Commun. Optim. Theory 2022 (2022) 6 https://doi.org/10.23952/cot.2022.6

MATHRES

Communications in Optimization Theory

Available online at http://cot.mathres.org

OPTIMAL CONTROL OF A CLASS OF SEMILINEAR SYSTEMS ON BANACH SPACES DRIVEN BY VECTOR MEASURES AND RELAXED CONTROLS

N.U. AHMED

University of Ottawa, Ottawa, Ontario, Canada

Dedicated in the memory of 100th birthday of Professor Jack Warga

Abstract. In this paper, we consider a class of semi linear systems on Banach spaces driven by relaxed controls (probability measure valued functions) and vector measures. We present existence of optimal control policies and develop necessary conditions of optimality whereby one can determine the optimal controls. Based on the necessary conditions of optimality, we present an algorithm including a proof of its convergence whereby the optimal policies can be constructed. Further, we consider non-convex control problems as special cases where the relaxed controls specialize to switching controls, generalizing the bang-bang principle.

Keywords. Existence of optimal controls; Necessary conditions of optimality; Relaxed controls; Semi-linear systems; Vector Measures.

1. System Dynamics and Optimization Problem

Let X,E be a pair of real Banach spaces and U a Polish space (complete separable metric space), and $\mathscr{M}_1(U)$ the space of probability measures on U. In this paper, we consider the following system on the Banach space X

$$dx = Axdt + \hat{F}(t,x,u_t)dt + G(t,x)\gamma(dt), x(0) = x_0, t \in I \equiv [0,T], \tag{1.1}$$

where A is a densely defined closed linear operator with domain and range in X generating a C_0-semigroup of operators $\{S(t), t \in I\} \subset \mathscr{L}(X)$, and $F : I \times X \times U \longrightarrow X$, and $G : I \times X \longrightarrow \mathscr{L}(E,X)$ are Borel measurable maps. The system is driven by a pair of controls denoted by $\{u,\gamma\}$, where $\hat{F}(t,x,u_t) \equiv \int_U F(t,x,\xi)u_t(d\xi)$ with u denoting the relaxed control, a measure valued function with values in $\mathscr{M}_1(U)$, and γ is an E valued vector measure containing impulsive controls as special cases.

Let $\mathscr{R}_{ad}$ denote the class of relaxed controls and $\mathscr{M}_{ad}$ the class of E-valued vector measures also considered as controls. In other words, the system is equipped with dual controls thereby

E-mail address: nahmed@uottawa.ca.

Received January 8, 2022; Accepted June 27, 2022.

increasing the reliability. The problem is to find a control $\nu = (u,\gamma) \in \mathscr{R}_{ad} \times \mathscr{M}_{ad}$ that minimizes the following cost functional

$$J(\nu) \equiv J(u,\gamma) = \int_I \ell(t,x(t))dt + \Phi(x(T)). \tag{1.2}$$

Since system (1.1) is subject to vector measures, which may contain Dirac measures, it is expected that the solutions may not be continuous. Thus $C(I,X)$ is not a suitable space for the solution trajectories, and hence we consider the space of (norm) bounded measurable functions on I with values in X. This is denoted by $B_\infty(I,X)$, which is endowed with norm topology

$$\| x \|_{B_\infty(I,X)} \equiv sup\{\| x(t) \|_X, t \in I\}.$$

With respect to this norm topology, it is a Banach space. Note that this space contains the class of piecewise continuous and bounded functions which in turn contains $C(I,X)$.

Remark 1.1. Warga's pioneering contribution to optimal control theory for non-convex control problems related to Functional-Differential equations with Radon measures is unique and inspiring, which is clearly reflected in his excellent book [9] and many of his publications. Here we consider optimal control problems for the systems governed by differential equations on Banach spaces determined by unbounded operators. Many of the results presented in this paper can be extended to Functional-Differential equations containing unbounded operators covering many results of Warga [9].

2. Admissible Controls

To solve the control problem, we need a full characterization of the set of admissible controls. Let U denote a compact Polish space, and $C(U)$ the Banach space of real valued continuous functions defined on U equipped with the standard supnorm topology. The topological dual of this space is given by the space of regular Borel measures $\mathscr{M}(U)$. Let $L_1(I,C(U))$ denote the Lebesgue-Bochner space of integrable functions defined on I with values in the Banach space $C(U)$. Since Banach space $C(U)$ does not satisfy Radon-Nikodym property [3] its topological dual is not given by $L_\infty(I,\mathscr{M}(U))$. However, it follows from the theory of lifting that its dual is given by the space of weak star measurable functions defined on I taking values in $\mathscr{M}(U)$ and denoted by $L^w_\infty(I,\mathscr{M}(U))$. The relaxed controls are weak star measurable functions defined on the interval I and taking values in $\mathscr{M}_1(U) \subset \mathscr{M}(U)$, where $\mathscr{M}_1(U)$ denotes the class of regular Borel probability measures on U. This class of controls is denoted by $\mathscr{R}_{ad}$, and it consists of weak star measurable functions defined on I with values in the space of probability measures $\mathscr{M}_1(U)$. Clearly, by virtue of Alaoglu's theorem, $\mathscr{R}_{ad} \subset L^w_\infty(I,\mathscr{M}(U))$ is a weak star compact set. For the second class of controls, let E be a real Banach space and consider the space of E valued finitely additive vector measures of bounded variations, denoted by $\mathscr{M}_{bfa}(\mathscr{A}_I,E)$, where $\mathscr{A}_I$ is the algebra of subsets of the set I. Endowed with the total variation norm, $\mathscr{M}_{bfa}(\mathscr{A}_I,E)$ is a Banach space. Here we consider the smaller class $\mathscr{M}_{bfa}(\Sigma_I,E)$. Let $\mathscr{M}_{ad}$ be a weakly compact subset of $\mathscr{M}_{bfa}(\Sigma_I,E)$ denoting the second class of admissible controls. Thus system (1.1) is furnished with dual controls. In other words, for admissible controls, we choose the set $\mathscr{K}_{ad} \equiv \mathscr{R}_{ad} \times \mathscr{M}_{ad}$ endowed with the product topology $\tau_p = \tau_{w^*} \times \tau_w$, where τ_{w^*} and τ_w denote the weak star and weak topologies respectively. An element of $\mathscr{K}_{ad}$ is denoted by the pair $\nu = (u,\gamma)$. In the study of optimal controls, we use the weak compactness of the admissible set $\mathscr{M}_{ad}$. Here we present a result characterizing weakly compact sets in the Banach space of bounded finitely additive measures $\mathscr{M}_{bfa}(\mathscr{A}_I, E)$.

Theorem 2.1. *Suppose both E and its dual E^* satisfy RNP (Radon Nikodym property). Then a set $\mathscr{M} \subset \mathscr{M}_{bfa}(\mathscr{A}_I, E)$ is relatively weakly compact if, and only if,*

(c1): the set $\mathscr{M}$ is bounded in norm (total variation norm),

(c2): there exists a nonnegative finitely additive measure $\mu \in \mathscr{M}^+_{bfa}(\mathscr{A}_I)$ such that $\lim_{\mu(\sigma)\to 0} |\gamma|(\sigma) = 0$ uniformly with respect to $\gamma \in \mathscr{M}$,

(c3): for every $\Delta \in \mathscr{A}_I$, the set $\{\gamma(\Delta), \gamma \in \mathscr{M}\}$ is a relatively weakly compact subset of E.

Proof. See Diestel Jr. [5, Corollary 6, p105]. □

Theorem 2.1 is due to Brooks and Dinculeanu [5, Corollary 6, p105] which itself is a generalization of the celebrated theorem due to Bartle-Dundord-Schwartz [5, Theorem 5, p105]. For details on vector measures, we refer to Diestel [5, 6] and Dunford [7].

Remark 2.2. In a recent paper [4], we considered vector measures only as controls for a class of systems governed by differential equations on Banach spaces. Here we extend this result and consider a combination of vector measures and probability measure valued functions (also known as relaxed controls) as admissible controls thereby covering non-convex control problems and increasing reliability of the system.

3. Existence and Regularity Properties of Solutions

In this section, we prove the existence and uniqueness of solutions of the system given by equation (1.1). Following this, we prove continuous dependence of solution with respect the controls. For this purpose we introduce the following basic assumptions,

Assumptions

(A): The operator A is the infinitesimal generator of a C_0-semigroup [1] of bounded linear operators $\{S(t) \in \mathscr{L}(X), t \in I\}$ in the Banach space X, and there exists a finite positive number M such that $\sup\{\| S(t) \|_{\mathscr{L}(X)}, t \in I\} \le M$.

