int_{\mathbf{R}^N} (u-v)^+ dx + \int_{\mathbf{R}^N} (a(u)-a(v))_x \xi(x) dx \le \int_{\mathbf{R}^N} (f-g)^+ dx.$$

Now, by the divergence theorem, we have

$$\int_{\mathbf{R}^N} (a(u)-a(v))_x \xi(x) dx = \int_{[u(x)>v(x)]} (a(u(x)) - a(v(x)))_x dx = 0$$

because

$$a(u) = a(v) \quad \text{on } \partial\{x;\, u(x) > v(x)\}.$$

(Here, ∂ denotes the boundary.) Hence, $\|(u-v)^+\|_{L^1(\mathbf{R}^N)} \le \|(f-g)^+\|_{L^1(\mathbf{R}^N)}$, as claimed. □

Proof of Theorem 3.8. Let us show first that $L^1(\mathbf{R}^N) \cap L^\infty(\mathbf{R}^N) \subset R(I + A_0)$. To this end, consider a sequence $\{a_\varepsilon\}$ of C^1 functions such that $a_\varepsilon(0) = 0$ and $a_\varepsilon \overset{\varepsilon \to 0}{\longrightarrow} a$ uniformly on compacta. For $f \in L^1(\mathbf{R}^N) \cap L^\infty(\mathbf{R}^N)$, let $u_\varepsilon \in H^1(\mathbf{R}^N) \cap L^1(\mathbf{R}^N) \cap L^\infty(\mathbf{R}^N)$ be the solution to equation (3.74). Note the estimates

$$\|u_\varepsilon\|_{L^1(\mathbf{R}^N)} \le \|f\|_{L^1(\mathbf{R}^N)},\ \|u_\varepsilon\|_{L^\infty(\mathbf{R}^N)} \le \|f\|_{L^\infty(\mathbf{R}^N)}, \tag{3.81}$$

which were proven earlier in Lemma 3.5. Also, multiplying (3.74) by u_ε and integrating on $\mathbf{R}^N$, we get

$$\|u_\varepsilon\|^2_{L^2(\mathbf{R}^N)} + \varepsilon\|\nabla u_\varepsilon\|^2_{L^2(\mathbf{R}^N)} \le C\|f\|^2_{L^2(\mathbf{R}^N)}. \tag{3.82}$$

Moreover, applying Lemma 3.4 to the functions $u = u_\varepsilon(x)$ and $v = v_\varepsilon(x+y)$, we get the estimate

$$\int_{\mathbf{R}^N} |u_\varepsilon(x+y) - u_\varepsilon(x)|dx \le \int_{\mathbf{R}^N} |f(x+y) - f(x)|dx, \qquad \forall y \in \mathbf{R}^N.$$

By the Kolmogorov's compactness criterion, these estimates imply that $\{u_\varepsilon\}$ is compact in $L^1_{\mathrm{loc}}(\mathbf{R}^N)$ and, therefore, there is a subsequence, which for simplicity again denoted u_ε, such that

$$\begin{aligned} u_\varepsilon &\to u \qquad && \text{strongly in every } L^1(B_R),\ \forall R > 0, \\ u_\varepsilon(x) &\to u(x), && \text{a.e. } x \in \mathbf{R}^N, \end{aligned} \tag{3.83}$$

where $B_R = \{x;\ \|x\| \le R\}$. We show that $u + A_0 u = f$.

Let $\varphi \in C_0^\infty(\mathbf{R}^N)$, $\varphi \ge 0$, and let $\alpha \in C^1(\mathbf{R})$ be such that $\alpha'' \ge 0$. We multiply equation (3.74) by $\alpha'(u_\varepsilon)\varphi$, and integrate on $\mathbf{R}^N$. Then, the integration by parts yields

$$\int_{\mathbf{R}^N} \alpha'(u_\varepsilon)u_\varepsilon \varphi\, dx - \int_{\mathbf{R}^N} (\alpha'(u_\varepsilon)\varphi)_x(a(u_\varepsilon) - a(k))dx + \varepsilon \int_{\mathbf{R}^N} \alpha''(u_\varepsilon)(\nabla u_\varepsilon)^2 \varphi\, dx$$
$$+ \varepsilon \int_{\mathbf{R}^N} (\nabla u_\varepsilon \cdot \nabla \varphi)\alpha'(u_\varepsilon)dx = \int_{\mathbf{R}^N} f\alpha'(u_\varepsilon)\varphi\, dx.$$

This yields

$$\int_{\mathbf{R}^N} (\alpha'(u_\varepsilon)u_\varepsilon\varphi + \varepsilon\alpha'(u_\varepsilon)\nabla u_\varepsilon \cdot \nabla\varphi - (\alpha'(u_\varepsilon)\varphi)_x(a(u_\varepsilon) - a(k)))dx$$
$$\le \int_{\mathbf{R}^N} f\alpha'(u_\varepsilon)\varphi\, dx.$$

Now, letting ε tend to zero, it follows by (3.81)–(3.83) that

$$\int_{\mathbf{R}^N} (\alpha'(u)u\varphi - (\alpha'(u)\varphi)_x(a(u) - a(k)))dx \le \int_{\mathbf{R}^N} f\alpha'(u)\varphi\, dx.$$

Next, we take $\alpha'(s) = \zeta(s-k)$, where ζ is of the form (3.26).

Then, letting $\zeta \to \mathrm{sign}_0$, we get the inequality

$$\int_{\mathbf{R}^N} \mathrm{sign}_0(u-k)[u\varphi - (a(u) - a(k))\varphi_x - f\varphi]dx \le 0.$$

On the other hand, because $\limsup_{|r|\to 0}(\|a(r)\|/|r|) < \infty$, we have that $a(u) \in L^1(\mathbf{R}^N)$. We have therefore shown that $f \in u + A_0 u$. Now, let $f \in L^1(\mathbf{R}^N)$, and let $f_n \in L^1(\mathbf{R}^N) \cap L^\infty(\mathbf{R}^N)$ be such that $f_n \to f$ in $L^1(\mathbf{R}^N)$ for $n \to \infty$. Let $u_n \in D(A_0)$ be the solution to the equation $u + A_0 u \ni f_n$. Because A_0 is accretive in $L^1(\mathbf{R}^N) \times L^1(\mathbf{R}^N)$, we see that $\{u_n\}$ is convergent in $L^1(\mathbf{R}^N)$. Hence, there is $u \in L^1(\mathbf{R}^N)$ such that

$$u_n \to u, \qquad v_n - u_n \to f \quad \text{in } L^1(\mathbf{R}^N),\ v_n \in A_0 u_n.$$

This implies that $f \in u + Au$. □

In particular, we have proved that for every $f \in L^1(\mathbf{R}^N)$ the first-order partial differential equation

$$u - \sum_{i=1}^{N} \frac{\partial}{\partial x_i} a_i(u) = f \quad \text{in } \mathbf{R}^N \tag{3.84}$$

has a unique generalized solution $u \in L^1(\mathbf{R}^N)$, and the map $f \to u$ is Lipschitz continuous in $L^1(\mathbf{R}^N)$.

Bibliographical Remarks

The general theory of nonlinear m-accretive operators in Banach spaces has been developed in the works of Kato [10] and Crandall and Pazy [6, 7] in connection with the theory of semigroups of nonlinear contractions and nonlinear Cauchy problem in Banach spaces, which is presented later on. The existence theory of semilinear elliptic equations in L^1 presented here is due to Bénilan, Brezis, and Crandall [3], and Brezis and Strauss [4].

The m-accretivity of operator associated with first-order linear equation in $\mathbf{R}^N$ (Theorem 3.8) was proven by Crandall [5] in connection with the conservation law equation which is discussed in Chapter 5.

References

1. V. Barbu, Continuous perturbation of nonlinear m-accretive operators in Banach spaces, *Boll. Unione Mat. Ital.*, **6** (1972), pp. 270–278.
2. V. Barbu, *Nonlinear Semigroups and Differential Equations in Banach Spaces*, Noordhoff, Leyden, 1976.
3. Ph. Bénilan, H. Brezis, M.G. Crandall, A semilinear equation in $L^1(\mathbf{R}^N)$, *Ann. Scuola Norm. Sup. Pisa*, **2** (1975), pp. 523–555.
4. H. Brezis, W. Strauss, Semilinear elliptic equations in L^1, *J. Math. Soc. Japan*, **25** (1973), pp. 565–590.
5. M.G. Crandall, The semigroup approach to first-order quasilinear equation in several space variables, *Israel J. Math.*, **12** (1972), pp. 108–132.
6. M.G. Crandall, A. Pazy, Semigroups of nonlinear contractions and dissipative sets, *J. Funct. Anal.*, **3** (1969), pp. 376–418.
7. M.G. Crandall, A. Pazy, On accretive sets in Banach spaces, *J. Funct. Anal.*, **5** (1970), pp. 204–217.
8. K. Deimling, *Nonlinear Functional Analysis*, Springer-Verlag, Berlin, 1975.
9. J. Dautray, J.L. Lions, *Mathematical Analysis and Numerical Methods for Science and Technology*, Springer-Verlag, Berlin, 1982.
10. T. Kato, Accretive operators and nonlinear evolution equations in Banach spaces, *Nonlinear Functional Analysis,* F. Browder (Ed.), American Mathematical Society, Providence, RI, 1970, pp. 138–161.
11. R.H. Martin, *Nonlinear Operators and Differential Equations in Banach Spaces*, John Wiley and Sons, New York, 1976.
12. G. Stampacchia, *Equations Elliptiques du Second Ordre à Coefficients Discontinues*, Les Presses de l'Université de Montréal, Montréal, 1966.

Chapter 4
The Cauchy Problem in Banach Spaces

Abstract This chapter is devoted to the Cauchy problem associated with nonlinear quasi-accretive operators in Banach spaces. The main result is concerned with the convergence of the finite difference scheme associated with the Cauchy problem in general Banach spaces and in particular to the celebrated Crandall–Liggett exponential formula for autonomous equations, from which practically all existence results for the nonlinear accretive Cauchy problem follow in a more or less straightforward way.

4.1 The Basic Existence Results

Mild Solutions

Let X be a real Banach space with the norm $\|\cdot\|$ and dual X^* and let $A \subset X \times X$ be a quasi-accretive set of $X \times X$, or in other terminology, $A : D(A) \subset X \to X$ is an operator (eventually multivalued) such that $A + \omega I$ is accretive for some $\omega \in \mathbf{R}$. We refer to Section 3.1 for definitions and basic properties of quasi-accretive (or ω-accretive) operators.

Consider the Cauchy problem

$$\begin{cases} \dfrac{dy}{dt}(t) + Ay(t) \ni f(t), & t \in [0,T], \\ y(0) = y_0, \end{cases} \tag{4.1}$$

where $y_0 \in X$ and $f \in L^1(0,T;X)$.

Definition 4.1. A strong solution to (4.1) is a function $y \in W^{1,1}((0,T];X) \cap C([0,T];X)$ such that

$$f(t) - \frac{dy}{dt}(t) \in Ay(t), \quad \text{a.e. } t \in (0,T), \ y(0) = y_0.$$

Here, $W^{1,1}((0,T];X) = \{y \in L^1(0,T;X);\ y' \in L^1(\delta,T;X), \forall \delta \in (0,T)\}$.

V. Barbu, *Nonlinear Differential Equations of Monotone Types in Banach Spaces*, Springer Monographs in Mathematics, DOI 10.1007/978-1-4419-5542-5_4,

It is readily seen that any strong solution to (4.1) is unique and is a continuous function of f and y_0. More precisely, we have:

Proposition 4.1. *Let A be ω-accretive, $f_i \in L^1(0,T;X)$, $y_0^i \in \overline{D(A)}$, $i=1,2$, and let $y_i \in W^{1,1}((0,T];X)$, $i=1,2$, be corresponding strong solutions to problem* (4.1). *Then,*

$$\begin{aligned}\|y_1(t)-y_2(t)\| &\le e^{\omega t}\|y_0^1-y_0^2\|+\int_0^t e^{\omega(t-s)}[y_1(s)-y_2(s),f_1(s)-f_2(s)]_s ds\\ &\le e^{\omega t}\|y_0^1-y_0^2\|+\int_0^t e^{\omega(t-s)}\|f_1(s)-f_2(s)\|ds, \qquad \forall t\in[0,T].\end{aligned} \tag{4.2}$$

Here (see Proposition 3.7)

$$[x,y]_s = \inf_{\lambda>0} \lambda^{-1}(\|x+\lambda y\| - \|x\|) = \max\{(y,x^*);\ x^* \in \Phi(x)\} \tag{4.3}$$

and $\|x\|\Phi(x) = J(x)$ is the duality mapping of X; that is, $\Phi(x) = \partial\|x\|$.

The main ingredient of the proof is the following chain differentiation rule lemma.

Lemma 4.1. *Let $y = y(t)$ be an X-valued function on $[0,T]$. Assume that $y(t)$ and $\|y(t)\|$ are differentiable at $t = s$. Then,*

$$\|y(s)\|\,\frac{d}{ds}\,\|y(s)\| = \left(\frac{dy}{ds}(s),w\right), \qquad \forall w \in J(y(s)). \tag{4.4}$$

Here, $J : X \to X^*$ is the duality mapping of X.

Proof. Let $\varepsilon > 0$. We have

$$(y(s+\varepsilon)-y(s),w) \le (\|y(s+\varepsilon)\| - \|y(s)\|)\|w\|, \qquad \forall w \in J(y(s)),$$

and this yields

$$\left(\frac{dy}{ds}(s),w\right) \le \frac{d}{ds}\|y(s)\|\,\|y(s)\|.$$

Similarly, from the inequality

$$(y(s-\varepsilon)-y(s),w) \le (\|y(s-\varepsilon)\| - \|y(s)\|)\|w\|,$$

we get

$$\left(\frac{d}{ds}y(s),w\right) \ge \frac{d}{ds}\|y(s)\|\,\|y(s)\|,$$

as claimed.

In particular, it follows by (4.4) that

$$\frac{d}{ds}\|y(s)\| = \left[y(s),\frac{dy}{ds}(s)\right]_s. \qquad \square \tag{4.5}$$

Proof of Proposition 4.1. We have

$$\frac{d}{ds}(y_1(s)-y_2(s))+Ay_1(s)-Ay_2(s)\ni f_1(s)-f_2(s),\quad \text{a.e. } s\in(0,T). \tag{4.6}$$

On the other hand, because A is ω-accretive, we have (see (3.16))

$$[y_1(s)-y_2(s),Ay_1(s)-Ay_2(s)]_s\geq-\omega\|y_1(s)-y_2(s)\|$$

and so, by (4.5) and (4.6), we see that

$$\frac{d}{ds}\|y_1(s)-y_2(s)\|\leq[y_1(s)-y_2(s),f_1(s)-f_2(s)]_s+\omega\|y_1(s)-y_2(s)\|,$$
$$\text{a.e. } s\in(0,T).$$

Then, integrating on $[0,t]$, we get (4.2), as claimed.

Proposition 4.1 shows that, as far as the uniqueness and continuous dependence of solution of data are concerned, the class of quasi-accretive operators A offers a suitable framework for the Cauchy problem. For this reason, such a nonlinear system is also called *quasi-accretive*. However, for the existence we must extend the notion of the solution for the Cauchy problem (4.1) from differentiable to continuous functions.

Definition 4.2. Let $f\in L^1(0,T;X)$ and $\varepsilon>0$ be given. An ε*-discretization* on $[0,T]$ of the equation $y'+Ay\ni f$ consists of a partition $0=t_0\leq t_1\leq t_2\leq\cdots\leq t_N$ of the interval $[0,t_N]$ and a finite sequence $\{f_i\}_{i=1}^N\subset X$ such that

$$t_i-t_{i-1}<\varepsilon\quad\text{for } i=1,\dots,N,\ T-\varepsilon<t_N\leq T, \tag{4.7}$$

$$\sum_{i=1}^N\int_{t_{i-1}}^{t_i}\|f(s)-f_i\|ds<\varepsilon. \tag{4.8}$$

We denote by $D_A^\varepsilon(0=t_0,t_1,\dots,t_N;f_1,\dots,f_N)$ this ε-discretization.

A $D_A^\varepsilon(0=t_0,t_1,\dots,t_N;f_1,\dots,f_N)$ *solution* to (4.1) is a piecewise constant function $z:[0,t_N]\to X$ whose values z_i on $(t_{i-1},t_i]$ satisfy the finite difference equation

$$\frac{z_i-z_{i-1}}{t_i-t_{i-1}}+Az_i\ni f_i,\qquad i=1,\dots,N. \tag{4.9}$$

Such a function $z=\{z_i\}_{i=1}^N$ is called an ε*-approximate solution* to the Cauchy problem (4.1) if it further satisfies

$$\|z(0)-y_0\|\leq\varepsilon. \tag{4.10}$$

Definition 4.3. A *mild solution* of the Cauchy problem (4.1) is a function $y\in C([0,T];X)$ with the property that for each $\varepsilon>0$ there is an ε-approximate solution z of $y'+Ay \ni f$ on $[0,T]$ such that $\|y(t)-z(t)\| \le \varepsilon$ for all $t \in [0,T]$ and $y(0)=x$.

Let us note that every strong solution $y \in C([0,T];X) \cap W^{1,1}((0,T];X)$ to (4.1) is a mild solution. Indeed, let $0=t_0 \le t_1 \le \cdots \le t_N$ be an ε-discretization of $[0,T]$ such that

$$\left\| \frac{d}{dt}y(t) - \frac{y(t_i)-y(t_{i-1})}{t_i-t_{i-1}} \right\| \le \varepsilon, \qquad t_i - t_{i-1} \le \delta,\ i=1,2,\dots,N,$$

and

$$\int_{t_{i-1}}^{t_i} \|f(t)-f(t_i)\|dt \le \varepsilon(t_i - t_{i-1}).$$

Then, the step function $z:[0,T] \to X$ defined by $z=y(t_i)$ on $(t_{i-1},t_i]$ is a solution to the ε-discretization D_A^ε $(0=t_0,t_1,\dots,t_n;f_1,\dots,f_n)$, and, if we choose the discretization $\{t_j\}$ so that $\|y(t)-y(s)\| \le \varepsilon$ for $t,s \in (t_{i-1},t_i)$, we have by (4.2) that $\|y(t)-z(t)\| \le \varepsilon$ for all $t \in [0,T]$, as claimed.

Theorem 4.1 below is the main result of this section.

Theorem 4.1. *Let A be ω-accretive, $y_0 \in \overline{D(A)}$, and $f \in L^1(0,T;X)$. For each $\varepsilon > 0$, let problem* (4.1) *have an ε-approximate solution. Then, the Cauchy problem* (4.1) *has a unique mild solution y. Moreover, there is a continuous function $\delta = \delta(\varepsilon)$ such that $\delta(0)=0$ and if z is an ε-approximate solution of* (4.1)*, then*

$$\|y(t)-z(t)\| \le \delta(\varepsilon) \quad \text{for } t \in [0,T-\varepsilon]. \tag{4.11}$$

Let $f,g \in L^1(0,T;X)$ and $y,\bar{y}$ be mild solutions to (4.1) *corresponding to f and g, respectively. Then,*

$$\begin{aligned} \|y(t)-\bar{y}(t)\| &\le e^{\omega(t-s)}\|y(s)-\bar{y}(s)\| \\ &\quad + \int_s^t e^{\omega(t-\tau)}[y(\tau)-\bar{y}(\tau), f(\tau)-g(\tau)]_s d\tau \\ &\qquad\qquad \text{for } 0 \le s < t \le T. \end{aligned} \tag{4.12}$$

This important result, which represents the core of the existence theory of evolution processes governed by accretive operators is proved below in several steps. It is interesting that, as Theorem 4.1 amounts to saying, the existence of a unique mild solution for (4.1) is the consequence of two assumptions on A: ω-accretivity and existence of an ε-approximate solution. The latter is implied by the quasi-m-accretivity or a weaker condition of this type. Indeed, we have

Theorem 4.2. *Let C be a closed convex cone of X and let A be ω-accretive in $X \times X$ such that*

$$D(A) \subset C \subset \bigcap_{0<\lambda<\lambda_0} R(I+\lambda A) \quad \text{for some } \lambda > 0. \tag{4.13}$$

Let $y_0 \in \overline{D(A)}$ and $f \in L^1(0,T;X)$ be such that $f(t) \in C$, a.e. $t \in (0,T)$. Then, problem (4.1) *has a unique mild solution y. If y and $\bar{y}$ are two mild solutions to* (4.1) *corresponding to f and g, respectively, then*

$$\begin{aligned}\|y(t) - \bar{y}(t)\| &\leq e^{\omega(t-s)}\|y(s) - \bar{y}(s)\| \\ &+ \int_s^t e^{\omega(t-\tau)}[y(\tau) - \bar{y}(\tau), f(\tau) - g(\tau)]_s d\tau \quad \textit{for } 0 \leq s < t \leq T.\end{aligned} \tag{4.14}$$

Proof. Let $f \in L^1(0,T;X)$ and let f_i be the nodal approximation of f; that is,

$$f_i = \frac{1}{t_i - t_{i-1}} \int_{t_{i-1}}^{t_i} f(s)ds, \qquad i = 1,2,...,N,$$

where $\{t_i\}_{i=1}^N$, $t_0 = 0$, is a partition of the interval $[0,t_N]$ such that $t_i - t_{i-1} < \varepsilon$, $t - \varepsilon < t_N < T$. By assumption (4.13), it follows that, for ε small enough, the function $z = z_i$ on $(t_{i-1}, t_i]$, $z_0 = y_0$, is well defined by (4.9) and it is an ε-approximate solution to (4.1). (It is readily seen by assumption (4.2) and the ω-accretivity of A that equation (4.9) has a unique solution $\{z_i\}_{i=0}^N$.) Thus, Theorem 4.1 is applicable and so problem (4.1) has a unique solution satisfying (4.14). □

In particular, by Theorem 4.2 we obtain the following.

Corollary 4.1. *Let A be quasi-m-accretive. Then, for each* $y_0 \in \overline{D(A)}$ *and* $f \in L^1(0,T;X)$ *there is a unique mild solution y to* (4.1).

In the sequel, we frequently refer to the map $(y_0, f) \to y$ from $\overline{D(A)} \times L^1(0,T;X)$ to $C([0,T];X)$ as the *nonlinear evolution associated with* A. It should be noted that, in particular, the range condition (4.13) holds if $C = X$ and A is ω-m-accretive in $X \times X$.

In the particular case when $f \equiv 0$, if A is ω-accretive and

$$R(I + \lambda A) \supset \overline{D(A)} \quad \text{for all small } \lambda > 0, \tag{4.15}$$

then we have, by Theorem 4.1:

Theorem 4.3 (Crandall and Liggett [24]). *Let A be ω-accretive, satisfying the range condition* (4.15) *and* $y_0 \in \overline{D(A)}$. *Then, the Cauchy problem*

$$\begin{aligned}&\frac{dy}{dt} + Ay \ni 0, \qquad t > 0, \\ &y(0) = y_0,\end{aligned} \tag{4.16}$$

has a unique mild solution y. Moreover,

$$y(t) = \lim_{n\to\infty} \left(I + \frac{t}{n}A\right)^{-n} y_0 \tag{4.17}$$

uniformly in t on compact intervals.

Indeed, in this case, if $t_0 = 0$, $t_i = i\varepsilon$, $i = 1,...,N$, then the solution z_ε to the ε-discretization $D_A^\varepsilon(0 = t_0, t_1, ..., t_N)$ is given by the iterative scheme

$$z_\varepsilon(t) = (I+\varepsilon A)^{-i} y_0 \quad \text{for } t \in ((i-1)\varepsilon, i\varepsilon].$$

Hence, by (4.11), we have

$$\|y(t) - (I+\varepsilon A)^{-i} y_0\| \le \delta(\varepsilon) \quad \text{for } (i-1)\varepsilon < t \le i\varepsilon,$$

which implies the exponential formula (4.17) with uniform convergence on compact intervals. We note that, in particular, the range conditions (4.13) and (4.15) are automatically satisfied if A is quasi-m-accretive; that is, if $\omega I + A$ is m-accretive for some real ω. The solution y to (4.16) given by exponential formula (4.17) is also denoted by $e^{-At} y_0$.

Corollary 4.2. *Let A be quasi-m-accretive and $y_0 \in \overline{D(A)}$. Then the Cauchy problem* (4.16) *has a unique mild solution y given by the exponential formula* (4.17).

We now apply Theorem 4.2 to the mild solutions $y = y(t)$ and $\bar{y} = x$ to the equations

$$y' + Ay \ni f \quad \text{in } (0,T),$$

and

$$y' + Ay \ni v \quad \text{in } (0,T),\ v \in Ax,$$

respectively. We have, by (4.14),

$$\|y(t) - x\| \le e^{\omega(t-s)} \|y(s) - x\| + \int_s^t [y(\tau) - x, f(\tau) - v]_s e^{\omega(t-\tau)} d\tau, \qquad \forall\, 0 \le s < t \le T,\ [x,v] \in A. \tag{4.18}$$

Such a function $y \in C([0,T];X)$ is called an *integral solution* to equation (4.1).

We may conclude, therefore, that under the assumptions of Theorem 4.2 the Cauchy problem (4.1) has an integral solution, which coincides with the mild solution of this problem. On the other hand, it turns out that the integral solution is unique (see Bénilan and Brezis [11]) and under the assumptions of Theorem 4.2 (in particular, if A is ω-m-accretive) these two notions coincide.

It should be mentioned that in finite-dimensional spaces, Theorem 4.1 reduces to the classical Peano convergence scheme for solutions to the Cauchy problem which is valid for any continuous operator A. However, in infinite dimensions there are classical counterexamples which show that continuity alone is not enough for the existence of solutions. On the other hand, in most of significant infinite-dimensional examples the operator A is not continuous. This is the case with nonlinear boundary value problems of parabolic or hyperbolic type where the domain $D(A)$ of operator A is a proper subset of X and so A is unbounded. More is said about this in Chapter 5.

If X is the Euclidean space $\mathbf{R}^N$ and $A = \psi : \mathbf{R}^N \to \mathbf{R}^N$ is a measurable and monotone function; that is,

$$(\psi(x) - \psi(y), x-y)_N \ge 0, \qquad \forall x,y \in \mathbf{R}^N,$$

where $(\cdot,\cdot)_N$ is the scalar product of $\mathbf{R}^N$, then the Cauchy problem

$$\frac{dy}{dt}(t) + \psi(y(t)) = 0, \qquad t \geq 0,$$
$$y(0) = y_0 \tag{4.19}$$

is not, generally, well posed.

This can be seen from the following elementary example

$$\frac{dy}{dt}(t) + \mathrm{sgn}_0 y(t) = 0, \qquad t \geq 0,\ y(0) = y_0,$$

where $\mathrm{sgn}_0 y = y/|y|$. However, if we replace ψ by the Filipov mapping

$$\widetilde{\psi}(x) = \bigcap_{\delta>0} \bigcap_{m(E)=0} \overline{\mathrm{conv}\ \psi(B_\delta(x) \setminus E)}, \qquad \forall x \in \mathbf{R}^N,$$

which, as seen in Proposition 2.5, is m-accretive in $\mathbf{R}^N \times \mathbf{R}^N$, then the corresponding Cauchy problem; that is,

$$\frac{dy}{dt}(t) + \widetilde{\psi}(y(t)) \ni 0, \qquad t \geq 0,$$
$$y(0) = y_0,$$

has by Theorem 4.1 a unique solution y. This is the so-called Filipov solution to (4.19) which exists locally even for nonmonotone functions ψ.

Let us now come back to the proof of Theorem 4.1.

Let z be a solution to an ε-discretization $D_A^\varepsilon(0 = t_1, t_1, \dots, t_N; f_1, \dots, f_N)$ and let w be a solution to $D_A^\varepsilon(0 = s_0, s_1, \dots, s_M; g_1, \dots, g_M)$ with the nodal values z_i and w_j, respectively. We set $a_{ij} = \|z_i - w_j\|$, $\delta_i = (t_i - t_{i-1})$, $\gamma_j = (s_j - s_{j-1})$.

We begin with the following estimate for the solutions to finite difference scheme (4.7)–(4.9).

Lemma 4.2. *For all* $1 \leq i \leq N,\ 1 \leq j \leq M$, *we have*

$$a_{ij} \leq \left(1 - \omega \frac{\delta_i \gamma_j}{\delta_i + \gamma_j}\right)^{-1} \left(\frac{\gamma_j}{\delta_i + \gamma_j} a_{i-1,j} + \frac{\delta_i}{\delta_i + \gamma_j} a_{i,j-1} + \frac{\delta_i \gamma_j}{\delta_i + \gamma_j} [z_i - w_j, f_i - g_j]_s\right). \tag{4.20}$$

Moreover, for all $[x, v] \in A$ *we have*

$$a_{i,0} \leq \alpha_{i,1} \|z_0 - x\| + \|w_0 - x\| + \sum_{k=1}^{i} \alpha_{i,k} \delta_k (\|f_k\| + \|v\|), \qquad 0 \leq i \leq N, \tag{4.21}$$

and

$$a_{0,j} \leq \beta_{j,1} \|w_0 - x\| + \|z_0 - x\| + \sum_{k=1}^{j} \beta_{j,k} \gamma_k (\|g_k + \|v\|), \qquad 0 \leq j \leq M, \tag{4.22}$$

where

$$\alpha_{i,k} = \prod_{m=k}^{i} (1-\omega\delta_m)^{-1}, \qquad \beta_{j,k} = \prod_{m=k}^{j} (1-\omega\gamma_m)^{-1}. \tag{4.23}$$

Proof. We have

$$f_i + \delta_i^{-1}(z_{i-1}-z_i) \in Az_i, \qquad g_j + \gamma_j^{-1}(w_{j-1}-w_j) \in Aw_j, \tag{4.24}$$

and, because A is ω-accretive, this yields (see (3.16))

$$[z_i - w_j, f_i + \delta_i^{-1}(z_{i-1}-z_i) - g_j - \gamma_j^{-1}(w_{j-1}-w_j)]_s \geq -\omega\|z_i - w_j\|.$$

Hence,

$$\begin{aligned} -\omega\|z_i - w_j\| &\leq [z_i - w_j, f_i - g_j]_s + \delta_i^{-1}[z_i - w_j, z_{i-1}-z_i]_s \\ &+ \gamma_j^{-1}[z_i - w_j, w_j - w_{j-1}]_s \\ &\leq [z_i - w_j, f_i - g_j]_s - \delta_i^{-1}(\|z_i - w_j\| - \|z_{i-1} - w_j\|) \\ &- \gamma_j^{-1}(\|z_i - w_j\| - \|z_i - w_{j-1}\|), \end{aligned}$$

and rearrranging we obtain (4.20).

To get estimates (4.21), (4.22), we note that, inasmuch as A is ω-accretive, we have (see (3.3))

$$\|z_i - x\| \leq (1-\delta_i\omega)^{-1}\|z_i - x + \delta_i(f_i + \delta_i^{-1}(z_{i-1}-z_i) - v)\|,$$

respectively,

$$\|w_j - x\| \leq (1-\gamma_j\omega)^{-1}\|w_j - x + \gamma_j(g_j + \gamma_j^{-1}(w_{j-1}-w_j) - v)\|,$$

for all $[x,v] \in A$. Hence,

$$\begin{aligned} \|z_i - x\| &\leq (1-\delta_i\omega)^{-1}\|z_{i-1} - x\| + (1-\delta_i\omega)^{-1}\delta_i(\|f_i\| + \|v\|) \\ \|w_j - x\| &\leq (1-\gamma_j\omega)^{-1}\|w_{j-1} - x\| + (1-\gamma_j\omega)^{-1}\gamma_j(\|g_i\| + \|v\|) \end{aligned}$$

and (4.21), (4.22) follow by a simple calculation. □

In order to get, by (4.20), explicit estimates for a_{ij}, we invoke a technique frequently used in stability analysis of finite difference numerical schemes.

Namely, consider the functions ψ and φ on $[0,T]$ that satisfy the linear first order hyperbolic equation

$$\frac{\partial \psi}{\partial t}(t,s) + \frac{\partial \psi}{\partial s}(t,s) - \omega\psi(t,s) = \varphi(t,s) \quad \text{for } 0 \leq t \leq T,\ 0 \leq s \leq T, \tag{4.25}$$

and the boundary conditions

$$\psi(t,s) = b(t-s) \quad \text{for } t=0 \text{ or } s=0, \tag{4.26}$$

where $b \in C([-T,T])$ and φ is defined later on.

There is a close relationship between equation (4.25) and inequality (4.20). Indeed, let us define the grid

$$D = \{(t_i, s_j);\ 0 = t_0 \le t_1 \le \cdots \le t_N < T,\ 0 = s_0 \le s_1 \le \cdots \le s_M < T\}$$

and approximate (4.25) by the difference equations

$$\begin{aligned} &\frac{\psi_{i,j} - \psi_{i-1,j}}{\delta_i} + \frac{\psi_{i,j} - \psi_{i,j-1}}{\gamma_j} - \omega\psi_{ij} = \varphi_{i,j} \\ &\qquad \text{for } i = 1,\ldots,N,\ j = 1,\ldots,M, \end{aligned} \tag{4.27}$$

where $\delta_i = t_i - t_{i-1}$, $\gamma_j = s_j - s_{j-1}$, and $\varphi_{i,j}$ is a piecewise constant approximation of φ defined below. After some rearrangement we obtain

$$\begin{aligned} \psi_{i,j} = \left(1 - \omega\,\frac{\delta_i\gamma_j}{\delta_i + \gamma_j}\right)^{-1} &\left(\tfrac{\gamma_j}{\delta_i+\gamma_j}\,\psi_{i-1,j} + \tfrac{\delta_i}{\delta_i+\gamma_j}\,\psi_{i,j-1} + \tfrac{\delta_i\gamma_j}{\delta_i+\gamma_j}\,\varphi_{i,j}\right), \\ &\qquad i = 1,\ldots,N,\ j = 1,\ldots,M. \end{aligned} \tag{4.28}$$

In the following we take

$$\varphi(t,s) = \|f(t) - g(s)\|, \qquad \varphi_{i,j} = \|f_i - g_j\|, \qquad i = 1,\ldots,N,\ j = 1,\ldots,M,$$

where f_i and g_j are the nodal approximations of $f, g \in L^1(0,T;X)$, respectively.

Integrating equations (4.25) and (4.26), via the characteristics method, we get

$$\begin{aligned} &\psi(t,s) = G(b,\varphi)(t,s) \\ &= \begin{cases} e^{\omega s} b(t-s) + \displaystyle\int_0^s e^{\omega(s-\tau)}\varphi(t-s+\tau,\tau)\,d\tau & \text{if } 0 \le s < t \le T, \\ e^{\omega t} b(t-s) + \displaystyle\int_0^t e^{\omega(t-\tau)}\varphi(\tau,s-t+\tau)\,d\tau & \text{if } 0 \le t < s \le T. \end{cases} \end{aligned} \tag{4.29}$$

We set $\Omega = (0,T) \times (0,T)$, and for every measurable function $\varphi : [0,T] \times [0,T] \to \mathbf{R}$ we set

$$\begin{aligned} \|\varphi\|_\Omega = \inf\{\|f\|_{L^1(0,T)} + \|g\|_{L^1(0,T)};\ |\varphi(t,s)| \le |f(t)| + |g(s)|,& \\ \text{a.e. } (t,s) \in \Omega\}.& \end{aligned} \tag{4.30}$$

Let $\Omega(\Delta) = [0,t_N] \times [0,s_M]$ and $B : [-s_M, t_N] \to \mathbf{R}$, $\phi : \Omega(\Delta) \to \mathbf{R}$ be piecewise constant functions; that is, here are $b_{i,j}, \phi_{i,j} \in \mathbf{R}$ such that $b(0) = B(0)$ and

$$\begin{aligned} B(r+s) &= b_{ij} && \text{for } t_{i-1} < r \le r_i,\ -s_j \le s < -s_{j-1}, \\ \phi(t,s) &= \phi_{i,j} && \text{for } (t,s) \in (t_{i-1},t_i] \times (s_{j-1},s_j]. \end{aligned}$$

Observe, by (4.29), via the Banach fixed point theorem, that if the mesh $m(\Delta) = \max\{(\delta_i, \gamma_j);\ i,j\}$ of Δ is sufficiently small, then the system (4.28) with the boundary value conditions

$$\psi_{i,j} = b_{i,j} \quad \text{for } i = 0 \text{ or } j = 0, \tag{4.31}$$

has a unique solution $\{\psi_{ij}\}$, $i = 1,...,N,\ j = 1,...,M$.

Denote by $\Psi = H_\Delta(B,\phi)$ the piecewise constant function on Ω defined by

$$\Psi = \psi_{i,j} \quad \text{on } (t_{i-1}, t_i] \times (s_{j-1}, s_j]; \tag{4.32}$$

that is, the solution to (4.28), (4.31).

Lemma 4.3 below provides the convergence of the finite difference scheme (4.27), (4.31) as $m(\Delta) \to 0$.

Lemma 4.3. *Let $b \in C([-T,T])$ and $\varphi \in L^1(\Omega)$ be given. Then,*

$$\|G(b,\varphi) - H_\Delta(B,\phi)\|_{L^\infty(\Omega(\Delta))} \to 0 \tag{4.33}$$

as

$$m(\Delta) + \|b - B\|_{L^\infty(-s_M,t_N)} + \|\varphi - \phi\|_{\Omega(\Delta)} \to 0.$$

Proof. In order to avoid a tedious calculus, we prove (4.33) in the accretive case only (i.e., $\omega = 0$).

Let us prove first the estimate

$$\|H_\Delta(B,\phi)\|_{L^\infty(\Omega(\Delta))} \le \|B\|_{L^\infty(-s_M,t_N)} + \|\phi\|_{\Omega(\Delta)}. \tag{4.34}$$

Indeed, we have $H_\Delta(B,\phi) = H_\Delta(B,0) + H_\Delta(0,\phi)$, and by (4.30), (4.32) we see that the values of $H_\Delta(B,0)$ are convex combinations of the values of B.

Hence,

$$\|H_\Delta(B,0)\|_{L^\infty(\Omega(\Delta))} \le \|B\|_{L^\infty(-s_M,t_N)}.$$

It remains to show that

$$\|H_\Delta(0,\phi)\|_{L^\infty(\Omega(\Delta))} \le \|\phi\|_{\Omega(\Delta)}.$$

By the definition (4.30) of the $\|\cdot\|_{\Omega(\Delta)}$-norm, we have

$$\|\phi\|_{\Omega(\Delta)} = \inf\left\{\sum_{i=1}^N \delta_i\alpha_i + \sum_{j=1}^M \gamma_j\beta_j;\ \alpha_i + \beta_j \ge |\phi_{i,j}|,\ \alpha_i, \beta_j \ge 0\right\}.$$

Now, let $g_{i,j} = \alpha_i + \beta_j \ge |\phi_{i,j}|$ and set

$$d_{i,j} = \sum_{k=1}^i \alpha_k\delta_k + \sum_{k=1}^j \beta_k\gamma_k.$$

It is readily seen that $\psi_{i,j} = d_{i,j}$ satisfy the system (4.28) where $\phi_{i,j} = g_{i,j}$. Hence, $d = H_\Delta(\widetilde{B}, g)$ provided $d_{i,j} = \widetilde{b}_{i,j}$ for $i = 0$ or $j = 0$, where $d = \{d_{i,j}\}$, $\widetilde{B} = \{\widetilde{b}_{i,j}\}$ and $g = \{g_{i,j}\}$. Inasmuch as $g_{i,j} \geq |\phi_{i,j}|$, we have

$$d = H_\Delta(\widetilde{B}, g) \geq H_\Delta(0, \phi) \geq |H_\Delta(0, \phi)$$

if $b_{i,j} \geq 0$. Hence,

$$\|H_\Delta(0,\phi)\|_{L^\infty(\Omega(\Delta))} \leq \|d\|_{L^\infty(\Omega(\Delta))} \leq \|\phi\|_{\Omega(\Delta)},$$

as claimed.

Now, let $\widetilde{\psi} = G(\widetilde{b}, \widetilde{\varphi})$ and assume first that $\widetilde{\psi}_{tt}, \widetilde{\psi}_{ss} \in L^\infty(\Omega)$. Then, by (4.25) we see that $\widetilde{\psi}_{i,j} = \widetilde{\psi}(t_i, s_j)$ satisfy the system

$$\frac{\widetilde{\psi}_{i,j} - \widetilde{\psi}_{i-1,j}}{\delta_i} + \frac{\widetilde{\psi}_{i,j} - \widetilde{\psi}_{i,j-1}}{\gamma_j} = \widetilde{\varphi}_{i,j} + e_{i,j}, \qquad \widetilde{\psi}_{i,0} = \widetilde{b}(t_i), \qquad \widetilde{\psi}_{0,j} = \widetilde{b}(-s_j),$$
$$i = 0, 1, ..., N, \ j = 0, 1, ..., M,$$

where $e = \{e_{ij}\}$ satisfies the estimate

$$|e_{ij}| \leq \gamma_j \|\widetilde{\psi}_{ss}\|_{L^\infty(\Omega)} + \delta_i \|\widetilde{\psi}_{tt}\|_{L^\infty(\Omega)}, \qquad \forall i, j.$$

Then, by (4.34), this yields

$$\begin{aligned}
&\|G(\widetilde{b}, \widetilde{\varphi}) - H_\Delta(B, \phi)\|_{L^\infty(\Omega(\Delta))} \\
&\leq \|B - \widetilde{b}\|_{L^\infty(-s_M, t_N)} + \|\widetilde{\varphi} - \phi\|_{\Omega(\Delta)} + \|e\|_{\Omega(\Delta)} \\
&\leq \|B - \widetilde{b}\|_{L^\infty(-s_M, t_N)} + \|\widetilde{\varphi} - \phi\|_{\Omega(\Delta)} \\
&\quad + Cm(\Omega)(\|\widetilde{\psi}_{tt}\|_{L^\infty(\Omega)} + \|\widetilde{\psi}_{ss}\|_{L^\infty(\Omega)}).
\end{aligned} \tag{4.35}$$

Now, let $\varphi \in L^1(\Omega)$, $b \in C([-T, T])$, and $\widetilde{b} \in C^2([-T, T])$, $\widetilde{\varphi} \in C^2(\widetilde{\Omega})$. Then, $\widetilde{\psi} = G(\widetilde{b}, \widetilde{\varphi})$ is smooth, and by (4.35) we have

$$\begin{aligned}
&\|G(b, \varphi) - H_\Delta(B, \phi)\|_{L^\infty(\Omega(\Delta))} \\
&\leq \|G(b, \varphi) - G(\widetilde{b}, \widetilde{\varphi})\|_{L^\infty(\Omega(\Delta))} + \|G(\widetilde{b}, \widetilde{\varphi}) - H_\Delta(B, \phi)\|_{L^\infty(\Omega(\Delta))} \\
&\leq 2\|b - \widetilde{b}\|_{L^\infty(-s_M, t_N)} + C\|\varphi - \widetilde{\varphi}\|_{\Omega(\Delta)} + \|B - b\|_{L^\infty(-s_M, t_N)} \\
&\quad + \|\widetilde{\varphi} - \phi\|_{\Omega(\Delta)} + Cm(\Delta)(\|\widetilde{\psi}_{tt}\|_{L^\infty(\Omega)} + \|\widetilde{\psi}_{ss}\|_{L^\infty(\Omega)}).
\end{aligned} \tag{4.36}$$

Given $\eta > 0$, we may choose $\widetilde{b}$ and $\widetilde{\varphi}$ such that $\|b - \widetilde{b}\|_{L^\infty(-s_M, t_N)}, \|\varphi - \widetilde{\varphi}\|_{\Omega(\Delta)} \leq \eta$. Then (4.36) implies (4.33), as desired. □

Proof of Theorem 4.1 (Continued). We apply Lemma 4.3, where $\varphi(t,s) = \|f(t) - g(s)\|$, $\phi = \{\phi_{i,j}\}$, $\phi_{i,j} = \|f_i - g_j\|$, $1 \leq j \leq M$, $1 \leq i \leq N$, f_i and g_j are the nodal values of f and g, respectively, and

$$B(t) = b_{i,0} \quad \text{for } t_{i-1} < t \le t_i,\ i = 1,...,N,$$
$$B(s) = b_{0,j} \quad \text{for } -s_j < s \le -s_{j-1},\ j = 1,...,M.$$

Here, $b_{i,0}$ is the right-hand side of (4.21) and $b_{0,j}$ is the right-hand side of (4.22). It is easily seen that, for $\varepsilon \to 0$,

$$B(t) \to b(t) = e^{\omega t}\|z_0 - x\| + \|w_0 - x\| + \int_0^t e^{\omega(t-\tau)}(\|f(\tau)\| + \|v\|)d\tau, \qquad \forall t \in [0,T],$$

and

$$B(s) \to b(-s) = e^{\omega s}\|w_0 - x\| + \|z_0 - x\| + \int_0^s e^{\omega(s-\tau)}(\|g(\tau)\| + \|v\|)d\tau, \qquad \forall s \in [-T,0].$$

By (4.8), we have

$$\|\varphi - \phi\|_{\Omega(\Delta)} \le 2\varepsilon$$

and, by Lemma 4.2,

$$a_{i,j} = \|z_i - w_j\| \le H_\Delta(B,\phi)_{i,j}, \qquad \forall i,j.$$

Then, by Lemma 4.3, we see that, for every $\eta > 0$, we have

$$\|z(t) - w(s)\| \le G(b,\varphi)(t,s) + \eta, \qquad \forall s,t \in [0,T], \tag{4.37}$$

as soon as $0 < \varepsilon < \nu(\eta)$.

If $f \equiv g$ and $z_0 = w_0$, then $G(b,\varphi)(t,t) = e^{\omega t}b(0) = 2e^{\omega t}\|z_0 - x\|$ and so, by (4.37),

$$\|z(t) - w(t)\| \le \eta + 2e^{\omega t}\|z - x\|, \qquad \forall x \in D(A),\ t \in [0,T],$$

for all $0 < \varepsilon \le \nu(\eta)$. Because $\|z_0 - s_0\| \le \varepsilon$, $y_0 \in \overline{D(A)}$, and x is arbitrary in $D(A)$, it follows that the sequence z_ε of ε-approximate solutions satisfies the Cauchy criterion and so $y(t) = \lim_{\varepsilon\to 0} z_\varepsilon(t)$ exists uniformly on $[0,T]$. Now, we take the limit as $\varepsilon \to 0$ in (4.36) with $s = t + h$, $g \equiv f$, and $z_0 = w_0 = y_0$. We get

$$\|y(t+h) - y(t)\| \le G(b,\varphi)(t+h,t) = e^{\omega t}(e^{\omega h} + 1)\|y_0 - x\|$$
$$+ \int_0^h e^{\omega(h-\tau)}(\|f(\tau)\| + \|v\|)d\tau + \int_0^t e^{\omega(t-\tau)}\|f(\tau+h) - f(\tau)\|d\tau, \qquad \forall [x,v] \in A,$$

and therefore y is continuous on $[0,T]$. □

Now, by (4.37) we have, for $f \equiv g,\ t = s$,

$$\|z(t) - y(t)\| \le \delta(\varepsilon), \qquad \forall t \in [0,T],$$

where z is any ε-approximate solution and $\delta(\varepsilon) \to 0$ as $\varepsilon \to 0$. Finally, we take $t = s$ in (4.37) and let ε tend to zero. Then, by (4.29), we get the inequality

$$\|y(t) - \bar{y}(t)\| \le e^{\omega t}\|y(0) - \bar{y}(0)\| + \int_0^t e^{\omega(t-\tau)}\|f(\tau) - g(\tau)\|d\tau.$$

To obtain (4.12), we apply inequality (4.37), where

$$\varphi(t,s) = [y(t) - \bar{y}(t), f(t) - g(s)]_s \quad \text{and} \quad t = s.$$

Then, by (4.29), we see that

$$G(h,\varphi)(t,t) = e^{\omega t}\|y(0) - \bar{y}(0)\| + \int_0^t e^{\omega(t-s)}[y(s) - \bar{y}(s), f(s) - g(s)]_s ds,$$

and so (4.12) follows for $s = 0$ and, consequently, for all $s \in (0,t)$.

Thus, the proof of Theorem 4.1 is complete.

The convergence theorem can be made more precise for the autonomous equation (4.16); that is, for $f \equiv 0$.

Corollary 4.3. *Let A be ω-accretive and satisfy condition* (4.15)*, and let $y_0 \in \overline{D(A)}$. Let y be the mild solution to problem* (4.16) *and let y_ε be an ε-approximate solution to* (4.16) *with $y_\varepsilon(0) = y_0$. Then,*

$$\|y_\varepsilon(t) - y(t)\| \le C_T(\|y_0 - x\| + |Ax|(\varepsilon + t^{1/2}\varepsilon^{1/2})), \qquad \forall t \in [0,T], \tag{4.38}$$

for all $x \in D(A)$. In particular, we have

$$\left\| y(t) - \left(I + \frac{t}{n}A\right)^{-n} y_0 \right\| \le C_T(\|y_0 - x\| + tn^{1/2}|Ax|) \tag{4.38$'$}$$

for all $t \in [0,T]$ and $x \in D(A)$. Here, C_T is a positive constant independent of x and y_0 and $|Ax| = \inf\{\|z\|;\ z \in Ax\}$.

Proof. The mappings $y_0 \to y$ and $y_0 \to y_\varepsilon$ are Lipschitz continuous with Lipschitz constant $e^{\omega T}$, thus it suffices to prove estimate (4.38) for $y_0 \in D(A)$.

By estimate (4.36), we have, for all $T > 0$,

$$\begin{aligned}&\|G(b,0) - H_\Delta(B,0)\|_{L^\infty(\Omega(\Delta))}\\ &\le \|b - \widetilde{b}\|_{L^\infty(-T,T)} + \|B - \widetilde{b}\|_{L^\infty(-T,T)} + C\varepsilon(\|\widetilde{\psi}_{tt}\|_{L^\infty(\Omega)} + \|\widetilde{\psi}_{ss}\|_{L(\Omega)}),\end{aligned}$$

where $\overline{\psi} = G(\widetilde{b},0)$, $\widetilde{b}$ is a sufficiently smooth function on $[-T,T]$, $\Omega = (0,T) \times (0,T)$, and C is independent of ε, b, and B. We apply this inequality for B and b as in the proof of Theorem 4.1; that is,

$$b(t) = \omega^{-1}(e^{\omega|t|} - 1)|Ax|, \qquad \forall t \in [-T,T].$$

Then, we have

$$b'(t) = e^{\omega|t|}|Ax|\,\mathrm{sign}\,t,$$

and we approximate the signum function $\mathrm{sign}\,t$ by

$$\theta(t) = \begin{cases} \dfrac{t}{\lambda} & \text{for } |t| \le \lambda, \\ \dfrac{t}{|t|} & \text{for } |t| > \lambda, \end{cases}$$

and so, we construct a smooth approximation $\widetilde{b}$ of b such that

$$\widetilde{b}(0) = 0, \widetilde{b}'(t) = e^{\omega|t|}Ax\theta(t),$$

and

$$\widetilde{b}''(t) = \omega\theta(t)|Ax|e^{\omega|t|} + \theta'(t)|Ax|e^{\omega|t|}.$$

Hence,

$$\sup\{|\widetilde{b}''(s)|;\ 0 \le s \le t\} \le e^{\omega|t|}|Ax|(\omega + \lambda^{-1})$$

and, therefore,

$$\begin{aligned} &\|b - \widetilde{b}\|_{L^\infty(-t,t)} + C\varepsilon(\|\widetilde{\psi}_{tt}\|_{L^\infty((0,t)\times(0,t))} + \|\widetilde{\psi}_{ss}\|_{L^\infty((0,t)\times(0,t))}) \\ &\le Ct\varepsilon|Ax|(1+\lambda^{-1}) + C\lambda|Ax|, \qquad \forall t \in [0,T], \end{aligned}$$

where C depends on T only.

Similarly, we have

$$\|B - \widetilde{b}\|_{L^\infty(-t,t)} \le C(\varepsilon + \lambda)|Ax|.$$

Finally,

$$\|G(b,0) - H_\Delta(B,0)\|_{L^\infty(\Omega_t(\Delta))} \le C(\varepsilon + \lambda + t\varepsilon\lambda^{-1})|Ax|,$$

where $\Omega_t = (0,t) \times (0,t)$. This implies that (see the proof of Theorem 4.1)

$$\|y_\varepsilon(t) - y(t)\| \le G(b,0)(t,t) + C|Ax|(\varepsilon + \lambda + t\varepsilon\lambda^{-1})$$

for all $t \in [0,T]$ and all $\lambda > 0$. For $\lambda = (t\varepsilon)^{1/2}$, this yields

$$\|y_\varepsilon(t) - y(t)\| \le C|Ax|(\varepsilon + t^{1/2}\varepsilon^{1/2}), \qquad \forall t \in [0,T],$$

which completes the proof. □

Regularity of Mild Solutions

A question of great interest is that of circumstances under which the mild solutions are strong solutions. One may construct simple examples which show that in a general Banach space this might be false. However, if the space is reflexive, then under natural assumptions on A, f, and y_ε the answer is positive.

Theorem 4.4. *Let X be reflexive and let A be closed and ω-accretive, and let A satisfy assumption* (4.13). *Let $y_0 \in D(A)$ and $f \in W^{1,1}([0,T];X)$ be such that $f(t) \in C$, $\forall t \in [0,T]$. Then, problem* (4.1) *has a unique mild strong solution y which is strong solution and $y \in W^{1,\infty}([0,T];X)$. Moreover, y satisfies the estimate*

$$\left\|\frac{dy}{dt}(t)\right\| \le e^{\omega t}|f(0)-Ay_0| + \int_0^t e^{\omega(t-s)}\left\|\frac{df}{ds}(s)\right\| ds, \quad \text{a.e. } t\in(0,T), \tag{4.39}$$

where $|f(0)-Ay_0| = \inf\{\|w\|;\ w \in f(0)-Ay_0\}$.

In particular, we have the following theorem.

Theorem 4.5. *Let X be a reflexive Banach space and let A be an ω-m-accretive operator. Then, for each $y_0 \in D(A)$ and $f \in W^{1,1}([0,T];X)$, problem* (4.1) *has a unique strong solution $y \in W^{1,\infty}([0,T];X)$ that satisfies estimate* (4.39).

