Zeitschrift für Analysis und ihre Anwendungen
Journal for Analysis and its Applications
Volume 22 (2003), No. 4, 863–898

Existence and Relaxation for Finite-Dimensional Optimal Control Problems Driven by Maximal Monotone Operators

N. S. Papageorgiou and F. Papalini

Abstract. In this paper we study the optimal control of a class of nonlinear finite-dimensional optimal control problems driven by a maximal monotone operator which is not necessarily everywhere defined. So our model problem incorporates systems monitored by variational inequalities. First we prove an existence theorem using the reduction method of Berkovitz and Cesari. This requires a convexity hypothesis. When this convexity condition is not satisfied, we have to pass to an augmented, convexified problem known as the "relaxed problem". We present four relaxation methods. The first uses Young measures, the second uses multi-valued dynamics, the third is based on Carathéodory's theorem for convex sets in $\mathbb{R}^N$ and the fourth uses lower semicontinuity arguments and Γ-limits. We show that they are equivalent and admissible, which roughly speaking means that the corresponding relaxed problem is in a sense the "closure" of the original one.

Keywords: *Maximal monotone operator, variational inequalities, reduction method, relaxed problem, Young measure, Carathéodory's theorem, multi-valued dynamics, multiple* Γ*-limits,* Γ*-regularization, admissible relaxation*

AMS subject classification: 49J24, 49J40, 49J45

1. Introduction

In this paper we develop the existence and relaxation theory for a large class of nonlinear finite-dimensional optimal control problems. The dynamic equation of our system involves a maximal monotone operator which is not necessarily everywhere defined. This way we incorporate in our framework variational

---

N. S. Papageorgiou: Nat. Techn. Univ., Dept. Math., Zografou Campus, Athens 15780, Greece; npapg@math.ntua.gr
F. Papalini: Univ. of Ancona, Dept. Math., Via Brecce Bianche, Ancona 60131, Italy

ISSN 0232-2064 / $ 2.50

inequalities (problems with unilateral constraints). Also, in the existence part of our work the control constraint set is state dependent, i.e. there is an a priori feedback in the system, a situation which is of interest in engineering problems. Our approach in the existence part is based on the "reduction" technique, which was developed by the fundamental works of Berkovitz [2 - 4] and Cesari [8, 9]. This method reveals the importance of convex structure in order to guarantee the existence of optimal pairs. If this convex structure is missing, in order to have an existence theory, we need to pass to an augmented, convexified problem, which on the one hand is required to remain close to the original problem and on the other hand must be "convex" enough in order to have a solution. This delicate balance is achieved by the so-called "relaxed problem". Relaxation is the process of embedding the original problem to a larger one with sufficient convex structure which guarantees the existence of optimal pairs. There is no unique approach to relaxation. However, a proper relaxation procedure should meet the following three basic criteria:

- (a) Every original state should also be relaxed state.
- (b) The set of original states must be dense in the set of relaxed states.
- (c) The relaxed problem has a solution and the values of the original and relaxed problems must be equal.

The first two requirements refer to the dynamics of the system, while the third concerns the objective (cost) functional. The condition that the values of the two problems are equal is often called "relaxability". If it is not satisfied, then it can be maintained that we have augmented the original optimal control problem too much. If the relaxation method meets the above three criteria, it is said to be admissible. In the second half of this work we present four different relaxation methods and show that under reasonable conditions on the data they are admissible. The first method uses Young measures (transition probabilities) as relaxed controls, the second uses multi-valued dynamics and it is an outgrowth of the reduction method of the existence theory, the third is based on Carathéodory's theorem for the convex sets in $\mathbb{R}^N$ and the fourth uses lower semicontinuity arguments and is based on the Γ-regularization of the extended cost functional. This last method uses the notion of Γ-limit which was developed by De Giorgi and can be found in the books of Buttazzo [7], Dal Maso [11] and Hu and Papageorgiou [17]. A well written introduction to the subject of relaxation of optimization problems can be found in the book of Roubicek [24].

2. Mathematical preliminaries

In this section, for the convenience of the reader, we present the main items of the mathematical background needed to follow this paper. Our main references are the books of Hu and Papageorgiou [17, 18].

Let (Ω, Σ) be a measurable space and X a separable Banach space. We will use the notations

$$P_{f(c)}(X) = \Big\{A \subset X : A \text{ non-empty, closed (and convex)}\Big\}$$
$$P_{(w)k(c)}(X) = \Big\{A \subset X : A \text{ non-empty, (weakly-)compact (and convex)}\Big\}.$$

A multifunction $F : \Omega \to P_f(X)$ is said to be measurable, if for all $x \in X$ the function

$$\omega \to d(x, F(\omega)) = \inf_{y \in F(\omega)} \|x - y\|$$

is measurable. Also, $F : \Omega \to 2^X \setminus \{\emptyset\}$ is said to be graph measurable, if

$$\mathrm{Gr}F = \big\{(\omega, x) \in \Omega \times X : x \in F(\omega)\big\} \in \Sigma \times B(X)$$

with $B(X)$ being the Borel σ-field of X. For a multifunction with values in $P_f(X)$, measurability implies graph measurability, while the converse is true if Σ is complete, i.e. if $\Sigma = \widehat{\Sigma}$ is the universal σ-field.

Let μ be a finite measure on (Ω, Σ). For $F : \Omega \to 2^X \setminus \{\emptyset\}$ and $1 \leq p \leq \infty$ we introduce the sets

$$S_F = \Big\{f : \Omega \to X \text{ measurable } : f(\omega) \in F(\omega) \ \mu\text{-a.e.}\Big\}$$

and

$$S_F^p = \Big\{f \in L^p(\Omega, X) : f(\omega) \in F(\omega) \ \mu\text{-a.e.}\Big\}$$

which may be empty. For a graph measurable multifunction F, the set S_F^p is non-empty if and only if $\inf_{x \in F(\omega)} \|x\| \leq \varphi(\omega)$ μ-a.e. with $\varphi \in L^p(\Omega)$. Moreover, if μ is non-atomic, then the set S_F^p is closed or convex if and only if, for μ-almost all $\omega \in \Omega$, $F(\omega)$ is closed or convex, respectively. Finally, the set S_F^p is decomposable, i.e. $\chi_A f_1 + \chi_{A^c} f_2 \in S_F^p$ for all $(A, f_1, f_2) \in \Sigma \times S_F^p \times S_F^p$.

Let Y and Z be Hausdorff topological spaces. A multifunction $G : Y \to 2^Z$ is said to be lower semicontinuous or upper semicontinuous, if for all $C \subset Z$ closed the set $G^+(C) = \{y \in Y : G(y) \subset C\}$ or $G^-(C) = \{y \in Y : G(y) \cap C \neq \emptyset\}$ is closed, respectively. An upper semicontinuous multifunction with closed values has a closed graph, provided Z is a regular topological space. The converse is true, if G is locally compact, i.e. for every $y \in Y$ there is a neighborhood U of Y such that $\overline{G(U)}$ is compact in Z. If Z is a metric space with a metric d, then G is lower semicontinuous if and only if the function $y \to d(z, G(y))$ is upper semicontinuous for any $z \in Z$. Also, when Z is a metric space on $P_f(Z)$, we can define a generalized metric, known as the Hausdorff metric, by setting

$$h(A,B) = \max\left[\sup_{a\in A} d(a,B), \sup_{b\in B} d(b,A)\right]$$

for $A, B \in P_f(Z)$.

If Z is a complete metric space, then so is $(P_f(Z), h)$. A multifunction $G : Y \to P_f(Z)$ is said to be h-continuous, if it is continuous from Y into the metric space $(P_f(Z), h)$. Finally, a multifunction G which is both upper and lower semicontinuous is said to be continuous. For compact-valued multifunctions h-continuity and continuity coincide.

As we already mentioned in the introduction, one of the relaxation methods is based on the notion of Young measure (transition probability). So let us give the definition. Let (Ω, Σ, μ) be a finite measure space, Y a compact metric space and $M(Y)$ the Banach space of all bounded Borel measures with the total variation norm. From the Riesz representation theorem, $C(Y)^* = M(Y)$. A Young measure is a function $\lambda : \Omega \to M_+^1(Y)$ (with $M_+^1(Y)$ the subset of $M(Y)$ of all probability measures on Y) such that, for every $C \in B(Y)$, $\omega \to \lambda(\omega)(C)$ is Σ-measurable. We denote the set of all Young measures from Ω into Y by $R(\Omega, Y)$. It is easy to check that $\lambda \in R(\Omega, Y)$ if and only if $\lambda : \Omega \to M_+^1(Y)$ is Σ-measurable when $M_+^1(Y)$ is furnished with the relative weak* topology (see [23]). We know that $M_+^1(Y)$ topologized this way is Polish, too (cf. [23: p. 46]). The weak topology of $M_+^1(Y)$ has a natural analog on $R(\Omega, Y)$. So let

$$\mathrm{Car}(\Omega \times Y)$$

denote the space of all L^1-Carathéodory integrands, i.e. the set of all functions $\varphi : \Omega \times Y \to \mathbb{R}$ such that

- for all $y \in Y$, $\omega \to \varphi(\omega, y)$ is Σ-measurable for μ-a.a. $\omega \in \Omega$
- $y \to \varphi(\omega, y)$ is continuous
- for μ-a.a. $\omega \in \Omega$ and all $y \in Y$, $|\varphi(\omega, y)| \le h(\omega)$ with $h \in L_+^1(\Omega)$.

Then the weak topology on $R(\Omega, Y)$ is the initial topology with respect to which the functionals

$$\lambda \to I_\varphi(\lambda) = \int_\Omega \int_Y \varphi(\omega, y)\lambda(\omega)\,(dy)d\mu \qquad \big(\varphi \in \mathrm{Car}(\Omega \times Y)\big)$$

are continuous. Recall that $\varphi : \Omega \times Y \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is a normal integrand if it is $\Sigma \times B(Y)$-measurable and, for all $\omega \in \Omega$, $y \to \varphi(\omega, y)$ is lower semicontinuous. A normal integrand can be approximated pointwise from below by Carathéodory integrands. So we can also define the weak topology on $R(\Omega, Y)$ as the initial topology that makes all functionals $\lambda \to I_\varphi(\lambda)$ lower semicontinuous as φ ranges over all non-negative normal integrands. We want to remark that

$$R(\Omega, Y) \subset L^\infty(\Omega, M(Y)_{w^*}) = L^1(\Omega, C(Y))^*$$

(see, for example, [18: p. 377] or [19: p. 99]). Then the weak topology of $R(\Omega, Y)$ coincides with the relative w^*-topology that it inherits from $L^\infty(\Omega, M(Y)_{w^*})$. Note that if the σ-field Σ is countably generated, then the space $L^1(\Omega, C(Y))$ is separable and so the weak*-topology on bounded subsets of $L^\infty(\Omega, M(Y)_{w^*})$, such as $R(\Omega, Y)$, is metrizable. This is actually the context in which we shall use the weak topology of $R(\Omega, Y)$ in this paper.

Now let X be a Banach space and let $A: X \to X^*$ be an operator. It is said to be "monotone" if $\langle Ax_1 - Ax_2, x_1 - x_2 \rangle \geq 0$ for all $x_1, x_2 \in X$, and it is said to be "maximal monotone", if the graph $\mathrm{Gr}A = \{[x, x^*] \in X \times X^* : x^* = A(x)\}$ of A is maximal with respect to inclusion among the graphs of all monotone maps.

Finally, if $C \subset X$, then the support function $\sigma(\cdot, C): X^* \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is defined by

$$\sigma(x^*, C) = \sup_{c \in C} (x^*, c).$$

It is well-known that $\sigma(\cdot, C)$ is sublinear and w^*-lower semicontinuous.

3. Existence theorem

The problem under consideration is the following:

$$\left.\begin{aligned} & J(x,u) = \int_0^b L(t, x(t), u(t))dt \to \inf = m \\ & \text{such that} \\ & -\dot{x}(t) \in A(x(t)) + f(t, x(t), u(t)) \text{ a.e. on } [0,b] \\ & x(0) = x_0, \ u \in S^1_{U(\cdot, x(\cdot))} \end{aligned}\right\}. \tag{1}$$

In problem (1),

$L: [0,b] \times \mathbb{R}^N \times \mathbb{R}^m \to \overline{\mathbb{R}} \cup \{+\infty\}$ is the cost integrand

$A: D(A) \subset \mathbb{R}^N \to 2^{\mathbb{R}^N}$ is a maximal monotone operator

$U: [0,b] \times \mathbb{R}^N \to 2^{\mathbb{R}^m}$ is the control constraint multifunction where it is not assumed that $D(A) = \mathbb{R}^N$. We emphasize that U is x-dependent (closed loop system). Also, $f : [0,b] \times \mathbb{R}^N \times \mathbb{R}^m \to \mathbb{R}^N$ is the so called "control vector field" and $x_0 \in \overline{D(A)}$ is the initial state. In problem (1) the function $x \in W^{1,1}([0,b],\mathbb{R}^N) \subset C([0,b],\mathbb{R}^N)$ is called "state" of the system and the function $u \in L^1([0,b],\mathbb{R}^m)$ is called "control" of the system. A state-control pair (x,u) which satisfies all the constraints of problem (1) is said to be admissible (feasible). By

$$P(x_0) \subset W^{1,1}([0,b],\mathbb{R}^N) \times L^1([0,b],\mathbb{R}^m)$$

we denote the set of all admissible pairs. An admissible pair (x,u) is optimal if $J(x,u) = m$.

