

Chapter 2

Set-Valued Mappings and Differential Inclusions

This chapter deals with two fundamental concepts for the control systems described by differential inclusions. They are set-valued mappings and differential inclusions. The first two sections introduce the set-valued mappings and the succeeding four sections involve the differential inclusions.

2.1 Set-Valued Mappings

The set-valued mapping is an important foundation in the theory of differential inclusions. From the word of “set-valued”, it is obvious that the mapping maps a variable to a set. Hence, the image of one variable may have more than one elements. In order to distinguish from the set-valued mapping, the mapping defined in Chap. 1, where the image of a variable only has one element, is often called by single-valued mapping in this chapter. This section will give the definition and elemental properties of set-valued mappings and extend the concepts of continuity and derivative to set-valued mappings.

2.1.1 Definition of the Set-Valued Mappings

X and Y are supposed to be two normed spaces. $\|\cdot\|_X$ and $\|\cdot\|_Y$ are the norms of X and Y , respectively. When there is no confusion, we often omit their subscripts.

Definition 2.1.1 $A \subset X$ is a set. A mapping $F : A \rightarrow Y$ is said to be a set-valued mapping defined on A , if for every $x \in A$, $F(x)$ is a subset of Y .

The set A is called by the domain of F . The set of $\{x; F(x) \neq \emptyset\}$ is called by the effective domain of F and is denoted by $\text{dom } F$. $F(x)$ is called by the image of x , the set $\bigcup_{x \in A} F(x) \subset Y$ is the range of F , and $\text{gra } F = \{(x, y); y \in F(x)\}$ is called by the graph of F . $\square$

From the definition, it seems to be better to define the set-valued mapping from A to the power set of Y , i.e., $F : A \rightarrow \mathcal{P}Y$. The definition can meet the requirement of single-valued mapping. The reason we have considered is that Y is a normed space, and we need to apply the norm of Y and topology induced by the norm.

From now on, capital letters, such as $F, G, \dots$ are used to express the set-valued mappings; and correspondingly, the lowercases, such as $f, g, \dots$, are to express the single-valued mappings. We assume in this book that $\text{dom } F \neq \emptyset$ for every set-valued mapping F .

We now present more remarks for the set-valued mappings.

Remark 1 If the domain of set-valued mapping $F : A \rightarrow Y$ is not equal to X , i.e., $A \neq X$, then we can extend the definition of F to $\bar{F}$ by

$$\bar{F}(x) = \begin{cases} F(x) & x \in A, \\ \emptyset & x \notin A. \end{cases}$$

By the alteration, the domain of $\bar{F}$ becomes X , but its effective domain of is exactly equal to that of F . Hence we always assert that every set-valued mapping is defined on whole space X . $\square$

Remark 2 For $y \in Y$, the inverse of y is defined as $\{x; y \in F(x)\}$ and is denoted by $F^{-1}(y)$. It is obvious that $F^{-1}(y)$ is a set of X . Let $G \subset Y$ be a set. There are two ways to define the inverse of G for a set-valued mapping F . The set $\{x; F(x) \subset G\} \subset X$ is defined as the strong inverse of G and denoted by $F^{-1}(G)(s)$ where the letter “s” means **strong**. The set $\{x; F(x) \cap G \neq \emptyset\} \subset X$ is called by the weak inverse of G and denoted by $F^{-1}(G)(w)$ where “w” stands for **weak**. It is obvious that $F^{-1}(G)(w) \supset F^{-1}(G)(s)$. At the most time, we apply the definition of weak inverse, hence, $F^{-1}(G)(w)$ is simplified as $F^{-1}(G)$. But we keep the notation of $F^{-1}(G)(s)$. $\square$

For these two kinds of inverses, we have the following equations and their proofs are left to readers as exercises.

$$F^{-1}(G^c)(w) = [F^{-1}(G)(s)]^c, \quad (2.1.1a)$$

$$F^{-1}(G^c)(s) = [F^{-1}(G)(w)]^c. \quad (2.1.1b)$$

In Eqs. (2.1.1a) and (2.1.1b), the complements at two sides are in different spaces. At the left side, $G^c = Y \setminus G$, but at the right side $[F^{-1}(\cdot)]^c = X \setminus F^{-1}(\cdot)$.

Remark 3 A set-valued mapping $F : X \rightarrow Y$ is with closed value if for every $x \in X$, $F(x)$ is a closed set of Y . If $\text{gra } F$ is a closed set of $X \times Y$, then $F(x)$ is called as a closed set-valued mapping or closed mapping for simplicity.

Replacing the word “closed” by “bounded”, we can establish the definitions of “set-valued mapping with bounded value” and “bounded set-valued mapping”. If it is replaced by “open” and “compact”, we have the corresponding statements and the details are omitted. $\square$

In fact, we have already applied the definition of set-valued mapping, for example, the epigraph $\text{epi } f$ is a graph of set-valued mapping, where the set-valued mapping is $F(x) = \{a; a \in \mathbb{R}, a \geq f(x)\}$. The another example is subdifferential $\partial f(x)$, it is also a set-valued mapping, although for some x it only has one element.

2.1.2 Continuities of Set-Valued Mappings

Giving definitions of continuities for set-valued mapping is a precise job. There are more than one kinds of continuities. Readers have to distinct and compare these statements, carefully.

Definition 2.1.2 Let $F : X \rightarrow Y$ be a set-valued mapping and $x_0 \in X$. F is upper semi-continuous at x_0 if for every open set $O_Y \subset Y$ which satisfies that $O_Y \supset F(x_0)$, there is a $\delta > 0$ such that $F(B(x_0, \delta)) \subset O_Y$. If F is upper semi-continuous at every point $x \in X$, then F is an upper semi-continuous set-valued mapping on X .

F is lower semi-continuous at x_0 if for every $y \in F(x_0)$ and every $\varepsilon > 0$, there is a $\delta > 0$ which may depend on y and ε such that for every $x \in B(x_0, \delta)$, $F(x) \cap B(y, \varepsilon) \neq \emptyset$. If F is lower semi-continuous at every point $x \in X$, then F is a lower semi-continuous set-valued mapping on X .

If F is both upper semi-continuous and lower semi-continuous at x_0 , then F is continuous at x_0 . If F is continuous at every point $x \in X$, then F is a continuous set-valued mapping on X . $\square$

Note that in Definition 2.1.2, the words “the space X ” can be replaced by “the set A ”. For example, we can say that F is an upper semi-continuous set-valued mapping on $A (\subset X)$.

For the upper semi-continuity and lower semi-continuity, we have the following equivalent statements.

Theorem 2.1.1 Let $F : X \rightarrow Y$ be a set-valued mapping. Then the following statements are equivalent.

- (1) F is upper semi-continuous.
- (2) For every open set $O_Y \subset Y$, $F^{-1}(O_Y)$ is an open set of X .
- (3) For every closed set $C_Y \subset Y$, $F^{-1}(C_Y)$ is a closed set of X .
- (4) Let $x \in X$ and $\{x_n\} \subset X$ be a sequence with $x_n \rightarrow x$. Then for every open set $O_Y \subset Y$ and $F(x) \subset O_Y$, there is an $N \in \mathbb{N}$, $F(x_n) \subset O_Y$ provided that $n > N$.

Proof (1) $\Rightarrow$ (2). Suppose that $O_Y \subset Y$ is an open set. If $x \in F^{-1}(O_Y)(s)$, then $F(x) \subset O_Y$. F is upper semi-continuous, by Definition 2.1.2, there is a neighborhood $B(x, \delta)$ such that $F(B(x, \delta)) \subset O_Y$, i.e., $B(x, \delta) \subset F^{-1}(O_Y)(s)$.

(2) $\Leftrightarrow$ (3). This is a direct result of Equations (2.1.1a) and (2.1.1b). The details are omitted.

(2) $\Rightarrow$ (4). Let $\{x_n\} \subset X$ be a sequence with $x_n \rightarrow x$. Let $O_Y \subset Y$ be an open set with $F(x) \subset O_Y$. Then by the conclusion (2), there is a $\delta > 0$, such that $B(x, \delta) \subset F^{-1}(O_Y)(s)$. Hence, there is an $N \in \mathbb{N}$, when $n > N$, $x_n \in B(x, \delta)$, i.e., $x_n \in F^{-1}(O_Y)(s)$, $F(x_n) \subset O_Y$.

(4) $\Rightarrow$ (1). If F is not upper semi-continuous at $x \in X$, then there is an open set $O_Y \subset Y$ such that $F(B(x, \delta)) \subset O_Y$ is not true. Let $\delta_n \downarrow 0$, there are an $x_n \in B(x, \delta_n)$, and a $y_n \in F(x_n)$, $y_n \notin O_Y$ for $n \in \mathbb{N}$. Obviously, $x_n \rightarrow x$. $y_n \in F(x_n)$, $y_n \notin O_Y$ for $n \in \mathbb{N}$. It contradicts to the conclusion (4). $\square$

For the lower semi-continuity, we have the following corresponding conclusions.

Theorem 2.1.2 Let $F : X \rightarrow Y$ be a set-valued mapping. Then the following statements are equivalent.

- (1) F is lower semi-continuous.
- (2) For every open set $O_Y \subset Y$, $F^{-1}(O_Y)(w)$ is an open set of X .
- (3) For every closed set $C_Y \subset Y$, $F^{-1}(C_Y)(s)$ is a closed set of X .
- (4) Let $x \in X$ and $\{x_n\} \subset X$ be a sequence with $x_n \rightarrow x$. Then for every open set $O_Y \subset Y$, $F(x) \cap O_Y \neq \emptyset$, there is an $N \in \mathbb{N}$, $F(x_n) \cap O_Y \neq \emptyset$ provided that $n > N$. $\square$

The proof of Theorem 2.1.2 is quite similar to that of Theorem 2.1.1 and is left to readers as an exercise.

There is an alternative way to define the continuity for the set-valued mappings. This, we call, is the ε -continuity.

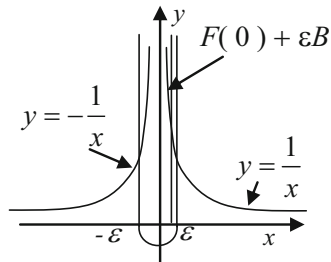
Definition 2.1.3 A set-valued mapping $F : X \rightarrow Y$ is ε -upper semi-continuous at $x_0 \in X$, if for every $\varepsilon > 0$, there is a $\delta > 0$ which may depend on ε , such that for every $x \in B(x_0, \delta)$, $F(x) \subset F(x_0) + \varepsilon B$. If for every $x \in X$, F is ε -upper semi-continuous at x , then F is an ε -upper semi-continuous mapping on X .

A set-valued mapping $F : X \rightarrow Y$ is ε -lower semi-continuous at $x_0 \in X$; if for every $\varepsilon > 0$, there is a $\delta > 0$ which may depend on ε , such that for every $x \in B(x_0, \delta)$, $F(x_0) \subset F(x) + \varepsilon B$. If for every $x \in X$, F is ε -lower semi-continuous at x , then F is an ε -lower semi-continuous mapping on X . $\square$

It is easy to see that the relation $F(x) \subset F(x_0) + \varepsilon B$ for every $x \in B(x_0, \delta)$ is equivalent to $F(B(x_0, \delta)) \subset B(F(x_0), \varepsilon)$, and $F(x_0) \subset F(x) + \varepsilon B$ for every $x \in B(x_0, \delta)$ is equivalent to $F(x_0) \subset \bigcap_{x \in B(x_0, \delta)} B(F(x), \varepsilon)$.

From $F(B(x_0, \delta)) \subset B(F(x_0), \varepsilon)$, it can be seen that if F is upper semi-continuous, then it is ε -upper semi-continuous, since $B(F(x_0), \varepsilon)$ is an open set. But the inverse statement may fail to be true. A counter example will be given later.

Fig. 2.1 The graph of $F(x)$ defined in Example 2.1.1



For the ε - lower semi-continuous, the situation is different. The inclusion $F(x_0) \subset F(x) + \varepsilon B$ implies that for every $y_0 \in F(x_0)$, there is a $y \in F(x)$ such that $\|y - y_0\| < \varepsilon$, i.e., $F(x) \cap B(y_0, \varepsilon) \neq \emptyset$. On the other hand, if $F(x) \cap B(y_0, \varepsilon) \neq \emptyset$, then there is a $y \in F(x)$ such that $\|y - y_0\| < \varepsilon$, i.e., $F(x_0) \subset F(x) + \varepsilon B$. We conclude that the lower semi-continuity is equivalent to the ε - lower semi-continuity.

Example 2.1.1 Consider $F : \mathbb{R} \rightarrow \mathbb{R}^2$ which maps every $x \in \mathbb{R}$ to the set $F(x) = \{(y_1, y_2) ; y_1 = x, y_2 \in [0, \infty)\}$, or, $F(x) = \{(x, y) ; x \in \mathbb{R}, y \in [0, \infty)\}$ for simplicity.

In the plane of $\mathbb{R}^2$, for every x the image of $F(x)$ is a radical line which starts at $(x, 0)$ and goes to positive infinite and is orthogonal to the x -axis (Fig. 2.1). It follows that $\cup F(x)$ is the upper half plane of $\mathbb{R}^2$.

We assert that F is not upper semi-continuous at every $x \in \mathbb{R}$. As an example, let us consider $F(0)$ and an open set $M = \{(x, y) ; |xy| < 1\} \subset \mathbb{R}^2$. It is obvious that $M \supset F(0)$. However, for every x , $x \neq 0$, M cannot contain $F(x)$. It follows that for every $\delta > 0$, the relation $M \supset F(B(0, \delta))$ is not true, i.e., $F(x)$ is not upper semi-continuous at $x = 0$.

We can prove that for every $x \in \mathbb{R}$, F is lower semi-continuous at x . It is because that for every $(x, y_0) \in F(x)$ and every $\varepsilon > 0$, we can take $\delta = \varepsilon$, then $F(B(x, \delta)) \cap B((x, y_0), \varepsilon) = B((x, y_0), \varepsilon)$.

By Definition 2.1.3, the mapping $F(x)$ is ε - upper semi-continuous. Because for every $\varepsilon > 0$, we can select $\delta = \frac{\varepsilon}{2}$, it follows that $F(x) = \{(x, y) ; y \in [0, \infty)\} \subset F(x_0) + \varepsilon B$ $\square$

In Example 2.1.1, the open set M cannot be written as the form of $F(0) + \varepsilon B$. This is the key issue for the failure of $F(x)$ to be an upper semi-continuous mapping.

In Remark 3 after Definition 2.1.1, we have defined a set-valued mapping with closed value and closed set-valued mapping. By these definitions, the following conclusions are obvious.

- (1) If $F : X \rightarrow Y$ is a set-valued mapping with open value, then it is an open set-valued mapping.
- (2) If $F : X \rightarrow Y$ is a closed set-valued mapping, then it is with closed value.

Readers can find the inverse instatements of the above conclusion are not valid. By using the concepts of continuities, we have the following conclusion.

Theorem 2.1.3 If $F : X \rightarrow Y$ is an upper semi-continuous set-valued mapping, then it is closed if and only if it is with closed value.

Proof By the above statement (2), it is sufficient to show if $F : X \rightarrow Y$ is with closed value then it is a closed mapping.

Suppose $\{(x_n, y_n), n = 1, 2, \dots\} \subset \text{gra } F$ such that $(x_n, y_n) \rightarrow (x_0, y_0)$. We are required to verify $(x_0, y_0) \in \text{gra } F$. By the condition of Theorem 2.1.1, F is upper semi-continuous; hence, it is ε -upper semi-continuous. For every $\varepsilon > 0$, there exists a $\delta > 0$, such that $F(x) \subset F(x_0) + \varepsilon B$ for every $x \in B(x_0, \delta)$. Because $x_n \rightarrow x_0$, there is an $N \in \mathbb{N}$, when $n > N$, $x_n \in B(x_0, \delta)$. It follows that $y_n \in F(x_n) \subset F(x_0) + \varepsilon B$ where $\varepsilon > 0$ can be selected arbitrary. Consequently, $y_0 \in \text{cl } F(x_0)$ since $y_n \rightarrow y_0$. F is with closed value, hence $F(x_0)$ is a closed set, $F(x_0) = \text{cl } F(x_0)$. Thus, $(x_0, y_0) \in \text{gra } F$ $\square$

Readers are suggested to state a similar conclusion for lower semi-continuous set-valued mappings.

Before giving Theorem 2.1.4, we prove a general result which can be treated as an extension of separation axiom in Topology.

Lemma 2.1.1 Let Y be a Banach space, $A \subset Y$ is a compact set, and $M \supset A$ is an open set. Then there is an $\varepsilon > 0$ such that $M \supset A + \varepsilon B$.

Proof M is an open set, hence its complement M^c is a closed set. $M^c \cap A = \emptyset$. Now consider the distance between M^c and A , $d(M^c, A)$. We conclude that $d(M^c, A) > 0$ since A is compact and M^c is closed. Let $\varepsilon = \frac{1}{2}d(M^c, A)$. $M^c \cap (A + \varepsilon B) = \emptyset$, i.e., $M \supset A + \varepsilon B$. $\square$

Theorem 2.1.4 If the set-valued mapping $F : X \rightarrow Y$ is with compact value, and Y is complete, then F is upper semi-continuous if and only if it is ε -upper semi-continuous.

Proof It is sufficient to show that under the conditions of Theorem 2.1.4, an ε -upper semi-continuous set-valued mapping is upper semi-continuous.

By the condition given by the theorem, for every $x \in X$, $F(x)$ is a compact set. Now let M be an open set which contains $F(x)$. Then by Lemma 2.1.1, there is an $\varepsilon > 0$ such that $F(x) + \varepsilon B \subset M$. $F(x)$ is ε -upper semi-continuous, hence there is a $\delta > 0$, such that for every $x_1 \in B(x, \delta)$, $F(x_1) \subset F(x) + \varepsilon B$. It follows $F(B(x, \delta)) \subset F(x) + \varepsilon B \subset M$. $\square$

In Example 2.1.1, $F(x)$ is ε -upper semi-continuous, but for every $x \in \mathbb{R}$, $F(x)$ is not a compact set. Hence, it fails to be upper semi-continuous.

We now turn to the operation of two mappings.

For two set-valued mappings $F(x)$ and $G(x)$, because their images are in a normed space, we can define linear operation $\alpha F(x) + \beta G(x)$. But at the most time, we prefer to define $F(x) \cup G(x)$ and $F(x) \cap G(x)$.

Theorem 2.1.5

- (1) If $F(x)$ and $G(x)$ are upper semi-continuous, then $F(x) \cup G(x)$ is upper semi-continuous;
- (2) If $F(x)$ and $G(x)$ are lower semi-continuous, then $F(x) \cup G(x)$ is lower semi-continuous.

Proof The theorem is verified by using the conclusion 4 in Theorem 2.1.1 and Theorem 2.1.2, respectively. Let $O_Y \subset Y$ be an open set. Let $\{x_n\} \subset X$ be a sequence and $x_n \rightarrow x$.

(1) If $O_Y \supset F(x) \cup G(x)$, then $O_Y \supset F(x)$. It follows from the upper semi-continuity of $F(x)$, $O_Y \supset F(x_n)$ for some N_1 and $n > N_1$. Similarly, there is an N_2 such that for $n > N_2$, $O_Y \supset G(x_n)$. Thus, when $n > \max\{N_1, N_2\}$, $O_Y \supset F(x_n) \cup G(x_n)$. The first conclusion is verified.

(2) If $O_Y \cap (F(x) \cup G(x)) \neq \emptyset$, then without loss of generality, we assume that $O_Y \cap F(x) \neq \emptyset$. It follows from the lower semi-continuity, $O_Y \cap F(x_n) \neq \emptyset$ for some N_1 and $n > N_1$. Thus, $O_Y \cap (F(x_n) \cup G(x_n)) \neq \emptyset$ when $n > N_1$. The second conclusion is also proved.

The situation for the intersections is more complicated. We present them in the following two theorems.

Theorem 2.1.6 Let $F(x)$ and $G(x)$ be two set-valued mappings from X to Y . If they are all lower semi-continuous and satisfy that:

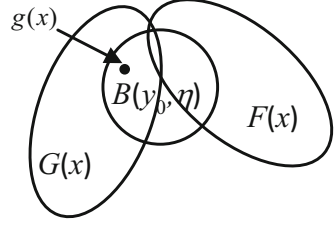
- (1) F and G have all convex values, i.e., $F(x)$ and $G(x)$ are convex sets for every $x \in X$.
- (2) $G(x)$ is locally bounded, i.e., for every $x_0 \in X$, there are $\delta > 0$ and $M > 0$ which may depend on x_0 such that if $x \in B(x_0, \delta)$ then $F(x) \subset MB_Y$ where B_Y is the open unit ball of Y .
- (3) There exists a $\gamma > 0$, such that for every $x \in X$ and every $z \in \gamma B_Y$, $F(x) \cap (G(x) + z) \neq \emptyset$.

Then $F(x) \cap G(x)$ is lower semi-continuous.

Proof Let $x_0 \in X$ and $y_0 \in F(x_0) \cap G(x_0)$. Because $G(x)$ is locally bounded, there are $\delta_1 > 0$ and $M > 0$ such that $G(x) \subset MB_Y$ for every $x \in B(x_0, \delta_1)$. For a given $\varepsilon > 0$, define $\eta = \frac{\varepsilon}{1+4M\gamma^{-1}}$. By the lower semi-continuity of $F(x)$ and $G(x)$, there exists a $\delta \leq \delta_1$, such that $F(x) \cap B(y_0, \eta) \neq \emptyset$, $G(x) \cap B(y_0, \eta) \neq \emptyset$ simultaneously for $x \in B(x_0, \delta)$.

Take a point $g(x) \in G(x) \cap B(y_0, \eta)$ for every $x \in B(x_0, \delta)$. The $g(x) \in B(y_0, \eta)$ and $F(x) \cap B(y_0, \eta) \neq \emptyset$, hence there is a $f(x) \in F(x)$ such that $\|g(x) - f(x)\| < 2\eta$, i.e., $g(x) \in F(x) + 2\eta z$ for some $z = z(x) \in B_Y$ (Fig. 2.2).

Fig. 2.2 The relation among $F(x)$, $G(x)$ and $B(y_0, \eta)$



By the third condition of Theorem 2.1.6, for this $z = z(x)$, there are $\bar{f}(x) \in F(x)$ and $\bar{g}(x) \in G(x)$, such that $\bar{g}(x) = \bar{f}(x) - \gamma z$. Now, we define $\lambda = \frac{\gamma}{\gamma+2\eta}$, it follows $1 - \lambda = \frac{2\eta}{\gamma+2\eta} = \frac{2\lambda\eta}{\gamma}$. At last

$$\begin{aligned} \lambda g(x) &\in \lambda F(x) + 2\lambda\eta z = \lambda F(x) + \frac{2\lambda\eta}{\gamma} (\bar{f}(x) - \bar{g}(x)) \\ &\subset \lambda F(x) + (1 - \lambda) F(x) - (1 - \lambda) \bar{g}(x). \end{aligned}$$

Because both $F(x)$ and $G(x)$ are convex, $\lambda g(x) + (1 - \lambda) \bar{g}(x) \in F(x) \cap G(x)$. On the other hand,

$$\begin{aligned} \|\lambda g(x) + (1 - \lambda) \bar{g}(x) - y_0\| &\leq \lambda \|g(x) - y_0\| + (1 - \lambda) \|\bar{g}(x) - y_0\| \\ &\leq \lambda\eta + (1 - \lambda) 2M \\ &= \lambda\eta + \frac{4M\lambda\eta}{\gamma} \\ &= \lambda\eta \left(1 + \frac{4M}{\gamma}\right) \\ &\leq \varepsilon, \end{aligned}$$

$\lambda g(x) + (1 - \lambda) \bar{g}(x) \in B(y_0, \varepsilon)$. Thus, $(F(x) \cap G(x)) \cap B(y_0, \varepsilon) \neq \emptyset$. □

This is a constructive proof, it gives an element in $(F(x) \cap G(x)) \cap B(y_0, \varepsilon)$. The next theorem is related to upper semi-continuous mappings.

Theorem 2.1.7 Let $F : X \rightarrow Y$ be an upper semi-continuous mapping. If the following conditions are satisfied:

- (1) F is with compact value.
- (2) $\text{gra } G$ is a closed set in $X \times Y$.

Then $F(x) \cap G(x)$ is upper semi-continuous on X .

Proof Let $O_Y \subset Y$ be an open set and $O_Y \supset F(x_0) \cap G(x_0)$ for $x_0 \in X$. Then the target of verification is to find a $\delta > 0$ such that $O_Y \supset F(x) \cap G(x)$ for every $x \in B(x_0, \delta)$.

If $O_Y \supset F(x_0)$, then by the upper semi-continuity of $F(x)$ at x_0 , there is a $\delta > 0$, such that $O_Y \supset F(x) \supset F(x) \cap G(x)$ for every $x \in B(x_0, \delta)$. The conclusion is proved. Hence it is sufficient to prove the case that $F(x_0) \not\subset O_Y$. If $F(x_0) \not\subset O_Y$, then $F(x_0) \cap O_Y^c \neq \emptyset$. Denote $K = F(x_0) \cap O_Y^c$, K is a compact set. We conclude that $K \cap G(x_0) = \emptyset$. Otherwise $G(x_0) \cap F(x_0) \cap O_Y^c \neq \emptyset$, it is contradicted to $O_Y \supset F(x_0) \cap G(x_0)$. Therefore, for every $y \in K$, $(x_0, y) \notin \text{gra } G$. $\text{gra } G$ is a closed set, there exist $\delta_y > 0$ and $\varepsilon_y > 0$ such that $(B(x_0, \delta_y) \times B(y, \varepsilon_y)) \cap \text{gra } G = \emptyset$. These $B(y, \varepsilon_y)$'s construct a open covering of K . There is a finite subcovering of K , i.e., $\bigcup_{i=1}^n B(y_i, \varepsilon_{y_i}) \supset K$. Denote $M = \bigcup_{i=1}^n B(y_i, \varepsilon_{y_i})$, M is an open set. Recall the construction of $B(y, \varepsilon_y)$, every ε_{y_i} corresponds to a δ_{y_i} . Let $\delta_1 = \min \{\delta_{y_i}; i = 1, 2, \dots, n\}$. For every $x \in B(x_0, \delta_1)$, $(\{x\} \times M) \cap \text{gra } G = \emptyset$, i.e., $M \cap G(x) = \emptyset$.

Because $M \supset K = F(x_0) \cap O_Y^c$, $O_Y \cup M \supset F(x_0)$. F is upper semi-continuous, for the open set $M \cup O_Y$, there is a δ_2 such that for every $x \in B(x_0, \delta_2)$, $O_Y \cup M \supset F(x)$. Let $\delta = \min(\delta_1, \delta_2)$. When $x \in B(x_0, \delta)$, it happens $O_Y \cup M \supset F(x)$ and $M \cap G(x) = \emptyset$. Thus,

$$\begin{aligned} F(x) \cap G(x) &\subset (M \cup O_Y) \cap G(x) \\ &= (M \cap G(x)) \cup (O_Y \cap G(x)) \\ &= O_Y \cap G(x) \subset O_Y. \end{aligned}$$

We have prove that $O_Y \supset F(x) \cap G(x)$ always holds in a neighborhood of x_0 . $\square$

In Sect. 1.4, we have mentioned that f is a single-valued upper semi-continuous function if and only if $-f$ is lower semi-continuous. But for set-valued mappings, there is no such a simple relation.

We now turn to deal with the Lipschitzian property of a set-valued mappings.

Definition 2.1.4 $F : X \rightarrow Y$ is a set-valued mapping and $x_0 \in X$. If there are two positive constants L and δ such that

$$F(x) \subset F(x_0) + L\|x - x_0\|_X B_Y \quad (2.1.2)$$

for every $x \in B(x_0, \delta)$ where B_Y is the open unit ball in Y . Then F is a locally Lipschitzian set-valued mapping at x_0 and L is its Lipschitzian constant.

F is said to be a global Lipschitzian set-valued mapping at x_0 , if the relation (2.1.2) is valid for all $x \in X$. Moreover, F is a (global) Lipschitzian set-valued mapping, if (2.1.2) is valid for any $x, x_0 \in X$. $\square$

If the relation (2.1.2) is satisfied for $F : X \rightarrow Y$, then for $y \in F(x)$, there exist $y_0 \in F(x_0)$ and $b \in B_Y$ such that $y = y_0 + L\|x - x_0\| b$, it is equivalent to

$$\|y - y_0\| = L\|x - x_0\| \|b\| < L\|x - x_0\|. \quad (2.1.3)$$

Conversely, if Inequality (2.1.3) holds for any two vectors $y \in F(x)$ and $y_0 \in F(x_0)$, then the relation (2.1.2) holds. Hence Inequality (2.1.3) is an equivalent condition for Lipschitzian set-valued mappings.

It is obvious that if $F : X \rightarrow Y$ is a local Lipschitzian set-valued mapping at $x_0 \in X$, then F is ε -upper semi-continuous at x_0 . The following example shows the Lipschitzian property of a set-valued mapping does not imply the upper semi-continuity. This is an important difference from the single-valued mapping.

Example 2.1.2 Let $F : \mathbb{R} \rightarrow \mathbb{R}^2$ be as follows: for $x_0 \in \mathbb{R}$, its image is

$$F(x_0) = \{(x, y) ; (x - x_0)^2 + y^2 < 1\}.$$

It is obvious for every $x_0 \in \mathbb{R}$, $F(x_0)$ is an open set. Hence, let $O_Y = F(x_0)$, then there is no $\delta > 0$ such that $F(B(x_0, \delta)) \subset O_Y$. However, F is a global Lipschitzian set-valued mapping, its Lipschitzian constant can be 1. $\square$

Theorem 2.1.8 Let $F : X \rightarrow Y$ be a Lipschitzian set-valued mapping, its Lipschitzian constant is L . Let $\bar{F}(x) : X \rightarrow Y$ be an extension of F with $\bar{F}(x) = \overline{\text{co}} F(x)$. Then $\bar{F}$ is still a Lipschitzian set-valued mapping whose Lipschitzian constant is still L .

Proof Let $x_0 \in X$ and consider $\bar{F}(x_0) = \overline{\text{co}} F(x_0) \subset Y$. By the property of closed hull, for $y_0 \in \bar{F}(x_0)$ and arbitrarily $\varepsilon > 0$, there exist $y_{i0} \in F(x_0)$, $i = 1, 2, \dots, s$ and λ_i , $i = 1, 2, \dots, s$ with $\lambda_i \geq 0$ and $\sum_{i=1}^s \lambda_i = 1$ such that $\|y_0 - \sum_{i=1}^s \lambda_i y_{i0}\| < \varepsilon$.

We now consider $x \in X$, and take s vectors y_i from $F(x)$, i.e., $y_i \in F(x)$, $i = 1, 2, \dots, s$. We have $\|y_0 - \sum_{i=1}^s \lambda_i y_i\| \leq \|y_0 - \sum_{i=1}^s \lambda_i y_{i0}\| + \|\sum_{i=1}^s \lambda_i y_{i0} - \sum_{i=1}^s \lambda_i y_i\| \leq \varepsilon + \sum_{i=1}^s \lambda_i \|y_{i0} - y_i\| < L \|x_0 - x\| + \varepsilon$ can be selected arbitrarily, consequently, $y_0 \in \overline{\text{co}} F(x) + L \|x - x_0\| B_Y$. $\square$

Theorem 2.1.8 illustrates the Lipschitzian property can be extended to its convex hull, and the Lipschitzian constant is invariant.

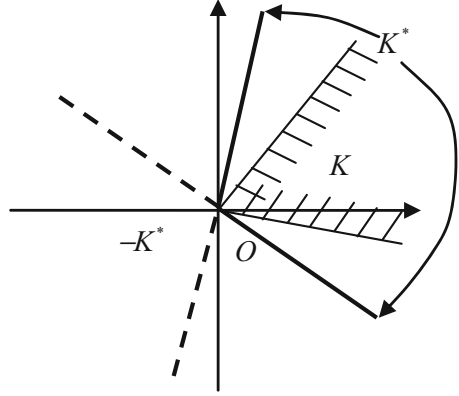
2.1.3 Tangent Cones and Normal Cones

Before considering the derivative of a set-valued mapping, we deal with the tangent cone and normal cone for a set-valued mapping.

1. Definition of cones

Recall the concept of the cone which is mentioned when we introduce the convex sets. A set K is called by a cone if $x \in K$ then $\lambda x \in K$ for every $\lambda \in \mathbb{R} (\geq 0)$. K is

Fig. 2.3 Cone K and conjugate cone K^* and polar cone $-K^*$



said to be a convex cone if and only if for $x, y \in K$ and arbitrarily $\lambda, \mu \in \mathbb{R} (\geq 0)$ such that $\lambda x + \mu y \in K$.

By the definition, it is easy to verify if K_1 and K_2 are two cones, then $K_1 \cup K_2$, $K_1 \cap K_2$, $K_1 + K_2$ and $K_1 \setminus K_2$ are also cones. If K_1 and K_2 are two convex cones, then $K_1 \cap K_2$ and $K_1 + K_2$ are convex cones.

Let K be a cone, the conjugate cone of K is denoted by K^*

$$K^* = \{x^*; \langle x^*, x \rangle \geq 0, \forall x \in K\}.$$

Moreover, $-K^*$ is called by the polar cone of K . For simplicity, we often apply N to denote a polar cone. It is obvious that both K^* and $-K^*$ are convex and closed no matter what is the cone K . Figure 2.3 gives an illustration of relations of K , K^* and $-K^*$ in $\mathbb{R}^2$.

By the definition of conjugate cone, it is direct to verify that $(K_1 + K_2)^* = K_1^* \cap K_2^*$ for two cones K_1 and K_2 .

In order to be convenient to application later, we deal with the support function of a cone. Let $K \subset \mathbb{R}^n$ be a cone. Then its support function $S(x, K)$ is defined by $S(x, K) = \sup_{x_K \in K} (\langle x, x_K \rangle)$. When $x^* \in -K^*$, then $\langle x^*, x_K \rangle \leq 0$ for every $x_K \in K$.

Since for all $\alpha \geq 0$, $\alpha x_K \in K$, $S(x^*, K) = \sup_{x_K \in K} (\langle x^*, \alpha x_K \rangle) \rightarrow 0, (\alpha \rightarrow 0)$.

When $x^* \notin -K^*$, there is a $x_K \in K$, such that $\langle x^*, x_K \rangle > 0$, then $S(x^*, K) = \sup_{x_K \in K} (\langle x^*, \alpha x_K \rangle) \rightarrow \infty, (\alpha \rightarrow \infty)$. Therefore, we can conclude that

$$S(x^*, K) = \delta(x^*, -K^*). \quad (2.1.4)$$

The fact can be explained from Fig. 2.3

The following is a fundamental result for a cone and its polar cone.

Theorem 2.1.9 Let $K \subset X$ be a closed and convex cone, N be the polar cone of K . $x \in X$ is a vector and $\pi(x, K)$ is the projection of x on K . Denote $\pi(x, N) = x - \pi(x, K)$. Then we have the following conclusions

- (1) $\pi(x, N) \in N$.
- (2) If $x \notin N$, then $\pi(x, N)$ is the projection of x on N .
- (3) $\|x\|^2 = \|\pi(x, K)\|^2 + \|\pi(x, N)\|^2$.

Proof (1) By Inequality (1.2.13), for every $y \in K$, we have

$$\langle x - \pi(x, K), y - \pi(x, K) \rangle \leq 0. \quad (2.1.5)$$

K is a cone, hence $0 \in K$, and by the definition of projection $\pi(x, K) \in K$, replacing y by 0 and $2\pi(x, K) \in K$, respectively, Inequality (2.1.5) yields, $\langle x - \pi(x, K), \pi(x, K) \rangle \geq 0$ and $\langle x - \pi(x, K), \pi(x, K) \rangle \leq 0$. Consequently,

$$\langle x - \pi(x, K), \pi(x, K) \rangle = 0. \quad (2.1.6)$$

Also by Eq. (2.1.6), Inequality (2.1.5) results in $\langle x - \pi(x, K), y \rangle \leq 0$ for every $y \in K$. It is just that $\pi(x, N) = x - \pi(x, K) \in N$ by the definition of polar cone. The first conclusion is verified.

(2) Let $v^* \in N$. We conclude that

$$\langle v^* - \pi(x, N), x - \pi(x, N) \rangle = \langle v^*, x - \pi(x, N) \rangle - \langle \pi(x, N), x - \pi(x, N) \rangle \leq 0.$$

It is because $\pi(x, K) = x - \pi(x, N)$, the first inner product is less than or equal to zero, and the second inner product is zero by Eq. (2.1.6). By Remark 2 of Lemma 1.2.2, $\pi(x, N)$ is the projection of x on N .

(3) We have

$$\begin{aligned} \|x\|^2 &= \|x - \pi(x, K) + \pi(x, K)\|^2 \\ &= \|x - \pi(x, K)\|^2 + \|\pi(x, K)\|^2 + 2\langle x - \pi(x, K), \pi(x, K) \rangle \\ &= \|x - \pi(x, K)\|^2 + \|\pi(x, K)\|^2. \end{aligned}$$

The conclusion 3 is now verified. □

2. Tangent cone and normal cone

In order to define the derivative for set-valued mappings, we consider tangent cone and normal cone of a set.

Definition 2.1.5 Let $A \subset X$ be a set and $x \in A$. The set

$$T(x, A) = \text{cl} \bigcup_{\lambda > 0} \frac{A - x}{\lambda} \quad (2.1.7)$$

is the tangent cone of A at x ; $-T^*(x, A)$, the polar cone of $T(x, A)$, is defined as the normal cone of A at x and denoted by $N(x, A)$. $\square$

Remark $T(x, A)$ is a cone. At first $x \in A$, hence $0 \in T(x, A)$. Moreover, if $y \in T(x, A)$, and there is a $\lambda > 0$ such that $y \in \frac{A-x}{\lambda}$, it follows that $\mu y \in \frac{A-x}{\lambda\mu-1}$, for every $\mu > 0$. For vectors in derived set of $\bigcup_{\lambda>0} \frac{A-x}{\lambda}$, the proof is similar and omitted. $\square$

If $X = \mathbb{R}^n$ and $x \in \text{int } A$, then $T(x, A) = \mathbb{R}^n$. It is a trivial case. Hence it is only meaningful to deal with the tangent cones at the boundary of A .

Lemma 2.1.1 Let A be a convex set, and two real numbers λ and μ satisfy $0 < \lambda < \mu$. Then for $x \in A$, $\frac{A-x}{\mu} \subset \frac{A-x}{\lambda}$.

Proof Let $y \in A$ be an arbitrarily element of A . Define $z = \left(1 - \frac{\lambda}{\mu}\right)x + \frac{\lambda}{\mu}y$, then $z \in A$ since A is convex. It is equivalent to that $y = \frac{\mu}{\lambda}z + \left(1 - \frac{\mu}{\lambda}\right)x$. Thus,

$$\frac{y-x}{\mu} = \mu^{-1} \left(\frac{\mu}{\lambda}z - \frac{\mu}{\lambda}x + x - x \right) = \frac{z-x}{\lambda} \in \frac{A-x}{\lambda}.$$

$\square$

Theorem 2.1.10 Let A be a convex set, and $x \in A$. Then

$$T(x, A) = \left\{ v; \lim_{\lambda \downarrow 0} \lambda^{-1} d(x + \lambda v, A) = 0 \right\}.$$

Before proving the theorem, we give an explanation. $d(x + \lambda v, A)$ is the distance of point $x + \lambda v$ to set A . The theorem asserts that, firstly, the tangent cone is a set-valued mapping and its effect domain is A ; secondly, if $d(x + \lambda v, A)$ is a higher infinitesimal when λ intends to zero, then the v belongs to the tangent cone. It meets well the definition of tangent line in calculus.

The proof of Theorem 2.1.10: If $v \in T(x, A)$, then by the Definition of 2.1.5, for every $\varepsilon > 0$, there is a positive real λ_0 such that $v \in \frac{A-x}{\lambda_0} + \varepsilon B_X$ where B_X is the unit ball of X . By Lemma 2.1.1, it follows that $v \in \frac{A-x}{\lambda} + \varepsilon B_X$ for every λ with $0 < \lambda \leq \lambda_0$. For the λ , we have

$$\lambda^{-1} d(x + \lambda v, A) = \lambda^{-1} \inf_{y \in A} \|x + \lambda v - y\| = \inf_{y \in A} \|v - \lambda^{-1}(y - x)\| \leq \varepsilon. \quad (2.1.8)$$

Because $\varepsilon > 0$ can be selected arbitrarily, $\lambda^{-1} d(x + \lambda v, A) \rightarrow 0$ as $\lambda \downarrow 0$.

