CHAPTER II

Nonlinear Stationary Problems

II.1. Banach Spaces

Banach space is a complete normed linear space X. Its *dual space* X' is the linear space of all continuous linear functionals $f : X \to \mathbb{R}$, and it has norm $\|f\|_{X'} \equiv \sup\{|f(x)| : \|x\| \le 1\}$; X' is also a Banach space. We shall denote the value of $f \in X'$ at $x \in X$ by either $f(x)$ or $\langle f, x\rangle$. Likewise from X' we construct the *bidual* or *second dual* $X'' = (X')'$. Furthermore, with each $x \in X$ we can define $\varphi(x) \in X''$ by $\varphi(x)(f) \equiv f(x)$, $f \in X'$; this satisfies clearly $\|\varphi(x)\| \le \|x\|$. Moreover, for each $x \in X$ there is an $f \in X'$ with $f(x) = \|x\|$ and $\|f\| = 1$, so it follows that $\|\varphi(x)\| = \|x\|$. Since φ is linear we see that $\varphi : X \to X''$ is a linear isometry of X onto a closed subspace of X''; we denote this by $X \hookrightarrow X''$. If φ is onto X'' we say X is *reflexive*: $X \cong X''$. Closed subspaces, duals and products of reflexive spaces are likewise reflexive.

Convergence in X is the usual *norm convergence* or *strong convergence*: a sequence $\{x_n\}$ in X converges to x if $\lim_{n\to\infty} \|x_n - x\| = 0$, and this is denoted by $x_n \to x$ or $\lim_{n\to\infty} x_n = x$. This is related to the strong topology on X with neighborhood basis $B(r) \equiv \{x \in X : \|x\| < r\}$, $r > 0$, at the origin. There is also a *weak topology* on X obtained from the base of neighborhoods $B(r; f_1, f_2, \dots, f_n) \equiv \{x \in X : |f_j(x)| < r,\ 1 \le j \le n\}$, $r > 0$, $f_j \in X'$. This is the weakest topology on X for which every $f \in X'$ is continuous; a (net or) sequence $\{x_n\}$ in X is *weakly convergent* to x if $\lim_{n\to\infty} f(x_n) = f(x)$ for every $f \in X'$, and this is denoted by $x_n \rightharpoonup x$ or $w - \lim_{n\to\infty} x_n = x$. Finally, if X happens to be a dual space, say $X = Y'$, there is also the *weak* topology* on X obtained from the neighborhood basis $B(r; f_1, \dots, f_n)$ as above for $r > 0$ but $f_j = \varphi(y_j)$, $1 \le j \le n$, i.e., $\widetilde{B}(r; y_1, \dots, y_n) = \{x \in X : |x(y_j)| < r,\ 1 \le j \le n\}$. This is precisely pointwise convergence, and it is weaker than the weak topology on X; it is equivalent to it exactly when X is reflexive.

Finally, we remark on compactness in the three topologies. First, the unit ball $B \equiv \{x \in X : \|x\| \le 1\}$ in X is compact if and only if $\dim(X) < \infty$. (The same holds for B' in X', of course.) Second, B' is weakly compact in X' if and only if X is reflexive, and, third, B' is always weakly* compact in X'. (This follows since B' is a closed subspace of a product of compact spaces.) The following very deep result is important in many applications; see [59, p.141].

THEOREM 1.1 (EBERLEIN-SHMULYAN). *A Banach space is reflexive if and only if it is sequentially weakly compact, i.e., every bounded sequence contains a weakly convergent subsequence.*

A set $S \subset X$ is *bounded* if $S \subset B(r)$ for some $r > 0 : x \in S \Rightarrow \|x\| < r$. Likewise it is *weakly bounded* if some multiple of each weak neighborhood base contains S : for each $f \in X'$ there is an r for which $x \in S \Rightarrow |f(x)| < r$.

THEOREM 1.2 (UNIFORM BOUNDEDNESS). *A set S in Banach space is weakly bounded if and only if it is bounded.*

Let X and Y be Banach spaces and $T : X \to Y$ a function. Then T is *continuous* (*weakly continuous*) if it is continuus with X and Y each endowed with their corresponding norm (resp., weak) topologies. Similar terminology is used for weak* when both are dual spaces, and the modifier "sequentially" means that sequences (not nets) are considered. T is called *demicontinuous* if it is continuous from X with norm convergence to Y with weak convergence, and it is called *completely continuous* if it is continuous from X with weak to Y with strong convergence.

$$\begin{array}{ccccc} & \nearrow & \text{weakly continuous} & \searrow & \\ \text{completely continuous} & & & & \text{demicontinuous} \\ & \searrow & \text{continuous} & \nearrow & \end{array}$$

The function T is *bounded* if S bounded in X implies the image $T(S)$ is bounded in Y, and it is *compact* if it is continuous and S bounded implies $T(S)$ is *precompact*, i.e., its closure is compact.

Assume hereafter that X is reflexive and $T : X \to Y$.

LEMMA 1.1. *If T is completely continuous, then it is compact.*

PROOF. If S is bounded in X then a sequence in $T(S)$ is given by $\{T(x_n)\}$ with $\{x_n\}$ in S; a subsequence $x_{n'} \rightharpoonup x$ in X so $T(x_{n'}) \to T(x)$ in Y. □

Conversely, let's suppose $T : X \to Y$ is compact. If $x_n \rightharpoonup x$ in X then $\{x_n\}$ is bounded, weakly and strongly. If it is *not* that $T(x_n) \to T(x)$, there is a subsequence $\{x_{n'}\}$ for which $T(x_{n'})$ remains outside a strong neighborhood of $T(x)$. Now $\{x_{n'}\}$ bounded implies $\{T(x_{n'}\}$ contains a convergent subsequence, say $T(x_{n''}) \to y$, and this is a contradiction if $y = T(x)$. Thus we have proved the

LEMMA 1.2. *If T is compact and weakly sequentially continuous, then T is completely sequentially continuous.*

Note that T being weakly sequentially continuous implies that T is bounded. Also we needed above only that the graph of T be *closed* in $X_w \times Y_s$. Here X_w denotes X with weak convergence, and Y_s is Y with strong convergence.

Finally, note that complete continuity or compactness is so severe a restriction that it never holds for the identity in an infinite-dimensional Hilbert space. That is, if $\{e_n\}$ is an orthonormal basis for Hilbert space X, then for each $x \in X$ we have $x = \sum_{n=1}^{\infty}(e_n, x)\, e_n$ and $\|x\|^2 = \sum_{n=1}^{\infty} |(e_n, x)|^2$, so $\lim_{n\to\infty}(e_n, x) = 0$ for each $x \in X \cong X'$. Thus $e_n \rightharpoonup 0$, but $\|e_n\| = 1$ so it is impossible for any subsequence to be strongly convergent.

Assume hereafter that $T : X \to Y$ is *linear*. If T is bounded then by linearity there is a $K > 0$: $\|T(x)\| \le K\|x\|$ for all $x \in X$ and (again, by linearity) this is equivalent to (uniform) continuity. This proves that T is bounded if and only if it is continuous. Recall the *dual* or *adjoint* of T is the linear $T' : Y' \to X'$ defined by $T'(f)(x) = f(T(x))$ for $f \in Y'$, $x \in X$. Thus, $T'(f) = f \circ T \in X'$. If the net $x_\alpha \rightharpoonup x$ in X and $f \in Y'$, then $f(T(x_\alpha)) \equiv (T'f)(x_\alpha) \to (T'f)(x) \equiv f(T(x))$, so $Tx_\alpha \rightharpoonup Tx$ in Y. This shows that T being continuous implies that T is weakly continuous. Suppose X is reflexive. If T is not bounded, there is a bounded sequence $\{x_n\}$ such that $\|Tx_n\| \to \infty$. But then there is a weakly convergent subsequence $\{x_{n'}\}$ for which $\{T(x_{n'})\}$ is not bounded and, hence, not weakly convergent. This shows that T being weakly continuous implies that T is bounded. Thus, for a linear operator in reflexive Banach space, we have the equivalence of continuity, weak continuity and boundedness.

II.2. Existence Theorems

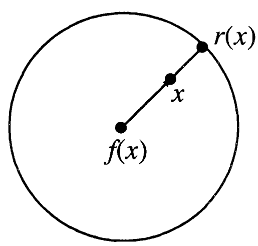
We begin with a fundamental fixed-point result; see [44, p. 82].

THEOREM (BROUWER). *If f is a continuous map of the closed unit ball B^n of $\mathbb{R}^n$ into itself, then there exists an $x \in B^n : f(x) = x$.*

There is another form of this result which is more geometric.

THEOREM. *There does not exist a continuous function r of B^n into its boundary ∂B^n with $r(x) = x \ \forall\ x \in \partial B^n$.*

Such a function r is called a *retract* of B^n. These two results are equivalent. To see this, let r be a retract of B^n. Define the antipodal map $a : \partial B^n \to \partial B^n$ by $a(x) = -x$. Then $a \circ r : B^n \to B^n$ has no fixed point, so the Brouwer Theorem implies the No-Retract Theorem. Conversely, suppose $f : B^n \to B^n$ has no fixed point: $f(x) \neq x$ for $x \in B^n$. Then for each $x \in B^n$, there is a unique $r(x) \in \partial B^n$: $r(x) = x + t(x - f(x))$ with $t \geq 0$, and r is a retract of B^n.



A special and useful case is the following intermediate-value theorem:

PROPOSITION 2.1. *Let $f : B^n \to \mathbb{R}^n$ be continuous and $(f(x), x) \geq 0$ for every $x \in \partial B^n$. Then f has a zero.*

PROOF. Otherwise, the map $x \mapsto -f(x)/\|f(x)\|$ of B^n into B^n has a fixed point $x_0 = -f(x_0)/\|f(x_0)\|$ for which $(f(x_0), x_0) = -\|f(x_0)\| < 0$. □

This is an example of how an *a-priori* estimate and continuity lead to the existence of a solution to a nonlinear equation, $f(x) = 0$. We shall obtain such results in the setting of Banach space where $f : K \to K$ is continuous and K is a compact convex set.

Let V be a reflexive Banach space; consider a function $\mathcal{A} : V \to V'$. We say that $\mathcal{A}$ is *monotone* if $\langle \mathcal{A}(u) - \mathcal{A}(v), u - v\rangle \geq 0 \ \forall\ u, v \in V$, and *hemicontinuous* if for each $u, v \in V$ the real-valued function $t \mapsto \mathcal{A}(u + tv)(v)$ is continuous. (Clearly this last condition is true if the restriction of $\mathcal{A}$ to each line segment is continuous into V' with weak convergence.)

DEFINITION. $\mathcal{A}$ is *type M* if $u_n \rightharpoonup u$, $\mathcal{A}u_n \rightharpoonup f$ and $\limsup \mathcal{A}u_n(u_n) \leq f(u)$ imply that $\mathcal{A}u = f$.

LEMMA 2.1. *If $\mathcal{A}$ is hemicontinuous and monotone then it is type M.*

PROOF. Let $\{u_n\}$ be given as above. By monotonicity, $\langle \mathcal{A}u_n - \mathcal{A}v, u_n - v\rangle \geq 0$, $v \in V$ so we obtain $\langle f - \mathcal{A}v, u - v\rangle \geq 0$ for all $v \in V$. For any $w \in V$ we set $v = u - tw$ with $t > 0$ and let t converge to zero; hemicontinuity implies $(f - \mathcal{A}u)(w) \geq 0$ for all $w \in V$, so $\mathcal{A}u = f$. □

LEMMA 2.2. *If $\mathcal{A}$ is type M and bounded then it is demicontinuous.*

PROOF. Let $u_n \to u$ in V. Since $\mathcal{A}$ is bounded there is a susequence $\{u_{n'}\}$ for which $\mathcal{A}(u_{n'}) \rightharpoonup f$. But $|\mathcal{A}u_{n'}(u_{n'}) - f(u)| \leq |(\mathcal{A}u_{n'} - f)(u)| + \|\mathcal{A}u_{n'}\| \, \|u_{n'} - u\| \to 0$ so $\lim_{n\to\infty} \mathcal{A}u_{n'}(u_{n'}) = f(u)$. Since $\mathcal{A}$ is type M it follows that $\mathcal{A}u = f$; since a subsequence as above can be extracted from *any* susequence of $\{u_n\}$, it follows that $\mathcal{A}u_n \rightharpoonup \mathcal{A}(u)$. □

COROLLARY 2.1. *Let V_0 be a finite dimensional space, $j : V_0 \to V$ an injection and $j' : V' \to V_0'$ the dual operator. Then $j'\mathcal{A}j : V_0 \to V_0'$ is continuous.*

THEOREM 2.1. *Let V be separable reflexive Banach space and $f \in V'$. Assume $\mathcal{A} : V \to V'$ is type M, bounded, and there is a $\rho > 0$ such that*

$$\mathcal{A}v(v) > f(v) \ \forall \ v \in V : \|v\| > \rho \ . \tag{2.1}$$

Then $f \in Rg(\mathcal{A})$: there exists $u \in V : \mathcal{A}u = f$.

PROOF. We use Galerkin's method to reduce the problem to the finite-dimensional case. Let $\{w_1, w_2, \dots\}$ be an independent set of vectors whose linear span is dense in V, and let V_m denote the linear span of $\{w_1, w_2, \dots, w_m\}$ for each $m \geq 1$. Let $j_m : \mathbb{R}^m \to V_m$ denote representation with respect to this basis.

Fix $m \geq 1$ and consider the approximate problem in V_m',

$$u_m \in V_m : \mathcal{A}(u_m)(w_j) = f(w_j) \ , \qquad 1 \leq j \leq m \ .$$

This is equivalent to finding $u_m = j_m(\hat{u}_m)$, where

$$\hat{u}_m \in \mathbb{R}^m : F(\hat{u}_m) \equiv (j_m' \mathcal{A} j_m)(\hat{u}_m) - j_m' f = 0 \ .$$

But $F : \mathbb{R}^m \to \mathbb{R}^m$ is continuous and satisfies $(F(u), u)_{\mathbb{R}^m} \geq 0$ for $\|u\| \geq \rho$ so F has a zero.

For each $m \geq 1$ we have a u_m as above and, hence, $\mathcal{A}u_m(u_m) = f(u_m)$. Again from (2.1) we find $\|u_m\| \leq \rho$, and $\mathcal{A}$ is bounded so there is a subsequence (denoted by $\{u_m\}$ again) for which $u_m \rightharpoonup u$, $\mathcal{A}(u_m) \rightharpoonup f$ and $\mathcal{A}u_m(u_m) = f(u_m) \to f(u)$. Since $\mathcal{A}$ is type M, $\mathcal{A}(u) = f$. □

DEFINITION. The function $\mathcal{A} : V \to V'$ is *coercive* if

$$\frac{\mathcal{A}u(u)}{\|u\|} \to \infty \quad \text{as } \|u\| \to \infty \ .$$

COROLLARY 2.2. *If $\mathcal{A}$ is type M, bounded and coercive on a separable reflexive Banach space to its dual then $\mathcal{A}$ is surjective.*

EXAMPLE. Let $A : \mathbb{R} \to \mathbb{R}$ be monotone, continuous and $f \in \mathbb{R}$. Then "$A(r)r > fr$ for large $|r|$" if and only if f belongs to the interior of $Rg(A)$.

REMARK. An operator $\mathcal{A}$ is called *locally bounded* if the image of any convergent sequence is bounded. The proof of Lemma 2.2 shows that this is sufficient with type M to obtain demicontinuity. Thus, for monotone $\mathcal{A}$, demicontinuity is equivalent to hemicontinuity and local boundedness. A deeper result of Browder and Rockafellar is that monotonicity implies local boundedness, so for monotone operators, hemicontinuity and demicontinuity are equivalent.

If uniqueness holds for the equation $\mathcal{A}(u) = f$, then the *original sequence* $\{u_m\}$ in the proof of Theorem 2.1 must converge (weakly) to u. This follows from an argument similar to the proof of Lemma 1.2. Sufficient conditions for uniqueness are the following.

DEFINITIONS. $\mathcal{A}$ is *strictly monotone* if $\langle \mathcal{A}u - \mathcal{A}v, u - v\rangle > 0$ for all $u \neq v$ in V and $\mathcal{A}$ is *strongly monotone* if there is a $c > 0$ for which

$$\langle \mathcal{A}u - \mathcal{A}v, u - v\rangle \geq c\|u - v\|_V^2 \ , \qquad u, v \in V \ .$$

Note that strongly monotone implies coercive and that in Theorem 2.1 we have $u_m \to u$ (strongly). Moreover, for monotone operators we have the following useful characterization of solutions.

PROPOSITION 2.2 (MINTY). *If $\mathcal{A} : V \to V'$ is monotone and hemicontinuous, then $\mathcal{A}u = f$ if and only if $\langle \mathcal{A}v - f, v - u\rangle \geq 0$ for all $v \in V$. If also $\mathcal{A}$ is bounded and satisfies (2.1), then the set of solutions $U \equiv \{u \in V : \mathcal{A}u = f\}$ is closed, convex, non-empty and bounded.*

PROOF. Clearly $\mathcal{A}u = f$ implies $\langle \mathcal{A}v - f, v - u\rangle \geq 0$ for $v \in V$. Conversely, set $v = u + tw$ and let t converge to zero to obtain $\langle \mathcal{A}u - f, w\rangle \geq 0$ for all $w \in V$. Thus $U = \bigcap_{v \in V} C_v$ where $C_v = \{u \in V : \langle \mathcal{A}v - f, v - u\rangle \geq 0\}$ is closed and convex. If $\mathcal{A}$ satisfies (2.1) then U is bounded and, by Theorem 2.1, $\mathcal{A}$ also bounded implies U is non-empty. □

DEFINITION. $\mathcal{A}$ is *maximal monotone* if it is monotone and $\mathcal{A}u = f$ if (and only if) $\langle \mathcal{A}v - f, v - u\rangle \geq 0$ for all $v \in V$.

Thus $\mathcal{A}$ has no monotone proper extension. From Proposition 2.2 and the proof of Lemma 2.1 it follows that monotone and hemicontinuous imply maximal monotone which in turn implies type M.

A useful variation on Theorem 2.1 is given next; although boundedness of $\mathcal{A}$ is relaxed, $\mathcal{A}$ is locally bounded by the preceding Remark.

THEOREM 2.2 (MINTY-BROWDER). *Let V be a separable reflexive Banach space and $f \in V'$. Assume $\mathcal{A} : V \to V'$ is monotone, demicontinuous and that (2.1) holds. Then $f \in Rg(\mathcal{A})$.*

PROOF. As before we obtain a sequence $u_m \in V_m$ such that $\langle \mathcal{A}u_m - f, w\rangle = 0$ for $w \in V_m$ and $u_m \rightharpoonup u$. If $v \in V_n$ and $m \geq n$, then $0 \leq \langle \mathcal{A}v - \mathcal{A}u_m, v - u_m\rangle = \langle \mathcal{A}v - f, v - u_m\rangle$ since $v - u_m \in V_m$, hence, $0 \leq \langle \mathcal{A}v - f, v - u\rangle$ for all $v \in \bigcup_{n \geq 1} V_n$.

The same holds for all $v \in V$ by demicontinuity since $\bigcup_{n\geq 1} V_n$ is dense in V, and the maximal monotonicity shows $\mathcal{A}u = f$. □

We have shown that a monotone continuous operator is type M and have proven an existence result for *equations* with these rather general operators. Next we give some examples of type M operators, one of which indicates that this class is very sensitive to perturbations, and then we introduce the more stable class of *pseudo-monotone operators* for which we can resolve a wider class of problems, namely, *inequalities*. In Section 6 we shall present a variety of examples of variational inequalities and then introduce a class of elliptic operators which are *quasimonotone*, that is, monotone in only the highest order terms.

Let $\mathcal{A} : V \to V'$ be a function, V a separable reflexive Banach space. We list some examples of conditions which imply $\mathcal{A}$ is type M, that is, if $u_n \rightharpoonup u$ in V, $\mathcal{A}(u_n) \rightharpoonup f$ in V' and $\limsup \mathcal{A}u_n(u_n) \leq f(u)$, then $\mathcal{A}(u) = f$.

EXAMPLE 2.A. Let $\mathcal{A}$ be *weakly closed*: $u_n \rightharpoonup u$ and $\mathcal{A}u_n \rightharpoonup f$ imply $\mathcal{A}u = f$. Clearly, $\mathcal{A}$ is then type M. Note that "bounded and weakly closed" is equivalent to "weakly continuous" in this space. Thus every continuous and linear operator from V to V' is type M; this holds without any monotonicity.

EXAMPLE 2.B. Type M plus *completely continuous* is type M: if $\mathcal{A}$ is type M and if B is completely continuous, then $\mathcal{A} + \mathcal{B}$ is type M. This is immediate since $u_n \rightharpoonup u$ implies $\mathcal{B}u_n \to \mathcal{B}u$ and so $\mathcal{B}u_n(u_n) \to \mathcal{B}u(u)$. Thus the type M property is stable under "compact perturbations" in this setting.

EXAMPLE 2.C. Type M plus *monotone, weakly continuous* is type M. Suppose $\mathcal{A}$ is type M and $\mathcal{B}$ is monotone, weakly continuous. Let $u_n \rightharpoonup u$, $(\mathcal{A}+\mathcal{B})(u_n) \rightharpoonup f$ and $\limsup(\mathcal{A}+\mathcal{B})(u_n)(u_n) \leq f(u)$. Since $\mathcal{B}$ is weakly continuous, $\mathcal{B}u_n \rightharpoonup \mathcal{B}u$ and so $\mathcal{A}u_n \rightharpoonup f - \mathcal{B}(u)$. Since $\mathcal{B}$ is monotone we have $\mathcal{A}(u_n)(u_n) \leq \mathcal{A}(u_n)(u) + (\mathcal{A} + \mathcal{B})(u_n)(u_n - u) - \mathcal{B}u(u_n - u)$, so $\limsup \mathcal{A}u_n(u_n) \leq (f - \mathcal{B}u)(u)$. But $\mathcal{A}$ is type M, so $\mathcal{A}(u) = f - \mathcal{B}(u)$.

Even with monotonicity it is difficult to relax the severe condition of weak continuity of the perturbation in the above. The next example shows that even Lipschitz continuity is not sufficient.

EXAMPLE 2.D. Let $V = V'$ be separable Hilbert space with orthonormal basis $\{e_n : n \geq 1\}$. Set $\mathcal{A} = -I$, the negative of the identity, and P the projection onto the unit ball $K = \{v \in V : \|v\| \leq 1\}$. Recall from I.2 that this projection is characterized by

$$Pu \in K : (Pu - u,\ Pu - v) \leq 0 \ \forall\ v \in K$$

Also, we note that for $u, v \in V$

$$\begin{aligned}(Pu - Pv,\ u - v) &= (Pu - Pv,\ u - Pu) + \|Pu - Pv\|^2 + (Pu - Pv,\ Pv - v)\\ &\geq \|Pu - Pv\|^2\ ,\end{aligned}$$

so P is a monotone contraction.

Choose $u_n = e_1 + e_n$, $n \geq 2$. Then $Pu_n = \frac{1}{\sqrt{2}}u_n$, $(A+P)(u_n) = (\frac{1}{\sqrt{2}} - 1)u_n$ so $u_n \rightharpoonup e_1$, $(A+P)(u_n) \rightharpoonup (\frac{1}{\sqrt{2}}-1)e_1 \equiv f$, and $((A+P)(u_n, u_n)) = 2(\frac{1}{\sqrt{2}}-1) < (f, e_1)$. However, $(A + P)(e_1) = 0 \neq f$, so we see that $A + P$ is not type M.

Given $\mathcal{A} : V \to V'$ and $f \in V'$ as above, suppose we wish to solve the *variational inequality* (cf. I.2.12)

$$(2.2) \qquad u \in K : \mathcal{A}u(v-u) \geq f(v-u) \ , \qquad v \in K \ ,$$

where K is a closed convex subset of V. In order to retain information on the value of the left side of this inequality when we take weak limits, for example, in a Galerkin approximation procedure as in the proof of Theorem 2.1, we introduce the following class of operators.

DEFINITION. Let V be a reflexive Banach space. An operator $\mathcal{A} : V \to V'$ is *pseudo-monotone* if $u_n \rightharpoonup u$ and $\limsup \mathcal{A}u_n(u_n - u) \leq 0$ imply $Au(u-v) \leq \liminf Au_n(u_n - v)$ for all $v \in V$.

Note in the above that from $v = u$ we obtain $\lim Au_n(u_n - u) = 0$ and so it follows that $Au(u-v) \leq \liminf Au_n(u-v)$ for all $v \in V$.

PROPOSITION 2.3. *If $\mathcal{A}$ is monotone and hemicontinuous then $\mathcal{A}$ is pseudo-monotone. If $\mathcal{A}$ is pseudo-monotone, then $\mathcal{A}$ is of type M.*

PROOF. Suppose $u_n \rightharpoonup u$ and $\limsup Au_n(u_n - u) \leq 0$. By monotonicity, $\mathcal{A}u_n(u_n-u) \geq \mathcal{A}u(u_n-u) \to 0$ so $\liminf \mathcal{A}u_n(u_n-u) \geq 0$; thus $\lim Au_n(u_n-u) = 0$. Let $v \in V$ and set $w = (1-t)u + tv$, $t > 0$. Then $u_n - w = t(u-v) + (u_n - u)$ and by monotonicity $0 \leq \langle Au_n - Aw,\ u_n - w\rangle$; this implies

$$tAu_n(u-v) \geq -Au_n(u_n - u) + Aw\big(t(u-v) + (u_n - u)\big) \ .$$

Taking the lim inf and dividing by t gives

$$\liminf Au_n(u-v) \geq Aw(u-v) \ .$$

Now let $t \downarrow 0$ and use hemicontinuity to get

$$\liminf Au_n(u-v) \geq Au(u-v) \ .$$

Since lim $Au_n(u_n - u) = 0$, we are done.

For the second part, let $u_n \rightharpoonup u$, $\mathcal{A}u_n \rightharpoonup f$ and lim sup $\mathcal{A}u_n(u_n) \leq f(u)$. Then for any $v \in V$ we have

$$\begin{aligned} \mathcal{A}u(u-v) &\leq \liminf \mathcal{A}u_n(u_n - v) \leq \limsup \mathcal{A}u_n(u_n - v) \\ &\leq f(u-v) \end{aligned}$$

so it follows that $\mathcal{A}u = f$. □

COROLLARY 2.3. *If $\mathcal{A}$ is pseudo-monotone and locally bounded then it is demi-continuous.*

PROOF. This is immediate from Lemma 2.2, but we can easily prove it directly. If $u_n \to u$ then (for some subsequence) we have $\mathcal{A}(u_n) \rightharpoonup f$ in V'. Thus, $\mathcal{A}u_n(u_n - u) \to 0$, so from the definition of pseudomonotone we obtain

$$\mathcal{A}u(u-v) \leq \liminf_{n\to\infty} \mathcal{A}u_n(u_n - v) = f(u-v) \ , \qquad v \in V \ ,$$

and this implies $\mathcal{A}u = f$. This holds for any subsequence, so we conclude $\mathcal{A}u_n \rightharpoonup \mathcal{A}u$ by the uniqueness of weak limits. □

PROPOSITION 2.4. *Let $\mathcal{A}$ and $\mathcal{B}$ be pseudo-monotone; then $\mathcal{A}+\mathcal{B}$ is pseudo-monotone.*

PROOF. Assume $u_n \rightharpoonup u$ and $\limsup(A+B)(u_n)(u_n-u) \leq 0$. We claim that $\limsup \mathcal{A}u_n(u_n-u) \leq 0$ and $\limsup \mathcal{B}u_n(u_n-u) \leq 0$. Otherwise, by symmetry $\limsup \mathcal{B}u_n(u_n-u) = \varepsilon > 0$ and, by passing to a subsequence, $\lim \mathcal{B}u_n(u_n-u) = \varepsilon$. This gives $\limsup \mathcal{A}u_n(u_n-u) = \limsup\{(A+B)u_n(u_n-u) - \mathcal{B}u_n(u_n-u)\} \leq 0-\varepsilon$. Since $\mathcal{A}$ is pseudo-monotone we have $\mathcal{A}u(u-v) \leq \liminf \mathcal{A}u_n(u_n-v)$ for all $v \in V$; setting $v = u$ gives

$$0 \leq \liminf \mathcal{A}u_n(u_n-u) \leq \limsup \mathcal{A}u_n(u_n-u) \leq -\varepsilon ,$$

a contradiction. With the claim established, the proof now follows by the super-additivity of the "lim inf". □

We have shown the class of pseudo-monotone operators is "intermediate" between monotone, hemicontinuous and type M, and that it is stable with respect to addition. Moreover, we note that this class is "strictly intermediate". It is easy to construct pseudo-monotone examples which are not monotone; in Example 2.D we had $(A+P)u_n(u_n-e_1) = \frac{1}{\sqrt{2}} - 1 < 0$, so $A+P$ is not pseudo-monotone. By Propositions 2.3 and 2.4 it follows that $\mathcal{A} = -I$ is not pseudo-monotone, but this is easy to see directly. In fact, $-I$ is pseudo-monotone in Hilbert space V if and only if $\dim(V) < \infty$. In general, for Hilbert space we have $-I$ pseudo-monotone if and only if weak convergence is equivalent to strong convergence.

THEOREM 2.3 (BREZIS). *Let V be a separable reflexive Banach space, K a closed, convex non-empty subset of V, $\mathcal{A} : V \to V'$ a bounded pseudo-monotone operator, and $f \in V'$. Assume there is a $v_0 \in K$ and $\rho > 0$ such that*

$$\mathcal{A}v(v-v_0) > f(v-v_0) \quad \forall\, v \in K\ ,\quad \|v\| \geq \rho\ . \tag{2.3}$$

Then there exists a solution of the variational inequality (2.2).

PROOF. We proceed in four steps. Assume K is bounded, so (2.3) is vacously true, and (i) solve the finite-dimensional problem by Brouwers fixed-point theorem, (ii) use boundedness of $\mathcal{A}$ to get an appropriate subsequence of "approximate" solutions, then (iii) use the pseudo-monotone property to show the weak limit is a solution. Finally, (iv) use (2.3) to eliminate the assumption that K is bounded.

First, let $\{w_1, w_2, \dots\}$ be dense in K, V_m be the linear span of $\{w_1, \dots, w_m\}$, and K_m be the convex hull of $\{w_1, \dots, w_m\}$ for $m \geq 1$. Let $j_m : V_m \to V$ and its dual $j'_m : V' \to V'_m$ denote injection and restriction, respectively. Note that $\cup\{K_m : m \geq 1\}$ is dense in K. Fix $m \geq 1$ and consider the finite-dimensional problem

$$u_m \in K_m : j'_m \mathcal{A} j_m u_m(v-u_m) \geq j'_m f(v-u_m)\ , \qquad v \in K_m\ .$$

This is equivalent to

$$u_m \in K_m : (u_m,\ v-u_m) \geq (u_m + j'_m f - j'_m \mathcal{A} j_m u_m,\ v-u_m)\ , \qquad v \in K_m$$

where we identify $V_m \cong \mathbb{R}^m \cong V'_m$, and this is equivalent to

$$u_m = P(u_m + j'_m f - j'_m \mathcal{A} j_m u_m)$$

where P is projection onto K_m in $\mathbb{R}^m$. But the function $u \mapsto P(u+j'_m f - j'_m \mathcal{A} j_m u)$ is continuous from any closed ball B containing K_m into B so the Brouwer fixed-point theorem shows this equation has a solution.

Since we assumed K is bounded, by passing to a subsequence, denoted by $\{u_m\}$, we have $u_m \rightharpoonup u \in K$, and $\|\mathcal{A}(u_m)\| \leq M$ since $\mathcal{A}$ is bounded. Let's show $\limsup \mathcal{A}u_m(u_m - u) \leq 0$. For any $\varepsilon > 0$ choose $\tilde{u} \in K_N$ with $\|u - \tilde{u}\| < \varepsilon$. Then $\mathcal{A}u_m(u_m - \tilde{u}) \leq f(u_m - \tilde{u})$ for $m \geq N$, since $K_N \subset K_m$, so we obtain

$$\limsup \mathcal{A}u_m(u_m - u) = \limsup\{\mathcal{A}u_m(u_m - \tilde{u}) + \mathcal{A}u_m(\tilde{u} - u)\} \leq (\|f\| + M)\varepsilon \ .$$

But $\varepsilon > 0$ is arbitrary so the claim is done.

Since $\mathcal{A}$ is pseudo-monotone it follows that $\mathcal{A}u(u - v) \leq \liminf \mathcal{A}u_m(u_m - v)$ for all $v \in V$. If $v \in K_n$ and $m \geq n$, then $\mathcal{A}u_m(u_m - v) \leq f(u_m - v)$ so $\mathcal{A}u(u-v) \leq f(u-v)$ for all $v \in K_n$, $n \geq 1$. But $\cup\{K_n : n \geq 1\}$ is dense in K, so u is a solution of (2.2).