If $C \subset P_{fc}(\mathbb{R}^N)$,

$$i_C(x) = \begin{cases} 0 & \text{if } x \in C \\ +\infty & \text{if } x \notin C \end{cases}$$

(the indicator function of C) and $A = \partial i_C$ (the subdifferential of i_C), then the dynamics of problem (1) is a "variational inequality". Note that $A(x) = \partial i_C(x) = N_C(x)$ is the normal cone to C at $x \in C$. If $C \in P_{kc}(\mathbb{R}^N)$, Cornet [10] proved that the variational inequality is equivalent to the projected differential inclusion

$$\left.\begin{aligned} -\dot{x}(t) &\in \operatorname{proj}\big(-f(t,x(t),u(t)); T_C(x(t))\big) \quad \text{a.e. on } [0,b] \\ x(0) &= x_0 \end{aligned}\right\}$$

with $T_C(x(t))$ being the tangent cone to C at $x(t)$ and $\operatorname{proj}(\cdot; T_C(x(t)))$ being the metric projection to the closed, convex set $T_C(x(t))$. Such projected differential inclusions are appropriate in the modeling of systems with state constraints. For such systems, in describing the effect of the constraints on the dynamical equation, it can be assumed in many cases that the velocity $\dot{x}(t)$ is projected at each time instant to the set of allowed directions towards the constraint set at the point $x(t)$. This is true for electrical networks with diode non-linearities (see, for example, Krasnoselskii and Pokrovskii [20]). Also, the projected inclusions are important in mathematical economics, in the analysis of resource allocation problems (see Henry [15]).

Our hypotheses on the data of (1) are the following:

H(A) $A : D(A) \subset \mathbb{R}^N \to 2^{\mathbb{R}^N}$ is a maximal monotone operator with $D(A)$ closed and the minimal section A^o is bounded on compact subsets of $D(A)$.

We want to recall that $A^o(x) = \operatorname{proj}(0; A(x))$ and, because A is maximal monotone, $A(x) \in P_{fc}(\mathbb{R}^N)$ for all $x \in D(A)$ (see [17]).

Remark 1. Note that hypothesis $H(A)$ is satisfied if $A = \partial i_C$ with $C \in P_{fc}(\mathbb{R}^N)$ (case of variational inequalities).

H(f) $f : [0,b] \times \mathbb{R}^N \times \mathbb{R}^m \to \mathbb{R}^N$ is a function such that:

(i) For all $(x,u) \in \mathbb{R}^N \times \mathbb{R}^m$, $t \to f(t,x,u)$ is measurable.

(ii) For a.a. $t \in [0,b]$, $(x,u) \to f(t,x,u)$ is continuous.

(iii) There exist $a, c \in L^1_+([0,b])$ such that $\|f(t,x,u)\| \leq a(t) + c(t)\|x\|$ for a.a. $t \in [0,b]$, all $x \in \mathbb{R}^N$ and all $u \in U(t,x)$.

H(U) $U : [0,b] \times \mathbb{R}^N \to P_k(\mathbb{R}^m)$ is a multifunction such that:

(i) For all $x \in \mathbb{R}^N$, $t \to U(t,x)$ is measurable.

(ii) For a.a. $t \in [0,b]$, $x \to U(t,x)$ is continuous.

(iii) There exists $c_1 > 0$ such that $\|u\| \leq c_1(1 + \|x\|)$ for a.a. $t \in [0,b]$, all $x \in \mathbb{R}^N$ and all $u \in U(t,x)$.

H(L) $L : [0,b] \times \mathbb{R}^N \times \mathbb{R}^m \to \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\}$ is an integrand such that:

(i) $(t,x,u) \to L(t,x,u)$ is measurable.

(ii) For almost all $t \in [0,b]$, $(x,u) \to L(t,x,u)$ is lower semicontinuous.

(iii) There exist $\phi \in L^1([0,b])$ and $c_2 > 0$ such that $\phi(t) - c_2\|x\| \leq L(t,x,u)$ for a.a. $t \in [0,b]$, all $x \in \mathbb{R}^N$ and all $u \in U(t,x)$.

Under these hypotheses, by considering the differential inclusion which results from the deparametrization of the problem (i.e. defining the multifunction $F(t,x) = f(t,x,U(t,x))$) and using the results of Papageorgiou [22] and Hu and Papageorgiou [16], we can check that if $x_0 \in \overline{D(A)} = D(A)$, then $P(x_0) \neq \emptyset$ and $P_1(x_0) = \text{proj}_{C([0,T],\mathbb{R}^N)} P(x_0) \subset C([0,b],\mathbb{R}^N)$ (the set of admissible states) is compact.

As we already mentioned in the introduction, in order to have an existence result for problem (1), we need a convexity hypothesis:

$\mathbf{H}_C$ For a.a. $t \in [0,b]$ and all $x \in \mathbb{R}^N$, the set

$$Q(t,x) = \left\{ (h,\eta) \in \mathbb{R}^N \times \mathbb{R} \;\middle|\; \begin{array}{l} h = f(t,x,u) \\ u \in U(t,x) \\ L(t,x,u) \leq \eta \end{array} \right\}$$

is convex.

Remark 2. If, for example, $f(t,x,u) = f_1(t,x) + f_2(t,x)u$ and $U(t,x) \in P_{kc}(\mathbb{R}^N)$ for all $(t,x) \in [0,b] \times \mathbb{R}^N$ and $L(t,x,\cdot)$ is convex for a.a. $t \in [0,b]$ and all $x \in \mathbb{R}^N$, then it is easy to see that hypothesis H_C is satisfied.

We employ the "reduction of variables method" due to Berkovitz [2, 3, 5] and Cesari [8, 9]. The idea of this method is simple and elegant. We use hypothesis H_C to pass to a control-free (deparametrized) variational problem (i.e. a calculus of variations problem) with convex structure, which we can solve using the "Direct Method".

Theorem 1. *If hypotheses* H(A), H(f), H(U), H(L), H_C *hold,* $x_0 \in \overline{D(A)} = D(A)$ *and* $m < +\infty$*, then problem* (1) *admits an optimal state-control pair.*

Proof. Let $\Gamma : [0,b] \times \mathbb{R}^N \times \mathbb{R}^N \to 2^{\mathbb{R}^m}$ be the multifunction defined by

$$\Gamma(t,x,v) = \Big\{ u \in U(t,x) : \, -v \in A(x) + f(t,x,u) \Big\}$$

which is the multifunction that gives all admissible controls which, at time $t \in [0,b]$ and when the state is $x \in \mathbb{R}^N$, generate the velocity $v \in \mathbb{R}^N$. Note that

$$\mathrm{Gr}\,\Gamma = \left\{ (t,x,v,u) \in [0,b] \times \mathbb{R}^N \times \mathbb{R}^N \times \mathbb{R}^m : \begin{array}{r} (t,x,u) \in \mathrm{Gr}U \\ (x, -v - f(t,x,u)) \in \mathrm{Gr}A \end{array} \right\}.$$

By modifying f, U and L on a Lebesgue-null subset of $[0,b]$, we may assume without any loss of generality that

- for all $(x,u) \in \mathbb{R}^N \times \mathbb{R}^m$, $t \to f(t,x,u)$ is measurable
- $(t,x,u) \to L(t,x,u)$ is Borel measurable
- $\mathrm{Gr}U \in B([0,b]) \times B(\mathbb{R}^N) \times B(\mathbb{R}^m)$.

Then because $\mathrm{Gr}A \subset \mathbb{R}^N \times \mathbb{R}^N$ is closed (because A is maximal monotone),

$$\mathrm{Gr}\,\Gamma \in B([0,b]) \times B(\mathbb{R}^N) \times B(\mathbb{R}^N) \times B(\mathbb{R}^m).$$

We introduce the function $p : [0,b] \times \mathbb{R}^N \times \mathbb{R}^N \to \overline{\mathbb{R}}$ defined by

$$p(t,x,v) = \inf_{u \in \Gamma(t,x,v)} L(t,x,u).$$

Note that $p(t,x,v)$ represents the minimum istantaneous cost needed to generate velocity $v \in \mathbb{R}^N$ when the state is $x \in \mathbb{R}^N$. As usual we use the convention $\inf \emptyset = +\infty$, and for this reason p is $\overline{\mathbb{R}}$-valued. In a series of Claims I - III below we establish the properties of p.

`Claim I`: $(t,x,v) \to p(t,x,v)$ *is Borel measurable.* Indeed, for every $\lambda \in \mathbb{R}$ we have

$$\begin{aligned} &\Big\{ (t,x,v) \in [0,b] \times \mathbb{R}^N \times \mathbb{R}^N : \, p(t,x,v) < \lambda \Big\} \\ &\quad = \mathrm{proj}_{T \times \mathbb{R}^N \times \mathbb{R}^N} \Big\{ (t,x,v,u) \in \mathrm{Gr}\,\Gamma : \, L(t,x,u) < \lambda \Big\} \\ &\quad \in B([0,b]) \times B(\mathbb{R}^N) \times B(\mathbb{R}^N) \end{aligned}$$

(see [16: p. 146]). So p is Borel measurable.

`Claim II`: *For all* $t \in [0,b], (x,v) \to p(t,x,v)$ *is lower semicontinuous.* Indeed, we need to show that for every $\lambda \in \mathbb{R}$ the sublevel set

$$K(\lambda) = \Big\{ (x,v) \in \mathbb{R}^N \times \mathbb{R}^N : \, p(t,x,v) \le \lambda \Big\}$$

is closed. To this end let $\{(x_n, v_n)\}_{n\geq 1} \subset K(\lambda)$ and assume that $x_n \to x$ and $v_n \to v$ in $\mathbb{R}^N$. We can find $u_n \in \Gamma(t, x_n, v_n)$ such that $L(t, x_n, u_n) \leq p(t, x_n, v_n) + \frac{1}{n}$. So $u_n \in U(t, x_n)$ and $-v_n \in A(x_n) + f(t, x_n, u_n)$ $(n \geq 1)$. By passing to a subsequence if necessary, we may assume that $u_n \to u$ in $\mathbb{R}^m$ and $u \in U(t, x)$ (see hypothesis H(U)/(ii)). Then $f(t, x_n, u_n) \to f(t, x, u)$ in $\mathbb{R}^N$ and, because $\mathrm{Gr}A \subset \mathbb{R}^N \times \mathbb{R}^N$ is closed, $-v \in A(x) + f(t, x, u)$ with $u \in U(t, x)$ and so $u \in \Gamma(t, x, v)$. Also, from hypothesis H(L)/(ii) we have $L(t, x, u) \leq \liminf_{n\to\infty} L(t, x_n, u_n) \leq \lambda$, hence $p(t, x, v) \leq \lambda$, i.e. $(x, v) \in K(\lambda)$ which proves the claim.

`Claim III`: *For all* $(t, x) \in [0, b] \times \mathbb{R}^N$, $v \to p(t, x, v)$ *is convex.* Indeed, note that

$$\begin{aligned} \operatorname{epi} p(t, x, \cdot) &= \Big\{ (v, \eta) \in \mathbb{R}^N \times \mathbb{R} : \, p(t, x, v) \leq \eta \Big\} \\ &= \bigcap_{\varepsilon > 0} \left\{ (v, \eta) \in \mathbb{R}^N \times \mathbb{R} \;\middle|\; \begin{array}{l} -v \in A(x) + f(t, x, u) \\ L(t, x, u) \leq \eta + \varepsilon \\ u \in U(t, x) \end{array} \right\}. \end{aligned}$$

By hypothesis H_C, each set in the intersection is convex and so $\operatorname{epi} p(t, x, \cdot)$ is convex, too. This proves the claim.

