Publ. Math. Debrecen
46 / 3-4 (1995), 215–237

Identification of parameters in systems governed by nonlinear evolution equations

By NIKOLAOS S. PAPAGEORGIOU (Athens)

Abstract. In this paper we study the problem of parameter identification for systems governed by nonlinear evolution equations. First we establish the existence of an optimal parameter value for two families of systems. The first family are those systems monitored by evolutions of the subdifferential type, while the second family consists of systems whose dynamics is described by differential variational inequalities. Then using the framework of evolution triples, we obtain necessary conditions for optimality. An example of a parabolic system is worked out in detail.

1. Introduction

In the analysis of models of engineering or physical systems, one of the major tasks is to determine the unknown parameters of the model, on the basis of the available field data. This problem is known as the "identification problem". In the recent years, the identification of distributed parameter systems has attracted the interest of many applied mathematicians and substantial progress was made in various aspects of the problem. In particular, techniques of optimal control have been used by Lions [15], Chavent [11], Ahmed [1], [2] and Banks–Reich–Rosen [6], [7], [8], who considered the parameter identification problem. Among the above authors, Ahmed [1], [2] made a systematic effort to develop a general theory of optimization and identification of linear and nonlinear systems monitored by evolution equations in Banach spaces. In [1] Ahmed presented the linear theory, while in his monograph [2] he also developed the nonlinear theory and related computational schemes. On the other

1980 *Mathematics Subject Classification* (1985 *Revision*): 49A20.
Key words and phrases: Parameter identification, subdifferential, R-convergence, differential variational inequality, adjoint equation, maximum principle, evolution triple, parabolic system.

hand, in a series of interesting papers, Banks–Reich–Rosen [6], [7], [8], developed the abstarct framework and the corresponding convergence theory for the Galerkin approximations for inverse problems involving the identification of nonlinear distributed parameter systems. Papers [6] and [7] deal primarily with autonomous systems, while paper [8] treats the temporally inhomogeneous case (nonautonomous systems), using a variational approach similar to the one used by Barbu [9] to prove his theorem III.4.2. In this paper, we do not limit ourselves to the Galerkin approximations of this systems under consideration, but instead as in Ahmed [2], we deal with general nonlinear systems involving a parameter to be identified in a ceratin optimal way. In Ahmed [2] (chapter 4) as well as in Banks–Reich–Rosen [6], [7], [8], the systems are defined using the "evolution triple" formalism (see Zeidler [24]) and satisfy a strong monotonicity (coercivity) property. In our existence theory (section 3), we consider a different class of systems that do not necessarily satisfy the coercivity property and in which the abstract nonlinear operator may be multivalued. However, we require that it is of subdifferential form. The parameter to be identified appears in all the data of the problem, including the subdifferential. Using an integral criterion for optimality, we are able to establish the existence of an optimal parameter value. Our formulation incorporates "differential variational inequalities", which appear in many applications, like economic dynamic systems (resource allocation problems; see Aubin–Cellina [4], Chapter 5, Section 6) and in theoretical mechanics, in the study of unilateral problems (see Moreau [16]). In general when we study systems with constraints, often in describing the effect of the constraint on the dynamics of the system, we can assume that the velocity $\dot{x}$ is projected at each instant on the set of allowed directions toward the constraint set at the point x. This leads us to a "different variational inequality" (see Aubin–Cellina [4]). This large class of nonlinear systems can not be treated using the framework adopted by Ahmed [2]. Furthermore, our forcing term f depends on the state x, while in Ahmed [2], Chapter 4, p. 96 is independent of x and finally our cost criterion is more general since it also depends on the "velocity" $\dot{x}$. In our hypotheses (see hypothesis $H(\varphi)(2)$, Section 3), we assume that the subdifferential operator as a set-valued function of the parameter is continuous in the "resolvent topology" (see Section 2). On the other hand, Ahmed [2], Theorem 4.2, p. 102 assumes that his abstract single-valued operator is strongly continuous in the parameter. Although the two continuity concepts are in general distinct, it seems to us that our hypothesis is more appropriate, since as it was illustrated in Attouch [3] (with examples from mechanics, optimization and optimal control), the resolvent convergence is more natural within the context of variational problems. Furthermore, our hypothesis is the natural extension to the nonlinear and set-valued situation of the hypothesis used in the linear theory by Ahmed [2] (see Theorem 1.11, p. 27). So our Theorem 3.1, can be viewed as a partial extension of Theorem 4.2, p. 102 of Ahmed

[2]. In Section 4, we turn our attention to a different class of nonlinear systems (this time defined using an evolution triple, nonautonomous, but with the maximal monotone operator being single-valued). For this new class of systems, we derive necessary conditions for optimality, without assuming any differentiability properties on the cost criterion. Our result (Theorem 4.1) extends Theorem 2.4, p. 54 of Ahmed [2] who considers autonomous systems, with state independent forcing term. Finally an example of a parabolic control system, with controls in the coefficients, is worked out in detail. To conclude this introduction we should mention that the works of Banks–Reich–Rosen [6], [7], [8] are more directly related to the recent continuous dependence results for evolution inclusions obtained by the author in [20] and [21].

2. Mathematical preliminaries

Let X be a Banach space and $\varphi : X \to \bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$. We will say that $\varphi(\cdot)$ is proper, if it is not identically $+\infty$. Assume that $\varphi(\cdot)$ is proper, convex and l.s.c. In the literature, this set of functions is denoted by $\Gamma_0(X)$. By $\operatorname{dom}\varphi$, we will denote the effective domain of $\varphi(\cdot)$; i.e. $\operatorname{dom}\varphi = \{x \in X : \varphi(x) < +\infty\}$. The subdifferential of $\varphi(\cdot)$ at x is the set $\partial\varphi(x) = \{x^* \in X^* : (x^*, y - x) \leq \varphi(y) - \varphi(x)$ for all $y \in \operatorname{dom}\varphi\}$ (here $(\cdot\,,\cdot)$ denotes the duality brackets of the pair (X, X^*)). If $\varphi(\cdot)$ is Gateaux differentiable at x, then $\partial\varphi(x) = \{\varphi'(x)\}$. Recall that $\partial\varphi : D(\partial\varphi) \subseteq X \to 2^{X^*}$ is a maximal monotone map. We will say that $\varphi(\cdot)$ is of compact type, if for every $\lambda \in \mathbb{R}$, the level set $\{x \in X : \|x\|^2 + \varphi(x) \leq \lambda\}$ is relatively compact.

Let $\{A_n, A\}_{n\geq 1} \subseteq 2^X\{\emptyset\}$. Denote by s- the strong topology on X and by w- the weak topology. We define:

$$\begin{aligned} s\text{-}\underline{\lim} A_n &= \{x \in X : x = s\text{-}\lim x_n,\ x_n \in A_n,\ n \geq 1\} = \\ &= \{x \in X : \lim d(x, A_n) = 0\} \end{aligned}$$

$$\begin{aligned} \text{and } w\text{-}\overline{\lim} A_n = \{x \in X : x = w\text{-}\lim x_{n_k},\ x_{n_k} \in A_{n_k}, \\ n_1 < n_2 < \dots < n_k < \dots\}. \end{aligned}$$

It is clear from the above definitions that we always have $s\text{-}\underline{\lim} A_n \subseteq w\text{-}\overline{\lim} A_n$. If $s\text{-}\underline{\lim} A_n = w\text{-}\overline{\lim} A_n = A$, then we say that the A_n's converge to A in the Kuratowski–Mosco sense, denoted by $A_n \xrightarrow{K-M} A$ (see Mosco [17]). If $\dim X < \infty$, then the weak and strong topologies on X coincide and so we recover the well-known Kuratowski mode of set convergence.

Let H be a Hilbert space. From a fundamental result of Minty, we know that an operator $A : H \to 2^H$ is maximal monotone if and only if for some $\lambda > 0$, $R(I + \lambda A) = H$. Then for every $\lambda > 0$, $J_\lambda = (I + \lambda A)^{-1}$:

$R(I+\lambda A)=H\to H$ is called the "resolvent of A". The resolvent map is nonexpansive and $J_\lambda x \xrightarrow{s} x$ as $\lambda\to 0^+$, for each $x\in D(A)=\{y\in X : A(y)\neq\emptyset\}$. Let $\mathcal{M}$ denote the set of all maximal monotone operators in H. The "topology of R-convergence" on $\mathcal{M}$, is the weakest topology that makes continuous the maps $\hat{J}_{\lambda,x}:\mathcal{M}\to H$ for every $\lambda>0$ and every $x\in H$, where $\hat{J}_{\lambda,x}(A)=(I+\lambda A)^{-1}x$. By $\mathcal{M}_R$ (or $\mathcal{M}_R(H)$) we will denote the set $\mathcal{M}$ equipped with the topology of R-convergence. If H is separable, then $\mathcal{M}_R$ is a Polish space (i.e. a separable, metrizable, complete space). Furthermore, we known that $A_n \xrightarrow{\mathcal{M}_R} A$ if and only if $GrA_n \xrightarrow{K-M} GrA$, where $GrA_n=\{(x,y)\in H\times H : y\in A_n x\}$ similarly for (GrA). For further details, we refer to Attouch [3].