Conversely, if $\lim_{\lambda \downarrow 0} \lambda^{-1} d(x + \lambda v, A) = 0$, then, for every $\varepsilon > 0$, there is a $\delta > 0$, when $\lambda < \delta$, we have Inequality (2.1.8). The last inequality means $v \in \text{cl} \frac{A-x}{\lambda} + \varepsilon B_X$, i.e., $v \in T(x, A)$. $\square$

Theorem 2.1.11 Let A be a convex set and $x \in A$. Then

$$N(x, A) = \{v^*; \langle v^*, y - x \rangle \leq 0, y \in A\}.$$

Proof By Definition 2.1.5, $A - x \subset T(x, A)$, and by the definition of normal cone $N(x, A) = -T^*(x, A)$, we have $\langle v^*, y - x \rangle \leq 0$ for every $y \in A$ and every $v^* \in N(x, A)$. Therefore, $N(x, A) \subset \{v^*; \langle v^*, y - x \rangle \leq 0, y \in A\}$.

Conversely, let $v \in T(x, A)$, then for $\varepsilon > 0$ there is a $\lambda > 0$ such that $v \in \frac{A-x}{\lambda} + \varepsilon B$. It can be rewritten by $\|v - \frac{y-x}{\lambda}\| \leq \varepsilon$ for some $y \in A$. Now, if $\langle v^*, y - x \rangle \leq 0$ holds for every $y \in A$, then

$$\begin{aligned} \langle v^*, v \rangle &= \left\langle v^*, v - \frac{y-x}{\lambda} + \frac{y-x}{\lambda} \right\rangle \\ &= \left\langle v^*, v - \frac{y-x}{\lambda} \right\rangle + \left\langle v^*, \frac{y-x}{\lambda} \right\rangle \\ &\leq \left\langle v^*, v - \frac{y-x}{\lambda} \right\rangle \\ &\leq \varepsilon \|v^*\|. \end{aligned}$$

We have applied Schwarz Inequality at the last step. Since the ε can be selected arbitrarily, $\langle v^*, v \rangle \leq 0$ for every $v \in T(x, A)$. It follows $v^* \in N(x, A)$. The conclusion is verified. $\square$

The last theorem in this subsection reveals the relation of $T(x, A)$ and $N(x, A)$, when they are treated as set-valued mappings.

Theorem 2.1.12 If $T(x, A)$ is compact and convex for every $x \in A \subset X$, then $\text{gra } N(x, A)$ is closed if and only if $T(x, A)$ is lower semi-continuous.

Proof The proof of the theorem applies the conclusion given in Problem 2 of this section. Let $(x_n, v_n^*) \in \text{gra } N(x, A)$ and $(x_n, v_n^*) \rightarrow (x, v^*)$. If we can prove $v^* \in N(x, A)$, then $\text{gra } N(x, A) (x \in A)$ is closed. Let $v \in T(x, A)$. $T(x, A)$ is lower semi-continuous, hence there is a $v_n \in T(x_n, A)$, $v_n \rightarrow v$. It follows that $\langle v_n, v_n^* \rangle \leq 0$. $0 \geq \lim_{n \rightarrow \infty} \langle v_n^*, v_n \rangle = \langle v^*, v \rangle$. Thus, $v^* \in N(x, A)$.

For a vector $v \in T(x, A)$, and a sequence $\{x_n\}$ with $x_n \rightarrow x$, we should construct a sequence $\{v_n\}$ such that $v_n \in T(x_n, A)$ and $v_n \rightarrow v$. Because $v \in T(x, A)$, and $T(x, A)$ is convex and compact, by using Theorem 2.1.9, we conclude there are $v_n = \pi(v, T(x_n, A)) \in T(x_n, A)$ and $v_n^* = \pi(v, N(x_n, A)) \in N(x_n, A)$ such that $v = v_n + v_n^*$. Thus, $\|v_n^*\| \leq \|v\|$, $\{v_n^*\}$ has a convergent subsequence, without loss of generality, we can assume $\{v_n^*\}$ is convergent, $v_n^* \rightarrow v^*$. $\text{gra } N(x, A)$ is closed, hence $v^* \in N(x, A)$. Furthermore, $v_n = v - v_n^* \rightarrow v - v^*$. If we can prove $v^* = 0$, then $v_n \rightarrow v$. The conclusion is verified.

By Eq. (2.1.6), $\langle v^*, v - v^* \rangle = \lim_{n \rightarrow \infty} \langle v_n^*, v - v_n^* \rangle = 0$, i.e.,

$$0 = \langle v^*, v - v^* \rangle = \langle v^*, v \rangle - \langle v^*, v^* \rangle.$$

Hence, $\|v^*\|^2 = \langle v^*, v^* \rangle = \langle v^*, v \rangle \leq 0$, $\|v^*\| = 0$. $\square$

2.1.4 Derivative of Set-Valued Mappings

This subsection starts with the discussion of tangent cone. At the most time, we restricted ourselves in $\mathbb{R}^n$.

Definition 2.1.6 Let $A \subset X$, and $x \in A$. The set defined as

$$T_+(x, A) = \left\{ v; \lim_{\lambda \downarrow 0} \lambda^{-1} d(x + \lambda v, A) = 0 \right\}$$

is called by the tangent cone of A at the x .

The set

$$T_-(x, A) = \left\{ v; \lim_{\lambda \downarrow 0} \inf \lambda^{-1} d(x + \lambda v, A) = 0 \right\}$$

is called by the contingent cone of A at the x . □

Recall the definition of limitation inferior of a function, if $\lim_{x \rightarrow x_0} \inf f(x) = c$ then (1) there exists a convergent $\{x_n\}$ sequence $x_n \rightarrow x_0$, such that $f(x_n) \rightarrow c$; (2) let $\{x_n\}$ be a convergent sequence, $x_n \rightarrow x_0$ and $f(x_n) \rightarrow \bar{c}$, then $\bar{c} \geq c$.

Remark 1 By Definition 2.1.6, it is obvious that both $T_+(x, A)$ and $T_-(x, A)$ are closed and $T_+(x, A) \subset T_-(x, A)$. □

Remark 2 If A is a convex set, then $\lambda^{-1} d(x + \lambda v, A)$ is decreasing with the decrease of λ . Hence, Theorem 2.1.10 concludes that the limitation $\lim_{\lambda \downarrow 0} \lambda^{-1} d(x + \lambda v, A)$ exists for every $v \in A$, and $T_+(x, A) = T_-(x, A) = T(x, A)$. Furthermore, $T_+(x, A) = T_-(x, A)$. □

Example 2.1.3 Consider a single-valued mapping

$$f(x) = \begin{cases} |x| \sin \frac{1}{|x|}, & x \neq 0, \\ 0, & x = 0, \end{cases}$$

and $A = \cup_{\text{epi}} f(x)$. Then the tangent cone $T_+(0, A)$ and the contingent cone $T_-(0, A)$ are shown in Fig. 2.4. □

Let $A \subset \mathbb{R}^n$ and $\alpha > 0$. Define a set $D_\alpha \subset \mathbb{R}^n \times \mathbb{R}$ as follows

$$D_\alpha = \{(x, x^0) \in \mathbb{R}^n \times \mathbb{R}, x^0 \geq \alpha d(x, A)\}.$$

In $\mathbb{R}^n \times \mathbb{R}$, the set D_α looks as a solid basin (Fig. 2.5). Hence D_α is called by the basin of A with the slope α .

Theorem 2.1.13 Let $A \subset \mathbb{R}^n$, $\alpha > 0$ and $\hat{x} \in A$. Then

$$T_+(\hat{x}, 0, D_\alpha) \supset \{(v, v^0) \in \mathbb{R}^n \times \mathbb{R}, v^0 \geq \alpha d(v, T_+(\hat{x}, A))\}.$$

Fig. 2.4 The tangent cone and contingent cone of Example 2.1.3

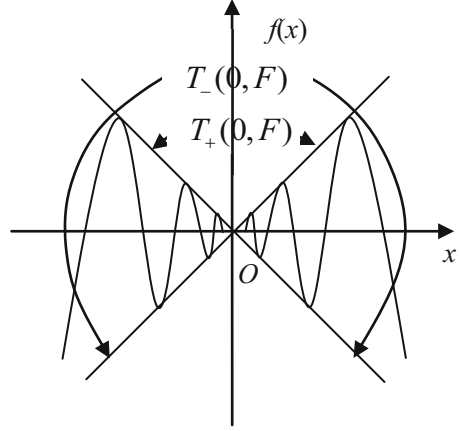
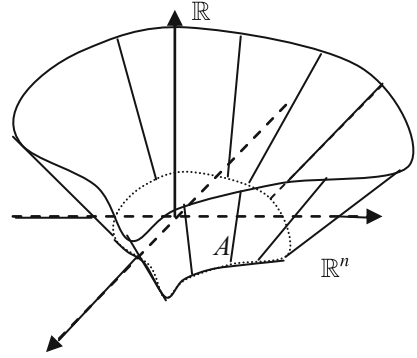


Fig. 2.5 The figure of D_α



Proof Let $v^0 \in \mathbb{R}$ and $v^0 \geq \alpha d(v, T_+(\hat{x}, A))$. Because $T_+(\hat{x}, A)$ is closed, there exists a $\omega \in T_+(\hat{x}, A)$ such that $d(v, T_+(\hat{x}, A)) = \|v - \omega\|$. Then $v^0 \geq \alpha \|v - \omega\|$. Moreover,

$$\alpha d(\hat{x} + \lambda v, A) \leq \alpha d(\hat{x} + \lambda \omega, A) + \alpha \lambda \|v - \omega\| \leq \alpha d(\hat{x} + \lambda \omega, A) + \lambda v^0.$$

Dividing both sides by λ , since $\omega \in T_+(\hat{x}, A)$, we have $\lim_{\lambda \downarrow 0} \lambda^{-1} \alpha d(\hat{x} + \lambda v, A) \leq v^0$.

It means $\lambda^{-1} \alpha d(\hat{x} + \lambda v, A) \leq v^0 + \varepsilon$ for small λ , i.e., $(\hat{x} + \lambda v, \lambda v^0 + \lambda \varepsilon) \in D_\alpha$. Equivalently,

$$\lim_{\lambda \downarrow 0} \lambda^{-1} d\left(\left(\begin{bmatrix} \hat{x} \\ 0 \end{bmatrix} + \begin{bmatrix} \lambda v \\ \lambda v^0 \end{bmatrix}\right), D_\alpha\right) \leq \varepsilon.$$

$\varepsilon > 0$ can be selected arbitrarily. Hence, $(v, v^0) \in T_+((\hat{x}, 0), D_\alpha)$. $\square$

Lemma 2.1.2 Let $F : X \rightarrow Y$ be a Lipschitzian mapping with Lipschitzian constant l . Then

$$d((x, y), \text{gra } F) \leq d(y, F(x)) \leq (1 + l) d((x, y), \text{gra } F).$$

Proof We have mentioned that in the Cartesian product space $X \times Y$, the norm can be defined as $(x, y) \in X \times Y$, $\|(x, y)\| = \|x\|_X + \|y\|_Y$. By the definition, there exists a $z \in F(x)$ such that $d(y, F(x)) \geq \|y - z\| - \varepsilon$ for every $\varepsilon > 0$. Hence, we have

$$d((x, y), \text{gra } F) \leq \|(x, y) - (x, z)\|_{X \times Y} = \|y - z\|_Y \leq d(y, F(x)) + \varepsilon.$$

Because $\varepsilon > 0$ can be selected arbitrarily, it leads to

$$d((x, y), \text{gra } F) \leq d(y, F(x)). \quad (2.1.9)$$

The left inequality is obtained.

Similarly, for $\varepsilon > 0$, there is $(x_0, y_0) \in \text{gra } F$ such that

$$d((x, y), \text{gra } F) \geq \|(x, y) - (x_0, y_0)\|_{X \times Y} - \varepsilon = \|x - x_0\|_X + \|y - y_0\|_Y - \varepsilon.$$

Because F is a Lipschitzian mapping, for this x , there is a $z \in F(x)$ such that $\|z - y_0\|_Y \leq l\|x - x_0\|_X$. Then,

$$\begin{aligned} d(y, F(x)) &\leq \|y - z\|_Y \\ &\leq \|y - y_0\|_Y + \|y_0 - z\|_Y \\ &\leq \|y - z\|_Y + l\|x - x_0\|_X \\ &\leq (1 + l)(\|y - z\|_Y + \|x - x_0\|_X) \\ &\leq (1 + l)(d((x, y), \text{gra } F) + \varepsilon). \end{aligned}$$

The right side inequality can be obtained since $\varepsilon > 0$ can be selected arbitrarily. $\square$

To end this section, we define the derivative for set-valued mappings. We start with a recall of single-valued function. Let $y = f(x)$ where $x, y \in \mathbb{R}$, be a function. In the plane of $\mathbb{R}^2$, the graph of $y = f(x)$ is a set of $\{(x, f(x)) ; x \in \mathbb{R}\}$, i.e., $\text{gra } f = \{(x, f(x))\}$. Let $f'(x_0)$ be the derivative of $f(x)$ at x_0 . Then, the tangent vector at $(x_0, f(x_0))$ is $(x, f'(x_0)x)$. It is clear that the set $\{(x, f'(x_0)x) ; x \in \mathbb{R} (\geq 0)\}$ forms a cone. The idea is extended to the set-valued mapping.

Definition 2.1.7 Let $F : X \rightarrow Y$ be a set-valued mapping. Suppose $(x_0, y_0) \in \text{gra } F$. Then a set-valued mapping $D_+F(x_0, y_0) : X \rightarrow Y$ is defined by

$$\text{gra } D_+F(x_0, y_0) = T_+((x_0, y_0), \text{gra } F).$$

$D_+F(x_0, y_0)$ is called by the derivative of F at (x_0, y_0) . Similarly, the set-valued mapping $D_-F(x_0, y_0) : X \rightarrow Y$ defined by

$$\text{gra } D_-F(x_0, y_0) = T_-((x_0, y_0), \text{gra } F)$$

is called by contingent derivative of F at (x_0, y_0) . □

The statement of $(x, y) \in \text{gra } D_+F(x_0, y_0)$ has an alternative explanation that $y \in D_+F(x_0, y_0)(x)$. By the notion, $D_+F(x_0, y_0)$ is qualified as the derivative of set-valued mapping F at (x_0, y_0) . We have to denote it by two variables x_0 and y_0 since for every x_0 there exist more than one $y \in F(x_0)$. Correspondingly, $D_+F(x_0, y_0)(x)$ is the differential. Similarly, $D_-F(x_0, y_0)(x)$ is the contingent differential of F at (x_0, y_0) . The following theorem gives an explanation of differential.

Theorem 2.1.14 Suppose $F : X \rightarrow Y$ is a Lipschitzian set-valued mapping, then

$$D_+F(x_0, y_0)(x) = \left\{ y; \lim_{\lambda \downarrow 0} \lambda^{-1} d(y_0 + \lambda y, F(x_0 + \lambda x)) = 0 \right\};$$

$$D_-F(x_0, y_0)(x) = \left\{ y; \lim_{\lambda \downarrow 0} \inf \lambda^{-1} d(y_0 + \lambda y, F(x_0 + \lambda x)) = 0 \right\}.$$

Proof Let l be the Lipschitzian constant of F . Then By Lemma 2.1.2, we have

$$\begin{aligned} d((x_0 + \lambda x, y_0 + \lambda y), \text{gra } F) &\leq d(y_0 + \lambda y, F(x_0 + \lambda x)) \\ &\leq (1 + l) d((x_0 + \lambda x, y_0 + \lambda y), \text{gra } F). \end{aligned} \quad (2.1.10)$$

Multiplying both side by λ^{-1} , and let $\lambda \downarrow 0$. If $(x, y) \in T_+((x_0, y_0), \text{gra } F)$, then

$$d((x_0 + \lambda x, y_0 + \lambda y), \text{gra } F) = 0.$$

Hence $\lim_{\lambda \downarrow 0} \lambda^{-1} d(y_0 + \lambda y, F(x_0 + \lambda x)) = 0$.

Conversely, if $\lim_{\lambda \downarrow 0} \lambda^{-1} d(y_0 + \lambda y, F(x_0 + \lambda x)) = 0$, then by using the first inequality of (2.1.10), we have $\lim_{\lambda \downarrow 0} d((x_0 + \lambda x, y_0 + \lambda y), \text{gra } F) = 0$, i.e., $(x, y) \in T_+((x_0, y_0), \text{gra } F)$.

The proof for the contingent derivative is similar to the case of tangent derivative. Hence it is omitted. □

Theorem 2.1.14 shows that the computation of tangent derivative for the set-valued mapping is quite similar to that of single-valued function.

Problems

1. $F : X \rightarrow Y$ is a set-valued mapping, $G \subset Y$ is a set. Prove the following equations.

$$F^{-1}(G^c)(w) = [F^{-1}(G)(s)]^c,$$

$$F^{-1}(G^c)(s) = [F^{-1}(G)(w)]^c.$$

2. Prove Theorem 2.1.2.
3. Let $F : X \rightarrow Y$ be set-valued mapping. Prove the following statements are equivalent:
 - (1) F is lower semi-continuous at x_0 .
 - (2) For every $y_0 \in F(x_0)$ and $\varepsilon > 0$, there is a $\delta > 0$ such that for every $x \in B(x_0, \delta)$ there exists a $y \in F(x) \cap B(y_0, \varepsilon)$.
 - (3) For every $y_0 \in F(x_0)$, there exists a sequence $\{x_n; x_n \in X\}$ such that $x_n \rightarrow x_0$, moreover, there is a sequence $\{y_n; y_n \in F(x_n)\}$ such that $y_n \rightarrow y_0$.
4. A set-valued mapping $F(x)$ is lower semi-continuous. Is it ε - lower semi-continuous? If you answer “no”, then give a counter example.
5. Prove the following statements:
 - (1) If the set-valued mapping $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is upper semi-continuous, and $g : \mathbb{R}^p \rightarrow \mathbb{R}^n$ is continuous single-valued function, then $F \circ g : \mathbb{R}^p \rightarrow \mathbb{R}^m$ is upper semi-continuous.
 - (2) If the set-valued mapping $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is lower semi-continuous, and $g : \mathbb{R}^p \rightarrow \mathbb{R}^n$ is continuous single-valued function, then $F \circ g : \mathbb{R}^p \rightarrow \mathbb{R}^m$ is lower semi-continuous.
6. Let $F : X \rightarrow Y$ be set-valued mapping. Prove the following statements are equivalent:
 - (1) F is ε - upper semi-continuous.
 - (2) $\lim_{x \rightarrow x_0} \sup F(x) = F(x_0)$ (By the definition of upper limitation, it is equivalent to $\{u_0; \exists x_n \rightarrow x_0, u_n \in F(x_n), u_n \rightarrow u_0\} \subset F(x_0)$);
 - (3) If $u \notin F(x_0)$, there exist open sets W and V such that $u \in W \subset Y, x_0 \in V \subset X$ and $V \cap F^{-1}(W) = \emptyset$.
7. Give examples to show that
 - (1) $F : X \rightarrow Y$ is an open set-valued mapping, but there is an $x_0 \in X$ such that $F(x_0)$ is not an open set.
 - (2) $F : X \rightarrow Y$ is a set-valued mapping with closed value, but F is not a closed mapping.

8. Let $F : X \rightarrow Y$ be set-valued mapping. Suppose $\text{gra } F$ is a closed set, and for every $x_0 \in X$ and $\delta > 0$, the set

$$M_\delta = \text{cl } \{F(x); x \in B(x_0, \delta)\}$$

is compact. Then F is upper semi-continuous.

9. Let $K_1, K_2 \subset \mathbb{R}^n$ be two closed and convex cones. If $\text{re int } K_1 \cap \text{re int } K_2 \neq \emptyset$, then $\text{cl}(K_1 \cap K_2) = \text{cl}K_1 \cap \text{cl}K_2$. Moreover, $(K_1 \cap K_2)^* = (\text{cl}K_1 \cap \text{cl}K_2)^* = \text{cl}(K_1^* + K_2^*)$.
10. Let $K_1, K_2 \subset \mathbb{R}^n$ be two closed and convex cones. Prove the following conclusions and further to discuss which statement is still valid if the condition of “closed” or “convex” is removed.
- (1) $(K_1 + K_2)^* = K_1^* \cap K_2^*$;
 - (2) $K_1^{**} = K_1$;
 - (3) If $\Lambda : \mathbb{R}^m \rightarrow \mathbb{R}^n$ is a linear mapping, and $\Lambda\mathbb{R}^m - K_1 = \mathbb{R}^n$, then $(\Lambda^{-1}K_1)^* = \Lambda^*K_1^*$ (If Λ is treated as a matrix, then Λ^* is its conjugate matrix);
 - (4) $\partial\delta(x, K_1) = -K_1^*$.
11. Let $x_i \in \mathbb{R}^n$, $i = 1, 2, \dots, m$ and denote $K = \{y; \langle y, x_i \rangle \geq 0, i = 1, 2, \dots, m\}$. Then K is a closed and convex cone. Moreover, $K = \overline{\text{con}} \{x_i, i = 1, 2, \dots, m\}$.
12. Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a set-valued mapping with convex and compact value (i.e., for every x , $F(x)$ is a convex and compact set). Then F is continuous if and only if the supporting mapping $x \rightarrow S(y^*, F(x))$ is continuous for every y^* .
13. Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a set-valued mapping. If $\text{gra } F$ is closed and the set

$$M_\delta = \text{cl } \{F(x); \|x - x_0\| < \delta\}$$

is compact for some $\delta > 0$, then F is upper semi-continuous at x_0 .

14. Let $F : \mathbb{R}^n \rightarrow A$ be a set-valued mapping, and $A \subset \mathbb{R}^m$ be a compact set. If $\text{gra } F$ is closed, then F is with closed value and is upper semi-continuous.

2.2 Selection of Set-Valued Mappings

In the first section of this chapter, we have introduced the set-valued mapping and pointed out that the image of a set-valued mapping $F(x)$ is a set for every x in the domain. Consequently, it is possible to select a $y \in F(x)$ for every $x \in \text{dom } F$, then a single-valued mapping $y = f(x)$ yields. It looks a very simple task. The task is called selection of a set-valued mapping. Alternatively, a selection is to give a way to construct a single-valued mapping $y = f(x)$ such that $f(x) \in F(x)$ for every $x \in A$. If there is an $x \in A$, $F(x)$ holds more than one elements, then the selection is not unique. The requirement is to choose a mapping $f(x)$ with some satisfied properties,

for example, it is a continuous mapping, or a Lipschitzian mapping, etc. Readers will find that obtaining a nice selection is a hard work. It is quite extravagant to obtain a Lipschitzian selection.

Usually, there are three kinds selection: continuous selection, measurable selection and approximate selection. This book only considers continuous selection and approximate selection. The later is to look for a sequence $\{f_n(x)\}$ such that the limitation of $f_n(x)$ belongs to $\text{cl } F(x)$. After the theory of selection, we spend some time to deal with the problem of fixed points which is a very important title in topology.

2.2.1 The Minimal Selection

Definition 2.2.1 $A \subset X$ is a set. If there is a vector $x_0 \in A$ such that $\|x_0\| = \min \{\|x\|, x \in A\}$, then the x_0 is called by a minimal norm element of A . The set of minimal norm elements is denoted by $m(A)$. $\square$

In general, we cannot guarantee the existence of minimal norm element for a set. But if the space is complete and the set is closed, then we can assure $m(A) \neq \emptyset$. Moreover, when the set A is convex and closed, then $m(A)$ holds only one element. If $m(A)$ has only one element, we call it the minimal norm element and also denote it by $m(A)$ for simplicity.

Suppose A is a convex and closed set. Because $m(A) = \arg \min_{x \in A} \|x\|$ ¹ We have $m(A) = \pi(0, A)$. By Remark 2 of Lemma 1.2.2, $\langle m(A) - x_A, m(A) \rangle \leq 0$ for every $x_A \in A$ and if there is an element x such that $\langle x - x_A, x \rangle \leq 0$ for every $x_A \in A$, then $x = m(A)$.

Theorem 2.2.1 Suppose that the set-valued mapping $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is continuous and with convex and closed value. Then $f(x) = m(F(x))$ is a continuous selection.

Proof By the condition of theorem, $F(x)$ is a convex and closed set for every $x \in \mathbb{R}^n$. Hence, $m(F(x))$ has only one element. We now prove that the single-valued mapping $x \mapsto m(F(x))$ is continuous.

Denote $y_0 = m(F(x_0))$. We first consider the case that $y_0 = 0 \in \mathbb{R}^m$. By the condition that $F(x)$ is continuous and is lower semi-continuous at x_0 . For every $\varepsilon > 0$, there is $\delta > 0$ such that if $x \in B(x_0, \delta)$, then $F(x) \cap B(0, \varepsilon) \neq \emptyset$. It means there is a $y(x) \in F(x) \cap B(0, \varepsilon)$. From the definition of the minimal norm element, we have $\|m(F(x))\| \leq \|y(x)\| \leq \varepsilon$, i.e., $f(x)$ is continuous at x_0 .

We now turn to the case where $y_0 \neq 0$. $F(x)$ is upper semi-continuous at x_0 . Hence, for every $\varepsilon > 0$, there is $\delta > 0$ such that if $x \in B(x_0, \delta)$, then

¹The notation $\arg \min_{x \in A} f(x)$ expresses a set. If $x_0 \in \arg \min_{x \in A} f(x)$ then $f(x_0) = \min_{x \in A} f(x)$. Similarly, we can define $\arg \max_{x \in A} f(x)$. The set $\arg \{f(x) = 0\}$ is the set of roots of $f(x) = 0$ and $x \in A$.

$$F(x) \subset F(x_0) + \frac{\varepsilon}{\|y_0\|} B_m. \quad (2.2.1)$$

There exists a $y(x_0) \in F(x_0)$ such that $m(F(x)) \in \left(y(x_0) + \frac{\varepsilon}{\|y_0\|} B_m\right)$, or equivalently, $\|m(F(x)) - y(x_0)\| < \frac{\varepsilon}{\|y_0\|}$. It follows $|\langle y_0, m(F(x)) - y(x_0) \rangle| \leq \|y_0\| \|m(F(x)) - y(x_0)\| < \varepsilon$.

By the remark given behind Definition 2.2.1, $\langle y_0 - y(x_0), y_0 \rangle \leq 0$, i.e., $0 < \langle y_0, y_0 \rangle \leq \langle y_0, y(x_0) \rangle$. It leads to

$$\begin{aligned} \langle y_0, m(F(x)) \rangle &= \left\langle y_0, y(x_0) \right\rangle + \langle y_0, m(F(x)) - y(x_0) \rangle \\ &\geq \left\langle y_0, y(x_0) \right\rangle - \varepsilon \\ &\geq \langle y_0, y_0 \rangle - \varepsilon, \end{aligned}$$

or,

$$\langle y_0, m(F(x)) - y_0 \rangle \geq -\varepsilon.$$

From the Problem 1 of this section, the relation (2.2.1) implies $\|m(F(x))\| \leq \|y_0\| + \frac{\varepsilon}{\|y_0\|}$. Thus,

$$\begin{aligned} \langle m(F(x)) - y_0, m(F(x)) - y_0 \rangle &= \langle m(F(x)), m(F(x)) - y_0 \rangle - \langle y_0, m(F(x)) - y_0 \rangle \\ &= \left\langle m(F(x)), m(F(x)) \right\rangle - \langle m(F(x)), y_0 \rangle \\ &\quad - \langle y_0, m(F(x)) - y_0 \rangle \\ &= \|m(F(x))\|^2 - \left\langle m(F(x)) - y_0 + y_0, y_0 \right\rangle \\ &\quad - \langle y_0, m(F(x)) - y_0 \rangle \\ &= \|m(F(x))\|^2 - \|y_0\|^2 - 2 \langle y_0, m(F(x)) - y_0 \rangle. \end{aligned}$$

On the other hand, $\|m(F(x))\|^2 \leq \left(\|y_0\| + \frac{\varepsilon}{\|y_0\|}\right)^2 = \|y_0\|^2 + 2\varepsilon + \left(\frac{\varepsilon}{\|y_0\|}\right)^2$. The above inequality leads to

$$\|m(F(x)) - y_0\|^2 \leq 4\varepsilon + \left(\frac{\varepsilon}{\|y_0\|}\right)^2.$$

We then conclude that $f(x) = m(F(x))$ is continuous at x_0 . $\square$

Corollary 2.2.1 If $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is continuous and with convex and closed value, then for every $a \in \mathbb{R}^n$, $f(x) = \pi(a, F(x))$ is a continuous selection.

Proof Let $\bar{x} = x - a$. Then $\pi(a, F(x)) = \pi(0, F(\bar{x} + a)) = \pi(0, \bar{F}(\bar{x})) = m(\bar{F}(\bar{x}))$. It is clear that $\bar{F}$ is continuous and with convex and closed value. $\bar{f}(\bar{x}) = m(\bar{F}(\bar{x}))$ is a continuous selection by Theorem 2.2.1. Hence, $f(x) = \bar{f}(x - a)$ is continuous. $\square$

Remark 1 Theorem 2.2.1 and Corollary 2.2.1 are still valid, if F is a set-valued mapping with convex and compact value from a normed space X to a Hilbert space Y . $\square$

Remark 2 Recall the proof of Theorem 2.2.1, we have applied the lower semi-continuity only for the case that $y_0 = 0 \in F(x_0)$. Hence, if $0 \notin F(x)$ then it is sufficient for Theorem 2.2.1 that F is upper semi-continuous and with convex and closed value. Accordingly, Corollary 2.2.1 can be rewritten as “there is a $a \in \mathbb{R}^m$, $a \notin F(x)$ for all $x \in \mathbb{R}^n$, then $F(x)$ holds a continuous selection provided that $F(x)$ is upper semi-continuous and with convex and closed value. $\square$

To extend the conclusion to a more general result, we give a lemma now.

Lemma 2.2.1 Let A be a convex and closed set of $\mathbb{R}^n$. Then the mapping of projection $x \mapsto \pi(x, A)$ is a Lipschitzian mapping and its Lipschitzian constant can be 1.

Proof Suppose $x_1, x_2 \in \mathbb{R}^n$. By Remark 2 of Lemma 1.2.2, we have $\langle \pi(x, A) - x_A, \pi(x, A) - x \rangle \leq 0$. Replacing x by x_1, x_2 and replacing x_A by $\pi(x_2, A)$ and $\pi(x_1, A)$, respectively, we obtain

$$\langle \pi(x_1, A) - \pi(x_2, A), \pi(x_1, A) - x_1 \rangle \leq 0 \text{ and } \langle \pi(x_2, A) - \pi(x_1, A), \pi(x_2, A) - x_2 \rangle \leq 0.$$

The second inequality can be rewritten as $\langle \pi(x_1, A) - \pi(x_2, A), x_2 - \pi(x_2, A) \rangle \leq 0$ and is added to the first inequality, we obtain

$$\langle \pi(x_1, A) - \pi(x_2, A), \pi(x_1, A) - x_1 + x_2 - \pi(x_2, A) \rangle \leq 0.$$

It leads to

$$\begin{aligned} \|\pi(x_1, A) - \pi(x_2, A)\|^2 &\leq \langle x_1 - x_2, \pi(x_1, A) - \pi(x_2, A) \rangle \\ &\leq \|x_1 - x_2\| \|\pi(x_1, A) - \pi(x_2, A)\|. \end{aligned}$$

When $\pi(x_1, A) \neq \pi(x_2, A)$, it leads to

$$\|\pi(x_1, A) - \pi(x_2, A)\| \leq \|x_1 - x_2\|. \quad (2.2.2)$$

Inequality (2.2.2) is also valid if $\pi(x_1, A) = \pi(x_2, A)$. The lemma is verified. $\square$

Theorem 2.2.2 Suppose that the set-valued mapping $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is continuous and with convex and closed value, and $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a continuous single-valued function, then $f(x) = \pi(g(x), F(x))$ is a continuous selection of F .

Proof $F(x)$ is convex and closed, it follows $f(x) = \pi(g(x), F(x))$ is a single valued mapping. Consider now the continuity of $f(x)$.

$$\begin{aligned} f(x_1) - f(x_2) &= \pi(g(x_1), F(x_1)) - \pi(g(x_2), F(x_2)) \\ &= \pi(g(x_1), F(x_1)) - \pi(g(x_1), F(x_2)) + \pi(g(x_1), F(x_2)) \\ &\quad - \pi(g(x_2), F(x_2)) . \end{aligned}$$

If we take $a = g(x_1)$, by Corollary 2.2.1

$$\pi(g(x_1), F(x_1)) - \pi(g(x_1), F(x_2)) \rightarrow 0, \quad (x_1 - x_2 \rightarrow 0);$$

On the other hand, if we treat $F(x_2)$ as set A , by Lemma 2.2.1, we have

$$\pi(g(x_1), F(x_2)) - \pi(g(x_2), F(x_2)) \rightarrow 0, \quad (x_1 - x_2 \rightarrow 0).$$

Consequently, $f(x_1) - f(x_2) \rightarrow 0$ as $x_1 - x_2 \rightarrow 0$, $f(x)$ is continuous. $\square$

The selection obtained by Theorem 2.2.1 is called by the minimal selection since the selected element is the vector with minimal norm. Because Theorem 2.2.2 deduces from Theorem 2.2.1, it is called as patulous minimal selection theorem or minimal selection theorem for simplicity. The conclusion is still valid if the $\mathbb{R}^n$ and $\mathbb{R}^m$ are replaced by Hilbert Spaces X and Y . There is an example to show the minimal selection is continuous but may fail to be Lipschitzian mapping. The readers who are interested in the example are referred to Aubin and Cellina (1984).

2.2.2 Michael Selection Theorem

The most famous result in selection theory of set-valued mapping is the Michael selection theorem. It is notable that in the proof of Theorem 2.2.2, we have applied the operation of inner production. Hence, if we try to apply Theorem 2.2.2, the image space should be Hilbert space. The Michael selection theorem improves the condition of Theorem 2.2.1, the image space is only a Banach space. In this subsection, we do not restrict ourselves in $\mathbb{R}^n$. Let us start with a conclusion of unit decomposition.

1. Unit decomposition

Decomposing the integer 1 into a summation of several nonnegative functions defined on a normed space is known as the unit decomposition, or the partition of 1. In the theory of functional analysis, there are many kinds unit decomposition. This book only introduces the Lipschitzian unit decomposition.

Let X be a normed space, $A \subset X$ be a set. $\{\Omega_i; i \in I\}$ is an open covering of A .² The open covering $\{\Omega_i; i \in I\}$ is said to be a locally finite open covering if for every $\Omega_{i_0} \in \{\Omega_i; i \in I\}$, there are only finite open sets $\Omega_j \in \{\Omega_i; i \in I\}$, $j = 1, 2, \dots, J$ such that $\Omega_{i_0} \cap \Omega_j \neq \emptyset$.

Let $\{\Omega_i; i \in I\}$ be an open covering of $A \subset X$ and $\{p_i(x); i \in I\}$, where $p_i(x) : A \rightarrow \mathbb{R} (\geq 0)$, be a set of functions. The set $\{p_i(x); i \in I\}$ is a unit decomposition attached to $\{\Omega_i; i \in I\}$, if it satisfies the following conditions:

- (1) $p_i(x) \geq 0$, $x \in \Omega_i$; $p_i(x) = 0$, $x \notin \Omega_i$;
- (2) For every $x \in A$, $\sum_{i \in I} p_i(x) = 1$.

If $p_i(x)$ is continuous for every $i \in I$, then $\{p_i(x); i \in I\}$ is said to be a continuous decomposition. If $p_i(x)$ is a Lipschitzian function for every $i \in I$, then $\{p_i(x); i \in I\}$ is said to be a Lipschitzian decomposition. If the open covering $\{\Omega_i; i \in I\}$ is determined, we often omit the words “attached to $\{\Omega_i; i \in I\}$ ” for simplicity.

Lemma 2.2.2 Let $A \subset \mathbb{R}^n$, and $\{\Omega_i; i \in I\}$ be an open covering which is nontrivial and locally finite. Then there exists a local Lipschitzian unit composition.

Proof For every $i \in I$, define a function $q_i(x) : A \rightarrow \mathbb{R} (\geq 0)$, $q_i(x) = d(x, A \setminus \Omega_i)$. By Problem 2 in this section, $q_i(x)$ is a Lipschitzian function and its Lipschitzian constant is 1. Moreover, $q_i(x) = 0$, $x \notin \Omega_i$ and $q_i(x) \geq 0$, $x \in \Omega_i$ since Ω_i is an open set. The $p_i(x)$ is defined as follows:

$$p_i(x) = \frac{q_i(x)}{\sum_{j \in I} q_j(x)}.$$

By the definition of $q_i(x)$, it is obvious that the set $\{p_i(x); i \in I\}$ satisfies the properties of unit decomposition. The remaining problem is that it is a locally Lipschitzian function. We now prove this fact.

Let $x_0 \in A$. Then there is an $i_0 \in I$, such that $x_0 \in \Omega_{i_0}$. Ω_{i_0} is an open set, there is a neighborhood $B(x_0, \varepsilon)$ such that $B(x_0, \varepsilon) \subset \Omega_{i_0}$. The open covering is locally finite; hence, there are only finite $\Omega_{j_k} \in \{\Omega_i; i \in I\}$, $k = 1, 2, \dots, N$ such that $B(x_0, \varepsilon) \cap \Omega_{j_k} \neq \emptyset$. It follows that when $x \in B(x_0, \varepsilon)$, there are at most N functions $q_{j_k}(x)$ are nonzero in $\{q_i(x); i \in I\}$. On the other hand, there exist $0 < m < M$, such that $m \leq \sum_{j \in I} q_j(x) \leq M$ on the compact set $\text{cl}B(x_0, \varepsilon)$. Thus, when $x_1, x_2 \in B(x_0, \varepsilon)$, we have

²If there is an Ω_{i_0} , $i_0 \in I$, such that $\text{cl } \Omega_{i_0} \supset A$, then the covering is trivial. We do not deal with such a special case.

$$\begin{aligned}
|p_i(x_1) - p_i(x_2)| &= \left| \frac{q_i(x_1)}{\sum_{j \in I} q_j(x_1)} - \frac{q_i(x_2)}{\sum_{j \in I} q_j(x_2)} \right| \\
&= \left| \frac{q_i(x_1) \sum_{j \in I} q_j(x_2) - q_i(x_2) \sum_{j \in I} q_j(x_1)}{\sum_{j \in I} q_j(x_1) \sum_{j \in I} q_j(x_2)} \right| \\
&\leq \frac{1}{m^2} \sum_{j \in I} |q_i(x_1) q_j(x_2) - q_i(x_2) q_j(x_1)| \\
&\leq \frac{1}{m^2} \sum_{j \in I} \left(|q_i(x_1) q_j(x_2) - q_i(x_1) q_j(x_1)| \right. \\
&\quad \left. + |q_i(x_1) q_j(x_1) - q_i(x_2) q_j(x_1)| \right) \\
&\leq \frac{1}{m^2} \sum_{j \in I} \left(|q_i(x_1)| |q_j(x_2) - q_j(x_1)| + |q_j(x_1)| |q_i(x_1) - q_i(x_2)| \right) \\
&\leq \frac{1}{m^2} \left(\sum_{j \in I} |q_j(x_2) - q_j(x_1)| \right) + (|q_i(x_1) - q_i(x_2)| \sum_{j \in I} |q_j(x_1)|) \\
&\leq \frac{1}{m^2} (N + MN) |x_2 - x_1| .
\end{aligned}$$

The conclusion is verified. $\square$

2. Michael selection Theorem

This subsection proves the main result of this section: Michael selection theory. We introduce the following two facts.

Lemma 2.2.3 Suppose that X and Y are two normed spaces, $\varphi : X \rightarrow Y$ is a continuous single-valued mapping, and $\varepsilon : X \rightarrow \mathbb{R} (> 0)$ is a continuous functional. If $G : X \rightarrow Y$ is a lower semi-continuous set-valued mapping, then the mapping defined by

$$F(x) = B(\varphi(x), \varepsilon(x)) \cap G(x)$$

is lower semi-continuous. $\square$

The proof of this lemma is a classical exercise of application of $\varepsilon - \delta$ language. Hence we leave it to readers.