Proof of Theorem 4.4. Let y be the mild solution to problem (4.1) provided by Theorem 4.2. We apply estimate (4.14), where $y(t) := y(t+h)$ and $g(t) := f(t+h)$. We get

$$\begin{aligned}\|y(t+h)-y(t)\| &\le \|y(h)-y(0)\|e^{\omega t} + \int_0^t \|f(s+h)-f(s)\|e^{\omega(t-s)}ds\\ &\le Ch + \|y(h)-y(0)\|e^{\omega t},\end{aligned}$$

because $f \in W^{1,1}([0,T];X)$ (see Theorem 1.18 and Remark 1.2). Now, applying the same estimate (4.14) to y and y_0, we get

$$\|y(h)-y_0\| \le \int_0^h \|f(s)-\xi\|e^{\omega(h-s)}ds \le \int_0^h |Ay_0 - f(s)|ds,$$
$$\forall \xi \in Ay_0,\ h \in [0,T].$$

We may conclude, therefore, that the mild solution y is Lipschitz on $[0,T]$. Then, by Theorem 1.17, it is, a.e., differentiable and belongs to $W^{1,\infty}([0,T];X)$. Moreover, we have

$$\left\|\frac{dy}{dt}(t)\right\| = \lim_{h\to 0}\frac{\|y(t+h)-y(t)\|}{h} \le e^{\omega t}|Ay_0 - f(0)| + \int_0^t \left\|\frac{df}{ds}(s)\right\| e^{\omega(t-s)}ds,$$
$$\text{a.e. } t \in (0,T).$$

Now, let $t \in [0,T]$ be such that

$$\frac{dy}{dt}(t) = \lim_{h\to 0}\frac{1}{h}(y(t+h)-y(t))$$

exists. By inequality (4.18), we have

$$\|y(t+h)-x\| \leq e^{\omega h}\|y(t)-x\| + \int_t^{t+h} e^{\omega(t+h-s)}[y(\tau)-x, f(\tau)-w]_s d\tau,$$
$$\forall [x,w] \in A.$$

Noting that

$$[v-x, u-v]_s \leq \|u-x\| - \|v-x\|, \qquad \forall u,v,x \in X,$$

we get

$$\begin{aligned} &[y(t)-x, y(t+h)-y(t)]_s \\ &\leq (e^{\omega h}-1)\|y(t)-x\| + \int_t^{t+h} e^{\omega(t+h-\tau)}[y(\tau)-x, f(\tau)-w]_s d\tau. \end{aligned}$$

Because the bracket $[u,v]_s$ is upper semicontinuous in (u,v), and positively homogeneous and continuous in v (see Proposition 3.7), this yields

$$\left[y(t)-x, \frac{dy}{dt}(t)\right]_s - \omega\|y(t)-x\| \leq [y(t)-x, f(t)-w]_s, \qquad \forall [x,w] \in A.$$

Taking into account part (v) of Proposition 3.7, this implies that there is $\xi \in J(y(t)-x)$ such that (J is the duality mapping)

$$\left(\frac{dy}{dt}(t) - \omega(y(t)-x) - f(t) - w, \xi\right) \leq 0. \tag{4.40}$$

Inasmuch as the function y is differentiable in t, we have

$$y(t-h) = y(t) - h\frac{d}{dt}y(t) + hg(h), \tag{4.41}$$

where $g(h) \to 0$ for $h \to 0$. On the other hand, by condition (4.13), for every h sufficiently small and positive, there are $[x_h, w_h] \in A$ such that

$$y(t-h) + hf(t) = x_h + hw_h.$$

Substituting successively in (4.30) and in (4.41) we get

$$(1-\omega h)\|y(t)-x_h\| \leq h\|g(h)\|, \qquad \forall h \in (0, \lambda_0).$$

Hence, $x_h \to y(t)$ and $w_h \to f(t) - dy(t)/dt$ as $h \to 0$. Because A is closed, we conclude that

$$\frac{dy}{dt}(t) + Ay(t) \ni f(t),$$

as claimed.

Remark 4.1. In particular, Theorems 4.1–4.5 remain true for equations of the form

$$\begin{cases} \dfrac{dy}{dt}(t) + Ay(t) + Fy(t) \ni f(t), & t \in [0,T], \\ y(0) = y_0, \end{cases} \tag{4.42}$$

where A is m-accretive in $X \times X$ and $F : X \to X$ is Lipschitzian. Indeed, in this case, as easily seen, the operator $A+F$ is quasi-m-accretive; that is, $A+F+\omega I$ is m-accretive for $\omega = \|F\|_{\mathrm{Lip}}$.

More can be said about the regularity of a strong solution to problem (4.1) if the space X is uniformly convex.

Theorem 4.6. *Let A be ω-m-accretive, $f \in W^{1,1}([0,T];X)$, $y_0 \in D(A)$ and let X be uniformly convex along with the dual X^*. Then, the strong solution to problem* (4.1) *is everywhere differentiable from the right, $(d^+/dt)y$ is right continuous, and*

$$\frac{d^+}{dt}y(t) + (Ay(t) - f(t))^0 = 0, \qquad \forall t \in [0,T), \tag{4.43}$$

$$\left\|\frac{d^+}{dt}y(t)\right\| \le e^{\omega t}\|(Ay_0 - f(0))^0\| + \int_0^t e^{\omega(t-s)}\left\|\frac{df}{ds}(s)\right\|ds, \qquad \forall t \in [0,T). \tag{4.44}$$

Here, $(Ay-f)^0$ is the element of minimum norm in the set $Ay - f$.

Proof. Because X and X^* are uniformly convex, Ay is a closed convex subset of X for every $x \in D(A)$ (see Section 3.1) and so, $(Ay(t)-f(t))^0$ is well defined.

Let $y \in W^{1,\infty}([0,T];X)$ be the strong solution to (4.1). We have

$$\frac{d}{dh}(y(t+h) - y(t)) + Ay(t+h) \ni f(t+h), \quad \text{a.e. } h > 0,\ t \in (0,T),$$

and because A is ω-accretive, this yields

$$\left(\frac{d}{dh}(y(t+h) - y(t)), \xi\right) \le \omega\|y(t+h)-y(t)\|^2 + (f(t+h) - \eta(t), \xi),$$
$$\forall \eta(t) \in Ay(t),$$

where $\xi = J(y(t+h) - y(t))$.

Then, by Lemma 4.1, we get

$$\|y(t+h) - y(t)\| \le \int_0^h e^{\omega(h-s)}\|\eta(t) - f(t+s)\|ds, \tag{4.45}$$

which yields

$$\left\|\frac{dy}{dt}(t)\right\| \le \|f(t) - \eta(t)\|, \qquad \forall \eta(t) \in Ay(t), \quad \text{a.e. } t \in (0,T).$$

In other words,

$$\left\|\frac{dy}{dt}(t)\right\| \le \|(Ay(t)-f(t))^0\|, \quad \text{a.e. } t\in(0,T),$$

and because $dy(t)/dt + Ay(t) \ni f(t)$, a.e. $t\in(0,T)$, we conclude that

$$\frac{dy}{dt}(t) + (Ay(t)-f(t))^0 = 0, \quad \text{a.e. } t\in(0,T). \tag{4.46}$$

Observe also that, for all h, y satisfies the equation

$$\frac{d}{dt}(y(t+h)-y(t)) + Ay(t+h) - Ay(t) \ni f(t+h)-f(t), \quad \text{a.e. in } (0,T).$$

Multiplying this equation by $J(y(t+h)-y(t))$ and using the ω-accretivity of A, we see by Lemma 4.1 that

$$\frac{d}{dt}\|y(t+h)-y(t)\| \le \omega\|y(t+h)-y(t)\| + \|f(t+h)-f(t)\|,$$
$$\text{a.e. } t, t+h \in (0,T),$$

and therefore

$$\begin{aligned} &\|y(t+h)-y(t)\| \\ &\le e^{\omega(t-s)}\|y(s+h)-y(s)\| + \int_s^t e^{\omega(t-\tau)}\|f(\tau+h)-f(\tau)\|d\tau. \end{aligned} \tag{4.47}$$

Finally,

$$\begin{aligned} \left\|\frac{dy}{dt}(t)\right\| \le e^{\omega(t-s)}\left\|\frac{dy}{ds}(s)\right\| + \int_s^t e^{\omega(t-\tau)}\left\|\frac{df}{d\tau}(\tau)\right\|d\tau, \\ \text{a.e. } 0<s<t<T. \end{aligned} \tag{4.48}$$

Similarly, multiplying the equation

$$\frac{d}{dt}(y(t)-y_0) + Ay(t) \ni f(t), \quad \text{a.e. } t\in(0,T),$$

by $J(y(t)-y_0)$ and, integrating on $(0,t)$, we get the estimate

$$\|y(t)-y_0\| \le \int_0^t e^{\omega(t-s)}\|(Ay_0-f(s))^0\|ds, \qquad \forall t\in[0,T], \tag{4.49}$$

and, substituting in (4.47) with $s=0$, we get

$$\begin{aligned} \left\|\frac{d}{dt}y(t)\right\| \le e^{\omega t}\|(Ay_0-f(0))^0\| + \int_0^t e^{\omega(t-s)}\left\|\frac{df}{ds}(s)\right\|ds, \\ \text{a.e. } t\in(0,T). \end{aligned} \tag{4.50}$$

Because A is demiclosed (see Proposition 3.4) and X is reflexive, it follows by (4.46) and (4.50) that $y(t) \in D(A)$, $\forall t \in [0,T]$, and

$$\|(Ay(t) - f(t))^0\| \leq C, \qquad \forall t \in [0,T]. \tag{4.51}$$

Let us show now that (4.46) extends to all $t \in [0,T]$. For t arbitrary but fixed in $[0,T]$, consider $h_n \to 0$ such that $h_n > 0$ for all n and

$$\frac{y(t+h_n) - y(t)}{h_n} \rightharpoonup \xi \quad \text{in } X \ \text{ as } n \to 0.$$

By (4.46) and the previous estimates, we see that

$$\|\xi\| \leq \|(Ay(t) - f(t))^0\|, \qquad \forall t \in [0,T], \tag{4.52}$$

and $\xi \in f(t) - Ay(t)$ because A is demiclosed. Indeed, we have

$$f(t) - \xi = w - \lim_{n\to\infty} \frac{1}{h_n} \int_t^{t+h_n} \eta(s)ds,$$

where $\eta \in L^\infty(0,T;X)$ and $\eta(t) \in Ay(t)$, $\forall t \in [0,T]$.

We set $\eta_n(s) = \eta(t + sh_n)$ and $y_n(s) = y(t + sh_n)$. If we denote again by A the realization of A in $L^2(0,T;X) \times L^2(0,T;X)$, we have $y_n \to y(t)$ in $L^2(0,T;X)$, $\eta_n \to f(t) - \xi$ weakly in $L^2(0,T;X)$.

Because A is demiclosed in $L^2(0,T;X) \times L^2(0,T;X)$ we have that $f(t) - \xi \in Ay(t)$, as claimed. Then, by (4.52) we conclude that $\xi = (Ay(t) - f(t))^0$ and, therefore,

$$\frac{d^+}{dt} y(t) = \lim_{h\downarrow 0} \frac{y(t+h) - y(t)}{h} = -(Ay(t) - f(t))^0, \qquad \forall t \in [0,T).$$

Next, we see by (4.47) that

$$\left\|\frac{d^+}{dt} y(t)\right\| \leq e^{\omega(t-s)} \left\|\tfrac{d^+}{dt} y(s)\right\| + \textstyle\int_s^t e^{\omega(t-\tau)} \left\|\tfrac{df}{d\tau}(\tau)\right\| d\tau, \qquad 0 \leq s \leq t \leq T. \tag{4.53}$$

Let $t_n \to t$ be such that $t_n > t$ for all n. Then, on a subsequence, again denoted by t_n,

$$\frac{d^+ y(t_n)}{dt} = -(Ay(t_n) - f(t_n))^0 \rightharpoonup \xi,$$

where $-\xi \in Ay(t) - f(t)$ (because A is demiclosed). On the other hand, it follows by (4.53) that

$$\|\xi\| \leq \limsup_{n\to\infty} \|(Ay(t_n) - f(t_n))^0\| \leq \|(Ay(t(-f(t))^0\|.$$

Hence, $\xi = -(Ay(t) - f(t))^0$ and $(d^+/dt)y(t_n) \to \xi$ strongly in X (because X is uniformly convex). We have, therefore, proved that $(d^+/dt)y(t)$ is right continuous on $[0,T)$, thereby completing the proof. □

In particular, it follows by Theorem 4.6 that, if A is quasi-m-accretive, $y_0 \in D(A)$, and X, X^* are uniformly convex, then the solution y to the autonomous problem (4.16) is everywhere differentiable from the right and

$$\frac{d^+}{dt} y(t) + A^0 y(t) = 0, \qquad \forall t \geq 0, \tag{4.54}$$

where A^0 is the minimal section of A. Moreover, the function $t \to A^0 y(t)$ is continuous from the right on $\mathbf{R}^+$.

It turns out that this result remains true under weaker conditions on A. Namely, one has the following.

Theorem 4.7. *Let A be ω-accretive, closed, and satisfy the condition*

$$\overline{\operatorname{conv} D(A)} \subset \bigcap_{0<\lambda<\lambda_0} R(I+\lambda A) \quad \textit{for some } \lambda_0 > 0. \tag{4.55}$$

Let X and X^* be uniformly convex. Then, for every $x \in D(A)$ the set Ax has a unique element of minimum norm A^0x, and for every $y_0 \in D(A)$ the Cauchy problem (4.16) *has a unique strong solution $y \in W^{1,\infty}([0,\infty);X)$, which is everywhere differentiable from the right and*

$$\frac{d^+}{dt} y(t) + A^0 y(t) = 0, \qquad \forall t \geq 0. \tag{4.56}$$

Moreover, the function $t \to A^0 y(t)$ is continuous from the right and

$$\left\| \frac{d^+}{dt} y(t) \right\| \leq e^{\omega t} \|A^0 y_0\|, \qquad \forall t \geq 0. \tag{4.57}$$

The result extends to nonhomogeneous equation (4.1) with $f \in W^{1,\infty}([0,T];X)$.

Proof. We assume first that A is demiclosed in $X \times X$.

Define the set $B \subset X \times Y$ by

$$Bx = \overline{\operatorname{conv} Ax}, \qquad x \in D(B) = D(A).$$

It is readily seen that B is ω-accretive. Moreover, by (4.55) it follows that

$$D(A) \subset \bigcap_{0<\lambda<\lambda_0} R(I+\lambda B).$$

Let $x \in D(A)$. Then, $x_\lambda = (I+\lambda A)^{-1}x$ and $y_\lambda = A_\lambda x$ are well defined for $0 < \lambda < \lambda_0$. Moreover, $\|A_\lambda x\| \leq |Ax| = \inf\{\|w\|;\ w \in Ax\}$ and $x_\lambda \to x$ for $\lambda \to 0$ (see Proposition 3.2). Let $\lambda_n \to 0$ be such that $A_{\lambda_n} x \rightharpoonup y$. Because $A_{\lambda_n} x \in Ax_{\lambda_n}$ and A is demiclosed, it follows that $y \in Ax$. On the other hand, we have

$$\|A_\lambda x\| = \|B_\lambda x\| \le |Bx| = \|B^0 x\|.$$

(B^0x exists and is unique because the set Bx is convex, closed, and X is uniformly convex.) This implies that $y = B^0x \in Ax$. Hence, Ax has a unique element of minimum norm A^0x. Then we may apply Theorem 4.6 to deduce that the strong solution y to problem (4.16) (which exists and is unique by Theorem 4.5) satisfies (4.56) and (4.57). (In the proof of Theorem 4.6, the quasi-m-accretivity has been used only to assure the existence of a strong solution, the demiclosedness of A, and the existence of A^0.)

To complete the proof, we turn now to the case where A is only closed. Let $\widetilde{A}$ be the closure of A in $X \times X_w$; that is, the smallest demiclosed extension of A. Clearly, $D(A) \subset D(\widetilde{A}) \subset \overline{D(A)}$ and $\widetilde{A}$ satisfies condition (4.55). Moreover, because the duality mapping J is continuous, it is easily seen that $\widetilde{A}$ is ω-accretive. Then, applying the first part of the proof, we conclude that problem

$$\frac{d^+u}{dt} + \widetilde{A}^0 u = 0 \qquad \text{in } [0,\infty),$$
$$u(0) = y_0,$$

has a unique solution u satisfying all the conditions of the theorem. To conclude the proof, it suffices to show that $D(\widetilde{A}) = D(A)$ and $\widetilde{A}^0 = A^0$.

Let $x \in D(\widetilde{A})$. Then, for each λ, there is $[x_\lambda, y_\lambda] \in A \subset \widetilde{A}$ such that

$$x = x_\lambda - \lambda y_\lambda \qquad \text{for } 0 < \lambda < \lambda_0.$$

We have $x_\lambda = (I + \lambda A)^{-1}x$ and $y_\lambda = A_\lambda x = \widetilde{A}_\lambda x$. Because $x \in D(\widetilde{A})$, we have that $x_\lambda \xrightarrow{\lambda \to 0} x$ and $\|y_\lambda\| \le |\widetilde{A}x| = \|\widetilde{A}^0 x\|$. As $\widetilde{A}$ is demiclosed and X is uniformly convex, this implies, by a standard device, that $y_\lambda \to \widetilde{A}^0 x$ as $\lambda \to 0$. Finally, because A is closed, this yields $\widetilde{A}^0 x \in Ax$ and $x \in D(A)$. Hence, $D(\widetilde{A}) = D(A)$ and $\widetilde{A}^0 x = A^0 x$, $\forall x \in D(A)$. The proof of Theorem 4.7 is complete. □

Remark 4.2. If the space X^* is uniformly convex, A is quasi-m-accretive, $f \in W^{1,1}([0,T];X)$, and $y_0 \in D(A)$, then the strong solution $y \in W^{1,\infty}([0,T];X)$ to problem (4.1) (see Theorem 4.4) can be obtained as

$$y(t) = \lim_{\lambda \to 0} y_\lambda(t) \quad \text{in } X, \text{ uniformly on } [0,T], \tag{4.58}$$

where $y_\lambda \in C^1([0,T];X)$ are the solutions to the Yosida approximating equation

$$\begin{cases} \dfrac{dy_\lambda}{dt}(t) + A_\lambda y_\lambda(t) = f(t), & t \in [0,T], \\ y_\lambda(0) = y_0, \end{cases} \tag{4.59}$$

where $A_\lambda = \lambda^{-1}(I - (I + \lambda A)^{-1})$ for $0 < \lambda < \lambda_0$. Here is the argument that also provides a simple proof of Theorem 4.4 in this special case. By Lemma 4.2, we have

$$\frac{1}{2}\,\frac{d}{dt}\,\|y_\lambda(t)-y_\mu(t)\|^2+(A_\lambda y_\lambda(t)-A_\mu y_\mu(t),J(y_\lambda(t)-y_\mu(t)))=0,$$
$$\text{a.e. } t\in(0,T),\quad \text{for all } \lambda,\mu\in(0,\lambda_0).$$

Inasmuch as A is ω-accretive and $A_\lambda y\in A(I+\lambda A)^{-1}y$, we get that

$$\begin{aligned}&\frac{1}{2}\,\frac{d}{dt}\,\|y_\lambda(t)-y_\mu(t)\|^2+(A_\lambda y_\lambda(t)-A_\mu y_\mu(t),J(y_\lambda(t)-y_\mu(t))\\&-J((I+\lambda A)^{-1}y_\lambda(t)-(1+\mu A)^{-1}y_\mu(t)))\\&\le\omega\|(1+\lambda A)^{-1}y_\lambda(t)-(1+\mu A)^{-1}y_\mu(t)\|^2,\quad \text{a.e. } t\in(0,T).\end{aligned}\tag{4.60}$$

On the other hand, multiplying the equation

$$\frac{d^2y_\lambda}{dt^2}+\frac{d}{dt}A_\lambda y_\lambda(t)=\frac{df}{dt},\quad \text{a.e. } t\in(0,T),$$

by $J(dy_\lambda/dt)$, it yields

$$\frac{1}{2}\,\frac{d}{dt}\,\left\|\frac{dy_\lambda}{dt}(t)\right\|^2\le\left\|\frac{df}{dt}(t)\right\|\,\left\|\frac{dy_\lambda}{dt}(t)\right\|+\omega\left\|\frac{dy_\lambda}{dt}(t)\right\|,\quad \text{a.e. } t\in(0,T),$$

because A_λ is ω-accretive. This implies that

$$\begin{aligned}\left\|\frac{dy_\lambda}{dt}(t)\right\|&\le e^{\omega t}\left\|\frac{dy_\lambda}{dt}(0)\right\|+\int_0^t e^{\omega(t-s)}\left\|\frac{df}{ds}(s)\right\|ds\\&\le e^{\omega t}|Ay_0-f(0)|+\int_0^t e^{\omega(t-s)}\left\|\frac{df}{ds}(s)\right\|ds.\end{aligned}\tag{4.61}$$

Hence, $\|A_\lambda y_\lambda(t)\|\le C$, $\forall\lambda\in(0,\lambda_0)$, and $\|y_\lambda(t)-(1+\lambda A)^{-1}y_\lambda(t)\|\le C\lambda$. Because J is uniformly continuous on bounded sets, it follows by (4.60) that

$$\begin{aligned}&\frac{1}{2}\,\frac{d}{dt}\,\|y_\lambda(t)-y_\mu(t)\|^2\le\omega\|(I+\lambda A)^{-1}y_\lambda(t)-(I+\mu A)^{-1}y_\mu(t)\|^2\\&+(\|A_\lambda y_\lambda(t)\|+\|A_\mu y_\mu(t)\|)\|J(y_\lambda(t)-y_\mu(t))-J((I+\lambda A)^{-1}y_\lambda(t)-(I+\mu A)^{-1}y_\mu(t))\|\\&\le\omega\|y_\lambda(t)-y_\mu(t)\|^2+C(\lambda+\mu)\\&+\|J(y_\lambda(t)-y_\mu(t))-J((I+\lambda A)^{-1}y_\lambda(t)-(1+\mu A)^{-1}y_\mu(t))\|,\end{aligned}$$

because $\|(I+\lambda A)^{-1}y_\lambda-y_\lambda\|=\lambda\|A_\lambda y_\lambda\|\le C\lambda$. Then, taking into account that J is uniformly continuous and that, by (4.59) and (4.61), $\{\|A_\lambda y_\lambda\|\}$ is bounded, the latter implies, via Gronwall's lemma, that $\{y_\lambda\}$ is a Cauchy sequence in the space $C([0,T];X)$ and $y(t)=\lim_{\lambda\to 0}y_\lambda(t)$ exists in X uniformly on $[0,T]$. Let $[x,w]$ be arbitrary in A and let $x_\lambda=x+\lambda w$. Multiplying equation (4.59) by $J(y_\lambda(t)-x_\lambda)$ and integrating on $[s,t]$, we get

$$\frac{1}{2}\|y_\lambda(t)-x_\lambda\|^2 \leq \frac{1}{2}\|y_\lambda(s)-x_\lambda\|^2 e^{\omega(t-s)} + \int_s^t e^{\omega(t-\tau)}(f(\tau)-w, J(y_\lambda(\tau)-x_\lambda))d\tau,$$

and, letting $\lambda \to 0$,

$$\frac{1}{2}\|y(t)-x\|^2 \leq \frac{1}{2}\|y(s)-x\|^2 e^{\omega(t-s)} + \int_s^t e^{\omega(t-\tau)}(f(\tau)-w, J(y_\lambda(\tau)-x))d\tau,$$

because J is continuous. This yields

$$\left(\frac{y(t)-y(s)}{t-s}, J(y(s)-x)\right) \leq \frac{1}{2}\|y(s)-x\|^2(e^{\omega(t-s)}-1)(t-s)^{-1} + \frac{1}{t-s}\int_s^t e^{\omega(t-\tau)}(f(\tau)-w, J(y_\lambda(\tau)-x))d\tau, \tag{4.62}$$

because, as seen earlier,

$$\frac{1}{2}\|y(t)-x\|^2 - \frac{1}{2}\|y(s)-x\|^2 \geq (y(t)-x, J(y(s)-x)).$$

By (4.61), we see that y is absolutely continuous on $[0,T]$ and $dy/dt \in L^\infty(0,T;X)$. Hence, y is, a.e., differentiable on $(0,T)$. If $s=t_0$ is a point where y is differentiable, by (4.62) we see that

$$\left(f(t_0)-\frac{dy}{dt}(t_0)-w+\omega(y(t_0)-x), J(y(t_0)-x)\right) \geq 0, \qquad \forall [x,w] \in A.$$

Because $A+\omega I$ is m-accretive, this implies that

$$f(t_0)-\frac{dy}{dt}(t_0) \in Ay(t_0).$$

Hence, y is the strong solution to problem (4.1).

Local Lipschitzian Perturbations

Consider the Cauchy problem

$$\begin{cases} \dfrac{dy}{dt}(t)+Ay(t)+Fy(t) \ni f(t), & t \in [0,T], \\ y(0)=y_0, \end{cases} \tag{4.63}$$

where A is quasi-m-accretive in $X \times X$ and $F : X \to X$ is locally Lipschitz; that is,

$$\|Fu - Fv\| \le L_R \|u - v\|, \qquad \forall u, v \in B_R,\ \forall R > 0, \tag{4.64}$$

where $B_R = \{u \in X;\ \|u\| \le R\}$.

We have the following.

Theorem 4.8. *Let X be a reflexive Banach space and let A be a quasi-m-accretive operator in X. Let $f \in W^{1,1}([0,T];X)$ and let $F : X \to X$ be locally Lipschitz. Then, for each $y_0 \in D(A)$ there is $T(y_0) \in (0,T)$ and a function $y \in W^{1,\infty}([0,T(y_0)];X)$ such that*

$$\begin{cases} \dfrac{dy}{dt}(t) + Ay(t) + Fy(t) \ni f(t), & a.e.\ t \in (0, T(y_0)), \\ y(0) = y_0. \end{cases} \tag{4.65}$$

Assume further that

$$(Fy, w) \ge -\gamma_1 \|y\|^2 + \gamma_2, \qquad \forall [y, w] \in J. \tag{4.66}$$

Then, the solution y to (4.65) *is global; that is, it exists on all of* $[0,T]$.

Proof. We truncate F on X as follows

$$F_R(y) = \begin{cases} F(y) & \text{if } \|y\| \le R \\ F\left(\dfrac{Ry}{\|y\|}\right) & \text{if } \|y\| > R \end{cases} \tag{4.67}$$

and notice that F_R is Lipschitz on X:

$$\|F_R(x) - F_R(y)\| \le L_R^1 \|x - y\|, \qquad \forall x, y \in X, \tag{4.68}$$

for some $L_R^1 > 0$. The latter is obvious if $\|x\|, \|y\| \le R$ or if $\|x\|, \|y\| > R$. If $\|x\| \le R$ and $\|y\| > R$, we have

$$\begin{aligned} &\|F_R(x) - F_R(y)\| = \left\| F(x) - F\left(\frac{Ry}{\|y\|}\right) \right\| \le L_R \left\| x - \frac{Ry}{\|y\|} \right\| \\ &\le L_R R^{-1} \|x\|y\| - Ry\| \le L_R R^{-1} \|R(x-y) + x(\|y\| - R)\| \le 2L_R \|x - y\|. \end{aligned} \tag{4.69}$$

Then, (4.69) implies that F_R is Lipschitz continuous and so $A + F_R$ is quasi-m-accretive. Hence for each $R > 0$ there is a unique strong solution y_R to equation

$$\begin{cases} \dfrac{dy_R}{dt}(t) + Ay_R(t) + F_R(y_R(t)) \ni f(t), & \text{a.e. } t \in (0,T), \\ y_R(0) = y_0. \end{cases} \tag{4.70}$$

Multiplying (4.70) by $w \in J(y_R)$ and using the quasi-accretivity of A, we get (without any loss of generality we assume that $0 \in A0$)

$$\frac{d}{dt} \|y_R(t)\| \le L_R^1 \|y_R(t)\| + \|f(t)\|, \quad \text{a.e. } t \in (0,T)$$

and therefore

$$\|y_R(t) \le e^{L_R^1 t}\|y_0\| + \int_0^t e^{L_R^1(t-s)}\|f(s)\|ds \le e^{L_R^1 t}\|y_0\| + \frac{M}{L_R^1}\,(e^{L_R^1 t} - 1), \qquad \forall t \in (0,T).$$

This yields

$$\|y_R(t)\| \le R$$

for $0 \le t \le T_R$ and $R > 0$ sufficiently large if $T_R > 0$ is suitably chosen.

Hence on $[0,T_R]$, $\|y_R(t)\| \le R$ and so equation (4.70) reduces on this interval to (4.63). This means that (4.63) has a unique solution y on $[0,T_R]$.

If we assume (4.66), then by (4.70) we see that

$$\frac{1}{2}\,\frac{d}{dt}\,\|y_R(t)\|^2 \le \gamma_1\|y_R(t)\|^2 + \gamma_2, \quad \text{a.e. } t \in (0,T).$$

Hence

$$\|y_R(t)\|^2 \le e^{2\gamma_1 t}\|y_0\|^2 + \frac{\gamma_2}{\gamma_1}\,(e^{2\gamma_1 T} - 1) \le R \quad \text{for } \ t \in [0,T]$$

if R is sufficiently large. Hence, for such R, y_R is the solution to (4.65) on all of $[0,T]$. □

The Cauchy Problem Associated with Demicontinuous Monotone Operators

We are given a Hilbert space H and a reflexive Banach space V such that $V \subset H$ continuously and densely. Denote by V' the dual space. Then, identifying H with its own dual, we may write

$$V \subset H \subset V'$$

algebraically and topologically.

The norms of V and H are denoted $\|\cdot\|$ and $|\cdot|$, respectively. We denote by (v_1,v_2) the pairing between $v_1 \in V'$ and $v_2 \in V$; if $v_1,v_2 \in H$, this is the ordinary inner product in H. Finally, we denote by $\|\cdot\|_*$ the norm of V' (which is the dual norm). In addition to these spaces, we are given a single-valued, monotone operator $A : V \to V'$. We assume that A is demicontinuous and coercive from V to V'.

We begin with the following simple application of Theorem 4.6.

Theorem 4.9. *Let $f \in W^{1,1}([0,T];H)$ and $y_0 \in V$ be such that $Ay_0 \in H$. Then, there exists one and only one function $y : [0,T] \to V$ that satisfies*

$$y \in W^{1,\infty}([0,T];H), \qquad Ay \in L^\infty(0,T;H), \tag{4.71}$$

$$\begin{cases} \dfrac{dy}{dt}\,(t) + Ay(t) = f(t), \quad \text{a.e. } t \in (0,T), \\ y(0) = y_0. \end{cases} \tag{4.72}$$

Moreover, y is everywhere differentiable from the right (in H) and

$$\frac{d^+}{dt}y(t)+Ay(t)=f(t), \qquad \forall t\in[0,T).$$

Proof. Define the operator $A_H : H \to H$,

$$A_H u = Au, \qquad \forall u \in D(A_H) = \{u \in V;\ Au \in H\}. \tag{4.73}$$

By hypothesis, the operator $u \to u + Au$ is monotone, demicontinuous, and coercive from V to V'. Hence, it is surjective (see, e.g., Corollary 2.1) and so, A_H is m-accretive (maximal monotone) in $H \times H$. Then, we may apply Theorem 4.6 to conclude the proof. □

Now, we use Theorem 4.9 to derive a classical existence result due to Lions [40].

Theorem 4.10. *Let $A : V \to V'$ be a demicontinuous monotone operator that satisfies the conditions*

$$(Au,u) \geq \omega\|u\|^p + C_1, \qquad \forall u \in V, \tag{4.74}$$

$$\|Au\|_* \leq C_2(1+\|u\|^{p-1}), \quad \forall u \in V, \tag{4.75}$$

where $\omega > 0$ and $p > 1$. Given $y_0 \in H$ and $f \in L^q(0,T;V')$, $1/p+1/q=1$, there exists a unique absolutely continuous function $y : [0,T] \to V'$ that satisfies

$$y \in C([0,T];H) \cap L^p(0,T;V) \cap W^{1,q}([0,T];V'), \tag{4.76}$$

$$\frac{dy}{dt}(t) + Ay(t) = f(t), \quad a.e.\ t \in (0,T), \qquad y(0) = y_0, \tag{4.77}$$

where d/dt is considered in the strong topology of V'.

Proof. Assume that $y_0 \in D(A_H)$ and $f \in W^{1,1}([0,T];H)$. By Theorem 4.9, there is $y \in W^{1,\infty}([0,T];H)$ with $Ay \in L^\infty(0,T;H)$ satisfying (4.77). Then, by assumption (4.74), multiplying equation by $y(t)$ (scalarly in H), we have

$$\frac{1}{2}\frac{d}{dt}|y(t)|^2 + \omega\|y(t)\|^p \leq \|f(t)\|_*\|y(t)\|, \quad \text{a.e. } t \in (0,T)$$

(see Theorem 1.18) and, therefore,

$$|y(t)|^2 + \int_0^t \|y(s)\|^p ds \leq C\left(|y_0|^2 + \int_0^t \|f(s)\|_*^q ds\right), \qquad \forall t \in [0,T]. \tag{4.78}$$

Then, by (4.75), we get

$$\int_0^T \left\|\frac{dy}{dt}(t)\right\|_*^q dt \leq C\left(|y_0|^2 + \int_0^T \|f(t)\|_*^q dt\right). \tag{4.79}$$

(We denote by C several positive constants independent of y_0 and f.) Let us show now that $D(A_H)$ is a dense subset of H. Indeed, if x is any element of H, we set $x =$

$(I+\varepsilon A_H)^{-1}x$ (I is the unity operator in H). Multiplying the equation $x_\varepsilon + \varepsilon A x_\varepsilon = x$ by x_ε, it follows by (4.74) and (4.75) that

$$|x_\varepsilon|^2 + \omega\varepsilon\|x_\varepsilon\|^p \le |x_\varepsilon|\,|x| + C\varepsilon, \qquad \forall \varepsilon > 0,$$

and

$$\|x_\varepsilon - x\|_* \le \varepsilon\|Ax\|_* \le C\varepsilon(\|x_\varepsilon\|^{p-1}+1), \qquad \forall \varepsilon > 0.$$

Hence, $\{x_\varepsilon\}$ is bounded in H and $x_\varepsilon \to x$ in V' as $\varepsilon \to 0$. Therefore, $x_\varepsilon \rightharpoonup x$ in H as $\varepsilon \to 0$, which implies that $D(A_H)$ is dense in H.

Now, let $y_0 \in H$ and $f \in L^q(0,T;V')$. Then, there are the sequences $\{y_0^n\} \subset D(A_H)$, $\{f_n\} \subset W^{1,1}([0,T];H)$ such that

$$y_0^n \to y_0 \quad \text{in } H, \qquad f_n \to f \quad \text{in } L^q(0,T;V'),$$

as $n \to \infty$. Let $y_n \in W^{1,\infty}([0,T];H)$ be the solution to problem (4.77), where $y_0 = y_0^n$ and $f = f_n$. Because A is monotone, we have

$$\frac{1}{2}\,\frac{d}{dt}\,|y_n(t)-y_m(t)|^2 \le (f_n(t)-f_m(t), y_n(t-y_m(t)), \quad \text{a.e. } t \in (0,T).$$

Integrating from 0 to t, we get

$$\begin{aligned}&|y_n(t)-y_m(t)|^2\\ &\le |y_n^0-y_m^0|^2+2\left(\int_0^t \|f_n(s)-f_m(s)\|_*^q ds\right)^{1/q}\left(\int_0^t \|y_m(s)-y_m(s)\|^p ds\right)^{1/p}.\end{aligned} \tag{4.80}$$

On the other hand, it follows by estimates (4.78) and (4.79) that $\{y_n\}$ is bounded in $L^p(0,T;V)$ and $\{dy_n/dt\}$ is bounded in $L^q(0,T;V')$. Then, it follows by (4.80) that $y(t) = \lim_{n\to\infty} y_n(t)$ exists in H uniformly in t on $[0,T]$. Moreover, extracting a further subsequence if necessary, we have

$$\begin{aligned} y_n &\to y \qquad \text{weakly in } L^p(0,T;V),\\ \frac{y_n}{dt} &\to \frac{dy}{dt} \qquad \text{weakly in } L^q(0,T;V'),\end{aligned}$$

where dy/dt is considered in the sense of V'-valued distributions on$(0,T)$. In particular, we have proved that $y \in C([0,T];H) \cap L^p(0,T;V) \cap W^{1,q}([0,T];V')$. It remains to prove that y satisfies, a.e., on $(0,T)$ equation (4.77).

Let $x \in V$ be arbitrary but fixed. Multiplying the equation

$$\frac{dy_n}{dt} + Ay_n = f_n, \quad \text{a.e. } t \in (0,T)$$

by $y_n - x$ and integrating on (s,t), we get

$$\frac{1}{2}(|y_n(t)-x|^2 - |y_n(s)-x|^2) \le \int_s^t (f_n(\tau)-Ax, y_n(\tau)-x)d\tau.$$

Letting $n \to \infty$, it yields

$$\frac{1}{2}(|y(t)-x|^2 - |y(s)-x|^2) \le \int_s^t (f(\tau) - Ax, y(\tau) - x)d\tau.$$

Hence,

$$\left(\frac{y(t)-y(s)}{t-s}, y(s)-x\right) \le \frac{1}{t-s}\int_s^t (f(\tau)-Ax, y(\tau)-x)d\tau. \tag{4.81}$$

We know that y is, a.e., differentiable from $(0,T)$ into V' and

$$f(t_0) = \lim_{h\downarrow 0} \frac{1}{h}\int_{t_0}^{t_0+h} f(s)ds, \quad \text{a.e. } t_0 \in (0,T).$$

Let t_0 be such a point where y is differentiable. By (4.81), it follows that

$$\left(\frac{dy}{dt}(t_0) - f(t_0) + Ax, y(t_0) - x\right) \le 0,$$

and because x is arbitrary in V and A is maximal monotone in $V \times V'$, this implies that

$$\frac{dy}{dt}(t_0) + Ay(t_0) = f(t_0),$$

as claimed. □

It should be noted that compared with Theorem 4.6 and the previous results on the Cauchy problem (4.1), Theorem 4.10 provides a strong solution (in the V'-sense) under quite weak conditions on initial data and the nonhomogeneous term f. However, this class of problems is confined to those that have a variational formulation in a dual pairing (V,V').

As we show later on in Section 4.3, Theorem 4.10 remains true for time-dependent operators $A(t): V \to V'$ satisfying assumptions (4.74) and (4.75).

Continuous Semigroups of Contractions

Definition 4.4. Let C be a closed subset of a Banach space X. A *continuous semigroup of contractions on* C is a family of mappings $\{S(t);\ t \ge 0\}$ that maps C into itself with the properties:

(i) $S(t+s)x = S(t)S(s)x,\ \forall x \in C,\ t,s \ge 0.$

(ii) $S(0)x = x,\ \forall x \in C.$

(iii) For every $x \in C$, the function $t \to S(t)x$ is continuous on $[0,\infty)$.

(iv) $\|S(t)x - S(t)y\| \le \|x-y\|,\ \forall t \ge 0,\ x,y \in C.$

More generally, if instead of (iv) we have

(v) $\|S(t)x - S(t)y\| \le e^{\omega t}\|x-y\|,\ \forall t \ge 0,\ x,y \in C,$

we say that $S(t)$ is a continuous ω-quasi-contractive semigroup on C.

The operator $A_0 : D(A_0) \subset C \to X$, defined by

$$A_0 x = \lim_{t \downarrow 0} \frac{S(t)x - x}{t}, \qquad x \in D(A_0), \tag{4.82}$$

where $D(A_0)$ is the set of all $x \in C$ for which the limit (4.82) exists, is called the *infinitesimal generator* of the semigroup $S(t)$.

As in the case of strongly continuous semigroups of linear continuous operators, there is a close relationship between the continuous semigroups of contractions and accretive operators. Indeed, it is easily seen that $-A_0$ is accretive in $X \times X$. More generally, if $S(t)$ is quasi-contractive, then $-A_0$ is ω-accretive. Keeping in mind the theory of C_0-semigroups of contractions, one might suspect that there is a one-to-one correspondence between the class of continuous semigroups of contractions and that of m-accretive operators.

As seen in Theorem 4.3, if X is a Banach space and A is an ω-accretive mapping satisfying the range condition (4.15) (in particular, if A is ω-m-accretive), then, for every $y_0 \in \overline{D(A)}$, the Cauchy problem (4.16) has a unique mild solution $y(t) = S_A(t)y_0 = e^{-At}y_0$ given by the exponential formula (4.17); that is,

$$S_A(t)y_0 = \lim_{n \to \infty} \left(I + \frac{t}{n} A \right)^{-n} y_0. \tag{4.83}$$

(For this reason, $S_A(t)$ is, sometimes, denoted by e^{-At}.) We have the following.

Proposition 4.2. *$S_A(t)$ is a continuous ω-quasi-contractive semigroup on $C = \overline{D(A)}$.*

Proof. It is obvious that conditions (ii)–(iv) are satisfied as a consequence of Theorem 4.3. To prove (i), we note that, for a fixed $s > 0$, $y_1(t) = S_A(t+s)x$ and $y_2(t) = S_A(t)S_A(s)x$ are both mild solutions to the problem

$$\begin{cases} \dfrac{dy}{dt} + Ay = 0, & t \geq 0, \\ y(0) = S_A(s)x, \end{cases}$$

and so, by uniqueness of the solution we have $y_1 \equiv y_2$.

Let us assume now that X, X^* are uniformly convex Banach spaces and that A is an ω-accretive set that is closed and satisfies condition (4.55):

$$\overline{\operatorname{conv} D(A)} \subset \bigcap_{0<\lambda<\lambda_0} R(I + \lambda A) \quad \text{for some } \lambda_0 > 0. \tag{4.84}$$

Then, by Theorem 4.7, for every $x \in D(A)$, $S_A(t)x$ is differentiable from the right on $[0, +\infty)$ and

$$-A^0 x = \lim_{t \downarrow 0} \frac{S_A(t)x - x}{t}, \qquad \forall x \in D(A).$$

Hence, $-A^0 \subset A_0$, where A_0 is the infinitesimal generator of $S_A(t)$. $\square$

As a matter of fact, we may prove in this case the following partial extension of Hille–Philips theorem in continuous semigroups of contractions. (See A. Pazy [45].)

Proposition 4.3. *Let X and X^* be uniformly convex and let A be an ω-accretive and closed set of $X \times X$ satisfying condition* (4.84). *Then, there is a continuous ω-quasi-contractive semigroup $S(t)$ on $\overline{D(A)}$, whose generator A_0 coincides with $-A^0$.*

Proof. For simplicity, we assume that $\omega = 0$. We have already seen that A^0 (the minimal section of A) is single-valued, everywhere defined on $D(A)$, and $-A_0 x = A^0 x$, $\forall x \in D(A)$. Here, A_0 is the infinitesimal generator of the semigroup $S_A(t)$ defined on $\overline{D(A)}$ by the exponential formula (4.17). We prove that $D(A_0) = D(A)$. Let $x \in D(A_0)$. Then

$$\limsup_{h \downarrow 0} \frac{\|S_A(t+h)x - S_A(t)x\|}{h} < \infty, \qquad \forall t \geq 0,$$

and, by the semigroup property (i), it follows that $t \to S_A(t)x$ is Lipschitz continuous on every compact interval $[0,T]$. Hence, $t \to S_A(t)x$ is a.e. differentiable on $(0,\infty)$ and

$$\frac{d}{dt}\, S_A(t)x = A_0 S_A(t)x, \quad \text{a.e. } t > 0.$$

Now, because $y(t) = S_A(t)x$ is a mild solution to (4.16), that is, a.e. differentiable and $(d/dt)y(0) = A_0 x$, it follows by Theorem 4.5 that $S_A(t)x$ is a strong solution to (4.16):

$$\frac{d}{dt}\, S_A(t)x + A^0 S_A(t)x = 0, \quad \text{a.e. } t > 0.$$

Now,

$$-A_0 x = \lim_{h \downarrow 0} \frac{1}{h} \int_0^h A^0 S_A(t)x\,dt,$$

and this implies as in the proof of Theorem 4.6 that $x \in D(A)$ and $-A_0 x \in Ax$ (as seen in the proof of Theorem 4.7, we may assume that A is demiclosed). This completes the proof. □

If X is a Hilbert space, it has been proven by Y. Komura [38] that every continuous semigroup of contractions $S(t)$ on a closed convex set $C \subset X$ is generated by an m-accretive set A; that is, there is an m-accretive set $A \subset X \times X$ such that $-A^0$ is an infinitesimal generator of $S(t)$. Moreover, the domain of the infinitesimal generator of a semigroup of contractions on a closed convex subset $C \subset X$ is dense in C. These remarkable results resemble the classical properties of semigroups of linear contractions in Banach spaces.

Remark 4.3. There is a simple way due to Dafermos and Slemrod [27] to transform the nonhomogeneous Cauchy problem (4.1) into a homogeneous problem. Let us assume that $f \in L^1(0,\infty;X)$ and denote by Y the product space $Y = X \times L^1(0,\infty;X)$ endowed with the norm

$$\|\{x,f\}\|_Y = \|x\| + \int_0^\infty \|f(t)\|dt, \;\; (x,f) \in Y.$$

Let $\mathscr{A} : Y \to Y$ be the (multivalued) operator

$$\mathscr{A}(x,f) = \{Ax - f(0), -f'\}, \qquad (x,f) \in D(\mathscr{A}),$$
$$D(\mathscr{A}) = D(A) \times W^{1,1}([0,\infty);X),$$

where $f' = df/dt$.

It is readily seen that if y is a solution to problem (4.1), then $Y(t) = \{y(t), f_t(s)\}$, where $f_t(s) = f(t+s)$ is the solution to the homogeneous Cauchy problem

$$\frac{d}{dt}Y(t) + \mathscr{A}Y(t) \ni 0, \qquad t \geq 0,$$
$$Y(0) = \{y_0, f\}.$$

On the other hand, if A is ω-m-accretive in $X \times X$, so is $\mathscr{A}$ in $Y \times Y$.

This result is, in particular, useful because it can lead (see Theorem 4.3) to an exponential representation formula for solutions to the nonautonomous equation (4.1) but we omit the details.

Remark 4.4. If A is m-accretive, $f \equiv 0$, and y_e is a stationary (equilibrium) solution to (4.1) (i.e., $0 \in Ay_e$), then we see by estimate (4.14) that the solution $y = y(t)$ to (4.1) is bounded on $[0,\infty)$. More precisely, we have

$$\|y(t) - y_e\| \leq \|y(0) - y_e\|, \qquad \forall t \geq 0.$$

Moreover, if A is strongly accretive (i.e., $A - \gamma I$ is accretive for some $\gamma > 0$), then

$$\|y(t) - y_e\| \leq e^{-\gamma t}\|y(0) - y_0\|, \qquad \forall t \geq 0,$$

which amounts to saying that the trajectory $\{y(t),\ t \geq 0\}$ approaches as $t \to \infty$ the equilibrium solution y_e of the system. This means that the dynamic system associated with (4.1) is *dissipative* and, in this sense, sometimes we refer to equations of the form (4.1) as *dissipative systems*.

Nonlinear Evolution Associated with Subgradient Operators

Here, we study problem (4.1) in the case where A is the subdifferential $\partial\varphi$ of a lower semicontinuous convex function φ from a Hilbert space H to $\overline{\mathbf{R}} = (-\infty, +\infty]$. In other words, consider the problem

$$\begin{cases} \dfrac{dy}{dt}(t) + \partial\varphi(y(t)) \ni f(t), \quad \text{in } (0,T), \\ y(0) = y_0, \end{cases} \tag{4.85}$$

in a real Hilbert space H with the scalar product $(\cdot,\cdot)$ and norm $|\cdot|$. It turns out that the nonlinear evolution generated by $A = \partial\varphi$ on $\overline{D(A)}$ has regularity properties that in the linear case are characteristic of analytic semigroups.

If $\varphi : H \to \overline{\mathbf{R}}$ is a lower semicontinuous, convex function, then its subdifferential $A = \partial\varphi$ is maximal monotone (equivalently, m-accretive) in $H \times H$ and $\overline{D(A)} = \overline{D(\varphi)}$ (see Theorem 2.8 and Proposition 2.3). Then, by Theorem 4.2, for every $y_0 \in \overline{D(A)}$ and $f \in L^1(0,T;H)$ the Cauchy problem (4.85) has a unique mild solution $y \in C([0,T];H)$, which is a strong solution if $y_0 \in D(A)$ and $f \in W^{1,1}([0,T];H)$ (Theorem 4.4).

Theorem 4.11 below amounts to saying that y remains a strong solution to (4.85) on every interval $[\delta,T]$ even if $y_0 \notin D(A)$ and f is not absolutely continuous. In other words, the evolution generated by $\partial\varphi$ has a smoothing effect on initial data and on the right-hand side f of (4.85). (Everywhere in the following, H is identified with its own dual.)

Theorem 4.11. *Let $f \in L^2(0,T;H)$ and $y_0 \in \overline{D(A)}$. Then the mild solution y to problem* (4.1) *belongs to $W^{1,2}([\delta,T];H)$ for every $0 < \delta < T$, and*

$$y(t) \in D(A), \qquad a.e.\ t \in (0,T), \tag{4.86}$$

$$t^{1/2}\frac{dy}{dt} \in L^2(0,T;H) \qquad \varphi(u) \in L^1(0,T), \tag{4.87}$$

$$\frac{dy}{dt}(t) + \partial\varphi(y(t)) \ni f(t), \qquad a.e.\ t \in (0,T). \tag{4.88}$$

Moreover, if $y_0 \in D(\varphi)$, then

$$\frac{dy}{dt} \in L^2(0,T;H), \qquad \varphi(y) \in W^{1,1}([0,T]). \tag{4.89}$$

The main ingredient of the proof is the following chain rule differentiation lemma.

Lemma 4.4. *Let $u \in W^{1,2}([0,T];H)$ and $g \in L^2(0,T;H)$ be such that $g(t) \in \partial\varphi(u(t))$, a.e., $t \in (0,T)$. Then, the function $t \to \varphi(u(t))$ is absolutely continuous on $[0,T]$ and*

$$\frac{d}{dt}\varphi(u(t)) = \left(g(t), \frac{du}{dt}(t)\right), \quad a.e.\ t \in (0,T). \tag{4.90}$$

Proof. Let φ_λ be the regularization of φ; that is,

$$\varphi_\lambda(u) = \inf\left\{\frac{|u-v|^2}{2\lambda} + \varphi(v);\ v \in H\right\}, \qquad u \in H,\ \lambda > 0.$$

We recall (see Theorem 2.9) that φ_λ is Fréchet differentiable on H and

$$\nabla\varphi_\lambda = (\partial\varphi)_\lambda = \lambda^{-1}(I - (I + \lambda\partial\varphi)^{-1}), \qquad \lambda > 0.$$

Obviously, the function $t \to \varphi_\lambda(u(t))$ is absolutely continuous (in fact, it belongs to $W^{1,2}([0,T];H)$) and

$$\frac{d}{dt}\,\varphi_\lambda(u(t)) = \left((\partial\varphi)_\lambda(u(t)), \frac{du}{dt}(t)\right), \quad \text{a.e. } t \in (0,T).$$

Hence,

$$\varphi_\lambda(u(t)) - \varphi_\lambda(u(s)) = \int_s^t \left((\partial\varphi)_\lambda(u(\tau)), \frac{du}{dt}(\tau)\right) d\tau, \qquad \forall s < t,$$

and, letting λ tend to zero, we obtain that

$$\varphi(u(t)) - \varphi(u(s)) = \int_s^t \left((\partial\varphi)^0(u(\tau)), \frac{du}{d\tau}(\tau)\right) d\tau, \qquad 0 \le s < t.$$

By the Lebesgue dominated convergence theorem, the function $t \to (\partial\varphi)^0(u(t))$ is in $L^2(0,T;H)$ and so $t \to \varphi(u(t))$ is absolutely continuous on $[0,T]$. ($(\partial\varphi)^0 = A^0$ is the minimal section of A.) Let t_0 be such that $\varphi(u(t))$ is differentiable at $t = t_0$. We have

$$\varphi(u(t_0)) \le \varphi(v) + (g(t_0), u(t_0) - v), \qquad \forall v \in H.$$

This yields, for $v = u(t_0 - \varepsilon)$,

$$\frac{d}{dt}\,\varphi(u(t_0)) \le \left(g(t_0), \frac{du}{dt}(t_0)\right).$$

Now, by taking $v = u(t_0 + \varepsilon)$ we get the opposite inequality, and so (4.90) follows. □

Proof of Theorem 4.11. Let x_0 be an element of $D(\partial\varphi)$ and $y_0 \in \partial\varphi(x_0)$. If we replace the function φ by $\widetilde{\varphi}(y) = \varphi(y) - \varphi(x_0) - (y_0, u - x_0)$, equation (4.85) reads

$$\frac{dy}{dt}(t) + \partial\widetilde{\varphi}(y(t)) \ni f(t) - y_0.$$

Hence, without any loss of generality, we may assume that

$$\min\{\varphi(u);\ u \in H\} = \varphi(x_0) = 0.$$

Let us assume first that $y_0 \in D(\partial\varphi)$ and $f \in W^{1,2}([0,T];H)$; that is, $df/dt \in L^2(0,T;H)$. Then, by Theorem 4.2, the Cauchy problem in (4.85) has a unique strong solution $y \in W^{1,\infty}([0,T];H)$. The idea of the proof is to obtain a priori estimates in $W^{1,2}([\delta,T];H)$ for y, and after this to pass to the limit together with the initial values and forcing term f.

To this end, we multiply equation (4.85) by $t(dy/dt)$. By Lemma 4.4, we have

$$t\left|\frac{dy}{dt}(t)\right|^2 + t\,\frac{d}{dt}\,\varphi(y(t)) = t\left(f(t), \frac{dy}{dt}(t)\right), \quad \text{a.e. } t \in (0,T).$$

Hence,

$$\int_0^T t\left|\frac{dy}{dt}(t)\right|^2 dt + T\varphi(y(T)) = \int_0^T t\left(f(t),\frac{dy}{dt}(t)\right)dt + \int_0^T \varphi(y(t))dt$$

and, therefore,

$$\int_0^T t\left|\frac{dy}{dt}(t)\right|^2 dt \le \int_0^T t|f(t)|^2 dt + 2\int_0^T \varphi(y(t))dt \tag{4.91}$$

because $\varphi \ge 0$ in H.