Now let $T=[0,r]$, let H be a separable Hilbert space and X a subspace of H, carrying the structure of a separable, reflexive Banach space, which embeds continuously and densely into H. Identifying H with its dual (pivot space), we have $X\to H\to X^*$, with all embeddings being continuous and dense. Such a triple of spaces (X,H,X^*) is known in the literature as "evolution triple". To have a concrete example in mind, let Z be a bounded domain in $\mathbb{R}^N$ and let $X=W_0^{m,p}(Z)$, $H=L^2(Z)$, $X^*=W^{-m,q}(Z)$, where $m\in\mathbb{N}$, $2\le p<\infty$, $\frac{1}{p}+\frac{1}{q}=1$. From the Sobolev embedding theorem, we know that (X,H,X^*) is an evolution triple and in addition all embeddings are compact. By $\|\cdot\|$ (resp. $|\cdot|$, $\|\cdot\|_*$), we will denote the norm of X (resp. of H,X^*). Also by $\langle\cdot,\cdot\rangle$ we will denote the duality brackets for the pair (X,X^*) and by $(\cdot,\cdot)$ the inner product of H. The two are compatible in the sense that $\langle\cdot,\cdot\rangle\,|_{X\times H}=(\cdot,\cdot)$. Let $W(T)=\{x\in L^2(X) : \dot{x}\in L^2(X^*)\}$. The derivative in this definition is understood in the sense of vector-valued distributions. This space equipped with the norm $\|x\|_{W(T)}=[\|x\|^2{}_{L^2(X)}+\|\dot{x}\|^2{}_{L^2(X^*)}]^{1/2}$ becomes a Banach space, which is separable, reflexive, being a closed subspace of the separable and reflexive Banach space $L^2(X)\times L^2(X^*)$. We know (see Zeidler [24], Proposition 23.23, p. 422), that $W(T)$ embeds in $C(T,H)$ continuously. So every element in $W(T)$ has a unique representative in $C(T,H)$. Also by virtue of Theorem 2.2, p. 19 of Barbu [9], every $x\in W(T)$ can be identified with an X^*-valued absolutely continuous function on T and $\dot{x}$ can be regarded as the strong (ordinary) derivative of $x:T\to X^*$. If $X\to H$ compactly, then $W(T)\to L^2(H)$ compactly (see Zeidler [24], p. 450) and in addition X is a Hilbert space too, then $W(T)\to C(T,H)$ compactly (see Nagy [18]).

3. Existence theorems

Let $T=[0,r]$, H a separable Hilbert space (the state space) and G a compact metric space (the parameter space). The system under consideration is governed by the following evolution equation of the subdifferential type, parametrized by the elements of G:

$$(*)_1 \qquad \begin{gathered} -\dot{x}(t)\in\partial\varphi(x(t),p)+f(t,x(t),p) \text{ a.e.} \\ x(0)=x_0(p). \end{gathered}$$

Our goal is, using the trajectories of $(*)_1$, to minimize the following integral functional:

$$(P_1) \qquad J(p)=\int_0^r L(t,x(p)(t),\dot{x}(p)(t),p)dt\to\inf_{p\in G}=m_1.$$

To this end we will need the following hypotheses on the data of our problem:

$\underline{H(\varphi)}$: $\varphi: H\times G\to\bar{\mathbb{R}}=\mathbb{R}\cup\{+\infty\}$ is a function s.t.

(*1*) for every $p\in G$, $\varphi(\cdot\,,p)\in\Gamma_0(H)$ and is of compact type uniformly in $p\in G$ (i.e. for every $\lambda\in\mathbb{R}$, $K_\lambda=\bigcup\limits_{p\in G}\{x\in H:$ $\|x\|^2+\varphi(x,p)\le\lambda\}$ is relatively compact in H),

(*2*) if $p_n\to p$ in G, then $\partial\varphi(\cdot\,,p_n)\to\partial\varphi(\cdot\,,p)$ in $\mathcal{M}_R(H)$,

(*3*) $\sup\limits_{p\in G}\varphi(x_0(p),p)<\infty$.

$\underline{H(f)}$: $f: T\times H\times G\to H$ is a map s.t.

(*1*) $t\to f(t,x,p)$ is measurable,

(*2*) $\|f(t,x,p)-f(t,x',p)\|\le k(t)\|x-x'\|$ a.e. for all $p\in G$ and with $k(\cdot)\in {L^1}_+$,

(*3*) $p\to f(t,x,p)$ is continuous,

(*4*) $\|f(t,x,p)\|\le a(t,p)+b(t,p)\|x\|$ a.e. with $a(\cdot\,,p),b(\cdot\,,p)\in {L^2}_+$, and $\sup\limits_{p\in G}\|a(\cdot\,,p)\|_2$, $\|b(\cdot\,,p)\|_2<\infty$.

$\underline{H_0}$: $x_0(\cdot)$ is continuous from G into $\operatorname{dom}\varphi(\cdot\,,p)\subseteq H$ and for every $p\in G$,

$$x_0(p)\in\bigcap_{p'\in G}\overline{\operatorname{dom}\partial\varphi(\cdot\,,p')}.$$

Note that under the above hypotheses, given a parameter $p \in G$, we know that $(*)_1$ admits a unique strong $x(p)(\cdot) \in C(T, H)$ (see Kravvaritis–Papageorgiou [14]).

We will also need the following hypothesis on the cost integrand

$\underline{H(L):}$ $L : T \times H \times H \times G \to \bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is an integrand s.t.

(*1*) $(t, x, y, p) \to L(t, x, y, p)$ is measurable,
(*2*) $(x, y, p) \to L(t, x, y, p)$ is l.s.c.,
(*3*) $y \to L(t, x, y, p)$ is convex,
(*4*) $\varphi(t) - M(\|x\| + \|y\|) \le L(t, x, y, p)$ a.e. for all $p \in G$ and with $\varphi(\cdot) \in L^1$, $M > 0$.

Theorem 3.1. *If hypotheses $H(\varphi)$, $H(f)$, H_0 and $H(L)$ hold, then there exists $p \in G$ s.t. $J(p) = m_1$.*

Proof. Let $\{p_n\}_{n \ge 1} \subseteq G$ be a minimizing sequence; i.e. $J(p_n) \downarrow m_1$. By passing to a subsequence if necessary, we may assume that $p_n \to p$ in G. Let $x_n = x(p_n) \in C(T, H)$ be the unique trajectory corresponding to the parameter $n \ge 1$. From Benilan's inequality (see for example Brezis [10], Lemma 3.1, p. 64), we have

$$\|x_n(t) - S(t, p_n)x_0(p_n)\| \le \int_0^t \|f(s, x_n(s), p_n)\| ds, \ t \in T,$$

where $\{S(t, p_n)\}_{t \in T}$ is the semigroup of nonlinear contractions generated by the maximal monotone operator $\partial\varphi(\cdot\,, p_n)$. So we get

$$\|x_n(t)\| \le \|S(t, p_n)x_0(p_n)\| + \int_0^t \|f(s, x_n(s), p_n)\| ds$$

$$\le \|S(t, p_n)x_0(p_n)\| + \int_0^t [a(s, p_n) + b(s, p_n)\|x_n(s)\|] ds.$$

Note that $\|S(t, p_n)x_0(p_n) - S(t, p)x_0(p)\| \le \|S(t, p_n)x_0(p_n) - S(t, p_n)x_0(p)\| + \|S(t, p_n)x_0(p) - S(t, p)x_0(p)\| \le \|x_0(p_n) - x_0(p)\| + \|S(t, p_n)x_0(p) - S(t, p)x_0(p)\| \to 0$ as $n \to \infty$ uniformly in $t \in T$ (see hypothesis H_0 and Theorem 4.2, p. 120 of Brezis [10]). Thus we can find $M_1 > 0$ s.t. for all $t \in T$ and all $n \ge 1$, we have $\|S(t, p_n)x_0(p_n)\| \le M_1$.