Another fact is about paracompactness. A set is said to be paracompact if for every open covering of the set there exists a fined open covering which is locally finite. Using mathematical language, it can be defined as follows. Let A be set, and $\{U_\alpha\}$ is an open covering of A . Then there exists another open covering $\{V_\beta\}$. $\{V_\beta\}$ is locally finite, and for every $V_\beta \in \{V_\beta\}$, there is a $U_\alpha \in \{U_\alpha\}$ such that $V_\beta \subset U_\alpha$. The $\{V_\beta\}$ is called a fineness of $\{U_\alpha\}$.

In 1948, A.H. Stone proved that a metric space is paracompact. The proof of this conclusion is quite complicated, we do not try to provide it in this book. The readers are referred to textbook for topology for the proof.

Theorem 2.2.3 (Michael selection theorem) Let X and Y be two normed spaces, and Y be complete. If F is a lower semi-continuous and with convex and closed value, then F exists a continuous selection.

Proof The proof of Theorem 2.2.3 consists of three steps. The critical step is construct a continuously single-valued mapping $\varphi_\varepsilon: X \rightarrow Y$ which satisfies $d(\varphi_\varepsilon(x), F(x)) < \varepsilon$. We now deal with the first step.

(1) For a given $\varepsilon > 0$, there is a continuously single-valued mapping $\varphi_\varepsilon: X \rightarrow Y$ such that $d(\varphi_\varepsilon(x), F(x)) < \varepsilon$.

$F: X \rightarrow Y$ is lower semi-continuous, hence, for a given $\varepsilon > 0$ and every $x \in X$, there is a $y_x \in F(x)$, such that $F(\bar{x}) \cap B(y_x, \varepsilon) \neq \emptyset$ provided that $\bar{x} \in B(x, \delta)$ for some $\delta > 0$. We now fix the vector y_x and the neighborhood $B(x, \delta)$ for every $x \in X$.

Thus the set $\{B(x, \delta), x \in X\}$ is an open covering of X . X is a normed space; consequently, by Stone theorem, $\{B(x, \delta), x \in X\}$ holds a fined open covering $\{U_i\}$ which is locally finite, i.e., for every U_i , there is a $B(x_i, \delta_i)$ such that $U_i \subset B(x_i, \delta_i)$. Lemma 2.2.2 shows that the $\{U_i\}$ provides a Lipschitzian unit partition $\{p_i(x)\}$.

For every $x \in X$, there are only finite $U_1, \dots, U_{N_x} \in \{U_i\}$, such that $x \in U_j$, $j = 1, 2, \dots, N_x$. Then, there is a $B(x_j, \delta_j)$ such that $U_j \subset B(x_j, \delta_j)$. From this x_j there is a corresponding y_{x_j} such that $y_{x_j} \in B(x_j, \delta_j)$ and $F(\bar{x}) \cap B(y_{x_j}, \varepsilon) \neq \emptyset$ provided that $\bar{x} \in B(x_j, \delta_j)$. We now define the following mapping.

$$\varphi_\varepsilon(x) = \sum_{j \in I} p_j(x) y_{x_j},$$

where I is the index set of i . $\varphi_\varepsilon(x)$ is then a local Lipschitzian mapping, so it is continuous. Because $x \in U_j \subset B(x_j, \delta_j)$, $F(x) \cap B(y_{x_j}, \varepsilon) \neq \emptyset$, i.e., $y_{x_j} \in F(x) + \varepsilon B_m$. $F(x)$ is a convex set, hence $\varphi_\varepsilon(x) = \sum_{j \in I} p_j(x) y_{x_j} \in F(x) + \varepsilon B_m$, i.e., $d(\varphi_\varepsilon(x), F(x)) < \varepsilon$ for every $x \in X$.

It is worth to note that at the above verification, we did not apply the condition that $F(x)$ is with closed value.

(2) A Cauchy sequence $\{f_n(x)\}$ is constructed where every $f_n(x)$ is continuous at X .

Let $\varepsilon = \frac{1}{2}$. By the construction method proposed in Step (1) we can obtain a Lipschitzian functional $\varphi_{\frac{1}{2}}(x)$ such that $d(\varphi_{\frac{1}{2}}(x), F(x)) \leq \frac{1}{2}$ for every $x \in X$. Now we denote $f_1(x) = \varphi_{\frac{1}{2}}(x)$. Let us consider the set-valued mapping $F_1(x) = B(f_1(x), (\frac{1}{2})^2) \cap F(x)$, $F_1(x)$ is lower semi-continuous by Lemma 2.2.2. Taking $\varepsilon = \frac{1}{4}$, by the method proposed in Step (1) again, a Lipschitzian functional $\varphi_{\frac{1}{4}}(x)$ can be constructed such that $d(\varphi_{\frac{1}{4}}(x), F_1(x)) < \frac{1}{4}$. Let $f_2(x) = \varphi_{\frac{1}{4}}(x)$. By definition,

$F_1(x) \subset F(x) \cap B\left(f_1(x), \frac{1}{4}\right)$, therefore $d(f_2(x), F(x)) \leq d(f_2(x), F_1(x)) < \frac{1}{4}$ and

$$\|f_1(x) - f_2(x)\| \leq d(f_1(x), F_1(x)) + d(f_2(x), F_1(x)) \leq \frac{1}{4} + \frac{1}{4} = \frac{1}{2}.$$

Inductively, when we construct $f_n(x)$ which satisfies that $d(f_n(x), F(x)) < \left(\frac{1}{2}\right)^n$ and $\|f_n(x) - f_{n-1}(x)\| \leq \left(\frac{1}{2}\right)^{n-1}$, we let $F_n(x) = B\left(f_n(x), \left(\frac{1}{2}\right)^{n+1}\right) \cap F(x)$, then $F_n(x)$ is lower semi-continuous by Lemma 2.2.2. There is a $f_{n+1}(x) = \varphi_{\frac{1}{2^{n+1}}}(x)$ which is Lipschitzian and satisfies $d(f_{n+1}(x), F_n(x)) < \left(\frac{1}{2}\right)^{n+1}$. By the definition of $F_n(x)$, $d(f_{n+1}(x), F(x)) \leq d(f_{n+1}(x), F_n(x)) < \left(\frac{1}{2}\right)^{n+1}$ and $\|f_{n+1}(x) - f_n(x)\| \leq \left(\frac{1}{2}\right)^n$. Thus,

$$\begin{aligned} \|f_{n+k}(x) - f_n(x)\| &\leq \|f_{n+k}(x) - f_{n+k-1}(x)\| + \cdots + \|f_{n+1}(x) - f_n(x)\| \\ &< \left(\frac{1}{2}\right)^{n+k-1} + \cdots + \left(\frac{1}{2}\right)^n \\ &< \left(\frac{1}{2}\right)^{n-1}. \end{aligned}$$

$\{f_n(x)\}$ is a Cauchy sequence.

(3) The selection $f(x)$ is obtained.

Let $f(x) = \lim_{n \rightarrow \infty} f_n(x)$. Then $f(x)$ is continuous. Moreover, the space Y is complete and $F(x)$ is a closed set; consequently, $f(x) \in F(x)$. $\square$

2.2.3 Lipschitzian Approximation

If f is a continuous single-valued mapping, then f may not be a Lipschitzian mapping. But its opposite statement is true. However, a Lipschitzian set-valued mapping may not be continuous. This subsection will prove that an upper semi-continuous set-valued mapping can be approximated by a sequence of Lipschitzian mappings.

Theorem 2.2.4 Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be an upper semi-continuous set-valued mapping. It is bounded and is with closed and convex value. Then there exist set-valued mappings $F_k : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $k = 1, 2, \dots$ which satisfy the following requirements.

- (1) $F_k : \mathbb{R}^n \rightarrow \mathbb{R}^m$, $k = 1, 2, \dots$ are all Lipschitzian mappings and are bounded and with closed and convex values.
- (2) $F(x) \subset \cdots \subset F_{k+1}(x) \subset F_k(x) \subset \cdots \subset F_1(x) \subset F_0(x)$.
- (3) For every $\varepsilon > 0$, there is a $K = K(x, \varepsilon)$ such $F_k(x) \subset F(x) + \varepsilon B_m$ provided that $k > K$.

Proof Suppose $F(x) \subset bB_m$ for every $x \in X$ where $b \in \mathbb{R} (> 0)$ and B_m is the open unit ball of $\mathbb{R}^n$. Denote $e_i = [0 \dots 1 \dots 0]$, i.e., its i th component is 1 and others are zeros. $\{e_i; i = 1, 2, \dots, n\}$ is a set of normal orthogonal basis of $\mathbb{R}^n$. Let $\mathbb{Z}^n = \underbrace{\mathbb{Z} \times \mathbb{Z} \times \dots \times \mathbb{Z}}_n \subset \mathbb{R}^n$ be the set in which components of every vector are all integers. The following proof consists of three steps.

(1) Constructing a sequence of Lipschitzian set-valued mappings $\{F_k(x)\}$.

Fix $\rho_0 = 1$ and define $\mathbb{Z}_0^n = \rho_0 \mathbb{Z}^n = \mathbb{Z}^n$. In addition, we define a set

$$O_0 = \{x; x = a_1 e_1 + a_2 e_2 + \dots + a_n e_n, a_i \in (-\rho_0, \rho_0)\},$$

O_0 is an open cube whose center is the origin and lengths of edges are all equal to $2\rho_0$. O_0 is obviously convex. $\mathbb{Z}_0^n + O_0 = \{z_0 + O_0, z_0 \in \mathbb{Z}_0^n\}$ is qualified as a locally finite open covering of $\mathbb{R}^n$. Then the open covering yields a Lipschitzian unit decomposition $p_{z_0}(x)$. For every $z_0 \in \mathbb{Z}_0^n$, we can have a set $C_{z_0} = \text{cl co } F(z_0 + 2O_0)$, then set-valued mapping $F_0(x)$ is defined as follows:

$$F_0(x) = \sum_{z_0 \in \mathbb{Z}_0^n} p_{z_0}(x) C_{z_0}.$$

Because $p_{z_0}(x)$ is a local Lipschitzian function, $F_0(x)$ is then a Lipschitzian set-valued mapping and with closed and convex value. $F(x) \subset bB_m$ leads to that $C_{z_0} \subset bB_m$, moreover, $F_0(x) \subset bB_m$ since $p_{z_0}(x)$ is a unit decomposition.

We now construct $F_1(x)$. Define $\rho_1 = \frac{1}{3}\rho_0$ and $\mathbb{Z}_1^n = \rho_1 \mathbb{Z}_0^n$. The element of $\mathbb{Z}_1^n$ can be denoted by $z_1 = \rho_1 z_0$ with $z_0 \in \mathbb{Z}_0^n$. A set O_1 is defined below

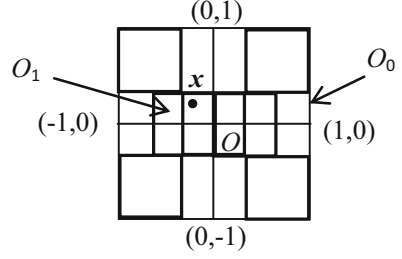
$$O_1 = \left\{ x; x = a_1 e_1 + a_2 e_2 + \dots + a_n e_n, a_i \in \left(-\frac{1}{3}, \frac{1}{3} \right) \right\}.$$

Then $\mathbb{Z}_1^n + O_1$ forms an open covering of $\mathbb{R}^n$ and it is locally finite. The open covering holds a unit partition $p_{z_1}(x)$, $z_1 \in \mathbb{Z}_1^n$. For every $z_1 \in \mathbb{Z}_1^n$, define a set $C_{z_1} = \text{cl co } F(z_1 + 2O_1)$ and a set-valued mapping $F_1(x)$ can be obtained as follows:

$$F_1(x) = \sum_{z_1 \in \mathbb{Z}_1^n} p_{z_1}(x) C_{z_1}.$$

$F_1(x) \subset bB_m$ and is also Lipschitzian and with closed and convex value.

Fig. 2.6 The relation of O_0 and O_1



Generally, in the k th step, we obtain a set-valued mapping $F_k(x)$, which is Lipschitzian and with closed and convex value, and is bounded by bB_m . We then turn to the $(k+1)$ th step. By defining $\rho_{k+1} = \frac{1}{3}\rho_k$, the set-valued mapping $F_{k+1}(x)$ can be defined with a similar way, i.e., we define $\mathbb{Z}_{k+1}^n = \rho_{k+1}\mathbb{Z}_0^n$ and O_{k+1} , then we obtain a unit partition $p_{z_{k+1}}(x)$, $z_{k+1} \in \mathbb{Z}_{k+1}^n$ and closed sets $C_{z_{k+1}} = \text{cl co } F(z_{k+1} + 2O_{k+1})$ for each $z_{k+1} \in \mathbb{Z}_{k+1}^n$, and at last we define

$$F_{k+1}(x) = \sum_{z_{k+1} \in \mathbb{Z}_{k+1}^n} p_{z_{k+1}}(x) C_{z_{k+1}}.$$

(2) Proof of $F(x) \subset F_{k+1}(x) \subset F_k(x)$.

If we can prove $F(x) \subset F_1(x) \subset F_0(x)$, then by a similar way, we can prove $F(x) \subset F_{k+1}(x) \subset F_k(x)$.

For a fixed $x \in \mathbb{R}^n$, there are only finite $z_{0i} \in \mathbb{Z}_0^n$, $i = 1, 2, \dots, N_{0x}$ such that $x \in z_{0i} + O_0$. Denote $Z_0(x) = \{z_{0i}; i = 1, 2, \dots, N_{0x}\}$ for the x . Similarly, there are only finite $z_{1j} \in \mathbb{Z}_1^n$, $j = 1, 2, \dots, N_{1x}$ such that $x \in z_{1j} + O_1$, and denote $Z_1(x) = \{z_{1j}; j = 1, 2, \dots, N_{1x}\}$. A scheme diagram for the relation of $Z_0(x)$ and $Z_1(x)$ is given in Fig. 2.6.

We now prove that for every $z_{0i} \in Z_0(x) \subset \mathbb{Z}_0^n$ and every $z_{1j} \in Z_1(x) \subset \mathbb{Z}_1^n$, $z_{1j} + 2O_1 \subset z_{0i} + 2O_0$.

Let $y \in z_{1j} + 2O_1$. Then $\|y - z_{1j}\| \leq 2\rho_1 = \frac{2}{3}\rho_0$. In addition, we have

$$\|z_{0i} - z_{1j}\| \leq \|z_{0i} - x\| + \|x - z_{1j}\| \leq \rho_0 + \rho_1 = \frac{4}{3}\rho_0.$$

Therefore, $\|y - z_{0i}\| \leq \|y - z_{1j}\| + \|z_{1j} - z_{0i}\| \leq \frac{2}{3}\rho_0 + \frac{4}{3}\rho_0 = 2\rho_0$, i.e., $y \in z_{0i} + 2O_1$.

The relation $z_{1j} + 2O_1 \subset z_{0i} + 2O_0$ implies $C_{z_1} \subset C_{z_0}$. C_{z_1} is convex; hence,

$$C_{z_1} = \sum_{z_0 \in Z_0(x)} p_{z_0}(x) C_{z_1} \subset \sum_{z_0 \in Z_0(x)} p_{z_0}(x) C_{z_0}.$$

It follows that,

$$\begin{aligned}
 F_1(x) &= \sum_{z_1 \in \mathbb{Z}_1^n} p_{z_1}(x) C_{z_1} = \sum_{z_1 \in Z_1(x)} p_{z_1}(x) C_{z_1} \\
 &\subset \sum_{z_1 \in Z_1(x)} p_{z_1}(x) \sum_{z_0 \in Z_0(x)} p_{z_0}(x) C_{z_0} \\
 &= \sum_{z_1 \in Z_1(x)} p_{z_1}(x) \sum_{z_0 \in \mathbb{Z}_0^n} p_{z_0}(x) C_{z_0} \\
 &= \sum_{z_1 \in Z_1(x)} p_{z_1}(x) F_0(x) \\
 &= F_0(x).
 \end{aligned}$$

Simultaneously, when $x \in \mathbb{R}^n$ and $z_{1j} \in Z_1(x)$, $x \in z_{1j} + O_1$ and $F(x) \subset C_{z_1}$.

$$F(x) \subset \sum_{z_1 \in Z_1(x)} p_{z_1}(x) C_{z_1} = \sum_{z_1 \in \mathbb{Z}_1^n} p_{z_1}(x) C_{z_1} = F_1(x).$$

(3) Convergence of $\{F_k(x)\}$.

At last, we verify the convergence of $F_k(x)$. Because $F(x)$ is upper semi-continuous for $x \in \mathbb{R}^n$ and $\varepsilon > 0$, there is a $\delta = \delta(x, \varepsilon)$ such that if $y \in B(x, \delta)$, $F(y) \subset F(x) + \frac{\varepsilon}{2} B_m$. For the x and ε , there is a $K = K(x, \varepsilon)$, $x + O_{K(x, \varepsilon)} \subset B(x, \delta)$. For every $k > K$, denote $Z_k(x) = \{z_{kj}; j = 1, 2, \dots, N_{kx}\}$, where $x \in z_{kj} + O_k$. Thus, we have

$$z_{kj} + 2O_k \subset x + 3O_k \subset x + O_K.$$

It illustrates that $y \in z_{kj} + 2O_k$ for every $z_{kj} \in Z_k(x)$, $F(y) \subset F(x) + \frac{\varepsilon}{2} B_m$. Furthermore,

$F(z_{kj} + 2O_k) \subset F(x) + \frac{\varepsilon}{2} B_m$. $C_{z_{kj}} = \text{clco} F_k(z_{kj} + 2O_k) \subset F(x) + \frac{\varepsilon}{2} \bar{B}_m \subset F(x) + \varepsilon B_m$. $F(x) + \varepsilon B_m$ is convex, hence, when $k > K$,

$$F_k(x) = \sum_{z_k \in \mathbb{Z}_k^n} p_{z_k}(x) C_{z_k} = \sum_{z_{k_j} \in Z_k(x)} p_{z_{k_j}}(x) C_{z_{k_j}} \subset F(x) + \varepsilon B_m.$$

The theorem is now verified. $\square$

From Problem 7 of this section the readers will conclude if the conditions given in Theorem 2.2.4 are satisfied, then for every $\varepsilon > 0$, there is a single-valued mapping $f_\varepsilon : \mathbb{R}^n \rightarrow \mathbb{R}^m$ which is continuous and $f_\varepsilon(x) \in F(x) + \varepsilon B$. The fact leads to the following definition.

Definition 2.2.2 Suppose that X and Y are two metric spaces, and $F : X \rightarrow Y$ is a set-valued mapping. If for every $\varepsilon > 0$, there is a single-valued mapping $f_\varepsilon : X \rightarrow Y$ such that $\text{gra } f_\varepsilon \subset \text{gra } F + \varepsilon B$, then F is approximately selectable, and f_ε is an ε -approximate selection. $\square$

From Definition 2.2.2, $f_\varepsilon(x) \in F(x) + \varepsilon B$ is only a sufficient condition for the approximate selection. Let us consider the following example.

Example 2.2.1 Consider a set-valued mapping $F : \mathbb{R} \rightarrow \mathbb{R}$

$$F(x) = \begin{cases} -1, & x < 0, \\ [-1, 1], & x = 0, \\ 1, & x > 0. \end{cases}$$

$F(x)$ is a monotonous mapping which will be studied at the end of this chapter. It is obvious that $F(x)$ has no a continuous selection. But the sigmoid function

$$f_\beta(x) = \frac{e^{\beta x} - e^{-\beta x}}{e^{\beta x} + e^{-\beta x}}$$

is continuous. It is direct to show that for $1 > \varepsilon > 0$, when $\beta > \frac{1}{2\varepsilon} \ln \frac{2-\varepsilon}{\varepsilon}$, $f_\varepsilon(x)$ is an ε -approximate selection (Fig. 2.7). $\square$

Example 2.2.1 really gives a new way for the selection since for every $1 > \varepsilon > 0$, except $x = 0$, $f_\varepsilon(x) \cap F(x) = \emptyset$. Furthermore, it can be seen that no matter how small ε is determined, there still is a $\delta > 0$, when $x \in (0, \delta)$, $d(f_\varepsilon(x), F(x)) > \frac{1}{2}$, or, in other words, $f_\varepsilon(x) \notin F(x) + \varepsilon B$.

Theorem 2.2.5 Let X be a metric space, Y be a Banach space, and $F : X \rightarrow Y$ be a set-valued mapping, which is upper semi-continuous and with closed and convex value. Then F holds an ε -approximate selection; moreover, the selection is a Lipschitzian mapping.

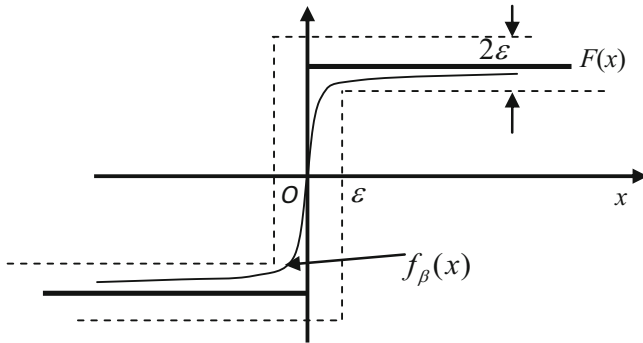


Fig. 2.7 ε -approximate selection of Example 2.2.1

Proof $F(x)$ is upper semi-continuous, then for a given $\varepsilon > 0$ and every $x_0 \in X$, there is a $\delta_1 = \delta_1(x_0, \varepsilon)$ such that when $x \in B(x_0, \delta_1)$, $F(x) \subset F(x_0) + \frac{\varepsilon}{2}B_Y$. Let $\delta(x_0) = \min(\frac{\varepsilon}{2}, \delta_1)$. Then these neighborhoods $B(x_0, \frac{\delta(x_0)}{4})$ form an open covering of X if x_0 goes over the space X . X is a metric space; hence, there exists a fined locally finite sub-covering $\{U_\alpha\}$. It follows that there is a Lipschitzian unit decomposition $p_i(x)$ attached to $\{U_\alpha\}$. Because $\{U_\alpha\}$ is a fined covering of $\{B(x_0, \frac{\delta(x_0)}{4}), x_0 \in X\}$, for every $U_i \in \{U_\alpha\}$, we can find a x_i such that $U_i \subset B(x_i, \frac{\delta(x_i)}{4})$. We now fix the x_i for U_i and construct a mapping f_ε as follows

$$f_\varepsilon(x) = \sum_i p_i(x) m(F(x_i)), \quad (2.2.3)$$

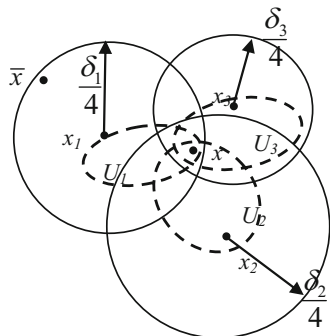
where $m(F(x_i))$ is the minimal norm element in $F(x_i)$. We analyze the properties of the f_ε . At first, because $p_i(x)$ is a Lipschitzian mapping, so is the f_ε . Secondly, because $\text{co gra } F$ is convex, $(x_i, m(F(x_i))) \in \text{co gra } F$ implies then $(x, f_\varepsilon(x)) \in \text{co gra } F$. At last, we prove that $f_\varepsilon(x) \in F(x_j) + \frac{\varepsilon}{2}B_Y$ for some $x_j \in X$.

$\{U_i\}$ is a locally finite open covering; consequently, for every $x \in X$, there are only finite nonzero terms in $\sum_i p_i(x) m(F(x_i))$. For a fixed x , let $I(x) = \{i; p_i(x) \neq 0\} = \{i; x \in U_i\}$, then $I(x)$ is a finite set. By the definition of U_i , an x_i can be found such that $U_i \subset B(x_i, \frac{\delta_i}{4})$. Define $\delta_{i_0} = \max\{\delta_i = \delta(x_i), i \in I(x)\}$, the δ_{i_0} corresponds to $x_{i_0} \in X$. Then for every $x_i, i \in I(x)$, $d(x, x_i) < \frac{\delta(x_i)}{4} \leq \frac{\delta(x_{i_0})}{4}$, and $d(x_{i_0}, x_i) \leq d(x_{i_0}, x) + d(x, x_i) < \frac{\delta(x_{i_0})}{2}$, furthermore, $U_i \subset B(x_{i_0}, \delta_{i_0}), i \in I(x)$. Hence, for $i \in I(x)$

$$m(F(x_i)) \in F(x_i) \subset F(x_{i_0}) + \frac{\varepsilon}{2}B_Y. \quad (2.2.4)$$

Figure 2.8 is used to illustrate the relations of x and U_i, x_i, δ_i . $x \in U_i \subset B(x_i, \frac{\delta_i}{4})$, then $d(x_i, x_2) \leq d(x_i, x) + d(x, x_2) < \frac{\delta_2}{2}$ (in Fig. $\delta_2 = \max\{\delta_i, i \in I(x)\}$), i.e., $x_i \in$

Fig. 2.8 The relation of x and U_i



$B\left(x_2, \frac{\delta_2}{2}\right)$, and a point $\bar{x} \in B(x_i, \delta_i)$, $d(\bar{x}, x_2) \leq d(\bar{x}, x_i) + d(x_i, x_2) < \frac{\delta_2}{4} + \frac{\delta_2}{2} < \delta_2$. Thus, $B(x_2, \delta_2) \supset B(x_i, \delta_i) \supset U_i$.

We have assumed that $F(x_{i_0})$ is convex, hence, from Relation (2.2.4)

$$f_\varepsilon(x) = \sum_i p_i(x) m(F(x_i)) \in F(x_{i_0}) + \frac{\varepsilon}{2} B_Y,$$

i.e., there exists a $y_{i_0} \in F(x_{i_0})$ such that $\|f_\varepsilon(x) - y_{i_0}\| < \frac{\varepsilon}{2}$. It follows that

$$d((x, f_\varepsilon(x)), \text{gra } F) \leq d((x, f_\varepsilon(x)), (x_{i_0}, y_{i_0})) \leq d(x, x_{i_0}) + d(f_\varepsilon(x), y_{i_0}) < \varepsilon.$$

□

We sum up the procedure of proof of Theorem 2.2.5. For a giving $x \in X$ and $\varepsilon > 0$, the target is to construct $f_\varepsilon(x)$ the image of x . A neighbourhood $B(x, \delta)$ is constructed firstly for the upper semi-continuous, then an open covering $\{B(x, \frac{\delta}{4})\}$ is applied to yield a locally finite fined open covering $\{U_\alpha\}$. $\{U_\alpha\}$ yields a Lipschitzian unit decomposition and a set $I(x)$ which is finite. By $i \in I(x)$, x_i is found. At last $m(F(x_i))$ is used to construct the $f_\varepsilon(x)$. To prove the conclusion, let $\delta_{i_0} = \max\{\delta_i, i \in I(x)\}$, then $B(x_{i_0}, \delta_{i_0}) \supset B(x_i, \frac{\delta_i}{4})$. It leads to $m(F(x_i)) \in F(x_{i_0}) + \frac{\varepsilon}{2} B_Y$ and $f_\varepsilon(x) \in F(x_{i_0}) + \frac{\varepsilon}{2} B_Y$.

It is meaningful to compare Theorem 2.2.5 with Theorem 2.2.4. Theorems 2.2.4 and 2.2.5 do not require the spaces X and Y to be finite dimensions. But, Theorem 2.2.5 does not require $F(x)$ is bounded, because the vector selected in Theorem 2.2.5 is $m(F(x_i))$ which holds determined meaning. However, Theorem 2.2.5 only provides an approximate selection.

A mapping $\varphi : X \rightarrow Y$ is said to be locally compact if for every $x \in X$ there is a compact set $K_x \subset Y$.³ and a $\delta > 0$ such that $\varphi(B(x, \delta)) \subset K_x$. Accordingly, we can define locally bounded. Let $\varphi_\varepsilon : X \rightarrow Y$ be mappings, $\varphi_\varepsilon(x)$ is locally equicontact if there is a $\varepsilon_0 > 0$, when $\varepsilon_0 > \varepsilon > 0$, there exists a δ which may depend on x but be independent of ε such that $\varphi_\varepsilon(B(x, \delta)) \subset K_x$.

Corollary 2.2.2 If all of the conditions of Theorem 2.2.5 are satisfied and the mapping $x \mapsto m(F(x))$ is locally compact, then the selection $f_\varepsilon(x)$ is locally equicontact.

Proof Because $x \mapsto m(F(x))$ is locally compact, by the definition of local compactness, there is an $\eta_x > 0$, such that $m(F(B(x, \eta_x))) \subset C_x$ where C_x is a compact set. Take $\varepsilon_0 = \frac{\eta_x}{2}$ and $\delta_0 = \frac{\eta_x}{4}$. Now let $\varepsilon \in (0, \varepsilon_0)$, we prove

³The subscript x is used to express that the compact set K depends on variable x . η_x has the same meaning.

$f_\varepsilon(B(x, \delta_0)) \subset K_x$. For every $\bar{x} \in B(x, \delta_0)$, the above proof asserts $f_\varepsilon(\bar{x}) = \sum p_i(\bar{x})m(F(\bar{x}_i))$, where $\bar{x}_i \in B(\bar{x}_i, \frac{\delta_i}{4}) \subset B(\bar{x}, \frac{\delta_{i0}}{2}) \subset B(\bar{x}, \frac{\delta_0}{2})$. The last containing relation comes from the fact that all $\delta(x) < \frac{\varepsilon}{2} < \frac{\varepsilon_0}{2} = \delta_0$. Furthermore, $\bar{x} \in B(x, \delta_0)$, hence $\bar{x}_i \in B(x, 2\delta_0) \subset B(x, \eta_x)$ and $m(F(\bar{x}_i)) \in C_x$. Let $K_x = \text{co}C_x$. Then K_x is compact and $f_\varepsilon(\bar{x}) \in K_x$. $\square$

2.2.4 Theorems for Fixed Points

To end this section, we deal with the fixed points. The existence of fixed points is an important issue in both topology and functional analysis. We will extend these results to set-valued mappings. We start with a new definition for selection.

Definition 2.2.3 Let $F : X \rightarrow Y$ be a set-valued mapping, and X, Y be two normed spaces. If there exist set-valued mappings $F_k : X \rightarrow Y, k = 1, 2, \dots$ such that

- (1) For every k , $\text{gra } F_k$ is a closed set, and F_k holds a continuous selection;
- (2) For every $x \in X, F_{k+1}(x) \subset F_k(x)$ and $\bigcap_{k=1}^{\infty} F_k(x) = F(x)$;
then the F is said to be σ -selectable and $\{F_k, k = 1, 2, \dots\}$ is a σ -selectable approximate sequence. $\square$

By Theorem 2.2.4, if $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is an upper semi-continuous mapping, and it is bounded and with closed and convex value, then F is σ -selectable.

Definition 2.2.4 Let $f : X \rightarrow X$ be a single-valued mapping. $x \in X$ is said to be a fixed point of f , if $f(x) = x$.

Let $F : X \rightarrow X$ be a set-valued mapping. $x \in X$ is said to be a fixed point of F , if $x \in F(x)$. $\square$

It is obvious, from Definition 2.2.4, that the concept of fixed point for set-valued mapping is exactly an extension of that for single-valued mapping.

Theorem 2.2.6 is known as Brouwer fixed point theorem and is a famous result.

Theorem 2.2.6 Let $A \subset \mathbb{R}^n$ be convex and compact set, $f : A \rightarrow A$ is a continuous single-valued mapping, then there is a $x \in A$ such that $f(x) = x$. $\square$

The conclusion was extended to general normed spaces by Schauder. Their proofs are quite complicated and omitted.

Theorem 2.2.7 Let $A \subset \mathbb{R}^n$ be a convex and compact set, $F : A \rightarrow A$ be a σ -selectable set-valued mapping, then there is a $x \in A$ such that $x \in F(x)$.

Proof Let $\{F_k(x)\}$ be the σ -selectable approximate sequence and $f_k(x)$ is a continuous selection of $F_k(x)$. Consider a continuous mapping from A to A defined as $x \mapsto \pi(f_k(x), A)$. By Theorem 2.2.6, there is a fixed point $x_k \in A$ such that $x_k = \pi(f_k(x_k), A)$, i.e., $\|f_k(x_k) - x_k\| = d(f_k(x_k), A)$.

Because $F_k \downarrow F$, when $m \leq k$, we have $f_k(x_k) \in F_k(x_k) \subset F_m(x_k)$ and

$$d(x_k, F_m(x_k)) \leq d(x_k, F_k(x_k)) \leq \|f_k(x_k) - x_k\| = d(f_k(x_k), A)$$

A is a compact set and $\{x_k\} \subset A$, hence, $\{x_k\}$ has an accumulation $\bar{x} \in A$. Without loss of generality, we assume $x_k \rightarrow \bar{x}$. By Problem 5 of the last section, F_m is an upper semi-continuous set-valued mapping; consequently, for a given $\varepsilon > 0$, we have $F_m(x_k) \subset F_m(\bar{x}) + \varepsilon B_n$ for some large $k (> m)$. Thus,

$$d(\bar{x}, F_m(\bar{x})) \leq \|\bar{x} - x_k\| + d(x_k, F_m(x_k)) + \varepsilon \leq \|\bar{x} - x_k\| + d(f_k(x_k), A) + \varepsilon \quad (2.2.5)$$

$F_m \downarrow F$, it leads to $F_m(\bar{x}) \subset F(\bar{x}) + \varepsilon B_n$ for some large m . Therefore,

$$f_k(x_k) \in F_k(x_k) \subset F_m(x_k) \subset F_m(\bar{x}) + \varepsilon B_n \subset F(\bar{x}) + 2\varepsilon B_n \subset A + 2\varepsilon B_n$$

i.e., $d(f_k(x_k), A) \leq 2\varepsilon$. Substituting it into Inequality (2.2.5) results in $d(\bar{x}, F_m(\bar{x})) \leq 4\varepsilon$. By $F_m(\bar{x}) \subset F(\bar{x}) + \varepsilon B_n$ again, $d(\bar{x}, F(\bar{x})) \leq 5\varepsilon$. Because ε can be selected arbitrarily, we obtain $\bar{x} \in F(\bar{x})$. $\square$

By Theorem 2.2.4, a direct corollary can be obtained.

Corollary 2.2.3 Let $A \subset \mathbb{R}^n$ be a convex and compact set, $F : A \rightarrow A$ be an upper semi-continuous set-valued mapping and with convex and closed value. Then F holds a fixed point on A . $\square$

If $\mathbb{R}^n$ is replaced by a Banach space Y , Corollary 2.2.3 is still valid and known as Kakutani fixed point theorem. We restate it below.

Theorem 2.2.8 Let Y be a Banach space and $A \subset Y$ be convex and compact set. If $F : A \rightarrow A$ is an upper semi-continuous set-valued mapping and with closed and convex value, then F holds a fixed point on A . $\square$

Problems

1. If $A_1, A_2 \subset \mathbb{R}^n$ and $A_1 \subset A_2 + \varepsilon B$, then for every $x_0 \in \mathbb{R}^n$, $d(x_0, A_1) \leq d(x_0, A_2) + \varepsilon$.
2. Let X be a normed space, $A \subset X$ be a set, then $d(x, A)$ is a Lipschitzian functional of $X \rightarrow \mathbb{R}$ and its Lipschitzian constant can be 1.
3. If $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a Lipschitzian set-valued mapping and with closed and convex value. Furthermore, for every $x \in \mathbb{R}^n$, $F(x)$ is bounded then $F(x)$ is ε -continuous.
4. Prove Lemma 2.2.3.
5. Let $A \subset \mathbb{R}^n$ be a convex and compact set, x_0 be an inner point of A . $f : A \rightarrow \mathbb{R}^n$ is a continuous single-valued mapping. If for every $x \in \text{bd } A$, $\|f(x) - x\| \leq \|x_0 - x\|$ then $x_0 \in f(A)$. (Hint: applying Brouwer fixed point theorem and Minkowski function.)

6. Give an example to show that $f_\varepsilon(x) \in F(x) + \varepsilon B$ is not a necessary condition for the approximate selection.
7. Suppose that all conditions given in Theorem 2.2.4 are satisfied, then for every $\varepsilon > 0$, there exists a single-valued mapping $f_\varepsilon : \mathbb{R}^n \rightarrow \mathbb{R}^n$ such that $f_\varepsilon(x) \in F(x) + \varepsilon B$. (Hint: showing the set-valued mapping constructed in Theorem 2.2.4 is also lower semi-continuous)
8. Try to verify Theorem 2.2.4 is still valid if all conditions given in Theorem 2.2.4 hold but F is replaced that $F : X \rightarrow Y$ where X and Y are all Banach spaces.

2.3 Differential Inclusions and Existence Theorems

This section deals with differential inclusion. It starts with the definition which is presented by a comparison with the differential equation. Then motivation of investigation of differential inclusions is presented. The main content of this section is existence theorems of solutions of differential inclusions. At last we extend the conclusion to time-delayed differential inclusions. The extension can show the advantages of differential inclusion theory.

2.3.1 Differential Equations and Differential Inclusions

The differential equation considered takes an explicit form that

$$\dot{x} = f(t, x) \quad (2.3.1)$$

where $f : [t_0, t_1] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a single-valued mapping, $t \in [t_0, t_1] \subset \mathbb{R} (\geq 0)$ is the interval of time, $x : [t_0, t_f] \rightarrow \mathbb{R}^n$ is a derivable function to be solved and $\dot{x}$ is its derivate related to time t . $[t_0, t_f]$ is the interval where the solution exists. t_0 is the initial time and $t_f \leq t_1$ is the final time. From geometrical viewpoint, $f(t, x)$ gives a vector field in $\mathbb{R}^n$. Solving Eq. (2.3.1) is equivalent to looking for a curve which, in time t , passes through $x \in \mathbb{R}^n$, with its tangent $f(t, x)$. Thus every vector in $\mathbb{R}^n$ yields a curve, i.e., a solution. If Eq. (2.3.1) satisfies the uniqueness condition, then these curves do not intersect. The fact illustrates the solution of Eq. (2.3.1) contains a vector of $\mathbb{R}^n$ as its parameter. Such a solution is called general solution. Usually for the initial time t_0 the starting vector $x(t_0)$ is assigned. When the initial condition is given, Eq. (2.3.1) becomes

$$\dot{x} = f(t, x), x(t_0) = x_0.$$

Solving this differential equation with initial condition is called Cauchy problem. The solution of Cauchy problem is called the special solution.

If f in Eq. (2.3.1) does not explicitly contain the variable t , i.e.,

$$\dot{x} = f(x) \quad (2.3.2)$$

the equation is then time-invariant. For the time-invariant differential equation, the initial time is usually fixed at $t_0 = 0$.

The primitive target for the investigation of differential equation is to get the general solution or special solution $x(t)$ in an analytic form. As well-known for almost all differential equations, we have no effective method to obtain their analytic solutions. Then researchers cleverly transferred their energy to study the qualitative properties of the solutions such as existence, uniqueness, stability, continuities on the parameters, and so on.

If the function $f(t, x)$ in the right side of Eq. (2.3.1) is replaced by a set-valued mapping $F : [t_0, t_1] \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, then Eq. (2.3.1) becomes

$$\dot{x} \in F(t, x) \quad (2.3.3)$$

Relation (2.3.3) is called differential inclusion and denoted by Inc. (2.3.3) from now on. Accordingly, if F is only a mapping from $\mathbb{R}^n$ to $\mathbb{R}^n$, then the differential inclusion is time-invariant. The book mostly considers the time-invariant case, and the inclusion extended from Eq. (2.3.2), i.e.,

$$\dot{x} \in F(x), \quad x(0) = x_0 \quad (2.3.4)$$

To solve Inc. (2.3.4) is called the Cauchy problem of differential inclusion. For the case, the initial time is fixed at $t_0 = 0$. We apply $x(t; [0, t_f], x_0)$ to denote a solution of Inc. (2.3.4). In the notation $x(t; [0, t_f], x_0)$, time t is the argument, x_0 is the initial condition, t_f is the final time and $[0, t_f]$ is the time interval where the solution exists. When there is no any confusion, we apply $x(t, x_0)$ or $x(t)$ to substitute the troublesome notation $x(t; [0, t_f], x_0)$ for simplicity. It is evident Inc. (2.3.4) holds many solution if $F(x)$ is a nontrivial set-valued mapping.

