Nonlinear Functional Analysis and Applications
Vol. 24, No. 4 (2019), pp. 837-864
ISSN: 1229-1595(print), 2466-0973(online)

http://nfaa.kyungnam.ac.kr/journal-nfaa

A CLASS OF NONLINEAR EVOLUTION EQUATIONS ON BANACH SPACES DRIVEN BY FINITELY ADDITIVE MEASURES AND ITS OPTIMAL CONTROL

N. U. Ahmed

Departments of EECS and Mathematics
University of Ottawa, Ottawa, Ontario, Canada
e-mail: nahmed@uottawa.ca

Abstract. In this paper we consider a class of nonlinear evolution equations on infinite dimensional Banach spaces driven by finitely additive measures generalizing the classical models of impulsive systems. We use measures as controls and prove existence of optimal controls and present necessary (and sufficient) conditions of optimality. Further, we prove a convergence theorem based on the necessary conditions of optimality. Using the general results we construct the necessary conditions of optimality for purely impulsive systems. In the final section we extend our results from signed measures to finitely additive vector measures taking values in infinite dimensional Banach spaces.

1. Introduction

In a recent paper [5] Ahmed and Wang considered a class of finite dimensional nonlinear systems driven by measures and then applied to purely impulsive systems and presented necessary conditions of optimality. These results were then applied to several control problems arising from ecology and space crafts. In this paper we consider infinite dimensional systems driven by finitely additive measures covering purely impulsive systems as a special case.

In the literature, an impulsive system is popularly described by a set of evolution equations on mutually disjoint intervals of time describing, on each interval, continuous evolution of the state followed by a jump. A much larger

[0]Received June 14, 2019. Revised July 3, 2019.
[0]2010 Mathematics Subject Classification: 34G20, 34H05, 49J27, 49J55, 93C20.
[0]Keywords: Systems driven by vector measures, impulsive systems, measures as controls, existence of solutions, existence of optimal controls, necessary conditions of optimality.

class of systems governed by differential equations or inclusions on Banach spaces and driven by vector measures taking values in infinite dimensional Banach spaces were introduced by the author in several papers [2, 3, 4, 6, 7]. Optimal control of such systems were considered using either relaxed controls or controls determined by vector measures. This class of systems cover the purely impulsive systems as special cases.

In this paper we consider another large class of infinite dimensional nonlinear dynamic systems on Banach spaces driven by space-time vector measures generalizing the class of impulsive systems in the literature. Here first we use signed measures as controls and then extend it to vector measures taking values in Banach spaces. We prove existence of optimal controls and then present necessary conditions of optimality whereby one can construct the optimal controls. Also a convergence theorem based on the necessary conditions of optimality is proved ensuring convergence of the algorithm.

For finite dimensional purely impulsive systems several computational techniques have been developed by Lin, Loxton, Teo and Wu [11, 13], Ahmed and Wang [5] and successfully applied in the area of finance [10], forest ecosystem management [12], and space craft attitude control [5]. For infinite dimensional systems driven by space-time vector measures one can use our algorithm to develop similar numerical techniques to construct optimal policies.

The rest of the paper is organized as follows. In Section 2, we present the general system dynamics. First, we consider the class of purely impulsive systems obtained from the general model by choosing purely discrete measures and prove existence of solutions including regularity properties thereof. Then we introduce the basic assumptions and consider the more general class of systems driven by finitely additive measures. We prove existence and uniqueness of solutions including their regularity properties. In Section 3, we consider optimal control problems, in particular, the Bolza problem. First, we introduce the class of admissible controls and state a theorem giving the necessary and sufficient conditions characterizing weakly compact sets in the space of finitely additive (control) measures. Then we prove continuous dependence of solutions with respect to control measures (weak to strong). Using this result we prove existence of optimal controls. In Section 4, we develop the necessary conditions of optimality whereby one can determine the optimal controls. In Section 5, we present a convergence theorem based on the necessary conditions of optimality developed in the previous section. This theorem guarantees the convergence of the sequence of controls constructed on the basis of the necessary conditions of optimality. Finally, in Section 6 we discuss some natural generalization of the results presented in this paper.

2. System dynamics

In general an impulsive system can be modelled by a differential equation on a Banach space X driven by finitely or countably additive vector measures. Let U be a Polish space (for example, a complete separable metric space) and $I \equiv [0,T]$ a closed bounded interval. Let $\Sigma_{I\times U}$ denote an algebra (or a field) of subsets of the set $I \times U$. And let $B_\infty(I\times U)$ denote the space of bounded measurable real valued functions. Furnished with the supnorm topology $B_\infty(I\times U)$ is a Banach space. It is known that the continuous (topological) dual of this space is given by the space of bounded finitely additive measures defined on $\Sigma_{I\times U}$ which may be denoted by $\mathcal{M}_{bfa}(\Sigma_{(I\times U)})$.

For convenience of the reader we recall the definition of the variation norm. Let D be a $\Sigma_{I\times U}$ measurable subset of the set $I \times U$ and let Π denote any finite disjoint $\Sigma_{I\times U}$ measurable partition of the set D. The total variation of μ on D, denoted by $|\mu|(D)$, is given by

$$|\mu|(D) \equiv \sup_{\Pi} \sum_{\sigma\in\Pi} |\mu|(\sigma)$$

where the sum is taken over the elements of the finite partition Π and the supremum is taken with respect to the class of all such finite partitions. The norm of the measure μ is then given by $\parallel \mu \parallel \equiv |\mu|(I\times U)$. Endowed with the total variation norm, $\mathcal{M}_{bfa}(\Sigma_{I\times U})$ is a Banach space. A continuous linear functional ℓ on $B_\infty(I\times U)$ has the representation through an element $\mu \in \mathcal{M}_{bfa}(\Sigma_{I\times U})$ giving

$$\ell(f) = \int_{I\times U} f(t,\xi)\mu(dt\times d\xi).$$

We can use these measures to develop mathematical models for dynamic systems which exhibit impulsive behavior. In general a system governed by any differential equation on a Banach space subject to or controlled by impulsive forces can be described as follows:

$$\begin{aligned} &dx(t) = Ax(t)dt + F(t,x(t))dt + \int_U G(t,x(t),\xi)\mu(dt\times d\xi), t\in I,\\ &x(0) = x_0, \end{aligned} \tag{2.1}$$

where A is the infinitesimal generator of a C_0 semigroup $\{S(t), t\geq 0\} \subset \mathcal{L}(X)$, and the functions $F : I\times X \longrightarrow X$ and $G : I\times X\times U \longrightarrow X$ are Borel measurable maps and $\mu \in \mathcal{M}_{bfa}(\Sigma_{I\times U})$. Using the semigroup and variation of constants formula (Duhamels formula) this differential equation can be written as an integral equation on the Banach space X as follows:

$$x(t) = S(t)x_0 + \int_0^t S(t-s)F(s,x(s))ds + \int_0^t \int_U S(t-s)G(s,x(s),\xi)\mu(ds \times d\xi), \quad t \in I. \tag{2.2}$$

2.1. **Systems Driven by Discrete Measures.** In case the measure μ has the form $\mu(dt \times d\xi) = \rho(dt)m_t(d\xi)$, the integral equation (2.2) takes the form

$$x(t) = S(t)x_0 + \int_0^t S(t-s)F(s,x(s))ds + \int_0^t \rho(ds) \int_U S(t-s)G(s,x(s),\xi)m_s(d\xi), \quad t \in I. \tag{2.3}$$

Further, if μ is a discrete measure given by a weighted sum of a finite number of Dirac measures as seen below,

$$\mu(dt \times d\xi)) = \sum a_i \delta_{t_i}(dt)\delta_{v_i}(d\xi), 0 < t_1 < t_2 < \cdots < t_\kappa < T,$$

for $a_i \in R, v_i \in U, \kappa \in N$, the integral equation (2.2) reduces to

$$x(t) = S(t)x_0 + \int_0^t S(t-s)F(s,x(s))ds + \sum_{t_i \le t} S(t-t_i)a_i G(t_i, x(t_i-), v_i), t \in I. \tag{2.4}$$

It is clear that the jump size at time t_i is determined by the following vector

$$\hat{G}_i(t_i, x(t_i-)) \equiv a_i G(t_i, x(t_i-), v_i), i = 1, 2, \cdots, \kappa. \tag{2.5}$$

Using this notation, we observe that equation (2.4) can be written as

$$x(t) = S(t)x_0 + \int_0^t S(t-s)F(s,x(s))ds + \sum_{t_i \le t} S(t-t_i)\hat{G}_i(t_i, x(t_i-)), \quad t \in I. \tag{2.6}$$

Letting $I_0 \equiv \{t_i, i = 1, 2, \cdots, \kappa\}$ denote the time instants at which jump occurs and recalling that $S(0) = I_d$, it is easy to verify that system (2.6) is the integral representation of the following system of equations:

$$\dot{x}(t) = Ax(t) + F(t,x(t)), x(0) = x_0, \quad t \in I \setminus I_0; \tag{2.7}$$

$$x(t_i+) = x(t_i-) + \hat{G}_i(t_i, x(t_i-)), \quad t_i \in I_0. \tag{2.8}$$

The jump size at time t_i, given by $\hat{G}_i$, depends on the choice of $v_i \in U$ and the state of the system $x(t_i-)$ just before the jump occurs. This seems to be natural.

Clearly, the solutions of equations (2.7)-(2.8), if they exist, are piecewise continuous and bounded. So this is the dynamics where the jump times are discrete and the control $\{v_i\}$, determining the jump sizes $\hat{G}_i$ at jump times $\{t_i\}$, can be chosen as desired from the set U. In classical impulsive systems, this is considered as control variables and can be chosen so as to optimize certain performance measures.

We consider the question of existence and regularity properties of solutions of the general impulsive system given by equation (2.1). First we consider the purely impulsive system consisting of equations (2.7)-(2.8).