Now let

$$\{(x_n, u_n)\}_{n\geq 1} \subset P(x_0) \subset C([0, b], \mathbb{R}^N) \times L^1([0, b], \mathbb{R}^m)$$

be a minimizing sequence for problem (1), i.e. $J(x_n, u_n) \downarrow m$. We already said that $P_1(x_0) \subset C([0, b], \mathbb{R}^N)$ is compact. So, from hypothesis H(A),

$$\sup_{t\in T} \sup_{n\geq 1} \|A^0(x_n(t))\| \leq c_3$$

for some $c_3 > 0$. But from [6: p. 69] we know that, for all $n \geq 1$ and a.a. $t \in [0, b]$,

$$\begin{aligned} \|\dot{x}_n(t)\| &= \|(-f(t, x_n(t), u_n(t)) - A(x_n(t)))^0\| \\ &\leq \| - f(t, x_n(t), u_n(t))\| + \|A^0(x_n(t))\| \\ &\leq \hat{a}(t) \end{aligned}$$

for some $\hat{a} \in L^1_+([0, b])$. So by the Dunford-Pettis theorem, $\{\dot{x}_n\}_{n\geq 1} \subset L^1([0, b], \mathbb{R}^N)$ is relatively weakly compact. Hence we may assume that $x_n \to x$ in $C([0, b], \mathbb{R}^N)$ (recall that $\{x_n\}_{n\geq 1} \subset P_1(x_0)$) and $\dot{x}_n \to \dot{x}$ weakly in $L^1([0, b], \mathbb{R}^N)$. Because of Claims I - III we can use the theorem of Olech [21] (see also Berkovitz [3]) and obtain

$$\begin{aligned} -\infty &< \int_0^b p(t, x(t), \dot{x}(t))\, dt \\ &\le \liminf_{n\to\infty} \int_0^b p(t, x_n(t), \dot{x}_n(t))\, dt \\ &\le \lim_{n\to\infty} J(x_n, u_n) \\ &= m \\ &< +\infty. \end{aligned}$$

By redefining, if necessary, $p(\cdot, x(\cdot), \dot{x}(\cdot))$ on a Lebesgue-null set in $[0, b]$, we may assume that, for all $t \in [0, b]$, $p(t, x(t), \dot{x}(t))$ is finite. Then via a straightforward application of the Yankov-von Neumann-Aumann selection theorem (see [17: p. 158]) we can find Borel measurable functions $u_k : [0, b] \to \mathbb{R}^m$ $(k \ge 1)$ such that

$$\left.\begin{aligned} u_k(t) &\in \Gamma(t, x(t), \dot{x}(t)) \\ L(t, x(t), u_k(t)) &\le p(t, x(t), \dot{x}(t)) + \tfrac{1}{k} \end{aligned}\right\} \qquad \text{(a.e. on } [0, b]).$$

Let

$$\hat{L}_k(t) = L(t, x(t), u_k(t)).$$

We have

$$\|\phi(t) - c_2\|x(t)\| \le \hat{L}_k(t) \le p(t, x(t), \dot{x}(t)) + \frac{1}{k} \qquad \text{(a.e. on } [0, b])$$

and so $\{\hat{L}_k\}_{k\ge1} \subset L^1([0, b])$ is uniformly integrable. Therefore we may assume that $\hat{L}_k \to \hat{L}$ weakly in $L^1([0, b])$ as $k \to \infty$, with $\hat{L} \in L^1([0, b])$. Also, if we set

$$h_k(t) = f(t, x(t), u_k(t)),$$

then $h_k \in L^1([0, b], \mathbb{R}^N)$ and from hypotheses H(f)/(iii) and H(U)/(iii) we have that $\{h_k\}_{k\ge1} \subset L^1([0, b], \mathbb{R}^N)$ is uniformly integrable. Thus we may assume that $h_k \to h$ weakly in $L^1([0, b], \mathbb{R}^N)$ as $k \to \infty$, with $h \in L^1([0, b], \mathbb{R}^N)$. Note that, for some $z_k \in L^1([0, b], \mathbb{R}^N)$ with $z_k(t) \in A(x(t))$ a.e. on $[0, b]$ $(k \ge 1)$,

$$(-\dot{x}(t) - z_k(t), \hat{L}_k(t)) \in Q(t, x(t)) \qquad \text{(a.e. on } [0, b]).$$

Therefore $-z_k(t) = \dot{x}(t) + h_k(t)$ a.e. on $[0, b]$ and so $z_k \to -\dot{x} - h$ weakly in $L^1([0, b], \mathbb{R}^N)$ as $k \to \infty$. From [17: p. 694] and using the closedness of $\mathrm{Gr}A \subset \mathbb{R}^N \times \mathbb{R}^N$ we obtain

$$-\dot{x}(t) - h(t) \in \limsup_{n\to\infty} A(x_n(t)) \subset A(x(t)) \qquad \text{(a.e. on } [0, b]),$$

i.e. $-\dot{x}(t) \in A(x(t)) + h(t)$ a.e. on $[0, b]$. Also, via Mazur's lemma, since $Q(t, x(t)) \in P_{fc}(\mathbb{R}^N \times \mathbb{R})$ (see hypothesis H_C), we have $(h(t), \hat{L}(t)) \in Q(t, x(t))$ a.e. on T. Another application of the Yankov-von Neumann-Aumann selection theorem gives a Borel measurable function $u : [0, b] \to \mathbb{R}^m$ such that

$$\left.\begin{aligned} u(t) &\in U(t, x(t)) \\ -\dot{x}(t) &\in A(x(t)) + f(t, x(t), u(t)) \\ L(t, x(t), u(t)) &\leq \hat{L}(t) \end{aligned}\right\} \qquad \text{(a.e. on } [0, b]).$$

So $(x, u) \in P(x_0)$. Moreover, since

$$\int_0^b \hat{L}_k(t)\, dt \leq \int_0^b p(t, x(t), \dot{x}(t))\, dt + \frac{b}{k}$$

we get

$$\int_0^b \hat{L}(t)\, dt \leq \int_0^b p(t, x(t), \dot{x}(t))\, dt \leq m$$

and so $J(x, u) \leq m$. But because $(x, u) \in P(x_0)$, we must have $J(x, u) = m$, i.e. (x, u) is an optimal state-control pair ∎

Reviewing the above proof, we realize that hypothesis H_C is crucial in proving the existence of an optimal state-control pair $(x, u) \in P(x_0)$. If hypothesis H_C fails, then we no longer can guarantee that problem (1) has a solution, because the limit of a minimizing sequence need not be admissible. To capture the asymptotic behaviour of the minimizing sequences, we need to augment the system in such a way so as to introduce the missing "convex structure". This is the object of study of relaxation theory, which we investigate in the next two sections.

4. Three relaxation methods

The first method of relaxation increases the set of admissible controls by considering Young measures. The idea behind this method is better understood in the case of $\mathbb{R}$-valued functions. It is well known that a sequence $\{u_n\}_{n \geq 1} \subset L^1([0, b])$ of controls, which converges weakly but not strongly to u, oscillates wildly around u. But in the limit function u all these fast oscillations are forgotten and only an average value is recorded. Certainly, this is not satisfactory if the control function enters in a nonlinear fashion in the dynamics of the system. We can not say that

$$f(\cdot, x_n(\cdot), u_n(\cdot)) \to f(\cdot, x(\cdot), u(\cdot)) \quad \text{weakly in } L^1([0, b], \mathbb{R}^N)$$

even if $x_n \to x$ in $C([0,b],\mathbb{R}^N)$. Then the idea is to assign as a limit of $\{u_n\}_{n\geq 1}$ not a usual $\mathbb{R}$-valued function, but a probability-valued function (a transition probability) $\lambda : [0,b] \to M_+^1(S)$ where $S \subset \mathbb{R}$ is the set where the control functions take their values. These considerations lead to the first relaxation method based on Young measures.

We will need to the following stronger hypotheses on the control constraint multifunction $U(t,x)$.

H(U)$_1$ $U : [0,b] \times \mathbb{R}^N \to P_k(\mathbb{R}^m)$ is a multifunction such that:

(i) For all $x \in \mathbb{R}^N$, $t \to U(t,x)$ is measurable.

(ii) For a.a. $t \in [0,b]$, $x \to U(t,x)$ is continuous.

(iii) There exists $r > 0$ such that $\|u\| \leq r$ for a.a. $t \in [0,b]$, all $x \in \mathbb{R}^N$ and all $u \in U(t,x)$.

In what follows,

$$\overline{B}_r = \{u \in \mathbb{R}^m : \|u\| \leq r\}.$$

We introduce the constraint set $\Sigma(t,x)$ for the controls by setting

$$\Sigma(t,x) = \Big\{\mu \in M_+^1(\overline{B}_r) : \mu(U(t,x)) = 1\Big\}.$$

Given a state function $x \in C([0,b],\mathbb{R}^N)$, the set of admissible relaxed controls is given by

$$S_{\Sigma(\cdot,x(\cdot))} = \Big\{\lambda \in R([0,b],\overline{B}_r) : \lambda(t) \in \Sigma(t,x(t)) \text{ a.e. on } [0,b]\Big\}.$$

Then the first relaxation of problem (1) based on Young measures is the optimal control problem

$$\left.\begin{aligned} & J_r^1(x,\lambda) = \int_0^b\int_{\overline{B}_r} L(t,x(t),u)\lambda(t)(du)dt \to \inf = m_r^1 \\ & \text{such that} \\ & -\dot{x}(t) \in A(x(t)) + \int_{\overline{B}_r} f(t,x(t),u)\lambda(t)(du) \quad \text{a.e. on } [0,b] \\ & x(0) = x_0,\ \lambda \in S_{\Sigma(\cdot,x(\cdot))} \end{aligned}\right\}. \tag{2}$$

Note that in (2) the control function enters linearly in the dynamics and in the cost functional. This gives problem (2) the desired "convex structure". Also, note that every original control $u \in S^1_{U(\cdot,x(\cdot))}$ can be viewed as a relaxed control by considering the corresponding Dirac transition probability $\delta_{u(\cdot)}$.

For problem (2), with no extra hypotheses, we can show that it has a solution.

Proposition 2. *If hypotheses* H(A), H(f), $\mathrm{H(U)}_1$, H(L) (*in which hypotheses* H(f)/(iii) *and* H(L)/(iii) *are true for all* $u \in \overline{B}_r$) *hold,* $x_0 \in \overline{D(A)} = D(A)$ *and* $m_r^1 < +\infty$, *then problem* (2) *admits an optimal relaxed pair* $(x, \lambda) \in C([0,b], \mathbb{R}^N) \times S_{\Sigma(\cdot, x(\cdot))}$.

Proof. Let $\{x_n, \lambda_n\}_{n\geq 1} \subset C([0,b], \mathbb{R}^N) \times R([0,b], \overline{B}_r)$ be a minimizing sequence for problem (2). We know that the sequence $\{x_n\}_{n\geq 1} \subset C([0,b], \mathbb{R}^N)$ is relatively compact. Also, by Alaoglu's theorem, the sequence $\{\lambda_n\}_{n\geq 1} \subset L^\infty([0,b], M(\overline{B}_r)_{w^*})$ is relatively w^*-compact. Because the predual $L^1([0,b], C(\overline{B}_r))$ is separable, $\{\lambda_n\}_{n\geq 1}$ is relatively w^*-sequentially compact. So we may assume that $x_n \to x$ in $C([0,b], \mathbb{R}^N)$ and $\lambda_n \to \lambda$ weakly* in $L^\infty([0,b], M(\overline{B}_r)_{w^*})$.

First we show that (x, λ) is admissible for problem (2). Set $\hat{f}_n(t)(u) = f(t, x_n(t), u)$ and $\hat{f}(t)(u) = f(t, x(t), u)$. We have

$$\begin{aligned} \big\|\hat{f}_n(t) - \hat{f}(t)\big\|_{C(\overline{B}_r)} &= \sup_{u\in\overline{B}_r} \big\|\hat{f}_n(t)(u) - \hat{f}(t)(u)\big\| \\ &= \big\|\hat{f}_n(t)(u_n) - \hat{f}(t)(u_n)\big\| \end{aligned} \qquad (n \geq 1)$$

for some $u_n \in \overline{B}_r$. We may assume that $u_n \to u$ in $\overline{B}_r$. Hence $\hat{f}_n(t) \to \hat{f}(t)$ in $C(\overline{B}_r)$, and by the dominated convergence theorem, $\hat{f}_n \to \hat{f}$ in $L^1([0,b], C(\overline{B}_r))$. If by $((\cdot,\cdot))$ we denote the duality brackets for the pair

$$\Big(L^1([0,b], C(\overline{B}_r)), L^\infty([0,b], M(\overline{B}_r)_{w^*})\Big),$$

for every $C \in B([0,b])$ we have $((\hat{f}_n, \chi_C \lambda_n)) \to ((\hat{f}, \chi_C \lambda))$ and subsequently

$$\int_C \int_{\overline{B}_r} f(t, x_n(t), u)\lambda_n(t)\,(du)dt \to \int_C \int_{\overline{B}_r} f(t, x(t), u)\lambda(t)\,(du)dt.$$

Also, as before we have $\dot{x}_n \to \dot{x}$ weakly in $L^1([0,b], \mathbb{R}^N)$ (see the proof of Theorem 1). Hence

$$\int_C \dot{x}_n(t)\,dt \to \int_C \dot{x}(t)\,dt.$$

Moreover, if $z_n \in S^1_{A(x_n(\cdot))}$ is such that

$$-\dot{x}_n(t) = z_n(t) + \int_{\overline{B}_r} f(t, x_n(t), u)\lambda_n(t)\,(du) \qquad \text{(a.e. on } [0,b]),$$

the sequence $\{z_n\}_{n\geq 1} \subset L^1([0,b], \mathbb{R}^N)$ is uniformly integrable and so we may assume that $z_n \to z$ weakly in $L^1([0,b], \mathbb{R}^N)$. Exploiting the closedness of GrA, $z \in S^1_{A(x(\cdot))}$. So in the limit, as $n \to \infty$, we obtain

$$-\int_C \dot{x}(t)\,dt \in \int_C A(x(t))\,dt + \int_C \int_{\overline{B}_r} f(t, x(t), u)\lambda(t)\,(du)dt$$

for all $C \in B([0,b])$, hence

$$\left.\begin{aligned} -\dot{x}(t) \in A(x(t)) + \int_{\overline{B}_r} f(t,x(t),u)\lambda(t)(du) \ \text{ a.e. on } [0,b] \\ x(0) = x_0 \end{aligned}\right\}.$$

We need to show that $\lambda \in S_{\Sigma(\cdot,x(\cdot))} = S_{\Sigma(x)}$. To this end, let $h \in L^1([0,b], C(\overline{B}_r))$. We have

$$((h,\lambda_n)) \le \sigma(h, S_{\Sigma(x_n)}) = \int_0^b \sigma\big(h(t), \Sigma(t,x_n(t))\big)dt$$

(see [17: p. 183]), and so

$$\begin{aligned} ((h,\lambda)) &\le \limsup_{n\to\infty} \int_0^b \sigma\big(h(t), \Sigma(t,x_n(t))\big)dt \\ &\le \int_0^b \limsup_{n\to\infty} \sigma\big(h(t), \Sigma(t,x_n(t))\big)dt. \end{aligned} \tag{3}$$