From the definition of the differential inclusion, one may suggest a way to solve the Cauchy problem of Inc. (2.3.4). Firstly, we try to find a selection $f(x) \in F(x)$, then solve the Cauchy problem of the differential equation $\dot{x} = f(x)$, $x(0) = x_0$. It is really a way that we used to prove many conclusions such as the existence of solutions; however, the scheme is almost impracticable to serve to solve the differential inclusions. At first, finding a satisfactory selection with an explicit description borders on a fantasy. Secondly, even if a simple selection has been obtained, we still lack a method to solve the Cauchy problem except some very special cases. Thirdly, even if a solution of the Cauchy problem is obtained, it still cannot analyze the properties of all solutions. Therefore such a way has been abandoned to solve Inc. (2.3.4) by almost all researchers.

There are two main differences from the investigation of the differential equation. At first, the solution of Inc. (2.3.4) is not unique; hence, we usually do not investigate its uniqueness of solution except some special form of Inc. (2.3.4).⁴ Secondly, the

⁴The monotonous differential inclusions are a special case which will be studied at the last section of this chapter.

solution $x(t)$ of a differential equation is required to be differentiable, and Eq. (2.3.1) should be held for every $t \in [t_0, t_f]$. But for the differential inclusion, the solution $x(t)$ is only required to be absolutely continuous; hence, it allows at some time $x(t)$ is not be differentiable. When $x(t)$ is not differentiable, the left side of Inc. (2.3.4) becomes meaningless. Hence, we only require Inc. (2.3.4) is valid for almost all t in the interval $[0, t_f]$, i.e., it is allowed Inc. (2.3.4) to fail on a set whose measure is zero. The relaxed requirement brings lots advantages for the research of differential inclusions.

$S_{[0,T]}(F, x_0)$ is used to denote the set of solutions of Inc. (2.3.4) where $T = \min\{t_f\}$ and $[0, T]$ is the common interval where the solutions all exist, i.e., $S_{[0,T]}(F, x_0) = \bigcup_{x \in F(x)} x(t; [0, T], x_0)$. Moreover, the set

$$S_{[0,T]}(F, C) = \bigcup_{x \in F(x)} \{x(t, [0, T], x_0), x_0 \in C\}$$

is denoted for the set of solutions whose initial values are in set $C \subset \mathbb{R}^n$. If the time interval is not a key issue in the discussing problem, we often omit the low subscript $[0, T]$.

The following example is given to illustrate the solutions of differential inclusion.

Example 2.3.1 Consider the following differential inclusion

$$\dot{x} \in [0.1, 0.3]x, \quad x(0) = 1. \quad (2.3.5)$$

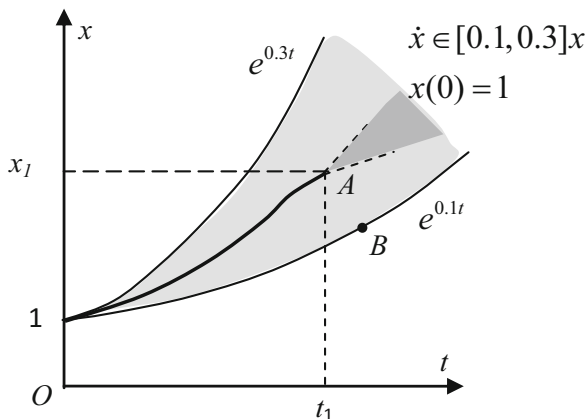
The solutions of Inclusion (2.3.5) are full of the shadow area in Fig. 2.9. But it is not true that every curve in the shadow area is its solution. The trajectories in $S(F, 1)$ hold the following features: (1) the curve is derivable at almost all points on the curve; (2) all trajectories start at $(0,1)$; (3) if one trajectory arrives at $A(t_1, x_1)$ (Fig. 2.9), then the trajectory will be restricted at the darker shadow area where the starting angle is restricted between $0.1x$ and $0.3x$. Hence the trajectory is always monotonically increasing; (4) two curves on the boundary of the shadow area, i.e., $x = e^{0.1t}$ and $x = e^{0.3t}$ are solutions, moreover, once a solution $x(t_1) = B$ where B (Fig. 2.9) is on the boundary, then the $x(t)$ is always at the boundary before t_1 . $\square$

We note that, throughout of the book, the set-valued mapping $F(x)$ in Inc. (2.3.4) is with convex and closed value. The integration adopted is in the meaning of Lebesgue, and $x(t) \in AC([0, T], \mathbb{R}^n)$ or $x(t) \in AC$ i.e., $x(t)$ is an absolutely continuous single-valued mapping.

2.3.2 Why Do We Propose Differential Inclusions?

Maybe there is a question for some readers: Why do we need to consider the differential inclusions? The differential equation is quite powerful so that lots of natural phenomena can be described by differential equations precisely. For

Fig. 2.9 The solutions area of Example 2.3.1



example, the motion of objects can be described by using Newton's laws which are described by differential equations. Another example is about the circuit. Because capacitance, inductance and operational amplifier all can be depicted by differential (or integral) equations, almost all circuits can be described by using differential equations. Kirchhoff's law provides a method to obtain the differential equation. Hence, do we really need differential inclusions?

The following examples are given to explain the reason why we really need to introduce differential inclusions.

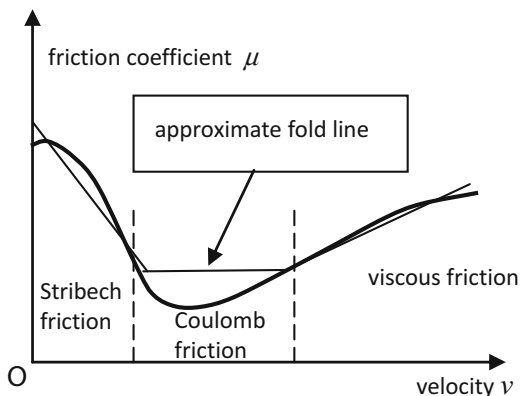
1. Need from real systems

It was known for us every mechanical system suffers from friction and every modern electronic system has to apply diodes. These two basic elements should be depicted by differential inclusions. We now give a detailed discussion with friction in mechanical system.

In 1902, Stribeck studied the dry friction (friction for the objects without using lubricant) of a friction pair. He found that the change of friction coefficient has four phases (Fig. 2.10), i.e., static friction which exists when no motion happens in the friction pair; Stribeck friction called by later researchers which happens in a very lower velocity, with the speed increasing the friction coefficient decreases; Coulomb friction, where the friction coefficient reaches its minimal value; viscous friction, where the friction efficient increases with the speed increasing. He then concluded that the friction is a very complicatedly nonlinear phenomenon. The most undetermined case is the first phase where the friction coefficient is zero but the function of friction exists.

In 2001, Glocker suggested the friction force λ should be depicted by the following set-valued mapping

$$\lambda \in F_f(\dot{q}) = -\mu\lambda_N \operatorname{sgn}(\dot{q}) + F_S(\dot{q})$$

Fig. 2.10 Stribeck plot

where q is the displacement of object, $\dot{q}$ is then the velocity, the sliding friction $F_S(\dot{q})$ is a single-valued function of $\dot{q}$, μ is its friction coefficient, λ_N is the pressure and $\text{sgn}(\dot{q})$ is the sign function defined as follows

$$\text{sgn}(\dot{q}) = \begin{cases} 1 & \dot{q} > 0, \\ [-1, 1] & \dot{q} = 0, \\ -1 & \dot{q} < 0. \end{cases} \quad (2.3.6)$$

Then the moving of the mechanical object which suffers from only dry friction can be described by the following equation

$$M\ddot{q} + D\dot{q} + Kq = Su + T\lambda \quad (2.3.7)$$

where M, D, K are the mass matrix, damping matrix and stiff matrix, respectively, q is generalized displacement which contains both translational motion and rotational motion, S is the input matrix and u is the control whose components are forces or moments, T is gain matrix for the friction and λ is the vector of friction where the i th component takes the form of

$$\lambda_i \in -\mu_i \lambda_{Ni} \text{sgn}(T_i^T \dot{q}) + F_{Si}(T_i^T \dot{q})$$

The relation is the friction that happens in the i th touch point. In the above relation, T_i is the i th column of T and $T_i^T \dot{q}$ is the relative sliding moving happens in the i th touch point.

Equation (2.3.7) can be rewritten by its state description

$$\begin{aligned} \dot{x} &= Ax + Gw + Bu, \\ z &= Hx, \\ y &= Cx, \\ w &\in -\varphi(z), \end{aligned} \quad (2.3.8)$$

where $x = \begin{bmatrix} q^T & \dot{q}^T \end{bmatrix}^T$ is the state, z and w are the input and output of set-valued mapping $\varphi(\cdot)$, respectively. The meanings of other matrices can be known from Eq. (2.3.7) and are omitted. Equation (2.3.8) is called Lur  differential inclusion which will be dealt with in Chap. 5.

In classical nonlinear control theory, the nonlinear phenomenon depicted by Eq. (2.3.6) is called by relay nonlinearity, which is very popular in engineering area. Hence, lots of engineering systems can be described by using Lur  differential inclusions.

2. Control system theory

Since 1980s, the development of differential inclusion is motivated by needs of control systems theory. Let us consider a state model of a control system. Generally its state equation is described by a differential equation

$$\dot{x} = f(x, u) \quad (2.3.9)$$

In the almost all textbooks of control theory, there is a paragraph of explanation attached to the equation: *in Eq. (2.3.9), x is the state which is usually an n -dimensional real vector, and $u \in U$ is an m -dimensional control input where U is the admissible control set.* The statement really suggests that if we apply differential inclusion, the state space model should be described as

$$\dot{x} \in F(x, U) = \bigcup_{u \in U} f(x, u) \quad (2.3.10)$$

Inc. (2.3.10) is usually called as state inclusion, by the state inclusion, the model of whole control system becomes

$$\begin{bmatrix} \dot{x} \\ y \end{bmatrix} \in \begin{bmatrix} F(x, U) \\ G(x, U) \end{bmatrix}$$

The second is the output mapping which is a set-valued mapping altered from output equation.

There are some effective investigation for the Inc. (2.3.10). Commonly, the control u is a mapping from time interval $[0, t_f] \subset \mathbb{R} (\geq 0)$ to $\mathbb{R}^r$; hence, U is a set of r -dimensional functions. Without loss of generality, it is commonly assumed that the U is a compact set of piecewise continuous functions. When U is a compact set, the continuity of $f(x, u)$ implies the continuity of set-valued mapping $F(x, U)$. Furthermore, if $f(x, u)$ is a uniform Lipschitzian mapping for every fixed u , then $F(x, U)$ is also a Lipschitzian set-valued mapping. These facts illustrate that $F(x, U)$ inherits lots of properties of $f(x, u)$, and $f(x, u)$ is only a continuous selection of $F(x, U)$. At problems of this section, we will give more properties of Inc. (2.3.10).

In the design of control system, the design target can be often transformed into performance of state $x(t)$. For example, the stability is that every trajectory of Inc. (2.3.10) converges to the origin (it is strongly asymptotically stable, to

see Sect. 2.5). Let $S(2.3.10)$ denote the set of solutions of Inc. (2.3.10), and $x(t)$ be the ideal trajectory satisfying required performance. For the open-loop control, the design of control system consists of two steps. The first step is to check whether or not $x(t) \in S(2.3.10)$; if the first step succeeds, then the second is to find a $u(t) \in U$, such that $\dot{x}(t) = f(x(t), u(t))$. In case $x(t) \notin S(2.3.10)$, we have two ways to complete the design. One is to extend $x(t)$ to a set $\{x_\lambda(t), \lambda \in \Lambda\}$ where x_λ 's are called admissible trajectories, and then to check whether $\{x_\lambda(t)\} \cap S(2.3.10) \neq \emptyset$. Then we try to find $x(t) \in \{x_\lambda(t)\}$ and $u(t) \in U$. The another is to enlarge the control set U such that $\{x_\lambda(t)\} \cap S(2.3.10) \neq \emptyset$. The introduction of slide mode control is a meaningful practice of extension of control ability. For the case of closed-loop control, the admissible control set U is replaced by $U(x)$ where $U(x)$ is the admissible feedback. Repeating the above statements, we have the design procedure of closed-loop system design. Of course, these only give an outline of system design, the check of $\{x_\lambda(t)\} \cap S(2.3.10) \neq \emptyset$ is very troublesome since we cannot obtain an explicit expression of $S(2.3.10)$. However, the analysis of $S(2.3.10)$ can give a guide for the design of controller.

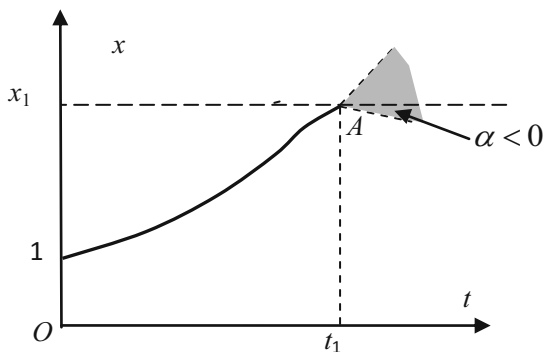
Let us continue to consider Example 2.3.1, the controlled differential inclusion is

$$\begin{aligned}\dot{x} &\in [0.1, 0.3]x + u, \quad u \in U, \\ x(0) &= 1.\end{aligned}$$

The design target is to stabilize the system by feedback, i.e., the trajectory will trend to the t -axis as the time goes. Hence, it is necessary to make the shadow area in Fig. 2.9 can contain an area which is approximate t -axis as the t increases. Suppose, at time t , the trajectory arrives at $A(t_1, x_1)$. If we require the trajectory tends to zero, then the derivative of trajectory has to be less than zero, i.e., $u + 0.1x_1 < 0$, the x_1 can be selected arbitrarily; hence $U \supset \{u; u < -0.1x, t > 0, x > 0\}$ is a necessary condition (Fig. 2.11).

The above discussion shows that using differential inclusion description can extend the design thought for control systems.

Fig. 2.11 The design of control set for Example 2.3.1



3. Discontinuous differential equations

We start with an example.

Example 2.3.2 Consider the Cauchy problem for the following discontinuous differential equation

$$\dot{x} = \begin{cases} 1 & x < 0, \\ -1 & x \geq 0, \end{cases} \quad x(0) = 0 \quad (2.3.11)$$

We consider the trajectory starting at the origin of the (t, x) plane. Because $\dot{x}(0) = -1$, the trajectory should enter the IV quadrant. However, as soon as the trajectory enters the IV quadrant, then $x < 0$ and $\dot{x} > 0$. The trajectory will leave the IV quadrant and enter the I quadrant. In the I quadrant, $x > 0$, and $\dot{x} < 0$, hence, the trajectory should return to the IV quadrant. It is really a trouble. The trajectory cannot also go ahead along with the t -axis since its derivative is nonzero. We conclude that the Cauchy problem has no solution. $\square$

In Example 2.3.2, as long as $x(0) \neq 0$, the Cauchy problem has a solution. If the trajectory arrives at the t -axis at a finite time, the trajectory then finalizes.

To make these equation solvable at interval $[0, \infty)$, Filippov presented an improved scheme which is now called Filippov theory.

For a differential equation $\dot{x} = f(x)$, we define a set-valued mapping

$$F(x) = \bigcap_{\delta > 0} \text{cl co } f(x + \delta B)$$

where B is the open unit ball of $\mathbb{R}^n$. $F(x)$ is called as the set-valued extension of $f(x)$, or the Filippov extension of $f(x)$. The Filippov extension holds the following properties: (1) $F(x)$ is with closed and convex value. Furthermore, if $f(x_0)$ is bounded, then F is upper semi-continuous at x_0 (Prob. 1–8 of this chapter). (2) If $f(x)$ is continuous at x , then $F(x) = \{f(x)\}$. In the next subsection, we will prove the above properties and prove that the Cauchy problem of differential inclusion

$$\dot{x} \in F(x), \quad x(0) = x_0$$

is always solvable. The solution of the above inclusion is called as Filippov solution of discontinuous differential equation $\dot{x} = f(x)$. The fact shows that Filippov extension is a powerful tool for discontinuous differential equations.

Let us return to Example 2.3.2. At $x = 0$, for every $\delta > 0$, $f(0 + \delta B) = \{-1, 1\}$. Hence, $\text{co } f(0 + \delta B) = [-1, 1]$, or $F(0) = [-1, 1]$. Then Eq. (2.3.11) becomes

$$\dot{x} \in F(x) = \begin{cases} -1 & x > 0, \\ [-1, 1] & x = 0, \\ 1 & x < 0. \end{cases}$$

$F(x)$ is continuous and its Cauchy problem $\dot{x} \in F(x)$, $x(0) = x_0$ is solvable. The Filippov solution is

$$x(t) = \begin{cases} x_0 - \operatorname{sgn}(x_0) t & t \in [0, |x_0|], \\ 0 & t \geq |x_0|, \end{cases}$$

where $\operatorname{sgn}(x)$ is the single-valued mapping and $\operatorname{sgn}(0) = 0$.

2.3.3 Existence Theorems of Solution of Differential Inclusions

This subsection deals with the existence of solution of differential inclusions. It is sufficient for us to show that $S(F, x_0) \neq \emptyset$ for Inc. (2.3.4). Because there are more than one definitions of continuity for set-valued mappings, we have to deal with the problem, separately.

In the theory of differential equation, the following conclusion is fundamental for the existence of the solution.

Consider the Cauchy problem of $\dot{x} = f(x)$, $x(0) = x_0$, if there is neighborhood of x_0 , $f(x)$ satisfies the Lipschitzian condition in the neighborhood, then there is a $\tau > 0$, the Cauchy problem is solvable at $[0, \tau]$. Furthermore, if there is neighborhood of $x(\tau)$, $f(x)$ satisfies Lipschitzian condition at that neighborhood, too. Then the solution can be extended.

By the method of using in the proof of Lemma 1.3.1, we can prove that if $f(x)$ is continuous at x_0 and x_0 is a inner point of domain of f , then $f(x)$ is a local Lipschitzian function at x_0 . Therefore, some textbooks revise the theorem as *if $f(x)$ is continuous at x_0 and x_0 is an inner point of domain, then the Cauchy problem $\dot{x} = f(x)$, $x(0) = x_0$ is locally solvable.*

We now turn to differential inclusion.

Theorem 2.3.1 Let X be a Hilbert space, and $F(x) : X \rightarrow X$ be an upper semi-continuous set-valued mapping and with closed and convex value. If the mapping $x \mapsto m(F(x))$ is locally compact, then the Cauchy problem of Inc. (2.3.4) has a solution.

Proof By Theorem 2.2.5, when F is an upper semi-continuous set-valued mapping and with closed and convex value, then F holds an approximate selection which is Lipschitzian. Hence for every $\varepsilon > 0$, there is a $f_\varepsilon(x)$ such that

$$f_\varepsilon(x) \in F(x) + \varepsilon B \quad (2.3.12)$$

Moreover, the selection has the form of

$$f_\varepsilon(x) = \sum_i p_i(x) m(F(x_i))$$

where $p_i(x)$ is a Lipschitzian unit decomposition, $m(F(x_i))$ is the minimal norm element of $F(x_i)$.

Let $\{\varepsilon_n\}$ be a monotonously decreasing series with that $\varepsilon_n \rightarrow 0$ ($n \rightarrow \infty$). For every ε_n there is an approximate selection $f_{\varepsilon_n}(x)$ (for simplicity, denoted by $f_n(x)$). $f_n(x)$ is a locally Lipschitzian mapping; hence, the Cauchy problem $\dot{x} = f_n(x)$, $x(0) = x_0$ has a solution $x_n(t)$.

By Corollary 2.2.2, $\{f_n(x)\}$ is locally equicontact, i.e., there is a $\delta > 0$ and $N \in \mathbb{N}$, such that for $n > N$, $f_n(B(x_0, \delta)) \subset K$, for a compact set K . It is equivalent to that $\dot{x}_n(t) \in K$, $t \in [0, T]$ for some $T > 0$. K is bounded hence the sequence $\{x_n(t)\}$ is equicontinuous at least for $n > N$. By Arzela-Ascoli theorem (to see Sect. 1.1 of this book), there are an absolutely continuous function $x(t)$ and a subsequence $\{x_{n_k}(t)\}$ of $\{x_n(t)\}$ such that $x_{n_k}(t)$ converges to $x(t)$ uniformly. By Theorem 1.1.7 (Alaoglu theorem), $\{\dot{x}_{n_k}(t)\}$ holds a weakly $*$ -convergent subsequence, for simplicity, without loss of generality; $\{\dot{x}_{n_k}(t)\}$ is the weakly $*$ -convergent sequence. Thus,

there is a $\delta > 0$ and $v(t) \in L^1[0, \delta)$ such that $\int_0^t \dot{x}_{n_k}(\tau) d\tau \rightarrow \int_0^t v(\tau) d\tau$.

On the other hand, $x_{n_k}(t) = x_0 + \int_0^t \dot{x}_{n_k}(\tau) d\tau$. Let k trend to infinite. We have

$x(t) = x_0 + \int_0^t v(\tau) d\tau$. Hence, $x(t) \in AC$ and $\dot{x}(t) = v(t)$ almost everywhere.

The approximate selection $f_{n_k}(x)$ satisfies that $\text{gra } f_{n_k} \subset \text{gra } F + \varepsilon_{n_k} B$, i.e.,

$$d[(x_{n_k}(t), f_{n_k}(x(t))), \text{gra } F(x)] \leq \varepsilon_{n_k}$$

Let k trend to infinite again. We have $d[(x(t), \dot{x}(t)), \text{gra } F(x)] = 0$, i.e., $\dot{x}(t) \in F(x(t))$. And $x_{n_k}(0) = x_0$, consequently, $x(0) = x_0$. Thus, we prove the theorem. $\square$

We now apply Theorem 2.3.1 to verify the existence of Filippov solution for discontinuous differential equations.

Theorem 2.3.2 Consider the Cauchy problem of differential equation

$$\dot{x} = f(x), x_0 \in \mathbb{R}^n,$$

where $f(x)$ is bounded, but may be discontinuous. Its Filippov extension is

$$F(x) = \bigcap_{\delta > 0} \text{clco } f(x + \delta B)$$

Then the Cauchy problem differential inclusion

$$\dot{x} \in F(x), x_0 \in \mathbb{R}^n,$$

exists a solution. Moreover, let $x(t, x_0)$ be any solution of the Fillippov extension, then $\dot{x}(t, x_0) = f(x(t, x_0))$ if $f(x)$ is continuous at $x(t, x_0)$.

Proof Because $f(x)$ is bounded by the definition $F(x) = \bigcap_{\delta > 0} \text{cl co } f(x + \delta B)$, $F(x)$ is bounded and with closed and convex value. The local compactness of $x \mapsto m(F(x))$ is guaranteed. We now verify the following two facts.

(1) If $f(x)$ is continuous at x_1 , then $F(x_1) = \{f(x_1)\}$.

Let $y_1 \neq f(x_1)$. Then we denote $\varepsilon = \|y_1 - f(x_1)\|$, there is a $\delta > 0$ such that $f(x_1 + \delta B) \subset B(f(x_1), \frac{\varepsilon}{2})$ since $f(x)$ is continuous at x_1 . Hence $y_1 \notin f(x_1 + \delta B)$. Let $z_i, i = 1, 2, \dots, n$ be arbitrarily n elements in $f(x_1 + \delta B)$. Then for $\lambda_i \geq 0, \sum_{i=1}^n \lambda_i = 1$, we have

$$\left\| \left(\sum_{i=1}^n \lambda_i z_i \right) - f(x_1) \right\| \leq \sum_{i=1}^n \lambda_i \|z_i - f(x_1)\| \leq \frac{\varepsilon}{2}$$

i.e., $y_1 \notin \text{co } f(x_1 + \delta B)$. Furthermore, for $z \in \text{co } f(x_1 + \delta B)$,

$$\|y_1 - z\| \geq \|y_1 - f(x_1)\| - \|f(x_1) - z\| \geq \frac{\varepsilon}{2}$$

i.e., $y_1 \notin \text{clco } f(x_1 + \delta B)$

However, for every $\delta > 0, f(x_1) \in f(x_1 + \delta B)$, moreover, $f(x_1) \in \text{clco } f(x_1 + \delta B)$. It follows $f(x_1) \in F(x_1)$.

(2) $F(x)$ is upper semi-continuous.

This is the result of Problem 8 in the first section of this chapter.

Applying Theorem 2.3.1, the Cauchy problem of differential inclusion $\dot{x} \in F(x), x(0) = x_0$ has a solution. Moreover, by the first fact proved in this theorem, we conclude $\dot{x}(t, x_0) = f(x(t, x_0))$ if $f(x)$ is continuous at $x(t, x_0)$. $\square$

Theorem 2.3.1 needs a quite strong condition, the local compactness of mapping $x \mapsto m(F(x))$, that is difficult to be checked. The following conclusion is about the Lipschitzian differential inclusion. From the conclusion, we can feel that the Lipschitzian condition is critical for both differential equation and differential inclusion.

We will give a theorem which is very useful in the theory of differential inclusion. Before stating the theorem, we present so-called Granwall inequality.

Lemma 2.3.1 Suppose that $x(t) \in AC([0, T], \mathbb{R})$. If

$$|\dot{x}(t)| \leq l(t)|x(t)| + \rho(t),$$

for two single-valued functions $l(t), \rho(t) \in L_1([0, T], \mathbb{R}(\geq 0))$. Then for $t \in [0, T]$, we have

$$|x(t)| \leq e^{\int_0^t l(s)ds} |x(0)| + \int_0^t \rho(s) e^{\int_0^t l(\tau)d\tau - \int_0^s l(\tau)d\tau} ds. \quad (2.3.13)$$

□

Lemma 2.3.1 is known as Gronwall inequality. It is a fundamental integral inequality. Reference (Cai et al. 2009) gives a proof for the inequality. We also leave it to readers as an example since the skill of proof is fundamental in math.

Theorem 2.3.3 Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a closed Lipschitzian set-valued mapping with Lipschitzian constant l , $M \subset AC([0, T], \mathbb{R}^n)$ be a set, and $r_0 : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a continuous single-valued mapping. If there is a function $\rho(t) \in L_1([0, T], \mathbb{R}(\geq 0))$ such that for every $x(t) \in M$,

$$d(\dot{x}(t), F(x(t))) \leq \rho(t),$$

and a real number $\delta \in \mathbb{R}(\geq 0)$ such that $\|r_0(x) - x\| \leq \delta$ for every $x \in \mathbb{R}^n$. Then, there is a mapping $r : M \rightarrow S_{[0, T]}(F, C)$ ⁵ where $C = \{r_0(x(0)), x(t) \in M\}$ is a set of $\mathbb{R}^n$. The mapping r holds the following properties:

1. r is continuous;
2. $r(x(t))(0) = r_0(x(0))$ for every $x(t) \in M$;
3. If $x(t) \in M \cap S(F, C)$ and $r_0(x(0)) = x(0)$, then $r(x(t)) = x(t)$;

4. Denote $\xi(t) = \delta e^{l|t|} + \left| \int_0^t e^{l(|t|-|s|)} \rho(s) ds \right|$,⁶ then

$$\|r(x(t)) - x(t)\| \leq \xi(t), \quad (2.3.14)$$

and

$$\left\| \frac{d}{dt} r(x(t)) - \dot{x}(t) \right\|_{\mathbb{R}^n} \leq l\xi(t) + \rho(t) \quad (2.3.15)$$

for every $t \in [0, T]$.

Proof This is a constructive proof, it consists of four steps.

(1) Construction of function sequence $\{x_k(t)\}$

Let $x(t) \in AC([0, T], \mathbb{R}^n)$. Then a mapping $\omega(x(t)) = \pi(\dot{x}(t), F(x(t)))$ is defined. Because

⁵Because $S_{[0, T]}(F, C) \subset AC([0, T], \mathbb{R}^n)$, r is a mapping from $AC([0, T], \mathbb{R}^n)$ to $AC([0, T], \mathbb{R}^n)$.

⁶In the current situation, the absolute value is useless. However, if the interval $[0, T]$ is replaced by $[-T, 0]$, all conclusions of the theorem are still valid. For convenience of later use, we here apply absolute $|t|$ to substitute t .

$$\begin{aligned}
\|\omega(x(t))\| &\leq \|\dot{x}(t)\| + \|\omega(x(t)) - \dot{x}(t)\| \\
&= \|\dot{x}(t)\| + \left\| \pi(\dot{x}(t), F(x(t))) - \dot{x}(t) \right\| \\
&= \|\dot{x}(t)\| + d(\dot{x}(t), F(x(t))) \\
&\leq 2\|\dot{x}(t)\| + d(0, F(x(t))) \\
&\leq 2\|\dot{x}(t)\| + d(0, F(x(0))) + l\|x(t) - x(0)\|,
\end{aligned}$$

where we have applied the fact that $F(x)$ is with closed value in the third step, the triangle inequality $d(\dot{x}(t), F(x(t))) \leq d(0, \dot{x}(t)) + d(0, F(x(t)))$ in the fourth step, and the Lipschitzian condition $F(x(t)) \subset F(x(0)) + l\|x(t) - x(0)\|B$ in the fifth step. $x(t) \in AC$, hence $\dot{x}(t) \in L_1[0, T]$, moreover, $\omega(x(t)) \in L_1([0, T], \mathbb{R}^n)$. It is obvious from the definition, if $x(t) \in S_{[0, T]}(F, C)$, then $\omega(x(t)) = \dot{x}(t)$.

Using the function $\omega(\cdot)$, a function sequence $\{x_k(t); k = 0, 1, 2, \dots\}$ is constructed as follows:

$$\begin{aligned}
x_0(t) &\in M, \\
x_1(t) &= r_0(x_0(0)) + \int_0^t \omega(x_0(\tau)) d\tau, \\
x_2(t) &= x_1(0) + \int_0^t \omega(x_1(\tau)) d\tau, \\
&\dots \\
x_{k+1}(t) &= x_k(0) + \int_0^t \omega(x_k(\tau)) d\tau, \\
&\dots
\end{aligned}$$

Note that by the definition of $x_k(t)$, we have $x_k(0) = r_0(x_0(0))$ and $x_k(t) \in AC$ for every $k \in \mathbb{N}$. Moreover, if $x(t) \in S_{[0, T]}(F, C)$ and $r_0(x(0)) = x(0)$, then $x_k(t) \equiv x(t)$ for every k .

(2) Convergence of sequences $\{x_k(t)\}$ and $\{\dot{x}_k(t)\}$

At first,

$$\|x_1(t) - x_0(t)\| \leq \|r_0(x_0(0)) - x_0(0)\| + \left\| \int_0^t (\omega(x_0(\tau)) - \dot{x}_0(\tau)) d\tau \right\| \leq \delta + \int_0^t \rho(\tau) d\tau \quad (2.3.16)$$

where the constant δ and the function $\rho(t)$ are those given by the theorem.

Secondly,

$$\|\dot{x}_1(t) - \dot{x}_0(t)\| = \|\omega(x_0(\tau)) - \dot{x}_0(t)\| = d(\dot{x}_0(t), F(x_0(t))) \leq \rho(t) \quad (2.3.17)$$

Now we extend the Inequalities (2.3.13) and (2.3.14) into the case of $k > 1$.

$$\begin{aligned}
 \|\dot{x}_{k+1}(t) - \dot{x}_k(t)\| &= \|\omega(x_k(t)) - \dot{x}_k(t)\| \\
 &= d(\dot{x}_k(t), F(x_k(t))) \\
 &\leq d(\dot{x}_k(t), F(x_{k-1}(t))) + l\|x_k(t) - x_{k-1}(t)\| \\
 &= l\|x_k(t) - x_{k-1}(t)\|.
 \end{aligned} \tag{2.3.18}$$

In the third step, we have applied triangle inequality and Lipschitzian condition. In the fourth step, we have applied the fact that $\omega(x_{k-1}(t)) = \pi(\dot{x}_{k-1}(t), F(x_{k-1}(t))) \in F(x_{k-1}(t))$.

We now prove that

$$\|x_{k+1}(t) - x_k(t)\| \leq \delta \frac{|lt|^k}{k!} + \int_0^t \frac{(l|t| - |\tau|)^k}{k!} \rho(\tau) d\tau \tag{2.3.19}$$

and

$$\|\dot{x}_{k+1}(t) - \dot{x}_k(t)\| \leq l\delta \frac{|lt|^{k-1}}{(k-1)!} + l \int_0^t \frac{(l|t| - |\tau|)^{k-1}}{(k-1)!} \rho(\tau) d\tau \tag{2.3.20}$$

By Inequality (2.3.18), Inequality (2.3.19) implies Inequality (2.3.20). Hence, it is sufficient to verify Inequality (2.3.19).

For the case of $k = 0$, Inequality (2.3.19) degenerates to Inequality (2.3.16). Hence the inequality holds for the case of $k = 0$. We assume that Inequality (2.3.19) holds for the case of k . Now we prove that it is also true for the case of $k + 1$. Since $k \geq 1$,

$$\begin{aligned}
 \|x_{k+1}(t) - x_k(t)\| &= \left\| \int_0^t (\dot{x}_{k+1}(\tau) - \dot{x}_k(\tau)) d\tau \right\| \\
 &\leq l \int_0^t \|x_k(\tau) - x_{k-1}(\tau)\| d\tau \\
 &\leq l \int_0^t \left(\delta \frac{|l\tau|^{k-1}}{(k-1)!} + \int_0^\tau \frac{(l|\tau| - |q|)^{k-1}}{(k-1)!} \rho(q) dq \right) d\tau \\
 &= \delta \frac{|lt|^k}{k!} + l \int_0^t \int_0^\tau \frac{(l|\tau| - |q|)^{k-1}}{(k-1)!} \rho(q) dq d\tau.
 \end{aligned}$$

$$\text{Because } \frac{d}{d\tau} \int_0^\tau \frac{(l|\tau| - |q|)^k}{k!} \rho(q) dq = \int_0^\tau l \frac{(l|\tau| - |q|)^{k-1}}{(k-1)!} \rho(q) dq,$$

$$\delta \frac{|lt|^k}{k!} + l \int_0^t \int_0^\tau \frac{(l|\tau| - |q|)^{k-1}}{(k-1)!} \rho(q) dq d\tau = \delta \frac{|lt|^k}{k!} + \int_0^t \frac{(l|\tau| - |q|)^k}{k!} \rho(q) dq.$$

Replacing q by τ , Inequality (2.3.19) is then verified.

Because $e^z = \sum_{k=0}^{\infty} \frac{z^k}{k!}$, for every real number z , $\sum_{k=m}^{\infty} \frac{z^k}{k!} \rightarrow 0$, ($m \rightarrow \infty$).

Therefore on the time interval $[0, T]$, both $\{x_k(t)\}$ and $\{\dot{x}_k(t)\}$ are Cauchy sequences.

(3) Definition of mapping r

From Step (2), for every $x_0(t) \in M$, we have constructed a sequence $\{x_k(t)\}$ such that $\dot{x}_k(t) \rightarrow v(t)$, ($k \rightarrow \infty$) for a $v(t) \in L_1$. Furthermore if we treat two functions are equivalent when there are identical except at a set whose measure is zero, then the limitation is unique. Now we define

$$r(x_0(t)) = r_0(x_0(0)) + \int_0^t v(\tau) d\tau \quad (2.3.21)$$

then $r(x_0(t)) \in AC$ and $r(x_0(t))(0) = r_0(x_0(0))$.

(4) Proof of the conclusions

From Inequality (2.3.18), we know $\lim_{k \rightarrow \infty} d(\dot{x}_k(t), F(x_k(t))) = 0$. Hence,

$$\|r(x_0(t)) - x_k(t)\| = \left\| \int_0^t (\dot{x}_k(\tau) - v(t)) d\tau \right\| \leq \int_0^t \|\dot{x}_k(\tau) - v(t)\| d\tau \rightarrow 0 \quad (2.3.22)$$

Moreover,

$$0 = \lim_{k \rightarrow \infty} d(\dot{x}_k(t), F(x_k(t))) = d(v(t), F(r(x_0(t)))) = d\left(\frac{d}{dt} r(x_0(t)), F(r(x_0(t)))\right)$$

F is with closed value, consequently, $\frac{dr(x_0(t))}{dt} \in F(r(x_0(t)))$ and $r(x_0(t)) \in S(F, r_0(x_0(0)))$, where we apply $r(x_0(t))(0) = r_0(x_0(0))$. The second conclusion of Theorem 2.3.3 is obtained.

By Inequality (2.3.19),

$$\|x_{k+1}(t) - x_0(t)\| \leq \delta e^{|lt|} + \int_0^t e^{|l|\tau| - |\tau|} \rho(\tau) d\tau = \xi(t)$$

Similarly, from Inequality (2.3.20), we obtained

$$\|\dot{x}_{k+1}(t) - \dot{x}_0(t)\| \leq l\delta e^{|t|} + l \int_0^t e^{|t|-\tau} \rho(\tau) d\tau + \rho(t) = l\xi(t) + \rho(t)$$

Thus,

$$\|r(x_0(t)) - x_0(t)\| \leq \|x_k(t) - x_0(t)\| + \|r(x_0(t)) - x_k(t)\| \leq \xi(t) + \|r(x_0(t)) - x_k(t)\|$$

Applying the last limitation in Inequality (2.3.22), then it implies $\|r(x_0(t)) - x_0(t)\| \leq \xi(t)$, i.e., Inequality (2.3.14) holds. By the same method, we can prove Inequality (2.3.15) holds. The Conclusion 4 of the theorem is then proved.

If $x_0(t) \in S(F, C)$, then $\omega(\dot{x}_0(t)) = \dot{x}_0(t)$, and $r_0(x_0(0)) = x_0(0)$. We conclude $x_1(t) = x_0(t)$. By such a way, we obtain $x_k(t) = x_0(t)$, $k = 1, 2, \dots$. Consequently, $v(t) = \dot{x}_0(t)$, and $r(x_0(t)) = x_0(t)$. The Conclusion 3 of the theorem is also verified.

To end the proof of the theorem, we consider the continuity of r . r is a mapping from M to the space of AC . If $x_0(t) \in M$, then

$$\|x_0(t)\|_{AC} = \|x_0(0)\|_{R^n} + \int_0^T \|\dot{x}_0(\tau)\|_{R^n} d\tau.$$

We can think that $x(t) \in B(x_0(t), \varepsilon)$, then $x(0) \in B(x_0(0), \varepsilon)$, $\|\dot{x}(t) - \dot{x}_0(t)\| < \varepsilon$.

At first, $\omega(x(t)) = \pi(\dot{x}(t), F(x(t)))$ is a continuous mapping from $AC([0, T])$ to

$AC([0, T])$ (to see Prob. 3 of this section). It follows that $y(t) = x(0) + \int_0^t \omega(x(\tau)) d\tau$

is a continuous mapping from $x(t)$ to $y(t)$. By the reason, for every k , the mapping from $x_0(t)$ to $x_k(t)$ is continuous. $\dot{x}_k(t)$ is uniformly convergent to $v(t)$. Hence, the mapping from $x_0(t)$ to $v(t)$ is continuous. At last the mapping r defined by Eq. (2.3.21) is continuous.

Thus, we complete the proof of the theorem. $\square$

Theorem 2.3.3 is a very important result in the theory of differential inclusion. It is a constructive proof for the solutions of differential inclusions. One can start with a set (or a function) in the space of absolutely continuous functions to obtain a solution of a differential inclusion. It is also a critical key to very problems. The following theorem is about the existence of the solution of Cauchy problem of differential inclusion, and the theorem is a natural result of Theorem 2.3.3.

Theorem 2.3.4 Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a Lipschitzian set-valued mapping. It is with closed value. Then the following Cauchy problem

$$\dot{x}(t) \in F(x(t)), \quad x(0) = x_0$$

has a solution.

Proof Let $M = \{0\}$, and $r_0 = x_0 + x$. Take $\rho(t) \equiv d(0, F(0))$ and $\delta = \|x_0\|$. By Theorem $r(0)(0) = r_0(0) = x_0$,⁷ and there exists a $x(t) = r(0) \in S_{[0,\infty]}(F, x_0)$. The $x(t)$ satisfies

$$\|x(t)\| \leq \|x_0\| e^{l|t|} + \int_0^t e^{l(|t|-|\tau|)} d(0, F(0)) d\tau.$$

□

Theorem 2.3.4 only verifies that the solution set $S_{[0,\infty]}(F, x_0)$ is not empty. The theorem does not provide all elements of $S_{[0,\infty]}(F, x_0)$. From the proofs of Theorems 2.3.3 and 2.3.4, we can find that the time interval where the solution exists depends on the interval where the Lipschitzian condition holds. When the set-valued mapping is only a local Lipschitzian mapping, then the above theorem gives a local result that there is a $\delta > 0$ such that $S_{[-\delta,\delta]}(F, x_0) \neq \emptyset$. The conclusion is very similar to that for differential equations.