It remains only to remove the assumption that K is bounded. Set $R = \max\{\|v_0\|, \rho\}$ and $K_R = \{v \in K : \|v\| \leq R\}$. Since K_R is bounded there is a

$$u_R \in K_R : \mathcal{A}(u_R)(v - u_R) \geq f(v - u_R) \ , \qquad v \in K_R \ .$$

Since $v_0 \in K_R$ it follows $\mathcal{A}(u_R)(u_R - v_0) \leq f(u_R - v_0)$ so (2.3) implies $\|u_R\| < \rho$. For any $v \in K$ set $v_t \equiv (1-t)u_R + tv$ with $t > 0$ so small that $v_t \in K_R$. Observe $v_t - u_R = t(v - u_R)$ so it follows that u_R is a solution of the original problem on K and the proof is finished. □

For the special class of monotone operators, we have a *weak characterization* of solutions as well as additional properties of the set of solutions.

COROLLARY 2.4. *If $\mathcal{A}$ is monotone, hemicontinuous, then u is a solution of the variational inequality* (2.2) *if and only if*

$$u \in K : \mathcal{A}v(v - u) \geq f(v - u) \ , \qquad v \in K \ . \tag{2.4}$$

If also $\mathcal{A}$ is bounded and satisfies (2.3) *then the set of solutions is closed, convex, non-empty and bounded.*

PROOF. Since $\mathcal{A}$ is monotone, every solution satisfies (2.4). Conversely, if u satisfies (2.4) and $v \in K$, we choose $0 < t < 1$ and $v_t = u + t(v-u) = (1-t)u + tv \in K$ as a test vector to obtain $v_t - u = t(v - u)$, hence,

$$\mathcal{A}(v_t)(v - u) \geq f(v - u) \ .$$

Now use hemicontinuity to let $t \to 0^+$; this shows the problems are equivalent. Clearly the solution set is convex and closed, since the set of $u \in K$ satisfying (2.4) for a fixed v is closed and convex and we may take intersections. Also (2.3) implies it is non-empty and bounded. □

We mention some easy but useful remarks. The condition (2.3) implies that any solution of the variational inequality satisfies the explicit bound $\|u\| < \rho$. Thus (2.3) is an *a-priori estimate* on solutions. It is vacously true if K is bounded; if $\mathcal{A}$ is coercive it holds for *every* $f \in V'$. Finally, if $\mathcal{A}$ is strictly monotone there is at most one solution of the variational inequality.

II.3. L^p Spaces

Let G be an open (or Lebesgue measurable) set in $\mathbb{R}^n$ and $1 \le p < \infty$. Then $L^p(G)$ is the set of equivalence classes of measurable functions $u : G \to \mathbb{R}$ such that $\int_G |u(x)|^p \, dx < \infty$, where we identify functions whose values equal a.e. on G. From the (crude) inequality, $|a+b|^p \le 2^p(a^p + b^p)$ for $a, b \ge 0$, it follows that L^p is a linear space. The space $L^\infty(G)$ consists of all (equivalence classes of) *essentially bounded* measurable functions, $u : G \to \mathbb{R}$, for each of which there is a $K < \infty$ with $|u(x)| \le K$ a.e. $x \in G$. The infimum of such K is denoted by $\|u\|_{L^\infty(G)}$; clearly $L^\infty(G)$ is a linear space.

We develop some fundamental inequalities.

LEMMA 3.1 (CAUCHY). *$ab \le \frac{\varepsilon a^2}{2} + \frac{b^2}{\varepsilon 2}$ for $a, b \ge 0$.*

PROOF. $(\sqrt{\varepsilon}\, a - \frac{1}{\sqrt{\varepsilon}}\, b)^2 \ge 0$. □

LEMMA 3.2 (JENSEN). *If $f : G \to \mathbb{R}$ is integrable and $\varphi : \mathbb{R} \to \mathbb{R}$ is convex, then*

$$\varphi\Big(\frac{1}{\mu(G)} \int_G f(x)\, dx\Big) \le \frac{1}{\mu(G)} \int_G \varphi(f(x))\, dx \ .$$

PROOF. Set $a = \frac{1}{\mu(G)} \int_G f$ and choose $m \in \mathbb{R} : m(t-a) + \varphi(a) \le \varphi(t)$ for all $t \in \mathbb{R}$. (That is, m is the slope of a line of support of the graph of φ.) Set $t = f(x)$ and integrate over G. □

LEMMA 3.3 (YOUNG). *$ab \le \frac{a^p}{p} + \frac{b^{p'}}{p'}$ for $a, b \ge 0$ and $\frac{1}{p} + \frac{1}{p'} = 1$.*

PROOF. Set

$$f(x) = \begin{cases} p \log a \ , & 0 \le x \le \frac{1}{p} \\ p' \log b \ , & \frac{1}{p} \le x \le 1 \end{cases} \ , \quad G = (0,1) \ , \quad \varphi(t) = e^t \ .$$ □

LEMMA 3.3A. $ab \le \frac{\varepsilon a^p}{p} + \frac{\varepsilon^{1-p'} b^{p'}}{p'}$.

LEMMA 3.4 (HÖLDER). *$\int_G |u(x)v(x)|\, dx \le \|u\|_{L^p} \|v\|_{L^{p'}}$, $1 \le p, p'$, $\frac{1}{p} + \frac{1}{p'} = 1$, where*

$$\|u\|_{L^p(G)} \equiv \Big(\int_G |u(x)|^p \, dx\Big)^{1/p} \quad \textit{for} \ \ 1 \le p < \infty \ .$$

PROOF. The cases $p = 1, +\infty$ are immediate. Otherwise,

$$\int |uv|\, dx \le \frac{\varepsilon}{p} \|u\|_{L^p}^p + \frac{1}{\varepsilon^{p'-1} p'} \|v\|_{L^{p'}}^{p'}$$

and we set $\varepsilon = \|v\|_{L^{p'}} / \|u\|_{L^p}^{p-1}$. □

LEMMA 3.5 (MINKOWSKI). $\|u + v\|_{L^p} \le \|u\|_{L^p} + \|v\|_{L^p}$.

PROOF. For $1 < p < \infty$,

$$\int |u+v|^{p-1}|u+v| \le \left(\int |u+v|^{(p-1)p'}\right)^{1-\frac{1}{p}} \left(\|u\|_{L^p} + \|v\|_{L^p}\right) .$$

The cases $p = 1, \infty$ are easy. □

Here is an *interpolation* inequality.

LEMMA 3.6. *Let* $1 \le p \le r \le q$, $\frac{1}{r} = \frac{\alpha}{p} + \frac{1-\alpha}{q}$ *and* $0 \le \alpha \le 1$. *Then*

$$\|u\|_{L^r} \le \|u\|_{L^p}^{\alpha} \|u\|_{L^q}^{1-\alpha} .$$

PROOF.

$$\begin{aligned}\|u\|_{L^r}^r &= \int |u|^{\alpha r + (1-\alpha) r} \\ &\le \left(\int |u|^{\alpha r \frac{p}{\alpha r}}\right)^{\frac{\alpha r}{p}} \left(\int |u|^{(1-\alpha) r \frac{q}{(1-\alpha) r}}\right)^{\frac{(1-\alpha)r}{q}} \\ &= \|u\|_{L^p}^{\alpha r} \|u\|_{L^q}^{(1-\alpha) r}\end{aligned}$$

since $\frac{\alpha r}{p} + \frac{(1-\alpha)r}{q} = 1$. □

The next inequality is an *imbedding* result.

LEMMA 3.7. *If* $\mu(G) < \infty$, $1 \le p \le q \le \infty$ *then* $L^q \subset L^p$ *and*

$$\|u\|_{L^p} \le \mu(G)^{\frac{1}{p} - \frac{1}{q}} \|u\|_{L^q} .$$

PROOF.

$$\|u\|_{L^p}^p = \int_G |u(x)|^p \cdot 1\, dx \le \|u\|_{L^q}^{p/q} \left(\int_G 1\right)^{1-\frac{p}{q}} .$$ □

The next result describes the intersection of the L^p spaces.

LEMMA 3.8. *If* $u \in L^p$, *for* $p \ge 1$, *then* $\lim_{q\to\infty} \|u\|_{L^q} = \|u\|_{L^\infty}$.

PROOF. If $0 < m < \|u\|_{L^\infty}$ and $A_m = \{x : |u(x)| > m\}$, then $\|u\|_{L^q} \ge m\, \mu(A_m)^{1/q}$ and $0 < \mu(A_m) < \infty$. Therefore $\underline{\lim}_{q\to\infty} \|u\|_{L^q} \ge m$, hence $\underline{\lim}_{q\to\infty} \|u\|_{L^q} \ge \|u\|_{L^\infty}$. If $\|u\|_{L^\infty} < \infty$ then $|u(x)|^q \le |u(x)|^p \|u\|_{L^\infty}^{q-p}$ so $\|u\|_{L^q} \le \|u\|_p^{p/q} \|u\|_{L^\infty}^{1-(p/q)}$ so $\overline{\lim}_{q\to\infty} \|u\|_{L^q} \le \|u\|_{L^\infty}$. □

PROPOSITION 3.1. *$L^p(G)$ is a Banach space,* $1 \le p \le \infty$.

PROOF. For $p = +\infty$, $\{u_n\}$ Cauchy in $L^\infty(G)$ implies there is a set A of zero measure such that for $x \in G \sim A$, $m, n \ge 1$

$$|u_n(x) - u_m(x)| \le \|u_n - u_m\|_{L^\infty} , \qquad |u_n(x)| \le \sup_m \|u_m\|_{L^\infty} .$$

Thus $u_n \to u$ uniformly on $G \sim A$; set $u(x) = 0$ for $x \in A$. Then $\|u\|_{L^\infty} \le \sup \|u_n\|_{L^\infty}$ and $\|u_n - u\|_{L^\infty} \to 0$ as $n \to \infty$.

Suppose $1 \le p < \infty$ and $\{u_n\}$ is Cauchy in L^p. There is a subsequence $\{u_{n_j}\}$ for which $\|u_{n_{j+1}} - u_{n_j}\|_{L^p} < \frac{1}{2^j}$. Set $v_m(x) \equiv \sum_{j=1}^m |u_{n_{j+1}}(x) - u_{n_j}(x)|$, so $\|v_m\|_{L^p} \le \sum_{j=1}^m \frac{1}{2^j} < 1$, and $v(x) = \lim_{m\to\infty} v_m(x)$ in $\mathbb{R}_\infty$. From Fatou's lemma,

$$\int_G |v(x)|^p\,dx \le \varliminf_{m\to\infty} \int_G |v_m(x)|^p\,dx \le 1 \quad \text{so} \quad v(x) < \infty \quad \text{a.e.}$$

and the series

$$\Big\{u_{n_1}(x) + \sum_{j=1}^{\infty}\big(u_{n_{j+1}}(x) - u_{n_j}(x)\big)\Big\}$$

is (absolutely) convergent to (definition) $u(x) \equiv \lim_{j\to\infty} u_{n_k}(x)$, a.e. on G. Set $u(x) = 0$ otherwise. Let $\varepsilon > 0$ and choose N: $m, n \ge N$ implies $\|u_n - u_m\|_{L^p} < \varepsilon$. Fatou's lemma shows for $n \ge N$ that

$$\begin{aligned}\int_G |u(x) - u_n(x)|^p\,dx &= \int_G \lim_{j\to\infty} |u_{n_j}(x) - u_n(x)|^p \\ &\le \varliminf_{j\to\infty} \int_G |u_{n_j}(x) - u_n(x)|^p\,dx \le \varepsilon^p\end{aligned}$$

so $u - u_n \in L^p$, (hence, $u \in L^p$,) and $u_n \to u$ in L^p. □

The proof of Proposition 3.1 also yields the following useful result.

COROLLARY 3.1. *Every Cauchy sequence in $L^p(G)$ has a subsequence which converges pointwise* a.e. *on G.*

We consider the approximation of functions in $L^p(G)$ by those which are smooth and vanish near the boundary, ∂G. For any real-valued function φ on G we define the *support* of φ to be the closure in G of the set $\{x \in G : \varphi(x) \ne 0\}$, and we denote it by $\operatorname{supp}(\varphi)$. For any pair of sets we denote by $K \subset\subset G$ that K is compact, G is open, and $K \subset G$. This implies there is a strictly positive distance between K and ∂G. Finally we define $C_0^\infty(G) = \{\varphi : G \to \mathbb{R} \mid \varphi$ is infinitely differentiable and $\operatorname{supp}(\varphi) \subset\subset G\}$. This linear space is frequently called the space of *test functions* because it is dense in many function spaces over G and it is easy to justify computations on such functions.

In order to construct smooth approximations of general functions, we specify a *regularizing sequence*: for each $\varepsilon > 0$, let $\varphi_\varepsilon \in C_0^\infty(\mathbb{R}^n)$ be given with

$$\varphi_\varepsilon \ge 0\ , \quad \operatorname{supp}\varphi_\varepsilon \subset \{x \in \mathbb{R}^n : |x| \le \varepsilon\}\ , \quad \int \varphi_\varepsilon = 1\ .$$

Such functions can be constructed in the form $\varphi_\varepsilon(x) = \varepsilon^{-n}\varphi_1(x/\varepsilon)$ from a single φ_1 as given above.

Let $f \in L^1_{\text{loc}}(G)$ and define the *mollifier* $\mathcal{M}_\varepsilon$ corresponding to $\{\varphi_\varepsilon\}$ by

$$(\mathcal{M}_\varepsilon f)(x) = f_\varepsilon(x) = \int_{\mathbb{R}^n} f(x-y)\varphi_\varepsilon(y)\,dy = \int_{\mathbb{R}^n} f(y)\varphi_\varepsilon(x-y)\,dy\ ,$$

where we have extended f to all of $\mathbb{R}^n$ as zero. From the first integral representation it follows that

$$\operatorname{supp} f_\varepsilon \subset \operatorname{supp} f + \{y : |y| \le \varepsilon\}\ .$$

That is, $f_\varepsilon(x) \ne 0$ only if $x \in \operatorname{supp}(f) + \{y : |y| \le \varepsilon\}$, $\operatorname{supp} f$ is closed and the ε-ball is compact, so their set sum is closed, hence, it contains $\operatorname{supp}(f_\varepsilon)$. From the second integral representation and Leibnitz' rule, it follows that $f_\varepsilon \in C^\infty(\mathbb{R}^n)$. Thus we have $\mathcal{M}_\varepsilon : L^1_{\text{loc}} \to C_0^\infty$ and the support of $\mathcal{M}_\varepsilon f$ is at most ε larger than that of f.

PROPOSITION 3.2. (a) *If $f \in C_0(G)$, then $f_\varepsilon \to f$ uniformly.* (b) *If $f \in L^p(G)$, $1 \le p < \infty$, then $\|f_\varepsilon\|_{L^p} \le \|f\|_{L^p}$ and $f_\varepsilon \to f$ in $L^p(G)$.*

PROOF. The proof of (a) follows from the calculation

$$\begin{aligned}|f_\varepsilon(x) - f(x)| &\le \int_{\mathbb{R}^n} |f(x-y) - f(x)|\varphi_\varepsilon(y)\,dy \\ &\le \sup\{|f(x-y) - f(x)| : y \in f \ , \ |y| \le \varepsilon\} \ .\end{aligned}$$

For (b) the case $p = 1$ follows from

$$\|f_\varepsilon\|_{L^1} \le \iint |f(x-y)|\varphi_\varepsilon(y)\,dx\,dy = \|f\|_{L^1} \ .$$

To prove this for $p > 1$, let $\psi \in C_0(G)$ and estimate

$$\begin{aligned}\left|\int_{\mathbb{R}^n} f_\varepsilon(x)\psi(x)\,dx\right| &\le \iint |f(x-y)\psi(x)|\,dx\ \varphi_\varepsilon(y)\,dy \\ &\le \int \|f\|_{L^p}\|\psi\|_{L^{p'}}\varphi_\varepsilon(y)\,dy\end{aligned}$$

and take the supremum over those ψ with $L^{p'}$-norm at most one. For the convergence proof, first let $\eta > 0$, then choose $g \in C_0(G)$ with $\|f - g\|_{L^p} < \eta$. Then we have $\|f_\varepsilon - g_\varepsilon\|_{L^p} < \eta$ and, hence, $\|f - f_\varepsilon\|_{L^p} < 2\eta + \|g - g_\varepsilon\|_{L^p}$. The last term goes to zero with ε, by part (a), so we are done. □

We consider the dual of $L^p(G)$. For each $v \in L^{p'}(G)$ define $\mathcal{R}_v(u) \equiv \int_G vu$, $u \in L^p(G)$. Then $\mathcal{R}_v \in (L^p)'$ and $\|\mathcal{R}_v\|_{(L^p)'} \le \|v\|_{L^{p'}}$ follows from Hölder's inequality, Lemma 3.4. We shall show $\mathcal{R} : L^{p'} \to (L^p)'$ is an isometry.

Suppose $1 < p \le \infty$ and set $u(x) \equiv |v(x)|^{p'-1}\operatorname{sgn} v(x)$, $x \in G$, where the *sign function* is defined by $\operatorname{sgn} w = 1$ if $w > 0$ and $\operatorname{sgn} w = -1$ if $w < 0$. Then $|u|^p = |v|^{p(p'-1)} = |v|^{p'}$ so $u \in L^p$ and

$$\mathcal{R}_v(u) = \int |v|^{p'} = \|v\|_{L^{p'}}^{p'(\frac{1}{p'}+\frac{1}{p})} = \|v\|_{L^{p'}}\|u\|_{L^p} \ .$$

Therefore $\|\mathcal{R}_v\|_{(L^p)'} = \|v\|_{L^{p'}}$. Suppose $p = 1$ and $\|v\|_{L^\infty} > 0$. Let $0 < \varepsilon < \|v\|_{L^\infty}$ and choose $A \subset G$ with $0 < \mu(A) < \infty$, $|v(x)| \ge \|v\|_{L^\infty} - \varepsilon$ for all $x \in A$. Set $u(x) = \operatorname{sgn} v(x)$, $x \in A$ and $u(x) = 0$ on $G \sim A$. Then $u \in L^1(G)$ and

$$\mathcal{R}_v(u) = \int_A |v(x)|\,dx \ge \mu(A)(\|v\|_{L^\infty} - \varepsilon) = \|u\|_{L^1}(\|v\|_{L^\infty} - \varepsilon) \ ,$$

so $\|\mathcal{R}_v\|_{(L^1)'} = \|v\|_{L^\infty}$.

This shows the first part of the following; the second is proved using the Radon-Nikodým Theorem.

THEOREM 3.1 (RIESZ). *The map $v \mapsto \mathcal{R}_v : L^{p'} \to (L^p)'$ is an isometry for $1 \le p \le \infty$; it is a bijection if $1 \le p < \infty$.*

COROLLARY 3.2. *L^p is reflexive if $1 < p < \infty$.*

PROOF. Denote by $\mathcal{R}_{p'} : L^p \to (L^{p'})'$ the Riesz map $\mathcal{R}_{p'}(u)v = \int_G uv\,dx$, and similarly by $\mathcal{R}_p : L^{p'} \to (L^p)'$ the map $\mathcal{R}_p v(u) = \int_G vu\,dx$. The dual of $\mathcal{R}_p$ is $\mathcal{R}_p' : (L^p)'' \to (L^{p'})'$ given by $\mathcal{R}_p'\varphi(v) = \varphi(\mathcal{R}_p v)$ for $\varphi \in (L^p)''$ and $v \in L^{p'}$.

Set $X = L^p(G)$ so $\varphi : X \to X''$ is given by $\varphi(u)(\mathcal{R}_p v) \overset{(\varphi)}{\equiv} \mathcal{R}_p v(u) = \mathcal{R}_{p'}u(v)$ $\forall\, u \in L^p$, $v \in L^{p'}$, hence, $\mathcal{R}_p'(\varphi(w)) = \mathcal{R}_{p'}(u)$ and $\varphi = (\mathcal{R}_p')^{-1} \circ \mathcal{R}_{p'}$ is onto.

We shall *identify* $L^p \cong (L^{p'})'$, $1 < p \leq \infty$, and $(L^p)'' \cong L^p$ for $1 < p < \infty$ as above. Thus certain functions on G are identified with linear functionals on $L^p(G)$ as prescribed above by $\mathcal{R}_p$ and φ, and these are *consistent*.

THEOREM 3.2 (NEMYTSKII). *Let $f : G \times \mathbb{R} \to \mathbb{R}$ satisfy the Caratheodory conditions*

(i) *$f(\cdot, r) : G \to \mathbb{R}$ is measurable for each $r \in \mathbb{R}$, and*

(ii) *$f(x, \cdot) : \mathbb{R} \to \mathbb{R}$ is continuous for* a.e. $x \in G$.

Let $1 \leq p, q < \infty$, $k \in L^q(G)$, and assume

$$|f(x,\xi)| \leq c|\xi|^{p/q} + k(x)\ , \quad \text{a.e. } x \in G\ , \quad \xi \in \mathbb{R}\ .$$

Then the operator defined by

$$F(u)(x) = f\big(x, u(x)\big)\ , \quad \text{a.e. } x \in G\ , \quad u \in L^p(G)$$

gives $F : L^p(G) \to L^q(G)$ which is bounded and strongly continuous.

PROOF. We may assume $f(x, r)$ is continuous in r at every $x \in G$. If $u(x) = \lim u_n(x)$ where each u_n is a simple function, then $(Fu)(x) = \lim(Fu_n)(x)$ for $x \in G$. Each $F(u_n)$ is the sum of a countable set of functions $f(x, c_j)\chi_j(x)$ where χ_j is the characteristic function of a measurable set, so each $F(u_n)$ and hence $F(u)$ is measurable. Also $\|F(u)\|_{L^q} \leq c\|u^{p/q}\|_{L^q} + \|k\|_{L^q} = c\|u\|_{L^p}^{p/q} + \|k\|_{L^q}$ so F is bounded from L^p to L^q.

To show F is continuous, let $u_n \to u$ in L^p and assume for the moment that $\{u_n\}$ has the properties of the subsequence $\{u_{n_j}\}$ in the proof of completeness of L^p above. Since $f(x, u)$ is continuous in u, it follows $\lim_{n\to\infty} F(u_n)(x) = F(u)(x)$ for a.e. $x \in G$. Also we have

$$\begin{aligned}|F(u_n)(x)| \leq c|u_n(x)|^{p/q} + k(x) &\leq c\Big(\sum_{n=1}^{\infty} |u_{n+1}(x) - u_n(x)| + |u_1(x)|\Big)^{p/q} + k(x) \\ &\equiv g(x) \text{ and } g \in L^q(G)\ .\end{aligned}$$

From Lebesgue's Theorem we obtain

$$\|F(u_n) - F(u)\|_{L^q}^q = \int \|F(u_n)(x) - F(u)(x)|^q\,dx \to 0$$

so $F(u_n) \to F(u)$ in L^q. Thus *every* subsequence (of the original sequence) has a subsequence for which $F(u_n)$ converges to $F(u)$ in L^q. □

COROLLARY 3.3. *If f satisfies the Caratheodory conditions and $|f(x,\xi)| \leq c|\xi|^{p-1} + k(x)$ where $k \in L^{p'}$, $1 < p < \infty$ then $F : L^p \to L^{p'} \cong (L^p)'$ is bounded and continuous. If also $f(x, \cdot) : \mathbb{R} \to \mathbb{R}$ is non-decreasing, then $F : L^p \to (L^p)'$ is monotone.*

PROOF. Choose $q = p'$, $\frac{1}{p} + \frac{1}{p'} = 1$ so $p/p' = p - 1$. Also

$$\langle F(u) - F(v),\ u - v\rangle = \int_G \Big(f\big(x, u(x)\big) - f\big(x, v(x)\big)\Big)\big(u(x) - v(x)\big)\, dx \geq 0 \ ,$$

since $(f(x,\xi) - f(x,\eta))(\xi - \eta) \geq 0$ for all $\xi, \eta \in \mathbb{R}$, $x \in G$. □

COROLLARY 3.4. *Let $N \geq 1$ and $f : G \times \mathbb{R}^N \to \mathbb{R}$ satisfy Caratheodory conditions as above: $f(x,\xi)$ is continuous in $\xi \in \mathbb{R}^N$ and measurable in $x \in G$. Let $1 \leq p, q < \infty$, $k \in L^q(G)$ and assume*

$$|f(x,\xi)| \leq c \sum_{j=1}^{N} |\xi_j|^{p/q} + k(x) \ , \ \text{a.e.} \ \ x \in G \ , \quad \xi \in \mathbb{R}^N \ .$$

Then $F(u)(x) \equiv f(x, u_1(x), u_2(x), \dots, u_N(x))$, $x \in G$, defines $F : [L^p(G)]^N \to L^q(G)$ continuous and bounded.

Operators constructed between L^p spaces as above by substitution are referred to as *Nemytskii operators*. As we have seen they are frequently continuous in the strong topology. Things are different in the weak topology.

EXAMPLE 3.A. The sequence $\varphi_n^{(x)} \equiv \sqrt{\frac{2}{\pi}} \sin(nx)$, $n \geq 1$, is an ortho-normal basis for $L^2(0,\pi)$; specifically, $\varphi_n \rightharpoonup 0$ in $L^2(0,\pi)$. Similarly, $\cos(nx) \rightharpoonup 0$ so from $\sin^2(x) = \frac{1}{2}(1 - \cos 2x)$ it follows $(\varphi_n)^2 \rightharpoonup \frac{1}{\pi}$ in $L^2(0,\pi)$. Thus, $u \mapsto u^2$ is not weakly continuous on $L^2(0,\pi)$. Furthermore, $\{\varphi_n\}$ is bounded in $L^p(0,\pi)$, $1 < p < \infty$, $C_0(0,\pi)$ is dense in $L^{p'}(0,\pi) \cong (L^p)'$ and $\int_G \varphi \varphi_n \to 0$ for all $\varphi \in C_0(0,\pi)$ as above, so $\varphi_n \rightharpoonup 0$ in every $L^p(0,\pi)$, $1 < p < \infty$, and $(\varphi_n)^2 \rightharpoonup \frac{1}{\pi}$ in every $L^q(0,\pi)$, $1 < q < \infty$. Hence $u \mapsto u^2$ is *not* weakly continuous $L^p \to L^q$ for *any* $1 < p, q < \infty$!

We shall give a sufficient condition for a subset of $L^p(\Omega)$ to be compact. Since this is a metric space it follows that a set is compact, i.e., sequentially compact, if and only if it is complete and *totally bounded*. A subset F of the space X is called an *ε-net* if F is finite and $X = \cup\{B_\varepsilon(a) : a \in F\}$. That is, F is finite and the set of all balls of radius centered at points of F is a covering of the space X. The space X is *totally bounded* if there is an ε-net for each $\varepsilon > 0$.

PROPOSITION 3.3. *The space X is compact if and only if it is complete and totally bounded.*

PROOF. The necessity is straight forward, and we only verify the sufficiency of these criteria. Since X is complete, to show compactness it suffices to show every sequence has a Cauchy subsequence. If we have a sequence $\{x_{11}, x_{12}, x_{13}, \dots\}$ given, then by total-boundedness there is a subsequence $\{x_{21}, x_{22}, x_{23}, \dots\}$, all of whose points lie in one sphere of radius $1/2$ (since there is a $1/2$-net which contains the original sequence). Likewise there is a subsequence $\{x_{31}, x_{32}, x_{33}, \dots\}$, all of whose points lie in one sphere of radius $1/3$. Continue selection subsequences in this manner and check that the diagonal sequence $\{x_{11}, x_{22}, x_{33}, \dots\}$ is Cauchy.

COROLLARY 3.5. *In a complete space, a closed set is compact if and only if it is totally bounded.*

These notions are particularly useful for establishing criteria for subsets of function spaces to be compact. An example which we use below is the following.

THEOREM 3.3 (ASCOLI). *Let K be a compact set (in a metric space) and let $\mathcal{F}$ be a bounded subset of functions in $C(K,\mathbb{R})$. If $\mathcal{F}$ is uniformly equicontinuous, i.e., for each $\varepsilon > 0$ there is a $\delta > 0$ such that $|f(x_1) - f(x_2)| < \varepsilon$ if $f \in \mathcal{F}$ and $d(x_1, x_2) < \delta$, then $\mathcal{F}$ is precompact in $C(K,\mathbb{R})$. That is, its closure $\overline{\mathcal{F}}$ is compact.*

THEOREM 3.4 (FRECHET-KOLMOGOROV). *Let G be open in $\mathbb{R}^n$ and $K \subset\subset G$. Let $\mathcal{F}$ be a bounded subset of $L^p(G)$, $1 \le p < \infty$. Suppose that for every $\varepsilon > 0$ there is a δ, $0 < \delta < \operatorname{dist}(K, \partial G)$ such that*

$$\left(\int_K |f(x+h) - f(x)|^p\,dx\right)^{1/p} < \varepsilon\ ,\ \textit{for all } h \in \mathbb{R}^n\ ,\ |h| < \delta\ ,\ f \in \mathcal{F}\ .$$

Then $\mathcal{F}|_K$ is precompact in $L^p(K)$.

PROOF. With no loss of generality, we assume G is bounded. Also, we let $\overline{f}$ denote the zero-extension to $\mathbb{R}^n$ of each function f on G. It follows that $\{\overline{f} : f \in \mathcal{F}\} \equiv \overline{\mathcal{F}}$ is bounded in $L^p(\mathbb{R}^n)$ and in $L^1(\mathbb{R}^n)$.

Fix $n > \frac{1}{\delta}$; then we have with $M_n = \mathcal{M}_{1/n}$

$$|M_n\overline{f}(x) = \overline{f}(x)| \le \int_{\mathbb{R}^n} |\overline{f}(x-y) - \overline{f}(x)|\varphi_n(y)\,dy\ ,$$

and writing $\varphi_n = \varphi_n^{1/p} \cdot \varphi_n^{1/p'}$ applying Hölder's inequality, Lemma 3.4, we obtain

$$|M_n\overline{f}(x) - \overline{f}(x)|^p \le \int_{|y|\le 1/n} |\overline{f}(x-y) - \overline{f}(x)|^p\varphi_n(y)\,dy\ .$$

Thus we have

$$\int_K |M_n\overline{f}(x) - \overline{f}(x)|^p\,dx \le \int_{|y|\le 1/n}\int_G |\overline{f}(x-y) - \overline{f}(x)|^p\,dx\ \varphi_n(y)\,dy < \varepsilon^p\ .$$

That is,

$$\|M_u\overline{f} - \overline{f}\|_{L^p(K)} < \varepsilon\ , \qquad n \ge \frac{1}{\delta}\ ,\ \overline{f} \in \overline{\mathcal{F}}\ .$$

Next consider $\mathcal{F}_n \equiv \{M_n\overline{f}|_K : \overline{f} \in \overline{\mathcal{F}}\}$ for a fixed n. Since φ_n is bounded,

$$\|M_n\overline{f}\|_{L^\infty(\mathbb{R}^n)} \le \|\varphi_n\|_{L^\infty(\mathbb{R}^n)}\|\overline{f}\|_{L^1(\mathbb{R}^n)}\ , \qquad f \in \mathcal{F}\ ,$$

and we have for $x_1, x_2 \in K$

$$|M_n\overline{f}(x_1) - M_n\overline{f}(x_2)| \le |x_1 - x_2|\ \|\varphi_n\|_{W^{1,\infty}}\|\overline{f}\|_{L^1}\ ,$$

so $\mathcal{F}_n$ is uniformly bounded and equicontinuous. From Ascoli's Theorem 3.3 it follows that $\mathcal{F}_n$ is precompact in $C(K,\mathbb{R})$, hence, in $L^p(K)$.

Finally, let $\varepsilon > 0$ be given and choose $n > 1/\delta$. Since $\mathcal{F}_n$ is precompact in $L^p(K)$ there is a finite collection of balls of radius ε which cover $\mathcal{F}_n$. The set of corresponding balls of radius 2ε then cover $\mathcal{F}|_K$. □

COROLLARY 3.6. *Let G be open in $\mathbb{R}^n$ and $\mathcal{F}$ a bounded subset of $L^p(G)$. Suppose for $\varepsilon > 0$ and compact $K \subset\subset G$ there is a $\delta > 0$, $\delta < \operatorname{dist}(K, \partial G)$, such that*

$$\int_K |f(x+y) - f(x)|^p\,dx < \varepsilon^p\ , \qquad |h| < \delta\ ,\ f \in \mathcal{F}\ ,$$

and for $\varepsilon > 0$ there is a compact $K \subset\subset G$, such that $\|f\|_{L^p(G\sim K)} < \varepsilon$, $f \in \mathcal{F}$. Then $\mathcal{F}$ is precompact in $L^p(G)$.

PROOF. Given $\varepsilon > 0$, fix K as above so that $\|f\|_{L^p(G\sim K)} < \varepsilon$, $f \in \mathcal{F}$. Thus $\mathcal{F}|_K$ is precompact in $L^p(K)$ and can be covered by a finite collection $B_\varepsilon(f_i)$, $f_i \in L^p(K)$, $1 \le i \le N$. Extend these as zero to get $\overline{f}_i \in L^p(G)$. Then it follows that $\mathcal{F} \subset \bigcup_{i=1}^N B_{2\varepsilon}(\overline{f}_i)$. □

We close with a remark on demicontinuity.

