Random Extremal Solutions of Differential Inclusions

Alberto Bressan and Vasile Staicu

Abstract. Given a Lipschitz continuous multifunction F on $\mathbb{R}^n$, we construct a probability measure on the set of all solutions to the Cauchy problem $\dot{x} \in F(x)$ with $x(0) = 0$. With probability one, the derivatives of these random solutions take values within the set $ext F(x)$ of extreme points for a.e. time t. This provides an alternative approach in the analysis of solutions to differential inclusions with non-convex right hand side.

1. Introduction

Let $x \mapsto F(x) \subset \mathbb{R}^n$ be a Lipschitz continuous multifunction with compact values. Consider the Cauchy problem

$$\dot{x}(t) \;\in\; F\big(x(t)\big) \qquad\qquad t \in [0, T], \tag{1.1}$$

$$x(0) \;=\; 0. \tag{1.2}$$

The main goal of this paper is to construct a probability measure P on the space of continuous functions $\mathcal{C}\big([0, T]; \mathbb{R}^n\big)$ which is supported on the set $\mathcal{F}$ of Carathéodory solutions of (1.1)-(1.2). Moreover, calling $\mathcal{F}^{ext}$ the family of solutions of

$$\dot{x}(t) \;\in\; ext\, F(x(t)) \tag{1.3}$$

with initial data (1.2), the following properties should hold:

(P1) With probability one, trajectories of (1.1)-(1.2) satisfy (1.3). Namely

$$P(\mathcal{F}^{ext}) = 1. \tag{1.4}$$

(P2) The support of P is dense on $\mathcal{F}$. Namely: for every open set $U \subset \mathcal{C}([0,T];\, \mathbb{R}^n)$,

$$U \cap \mathcal{F} \neq \emptyset \qquad \Longleftrightarrow \qquad P(U) > 0. \tag{1.5}$$

In the special case where the sets $F(x)$ are segments in the plane, a probability measure satisfying (1.4)-(1.5) was first constructed in [11]. In the present paper we develop a different approach, which applies to any Lipschitz continuous multifunction with compact values in $\mathbb{R}^n$. In the following, $\mathcal{K}_n$ denotes the family of all nonempty, compact subsets of $\mathbb{R}^n$, endowed with the Hausdorff metric. Our main result is

Theorem 1. *Let $F : \mathbb{R}^n \mapsto \mathcal{K}_n$ be a Lipschitz continuous multifunction. Then there exists a probability measure P on the set $\mathcal{F}$ of Carathéodory solutions of (1.1)-(1.2) such that the properties (P1)-(P2) hold.*

Extremal trajectories of differential inclusions are of interest for various applications [12, 16]. Since the seminal paper by A. Cellina [10], Baire category has provided a key tool for proving the existence of solutions of (1.3). This approach consists of two main steps.

(i) Replacing $F(x)$ by its closed convex hull, prove that the convexified problem

$$\dot{x}(t) \in \overline{co}F(x(t)), \tag{1.6}$$

has a nonempty, closed set of solutions.

(ii) Prove that the family of solutions to (1.3) is a subset of second category of the family of all solutions to (1.6). Hence, by Baire's theorem, it is nonempty.

Various applications of this technique can be found in [5, 7, 8, 12, 13, 14, 15, 16, 24]. We remark that in [5] a "set-valued" extension of the Baire category theorem was introduced, replacing "points" with "compact sets". This can be applied also in cases where F is not Lipschitz continuous and the set $\mathcal{F}^{ext}$ of solutions to (1.3) is not dense on $\mathcal{F}$. In addition, the recent paper [4] proved that, for a generic (in the sense of Baire category) continuous guiding function $t \mapsto \omega(t) \in \mathbb{R}^n$, all trajectories of

$$\dot{x}(t) \in F^{\omega(t)}(x(t)), \qquad x(0) = 0, \tag{1.7}$$

satisfy (1.3) as well. Here

$$F^{\omega}(x) \doteq \left\{ y \in F(x); \ \langle y, \omega \rangle = \max_{y' \in F(x)} \langle y', \omega \rangle \right\} \tag{1.8}$$

is the set of vectors in $F(x)$ maximizing the inner product with ω.

As shown in Oxtoby's book [20], measure theory and Baire category share many similarities. A natural question, proposed by the first author some years ago, is whether solutions of the differential inclusion (1.3) can be obtained by an alternative argument, replacing Baire category with probability theory.

We recall the approach followed in [11]. Without loss of generality, assume that the multifunction F has convex values. Let ω denote Brownian motion on the surface of the unit sphere $S^{n-1} \subset \mathbb{R}^n$, with initial data uniformly distributed on S^{n-1}. Of course, by symmetry this implies that the random variable $\omega(t)$ is still uniformly distributed on S^{n-1} for every $t > 0$. The space $\mathcal{W}$ of all such Brownian paths is used as the basic probability space. For each sample path $t \mapsto \omega(t)$, consider the set of solutions $\mathcal{F}^{\omega(\cdot)}$ of the differential inclusion (1.7). Assume that, for a.e. path $\omega(\cdot) \in \mathcal{W}$, the Cauchy problem (1.7) has a unique solution $x^\omega(\cdot)$, with derivative contained in $ext F(x^\omega(t))$. Then the push-forward of the probability measure on $\mathcal{W}$ under the map $\omega(\cdot) \mapsto x^\omega(\cdot)$ yields a probability measure on $\mathcal{F}$ with the desired properties.

