

Dipartimento di / Department of

Pure and Applied Mathematics

Dottorato di Ricerca in / PhD program Mathematics Ciclo / Cycle 29

Curriculum in (se presente / if it is)

TITOLO TESI / THESIS TITLE

Metrically Regular Generalized Equations:
A Case Study in Electronic Circuits

Cognome / Surname Mehrabi nezhad Nome / Name Iman

Matricola / Registration number 787920

Tutore / Tutor: Prof.ssa Rita Pini

Cotutore / Co-tutor: Prof. Amos Uderzo
(se presente / if there is one)

Supervisor: Prof.ssa Rita Pini, Prof. Amos Uderzo
(se presente / if there is one)

Coordinatore / Coordinator: Prof. Roberto Paoletti

ANNO ACCADEMICO / ACADEMIC YEAR 2015-2016

UNIVERSITÀ DEGLI STUDI DI MILANO - BICOCCA



DOCTORAL THESIS

**Metrically Regular Generalized Equations:
A Case Study in Electronic Circuits**

Author:
Iman MEHRABI NEZHAD

Supervisor:
Prof.ssa Rita PINI
Prof. Amos UDERZO

*A thesis submitted in fulfilment of the requirements
for the degree of Doctor of Philosophy*

in the

Department of Mathematics and Applications

UNIVERSITÀ DEGLI STUDI DI MILANO - BICOCCA

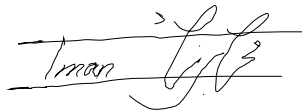
March 2017

Declaration of Authorship

I, Iman MEHRABI NEZHAD, declare that this thesis titled, “Metrically Regular Generalized Equations: A Case Study in Electronic Circuits” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed: _____

Handwritten signature of Iman Mehrabi Nezhad in black ink, written over a horizontal line. The signature is cursive and includes the name 'Iman' followed by a stylized set of initials.

Date: 10/March/2017

Università degli studi di Milano - Bicocca

Abstract

Department of Mathematics and Applications

Doctor of Philosophy

Metrically Regular Generalized Equations: A Case Study in Electronic Circuits

by Iman MEHRABI NEZHAD

In this thesis, set-valued maps are considered to model the $i-v$ characteristics of semiconductors like diode, and transistor. Using the circuit theory laws, a generalized equation is obtained. The main concern of the thesis is to investigate how perturbing the input signal will affect the output variables. The problem is studied in two cases: the static case, where the input signal is a DC source; and the dynamic case, where there exists an AC source in the circuit.

In the static case, the problem can be reduced to the existence or absence of local stability properties of the solution map, like the Aubin property, isolated calmness, and calmness, or metric regularity for the inverse map. Some tools from variational analysis are used to provide necessary and/or sufficient conditions that guarantee these properties.

In the dynamic case, those pointwise results are used to obtain descriptions for regularity properties of the solution trajectories in function spaces.

Acknowledgements

Imagine each human as a sphere in \mathbb{R}^n (with $n \geq 4$), whose mass is concentrated in a human body shape at the center of this sphere and his/her effect radiates centrifugally. Thus, whatever a human being does, would affect all the other humans, more or less. Of course, such an effect may decrease when you get farther from the center, but it never reaches zero.

You may define an “effective radius”, inside which, people are physically close enough to see each other, to start a conversation, to communicate and share experiences and emotions, ..., to live with each other. Such an effective radius, if it exists, is different from one person to another. Also there could be some disagreements in defining this radius. You may see someone, everyday for a while, but never have the courage to say something to him/her. On the other hand, it is possible to read a note and feel a great and in depth sympathy and closeness to the person who has written it.

Such a model, if one accepts, can explain many things in human relations and particularly, the difficulty I may encounter to write an acknowledgement. There has been many people in my life, who had helped me to be where I am. Consciously or unconsciously, their manners and deeds have affected me, even though I may not know them in person. To be honest, I should thank all of humans.

After this general introduction, I may name a few people who had stronger influence on me, at least in the last three years, which would not mean being rude or ungrateful to those I may not name here.

Above all, I want to express my sincere appreciation to my advisers, Prof.ssa Rita Pini, and Prof. Amos Uderzo. English was not the mother tongue of me nor theirs, but their kindness allows me to feel powerful enough to talk, their patience provides me a safety to ask my questions, their expertise permits me to run faster in my path, and their openness to new experiences enables me to bridge a gap between my bachelor and master studies.

I learned a lot from their mathematical knowledge and manner.

I would like to thank Prof. Radek Cibulka for his kindness and trust to give me drafts of his unpublished works; his patience to answer my endless questions and doubts; his hospitality to let me visit him in Pilsen; and his ideas to enrich my work. The spark of most of the new results in this thesis has been produced during my stay in Pilsen. Tomas Roubal also helped me to have a more pleasant and fruitful visiting period, thanks to his valuable friendship, which I hope lasts for a life-time.

During my stay in Italy, I had very good moments with my friends that I may not forget. Thanks to Simone, Davide, Chiara, Elena, Jessica, Daniela, Federico, Morteza, Alberto, Mani, Masoud, Maryam, Reza, ... and all the others.

Last but not least, I shall thank my family: my parents and my sister for their continuous love and support; my brothers who empowered and encouraged me for higher education.

Contents

Declaration of Authorship	i
Abstract	ii
Acknowledgements	iii
Contents	v
List of Figures	vii
Introduction	1
1 Mathematical Preliminaries	9
1.1 Multifunctions: First Properties	9
1.2 Local Stability Properties	19
1.3 Regularity Terms and Their Relation with Local Stability Properties	24
1.4 A Review on Perturbation Results	33
2 An Introduction to the Electrical Problem	43
2.1 Introduction to the Theory of Electrical Circuits	43
2.2 Formulating the Problem	56
2.2.1 Alternative Formulation	66
2.3 Examples	66
3 Tools of Variational Analysis for Studying the Local Stability of the Solution Mapping in the Static Case	74
3.1 Variational Geometry	75
3.1.1 Calculus Rules	85
3.1.2 Inner and Outer Norms	89
3.1.3 Subdifferentials	98
3.2 Results about Aubin Property	100
3.3 Results about Isolated Calmness	111
3.3.1 Isolated Calmness Without Injectivity Assumption	119
3.4 Results about Calmness	123
3.5 Results in terms of Metric Regularity	129
3.5.1 Strong Metric (Sub-) Regularity with a Nonsmooth Single-valued Part	135

4 Metric Regularity in the Case of Time-Varying Sources	142
4.1 Obtaining a Proper Model	143
4.1.1 Review on an Existence Theorem	147
4.2 A Study about the Solution Trajectories	154
4.2.1 Continuity of Solution Trajectories	156
4.2.2 Uniform Strong Metric Regularity	159
4.2.3 Perturbations of the Input Signal	167
References	174
Notation	180
Index	182

List of Figures

1.1	The role of neighborhood V in graphical localization	16
1.2	A map which is open but not linearly open at $\bar{x} = (0, 0)$ for $\bar{y} = 0$	28
2.1	A simple electrical circuit	44
2.2	General black box viewpoint of components	45
2.3	Circuit components: Resistor	45
2.4	Circuit components: Capacitor	46
2.5	Circuit components: Inductor	46
2.6	Circuit components: Voltage Source	47
2.7	Circuit components: Current Source	47
2.8	Circuit components: Diode	48
2.9	Diode operation: (a) Current flow is permitted; the diode is forward biased. (b) Current flow is prohibited; the diode is reversed biased.	48
2.10	A circuit with Ideal Diode, its $i - v$ characteristic, and equivalent circuits in the reverse and forward directions.	49
2.11	Circuit components: Zener Diode	49
2.12	Circuit components: LED	50
2.13	Circuit components: Shockley diode. from left to right: schematic, physical diagram, equivalent schematic, circuit symbol, and $i - v$ characteristic	50
2.14	Circuit components: DIAC. from left to right: equivalent circuit, schematic symbol, and $i - v$ characteristic	51
2.15	Circuit components: SCR. from left to right: physical diagram, equivalent schematic, and circuit symbol	52
2.16	Circuit components: BJT transistor: (a) PNP schematic symbol, (b) physical layout (c) NPN symbol, (d) layout.	52
2.17	An <i>npn</i> BJT biased in active mode	53
2.18	The Ebers-Moll (EM) model of the <i>npn</i> transistor (left), and the <i>pnp</i> transistor (right)	53
2.19	A simple circuit that divides voltage between a bias resistor and a load	55
2.20	A simple electrical circuit, the schematic and $i - v$ characteristic of Zener Diode	56
2.21	Approximation of the Diode's $i - v$ characteristic	57
2.22	A circuit with Tunnel diode (left) and the graphical analysis of it (right)	60
2.23	Cascading	62
2.24	Simplifying the study of the circuit by splitting the layers	62
2.25	A simplified $i - v$ characteristic of a DIAC and the equivalent maps	63
2.26	A circuit with two loops	63
2.27	A circuit with AC voltage source	65
2.28	A simple AC to DC Rectifier circuit	65

2.29	Driving an LED with AC	67
2.30	A simple circuit with DIAC	68
2.31	The graph of Φ for different values of parameters R , a and V (left: $a = 200, R = 5, V = 10$; right: $a = 10, R = 20, V = 10$)	69
2.32	A circuit with SCR and Zener Diode	70
2.33	Loops and variables definition in sampling gate circuit	72
2.34	Sampling gate circuit	72
3.1	Calculating contingent and paratingent cones for a set	76
3.2	Calculating contingent and paratingent cones for a convex set	77
3.3	Regular and limiting normal cones to Ω at $\bar{x} = (0, 0)$	79
3.4	Calculating the Aubin property in a circuit with DIAC	105
3.5	Calculating the Aubin property in a circuit with SCR and Zener Diode.	107
3.6	A part of the set $M_2(N((\bar{z}_2, \bar{v}_2); \text{gph } F_2))$	110
3.7	Calculating isolated calmness in a circuit without the (A1) assumption	122
3.8	Graph of F with normal and tangent cones at $(0, -1)$	127
3.9	A simple circuit with SCR	140
3.10	General form of the map J_A in Example 3.5.14	141
4.1	A regulator circuit with multiple DC sources	143
4.2	Solution mapping for the circuit in Figure 4.1 (left), graphical method to find the output for a typical input (right)	145
4.3	Difficulty in defining isolated trajectories	158
4.4	The side effects of changing radii independently	160

Introduction

The definition of a good mathematical problem is the mathematics it generates rather than the problem itself.

Andrew J. Wiles

It would be useful to compare learning mathematics with practising arts like painting and drawing. Imagine you want to draw a landscape. As the first steps in mastering as a painter are to watch carefully, and try to make a simplified sketch of what the painter is seeing, then going slowly toward the details to create more accurate copy of the landscape (in other words, closer to what is “real”) by improving his/her knowledge of the painting instruments, mastering hand skills, and an endless process of trial and error to become better and better, so should be the approach to use mathematics in studying the nature. Modelling a natural phenomenon is a way to study it in a more abstract way than doing experiments on it (if possible), and a goal of modelling is to “predict” the behaviour of the phenomenon under study with respect to changes of different variables in the model. The more appropriate a model would be, a *better* prediction of the phenomenon it would provide, and by “better” one can think of different desired properties: a wider range of change for the input variables, an easier way to find the equilibrium point, a more detailed view of small changes, or ...

Our study starts with modelling electrical circuits. As it would be explained with more details in Chapter 2, every component in the circuit could be understood as a relation between the current passing through it and the drop of potential difference (voltage, for short) over it. For many components such a relation is a function, but some electronic components like diodes, and transistors, need a set-valued map to describe their $i - v$ characteristic. Understanding the behaviour of a circuit depends on knowing the current of each branch and voltage of each component in the circuit, and to do so, we need to use the $i - v$ characteristics of components, together with Kirchhoff’s current and voltage laws to form a system of n equations, n variables, and then, solve it. When there exists

a set-valued map in the model, we will obtain a *generalized equation* like:

$$0 \in f(x) + F(x), \quad (1)$$

where $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, and $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ could be determined specifically for each circuit. Generalized equations are interesting structures in their own. First, observe that $F(\cdot) \equiv \{0\}$, reduces (1) to an ordinary equation $0 = f(x)$, hence the name. Second, consider a closed and convex set Ω in \mathbb{R}^n , and let F be the normal cone (in the sense of convex analysis) to Ω at a point $x \in \mathbb{R}^n$, that is,

$$F(x) = N(x; \Omega) = \begin{cases} \{v \mid \langle v, c - x \rangle \leq 0 \text{ for all } c \in \Omega\} & x \in \Omega, \\ \emptyset & x \notin \Omega. \end{cases} \quad (2)$$

Geometrically, this is the cone of all outward normals to Ω at x . Note that, if (1) holds with this particular F , then the sum on the right hand side is non-empty (it contains 0), so $N(x; \Omega) \neq \emptyset$, which means $x \in \Omega$. Also, $-f(x)$ must belong to $N(x; \Omega)$, so for each $c \in \Omega$,

$$\langle -f(x), c - x \rangle \leq 0.$$

Thus, we can see that (1) holds if and only if x satisfies the so called *variational inequality*:

$$x \in \Omega, \text{ and } \langle f(x), c - x \rangle \geq 0 \text{ for each } c \in \Omega, \quad (3)$$

and this, geometrically, means that $f(x)$ is an inward normal to Ω at x . Robinson in [59–62] has studied this particular type of generalized equations in details and found the setting of generalized equations as an appropriate way to express and analyse problems in complementarity, mathematical programming, and variational inequalities.

It is worthwhile mentioning that, although one could write (1) as $0 \in G(x)$, by defining a new set-valued map $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ as $G(x) = f(x) + F(x)$, keeping the single-valued and set-valued parts separated is more useful in practice, as it turns out that f is often a fairly smooth function, while F involves “corners”.

In the study of electrical circuits, power supplies (that is, both current and voltage sources) play an important role. Not only their failure in providing the minimum voltage level for other components to work would be a problem, but also small changes in the provided voltage level will affect the whole circuit and the goal it has been designed for. This small changes around a desired value could happen mainly because of failure in precise measurements, ageing process, and thermal effects (which are explained in Chapter 2, see also [16, 68]).

Thus, based on the type of voltage/current sources in the circuits, we consider two different cases:

1. static case:

This is the situation when the signal sources in the circuit are DC (that is, its value is not changing with respect to time). For practical reasons, we would prefer to rewrite (1) as $p \in f(x) + F(x)$, where $p \in \mathbb{R}^n$ is a fixed vector representing the voltage or current sources in the circuit. Then, we define the corresponding *solution mapping* as follows:

$$p \mapsto S(p) := \{x \in \mathbb{R}^n \mid p \in f(x) + F(x)\}. \quad (4)$$

In this framework, the desire to have small deviations of x with respect to perturbations of p around a presumed point $(\bar{p}, \bar{x}) \in \text{gph } S$ could be investigated as the local stability properties (like the Aubin property, isolated calmness, and calmness) of S at \bar{p} for \bar{x} . Or equivalently, we can ask for metric regularity features of $\Phi = f + F$ at \bar{x} for \bar{p} . The details of this analogy is given in Chapter 2, while the definition of these local properties and their equivalences (one for S , the other for $\Phi = S^{-1}$) is provided in Chapter 1. We would provide handy theorems to check these properties in Chapter 3.

Let us mention that by the term *metrically regular generalized equation*, we refer to a generalized equation 1 where the right hand side is a metrically regular map. In general, the stability properties under parameter perturbations is a very important topic in engineering, not only for the determination of the behaviour of a system with respect to perturbations but also for the construction of algorithms for the numerical simulation of the problem.

2. dynamic case:

When an AC signal source (that is, its value is a function of time) is in the circuit, problem could be more complicated. First of all, all the other variables of the model would become a function of time, too. Second, it is not appropriate any more to formulate the solution mapping as $S(p)$. One can consider a parametric generalized equation like:

$$0 \in h(t, x) + F(x), \quad (5)$$

where $h : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ now depends on a scalar parameter $t \in [0, 1]^1$, and define the solution mapping corresponding to (5) as

$$S : t \mapsto S(t) = \{x \in \mathbb{R}^n \mid h(t, x) + F(x) \ni 0\}, \quad (6)$$

where $h(t, x) = -p(t) + f(x)$ corresponding to the previous notation. The third

¹ In fact, t can belong to any finite interval like $[0, T]$ for a $T > 0$. The starting point $t = 0$ is considered as the moment that the circuit starts working, in other words, when the circuit is connected to the signal sources and is turned on with a key. We keep the time interval as $[0, 1]$ in the entire thesis for simplicity.

difficulty rises here; the study of the effects of perturbations of p is not equivalent any more to searching the local stability properties of S . The *strongly regular point* criteria of Robinson [60] comes into play now, which guarantees a good behaviour of the problem solutions (cf. Theorem 4.1.1 and the explanations therein).

One can encounter this framework with an overlook to the previous case, and consider this problem as an iteration process, that is, for any $t \in [0, 1]$ we have a static case problem. This approach is well known and well studied in the literature, both as a pointwise study (see for example [10, 13, 47, 49, 52, 69, 70]), or as a numerical method and for designing algorithms (see for example [5, 6, 28, 37, 38] and references therein).

However, we have an intuition (coming from experimental observations and simulation results) that the variable x could be seen as a function (of time) in this case, and there is some relation between the signal source function (we call it input signal) and this function (we call it output signal). Thus, instead of looking at the sets $S(t)$, we focus on *solution trajectories*, functions like $x : [0, 1] \rightarrow \mathbb{R}^n$ such that

$$x(t) \in S(t), \quad \text{for all } t \in [0, 1], \quad (7)$$

that is, $x(\cdot)$ is a selection for S over $[0, 1]$; and search for their regularity properties in function spaces. This study would be the main concern of Chapter 4 which is mostly our contribution to the subject.

There, we provide a smoothness relation between the input signal and the output signal, and study the perturbation effect of the input signal on solution trajectories.

All the statements, definitions, and examples from books or papers has been cited, though some explanations and changes has been added to adapt them with our setting. Most of the proofs have been rewritten, partly or entirely, to make it more easier to follow². Figures, especially in Chapter 2, are obtained from internet, and we did not consider them something crucial to cite. However, when they contain an original idea, they have been explained in the context and truly cited. *Wikipedia* is always a good source to find appropriate photos for electronic components.