Next, we use the obvious inequality

$$\varphi(y(t)) \le (w(t), y(t)-x_0), \qquad \forall w(t) \in \partial\varphi(y(t))$$

to get

$$\varphi(y(t)) \le \left(f(t) - \frac{dy}{dt}(t), y(t)-x_0\right), \quad \text{a.e. } t \in (0,T),$$

which yields

$$\int_0^T \varphi(y(t))dt \le \frac{1}{2}|y(0)-x_0|^2 + \int_0^T |f(t)|\,|y(t)-x_0|dt.$$

Now, multiplying equation (4.85) by $y(t)-x_0$ and integrating on $[0,t]$, yields

$$|y(t)-x_0| \le |y(0)-x_0| + \int_0^t |f(s)|ds, \qquad \forall t \in [0,T].$$

Hence,

$$2\int_0^T \varphi(y(t))dt \le \left(|y(0)-x_0| + \int_0^T |f(t)|dt\right)^2. \tag{4.92}$$

Now, combining estimates (4.91) and (4.92), we get

$$\int_0^T t\left|\frac{dy}{dt}(t)\right|^2 dt \le \int_0^T t|f(t)|^2 dt + 2\left(|y_0-x_0| + \int_0^T |f(t)|dt\right)^2. \tag{4.93}$$

Multiplying equation (4.85) by dy/dt, we get

$$\left|\frac{dy}{dt}(t)\right|^2 + \frac{d}{dt}\varphi(y(t)) = \left(f(t),\frac{dy}{dt}(t)\right), \quad \text{a.e. } t \in (0,T).$$

Hence,

$$\frac{1}{2}\int_0^t \left|\frac{dy}{dt}(s)\right|^2 ds + \varphi(y(t)) \le \frac{1}{2}\int_0^t |f(s)|^2 ds + \varphi(y_0). \tag{4.94}$$

Now, let us assume that $y_0 \in \overline{D(\partial\varphi)}$ and $f \in L^2(0,T;H)$. Then, there exist subsequences $\{y_0^n\} \subset D(\partial\varphi)$ and $\{f_n\} \subset W^{1,2}([0,T];H)$ such that $y_0^n \to y_0$ in H and $f_n \to f$ in $L^2(0,T;H)$ as $n \to \infty$. Denote by $y_n \in W^{1,\infty}([0,T];H)$ the corresponding solutions to (4.86). Because $\partial\varphi$ is monotone, we have (see Proposition 4.1)

$$|y_n(t) - y_m(t)| \le |y_0^n - y_0^m| + \int_0^t |f_n(s) - f_m(s)|ds.$$

Hence, $y_n \to y$ in $C([0,T];H)$. On the other hand, this clearly implies that

$$\frac{dy_n}{dt} \to \frac{dy}{dt} \quad \text{in } \mathscr{D}'(0,T;H),$$

(i.e., in the sense of vectorial H-valued distributions on $(0,t)$), and, by estimate (4.93), it follows that $t^{1/2}(dy/dt) \in L^2(0,T;H)$. Hence, y is absolutely continuous on every interval $[\delta,T]$ and $y \in W^{1,2}([\delta,T];H)$ for all $0 < \delta < T$.

Moreover, by estimate (4.92), written for $y = y_n$, we deduce by virtue of Fatou's lemma that $\varphi(y) \in L^1(0,T)$ and

$$\int_0^T \varphi(y(t))dt \le \liminf_{n\to\infty} \int_0^T \varphi(y_n(t))dt \le \left(|y_0 - x| + \int_0^T |f(t)|dt\right)^2.$$

We may infer, therefore, that y satisfies estimates (4.92) and (4.93). Moreover, y satisfies equation (4.85). Indeed, we have

$$\frac{1}{2}|y_n(t) - x|^2 \le \frac{1}{2}|y_n(s) - x|^2 + \int_s^t (f_n(\tau) - w, y_n(\tau) - x)d\tau$$

for all $0 \le x < t \le T$ and all $[x,w] \in \partial\varphi$. This yields for all $0 \le s < t \le T$ and all $[x,w] \in \partial\varphi$,

$$\frac{1}{2}(|y(t) - x|^2 - |y(s) - x|^2) \le \int_s^t (f(\tau) - w, y(\tau) - x)d\tau$$

and, therefore,

$$\left(\frac{y(t) - y(s)}{t - s}, y(s) - x\right) \le \frac{1}{t-s}\int_s^t (f(\tau) - w, y(\tau) - x)d\tau.$$

Letting $s \to t$, we get, a.e. $t \in (0,T)$,

$$\left(\frac{dy}{dt}(t), y(t) - x\right) \le (f(t) - w, y(t) - x)$$

for all $[x,w] \in A$, and because $A = \partial\varphi$ is maximal monotone, this implies that $y(t) \in D(A)$ and $(d/dt)y(t) \in f(t) - Ay(t)$, a.e. $t \in (0,T)$, as desired.

Assume now that $y_0 \in D(\varphi)$. We choose in this case $y_0^n = (I + n^{-1}\partial\varphi)^{-1}y_0 \in D(\partial\varphi)$ and note that $y_0^n \to y_0$ as $n \to \infty$, and

$$\varphi(y_0^n) \le \varphi(y_0) + (\partial\varphi_n(y_0), (I + n^{-1}\partial\varphi)^{-1}y_0 - y_0) \le \varphi(y_0), \qquad \forall n \in \mathbf{N}^*.$$

Then, by estimate (4.94), we have

$$\frac{1}{2}\int_0^t \left|\frac{dy_n}{ds}(s)\right|^2 ds + \varphi(y_n(t)) \le \frac{1}{2}\int_0^t \left|\frac{df_n}{ds}(s)\right|^2 ds + \varphi(y_0)$$

and, letting $n \to \infty$, we find the estimate

$$\frac{1}{2}\int_0^t \left|\frac{dy}{dt}(s)\right|^2 ds + \varphi(y(t)) \le \frac{1}{2}\int_0^t \left|\frac{df}{ds}(s)\right|^2 ds + \varphi(y_0), \qquad t \in [0,T], \quad (4.95)$$

because $\{dy_n/dt\}$ is weakly convergent to dy/dt in $L^2(0,T;H)$ and φ is lower semicontinuous in H. This completes the proof of Theorem 4.11.

In the sequel, we denote by $W^{1,p}((0,T];H)$, $1 \le p \le \infty$, the space of all $y \in L^p(0,T;H)$ such that $dy/dt \in L^p(\delta,T;H)$ for every $\delta \in (0,T)$.

Theorem 4.12. *Assume that* $y_0 \in \overline{D(A)}$ *and* $f \in W^{1,1}([0,T];H)$. *Then, the solution* y *to problem* (4.85) *satisfies*

$$t\,\frac{dy}{dt} \in L^\infty(0,\infty;H),\; y(t) \in D(A), \qquad \forall t \in (0,T], \quad (4.96)$$

$$\frac{d^+}{dt}\,y(t) + (Ay(t) - f(t))^0 = 0, \qquad \forall t \in (0,T]. \quad (4.97)$$

Proof. By equation (4.85), we have

$$\frac{d}{dt}\,|y(t+h) - y(t)| \le |f(t+h) - f(t)|, \quad \text{a.e. } t, t+h \in (0,T).$$

Hence,

$$\left|\frac{dy}{dt}(t)\right| \le \left|\frac{dy}{ds}(s)\right| + \int_s^t \left|\frac{df}{dt}(\tau)\right| d\tau, \quad \text{a.e. } 0 < s < t < T. \quad (4.98)$$

This yields

$$\frac{1}{2}\,s\left|\frac{dy}{dt}(t)\right|^2 \le s\left|\frac{dy}{ds}(s)\right|^2 + s\left(\int_s^t \left|\frac{df}{d\tau}(\tau)\right| d\tau\right)^2, \quad \text{a.e. } 0 < s < t < T.$$

Then, integrating from 0 to t and using estimate (4.93), we get

$$t\left|\frac{dy}{dt}(t)\right| \le \left(\int_0^t s|f(s)|^2 ds + 2\left(|y(0) - x_0| + \int_0^t |f(s)|ds\right)^2 + \frac{t^2}{2}\left(\int_0^t \left|\frac{df}{d\tau}(\tau)\right| d\tau\right)^2\right)^{1/2}, \quad \text{a.e. } t \in (0,T). \quad (4.99)$$

In particular, it follows by (4.99) that

$$\limsup_{\substack{h\to 0\\ h>0}} \left|\frac{y(t+h) - y(t)}{h}\right| < \infty, \qquad \forall t \in [0,T].$$

Hence, the weak closure E of

$$\left\{\frac{(y(t+h)-y(t))}{h}\right\} \qquad \text{for } h\to 0$$

is nonempty for every $t\in[0,T)$. Let η be an element of E. We have proved earlier the inequality

$$\left(\frac{y(t+h)-y(t)}{h}, y(t)-x\right) \le \frac{1}{h}\int_t^{t+h}(f(\tau)-w, y(\tau)-x)d\tau$$

for all $[x,w]\in\partial\varphi$ and $t, t+h\in(0,T)$. This yields

$$(\eta, y(t)-x) \le (f(t)-w, y(t)-x), \qquad \forall t\in(0,T),$$

and, because $[x,w]$ is arbitrary in $\partial\varphi$, we conclude, by maximal monotonicity of A, that $y(t)\in D(A)$ and $f(t)-\eta\in Ay(t)$. Hence, $y(t)\in D(A)$ for every $t\in(0,T)$. Then, by Theorem 4.6, it follows that

$$\frac{d^+}{dt}y(t) + (Ay(t)-f(t))^0 = 0, \qquad \forall t\in(0,T), \tag{4.100}$$

because, for every $\varepsilon>0$ sufficiently small, $y(\varepsilon)\in D(A)$ and so (4.100) holds for all $t>\varepsilon$. □

In particular, it follows by Theorem 4.12 that the semigroup $S(t)=e^{-At}$ generated by $A=\partial\varphi$ on $\overline{D(A)}$ maps $\overline{D(A)}$ into $D(A)$ for all $t>0$ and

$$t\left|\frac{d^+}{dt}S(t)y_0\right| \le C, \qquad \forall t>0.$$

More precisely, we have the following.

Corollary 4.4. *Let $S(t)=e^{-At}$ be the continuous semigroup of contractions generated by $A=\partial\varphi$ on $\overline{D(A)}$. Then, $S(t)\overline{D(A)}\subset D(A)$ for all $t>0$, and*

$$\left|\frac{d^+}{dt}S(t)y_0\right| = |A^0S(t)y_0| \le |A^0x| + \frac{1}{t}|x-y_0|, \qquad \forall t>0, \tag{4.101}$$

for all $y_0\in\overline{D(A)}$ and $x\in D(A)$.

Proof. Multiplying equation (4.85) (where $f\equiv 0$) by $t(dy/dt)$ and integrating on $(0,t)$, we get

$$\int_0^t s\left|\frac{dy}{ds}(s)\right|^2 ds + t\varphi(y(t)) \le \int_0^T \varphi(y(s))ds, \qquad \forall t>0.$$

Next, we multiply the same equation by $y(t)-x$ and integrate on $(0,t)$. We get

$$\frac{1}{2}|y(t)-x|^2+\int_0^t \varphi(y(s))ds \le \frac{1}{2}|y(0)-x|^2+t\varphi(x).$$

Combining these two inequalities, we obtain

$$\begin{aligned}\int_0^t s\left|\frac{dy}{ds}(s)\right|^2 &\le \frac{1}{2}(|y(0)-x|^2-|y(t)-x|^2+t(\varphi(x)-\varphi(y(t)))\\ &\le \frac{1}{2}(|y(0)-x|^2-|y(t)-x|^2+t(A^0x,x-y(t))\\ &\le \frac{1}{2}|y(0)-x|^2+\frac{t^2|A^0x|^2}{2}, \qquad \forall t>0.\end{aligned}$$

Because, by formula (4.98) the function $t \to |(d/dt)y(t)|$ (and consequently $t \to |(d^+/dt)y(t)|$) is monotonically decreasing, this implies (4.101). □

Remark 4.5. Theorems 4.11 and 4.12 clearly remain true for equations of the form

$$\begin{cases}\dfrac{dy}{dt}(t)+\partial\varphi(y(t))-\omega y(t) \ni f(t), \ \text{a.e. in } (0,T),\\ y(0)=y_0,\end{cases}$$

where $\omega \in \mathbf{R}$ and also for Lipschitzian perturbations of $\partial\varphi$. The proof is exactly the same and so it is omitted.

A nice feature of nonlinear semigroups generated by subdifferential operators in Hilbert space is their longtime behavior. Namely, one has the following result due to Bruck [18].

Theorem 4.13. *Let $A=\partial\varphi$, where $\varphi : H \to (-\infty,+\infty]$ is a convex l.s.c. function such that $(\partial\varphi)^{-1}(0) \ne \emptyset$. Then, for each $y_0 \in \overline{D(A)}$ there is $\xi \in (\partial\varphi)^{-1}(0)$ such that*

$$\xi = w\text{-}\lim_{t\to\infty} e^{-At}y_0. \tag{4.102}$$

Proof. If we multiply the equation

$$\frac{d}{dt}y(t)+Ay(t) \ni 0, \quad \text{a.e. } t>0,$$

by $y(t)-y_0$, where $x \in (\partial\varphi)^{-1}(0)$, we obtain that

$$\frac{1}{2}\frac{d}{dt}|y(t)-x|^2 \le 0, \quad \text{a.e. } t>0,$$

because $A=\partial\varphi$ and, therefore, $(Ay(t),y(t)-x) \ge 0$, $\forall t \ge 0$. This implies that $\{y(t)\}_{t\ge 0}$ is bounded and we denote by K the so-called weak ω-limit set associated with the trajectory $\{y(t)\}_{t\ge 0}$; that is,

$$K=\left\{w\text{-}\lim_{t_n\to\infty} y(t_n)\right\}.$$

Let us notice that $K \subset (\partial\varphi)^{-1}(0)$. Indeed, if $y(t_n) \rightharpoonup \xi$, for some $\{t_n\} \to \infty$, then we see by (4.101) that

$$\lim_{n\to\infty} \frac{dy}{dt}(t_n) = 0$$

and because A is demiclosed, this implies that $0 \in A\xi$ (i.e., $\xi \in A^{-1}(0) = (\partial\varphi)^{-1}(0)$). On the other hand, $t \to |y(t) - x|^2$ is decreasing for each $x \in (\partial\varphi)^{-1}(0)$ and, in particular, for each $x \in K$.

Let ξ_1, ξ_2 be two arbitrary elements of K given by

$$\xi_1 = w\text{-}\lim_{n'\to\infty} y(t_{n'}), \qquad \xi_2 = w\text{-}\lim_{n''\to\infty} y(t_{n''}),$$

where $t_{n'} \to \infty$ and $t_{n''} \to \infty$ as $n' \to \infty$ and $n'' \to \infty$, respectively.

Because $\lim_{t\to\infty} |y(t) - x|^2$ exists for each $x \in K \subset (\partial\varphi)^{-1}(0)$, we have

$$\lim_{n'\to\infty} |y(t_{n'}) - \xi_1|^2 = \lim_{n''\to\infty} |y(t_{n''}) - \xi_1|^2,$$
$$\lim_{n''\to\infty} |y(t_{n''}) - \xi_2|^2 = \lim_{n'\to\infty} |y(t_{n'}) - \xi_2|^2.$$

The latter implies by an elementary calculation that $|\xi_1 - \xi_2|^2 = 0$. Hence, K consists of a single point and this completes the proof of (4.102). □

Remark 4.6. In particular, it follows by Theorem 4.13 that, for each $y_0 \in \overline{D(A)}$, the solution $y(t) = e^{-At}y_0$, $A = \partial\varphi$ is weakly convergent to an equilibrium point $\xi \in \arg\min_{u\in H} \varphi(u)$ of system (4.14). There is a discrete version which asserts that the sequence $\{y_n\}$ defined by

$$y_{n+1} = y_n - h\partial\varphi(y_{n+1}), \qquad n = 0, 1, \dots, \ h > 0,$$

is weakly convergent in H to an element $\xi \in (\partial\varphi)^{-1}(0)$; that is, to a minimum point for φ on H. The proof is completely similar. This discrete version of Theorem 4.13, known in convex optimization as the steepest descent algorithm is at the origin of a large category of gradient type algorithms.

Remark 4.7. If, under assumptions of Theorem 4.13, the trajectory $\{y(t)\}_{t\geq 0}$ is relatively compact in H (this happens for instance if each level set $\{x;\ \varphi(x) \leq \lambda\}$ is compact), then (4.102) is strengthening to

$$y(t) = e^{-At}y_0 \to \xi \quad \text{strongly in } H \text{ as } t \to \infty.$$

The longtime behavior of trajectories $\{y(t);\ t > 0\}$ to nonlinear equation (4.1) and their convergence for $t \to \infty$ to an equilibrium solution $\xi \in A^{-1}(0)$ is an important problem largely studied in the literature by different methods including dynamic topology (the Lasalle principle) or by accretivity arguments of the type presented above. Without entering into details we refer to the works of Dafermos and Slemrod [27], Haraux [31] and also to the book of Moroşanu [42].

The Reflection Problem on Closed Convex Sets

Let A be a self-adjoint positive operator in Hilbert space H and let K be a closed convex subset of H. Then, the function $\varphi : H \to \overline{\mathbf{R}}$ defined by

$$\varphi(u) = \begin{cases} \frac{1}{2}(Au,u) + I_K(u), & \forall u \in K \cap D(A^{1/2}), \\ +\infty, & \text{otherwise} \end{cases}$$

(I_K indicator function of K) is convex and l.s.c. Moreover, if there is $h \in H$ such that

$$(I+\lambda A)^{-1}(x+\lambda h) \in K, \qquad \forall \lambda > 0,\ x \in K,$$

then $A + \partial I_K$ is maximal monotone (see Theorem 2.11) and so $\partial\varphi = A + \partial I_K$ with $D(\partial\varphi) = D(A) \cap K$.

For this special form of φ, equation (4.85) reduces to the variational inequality

$$\begin{cases} \left(\frac{dy}{dt}(t) + Ay(t) - f(t), y(t) - z\right) \le 0, & \forall z \in K,\ t \in (0,T), \\ y(0) = y_0, \qquad y(t) \in K, & \forall t \in [0,T], \end{cases} \tag{4.103}$$

which is similar to that considered in Section 2.3.

A more general situation is discussed in Section 5.2 below. Here, we confine ourselves to noting that the solution $y \in W^{1,2}([0,T];H)$ to (4.103), which exists and is unique for $y_0 \in K$ and $f \in L^2(0,T;H)$, satisfies the system

$$\begin{cases} \frac{dy}{dt}(t) + Ay(t) = f(t) & \text{if } y(t) \in \overset{\circ}{K}, \\ \frac{dy}{dt}(t) + Ay(t) = -\eta_K(t) + f(t) & \text{if } y(t) \in \partial K, \end{cases}$$

where $\eta_K(t) \in N_K(y(t))$, the normal cone to K on the boundary ∂K. (Here, $\overset{\circ}{K}$ is the interior of K if nonempty.) For instance, if $K = \{u \in H;\ |u| \le \rho\}$, then we have

$$\begin{cases} \frac{dy}{dt}(t) + Ay(t) = f(t) & \text{on } \{t;\ |y(t)| < \rho\}, \\ \frac{dy}{dt}(t) + Ay(t) = -\lambda y(t) + f(t) & \text{on } \{t;\ |y(t)| = \rho\}, \end{cases}$$

for some $\lambda \ge 0$. The parameter λ must be viewed as a Lagrange multiplier that arises from constraint $y(t) \in K$, $\forall t \ge 0$.

For this reason, problem (4.103) is also called the *reflection problem on K* associated with linear equation $dy/dt + Ay = 0$ and under this interpretation it is relevant not only in the dynamic theory of free boundary problems, but also in the theory of stochastic processes with optimal stopping time arising in the theory of financial markets (see, e.g., Barbu and Marinelli [8]).

The Brezis–Ekeland Variational Principle

It turns out that the Cauchy problem (4.85) can be equivalently represented as a minimization problem in the space $L^2(0,T;H)$ or $W^{1,2}([0,T];H)$ which is quite surprising because, in general, the Cauchy problem is not of variational type.

In fact, if $\varphi : H \to \overline{\mathbf{R}}$ is convex, l.s.c., and φ^* is its conjugate function we have by Proposition 1.5 that

$$\varphi(y) + \varphi^*(p) \geq (y,p), \qquad \forall y,p \in H,$$

with equality if and only if $p \in \partial\varphi(y)$. Then, we may equivalently write (4.85) as

$$\frac{dy}{dt}(t) + z(t) = f(t), \qquad \varphi(y(t)) + \varphi^*(z(t)) = (y(t),z(t)), \quad \text{a.e. } t \in (0,T),$$
$$y(0) = y_0.$$

Hence, if $y \in W^{1,2}([0,T];H)$ is the solution to (4.85), where $y_0 \in D(\varphi)$ (see Theorem 4.11), then we have

$$\varphi(y(t)) + \varphi^*\left(f(t) - \frac{dy}{dt}(t)\right) = \left(y(t), f(t) - \frac{dy}{dt}(t)\right), \quad \text{a.e. } t \in (0,T),$$

and the latter is equivalent to (4.85). This yields

$$\int_0^T \left(\varphi(y(t)) + \varphi^*\left(f(t) - \frac{dy}{dt}(t)\right) - (y(t),f(t))\right) dt + \frac{1}{2}|y(T)|^2 - \frac{1}{2}|y_0|^2 = 0$$

and we have also that

$$\begin{aligned} y = \arg\min \Bigg\{ &\int_0^T \left[\varphi(\theta(t)) + \varphi^*\left(f(t) - \frac{d\theta}{dt}(t)\right) - (\theta(t),f(t))\right] dt \\ &+ \frac{1}{2}|\theta(T)|^2 - \frac{1}{2}|y_0|^2; \quad \theta \in W^{1,2}([0,T];H),\ \theta(0) = y_0 \Bigg\}. \end{aligned} \tag{4.104}$$

This means that the Cauchy problem (4.85) is equivalent to the minimization problem (4.104). This is the *Brezis–Ekeland principle* and it reveals an interesting connection between the subpotential Cauchy problem and convex optimization, which found many interesting applications in the theory of variational inequalities (see, e.g., Stefanelli [51], and Visintin [53]).

However, the function $\Phi : W^{1,2}([0,T];H) \to \overline{\mathbf{R}}$, defined by the right-hand side of (4.104), is convex and lower semicontinuous but, in general, not coercive (this happens if $D(\varphi) = H$ only) and so, one cannot derive Theorem 4.11 directly from the existence of a minimizer y in problem (4.104).

4.2 Approximation and Structural Stability of Nonlinear Evolutions

The Trotter–Kato Theorem for Nonlinear Evolutions

One might expect the solution to Cauchy problem (4.1) to be continuous with respect to the operator A, that is, with respect to small structural variations of the problem. We show below that this indeed happens in a certain precise sense and for a certain notion of convergence defined in the space of quasi-m-accretive operators.

Consider in a general Banach space X a sequence A_n of subsets of $X \times X$. The subset of $X \times X$, $\liminf A_n$ is defined as the set of all $[x,y] \in X \times X$ such that there are sequences $x_n, y_n, y_n \in A_n x_n, x_n \to x$ and $y_n \to y$ as $n \to \infty$. If A_n are quasi-m-accretive, there is a simple resolvent characterization of $\liminf A_n$. (See Attouch [1, 2].)

Proposition 4.4. *Let $A_n + \omega I$ be m-accretive for $n = 1,2\ldots$. Then $A \subset \liminf A_n$ if and only if*

$$\lim_{n\to\infty} (I+\lambda A_n)^{-1}x = (I+\lambda A)^{-1}x, \qquad \forall x \in X, \tag{4.105}$$

for $0 < \lambda < \omega^{-1}$.

Proof. Assume that (4.105) holds and let $[x,y] \in A$ be arbitrary but fixed. Then, we have

$$(I+\lambda A)^{-1}(x+\lambda y) = x, \qquad \forall \lambda \in (0,\omega^{-1})$$

and, by (4.105),

$$(I+\lambda A_n)^{-1}(x+\lambda y) \to (I+\lambda A)^{-1}(x+\lambda y) = x.$$

In other words, $x_n = (I+\lambda A_n)^{-1}(x+\lambda y) \to x$ as $n \to \infty$ and $x_n + \lambda y_n = x + \lambda y$, $y_n \in Ax_n$. Hence, $y_n \to y$ as $n \to \infty$, and so $[x,y] \in \liminf A_n$.

Conversely, let us assume now that $A \subset \liminf A_n$. Let x be arbitrary in X and let $x_0 = (I+\lambda A)^{-1}x$; that is,

$$x_0 + \lambda y_0 = x, \quad \text{where } y_0 \in Ax_0.$$

Then, there are $[x_n, y_n] \in A_n$ such that $x_n \to x_0$ and $y_n \to y_0$ as $n \to \infty$. We have

$$x_n + \lambda y_n = z_n \to x_0 + \lambda y_0 = x \quad \text{as } n \to \infty.$$

Hence,

$$(I+\lambda A_n)^{-1}x \to x_0 = (I+\lambda A)^{-1}y_0 \quad \text{for } 0 < \lambda < \omega^{-1},$$

as claimed. □

In the literature, such a convergence is called *convergence in the sense of graphs*.

Theorem 4.14 below is the nonlinear version of the Trotter–Kato theorem from the theory of C_0-semigroups and, roughly speaking, it amounts to saying that if A_n

is convergent to A in the sense of graphs, then the dynamic (evolution) generated by A_n is uniformly convergent to that generated by A (see Pazy [45]).

Theorem 4.14. *Let A_n be ω-m-accretive in $X \times X$, $f^n \in L^1(0,T;X)$ for $n = 1,2,\ldots$ and let y_n be mild solution to*

$$\frac{dy_n}{dt}(t) + A_n y_n(t) \ni f^n(t) \quad in \ [0,T], \quad y_n(0) = y_0^n. \tag{4.106}$$

Let $A \subset \liminf A_n$ and assume that

$$\lim_{n\to\infty}\left(\int_0^T \|f^n(t) - f(t)\|dt + \|y_0^n - y_0\|\right) = 0. \tag{4.107}$$

Then, $y_n(t) \to y(t)$ uniformly on $[0,T]$, where y is the mild solution to problem (4.106).

Proof. Let $D_{A^n}^{\varepsilon}(0 = t_0, t_1, \ldots, t_N; f_1^n, \ldots, f_N^n)$ be an ε-discretization of problem (4.106) and let $D_A^{\varepsilon}(0 = t_0, t_1, \ldots, t_n; f_1, \ldots, f_N)$ be the corresponding ε-discretization for (4.1). We take $t_i = i\varepsilon$ for all i. Let $y_{\varepsilon,n}$ and y_ε be the corresponding ε-approximate solutions; that is,

$$y_{\varepsilon,n}(t) = y_{\varepsilon,n}^i, \ y_\varepsilon(t) = y_\varepsilon^i \quad \text{for } t \in (t_{i-1}, t_i],$$

where $y_{\varepsilon,n}^0 = y_0^n$, $y_\varepsilon^0 = y_0$, and

$$y_{\varepsilon,n}^i + \varepsilon A_n y_{\varepsilon,n}^i \ni y_{\varepsilon,n}^{i-1} + \varepsilon f_i^n, \qquad i = 1,\ldots,N, \tag{4.108}$$

$$y_\varepsilon^i + \varepsilon A y_\varepsilon^i \ni y_\varepsilon^{i-1} + \varepsilon f_i, \qquad i = 1,\ldots,N. \tag{4.109}$$

By the definition of $\liminf A_n$, for every $\eta > 0$ there is $[\bar{y}_{\varepsilon,n}^i, w_{\varepsilon,n}^i] \in A_n$ such that

$$\|\bar{y}_{\varepsilon,n}^i - y_\varepsilon^i\| + \|w_{\varepsilon,n}^i - w_\varepsilon^i\| \le \eta \quad \text{for } n \ge \delta(\eta,\varepsilon). \tag{4.110}$$

Here, $w_\varepsilon^i = (1/\varepsilon)(y_\varepsilon^{i-1} + \varepsilon f_i - y_\varepsilon^i) \in A y_\varepsilon^i$. Then, using the ω-accretivity of A_n, by (4.108)–(4.110) it follows that

$$\|\bar{y}_{\varepsilon,n}^i - y_{\varepsilon,n}^i\| \le (1-\varepsilon\omega)^{-1}\|\bar{y}_{\varepsilon,n}^{i-1} - y_{\varepsilon,n}^{i-1}\| + \varepsilon(1-\varepsilon\omega)^{-1}\|f_i^n - f_i\| + C\varepsilon\eta, \qquad \forall i,$$

for $n \ge \delta(\eta,\varepsilon)$. This yields

$$\|\bar{y}_{\varepsilon,n}^i - y_{\varepsilon,n}^i\| \le C\eta + C\varepsilon\sum_{k=1}^{i}(1-\varepsilon\omega)^{-k}\|f_k^n - f_k\|, \qquad i = 1,\ldots,N.$$

Hence,

$$\|y_{\varepsilon,n}^i - y_\varepsilon^i\| \le C\eta + C\varepsilon\sum_{k=1}^{i}(1-\varepsilon\omega)^{-k}\|f_k^n - f_k\|, \qquad i = 1,\ldots,N,$$

for $n \ge \delta(\varepsilon,\eta)$.

We have shown, therefore, that, for $n \geq \delta(\varepsilon,\eta)$,

$$\|y_{\varepsilon,n}(t) - y_\varepsilon(t)\| \leq C\left(\eta + \int_0^T \|f^n(t) - f(t)\|dt\right), \qquad \forall t \in [0,T], \tag{4.111}$$

where C is independent of n and ε.

Now, we have

$$\begin{aligned} \|y_n(t) - y(t)\| \leq \|y_n(t) - y_{\varepsilon,n}(t)\| + \|y_{\varepsilon,n}(t) - y_\varepsilon(t)\| + \|y_\varepsilon(t) - y(t)\|, \\ \forall t \in [0,T). \end{aligned} \tag{4.112}$$

Let η be arbitrary but fixed. Then, by Theorem 4.1, we have

$$\|y_\varepsilon(t) - y(t)\| \leq \eta, \qquad \forall t \in [0,T], \quad \text{if } 0 < \varepsilon < \varepsilon_0(\eta).$$

Also, by estimate (4.37) in the proof of Theorem 4.1, we have

$$\|y_{\varepsilon,n}(t) - y_n(t)\| \leq \eta, \qquad \forall t \in [0,T],$$

for all $0 < \varepsilon < \varepsilon_1(\eta)$, where $\varepsilon_1(\eta)$ does not depend on n. Thus, by (4.111) and (4.112), we have

$$\|y_n(t) - y(t)\| \leq C\left(\eta + \int_0^T \|f^n(t) - f(t)\|dt\right), \qquad \forall t \in [0,T]$$

for n sufficiently large and any $\eta > 0$. □

Corollary 4.5. *Let A be ω-m-accretive, $f \in L^1(0,T;X)$, and $y_0 \in \overline{D(A)}$. Let $y_\lambda \in C^1([0,T];X)$ be the solution to the approximating Cauchy problem*

$$\frac{dy}{dt}(t) + A_\lambda y(t) = f(t) \quad \text{in } [0,T], \qquad y(0) = y_0,\ 0 < \lambda < \frac{1}{\omega}, \tag{4.113}$$

where $A_\lambda = \lambda^{-1}(I - (I + \lambda A)^{-1})$. Then, $\lim_{\lambda\to 0} y_\lambda(t) = y(t)$ uniformly in t on $[0,T]$, where y is the mild solution to problem (4.1).

Proof. It is easily seen that $A \subset \liminf_{\lambda\to 0} A_\lambda$. Indeed, for $\alpha \in (0,1/\omega)$ we set

$$x_\lambda = (I + \alpha A_\lambda)^{-1}x,\ \ u = (I + \alpha A)^{-1}x, \qquad \forall \lambda > 0.$$

After some calculation, we see that

$$x_\lambda + \alpha A\left(\left(1 + \frac{\lambda}{\alpha}\right)x_\lambda - \frac{\lambda}{\alpha}x\right) \ni x.$$

Subtracting this equation from $u + \alpha Au \ni x$ and using the ω-accretivity of A, we get

$$\|x_\lambda - u\|^2 \leq \alpha\omega\left\|\left(1 + \frac{\lambda}{\alpha}\right)x_\lambda - \frac{\lambda}{\alpha}x - u\right\|^2 + \frac{\lambda}{\alpha}(x_\lambda - u, x - x_\lambda).$$

Hence, $\lim_{\lambda\to 0} x_\lambda = u = (I+\alpha A)^{-1}x$ for $0<\alpha<1/\lambda$, and so we may apply Theorem 4.14. □

Remark 4.8. If X is a Hilbert space and $S_n(t)$ is the semigroup generated by A_n on X, then, according to a result due to H. Brezis, condition (4.105) is equivalent to the following one. For every $x \in \overline{D(A)}$, $\exists \{x_n\} \subset D(A_n)$ such that $x_n \to x$ and $S_n(t)x_n \to S(t)x$, $\forall t>0$, where $S(t)$ is the semigroup generated by A on $\overline{D(A)}$.

Theorem 4.14 is useful in proving the stability and convergence of a large class of approximation schemes for problem (4.1). For instance, if A is a nonlinear partial differential operator on a certain space of functions defined on a domain $\Omega \subset \mathbf{R}^m$, then very often the A_n arise as finite element approximations of A on a subspace X_n of X. Another important class of convergence results covered by this theorem is the homogenization problem (see, e.g., Attouch [2] and references given there).

Nonlinear Chernoff Theorem and Lie–Trotter Products

We prove here the nonlinear version of the famous Chernoff theorem (see Chernoff [21]), along with some implications for the convergence of the Lie–Trotter product formula for nonlinear semigroups of contractions.

Theorem 4.15. *Let X be a real Banach space, A be an accretive operator satisfying the range condition* (4.15)*, and let $C=\overline{D(A)}$ be convex. For each $t>0$, let $F(t): C\to C$ satisfy:*

(i) $\quad \|F(t)x - F(t)u\| \le \|x-u\|, \qquad \forall x,y \in C \quad \text{and} \quad t\in[0,T].$

(ii) $\quad \displaystyle\lim_{t\downarrow 0}\left(I+\lambda\,\frac{I-F(t)}{t}\right)^{-1} x = (I+\lambda A)^{-1}x, \qquad \forall x\in C,\ \lambda>0.$

Then, for each $x\in C$ and $t>0$,

$$\lim_{n\to 0}\left(F\left(\frac{t}{n}\right)\right)^n x = S_A(t)x, \tag{4.114}$$

uniformly in t on compact intervals.

Here, $S_A(t)$ is the semigroup generated by A on $C=\overline{D(A)}$. (See (4.82).) It should be said that in the special case where $F(t)=(I+tA)^{-1}$, Theorem 4.15 reduces to the exponential formula (4.17) in Theorem 4.3.

The main ingredient of the proof is the following convergence result.

Proposition 4.5. *Let $C\subset X$ be nonempty, closed, and convex, let $F: C\to C$ be a nonexpansive operator, and let $h>0$. Then, the Cauchy problem*

$$\frac{du}{dt} + h^{-1}(I-F)u = 0, \qquad u(0)=x\in C, \tag{4.115}$$

has a unique solution $u\in C^1([0,\infty);X)$, such that $u(t)\in C$, for all $t\ge 0$.

Moreover, the following estimate holds

$$\|F^n x - u(t)\| \le \left(\left(n - \frac{t}{h}\right)^2 + n\right)^{1/2} \|x - Fx\|, \qquad \forall t \ge 0, \tag{4.116}$$

for all $n \in \mathbf{N}$. *In particular, for* $t = nh$ *we have*

$$\|F^n x - u(nh)\| \le n^{1/2} \|x - Fx\|, \qquad n = 1, 2, \ldots, t \ge 0. \tag{4.117}$$

Proof. The initial value problem (4.115) can be written equivalently as

$$u(t) = e^{-(t/h)} x + \int_0^t e^{-((t-s)/h)} F u(s) ds, \qquad \forall t \ge 0,$$

and it has a unique solution $u(t) \in C$, $\forall t \ge 0$, by the Banach fixed point theorem. Making the substitution $t \to t/h$, we can reduce the problem to the case $h = 1$.

Multiplying equation (4.115) by $J(u(t) - x)$, where $J : X \to X^*$ is the duality mapping, we get

$$\frac{d}{dt} \|u(t) - x\| \le \|Fx - x\|, \quad \text{a.e. } t > 0,$$

because $I - F$ is accretive. Hence,

$$\|u(t) - x\| \le t \|Fx - x\|, \qquad \forall t \ge 0. \tag{4.118}$$

On the other hand, we have

$$u(t) - F^n x = e^{-t}(x - F^n x) + \int_0^t e^{s-t}(Fu(s) - F^n x) ds$$

and

$$\|x - F^n x\| \le \sum_{k=1}^{n} \|F^{k-1} x - F^k x\| \le n \|x - Fx\|, \qquad \forall n.$$

Hence,

$$\|u(t) - F^n x\| \le n e^{-t} \|x - Fx\| + \int_0^t e^{s-t} \|u(s) - F^{n-1} x\| ds.$$

We set $\varphi_n(t) = \|u(t) - F^n x\| \, \|x - Fx\|^{-1} e^t$. Then, we have

$$\varphi_n(t) \le n + \int_0^t \varphi_{n-1}(x) ds, \qquad \forall t \ge 0, \ n = 1, 2, \ldots, \tag{4.119}$$

and, by (4.118), we see that

$$\varphi_0(t) \le t e^t, \qquad \forall t \ge 0. \tag{4.120}$$

Solving iteratively (4.119) and (4.120), we get

$$\begin{aligned}\varphi_n(t) &\le \sum_{k=1}^{n} \frac{kt^{n-k}}{(n-k)!} + \frac{1}{(n-1)!}\int_0^t (t-s)^{n-1}\varphi_0(s)ds \\ &= \sum_{k=1}^{n} \frac{kt^{n-k}}{(n-k)!} + \frac{1}{(n-1)!}\int_0^t (t-s)^{n-1}\sum_{j=1}^{\infty}\frac{s^{j+1}}{j!}ds \\ &= \sum_{k=1}^{n} \frac{kt^{n-k}}{(n-k)!} + \sum_{j=0}^{\infty}\frac{1}{(n-1)!j!}\int_0^t (t-s)^{n-1}s^{j+1}ds.\end{aligned}$$

Because

$$\int_0^t (t-s)^{n-1}s^{j+1}ds = \frac{t^{n+j+1}(j+1)!(n-1)!}{(n+j+1)!},$$

we obtain that

$$\begin{aligned}\varphi_n(t) &\le \sum_{k=0}^{n} \frac{(n-k)t^k}{k!} + \sum_{j=0}^{\infty}\frac{(j+1)t^{n+j+1}}{(n+j+1)!} = \sum_{k=0}^{\infty}\frac{(n-k)t^k}{k!} \\ &= \sum_{k=0}^{\infty}\frac{t^k}{k!}|n-k| \le \left(\sum_{k=0}^{\infty}\frac{(n-k)^2t^k}{k!}\right)^{1/2} e^{t/2}.\end{aligned}$$

Hence,

$$\varphi_n(t) \le e^t((n-t)^{-1}+t)^{1/2}, \qquad \forall t \ge 0,$$

as claimed. □

Proof of Theorem 4.15. We set $A_h = h^{-1}(I-F(h))$ and denote by $S_h(t)$ the semigroup generated by A_h on $C = \overline{D(A)}$ (Theorem 4.3). We also use the standard notation

$$J_\lambda = (I+\lambda A)^{-1}, \qquad J_\lambda^h = (I+\lambda A_h)^{-1}.$$

Because $J_\lambda^h x \to J_\lambda x$, $\forall x \in C$, as $h \to 0$, it follows by Theorem 4.14 that, for every $x \in C$,

$$S_h(t)x \to S_A(t)x \quad \text{uniformly in } t \text{ on compact intervals.} \tag{4.121}$$

Next, by Proposition 4.5, we have that

$$\begin{aligned}\|S_h(nh)x - F^n(h)x\| &\le \|S_h(nh)J_\lambda^h x - F^n(h)J_\lambda^h x\| + 2\|x - J_\lambda^h x\| \\ &\le \|x - J_\lambda^h x\|(2+\lambda^{-1}hn^{1/2}).\end{aligned}$$

Now, we fix $x \in D(A)$ and $h = n^{-1}t$. Then, the previous inequality yields

$$\begin{aligned}&\left\|S_{t/n}(t)x - F^n\left(\frac{t}{n}\right)x\right\| \le (2+\lambda^{-1}tn^{-(1/2)})(\|x-J_\lambda x\| + \|J_\lambda^{t/n}x\|) \\ &\le (2+\lambda^{-1}tn^{-(1/2)})(\lambda|Ax| + \|J_\lambda^{t/n}x - J_\lambda x\|), \qquad \forall t>0,\ \lambda>0.\end{aligned}$$

Finally,

$$\left\| S_{t/n}(t)x - F^n\left(\frac{t}{n}\right)x \right\| \le 2\lambda|Ax| + tn^{-(1/2)}|Ax| \\ + (2+\lambda^{-1}tn^{-(1/2)})\|J_\lambda^{t/n}x - J_\lambda x\|, \\ \forall t > 0,\ \lambda > 0. \tag{4.122}$$

Now, fix $\lambda > 0$ such that $2\lambda|Ax| \le \varepsilon/3$. Then, by (ii), we have

$$(2+\lambda^{-1}tn^{-(1/2)})\|J_\lambda^{t/n}x - J_\lambda x\| \le \frac{\varepsilon}{3} \quad \text{for } n > N(\varepsilon),$$

and so, by (4.121) and (4.122), we conclude that, for $n \to \infty$,

$$F^n\left(\frac{t}{n}\right)x \to S_A(t)x \quad \text{uniformly in } t \text{ on every } [0,T]. \tag{4.123}$$

Now, because

$$\|S_A(t)x - S_A(t)y\| \le |x-y|, \qquad \forall t \ge 0,\ x,y \in C,$$

and

$$\left\| F^n\left(\frac{1}{n}\right)x - F^n\left(\frac{t}{n}\right)y \right\| \le \|x-y\|, \qquad \forall t \ge 0,\ x,y \in C,$$

(4.123) extends to all $x \in \overline{D(A)} = C$. The proof of Theorem 4.15 is complete.

Remark 4.9. The conclusion of Theorem 4.15 remains unchanged if A is ω-accretive, satisfies the range condition (4.15), and $F(t) : C \to C$ are Lipschitzian with Lipschitz constant $L(t) = 1 + \omega t + o(t)$ as $t \to 0$. The proof is essentially the same and relies on an appropriate estimate of the form (4.117) for Lipschitz mappings on C.

Given two m-accretive operators $A, B \subset X \times X$ such that $A+B$ is m-accretive, one might expect that

$$S_{A+B}(t)x = \lim_{n\to\infty}\left(S_A\left(\frac{t}{n}\right)S_B\left(\frac{t}{n}\right)\right)^n x, \qquad \forall t \ge 0, \tag{4.124}$$

for all $x \in \overline{D(A)\cap D(B)}$. This is the *Lie–Trotter product* formula and one knows that it is true for C_0-semigroups of contractions and in other situations (see Pazy [45], p. 92). It is readily seen that (4.124) is equivalent to the convergence of the fractional step method scheme for the Cauchy problem

$$\begin{cases} \dfrac{dy}{dt} + Ay + By \ni 0 \quad \text{in } [0,T], \\ y(0) = y_0; \end{cases} \tag{4.125}$$

that is,

$$\begin{cases} \dfrac{dy}{dt} + Ay \ni 0 & \text{in } [i\varepsilon,(i+1)\varepsilon],\ i=0,1,...,N-1,\ T=N\varepsilon, \\ y^+(i\varepsilon) = z(\varepsilon), & i=0,1,...,N-1, \\ y^+(0) = y_0, \end{cases} \tag{4.126}$$

$$\begin{aligned} &\frac{dz}{dt} + Bz \ni 0 \quad \text{in } [0,\varepsilon], \\ &z(0) = y^-(i\varepsilon). \end{aligned} \tag{4.127}$$

In a general Banach space, the Lie–Trotter formula (4.124) is not convergent even for regular operators B unless $S_A(t)$ admits a graph infinitesimal generator A: for all $[x,y] \in A$ there is $x_h \to x$ as $h \to 0$ such that $h^{-1}(x_h - S_A(h)x) \to y$ (Bénilan and Ismail [12]). However, there are known several situations in which formula (4.124) is true and one is described in Theorem 4.16 below.

Theorem 4.16. *Let X and X^* be uniformly convex and let A,B be m-accretive single-valued operators on X such that $A+B$ is m-accretive and $S_A(t),S_B(t)$ map $\overline{D(A)\cap D(B)}$ into itself. Then,*

$$S_{A+B}(t)x = \lim_{n\to\infty}\left(S_A\left(\frac{t}{n}\right)S_B\left(\frac{t}{n}\right)\right)^n x, \qquad \forall x \in \overline{D(A)\cap D(B)}, \tag{4.128}$$

and the limit is uniform in t on compact intervals.

Proof. We verify the hypotheses of Theorem 4.15, where $F(t) = S_A(t)S_B(t)$ and $C = \overline{D(A)\cap D(B)}$. To prove (ii), it suffices to show that

$$\lim_{t\downarrow 0} \frac{x - F(t)x}{x} = Ax + Bx, \qquad \forall x \in D(A)\cap D(B). \tag{4.129}$$

Indeed, if

$$x_t = \left(I + \lambda\, \frac{I - F(t)}{t}\right)^{-1} x$$

and

$$x_0 = (I + \lambda(A+B))^{-1}x,$$

then we have

$$x_t + \frac{\lambda}{t}\,(x_t - F(t)x_t) = x \tag{4.130}$$

and, respectively,

$$x_0 + \lambda Ax_0 + \lambda Bx_0 = x. \tag{4.131}$$

Subtracting (4.130) from (4.131), we may write

$$x_t - x_0 + \frac{\lambda}{t}\,((I - F(t))x_t - (I + F(t)x_0)) + \lambda\left(Ax_0 + Bx_0 - \frac{x_0 - F(t)x_0}{t}\right) = 0.$$

Multiplying this by $J(x_t - x_0)$, where J is the duality mapping of X, and using (4.129) and the accretiveness of $I - F(t)$, it follows that

$$\lim_{t\downarrow 0} \|x_t - x_0\| \le \lambda \lim_{t\downarrow 0} \left\| Ax_0 + Bx_0 - \frac{x_0 - F(t)x_0}{t} \right\| = 0.$$

Hence, $\lim_{t\downarrow 0} x_t = x_0$, which implies (ii).

To prove (4.129), we write $t^{-1}(x - F(t)x)$ as

$$t^{-1}(x - F(t)x) = t^{-1}(x - S_A(t)x) + t^{-1}(S_A(t)x - S_A(t)S_B(t)x).$$

Because $t^{-1}(x - S_A(t)x) \to Ax$ as $t \to 0$ (Theorem 4.7), it remains to prove that

$$z_t = t^{-1}(S_A(t)x - S_A(t)S_B(t)x) \to Bx \quad \text{as } t \to 0. \tag{4.132}$$

Because $S_A(t)$ is nonexpansive, we have

$$\|z_t\| \le t^{-1}\|S_B(t)x - x\| \le \|Bx\|, \qquad \forall t > 0. \tag{4.133}$$

On the other hand, inasmuch as $I - S_A(t)$ is accretive, we have

$$\left(\frac{u - S_A(t)u}{t} + \frac{S_A(t)x - S_B(t)x}{t} - z_t, J(u - S_A(t)x) \right) > 0, \qquad \forall u \in C,\ t > 0. \tag{4.134}$$

Let $t_n \to 0$ be such that $z_{t_n} \rightharpoonup z$. Then, by (4.134), we have that

$$(Au + Bx - Ax - z, J(u - x)) \ge 0, \qquad \forall u \in D(A),$$

because $J : X \to X^*$ is continuous and

$$t^{-1}(x - S_B(t)x) \to Bx, \qquad t^{-1}(x - S_A(t)x) \to Ax.$$

Inasmuch as A is m-accretive, this implies that $Ax + z - Bx = Ax$ (i.e., $z = Bx$). On the other hand, by (4.133), recalling that X is uniformly convex, it follows that $z_{t_n} \to Bx$ (strongly). Then, (4.132) follows, and the proof of Theorem 4.16 is complete. □

Remark 4.10. Theorem 4.16, which is essentially due to Brezis and Pazy [16] was extended by Kobayashi [35] to multivalued operators A and B in a Hilbert space H. More precisely, if A, B and $A + B$ are maximal monotone and if there is a nonempty closed convex set $C \subset \overline{D(A) \cap D(B)}$ such that $(I + \lambda A)^{-1}C \subset C$ and $(I + \lambda B)^{-1}C \subset C$, $\forall \lambda > 0$, then

$$S_{A+B}(t)x = \lim_{n\to\infty} \left(S_A\left(\frac{t}{n}\right) S_B\left(\frac{t}{n}\right) \right)^n x, \qquad \forall x \in C,$$

uniformly in t on compact intervals. For some extensions to Banach spaces we refer to Reich [49].

4.3 Time-Dependent Cauchy Problems

This section is concerned with the evolution problem

$$\begin{cases} \dfrac{dy}{dt}(t) + A(t)y(t) \ni f(t), & t \in [0,T], \\ y(0) = y_0, \end{cases} \tag{4.135}$$

where $\{A(t)\}_{t\in[0,T]}$ is a family of quasi-m-accretive operators in $X \times X$.

The existence problem for (4.135) is a difficult one and not completely solved even for linear operators $A(t)$. In general, one cannot expect a positive and convenient answer to the existence problem for (4.135) if one takes into account that in most applications to partial differential equations the domain $D(A(t))$ might not be independent of time. However, we can identify a few classes of time-dependent problems for which the Cauchy problem (4.135) is well posed.

Nonlinear Demicontinuous Evolutions in Duality Pair of Spaces

Let V be a reflexive Banach space and H be a real Hilbert space identified with its own dual such that $V \subset H \subset V'$ algebraically and topologically. The existence result given below is the time-dependent analogue of Theorem 4.10.

Theorem 4.17. *Let $\{A(t);\ t \in [0,T]\}$ be a family of nonlinear, monotone, and demicontinuous operators from V to V' satisfying the assumptions:*

(i) *The function $t \to A(t)u(t)$ is measurable from $[0,T]$ to V' for every measurable function $u : [0,T] \to V$.*

(ii) $(A(t)u,u) \ge \omega\|u\|^p + C_1,\ \forall u \in V,\ t \in [0,T]$.

(iii) $\|A(t)u\|_{V'} \le C_1(1+\|u\|^{p-1}),\ \forall u \in V,\ t \in [0,T]$, *where $\omega > 0$, $p > 1$.*

Then, for every $y_0 \in H$ and $f \in L^q(0,T;V')$, $1/p+1/q=1$, there is a unique absolutely continuous function $y \in W^{1,q}([0,T];V')$ that satisfies

$$\begin{aligned} y \in C([0,T];H) \cap L^p(0,T;V), \\ \frac{dy}{dt}(t) + A(t)y(t) = f(t), \quad \text{a.e. } t \in (0,T), \\ y(0) = y_0. \end{aligned} \tag{4.136}$$

Proof. For the sake of simplicity, we assume first that $p \ge 2$. Consider the spaces

$$\mathscr{V} = L^p(0,T;V), \qquad \mathscr{H} = L^2(0,T;H), \qquad \mathscr{V}' = L^q(0,T;V').$$

Clearly, we have

$$\mathscr{V} \subset \mathscr{H} \subset \mathscr{V}'$$

algebraically and topologically.

Let $y_0 \in H$ be arbitrary and fixed and let $B : \mathscr{V} \to \mathscr{V}'$ be the operator

$$Bu = \frac{du}{dt}, \qquad u \in D(B) = \left\{ u \in \mathscr{V};\ \frac{du}{dt} \in \mathscr{V}',\ u(0) = y_0 \right\},$$

where d/dt is considered in the sense of vectorial distributions on $(0,T)$. We note that $D(B) \subset W^{1,q}(0,T;V') \cap L^q(0,T;V) \subset C([0,T];H)$, so that $y(0) = y_0$ makes sense.

Let us check that B is maximal monotone in $\mathscr{V} \times \mathscr{V}'$. Because B is clearly monotone, by virtue of Theorem 2.3, it suffices to show that $R(B + \Phi_p) = \mathscr{V}'$, where

$$\Phi_p(u(t)) = F(u(t))\|u(t)\|^{p-2}, \qquad u \in \mathscr{V},$$

and $F : V \to V'$ is the duality mapping of V. Indeed, for every $f \in \mathscr{V}'$ the equation

$$Bu + \Phi_p(u) = f,$$

or, equivalently,

$$\frac{du}{dt} + F(u)\|u\|^{p-2} = f \quad \text{in } [0,T],\ u(0) = y_0,$$

has, by virtue of Theorem 4.10, a unique solution

$$u \in C([0,T];H) \cap L^p(0,T;V), \qquad \frac{du}{dt} \in L^q(0,T;V').$$

(Renorming the spaces V and V', we may assume that V and V' are strictly convex and F is demicontinuous and that so is the operator $u \to F(u)\|u\|^{p-2}$.) Hence, B is maximal monotone in $\mathscr{V} \times \mathscr{V}'$.

Define the operator $A_0 : \mathscr{V} \to \mathscr{V}'$ (the realization of A in pair $\mathscr{V}, \mathscr{V}'$) by

$$(A_0 u)(t) = A(t)u(t), \quad \text{a.e. } t \in (0,T).$$

Clearly, A_0 is monotone, demicontinuous, and coercive from $\mathscr{V}$ to $\mathscr{V}'$ because so is $A(t) : V \to V'$.

Then, by Corollaries 2.2 and 2.6, $A_0 + B$ is maximal monotone and surjective. Hence, $R(A_0 + B) = \mathscr{V}'$, which completes the proof.

The proof in the case $1 < p < 2$ is completely similar if we take $\mathscr{V} = L^p(0,T;V) \cap L^2(0,T;H)$ and replace $A(t)$ by $A(t) + \lambda I$ for some $\lambda > 0$. The details are left to the reader. □

Remark 4.11. It should be said that Theorem 4.17 applies neatly to the parabolic boundary value problem

$$\begin{aligned}
\frac{\partial y}{\partial t}(x,t) - \sum_{|\alpha|\le m} D^\alpha(A_\alpha(t,x,y,D^\beta y)) &= f(x,t), && (x,t)\in\Omega\times(0,T)\\
y(x,0) &= y_0(x), && x\in\Omega\\
D^\beta y &= 0 && \text{on } \partial\Omega \text{ for } |\beta|<m,
\end{aligned}$$

where $A_\alpha : [0,T]\times\Omega\times\mathbf{R}^{mN}\to\mathbf{R}^{mN}$ are measurable in (t,x), continuous in other variables and satisfy for each $t\in[0,T]$ assumptions (i)–(iii) in Remark 2.6.