LEMMA 3.9. *If $u_n \rightharpoonup u$ in $L^p(G)$, $1 \le p \le +\infty$ and $u_n(x) \to v(x)$ a.e. $x \in G$ then $u = v$.*

PROOF. Let $G_0 \subset G$ with $\mu(G_0) < \infty$. Then Egorov's Theorem implies that for each $\varepsilon > 0$ there is a measurable $A \subset G_0$ with $\mu(G_0 \sim A) < \varepsilon$ and $u_n \to v$ uniformly on A. For each measurable $B \subset A$, $\int_B u_n \to \int_B v$ by uniform convergence. Also $\chi_B \in L^{p'}$ since $\mu(B) < \infty$ so by weak convergence $\int_B u_n = \int_B \chi_B \cdot u_n \to \int_G \chi_B u = \int_B u$, and so $\int_B (u - v) = 0$ for every $B \subset A$. It follows that $u(x) = v(x)$ a.e. in G.□

PROPOSITION 3.4. *If $\{u_n\}$ is bounded in $L^p(G)$, $1 < p < \infty$, and if $u_n(x) \to v(x)$ a.e. $x \in G$, then $u_n \rightharpoonup v$.*

PROOF. Otherwise there is an $\varepsilon > 0$, $f \in (L^p)'$ and subsequence $\{u_{n_j}\}$ for which $|f(u_{n_j} - v)| \ge \varepsilon$, $j \ge 1$. Pick a further subsequence $u_{n'_j} \rightharpoonup u$ in L^p and note from above $u = v$, a contradiction. □

The point of Proposition 3.4 is that we need only to have the function $F(u) = f(\cdot, u(\cdot))$ to be bounded and to have $f(x, \cdot)$ continuous in order to get $F : L^p \to L^q$ to be demicontinuous.

II.4. Sobolev Spaces

We shall introduce certain spaces of functions which with their derivatives belong to $L^p(G)$ and describe the sense in which they have values on ∂G. First we construct a space which is very large and contains all its derivatives. Recall that $C_0^\infty(G)$ is dense in every $L^p(G)$, $1 \le p < \infty$. We call $C_0^\infty(G)$ the space of *test functions* on G.

DEFINITION. $\mathcal{D}^* = C_0^\infty(G)^*$, the algebraic dual of $C_0^\infty(G)$, is the linear space consisting of all linear functionals on $C_0^\infty(G)$. These linear functionals are called *generalized functions* on G.

EXAMPLE 4.A. For each $f \in L^1_{\text{loc}}(G)$ (that is, $f|_K \in L^1(K)$ for each $K \subset\subset G$) we define $\tilde{f} \in \mathcal{D}^*$ by $\tilde{f}(\varphi) = \int_G f\varphi\,dx$, $\varphi \in C_0^\infty(G)$. Then $f \mapsto \tilde{f}$ is an injection by which we identify $L^1_{\text{loc}}(G) \subset \mathcal{D}^*$.

EXAMPLE 4.B. For each $f \in L^{p'}$, $1 \le p < \infty$, we have given $\mathcal{R}f \in (L^p)'$ by $\mathcal{R}f(v) = \int_G fv\,dx$, $v \in L^p(G)$ and have agreed to identify $f \cong \mathcal{R}f$. If we denote by $i : C_0^\infty(G) \to L^p(G)$ the identity with dense range, and by $i^* : (L^p)' \to \mathcal{D}^*$ the algebraic dual which is one-to-one (it is just restriction to $C_0^\infty(G)$), then we find $i^*(\mathcal{R}f) = \mathcal{R}f \circ i = \tilde{f}$, so these identifications are *compatible*.

$$\begin{array}{ccc} (L^p)' & \xrightarrow{i^*} & \mathcal{D}^* \\ \uparrow\mathcal{R} & & \uparrow \\ L^{p'} & \longrightarrow & L^1_{\text{loc}} \end{array}$$

EXAMPLE 4.C. Let $x_0 \in G$ and define $\delta_{x_0} \in \mathcal{D}^*$ by

$$\delta_{x_0}(\varphi) = \varphi(x_0) , \qquad \varphi \in C_0^\infty(G) .$$

Then $\delta_{x_0} \notin L^1_{\text{loc}}$; similarly we can construct $\delta_S(\varphi) = \int_S \varphi \, ds$ for any lower-dimensional manifold S in G.

Suppose $f \in L^1_{\text{loc}}(G)$ and for a.e. $(x_1, \ldots, x_{n-1})$ that $x \mapsto f(x_1, \ldots, x_{n-1}, x)$ is absolutely continuous with derivative $D_n f \in L^1_{\text{loc}}$. For each $\varphi \in C_0^\infty(G)$

$$\widetilde{D_n f}(\varphi) \equiv \int_G (D_n f)\varphi = -\int_G f(D_n\varphi) \equiv -\tilde{f}(D_n\varphi) = -D_n^*\tilde{f}(\varphi)$$

where $D_n^* : \mathcal{D}^* \to \mathcal{D}^*$ is the adjoint of $D_n : C_0^\infty \to C_0^\infty$. Thus $\widetilde{D_n f} = -D_n^*\tilde{f}$ for all such f and this *forces* the following.

DEFINITION. The partial derivative of the generalized function $T \in \mathcal{D}^*$ in the j^{th}-coordinate direction is the generalized function $\partial_j T \equiv -D_j^* T$; that is,

$$\partial_j T(\varphi) = -T(D_j\varphi) \quad \text{for} \quad \varphi \in C_0^\infty(G) .$$

When $T = \tilde{f}$ as above, we call $\partial_j T = \partial_j \tilde{f}$ the *generalized derivative* of f. Of course we rigged this so that $\partial_j \tilde{f} = \widetilde{D_j f}$ so this extended notion is compatible with the usual notion of derivative. Moreover, we can repeatedly differentiate every element of $\mathcal{D}^*$, hence, every $f \in L^1_{\text{loc}}(G)$.

LEMMA 4.1. *Let $D = \{u \in L^p(G) : \partial_j u \in L^q(G)\}$, $1 \le p, q \le \infty$, and let $\partial_j : D \to L^q(G)$ be the corresponding linear function. Then ∂_j has closed graph in $L^p \times L^q$.*

PROOF. If $u_n \to u$ in L^p and $\partial_j u_n \to v$ in L^q, then $\partial_j u_n(\varphi) = -u_n(D_j\varphi) \to -u(D_j\varphi) = \partial_j u(\varphi)$ as $n \to \infty$ so $\partial_j u = v$. □

DEFINITION. Let G be a domain in $\mathbb{R}^n$, $1 \le p \le \infty$, $0 \le m =$ integer. $W^{m,p}(G)$ is the linear space of all $u \in L^p(G)$ for which $\partial^\alpha u \in L^p(G)$ for all multi-indices $\alpha = (\alpha_1, \ldots, \alpha_n)$ of non-negative integers with order $|\alpha| = \sum_{j=1}^n \alpha_j \le m$, where $\partial^\alpha = \partial_1^{\alpha_1}\partial_2^{\alpha_2} \ldots \partial_n^{\alpha_n}$. With the corresponding norms given by

$$\|u\|_{m,p} = \left(\sum_{|\alpha| \le m} \|\partial^\alpha u\|^p_{L^p(G)} \right)^{1/p} , \qquad 1 \le p < \infty \text{ and}$$

$$\|u\|_{m,\infty} = \max_{|\alpha| \le m} \|\partial^\alpha u\|_{L^\infty(G)} ,$$

each is the *Sobolev space* of order m in $L^p(G)$.

We digress in order to connect this notion to related ones that frequently occur in the literature. For any pair $u, v \in L^1_{\text{loc}}(G)$, $v = D^\alpha u$ in the *weak sense* and v is called the α^{th} *weak derivative* of u if

$$\int_G u D^\alpha \varphi = (-1)^{|\alpha|} \int_G v\varphi , \qquad \varphi \in C_0^\infty(G) .$$

This is clearly equivalent to $\widetilde{v} = \partial^\alpha \widetilde{u}$, so weak and generalized derivatives are equivalent notions.

Certain properties are obvious. The multiple derivative $\partial^\alpha u$ is independent of the order of differentiation, since this is true of $D^\alpha\varphi$. Similarly it follows that $\partial^\alpha\partial^\beta = \partial^{\alpha+\beta}$. Finally, let v be the α^{th} weak derivative of u in $L^1_{\text{loc}}(G)$ as above and consider the mollifier $\mathcal{M}_\varepsilon$ constructed in Section 3. From the definition of derivative we obtain

$$D^\alpha \mathcal{M}_\varepsilon u(x) = \int_G u(y) D_x^\alpha \rho_\varepsilon(x-y)\,dy =$$

$$(-1)^{|\alpha|}\int_G u(y) D_y^\alpha \rho_\varepsilon(x-y)\,dy = \mathcal{M}_\varepsilon \partial^\alpha u(x)\ , \quad x \in G\ , \quad 0<\varepsilon<\operatorname{dist}(x,\partial G)\ .$$

That is, $\mathcal{M}_\varepsilon\partial^\alpha = D^\alpha\mathcal{M}_\varepsilon$ in the interior of G. Thus, if $v = \partial^\alpha u$ in $L^1_{\text{loc}}(G)$, then by setting $u_n \equiv \mathcal{M}_{\frac{1}{n}} u$ we obtain (with $u = 0$ outside of G) a sequence $u_n \in C^\infty(G)$ such that for any compact $K \subset G$

$$u_n \to u \text{ and } D^\alpha u_n \to v \text{ in } L^1(K)$$

as $n \to \infty$. In this case, v is called the *strong derivative* of u. From Lemma 4.1 it follows that every strong derivative is also a weak or generalized derivative, so all three notions are equivalent.

Finally, we shall show that smooth functions are dense in $W^{m,p}(G)$, $1 \le p < \infty$, so this space is equivalently obtained as the completion of $C^\infty(G) \cap W^{m,p}(G)$.

THEOREM 4.1 (MEYERS-SERRIN). *If $1 \le p < \infty$, then $C^\infty(G) \cap W^{m,p}(G)$ is dense in $W^{m,p}(G)$.*

PROOF. For each integer $n \ge 1$ set

$$G_n = \left\{x \in G : |x| < n \text{ and } \operatorname{dist}(x,\partial G) > \frac{1}{n}\right\}$$

and let $G_0 = G_{-1} = \varphi$. Then the sequence $\Omega_n = G_{n+1} \sim \overline{G}_{n-1}$ is an open cover of G. For each $n \ge 1$ let $\beta_j \in C_0^\infty(\Omega_j)$ with $\beta_j \ge 0$ and $\sum_{j=1}^\infty \beta_j(x) = 1$ for $x \in G$. Such a collection $\{\beta_j\}$ is called a *partition-of-unity* subordinate to the open cover $\{\Omega_j\}$. Let $u \in W^{m,p}(G)$ and $\varepsilon > 0$. Since $\operatorname{supp}(\beta_n u) \subset \Omega_n$, there is an $a_n > 0$ such that the mollifier $M_n \equiv \mathcal{M}_{a_n}$ satisfies

$$\operatorname{supp} M_n(\beta_n u) \subset \bigcup_{j=-1}^{1} \Omega_{n+j}$$

and

$$\|M_n(\beta_n u) - \beta_n u\|_{m,p} < \frac{\varepsilon}{2^n}\ .$$

Define $w = \sum_{n=1}^\infty M_n(\beta_n u)$ and note that for $x \in \Omega_n$ we have

$$w(x) = \sum_{j=-1}^{1} M_{n+j}(\beta_{n+j}u)\ .$$

Hence, $w \in C^\infty(\Omega)$ and

$$\|u-w\|_{m,p} \le \sum_{n=1}^\infty \sum_{j=-1}^{1} \|M_{n+j}(\beta_{n+j}u) - \beta_n u\|_{m,p} < 3\varepsilon\ ,$$

so the desired approximation is achieved. □

The domain G has the *segment property* if for each $x \in \partial G$ there is an open neighborhood N_x and vector $\vec{v} \neq \vec{0}$ such that $z \in \overline{G} \cap N_x$, $0 < t < 1$ imply $z + t\vec{v} \in G$. Such a domain lies on one side (locally) of its $(n-1)$-dimensional boundary. By mollifier arguments similar to those above, one may prove the following approximation property:

If $1 \leq p < \infty$ then the subspace of restrictions to G of functions in $C_0^\infty(\mathbb{R}^n)$ is dense in $W^{m,p}(G)$.

Here are some elementary properties of the Sobolev spaces.

PROPOSITION 4.1. *$W^{m,p}(G)$ is a Banach space.*

PROOF. If $\{u_n\}$ is Cauchy in $W^{m,p}$ and $|\alpha| \leq m$ then $\{\partial^\alpha u_n\}$ is Cauchy in L^p, hence, $u_n \to u$ and $\partial^\alpha u_n \to v_\alpha$ in $L^p(G)$. Since ∂^α has a closed graph, $v_\alpha = \partial^\alpha u$.□

By considering the map $u \mapsto (\partial^\alpha u)_{|\alpha| \leq m}$ of $W^{m,p}$ into the product space $L_p^N(G) = L^p(G)^N$, N being the number of multi-indices with order $|\alpha| \leq m$, we find that $W^{m,p}$ is isomorphic to a subspace of a product of L^p spaces.

COROLLARY 4.1. *If $1 \leq p < \infty$, then $W^{m,p}$ is separable and every $T \in (W^{m,p})'$ is given by*

$$T(u) = \sum_{|\alpha| \leq m} \int_G t_\alpha \partial^\alpha u \, dx \ , \qquad u \in W^{m,p}$$

where $\tilde{t} = (t_\alpha)$ belongs to $[L^{p'}]^N$. If $1 < p < \infty$ then $W^{m,p}$ is reflexive.

PROPOSITION 4.2. *A sequence $\{u_n\}$ in $W^{m,p}$, $1 < p < \infty$, is weakly convergent if and only if $\{\partial^\alpha u_n\}$ is weakly convergent in L^p for all α, $|\alpha| \leq m$; in this case we have $u_n \rightharpoonup u$ in $W^{m,p}$ if and only if each $\partial^\alpha u_n \rightharpoonup \partial^\alpha u$ in L^p.*

PROOF. We may assume all sequences are bounded. If $u_n \rightharpoonup u$ in $W^{m,p}$ then for $\varphi \in C_0^\infty(G)$ we have

$$\partial^\alpha u_n(\varphi) = (-1)^{|\alpha|} u_n(\partial^\alpha \varphi) \longrightarrow (-1)^{|\alpha|} u(\partial^\alpha \varphi) = \partial^\alpha u(\varphi)$$

and $\{\partial^\alpha u_n\}$ is bounded in L^p with $C_0^\infty(G)$ dense in $L^{p'}$, so $\partial^\alpha u_n \rightharpoonup \partial^\alpha u$ in L^p.

Conversely, let $u_n \rightharpoonup u$ and $\partial^\alpha u_n \rightharpoonup v_\alpha$ in L^p; each ∂^α is weakly closed, so $\partial^\alpha u_n \rightharpoonup \partial^\alpha u$. The result follows from the representation of $(W^{m,p})'$ given in Corollary 4.1. □

The dual of $W^{m,p}(G)$ is generally more than a space of generalized functions on G. Certainly the restriction to $C_0^\infty(G)$ of an element of $(W^{m,p})'$ belongs to $\mathcal{D}^*$; however, this restriction is *not* injective because $C_0^\infty(G)$ is (in general) not dense in $W^{m,p}(G)$. The problem is that $W^{m,p}$ elements may have non-zero *boundary values*, as we shall see below.

DEFINITION. $W_0^{m,p}(G)$ is the closure of $C_0^\infty(G)$ in $W^{m,p}(G)$.

COROLLARY 4.2. *If $1 \leq p < \infty$, then every $T \in (W_0^{m,p}(G))'$ is given by $T = \sum_{|\alpha| \leq m} \partial^\alpha f_\alpha$ where each $f_\alpha \in L^{p'}(G)$. Conversely, every linear combination of at most m^{th}-order derivatives of $L^{p'}$ functions is a generalized function in $(W_0^{m,p})'$.*

We define the *trace operator*, i.e., restriction to values on the boundary, in the special case of a half-space. Let $G = \mathbb{R}_+^n = \{x = (x', x_n) : x_n > 0\} \subset \mathbb{R}^n$

so $\partial G = \mathbb{R}^{n-1}$. For any $\varphi \in C_0^\infty(\mathbb{R}^n)$ we define $\gamma\varphi = \varphi|_{\partial G}$ and note that the restrictions to G of such functions are dense in $W^{1,p}(G)$, $1 \le p < \infty$. We compute

$$|\varphi(x',0)|^p = -\int_0^\infty D_n|\varphi|^p\,dx_n \le p\int_0^\infty |\varphi|^{p-1}|D_n\varphi| \quad (p-1 = p/p')$$

$$\le p\|D_n\varphi\|_{L^p(\mathbb{R}^+)}\big\|\,|\varphi|^{p/p'}\big\|_{L^{p'}} = p\|D_n\varphi\|_{L^p(\mathbb{R}^+)}\|\varphi\|_{L^p(\mathbb{R}^+)}^{p/p'} ,$$

apply Young's inequality, and integrate over $\mathbb{R}^{n-1}$ to get successively

$$|\varphi(x',0)|^p \le \|D_n\varphi(x',\cdot)\|_{L^p(\mathbb{R}_+)}^p + \|\varphi(x',\cdot)\|_{L^p}^p\ (p-1)$$

$$\|\gamma\varphi\|_{L^p(\mathbb{R}^{n-1})}^p \le \|D_n\varphi\|_{L^p(\mathbb{R}^n_+)}^p + (p-1)\|\varphi\|_{L^p(\mathbb{R}^n_+)}^p .$$

Thus, the trace operator $\gamma : C_0^\infty(\mathbb{R}^n)|_{\mathbb{R}^n_+} \to C_0^\infty(\mathbb{R}^{n-1})$ has a unique extension by continuity to

$$\gamma : W^{1,p}(\mathbb{R}^n_+) \longrightarrow L^p(\mathbb{R}^{n-1}) , \qquad 1 \le p < \infty .$$

Note that the formula

$$u(x',x_n) = \int_0^{x_n} \partial_n u(x',s)\,ds + \gamma u(x') , \qquad x' \in \mathbb{R}^{n-1} ,\ x_n > 0$$

extends to $u \in W^{1,p}(\mathbb{R}^n_+)$ from smooth functions.

Let's characterize the kernel of γ, $\ker(\gamma)$. This is a closed subspace of $W^{1,p}(\mathbb{R}^n_+)$ which contains $C_0^\infty(\mathbb{R}^n_+)$. Suppose $u \in \ker(\gamma)$. We shall show there is a sequence in $C_0^\infty(\mathbb{R}^n_+)$ which converges in $W^{1,p}(\mathbb{R}^n_+)$ to u. Pick a sequence $\theta_j : \mathbb{R} \to \mathbb{R}$ by $\theta_j(t) = 0$, $t \le 1/j$, $\theta_j(t) = 1$ for $t \ge 2/j$, and $\theta_j(t) = j(t-1/j)$ otherwise, and define $u_j(x',x_n) = \theta_j(x_n)u(x)$. Each u_j is in the closure of $C_0^\infty(\mathbb{R}^n_+)$ (by a convolution approximation) so it suffices to show $\lim_{j\to\infty} u_j = u$ in $W^{1,p}(\mathbb{R}^n_+)$.

By Lebesgue's dominated convergence theorem it follows that $u_j \to u$ in L^p and for each k with $1 \le k \le n-1$ that $\partial_k(u_j) = \theta_j(\partial_k u) \to \partial_k u$ in $L^p(\mathbb{R}^n_+)$ as $j \to \infty$. Similarly, $\theta_j\partial_n u \to \partial_n u$ and $\partial_n(\theta_j u) = \theta_j(\partial_n u) + \theta_j' u$, so it suffices to show $\theta_j' u \to 0$ in $L^p(\mathbb{R}^n_+)$. Since $\gamma u = 0$ in $L^p(\mathbb{R}^{n-1})$ our formula above gives

$$u(x',x_n) = \int_0^{x_n} \partial_n u(x',t)\,dt , \qquad x' \in \mathbb{R}^{n-1} ,\ x_n > 0 ,$$

so we compute

$$|u(x',x_n)| \le \left(\int_0^{x_n} |\partial_n u(x',t)|^p\,dt\right)^{1/p} x_n^{1/p'} , \quad \text{(Hölder)}$$

$$\int_0^\infty |\theta_j'(x_n)u(x',x_n)|^p\,dx_n \le \int_0^{2/j}\int_0^{x_n} |\partial_n u(x',t)|^p\,dt\ x_n^{p-1} j^p\,dx_n$$

$$= j^p(2/j)^{p-1}\int_0^{2/j}\int_0^{x_n} |\partial_n u(x',t)|^p\,dt\,dx_n$$

$$= 2^{p-1} j\int_0^{2/j}\int_t^{2/j} |\partial_n u(x',t)|^p\,dx_n\,dt$$

$$\le 2^{p-1} j(2/j)\int_0^{2/j} |\partial_n u(x',t)|^p\,dt .$$

Integration over $\mathbb{R}^{n-1}$ gives

$$\|\theta_j' u\|_{L^p(\mathbb{R}^{n-1})}^p \le 2^p \int_0^{2/j} \int_{\mathbb{R}^{n-1}} |\partial_n u|^p \, dx \ ,$$

and this converges to zero as $j \to \infty$. This proves the following.

LEMMA 4.2. *The kernel of* $\gamma : W^{1,p}(\mathbb{R}^n_+) \to L^p(\mathbb{R}^{n-1})$ *is the closure in* $W^{1,p}(\mathbb{R}^n_+)$ *of* $C_0^\infty(\mathbb{R}^n_+) : \ker(\gamma) = W_0^{1,p}(\mathbb{R}^n_+)$.

Finally we shall describe the extension of the trace operator to a bounded open domain $G \subset \mathbb{R}^n$. Assume ∂G is an $n-1$ dimensional C^m *manifold*. Letting $Q = \{y \in \mathbb{R}^n : |y_i| \le 1\}$, $Q_0 = \{y \in Q : y_n = 0\}$ and $Q_+ = \{y \in Q : y_n > 0\}$, we state this last condition as follows. There is a collection of open bounded sets $\{G_j : 1 \le j \le N\}$ with $\cup\{G_j : 1 \le i \le N\} \supset \partial G$ and corresponding $\varphi_j \in C^m(Q, G_j)$ which are bijections of Q, Q_+ and Q_0 onto G_j, $G_j \cap G$ and $G_j \cap \partial G$, respectively, and each Jacobian $J(\varphi_j)$ is positive. (Each pair (φ_j, G_i) is a *coordinate patch*.) Let $G_0 = G$. We can construct $\beta_j \in C_0^\infty(G_j)$, $0 \le j \le N$, with $\beta_j(x) \ge 0$, and $\sum_{j=0}^N \beta_j(x) = 1$ for $x \in \overline{G}$. Thus, $\{\beta_j : 1 \le j \le N\}$ is a *partition-of-unity* subordinate to the open cover $\{G_j : 1 \le j \le N\}$ of ∂G, and $\{\beta_j : 0 \le j \le N\}$ is a partition-of-unity subordinate to the open cover $\{G_j : 0 \le j \le N\}$ of $\overline{G}$.

If f is a function defined on ∂G then

$$\int_{\partial G} f \equiv \sum_{j=1}^N \int_{\partial G \cap G_j} \beta_j f \, ds \equiv \sum_{j=1}^n \int_{Q_0} (\beta_j f) \circ \varphi_j(y', 0) J_j(y') \, dy'$$

where $s = \varphi_j(y', 0)$ and

$$J_j(y') = \left\{ \sum_{k=1}^N \left(\frac{\partial(s_1, \dots, \hat{s}_k, \dots, s_n)}{\partial(y_1, \dots, y_{n-1})} \right)^2 \Bigg|_{y_n = 0} \right\}^{1/2}$$

By the smoothness property, $|J_j(y')| \le K$, $1 \le j \le N$, $y' \in Q_0$, since $m \ge 1$. Finally, we construct the trace on ∂G as indicated.

$$\begin{array}{cccc}
W^{1,p}(G) & \longrightarrow & W_0^{1,p}(G) \times W^{1,p}(Q_+)^N & \\
u = \sum\limits_{j=0}^N \beta_j u & \longmapsto & \beta_0 u \ , \ (\beta_j u) \circ \varphi_j \ , & 1 \le j \le N \\
 & & \downarrow & \\
\gamma(u) \equiv \sum\limits_{j=1}^N \beta_j \gamma_j u & \longleftarrow & \gamma_j(\beta_j u \circ \varphi_j) = (\beta_j \gamma_j(u)) \circ \varphi_j \ , & 1 \le j \le N \\
L^p(\partial G) & & L^p(Q_0)^N &
\end{array}$$

That $\gamma(u) \in L^p(\partial G)$ follows from the estimates

$$\begin{aligned}\int_{\partial G} |\gamma(u)|^p \, ds &\le \sum_{j=1}^N \int_{\partial G\cap G_j} |\gamma_j u|^p \, ds \le K \sum_{j=1}^N \int_{Q_0} |\gamma_j(u)\circ\varphi_j|^p_{L^p} \, dy' \\ &\le K\cdot K_p^p \sum_{j=1}^N \|u\circ\varphi_j\|^p_{W^{1,p}(Q_+)} \le K\cdot K_p^p K_\varphi^p \sum_{j=1}^N \|u\|^p_{W^{1,p}(G\cap G_j)} \\ &\le K\cdot K_p^p\cdot K_\varphi^p\cdot N\|u\|^p_{W^{1,p}(G)} \ ,\end{aligned}$$

where K is the maximum of all Jacobians, K_p is the norm of trace from half-space, and K_φ is the largest norm in $W^{1,p}$ under change of variables $\tilde{\varphi}_j : W^{1,p}(Q_+) \to W^{1,p}(G\cap G_j)$. Clearly, if $u \in C(\overline{G}) \cap W^{1,p}(G)$ then $\gamma(u) = u|_{\partial G}$. We summarize the above as follows.

PROPOSITION 4.3. *Let $1 \le p < \infty$ and assume G is a bounded domain in $\mathbb{R}^n$ whose boundary ∂G is a C^1 manifold of dimension $n-1$. Then the linear trace operator $\gamma : W^{1,p}(G) \to L^p(\partial G)$ is continuous and uniquely determined by $\gamma(u) = u|_{\partial G}$ on those $u \in C^1(\overline{G})$. For any $u \in W^{1,p}(G)$, $\gamma(u) = 0$ in $L^p(\partial G)$ if and only if $u \in W_0^{1,p}(G)$.*

We consider the *extension* of functions in $W^{1,p}(G)$ to all of $\mathbb{R}^n$ in the situation of Proposition 4.3. First, one shows for the special case of a cube, Q, that

$$Eu(x) = \begin{cases} u(x', x_n)\ , & x_n \ge 0 \\ u(x', -x_n)\ , & x_n < 0\ . \end{cases}$$

determines an extension operator $E : W^{1,p}(Q_+) \to W^{1,p}(Q)$ for which $E(u)|_{Q_+} = u$, and $\|E(u)\|_{W^{1,p}(Q)} = 2\|u\|_{W^{1,p}(Q_+)}$. We apply this operator to each of the functions $(\beta_j u)\circ\varphi_j \in W^{1,p}(Q_+)$, and note that each $E(\beta_i u\circ\varphi_i)$ has support within Q. Thus we obtain an extension operator $E_G : W^{1,p}(G) \to W_0^{1,p}(\tilde{G})$ as above where $\tilde{G}$ is any open set containing $\cup\{G_j : 0 \le j \le N\}$. Of course, any open $\tilde{G} \supset G$ could have been specified first and then the partition $\{G_i\}$ chosen as before with $G_j \subset \tilde{G}$. This leads to the following.

COROLLARY 4.3. *Let G be given as above and $\tilde{G}$ be open, $\tilde{G} \supset G$. There exists an extension operator $E_G : W^{1,p}(G) \to W_0^{1,p}(\tilde{G})$ for which $E_G u(x) = u(x)$ for $x \in G$ and*

$$\|E_G(u)\|_{W^{1,p}(\tilde{G})} \le C(G,\tilde{G})\|u\|_{W^{1,p}(G)}\ , \quad u \in W^{1,p}(G)\ .$$

THEOREM 4.2 (RELLICH-KANDOROCHOV). *Let G be a bounded domain in $\mathbb{R}^n$ and $1 \le p < \infty$. Then the imbedding $W_0^{1,p}(G) \to L^p(G)$ is compact. If also the boundary ∂G is a C^1 manifold of dimension $n-1$, then the imbedding $W^{1,p}(G) \to L^p(G)$ is compact.*

PROOF. Choose an open set $\tilde{G}$ with $K \equiv \overline{G} \subset\subset \tilde{G}$, $\overline{G}$ denoting the closure of G. For each $\varphi \in C_0^\infty(G)$, the zero-extension belongs to $C_0^\infty(\tilde{G})$, and we denote it by φ. For any $h \in \mathbb{R}^n$ with $|h| < \operatorname{dist}(G, \partial\tilde{G})$ we have for $x \in G$

$$\varphi(x+h) - \varphi(x) = \int_0^1 \frac{d}{dt}\varphi(x+th)\,dt = \int_0^1 h \cdot \vec{\nabla}\varphi(x+th)\,dt\ ,$$

$$|\varphi(x+h) - \varphi(x)|^p \le |h|^p \int_0^1 |\vec{\nabla}\varphi(x+th)|^p\,dt\ .$$

Since $G + th \subset \tilde{G}$ for $t \in [0,1]$, we obtain from a change of order of integration

$$\begin{aligned}\int_G |\varphi(x+h) - \varphi(x)|^p\,dx &\le |h|^p \int_0^1 \left\{ \int_{G+th} |\vec{\nabla}\varphi(y)|^p\,dy \right\} dt \\ &\le |h|^p \int_{\tilde{G}} |\vec{\nabla}\varphi(y)|^p\,dy\ .\end{aligned}$$

This estimate holds for all $\varphi \in W_0^{1,p}(G)$. Let $\mathcal{F}$ be the unit ball in $W_0^{1,p}(\tilde{G})$. From the Frechet-Kolmogorov Theorem 3.4 it follows that $\mathcal{F}|_G$ is compact in $L^p(G)$, but $\mathcal{F}|_G$ is contained in the unit ball in $W_0^{1,p}(G)$. □

For the case of a domain G with smooth boundary, note that the composite operator,

$$W^{1,p}(G) \xrightarrow{E_G} W_0^{1,p}(\tilde{G}) \longrightarrow L^p(\tilde{G}) \longrightarrow L^p(G)$$

consisting of the continuous extension, the compact imbedding, and the continuous restriction, is necessarily compact.

We cite finally the following result from [1, p. 97]. It is very important for many types of nonlinear problems, especially those in which one needs to show that certain products of functions belong to an L^p class. The domain G satisfies the *cone condition* if there is a fixed cone K such that at any point $y \in \partial G$ one can place the vertex at y with $K - y$ lying within G. This is true, for example, in the case of polyhedra and of bounded domains with ∂G being a C^1 manifold.

THEOREM 4.3 (SOBOLEV). *Let G be a bounded domain which satisfies the cone condition in $\mathbb{R}^n$, $1 \le p < \infty$, and let $m \ge 0$ be an integer.*

If $m < \frac{n}{p}$, then

$$W^{m,p}(G) \to L^q(G)\ , \qquad p \le q \le \frac{np}{n - mp}\ .$$

If $m = \frac{n}{p}$, then the above holds for $p \le q < \infty$. If $p = 1$, this holds for $q = \infty$, and

$$W^{n,1}(G) \to C_b(G)\ ,$$

where $C_b(G)$ is the space of continuous and bounded functions on G with the sup norm. If $mp > n$, then

$$W^{m,p}(G) \to C_b(G)\ .$$

Assume ∂G is locally Lipschitz. If $m - 1 < \frac{n}{p} < m$, then

$$W^{m,p}(G) \to C^\lambda(\bar{G})\ , \qquad 0 < \lambda \le m - \frac{n}{p}\ ,$$

where $C^{\lambda}(\bar{G})$ is the space of (uniformly) continuous functions on $\bar{G}$ which satisfy a Hölder condition of exponent λ, $0 < \lambda \leq 1$. If $m - 1 = \frac{n}{p}$, then the above holds for all $0 < \lambda < 1$; if also $p = 1$, it holds for $\lambda = 1$.

II.5. Elliptic Boundary-Value Problems

We shall begin with an elementary but instructive model problem, a nonlinear Dirichlet problem. This will be resolved by our general existence theorems applied to operators from a Sobolev space to its dual. This will lead to additional general results as well as applications.

EXAMPLE 5.A. Let G be a domain in $\mathbb{R}^n$, $1 < p < \infty$, and define $a : \mathbb{R} \to \mathbb{R}$ by $a(\xi) = |\xi|^{p-1} \operatorname{sgn} \xi = |\xi|^{p-2}\xi$. Also let $F \in L^{p'}(G)$, where $1/p + 1/p' = 1$ as usual, and consider the nonlinear *Dirichlet problem* of finding a function $u : G \to \mathbb{R}$ which satisfies

$$-\sum_{j=1}^{n} \partial_j a\big(\partial_j u(x)\big) = F(x) , \qquad x \in G , \tag{5.1.a}$$

$$u(s) = 0 \quad , \qquad s \in \partial G , \tag{5.1.b}$$

in some sense. For example, since $a(\cdot)$ determines a Nemytskii operator from $L^p(G)$ to $L^{p'}(G)$ we seek $u \in W^{1,p}(G)$; then the condition (5.1.b) on ∂G is meaningful in the sense of trace. Finally, the partial differential equation (5.1.a) should hold (at least) in the sense of $\mathcal{D}^*(G)$, i.e.,

$$\sum_{j=1}^{n} \int_G a\big(\partial_j u(x)\big)\partial_j \varphi(x)\, dx = \int_G F(x)\varphi(x)\, dx , \qquad \varphi \in C_0^\infty(G) ,$$

and this is equivalent to requiring this identity to hold for all $\varphi \in W_0^{1,p}(G)$.

