The bang–bang theorem via Baire category: a dual approach

Alberto Bressan, Marco Mazzola and Khai T. Nguyen

Abstract. The paper develops a new approach to the classical bang–bang theorem in linear control theory, based on Baire category. Among all controls which steer the system from the origin to a given point $\bar{x}$, we consider those which minimize an auxiliary linear functional ϕ . For all ϕ in a residual set, we show that the minimizing control is unique and takes values within a set of extreme points.

Mathematics Subject Classification. 34A60, 54E52.

1. Introduction

For $t \in [0, T]$, consider the Cauchy problem

$$\dot{x}(t) \in F(x(t)), \quad (1.1)$$

$$x(0) = 0, \quad (1.2)$$

where $x \mapsto F(x) \subset \mathbb{R}^n$ is a bounded, Hausdorff continuous multifunction with compact convex values. Call $\mathcal{F} \subset \mathcal{C}([0, T]; \mathbb{R}^n)$ the set of Carathéodory solutions of (1.1). Moreover, call $\mathcal{F}^{ext}$ the set of trajectories of

$$\dot{x}(t) \in extF(x(t)), \quad (1.3)$$

where the time derivative takes values within the set of extreme points of $F(x)$.

If F is Lipschitz continuous w.r.t. the Hausdorff distance, starting with the seminal paper by Cellina [6], it is now well known that the set of extremal solutions $\mathcal{F}^{ext}$ is a *residual* subset of $\mathcal{F}$, i.e. it contains the intersection of countably many open dense subsets [8, 9]. By an application of Baire's theorem, this implies that the set $\mathcal{F}^{ext}$ is nonempty and every solution of (1.1) can be uniformly approximated by solutions of (1.3).

In [3] an alternative approach was developed, still based on Baire category but from a dual point of view.

For every $w \in \mathbb{R}^n$, consider the compact, convex subset of vectors in $F(x)$ which maximize the inner product with w , namely

$$F^w(x) \doteq \left\{ y \in F(x); \langle y, w \rangle = \max_{y' \in F(x)} \langle y', w \rangle \right\}. \quad (1.4)$$

For each continuous path $t \mapsto w(t)$, the multifunction $F^w(t, x) \doteq F^{w(t)}(x)$ is upper semicontinuous with compact, convex values. Hence the Cauchy problem

$$\dot{x}(t) \in F^{w(t)}(x(t)), \quad x(0) = 0, \quad (1.5)$$

has a non-empty, compact set of solutions $\mathcal{F}^w \subset \mathcal{C}([0, T]; \mathbb{R}^n)$. The main result in [3] shows that for “almost all” functions $w \in \mathcal{C}([0, T]; \mathbb{R}^n)$, in the Baire category sense, all solutions of (1.5) are also solutions of (1.3).

Theorem 1. *Let F be a bounded, Hausdorff continuous multifunction on $\mathbb{R}^n$, with compact convex values. Then the set*

$$W \doteq \{w(\cdot); \mathcal{F}^w \subseteq \mathcal{F}^{ext}\} \quad (1.6)$$

is a residual subset of $\mathcal{C}([0, T]; \mathbb{R}^n)$.

Notice that here the Lipschitz continuity of F is not required. Incidentally, this yields yet another proof of the classical theorem of Filippov [10], on the existence of solutions to differential inclusions with continuous, non-convex valued right hand side.

The purpose of the present paper is to explore whether a “dual” Baire category approach can be applied also to some boundary value problems (or variational problems) without convexity assumptions. A general formulation of these two approaches is as follows.

Consider a non-convex problem P , and let $\mathcal{S}$ be the set of solutions to a suitable convexified problem $\hat{P}$. Under natural assumptions, $\mathcal{S}$ will be a nonempty, closed subset of a Banach space, hence a complete metric space. Moreover, one can identify a set $\mathcal{S}^{ext} \subset \mathcal{S}$ of “extremal solutions” which solve the original non-convex problem P .

- **Direct approach** Show that the set $\mathcal{S}^{ext}$ of extremal solutions is residual in $\mathcal{S}$.
- **Dual approach** Consider a family of constrained optimization problems

$$\min_{u(\cdot) \in \mathcal{S}} J^w(u), \quad (1.7)$$

where the functional J^w depends on an auxiliary function w , ranging in a Banach space W . For each $w \in W$, call $\mathcal{S}^w$ the set of minimizers. Show that the set $\{w \in W; \mathcal{S}^w \subseteq \mathcal{S}^{ext}\}$ is residual in W .