(F): $F : I \times X \times U \longrightarrow X$ is a Borel measurable measurable map in all the variables and continuous in the last two arguments satisfying the following growth and Lipschitz properties:

There exists a nonnegative constant K such that

$(\mathbf{F1}): \; \| F(t,x,\xi) \|_X \le K(1 + \| x \|_X) \; \forall \, (t,x) \in I \times X,$ uniformly with respect to $\xi \in U$.

$(\mathbf{F2}): \; \| F(t,x,\xi) - F(t,y,\xi) \|_X \le K \| x - y \|_X \;\; \forall \, (t,x,y) \in I \times X \times X,$ uniformly with respect to $\xi \in U$.

(G): $G : I \times X \longrightarrow \mathscr{L}(E,X)$ is a Borel measurable measurable map in all the variables and continuous in the last argument satisfying the following growth and Lipschitz properties:

There exists a nonnegative constant L such that

$$(\mathbf{G1}): \; \| G(t,x) \|_{\mathscr{L}(E,X)} \le L(1 + \| x \|_X) \; \forall \, (t,x) \in I \times X,$$

$$(\mathbf{G2}): \; \| G(t,x) - G(t,y) \|_{\mathscr{L}(E,X)} \le L \| x - y \|_X \;\; \forall \, (t,x,y) \in I \times X \times X.$$

Using the above assumptions, we prove the following result.

Theorem 3.1. *Consider the system (1.1) and suppose the assumptions (A),(F), and (G) hold. Then, for each initial state $x_0 \in X$ and control $\nu = (u,\gamma) \in \mathscr{K}_{ad}$, the system governed by the evolution equation (1.1) has a unique mild solution $x \in B_\infty(I,X)$.*

Proof. By a mild solution of equation (1.1), we mean a solution of the following integral equation

$$x(t) = S(t)x_0 + \int_0^t S(t-s)\hat{F}(s,x(s),u_s)ds + \int_0^t S(t-s)G(s,x(s))\gamma(ds), t \in I. \tag{3.1}$$

To prove the existence and uniqueness of solutions of the above integral equation, we use the Banach fixed point theorem. For any given control $\nu = (u,\gamma) \in \mathscr{K}_{ad}$, let us introduce the operator Γ on the Banach space $B_\infty(I,X)$ as follows

$$(\Gamma x)(t) = S(t)x_0 + \int_0^t S(t-s)\hat{F}(s,x(s),u_s)ds + \int_0^t S(t-s)G(s,x(s))\gamma(ds), t \in I. \tag{3.2}$$

First, we verify that Γ maps $B_\infty(I,X)$ into itself. By virtue of the assumptions (A), $(F1)$, $(G1)$, and triangle inequality, it follows from the expression (3.2) that

$$\begin{aligned} &\| (\Gamma x)(t) \|_X \leq M \| x_0 \| + M(KT + L|\gamma|(I)) \\ &+ MK \int_0^t \| x(s) \|_X \, ds + ML \int_0^t \| x(s) \|_X \, |\gamma|(ds), t \in I, \end{aligned} \tag{3.3}$$

where $|\gamma|(\cdot)$ denotes the nonnegative finitely additive measure induced by the variation of the vector measure γ. For any $\Delta \in \Sigma_I$, the measure $|\gamma|(\Delta)$ induced by the variation is given by

$$|\gamma|(\Delta) \equiv \sup_{\Pi} \sum_{\sigma \in \Pi} \| \gamma(\sigma) \|_E,$$

where Π denotes any partition of the set Δ by a finite number of disjoint Σ_I measurable sets contained in Δ. The summation is taken over all $\sigma \in \Pi$, and the supremum is taken over all such disjoint Σ_I measurable partitions of the set Δ. Clearly, it follows from the above inequality (3.3) that

$$\begin{aligned} &\sup\{\| (\Gamma x)(t) \|, t \in I\} \leq M \| x_0 \| + M(KT + L|\gamma|(I)) \\ &+ M(KT + L|\gamma|(I)) \sup\{\| x(s) \|, s \in I\}. \end{aligned}$$

Thus for $C \equiv M(KT + L|\gamma|(I))$, we arrive at the following inequality

$$\| \Gamma x \|_{B_\infty(I,X)} \leq M \| x_0 \| + C \left(1 + \| x \|_{B_\infty(I,X)}\right).$$

Hence, for every $x \in B_\infty(I,X)$, $\Gamma x \in B_\infty(I,X)$ indicates that Γ maps $B_\infty(I,X)$ into itself. Next, we prove that Γ has a unique fixed point in $B_\infty(I,X)$. From here on the proof is quite similar to that of [2, Theorem 2.1]. Using the expression (3.2) for $x, y \in B_\infty(I,X)$ and the assumptions (A),(F2) and (G2), one can easily verify that

$$\begin{aligned} &\| (\Gamma x)(t) - (\Gamma y)(t) \|_X \\ &\leq MK \int_0^t \| x(s) - y(s) \|_X \, ds + ML \int_0^t \| x(s) - y(s) \|_X \, |\gamma|(ds) \\ &= \int_0^t \| x(s) - y(s) \| \, d\beta(s), t \in I. \end{aligned} \tag{3.4}$$

where $\beta(t) \equiv \int_0^t MKds + \int_0^t ML|\gamma|(ds), t \in I$. Since $M,K,L > 0$ and $|\gamma|(\cdot)$ is a nonnegative finitely additive measure of bounded variation, it is clear that the function β is a nonnegative monotone increasing function of bounded variation. A function of bounded variation is differentiable almost every where and its differential is integrable. Thus there exists a function $g \in L_1^+(I)$ such that $g(t) > \dot{\beta}(t)$ *a.e* and $\beta_o(t) \equiv \int_0^t g(s)ds \geq \beta(t+)$ for all $t \in I$. Clearly, β_o is a continuous, nonnegative, increasing function of bounded variation on I dominating β for all $t \in I$, and $\dot{\beta}_o(t) > \dot{\beta}(t)$ for almost all $t \in I$. Hence it follows from the inequality (3.4) that

$$\| (\Gamma x)(t) - (\Gamma y)(t) \|_X \leq \int_0^t \| x(s) - y(s) \|_X \, d\beta_o(s), t \in I. \tag{3.5}$$

Define

$$\rho_t(x,y) \equiv \sup\{\| x(s) - y(s) \|_{R^n}, 0 \leq s \leq t\} \text{ for } t \in I.$$

Clearly, for $t = T$, $\rho_T(x,y) \equiv \rho(x,y) = \| x - y \|_{B_\infty(I,X)}$ defines a metric on $B_\infty(I,X)$. Hence $(B_\infty(I,X), \rho)$ is a complete metric space. Using the metric ρ_t and the inequality (3.5), one can easily verify that

$$\rho_t(\Gamma x, \Gamma y) \leq \int_0^t \rho_s(x,y) d\beta_o(s), \ \forall\, t \in I.$$

For any $m \in N$, let Γ^m denote the m-fold composition of the operator Γ, that is, $\Gamma^m = \Gamma \circ \Gamma, \cdots, \circ\Gamma$. Clearly, it follows from the above inequality that the second iterate of Γ satisfies the following inequality

$$\rho_t(\Gamma^2 x, \Gamma^2 y) \leq \int_0^t \rho_s(\Gamma x, \Gamma y) d\beta_o(ds) \leq \rho_t(x,y)(\beta_o(t))^2/2, t \in I.$$

Following this procedure, after m iterations, we arrive at the following inequality

$$\rho_t(\Gamma^m x, \Gamma^m y) \leq \rho_t(x,y)(\beta_o(t))^m/m!, \ \forall\, t \in I.$$

Then using the metric ρ, which is equivalent to the norm on the Banach space $B_\infty(I,X)$, we arrive at the following expression

$$\| \Gamma^m x - \Gamma^m y \|_{B_\infty(I,X)} \leq \alpha_m \| x - y \|_{B_\infty(I,X)}, \ m \in N,$$

where $\alpha_m = (\beta_o(T))^m/m!$. Since $\beta_o(T)$ is finite, it is clear that for $m_o \in N$ sufficiently large, $\alpha_{m_o} < 1$, and hence the operator Γ^{m_o} is a contraction. Therefore it follows from the Banach fixed point theorem that Γ^{m_o} has a unique fixed point $x^o \in B_\infty(I,X)$, that is, $\Gamma^{m_o} x^o = x^o$. Hence

$$\| \Gamma x^o - x^o \| = \| \Gamma(\Gamma^{m_o} x^o) - (\Gamma^{m_o} x^o) \| = \| \Gamma^{m_o}(\Gamma x^o) - \Gamma^{m_o}(x^o) \| < \alpha_{m_o} \| \Gamma x^o - x^o \| .$$

Since $0 < \alpha_{m_o} < 1$, this inequality holds if and only if $\Gamma x^o = x^o$. Thus x^o is the unique fixed point of the operator Γ. This proves the existence and uniqueness of solutions of equation (3.1) and hence the existence and uniqueness of a mild solution of equation (1.1). □

As a corollary of the above theorem, we have the following result.