Then we apply Theorem 4.17 for $V = W_0^{m,p}(\Omega), V' = W^{-m,q}(\Omega)$ and $A(t) : V\to V'$ defined by

$$(A(t)y,z) = \sum_{|\alpha|\le m}\int_\Omega A_\alpha(t,x,y(x),D^\beta y(x))\cdot D^\alpha y(x)dx, \qquad \forall y,z\in W_0^{m,p}(\Omega).$$

Hence, for $f\in L^q(0,T;W^{-m,q}(\Omega))$, $y_0\in L^2(\Omega)$, there is a unique solution

$$\begin{aligned}
&y\in L^p(0,T;W_0^{m,p}(\Omega))\cap C([0,T];L^2(\Omega))\\
&\frac{dy}{dt}\in L^q(0,T;W^{-m,q}(\Omega)).
\end{aligned}$$

Subpotential Time-Dependent Evolutions

Let $X = H$ be a real Hilbert space and $A(t) = \partial\varphi(t,y)$, $t\in[0,T]$, where $\varphi(t) : H\to\overline{\mathbf{R}} = (-\infty,\infty]$ is a family of convex and lower semicontinuous functions satisfying the following conditions.

(k) For each measurable function $y : [0,T]\to H$, the function $t\to\varphi(t,y(t))$ is measurable on $(0,T)$.

(kk) $\varphi(t,y)\le\varphi(s,y)+\alpha|t-s|(\varphi(s,y)+|y|^2+1)$ for all $y\in H$ and $0\le s\le t\le T$.

Here α is a nonnegative constant.

We note that, in particular, assumption (kk) implies that $D\varphi(s,\cdot)\subset D\varphi(t,\cdot)$ for all $0\le s\le t\le T$. A standard example of such a family $\{\varphi(t,\cdot)\}_t$ is

$$\varphi(t,\cdot) = I_{K(t)}, \qquad t\in[0,T],$$

where $\{K(t)\}_t$ is an increasing family of closed convex subsets such that the function $t\to P_{K(t)}y(t)$ is measurable for each measurable function $y : [0,T]\to H$. Here, $P_{K(t)} = (I+\lambda\partial I_{K(t)})^{-1}$ is the projection operator on $K(t)$ and the last assumption implies of course (k) for $\varphi(t) = I_{K(t)}$.

Theorem 4.18. *Assume that* $\varphi : [0,T]\times H\to\overline{\mathbf{R}} = (-\infty,\infty]$ *satisfies hypotheses* (k), (kk). *Then, for each* $y_0\in D(\varphi(0,\cdot))$ *and* $f\in L^2(0,T;H)$, *there is a unique pair of functions* $y\in W^{1,2}([0,T];H)$ *and* $\eta\in L^2(0,T;H)$ *such that*

$$\begin{aligned} \eta(t) \in \partial\varphi(t,y(t)), &\quad a.e.\ t \in (0,T), \\ \frac{dy}{dt}(t) + \eta(t) = f(t), &\quad a.e.\ t \in (0,T), \\ y(0) = y_0. & \end{aligned} \tag{4.137}$$

This means that y is solution to (4.135), where $A(t) = \partial\varphi(t,\cdot)$.

Proof. It suffices to prove the existence in the sense of (4.137) for the equation

$$\begin{aligned} \frac{dy}{dt}(t) + \partial\varphi(t,y(t)) + \lambda_0 y(t) \ni f(t), &\quad \text{a.e. } t \in (0,T), \\ y(0) = y_0, & \end{aligned} \tag{4.138}$$

where $\lambda_0 > 0$ is arbitrary but fixed. Indeed, by the substitution $e^{\lambda_0 t} y \to y$, equation (4.138) reduces to

$$\frac{dy}{dt}(t) + e^{\lambda_0 t}\partial\varphi(t, e^{-\lambda_0 t} y(t)) \ni e^{\lambda_0 t} f(t), \qquad t \in [0,T);$$

that is,

$$\frac{dy}{dt} + \partial\widetilde{\varphi}(t,y) \ni e^{\lambda_0 t} f, \qquad t \in (0,T),$$

where $\widetilde{\varphi}(t,y) = e^{2\lambda_0 t}\varphi(t, e^{-\lambda_0 t} y)$ and $e^{\lambda_0 t}\partial\varphi(t, e^{-\lambda_0 t} y) = \partial\widetilde{\varphi}(t,y)$.

Clearly, $\widetilde{\varphi}$ satisfies assumptions (k), (kk).

Now, we may rewrite equation (4.138) in the space $\mathscr{H} = L^2(0,T;H)$ as

$$By + \mathscr{A}y + \lambda_0 y \ni f, \tag{4.139}$$

where

$$\begin{aligned} By &= \frac{dy}{dt}, \quad D(B) = \{y \in W^{1,2}([0,T];H)\ y(0) = y_0\}, \\ \mathscr{A}y &= \{\eta \in L^2(0,T;H);\ \eta(t) \in \partial\varphi(t,y(t)), \quad \text{a.e. } t \in (0,T)\}, \\ D(\mathscr{A}) &= \{y \in L^2(0,T;H),\ \exists \eta \in L^2(0,T;H),\ \eta(t) \in \partial\varphi(t,y(t)), \\ &\qquad \text{a.e. } t \in (0,T)\}. \end{aligned}$$

Because, as easily seen, $\mathscr{A}$ is maximal monotone in $\mathscr{H} \times \mathscr{H}$ and $\mathscr{A} \subset \partial\varphi$, we infer that $\mathscr{A} = \partial\phi$, where $\phi : \mathscr{H} \to (-\infty, +\infty]$ is the convex function

$$\phi(y) = \int_0^T \varphi(t,y(t))dt. \tag{4.140}$$

By assumption (k), it follows via Fatou's lemma that ϕ is also lower semicontinuous and nonidentically $+\infty$ on $\mathscr{H}$. (The latter follows by (kk).)

To prove the existence for equation (4.138) (equivalently (4.139)), we apply Proposition 3.9. To this end it suffices to check the inequality

$$\phi((I+\lambda B)^{-1}y) \leq \phi(y) + C\lambda(\phi(y) + |y|^2_{\mathscr{H}} + 1), \qquad \forall y \in \mathscr{H}. \tag{4.141}$$

We notice that

$$(I+\lambda B)^{-1}y = e^{-(t/\lambda)}y_0 + \frac{1}{\lambda}\int_0^t e^{-(t-s)/\lambda}y(s)ds, \qquad \forall \lambda > 0,\ t \in (0,T),$$

and this yields (by convexity of $y \to \varphi(t,y)$ and by (kk))

$$\begin{aligned}
\phi((I+\lambda B)^{-1}y) &= \int_0^T \varphi\left(t, e^{-(t/\lambda)}y_0 + \frac{1}{\lambda}\int_0^t e^{-(t-s)/\lambda}y(s)ds\right)dt \\
&\leq \int_0^T \left(e^{-(t/\lambda)}\varphi(t,y_0) + \frac{1}{\lambda}\int_0^t e^{-(t-s)/\lambda}\varphi(t,y(s))ds\right)dt \\
&\leq C\lambda(1-e^{-(T/\lambda)})\varphi(0,y_0) + \alpha T(\varphi(0,y_0) + |y_0|^2 + 1)) \\
&\quad + \frac{1}{\lambda}\int_0^T dt \int_0^t e^{-(t-s)/\lambda}\varphi(s,y(s))ds \\
&\quad + \frac{\alpha}{\lambda}\int_0^T dt \int_0^t e^{-(t-s)/\lambda}(\varphi(s,y(s)) + 1 + |y(s)|^2)|t-s|ds \\
&\leq \frac{1}{\lambda}\int_0^T \varphi(s,y(s))ds \int_s^T e^{-(t-s)/\lambda}dt \\
&\quad + \frac{\alpha}{\lambda}\int_0^T (\varphi(s,y(s)) + |y(s)|^2)ds \int_s^T e^{-(t-s)/\lambda}|t-s|dt \\
&\quad + C\lambda(\varphi(0,y_0) + |y_0|^2 + 1 \\
&\leq \phi(y) + C\lambda(\varphi(0,y_0) + \phi(y) + |y|^2_{\mathscr{H}} + 1). \qquad \square
\end{aligned}$$

Time-Dependent *m*-Accretive Evolution

We consider here equation (4.135) under the following assumptions.

(j) $\{A(t)\}_{t\in[0,T]}$ is a family of *m*-accretive operators in X such that, for all $\lambda > 0$,

$$\begin{aligned}
\|A_\lambda(t)y - A_\lambda(s)y\| &\leq C|t-s|(\|A_\lambda(t)y\| + \|y\| + 1), \\
&\qquad \forall y \in X,\ \forall s,t \in [0,T].
\end{aligned} \tag{4.142}$$

Here, $A_\lambda(t)$ is the Yosida approximation of $y \to A(t,y)$. (See (3.1).)

Unlike the previous situations considered here, condition (4.142) has the unpleasant consequence that the domain of $A(t)$ is independent of t; that is, $D(A(t)) \equiv D(A(0))$, $\forall t \in [0,T]$. This assumption is, in particular, too restrictive if we want to treat partial differential equations with time-dependent boundary value conditions, but it is, however, satisfied in a few significant cases involving partial differential equations with smooth time-dependent nonlinearities.

Theorem 4.19. *Assume that X is a reflexive Banach space with uniformly convex dual X^*. If $\{A(t)\}$ satisfies assumption* (j)*, then, for each $f \in W^{1,1}([0,T];X)$ and $y_0 \in D \equiv D(A(t))$, there is a unique function $y \in W^{1,\infty}([0,T];X)$ such that*

$$\begin{cases} \dfrac{dy}{dt}(t) + A(t)y(t) \ni f(t), \quad a.e.\ t \in (0,T), \\ y(0) = y_0. \end{cases} \tag{4.143}$$

Proof. We start, as usual, with the approximating equation

$$\begin{aligned} \frac{dy_\lambda}{dt} + A_\lambda(t)y_\lambda(t) &= f(t), \quad t \in (0,T), \\ y_\lambda(0) &= y_0, \end{aligned} \tag{4.144}$$

which has a unique solution $y_\lambda \in C^1([0,T];X)$. By (4.142) and (4.144) and the accretivity of $A_\lambda(t)$, we see that

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}\|y_\lambda(t+h) - y_\lambda(t)\|^2 \\ &\le (A_\lambda(t+h)y_\lambda(t) - A_\lambda(t)y_\lambda(t), J(y_\lambda(t+h) - y_\lambda(t))) \\ &\le C|h|\,\|y_\lambda(t+h) - y_\lambda(t)\|(\|A_\lambda(t)y_\lambda(t)\| + \|y_\lambda(t)\| + 1), \qquad \forall t, t+h \in [0,T]. \end{aligned}$$

This yields

$$\begin{aligned} &\|y_\lambda(t+h) - y_\lambda(t)\| \\ &\le C\int_0^t (\|A_\lambda(s)y_\lambda(s)\| + \|y_\lambda(s)\| + 1)ds + \|y_\lambda(h) - y_0\|. \end{aligned} \tag{4.145}$$

On the other hand, we have

$$\begin{aligned} \frac{1}{2}\frac{d}{dt}\|y_\lambda(h) - y_\lambda(0)\|^2 = &-(A_\lambda(t)y_\lambda(t), J(y_\lambda(t) - y_0)) \\ &+ (f(t), J(y_\lambda(t) - y_0)), \quad \text{a.e. } t \in (0,T), \end{aligned}$$

and therefore

$$\begin{aligned} \|y_\lambda(h) - y_0\| &\le \int_0^h \|A_\lambda(s)y_0\|ds + \|f\|_{L^\infty(0,T;H)}h \\ &\le h(\|A_\lambda(0)y_0\| + \|f\|_{L^\infty(0,T;H)}). \end{aligned}$$

Then, substituting into (4.144) and letting $h \to 0$, we obtain that

$$\begin{aligned} \left\|\frac{dy_\lambda}{dt}(t)\right\| \le C\Big(&\int_0^t (\|A_\lambda(s)y_\lambda(s)\| + \|y_\lambda(s)\| + 1)ds \\ &+ \|A^0(0)y_0\| + \|f\|_{L^\infty(0,T;H)}\Big), \qquad \forall \lambda > 0. \end{aligned} \tag{4.146}$$

On the other hand, by (4.144) we also have that

$$\|y_\lambda(t)\| \le C, \qquad \forall t \in [0,T],\ \lambda > 0.$$

By (4.144) and (4.146), we get via Gronwall's lemma that

$$\left\| \frac{dy_\lambda}{dt}(t) \right\| + \|A_\lambda(t) y_\lambda(t)\| \le C, \qquad \forall \lambda > 0,\ t \in [0,T]. \tag{4.147}$$

Then, by (4.147) we find as in the proof of Theorem 4.6 that the sequence $\{y_\lambda\}_\lambda$ is Cauchy in $C([0,T];X)$ and $y = \lim_{\lambda \to 0} y_\lambda$ is the solution to (4.143). The details are left to the reader. □

4.4 Time-Dependent Cauchy Problem Versus Stochastic Equations

The above methods apply as well to stochastic differential equations in Hilbert spaces with additive Gaussian noise because, as we show below, these equations can be reduced to time-dependent deterministic equations depending on a random parameter. Below we treat only two problems of this type and refer to standard monographs for complete treatment.

Consider the stochastic differential equation in a separable Hilbert space H,

$$\begin{cases} dX(t) + AX(t)dt = B\,dW(t), & t \ge 0, \\ X(0) = x. \end{cases} \tag{4.148}$$

Here $A : D(A) \subset H \to H$ is a quasi-m-accretive operator in H, $B \in L(U,H)$, where U is another Hilbert space and $W(t)$ is a cylindrical Wiener process in U defined on a probability space $\{\Omega, \mathscr{F}, \mathbb{P}\}$. This means that

$$W(t) = \sum_{k=1}^{\infty} \beta_k(t) e_k,$$

where $\{e_k\}_k$ is an orthonormal basis in U and $\{\beta_k\}_k$ is a sequence of mutually independent Brownian motions on $\{\Omega, \mathscr{F}, \mathbb{P}\}$. Denote by $\mathscr{F}_t$ the σ-algebra generated by $\beta_k(s)$ for $s \le t$, $k \in \mathbf{N}$ (also called *filtration*).

By solution to (4.148) we mean a stochastic process $X = X(t)$ on $\{\Omega, \mathscr{F}, \mathbb{P}\}$ adapted to $\mathscr{F}_t$; that is, $X(t)$ is measurable with respect to the σ-algebra $\mathscr{F}_t$, and satisfies equation

$$X(t) = x - \int_0^t AX(s)ds + \int_0^t B\,dW(s)ds, \qquad \forall t \ge 0,\ \mathbb{P}\text{-a.s.}, \tag{4.149}$$

where the integral $\int_0^t B\,dW(s)$ is considered in the sense of Ito (see Da Prato [28], Da Prato and Zabczyk [29], and Prévot and Roeckner [48]) for the definition and basic existence results for equation (4.149).

A standard way to study the existence for equation (4.148) is to reduce it via substitution

$$y(t) = X(t) - BW(t)$$

to the random differential equation

$$\begin{cases} \dfrac{d}{dt} y(t,\omega) + A(y(t,\omega) + BW(t,\omega)) = 0,\ t \geq 0,\ \mathbb{P}\text{-a.s.},\ \omega \in \Omega, \\ y(0,\omega) = x. \end{cases} \tag{4.150}$$

For almost all $\omega \in \Omega$ (i.e., P-a.s.), (4.150) is a deterministic time-dependent equation in H of the form (4.135); that is,

$$\begin{cases} \dfrac{dy}{dt}(t) + A(t)y(t) = 0, \qquad t \geq 0, \\ y(0) = x, \end{cases}$$

where $A(t)y = A(y + BW(t))$. This fact explains why one cannot expect a complete theory of existence similar to that from the deterministic case. In fact, because the Wiener process $t \to W(t)$ does not have bounded variation, Theorems 4.18 and 4.19 are inapplicable in the present situation. More appropriate for this scope is, however, Theorem 4.17 which requires no regularity in t for $A(t)$.

Then, we assume that V is a reflexive Banach space continuously embedded in H and so we have

$$V \subset H \subset V'$$

algebraically and topologically, where V' is the dual space of V.

Let $A : V \to V'$ satisfy the conditions of Theorem 4.10:

(ℓ) A is a demicontinuous monotone operator and

$$(Au, u) \geq \gamma \|u\|_V^p + C_1, \qquad \forall u \in V,$$

$$\|Au\|_{V'} \leq C_2(1 + \|u\|_V^{p-1}), \qquad \forall u \in V,$$

where $\gamma > 0$ and $p > 1$.

Then, we have the following theorem.

Theorem 4.20. *Assume that A satisfies hypothesis* (ℓ) *and that*

$$BW \in L^p(0,T;V),\ \mathbb{P}\text{-a.s.} \tag{4.151}$$

Then, for each $x \in H$, equation (4.150) *has a unique adapted solution $X = X(t,\omega) \in L^p(0,T;V) \cap C([0,T];H)$, a.e. $\omega \in \Omega$.*

Proof. One simply applies Theorem 4.17 to the operator $A(t)y = A(y + BW(t))$ and check that conditions (i)–(iii) are satisfied under hypotheses (ℓ) and (4.151).

Thus, one finds a solution $X = X(t, \omega)$ to (4.150) that satisfies (4.76) for $\mathbb{P}$-almost all $\omega \in \Omega$. Taking into account that, as seen earlier, such a solution can be obtained as the limit of solutions y_λ to the approximating equations

$$\begin{cases} \dfrac{d}{dt} y_\lambda + A_\lambda(y_\lambda + BW) = 0, & t \in (0,T), \\ y_\lambda(0) = x, \end{cases}$$

where A_λ is the Yosida approximation of $A\big|_H$ (the restriction of the operator A to H), we may conclude that X is adapted with respect to the filtration $\{\mathscr{F}_t\}$. One might also prove H-continuity of $t \to X(t,\omega)$ by the methods of Krylov and Rozovski [39] (see also Prévot and Roeckner [48]), which completes the proof. In particular, Theorem 4.20 applies to parabolic stochastic differential equations of the type mentioned in Remark 4.11. □

It should be said, however, that this variational framework covers only a small part of stochastic partial differential equations because most of them cannot be written in this variational setting and so, in general, other arguments should be involved. This is the case, for instance, with the reflection problem for stochastic differential equations in a Hilbert space H. Namely, for the equation

$$\begin{aligned} dX(t) + (AX(t) + F(X(t)) + \partial I_K(X(t)))dt &\ni \sqrt{Q}\, dW(t), \\ X(0) &= x \in K, \end{aligned} \tag{4.152}$$

where K is a closed convex subset of H such that $0 \in \overset{\circ}{K}$ and

(j) $A : D(A) \subset H \to H$ is a linear self-adjoint operator on H such that A^{-1} is compact and $(Ax, x) \geq \delta |x|^2$, $\forall x \in D(A)$, for some $\delta > 0$.

(jj) $Q : H \to H$ is a linear, bounded, positive, and self-adjoint operator on H such that $Qe^{-tA} = e^{-tA}Q$ for all $t \geq 0$, $Q(H) \subset D(A)$ and $\mathrm{Tr}[AQ] < \infty$.

(jjj) $F : H \to H$ is a Lipschitzian mapping such that, for some $\gamma > 0$, we have

$$(F(x), x) \geq -\gamma |x|^2, \qquad \forall x \in H.$$

(jv) W is a cylindrical Wiener process on H of the form

$$W(t) = \sum_{k=1}^{\infty} \mu_k \beta_k(t) e_k, \qquad t \geq 0,$$

where $\{\beta_k\}$ is a sequence of mutually independent real Brownian motions on filtered probability spaces $(\Omega, \mathscr{F}, \{\mathscr{F}_t\}_{t \geq 0}, \mathbb{P})$ (see [28]) and $\{e_k\}$ is an orthonormal basis in H taken as a system of eigenfunctions for A.

We denote, as usual, by $C([0,T]; H)$ the space of all continuous functions from $[0,T]$ to H and by $BV([0,T]; H)$ the space of all functions with bounded variation from $[0,T]$ to H. We set $V = D(A^{1/2})$ with the norm $\|\cdot\|$ and denote by V' the dual of V in the pairing induced by the scalar product $(\cdot,\cdot)$ of H. By $C_W([0,T];H)$, $L^2_W([0,T];V)$, $L^2_W([0,T];V')$ we denote the standard spaces of adapted processes on $[0,T]$ (see [28, 29]).

Denote by W_A the stochastic convolution,

$$W_A(t) = \int_0^t e^{-A(t-s)}\sqrt{Q}\,dW(s)$$

and note that (4.152) can be rewritten as

$$\begin{cases} \dfrac{d}{dt}Y(t)+AY(t)+F(Y(t)+W_A(t))+\partial I_K(Y(t)+W_A(t)) \ni 0, \\ \qquad\qquad\qquad\qquad \forall t\in(0,T),\ \mathbb{P}\text{-a.s. } \omega\in\Omega \\ Y(0)=x, \end{cases} \tag{4.152$'$}$$

where $Y(t) = X(t) - W_A(t)$.

Definition 4.5. The adapted process $X \in C_W(0,T];H)\cap L^2_W(0,T;V)$ is said to be a solution to (4.152) if there are functions $Y \in C_W([0,T];H)\cap L^2_W(0,T;V)$ and $\eta \in BV([0,T];H)$ such that $X(t) = Y(t)+W_A(t) \in K$, a.e. in $\Omega\times(0,T)$ and

$$Y(t)+\int_0^t (AY(s)+F(X(s)))ds+\eta(t) = x, \qquad \forall t\in[0,T],\ \mathbb{P}\text{-a.s.} \tag{4.153}$$

$$\int_0^t (d\eta(s), X(s)-Z(s))ds \ge 0, \qquad \forall Z\in C([0,T];K),\ \mathbb{P}\text{-a.s.} \tag{4.154}$$

Here $\int_0^t (d\eta(s),X(s)-Z(s))ds$ is the Stieltjes integral with respect to η.

Theorem 4.21 below is an existence result for equation (4.152) (equivalently, (4.152)$'$) and is given only to illustrate how the previous methods work in the case of stochastic infinite-dimensional equations.

Theorem 4.21. *Under the above hypotheses there is a unique strong solution to equation* (4.152)*.*

Proof. Existence. We start with the approximating equation

$$\begin{cases} dX_\varepsilon + (AX_\varepsilon + F(X_\varepsilon)+\beta_\varepsilon(X_\varepsilon))dt = \sqrt{Q}\,dW, \\ X_\varepsilon(0) = x, \end{cases} \tag{4.155}$$

where β_ε is the Yosida approximation of ∂I_K,

$$\beta_\varepsilon(x) = \frac{1}{\varepsilon}(x-\Pi_K(x)), \qquad \forall x\in H,\ \varepsilon>0,$$

and Π_K is the projection on K.

Equation (4.155) has a unique strong solution $X_\varepsilon \in C_W([0,T];H)$ such that $Y_\varepsilon := X_\varepsilon - W_A$ belongs to $L^2_W(0,T;H)$. As seen above, we can rewrite (4.155) as

$$\begin{cases} \dfrac{dY_\varepsilon}{dt} + AY_\varepsilon + F(X_\varepsilon) + \beta_\varepsilon(X_\varepsilon) = 0, \\ Y_\varepsilon(0) = x, \end{cases} \tag{4.156}$$

which is considered here for a fixed $\omega \in \Omega$. Because $0 \in \overset{\circ}{K}$, there is $\rho > 0$ such that $(\beta_\varepsilon(x), x - \rho\theta) \geq 0$, $\forall \theta \in H$, $|\theta| = 1$. This yields $\rho|\beta_\varepsilon(x)| \leq (\beta_\varepsilon(x), x)$, $\forall x \in H$.

Step 1. There exists $C = C(\omega) > 0$ such that

$$|Y_\varepsilon(t)|^2 + \int_0^t \|Y_\varepsilon(s)\|^2 ds + \int_0^t |\beta_\varepsilon(X_\varepsilon(s))| ds \leq C. \tag{4.157}$$

Indeed, multiplying (4.156) scalarly in H by $Y_\varepsilon(s)$ and integrating over $(0,t)$ yields

$$\begin{aligned} &\frac{1}{2}|Y_\varepsilon(t)|^2 + \int_0^t \|Y_\varepsilon(s)\|^2 ds + \rho \int_0^t |\beta_\varepsilon(X_\varepsilon(s))| ds \\ &\leq \frac{1}{2}|x|^2 + \gamma \int_0^t |X_\varepsilon(s)|^2 ds + \int_0^t (F(X_\varepsilon(s)) + \beta_\varepsilon(X_\varepsilon(s)), W_A(s)) ds. \end{aligned} \tag{4.158}$$

In order to estimate the last term in formula (4.158), we choose a decomposition $0 < t_1 < \cdots < t_N = t$ of $[0,t]$ such that, for $t, s \in [t_{i-1}, t_i]$, we have

$$|W_A(t) - W_A(s)| \leq \frac{\rho}{2}.$$

This is possible because W_A is $\mathbb{P}$-a.s. continuous in H, and so we may assume that

$$\sup_{t \in [0,T]} |W_A(t+h) - W_A(t)| \leq \delta(h) \to 0 \quad \text{as } h \to 0,$$

because by (jj) it follows that W_A is $\mathbb{P}$-a.s. continuous in H (see Da Prato [28]).

Then, we write

$$\begin{aligned} \int_0^t (\beta_\varepsilon(X_\varepsilon(s)), W_A(s)) ds = &\sum_{i=1}^N \int_{t_{i-1}}^{t_i} (\beta_\varepsilon(X_\varepsilon(s)), W_A(s) - W_A(t_i)) ds \\ &+ \sum_{i=1}^N \left(W_A(t_i), \int_{t_{i-1}}^{t_i} \beta_\varepsilon(X_\varepsilon(s)) ds \right). \end{aligned}$$

Consequently,

$$\begin{aligned} &\int_0^t (\beta_\varepsilon(X_\varepsilon(s)), W_A(s)) ds \leq \frac{\rho}{2} \int_0^t |\beta_\varepsilon(X_\varepsilon(s))| ds \\ &+ \left| \sum_{i=1}^N \left(W_A(t_i), \int_{t_{i-1}}^{t_i} (AY_\varepsilon(s) + F(X_\varepsilon(s))) ds + Y_\varepsilon(t_i) - Y_\varepsilon(t_{i-1}) \right) \right|. \end{aligned}$$

Now, using the estimate

$$\left(W_A(t_i), \int_{t_{i-1}}^{t_i} AY_\varepsilon(s)ds\right) \le C\int_{t_{i-1}}^{t_i} \|Y_\varepsilon(s)\|^2 ds,$$

we get (4.157).

We now prove that the sequence $\{Y_\varepsilon\}$ is equicontinuous in $C([0,T];H)$. Let $h>0$, then we have

$$\begin{aligned}&\frac{d}{dt}\left(Y_\varepsilon(t+h)-Y_\varepsilon(t)\right)+A(Y_\varepsilon(t+h)-Y_\varepsilon(t))\\ &+F(X_\varepsilon(t+h))-F(X_\varepsilon(t))+\beta_\varepsilon(X_\varepsilon(t+h))-\beta_\varepsilon(X_\varepsilon(t))=0.\end{aligned}$$

By the monotonicity of β_ε and because F is Lipschitz continuous, we have

$$|Y_\varepsilon(t+h)-Y_\varepsilon(t)| \le C\delta(h), \qquad \forall t\in[0,T],\ h>0,\ \varepsilon>0.$$

So, $\{Y_\varepsilon\}$ is equi-continuous. To apply the Ascoli-Arzelà theorem, we have to prove that, for each $t\in[0,T]$, the set $\{Y_\varepsilon(t)\}_{\varepsilon>0}$ is pre-compact in H. To prove this, choose for any $\varepsilon>0$ a sequence $\{f_n^\varepsilon\}\subset L^2(0,T;V)$ such that

$$|f_n^\varepsilon-\beta_\varepsilon(Y_\varepsilon+W_A)|_{L^1(0,T;H)} \le \frac{1}{n}, \qquad n\in\mathbf{N}.$$

On the other hand, for each $n\in\mathbf{N}$, the set

$$M_n := \left\{\int_0^t e^{-A(t-s)} f_n^\varepsilon ds + e^{-At}x : \varepsilon>0\right\}$$

is compact in H because $\{f_n^\varepsilon\}$ is bounded in $L^2(0,T;H)$ for each $n\in\mathbf{N}$. This implies that, for any $\delta>0$, there are $N(n)\in\mathbf{N}$ and $\{u_i^n\}_{i=1,\dots,N(n)}\subset H$ such that

$$\bigcup_{i=1}^{N(n)} B(u_i^n,\delta)\supset M_n.$$

Therefore,

$$\left\{Y_\varepsilon(t) := \int_0^t e^{-A(t-s)} f_n^\varepsilon ds + e^{-At}x : \varepsilon>0\right\} \subset \bigcup_{i=1}^{N(n)} B(u_i^n,\delta+n^{-1}).$$

Hence, the set $\{Y_\varepsilon(t)\}_{\varepsilon>0}$ is precompact in H, as claimed. Then, by the Ascoli–Arzelà theorem we infer that on a subsequence, $Y_\varepsilon\to Y$ strongly in $C([0,T];H)$ and weakly in $L^2(0,T;V)$. Moreover, thanks to Helly's theorem (see [9]), we have that there is $\eta\in BV([0,T];H)$ such that, for $\varepsilon\to 0$,

$$\int_0^t \beta_\varepsilon(X_\varepsilon(s))ds \to \eta(t) \quad \text{weakly in } H, \qquad \forall t\in[0,T],$$

which implies that

$$\int_0^t (\beta_\varepsilon(X_\varepsilon(s)), Z(s))ds \to \int_0^t (d\eta(s), Z(s))ds, \qquad \forall Z \in C([0,T];K).$$

Letting $\varepsilon \to 0$ into the identity

$$Y_\varepsilon(t) + \int_0^t (AY_\varepsilon(s + F(Y_\varepsilon(s)))ds + \int_0^t \beta_\varepsilon(Y_\varepsilon(s) + W_A(s)))ds = x,$$

we see that (Y, η) satisfy (4.153).

Finally, by the monotonicity of β_ε we have (recall that $\beta_\varepsilon(Z(s)) = 0$),

$$(\beta_\varepsilon(Y_\varepsilon(s) + W_A(s)), Y_\varepsilon(s) + W_A(s) - Z(s)) \geq 0, \qquad \forall Z \in C([0,T];K),$$

and so (4.154) holds.

Uniqueness. Assume that $(Y_1, \eta_1), (Y_2, \eta_2)$ are two solutions. Then, we have

$$\int_0^t (d(\eta_1(s) - \eta_2(s)), Y_1(s) - Y_2(s))ds \geq 0, \qquad \forall t \in [0,T].$$

This yields

$$\int_0^t \Big(d(Y_1(s) - Y_2(s)) + \int_0^s (A(Y_1(\tau) - Y_2(\tau)) + F(X_1(\tau) - F(X_2(\tau)))d\tau, Y_1(s) - Y_2(s)) \Big) \leq 0$$

and, by integration, we obtain that

$$\frac{1}{2}|Y_1(t) - Y_2(t)|^2 + \int_0^t (A(Y_1 - Y_2) + F(X_1) - F(X_2), Y_1 - Y_2)ds \leq 0,$$

$\forall t \in [0,T]$, which implies via Gronwall's lemma that $Y_1 = Y_2$.

In particular, the latter implies that the sequence $\{\varepsilon\}$ founded before is independent of ω and so, there is indeed a unique pair satisfying Definition 4.5. (For proof details, we refer to Barbu and Da Prato [6].) □

Remark 4.12. The above argument can be formalized to treat more general equations of the form $(4.152)'$ and, in particular, the so-called variational inequalities with singular inputs (see Barbu and Răşcanu [7]). In the literature, such a problem is also called the Skorohod problem (see, e.g., Cépa [20]).

Bibliographical Remarks

The existence theory for the Cauchy problem associated with nonlinear *m*-accretive operators in Banach spaces begins with the influential pioneering papers of Komura [37, 38] and Kato [32] in Hilbert spaces. The theory was subsequently extended in a more general setting by several authors mentioned below.

The main result of Section 4.1 is due to Crandall and Evans [23] (see also Crandall [22]), and Theorem 4.3 has been previously proved by Crandall and Liggett [24]. The existence and uniqueness of integral solutions for problem (4.1) (see Theorem 4.18) is due to Bénilan [10]. Theorems 4.5 and 4.6 were established in a particular case in Banach space by Komura [37] (see also Kato [32]) and later extended in Banach spaces with uniformly convex duals by Crandall and Pazy [25, 26]. Note that the generation theorem, 4.3 remains true for m-accretive operators satisfying the extended range condition (Kobayashi [35])

$$\liminf_{h\downarrow 0} \frac{1}{h} d(x, R(I+\lambda A)) = 0, \qquad \forall x \in \overline{D(A)},$$

$d(x,K)$ is the distance from x to K.

The basic properties of continuous semigroups of contractions have been established by Komura [38], Kato [33], and Crandall and Pazy [25, 26]. For other significant results of this theory, we refer the reader to the author's book [5]. (See also Showalter [50].) The results of Section 4.4 are due to Brezis [13, 14]. Other results related to the smoothing effect of nonlinear semigroups are given in the book by Barbu [5].

Convergence results of the type presented in Section 4.2 were obtained by Brezis and Pazy [16], Kobayashi and Myadera [36], and Goldstein [30].

Time-dependent differential equations of subdifferential type under conditions given here (Section 4.3) were studied by Moreau [41], Peralba [47], Kenmochi [34], and Attouch and Damlamian [3].

Other special problems related to evolutions generated by nonlinear accretive operators are treated in Vrabie's book [54]. We mention in this context a characterization of compact semigroups of nonlinear contractions and evolutions generated by operators of the form $A+F$, where A is m-accretive and F is upper semicontinuous and compact. For other results such as asymptotic behavior and existence of periodic and almost periodic solutions to problem (4.1), we refer the reader to the monographs of Haraux [31] and Moroşanu [42].

We have omitted from our presentation the invariance and viability results related to nonlinear contraction semigroups on closed subsets. We mention in this context the books of Aubin and Cellina [4], Pavel [43, 44] and the recent monograph of Cârjă, Necula, and Vrabie [19], which contains detailed results and complete references on this subject.

References

1. H. Attouch, Familles d'opérateurs maximaux monotones et mesurabilité, *Annali Mat. Pura Appl.*, **CXX** (1979), pp. 35–111.
2. H. Attouch, *Variational Convergence for Functions and Operators*, Pitman, Boston, 1984.
3. H. Attouch, A. Damlamian, Problèmes d'évolution dans les Hilbert et applications, *J. Math. Pures Appl.*, **54** (1975), pp. 53–74.
4. P. Aubin, A. Cellina, *Differential Inclusions*, Springer-Verlag, Berlin, 1984.

5. V. Barbu, *Nonlinear Semigroups and Differential Equations in Banach Spaces*, Noordhoff, Leyden, 1976.
6. V. Barbu, G. Da Prato, Some results for the reflection problems in Hilbert spaces, *Control Cybern.*, **37** (2008), pp. 797–810.
7. V. Barbu, A. Răşcanu, Parabolic variational inequalities with singular inputs, *Differential Integral Equ.*, **10** (1997), pp. 67–83.
8. V. Barbu, C. Marinelli, Variational inequalities in Hilbert spaces with measures and optimal stopping problems, *Appl. Math. Optimiz.*, **57** (2008), pp. 237–262.
9. V. Barbu, T. Precupanu, *Convexity and Optimization in Banach Spaces*, D. Reidel, Dordrecht, 1987.
10. Ph. Bénilan, *Equations d'évolution dans un espace de Banach quelconque et applications*, Thèse, Orsay, 1972.
11. Ph. Bénilan, H. Brezis, Solutions faibles d'équations d'évolution dans les espaces de Hilbert, *Ann. Inst. Fourier*, **22** (1972), pp. 311–329.
12. Ph. Bénilan, S. Ismail, Générateurs des semigroupes nonlinéaires et la formule de Lie-Trotter, *Annales Faculté de Sciences*, Toulouse, **VII** (1985), pp. 151–160.
13. H. Brezis, *Opérateurs maximaux monotones et semigroupes de contractions dans les espaces de Hilbert*, North Holland, Amsterdam, 1975.
14. H. Brezis, Propriétés régularisantes de certaines semi-groupes nonlinéaires, *Israel J. Math.*, **9** (1971), pp. 513–514.
15. H. Brezis, Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations, *Contributions to Nonlinear Functional analysis*, E. Zarantonello (Ed.), Academic Press, New York, 1971.
16. H. Brezis, A. Pazy, Semigroups of nonlinear contrctions on convex sets, *J. Funct. Anal.*, **6** (1970), pp. 367–383.
17. H. Brezis, A. Pazy, Convergence and approximation of semigroups of nonlinear operators in Banach spaces, *J. Funct. Anal.*, **9** (1971), pp. 63–74.
18. R. Bruck, Asymptotic convergence of nonlinear contraction semigroups in Hilbert space, *J. Funct. Anal.*, **18** (1975), pp. 15–26.
19. O. Cârjă, M. Necula, I.I. Vrabie, *Viability, Invariance and Applications*, North–Holland Math. Studies, Amsterdam, 2007.
20. E. Cépa, Problème de Skorohod multivoque, *Ann. Probab.*, **26** (1998), pp. 500–532.
21. P. Chernoff, Note on product formulas for opertor semi-groups, *J. Funct. Anal.*, **2** (1968), pp. 238–242.
22. M.G. Crandall, Nonlinear semigroups and evolutions generated by accretive operators, *Nonlinear Functional Analysis and Its Applications*, pp. 305–338, F. Browder (Ed.), American Mathematical Society, Providence, RI, 1986.
23. M.G. Crandall, L.C. Evans, On the relation of the operator $\partial/\partial s + \partial/\partial t$ to evolution governed by accretive operators, *Israel J. Math.*, **21** (1975), pp. 261–278.
24. M.G. Crandall, T.M. Liggett, Generation of semigroups of nonlinear transformations in general Banach spaces, *Amer. J. Math.*, **93** (1971), pp. 265–298.
25. M.G. Crandall, A. Pazy, Semigroups of nonlinear contractions and dissipative sets, *J. Funct. Anal.*, **3** (1969), pp. 376–418.
26. M.G. Crandall, A. Pazy, Nonlinear evolution equations in Banach spaces, *Israel J. Math.*, **11** (1972), pp. 57–94.
27. C. Dafermos, M. Slemrod, Asymptotic behaviour of nonlinear contraction semigroups, *J. Funct. Anal.*, **12** (1973), pp. 96–106.
28. G. Da Prato, *Kolmogorov Equations for Stochastic PDEs*, Birhkäuser Verlag, Basel, 2004.
29. G. Da Prato, J. Zabczyk, *Stochastic Equations in Infinite Dimensions*, Cambridge University Press, Cambridge, UK, 1992.
30. J. Goldstein, Approximation of nonlinear semigroups and ev olution equations, *J. Math. Soc. Japan*, **24** (1972), pp. 558–573.
31. A. Haraux, *Nonlinear Evolution Equations. Global Behaviour of solutions*, Lecture Notes in Mathematics, Springer-Verlag, Berlin, 1981.

32. T. Kato, Nonlinear semigroups and evolution equations, *J. Math. Soc. Japan*, **19** (1967), pp. 508–520.
33. T. Kato, Accretive operators and nonlinear evolution equations in Banach spaces, *Nonlinear Functional Analysis*, Proc. Symp. Pure Math., vol. **13**, F. Browder (Ed.), American Mathematical Society, (1970), pp. 138–161.
34. N. Kenmochi, Nonlinear parabolic variational inequalities with time-dependent constraints, *Proc. Japan Acad.*, **53** (1977), pp. 163–166.
35. Y. Kobayashi, Difference approximation of Cauchy problem for quasi-dissipative operators and generation of nonlinear semigroups, *J. Math. Soc. Japan*, **27** (1975), pp. 641–663.
36. Y. Kobayashi, I. Miyadera, Donvergence and approximation of nonlinear semigroups, *Japan-France Seminar*, pp. 277–295, H. Fujita (Ed.), Japan Soc. Promotion Sci., Tokyo, 1978.
37. Y. Komura, Nonlinear semigroups in Hilbert spaces, *J. Math. Soc. Japan*, **19** (1967), pp. 508–520.
38. Y. Komura, Differentiability of nonlinear semigroups, *J. Math. Soc. Japan*, **21** (1969), pp. 375–402.
39. N. Krylov, B. Rozovski, *Stochastic Evolution Equations*, Plenum, New York, 1981.
40. J.L. Lions, *Quelques Méthodes de Resolution des Problèmes aux Limites Nonlinéaires*, Dunod, Paris, 1969.
41. J.J. Moreau, Evolution problem associated with a moving convex set associated with a moving convex set in a Hilbert space, *J. Differential Equ.*, **26** (1977), pp. 347–374.
42. G. Moroşanu, *Nonlinear Evolution Equations and Applications*, D. Reidel, Dordrecht, 1988.
43. N. Pavel, *Differential Equations, Flow Invariance and Applications*, Research Notes Math., **113**, Pitman, Boston, 1984.
44. N. Pavel, *Nonlinear Evolution Equations, Operators and Semigroups*, Lecture Notes, **1260**, Springer-Verlag, New York, 1987.
45. A. Pazy, *Semigroups of Linear Operators and Applications,* Springer-Verlag, New York, 1979.
46. A. Pazy, The Lyapunov method for semigroups of nonlinear contractions in Banach spaces, *J. Analyse Math.*, **40** (1982), pp. 239–262.
47. J.C. Peralba, Un problème d'évolution relativ à un opérateur sous-différentiel dépendent du temps, *C.R.A.S. Paris*, **275** (1972), pp. 93–96.
48. C. Prévot, M. Roeckner, *A Concise Course on Stochastic Partial Differential Equations*, Lect. Notes Math., **1905**, Springer, New York, 2007.
49. S. Reich, Product formulas, nonlinear semigroups and accretive operators in Banach spaces, *J. Funct. Anal.*, **36** (1980), pp. 147–168.
50. R.E. Showalter, *Monotone Operators in Banach Spaces and Nonlinear Partial Differential Equations*, American Mathematical Society, Providence, RI, 1977.
51. U. Stefanelli, The Brezis-Ekeland principle for doubly nonlinear equations, *SIAM J. Control Optim.*, **8** (2008), pp. 1615–1642.
52. L. Véron, Effets régularisant de semi-groupes non linéaire dans des espaces de Banach, *Ann. Fc. Sci. Toulouse Math.*, **1** (1979), pp. 171–200.
53. A. Visintin, Extension of the Brezis-Ekeland-Nayroles principle to monotone operators (to appear).
54. I.I. Vrabie, *Compactness Methods for Nonlinear Evolutions*, Pitman Monographs and Surveys in Pure and Applied Mathematics, Second Edition, **75**, Addison Wesley and Longman, Reading, MA, 1995.

Chapter 5
Existence Theory of Nonlinear Dissipative Dynamics

Abstract In this chapter we present several applications of general theory to nonlinear dynamics governed by partial differential equations of dissipative type illustrating the ideas and general existence theory developed in the previous section. Most of significant dynamics described by partial differential equations can be written in the abstract form (4.1) with appropriate quasi-m-accretive operator A and Banach space X. The boundary value conditions are incorporated in the domain of A. The whole strategy is to find the appropriate operator A and to prove that it is quasi-m-accretive. The main emphasis here is on parabolic-like boundary value problems and the nonlinear hyperbolic equations although the area of problems covered by general theory is much larger.

5.1 Semilinear Parabolic Equations

The classical linear heat (or diffusion) equation perturbed by a nonlinear potential $\beta = \beta(y)$, where y is the state of system, is the simplest form of semilinear parabolic equation arising in applications and is treated below. The nonlinear potential β might describe exogeneous driving forces intervening over diffusion process or might induce unilateral state constraints.

The principal motivation for choosing multivalued functions β in examples below is to treat problems with a free (or moving) boundary as well as problems with discontinuous monotone nonlinearities. In the latter case, filling the jumps $[\beta(r_0 - 0), \beta(r_0 + 0)]$ of function β, we get a maximal monotone multivalued graph $\beta \subset \mathbf{R} \times \mathbf{R}$ for which the general existence theory applies.

To be more specific, assume that β is a maximal monotone graph such that $0 \in D(\beta)$, and Ω is an open and bounded subset of $\mathbf{R}^N$ with a sufficiently smooth boundary $\partial\Omega$ (for instance, of class C^2). Consider the parabolic boundary value problem

V. Barbu, *Nonlinear Differential Equations of Monotone Types in Banach Spaces*, Springer Monographs in Mathematics, DOI 10.1007/978-1-4419-5542-5_5,

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y + \beta(y) \ni f & \text{in } \Omega \times (0,T) = Q, \\ y(x,0) = y_0(x) & \forall x \in \Omega, \\ y = 0 & \text{on } \partial\Omega \times (0,T) = \Sigma, \end{cases} \tag{5.1}$$

where $y_0 \in L^2(\Omega)$ and $f \in L^2(\Omega)$ are given.

We may represent problem (5.1) as a nonlinear differential equation in the space $H = L^2(\Omega)$:

$$\begin{cases} \dfrac{dy}{dt}(t) + Ay(t) \ni f(t), & t \in [0,T], \\ y(0) = y_0, \end{cases} \tag{5.2}$$

where $A : L^2(\Omega) \to L^2(\Omega)$ is the operator defined by

$$\begin{aligned} Ay &= \{z \in L^2(\Omega);\ z = -\Delta y + w,\ w(x) \in \beta(y(x)),\ \text{a.e. } x \in \Omega\}, \\ D(A) &= \{y \in H_0^1(\Omega) \cap H^2(\Omega);\ \exists w \in L^2(\Omega),\ w(x) \in \beta(y(x)),\ \text{a.e. } x \in \Omega\}. \end{aligned} \tag{5.3}$$

Here, $(d/dt)y$ is the strong derivative of $y : [0,T] \to L^2(\Omega)$ and

$$\Delta y = \sum_{i=1}^{N} (\partial^2 y / \partial x_i^2)$$

is considered in the sense of distributions on Ω.

As a matter of fact, it is readily seen that if y is absolutely continuous from $[a,b]$ to $L^1(\Omega)$, then $dy/dt = \partial y/\partial t$ in $\mathscr{D}'((a,b);\ L^1(\Omega))$, and so a strong solution to equation (5.2) satisfies this equation in the sense of distributions in $(0,T) \times \Omega$. For this reason, whenever there is no any danger of confusion we write $\partial y/\partial t$ instead of dy/dt.

Recall (see Proposition 2.8) that A is maximal monotone (i.e., m-accretive) in $L^2(\Omega) \times L^2(\Omega)$ and $A = \partial\varphi$, where

$$\varphi(y) = \begin{cases} \dfrac{1}{2} \displaystyle\int_\Omega |\nabla y|^2 dx + \int_\Omega g(y) dx, & \text{if } y \in H_0^1(\Omega),\ g(y) \in L^1(\Omega), \\ +\infty, & \text{otherwise,} \end{cases}$$

and $\partial g = \beta$. Moreover, we have

$$\|y\|_{H^2(\Omega)} + \|y\|_{H_0^1(\Omega)} \le C(\|A^0 y\|_{L^2(\Omega)} + 1), \qquad \forall y \in D(A). \tag{5.4}$$

Writing equation (5.1) in the form (5.2), we view its solution y as a function of t from $[0,T]$ to $L^2(\Omega)$. The boundary conditions that appear in (5.1) are implicitly incorporated into problem (5.2) through the condition $y(t) \in D(A)$, $\forall t \in [0,T]$.

The function $y : \Omega \times [0,T] \to \mathbf{R}$ is called a *strong solution* to problem (5.1) if $y : [0,T] \to L^2(\Omega)$ is continuous on $[0,T]$, absolutely continuous on $(0,T)$, and satisfies

$$\begin{cases} \dfrac{d}{dt}y(x,t) - \Delta y(x,t) + \beta(y(x,t)) \ni f(x,t), & \text{a.e. } t \in (0,T),\ x \in \Omega, \\ y(x,0) = y_0(x), & \text{a.e. } x \in \Omega, \\ y(x,t) = 0, & \text{a.e. } x \in \partial\Omega,\ t \in (0,T). \end{cases} \tag{5.5}$$

Proposition 5.1. *Let* $y_0 \in L^2(\Omega)$ *and* $f \in L^2(0,T;L^2(\Omega)) = L^2(Q)$ *be such that* $y_0(x) \in \overline{D(\beta)}$, *a.e.* $x \in \Omega$. *Then, problem* (5.1) *has a unique strong solution*

$$y \in C([0,T];L^2(\Omega)) \cap W^{1,1}((0,T];L^2(\Omega))$$

that satisfies

$$t^{1/2}y \in L^2(0,T;H_0^1(\Omega) \cap H^2(\Omega)),\ t^{1/2}\frac{dy}{dt} \in L^2(0,T;L^2(\Omega)). \tag{5.6}$$

If, in addition, $f \in W^{1,1}([0,T];L^2(\Omega))$, *then* $y(t) \in H_0^1(\Omega) \cap H^2(\Omega)$ *for every* $t \in (0,T]$ *and*

$$t\ \frac{dy}{dt} \in L^\infty(0,T;L^2(\Omega)). \tag{5.7}$$

If $y_0 \in H_0^1(\Omega),\ g(y_0) \in L^1(\Omega)$, *and* $f \in L^2(0,T;L^2(\Omega))$, *then*

$$\frac{dy}{dt} \in L^2(0,T;L^2(\Omega)), \qquad y \in L^\infty(0,T;H_0^1(\Omega)) \cap L^2(0,T;H^2(\Omega)). \tag{5.8}$$

Finally, if $y_0 \in D(A)$ *and* $f \in W^{1,1}([0,T];L^2(\Omega))$, *then*

$$\frac{dy}{dt} \in L^\infty(0,T;L^2(\Omega)), \qquad y \in L^\infty(0,T;H^2(\Omega) \cap H_0^1(\Omega)) \tag{5.9}$$

and

$$\frac{d^+}{dt}y(t) + (-\Delta y(t) + \beta(y(t)) - f(t))^0 = 0, \qquad \forall t \in [0,T]. \tag{5.10}$$

Proof. This is a direct consequence of Theorems 4.11 and 4.12, because, as seen in Proposition 2.8, we have

$$\overline{D(A)} = \{u \in L^2(\Omega);\ u(x) \in \overline{D(\beta)}, \quad \text{a.e. } x \in \Omega\}.$$

In particular, it follows that for $y_0 \in H_0^1(\Omega)$, $g(y_0) \in L^1(\Omega)$, and $f \in L^2(\Omega \times (0,T))$, the solution y to problem (5.1) belongs to the space

$$H^{2,1}(Q) = \left\{ y \in L^2(0,T;H^2(\Omega)),\ \frac{\partial y}{\partial t} \in L^2(Q) \right\}, \qquad Q = \Omega \times (0,T).$$

Problem (5.1) can be studied in the L^p setting, $1 \le p < \infty$ as well, if one defines the operator $A : L^p(\Omega) \to L^p(\Omega)$ as

$$Ay = \{z \in L^p(\Omega);\ z = -\Delta y + w,\ w(x) \in \beta(y)),\ \text{a.e. } x \in \Omega\}, \tag{5.11}$$

$$D(A) = \{y \in W_0^{1,p}(\Omega) \cap W^{2,p}(\Omega);\ w \in L^p(\Omega) \text{ such that} \\ w(x) \in \beta(y(x)),\ \text{a.e. } x \in \Omega\} \quad \text{if } p > 1, \tag{5.12}$$

$$D(A) = \{y \in W_0^{1,1}(\Omega);\ \Delta y \in L^1(\Omega),\ \exists w \in L^1(\Omega) \text{ such that} \\ w(x) \in \beta(y(x)),\ \text{a.e. } x \in \Omega\} \quad \text{if } p = 1. \tag{5.13}$$

As seen earlier (Theorem 3.2), the operator A is m-accretive in $L^p(\Omega) \times L^p(\Omega)$ and so, also in this case, the general existence theory is applicable. □

Proposition 5.2. *Let $y_0 \in D(A)$ and $f \in W^{1,1}([0,T];L^p(\Omega))$, $1 < p < \infty$. Then, problem* (5.1) *has a unique strong solution*

$$y \in C([0,T];L^p(\Omega)),$$

that satisfies

$$\frac{d}{dt}y \in L^\infty(0,T;L^p(\Omega)),\ y \in L^\infty(0,T;W_0^{1,p}(\Omega) \cap W^{2,p}(\Omega)) \tag{5.14}$$

$$\frac{d^+}{dt}y(t) + (-\Delta y(t) + \beta(y(t)) - f(t))^0 = 0, \qquad \forall t \in [0,T]. \tag{5.15}$$

If $y_0 \in \overline{D(A)}$ and $f \in L^1(0,T;L^p(\Omega))$, then (5.1) *has a unique mild solution*

$$y \in C([0,T];L^p(\Omega)).$$

Proof. Proposition 5.2 follows by Theorem 4.6 (recall that $X = L^p(\Omega)$ is uniformly convex for $1 < p < \infty$). □

Next, by Theorem 4.1 we have the following.

Proposition 5.3. *Assume $p = 1$. Then, for each $y_0 \in \overline{D(A)}$ and $f \in L^1(0,T;L^1(\Omega))$, problem* (5.1) *has a unique mild solution $y \in C([0,T];L^1(\Omega))$; that is,*

$$y(t) = \lim_{\varepsilon \to 0} y_\varepsilon(t),$$

where y_ε is the solution to the finite difference scheme

$$\begin{aligned} & y_\varepsilon^{i+1} = y_\varepsilon^i + \varepsilon\Delta y_\varepsilon^{i+1} - \varepsilon\beta(y_\varepsilon^{i+1}) + \int_{i\varepsilon}^{(i+1)\varepsilon} f(t)dt && \text{in } \Omega,\ i = 0,1,\dots,m, \\ & && m = \left[\tfrac{T}{\varepsilon}\right] + 1, \\ & y_\varepsilon^{i+1} \in H_0^1(\Omega) && \\ & y_\varepsilon(t) = y_\varepsilon^i && \text{for } t \in (i\varepsilon,(i+1)\varepsilon). \end{aligned}$$

Because the space $X = L^1(\Omega)$ is not reflexive, the mild solution to the Cauchy problem (5.2) in $L^1(\Omega)$ is only continuous as a function of t, even if y_0 and f are regular. However, also in this case we have a regularity property of mild solutions; that is, a smoothing effect on initial data, which resembles the case $p = 2$.