We apply these ideas to derive an alternative proof of the classical bang–bang principle. Namely, consider the linear control system in $\mathbb{R}^n$

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \quad u(t) \in \Omega. \quad (1.8)$$

Here $A(t)$ and $B(t)$ are $n \times n$ and $n \times m$ matrices respectively, while $\Omega \subset \mathbb{R}^m$ is a compact convex set. Together with (1.8) we consider a system where the control takes values in the extreme points:

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \quad u(t) \in ext \Omega. \quad (1.9)$$

Given initial and terminal conditions

$$x(0) = 0, \quad x(T) = \bar{x}, \quad (1.10)$$

assume that the boundary value problem (1.8)–(1.10) has a solution. Then by the bang–bang theorem [5, 7, 11] the problem (1.9)–(1.10) has a solution as well.

The standard proof of the bang–bang theorem relies on Lyapunov’s theorem, providing the convexity of the range of a non-atomic vector measure. An alternative approach, based on Baire category, was developed in [4]. Call

$$\mathcal{S} \doteq \{x : [0, T] \mapsto \mathbb{R}^n; \quad x(0) = 0, \quad x(T) = \bar{x}, \quad \dot{x}(t) = A(t)x(t) + B(t)u(t) \\ \text{for some measurable control } u : [0, T] \mapsto \Omega\}.$$

$$\mathcal{S}^{ext} \doteq \{x : [0, T] \mapsto \mathbb{R}^n; \quad x(0) = 0, \quad x(T) = \bar{x}, \quad \dot{x}(t) = A(t)x(t) + B(t)u(t) \\ \text{for some measurable control } u : [0, T] \mapsto ext \Omega\}.$$

As proved in [4], one has

Theorem 2. *Let A, B be bounded, measurable, matrix-valued functions, and let $\Omega \subset \mathbb{R}^m$ be a compact convex set. Assume that $\mathcal{S} \neq \emptyset$. Then $\mathcal{S}$ is compact in $\mathcal{C}([0, T]; \mathbb{R}^n)$ and $\mathcal{S}^{ext}$ is a residual subset of $\mathcal{S}$. In particular, $\mathcal{S}^{ext}$ is nonempty.*

In this paper we develop a “dual” approach, also based on Baire category. Let $\mathcal{U}$ be the set of all measurable functions $u : [0, T] \mapsto \Omega$ such that the solution to the Cauchy problem

$$\dot{x}(t) = A(t)x(t) + B(t)u(t), \quad x(0) = 0 \tag{1.11}$$

satisfies the terminal condition

$$x(T) = \bar{x}. \tag{1.12}$$

Throughout the following, we shall assume that $\mathcal{U}$ is non-empty. Given any continuous function $w \in \mathcal{C}([0, T]; \mathbb{R}^m)$, consider the constrained optimization problem

$$\min_{u \in \mathcal{U}} \int_0^T \langle w(t), u(t) \rangle dt. \tag{1.13}$$

We show that, for “almost all” continuous functions $w \in \mathcal{C}([0, T]; \mathbb{R}^m)$, in the sense of Baire category, the problem (1.13) has a unique minimizer. This minimizer is a bang–bang control, i.e. it takes values within the set of extreme points of Ω .

Theorem 3. *Let A, B be bounded, measurable, matrix-valued functions, and let $\Omega \subset \mathbb{R}^m$ be a compact convex set. Let $\mathcal{W} \subseteq \mathcal{C}([0, T]; \mathbb{R}^m)$ be the set of all continuous functions w such that the variational problem (1.13) has a unique minimizer, satisfying $u(t) \in ext \Omega$ for a.e. $t \in [0, T]$. Then $\mathcal{W}$ is residual in $\mathcal{C}([0, T]; \mathbb{R}^m)$.*

A proof of the theorem will be given in the next section. For the basic theory of multifunctions and differential inclusions we refer to the classical monograph [1]. Definitions and basic properties of Young measures, mentioned in the proof, can be found in [13–15].

2. Proof of Theorem 3

1. In the first step we define for any $\varepsilon > 0$ and $N \geq 1$ a set $\mathcal{W}_{\varepsilon,N} \subseteq \mathcal{C}([0, T]; \mathbb{R}^m)$. We will show afterwards that each $\mathcal{W}_{\varepsilon,N}$ is open and dense in $\mathcal{C}([0, T]; \mathbb{R}^m)$ and that the intersection of all these sets is contained in $\mathcal{W}$.