We claim that $\Sigma(t,\cdot)$, as a multifunction into $M_+^1(\overline{B}_r)$ with the weak topology, has a closed graph. Recall that $M_+^1(\overline{B}_r)$, topologized this way, is compact and metrizable (see [23: p. 45]). Moreover, the weak topology coincides with the relative w^*-topology, it inherits from the Banach space $M(\overline{B}_r)$. To prove the claim, we consider a sequence $\{(v_n,\mu_n)\}_{n\ge1} \subset Gr\Sigma(t,\cdot)$ and assume that $v_n \to v$ in $\mathbb{R}^N$ and $\mu_n \to \mu$ weakly in $M_+^1(\overline{B}_r)$. By virtue of hypothesis H(U)$_1$/(ii), given $0 < \varepsilon < r$ we can find $n_0 = n_0(\varepsilon) \ge 1$ such that $U(t,v_n) \subset U(t,v) + \varepsilon\overline{B}_1$ for all $n \ge n_0$, hence $\mu_n(U(t,v_n)) \le \mu_n(U(t,v)) + \varepsilon$ for all $n \ge n_0$ and so

$$1 \le \limsup_{n\to\infty} \mu_n(U(t,v_n)) + \varepsilon \le \mu(U(t,v)) + \varepsilon$$

by the Portmanteau theorem (see [23: p. 40]). Let $\varepsilon \downarrow 0$ to conclude that $1 \le \mu(U(t,v))$, hence $\mu(U(t,v)) = 1$ and so $(v,\mu) \in \mathrm{Gr}\Sigma(t,\cdot)$ which proves the claim. Using this fact we have

$$\limsup_{n\to\infty} \sigma\big(h(t), \Sigma(t,x_n(t))\big) \le \sigma\big(h(t), \Sigma(t,x(t))\big)$$

a.e. on $[0,b]$ and so from (3) we obtain

$$((h,\lambda)) \le \int_0^b \sigma\big(h(t), \Sigma(t,x(t))\big)dt = \sigma(h, S_{\Sigma(x)})$$

for all $h \in L^1\big([0,b], C(\overline{B}_r)\big)$, hence $\lambda \in S_{\Sigma(x)}$, i.e. (x,λ) is admissible for problem (2).

Next, let $L_k : [0,b] \times \mathbb{R}^N \times \overline{B}_r \to \mathbb{R}$ be Carathéodory integrands such that $\psi(t) - c_2\|x\| \le L_k(t,x,u) \le k$ for a.a. $t \in [0,b]$, all $x \in \mathbb{R}^N$ and all $u \in \overline{B}_r$ and $L_k \uparrow L$ as $k \to \infty$ (see, for example, [18: p. 279]). So let $\hat{L}_{k,n}(t,u) = L_k(t,x_n(t),u)$ and $\hat{L}_k(t,u) = L_k(t,x(t),u)$. Evidently, for all $k \ge 1$, $\{\hat{L}_{k,n}, \hat{L}_k\}_{n\ge 1} \subset L^1([0,T],C(\overline{B}_r))$, and as we did for the $\hat{f}_n$'s, we can show that $\hat{L}_{k,n} \to \hat{L}_k$ in $L^1([0,T],C(\overline{B}_r))$ as $n \to \infty$. So $((\hat{L}_{k,n},\lambda_n)) \to ((\hat{L}_k,\lambda))$ as $n \to \infty$, while from the monotone convergence theorem

$$((\hat{L}_k,\lambda)) \uparrow \int_0^b\int_{\overline{B}_r} L(t,x(t),u)\lambda(t)\,(du)dt$$

as $k \to \infty$. Then from [1: p. 32] we can find a sequence $n \to k(n)$ increasing (not necessarily strictly) to $+\infty$ such that

$$((\hat{L}_{k(n),n},\lambda_n)) \to \int_0^b\int_{\overline{B}_r} L(t,x(t),u)\lambda(t)\,(du)dt.$$

But note that

$$((\hat{L}_{k(n),n},\lambda_n)) \le \int_0^b\int_{\overline{B}_r} L(t,x_n(t),u)\lambda_n(t)\,(du)dt = J_r^1(x_n,\lambda_n).$$

Hence

$$\int_0^b\int_{\overline{B}_r} L(t,x(t),u)\lambda(t)\,(du)dt \le m_r^1.$$

Because (x,λ) is admissible for problem (2), we conclude that $J_r^1(x,\lambda) = m_r^1$ ∎

Since there are more relaxed controls than originals ones, we have $m_r^1 \le m$. In principle strict inequality is possible, including the extreme case in which $m_r^1 < +\infty$ and $m = +\infty$. This can happen if, for example, there is a target set which can be reached by relaxed trajectories but not by original ones. We want to have that $m_r^1 = m$ or otherwise it can be said that the relaxed problem generalizes the original one too much and we can not find an ε-optimal control among the original (physically realizable) ones. Such a relaxation method is for all practical purposes "non-admissible". To avoid having a "relaxation gap" we need to strengthen our hypotheses on the data.

$\mathbf{H(U)}_2$ $U : [0,b] \to P_k(\mathbb{R}^m)$ is a measurable multifunction such that there exists $r > 0$ with the property that $\|u\| \le r$ for all $u \in U(t)$.

$\mathbf{H(f)}_1$ $f : [0,b] \times \mathbb{R}^N \times \mathbb{R}^m \to \mathbb{R}^N$ is a function such that:
(i) For all $(x,u) \in \mathbb{R}^N \times \mathbb{R}^m$, $t \to f(t,x,u)$ is measurable.
(ii) There exists $k \in L_+^1([0,T])$ such that $\|f(t,x,u) - f(t,x',u)\| \le k(t)\|x - x'\|$ for a.a. $t \in [0,b]$, all $x,x' \in \mathbb{R}^N$ and all $u \in U(t)$.

(iii) For a.a. $t \in [0,b]$ and all $x \in \mathbb{R}^N$, $u \to f(t,x,u)$ is continuous.
(iv) There exist $a, c \in L^1_+([0,b])$ such that $\|f(t,x,u)\| \le a(t) + c(t)\|x\|$ for a.a. $t \in [0,b]$, all $x \in \mathbb{R}^N$ and all $u \in U(t)$.

H(L)$_1$ $L : [0,b] \times \mathbb{R}^N \times \overline{B}_r \to \mathbb{R}$ is an integrand such that:
(i) For all $(x,u) \in \mathbb{R}^N \times \overline{B}_r$, $t \to L(t,x,u)$ is measurable.
(ii) For a.a. $t \in [0,b]$, $(x,u) \to L(t,x,u)$ is continuous.
(iii) For every $n \ge 1$, there exists $\psi_n \in L^1_+([0,b])$ such that $|L(t,x,u)| \le \psi_n(t)$ for a.a. $t \in [0,b]$, all $x \in \mathbb{R}^N$ with $\|x\| \le n$ and all $u \in \overline{B}_r$.

Remark 3. Hypothesis H(U)$_2$ implies that there is no feedback in the system (open loop system).

With these stronger hypotheses we shall show the admissibility of the first relaxation method. We start with an auxiliary result which follows from a powerful result about the extremal structure of a measurable multifunction (see [17: pp. 191 - 192]).

Lemma 3. *If hypothesis* H(U)$_2$ *holds, then* $\overline{S}^w_U = S_\Sigma$ *where* $\overline{S}^w_U$ *denotes the closure of* S_U *in* $R([0,b], \overline{B}_r)$ *with the weak topology.*

Proof. By definition,

$$\Sigma(t) = \Big\{ \mu \in M^1_+(\overline{B}_r) : \mu(U(t)) = 1 \Big\}.$$

We claim that Σ is graph measurable. For $G \in B([0,b]) \times B(\overline{B}_r)$, let $\eta_G : [0,b] \times M^1_+(\overline{B}_r) \to [0,1]$ be defined by $\eta_G(t,\mu) = (\delta_t \otimes \mu)(G)$, where δ_t is the Dirac probability measure concentrated on t. Let

$$\mathcal{I} = \Big\{ G \in B([0,b]) \times B(\overline{B}_r) : \eta_G \text{ is measurable} \Big\}$$

(recall that on $M^1_+(\overline{B}_r)$ we consider the weak topology). For any $D \in B([0,b])$ and $C \in B(\overline{B}_r)$ we have

$$\eta_{D \times C}(t,\mu) = \chi_D(t)\mu(C) = \chi_D(t)\varphi_C(\mu)$$

where $\varphi_C : M^1_+(\overline{B}_r) \to [0,1]$ is defined by $\varphi_C(\mu) = \mu(C)$. From the Portmanteau theorem, φ_C is measurable. Hence $\eta_{D\times C}$ is measurable and so $D \times C \in \mathcal{I}$. It is easy to see that $\mathcal{I}$ is actually a field and a monotone class, so it is a σ-field. Therefore $\mathcal{I} = B([0,b]) \times B(\overline{B}_r)$. So Gr$U \in \mathcal{I}$ and let

$$\begin{aligned} G_1 &= \Big\{ (t,\mu) \in [0,b] \times M^1_+(\overline{B}_r) : (\delta_t \otimes \mu)(GrU) = 1 \Big\} \\ &= \eta^{-1}_{GrU}(1) \\ &\in B([0,b]) \times B(M^1_+(\overline{B}_r)). \end{aligned}$$

By Fubini's theorem,

$$(\delta_t \otimes \mu)(\mathrm{Gr}U) = \int_0^b \int_{\overline{B}_r} \chi_{U(s)}(u)\,\mu(du)\delta_t(ds) = \mu(U(t)).$$

So

$$G_1 = \mathrm{Gr}\Sigma \in B([0,b]) \times B(M_+^1(\overline{B}_r)),$$

i.e. Σ is graph measurable. Then, from [17: p. 191], $\operatorname{ext} S_\Sigma = S_{\operatorname{ext}\Sigma}$ and $\operatorname{ext}\Sigma(t) = \{\delta_u : u \in U(t)\}$. So $\operatorname{ext} S_\Sigma = S_U$. Finally, use [17: p. 192/Proposition II.4.7] to obtain the desired density result ∎

Using this lemma we can prove the admissibility of the first relaxation method. In what follows, by $P_r^1(x_0)$ we denote the admissible relaxed pairs of problem (2) and $P_{r_1}^1(x_0) = \mathrm{proj}_{C([0,T],\mathbb{R}^N)} P_r^1(x_0)$.

Theorem 4. *If hypotheses* H(A), H(f)$_1$, H(U)$_2$, H(L)$_1$ *hold,* $x_0 \in \overline{D(A)} = D(A)$ *and* $m < +\infty$*, then* $P_{r_1}^1(x_0) = \overline{P_1(x_0)}^{C([0,T],\mathbb{R}^N)}$ *and* $m_r^1 = m$*.*

Proof. From Proposition 2 we know that problem (2) has a solution $(x,\lambda) \in C([0,b],\mathbb{R}^N) \times R([0,b],\overline{B}_r)$. Using Lemma 3 we can find a sequence $\{u_n\}_{n\geq 1} \subset S_U^1$ such that $\delta_{u_n} \to \lambda$ weakly* in $L^\infty([0,b], M(\overline{B}_r)_{w^*})$ (recall that the weak topology of $R([0,b],\overline{B}_r)$ coincides with the relative w^*-topology inherited from $L^\infty([0,b], M(\overline{B}_r)_{w^*})$, see Section 2). Because of hypothesis H(f)$_1$/(ii), every original control $u_n \in S_U^1$ $(n \geq 1)$ generates a unique state $x_n \in C([0,b],\mathbb{R}^N)$. We may assume that $x_n \to \hat{x}$ in $C([0,b],\mathbb{R}^N)$. If

$$\hat{f}_n(t) = \int_{\overline{B}_r} f(t, x_n(t), u)\delta_{u_n(t)}(du)$$

$$\hat{f}(t) = \int_{\overline{B}_r} f(t, \hat{x}(t), u)\lambda(t)(du)$$

we have $\hat{f}_n \to \hat{f}$ weakly in $L^1([0,T],\mathbb{R}^N)$. Also, from the proof of Theorem 1 we know that the sequence $\{\dot{x}_n\}_{n\geq 1} \subset L^1([0,b],\mathbb{R}^N)$ is relatively weakly compact and so we may assume that $\dot{x}_n \to v$ weakly in $L^1([0,b],\mathbb{R}^N)$. Clearly, $v = \dot{\hat{x}}$. As in the proof of Theorem 1 we can check that in the limit, as $n \to \infty$, we get

$$\left.\begin{aligned} &-\dot{\hat{x}}(t) \in A(\hat{x}(t)) + \int_{\overline{B}_r} f(t,\hat{x}(t),u)\lambda(t)(du) \quad \text{a.e. on } [0,b] \\ &x(0) = x_0 \end{aligned}\right\}.$$

Since the relaxed control $\lambda \in S_\Sigma$ generates a unique state, we must have $\hat{x} = x$. There is

$$J(x_n,u_n) = J_r^1(x_n,\delta_{u_n}) \to J_r^1(x,\lambda) = m_r^1$$

and so $m \leq m_r^1$. Since the opposite inequality is always true, $m = m_r^1$. Moreover, from the above argument it is clear that $P_{r_1}^1(x_0) = \overline{P_1(x_0)}^{C([0,T],\mathbb{R}^N)}$ ∎

Remark 4. The relaxation method based on Young measures was initiated by Gamkrelidze [14] and Warga [25].