The thesis is organized in four chapters, and two appendices. We will provide a brief review of their contents here.

In Chapter 1, the main mathematical preliminaries of the thesis are presented. Most of the definitions and notations that would be used throughout this thesis are introduced in this chapter. After mentioning basic concepts of set-valued maps in Section 1.1, we begin to present some local properties for a general set-valued map in the next two sections.

² This is a risky note, as it uncovers the level of my mathematical knowledge by showing what was assumed not clear or difficult to understand. I am eager to take the risk and also cheerful to learn more.

We gather these properties under the general name of *local stability* properties in Section 1.2 and *regularity* in Section 1.3.

We tried to familiarize the reader with these properties by providing several examples to distinguish the delicate differences in the definitions and by describing the relation between these notions (see, for example, Theorems 1.3.5, 1.3.7, and 1.3.14).

Section 1.4 will be an omen of what would be the main question of this thesis: we have a set-valued map F with a particular stability-like³ property. We consider a perturbing function f affecting F in a summation form (the reason for using this model will be discussed in Chapter 2). Then, we wonder under which conditions the new set-valued map $G = f + F$, would still have the stability-like property?

We can also model a closely related situation: we start from $f + F$, and then the perturbed model would be of the form $g + F$. Having the set-valued map fixed, one might think of a possible relation between g and f . Then, we get the idea to approximate f in the model which may not be smooth enough, with a smooth function. This is the reason we introduce *approximations* and *estimators* in that section.

Chapter 2 starts by a short review on the theory of electrical circuits, and a brief introduction to some electronic components, which is a wide and deep topic. It would be hard to talk about so many things in few pages, and it might cause more confusion than clarification; but since the aim of this chapter is only to provide the framework of how to model properly a real world situation into a generalized equations setting, we tried to avoid any further details.

Some engineering terms may not be explained completely, importance and necessity of studying some circuits are not highlighted well enough, and not all interesting examples are included in the text, but this is the price one should pay eagerly to write a thesis in pure mathematics about a concrete problem.

Although we have been inspired by the works already done in this direction (see, for example, [3, 4, 7, 66]), we found some ambiguity and misinterpretations in the modelling of some circuits in the literature, and thus, tried to write a self-reliant chapter about it. Most of the materials in this chapter are new or, at least, reorganized in a new form to serve the purpose of this thesis⁴.

³ We use this phrase in our general descriptions in order to refer to any of the local stability properties of Section 1.2 (that is, Aubin property, calmness, linear openness, and isolated calmness) or any of the regularity terms of Section 1.3 (that is, metric regularity, strong metric regularity, metric sub-regularity, and strong metric sub-regularity).

The reason behind this choice is that, first of all, these properties provide a local and not general description of the map under study. Secondly, the term “stability” has already been used in analysis, dynamical systems, and many other areas of mathematics with certain definitions, which is different from what we want to study here.

⁴ This sentence should not be understood in the way that we have discovered the Kirchhoff’s laws, or invented the diodes, or things like this which could be found in electronic books, but simply means that we noticed the $i - v$ relation of diodes and other semiconductors as set-valued maps and used circuit theory to form a proper generalized equation, not an approximated equation, nor a variational

Section 2.2 has several important aims. Firstly, it provides the general form of generalized equations that would be considered in the rest of the thesis as an outcome of modelling process. Secondly, it provides physical explanation for the importance of studying the small perturbations of source signals. And in third place, it gives a meaning to the stability-like properties introduced in Chapter 1: in the static case, the question about sensitivity of the circuit to small perturbations of the power source, is translated to the question whether the solution mapping of the obtained generalized equation has some stability-like properties or not. For the dynamic case, which is the situation where sources are changing with time known as AC voltages or currents, we only present the framework of the generalized equation. How they could be related to stability-like concepts would be the subject of Chapter 4.

This chapter ends with a review on an alternative formulation method in Subsection 2.2.1, and examples of circuits we would like to discuss stability of the solution mappings related to them. There, we tried to explain how this formulation could open a door to variational inequality study, and why this model could not be used for all circuits (see for instance, the circuit in Example 2.3.1).

Chapter 3 starts with introducing some tools from variational analysis. Starting from a geometrical interpretation of derivatives, the concept of graphical derivative (with three different approaches) for set-valued maps is introduced in Section 3.1, and some calculus rules for these derivatives is provided in subsection 3.1.1. Although there are many results in this area (see for example [14, 32, 51, 53, 55, 72]), we only present the most needed ones and adapt them to our setting.

Then we provide two norm-like maps that assign an extended real number (i.e. $\mathbb{R} \cup \{-\infty, +\infty\}$) to a *positively homogeneous* set-valued map (cf. Definitions 3.1.28, and 3.1.30) in Subsection 3.1.2. These definitions allow us to characterize the stability-like properties of a set-valued map in terms of the *inner* and *outer norm* of their graphical derivatives (see for example Proposition 3.1.35, and Theorems 3.1.37, 3.1.38, and 3.1.42). Subsection 3.1.3 will introduce *subdifferentials* and *generalized Jacobians* for functions that may not be differentiable. It might be better to introduce subdifferentials before graphical derivatives. Firstly, because they extend the idea of differentiability to a non-smooth function in a graphical way. Secondly, a subdifferential is usually a set-valued map. So, one might be more ready for defining the concept of graphical derivatives for set-valued maps afterwards, but we preferred to keep them where they are as we only use them in Section 3.4 and Subsection 3.5.1 to show a possible way to continue the stability study when dealing with a generalized equation with a non-smooth single-valued part.

The remaining sections of this chapter have a common structure. Each section will discuss inequality. It is the idea of viewing the old problem from a new stand point.

one of the stability-like properties for the specific generalized equation $f(\cdot) + BF(C\cdot) \ni p$, with the following assumptions:

- (A1) B is injective;
- (A2) f is continuously differentiable in \mathbb{R}^n ;
- (A3) F has closed graph;
- (A4) C is surjective; and
- (A5) there are $F_i : \mathbb{R} \rightrightarrows \mathbb{R}$, $i \in \{1, \dots, m\}$ such that $F(x) = \prod_{i=1}^m F_i(x_i)$ whenever $x = (x_1, \dots, x_m)^T \in \mathbb{R}^m$.

We try to use the norm characterization obtained for an arbitrary set-valued mapping in Subsection 3.1.2 for each property, and calculus rules of Subsection 3.1.1 to go step by step toward a pointwise easy-to-check criteria for the local stability of the solution mapping.

The main theorem of each section uses only assumptions (A1) - (A3) (see, for example, Theorem 3.2.4, and 3.3.3). Results using (A4) and/or (A5) are expressed as corollaries (see, for instance, Corollary 3.2.8, and 3.3.4). Although in many circuits these assumptions may hold, there are specific situations where these conditions may not be satisfied. Among the first three assumptions, (A3) is not such a strong requirement and holds for the $i - v$ characteristics of semiconductors like diodes in our study. In Subsection 3.3.1 we try to consider the case where B is not injective, and use the following condition instead of (A1) to provide some similar statements for isolated calmness in this case (cf. Theorem 3.3.10).

($\widetilde{A1}$) Suppose that there is $\bar{v} \in F(C\bar{z})$ such that

$$\bar{p} = f(\bar{z}) + B\bar{v} \quad \text{and} \quad \bigcup_{t>0} \overline{\text{rge } \frac{F_C - \bar{v}}{t}} \cap \ker B = \{0_{\mathbb{R}^m}\}.$$

The process is almost the same as before. Example 3.3.12 shows such a situation in a circuit.

In Subsection 3.5.1 we focus on functions f which are not smooth enough to satisfy (A2). A possible approach would be considering the generalized Jacobians and replace (A2) with

($\widetilde{A2}$) f is locally Lipschitz continuous on \mathbb{R}^n .

In Section 3.5, using the relations between the regularity terms and local stability properties of the inverse map expressed in Section 1.3, we restate the results of previous sections in terms of metric regularities in Theorems 3.5.6, and 3.5.13. Not only gathering all separated statements together would ease the future referring, but also expressing results in terms of metric regularity would be more useful in view of the method we choose for

our study in Chapter 4.

In Chapter 4 we try to answer the question of perturbation effect when the vector p becomes a function of time, which corresponds to the electronic circuits working with AC sources. In Section 4.1 we obtain a proper model for the electrical circuit in this case, and explain the shift of our goal from studying the solution mapping to solution trajectories. In Subsection 4.1.1, we express an important theorem (cf. Theorem 4.1.2) to build our structure. Theorem 4.1.2 and its backbone theorem (Theorem 1.4.14), were first stated in [25], and then appear in [29] with a small modification. Being assured that there does exist a solution trajectory, we search for the smoothness and other properties of the trajectories.

Section 4.2 starts with a proposition that has a simple proof, but contains a novel idea which considerably eases the study of the problem in the time varying case (cf. Proposition 4.2.1). In Subsection 4.2.1, we provide results that highlight smoothness dependence of trajectories on the input signal (cf. Proposition 4.2.4, and Corollary 4.2.6).

In Subsection 4.2.2, we deviate from the study about solution trajectories for a while to present a uniform strong metric regularity result along a trajectory, with assumptions weaker than Theorem 4.1.2. The reason we present Theorems 4.2.10, and 4.2.12 in this subsection is that they are necessary for obtaining the results about perturbation effect in Subsection 4.2.3.

The final result of this chapter would be Theorem 4.2.15 that guarantees the existence of a solution trajectory for the perturbed problem which is continuous, and whose distance from the solution trajectory of the (non-perturbed) generalized equation is controlled by the distance between the input signal and its perturbed function.

Chapter 1

Mathematical Preliminaries

Life is the only art that we are required to practice without preparation, and without being allowed the preliminary trials, the failures and botches, that are essential for training.

Lewis Mumford

The main mathematical preliminaries of this thesis are presented in this chapter. We start with mentioning basic concepts of set-valued maps in Section 1.1, and then present some local properties for a general set-valued map under the general name of *local stability* properties in Section 1.2.

In Section 1.3 we will introduce the regularity terms, and will clarify how these concepts are related to local stability properties. The last section of this chapter, Section 1.4, will be devoted to the study of perturbing a set-valued map F , having a particular stability-like property, with a function f . We wonder under which conditions the new set-valued map $G = f + F$, would still have the stability-like property.

1.1 Multifunctions: First Properties

In this section, we review the most needed definitions, terminology, and notations, that we will use throughout the thesis.

In order to avoid confusion, after this section we would refer to multifunctions as set-valued maps and use “function” only for single-valued maps. The different notation will also help us to separate the maps more easily.

Definition 1.1.1. (Set-valued Maps) [67, p. 63] Let X and Y be vector spaces. A mapping $F : X \rightarrow 2^Y$, which associates to $x \in X$ a (possibly empty) subset $F(x)$ of Y ,

is called a *multifunction*¹ or *set-valued mapping*, and is denoted by $F : X \rightrightarrows Y$.

The *graph*, *domain*, and *range* of F are defined, respectively, by

$$\begin{aligned} \text{gph } F &:= \{(x, y) \in X \times Y \mid x \in X, y \in F(x)\}, \\ \text{dom } F &:= \{x \in X \mid F(x) \neq \emptyset\}, \\ \text{rge } F &:= \{y \in Y \mid y \in F(x) \text{ for some } x \in X\}. \end{aligned}$$

If A is a subset of X , we write $F(A) := \bigcup_{x \in A} F(x)$.

Note 1.1.2. Although the above and many other definitions in this thesis could be expressed in general vector spaces (see, for example, [54]), we keep our attention only on \mathbb{R}^n for some positive natural number n .

Note 1.1.3. The multifunction F is said to be *closed* or *convex* if $\text{gph } F$ is closed or convex, respectively. We call F *closed-valued* or *bounded-valued* if $F(x)$ is, respectively, a closed, or a bounded subset of Y for any $x \in X$.

Notice that a closed multifunction is closed-valued but the converse is not true, for example consider the map $F : \mathbb{R} \rightrightarrows \mathbb{R}$ with $F(0) = [-1, 0]$ and $F(x) = [0, 1]$ for $x \neq 0$.

Remark 1.1.4. When F assigns more than one element to x we say it is *multi-valued* at x , and when it assigns no element at all, it is *empty-valued* at x . When it assigns exactly one element y to x , it is *single-valued* at x , in which case we allow ourselves to write $F(x) = y$ instead of $F(x) = \{y\}$.

Thus, a mapping $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ can be identified with the (single-valued) multifunction $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ defined by $F(x) = \{f(x)\}$, $x \in \mathbb{R}^n$. Hence the name *multifunction* is explained. Moreover, concepts defined below for set-valued mappings will be also applicable to a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ according to this identification.

Though, at the beginning the concept of set-valued maps may seem to be superficial and an unnecessary complication; the following example shows that it is deeply hidden in the heart of analysis, as in many cases after formulating the problem in terms of a relation (or function) between the input and output data, we are interested in the properties of the inverse map.

Example 1.1.5. (Inverse Functions) Consider a (single valued) function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ as a relation that describes the behaviour of system, a relation between inputs $x \in \mathbb{R}^n$ and outputs $y \in \mathbb{R}^m$.

¹ **Historical note.** The usage of terms *multifunction*, and *multivalued function* as a sort of allowing “function” in mathematics to mean also “multivalued function”, had some ups and downs in history and finally dropped out of usage at some point in the first half of the twentieth century.

It seems that the theory of multivalued functions (with this name) was fairly systematically developed for the first time in 1959 in the book “*Topological Spaces: including a treatment of multi-valued functions, vector spaces, and convexity*” by the French mathematician Claude Berge (cf. [12]).

In many applications we are interested in finding the accurate input which gives us a desired output. Mathematically speaking, we are interested in the *solution mapping*

$$S(y) := \{x \in \mathbb{R}^n \mid y = f(x)\}.$$

This map is generally a set-valued map and if we have the further assumption on f to be injective, it would be single valued and known as the inverse function $S = f^{-1}$.

In fact, the inverse of a set-valued map $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ at a point $\bar{y} \in \mathbb{R}^m$ is defined as

$$F^{-1}(\bar{y}) := \{x \in \mathbb{R}^n \mid \bar{y} \in F(x)\}.$$

In this manner, a function f always has an inverse f^{-1} as a set-valued mapping. The question of an *inverse function* comes down then, to passing to some “piece” of the graph of f^{-1} . Soon, we will define selection and localization for a set-valued map (ref. Definitions 1.1.14 and 1.1.19) to describe this “piecing” process.

Terminology and Notation [29, p. 7]

- (a) In working with \mathbb{R}^n we will denote by $\|x\|$ the Euclidean norm associated with the canonical inner product

$$\langle x, x' \rangle = \sum_{j=1}^n x_j x'_j \quad \text{for } x = (x_1, \dots, x_n) \quad \text{and} \quad x' = (x'_1, \dots, x'_n),$$

namely

$$\|x\| = \sqrt{\langle x, x \rangle} = \left[\sum_{j=1}^n x_j^2 \right]^{1/2}.$$

- (b) The closed ball around \bar{x} with radius r is $\mathbb{B}_r(\bar{x}) = \{x \mid \|x - \bar{x}\| \leq r\}$. We denote the closed unit ball $\mathbb{B}_1(0)$ by \mathbb{B} .
- (c) A *neighborhood* of \bar{x} is any set U such that $\mathbb{B}_r(\bar{x}) \subset U$ for some $r > 0$. (Often the neighborhoods can conveniently be taken to be closed balls themselves.)
- (d) The *interior* and *closure* of a set $C \subset \mathbb{R}^n$ will be denoted by $\text{int } C$ and $\text{cl } C$. Thus, $\text{int } \mathbb{B}_r(x)$ will indicate the open ball around \bar{x} with radius r .
- (e) The *distance* from a point $x \in \mathbb{R}^d$ to a set B is denoted by $d(x, B)$ and defined as

$$d(x, B) := \inf_{y \in B} d(x, y),$$

with the convention that $d(x, \emptyset) = +\infty$.

The *Hausdorff distance* of two non-empty sets A and B is then defined as

$$d_H(A, B) := \max\{e(A, B), e(B, A)\}$$

where $e(A, B)$ is the *excess* of A beyond B defined as

$$e(A, B) := \sup_{x \in A} d(x, B) = \sup_{x \in A} \inf_{y \in B} d(x, y), \quad (1.1)$$

with the following convention that

$$e(\emptyset, B) = \begin{cases} 0 & \text{when } B \neq \emptyset, \\ \infty & \text{otherwise.} \end{cases}$$

(f) Given two set-valued mappings $F : X \rightrightarrows Y$ and $G : Y \rightrightarrows Z$, the *restriction* $G|_F : Y \rightrightarrows Z$ of G on F is defined by

$$G|_F(y) := \begin{cases} G(y) & \text{if } y \in F(x), \\ \emptyset & \text{otherwise.} \end{cases} \quad (1.2)$$

We shortly review three basic properties of functions, mostly for the sake of notation and ease of future reference. The set-valued counterparts of calmness and Lipschitz continuity would be introduced in next section, while the derivative definitions for set-valued maps are postponed to Chapter 3.

Definition 1.1.6. (Calmness) [29, p. 25] A function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be *calm* at \bar{x} relative to a set D in \mathbb{R}^n if $\bar{x} \in D \cap \text{dom } f$ and there exists a constant $\kappa \geq 0$ such that

$$\|f(x) - f(\bar{x})\| \leq \kappa \|x - \bar{x}\| \quad \text{for all } x \in D \cap \text{dom } f. \quad (1.3)$$

The calmness property (1.3) can alternatively be expressed in the form of the inclusion

$$f(x) \in f(\bar{x}) + \kappa \|x - \bar{x}\| \mathbb{B} \quad \text{for all } x \in D \cap \text{dom } f.$$

That expression connects with the generalization of the definition of calmness to set-valued mappings, which we will discuss at length in next section.

The *calmness modulus* of f at \bar{x} , denoted by $\text{clm}(f; \bar{x})$, is the infimum of the set of values $\kappa \geq 0$ for which there exists a neighborhood D of \bar{x} such that (1.3) holds.