Given the compact convex set $\Omega \subset \mathbb{R}^m$, consider the function $\varphi : \mathbb{R}^m \mapsto \mathbb{R} \cup \{-\infty\}$ defined by

$$\varphi(y) \doteq \max \left\{ \int_0^1 |f(s) - y|^2 ds : f : [0, 1] \mapsto \Omega, \int_0^1 f(s) ds = y \right\},$$

with the provision that $\varphi(y) \doteq -\infty$ if $y \notin \Omega$.

Otherwise stated, $\varphi(y)$ is the maximum variance among all probability measures supported on Ω , whose barycenter is at y . As proved in [2], the function φ is upper semicontinuous and concave on Ω . Moreover, we have the equivalence

$$\varphi(y) = 0 \quad \Longleftrightarrow \quad y \in \text{ext } \Omega. \quad (2.1)$$

A measurable control $u : [0, T] \mapsto \Omega$ takes values in $\text{ext } \Omega$ for a.e. time t iff

$$\int_0^T \varphi(u(t)) dt = 0. \quad (2.2)$$

The solution to the Cauchy problem (1.11) can be written as

$$x(t) = \int_0^t M(t, s) B(s) u(s) ds,$$

where $M(t, s)$ is the $n \times n$ matrix fundamental solution to the linear homogeneous system $\dot{x} = A(t)x$. In other words,

$$\frac{\partial}{\partial t} M(t, s) = A(t)M(t, s), \quad M(s, s) = I_n,$$

where I_n is the $n \times n$ identity matrix. Setting $G(s) \doteq M(T, s)B(s)$, for any $w \in \mathcal{C}([0, T]; \mathbb{R}^m)$ the optimization problem (1.13) can be reformulated as follows.

(OP) Find $u : [0, T] \mapsto \Omega$ which minimizes the integral

$$J^w(u) \doteq \int_0^T \langle w(t), u(t) \rangle dt \quad (2.3)$$

subject to

$$\int_0^T G(t) u(t) dt = \bar{x}. \quad (2.4)$$

For any $\varepsilon > 0$ and $N \geq 1$, consider the set $\mathcal{W}_{\varepsilon,N} \subseteq \mathcal{C}([0, T]; \mathbb{R}^m)$ of all functions w with the following property. If

$$u(t) \in \arg\min_{\omega \in \Omega} \langle \omega, w(t) - \lambda G(t) \rangle \quad (2.5)$$

for some row vector $\lambda = (\lambda_1, \dots, \lambda_n) \in [-N, N]^n$ and all $t \in [0, T]$, then

$$\int_0^T \varphi(u(t)) dt < \varepsilon. \quad (2.6)$$

In the next two steps we will prove that each $\mathcal{W}_{\varepsilon,N}$ is open and dense, hence the intersection

$$\mathcal{W}^\sharp \doteq \bigcap_{\varepsilon>0, N\geq 1} \mathcal{W}_{\varepsilon,N} \quad (2.7)$$

is residual in $\mathcal{C}([0, T]; \mathbb{R}^m)$.

2. We show that each set $\mathcal{W}_{\varepsilon,N}$ is open in $\mathcal{C}([0, T]; \mathbb{R}^m)$. Indeed, consider a sequence $(w_i)_{i\geq 1}$ converging uniformly to w and such that $w_i \notin \mathcal{W}_{\varepsilon,N}$ for every i . Then there exist sequences $\lambda^i \in [-N, N]^n$ and $u_i : [0, T] \mapsto \Omega$ satisfying

$$u_i(t) \in \operatorname{argmin}_{\omega \in \Omega} \langle \omega, w_i(t) - \lambda^i G(t) \rangle \quad \text{for all } t \in [0, T] \quad (2.8)$$

and

$$\int_0^T \varphi(u_i(t)) dt \geq \varepsilon. \quad (2.9)$$

By taking subsequences, we can assume that $\lambda^i \rightarrow \lambda \in [-N, N]^n$ and u_i converges weakly in $\mathbf{L}^2([0, T], \mathbb{R}^m)$ to some function u . By the linearity of the functional in (2.5), and the convexity of the set Ω , this limit function $u(\cdot)$ satisfies (2.5), for a.e. $t \in [0, T]$.

By the upper semicontinuity and the concavity of φ in Ω it follows

$$\int_0^T \varphi(u(t)) dt \geq \varepsilon. \quad (2.10)$$

Hence $w \notin \mathcal{W}_{\varepsilon,N}$, showing that $\mathcal{W}_{\varepsilon,N}$ is open in $\mathcal{C}([0, T]; \mathbb{R}^m)$.