Theorem 2.1. *Suppose A is the infinitesimal generator of a C_0-semigroup $S(t), t \geq 0$, of operators in $\mathcal{L}(X)$, and $F : I \times X \longrightarrow X$ is Borel measurable and uniformly Lipschitz in $x \in X$ having at most linear growth and the function $G : I \times X \times U \longrightarrow X$ is continuous. Then the system* (2.7)-(2.8) *has a unique mild solution which is piecewise continuous.*

Proof. The proof is classical. Considering the interval $I_1 \equiv (0, t_1]$, it follows from Lipschitz continuity and the linear growth property of F and the semigroup $S(t), t \geq 0$, that equation (2.7) has a unique mild solution $\varphi_1 \in C([0, t_1], X)$ satisfying the integral equation

$$\varphi_1(t) = S(t)x_0 + \int_0^t S(t-s)F(s, \varphi_1(s))ds, t \in I_1 \equiv [0, t_1)$$

having the left hand limit, $\lim_{t \uparrow t_1} \varphi(t) = \varphi(t_1) \equiv x(t_1-)$. According to equation (2.8) the system makes a jump at t_1 resulting in the state

$$x(t_1+) = x(t_1-) + \hat{G}_1(t_1, x(t_1-)) = \varphi_1(t_1) + \hat{G}_1(t_1, \varphi_1(t_1)).$$

Following this jump, again the system evolves continuously according to

$$\dot{x}(t) = Ax(t) + F(t, x(t)), \ x(t_1) = x(t_1+), t \in (t_1, t_2],$$

where the initial state is given by the expression as described above. Again, by virtue of the Lipschitz and (at most) linear growth property of F, and the fact that A is the generator of the semigroup $S(t), t \geq 0$, this equation has a unique mild solution $\varphi_2 \in C((t_1, t_2], X)$ with $\varphi_2(t_1) = x(t_1+)$. Taking the limit of φ_2 from the left up to the next jump time t_2, we have $\varphi_2(t_2) \equiv x(t_2-)$. As a result of the jump governed by equation (2.8), the state takes value

$$x(t_2+) = \varphi_2(t_2) + \hat{G}_2(t_2, \varphi_2(t_2)).$$

Continuing this process one reaches the time instant t_κ where, after the jump, the state takes the value

$$x(t_\kappa+) = \varphi_\kappa(t_\kappa) + \hat{G}_\kappa(t_\kappa, \varphi_\kappa(t_\kappa)).$$

Thereafter, the system evolves according to differential equation

$$\dot{x}(t) = Ax(t) + F(t, x(t)), x(t_\kappa) = x(t_\kappa+), \;\; t \in (t_\kappa, T]$$

from the initial state as shown. It has the unique mild solution given by $\varphi_{\kappa+1} \in C((t_\kappa, T], X)$ with the terminal state given by $x(T) = \varphi_{\kappa+1}(T)$. Concatenating the pieces maintaining time order one obtains the solution trajectory $\varphi = \varphi_{\kappa+1} o \varphi_\kappa o \cdots \varphi_1$. It is clear from the construction that the solution of the system (2.7)-(2.8) is piecewise continuous and bounded. More precisely it is bounded in supnorm and continuous from the right having left hand limits. This completes the proof. □

Thus we have proved the existence of solution of equation (2.1) under the assumption that the measure is discrete having the special structure given by a sum of weighted Dirac measures

$$\mu(dt \times d\xi) \equiv \sum a_i \delta_{v_i}(d\xi)\delta_{t_i}(dt)$$

with total variation norm $\parallel \mu \parallel = \sum |a_i| < \infty$. Next we consider the general model given by equation (2.1).

2.2. **Systems Driven by General Measures.** Here we consider the general model driven by finitely additive bounded measures. Let U be a Polish space (complete separable metric space)not necessarily compact and $\mathcal{M}_{ad}(\Sigma_{I\times U}) \subset \mathcal{M}_{bfa}(\Sigma_{I\times U})$ be a nonempty bounded set denoting the set of admissible control measures. Later, we state more precise characterization of this set. Let $B_\infty(I, X)$ denote the space of bounded Borel measurable functions defined on I and taking values from the Banach space X. Endowed with the sup norm topology $\parallel z \parallel_{B_\infty(I,X)} = \sup\{\parallel z(t) \parallel_X, t \in I\}$, this is a Banach space.

We need the following basic assumptions.

- (A1) The operator A is the infinitesimal generator of a C_0-semigroup $\{S(t), t \geq 0\}$ of bounded linear operators in X.
- (A2) $F : I \times X \longrightarrow X$ is Borel measurable and there exists a constant $K_1 > 0$ such that

$$(1) : \parallel F(t, x) \parallel_X \leq K_1(1 + \parallel x \parallel_X), \;\; x \in X, \;\; t \in I,$$
$$(2) : \parallel F(t, x) - F(t, y) \parallel_X \leq K_1 \parallel x - y \parallel, \;\; x, y \in X, \;\; t \in I.$$

(A3) $G : I \times X \times U \longrightarrow X$ is Borel measurable and there exists a bounded measurable function $K_2 : U \longrightarrow R_0 \equiv [0, \infty)$ and a nonnegative bounded finitely additive nonatomic measure $\nu \in \mathcal{M}^+_{bfa}(\Sigma_I)$ such that

$$(1) : \| G(t, x, \xi)) \|_X \leq K_2(\xi)(1 + \| x \|_X),\ x \in X,\ t \in I,\ \xi \in U,$$

$$(2) : \| G(t, x, \xi) - G(t, y, \xi) \|_X \leq K_2(\xi) \| x - y \|_X,\ x, y \in X,\ t \in I,\ \xi \in U,$$

$$(3) : \int_{\Delta \times U} K_2(\xi)|\mu|(dt \times d\xi) \leq \nu(\Delta),\ \text{for each } \Delta \in \Sigma_I \text{ and for all } \mu \in \mathcal{M}_{ad}(\Sigma_{I \times U}).$$

Theorem 2.2. *Consider the evolution equation* (2.1) *with the control measure* $\mu \in \mathcal{M}_{ad}(\Sigma_{I \times U})$ *and suppose the assumptions* (A1), (A2), (A3) *hold. Then, for every* $x_0 \in X$, *the system* (2.1) *has a unique mild solution* $x \in B_\infty(I, X)$.

Proof. For proof we use Banach fixed point theorem. For any given $x_0 \in X$ and $\mu \in \mathcal{M}_{ad}(\Sigma_{I \times U})$, we use the semigroup $S(t), t \geq 0$, and define the operator Γ on $B_\infty(I, X)$ as follows,

$$\Gamma(x)(t) \equiv S(t)x_0 + \int_0^t S(t-s)F(s, x(s))ds + \int_0^t \int_U S(t-s)G(s, x(s), \xi)\mu(ds \times d\xi),\quad t \in I. \tag{2.9}$$

Under the given assumptions, we show that Γ maps $B_\infty(I, X)$ to itself. Since $S(t), t \geq 0$, is a C_0-semigroup on X and I is a finite interval there exists a finite positive number M such that $\sup\{\| S(t) \|_{\mathcal{L}(X)}, t \in I\} \leq M$. Computing the norm of $\Gamma(x)(t)$ and using the assumptions (A2) and (A3) it follows from triangle inequality that for each $t \in I$,

$$\| \Gamma(x)(t) \|_X \leq M \| x_0 \|_X + MK_1 t(1 + \sup_{0 \leq s \leq t} \| x(s) \|_X) + M(1 + \sup_{0 \leq s \leq t} \| x(s) \|_X) \int_0^t \int_U K_2(\xi)|\mu|(ds \times d\xi). \tag{2.10}$$

Using the assumption (A3), related to K_2, it follows from the above inequality that for all $t \in I$, we have

$$\| \Gamma(x)(t) \|_X \leq M \| x_0 \|_X + M\left(K_1 t + \int_0^t \nu(ds)\right)(1 + \sup_{0 \leq s \leq t} \| x(s) \|_X). \tag{2.11}$$

Since $I \equiv [0, T]$ is a finite interval it follows from the above inequality that

$$\| \Gamma(x) \|_{B_\infty(I,X)} \leq M \| x_0 \|_X + M\big(K_1 T + \nu(I)\big)\big(1 + \| x \|_{B_\infty(I,X)}\big). \tag{2.12}$$

This shows that the operator Γ maps $B_\infty(I,X)$ to $B_\infty(I,X)$. Next we verify that Γ is a contraction. Let $x, y \in B_\infty(I,X)$ satisfying $x(0) = y(0) = x_0$. Using the expression (2.9) it is easy to verify that

$$\begin{aligned} &\parallel \Gamma(x)(t) - \Gamma(y)(t) \parallel_X \\ &\leq MK_1 \int_0^t \parallel x(s) - y(s) \parallel_X ds \\ &\quad + M \int_0^t \int_U K_2(\xi) \parallel x(s) - y(s) \parallel_X |\mu|(ds \times d\xi), \quad t \in I. \end{aligned} \tag{2.13}$$

Using the assumption **(A2)** related to K_2, this can be rewritten as

$$\begin{aligned} \parallel \Gamma(x)(t) - \Gamma(y)(t) \parallel_X \; &\leq \; MK_1 \int_0^t \parallel x(s) - y(s) \parallel_X ds \\ &\quad + M \int_0^t \parallel x(s) - y(s) \parallel_X \nu(ds). \end{aligned} \tag{2.14}$$

Define the function $\beta(t) \equiv M \int_0^t K_1 ds + M \int_0^t \nu(ds), t \in I$. Since $K_1 > 0$ and $\nu \in \mathcal{M}_{bfa}^+(\Sigma_I)$, it is clear that β is a nonnegative increasing function of bounded total variation on I. Using this function, the expression (2.14) can be rewritten as

$$\parallel \Gamma(x)(t) - \Gamma(y)(t) \parallel_X \leq \int_0^t \parallel x(s) - y(s) \parallel_X d\beta(s). \tag{2.15}$$

For any pair $x, y \in B_\infty(I,X)$ and $t \in I$, define

$$\rho_t(x,y) \equiv \sup\{\parallel x(s) - y(s) \parallel_X, \;\; 0 \leq s \leq t\}$$

and note that $\rho_T(x,y) = \parallel x - y \parallel_{B_\infty(I,X)}$. Using this notation one can easily verify that the inequality (2.15) is equivalent to the following inequality.

$$\rho_t(\Gamma(x), \Gamma(y)) \; \leq \int_0^t \rho_s(x,y) \; d\beta(s), \;\; t \in I. \tag{2.16}$$

Considering the second iteration of the operator Γ (i.e $\Gamma^2 \equiv \Gamma o \Gamma$) it follows from the above expression and the fact that $t \longrightarrow \rho_t(x,y)$ is a nondecreasing function of $t \geq 0$, that, for each $t \in I$, we have

$$\begin{aligned} \rho_t(\Gamma^2(x), \Gamma^2(y)) \; &\leq \int_0^t \rho_s(\Gamma(x), \Gamma(y)) d\beta(s) \\ &\leq \int_0^t \left(\int_0^s \rho_\theta(x,y) d\beta(\theta) \right) d\beta(s). \end{aligned}$$

