INDUSTRIAL MATHEMATICS,
The Journal of the Industrial Mathematics Society, Volume 46, Part 2, 1996

Some Remarks on Solutions of Discontinuous Differential Equations Applied in Automatic Control

Piotr Kulczycki
Faculty of Electrical Engineering
Cracow University of Technology
PL-31-155 Cracow, ul. Warszawska 24, Poland

ABSTRACT

In many tasks of contemporary engineering, the classical notion of the solution of a differential equation turns out to be insufficient. The present paper is devoted to generalized solutions of ordinary differential equations, utilized to analyze equations with a discontinuous right-hand side. In particular, the most frequently used types have been discussed: Caratheodory's, Filippov's, and Krasovski's. Based on examples from the area of automatic control, there are pointed out a number of aspects connected with the applications of the aforementioned notions in modern technical problems.

1. INTRODUCTION

The steadily increasing demands placed on modern production systems are enforcing the use of differential equations with a discontinuous right-hand side to describe the dynamics of devices being designed. This results, for example, from the application of ever more precise models of motion resistances, or from the alternating generation of the extreme values of admissible controls—the so-called "bang-bang control" (see e.g. [5]). The classical solution of a differential equation (i.e. a differentiable function fulfilling the equation at every point of its domain [4]), whose existence is guaranteed by the continuity of the right-hand side of the differential equation, may no longer exist even in the simplest forms of the discontinuities that appear. A well known example is the equation

$$\dot{x}_1(t) = 1, \tag{1.1}$$

This paper was prepared during the author's stay at the Department of Control Engineering, Aalborg University, Denmark.

$$\dot{x}_2(t) = \begin{cases} -1 & \text{for } x_2(t) \geq 0 \\ 1 & \text{for } x_2(t) < 0, \end{cases} \tag{1.2}$$

whose solution exists only until it reaches the axis x_1. In such cases, it becomes necessary to make use of the notion of generalized solutions. The conceptions most often applied for these purposes, those of Caratheodory, Filippov, and Krasovski, considered together with the practical aspects in tasks of optimal control, will be the subject of the investigations presented in this paper.

2. DEFINITIONS

Definition 2.1. Let T be an interval with nonempty interior. Consider a differential equation

$$\dot{x}(t) = g(x(t), t), \tag{2.1}$$

where $x : T \to \mathbb{R}^n$, and a mapping $g : \mathbb{R}^n \times T \to \mathbb{R}^n$ is discontinuous.

The function x, absolutely continuous on every compact subinterval of the set T, is a solution of differential equation (2.1):

- in the Caratheodory sense (C-solution), if it fulfills equation (2.1) almost everywhere in T,
- in the Filippov sense (F-solution), if

$$\dot{x}(t) \in F[g](x(t), t) \text{ almost everywhere in } T, \tag{2.2}$$

- in the Krasovski sense (K-solution), if

$$\dot{x}(t) \in K[g](x(t), t) \text{ almost everywhere in } T, \tag{2.3}$$

where the Filippov and Krasovski operators are defined, respectively, by

$$F[g](x(t), t) = \bigcap_{e>0} \bigcap_{Z \subset \mathbb{R}^n : m(Z)=0} \operatorname{conv}\left[g((x(t) + eB) \setminus Z, t)\right], \tag{2.4}$$

$$K[g](x(t), t) = \bigcap_{e>0} \operatorname{conv}\left[g(x(t) + eB, t)\right], \tag{2.5}$$

B denotes the open unit ball in the space $\mathbb{R}^n$, m is the n-dimensional Lebesgue measure, and conv[C] means the convex closed hull of the set C.

For practical purposes, the C-solution precisely states the conception of "joining" classical solutions [4], and is most frequently used when the function g is discontinuous only with respect to the variable t. Just as in the case of the classical solution, the derivative of the C-solution is dependent (besides the variable t) only on the present value of the solution. By contrast, in accordance with the necessity in contemporary engineering to take into account the influence of inaccuracies and approximations, the derivative of the K-solution depends on all the points in the neighborhood of the present value of the solution. Continuing the concept of taking errors into account, the F-solution also omits zero-measure sets, which are unimportant from a practical point of view.