3. In this step we prove that each $\mathcal{W}_{\varepsilon,N}$ is dense in $\mathcal{C}([0, T]; \mathbb{R}^m)$. Consider the function $\Phi : \mathbb{R}^m \times \mathbb{R}^m \mapsto \mathbb{R}$ defined by

$$\Phi(v, w) \doteq \max \left\{ \varphi(u); u \in \operatorname{argmin}_{\omega \in \Omega} \langle \omega, w - v \rangle \right\}. \quad (2.11)$$

Observing that the map Φ is upper semicontinuous, for every $\eta > 0$ we can consider the Lipschitz continuous approximation

$$\Phi^\eta(v, w) \doteq \max \{ \Phi(v', w') - \eta|v' - v| - \eta|w' - w|; (v', w') \in \mathbb{R}^m \times \mathbb{R}^m \}. \quad (2.12)$$

Notice that for any $\eta' \leq \eta$ it holds

$$\Phi^{\eta'}(v, w) \geq \Phi^\eta(v, w), \quad \text{for all } (v, w) \in \mathbb{R}^m \times \mathbb{R}^m. \quad (2.13)$$

For any fixed $v \in \mathbb{R}^m$, it is known that $\Phi(v, w) = 0$ for almost every $w \in \mathbb{R}^m$. Indeed (see [12]) $\Phi(v, w) = 0$ if and only if the minimum

$$\psi(w) \doteq \min_{\omega \in \Omega} \langle \omega, w - v \rangle$$

is attained at a single point. This is true if and only if the Lipschitz map $z \mapsto \psi(z)$ is differentiable at the point $z = w$. By Rademacher's theorem, this holds for a.e. $w \in \mathbb{R}^n$.

We claim that for every $\delta > 0$ and $K, M > 0$ there exists $\eta > 0$ sufficiently large so that

$$\int_{B(0,K)} \Phi^\eta(v, w) dw < \delta \quad \text{for all } v \in \mathbb{R}^m, \quad |v| \leq M. \quad (2.14)$$

Otherwise, let $(v_k, \eta_k)_{k \geq 1}$ be a sequence of vectors in $\mathbb{R}^m \times \mathbb{R}$ such that $\lim_{k \rightarrow \infty} \eta_k = +\infty$, $\lim_{k \rightarrow \infty} v_k = \bar{v}$, and

$$\int_{B(0,K)} \Phi^{\eta_k}(v_k, w) dw \geq \delta \quad \text{for all } k \geq 1. \quad (2.15)$$

By (2.13), for any $\eta > 0$ there is $K_\eta > 0$ such that

$$\int_{B(0,K)} \Phi^\eta(v_k, w) dw \geq \delta, \quad \text{for all } k \geq K_\eta.$$

The Lipschitz continuity of Φ^η implies

$$\int_{B(0,K)} \Phi^\eta(\bar{v}, w) dw = \lim_{k \rightarrow \infty} \int_{B(0,K)} \Phi^\eta(v_k, w) dw \geq \delta.$$

Therefore, letting $\eta \rightarrow +\infty$ we obtain

$$\int_{B(0,K)} \Phi(\bar{v}, w) dw \geq \delta. \quad (2.16)$$

This yields a contradiction because $\Phi(\bar{v}, w) = 0$ for almost every w .

Next, choose a constant $L > 0$ such that

$$\varphi(u) \leq L \quad \text{for all } u \in \Omega. \quad (2.17)$$

Relying on Lusin's theorem, we can replace the measurable matrix-valued function G with a continuous function $\widehat{G}$ such that

$$\text{meas}(\{t; \widehat{G}(t) \neq G(t)\}) < \frac{\varepsilon}{4L}. \quad (2.18)$$

Let any radius $\rho > 0$ and any $\tilde{w} \in \mathcal{C}([0, T]; \mathbb{R}^m)$ be given. By the previous claim, we can find $\eta > 0$ large enough so that

$$\int_{B(\tilde{w}(t), \rho)} \Phi^\eta(\lambda \widehat{G}(t), w) dw \leq \frac{\varepsilon}{4T}, \quad (2.19)$$

for every $t \in [0, T]$ and $\lambda \in [-N, N]^n$. Here and in the sequel, $\int_S f dx$ denotes the average value of the function f over the set S . From (2.19) we deduce

$$\int_0^T \int_{B(\tilde{w}(t), \rho)} \Phi^\eta(\lambda \widehat{G}(t), w) dw dt \leq \frac{\varepsilon}{4}. \quad (2.20)$$

Consider a sequence of continuous functions $w_\nu \in \mathcal{C}([0, T]; \mathbb{R}^m)$, weakly converging to $\tilde{w}$, such that

- (i) $w_\nu(t) \in B(\tilde{w}(t), \rho)$, for all $\nu \geq 1$ and $t \in [0, T]$,
- (ii) as $\nu \rightarrow \infty$, the limit is described by the family of Young measures $\{\mu_t; t \in [0, T]\}$, where μ_t is the probability measure uniformly distributed on the ball $B(\tilde{w}(t), \rho)$.