Next we present the second relaxation method. It is motivated by the reduction method used in the proof of Theorem 1. According to this method we deparametrize the dynamics (i.e. remove the control variable) and pass to a set-valued dynamical system. For this reason we call this method the "multi-valued relaxation method". Since $p(t,x,\cdot)$ is not convex in general, we consider its second conjugate in the sense of convex analysis. The resulting variational problem is a calculus of variations problem. More precisely, the second relaxed problem corresponding to problem (1) is

$$\left.\begin{array}{l} J_r^2(x) = \displaystyle\int_0^b p^{**}\big(t, x(t), \dot{x}(t)\big)\, dt \to \inf = m_r^2 \\ \text{such that} \\ -\dot{x}(t) \in A(x(t)) + \overline{\text{conv}} F(t, x(t)) \quad \text{a.e. on } [0,b] \\ x(0) = x_0 \end{array}\right\} \tag{4}$$

where $F(t,x) = f(t,x,U(t))$.

First we want to clarify the relation between the two relaxed problems (2) and (4). This requires a closer look on the dynamics and cost integrands of the two problems. As before, by $P_{r_1}^2(x_0) \subset C([0,b],\mathbb{R}^N)$ we denote the set of states of problem (4).

Proposition 5. *If hypotheses* H(A), H(f), H(U)$_2$ *hold and* $x_0 \in \overline{D(A)} = D(A)$*, then* $P_{r_1}^1(x_0) = P_{r_1}^2(x_0)$*, i.e. problems* (2) *and* (4) *share the same set of states.*

Proof. First we show that, for all $(t,x) \in [0,b]\times\mathbb{R}^N$,

$$\overline{\text{conv}} F(t,x) = \left\{ \int_{\overline{B}_r} f(t,x,u)\lambda(du) : \lambda \in \Sigma(t) \right\}.$$

Note that the right-hand side of the claimed equality is convex. We show that it is also closed. To this end let

$$\{y_n\}_{n\geq 1} \subset \left\{ \int_{\overline{B}_r} f(t,x,u)\lambda(du) : \lambda \in \Sigma(t) \right\}$$

and assume that $y_n \to y$ in $\mathbb{R}^N$. We have

$$y_n = \int_{\overline{B}_r} f(t,x,u)\lambda_n(du)$$

with $\lambda_n \in \Sigma(t)$. Recall that $M_+^1(\overline{B}_r)$ furnished with the weak topology (equivalently, with the relative w^*-topology of $M(\overline{B}_r)$) is compact metrizable. So we may assume that $\lambda_n \to \lambda$ weakly in $M_+^1(\overline{B}_r)$ and $\lambda \in \Sigma(t)$ (see the proof of Proposition 2). Then

$$\int_{\overline{B}_r} f(t,x,u)\lambda_n(du) \to \int_{\overline{B}_r} f(t,x,u)\lambda(du)$$

(since $f(t,x,\cdot) \in C(\overline{B}_r)$) and so $y = \int_{\overline{B}_r} f(t,x,u)\lambda(du)$ with $\lambda \in \Sigma(t)$. This proves the closedness of the set $\{\int_{\overline{B}_r} f(t,x,u)\lambda(du) : \lambda \in \Sigma(t)\}$. By taking $\lambda = \delta_u$ with $u \in U(t)$ we obtain

$$\overline{\mathrm{conv}} F(t,x) \subset \left\{ \int_{\overline{B}_r} f(t,x,u)\lambda(du) : \lambda \in \Sigma(t) \right\}.$$

Next, let $y = \int_{\overline{B}_r} f(t,x,u)\lambda(du)$ with $\lambda \in \Sigma(t)$. From [23: p. 44], we can find a sequence of discrete probabilities $\lambda_n = \sum_{k=1}^{M_n} a_k \delta_{u_k}$, with $a_k \in [0,1]$ such that $\sum_{k=1}^{M_n} a_k = 1$ and $u_k \in U(t)$ such that $\lambda_n \to \lambda$ weakly in $M_+^1(\overline{B}_r)$. So

$$y_n = \int_{\overline{B}_r} f(t,x,u)\lambda_n(du) = \sum_{k=1}^{M_n} a_k f(t,x,u_k) \in \overline{\mathrm{conv}} F(t,x)$$

and $y_n \to y$. Hence

$$\left\{ \int_{\overline{B}_r} f(t,x,u)\lambda(du) : \lambda \in \Sigma(t) \right\} \subset \overline{\mathrm{conv}} F(t,x)$$

and thus finally equality holds.

Then we claim that, for all $x \in C([0,T],\mathbb{R}^N)$,

$$S^1_{\overline{\mathrm{conv}} F(\cdot,x(\cdot))} = \left\{ y \in C([0,T],\mathbb{R}^N) : y(t) = \int_{\overline{B}_r} f(t,x(t),u)\lambda(t)(du), \lambda \in S_\Sigma \right\}. \tag{5}$$

Denote the left- and right-hand sides of (5) by $V_1(x)$ and $V_2(x)$, respectively. It is clear from the equality

$$\overline{\mathrm{conv}} F(t,x) = \left\{ \int_{\overline{B}_r} f(t,x,u)\lambda(du) : \lambda \in \Sigma(t) \right\}$$

that $V_2(x) \subset V_1(x)$. Suppose $y \in V_1(x)$ and let $K : T \to 2^{M_+^1(\overline{B}_r)} \setminus \{\emptyset\}$ be defined by

$$K(t) = \left\{ \lambda \in \Sigma(t) : y(t) = \int_{\overline{B}_r} f(t,x(t),u)\lambda(du) \right\}$$

(because of the equality $\overline{\text{conv}}F(t,x) = \{\int_{\overline{B}_r} f(t,x,u)\lambda(du) : \lambda \in \Sigma(t)\}$, K has non-empty values). Let

$$\beta(t,\lambda) = y(t) - \int_{\overline{B}_r} f(t,x(t),u)\lambda(du).$$

Evidently, β is a Carathéodory function (i.e. measurable in $t \in T$ and continuous in $\lambda \in M_+^1(\overline{B}_r)$). Therefore it is jointly measurable and so

$$\begin{aligned}\text{Gr}K &= \Big\{(t,\lambda) \in T \times M_+^1(\overline{B}_r) : \beta(t,\lambda) = 0\Big\} \cap \text{Gr}\Sigma \\ &\in B([0,b]) \times B(M_+^1(\overline{B}_r))\end{aligned}$$

(recall that Σ is graph measurable, see the proof of Lemma 3). Apply the Yankov-von Neumann-Aumann selection theorem (see [17: p. 158]) to obtain a Borel measurable function $\lambda \to M_+^1(\overline{B}_r)$ such that $\lambda(t) \in K(t)$ a.e. on $[0,b]$. So $\lambda \in S_\Sigma$ and $y(t) = \int_{\overline{B}_r} f(t,x(t),u)\lambda(t)(du)$ which proves (5). From (5) it follows at once that $P_{r_1}^1(x_0) = P_{r_1}^2(x_0)$ ∎

Proposition 6. *If hypotheses* H(A), H(f), H(U)$_2$, H(L)$_1$ *hold,* $x_0 \in \overline{D(A)} = D(A)$ *and* $(t,x,v) \in [0,b] \times \mathbb{R}^N \times \mathbb{R}^N$ *is such that*

$$-v \in A(x) + \bigg\{ \int_{\overline{B}_r} f(t,x,u)\lambda(du) : \lambda \in \Sigma(t)\bigg\},$$

then there exists $\lambda_0 \in \Sigma(t)$ *such that*

$$\begin{aligned}-v &\in A(x) + \int_{\overline{B}_r} f(t,x,u)\lambda_0(du) \\ p^{**}(t,x,v) &= \int_{\overline{B}_r} L(t,x,u)\lambda_0(du).\end{aligned}$$

Proof. By definition,

$$p^{**}(t,x,v) = \inf\Big\{\eta \in \overline{\mathbb{R}} = \mathbb{R} \cup \{+\infty\} : (v,\eta) \in \overline{\text{conv}}\,\text{epi}\,p(t,x,\cdot)\Big\}.$$

Set

$$H(t,x) = \Big\{\big(A(x) + f(t,x,u), L(t,x,u)\big) : u \in U(t)\Big\} \subset \mathbb{R}^N \times \overline{\mathbb{R}}.$$

We have

$$p^{**}(t,x,v) = \inf\Big\{\eta \in \overline{\mathbb{R}} : (-v,\eta) \in \overline{\text{conv}}H(t,x)\Big\}.$$

From the proof of Proposition 5 we know that

$$\begin{aligned} V &= \overline{\text{conv}} H(t,x) \\ &= \left\{ \left(A(x) + \int_{\overline{B}_r} f(t,x,u)\lambda(du), \int_{\overline{B}_r} L(t,x,u)\lambda(du) \right) : \lambda \in \Sigma(t) \right\}. \end{aligned}$$

Hence we obtain

$$p^{**}(t,x,v) = \inf \left\{ \int_{\overline{B}_r} L(t,x,u)\lambda(du) \;\middle|\; \begin{array}{c} \lambda \in \Sigma(t) \\ -v \in A(x) + \int_{\overline{B}_r} f(t,x,u)\lambda(du) \end{array} \right\}.$$

We introduce the set

$$\Sigma_1(t,x,v) = \left\{ \lambda \in \Sigma(t) : -v \in A(x) + \int_{\overline{B}_r} f(t,x,u)\lambda(du) \right\} \neq \emptyset.$$

This set is weakly closed in $\Sigma(t) \subset M_+^1(\overline{B}_r)$, hence it is weakly compact. Moreover, the map $\lambda \to \int_{\overline{B}_r} L(t,x,u)\lambda(du)$ is lower semicontinuous on $M_+^1(\overline{B}_r)$ (see Section 2). So by the Weierstrass theorem we can find $\lambda_0 \in \Sigma_1(t,x,v)$ such that $p^{**}(t,x,v) = \int_{\overline{B}_r} L(t,x,u)\lambda_0(du)$ ∎

Remark 5. According to the proof of Proposition 6,

$$p^{**}(t,x,v) = \inf \left\{ \int_{\overline{B}_r} L(t,x,u)\lambda(du) \;\middle|\; \begin{array}{c} \lambda \in \Sigma(t) \\ -v \in A(x) + \int_{\overline{B}_r} f(t,x,u)\lambda(du) \end{array} \right\}$$

with the usual convention $\inf \emptyset = +\infty$. Moreover, the above infimum is actually attained.

Proposition 7. *If hypotheses* H(A), H(f), H(U)_2, H(L)_1 *hold and* $x_0 \in \overline{D(A)} = D(A)$, *then* $p^{**} : [0,b] \times \mathbb{R}^N \times \mathbb{R}^N \to \overline{\mathbb{R}}$ *is superpositionally measurable (i.e. if* $x, v : [0,b] \to \mathbb{R}^N$ *are measurable, then so is* $t \to p^{**}(t,x(t),v(t)))$ *and* $(x,v) \to p^{**}(t,x,v)$ *is lower semicontinuous.*

Proof. Let $x, v : [0,b] \to \mathbb{R}^N$ be two measurable functions. By approximating from below $L(t,x(t),\cdot)$ with Carathéodory functions (which are jointly measurable), we can see that

$$(t,\lambda) \to \int_{\overline{B}_r} L(t,x(t),u)\lambda(du)$$

is jointly measurable. Also,

$$\Sigma_2(t) = \left\{ \lambda \in \Sigma(t) : -v(t) \in A(x(t)) + \int_{\overline{B}_r} f(t,x(t),u)\lambda(du) \right\}$$

is graph measurable (since Σ is). Because

$$p^{**}(t,x(t),v(t)) = \inf\left\{\int_{\overline{B}_r} L(t,x(t),u)\lambda(du) : \lambda \in \Sigma_2(t)\right\},$$

from [17: p. 161] it follows that $t \to p^{**}(t,x(t),v(t))$ is measurable, i.e. p^{**} is superpositionally measurable.

For the lower semicontinuity of $p^{**}(t,\cdot,\cdot)$, we need to show that for every $\eta \in \mathbb{R}$ the sublevel set

$$\Gamma(\eta) = \Big\{(x,v) \in \mathbb{R}^N \times \mathbb{R}^N : p^{**}(t,x,v) \le \eta\Big\}$$

is closed. So let $x_n \to x$ and $v_n \to v$ in $\mathbb{R}^N$ with $(x_n,v_n) \in \Gamma(\eta)$ and $\eta \ge 1$. We know (see Proposition 6 and Remark 5) that

$$p^{**}(t,x_n,v_n) = \int_{\overline{B}_r} L(t,x_n,u)\lambda_n(du)$$

with $\lambda_n \in \Sigma(t)$ and

$$-v_n \in A(x_n) + \int_{\overline{B}_r} f(t,x_n,u)\lambda_n(du).$$

Evidently, we may assume that $\lambda_n \to \lambda$ weakly in $M_+^1(\overline{B}_r)$. Then as in the proof of Proposition 2 we have

$$\int_{\overline{B}_r} L(t,x,u)\lambda(du) \le \liminf_{n\to\infty} \int_{\overline{B}_r} L(t,x_n,u)\lambda_n(du)$$

while

$$-v \in A(x) + \int_{\overline{B}_r} f(t,x,u)\lambda(du).$$

Therefore

$$p^{**}(t,x,v) \le \int_{\overline{B}_r} L(t,x,u)\lambda(du) \le \eta$$

which proves the lower semicontinuity of $p^{**}(t,\cdot,\cdot)$ ∎

These propositions lead at once to the following theorem which compares the two relaxed problems (2) and (4) and shows that they are equivalent.

Theorem 8. *If hypotheses* H(A), H(f), H(U)$_2$, H(L)$_1$ *hold and* $x_0 \in \overline{D(A)} = D(A)$, *then problem* (4) *has a solution and* $m_r^2 = m_r^1$.