By assumption the measure ν is nonatomic and hence $\nu(\{0\}) = 0$ and consequently $\beta(0) = 0$. Thus it follows from the above inequality that

$$\rho_t(\Gamma^2(x), \Gamma^2(y) \leq \int_0^t \rho_s(x,y)\beta(s)d\beta(s), \quad t \in I \tag{2.17}$$

and hence we have

$$\rho_t(\Gamma^2(x), \Gamma^2(y)) \leq \rho_t(x,y)\big(\beta^2(t)/2\big), \quad t \in I. \tag{2.18}$$

Continuing this process of iteration m times we arrive at the following inequality

$$\rho_t(\Gamma^m(x), \Gamma^m(y)) \leq \rho_t(x,y)\big(\beta^m(t)/m!\big), \quad t \in I. \tag{2.19}$$

Thus, for $t = T$, we have

$$\| \Gamma^m(x) - \Gamma^m(y) \|_{B_\infty(I,R^n)} \leq \; \alpha_m \| x - y \|_{B_\infty(I,R^n)} \tag{2.20}$$

where $\alpha_m = ((\beta(T))^m/m!)$. Since $\beta(T)$ is finite, for $m \in N$ sufficiently large, $\alpha_m < 1$ and hence the m-th iterate of the operator Γ is a contraction. Thus it follows from Banach fixed point theorem that Γ^m has a unique fixed point $x^* \in B_\infty(I,X)$. Using this fact one can easily verify that x^* is also the unique fixed point of the operator Γ itself. This proves the existence of a unique mild solution of equation (2.1) in the Banach space $B_\infty(I,X)$. □

Under the assumptions of Theorem 2.2, along with an additional assumption on the set of admissible control measures $\mathcal{M}_{ad}(\Sigma_{I\times U})$, we show that the solution set is a bounded subset of the Banach space $B_\infty(I,X)$. For $\mu \in \mathcal{M}_{bfa}(\Sigma_{I\times U})$, let $x(\mu)$ denote the mild solution of the evolution equation (2.1).

Corollary 2.3. *Consider the system* (2.1) *and suppose the assumptions of Theorem 2.2 hold and that the inequality* (A3)-3 *holds uniformly with respect to the admissible set of control measures* $\mathcal{M}_{ad}(\Sigma_{I\times U})$. *Then the solution set*

$$\mathcal{S} \equiv \{x \in B_\infty(I,X) : x = x(\mu) \text{ for some } \mu \in \mathcal{M}_{ad}(\Sigma_{I\times U})\} \tag{2.21}$$

is a bounded subset of $B_\infty(I,X)$.

Proof. It follows from Theorem 2.2 that, for each $\mu \in \mathcal{M}_{ad}(\Sigma_{I\times U})$, the evolution equation (2.1) has a unique mild solution $x(\mu) \in B_\infty(I,X)$. Thus $x(\mu)$ satisfies the following integral equation

$$x(\mu)(t) = S(t)x_0 + \int_0^t S(t-s)F(s,x(\mu)(s))ds + \int_0^t \int_U S(t-s)G(s,x(\mu)(s),\xi)\mu(ds\times d\xi), \quad t \in I. \tag{2.22}$$

By taking the norm on either side and using the assumptions (A2) and (A3) it follows from triangle inequality that

$$\begin{aligned} \| x(\mu)(t) \| &\leq M\Big(\| x_0 \| + K_1 T + \int_0^T \int_U K_2(\xi)|\mu|(ds \times d\xi)\Big) \\ &\quad + \int_0^t \| x(\mu)(s) \| \, d\beta(s) \\ &\leq C + \int_0^t \| x(\mu)(s) \| \, d\beta(s), \;\; t \in I, \end{aligned} \tag{2.23}$$

for all $\mu \in \mathcal{M}_{ad}(\Sigma_{I\times U})$, where $C \equiv M\big(\| x_0 \| + K_1 T + \nu(I)\big)$. By virtue of assumption (A3)-3 guaranteeing the uniform boundedness of the admissible set of controls relating the function K_2 and the measure ν, we have $C < \infty$. Using generalized Gronwall inequality [2, Lemma 5, pp.268] applied to (2.23) and recalling that β is a nonnegative monotone increasing function of time, one can easily verify that

$$\| x(\mu)(t) \| \leq C + Ce^{\beta(t)} \int_0^t d\beta(t), \;\; t \in I, \; \forall \, \mu \in \mathcal{M}_{ad}(\Sigma_{I\times U}). \tag{2.24}$$

Hence, we have

$$\begin{aligned} &\sup\{\| x(\mu) \|_{B_\infty(I,X)}, \mu \in \mathcal{M}_{ad}(\Sigma_{I\times U})\} \\ &\leq C\big(1 + \beta(T)\exp(\beta(T))\big) < \infty. \end{aligned} \tag{2.25}$$

Thus the solution set $\mathcal{S}$ is a bounded subset of $B_\infty(I, X)$. □

3. Optimal control

To consider optimal control problems we need a more detailed characterization of admissible set of control measures. In order to consider more general cases later, we introduce a broader class of controls. Let Y denote a real Banach space, U a Polish space, and $\mathcal{M}_{bfa}(\Sigma_{I\times U}, Y)$ the space of finitely additive measures defined on a field (an algebra) $\Sigma \equiv \Sigma_{I\times U}$ of subsets of the set $I \times U$ and taking values in the Banach space Y. For admissible controls we choose a subset $\mathcal{M}_{ad}(\Sigma_{I\times U}, Y) \subset \mathcal{M}_{bfa}(\Sigma_{I\times U}, Y)$ satisfying the following conditions:

- (a1): $\mathcal{M}_{ad}$ is a bounded set: $\sup\{\| \mu \|, \mu \in \mathcal{M}_{ad}(\Sigma_{I\times U}, Y)\} < \infty$.
- (a2): There exists a finitely additive nonnegative measure $m \in M^+_{bfa}(\Sigma_{I\times U})$ such that for every $B \subset U \times I$ and $B \in \Sigma \equiv \Sigma_{I\times U}$, $\lim_{m(B)\to 0} |\mu|(B) = 0$ uniformly with respect to $\mu \in \mathcal{M}_{ad}(\Sigma_{I\times U}, Y)$.
- (a3): For every $B \in \Sigma$, the set $\{\mu(B), \mu \in \mathcal{M}_{ad}(\Sigma_{I\times U}, Y)\}$ is a relatively weakly compact subset of Y.

Theorem 3.1. (Brooks and Dinculeanu) *Suppose the Banach space Y and its dual Y^* have RNP (Radon Nikodym property) and suppose the assumptions* (a1)-(a3) *hold. Then, the set $\mathcal{M}_{ad}(\Sigma_{I\times U}, Y)$ is a (relatively) weakly compact subset of $\mathcal{M}_{bfa}(\Sigma_{I\times U}, Y)$.*

Proof. See Diestel and Uhl Jr. [8, Corollary IV.2.6, p106]. □

In the finite dimensional case, for example $Y = R^d$, $1 \leq d < \infty$, the condition (a3) is superfluous since it is implied by condition (a1).

Theorem 3.1 is a very general result on the characterization of weakly compact sets in the space of finitely additive vector measures [8, Theorem IV. 2.5, Corollary IV. 2.6, pp.106] taking values in a Banach space. The Corollary IV.2.6 is due to Brooks and Dinculeanu which generalizes a celebrated result due to Bartle-Dunford-Schwartz for countably additive vector measures [8, Theorem IV.2.5].

First, we consider the spacial case, $Y = R$, and later in section 6, we state some results where Y is an infinite dimensional Banach space.

Now we introduce the objective (cost) functional.

$$J(\mu) \equiv \int_{I\times U} \ell(t, x(\mu)(t), \xi) m(dt \times d\xi) + \Phi(x(\mu)(T)), \tag{3.1}$$

where $m \in \mathcal{M}_{ad}^{+}(\Sigma_{I\times U}) \subset \mathcal{M}_{ad}(\Sigma_{I\times U})$ and $x(\mu) \in B_\infty(I, X)$ is the mild solution of the system equation (2.1) or equivalently the associated integral equation (2.2). The objective is to find a control measure $\mu \in \mathcal{M}_{ad}(\Sigma_{I\times U})$ that minimizes the cost functional (3.1) subject to the dynamic constraint (2.1). For this we need optimality conditions and this depends on the question of existence of optimal controls.

Before we can prove the existence of optimal control we need the following important result on continuity of the control to solution map $\mu \longrightarrow x(\mu)$. This is presented in the following theorem.

Theorem 3.2. *Consider the system* (2.1) *with the operator A being the generator of a compact C_0-semigroup $S(t), t > 0$, on X, and the assumptions of Theorem* 2.2 *and Corollary* 2.3 *hold with $\nu \in \mathcal{M}_{bfa}^{+}(\Sigma_I)$ being nonatomic. Then the map $\mu \longrightarrow x(\mu)$ from $\mathcal{M}_{ad}(\Sigma_{I\times U})$ to $B_\infty(I, X)$ is continuous with respect to the relative weak topology on $\mathcal{M}_{ad}(\Sigma_{I\times U})$ and the norm topology on $B_\infty(I, X)$.*

Proof. Let $\{\mu^n, \mu^o\} \in \mathcal{M}_{ad}(\Sigma_{I\times U})$ and suppose $\mu^n \xrightarrow{w} \mu^o$. Let $x^n \equiv x(\mu^n)$ and $x^o \equiv x(\mu^o)$ denote the unique mild solutions of equation (2.1) corresponding to the same initial state, $x(\mu^n)(0) = x(\mu^o)(0) = x_0$ and driving measures μ^n

and μ^o respectively. Clearly, this means that $\{x^n, x^o\}$ satisfy the following integral equations:

$$\begin{aligned} x^n(t) &= S(t)x_0 + \int_0^t S(t-s)F(s,x^n(s))ds \\ &+ \int_0^t \int_U S(t-s)G(s,x^n(s),\xi)\mu^n(ds \times d\xi), \;\; t \in I, \end{aligned} \tag{3.2}$$

$$\begin{aligned} x^o(t) &= S(t)x_0 + \int_0^t S(t-s)F(s,x^o(s))ds \\ &+ \int_0^t \int_U S(t-s)G(s,x^o(s),\xi)\mu^o(ds \times d\xi), \;\; t \in I, \end{aligned} \tag{3.3}$$

where $\{x^n(t) \equiv x(\mu^n)(t), x^o(t) \equiv x(\mu^o)(t), t \in I\}$. Subtracting the expression (3.3) from (3.2) term by term and suitably rearranging terms we obtain the following identity

$$\begin{aligned} x^n(t) - x^o(t) &= \int_0^t S(t-s)[F(s,x^n(s)) - F(s,x^o(s))]ds \qquad (3.4) \\ &+ \int_0^t \int_U S(t-s)[G(s,x^n(s),\xi) - G(s,x^o(s),\xi)]\mu^n(ds \times d\xi) \\ &+ \int_0^t \int_U S(t-s)G(s,x^o(s),\xi)\big(\mu^n - \mu^o\big)(ds \times d\xi), \;\; t \in I. \end{aligned}$$