Hence we get

$$\|x_n(t)\| \le M_1 + \sqrt{r}\sup_{n\ge 1}\|a(\cdot\,,p_n)\|_2 + \int\limits_0^r b(s,p_n)\|x_n(s)\|ds.$$

Invoking Gronwall's inequality, we get

$$\|x_n(t)\| \le (M_1 + \sqrt{r}\sup_{n\ge 1}\|a(\cdot\,,p_n)\|_2)\exp\left(\sqrt{r}\sup_{n\ge 1}\|b(\cdot\,,p_n\|_2\right) = M_2 < \infty,$$

for all $t \in T$ and all $n \ge 1$ (see hypothesis $H(f)(4)$).

Also from Theorem 3.6, p. 72 of Brezis [10], we get

$$\begin{aligned}\int\limits_0^r \|\dot{x}_n(t)\|^2 dt &\le \left(\int\limits_0^r \|f(t,x_n(t),p_n)\|^2 dt\right)^{1/2} + \sqrt{\varphi(x_0(p_n),p_n)}\\ &\le \|a(\cdot\,,p_n) + b(\cdot\,,p_n)M_2\|_2 + \sqrt{\varphi(x_0(p_n),p_n)}\\ &\le \|a(\cdot\,,p_n)\|_2 + M_2\|b(\cdot\,,p_n)\|_2 + \sqrt{\varphi(x_0(p_n),p_n)}\\ &\Longrightarrow \|\dot{x}_n\|_{L^2(H)} \le M_3 < \infty\end{aligned}$$

for some $M_3 > 0$ and for all $n \ge 1$ (see hypotheses $H(\varphi)(3)$ and $H(f)(4)$. Since $x_n(\cdot)$ is a strong solution, by Lebesgue's differentiation theorem, we have

$$x_n(t') - x_n(t) = \int\limits_t^{t'} \dot{x}_n(s)ds$$

$$\Longrightarrow \|x_n(t') - x_n(t)\| \le \int\limits_t^{t'} \|\dot{x}_n(s)\|ds \le \int\limits_0^r \chi_{[t,t']}(s)\|\dot{x}_n(s)\|ds \le (\sqrt{t'-t})M_3$$

$$\Longrightarrow \{x_n(\cdot)\}_{n\ge 1} \subseteq C(T,H) \text{ is equicontinuous.}$$

Furthermore invoking once again Theorem 3.6, p. 72 of Brezis [10], we have

$$\|\dot{x}_n(t)\|^2 + \frac{d}{dt}\varphi(x_n(t),p_n) = (f(t,x_n(t),p_n),\dot{x}_n(t)) \text{ a.e.}$$

$$\Longrightarrow \varphi(x_n(t), p_n) \le \int_0^t \|f(s, x_n(s), p_n)\| \cdot \|\dot{x}_n(s)\| ds + \varphi(x_0(p_n), p_n)$$
$$\le (\|a(\cdot\,, p_n)\|_2 + M_2\|b(\cdot\,, p_n)\|_2)M_3 + \varphi(x_0(p_n), p_n).$$

So from hypotheses $H(\varphi)(3)$ and $H(f)(4)$, we get that there exists $M_4 > 0$ s.t.

$$\varphi(x_n(t), p_n) \le M_4$$

for all $t \in T$ and all $n \ge 1$. Therefore we have:

$$\{x_n(t)\}_{n\ge1} \subseteq \bigcup_{n\ge1} \{y \in H : \|y\|^2 + \varphi(y, p_n) \le M_2{}^2 + M_4 = M_5\}$$
$$\Longrightarrow \{\overline{x_n(t)}\}_{n\ge1} \text{ is compact (see hypothesis } H(\varphi)(1)).$$

Thus applying the Arzela-Ascoli theorem, we deduce that $\{x_n(\cdot)\}_{n\ge1}$ is relatively compact in $C(T, H)$. So by passing to a subsequence if necessary, we get that $x_n \xrightarrow{s} x$ in $C(T, H)$. Then because of hypothesis $H(f)$, we have

$$\int_0^r \|f(t, x_n(t), p_n) - f(t, x(t), p)\| dt \to 0 \text{ as } n \to \infty.$$

So from Theorem 3.16, p. 102 of Brezis [10], we get that

$$-\dot{x}(t) \in \partial\varphi(x(t), p) + f(t, x(t), p) \text{ a.e.}$$
$$x(0) = x_0(p).$$

Also since $\|\dot{x}_n\|_{L^2(H)} \le M_3$, we may assume that $\dot{x}_n \xrightarrow{w} y$ in $L^2(H)$. Clearly $y = \dot{x}$. Then because of hypothesis $H(L)$, we can apply Theorem 2.1 of Balder [5] and get that

$$\int_0^r L(t, x(t), \dot{x}(t), p) dt \le \underline{\lim} \int_0^r L(t, x_n(t), \dot{x}_n(t), p_n) dt = \underline{\lim} J(p_n) = m_1.$$

But we already know that $x(\cdot) = x(p)(\cdot) \in C(T, H)$ solves $(*)_1$ for the parameter value $p \in G$. So $J(p) = m_1$. $\square$

Remark. A similar problem was considered by Ahmed [2] (see Theorem 4.8) for systems driven by m-accretive operators with $f = 0$ and L independent of the derivative $\dot{x}$.

Another closely related family of problems that can be treated by similar methods are the so-called "differential variational inequalities", that we encounter in mathematical economics and mechanics. Recall that if $K \subseteq \mathbb{R}^N$ is nonempty and closed, and $x \in K$, then the tangent cone to K at x denoted by $N_K(x)$ is defined by $N_K(x) = \partial\delta_K(x)$, where $\delta_K(\cdot)$ is the convex function $\delta_K(z) = 0$ if $z \in K$ and $+\infty$ otherwise (indicator function), and $\partial\delta_K(\cdot)$ is its subdifferential in the sense of convex analysis.

The system under consideration has state space $\mathbb{R}^N$ and is monitored by the following "differential variational inequality"

$$(*)_2 \qquad \begin{aligned} -\dot{x}(t) \in N_{K(t,p)}(x(t)) + f(t, x(t), p) \text{ a.e.} \\ x(0) = x_0(0). \end{aligned}$$

Denote by $P_{kc}(\mathbb{R}^N)$ the space of all nonempty, compact and convex subsets of $\mathbb{R}^N$. Equipped with the Hausdorff metric $h(A, B) = \max[\sup_{a\in A} d(a, B) \sup_{b\in B} d(b, A)], P_{kc}(\mathbb{R}^N)$ becomes a Polish space (i.e. is separable and complete).

We will need the following hypotheses on the data of $(*)_2$.

$\underline{H(K)}$: $K : T \times G \to P_{kc}(\mathbb{R}^N)$ is a multifunction s.t. $K(t, \cdot)$ is h-continuous and for all $(t, s) \in T \times T$, $t \geq s$ and all $p \in G$, $h(K(t, p), K(s, p)) \leq \int_s^t \eta(\tau)d\tau$, with $\eta(\cdot) \in L^2_+$.

$\underline{H(f)_1}$: $f : T \times \mathbb{R}^N \times G \to \mathbb{R}^N$ is a function s.t.

(*1*) $t \to f(t, x, p)$ is measurable,

(*2*) $\|f(t, x', p) - f(t, x, p)\| \leq k(t)\|x' - x\|$ a.e. for all $p \in G$ and with $k(\cdot) \in L^1_+$,

(*3*) $p \to f(t, x, p)$ is continuous,

(*4*) $\|f(t, x, p)\| \leq a(t, p) + b(t, p)\|x\|$ a.e. for all $p \in G$, with $a(\cdot\,, p), b(\cdot\,, p) \in L^2_+$ and s.t. $\sup_{p\in G} \|a(\cdot\,, p)\|_2, \|b(\cdot\,, p)\|_2 < \infty$.

$\underline{H_0^1}$: $x_0 : G \to \mathbb{R}^N$ is continuous and for all $p \in G$, $x_0(p) \in K(p)$.

Again our goal is to minimize the integral cost functional $J_1(p) = \int_0^r L(t, x(p)(t), \dot{x}(p)(t), p)dt$, using the trajectories of system $(*)_2$. The hypothesis on the integrand $L : T \times \mathbb{R}^N \times \mathbb{R}^N \times G \to \bar{\mathbb{R}}$ is the same as before, namely:

$\underline{H(L)_1:}$ $L : T \times \mathbb{R}^N \times \mathbb{R}^N \times G \to \bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is an integrand s.t.