Corollary 3.2. *Consider system (1.1) and suppose assumptions (A), (F1), and (G1) hold, and that both F and G are locally Lipschitz in the state variable, the former uniformly with respect to the set U. Then, for each initial state $x_0 \in X$ and control $\nu = (u, \gamma) \in \mathscr{K}_{ad}$, the system has a unique mild solution $x \in B_\infty(I,X)$.*

Proof. Using assumption (A) and growth properties (F1) and (G1), one can easily verify that every mild solution of equation (1.1), or equivalently the integral equation (3.1), if one exists, satisfies the following inequality

$$\| x(t) \|_X \leq C + \int_0^t \| x(s) \|_X \rho(ds), t \in I, \tag{3.6}$$

where $C \equiv M[\| x_0 \|_X + (KT + L|\gamma|(I))]$, and ρ is a positive measure defined on Σ_I given by $\rho(\sigma) = \int_\sigma MKds + \int_\sigma ML|\gamma|(ds), \sigma \in \Sigma_I$. By virtue of generalized Gronwall inequality, due to the author [3, Lemma 5, p268], it follows from the expression (3.6) that

$$\| x \|_{B_\infty(I,X)} \leq C\exp\{\rho(I)\} \leq C\exp M\{KT + L|\gamma|(I)\}.$$

Since $\mathscr{K}_{ad}$ is compact in the product topology τ_p, it is bounded, and hence there exists a positive number $r_o \in R$ such that $C\exp\{\rho(I)\} \leq r_o$ uniformly with respect to $(u,\gamma) = \nu \in \mathscr{K}_{ad}$. Let $B_{r_o} \subset X$ denote the closed ball of radius r_o centered at the origin. By assumption both F and G are locally Lipschitz and hence there exist positive constants K_{r_o}, L_{r_o} such that

$$\overline{(F2)} : \| F(t,x,\xi) - F(t,y,\xi) \|_X \leq K_{r_o} \| x - y \|_X, \text{uniformly in } \xi \in U, \forall\, x,y \in B_{r_o},$$
$$\overline{(G2)} : \| G(t,x) - G(t,y) \|_{\mathscr{L}(E,X)} \leq L_{r_o} \| x - y \|_X,\ \forall\, x,y \in B_{r_o}.$$

We define the set

$$X_o \equiv \{x \in B_\infty(I,X) : x(t) \in B_{r_o}\ \forall\, t \in I\}.$$

Clearly, this set is closed with respect to the norm topology on $B_\infty(I,X)$. Hence, using the metric ρ as seen in Theorem 3.1, one can verify that $X_o \equiv (X_o, \rho)$, equipped with relative metric topology ρ, is a complete metric space. Considering the operator Γ (as defined in Theorem 3.1) on X_o, and following the same procedure as in Theorem 3.1, one can verify that it has a unique fixed point $x^o \in X_o$. This completes the proof. □

Remark 3.3. The reader can find some interesting results in the recent book of Sofonea and Migorski [8] on the questions of existence and uniqueness of solutions of functional differential equations based on fixed point theorems for history dependent operators.

4. Existence of Optimal Controls

Throughout the rest of the paper we use the notation $\| \cdot \|$ for $\| \cdot \|_X$ unless otherwise stated. In order to prove the existence of optimal control we use the continuity of the control to solution map, $\nu \longrightarrow x(\nu)$, with respect to given topologies. Precisely we prove the following result.

Theorem 4.1. *Consider control system (1.1) and suppose that the assumptions of Theorem 3.1 hold and the semigroup $S(t), t > 0$, generated by A is compact. Suppose that $\mathscr{M}_{ad}$ satisfies the assumptions of Theorem 2.1 with μ being non-atomic, and further there exists a nonnegative finitely additive non-atomic measure μ_o uniformly dominating $\mathscr{M}_{ad}$ set wise. Let $\mathscr{K}_{ad} \equiv \mathscr{R}_{ad} \times \mathscr{M}_{ad}$ denote the set of admissible controls endowed with the product topology τ_p. Then the control to solution map $\mathscr{K}_{ad} \ni \nu \longrightarrow x(\nu) \in B_\infty(I,X)$ is continuous with respect to the product topology τ_p on $\mathscr{K}_{ad}$ and the norm topology on $B_\infty(I,X)$.*

Proof. Let $\{\nu^k\} \in \mathscr{K}_{ad}$ be a sequence of controls and $\{x^k\} \in B_\infty(I,X)$ the corresponding sequence of solutions to integral equation (3.1). Suppose $\nu^k \xrightarrow{\tau_p} \nu^o$, and let $x^o \in B_\infty(I,X)$ denote the solution corresponding to the control ν^o. We prove that $x^k \xrightarrow{s} x^o$ in $B_\infty(I,X)$. Clearly, x^k and x^o satisfy the following integral equations,

$$x^k(t) = S(t)x_0 + \int_0^t S(t-s)\hat{F}(s,x^k(s),u_s^k)ds + \int_0^t S(t-s)G(s,x^k(s))\gamma^k(ds), t \in I, \quad (4.1)$$

$$x^o(t) = S(t)x_0 + \int_0^t S(t-s)\hat{F}(s,x^o(s),u_s^o)ds + \int_0^t S(t-s)G(s,x^o(s))\gamma^o(ds), t \in I. \quad (4.2)$$

Subtracting equation (4.2) form equation (4.1) term by term and rearranging terms suitably, we obtain

$$\begin{aligned} x^k(t) - x^o(t) &= \int_0^t S(t-s)[\hat{F}(s,x^k(s),u_s^k) - \hat{F}(s,x^o(s),u_s^k)]ds \\ &+ \int_0^t S(t-s)[\hat{F}(s,x^o(s),u_s^k) - \hat{F}(s,x^o(s),u_s^o)]ds \\ &+ \int_0^t S(t-s)[G(s,x^k(s)) - G(s,x^o(s))]\gamma^k(ds) \\ &+ \int_0^t S(t-s)[G(s,x^o(s))](\gamma^k(ds) - \gamma^o(ds)), t \in I. \end{aligned} \quad (4.3)$$

Defining

$$e_1^k(t) \equiv \int_0^t S(t-s)[\hat{F}(s,x^o(s),u_s^k) - \hat{F}(s,x^o(s),u_s^o)]ds,$$

$$= \int_0^t \int_U S(t-s)F(s,x^o(s),\xi)[u_s^k(d\xi) - u_s^o(d\xi)]ds, t \in I, \quad (4.4)$$

$$e_2^k(t) \equiv \int_0^t S(t-s)G(s,x^o(s))(\gamma^k(ds) - \gamma^o(ds)), t \in I, \quad (4.5)$$

computing the norm on either side of the expression (4.3), and using the assumptions (F1) and (G1), and triangle inequality, one can easily verify that

$$\begin{aligned} \| x^k(t) - x^o(t) \| &\leq \int_0^t MK \| x^k(s) - x^o(s) \| ds \\ &+ \int_0^t ML \| x^k(s) - x^o(s) \| |\gamma^k|(ds) + \| e_1^k(t) \| + \| e_2^k(t) \| \; t \in I. \end{aligned}$$

Using the nonnegative measure μ_o, set wise dominating the admissible class $\mathscr{M}_{ad}$, in the above inequality we arrive at the following inequality

$$\begin{aligned} \| x^k(t) - x^o(t) \| &\leq \int_0^t K \| x^k(s) - x^o(s) \| ds \\ &+ \int_0^t L \| x^k(s) - x^o(s) \| \mu_o(ds) + \| e_1^k(t) \| + \| e_2^k(t) \| \; t \in I. \end{aligned}$$