Proposition 5.4. *Let $\beta : \mathbf{R} \to \mathbf{R}$ be a maximal monotone graph, $0 \in \overline{D(\beta)}$, and $\beta = \partial g$. Let $f \in L^2(0,T;L^\infty(\Omega))$ and $y_0 \in L^1(\Omega)$ be such that $y_0(x) \in \overline{D(\beta)}$, a.e. $x \in \Omega$. Then, the mild solution $y \in C([0,T];L^1(\Omega))$ to problem* (5.1) *satisfies*

$$\|y(t)\|_{L^\infty(\Omega)} \le C\left(t^{-(N/2)}\|y_0\|_{L^1(\Omega)} + \int_0^t \|f(s)\|_{L^\infty(\Omega)}ds\right), \tag{5.16}$$

$$\begin{aligned}&\int_0^T\int_\Omega (t^{(N+4)/2}y_t^2 + t^{(N+2)/2}|\nabla y|^2)dx\,dt + T^{(N+4)/2}\int_\Omega |\nabla y(x,T)|^2dx \\ &\le C\left(\left(\|y_0\|_{L^1(\Omega)}^{4/(N+2)} + \int_0^T\int_\Omega |f|dx\,dt\right)^{(N+2)/2} + T^{(N+4)/2}\int_0^T\int_\Omega f^2dx\,dt\right).\end{aligned} \tag{5.17}$$

Proof. Without loss of generality, we may assume that $0 \in \beta(0)$. Also, let us assume first that $y_0 \in H_0^1(\Omega) \cap H^2(\Omega)$. Then, as seen in Proposition 5.1, problem (5.1) has a unique strong solution such that $t^{1/2}y_t \in L^2(Q)$, $t^{1/2}y \in L^2(0,T;H_0^1(\Omega) \cap H^2(\Omega))$:

$$\begin{cases} \dfrac{\partial y}{\partial t}(x,t) - \Delta y(x,t) + \beta(y(x,t)) \ni f(x,t), & \text{a.e. } (x,t) \in Q, \\ y(x,0) = y_0(x), & x \in \Omega, \\ y = 0, & \text{on } \Sigma. \end{cases} \tag{5.18}$$

Consider the linear problem

$$\begin{cases} \dfrac{\partial z}{\partial t} - \Delta z = \|f(t)\|_{L^\infty(\Omega)} & \text{in } Q, \\ z(x,0) = |y_0(x)|, & x \in \Omega, \\ z = 0, & \text{on } \Sigma. \end{cases} \tag{5.19}$$

Subtracting these two equations and multiplying the resulting equation by $(y-z)^+$, and integrating on Ω we get

$$\begin{aligned}&\frac{1}{2}\frac{d}{dt}\|(y-z)^+\|_{L^2(\Omega)}^2 + \int_\Omega |\nabla(y-z)^+|^2dx \le 0, && \text{a.e. } t \in (0,T), \\ &(y-z)^+(0) \le 0 && \text{in } \Omega,\end{aligned}$$

because $z \ge 0$ and β is monotonically increasing. Hence, $y(x,t) \le z(x,t)$, a.e. in Q and so $|y(x,t)| \le z(x,t)$, a.e. $(x,t) \in Q$. On the other hand, the solution z to problem (5.19) can be represented as

$$z(x,t) = S(t)(|y_0|)(x) + \int_0^t S(t-s)(\|f(s)\|_{L^\infty(\Omega)})ds, \quad \text{a.e. } (x,t) \in Q,$$

where $S(t)$ is the semigroup generated on $L^1(\Omega)$ by $-\Delta$ with Dirichlet homogeneous conditions on $\partial\Omega$. We know, by the regularity theory of $S(t)$ (see also Theorem 5.4 below), that

$$\|S(t)u_0\|_{L^\infty(\Omega)} \le Ct^{-(N/2)}\|u_0\|_{L^1(\Omega)}, \qquad \forall u_0 \in L^1(\Omega),\ t>0.$$

Hence,

$$|y(x,t)| \le Ct^{-(N/2)}\|y_0\|_{L^1(\Omega)} + \int_0^t \|f(s)\|_{L^\infty(\Omega)}ds, \qquad (t,x)\in Q. \tag{5.20}$$

Now, for an arbitrary $y_0 \in L^1(\Omega)$ such that $y_0 \in \overline{D(\beta)}$, a.e. in Ω, we choose a sequence $\{y_0^n\} \subset H_0^1(\Omega)\cap H^2(\Omega)$, $y_0^n \in \overline{D(\beta)}$, a.e. in Q, such that $y_0^n \to y_0$ in $L^1(\Omega)$ as $n\to\infty$. (We may take, for instance, $y_0^n = S(n^{-1})(1+n^{-1}\beta)^{-1}y_0$.) If y_n is the corresponding solution to problem (5.1), then we know that $y_n \to y$ strongly in $C([0,T];L^1(\Omega))$, where y is the solution with the initial value y_0. By (5.20), it follows that y satisfies estimate (5.16).

Because $y(t)\in L^\infty(\Omega)\subset L^2(\Omega)$ for all $t>0$, it follows by Proposition 5.1 that $y\in W^{1,2}([\delta,T];L^2(\Omega))\cap L^2(\delta,T;H_0^1(\Omega)\cap H^2(\Omega))$ for all $0<\delta<T$ and it satisfies equation (5.18), a.e. in $Q=\Omega\times(0,T)$. (Arguing as before, we may assume that $y_0\in H_0^1(\Omega)\cap H^2(\Omega)$ and so $y_t, y\in L^2(0,T;L^2(\Omega))$.) To get the desired estimate (5.17), we multiply equation (5.18) by $y_t t^{k+2}$ and integrate on Q to get

$$\begin{aligned}\int_0^T\int_\Omega t^{k+2}y_t^2\,dx\,dt + \frac{1}{2}\int_0^T\int_\Omega t^{k+2}|\nabla y|_t^2\,dx\,dt + \int_0^T\int_\Omega t^{k+2}\frac{\partial}{\partial t}g(y)dx\,dt\\ = \int_0^T\int_\Omega t^{k+2}y_t f\,dx\,dt,\end{aligned}$$

where $y_t = \partial y/\partial t$ and $\partial g = \beta$. This yields

$$\begin{aligned}\int_Q t^{k+2}y_t^2dx\,dt + \frac{T^{k+2}}{2}\int_\Omega |\nabla y(x,T)|^2dx + T^{k+2}\int_\Omega g(y(x,T))dx\\ \le \frac{k+2}{2}\int_Q t^{k+1}|\nabla y|^2dx\,dt + (k+2)\int_Q t^{k+1}g(y)dx\,dt\\ + \frac{1}{2}\int_0^T t^{k+2}y_t^2dx\,dt + \frac{1}{2}\int_Q t^{k+2}f^2dx\,dt.\end{aligned}$$

Hence,

$$\begin{aligned}&\int_Q t^{k+2}y_t^2dx\,dt + T^{k+2}\int_\Omega |\nabla y(x,T)|^2dx\\ &\le (k+2)\int_Q t^{k+1}|\nabla y|^2dx\,dt + 2(k+2)\int_Q t^{k+1}\beta(y)dx + T^{k+2}\int_Q f^2dx\,dt.\end{aligned}$$

(If β is multivalued, then $\beta(y)$ is of course the section of $\beta(y)$ arising in (5.18).)

Finally, writing $\beta(y)y$ as $(f+\Delta y - y_t)y$ and using Green's formula, we get

$$\begin{aligned}
&\int_Q t^{k+2} y_t^2 dx\,dt + T^{k+2}\int_\Omega |\nabla y(x,T)|^2 dx + \int_Q t^{k+1}|\nabla y|^2 dx\,dt \\
&\le (k+2)(k+1)\int_Q y^2 t^k dx\,dt \\
&\quad + T^{k+2}\int_Q f^2 dx\,dt + 2(k+2)\int_Q t^{k+1}|f|\,|y| dx\,dt \\
&\le C\left(\int_Q t^k y^2 dx\,dt + T^{k+2}\int_Q f^2 dx\,dt\right).
\end{aligned} \tag{5.21}$$

Next, we have, by the Hölder inequality

$$\int_\Omega y^2 dx \le \|y\|_{L^p(\Omega)}^{(N-2/N+2)} \|y\|_{L^1(\Omega)}^{4/(N+2)}$$

for $p = 2N(N-2)^{-1}$. Then, by the Sobolev embedding theorem,

$$\int_\Omega |y(x,t)|^2 dx \le \left(\int_\Omega |\nabla y(x,t)|^2 dx\right)^{N/(N+2)} \left(\int_\Omega |y(x,t)| dx\right)^{4/(N+2)}. \tag{5.22}$$

On the other hand, multiplying equation (5.18) by sign y and integrating on $\Omega \times (0,t)$, we get

$$\|y(t)\|_{L^1(\Omega)} \le \|y_0\|_{L^1(\Omega)} + \int_0^t \int_\Omega |f(x,s)| dx\,ds, \qquad t \ge 0,$$

because, as seen earlier (Section 3.2),

$$\int_\Omega \Delta y \operatorname{sign} y\, dx \le 0.$$

Then, by estimates (5.21) and (5.22), we get

$$\begin{aligned}
&\int_Q t^{k+2} y_t^2 dx\,dt + T^{k+2}\int_\Omega |\nabla y(x,T)|^2 dx + \int_Q t^{k+1}|\nabla y(x,t)|^2 dx\,dt \\
&\le C\left(\left(\|y_0\|_{L^1(\Omega)}^{4/(N+2)} + \int_0^T \int_\Omega |f(x,t)| dx\,dt\right)\right. \\
&\quad \left.\times \int_0^t t^k \|\nabla y(t)\|_{L^2(\Omega)}^{2N/(N+2)} dt + T^{k+2}\int_Q f^2 dx\,dt\right).
\end{aligned}$$

On the other hand, we have, for $k = N/2$,

$$\int_0^T t^k |\nabla y(t)|^{2N/(N+2)} dt \le \left(\int_0^T t^{k+1}|\nabla y(t)|^2 dt\right)^{N/(N+2)} T^{2/(N+2)}.$$

Substituting in the latter inequality, we get after some calculation involving the Hölder inequality

$$\begin{aligned}\int_Q t^{(N+4)/2} y_t^2 \, dx\, dt &+ \int_Q t^{(N+2)/2} |\nabla y(x,t)|^2 dx\, dt \\ &+ T^{(N+4)/2} \int_\Omega |\nabla y(x,T)|^2 dx \\ &\le C_1 \left(\|y_0\|_{L^1(\Omega)}^{4/(N+2)} + \int_Q |f(x,t)| dx\, dt \right)^{(N+2)/2} \\ &+ C_2 T^{(N+4)/2} \int_Q f^2(x,t) dx\, dt,\end{aligned} \tag{5.23}$$

as claimed. □

In particular, it follows by Proposition 5.4 that the semigroup $S(t)$ generated by A (defined by (5.11) and (5.13) on $L^1(\Omega)$ has a smoothing effect on initial data; that is, for all $t > 0$ it maps $L^1(\Omega)$ into $D(A)$ and is differentiable on $(0, \infty)$.

In the special case where

$$\beta(r) = \begin{cases} 0 & \text{if } r > 0, \\ \mathbf{R}^- & \text{if } r = 0, \end{cases}$$

problem (5.1) reduces to the parabolic variational inequality (the obstacle problem)

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y = f & \text{in } \{(x,t);\ y(x,t) > 0\}, \\ y \ge 0,\ \dfrac{\partial y}{\partial t} - \Delta y \ge f & \text{in } Q, \\ y(x,0) = y_0(x) & \text{in } \Omega,\ \ y = 0 \quad \text{on } \partial\Omega \times (0,T) = \Sigma. \end{cases} \tag{5.24}$$

This is a problem with free (moving) boundary that is discussed in detail in the next section.

We also point out that Proposition 5.1 remains true for equations of the form

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y + \beta(x,y) \ni f & \text{in } Q, \\ y(x,0) = y_0(x) & \text{in } \Omega, \\ y = 0 & \text{on } \Sigma, \end{cases}$$

where $\beta : \Omega \times \mathbf{R} \to 2^{\mathbf{R}}$ is of the form $\beta(x,y) = \partial_y g(x,y)$ and $g : \Omega \times \mathbf{R} \to \mathbf{R}$ is a normal convex integrand on $\Omega \times \mathbf{R}$ sufficiently regular in X and with appropriate polynomial growth with respect to y. The details are left to the reader.

Now, we consider the equation

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y = f & \text{in } \Omega \times (0,T) = Q, \\ \dfrac{\partial}{\partial \nu} y + \beta(y) \ni 0 & \text{on } \Sigma, \\ y(x,0) = y_0(x) & \text{in } \Omega, \end{cases} \tag{5.25}$$

where $\beta \subset \mathbf{R} \times \mathbf{R}$ is a maximal monotone graph, $0 \in D(\beta)$, $y_0 \in L^2(\Omega)$, and $f \in L^2(Q)$. As seen earlier (Proposition 2.9), we may write (5.25) as

$$\begin{cases} \dfrac{dy}{dt}(t) + Ay(t) = f(t) & \text{in } (0,T), \\ y(0) = y_0, \end{cases}$$

where $Ay = -\Delta y$, $\forall y \in D(A) = \{y \in H^2(\Omega);\ 0 \in \partial y/\partial \nu + \beta(y), \text{ a.e. on } \partial\Omega\}$.

More precisely, $A = \partial\varphi$, where $\varphi : L^2(\Omega) \to \overline{\mathbf{R}}$ is defined by

$$\varphi(y) = \frac{1}{2}\int_\Omega |\nabla y|^2 dx + \int_{\partial\Omega} j(y)d\sigma, \qquad \forall y \in L^2(\Omega),$$

and $\partial j = \beta$.

Then, applying Theorems 4.11 and 4.12, we get the following.

Proposition 5.5. *Let $y_0 \in \overline{D(A)}$ and $f \in L^2(Q)$. Then, problem* (5.25) *has a unique strong solution $y \in C([0,T];L^2(\Omega))$ such that*

$$t^{1/2}\,\frac{dy}{dt} \in L^2(0,T;L^2(\Omega)),$$
$$t^{1/2} y \in L^2(0,T;H^2(\Omega)).$$

If $y_0 \in H^1(\Omega)$ and $j(y_0) \in L^1(\Omega)$, then

$$\frac{dy}{dt} \in L^2(0,T;L^2(\Omega)),$$
$$y \in L^2(0,T;H^2(\Omega)) \cap L^\infty(0,T;H^1(\Omega)).$$

Finally, if $y_0 \in D(A)$ and $f, \partial f/\partial t \in L^2(\Omega)$, then

$$\frac{dy}{dt} \in L^\infty(0,T;L^2(\Omega)),$$
$$y \in L^\infty(0,T;H^2(\Omega))$$

and

$$\frac{d^+}{dt} y(t) - \Delta y(t) = f(t), \qquad \forall t \in [0,T].$$

It should be mentioned that one uses here the estimate (see (2.65))

$$\|u\|_{H^2(\Omega)} \le C(\|u - \Delta u\|_{L^2(\Omega)} + 1), \qquad \forall u \in D(A).$$

An important special case is

$$\beta(r) = \begin{cases} 0 & \text{if } r > 0, \\ (-\infty, 0] & \text{if } r = 0. \end{cases}$$

Then, problem (5.25) reads as

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y = f & \text{in } Q, \\ y\,\dfrac{\partial y}{\partial \nu} = 0,\ y \geq 0,\ \dfrac{\partial y}{\partial \nu} \geq 0 & \text{on } \Sigma, \\ y(x,0) = y_0(x) & \text{in } \Omega. \end{cases} \tag{5.26}$$

A problem of this type arises in the control of a heat field. More generally, the thermostat control process is modeled by equation (5.26), where

$$\beta(r) = \begin{cases} a_1(r-\theta_1) & \text{if } -\infty < r < \theta_1, \\ 0 & \text{if } \theta_1 \leq r \leq \theta_2, \\ a_2(r-\theta_2) & \text{if } \theta_2 < r < \infty, \end{cases}$$

$a_i \geq 0,\ \theta_1 \in \mathbf{R},\ i = 1,2$. In the limit case, we obtain (5.26).

The black body radiation heat emission on $\partial\Omega$ is described by equation (5.26), where β is given by (the Stefan–Boltzman law)

$$\beta(r) = \begin{cases} \alpha(r^4 - y_1^4) & \text{for } r \geq 0, \\ -\alpha y_1^4 & \text{for } r < 0, \end{cases}$$

and, in the case of natural convection heat transfer,

$$\beta(r) = \begin{cases} ar^{5/4} & \text{for } r \geq 0, \\ 0 & \text{for } r < 0. \end{cases}$$

Note, also, that the Michaelis–Menten dynamic model of enzyme diffusion reaction is described by equation (5.1) (or (5.25)), where

$$\beta(r) = \begin{cases} \dfrac{r}{\lambda(r+k)} & \text{for } r > 0, \\ (-\infty, 0] & \text{for } r = 0, \\ \emptyset & \text{for } r < 0, \end{cases}$$

where λ, k are positive constants.

We note that more general boundary value problems of the form

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y + \gamma(y) \ni f & \text{in } Q, \\ y(x,0) = y_0(x) & \text{in } \Omega, \\ \dfrac{\partial y}{\partial \nu} + \beta(y) \ni 0 & \text{on } \Sigma, \end{cases}$$

where β and γ are maximal monotone graphs in $\mathbf{R}\times\mathbf{R}$ such that $0\in D(\beta)$, $0\in D(\gamma)$ can be written in the form (5.2) where $A=\partial\varphi$ and $\varphi: L^2(\Omega)\to\overline{\mathbf{R}}$ is defined by

$$\varphi(y)=\begin{cases}\frac{1}{2}\int_\Omega |\nabla y|^2 dx+\int_\Omega g(y)dx+\int_{\partial\Omega} j(y)d\sigma & \text{if } y\in H^1(\Omega),\\ +\infty & \text{otherwise,}\end{cases}$$

and $\partial g=\gamma$, $\partial j=\beta$.

We may conclude, therefore, that for $f\in L^2(\Omega)$ and $y_0\in H^1(\Omega)$ such that $g(y_0)\in L^1(\Omega)$, $j(y_0)\in L^1(\partial\Omega)$ the preceding problem has a unique solution $y\in W^{1,2}([0,T];L^2(\Omega))\cap L^2(0,T;H^2(\Omega))$.

On the other hand, semilinear parabolic problems of the form (5.1) or (5.25) arise very often as feedback systems associated with the linear heat equation. For instance, the feedback relay control

$$u=-\rho\ \text{sign}\ y, \tag{5.27}$$

where

$$\text{sign}\, r=\begin{cases}\frac{r}{|r|} & \text{if } r\neq 0,\\ [-1,1] & \text{if } r=0,\end{cases}$$

applied to the controlled heat equation

$$\begin{cases}\frac{\partial y}{\partial t}-\Delta y=u & \text{in } \Omega\times\mathbf{R}^+,\\ y=0 & \text{on } \partial\Omega\times\mathbf{R}^+,\\ y(x,0)=y_0(x) & \text{in } \Omega\end{cases} \tag{5.28}$$

transforms it into a nonlinear equation of the form (5.1); that is,

$$\begin{cases}\frac{\partial y}{\partial t}-\Delta y+\rho\ \text{sign}\, y\ni 0 & \text{in } \Omega\times\mathbf{R}^+,\\ y=0 & \text{on } \partial\Omega\times\mathbf{R}^+,\\ y(x,0)=y_0(x) & \text{in } \Omega.\end{cases} \tag{5.29}$$

This is the closed-loop system associated with the feedback law (5.27) and, according to Proposition 5.4, for every $y_0\in L^1(\Omega)$, it has a unique strong solution $y\in C(\mathbf{R}^+;L^2(\Omega))$ satisfying

$$y(t)\in L^\infty(\Omega), \qquad \forall t>0,$$
$$t^{(N+4)/4}y_t\in L^2_{\text{loc}}(\mathbf{R}^+;L^2(\Omega)), \qquad t^{(N+2)/4}y\in L^2_{\text{loc}}(\mathbf{R}^+;H^1(\Omega)).$$

(Of course, if $y_0\in L^2(\Omega)$, then y has sharper properties provided by Proposition 5.1.)

Let us observe that the feedback control (5.27) belongs to the constraint set $\{u \in L^\infty(\Omega \times \mathbf{R}^+); \|u\|_{L^\infty(\Omega\times\mathbf{R}^+)} \le \rho\}$ and steers the initial state y_0 into the origin in a finite time T. Here is the argument. We assume first that $y_0 \in L^\infty(\Omega)$ and consider the function $w(x,t) = \|y_0\|_{L^\infty(\Omega)} - \rho t$. On the domain $\Omega \times (0, \rho^{-1}\|y_0\|_{L^\infty(\Omega)}) = Q_0$, we have

$$\begin{cases} \dfrac{\partial w}{\partial t} - \Delta w + \rho \text{ sign } w \ni 0 & \text{in } Q_0, \\ w(0) = \|y_0\|_{L^\infty(\Omega)} & \text{in } \Omega, \\ w \ge 0 & \text{on } \partial\Omega \times (0, \rho^{-1}\|y_0\|_{L^\infty(\Omega)}). \end{cases} \tag{5.30}$$

Then, subtracting equations (5.29) and (5.30) and multiplying by $(y-w)^+$ (or, simply, applying the maximum principle), we get

$$(y-w)^+ \le 0 \quad \text{in } Q_0.$$

Hence, $y \le w$ in Q_0. Similarly, it follows that $y \ge -w$ in Q_0 and, therefore,

$$|y(x,t)| \le \|y_0\|_{L^\infty(\Omega)} - \rho t, \qquad \forall (x,t) \in Q_0.$$

Hence, $y(t) \equiv 0$ for all $t \ge T = \rho^{-1}\|y_0\|_{L^\infty(\Omega)}$. Now, if $y_0 \in L^1(\Omega)$, then inserting in system (5.28) the feedback control

$$u(t) = \begin{cases} 0 & \text{for } 0 \le t \le \varepsilon, \\ -\rho \text{ sign } y(t) & \text{for } t > \varepsilon, \end{cases}$$

we get a trajectory $y(t)$ that steers y_0 into the origin in the time

$$T(y_0) < \varepsilon + \rho^{-1}\|y(\varepsilon)\|_{L^\infty(\Omega)} \le \varepsilon + C(\rho\varepsilon^{N/2})^{-1}\|y_0\|_{L^1(\Omega)},$$

where C is independent of ε and y_0 (see estimate (5.16)). If we choose $\varepsilon > 0$ that minimizes the right-hand side of the latter inequality, then we get

$$T(y_0) \le \left(\frac{CN}{2\rho}\|y_0\|_{L^1(\Omega)}\right)^{2/(N+2)} + \left(\frac{N}{2}\right)^{-(N/(N+2))} \left(\frac{C}{\rho}\|y_0\|_{L^1(\Omega)}\right)^{2/(N+2)}.$$

We have, therefore, proved the following null controllability result for system (5.28).

Proposition 5.6. *For any $y_0 \in L^1(\Omega)$ and $\rho > 0$ there is $u \in L^\infty(\Omega \times \mathbf{R}^+)$, $\|u\|_{L^\infty(\Omega\times\mathbf{R}^+)} < \rho$, that steers y_0 into the origin in a finite time $T(y_0)$.*

Remark 5.1. Consider the nonlinear parabolic equation

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y + |y|^{p-1}y = 0, & \text{in } \Omega \times \mathbf{R}^+, \\ y(x,0) = y_0(x), & x \in \Omega, \\ y = 0, & \text{on } \partial\Omega \times \mathbf{R}^+, \end{cases} \tag{5.31}$$

where $0 < p < (N+1)/N$ and $y_0 \in L^1(\Omega)$. By Proposition 5.4, we know that the solution y satisfies the estimates

$$\|y(t)\|_{L^\infty(\Omega)} \leq Ct^{-(N/2)}|y_0\|_{L^1(\Omega)},$$
$$\|y(t)\|_{L^1(\Omega)} \leq C\|y_0\|_{L^1(\Omega)},$$

for all $t > 0$.

Now, if y_0 is a bounded Radon measure on Ω; that is, $y_0 \in M(\Omega) = (C_0(\overline{\Omega}))^*$ ($C_0(\overline{\Omega})$ is the space of continuous functions on $\overline{\Omega}$ that vanish on $\partial\Omega$), there is a sequence $\{y_0^j\} \subset C_0(\Omega)$ such that $\|y_0^j\|_{L^1(\Omega)} \leq C$ and $y_0^j \to y_0$ weak-star in $M(\Omega)$. Then, if y^j is the corresponding solution to equation (5.31) it follows from the previous estimates that (see Brezis and Friedman [17])

$$y^j \to y \qquad \text{in } L^q(Q), \qquad 1 < q < \frac{N+2}{N},$$
$$|y^j|^{p-1}y^j \to |y|^{p-1}y \quad \text{in } L^1(Q).$$

This implies that y is a generalized (mild) solution to equation (5.31).

If $p > (N+2)/N$, there is no solution to (5.31).

Remark 5.2. Consider the semilinear parabolic equation (5.1), where β is a continuous monotonically increasing function, $f \in L^p(Q)$, $p > 1$, and $y_0 \in W_0^{p,2-(2/p)}(\Omega)$, $g(y_0) \in L^1(\Omega)$, $g(r) = \int_0^r |\beta(s)|^{p-2}\beta(s)ds$. Then, the solution y to problem (5.1) belongs to $W_p^{2,1}(Q)$ and

$$\|y\|^p_{W_p^{2,1}(Q)} \leq C\left(\|f\|^p_{L^p(\Omega)} + \|y_0\|^p_{W_0^{p,2-(2/p)}(\Omega)} + \int_\Omega g(y_0)dx\right).$$

Here, $W_p^{2,1}(Q)$ is the space

$$\left\{y \in L^p(Q);\ \frac{\partial^{r+s}}{\partial t^r \partial x^s} y \in L^p(Q), 2r+s \leq 2\right\}.$$

For $p = 2$, $W_2^{2,1}(Q) = H^{2,1}(Q)$.

Indeed, if we multiply equation (5.1) by $|\beta(y)|^{p-2}\beta(y)$ we get the estimate (as seen earlier in Proposition 5.1, for f and y_0 smooth enough this problem has a unique solution $y \in W^{1,\infty}([0,T];L^p(\Omega))$, $y \in L^\infty(0,T;W^{2,p}(\Omega))$)

$$\begin{aligned}
&\int_\Omega g(y(x,t))dx + \int_0^t \int_\Omega |\beta(y(x,s))|^p dxds \\
&\leq \int_0^t \int_\Omega |\beta(y(x,s))|^{p-1}|f(x,s)|dxds + \int_\Omega g(y_0(x))dx \\
&\leq \left(\int_0^t \int_\Omega |\beta(y(x,s))|^p dxds\right)^{1/q} \left(\int_0^t \int_\Omega |f(x,s)|^p dxds\right)^{1/p},
\end{aligned}$$

where $1/p + 1/q = 1$. In particular, this implies that

$$\|\beta(y)\|_{L^p(Q)} \leq C(\|f\|_{L^p(Q)} + \|g(y_0)\|_{L^1(\Omega)})$$

and by the L^p estimates for linear parabolic equations (see, e.g., Ladyzenskaya, Solonnikov, and Ural'ceva [31] and Friedman [27]) we find the estimate (5.34), which clearly extends to all $f \in L^p(Q)$ and $y_0 \in W_0^{p,2-(2/p)}(\Omega)$, $g(y_0) \in L^1(\Omega)$.

Nonlinear Parabolic Equations of Divergence Type

Several physical diffusion processes are described by the continuity equation

$$\frac{\partial y}{\partial t} + \mathrm{div}_x\, \mathbf{q} = f,$$

where the flux $\mathbf{q}$ of the diffusive material is a nonlinear function β of local density gradient ∇y. Such an equation models nonlinear interaction phenomena in material science and in particular in mathematical models of crystal growth as well as in image processing (see Section 2.4). This class of problems can be written as

$$\begin{cases} \dfrac{\partial y}{\partial t}(x,t) - \mathrm{div}_x\, \beta(\nabla(y(x,t))) \ni f(x,t), & x \in \Omega,\ t \in (0,T), \\ y = 0 & \text{on } \partial\Omega \times (0,T), \\ y(x,0) = y_0(x), & x \in \Omega, \end{cases} \tag{5.32}$$

where $\beta : \mathbf{R}^N \to \mathbf{R}^N$ is a maximal monotone graph satisfying conditions (2.138) and (2.139) (or, in particular, conditions (2.134) and (2.135) of Theorem 2.15).

In the space $X = L^2(\Omega)$ consider the operator A defined by (2.155) and thus represent (5.32) as a Cauchy problem in X; that is,

$$\begin{cases} \dfrac{dy}{dt}(t) + Ay(t) \ni f(t), \ t \in (0,T), \\ y(0) = y_0. \end{cases} \tag{5.33}$$

In Section 2.4, we studied in detail the stationary version of (5.37) (i.e., $Ay = f$) and we have proven (Theorem 2.18) that A is maximal monotone (m-accretive) and so, by Theorem 4.6, we obtain the following existence result.

Proposition 5.7. *Let* $f \in W^{1,1}([0,T];L^2(\Omega))$, $y_0 \in W_0^{1,p}(\Omega)$ *be such that* $\mathrm{div}\, \eta_0 \in L^2(\Omega)$ *for some* $\eta_0 \in (L^q(\Omega))^N$, $\eta_0 \in \beta(\nabla y_0)$, *a.e. in* Ω. *Then, there is a unique strong solution* y *to* (5.32) (*equivalently to* (5.33)) *such that*

$$y \in L^\infty(0,T;W_0^{1,p}(\Omega)) \cap W^{1,\infty}([0,T];L^2(\Omega))$$
$$\frac{d^+}{dt} y(t) - \mathrm{div}_x\, \eta(t) = f(t), \qquad \forall t \in [0,T],$$

where $\eta \in L^\infty(0,T;L^2(\Omega))$, $\eta(t,x) \in \beta(\nabla y(x,t))$, *a.e.* $(x,t) \in \Omega \times (0,T) = Q$. *Moreover, if* $\beta = \partial j$, *then the strong solution* y *exists for all* $y_0 \in L^2(\Omega)$ *and* $f \in L^2(Q)$.

The last part of Proposition 5.7 follows by Theorem 4.11, because, as seen earlier in Theorem 2.18, in this latter case $A = \partial\varphi$.

Now, if we refer to Theorem 2.19 and Remark 2.4 we may infer that Proposition 5.7 remains true under conditions $\beta = \partial j$ and (2.161) and (2.162). We have, therefore, the following.

Proposition 5.8. *Let* β *satisfy conditions* (2.161) and (2.162). *Then, for each* $y_0 \in L^2(\Omega)$ *and* $f \in L^2(0,T;L^2(\Omega))$ *there is a unique strong solution to* (5.32) *or to the equation with Neumann boundary conditions* $\beta(\nabla y(x)) \cdot \nu(x) = 0$ *in the following weak sense,*

$$\begin{aligned}
&\frac{d}{dt}\int_\Omega y(x,t)v(x)dx + \int_\Omega \eta(x,t)\cdot\nabla v(x)dx = \int_\Omega f(x,t)v(x)dx, \qquad \forall v \in C^1(\overline{\Omega}),\\
&\eta(x,t) \in \beta(\nabla y(x,t)), \quad \text{a.e. } (x,t) \in \Omega \times (0,T),\\
&y(x,0) = y_0(x).
\end{aligned}$$

Now, if we refer to the singular diffusion boundary value problem

$$\begin{cases}
\dfrac{\partial y}{\partial t} - \operatorname{div}_x(\operatorname{sign}(\nabla y)) \ni f & \text{in } \Omega \times (0,T),\\
y = 0 & \text{on } \partial\Omega \times (0,T),\\
y(x,0) = y_0(x),
\end{cases}$$

it has for each $y_0 \in BV^0(\Omega)$ a unique strong solution

$$y \in W^{1,2}([0,T];L^2(\Omega)) \cap C([0,T];L^2(\Omega))$$

with $\|Dy(t)\| \in W^{1,\infty}([0,T])$ (similarly for the case of Neumann boundary conditions).

Indeed, as seen earlier, it can be written as a first-order equation of subgradient type in $L^2(\Omega)$,

$$\begin{cases}
\dfrac{dy}{dt}(t) + \partial\varphi(y(t)) \ni f(t), \quad t \in (0,T),\\
y(0) = y_0,
\end{cases}$$

where φ is given by (2.182). Then, the existence follows by Theorem 4.11.

By (2.149) and the Trotter–Kato theorem (see Theorem 4.14), we know that the solution y is the limit in $C([0,T];L^2(\Omega))$ of solution y_ε to the problem

$$\begin{cases}
\dfrac{\partial y_\varepsilon}{\partial t} - \varepsilon\Delta y_\varepsilon - \operatorname{div}_x\beta_\varepsilon(\nabla y_\varepsilon) = f \quad \text{in } \Omega \times (0,T)\\
y_\varepsilon = 0 \quad \text{on } \partial\Omega; \qquad y_\varepsilon(x,0) = y_0(x),
\end{cases}$$

where β_ε is the Yosida approximation of $\beta = \mathrm{sign}$.

As noticed earlier, this equation is relevant in image restoration techniques and crystal-faceted growth theory. In particular, for $f(t) \equiv f_e \in L^2(\Omega)$ it follows by Theorem 4.13 that

$$\lim_{t\to\infty} y(t) = y_e \quad \text{strongly in } L^2(\Omega),$$

where y_e is an equilibrium solution; that is, $\partial\varphi(y_e) \ni f_e$.

In image processing, the solution $y = y(\cdot,t)$ might be seen as a family of restored images with the scale parameter t. The parabolic equation (5.32) itself acts as a filter that processes the original corrupted version $f = f(x)$.

Semilinear Parabolic Equation in $\mathbf{R}^N$

We consider here equation (5.1) in $\Omega = \mathbf{R}^N$; that is,

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y + \beta(y) \ni f & \text{in } (0,T)\times\mathbf{R}^N, \\ y(0,x) = y_0(x) & x \in \mathbf{R}^N, \\ y(t,\cdot) \in L^1(\mathbf{R}^N) & \forall t \in (0,T). \end{cases} \tag{5.34}$$

With respect to the case of bounded domain Ω previously studied, this problem presents some peculiarities and the more convenient functional space to study it is $L^1(\mathbf{R}^N)$.

We write (5.34) as a differential equation in $X = L^1(\mathbf{R}^N)$ of the form

$$\begin{cases} \dfrac{dy}{dt}(t) + Ay(t) \ni f(t), & t \in (0,T), \\ y(0) = y_0, \end{cases}$$

where $A : D(A) \subset L^1(\mathbf{R}^N) \to \mathbf{R}^N$ is defined by

$$\begin{aligned} Ay &= \{z \in L^1(\mathbf{R}^N);\ z = -\Delta y + w,\ w \in \beta(y),\ \text{a.e. in } \mathbf{R}^N\}, \\ D(A) &= \{y \in L^1(\mathbf{R}^N);\ \Delta y \in L^1(\mathbf{R}^N),\ \exists w \in L^1(\mathbf{R}^N), \\ &\qquad \text{such that } w(x) \in \beta(y(x)),\ \text{a.e. } x \in \mathbf{R}^N\}. \end{aligned}$$

By Theorem 3.3 we know that, if $N = 1,2,3$, then A is m-accretive in $L^1(\mathbf{R}^N) \times L^1(\mathbf{R}^N)$.

Then, by Theorem 4.1, which neatly applies to this situation, we get the following existence result.

Proposition 5.9. *Let $y_0 \in L^1(\mathbf{R}^N)$ and $f \in L^1(0,T;\mathbf{R}^N)$ be such that $\Delta y_0 \in L^1(\mathbf{R}^N)$ and $\exists w \in L^1(\mathbf{R}^N)$, $w(x) \in \beta(y_0(x))$, a.e. $x \in \mathbf{R}^N$. Then, problem* (5.34) *has a unique mild solution $y \in C([0,T];L^1(\mathbf{R}^N))$. In other words,*

$$y(t) = \lim_{\varepsilon \to 0} y_\varepsilon(t) \quad \textit{strongly in } L^1(\mathbf{R}^n) \textit{ for each } t \in [0,T], \tag{5.35}$$

where y_ε is the solution to the finite difference scheme

$$\begin{aligned} & y_\varepsilon(t) = y_\varepsilon^i \quad \textit{for } t \in (i\varepsilon, (i+1)\varepsilon), \qquad i = 0,1,\dots,M, \\ & y_\varepsilon^{i+1} - y_\varepsilon^i - \varepsilon \Delta y_\varepsilon^{i+1} + \varepsilon \beta(y_\varepsilon^{i+1}) \ni \int_{i\varepsilon}^{(i+1)\varepsilon} f(t)dt \quad \textit{in } \mathbf{R}^n, \\ & y_\varepsilon^i \in L^1(\mathbf{R}^N), \qquad i = 0,1,\dots,M = \left[\tfrac{T}{\varepsilon}\right]. \end{aligned} \tag{5.36}$$

5.2 Parabolic Variational Inequalities

An important class of multivalued nonlinear parabolic-like boundary value problem is the so-called parabolic variational inequalities which we briefly present below in an abstract setting.

Here and throughout in the sequel, V and H are real Hilbert spaces such that V is dense in H and $V \subset H \subset V'$ algebraically and topologically. We denote by $|\cdot|$ and $\|\cdot\|$ the norms of H and V, respectively, and by $(\cdot,\cdot)$ the scalar product in H and the pairing between V and its dual V'. The norm of V' is denoted $\|\cdot\|_*$. The space H is identified with its own dual.

We are given a linear continuous and symmetric operator A from V to V' satisfying the coercivity condition

$$(Ay,y) + \alpha |y|^2 \geq \omega \|y\|^2, \qquad \forall y \in V, \tag{5.37}$$

for some $\omega > 0$ and $\alpha \in \mathbf{R}$. We are also given a lower semicontinuous convex function $\varphi : V \to \overline{\mathbf{R}} = (-\infty, +\infty]$, $\varphi \not\equiv +\infty$.

For $y_0 \in V$ and $f \in L^2(0,T;V')$, consider the following problem.

Find $y \in L^2(0,T;V) \cap C([0,T];H) \cap W^{1,2}([0,T];V')$ such that

$$\begin{cases} (y'(t) + Ay(t), y(t) - z) + \varphi(y(t)) - \varphi(z) \leq (f(t), y(t) - z), \\ \hfill \text{a.e. } t \in (0,T),\ \forall z \in V, \\ y(0) = y_0. \end{cases} \tag{5.38}$$

Here, $y' = dy/dt$ is the strong derivative of the function $y : [0,T] \to V'$. In terms of the subgradient mapping $\partial\varphi : V \to V'$, problem (5.38) can be written as

$$\begin{cases} y'(t) + Ay(t) + \partial\varphi(y(t)) \ni f(t), \quad \text{a.e. } t \in (0,T), \\ y(0) = y_0. \end{cases} \tag{5.39}$$

This is an abstract variational inequality of parabolic type. In applications to partial differential equations, V is a Sobolev subspace of $H = L^2(\Omega)$ (Ω is an open subset of $\mathbf{R}^N$), A is an elliptic operator on Ω, and the unknown function $y : \Omega \times [0,T] \to \mathbf{R}$ is viewed as a function of t from $[0,T]$ to $L^2(\Omega)$.

As seen earlier in Section 4.1, in the special case where $\varphi = I_K$ is the indicator function of a closed convex subset K of V; that is,

$$\varphi(y) = 0 \quad \text{if } y \in K, \qquad \varphi(y) = +\infty \quad \text{if } y \notin K, \tag{5.40}$$

the variational inequality (5.38) reduces to the reflection problem

$$\begin{cases} y(t) \in K, & \forall t \in [0,T], \\ (y'(t) + Ay(t), y(t) - z) \le (f(t), y(t) - z), & \text{a.e. } t \in (0,T), \forall z \in K, \\ y(0) = y_0. \end{cases} \tag{5.41}$$

Regarding the existence for problem (5.38), we have the following.

Theorem 5.1. *Let $f \in W^{1,2}([0,T];V')$ and $y_0 \in V$ be such that*

$$\{Ay_0 + \partial\varphi(y_0) - f(0)\} \cap H \neq \emptyset. \tag{5.42}$$

Then, problem (5.38) *has a unique solution $y \in W^{1,2}([0,T];V) \cap W^{1,\infty}([0,T];H)$ and the map $(y_0, f) \to y$ is Lipschitz from $H \times L^2(0,T;V')$ to $C([0,T];H) \cap L^2(0,T;V)$. If $f \in W^{1,2}([0,T];V')$ and $\varphi(y_0) < \infty$, then problem* (5.38) *has a unique solution $y \in W^{1,2}([0,T];H) \cap C_w([0,T];V)$. If $f \in L^2(0,T;H)$ and $\varphi(y_0) < \infty$, then problem* (5.38) *has a unique solution $y \in W^{1,2}([0,T];H) \cap C_w([0,T];V)$, that satisfies*

$$y'(t) = (f(t) - Ay(t) - \partial\varphi(y(t)))^0, \quad \text{a.e. } t \in (0,T).$$

Here $C_w([0,T];V)$ is the space of weakly continuous functions from $(0,T)$ to V; that is, from $(0,T)$ to V endowed with the weak topology.

Proof. Consider the operator $L : D(A) \subset H \to H$,

$$\begin{aligned} Ly &= \{Ay + \partial\varphi(y)\} \cap H, \qquad \forall y \in D(L), \\ D(L) &= \{y \in V;\ \{Ay + \partial\varphi(y)\} \cap H \neq \emptyset\}. \end{aligned}$$

Note that $\alpha I + L$ is maximal monotone in $H \times H$ (I is the identity operator in H). Indeed, by hypothesis (5.37), the operator $\alpha I + A$ is continuous and positive definite from V to V'. Because $\partial\varphi : V \to V'$ is maximal monotone we infer by Theorem 2.6 (or by Corollary 2.6) that $\alpha I + L$ is maximal monotone from V to V' and, consequently, in $H \times H$.

Then, by Theorem 4.6, for every $y_0 \in D(L)$ and $g \in W^{1,1}([0,T];H)$ the Cauchy problem

$$\begin{cases} \dfrac{dy}{dt}(t) + Ly(t) \ni g(t), & \text{a.e. in } (0,T), \\ y(0) = y_0, \end{cases}$$

has a unique strong solution $y \in W^{1,\infty}([0,T];H)$. Let us observe that $\partial\varphi_\alpha = \alpha I + L$, where $\varphi_\alpha : H \to \overline{\mathbf{R}}$ is given by

$$\varphi_\alpha(y) = \frac{1}{2}(Ay + \alpha y, y) + \varphi(y), \qquad \forall y \in H. \tag{5.43}$$

Indeed, φ_α is convex and lower semicontinuous in H because

$$\lim_{\|y\|\to\infty} \frac{\varphi_\alpha(y)}{\|y\|} = \infty$$

and φ_α is lower semicontinuous on V.

On the other hand, it is readily seen that $\alpha I + L \subset \partial\varphi_\alpha$, and because $\alpha I + L$ is maximal monotone, we infer that $\alpha I + L = \partial\varphi_\alpha$, as claimed. In particular, this implies that $\overline{D(L)} = \overline{D(\varphi_\alpha)} = \overline{D(\varphi)}$ (in the topology of H).

Now, let $y_0 \in V$ and $f \in W^{1,2}([0,T];V')$, satisfying condition (5.42).

Let $\{y_0^n\} \subset D(L)$ and $\{f_n\} \subset W^{1,2}([0,T];H)$ be such that

$$\begin{aligned} y_0^n &\to y_0 && \text{strongly in } H, \text{ weakly in } V,\\ f_n &\to f && \text{strongly in } L^2(0,T;V'),\\ \frac{d}{dt}f_n &\to \frac{df}{dt} && \text{strongly in } L^2(0,T;V'). \end{aligned}$$

Let $y_n \in W^{1,\infty}([0,T];H)$ be the corresponding solution to the Cauchy problem

$$\begin{cases} \dfrac{dy_n}{dt}(t) + Ly_n(t) \ni f_n(t), & \text{a.e. in } (0,T),\\ y_n(0) = y_0^n. \end{cases} \tag{5.44}$$

If we multiply (5.44) by $y_n - y_0$ and use condition (5.37), we get

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}|y_n(t) - y_0|^2 + \omega\|y_n(t) - y_0\|^2\\ &\le \alpha|y_n(t) - y_0|^2 + (f_n(t) - \xi, y_n(t) - y_0), \quad \text{a.e. } t \in (0,T), \end{aligned} \tag{5.45}$$

where $\xi \in Ay_0 + \partial\varphi(y_0) \subset V'$. After some calculation involving Gronwall's lemma, this yields

$$|y_n(t) - y_0|^2 + \int_0^t \|y_n(s) - y_0\|^2 ds \le C, \qquad \forall n \in \mathbf{N},\ t \in [0,T]. \tag{5.46}$$

Now, we use the monotonicity of $\partial\varphi$ along with condition (5.37) to get by (5.44) that

$$\frac{1}{2}\frac{d}{dt}|y_n(t)-y_m(t)|^2+\omega\|y_n(t)-y_m(t)\|^2$$
$$\leq \alpha|y_n(t)-y_m(t)|^2+\|f_n(t)-f_m(t)\|_*\|y_n(t)-y_m(t)\|, \quad \text{a.e. } t\in(0,T).$$

Integrating on $(0,t)$, and using Gronwall's lemma, we obtain the inequality

$$|y_n(t)-y_m(t)|^2+\int_0^T\|y_n(t)-y_m(t)\|^2dt$$
$$\leq C\left(|y_0^n-y_0^m|^2+\int_0^t\|f_n(t)-f_m(t)\|^2dt\right).$$

Thus, there is $y\in C([0,T];H)\cap L^2(0,T;V)$ such that

$$y_n\to y \quad \text{in } C([0,T];H)\cap L^2(0,T;V). \tag{5.47}$$

Now, again using equation (5.44), we get

$$\frac{1}{2}\frac{d}{dt}|y_n(t+h)-y_n(t)|^2+\omega\|y_n(t+h)-y_n(t)\|^2$$
$$\leq \alpha|y_n(t+h)-y_n(t)|^2+\|f_n(t+h)-f_n(t)\|_*\|y_n(t+h)-y_n(t)\|,$$

for all $t,h\in(0,T)$ such that $t+h\in(0,T)$. This yields

$$|y_n(t+h)-y_n|^2+\int_0^{T-h}\|y_n(t+h)-y_n(t)\|^2dt$$
$$\leq C\left(|y_n(h)-y_0^n|^2+\int_0^{T-h}\|f_n(t+h)-f_n(t)\|_*^2dt\right)$$

and, letting n tend to $+\infty$,

$$|y(t+h)-y(t)|^2+\int_0^{T-h}\|y(t+h)-y(t)\|^2dt$$
$$\leq C\left(|y(h)-y_0|^2+\int_0^{T-h}\|f(t+h)-f(t)\|_*^2dt\right), \tag{5.48}$$
$$\forall t\in[0,T-h].$$

Next, by (5.45) we see that, if $\xi\in Ay_0+\partial\varphi(y_0)$ is such that $f(0)-\xi\in H$, then we have

$$\frac{1}{2}\frac{d}{dt}|y_n(t)-y_0|^2+\omega\|y_n(t)-y_0\|^2$$
$$\leq \alpha|y_n(t)-y_0|^2+\|f_n(t)-f_n(0)\|_*\|y_n(t)-y_0^n\|+|f_n(0)-\xi|\,|y_n(t)-y_0^n|.$$

Integrating and letting $n\to\infty$, we get by the Gronwall inequality

$$|y(t)-y_0| \le C\left(\int_0^t \|f(s)-f(0)\|_* ds + |f(0)-\xi|t\right), \qquad \forall t \in [0,T].$$

This yields, eventually with a new positive constant C,

$$|y(t)-y_0| \le Ct, \qquad \forall t \in [0,T].$$

Along with (5.48), the latter inequality implies that y is H-valued, absolutely continuous on $[0,T]$, and

$$|y'(t)|^2 + \int_0^t \|y'(t)\|^2 dt \le C\left(|y_0|^2 + \int_0^T \|f'(t)\|_*^2 dt + 1\right), \quad \text{a.e. } t \in (0,T),$$

where $y' = dy/dt,\ f' = df/dt$. Hence, $y \in W^{1,\infty}([0,T];H) \cap W^{1,2}([0,T];V)$.

Let us show now that y satisfies equation (5.38) (equivalently, (5.39)). By (5.44), we have

$$\frac{1}{2}\frac{d}{dt}|y_n(t)-z|^2 \le (f_n(t)-\alpha y_n(t)-\eta, y_n(t)-z), \quad \text{a.e. } t \in (0,T),$$

where $z \in D(L)$ and $\eta \in Lz$. This yields

$$\frac{1}{2}(|y_n(t+\varepsilon)-z|^2 - |y_n(t)-z|^2 \le \int_t^{t+\varepsilon} (f_n(s)+\alpha y_n(s)-\eta, y_n(s)-z))ds$$

and, letting $n \to \infty$,

$$\frac{1}{2}(|y(t+\varepsilon)-z|^2 - |y(t)-z|^2) \le \int_t^{t+\varepsilon} (f(s)+\alpha y(s)-\eta, y(s)-z)\,ds.$$

Finally, this yields

$$(y(t+\varepsilon)-y(t), y(t)-z) \le \int_t^{t+\varepsilon} (f(s)+\alpha y(s)-\eta, y(s)-z)ds.$$

Because y is, a.e., H-differentiable on $(0,T)$, we get

$$(y'(t)-\alpha y(t)+\eta-f(t), y(t)-z) \le 0, \quad \text{a.e. } t \in (0,T),$$

for all $[z,\eta] \in L$. Now, because L is maximal monotone in $H \times H$, we conclude that

$$f(t) \in y'(t) + Ly(t), \quad \text{a.e. } t \in (0,T),$$

as desired.

Now, if $(y_0^i, f_i),\ i=1,2$, satisfy condition (5.42) and the y_i are the corresponding solutions to equation (5.39), by assumption (5.37) it follows that

$$
\begin{aligned}
&|y_1(t)-y_2(t)|^2+\int_0^T \|y_1(t)-y_2(t)\|^2 dt \\
&\le C\left(|y_0^1-y_0^2|^2+\int_0^T \|f_1(t)-f_2(t)\|_*^2 dt\right), \qquad \forall t\in[0,T].
\end{aligned}
$$

Now, assume that $f\in W^{1,2}([0,T];V')$ and $y_0\in D(\varphi)$. Then, as seen earlier, we may rewrite equation (5.39) as

$$
\begin{cases} y'(t)+\partial\varphi_\alpha(y(t))-\alpha y(t)\ni f(t), & \text{a.e. } t\in(0,T),\\ y(0)=y_0, \end{cases} \tag{5.49}
$$

where $\varphi_\alpha : H\to\overline{\mathbf{R}}$ is defined by (5.43). For $f=f_n$ and $y_0=y_0^n$, $y=y_n$, we have the estimate

$$
|y_n'(t)|^2+\frac{d}{dt}\,\varphi_\alpha(y_n(t))-\frac{\alpha}{2}\frac{d}{dt}\,|y_n(t)|^2\le (f_n(t),y_n'(t)), \quad \text{a.e. } t\in(0,T).
$$

This yields

$$
\int_0^T |y_n'(t)|^2 dt+\varphi_\alpha(y_n(t))\le (f_n(0),y_n^0)+\int_0^T \|f_n'(t)\|_*\|y_n(t)\|dt-\frac{\alpha}{2}\,|y_n^0|^2.
$$

Finally,

$$
\int_0^T |y_n'(t)|^2 dt+\|y_n(t)\|^2\le C(\|f_n\|_{W^{1,2}([0,T];V')}+|y_n^0|^2)\le C.
$$

Then, arguing as before, we see that the function y given by (5.47) belongs to $W^{1,2}([0,T];H)\cap L^\infty(0,T;V)$ and is a solution to equation (5.38).

Because $y\in C([0,T];H)\cap L^\infty(0,T;V)$, it is readily seen that y is weakly continuous from $[0,T]$ to V.

If $f\in L^\infty(0,T;H)$ and $y_0\in D(\varphi_\alpha)$, we may apply Theorem 5.1 to equation (5.49) to arrive at the same result. □

Theorem 5.2. *Let $y_0\in K$ and $f\in W^{1,2}([0,T];V')$ be given such that*

$$
(f(0)-Ay_0-\xi_0,y_0-v)\ge 0, \qquad \forall v\in K, \tag{5.50}
$$

for some $\xi_0\in H$.

Then, (5.41) *has a unique solution $y\in W^{1,\infty}([0,T];H)\cap W^{1,2}([0,T];V)$.*

If $y_0\in K$ and $f\in W^{1,2}([0,T];V')$, then system (5.41) *has a unique solution $y\in W^{1,2}([0,T];H)\cap C_w([0,T];V)$. If $f\in L^2(0,T;H)$ and $y_0\in K$, then* (5.41) *has a unique solution $y\in W^{1,2}([0,T];H)\cap C_w([0,T];V)$. Assume in addition that*

$$
(Ay,y)\ge\omega\|y\|^2, \qquad \forall y\in V, \tag{5.51}
$$

for some $\omega>0$, and that there is $h\in H$ such that

$$(I+\varepsilon A_H)^{-1}(y+\varepsilon h)\in K, \qquad \forall \varepsilon>0,\ \forall y\in K. \tag{5.52}$$

Then, $Ay\in L^2(0,T;H)$.

Proof. The first part of the theorem is an immediate consequence of Theorem 5.1. Now, assume that $f\in L^2(0,T;H)$, $y_0\in K$, and conditions (5.51) and (5.52) hold. Let $y\in W^{1,2}([0,T];H)\cap C_w([0,T];V)$ be the solution to (5.41). If in (5.41) we take $z=(I+\varepsilon A_H)^{-1}(y+\varepsilon h)$ (we recall that $A_Hy=Ay\cap H$), we get

$$\begin{aligned}&(y'(t)+A(t),A_\varepsilon(t)-(I+\varepsilon A_H)^{-1}h)\\ &\le (f(t),A_\varepsilon y(t)-(I+\varepsilon A_H)^{-1}h), \quad \text{a.e. } t\in(0,T),\end{aligned}$$

where $A_\varepsilon=A(I+\varepsilon A_H)^{-1}=\varepsilon^{-1}(I-(I+\varepsilon A_H)^{-1})$. Because, by monotonicity of A,

$$(Ay,A_\varepsilon y)\ge |A_\varepsilon y|^2, \qquad \forall y\in D(A_H)=\{y;\ Ay\in H\}$$

and

$$\frac{1}{2}\,\frac{d}{dt}\,(A_\varepsilon y(t),y(t))=(y'(t),A_\varepsilon(t)), \quad \text{a.e. } t\in(0,T),$$

we get

$$\begin{aligned}&(A_\varepsilon y(t),y(t))+\int_0^t |A_\varepsilon y(s)|^2ds\\ &\le (A_\varepsilon y_0,y_0)+2\int_0^t (A_\varepsilon y(s)-(I+\varepsilon A_H)^{-1}f(s),h)ds\\ &+\int_0^t |f(s)|^2ds+2(y(t)-y_0,(I+\varepsilon A_H)^{-1}h), \quad \text{a.e. } t\in(0,T).\end{aligned}$$

Hence,

$$\int_0^T |A_\varepsilon y(t)|^2dt+(A_\varepsilon(t),y(t))\le C, \qquad \forall \varepsilon>0,\ t\in[0,T],$$

and, by Proposition 2.3, we conclude that $Ay\in L^2(0,T;H)$, as claimed. □

Now, we prove a variant of Theorem 5.1 in the case where $\varphi: V\to \overline{\mathbf{R}}$ is lower semicontinuous on H. (It is easily seen that this happens, for instance, if $\varphi(u)/\|u\|\to+\infty$ as $\|u\|\to\infty$.