(*1*) $(t, x, y, p) \to L(t, x, y, p)$ is measurable,

(*2*) $(x, y, p) \to L(t, x, y, p)$ is l.s.c.,

(3) $y \to L(t,x,y,p)$ is convex,

(4) $\varphi(t) - M(\|x\| + \|y\|) \leq L(t,x,y,p)$ a.e. for all $p \in G$, with $\varphi(\cdot) \in L^1$ and $M > 0$.

Theorem 3.2. *If hypotheses $H(K)$, $H(f)_1$, H_0^1 and $H(L)_1$ hold, then there exists $p \in G$ s.t. $J_1(p) = m_1$.*

Proof. Again let $\{p_n\}_{n\geq 1}$ be a minimizing sequence; i.e. $J(p_n) \downarrow m_1$. As before let $x_n(\cdot) = x(p_n)(\cdot) \in C(T, \mathbb{R}^N)$. From Moreau [14], we know that

$$\|\dot{x}_n(t)\| \leq \eta(t) + 2\|f(t, x_n(t), p_n)\| \text{ a.e.}$$

$$\Longrightarrow \|\dot{x}_n(t)\| \leq \eta(t) + 2a(t,p_n) + b(t,p_n)\|x_n(t)\| \text{ a.e.}$$

But note that $x_n(t) \in K(T,G) = \bigcup_{(t,p)\in T\times G} K(t,p) \in P_k(\mathbb{R}^N)$ (see hypothesis $H(K)$ and Theorem 7.4.2, p. 90 of Klein-Thompson [11]). So we can find $M_1 > 0$ s.t. for all $t \in T$ and all $n \geq 1$, we have, $\|x_n(t)\| \leq M_1$. Then

$$\|\dot{x}_n(t)\| \leq \eta(t) + 2a(t,p_n) + 2b(t,p_n)M_1 \text{ a.e.}$$

So using hypothesis $H(f)_1(4)$, we can find $M_2 > 0$ s.t. for all $n \geq 1$

$$\|\dot{x}_n\|_{L^2(T,\mathbb{R}^N)} \leq M_2.$$

Using this, we get that for all $t, t' \in T$ and all $n \geq 1$, we have

$$\|x_n(t') - x_n(t)\| \leq \int_t^{t'} \|\dot{x}_n(s)\| ds \leq (\sqrt{t'-t})M_2$$

$$\Longrightarrow \{x_n(\cdot)\}_{n\geq 1} \subseteq C(T,\mathbb{R}^N) \text{ is equicontinuous and bounded.}$$

Invoking the Arzela–Ascoli theorem, we deduce that $\{x_n(\cdot)\}_{n\geq 1}$ is relatively compact in $C(T,\mathbb{R}^N)$. So by passing to subsequence if necessary, we may assume that $x_n \xrightarrow{s} x$ in $C(T,\mathbb{R}^N)$ and $\dot{x}_n \xrightarrow{w} \dot{x}$ in $L^2(T,\mathbb{R}^N)$. Then $\int_0^b \|f(t,x_n(t),p_n) - f(t,x(t),p)\|dt \to 0$ as $n \to \infty$. Also note that because of hypothesis $H(K)$, $K(t,p_n) \xrightarrow{K} K(t,p)$ where $\xrightarrow{K}$ denotes the Kuratowski mode of set convergence) $\Rightarrow \delta_{K(t,p_n)} \xrightarrow{r} \delta_{K(t,p)}$ (where $\xrightarrow{r}$ denotes the epigraphical convergence, i.e. $\operatorname{epi}\delta_{K(t,p_n)} \xrightarrow{K} \operatorname{epi}\delta_{K(t,p)}$, with

$\operatorname{epi}\delta_{K(t,p_n)} = \{(x,\lambda) \in \mathbb{R}^N \times \mathbb{R} : \delta_{K(t,p_n)}(x) \le \lambda\}$ and similarly for $\operatorname{epi}\delta_{K(t,p)}) \Rightarrow \partial\delta_{K(t,p_n)} = N_{K(t,p_n)} \to \partial\delta_{K(t,p)} = N_{K(t,p)}$ in $\mathcal{M}_R(\mathbb{R}^N)$ (see Theorems 3.62 and 3.66 of Attouch [3]). $\Rightarrow GrN_{K(t,p_n)} \xrightarrow{K} GrN_{K(t,p)}$. Thus using Theorem 3.1 of [19], we get that $x(\cdot) = x(p)(\cdot) \in C(T, \mathbb{R}^N)$; i.e. $-\dot{x}(t) \in N_{K(t,p)}(x(t)) + f(t, x(t), p)$ a.e. $x(0) = x_0(p)$. Then via Balder's lower semicontinuity theorem, we get $J(p) = m_1 \Rightarrow p$ is the desired optimal parameter value. $\square$

4. Necessary conditions

In this section we derive necessary conditions for parameter optimality for a somewhat different class of systems. The maximal monotone operator is no longer multivalued, but we allow it to be time dependent. We work within the framework of evolution triples.

So let $T = [0,r]$ and let (X, H, X^*) be an evolution triple, with all embeddings being compact. Also G (the parameter set), is a nonempty, bounded, closed and convex set in a Banach space Y. The problem under consideration is the following:

$$(*)_3 \qquad \begin{aligned} J(x,p) = \ell(x(b)) + \int_0^r L(t, x(t), p)dt \to \inf = m_2 \\ s.t.\, \dot{x}(t) + A(t,p)x(t) = f(t, x(t), p) \ a.e. \\ x(0) = x_0. \end{aligned}$$

We will need the following hypotheses on the data:

$\underline{H(A)}$: $A : T \times G \to \mathcal{L}(X, X^*)$ is a map s.t.

(*1*) $t \to A(t,p)x$ is measurable,
(*2*) $p \to A(t,p)x$ is continuous,
(*3*) $\|A(t,p)x\|_* \le c\|x\|$ for all $(t,p) \in T \times G$, with $c > 0$,
(*4*) $\langle A(t,p)x, x\rangle \ge c_1\|x\|^2$ for all $(t,p) \in T \times G$, with $c_1 > 0$,
(*5*) $p \to A(t,p)x$ is weakly Gateaux differentiable and $\|A(t,p') - A(t,p)\|_* \le k_1(t)\|p' - p\|_Y$ a.e. with $k_1(\cdot) \in L^2_+$.

$\underline{H(f)_1}$: $f : T \times H \times Y \to H$ is a map s.t.

(*1*) $t \to f(t,x,p)$ is measurable,

(2) $|f(t,x',p')-f(t,x,p)| \le k_2(t)[\,|x'-x|+\|p'-p\|_Y]$ a.e. with $k_2(\cdot)\in L^2_+$,

(3) $f(t,\cdot,\cdot)$ is continuously Frechet differentiable, for every $h\in H$

$$(f'_x(t,x,p)h,h)\le 0 \ a.e.$$

and $|f'_x(t,x,p)\|_{\mathcal{L}(H)'}\|f'_p(t,x,p)\|_{\mathcal{L}(Y,H)} \le a_1(t)+b_1(|x|+\|p\|_Y)$ a.e. with $a_1(\cdot)\in L^2_+$, $b_1>0$,

(4) $|f(t,x,p)| \le a_2(t)+b_2(|x|+\|p\|_Y)$ a.e. with $a_2(\cdot)\in L^2_+$, $b_2>0$.

$\underline{H(L)_1}:$ $L: T\times H\times Y\to\mathbb{R}$ is a function s.t.

(1) $t\to L(t,x,p)$ is measurable,

(2) $(x,p)\to L(t,x,p)$ is continuous, convex,

(3) $\sup[L(t,x,p) : |x|\le n, \|p\|_Y\le n]\le\varphi_n(t)$ a.e. with $\varphi_n(\cdot)\in L^1_+$.

$\underline{H(\ell)}:$ $\ell: H\to\mathbb{R}$ is continuously, Frechet differentiable.

Suppose $\hat{p}\in G$ is the optimal parameter for $(*)_3$ and $\hat{x}(\cdot)=\hat{x}(p)(\cdot)\in W(T)\to C(T,H)$ the corresponding optimal state. By $\partial_x J(\hat{x},\hat{p})$ (resp. $\partial_p J(\hat{x},\hat{p})$) we denote the subdifferential of $J(\cdot,\hat{p})$ at $\hat{x}$ (resp. of $J(\hat{x},\cdot)$ at $\hat{p}$).