This approach was implemented in [11], in the special case where each set $F(x)$ is a segment in the plane. It is doubtful whether this technique can be extended to general Lipschitz continuous multifunctions F on $\mathbb{R}^n$. The key issue here is the behavior of the extremal faces F^ω. For the method to work, the guiding functions $\omega(\cdot)$ should "wiggle" much faster than the unit normal vectors to the flat faces of the sets $F(x)$. In the case considered in [11], the set

$$Q(x) \;=\; \big\{\omega \in \mathbb{R}^2 \,;\; |\omega| = 1, \; F^\omega(x) \text{ contains more than one point}\big\}$$

varies with x in a Lipschitz continuous way. Since Brownian paths are only Hölder continuous, one can show that for a.e. random path $\omega(\cdot)$, the set of times

$$\big\{t \in [0,T] \,;\; \omega(t) \in Q(x^\omega(t))\big\} \tag{1.9}$$

has measure zero. As a consequence, the analysis in [11] shows that the differential inclusion (1.7) has single-valued right hand side for a.e. time t, and a unique solution. However, for a general Lipschitz continuous multifunction F on $\mathbb{R}^n$, the corresponding set $Q(x)$ of vectors normal to its flat faces varies with x in a highly irregular way, possibly not Lipschitz continuous. It is far from clear whether the set of times (1.9) has measure zero, for a.e. Brownian path $\omega(\cdot)$.

For the above reason, we develop here an entirely different approach, based on piecewise Lipschitz approximate selections. The construction of the probability measure is described in Section 2, while the remaining two sections contain details of the proof.

We remark that, in the case where the multifunction F is only continuous (but not Lipschitz), a well known counterexample by Plis [22] shows that the set of solutions of (1.1)-(1.2) may not be dense on the set of solutions to (1.3)-(1.2). Because of this, there is no hope of achieving both properties (P1)-(P2). However, we conjecture that the statement of Theorem 1 remains valid for a general continuous multifunction F, provided that (P2) is replaced by the weaker property

(P2′) *Let f be a continuous map such that $f(x) \in \overline{co} F(x)$ for all $x \in \mathbb{R}^n$, and let $U \subseteq \mathcal{C}([0,T]\,;\,\mathbb{R}^n)$ be an open set containing all solutions to the*

Cauchy problem

$$\dot{x} \;=\; f(x)\,, \qquad\qquad x(0) = 0\,. \qquad\qquad (1.10)$$

Then $P(U) > 0$.

2. Construction of the probability measure

We denote by $B = B_1$ the closed unit ball in $\mathbb{R}^n$ with boundary $\partial B = S^{n-1}$, while σ is the normalized, $n-1$-dimensional surface measure on ∂B. More generally, we write $B_r = \{x \in \mathbb{R}^n \,;\, |x| \le r\}$ for the closed ball with radius r centered at the origin. By $\overline{A}$, $ext A$ and $\overline{co}A$ we denote respectively the closure, the set of extreme points, and the closed convex hull of a set A. A general introduction to the theory of multifunctions and differential inclusions can be found in the classical monograph [1]. For basic notions of probability theory we refer to [19].

The first ingredient in our construction is the Laplace-Beltrami operator Δ on the surface of the unit sphere $\partial B \subset \mathbb{R}^n$. For each unit vector $w \in \partial B$, we let $G(z,t,w)$ be the solution to

$$G_t = \Delta G\,, \qquad\qquad G(z,0) = \delta_w\,. \qquad\qquad (2.1)$$

Here δ_w denotes the Dirac distribution consisting of a unit mass at the point w. In other words, $G(\cdot,t,w)$ is the density at time t (w.r.t. the normalized surface measure σ on ∂B) of the probability measure describing a diffusion process on ∂B, starting at the point w.

By standard properties of parabolic equations, it is well known that $z \mapsto G(z,t,w)$ is smooth and strictly positive for all $t > 0$. Moreover, the Markov semigroup property holds:

$$G(z, s+t, w) = \int_{\partial B} G(z,t,w')G(w',s,w)\,dw' \text{ for all } z, w \in \partial B\,, \ s, t > 0\,. \tag{2.2}$$

The second ingredient is the measure of curvature on the boundary ∂K of a compact convex set $K \subset \mathbb{R}^n$. For any unit vector $\omega \in \partial B$, let $\psi^K(\omega) \in K$ be the point such that

$$\langle \psi^K(\omega), \omega \rangle \;=\; \max_{y \in K}\, \langle y, \omega \rangle\,. \qquad\qquad (2.3)$$

It is well known [23] that for a.e. $\omega \in \partial B$ the point $\psi^K(\omega)$ is uniquely defined, and is an extreme point of K. The curvature measure on ∂K is defined by

$$\nu^K(A) \;=\; \sigma\Big(\{\omega \in \partial B\,;\ \psi^K(\omega) \in A\}\Big)$$

for every Borel set $A \subseteq \mathbb{R}^n$. By definition, ν^K is the push-forward of the normalized surface measure σ by the map $\psi^K : \partial B \mapsto K$. Observe that ν^K is a probability measure on K supported on the closure $\overline{ext K}$ of the set of extreme points of K.

More generally, let f be a continuous function defined on the surface of the unit sphere, such that

$$f \geq 0, \qquad \int_{\partial B} f \, d\sigma = 1. \qquad (2.4)$$

Then we can consider the probability measure $\mu = f \cdot \sigma$ on ∂B having density f w.r.t. the normalized surface measure σ. In turn, for every compact set K, the push-forward $\psi^K \circ \mu$ is a probability measure on K, with support contained in $\overline{ext}\, K$.

Lemma 1. *Let $f : \partial B \mapsto \mathbb{R}$ be a continuously differentiable function which satisfies (2.4), and consider the measure $\mu \doteq f \cdot \sigma$. Then the barycenter $E^\mu[\psi^K]$ of the probability measure $\psi^K \circ \mu$ on K is a Lipschitz continuous function of K. More precisely, there exists a Lipschitz constant L depending only on f such that*

$$\left| E^\mu[\psi^K] - E^\mu[\psi^{K'}] \right| \leq L \cdot d_H(K, K'). \qquad (2.5)$$

for any two compact convex sets $K, K' \subset \mathbb{R}^n$.