Proposition 5.10. *Let* $A:V\to V'$ *be a linear, continuous, symmetric operator satisfying condition* (5.37) *and let* $\varphi: H\to\overline{\mathbf{R}}$ *be a lower semicontinuous convex function. Furthermore, assume that there is* C *independent of* ε *such that either*

$$(Ay,\nabla\varphi_\varepsilon(y))\ge -C(1+|\nabla\varphi_\varepsilon(y)|)(1+|y|), \qquad \forall y\in D(A_H), \tag{5.53}$$

or

$$\varphi((I+\varepsilon A_H)^{-1}(y+\varepsilon h))\le\varphi(y)+C, \qquad \forall\varepsilon>0,\ \forall y\in H, \tag{5.54}$$

for some $h\in H$*, where* $A_\alpha=\alpha I+A_H$.

Then, for every $y_0 \in \overline{D(\varphi) \cap V}$ *and every* $f \in L^2(0,T;H)$, *problem* (5.41) *has a unique solution* $y \in W^{1,2}((0,T];H) \cap C([0,T];H)$ *such that* $t^{1/2}y' \in L^2(0,T;H)$, $t^{1/2}Ay \in L^2(0,T;H)$. *If* $y_0 \in D(\varphi) \cap V$, *then* $y \in W^{1,2}([0,T];H) \cap C([0,T];V)$. *Finally, if* $y_0 \in D(A_H) \cap D(\partial\varphi)$ *and* $f \in W^{1,1}([0,T];H)$, *then* $y \in W^{1,\infty}([0,T];H)$.

Here, φ_ε is the regularization of φ.

Proof. As seen previously, the operator

$$A_\alpha y = \alpha y + Ay, \qquad \forall y \in D(A_\alpha) = D(A_H),$$

is maximal monotone in $H \times H$. Then, by Theorem 2.6 (if condition (5.53) holds) and, respectively, Theorem 2.1 (under assumption (5.54)), $A_\alpha + \partial\varphi$ is maximal monotone in $H \times H$ and

$$|A_\alpha y| \le C(|(A_\alpha + \partial\varphi)^0(y)| + |y| + 1), \qquad \forall y \in D(A_H) \cap D(\partial\varphi).$$

Moreover, $A_\alpha + \partial\varphi = \partial\varphi^\alpha$, where (see (5.43))

$$\varphi^\alpha(y) = \frac{1}{2}\,(Ay,y) + \varphi(y) + \frac{\alpha}{2}\,|y|^2, \qquad \forall y \in V,$$

and writing equation (5.39) as

$$y' + \partial\varphi^\alpha(y) - \alpha y \ni f, \qquad \text{a.e. in } (0,T),$$
$$y(0) = y_0,$$

it follows by Theorem 4.1 that there is a strong solution y to equation (5.43) satisfying the conditions of the theorem. Note, for instance, that if $y_0 \in D(\varphi) \cap V$, then $y \in W^{1,2}([0,T];H)$ and $\varphi^\alpha(y) \in W^{1,1}([0,T])$. Because y is continuous from $[0,T]$ to H and bounded in V, this implies that y is weakly continuous from $[0,T]$ to V. Now, because $t \to \varphi^\alpha(y(t))$ is continuous and $\varphi : H \to \overline{\mathbf{R}}$ is lower semicontinuous, we have

$$\lim_{t_n \to t} (Ay(t_n), y(t_n)) \le (Ay(t), y(t)), \qquad \forall t \in [0,T],$$

and this implies that $y \in C([0,T];V)$, as claimed. □

Corollary 5.1. *Let* $A : V \to V'$ *be a linear, continuous, and symmetric operator satisfying condition* (5.37) *and let* K *be a closed convex subset of* H *with*

$$(I + \varepsilon A_\alpha)^{-1}(y + \varepsilon h) \in K, \qquad \forall \varepsilon > 0,\ \forall y \in K, \tag{5.55}$$

for some $h \in H$. *Then, for every* $y_0 \in K$ *and* $f \in L^2(0,T;H)$, *the variational inequality* (5.41) *has a unique solution*

$$y \in W^{1,2}([0,T];H) \cap C([0,T];V) \cap L^2(0,T;D(A_H)).$$

Moreover, one has

$$\frac{dy}{dt}(t) + (A_H y(t) - f(t) - N_K(y(t)))^0 = 0, \quad a.e.\ t \in (0,T),$$

where $N_K(y) \subset L^2(\Omega)$ *is the normal cone at* K *in* y*.*

The parabolic variational inequalities represent a rigorous and efficient way to treat dynamic diffusion problems with a free or moving boundary. As an example, consider the *obstacle parabolic problem*

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y = f & \text{in } \{(x,t) \in Q;\ y(x,t) > \psi(x)\}, \\ \dfrac{\partial y}{\partial t} - \Delta y \geq f & \text{in } Q = \Omega \times (0,T), \\ y(x,t) \geq \psi(x) & \forall (x,t) \in Q, \\ \alpha_1 y + \alpha_2 \dfrac{\partial y}{\partial \nu} = 0 & \text{on } \Sigma = \partial\Omega \times (0,T), \\ y(x,0) = y_0(x) & x \in \Omega, \end{cases} \tag{5.56}$$

where Ω is an open bounded subset of $\mathbf{R}^N$ with smooth boundary (of class $C^{1,1}$, for instance), $\psi \in H^2(\Omega)$, and $\alpha_1, \alpha_2 \geq 0$, $\alpha_1 + \alpha_2 > 0$.

This is a problem of the form (5.41), where

$$H = L^2(\Omega), \qquad V = H^1(\Omega),$$

and $A \in L(V,V')$ is defined by

$$(Ay,z) = \int_\Omega \nabla y \cdot \nabla z\, dx + \frac{\alpha_1}{\alpha_2} \int_{\partial\Omega} yz\, d\sigma, \qquad \forall y,z \in H^1(\Omega), \tag{5.57}$$

if $\alpha_2 \neq 0$, or

$$(Ay,z) = \int_\Omega \nabla y \cdot \nabla z\, dx, \qquad \forall y,z \in H_0^1(\Omega), \tag{5.58}$$

if $\alpha_2 = 0$. (In this case, $V = H_0^1(\Omega)$, $V' = H^{-1}(\Omega)$.)

The set $K \subset V$ is given by

$$K = \{y \in H^1(\Omega);\ y(x) \geq \psi(x), \quad \text{a.e. } x \in \Omega\}, \tag{5.59}$$

and condition (5.55) is satisfied if

$$\alpha_1 \psi + \alpha_2 \frac{\partial \psi}{\partial \nu} \leq 0, \quad \text{a.e. on } \partial\Omega. \tag{5.60}$$

Note also that $A_H : D(A_H) \subset L^2(\Omega) \to L^2(\Omega)$ is defined by

$$A_H y = -\Delta y, \quad \text{a.e. in } \Omega, \qquad \forall y \in D(A_H),$$
$$D(A_H) = \left\{ y \in H^2(\Omega);\ \alpha_1 y + \alpha_2 \frac{\partial y}{\partial \nu} = 0, \quad \text{a.e. on } \partial\Omega \right\},$$

and

$$\|y\|_{H^2(\Omega)} \le C(\|A_H y\|_{L^2(\Omega)} + \|y\|_{L^2(\Omega)}), \qquad \forall y \in D(A_H),$$

Then, we may apply Corollary 5.1 to get the following.

Corollary 5.2. *Let $f \in L^2(Q)$, $y_0 \in H^1(\Omega)$ ($y_0 \in H_0^1(\Omega)$ if $\alpha_2 = 0$) be such that $y_0 \ge \psi$, a.e. in Ω. Assume also that $\psi \in H^1(\Omega)$ satisfies condition* (5.60). *Then, problem* (5.56) *has a unique solution*

$$y \in W^{1,2}([0,T]; L^2(\Omega)) \cap L^2(0,T; H^2(\Omega)) \cap C([0,T]; H_0^1(\Omega)).$$

Noting that

$$N_K(y) = \{v \in L^2(\Omega);\ v(x) \in \beta(y(x) - \psi(x)), \quad \text{a.e. } x \in \Omega\},$$

where $\beta : \mathbf{R} \to 2^{\mathbf{R}}$ is given by

$$\beta(r) = \begin{cases} 0 & r > 0, \\ \mathbf{R}^- & r = 0, \\ \emptyset & r < 0, \end{cases}$$

it follows by Corollary 5.1 that the solution y satisfies the equation

$$\frac{d}{dt} y(t) + (-\Delta y(t) + \beta(y(t) - \psi) - f(t))^0 = 0, \quad \text{a.e. } t \in (0,T).$$

Hence, the solution y to problem (5.56) given by Corollary 5.2 satisfies the system

$$\begin{cases} \dfrac{\partial}{\partial t} y(x,t) - \Delta y(x,t) = f(x,t), \quad \text{a.e. in } \{(x,t) \in Q;\ y(x,t) > \psi(x)\}, \\ \dfrac{\partial}{\partial t} y(x,t) = \max\{f(x,t) + \Delta\psi(x), 0\}, \quad \text{a.e. in } \{(x,t);\ y(x,t) = \psi(x)\}, \end{cases} \tag{5.61}$$

because $y(\cdot,t) \in H^2(\Omega)$ and so $\Delta y(x,t) = \Delta\psi(x)$, a.e. in $\{y(x,t) = \psi(x)\}$.

It follows, also, that the solution y to the obstacle problem (5.56) is given by

$$y(t) = \lim_{\varepsilon \to 0} y_\varepsilon(t) \quad \text{in } C([0,T]; L^2(\Omega)),$$

where y_ε is the solution to the penalized problem

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y - \dfrac{1}{\varepsilon}(y - \psi)^- = f & \text{in } Q, \\ y(x,0) = y_0(x) & \text{in } \Omega, \\ \alpha_1 y + \alpha_1 \dfrac{\partial y}{\partial \nu} = 0 & \text{on } \Sigma. \end{cases} \tag{5.62}$$

Now, let us consider the obstacle problem (5.56) with nonhomogeneous boundary conditions; that is,

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y = f & \text{in } \{(x,t) \in Q;\ y(x,t) > \psi(x)\}, \\ \dfrac{\partial y}{\partial t} - \Delta y \geq f,\ y \geq 0 & \text{in } Q, \\ \alpha y + \dfrac{\partial y}{\partial \nu} = g & \text{on } \Sigma_1 = \Gamma_1 \times (0,T), \\ y = 0 & \text{on } \Sigma_2 = \Gamma_2 \times (0,T), \\ y(x,0) = y_0(x) & \text{on } \Omega, \end{cases} \tag{5.63}$$

where $\partial\Omega = \Gamma_1 \cup \Gamma_2,\ \Gamma_1 \cap \Gamma_2 = \emptyset$, and $g \in L^2(\Sigma_1)$.

If we take

$$V = \{y \in H^1(\Omega); \qquad y = 0 \quad \text{on } \Gamma_2\},$$

define $A : V \to V'$ by

$$(Ay, z) = \int_\Omega \nabla y \cdot \nabla z\, dx + \alpha \int_{\Gamma_1} yz\, dx, \qquad \forall y, z \in V,$$

and $f_0 : [0,T] \to V'$ by

$$(f_0(t), z) = \int_{\Gamma_1} g(x,t) z(x) dx, \qquad \forall z \in V,$$

we may write problem (5.63) as

$$\begin{aligned} \left(\frac{dy}{dt}(t) + Ay(t), y(t) - z\right) &\leq (F(t), y(t) - z), \qquad \forall z \in K,\ \text{a.e. } t \in (0,T), \\ y(0) &= y_0, \end{aligned} \tag{5.64}$$

where $F = f + f_0 \in L^2(0,T;V')$ and K is defined by (5.59).

Equivalently,

$$\begin{aligned} \int_\Omega \frac{\partial y}{\partial t}(x,t)(y(x,t) - z(x))dx &+ \int_\Omega \nabla y(x,t) \cdot \nabla(y(x,t) - z(x))dx \\ &+ \alpha \int_{\Gamma_1} f(x,t)(y(x,t) - z(x))dx \\ &\leq \int_\Omega f(x,t)(y(x,t) - z(x))dx \\ &+ \int_{\Gamma_1} g(x,t)(y(x,t) - z(x))dx, \\ &\qquad\qquad \forall z \in K,\ t \in [0,T]. \end{aligned} \tag{5.65}$$

Applying Theorem 5.2, we get the following.

Corollary 5.3. *Let* $f \in W^{1,2}([0,T];L^2(\Omega))$*,* $g \in W^{1,2}([0,T];L^2(\Gamma_1))$*, and* $y_0 \in K$*. Then, problem* (5.65) *has a unique solution*

$$y \in W^{1,2}([0,T];V) \cap C_w([0,T];V).$$

If, in addition,

$$\begin{cases} \dfrac{\partial y_0}{\partial \nu} + \alpha y_0 = g(x,0), & \text{a.e. on } \{x \in \Gamma_1;\ y_0(x) > \psi(x)\}, \\ \dfrac{\partial \psi}{\partial \nu} + \alpha \psi \le g(x,0), & \text{a.e. on } \{x \in \Gamma_1;\ y_0(x) = \psi(x)\}, \end{cases} \tag{5.66}$$

then $y \in W^{1,2}([0,T];V) \cap W^{1,\infty}([0,T];L^2(\Omega))$.

(We note that condition (5.66) implies (5.50).)

It is readily seen that the solution y to (5.65) satisfies (5.63) in a certain generalized sense. Indeed, assuming that the set $E = \{(x,t);\ y(x,t) > \psi(x)\}$ is open and taking $z = y(x,t) \pm \rho\varphi$ in (5.65), where $\varphi \in C_0^\infty(E)$ and ρ is sufficiently small, we see that

$$\frac{\partial y}{\partial t} - \Delta y = f \quad \text{in } \mathscr{D}'(E). \tag{5.67}$$

It is also obvious that

$$\frac{\partial y}{\partial t} - \Delta y \ge f \quad \text{in } \mathscr{D}'(Q). \tag{5.68}$$

Regarding the boundary conditions, by (5.65), (5.67), and (5.68), it follows that

$$\frac{\partial y}{\partial \nu} + \alpha y = g \quad \text{in } \mathscr{D}'(E \cap \Sigma_1),$$

respectively,

$$\frac{\partial y}{\partial \nu} + \alpha y \ge g \quad \text{in } \mathscr{D}'(\Sigma_1).$$

In other words,

$$\begin{cases} \dfrac{\partial y}{\partial \nu} + \alpha y = g & \text{on } \{(x,t) \in \Sigma_1;\ y(x,t) > \psi(x)\}, \\ \dfrac{\partial \psi}{\partial \nu} + \alpha \psi \ge g & \text{on } \{(x,t) \in \Sigma_1;\ y(x,t) = \psi(x)\}. \end{cases}$$

Hence, if g satisfies the compatibility condition

$$\frac{\partial \psi}{\partial \nu} + \alpha \psi \le g \quad \text{on } \Sigma_1,$$

then the solution y to problem (5.65) satisfies the required boundary conditions on Σ_1.

Also in this case, the solution y given by Corollary 5.3 can be obtained as the limit as $\varepsilon \to 0$ of the solution y_ε to the equation

$$\begin{cases} \dfrac{\partial y_\varepsilon}{\partial t} - \Delta y_\varepsilon + \beta_\varepsilon(y_\varepsilon - \psi) = f & \text{in } \Omega \times (0,T), \\ y_\varepsilon(x,0) = y_0(x) & \text{in } \Omega, \\ \dfrac{\partial y_\varepsilon}{\partial \nu} + \alpha y_\varepsilon = g \quad \text{on } \Sigma_1, & y_\varepsilon = 0 \quad \text{on } \Sigma_2, \end{cases} \tag{5.69}$$

where

$$\beta_\varepsilon(r) = -\left(\frac{1}{\varepsilon}\right) r^-, \qquad \forall r \in \mathbf{R}.$$

If $Q^+ = \{(x,t) \in Q;\ y(x,t) > \psi(x)\}$, we may view y as the solution to the free boundary problem

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta y = f & \text{in } Q^+, \\ y(x,0) = y_0(x) & \text{in } \Omega, \\ \alpha_1 y + \alpha_2 \dfrac{\partial y}{\partial \nu} = 0 \quad \text{on } \Sigma, & y = \psi,\ \dfrac{\partial y}{\partial \nu} = \dfrac{\partial \psi}{\partial \nu} \quad \text{on } \partial Q^+(t), \end{cases} \tag{5.70}$$

where $\partial Q^+(t)$ is the boundary of the set $Q^+(t) = \{x \in \Omega;\ y(x,t) > \psi(x)\}$. We call $\partial Q^+(t)$ the *moving boundary* and ∂Q^+ the *free boundary* of problem (5.70).

In problem (5.70), the noncoincidence set Q^+ as well as the free boundary ∂Q^+ is not known a priori and represents unknowns of the problem. In problem (5.41) or (5.65), the free boundary does not appear explicitly, but in this formulation the problem is nonlinear and multivalued.

Perhaps the best-known example of a parabolic free boundary problem is the classical Stefan problem, which we briefly describe in what follows and which has provided one of the principal motivations of the theory of parabolic variational inequalities.

The Stefan Problem

This problem describes the conduction of heat in a medium involving a phase charge. To be more specific, consider a unit volume of ice Ω at temperature $\theta < 0$. If a uniform heat source of intensity F is applied, then the temperature increases at rate E/C_1 until it reaches the melting point $\theta = 0$. Then, the temperature remains at zero until ρ units of heat have been supplied to transform the ice into water (ρ is the latent heat). After all the ice has melted the temperature begins to increase at the rate h/C_2 (C_1 and C_2 are specific heats of ice and water, respectively). During the process, the variation of the internal energy $e(t)$ is therefore given by

$$e(t) = C(\theta(t)) + \rho H(\theta(t)),$$

where

$$C(\theta) = \begin{cases} C_1\theta & \text{for } \theta \le 0, \\ C_2\theta & \text{for } \theta > 0, \end{cases}$$

and H is the Heaviside graph

$$H(\theta) = \begin{cases} 1 & \theta > 0, \\ [0,1] & \theta = 0, \\ 0 & \theta < 0. \end{cases}$$

In other words, we have

$$e = \gamma(\theta) = \begin{cases} C_1\theta & \text{if } \theta < 0, \\ [0,\rho] & \text{if } \theta = 0, \\ C_2\theta + \rho & \text{if } \theta > 0. \end{cases} \tag{5.71}$$

The function γ is called the *enthalpy* of the system.

Now, let $Q = \Omega \times (0,\infty)$ and denote by Q_-, Q_+, Q_0 the regions of Q, where $\theta < 0$, $\theta > 0$, and $\theta = 0$, respectively. We set $S_+ = \partial Q_+$, $S_- = \partial Q_-$, and $S = S_+ \cup S_-$.

If $\theta = \theta(x,t)$ is the temperature distribution in Q and $q = q(x,t)$ the heat flux, then, according to the Fourier law,

$$q(x,t) = -k\nabla\theta(x,t), \tag{5.72}$$

where k is the thermal conductivity. Consider the function

$$K(\theta) = \begin{cases} k_1\theta & \text{if } \theta < 0, \\ k_2\theta & \text{if } \theta > 0, \end{cases}$$

where k_1, k_2 are the thermal conductivity of the ice and water, respectively.

If f is the external heat source, then the conservation law yields

$$\frac{d}{dt}\int_{\Omega^*} e(x,t)dx = -\int_{\partial\Omega^*}(q(x,t),\nu)d\sigma + \int_{\Omega^*} F(x,t)dx$$

for any subdomain $\Omega^* \times (t_1,t_2) \subset Q$ (ν is the normal to $\partial\Omega^*$) if e and q are smooth. Equivalently,

$$\begin{aligned} &\int_{\Omega^*} e_t(x,t)dx + \int_{S\cap\Omega^*} [|e(t)|]V(t)dt \\ &= -\int_{\Omega^*} \operatorname{div} q(x,t)dx + \int_{\partial\Omega^*\cap S} [|(q(t),\nu)|]d\sigma + \int_{\Omega^*} F(x,t)dx, \end{aligned}$$

where $V(t) = -N_t\|N_t\|$ is the true velocity of the interface S ($N = (N_1,N_2)$ is the unit normal to S) and $[|\cdot|]$ is the jump along S.

The previous inequality yields

$$\begin{aligned} &\frac{\partial}{\partial t} e(x,t) + \operatorname{div} q(x,t) = F(x,t) && \text{in } Q \setminus S, \\ &[|e(t)|] N_x + [|(q(t), N_t|] = 0 && \text{on } S. \end{aligned} \tag{5.73}$$

Taking into account equations (5.71)–(5.73), we get the system

$$\begin{cases} C_1 \dfrac{\partial \theta}{\partial t} - k_1 \Delta \theta = f & \text{in } Q_-, \\ C_2 \dfrac{\partial \theta}{\partial t} - k_2 \Delta \theta = f & \text{in } Q_+, \end{cases} \tag{5.74}$$

$$\begin{cases} (k_2 \nabla \theta^+ - k_1 \nabla \theta^-) \cdot N_x = \rho N_t & \text{on } S, \\ \theta^+ = \theta^- = 0 & \text{on } S. \end{cases} \tag{5.75}$$

If we represent the interface S by the equation $t = \sigma(x)$, then (5.75) reads

$$\begin{cases} (k_1 \nabla \theta^+ - k_2 \nabla \theta^-) \cdot \nabla \sigma = -\rho & \text{in } S, \\ \theta^+ = \theta^- = 0. \end{cases} \tag{5.76}$$

The usual boundary and initial value conditions can be associated with equations (5.74) and (5.76), for instance,

$$\theta = 0 \qquad \text{in } \partial\Omega \times (0,T), \tag{5.77}$$

$$\theta(x,0) = \theta_0(x) \qquad \text{in } \Omega, \tag{5.78}$$

or Neumann boundary conditions on $\partial\Omega$.

This is the classical two-phase Stefan problem. Here, we first study with the methods of variational inequalities a simplified model described by the one-phase Stefan problem

$$\begin{cases} \dfrac{\partial \theta}{\partial t} - \Delta \theta = 0 & \text{in } Q_+ = \{(x,t) \in Q;\ \sigma(x) < t < T\}, \\ \theta = 0 & \text{in } Q_- = \{(x,t) \in Q;\ 0 < t < \sigma(x)\}, \\ \nabla_x(x,t) \cdot \nabla \sigma(x) = -\rho & \text{on } S = \{(x,t);\ t = \sigma(x)\}, \\ \theta = 0 & \text{in } S \cup Q_-, \\ \theta \geq 0 & \text{in } Q_+. \end{cases} \tag{5.79}$$

These equations model the melting of a body of ice $\Omega \subset \mathbf{R}^3$ maintained at $\theta^0 C$. Therefore, assume that $\partial\Omega = \Gamma_1 \cup \Gamma_2$, where Γ_1 and Γ_2 are disjoint and Γ_1 is in contact with a heating medium with temperature θ_1; $t = \sigma(x)$ is the equation of the interface (moving boundary) S_t, which separates the liquid phase (water) and solid (ice). Thus, to equations (5.79) we must add the boundary conditions

$$\begin{cases} \dfrac{\partial \theta}{\partial \nu} + \alpha(\theta - \theta_1) = 0 & \text{on } \Sigma_1 = \Gamma_1 \times (0,T), \\ \theta = 0 & \text{on } \Sigma_2 = \Gamma_2 \times (0,T) \end{cases} \tag{5.80}$$

and the initial value conditions

$$\theta(x,0) = \theta_0(x) > 0,\ \forall x \in \Omega_0, \qquad \theta(x,0) = 0,\ \forall x \in \Omega \setminus \Omega_0. \tag{5.81}$$

There is a simple device due to G. Duvaut [21] that permits us to reduce problem (5.79)–(5.81) to a parabolic variational inequality. To this end, consider the function

$$y(x,t) = \begin{cases} \displaystyle\int_{\sigma(x)}^{t} \theta(x,s)ds & \text{if } x \in \Omega \setminus \Omega_0,\ t > \sigma(x), \\ \displaystyle\int_0^t \theta(x,s)ds & \text{if } x \in \Omega_0,\ t \in [0,T], \\ 0 & \text{if } (x,t) \in Q_-, \end{cases} \tag{5.82}$$

and let

$$f_0(x,t) = \begin{cases} -\rho & \text{if } x \in \Omega \setminus \Omega_0,\ 0 < t < T, \\ \theta_0(x) & \text{if } x \in \Omega_0,\ 0 < t < T. \end{cases} \tag{5.83}$$

Lemma 5.1. *Let $\theta \in H^1(Q)$ and $\sigma \in H^1(\Omega)$. Then,*

$$\frac{\partial y}{\partial t} - \Delta y = f_0 \chi \quad in\ \mathscr{D}'(Q), \tag{5.84}$$

where χ is the characteristic function of Q_+.

Proof. By (5.82), we have

$$\frac{\partial y}{\partial t}\,(\varphi) = \int_{Q_+} \theta(x,t)\varphi(x,t)dxdt, \qquad \forall \varphi \in C_0^\infty(Q).$$

On the other hand, we have

$$\begin{aligned} (y_x, \varphi) &= -y(\varphi_x) \\ &= -\int_{\Omega\setminus\Omega_0} dx \int_{\sigma(x)}^T \varphi_x(x,t)dt \int_{\sigma(x)}^t \theta(x,s)ds \\ &\quad - \int_{\Omega_0} dx \int_0^T \varphi_x(x,t)dt \int_0^t \theta(x,s)ds \\ &= -\int_{\Omega\setminus\Omega_0} dx \operatorname{div}\left(\int_{\sigma(x)}^T \varphi(x,t)dt \int_{\sigma(x)}^t \theta(x,s)ds\right) \\ &= \int_{\Omega\setminus\Omega_0} dx \left(\int_{\sigma(x)}^T \varphi(x,t)dt \int_{\sigma(x)}^t \theta_x(x,s)ds\right) \\ &\quad - \int_{\Omega_0} dx \operatorname{div}\left(\int_0^T \varphi(x,t)dt \int_0^t \theta(x,s)ds\right) \end{aligned}$$

$$= \int_{\Omega \setminus \Omega_0} dx \int_{\sigma(x)}^{T} \varphi(x,t)dt \int_{\sigma(x)}^{t} \theta_x(x,s)s$$
$$+ \int_{\Omega_0} dx \int_0^T \varphi(x,t)dt \int_{\sigma(x)}^{t} \theta_x(x,s)ds.$$

(Here, $y_x = \nabla_x y,\ \varphi_x = \nabla_x \varphi$.) This yields

$$\Delta y(\varphi) = -y_x(\varphi_x) = -\int_{\Omega \setminus \Omega_0} dx \int_{\sigma(x)}^{T} \varphi_x(x,t)dt \cdot \int_{\sigma(x)}^{T} \theta_x(x,s)ds$$
$$-\int_{\Omega_0} dx \int_0^T \varphi_x(x,t)dt \cdot \int_0^t \theta_x(x,s)ds$$

and, by the divergence formula, we get

$$\Delta y(\varphi) = \int_{\Omega \setminus \Omega_0} dx \int_{\sigma(x)}^{T} dt \left(\int_{\sigma(x)}^{t} \Delta\theta(x,s)ds\, \varphi(x,t) \right)$$
$$+ \int_{\Omega_0} ds \int_0^T dt \left(\int_0^t \Delta\theta(x,s)ds\, \varphi(x,t) \right), \qquad \forall \varphi \in C_0^\infty(Q),$$

because $\nabla_x \theta(x, \sigma(x)) \cdot \nabla\sigma(x) = -\rho$, $\forall x \in \Omega \setminus \Omega_0$. Then, by equations (5.79), we see that

$$\left(\frac{\partial y}{\partial t} - \Delta y \right)(\varphi) = -\int_{\Omega \setminus \Omega_0} dx \int_{\sigma(x)} dt \left(\int_{\sigma(x)}^{t} \theta_t(x,s)ds - \theta(s,t) \right) \varphi(x,t)$$
$$-\int_{\Omega_0} dx \int_0^T dt \left(\int_0^t \theta_t(x,s)ds - \theta(x,t) \right) \varphi(x,t)$$
$$-\rho \int_{\Omega \setminus \Omega_0} dx \int_{\sigma(x)}^{T} \varphi(x,t)dt$$
$$= \int_{Q_+} f_0(x,t)\varphi(x,t)dx\,dt,$$

as claimed. □

By Lemma 5.1 we see that the function y satisfies the obstacle problem

$$\begin{cases} y \geq 0,\ \dfrac{\partial y}{\partial t} - \Delta y \geq f_0 & \text{in } Q, \\ \dfrac{\partial y}{\partial t} - \Delta y = f_0 & \text{in } \{(x,t) \in Q;\ y(x,t) > 0\}, \\ y = 0 & \text{in } \{(x,t) \in Q;\ \sigma(x) > t\}, \end{cases} \tag{5.85}$$

and the boundary value conditions

$$\frac{\partial}{\partial \nu} \frac{\partial y}{\partial t} = -\alpha \left(\frac{\partial y}{\partial t} - \theta_1 \right) \quad \text{on } \Sigma_1, \qquad \frac{\partial y}{\partial t} = 0 \quad \text{on } \Sigma_2, \tag{5.86}$$

(see (5.80) and (5.82)). Then, by Corollary 5.2, we have the following.

Corollary 5.4. *Let* $\theta_1 \in L^2(\Sigma_1)$ *be given. Then, problem* (5.85) and (5.86) *has a unique* (*generalized*) *solution* $y \in W^{1,\infty}([0,T];L^2(\Omega)) \cap W^{1,2}([0,T];H^1(\Omega))$.

Keeping in mind that $S_t = \partial\{(x,t);\ y(x,t) = 0\}$, we can derive from Corollary 5.4 an existence result for the one-phase Stefan problem (5.79)–(5.81).

Other mathematical models for physical problems involving a free boundary such as the oxygen diffusion in an absorbing tissue (Elliott and Ockendon [23]) or electrochemical machining processes lead by similar devices to parabolic variational inequalities of the same type. It should be mentioned also that dynamics of elasto-plastic materials as well as the phase transition in systems composed of different metals are better described by parabolic variational inequalities, eventually combined with linear hyperbolic equations. This is the case for instance with Fremond's model of thermomechanical dynamics of shape memory delay. The phase transition often manifests a hysteretic behavior due to irreversible changes in process dynamics and the study of hypothesis models is another source of variational inequalities although the hysteresis operator, in general, is not monotone in the sense described above. However, some standard hysteresis equations (stop and play, for instance) are expressed in terms of variational inequalities. (We refer to Visintin book's [42] for a treatment of these problems.)

5.3 The Porous Media Diffusion Equation

The nonlinear diffusion equation models the dynamic of density in a substance undergoing diffusion described by Fick's first law (or Darcy's law). It also models phase transition dynamics (the Stefan problem) or other physical processes that are of diffusion type (heat propagation, filtration, or dynamics of biological groups). Such an equation can be schematically written as

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta\beta(y) \ni f & \text{in } \Omega \times (0,T) = Q, \\ \beta(y) = 0 & \text{on } \partial\Omega \times (0,T) = \Sigma, \\ y(x,0) = y_0(x) & \text{in } \Omega, \end{cases} \tag{5.87}$$

where Ω is a bounded and open subset of $\mathbf{R}^N$ with smooth boundary, and $\beta : \mathbf{R} \to 2^{\mathbf{R}}$ is a maximal monotone graph in $\mathbf{R} \times \mathbf{R}$ such that $0 \in D(\beta)$.

The steady-state equation associated with (5.87) is just the stationary porous media equation studied in Sections 2.2 and 3.2.

The function $y \in C([0,T];L^1(\Omega))$ is called a generalized solution to problem (5.87) if

$$\int_Q (y\varphi_t + \beta(y)\Delta\varphi)\,dx\,dt + \int_Q f\varphi\,dx\,dt + \int_\Omega y_0\varphi(x,0)\,dx = 0 \tag{5.88}$$

for all $\varphi \in C^{2,1}(\overline{Q})$ such that $\varphi(x,T) = 0$ in Ω and $\varphi = 0$ on Σ.

Let us first briefly describe some specific diffusive-like problems that lead to equations of this type.

1. *The flow of gases in porous media.* Let y be the density of a gas that flows through a porous medium that occupies a domain $\Omega \subset \mathbf{R}^3$ and let $\bar{v}$ be the pore velocity. If p denotes the pressure, we have $p = p_0 y^\alpha$ for $\alpha \geq 1$. Then, the conservation law equation

$$k_1 \frac{\partial y}{\partial t} + \operatorname{div}(y\bar{v}) = 0$$

combined with Darcy's law

$$\gamma\bar{v} = -k_2 \nabla p$$

(k_1 is the porosity of the medium, k_2 the permeability, and γ the viscosity) yields the *porous medium equation*

$$\frac{\partial y}{\partial t} - \delta \Delta y^{\alpha+1} = 0 \quad \text{in } Q, \tag{5.89}$$

where

$$\delta = k_2 p_0 (k_1(\alpha+1)\gamma)^{-1}.$$

Equation (5.89) is also relevant in the study of other mathematical models, such as population dynamics. The case where $-1 < \alpha < 0$ is that of fast diffusion processes arising in physics of plasma. In particular, the case

$$\beta(x) = \begin{cases} \log x & \text{for } x > 0 \\ -\infty & \text{for } x \leq 0 \end{cases}$$

emerges from the central limit approximation to Carleman's model of Boltzman equations. Nonlinear diffusion equations of the form (5.87) perturbed by a term of transport; that is,

$$\frac{\partial y}{\partial t} - \Delta\beta(y) + \operatorname{div} K(y) \ni f$$

with appropriate boundary conditions arise in the dynamics of underground water flows and are known in the literature as the Richards equation. The special case

$$\beta(y) = \begin{cases} \beta_0(y) & \text{for } y < y_s, \\ [\beta_0(y_s), +\infty) & \text{for } y = y_s, \\ \emptyset & \text{for } y > y_s, \end{cases}$$

where $\beta_0 : \mathbf{R} \to \mathbf{R}$ is a continuous and monotonically increasing function, models the dynamics of saturated–unsaturated underground water flows. The treatment of such an equation with methods of nonlinear accretive differential equations is given in Marinoschi [34, 35].

2. *Two-phase Stefan problem.* We come back to the two-phase Stefan problem (5.74), (5.75), (5.77), (5.78); that is

$$\begin{cases} C_1\theta_t - k_1\Delta\theta = f & \text{in } Q_-\{(x,t);\ \theta(x,t)<0\} \\ C_2\theta_t - k_2\Delta\theta = f & \text{in } Q_+ = \{(x,t);\ \theta(x,t)>0\}, \\ (k_1\nabla\theta^+ - k_2\nabla\theta^-)\cdot\nabla\sigma(x) = -\rho & \text{on } S, \end{cases} \tag{5.90}$$

where $t=\sigma(x)$ is the equation of the interface S.

We may write system (5.90) as

$$\frac{\partial}{\partial t}\gamma(\theta) - \Delta K(\theta) \ni f \quad \text{in } Q, \tag{5.91}$$

where $\gamma:\mathbf{R}\to 2^{\mathbf{R}}$ is given by (5.71). Indeed, for every test function $\varphi\in C_0^\infty(Q)$ we have

$$\begin{aligned}
&\left(\frac{\partial}{\partial t}\gamma(\theta) - \Delta K(\theta)\right)(\varphi) \\
&= -\int_Q (\gamma(\theta)\varphi_t + K(\theta)\Delta\varphi)\,dx\,dt \\
&= C_1\int_{Q_-}\theta_t\varphi\,dx\,dt + C_2\int_{Q_+}\theta_t\,dx\,dt - k_1\int_{Q_-}\varphi\Delta\theta\,dx\,dt \\
&\quad -k_2\int_{Q_+}\varphi\Delta\theta\,dx\,dt + \int_S\left(k_2\frac{\partial\theta^+}{\partial\nu} - k_1\frac{\partial\theta^-}{\partial\nu}\right)\varphi\,ds - \rho\int_{Q_+}\varphi_t\,dx\,dt \\
&= \int_{Q_-}(C_1\theta_t - k_1\Delta\theta)\varphi\,dx\,dt + \int_{Q_+}(C_2\theta_t - k_2\Delta\theta)\varphi\,dx\,dt \\
&\quad + \int_S((k_2\nabla\theta^+ - k_1\nabla\theta^-)\cdot\nabla\sigma + \rho)dx = 0.
\end{aligned} \tag{5.92}$$

If we denote by β the function $\gamma^{-1}K$; that is,

$$\beta(r) = \begin{cases} k_1C_1^{-1}r & \text{for } r<0, \\ 0 & \text{for } 0\le r<\rho, \\ k_2C_2^{-1}(r-\rho) & \text{for } r\ge\rho, \end{cases} \tag{5.93}$$

we may write (5.91) in the form (5.87).

Problem (5.87) can be treated as a nonlinear accretive Cauchy problem in two functional spaces: $H^{-1}(\Omega)$ and $L^1(\Omega)$.

3. *The Hilbert space approach.* In the space $H^{-1}(\Omega)$, consider the operator

$$\begin{aligned} A = \{\ [y,w]\in(H^{-1}(\Omega)\cap L^1(\Omega))\times H^{-1}(\Omega);\ w=-\Delta v, \\ v\in H_0^1(\Omega),\ v(x)\in\beta(y(x)),\quad \text{a.e. } x\in\Omega\}. \end{aligned}$$

We assume that

$$\beta^{-1} \text{ is everywhere defined and bounded on the bounded subsets of } \mathbf{R}. \tag{5.94}$$

Then, by Proposition 2.10, A is maximal monotone in $H^{-1}(\Omega) \times H^{-1}(\Omega)$. More precisely, $A = \partial\varphi$, where $\varphi : H^{-1}(\Omega) \to \overline{\mathbf{R}}$ is defined by

$$\varphi(y) = \begin{cases} \displaystyle\int_\Omega j(y(x))dx & \text{if } y \in L^1(\Omega) \cap H^{-1}(\Omega),\ j(y) \in L^1(\Omega), \\ +\infty & \text{otherwise,} \end{cases}$$

where $\partial j = \beta$.

Then, we may write problem (5.87) as

$$\begin{aligned} \frac{dy}{dt} + Ay \ni f \qquad & \text{in } (0,T), \\ y(0) = y_0, & \end{aligned} \tag{5.95}$$

and so, by Theorem 4.11, we obtain the following existence result.

Theorem 5.3. *Let β be a maximal monotone graph in $\mathbf{R} \times \mathbf{R}$ satisfying condition (5.94). Let $f \in L^1(0,T;H^{-1}(\Omega))$ and let $y_0 \in H^{-1}(\Omega) \cap L^1(\Omega)$ be such that $y_0(x) \in \overline{D(\beta)}$, a.e. $x \in \Omega$. Then, there is a unique pair of functions $y \in C([0,T];H^{-1}(\Omega)) \cap W^{1,2}(0,T;H^{-1}(\Omega))$ and $v : Q \to \mathbf{R}$, such that $v(t) \in H_0^1(\Omega)$, $\forall t \in [0,T]$ satisfying*

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta v = f, & \text{a.e. in } Q = \Omega \times (0,T), \\ v(x,t) \in \beta(y(x,t)), & \text{a.e. } (x,t) \in Q, \\ y(x,0) = y_0(x), & \text{a.e. in } \Omega. \end{cases} \tag{5.96}$$

$$t^{1/2}\, \frac{\partial y}{\partial t} \in L^2(0,T;H^{-1}(\Omega)), \qquad t^{1/2} v \in L^2(0,T;H_0^1(\Omega)). \tag{5.97}$$

Moreover, if $j(y_0) \in L^1(\Omega)$, then

$$\frac{\partial y}{\partial t} \in L^2(0,T;H^{-1}(\Omega)), \qquad v \in L^2(0,T;H_0^1(\Omega)). \tag{5.98}$$

If $y_0 \in D(A)$ and $f \in W^{1,1}([0,T];H^{-1}(\Omega))$, then

$$\frac{\partial y}{\partial t} \in L^\infty(0,T;H^{-1}(\Omega)), \qquad v \in L^\infty(0,T;H_0^1(\Omega)). \tag{5.99}$$

We note that the derivative $\partial y/\partial t$ in (5.96) is the strong derivative dy/dt of the function $t \to y(\cdot,t)$ from $[0,T]$ into $H^{-1}(\Omega)$, and it coincides with the derivative $\partial y/\partial t$ in the sense of distributions on Q. It is readily seen that the solution y (see Theorem 5.3) is a generalized solution to (5.87) in the sense of definition (5.88).

4. *The L^1-approach.* In the space $X = L^1(\Omega)$, consider the operator

$$A = \{[y,w] \in L^1(\Omega) \times L^1(\Omega);\ w = -\Delta v,\\ v \in W_0^{1,1}(\Omega),\ v(x) \in \beta(y(x)),\quad \text{a.e. } x \in \Omega\}. \tag{5.100}$$

We have seen earlier (Theorem 3.5) that A is m-accretive in $L^1(\Omega) \times L^1(\Omega)$. Then, applying the general existence Theorem 4.2, we obtain the following.

Proposition 5.11. *Let β be a maximal monotone graph in $\mathbf{R} \times \mathbf{R}$ such that $0 \in \beta(0)$. Then, for every $f \in L^1(0,T;L^1(\Omega))$ and every $y_0 \in L^1(\Omega)$, such that $y_0(x) \in \overline{D(\beta)}$, a.e. $x \in \Omega$, the Cauchy problem*

$$\begin{cases} \dfrac{dy}{dt}(t) + Ay(t) \ni f(t) & \text{in } (0,T), \\ y(0) = y_0, \end{cases} \tag{5.101}$$

has a unique mild solution $y \in C([0,T];L^1(\Omega))$.

We note that $\overline{D(A)} = \{y_0 \in L^1(\Omega);\ y_0(x) \in \overline{D(\beta)}, \text{ a.e. } x \in \Omega\}$.

Indeed, $(1+\varepsilon\beta)^{-1}y_0 \to y_0$ in $L^1(\Omega)$ as $\varepsilon \to 0$, if $y_0 \in \overline{D(\beta)}$, a.e. $x \in \Omega$, and $(I+\varepsilon A)^{-1}y_0 \to y_0$ if $j(y_0) \in L^1(\Omega)$.

Proposition 5.11 amounts to saying that

$$y(t) = \lim_{\varepsilon \to 0} y_\varepsilon(t) \quad \text{in } L^1(\Omega), \text{ uniformly on } [0,T],$$

where y_ε is the solution to the difference equations

$$\begin{cases} \dfrac{1}{\varepsilon}(y_\varepsilon(t) - y_\varepsilon(t-\varepsilon)) - \Delta v_\varepsilon(t) = f_\varepsilon(t) & \text{in } \Omega \times (0,T), \\ v_\varepsilon(x,t) \in \beta(y_\varepsilon(x,t)), & \text{a.e. in } \Omega \times (0,T), \\ v_\varepsilon = 0 & \text{on } \partial\Omega \times (0,T), \\ y_\varepsilon(t) = y_0 & \text{for } t \le \varepsilon,\ x \in \Omega. \end{cases} \tag{5.102}$$

The function $t \to v_\varepsilon(t) \in W_0^{1,1}(\Omega)$ is piecewise constant on $[0,T]$ and $f_\varepsilon(t) = f_i$, $\forall t \in [i\varepsilon,(i+1)\varepsilon]$ is a piecewise constant approximation of $f : [0,T] \to L^1(\Omega)$.

By (5.102), it is readily seen that y is a generalized solution to problem (5.87). In particular, it follows by Proposition 5.11 that the operator A defined by (5.100) generates a semigroup of nonlinear contractions $S(t) : \overline{D(A)} \to \overline{D(A)}$. This semigroup is not differentiable in $L^1(\Omega)$, but in some special situations it has regularity properties comparable with those of the semigroup generated by the Laplace operator on $L^2(\Omega)$ under Dirichlet boundary conditions. In fact, we have the following smoothing effect of nonlinear semigroup $S(t)$ with respect to the initial data.

Theorem 5.4. *Let $\beta \in C^1(\mathbf{R} \setminus \{0\}) \cap C(\mathbf{R})$ be a monotone function satisfying the conditions*

$$\beta(0) = 0, \qquad \beta'(r) \ge C|r|^{\alpha-1}, \qquad \forall r \ne 0, \tag{5.103}$$

where $\alpha > 0$ if $N \le 2$ and $\alpha > (N-2)/N$ if $N \ge 3$. Then, $S(t)(L^1(\Omega)) \subset L^\infty(\Omega)$ for every $t > 0$,

$$\|S(t)y_0\|_{L^\infty(\Omega)} \le Ct^{-(N/(N\alpha+2-N))}\|y_0\|_{L^1(\Omega)}^{2/(2+N(\alpha-1))}, \qquad \forall t>0, \tag{5.104}$$

and $S(t)(L^p(\Omega)) \subset L^p(\Omega)$ *for all* $t>0$ *and* $1 \le p < \infty$.

Proof. First, we establish the estimates

$$\begin{aligned} \|(I+\lambda A)^{-1}f\|_p^p + C\lambda\left(\int_\Omega |(I+\lambda A)^{-1}f|^{((p+\alpha-1)N)/(N-2)}dx\right)^{(N-2)/N} \\ \le \|f\|_p^p, \qquad \forall f \in L^p(\Omega),\ \lambda>0, \end{aligned} \tag{5.105}$$

for $N>2$, and

$$\begin{aligned} |(I+\lambda A)^{-1}f\|_p + C\lambda\left(\int_\Omega |(I+\lambda A)^{-1}f|^{(p+1-\alpha)q}dx\right)^{1/q} \le \int_\Omega |f|^p dx, \\ \forall q{>}1, \end{aligned} \tag{5.106}$$

if $N=2$. Here $\|\cdot\|_p$ is the L^p norm in Ω, C is independent of $p \ge 1$, and A is the operator defined by (5.100).

We set $u=(I+\lambda A)^{-1}f$; that is,

$$\begin{cases} u-\lambda\Delta\beta(u)=f & \text{in } \Omega, \\ \beta(u)=0 & \text{on } \partial\Omega. \end{cases} \tag{5.107}$$

We recall that $\beta(u) \in W_0^{1,q}(\Omega)$, where $1<q<N/(N-2)$ (see Corollary 3.1).

Multiplying equation (5.107) by $|u|^{p-1}\operatorname{sign} u$ and integrating on Ω, we get

$$\int_\Omega |u|^p dx + \lambda p(p-1)\int_\Omega \beta'(u)|u|^{p-2}|\nabla u|^2 dx \le \int_\Omega |f|^p dx.$$

Now, using the identity

$$|u|^{p+\alpha-3}|\nabla u|^2 = \frac{4}{(p+\alpha-1)^2}\left|\nabla|u|^{(p+\alpha-1)/2}\right|^2, \quad \text{a.e. in } \Omega$$

and condition (5.103), we get

$$\int_\Omega |u|^p dx + \frac{4\lambda p(p-1)}{(p+\alpha-1)^2}\int_\Omega \left|\nabla|u|^{(p+\alpha-1)/2}\right|^2 dx \le C\int_\Omega |f|^p dx. \tag{5.108}$$

On the other hand, by the Sobolev embedding theorem

$$\int_\Omega \left|\nabla|u|^{(p+\alpha-1)/2}\right|^2 dx \le C\left(\int_\Omega |u|^{(p+\alpha-1)N/(N-2)}dx\right)^{(N-2)/N} \quad \text{if } N>2,$$

and

$$\int_{\Omega} \left|\nabla |u|^{(p+\alpha-1)/2}\right|^2 dx \le C \left(\int_{\Omega} |u|^{(p+\alpha-1)/q} dx\right)^{1/q}, \qquad \forall q > 1,$$

for $N = 2$. Then, substituting these inequalities into (5.108), we get (5.105) and (5.106), respectively.

We set $J_\lambda = (I + \lambda A)^{-1}$ and

$$\varphi(u) = \|u\|_p^p, \qquad \psi(u) = C\|u\|_{(p+\alpha-1)N/(N-2)}^{p+\alpha-1}.$$

Then, inequality (5.105) can be written as

$$\varphi(J_\lambda f) + \lambda\,\psi(J_\lambda f) \le \varphi(f), \qquad \forall f \in L^p(\Omega).$$

This yields

$$\varphi(J_\lambda^k f) + \lambda\,\psi(J_\lambda^k f) = \varphi(J_\lambda^{k-1}), \qquad \forall k.$$

Summing these equations from $k = 1$ to $k = n$, and taking $\lambda = t/n$, yields

$$\varphi(J_{t/n}^n f) + \sum_{k=1}^{n} \frac{1}{n}\ \psi(J_{t/n}^k f) = \varphi(f).$$

Recalling that, by Theorem 4.3, $J_{t/n}^n f \to S(t)$ for $n \to \infty$, the latter equation implies that

$$\varphi(S(t)f) + \int_0^t \psi(S(\tau)f)d\tau = \varphi(f), \qquad \forall t \ge 0. \tag{5.109}$$

In particular, it follows that the function $t \to \varphi(S(t)f)$ is decreasing and so is $t \to \psi(S(t)f)$. Then, by (5.109), we see that $\varphi(S(t)f) + t\psi(S(t)f) \le \varphi(f)$, $\forall t > 0$; that is,

$$\|S(t)f\|_p^p + Ct\|S(t)f\|_{(p+\alpha-1)N/(N-2)}^{p+\alpha-1} \le \|f\|_p^p, \qquad \forall t > 0, \tag{5.110}$$

where C is independent of p and f.

Let p_n be inductively defined by

$$p_{n+1} = (p_n + \alpha - 1)\,\frac{N}{N-2}.$$

Then, by (5.110), we see that

$$\|S(t_{n+1})f\|_{p_{n+1}}^{(N/(N-2))p_{n+1}} \le \frac{\|S(t_n)f\|_{p_n}^{p_n}}{C(t_{n+1} - t_n)},$$

where $t_0 = 0$ and $t_{n+1} > t_n$. Choosing $t_{n+1} - t_n = t/(2^{n+1})$, we get after some calculation that

$$\limsup_{n\to\infty} \|S(t)f\|_{p_{n+1}}^{((N-2)/N)np_{n+1}} \le C\|f\|_{p_0} \left(\frac{2}{t}\right)^{\mu}, \qquad \forall t > 0,$$

where $\mu = N/2$, because p_n is given by

$$p_n = \left(\frac{N}{N-2}\right)^n p_0 + \frac{N\alpha}{2(N-2)}\left(\left(\frac{N}{N-2}\right)^n - 1\right)$$

(here, we have used the fact that $\alpha > (N-2)/N$), we get the final estimate

$$\|S(t)f\|_\infty \le C\|f\|_{p_0}^{2p_0/(2p_0+N(\alpha-1))} t^{-(N/(2p_0+N(\alpha-1)))}, \qquad \forall p_0 \ge 1,$$

as claimed.

The case $N = 2$ follows similarly. Moreover, by inequality (5.105) and the exponential formula defining $S(t)$, it follows that

$$\|S(t)f\|_p \le \|f\|_p, \qquad \forall p \in L^p(\Omega), \qquad t \ge 0.$$

This completes the proof of Theorem 5.4. □

The Porous Media Equation in $\mathbf{R}^N$

Consider now equation (5.87) in $\Omega = \mathbf{R}^N$, for $N = 1,2,3$:

$$\begin{cases} \dfrac{\partial y}{\partial t} - \Delta\beta(y) \ni f & \text{in } \mathbf{R}^N \times (0,T), \\ y(0,x) = y_0(x), & x \in \mathbf{R}^N, \\ \beta(y(t)), y(t) \in L^1(\mathbf{R}^n), & \forall t \in [0,T]. \end{cases} \tag{5.111}$$

where $\partial/\partial t$ and Δ are taken in the sense of distributions on $(0,T) \times \mathbf{R}^N$ (see (5.88)). We may rewrite equation (5.111) in the form (5.83) on the space $X = L^1(\mathbf{R}^N)$, where

$$Ay = \{-\Delta w;\ w(x) \in \beta(y(x)),\ \text{a.e. } x \in \Omega,\ w, \Delta w \in L^1(\mathbf{R}^N)\},\ \forall y \in D(A),$$
$$D(A) = \{y \in L^1(\mathbf{R}^N);\ \exists w \in L^1(\mathbf{R}^N),\ \Delta w \in L^1(\mathbf{R}^N),\ w(x) \in \beta(y(x)),\ \text{a.e. } x \in \mathbf{R}^N\},$$

where Δw is taken in the sense of distributions. Here β is a maximal monotone graph in $\mathbf{R} \times \mathbf{R}$ such that $0 \in \beta(0)$ and $0 \in \operatorname{int} D(\beta)$ if $N = 1,2$. Then, as shown earlier in Theorem 3.7, A is m-accretive in $L^1(\mathbf{R}^N) \times \mathbf{R}^N$ and so, by Theorem 4.1, we obtain the following.

Proposition 5.12. *Assume that $f \in L^1(0,T;L^1(\mathbf{R}^N))$ and $y_0 \in L^1(\mathbf{R}^N)$ is such that $\exists w \in L^1(\mathbf{R}^N)$, $\Delta w \in L^1(\mathbf{R}^N)$, $w(x) \in \beta(y_0(x))$, a.e. $x \in \mathbf{R}^N$. Then, problem* (5.111) *has a unique mild solution $y \in C([0,T];L^1(\mathbf{R}^N))$.*

Remark 5.3. The continuity of solutions to (5.111) with respect to φ is studied in the work of Bénilan and Crandall [9]. In this context, we mention also the work of Brezis and Crandall [16] and Alikakos and Rostamian [1].

Localization of Solutions to Porous Media Equations

A nice feature of solutions to the porous media equation are finite time extinction for the fast diffusion equation (i.e., $\beta(y) = y^\alpha$, $0 < \alpha < 1$), and propagation with finite velocity for the low diffusion equation (i.e., $1 < \alpha < \infty$). We refer the reader to the work of Pazy [36] and to the recent book of Antontsev, Diaz, and Shmarev [2] for detailed treatment of this phenomena. (See also the Vasquez monograph [40] for a detailed study of the localization of solutions to a porous media equation.) Here, we briefly discuss the extinction in finite time.