We denote the last term on the right hand side of the above expression by e_n giving

$$e_n(t) \equiv \int_0^t \int_U S(t-s)G(s,x^o(s),\xi)\big(\mu^n - \mu^o\big)(ds \times d\xi), \;\; t \in I. \tag{3.5}$$

Evaluating the X norm on either side of the expression (3.5) and using the basic assumptions (A2) and (A3) and triangle inequality we obtain the following inequality

$$\begin{aligned} \| x^n(t) - x^o(t) \|_X &\leq \int_0^t MK_1 \| x^n(s) - x^o(s) \|_X \, ds \\ &+ \int_0^t \int_U MK_2(\xi) \| x^n(s) - x^o(s) \|_X \, |\mu^n|(ds \times d\xi) \\ &+ \| e_n(t) \|_X, \;\; t \in I. \end{aligned} \tag{3.6}$$

Hence, using the assumption (A3) related to the function K_2 and the uniform (with respect to the admissible set $\mathcal{M}_{ad}$) dominating property of the related measure $\nu \in \mathcal{M}^+_{bfa}(\Sigma_I)$, it follows from the above inequality that

$$\| x^n(t) - x^o(t) \|_X \leq \int_0^t MK_1 \| x^n(s) - x^o(s) \|_X \, ds + M \int_0^t \| x^n(s) - x^o(s) \|_X \, \nu(ds) + \| e_n(t) \|_X, \quad t \in I. \tag{3.7}$$

Using the function β (which is a positive monotone increasing function of bounded variation) as defined immediately following the inequality (2.14), we can rewrite the inequality (3.7) as follows:

$$\| x^n(t) - x^o(t) \|_X \leq \int_0^t \| x^n(s) - x^o(s) \|_X \, d\beta(s) + \| e_n(t) \|_X, t \in I. \tag{3.8}$$

Defining $\varphi_n(t) \equiv \| x^n(t) - x^o(t) \|$, $t \in I$, again it follows from generalized Gronwall inequality [2, Lemma 5, pp.268] that

$$\begin{aligned} \varphi_n(t) &\leq \| e_n(t) \|_X + \int_0^t exp \left\{ \int_s^t d\beta(\theta) \right\} \| e_n(s) \|_X \, d\beta(s), \\ &\leq \| e_n(t) \| + e^{\beta(t)} \int_0^t \| e_n(s) \|_X \, d\beta(s), \quad t \in I. \end{aligned} \tag{3.9}$$

It suffices to show that $e_n(t)$, given by the expression (3.5), converges to zero strongly in X uniformly on I. Here we use the compactness of the semigroup $S(t), t > 0$, and the weak convergence of μ^n to μ^o. For any $\varepsilon > 0$ we can rewrite the expression (3.5) as

$$e_n(t) = e_n^{(1)}(t) + e_n^{(2)}(t), \quad t \in I$$

where

$$\begin{aligned} e_n^{(1)}(t) &\equiv S(\varepsilon) \Big(\int_0^{t-\varepsilon} \int_U S(t-\varepsilon-s) G(s, x^o(s), \xi) \big(\mu^n - \mu^o\big)(ds \times d\xi) \Big) \\ e_n^{(2)}(t) &= \int_{t-\varepsilon}^t \int_U S(t-s) G(s, x^o(s), \xi) (\mu^n - \mu^o)(ds \times d\xi), \quad t \in I. \end{aligned}$$

Referring to the first term $e_n^{(1)}$, it follows from weak convergence of μ^n to μ^o that the integral within the round bracket converges weakly to zero. Since by assumption the semigroup is compact, the operator $S(\varepsilon)$ is compact and hence the first term converges strongly to zero uniformly with respect to $t \in I$. In other words, $\lim_{n\to\infty} \sup\{\| e_n^{(1)}(t) \|_X, t \in I\} = 0$. Considering the second term $e_n^{(2)}$ and computing its norm and recalling that the assumption (A3)-3 holds uniformly with respect to the set of admissible controls $\mathcal{M}_{ad}$, we obtain the following estimate

$$\| e_n^{(2)}(t) \|_X \leq 2M \int_{t-\varepsilon}^{t} \big(1+ \| x^o(s) \|_X\big)\nu(ds), \ \ t \in I.$$

Since $x^o \in B_\infty(I, X)$ and the measure ν is nonatomic, the above integral converges to zero as $\varepsilon \downarrow 0$ uniformly on I. Thus $\| e_n \|_{B_\infty(I,X)} \xrightarrow{s} 0$, and hence it follows from Lebesgue bounded convergence theorem that the expression on the right hand side of the inequality (3.9) converges to zero uniformly with respect to $t \in I$. Hence $\varphi_n(t) \longrightarrow 0$ uniformly in $t \in I$. In other words, $x^n \longrightarrow x^o$ in the norm topology of $B_\infty(I, X)$. This proves the continuity of the map $\mu \longrightarrow x(\mu)$ in the sense as stated in the theorem. □

Now we are prepared to consider the question of existence of optimal controls. This is presented in the following theorem.

Theorem 3.3. *Consider the system* (1) *and suppose the assumptions of Theorem* 3.2 *hold and that the set of admissible control measures* $\mathcal{M}_{ad}(\Sigma_{I\times U})$ *is a weakly compact subset of* $\mathcal{M}_{bfa}(\Sigma_{I\times U})$ *and the objective functional is given by*

$$J(\mu) \equiv \int_{I\times U} \ell(t, x(t), \xi)m(dt \times d\xi) + \Phi(x(T)) \tag{3.10}$$

where $m \in \mathcal{M}_{ad}^{+}(\Sigma_{I\times U})$ *and* $x(t) \equiv x(\mu)(t)$, $t \in I$, *is the mild solution of the evolution equation* (2.1) *corresponding to the control measure* $\mu \in \mathcal{M}_{ad}$. *Suppose the functions* ℓ *and* Φ *satisfy the following assumptions:*

(1) $\ell : I \times X \times U \longrightarrow R$ *is nonnegative, Borel measurable in all the arguments, and lower semicontinuous in the second argument* $x \in X$ *uniformly with respect to* $(t, \xi) \in I \times U$, *and* m*-integrable on* $I \times U$ *uniformly with respect to* x *in bounded subsets of* X.

(2) $\Phi : X \longrightarrow R$ *is nonnegative and lower semicontinuous.*

Then, there exists an optimal control measure at which J *attains its minimum.*

Proof. Since $\mathcal{M}_{ad}(\Sigma_{I\times U})$ is weakly compact, it suffices to prove that the map $\mu \longrightarrow J(\mu)$ is weakly lower semicontinuous on $\mathcal{M}_{ad}$. Let $\mu^n \xrightarrow{w} \mu^o$ in $\mathcal{M}_{ad}(\Sigma_{I\times U})$. It follows from Theorem 3.2 that,(along a subsequence if necessary which may be relabeled as the original sequence), $x(\mu^n) \xrightarrow{s} x(\mu^o)$ in the Banach space $B_\infty(I, X)$. Thus it follows from lower semicontinuity of ℓ and Φ in $x \in X$ that

$$\ell(t, x^o(t), \xi) \leq \underline{\lim}\, \ell(t, x^n(t), \xi), \tag{3.11}$$

$$\Phi(x^o(T)) \leq \underline{\lim}\, \Phi(x^n(T)) \tag{3.12}$$

for m-almost all $(t,\xi) \in I \times U$. By Corollary 2.3, the solution set $\mathcal{S}$ is bounded and, since ℓ is m-integrable on $I \times U$ uniformly with respect to x in bounded subsets of X, both sides of the first inequality are m-integrable. Hence it follows from the inequality (3.11) that

$$\int_{I\times U} \ell(t, x^o(t), \xi)\; m(dt\times d\xi) \le \int_{I\times U} \underline{\lim}\, \ell(t, x^n(t), \xi)\; m(dt\times d\xi). \quad (3.13)$$

Using Fatou's Lemma, it follows from the above inequality that

$$\int_{I\times U} \ell(t, x^o(t), \xi)\; m(dt\times d\xi) \le \underline{\lim} \int_{I\times U} \ell(t, x^n(t), \xi)\; m(dt\times d\xi). \quad (3.14)$$

Summing (3.12) and (3.14) we conclude that $J(\mu^o) \le \underline{\lim}\, J(\mu^n)$. This proves that J is weakly lower semicontinuous on $\mathcal{M}_{ad}(\Sigma_{I\times U})$ and since $\mathcal{M}_{ad}(\Sigma_{I\times U})$ is weakly compact we conclude that there exists a $\mu^o \in \mathcal{M}_{ad}(\Sigma_{I\times U})$ at which J attains its minimum. This completes the proof. □

Remark 3.4. Since the maps $\mu \longrightarrow x(\mu) \longrightarrow J(\mu)$ are not convex, we cannot expect uniqueness of the control measure. However, using the fact that J is weakly lower semi-continuous on $\mathcal{M}_{ad}$, one can prove that the set of optimizers

$$\mathcal{O}_p \equiv \left\{ \mu \in \mathcal{M}_{ad} : J(\mu) = \inf\{J(\varrho), \varrho \in \mathcal{M}_{ad}\} \right\}$$

is a weakly closed subset of $\mathcal{M}_{ad}$ and hence a weakly compact subset of $\mathcal{M}_{ad}$.

4. Necessary conditions of optimality

In the preceding section we proved existence of optimal control policies. Here we present the necessary conditions of optimality whereby one can determine the optimal controls. For necessary conditions of optimality we need stronger regularity properties for the functions $\{F, G, \ell, \Phi\}$. For any Banach space E, let $L_1(m, E)$ denote the space of Bochner m-integrable functions on $I \times U$ with values in E. Again, for a technical reason, we have to limit the Banach space X used for the state space. It will be clear in the proof of the following result.