For a proof, see Proposition 2.4 in [17].

In the special case $f \equiv 1$, the point $Steiner(K) \doteq E^\sigma[\psi^K]$ obtained by taking the barycenter of the measure of curvature on ∂K is called the *Steiner point* of K. It is well known that the map $K \mapsto Steiner(K)$ is uniformly Lipschitz continuous w.r.t. the Hausdorff distance, with a constant depending only on the dimension n of the space.

We shall repeatedly use Lemma 1 to provide the Lipschitz continuity of selections having the form

$$f^{\delta,w}(x) \doteq E^\mu[\psi^{F(x)}] \in F(x), \qquad (2.6)$$

where μ is the probability measure having density $G(\cdot, \delta, w)$ w.r.t. the uniform surface measure on ∂B.

The third ingredient in our construction is the Choquet-type functional $\phi : \mathbb{R}^n \times \mathcal{K}_n \mapsto \mathbb{R} \cup \{-\infty\}$, defined by

$$\phi(y, K) \doteq \sup \left\{ \int_0^1 |w(s) - y|^2 \, ds \, ; \ w : [0, 1] \mapsto K, \ \int_0^1 w(s) \, ds = y \right\}, \qquad (2.7)$$

with the understanding that $\phi(y, K) = -\infty$ if $y \notin \overline{co}K$. One can interpret $\phi(y, K)$ as the maximum variance among all random variables supported inside K whose mean value is y. We observe that $\phi(y, K) = \phi(y, \overline{co}K)$. The following results were proved in [3].

Lemma 2. *The map $(y, K) \mapsto \phi(y, K)$ is upper semicontinuous in both variables; for each fixed $K \in \mathcal{K}_n$ the function $y \mapsto \phi(y, K)$ is strictly concave down on $\overline{co}K$. Moreover, one has*

$$\phi(y, K) = 0 \qquad \text{if and only if} \qquad y \in extK. \qquad (2.8)$$

Following [3], we consider the functional

$$\Phi(x) \ \dot= \ \int_0^T \phi\Big(\dot x(t),\, F(x(t))\Big)\, dt\,, \tag{2.9}$$

defined for any solution $x(\cdot)$ of (1.1). Using Lemma 2 one obtains

Lemma 3. *Let P be a probability measure on the set $\mathcal{F}$ of solutions to (1.1). Then $P(\mathcal{F}^{ext}) = 1$ if and only if*

$$E^P[\Phi(x)] \ = \ 0\,. \tag{2.10}$$

Indeed, (2.10) holds if and only if $\Phi(x) = 0$ for P-a.e. $x \in \mathcal{F}$. By Lemma 2 this happens if and only if a.e. trajectory $x(\cdot)$ (w.r.t. the probability P) satisfies (1.3) at a.e. $t \in [0,T]$. □

We are now ready to construct a probability measure P on $\mathcal{F}$ with the desired properties (1.4)-(1.5). This will be obtained as a limit of a sequence of probability measures P_N defined inductively. We start with an increasing sequence of integer numbers

$$1 \ = \ n_1 \ < \ n_2 \ < \ n_3 \ < \ \cdots \tag{2.11}$$

For each k we have a partition of the interval $[0,T]$ into n_k equal subintervals:

$$I_{k,\ell} \ \dot= \ ]t_{k,\ell-1},\, t_{k,\ell}]\,, \qquad t_{k,\ell} = \frac{\ell T}{n_k}\,. \tag{2.12}$$

We assume that n_{k+1} is a multiple of n_k, so that each partition is a refinement of the previous one. In addition, we consider a decreasing sequence of real numbers

$$1 \ = \ \beta_1 \ > \ \beta_2 \ > \ \beta_3 \ > \ \cdots \tag{2.13}$$

with $\beta_k \to 0$ as $k \to \infty$. Our basic probability space will be a product of countably many copies of the unit sphere $\partial B = S^{n-1}$. More precisely, for $k \geq 1$ let

$$\Omega_k \ \dot= \ \underbrace{\partial B \times \cdots \times \partial B}_{n_k \text{ times}}\,, \qquad \Omega \ \dot= \ \prod_{k \geq 1} \Omega_k\,. \tag{2.14}$$

Points in Ω will be written as $w = (w_1, w_2, \ldots) \in \Omega$, with $w_k \in \Omega_k$ for every $k \geq 1$. In turn, points in Ω_k are denoted by $w_k = (w_{k,1}, \ldots, w_{k,n_k})$, with $w_{k,\ell} \in \partial B$ for every k, ℓ.

The probability distribution P on the space Ω is defined by induction on k.

(i) If $A \subseteq \Omega_1 = \partial B$, then

$$P\{w_1 \in A\} \ = \ \sigma(A).$$

(ii) Let the probability distribution of $w_{j,\ell}$ be given, for all $j \leq k - 1$. Then the probability distribution of $w_{k,\ell}$ is defined as follows.

    - The random variables $w_{k,\ell}$, $\ell = 1, \ldots, n_k$ are mutually independent, given $w_{k-1,\ell'}$, $\ell' = 1, \ldots, n_{k-1}$.

    - If $I_{k,\ell} \subset I_{k-1,\ell'}$, then, given $w_{k-1,\ell'}$, the random variable $w_{k,\ell}$ is distributed on the unit sphere with density $G\left(\cdot,\ \beta_{k-1} - \beta_k,\ w_{k-1,\ell'}\right)$.

Next, we construct a map $\Lambda : \Omega \mapsto \mathcal{F}$ as a pointwise limit of maps $\Lambda_k : \Omega \mapsto \mathcal{F}$, defined as follows.