Hence

$$\| x^k(t) - x^o(t) \| \leq \int_0^t \| x^k(s) - x^o(s) \| m_o(ds) + \| e_1^k(t) \| + \| e_2^k(t) \| \; t \in I,$$

where the measure m_o is given by $m_o(D) \equiv \int_D K ds + \int_D L\mu_o(ds), D \in \Sigma_I$. It follows from generalized Gronwall inequality [3, Lemma 5, p268] that

$$\| x^k(t) - x^o(t) \| \leq \varphi_k(t) + \exp\big(m_o(I)\big) \int_0^t \varphi_k(s) m_o(ds), t \in I, \quad (4.6)$$

where

$$\varphi_k(t) \equiv \parallel e_1^k(t) \parallel + \parallel e_2^k(t) \parallel, t \in I. \tag{4.7}$$

Using the growth properties of $\{F,G\}$, the bound of the semigroup $S(t), t \in I$, and the fact that the set of admissible controls $\mathscr{K}_{ad}$ is a bounded subset of $L_\infty^w(I, \mathscr{M}(U)) \times \mathscr{M}_{bfa}(\Sigma_I, E)$, one can verify that both $\{e_1^k\}$ and $\{e_2^k\}$ given by the expressions (4.4) and (4.5) are uniformly norm bounded on I. In other words $\{e_1^k, e_2^k\}$ is contained in a bounded subset of $B_\infty(I,X) \subset L_1(I,X)$. Define

$$f_k(t) \equiv \hat{F}(t, x^o(t), u_t^k) - \hat{F}(t, x^o(t), u_t^o), \quad t \in I.$$

It is clear from the growth property (F1) and the fact that $x^o \in B_\infty(I,X)$ that $f_k \in L_1(I,X)$. Since u^k converges in the weak star topology to u^o in $\mathscr{R}_{ad} \subset L_\infty^w(I, \mathscr{M}(U))$, it is clear that $f_k \xrightarrow{w} 0$ in $L_1(I,X)$. Thus it follows from compactness of the semigroup of operators $\{S(t), t > 0\}$ that, for each $t \in I$,

$$e_1^k(t) = \int_0^t S(t-s) f_k(s) ds \xrightarrow{s} 0 \text{ in } X.$$

For the second term, let us define $G_o(t) \equiv G(t, x^o(t)), t \in I$. It follows from the growth property (G1) and the fact that $x^o \in B_\infty(I,X)$, that $G_o \in B_\infty(I, \mathscr{L}(E,X))$. Since $\gamma^k \xrightarrow{w} \gamma^o$ in $\mathscr{M}_{bfa}(\Sigma_I, E)$ with the weak limit $\gamma^o \in \mathscr{M}_{ad}$, it is clear that the corresponding Radon-Nikodym derivatives $g^k \xrightarrow{w} g^o$ in $L_1(\mu, E)$. Thus the function e_2^k given by the expression (4.5) can be rewritten as follows

$$\begin{aligned} e_2^k(t) &\equiv \int_0^t S(t-s) G_o(s)[\gamma^k(ds) - \gamma^o(ds)] \\ &= \int_0^t S(t-s) G_o(s)[g^k(s) - g^o(s)]\mu(ds), t \in I. \end{aligned}$$

Hence, again it follows from compactness of the semigroup $\{S(t), t > 0\}$ and non-atomicity of the measure μ that, for all $t \in I$,

$$e_2^k(t) \equiv \int_0^t S(t-s) G_o(s)[g^k(s) - g^o(s)]\mu(ds) \xrightarrow{s} 0 \text{ in } X.$$

Thus it follows from the expression (4.7) that $\varphi_k(t) \to 0$ for each $t \in I$. Then by virtue of Lebesgue bounded convergence theorem, we conclude that the expression on the righthand side of the inequality (4.6) converges to zero for all $t \in I$. Thus we have proved that

$$\lim_{k \to \infty} \parallel x^k(t) - x^o(t) \parallel_X = 0, \; \forall \, t \in I.$$

Hence $x^k \xrightarrow{s} x^o$ in $B_\infty(I,X)$. This completes the proof. □

Remark 4.2. It would be interesting to explore the possibility of relaxing the compactness assumptions of the semigroup $S(t), t > 0$, thereby broadening the scope of applications.

Now we are in a position to consider the question of existence of optimal control.

Theorem 4.3. *Consider the system (1.1) with the cost functional (1.2) and admissible controls $\mathscr{K}_{ad}$, and suppose that the assumptions of Theorem 4.1 hold. Suppose that ℓ is Borel measurable* *in all its arguments and lower semi continuous in the state variable;, and the function Φ is also lower semi continuous in the state variable satisfying the following inequalities:*

$$(i) : |\ell(t,x)| \leq \alpha_1(t) + \alpha_2 \parallel x \parallel_X^p,\ p \in [1,\infty), \alpha_1 \in L_1^+(I) \text{ and } \alpha_2 > 0, \text{ finite} \tag{4.8}$$

$$(ii) : |\Phi(x)| \leq \alpha_3 + \alpha_4 \parallel x \parallel_X^p, \alpha_3, \alpha_4 \geq 0 \text{ finite}. \tag{4.9}$$

Then there exists an optimal control.

Proof. Since the set of admissible controls $\mathscr{K}_{ad}$ is compact in the product topology τ_p, it suffices to prove that $\nu \longrightarrow J(\nu)$ is lower semi continuous on it. Let $\{\nu^k\} = \{(u^k, \gamma^k)\}$ be a sequence in $\mathscr{K}_{ad}$, and suppose $\nu^k \xrightarrow{\tau_p} \nu^o = (u^o, \gamma^o)$. Let $\{x^k\}$ and x^o denote the mild solutions of equation (1.1) corresponding to the controls $\{\nu^k\}$ and ν^o, respectively. It follows from Theorem 4.1 that $x^k \xrightarrow{s} x^o$ in the norm topology of the Banach space $B_\infty(I,X)$. By virtue of lower semi continuity of both ℓ and Φ in the state variable, we have

$$\begin{aligned} &(i) : \ell(t, x^o(t)) \leq \underline{\lim}\, \ell(t, x^k(t)), a.e\ t \in I, \\ &(ii) : \Phi(x^o(T)) \leq \underline{\lim}\, \Phi(x^k(T)). \end{aligned} \tag{4.10}$$

Since $x^o \in B_\infty(I,X)$, it follows from the absolute growth property (4.8) that $\ell(\cdot, x^o(\cdot))$ is bounded from below by an integrable function. Thus it follows from extended Fatou's lemma that

$$\int_I \ell(t, x^o(t))dt \leq \underline{\lim}_{k\to\infty} \int_I \ell(t, x^k(t))dt. \tag{4.11}$$

Since sum of lower semi continuous functionals is lower semi continuous, it follows from (4.11) and (4.10) that

$$\int_I \ell(t, x^o(t))dt + \Phi(x^o(T)) \leq \underline{\lim}_{k\to\infty} \left\{ \int_I \ell(t, x^k(t))dt + \Phi(x^k(T)) \right\}.$$

Hence $J(\nu^o) \leq \underline{\lim}\, J(\nu^k)$ proving lower semi continuity of J in the product topology τ_p. Since $\mathscr{K}_{ad}$ is compact in this topology, J attains its minimum on it. This proves the existence of optimal control. □

Remark 4.4. The cost functional given by the expression (1.2) does not include the cost of control. A reasonable model for the cost of control is given by

$$\begin{aligned} &J_c(u) \equiv J_1(u) + J_2(\gamma) \equiv \varphi_1(< h,u >) + \parallel \gamma \parallel, \\ &\text{with } < h,u > \equiv \int_{I\times U} h(t,\xi) u_t(d\xi)dt, \\ &\textit{and } \parallel \gamma \parallel = |\gamma|(I), \text{ variation norm} \end{aligned} \tag{4.12}$$

where $h \geq 0$ and $h \in L_1(I, C(U))$. The function φ_1 is real valued continuous, nonnegative, monotone increasing in it's argument satisfying $\varphi_1(0) = 0$. One can verify that $u \longrightarrow J_1(u)$ is weak star lower semi continuous. Since the norm in any Banach space is weakly lower semicontinuous, $J_2(\gamma)$ is weakly lower semi continuous. Thus the sum

$$\tilde{J}(\nu) = J(\nu) + J_c(\nu) \tag{4.13}$$

is lower semi continuous in the product topology τ_p. Hence the system (1.1) with the cost functional $\tilde{J}$ has an optimal control.