Theorem 4.1. *If hypotheses $H(A), H(f)_2, H(L)_1$ and $H(\ell)$ hold, then there exist $\psi(\cdot)\in W(T)$, $x^*\in\partial_x J(\hat{x},\hat{p})\subseteq L^2(H)$ and $p^*\in\partial_p J(\hat{x},\hat{p})\subseteq Y^*$ s.t.*

(1) $\dot{\hat{x}}(t)+A(t,p)\hat{x}(t)=f(t,\hat{x}(t),\hat{p})$ *a.e.* $\hat{x}(0)=x_0$

(2) $-\dot{\psi}(t)+A(t,\hat{p})^*\psi(t)-f'_x(t,\hat{x}(t),\hat{p})\psi(t)=x^*(t)$ *a.e.*
$\psi(b)=\ell_x(\hat{x}(b))$ *"adjoint equation"*

(3) $0 \le \int_0^b \langle\psi(t),(f'_p(t,\hat{x}(t),\hat{p})-A'_p(t,\hat{x}(t),\hat{p}))(p'-\hat{p}) > dt + (p^*,p'-\hat{p})_{Y^*,Y}$ *for all $p'\in G$ "maximum principle".*

Proof. Consider the map $p\to x(p)$ from G into $W(T)$, which to every parameter value $p\in G$ assigns the corresponding unique trajectory $x(p)(\cdot)\in W(T)$ of system $(*)_3$. We claim that this map is weakly Gateaux differentiable from above (i.e. as $\lambda\to 0^+$) at the point $\hat{p}\in G$.

To this end, let $p_\lambda = \hat{p} + \lambda(p' - \hat{p}) \in G$ for $p' \in G$, $\lambda \in [0,1]$ and let $x_\lambda(\cdot) = x(p_\lambda)(\cdot) \in W(T)$. We have:

$$\dot{x}_\lambda(t) + A(t, p_\lambda)x_\lambda(t) = f(t, x_\lambda(t), p_\lambda) \ a.e., \ x_\lambda(0) = x_0$$
$$\text{and } \dot{\hat{x}}(t) + A(t, \hat{p})\hat{x}(t) = f(t, \hat{x}(t), \hat{p}) \ a.e., \ \hat{x}(0) = x_0.$$

Subtract the second equation from the first and then act on the resulting equation with $x_\lambda(t) - x(t)$. We get, using the monotonicity of $A(t, p_\lambda)(\cdot)$:

$$(1)\ \langle \dot{x}_\lambda - \dot{\hat{x}}(t), x_\lambda(t) - \hat{x}(t)\rangle + \langle A(t,p_\lambda)x_\lambda(t) - A(t,\hat{p})\hat{x}(t), x_\lambda(t) - \hat{x}(t)\rangle$$
$$= (f(t, x_\lambda(t), p_\lambda) - f(t, \hat{x}(t), \hat{p}), x_\lambda(t) - \hat{x}(t)) \text{ a.e.}$$
$$\Longrightarrow \ \frac{1}{2}\frac{d}{dt}|x_\lambda(t) - \hat{x}(t)|^2 + c\|x_\lambda(t) - \hat{x}(t)\|^2$$
$$\leq \langle A(t,\hat{p})\hat{x}(t) - A(t,p_\lambda)\hat{x}(t), x_\lambda(t) - \hat{(t)}\rangle +$$
$$+(f(t, x_\lambda(t), p_\lambda) - f(t, \hat{x}(t), \hat{p}), x_\lambda(t) - \hat{x}(t)) \text{ a.e.}$$
$$\Longrightarrow \ |x_\lambda(t) - \hat{x}(t)|^2 + 2c\int_0^t \|x_\lambda(s) - \hat{x}(s)\|^2 ds$$
$$\leq 2\int_0^t \langle A(s,\hat{p})\hat{x}(s) - A(s,p_\lambda)\hat{x}(s), x_\lambda(s) - \hat{x}(s)\rangle ds$$
$$+2\int_0^t (f(s, x_\lambda(s), p_\lambda) - f(s, \hat{x}(s), \hat{p}), x_\lambda(s) - \hat{x}(s))ds$$
$$\leq 2\int_0^t \|A(s,\hat{p})\hat{x}(s) - A(s,p_\lambda)\hat{x}(s)\|_* \|x_\lambda(s) - \hat{x}(s)\| ds$$
$$+2\int_0^t |f(s, x_\lambda(s)p_\lambda) - f(s, \hat{x}(s), \hat{p})| \cdot |x_\lambda(s) - \hat{x}(s)| ds$$
$$\leq 2\int_0^t k_1(s)\lambda\|p' - \hat{p}\|_Y \|x_\lambda(s) - \hat{x}(s)\| ds$$
$$2\int_0^t k_2(s)|x_\lambda(s) - \hat{x}(s)|^2 ds + 2\int_0^t k_2(s)\|p' - p\|_Y |x_\lambda(s) - \hat{x}(s)| ds.$$

Applying Cauchy's inequality with $\epsilon > 0$ on the right hand side of the above inequality, we get

$$
\begin{aligned}
&2\int_0^t k_1(s)\lambda\|p' - \hat{p}\|_Y \|x_\lambda(s) - \hat{x}(s)\|ds \\
(2) \qquad &\le 2|G|\epsilon \int_0^t k_1(s)^2 ds + \frac{1}{\epsilon}\int_0^t \|x_\lambda(s) - \hat{x}(s)\|^2 ds
\end{aligned}
$$

where $|G| = \sup\{\|p\|_Y : p \in G\}$. Let $\epsilon = \frac{1}{2c}$ in (2). We get:

$$
\begin{aligned}
&2\int_0^t k_1(s)\lambda\|p' - \hat{p}\|_Y \|x_\lambda(s) - \hat{x}(s)\|ds \\
&\le \frac{|G|}{c}\|k_1\|_2^2 + 2c\int_0^t \|x_\lambda(s) - \hat{x}(s)\|^2 ds
\end{aligned}
$$

$$
\begin{aligned}
\implies |x_\lambda(t) - x(t)|^2 &\le \frac{|G|}{c}\|k_1\|_2^2 + 2\int_0^t k_2(s)|x_\lambda(s) - \hat{x}(s)|^2 ds \\
&+4|G|\int_0^t k_2(s)^2 ds + 4|G|\int_0^t |x_\lambda(s) - \hat{x}(s)|^2 ds
\end{aligned}
$$

$$
\implies |x_\lambda(t) - x(t)|^2
$$

$$
\begin{aligned}
&\le \frac{|G|}{c}\|k_1\|_2^2 + 4|G| \ \|k_2\|_2^2 + \int_0^t (2k_2(s) + 4|G|)|x_\lambda(s) - \hat{x}(s)|^2 ds \\
&= M_1 + \int_0^t k_3(s)|x_\lambda(s) - \hat{x}(s)|^2 ds
\end{aligned}
$$

where $M_1 = \frac{|G|}{c}\|k_1\|_2^2 + 4|G| \ \|k_2\|_2^2 > 0$ and $k_3(\cdot) = 2k_2(\cdot) + 4|G| \in L^2$. So we have:

$$
\frac{|x_\lambda(t) - \hat{x}(t)|^2}{\lambda^2} \le M_1 + \int_0^t k_3(s)\frac{|x_\lambda(s) - \hat{x}(s)|^2}{\lambda^2} ds.
$$

Invoking Gronwall's inequality, we deduce that there exists $M_2>0$ s.t. for all $\lambda \in (0,1)$ we have

$$(3) \qquad \frac{|x_\lambda(t)-\hat{x}(t)|^2}{\lambda^2} \le M_2 \text{ for all } t \in T.$$

Next on inequality (2) above, let $\epsilon = \frac{1}{c}$. We get

$$2\int_0^t k_1(s)\lambda\|p'-p\|_Y\|x_\lambda(s)-\hat{x}(s)\|ds \le \frac{2|G|}{c}\|k_1\|_2^2 + c\int_0^t \|x_\lambda(s)-\hat{x}(s)\|^2 ds.$$

Use this inequality in (1) to get

$$\begin{aligned} &c\int_0^t \|x_\lambda(s)-\hat{x}(s)\|^2 ds \\ &\le \frac{2|G|}{c}\|k_1\|_2^2 + 4|G|\,\|k_2\|_2^2 + \int_0^t k_3(s)\|x_\lambda(s)-\hat{x}(s)\|^2 ds \\ \Longrightarrow &\int_0^b \frac{\|x_\lambda(t)-\hat{x}(t)\|^2}{\lambda^2}dt \le M_3 + \int_0^b k_3(s)\frac{|x_\lambda(s)-\hat{x}(s)|^2}{\lambda^2}ds \\ &\le M_3 + \int_0^b k_3(s)M_2 ds \text{ (see (3) above),} \end{aligned}$$

where $M_3 = \frac{2|G|}{c}\|k_1\|_2^2 + 4|G|\,\|k_2\|_2^2$ and $k_3(\cdot) \in L^2_+$ as above. Thus we deduce that there exists $M_4 > 0$ s.t. for all $\lambda \in (0,1)$ we have

$$(4) \qquad \left\|\frac{x_\lambda - \hat{x}}{\lambda}\right\|_{L^2(X)} \le M_4.$$