The linear case ($p = 2$) as discussed in Chapter I suggests that the Dirichlet problem (5.1) can be formulated as an operator equation. Thus, define $V = W_0^{1,p}(G)$ and

$$\mathcal{A}u(v) \equiv \sum_{j=1}^{n} \int_G a\big(\partial_j u(x)\big)\partial_j v(x)\, dx , \qquad u, v \in V .$$

This is meaningful by the estimates

$$|\mathcal{A}u(v)| \leq \sum_{j=1}^{n} \Big(\int_G |\partial_j u|^{\overbrace{p'(p-1)}^{p}}\Big)^{1/p'} \Big(\int_G |\partial_j v|^p\Big)^{1/p}$$

$$\leq \Big(\sum_{j=1}^{n} \|\partial_j u\|_{L^p}^p\Big)^{1/p'} \Big(\sum_{j=1}^{n} \|\partial_j v\|_{L^p}^p\Big)^{1/p} \leq \|u\|_{W^{1,p}}^{p-1} \|v\|_{W^{1,p}}$$

from which it follows that $\mathcal{A} : V \to V'$ is *bounded*. Moreover, $\mathcal{A}$ is the sum of terms which are composition of $\partial_j : W_0^{1,p} \to L^p$, $a(\cdot) : L^p \to L^{p'}$, and $\partial_j^* = -\partial_j : L^{p'} \to (W_0^{1,p})'$, each of which is continuous, so $\mathcal{A} : V \to V'$ is *continuous*. For a pair $u, v \in V$ we have

$$\langle \mathcal{A}u - \mathcal{A}v,\ u - v\rangle = \sum_{j=1}^{n} \int_G \Big(a\big(\partial_j u(x)\big) - a\big(\partial_j v(x)\big)\Big)\big(\partial_j u(x) - \partial_j v(x)\big)\, dx\ ,$$

and the integrands are non-negative a.e. in G because $a(\cdot)$ is monotone; thus, $\mathcal{A}$ is *monotone* and type M. Our Dirichlet problem (5.1) is of the form $\mathcal{A}u = f$ in V', where $f \in V'$ is defined by

$$f(v) = \int_G F(x)v(x)\, dx\ , \qquad v \in V\ ,$$

and we need only to verify that $\mathcal{A}$ is V-coercive in order to apply our general existence theorems. Note that

$$\mathcal{A}v(v) = \sum_{j=1}^{n} \|\partial_j v\|_{L^p}^p\ .$$

LEMMA 5.1 (POINCARÉ). *Assume G is bounded in the first coordinate direction:*

$$\sup\{|x_1| : x = (x_1, \dots, x_n) \in G\} \equiv K < \infty\ .$$

Then $\|v\|_{L^p(G)} \le pK\|\partial_1 v\|_{L^p(G)}$ for all $v \in W_0^{1,p}(G)$.

PROOF. It suffices to consider $\varphi \in C_0^\infty(G)$ and note

$$\partial_1\big(x_1|\varphi(x)|^p\big) = |\varphi(x)|^p + x_1 p|\varphi(x)|^{p-1}\operatorname{sgn}\big(\varphi(x)\big)\partial_1\varphi(x)\ .$$

Integrating this identity over G and using the divergence theorem yields the estimate

$$\int_G |\varphi(x)|^p\, dx \le pK\|\varphi\|_{L^p}^{p-1}\|\partial_1\varphi\|_{L^p}$$

and, hence, the desired result. □

COROLLARY 5.1. *If G is bounded in some direction, there exists $c > 0$ for which*

$$\sum_{j=1}^{n} \|\partial_j v\|_{L^p(G)}^p \ge c\|v\|_{W^{1,p}(G)}^p\ , \qquad v \in W_0^{1,p}(G)\ .$$

We remark that Lemma 5.1 holds if $v \in W^{1,p}(G)$ and $\gamma v(s) = 0$ for those $s \in \partial G$ with $s_1\nu_1(s) > 0$, where $\vec{\nu} = (\nu_1, \dots, \nu_n)$ is the unit outward normal on ∂G. The problem (5.1) extends easily to the following.

PROPOSITION 5.1 (BROWDER-VISHIK). *Let G be a domain in $\mathbb{R}^n$, bounded in some direction, $1 < p < \infty$. Assume given the functions $A_j : G \times \mathbb{R}^{n+1} \to \mathbb{R}$ for $j = 0, 1, \dots, n$ which satisfy*

(i) *$A_j(x,\xi)$ is measurable in x and continuous in ξ,*

(ii) $|A_j(x,\xi)| \le c\sum_{j=0}^n |\xi_j|^{p-1} + k(x)$, $\qquad \xi \in \mathbb{R}^{n+1}$, $x \in G$

(iii) $\sum_{j=0}^n (A_j(x,\xi) - A_j(x,\eta))(\xi_j - \eta_j) \ge 0$, $\qquad \xi, \eta \in \mathbb{R}^{n+1}$, $x \in G$

(iv) $\sum_{j=0}^n A_j(x,\xi)\xi_j \ge c_0\sum_{j=1}^n |\xi_j|^p - k(x)$, $\qquad \xi \in \mathbb{R}^{n+1}$, $x \in G$

where $c_0 > 0$, $k \in L^{p'}(G)$. Let $F_0, F_1, \ldots, F_n$ be given in $L^{p'}(G)$. Then the set of solutions to the generalized Dirichlet problem

$$u \in W_0^{1,p}(G) : \sum_{j=0}^{n} \int_G A_j\big(x, u(x), \vec{\nabla} u(x)\big)\partial_j v(x)\, dx \tag{5.2}$$

$$= \sum_{j=0}^{n} \int_G F_j(x)\partial_j v(x)\, dx \text{for all } \ v \in W_0^{1,p}(G)$$

is closed, convex, bounded and non-empty.

PROOF. From (i) and (ii) we define $\mathcal{A} : W_0^{1,p}(G) \to W_0^{1,p}(G)'$ by the left side of (5.2) and note that it is continuous with bound given by $\|\mathcal{A}(u)\|_{(W_0^{1,p})'} \leq c\|u\|_{W_0^{1,p}}^{p-1} + \|k\|_{L^{p'}}(n+1)^{1/p'}$. That $\mathcal{A}$ is monotone and coercive follow from (iii) and (iv), respectively. □

Non-homogeneous boundary conditions can be attained routinely in the above situation. We need only consider the translation $\mathcal{A}_0(u) = \mathcal{A}(u + u_0)$ where $u_0 \in W^{1,p}(G)$ is prescribed. Alternatively, we may set $K = u_0 + W_0^{1,p}(G)$ and use Theorem 2.3. The partial differential equation solved in $\mathcal{D}^*(G)$ in (5.2) is

$$-\sum_{j=0}^{n} \partial_j A_j\big(x, u(x), \vec{\nabla} u(x)\big) = F_0(x) - \sum_{j=1}^{n} \partial_j F_j(x)$$

which is *quasilinear* with "first-order" distributions on the right-side — rather non-regular and not necessarily even functions. The easiest examples of A_j's satisfying (iii) are those which depend only on x, ξ_j, i.e., $A_j(x,\xi) = c_j(x)a_j(\xi_j)$ with $0 \leq c_j \in L^\infty(G)$ and $a_j : \mathbb{R} \to \mathbb{R}$ monotone and continuous. We shall relax these implicit restrictions in the class of *quasimonotone* operators in Section 6.

Finally, note that if the inequality in (iii) is strict for $\xi \neq \eta$, then whenever $u, v \in W_0^{1,p}(G)$ with $\mathcal{A}u = \mathcal{A}v$, hence, $\langle \mathcal{A}(u) - \mathcal{A}(v),\ u - v\rangle = 0$, then

$$\sum_{j=0}^{n}\Big(A_j\big(x, u(x), \vec{\nabla} u(x)\big) - A_i\big(x, v(x), \vec{\nabla} v(x)\big)\Big)\partial_j\big(u(x) - v(x)\big) = 0\ , \quad \text{a.e.} \quad x \in G\ ,$$

hence, $(u(x), \vec{\nabla} u(x)) = (v(x), \nabla v(x))$ in $\mathbb{R}^{n+1}$ for a.e. x. That is, there is *at most* one solution. But this rather special restriction can be considerably relaxed!

We pass on to different boundary-value problems for similar quasilinear equations. For simplicity we consider the following special case:

Let G be a bounded domain in $\mathbb{R}^n$ with ∂G a C^1 manifold, $1 < p < \infty$, and $a(\xi) = |\xi|^{p-1}\operatorname{sgn}\xi$, $b(\cdot) \in L^\infty(G)$, $c(\cdot) \in L^\infty(\partial G)$, both $b(\cdot)$ and $c(\cdot)$ are non-negative, and define

$$\mathcal{A}u(v) = \int_G \left\{\sum_{j=1}^{n} a\big(\partial_j u(x)\big)\partial_j v(x) + b(x)a\big(u(x)\big)v(x)\right\} dx \tag{5.3}$$

$$+ \int_{\partial G} c(s)a\big(u(s)\big)v(s)\, ds\ , \qquad u, v \in W^{1,p}(G)\ .$$

Here we denote $u(s) = (\gamma u)(s)$, where $\gamma : W^{1,p}(G) \to L^p(\partial G)$ is the trace operator. Let $F \in L^{p'}(G)$, $F_0 \in L^{p'}(\partial G)$ and set

$$f(v) = \int_G F(x)v(x)\,dx + \int_{\partial G} F_0(s)\gamma v(s)\,ds\ , \qquad v \in W^{1,p}(G)\ .$$

Although $\mathcal{A}$, f are defined on all of $W^{1,p}(G)$, we shall frequently consider their restrictions to a given closed subspace V, with $W_0^{1,p}(G) \subset V \subset W^{1,p}(G)$. The extreme case $V = W_0^{1,p}$ (with $b = 0$, $c = 0$, $F_0 = 0$) gave us the original model problem, the *Dirichlet problem* (5.1) known as the boundary-value problem of *first type*. In this case, the boundary conditions are obtained from the inclusion, $u \in V$, and these are called the *stable boundary conditions*. That is, they are imposed by the definition of the closed subspace V.

EXAMPLE 5.B. Choose $V = W^{1,p}(G)$, the other extreme case. Then $\mathcal{A} : V \to V'$ is monotone, continuous and bounded; if $b(x) \geq b_0 > 0$ then $\mathcal{A}$ is strictly-monotone and V-coercive. Thus, there is a unique

$$u \in V : \mathcal{A}u(v) = f(v)\ , \qquad v \in V\ , \tag{5.4}$$

and we shall characterize it as a partial differential equation and boundary condition. Note that there are no boundary constraints imposed in this case by the inclusion, $u \in V$. First, since $C_0^\infty(G) \subset V$ we obtain

$$-\sum_{j=1}^{n} \partial_j a(\partial_j u) + b(x)a(u) = F$$

in $L^{p'}(G) \subset \mathcal{D}^*(G)$. Substituting this back into the functional equation yields

$$\int_G \Big\{ \sum_{j=1}^{n} a(\partial_j u)\partial_j v + \sum_{j=1}^{n} \partial_j \big(a(\partial_j u)\big)v \Big\} + \int_{\partial G} \big(c(s)a(u) - F_0\big)v\,ds = 0\ , \quad v \in V$$

Now (formally, unless u is smoother) this first integrand is $\sum_{j=1}^{n} \partial_j(a(\partial_j u)v)$ and the divergence theorem gives

$$\int_{\partial G} \Big(\sum_{j=1}^{n} a(\partial_j u)\nu_j + c(s)a(u) - F_0 \Big) \gamma v(s)\,ds = 0\ , \qquad v \in V\ .$$

But $Rg(\gamma)$ is dense in $L^p(\partial G)$, so it follows that a (sufficiently smooth) solution of (5.4) is characterized by

(5.5.a)

$$u \in W^{1,p}(G) : -\sum_{j=1}^{n} \partial_j a(\partial_j u) + b(x)a(u) = F \ \text{ in } \ L^{p'}(G)$$

$$\sum_{j=1}^{n} a(\partial_j u)\nu_j + c(s)a\big(u(s)\big) = F_0 \ \text{ in } \ L^{p'}(\partial G)\ . \tag{5.5.b}$$

When $c \equiv 0$ this is a nonlinear *Neumann problem*, a boundary-value problem of *second type*, while for $c > 0$ it is a *Robin problem*, a boundary condition of *third type*.

EXAMPLE 5.C. Problems of *mixed type* (e.g., first and third type) arise as follows. Let Γ_1 be a subset of ∂G and $\Gamma_2 = \partial G \sim \Gamma_1$, the complement. Choose $V = \{u \in W^{1,p}(G) : \gamma u|_{\Gamma_1} = 0\}$. Then the solution to the functional equation (5.4) satisfies the stable boundary condition $u|_{\Gamma_1} = 0$ since $u \in V$, and it also satisfies the third type complementary boundary condition (5.5.b) on Γ_2, since $Rg\{\gamma|_V\}$ is dense in $L^{p'}(\Gamma_2)$.

EXAMPLE 5.D. Finally, we mention a boundary condition of *non-local* character. For this we choose

$$V = \{\ v \in W^{1,p}(G) : \gamma(v) \ \text{ is constant in } \ L^p(\partial G)\ \}\ .$$

Since the (one-dimensional) subspace of constant functions is closed in $L^p(\partial G)$ and γ is continuous, it follows that V is closed in $W^{1,p}(G)$; as before, $\mathcal{A} : V \to V'$ is strictly-monotone, continuous, bounded and coercive. If $\mathcal{A}u = f$, as above, then we find

$$u \in W^{1,p}(G)\ ,\ -\sum_{j=1}^{n} \partial_j a(\partial_j u) + b(x)a(u) = F \ \text{ in } \ L^{p'}(G) \tag{5.6.a}$$

$$\gamma u(s) = c_0\ ,\ s \in \partial G\ ,\ \text{for some } c_0 \in \mathbb{R}\ ,\ \text{and} \tag{5.6.b}$$

$$\int_{\partial G} \sum_{j=1}^{n} a(\partial_j u)\nu_j\, ds + a(c_0) \int_{\partial G} c(s)\, ds = \int_{\partial G} F_0 \ \text{ in } \ \mathbb{R}\ . \tag{5.6.c}$$

This is the *Adler problem*, a boundary-value problem of *fourth type*. Note that the constant c_0 is not known a-priori; otherwise, this would be an over-determined Dirichlet problem. Note that (5.6.b) is the stable boundary condition and (5.6.c) is the complementary boundary condition.

In each of the preceding problems we used the same operator $\mathcal{A}$ but varied the space V; $\mathcal{A} : V \to V'$ is monotone, continuous and bounded. It is not necessarily coercive: if $b(\cdot)$ and $c(\cdot)$ are both zero, then $\mathcal{A}u(u) = 0$ whenever u is constant. If V contains non-zero constant functions then $\mathcal{A}$ is *not* coercive on V. We shall show these are the *only* ways this $\mathcal{A}$ can fail to be coercive!

We shall use the Rellich-Kandorochov Theorem 4.2 to implement the following general situation.

PROPOSITION 5.2. *Suppose p, q, r are seminorms on a linear space V and that $\|x\| \equiv p(x) + r(x)$, $|x| \equiv p(x) + q(x)$ are norms. Let $B = \{V, \|\cdot\|\}$, $B_1 = \{V, |\cdot|\}$, $B_0 = \{V, r(\cdot)\}$ and assume $B \hookrightarrow B_1$ is continuous, $B \hookrightarrow B_0$ is compact, and B is reflexive. Then $B_1 \hookrightarrow B$ is continuous, hence, $\|\cdot\|$ and $|\cdot|$ are equivalent norms.*

PROOF. Otherwise there is a sequence $\{v_n\}$ such that $|v_n| \to 0$ and $\|v_n\| = 1$ for $n \geq 1$. Since $\{v_n\}$ is bounded in B, some subsequence (with the same notation) satisfies $v_n \to v$ in B_0, $v_n \rightharpoonup v$ in B. But then $v_n \rightharpoonup v$ in B_1 so $v = 0$. Since $v_n \to 0$ in B_0 and B_1, we have $v_n \to 0$ in B. □

As an immediate application take V a closed subspace of $W^{1,p}(G)$, $1 < p < \infty$, and

$$p(v) = \left(\sum_{j=1}^{n} \|\partial_j v\|_{L^p(G)}^p\right)^{1/p}, \quad r(v) = \|v\|_{L^p(G)}$$

$$q(v) = \left(\int_G b(x)|v(x)|^p \, dx\right)^{1/p} + \left(\int_{\partial G} c(s)|v(s)|^p \, ds\right)^{1/p}, \quad v \in V .$$

where $b \in L^\infty(G)$, $c \in L^\infty(\partial G)$ are non-negative. Then Proposition 5.2 applies if $p + q$ is a norm, that is, if one of $b(\cdot)$ or $c(\cdot)$ is non-identically-zero or if zero is the only constant function in V. Thus, in either case, it follows that $\mathcal{A} : V \to V'$ is V-coercive and so we obtain existence of solutions.

Finally, we shall extract the essential points of our disection of $\mathcal{A} : V \to V'$ into a partial differential equation (5.5.a) and a complementary boundary condition (5.5.b). This construction results in an *abstract Green's theorem.*

Assume V, B are Banach spaces and $\gamma : V \to B$ is a strict homomorphism with kernel given by $\ker(\gamma) = V_0$. Thus the quotient map is an isomorphism of V/V_0 onto B, and it follows that the dual map $\gamma^*(g) = g \circ \gamma$ defines an isomorphism of B' onto $V_0^\perp \equiv \{f \in V' : f|_{V_0} = 0\}$, the *annihilator* of V_0 in V'.

EXAMPLE 5.E. Take $V = W^{1,p}(G)$, $\gamma =$ trace, and $B \equiv Rg(\gamma)$; it is known that $B = W^{1-1/p,p}(\partial G)$. We have shown that $V_0 = W_0^{1,p}(G)$; since $C_0^\infty(G)$ is dense in V_0 it follows $V_0' \subset \mathcal{D}^*(G)$.

Assume $m_0 : V \times V \to \mathbb{R}$ is a continuous semi-scalar-product, $|\cdot|$ is a continuous seminorm on V and define $|v|_{\mathcal{X}} \equiv m_0(v,v)^{1/2} + |v|$, $v \in V$; $\mathcal{X} = \{V, |\cdot|_{\mathcal{X}}\}$ is a seminorm space, and $\mathcal{X}' \subset V'$. Assume V_0 is dense in $\mathcal{X}$; then $\mathcal{X}' \subset V_0'$. We call $\mathcal{X}'$ a *pivot space* since its elements belong to both V_0' and V'. Let $M_1 : B \to B'$ be given.

EXAMPLE 5.E (CONTINUED). Set $m_0(u,v) = \int_G m_0(x)u(x)v(x)\,dx$ where $m_0(x) \geq 0$ and $m_0 \in L^{p/p-2}(G)$, $p \geq 2$. Then $\mathcal{X}' = \{m_0^{1/2}u : u \in L^2(G)\}$. Similarly construct $M_1\varphi(\psi) = \int_{\partial G} m_1(s)\varphi(s)\psi(s)\,ds$, $\varphi, \psi \in B$.

Define $\mathcal{M} : V \to V'$ by $\mathcal{M}u(v) = m_0(u,v) + M_1(\gamma u)(\gamma v)$, for $u, v \in V$, and $M : V \to V_0'$ by $Mu = \mathcal{M}(u)|_{V_0}$. Thus $Mu \in \mathcal{X}'$ and $Mu(v) = m_0(u,v)$ for $u, v \in V$, since V_0 is dense in $\mathcal{X}$. This gives the identity $\mathcal{M}u(v) = Mu(v) + M_1(\gamma u)(\gamma v)$, $u, v \in V$, hence, $\mathcal{M}(u) = Mu + \gamma' M_1 \gamma(u)$, $u \in V$, a trivial decomposition of this element of V' as the sum of one each from V_0' and from $V_0^\perp$.

Let $\mathcal{A} : V \to V'$ be given and define the *formal operator* $A(u) = \mathcal{A}(u)|_{V_0}$ to be the indicated restriction and consider it on the domain $D \equiv \{u \in V : A(u) \in \mathcal{X}'\}$. Then for each $u \in D$, $\mathcal{A}(u) - A(u) \in V_0^\perp$ in V'. Since $Rg(\gamma') = V_0^\perp$, there is a unique $\partial_A u \in B'$ for which $\mathcal{A}(u) - A(u) = \gamma'(\partial_A u)$. This defines the operator $\partial_A : D \to B'$ for which

$$\mathcal{A}u(v) = Au(v) + \partial_A u(\gamma v), \quad u \in D, \quad v \in V . \tag{5.7}$$

This is the abstract *Green's formula* for $\mathcal{A}$, and ∂_A is the corresponding *boundary operator*. Such formulas will be very useful in a vast variety of situations, and they will be used frequently hereafter. We summarize this elementary construction in the following.

PROPOSITION 5.3. *Assume $\gamma : V \to B$ is a strict surjective homomorphism between Banach spaces and V_0 is the kernel of γ. Assume $m_0 : V \times V \to \mathbb{R}$ is a continuous semi-scalar-product and $|\cdot|$ is a continuous seminorm on V; let $|v|_{\mathcal{X}} = m_0(v,v)^{1/2} + |v|$, and assume V_0 is dense in the seminorm space $\mathcal{X} \equiv \{V, |\cdot|_{\mathcal{X}}\}$. Thus $\mathcal{X}' \subset V'$ and $\mathcal{X}' \hookrightarrow V_0'$. Assume that operators $M_1 : B \to B'$ and $\mathcal{A} : V \to V'$ are given; define $M \in \mathcal{L}(V, \mathcal{X}')$ and $\mathcal{M} : V \to V'$ by $Mu(v) = m_0(u,v)$, $\mathcal{M}u(v) = m_0(u,v) + M_1(\gamma u)(\gamma v)$, $u, v \in V$, so that $\mathcal{M} = M + \gamma' M_1 \gamma$. Define the formal part $A : V \to V_0'$ of $\mathcal{A}$ by $Au = \mathcal{A}u|_{V_0}$, $u \in V$, and set $D = \{u \in V : Au \in \mathcal{X}'\}$. Then there is a unique $\partial_A : D \to B'$ for which (5.7) holds.*

EXAMPLE 5.E (CONTINUED). Let the operator $\mathcal{A}$ on $V = W^{1,p}(G)$ be given by (5.3). Then the formal part is $A(u) = -\sum_{j=1}^{n} \partial_j a(\partial_j u) + b(x)a(u)$ and D is the set of $u \in W^{1,p}$ for which the partial differential equation takes values in $\mathcal{X}'$, the space determined above by $m_0(x)$. (If $m_0(x) = 1$, then $\mathcal{X}' = L^2(G)$.) Although the boundary operator is defined on all of D, we can compute its value on the smooth functions $u \in D$ by the divergence theorem as

$$\partial_A u(\psi) = \int_{\partial G} \Big(\sum_{j=1}^{n} a(\partial_j u)\nu_j + c(s)a\big(u(s)\big) \Big) \psi(s)\, ds\ , \quad \psi \in Rg(\gamma)\ ,$$

just as above. In fact, ∂_A is given on D by the identity (5.7) once we verify that its value depends only on γv.

We show how the preceding construction is useful in representing the solution of an abstract functional equation as a formal part (the partial differential equation) and a complimentary part (the boundary condition in $Rg(\gamma)$); these conditions are in addition to those imposed directly by the space V. Given the spaces and operators above, let $F \in \mathcal{X}'$, $g \in B'$ and define $f \in V'$ by $f = F + \gamma'(g)$. Consider the functional equation

$$u \in V : \lambda \mathcal{M}(u) + \mathcal{A}(u) = f \ \text{ in } \ V'\ . \tag{5.8}$$

That is,

$$\lambda\big(Mu(v) + M_1(\gamma u)(\gamma v)\big) + \mathcal{A}u(v) = F(v) + g(\gamma v)\ , \qquad v \in V\ .$$

From our abstract Green's theorem it follows that this is equivalent to

$$u \in D\ , \quad \lambda M(u) + A(u) = F \ \text{ in } \ \mathcal{X}'\ , \ \text{ and} \tag{5.9.a}$$

$$\lambda M_1(\gamma u) + \partial_A(u) = g \ \text{ in } \ B'\ . \tag{5.9.b}$$

This is the desired decomposition of the functional equation (5.8) into a partial differential equation (5.9.a) and a complementary boundary condition (5.9.b).

An Elliptic System.

All of our examples above of boundary-value problems for an elliptic equation can be extended in various ways to the case of elliptic systems. However it is often the case that the structure of these systems does *not* permit the direct application of our general existence theorems, especially with respect to monotone or coercive properties. Here we shall describe a model system to which our theory does directly apply. This will be done in the abstract setting developed above. We shall then show by examples that this model system contains a variety of systems of boundary-value problems, including those which are coupled along a common boundary, or throughout an interior region, or the case of an elliptic problem in the tangent space on a boundary or interior submanifold of a domain which is coupled to a problem on the full domain.

Suppose V_1, V_2, U are reflexive Banach spaces. For $j = 1, 2$, let $\lambda_j \in \mathcal{L}(V_j, U)$ with dual $\lambda_j' \in \mathcal{L}(U', V_j')$ and assume operators $\mathcal{A}_j : V_j \to V_j'$ and $\mathcal{B} : U \to U'$ are given. On the product space $V = V_1 \times V_2$ with elements denoted by $u = [u_1, u_2]$, $u_j \in V_j$, we define $\lambda : V \to U$ by $\lambda u = \lambda_1 u_1 - \lambda_2 u_2$ and $\mathcal{A} : V \to V'$ by

$$\mathcal{A}u(v) = \mathcal{A}_1 u_1(v_1) + \mathcal{A}_2 u_2(v_2) + \mathcal{B}(\lambda u + u_0)(\lambda v) \ , \qquad u, v \in V \ , \tag{5.9}$$

where $u_0 \in U$ is given. Thus, $\mathcal{A}u = [\mathcal{A}_1 u_1,\ \mathcal{A}_2 u_2] + \lambda' \mathcal{B}(\lambda u + u_0)$, $u \in V$, where $\lambda' g = [\lambda_1' g, -\lambda_2' g] \in V'$ is the dual of λ. For any $f = [f_1, f_2] \in V'$, the equation $\mathcal{A}u = f$ is equivalent to the system

$$u_1 \in V_1 : \mathcal{A}_1 u_1 + \lambda_1' \mathcal{B}(\lambda_1 u_1 - \lambda_2 u_2 + u_0) = f_1 \text{ in } V_1' \ , \tag{5.10.a}$$

$$u_2 \in V_2 : \mathcal{A}_2 u_2 - \lambda_2' \mathcal{B}(\lambda_1 u_1 - \lambda_2 u_2 + u_0) = f_2 \text{ in } V_2' \ . \tag{5.10.b}$$

The special form of the coupling λ is motivated by models in which exchange flux or forces are determined by *differences*. However, we need here only that $\lambda \in \mathcal{L}(V, U)$. In order to apply the results of Section 2 to such a system, the following observations are necessary.

PROPOSITION 5.4. *If each of the operators $\mathcal{A}_1$, $\mathcal{A}_2$, $\mathcal{B}$ is either monotone, hemicontinuous, demicontinuous, continuous, or bounded, then the operator $\mathcal{A}$, as given by* (5.9), *has the same property. If both of $\mathcal{A}_1$, $\mathcal{A}_2$ are Type-M and if $\lambda' \mathcal{B} \lambda$ is either completely continuous or monotone and weakly continuous, then $\mathcal{A}$ is Type-M.*

PROOF. The verifications of all but the last property are immediate. The Type-M property follows from Example 2.B and Example 2.C. □

EXAMPLE 5.F. Let G_1 and G_2 be disjoint bounded domains in $\mathbb{R}^n$ which share a common portion Γ of their boundaries: Γ is a manifold of dimension $n - 1$ with $\Gamma \subset \partial G_1 \cap \partial G_2$. Let $U = L^p(\Gamma)$ and $V_j = W^{1,p}(G_j)$ for $j = 1, 2$, and set $\lambda_j = \gamma_j|_\Gamma$, the restriction of the traces to Γ. Choose for simplicity (cf., (5.3))

$$\mathcal{A}_j u(v) = \int_{G_j} \Big\{ \sum_{i=1}^{n} a(\partial_i u) \partial_i v \Big\} \, dx \ , \qquad u, v \in V_j$$

so the formal operators are given by

$$A_1 u = A_2 u = -\sum_{j=1}^{n} \partial_j a(\partial_j u) \quad \text{in } V_j'$$

and the boundary operators are

$$\partial_{A_j} u = \sum_{i=1}^{n} a(\partial_i u) \cdot \nu_i^j \ , \quad j = 1, 2 \ ,$$

where $\vec{\nu}^j$ is the unit outward normal on ∂G_j, hence, $\vec{\nu}^1 = -\vec{\nu}^2$ on Γ. Let $F_j \in L^{p'}(G_j)$, $g_j \in L^{p'}(\partial G_j)$ and define

$$f_j(v) = \int_G F_j v \, dx + \int_{\partial G_j} g_j \gamma_j v \, ds \ , \qquad v \in V_j \ ,$$

for $j = 1, 2$. Let $b : \mathbb{R} \to \mathbb{R}$ be continuous with $|b(\xi)| \le c(|\xi|^{p-1} + 1)$ and define $\mathcal{B} : U = L^p(\Gamma) \to U'$ by Nemytskii's Theorem 3.2 as $\mathcal{B}(\psi)(x) = b(\psi(x))$, $x \in \Gamma$. Then the system (5.10) is the generalized formulation of the *boundary-coupled* elliptic system

$$(5.11.a) \quad A_j(u_j) = F_j \text{ in } L^{p'}(G_j) \ , \partial_{A_j} u_j = g_j \text{ in } L^{p'}(\partial G_j \sim \Gamma) \ , \quad j = 1, 2$$

$$(5.11.b) \quad \partial_{A_1} u_1 + b(\gamma_1 u_1 - \gamma_2 u_2) = g_1 \ , \partial_{A_2} u_2 - b(\gamma_1 u_1 - \gamma_2 u_2) = g_2 \text{ in } L^{p'}(\Gamma) \ .$$

Such problems arise as the description of diffusion in two regions G_1, G_2 with a distributed flux $b(\gamma_1 u_1 - \gamma_2 u_2)$ across their common boundary interface Γ, or in elasticity where G_1, G_2 are membranes with a distributed elastic constraint along the interface Γ.

EXAMPLE 5.G. Let $G_0 \subset G$ be bounded domains in $\mathbb{R}^n$ and set $V_j = W_0^{1,p}(G)$, $U = L^p(G_0)$, and let $\lambda_j : V_j \to U$ be the restriction to G_0 of the imbedding. The dual $\lambda_j' : U' \to V_j'$ is the zero-extension operator. Define $\mathcal{A}_j$, A_j, ∂_{A_j} as above on $G_j = G$, $j = 1, 2$. Likewise define f_j as above, but with $g_j = 0$, and construct $\mathcal{B} : L^p(G_0) \to L^{p'}(G_0)$ from $b(\cdot)$ as above. With this data the system (5.10) corresponds to the *interior-coupled* system.

$$(5.12.a) \quad A_1 u_1 + \lambda_1' b(u_1 - u_2) = F_1 \text{ in } L^{p'}(G) \ , \qquad u_1 = 0 \text{ in } L^p(\partial G) \ ,$$

$$(5.12.b) \quad A_2 u_2 - \lambda_2' b(u_1 - u_2) = F_2 \text{ in } L^{p'}(G) \ , \qquad u_2 = 0 \text{ in } L^p(\partial G) \ .$$

Such problems arise as double porosity models of diffusion through a composite media. Each equation results from conservation of fluid within each of the two components of the medium, and the second term corresponds to the flux exchanged between the two component materials. The description of displacements of a parallel pair of membranes which are elastically connected on G_0 is also given in this form.