We give several remarks to Theorem 2.3.3.

Remark 1 In Theorem 2.3.3, we did not require that $F(x)$ is with convex value. We only require $F(x)$ is with closed value. □

Remark 2 Theorem 2.3.3 gives an algorithm to calculate solutions from a set of absolutely continuous functions. It is an approximate method. The key is to calculate the projection of $\dot{x}(t)$ on the set $F(x(t))$. From the theoretic viewpoint, the calculation is operable. □

Remark 3 Inequalities (2.3.14) and (2.3.15) give the errors between $x_0(t) \in M$, $\dot{x}_0(t)$, and its corresponding solution in $S_{[0,\infty]}(F, x_0)$ and its derivate. These errors depend on δ and $\rho(t)$. If $M = \{x_0(t)\}$, i.e., M contains only one function, and $x_0(0) = x_0$, then δ can be zero. Therefore, the less of the distance between $\dot{x}_0(t)$ and $F(x_0(t))$ is, the faster convergence of $x_0(t)$ to $r(x_0(t))$. □

Remark 4 As mentioned before, Theorem 2.3.3 allows that $T < 0$. This is the reason why we apply the notations of $|t|$ and $|s|$. □

As an application of Theorem 2.3.4, we can estimate the boundary of a solution. We have

$$\|x(t, x_0)\| = \|r(0)\| \leq \|x_0\| e^{l|t|} + \int_0^t e^{l(|t|-|s|)} d(0, F(0)) ds \quad (2.3.23)$$

⁷In the notation $r(0)(0)$, the zero in $r(0)$ is the zero function in M , i.e., a constant function whose value is always zero. The another 0 is the initial time.

We have $d(0, F(0)) = \|m(F(0))\|$. Suppose $t > 0$, then Inequality (2.3.23) leads to $\|x(t, x_0)\| \leq \|x_0\| e^{lt} + \frac{1}{l} e^{lt-1} m(F(0))$. The highest increasing speed is e^{lt} which is similar to the corresponding conclusion for differential equation.

2.3.4 The Existence of Solutions of Time-delayed Differential Inclusions

Before ending this section, we spend some time to deal with the extension of the existence theorem for the solution of differential inclusions. In the theory of differential inclusions, it is only required that Inc. (2.3.4) holds only almost everywhere, i.e., it is allowed the inclusion does not hold on a set whose measure is zero. The property makes great convenience for the extension of the existence theorem. As an example, we deal with the time-delayed differential inclusion since time-delayed is very popular for almost all dynamic systems.

Suppose $F : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a closed Lipschitzian set-valued mapping. We consider the Cauchy problem for the following differential inclusion with time-delay.

$$\dot{x}(t) \in F(x(t), x(t - \tau)); \quad x(t) = \phi(t), \quad t \in [-\tau, 0], \quad (2.3.24)$$

where $\tau \in \mathbb{R} (\geq 0)$ is a given constant and used to express the time delay, and $\phi(t) \in AC([-\tau, 0])$ is the initial condition. Because $F : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a Lipschitzian set-valued mapping, there exist two constants l_1 and l_2 such that

$$d(F(x_1, y_1), F(x_2, y_2)) \leq l_1 \|x_1 - x_2\| + l_2 \|y_1 - y_2\|$$

It implies $d(F(x, y), F(0, 0)) \leq l_1 \|x\| + l_2 \|y\|$, or

$$d(0, F(x, y)) \leq l_1 \|x\| + l_2 \|y\| + l_3, \quad (2.3.25)$$

where $l_3 = d(0, F(0, 0))$.

Theorem 2.3.5 Let $F : \mathbb{R}^n \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a closed Lipschitzian set-valued mapping. Then Inc. (2.3.24) has a solution.

Proof Consider the following Cauchy problem,

$$\dot{x}(t) \in F(x(t), \phi(t - \tau)); \quad x(0) = \phi(0).$$

On the time interval $[0, \tau]$, $\phi(t - \tau)$ is a absolutely continuous function, hence, $F(x(t), \phi(t - \tau))$ can be treated as a time-varying differential inclusion. By an appropriate extension of Theorem 2.3.3, the set of solutions $S_{[0, \tau]}(F(x(t), \phi(t - \tau)), \phi(0))$ is not a empty set.

For a solution $x(t) \in S_{[0,\tau]}(F(x(t), \phi(t-\tau)), \phi(0))$, we define

$$x_\tau(t) = \begin{cases} \phi(t), & t \in [-\tau, 0], \\ x(t), & t \in [0, \tau], \end{cases}$$

then $x_\tau(t) \in S_{[0,\tau]}(2.3.23)$ where, for simplicity, we apply the notation $S_{[0,\tau]}(2.3.24)$ for the set of solutions of Inc. (2.3.24). By Inequality (2.3.23), $S_{[0,\tau]}(2.3.24)$ is a bounded set (the fact can be also verified by Inequality (2.3.25), we leave it as an problem).

Now for every $x_\tau(t) \in S_{[0,\tau]}(2.3.24)$, we consider the Cauchy problem

$$\dot{x}(t) \in F(x(t), x_\tau(t-\tau)); \quad x(\tau) = x_\tau(\tau), \quad (2.3.26)$$

with initial time is τ . On the time interval $[\tau, 2\tau]$, $x_\tau(t-\tau)$ is a determined absolutely continuous function, hence by a similar discussion to the first step, $S_{[\tau,2\tau]}(F(x(t), x_\tau(t-\tau)), x_\tau(\tau))$ is nonempty.

For every $x(t) \in S_{[\tau,2\tau]}(F(x(t), x_\tau(t-\tau)), x_\tau(\tau))$, we define

$$x_{2\tau}(t) = \begin{cases} x_\tau(t), & t \in [-\tau, \tau], \\ x(t), & t \in [\tau, 2\tau], \end{cases}$$

then $x_{2\tau}(t) \in S_{[0,2\tau]}(2.3.23)$. $\{x_{2\tau}(\tau); x_{2\tau}(t) \in S_{[0,2\tau]}(2.3.23)\}$ is still a bounded set, thus, step by step, we can extend the solution to $[0, \infty)$. $\square$

Note that $x_\tau(t)$ is defined on $[-\tau, \tau]$. At the point $t = 0$, $x_\tau(t)$ may be not differentiable. The fact meets with the requirement of solutions of differential inclusion.

Problems

1. (1) Suppose $f(x, u)$ is continuous for every fixed $u \in U$, then $F(x, U)$ is lower semi-continuous;
 (2) Suppose $f(x, u)$ is continuous at every pair (x, u) and U is compact, then $F(x, U)$ is continuous.
2. Suppose $\{f_i(x, u), i = 1, 2, \dots\}$ is series of continuous mappings from normed space X to Banach space Y , $\{U_i, i = 1, 2, \dots\}$ is a series of compact and connect subsets of Y . If they satisfy that
 (1) $F(x) \subset \dots \subset f_{n+1}(x, U_{n+1}) \subset f_n(x, U_n) \dots \subset f_1(x, U_1)$ for every $x \in X$;
 (2) For every $\varepsilon > 0$, there is an $N \in \mathbb{N}$, when $n > N$, it holds $f_n(x, U_n) \subset F(x) + \varepsilon B$ where B is the unit ball in Y .

Then $F(x)$ is upper semi-continuous and with compact and connect value.
 And

$$F(x) = \bigcap_{i=1}^{\infty} f_i(x, U_i)$$

3. Suppose $F(x)$ is the Filippov extension of $f(x)$. If $f(x)$ is only upper semi-continuous at x_0 , then what kind continuity can hold the $F(x_0)$? If $f(x)$ is only lower semi-continuous at x_0 , what is the $F(x_0)$?
4. Prove the Gronwell inequality (Hint: using the comparison principle of differential equation).
5. Prove that under the conditions of Theorem 2.3.3, Inequality (2.3.15) holds.
6. Suppose $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is closed. Prove that $\omega(x(t)) = \pi(\dot{x}(t), F(x(t)))$ which is defined in the proof of Theorem 2.3.3 is a continuous mapping from $L_1[0, T]$ to $L_1[0, T]$.
7. Prove the continuity of mapping r defined by Eq. (2.3.21) by using $\varepsilon - \delta$ language.
8. Suppose $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a closed and Lipschitzian set-valued mapping and $v_0 \in F(x_0)$. Then there is a solution $x(t) \in S_{[0, T]}(F, x_0)$ such that $\dot{x}(0) = v_0$. The fact means that we can appoint the initial value of the solution as well as the value of its derivate.
9. Prove that $\{x_\tau(\tau), x_\tau(t) \in S_{[0, \tau]}(2.3.23)\}$ which is defined in the proof of Theorem 2.3.5 is bounded.
10. Repeat the proof of Theorem 2.3.3 if $M = \{x_1(t)\}$ with $x_1(t) \in S_{[0, T]}(F, x_0)$ but the image of M is $S_{[0, T]}(F, x_1)$, $x_1 \neq x_0$.

2.4 Qualitative Analysis for Differential Inclusions

We carry through the qualitative analysis of differential inclusions in this section. The study can be treated as an inheritance and development of the similar study for differential equations since they consider almost the same questions, but methods used and conclusions are quite different. This section only considers the local properties of differential inclusions. We start with those differential inclusions which hold Lipschitzian property, then turn to those which are upper semi-continuous, at last we deal with the problem of convexification of differential inclusions.

2.4.1 Qualitative Analysis for Lipschitzian Differential Inclusions

Lipschitzian differential inclusions are studied qualitatively by using Theorem 2.3.3 which is powerful tool in the theory of differential inclusions as mentioned before.

Suppose $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is a Lipschitzian set-valued mapping, it is with closed and convex value. By Theorem 2.3.4, the Cauchy problem of differential inclusion

$$\dot{x}(t) \in F(x(t)), \quad x(0) = x_0 \quad (2.4.1)$$

holds solutions on the time interval $[0, T]$ (and/or $[-T, 0]$ ⁸). $S_{[-T, 0]}(F, x_0)$ is the solution set of Inc. (2.4.1). In this subsection, three qualitative conclusions are presented for $S_{[0, T]}(F, x_0)$.

Theorem 2.4.1 Let $x_1(t), x_2(t) \in S_{[0, T]}(F, x_0)$ be two solutions. Then there is a continuous mapping $\phi : [0, 1] \rightarrow S_{[0, T]}(F, x_0)$ such that $\phi(0) = x_1(t)$ and $\phi(1) = x_2(t)$.

Proof Define $M = \{(1 - \lambda)x_1(t) + \lambda x_2(t); \lambda \in [0, 1]\}$, then $M \subset AC([0, T], \mathbb{R}^n)$. To use Theorem 2.3.3, we check the distance $d((1 - \lambda)\dot{x}_1(t) + \lambda\dot{x}_2(t), F((1 - \lambda)x_1(t) + \lambda x_2(t)))$.

$$\begin{aligned}
 & d((1 - \lambda)\dot{x}_1(t) + \lambda\dot{x}_2(t), F((1 - \lambda)x_1(t) + \lambda x_2(t))) \\
 & \leq d((1 - \lambda)\dot{x}_1(t) + \lambda\dot{x}_2(t), \dot{x}_1(t)) + d(\dot{x}_1(t), F((1 - \lambda)x_1(t) + \lambda x_2(t))) \\
 & \leq \lambda \|\dot{x}_2(t) - \dot{x}_1(t)\| + d(\dot{x}_1(t), F((1 - \lambda)x_1(t) + \lambda x_2(t))) \\
 & \leq \lambda \|\dot{x}_2(t) - \dot{x}_1(t)\| + d(\dot{x}_1(t), F(x_1)) + d(F(x_1), F((1 - \lambda)x_1(t) + \lambda x_2(t))) \\
 & \leq \lambda \|\dot{x}_2(t) - \dot{x}_1(t)\| + l\lambda \|x_2(t) - x_1(t)\| \\
 & \leq \|\dot{x}_2(t) - \dot{x}_1(t)\| + l\|x_2(t) - x_1(t)\|,
 \end{aligned}$$

where we have use the triangle inequality in the second and the fourth steps, respectively, and the facts that $d(\dot{x}_1, F(x_1)) = 0$, Lipschitzian condition and $\lambda \leq 1$.

Denote $\rho(t) = \|\dot{x}_2(t) - \dot{x}_1(t)\| + l\|x_2(t) - x_1(t)\|$. Then $\rho(t) \in L_1([0, T], \mathbb{R})$ because $x_1(t), x_2(t) \in AC([0, T], \mathbb{R}^n)$. Let $r_0(x) = x$. By Theorem 2.3.3, there is a continuous mapping $r : M \rightarrow S_{[0, T]}(F, x_0)$. Because $x_1(t), x_2(t) \in S_{[0, T]}(F, x_0)$, by Theorem 2.3.3 again, $r(x_1(t)) = x_1(t)$, $r(x_2(t)) = x_2(t)$. Now we define

$$\phi(\lambda) = r((1 - \lambda)x_1(t) + \lambda x_2(t))$$

$\phi(\lambda)$ is continuous and $\phi(\lambda) \in S_{[0, T]}(F, x_0)$ for all $\lambda \in [0, 1]$. It follows $\phi(0) = x_1(t)$ and $\phi(1) = x_2(t)$. We have complete the proof. $\square$

We have no reason to assert that the set $S_{[0, T]}(F, x_0)$ is convex, consequently, $(1 - \lambda)x_1(t) + \lambda x_2(t)$ may fail to be a solution of Inc. (2.4.1).

Theorem 2.4.1 is usually called as arc connectedness theorem. It illustrates that if we treat two solutions $x_1(t)$ and $x_2(t)$ as two points of the vector space of absolutely continuous functions, then there is an arc which connects the two points such that every point at the arc is also the solution in $S_{[0, T]}(F, x_0)$. In $\mathbb{R} \times \mathbb{R}^n$, every solution in $S_{[0, T]}(F, x_0)$ can be drawn as a curve, the conclusion illustrate that there exists a

⁸We only consider the case of $[0, T]$ below. Readers are suggested to give and prove the similar conclusions for the case of $S_{[-T, 0]}(F, x_0)$ as exercises.

curved surface which links with $x_1(t)$ and $x_2(t)$, and every point in the surface has passed at least one solution in $S_{[0,T]}(F, x_0)$ (Fig. 2.12).

Definition 2.4.1 Let F be set-valued mapping, the following set

$$R_{[0,T]}(F, x_0) = \{x(T); x(t) \in S_{[0,T]}(F, x_0)\}$$

is the T -reachable set of Inc. (2.4.1) with the starting point x_0 . When T and x_0 are in unambiguous, the set is called reachable set of Inc. (2.4.1). $\square$

By Definition 2.4.1, we can restate Theorem 2.4.1 as follows: in $\mathbb{R}^n$, $R_{[0,t]}(F, x_0)$ is arc wise connected for every $t \in [0, T]$.

Theorem 2.4.2 Let $x(t) \in S_{[0,T]}(F, x_0)$ be a solution of Inc. (2.4.1), and $x(T) = x^*$. If $x^* \in \text{bd } R_{[0,T]}(F, x_0)$, then for every $t_1 \in [0, T]$, $x(t_1) \in \text{bd } R_{[0,t_1]}(F, x_0)$.

Proof It is proved by contradiction. If there is a $t_1 \in (0, T)$ such that $x(t_1) \notin \text{bd } R_{[0,t_1]}(F, x_0)$, then by the definition of boundary, there is an $\varepsilon > 0$ such that $B(x(t_1), \varepsilon) \subset S_{[0,t_1]}(F, x_0)$.

On the other hand, Because $x^* \in \text{bd } R_{[0,T]}(F, x_0)$, there is a $y \in \mathbb{R}^n$ such that $y \notin R_{[0,T]}(F, x_0)$ and $\|y - x^*\| \leq \varepsilon e^{-|t_1 - T|}$. To apply Theorem 2.3.3, let $M = \{x(t)\}$ and $r_0(x) = x - x^* + y$, then $\rho(t) = 0$ and $\delta = \varepsilon e^{-|t_1 - T|}$. Theorem 2.3.3 asserts there is a solution $y(t) = r(x(t)) \in S_{[t_1, T]}(F, y_1)$ ⁹ where $y_1 = y(t_1)$ which satisfies $y(T) = y$ (Fig. 2.13).

Fig. 2.12 The shadow part is contained in $S_{[0,T]}(F, x_0) \subset \mathbb{R} \times \mathbb{R}^n$

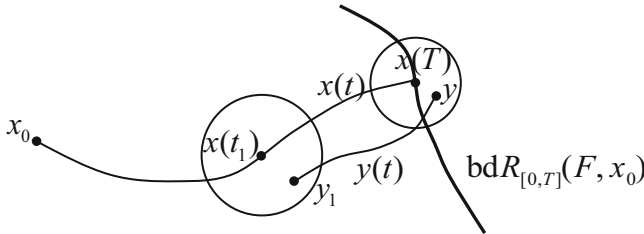
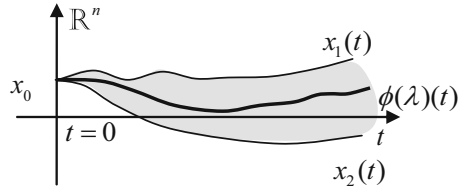


Fig. 2.13 Illustration of Theorem 2.4.2

⁹We apply Theorem 2.3.3 on the time interval $[T - t_1, T]$. The terminal is fixed at $y(T) = y$. The solution is $y(t) \in S_{[T-t_1, T]}(F, y_1)$, where $y_1 = y(T - t_1)$.

Using Inequality (2.3.14) in Theorem 2.3.3, we have $\|y(t_1) - x_1(t_1)\| \leq \|y - x^*\| e^{l(T-t_1)} < \varepsilon$. Hence, $y(t_1) \in R_{[0,t_1]}(F, x_0)$, i.e., there is a solution $x_2(t) \in S_{[0,t_1]}(F, x_0)$ such that $x_2(t_1) = y(t_1)$. We define

$$x_1(t) = \begin{cases} x_2(t) & 0 \leq t \leq t_1, \\ y(t) & t_1 \leq t \leq T. \end{cases}$$

Then $x_1(t) \in S_{[0,T]}(F, x_0)$, and $x_1(T) = y$. It is in conflict with the assumption $y \notin R_{[0,T]}(F, x_0)$. $\square$

Theorem 2.4.2 illustrates that if $x(T)$ is on the boundary of $R_{[0,T]}(F, x_0)$, then the trajectory $x(t) \in S_{[0,T]}(F, x_0)$ has the property that for every $t_1 \in [0, T]$, $x(t_1)$ is at the boundary of $R_{[0,t_1]}(F, x_0)$. The readers are suggested to recall Example 3.2.1, then you can have a deeper understanding.

$S_{[0,T]}(F, x_0)$ is the solution set of Inc. (2.4.1), hence $S_{[0,T]}(F, x_0)$ is a set in normed space $AC([0, T], \mathbb{R}^n)$. We can define a set-valued mapping from $\mathbb{R}^n$ to $AC([0, T], \mathbb{R}^n)$ by $x \mapsto S_{[0,T]}(F, x)$. This is a mapping from the initial condition to the solution set. The following theorem is given for the continuity of the mapping.

Theorem 2.4.3 The set-valued mapping $x \mapsto S_{[0,T]}(F, x)$ is Lipschitzian. Moreover, if $x_1(t) \in S_{[0,T]}(F, x_0)$, then there is a continuous selection $\phi(x)$ of $S_{[0,T]}(F, x)$ such that $\phi(x_0) = x_1(t)$.

Proof Let $x_1, x_2 \in \mathbb{R}^n$. By Theorem 2.3.4, $S_{[0,T]}(F, x_1) \neq \emptyset$. $x_1(t)$ is an arbitrary solution in $S_{[0,T]}(F, x_1)$. Let $M = \{x_1(t)\}$ and $r_0(x) = x - x_1 + x_2$. Then, by Theorem 2.3.3, there is a mapping $r : M \rightarrow S_{[0,T]}(F, x_2)$. By using the notations defined in Theorem 2.3.3, we have $\rho(t) = 0$ and $\delta = \|x_2 - x_1\|_{\mathbb{R}^n}$. Let $r(x_1(t)) = x_2(t)$. Then by Inequality (2.3.14), we have

$$\|\dot{x}_1(t) - \dot{x}_2(t)\|_{\mathbb{R}^n} \leq \|x_1 - x_2\|_{\mathbb{R}^n} l e^{lt}$$

Using the norm defined in AC , we have

$$\|x_1(t) - x_2(t)\|_{AC} = \|x_1 - x_2\|_{\mathbb{R}^n} + \int_0^T \|\dot{x}_1(t) - \dot{x}_2(t)\|_{\mathbb{R}^n} dt \leq e^{lT} \|x_1 - x_2\|_{\mathbb{R}^n}$$

The e^{lT} is independent of the selection of $x_1(t)$ and $x_2(t)$, hence $S_{[0,T]}(F, x)$ is Lipschitzian globally, i.e.,

$$S_{[0,T]}(F, x_1) \subset S_{[0,T]}(F, x_2) + e^{lT} \|x_1 - x_2\|_{\mathbb{R}^n} B.$$

For the second conclusion, let $M = \{x_1(t) + x - x_0\}$ ¹⁰ and $r_0(x) = x$, then there is a continuous mapping $r : M \rightarrow S_{[0,T]}(F, x)$. Let $\phi(x) = r(x_1(t) + x - x_0)$. Then the $\phi(x)$ is continuous, $\phi(x) \in S_{[0,T]}(F, x)$ and $\phi(x_0) = r(x_1(t)) = x_1(t)$. $\square$

2.4.2 Qualitative Analysis for Upper Semi-continuous Differential Inclusions

This subsection turns to upper semi-continuous differential inclusions and deals with the same problems as the last subsection. Because failure of Lipschitz condition, we cannot apply Theorem 2.3.3, these conclusions are established by a different way. Readers are suggested to compare these conclusions as well as the proof procedures.

Consider the Cauchy problem of the following differential inclusion

$$\dot{x}(t) \in F(x(t)), \quad x(0) = x_0 \quad (2.4.2)$$

where $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is an upper semi-continuous set-valued mapping and with closed and convex value. We further assume that the F is bounded, i.e., there is a $b \in \mathbb{R} (> 0)$ such that for every $x \in \mathbb{R}^n$, $F(x) \subset bB$. Comparing with Inc. (2.4.1), the set-valued mapping F in Inc. (2.4.2) is required to be bounded and with convex value.

Two lemmas are given firstly.

Lemma 2.4.1 Suppose $A \subset \mathbb{R}^n$ is a convex and compact set, $f(t) \in L_1([0, T], \mathbb{R}^n)$.

If for almost all $t \in [0, T]$, $f(t) \in A$, then $\frac{1}{T} \int_0^T f(t) dt \in A$. $\square$

The lemma can be treated as an extension of mean value theorem of Riemann integration. It can be proved by the definition of Lebesgue integration, and we omit it here. It is only pointed out that the condition that A is a convex and compact set is necessary.

Lemma 2.4.2 Suppose that $A \subset \mathbb{R}^n$ is a convex and compact set, $x_k(t) \in AC([0, T], \mathbb{R}^n)$, $k = 1, 2, \dots$, and for almost all $t \in [0, T]$, $\dot{x}_k(t) \in A$. If $\lim_{k \rightarrow \infty} x_k(t) = x(t)$ for every $t \in [0, T]$, then $x(t) \in AC([0, T], \mathbb{R}^n)$ and $\dot{x}(t) \in A$ for almost all $t \in [0, T]$.

¹⁰This M is not a solution set.

Proof A is compact, so is bounded. Because $\dot{x}_k(t) \in A$, there exists a $l \in \mathbb{R}^+$ such that $\|\dot{x}_k(t)\| \leq l$ for almost all $t \in [0, T]$ and all $k = 1, 2, \dots$. Consequently,

$$\|x_k(t_2) - x_k(t_1)\| \leq \int_{t_1}^{t_2} \|\dot{x}_k(\tau)\| d\tau \leq l|t_2 - t_1|,$$

for all $t_1, t_2 \in [0, T]$. It implies the sequence $\{x_k(t), t \in [0, T], k = 1, 2, \dots\}$ is equicontinuous. Now let $k \rightarrow \infty$, we have that $x(t)$ is a Lipschitzian mapping, and $x(t) \in AC([0, T], \mathbb{R}^n)$.

Moreover, by Lemma 2.4.1,

$$\frac{x_k(t+h) - x_k(t)}{h} = \frac{1}{h} \int_t^{t+h} \dot{x}_k(\tau) d\tau \in A,$$

let $k \rightarrow \infty$, we obtain $\frac{x(t+h) - x(t)}{h} \in A$ for $t \in [0, T-h]$ because of the compactness of A . By the compactness again, $\dot{x}(t) = \lim_{h \rightarrow 0} \frac{x(t+h) - x(t)}{h} \in A$ for all $t \in [0, T]$. $\square$

The proof of Lemma 2.4.2 illustrates the compactness plays a critical role in the convergence. It also illustrates that if a sequence of absolutely continuous functions whose values are on a compact set is convergent, then the limitation is absolutely continuous.

We now give the main conclusion for upper semi-continuous differential inclusions which holds a similar position to that of Theorem 2.3.3.

Theorem 2.4.4 Let $F, F_k : \mathbb{R}^n \rightarrow \mathbb{R}^n$ are all upper semi-continuous and with convex and closed values. Moreover, they satisfy the following conditions:

There is a $b \in \mathbb{R} (> 0)$ such that

- (1) $F(x) \subset \dots \subset F_{k+1}(x) \subset F_k(x) \subset \dots \subset F_0(x) \subset bB$ for every $x \in \mathbb{R}^n$;
- (2) For every $\varepsilon > 0$ and $x \in \mathbb{R}^n$, there is a $k_0 = k_0(x, \varepsilon) \in \mathbb{N}$, when $k > k_0$, $F_k(x) \subset F(x) + \varepsilon B$;
- (3) $x_k(t) \in S_{[0,T]}(F_k)$ and $x_k(t)$ converges to $x(t)$ uniformly.

Then $x(t) \in S_{[0,T]}(F)$.

Proof Because $\dot{x}_k(t) \in F_k(x_k(t)) \subset bB$ and $x_k(t)$ converges uniformly to $x(t)$, we conclude $x(t) \in AC$ by Lemma 2.4.2. Suppose $x(t)$ is derivable at $t_0 \in [0, T]$ and denote $x_0 = x(t_0)$. Using the second condition of the Theorem 2.4.4, for every $\varepsilon > 0$, we have a $\bar{k} > k_0(x_0, \varepsilon)$ such that $F_{\bar{k}}(x_0) \subset F(x_0) + \varepsilon B$. $F_{\bar{k}}(x_0)$ is an upper semi-continuous set-valued mapping, hence, there is an $\eta > 0$, when $x \in B(x_0, \eta)$,

$$F_{\bar{k}}(x) \subset F_{\bar{k}}(x_0) + \varepsilon B \subset F(x_0) + 2\varepsilon B. \quad (2.4.3)$$

On the other hand, $x_k(t)$ converges uniformly to $x(t)$, there exist $k_1 \geq \bar{k}$ and γ when $k > k_1$ and $|t - t_0| < \gamma$, we have

$$\|x_k(t) - x(t)\| < \frac{\eta}{2}, \quad \|x(t) - x(t_0)\| < \frac{\eta}{2}.$$

By Relation (2.4.3),

$$\dot{x}_k(t) \in F_k(x_k(t)) \subset F_{k_0}(x_k(t)) \subset F_{k_0}(x(t_0)) + \varepsilon B \subset F(x_0) + 2\varepsilon B.$$

Using Lemma 2.4.2 again, $\dot{x}(t) \in F(x_0) + 2\varepsilon \bar{B}$ for all $|t - t_0| < \gamma$. Let $t = t_0$. Then $\dot{x}(t_0) \in F(x_0) + 2\varepsilon B$. The ε can be selected arbitrary and $F(x_0)$ is closed, consequently, $\dot{x}(t_0) \in F(x_0)$. $\square$

Remark If the condition that $F_k : \mathbb{R}^n \rightarrow \mathbb{R}^n$ are upper semi-continuous is replaced by $F_k : \mathbb{R}^n \rightarrow \mathbb{R}^n$ are Lipschitzian mappings, then it can be proved that Relation (2.4.3) still holds, hence, the theorem is still valid.

By Theorem 2.2.4 and the above remark, Theorem 2.4.4 supports the following conclusion, the detailed proof is left to readers.

Theorem 2.4.5 Consider Inc. (2.4.2) i.e., $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is an bounded, upper semi-continuous set-valued mapping and with closed and convex value. Then its Cauchy problem is solvable. $\square$

Corollary 2.4.1 Suppose $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is upper semi-continuous (or locally Lipschitzian) and with closed and convex value, and F is bounded, i.e., there is a b such that for every $x \in \mathbb{R}^n$ $F(x) \subset bB$. Then $S_{[0,T]}(F, x_0)$ is a compact set in AC .

Proof Because $F(x) \subset bB$, every sequence of $S_{[0,T]}(F, x_0)$ is equicontinuous. By Arzela-Ascoli theorem (Sect. 1.1), the sequence holds a convergent subsequence. By Theorem 2.4.4, replacing all F_k by F , we conclude the limitation of the subsequence belongs also to $S_{[0,T]}(F, x_0)$. $\square$

Corollary 2.4.1 can be extended to the set $S_{[0,T]}(F, x_0)$ provided that $C \subset \mathbb{R}^n$ is a compact set in $\mathbb{R}^n$ (Problem 1).

We now apply Theorem 2.4.4 to complete the corresponding qualitative analysis for Inc. (2.4.2).

A set A is said to be unconnected if there are two closed sets A_1 and A_2 which satisfy $A_1 \cap A_2 = \emptyset$, and $A_1 \cap A \neq \emptyset$ and $A_2 \cap A \neq \emptyset$ such that $A_1 \cup A_2 \supset A$. Otherwise, A is connected.

Theorem 2.4.6 Suppose that $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is bounded, upper semi-continuous and with closed and convex value. Then for every $x_0 \in \mathbb{R}^n$, $S_{[0,T]}(F, x_0)$ is a connected set of the space $C([0, T], \mathbb{R}^n)$.¹¹

¹¹Note that the connectedness is considered in normed space C , and is not in $\mathbb{R} \times \mathbb{R}^n$. In $\mathbb{R} \times \mathbb{R}^n$, all solutions in $S_{[0,T]}(F, x_0)$ start at $(0, x_0)$, hence it has to be connected.

Proof The conclusion is proved by using contradiction. If there are two nonempty closed sets S_1 and S_2 which satisfy $S_1 \cap S_2 = \emptyset$, $S_1 \cup S_2 \supset S_{[0,T]}(F, x_0)$, and $S_i \cap S_{[0,T]}(F, x_0) \neq \emptyset$, for $i = 1, 2$. Denote $x_i(t) \in S_{[0,T]}(F, x_0) \cap S_i$, $i = 1, 2$. By Corollary 2.4.1, $S_{[0,T]}(F, x_0)$ is compact, hence S_1 and S_2 are all bounded. We have

$$\inf \left\{ \|x_1(t) - x_2(t)\| ; x_1(t) \in S_{[0,T]}(F, x_0) \cap S_1, x_2(t) \in S_{[0,T]}(F, x_0) \cap S_2 \right\} = 2\delta > 0$$

A function is defined as follows:

$$\varphi : C([0, T], \mathbb{R}^n) \rightarrow \mathbb{R}, \quad \varphi(x(t)) = d(x(t), S_1) - \delta$$

When $x(t) \in S_1$, $\varphi(x(t)) = -\delta$; and $x(t) \in S_2$, $\varphi(x(t)) \geq \delta$. We now fix a $x_1(t) \in S_1 \cap S_{[0,T]}(F, x_0)$ and a $x_2(t) \in S_2 \cap S_{[0,T]}(F, x_0)$.

By Theorem 2.2.4, there are $F_k : \mathbb{R}^n \rightarrow \mathbb{R}^n$, $k = 1, 2, \dots$, which are locally Lipschitzian set-valued mappings and are with such that convex and closed values. Moreover, they satisfy

- (1) For every $x \in \mathbb{R}^n$, $F(x) \subset \dots \subset F_{k+1}(x) \subset F_k(x) \subset \dots \subset F_0(x) \subset bB$;
- (2) For every $\varepsilon > 0$, there exists a $k_0(x, \varepsilon)$, when $k > k_0(x, \varepsilon)$, we have $F_k(x) \subset F(x) + \varepsilon B$.

When $k > k_0(x_0, \varepsilon)$, we consider the Cauchy problem that $\dot{x} \in F_k(x)$, $x(0) = x_0$, its solution set $S_{[0,T]}(F_k, x_0) \neq \emptyset$ by Theorem 2.3.4. Using Theorem 2.4.1, there is a continuous mapping $\phi_k : [0, 1] \rightarrow S_{[0,T]}(F_k, x_0)$ such that $\phi_k(0) = x_1(t)$ and $\phi_k(1) = x_2(t)$. Consider the compound function: $\varphi \circ \phi_k : [0, 1] \rightarrow \mathbb{R}$, we have $\varphi \circ \phi_k(0) < 0$ and $\varphi \circ \phi_k(1) > 0$. By the property of continuous function, there is a $\lambda_k \in (0, 1)$ such that $\varphi \circ \phi_k(\lambda_k) = 0$. Denote $x_k(t) = \phi_k(\lambda_k)$, then $x_k(t) \in S_{[0,T]}(F_k, x_0)$. Because sequence $\{\dot{x}_k(t)\} \subset bB$, the sequence $\{x_k(t)\}$ has a convergent subsequence. Without loss of generality, we assume that $x_k(t) \rightarrow x(t)$ by the norm of the space of continuous functions. Theorem 2.4.4 then asserts that $x(t) \in S_{[0,T]}(F, x_0)$. From $\varphi(x_k(t)) = 0$ and the continuity of φ , we have $\varphi(x(t)) = 0$. Therefore, $x(t) \notin S_1 \cup S_2$, it is contradictive to the hypothesis of $S_1 \cup S_2 \supset S_{[0,T]}(F, x_0)$. Hence $S_{[0,T]}(F, x_0)$ is connected. $\square$

Theorem 2.4.7 Suppose that $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is bounded, upper semi-continuous and with closed and convex value. Let $x^* \in \text{bd } R_{[0,T]}(F, x_0)$. Then there is an $x_1(t) \in S_{[0,T]}(F, x_0)$ which holds the properties: (1) $x_1(T) = x^*$, (2) for every $t \in [0, T]$, $x_1(t) \in \text{bd } R_{[0,t]}(F, x_0)$.

Proof The proof consists of two parts. At first, we construct a solution $x(t) \in S_{[0,T]}(F, x_0)$, then we prove the solution satisfies the requirements of the theorem.

- (1) By Theorem 2.2.4, there is a sequence $\{F_k(x)\}$ such that for every $x \in \mathbb{R}^n$

$$F(x) \subset \dots \subset F_{k+1}(x) \subset F_k(x) \subset \dots \subset F_0(x) \subset bB;$$

And for every $\varepsilon > 0$, there exists a $k_0(x, \varepsilon)$, when $k > k_0(x, \varepsilon)$, $F_k(x) \subset F(x) + \varepsilon B$.

By Corollary 2.4.1, $S_{[0,T]}(F_k, x_0)$ is compact. Hence, there exists a $x_k \in R_{[0,T]}(F_k, x_0)$ such that $d(x^*, R_{[0,T]}(F_k, x_0)) = \|x_k - x^*\|$. It is obvious that $x_k \in \text{bd}R_{[0,T]}(F_k, x_0)$. By Theorem 2.4.2, there exists an $x_k(t) \in S_{[0,T]}(F_k, x_0)$ such that $x_k(T) = x_k$, moreover, for every $t \in [0, T]$, $x_k(t) \in \text{bd}R_{[0,t]}(F_k, x_0)$. By the same reasoning as used in Theorem 2.4.6, we can assume that $x_k(t) \rightarrow x(t)$ by the norm of the space of continuous functions, and $x(t) \in S_{[0,T]}(F, x_0)$.

(2) We now prove that the $x(t)$ satisfies the requirements of theorem. At first, we assert that $x(T) = x^*$. The conclusion is verified by contradiction. If $\delta = \|x(T) - x^*\| > 0$, then there is a k_0 such that for every $k \geq k_0$, $\|x_k - x^*\| \geq \frac{\delta}{2}$ where $x_k = x_k(T) \in \text{bd}R_{[0,T]}(F_k, x_0)$. By the condition of theorem $x^* \in R_{[0,T]}(F, x_0) \subset R_{[0,T]}(F_k, x_0)$, hence $x^* \notin \text{bd}R_{[0,T]}(F_k, x_0)$ (recall the selection of x_k). It follows x^* is an inner point of $R_{[0,T]}(F_k, x_0)$, consequently, there is an $\varepsilon > 0$ such that $x^* + \varepsilon B \subset R_{[0,T]}(F_k, x_0)$ for all $k \geq k_0$. If $y \in B(x^*, \varepsilon)$, then we have a $y_k(t) \in S_{[0,T]}(F_k, x_0)$ and $y = y_k(T)$. By the proof procedure applied above, we can assume $\{y_k(t)\}$ converges uniformly to $y(t) \in S_{[0,T]}(F, x_0)$. That $y \in R_{[0,T]}(F, x_0)$ implies that x^* is not in on the boundary of $R_{[0,T]}(F_k, x_0)$. It contradict condition of the theorem.

At last, we show that $x(t)$ has the property that $x(t) \in \text{bd}R_{[0,t]}(F, x_0)$ for every $t \in [0, T]$. Otherwise, this conclusion is not true, i.e., there is a $t_1 \in [0, T]$, such that $x(t_1) + \varepsilon B \subset R_{[0,t_1]}(F, x_0)$. Then for k which is large enough, we have

$$x_k(t_1) + \frac{\varepsilon}{2}B \subset x(t_1) + \varepsilon B \subset R_{[0,t_1]}(F, x_0) \subset R_{[0,t_1]}(F_k, x_0)$$

It conflicts to Theorem 2.4.2 since F_k is a Lipschitzian set-valued mapping. □

Remark Theorem 2.4.6 and Theorem 2.4.7, in the theory of differential equations in manifold, are known as Kneser theorem and Hukuhara theorem, respectively.

Readers should distinguish Theorem 2.4.2 from Theorem 2.4.7. Theorem 2.4.2 asserts that if $x_1(T) = x^*$, then for every $t \in [0, T]$, $x_1(t) \in \text{bd} R_{[0,t]}(F, x_0)$. But, Theorem 2.4.6 states that there may be several solutions in $S_{[0,T]}(F, x_0)$ whose terminals are all x^* . Among these solutions there is certainly a $x_1(t)$ such that $x_1(T) = x^*$ and $x_1(t) \in \text{bd} R_{[0,t]}(F, x_0)$ for $t \in [0, T]$.

Example 2.4.1 A set-valued mapping $F(x)$ is defined as follows:

$$F(x) = \begin{cases} \bar{B}_2 & x_2 < 0, \\ 6\bar{B}_2 & x_2 = 0, \\ 0 & x_2 > 0. \end{cases} \quad (2.4.4)$$

where $x = [x_1 \ x_2]^T \in \mathbb{R}^{2^{12}}$ and $\bar{B}_2$ is the closed unit ball in $\mathbb{R}^2$. Let the initial condition be $x(0) = [0 \ -4]^T$ and terminal be $x(T) = [3 \ 0]^T$, and time interval is $[0, 5]$ (i.e., $T = 5$).

¹²The superscript T means transposition.

It is easy to check that the $F(x)$ is an upper semi-continuous and with convex and closed value. We firstly illustrate that $F(x)$ is not a Lipschitzian mapping. If $x = [1 \ 0]^T$, then $F(x) = 6\bar{B}_2$; but if $y = [1 \ \varepsilon]^T$ for every $\varepsilon > 0$, then $F(y) \equiv 0$. If we require that

$$6\bar{B}_2 = F(x) \subset F(y) + l\|x - y\| B_2 = l\varepsilon B_2$$

then $l\varepsilon > 6$. When $\varepsilon \rightarrow 0$, then $l \rightarrow \infty$, hence $F(x)$ is not a Lipschitzian mapping.