Proposition 2.2.
(A) The C-solution is a K-solution (since $g(x(t),t) \in K[g](x(t),t)$).
(B) The F-solution is a K-solution (because $F[g](x(t),t) \subset K[g](x(t),t)$).

Example 2.3. Consider the following differential equation with an initial condition

$$\dot{x}_1(t) = \begin{cases} 1 & \text{for } x_2(t) > 0 \\ -1 & \text{for } x_2(t) = 0, \\ 1 & \text{for } x_2(t) < 0 \end{cases} \quad x_1(0) = x_{01}, \tag{2.6}$$

$$\dot{x}_2(t) = \begin{cases} -1 & \text{for } x_2(t) > 0 \\ 0 & \text{for } x_2(t) = 0, \\ 1 & \text{for } x_2(t) < 0 \end{cases} \quad x_2(0) = x_{02}, \tag{2.7}$$

where $x_{01}, x_{02} \in \mathbb{R}$ (Fig. 1).

First, the case $x_{02} = 0$ will be considered. The function $[x_{01} - t,\ 0]^T$ is then a C-solution, while the mapping $[x_{01} + t,\ 0]^T$ constitutes an F-solution (for values of the Filippov operator, see Fig. 1). In turn, every absolutely continuous function of the form $[x_1(t),\ 0]^T$, such that $x_1(0) = x_{01}$ and the condition $|\dot{x}_1(t)| \leq 1$ is fulfilled at the points of existence of the derivative, represents a K-solution (see also Fig. 1).

In the case $x_{02} \neq 0$, however, the C-, F-, and K-solutions are given, until

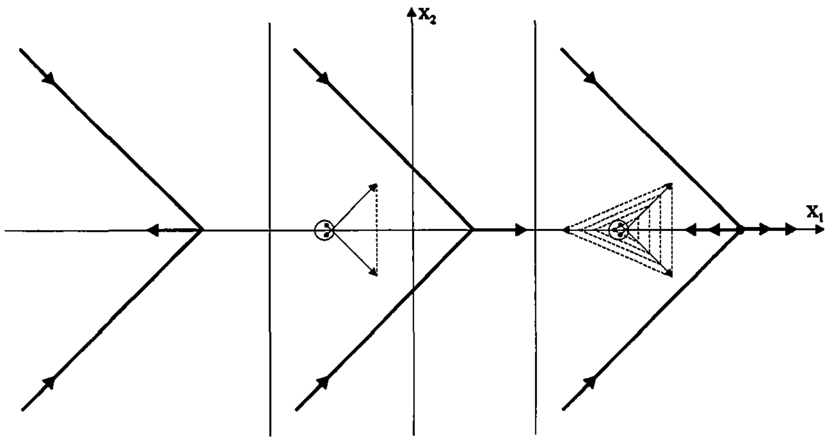

Figure 1. C-, F-, and K-solutions, as well as Filippov and Krasovski operators for equation (2.6)–(2.7).

the x_1 axis is crossed, as $[x_{01} + t,\ x_{02} - t \cdot \mathrm{sgn}(x_{02})]^T$. After this axis is reached, these solutions may be extended, in accordance with the rules presented in the preceding paragraph.

The above example indicates that there is a lack of relation between C- and F-solutions, whereas K-solutions often constitute an excessively broad class.

A detailed discussion of these issues is found in [1,2].

3. EXAMPLES OF APPLICATIONS

In this section, examples from the area of optimal control will be shown to illustrate various aspects of the application of the conceptions presented in Section 2.

Example 3.1. (See also [3].) Suppose a dynamical system, submitted to the two-dimensional control $u \equiv [u_1, u_2]^T$, and described by a differential equation with an initial condition

$$\dot{x}_1(t) = x_2(t) + u_1(t), \quad x_1(0) = x_{01}, \tag{3.1}$$

$$\dot{x}_2(t) = -x_1(t) + u_2(t), \quad x_2(0) = x_{02}, \tag{3.2}$$

where $x_{01}^2 + x_{02}^2 \leq 2$. Consider the optimal control problem consisting in reaching the target set, which is the origin of coordinates, and minimizing the value of the functional

$$J(u) = \int_0^{t_k} \left[(x_1(t)-1)^2 + x_2(t)^2 - 1\right]^2 \left[u_1(t)^2 + u_2(t)^2\right] dt, \tag{3.3}$$

where t_k means the time to reach the origin. (Note that the equation $(x_1(t) - 1)^2 + x_2(t)^2 - 1 = 0$ represents a unit circle with its center at the point $[1, 0]^T$, which will be marked hereafter with R.)