EXAMPLE 5.H. Let G be a bounded domain in $\mathbb{R}^n$ and let $\Gamma \subset \partial G$ be a smooth manifold of dimension $n-1$. Set $V_1 = W^{1,p}(G)$, $V_2 = W^{1,p}(\Gamma)$, and $U = L^p(\Gamma)$, and define $\lambda_1 = \gamma|_\Gamma$, the trace restricted to Γ, $\lambda_2 = I$, the identity. Define $\mathcal{A}_1$ as before on G and construct $\mathcal{A}_2$ similarly with *tangential coordinates* in Γ. (For simplicity, we could assume $\Gamma \subset \mathbb{R}^{n-1}$, i.e., that Γ is *flat*.) Then A_2 is an elliptic operator in the tangential variables and ∂_{A_2} is characterized as a derivative in the direction of the outward normal to $\partial\Gamma$. Choose $\mathcal{B} : L^p(\Gamma) \to L^{p'}(\Gamma)$ as in Example 5.F. Assume $F_1 \in L^{p'}(G)$, $F_2 \in L^{p'}(\Gamma)$, $g_1 \in L^{p'}(\partial G)$, $g_2 \in L^{p'}(\partial\Gamma)$ are given and define

$$f_1(v) = \int_G F_1 v \, dx + \int_{\partial G} g_1 \gamma v \, ds \ , \qquad v \in V_1 \ ,$$

$$f_2(w) = \int_\Gamma F_2 w \, ds + \int_{\partial\Gamma} g_2 \gamma_\Gamma w \, dt \ , \qquad w \in V_2 \ ,$$

where $\gamma_\Gamma : W^{1,p}(\Gamma) \to L^p(\partial\Gamma)$ is the trace operator. A solution of the system (5.10) with such data is characterized by

$$(5.13.a) \quad A_1 u_1 = F_1 \text{ in } L^{p'}(G) \ , \quad \partial_{A_1} u_1 = g_1 \text{ in } L^{p'}(\partial G \sim \Gamma) \ ,$$

$$(5.13.b) \quad \partial_{A_1} u_1 + b(\gamma u_1 - u_2) = g_1 \text{ in } L^{p'}(\Gamma) \ ,$$

$$(5.13.c) \quad A_2 u_2 - b(\gamma u_1 - u_2) = F_2 \text{ in } L^{p'}(\Gamma) \ , \quad \partial_{A_2} u_2 = g_2 \text{ in } L^{p'}(\partial\Gamma) \ .$$

Here (5.13.a) is a boundary-value problem in the region G and (5.13.c) is a boundary-value problem for an *elliptic equation in the tangent space*. The exchange flux, $b(\gamma u_1 - u_2)$, occurs in the first as a sink on the boundary and in the second as a source in the interior. Such a problem describes diffusion in a region on part of whose boundary is a singular region of such high permeability that there is substantial diffusion tangential to the boundary. Similar problems occur when there is an $(n-1)$-dimensional submanifold in the interior within which diffusion occurs in the tangential directions. Such models are used to describe flow in narrow fractures in a medium.

EXAMPLE 5.I. Let G be a bounded domain in $\mathbb{R}^n$ for which the hypersurface $\Gamma = \{x = (x', x_n) \in G : x_n = 0\}$ is a domain in $\mathbb{R}^{n-1}$. Set $G_1 = \{x \in G : x_n < 0\}$ and $G_2 = \{x \in G : x_n > 0\}$, and define V_j, λ_j, $\mathcal{A}_j$, f_j for $j = 1, 2$ as in Example 5.F. Choose $U = L^p(\Gamma) \times W_0^{1,p}(\Gamma)$. Let $b_j : \mathbb{R} \to \mathbb{R}$ be continuous with $|b_j(\xi)| \leq c(|\xi|^{p-1} + 1)$, $j = 1, 2$, and define $\mathcal{B}_1 : L^p(\Gamma) \to L^{p'}(\Gamma)$ by $\mathcal{B}_1(\varphi)(s) = b_1(\varphi(s))$, $s \in \Gamma$, and $\mathcal{B}_2 : W_0^{1,p}(\Gamma) \to W_0^{1,p}(G)'$ by

$$\mathcal{B}_2\varphi(\psi) = \int_\Gamma \Big(\sum_{j=1}^{n-1} b_2(\partial_j\varphi)\partial_j\psi\Big)\, ds\ , \qquad \varphi, \psi \in W_0^{1,p}(\Gamma)\ .$$

Thus the formal part is given by $B_2\varphi = -\sum_{j=1}^{n-1} \partial_j b_2(\partial_j\varphi)$. We define $\mathcal{B} : U \to U'$ by

$$\mathcal{B}\varphi(\psi) = \mathcal{B}_1\varphi_1(\psi_1) + \mathcal{B}_2\varphi_2(\psi_2)\ , \quad \varphi = [\varphi_1, \varphi_2]\ , \quad \psi = [\psi_1, \psi_2] \in U\ .$$

Finally, let $\alpha \in [0, 1]$ and define the space

$$V \equiv \{v = [v_1, v_2] \in V_1 \times V_2 : \alpha\lambda_1 v_1 + (1-\alpha)\lambda_2 v_2 \in W_0^{1,p}(\Gamma)\}\ .$$

Note that V is a Banach space with the norm

$$\|v\|_V = \|v_1\|_{V_1} + \|v_2\|_{V_2} + \|\alpha\gamma_1 v_1 + (1-\alpha)\gamma_2 v_2\|_{W_0^{1,p}(\Gamma)}\ .$$

Define $\lambda \in \mathcal{L}(V, U)$ by

$$\lambda v = [\lambda_1 v_1 - \lambda_2 v_2\ , \alpha\lambda_1 v_1 + (1-\alpha)\lambda_2 v_2]\ , \qquad v \in V\ ,$$

and $\mathcal{A} : V \to V'$ by (5.9). The results of Proposition 5.4 hold for this case, and the equation $\mathcal{A}u = f$ is equivalent to the system

$$\begin{aligned} u_1 \in V_1 &: \mathcal{A}_1 u_1 + \lambda_1'\mathcal{B}_1(\lambda_1 u_1 - \lambda_2 u_2) + \alpha\lambda_1'\mathcal{B}_2(w) = f_1 \text{ in } V_1'\ , \\ w &\equiv \alpha\lambda_1 u_1 + (1-\alpha)\lambda_2 u_2 \in W_0^{1,p}(\Gamma)\ , \\ u_2 \in V_2 &: \mathcal{A}_2 u_2 - \lambda_2'\mathcal{B}_1(\lambda_1 u_1 - \lambda_2 u_2) + (1-\alpha)\lambda_2'\mathcal{B}_2(w) = f_2 \text{ in } V_2'\ . \end{aligned}$$

From the abstract Green's theorem we find that this system is characterized by

$$\text{(5.14.a)} \quad A_j(u_j) = F_j \text{ in } L^{p'}(G_j)\ , \quad \partial_{A_j} u_j = g_j \text{ in } L^{p'}(\partial G_j \sim \Gamma)\ , \quad j = 1, 2\ ,$$

$$\text{(5.14.b)} \quad \begin{aligned} &\partial_{A_1} u_1 + b_1(\gamma_1 u_1 - \gamma_2 u_2) + \alpha B_2(w) = g_1 \text{ and} \\ &\partial_{A_2} u_2 - b_1(\gamma_1 u_1 - \gamma_2 u_2) + (1-\alpha)B_2(w) = g_2\ , \text{ in } L^{p'}(\Gamma)\ , \end{aligned}$$

$$\text{(5.14.c)} \quad w = \alpha\gamma_1 u_1 + (1-\alpha)\gamma_2 u_2)\ , \text{ in } W_0^{1,p}(\Gamma)\ .$$

The pair of equations in $L^{p'}(\Gamma)$ can be written in the equivalent form

$$(5.14.b') \qquad B_2(w) = -\partial_{A_1} u_1 - \partial_{A_2} u_2 + g_1 + g_2 \ ,$$
$$(1-\alpha)[\partial_{A_1} u_1 - g_1 + b_1(\gamma_1 u_1 - \gamma_2 u_2)] = \alpha[\partial_{A_2} u_2 - g_2 - b_1(\gamma_1 u_1 - \gamma_2 u_2)] \ ,$$

and this alternate form is useful for the applications.

The system (5.14) occurs as a model of diffusion in two regions which are connected by a narrow fissure Γ along their common boundary. Thus, suppose we have two regions G_1, G_2 of different materials and assume the fraction of each material in the fissure varies linearly with respect to the distance from the region. Let $\alpha \in [0,1]$ denote the location of the interface Γ as a function of the fraction of the width from G_2, so that $\alpha = 0, 1$ correspond respectively to ∂G_2 and ∂G_1. We assume the concentration in the fracture varies linearly with the transverse distance, so the concentration w on Γ is given by (5.14.c). The diffusion in the regions G_1, G_2 and on their boundaries away from Γ is described by (5.14.a). The flow in the fissure system is modelled by a transverse component and a tangential flow along Γ. The first equation of (5.14.b$'$) gives the tangential flow field by the total flux coming into the fissure along ∂G_1 and ∂G_2. According to (5.14.c), the transverse gradient

$$\frac{1}{1-\alpha}(\gamma_1 u_1 - w) = \frac{1}{\alpha}(w - \gamma_2 u_2) = \gamma_1 u_1 - \gamma_2 u_2$$

is independent of α, $0 < \alpha < 1$. That part of the transverse flux induced in the fissure by this concentration difference is $b_1(\gamma_1 u_1 - \gamma_2 u_2)$; the second equation of (5.14b$'$) fixes the proportion of total flux that enters the fissure from the two sides.

II.6. Variational Inequalities and Quasimonotone Operators

Our objectives here are to illustrate some applications of Theorem 2.3. We begin by constructing a variety of examples of variational inequalities that can be resolved in the form (2.2). Then we show that the class of pseudo-monotone operators contains many quasilinear elliptic equations which are monotone only in their principle part and for which the remaining lower order terms comprise a compact perturbation. This provides a substantial extension of the problems for which existence results can be obtained by the methods of Section 2.

We first describe some examples of variational inequalities for elliptic boundary-value problems. For simplicity we shall take the special class of operators $\mathcal{A} : V \to V'$ given by (5.3); all of the remarks below apply as well to the more general classes of *monotone* operators as given in Proposition 5.1 and the *quasimonotone* case described below. Thus we consider here

$$(6.1.a) \qquad \mathcal{A}u(v) = \int_G \Big\{ \sum_{j=1}^N a(\partial_j u(x)) \partial_j v(x) + b(x) a(u(x)) v(x) \Big\}\, dx + \int_{\partial G} c(s) a(\gamma u(s)) \gamma v(s)\, ds \ , \qquad u, v \in V \ ,$$

where V is a closed subspace of $W^{1,p}(G)$, $1 < p < \infty$, G is a bounded domain in $\mathbb{R}^N$, and $a(\xi) = |\xi|^{p-2}\xi$. Let $F \in L^{p'}(G)$, $F_0 \in L^{p'}(\partial G)$ and define $f \in V'$ by

$$(6.1.b) \qquad f(v) = \int_G F(x) v(x)\, dx + \int_{\partial G} F_0(s) \gamma v(s)\, ds \ , \qquad v \in V \ .$$

Here $\gamma : V \to L^p(\partial G)$ is the trace operator, and its range is denoted by $B = \gamma[V]$. For a given closed, convex, non-empty subset K of V we consider the variational inequality (2.2). Note that if $K = V$ this is just the equation $\mathcal{A}(u) = f$ in V', and if K is a *cone* (i.e., $t(x+y) \in K$ for $x, y \in K$, $t \geq 0$), then (2.2) is equivalent to

$$u \in K : \mathcal{A}u(v) \geq f(v) \ , \quad v \in K \ , \quad \text{and} \quad \mathcal{A}u(u) = f(u) \ . \tag{6.2}$$

This follows from (2.2) by replacing v with $v + u$ and then setting $v = 0$, and it easily implies (2.2) by subtraction. This characterization of solutions of (2.2) in the case of a cone will be useful below.

EXAMPLE 6.A. BOUNDARY CONSTRAINT. Consider the subset of $L^{p'}(\partial G)$ given by $C \equiv \{\psi \in B : \psi(s) \geq 0, \text{ a.e. } s \in \partial G\}$. This set is closed and convex, and, since γ is continuous and linear, it follows that $K \equiv \{v \in V : \gamma v \in C\}$ is also closed and convex. Furthermore, C and K are cones, so any solution of (2.2) is also a solution of (6.2). To characterize such a solution u, we first set $v = \pm\varphi$ in (6.2), $\varphi \in C_0^\infty(G)$, to obtain

$$u \in V : Au = F \text{ in } L^{p'}(G) \tag{6.3.a}$$

where A is the formal part of $\mathcal{A}$ given by

$$Au = -\sum_{j=1}^{N} \partial_j a(\partial_j u) + b(x)a(u) \ .$$

Thus we can apply the abstract Green's formula (5.7) to obtain from (6.2)

$$\gamma u \in B \ , \langle \partial_A u - F_0, \psi \rangle \geq 0 \text{ for } \psi \in C \ , \quad \text{and} \quad (\partial_A u - F_0)\gamma u = 0 \ . \tag{6.3.b}$$

That is, u is a solution of the elliptic equation (6.3.a) which satisfies in a generalized sense the boundary conditions

$$\begin{cases} \gamma u \geq 0 \ , \ \text{a.e. in } L^p(\partial G) \ , \\ \partial_A u \geq F_0 \ \text{a.e. in } L^{p'}(\partial G) \ , \ \text{and} \\ (\partial_A u - F_0)\gamma u = 0 \ \text{a.e. in } L^1(\partial G) \ . \end{cases} \tag{6.3.b$'$}$$

These conditions hold if $\partial_A u \in L^{p'}(\partial G)$, whereas we have in general only $\partial_A u \in B'$. If u is sufficiently smooth then this is the case and we have $\partial_A u$ given by

$$\partial_A u(s) = \sum_{j=1}^{n} a(\partial_j u)\nu_j + c(s)a(u(s)) \ , \ \text{a.e. } s \in \partial G \ ,$$

as in Section 5. These pointwise constraints of (6.3.b$'$) are meaningful if we can deduce from regularity theory for solutions of (6.3.a) that $u \in W^{2,p}(G)$. However, (6.3.b) is meaningful without any such additional information on the solution u, and it is equivalent to

$$\gamma u \geq 0 \text{ in } B \ , \quad \partial_A u - F_0 \geq 0 \text{ in } B' \ , \quad \text{and} \quad \langle \partial_A u - F_0, u \rangle = 0 \ ,$$

where the cone C gives the ordering on B and in B'.

The classical form of (6.3) is known as the *Signorini problem* of elasticity. This describes the equilibrium position of an elastic body which is supported at its boundary by a rigid frictionless constraint surface corresponding to $u = 0$. The boundary ∂G is partitioned *by the solution* into two subregions: Γ_0, on which $\gamma u = 0$

and $\partial_A u \geq F_0$, and Γ_+, on which $\gamma u > 0$ and $\partial_A u = F_0$. The functional $\partial_A u - F_0$ is just the force exerted by the constraint on the body; this force is active only on that subregion Γ_0. If Γ_0 or Γ_+ were known, the solution could be more easily obtained by solving the corresponding mixed Dirichlet-Neumann boundary value problem. But finding this unknown Γ_0 is the essential problem here, analogous to various problems with a "free boundary."

Problems with constraint (only) in the boundary conditions can be characterized rather elegantly in the abstract situation of Proposition 5.3. Thus, assume we have a generalized trace $\gamma : V \to B$ with kernel V_0 and range B, a seminorm $|v|_{\mathcal{X}} = m_0(v,v)^{1/2} + |v|$ given by a continuous seminorm and semi-scalar-product on V, so that $\mathcal{X} \equiv \{V, |\cdot|_{\mathcal{X}}\}$ has dual $\mathcal{X}' \subset V'$, and that V_0 is dense in $\mathcal{X}$, so that we can identify the pivot space $\mathcal{X}' \hookrightarrow V_0'$ by restriction. Let $M_1 : B \to B'$ be given and set

$$\mathcal{M}u(v) = m_0(u,v) + M_1(\gamma u)(\gamma v) \ , \qquad u, v \in V$$

to obtain $\mathcal{M} = M + \gamma' M_1 \gamma$, where $M : V \to V_0'$ is the formal operator obtained from $\mathcal{M}$. Let $\mathcal{A} : V \to V'$ be given, denote its formal part by Au, $Au = \mathcal{A}u|_{V_0}$, and then we have the abstract Green's formula

$$\mathcal{A}u(v) = Au(v) + \partial_A u(\gamma v) \ , \qquad u \in D \ , \ v \in V \ ,$$

where $D = \{u \in V : A(u) \in \mathcal{X}'\}$ is the domain of the dual boundary operator $\partial_A : D \to B'$.

Assume we have $F \in \mathcal{X}'$ and $g \in B'$ given; define $f \in V'$ by $f(v) = F(v) + g(\gamma v)$, $v \in V$. Let C be a closed convex set in B for which $K \equiv \{v \in V : \gamma v \in C\}$ is also non-empty. Consider a solution of the variational inequality

$$u \in K : (\lambda \mathcal{M} u + \mathcal{A} u)(v - u) \geq f(v - u) \ , \qquad v \in K \ . \tag{6.4}$$

Since $K + V_0 = K$ and γ maps K onto C, it follows easily that this is equivalent to

$$u \in V : \lambda M u + Au = F \text{ in } \mathcal{X}' \tag{6.4.a}$$

$$\gamma u \in C, (\lambda M_1 u + \partial_A u)(\psi - \gamma u) \geq g(\psi - \gamma u) \ , \qquad \psi \in C \ . \tag{6.4.b}$$

That is, $u \in D$ and the variational inequality on V' is naturally decoupled into an equation in $\mathcal{X}'$ and a variational inequality on B'.

Finally, we note that the non-homogeneous Dirichlet problem is obtained by setting $C = \{g_0\}$ for a given boundary condition $g_0 \in B$. More generally, we get the mixed Dirichlet-Neumann problem by setting $C = \{\psi \in B : \psi = g_0 \text{ a.e. on } \Gamma_0\}$ for a given $\Gamma_0 \subset \partial G$.

EXAMPLE 6.B. INTERIOR CONSTRAINT. In order to simplify the situation, we choose hereafter $V = W_0^{1,p}(G)$ and set $c = 0$, $F_0 = 0$ in (6.4). Thus we obtain null Dirichlet conditions on ∂G. Define the closed cone

$$K = \{v \in W_0^{1,p}(G) : v(x) \geq 0 \ , \ \text{ a.e. } x \in G\}$$

and consider (6.2). A solution is characterized by

$$u \in W_0^{1,p}(G) : u \geq 0 \ , \quad A(u) - F \geq 0 \ , \quad (A(u) - F)u = 0 \ . \tag{6.5}$$

The first inequality holds a.e. in G and the second in the dual space V'. The domain G is partitioned by the solution u into two regions

$$G_0 = \{x \in G : u(x) = 0\} \ , \quad G_+ = \{x \in G : u(x) > 0\}$$

and we have (formally)

$$(6.6) \qquad Au = F \text{ in } G_+ \ , \quad \gamma u = 0 \text{ on } \partial G_+ \ , \quad \partial_A u = 0 \text{ on } \partial G_+ \cap \overline{G_0} \ .$$

If G_+ is known then u can be obtained from the Dirichlet problem implicit in (6.6). However the difficulty is to find such a G_+ for which the third condition in (6.9) holds. Thus the unknown boundary, ∂G_+, is to be determined from the *pair* of boundary conditions in (6.6).

Such a problem arises in elasticity theory to describe the equilibrium position of a membrane ($N = 2$) which is loaded by a distributed force $F(x)$ and constrained below by an obstacle. According to (6.5) we see the vertical force exerted *upward* by the constraint is $Au - F \geq 0$, and this is non-zero only on G_0, the set on which the constraint is in contact with the membrane. The essential difficulties in this problem involve the regularity of the solution and of the *free surface* which separates G_0 and G_+.

EXAMPLE 6.C. GRADIENT CONSTRAINT. With the same space V and operator $\mathcal{A}$ as above, we now consider the closed, convex and *bounded* set

$$K = \{v \in V : \|\vec{\nabla} v(x)\|_{\mathbb{R}^n} \leq 1 \ , \ \text{a.e. } x \in G\} \ .$$

A solution u of (2.2) corresponds formally to the two sets

$$G_0 = \{x \in G : \|\vec{\nabla} u(x)\| < 1\} \ , \quad G_1 = \{x \in G : \|\vec{\nabla} u(x)\| = 1\}$$

and we have $Au = F$ in G_0 with both of u, $\vec{\nabla} u$ continuous across the "free surface" which separates G_0 and G_1. This corresponds to the *elastic-plastic torsion problem* in mechanics in which the stress $\vec{\nabla} u$ is pointwise bounded. The domain G is divided into the *elastic region* G_0 where the usual equilibrium equation holds, and a *plastic* region in which the stress is at the threshold at which the material begins to flow. That is, $\|\vec{\nabla} u(x)\| > 1$ in the plastic region corresponds to "yielding" of the material to flow. A solution is clearly Lipschitz continuous. The regularity problem is to show that $\vec{\nabla} u$ is smooth across the free surface.

EXAMPLE 6.D. GLOBAL CONSTRAINT. All of the examples above have involved constraints that are applied pointwise, i.e., locally. Here we consider the closed cone

$$K = \left\{ v \in W_0^{1,p}(G) : \int_G v \, dx \geq 0 \right\}$$

and consider a solution of (6.2). In order to characterize a solution we consider the following.

LEMMA 6.1. *Let $\mathbb{B}$ be a Banach space, $f \in \mathbb{B}'$ and set $C = \{v \in \mathbb{B} : f(v) \geq 0\}$. Then*

$$g \in \mathbb{B}' : g(v) \geq 0 \ , \qquad v \in C$$

if and only if $g = cf$ for some $c \geq 0$.

PROOF. If $v \in \ker(f)$, then $\pm v \in C$ and thus $g(v) = 0$. Thus $\ker(f) \subset \ker(g)$ and the result follows since these kernels are hyperplanes in $\mathbb{B}$. □

From Lemma 6.1 we see that a solution of (6.5) is characterized by

$$u \in W_0^{1,p}(G) : \int_G u \geq 0 \ , \quad A(u) - F = c \geq 0 \ , \quad c \cdot \int_G u = 0 \ . \tag{6.7}$$

(This could be shown directly by choosing $v = \partial_j \varphi$, $\varphi \in C_0^\infty(G)$, $1 \leq j \leq n$, since then such a $v \in C$.) Such a problem in elasticity would correspond to the equilibrium displacement u of a membrane ($N = 2$) which is loaded by a distributed force $F(x)$ and constrained below by a minimum volume, such as that of an enclosed incompressible fluid. The *uniform* constraint load c of the fluid on the membrane is non-negative, and it is necessarily zero unless the constraint is attained, i.e., $\int_G u = 0$.

Systems of variational inequalities to which our theory immediately applies can be constructed as in Section 5. Specifically, in the situation of Proposition 5.4, the variational inequality (2.2) is equivalent to a system of the form (5.10). If the constraint set is merely a product, $K = K_1 \times K_2$, with K_j convex in V_j, $j = 1, 2$, this system has the structure of a pair of inequalities on the respective spaces. However, more interesting convex sets in the product are obtained from constraints on the *pair* of components. A good illustration is the following.

LEMMA 6.2. *Let $C = \{v = [v_1, v_2] \in \mathbb{R}^2 : v_1 \geq v_2\}$. Then for $u, f \in \mathbb{R}^2$,*

$$u \in C : f(v) \geq 0 \ , \qquad v \in C$$

is equivalent to

$$u_1 - u_2 \geq 0 \ , \quad f_1 + f_2 = 0 \ , f_1 \geq 0 \ , \quad f_1(u_1 - u_2) = 0 \ .$$

Thus we have either $u_1 = u_2$ and $f_1 = -f_2 \geq 0$, or else $u_1 > u_2$ and $f_1 = f_2 = 0$.

This result follows from the observation that $f(v) = f_1(v_1 - v_2) + (f_1 + f_2)v_2$ with arbitrary $v_2 \in \mathbb{R}$.

If we choose our operators, spaces and data as in (5.11) and set

$$K = \{v \in W^{1,p}(G_1) \times W^{1,p}(G_2) : \gamma_1 v_1 \geq \gamma_2 v_2 \text{ in } L^p(\Gamma)\} \ ,$$

then the variational inequality (6.2) is equivalent to

$$A_j(u_j) = F_j \text{ in } L^{p'}(G_j) \ , \ \partial_{A_j} u_j = g_j \text{ in } L^{p'}(\partial G_j \sim \Gamma) \ , \qquad j = 1, 2 \ , \tag{6.8.a}$$

$$\partial_{A_1} u_1 + \partial_{A_2} u_2 = g_1 + g_2 \text{ in } L^{p'}(\Gamma) \tag{6.8.b}$$

$$\begin{aligned} &\gamma_1 u_1 \geq \gamma_2 u_2 \ , \ \partial_{A_1} u_1 + b(\gamma_1 u_1 - \gamma_2 u_2) - g_1 \geq 0 \ , \\ &\langle \partial_{A_1} u_1 + b(\gamma_1 u_1 - \gamma_2 u_2) - g_2 \ , \ \gamma_1 u_1 - \gamma_2 u_2 \rangle = 0 \ . \end{aligned} \tag{6.8.c}$$

Such a problem corresponds to the displacements of a pair of elastic membranes on adjoining regions for which the first is constrained above the second along the common boundary Γ by the force $\partial_{A_1} u_1 + b(\gamma_1 u_1 - \gamma_2 u_2) - g_1$ between them. This force is non-zero only on the subset of Γ where $\gamma_1 u_1 = \gamma_2 u_2$, where they are in contact.

Similarly if we choose our operators, spaces and data as in (5.12) and set

$$K = \{v \in W_0^{1,p}(G) \times W_0^{1,p}(G) : v_1 \geq v_2 \text{ in } L^{p'}(G)\} \ ,$$

then (6.2) is formally equivalent to

$$\text{(6.9.a)} \qquad u_1, u_2 \in W_0^{1,p}(G) : A_1 u_1 + A_2 u_2 = F_1 + F_2 \text{ in } L^{p'}(G) ,$$

$$\text{(6.9.b)} \qquad u_1 \geq u_2 \ , \ A_1 u_1 \lambda_1' b(u_1 - u_2) \geq F_1 \ ,$$
$$\langle A_1 u_1 + \lambda_1' b(u_1 - u_2) - F_1 \ , u_1 - u_2 \rangle = 0 \ .$$

Such a problem corresponds to the description of a pair of parallel membranes for which the first is bounded below by the second. The force of the second on the first is $A_1 u_1 + \lambda' b(u_1 - u_2) - F_1$; it is non-negative everywhere and it is non-zero only where $u_1 = u_2$, i.e., where the membranes are in contact.

We next construct a class of nonlinear elliptic operators in divergence form which are monotone only in those terms in the principle part, i.e., the highest order terms. The remaining terms must be of lower order, and this is characterized by a compactness assumption.

Let G be bounded domain in $\mathbb{R}^N$, V a closed subspace of $W^{1,p}(G)$ which contains $W_0^{1,p}(G)$, $1 < p < \infty$, and assume the imbedding $V \hookrightarrow L^p(G)$ is compact. Assume given the functions $a_j : G \times \mathbb{R} \times \mathbb{R}^N \to \mathbb{R}$ for $0 \leq j \leq N$ which satisfy

(P_1) $\quad a_j(x, \eta, \xi)$ is measurable in $x \in G$ and continuous in $(\eta, \xi) \in \mathbb{R} \times \mathbb{R}^N$,

(P_2) $\quad |a_j(x,\eta,\xi)| \leq c\big(k(x) + |\eta|^{p-1} + \|\xi\|^{p-1}\big)$, a.e. $x \in G$, $\eta \in \mathbb{R}$, $\xi \in \mathbb{R}^N$,

where $k \in L^{p'}(G)$, $1/p + 1/p' = 1$. Define $\mathcal{A} : V \times V \to V'$ by

$$\text{(6.10)} \qquad \mathcal{A}(u,v)(w) = \mathcal{A}_1(u,v)(w) + \mathcal{A}_0(u)(w) \ ,$$
$$\mathcal{A}_1(u,v)(w) = \int_G \left\{ \sum_{j=1}^N a_j(x, u(x), \vec{\nabla} v(x)) \partial_j w(x) \right\} dx \ ,$$
$$\mathcal{A}_0(u)(w) = \int_G a_0(x, u(x), \vec{\nabla} u(x)) w(x) \, dx \ , \quad u, v, w \in V \ .$$

This operator satisfies the estimate

$$\|\mathcal{A}(u,v)\|_{V'} \leq C\{\|k\|_{L^{p'}} + \|u\|_V^{p-1} + \|v\|_V^{p-1}\} \ ,$$

so it follows from Nemytskii's Theorem 3.2 that $\mathcal{A}$ is continuous. Assume also that

$$(P_3) \qquad \sum_{j=1}^N (a_j(x,\eta,\xi) - a_j(x,\eta,\tilde{\xi}))(\xi_j - \tilde{\xi}_j) > 0$$
$$\text{for a.e. } x \in G \ , \ \eta \in \mathbb{R} \ , \ \xi \neq \tilde{\xi} \in \mathbb{R}^N \ ,$$

$$(P_4) \qquad \frac{\sum_{j=1}^N a_j(x,\eta,\xi)\xi_j}{\|\xi\| + \|\xi\|^{p-1}} \to +\infty \text{ as } \|\xi\| \to +\infty,$$
$$\text{uniformly for } \eta \text{ bounded, at a.e. } x \in G \ .$$

From P_3 it follows that $\mathcal{A}(u,v)$ is monotone in v:

$$\langle \mathcal{A}_1(u, v_1) - \mathcal{A}_1(u, v_2), v_1 - v_2 \rangle \geq 0 \ , \qquad u, v_1, v_2 \in V \ .$$

Our objective is to examine the dependence of $\mathcal{A}(u,u)$ on u with respect to *weak convergence*.

First consider $u_n \rightharpoonup u$ and $w_n \rightharpoonup w$ in V and a fixed $v \in V$. Then we have $u_n \to u$ and $w_n \to w$ while $\partial_j u_n \rightharpoonup \partial_j u$ and $\partial_j w_n \rightharpoonup \partial_j w$ in $L^p(G)$. Thus each of $a_j(\cdot, u_n, \vec{\nabla} v) \to a_j(\cdot, u, \nabla v)$ in $L^{p'}(G)$, so

$$\lim_{n\to\infty} \mathcal{A}_1(u_n, v)w_n = \mathcal{A}_1(u,v)w \ .$$

That is, the first part or "higher order" term is harmless.

Consider the second part, $\mathcal{A}_0$. If $w = 0$, it follows that $\mathcal{A}_0(u_n)w_n \to 0$ since $\{\mathcal{A}_0(u_n)\}$ is bounded in $L^{p'}$. Specifically we have $\lim \mathcal{A}(u_n, v)(u_n - u) = 0$, and this leads to the following result:

if $u_n \rightharpoonup u$ in V, $v \in V$, and $\mathcal{A}(u_n, v) \rightharpoonup f$ in V', then $\mathcal{A}(u_n, v)(u_n) \to f(u)$.

Assume further that $\limsup_{n\to\infty} \mathcal{A}(u_n, u_n)(u_n - u) \le 0$, as suggested by the definition of pseudomonotone. From above we have $\lim \mathcal{A}(u_n, u)(u_n - u) = 0$ so we obtain

$$\begin{aligned} \limsup \langle \mathcal{A}(u_n,u_n) - \mathcal{A}(u_n,u), u_n - u\rangle &= \\ \lim \langle \mathcal{A}(u_n,u_n) - \mathcal{A}(u_n,u), u_n - u\rangle &= \\ \lim_{n\to\infty} \int_G F_n(x)\,dx &= 0 \end{aligned}$$

because the integrands satisfy

$$\begin{aligned} F_n(x) &\equiv \sum_{j=1}^{N} \Big(a_j\big(x, u_n(x), \vec{\nabla} u_n(x)\big) - a_j\big(x, u_n(x), \vec{\nabla} u(x)\big)\Big)\big(\partial_j u_n(x) - \partial_j u(x)\big) \\ &\ge 0, \ \text{a.e. } x \in G \ , \end{aligned}$$

by P_3. Thus $F_n \to 0$ in $L^1(G)$ and by passing to a subsequence we obtain $F_n(x) \to 0$ at a.e. $x \in G$. Pick an $x \in G$ for which we have P_2, P_3, P_4, $u_n(x) \to u(x)$ and $F_n(x) \to 0$. From P_2 we have

$$F_n(x) \ge \sum_{j=1}^{N} a_j\big(x, u_n(x), \nabla u_n(x)\big)\partial_j u_n(x) - C_x\big(1 + \|\vec{\nabla} u_n(x)\|_{\mathbb{R}^N}^{p-1} + \|\vec{\nabla} u_n(x)\|_{\mathbb{R}^N}\big)$$

and by P_4 it follows that $\|\vec{\nabla} u_n(x)\|_{\mathbb{R}^N}$ is bounded. Let ξ be an accumulation point of $\{\vec{\nabla} u_n(x)\}$. Then in the limit we obtain

$$\sum_{j=1}^{N} \Big(a_j\big(x, u(x), \xi\big) - a_j\big(x, u(x), \vec{\nabla} u(x)\big)\Big)\big(\xi_j - \partial_j u(x)\big) = 0 \ ,$$

so P_3 shows that $\xi = \vec{\nabla} u(x)$. Applying the preceding argument to an arbitrary subsequence shows that $\vec{\nabla} u_n(x) \to \vec{\nabla} u(x)$. From P_1 it follows that

$$a_0\big(x, u_n(x), \vec{\nabla} u_n(x)\big) \to a_0\big(x, u(x), \vec{\nabla} u(x)\big) \ , \ \text{a.e. } x \in G \ .$$

Since $\mathcal{A}_0(u_n)$ is bounded in $L^{p'}(G)$, it follows from Proposition 3.4 that $\mathcal{A}_0(u_n) \rightharpoonup \mathcal{A}_0(u)$ in $L^{p'}(G)$. The preceding proves the following.