For a fixed $k \geq 1$, let $w_k = (w_{k,1}, \ldots, w_{k,n_k})$ be given. We then define $\Lambda_k : \Omega_k \mapsto \mathcal{F}$ to be the map such that

$$\Lambda_k(w_k) \;=\; \Lambda_k(w_{k,1}, \ldots, w_{k,n_k}) \;=\; x(\cdot), \tag{2.15}$$

where $x : [0, T] \mapsto \mathbb{R}^n$ is the solution to

$$x(0) = 0, \qquad \dot{x}(t) \;=\; f^{\beta_k, w_{k,j}}(x) \qquad \text{if} \quad t \in I_{k,j}\,, \ j = 1, \ldots, n_k\,. \tag{2.16}$$

Here the vector field $f^{\beta_k, w_{k,j}}(x)$ is the selection from the convex-valued multifunction F, defined at (2.6). By Lemma 1 the right hand side of the ODE in (2.16) is a Lipschitz continuous function of x. Hence the solution exists and is unique.

Finally, we define the map $\Lambda : \Omega \mapsto \mathcal{F}$ by setting

$$\Lambda(w) \;\doteq\; \lim_{k \to \infty} \Lambda_k(w_k)\,. \tag{2.17}$$

The probability distribution on $\mathcal{F}$ is obtained as the push-forward of the probability distribution P on Ω via the map Λ.

In the next section we shall prove that, if the sequences n_k, β_k are suitably chosen, then the limit in (2.17) exists almost surely, and the properties (P1)-(P2) are satisfied.

To help the reader, we give here an intuitive explanation of the above construction. Recalling (2.6), for any given $\beta > 0$ and $w \in \partial B$, one has a Lipschitz continuous selection $x \mapsto f^{\beta,w}(x) \in F(x)$. For a.e. unit vector w, the point $\psi^{F(x)}(w)$ maximizing the inner product with w is an extreme point of $F(x)$. Recalling (2.7) we thus have

$$\lim_{\beta \to 0} \phi\big(f^{\beta,w}(x);\ F(x)\big) \;=\; \phi\big(\psi^{F(x)}(w);\ F(x)\big) \;=\; 0.$$

By choosing β_k sufficiently small, we thus guarantee that the the speed $\dot{x}$ of our random solution stays close to an extreme point most of the time.

To achieve convergence as $k \to \infty$ we observe that, if $0 < \beta' < \beta$, then by the property (2.2) of the heat kernel, one has

$$f^{\beta,w}(x) \;=\; \int_{\partial B} f^{\beta',w'}(x)\, d\mu(w'),$$

where μ is the probability measure on the unit sphere ∂B with density $G(\cdot, \beta - \beta', w)$ w.r.t. the uniform surface measure σ. This fact can be exploited as follows. For a given integer N, let $w_1, \ldots, w_N \in S^{n-1}$ be independent random variables, with probability distribution having density $G(\cdot, \beta' - \beta, w)$

w.r.t. the uniform surface measure. Denote by $x_N(\cdot)$ the solution to the random differential inclusion

$$x(0) = 0, \qquad \dot{x}(t) \; = \; f^{\beta, w_j}(x(t)) \qquad t \in \left] \frac{(j-1)\tau}{N}, \; \frac{j\tau}{N} \right].$$

Then we have the convergence

$$\lim_{N \to \infty} E \left[\sup_{t \in [0, \tau]} |x_N(t) - \bar{x}(t)| \right] \; = \; 0,$$

where $\bar{x}$ is the solution to

$$x(0) = 0, \qquad \dot{x}(t) \; = \; f^{\beta, w}(x(t)).$$

Relying on the above analysis, choosing the integers n_k in (2.11)-(2.12) such that $n_k \to \infty$ fast enough, we shall achieve the almost sure convergence of our random solutions.

3. The main estimates

Since $extF = ext(\overline{co}F)$, without loss of generality we can assume that all sets $F(x)$ are compact, convex. Moreover, we shall assume the uniform bound

$$F(x) \subseteq B_1 \qquad \text{for all } x \in B_T. \tag{3.1}$$

For every solution of (1.1)-(1.2), this implies

$$x(t) \; \in \; B_T \qquad \text{for all } t \in [0, T]. \tag{3.2}$$

To motivate the next lemma we observe that, if $\delta > 0$ is very small then the barycenter $f^{\delta, w}(x)$ of the measure μ in (2.6) will be very close to $\psi^{F(x)}(w)$. In particular, if $\psi^{F(x)}(w) \in extF(x)$ (which is true for a.e. w), then

$$\phi(f^{\delta, w}(x); \; F(x)) \approx 0.$$

Lemma 4. *For any $\varepsilon, \varepsilon' > 0$, there exists $\bar{\beta} > 0$ such that the following holds.*

For every $x \in B_T$, $w' \in \partial B$, $0 < \beta \leq \bar{\beta}$, one has

$$E^w \left[\sup_{|y - x| < \beta} \phi(f^{\beta, w}(y); \; F(y)) \right] \; < \; \varepsilon. \tag{3.3}$$

Here E^w denotes expectation w.r.t. the random variable $w \in \partial B$, having density $G(\cdot, \varepsilon', w')$ w.r.t. the uniform surface measure σ on ∂B.

Proof. If the conclusion does not hold, we can find $\varepsilon_0 > 0$ and sequences $x_n \in B_T$, $w'_n \in \partial B$ and $\beta_n \to 0$ such that

$$E_n^w \left[\sup_{|y - x_n| < \beta_n} \phi(f^{\beta_n, w}(y); \; F(y)) \right] \; \geq \; \varepsilon_0 \qquad \text{for all } n \geq 1. \tag{3.4}$$

Here E_n^w denotes expectation w.r.t. the random variable $w \in \partial B$, having density $G(\cdot, \varepsilon', w'_n)$.

By possibly taking a subsequence, we can assume the convergence $w'_n \to w'$, $x_n \to x$.