Now

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 0 \\ -4 + \frac{8t}{9} \end{bmatrix}, t \in \left[0, \frac{9}{2}\right], \quad \begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} 6\left(t - \frac{9}{2}\right) \\ 0 \end{bmatrix}, \quad t \in \left[\frac{9}{2}, 5\right]$$

is a solution of differential inclusion $\dot{x} \in F(x)$ where $F(x)$ is given in Eq. (2.4.4). The solution satisfies that

$$\begin{bmatrix} x_1(0) \\ x_2(0) \end{bmatrix} = \begin{bmatrix} 0 \\ -4 \end{bmatrix}, \quad \begin{bmatrix} x_1(5) \\ x_2(5) \end{bmatrix} = \begin{bmatrix} 3 \\ 0 \end{bmatrix}$$

But when $t \in \left[0, \frac{9}{2}\right]$, $\begin{bmatrix} \dot{x}_1(t) \\ \dot{x}_2(t) \end{bmatrix} = \begin{bmatrix} 0 \\ \frac{8}{9} \end{bmatrix}$ which does not lay on the boundary of B_2 .

There is another solution

$$\begin{bmatrix} x_1(t) \\ x_2(t) \end{bmatrix} = \begin{bmatrix} \frac{3}{5}t \\ -4 + \frac{4}{5}t \end{bmatrix}, \quad t \in [0, 5],$$

which also satisfies the boundary conditions. The solution is always on the boundary. $\square$

Theorem 2.4.8 Suppose $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a bounded, upper semi-continuous set-valued mapping, and with closed and convex value. If we treat $S_{[0,T]}(F, x_0)$ as a set-valued mapping which maps $\mathbb{R}^n$ to $AC([0, T], \mathbb{R}^n)$. Then the mapping is upper semi-continuous.

Proof We prove a more general conclusion that if $C \subset \mathbb{R}^n$ is a compact set, then $S_{[0,T]}(F, C)$ is a compact set of $AC([0, T], \mathbb{R}^n)$.

Suppose $\{x_k(x)\} \subset S_{[0,T]}(F, C)$ is a sequence. Because $F(C) \subset bB$ for some constant b , $\|\dot{x}_k(x)\| \leq b$ is true for all k and almost all x . The fact implies that $\{x_k(x)\}$ is equicontinuous. There exists a convergent subsequence by Arzela-Ascoli theorem. Without loss of generality, we assume $\{x_k(x)\}$ is uniformly convergence, i.e., there is a function $x_0(t) \in AC([0, T], \mathbb{R}^n)$ such that $x_{k_j}(t) \rightarrow x_0(t)$, we now prove that $x_0(t) \in S_{[0,T]}(F, C)$.

Suppose $\dot{x}_k(t) \in F_k(x_k(t))$ where $F_k \equiv F$, $k = 1, 2, \dots$. Then by Theorem 2.4.4, $x_0(t) \in S_{[0,T]}(F, C)$. $S_{[0,T]}(F, C)$ is compact by Corollary 2.4.1. Moreover, by the Problem 9 in Sect. 2.1, we conclude that the mapping $x_0 \mapsto S_{[0,T]}(F, x_0)$ is upper semi-continuous. $\square$

Note that in Theorem 2.4.8, the condition that $F(x)$ is convex is necessary.

2.4.3 Convexification of Differential Inclusions and Relaxed Theorem

At last, we deal with the convexification of differential inclusions in this section. It have been found that convexity is a critical condition for the existence of the solutions for differential inclusions. One may ask such a question *if the set $F(x)$ is not convex, what can we do?* The best answer currently is to make $F(x)$ convex.

Let $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a set-valued mapping, and $\overline{\text{co}} F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be the closure of convex hull of F . Then the differential inclusion

$$\dot{x} \in \overline{\text{co}} F(x) \quad (2.4.5)$$

is the convexification of the original differential inclusion

$$\dot{x} \in F(x). \quad (2.4.6)$$

For the sake of convenience, the notation $\overline{\text{co}} F$ is simplified as $\bar{F}$, and the solution of Inc. (2.4.5) as $\bar{x}(t)$. If the initial conditions are equal, i.e., $\bar{x}(0) = x(0) = x_0$, then the solution of Inc. (2.4.6) is certainly the solution of Inc. (2.4.5). In the study of convexification, the most important problem is to present the conditions under which the solutions of Inc. (2.4.5) is dense at the solutions of Inc. (2.4.6).

The main target of this subsection is to prove Theorem 2.4.9. The proof is quite similar to that of the existence of solution for differential equations by using Euler polygonal lines. But the construction is much complicated.

Theorem 2.4.9 Let $x_0 \in \mathbb{R}^n$, $F : B(x_0, b) \rightarrow \mathbb{R}^n$ be a Lipschitzian set-valued mapping with Lipschitzian constant l . F is bounded and with closed value. Then there is a $T > 0$ such that the solutions of Inc. (2.4.5) is dense at the solutions of Inc. (2.4.6) on the time interval $I = [-T, T]$.

Proof The theorem is equivalent to prove that for every $\varepsilon > 0$ and every $\bar{x}(t, x_0)$ which is a solution of Inc. (2.4.5), then there is a solution $x(t, x_0)$ of Inc. (2.4.6) such that $\|x(t, x_0) - \bar{x}(t, x_0)\| < \varepsilon$ for every $t \in I$. Both $\bar{x}(t, x_0)$ and $x(t, x_0)$ are in $B(x_0, b)$ for each $t \in [-T, T]$.

Because F is bounded, there is a constant M such that $F(B(x_0, b)) \subset M\bar{B}$, where $\bar{B}$ is the closed unit ball of $\mathbb{R}^n$. It implies $\bar{F}(B(x_0, b)) \subset M\bar{B}$. F and $\bar{F}$ are all Lipschitzian mappings, hence, Inc. (2.4.5) and Inc. (2.4.6) exist solutions.

If for an $\varepsilon > 0$ given arbitrarily, we can find a $z(t) \in AC$, $z(0) = x_0$ (for convenience, it is denoted by $z(t, x_0)$) such that for $t \in I$,

$$\|\bar{x}(t, x_0) - z(t, x_0)\| < \frac{\varepsilon}{2} \quad (2.4.7)$$

and

$$d(\dot{z}(t, x_0), F(z(t, x_0))) \leq \frac{l\varepsilon}{2(e^{lT} - 1)} \quad (2.4.8)$$

then by the notations used in Theorem 2.3.3, we have $\delta = 0$ and $\rho(t) = \frac{l\varepsilon}{2(e^{lT} - 1)}$. Theorem 2.3.3 asserts that there is a solution $x \in S_{[0, T]}(F, x_0)$ such that

$$\|x(t, x_0) - z(t, x_0)\| \leq \int_0^t e^{l(t-s)} \frac{l\varepsilon}{2(e^{lT} - 1)} ds \leq \frac{\varepsilon}{2}$$

It follows $\|x(t, x_0) - \bar{x}(t, x_0)\| < \varepsilon$. The theorem is then verified. Hence, the key point is to construct a $z(t, x_0)$ which satisfies Inequalities (2.4.7) and (2.4.8).

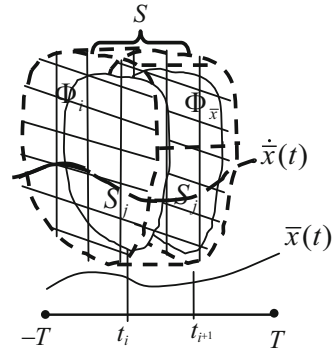
(1) We now construct $z(t, x_0)$ by the following two steps.

(i) The time interval I is divided into $2q$ equal sections, i.e., the length of each section is $\frac{T}{q}$. $I_i = [t_i, t_{i+1}]$ is denoted for the i th sub-interval and $\Phi_i = \bar{F}(\bar{x}(t_i, x_0)) \subset \mathbb{R}^n$ is the image of $\bar{F}(\bar{x}(t_i, x_0))$ which is closed and convex. It is direct to show $\bar{F}(\bar{x}(t, x_0))$ is a Lipschitzian mapping and its Lipschitzian constant can be lM (to see Problem 8 of this section). It follows that for all $t \in I_i$, $\bar{F}(\bar{x}(t, x_0)) \subset \Phi_i + \frac{lMT}{q}\bar{B}$. We denote $S = \Phi_i + \frac{lMT}{q}\bar{B}$, S is bounded, closed and convex. The S can be contained by a union of finite lattice cubes¹³, i.e., $S \subset \bigcup_{j=1}^J S_j$ where S_j 's are lattice cubes with equal side length. We can require that the diameter of the lattice cubes is less than a given positive number ζ . Figure 2.14 is given to explain the set S and these lattice cubes. The line in the lowest position is the segment I , where $\bar{x}(t, x_0)$ is defined. $\Phi_{\bar{x}}$ is the image of the set-valued mapping $\bar{F}(\bar{x}(t, x_0))$, $\Phi_{\bar{x}}$ and $\Phi_i = \bar{F}(\bar{x}(t_i, x_0))$ are surrounded by full line (both $\Phi_{\bar{x}}$ and Φ_i are slices in Fig. 2.14). S is a stereoscopic cube and is surrounded by dash lines. S contains Φ_i .¹⁴ These straight lines separate S into lattice cubes. S_j is one lattice cube, whose diameter is its diagonal line (no drawing). We can increase the density of these lattice vertexes so that the diameter is less than any pointed real number ζ . $\bar{x}(t, x_0) \in \bar{F}(\bar{x}(t, x_0))$, hence $\bar{x}(t, x_0)$ passes through $\Phi_{\bar{x}}$ in only one S_j .

¹³A lattice cube is a regular cube whose vertexes are rational numbers $\frac{q}{m}$, $q \in \{0, \pm 1, \pm 2, \dots\}$ and length of side is $\frac{1}{m}$.

¹⁴ S looks to like a sandwich and Φ_i likes a piece of ham in the sandwich.

Fig. 2.14 Diagrammatic drawing for the proof of Theorem 2.4.9



A set $E_j = \{t; t \in I_i, \dot{\bar{x}}(t, x_0) \in S_j\}$ and a step function $\xi(t) = \sum_j \xi_j \chi_j(t)$ are defined where $\xi_j \in S_j \cap S$ is a determined vector and $\chi_j(t)$ is the characteristic function of E_j , i.e.,

$$\chi_j(t) = \begin{cases} 1, & t \in E_j, \\ 0, & t \notin E_j. \end{cases}$$

It is obvious that $d(\xi_j, \Phi_i) \leq \frac{lMT}{q}$, $\|\dot{\bar{x}}(t, x_0) - \xi(t)\| \leq \zeta, t \in I_i$

(ii) The function $\xi(t) = \sum_j \xi_j \chi_j(t)$ is a simple function of S_j , we now construct a simple function for $F(\bar{x}(t_i, x_0))$.

Recall $\xi_j \in S_j \cap S \subset S = \Phi_i + \frac{lMT}{q}\bar{B}$. $\Phi_i = \bar{F}(\bar{x}(t_i, x_0))$ is the closed convex hull of $F(\bar{x}(t_i, x_0))$, hence, there exist k vectors $z_{jh} \in F(\bar{x}(t_i, x_0))$ $h = 1, 2, \dots, k$, and $k \leq n + 1$, such that

$$\left\| \xi_j - \sum_{h=1}^k \alpha_{jh} z_{jh} \right\| \leq \frac{2lMT}{q}, \quad \sum_{h=1}^k \alpha_{jh} = 1, \alpha_{jh} > 0. \quad (2.4.9)$$

We now define $\psi(t) = \int_{t_i}^t \chi_j(s) ds$, the $\psi(t)$ is monotonous and $\psi(t_{i+1}) = m(E_j)$,

where $m(E_j)$ is the Lebesgue measure of the set E_j . Real numbers $t_i < \tau_1 < \tau_2 < \dots < \tau_k$ can be obtained

$$\tau_h = \sup \left\{ t; \psi(t) \leq m(E_j) \sum_{p=1}^h \alpha_{jp} \right\}$$

These τ_h 's, $h = 1, 2, \dots, k$ define k intervals $[\tau_0, \tau_1], (\tau_1, \tau_2], \dots, (\tau_{k-1}, \tau_k]$ with $\tau_0 = t_i$ and $\tau_k = t_{i+1}$. Set E_{jh} is then defined as $E_{jh} = E_j \cap (\tau_{h-1}, \tau_h]$, $h = 1, 2, \dots, k$, then these E_{jh} 's hold the properties that $E_{jh_1} \cap E_{jh_2} = \emptyset$, $h_1 \neq h_2$ and $\bigcup_{h=1}^k E_{jh} = E_j$. Let $\chi_{jh}(t)$ be the characteristic function of E_{jh} . Then on the set E_j a simple function $\zeta_j(t)$ is defined as follows

$$\zeta_j(t) = \sum_h z_{jh} \chi_{jh}(t)$$

where $z_{jh} \in F(\bar{x}(t_i, x_0))$ defined by Inequality (2.4.9). Define now $\zeta(t) = \sum_j \zeta_j(t)$ and finally

$$z(t, x_0) = x_0 + \int_0^t \zeta(s) ds \quad (2.4.10)$$

(2) We prove that $z(t, x_0)$ defined by Eq. (2.4.10) can meet the requirements of Inequalities (2.4.7) and (2.4.8).

At first, by the construction of $z(t, x_0)$, we have $\dot{z}(t, x_0) = \zeta(t) \in F(B(x_0, b)) \subset M\bar{B}$, and $z(t, x_0)$ is a Lipschitzian function and its Lipschitzian constant can be M , the same as $\bar{x}(t, x_0)$. For every $t \in [-T, T]$, there is an subinterval I_i , such that $t \in I_i$, we have

$$\|\bar{x}(t, x_0) - z(t, x_0)\| \leq \|\bar{x}(t, x_0) - \bar{x}(t_i, x_0)\| + \|\bar{x}(t_i, x_0) - z(t_i, x_0)\| + \|z(t_i, x_0) - z(t, x_0)\|.$$

By the Lipschitzian condition, we have

$$\|\bar{x}(t, x_0) - \bar{x}(t_i, x_0)\| \leq \frac{MT}{q}, \quad \|z(t_i, x_0) - z(t, x_0)\| \leq \frac{MT}{q}.$$

Recall the construction that $\xi(t) = \sum_j \xi_j \chi_j(t)$ and $\zeta(t) = \sum_j \sum_k z_{jk} \chi_{jk}(t)$,

consequently, $\int_{I_i} \xi(s) ds = \sum_j m(E_j) \xi_j$, and

$$\int_{I_i} \zeta(s) ds = \int_{I_i} \sum_j \sum_k z_{jk} \chi_{jk}(s) ds = \sum_j \sum_k z_{jk} \int_{I_i} \chi_{jk}(s) ds = \sum_j \sum_k z_{jk} \alpha_{jk} m(E_j).$$

Therefore,

$$\begin{aligned}
 \left\| \int_{I_i} \xi(s) ds - \int_{I_i} \rho(s) ds \right\| &\leq \sum_j m(E_j) \left\| \xi_j - \sum_k \alpha_{jk} z_{jk} \right\| \\
 &\leq \sum_j m(E_j) \frac{2lMT}{q} \\
 &\leq m(I_i) \frac{2lMT}{q} \\
 &= \frac{2lMT^2}{q^2}.
 \end{aligned}$$

At the time t_i , we have

$$\begin{aligned}
 \left\| z(t_i, x_0) - x_0 - \int_0^{t_i} \xi(s) ds \right\| &= \left\| \int_0^{t_i} \rho(s) ds - \int_0^{t_i} \xi(s) ds \right\| \\
 &\leq \sum_{k=1}^i \int_{I_k} \|\rho(s) - \xi(s)\| ds \\
 &\leq \frac{4lMT^2}{q}.
 \end{aligned}$$

On the other hand,

$$\left\| \bar{x}(t_i, x_0) - x_0 - \int_0^{t_i} \xi(s) ds \right\| = \left\| \int_0^{t_i} \dot{\bar{x}}(s, x_0) - \xi(s) ds \right\| \leq \int_0^{t_i} \|\dot{\bar{x}}(s, x_0) - \xi(s)\| ds \leq \zeta T.$$

The both inequalities lead to $\|\bar{x}(t_i, x_0) - z(t_i, x_0)\| \leq \zeta T + \frac{4lMT^2}{q}$.

Summing up the above procedure, we have

$$\|\bar{x}(t, x_0) - z(t, x_0)\| \leq \frac{2MT}{q} + \zeta T + \frac{4lMT^2}{q} \quad (2.4.11)$$

for every $t \in [-T, T]$.

Recall that S_j is a lattice cube. If its length of side is $\frac{1}{q}$, then its diameter is $\frac{\sqrt{n}}{q}$. Inequality (2.4.11) leads to

$$\|\bar{x}(t, x_0) - z(t, x_0)\| \leq \frac{\lambda_1}{q} \quad (2.4.12)$$

where $\lambda_1 = (2M + \sqrt{n} + 4lMT) T$.

Because $\dot{z}(t, x_0) \in F(\bar{x}(t_i, x_0))$ for $t \in I_i$, $d(\dot{z}(t), F(\bar{x}(t, x_0))) \leq \frac{lMT}{q}$. By Inequality (2.4.12), we have $F(x(t)) \subset F(z(t)) + \frac{\lambda_1 l}{q}B$. Using triangle inequality, at last

$$d(\dot{z}(t), F(z(t))) \leq \frac{lMT}{q} + \frac{\lambda_1 l}{q}$$

Therefore, when

$$q = \min \left\{ \left\lfloor \frac{2\lambda_1}{\varepsilon} \right\rfloor, \left\lfloor \frac{2(e^{lT} - 1)(MT + \lambda_1)}{\varepsilon l} \right\rfloor \right\} + 1$$

where $\lfloor \alpha \rfloor$ is used to express the largest integer which is less or equal to α . Inequalities (2.4.7) and (2.4.8) are both satisfied. We complete the proof. $\square$

Theorem 2.4.9 is also named by relaxation theorem. We emphasize that the Lipschitzian condition for set-valued mapping F is necessary for the validation of the theorem.

Problems

1. Suppose $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is an upper semi-continuous and bounded set-valued mapping. If it is with convex and closed value, then prove $S_{[0,T]}(F, C)$ is a compact set in AC where $C \subset \mathbb{R}^n$ is a compact set. Furthermore, $R_{[0,T]}(F, x_0)$ is also compact.
2. Prove Lemma 2.4.1.
3. Prove Theorem 2.4.5.
4. Suppose $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a Lipschitzian set-valued mapping, and is with closed value. By using Theorem 2.3.3, prove the former part of Theorem 2.4.3, i.e., if we treat $S_{[0,T]}(F, x_0)$ as a set-valued mapping from $\mathbb{R}^n$ to AC $([0, T], \mathbb{R}^n)$, then the mapping is also Lipschitzian.
5. Suppose $F(x) \equiv \{-1, 1\}$. Prove the F is upper semi-continuous and with closed value, but is not with convex value. Consider the differential inclusion $\dot{x} \in F(x), x(0) = 0$. Prove that a saw-toothed function

$$x_k(t) = \begin{cases} t - \frac{i}{2^k}, & \frac{i}{2^k} \leq t < \frac{i+1}{2^k}, \\ \frac{i+2}{2^k} - t, & \frac{i+1}{2^k} \leq t < \frac{i+2}{2^k}, \end{cases} \quad i = 0, 1, 2, \dots$$

is its solution for every $k \in \mathbb{N}$. The limitation of $x_k(t)$, $x(t) \equiv 0$ when $k \rightarrow \infty$, is not a solution. Hence, Theorem 2.4.7 fails.

6. Suppose $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a bounded and upper semi-continuous mapping, and with closed and convex value. $C \subset \mathbb{R}^n$ is a compact set. We construct a mapping $\text{gra}(F) \rightarrow \text{gra}_{S_{[0,T]}}(F, C)$. Prove that the mapping is continuous, i.e., for every $\varepsilon > 0$, there exists a $\delta > 0$, if $\text{gra}G \subset \text{gra}F + \delta B$ then

$$\text{gra}_{S_{[0,T]}}(G, C) \subset \text{gra}_{S_{[0,T]}}(F, C) + \varepsilon B.$$

7. $F : \Lambda \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, $(\lambda, x) \mapsto F(\lambda, x)$ is a set-valued mapping with parameter λ . Suppose F is an upper semi-continuous mapping, and with convex and closed value for every (λ, x) . Then for every $\varepsilon > 0$, there is a $\delta > 0$ such that when $|\lambda - \lambda_0| < \delta$

$$\text{gra}S_{[0,T]}(F(\lambda, x(t)), x_0) \subset \text{gra}S_{[0,T]}(F(\lambda_0, x(t)), x_0) + \varepsilon B.$$

8. Suppose $F : \mathbb{R}^2 \rightarrow \mathbb{R}^2$ is as

$$F\left(\begin{bmatrix} x_1 \\ x_2 \end{bmatrix}\right) = \left[\begin{array}{c} \{-1, 1\} \\ \sqrt{|x_2|} + |x_1| \end{array} \right]$$

Prove the following conclusions.

- 1) F is with compact value;
- 2) F is not a Lipschitzian mapping;
- 3) Consider the Cauchy problem for $\dot{x}(t) \in F(x(t))$, $x(0) = 0$. Prove that the relaxation theorem (Theorem 2.4.9) is not true for the example.
9. Suppose $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a Lipschitzian mapping with Lipschitzian constant l , and $F(x) \subset MB$ for every $x \in \mathbb{R}^n$. If $\dot{x}(t) \in F(x(t))$ be a solution of the differential inclusion, then $t \mapsto F(x(t))$ is a Lipschitzian set-valued mapping whose Lipschitzian constant can be lM .

2.5 Stability of Differential Inclusions

In both theories of differential equations and control systems, stability is always a very fundamental issue. This section deals with the theory of stability for differential inclusions, and the elemental method is the Lyapunov direct method. To introduce the method, we define Dini derivatives firstly. Then the stabilities for differential inclusion are defined. At last, the theorems for stabilities are presented by using Lyapunov functions.

2.5.1 Dini Derivatives

This subsection deals with the concept of Dini derivatives and their main properties.

Suppose $f : [\alpha, \beta] \rightarrow \mathbb{R}$ is a single-valued function and $x_0 \in (\alpha, \beta)$, then

$$D^+f(x_0) = \limsup_{\Delta x \downarrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

$$D_+f(x_0) = \liminf_{\Delta x \downarrow 0} \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

$$D^-f(x_0) = \lim_{\Delta x \uparrow 0} \sup \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

$$D_-f(x_0) = \lim_{\Delta x \uparrow 0} \inf \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$$

are defined to be the right upper Dini derivative, the right lower Dini derivative, the left upper Dini derivative and the left lower Dini derivative, of $f(x)$ at x_0 , respectively.

By the decrease of Δx , $\sup \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x}$ decreases monotonously, the right upper Dini derivative is existed no matter what is the function $f(x)$. Of course, $D^+f(x_0)$ may be equal to infinite. By the same reason, we can conclude that $D_+f(x_0)$, $D^-f(x_0)$ and $D_-f(x_0)$ all exist, of course they may be equal to infinite. However, if $f(x)$ is a Lipschitzian function, then we can proof the four derivatives all have finite values.

Lemma 2.5.1 If $f : (\alpha, \beta) \rightarrow \mathbb{R}$ is a continuous function, then $f(x)$ is a monotonously increasing if and only if one of the following conditions holds at every $x_0 \in (\alpha, \beta)$:

- (1) $D^+f(x_0) \geq 0$;
- (2) $D_+f(x_0) \geq 0$;
- (3) $D^-f(x_0) \geq 0$;
- (4) $D_-f(x_0) \geq 0$.

Proof We only prove the first condition, the others are similar and left to readers.

If $f(x)$ is monotonously increasing on (α, β) , then $f(x_0 + \Delta x) - f(x_0) \geq 0$ for every $x_0 \in (\alpha, \beta)$. Hence, when $\Delta x > 0$, $\sup \frac{f(x_0 + \Delta x) - f(x_0)}{\Delta x} \geq 0$, it follows $D^+f(x_0) \geq 0$.

We now prove the sufficiency. At first, we consider the case that $D^+f(x_0) > 0$. If there are $x_1, x_2 \in (\alpha, \beta)$. $x_1 < x_2$, but $f(x_1) > f(x_2)$. Then there exists a constant M such that $f(x_1) > M > f(x_2)$. We define a set $S = \{x; x \in [x_1, x_2], f(x) \geq M\}$ which is a subset of (α, β) . $S \neq \emptyset$ since $(x_1, x_1 + \delta_{x_1}) \subset S$ for some small $\delta_{x_1} > 0$. Let $\xi = \sup S$. Then $\xi < x_2 \leq \beta$ and $f(\xi) = M$ since $f(x)$ is continuous at $\xi \in (\alpha, \beta)$. For $x \in (\xi, \xi + \delta)$ where δ is a small positive number, $f(x) < M = f(\xi)$. Thus, $\frac{f(x) - f(\xi)}{x - \xi} < 0$, $x \in (\xi, \xi + \delta)$. It leads to $D^+f(\xi) \leq 0$. A contradiction appears.

If there is some $x_0 \in (\alpha, \beta)$, such that $D^+f(x_0) = 0$, then we consider a function $g(x) = f(x) + \varepsilon x$ with $\varepsilon > 0$. It is obvious that $D^+g(x_0) = D^+f(x_0) + \varepsilon > 0$. Then $g(x)$ is an increasing function, i.e., $f(x_1) + \varepsilon x_1 = g(x_1) \leq g(x_2) = f(x_2) + \varepsilon x_2$ if $x_1 < x_2$. Or equivalently, $f(x_1) \leq f(x_2) + \varepsilon(x_2 - x_1)$. Because $\varepsilon > 0$ can be selected arbitrarily, we have $f(x_1) \leq f(x_2)$. $\square$

Remark 1 Theorem 2.5.1 is still valid if the domain $[\alpha, \beta]$ is replaced by (α, β) if at both terminals α and β only the single-side derivatives are considered. $\square$

Remark 2 In general, if $f(x)$ is monotonously increasing on $B(x_0, \delta)$ we cannot conclude that $D^+f(x_0) > 0$. $\square$

Remark 3 Theorem 2.5.1 illustrates that for the verification of monotonicity of a continuous function the four Dini derivatives have the same function. A further conclusion is if $f(x)$ is monotonous on a $(\alpha, \beta) \subset \mathbb{R}$, then for almost all point in (α, β) the four derivatives are equal and take finite values. $\square$

Lemma 2.5.2 Suppose $f, g : [\alpha, \beta] \rightarrow \mathbb{R}$ are two continuous functions, if for every $x \in (\alpha, \beta) D^+f(x) \leq g(x)$, then

$$f(\beta) - f(\alpha) \leq \int_{\alpha}^{\beta} g(x) dx$$

Proof Because $D^+f(\alpha) \leq g(\alpha)$, for every $\varepsilon > 0$, there exists a $\delta_1 = \delta(\alpha, \varepsilon) > 0$ such that $\alpha + \delta_1 < \beta$, and for every $\Delta x \in [0, \delta_1]$,

$$\frac{f(\alpha + \Delta x) - f(\alpha)}{\Delta x} \leq g(\alpha) + \varepsilon$$

Let $\Delta x = \delta_1$. Then the inequality leads to $f(\alpha + \delta_1) - f(\alpha) \leq \delta_1 g(\alpha) + \delta_1 \varepsilon$.

Similarly, by $D^+f(\alpha + \delta_1) \leq g(\alpha + \delta_1)$, there is a $\delta_2 = \delta(\alpha + \delta_1, \varepsilon) > 0$ such $\alpha + \delta_1 + \delta_2 < \beta$ and

$$f(\alpha + \delta_1 + \delta_2) - f(\alpha + \delta_1) \leq \delta_2 g(\alpha + \delta_1) + \delta_2 \varepsilon$$

The procedure can be repeated before $\alpha + \sum_{i=1}^K \sigma_i = \beta$.¹⁵ In the $(k+1)$ th step, we obtain

$$f\left(\alpha + \sum_{i=1}^{k+1} \delta_i\right) - f\left(\alpha + \sum_{i=1}^k \delta_i\right) \leq \delta_{k+1} g\left(\alpha + \sum_{i=1}^k \delta_i\right) + \delta_{k+1} \varepsilon, 1 \leq k \leq K.$$

Summing up these inequalities yields

$$f(\beta) - f(\alpha) \leq \sum_{i=1}^K \delta_i g\left(\alpha + \sum_{k=0}^{i-1} \delta_k\right) + (\beta - \alpha) \varepsilon \quad (2.5.1)$$

¹⁵It is reasonable because on the interval $[\alpha, \beta]$, $f(x)$ is equicontinuous.

where $\delta_0 = 0$. Because g is a continuous function, it is Lebesgue integrable. When $\max \delta_i \rightarrow 0$, the right side of Inequality (2.5.1) converges to $\int_{\alpha}^{\beta} g(x)dx + (\beta - \alpha) \varepsilon$.

The ε is selected arbitrarily, hence, the conclusion is verified. $\square$

We now extend the definition of Dini derivatives to $\mathbb{R}^n$. Because there are infinite directions for a point in $\mathbb{R}^n$, we have to revise the definition.

Definition 2.5.1 Suppose $f : \mathbb{R}^n \rightarrow \mathbb{R}$ is a single-valued mapping, $x_0, v \in \mathbb{R}^n$, then

$$D^+f(x_0)(v) = \lim_{h \downarrow 0} \sup_{v' \rightarrow v} \frac{f(x_0 + hv') - f(x_0)}{h}$$

$$D^-f(x_0)(v) = \lim_{h \downarrow 0} \inf_{v' \rightarrow v} \frac{f(x_0 + hv') - f(x_0)}{h}$$

are called to be the upper Dini derivative and lower Dini derivative, respectively, of f at x_0 along with the vector v . $\square$

Readers may find that the definition of Dini derivatives is quite similar to the direction derivative of the convex function defined in Sect. 1.3 except the limitation of $v' \rightarrow v$ which leads to supremum and infimum. The introduction of supremum and infimum assures the existence of Dini derivatives in $\mathcal{R}$, the set of closed real numbers. However, if f is a Lipschitzian function, then Dini derivatives are all finite.

Readers may find that notations used for n -dimensional function are quite different from those for $\mathbb{R}$. For example, by using the notating method for multivariable case, the upper Dini derivative $D^+f(x_0)$ given at the beginning of this subsection should be changed to $D^+f(x_0)(1)$. In the notation $D^+f(x_0)(1)$, the superscript “+” means supremum that differs from that in $D^+f(x_0)$ where means convergence from right side, and “1” means the right limitation. By such a notating method $D^-f(x_0)$ should be changed to $D^+f(x_0)(-1)$, $D_+f(x_0)$ should to $D^-f(x_0)(1)$, and $D_-f(x_0)$ to $D^-f(x_0)(-1)$.

The following theorem plays a basic role in the direct Lyapunov method.

Theorem 2.5.1 Suppose $x(t) : [0, T] \rightarrow \Omega \subset \mathbb{R}^n$ is a solution of n -dimensional differential equation $\dot{x} = f(x)$. Suppose $V : \mathbb{R}^n \rightarrow \mathbb{R}$ is a continuously single-valued and Lipschitzian function. Then the compounded function $V(x(t)) : [0, T] \subset \mathbb{R} \rightarrow \mathbb{R}$ satisfies that

$$D^+V(x(t))(1) = \lim_{h \downarrow 0} \sup \frac{V(x(t) + hf(x(t))) - V(x(t))}{h}$$

$$D^-V(x(t))(1) = \lim_{h \downarrow 0} \inf \frac{V(x(t) + hf(x(t))) - V(x(t))}{h}$$

Proof By definition of upper Dini derivative, we have

$$\begin{aligned}
 D^+ V(x(t)) (1) &= \lim_{h \downarrow 0} \sup \frac{V(x(t+h)) - V(x(t))}{h} \\
 &= \lim_{h \downarrow 0} \sup_{o(h)} \frac{V(x(t) + hf(x(t)) + o(h)) - V(x(t))}{h} \\
 &\leq \lim_{h \downarrow 0} \sup_{o(h)} \frac{V(x(t) + hf(x(t)) + L|o(h)| - V(x(t))}{h} \\
 &= \lim_{h \downarrow 0} \sup \frac{V(x(t) + hf(x(t)) - V(x(t))}{h},
 \end{aligned}$$

where $x(t+h) = x(t) + \dot{x}(t)h + o(h) = x(t) + hf(x(t)) + o(h)$ is the Taylor expansion. $o(h)$ is an infinitesimal with higher order than h . At the third step, we apply mean value theorem where L is the Lipschitzian constant. On the other hand,

$$\begin{aligned}
 D^+ V(x(t)) (1) &= \lim_{h \downarrow 0} \sup \frac{V(x(t) + hf(x(t)) + o(h)) - V(x(t))}{h} \\
 &\geq \lim_{h \downarrow 0} \sup \frac{V(x(t) + hf(x(t)) - L|o(h)| - V(x(t))}{h} \\
 &= \lim_{h \downarrow 0} \sup \frac{V(x(t) + hf(x(t)) - V(x(t))}{h}.
 \end{aligned}$$

The first equation is verified. The second one can be verified by a similar way. It is omitted. $\square$

Theorem 2.5.1 illustrates that Dini derivatives can be applied in direct Lyapunov method as the same as general derivative.

2.5.2 Definitions of Stability of Differential Inclusions

The stability of differential inclusions is defined by following the Lyapunov theory for differential equations. Hence, we start with the concept of equilibrium.

Suppose $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a set-valued mapping. If $0 \in F(x_0)$ for an $x_0 \in \mathbb{R}^n$, then x_0 is called as an equilibrium of the differential inclusion $\dot{x}(t) \in F(x(t))$ since $x(t) \equiv x_0$ is a solution of the differential inclusion. It is clear that a differential inclusion may hold infinite equilibriums.

If $0 \in F(x_0)$, then we can make a coordinate transformation $y = x - x_0$ and denote $\widehat{F}(y) = F(y + x_0)$. By such a transformation, $0 \in \widehat{F}(0)$, i.e., $y(t) \equiv 0$ is an equilibrium of differential inclusion $\dot{y}(t) \in \widehat{F}(y(t))$. Therefore, we always assume in the discussion of stability that the equilibrium is the origin.

We now introduce the **K**-class functions and **L**-class functions.

$\alpha : \mathbb{R}(\geq 0) \rightarrow \mathbb{R}(\geq 0)$ is a function of **K**-class if $\alpha(0) = 0$ and it is continuous and strictly monotonously increasing; additionally, if $\alpha(t) \rightarrow \infty (t \rightarrow \infty)$, then it is a function of **KR**-class. The facts are denoted by $\alpha \in \mathbf{K}$ and $\alpha \in \mathbf{KR}$, respectively.

$\beta : \mathbb{R}(\geq 0) \rightarrow \mathbb{R}(\geq 0)$ is a function of **L**-class if it is continuous and strictly monotonously decreasing, and $\beta(\infty) = 0$. The fact is denoted by $\beta \in \mathbf{L}$.

The two concepts can be compounded. For example, $k : \mathbb{R}(\geq 0) \times \mathbb{R}(\geq 0) \rightarrow \mathbb{R}(\geq 0)$ is a **KL**-class function if we fix the second variable it is a function of **K**-class, and if we fix the first variable it is in class **L**. Similarly, we have classes **KKL**, **KLL** and **KRL**, etc.

We list several fundamental properties of **K**-class functions. If $\alpha_1, \alpha_2 \in \mathbf{K}$, then $\alpha_1 + \alpha_2 \in \mathbf{K}$, $\alpha_1(\alpha_2) \in \mathbf{K}$ and $k\alpha_1 \in \mathbf{K}$ for $k > 0$. If $\alpha \in \mathbf{K}$ then α is invertible and its inverse $\alpha^{-1} \in \mathbf{K}$. The discussion about **L**-class functions is left to readers.

$V : \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be positively definite if it satisfies that $V(x) > 0$ for $x \neq 0$ and $V(0) = 0$. The $V(x)$ is said to be semi-positively definite, if $V(0) = 0$ and $V(x) \geq 0$ for all $x \in \mathbb{R}^n$. Additionally, if $V(x)$ satisfies $V(x) \rightarrow \infty (x \rightarrow \infty)$ then the function is called by positively (semi-positively) definite infinity. A function $W : \mathbb{R}^n \rightarrow \mathbb{R}$ is said to be negatively (semi-negatively) definite if $-W$ is positively (semi-positively) definite. We can also define negatively (semi-negatively) definite infinite. The positively definite and negatively definite functions play important roles in the study of stability by using Lyapunov method. Usually, a positively definite function $V(x)$ is also called V -function for simplicity.

Lemma 2.5.3 If $V(x)$ is continuous and positively definite, then there are $\alpha_1, \alpha_2 \in \mathbf{K}$ such that

$$\alpha_2(\|x\|) \leq V(x) \leq \alpha_1(\|x\|)$$

Additionally, if $V(x)$ is positively definite infinity, then $\alpha_1, \alpha_2 \in \mathbf{KR}$.

Proof We prove the first part, and the second part can be proved by a similar way. Hence, we omit it.

Define $\phi_m(\|x\|) = \min_{\|r\| \geq \|x\|} V(r)$, then $\phi_m(\lambda) : \mathbb{R} \rightarrow \mathbb{R}$ is monotonously increasing and is continuous since $V(x)$ is continuous. It is obvious that $\phi_m(\|x\|) \leq V(x)$. Let $\alpha_2(\|x\|) = \frac{\|x\|}{\|x\|+1} \phi_m(\|x\|)$. Then $\alpha_2 \in \mathbf{K}$ and $\alpha_2(\|x\|) < \phi_m(\|x\|) \leq V(\|x\|)$.

Define $\phi_M(\|x\|) = \max_{\|r\| \leq \|x\|} V(r)$, then $\phi_M(\lambda) : \mathbb{R} \rightarrow \mathbb{R}$ is also monotonously increasing and continuous. It is obvious that $\phi_M(\|x\|) \geq V(x)$. Let $\alpha_1(\|x\|) = \phi_M(\|x\|) + \|x\|$. Then $\alpha_1 \in \mathbf{K}$ and $\alpha_1(\|x\|) > \phi_M(\|x\|) \geq V(\|x\|)$. $\square$

The following lemma is fundamental for the stability study by using **K**-class functions and **L**-class functions. The proof is referred to (Sontag 1989) and is omitted.

Lemma 2.5.4 Suppose $\alpha \in \mathbf{K}$, $y : \mathbb{R} \rightarrow \mathbb{R}$ is the solution of the following Cauchy problem

$$\dot{y} = -\alpha(y), \quad y(t_0) = y_0 \quad (2.5.2)$$

where $y_0 > 0$, then there exists a $\beta \in \mathbf{KL}$ such that $|y(t)| \leq \beta(y_0, t - t_0)$. $\square$

We now turn to deal with the stability of differential inclusions. Let $S(F, x_0)$ ¹⁶ be the set of solutions of Cauchy problem of differential inclusion

$$\dot{x}(t) \in F(x(t)) \quad x(0) = x_0 \quad (2.5.3)$$

Note that we have assumed throughout at this chapter that $0 \in F(0)$, and the investigation of stability is carried out about the origin.

Definition 2.5.2 Consider the following differential inclusion

$$\dot{x}(t) \in F(x(t)) \quad (2.5.4)$$

If for every $\varepsilon > 0$, there exists a $\delta = \delta(\varepsilon)$ such that:

- (1) When $\|x_0\| < \delta$, for every $x(t, x_0) \in S(F, x_0)$, $\|x(t, x_0)\| < \varepsilon$,¹⁷ then Inc. (2.5.4) is strongly stable with regard to the equilibrium $x(t) \equiv 0$;
- (2) If there is an $x(t, x_0) \in S(F, x_0)$ such that $\|x(t, x_0)\| < \varepsilon$, then Inc. (2.5.4) is weakly stable with regard to the equilibrium $x(t) \equiv 0$;
- (3) With regard to the equilibrium $x(t) \equiv 0$, Inc. (2.5.4) is unstable if it is not weakly stable. $\square$

Definition 2.5.3

- (1) If Inc. (2.5.4) is strongly stable, additionally, for every $x(t, x_0) \in S(F, x_0)$ such that

$$\lim_{t \rightarrow \infty} x(t, x_0) = 0$$

then it is strongly asymptotically stable with regard to the equilibrium $x(t) \equiv 0$;

- (2) If Inc. (2.5.4) is weakly stable, additionally, there is an $x(t, x_0) \in S(F, x_0)$ such that

$$\lim_{t \rightarrow \infty} x(t, x_0) = 0$$

¹⁶When we deal with the stability of differential equation or inclusion, the existence interval of a solution is always $[t_0, \infty)$, where t_0 is the initial time. Hence the subscript $[t_0, \infty)$ is omitted.