Theorem 4.1. *Let X be a separable reflexive Banach space. Suppose the assumptions of Theorem* 3.3 *remain in force and that the set $\mathcal{M}_{ad}(\Sigma_{I\times U})$ is also convex. Further, suppose the pair $\{F, G\}$ is once Gâteaux differentiable in the state variable with the Gâteaux derivatives being continuous and bounded, and the functions $\{\ell, \Phi\}$ appearing in the objective functional* (3.10) *are once continuously Gâteaux differentiable with respect to the state variable satisfying*

$\ell_x(\cdot, x^o(\cdot), \cdot) \in L_1(m, X^*)$ and $\Phi_x(\cdot) \in X^*$. Then, in order for the control state pair $\{\mu^o, x^o\} \in \mathcal{M}_{ad}(\Sigma_{I\times U}) \times B_\infty(I, X)$ to be optimal, it is necessary that there exists a $\psi \in B_\infty(I, X^*)$ such that the triple $\{\mu^o, x^o, \psi\}$ satisfies the following system of evolution equations, and the inequality:

$$\begin{aligned} dx^o(t) &= Ax^o dt + F(t, x^o(t))dt + \int_U G(t, x^o(t), \xi)\mu^o(dt \times d\xi), x(0) \\ & \quad x^0(0) = x_0, \end{aligned} \tag{4.1}$$

$$\begin{aligned} -d\psi(t) &= A^*\psi dt + DF^*(t, x^o(t))\psi(t)dt \\ &+ \int_U DG^*(t, x^o(t), \xi)\ \psi(t)\ \mu^o(dt \times d\xi) \\ &+ \int_U \ell_x(t, x^o(t), \xi)\ m(dt \times d\xi),\ \psi(T) = \Phi_x(x^o(T)), \end{aligned} \tag{4.2}$$

$$\int_{I\times U} < \psi(t), G(t, x^o(t), \xi) >_{X^*,X} (\mu - \mu^o)(dt \times d\xi) \geq 0,\ \forall\ \mu \in \mathcal{M}_{ad}. \tag{4.3}$$

Proof. Suppose $\mu^o \in \mathcal{M}_{ad}(\Sigma_{I\times U})$ is optimal and $\mu \in \mathcal{M}_{ad}(\Sigma_{I\times U}) \equiv \mathcal{M}_{ad}$ any other element, and $\varepsilon > 0$. By convexity of $\mathcal{M}_{ad}$, it is clear that $\mu^\varepsilon \equiv \mu^o + \varepsilon(\mu - \mu^o) \in \mathcal{M}_{ad}$ for all $\varepsilon \in [0, 1]$. Then, by optimality of μ^o, it is evident that

$$J(\mu^\varepsilon) \geq J(\mu^o)\ \forall \mu \in \mathcal{M}_{ad},\ \text{and } \varepsilon \in [0, 1].$$

Hence

$$(1/\varepsilon)(J(\mu^\varepsilon) - J(\mu^o)) \geq 0\ \forall \mu \in \mathcal{M}_{ad}\ \text{ and } \varepsilon \in (0, 1]. \tag{4.4}$$

Let $\{x^\varepsilon, x^o\} \in B_\infty(I, X)$ denote the mild solutions of the state equation (2.1) corresponding to the control measures $\{\mu^\varepsilon, \mu^o\}$ respectively. In other words, $\{x^\varepsilon, x^o\}$ satisfy the following integral equations

$$\begin{aligned} x^\varepsilon(t) = S(t)x_0 &+ \int_0^t S(t-s)F(s, x^\varepsilon(s))ds \\ &+ \int_0^t \int_U S(t-s)G(s, x^\varepsilon(s), \xi)\mu^\varepsilon(ds \times d\xi),\ \ t \in I, \end{aligned} \tag{4.5}$$

$$\begin{aligned} x^o(t) = S(t)x_0 &+ \int_0^t S(t-s)F(s, x^o(s))ds \\ &+ \int_0^t \int_U S(t-s)G(s, x^o(s), \xi)\mu^o(ds \times d\xi),\ \ t \in I. \end{aligned} \tag{4.6}$$

It is evident that $\mu^\varepsilon \xrightarrow{w} \mu^o$. In fact it follows from the construction of μ^ε that this convergence also holds in the strong sense (total variation norm).

Clearly, it follows Theorem 3.2 that $x^\varepsilon \xrightarrow{s} x^o$ in $B_\infty(I,X)$. Further, subtracting equation (4.6) from equation (4.5) term by term and computing the difference quotient $(1/\varepsilon)(x^\varepsilon(t) - x^o(t))$ and letting $\varepsilon \downarrow 0$, and denoting the limit by y, if one exists, we have

$$y(t) \equiv \lim_{\varepsilon\downarrow 0}(1/\varepsilon)(x^\varepsilon(t) - x^o(t)), \;\; t \in I. \tag{4.7}$$

One can easily verify that y satisfies the following evolution equation in the mild sense

$$\begin{aligned} dy(t) = Aydt + DF(t,x^o(t))y(t)dt + \int_U DG(t,x^o(t),\xi)y(t)\mu^o(dt\times d\xi) \\ + \int_U G(t,x^o(t),\xi)(\mu-\mu^o)(dt\times d\xi), \;\; y(0)=0, \;\; t\in I. \end{aligned} \tag{4.8}$$

This is a linear differential equation in y and can be written compactly as

$$dy = Aydt + B(t)y(t)dt + \mathbf{\Lambda}(dt)y + \vartheta_\mu(dt), \;\; y(0)=0, \;\; t\in I, \tag{4.9}$$

where

$$\begin{aligned} &B(t) \equiv DF(t,x^o(t)), \;\; t\in I, \\ &\mathbf{\Lambda}(\Delta) \equiv \int_{\Delta\times U} DG(t,x^o(t),\xi)\mu^o(dt\times d\xi), \text{ for each } \Delta\in\Sigma_I, \\ &\vartheta_\mu(\Delta) \equiv \int_{\Delta\times U} G(t,x^o(t),\xi)(\mu-\mu^o)(dt\times d\xi), \text{ for each } \;\Delta\in\Sigma_I. \end{aligned}$$

Here Σ_I denotes a field of subsets of the set I. Since, under the given assumptions, both F and G are continuously Gâteaux differentiable in the state variable with the Gâteaux derivatives being bounded on bounded sets and $x^o \in B_\infty(I,X)$, and $\mu^o,\mu \in \mathcal{M}_{ad}(\Sigma_{I\times U})$, it is clear that B is a bounded operator valued function with values in $\mathcal{L}(X)$, and $\mathbf{\Lambda}(\Delta), \Delta\in\Sigma_I$, is also a bounded operator valued measure belonging to $\mathcal{M}_{bfa}(\Sigma_I,\mathcal{L}(X))$, and $\vartheta_\mu \in M_{bfa}(\Sigma_I,X)$ is a bounded finitely additive X valued vector measure. This is a Banach space with respect to the total variation norm. Using Banach fixed point theorem, as in Theorem 2.2, one can verify that the variational equation (4.9), and hence (4.8), has a unique mild solution $y\in B_\infty(I,X)$ and the limit in (4.7) is well defined. Thus the map,

$$\vartheta_\mu \longrightarrow y \tag{4.10}$$

from $\mathcal{M}_{bfa}(\Sigma_I,X)$ to $B_\infty(I,X)$, is a continuous linear map and hence bounded. On the other hand, computing the difference quotient (4.4), letting $\varepsilon\downarrow 0$, we obtain the Gâteaux differential of J at μ^o in the direction $\mu - \mu^o$ as follows:

$$\begin{aligned} dJ(\mu^o, \mu - \mu^o) &= \lim_{\varepsilon \downarrow 0} (1/\varepsilon)(J(\mu^\varepsilon) - J(\mu^o)) \\ &= \int_{I\times U} < \ell_x(t, x^o(t), \xi), y(t) >_{X^*,X} m(dt \times d\xi) \\ &\quad + < \Phi_x(x^o(T)), y(T) >_{X^*,X} . \end{aligned} \tag{4.11}$$

By optimality of μ^o, it follows from (4.4) that

$$dJ(\mu^o, \mu - \mu^o) \geq 0, \ \forall \, \mu \in \mathcal{M}_{ad}. \tag{4.12}$$

By assumption, $\ell_x(\cdot, x^o(\cdot), \cdot) \in L_1(m, X)$ and $\Phi_x(x^o(T)) \in X^*$. Combining this with the fact that equation (4.9) has a unique solution $y \in B_\infty(I, X)$, we conclude that the functional L, given by

$$\begin{aligned} L(y) \ &\equiv \ \int_{I\times U} < \ell_x(t, x^o(t), \xi), y(t) >_{X^*,X} m(dt \times d\xi) \\ &\quad + < \Phi_x(x^o(T)), y(T) >_{X^*,X}, \end{aligned} \tag{4.13}$$

is a well defined bounded linear functional on the Banach space $B_\infty(I, X)$. Thus $y \longrightarrow L(y)$ is a continuous linear functional on $B_\infty(I, X)$ and hence it follows from (4.10) that the composition map

$$\vartheta_\mu \longrightarrow y \longrightarrow L(y) \equiv \tilde{L}(\vartheta_\mu) \tag{4.14}$$

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

$$\tilde{L}(\vartheta_\mu) =<< \psi, \vartheta >>_{\mathcal{M}^*_{bfa}(\Sigma_I,X),\mathcal{M}_{bfa}(\Sigma_I,X)} \equiv \int_I < \psi(t), \vartheta_\mu(dt) >_{X^*,X}, \tag{4.15}$$

where $\mathcal{M}^*_{bfa}(\Sigma_I, X)$ denotes the dual of the Banach space $\mathcal{M}_{bfa}(\Sigma_I, X)$. Since by our assumption X is reflexive, under the canonical embedding of a Banach space into its bidual, we have $B_\infty(I, X^*) \hookrightarrow \mathcal{M}^*_{bfa}(\Sigma_I, X)$. It is known that a reflexive Banach space X is separable if and only if X^* is separable. Thus our assumption implies that X^* is also separable. Hence, by Pettis measurability theorem as seen in Dunford and Schwartz [9, Theorem III.6.11, p149] the elements of $B_\infty(I, X^*)$ are also strongly measurable functions with values in X^*. So the duality pairing in (4.15) is also well defined for $\psi \in B_\infty(I, X^*)$. Later we show that actually ψ does belong to this smaller space. Using the expression for ϑ_μ in equation (4.15) we obtain

$$\tilde{L}(\vartheta_\mu) = \int_{I\times U} < \psi(t), G(t, x^o(t), \xi) >_{X^*,X} (\mu - \mu^o)(d\xi \times dt). \tag{4.16}$$

It follows from (4.11)-(4.14) and (4.16) that

$$\int_{I\times U} < \psi(t), G(t,x^o(t),\xi) >_{X^*,X} (\mu - \mu^o)(d\xi \times dt) \ \geq 0,\ \forall\, \mu \in \mathcal{M}_{ad}. \quad (4.17)$$