Note 1.1.7. According to this definition, if \bar{x} is an isolated point, we have $\text{clm}(f; \bar{x}) = 0$. As long as \bar{x} is not an isolated point of $\text{dom } f$, the calmness modulus satisfies

$$\text{clm}(f; \bar{x}) = \limsup_{\substack{x \in D \cap \text{dom } f, \\ x \rightarrow \bar{x} \\ x \neq \bar{x}}} \frac{\|f(x) - f(\bar{x})\|}{\|x - \bar{x}\|}.$$

When f is not calm at \bar{x} , from the definition we get $\text{clm}(f; \bar{x}) = \infty$. In this way,

$$f \text{ is calm at } \bar{x} \iff \text{clm}(f; \bar{x}) < \infty.$$

Definition 1.1.8. (Lipschitz Continuous Functions)

[29, p. 29] A function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be *Lipschitz continuous* relative to a set D , or on a set D , if $D \subset \text{dom } f$ and there exists a constant $\kappa \geq 0$ (Lipschitz constant) such that

$$\|f(x') - f(x)\| \leq \kappa \|x' - x\| \quad \text{for all } x', x \in D \cap \text{dom } f. \quad (1.4)$$

It is said to be Lipschitz continuous *around* \bar{x} when this inequality holds for some neighborhood D of \bar{x} .

The *Lipschitz modulus* of f at \bar{x} , denoted by $\text{lip}(f; \bar{x})$, is the infimum of the set of values of κ for which there exists a neighborhood D of \bar{x} such that (1.4) holds. Equivalently,

$$\text{lip}(f; \bar{x}) = \limsup_{\substack{x, x' \rightarrow \bar{x} \\ x \neq x'}} \frac{\|f(x') - f(x)\|}{\|x' - x\|}.$$

Note 1.1.9. Note that, by this definition, for the Lipschitz modulus we have $\text{lip}(f; \bar{x}) = \infty$ precisely in the case where, for every $\kappa > 0$ and every neighborhood D of \bar{x} , there are points $x', x \in D$ violating (1.4). Thus,

$$f \text{ is Lipschitz continuous around } \bar{x} \iff \text{lip}(f; \bar{x}) < \infty.$$

For an open set C , a function f is *locally Lipschitz continuous* on C exactly when $\text{lip}(f; \bar{x}) < \infty$ for every $x \in C$.

Every continuously differentiable function on an open set C is locally Lipschitz continuous on C .

Note 1.1.10. A function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is differentiable at a point \bar{x} , when $\bar{x} \in \text{int dom } f$ and there is a linear mapping $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ with the property that for every $\epsilon > 0$ there exists $\delta > 0$ with

$$\|f(\bar{x} + h) - f(\bar{x}) - Ah\| \leq \epsilon \|h\| \quad \text{for every } h \in \mathbb{R}^n \text{ with } \|h\| < \delta.$$

If such a mapping A exists at all, it is unique; it is denoted by $Df(\bar{x})$ and is called the *derivative* of f at \bar{x} . The $m \times n$ matrix that represents the derivative $Df(\bar{x})$ is called the *Jacobian* of f at \bar{x} and is denoted by $\nabla f(\bar{x})$.

In distinguishing between $Df(\bar{x})$ as a linear mapping and $\nabla f(\bar{x})$ as its matrix, one can guard better against ambiguities which may arise in some situations. Also it would provide a better form for representing some results coming later (see for instance, Example 3.1.17 and Theorem 3.1.25).

Definition 1.1.11. (Strict Differentiability)

[29, p. 34] A function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be *strictly differentiable* at a point \bar{x} if there is a linear mapping $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$ such that

$$\text{lip}(e; \bar{x}) = 0 \quad \text{for } e(x) := f(x) - [f(\bar{x}) + A(x - \bar{x})]. \quad (1.5)$$

In particular, in this case we have that $\text{clm}(e; \bar{x}) = 0$ and hence f is differentiable at \bar{x} with $A = \nabla f(\bar{x})$, but strictness imposes a requirement on the difference

$$e(x) - e(x') = f(x) - [f(x') + \nabla f(\bar{x})(x - x')]$$

also when $x' \neq \bar{x}$. Specifically, it demands for each $\epsilon > 0$, the existence of a neighborhood U of \bar{x} such that

$$\|f(x) - [f(x') + \nabla f(\bar{x})(x - x')]\| \leq \epsilon \|x - x'\| \quad \text{for every } x, x' \in U. \quad (1.6)$$

Example 1.1.12. For the function $f : \mathbb{R} \rightarrow \mathbb{R}$ defined as below, simple calculations show that f is differentiable at $\bar{x} = 0$, but there is no $\epsilon > 0$ such that (1.6) holds.

$$f(x) := \begin{cases} x^2 \sin\left(\frac{1}{x}\right) & x \neq 0, \\ 0 & x = 0, \end{cases}$$

To observe that, consider the sequences $x_n := \frac{1}{(n + \frac{1}{2})\pi}$, and $x'_n := \frac{1}{(n + \frac{3}{2})\pi}$. Therefore, f is not strictly differentiable at 0.

Remark 1.1.13. The following statements which could be easily obtained from the above definition, are useful in the sequel, especially when dealing with coderivatives.

- Every function f that is continuously differentiable in a neighborhood of \bar{x} is strictly differentiable at \bar{x} (cf. [29, p. 35]).
- Every mapping f strictly differentiable at \bar{x} is Lipschitz continuous around \bar{x} , or *locally Lipschitzian* around this point ([54, p. 19]), that is, there is a neighborhood U of \bar{x} and a constant $l \geq 0$ such that

$$\|f(x) - f(u)\| \leq l \|x - u\| \quad \text{for all } x, u \in U. \quad (1.7)$$

Definition 1.1.14. (Selection) [67, p. 64] Given a set-valued map $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be a *selection* of F if $f(x) \in F(x)$ for each $x \in \text{dom } F$.

Example 1.1.15. Consider the set-valued mapping $F : \mathbb{R} \rightrightarrows \mathbb{R}$ defined for every $x \in \mathbb{R}_+$ with $F(x) := \{y \in \mathbb{R} \mid y \geq x^3\}$; whose graph is the *epigraph* of the function $f : \mathbb{R} \rightarrow \mathbb{R}$,

$f(x) = x^3$ for $x \geq 0$, that is, $\text{gph } F = \text{epi } f$. Then, the function s defined as

$$s(x) := \begin{cases} x^2 & 0 \leq x \leq 1, \\ x^4 & x > 1, \end{cases}$$

is a *continuous* selection of F .

Definition 1.1.16. (Graphical Localization)

[29, p. 4] Given $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a pair $(\bar{x}, \bar{y}) \in \text{gph } F$, a *graphical localization* of F at \bar{x} for \bar{y} is a set-valued mapping \tilde{F} such that

$$\text{gph } \tilde{F} = (U \times V) \cap \text{gph } F \text{ for some neighborhoods } U \text{ of } \bar{x} \text{ and } V \text{ of } \bar{y},$$

so that

$$\tilde{F} : x \mapsto \begin{cases} F(x) \cap V & \text{when } x \in U \\ \emptyset & \text{otherwise.} \end{cases}$$

Note 1.1.17. (a) The inverse of \tilde{F} has the form

$$\tilde{F}^{-1}(y) = \begin{cases} F^{-1}(y) \cap U & \text{when } y \in V \\ \emptyset & \text{otherwise,} \end{cases}$$

and therefore, is a graphical localization of the set-valued mapping F^{-1} at \bar{y} for \bar{x} .

(b) The neighborhoods U and V can often be taken, conveniently, as closed balls $\mathbb{B}_a(\bar{x})$ and $\mathbb{B}_b(\bar{y})$.

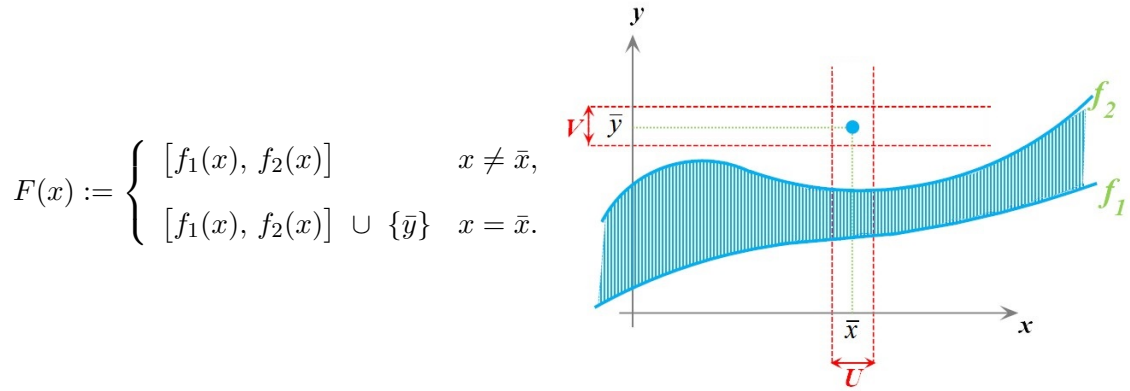
(c) The domain of a graphical localization \tilde{F} of F with respect to U and V may differ from $U \cap \text{dom } F$, and in general depends on the choice of V (like the situation described in the following example).

To avoid this, specially when one deals with the inverse mapping, some authors will consider a slightly different definition which would be *graphical localization around* \bar{x} for \bar{y} that takes into account only neighborhoods $U \subset \text{dom } \tilde{F}$ in the above definition (see, for example, [20]).

This condition guarantees that \bar{x} is an interior point of $\text{dom } \tilde{F}$.

Example 1.1.18. Consider the set valued mapping $F : \mathbb{R} \rightrightarrows \mathbb{R}$ defined in Figure (1.1), in terms of smooth functions f_1 , and f_2 .

One can easily check that based on the chosen V , for any $x \in U \setminus \{\bar{x}\}$, $\tilde{F}(x) = \emptyset$, and $\{\bar{x}\} = \text{dom } \tilde{F} \neq U \cap \text{dom } F$.

FIGURE 1.1: The role of neighborhood V in graphical localization

Definition 1.1.19. (Single-valued Localization) [29, p. 4] By a *single-valued localization* of F at \bar{x} for \bar{y} will be meant a graphical localization that is a function, its domain not necessarily being a neighborhood of \bar{x} .

The case where the domain is indeed a neighborhood of \bar{x} will be indicated by referring to as single-valued localization of F **around** \bar{x} for \bar{y} instead of just **at** \bar{x} for \bar{y} .

Example 1.1.20. (Classical Inverse Function Theorem) [29, p. 11] Considering a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$, one might be interested in the possibility of having a single-valued inverse map. The problem is well studied in the literature under the title of “Inverse Mapping Theorems”.

For future reference, we would like to mention an old classical theorem of this type.

Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be continuously differentiable in a neighborhood of a point \bar{x} and let $\bar{y} := f(\bar{x})$. If $\nabla f(\bar{x})$ is non-singular, then f^{-1} has a single-valued localization s around \bar{y} for \bar{x} . Moreover, the function s is continuously differentiable in a neighborhood V of \bar{y} , and its Jacobian satisfies

$$\nabla s(y) = \nabla f(s(y))^{-1} \quad \text{for every } y \in V.$$

Example 1.1.21. (Implicit Functions) [29, p. 4] In passing from inverse functions to implicit functions, we need to pass from an equation $f(x) = y$ to one of the form

$$g(p, x) = 0 \quad \text{for a function } g : \mathbb{R}^d \times \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad (1.8)$$

in which p acts as a parameter.

The question is no longer about inverting f , but the framework of set-valuedness is valuable nonetheless because it allows us to immediately introduce the *solution mapping*

$$S : \mathbb{R}^d \rightrightarrows \mathbb{R}^n \quad \text{with } S(p) = \{x \mid g(p, x) = 0\}. \quad (1.9)$$

We can then look at pairs (\bar{p}, \bar{x}) in $\text{gph } S$ and ask whether S has a single-valued localization s around \bar{p} for \bar{x} . Such a localization is exactly what constitutes an implicit function coming out of the equation².

Definition 1.1.22. (Monotone Mappings)

[29, p. 195] [11, p. 104] A mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is said to be *monotone* if

$$\langle y' - y, x' - x \rangle \geq 0 \quad \text{whenever} \quad (x', y'), (x, y) \in \text{gph } F. \quad (1.10)$$

It is called *maximal monotone* when no more points can be added to $\text{gph } F$ without running into a violation of (1.10).

In other words, a monotone set-valued map F is *maximal* if there is no other monotone set-valued map whose graph strictly contains the graph of F .

F is called *locally monotone* at \bar{x} for \bar{y} if $(\bar{x}, \bar{y}) \in \text{gph } F$ and for some neighborhood W of (\bar{x}, \bar{y}) , one has

$$\langle y' - y, x' - x \rangle \geq 0 \quad \text{whenever} \quad (x', y'), (x, y) \in \text{gph } F \cap W. \quad (1.11)$$

Example 1.1.23. Consider the set-valued mappings $F_1, F_2 : \mathbb{R} \rightrightarrows \mathbb{R}$ defined as

$$F_1(x) := \begin{cases} x^3 - 1 & x < 0, \\ \{-1, +1\} & x = 0, \\ x^3 + 1 & x > 0. \end{cases} \quad \text{and} \quad F_2(x) := \begin{cases} x^3 - 1 & x < 0, \\ [-1, +1] & x = 0, \\ x^3 + 1 & x > 0. \end{cases}$$

It is easy to check that both mappings are monotone. Since $\text{gph } F_1 \subset \text{gph } F_2$, one concludes that F_1 is not a maximal monotone map.

Remark 1.1.24. The following statements, which could be easily obtained from the above definition, are used in the sequel.

- If A and B are monotone maps and $\lambda, \mu > 0$ are scalars, then $\lambda A + \mu B$ is also monotone (cf. [11, p. 105]).

² **Historical note.** A very early result in this direction was introduced and proved by Ulisse Dini (1845–1918) in his lecture notes of 1877–78, which is now known as the *classical implicit function theorem* or *Dini's theorem*; though the set-valued solution mapping S in (1.9) never enters the picture directly. We express the theorem with our notations as below:

Dini classical implicit function theorem. Let the function $f : \mathbb{R}^d \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ in (1.8) be continuously differentiable in a neighborhood of (\bar{p}, \bar{x}) and such that $f(\bar{p}, \bar{x}) = 0$, and let the partial Jacobian of f with respect to x at (\bar{p}, \bar{x}) , namely $\nabla_x f(\bar{p}, \bar{x})$, be non-singular. Then the solution mapping S defined in (1.9) has a single-valued localization s around \bar{p} for \bar{x} which is continuously differentiable in a neighborhood Q of \bar{p} with Jacobian satisfying

$$\nabla s(p) = -\nabla_x f(p, s(p))^{-1} \nabla_p f(p, s(p)) \quad \text{for every } p \in Q.$$

- Since monotonicity is a property bearing on the graph of F , a set-valued map F is monotone (or maximal monotone) if and only if its inverse F^{-1} is monotone (or maximal monotone) (cf. [11, p. 105]).
- A necessary and sufficient condition for a set-valued map F to be maximal monotone is that the property

$$\langle u - v, x - y \rangle \geq 0, \quad \text{for all } (y, v) \in \text{gph } F$$

is equivalent to $u \in F(x)$ (cf. [11, p. 107]).

Example 1.1.25. (Single-valued Localization and Selection for Locally Monotone Maps) [20] This example will discuss an interesting property of locally monotone maps. The claim is the following:

A set-valued mapping $S : \mathbb{R}^l \rightrightarrows \mathbb{R}^l$, which is locally monotone at $(\bar{y}, \bar{x}) \in \text{gph } S$, has a single-valued Lipschitz continuous localization around \bar{y} for \bar{x} if and only if it has a Lipschitz continuous selection around \bar{y} for \bar{x} .

Find W such that (1.11) holds. Let s be a local selection for S , defined on $\text{int } \mathbb{B}_r(\bar{y})$ for some $r > 0$; and Lipschitz continuous there such that $\text{int } \mathbb{B}_r(\bar{y}) \times \text{int } \mathbb{B}_{\kappa r}(\bar{x}) \subset W$, where $\kappa > 0$ is the corresponding Lipschitz constant.

Fix any $y \in \text{int } \mathbb{B}_r(\bar{y})$. By definition, $s(\bar{y}) = \bar{x}$, and

$$\|s(y) - s(\bar{y})\| \leq \kappa \|y - \bar{y}\| \leq \kappa r,$$

implies that $s(y) \in \text{int } \mathbb{B}_{\kappa r}(\bar{x})$. Therefore, the point $s(y)$ lies in $S(y) \cap \text{int } \mathbb{B}_{\kappa r}(\bar{x})$. It suffices to show that the latter set is singleton. Suppose that this is not the case. Find $x \in \mathbb{R}^l$ such that

$$x \in S(y) \cap \text{int } \mathbb{B}_{\kappa r}(\bar{x}) \quad \text{with} \quad x \neq s(y).$$

Let $b := \|x - s(y)\| > 0$, and $c := \frac{x - s(y)}{b}$. Thus,

$$\begin{aligned} \langle x, c \rangle &= \left\langle x, \frac{x - s(y)}{b} \right\rangle = \frac{1}{b} \langle x, x - s(y) \rangle = \frac{1}{b} \langle x - s(y) + s(y), x - s(y) \rangle \\ &= \frac{1}{b} \langle x - s(y), x - s(y) \rangle + \frac{1}{b} \langle s(y), x - s(y) \rangle \\ &= b + \langle s(y), c \rangle \end{aligned} \tag{1.12}$$

Find $\tau > 0$ such that $\kappa\tau < b$ and $y + \tau c \in \text{int } \mathbb{B}_r(\bar{y})$. Since $\|c\| = 1$, the Cauchy-Schwartz inequality and the Lipschitz continuity of s imply that

$$\langle s(y + \tau c) - s(y), c \rangle \leq \|s(y + \tau c) - s(y)\| \|c\| \leq \kappa \|(y + \tau c) - y\| = \kappa\tau. \tag{1.13}$$

Since $(y + \tau c, s(y + \tau c))$ and (y, x) are in $\text{gph } S \cap W$, local monotonicity of S , Condition (1.11), reveals that $0 \leq \langle s(y + \tau c) - x, y + \tau c - y \rangle = \tau \langle s(y + \tau c) - x, c \rangle$. Thus,

$$\langle s(y + \tau c), c \rangle \geq \langle x, c \rangle. \quad (1.14)$$

Now, one can use these equations to obtain

$$b + \langle s(y), c \rangle \stackrel{(1.12)}{=} \langle x, c \rangle \stackrel{(1.14)}{\leq} \langle s(y + \tau c), c \rangle \stackrel{(1.13)}{\leq} \langle s(y), c \rangle + \kappa\tau < \langle s(y), c \rangle + b.$$

We arrived at a contradiction, therefore $S(y) \cap \text{int } \mathbb{B}_{\kappa r}(\bar{x}) = \{s(y)\}$ for each $y \in \text{int } \mathbb{B}_r(\bar{y})$. The opposite implication is trivial.