LEMMA 6.3. *Assume* $u_n \rightharpoonup u$ *and* $w_n \rightharpoonup w$ *in* V, *and let* $\lim_{n\to\infty} \sup \mathcal{A}(u_n, u) \cdot (u_n - u) \le 0$. *Then it follows that*

$$\lim_{n\to\infty} \mathcal{A}(u_n, u_n)(u_n - u) = \lim_{n\to\infty} \mathcal{A}(u_n, u)(u_n - u) = 0$$

and

$$\lim_{n\to\infty} \mathcal{A}(u_n, v)w_n = \mathcal{A}(u, v)w \ , \qquad v \in V \ .$$

THEOREM 6.1 (LERAY-LIONS). *Let* G *be a bounded domain in* $\mathbb{R}^N$, $1 < p < \infty$, *and* V *a closed subspace,* $W_0^{1,p}(G) \subset V \subset W^{1,p}(G)$. *Assume that* $V \hookrightarrow L^p(G)$ *is compact. Let the operator* $\mathcal{A} : V \to V'$ *be given by* $\mathcal{A}(u) = \mathcal{A}(u, u)$ *and* (6.13), *and assume* P_1, P_2, P_3 *and* P_4 *hold. Then* $\mathcal{A}$ *is pseudo-monotone.*

PROOF. Let $u_n \rightharpoonup u$ in V and $\limsup \mathcal{A}(u_n)(u_n - u) \le 0$. Let $v \in V$, $0 < t < 1$, and set $v_t = (1-t)u + tv$. Then P_3 gives

$$\langle \mathcal{A}(u_n, u_n) - \mathcal{A}(u_n, v_t), u_n - v_t \rangle \ge 0$$

and writing $u_n - v_t = u_n - u + t(u - v)$ shows that

$$t\langle \mathcal{A}(u_n) - \mathcal{A}(u_n, v_t), u - v \rangle \ge \langle \mathcal{A}(u_n, v_t) - \mathcal{A}(u_n), u_n - u \rangle$$

and the preceding Lemma shows the right side converges to zero as $n \to \infty$. Take the lim inf, divide by $t > 0$, and then let $t \to 0$ to obtain

$$\liminf \mathcal{A}(u_n)(u_n - v) \ge \mathcal{A}(u)(u - v)$$

as required. □

In the situation of Theorem 6.1, one needs only coercivity to establish the existence of a solution to the corresponding variational inequality or equation. There are various combinations of hypotheses which achieve this; a rather general one is the following.

LEMMA 6.4. *Assume* P_1, P_2 *and*

$$(P_5) \qquad \sum_{j=1}^{N} a_j(x, \eta, \xi)\xi_j \ge c_q \|\xi\|^p - K_q\big(K(x) + |\eta|^q\big) \ ,$$
$$|a_0(x, \eta, \xi)| \le K_q\big(k(x) + |\eta|^{q-1} + \|\xi\|^{q-1}\big)$$

for some q, $1 \le q < p$, *and* $K(\cdot) \in L^1(G)$. *Let* $V \le \{v \in W^{1,p}(G) : \gamma v(s) = 0$, *a.e.* $s \in \Gamma_0\}$, *where* Γ_0 *has positive measure in* ∂G. *Then* $\mathcal{A}$ *is* V*-coercive.*

PROOF. From P_5 we obtain

$$\mathcal{A}(v, v)(v) \ge c_q \|\vec{\nabla} v\|_{L^p}^p - C\{1 + \|v\|_{L^p} + \|v\|_{L^q}^q + \|\vec{\nabla} v\|_{L^q}^q\} \ .$$

From Proposition 5.2 it follows that the first term is equivalent to $\|v\|_{W^{1,p}}^p$, and Young's Lemma 3.3.A shows that for each $\varepsilon > 0$ there is a $C_\varepsilon > 1$ for which

$$|s|^q \le \varepsilon |s|^p + C_\varepsilon \ , \qquad s \in \mathbb{R} \ .$$

These remarks imply that for some $c_0 > 0$, $C_1 > 1$,

$$\mathcal{A}(v, v)(v) \ge c_0 \|v\|_{W^{1,p}}^p - C_1(\|v\|_{L^p} + 1) \ , \qquad u \in V \ . \qquad □$$

REMARK. With the Sobolev embedding Theorem 4.3 the similar result can be obtained with $q \le p + \delta(N)$ for some $\delta(N) > 0$.

Let's characterize the solution of the variational equations for the operator (6.10) as a partial differential equation in G and boundary conditions on ∂G. For this we use the abstract Green's formula (5.7). Thus we choose V to be a subspace of $W^{1,p}(G)$ which contains $V_0 \equiv W_0^{1,p}(G)$ and $B \subset L^p(\partial G)$ to be the range of the trace operator on V, i.e. $\gamma : V \to B$ is surjective. Choose q, $1 \le q \le p$, and $|v|_{\mathcal{X}} = \|v\|_{L^q(G)}$ for $v \in V$. Then $V \hookrightarrow \mathcal{X}$ is continuous and V_0 is dense in $\mathcal{X}$.

Consider $\mathcal{A} : V \to V'$ given by $\mathcal{A}(u) = \mathcal{A}(u,u)$ as above, where $\mathcal{A}(\cdot,\cdot)$ is defined by (6.10). That is,

$$\mathcal{A}(u)(v) = \int_G \left\{ \sum_{j=1}^N a_j(x,u,\vec{\nabla} u)\partial_j v + a_0(x,u,\vec{\nabla} u)v \right\} dx \ , \qquad u,v \in V \ .$$

The formal operator, the restriction to V_0, is given as the generalized function

$$A(u) = \mathcal{A}(u)|_{V_0} = -\sum_{j=1}^N \partial_j a_j(\cdot,u,\vec{\nabla} u) + a_0(\cdot,u,\vec{\nabla} u) \tag{6.11}$$

in V_0'. The domain of the Green's operator is $D \equiv \{u \in V : A(u) \in L^{q'}(G)\}$, and $\partial_A : D \to B' \subset L^{p'}(\partial G)$ satisfies

$$\mathcal{A}u(u) = Au(v) + \partial_A u(\gamma v), \qquad u \in D \ , \ v \in V \ .$$

If we further require that each $a_j(\cdot,u,\vec{\nabla} u) \in W^{1,p}(G)$, ∂G is smooth, and that $u \in W^{2,p}(G)$, respectively, we obtain in succession

$$\begin{aligned}
&\int_G \Big(\partial_j a_j\big(x,u(x),\vec{\nabla} u(x)\big) \cdot v(x) + a_j\big(x,u(x),\vec{\nabla} u(x)\big) \cdot \partial_j v(x) \Big)\, dx \\
&\quad = \int_G \partial_j \big(a_j(\cdot,u,\vec{\nabla} u) \cdot v\big)\, dx = \int_{\partial G} \gamma a_j(\cdot,u,\vec{\nabla} u)\nu_j \gamma(v)\, ds \\
&\quad = \int_{\partial G} a_j\big(s,\gamma u(s),\gamma(\vec{\nabla} u)\big)\nu_j(s)\gamma v(s)\, ds \ .
\end{aligned}$$

Thus the Green's operator is given by

$$\partial_A(u)(s) = \sum_{j=1}^N a_j\big(s,\gamma u(s),\vec{\nabla} u(s)\big)\nu_j(s) \tag{6.12}$$

where each term is in $L^{p'}(\partial G)$. Let's choose $V = \{v \in W^{1,p}(G) : \gamma v = 0 \text{ in } L^p(\Gamma_0)\}$ where Γ_0 has positive measure in ∂G. Denote the remainder of the boundary by $\Gamma_1 = \partial G \sim \Gamma_0$. Thus each term of (6.12) is in $L^{p'}(\Gamma_1)$.

Let the data $F \in L^{q'}(G) = \mathcal{X}'$, $g \in L^{p'}(\Gamma_1) \subset B'$ be given and define

$$f(v) = \int_G F(x)v(x)\, dx + \int_{\Gamma_1} g(s)\gamma v(s)\, ds \ , \qquad v \in V \ .$$

Then $f \in V'$ and we have $u \in V$, $\mathcal{A}(u) = f$, if and only if

$$\begin{aligned}
&u \in W^{1,p}(G) : A(u) = F \text{ in } L^{q'}(G) \ , \\
&\gamma u = 0 \text{ in } L^p(\Gamma_0) \ , \partial_A u = g \text{ in } L^{p'}(\Gamma_1) \ ,
\end{aligned}$$

where A and ∂_A are given by (6.14) and (6.15). The existence of such a solution follows from Theorem 2.3 in the special case of $K = V$. Similar results are immediate for variational inequalities by making appropriate choices for K as in the Examples above.

II.7. Convex Functions

For the case of a continuous and symmetric linear operator $\mathcal{A} : V \to V'$, we saw in Section I.2 that the problem of solving the linear equation $\mathcal{A}u = f$ was equivalent to minimizing the quadratic convex function $\varphi : V \to \mathbb{R}$ given by $\varphi(u) = \frac{1}{2}\mathcal{A}u(u) - f(u)$. This occurs because $\mathcal{A}$ is the derivative of the first term in φ. Here we shall consider a rather general class of convex functions and find that their derivatives are monotone and hemicontinuous functions (when they exist). More important, the *subgradient* $\partial\varphi$ of a convex function φ extends the notion of derivative to non-smooth functions, and it leads us to the notion of *multi-valued* operators or *relations*. These will be very useful in various contexts. Moreover, we shall extend our existence theorems to include the class of operator equations of the form

$$u \in V : \mathcal{A}(u) + \partial\varphi(u) \ni f \quad \text{in} \quad V'$$

where $\mathcal{A}$ is pseudo-monotone.

Let V be a Banach space and $\varphi : V \to \mathbb{R}_\infty \equiv (-\infty, +\infty]$ an extended real-valued function. Then φ is *convex* if

$$\varphi\big(tu + (1-t)v\big) \le t\varphi(u) + (1-t)\varphi(v) \ , \qquad u, v \in V \ , \quad 0 \le t \le 1 \ .$$

It is *proper* if $\varphi(u) < \infty$ for some $u \in V$; its *effective domain* is $\text{dom}(\varphi) = \{u \in V : \varphi(u) < \infty\}$. For any set $S \subset V$ we define the *indicator function* by $I_S(u) = 0$ if $u \in S$, $I_S(u) = +\infty$ if $u \notin S$. Then I_S is convex if and only if S is convex, and I_S is proper if and only if S is non-empty.

A problem of general interest is to minimize a convex function φ on a convex set S. This is equivalent to minimizing $\varphi + I_S$ over the whole space V, so it is useful to permit $\varphi(u) = +\infty$. Suppose we had a convex function φ with $\varphi(u) = -\infty$ at some $u \in V$. Then for any $v \in V$ the convexity of φ shows that we can find a $\bar{t} \in [0,1]$ such that $\varphi = -\infty$ at $tu + (1-t)v$ for $0 \le t < \bar{t}$ and $\varphi = +\infty$ for $\bar{t} < t \le 1$. Such functions are too special to be of interest to us, and they unduly complicate our calculations. Hence, we never allow "$-\infty$" as a value for our functions.

The *epigraph* of $\varphi : V \to \mathbb{R}_\infty$ is given by

$$\text{epi}(\varphi) \equiv \big\{(u, a) \in V \times \mathbb{R} : \varphi(u) \le a\big\} \ .$$

Note that $u \in \text{dom}(\varphi)$ if and only $(u, a) \in \text{epi}(\varphi)$ for some $a \in \mathbb{R}$. We summarize some elementary properties as follows.

PROPOSITION 7.1. *If φ is convex and $\lambda \ge 0$, then $\lambda\varphi$ is convex. If φ_1 and φ_2 are convex, then $\varphi_1 + \varphi_2$ is convex. If each φ_α is convex, $\alpha \in A$, then* $\sup \varphi_\alpha$ *is convex and* $\text{epi}(\sup_{\alpha \in A} \varphi_\alpha) = \cap\{\text{epi}(\varphi_\alpha) : \alpha \in A\}$. *The function $\varphi : V \to \mathbb{R}_\infty$ is convex, proper, and lower-semi-continuous if and only if* $\text{epi}(\varphi)$ *is, respectively, convex, non-empty, and closed in $V \times \mathbb{R}$.*

PROOF. The first three statements are elementary. To prove the fourth, assume φ is convex and let $(u,a),(v,b) \in \text{epi}(\varphi)$. Then $\varphi(u) \le a < \infty$, $\varphi(v) \le b < \infty$ and for all $t \in [0,1]$ we have $\varphi(tu + (1-t)v) \le ta + (1-t)b$, so

$$t(u,a) + (1-t)(v,b) = \big(tu + (1-t)v\ ,\ ta + (1-t)b\big) \in \text{epi}(\varphi)\ .$$

Conversely, if $\text{epi}(\varphi)$ is convex then $\text{dom}(\varphi) = P_V(\text{epi}(\varphi))$ is convex and we need only to consider $u, v \in \text{dom}(\varphi)$. But $\varphi(u) = a < \infty$ and $\varphi(v) = b < \infty$ so from

$$t(u,a) + (1-t)(v,b) = \big(tu + (1-t)v\ ,\ ta + (1-t)b\big) \in \text{epi}(\varphi)$$

we obtain $\varphi(tu + (1-t)v) \le t\varphi(u) + (1-t)\varphi(v)$. The remaining equivalences are elementary. □

The continuity of a convex function follows from local upper-boundedness.

LEMMA 7.1. *If $\varphi : V \to \mathbb{R}_\infty$ is convex and upper-bounded on a neighborhood of a point, then φ is continuous at that point.*

PROOF. By translation we obtain an open sphere $S = \{u \in V : \|u\| < r\}$ such that $\varphi(u) \le 1$ for $u \in S$, $\varphi(0) = 0$. Set $S_\varepsilon = \varepsilon S = \{u \in V : \|u\| < \varepsilon r\}$ with $\varepsilon > 0$. Then for each $u \in S_\varepsilon$ we have $-u/\varepsilon \in S$, so $0 = \varphi(0) = \varphi\big(u/(1+\varepsilon) + \big(1 - 1/(1+\varepsilon)\big)(-u/\varepsilon)\big) \le \varphi(u)/(1+\varepsilon) + \big(1 - 1/(1+\varepsilon)\big)\varphi(-u/\varepsilon)$, hence, $\varphi(u) \ge -\varepsilon\varphi(-u/\varepsilon) \ge -\varepsilon$. Also, $u/\varepsilon \in S$ implies $\varphi(u) \le \varepsilon\varphi(u/\varepsilon) + (1-\varepsilon)\varphi(0) \le \varepsilon$ so we have $|\varphi(u)| \le \varepsilon$ for all $u \in S_\varepsilon$. □

PROPOSITION 7.2. *If the convex $\varphi : V \to \mathbb{R}_\infty$ is upper-bounded on a neighborhood of some point, then φ is continuous at each point of the interior of* $\text{dom}(\varphi)$.

PROOF. Let S be as above and $v \in \text{int}(\text{dom}(\varphi))$. Choose $\rho > 1$ such that $\rho v \in \text{dom}(\varphi)$. Then for each $u \in V$ of the form $u = v + (1 - \frac{1}{\rho})w$ with $w \in S$ we have $u = \frac{1}{\rho}(\rho v) + (1 - \frac{1}{\rho})w$ and so $\varphi(u) \le \frac{1}{\rho}\varphi(\rho v) + (1 - \frac{1}{\rho})$. That is, φ is upper-bounded on a neighborhood of v. □

The proof of Lemma 7.1 shows that φ is actually locally Lipschitz in the $\text{int}(\text{dom}(\varphi))$. Furthermore, it follows that any convex function φ on a finite-dimensional space is continuous on $\text{int}(\text{dom}(\varphi))$. (Take any n-dimensional simplex in $\text{int}(\text{dom}(\varphi))$ and observe that φ is bounded there by its values on the vertices.)

PROPOSITION 7.3. *A proper convex* l.s.c. *φ on a Banach space V is continuous on* $\text{int}(\text{dom}(\varphi))$.

PROOF. We may suppose $0 \in \text{int}(\text{dom}(\varphi))$ and $\varphi(0) < a$ for some $a \in \mathbb{R}$. The level set $S = \{v \in V : \varphi(v) \le a\}$ is closed, convex and so also is $B \equiv S \cap (-S)$. The set B is "balanced": $v \in B$ and $|\lambda| \le 1$ imply $\varphi(\lambda v) = \varphi(\lambda v + (1-\lambda)0) \le a$. For any $v \in V$ there is a $\lambda_0 > 0$ such that $|\lambda| \le \lambda_0$ implies $\lambda v \in \text{int}(\text{dom}(\varphi))$ and by the above remark the restriction of φ to $[-\lambda_0 v,\ \lambda_0 v]$ is continuous. Thus it is continuous at the origin, so there is a $\lambda_v > 0$ such that $0 \le \varphi(\lambda v) - \varphi(0) < a - \varphi(0)$ for $|\lambda| \le \lambda_v$. That is, B is "absorbent": for $v \in V$ there is a $\lambda_v > 0 : |\lambda| \le \lambda_v$ implies $\lambda v \in B$. Such a set ... called a *barrel* ... is necessarily a neighborhood. To see this, let B be a barrel. Since B is absorbing, $V = \cup\{nB : n \ge 1\}$. Since V is a complete metric space, it is not of first category, so some nB, hence, B contains an interior point x_0. If $x_0 = 0$ we are done. Otherwise, $-x_0 \in \operatorname{int}(B)$, since it is balanced, and by convexity we have $0 = 1/2x_0 + 1/2(-x_0)$ in the interior of B. □

The (*directional derivative* of $\varphi : V \to \mathbb{R}_\infty$ at $u \in \operatorname{dom}(\varphi)$ in the direction v is the (one-sided) limit

$$\varphi'(u,v) \equiv \lim_{t \downarrow 0} \frac{1}{t}\big(\varphi(u+tv) - \varphi(u)\big)$$

where it exists. If φ is convex and $u \in \operatorname{dom}(\varphi)$ then for $v \in V$ and $0 < s \le t$ we have $\varphi(u+sv) = \varphi(\frac{s}{t}(u+tv) + (1-\frac{s}{t})u) \le \frac{s}{t}\varphi(u+tv) + (1-\frac{s}{t})\varphi(u)$, hence,

$$\frac{\varphi(u+sv) - \varphi(u)}{s} \le \frac{\varphi(u+tv) - \varphi(u)}{t} .$$

Thus the difference-quotient is monotone in t and necessarily has a limit in $[-\infty, +\infty]$. The *G-differential* of $\varphi : V \to \mathbb{R}_\infty$ at $u \in \operatorname{dom}(\varphi)$ is an $f \in V'$ for which $f(v) = \varphi'(u,v)$ for all $v \in V$. Such an f is unique, it is denoted by $\varphi'(u)$, and we say φ is *G-differentiable* at u.

PROPOSITION 7.4 (KACHUROVSKII). *Let K be convex in V and let $\varphi : V \to \mathbb{R}_\infty$ be G-differentiable at each $u \in K$, $K = \operatorname{dom}(\varphi)$. The following are equivalent:*
(a) *φ is convex,*
(b) *$\varphi'(u)(v-u) \le \varphi(v) - \varphi(u)$ for all $u, v \in K$, and*
(c) *$(\varphi'(u) - \varphi'(v))(u-v) \ge 0$ for all $u, v \in K$.*

PROOF. Since $\varphi(u+t(v-u) \le t\varphi(v) + (1-t)\varphi(u)$ it follows that (a) implies (b). By adding to (b) the corresponding inequality for $\varphi'(v)$ it follows that (b) implies (c). To show that (c) implies (a), let $u, v \in K$ and define $g(t) = \varphi(tu + (1-t)v)$ for $0 \le t \le 1$. Then $g'(t) = \varphi'(v + t(u-v))(u-v)$ and for $0 \le s < t \le 1$ we have

$$\big(g'(t) - g'(s)\big)(t-s) = \Big(\varphi'\big(v + t(u-v)\big) - \varphi'\big(v + s(u-v)\big)\Big)(t-s)(u-v) \ge 0$$

so g' is non-decreasing on $[0,1]$. By the mean-value theorem there follows

$$\frac{g(t) - g(0)}{t - 0} \le \frac{g(1) - g(t)}{1-t} , \qquad 0 < t < 1 ,$$

hence, $g(t) \le tg(1) + (1-t)g(0)$, and this implies (a). □

Note that since g' is a monotone derivative it is necessarily continuous. Thus the proof shows that the derivative of a convex function is monotone and hemicontinuous.

An *affine function* on V is a function given in the form

$$\ell(v) = c + u^*(v) , \qquad v \in V$$

where $(u^*, c) \in V' \times \mathbb{R}$. The graph of ℓ is then

$$\operatorname{gr}(\ell) = \{\ (v,t) \in V \times \mathbb{R} : \ell(v) = t\ \} = \{\ (v,t) \in V \times \mathbb{R} : -u^*(v) + t = c\ \} ,$$

a *hyperplane* in $V \times \mathbb{R}$, given by $h^* = c$ where $h^*(v,t) = -u^*v + 1t = \langle(-u^*, 1)(v,t)\rangle$; it is non-vertical and identified with the *normal* $(-u^*, 1) \in V' \times \mathbb{R}$. The meaning of part (b) of Proposition 7.4 is that the continuous affine function

$$\ell(v) \equiv \varphi(u) + \varphi'(u)(v-u) , \qquad v \in V$$

is everywhere below epi(φ) and coincides with it at $v = u$. That is, the graph of ℓ is a hyperplane of support to epi(φ) at u. In examples we shall see a correspondence between the smoothness of φ and the uniqueness of such supporting hyperplanes.

DEFINITION. Let $\varphi : V \to \mathbb{R}_\infty$ be convex and proper. The *subdifferential* of φ at $u \in \text{dom}(\varphi)$ is the set of all functionals $u^* \in V'$ such that

$$u^*(v - u) \leq \varphi(v) - \varphi(u) \ , \qquad v \in V \ ,$$

and is denoted by $\partial\varphi(u)$. Each such $u^* \in \partial\varphi(u)$ is also called a subdifferential of φ at u, and when $\partial\varphi(u) \neq \varphi$ we say φ is *subdifferentiable* at u.

We consider separately the existence and uniqueness of a subdifferential.

PROPOSITION 7.5. *If $\varphi : V \to \mathbb{R}_\infty$ is convex, proper, and continuous at $u \in V$, then $\partial\varphi(u)$ is closed, convex, bounded, and non-empty.*

PROOF. It is immediate that $\partial\varphi(u)$ is closed and convex, even in general. To show it is bounded, choose $\delta > 0$ so that $\|v - u\| < \delta$ implies $|\varphi(v) - \varphi(u)| \leq 1$. For $u^* \in \partial\varphi(u)$ choose v with $\|v\| = 1$. Then

$$1 \geq \varphi(u + \delta v) - \varphi(u) \geq u^*(\delta v) = \delta u^*(v)$$

so we have $\|u^*\| \leq 1/\delta$.

It remains to show there exists a subgradient at $u \in \text{dom}(\varphi)$. Since φ is continuous at u it follows that epi(φ) has non-empty interior; it contains $S_\delta(u) \times (\varphi(u) + 1, \ +\infty)$ in $V \times \mathbb{R}$. Thus the convex body epi(φ) lies on one side of a hyperplane containing $(u, \varphi(u))$, a point not in the interior of epi(φ). That is, there exist $u^* \in V'$ and $a \in \mathbb{R}$ such that

$$-u^*(v) + at \geq c \ \text{ for all } \ (v,t) \in \text{epi}(\varphi)$$

and $-u^*(u) + a\varphi(u) = c$. Now if $a = 0$ we get $u^*(u - v) \geq 0$ for all $v \in \text{dom}(\varphi)$ and since $u \in \text{int}(\text{dom}(\varphi))$ this means $u^* = 0$, a contradiction. Thus we may assume $a = 1$ above, hence,

$$-u^*(v) + t \geq c = -u^*(u) + \varphi(u) \ \text{ for } \ t \geq \varphi(v) \ ,$$

and this shows with $t = \varphi(v)$ that $u^* \in \partial\varphi(u)$. □

PROPOSITION 7.6. *Let $\varphi : V \to \mathbb{R}_\infty$ be convex and proper. If φ is G-differentiable at $u \in \text{int}(\text{dom}(\varphi))$, then $\partial\varphi(u) = \{\varphi'(u)\}$. If φ is somewhere continuous and $\partial\varphi(u)$ is a singleton, then φ is G-differentiable at u.*

PROOF. Let $u^* \in \partial\varphi(u)$ so $u^*(v - u) \leq \varphi(v) - \varphi(u)$, $v \in V$. Set $v = u + tw$ and let $t \downarrow 0$ to obtain $u^*(w) \leq \varphi'(u)(w)$ for all $w \in V$. Then $u^* = \varphi'(u)$. Since φ is convex we have for each $v \in V$

$$\varphi(u) + t\varphi'(u,v) \leq \varphi(u + tv) \ \text{ for } \ t \in \mathbb{R} \ ,$$

so the affine subset $\{(u + tv, \ \varphi(u) + t\varphi'(u,v)) : t \in \mathbb{R}\}$ in $V \times \mathbb{R}$ is disjoint from the convex, open and non-empty interior of epi(φ). Thus there is a closed hyperplane containing this affine set and disjoint from int(epi(φ)). This hyperplane is the graph of a continuous affine function below epi(φ) and is exact at $(u, \varphi(u))$. The corresponding functional is the unique subdifferential and agrees with $\varphi'(u,v)$ for all $v \in V$. □

We briefly consider some calculus with subgradients. First note that if $\lambda > 0$ then $\partial(\lambda\varphi) = \lambda\partial\varphi$ for the proper convex φ.

PROPOSITION 7.7. *Let φ_1 and φ_2 be convex functions and suppose there is a point in* $\mathrm{dom}(\varphi_1) \cap \mathrm{dom}(\varphi_2)$ *at which φ_1 is continuous. Then*

$$\partial(\varphi_1 + \varphi_2) = \partial\varphi_1 + \partial\varphi_2 \ .$$

PROOF. It is clear that $\partial\varphi_1 + \partial\varphi_2 \subset \partial(\varphi_1 + \varphi_2)$ always holds. Suppose that $u^* \in \partial(\varphi_1 + \varphi_2)(u)$. That is,

$$\varphi_1(v) - \varphi_1(u) - u^*(v - u) \geq \varphi_2(u) - \varphi_2(v) \ , \qquad v \in V \ ,$$

so the two sets

$$E \equiv \{\ (v,t) \in V \times \mathbb{R} : \varphi_1(v) - \varphi_1(u) - u^*(v-u) \leq t \ \} \ ,$$
$$F \equiv \{\ (v,t) \in V \times \mathbb{R} : \varphi_2(u) - \varphi_2(v) \geq t \ \}$$

can have only boundary points in common. Moreover, E is convex with non-empty interior, since it is the epigraph of a convex function somewhere continuous. F is the reflection of $\mathrm{epi}(\varphi_2)$, so F is convex. Thus there is a closed hyperplane which separates E and F. It is non-vertical, since otherwise it would separate $\mathrm{dom}(\varphi_1)$ from $\mathrm{dom}(\varphi_2)$, so there exist $u_2^* \in V'$ and $c \in \mathbb{R}$ such that

$$\varphi_1(v) - \varphi_1(u) - u^*(v-u) \geq -u_2^*(v) + c \geq \varphi_2(u) - \varphi_2(v) \ , \qquad v \in V \ .$$

Setting $v = u$ shows $c = u_2^*(u)$ so we obtain

$$u^* - u_2^* \equiv u_1^* \in \partial\varphi_1(u) \ , \quad u_2^* \in \partial\varphi_2(u) \ , \quad u^* = u_1^* + u_2^*$$

as desired. □

Suppose that $\varphi : W \to \mathbb{R}_\infty$ is convex on the linear space W, that $\Lambda : V \to W$ is linear, and $\mathrm{dom}(\varphi) \cap Rg(\Lambda)$ is non-empty. Then the composite $\varphi \circ \Lambda$ is convex and proper on V. For each $w^* \in \partial\varphi(\Lambda u)$ we have

$$w^*(w - \Lambda u) \leq \varphi(w) - \varphi(\Lambda u) \ , \qquad w \in W \ ,$$

hence, denoting the dual operator by $\Lambda' : W' \to V'$, we have

$$\Lambda' w^*(v - u) = w^*(\Lambda v - \Lambda u) \leq \varphi \circ \Lambda(v) - \varphi \circ \Lambda(u) \ , \qquad v \in V \ .$$

That is, $\Lambda' w^* \in \partial(\varphi \circ \Lambda)(u)$ for each $w^* \in \partial\varphi(\Lambda(u))$:

$$\Lambda' \cdot \partial\varphi\Lambda \subset \partial(\varphi \circ \Lambda) \ .$$

The reverse inclusion and *chain rule* follows by a separation argument.

PROPOSITION 7.8. *Let $\varphi : W \to \mathbb{R}_\infty$ be convex, $\Lambda : V \to W$ be continuous and linear, and assume φ is continuous at some point of $Rg(\Lambda)$. Then $\partial(\varphi \circ \Lambda) = \Lambda' \cdot \partial\varphi \cdot \Lambda$.*

PROOF. (Continued): Let $u^* \in \partial(\varphi \circ \Lambda)(u)$ so

$$u^*(v-u) + \varphi(\Lambda u) \le \varphi(\Lambda v) , \qquad v \in V .$$

The affine set $S = \{(\Lambda v,\ u^*(v-u)+\varphi(\Lambda u)) : v \in V\}$ in $W \times \mathbb{R}$ meets $\mathrm{epi}(\varphi)$ only at boundary points, and $\mathrm{int}(\mathrm{epi}\,\varphi)$ is non-empty, so there is a non-vertical hyperplane $-w^*(w) + t = c$, $w \in W$ determined by $w^* \in W'$ and $c \in \mathbb{R}$, which is disjoint from $\mathrm{epi}(\varphi)$ and contains S. Thus,

$$-w^*(\Lambda v) + u^*(v-u) + \varphi(\Lambda u) = c , \qquad v \in V .$$

Setting $v = u$ shows $c = -w^*(\Lambda u) + \varphi(\Lambda u)$, so

$$w^*\big(\Lambda(v-u)\big) = u^*(v-u) , \qquad v \in V ,$$

and we have $u^* = \Lambda'(w^*)$. Finally, this hyperplane is below $\mathrm{epi}(\varphi)$ so

$$w^*(w - \Lambda u) + \varphi(\Lambda u) \le \varphi(w) , \qquad w \in W .$$

That is, $w^* \in \partial\varphi(\Lambda u)$ as we desired. □

We show with an elementary example that $Rg(\Lambda) \cap \mathrm{dom}(\varphi) \neq \varphi$ is not sufficient for the above. Consider $\varphi(s) = -\sqrt{s}$, $s \ge 0$, and $\varphi(s) = +\infty$ for $s < 0$. If $\Lambda = 0$, $W = V = \mathbb{R}$, then $\partial\varphi(0) = \varphi$ but $\partial(\varphi \circ \Lambda(0)) = \{0\}$. The continuity of $\varphi \circ \Lambda$ was used to show the hyperplane is non-vertical, i.e., we use $\mathrm{int}\,\mathrm{dom}(\varphi \circ \Lambda) \neq \varphi$. Specifically, suppose that $-w^*(w) + at \ge c \ \forall\ \varphi(w) \le t$. If $a = 0$ then we obtain $w^*(\Lambda v - \Lambda u) \le 0 \ \forall\ v \in \mathrm{dom}(\varphi \circ \Lambda)$ and thus $w^* = 0$, a contradiction.

Derivatives arise naturally in minimization problems. Consider a convex set $K \subset V$ and a convex, proper function $\varphi : K \to \mathbb{R}_\infty$ with $\mathrm{dom}(\varphi) \subset K$. Then the minimization problem

$$u \in K : \varphi(u) \le \varphi(v) , \ \text{ all } \ v \in K$$

is equivalent to $0 \in \partial\varphi(u)$. Another useful but less obvious criterion for a minimization is given as follows. Let $\widetilde{K} = \mathrm{epi}(\varphi) \cap (K \times \mathbb{R}) = \{(x,t) \in V \times \mathbb{R} : x \in K,\ \varphi(x) \le t\}$ and $\tilde{\varphi}(x,t) = t$ for $(x,t) \in V \times \mathbb{R}$. Then u minimizes φ on K if and only if

$$u \in K : \varphi(u) \le t \ \text{ for all } \ v \in K \ \text{ and } \ t \ge \varphi(v) ,$$

hence,

$$\big(u, \varphi(u)\big) \in \widetilde{K} : \tilde{\varphi}\big(u, \varphi(u)\big) \le \tilde{\varphi}(v,t) \ \text{ for } \ (v,t) \in \widetilde{K} .$$

Note that $\tilde{\varphi}$ is *differentiable*: $\tilde{\varphi}'((u,t)) = (0,1) \in V' \times \mathbb{R}$.