The existence of a sequence $(w_\nu)_{\nu \geq 1}$ satisfying the properties (i)–(ii) follows from a standard construction, which we briefly review. Consider a sequence of points $y_j \in B(0, \rho)$ uniformly distributed on the ball centered at the origin with radius ρ . That means

$$\lim_{\nu \rightarrow \infty} \frac{1}{\nu} \sum_{j=1}^{\nu} f(y_j) = \int_{B(0, \rho)} f \, dx, \quad (2.21)$$

for every continuous function $f : \mathbb{R}^m \mapsto \mathbb{R}$. For a given $\nu \geq 1$, divide the interval $[0, T]$ into ν^2 subintervals, inserting the times $t_\ell = h\ell$, with $h = T/\nu^2$. Consider the piecewise continuous function

$$z_\nu(t) \doteq \tilde{w}(t) + y_j \quad \text{if } t \in [t_{\ell-1}, t_\ell] \quad \text{with } \ell = m\nu + j \text{ for some integer } m.$$

Then choose a continuous function w_ν such that

$$\text{meas}(\{t; w_\nu(t) \neq z_\nu(t)\}) < \frac{1}{\nu}.$$

This sequence satisfies the required properties.

Since Φ^η is continuous, this yields

$$\lim_{\nu \rightarrow \infty} \int_0^T \Phi^\eta(\lambda \hat{G}(t), w_\nu(t)) \, dt = \int_0^T \int_{B(\tilde{w}(t), \rho)} \Phi^\eta(\lambda \hat{G}(t), w) \, dw \, dt.$$

Choosing $w = w_\nu$ for some ν sufficiently large, in view of (2.20) we obtain

$$\int_0^T \Phi^\eta(\lambda \hat{G}(t), w(t)) \, dt \leq \frac{\varepsilon}{3}. \quad (2.22)$$

In turn, by (2.17) and (2.18) and the obvious inequality $\Phi \leq \Phi^\eta$, this implies

$$\int_0^T \Phi(\lambda G(t), w(t)) \, dt \leq L \frac{\varepsilon}{4L} + \int_0^T \Phi^\eta(\lambda \hat{G}(t), w(t)) \, dt \leq \frac{\varepsilon}{4} + \frac{\varepsilon}{3}. \quad (2.23)$$

Since $\|w - \tilde{w}\|_C \leq \rho$, and $\rho > 0$ can be chosen arbitrarily small, this proves the density of $\mathcal{W}_{\varepsilon, N}$.

4. To conclude the proof of Theorem 3, it remains to show that the intersection $\mathcal{W}^\#$ in (2.7) is a subset of $\mathcal{W}$.

Assume that $w \in \mathcal{W}^\#$. We claim that the optimization problem **(OP)** has a unique solution $u : [0, T] \mapsto \text{ext } \Omega$. Indeed, assume that u_1, u_2 are two distinct solutions. Then $u(t) = \frac{u_1(t) + u_2(t)}{2}$ is also a solution. Set $v(t) \doteq \frac{u_1(t) - u_2(t)}{2}$ and consider the reachable subspace

$$Y \doteq \left\{ \int_0^T G(t)v(t)\theta(t) \, dt; \quad \theta : [0, T] \mapsto \mathbb{R} \text{ measurable} \right\} \subseteq \mathbb{R}^n. \quad (2.24)$$

Consider the control system on the product space $\mathbb{R} \times Y$, with variables (y_0, y) .

$$\begin{cases} \dot{y}_0(t) = \langle w(t), v(t) \rangle \theta(t), & \begin{cases} y_0(0) = 0, \\ y(0) = 0. \end{cases} \\ \dot{y}(t) = G(t)v(t)\theta(t). \end{cases}$$