1.2 Local Stability Properties

This section is devoted to introducing the main local properties we would consider in this thesis. These properties which mostly has a well known counterpart for single-valued maps, let us stud the local behaviour of a set-valued map.

Later in Chapter 2, we will see how slight modifications in formulating the problem will result to each of these properties.

Definition 1.2.1. (Locally Closed Set) [29, p. 172] A set C is said to be *locally closed* at $x \in C$ if there exists a neighborhood U of x such that the intersection $C \cap U$ is closed.

It could be equivalently defined as the existence of a scalar $r > 0$ such that the set $C \cap \mathbb{B}_r(x)$ is closed.

Definition 1.2.2. (Aubin Property)

[29, p. 172] A mapping $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is said to have the *Aubin property*³ at $\bar{y} \in \mathbb{R}^m$ for $\bar{x} \in \mathbb{R}^n$ if $\bar{x} \in S(\bar{y})$, the graph of S is locally closed at (\bar{y}, \bar{x}) , and there is a constant $\kappa \geq 0$ together with neighborhoods U of \bar{x} and V of \bar{y} such that

$$e(S(y') \cap U, S(y)) \leq \kappa \|y' - y\| \quad \text{for all } y', y \in V, \quad (1.15)$$

³ **Historical note.** Actually, the property that was introduced by Aubin in [11] under the name of “pseudo-Lipschitz”, with the formulation (1.16). Under closedness of the values of the mapping both formulations and their constants agree, see [8]; but without this assumption, the constant κ in (1.16) might be slightly larger than κ in (1.15). For instance, consider the following mapping $S : \mathbb{R} \rightrightarrows \mathbb{R}$ defined as

$$S(y) := \begin{cases} (-\infty, y] & \text{for } y \text{ rational,} \\ (-\infty, y) & \text{for } y \text{ irrational.} \end{cases}$$

Here (1.15) holds around $(0, 0)$ for $\kappa = 1$ while (1.16) is only valid for $\kappa > 1$.

or equivalently, there exist κ, U and V , as described, such that

$$S(y') \cap U \subset S(y) + \kappa \|y' - y\| \mathbb{B} \quad \text{for all } y', y \in V. \quad (1.16)$$

The infimum of κ over all such combinations of κ, U , and V , is called the *Lipschitz modulus* of S at \bar{y} for \bar{x} and is denoted by $\text{lip}(S; \bar{y} | \bar{x})$.

Note 1.2.3. (a) The absence of this property is signalled by $\text{lip}(S; \bar{y} | \bar{x}) = \infty$.

(b) When S is single-valued on a neighborhood of \bar{y} , the Lipschitz modulus $\text{lip}(S; \bar{y} | S(\bar{y}))$ equals the usual Lipschitz modulus $\text{lip}(S; \bar{y})$ for functions.

(c) It is not claimed that (1.15) and (1.16) are themselves equivalent, although this is true when $S(y)$ is closed for every $y \in V$. Nonetheless, the infimum furnishing $\text{lip}(S; \bar{y} | \bar{x})$ is the same whichever formulation is adopted.

Remark 1.2.4. A mapping $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is said to be *Lipschitz continuous* relative to a (non-empty) set D in \mathbb{R}^m if $D \subset \text{dom } S$, S is closed-valued on D , and there exists $\kappa \geq 0$ (Lipschitz constant) such that

$$S(y') \subset S(y) + \kappa \|y' - y\| \mathbb{B} \quad \text{for all } y', y \in D,$$

or equivalently, there exists $\kappa \geq 0$ such that

$$d(x, S(y)) \leq \kappa d(y, S^{-1}(x) \cap D) \quad \text{for all } x \in \mathbb{R}^n \text{ and } y \in D.$$

The similarity between *Aubin property* and this definition is the reason why some authors refer to having the *Aubin property* as being “*Lipschitz-like*” or “*pseudo-Lipschitz*”⁴.

Example 1.2.5. [29, p. 172] Consider the set-valued mapping $S : \mathbb{R} \rightrightarrows \mathbb{R}$ defined as

$$S(y) = \begin{cases} \{0, 1 + \sqrt{y}\} & \text{for } y \geq 0, \\ 0 & \text{for } y < 0. \end{cases} \quad (1.17)$$

At 0, the value $S(0)$ consists of two points, 0 and 1. This mapping has the Aubin property at 0 for 0 but not at 0 for 1. To see the latter, consider sequences $y_n = \frac{1}{n}$, and $y'_n = \frac{2}{n}$ for $n \in \mathbb{N}$. For any neighborhood V of $\bar{y} = 0$, there is $N \in \mathbb{N}$ such that for $n > N$, points y_n , and y'_n are inside V . Then, no matter how small would be the neighborhood U of $\bar{x} = 1$, $S(y'_n) \cap U$ would include $1 + \sqrt{y'_n}$ and thus, $e(S(y'_n) \cap U, S(y_n)) = \frac{y'_n - y_n}{\sqrt{y'_n} + \sqrt{y_n}}$.

⁴ This definition is in fact an equivalent way to define Lipschitz continuity for set-valued maps, based on the Proposition 3C.1 [29, p.161]. This form suits better for the purpose of comparison.

To have the Aubin property, κ must satisfy the following inequality

$$\frac{\sqrt{n}}{1 + \sqrt{2}} \leq \kappa,$$

which is absurd. This example shows that the Aubin property is tied to a particular point in the graph of the mapping.

The Aubin property could alternatively be defined with one variable “free,” as shown in the next proposition.

Proposition 1.2.6 (Alternative Description of Aubin Property).

[29, p. 176] *A mapping $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ has the Aubin property at \bar{y} for \bar{x} with constant $\kappa > 0$ if and only if $\bar{x} \in S(\bar{y})$, $\text{gph } S$ is locally closed at (\bar{y}, \bar{x}) , and there exist neighborhoods U of \bar{x} and V of \bar{y} such that*

$$e(S(y') \cap U, S(y)) \leq \kappa \|y' - y\| \quad \text{for all } y' \in \mathbb{R}^m \text{ and } y \in V. \quad (1.18)$$

Proof. Clearly, (1.18) implies (1.15). Assume (1.15) with corresponding U and V and choose positive a and b such that $\mathbb{B}_a(\bar{x}) \subset U$ and $\mathbb{B}_b(\bar{y}) \subset V$. Let $0 < a' < a$ and $0 < b' < b$ be such that

$$2\kappa b' + a' \leq \kappa b. \quad (1.19)$$

For any $y \in \mathbb{B}_{b'}(\bar{y})$ we have from (1.15) that

$$d(\bar{x}, S(y)) \leq e(S(\bar{y}) \cap U, S(y)) \leq \kappa \|y - \bar{y}\| \leq \kappa b',$$

hence

$$e(\mathbb{B}_{a'}(\bar{x}), S(y)) \leq \kappa b' + a'. \quad (1.20)$$

Take any $y' \in \mathbb{R}^m$. If $y' \in \mathbb{B}_b(\bar{y})$ the inequality in (1.18) comes from (1.15) and there is nothing more to prove. Assume $\|y' - \bar{y}\| > b$. Then $\|y - y'\| > b - b'$ and from (1.19),

$$\kappa b' + a' \leq \kappa(b - b') \leq \kappa\|y - y'\|.$$

Using this in (1.20) we obtain

$$e(\mathbb{B}_{a'}(\bar{x}), S(y)) \leq \kappa \|y' - y\| \quad (1.21)$$

and since $S(y') \cap \mathbb{B}_{a'}(\bar{x})$ is obviously a subset of $\mathbb{B}_{a'}(\bar{x})$, we come again to (1.18). \square

The following proposition will highlight an interesting fact of a map with Aubin property, which we will use later in Chapter 3.

Proposition 1.2.7 (Local Non-emptiness).

[29, p. 173] If $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ has the Aubin property at \bar{y} for \bar{x} , then for every neighborhood U of \bar{x} there exists a neighborhood V of \bar{y} such that $S(y) \cap U \neq \emptyset$ for all $y \in V$.

Proof. The inclusion (1.16) for $y' = \bar{y}$ yields

$$\bar{x} \in S(y) + \kappa \|y - \bar{y}\| \mathbb{B} \quad \text{for every } y \in V,$$

which is the same as

$$(\bar{x} + \kappa \|y - \bar{y}\| \mathbb{B}) \cap S(y) \neq \emptyset \quad \text{for every } y \in V.$$

That is, $S(y)$ intersects every neighborhood of \bar{x} when y is sufficiently close to \bar{y} . \square

The next property would be the “one-variable” version of the Aubin property. We have already introduced a calm function (cf. Definition 1.1.6); it is now time to define the set-valued counterpart.

Definition 1.2.8. (Calmness)

[29, p. 197] A mapping $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is said to be *calm* at \bar{y} for \bar{x} if $(\bar{y}, \bar{x}) \in \text{gph } S$, and there is a constant $\kappa \geq 0$ along with neighborhoods U of \bar{x} and V of \bar{y} such that

$$e(S(y) \cap U, S(\bar{y})) \leq \kappa \|y - \bar{y}\| \quad \text{for all } y \in V. \quad (1.22)$$

Equivalently, the property in (1.22) can be also written as

$$S(y) \cap U \subset S(\bar{y}) + \kappa \|y - \bar{y}\| \mathbb{B} \quad \text{for all } y \in V \quad (1.23)$$

although perhaps with larger constant κ .

The infimum of κ over all such combinations of κ , U and V is called the *calmness modulus* of S at \bar{y} for \bar{x} and is denoted by $\text{clm}(S; \bar{y}|\bar{x})$.

Note 1.2.9. (a) The absence of this property is signalled by $\text{clm}(S; \bar{y}|\bar{x}) = \infty$.

(b) As in the case of the Lipschitz modulus, it is not claimed that (1.22) and (1.23) are themselves equivalent; anyway, the infimum furnishing $\text{clm}(S; \bar{y}|\bar{x})$ is the same with respect to either of them.

Definition 1.2.10. (Isolated Calmness)

[29, p. 201] A mapping $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is said to have the *isolated calmness* property if it is calm at \bar{y} for \bar{x} and, in addition, S has a graphical localization at \bar{y} for \bar{x} that is single-valued at \bar{y} itself (with value \bar{x}). Specifically, this refers to the existence of a

constant $\kappa \geq 0$ and neighborhoods U of \bar{x} and V of \bar{y} such that

$$\|x - \bar{x}\| \leq \kappa \|y - \bar{y}\| \quad \text{when } x \in S(y) \cap U \quad \text{and } y \in V^5. \quad (1.24)$$

Note 1.2.11. (a) Observe that in this definition $S(\bar{y}) \cap U$ is a singleton, namely the point \bar{x} , so \bar{x} is an isolated point in $S(\bar{y})$, hence the terminology.

(b) Isolated calmness can equivalently be defined as the existence of a (possibly slightly larger) constant κ and neighborhoods U of \bar{x} and V of \bar{y} such that

$$S(y) \cap U \subset \bar{x} + \kappa \|y - \bar{y}\| \mathbb{B} \quad \text{when } y \in V. \quad (1.25)$$

The last property of this subsection would be Linear openness, for which, we remind the definition of an open map.

Definition 1.2.12. (Openness) [29, p. 180] A mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is said to be *open* at \bar{x} for \bar{y} if $\bar{y} \in F(\bar{x})$, and for every neighborhood U of \bar{x} , $F(U)$ is a neighborhood of \bar{y} .

Definition 1.2.13. (Linear Openness)

[29, p. 180] A mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is said to be *linearly open* at \bar{x} for \bar{y} when $\bar{y} \in F(\bar{x})$, the graph of F is locally closed at (\bar{x}, \bar{y}) , and there is a constant $\kappa > 0$ together with neighborhoods U of \bar{x} and V of \bar{y} such that

$$F(x + \kappa r \text{ int } \mathbb{B}) \supset [F(x) + r \text{ int } \mathbb{B}] \cap V \quad \text{for all } x \in U \quad \text{and all } r > 0. \quad (1.26)$$

Note 1.2.14. Linear openness is in fact, a particular case of openness and follows from (1.26) for $x = \bar{x}$, and the following example shows the reverse is not true in general. Linear openness postulates openness *around* the reference point with balls having proportional radii.

For example, consider the function $f : \mathbb{R} \rightarrow \mathbb{R}$ defined as $f(x) = x^3$ for every $x \in \mathbb{R}$. Obviously, f is open around the origin, but in order to have the linear openness at $\bar{x} = 0$ for $\bar{y} = 0$, we must find $\kappa > 0$ together with neighborhoods U of \bar{x} and V of \bar{y} such that (1.26) holds.

In particular, let $x = \bar{x} = 0$. Thus, $f(x) + r \text{ int } \mathbb{B} = \text{int } \mathbb{B}_r(0)$, and we can also compute $f(x + \kappa r \text{ int } \mathbb{B}) = \text{int } \mathbb{B}_{\kappa^3 r^3}(0)$. Hence, we must have $\kappa^3 > \frac{1}{r^2}$, which implies $\kappa \rightarrow \infty$ as $r \rightarrow 0$. Therefore, it does not have the linear openness property at the origin.

⁵ Isolated calmness was formally introduced by A. L. Dontchev in [24] under the name “local upper Lipschitz continuity at a point”.

1.3 Regularity Terms and Their Relation with Local Stability Properties

In this section, we will introduce a regularity concept in terms of the distance of sets, hence a “metric” regularity. The first benefit of these metric regularity terms (the four different versions) is to gather closely related properties which describe the local behaviour of a set-valued map under a single roof, that is metric regularity.

The small differences between these definitions and their relation with local stability properties defined in the previous section is studied throughout this section. It is worth mentioning that Ioffe in [43, 44] has provided a comprehensive note on metric regularity, with many old and new results.

Definition 1.3.1. (Metric Regularity)

[29, p. 176] A mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is said to be *metrically regular*⁶ at \bar{x} for \bar{y} when $\bar{y} \in F(\bar{x})$, the graph of F is locally closed at (\bar{x}, \bar{y}) , and there is a constant $\kappa \geq 0$ together with neighborhoods U of \bar{x} and V of \bar{y} such that

$$d(x, F^{-1}(y)) \leq \kappa d(y, F(x)) \quad \text{whenever } (x, y) \in U \times V. \quad (1.27)$$

The infimum of κ over all such combinations of κ , U and V is called the *regularity modulus* of F at \bar{x} for \bar{y} and is denoted by $\text{reg}(F; \bar{x}|\bar{y})$.

Note 1.3.2. The absence of metric regularity is signalled by $\text{reg}(F; \bar{x}|\bar{y}) = \infty$.

Example 1.3.3. (Metric Regularity for a Bounded Linear Operator) [8]

The concept of metric regularity goes back to the classical Banach open mapping principle. It is a well known result by Lyusternik and Graves [34] that a stronger property than the usual openness can be deduced from the open mapping principle (which we already introduced it as linear openness, see Definition 1.2.13).

For any bounded linear mapping A between two Banach spaces X and Y , the Banach open mapping theorem states the equivalence between:

- (a) A is surjective;
- (b) A is an open map (at every point);
- (c) there exists $\kappa > 0$ such that for every $y \in Y$ there exists $x \in X$ with

$$y = Ax \quad \text{and} \quad \|x\| \leq \kappa \|y\|.$$

⁶ **Historical note.** Although the term “regularity” has been already used in the literature to describe a similar property (see, for example [42, 60]), it was firstly Borwein [15] who called this property *metric regularity*. Let f be a C^1 function from a Banach space X (or from an open set in X) into another Banach space Y . It was usually called *regular* at z if $f'(z)$ maps X onto Y .

Consider a linear function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ which is surjective; then, according to (c), there exists $\kappa > 0$ such that

$$d(0, f^{-1}(y)) \leq \kappa \|y\| \quad \text{for all } y \in \mathbb{R}^m = \text{rge } f.$$

Because of linearity, for any $x \in \mathbb{R}^n$ we have $f^{-1}(y - f(x)) = f^{-1}(y) - x$, and then

$$d(x, f^{-1}(y)) = d(0, f^{-1}(y - f(x))) \leq \kappa \|y - f(x)\| = \kappa d(y, f(x)).$$

This is a global form of the metric regularity condition for the function f . Hence, we showed that, in terms of metric regularity, the equivalence between (a) and (c) can be stated as follows: f is surjective if and only if it is metrically regular at any point $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$.

Remark 1.3.4. [29, p. 177] Metric regularity is a valuable concept in its own right, especially for numerical purposes (see also [26]). For a general set-valued mapping F and a vector y , it gives an estimate for how far a point x is from being a solution to the generalized equation $F(x) \ni y$ in terms of the “residual”⁷ $d(y, F(x))$.

To be specific, let \bar{x} be a solution of the inclusion $\bar{y} \in F(x)$, let F be metrically regular at \bar{x} for \bar{y} , and let x_a and y_a be approximations to \bar{x} and \bar{y} , respectively. Then from (1.27), one can deduce that the distance from x_a to the set of solutions of the inclusion $y_a \in F(x)$ (i.e. $F^{-1}(y_a)$) is bounded by the constant κ times the residual $d(y_a, F(x_a))$. In applications, the residual is typically easy to compute or estimate, whereas finding a solution might be considerably more difficult. Metric regularity says that there exists a solution to the inclusion $y_a \in F(x)$ at distance from x_a proportional to the residual. In particular, if we know the rate of convergence of the residual to zero, then we will obtain the rate of convergence of approximate solutions to an exact one.

⁷ **Historical note.** The term *residual function* seems to have an origin in the theory of Error Bounds, as explicitly discussed in [35, chapter 6].