Remark 4.5. Here we state some open problems that arise only in infinite dimensional setting. (1): For proof of continuity of the control to solution map we assumed the semigroup $\{S(t), t > 0\}$ to be compact and the measure μ non-atomic. It will be useful to relax these assumptions to broaden the scope of application. (2): Similarly, in order to admit discrete measures, it is important to drop the assumption of non-atomicity of the control measure $\mu \in \mathscr{M}^+_{bfa}(\Sigma_I)$. It is evident that according to our approach discreet measures can be admitted if $G_o \in B_\infty(I, \mathscr{K}(E, X))$, and the semigroup $\{S(t), t > 0\}$ is compact, where $\mathscr{K}(E, X)$ is the Banach space of compact operators.

5. Necessary Conditions of Optimality

In the preceding section, we have proved the existence of optimal controls. In order to determine such controls, we need necessary conditions of optimality. Here in this section, we present the necessary conditions optimality. Before we consider this, it is necessary to introduce some notations. Let $\mathscr{Z}$ be any real Banach space with $\mathscr{Z}^*$ denoting it's topological dual. The duality pairing between such spaces is denoted by $< z, z^* >_{\mathscr{Z},\mathscr{Z}^*}$ or $< z^*, z >_{\mathscr{Z}^*,\mathscr{Z}}$.

Theorem 5.1. *Consider the system* (1.1) *defined on a reflexive Banach space* X *with the cost functional given by* (1.2) *and the admissible controls* $\mathscr{K}_{ad} \equiv \mathscr{R}_{ad} \times \mathscr{M}_{ad}$ *endowed with the product topology* τ_p*, and suppose both* $\mathscr{R}_{ad}$ *and* $\mathscr{M}_{ad}$ *are convex. Suppose that the assumptions of Theorem 4.3 hold and further both* F *and* G *are continuously Gâteaux differentiable in the state variable having bounded Gâteaux derivatives, and the cost integrands* $\{\ell, \Phi\}$ *are also continuously Gâteaux differentiable in the state variable satisfying* $\ell_x(\cdot, x(\cdot)) \in L_1(I, X^*)$ *and* $\Phi_x(x(T)) \in X^*$ *for each* $x \in B_\infty(I, X)$*. Then, for* $\nu^o = (u^o, \gamma^o) \in \mathscr{K}_{ad}$ *to be an optimal control with* $x^o \in B_\infty(I, X)$ *the corresponding mild solution of the system* (1.1)*, it is necessary that there exists a function* $\psi \in B_\infty(I, X^*)$ *such that the triple* $\{\nu^o, x^o, \psi\}$ *satisfy the following inequality and the evolution equations:*

$$\begin{aligned} dJ(\nu^o; \nu - \nu^o) &= dJ((u^o, \gamma^o); (u - u^o, \gamma - \gamma^o)) \\ &= \int_0^T < \psi(t), \hat{F}(t, x^o(t); u_t - u^o_t) >_{X^*,X} dt \\ &+ < \psi(t), G(t, x^o(t))[\gamma(dt) - \gamma^o(dt)] >_{X^*,X} \geq 0, \ \forall\, \nu \in \mathscr{K}_{ad} \end{aligned} \tag{5.1}$$

where

$$\begin{aligned} &< \psi(t), \hat{F}(t, x^o(t); u_t - u^o_t) >_{X^*,X} \\ &= \int_U \langle \psi(t), F(t, x^o(t), \xi) \rangle_{X^*,X} [u_t(d\xi) - u^o_t(d\xi)], \end{aligned}$$

$$\begin{aligned} &-d\psi = A^*\psi dt + (D\hat{F})^*(t, x^o(t), u^o_t))\psi dt + \Lambda^*(t, x^o(t); \psi(t))\gamma^o(dt) + \ell_x(t, x^o(t))dt, \\ &\psi(T) = \Phi_x(x^o(T)), t \in I, \end{aligned} \tag{5.2}$$

$$dx^o(t) = Ax^o dt + \hat{F}(t, x^o(t), u^o_t)dt + G(t, x^o(t))\gamma^o(dt), x^o(0) = x_0, t \in I, \tag{5.3}$$

where the operator valued function Λ *is given by*

$$< \Lambda(t, x^o(t); x^*)x, e >_{E^*,E} = < DG^*(t, x^o(t); x)x^*, e >_{E^*,E} .$$

Proof. Under the given assumptions, it follows from Theorem 4.3 that an optimal control exists. So let $\nu^o = (u^o, \gamma^o) \in \mathscr{K}_{ad}$ denote the optimal control, and $\nu = (u, \gamma) \in \mathscr{K}_{ad}$ be any other control. For any $\varepsilon \in [0,1]$ define $\nu^\varepsilon = (u^\varepsilon, \gamma^\varepsilon) = (u^o + \varepsilon(u - u^o), \gamma^o + \varepsilon(\gamma - \gamma^o))$. Clearly, it follows from convexity of the set of admissible controls that $\nu^\varepsilon \in \mathscr{K}_{ad}$ and from the optimality of ν^o that

$$J(\nu^\varepsilon) \geq J(\nu^o), \ \forall\ \varepsilon \in [0,1] \text{ and } \nu \in \mathscr{K}_{ad}.$$

Computing the difference quotient and letting $\varepsilon \to 0$, we obtain the Gâteaux differential dJ of J at ν^o in the direction $\nu - \nu^o$, satisfying

$$dJ(\nu^o; \nu - \nu^o) \geq 0\ \forall\ \nu \in \mathscr{K}_{ad}. \tag{5.4}$$

Let $\{x^\varepsilon, x^o\} \in B_\infty(I, X)$ denote the mild solutions of the system equation (1.1) corresponding to the controls ν^ε and ν^o, respectively. Let $y \in B_\infty(I, X)$ be defined by

$$y(t) = \lim_{\varepsilon \to} (1/\varepsilon)[x^\varepsilon(t) - x^o(t)], t \in I.$$

Let us verify that y is given by the mild solution of a variational equation. Note that $\{x^\varepsilon, x^o\}$ satisfy the following integral equations,

$$x^\varepsilon(t) = S(t)x_0 + \int_0^t S(t-s)\hat{F}(s, x^\varepsilon(s), u_s^\varepsilon)ds + \int_0^t S(t-s)G(s, x^\varepsilon(s))\gamma^\varepsilon(ds), t \in I, \tag{5.5}$$

$$x^o(t) = S(t)x_0 + \int_0^t S(t-s)\hat{F}(s, x^o(s), u_s^o)ds + \int_0^t S(t-s)G(s, x^o(s))\gamma^o(ds), t \in I. \tag{5.6}$$

Subtracting equation (5.6) from equation (5.5) term by term and dividing by ε, and then letting $\varepsilon \to 0$ while using the assumptions on Gâteax differentiability of F and G in the state variable, it is easy to verify that y satisfies the following integral equation

$$\begin{aligned} y(t) = &\int_0^t S(t-s)D\hat{F}(s, x^o(s); u_s^o)y(s)ds + \int_0^t S(t-s)DG(s, x^o(s); y(s))\gamma^o(ds) \\ &+ \int_0^t S(t-s)\hat{F}(s, x^o(s); u_s - u_s^o)ds \\ &+ \int_0^t S(t-s)G(s, x^o(s))[\gamma(ds) - \gamma^o(ds)], t \in I. \end{aligned} \tag{5.7}$$

Note that, for each $s \in I$, $DG(s, x^o(s); y(s))$ denotes the Gâteaux differential of G with respect to the state variable evaluated at $x^o(s)$ in the direction $y(s)$. It follows from the assumption on its continuous Gâteaux differentiability that, for any given $(s, x) \in I \times X$, $z \longrightarrow DG(s, x; z)$ is an element of $\mathscr{L}(X, \mathscr{L}(E, X))$. In other words, for any fixed $(s, x, z) \in I \times X \times X$, $DG(s, x; z) \in \mathscr{L}(E, X)$. Clearly it follows from integral equation (5.7) that y satisfies the following measure driven linear differential equation

$$\begin{aligned} dy = &Aydt + D\hat{F}(t, x^o(t); u_t^o)ydt + DG(t, x^o(t); y)\gamma^o(dt) \\ &+ \hat{F}(t, x^o(t); u_t - u_t^o)dt + G(t, x^o(t))[\gamma(dt) - \gamma^o(dt)], y(0) = 0, t \in I. \end{aligned} \tag{5.8}$$