Finally let $h \in L^2(X)$. We have

$$\begin{aligned} &\langle \dot{x}_\lambda(t) - \dot{\hat{x}}(t), h(t)\rangle + \langle A(t,p_\lambda)x_\lambda(t) - A(t,\hat{p})\hat{x}(t), h(t)\rangle \\ &\quad = (f(t,x_\lambda(t),p_\lambda) - f(t,\hat{x}(t),\hat{p}), h(t)) \text{ a.e.} \\ \Longrightarrow\ &\langle \dot{x}_\lambda(t) - \dot{\hat{x}}(t), h(t)\rangle \le [c\|x_\lambda(t)-\hat{x}(t)\| + k_1(t)\|p_\lambda - \hat{p}\|_Y + \\ &\quad +\beta k_2(t)|x_\lambda(t)-\hat{x}(t)|]\cdot\|h(t)\| \text{ a.e.} \end{aligned}$$

where $\beta > 0$ is such that $|\cdot| \leq \beta\|\cdot\|$. Such a $\beta > 0$ exists since by hypothesis X embeds into H continuously. Then we get

$$\left\langle \frac{\dot{x}_\lambda(t) - \dot{\hat{x}}(t)}{\lambda}, h(t) \right\rangle$$
$$\leq \left[c\frac{\|x_\lambda(t) - \hat{x}(t)\|}{\lambda} + k_1(t)\|p' - \hat{p}\|_Y + \beta k_2(t)\frac{|x_\lambda(t) - \hat{x}(t)|}{\lambda} \right] \cdot \|h(t)\| \text{ a.e.}$$

Denote by $((\cdot\,,\cdot))_0$ the duality brackets for the pair $(L^2(X), L^2(X^*))$. We have after integration and by applying on the right hand side Hölder's inequality:

$$\left(\left(\frac{\dot{x}_\lambda - \dot{\hat{x}}}{\lambda}, h\right)\right)_0 \leq \left[cM_4 + 2\sqrt{r}|G|\ \|k_1\|_2 + \beta M_2^{\frac{1}{2}} \|k_2\|_2 \right] \cdot \|h\|_{L^2(X)}.$$

Since $h \in L^2(X)$ was arbitrary, we deduce that there exists $M_5 > 0$ s.t. for all $\lambda \in (0,1)$

$$\left\| \frac{\dot{x}_\lambda - \dot{\hat{x}}}{\lambda} \right\|_{L^2(X^*)} \leq M_5. \tag{5}$$

From (4) and (5) above, we deduce that $\left\{\frac{x_\lambda - \hat{x}}{\lambda}\right\}_{\lambda \in (0,1)} \subseteq W(T)$ is bounded hence relatively weakly compact. So if $\lambda_n \to 0^+$, by passing to a subsequence if necessary, we may assume that $y_n = \frac{x_n - \hat{x}}{\lambda_n} \xrightarrow{w} v$ in $W(T)$. Thus for every $h \in L^2(X)$, we have

$$((\dot{y}_n, h))_0 + \left(\left(\frac{\hat{A}(p_n)x_n - \hat{A}(\hat{p})\hat{x}}{\lambda_n}, h\right)\right)_0 + \left(\left(\frac{\hat{f}(x_n, p_n) - \hat{f}(\hat{x}, \hat{p})}{\lambda_n}, h\right)\right)_0$$

where $\hat{A}(p')y(t) = A(t,p')y(t)$ and $\hat{f}(y,p')(t) = f(t,y(t),p')$ (i.e. $\hat{A}$ and $\hat{f}$ are the Nemitsky (superposition) operators, corresponding to the maps $A(t,x)$ and $f(t,x,p)$ respectively). We have:

$$((\dot{y}_n, y))_0 \to ((\dot{v}, h))_0 \text{ as } n \to \infty.$$

$$\text{Also } \left(\left(\frac{\hat{A}(p_n)x_n - \hat{A}(\hat{p})\hat{x}}{\lambda_n} h\right)\right)_0$$

$$= \left(\left(\frac{\hat{A}(p_n)x_n - \hat{A}(p_n)\hat{x}}{\lambda_n}, h\right)\right)_0 + \left(\left(\frac{\hat{A}(p_n)\hat{x} - \hat{A}(\hat{p})\hat{x}}{\lambda_n}, h\right)\right)_0$$

$$= ((\hat{A}(p_n)y_n, h))_0 = \left(\left(\frac{\hat{A}(p_n)\hat{x} - \hat{A}(p)\hat{x}}{\lambda_n}, h\right)\right)_0$$

$$\to ((\hat{A}(p)v, h))_0 + ((A'_p(\hat{x},\hat{p})(p'-\hat{p}), h))_0 \text{ as } n \to \infty$$

where $\hat{A}'_p(\hat{x},\hat{p})(p'-\hat{p})(t) = A(t,\hat{x}(t),\hat{p})(p'-\hat{p})$.

Finally note that $x_n \xrightarrow{w} \hat{x}$ in $W(T)$ and since $W(T) \to L^2(H)$ compactly, we have that $x_n \xrightarrow{s} \hat{x}$ in $L^2(H)$. So from hypothesis $H(f)$ and the total differential rule, we have that

$$\left(\left(\frac{\hat{f}(x_n,p_n) - \hat{f}(\hat{x},\hat{p})}{\lambda_n}, h\right)\right)_0 \to ((\hat{f}'_x(\hat{x},\hat{p})v + \hat{f}'_p(\hat{x},\hat{p})(p'-\hat{p}), h))_0$$

where $\hat{f}'_x(\hat{x},\hat{p})v(t) = f'_x(t,\hat{x}(t),\hat{p})v(t)$ and $\hat{f}'_p(\hat{x},\hat{p})(p'-\hat{p})(t) =$
$= f'_p(t,\hat{x}(t),\hat{p})(p'-\hat{p})$. Since $h \in L^2(X)$ was arbitrary, we get

$$\begin{aligned} &\dot{v}(t) + A(t,\hat{p})v(t) + A'_p(t),\hat{x}(t),\hat{p})(p'-\hat{p}) \\ &= f'_x(t,\hat{x}(t),\hat{p})v(t) + f'_p(t,\hat{x}(t),\hat{p})(p'-\hat{p}) \text{ a.e. } v(0) = 0. \end{aligned}$$

Because of hypothesis $H(f)$ and Theorem 30.A, p. 771 of Zeidler [24], we know that the above evolution equation has a unique solution $v(\cdot) \in W(T)$. This then establishes the uniqueness of the weak limit as $\lambda \to 0^+$ of $\frac{x_\lambda - \hat{x}}{\lambda}$, and so we have proved the weak differentiability from above of $p \to x(p)$.

Since by hypothesis $(\hat{x},\hat{p}) \in W(T) \times G$ is optimal and since $y_n(t) \xrightarrow{s} v(t)$ a.e. (this being a consequence of the compact embedding of $W(T)$ in $L^2(H)$), we get

$$0 \le \int_0^b L'(t,\hat{x}(t),\hat{p})(v(t),p'-\hat{p})dt + \ell_x(\hat{x}(b))(v(b))$$

for all $p' \in G$, where L' denotes the directional derivative of the convex function $L(t,\cdot,\cdot)$ (see Clarke [12] and recall that a continuous convex function is locally Lipschitz).