¹⁷The norm is $\|x(t, x_0)\|_{\mathbb{R}} = \sqrt{\max_t x_1^2(t, x_0) + \max_t x_2^2(t, x_0) + \cdots + \max_t x_n^2(t, x_0)}$. It is a function of t . For the sake of convenience we omit the subscript $\mathbb{R}$.

then it is weakly asymptotically stable with regard to the equilibrium $x(t) \equiv 0$. $\square$

Remark The stabilities of Inc. (2.5.4) can be equivalently defined by using **K**-class and **KL**-class functions. For example, with regard to the equilibrium $x(t) \equiv 0$ Inc. (2.5.4) is strongly stable if there is an $\alpha \in \mathbf{K}$, such that for every $x(t, x_0) \in S(F, x_0)$ $\|x(t, x_0)\| \leq \alpha(\|x_0\|)$ in a neighborhood of the origin of $\mathbb{R}^n$. With regard to the equilibrium $x(t) \equiv 0$ Inc. (2.5.4) is strongly asymptotically stable if there is a $\beta \in \mathbf{KL}$, such that for every $x(t, x_0) \in S(F, x_0)$, $\|x(t, x_0)\| \leq \beta(\|x_0\|, t)$ in a neighborhood of the origin of $\mathbb{R}^n$. Readers can try to give the definitions for weak stability and weakly asymptotical stability for Inc. (2.5.4) by using **K**-class and **KL**-class functions. $\square$

Usually, the word “strongly” in strongly stable and strongly asymptotically stable is omitted. Furthermore, we only consider the stability with regard to the equilibrium $x(t) \equiv 0$, hence the sentence “with regard to the equilibrium $x(t) \equiv 0$ ” is also omitted. We just call that Inc. (2.5.4) is stable, weakly stable, asymptotically stable, weakly asymptotically stable, and unstable, respectively.

2.5.3 Lyapunov-like Criteria for Stability of Differential Inclusions

The direct Lyapunov method is the uppermost method used in the investigation of differential equations as well as of differential inclusions. This subsection presents several Lyapunov-like criteria for differential inclusions.

We continue to consider Inc. (2.5.4). It is assumed that the set-valued mapping $F(x)$ is upper semi-continuous, and with convex and compact value.

Theorem 2.5.2 If there exist a positive constant η , a positively definite and continuous function $V : \mathbb{R}^n \rightarrow \mathbb{R}$, and a semi-negatively definite function $W : \mathbb{R}^n \rightarrow \mathbb{R}$ such that for every $\|x\| \leq \eta$ and $v \in F(x)$,

$$D^+V(x)(v) \leq W(x)$$

then Inc. (2.5.4) is stable.

Proof Let $x(t, x_0) \in S(F, x_0)$ be a solution. We now fix the time t and consider Dini derivative

$$D^+V(x(t, x_0))(1) = \lim_{h \downarrow 0} \sup \frac{V(x(t+h, x_0)) - V(x(t, x_0))}{h}$$

By the definition of supremum, there is sequence $\{h_n, n = 1, 2, \dots\}$ such that $h_n \downarrow 0$ ($n \rightarrow \infty$) and

$$D^+V(x(t, x_0))(1) = \lim_{n \rightarrow \infty} \frac{V(x(t+h_n, x_0)) - V(x(t, x_0))}{h_n}$$

Denote $v_n = \frac{x(t+h_n, x_0) - x(t, x_0)}{h_n}$. Then $x(t+h_n, x_0) = x(t, x_0) + h_n v_n$. Because $\dot{x}(t, x_0) \in F(x(t, x_0))$, for $\gamma > 0$ given arbitrarily, there exists an N , when $n > N$

$$v_n = \frac{x(t+h_n, x_0) - x(t, x_0)}{h_n} \in F(x(t, x_0)) + \gamma \bar{B}$$

Because $F(x(t, x_0))$ is a compact set, $\{v_n\}$ holds a convergent subsequence. Without loss of generality, we can assume that $v_n \rightarrow v$, then $v \in F(x(t, x_0))$. Thus,

$$\begin{aligned} D^+ V(x(t, x_0))(1) &= \limsup_{n \rightarrow \infty} \frac{V(x(t+h_n, x_0)) - V(x(t, x_0))}{h_n} \\ &= \limsup_{n \rightarrow \infty} \frac{V(x(t, x_0) + h_n v_n) - V(x(t, x_0))}{h_n} \\ &\leq D^+ V(x(t, x_0))(v) \\ &\leq W(x(t, x_0)) . \end{aligned}$$

This inequality is obtained from the condition that $D^+ V(x)(v) \leq W(x)$ for every $v \in F(x)$. This is a very useful conclusion.

By Lemma 2.5.2, for $t > 0$, we have

$$V(x(t, x_0)) - V(x(0)) \leq \int_0^t W(x(s, x_0)) ds < 0 \quad (2.5.5)$$

or $V(x(t, x_0)) \leq V(x_0)$. Using Lemma 2.5.3, two functions $\alpha_1, \alpha_2 \in \mathbf{K}$ can be found such that $\alpha_2(\|x\|) \leq V(x) \leq \alpha_1(\|x\|)$. For every $\eta \geq \varepsilon > 0$, we can find a $\delta > 0$ such that $\alpha_1(\delta) \leq \alpha_2(\varepsilon)$. Now if $\|x_0\| < \delta$, then

$$\alpha_2(\|x(t, x_0)\|) \leq V(x(t, x_0)) \leq V(x_0) \leq \alpha_1(\|x_0\|) < \alpha_1(\delta) \leq \alpha_2(\varepsilon)$$

The last inequality implies $\|x(t, x_0)\| < \varepsilon$ for every $t \geq 0$, i.e., Inc. (2.5.4) is stable. $\square$

In the proof of Theorem 2.5.2, we have established a useful inequality that

$$D^+ V(x(t))(1) \leq D^+ V(x(t))(v) \quad (2.5.6)$$

where $v \in F(x(t))$. The left side is $D^+ V(x(t))(1)$ whose argument is time t , but in $D^+ V(x(t))(v)$ the arguments are vector $x(t)$ and vector v .

From the proof of Theorem 2.5.2, we can find that the upper semi-continuity and convexity compactness are only used to assure the existence of solutions. If we are sure that the solutions exist, then we only need the compactness to assure the convergence.

Theorem 2.5.3 If there exist a positive constant η , a positively definite and continuous function $V : \mathbb{R}^n \rightarrow \mathbb{R}$, and a negatively definite function $W : \mathbb{R}^n \rightarrow \mathbb{R}$ such that for every $\|x\| \leq \eta$ and $v \in F(x)$,

$$D^+V(x)(v) \leq W(x)$$

then Inc. (2.5.4) is asymptotically stable.

Proof The conditions of Theorem 2.5.3 implies those of Theorem 2.5.2, hence, for every $\varepsilon \in [0, \eta]$, there exists a $\delta > 0$ such that if $\|x_0\| < \delta$, then every $x(t, x_0) \in S(F, x_0)$ satisfies that $\|x(t, x_0)\| < \varepsilon$. We only need to prove that $x(t, x_0) \rightarrow 0$ ($t \rightarrow \infty$).

The conclusion is verified by reducing a contradiction. If there is a solution $\bar{x}(t, x_0) \in S(F, x_0)$ where $\|x_0\| < \delta$ and $\|\bar{x}(t, x_0)\| < \varepsilon$, but $\bar{x}(t, x_0)$ is not convergent to the origin. Therefore, there exist a constant l with $\varepsilon > l > 0$ and a sequence $\{t_n\} \subset \mathbb{R}$ (> 0) with $t_n \rightarrow \infty$ ($n \rightarrow \infty$) such that $\|\bar{x}(t_n, x_0)\| \geq l$. Now we replace ε by this l , and repeat the proof of Theorem 2.5.2. Then we can obtain a $\bar{\delta} > 0$, when $\|\bar{x}_0\| < \bar{\delta}$, $\|x(t, \bar{x}_0)\| \leq l$ for every $x(t, \bar{x}_0) \in S(F, \bar{x}_0)$. Thus, in the solution $\bar{x}(t, x_0)$, x_0 has to satisfy that $\bar{\delta} \leq \|x_0\| \leq \delta$, and $\bar{x}(t, x_0)$ to satisfy $\bar{\delta} \leq \|\bar{x}(t, x_0)\| \leq \varepsilon$.

Let $-\lambda$ be the maximum of $W(x)$ in the area $\bar{\delta} \leq \|x\| \leq \eta$. Then $\lambda > 0$. Inequality (2.5.5) leads to

$$V(\bar{x}(t, x_0)) - V(\bar{x}(0, x_0)) \leq \int_0^t W(\bar{x}(s, x_0)) ds \leq -\lambda t \quad (2.5.7)$$

It follows that $V(\bar{x}(t, x_0)) \leq -\lambda t + V(\bar{x}(0, x_0))$. The left side will intend to negative infinite. This is impossible since $V(x)$ is positive definite. Consequently, every solution converges to the origin. $\square$

The next theorem is about instability.

Theorem 2.5.4 If there exist a positive constant η , a continuous function $V : \mathbb{R}^n \rightarrow \mathbb{R}$ with $V(0) = 0$ and a negatively definite function $W : \mathbb{R}^n \rightarrow \mathbb{R}$ such that for every $\|x\| \leq \eta$ and $v \in F(x)$,

$$D^+V(x)(v) \leq W(x)$$

Moreover, for every $\delta > 0$, there is at least one $\|x\| < \delta$ such that $V(x) < 0$. Then Inc. (2.5.4) is unstable.

Proof The condition of Theorem 2.5.4 implies $\min_{\|x\| \leq \delta} V(x) < 0$ for $\delta > 0$ given arbitrarily. We denote $x_\delta = \arg \min_{\|x\| \leq \delta} V(x)$, and consider the solution $x(t, x_\delta)$. By the same procedure did in Theorem 2.5.2, when $\|x(t, x_\delta)\| \leq \eta$, Inequality (2.5.5) holds, i.e.,

$$V(x(t, x_\delta)) - V(x_\delta) = V(x(t, x_\delta)) - V(x(0)) \leq \int_0^t W(x(s, x_\delta)) ds \leq 0$$

hence, $V(x(t, x_\delta)) \leq V(x_\delta)$ for $t \in \{t \geq 0; \|x(t, x_\delta)\| \leq \eta\}$. Because $x_\delta = \arg \min_{\|x\| \leq \delta} V(x)$, it is certain that $\|x(t, x_\delta)\| \geq \delta$. Now we denote $-\lambda = \max_{\delta \leq \|x\| \leq \eta} W(x)$, then $\lambda > 0$. By a similar process of Inequality (2.5.7), we can obtain

$$V(x(t, x_\delta)) \leq V(x_\delta) - \lambda t \quad (2.5.8)$$

If for every t , $\|x(t, x_\delta)\| \leq \eta$, then Inequality (2.5.8) leads to $V(x(t, x_\delta)) \rightarrow -\infty$ ($t \rightarrow \infty$). It contradicts to the fact $V(x)$ is continuous when $\|x\| \leq \eta$. Consequently, there is a T such that $\|x(T, x_\delta)\| = \eta$. It means the solution $x(t, x_\delta)$ is unstable. $\square$

The last two theorems of this section are concerned with the weak stability. We prove firstly a lemma for $D^-V(x)(v)$.

Lemma 2.5.5 Suppose the set-valued mapping $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is bounded and Lipschitzian, and with closed and convex value. Suppose $V : \mathbb{R}^n \rightarrow \mathbb{R}$ and $W : \mathbb{R}^n \rightarrow \mathbb{R}$ are two Lipschitzian mappings. Then the following two conclusions are equivalent.

- (1) For every $x \in \mathbb{R}^n$, there is a $v \in F(x)$ such that $D^-V(x)(v) \leq W(x)$;
- (2) For every $x_0 \in \mathbb{R}^n$, there is an $x(t, x_0) \in S_{[0, \infty)}(F, x_0)$ such that for $t_2 \geq t_1 \geq 0$,

$$V(x(t_2, x_0)) - V(x(t_1, x_0)) \leq \int_{t_1}^{t_2} W(x(s, x_0)) ds$$

Proof We firstly prove that the first conclusion implies the second one.

We define a function $V_\varepsilon(x(t, x_0), t) : \mathbb{R}^n \times \mathbb{R} \rightarrow \mathbb{R}$ for $x(t, x_0) \in S_{[0, \infty)}(F, x_0)$ as follows

$$V_\varepsilon(x(t, x_0), t) = V(x(t, x_0)) - V(x_0) - \int_0^t W(x(s, x_0)) ds - \varepsilon t \quad (2.5.9)$$

where $\varepsilon > 0$ and $x_0 \in \mathbb{R}^n$ are treated as parameters. In any bonded area, $V_\varepsilon(x(t, x_0), t)$ is Lipschitzian for both variables $x(t, x_0)$ and t . It is direct to verify that $V_\varepsilon(x(0), 0) = 0$. The following proof consists of two steps.

(i) For a given $T > 0$, there is an $x(t) \in S_{[0, T]}(F, x_0)$ such that for every $t \in [0, T]$ $V_\varepsilon(x(t), t) \leq \varepsilon$.

For every $x(t) \in S_{[0, T]}(F, x_0)$, we define a subset $\theta(x(t))$ of $[0, T]$ as follows

$$\theta(x(t)) = \left\{ t; V_\varepsilon(x(t), t) \leq 0, \max_{s \in [0, t]} V_\varepsilon(x(s), s) \leq \varepsilon \right\}$$

Because $V_\varepsilon(x(0), 0) = 0$, $0 \in \theta(x(t))$, i.e., $\theta(x(t)) \neq \emptyset$. Denote $t_x = \sup \theta(x(t))$, since $V_\varepsilon(x(t), t)$ is continuous for t , we have $V_\varepsilon(x(t_x), t_x) = 0$ if $t_x \neq T$. We further define

$$\Theta = \{t_x; x(t) \in S_{[0,T]}(F, x_0)\}$$

and $\hat{t} = \sup \Theta$.

We now try to verify that there is an $\hat{x}(t) \in S_{[0,T]}(F, x_0)$ such that $\hat{t} = t_{\hat{x}}$.

By the definition of supremum, there is $\{t_k\}_k \subset \Theta$ such that $t_k \uparrow \hat{t} (k \rightarrow \infty)$. For every t_k , there is an $x_k(t) \in S_{[0,T]}(F, x_0)$ such that $t_k = t_{x_k} = \sup \theta(x_k(t))$. $F(x)$ is bounded for every x , so is $\{\dot{x}_k(t)\}$. It implies the series $\{x_k(t)\}$ is equicontinuous, hence it holds a convergent subseries. Without loss of generality, we can assume $x_k(t) \rightarrow \hat{x}(t)$, $k \rightarrow \infty$. By Corollary 2.4.1, we conclude that $\hat{x}(t) \in S_{[0,T]}(F, x_0)$. $V_\varepsilon(x, t)$ is continuous for x , hence, $\hat{t} = t_{\hat{x}}$.

The fact $\hat{t} = T$ is verified by contradiction. If $\hat{t} < T$, then $V_\varepsilon(\hat{x}(\hat{t}), \hat{t}) = 0$. For this $\hat{x}(t)$, there is a $v \in F(\hat{x}(\hat{t}))$ such that $D^-V(\hat{x}(\hat{t}))(v) \leq W(\hat{x}(\hat{t}))$ by the first condition of the lemma. Consequently,

$$\begin{aligned} D^-V_\varepsilon(\hat{x}(\hat{t}, x_0), \hat{t})(v, 1) &= D^-[V(\hat{x}(\hat{t}, x_0)) - V(x_0)](v) \\ &\quad - D^- \left[\int_0^{\hat{t}} W(\hat{x}(s, x_0)) ds + \varepsilon t \right] \quad (1) \\ &= D^-V(\hat{x}(\hat{t}, x_0))(v) - \lim_{\Delta t \rightarrow 0} \int_{\hat{t}}^{\hat{t} + \Delta t} W(\hat{x}(s, x_0)) ds - \varepsilon \\ &\leq -\varepsilon. \end{aligned} \tag{2.5.10}$$

A function $y(t)$ is now defined

$$y(t) = \hat{x}(\hat{t}) + (t - \hat{t})v, \quad \hat{t} \leq t \leq T$$

We now apply Theorem 2.3.3, and the initial time is $t_0 = \hat{t}$. Using the notations of Theorem 2.3.3, $\delta = 0$ and

$$\begin{aligned} d(\dot{y}(t), F(y(t))) &= d(v, F(\hat{x}(\hat{t}) + (t - \hat{t})v)) \\ &\leq d(F(\hat{x}(\hat{t})), F(\hat{x}(\hat{t}) + (t - \hat{t})v)) \\ &\leq l \|v\| \|t - \hat{t}\|. \end{aligned}$$

It implies that $\rho(t)$ can be $l \|v\| (t - \hat{t})$. Theorem 2.3.3 concludes that there exists an $x(t) \in S_{[\hat{t}, T]}(F, \hat{x}(\hat{t}))$ such that

$$\|y(t) - x(t)\| \leq \int_{\hat{t}}^t e^{l(t-s)} \cdot l \|v\| (s - \hat{t}) ds = \|v\| \left(\frac{1}{l} e^{l(t-\hat{t})} - \frac{1}{l} - (t - \hat{t}) e^{l(t-\hat{t})} \right)$$

For convenience, we denote $\alpha(t - \hat{t}) = \int_{\hat{t}}^t e^{l(t-s)} \cdot l \|v\| (s - \hat{t}) ds$. It is direct to obtain

$$\lim_{t \downarrow \hat{t}} \frac{\alpha(t - \hat{t})}{t - \hat{t}} = 0 \quad (2.5.11)$$

Let L be the Lipschitzian constant of V_ε . Then

$$V_\varepsilon(x(t), t) \leq V_\varepsilon(y(t), t) + L\alpha(t - \hat{t}) \quad (2.5.12)$$

By Inequality (2.5.10), we have

$$\lim_{t \downarrow \hat{t}} \frac{V_\varepsilon(\hat{x}(\hat{t}) + (t - \hat{t})v, t) - V_\varepsilon(\hat{x}(\hat{t}), \hat{t})}{t - \hat{t}} \leq -\varepsilon \quad (2.5.13)$$

Because $V_\varepsilon(\hat{x}(\hat{t}), \hat{t}) = 0$, Inequality (2.5.13) implies there is a $\tilde{t} > \hat{t}$ such that for every $t \in [\hat{t}, \tilde{t}]$,

$$V_\varepsilon(y(t), t) = V_\varepsilon(\hat{x}(\hat{t}) + (t - \hat{t})v, t) \leq -\frac{\varepsilon}{2}(t - \hat{t})$$

We can further require that for this $\tilde{t}$

$$\alpha(t - \hat{t}) \leq \frac{\varepsilon}{2L}(t - \hat{t})$$

when $t \in [\hat{t}, \tilde{t}]$. Substituting these two inequalities to Inequality (2.5.12), we obtain $V_\varepsilon(x(\tilde{t}), \tilde{t}) \leq 0$.

On the other hand,

$$\begin{aligned} V_\varepsilon(x(t), t) &\leq V_\varepsilon(y(t), t) + L\alpha(t - \hat{t}) \\ &= V_\varepsilon(\hat{x}(\hat{t}) + (t - \hat{t})v, t) + L\alpha(t - \hat{t}) \\ &\leq V_\varepsilon(\hat{x}(\hat{t}), \hat{t}) + L\|v\|(t - \hat{t}) + L(t - \hat{t}) + L\alpha(t - \hat{t}) \\ &= L\|v\|(t - \hat{t}) + L\alpha(t - \hat{t}) + L(t - \hat{t}) . \end{aligned}$$

If $t - \hat{t}$ is small enough, we can require $V_\varepsilon(x(t), t) \leq \varepsilon$. The above discussion implies that for the function defined by

$$\tilde{x}(t) = \begin{cases} \hat{x}(t) & t \leq \hat{t}, \\ x(t) & t \geq \hat{t}, \end{cases}$$

$\tilde{x}(t) \in S_{[0,T]}(F, x_0)$, and $\sup \theta(\tilde{x}(t)) > \hat{t}$. It contradicts to $\hat{t} = \sup \Theta$. Hence $\hat{t} = T$.

(ii) There is an $x(t) \in S_{[0,T]}(F, x_0)$ such that Conclusion (2) of theorem is true.

Let $\{\varepsilon_n\}$ be a monotone series such that $\varepsilon_n \downarrow 0$ ($n \rightarrow \infty$). By the conclusion of (i), for every ε_n , there is an $x_n(t) \in S_{[0,T]}(F, x_0)$ such that $\max_{s \in [0,T]} V_{\varepsilon_n}(x_n(s), s) \leq \varepsilon_n$.

By the reason applied in the proof of (i), we can assume $x_n(t) \rightarrow x(t)$ ($n \rightarrow \infty$) and then $x(t) \in S_{[0,T]}(F, x_0)$. By this $x(t)$, $\max_{s \in [0,T]} V_{\varepsilon_n}(x_n(s), s) \leq \varepsilon_n$ leads to $V_0(x(t), t) \leq 0$, i.e.,

$$V(x(t)) - V(x_0) - \int_0^t W(x(s)) ds \leq 0$$

If we replace the initial time by t_0 , Then conclusion (2) is verified.

We now prove the second conclusion implies the first one.

Because F is a Lipschitzian set-valued mapping, it is also ε -semi-upper continuous. Let $x(t) \in S_{[0,T]}(F, x_0)$. Then, for every integer k , we can find a $t_k > 0$ such that for $t \in [0, t_k]$

$$F(x(t)) \subset F(x(0)) + \frac{1}{k}B$$

We can further require $t_k \downarrow 0$ ($k \rightarrow \infty$). Consider now

$$\frac{x(t_k) - x(0)}{t_k} = \frac{1}{t_k} \int_0^{t_k} \dot{x}(s) ds \in F(x(0)) + \frac{1}{k}B$$

$F(x(0))$ is a bounded and closed set, hence $\left\{ \frac{x(t_k) - x(0)}{t_k} \right\}$ holds a convergent subseries.

Without loss of generality, we can assume $\frac{x(t_k) - x(0)}{t_k} \rightarrow v$ ($k \rightarrow \infty$). Then $v \in F(x(0))$. The first conclusion of the lemma illustrates that

$$V(x(t_k)) - V(x_0) = V\left(x_0 + t_k \frac{x(t_k) - x(0)}{t_k}\right) - V(x(0)) \leq \int_0^{t_k} W(x(s)) ds$$

By dividing both sides with t_k , and letting $k \rightarrow \infty$, the definition of Dini derivative leads to

$$D^-V(x_0)(v) \leq W(x_0)$$

The lemma is completely verified. □

We apply Lemma 2.5.5 to prove the following theorem.

Theorem 2.5.5 Suppose the set-valued mapping $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is bounded and Lipschitzian, and is with closed and convex value. If there exist a positive constant η , a positive definite and Lipschitzian function $V : \mathbb{R}^n \rightarrow \mathbb{R}$ and a semi-negatively definite function $W : \mathbb{R}^n \rightarrow \mathbb{R}$ such that for every $\|x\| \leq \eta$ there exists a $v \in F(x)$ such that

$$D^-V(x)(v) \leq W(x)$$

Then Inc. (2.5.4) is weakly stable.

Proof By Lemma 2.5.5, for every $x_0 \in \mathbb{R}^n$ with $\|x_0\| \leq \eta$, there is an $x(t) \in S_{[0,T]}(F, x_0)$ such that

$$V(x(t)) - V(x_0) \leq \int_0^t W(x(s)) ds \leq 0$$

V is positive definite, hence, there are $\alpha_1, \alpha_2 \in \mathbf{K}$ such that $\alpha_2(\|x\|) \leq V(x) \leq \alpha_1(\|x\|)$. It follows that

$$\alpha_2(\|x(t)\|) \leq V(x(t)) \leq V(x_0) \leq \alpha_1(\|x_0\|)$$

For every $\varepsilon > 0$, if $\|x_0\| < \min\{\alpha_1^{-1}\alpha_2(\varepsilon), \eta\}$, then $\alpha_2(\|x(t)\|) < \alpha_2(\varepsilon)$, or $\|x(t)\| < \varepsilon$. $\square$

At last, we give a theorem about weakly asymptotical stability. It can be proved by a similar way to Theorem 2.5.3. But we prefer to provide a new method by using Lemma 2.5.4.

Theorem 2.5.6 Suppose the set-valued mapping $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is bounded and Lipschitzian, and is with closed and convex value. If there exist a positive constant η , a positive definite and Lipschitzian function $V : \mathbb{R}^n \rightarrow \mathbb{R}$ and a negatively definite function $W : \mathbb{R}^n \rightarrow \mathbb{R}$ such that for every $\|x\| \leq \eta$ there exists a $v \in F(x)$ such that

$$D^-V(x)(v) \leq W(x)$$

Then Inc. (2.5.4) is weakly asymptotically stable.

Proof By Theorem 2.5.5, the conditions of Theorem 2.5.6 imply that for every $x_0 \in \mathbb{R}^n$ with $\|x_0\| \leq \eta$, there is a stable solution $x(t) \in S_{[0,T]}(F, x_0)$. We now verify the solution is also asymptotically stable.

By Lemma 2.5.5, we have

$$V(x(t)) \leq V(x_0) + \int_0^t W(x(s)) ds \quad (2.5.14)$$

By Lemma 2.5.3, there are $\alpha_1, \alpha_2, \alpha_3 \in \mathbf{K}$ such that $\alpha_2(\|x\|) \leq V(x) \leq \alpha_1(\|x\|)$ and $W(x) \leq -\alpha_3(\|x\|)$.

Inequality (2.5.14) is equivalent to the following Cauchy problem

$$\frac{d(V(x(t)))}{dt} \leq W(x(t)), \quad V(x(0)) = V(x_0)$$

We also have

$$W(x(t)) \leq -\alpha_3(\|x(t)\|) \leq -\alpha_3(\alpha_1^{-1}(V(x(t))))$$

$\alpha_3\alpha_1^{-1}(\cdot)$ is a function in $\mathbf{K}$ -class, by Lemma 2.5.4, there is $\beta \in \mathbf{KL}$ such that $V(x(t)) \leq \beta(t, V(x_0))$, or $\|x(t)\| \leq \alpha_2^{-1}(\beta(t, V(x_0)))$. The proof is completed. $\square$

Problems

1. $f(x), g(x): \mathbb{R} \rightarrow \mathbb{R}$ are single-valued functions. Prove the following relations

- (1) $D^+(f(x) + g(x)) \leq D^+f(x) + D^+g(x)$
- (2) $D^+(f(x) + g(x)) \geq D^+f(x) + D_+g(x)$

By the above conclusions, establish the corresponding conclusions for $D_+(f(x) + g(x))$, $D^-(f(x) + g(x))$ and $D_-(f(x) + g(x))$.

2. $f(x): \mathbb{R} \rightarrow \mathbb{R}$ is a single-valued function. Prove the following conclusions.

- (1) If $f(x_0) = \max\{f(x), x \in B(x_0, \delta)\}$, then $D^+(f(x_0)) \leq 0 \leq D_-(f(x_0))$,
- (2) If $f(x_0) = \min\{f(x), x \in B(x_0, \delta)\}$, then $D^-(f(x_0)) \leq 0 \leq D_+(f(x_0))$,

3. If $g(x)$ is differentiable, then the following equations hold.

$$D^+(f(x)g(x)) = f(x)\dot{g}(x) + g(x)D^+f(x), \text{ if } g(x) \geq 0;$$

$$D^+(f(x)g(x)) = f(x)\dot{g}(x) + g(x)D_+f(x), \text{ if } g(x) \leq 0.$$

For $D_+(f(x)g(x))$, $D^-(f(x)g(x))$ and $D_-(f(x)g(x))$, present similar conclusions.

4. Let $f(x), g(x)$ be continuous functions, and $h(x) = \max(f(x), g(x))$. Prove that if $D^+f(x) \leq 0$ and $D^+g(x) \leq 0$, then $D^+h(x) \leq 0$.

5. Prove the following conclusion given in Theorem 2.5.1 that

$$D^-f(x(t))(1) = \liminf_{h \downarrow 0} \frac{f(x(t) + hf(x(t))) - f(x(t))}{h}$$

6. If β_1 and β_2 are both $\mathbf{L}$ -class functions, which of the following functions are in $\mathbf{L}$ -class?

- (1) $\beta_1 + \beta_2$,
- (2) $a\beta_1 (a > 0)$,
- (3) $\beta_1(\beta_2(\cdot))$

7. If $\alpha(t)$ is a **K**-class function, and $\beta(t)$ is a **L**-class function, then $\beta(\alpha(t))$ and $\alpha(\beta(t))$ are all **L**-class functions.
8. The following statement is an alternative version of Theorem 2.5.3. Prove the statement.
 “If there are a constant $\eta > 0$, a positive and continuous function $V : \mathbb{R}^n \rightarrow \mathbb{R}$, and a **K**-class function α such that for every $\|x\| \leq \eta$ and $v \in F(x)$ the inequality $D^+V(x)(v) \leq -\alpha(\|x\|)$ holds, then Inc. (2.5.4) is asymptotically stable.”
9. Try to use **K**-class function to alter Theorem 2.5.6 by imitating Problem 8.
10. Prove Theorem 2.5.6.
11. Consider discrete differential inclusion

$$x_{k+1} \in F(x_k), \quad k = 0, 1, 2, \dots$$

where $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a set-valued mapping. Suppose $0 \in F(0)$.

The discrete differential inclusion is said to (strong) stable, if for every $\varepsilon > 0$ there exists a $\delta > 0$ such that if $\|x_0\| < \delta$, $\{x_k, k = 1, 2, \dots\} \subset \varepsilon B$ where B is the unit ball in $\mathbb{R}^n$. Additionally, if $x_k \rightarrow 0$, then it is asymptotically stable.

Give a condition under which the discrete differential inclusion is stable, and prove your condition is sufficient.

12. Consider the discrete differential inclusion given in Problem 11. Prove that if there are a positive and continuous function $V : \mathbb{R}^n \rightarrow \mathbb{R}$ and $\eta > 0, \lambda \in (0, 1)$, when $\|x\| \leq \eta$,

$$\inf_{v \in F(x)} V(v) \leq \lambda V(x)$$

then the discrete differential inclusion is weakly asymptotically stable.

Try to give a condition for weakly stable and prove it.

13. If the set-valued mapping $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is bounded and Lipschitzian, and with closed and convex value. Suppose $V : \mathbb{R}^n \rightarrow \mathbb{R}$ is a Lipschitzian function. If for every $x \in \mathbb{R}^n$, there is a $v \in F(x)$ such that $D^-V(x)(v) \leq W(x)$, then for every $T > 0$, there is a solution $x(t) \in S_{[0,T]}(F, x_0)$ such that $V(x(t_1)) \leq V(x(t_2))$ provided that $T \geq t_1 \geq t_2 \geq 0$.

2.6 Monotonous Differential Inclusions

Monotonicity is a simple variation in the natural world. In mathematics, many investigations start with monotonous phenomenon, such as linear mappings, monotonous sequences, monotonous functions. This section deals with monotonous differential inclusions. The most interesting conclusion is that the solution of Cauchy problem of a maximal monotonous differential inclusion is unique. From this viewpoint the differential inclusion is quite similar to the differential equation. The authors thought

that the most significance is the procedure of investigation. It shows a special project always needs some special techniques. We will start with Minty theorem, a precise proof is presented. Then we consider the Yosida approximation. At last we prove the main conclusion for the maximal monotonous differential inclusion.

2.6.1 Monotonous Set-valued Mappings and Their Properties

A differential inclusion is monotonous if the set-valued mapping in the right side of the inclusion is monotonous. Hence this subsection deals with monotonous set-valued mappings.

$f : \mathbb{R} \rightarrow \mathbb{R}$ is a single-valued mapping, f is said to be monotonous if $x_1 > x_2$ then $f(x_1) \geq f(x_2)$.¹⁸ In the classical control theory, most nonlinear phenomena considered are monotonous, such as step nonlinearity, dead band nonlinearity, saturation nonlinearity. In order to extend the concept to the n -dimensional space or general Hilbert spaces,¹⁹ we rewrite the requirement $x_1 > x_2 \Rightarrow f(x_1) \geq f(x_2)$ by $(x_1 - x_2)(f(x_1) - f(x_2)) \geq 0$, the restriction $x_1 > x_2$ is removed.

A single-valued mapping $f : X \rightarrow X$, where X is a Hilbert space, is monotonous if $\langle (x_1 - x_2), (f(x_1) - f(x_2)) \rangle \geq 0$ for arbitrary two vectors $x_1, x_2 \in X$ where $\langle \cdot, \cdot \rangle$ is the inner product. Sometimes, the inner product is also denoted by $(x_1 - x_2)^T (f(x_1) - f(x_2))$, the superscript “T” means transposition.

Definition 2.6.1 $F : X \rightarrow X$ is a monotonous set-valued mapping, if for arbitrary two vectors $x_1, x_2 \in X$, $y_1 \in F(x_1)$ and $y_2 \in F(x_2)$ then $\langle (x_1 - x_2), (y_1 - y_2) \rangle \geq 0$;

If $\langle (x_1 - x_2), (y_1 - y_2) \rangle > 0$ for arbitrary $x_1 \neq x_2$, $y_1 \in F(x_1)$ and $y_2 \in F(x_2)$, the set-valued F is strictly monotonous;

F is a maximally monotonous set-valued mapping, if there is another set-valued mapping $F_1 : X \rightarrow X$ which holds the property that $F_1(x) \supset F(x)$ for every $x \in X$, then $F_1 = F$. $\square$

Remark 1 In Definition 2.6.1, the space X can be replaced by a set $A \subset X$. we can say a set-valued mapping is monotonous at set A and maximally monotonous at set A . $\square$

Remark 2 If $F : X \rightarrow X$ is a monotonous set-valued mapping, then $F^{-1} : X \rightarrow X$ is also a monotonous set-valued mapping. $\square$

Remark 3 If F, G are two monotonous set-valued mappings, $\lambda, \mu \geq 0$ are two real numbers, then $\lambda F + \mu G$ is also a monotonous set-valued mapping. $\square$

¹⁸The terminology used is somewhat different from the common meaning. In calculus, the monotonicity defined is called monotonous increase.

¹⁹In this section we do not restrict ourselves in a finite dimensional space, $\mathbb{R}^n$ can be replaced by a Hilbert space X .

Lemma 2.6.1 Let $F : X \rightarrow X$ be a set-valued mapping. Then F is monotonous if and only if for every positive real number $\lambda \geq 0$ and any $y_1 \in F(x_1)$, $y_2 \in F(x_2)$

$$\|x_1 - x_2\| \leq \|(x_1 - x_2) + \lambda(y_1 - y_2)\|$$

Proof Because

$$\|(x_1 - x_2) + \lambda(y_1 - y_2)\|^2 = \|x_1 - x_2\|^2 + \lambda^2 \|y_1 - y_2\|^2 + 2\lambda \langle (x_1 - x_2), (y_1 - y_2) \rangle$$

If the condition of Lemma 2.6.1 holds, then

$$\lambda^2 \|y_1 - y_2\|^2 + 2\lambda \langle (x_1 - x_2), (y_1 - y_2) \rangle \geq 0 \quad (2.6.1)$$

We assert Inequality (2.6.1) implies $\langle (x_1 - x_2), (y_1 - y_2) \rangle \geq 0$, i.e., F is monotonous. Otherwise, $\langle (x_1 - x_2), (y_1 - y_2) \rangle < 0$, then there exists a $\lambda > 0$, such that

$$\lambda \|y_1 - y_2\|^2 + \langle (x_1 - x_2), (y_1 - y_2) \rangle < 0$$

It conflicts to Inequality (2.6.1).

If f is monotonous, then $\langle (x_1 - x_2), (y_1 - y_2) \rangle \geq 0$. It implies Inequality (2.6.1), so is

$$\|(x_1 - x_2) + \lambda(y_1 - y_2)\| \geq \|x_1 - x_2\|$$

□

The advantage of Lemma 2.6.1 is that it does not involve inner product. Thus, some property of monotonous mappings may be extended to the Banach space.

The following example gives the difference of monotonous mapping and the maximal monotonous mapping.

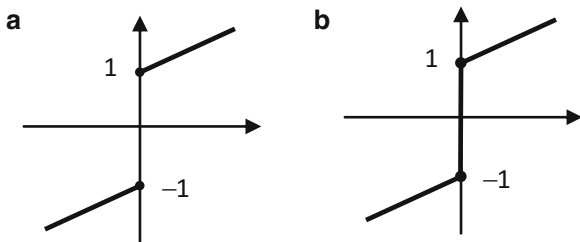
Example 2.6.1 Consider the following two set valued mappings

$$F_1(x) = \begin{cases} 0.5x + 1, & x > 0, \\ \{-1, 1\}, & x = 0, \\ 0.5x - 1, & x < 0; \end{cases} \quad F_2(x) = \begin{cases} 0.5x + 1, & x > 0, \\ [-1, 1], & x = 0, \\ 0.5x - 1, & x < 0. \end{cases}$$

Their graphs are given in Fig. 2.15.

$F_1(x)$ is not maximal since $F_2(0) \supset F_1(0)$ and $F_2(0) \neq F_1(0)$. But $F_2(x)$ is maximal since adding any a point in the $\mathbb{R}^2$ (Fig. 2.15b) will destroy the monotonicity. □

Fig. 2.15 Two monotonous set-valued mappings (a) Graph of $F_1(x)$ (b) Graph of $F_2(x)$



The above example directly leads to the following lemma, we omit its proof.

Lemma 2.6.2 Let $F : X \rightarrow X$ be a monotonous set-valued mapping. It is maximal if the conditions that $\langle u - x, v - y \rangle \geq 0$ and some $(x, y) \in \text{gra } F$ imply $v \in F(u)$ $\square$

The following theorem presents the fundamental properties of maximally monotonous set-valued mappings.

Theorem 2.6.1 If $F : X \rightarrow X$ is a maximally monotonous set-valued mapping, then

- (1) F is with closed and convex value;
- (2) Suppose $x_n \rightarrow x$ ($n \rightarrow \infty$), and $y_n \in F(x_n)$ such that $\langle y_n, z \rangle \rightarrow \langle y, z \rangle$ for every $z \in X$, then $y \in F(x)$.

Proof (1) By Lemma 2.6.2, we have

$$F(x) = \bigcap_{(u,v) \in \text{gra } F} \{y \in X, \langle x - u, y - v \rangle \geq 0\}$$

Let $(u, v) \in \text{gra } F$. Then the set $\{y; \langle x - u, y - v \rangle \geq 0\}$ is a closed and convex for any $x \in X$. Therefore as an intersection, $F(x)$ is closed and convex.

(2) By the conditions of theorem we have $\langle x_n - u, y_n - v \rangle \rightarrow \langle x - u, y - v \rangle$. Because $F : X \rightarrow X$ is a monotonous set-valued mapping and $y_n \in F(x_n)$, for every $(u, v) \in \text{gra } F$ $\langle x_n - u, y_n - v \rangle \geq 0$. It follows $\langle x - u, y - v \rangle \geq 0$. F is maximal, by Lemma 2.6.1, $y \in F(x)$. The theorem is then verified. $\square$

In functional analysis, the fact that $\langle y_n, z \rangle \rightarrow \langle y, z \rangle$ for every $z \in X$ is called that y_n weakly converges to y . The conclusion of (2) can also be stated as “if $x_n \rightarrow x$ strongly, $y_n \rightarrow y$ weakly, then $y \in F(x)$ ”. The conclusion is then called “strong-weak compactness”.

2.6.2 Minty Theorem

This subsection proves an important result of maximally monotonous set-valued mappings. It is called Minty theorem.