Thus we have proved the necessary condition (4.3). Next we prove the necessary condition given by (4.2). Using the variational equation (4.8) in the above expression and integrating by parts (justified later) and using Fubini's theorem one can formally derive the following identity,

$$\begin{aligned} \tilde{L}(\vartheta_\mu) \ &= \ < \psi(T), y(T) >_{X^*,X} - \int_0^T < y(t), d\psi(t) + A^*\psi(t)dt >_{X,X^*} \\ &- \int_0^T < y(t), DF^*(t,x^o(t))\psi(t) >_{X,X^*} dt \qquad (4.18) \\ &- \int_0^T < y(t), \left(\int_U DG^*(t,x^o(t),\xi)\psi(t)\mu^o(dt\times d\xi)\right) >_{X,X^*} . \end{aligned}$$

Since the identity (4.14), expressed by $\tilde{L}(\vartheta_\mu) = L(y)$, must hold it follows from the above expression that ψ must satisfy the following equations:

$$\psi(T) = \Phi_x(x^o(T)), \qquad (4.19)$$

$$d\psi + A^*\psi dt + DF^*(t,x^o(t))\psi(t)dt + \int_U DG^*(t,x^o(t),\xi)\psi(t)\mu^o(dt\times d\xi)$$

$$= - \int_U \ell_x(t,x^o(t),\xi)m(dt\times d\xi), \ \ t \in I. \qquad (4.20)$$

This is precisely the necessary condition given by equation (4.2). Equation (4.1) is the given dynamic system with x^o being the solution corresponding to the optimal control measure μ^o and hence nothing to prove. To complete the proof, it remains to show that the adjoint variable ψ, whose existence was asserted by the duality pairing (see equation (4.15)), is actually given by the mild solution of the evolution equation (4.20). Equation (4.20), or equivalently (4.2), is a linear evolution equation on the Banach space X^* with the terminal condition (4.19) (instead of initial condition) and called the adjoint evolution equation. This equation can be written in the compact form as follows,

$$\begin{aligned} &-d\psi = A^*\psi(t)dt + B^*(t)\psi(t)dt + \Lambda^*(dt)\psi(t) + \gamma_m(dt), \\ &\psi(T) = \Phi_x(x^o(T)) \equiv \psi_o(T), \ \ t \in I, \qquad (4.21) \end{aligned}$$

where A^* is the conjugate (adjoint) of the semigroup generator A, $B^*(t)$ is the adjoint of the bounded operator valued function $B(t) \in \mathcal{L}(X)$ and Λ^* is the adjoint of the operator valued measure $\Lambda \in \mathcal{M}_{bfa}(\Sigma_I, \mathcal{L}(X))$, all defined immediately following equation (4.9). The set function $\gamma_m(\cdot)$ is given by

$$\gamma_m(\sigma) = \int_{\sigma\times U} \ell_x(t,x^o(t),\xi)m(dt\times d\xi), \ \text{for each } \sigma \in \Sigma_I.$$

Since by our assumption $\ell_x(\cdot, x^o(\cdot), \cdot) \in L_1(m, X^*)$, this is a bounded finitely additive X^* valued vector measure. The mild solution of the adjoint evolution equation (4.21) (if one exists) is given by the solution of the following backward integral equation on the dual space X^*,

$$\begin{aligned} \psi(t) &= S^*(T-t)\Phi_x(x^o(T)) + \int_t^T S^*(s-t)B^*(s)\psi(s)ds \\ &\quad + \int_t^T S^*(s-t)\Lambda^*(ds)\psi(s) + \int_t^T S^*(s-t)\gamma_m(ds), \ \ t \in I. \end{aligned} \tag{4.22}$$

We prove that this equation has a unique solution ψ in the smaller space $B_\infty(I, X^*)$ which is a subset of $\mathcal{M}^*_{bfa}(\Sigma_I, X)$ even though the argument based on duality pairing says that $\psi \in \mathcal{M}^*_{bfa}(\Sigma_I, X)$. Before we consider the question of existence, we show that if ψ is any solution of the integral equation (4.22) it must necessarily belong to $B_\infty(I, X^*)$. First we note that, since X is a reflexive Banach space, the adjoint semigroup $S^*(t), t \geq 0$, is also a C_0-semigroup [1, Theorem 2.4.4, p51] and belongs to $\mathcal{L}(X^*)$ and for any finite interval $I \equiv [0, T]$ there exists a finite positive number M such that $\sup\{\| S(t) \|_{\mathcal{L}(X)} = \| S^*(t) \|_{\mathcal{L}(X^*)}, t \in I\} \leq M$. For convenience of notation define

$$\eta(t) \equiv S^*(t)\Phi_x(x^o(T)) + \int_t^T S^*(s-t)\gamma_m(ds), \ \ t \in I.$$

Taking the X^* norm of η and recalling that $\ell^o_x(\cdot,\cdot) \equiv \ell_x(\cdot, x^o(\cdot), \cdot) \in L_1(m, X^*)$, it is clear that

$$\sup\{\| \eta(t) \|_{X^*}, t \in I\} \leq M\{\| \Phi_x(x^o(T)) \|_{X^*} + \| \ell^o_x \|_{L_1(m,X^*)}\} \equiv C < \infty.$$

Thus $\eta \in B_\infty(I, X^*)$. By assumption both F and G are continuously Gâteaux differentiable with the G-derivatives being bounded. Thus taking X^* norm of ψ, it follows from (4.22) that

$$\begin{aligned} \| \psi(t) \|_{X^*} &\leq C + MK_1 \int_t^T \| \psi(s) \|_{X^*} \, ds \\ &\quad + M \int_t^T \| \psi(s) \|_{X^*} \left(\int_U K_2(\xi) |\mu^o|(ds \times d\xi \right), \ \ t \in I. \end{aligned} \tag{4.23}$$

Recalling the definition of β, a nonnegative monotone increasing function of bounded total variation as seen in equations (2.14)-(2.15), it follows from (4.23) that

$$\| \psi(t) \|_{X^*} \leq C + \int_t^T \| \psi(s) \|_{X^*} \, d\beta(s), \ \ t \in I. \tag{4.24}$$

Hence it follows from generalized Gronwall inequality that

$$\| \psi \|_{B_\infty(I,X^*)} \equiv \sup\{\| \psi(t) \|_{X^*}, t \in I\} \leq C \exp(\beta(T)).$$

Thus, if the integral equation (4.22) has a solution, it must belong to $B_\infty(I, X^*)$.

Now we prove that equation (4.22) has a unique solution in $B_\infty(I, X^*)$. In view of equation (4.22), let us introduce the operator Ξ as follows

$$\begin{aligned}(\Xi\psi)(t) \quad &\equiv \quad \eta(t) + \int_t^T S^*(s-t)B^*(s)\psi(s)ds \\ &\quad + \int_t^T S^*(s-t)\Lambda^*(ds)\psi(s), \quad t \in I. \end{aligned} \tag{4.25}$$

Since $\eta \in B_\infty(I, X^*)$ one can readily verify that $\Xi : B_\infty(I, X^*) \longrightarrow B_\infty(I, X^*)$. We prove that this operator has a unique fixed point in $B_\infty(I, X^*)$. Let $\psi_1, \psi_2 \in B_\infty(I, X^*)$. Clearly

$$\begin{aligned}(\Xi\psi_1)(t) - (\Xi\psi_2)(t) \quad &= \quad \int_t^T S^*(s-t)B^*(s)[\psi_1(s) - \psi_2(s)]ds \qquad (4.26) \\ &\quad + \int_t^T S^*(s-t)\Lambda^*(ds)[\psi_1(s) - \psi_2(s)], \quad t \in I.\end{aligned}$$

Evaluating the X^* norm of the difference $(\Xi\psi_1)(t) - (\Xi\psi_2)(t)$ and using the assumptions (A2)-(A3) we obtain the following inequality

$$\| (\Xi\psi_1)(t) - (\Xi\psi_2)(t) \|_{X^*} \leq \int_t^T \| \psi_1(s) - \psi_2(s) \|_{X^*} \alpha(ds) \tag{4.27}$$

where $\alpha(ds) \equiv [MK_1 ds + M\nu(ds)]$. Since by assumption the measure ν is nonatomic, it is clear that the measure α is nonatomic. We partition the interval $I \equiv [0, T]$ into n disjoint intervals of equal length $\Delta = T/n$ giving $I \equiv [0, T] = \left(\bigcup_{k=1}^n D_{k\Delta}\right) \bigcup \{T\}$ where $D_{k\Delta}$ is given by the (left closed right open) interval $D_{k\Delta} \equiv [T - k\Delta, T - (k-1)\Delta)$. Since α is nonatomic, we can choose n large enough so that $\alpha(D_{k\Delta}) < 1$ for each $k = 1, 2, \cdots, n$. Now, considering the restriction of the operator Ξ on to the interval $\overline{D}_\Delta$, it follows from the expression (4.27) that

$$\| (\Xi\psi_1) - (\Xi\psi_2) \|_{B_\infty(\overline{D}_\Delta, X^*)} \leq \alpha(D_\Delta) \| \psi_1 - \psi_2 \|_{B_\infty(\overline{D}_\Delta, X^*)} .$$

Thus the operator Ξ, restricted to each of the sequence of Banach spaces $B_\infty(\overline{D}_{k\Delta}, X^*)$ $k = 1, 2 \cdots n$, is a contraction and hence it follows from Banach fixed point theorem that the operator Ξ has a unique fixed point in $B_\infty(I, X^*)$.