Let X be a given subset of \mathbb{R}^n . We are interested in obtaining inequalities that bound the distance $d(x, X)$ from points $x \in S$ (a given set) to X in terms of a computable nonnegative-valued function $r : S \cup X \rightarrow \mathbb{R}_+$ whose zeros coincide with the elements of X ; that is,

$$r(x) = 0 \Leftrightarrow x \in X.$$

Given a residual function r of the set X and a subset (called *test set*) S of \mathbb{R}^n , we wish to establish the existence of positive scalars c_1, c_2, γ_1 , and γ_2 such that

$$c_1 r(x)^{\gamma_1} \leq d(x, X) \leq c_2 r(x)^{\gamma_2}, \quad \forall x \in S.$$

Such an inequality is called an *error bound* for the set X with residual r and with respect to S . Since the residual function is a computable quantity whereas the exact distance function is not (because we do not know the elements of the set X), the former function can therefore be used as a surrogate of the latter function for computational purposes; such as in the design of solution algorithms for computing an (approximate) element of X .

Apart from the importance of metric regularity mentioned in the previous remark, the next two theorems reveal another aspect of importance of this concept, which would be its relation with the Aubin property and linear openness.

Theorem 1.3.5 (Equivalence of Metric Regularity and the Aubin Property of the Inverse). [29, p. 177]

A set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is metrically regular at \bar{x} for \bar{y} with a constant $\kappa > 0$ if and only if its inverse $F^{-1} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ has the Aubin property at \bar{y} for \bar{x} with constant $\kappa > 0$, i.e. there exist neighborhoods U of \bar{x} and V of \bar{y} such that

$$e(F^{-1}(y') \cap U, F^{-1}(y)) \leq \kappa \|y' - y\| \quad \text{for all } y', y \in V. \quad (1.28)$$

Thus,

$$\text{lip}(F^{-1}; \bar{y}|\bar{x}) = \text{reg}(F; \bar{x}|\bar{y}). \quad (1.29)$$

Proof. Clearly, the local closedness of the graph of F at (\bar{x}, \bar{y}) is equivalent to the local closedness of the graph of F^{-1} at (\bar{y}, \bar{x}) .

Let $\kappa > \text{reg}(F; \bar{x}|\bar{y})$; then there are positive constants a and b such that (1.27) holds with $U = \mathbb{B}_a(\bar{x})$, $V = \mathbb{B}_b(\bar{y})$ and with this κ . Without loss of generality, assume $b < \frac{a}{\kappa}$ (to be more precise, one can take $b_0 < \min\{b, a/\kappa\}$). Choose $y, y' \in \mathbb{B}_b(\bar{y})$. If $F^{-1}(y) \cap \mathbb{B}_a(\bar{x}) = \emptyset$, then $d(\bar{x}, F^{-1}(y)) > a$. But then the inequality (1.27) with $x = \bar{x}$ yields

$$a \leq d(\bar{x}, F^{-1}(y)) \leq \kappa d(y, F(\bar{x})) \leq \kappa \|y - \bar{y}\| \leq \kappa b < a,$$

a contradiction. Hence there exists $x \in F^{-1}(y) \cap \mathbb{B}_a(\bar{x})$, and for any such x we have from (1.27) that

$$d(x, F^{-1}(y')) \leq \kappa d(y', F(x)) \leq \kappa \|y - y'\|. \quad (1.30)$$

Taking the supremum with respect to $x \in F^{-1}(y) \cap \mathbb{B}_a(\bar{x})$ we obtain (1.28) with $U = \mathbb{B}_a(\bar{x})$ and $V = \mathbb{B}_b(\bar{y})$, and therefore (considering the infimum in the definition of Lipschitz modulus)

$$\text{reg}(F; \bar{x}|\bar{y}) \geq \text{lip}(F^{-1}; \bar{y}|\bar{x}). \quad (1.31)$$

Conversely, suppose there are neighborhoods U of \bar{x} and V of \bar{y} along with a constant κ such that (1.28) holds. Take U and V smaller if necessary so that, according to Proposition (1.2.6), we have

$$e(F^{-1}(y') \cap U, F^{-1}(y)) \leq \kappa \|y' - y\| \quad \text{for all } y' \in \mathbb{R}^m \text{ and } y \in V. \quad (1.32)$$

Let $x \in U$ and $y \in V$. If $F(x) \neq \emptyset$, then for any $y' \in F(x)$ we have $x \in F^{-1}(y') \cap U$. From (1.32), we obtain

$$d(x, F^{-1}(y)) \leq e(F^{-1}(y') \cap U, F^{-1}(y)) \leq \kappa \|y - y'\|. \quad (1.33)$$

This holds for any $y' \in F(x)$, hence, by taking the infimum with respect to $y' \in F(x)$ in the last expression we get

$$d(x, F^{-1}(y)) \leq \kappa d(y, F(x)). \quad (1.34)$$

(If $F(x) = \emptyset$, then because of the convention $d(y, \emptyset) = \infty$, this inequality holds automatically.) Hence, F is metrically regular at \bar{x} for \bar{y} with a constant κ . Then we have $\kappa \geq \text{reg}(F; \bar{x}|\bar{y})$ and hence $\text{reg}(F; \bar{x}|\bar{y}) \leq \text{lip}(F^{-1}; \bar{y}|\bar{x})$. This inequality together with (1.31) results in (1.29). \square

Note 1.3.6. [29, p. 180] From the equivalence of metric regularity of F at \bar{x} for \bar{y} and the Aubin property of F^{-1} at \bar{y} for \bar{x} , and Proposition 1.2.7, we obtain that if a mapping F is metrically regular at \bar{x} for \bar{y} , then F is open at \bar{x} for \bar{y} .

Metric regularity is actually equivalent to the stronger version of the openness property as mentioned in the following theorem:

Theorem 1.3.7 (Equivalence of Linear Openness and Metric Regularity).

[29, p. 180] *A set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is linearly open at \bar{x} for \bar{y} if and only if F is metrically regular at \bar{x} for \bar{y} . In this case the infimum of κ for which (1.26) holds is equal to $\text{reg}(F; \bar{x}|\bar{y})$.*

Proof. Both properties require local closedness of the graph of F at the reference point. Let (1.26) hold. Choose $y \in V$ and $x' \in U$. Let $y' \in F(x')$ (if there is no such y' there is nothing to prove). Since $y = y' + \|y - y'\|w$ for some $w \in \mathbb{B}$, denoting $r = \|y - y'\|$, for every $\epsilon > 0$ we have $y \in (F(x') + r(1 + \epsilon) \text{int } \mathbb{B}) \cap V$. From (1.26), there exists $x \in F^{-1}(y)$ with $\|x - x'\| \leq \kappa(1 + \epsilon)r = \kappa(1 + \epsilon)\|y' - y\|$. Then

$$d(x', F^{-1}(y)) \leq \kappa(1 + \epsilon)\|y' - y\|.$$

Taking infimum with respect to $y' \in F(x')$ on the right and passing to zero with ϵ (since the left side does not depend on ϵ), we obtain that F is metrically regular at \bar{x} for \bar{y} with constant κ .

For the converse, let $x \in U$, $r > 0$, and let $y' \in (F(x) + r \text{int } \mathbb{B}) \cap V$. Then there exists $y \in F(x)$ such that $\|y - y'\| < r$.

If $y = y'$ then $y' \in F(x) \subset F(x + \kappa r \text{int } \mathbb{B})$, which yields (1.26) with constant κ .

Let $y \neq y'$. From the equivalence of metric regularity of F and Aubin property of F^{-1} (proved in Theorem 1.3.5), and by using the characterization of the Aubin property given in Proposition (1.2.6) we obtain

$$d(x, F^{-1}(y')) \leq e(F^{-1}(y) \cap U, F^{-1}(y')) \leq \kappa \|y - y'\| < \kappa r.$$

Then there exists $x' \in F^{-1}(y')$ such that $\|x - x'\| < \kappa r$. But then

$$y' \in F(x') \subset F(x + \kappa r \operatorname{int} \mathbb{B}),$$

which again yields (1.26) with constant κ . □

Example 1.3.8. Consider the function $f : \mathbb{R}^2 \rightarrow \mathbb{R}$, defined as:

$$f(x_1, x_2) = \begin{cases} x_2 + x_1^2 & x_2 \geq 0, \\ x_2 - x_1^2 & \text{otherwise.} \end{cases}$$

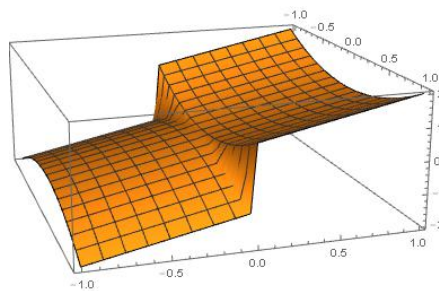


FIGURE 1.2: A map which is open but not linearly open at $\bar{x} = (0, 0)$ for $\bar{y} = 0$

This function does not have the linear openness property at $\bar{x} = (0, 0)$ for $\bar{y} = 0$, but it is open at that point. Indeed, one can consider the neighborhood U of \bar{x} as $U = \mathbb{B}_{r_1}(0) \times \mathbb{B}_{r_2}(0)$ for $r_1, r_2 > 0$, and observe that

$$\|f(x) - 0\| \leq r_2 + r_1^2 \quad \text{for any } x = (x_1, x_2) \in \mathbb{B}_{r_1}(0) \times \mathbb{B}_{r_2}(0).$$

Thus, $f(U)$ is a neighborhood of 0, and f is open at $\bar{x} = (0, 0)$ for $\bar{y} = 0$.

To see that it is not linearly open at the reference point, in view of Theorem 1.3.7, we would assume by contradiction that it is metrically regular at \bar{x} for \bar{y} . Thus, there should exist a constant $\kappa > 0$, together with neighborhoods U of \bar{x} and V of \bar{y} such that

$$d(x, f^{-1}(y)) \leq \kappa d(y, f(x)) \quad \text{for any } x \in U, y \in V.$$

Assume that $y = \bar{y} = 0$, $U = \mathbb{B}_r((0, 0))$, with a fixed $r > 0$, and the sequence of points $x_n = (\frac{1}{n}, 0)$, with $n \in \mathbb{N}$. Then, for n large enough, $x_n \in U$, and we have $d(y, f(x_n)) = \frac{1}{n^2}$. On the other hand, $d(x_n, f^{-1}(y)) = d(\frac{1}{n}, 0) = \frac{1}{n}$. This implies that $\kappa = n \rightarrow \infty$, which is a contradiction.

Definition 1.3.9. (Strong Metric Regularity)

[29, p. 194] A mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ with $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ whose inverse F^{-1} has

a Lipschitz continuous single-valued localization around \bar{y} for \bar{x} will be called *strongly metrically regular* at \bar{x} for \bar{y} .

Note 1.3.10. The terminology of strong metric regularity offers a way of gaining new perspectives on earlier results by translating them into the language of metric regularity. Indeed, strong metric regularity is just metric regularity plus the existence of a single-valued localization of the inverse.

The following theorem, provides a tool to check strong metric regularity via metric regularity, in examples. Indeed, it suggests that having metric regularity in hand, it only suffices to check the localization of the inverse map for not being multi-valued. In view of Theorem 1.3.5, it also enlighten the relation between strong metric regularity and the Aubin property.

Proposition 1.3.11 (Single-Valued Localizations and Metric Regularity).

[29, p. 192] For a mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a pair $(\bar{x}, \bar{y}) \in \text{gph } F$, the following properties are equivalent:

- (a) F^{-1} has a Lipschitz continuous single-valued localization s around \bar{y} for \bar{x} ;
- (b) F is metrically regular at \bar{x} for \bar{y} and F^{-1} has a localization at \bar{y} for \bar{x} that is nowhere multivalued.

Indeed, in the circumstances of (b) the localization s in (a) has $\text{lip}(s; \bar{y}) = \text{reg}(F; \bar{x}|\bar{y})$.

There are certain situations in which, one can not distinguish between metric regularity and strong metric regularity; one of these situations is described in the following proposition.

Proposition 1.3.12 (Strong Metric Regularity of Locally Monotone Mappings). [29, Theorem 3G.5, p. 195]

If a mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ that is locally monotone at \bar{x} for \bar{y} is metrically regular at \bar{x} for \bar{y} , then it must be strongly metrically regular at \bar{x} for \bar{y} .

The “one-variable ” version of metric regularity is defined as metric sub-regularity. Considering Theorem 1.3.5, and definition of calmness and isolated calmness, one may think of a relation between metric sub-regularity and the calmness of the inverse map. Such a relation actually exists and will be expressed afterwards.

Definition 1.3.13. (Metric Sub-Regularity)

[29, p. 198] A mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is called *metrically sub-regular* at \bar{x} for \bar{y} if $(\bar{x}, \bar{y}) \in \text{gph } F$ and there exists $\kappa \geq 0$ along with neighborhoods U of \bar{x} and V of \bar{y} such that

$$d(x, F^{-1}(\bar{y})) \leq \kappa d(\bar{y}, F(x) \cap V) \quad \text{for all } x \in U. \quad (1.35)$$

The infimum of all κ for which (1.35) holds is the *modulus* of metric sub-regularity, denoted by $\text{subreg}(F; \bar{x}|\bar{y})$.

The absence of metric sub-regularity is signaled by $\text{subreg}(F; \bar{x}|\bar{y}) = \infty$.

The main difference between metric sub-regularity and metric regularity is that the data input \bar{y} is now fixed and not perturbed to a nearby y .

It is worth mentioning that, since $d(\bar{y}, F(x)) \leq d(\bar{y}, F(x) \cap V)$, it is clear that sub-regularity is a weaker condition than metric regularity, and

$$\text{subreg}(F; \bar{x}|\bar{y}) \leq \text{reg}(F; \bar{x}|\bar{y}).$$

Theorem 1.3.14 (Characterization by Calmness of the Inverse). [29, p. 198]

For a mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, let $\bar{y} \in F(\bar{x})$. Then F is metrically sub-regular at \bar{x} for \bar{y} if and only if its inverse $F^{-1} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ is calm at \bar{y} for \bar{x} , in which case

$$\text{clm}(F^{-1}; \bar{y}|\bar{x}) = \text{subreg}(F; \bar{x}|\bar{y}).$$

Proof. First assume that F^{-1} is calm, i.e., there exist a constant $\kappa > 0$ and neighborhoods U of \bar{x} and V of \bar{y} such that

$$F^{-1}(y) \cap U \subset F^{-1}(\bar{y}) + \kappa \|y - \bar{y}\| \mathbb{B} \quad \text{for all } y \in V. \quad (1.36)$$

Let $x \in U$. If $F(x) \cap V = \emptyset$, then the right side of (1.35) is ∞ and we are done. If not, having $x \in U$ and $y \in F(x) \cap V$ is the same as having $x \in F^{-1}(y) \cap U$ and $y \in V$. For such x and y , the inclusion in (1.36) requires the ball $x + \kappa \|y - \bar{y}\| \mathbb{B}$ to have non-empty intersection with $F^{-1}(\bar{y})$. Then $d(x, F^{-1}(\bar{y})) \leq \kappa \|y - \bar{y}\|$. Thus, for any $x \in U$, we must have $d(x, F^{-1}(\bar{y})) \leq \inf_y \{ \kappa \|y - \bar{y}\| : y \in F(x) \cap V \}$ which is (1.35). This shows that (1.36) implies (1.35) and that

$$\inf\{\kappa \mid U, V, \kappa \text{ satisfying (1.36)}\} \geq \inf\{\kappa \mid U, V, \kappa \text{ satisfying (1.35)}\}$$

the latter being by definition $\text{subreg}(F; \bar{x}|\bar{y})$.

For the opposite direction, we have to demonstrate that if $\text{subreg}(F; \bar{x}|\bar{y}) < \kappa < \infty$, then (1.36) holds for some choice of neighborhoods U and V . Consider any κ' with $\text{subreg}(F; \bar{x}|\bar{y}) < \kappa' < \kappa$. For this κ' , there exist U and V such that

$$d(x, F^{-1}(\bar{y})) \leq \kappa' d(\bar{y}, F(x) \cap V) \quad \text{for all } x \in U.$$

Then we have $d(x, F^{-1}(\bar{y})) \leq \kappa \|y - \bar{y}\|$ when $x \in U$ and $y \in F(x) \cap V$, or equivalently $y \in V$ and $x \in F^{-1}(y) \cap U$.

Fix $y \in V$. If $y = \bar{y}$ there is nothing to prove; let $y \neq \bar{y}$. If $x \in F^{-1}(y) \cap U$, then

$d(x, F^{-1}(\bar{y})) \leq \kappa' \|y - \bar{y}\| < \kappa \|y - \bar{y}\|$. Then there must be a point $x' \in F^{-1}(\bar{y})$ having $\|x' - x\| < \kappa \|y - \bar{y}\|$. Hence we have (1.36), as required, and the proof is complete. \square

Note that, undesirably, the property of metric sub-regularity (and hence calmness) is not stable under smooth perturbation, even for convex multifunctions, as demonstrated by the following example.

Example 1.3.15. (Instability of Metric Sub-Regularity under Perturbation)

[39, Example 2.1, p. 1441] Consider the convex multifunction $G : \mathbb{R} \rightrightarrows \mathbb{R}$, defined as $G(x) := \mathbb{R}_+$ for every $x \in \mathbb{R}$. At the reference point $(\bar{x}, \bar{y}) = (0, 0)$, we would have $G^{-1}(\bar{y}) = \{x \mid 0 \in G(x)\} = \mathbb{R}$. Then, $d(x, G^{-1}(\bar{y})) = d(\bar{y}, G(x)) = 0$ holds for every $x \in \mathbb{R}$, showing G is metrically sub-regular at $(0, 0)$.

On the other hand, let $\varphi : \mathbb{R} \rightarrow \mathbb{R}_+$ denote any convex function differentiable at 0 fulfilling $\varphi(0) = \varphi'(0) = 0$, and $\varphi(t) > 0$ for every $t \neq 0$. For arbitrarily fixed $\epsilon > 0$, let $\tilde{G}(x) := G(x) + \epsilon\varphi(x)$. Then \tilde{G} is still a convex multifunction, but it is no longer metrically sub-regular at $(0, 0)$. In fact, $\tilde{G}^{-1}(\bar{y}) = \{0\}$, and $d(x, \tilde{G}^{-1}(\bar{y})) = d(x, 0) = |x|$. Also we can compute $d(\bar{y}, \tilde{G}(x)) = \epsilon\varphi(x)$. Since

$$\lim_{x \rightarrow 0} \frac{|x|}{\varphi(x)} = \infty,$$

the metric sub-regularity condition (1.35) does not hold for any finite κ .