But from Rockafellar [23], we know that

$$J'(\hat{x},\hat{p})(v,p'-\hat{p}) = \int_0^b L'(t,\hat{x}(t),\hat{p})(v(t),p'-\hat{p})dt$$
$$= \sigma([v,p'-\hat{p}],\partial J(\hat{x},\hat{p}))$$

where $\partial J(\hat{x},\hat{p})$ is the convex subdifferential of the convex functional $J(\cdot\,,\cdot)$ at $(\hat{x},\hat{p})$ and $\sigma(\cdot\,,\partial J(\hat{x},\hat{p}))$ its support function (i.e. $\sigma([v,p'-\hat{p}],\partial J(\hat{x},\hat{p})) = \sup[(y^*,v)_{L^2(H)} + (g^*,p'-\hat{p})_{Y^*,Y} : [y^*,g^*] \in \partial J(\hat{x},\hat{p}) \subseteq L^2(H)\times Y^*])$. Note that since $J(\cdot,\cdot)$ is continuous and convex, $\partial J(\hat{x},\hat{p})$ is w^*-compact in $L^2(H)\times Y^*$ and so the supremum above is attained. Furthermore from Proposition 2.3.15, p. 48 of Clarke [12], we have $\partial J(\hat{x},\hat{p}) \subseteq \partial_x J(\hat{x},\hat{p}) \times \partial_p J(\hat{x},\hat{p})$ (here $\partial_x J(\hat{x},\hat{p})$ and $\partial_p J(\hat{x},\hat{p})$ are the subdifferentials in the sense of convex analysis of $J(\cdot,\hat{p})$ and $J(\hat{x},\cdot)$ respectively). Thus we can find $x^* \in \partial_x J(\hat{x},\hat{p}) \subseteq L^2(H)$ and $p^* \in \partial J_p(\hat{x},\hat{p}) \subseteq Y^*$ s.t.

$$0 \le \int_0^b (x^*(t),v(t))dt + (p^*,p'-\hat{p})_{Y^*,Y} + \ell'_x(\hat{x}(b))v(b).$$

Introduce the adjoint equation

$$-\dot{\varphi}(t) + A(t,\hat{p})^*\psi(t) - f'_x(t,\hat{x}(t),\hat{p})^*\psi(t) = x^*(t) \text{ a.e.}$$
$$\psi(b) = \ell'_x(\hat{x}(b)).$$

Because of hypotheses $H(A)$, $H(f)$ and Theorem 30.A, p. 771 of Zeidler [24], this evolution equation has a unique solution $\psi(\cdot) \in W(T)$. Then we have

$$0 \le \int_0^b \langle -\dot{\psi}(t) + A(t,\hat{p})^*\psi(t) - f'_x(t,\hat{x}(t),\hat{p})^*\psi(t), v(t)\rangle dt$$
$$+ (p^*,p'-\hat{p})_{Y^*,Y} + \ell'_x(\hat{x}(b))v(b).$$

Using the integration by parts formula for functions in $W(T)$ (see

Zeidler [24], Proposition 23.23, p. 422–423), we get

$$0 \leq \int_0^b \langle \psi(t), \dot{v}(t) \rangle dt - \ell'_x(\hat{x}(b))v(b) + \int_0^b \langle \psi(t), A(t,\hat{p})v(t) \rangle dt$$

$$- \int_0^b (\psi(t), f'_x(t, \hat{x}(t), \hat{p})v(t))dt + (p^*, p' - \hat{p}) + \ell'_x(\hat{x}(b))v(b)$$

$$\Longrightarrow \; 0 \leq \int_0^b \langle \psi(t), f'_p(t, \hat{x}(t), \hat{p})(p' - \hat{p}) - A'_p(t, \hat{x}(t), \hat{p})(p' - \hat{p}) \rangle dt$$

$$+ (p^*, p' - \hat{p}) \text{ for all } p' \in G$$

which completes the proof of the theorem. □

5. An example

In this section, we work in detail an example of parameter identification in parabolic distributed parameter system.

So let $T = [0, r]$ and Z a bounded domain in $\mathbb{R}^N$ with smooth boundary $\partial Z = \Gamma$. Also let G be a compact metric space (the parameter space).

The problem under consideration is the following:

$$J(p) = \int_0^r \int_Z L(t, z, x(t,z), p) dz\, dt \to \inf = m_3$$

$(*)_4$

$$\text{s.t. } \frac{\partial x}{\partial t} - \sum_{i,j=1}^{N} \frac{\partial}{\partial z_j} \left(u(z_1, p) \frac{\partial x(t,z)}{\partial z_i} \right) + f(t, z, x(t,z), p) = 0$$

$$\text{a.e. on } T \times Z \quad x|_{T \times \Gamma} = 0, \; x(0, z, p) = x_0(z, p).$$

We will need the following hypotheses on the data of $(*)_4$.

$H(u)$: For every $(z, p) \in \mathbb{R}^N \times G$, $m_0 |z|^2 \leq \sum_{i,j=1}^{N} u(z_1, p) z_i z_j \leq M_0 |z|^2$

$0 < m_0 < M_0$, $|u(z,p)| \leq k$, $z_1 \in Z_1 = \mathrm{proj}_1 Z$ and if $p_n \to p$ in G, then $\frac{1}{u(\cdot, p_n)} \xrightarrow{w} \frac{1}{u(\cdot, p)}$ in $L^2(Z_1)$.

$H(f)_2$: $f : T \times Z \times \mathbb{R} \times G \to \mathbb{R}$ is a function s.t.

(1) $(t, z) \to f(t, z, x, p)$ is measurable,

(2) $|f(t,z,x',p) - f(t,z,x,p)| \le k(t,z)|x'-x|$ a.e. for all $p \in G$ and with $k(\cdot,\cdot) \in L^1(T \times Z)$,

(3) $p \to f(t,z,p)$ is continuous,

(4) $|f(t,z,x,p)| \le a(t,z,p) + b(t,z,p)|x|$ a.e with $a(\cdot,\cdot,p)$, $b(\cdot,\cdot,p) \in L^2(T \times Z)$ s.t. $\sup\limits_{p\in G} \|a(\cdot,\cdot,p)\|_2 \|b(\cdot,\cdot,p)\|_2 < \infty$.

$\underline{H(L)_2}$: $L : T \times Z \times \mathbb{R} \times G \to \bar{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is an integrand s.t.

(1) $(t,z,x,p) \to L(t,z,x,p)$ is measurable,

(2) $(x,p) \to L(t,z,x,p)$ is l.s.c,

(3) $x \to L(t,z,x,p)$ is convex,

(4) $\varphi(t,z) - M(z)|x| \le L(t,z,x,p)$ a.e. for all $p \in G$ and with $\varphi \in L^1(T \times Z)$, $M \in L^\infty_+(Z)$.

$\underline{H_0^1}$: $x_0(\cdot,p) \in H_0^1(Z)$ and $\sup\limits_{p\in G} \|x_0(\cdot,p)\|_{H_0^1(Z)} < \infty$.

Let $H = L^2(Z)$ and define $\varphi : L^2(H) \times G \to \mathbb{R}$ by

$$\varphi(x,p) = \int\limits_Z \sum_{i,j=1}^N u(z_1,p) \frac{\partial x}{\partial z_i} \frac{\partial x}{\partial z_j} dz, \text{ if } x \in H_0^1(Z)$$
$$= +\infty \text{ if } x \in L^2(Z) H_0^1(Z).$$

Also let $a(p) : H_0^1(Z) \times H_0^1(Z) \to \mathbb{R}$ be the Dirichlet form defined by

$$a(p)(x,y) = \int\limits_Z \sum_{i,j=1}^N u(z_1,p) D_i x D_j y \, dz$$

where $D_i = \frac{\partial}{\partial z_i}$, $D_j = \frac{\partial}{\partial z_j}$. Then this bilinear form $a(p)$ $(\cdot,\cdot)$ is clearly continuous (i.e. $|a(p)(x,y)| \le c\|x\|_{H_0^1(Z)} \ \|y\|_{H_0^1(Z)}, \ c > 0$), and so there exists $A(p) \in \mathcal{L}(H_0^1(Z), \ H^{-1}))$ s.t. $a(p)(x,y) = \langle A(p)x, y\rangle$, with $\langle\cdot,\cdot\rangle$ denoting the duality brackets of the pair $(H_0^1(Z), H^{-1}(Z))$. The operator is the energetic extension of $\partial\varphi(\cdot,p)$ and $\partial\varphi(x,p) = A(p)x$ for $x \in \operatorname{dom} \partial\varphi(\cdot,p) = \{x \in H_0^1(Z) : \partial\varphi(x,p) = \nabla\varphi(x,p) \in L^2(Z) = H\}$, which is dense in H.

From Theorem 29 of Zhikov–Kozlov–Oleinik [21], we know that if $p_n \to p$ in $G \Rightarrow A(p_n) \xrightarrow{G} A(p) \Rightarrow \partial\varphi(\cdot,p_n) \to \partial\varphi(\cdot,p)$ in $\mathcal{M}_R(H)$ (see Attouch [3], Proposition 3.69, p. 379).

Observe that for every $x \in H_0^1(Z)$, we have

$$\varphi(x,p) \ge m_0 \|\nabla x\|_2^2$$

for all $p \in G$. So if $\lambda > 0$, the set

$$H_\lambda(p) = \{x \in L^2(Z) : \|x\|_2^2 + \varphi(x,p) \leq \lambda\}$$

is bounded in $H_0^1(Z)$, uniformly in $p \in G \Rightarrow \bigcup\limits_{p\in G} H_\lambda(p)$ is bounded in $H_0^1(Z)$. But by the Sobolev embedding theorem $H_0^1(Z) \to L^2(Z) = H$ compactly. So $\overline{\bigcup\limits_{p\in G} H_\lambda(p)}$ is compact in $L^2(Z)$. Furthermore because of hypothesis H_0^1, we have $\sup\limits_{p\in G} \varphi(x_0(p),p) < \infty$. So we have satisfied hypotheses $H(\varphi)$ and H_0 of Section 3.