By strengthening our hypotheses on f, we can guarantee the admissibility of this second relaxation method. More precisely, we have

Theorem 9. *If hypotheses* H(A), H(f)$_1$, H(U)$_2$, H(L)$_1$ *hold,* $x_0 \in \overline{D(A)} = D(A)$ *and* $m < +\infty$*, then* $P^2_{r_1}(x_0) = \overline{P_1(x_0)}^{C([0,T],\mathbb{R}^N)}$ *and* $m = m^1_r = m^2_r$.

Remark 6. The second relaxation method (the multi-valued relaxation method) was initiated with the work of Filippov [13].

Now we shall present the third relaxation method, which is based on Carathéodory's theorem for convex sets in $\mathbb{R}^N$. Recall that, according to this theorem, if $C \subset \mathbb{R}^N$, then every point of $\operatorname{conv} C$ is a convex combination of no more than $N+1$ distinct points of C. Motivated by this theorem, we introduce the following relaxed problem in which $\hat{u} = (u_k)_{k=1}^{N+1}$ and $\hat{\gamma} = (\gamma_k)_{k=1}^{N+1}$:

$$\left.\begin{array}{l} J^3_r(x,\hat{u},\hat{\gamma}) = \displaystyle\int_0^b \sum_{k=1}^{N+1} \gamma_k(t) L\big(t, x(t), u_k(t)\big)\, dt \to \inf = m^3_r \\ \text{such that} \\ -\dot{x}(t) \in A(x(t)) + \displaystyle\sum_{k=1}^{N+1} \gamma_k(t) f\big(t, x(t), u_k(t)\big) \text{ a.e. on } [0,b] \\ x(0) = x_0,\ u_k \in S^1_U,\ \gamma_k : T \to [0,1] \text{ measurable and } \displaystyle\sum_{k=1}^{N+1} \gamma_k(t) = 1 \end{array}\right\}. \qquad (6)$$

By $P^3_{r_1}(x_0) \subset C([0,T],\mathbb{R}^N)$ we denote the set of states of problem (6).

The next theorem shows that this new relaxed problem is actually equivalent to the previous ones.

Theorem 10. *If hypotheses* H(A), H(f), H(U)$_2$, H(L)$_1$ *hold and* $x_0 \in \overline{D(A)} = D(A)$*, then* $P^1_{r_1}(x_0) = P^2_{r_1}(x_0) = P^3_{r_1}(x_0)$ *and* $m^3_r = m^2_r = m^1_r$.

Proof. Evidently, $P^1_{r_1}(x_0) \subset P^2_{r_1}(x_0) = P^3_{r_1}(x_0)$. On the other hand, from Carathéodory's theorem mentioned above, if $H(t,x)$ is as in the proof of Proposition 6, we have

$$\overline{\operatorname{conv}} H(t,x) = \left\{ A(x(t)) + \sum_{k=1}^{N+1} \gamma_k f(t,x(t),u_k), \sum_{k=1}^{N+1} \gamma_k L(t,x(t),u_k) \quad \begin{array}{l} \gamma_k \in [0,1] \\ \displaystyle\sum_{k=1}^{N+1} \gamma_k = 1 \\ \{u_k\}_{k=1}^{N+1} \subset U(t) \end{array} \right\}. \qquad (7)$$

From this equality and the proof of Proposition 6 it follows easily that

$$P^1_{r_1}(x_0) = P^2_{r_1}(x_0) = P^3_{r_1}(x_0).$$

Moreover, since

$$p^{**}(t,x(t),\dot{x}(t)) = \inf\Big\{\eta \in \overline{\mathbb{R}} : (-\dot{x}(t),\eta) \in \overline{\operatorname{conv}} H(t,x(t))\Big\}$$

a.e. on $[0,b]$, from (7) we have

$$p^{**}(t,x(t),\dot{x}(t)) = \inf\left\{ \sum_{k=1}^{N+1}\gamma_k L(t,x(t),u_k) \;\middle|\; \begin{array}{l} u_k\in U(t),\ \gamma_k\in[0,1],\ \sum_{k=1}^{N+1}\gamma_k = 1 \\ -\dot{x}(t)\in A(x(t))+\sum_{k=1}^{N+1}\gamma_k f(t,x(t),u_k) \end{array} \right\}$$

a.e. on $[0,b]$. Since the infimum is that of a lower semicontinuous function on a compact set, it is attained. Then an easy measurable selection argument generates measurable $\gamma_k : T\to[0,1]$ and $u_k\in S_U^1$ such that

$\sum_{k=1}^{N+1}\gamma_k(t)=1$

$-\dot{x}(t)\in A(x(t))+\sum_{k=1}^{N+1}\gamma_k(t)f(t,x(t),u_k)$ a.e. on $[0,b]$ and

$p^{**}(t,x(t),\dot{x}(t))=\sum_{k=1}^{N+1}\gamma_k(t)L(t,x(t),u_k(t))$ a.e. on $[0,b]$.

It is now clear that Theorem 8 implies $m_r^3=m_r^2=m_r^1$ ∎

As before, with stronger hypotheses we have admissibility of this relaxation method:

Theorem 11. *If hypotheses* H(A), H(f)$_1$, H(U)$_2$, H(L)$_1$ *hold,* $x_0\in\overline{D(A)} = D(A)$ *and* $m<+\infty$*, then* $P_{r_1}^3(x_0)=\overline{P_1(x_0)}^{C([0,T],\mathbb{R}^N)}$ *and* $m=m_r^1=m_r^2=m_r^3$.

Remark 7. This third relaxation method was first suggested by Ekeland and Temam [12] for a more restricted family of systems.

5. Relaxation via Γ-regularization

In this section we present a fourth relaxation method, quite distinct from the other three, based on semicontinuity arguments and developed by Buttazzo [7]. Roughly speaking the idea is the following. In the cost functional $J(x,u)$ we incorporate all the constraints of the problem (dynamic and non-dynamic) by adding to $J(x,u)$ the indicator function i_Λ of the set Λ of all constraints. Denote the resulting unconstrained cost functional by $H(x,u)$. The new relaxed problem is then obtained by producing the lower semicontinuous envelope (Γ-regularization) of H. The implementation of this method uses the so-called "Γ-limits" (see [7: p. 176]).

Definition. Let X_1,X_2 be two Hausdorff topological spaces and $\varphi: X_1\times X_2\to\overline{\mathbb{R}}=\mathbb{R}\cup\{+\infty\}$ a proper function. In what follows, by $Z(+)$ we denote the "sup" operation and by $Z(-)$ the "inf" operation. Let $(x_1, x_2) \in X_1 \times X_2$ and, for $k = 1, 2$, let S_k be the set of all sequences in X_k which converge to x_k. Also, let α_k be one of the symbols $+$ or $-$. We define

$$\Gamma_{seq}(X_1^{\alpha_1}, X_2^{\alpha_2})\varphi(x_1, x_2) = Z(\alpha_1)_{(x_1^n)\in S_1} Z(\alpha_2)_{(x_2^n)\in S_2} Z(-\alpha_1)_{\bar{n}\geq 1} Z(\alpha_1)_{n\geq \bar{n}} \varphi(x_1^n, x_2^n).$$

If Γ_{seq} is independent of the symbols $+$ or $-$ in one of the spaces, then the symbol is omitted. For example, if

$$\Gamma_{seq}(X_1^-, X_2^+)\varphi(x_1, x_2) = \Gamma_{seq}(X_1^+, X_2^+)\varphi(x_1, x_2),$$

then we write $\Gamma_{seq}(X_1, X_2^+)\varphi(x_1, x_2)$.

Remark 8. If $X_1 \times X_2$ is metrizable or the compact subsets of $X_1 \times X_2$ are metrizable and φ is coercive, then $\Gamma_{seq}(X_1^-, X_2^-)\varphi = \overline{\varphi}$, where $\overline{\varphi}$ is the lower semicontinuous envelope of φ. The problem under consideration is now the following:

$$\left.\begin{array}{l} J(x,u) = \displaystyle\int_0^b L(t, x(t), u(t))dt \to \inf = m \\ \text{such that} \\ -\dot{x}(t) \in A(x(t)) + C(t, x(t))g(t, u(t)) \text{ a.e. on } [0, b] \\ x(0) = x_0,\ u \in S_U^1 \end{array}\right\}. \tag{8}$$

As before, $A:\ D(A) \subset \mathbb{R}^N \to 2^{\mathbb{R}^N}$ is a maximal monotone operator satisfying hypothesis H(A) and $U:\ T \to P_k(\mathbb{R}^m)$ is a multifunction satisfying hypothesis $\mathrm{H(U)}_2$. The precise hypotheses on the other data of problem (8) are the following:

H(C) $C:\ [0, b] \times \mathbb{R}^N \to \mathbb{R}^{N\times k}$ is a function such that:

(i) For all $x \in \mathbb{R}^N$, $t \to C(t, x)$ is measurable.

(ii) There exists $\gamma \in L_+^\infty([0, b])$ such that, for a.a. $t \in [0, b]$ and all $x, z \in \mathbb{R}^N$, $\|C(t, x) - C(t, z)\| \leq \gamma(t)\|x - z\|$.

(iii) There exist $a, c \in L_+^\infty([0, b])$ such that $\|C(t, x)\| \leq a(t) + c(t)\|x\|$ for a.a. $t \in [0, b]$ and all $x \in \mathbb{R}^N$.

H(g) $g:\ [0, b] \times \mathbb{R}^m \to \mathbb{R}^k$ is a measurable function such that:

(i) For all $u \in \mathbb{R}^m$, $t \to g(t, u)$ is measurable.

(ii) For a.a. $t \in [0, b]$, $u \to g(t, u)$ is continuous.

(iii) There exists $\xi \in L_+^1([0, b])$ such that, for a.a. $t \in [0, b]$ and all $u \in \mathbb{R}^m$, $\|g(t, u)\| \leq \xi(t)$.

H(L)$_2$ $L:\ [0, b] \times \mathbb{R}^N \times \mathbb{R}^m \to \mathbb{R}$ is an integrand such that:

(i) For all $(x,u) \in \mathbb{R}^N \times \mathbb{R}^m$, $t \to L(t,x,u)$ is measurable.

(ii) There exists $k \in L^1([0,b])$ such that, for a.a. $t \in [0,b]$, all $x,z \in \mathbb{R}^N$ and all $u \in \mathbb{R}^m$, $|L(t,x,u) - L(t,z,u)| \leq k(t)\|x-z\|$, and there exist $r_1 \in L^1([0,b])$ and $\beta : \mathbb{R}^+ \to \mathbb{R}^+$ convex such that $\lim_{s\to+\infty} \frac{\beta(s)}{s} = +\infty$ and, for a.a. $t \in [0,b]$ and all $u \in \mathbb{R}^m$, $\beta(\|u\|) - r_1(t) \leq L(t,0,u)$.

(iii) For a.a. $t \in [0,b]$ and all $x \in \mathbb{R}^N$, $u \to (t,x,u)$ is continuous.

(iv) For every $n \geq 1$ there exist $\psi_n \in L^1_+([0,b])$ such that $|L(t,x,u)| \leq \psi_n(t)$ for a.a. all $t \in [0,b]$, all $x \in \mathbb{R}^N$ with $\|x\| \leq n$ and all $u \in \mathbb{R}^m$.

As we already mentioned, the idea of this relaxation method is to consider the extended cost functional

$$H(x,u) = J(x,u) + i_\Lambda(x,u),$$

where Λ is the set of all admissible control pairs for problem (8), and to determine its lower semicontinuous envelope $\overline{H}$

- on $W^{1,1}([0,b],\mathbb{R}^N)$ equipped with the $C([0,b],\mathbb{R}^N)$-topology
- and on $L^1([0,b],\mathbb{R}^m)$ with the weak topology

(the last denoted henceforth by $L^1([0,b],\mathbb{R}^m)_w$). So our goal is to find

$$\Gamma_{seq}\big((W^{1,1}([0,b],\mathbb{R}^N),\|\cdot\|_\infty)^-, L^1([0,b],\mathbb{R}^m)_w^-\big)H.$$

To do this we employ the so-called "auxiliary variable method" of Buttazzo (see [7]).

So in what follows we set

$$X = \big(W^{1,1}([0,b],\mathbb{R}^N),\|\cdot\|_\infty\big), \quad W = L^1([0,b],\mathbb{R}^m)_w, \quad V = L^1([0,b],\mathbb{R}^k)_w.$$

According to [7: p. 184] the relaxation (Γ-regularization) of H is reduced to the relaxation (Γ-regularization) in $X \times (W \times V)$ of the functional

$$G(x,u,v) = \Phi(x,u,v) + i_\Delta(x,u,v)$$

where

$$\Phi(x,u,v) = \int_0^b \varphi\big(t,x(t),u(t),v(t)\big)dt$$

with

$$\varphi(t,x,u,v) = L(t,x,u) + i_{\{v=g(t,u),u\in U(t)\}}$$

and

$$\Delta = \left\{ (x,u,v) \in X \times V \times W \;\middle|\; \begin{array}{l} -\dot{x}(t) \in A(x(t)) + C(t,x(t))v(t) \text{ a.e.} \\ x(0) = x_0 \end{array} \right\}.$$

So we need to compute the functional

$$\Gamma_{seq}(X^-, (W \times V)^-)G(x, u, v).$$

Using the definition of Γ_{seq}-limit we can easily check that

$$\begin{aligned}&\Gamma_{seq}(X^-, (W \times V)^-)G(x, u, v)\\&\quad = \Gamma_{seq}(X, (W \times V)^-)\Phi(x, u, v) + \Gamma_{seq}(X^-, W \times V)i_\Delta(x, u, v).\end{aligned}$$

From [7: p. 74] we know that for every

$$(x, u, v) \in W^{1,1}([0, T], \mathbb{R}^N) \times L^1([0, T], \mathbb{R}^m) \times L^1([0, T], \mathbb{R}^k)$$

we have

$$\Gamma_{seq}(X, (W \times V)^-)\Phi(x, u, v) = \int_0^b \varphi^{**}(t, x(t), u(t), v(t))\, dt$$

with φ^{**} being the second conjugate of $\varphi(t, x, \cdot, \cdot)$ in the sense of convex analysis. So it remains to calculate

$$\Gamma_{seq}(X^-, W \times V)i_\Delta(x, u, v).$$

Proposition 12. *If hypotheses* H(A), H(C) *hold and* $x_0 \in \overline{D(A)} = D(A)$, *then* $\Gamma_{seq}(X^-, W \times V)i_\Delta(x, u, v) = i_\Delta(x, u, v)$.