Proposition 5.13. *Let $y \in C([0,\infty); L^1(\Omega) \cap H^{-1}(\Omega))$ be the solution to equation*

$$\frac{\partial y}{\partial t} - \mu\Delta(|y|^\alpha \operatorname{sign} y) = 0 \quad \text{in } \ \Omega \times (0,\infty), \tag{5.112}$$

where $y_0 \in H^{-1}(\Omega) \cap L^1(\Omega)$, $\mu > 0$, $0 < \alpha < 1$ if $N = 1,2$ and $1/5 \le \alpha < 1$ if $N = 3$. Then,

$$y(x,t) = 0 \quad \text{for } \ t \ge T(y_0),$$

where

$$T(y_0) = \frac{|y_0|_{-1}^{1-\alpha}}{\mu\gamma^{1+\alpha}}\,\cdot$$

If $\alpha = 0$ and $N = 1$, then $y(x,t) = 0$ for $t \ge (|y_0|_{-1})/\mu\gamma$.

Proof. Assume first that $N > 1$. As seen earlier, the equation has a unique smooth solution $y \in W^{1,2}([0,T]; H^{-1}(\Omega))$ for each $T > 0$. Multiplying scalarly in $H^{-1}(\Omega)$ equation (5.112) by y and integrating on $(0,T)$, we obtain

$$\frac{1}{2}\,\frac{d}{dt}\,|y(t)|_{-1}^2 + \mu\int_\Omega |y(s,x)|^{\alpha+1}dx = 0, \qquad \forall t \ge 0.$$

Now, by the Sobolev embedding theorem (see Theorem 1.4), we have

$$\gamma|y(s)|_{-1} \le |y(s)|_{L^{\alpha+1}(\Omega)} \quad \text{for all } \alpha > 0 \text{ if } N = 1,2 \text{ and for } \alpha \ge \frac{N-2}{N+2} \text{ if } N \ge 3.$$

(Here, $|\cdot|_{-1}$ is the $H^{-1}(\Omega)$ norm.) This yields

$$\frac{d}{dt}\,|y(t)|_{-1}^2 + 2\mu\gamma^{\alpha+1}|y(t)|_{-1}^{\alpha+1} \le 0, \qquad \forall t \ge 0,$$

and therefore

$$\frac{d}{dt}\,|y(t)|_{-1}^{1-\alpha} + \mu\gamma^{1+\alpha} \le 0, \quad \text{a.e. } t > 0.$$

Hence,

$$|y(t)|_{-1} = 0 \quad \text{ for } t \ge \frac{|y_0|_{-1}^{1-\alpha}}{\mu\gamma^{1+\alpha}}.$$

If $N = 1$, then, multiplying scalarly in $H^{-1}(\Omega)$ equation (5.112) by $y(t)$, we get

$$\frac{1}{2}\frac{d}{dt}|y(t)|_{-1}^2+\mu|y(t)|_{L^1(\Omega)}\le 0,\quad \text{a.e. } t>0.$$

This yields (we have $|y|_{L^1(\Omega)}\ge\gamma|y_0|_{-1}$):

$$|y(t)|_{-1}+\mu\gamma t\le|y_0|_{-1},\qquad \forall t\ge 0$$

and, therefore,

$$|y(t)|_{-1}=0\quad \text{for } t\ge\frac{|y_0|_{-1}}{\mu\gamma}.\qquad\square$$

Remark 5.4. The extinction in finite time is a significant nonlinear behavior of solutions to fast diffusion porous media equations and this implies that the diffusion process reaches its critical state (which is zero in this case) in finite time. The case $\alpha=0$ models an important class of diffusion processes with self-organized criticality, the so-called Bak's sand-pile model.

5.4 The Phase Field System

Consider the parabolic system

$$\begin{cases}\dfrac{\partial}{\partial t}\theta(t,x)+\ell\dfrac{\partial\varphi}{\partial t}(t,x)-k\Delta\theta(t,x)=f_1(t,x), & \text{in } Q=\Omega\times(0,T),\\ \dfrac{\partial}{\partial t}\varphi(t,x)-\alpha\Delta\varphi(t,x)-\kappa(\varphi(t,x)-\varphi^3(t,x)) & \\ \qquad\qquad\qquad +\delta\theta(t,x)=f_2(t,x), & \text{in } Q,\\ \theta(0,x)=\theta_0(x),\qquad \varphi(0,x)=\varphi_0(x), & x\in\Omega,\\ \theta=0,\qquad \varphi=0, & \text{on } \partial\Omega\times(0,T),\end{cases}\tag{5.113}$$

where $\ell,k,\alpha,\kappa,\delta$ are positive constants. This system, called in the literature the *phase-field system*, was introduced as a model of a phase transition process in physics and, in particular, the melting and solidification phenomena. (See Caginalp [18].) In this latter case, $\theta=\theta(t,x)$ is the temperature, whereas φ is the phase-field transition function. The two-phase Stefan problem presented above can be viewed as a particular limit case of this model. In fact, it can be obtained from the two-phase Stefan model of phase transition by the following heuristic argument.

As seen earlier, the two-phase Stefan problem (5.74) and (5.75) can be rewritten as

$$\frac{\partial}{\partial t}\gamma(\theta)-\Delta K(\theta)=f\quad \text{in } \mathscr{D}'(\Omega\times(0,T)),$$

where γ is the multivalued graph (5.71); that is, $\gamma=C+\rho H$. Equivalently,

$$\frac{\partial}{\partial t}\varphi(\theta)\theta-\Delta K(\theta)=f\quad \text{in } \mathscr{D}'(\Omega\times(0,T)),\tag{5.114}$$

where $\varphi : \mathbf{R} \to \mathbf{R}$ is given by the graph

$$\varphi(\theta) = \begin{cases} C_1 & \text{if } \theta < 0, \\ C_2 + \dfrac{\rho}{\theta} & \text{if } \theta > 0. \end{cases} \tag{5.115}$$

The idea behind Caginalp's model of phase transition is to replace the multivalued graph φ by a function $\varphi = \varphi(t,x)$, called the *phase function* and equation (5.114) by

$$\varphi \frac{\partial \theta}{\partial t} + \theta \frac{\partial \varphi}{\partial t} - \Delta K(\theta) = f. \tag{5.116}$$

The phase function φ should be interpreted as a measure of phase transition and more precisely as the proportion related to the first phase and the second one. For instance, in the case of liquid–solid transition, one has, formally, $\varphi \geq 1$ in the liquid zone $\{(t,x);\ u(t,x) > 0\}$ and $\varphi < 0$ in the solid zone $\{(t,x);\ u(t,x) < 0\}$. In general, however, φ remains in an interval $[\varphi_*, \varphi^*]$ which is determined by the specific physical model. This is the reason why φ is taken as the solution to a parabolic equation of the Ginzburg–Landau type

$$\frac{\partial \varphi}{\partial t} - \alpha \Delta \varphi - \kappa(\varphi - \varphi^3) + \delta \theta = f_2, \tag{5.117}$$

which is the basic mathematical model of phase transition. Equations (5.116) and (5.117) lead, after further simplifications, to system (5.113).

As regards the existence in problem (5.113), we have the following.

Theorem 5.5. *Assume that* $\varphi_0, \theta_0 \in H_0^1(\Omega) \cap H^2(\Omega)$, $\Omega \subset \mathbf{R}^N$, $N = 1,2,3$, *and that* $f_1, f_2 \in W^{1,2}([0,T]; L^2(\Omega))$. *Then, there is a unique solution* (θ, φ) *to system* (5.113) *satisfying*

$$(\theta, \varphi) \in (W^{1,\infty}([0,T]; L^2(\Omega)))^2 \cap (L^\infty(0,T; H_0^1(\Omega) \cap H^2(\Omega)))^2. \tag{5.118}$$

Proof. We set $y = \theta + \ell\varphi$ and reduce system (5.113) to

$$\begin{cases} \dfrac{\partial}{\partial t} y - k\Delta y + k\ell\Delta\varphi = f_1 & \text{in } Q, \\ \dfrac{\partial}{\partial t} \varphi - \alpha\Delta\varphi - \kappa(\varphi - \varphi^3) + \delta(y - \ell\varphi) = f_2 & \text{in } Q, \\ y(0) = y_0 = \theta_0 + \ell\varphi_0, \qquad \varphi(0) = \varphi_0 \text{ in } \Omega, \qquad y = \varphi = 0 & \text{on } \Sigma. \end{cases} \tag{5.119}$$

In the space $X = L^2(\Omega) \times L^2(\Omega)$ consider the operator $A : X \to X$,

$$A \begin{pmatrix} y \\ \varphi \end{pmatrix} = \begin{pmatrix} -k\Delta y + k\ell\Delta\varphi \\ -\alpha\Delta\varphi - \kappa(\varphi - \varphi^3) + \delta(y - \ell\varphi) \end{pmatrix}$$

with the domain $D(A) = \{(y,\varphi) \in (H^2(\Omega) \cap H_0^1(\Omega))^2;\ \varphi \in L^6(\Omega)\}$. Then, system (5.119) can be written as

$$\begin{cases} \dfrac{d}{dt}\begin{pmatrix} y \\ \varphi \end{pmatrix} + A\begin{pmatrix} y \\ \varphi \end{pmatrix} = \begin{pmatrix} f_1 \\ f_2 \end{pmatrix}, & t \in (0,T), \\ \begin{pmatrix} y \\ \varphi \end{pmatrix}(0) = \begin{pmatrix} y_0 \\ \varphi_0 \end{pmatrix}. \end{cases} \tag{5.120}$$

In order to apply Theorem 4.4 to (5.120), we check that A is quasi-m-accretive in X. To this aim we endow the space $X = L^2(\Omega) \times L^2(\Omega)$ with an equivalent Hilbertian norm provided by the scalar product

$$\left\langle \begin{pmatrix} y \\ \varphi \end{pmatrix}, \begin{pmatrix} \widetilde{y} \\ \widetilde{\varphi} \end{pmatrix} \right\rangle = a(y,\widetilde{y})_{L^2(\Omega)} + (\varphi,\widetilde{\varphi})_{L^2(\Omega)},$$

where $a = \alpha/k\ell^2$. Then, as easily seen, we have

$$\left\langle A\begin{pmatrix} y \\ \varphi \end{pmatrix} - A\begin{pmatrix} y^* \\ \varphi^* \end{pmatrix}, \begin{pmatrix} y \\ \varphi \end{pmatrix} - \begin{pmatrix} y^* \\ \varphi^* \end{pmatrix} \right\rangle$$
$$\geq \eta(\|\nabla(y-y^*)\|^2_{L^2(\Omega)} + \|\nabla(\varphi-\varphi^*)\|^2_{L^2(\Omega)}) - \omega(\|y-y^*\|^2_{L^2(\Omega)} + \|\varphi-\varphi^*\|^2_{L^2(\Omega)}),$$

for some $\omega, \eta > 0$. Clearly, this implies that A is quasi-accretive; that is, $A + \omega I$ is accretive.

Now, consider for $g_1, g_2 \in L^2(\Omega)$ the equation

$$\lambda\begin{pmatrix} y \\ \varphi \end{pmatrix} + A\begin{pmatrix} y \\ \varphi \end{pmatrix} = \begin{pmatrix} g_1 \\ g_2 \end{pmatrix}; \tag{5.121}$$

that is,

$$\begin{cases} \lambda y - k\Delta y + k\ell\Delta\varphi = g_1 & \text{in } \Omega, \\ \lambda\varphi - \alpha\Delta\varphi - \kappa(\varphi - \varphi^3) + \delta(y - \ell\varphi) = g_2, & \\ y = \varphi = 0 & \text{on } \partial\Omega. \end{cases} \tag{5.122}$$

System (5.122) can be equivalently rewritten as

$$\begin{pmatrix} \lambda y \\ (\lambda - \kappa - \ell\delta)\varphi + \delta y \end{pmatrix} + A_0\begin{pmatrix} y \\ \varphi \end{pmatrix} + F\begin{pmatrix} y \\ \varphi \end{pmatrix} = \begin{pmatrix} q_1 \\ q_2 \end{pmatrix}, \tag{5.123}$$

where $F, A_0 : L^2(\Omega) \times L^2(\Omega) \to L^2(\Omega) \times L^2(\Omega)$ are given by

$$A_0 \begin{pmatrix} y \\ \varphi \end{pmatrix} = \begin{pmatrix} -k\Delta y + k\ell\Delta\varphi \\ -\alpha\Delta\varphi \end{pmatrix}$$

$$D(A_0) = (H^2(\Omega) \times H_0^1(\Omega))^2$$

and

$$F \begin{pmatrix} y \\ \varphi \end{pmatrix} = \begin{pmatrix} 0 \\ \kappa\varphi^3 \end{pmatrix}$$

$$D(F) = L^2(\Omega) \times L^6(\Omega).$$

By the Lax–Milgram lemma (Lemma 1.3), it is easily seen that A_0 is m-accretive and coercive in $X = L^2(\Omega) \times L^2(\Omega)$. On the other hand, F is quasi-m-accretive and

$$\left\langle A_0 \begin{pmatrix} y \\ \varphi \end{pmatrix}, F \begin{pmatrix} y \\ \varphi \end{pmatrix} \right\rangle \geq 0, \qquad \forall \begin{pmatrix} y \\ \varphi \end{pmatrix} \in D(A_0).$$

Hence, by Proposition 3.8, $A_0 + F$ is quasi-m-accretive and this implies that (5.123) has a solution for λ sufficiently large. □

Remark 5.5. The liquid and solid regions in the case of a melting solidification problem are those that remain invariant by the flow $t \to (\theta(t), \varphi(t))$. This is one way of determining in specific physical models the range interval $[\varphi_*, \varphi^*]$ of phase-field function φ. A more general nonlinear phase-field model is proposed and studied by Bonetti, Colli, Fabrizio, and Gilardi [12] in connection with a phase transition model proposed by Fremond [26]. More precisely, under our notation this system is of the following form

$$\begin{cases} \dfrac{\partial u}{\partial t} - \dfrac{\partial}{\partial t}(G(\varphi)) - \lambda\Delta \log u = f, \\ \mu \dfrac{\partial \varphi}{\partial t} - \nu\Delta\varphi + F'(\varphi) + uG'(\varphi) = 0, \end{cases}$$

and the above functional treatment applies as well to this general problem.

5.5 The Equation of Conservation Laws

We consider here the Cauchy problem

$$\begin{cases} \dfrac{\partial y}{\partial t} + \displaystyle\sum_{i=1}^{N} \dfrac{\partial}{\partial x_i} a_i(y) = 0 & \text{in } \mathbf{R}^N \times \mathbf{R}^+, \\ y(x,0) = y_0(x), & x \in \mathbf{R}^N, \end{cases} \tag{5.124}$$

where $a = (a_1, ..., a_N)$ is a continuous map from $\mathbf{R}$ to $\mathbf{R}^N$ satisfying the condition

$$\limsup_{|r|\to 0} \frac{\|a(r)\|}{|r|} < \infty,$$

and $y_0 \in L^1(\mathbf{R}^N)$.

This equation can be treated as a nonlinear Cauchy problem in the space $X = L^1(\mathbf{R}^N)$. In fact, we have seen earlier (Theorem 3.8) that the first-order differential operator $y \to \sum_{i=1}^N (\partial/\partial x_i)\, a_i(y)$ admits an m-accretive extension $A \subset L^1(\mathbf{R}^N) \times L^1(\mathbf{R}^N)$ defined as the closure in $L^1(\mathbf{R}^N) \times L^1(\mathbf{R}^N)$ of the operator A_0 given by Definition 3.2.

Then, by Theorem, 4.3, the Cauchy problem

$$\begin{cases} \dfrac{dy}{dt} + Ay \ni 0 \quad \text{ in } \ (0,+\infty), \\ \\ y(0) = y_0, \end{cases}$$

has for every $y_0 \in \overline{D(A)}$ a unique mild solution $y(t) = S(t)y_0$ given by the exponential formula (4.17) or, equivalently,

$$y(t) = \lim_{\varepsilon \to 0} y_\varepsilon(t) \quad \text{uniformly on compact intervals,}$$

where y_ε is the solution to difference equation

$$\begin{aligned} \varepsilon^{-1}(y_\varepsilon(t) - y_\varepsilon(t-\varepsilon)) + Ay_\varepsilon(t) &= 0 \qquad \text{for } \ t > \varepsilon, \\ y_\varepsilon(t) &= y_0 \qquad \text{for } \ t < 0. \end{aligned} \tag{5.125}$$

We call such a function $y(t) = S(t)y_0$ a *semigroup solution* or *mild solution* to the Cauchy problem (5.124).

We see in Theorem 5.6 below that this solution is in fact an entropy solution to the equation of conservation laws.

Theorem 5.6. *Let $y = S(t)y_0$ be the semigroup solution to problem* (5.124). *Then,*

(i) $S(t)L^p(\mathbf{R}^N) \subset L^p(\mathbf{R}^N)$ *for all* $1 \le p < \infty$ *and*

$$\|S(t)y_0\|_{L^p(\mathbf{R}^N)} \le \|y_0\|_{L^p(\mathbf{R}^N)}, \qquad \forall y_0 \in \overline{D(A)} \cap L^p(\mathbf{R}^N). \tag{5.126}$$

(ii) *If* $y_0 \in \overline{D(A)} \cap L^\infty(\mathbf{R}^N)$, *then*

$$\begin{aligned} &\int_0^T \int_{\mathbf{R}^N} \big(|y(x,t) - k|\varphi_t(x,t) \\ &+ \operatorname{sign}_0(y(x,t)-k)(a(y(x,t)) - a(k)) \cdot \varphi_x(x,t)\big)\,dx\,dt \ge 0 \end{aligned} \tag{5.127}$$

for every $\varphi \in C_0^\infty(\mathbf{R}^N \times (0,T))$ *such that* $\varphi \ge 0$, *and all* $k \in \mathbf{R}^N$ *and* $T > 0$.

Here $\varphi_t = \partial\varphi/\partial t$ and $\varphi_x = \nabla_x \varphi$.

Inequality (5.127) is Kruzkhov's [30] definition of entropy solution to the Cauchy problem (5.124) and its exact significance is discussed below.

Proof of Theorem 5.6. Because, as seen in the proof of Theorem 3.8, $(I+\lambda A)^{-1}$ maps $L^p(\mathbf{R}^N)$ into itself and

$$\|(I+\lambda A)^{-1}u\|_{L^p(\mathbf{R}^N)} \le \|u\|_{L^p(\mathbf{R}^N)}, \qquad \forall \lambda > 0,\ u \in L^p(\mathbf{R}^N) \text{ for } 1 \le p \le \infty,$$

we deduce (i) by the exponential formula (4.17).

To prove inequality (5.126), consider the solution y to equation (5.125), where $y_0 \in L^1(\mathbf{R}^N) \cap L^\infty(\mathbf{R}^N)$ and $A_0 = A$. (Recall that $L^1(\mathbf{R}^N) \cap L^\infty(\mathbf{R}^N) \subset R(I+\lambda A)^{-1}$ for all $\lambda > 0$.) Then, $\|y_\varepsilon(t)\|_{L^p(\mathbf{R}^N)} \le \|y_0\|_{L^p(\mathbf{R}^N)}$ for $p = 1, \infty$ and so, by Definition 3.2 and by (5.125), we have

$$\begin{aligned} &\int_{\mathbf{R}^N} (\operatorname{sign}_0(y_\varepsilon(x,t)-k)(a(y_\varepsilon(x,t))-a(k))) \cdot \varphi_x(x,t) \\ &+\varepsilon(y_\varepsilon(x,t-\varepsilon)-y_\varepsilon(x,t))\operatorname{sign}_0(y_\varepsilon(x,t)-k)\varphi(x,t))dx \ge 0, \\ &\qquad \forall k \in \mathbf{R},\ \varphi \in C_0^\infty(\mathbf{R}^N \times (0,T)),\ \varphi \ge 0,\ t \in (0,T). \end{aligned} \tag{5.128}$$

On the other hand, we have

$$\begin{aligned} &(y_\varepsilon(x,t-\varepsilon)-y_\varepsilon(x,t))\operatorname{sign}_0(y_\varepsilon(x,t)-k) \\ &= (y_\varepsilon(x,t-\varepsilon)-k)\operatorname{sign}_0(y_\varepsilon(x,t)-k) - (y_\varepsilon(x,t)-k)\operatorname{sign}_0(y_\varepsilon(x,t)-k) \\ &\le z_\varepsilon(x,t-\varepsilon) - z_\varepsilon(x,t), \end{aligned}$$

where $z_\varepsilon(x,t) = |y_\varepsilon(x,t)-k|$.

Substituting the latter into (5.128) and integrating on $\mathbf{R}^N \times [0,T]$, we get

$$\begin{aligned} &\int_0^T \int_{\mathbf{R}^N} (\operatorname{sign}_0(y_\varepsilon(x,t)-k)(a(y_\varepsilon(x,t))-a(k)) \cdot \varphi_x(x,t) \\ &+\varepsilon^{-1}(z_\varepsilon(x,t-\varepsilon)-z_\varepsilon(x,t))\varphi(x,t))dxdt \ge 0. \end{aligned}$$

This yields

$$\begin{aligned} &\int_0^T \int_{\mathbf{R}^N} (\operatorname{sgn}_0(y_\varepsilon(x,t)-k)(a(y_e(x,t))-a(k)) \cdot \varphi_x(x,t))dxdt \\ &-\varepsilon^{-1}\int_0^\varepsilon \int_{\mathbf{R}^N} |y_\varepsilon(x,t)-k|\varphi(x,t)dxdt + \varepsilon^{-1}\int_0^T \int_{\mathbf{R}^N} z_\varepsilon(x,t)\varphi(x,t)dxdt \\ &+\varepsilon^{-1}\int_{T-\varepsilon}^T \int_{\mathbf{R}^N} z_\varepsilon(x,t)(\varphi(x,t+\varepsilon)-\varphi(x,t))dxdt \ge 0. \end{aligned}$$

Now, letting ε tend to zero, we get (5.127) because $y_\varepsilon(t) \to y(t)$ uniformly on $[0,T]$ in $L^1(\mathbf{R}^N)$ and $\varepsilon^{-1}(z_\varepsilon(x,t-\varepsilon)-z_\varepsilon(x,t)) \to |y(x,t)-k|$. This completes the proof of Theorem 5.5.

As mentioned earlier, equation (5.124) is known in the literature as the *equation of conservation laws* and has a large spectrum of applications in mechanics and was extensively studied in recent years. A function $\eta : \mathbf{R} \to \mathbf{R}$ is called an *entropy* of system (5.124) if there is a function $q : \mathbf{R} \to \mathbf{R}^n$ (the *entropy flux* associated with entropy η) such that $\nabla^2 q \geq 0$ and

$$\nabla q_j(y) = \nabla \eta(y) \cdot \nabla a_j(y), \qquad \forall y \in \mathbf{R}^N,\ j = 1,...,N.$$

(Such a pair (η, q) is called an *entropy pair*.)

The bounded measurable function $y : [0,T] \times \mathbf{R}^N \to \mathbf{R}$ is called an *entropy solution* to (5.124) if, for all convex entropy pairs (η, q),

$$\frac{\partial}{\partial t}\eta(y(t,x)) + \mathrm{div}_x\, q(y(t,x)) \leq 0 \quad \text{in } \mathscr{D}'(\mathbf{R}^N \times (0,T));$$

that is,

$$\int_0^T \int_{\mathbf{R}^N} (\eta(y(t,x))\varphi_t(t,x) + q(y(t,x)) \cdot \varphi_x(t,x))dtdx \geq 0$$

for all $\varphi \in C_0^\infty((0,T) \times \mathbf{R}^N),\ \varphi \geq 0$.

If take $\eta(y) \equiv |y-k|$ and $q(y) \equiv \mathrm{sign}_0(y-k)(a(y)-a(k))$, we see that y satisfies equation (5.127). The existence and uniqueness of the entropy solution were proven by S. Kruzkhov [30]. (See also Bénilan and Kruzkhov [11] for some recent results.) Recalling that the resolvent $(I+\lambda A)^{-1}$ of the operator A can be approximated by the family of approximating equation (3.74), one might deduce via the Trotter–Kato Theorem 4.14 that the entropy solution y can also be obtained as the limit for $\varepsilon \to 0$ to solutions y_ε to the parabolic nonlinear equation

$$\frac{\partial y}{\partial t} - \varepsilon \Delta y + (a(y))_x = 0,$$

in $\mathbf{R}^N$ which is related to Hopf's viscosity solution approach to nonlinear conservation laws equations.

5.6 Semilinear Wave Equations

The linear wave equation perturbed by a nonlinear term in speed can be conveniently written as a first order differential equation in an appropriate Hilbert space defined below and treated so by the general existence theory developed in Chapter 4.

We are given two real Hilbert spaces V and H such that $V \subset H \subset V'$ and the inclusion mapping of V into H is continuous and densely defined. We have denoted by V' the dual of V and H is identified with its own dual. As usual, we denote by $\|\cdot\|$ and $|\cdot|$ the norms of V and H, respectively, and by $(\cdot,\cdot)$ the duality pairing between V and V' and the scalar product of H.

We consider the second-order Cauchy problem

$$\frac{d^2y}{dt^2} + Ay + B\left(\frac{dy}{dt}\right) \ni f, \qquad y(0) = y_0, \qquad \frac{dy}{dt}(0) = y_1, \tag{5.129}$$

where A is a linear continuous and symmetric operator from V to V' and $B \subset V \times V'$ is maximal monotone operator. We assume further that

$$(Ay, y) + \alpha|y|^2 \geq \omega\|y\|^2, \qquad \forall y \in V, \tag{5.130}$$

where $\omega > 0$ and $\alpha \in \mathbf{R}$.

One principal motivation and model for equation (5.129) is the nonlinear hyperbolic boundary value problem

$$\begin{cases} \dfrac{\partial^2 y}{\partial t^2} - \Delta y + \beta\left(\dfrac{\partial y}{\partial t}\right) \ni f(x,t) & \text{in } \Omega \times (0,T), \\ y = 0 & \text{on } \partial\Omega \times (0,T), \\ y(x,0) = y_0(x), \qquad \dfrac{dy}{dt}(x,0) = y_1(x) & \text{in } \Omega, \end{cases} \tag{5.131}$$

where β is a maximal monotone graph in $\mathbf{R} \times \mathbf{R}$ and Ω is a bounded open subset of $\mathbf{R}^N$ with a smooth boundary.

As regards problem (5.129), we have the following existence result.

Theorem 5.7. *Let $f \in W^{1,1}([0,T];H)$ and $y_0 \in V$, $y_1 \in D(B)$ be given such that*

$$\{Ay_0 + By_1\} \cap H \neq \emptyset. \tag{5.132}$$

Then, there is a unique function $y \in W^{1,\infty}([0,T];V) \cap W^{2,\infty}([0,T];H)$ that satisfies

$$\begin{cases} \dfrac{d^+}{dt}\left(\dfrac{dy}{dt}\right)(t) + Ay(t) + B\left(\dfrac{d^+}{dt}\, y(t)\right) \ni f(t), \qquad \forall t \in [0,T], \\ y(0) = y_0, \qquad \dfrac{dy}{dt}(0) = y_1, \end{cases} \tag{5.133}$$

where $d^+/dt(dy/dt)$ is considered in the topology of H and $(d^+/dt)y$ in V.

Proof. Let $X = V \times H$ be the Hilbert space with the scalar product

$$\langle U_1, U_2\rangle = (Au_1, u_2) + \alpha(u_1, u_2) + (v_1, v_2),$$

where $U_1 = [u_1, v_1]$, $U_2 = [u_2, v_2]$.

In the space X, define the operator $\mathscr{A} : D(\mathscr{A}) \subset X \to X$ by

$$\begin{cases} D(\mathscr{A}) = \{[u,v] \in V \times H; \{Au + Bv\} \cap H \neq \emptyset\}, \\ \mathscr{A}[u,v] = [-v; \{Au + Bv\} \cap H] + \sigma[u,v], \;\; [u,v] \in D(\mathscr{A}), \end{cases} \tag{5.134}$$

where

$$\sigma = \sup\left\{ \frac{\alpha(u,v)}{((Au,u)+\alpha|u|^2+|v|^2)};\ u \in V,\ v \in H \right\}.$$

We may write equation (5.129) as a first-order differential system

$$\begin{cases} \dfrac{dy}{dt} - z = 0 \quad \text{in } (0,T), \\ \dfrac{dy}{dt} + Ay + Bz \ni f. \end{cases}$$

Equivalently,

$$\begin{cases} \dfrac{dt}{dt} U(t) + \mathscr{A}U(t) - \sigma U(t) \ni F(t), \qquad t \in (0,T), \\ U(0) = U_0, \end{cases} \tag{5.135}$$

where

$$U(t) = [y(t), z(t)], \qquad F(t) = [0, f(t)], \qquad U_0 = [y_0, y_1].$$

It is easily seen that $\mathscr{A}$ is monotone in $X \times X$. Let us show that it is maximal monotone; that is, $R(I+\mathscr{A}) = V \times H$, where I is the unity operator in $V \times H$. To this end, let $[g,h] \in V \times H$ be arbitrarily given. Then, the equation $U + \mathscr{A}U \ni [g,h]$ can be written as

$$\begin{cases} y - z + \sigma y = g, \\ z + Ay + Bz + \sigma z \ni h. \end{cases}$$

Substituting $y = (1+\sigma)^{-1}(z+g)$ in the second equation, we obtain

$$(1+\sigma)z + (1+\sigma)^{-1}Az + Bz \ni h - (1+\sigma)^{-1}Ag.$$

Under our assumptions, the operator $z \xrightarrow{\Gamma} (1+\sigma)z + (1-\sigma)^{-1}Az$ is continuous, positive, and coercive from V to V'. Then, $R(\Gamma + B) = V'$ (see Corollary 2.6, and so the previous equation has a solution $z \in D(B)$ and a fortiori $[g,h] \in R(I+\mathscr{A})$.

Then, the conclusions of Theorem 5.7 follow by Theorem 4.6 because there is a unique solution $U \in W^{1,\infty}([0,T]; V \times H)$ to problem (5.135) satisfying

$$\frac{d^+}{dt} U(t) + \mathscr{A}U(t) - \sigma U(t) \ni F(t), \qquad \forall t \in [0,T):$$

$$\begin{cases} \dfrac{d^+}{dt} y(t) = z(t), & \forall t \in [0,T), \\ \dfrac{d^+}{dt} z(t) + Ay(t) + B(z(t)) \ni f(t), & \forall t \in [0,T), \end{cases}$$

where $(d^+/dt)y$ is in the topology of V whereas $(d^+/dt)z$ is in the topology of H. □

The operator B that arises in equation (5.129) might be multivalued. Moreover, if $B = \partial\varphi$, where $\varphi : V \to \overline{\mathbf{R}}$ is a lower semicontinuous convex function, problem (5.129) reduces to a variational inequality of hyperbolic type.

In order to apply Theorem 5.7 to the hyperbolic problem (5.131), we take $V = H_0^1(\Omega)$, $H = L^2(\Omega)$, $V' = H^{-1}(\Omega)$, $A = -\Delta$, and $B : H_0^1(\Omega) \to H^{-1}(\Omega)$ defined by $B = \partial\varphi$, where $\varphi : H_0^1(\Omega) \to \overline{\mathbf{R}}$ is the function

$$\varphi(y) = \int_\Omega j(y(x))dx, \qquad \forall y \in H_0^1(\Omega),\ \beta = \partial j. \tag{5.136}$$

The operator B is an extension of the operator $(B_0 y)(x) = \{w \in L^2(\Omega);\ w(x) \in \beta(y(x)),\ \text{a.e. } x \in \Omega\}$, from $H_0^1(\Omega)$ to $H^{-1}(\Omega)$. It should be said that, in general, the operator B does not coincide with B_0. The simplest example is $j(r) = 0$ if $0 \le r \le 1$, $j(r) = +\infty$ otherwise. In this case, $\partial\varphi = \partial I_K$, where $K = \{y \in H_0^1(\Omega); 0 \le y(x) \le 1$, a.e., $x \in \Omega\}$. Then $\mu \in \partial\varphi(y)$ satisfies $\mu(y - z) \ge 0$, $\forall z \in K$ and, therefore, $\mu(\varphi) = 0$ for all $\varphi \in C_0^\infty(\Omega)$. Hence, μ is a measure with support on $\partial\Omega$. More generally (see Brezis [13]), if φ is defined by (5.136), then $\mu \in \partial\varphi(y) \in H^{-1}(\Omega)$, and then μ is a bounded measure on Ω and $\mu = \mu_a dx + \mu_s$ where the absolutely continuous part $\mu_a \in L^1(\Omega)$ has the property that $\mu_a(x) \in \beta(y(x))$, a.e. $x \in \Omega$. However, if $D(\beta) = \mathbf{R}$, then, by Lemma 2.2, if $\mu \in H^{-1}(\Omega) \cap L^1(\Omega)$ is such that $\mu(x) \in \beta(y(x))$, a.e. $x \in \Omega$, then $\mu \in By$.

Then, by Theorem 5.7, we get the following.

Corollary 5.5. *Let β be a maximal monotone graph in $\mathbf{R} \times \mathbf{R}$ and let $B = \partial\varphi$, where φ is defined by* (5.136). *Let $y_0 \in H_0^1(\Omega) \cap H^2(\Omega)$, $y_1 \in H_0^1(\Omega)$, and $f \in L^2(Q)$ be such that $\partial f/\partial t \in L^2(Q)$ and*

$$\mu_0(x) \in \beta(y_1(x)), \quad \text{a.e. } x \in \Omega \quad \text{for some } \mu_0 \in L^2(\Omega). \tag{5.137}$$

Then, there is a unique function $y \in C([0,T]; H_0^1(\Omega))$ such that

$$\frac{\partial y}{\partial t} \in C([0,T]; L^2(\Omega)) \cap C([0,T]; H_0^1(\Omega)), \qquad \frac{\partial^2 y}{dt^2} \in L^\infty(0,T; L^2(\Omega)) \tag{5.138}$$

$$\begin{cases} \dfrac{d^+}{dt}\dfrac{\partial y}{\partial t}(t) - \Delta y(t) + B\left(\dfrac{\partial}{\partial t} y(t)\right) \ni f(t), & \forall t \in [0,T), \\ y(x,0) = y_0(x), \qquad \dfrac{\partial y}{\partial t}(x,0) = y_1(x), & \text{in } \Omega, \\ y = 0, & \text{on } \partial\Omega \times (0,T). \end{cases} \tag{5.139}$$

Assume further that $D(\beta) = \mathbf{R}$. Then, $\Delta y(t) \in L^1(\Omega)$ for all $t \in [0,T)$ and

$$\frac{d^+}{dt}\frac{dy}{dt}(x,t) - \Delta y(x,t) + \mu(x,t) = f(x,t), \qquad x \in \Omega,\ t \in [0,T), \tag{5.140}$$

where $\mu(x,t) \in \beta((\partial y/\partial t)(x,t))$, a.e. $x \in \Omega$.

(We note that condition (5.139) implies (5.132).)

Problems of the form (5.131) arise in wave propagation and description of the dynamics of an elastic solid. For instance, if $\beta(r) = r|r|$, this equation models the behavior of an elastic membrane with the resistance proportional to the velocity.

If $j(r) = |r|$, then $\beta(r) = \text{sign } r$ and so equation (5.139) is of multivalued type.

As another example, consider the unilateral hyperbolic problem

$$\begin{cases} \dfrac{\partial^2 y}{\partial t^2} = \Delta y + f & \text{in } \left\{(x,t) \in Q;\ \dfrac{\partial y}{\partial t}(x,t) > \psi(x)\right\}, \\ \dfrac{\partial^2 y}{\partial t^2} \geq \Delta y + f,\ \dfrac{\partial}{\partial t} y \geq \psi & \text{in } Q, \\ y = 0 & \text{on } \partial\Omega \times [0,T), \\ y(x,0) = y_0(x),\ \dfrac{\partial y}{\partial t}(x,0) = y_1(x) & \text{in } \Omega, \end{cases} \tag{5.141}$$

where $\psi \in H^2(\Omega)$ is such that $\psi \leq 0$, a.e. on $\partial\Omega$. This is a reflection-type problem for the linear wave equation with constraints on velocity that exhibits a free boundary type behavior with moving boundary.

Clearly, we may write this variational inequality in the form (5.129), where $V = H_0^1(\Omega)$, $H = L^2(\Omega)$, $A = -\Delta$, and $B \subset H_0^1(\Omega) \times H^{-1}(\Omega)$ is defined by

$$Bu = \{w \in H^{-1}(\Omega);\ (w, u - v) \geq 0,\ \forall v \in K\}$$

for all $u \in D(B) = K = \{u \in H_0^1(\Omega);\ u \geq \psi, \text{ a.e. in } \Omega\}$.

By Theorem 5.7, we have therefore the following existence result for problem (5.141).

Corollary 5.6. *Let $f, f_t \in L^2(Q)$ and $y_0 \in H_0^1(\Omega) \cap H^2(\Omega)$, $y_1 \in H_0^1(\Omega)$ be such that $y_1(x) \geq \psi(x)$, a.e. $x \in \Omega$. Then, there is a unique function $y \in W^{1,\infty}([0,T]; H_0^1(\Omega))$ with $\partial y/\partial t \in W^{1,\infty}([0,T]; L^2(\Omega))$ satisfying*

$$\begin{cases} \displaystyle\int_\Omega \left(\frac{d^+}{dt}\frac{\partial y}{\partial t}(x,t)\left(\frac{\partial y}{\partial t}(x,t) - u(x)\right) + \nabla y(x,t) \cdot \nabla\left(\frac{\partial y}{\partial t}(x,t) - u(x)\right)\right) dx \\ \qquad \displaystyle\leq \int_\Omega f(x,t)\left(\frac{\partial y}{\partial t}(x,t) - u(x)\right) dx, \qquad \forall u \in K,\ \forall t \in [0,T), \\ y(x,0) = y_0(x), \qquad \dfrac{\partial y}{\partial t}(x,0) = y_1(x), \qquad \forall x \in \Omega. \end{cases} \tag{5.142}$$

Problem (5.142) is a variational (or weak) formulation of the free boundary problem (5.141).

The Klein–Gordon Equation

We consider now the hyperbolic boundary value problem

$$\begin{cases} \dfrac{\partial^2 y}{\partial t^2} - \Delta y + g(y) = f & \text{in } \Omega \times (0,T) = Q, \\ y(x,0) = y_0(x), \qquad \dfrac{\partial y}{\partial t}(x,0) = y_1(x) & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega \times (0,T) = \Sigma, \end{cases} \tag{5.143}$$

where Ω is a bounded and open subset of $\mathbf{R}^N$, with a sufficiently smooth boundary (of class C^2, for instance), and $g \in W^{1,\infty}(\mathbf{R})$ satisfies the following conditions.

(i) $|g'(r)| \le L(1+|r|^p)$, a.e. $r \in \mathbf{R}$, where $0 \le p \le 2/(N-2)$ if $N > 2$, and p is any positive number if $1 \le N \le 2$;
(ii) $rg(r) \ge 0$, $\forall r \in \mathbf{R}$.

In the special case where $g(y) = \mu |y|^\rho y$, assumptions (i) and (ii) are satisfied for $0 < \rho \le 2/(N-2)$ if $N > 2$, and for $\rho \ge 0$ if $N \le 2$. For $\rho = 2$, this is the classical Klein–Gordon equation, arising in the quantum field theory (see Reed and Simon [37]).

In the sequel, we denote by ψ the primitive of g, which vanishes at 0: $\psi(r) = \int_0^r g(t)dt$, $\forall r \in \mathbf{R}$.

Theorem 5.8. *Let* $f, (\partial f/\partial t) \in L^2(Q)$ *and* $y_0 \in H_0^1(\Omega) \cap H^2(\Omega)$, $y_1 \in H_0^1(\Omega)$ *be such that* $\psi(y_0) \in L^1(\Omega)$. *Then, under assumptions* (i) and (ii) *there is a unique function y that satisfies*

$$\begin{cases} y \in L^\infty(0,T; H_0^1(\Omega) \cap H^2(\Omega)) \cap C^1([0,T]; H_0^1(\Omega)), \\ \dfrac{\partial y}{\partial t} \in C([0,T]; H_0^1(\Omega)), \qquad \dfrac{\partial^2 y}{\partial t^2} \in L^\infty(0,T; L^2(\Omega)), \\ \psi(y) \in L^\infty(0,T; L^1(\Omega)), \end{cases} \tag{5.144}$$

and

$$\begin{cases} \dfrac{\partial^2 y}{\partial t^2} - \Delta y + g(y) = f, & \text{a.e. in } Q, \\ y(x,0) = y_0(x), \qquad \dfrac{\partial y}{\partial t}(x,0) = y_1(x), & \text{a.e. } x \in \Omega. \end{cases} \tag{5.145}$$

Proof. As in the previous case, we write equation (5.143) as a first-order differential equation in $X = H_0^1(\Omega) \times L^2(\Omega)$; that is,

$$\frac{\partial y}{\partial t} - z = 0, \qquad \frac{dz}{dt} - \Delta y + g(y) = f \quad \text{in } [0,T]. \tag{5.146}$$

Equivalently,

$$\begin{cases} \dfrac{d}{dt} U(t) + A_0 U(t) + GU(t) = F(t), & t \in [0,T], \\ U(0) = [y_0, y_1], \end{cases} \tag{5.147}$$

where $U(t) = [y(t), z(t)]$, $G(U) = [0, g(y)]$, $A_0 U = [-z, -\Delta y]$, and $F(t) = [0, f(t)]$.

The space $X = H_0^1(\Omega) \times L^2(\Omega)$ is endowed with the usual norm:

$$\|U\|_X^2 = \|y\|_{H_0^1(\Omega)}^2 + \|z\|_{L^2(\Omega)}^2, \qquad U = [y,z].$$

It should be said that although the operator $A_0 + G$ is not quasi-m-accretive in the space X, the Cauchy problem (5.147) can be treated with the previous method.

We note first that the operator G is locally Lipschitz on X. Indeed, we have

$$\|G(y_1,z_1) - G(y_2,z_2)\|_X = \|g(y_1) - g(y_2)\|_{L^2(\Omega)}.$$

On the other hand, we have

$$\begin{aligned}|g(y_1){-}g(y_2)| &\le \left|\int_0^1 g'(\lambda y_1 + (1-\lambda)y_2)d\lambda(y_1 - y_2)\right|\\ &\le L|y_1 - y_2|\int_0^1 (1 + |\lambda(y_1 - y_2) + y_2|^p)d\lambda\\ &\le C|y_1 - y_2|(\max(|y_1|^p, |y_2|^p) + 1), \qquad \forall y_1, y_2 \in \mathbf{R}.\end{aligned}$$

Hence, for any $z \in L^2(\Omega)$ and $y_i \in H_0^1(\Omega)$, $i = 1,2$, we have

$$\begin{aligned}&\int_\Omega z(x)(g(y_1(x)) - g(y_2(x)))dx\\ &\le C\int_\Omega |z(x)|\,|y_1(x) - y_2(x)|(\max(|y_1(x)|^p, |y_2(x)|^p) + 1)dx\end{aligned}$$

and, therefore, by the Hölder inequality,

$$\begin{aligned}\int_\Omega z(g(y_1) - g(y_2))dx \le C\|z\|_{L^2(\Omega)}\|y_1{-}y_2\|_{L^\beta(\Omega)} \max(\|y_1\|_{L^{2p}(\Omega)}^p, \|y_2\|_{L^{2p}(\Omega)}^p)\\ +C\|z\|_{L^2(\Omega)}\|y_1{-}y_2\|_{L^2(\Omega)},\end{aligned}$$

where

$$\frac{1}{\beta} + \frac{1}{\delta} + \frac{1}{2} = 1.$$

Now, we take in the latter inequality $\delta = N$ and $\beta = 2N/(N-2)$. We get

$$\begin{aligned}&\|g(y_1){-}g(y_2)\|_2\\ &\le C\|y_1{-}y_2\|_{2N/(N-2)} \max(\|y_1\|_{Np}^p, \|y_2\|_{Np}^p){+}C\|y_1{-}y_2\|_2, \qquad \forall y_1, y_2 \in H_0^1(\Omega).\end{aligned}$$

Then, by the Sobolev embedding theorem and assumption (i), we have

$$\begin{aligned}\|y_i\|_{Np} &\le C_i\|y_i\|_{H_0^1(\Omega)}, \qquad i = 1,2,\\ \|y_1 - y_2\|_{2N/(N-2)} &\le C_0\|y_1 - y_2\|_{H_0^1(\Omega)}.\end{aligned}$$

(We have denoted by $\|\cdot\|_p$ the L^p norm.) This yields

$$\|g(y_1)-g(y_2)\|_2 \le C\|y_1-y_2\|_{H_0^1(\Omega)}(\max(\|y_1\|^p_{H_0^1(\Omega)}, \|y_2\|^p_{H_0^1(\Omega)})+1)$$

and, therefore,

$$\begin{aligned}&\|G(y_1,z_1)-G(y_2,z_2)\|_X \\ &\le C\|y_1-y_2\|_{H_0^1(\Omega)}(1+\max(\|y_1\|^p_{H_0^1(\Omega)}, \|y_2\|^p_{H_0^1(\Omega)})), \\ &\qquad \forall y_1,y_2 \in H_0^1(\Omega),\end{aligned} \tag{5.148}$$

as claimed. □

To prove the existence of a local solution, we use the truncation method presented in Section 4.1 (see Theorem 4.8).

Let $r > 0$ be arbitrary but fixed. Define the operator $\widetilde{G} : X \to X$,

$$\widetilde{G}(y,z) = \begin{cases} G(y,z) & \text{if } \|y\|_{H_0^1(\Omega)} \le r, \\ G\left(r\,\dfrac{y}{\|y\|_{H_0^1(\Omega)}}, z\right) & \text{if } \|y\|_{H_0^1(\Omega)} > r. \end{cases}$$

By (5.148), we see that the operator $\widetilde{G}$ is Lipschitz on X. Hence, A_0+G is ω-m-accretive on X and, by Theorem 4.6, we conclude that the Cauchy problem

$$\begin{cases} \dfrac{d}{dt}U(t)+A_0U(t)+\widetilde{G}U(t)=F(t), & \text{a.e. } t\in(0,T), \\ U(0)=[y_0,y_1], \end{cases} \tag{5.149}$$

has a unique solution $U \in W^{1,\infty}([0,T];X)$. This implies that there is a unique $y \in W^{1,\infty}([0,T];H_0^1(\Omega))$ with $dy/dt \in W^{1,\infty}([0,T];L^2(\Omega))$ such that

$$\begin{cases} \dfrac{d^2y}{dt^2}(t)-\Delta y(t)+\widetilde{g}(y(t))=f(t), & \text{a.e. } t\in(0,T), \\ y(0)=y_0,\ \dfrac{dy}{dt}(0)=y_1 & \text{in } \Omega, \end{cases} \tag{5.150}$$

where $\widetilde{g} : H_0^1(\Omega) \to L^2(\Omega)$ is defined by

$$\widetilde{g}(y) = \begin{cases} g(y) & \text{if } \|y\|_{H_0^1(\Omega)} \le r, \\ g\left(r\,\dfrac{y}{\|y\|_{H_0^1(\Omega)}}\right) & \text{if } \|y\|_{H_0^1(\Omega)} > r. \end{cases}$$

Choose r sufficiently large such that $\|y_0\|_{H_0^1(\Omega)} < r$. Then, there is an interval $[0,T_r]$ such that $\|y(t)\|_{H_0^1(\Omega)} \le r$ for $t\in[0,T_r]$ and $\|y(t)\|_{H_0^1(\Omega)} > r$ for $t > T_r$. We have therefore

$$\frac{\partial^2 y}{\partial t^2} - \Delta y + g(y) = f \quad \text{in } \Omega\times(0,T_r),$$

and multiplying this by y_t and integrating on $\Omega \times (0,t)$, we get the energy equality

$$\begin{aligned}&\|y_t(t)\|_2^2 + \|y(t)\|^2_{H_0^1(\Omega)} + 2\int_\Omega \psi(y(x,t))dx \\ &= \|y_1\|_2^2 + \|y_0\|^2_{H_0^1(\Omega)} + 2\int_\Omega \psi(y_0(x))dx + 2\int_0^t \int_\Omega f y_s \, dx\, ds.\end{aligned}$$

Because $\psi(y) \ge 0$ and $\psi(y_0) \in L^1(\Omega)$, by Gronwall's lemma we see that

$$\|y_t(t)\|_2 \le (\|y_1\|_2^2 + \|y_0\|^2_{H_0^1(\Omega)} + 2\|\psi(y_0)\|_{L^1(\Omega)})^{1/2} + \int_0^{T_r} \|f(s)\|_2 ds$$

and, therefore,

$$\begin{aligned}&\|y_t(t)\|_2^2 + \|y(t)\|^2_{H_0^1(\Omega)} + 2\int_\Omega \psi(y(x,t))dx \\ &\le \|y_1\|_2^2 + \|y_0\|^2_{H_0^1(\Omega)} + 2\int_\Omega \psi(y_0)dx + \left(\int_0^t \|f(s)\|_2^2 ds\right)^{1/2} \\ &\quad \times \left((\|y_1\|_2^2 + \|y_0\|^2_{H_0^1(\Omega)} + 2\|\psi(y_0)\|_{L^1(\Omega)})^{1/2} + \int_0^{T_r} \|f(s)\|_2 ds \right).\end{aligned}$$

The latter estimate shows that, given $y_0 \in H_0^1(\Omega)$, $y_1 \in L^2(\Omega)$, $T > 0$, and $f \in L^2(Q_T)$, there is a sufficiently large r such that $\|y(t)\|_{H_0^1(\Omega)} \le r$ for $t \in [0,T]$. We may infer, therefore, that for r large enough the function y found as the solution to (5.150) is, in fact, a solution to equation (5.145) satisfying all the conditions of Theorem 5.8.

The uniqueness of y satisfying (5.144) and (5.145) is the consequence of the fact that such a function is the solution (along with $z = \partial y/\partial t$) to the ω-accretive differential equation (5.149).

By the previous proof, it follows that, if one merely assumes that

$$y_0 \in H_0^1(\Omega), \qquad y_1 \in L^2(\Omega), \qquad \psi(y_0) \in L^1(\Omega),$$

then there is a unique function $y \in C([0,T]; H_0^1(\Omega))$, $\partial y/\partial t \in C([0,T]; L^2(\Omega))$, that satisfies equation (5.143) in a mild sense. However, if $\psi(y_0) \notin L^1(\Omega)$ or, if one drops assumption (ii), then the solution to (5.143) exists locally in time, only; that is, in a neighborhood of the origin.

Under appropriate assumptions on g and β, the above existence results extend to equations of the form

$$\begin{cases} \dfrac{\partial^2 y}{\partial t^2} - \Delta y + \beta\left(\dfrac{\partial y}{\partial t}\right) + g(y) = f & \text{in } Q, \\ y(x,0) = y_0(x), \qquad \dfrac{\partial y}{\partial t}(x,0) = y_1(x) & \text{in } \Omega, \\ y = 0 & \text{on } \partial\Omega \times (0,T). \end{cases}$$

(See Haraux [28].) In Barbu, Lasiecka and Rammaha [5], the local and global existence of generalized solutions is studied in the case of more general equations of the form

$$\frac{\partial^2 y}{\partial t^2} - \Delta y + |y|^k \beta\left(\frac{\partial y}{\partial t}\right) = |y|^{p-1} y \quad \text{in } \Omega \times (0,T),$$

where $\beta(r) \le C_0 r^m$, $\int_0^r \beta(s)ds \ge C r^{m+1}$, $0 \le k < N/(N+2)$, $1 < p < \infty$.

It turns out that, if $1 < p \le k+m$, then there is a global solution but every solution is only local and blows up if p is greater than $m+k$. For other recent results in this context we refer also to the work of Serrin, Todorova, and Vitillaro [38].

5.7 Navier–Stokes Equations

The classical Navier–Stokes equations

$$\begin{cases} y_t(x,t) - \nu_0 \Delta y(x,t) + (y \cdot \nabla) y(x,t) = f(x,t) + \nabla p(x,t), \\ \hfill x \in \Omega,\ t \in (0,T) \\ (\nabla \cdot y)(x,t) = 0, \hfill \forall (x,t) \in \Omega \times (0,T) \\ y = 0 \hfill \text{on } \partial\Omega \times (0,T) \\ y(x,0) = y_0(x), \hfill x \in \Omega \end{cases} \tag{5.151}$$

describe the non-slip motion of a viscous, incompressible, Newtonian fluid in an open domain $\Omega \subset \mathbf{R}^N$, $N = 2,3$. Here $y = (y_1, y_2, \ldots, y_N)$ is the velocity field, p is the pressure, f is the density of an external force, and $\nu_0 > 0$ is the viscosity of the fluid.

We have used the following standard notation

$$\begin{cases} \nabla \cdot y = \operatorname{div} y = \sum_{i=1}^N D_i y_i, \qquad D_i = \dfrac{\partial}{\partial x_i}, \qquad i = 1, \ldots, N \\ (y \cdot \nabla) y = \sum_{i=1}^N y_i D_i y_j, \qquad j = 1, \ldots, N. \end{cases}$$

By a classical device due to J. Leray, the boundary value problem (5.151) can be written as an infinite-dimensional Cauchy problem in an appropriate function space on Ω. To this end we introduce the following spaces

$$H = \{y \in (L^2(\Omega))^N;\ \nabla \cdot y = 0,\ y \cdot \nu = 0 \text{ on } \partial\Omega\} \tag{5.152}$$

$$V = \{y \in (H_0^1(\Omega))^N;\ \nabla \cdot y = 0\}. \tag{5.153}$$

Here ν is the outward normal to $\partial\Omega$.

The space H is a closed subspace of $(L^2(\Omega))^N$ and it is a Hilbert space with the scalar product

$$(y,z) = \int_\Omega y \cdot z\, dx \tag{5.154}$$

and the corresponding norm $|y| = \left(\int_\Omega |y|^2 dx\right)^{1/2}$. (We denote by the same symbol $|\cdot|$ the norm in $\mathbf{R}^N$, $(L^2(\Omega))^N$, and H, respectively.) The norm of the space V is denoted by $\|\cdot\|$:

$$\|y\| = \left(\int_\Omega |\nabla y(x)|^2 dx\right)^{1/2}. \tag{5.155}$$

We denote by $P : (L^2(\Omega))^N \to H$ the orthogonal projection of $(L^2(\Omega))^N$ onto H (the Leray projector) and set

$$a(y,z) = \int_\Omega \nabla y \cdot \nabla z\, dx, \qquad \forall y,z \in V. \tag{5.156}$$

$$A = -P\Delta,\ D(A) = (H^2(\Omega))^N \cap V. \tag{5.157}$$

Equivalently,

$$(Ay,z) = a(y,z), \qquad \forall y,z \in V. \tag*{(5.157)$'$}$$

The *Stokes operator* A is self-adjoint in H, $A \in L(V,V')$ (V' is the dual of V with the norm denoted by $\|\cdot\|_{V'}$) and

$$(Ay,y) = \|y\|^2, \qquad \forall y \in V. \tag{5.158}$$

Finally, consider the trilinear functional

$$b(y,z,w) = \int_\Omega \sum_{i,j=1}^N y_i D_i z_j w_j\, dx, \qquad \forall y,z,w \in V \tag{5.159}$$

and we denote by $B : V \to V'$ the nonlinear operator defined by

$$By = P(y \cdot \nabla)y \tag{5.160}$$

or, equivalently,

$$(By,w) = b(y,y,w), \qquad \forall w \in V. \tag*{(5.160)$'$}$$

Let $f \in L^2(0,T;V')$ and $y_0 \in H$. The function $y : [0,T] \to H$ is said to be a *weak solution* to equation (5.151) if

$$y \in L^2(0,T;V') \cap C_w([0,T];H) \cap W^{1,1}([0,T];V') \tag{5.161}$$

$$\begin{cases} \dfrac{d}{dt}(y(t),\psi) + \nu_0 a(y(t),\psi) + b(y(t),y(t),\psi) = (f(t),\psi), & \text{a.e. } t \in (0,T), \\ y(0) = y_0, \qquad \forall \psi \in V. \end{cases} \tag{5.162}$$

(Here $(\cdot,\cdot)$ is, as usual, the pairing between V, V' and the scalar product of H.)