Suppose $\{\Omega_\lambda, \lambda \in \Lambda\}$ is a class of sets²⁰ where Λ is set of indexes. $\{\Omega_\lambda, \lambda \in \Lambda\}$ is said to hold the property of finite intersection: If for any finite subset $\{\Omega_{\lambda_1}, \dots, \Omega_{\lambda_n}\}$ of $\{\Omega_\lambda, \lambda \in \Lambda\}$, their intersection $\bigcap_{k=1}^n \Omega_{\lambda_k}$ is nonempty, then the intersection $\bigcap_{\lambda \in \Lambda} \Omega_\lambda \neq \emptyset$. In topology there is a theorem as follows: Suppose $\{\Omega_\lambda, \lambda \in \Lambda\}$ is a class and Ω_λ 's all are closed and contained in a compact set K . Then the class $\{\Omega_\lambda, \lambda \in \Lambda\}$ holds finite intersection property. The property is equivalent to finite covering theorem. The detailed proof will not be given in this book.

Before presenting Minty theorem, we need two lemmas.

Lemma 2.6.3 Suppose K and S are two sets of X , and K is compact. $\phi : K \times S \rightarrow \mathbb{R}$ is a single-valued functional. If for every $y_0 \in S$, $\phi(x, y_0)$ is lower semi-continuous. Suppose M is the class consisting of all finite subsets of S . Then there is a $\bar{x} \in K$ such that

$$\sup_{y \in S} \phi(\bar{x}, y) \leq \nu = \sup_{E \in M} \inf_{x \in K} \max_{y \in E} \phi(x, y)$$

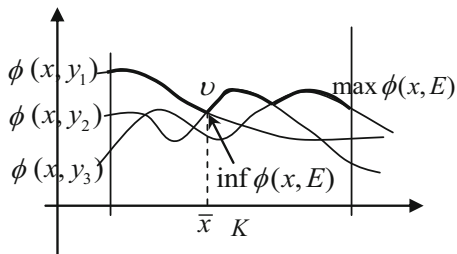
Proof For a fixed $y \in S$, we define a set $\Omega_y = \{x \in K, \phi(x, y) \leq \nu\}$ where ν is real number defined by the theorem. From the definition of ν , we conclude that $\Omega_y \neq \emptyset$. The functional $\phi(x, y)$ is lower semi-continuous, hence, Ω_y is closed. $\Omega_y \subset K$, K is compact, so is the Ω_y .

Denote $\phi(x, E) = \max_{y \in E} \phi(x, y)$. If E is fixed, then $\phi(x, E)$ is a functional of $x \in K$. It can be verified that $\phi(x, E)$ is a lower semi-continuous functional (to see Sect. 1.4). The functional $\phi(x, E)$ possesses such a property that for $x \in K$ and $y_i \in E$, $\phi(x, E) \geq \phi(x, y_i)$. K is compact, there is a $\bar{x}_E \in K$ such that $\phi(\bar{x}_E, E) = \min_{x \in K} \phi(x, E) = \inf_{x \in K} \phi(x, E) \leq \nu$ for every E . Combining the above reasoning, $\phi(\bar{x}_E, y_i) \leq \phi(\bar{x}_E, E) \leq \nu$. Furthermore, $\bar{x}_E \in \bigcap_{y \in E} \Omega_y$. The fact supports that $\bigcap_{y \in E} \Omega_y \neq \emptyset$ for all finite set E , i.e., it holds finite intersection property. K is compact, therefore, $\bigcap_{y \in S} \Omega_y \neq \emptyset$. There is a $\bar{x} \in \bigcap_{y \in S} \Omega_y$. By the definition of Ω_y , this $\bar{x}$ has the property that $\phi(\bar{x}, y) \leq \nu$, $y \in S$. The conclusion is proved. $\square$

The statement of Lemma 2.6.3 is quite complicated, Fig. 2.16 is used to explain the conclusion. In Fig. 2.16, $E = \{y_1, y_2, y_3\}$, three slimsy lines are $\phi(x, y_1)$, $\phi(x, y_2)$ and $\phi(x, y_3)$, the thick line is $\phi(x, E)$. If $S = E$, then the ν and $\bar{x}$ given in the figure meet the requirement of lemma.

²⁰In this book we call the set of sets to be class in order to distinguish the set. For example the power set can call as the class of subsets.

Fig. 2.16 Schematic diagram of $\inf \max \phi(x, y_i)$, $(\bar{x}, v)$ holds the min-max property that in $\phi(x, E)$ it is minimum, in $\phi(\bar{x}, S)$ is maximum



Lemma 2.6.4 Suppose that K is a convex and compact set in X . If $\phi : X \times X \rightarrow \mathbb{R}$ satisfies that:

- (1) For every $y \in K$, the functional $\phi(\cdot, y)$ is lower semi-continuous;
- (2) For every $x \in K$, the functional $\phi(x, \cdot)$ is a concave functional (i.e., $-\phi(x, \cdot)$ is a convex functional);
- (3) For every $y \in K$, $\phi(y, y) \leq 0$.

Then there is a $\bar{x} \in K$ c $\phi(\bar{x}, y) \leq 0$ holds for every $y \in K$. □

Lemma 2.6.4 is known as Fan Ky inequality. To prove this inequality, we should apply the unit decomposition and fixed point theorem. This book does not try to prove the inequality, readers are referred to (Filippov 1988). It would like to point out the inequality holds in Banach space.

We are ready to prove Minty theorem. In the theorem I is used to denote the identical mapping in X . In order to make the proof convenience, we restrict ourselves in $\mathbb{R}^n$. A remark is given behind the proof to point out the difference if we deal with the theorem in a general Hilbert space.

Theorem 2.6.2 $F : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a monotonous set-valued mapping. Then the F is maximal if and only if $I + F$ is a surjective, i.e., the range of $I + F$ is $\mathbb{R}^n$.

Proof The proof for necessity consists of two steps.

- (1) An alternative statement

The conclusion that $I + F$ is a surjective is equivalent to that for every $y \in \mathbb{R}^n$, there is an $\bar{x} \in \mathbb{R}^n$ such that $y \in \bar{x} + F(\bar{x})$. Let $A = -y + F$. Then the conclusion is further equivalent that $0 \in \bar{x} + A(\bar{x})$. By Problem 2 of this section, we conclude the F is maximal is equivalent to that A is maximal. Now, it is sufficient for a maximally monotonous set-valued mapping A to prove that there is an $\bar{x} \in \mathbb{R}^n$ such that $-\bar{x} \in A(\bar{x})$. Using Lemma 2.6.1, it is sufficient to verify $(\bar{x} - u, -\bar{x} - v) \geq 0$, or $(\bar{x} - u, \bar{x} + v) \leq 0$ for every $(u, v) \in \text{gra } A$.

Define a continuous functional $\phi(x, (u, v))$ ²¹ as follows:

$$\phi(x, (u, v)) = \langle x - u, x + v \rangle = \|x\|^2 + \langle x, v - u \rangle - \langle u, v \rangle \quad (2.6.2)$$

²¹In $\phi(x, (u, v))$, v is not a dependent argument, it depends on u by $v \in A(u)$. Hence in $\phi(x, (u, v))$ the dependent variables are u and x , v can be treated as a parameter.

Then, we are only needed to prove there is an $\bar{x} \in X$, such that $\phi(\bar{x}, (u, v)) \leq 0$ for every $(u, v) \in \text{gra } A$.

(2) The existence of $\bar{x}$

We apply Lemma 2.6.4 to verify the existence of $\bar{x}$. Let $u_0 \in \mathbb{R}^n$ and $v_0 \in A(u_0)$. Then define a set $\Omega_{(u_0, v_0)} = \{x; \phi(x, (u_0, v_0)) \leq 0\}$. The set $\Omega_{(u_0, v_0)}$ holds the following properties.

(i) As a quadratic function, $\phi(x, (u_0, v_0))$ has its minimum, hence $\Omega_{(u_0, v_0)}$ is a closed set;

(ii) $\phi(u_0, (u_0, v_0)) = 0$, $u_0 \in \Omega_{(u_0, v_0)}$; and $\phi(x, (u_0, v_0)) \rightarrow +\infty$ ($x \rightarrow \infty$), hence $\Omega_{(u_0, v_0)}$ is a nonempty and bounded set;

(iii) By the discussion in Sect. 1.3, $\phi(x, (u_0, v_0))$ is a convex function. Its epigraph is convex, so is the $\Omega_{(u_0, v_0)}$.

We now construct a compact set K . Let $u_i \in \mathbb{R}^n$, $i = 1, 2, \dots, m$ be m vectors, and we take arbitrarily $v_i \in A(u_i)$. Denote $K = \text{co} \{u_i; i = 1, 2, \dots, m\}$.²² Then K is compact and convex. Fix $x = x_0 \in K$ and prove that $\phi(x_0, (u, v))$ is concave. In order to use Theorem 1.3.3 (2), calculating

$$\nabla_{(u,v)}(-\phi) = \left(-\frac{\partial \phi}{\partial u} \quad -\frac{\partial \phi}{\partial v} \right) = ((x^T + v^T) \quad (-x^T + u^T))$$

Hence

$$\begin{aligned} & \langle (u_2 - u_1), \nabla_u(-\phi)(u_2) - \nabla_u(-\phi)(u_1) \rangle + \langle (v_2 - v_1), \nabla_v(-\phi)(u_2) - \nabla_v(-\phi)(u_1) \rangle \\ &= \langle (u_2 - u_1), (v_2 - v_1) \rangle + \langle (v_2 - v_1), (u_2 - u_1) \rangle \\ &\geq 0. \end{aligned} \tag{2.6.3}$$

The last inequality comes from the fact of monotonicity.

To apply Lemma 2.6.4, we are lastly required to verify $\phi(u, (u, v)) \leq 0$. We now use Lemma 2.6.3 to verify the conclusion. Denote $M = \{(u_j, v_j), j = 1, 2, \dots, m\}$, M is a finite set contained by $\text{gra } A$. By Lemma 2.6.3, we have

$$\begin{aligned} \sup_{(u,v) \in \text{gra } A} \phi(\bar{x}, (u, v)) &\leq \sup_{M \subset \text{gra } A} \inf_{x \in K} \max_{(u,v) \in M} \phi(x, (u, v)) \\ &\leq \sup_{M \subset \text{gra } A} \inf_{x \in K} \max_{\lambda \in \Lambda} \sum_j \lambda_j \phi(x, (u_j, v_j)), \end{aligned} \tag{2.6.4}$$

²²Without loss of generality, we can select $u_i, i = 1, 2, \dots, n$ such that they are linear independent. Therefore, $\text{Int } K \neq \emptyset$.

where $\Lambda = \left\{ \lambda = (\lambda_1 \ \lambda_2 \ \dots \ \lambda_m)^T; \lambda_i \geq 0, \sum_{i=1}^m \lambda_i = 1 \right\} \subset \mathbb{R}^m$. Inequality (2.6.4) is valid since

$$\phi(x, (u_i, v_i)) \in \left\{ \sum_j \lambda_j \phi(x, (u_j, v_j)), \lambda \in \Lambda \right\}$$

In Inequality (2.6.4), $x \in K$, hence $x = \sum_{k=1}^m \mu_k u_k$, with $\mu = (\mu_1 \ \mu_2 \ \dots \ \mu_m)^T \in \Lambda$.

Thus,

$$\sum_j \lambda_j \phi(x, (u_j, v_j)) = \sum_j \lambda_j \phi\left(\sum_k \mu_k u_k, (u_j, v_j)\right)$$

For simplicity, we denote $\phi_M(\mu, \lambda) = \sum_j \lambda_j \phi\left(\sum_k \mu_k u_k, (u_j, v_j)\right)$. Using the notation, $\phi(u, (u, v)) \leq 0$ becomes $\phi_M(\mu, \mu) \leq 0$.

Because

$$\begin{aligned} \phi\left(\sum_k \mu_k u_k, (u_j, v_j)\right) &= \left\langle \sum_k \mu_k u_k - u_j, \sum_k \mu_k u_k + v_j \right\rangle \\ &= \sum_k \mu_k \left\langle u_k - u_j, \sum_k \mu_k u_k + v_j \right\rangle \\ &= \sum_k \mu_k \left(\langle u_k - u_j, v_j \rangle + \left\langle u_k - u_j, \sum_k \mu_k u_k \right\rangle \right), \end{aligned}$$

it follows

$$\begin{aligned} \phi_M(\mu, \lambda) &= \sum_j \lambda_j \phi\left(\sum_k \mu_k u_k, (u_j, v_j)\right) \\ &= \sum_j \lambda_j \sum_k \mu_k \langle u_k - u_j, v_j \rangle + \sum_j \lambda_j \sum_k \mu_k \left\langle u_k - u_j, \sum_k \mu_k u_k \right\rangle. \end{aligned}$$

Let $\lambda = \mu$. Then the second term of the above equation

$$\sum_{j,k=1}^m \mu_j \mu_k \left\langle u_k - u_j, \sum_k \mu_k u_k \right\rangle = 0$$

since both $\mu_j \mu_k \left\langle u_k - u_j, \sum_k \mu_k u_k \right\rangle$ and $\mu_k \mu_j \left\langle u_j - u_k, \sum_k \mu_k u_k \right\rangle$ appear in the summation, and they hold identical absolute values but opposite signs. We now turn to the first term

$$\begin{aligned} \sum_{j,k=1}^m \mu_j \mu_k \langle u_k - u_j, v_j \rangle &= \frac{1}{2} \sum_{j,k=1}^m \mu_j \mu_k \langle u_k - u_j, v_j \rangle + \frac{1}{2} \sum_{j,k=1}^m \mu_j \mu_k \langle u_k - u_j, v_j \rangle \\ &= \frac{1}{2} \sum_{j,k=1}^m \mu_j \mu_k \langle u_k - u_j, v_j \rangle + \frac{1}{2} \sum_{j,k=1}^m \mu_k \mu_j \langle u_j - u_k, v_k \rangle \\ &= \frac{1}{2} \sum_{j,k=1}^m \mu_j \mu_k \langle u_k - u_j, v_j - v_k \rangle \leq 0. \end{aligned}$$

The last inequality comes from the monotonicity. Therefore, $\phi_M(\mu, \mu) \leq 0$. By Lemma 2.6.4, there is a $\bar{\mu} \in \Lambda$ such that for every $\lambda \in \Lambda$ $\phi_M(\bar{\mu}, \lambda) \leq 0$.

Furthermore,

$$\inf_{\mu \in \Lambda} \max_{\lambda \in \Lambda} \sum_j \phi_M(\mu, \lambda) \leq \max_{\lambda \in \Lambda} \sum_j \phi_M(\bar{\mu}, \lambda) \leq 0$$

consequently, $\phi(\bar{x}, (u, v)) \leq 0$ for every $(u, v) \in \text{gra } A$.

Sufficiency. Consider two arbitrary vectors $x, y \in \mathbb{R}^n$. Because $I + F$ is surjective, there is a $x_0 \in \mathbb{R}^n$ such that $y + x \in x_0 + F(x_0)$. Hence, there is a $y_0 \in F(x_0)$, such that $y + x = x_0 + y_0$. Thus, we have

$$\|y - y_0\|^2 = \langle y - y_0, y - y_0 \rangle = -\langle y - y_0, x - x_0 \rangle \quad (2.6.5)$$

If $\langle y - y_0, x - x_0 \rangle \geq 0$, then Eq. (2.6.5) leads to $\|y - y_0\| \leq 0$. We obtain $y = y_0$. By using the equation $y + x = x_0 + y_0$, we have $x = x_0$. Therefore $y \in F(x)$. By Lemma 2.6.2, F is maximal. $\square$

Remark If X is a Hilbert space, the proof of Theorem 2.6.2 is similar except the concavity of $\phi(x_0, (u, v))$. In $\mathbb{R}^n$, we can apply partial derivative to verify the concavity of $\phi(x_0, (u, v))$. However, in a Hilbert space we have no such an operation. This conclusion is proved by using weak topology in Hilbert space. To complete such a proof, we need introduce some concepts which are beyond the target of this book. $\square$

2.6.3 Yosida Approximation

In this subsection, we try to construct a series of single-valued mappings to approximate a maximally monotonous set-valued mapping.

Let A be a maximally monotonous set-valued mapping, and let $\lambda \in \mathbb{R} (> 0)$, then by Minty theorem (Theorem 2.6.2) $I + \lambda A$ is invertible, i.e., for every $x \in X$, $(I + \lambda A)^{-1}(x) \neq \emptyset$.

Definition 2.6.2 Suppose A is a maximally monotonous set-valued mapping, $\lambda \in \mathbb{R} (> 0)$, then $J_\lambda = (I + \lambda A)^{-1}$ is the resolvent of $I + \lambda A$. $A_\lambda = \frac{1}{\lambda} (I - J_\lambda)$ ²³ is the Yosida λ -approximation of A . $\square$

The following theorem presents the fundamental property of J_λ .

Theorem 2.6.3 Suppose $A : X \rightarrow X$ is a maximally monotonous set-valued mapping and $\lambda \in \mathbb{R} (> 0)$, then the resolvent $J_\lambda = (I + \lambda A)^{-1}$ is a single-valued mapping. It is Lipschitzian mapping and its Lipschitzian constant can be 1.

Proof Let $y \in X$. Then there is an $x \in X$ such that $y \in x + \lambda A(x)$ i.e., $x \in J_\lambda(y)$. The effective domain of J_λ is nonempty for every $\lambda \in \mathbb{R} (> 0)$.

Now let $y_1, y_2 \in X$. Then there are $x_1, x_2 \in X$ such that $y_i \in x_i + \lambda A(x_i)$, $i = 1, 2$, i.e., there are $v_i \in A(x_i)$, $i = 1, 2$ such that $y_i = x_i + \lambda v_i$. Hence,

$$y_1 - y_2 = x_1 - x_2 + \lambda (v_1 - v_2)$$

It follows

$$\begin{aligned} \|y_1 - y_2\|^2 &= \|x_1 - x_2\|^2 + \lambda^2 \|v_1 - v_2\|^2 + 2\lambda \langle x_1 - x_2, v_1 - v_2 \rangle \geq \|x_1 - x_2\|^2 \\ &\quad + \lambda^2 \|v_1 - v_2\|^2 \end{aligned} \tag{2.6.6}$$

Thus, $\|x_1 - x_2\| \leq \|y_1 - y_2\|$. If $y_1 = y_2$, then $x_1 = x_2$. It implies that J_λ is a single-valued mapping. It is Lipschitzian and its Lipschitzian constant can be 1. $\square$

In terminology, if Lipschitzian constant is 1, then the mapping is said to be non-expansive, i.e., the distance of two images is less than or equal to that of arguments. If it is less than 1, the mapping is a compressed mapping or contraction.

By the relation between J_λ and A_λ , the following corollary can be verified.

Corollary 2.6.1 A_λ is a Yosida approximation of A . Then A_λ holds the following properties.

- (1) A_λ is a single-valued mapping;
- (2) A_λ is a Lipschitzian mapping and its Lipschitzian constant can be $\frac{1}{\lambda}$
- (3) For every $x \in X$, $A_\lambda(x) \in A(J_\lambda(x))$;
- (4) A_λ is a maximally monotonous mapping.

²³In order to distinguish the inverse of a mapping, we apply $\frac{a}{b}$ to express a fraction.

Proof (1) Because $A_\lambda = \frac{1}{\lambda} (I - J_\lambda)$, J_λ is single-valued mapping, so is A_λ .

(2) Applying the notations used in the proof of Theorem 2.6.3, for every $y \in X$, there are $x = J_\lambda(y)$, $v \in A(x)$ such that $v = \frac{1}{\lambda} (y - x)$. Then

$$A_\lambda(y) = \frac{1}{\lambda} (I - J_\lambda)y = \frac{1}{\lambda} (y - x) = v$$

Now let $v_i = A_\lambda(y_i)$, $i = 1, 2$. By Inequality (2.6.6), we have

$$\|y_1 - y_2\|^2 \geq \lambda^2 \|v_1 - v_2\|^2 = \lambda^2 \|A_\lambda(y_1) - A_\lambda(y_2)\|^2$$

Therefore, the Lipschitzian constant of A_λ can be $\frac{1}{\lambda}$.

(3) Because $(I + \lambda A)(I + \lambda A)^{-1} = (I + \lambda A)J_\lambda$, for every $x \in X$, we have $x \in (I + \lambda A)J_\lambda(x) = J_\lambda(x) + \lambda A(J_\lambda(x))$. By the definition, $\lambda A_\lambda + J_\lambda = I$, i.e., $x = J_\lambda(x) + \lambda A_\lambda(x)$. Comparing these two results, it is clear $A_\lambda(x) \in A(J_\lambda(x))$.

(4) We only prove that A_λ is monotonous, then by Problem 1 of this section, it is maximal. Consider now

$\langle A_\lambda(y_1) - A_\lambda(y_2), y_1 - y_2 \rangle = \langle A_\lambda(y_1) - A_\lambda(y_2), J_\lambda(y_1) - J_\lambda(y_2) \rangle + \lambda \|A_\lambda(y_1) - A_\lambda(y_2)\|^2$ where we have applied $y = \lambda A_\lambda(y) + J_\lambda(y)$. From conclusion (3) above, $A_\lambda(x) \in A(J_\lambda(x))$. A is monotonous, hence the first term is nonnegative. Thus, $\langle A_\lambda(y_1) - A_\lambda(y_2), y_1 - y_2 \rangle \geq 0$ for $y_1, y_2 \in X$. A_λ is monotonous by Definition 2.6.1. $\square$

The further discussion needs the concept $m(A(x))$. Recall that $m(A(x))$ is the minimal norm element in $A(x)$. If $A(x)$ is convex, $m(A(x))$ is the unique projection of the origin on $A(x)$. The following theorem presents the relations between A_λ and $m(A(x))$.

Theorem 2.6.4 Suppose A is a maximally monotonous set-valued mapping. Then for every $x \in X$ the following conclusions hold.

- (1) $\|m(A(x)) - A_\lambda(x)\|^2 \leq \|m(A(x))\|^2 - \|A_\lambda(x)\|^2$
- (2) $\lim_{\lambda \rightarrow 0} J_\lambda(x) = x$
- (3) $\lim_{\lambda \rightarrow 0} A_\lambda(x) = m(A(x))$

Proof (1) Because

$$\begin{aligned} \|m(A(x)) - A_\lambda(x)\|^2 &= \|m(A(x))\|^2 + \|A_\lambda(x)\|^2 - 2 \langle A_\lambda(x), m(A(x)) \rangle \\ &= \|m(A(x))\|^2 - \|A_\lambda(x)\|^2 - 2 \langle A_\lambda(x), m(A(x)) - A_\lambda(x) \rangle, \end{aligned}$$

A is monotonous, and $m(A(x)) \in A(x)$, $A_\lambda(x) \in A(J_\lambda(x))$, hence

$$\langle A_\lambda(x), m(A(x)) - A_\lambda(x) \rangle = \frac{1}{\lambda} \langle x - J_\lambda(x), m(A(x)) - A_\lambda(x) \rangle \geq 0$$

The conclusion follows directly.

(2) By conclusion (1) above

$$\|x - J_\lambda(x)\| = \lambda \|A_\lambda(x)\| \leq \lambda \|m(A(x))\|$$

It leads to

$$\lim_{\lambda \rightarrow 0} J_\lambda(x) = x \quad (2.6.7)$$

(3) The proof of this conclusion seems to be complicated, and is divided into three parts.

(i) For every $x \in X$, the relation $y \in A(x - \lambda y)$ has a unique solution $y = A_\lambda(x)$.

Let $y = A_\lambda(x)$. Then $x - \lambda y = J_\lambda(x)$. It follows $A(x - \lambda y) = A(J_\lambda(x))$. By Corollary 2.6.1 (3), $y = A_\lambda(x) \in A(J_\lambda(x)) = A(x - \lambda y)$, i.e., $A_\lambda(x)$ is a solution of relation $y \in A(x - \lambda y)$.

On the other hand, if y_0 is a solution of $y \in A(x - \lambda y)$. Denote $z = x - \lambda y_0$, then $y_0 \in A(z)$, i.e., $x \in z + \lambda A(z)$. By the definition of J_λ , it follows $z = J_\lambda(x)$. $y_0 \in A(z) = A(J_\lambda(x)) = A_\lambda(x)$. A_λ is a single-valued mapping, hence $y_0 = A_\lambda(x)$.

(ii) If $x \in X$ is fixed, then $\|A_\lambda(x)\|$ is a monotonous function of λ .

By the conclusion (2), if $x \in X$ is given, then $y_0 = A_{\lambda+\mu}(x)$ is the unique solution of relation $y \in A(x - (\lambda + \mu)y)$. Let us rewrite $y_0 \in A(x - (\lambda + \mu)y_0) = A((x - \mu y_0) - \lambda y_0)$, hence y_0 is also the solution of $y \in A_\lambda(x - \mu y)$. By Corollary 2.6.1 (4), $A_\lambda(x)$ is maximally monotonous. If we repeat the process of (i), then we can obtain $y_0 = (A_\lambda)_\mu(x)$ is the unique solution of relation $y \in A_\lambda(x - \mu y)$. Consequently $y_0 = A_{\lambda+\mu}(x) = (A_\lambda)_\mu(x)$ for every $x \in X$.

By conclusion (1) of this theorem, $\|m(A(x)) - A_\lambda(x)\|^2 \leq \|m(A(x))\|^2 - \|A_\lambda(x)\|^2$. Replacing $A(x)$ and $A_\lambda(x)$ by $A_\mu(x)$ and $A_{\lambda+\mu}(x)$, respectively, we obtain

$$\|A_\mu(x) - A_{\mu+\lambda}(x)\|^2 \leq \|A_\mu(x)\|^2 - \|A_{\mu+\lambda}(x)\|^2 \quad (2.6.8)$$

where $m(A_\lambda(x))$ has been replaced by $A_\lambda(x)$. This is correct because $A_\lambda(x)$ is a single-valued mapping. Inequality (2.6.8) implies that the decreasing of λ leads to the monotonously increasing of $\|A_\lambda(x)\|^2$.

(iii) The convergence of $A_\lambda(x)$.

By conclusion (1) of this theorem, for every $x \in X$, $\{\|A_\lambda(x)\|\}$ holds an upper boundary $\|m(A(x))\|$. We now let $\lambda = \frac{1}{k}$, $k \in \mathbb{N}$, by Inequality (2.6.8), $\{A_\lambda(x)\}$ is a Cauchy series. Hence, there exists a $v_x \in X$ such that $v_x = \lim_{\lambda \rightarrow 0} A_\lambda(x)$. The v_x has the following properties:

(i) $A_\lambda(x) \in A(J_\lambda(x))$ and $\lim_{\lambda \rightarrow 0} J_\lambda(x) = x$, $A(x)$ is closed (Theorem 2.6.1), hence $v_x \in A(x)$;

(ii) $\|A_\lambda(x)\| \leq \|m(A(x))\|$ by conclusion (1), consequently $\|v_x\| \leq \|m(A(x))\|$.

The two facts lead to $v_x = m(A(x))$. $\square$

Actually, $A_\lambda(x)$ is a approximate selection of $A(x)$. Let us write

$$\begin{aligned} (x, A_\lambda(x)) &= (J_\lambda(x), A_\lambda(x)) + (x - J_\lambda(x), 0) \\ &\in (J_\lambda(x), A(J_\lambda(x))) + \|x - J_\lambda(x)\| B, \end{aligned}$$

In the proof of Theorem 2.6.4 (2), we have $\|x - J_\lambda(x)\| \leq \lambda \|m(A(x))\|$, hence

$$(x, A_\lambda(x)) \in \text{gra } A + \lambda \|mA(x)\| B$$

2.6.4 Maximally Monotonous Differential Inclusions

To end this section, we prove that the solution of a maximally monotonous differential inclusion is uniquely existed. The proof is completed by using Yosida approximation. We now give a lemma about weak convergence.

Lemma 2.6.5 Suppose X is a Hilbert space, and $x_n \in X$, $n = 1, 2, \dots$. If x_n weakly converges to x_0 , then

- (1) $\|x_0\| \leq \liminf_{n \rightarrow \infty} \|x_n\|$
- (2) Additionally, if $\lim_{n \rightarrow \infty} \|x_n\| = \|x_0\|$, then $\lim_{n \rightarrow \infty} \|x_n - x_0\| = 0$, i.e., x_n strongly converges to x_0 .

Proof (1) $\|x_0\|^2 = \lim_{n \rightarrow \infty} \langle x_n, x_0 \rangle = \lim_{n \rightarrow \infty} \|x_0\| \|x_n\| \cos \theta_n$ where θ_n is the angle yielded by x_0 and x_n . Because $\|x_n\| \cos \theta_n \leq \|x_n\|$, the equation implies $\|x_0\| \leq \liminf_{n \rightarrow \infty} \|x_n\|$.

(2) $\|x_n - x\|^2 = \langle x_n - x_0, x_n - x_0 \rangle = \langle x_n, x_n - x_0 \rangle - \langle x_0, x_n - x_0 \rangle$. Its second term converges to zero by the weak convergence of x_n . Consider the first term

$$\langle x_n, x_n - x_0 \rangle = \langle x_n, x_n \rangle - \langle x_n, x_0 \rangle \rightarrow \|x_n\|^2 - \|x_0\|^2 = 0$$

The conclusion is verified. □

We now consider the following Cauchy problem

$$\dot{x}(t) = -A(x(t)), \quad x(0) = x_0 \quad (2.6.9)$$

where A is a maximally monotonous set-valued mapping with effective domain $\text{dom}(A) \subset X$, $x_0 \in \text{dom}(A)$.

Theorem 2.6.5 The solution of Cauchy problem (2.6.9) exists uniquely. Let $x(t, x_0)$ be the solution. Then it is also the solution of

$$\dot{x}(t) = -m(A(x(t))), \quad x(0) = x_0 \quad (2.6.10)$$

for almost all $t \in [0, \infty)$.

Furthermore, the solution $x(t, x_0)$ holds the following properties

- (1) $\|\dot{x}(t, x_0)\|$ is monotonously decreasing;
- (2) The right derivative $\dot{x}^+(t, x_0) = \lim_{h \downarrow 0} \frac{x(t+h, x_0) - x(t, x_0)}{h}$ is rightly continuous;
- (3) Let $x_i(t, x_{i0})$, $i = 1, 2$ be two solutions with initial conditions x_{10} and x_{20} , respectively. Then

$$\|x_1(t, x_{10}) - x_2(t, x_{20})\| \leq \|x_{10} - x_{20}\| = \|x_1(0) - x_2(0)\| \quad (2.6.11)$$

for every $t \in [0, \infty)$.

Proof The proof of Theorem 2.6.5 consists of 6 parts.

(1) We start to prove the conclusion (3). For $t \in [0, \infty)$, we have

$$\frac{d}{dt} \frac{1}{2} \|x_1(t, x_{10}) - x_2(t, x_{20})\|^2 = \langle \dot{x}_1(t, x_{10}) - \dot{x}_2(t, x_{20}), x_1(t, x_{10}) - x_2(t, x_{20}) \rangle \leq 0$$

The last inequality holds because of monotonicity of A . Furthermore,

$$\begin{aligned} 0 &\geq \int_0^t \langle \dot{x}_1(\tau, x_{10}) - \dot{x}_2(\tau, x_{20}), x_1(\tau, x_{10}) - x_2(\tau, x_{20}) \rangle d\tau \\ &= \frac{1}{2} \left(\|x_1(t, x_{10}) - x_2(t, x_{20})\|^2 - \|x_1(0) - x_2(0)\|^2 \right). \end{aligned}$$

Conclusion (3) is verified. It also implies the uniqueness of the solution.

(2) Denote $y(t) = x(t+h)$ where $x(t)$ is the solution of Cauchy problem (2.6.9), and consider the following Cauchy problem

$$\dot{y}(t) = -A(y(t)), \quad y(0) = x(h).$$

By Inequality (2.6.10), we obtain

$$\left\| \frac{y(t) - x(t)}{h} \right\| \leq \left\| \frac{y(0) - x(0)}{h} \right\|.$$

It is equivalent to

$$\left\| \frac{x(t+h) - x(t)}{h} \right\| \leq \left\| \frac{x(h) - x(0)}{h} \right\|,$$

and then let $h \rightarrow 0$. When their derivatives all exist, we have $\|\dot{x}(t)\| \leq \|\dot{x}(0)\|$. If we replace t and 0 by $t_0 + h$ and t_0 , then the above discussion leads to $\|\dot{x}(t_0 + h)\| \leq \|\dot{x}(t_0)\|$. The conclusion (1) is verified.

(3) We now prove the existence of Inc. (2.6.8). A is maximally monotonous, hence $A(x)$ is with closed and convex value for every $x \in \text{dom } A$ (Theorem 2.6.1). Consider Cauchy problem of the following differential equation

$$\dot{x}_\lambda(t) = -A_\lambda(x_\lambda(t)), \quad x_\lambda(0) = x_0 \quad (2.6.12)$$

We have proved $-A_\lambda$ is a Lipschitzian mapping (Corollary 2.6.1 (2)), the above Cauchy problem exists unique solution. By the first conclusion of this theorem, $\|\dot{x}_\lambda(t)\|$ is monotonously decreasing by the t increasing. Hence,

$$\|\dot{x}_\lambda(t)\| \leq \|\dot{x}_\lambda(0)\| = \|A_\lambda(x_0)\| \leq \|m(A(x_0))\| \quad (2.6.13)$$

the last inequality comes from the proof of Theorem 2.6.4 (3). $\{x_\lambda(t)\}$ is a set of equicontinuous functions for all $\lambda > 0$. Consequently, series $\{x_{\frac{1}{k}}(t); k \in \mathbb{N}\}$ holds a convergent subsequence. Without loss of generality, we assume $\lim_{k \rightarrow \infty} x_{\frac{1}{k}}(t) = x(t)$ for $x(t) \in AC([0, \infty), \mathbb{R}^n)$ uniformly. Because $\{\dot{x}_\lambda(t)\}$ is contained in a bounded set, by Theorem 1.1.7, we can assume $\dot{x}_{\frac{1}{k}}(t) \rightarrow \dot{x}(t) (k \rightarrow \infty)$ weakly.

Define $\mathcal{A} : x_\lambda(t) \rightarrow -A_\lambda(x_\lambda(t))$, $\mathcal{A}$ is an operator from $L^2[0, T]$ to $L^2[0, T]$. We can prove $\mathcal{A}$ is maximally monotonous.²⁴ Because $x_{\frac{1}{k}}(t) \rightarrow x(t)$ strongly and $A_{\frac{1}{k}}(x_{\frac{1}{k}}(t)) = \dot{x}_{\frac{1}{k}}(t) \rightarrow \dot{x}(t)$ weakly, by Theorem 2.6.1 (2), $\dot{x}(t) \in -A(x(t)) = -A(x(t))$. The existence is verified.

(4) $\|m(A(x(t)))\|$ is monotonously decrease as t increases.

At the Part (3) of this proof, we have proved that $\{A_\lambda(x_\lambda(t))\}$ (or $\{A_{\frac{1}{k}}(x_{\frac{1}{k}}(t))\}$) weakly converges to $\dot{x}(t)$ as $\lambda \rightarrow 0 (k \rightarrow \infty)$. For convenience, we denote $v(t) = \dot{x}(t)$, then $v(t) \in A(x(t))$. It follows $\|v(t)\| \geq \|m(A(x(t)))\|$. On other hand, $\{A_\lambda(x_\lambda(t))\}$ is weakly convergent to $v(t)$, hence, by Lemma 2.6.5,

$$\|m(A(x(t)))\| \leq \|v(t)\| \leq \liminf_{\lambda \rightarrow 0} \|A_\lambda(x_\lambda(t))\| \leq \|mA(x(0))\| \quad (2.6.14)$$

the last inequality is obtained from (2.6.12). Consequently, $\|m(A(x(0)))\| \geq \|m(A(x(t)))\|$. The conclusion implies that $\|m(A(x(t)))\|$ is monotonously decrease.

(5) $\|m(A(x(t)))\|$ is continuous from right.

²⁴Note $A: X \rightarrow X$, but the operator $\mathcal{A}: L^2[0, T] \rightarrow L^2[0, T]$. The norm and inner product considered are in space $L^2[0, T]$. It is direct to show the maximal monotonicity of $\mathcal{A}$ implies that of A .

We now prove for any sequence $\{t_n\}$, $t_n \downarrow t_0$, then $m(A(x(t_n))) \rightarrow m(A(x(t_0)))$. By Part (4) of the proof, the series $\{m(A(x(t_n)))\}$ is monotonously increasing since $\{t_n\}$ decreases. Furthermore, by Inequality (2.6.13),

$$\|m(A(x(t_n)))\| \leq \|v(t_n)\| \leq \|m(A(x(t_0)))\|$$

There is a subsequence of $m(A(x(t_n)))$ and the subsequence is weakly convergent to a vector v_0 . We can assume that $m(A(x(t_n)))$ is weakly convergent to v_0 . Repeating the proof of Inequality (2.6.12), inequalities can be obtained as follows

$$\|m(A(x_0))\| \leq \|v_0\| \leq \lim_{n \rightarrow \infty} \inf \|m(A(x(t_n)))\| \leq \|m(A(x(t_0)))\| = \|m(A(x_0))\|$$

By Lemma 2.6.5, $m(A(x(t_n)))$ converges strongly to $m(A(x_0))$.

(6) At last, the solution of Inc. (2.6.8) is identical to the solution of Eq. (2.6.9)

For almost all $t \in [0, T]$ and $h > 0$, we have

$$\|x(t+h) - x(t)\| = \left\| \int_t^{t+h} \dot{x}(\tau) d\tau \right\| \leq \int_t^{t+h} \|\dot{x}(\tau)\| d\tau \leq h \|m(A(x(t)))\|,$$

the last inequality is obtained from the fact the $\|m(A(x(t)))\|$ decreases monotonously. $\|\dot{x}(t)\|$ is continuous from right, when $h \downarrow 0$, $\left\| \frac{x(t+h) - x(t)}{h} \right\| \rightarrow \|\dot{x}(t)\|$. The above inequality implies $\|\dot{x}(t)\| \leq \|m(A(x(t)))\|$. Because $\dot{x}(t) \in -A(x(t))$ and $A(x)$ is closed, $\dot{x}(t) = -m(A(x(t)))$.

Thus, we complete proof for all conclusions of the theorem. $\square$

Note that in the theorems given above, the norms $\|x(t)\|$, $\|\dot{x}(t)\|$ and others are functions of time t . The $x(t)$ is an trajectory in space X .

The solution of Inc. (2.6.8) which satisfies equation $\dot{x}(t) = -m(A(x(t)))$ is called slow solution or mild solution.

Problems

1. $f: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a monotonously single-valued mapping, if f is continuous then it is maximal.
2. $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is maximally monotonous, prove the following conclusions:
 - (1) $y_0 \in \mathbb{R}^n$ is a given vector, then $y_0 + F$ is maximally monotonous.
 - (2) $\alpha > 0$ is a positive real number, then αF is maximally monotonous.
3. $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a monotonous mapping. Prove that F is monotonous if and only if the inverse $F^{-1}: \mathbb{R}^n \rightarrow \mathbb{R}^n$, which may be a set-valued mapping, is monotonous.
4. If $F: \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a convex mapping then its subdifferential $\partial F(x)$ is monotonous.

5. Assume, in Theorem 2.6.2, $X = \mathbb{R}^n$, prove that the set $\{x; \phi(x, (u, v)) \leq 0\}$ defined in the proof of Theorem 2.6.2 is compact and convex.
6. Let $x_\lambda(t)$ be the solution of Eq. (2.6.11). Then $J_\lambda(x_\lambda(t))$ converges to $x(t)$ uniformly in the proof of Theorem 2.6.5.
7. If $V : \mathbb{R}^n \rightarrow \mathbb{R}(\infty)$ is a lower semi-continuous and convex function, then $\partial V(x)$ is maximally monotonous. Moreover, the solution of Cauchy problem $\dot{x}(t) = -\partial V(x), x(0) = x_0$ satisfies the equation $\frac{d}{dt} V(x(t)) + \|\dot{x}(t)\|^2 = 0$.

References

- Aubin J-P, Cellina A (1984) Differential inclusions – set-valued maps and viability theory [M]. Springer, Berlin
- Cai X, Liu L, Zhang W (2009) Saturated control design for linear differential inclusions subject to disturbance. Nonlinear Dyn 58(3):487–496
- Filippov AF (1988) Differential equations with discontinuous right-hand sides [M]. Kluwer Academic Publishers, Dordrecht
- Sontag ED (1989) Smooth stabilization implies coprime factorization [J]. IEEE Trans Autom Control 34(4):435–442

Theory of Control Systems Described by Differential
Inclusions

Han, Z.; Cai, X.; Huang, J.

2016, XI, 344 p. 50 illus., 5 illus. in color., Hardcover

ISBN: 978-3-662-49243-7