If $\psi^{F(x)}(w)$ is an extreme point of $F(x)$, then

$$\lim_{n\to\infty} f^{\beta_n,w}(x) = \psi^{F(x)}(w). \tag{3.5}$$

In turn, the upper semicontinuity of ϕ in (2.7) yields

$$\limsup_{n\to\infty} \sup_{|y-x|<\beta_n} \phi(f^{\beta_n,w}(y)) \leq \phi\left(\psi^{F(x)}(w),\ F(x)\right). \tag{3.6}$$

For a.e. $w \in \partial B$, $\psi^{F(x)}(w)$ is an extreme point of $F(x)$, hence the right hand side of (3.6) vanishes. By the Lebesgue dominated convergence theorem we conclude

$$\lim_{n\to\infty} E_n^w\left[\sup_{|y-x_n|<\beta_n} \phi\left(f^{\beta_n,w}(y);\ F(y)\right)\right] = E_n^w\left[\phi\left(\psi^{F(x)}(w),\ F(x)\right)\right] = 0.$$

This yields a contradiction with (3.4), proving the lemma. $\qquad\square$

The next lemma will be used to prove the convergence of the probability measures in (2.17). At this stage, the Lipschitz continuity of the multifunction F is essential.

Lemma 5. *Let the multifunction F be Lipschitz continuous. Assume $0 < \beta' < \beta$ and let any $w_1, \ldots, w_n \in \partial B$ be given. Let $z : [0,T] \mapsto B_T$ be the solution to the Cauchy problem*

$$z(0) = 0,\ \dot{z}(t) = f^{\beta,w_k}(z(t))\ if\ t \in I_{n,k} \doteq \left]\frac{k-1}{n}T,\ \frac{k}{n}T\right]. \tag{3.7}$$

Then, for any $\varepsilon_0 > 0$, there exist an integer m which is a multiple of n and such that the following holds.
Let $w'_1, \ldots, w'_m \in \partial B$ be independent random variables such that, if $I_{m,j} \subset I_{n,k}$, then w'_j has density $G(\cdot, \beta - \beta', w_k)$ w.r.t. the uniform surface measure σ. Consider the solution $x : [0,T] \mapsto B_T$ of

$$x(0) = 0,\qquad \dot{x}(t) = f^{\beta',w'_j}(x(t))\ if\ t \in I_{m,j} \doteq \left]\frac{j-1}{m}T,\ \frac{j}{m}T\right]. \tag{3.8}$$

Then, taking the expectation w.r.t. the distribution of $w'_1, \ldots, w'_m$, one has

$$E\left[\sup_{t\in[0,T]} |x(t) - z(t)|\right] \leq \varepsilon_0. \tag{3.9}$$

Proof. 1. Without loss of generality we can assume that, for every $x \in B_T \doteq \{x \in \mathbb{R}^n;\ |x| \leq T\}$, the set $F(x)$ is contained inside the closed unit ball B_1. As a consequence, for $t \in [0,T]$ all solutions of (3.7) and (3.8) take values inside the ball B_T. Let $L \geq 1$ be a common Lipschitz constant for all

vector fields $f^{\beta,w}$ and $f^{\beta',w'}$ on B_T. Notice that L depends only on β, β', and on the multifunction F. Moreover, let

$$\eta(t) \doteq \frac{\varepsilon}{L}(e^{Lt} - 1)$$

be the solution to the scalar Cauchy problem

$$\dot{\eta} = L\eta + \varepsilon, \qquad \eta(0) = 0.$$

Since all solutions x, z remain uniformly bounded, to prove (3.9) it suffices to show that, for any $\varepsilon > 0$, we can choose m large enough so that

$$\text{Prob.}\left\{|x(t) - z(t)| > \eta(t) + \varepsilon \quad \text{for some } t \in [0,T]\right\} \leq \varepsilon. \qquad (3.10)$$

Notice that here the solution $z(\cdot)$ is given, while $x(\cdot)$ is a random solution determined by $w'_1, \ldots, w'_m$. The probability of the set on the left hand side of (3.10) depends on the distribution of the random variables w'_j in (3.8).

2. Choose an integer N_ε such that N_ε is a multiple of n and moreover

$$\frac{2LT}{N_\varepsilon} < \frac{\varepsilon}{2}. \qquad (3.11)$$

Then choose $\delta > 0$ small enough so that

$$N_\varepsilon \delta < \varepsilon. \qquad (3.12)$$

For any unit vector w, call μ^w the measure with density $G(\cdot, \beta' - \beta, w)$ w.r.t. to the uniform surface measure σ. By the law of large numbers, there exists N large enough so that $T/N < \varepsilon$ and moreover the following holds. If $w'_1, \ldots, w'_N$ are independent variables, distributed on the unit sphere according to the same probability measure μ^w, then

$$\text{Prob.}\left\{\left|\frac{1}{N}\sum_{i=1}^{N} f^{w'_i}(x) - \int f^{w'}(x)d\mu^w\right| \geq \frac{\varepsilon}{2}\right\} < \delta, \qquad (3.13)$$

Notice that, by a compactness argument, the integer N can be chosen so that (3.13) remains uniformly valid for every $w \in \partial B_1$ and every $x \in B_T$.