Equation (5.162) can be equivalently written as

$$\begin{cases} \dfrac{dy}{dt}(t) + \nu_0 Ay(t) + By(t) = f(t), & \text{a.e. } t \in (0,T) \\ y(0) = y_0 \end{cases} \tag{5.163}$$

where dy/dt is the strong derivative of function $y : [0,T] \to V'$.

The function y is said to be the *strong solution* to (5.151) if $y \in W^{1,1}([0,T];H) \cap L^2(0,T;D(A))$ and (5.163) holds with $dy/dt \in L^1(0,T;H)$ the strong derivative of function $y : [0,T] \to H$.

There is a standard approach to existence theory for the Navier–Stokes equation (5.163) based on the Galerkin approximation scheme (see, e.g., Temam [39]). The method we use here relies on the general results on the nonlinear Cauchy problem of monotone type developed before and, although it leads to a comparable result, it provides a new insight into existence theory of this problem.

It should be said that equation (5.163) is not of monotone type in H, but it can be treated, however, into this framework by an argument described below.

Before proceeding with the existence for problem (1.1), we pause briefly to present some fundamental properties of the trilinear functional b defining the inertial operator B (see Constantin and Foias [19], Temam [39]).

Proposition 5.14. *Let* $1 \le N \le 3$. *Then*

$$b(y,z,w) = -b(y,w,z), \ \forall y,z,w \in V \tag{5.164}$$

$$|b(y,z,w)| \le C\|y\|_{m_1}\|z\|_{m_2+1}\|w\|_{m_3}, \ \forall u \in V_{m_1}, \ v \in V_{m_2}, \ w \in V_{m_3} \tag{5.165}$$

where $m_i \ge 0, \ i = 1,2,3$ *and*

$$\begin{aligned} m_1 + m_2 + m_3 &\ge \frac{N}{2} \quad \text{if } m_i \ne \frac{N}{2}, \quad \forall i = 1,2,3, \\ m_1 + m_2 + m_3 &> \frac{N}{2} \quad \text{if } m_i = \frac{N}{2}, \quad \text{for some } i = 1,2,3. \end{aligned} \tag{5.166}$$

Here $V_{m_i} = V \cap (H_0^{m_i}(\Omega))^N$.

Proof. It suffices to prove (5.165) for $y,z,w \in \{y \in (C_0^\infty(\Omega))^N;\ \nabla \cdot y = 0\}$. We have

$$\begin{aligned} b(y,z,w) &= \int_\Omega y_i D_i z_j w_j dx = \int_\Omega (y_i D_i(z_j w_j) - y_i D_i w_j z_j) dx \\ &= -\int_\Omega y_i D_i w_j z_j dx = -b(y,z,w) \end{aligned}$$

because $\nabla \cdot y = 0$. By Hölder's inequality we have

$$|b(y,z,w)| \le |y_i|_{q_1} |D_i z_j|_{q_2} |w_j|_{q_3}, \qquad \frac{1}{q_1} + \frac{1}{q_2} + \frac{1}{q_3} \le 1. \tag{5.167}$$

(Here $|\cdot|_q$ is the norm of $L^q(\Omega)$.) On the other hand, by the Sobolev embedding theorem we have (see Theorem 1.5)

$$H^{m_i}(\Omega) \subset L^{q_i}(\Omega) \quad \text{for } \frac{1}{q_i} = \frac{1}{2} - \frac{m_i}{N}$$

if $m_i < N/2$. Then, (5.167) yields

$$|b(y,z,w)| \le C\|y\|_{m_1}\|z\|_{m_2+1}\|w\|_{m_3}$$

if $m_i < N/2$, $i = 1,2,3$.

If one m_i is larger than $N/2$ the previous inequality still remains true because, in this case,

$$H^{m_i}(\Omega) \subset L^\infty(\Omega).$$

If $m_i = N/2$ then

$$H^{m_i}(\Omega) \subset \bigcap_{q>2} L^q(\Omega)$$

and so (5.167) holds for $1/q_2 + 1/q_3 < 1$ and $q_1 = \varepsilon$ where

$$\frac{1}{\varepsilon} = 1 - \frac{1}{q_2} - \frac{1}{q_3}\cdot$$

Then (5.165) follows for $m_1 + m_2 + m_3 > N/2$ as claimed.

We have also the interpolation inequality

$$\|u\|_m \le c\|u\|_\ell^{1-\alpha}\|u\|_{\ell+1}^\alpha, \quad \text{for } \alpha = m - \ell \in [0,1]. \tag{5.168}$$

In particular, it follows by Proposition 5.14 that B is continuous from V to V'. Indeed, we have

$$(By - Bz, w) = b(y, y-z, w) + b(y-z, z, w), \qquad \forall w \in V$$

and this yields (notice that $\|\cdot\| = \|\cdot\|_1$ and $|Ay| = |y|_2$)

$$|(By - Bz, w)| \le C(\|y\|\|y-z\|\|w\| + \|y-z\|\|z\|\|w\|).$$

Hence

$$\|By - Bz\|_{V'} \le C\|y-z\|(\|y\| + \|z\|), \qquad \forall y,z \in V. \tag{5.169}$$

We would like to treat (5.163) as a nonlinear Cauchy problem in the space H. However, because the operator $\nu_0 A + B$ is not quasi-m-accretive in H, we first consider a quasi-m-accretive approximation of the form taken in the proof of Theorem 4.8.

For each $M > 0$ define the operator $B_M : V \to V'$ (see (4.67))

$$B_M y = \begin{cases} By & \text{if } \|y\| \le M, \\ \dfrac{M^2}{\|y\|^2} By & \text{if } \|y\| > M, \end{cases}$$

and consider the operator $\Gamma_M : D(\Gamma_M) \subset H \to H$

$$\Gamma_M = \nu_0 A + B_M, \qquad D(\Gamma_M) = D(A). \tag{5.170}$$

Let us show that Γ_M is well defined. Indeed, we have

$$|\Gamma_M y| \leq \nu_0 |Ay| + |B_M y|, \qquad \forall y \in D(A).$$

On the other hand, by (5.165) for $m_1 = 1,\ m_2 = 1/2,\ m_3 = 0$, we have for $\|y\| \leq M$

$$|(B_M y, w)| = |b(y,y,w)| \leq C\|y\|^{3/2}|Ay|^{1/2}|w|$$

because $\|y\|_{3/2} \leq \|y\|^{1/2}|Ay|^{1/2}$. Hence

$$|B_M y| \leq C|Ay|^{1/2}\|y\|^{3/2}, \qquad \forall y \in D(A).$$

Similarly, we get for $\|y\| > M$

$$|B_M y| \leq \frac{CM^2}{\|y\|^2}\,|Ay|^{1/2}\|y\|^{3/2} \leq C|Ay|^{1/2}\|y\|^{3/2}.$$

This yields

$$|\Gamma_M y| \leq \nu_0|Ay| + C|Ay|^{1/2}\|y\|^{3/2}, \qquad \forall y \in D(A) \tag{5.171}$$

as claimed. □

Lemma 5.2. *There is α_M such that $\Gamma_M + \alpha_M I$ is m-accretive in $H \times H$.*

Proof. We show first that for each $\nu > 0$

$$((\Gamma_M + \lambda)y - (\Gamma_M + \lambda)z, y - z) \geq \frac{\nu}{2}\,\|y - z\|^2, \qquad \forall y, z \in D(A),\ \text{ for } \lambda \geq C_M^\nu.$$

To this end we prove that

$$|(B_M y - B_M z, y - z)| \leq \frac{\nu}{2}\,\|y - z\|^2 + C_M|y - z|^2. \tag{5.172}$$

We treat only the case $N = 3$ because $N = 2$ follows in a similar way.

Let $\|y\|, \|z\| \leq M$. Then we have

$$\begin{aligned}(B_M y - B_M z, y - z) &= (By - Bz, y - z) = b(y,y,y-z) - b(z,z,y-z)\\ &= b(y-z,y,y-z) + b(z,y-z,y-z) = b(y-z,y,y-z).\end{aligned}$$

Hence, by Proposition 5.14, for $m_1 = 1,\ m_2 = 0,\ m_3 = 1/2$ we have

$$\begin{aligned}|(B_My - B_Mz, y-z)| &= |b(y-z,y,y-z)| \le C\|y-z\|\|y\|\|y-z\|_{1/2}\\ &\le C\|y-z\|^{3/2}\|y\||y-z|^{1/2}\\ &\le C_M\|y-z\|^{3/2}|y-z|^{1/2}\\ &\le \frac{\nu}{2}\|y-z\|^2 + C_M|y-z|^2\end{aligned}$$

as desired.

Now consider the case where $\|y\| > M$, $\|z\| > M$. We have

$$\begin{aligned}&(B_My - B_Mz, y-z)\\ &= \frac{M^2}{\|y\|^2}\left(b(y,y,y-z) - b(z,z,y-z)\right) + \left(\frac{M^2}{\|y\|^2} - \frac{M^2}{\|z\|^2}\right) b(z,z,y-z)\\ &= \frac{M^2}{\|y\|^2} b(y-z,y,y-z) + M^2\left(\frac{\|z\|^2 - \|y\|^2}{\|y\|^2\|z\|^2}\right) b(z,z,y-z).\end{aligned}$$

This yields

$$\begin{aligned}|(B_My - B_Mz, y-z)| &\le \frac{CM^2}{\|y\|}\|y-z\|^{3/2}|y-z|^{1/2}\\ &\quad + \frac{CM^2}{\|y\|^2\|z\|^2}\left|\,\|z\|^2 - \|y\|^2\right|\|z\|\,\|y-z\|_{1/2}\\ &\le \frac{\nu}{2}\|y-z\|^2 + C_M^1|y-z|^2.\end{aligned}$$

Assume now that $\|y\| > M,\ \|z\| \le M$. We have

$$\begin{aligned}&|(B_My - B_Mz, y-z)| = \left|\frac{M^2}{\|y\|^2} b(y,y,y-z) - b(z,z,y-z)\right|\\ &\le \left|\frac{M^2}{\|y\|^2} - 1\right| |b(z,z,y-z)| + \frac{M^2}{\|y\|^2}|b(y,y,y-z) - b(z,z,y-z)|\\ &\le C\frac{\|y\|^2 - M^2}{\|y\|^2}\|z\|^2\|y-z\|^{1/2}|y-z|^{1/2} + \frac{M^2}{\|y\|^2}|b(y-z,y,y-z)|\\ &\le C_M^1\|y-z\|^{3/2}|y-z|^{1/2}\end{aligned}$$

which again implies (5.172), as claimed.

We note also that by (5.169) it follows that

$$\|B_My - B_Mz\|_{V'} \le C\|y-z\|(\|y\| + \|z\|), \qquad \forall y,z \in V, \tag{5.173}$$

where C is independent of M.

Let us now proceed with the proof of α_M-m-accretivity of Γ_M. Consider the operator

$$\begin{aligned} F_M u &= \nu_0 A u + B_M u + \alpha_M u, \ \forall u \in D(F_M) \\ D(F_M) &= \{u \in V;\ \nu_0 A u + B_M u \in H\}. \end{aligned} \tag{5.174}$$

By (5.172) we see that for $\alpha_M \geq C_M$ the operator $u \to \nu_0 A u + B_M u + \alpha_M u$ is monotone, coercive, and continuous from V to V'. Hence its restriction to H; that is, F_M is maximal monotone (m-accretive) in $H \times H$. To complete the proof it suffices to show that $D(F_M) = D(A)$ for α_M large enough. (Clearly $D(A) \subset D(F_M)$.)

Note first that by (5.165) we have

$$|(B_M y, w)| \leq C|b(y,y,w)| \leq C\|y\|\|y\|_{3/2}|w|, \qquad \forall w \in H,$$

and this yields by interpolation (see (5.168))

$$|B_M(y)| \leq C\|y\|^{3/2}|Ay|^{1/2} \leq C_M|Ay|^{1/2}.$$

Hence

$$|Ay| \leq \frac{1}{\nu_0}(|\Gamma_M y| + |B_M y|) \leq \frac{1}{\nu_0}(|\Gamma_M y| + C_M|Ay|^{1/2}), \qquad \forall y \in D(A);$$

that is,

$$|Ay| \leq C_M(|\Gamma_M y| + 1), \qquad \forall y \in D(A). \tag{5.175}$$

Now we consider the operators

$$\begin{aligned} &F_M^1 = \nu_0(1-\varepsilon)A, && D(F_M^1) = D(A) \\ &F_M^2 = \varepsilon\nu_0 A + B_M + \alpha_M I, && D(F_M^2) = \{u \in V;\ \varepsilon\nu_0 A u + B_M u \in H\}, \end{aligned}$$

where α_M is large enough so that F_M^2 is m-accretive in $H \times H$. (We have seen above that such an α_M exists.)

We have

$$\begin{aligned} \left|F_M^2(y)\right| &\leq \varepsilon\nu_0|Ay| + |B_M y| + \alpha_M|y| \\ &\leq \varepsilon\nu_0|Ay| + C_M|Ay|^{1/2} + \alpha_M|y| \leq \varepsilon(1+\delta)|Ay| + \alpha_M|y| + C_M^1 \\ &\leq \frac{\varepsilon(1+\delta)}{\nu_0(1-\varepsilon)}\left|F_M^1(y)\right| + \alpha_M|y| + C_M^1,\ \forall y \in D(A) = D(F_M^1). \end{aligned}$$

Thus for ε small enough it follows by Proposition 3.9 that $F_M^1 + F_M^2$ with the domain $D(A)$ is m-accretive in $H \times H$. Because $F_M = F_M^1 + F_M^2$ on $D(A) \subset D(F_M)$ we infer that $D(F_M) = D(A)$ as claimed. □

For each $M > 0$ consider the equation

$$\begin{cases} \dfrac{dy}{dt}(t) + \nu_0 A y(t) + B_M y(t) = f(t), \qquad t \in (0,T) \\ y(0) = y_0. \end{cases} \tag{5.176}$$

Proposition 5.15. *Let $y_0 \in D(A)$ and $f \in W^{1,1}([0,T];H)$ be given. Then there is a unique solution $y_M \in W^{1,\infty}([0,T];H) \cap L^\infty(0,T;D(A)) \cap C([0,T];V)$ to equation* (5.176)*. Moreover, $(d^+/dt)y_M(t)$ exists for all $t \in [0,T)$ and*

$$\frac{d^+}{dt} y_M(t) + \nu_0 A y_M(t) + B_M y_M(t) = f(t), \qquad \forall t \in [0,T). \tag{5.177}$$

Proof. This follows by Theorem 4.4. Because $\Gamma_M y_M = \nu_0 A y_M + B_M y_M \in L^\infty(0,T;H)$, by (5.175) we infer that $A y_M \in L^\infty(0,T;H)$. As $dy_M/dt \in L^\infty(0,T;H)$, we conclude also that $y_M \in C([0,T];V) \cap L^\infty(0,T;D(A))$, as claimed. □

Now we are ready to formulate the main existence result for the strong solutions to Navier–Stokes equation (5.151) ((5.151)′).

Theorem 5.9. *Let $N = 2,3$ and $f \in W^{1,1}([0,T];H)$, $y_0 \in D(A)$ where $0 < T < \infty$. Then there is a unique function $y \in W^{1,\infty}([0,T^*);H) \cap L^\infty(0,T^*;D(A)) \cap C([0,T^*];V)$ such that*

$$\begin{cases} \dfrac{dy(t)}{dt} + \nu_0 A y(t) + B y(t) = f(t), & \text{a.e. } t \in (0,T^*), \\ y(0) = y_0, \end{cases} \tag{5.178}$$

for some $T^* = T^*(\|y_0\|) \le T$. If $N = 2$ then $T^* = T$. Moreover, $y(t)$ is right differentiable and

$$\frac{d^+}{dt} y(t) + \nu_0 A y(t) + B y(t) = f(t), \qquad \forall t \in [0,T^*). \tag{5.179}$$

Proof. The idea of the proof is to show that for M sufficiently large the flow $y_M(t)$, defined by Proposition 5.15, is independent of M on each interval $[0,T]$ if $N = 2$ or on $[0,T(y_0)]$ if $N = 3$. Let y_M be the solution to (5.176); that is,

$$\begin{cases} \dfrac{dy_M}{dt}(t) + \nu_0 A y_M(t) + B_M y_M(t) = f(t), & \text{a.e. } t \in (0,T), \\ y(0) = y_0. \end{cases} \tag{5.180}$$

If we multiply (5.180) by y_M and integrate on $(0,t)$, we get

$$|y_M(t)|^2 + \nu_0 \int_0^t \|y_M(s)\|^2 ds \le C\left(|y_0|^2 + \frac{1}{\nu_0}\int_0^T |f(t)|^2 dt\right), \qquad \forall M.$$

Next, we multiply (5.180) (scalarly in H) by $A y_M(t)$. We get

$$\frac{1}{2}\,\frac{d}{dt}\,\|y_M(t)\|^2 + \nu_0 |A y_M(t)|^2 \le |(B_M y_M(t), A y_M(t))| + |f(t)||A y_M|,$$
$$\text{a.e. } t \in (0,T).$$

This yields

$$\begin{aligned}&\|y_M(t)\|^2+\nu_0\int_0^t |Ay_M(s)|^2ds\\&\le C\left(\|y_0\|^2+\frac{1}{\nu_0}\int_0^T |f(t)|^2dt+\int_0^t |(B_My_M,Ay_M)|ds\right).\end{aligned} \tag{5.181}$$

On the other hand, for $N=3$, by (5.165) we have (the case $N=2$ is treated separately below)

$$\begin{aligned}|(B_My_M,Ay_M)| &< |b(y_M,y_M,Ay_M)|\\&\le C\|y_M\|\|y_M\|_{3/2}|Ay_M|\\&\le C\|y_M\|^{3/2}|Ay_M|^{3/2},\quad \text{a.e. } t\in(0,T).\end{aligned}$$

(Everywhere in the following C is independent of M,ν_0.) Then, by (5.181) we have

$$\begin{aligned}&\|y_M(t)\|^2+\nu_0\int_0^t |Ay_M(s)|^2ds\\&\le C\left(\|y_0\|^2+\frac{1}{\nu_0}\int_0^T |f(t)|^2dt+\int_0^t |Ay_M(s)|^{3/2}\|y_M(s)\|^{3/2}ds\right)\\&\le C\left(\|y_0\|^2+\frac{1}{\nu_0}\int_0^T |f(t)|^2dt+\frac{1}{\nu_0}\int_0^t \|y_M(s)\|^6ds\right)+\frac{\nu}{2}\int_0^t |Ay_M(s)|^2ds,\\&\qquad\qquad \forall t\in[0,T].\end{aligned}$$

Finally,

$$\begin{aligned}&\|y_M(t)\|^2+\frac{\nu_0}{2}\int_0^t |Ay_M(s)|^2ds\\&\le C_0\left(\|y_0\|^2+\frac{1}{\nu_0}\int_0^T |f(s)|^2ds+\frac{1}{\nu_0}\int_0^t \|y_M(s)\|^6ds\right).\end{aligned} \tag{5.182}$$

Next, we consider the integral inequality

$$\|y_M(t)\|^2\le C_0\left(\|y_0\|^2+\frac{1}{\nu_0}\int_0^T |f(s)|^2ds+\frac{1}{\nu_0}\int_0^t \|y_M(s)\|^6ds\right). \tag{5.183}$$

We have

$$\|y_M(t)\|^2\le\varphi(t),\qquad \forall t\in(0,T),$$

where

$$\begin{aligned}\varphi' &\le \frac{C_0}{\nu_0}\,\varphi^3,\qquad \forall t\in(0,T)\\ \varphi(0)&=C_0\left(\|y_0\|^2+\frac{1}{\nu_0}\int_0^T |f(s)|^2ds\right).\end{aligned}$$

This yields

$$\varphi(t)\le\left(\frac{\nu_0\varphi^3(0)}{\nu_0-3t\varphi^3(0)}\right)^{1/3},\qquad \forall t\in\left(0,\frac{\nu_0}{3\varphi^3(0)}\right).$$

Hence

$$\|y_M(t)\|^2 \le \left(\frac{\nu_0 \varphi^3(0)}{\nu_0 - 3t\varphi^3(0)}\right)^{1/3}, \qquad \forall t \in (0,T^*), \tag{5.184}$$

where

$$T^* = \frac{\nu_0}{3C_0^3\left(\|y_0\|^2 + \frac{1}{\nu_0}\int_0^T |f(s)|^2 ds\right)^3}.$$

Then, by (5.182) we get

$$\|y_M(t)\|^2 + \frac{\nu_0}{2}\int_0^t |Ay_M(s)|^2 ds \le C_1(\delta)\left(\|y_0\|^2 + \frac{1}{\nu_0}\int_0^T |f(t)|^2 dt\right), \quad 0 < t < T^* - \delta. \tag{5.185}$$

For $N = 2$, we have (see (5.165))

$$\begin{aligned} |(B_M y_M, Ay_M)| &\le C|y_M|^{1/2}\|y_M\||Ay_M|^{3/2} \\ &\le \frac{\nu_0}{2}|Ay_M|^2 + \frac{C}{\nu_0}\|y_M\|^4. \end{aligned}$$

This yields

$$\begin{aligned} &\|y_M(t)\|^2 + \frac{\nu_0}{2}\int_0^t |Ay_M(s)|^2 ds \\ &\le C\left(\|y_0\|^2 + \frac{1}{\nu_0}\int_0^T |f(t)|^2 dt + \frac{1}{\nu_0}\int_0^t \|y_M(s)\|^4 ds\right). \end{aligned}$$

Then, by (5.182) and the Gronwall lemma, we obtain

$$\|y_M(t)\|^2 + \frac{\nu_0}{2}\int_0^t |Ay_M(s)|^2 ds \le C\left(\|y_0\|^2 + \frac{1}{\nu_0}\int_0^T |f(t)|^2 dt\right), \quad \forall t \in (0,T). \tag{5.186}$$

By (5.184), (5.186) we infer that for M large enough, $\|y_M(t)\| \le M$ on $(0,T^*)$ if $N = 3$ or on the whole of $(0,T)$ if $N = 2$.

Hence $B_M y_M = By_M$ on $(0,T^*)$ (respectively on $(0,T)$) and so $y_M = y$ is a solution to (5.178). This completes the proof of existence.

Uniqueness. If y_1, y_2 are two solutions to (5.178), we have

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}|y_1(t) - y_2(t)|^2 + \nu_0\|y_1(t) - y_2(t)\|^2 \\ &\le |(B(y)(t) - By_2(t), y_1(t) - y_2(t))| \\ &= |b(y_1(t), y_1(t), y_1(t) - y_2(t)) - b(y_2(t), y_2(t), y_1(t) - y_2(t))| \\ &\le C\|y_1(t) - y_2(t)\|^2(\|y_1(t)\| + \|y_2(t)\|), \quad \text{a.e. } t \in (0,T^*). \end{aligned}$$

Hence, $y_1 \equiv y_2$.

It is useful to note that the solution y to (5.178) satisfies the estimates

$$|y(t)|^2 + \nu_0 \int_0^t \|y(s)\|^2 ds \le C \left(|y_0|^2 + \frac{1}{\nu_0} \int_0^T |f(s)|^2 ds \right) \tag{5.187}$$

and (for $N = 3$)

$$\begin{aligned} &\|y(t)\|^2 + \nu_0 \int_0^t |Ay(s)|^2 ds \\ &\le C \left(\|y_0\|^2 + \frac{1}{\nu_0} \int_0^{T^*} |f(t)|^2 dt \right) \left(\int_0^t \frac{ds}{T^* - t} + 1 \right), \qquad t \in (0, T^*), \end{aligned} \tag{5.188}$$

whereas, for $N = 2$,

$$\begin{aligned} \|y(t)\|^2 + \nu_0 \int_0^t |Ay(s)|^2 ds \le C \left(\|y_0\|^2 + \frac{1}{\nu_0} \int_0^T |f(t)|^2 dt \right), \\ \forall t \in (0, T), \end{aligned} \tag{5.189}$$

where C is independent of y_0 and f.

If $N = 2$, we have a sharper estimate for y. Indeed, if we multiply (5.178) by tAy and integrate on $(0, t)$, we get after integration by parts

$$\begin{aligned} &\frac{t}{2} \|y(t)\|^2 + \nu_0 \int_0^t s|Ay(s)|^2 ds \\ &= -\int_0^t (sb(y(s), y(s), Ay(s)) - s(f(s), Ay(s)))ds + \frac{1}{2} \int_0^t \|y(s)\|^2 ds \\ &\le C \int_0^t s|Ay(s)|^{3/2} |y(s)|^{1/2} \|y(s)\| ds + \frac{\nu_0}{2} \int_0^t s|Ay(s)|^2 ds \\ &\quad + \frac{1}{2} \int_0^t s|f(s)|^2 ds + \frac{1}{2} \int_0^t \|y(s)\|^2 ds. \end{aligned}$$

Then, by (5.188), we get the estimate

$$\begin{aligned} t\|y(t)\|^2 + \nu_0 \int_0^t s|Ay(s)|^2 ds \le C \left(|y_0|^2 + \frac{1}{\nu_0} \int_0^T |f(t)|^2 dt \right), \\ \forall t \in (0, T). \end{aligned} \tag{5.190}$$

Estimates (5.186), (5.188), and (5.190) suggest that equation (5.151) could have a strong solution y under weaker assumptions on y_0 and f. We show below that this is indeed the case. □

Theorem 5.10. *Let* $y_0 \in H,\ f \in L^2(0, T; H),\ T > 0$, *and* $N = 2$. *Then there is a unique solution*

$$y \in C(]0,T];V) \cap C_w([0,T];H) \cap L^2(0,T;V),$$
$$t^{1/2}y \in L^2(0,T;D(A)) \cap L^\infty(0,T;V),$$
$$t^{1/2}\,\frac{dy}{dt} \in L^2(0,T;H), \qquad \frac{dy}{dt} \in L^{2/(1+\varepsilon)}(0,T;V')$$

to equation (5.178)*; that is,*

$$\begin{cases} \dfrac{dy}{dt}(t) + \nu_0 Ay(t) + By(t) = f(t), \quad a.e.\ t \in (0,T) \\ y(0) = y_0. \end{cases} \tag{5.191}$$

If $y_0 \in V$, then $y \in L^\infty(0,T;V) \cap L^2(0,T;D(A))$.

Proof. Let $\{y_0^j\} \subset D(A)$ and $\{f_j\} \subset W^{1,1}([0,T];H)$ be such that

$$\begin{aligned} y_0^j &\to y_0 \quad \text{strongly in } H,\\ f_j &\to f \quad \text{strongly in } L^2(0,T;H). \end{aligned}$$

By (5.187), (5.190), we have

$$|y_j(t)|^2 + \int_0^T \|y_j(t)\|^2 dt + t\|y_j(t)\|^2 + \int_0^T t|Ay_j(t)|^2 dt \le C, \qquad t \in (0,T).$$

Then, by (5.165), we obtain that

$$\int_0^T \|By_j(t)\|_{V'}^{2/(1+\varepsilon)} dt + \int_0^T t|By_j(t)|^2 dt \le C, \qquad \forall \varepsilon > 0$$

because

$$|(By_j,\varphi)| = |b(y_j,y_j,\varphi)| \le C|y_j|^{1/2}\|y_j\||Ay_j|^{1/2}|\varphi|$$

and

$$|(By_j,\varphi)| \le C\|y_j\|_\varepsilon \|y_j\|\|\varphi\|.$$

This yields

$$\begin{aligned} |By_j| &\le C|y_j|^{1/2}\|y_j\||Ay_j|^{1/2},\\ \|By_j\|_{V'} &\le C\|y_j\|_\varepsilon\|y_j\| \le C\|y_j\|^{1+\varepsilon}|y_j|^{1-\varepsilon}. \end{aligned}$$

Hence

$$\int_0^T \left(\left\| \frac{dy_j(t)}{dt} \right\|_{V'}^{2/(1+\varepsilon)} + t\left| \frac{dy_j(t)}{dt} \right|^2 \right) dt \le C.$$

Because the embeddings $D(A) \subset V \subset H \subset V'$ are compact, it follows by the Ascoli–Arzelà theorem that on a subsequence, again denoted y_j, we have

$$\begin{aligned}
y_j(t) &\xrightarrow{j\to\infty} y(t) && \text{in } C([0,T];V') \\
y_j &\longrightarrow y && \text{weak-star in } L^\infty(0,T;H), \\
& && \text{weakly in } L^2(0,T;V), \\
\sqrt{t}\,\frac{dy_j}{dt} &\longrightarrow \sqrt{t}\,\frac{dy}{dt} && \text{weakly in } L^2(0,T;H) \\
Ay_j &\longrightarrow Ay && \text{weakly in } L^2(0,T;V'), \\
\sqrt{t}\,y_j &\longrightarrow \sqrt{t}\,y && \text{weak-star in } L^\infty(0,T;V), \\
& && \text{weakly in } L^2(0,T;D(A)).
\end{aligned}$$

Moreover, by the Aubin compactness theorem, we have

$$\begin{aligned}
\sqrt{t}\,y_j(t) &\longrightarrow \sqrt{t}\,y(t) && \text{uniformly in } H \text{ on } [0,T] \\
\sqrt{t}\,y_j &\longrightarrow \sqrt{t}\,y && \text{strongly in } L^2(0,T;V).
\end{aligned}$$

Next, we have

$$\begin{aligned}
\big|(By_j(t)-By(t),\varphi)\big| &\le \big|b(y_j(t)-y(t),y_j(t),\varphi)\big| + \big|b(y(t),y_j(t)-y(t),\varphi)\big| \\
&\le C\big|y_j(t)-y(t)\big|^{1/2}\big\|y_j(t)-y(t)\big\|^{1/2}\big|Ay_j(t)\big|^{1/2}\big\|y_j(t)\big\|^{1/2}|\varphi| \\
&\quad + C\|y(t)\|^{1/2}\big\|y_j(t)-y(t)\big\|^{1/2}|y(t)|^{1/2}\big|A(y_j(t)-y(t))\big|^{1/2}|\varphi|.
\end{aligned}$$

Hence,

$$\begin{aligned}
\big|By_j(t)-By(t)\big| \le C\big\|y_j(t)-y(t)\big\|^{1/2}\big(\big|Ay_j(t)\big|^{1/2}\big|y_j(t)-y(t)\big|^{1/2}\big|y_j(t)\big|^{1/2} \\
+ \|y(t)\|^{1/2}\big|A(y_j(t)-y(t))\big|^{1/2}\big|y_j(t)\big|^{1/2}\big).
\end{aligned}$$

We have, therefore,

$$\int_0^T t^2\big|By_j(t)-By(t)\big|^2 dt \to 0 \quad \text{as } j\to\infty.$$

Letting $j\to\infty$, we conclude that y satisfies, a.e. on $(0,T)$, equation (5.191) and that

$$\begin{aligned}
t\|y(t)\|^2+|y(t)|^2+\int_0^T (\|y(t)\|^2+t|Ay(t)|^2)dt \le C, \\
\int_0^T \left(\left\|\frac{dy}{dt}(t)\right\|_{V'}^{2/(1+\varepsilon)} + t\left|\frac{dy}{dt}(t)\right|^2\right) dt \le C,
\end{aligned}$$

where d/dt is considered in the sense of distributions.

If $y_0 \in V$, then we have

$$\big\|y_j(t)\big\|^2+\nu_0\int_0^T \big|Ay_j(t)\big|^2 dt \le C$$

and this implies the last part of the theorem. This completes the proof. (The uniqueness follows as in the proof of Theorem 5.9.) □

Theorem 5.11. *Let $N = 3$, $y_0 \in V$, and $f \in L^2(0,T;H)$. Then there is*

$$T_0^* = T(\|y_0\|, \|f\|_{L^2(0,T;H)})$$

such that on $(0,T_0^*)$ equation (5.151) *has a unique solution*

$$y \in L^\infty(0,T_0^*;V) \cap L^2(0,T_0^*;D(A)) \cap C([0,T_0^*];H)$$
$$\frac{dy}{dt} \in L^2(0,T_0^*;H), \qquad By \in L^2(0,T_0^*;H).$$

Proof. Let $\{y_0^j\}$ and $\{f_j\}$ be as in the proof of Theorem 5.10 ($y_0^j \to y_0$ in V this time.) By the above estimates (see (5.188)), we have

$$\|y_j(t)\|^2 + \nu_0 \int_0^{T_0^*} |Ay_j(t)|^2 dt \le C\left(\|y_0\|^2 + \frac{1}{\nu_0}\int_0^T |f(s)|^2 ds\right), \qquad \forall t \in [0,T_0^*),$$

where $T_0^* < T^* < T$.

We also have (see (5.165))

$$|By_j(t)| \le C\|y_j(t)\|^{3/2}|Ay_j(t)|^{1/2}|y_j(t)|^{1/2} \le C_1|Ay_j(t)|^{1/2}, \qquad \forall t \in (0,T_0^*).$$

Hence,

$$\int_0^{T_0^*}\left(|By_j(t)|^2 + \left|\frac{dy_j}{dt}(t)\right|^2\right) dt \le C.$$

Hence, on a subsequence

$$\begin{aligned} y_j(t) &\to y(t) && \text{strongly in } H \text{ uniformly on } [0,T] \\ & && \text{weak-star in } L^\infty(0,T;V) \\ \frac{dy_j}{dt} &\to \frac{dy}{dt} && \text{weakly in } L^2(0,T;H) \\ Ay_j &\to Ay && \text{weakly in } L^2(0,T;H) \\ By_j &\to \eta && \text{weakly in } L^2(0,T;H). \end{aligned}$$

Moreover, by the Aubin compactness theorem we have $y_j \to y$ strongly in $L^2(0,T;V)$. Note also that, by (5.165), we have

$$|(By_j - By, \varphi)| \le C(\|y_j - y\|^{3/2}|A(y_j - y)|^{1/2} + \|y_j - y\|\,\|y\|_{3/2})|\varphi|.$$

Hence,

$$\int_0^T \big|By_j - By\big|dt \le C\left(\int_0^T \|y_j - y\|^2 dt\right)^{1/2}\left(\left(\int_0^T \|y_j - y\|\big|A(y_j - y)\big|dt\right)^{1/2}\right.$$
$$\left. + \int_0^T |Ay|^{1/2}\|y\|^{3/2}dt\right) \le C\int_0^T \|y_j - y\|^2 dt \to 0 \quad \text{as } j \to 0$$

and, therefore,

$$By_j \to By \quad \text{strongly in } L^1(0,T;H),$$

which implies that $\eta = By$. Hence, y is a strong solution on $(0,T_0^*)$. The uniqueness is immediate. □

The main existence result for a weak solution to equation (5.151) ((5.151)$'$) is Leray's theorem below.

Theorem 5.12. *Let $y_0 \in H$, $f \in L^2(0,T;V')$. Then there is at least one weak solution y^* to equation* (5.151). *Moreover,*

$$\frac{dy^*}{dt} \in L^{4/3}(0,T;V') \qquad \textit{for } N = 3. \tag{5.192}$$

$$\frac{dy^*}{dt} \in L^{2/(1+\varepsilon)}(0,T;V') \quad \textit{for } N = 2. \tag{5.193}$$

If $N = 2$, there is a unique weak solution satisfying (5.193).

Proof. We return to approximating equation (5.176) and note the estimates

$$|y_M(t)|^2 + \int_0^T \|y_M(t)\|^2 dt \le C\left(|y_0|^2 + \int_0^T |f(t)|_*^2 dt\right). \tag{5.194}$$

(For simplicity, we denote below by $|\cdot|_*$ the norm $\|\cdot\|_{V'}$ of V'.) We also have by (5.165)

$$|(B_M y_M(t), w)| \le C\|y_M(t)\|_{1/2}\|y_M(t)\|\|w\| \le C|y_M(t)|^{1/2}\|y_M(t)\|^{3/2}\|w\|.$$

Hence, $|B_M y_M|_* \le C\|y_M\|^{3/2}|y_M|^{1/2}$ and, therefore,

$$\int_0^T |B_M y_M(t)|_*^{4/3} dt \le C\left(|y_0|^2 + \int_0^T |f(t)|_*^2 dt\right) \tag{5.195}$$

$$\int_0^T \left|\frac{dy_M}{dt}(t)\right|_*^{4/3} dt \le C\left(|y_0|^2 + \int_0^T |f(t)|_*^2 dt\right). \tag{5.196}$$

For $N = 2$ we have (see (5.165)) for $m_1 = \varepsilon$, $m_2 = 0$, $m_3 = 1$,

$$|B_M y_M(t)|_* \le C|y_M(t)|^{1-\varepsilon}\|y_M(t)\|^{1+\varepsilon} \le C_1\|y_M(t)\|^{1+\varepsilon}.$$

Hence,

$$\int_0^T \left(\left| \frac{dy_M}{dt} \right|_*^{2/(1+\varepsilon)} + |B_M y_M|_*^{2/(1+\varepsilon)} \right) dt \le C \quad \text{for } N = 2. \tag{5.197}$$

Assume now that $y_0 \in H$ and $f \in L^2(0,T;V')$.

Let $y_0^j \in D(A)$ and $\{f_j\} \subset W^{1,1}([0,T];H)$ be such that

$$y_0^j \to y_0 \ \text{ in } H, \qquad f_j \to f \ \text{ in } L^2(0,T;V').$$

Let y_j be the corresponding solution to equation $(5.151)'$. By estimates (5.195)–(5.197), we have for a constant C independent of M,

$$\int_0^T \left(\|y_j\|^2 + \left\| \frac{dy_j}{dt} \right\|_*^{4/3} + |B_M y_j|_*^{4/3} \right) dt + |y_j(t)|^2 \le C \tag{5.198}$$

if $N = 3$, and

$$\int_0^T \left(\|y_j(t)\|^2 + \left\| \frac{dy_j}{dt} \right\|_*^{2/(1+\varepsilon)} + |B_M y_j|_*^{2/(1+\varepsilon)} \right) dt + |y_j(t)|^2 \le C \tag{5.199}$$

if $N = 2$.

Hence, on a subsequence we have

$$\begin{aligned}
y_j &\to y_M && \text{weakly in } L^2(0,T;V)\\
Ay_j &\to Ay_M && \text{weakly in } L^2(0,T;V')\\
\frac{dy_j}{dt} &\to \frac{dy_M}{dt} && \text{weakly in } L^{4/3}(0,T;V') \text{ if } N = 3\\
& && \text{weakly in } L^{2/(1+\varepsilon)}(0,T;V') \text{ if } N = 2\\
B_M y_j &\to \eta_M && \text{weakly in } L^{4/3}(0,T;V') \text{ if } N = 3\\
& && \text{weakly in } L^{2/(1+\varepsilon)}(0,T;V') \text{ if } N = 2.
\end{aligned}$$

Moreover, recalling inequality (5.172) we get

$$\begin{aligned}
&\frac{1}{2}\frac{d}{dt}|y_j(t) - y_k(t)|^2 + \frac{\nu_0}{2}\|y_j(t) - y_k(t)\|^2\\
&\le \alpha_M |y_j(t) - y_k(t)|^2 + |f_j(t) - f_k(t)| \|y_j(t) - y_k(t)\|_*.
\end{aligned}$$

By Gronwall's lemma we have

$$|y_j(t) - y_k(t)|^2 \le \left| y_0^j - y_0^k \right|^2 + C \int_0^T |f_j(t) - f_k(t)|_*^2 dt$$

and, therefore,

$$\int_0^T \|y_j(t) - y_k(t)\|^2 dt \le C\left(\left|y_0^j - y_0^k\right|^2 + \int_0^T |f_j(t) - f_k(t)|_*^2 dt\right).$$

Hence,

$$y_j \to y_M \quad \text{strongly in } L^2(0,T;V) \cap C([0,T];H).$$

Clearly, we have

$$\begin{cases} \dfrac{dy_M}{dt}(t) + \nu A y_M(t) + \eta_M(t) = f(t), \quad \text{a.e. } t \in (0,T) \\ y_M(0) = y_0. \end{cases}$$

On the other hand, by (5.165), where $m_1 = 1$, $m_2 = 0$, $m_3 = 1$, it follows that

$$\left|B_M y_j - B_M y_M\right|_* \le C\|y_j - y_j\|(\|y_j\| + \|y_M\|).$$

Hence,

$$B_M y_j \to B_M y_M = \eta_M \quad \text{strongly in } L^1(0,T;V').$$

We have shown therefore that for each $y_0 \in H$ and $f \in L^2(0,T;V')$ the equation

$$\begin{cases} \dfrac{dy_M}{dt}(t) + \nu A y_M(t) + B_M y_M(t) = f(t), \quad \text{a.e. } t \in (0,T) \\ y_M(0) = y_0 \end{cases} \tag{5.200}$$

has a solution $y_M \in L^2(0,T;V) \cap C([0,T];H)$ with $dy_M/dt \in L^{4/3}(0,T;V')$ if $N = 3$, $dy_M/dt \in L^{2/(1+\varepsilon)}(0,T;V')$ if $N = 2$. Moreover, y_M satisfies estimates (5.194)–(5.196).

Now, we let $M \to \infty$. Then on a subsequence, again denoted M, we have

$$\begin{aligned}
y_M &\to y^* && \text{weak-star in } L^\infty(0,T;H) \\
& && \text{weakly in } L^2(0,T;V) \\
\frac{dy_M}{dt} &\to \frac{dy^*}{dt} && \text{weakly in } L^{4/3}(0,T;V') \text{ if } N = 3 \\
& && \text{weakly in } L^{2/(1+\varepsilon)}(0,T;V') \text{ if } N = 2 \\
Ay_M &\to Ay^* && \text{weakly in } L^2(0,T;V') \\
B_M y_M &\to \eta && \text{weakly in } L^{4/3}(0,T;V') \text{ if } N = 3 \\
& && \text{weakly in } L^{2/(1+\varepsilon)}(0,T;V') \text{ if } N = 2.
\end{aligned}$$

We have

$$\begin{cases} \dfrac{dy^*}{dt}(t) + \nu_0 A y^*(t) + \eta(t) = f(t), \quad \text{a.e. in } (0,T) \\ y^*(0) = y_0. \end{cases} \tag{5.201}$$

To conclude the proof it remains to be shown that $\eta(t) = By^*(t)$, a.e. $t \in (0,T)$.

We note first that, by Aubin's compactness theorem, for $M \to \infty$,

$$y_M \to y^* \quad \text{strongly in } L^2(0,T;H).$$

We note also that by (5.194) we have $m\{t;\ \|y_M(t)\| > M\} \le C/M^2$.

Let $\varphi \in L^\infty(0,T;\mathscr{V})$. Then, we have

$$\begin{aligned}
&\int_0^T |(B_M y_M - By^*, \varphi)|dt \\
&\le \int_{E_M} |(By_M - By^*, \varphi)|dt + C\int_{E_M^c} \|\varphi\|(|y_M|^{1/2}\|y_M\|^{3/2} + |y^*|^{1/2}\|y^*\|^{3/2})dt,
\end{aligned}$$

where $E_M = \{t;\ \|y_M(t)\| > M\}$. Hence, by estimates (5.194) we have

$$\begin{aligned}
&\int_0^T |(B_M y_M - By^*, \varphi)|dt \\
&\le \int_0^T (|b(y_M - y^*, y_M, \varphi)| + |b(y^*, y_M - y^*, \varphi)|)dt + CM^{-2}\|\varphi\|_{L^\infty(0,T;V)}.
\end{aligned}$$

Recalling that $y_M \to y^*$ strongly in $L^2(0,T;H)$ and weakly in $L^2(0,T;V)$, we get

$$\lim_{M\to\infty} \int_0^T (B_M y_M - By^*, \varphi)dt = 0, \qquad \forall \varphi \in L^2(0,T;\mathscr{V}),$$

where $\mathscr{V} = \{\varphi \in C_0^\infty(\Omega);\ \operatorname{div}\varphi = 0\}$. Hence, $\eta = By^*$ and this concludes the proof.

If $N = 2$, the solution is unique. Indeed, for two such solutions y_1, y_2 we have

$$\frac{1}{2}\,\frac{d}{dt}|y_1 - y_2|^2 + \nu_0\|y_1 - y_2\|^2 + b(y_1 - y_2, y_1, y_1 - y_2) = 0, \quad \text{a.e. } t \in (0,T).$$

This yields

$$\begin{aligned}
\frac{1}{2}\,\frac{d}{dt}|y_1 - y_2|^2 + \nu_0\|y_1 - y_2\|^2 &\le C\|y_1 - y_2\|_{1/2}\|y_1\|\|y_1 - y_2\|_{1/2} \\
&\le C|y_1 - y_2|\|y_1 - y_2\|\|y_1\|.
\end{aligned}$$

By Gronwall's lemma, we get $y_1 = y_2$. □

Remark 5.6. The existence results presented in this section are classic and can be found in a slightly different form in the monographs of Temam [39], Constantin and Foias [19]. However, the semigroup approach used here is new and it closely follows the work of Barbu and Sritharan [6].

Perhaps the main advantage of the semigroup approach is that one can apply the general theory developed in Chapter 4 to get existence, regularity, and approximation results for Navier–Stokes equations.

In fact, as shown earlier, the Navier–Stokes flow $t \to y(t)$ is the restriction to $[0,T]$ of the flow $t \to y_M(t)$ generated by an equation of quasi-m-accretive type.

Bibliographical Remarks

There is an extensive literature on semilinear parabolic equations, parabolic variational inequalities, and the Stefan problem (see Lions [33], Duvaut and Lions [22], Friedman [27]and Elliott and Ockendon [23] for significant results and complete references on this subject). Here, we were primarily interested in the existence results that arise as direct consequences of the general theory developed previously, and we tried to put in perspective those models of free boundary problems that can be formulated as nonlinear differential equations of accretive type. The L^1-space semigroup approach to the nonlinear diffusion equation was initiated by Bénilan [8] (see also Konishi [29]), and the $H^{-1}(\Omega)$ approach is due to Brezis [15]. The smoothing effect of the semigroup generated by the semilinear elliptic operator in $L^1(\Omega)$ (Proposition 5.5) is due to Evans [24, 25]. The analogous result for the nonlinear diffusion operator in $L^1(\Omega)$ (Theorem 5.4) was first established by Bénilan [8], and Véron [41], but the proof given here is essentially due to Pazy [36]. For other related contributions to the existence and regularity of solutions to the porous medium equation, we refer to Bénilan, Crandall, and Pierre [10], and Brezis and Crandall [16]. The semigroup approach to the conservation law equation (Theorem 5.6) is due to Crandall [20]. Theorem 5.7 along with other existence results for abstract hyperbolic equations has been established by Brezis [15] (see also Haraux's book [28] and Barbu [4]). The semigroup approach to Navier–Stokes equations was developed in the works of Barbu [3] and Barbu and Sritharan [6] (see also Barbu and Sritharan [7] and Lefter [32] for other results in this direction).

References

1. N. Alikakos, R. Rostamian, Large time behaviour of solutions of Neumann boundary value problems for the porous medium equations, *Indiana Univ. Math. J.*, **39** (1981), pp. 749–785.
2. A.N. Antontsev, J.I. Diaz, S. Shmarev, *Energy Methods for Free Boundary Problems*, Birkhäuser, Basel, 2002.
3. V. Barbu, *Analysis and Control of Nonlinear Infinite Dimensional Systems*, Academic Press, Boston, 1993
4. V. Barbu, *Nonlinear Semigroups and Differential Equations in Banach Spaces*, Noordhoff, Leyden, 1976.
5. V. Barbu, I. Lasiecka, M. Rammaha, Blow up of generalized solutions to wave equations with nonlinear degenerate damping and source term, *Trans. Amer. Math. Soc.*, **357** (2005), pp. 2571–2611.
6. V. Barbu, S. Sritharan, Flow invariance preserving feedback controllers for the Navier–Stokes equations, *J. Math. Anal. Appl.* **255** (2001), 281–307.
7. V. Barbu, S. Sritharan, *m*-accretive quantization of vorticity equation, Semigroup of operators: Theory and applications, *Progress in Nonliner Differentiable Equations*, **42**, 2000, pp. 296–303, Birkhäuser, Basel.
8. Ph. Bénilan, Opérateurs accétifs et semigroupes dans les espaces L^p, $1 \leq p \leq \infty$, *Functional Analysis and Numerical Analysis*, pp. 15–51, T. Fuzita (Ed.), Japan Soc., Tokyo, 1978.
9. Ph. Bénilan, M.G. Crandall, The continuous dependence on φ of solutions of $u_t - \Delta\varphi(u) = 0$, *Indiana Univ. Math. J.,* **30** (1981), pp. 161–177.

10. Ph. Benilan, M.G. Crandall, M. Pierre, Solutions of the porous medium equations in $\mathbf{R}^N$ under optimal conditions on initial values, *Indiana Univ. Math. J.*, **33** (1984), pp. 51–87.
11. Ph. Benilan, N. Kruzkhov, Conservation laws with continuous flux conditions, *Nonlinear Differential Eqs. Appl.*, **3** (1996), pp. 395–419.
12. E. Bonetti, P. Colli, M. Fabrizio, G. Gilardi, Existence and boundedness of solutions for a singular phase field system, *J. Differential Eqs.*, **246** (2009), pp. 3260–3295.
13. H. Brezis, Intégrales convexes dans les espaces de Sobolev, *Israel J. Math.*, **13** (1972), pp. 9–23.
14. H. Brezis, Problèmes unilatéraux, *J. Math. Pures Appl.*, **51** (1972), pp. 1–168.
15. H. Brezis, Monotonicity methods in Hilbert spaces and some applications to nonlinear partial differential equations, *Contributions to Nonlinear Functional Analysis*, E. Zarantonello (Ed.), Academic Press, New York, 1971.
16. H. Brezis, M.G. Crandall, Uniqueness of solutions of the initial-value problem for $u_t - \Delta\varphi(u) = 0$, *J. Math. Pures Appl.*, **58** (1979), pp. 153–163.
17. H. Brezis, A. Friedman, Nonlinear parabolic equations involving measures as initial conditions, *J. Math. Pures Appl.*, **62** (1983), pp. 73–97.
18. G. Caginalp, An analysis of a phase field model of a free boundary, *Arch. Ration. Mech. Anal.*, **92** (1996), pp. 206–245.
19. P. Constantin, C. Foias, *Navier-Stokes Equations*, The University of Chicago Press, Chicago, 1988.
20. M.G. Crandall, The semigroup approach to the first order quasilinear equations in several space variables, *Israel J. Math.*, **12** (1972), pp. 108–132.
21. G. Duvaut, Résolution d'un problème de Stefan, *C.R. Acad. Sci. Paris*, **267** (1973), pp. 1461–1463.
22. G. Duvaut, J.L. Lions, *Inequalities in Mechanics and Physics,* Springer-Verlag, Berlin, 1976.
23. C.M. Elliott, J.R. Ockendon, *Weak and Variational Methods for Moving Boundary Value Problems*, Pitman, London, 1992.
24. L.C. Evans, Applications of nonlinear semigroup theory to certain partial differential equations, *Proc. Symp. Nonlinear Evolution Equations*, M.G. Crandall (Ed.), Academic Press, New York (1978), pp. 163–188.
25. L.C. Evans, Differentiability of a nonlinear semigroup in L^1, *J. Math. Anal. Appl.*, **60** (1977), pp. 703–715.
26. M. Fremond, *Non-Smooth Thermo-Mechanics*, Springer-Verlag, Berlin, 2002.
27. A. Friedman, *Variational Principles and Free-Boundary Problems,* John Wiley, New York, 1983.
28. A. Haraux, *Semilinear Hyperbolic Problems in Bounded Domains*, Mathematical Reports, vol. **3**, Paris, 1989.
29. Y. Konishi, On the nonlinear semigroups associated with $u_t = \Delta\beta(u)$ and $\phi(u_t) = \Delta u$, *J. Math. Soc.*, **25** (1973), pp. 622–627.
30. S.N. Kružkov, First-order quasilinear equations in several independent variables, *Math. Sbornik*, **10** (1970), pp. 217–236.
31. O.A. Ladyzenskaya, V.A. Solonnikov, N.N. Ural'ceva, *Linear and Quasilinear Equations of Parabolic Type*, American Mathematical Society Transl., American Mathematical Society, Providence, RI, 1968.
32. I. Lefter, Navier-Stokes equations with potentials, *Abstract Appl. Anal.*, **30** (2007), p. 30.
33. J.L. Lions, *Quelques Méthodes de Résolution des Problèmes aux Limites Nonlinéaires*, Dunod, Gauthier–Villars, Paris 1969.
34. G. Marinoschi, A free boundary problem describing the saturated-unsaturated flow in porous medium, *Abstract App. Anal.*, 2005, pp. 813–854.
35. G. Marinoschi, *Functional Approach to Nonlinear Models of Water Flows in Soils*, Springer, New York, 2006.
36. A. Pazy, The Lyapunov method for semigroups of nonlinear contractions in Banach spaces, *Journal d'Analyse Mathématiques*, **40** (1982), pp. 239–262.
37. M. Reed, B. Simon, *Methods of Modern Mathematical Physics*, American Mathematical Society, New York, 1979.

38. J. Serrin, G. Todorova, E. Vitillaro, Existence for a nonlinear wave equation with damping and source terms, *Diff. Integral Eqs.*, **16** (2003), pp. 13–50.
39. R. Temam, *Navier–Stokes Equations and Nonlinear Functional Analysis*, SIAM Philadelphia, 1983.
40. J.L. Vasquez, *The Porous Medium Equation*, Oxford University Press, Oxford, UK, 2006.
41. V. Veron, Effects regularisants de semigroupes nonlinéaire dans les espaces de Banach, *Annales Faculté Sciences Toulouse*, **1** (1979), pp. 171–200.
42. A. Visintin, *Differential Models of Hysteresis*, Springer-Verlag, Berlin, 1994.

Index