Next let $\hat{f} : T \times H \times G \to H$ be defined by

$$\hat{f}(t,x,p)(z) = f(t,z,x(z),p),$$

i.e. $\hat{f}(t,x,p)$ is the Nemitsky operator corresponding to $f(t,z,x,p)$. Then using hypothesis $H(f)_2$ and Krasnosel'skii's theorem, we conclude that $\hat{f}(t,x,p)$ satisfies hypothesis $H(f)$.

Finally let $\hat{L} : T \times L^2(Z) \times G \to \bar{\mathbb{R}}$ be defined by

$$\hat{L}(t,x,p) = \int\limits_Z L(t,z,x(z),p)dz.$$

From Theorem 1 of Pappas [18], we know that we can find $L_k : T\times Z\times\mathbb{R}\times G \to \mathbb{R}$ Caratheodory integrands (i.e., $L(\cdot\,,\cdot,x,p)$ is measurable, $L(t,z,\cdot,\cdot)$ is continuous), $\varphi(t,z) - V(z)|x| \leq L_k(t,z,x,p) \leq k$ for all $p \in G$ and $L_k \uparrow L$ as $k \to \infty$. Set

$$\hat{L}_k(t,x,p) = \int\limits_Z L_k(t,z,x(z),p)dz.$$

It is clear that $\hat{L}_k(t,x,p)$ is measurable in t, continuous in (x,p), thus jointly measurable. Furthermore from the monotone convergence theorem, we have $\hat{L}_k \uparrow \hat{L} \Rightarrow \hat{L}$ is jointly measurable. Also from Theorem 2.1 of Balder [5], we get that $\hat{L}(t,\cdot,\cdot)$ is l.s.c., while $\hat{L}(t,\cdot,p)$ is convex. Finally, $\hat{\varphi}(t) - \hat{M}\|x\|_{L^2(Z)} \leq \hat{L}(t,x,p)$ a.e. with $\hat{\varphi}(t) = \|\varphi(t,\cdot)\|_1$ and $\hat{M} = \|M(\cdot)\|_\infty$. So we have satisfied hypothesis $H(L)$.

New rewrite $(*)_4$ in the following equivalent abstract form:

$$(*)_4' \qquad \begin{aligned} &\hat{J}(x,u) = \int_0^r \hat{L}(t,x(t),p)dt \to \inf = m_3 \\ &s.t. \ -\dot{x}(t) \in \partial\varphi(x(t),p) + \hat{f}(t,x,(t),p) \\ &\qquad\qquad x(0) = x_0(p). \end{aligned}$$

We have checked that this system satisfies all hypotheses of Theorem 3.1. So invoking that theorem, we get:

Theorem 5.1. *If hypotheses $H(u)$, $H(f)_2$, $H(L)_2$ and H_0^1 hold, then problem $(*)_4$ admits a solution.*

Remark. The hypothesis $H(u)$ is of course satisfied if $u(z,p_n) \to u(z,p)$ a.e. However more generally, our hypothesis allows oscillatory behavior in the coefficients. Such problems arise often in applications. We particularly mention the problem of homogenization (i.e., the problem of finding the physical properties of a homogeneous material whose overall response is close to that of the periodic material, when the size ϵ of the periodicity cell goes to zero; (see Attouch [3])) and the problem of shape optimization and sensitivity analysis of unilateral problems (see Sokolowski–Zolezio [26]).

Acknowledgement. The author would like to thank the anonymous referee for his (her) constructive criticisms.

References

[1] N.U. Ahmed, Identification of linear operators in differential equations on Banach space, Operator Methods for Optimal Control Problems (S–J Lee, ed.), vol. 108, *Marcel Dekker*, *New York*, 1987, pp. 1–35.

[2] N.U. Ahmed, Optimization and Identification of Systems Governed by Evolution Equations on Banach Space, Pitman Research Notes in Mathematical Series, vol. 184, *Longman, Essex, U.K.*, 1988.

[3] H. Attouch, Variational Convergence for Functionals and Operators, *Pitman*, *London*, 1984.

[4] J.-P. Aubin and A. Cellina, Differential Inclusions, *Springer*, *Berlin*, 1984.

[5] E. Balder, Necessary and sufficient conditions for L_1-strong-weak lower semicontinuity of integral functionals, *Nonl. Anal.-TMA* **11** (1987), 1399–1404.

[6] H.T. Banks, S. Reich and I. Rosen, Parameter estimation of nonlinear distributed systems-Approximation theory and convergence results, *Appl. Math. Letters* **1** (1988), 211–216.

[7] H.T. Banks, S. Reich and I.G. Rosen, An Approximation theory for the identification of nonlinear distributed parameter systems, *SIAM J. Control and Optimization* **28** (1990), 552–569.

[8] H.T. Banks, S. Reich and I.G. Rosen, Galerkin approximation for inverse problems for nonautonomous nonlinear distributed systems, *Applied Math. and Optimization* **24** (1991), 233–256.

[9] V. Barbu, Nonlinear Semigroups and Differential Equations in Banach Spaces, Noordhoff International Publishing, *Leiden*, *The Netherlands*, 1976.
[10] H. Brezis, Operateures Maximaux Monotones, *North Holland*, *Amsterdam*, 1973.
[11] G. Chavent, On the identification of distributed parameter systems, Proc. of the 5th IFAC Symposium on Identification and System Parameter Estimation, *Darmstadt, Federal Republic of Germany*, 1979.
[12] F. H. Clarke, Optimization and Nonsmooth Analysis, *Wiley*, *New York*, 1983.
[13] E. Klein and A. Thompson, Theory of Correspondences, *Wiley*, *New York*, 1984.
[14] D. Kravvaritis and N. S. Papageorgiou, Multivalued perturbations of subdifferential type evolution equations in Hilbert spaces, *J. Diff. Equations* **76** (1988), 238–255.
[15] J.-L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, *Springer*, *Berlin*, 1971.
[16] J.-J. Moreau, Evolution problem associated with a moving convex set in a Hilbert space, *J. Diff. Equations* **26** (1977), 347–374.
[17] U. Mosco, Convergence of convex sets and solutions of variational inequalities, *Advances in Math.* **3** (1969), 510–585.
[18] E. Nagy, A theorem on compact embedding for functions with values in an infinite dimensional Hilbert space, *Annales Univ. Sci. Section Math.* **23** (1980), 243–245, *Budapest.*
[19] N. S. Papageorgiou, Convergence theorems for Banach space valued integrable multifunctions, *Intern. J. Math. and Math. Sci.* **10** (1987), 433–442.
[20] N. S. Papageorgiou, Continuous dependence results for a class of evolution inclusion, *Proceedings of the Edinburgh Math. Soc.* **35** (1992), 139–158.
[21] N. S. Papageorgiou, On the solution set of nonlinear evolution inclusions depending on a parameter, *Publicationes Math. Debrecen* **44** (1994), 31–49.
[22] G. Pappas, An approximation result for normal integrands and application to relaxed controls theory, *J. Math. Anal. Appl.* **93** (1983), 132–141.
[23] R. T. Rockafellar, Conjugate Duality and Optimization, *SIAM Publications*, *Philadelphia*, 1974.
[24] E. Zeidler, Nonlinear Functional Analysis and its Applications II, *Springer*, *New York*, 1990.
[25] V. Zhikov, S. Kozlov and O. Oleinik, G-convergence of parabolic operators, *Russian Math. Surveys* **36** (1981), 9–60.
[26] J. Sokolowski and J.-P. Zolezio, Shape sensitivity analysis of unilateral problems, *SIAM J. Math. Anal.* **18** (1987), 1416–1437.

NIKOLAOS S. PAPAGEORGIOU
NATIONAL TECHNICAL UNIVERSITY
DEPARTMENT OF MATHEMATICS
ZOGRAFOU CAMPUS
ATHENS 15780, GREECE (PERMANENT ADDRESS)

FLORIDA INSTITUTE OF TECHNOLOGY
DEPARTMENT OF APPLIED MATHEMATICS
150 W. UNIVERSITY BLVD.
MELBOURNE, FLORIDA 32901–6988, USA

(Received June 23, 1993; revised August 1, 1994)