Proof. According to the definition of the Γ_{seq}-limit, we need to prove the following two properties:

(a) If $(x_n, u_n, v_n) \to (x, u, v)$ in $X \times W \times V$ and $(x_n, u_n, v_n) \in \Delta$ for all $n \geq 1$, then $(x, u, v) \in \Delta$.

(b) If $(x, u, v) \in \Delta$, then for all $(u_n, v_n) \to (u, v)$ in $W \times V$ we can find a sequence $x_n \to x$ in X such that $(x_n, u_n, v_n) \in \Delta$ for all $n \geq 1$ large.

First we show that property (a) holds. So let $x_n \to x$ in $C([0, b], \mathbb{R}^N)$, $v_n \to v$ weakly in $L^1([0, b], \mathbb{R}^k)$ and

$$\left.\begin{aligned}&-\dot{x}_n(t) \in A(x_n(t)) + C(t, x_n(t))v_n(t) \quad \text{a.e. on } [0, b]\\&x_n(0) = x_0\end{aligned}\right\}.$$

Note that $C(\cdot, x_n(\cdot))v_n(\cdot) \to C(\cdot, x(\cdot))v(\cdot)$ weakly in $L^1([0, b], \mathbb{R}^N)$ (hypothesis H(C)). Also, recall that $\{\dot{x}_n\}_{n \geq 1} \subset L^1([0, b], \mathbb{R}^N)$ is relatively weakly compact. So we may assume that $\dot{x}_n \to w$ weakly in $L^1([0, b], \mathbb{R}^N)$ and clearly $w = \dot{x}$. Moreover, from [17: p. 694] we have

$$-\dot{x}(t) - C(t, x(t))v(t) \in \overline{\text{conv}} \limsup_{n \to \infty} A(x_n(t)) \subset A(x(t))$$

a.e. on $[0,b]$, the last inclusion following from the fact that GrA is closed and A has closed and convex values on $\overline{D(A)} = D(A)$ (since it is maximal monotone). Hence

$$\left.\begin{aligned} &-\dot{x}(t) \in A(x(t)) + C(t,x(t))v(t) \ \text{ a.e. on } [0,b] \\ &x(0) = x_0 \end{aligned}\right\}.$$

This proves property (a).

Next we show that property (b) holds. So let $(x,u,v) \in \Delta$ and suppose that $v_n \to v$ weakly in $L^1([0,b],\mathbb{R}^k)$. Let $y_n \in W^{1,1}([0,b],\mathbb{R}^N) \subset C([0,b],\mathbb{R}^N)$ be the unique solution of

$$\left.\begin{aligned} &-\dot{y}_n(t) \in A(y_n(t)) + C(t,x(t))v_n(t) \ \text{ a.e. on } [0,b] \\ &y_n(0) = x_0 \end{aligned}\right\}.$$

We know (see [16, 22]) that $\{y_n\}_n \subset C([0,b],\mathbb{R}^N)$ is relatively compact, and so we may assume that $y_n \to y$ in $C([0,b],\mathbb{R}^N)$. Because $C(\cdot,x(\cdot))v_n(\cdot) \to C(\cdot,x(\cdot))v(\cdot)$ weakly in $L^1([0,b],\mathbb{R}^N)$, so as above in the limit as $n \to \infty$ we obtain

$$\left.\begin{aligned} &-\dot{y}(t) \in A(y(t)) + C(t,x(t))v(t) \ \text{ a.e. on } [0,b] \\ &y(0) = x_0 \end{aligned}\right\}$$

and hence $y = x$. Therefore $y_n \to x$ in $C([0,b],\mathbb{R}^N)$.

Now let $x_n \in W^{1,1}([0,b],\mathbb{R}^N) \subset C([0,b],\mathbb{R}^N)$ be the unique solution of

$$\left.\begin{aligned} &-\dot{x}_n(t) \in A(x_n(t)) + C(t,x_n(t))v_n(t) \ \text{ a.e. on } [0,b] \\ &x_n(0) = x_0 \end{aligned}\right\}.$$

Exploiting the monotonicity of A, we get

$$\begin{aligned} &\Big(-\dot{x}_n(t) + \dot{y}_n(t), y_n(t) - x_n(t)\Big)_{\mathbb{R}^N} \\ &\qquad \le \Big(C(t,x_n(t))v_n(t) - C(t,x(t))v_n(t), y_n(t) - x_n(t)\Big)_{\mathbb{R}^N} \end{aligned}$$

a.e. on $[0,b]$, so

$$\begin{aligned} &\frac{1}{2}\frac{d}{dt}\|x_n(t) - y_n(t)\|^2 \\ &\qquad \le \Big(C(t,x_n(t))v_n(t) - C(t,x(t))v_n(t), y_n(t) - x_n(t)\Big)_{\mathbb{R}^N} \end{aligned}$$

a.e. on T and therefore

$$\begin{aligned}&\|x_n(t)-y_n(t)\|^2\\&\leq 2\int_0^t \big\|\big(C(s,x_n(s))-C(s,x(s))\big)v_n(s)\big\|\|y_n(s)-x_n(s)\|\,ds\end{aligned}$$

for all $t\in[0,b]$. Using [6: p. 157/Lemma A.5] and hypothesis H(C)/(ii), for all $t\in[0,b]$ we obtain

$$\begin{aligned}\|x_n(t)-y_n(t)\|&\leq 2\int_0^t \big\|\big(C(s,x_n(s))-C(s,x(s))\big)v_n(s)\big\|ds\\&\leq 2\int_0^t \gamma(s)\|x_n(s)-x(s)\|ds.\end{aligned}$$

So

$$\begin{aligned}\|x_n(t)-y_n(t)\|\leq{}& 2\int_0^t \gamma(s)\|x_n(s)-y_n(s)\|\,\|v_n(s)\|ds\\&+2\int_0^t \gamma(s)\|y_n(s)-x(s)\|\,\|v_n(s)\|ds\end{aligned}$$

for all $t\in[0,b]$. We know that $y_n\to x$ in $C([0,b],\mathbb{R}^N)$ and $v_n\to v$ weakly in $L^1([0,b],\mathbb{R}^k)$. So given $\varepsilon>0$ we can find $n_o=n_o(\varepsilon)\geq 1$ such that, for all $n\geq n_o$,

$$\|x_n(t)-y_n(t)\|\leq \varepsilon+2\int_0^t \gamma(s)\|x_n(s)-y_n(s)\|\,\|v_n(s)\|ds$$

for all $t\in[0,b]$ which implies that (using the Gronwall inequality) $\|x_n(t)-y_n(t)\|\leq \hat{k}\varepsilon\|\gamma\|_\infty$ for all $n\geq n_o$ and all $t\in[0,b]$, for some $\hat{k}>0$. Therefore $x_n-y_n\to 0$ in $C([0,b],\mathbb{R}^N)$ and so $x_n\to x$ in $C([0,b],\mathbb{R}^N)$. This proves property (b) and so we have proved the proposition ∎

Therefore we can write

$$\overline{H}(x,u)=\inf\left\{\int_0^b \varphi^{**}\ t,x(t),u(t),v(t)\ dt \quad \begin{matrix} v\in L^1([0,b],\mathbb{R}^k)\\ -\dot{x}(t)\in A(x(t))+C(t,x(t))v(t)\ \text{a.e.}\\ x(0)=x_0\end{matrix}\right\}. \tag{9}$$

Recall that

$$\varphi(t,x,u,v)=L(t,x,u)+i_{\{v=g(t,u),u\in U(t)\}}$$

and the double convex conjugation is with respect to the variables u and v. Of course, (9) is not a satisfactory formulation of the relaxed problem because of the presence of the auxiliary variable v. So our goal is to eliminate this auxiliary variable. To this end let

$$E(t,u) = \Big\{ v \in \mathbb{R}^k : (u,v) \in \overline{\operatorname{conv}} \operatorname{Gr} g(t,\cdot),\ u \in U(t) \Big\}$$

and set

$$L_o(t,x,u,z) = \inf \Big\{ \varphi^{**}(t,x,u,v) : -z \in A(x) + C(t,x)v \Big\}$$

and

$$\Delta_o = \left\{ (x,u) \in X \times W \;\middle|\; \begin{array}{l} -\dot{x}(t) \in A(x(t)) + C(t,x(t))E(t,u(t)) \text{ a.e.} \\ x(0) = x_0 \end{array} \right\}.$$

Using these items we can now formulate the relaxed problem which corresponds to this fourth relaxation method.

Proposition 13. *If hypotheses* H(A), H(C), H(g), $\mathrm{H(U)}_2$, $\mathrm{H(L)}_2$ *hold and* $x_0 \in \overline{D(A)} = D(A)$, *then*

$$\begin{aligned} \overline{H}(x,u) &= \Gamma_{seq}(X^-, W^-)H(x,u) \\ &= \int_0^b L_o\big(t, x(t), u(t), \dot{x}(t)\big)dt + i_{\Delta_o}(x,u). \end{aligned}$$

Proof. Note that if $\varphi^{**}(t,x,u,v) < +\infty$, then $v \in E(t,u)$. So we have

$$\overline{H}(x,u) = \inf \left\{ \int_0^b \varphi^{**}\big(t, x(t), u(t), v(t)\big)dt \;\middle|\; \begin{array}{l} v \in L^1([0,b], \mathbb{R}^k) \\ -\dot{x}(t) \in A(x(t)) + C(t,x(t))v(t) \text{ a.e. on } T \end{array} \right\} + i_{\Delta_0}(x,u).$$

For every $v \in L^1([0,b], \mathbb{R}^k)$ with $-\dot{x}(t) \in A(x(t)) + C(t,x(t))v(t)$ a.e. on $[0,b]$ we have

$$L_o\big(t, x(t), u(t), \dot{x}(t)\big) \le \varphi^{**}\big(t, x(t), u(t), v(t)\big)$$

a.e. on $[0,b]$ and so

$$\int_0^b L_o\big(t, x(t), u(t), \dot{x}(t)\big)dt + i_{\Delta_o}(x,u) \le \overline{H}(x,u).$$

So we need to show that the opposite inequality also holds. To this end let $(x,u) \in \Delta_o$ and suppose that $\int_0^b L_o\big(t, x(t), u(t), \dot{x}(t)\big)dt < +\infty$ (otherwise there is nothing to prove). Let

$$K(t) = \Big\{ v \in \mathbb{R}^k : -\dot{x}(t) \in A(x(t)) + C(t,x(t))v \Big\}.$$

Note that $\eta : [0,b] \times \mathbb{R}^k \to \mathbb{R}^N \times \mathbb{R}^N$ defined by

$$\eta(t,v) = \big(x(t), -\dot{x}(t) - C(t,x(t))v\big)$$

is a Carathéodory function (i.e. measurable in $t \in [0,b]$ and continuous in $v \in \mathbb{R}^k$), hence it is jointly measurable. Since GrA is closed, we have

$$\mathrm{Gr}K = \Big\{(t,v) \in [0,b] \times \mathbb{R}^k : \eta(t,v) \in GrA\Big\} \in B([0,b]) \times B(\mathbb{R}^k),$$

i.e. the multifunction K is graph measurable. Then

$$\begin{aligned}
&\inf \left\{ \int_0^b \varphi^{**}\; t, x(t), u(t), v(t)\; dt \quad \begin{matrix} v \in L^1([0,b],\mathbb{R}^k) \\ -\dot{x}(t) \in A(x(t)) + C(t,x(t))v(t) \text{ a.e.} \end{matrix} \right\} \\
&\quad = \inf \quad \int_0^b \varphi^{**}\; t, x(t), u(t), v(t)\; dt : v \in S^1_K \\
&\quad = \int_0^b \inf \Big\{ \varphi^{**}(t, x(t), u(t), v) : \; v \in K(t) \Big\} dt \\
&\quad = \int_0^b L_o\; t, x(t), u(t), \dot{x}(t)\; dt
\end{aligned}$$

(see [17: p. 183]). Therefore

$$\overline{H}(x,u) \le \int_0^b L_o\big(t, x(t), u(t), \dot{x}(t)\big) dt + i_{\Delta_o}(x,u).$$

We conclude that

$$\overline{H}(x,u) = \int_0^b L_o\big(t, x(t), u(t), \dot{x}(t)\big) dt + i_{\Delta_o}(x,u).$$

and the proposition is proved ∎

Proposition 13 gives us the fourth relaxed problem, which is

$$\int_0^b L_o\big(t, x(t), u(t), \dot{x}(t)\big) dt + i_{\Delta_o}(x,u) \to \inf = m_r^4. \tag{10}$$

Invoking [7: p. 16/Proposition 1.3.1] we obtain the solvability of problem (10).