Hence, the integral equation (4.22) has a unique solution $\psi \in B_\infty(I, X^*)$. It remains to justify the integration by parts used in the expression (4.18). For this, we use the Yosida approximation [1], $A_\lambda \equiv \lambda AR(\lambda, A)$, of the unbounded operator A with $\lambda \in \rho(A)$, the resolvent set of A. This is a family $\{A_\lambda, \lambda \in$

$\rho(A)\}$ of bounded linear operators belonging to $\mathcal{L}(X)$ with the corresponding semigroups $\{S_\lambda(t), t \geq 0\} \subset \mathcal{L}(X)$. As $\lambda \to \infty$, the (generalized) sequence $\{A_\lambda\}$ converges to the unbounded operator A in the strong operator topology on $D(A)$. Using this operator $A_\lambda \in \mathcal{L}(X)$ (in place of the unbounded operator A) in the derivation, one arrives at the necessary conditions of optimality (4.1)-(4.3) as stated in the theorem with x^o replaced by x^o_λ, ψ replaced by ψ_λ and μ^o replaced by μ^o_λ for $\lambda \in \rho(A)$. Then using the fact that the corresponding semigroups $\{S_\lambda(t), t \geq 0\}$, as well as its adjoint counterparts $\{S^*_\lambda(t), t \geq 0\}$, (since X is reflexive) converge in the strong operator topology to $S(t), t \geq 0$, and $S^*(t), t \geq 0$, respectively uniformly on bounded intervals [Ahmed,1, Remark 2.2.9, Corollary 2.2.10, Lemma 2.3.1], one arrives at the conclusion that as $\rho(A) \ni \lambda \to \infty$, $\mu^o_\lambda \xrightarrow{w} \mu^o$ in $\mathcal{M}_{ad}$, $x^o_\lambda \xrightarrow{s} x^o$ in $B_\infty(I, X)$ and $\psi_\lambda \xrightarrow{s} \psi$ in $B_\infty(I, X^*)$. Using these results and letting $\lambda \to \infty$ one arrives at the necessary conditions of optimality as stated in the theorem. This completes the proof. □

Remark 4.2 (Sufficient Condition) The necessary conditions of optimality (4.1)-(4.3) given by Theorem 4.1 are also sufficient. Let $\mu^o \in \mathcal{M}_{ad}$ satisfy the necessary conditions with $\{x^o, \psi^o\}$ being the corresponding mild solutions of the evolution equations (4.1) and (4.2) respectively. Using Lagrange formula one can verify that for any $\mu \in \mathcal{M}_{ad}$,

$$J(\mu) = J(\mu^o) + dJ(\mu^o, \mu - \mu^o) + o(\| \mu - \mu^o \|),$$

where $dJ(\mu^o, \mu-\mu^o)$ denotes the Gâteaux differential of J at μ^o in the direction $(\mu - \mu^o)$ and $\| \mu - \mu^o \|$ denotes the total variation norm. In the course of the proof of the necessary conditions of optimality, we have seen that dJ is given by

$$dJ(\mu^o, \mu - \mu^o) = \int_{I\times U} < \psi^o(t), G(t, x^o(t), \xi) >_{X^*,X} (\mu - \mu^o)(dt \times d\xi).$$

By virtue of the necessary condition (4.3) we have

$$\int_{I\times U} < \psi^o(t), G(t, x^o(t), \xi) >_{X^*,X} (\mu - \mu^o)(dt \times d\xi) \geq 0, \ \forall \mu \in \mathcal{M}_{ad}.$$

Thus $J(\mu) \geq J(\mu^o)$ for all $\mu \in \mathcal{M}_{ad}$ and hence μ^o is optimal.

For purely impulsive systems given by equations (2.7)-(2.8) with the cost functional given by (3.1) one can easily derive the necessary conditions of optimality from Theorem 4.1 as a corollary. Here the set of admissible controls is given by a family of discrete measures. Let $\mathcal{A}$ be a closed bounded subset of the real line and V a compact subset of the Polish space U. For the set of admissible controls, we choose the following family of discrete signed measures:

$$\mathcal{M}_\delta \equiv \left\{ \mu \in \mathcal{M}_{bfa}(\Sigma_{I\times U}) : \mu(dt \times d\xi) = \sum_{i=1}^{\kappa} a_i \delta_{t_i}(dt)\delta_{v_i}(d\xi) : a_i \in \mathcal{A}, v_i \in V \right\},$$

where $t_i \in I_0 \equiv \{t_i, i = 1, 2, \cdots, \kappa, 0 < t_1 < t_2 \cdots t_\kappa < T\}$. It is clear that the total variation norm of any element in this family is given by $\| \mu \| = \sum_{i=1}^{\kappa} |a_i|$. In the case of general system model considered in section 6, $a_i \in \mathcal{A}$, where $\mathcal{A}$ is a weakly compact and convex subset of the Banach space Y. The necessary conditions of optimality corresponding to the admissible set $\mathcal{M}_\delta$ is given by the following corollary.

Corollary 4.2. *Consider the system* (2.7)-(2.8) *with the cost functional* (3.1). *In order that a control measure* $\mu^o(\equiv \sum a_i^o \delta_{t_i}(dt)\delta_{v_i^o}(d\xi)) \in \mathcal{M}_\delta$ *and the corresponding solution* $x^o \in B_\infty(I, X)$ *(of equations* (2.7)-(2.8)*) be optimal, it is necessary that there exists a* $\psi^o \in B_\infty(I, X^*)$ *such that the triple* $\{\mu^o, x^o, \psi^o\}$ *satisfies the following equations and inequalities:*

$$dx^o = Ax^o dt + F(t, x^o(t))dt, x(0) = x_0, t \in I \setminus I_0, \tag{4.28}$$

$$x^o(t_i+) = x^o(t_i-) + a_i^o G(t_i, x^o(t_i-), v_i^o), \quad i = 1, 2, \cdots, \kappa, \tag{4.29}$$

$$\begin{aligned} -d\psi^o \;=\; & A^*\psi^o dt + DF^*(t, x^o(t))\psi^o(t)dt \\ & + \int_U \ell_x(t, x^o(t), \xi)m(dt \times d\xi), \quad \psi^o(T) = \Phi_x(x^o(T)), \;\; t \in I \setminus I_0, \end{aligned} \tag{4.30}$$

$$\psi^o(t_i-) = \psi^o(t_i+) + a_i^o DG^*(t_i, x^o(t_i-), v_i^o)\psi^o(t_i+), \quad i = 1, 2, \cdots, \kappa, \tag{4.31}$$

$$\begin{aligned} & \sum_i a_i < \psi^o(t_i-), G(t_i, x^o(t_i-), v_i) >_{X^*,X} \\ & \geq \sum_i a_i^o < \psi^o(t_i-), G(t_i, x^o(t_i-), v_i^o) >_{X^*,X}, \;\; \forall\, (a_i, v_i) \in \mathcal{A} \times V. \end{aligned} \tag{4.32}$$

Proof. Proof readily follows from Theorem 4.1 by choosing $\mathcal{M}_\delta$ as the set of admissible controls. □

Finite dimensional version of the necessary conditions of optimality given by Corollary 4.3 has been applied to attitude control of geosynchronous Satellites and Population dynamics (prey-predator model) [5].

5. A convergence theorem

For convenience of notation, let us set $\mathcal{Z} \equiv B_\infty(I \times U)$ and $\mathcal{Z}^* \equiv (B_\infty(I \times U))^* = \mathcal{M}_{bfa}(\Sigma_{I\times U})$. Define the duality map $D : \mathcal{Z} \setminus \{0\} \longrightarrow \mathcal{Z}^*$ by

$$D(\eta) \equiv \{\mu \in \mathcal{Z}^* :< \mu, \eta >_{\mathcal{Z}^*,\mathcal{Z}} = \| \eta \|^2_{\mathcal{Z}} = \| \mu \|^2_{\mathcal{Z}^*}\}, \ \ \eta \in \mathcal{Z}.$$

By virtue of Hahn-Banach theorem, the duality set $D(\eta) \neq \emptyset$. In general, this is a multi valued map and for each $\eta(\neq 0) \in \mathcal{Z}$, $D(\eta)$ is a weak star closed convex subset of $\mathcal{Z}^*$ and it is also demicontinuous. It is single valued only if the unit ball of $\mathcal{Z}^*$ is strictly convex. In our particular case, the unit ball of $\mathcal{Z}^*$ is not strictly convex.

Now we can introduce an algorithm based on the necessary conditions of optimality given by Theorem 4.1.

Step 1: Choose an arbitrary $\mu^1 \in \mathcal{M}_{ad}(\Sigma_{I\times U}) \subset \mathcal{M}_{bfa}(\Sigma_{I\times U})$ and solve the state equation

$$dx = Axdt + F(t,x)dt + \int_U G(t,x,\xi)\mu^1(dt \times d\xi), \ \ x(0) = x_0, \ \ t \in I, \quad (5.1)$$

(in the mild sense) giving the solution $x^1 \in B_\infty(I, X)$.

Step 2: Using the pair $\{\mu^1, x^1\}$, solve the following adjoint equation,

$$\begin{aligned} -d\psi &= A^*\psi dt + DF^*(t, x^1(t))\psi(t)dt + \int_U DG^*(t, x^1(t), \xi)\psi(t)\mu^1(dt \times d\xi) \\ &\quad + \int_U \ell_x(t, x^1(t), \xi)\mu^1(dt \times d\xi), \psi(T) = \Phi_x(x^1(T)), \ \ t \in I, \qquad (5.2) \end{aligned}$$

(in the mild sense) giving $\psi^1 \in B_\infty(I, X^*)$. Now we have the triple $\{\mu^1, x^1, \psi^1\}$.

Step 3: Use the triple $\{\mu^1, x^1, \psi^1\}$ in the necessary condition of optimality given by equation (4.3) yielding the following inequality,

$$\int_{I\times U} [< \psi^1(t), G(t, x^1(t), \xi) >_{X^*,X} (\mu - \mu^1)(d\xi \times dt) \geq 0, \ \forall \mu \in \mathcal{M}^+_{ad}. \quad (5.3)$$

If the inequality is satisfied, the measure μ^1 is the optimal control. If NOT go to Step 4.

Step 4: Define the function,

$$\eta^1(t, \xi) \equiv < \psi^1(t), G(t, x^1(t), \xi) >_{X^*,X}, (t, \xi) \in I \times U,$$

which clearly belongs to $B_\infty(I \times U)$. Choose any element $m^1 \in D(\eta^1)$. Then choose an $\varepsilon > 0$ sufficiently small so that $\mu^2 \equiv \mu^1 - \varepsilon m^1 \in \mathcal{M}_{ad}(\Sigma_{I\times U})$. Then, using Lagrange formula and computing $J(\mu^2)$ we obtain

$$\begin{aligned} J(\mu^2) &= J(\mu^1) + dJ(\mu^1;\mu^2-\mu^1) + o(\varepsilon) \\ &= J(\mu^1) + <\eta^1, \mu^2-\mu^1>_{\mathcal{Z},\mathcal{Z}^*} + o(\varepsilon) \\ &= J(\mu^1) - \varepsilon <\eta^1, m^1>_{\mathcal{Z},\mathcal{Z}^*} + o(\varepsilon) \\ &= J(\mu^1) - \varepsilon \parallel \eta^1 \parallel_{\mathcal{Z}}^2 + o(\varepsilon) = J(\mu^1) - \varepsilon \parallel m^1 \parallel_{\mathcal{Z}^*}^2 + o(\varepsilon). \end{aligned} \tag{5.4}$$

Hence, for $\varepsilon > 0$ sufficiently small, we have $J(\mu^2) < J(\mu^1)$. This process is repeated by returning to Step 1 with μ^2 replacing μ^1. This is continued until a prescribed stopping criterion is satisfied. Thus we have proved the following result.