Let $\mathcal{A} : V \to V'$ and $f \in V'$. Recall that if K is a closed convex non-empty subset of V and $\varphi = I_K$, the indicator function of K, then the variational inequality

$$u \in K : \langle f - \mathcal{A}(u),\ v-u \rangle \le 0 , \qquad v \in K \tag{7.1}$$

is equivalent to

$$u \in V : \langle f - \mathcal{A}(u),\ v-u \rangle \le \varphi(v) - \varphi(u) , \qquad v \in V . \tag{7.2}$$

That is, (7.1) is equivalent to

$$\mathcal{A}(u) + \partial\varphi(u) \ni f .$$

Consider a general convex, proper function φ and let $\widetilde{V} = V \times \mathbb{R}$, $\widetilde{K} = \mathrm{epi}(\varphi)$, $\tilde{f} = (f, 0) \in \widetilde{V}'$ and $\widetilde{\mathcal{A}}(v,t) = (\mathcal{A}v, 1) \in \widetilde{V}'$ for $(u,t) \in \widetilde{V}$. Then $\langle \tilde{f} - \widetilde{\mathcal{A}}\tilde{u},\ \tilde{v} - \tilde{u}\rangle = \langle f - \mathcal{A}(u),\ v - u\rangle - (t - s)$ for

$$\tilde{u} = (u, s)\ , \quad \tilde{v} = (v, t) \in \widetilde{V}\ .$$

If u is a solution of (7.2) then $\tilde{u} \equiv \big(u, \varphi(u)\big) \in \widetilde{K}$ and

$$\langle \tilde{f} - \widetilde{\mathcal{A}}(\tilde{u}),\ \tilde{v} - \tilde{u}\rangle = \langle f - \mathcal{A}(u),\ v - u\rangle - \big(t - \varphi(u)\big) \le 0$$

whenever $t \ge \varphi(v)$, $v \in V$, so

$$\tilde{u} \in \widetilde{K} : \langle \tilde{f} - \widetilde{\mathcal{A}}(\tilde{u}),\ \tilde{v} - \tilde{u}\rangle \le 0\ , \qquad \tilde{v} \in \widetilde{K}\ . \tag{7.3}$$

Conversely, if $\tilde{u} = (u, s)$ is a solution of (7.3), then for all $v \in V$, $t \ge \varphi(v)$

$$\langle f - \mathcal{A}(u),\ v - u\rangle \le t - s \le t - \varphi(u)$$

so (7.2) follows. These remarks prove the following.

LEMMA 7.2. *u is a solution of* (7.2) *if and only if* $(u, \varphi(u))$ *is a solution of* (7.3).

By combining Lemma 7.2 with Theorem 2.3 we obtain the following.

THEOREM 7.1. *Let V be a separable reflexive Banach space, $\mathcal{A} : V \to V'$ be bounded and pseudo-monotone, and $f \in V'$. Let $\varphi : V \to \mathbb{R}_\infty$ be proper, convex, lower semi-continuous, and assume there is a $v_0 \in \mathrm{dom}(\varphi)$ and $R > 0$ such that*

$$\langle \mathcal{A}v - f,\ v - v_0\rangle + \varphi(v) - \varphi(v_0) > 0 \quad \text{for} \quad \|v\| \ge R\ . \tag{7.4}$$

Then there exists a solution of

$$u \in V : \langle f - \mathcal{A}(u),\ v - u\rangle \le \varphi(v) - \varphi(u)\ , \qquad v \in V\ .$$

PROOF. We need only to check that the coercivity condition (7.4) implies the corresponding estimate in Theorem 2.3. Note that with $\tilde{v}_0 \equiv (v_0, \varphi(v_0))$ we have for each $\tilde{v} = (v,t) \in K \equiv \mathrm{epi}(\varphi)$

$$\begin{aligned}\langle \widetilde{\mathcal{A}}\tilde{v} - \tilde{f},\ \tilde{v} - \tilde{v}_0\rangle &= \langle \mathcal{A}v - f,\ v - v_0\rangle + t - \varphi(v_0)\\ &\ge \langle \mathcal{A}v - f,\ v - v_0\rangle + \varphi(v) - \varphi(v_0)\ .\end{aligned}$$

If $\langle \widetilde{\mathcal{A}}(\tilde{v}) - \tilde{f},\ \tilde{v} - \tilde{v}_0\rangle \le 0$, then $(*)$ implies $\|v\| < R$, so

$$\begin{aligned}&t \le \sup_{\|w\| \le R} \langle f - \mathcal{A}(w),\ w - v_0\rangle + \varphi(v_0) \quad \text{and}\\ &t \ge \inf\{\ \varphi(w) : \|w\| \le R\ \}\ ,\end{aligned}$$

i.e., $|t| \le C(R)$. That is, $\langle \widetilde{\mathcal{A}}(\tilde{v}) - \tilde{f},\ \tilde{v} - \tilde{v}_0\rangle \le 0$ implies $\|\tilde{v}\|^2 < R^2 + C(R)^2 \equiv \rho^2$ as desired. □

COROLLARY 7.1. *If there is a $v_0 \in \mathrm{dom}(\varphi)$ such that*

$$\lim_{\|v\| \to \infty} \left(\frac{\mathcal{A}v(v - v_0) + \varphi(v)}{\|v\|} \right) = +\infty\ ,$$

then $Rg(\mathcal{A} + \partial\varphi) = V'$.

II.8. Examples

EXAMPLE 8.A. CONVEX FUNCTIONS ON $\mathbb{R}$.

Let $F : \mathbb{R} \to \overline{\mathbb{R}}$ be a monotone function. It is then continuous at all but a countable number of points. Set $F^-(x) = \lim_{y\to x^-} F(y)$, $F^+(x) = \lim_{y\to x^+} F(y)$, so $F^- = F^+$ a.e. on $\mathrm{dom}(F)$. We then define

$$\varphi(x) = \int_{x_0}^{x} F^-(s)\,ds = \int_{x_0}^{x} F^+(s)\,ds \ , \tag{8.1}$$

where $x_0 \in \mathrm{dom}(F) \subset \mathrm{dom}(\varphi)$ is given. Note that φ is convex and there is an interval $(a,b) \subset \mathrm{dom}(F) \subset \mathrm{dom}(\varphi) \subset [a,b]$. The subgradient is characterized by

$$y \in \partial\varphi(x) \ \text{ if and only if } \ y(\xi - x) \leq \int_x^\xi F^- \ \text{ for } \ \xi \in \mathbb{R}\ .$$

To see this, note first that if $y \leq F^+(x)$ then $x \leq \xi$ implies $y(\xi - x) \leq \int_x^\xi F^+$, and if $y \geq F^-(x)$ then $\xi \leq x \Rightarrow \int_\xi^x F^- \leq y(x-\xi)$ which implies the above. Hence, $y \in [F^-(x),\ F^+(x)]$ implies $y \in \partial\varphi(x)$. The converse follows, since F^- is left-continuous and F^+ is right-continuous, so

$$\partial\varphi(x) = [F^-(x),\ F^+(x)]\ , \qquad x \in \mathbb{R}\ .$$

EXAMPLE 8.B. CONVEX INTEGRANDS.

PROPOSITION 8.1. *Let $\varphi : \mathbb{R} \to \mathbb{R}_\infty$ be proper, convex, lower semi-continuous and either $0 = \varphi(0) = \min(\varphi)$ or the measurable $\Omega \subset \mathbb{R}^n$ has finite measure. Define $\Phi : L^p(\Omega) \to \mathbb{R}_\infty$, $1 \leq p < \infty$, by*

$$\Phi(u) = \int_\Omega \varphi\big(u(x)\big)\,dx \ \ \textit{if} \ \ \varphi(u) \in L^1(\Omega)\ , \quad +\infty \ \textit{otherwise.} \tag{8.2}$$

Then Φ is proper, convex, lower semi-continuous, and $f \in \partial\Phi(u)$ if and only if

$$f \in L^{p'}(\Omega)\ , u \in L^p(\Omega) \ \textit{and} \ f(x) \in \partial\varphi\big(u(x)\big)\ , \ \text{a.e.} \ \ x \in \Omega\ .$$

PROOF. Each set $\{y : \varphi(y) > a\}$ is open and u is measurable, so $\{x : \varphi(u(x)) > a\}$ is measurable for each $a \in \mathbb{R}$, hence $\varphi \circ u$ is measurable. Let's show Φ is LSC. Since φ has an affine lower bound, we may assume it is non-negative. (Otherwise, we add that lower bound to obtain a non-negative φ.) If $u_n \to u$ in L^p and $\Phi(u_n) \leq r$ for $n \geq 1$, there is a subsequence for which (after a change of notation) we have $u_n(x) \to u(x)$ a.e. $x \in \Omega$. Since φ is LSC, $\varphi(u(x)) \leq \lim_{n\to\infty} \inf \varphi(u_n(x))$, a.e. x. Fatou's Lemma shows $\Phi(u) \leq \lim_{n\to\infty} \inf \Phi(u_n) \leq r$.

Suppose $f \in L^{p'}$, $u \in L^p$ and $f(x) \in \partial\varphi(u(x))$, a.e. x. Then $f(x)(v(x) - u(x)) \leq \varphi(v(x)) - \varphi(u(x))$ and $\varphi \circ u \in L^1$; an integration shows $f \in \partial\Phi(u)$. Conversely, let

$$\int_\Omega f(x)\big(v(x) - u(x)\big)\,dx \leq \int_\Omega \Big(\varphi\big(v(x)\big) - \varphi\big(u(x)\big)\Big)\,dx\ , \qquad v \in L^p\ .$$

For each measurable $M \subset \Omega$ we set $w(x) = v(x)$ if $x \in M$ and $w(x) = u(x)$ if $x \in \Omega \sim M$. Then $w \in L^p$ and

$$\int_M \Big\{\ f(x)\big(v(x) - u(x)\big) - \varphi\big(v(x)\big) + \varphi\big(u(x)\big)\ \Big\} \leq 0$$

for each such M, so we have

$$f(x)\big(v(x) - u(x)\big) \le \varphi\big(v(x)\big) - \varphi\big(u(x)\big) \ , \quad \text{a.e.} \quad x \in \Omega \ ,$$

for each $v \in L^p$, so $f(x) \in \partial\varphi(u(x))$, a.e. x. □

EXAMPLE 8.C. BOUNDARY FUNCTIONALS.

Let $1 < p < \infty$, G be a bounded domain in $\mathbb{R}^n$, the boundary ∂G be a C^1 manifold of dimension $n-1$, and denote by γ the trace operator from $W^{1,p}(G)$ into $L^p(\partial G)$. Let $\varphi : \mathbb{R} \to \mathbb{R}$ be convex and continuous and suppose it satisfies

$$|\varphi(s)| \le C(|s|^p + 1) \ , \qquad s \in \mathbb{R} \ .$$

Thus we define

$$\Phi(u) = \int_{\partial G} \varphi\big(\gamma u(s)\big)\, ds \ , \qquad u \in W^{1,p}(G) \ . \tag{8.3}$$

Since Φ is a composite of a convex and a linear function as in Proposition 7.8, the function Φ is convex and continuous, and its subgradient is given by

$$\partial\Phi(u) = \gamma'\big(\partial\varphi(\gamma u)\big) \ , \qquad u \in W^{1,p} \ ,$$

where the dual $\gamma' : L^{p'}(\partial G) \to (W^{1,p})'$ is given by

$$\gamma' g(v) = \int_{\partial G} g(s)\gamma v(s)\, ds \ , \qquad v \in W^{1,p} \ ,$$

for $g \in L^{p'}(\partial G)$. That is, $F \in \partial\Phi(u)$ if and only if there is a $g \in \partial\varphi(\gamma u)$ in $L^{p'}(\partial G)$ for which

$$F(v) = \int_{\partial G} g(s)\gamma v(s)\, ds \ , \qquad v \in W^{1,p} \ .$$

EXAMPLE 8.D. DIRICHLET INTEGRANDS.

For each integer k, $0 \le k \le n$, let there be given a continuous, convex function $\varphi_k : \mathbb{R} \to \mathbb{R}$ which satisfies

$$\varphi_k(s) \le C(|s|^p + 1) \ , \qquad s \in \mathbb{R} \ ,$$

for some p, $1 < p < \infty$. Define

$$\Phi(u) = \sum_{k=0}^{n} \int_G \varphi_k\big(\partial_k u(x)\big)\, dx \ , \qquad u \in W^{1,p}(G) \tag{8.4}$$

where G is a domain in $\mathbb{R}^n$. Note that Φ is a sum of continuous, convex functions on the Sobolev space $W^{1,p}(G)$, each of which is the composition of a convex integrand on $L^p(G)$ following the continuous linear $\partial_k : W^{1,p} \to L^p$. By Propositions 7.7 and 7.8, the subgradient can be computed term-by-term and is given formally by

$$\partial\Phi(u) = \sum_{k=0}^{n} \partial_k'\big(\partial\varphi_k(\partial_k u)\big) \ , \qquad u \in W^{1,p}(G) \ ,$$

where the dual $\partial_k' : L^{p'} \to (W^{1,p})'$ is given by

$$\partial_k' f(v) = \int_G f(x)\partial_k v(x)\, dx \ , \qquad v \in W^{1,p}(G) \ ,$$

for $1 \leq k \leq n$ and $f \in L^{p'}$; $\partial_0' = \partial_0$ is the identity. To be precise, we have $F \in \partial\Phi(u)$ if and only if there exists $f_k \in \partial\varphi_k(\partial_k u)$ in $L^{p'}$ for each k, $0 \leq k \leq n$, for which

$$F(v) = \sum_{k=0}^{n} \int_G f_k(x)\partial_k v(x)\,dx\ , \qquad v \in W^{1,p}(G)\ .$$

By restricting this functional to $V_0 = W_0^{1,p}(G)$ we see the *formal part* is the distribution

$$F_0 = -\sum_{k=1}^{n} \partial_k f_k + f_0 \in V_0'$$

and we denote this by

$$(\partial\Phi)_0(u) = -\sum_{k=1}^{n} \partial_k \partial\varphi_k(\partial_k u) + \partial\varphi_0(u)\ .$$

Note that by the classical Green's Theorem if the boundary of G is smooth and each $\partial_k f_k \in L^{p'}$, $1 \leq k \leq n$, then

$$F(v) - F_0(v) = \int_{\partial G} \left\{ \sum_{k=1}^{n} f_k(s)\nu_k(s) \right\} v(s)\,ds\ , \qquad v \in V\ .$$

As in Section 5 we construct an *abstract Green's Theorem* for $\partial\Phi$ for which

(8.5) $$\partial\Phi(u) = (\partial\Phi)_0(u) + \gamma'(\partial_\Phi(u))\ , \qquad u \in D\ ,$$

where $D = \{u \in W^{1,p}(G) : (\partial\Phi)_0(u) \cap L^{p'}(G) \neq \emptyset\}$ and $\partial_\Phi : D \to B'$ is the corresponding *boundary operator* onto the dual of $B = Rg(\gamma)$. Finally we remark that it is the sum of a Dirichlet integrand and a boundary functional that frequently arises in applications.

Let's consider again the general problem of minimizing the proper, convex $\varphi : V \to \mathbb{R}_\infty$. Note that u is a solution of

$$u \in V : \varphi(u) \leq \varphi(v)\ , \qquad v \in V$$

if and only if $0 \in \partial\varphi(u)$. Also, if K is a convex subset of V which contains a $v_0 \in \mathrm{dom}(\varphi)$, then u is a solution of

$$u \in K : \varphi(u) \leq \varphi(v)\ , \qquad v \in K$$

if and only if u minimizes $\varphi + I_K$ over V, and this is equivalent to $0 \in \partial(\varphi + I_K)(u)$. In addition, if φ is continuous at some $v_0 \in K$ then u minimizes φ on K if and only $0 \in \partial\varphi(u) + \partial I_K(u)$ by Proposition 7.7, and this is equivalent to

$$u \in K\ , \quad f \in \partial\varphi(u)\ , \quad f(v-u) \geq 0\ , \quad \text{all } v \in K\ .$$

Thus, one obtains a characterization of solutions of variational inequalities for multivalued operators (subgradients) as minima of convex functions on the whole space or as minima on the convex set. The following is an easy sufficient condition for the existence of minimizers.

THEOREM 8.1 (WEIERSTRASS). *If K is convex and closed in a reflexive Banach space V, $\varphi : V \to \mathbb{R}_\infty$ is lower-semi-continuous and convex with* $\mathrm{dom}(\varphi) \cap K \neq \emptyset$*, and $\varphi(v) \to \infty$ as $\|v\| \to \infty$ with $v \in K$, then φ attains its infimum on K.*

In most applications the convex functional is identified as the *energy* associated with the state u of a system. As in the preceding examples this is frequently a *local* function of u, i.e., it depends on the values of u or its derivatives at each point in the domain G. Our next example is non-local: the function depends on the "total energy" or "total flux" of the system.

PROPOSITION 8.2. *Let A, B be continuous symmetric monotone linear operators from V to V' and $a, b \in \mathbb{R}$. The function*

$$\varphi(u) = 1/2 \max\{\ Au(u) + a,\ Bu(u) + b\ \}\ , \qquad u \in V\ ,$$

is convex and continuous. Its subgradient is given by

$$\partial\varphi(u) = \begin{cases} \{A(u)\} & \text{if} \quad Au(u) + a > Bu(u) + b, \\ \{\lambda A(u) + (1-\lambda)Bu\ ,\ 0 \le \lambda \le 1\} & \text{if} \quad Au(u) + a = Bu(u) + b, \\ \{B(u)\} & \text{if} \quad Au(u) + a < Bu(u) + b. \end{cases}$$

PROOF. We need only to verify the computations of the subgradient, and the first and last cases follow from G-differentiability of φ in the respective regions. In the middle case we obtain

$$t^{-1}\big(\varphi(u+tv) - \varphi(u)\big) = \max\Big\{\ Au(v) + \frac{t}{2}\ Av(v)\ ,\ Bu(v) + \frac{t}{2}\ Bv(v)\ \Big\}\ ,$$

so we have the equivalence of $f \in \partial\varphi(u)$,

$$f(v) \le t^{-1}\big(\varphi(u+tv) - \varphi(u)\big)\ , \qquad v \in V\ ,\ t > 0\ , \tag{8.6}$$

and of

$$f(v) \le \max\{\ Au(v),\ Bu(v)\ \}\ , \qquad v \in V\ . \tag{8.6'}$$

This last condition is equivalent to $f = \lambda Au + (1-\lambda)Bu$ for some λ, $0 \le \lambda \le 1$. For this we use the following.

LEMMA 8.1. *Let $f, f_1, f_2 \in V^*$. Then f is a linear combination of f_1 and f_2 if and only if the* $\ker(f) \supset \ker(f_1) \cap \ker(f_2)$.

PROOF. The "only if" is clear. Suppose $\ker(f) \supset \ker(f_1)$. If f_1 is not identically zero, $f_1(v_0) \neq 0$, then $f(v) = (f(v_0)/f_1(v_0))f_1(v)$, since $\ker(f_1)$ has codimension 1. That is, f is a multiple of f_1. Suppose $\ker(f) \supset \ker(f_1) \cap \ker(f_2)$. For each $v \in \ker(f_2)$ we have $f(v) = 0$ whenever $f_1(v) = 0$. Thus there is an $a_1 \in \mathbb{R} : f(v) = a_1 f_1(v)$ for $v \in \ker(f_2)$. That is $f - a_1 f_1$ vanishes on $\ker(f_2)$ so as above $f - a_1 f_1 = a_2 f_2$ for some $a_2 \in \mathbb{R}$. □

To finish Proposition 8.2, we note that the condition (8.6′) on f implies

$$\min\{\ Au(v),\ Bu(v)\ \} \le a_1 Au(v) + a_1 Bu(v) \le \max\{\ Au(v),\ Bu(v)\ \}$$

where $f_1 = Au$, $f_2 = Bu$. □

EXAMPLE 8.E. ENERGY-DEPENDENT ELLIPTIC EQUATIONS.

Choose $V = W_0^{1,2}(G)$, $f(v) = \int_G Fv$, $v \in V$, where F is given in $L^2(G)$ and define

$$\varphi(v) = \frac{1}{2}\int_G |\vec{\nabla} v|^2 + \frac{1}{2}\max\left\{1, \int_G |v|^2\right\}, \qquad v \in V . \tag{8.7}$$

Then $\partial\varphi(u) \ni f$ is equivalent to

$$u \in W_0^{1,2}(G) : -\Delta u + \operatorname{sgn}^+\Big(\int_G |u|^2\,dx - 1\Big)u \ni F .$$

EXAMPLE 8.F. FLUX-DEPENDENT EQUATION.

Define V and f as before but set

$$\varphi(v) = \frac{1}{2}\max\left\{1, \int_G |\vec{\nabla} v|^2\right\}, \qquad v \in V . \tag{8.8}$$

Then $\partial\varphi(u) \ni f$ is characterized by

$$u \in W_0^{1,2}(G) : -\operatorname{sgn}^+\Big(\int_G |\nabla u|^2 - 1\Big)\Delta u \ni F .$$

In these examples we have used the real-valued subgradient

$$\operatorname{sgn}^+(r) = \begin{cases} \{0\}, & \text{if } r < 0, \\ [0,1], & \text{if } r = 0, \\ \{1\}, & \text{if } r > 0. \end{cases}$$

This is the *positive sign* or *Heaviside relation*. Also see Proposition 8.6 below.

We consider again the relationship between subgradients and the directional derivative,

$$\varphi'(u,v) = \lim_{t\to 0^+} t^{-1}\big(\varphi(u+tv) - \varphi(u)\big) .$$

This will be used to study the differentiability of the norm and compositions with it. Thus let V be a normed linear space and assume $\varphi : V \to \mathbb{R}_\infty$ is proper and convex.

LEMMA 8.2. *The following are equivalent for an $f \in V'$:*

(a) $f \in \partial\varphi(u)$, *i.e.*, $f(v-u) \le \varphi(v) - \varphi(u)$, $v \in V$,

(b) $f(v-u) \le \varphi'(u, v-u)$, $v \in V$, *and*

(c) $f(v-u) \le t^{-1}(\varphi(u+t(v-u)) - \varphi(u))$, $0 < t \le 1$, $v \in V$.

PROOF. If we set $v = u + t(w-u)$ in (a) we obtain

$$f(w-u) \le t^{-1}\Big(\varphi\big(u+t(w-u)\big) - \varphi(u)\Big) \le \varphi(w) - \varphi(u) , \qquad 0 < t \le 1 .$$

Recall that the difference quotient is monotone in t, so the limit as $t \to 0^+$ exists in $\mathbb{R}_\infty$. □

For the special case of the norm, $\varphi(v) = \|v\|$, on V we obtain yet another equivalent statement,

(d) $\|f\| = 1$ and $f(u) = \|u\|$.

This equivalence is straightforward and will be obtained more generally below, but first we characterize the G-differentiability of the norm. Note by Proposition 7.6 this occurs exactly when the subgradient is a singleton.

DEFINITION. A normed space is *strictly convex* if $u \neq v$, $\|u\| = \|v\| = 1$ and $0 < t < 1$ imply $\|tu + (1-t)v\| < 1$.

PROPOSITION 8.3. *The following are equivalent:*
(a) *V is strictly convex.*
(b) *$u \neq v$, $\|u\| = \|v\| = 1$ imply $\|tu + (1-t)v\| < 1$ for some $t \in (0,1)$.*
(c) *Any convex subset of the unit sphere contains at most one point.*
(d) *Any non-zero $f \in V'$ takes its maximum value on the unit sphere at most once.*

PROOF. That (b) $\Rightarrow$ (a) follows by an easy convexity argument; we check that (a) $\Rightarrow$ (c) $\Rightarrow$ (b), so the first three are equivalent. If $\|u\| = \|v\| = 1$ and $f(u) = f(v) = \|f\| \neq 0$, then for $t \in (0,1)$

$$\|f\| = f\big(tu + (1-t)v\big) \leq \|f\| \, \|tu + (1-t)v\|$$

so $\|tu + (1-t)v\| = 1$ and $u = v$ by (a). Thus (a) $\Rightarrow$ (d). Conversely, if $\|u\| = \|v\| = \|1/2(u+v)\| = 1$, there is an $f \in V' : \|f\| = 1$ and $f(u+v) = 2$. Since $f(u) \leq 1$ and $f(v) \leq 1$ we have $f(u) = f(v)$ so $u = v$ by (d). □

COROLLARY 8.1. *The norm on V is G-differentiable at each $u \neq 0$ if and only if the dual space V' is strictly convex.*

PROOF. Each non-zero $u \in V \subset V''$ attains its maximum on the unit sphere of V' at most once by (d) of Lemma 2 and of Proposition 8.3. □

The next two results will be useful for obtaining estimates on solutions of evolution equations.

PROPOSITION 8.4. *Let $u, v \in V$. Then $\|u + tv\| \geq \|u\|$ for all $t > 0$ if and only if there is an $f \in \partial\varphi(u)$ such that $f(v) \geq 0$, where $\varphi(w) = \|w\|$ above.*

PROOF. Let f be as given. Then $\|f\| \, \|u\| \leq f(u + tv) \leq \|f\| \, \|u + tv\|$ so the result follows if $\|f\| \neq 0$, and it is trivial if $\|f\| = 0$. Conversely, if $\|u\| \leq \|u + tv\|$ for $t > 0$ we select for each $t > 0$ an $f_t \in \partial\varphi(u + tv)$. Then

$$\|u\| \leq \|u + tv\| = f_t(u + tv) = f_t(u) + tf_t(v) \leq \|u\| + tf_t(v) \leq \|u\| + t\|v\| \ ,$$

so we have $f_t(v) \geq 0$, $t > 0$, and $\lim_{t \downarrow 0} f_t(u) = \|u\|$. Since the unit sphere of V' is w^*-compact, there is a subsequence $\{f_{t_n}\}$ w^*-convergent to an $f \in V'$. Then $|f| \leq 1$ and $f(u) = \|u\|$, so $\|f\| = 1$ and $f \in \partial\varphi(u)$ with $f(u) \ni 0$. □

The derivative of the norm of a function is computed as follows.

PROPOSITION 8.5. *If the function $u : [0,T) \to V$ is right-differentiable at $t \in [0,T)$, i.e.,*

$$u^+(t) = \lim_{h \to 0^+} h^{-1}\big(u(t+h) - u(t)\big) \ \ in \ \ V \ ,$$

then so also is $\varphi(u(t)) \equiv \|u(t)\|$ and

$$\frac{d^+}{dt}\|u(t)\| = \varphi'\big(u(t), \ u^+(t)\big) \ .$$

PROOF. For sufficiently small $h > 0$ we have

$$\begin{aligned}\big|\big(\|u(t+h)\| - \|u(t)\|\big) - \big(\|u(t) + hu^+(t)\| - \|u(t)\|\big)\big| \\ = \big|\ \|u(t+h)\| - \|u(t) + hu^+(t)\|\ \big| \\ \leq \|u(t+h) - u(t) - hu^+(t)\|\ .\end{aligned}$$

Divide by $h > 0$ and take the limit as $h \to 0^+$. □

Next we compute the subgradient of a convex function of the norm. See (8.6) and (8.7) for examples.

PROPOSITION 8.6. *Let $\Phi : \mathbb{R}_+ \to \mathbb{R}_+$ be convex, $\Phi(0) = 0$, and V a normed space. Then $f \in V'$ belongs to the subgradient of $\Phi(\|\cdot\|)$ at $u \in V$, i.e.,*

(a) $$f(v - u) \leq \Phi(\|v\|) - \Phi(\|u\|)\ , \qquad v \in V$$

if and only if

(b) $$f(u) = \|f\|\ \|u\| \quad \textit{and} \quad \|f\| \in \partial\Phi(\|u\|)\ .$$

PROOF. Suppose (a) and set $v = \|u\|w$ for any unit vector w. Thus we obtain $\|u\|f(w) \leq f(u)$, $\|w\| = 1$, and thereby the first part of (b). If $u \neq 0$, set $v = (t/\|u\|)u$ in (a) to obtain

$$\|f\|(t - \|u\|) \leq f(v - u) \leq \Phi(t) - \Phi(\|u\|)\ , \qquad t \in \mathbb{R}\ .$$

If $u = 0$ this follows by letting $v = tw$ in (a) with $\|w\| = 1$, so (b) follows from (a). Conversely, (b) implies

$$f(v - u) \leq \|f\|\ (\|v\| - \|u\|) \leq \Phi(\|v\|) - \Phi(\|u\|)\ , \qquad v \in V\ .$$ □

We note that if $\Phi(r) = r$ we obtain part (d) of Lemma 8.2. Also, if $\Phi(r) = \frac{1}{2}r^2$ and V is a Hilbert space with scalar-product $(\cdot, \cdot)$, then f is given by $f(v) = (u, v)$, $v \in V$, so the derivative of $\frac{1}{2}(\|\cdot\|^2)$ is just the *Riesz map* from V to V'.

DEFINITION. For any normed space V the map $J : V \to V'$,

$$J(u) = \{\ f \in V' : f(u) = \|f\|^2 = \|u\|^2\ \}$$

is the *normalized duality map*.

From Proposition 8.6 it follows that $J(u)$ is closed, convex and non-empty. When V' is strictly convex, J is a function (single-valued). More generally, we have the following.

DEFINITION. Let $\xi : \mathbb{R}_+ \to \mathbb{R}_+$ be continuous, monotone, surjective and $\xi(0) = 0$. The multi-valued operator $J_\xi : V \to V'$

$$J_\xi(u) = \{\ f \in V' : f(u) = \|f\|\ \|u\|\ ,\ \|f\| = \xi(\|u\|)\ \}$$

is the corresponding *duality map* with *gauge* ξ.

As before, $J_\xi(u)$ is closed, convex, non-empty and J_ξ is a function when V' is strictly convex. To see this, note that for $f, g \in J_\xi(u)$ and $0 < t < 1$ we have $tf + (1-t)g \in V'$ and

$$\langle tf + (1-t)g, u\rangle = (t + (1-t))\xi(\|u\|)\|u\|\ ,$$

so $J_\xi(u)$ is convex. But by Proposition 8.3.c, it follows that $J_\xi(u)$ is a singleton. Since J_ξ is a G-differential, it is monotone. More precisely, for $u, v \in V$

$$\langle J(u) - J(v),\ u - v\rangle \geq \big(\xi(\|u\|) - \xi(\|v\|)\big)(\|u\| - \|v\|)\ ,$$

and a corresponding estimate holds in the multi-valued case where V' is not strictly convex.

PROPOSITION 8.7. *If V is a reflexive Banach space and V' is strictly convex, then $J_\xi : V \to V'$ is monotone and demicontinuous.*

PROOF. Let $u_n \to u$ in V. Then $\|J_\xi(u_n)\| = \xi(\|u_n\|) \to \xi(\|u\|)$ since ξ is continuous, and we may suppose $J_\xi(u_n) \rightharpoonup f$ in V'. Then $\|f\| \leq \xi(\|u\|)$ and

$$f(u) = \lim J(u_n)(u_n) = \lim \xi(\|u_n\|)\|u_n\| = \xi(\|u\|)\|u\|\ .$$

Finally, $f(u) \leq \|f\|\ \|u\|$, so $\|f\| \geq \xi(\|u\|)$ and we have $f = J_\xi(u)$. □

DEFINITION. A normed space V is *uniformly convex* if for each ε, $0 < \varepsilon < 2$, there exists a $\delta > 0$ such that if $\|u\| \leq 1$, $\|v\| \leq 1$ and $\|u - v\| \geq \varepsilon$, then $\|u + v\| \leq 2(1 - \delta)$.

This is equivalent to requiring that if $\|u_n\| \leq 1$, $\|v_n\| \leq 1$ and $\|u_n + v_n\| \to 2$, then $\|u_n - v_n\| \to 0$. From the parallelogram law it follows that every Hilbert space is uniformly convex.

PROPOSITION 8.8. *If V' is uniformly convex, then J_ξ is uniformly continuous on each bounded set in V. That is, for $\varepsilon > 0$ and $M > 0$ there is a $\delta > 0$ such that if $\|u\|, \|v\| \leq M$ and $\|u - v\| < \delta$, then $\|J_\xi(u) - J_\xi(v)\| < \varepsilon$.*

PROOF. We give the argument for the normalized case, $\xi(s) = s$, but the general case is similarly obtained. Assume there are sequences in V with $\|u_n\| \leq M$, $\|v_n\| \leq M$, and $\|u_n - v_n\| \to 0$. If $u_n \to 0$, then $v_n \to 0$ and we have $\|Ju_n\| = \|u_n\| \to 0$ and $\|Jv_n\| = \|v_n\| \to 0$, so $\|Ju_n - Jv_n\| \to 0$ and we are done.

Hence, we may assume (by passing to subsequences denoted similarly) that $\|u_n\| \geq \alpha > 0$ and $|v_n\| \geq \alpha$. Set $x_n = u_n/\|u_n\|$ and $y_n = v_n/\|v_n\|$ so that $\|x_n\| = \|y_n\| = 1$ and

$$x_n - y_n = (u_n - v_n)/\|u_n\| + (\|u_n\|^{-1} - \|v_n\|^{-1})v_n \to 0\ .$$

Since $\|J(x_n)\| = \|J(y_n)\| = 1$ we have

$$\begin{aligned} 2 &\geq \|J(x_n) + J(y_n)\| \geq \langle Jx_n + Jy_n, x_n\rangle \\ &= \langle Jx_n, x_n\rangle + \langle Jy_n, y_n\rangle + \langle Jy_n, x_n - y_n\rangle \geq 1 + 1 - \|x_n - y_n\|\ , \end{aligned}$$

and this shows $\lim\limits_{n\to\infty} \|J(x_n) + J(y_n)\| = 2$. Since V' is uniformly convex it follows that $\|J(x_n) - J(y_n)\| \to 0$. We also have

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

and this converges to zero. □

We conclude with some examples of duality maps that are useful in certain applications.