Of course we can simply take $\varphi(x) = x^2$, but as a more subtle choice for the function φ we can take the function $\varphi \in C^\infty(\mathbb{R})$ defined by $\varphi(0) = \varphi'(0) = \varphi''(0) = 0$, $\varphi''(x) = e^{-x^{-2}}$ for every $x \neq 0$, which has the property that all derivatives vanish at 0, that is, $\varphi^{(i)}(0) = 0 \forall i \in \mathbb{N}$.

Then, any “derivative-like” tool offered for checking metric sub-regularity would be desired to have the property that the derivatives of the multifunctions G and \tilde{G} are different (so the metric sub-regularity of G and the lack of this property for \tilde{G} could be detectable). But G and \tilde{G} differ only by a C^∞ -function, where all derivatives vanish at \bar{x} and so the usual calculus rules (that is, the derivative of the sum being equal to the sum of derivatives) cannot be valid.

Definition 1.3.16. (Strong Metric Sub-Regularity)

[29, p. 202] A mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is said to be *strongly metrically sub-regular* at \bar{x} for \bar{y} if $(\bar{x}, \bar{y}) \in \text{gph } F$ and there is a constant $\kappa \geq 0$ along with neighborhoods U of \bar{x} and V of \bar{y} such that

$$\|x - \bar{x}\| \leq \kappa d(\bar{y}, F(x) \cap V) \quad \text{for all } x \in U^8. \quad (1.37)$$

⁸ The strong metric sub-regularity property was first considered without giving it a name, by Rockafellar [64]. The name “strong metric sub-regularity” was first used in [27] where its equivalence with the isolated calmness was proved.

Clearly, the infimum of κ for which (1.37) holds is equal to $\text{subreg}(F; \bar{x}|\bar{y})$.

Theorem 1.3.17 (Characterization by Isolated Calmness of the Inverse Map).

[29, p. 202] *A mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is strongly metrically sub-regular at \bar{x} for \bar{y} if and only if its inverse F^{-1} has the isolated calmness property at \bar{y} for \bar{x} .*

Specifically, for any $\kappa > \text{subreg}(F; \bar{x}|\bar{y})$ there exist neighborhoods U of \bar{x} and V of \bar{y} such that

$$F^{-1}(y) \cap U \subset \bar{x} + \kappa \|y - \bar{y}\| \mathbb{B} \quad \text{when } y \in V. \quad (1.38)$$

Moreover, the infimum of all κ such that the inclusion (1.38) holds for some neighborhoods U and V actually equals $\text{subreg}(F; \bar{x}|\bar{y})$.

Proof. Assume first that F is strongly metrically sub-regular at \bar{x} for \bar{y} . Let $\kappa > \text{subreg}(F; \bar{x}|\bar{y})$. Then there are neighborhoods U for \bar{x} and V for \bar{y} such that (1.37) holds with the indicated κ . Consider any $y \in V$. If $F^{-1}(y) \cap U = \emptyset$, then (1.38) holds trivially. If not, let $x \in F^{-1}(y) \cap U$. This entails $y \in F(x) \cap V$, hence $d(\bar{y}, F(x) \cap V) \leq \|y - \bar{y}\|$ and consequently $\|x - \bar{x}\| \leq \kappa \|y - \bar{y}\|$ by (1.37). Thus, $x \in \bar{x} + \kappa \|y - \bar{y}\| \mathbb{B}$, and we conclude that (1.38) holds. Also, we see that $\text{subreg}(F; \bar{x}|\bar{y})$ is not less than the infimum of all κ such that (1.38) holds for some choice of U and V .

For the converse, suppose (1.38) holds for some κ and neighborhoods U and V . Consider any $x \in U$. If $F(x) \cap V = \emptyset$ the right side of (1.37) is ∞ and there is nothing more to prove. If not, for an arbitrary $y \in F(x) \cap V$ we have $x \in F^{-1}(y) \cap U$, and therefore $x \in \bar{x} + \kappa \|y - \bar{y}\| \mathbb{B}$ by (1.38), which means $\|x - \bar{x}\| \leq \kappa \|y - \bar{y}\|$. This being true for all $y \in F(x) \cap V$, we must have $\|x - \bar{x}\| \leq \kappa d(\bar{y}, F(x) \cap V)$. Thus, (1.37) holds, and in particular we have $\kappa \geq \text{subreg}(F; \bar{x}|\bar{y})$. Therefore, the infimum of κ in (1.38) equals $\text{subreg}(F; \bar{x}|\bar{y})$. \square

The relation of local stability properties in Section 1.2 and the metric regularities⁹ of this section, is summarized in the following chart.

for $F :$	SMR	\subset	MR	\subset	MSR	\supset	$SMSR$
			\cong		\cong		\cong
for $F^{-1} :$			$Aubin\ Property$	\subset	$Calmness$	\supset	$Isolated\ Calmness$

⁹ Referring to all four definitions of metric regularity (shown with MR in the figure, cf. Definition 1.3.1), strong metric regularity (abbreviated as SMR, cf. Definition 1.3.9), metric sub-regularity (abbreviated as MSR, cf. Definition 1.3.13), and strong metric sub-regularity (abbreviated as SMSR, cf. Definition 1.3.16).

1.4 A Review on Perturbation Results

This section will contain results about perturbing a set-valued map that has one of the stability-like properties mentioned in previous sections, with a single-valued function. The general idea is to provide conditions under which, the sum of two maps still has the stability-like property. In Chapter 2, we will give a physical interpretation to this problem.

In order to study the perturbation problem, we need to somehow approximate the non-smooth function and replace it with another smooth function which is close enough to it. This would be the essence of the following three definitions.

Definition 1.4.1. (First-Order Approximations of Functions)

[29, p. 39] Consider a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and a point $\bar{x} \in \text{int dom } f$. A function $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ with $\bar{x} \in \text{int dom } h$ is a *first-order approximation* to f at \bar{x} if $h(\bar{x}) = f(\bar{x})$ and

$$\text{clm}(e; \bar{x}) = 0 \quad \text{for } e(x) = f(x) - h(x),$$

which can also be written as $f(x) = h(x) + o(\|x - \bar{x}\|)$.

It is a *strict first-order approximation* if the stronger condition holds that

$$\text{lip}(e; \bar{x}) = 0 \quad \text{for } e(x) = f(x) - h(x).$$

Example 1.4.2. When the function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is differentiable around a point $\bar{x} \in \text{int dom } f$, one can define a first-order approximation as $h(x) = f(\bar{x}) + \nabla f(\bar{x})(x - \bar{x})$. First, observe that $h(\bar{x}) = f(\bar{x})$. Then, for every $\epsilon > 0$ one can find $\delta > 0$ such that

$$\|f(x) - h(x)\| = \|f(x) - f(\bar{x}) - \nabla f(\bar{x})(x - \bar{x})\| \leq \epsilon \|x - \bar{x}\| \quad \text{for every } x \in \mathbb{B}_\delta(\bar{x}).$$

Definition 1.4.3. (Estimators)

[29, p. 41] Consider a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ and a point $\bar{x} \in \text{int dom } f$. A function $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ with $\bar{x} \in \text{int dom } h$ is an *estimator* of f at \bar{x} with constant μ if $h(\bar{x}) = f(\bar{x})$ and

$$\text{clm}(e; \bar{x}) \leq \mu < \infty \quad \text{for } e(x) = f(x) - h(x),$$

which can also be written as $\|f(x) - h(x)\| \leq \mu \|x - \bar{x}\| + o(\|x - \bar{x}\|)$.

It is a *strict estimator* if the stronger condition holds that

$$\text{lip}(e; \bar{x}) \leq \mu < \infty \quad \text{for } e(x) = f(x) - h(x).$$

Note 1.4.4. (a) In this terminology, a first-order approximation is simply an estimator with constant $\mu = 0$. Through that, any result involving estimators can immediately be

specialized to a result about first-order approximations.

(b) Although one can consider the idea of using estimators as a substitute for differentiability in some sense, estimators can be of interest even when differentiability is present. For instance, in the case of a function f that is strictly differentiable at \bar{x} a strict estimator of f at \bar{x} with constant μ is furnished by $h(x) = f(\bar{x}) + A(x - \bar{x})$ for any matrix A with $\|\nabla f(\bar{x}) - A\| \leq \mu$ ¹⁰. We will use this idea in Subsection 3.5.1.

When dealing with multivariable functions, it is possible to consider one variable as a parameter (which is actually the case in Chapter 4), and thus, we need to find the approximation with respect to one variable, while the others are fixed. To do so, we introduce *partial* estimators and approximations in the following way.

Definition 1.4.5. (Partial First-Order Estimators and Approximations)

[29, p. 49] For $f : \mathbb{R}^d \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ and a point $(\bar{p}, \bar{x}) \in \text{int dom } f$, a function $h : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be an *estimator* of f with respect to x uniformly in p at (\bar{p}, \bar{x}) with constant μ if $h(\bar{x}) = f(\bar{x}, \bar{p})$ and

$$\widehat{\text{clm}}_x(e; (\bar{p}, \bar{x})) \leq \mu < \infty \quad \text{for } e(p, x) = f(p, x) - h(x).$$

It is a *strict estimator* in this sense if the stronger condition holds that

$$\widehat{\text{lip}}_x(e; (\bar{p}, \bar{x})) \leq \mu < \infty \quad \text{for } e(p, x) = f(p, x) - h(x).$$

In which, $\widehat{\text{clm}}_x$ is the uniform partial calmness modulus¹¹ and similarly, $\widehat{\text{lip}}_x$ is the uniform partial Lipschitz modulus.

In the case of $\mu = 0$, such an estimator is called a *partial first-order approximation*.

The first theorem of this section is about perturbing a strongly metrically regular set-valued map with a Lipschitz function. We would see under suitable conditions for regularity modulus and Lipschitz constant, the new map would remain strongly metrically

¹⁰ Throughout this thesis, we consider matrices as linear operators and thus, $\|A\|$ is the operator norm of A induced by the Euclidean norm,

$$\|A\| = \max_{\|x\| \leq 1} \|Ax\|.$$

¹¹ (**Notation Review: Partial Calmness**) A function $f : \mathbb{R}^d \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to be calm with respect to x at $(\bar{p}, \bar{x}) \in \text{dom } f$ when the function Φ with values $\Phi(x) = f(\bar{p}, x)$ is calm at \bar{x} . Such calmness is said to be uniform in p at (\bar{p}, \bar{x}) when there exists a constant $\kappa \geq 0$ and neighborhoods Q of \bar{p} and U of \bar{x} such that actually

$$\|f(p, x) - f(p, \bar{x})\| \leq \kappa \|x - \bar{x}\| \quad \text{for all } (p, x) \in (Q \times U) \cap \text{dom } f.$$

Correspondingly, the partial calmness modulus of f with respect to x at (\bar{p}, \bar{x}) is denoted as $\text{clm}_x(f; (\bar{p}, \bar{x}))$, while the uniform partial calmness modulus is

$$\widehat{\text{clm}}_x(f; (\bar{p}, \bar{x})) := \limsup_{\substack{x \rightarrow \bar{x}, p \rightarrow \bar{p} \\ (p, x) \in \text{dom } f, x \neq \bar{x}}} \frac{\|f(p, x) - f(p, \bar{x})\|}{\|x - \bar{x}\|}$$

regular. In order to state and prove this theorem, we use the idea presented in Proposition 1.3.11, and go in two steps.

First, we would prove the single-valuedness (non-multivaluedness, to be more precise) of the new map in the following proposition. Next, we would prove the metric regularity of the new map in Theorem 1.4.7. Then, we would sum up everything in Theorem 1.4.9.

Proposition 1.4.6 (Stability of Single-Valuedness Under Perturbation).

[29, p. 193] *Let ν and λ be positive constants such that $\nu\lambda < 1$. Consider a mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a pair $(\bar{x}, \bar{y}) \in \text{gph } F$, such that F^{-1} has a Lipschitz continuous single-valued localization s around \bar{y} for \bar{x} with $\text{lip}(s; \bar{y}) < \lambda$. Consider also a function $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ with $\bar{x} \in \text{int dom } g$ and such that $\text{lip}(g; \bar{x}) < \nu$.*

Then the mapping $(g + F)^{-1}$ has a localization around $g(\bar{x}) + \bar{y}$ for \bar{x} which is nowhere multivalued.

Theorem 1.4.7 (Inverse Mapping Theorem with Metric Regularity).

[29, p. 184] *Consider a mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, a point $(\bar{x}, \bar{y}) \in \text{gph } F$, and a function $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ with $\bar{x} \in \text{int dom } g$. Let κ and μ be nonnegative constants such that*

$$\kappa\mu < 1, \text{reg}(F; \bar{x}|\bar{y}) \leq \kappa \text{ and } \text{lip}(g; \bar{x}) \leq \mu.$$

Then

$$\text{reg}(g + F; \bar{x}|g(\bar{x}) + \bar{y}) \leq \frac{\kappa}{1 - \kappa\mu}.$$

Note 1.4.8. Although formally there is no inversion of a mapping in Theorem 1.4.7, if this result is stated equivalently in terms of the Aubin property of the inverse mapping F^{-1} , it fits then into the pattern of the inverse function theorem paradigm ¹².

The theorem can also be viewed as a result concerning stability of metric regularity under perturbations by functions with small Lipschitz constants.

Theorem 1.4.9 (Inverse Function Theorem with Strong Metric Regularity).

[29, p. 194] *Let κ and μ be nonnegative constants such that $\kappa\mu < 1$. Consider a mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and any $(\bar{x}, \bar{y}) \in \text{gph } F$ such that F is strongly metrically regular at \bar{x} for \bar{y} with $\text{reg}(F; \bar{x}|\bar{y}) \leq \kappa$ and a function $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ with $\bar{x} \in \text{int dom } g$ and $\text{lip}(g; \bar{x}) \leq \mu$.*

Then the mapping $g + F$ is strongly metrically regular at \bar{x} for $g(\bar{x}) + \bar{y}$. Moreover,

$$\text{reg}(g + F; \bar{x}|g(\bar{x}) + \bar{y}) \leq \frac{\kappa}{1 - \kappa\mu}.$$

¹² We can actually deduce the classical inverse function Theorem (1.1.20) from Theorem 1.4.7. Indeed, let $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be a smooth function around \bar{x} and let $\nabla f(\bar{x})$ be non-singular.

Then $F(x) := Df(\bar{x})(x - \bar{x})$ is metrically regular everywhere and specially at \bar{x} for 0. Moreover, the function $g(x) := f(x) - Df(\bar{x})(x - \bar{x})$ is Lipschitz continuous at \bar{x} with modulus $\text{lip}(g; \bar{x}) = 0$.

Thus, from Theorem 1.4.7, we obtain that $g + F = f$ is metrically regular at \bar{x} for $f(\bar{x})$ with modulus $\text{reg}(f; \bar{x}) = \text{reg}(F; \bar{x}|0)$. But then f must be open (cf. Theorem 1.3.7; in fact, f is open at linear rate). So, the inverse map has a localization around $f(\bar{x})$ which is single-valued and Lipschitz continuous.

Proof. We try to prove the equivalent statement (that is (b) in Proposition 1.3.11), so we have to show

- (1) $g + F$ is strongly metrically regular at \bar{x} for $\bar{y} + g(\bar{x})$; and
- (2) $(g + F)^{-1}$ has a localization at $\bar{y} + g(\bar{x})$ for \bar{x} that is nowhere multivalued.

Our hypothesis that F is strongly metrically regular at \bar{x} for \bar{y} implies that a graphical localization of F^{-1} around (\bar{y}, \bar{x}) is single-valued near \bar{y} . Further, by fixing $\lambda > \kappa$ such that $\lambda\mu < 1$ and using Proposition 1.3.11, we can get neighborhoods U of \bar{x} and V of \bar{y} such that for every $y \in V$ the set $F^{-1}(y) \cap U$ consists of exactly one point, which we may denote by $s(y)$ and know that the function $s : y \mapsto F^{-1}(y) \cap U$ is Lipschitz continuous on V with Lipschitz constant λ . Let $\mu < \nu < \lambda^{-1}$ and choose a neighborhood $U' \subset U$ of \bar{x} on which g is Lipschitz continuous with constant ν . Applying Proposition 1.4.6, we obtain that the mapping $(g + F)^{-1}$ has a localization around $g(\bar{x}) + \bar{y}$ for \bar{x} which is nowhere multivalued.

On the other hand, we know from Theorem 1.4.7 that for such g the mapping $g + F$ is metrically regular at $g(\bar{x}) + \bar{y}$ for \bar{x} . Applying Proposition 1.3.11 once more, we get that $(g + F)^{-1}$ has a Lipschitz single-valued localization s around $g(\bar{x}) + \bar{y}$ for \bar{x} with modulus

$$\text{lip}(s; \bar{y} + g(\bar{x})) = \text{reg}(g + F; \bar{x} | g(\bar{x}) + \bar{y}) < \frac{\lambda}{1 - \lambda\nu}$$

As λ and ν could be arbitrarily close to κ and μ , respectively, the modulus criteria is satisfied. \square

Now, we will consider the situation that working with the function f in the sum $f + F$ is not easy for some reasons. So, we would like to use an approximation function h , instead. It would be important to know whether the regularity properties obtained for $h + F$, contain any information about the regularity of $f + F$ or not.

The following proposition is needed to prove the theorem afterwards, which will be one of the important tools in order to give a meaning to the usage and conditions of the auxiliary map G_t in Chapter 4. First, we will prove the parametric version of this result, and then express the “inverse function theorem” version in Theorem 1.4.12.

Proposition 1.4.10 (Contraction Mapping Principle for Composition).