Thus, if the control u is given by

$$u(t) = \begin{bmatrix} 0 \\ 0 \end{bmatrix}, \tag{3.4}$$

then the C-, F-, and K-solutions of the differential equation created in this manner, i. e. (3.1)–(3.2) and (3.4), take on the form shown in Fig. 2A. However, for the case

$$u(t) = u_r(x(t)) = \begin{bmatrix} -2x_2(t) \\ 2x_1(t) - 1 \end{bmatrix}, \tag{3.5}$$

where $x \equiv [x_1, x_2]^T$, the solutions of equation (3.1)–(3.2), (3.5) are illustrated in Fig. 2B.

Now, let a control be the following combination of the preceding ones:

$$u^*(t) = u_r^*(x(t)) = \begin{cases} \begin{bmatrix} 0 \\ 0 \end{bmatrix} & \text{if } x(t) \notin R, \\ \begin{bmatrix} -2x_2(t) \\ 2x_1(t) - 1 \end{bmatrix} & \text{if } x(t) \in R. \end{cases} \tag{3.6}$$

For $x(t) \notin R$, formula (3.6) is equivalent to (3.4), therefore the solutions are brought to the circle R (see Fig. 2A), and afterwards, being permanently contained in that circle, reach the origin (compare dependence (3.6) with (3.5), and see Fig. 2B). In the first case, $u_1(t)^2 + u_2(t)^2 = 0$; in the second, $(x_1(t)-1)^2 + x_2(t)^2 - 1 = 0$. Therefore, $J(u^*) = 0$. Since the values of the functional J are nonnegative, the control u^* defined by formula (3.6) is optimal.

However, in order to confirm the formal correctness of such a result, a

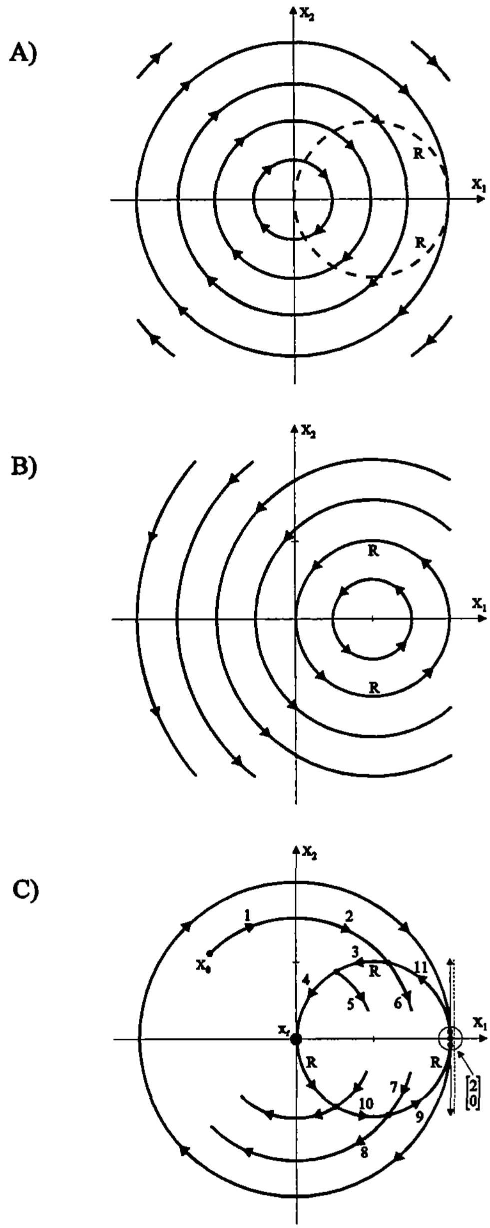

Figure 2. C-, F-, and K-solutions of equation (3.1)–(3.2) if the function u is given respectively by formulas A- (3.4), B- (3.5), C- (3.6).

detailed analysis of the solutions of the obtained equation (3.1)–(3.2), (3.6) should be performed. Namely, the C-solution can indeed be brought to the circle R, and then along it to the target set (parts 1-2-3-4 in Fig. 2C), but, since for a single t the differential equation need not be satisfied, that solution may "leave" the circle R (parts 3-5). In the same way, it is possible that the circle R will be crossed many times (parts 2–6 and 7–8); then the C-solution may also begin to "move" along this circle (parts 2–3 and 7–9), and afterwards "leave" it again (parts 3–5, 10–8 or 11–6). The above phenomena can be repeated in any sequence.