Theorem 5.1. (Convergence Theorem) *Suppose the Necessary conditions of optimality given by Theorem 4.1 hold. Then there exists a sequence* $\{\mu^n\} \in \mathcal{M}_{ad}(\Sigma_{I\times U}) \subset \mathcal{M}_{bfa}(\Sigma_{I\times U})$ *along which the cost functional* J *converges monotonically to its minimum.*

Proof. Following the steps as presented above, we can construct a sequence of control measures $\{\mu^n\} \in \mathcal{M}_{ad}(\Sigma_{I\times U})$ such that $J(\mu^1) > J(\mu^2) > J(\mu^3) > \cdots J(\mu^n) \cdots$. This is a monotone decreasing sequence. Since, by our assumption, ℓ and Φ are nonnegative, $J(\mu) \geq 0$ for all $\mu \in \mathcal{M}_{ad}(\Sigma_{I\times U})$. Thus there exists a nonnegative real number m_o such that $\lim_{n\to\infty} J(\mu^n) \longrightarrow m_o$. This completes the proof. □

6. Extensions to more general system models

(E1): The results presented in the preceding sections can be easily extended to a more general class of systems driven by vector measures as stated below,

$$dx = Axdt + F(t,x(t))dt + \int_U \mathcal{G}(t,x(t),\xi)\mu(dt\times d\xi), \quad x(0) = x_0, \tag{6.1}$$

where A and F are the same as for system (2.1), and $\mathcal{G} : I \times X \times U \longrightarrow \mathcal{L}(Y,X)$ and $\mu \in \mathcal{M}_{bfa}(\Sigma_{I\times U}, Y)$ the space of finitely additive Y valued vector measures. The assumption $(A3)$ is modified as follows:

$\overline{\textbf{(A3)}}$: $\mathcal{G} : I \times X \times U \longrightarrow \mathcal{L}(Y,X)$ is Borel measurable in the uniform operator topology on $\mathcal{L}(Y,X)$, and there exists a bounded measurable function $K_2 : U \longrightarrow R_0 \equiv [0,\infty)$ and a nonnegative bounded finitely additive nonatomic measure $\nu \in \mathcal{M}_{bfa}^+(\Sigma_I)$ such that for all $x, y \in X, t \in I$ and $\xi \in U$

(1) $\parallel G(t,x,\xi)) \parallel_{\mathcal{L}(Y,X)} \leq K_2(\xi)(1+ \parallel x \parallel_X)$,

(2) $\parallel G(t,x,\xi) - G(t,y,\xi) \parallel_{\mathcal{L}(Y,X)} \leq K_2(\xi) \parallel x-y \parallel_X$,

(3) $\int_{\Delta\times U} K_2(\xi)|\mu|(dt\times d\xi) \leq \nu(\Delta), \quad \Delta \in \Sigma_I, \quad \mu \in \mathcal{M}_{ad}(\Sigma_{I\times U}, Y)$.

In the preceding sections we used scalar valued finitely additive measures as controls. In this section we show that all our results remain valid for vector valued finitely additive measures as controls. Here the set of admissible controls is given by the set

$$\mathcal{M}_{ad}(\Sigma_{I\times U}, Y) \subset \mathcal{M}_{bfa}(\Sigma_{I\times U}, Y)$$

which is assumed to satisfy the compactness criterion of Brooks and Dinculeanu as stated in Theorem 3.1.

(G1) Under the assumptions (A1),(A2), $\overline{(A3)}$, Theorem 2.2 and Corollary 2.3 remain valid. The proof is identical with minor changes in the wordings.

(G2) Under the assumptions (A1),(A2), $\overline{(A3)}$, with $\mathcal{M}_{ad}(\Sigma_{I\times U})$ replaced by $\mathcal{M}_{ad}(\Sigma_{I\times U}, Y)$, Theorem 3.2 and Theorem 3.3 remain valid.

(G3) Under the assumptions (A1),(A2), $\overline{(A3)}$, and convexity assumption for the set of admissible controls, $\mathcal{M}_{ad}(\Sigma_{I\times U}, Y)$, the necessary conditions of optimality (Theorem 4.1) remain valid with slight changes in notation as follows:

$$\begin{aligned} dx^o(t) &= Ax^o dt + F(t, x^o(t))dt + \int_U \mathcal{G}(t, x^o(t), \xi)\mu^o(dt \times d\xi), \\ &\quad x(0) = x_0, \end{aligned} \tag{6.2}$$

$$\begin{aligned} -d\psi(t) &= A^*\psi dt + DF^*(t, x^o(t))\psi(t)dt \\ &\quad + \int_U D\mathcal{G}^*(t, x^o(t), \xi; \psi(t))\ \mu^o(dt \times d\xi) \\ &\quad + \int_U \ell_x(t, x^o(t), \xi)\ m(dt \times d\xi),\ \ \psi(T) = \Phi_x(x^o(T)), \end{aligned} \tag{6.3}$$

$$h(\mu) \equiv \int_{I\times U} < \mathcal{G}^*(t, x^o(t), \xi)\psi(t), (\mu - \mu^o)(dt \times d\xi) >_{Y^*,Y} \geq 0 \tag{6.4}$$

for all $\mu \in \mathcal{M}_{ad}$, where $X^* \ni x^* \longrightarrow D\mathcal{G}^*(t, x^o(t), \xi; x^*)$ is linear. In other words, $D\mathcal{G}^*(t, x^o(t), \xi; \cdot) \in \mathcal{L}(X^*, \mathcal{L}(X^*, Y^*))$.

The duality pairing in the expression (6.4) is better understood as

$$h(\mu) \equiv << \mathcal{G}^*\psi, \mu - \mu^o >>_{B_\infty(I\times U, Y^*), \mathcal{M}_{bfa}(\Sigma_{I\times U}, Y^{**})}\ .$$

Since $\mu - \mu^o \in \mathcal{M}_{bfa}(\Sigma_{I\times U}, Y) \subset \mathcal{M}_{bfa}(\Sigma_{I\times U}, Y^{**})$, the pairing is well defined. If Y is also assumed to be reflexive, it is clear that Y^* is also reflexive. It is known [8] that reflexive Banach spaces have RNP (Radon-Nikodym property). Thus, Theorem 3.1 holds for the pair $\{Y, Y^*\}$ and hence the pairing in the expression (6.4) holds for the pair $B_\infty(I \times U, Y^*)$ and $\mathcal{M}_{bfa}(\Sigma_{I\times U}, Y)$.

(E2): Another more general system model can be described by the following evolution equation

$$dx = Ax(t)dt + F(t, x(t))\gamma(dt) + \int_U \mathcal{G}(t, x(t), \xi)\mu(dt \times d\xi), \quad x(0) = x_0, \ (6.5)$$

where $\{A, \mathcal{G}, \mu\}$ are as in (E1) and $\gamma \in \mathcal{M}_{bfa}(\Sigma_I, E)$ is a fixed finitely additive E-valued vector measure having bounded total variation and $F : I \times X \longrightarrow \mathcal{L}(E, X)$ where E is another real Banach space. The measure $\mu \in \mathcal{M}_{ad}(\Sigma_{I\times U}, Y)$ plays the role of controls. Given that γ is nonatomic, all the results presented in this paper also apply to this case.

Some Open Problems:

(P1): In (A3), we assumed that the dominating measure ν belonging to $B^+_{bfy}(\Sigma_I)$ is nonatomic, and also in (E2) we assumed the measure $\gamma \in \mathcal{M}_{bfa}(\Sigma_I, E)$ to be nonatomic. For a wider scope of applications, in particular where it may be necessary to include measures containing both regular and singular components, it is important to relax this assumption.

(P2): For purely impulsive systems, numerical techniques based on the necessary conditions of optimality (4.28)-(4.32) are easily developed and it is classical [5, 11, 13]. For the arbitrary measure driven systems, like (2.1) and (6.1), it is expected that one can use the necessary conditions of optimality (4.1)-(4.3) (Theorem 4.1) to develop suitable numerical techniques. However, to the best of knowledge of the author, no such technique is currently available in the literature.

References

[1] N.U. Ahmed, *Semigroup Theory With Applications to systems and Control,* Pitman research Notes in Mathematics Series (1991) , Vol. 246, Longman scientific and technical, U.K (copublished in U.S by John Wiley & Sons. Inc, New York).

[2] N.U. Ahmed, *Some remarks on the dynamics of impulsive systems in Banach spaces,* Dyna. of Contin., Disc. and Impul. Syst.Series A, **8** (2001), 261-274.

[3] N.U. Ahmed, *Impulsive evolution equations in infinite dimensional spaces,* Dyna. of Contin., Disc. and Impul. Syst.Series A, **10** (2003), 11-24.

[4] N.U. Ahmed, *Some Recent Developments in Systems and Control Theory on Infinite Dimensional Banach Spaces*: Part 1 (Part 2) , Optimization and Control Techniques and Applications, Springer, Berlin, Heidelberg, 2014, 3-24 (25-44).

[5] N.U. Ahmed and Shi'an Wang, *Measure Driven Nonlinear Dynamic Systems with Applications to Optimal Impulsive Controls,* Submitted.

[6] N.U. Ahmed, *Existence of optimal controls for a general class of impulsive systems on Banach spaces,* SIAM J. on Cont. and Opti., **42** (2003), 669-685.

[7] N.U. Ahmed, *Optimal relaxed controls for systems gverned by impulsive differential inclusions,* Nonlinear Funct. Anal. Appl., **10**(3) (2005), 427-460

[8] J. Diestel and J.J. Uhl, Jr., *Vector Measures,* Amer. Math. Soc. Providence, Rhode Island 1978.

[9] N. Dunford and J.T. Schwartz, *Linear Operators,* Part 1: General Theory, Interscience publishers Inc., New York. (second printing) 1964,

[10] R. Korn, *Portfolio optimization with strictly positive transaction costs and impulse control,* Finance and Stochastics, **2** (1998), 85-114.

[11] Q. Lin, R. Loxton, K.L. Teo, and Y.H. Wu, *A new computational method for optimizing nonlinear impulsive systems,* Dyna. of Conti. Disc. and Impul. Syst. Series B, **18** (2011), 59-76.

[12] C.E. Neuman and V. Costanza,*Deterministic impulse control in native forest ecosystems management,* J. of Opti. Theory and Appl., **66** (1990), 173-196.

[13] C.Z. Wu and K.L. Teo, *Global impulsive optimal control computation,* J. of Indus. and Manag. Opti., **2** (2006), 435-450.