3. Choosing $m = N_\varepsilon N$, we claim that (3.9) holds. Indeed, consider the times

$$\tau_k \doteq k\frac{T}{N_\varepsilon}, \qquad k = 0, 1, \ldots, N_\varepsilon.$$

By induction, assume that for some $k = 0, \ldots, N_\varepsilon - 1$ one has

$$|x(\tau_k) - z(\tau_k)| \leq \eta(\tau_k). \qquad (3.14)$$

Then
$$|x(\tau_{k+1}) - z(\tau_{k+1})|$$

$$\leq |x(\tau_k) - z(\tau_k)| + (\tau_{k+1} - \tau_k)\left|\frac{1}{N}\sum_{i=1}^{N} f^{w'_i}(x(\tau_k)) - \int f^{w'}(x(\tau_k))d\mu^w\right|$$
$$+(\tau_{k+1} - \tau_k)\left|\int f^{w'}(x(\tau_k))d\mu^w - \int f^{w'}(z(\tau_k))d\mu^w\right| + 2L(\tau_{k+1} - \tau_k)^2$$

$$\leq \eta(\tau_k) + \frac{T}{N_\varepsilon}\left|\frac{1}{N}\sum_{i=1}^{N} f^{w'_i}(x(\tau_k)) - \int f^{w'}(x(\tau_k))d\mu^w\right| + \frac{T}{N_\varepsilon}L\left[\eta(\tau_k) + \frac{2T}{N_\varepsilon}\right].$$

By (3.13) and (3.11), with probability $\geq 1 - \delta$ we thus have

$$|x(\tau_{k+1}) - z(\tau_{k+1})| \leq \eta(\tau_k) + \frac{T}{N_\varepsilon}\frac{\varepsilon}{2} + \frac{T}{N_\varepsilon}\left[L\eta(\tau_k) + \frac{\varepsilon}{2}\right] \leq \eta(\tau_{k+1}). \quad (3.15)$$

Since $x(0) = z(0) = 0$, with probability $\geq 1 - N_\varepsilon\delta$ we thus have

$$|x(\tau_k) - z(\tau_k)| \leq \eta(\tau_k) \qquad \text{for all } k = 0, 1, \ldots, N_\varepsilon. \quad (3.16)$$

Recalling that $|\dot{x}| \leq 1$, $\dot{z} \leq 1$ and $2T/N_\varepsilon < \varepsilon$, from (3.16) it follows

$$\text{Prob. }\{|x(t) - z(t)| > \eta(t) + \varepsilon \quad \text{for some } t \in [0, T]\}$$

$$\leq \text{ Prob. }\{|x(\tau_k) - z(\tau_k)| > \eta(\tau_k) \text{ for some } k = 1, \ldots, N_\varepsilon\} \quad (3.17)$$

$$\leq N_\varepsilon\delta \leq \varepsilon.$$

This proves (3.10), and hence the lemma. $\qquad\square$

The next result is an integral version of the upper semicontinuity property of the functional ϕ introduced at (2.7).

Lemma 6. *Let (Ω, μ) be a probability space and consider a measurable map $\Lambda : \omega \mapsto z^\omega(\cdot)$ from Ω into $\mathcal{F}$. Then the push-forward of the measure μ by the map Λ yields a probability distribution on the set $\mathcal{F}$ of solutions to (1.1)-(1.2). Assume that*

$$E^\mu\left[\int_0^T \phi(\dot{z}(t), F(z(t)))\,dt\right] < \varepsilon. \quad (3.18)$$

Then there exists $\delta > 0$ such that the following holds. If $\widetilde{\Lambda} : \omega \mapsto x^\omega(\cdot)$ is another map from Ω into $\mathcal{F}$ such that

$$E^\mu\left[\sup_{t \in [0,T]} |x^\omega(t) - z^\omega(t)|\right] \leq \delta, \quad (3.19)$$

then

$$E^\mu\left[\int_0^T \phi(\dot{x}^\omega(t), F(x^\omega(t)))\,dt\right] < 2\varepsilon. \quad (3.20)$$

Proof. By assumption, the functional

$$y(\cdot) \mapsto \Phi(y) \doteq \int_0^T \phi(\dot{y}(t), F(y(t)))\,dt$$

introduced at (2.9) is upper semicontinuous on the compact set of trajectories $\mathcal{F} \subset \mathcal{C}^0([0,T]; \mathbb{R}^n)$. Hence there exists a decreasing sequence of continuous functionals $(\Phi_\ell)_\ell \geq 1$ such that

$$\Phi(y) = \inf_{\ell \geq 1} \Phi_\ell(y),$$

for all $y \in \mathcal{F}$.

The assumption (3.18) implies that, for some k,

$$E^\mu \left[\Phi_k(z) \right] < \varepsilon. \tag{3.21}$$

Since the functional Φ_k is continuous and uniformly bounded on $\mathcal{F}$, we can find $\delta_k > 0$ and M such that

$$\|y - z\|_{\mathcal{C}^0} \leq \delta_k \implies |\Phi_k(y) - \Phi_k(z)| < \frac{\varepsilon}{2}, \tag{3.22}$$

$$\Phi_k(y) \leq M,$$

for all $y, z \in \mathcal{F}$. Choose $\delta_k > 0$ as so that (3.22) holds and assume that $\|x - z\|_{\mathcal{C}^0} \leq \delta_k$. Then by (3.21) we obtain

$$
\begin{aligned}
&E^\mu \left[\int_0^T \phi(\dot{x}(t), F(x(t)))\, dt \right] \\
&= E^\mu \left[\Phi(x) \right] \\
&\leq E^\mu \left[\Phi_k(x) \right] \leq \int \Phi_k(z^\omega)\, d\mu(\omega) + \int |\Phi_k(x^\omega) - \Phi_k(z^\omega)|\, d\mu(\omega) \\
&< \varepsilon + \tfrac{\varepsilon}{2} + 2M \cdot \mu \left(\{\omega \in \Omega;\ \|x^\omega - z^\omega\|_{\mathcal{C}^0([0,T])} > \delta_k \} \right).
\end{aligned}
\tag{3.23}
$$

We now observe that

$$E^\mu \left[\sup_{t \in [0,T]} |x^\omega(t) - z^\omega(t)| \right] \geq \delta_k \cdot \mu \left(\{\omega \in \Omega;\ \|x^\omega - z^\omega\|_{\mathcal{C}^0([0,T])} > \delta_k \} \right).$$

Hence, if in (3.19) we choose $\delta < \varepsilon \delta_k / 4M$, then the right hand side of (3.23) is smaller than 2ε. This yields (3.20). $\qquad \square$

4. Proof of Theorem 1

To prove Theorem 1 we need to show that, by suitably choosing the integers n_k and the numbers β_k, the construction described in Section 2 yields in the limit a probability distribution on $\mathcal{F}$ satisfying (P1)-(P2).