[29, p. 86] Consider a function $\varphi : \mathbb{R}^d \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ and a point $(\bar{p}, \bar{x}) \in \text{int dom } \varphi$ and let the scalars $\nu \geq 0, b \geq 0, a > 0$, and the set $Q \subset \mathbb{R}^d$ be such that $\bar{p} \in Q$ and

$$\begin{cases} \|\varphi(p, x') - \varphi(p, x)\| \leq \nu \|x - x'\| & \text{for all } x', x \in \mathbb{B}_a(\bar{x}) \text{ and } p \in Q, \\ \|\varphi(p, \bar{x}) - \varphi(\bar{p}, \bar{x})\| \leq b & \text{for all } p \in Q \end{cases} \quad (1.39)$$

Consider also a set-valued mapping $M : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ with $(\bar{y}, \bar{x}) \in \text{gph } M$ in which $\bar{y} := \varphi(\bar{p}, \bar{x})$, such that for each $y \in \mathbb{B}_{\nu a + b}(\bar{y})$ the set $M(y) \cap \mathbb{B}_a(\bar{x})$ consists of exactly one

point, denoted by $r(y)$, and suppose that the function

$$r : y \mapsto M(y) \cap \mathbb{B}_a(\bar{x}) \quad \text{for } y \in \mathbb{B}_{\nu a+b}(\bar{y}) \quad (1.40)$$

is Lipschitz continuous on $\mathbb{B}_{\nu a+b}(\bar{y})$ with a Lipschitz constant λ . In addition, suppose that the following relations hold

- (a) $\lambda\nu < 1$;
- (b) $\lambda\nu a + \lambda b \leq a$.

Then for each $p \in Q$ the set $\{x \in \mathbb{B}_a(\bar{x}) \mid x \in M(\varphi(p, x))\}$ consists of exactly one point, and the associated function

$$s : p \mapsto \{x \mid x = M(\varphi(p, x)) \cap \mathbb{B}_a(\bar{x})\} \quad \text{for } p \in Q \quad (1.41)$$

satisfies

$$\|s(p') - s(p)\| \leq \frac{\lambda}{1 - \lambda\nu} \|\varphi(p', s(p)) - \varphi(p, s(p))\| \quad \text{for all } p', p \in Q. \quad (1.42)$$

Theorem 1.4.11 (Robinson Theorem Extended Beyond Differentiability).

[29, p. 86] For $f : \mathbb{R}^d \times \mathbb{R}^n \rightarrow \mathbb{R}^m$ and $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, consider a generalized equation $f(p, x) + F(x) \ni 0$ with the solution mapping defined as $S : t \mapsto \{x \mid f(p, x) + F(x) \ni 0\}$.

Let \bar{p} and \bar{x} be such that $\bar{x} \in S(\bar{p})$. Assume that:

- (a) $f(\cdot, \bar{x})$ is continuous at \bar{p} , and h is a strict estimator of f with respect to x uniformly in p at (\bar{p}, \bar{x}) with constant μ ;
- (b) the inverse G^{-1} of the mapping $G = h + F$, for which $G(\bar{x}) \ni 0$, has a Lipschitz continuous single-valued localization σ around 0 for \bar{x} with $\text{lip}(\sigma; 0) \leq \kappa$ for a constant κ such that $\kappa\mu < 1$.

Then S has a single-valued localization s around \bar{p} for \bar{x} which is continuous at \bar{p} , and moreover for every $\epsilon > 0$ there is a neighborhood Q of \bar{p} such that

$$\|s(p') - s(p)\| \leq \frac{\kappa + \epsilon}{1 - \kappa\mu} \|f(p', s(p)) - f(p, s(p))\| \quad \text{for all } p', p \in Q. \quad (1.43)$$

Proof. For an arbitrary $\epsilon > 0$, choose any $\lambda > \text{lip}(\sigma; 0)$ and $\nu > \mu$ such that $\lambda\nu < 1$ and

$$\frac{\lambda}{1 - \lambda\nu} \leq \frac{\kappa + \epsilon}{1 - \kappa\mu},$$

as is possible under the assumption that $\kappa\mu < 1$. Let a, b and c be positive numbers such that

$$\begin{aligned} \|\sigma(y) - \sigma(y')\| &\leq \lambda \|y - y'\| && \text{for } y, y' \in \mathbb{B}_{\nu a+b}(0), \\ \|e(p, x') - e(p, x)\| &\leq \nu \|x - x'\| && \text{for } x, x' \in \mathbb{B}_a(\bar{x}), p \in \mathbb{B}_c(\bar{p}), \end{aligned} \quad (1.44)$$

where $e(p, x) = f(p, x) - h(x)$, the first inequality is guaranteed by Lipschitz continuity of σ , and the second inequality comes from the definition of estimator.

Continuity of $f(\cdot, \bar{x})$ at \bar{p} implies that

$$\|f(p, \bar{x}) - f(\bar{p}, \bar{x})\| \leq b \quad \text{for } p \in \mathbb{B}_c(\bar{p}). \quad (1.45)$$

Take b smaller if necessary so that $b\lambda < a(1 - \lambda\nu)$, and accordingly adjust c to ensure having the last inequality.

Now apply Proposition 1.4.10 with $\varphi = -e$, $Q = \mathbb{B}_c(\bar{p})$, $M = G^{-1} = (h + F)^{-1}$, $\bar{y} = 0$, and $r = \sigma$. The first condition in (1.39) would be the second statement in (1.44). The second condition in (1.39) will be

$$\|e(p, \bar{x}) - e(\bar{p}, \bar{x})\| = \|f(p, \bar{x}) - h(\bar{x}) - f(\bar{p}, \bar{x}) + h(\bar{x})\| \leq b$$

which holds because of (1.45).

Extra conditions of (a) and (b) of the proposition hold true trivially. Thus, one obtains that for any $p \in \mathbb{B}_c(\bar{p})$, the map

$$p \mapsto s^*(p) := \{x \mid x = [(h + F)^{-1}(-e(p, x))] \cap \mathbb{B}_a(\bar{x})\}$$

is single-valued with the estimate mentioned in (1.42). The only point to check out is that whether s^* is related to the solution mapping S or not, that is, if we have $s^*(p) \in S(p)$ for any $p \in \mathbb{B}_c(\bar{p})$.

Let $x = s^*(p)$. Then $x = (h + F)^{-1}(-e(p, x)) \cap \mathbb{B}_a(\bar{x})$.

$$\begin{aligned} x \in (h + F)^{-1}(-e(p, x)) &\Leftrightarrow -e(p, x) \in (h + F)(x) \Leftrightarrow h(x) - f(p, x) \in h(x) + F(x) \\ &\Leftrightarrow 0 \in f(p, x) + F(x) \Leftrightarrow x \in S(p) \end{aligned}$$

So, s^* is a single-valued localization for S around \bar{p} for \bar{x} . The estimate in (1.42) also implies the continuity of s^* at \bar{p} , since $f(\cdot, \bar{x})$ is assumed to be continuous at \bar{p} by assumption (a). \square

The inverse function version of Theorem 1.4.11 has the following simpler form. Let us note that in view of Definition 1.3.9, this theorem is nothing more than a translation of Theorem 1.4.9 from the strong metric regularity term to the Lipschitz continuity (for the localized inverse map).

Theorem 1.4.12 (Inverse Function Theorem for Set-Valued Mappings).

[29, p. 89] Consider a mapping $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ with $G(\bar{x}) \ni \bar{y}$ and suppose that G^{-1} has a Lipschitz continuous single-valued localization σ around \bar{y} for \bar{x} with $\text{lip}(\sigma; \bar{y}) \leq \kappa$ for a constant κ . Let $g : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be Lipschitz continuous around \bar{x} with Lipschitz constant μ such that $\kappa\mu < 1$.

Then the mapping $(g+G)^{-1}$ has a Lipschitz continuous single-valued localization around $\bar{y} + g(\bar{x})$ for \bar{x} with Lipschitz constant $\kappa/(1 - \kappa\mu)$.

The last Theorem of this section would discuss the perturbation effect on metric regularity and strong metric regularity in terms of the graph of the new map. It would be an important tool in our study and we may refer to it several times in Chapter 4.

We need the following statement to prove the theorem afterwards.

Theorem 1.4.13 (Contraction Mapping Principle for Set-Valued Mappings).

[29, p. 313] Let X be a complete metric space with metric ρ , and consider a set-valued mapping $\Phi : X \rightrightarrows X$ and a point $\bar{x} \in X$. Suppose that there exist scalars $a > 0$ and $\lambda \in (0, 1)$ such that the set $\text{gph } \Phi \cap (\mathbb{B}_a(\bar{x}) \times \mathbb{B}_a(\bar{x}))$ is closed and

- (a) $d(\bar{x}, \Phi(\bar{x})) < a(1 - \lambda)$;
- (b) $e(\Phi(u) \cap \mathbb{B}_a(\bar{x}), \Phi(v)) \leq \lambda \rho(u, v)$ for all $u, v \in \mathbb{B}_a(\bar{x})$.

Then Φ has a fixed point in $\mathbb{B}_a(\bar{x})$; that is, there exists $x \in \mathbb{B}_a(\bar{x})$ such that $x \in \Phi(x)$.

Theorem 1.4.14 (Perturbed [Strong] Metric Regularity). [29, p. 325]

Let X, Y be Banach spaces. Consider a mapping $F : X \rightrightarrows Y$ and a point $(\bar{x}, \bar{y}) \in \text{gph } F$ at which F is metrically regular, [that is, there exist positive constants a, b , and a nonnegative κ such that the set $\text{gph } F \cap (\mathbb{B}_a(\bar{x}) \times \mathbb{B}_b(\bar{y}))$ is closed and

$$d(x, F^{-1}(y)) \leq \kappa d(y, F(x)) \quad \text{for all } (x, y) \in \mathbb{B}_a(\bar{x}) \times \mathbb{B}_b(\bar{y}).]$$

Let $\mu > 0$ be such that $\kappa\mu < 1$ and let $\kappa' > \kappa/(1 - \kappa\mu)$.

Then for every positive α and β such that

$$\alpha \leq a/2, \quad 2\mu\alpha + 2\beta \leq b, \quad \text{and} \quad 2\kappa'\beta \leq \alpha \tag{1.46}$$

and for every function $g : X \rightarrow Y$ satisfying

$$\|g(\bar{x})\| \leq \beta \tag{1.47}$$

and

$$\|g(x) - g(x')\| \leq \mu \|x - x'\| \quad \text{for every } x, x' \in \mathbb{B}_{2\alpha}(\bar{x}), \tag{1.48}$$

the mapping $g + F$ has the following property:

for every $y, y' \in \mathbb{B}_\beta(\bar{y})$ and every $x \in (g + F)^{-1}(y) \cap \mathbb{B}_\alpha(\bar{x})$ there exists $x' \in (g + F)^{-1}(y')$ such that

$$\|x - x'\| \leq \kappa' \|y - y'\|. \tag{1.49}$$

In addition, if the mapping F is strongly metrically regular at \bar{x} for \bar{y} ; [that is, the mapping $y \mapsto F^{-1}(y) \cap \mathbb{B}_a(\bar{x})$ is single-valued and Lipschitz continuous on $\mathbb{B}_b(\bar{y})$ with a Lipschitz constant κ], then for μ, κ', α , and β as above and any function g satisfying (1.47) and (1.48), the mapping $y \mapsto (g + F)^{-1}(y) \cap \mathbb{B}_\alpha(\bar{x})$ is a Lipschitz continuous function on $\mathbb{B}_\beta(\bar{y})$ with a Lipschitz constant κ' .

Proof. We consider two different cases.

Case 1. *with metric regularity;*

Choose μ and κ' as required and then α and β to satisfy (1.46). For any $x \in \mathbb{B}_{2\alpha}(\bar{x})$ and $y \in \mathbb{B}_\beta(\bar{y})$, using (1.47), (1.48) and the triangle inequality, we obtain

$$\begin{aligned} \|-g(x) + y - \bar{y}\| &\leq \|g(\bar{x})\| + \|g(\bar{x}) - g(x)\| + \|y - \bar{y}\| \\ &\leq \beta + \mu \|x - \bar{x}\| + \beta \leq 2\beta + 2\mu\alpha \leq b, \end{aligned} \tag{1.50}$$

where the last inequality follows from the second inequality in (1.46). Fix $y' \in \mathbb{B}_\beta(\bar{y})$ and consider the mapping

$$\mathbb{B}_\alpha(\bar{x}) \ni x \mapsto \Phi_{y'}(x) := F^{-1}(-g(x) + y').$$

Let $y \in \mathbb{B}_\beta(\bar{y})$, $y \neq y'$ and let $x \in (g + F)^{-1}(y) \cap \mathbb{B}_\alpha(\bar{x})$. We will apply Theorem (1.4.13) with the complete metric space X identified with the closed ball $\mathbb{B}_\alpha(\bar{x})$ to show that there is a fixed point $x' \in \Phi_{y'}(x')$ in the closed ball centered at x with radius

$$r := \kappa' \|y - y'\| \tag{1.51}$$

From the third inequality in (1.46), we obtain $r \leq \kappa'(2\beta) \leq \alpha$. Hence, from the first inequality in (1.46) we get $\mathbb{B}_r(x) \subset \mathbb{B}_\alpha(\bar{x})$.

Let $(x_n, z_n) \in \text{gph } \Phi'_y \cap (\mathbb{B}_r(x) \times \mathbb{B}_r(x))$ and $(x_n, z_n) \rightarrow (\tilde{x}, \tilde{z})$.

From (1.50), $\|-g(x_n) + y' - \bar{y}\| \leq b$; also note that $\|z_n - \bar{x}\| \leq a$. Using closedness of $\text{gph } F \cap (\mathbb{B}_a(\bar{x}) \times \mathbb{B}_b(\bar{y}))$ (by metric regularity assumption) and passing to the limit we obtain that $(\tilde{x}, \tilde{z}) \in \text{gph } \Phi'_y \cap (\mathbb{B}_r(x) \times \mathbb{B}_r(x))$, hence this set is closed.

Since $y \in g(x) + F(x)$ and (x, y) satisfies (1.50), from the assumed metric regularity of F we have

$$\begin{aligned} d(x, \Phi_{y'}(x)) &= d(x, F^{-1}(-g(x) + y')) \leq \kappa d(-g(x) + y', F(x)) \\ &= \kappa d(y', g(x) + F(x)) \leq \kappa \|y - y'\| \\ &< \kappa' \|y - y'\| (1 - \kappa\mu) = r(1 - \kappa\mu). \end{aligned}$$

For any $u, v \in \mathbb{B}_r(x)$, using (1.48), we have

$$\begin{aligned} e\left(\Phi_{y'}(u) \cap \mathbb{B}_r(x), \Phi_{y'}(v)\right) &\leq e\left(F^{-1}(-g(u) + y') \cap \mathbb{B}_a(\bar{x}), F^{-1}(-g(v) + y')\right) \\ &\leq \kappa \|g(u) - g(v)\|, \leq \kappa\mu \|u - v\|. \end{aligned}$$

Applying Theorem 1.4.13 to the mapping $\Phi_{y'}$, with \bar{x} identified with x and constants $a = r$ and $\lambda = \kappa\mu$, we obtain the existence of a fixed point $x' \in \Phi_{y'}(x')$, which is equivalent to $x' \in (g + F)^{-1}(y')$, within distance r given by (1.51) from x . This proves (1.49).

Case 2. *with strong metric regularity;*

For the second part of the theorem, suppose that $y \mapsto s(y) := F^{-1}(y) \cap \mathbb{B}_a(\bar{x})$ is a Lipschitz continuous function on $\mathbb{B}_b(\bar{y})$ with a Lipschitz constant κ . Choose μ, κ', α and β as in the statement and let g satisfy (1.47) and (1.48).

For any $y \in \mathbb{B}_\beta(\bar{y})$, since $\bar{x} \in (g + F)^{-1}(\bar{y} + g(\bar{x})) \cap \mathbb{B}_\alpha(\bar{x})$, from (1.49) we obtain that there exists $x \in (g + F)^{-1}(y)$ such that

$$\|x - \bar{x}\| \leq \kappa' \|y - \bar{y} - g(\bar{x})\|.$$

Since $\|y - \bar{y} - g(\bar{x})\| \leq 2\beta$, by (1.46) we get $\|x - \bar{x}\| \leq \alpha$, i.e., $(g + F)^{-1}(y) \cap \mathbb{B}_\alpha(\bar{x}) \neq \emptyset$. Hence the domain of the mapping $(g + F)^{-1}(y) \cap \mathbb{B}_\alpha(\bar{x})$ contains $\mathbb{B}_\beta(\bar{y})$.

If $x \in (g + F)^{-1}(y) \cap \mathbb{B}_\alpha(\bar{x})$, then

$$x \in F^{-1}(y - g(x)) \cap \mathbb{B}_\alpha(\bar{x}) \subset F^{-1}(y - g(x)) \cap \mathbb{B}_\alpha(\bar{x}) = s(y - g(x))$$

since $y - g(x) \in \mathbb{B}_b(\bar{y})$ according to (1.50). Hence,

$$F^{-1}(y - g(x)) \cap \mathbb{B}_\alpha(\bar{x}) = s(y - g(x)) = x. \quad (1.52)$$

Let $y, y' \in \mathbb{B}_\beta(\bar{y})$, and define a map σ on $\mathbb{B}_\beta(\bar{y})$ with

$$y \mapsto \sigma(y) := (g + F)^{-1}(y) \cap \mathbb{B}_\alpha(\bar{x}).$$

Utilizing the equality $\sigma(y) = s(-g(\sigma(y)) + y)$ which comes from (1.52), we have

$$\begin{aligned} \|\sigma(y) - \sigma(y')\| &= \|s(-g(\sigma(y)) + y) - s(-g(\sigma(y')) + y')\| \\ &\leq \kappa \|g(\sigma(y)) - g(\sigma(y'))\| + \kappa \|y - y'\| \\ &\leq \kappa\mu \|\sigma(y) - \sigma(y')\| + \kappa \|y - y'\|. \end{aligned} \quad (1.53)$$

If $y = y'$, taking into account that $\kappa\mu < 1$ we obtain that $\sigma(y)$ must be equal to $\sigma(y')$. Hence, the mapping $y \mapsto \sigma(y) := (g + F)^{-1}(y) \cap \mathbb{B}_\alpha(\bar{x})$ is single-valued. From (1.53) and

(1.49) this function satisfies

$$\|\sigma(y) - \sigma(y')\| \leq \kappa' \|y - y'\|.$$

The proof is complete. □

Note 1.4.15. If $g(\bar{x}) \neq 0$ then (\bar{x}, \bar{y}) may be not in the graph of $g + F$ and we can not claim that $g + F$ is (strongly) metrically regular at \bar{x} for \bar{y} . This could be handled easily by choosing a new function \tilde{g} with $\tilde{g}(x) = g(x) - g(\bar{x})$.