For the convenience of presentation, let us introduce the following vector measures

$$\mu_1(\sigma) \equiv \int_\sigma \hat{F}(t,x^o(t);u_t - u_t^o)dt, \sigma \in \Sigma_I, \text{ and}$$
$$\mu_2(\sigma) \equiv \int_\sigma G(t,x^o(t))[\gamma(dt) - \gamma^o(dt)], \sigma \in \Sigma_I.$$

Using the growth property (**F1**) of F and the facts that $x^o \in B_\infty(I,X)$ and $\{u,u^o\} \in \mathscr{M}_{ad} \subset L_\infty^w(I,\mathscr{M}_1(U))$, one can verify that $\mu_1 \in \mathscr{M}_{ca}(\Sigma_I,X)$. Since the measure $\gamma - \gamma^o$ is finitely additive having bounded variations, and G has the growth property (**G1**) and $x^o \in B_\infty(I,X)$, we conclude that $\mu_2 \in \mathscr{M}_{bfa}(\Sigma_I,X)$. Thus the differential equation (5.8) can be written compactly as follows,

$$dy = Aydt + D\hat{F}(t,x^o(t);u_t^o)ydt + DG(t,x^o(t);y)\gamma^o(dt) + \mu_1(dt) + \mu_2(dt),$$
$$y(0) = 0,\ t \in I. \tag{5.9}$$

Clearly, the sum $\mu \equiv \mu_1 + \mu_2 \in \mathscr{M}_{bfa}(\Sigma_I,X)$. By virtue of our assumptions on F and G, it is clear that for the given $\{u^o,\gamma^o,x^o\}$, $D\hat{F}$ and DG are bounded linear operator valued functions on X. More precisely, for the given $\{u^o,\gamma^o,x^o\}$, $D\hat{F} \in \mathscr{L}(X)$ and $DG \in \mathscr{L}(E,X)$. Thus (5.9) is a linear evolution equation on the Banach space X driven by the vector measure $\mu \equiv \mu_1 + \mu_2$. Hence, as a special case of Theorem 3.1, it is immediate that equation (5.9) has a unique mild solution $y \in B_\infty(I,X)$. Thus $\mu \longrightarrow y$ is a bounded linear map from the Banach space $\mathscr{M}_{bfa}(\Sigma_I,X)$ to the Banach space $B_\infty(I,X)$ and hence continuous. We denote this map by Ξ giving $y = \Xi\mu$. Considering the cost functional $J(\nu)$ given by (1.2) corresponding to the controls ν^ε and ν^o, respectively and computing its directional derivative at ν^o in the direction $\nu - \nu^o$, one can verify that

$$dJ(\nu^o;\nu-\nu^o) = \int_0^T < \ell_x(t,x^o(t)),y(t) >_{X^*,X} dt + < \Phi_x(x^o(T)),y(T) >_{X^*,X} \equiv L(y). \tag{5.10}$$

In view of $y \in B_\infty(I,X)$, and our assumption $\ell_x(\cdot,x^o(\cdot)) \in L_1(I,X^*)$ and $\Phi_x(x^o(T)) \in X^*$, it is clear that $y \longrightarrow L(y)$ is a continuous linear functional on $B_\infty(I,X)$. Thus

$$\mu \longrightarrow y \longrightarrow L(y) = L(\Xi\mu) \equiv \tilde{L}(\mu) \tag{5.11}$$

is a continuous linear functional on $\mathscr{M}_{bfa}(\Sigma_I,X)$. Hence there exists a $\psi \in (\mathscr{M}_{bfa}(\Sigma_I,X))^*$, the topological dual of the Banach space $\mathscr{M}_{bfa}(\Sigma_I,X)$ such that

$$\tilde{L}(\mu) = < \psi,\mu > = < \psi,\mu_1 + \mu_2 > . \tag{5.12}$$

The duality bracket in the above expression is between $(\mathscr{M}_{bfa}(\Sigma_I,X))^*$ and $\mathscr{M}_{bfa}(\Sigma_I,X)$. It follows from the canonical embedding of any Banach space into its bidual that $B_\infty(I,X^*) \subset (B_\infty(I,X^*))^{**} = (\mathscr{M}_{bfa}(\Sigma_I,X^{**}))^*$. Since X is a reflexive Banach space, we have $B_\infty(I,X^*) \subset (\mathscr{M}_{bfa}(\Sigma_I,X))^*$. Thus for any $\psi \in B_\infty(I,X^*)$ the duality pairing in (5.12) is also well defined. In the sequel, we prove that actually ψ belongs to this smaller space, that is, $\psi \in B_\infty(I,X^*)$. It follows from the expressions for the measures μ_1 and μ_2 that the functional (5.12) can be rewritten as follows

$$\tilde{L}(\mu) = < \psi,\mu_1 + \mu_2 > = \int_I < \psi(t),F(t,x^o(t);u_t - u_t^o) >_{X^*,X} dt$$
$$+ \int_I < \psi(t),G(t,x^o(t))[\gamma(dt) - \gamma^o(dt)] >_{X^*,X} . \tag{5.13}$$

Thus the necessary condition (5.1) follows from inequality (5.4), and equations (5.11) and (5.13). Next, computing the variation of the scalar product $< y(t), \psi(t) >=< y(t), \psi(t) >_{X,X^*}$ and integrating by parts, we obtain

$$\int_0^T d < y(t), \psi(t) >= \int_0^T < dy, \psi > + \int_0^T < y, d\psi > .$$

Since y satisfies the variational equation (5.9) with $y(0) = 0$, it follows from the above expression and elementary algebraic operations using integration by parts and adjoint operations that

$$\begin{aligned} < y(T), \psi(T) >= &\int_0^T < y(t), A^*\psi(t) + (D\hat{F})^*(t, x^o(t); u_t^o)\psi(t) > dt \\ &+ \int_0^T < \psi(t), (DG)(t, x^o(t); y(t))\gamma^o(dt) > + \int_0^T < y(t), d\psi(t) > \\ &+ \int_0^T < \psi(t), F(t, x^o(t); u_t - u_t^o) > dt + < \psi(t), G(t, x^o(t))(\gamma - \gamma^o)(dt > . \quad (5.14)\end{aligned}$$

Considering the integrand of the second term on the righthand side of the above expression, we observe that it has the form

$$< x^*, DG(t, x^o(t); x)e >_{X^*,X}, x^* \in X^*, x \in X, \text{ and } e \in E. \quad (5.15)$$

This is a trilinear form on $X^* \times X \times E$, with $x \to DG(t, x^o(t); x)$ being a bounded linear operator from X to $\mathscr{L}(E, X)$. Thus for the given x^o, $DG^*(t, x^o(t); x) \in \mathscr{L}(X^*, E^*)$ $\forall\ t \in I$. Hence the above trilinear form can be expressed as $< DG^*(t, x^o(t); x)x^*, e >_{E^*,E}$, $\forall x^* \in X^*$, $x \in X, e \in E$. At this point, by interchanging the order of operation in the above expression, we introduce a linear operator valued function $\zeta \longrightarrow \Lambda(t, x^o(t); \zeta)$ from X^* to $\mathscr{L}(X, E^*)$ given by

$$< DG^*(t, x^o(t); x)x^*, e >_{E^*,E} \equiv < \Lambda(t, x^o(t); x^*)x, e >_{E^*,E} .$$

Since, by assumption, G is continuously Gâteaux differentiable in the state variable having bounded Gâteaux derivative, $\Lambda(t, x^o(t); x^*)$ is a well defined bounded linear operator valued function with values in $\mathscr{L}(X, E^*)$. Hence it's adjoint $\Lambda^*(t, x^o(t); x^*) \in \mathscr{L}(E^{**}, X^*)$ leading to the expression

$$\begin{aligned} &< DG^*(t, x^o(t); x)x^*, e >_{E^*,E} = < \Lambda(t, x^o(t); x^*)x, e >_{E^*,E} \\ &= < x, \Lambda^*(t, x^o(t); x^*)e >_{X,X^*} . \quad (5.16)\end{aligned}$$