The form of the Filippov operator causes the zero-measure sets (of which the circle R is one) not to influence the F-solutions. Therefore formula (3.6) is here equivalent to (3.4), and thus the F-solutions take on the form shown in Fig. 2A. The origin is not reached by the F-solutions at all.

Finally, in accordance with Propositions 2.2, the K-solutions contain the C-solutions described above. Nevertheless, the set of the values of the Krasovski operator at the point $[2, 0]^T$ includes zero as well (see Fig. 2C), and therefore the K-solution may additionally "stop" at this point. Afterwards it can also "leave."

The optimal control problem considered above, seemingly completed by the pointing out of formula (3.6), turns out to be obviously unacceptable, once the task of the solutions of the differential equation has been carefully examined.

Example 3.2. Suppose a dynamical system, submitted to the one-dimensional control u with values bounded to the interval $[-1, 1]$, and described by a differential equation with an initial condition

$$\dot{x}_1(t) = x_2(t), \qquad x_1(0) = x_{01}, \tag{3.7}$$

$$\dot{x}_2(t) = u(t) - w \cdot \operatorname{sgn}(x_2(t)), \qquad x_2(0) = x_{02}, \tag{3.8}$$

where $x_{01}, x_{02} \in \mathbb{R}$ and $w \in (0,1)$. Consider the time-optimal control problem consisting of reaching the fixed target set $x_f = [x_{f1}, x_{f2}]^T \in \mathbb{R}^2$ in a minimal and finite time.

Assume that x_+ and x_- are C-solutions of differential equation (3.7)–(3.8), with the conditions $x_+(0) = x_-(0) = x_f$, defined on the interval $(-\infty, 0]$, and generated by the control $u \equiv +1$ or $u \equiv -1$, respectively. Let

$$Q_+ = \{x_+(t) \text{ for } t < 0\}, \tag{3.9}$$

$$Q_- = \{x_-(t) \text{ for } t < 0\} \tag{3.10}$$

(Fig. 3); therefore, these are the sets of all points which can be brought to the origin by the control $u \equiv +1$ or $u \equiv -1$, respectively. Let also

$$R_+ = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ such that there exists } \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} \in Q, \text{ with } x_1 < x_1^* \text{ and } x_2 = x_2^* \right\}, \quad (3.11)$$

$$R_- = \left\{ \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \text{ such that there exists } \begin{bmatrix} x_1^* \\ x_2^* \end{bmatrix} \in Q, \text{ with } x_1 > x_1^* \text{ and } x_2 = x_2^* \right\}, \quad (3.12)$$

where $Q = Q_- \cup \{x_f\} \cup Q_+$. The discussions often encountered in the literature regarding the optimality of the control expressed by the formula

$$u(t) = u_r(x(t)) = \begin{cases} -1 & \text{if } x(t) \in (R_- \cup Q_-) \\ +1 & \text{if } x(t) \in (R_+ \cup Q_+), \end{cases} \quad (3.13)$$

where $x \equiv [x_1, x_2]^T$, have been based on the conviction that the solutions of differential equation (3.7)–(3.8), (3.13) are brought to the curve Q, and being thereafter permanently contained in it, reach the target set in a finite time. Nevertheless, the detailed analysis of the solutions of that equation, conducted in the case $x_{f2} \neq 0$, calls into question the correctness of such a procedure.