1. Assume that the integers $n_1 < \cdots < n_{k-1}$ and the positive constants $\beta_1 > \beta_2 > \cdots > \beta_{k-1}$ have already been chosen, satisfying

$$E \left[\int_0^T \phi\left(\dot{x}_{k-1}(t),\ F(x_{k-1}(t))\right) dt \right] \leq 2^{1-k}, \tag{4.1}$$

Using Lemma 6 with z replaced by x_{k-1}, we can find $\delta_k > 0$ with $\delta_k \leq \delta_{k-1}/2$ such that

$$E \left[\sup_{t \in [0,T]} |x(t) - x_{k-1}(t)| \right] \leq 2\delta_k \tag{4.2}$$

implies

$$E\left[\int_0^T \phi(\dot{x}(t), F(x(t)))\,dt\right] \leq 2^{2-k}. \tag{4.3}$$

Using Lemma 4 and Lemma 5, we can choose $\beta_k \in]0, \beta_{k-1}[$ and an integer n_k, with n_k/n_{k-1} integer, such that the following holds. Calling $x_k(\cdot)$ the solution to the Cauchy problem with random right hand side (2.16), one has

$$E\left[\int_0^T \phi(\dot{x}_k(t), \ F(x_k(t)))\,dt\right] \leq 2^{-k}, \tag{4.4}$$

$$E\left[\sup_{t\in[0,T]} |x_k(t) - x_{k-1}(t)|\right] \leq \delta_k, \tag{4.5}$$

2. It remains to show that with the above choices the properties (P1)-(P2) are satisfied. To prove (P1) we observe that, by (4.5) the sequence of random variables $x_k = \Lambda_k(w_k) \in \mathcal{F}$ is Cauchy, and converges to a random variable $x \in \mathcal{F}$ with probability one. Moreover, this limit x satisfies (4.2) for every k. Hence (4.3) holds. Since k is arbitrary, this shows that $w \in \mathcal{F}^{ext}$ with probability one. This proves (P1).

3. Toward a proof of (P2), let $\mathcal{U} \subset \mathcal{C}([0,T]; \mathbb{R}^n)$ be an open set containing some solution of the differential inclusion, say $z(\cdot) \in \mathcal{F}$. Since $\mathcal{U}$ is open, there exist a radius $\rho > 0$ such that

$$\sup_{t\in[0,T]} |x(t) - z(t)| \leq 2\rho \qquad \Longrightarrow \qquad x(\cdot) \in \mathcal{U}.$$

We claim that, for every k sufficiently large,

$$\eta_k \doteq \mathrm{Prob.}\left(\sup_{t\in[0,T]} |x_k(t) - z(t)| \leq \rho\right) > 0. \tag{4.6}$$

Indeed, for any $\beta > 0$ consider the compact convex sets

$$F^\beta(x) \doteq \overline{co}\left\{f^{\beta,w}(x); \ w \in \partial B_1\right\} \subseteq F(x).$$

As $\beta \to 0$ one has the convergence

$$F^\beta(x) \ \to \ F(x)$$

in the Hausdorff distance, uniformly for $x \in B_T$. We regard the equation

$$\dot{x}(t) \ = \ f^{\beta,w(t)}(x(t)) \tag{4.7}$$

as a control system, where $t \mapsto w(t) \in \partial B_1$ plays the role of the control function. By standard approximation results, we can choose $\beta^\sharp > 0$ sufficiently small and a piecewise constant control $t \mapsto w^\sharp(t)$ such that the solution $x^\sharp(\cdot)$ of

$$\dot{x}(t) \ = \ f^{\beta^\sharp, w^\sharp(t)}(x(t)), \qquad x(0) = 0, \tag{4.8}$$

satisfies

$$\sup_{t\in[0,T]} |x^\sharp(t) - z(t)| < \frac{\rho}{3}. \tag{4.9}$$

In turn, for every k sufficiently large, by a standard approximation result [9] there exists a piecewise constant control $\tilde{w} : [0, T] \mapsto \partial B_1$, say

$$\tilde{w}(t) = \tilde{w}_\ell \qquad t \in I_{k,\ell}, \ \ell = 1, \ldots, n_k,$$

with $I_{k,\ell}$ as in (2.12), such that the following holds. The solution $\tilde{x}(\cdot)$ of

$$\dot{x}(t) = f^{\beta_k, \tilde{w}(t)}(x(t)), \qquad x(0) = 0,$$

satisfies

$$\sup_{t\in[0,T]} \left|\tilde{x}(t) - x^\sharp(t)\right| < \frac{\rho}{3}. \tag{4.10}$$

By continuity, there exists $\epsilon > 0$ such that, if $w(\cdot)$ is any other piecewise constant control such that

$$w(t) = w_\ell \qquad t \in I_{k,\ell},$$

and $|w_\ell - \tilde{w}_\ell| < \epsilon$ for every ℓ, then the corresponding solution $x(\cdot)$ satisfies

$$\sup_{t\in[0,T]} |x(t) - \tilde{x}(t)| < \frac{\rho}{3}. \tag{4.11}$$

Observing that for every $\epsilon > 0$ one has

$$\text{Prob.}\Big\{|w_\ell - \tilde{w}_\ell| < \epsilon \qquad \text{for every } \ell = 1, \ldots, n_k\Big\} > 0,$$

combining the three inequalities (4.9)–(4.11) we prove the claim (4.6).