Clearly, it follows from the canonical embedding $E \hookrightarrow E^{**}$, that the above expression is well defined for all $e \in E$. Hence it follows from the expressions (5.15) and (5.16) that

$$< x^*, DG(t, x^o(t); x)e >_{X^*,X} = < x, \Lambda^*(t, x^o(t); x^*)e >_{X,X^*}, x^* \in X^*, x \in X, \text{ and } e \in E.$$

Using the above identity, we can rewrite the expression (5.14) as follows

$$\begin{aligned} < y(T), \psi(T) >= &\int_0^T < y(t), A^*\psi(t) + (D\hat{F})^*(t, x^o(t); u_t^o)\psi(t) > dt \\ &+ \int_0^T < y(t), \Lambda^*(t, x^o(t); \psi(t))\gamma^o(dt) > + \int_0^T < y(t), d\psi(t) > \\ &+ \int_0^T < \psi(t), F(t, x^o(t); u_t - u_t^o) > dt + < \psi(t), G(t, x^o(t))(\gamma - \gamma^o)(dt > . \quad (5.17)\end{aligned}$$

Setting

$$-d\psi = A^*\psi dt + D\hat{F}^*(t,x^o(t);u_t^o)\psi(t)dt + \Lambda^*(t,x^o(t);\psi(t))\gamma^o(dt) + \ell_x(t,x^o(t))dt, t \in I,$$
$$\psi(T) = \Phi_x(x^o(T)), \tag{5.18}$$

and using these identities in the expression (5.17), we obtain

$$< y(T), \Phi_x(x^o(T)) > + \int_0^T < y(t), \ell_x(t,x^o(t)) > dt$$
$$= \int_0^T < \psi(t), F(t,x^o(t);u_t - u_t^o) > dt + \int_0^T < \psi(t), G(t,x^o(t))(\gamma - \gamma^o)(dt > .$$

It is evident from the expressions (5.10) and (5.13) that the left hand side of the above expression coincides with $L(y)$ while the right hand expression coincides with $\tilde{L}(\mu)$. Thus this satisfies the required identity (5.11). Hence the necessary condition (5.2) follows from equation (5.18). Equation (5.3) is the given dynamic system subject to optimal control ν^o, so nothing to prove. It remains to verify that (5.2) has a unique solution $\psi \in B_\infty(I,X^*)$. This is a linear backward evolution equation on the dual space X^*. By reversing the flow of time, one can write this as a linear forward evolution equation on X^*. Then following the same procedure as seen in Theorem 3.1, one can prove that it has a unique mild solution $\psi \in B_\infty(I,X^*)$. This completes the proof of all the necessary conditions as stated in the theorem. □

Remark 5.2. Necessary conditions of the optimality with either relaxed controls only, or vector measures alone, easily follow from Theorem 5.1 by imposing appropriate restrictions.

Given the necessary conditions of optimality, it is important to develop a computational technique whereby one can determine the optimal policies. Here we present an algorithm and a related convergence theorem.

Theorem 5.3. *Consider the system* (1.1) *with the cost functional* (1.2) *and admissible controls* $\mathscr{K}_{ad}$*, and suppose that the assumptions of Theorem 4.1 hold and that E is a reflexive Banach space. Then there exists (and one can construct) a sequence of controls* $\{\nu^k\} \subset \mathscr{K}_{ad}$ *along which the cost functional* $J(\nu^k)$ *monotonically converges (possibly) to a local minimum.*

Proof. **Step 1.** Let $\nu^1 \in \mathscr{K}_{ad}$ be an arbitrary element and $x^1 \in B_\infty(I,X)$ be the corresponding solution of the system equation (5.3) with ν^o replaced by ν^1. This gives us the pair $\{\nu^1, x^1\}$.

Step 2. We use this pair in the adjoint system (5.2) replacing the pair $\{\nu^o, x^o\}$ and solve this equation giving $\psi^1 \in B_\infty(I,X^*)$ and yielding the triple $\{\nu^1, x^1, \psi^1\}$.

Step 3. We use (5.1) and replace $\{\nu^o, x^o, \psi\}$ by $\{\nu^1, x^1, \psi^1\}$. If ν^1 satisfies the inequality $dJ(\nu^1; \nu - \nu^1) \geq 0$, $\forall\ \nu \in \mathscr{K}_{ad}$, the algorithm ends and ν^1 determines the minimum. It is not necessary to verify this rare situation and can be ignored. For the next step, we rewrite the inequality (5.1) as follows:

$$\begin{aligned} dJ(\nu^1;\nu-\nu^1) &= \int_0^T <\psi^1(t),\hat{F}(t,x^1(t);u_t-u_t^1)>_{X^*,X} dt \\ &+\int_0^T <\psi^1(t),G(t,x^1(t))[\gamma(dt)-\gamma^1(dt)]>_{X^*,X} \\ &=\int_0^T\int_U \eta^1(t,\xi)[u_t(d\xi)-u_t^1(d\xi)]dt \\ &+\int_0^T <\zeta^1(t),[\gamma(dt)-\gamma^1(dt)]>_{E^*,E} . \end{aligned} \tag{5.19}$$

where

$$\eta^1(t,\xi)\equiv< F(t,x^1(t),\xi),\psi^1(t)>_{X,X^*}, t\in I,\xi\in U$$
$$\text{and } \zeta^1(t)\equiv G^*(t,x^1(t))\psi^1(t)\in E^*, t\in I.$$

Since $x^1\in B_\infty(I,X)$, $\psi^1\in B_\infty(I,X^*)$ and F is continuous in its third argument on U and satisfies the growth property (F1) independently of $\xi\in U$, it is clear that η^1 is a Borel measurable function on I with values in $C(U)$ and bounded in norm, and hence $\eta^1\in L_1(I,C(U))$. Similarly, it follows from the growth property (G1) of G that $\zeta^1\in B_\infty(I,E^*)$. For convenience of presentation, we denote $L_1(I,C(U))=\mathscr{Z}, L_\infty^w(I,\mathscr{M}(U))=\mathscr{Z}^*$ and $B_\infty(I,E^*)=Y, \mathscr{M}_{bfa}(\Sigma_I,E^{**})=\mathscr{M}_{bfa}(\Sigma_I,E)=Y^*$, the later following from the fact that E is reflexive. We introduce the two related duality maps $\mathscr{D}_1,\mathscr{D}_2$, as follows:

$$\text{for any } 0\neq z\in\mathscr{Z}, \mathscr{D}_1(z)\equiv\{z^*\in\mathscr{Z}^*:<z^*,z>=\parallel z\parallel^2=\parallel z^*\parallel^2\},$$
$$\text{for any } 0\neq y\in Y, \mathscr{D}_2(y)\equiv\{y^*\in Y^*:<y^*,y>=\parallel y\parallel^2=\parallel y^*\parallel^2\}.$$

By virtue of Hahn-Banach theorem, these duality maps are nonempty and generally multi valued. It is also known that these are weak star closed convex subsets of the respective dual spaces and that they are upper semi-continuous (multi valued) maps given that the dual spaces are equipped with the weak star topologies. With this preparation, we proceed with the next step.

Step 4. For $\varepsilon>0$ sufficiently small, define

$$u^2=u^1-\varepsilon v^1,\quad v^1\in\mathscr{D}_1(\eta^1),$$
$$\gamma^2=\gamma^1-\varepsilon\lambda^1,\quad \lambda^1\in\mathscr{D}_2(\zeta^1),$$

so that $\nu^2\equiv(u^2,\gamma^2)\in\mathscr{K}_{ad}$. Replacing $\nu\equiv\{u,\gamma\}$ by $\nu^2\equiv\{u^2,\gamma^2\}$ in the expression (5.19), we obtain

$$\begin{aligned} dJ(\nu^1;\nu^2-\nu^1) &= \int_0^T\int_U \eta^1(t,\xi)[u_t^2(d\xi)-u_t^1(d\xi)]dt \\ &+\int_0^T <\zeta^1(t),[\gamma^2(dt)-\gamma^1(dt)]>_{E^*,E} \\ &=<\eta^1,u^2-u^1>_{\mathscr{Z},\mathscr{Z}^*}+<\zeta^1,\gamma^2-\gamma^1>_{Y,Y^*} \\ &=-\varepsilon<\eta^1,v^1>_{\mathscr{Z},\mathscr{Z}^*}-\varepsilon<\zeta^1,\lambda^1>_{Y,Y^*} \\ &=-\varepsilon\{\parallel\eta^1\parallel_{\mathscr{Z}}^2+\parallel\zeta^1\parallel_Y^2\} \\ &=-\varepsilon\{\parallel v^1\parallel_{\mathscr{Z}^*}^2+\parallel\lambda^1\parallel_{Y^*}^2\}. \end{aligned} \tag{5.20}$$

By use of Lagrange formula, we can express $J(\nu^2)$ in terms of $J(\nu^1)$ and its Gâteaux differential (5.20) as follows:

$$\begin{aligned} J(\nu^2) &= J(\nu^1) + dJ(\nu^1; \nu^2 - \nu^1) + o(\varepsilon) \\ &= J(\nu^1) - \varepsilon\{\parallel \eta^1 \parallel_{\mathscr{X}}^2 + \parallel \zeta^1 \parallel_Y^2\} + o(\varepsilon). \end{aligned}$$

Hence, for $\varepsilon > 0$ sufficiently small, $J(\nu^2) < J(\nu^1)$.