Theorem 14. *If hypotheses* H(A), H(C), H(g), H(U)$_2$, H(L)$_2$ *hold and* $x_0 \in \overline{D(A)} = D(A)$*, then problem* (10) *admits an optimal pair* $(x, u) \in W^{1,1}([0, b], \mathbb{R}^N) \times L^1([0, b], \mathbb{R}^m)$.

What is important to us is the admissibility of this new relaxation method and the relation of the resulting relaxed problem and the other three problems established in Section 4. By showing that this fourth relaxation method is equivalent to the other three, we shall have the answer to both questions. The task is non-trivial since the procedure in this relaxation method is quite distinct from the other three, and so a priori it is not at all clear that there is any relation between them.

From Section 4 we know that to problem (8) we can associate the relaxed problem

$$\left.\begin{aligned} & J_r^1(x,\lambda) = \int_0^b \int_{\overline{B}_r} L(t, x(t), u)\lambda(t)(du)dt \to \inf = m_r^1 \\ & \text{such that} \\ & -\dot{x}(t) \in A(x(t)) + C(t, x(t)) \int_{\overline{B}_r} g(t,u)\lambda(t)(du) \text{ a.e. on } [0,b] \\ & x(0) = x_0, \ \lambda \in S_\Sigma \end{aligned}\right\}. \tag{11}$$

In what follows, given $u \in S_U^1$ we can define the "barycenter" of u to be the set

$$\mathrm{Bar}(u) = \left\{ \lambda \in S_\Sigma \,\middle|\, u(t) = \int_{\overline{B}_r} u\lambda(t)(du) \ \text{ a.e. on } [0,b] \right\}.$$

Proposition 15. *If hypotheses* H(A), H(C), H(g), H(U)$_2$, H(L)$_2$ *hold and* $x_0 \in \overline{D(A)} = D(A)$*, then*

$$\overline{H}(x,u) = \min \left\{ J_r^1(x,\lambda) \;\; \begin{aligned} & -\dot{x}(t) \in A(x(t)) + C(t,x(t)) \int_{\overline{B}_r} g(t,u)\lambda(t)(du) \text{ a.e.} \\ & x(0) = x_0 \\ & \lambda \in \mathrm{Bar}(u) \end{aligned} \right\}.$$

Proof. Let $(x, \lambda) \in W^{1,1}([0, b], \mathbb{R}^N) \times S_\Sigma$ be an admissible state-control pair for problem (11) such that $\lambda \in \mathrm{Bar}(u)$. From Lemma 3 we can find a sequence $\{u_n\}_{n\geq 1} \subset S_U^1$ such that $\delta_{u_n} \to \lambda$ in $R([0, b], \overline{B}_r)$. Let $x_n \in W^{1,1}([0, b], \mathbb{R}^N)$ be the unique state for problem (8) generated by the control u_n. We know that the sequence $\{x_n\}_{n\geq 1} \subset C([0, b], \mathbb{R}^N)$ is relatively compact and so, by passing to a subsequence if necessary, we may assume that $x_n \to x$

in $C([0,b],\mathbb{R}^N)$. For all $n \geq 1$ we have $u_n(t) = \int_{\overline{B}_r} u\delta_{u_n(t)}(du)$ and for all $D \in B([0,b])$ we have

$$\int_D \int_{\overline{B}_r} u\delta_{u_n(t)}(du)dt \to \int_D \int_{\overline{B}_r} u\lambda(t)(du)dt.$$

If $u(t) = \int_{\overline{B}_r} u\lambda(t)(du)dt$, we have

$$\int_D u_n(t)dt \to \int_D u(t)dt$$

for all $D \in B([0,b])$ and so $u_n \to u$ weakly in $L^1([0,b],\mathbb{R}^m)$ with $u \in S^1_U$. Also, because (x_n,u_n) is admissible for problem (8), $H(x_n,u_n) = J(x_n,u_n)$ and so

$$\liminf_{n\to\infty} H(x_n,u_n) = \liminf_{n\to\infty} J(x_n,u_n) = \liminf_{n\to\infty} J^1_r(x_n,\delta_{u_n}).$$

Using hypothesis $\mathrm{H(L)}_2$/(ii) we can see that $J^1_r(x_n,\delta_{u_n}) \to J^1_r(x,\lambda)$ and so $\liminf_{n\to\infty} H(x_n,u_n) = J^1_r(x,\lambda)$. Therefore

$$\overline{H}(x,u) \leq \inf\left\{J^1_r(x,\lambda) : (x,\lambda) \in P^1_r(x_o), \lambda \in \mathrm{Bar}(u)\right\}. \tag{12}$$

On the other hand, if $\overline{H}(x,u) < +\infty$, from the definition of

$$\overline{H}(x,u) = \Gamma_{seq}(X^-,W^-)H(x,u),$$

given $\varepsilon > 0$ we can find a sequence $\{(x_n,u_n)\}_{n\geq 1} \subset \Lambda$ such that

$$\begin{aligned} x_n &\to x \ \text{ in } C^1([0,b],\mathbb{R}^N) \\ u_n &\to u \text{ weakly in } L^1([0,b],\mathbb{R}^m) \\ \liminf_{n\to\infty} J(x_n,u_n) &\leq \overline{H}(x,u) + \varepsilon. \end{aligned}$$

By passing to a subsequence if necessary, we may assume that $\delta_{u_n} \to \lambda$ weakly in $R([0,b],\overline{B}_r)$. For every $D \in B([0,b])$ we have

$$\int_D u_n(t)\,dt = \int_D \int_{\overline{B}_r} u\delta_{u_n(t)}(du)dt \to \int_D \int_{\overline{B}_r} u\lambda(t)(du)dt$$
$$\int_D u_n(t)\,dt \to \int_D u(t)\,dt.$$

Hence $u(t) = \int_{\overline{B}_r} u\lambda(t)(du)$ a.e. on T and so $\lambda \in \mathrm{Bar}(u)$. Also, by virtue of hypothesis $\mathrm{H(L)}_2$, $J(x_n,u_n) \to J^1_r(x,\lambda)$ and so $J^1_r(x,\lambda) \leq \overline{H}(x,u) + \varepsilon$ with $\lambda \in \mathrm{Bar}(u)$ and $(x,\lambda) \in P^1_r(x_o)$. Since $\varepsilon > 0$ was arbitrary, it follows that

$$\inf\left\{J^1_r(x,\lambda) : (x,\lambda) \in P^1_r(x_o),\ \lambda \in \mathrm{Bar}(u)\right\} \leq \overline{H}(x,u). \tag{13}$$

From (12) and (13) we conclude that equality must hold. Moreover, it is clear that the infimum is attained ∎

From Proposition 15 we obtain at once the equivalence of problems (10) and (11).

Theorem 16. *If hypotheses* H(A), H(C), H(g), $\mathrm{H(U)}_2$, $\mathrm{H(L)}_2$ *hold and* $x_0 \in \overline{D(A)} = D(A)$, *then the relaxed problems* (10) *and* (11) *are equivalent. More precisely, if* $(x,u) \in \Lambda$ *solves problem* (10), *then there exists* $\lambda \in \mathrm{Bar}(u)$ *such that* $(x,\lambda) \in P_r^1(x_o)$ *solves problem* (11) *and conversely, if* $(x,\lambda) \in P_r^1(x_o)$ *solves problem* (11), *then we can find* $u \in S_U^1$ *such that* $\lambda \in \mathrm{Bar}(u)$ *and* (x,u) *is a solution of problem* (10). *Moreover, we have* $m_r^4 = m_r^1 = m$.

From Section 4 we know that we can have two more relaxed problems, which are

$$\left.\begin{aligned} & J_r^2(x) = \int_0^b p^{**}(t,x(t),\dot{x}(t))dt \to \inf = m_r^2 \\ & \text{such that} \\ & -\dot{x}(t) \in A(x(t)) + \overline{\mathrm{conv}} F(t,x(t)) \text{ a.e. on } [0,b] \\ & x(0) = x_0 \end{aligned}\right\} \tag{14}$$

where $F(t,x) = C(t,x)g(t,U(t))$ and

$$\left.\begin{aligned} & J_r^3(x,\hat{u},\hat{\gamma}) = \int_0^b \sum_{k=1}^{N+1} \gamma_k(t) L\big(t,x(t),u_k(t)\big)dt \to \inf m_r^3 \\ & \text{such that} \\ & -\dot{x}(t) \in A(x(t)) + C(t,x(t)) \sum_{k=1}^{N+1} \gamma_k(t) g(t,u_k(t)) \text{ a.e. on } [0,b] \\ & x(0) = x_0, u_k \in S_U^1, \gamma_k : T \to [0,1] \text{ measurable}, \sum_{k=1}^{N+1} \gamma_k(t) = 1 \end{aligned}\right\}. \tag{15}$$

Combining Theorem 16 with the results of Section 4, we obtain

Theorem 17. *If hypotheses* H(A), H(C), H(g), $\mathrm{H(U)}_2$, $\mathrm{H(L)}_2$ *hold and* $x_0 \in \overline{D(A)} = D(A)$, *then problems* (10), (11), (14) *and* (15) *are equivalent* (*in the sense of Theorem* 16) *and* $m_r^1 = m_r^2 = m_r^3 = m_r^4 = m$.

Remark 9. Buttazzo in [7] did not investigate the relation of his relaxation method with the other methods existing in the literature. So our work in this section in addition of extending the work of Buttazzo [7] (since our model system is more general) it also complements it. Finally, note that our general framework in this paper also incorporates gradient systems with non-smooth potential. This is the case when $A = \partial\varphi$ with $\varphi : \mathbb{R}^N \to \mathbb{R}$ continuous, convex but not necessarily differentiable.

References

[1] Attouch, A.: *Variational Convergence for Functions and Operators.* London: Pitman 1984.

[2] Berkovitz, L. D.: *Existence theorems in problems of optimal control.* Studia Math. 44 (1972), 275 – 285.

[3] Berkovitz, L. D.: *Optimal Control Theory.* New York: Springer-Verlag 1974.

[4] Berkovitz, L. D.: *Lower semicontinuity of integral functional.* Trans. Amer. Math. Soc. 192 (1974), 51 – 57.

[5] Berkovitz, L. D.: *Existence theory for optimal control problems.* In: Optimal Control and Differential Equations (eds.: A. B. Schwartzkopf et al.). New York: Acad. Press 1978, pp. 107 – 130.

[6] Brezis, H.: *Operateurs Maximaux Monotones.* Amsterdam: North-Holland 1973.

[7] Buttazzo, G.: *Semicontinuity, Relaxation and Representation in the Calculus of Variations.* Harlow, Essex (UK): Longman 1989.

[8] Cesari, L.: *Existence theorems for weak and usual optimal solutions in Lagrange problems with unilateral constraints*, Parts I and II. Trans. Amer. Math. Soc. 124 (1966), 369 – 412 and 413 – 429.

[9] Cesari, L.: *Optimization Theory and Applications.* New York: Springer-Verlag 1983.

[10] Cornet, B.: *Existence of slow solutions for a class of differential inclusions.* J. Math. Anal. Appl. 96 (1983), 130 – 147.

[11] Dal Maso, G.: *An Introduction to* Γ*-Convergence.* Boston: Birkhäuser 1993.

[12] Ekeland, I. and R. Temam: *Convex Analysis and Variational Problems.* Amsterdam: North-Holland 1976.

[13] Filippov, A. F.: *On certain questions in the theory of optimal control.* SIAM J. Control 1 (1962), 76 – 84.

[14] Gamkrelidze, R. V.: *Principles of Optimal Control Theory.* New York: Plenum Press 1978.

[15] Henry, C.: *Differential equations with discontinuous right hand side for planning procedures.* J. Econ. Theory 4 (1972), 545 – 551.

[16] Hu, S. and N. S. Papageorgiou: *Time-Dependent Subdifferential Evolution Inclusions and Optimal Control* (Memoirs of the Amer. Math. Soc.: Vol. 632). Providence (RI, USA): Amer. Math. Soc. 1998.

[17] Hu, S. and N. S. Papageorgiou: *Handbook of Multivalued Analysis.* Volume I: *Theory.* Dordrecht (The Netherlands): Kluwer 1997.

[18] Hu, S. and N. S. Papageorgiou: *Handbook of Multivalued Analysis.* Volume II: *Applications.* Dordrecht (The Netherlands): Kluwer 2000.

[19] Ionescu-Tulcea, A. and Ionescu-Tulcea, C.: *Topics in the Theory of Lifting.* Berlin: Springer-Verlag 1969.

[20] Krasnoselskii, M. A. and A. V. Pokrovskii: *Systems with Hysteresis.* Berlin: Springer-Verlag 1988.

[21] Olech, C.: *A characterization of L_1-weak lower semicontinuity of integral functionals.* Bull. Acad. Polon. Sci. 25 (1977), 135 – 142.

[22] Papageorgiou, N. S.: *Differential inclusions with state constraints.* Proc. Edinburgh Math. Soc. 32 (1988), 81 – 97.

[23] Parthasarathy, K. R.: *Probability Measures on Metric Spaces.* New York: Acad. Press 1967.

[24] Roubicek, J.: *Relaxation in Optimization Theory and Variational Calculus.* Berlin: de Gruyter 1997.

[25] Warga, J.: *Optimal Control of Differential and Functional Equations.* New York: Acad. Press 1972.

Received 14.03.2003