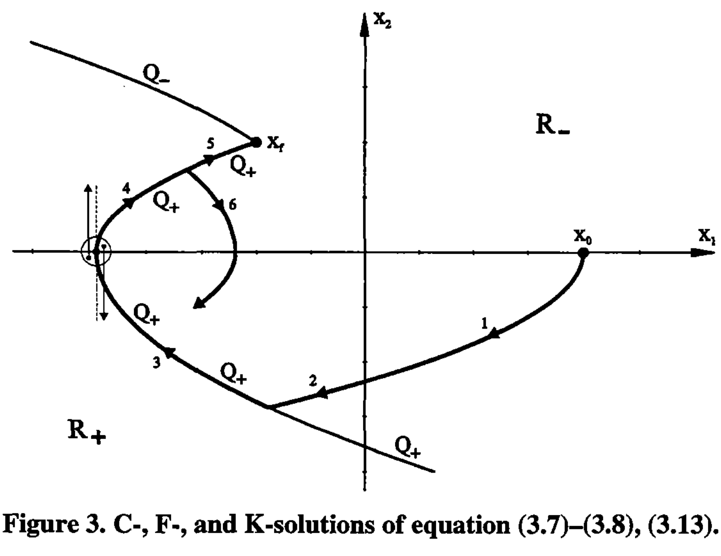

Figure 3. C-, F-, and K-solutions of equation (3.7)–(3.8), (3.13).

The C-solution is indeed brought to the curve Q, and, along it, the solution can reach the target set (parts 1-2-3-4-5 in Fig. 3). However, since, for a single t, the differential equation may not be satisfied, that solution can also "leave" the curve Q between the axis x_1 and the target set, penetrating the region R_- (parts 4–6 in Fig. 3). The cycle that appears in this manner may repeat itself any number of times.

The C-solutions described above are also F- and K-solutions. However, at the point where the curve Q crosses the axis x_1, the values of the Filippov and Krasovski operators include zero as well (see Fig. 3), and therefore the F- and K-solutions may "stop" at that point. Afterwards, they may also "leave."

The nonuniquenss of the solutions demonstrated above causes the time to reach the target to be a set rather than a number. Moreover, there exist solutions which do not reach the target at all in a finite time. Without additional explicitness, the acknowledgement of control (3.13) as time-optimal turns out to be therefore groundless.

The task considered in Example 3.2 is one of the basic problems in the practice of optimal control, especially in the area of industrial robots and manipulators [6]. Differential equation (3.7)–(3.8), equivalent to the form

$$\ddot{y}(t) = u(t) - w \cdot sgn(\dot{y}(t)), \quad y(0) = x_{01}, \quad \dot{y}(0) = x_{02}, \tag{3.14}$$

expresses the second law of Newtonian mechanics, when the term $-w \cdot \mathrm{sgn}(\dot{y}(t))$ represents an exemplary, simple but discontinuous model of motion resistances. (Note that the form of control (3.13) introduces a second factor determining the discontinuity of the right-hand side.) The absence of a correct analysis of the solutions of the differential equation is the lot of many advanced publications, even though these solutions are in fact the basis for drawing conclusions about the evolution of the state of the system under control. The prevailing tendency towards mathematical precision in technical problems will doubtless lead in the direction of making the issues presented above generally familiar in processes of designing engineering devices.

4. REFERENCES

1. A.F. Filippov, "Differential Equations with Discontinuous Right-Hand Side," *Matematicheskiy Sbornik*, 51 (1960), 1, 99–128, in Russian; English translation: *Transactions of the American Mathematical Society, Ser. 2*, 42 (1964), 199–231.
2. O. Hajek, "Discontinuous Differential Equations, Parts I–II," *Journal of Differential Equations*, 32 (1979), 149–185.

3. H. Hermes, "Discontinuous Vector Fields and Feedback Control," in *Differential Equations and Dynamical Systems*, ed. J.K. Hale and J.P. LaSalle, Academic Press, New York, 1967, 155–165.
4. J.H. Hubbard and B.H. West, *Differential Equations, A Dynamical Systems Approach, Part I: Ordinary Differential Equations*, Springer-Verlag, New York, 1991.
5. P. Kulczycki, "Almost Certain Time-Optimal Positional Control," *IMA Journal of Mathematical Control & Information*, 13 (1996), 63–77.
6. P. Kulczycki, "Time-Optimal Positional Feedback Controller for Random Systems," *Proceedings of the National Science Council, ROC(A),* 20 (1996), 79–89.