4. We now recall that, given the random variables $w_{k,1}, \ldots, w_{k,n_k} \in \partial B_1$, the random trajectory $x_k(\cdot)$ is uniquely determined by solving the Cauchy problem (2.16). For every given values of the random variables $w_{k,\ell}$, our construction yields an estimate on the conditional expectation:

$$E\left[\|x(\cdot) - x_k(\cdot)\|_{\mathcal{C}^0}\right] \leq 2^{1-k}.$$

In particular, we have

$$E\left[\|x(\cdot) - x_k(\cdot)\|_{\mathcal{C}^0} \ \Big| \ |w_{k,\ell} - \tilde{w}_\ell| < \epsilon \text{ for every } \ell = 1, \ldots, n_k\right] \leq 2^{1-k}. \tag{4.12}$$

Choosing k large enough so that $2^{1-k} < \rho/2$, we obtain a corresponding estimate on the conditional probability:

$$\text{Prob.}\left(\|x(\cdot) - x_k(\cdot)\|_{\mathcal{C}^0} \leq \rho \ \Big| \ |w_{k,\ell} - \tilde{w}_\ell| < \epsilon \text{ for every } \ell = 1, \ldots, n_k\right) > \frac{1}{2}.$$

In turn, this implies

$$\text{Prob.}\{x(\cdot) \in \mathcal{U}\} \geq \text{Prob.}\left(\|x(t) - z(t)\|_{\mathcal{C}^0} \leq 2\rho\right)$$

$$\geq \text{Prob.}\left(|w_{k,\ell} - \tilde{w}_\ell| < \epsilon \text{ for every } \ell = 1, \ldots, n_k\right)$$
$$\times \text{Prob.}\left(\|x(\cdot) - x_k(\cdot)\|_{\mathcal{C}^0} \leq \rho \ \Big| \ |w_{k,\ell} - \tilde{w}_\ell| < \epsilon \text{ for every } \ell = 1, \ldots, n_k\right)$$

$$> 0,$$

completing the proof. □

References

[1] J. P. Aubin and A. Cellina, *Differential inclusions. Set-valued maps and viability theory.* Springer-Verlag, Berlin, 1984.

[2] A. Bressan, On a bang-bang principle for nonlinear systems. *Boll. Un. Mat. Italiana, Suppl.* (1980), 53-59.

[3] A. Bressan, The most likely path of a differential inclusion. *J. Differential Equations* **88** (1990), 155-174.

[4] A. Bressan, Extremal solutions to differential inclusions via Baire category: a dual approach. *J. Differential Equations* **255** (2013), 2392–2399.

[5] A. Bressan and G. Colombo, Generalized Baire category and differential inclusions in Banach spaces. *J. Differential Equations* **76** (1988), 135-158.

[6] A. Bressan and G. Crasta, Extremal selections of multifunctions generating a continuous flow. *Annal. Polon. Mat.* **LX.2** (1994), 101-117.

[7] A. Bressan and F. Flores, On total differential inclusions. *Rend. Sem. Mat. Univ. Padova* **92** (1994), 9-16.

[8] A. Bressan and B. Piccoli, A Baire category approach to the bang-bang property, *J. Differential Equations* **116** (1995), 318-337.

[9] A. Bressan and B. Piccoli,*Introduction to the Mathematical Theory of Control*, AIMS Series in Applied Mathematics, Springfield Mo. 2007.

[10] A. Cellina, On the differential inclusion $x' \in [-1, 1]$, *Atti Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* **69** (1980), 1-6.

[11] G. Colombo and V. Goncharov, Brownian motion and exposed solutions of differential inclusions. *Nonlin. Diff. Equat. Appl.*, **20** (2013), 323–343.

[12] B. Dacorogna and P. Marcellini, General existence theorems for Hamilton-Jacobi equations in the scalar and vectorial cases. *Acta Mathematica* **178** (1997), 1-37.

[13] B. Dacorogna and P. Marcellini, Cauchy-Dirichlet problem for first order nonlinear systems. *J. Funct. Analysis* **152** (1998), 404-446.

[14] F. S. De Blasi and G. Pianigiani, A Baire category approach to the existence of solutions of multivalued differential equations in Banach spaces. *Funkcial. Ekvac.* **25** (1982), 153–162.

[15] F. S. De Blasi and G. Pianigiani, Differential inclusions in Banach spaces. *J. Differential Equations* **66** (1987), no. 2, 208–229.

[16] C. De Lellis and L. Székelyhidi, The Euler equations as a differential inclusion. *Ann. Math.* **170** (2009), 1417–1436.

[17] D. Dentcheva, Regular Castaing representations of multifunctions with appli-
cations to stochastic programming. *SIAM J. Optim.* **10** (2000), 732Ã749.

[18] A. F. Filippov, The existence of solutions of multivalued differential equations.
Math. Notes **10** (1971), 608–611.

[19] L. Koralov and Y. Sinai, *Theory of Probability and Random Processes,* Second
edition. Universitext. Springer, Berlin, 2007.

[20] J. C. Oxtoby, *Measure and category. A survey of the analogies between topo-
logical and measure spaces.* Springer-Verlag, New York-Berlin, 1971.

[21] G. Pianigiani, Differential inclusions: the Baire category method, in *Methods of
Non-convex Analysis,* A. Cellina ed., Springer Lecture Notes in Mathematics,
Volume 1446 (1990), pp.104–136 .

[22] A. Plis, Trajectories and quasitrajectories of an orientor field. *Bull. Acad.
Polon. Sci. Sér. Sci. Math. Astronom. Phys.* **11** (1963), 369-370.

[23] R. T. Rockafellar, *Convex Analysis*, Princeton University Press, Princeton
(1970).

[24] A. Tolstonogov, *Differential Inclusions in a Banach Space*, Kluwer, Dordrecht,
2000.

Alberto Bressan
Department of Mathematics
Penn State University
University Park
Pa. 16802, U.S.A.
e-mail: `bressan@math.psu.edu`

Vasile Staicu
CIDMA and Department of Mathematics
University of Aveiro
3810-193, Aveiro
Portugal
e-mail: `vasile@ua.pt`