Step 5. To continue, we use ν^2 and return to Step 1 to repeat the process. It is clear that following this procedure we can construct a sequence of controls $\{\nu^k\}$ along which the cost functional monotonically decreases. By virtue of Gronwall inequality, it follows from the growth properties **(F1)** and **(G1)** and the boundedness of the set of admissible controls that the set of admissible solutions is contained in a bounded subset of $B_\infty(I,X)$. Hence by virtue of the growth properties (4.8) and (4.9) of ℓ and Φ respectively, we conclude that $\{J(\nu^k)\}$ is bounded away from $-\infty$. It follows from these facts that $J(\nu^k)$ converges monotonically possibly to a local minimum $m_0 > -\infty$. In applications, the process is terminated as soon as a prescribed stopping criteria is satisfied. This completes the proof. □

A Special Class of Relaxed Controls: An important class of relaxed controls is given by switching controls similar to Bang-Bang controls. Suppose that U consists of a finite set of distinct points given by $U \equiv \{e_1, e_2, e_3, \cdots, e_m\}$. Here the controller switches from one node to another as required by optimal control policy. Evidently, this is a non-convex set. And so Pontryagin minimum principle with controls from the class of bounded measurable functions with values in U, denoted by $BM(I,U)$, does not hold. However, the necessary conditions of optimality given by Theorem 5.1 do hold. Here $\mathscr{M}_1(U)$ is a finite state probability space and the controls have the following form

$$u_t(d\xi) = \sum_{1=1}^{m} u_t(e_i)\delta_{e_i}(d\xi) \equiv \sum_{i=1}^{m} p_i(t)\delta_{e_i}(d\xi), \tag{5.21}$$

where $p_i(t) \geq 0, i = 1, 2, \cdots, m$, and $\sum_{i=1}^m p_i(t) = 1 \; \forall \, t \in I$. Let us consider the simplex

$$\mathscr{S}_m \equiv \{q_i \geq 0, i = 1, 2, 3, \cdots, m; \text{ and } \sum_{i=1}^{m} q_i = 1\}$$

with vertices given by the point masses $\{\delta_{e_i}, i = 1, 2, \cdots m\}$. We denote this set of controls by $\mathscr{C}_{ad} \subset \mathscr{R}_{ad}$. For a given set of vertices, it is clear that $\mathscr{C}_{ad}$ is isomorphic to the class of bounded measurable functions defined on I and taking values in the simplex $\mathscr{S}_m$. In particular, it is a bounded weak star closed convex subset of $L_\infty(I, R^m)$ and hence by Alaoglu's theorem [7] it is weak star compact. Thus the existence of optimal control readily follows from Theorem 4.1. With this modification, the full set of admissible controls is now given by $\mathbf{\Gamma}_{ad} \equiv \mathscr{C}_{ad} \times \mathscr{M}_{ad}$. Corresponding to this set we have the following result as a corollary of Theorem 5.1.

Corollary 5.4. *Consider the system* (1.1) *with the cost functional* (1.2) *and admissible controls* $\mathbf{\Gamma}_{ad} \equiv \mathscr{C}_{ad} \times \mathscr{M}_{ad}$ *. Suppose that the assumptions of Theorem 5.1 hold. Then, for the control* $\nu^o = (u^o, \gamma^o) \in \mathbf{\Gamma}_{ad}$ *and the corresponding solution* $x^o \in B_\infty(I,X)$ *of the system* (1.1) *to be optimal, it is necessary that there exists a* $\psi \in B_\infty(I, X^*)$ *such that the triple* $\{\nu^o, x^o, \psi\}$ *satisfy* *the following inequality and the differential equations,*

$$\begin{aligned} dJ(\nu^o;\nu-\nu^o) &= dJ((u^o,\gamma^o);(u-u^o,\gamma-\gamma^o)) \\ &= \int_0^T \sum_{i=1}^m [p_i(t)-p_i^o(t)] < \psi(t), F(t,x^o(t),e_i) >_{X^*,X} dt \\ &+ < \psi(t), G(t,x^o(t))[\gamma(dt)-\gamma^o(dt)] >_{X^*,X} \geq 0, \; \forall\, \nu \in \mathbf{\Gamma}_{ad} \end{aligned} \tag{5.22}$$

$$\begin{aligned} -d\psi &= \sum_{i=1}^m p_i^o(t) DF^*(t,x^o(t),e_i)\psi dt + \Lambda^*(t,x^o(t);\psi(t))\gamma^o(dt) + \ell_x(t,x^o(t))dt, \\ \psi(T) &= \Phi_x(x^o(T)), t \in I, \end{aligned} \tag{5.23}$$

$$dx^o(t) = \sum_{i=1}^m p_i^o(t) F(t,x^o(t),e_i))dt + G(t,x^o(t))\gamma^o(dt), x^o(0) = x_0, t \in I. \tag{5.24}$$

Proof. Here $\mathscr{M}_1(U)$ is a finite state probability space, and the controls have the form given by the expression (5.21). Let $u^o \in \mathscr{C}_{ad}$ be given by the expression $u_t^o(d\xi) = \sum_{i=1}^m p_i^o(t)\delta_{e_i}(d\xi), t \in I$, and $u \in \mathscr{C}_{ad}$ any other control given by $u_t(d\xi) = \sum_{i=1}^m p_i(t)\delta_{e_i}(d\xi), t \in I$. Let $\nu^o = \{u^o,\gamma^o\} \in \mathbf{\Gamma}_{ad}$ denote the optimal control and $\nu = \{u,\gamma\} \in \mathbf{\Gamma}_{ad}$ any other admissible control. Using this class of controls in the necessary conditions of optimality given by Theorem 5.1, it is easy to verify the inequality (5.22) and the equations (5.23) and (5.24). This completes the proof. □

Remark 5.5. Using Corollary 5.4 and following similar steps as given in the proof of Theorem 5.3, one can construct a sequence of controls from the admissible set $\mathbf{\Gamma}_{ad}$ along which the cost functional monotonically converges to a local minimum.

REFERENCES

[1] N.U. Ahmed, Semigroup Theory With Applications to Systems and Control, Pitman Research Notes in Mathematics Series, Longman Scientific and Technical, UK (Copublished with John Wiley, New York, USA) 1991.

[2] N.U. Ahmed, S. Wang, Measure driven nonlinear dynamic systems with applications to optimal impulsive controls, J. Optim. Theory Appl. 188 (2021) 26-51.

[3] N.U. Ahmed, Some remarks on the dynamics of impulsive systems in Banach spaces,Discrete Contin. Dyn. Syst. Ser. B 8 (2001) 261-274.

[4] N.U. Ahmed, Vector Measures Applied to Optimal Control for a Class of Evolution Equations on Banach Spaces, Commun. Korean Math. Soc. 35(4),1329-1352, 2020.

[5] J. Diestel, J.J. Uhl Jr. Vector Measures, American Mtehmatical Society, Providence, 1977.

[6] J. Diestel, Sequences and Series in Banach Spaces, Springer, New York, 1984.

[7] N. Dunford, J.T. Schwartz, Linear Operators, Part 1, Interscience Publishers Inc., New York, 1958,

[8] M. Sofonea, S. Migorski, Variational–Hemivariational Inequalities with Applications, Chapman & Hall/CRC, Monographs and Research Notes in Mathematics, Boca Raton, 2018.

[9] J. Warga, Optimal Control of Differential and Functional Equations, Academic Press, New York, 1972.