Then the control $\theta^*(t) \equiv 0$ provides an optimal solution to the problem

$$\text{minimize: } y_0(T)$$

$$\text{subject to} \quad \theta(t) \in [-1, 1], \quad y(T) = 0.$$

The necessary conditions for optimality yield the existence of a nonzero vector $\lambda = (\lambda_0, \lambda) \in \mathbb{R} \times Y$ such that

$$\theta^*(t) = \arg \min_{\theta \in [-1, 1]} (\lambda_0 \langle w(t), v(t) \rangle - \langle \lambda, G(t)v(t) \rangle) \theta$$

for a.e. $t \in [0, T]$. If $\lambda_0 = 0$, then the non-zero vector $\lambda \in Y$ has the property

$$\langle \lambda, G(t)v(t) \rangle = 0 \quad \text{for a.e. } t \in [0, T].$$

But this contradicts the definition (2.24) of Y .

If $\lambda_0 \neq 0$, by a normalization we can assume $\lambda_0 = 1$. Then there exists a vector $\lambda \in Y$ such that

$$\langle w(t), v(t) \rangle - \langle \lambda, G(t)v(t) \rangle = 0 \quad \text{for a.e. } t \in [0, T].$$

This implies

$$u(t) \in \operatorname{argmin}_{\omega \in \Omega} \langle \omega, w(t) - \lambda G(t) \rangle$$

while

$$\int_0^T \varphi(u(t)) dt \geq \int_0^T |v(t)|^2 dt > 0.$$

This contradicts the assumption $w \in \mathcal{W}^\sharp$. We thus conclude that $\mathcal{W}^\sharp \subseteq \mathcal{W}$. By (2.7), $\mathcal{W}^\sharp$ is residual, hence the same is true for $\mathcal{W}$. $\square$

References

- [1] Aubin, J.P., Cellina, A.: *Differential Inclusions*. Springer, Berlin (1984)
- [2] Bressan, A.: The most likely path of a differential inclusion. *J. Differ. Equ.* **88**, 155–174 (1990)
- [3] Bressan, A.: Extremal solutions to differential inclusions via Baire category: a dual approach. *J. Differ. Equ.* **255**, 2392–2399 (2013)
- [4] Bressan, A., Piccoli, B.: A Baire category approach to the bang–bang property. *J. Differ. Equ.* **116**, 318–337 (1995)
- [5] Bressan, A., Piccoli, B.: *Introduction to the mathematical theory of control*. In: AIMS Series in Applied Mathematics. Springfield, Mo (2007)
- [6] Cellina, A.: On the differential inclusion $x' \in [-1, 1]$. *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* **69**, 1–6 (1980)
- [7] Cesari, L.: *Optimization—theory and applications*. In: *Problems with Ordinary Differential Equations*. Springer, New York (1983)
- [8] De Blasi, F.S., Pianigiani, G.: The Baire category method in existence problems for a class of multivalued differential equations with nonconvex right-hand side. *Funkcial. Ekvac.* **28**, 139–156 (1985)

- [9] De Blasi, F.S., Pianigiani, G.: Baire category and boundary value problems for ordinary and partial differential inclusions under Carathéodory assumptions. *J. Differ. Equ.* **243**, 558–577 (2007)
- [10] Filippov, A.F.: The existence of solutions of multivalued differential equations. (Russian) *Mat. Zametki* **10**, 307–313 (1971). (**English translation in Math. Notes** **10**, 608–611 (1971))
- [11] Hermes, H., LaSalle, J.P.: *Functional Analysis and Time Optimal Control*. Academic Press, New York (1969)
- [12] Rockafellar, R.T.: *Convex Analysis*. Princeton University Press, Princeton (1970)
- [13] Sychev, M.A.: A new approach to Young measure theory, relaxation and convergence in energy. *Ann. Inst. Poincaré Anal. Non Linéaire* **16**, 773–812 (1999)
- [14] Valadier, M.: Young measures. In: Cellina, A. (ed.) *Methods of Nonconvex Analysis*. *Lecture Notes in Mathematics*, vol. 1446, pp. 152–188. Springer, Berlin (1990)
- [15] Young, L.C.: *Lectures on the Calculus of Variations and Optimal Control Theory*. Saunders Co., Philadelphia (1969)

Alberto Bressan, Khai T. Nguyen
Department of Mathematics
Penn State University
State College, USA
e-mail: ktn2@psu.edu

Alberto Bressan
e-mail: bressan@math.psu.edu

Marco Mazzola
Université Pierre et Marie Curie
Paris VI, France
e-mail: marco.mazzola@imj-prg.fr

Received: 22 November 2015.

Accepted: 11 July 2016.