$\underline{L^p(\Omega)}$. Let $1 \le p < \infty$ and Ω measurable in $\mathbb{R}^n$. The duality map with gauge $\xi(r) = r^{p-1}$ is the subgradient of the convex function

$$\frac{1}{p}\|v\|_{L^p}^p = \frac{1}{p}\int_\Omega |v(x)|^p\,dx\ , \qquad v \in L^p(\Omega)\ ,$$

and by Proposition 8.1 this is characterized by

$$f \in J_\xi(u) \iff \int_\Omega f(x)u(x)\,dx = \|f\|_{L^{p'}}\|u\|_{L^p} \text{ and } \|f\|_{L^{p'}} = \|u\|_{L^p}^{p-1}\ ,$$
$$\iff f(x) \in |u(x)|^{p-1}\operatorname{sgn}\big(u(x)\big)\ , \quad \text{a.e. } x \in \Omega$$

for $f \in L^{p'}(\Omega)$ and $u \in L^p(\Omega)$.

$\underline{W_0^{1,p}(\Omega)}$. Let $1 < p < \infty$ and G a bounded domain in $\mathbb{R}^n$. Then by Poincaré's Lemma 5.1, $W_0^{1,p}(G)$ is a Banach space with norm

$$\|v\| = \left(\sum_{j=1}^n \|\partial_j v\|_{L^p}^p\right)^{1/p}\ .$$

The duality map with gauge $\xi(r) = r^{p-1}$ is (a subgradient) characterized by

$$f \in J_\xi(u) \iff f(u) = \|f\|_{W^{-1,p'}}\|u\|_{W_0^{1,p}} \text{ and } \|f\|_{W^{-1,p'}} = (\|u\|_{W_0^{1,p}})^{p-1}$$
$$\iff f = -\sum_{j=1}^n \partial_j\big(|\partial_j u|^{p-1}\operatorname{sgn}(\partial_j u)\big) \text{ in } \mathcal{D}^*$$

for $f \in W^{-1,p'} = (W_0^{1,p})'$, see Section 4, and $u \in W_0^{1,p}(G)$.

II.9. Elliptic Equations in L^1

Let G be a bounded domain in $\mathbb{R}^n$ with smooth boundary ∂G. Consider the linear elliptic differential operator given by

$$Au \equiv -\sum_{i,j=1}^n \partial_j(a_{ij}\partial_i u) + \sum_{i=1}^n \partial_i(a_i u) + au$$

where the coefficients satisfy

$$a_{ij}\ ,\quad a_i \in C^1(\bar{G}) \quad\text{and}\quad a \in L^\infty(G)\ ,$$
$$a(x) \ge 0 \quad\text{and}\quad a(x) + \sum_{i=1}^n \partial_i a_i(x) \ge 0\ ,$$
$$\sum_{i,j=1}^n a_{ij}(x)\xi_i\xi_j \ge c_0|\xi|^2\ , \qquad \xi \in \mathbb{R}^n$$

for a.e. $x \in G$ and $c_0 > 0$. The meaning of the *Dirichlet problem* for Au is different in each $L^p(G)$. The appropriate definition of $A_p u = f \in L^p(G)$ for $1 \le p < \infty$ is that

$$u \in W_0^{1,p}(G) : \int_G \Big(\sum_{i,j=1}^n a_{ij}\partial_i u \partial_j v + \Big(\sum_{i=1}^n \partial_i(a_i u) + au \Big) v \Big) = \int_G fv \ ,$$
$$v \in W_0^{1,p'}(G) \ .$$

The natural domain is $D(A_p) = \{u \in W_0^{1,p}(G) : A_p u \in L^p(G)\}$. We cite the following fundamental results for these operators

THEOREM 9.1 (AGMON-DOUGLIS-NIRENBERG). *The operator A_p is closed and densely-defined in $L^p(G)$ with domain $D(A_p) = W_0^{1,p}(G) \cap W^{2,p}(G)$ for $1 < p < \infty$, and $(I + \lambda A_p)^{-1}$ is a contraction for each $\lambda > 0$.*

For $1 \le i \le n$ we have

$$\int_G \partial_i(a_i u)u = -\int_G a_i u \partial_i u = (1/2) \int_G (\partial_i a_i) u^2 \ ,$$

so it follows that

$$\int_G \Big(\sum_{i=1}^n \partial_i(a_i u)u + au^2 \Big) = \int_G \Big((1/2) \sum_{i=1}^n \partial_i a_i + a \Big) u^2 \ge 0 \ .$$

This shows $(A_2 u, v)_{L^2}$ is coercive. Thus, we see that $-A_p$ generates a contraction semigroup on $L^p(G)$, and the coercivity estimate shows A_p is a surjection.

The case $p = 1$ is different. The spaces L^1 and L^∞ are not mutually dual so we cannot employ the adjoint arguments used in the proof of Theorem 9.1; also see Lemma 9.1. Since there is not a duality map $L^\infty \to L^1$, a-priori estimates are more difficult to obtain. As before, we define $D(A_1) = \{u \in W_0^{1,1}(G) : A_1 u \in L^1(G)\}$ where $A_1 u = f \in L^1(G)$ means

$$u \in W_0^{1,1}(G) : \int_G \Big(\sum_{i,j=1}^n a_{ij}\partial_i u \partial_j v - \sum_{i=1}^n a_i u \partial_i v + auv \Big) = \int_G fv \ , \qquad v \in W_0^{1,\infty}(G) \ .$$

We shall prove the following.

PROPOSITION 9.1.
(a) *$D(A_1)$ is dense, and $(I + \lambda A_1)^{-1}$ is a contraction on L^1 for each $\lambda > 0$.*
(b) *$D(A_1) \subset W_0^{1,q}$ for $1 \le q < n/(n-1)$ and there is a $c(q) > 0$: $c(q)\|u\|_{W^{1,q}} \le \|A_1 u\|_{L^1}$ for $u \in D(A_1)$.*
(c) *A_1 is the L^1-closure $\bar{A}_2$ of A_2.*
(d) *$\sup_G (I + \lambda A_1)^{-1} f \le \max\{0, \sup_G f\}$ for each $\lambda > 0$ and $f \in L^1$, that is $\|[(I + \lambda A_1)^{-1} f\|_{L^\infty(G)} \le \|f^+\|_{L^\infty(G)}$ where $x^+ = \max\{0, x\}$ denotes the positive part of $x \in \mathbb{R}$.*

LEMMA 9.1. *$A_1 \supset \bar{A}_2$ and $\bar{A}_2$ satisfies (b).*

PROOF. From the Sobolev Theorem 4.3, we obtain $W^{2,1} \subset W^{1,q}$, so $D(A_2) = W_0^{1,2} \cap W^{2,2} \subset W^{1,q}$. Also $W_0^{1,2} \subset W_0^{1,1}$ so $A_2 \subset A_1$. That is, we have $D(A_2) \subset D(A_1)$. Consider the adjoint equation

$$-\sum_{i,j=1}^n \partial_i(a_{ij}\partial_j v) - \sum_{i=1}^n a_i \partial_i v + av = -\sum_{i=1}^n \partial_i h_i \ .$$

Stampacchia (1963) showed there is a unique solution $v \in H_0^1 \cap L^\infty$ if each $h_i \in L^p$, $p > n$: for each $w \in H_0^1$

$$\int_G \Big(\sum_{i,j=1}^n a_{ij}\partial_j v \partial_i w - \sum a_i \partial_i v w + a v w \Big) = \int_G \sum_{i=1}^n h_i \partial_i w$$

and it satisfies

$$\|v\|_{L^\infty} \le C \sum_{i=1}^n \|h_i\|_{L^p} \ .$$

By choosing $w = u \in D(A_2)$ we obtain

$$\sum_{i=1}^n \int_G h_i \partial_i u = \int_G v A_2 u \le \|v\|_{L^\infty} \|A_2 u\|_{L^1} \le C \sum_{i=1}^n \|h_i\|_{L^p} \|A_2 u\|_{L^1} \ .$$

Since $h_1, \dots, h_n$ is arbitrary in $(L^p)^n$ it follows that

$$\sum_{i=1}^n \|\partial_i u\|_{L^q} \le C \|A_2 u\|_{L^1} \ , \qquad q = p/(p-1) < n/(n-1) \ .$$

Since the graph of A_1 is closed in $W^{1,1} \times L^1$, it follows that $\overline{A}_2 \subset A_1$. □

LEMMA 9.2. *Let $\varphi : \mathbb{R} \to \mathbb{R}$ be Lipschitz, monotone and $\varphi(0) = 0$. Then*

$$\big(A_2 u, \varphi_\varepsilon(u)\big)_{L^2} \ge 0 \quad \textit{for} \quad u \in D(A_2) \ .$$

PROOF. We have

$$\big(A_2 u, \varphi(u)\big)_{L^2} = \int_G \Big(\sum_{i,j=1}^n (a_{ij}\partial_i u \partial_j u)\varphi'(u) + \sum_{i=1}^n \partial_i(a_i u)\varphi(u) + a u \varphi(u) \Big)$$

and the first term is non-negative by ellipticity of A and monotonicity of φ. To estimate the other terms set $\zeta(t) = \int_0^t s\varphi'(s)\,ds$ and note $0 \le \zeta(t) = t\varphi(t) - \int_0^t \varphi \le t\varphi(t)$. The sum of the second and third terms is

$$\int_G \Big(-\sum_{i=1}^n a_i u \varphi'(u) \partial_i u + a u \varphi(u) \Big) \ge \int_G \Big(-\sum_{i=1}^n a_i \partial_i \zeta(u) + a\zeta(u) \Big)$$

$$= \int_G \Big(\sum_{i=1}^n \partial_i a_i + a \Big) \zeta(u) \ge 0 \ . \qquad \square$$

LEMMA 9.3. *$\bar{A}_2$ satisfies (a).*

PROOF. Let $f \in L^2$, $u = (I + \lambda A_2)^{-1} f$ so $u + \lambda A_2 u = f$. For $\varepsilon > 0$ define $\varphi_\varepsilon(u) = \frac{1}{\varepsilon}(u - (I + \varepsilon\,\mathrm{sgn})^{-1} u) = \mathrm{sgn}(u)$ for $|u| \ge \varepsilon$ and $= u/\varepsilon$ for $|u| \le \varepsilon$. Multiply by $\varphi_\varepsilon(u)$ and integrate: $\int_G u\varphi_\varepsilon(u) \le \int_G f\varphi_\varepsilon(u) \le \|f\|_{L^1}$. By letting $\varepsilon \to 0^+$ we obtain $\|u\|_{L^1} \le \|f\|_{L^1}$, hence, $(I + \lambda A_2)^{-1}$ is a contraction on L^2.

Finally note that if $f_n \in L^2$ and $f_n \to f \in L^1$ with convergence in L^1, then $u_n \equiv (I + \lambda A_2)^{-1} f_n$ converges in L^1 to $u = (I + \lambda \bar{A}_2)^{-1} f$ and $\|u\|_{L^1} \le \|f\|_{L^1}$. □

LEMMA 9.4. *A_1 is one-to-one.*

PROOF. For each $g \in L^p$, $p > n$, there is a solution v of the adjoint problem $v \in W_0^{1,p} \cap W^{2,p}$:

$$\int_G \Big(\sum_{i,j=1}^{n} a_{ij} \partial_j v \partial_i w - \sum_{i=1}^{n} a_i \partial_i v w + a v w \Big) = \int_G g w \ , \qquad w \in C_0^\infty \ .$$

Since $W^{2,p} \subset C^1(\bar{G})$ by Theorem 4.3, we can let $w \in W_0^{1,1}$. Taking $w = u$ where $A_1 u = 0$ we obtain $\int_G gu = \int_G v A_1 u = 0$ for g as above, so $u = 0$. □

COROLLARY 9.1. *$I + \lambda A_1$ is one-to-one for each $\lambda > 0$.*

We have shown $\bar{A}_2 \subset A_1$, $I + \lambda \bar{A}_2$ is onto and $I + \lambda A_1$ is one-to-one. Thus $\bar{A}_2 = A_1$ and all of Proposition 9.1 is proved except (d). The proof of this maximum principle is the same as the L^1-estimate in (a).

LEMMA 9.2A. *For $\varepsilon > 0$ define $\varphi_\varepsilon^+(u) = \frac{1}{\varepsilon}\big(u - (I + \varepsilon \operatorname{sgn}^+)^{-1} u\big) = \operatorname{sgn}^+(u)$ for $u \geq \varepsilon$ or $u \leq 0$ and $\varphi_\varepsilon^+(u) = u/\varepsilon$ for $0 \leq u \leq \varepsilon$. Then*

$$\big(A_2 u \ , \ \varphi_\varepsilon^+(u-k)\big)_{L^2} \geq 0 \quad \textit{for} \quad u \in D(A_2) \ , \qquad k \geq 0 \ .$$

LEMMA 9.3A. *A_2 satisfies (d).*

PROOF. If $(I + \lambda A_2)u = f \in L^2$ then $u - k + \lambda A_2 u = f - k$. Multiply by $\varphi_\varepsilon^+(u-k)$, integrate and let $\varepsilon \to 0^+$ to obtain

$$\int_G (u-k)^+ \leq \int_G (f^+ - k)^+ \ ,$$

where $r^+ = r \operatorname{sgn}^+(r)$. Setting $k = \sup f^+$ shows $(u-k)^+ = 0$, and so we obtain

$$\|u^+\|_{L^\infty} \leq \|f^+\|_{L^\infty} \ . \qquad \square$$

COROLLARY 9.2. *A_1 satisfies (d).*

PROOF. Let $f \in L^1$ and choose $f_n \in L^2$ such that $f_n \to f$ in L^1, $f_n(x) \to f(x)$, $f_n(x) \leq f(x)^+$, and $u_n(x) = (I + \lambda A_2)^{-1} f_n(x) \to u(x)$ a.e. $x \in G$. Then let $n \to \infty$ in $u_n(x) \leq \|f_n^+\|_{L^\infty} \leq \|f^+\|_{L^\infty}$. □

We shall extend Lemma 9.2 and Lemma 9.2A to more general monotone functions. Also, the *negative part* of any $x \in \mathbb{R}$ will be denoted by $x^- = \min\{0, x\}$. A useful estimate is given by the following.

PROPOSITION 9.2. *Let $T : L^1(G) \to L^1(G)$ satisfy*
(a) *$\|Tu - Tv\|_{L^1} \leq \|u - v\|_{L^1}, u, v \in L^1$, and*
(b) *$-\|u^-\|_{L^\infty} \leq Tu(x) \leq \|u^+\|_{L^\infty}$ a.e. $x \in G$. Let $j : \mathbb{R} \to \mathbb{R}_+$ be convex, lower-semi-continuous and $j(0) = 0$. Then for each $u \in L^1(G)$*

$$\int_G j\big(Tu(x)\big)\, dx \leq \int_G j\big(u(x)\big)\, dx \ .$$

PROOF. Consider the special case $j_1(r) = (r-t)^+$ with $t \geq 0$. Set $v(x) = \min\{u(x), t\}$ so $v \in L^1$ and $|u(x) - v(x)| = (u(x)-t)^+$. Also $Tv(x) \leq \|v^+\|_{L^\infty} \leq t$ so

$$\big(Tu(x) - t\big)^+ \leq \big(Tu(x) - Tv(x)\big)^+ \leq |Tu(x) - Tv(x)| \ ,$$

hence

$$\int_G \big(Tu(x) - t\big)^+ \, dx \leq \|u - v\|_{L^1} = \int_G \big(u(x) - t\big)^+ \, dx$$

as desired. The same holds for $u \mapsto -T(-u)$ and so for $t \leq 0$ we have

$$\int_G \big(-Tu(x) + t\big)^+ \, dx \leq \int_G \big(-u(x) + t\big)^+ \, dx \ .$$

These may be summarized by

$$\int_G \Big[t\big(Tu(x) - t\big)\Big]^+ \, dx \leq \int_G \Big[t\big(u(x) - t\big)\Big]^+ dx \ , \qquad t \in \mathbb{R} \ . \tag{9.1}$$

Consider the smooth case with $0 \leq j'' \in L^\infty(\mathbb{R})$ and, hence, $j'(0) = 0$. First check that

$$j(s) = \int_{-\infty}^{\infty} j''(t) \big[\mathrm{sgn}(t)(s-t)\big]^+ \, dt \ .$$

Then it follows that we can multiply (9.1) by $j''(t)/|t|$ and integrate via Tonelli to obtain the desired estimate.

Finally, for any function j as given in the Proposition 9.2 we construct a sequence of smooth functions as above, e.g., $j_\varepsilon(r) = \inf_{t \in \mathbb{R}}\{(1/2\varepsilon)|r-t|^2 + j(t)\}$, which converge monotonically to j from below. Then

$$\int_G j_\varepsilon(Tu) \leq \int_G j_\varepsilon(u) \leq \int_G j(u) \ , \qquad \varepsilon > 0 \ ,$$

and this implies the desired estimate. □

COROLLARY 9.3. *$\|Tu\|_{L^p} \leq \|u\|_{L^p}$, $\|Tu^+\|_{L^p} \leq \|u^+\|_{L^p}$ for $p \geq 1$.*

PROOF. Take $j(r) = r^p$ and $(r^+)^p$. □

The next result is central and roughly asserts that $\int_G Au \cdot \beta(u) \, dx \geq 0$ for any maximal monotone graph $\beta \subset \mathbb{R} \times \mathbb{R}$ which contains the origin.

PROPOSITION 9.3. *Let β be a maximal monotone graph in $\mathbb{R} \times \mathbb{R}$ and $0 \in \beta(0)$; let A satisfy (a) and (d) of Proposition 9.1; let $1 \leq p \leq \infty$ and $1/p + 1/p' = 1$. Then for each pair $u \in L^p, v \in L^{p'}$ with $Au \in L^p$ and $v(x) \in \beta(u(x))$ a.e. $x \in G$*

$$\int_G Au(x) v(x) \, dx \geq 0 \ .$$

PROOF. Let the convex lower-semi-continuous $j : \mathbb{R} \to \mathbb{R}_\infty$ be the indefinite integral of β with $j(0) = 0$; that is, $\partial j = \beta$. (See Example 8.A.) For each $\lambda > 0$ set $T_\lambda = (I + \lambda A)^{-1}$, so (a) and (b) of Proposition 9.2 are satisfied. Note that $\lambda T_\lambda A = I - T_\lambda$ on $D(A)$ so $T_\lambda \to I$ pointwise on L^1. Since $v(x) \in \partial j(u(x))$, a.e. $x \in G$, we have $v(x)(0 - u(x)) \leq j(0) - j(u(x))$, hence, $0 \leq j(u(x)) \leq v(x)u(x)$, so $j \circ u \in L^1$. Likewise

$$-\lambda v(x)\big(T_\lambda Au(x)\big) = v(x)\big(T_\lambda u(x) - u(x)\big) \leq j\big(T_\lambda u(x)\big) - j\big(u(x)\big) \ , \quad \text{a.e. } \ x \in G \ ,$$

so from Proposition 9.2 we obtain

$$\int_G T_\lambda Au(x)v(x)\,dx \ge 0\ , \qquad \lambda > 0\ .$$

If $p = 1$ the result follows from the strong convergence $T_\lambda Au \to Au$ in L^1. If $1 < p < \infty$ then Proposition 9.2 shows $\|T_\lambda Au\|_{L^p} \le \|Au\|_{L^p}$ and the result follows from the weak convergence $T_\lambda Au \rightharpoonup Au$ in L^p. If $p = \infty$ then $\|T_\lambda Au\|_{L^\infty} \le \|Au\|_{L^\infty}$; a subsequence a.e. convergent gives the result by dominated convergence. □

We have shown that the operator A_1 satisfies the hypotheses on the abstract linear operator A in $L^1(G)$ in the following result. As a consequence, for any maximal monotone graph β with $0 \in \beta(0)$ it follows by setting $\beta = \alpha^{-1}$ that the composite nonlinear operator $A_1 \circ \beta$ is m-accretive in $L^1(G)$, and its resolvent equation satisfies some useful *comparison estimates*. Additional examples will be given below.

THEOREM 9.2 (BREZIS-STRAUSS). *Let α be a maximal monotone graph in $\mathbb{R} \times \mathbb{R}$ and $0 \in \alpha(0)$. Let $A : D(A) \to L^1(G)$ be linear and satisfy the following:*
(i) *$D(A)$ is dense and $(I + \lambda A)^{-1}$ is a contraction in L^1 for each $\lambda > 0$;*
(ii) *$\sup_G (I + \lambda A)^{-1} f \le (\sup_G f)^+ = \|f^+\|_{L^\infty}$ for $f \in L^1$ and $\lambda > 0$;*
(iii) *there is a $c > 0$ such that*

$$c\|u\|_{L^1} \le \|Au\|_{L^1} \quad \textit{for} \quad u \in D(A).$$

Then for each $f \in L^1$ there is a unique pair $u \in D(A)$, $v \in L^1$ such that

$$Au + v = f \quad \textit{and} \quad v(x) \in \alpha\big(u(x)\big)\ ,\ \textit{a.e}\quad x \in G\ . \tag{9.2}$$

If u_1, v_1 and u_2, v_2 are solutions corresponding to f_1, f_2 as above, then

$$\|(v_1 - v_2)^+\|_{L^1} \le \|(f_1 - f_2)^+\|_{L^1}\ ,\ \|(v_1 - v_2)^-\|_{L^1} \le \|(f_1 - f_2)^-\|_{L^1}\ , \tag{9.3}$$

and, hence,

$$\|v_1 - v_2\|_{L^1} \le \|f_1 - f_2\|_{L^1}\ . \tag{9.4}$$

If $f_1 \ge f_2$ a.e. then $v_1 \ge v_2$ a.e. on G.

PROOF. The last claim follows from the second estimate in (9.3), and (9.4) follows by adding (9.3). By symmetry it suffices to verify the first estimate in (9.3). Multiply the equation

$$A(u_1 - u_2) + (v_1 - v_2) = f_1 - f_2$$

by $v(x) \equiv \operatorname{sgn}_0^+(u_1(x) - u_2(x) + v_1(x) - v_2(x))$ where $\operatorname{sgn}_0^+(r) = 1$ for $r > 1$ and $= 0$ otherwise. Note $v(x) \in \operatorname{sgn}^+(u_1(x) - u_2(x)) \cap \operatorname{sgn}^+(v_1(x) - v_2(x))$ since α is monotone. From Proposition 9.3 we obtain

$$\|(v_1 - v_2)^+\|_{L^1} = \int_G (v_1 - v_2)v \le \int_G (f_1 - f_2)v \le \int_G (f_1 - f_2)^+ v \le \|(f_1 - f_2)^+\|_{L^1}\ .$$

It remains to show $Rg(A + \alpha)$ is all of L^1. To show the range is closed, let $Au_n + v_n = f_n \to f$ in L^1 with $v_n \in \alpha(u_n)$ for $n \ge 1$. Then (9.4) implies $v_n \to v$ in L^1 and (iii) implies $u_n \to u$ in L^1 with $Au + v = f$. Since $(I + \alpha)^{-1}$ is a Lipschitz function we may take the limit in $u_n = (I + \alpha)^{-1}(u_n + v_n)$ to obtain

$u = (I+\alpha)^{-1}(u+v)$, hence, $v \in \alpha(u)$. It suffices now to show $Rg(A+\alpha)$ is dense in L^1.

In order to solve (9.2) we shall *regularize* both operators, replacing A by $A+\varepsilon I$ and α^{-1} by $\alpha^{-1}+\lambda I$ where $\varepsilon, \lambda > 0$. Note the equivalence of $(\lambda I + \alpha^{-1})(y) \ni x$, $y \in \alpha(x-\lambda y)$, $x - \lambda y + \lambda\alpha(x-\lambda y) \ni x$ and $y = (1/\lambda)(I-(I+\lambda\alpha)^{-1})(x)$. Thus we set $\alpha_\lambda \equiv (1/\lambda)(I-(I+\lambda\alpha)^{-1})$, the monotone Lipschitz (Yosida) approximation of α, and consider the equation

$$(9.5) \qquad \varepsilon u + Au + \alpha_\lambda(u) = f \ .$$

This is equivalent to

$$u = (1+\lambda\varepsilon)^{-1}\big(I + (\lambda/1+\varepsilon\lambda)A\big)^{-1}\big(\lambda f + (I+\lambda\alpha)^{-1}u\big) \ ,$$

and so this strict contraction on the right side guarantees there is a solution in L^1. But we shall need more ... estimates in L^2 ... so we consider $L^1 \cap L^\infty$ with the norm $\|u\| = \|u\|_{L^1} + \|u\|_{L^\infty}$. If $f \in L^1 \cap L^\infty$ and $\varepsilon > 0$ is fixed, then for each $\lambda > 0$ there is a solution $u_\lambda \in L^1 \cap L^\infty$ of (9.5) obtained as a fixed-point in this space, and it satisfies $\|u_\lambda\| \le (1+\lambda\varepsilon)^{-1}(\lambda\|f\| + \|u_\lambda\|)$, hence, $\|u_\lambda\| \le (1/\varepsilon)\|f\|$. Next multiply (9.5) by $\operatorname{sgn}(u_\lambda)$ and integrate to obtain from Proposition 9.3 $\|\alpha_\lambda(u_\lambda)\|_{L^1} \le \|f\|_{L^1}$. Similarly get an L^p estimate and let $p \to \infty$ to give $\|\alpha_\lambda(u_\lambda)\|_{L^\infty} \le \|f\|_{L^\infty}$ for each $\lambda > 0$. This shows that $\{\alpha_\lambda(u_\lambda)\}$ is bounded in L^2, and we next deduce that $\{u_\lambda\}$ is Cauchy in L^2. Take the difference of (9.5) at λ, μ and scalar product with $u_\lambda - u_\mu$ to obtain from Proposition 9.3

$$\varepsilon\|u_\lambda - u_\mu\|^2_{L^2} + \big(\alpha_\lambda(u_\lambda) - \alpha_\mu(u_\mu) \ , \ u_\lambda - u_\mu\big)_{L^2} \le 0 \ .$$

Substitute in the right term

$$u_\lambda - u_\mu = \big(I-(I+\lambda\alpha)^{-1}\big)u_\lambda + \big((I+\lambda\alpha)^{-1}u_\lambda - (I+\mu\alpha)^{-1}u_\mu\big) - \big(I-(I+\mu\alpha)^{-1}\big)u_\mu$$

and note the middle product is non-negative because $\alpha_\lambda \in \alpha(I+\lambda\alpha)^{-1}$; this gives

$$\varepsilon\|u_\lambda - u_\mu\|^2_{L^2} + \big(\alpha_\lambda(u_\lambda) - \alpha_\mu(u_\mu) \ , \ \lambda\alpha_\lambda(u_\lambda) - \mu\alpha_\mu(u_\mu)\big) \le 0 \ .$$

Since $\|\alpha_\lambda(u_\lambda)\|_{L^2}$ is bounded, it follows that $\{u_\lambda\}$ is Cauchy in L^2. From Lemma 9.5 below it also follows that $\{\alpha_\lambda(u_\lambda)\}$ is Cauchy in L^2. By passing to a subsequence we obtain $u_\lambda \to u$ and $\alpha_\lambda(u_\lambda) \to v$ in $L^2(G)$ with $u, v \in L^1 \cap L^\infty$. As before we find $v(x) \in \alpha(u(x))$ a.e. and $(\varepsilon I + A)u_\lambda \to f - v$ in $L^2(G)$. Set $w = (\varepsilon I + A)^{-1}(f-v)$; then $w \in D(A) \cap L^\infty$ and $(\varepsilon I + A)(u_\lambda - w) \to 0$ and $u_\lambda \to w = u$. Thus

$$\varepsilon u + Au + v = f \ , \qquad v \in \alpha(u) \ .$$

Finally, let $f \in L^1$ and $f_\varepsilon \in L^1 \cap L^\infty$ with $f_\varepsilon \to f$ in L^1. For $\varepsilon > 0$ let u_ε be the solution as above of

$$\varepsilon u_\varepsilon + Au_\varepsilon + v_\varepsilon = f_\varepsilon \ , \qquad v_\varepsilon \in \alpha(u_\varepsilon) \ .$$

As before we get $\varepsilon\|u_\varepsilon\|_{L^1} + \|v_\varepsilon\|_{L^1} \le \|f_\varepsilon\|_{L^1}$ from Proposition 9.3 and then (iii) gives

$$c\|u_\varepsilon\|_{L^1} \le 2\|f_\varepsilon\|_{L^1} \le 2\|f\|_{L^1} \ ,$$

hence, $\varepsilon u_\varepsilon \to 0$ and $f = \lim_{\varepsilon\to 0}(f_\varepsilon - \varepsilon u_\varepsilon)$ belongs to the closure of $Rg(A+\alpha)$. This finishes the proof of Theorem 9.2 except for the following general result. □

LEMMA 9.5. *Let $\{z_\lambda\}$ be given in a scalar-product space for $\lambda > 0$ and assume*

$$(z_\lambda - z_\mu \ , \ \lambda z_\lambda - \mu z_\mu) \le 0 \ , \qquad \lambda, \mu > 0 \ .$$

Then $\lambda \mapsto \|z_\lambda\|$ is monotone decreasing; if $\{z_\lambda\}$ is bounded then it is Cauchy for $\lambda \to 0^+$.

PROOF. From the calculation

$$0 \ge 2(z_\lambda - z_\mu \ , \ \lambda z_\lambda - \mu z_\mu) = (\lambda + \mu)\|z_\lambda - z_\mu\|^2 + (\lambda - \mu)\big(\|z_\lambda\|^2 - \|z_\mu\|^2\big)$$

we obtain

$$(\lambda + \mu)\|z_\lambda - z_\mu\|^2 \le (\lambda - \mu)\big(\|z_\mu\|^2 - \|z_\lambda\|^2\big)$$

so $\lambda > \mu$ implies $\|z_\lambda\| \le \|z_\mu\|$. Also we have

$$\|z_\lambda - z_\mu\|^2 \le \|z_\mu\|^2 - \|z_\lambda\|^2 \ , \qquad \lambda > \mu > 0 \ ,$$

so boundedness implies the Cauchy condition. □

EXAMPLE 9.A. This first order operator is the L^1 realization of Example I.4.B. Set $D(A) = \{v \in W^{1,1}(a,b) : v(a) = cv(b)\}$, where $a < b$ and $0 \le c < 1$. Define $A = \partial$ on $D(A)$. It is easy to check that there is a solution $u \in D(A)$ of $(I + \lambda A)(u) = f$ for each $f \in L^1(a,b)$ and $\lambda > 0$. For any such solution, we multiply the equation by $\text{sgn}(u)$ and integrate to get $\|u\|_{L^1(a,b)} \le \|f\|_{L^1(a,b)}$, so (i) holds. Multiply the identity

$$u(x) - k + \lambda A u(x) = f(x) - k$$

by $\text{sgn}^+(u(x) - k)$, integrate, and set $k = \|f^+\|_{L^\infty(a,b)}$ to obtain

$$\|(u-k)^+\|_{L^1(a,b)} + (u(b) - k)^+ - (cu(b) - k)^+ \le 0 \ .$$

Since $k \ge 0$, $c \ge 0$ and the *positive part* function $(\cdot)^+$ is monotone, we have

$$\begin{aligned}(u(b) - k)^+ - (cu(b) - k)^+ &\ge (u(b) - k)^+ - (cu(b) - ck)^+ \\ &\ge (1 - c)(u(b) - k)^+ \ge 0 \ ,\end{aligned}$$

so $(u - k)^+ = 0$ and we have (ii). To check (iii), we let $Au = f$. Multiply by $\text{sgn}(u)$ and integrate to get

$$|u(x)| \le |u(a)| + \|f\|_{L^1(a,b)} \ , \qquad a \le x \le b \ .$$

Integrating $Au = f$ and using the boundary condition give $u(a) = \frac{c}{1-c}\|f\|_{L^1(a,b)}$, so with the above we obtain (iii) from

$$|u(x)| \le \frac{1}{1-c}\|f\|_{L^1(a,b)} \ , \qquad a \le x \le b \ .$$

Theorem 9.2 asserts that for any maximal monotone β in $\mathbb{R} \times \mathbb{R}$ with $0 \le \beta(0)$, the operator $\partial \circ \beta$ is m-accretive in $L^1(a,b)$. Specifically, if $f \in L^1(a,b)$ there is a unique pair

$$v \in L^1(a,b) \quad , \quad u \in W^{1,1}(a,b) \ ,$$

for which $u(a) = cu(b)$ and

$$v(x) + \partial u(x) = f(x) \quad , \quad u(x) \in \beta(v(x)) \ , \qquad \text{a.e. } x \in (a,b) \ ,$$

and the mapping $f \mapsto v$ is an $L^1(a,b)$-contraction. This operator will arise in the study of *scalar conservation laws*. Other such first order operators can be so constructed; one can delete the second order terms from A_1 above, for example.

EXAMPLE 9.B. We record here a special case of Proposition 9.1. Let β be given as above. Then for each $f \in L^1(G)$ there is a unique pair

$$v \in L^1(G) \quad , \quad u \in W_0^{1,1}(G)$$

for which the Laplacian $\Delta u \in L^1(G)$ and

$$v(x) - \Delta u(x) = f(x) \quad , \quad u(x) \in \beta(v(x)) \ , \qquad \text{a.e. } \ x \in G \ ,$$

and the mapping $f \mapsto v$ is a contraction in $L^1(G)$. The operator $-\Delta \circ \beta$ corresponds to the stationary problem for the *porous medium equation*, and one can construct similar operators by varying either the linear elliptic part, $-\Delta$, as above or the boundary conditions.