Chapter 2

An Introduction to the Electrical Problem

Everything should be made as simple as possible, but not simpler.

*Albert Einstein, a.e.
(almost every-word!)*

In this chapter, we will provide a short review on the theory of electrical circuits, and a brief introduction to some electronic components (in Section 2.1). Then, in Section 2.2 a general form of generalized equations that would be considered in the rest of the thesis as an outcome of modelling process would be obtained.

Also, it provides physical explanation for the importance of studying the small perturbations of source signals. Another important role of this section is to provide a meaning for the stability-like properties introduced in Chapter 1: in the static case, the question about sensitivity of the circuit to small perturbations of the power source, is translated to the question whether the solution mapping of the obtained generalized equation has some stability-like properties or not.

This chapter ends with a review on the examples that we will discuss in depth in Chapter 3. Here, we just provide the mathematical model and general description of solutions.

2.1 Introduction to the Theory of Electrical Circuits

In this section, we will introduce the most basic concepts in the circuit theory in order to provide the appropriate language for formulation of the physical problem, which would be presented in the next section.

Mostly, the information about the electrical components and circuit theory are drawn from [23], while for the electronic details we refer to [68], and [48]. Although the way we present these physical concepts is a bit different from their approach and purpose, it would be a relaxing information to know where to refer in case of any ambiguity in the physical meaning, or for the sake of curiosity.

We hope that this short note legitimizes why we do not give the detail referencing in this chapter, unlike the rest of this thesis.

Circuits Without going into technical details, one can think of an electrical circuit (like the one shown in Figure 2.1) as a bunch of “components” connected together with wires. The special forms of connecting these components and the possible choice of different components in each combination, is due to the specific duty that each circuit is going to serve¹.

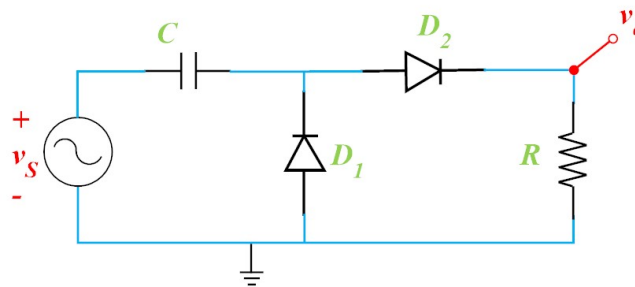


FIGURE 2.1: A simple electrical circuit

The word *component*, in this context, refers to different materials which exhibit a particular electrical behaviour under an electromagnetic force. The language of circuit theory has two variables:

- current, shown with I or i , and measured in amperes;
- voltage, indicated with V or v , and measured in volts².

Thus, for describing a component, we would look at the current passing through it, and the electric potential difference between its terminals, that is voltage over it. The behaviour of the component under different voltages dropped over its terminal, or equivalently, under various currents passing through it is usually described with a graph in

¹ In other words, one has a goal and then tries to design a circuit, that is a combination of components and how to connect them, that provides the desired aim. It is referred to as *circuit design*, in the literature. The other approach would be that given a particular circuit, one tries to use electrical rules and some tools to understand the goal of that circuit. It is referred to as *circuit analysis*, or *synthesis*, in the literature.

² The difference between small and capital letter symbols is that whether they indicate the quantities changing with time (known as *alternative current* in the literature, with the abbreviation AC) or constant quantities with respect to time (known as *direct current* in the literature, with the abbreviation DC), respectively.

the $i - v$ plane, and referred to as the $i - v$ characteristic of the component.

In order to avoid confusion, a subscript indicates which component in the circuit we are talking about (for example, v_s in the above circuit indicates the voltage of Source).

In circuit theory, we do not know nor are interested in the physical phenomenon causing such a behaviour, and consider a component as a black box that could be described with a map, that is, a relation between current and voltage.



FIGURE 2.2: General black box viewpoint of components

In the sequel, we will use “electrical” and “electronic” quite very often in referring to circuits, components, or different phrases. It is better to say few words about it now. An *electronic component* is usually made of semiconductors, which are materials that exhibit electrical behaviour somewhere between that of insulators and that of conductors³. Examples of semiconductors are silicon and germanium, and Diodes and BJT Transistors are examples of electronic components.

An *electronic circuit*, is simply a circuit with at least one electronic component. So, if one wants to be with mathematical precision in these engineering concepts, he/she can say the set of electronic circuits is a subset of electrical circuits.

Components We will now review those fundamental components of circuits that would be used in different examples of this thesis.

Resistor

In the following figure you can see the schematic, and $i - v$ characteristic of the most common two-terminal electrical component, which has a linear map (i. e. $v_R = Ri_R$).

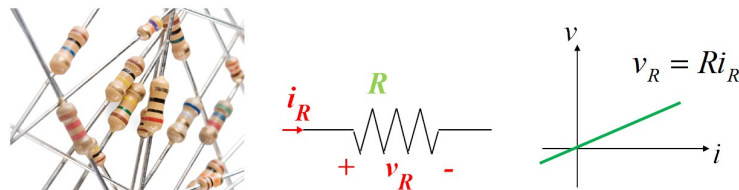


FIGURE 2.3: Circuit components: Resistor

³ Conductors present very low resistance to the flow of current, whereas insulators conduct very little current even when a large potential difference is applied. A semiconductor exhibits intermediate conductivity because it has more available charge carriers than an insulator but fewer than a conductor.

In electronic circuits, resistors are used to reduce current flow, adjust signal levels, to divide voltages, to bias active elements⁴, and to model a typical consumer, among other uses.

Capacitor

A capacitor is a passive two-terminal electrical component that stores electrical energy in an electric field. *Capacitance* is defined as the ratio of the electric charge Q on each conductor to the potential difference V between them. Unlike a resistor, an ideal capacitor does not dissipate energy.

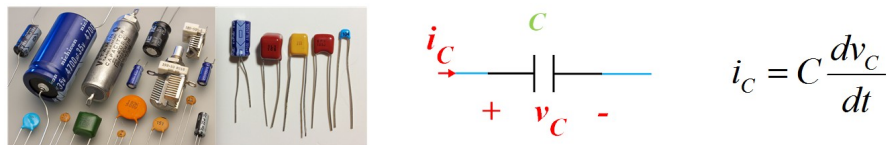


FIGURE 2.4: Circuit components: Capacitor

Capacitors are widely used in electronic circuits for blocking direct current while allowing alternating current to pass. In electric power transmission systems, they stabilize voltage and power flow. The property of energy storage in capacitors was exploited as dynamic memory in early digital computers.

Inductor

Inductors are components that store electrical energy in a magnetic field when electric current is flowing through it. An inductor typically consists of an electric conductor, such as a wire, that is wound into a coil, and it is characterized by its inductance, which is the ratio of the voltage to the rate of change of current. An “ideal inductor” has inductance, but no resistance or capacitance, and does not dissipate or radiate energy.

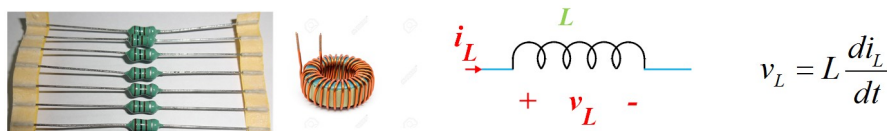


FIGURE 2.5: Circuit components: Inductor

Inductors are widely used in alternating current (AC) electronic equipment, particularly in radio equipment. They are used to block AC while allowing DC to pass; inductors designed for this purpose are called *chokes*. They are also used in electronic filters to separate signals of different frequencies, and in combination with capacitors to make tuned circuits, used to tune radio and TV receivers.

⁴ In electronics, *biasing* means establishing predetermined voltages or currents at various points of an electronic circuit in order to provide proper operating conditions in electronic components. Many electronic devices like transistors whose function is processing time-varying (AC) signals also require a steady (DC) current or voltage to operate correctly, a *bias*.

Voltage Source

A voltage source is a two terminal device which can maintain a fixed voltage drop across its terminals. An ideal voltage source can maintain the fixed voltage independent of the load resistance or the output current. However, a real-world voltage source cannot supply unlimited current.

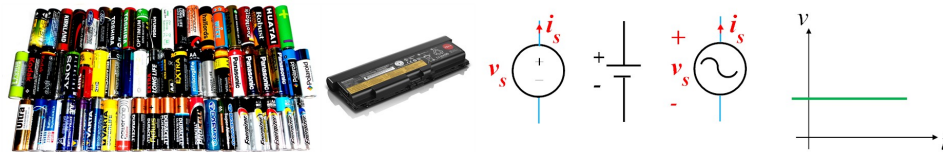


FIGURE 2.6: Circuit components: Voltage Source

Figure 2.6 shows the different schematics for DC and AC voltage sources. Real-world sources of electrical energy, such as batteries, generators, and power systems, can be modelled for analysis purposes as a combination of an ideal voltage source and additional combinations of impedance elements.

Most sources of electrical energy (mains electricity⁵, a battery, etc.) are best modelled as voltage sources, though in theory we can introduce current sources as components, too.

Current Source

A current source is an electronic circuit that delivers or absorbs an electric current which is independent of the voltage across it. A current source is the dual of a voltage source. The term, *constant-current sink*, is sometimes used for sources fed from a negative voltage supply.

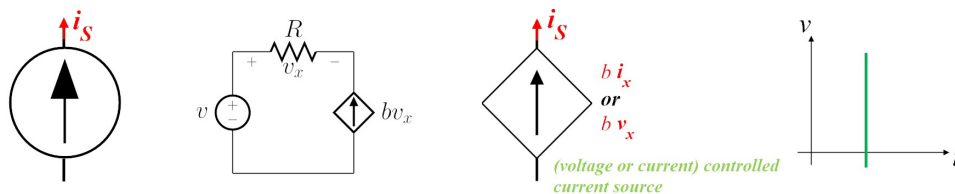


FIGURE 2.7: Circuit components: Current Source

If the current through an ideal current source can be specified independently of any other variable in a circuit, it is called an *independent* current source. Conversely, if the current through an ideal current source is determined by some other voltage or current in a circuit, it is called a *dependent* or *controlled current source*.

The current value could be controlled by the voltage over another component, or the

⁵ *Mains electricity* is the general-purpose alternating-current (AC) electric power supply. Mains electricity is the form of electrical power that is delivered to homes and businesses, and it is the form of electrical power that consumers use when they plug kitchen appliances, televisions and electric lamps into wall sockets.

current of another branch in the circuit. Soon, we would see a usage of current-controlled current sources in modelling of transistors.

Diode

A diode is an electrical device allowing current to move through it in one direction with far greater ease than in the other. Diode behaviour is analogous to the behaviour of a hydraulic device called *check valve* (a check valve allows fluid flow through it in only one direction, also known as one-way valve). The most common kind of diode in modern circuit design is the semiconductor diode, although other diode technologies exist.

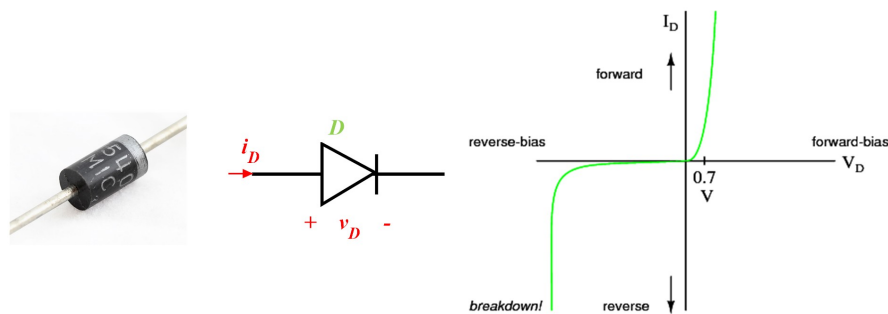


FIGURE 2.8: Circuit components: Diode

When placed in a simple battery-lamp circuit, the diode will either allow or prevent current through the lamp, depending on the polarity of the applied voltage (Figure 2.9). The essential difference between forward-bias and reverse-bias is the polarity of the voltage dropped across the diode.

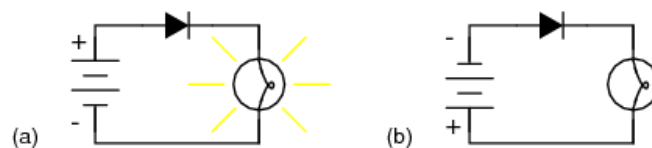


FIGURE 2.9: Diode operation: (a) Current flow is permitted; the diode is forward biased. (b) Current flow is prohibited; the diode is reversed biased.

Thus, if the “*ideal diode*” is reverse biased, the current flowing through it is zero. This ideal diode starts conducting at 0 volts and for any positive voltage an infinite current flows and the diode acts like a short circuit. The following figure (Figure 2.10), shows a simple circuit using this component with the $i - v$ characteristic of this diode and the two working models obtained from that behaviour.

The first step toward practical model for the diode could be obtained by adding a voltage source to an ideal diode in series, compensating the small voltage drop over the component. Next, one can consider a small resistance, again added in series to recover the small slope in the $i - v$ characteristic of Figure 2.8. The final step would

to hundreds of volts. This Zener voltage changes slightly with temperature, and like common carbon-composition resistor values, may be anywhere from 5% to 10% in error from the manufacturer's specifications.

LED

Some semiconductor junctions, composed of special chemical combinations, emit radiant energy within the spectrum of visible light as the electrons change energy levels. Simply put, these junctions glow when forward biased. A diode intentionally designed to glow like a lamp is called a *light-emitting diode*, or *LED*.

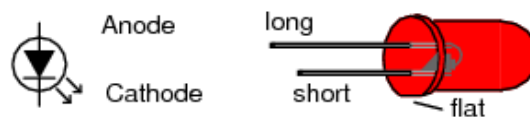


FIGURE 2.12: Circuit components: LED

Despite the different schematic symbol of LEDs shown in Figure 2.12, there is no other difference (from the mathematical point of view) in analysing a circuit with standard rectifying diodes or with an LED.

It is worth mentioning that LEDs are more sensitive to voltage change, and it is reflected in the light they emit. Thus, for the design problems one should try to stabilize the voltage drop over them (cf. Example 2.3.1).

Shockley Diode

Shockley diodes are four-layer *pnpn* diodes, which were one of the first semiconductor devices invented. The mathematical (and also physical) interesting point about these family is that their $i - v$ characteristic exhibits *hysteresis*, the property whereby a system fails to return to its original state after some cause of state change has been removed.

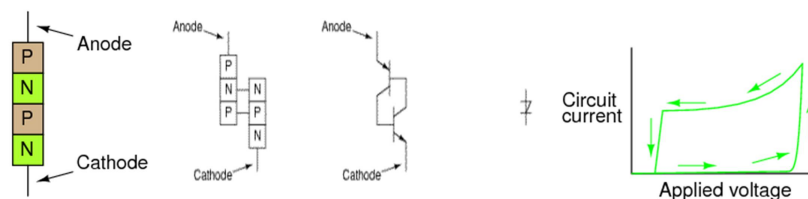


FIGURE 2.13: Circuit components: Shockley diode. from left to right: schematic, physical diagram, equivalent schematic, circuit symbol, and $i - v$ characteristic

Figure 2.13 also gives the idea that how does a Shockley diodes behaves as a pair of interconnected *pnp* and *npn* transistors. We only refer to two components with this physical structure:

DIAC

Like all diodes, Shockley diodes are unidirectional devices; that is, these only conduct current in one direction. If bidirectional (AC) operation is desired, two Shockley diodes may be joined in parallel facing different directions to form a DIAC. The term DIAC is an acronym of “diode for alternating current”.

When breakdown occurs, the diode enters a region of negative dynamic resistance, leading to a decrease in the voltage drop across the diode and, usually, a sharp increase in current through the diode. The diode remains in conduction until the current through it drops below a value characteristic for the device, called the holding current, I_H . Below this value, the diode switches back to its high-resistance, non-conducting state. This behaviour is bidirectional, meaning typically the same for both directions of current.

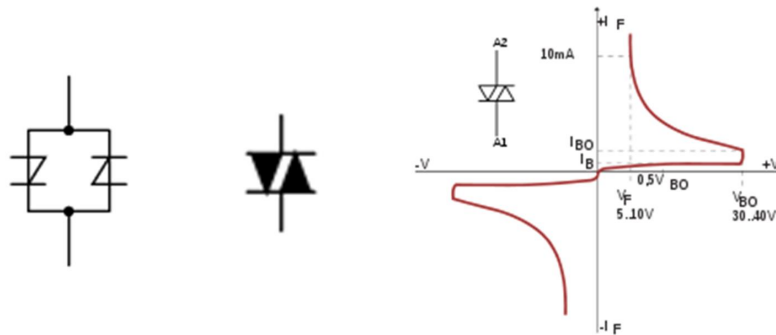


FIGURE 2.14: Circuit components: DIAC. from left to right: equivalent circuit, schematic symbol, and $i-v$ characteristic

DIACs are also called “symmetrical trigger diodes” due to the symmetry of their characteristic curve. Since DIACs are bidirectional devices, their terminals are not labelled as anode and cathode but as A1 and A2 or main terminal MT1 and MT2. DIACs are widely used in light dimmers, starter circuits for fluorescent lamps, and in conjunction with TRIACs to equalise their switching characteristics.

Silicon-Controlled Rectifier (SCR)

In order to expand the usefulness of Shockley diodes, one can equip them with another means of latching⁶. In doing so, each of the $n-p-n$ and $p-n-p$ junctions/transistors becomes true amplifying device, and we refer to these components as silicon-controlled rectifiers, or SCRs. The progression from Shockley diode to SCR is achieved with one small addition, actually nothing more than a third wire connection to the existing $p-n-p$ structure. This extra terminal is called the gate, and it is used to trigger the device into conduction (latch it) by the application of a small voltage.

⁶ This term is used to describe the Shockley diode “on” state. To get a Shockley diode to latch, the applied voltage must be increased until break-over is attained.

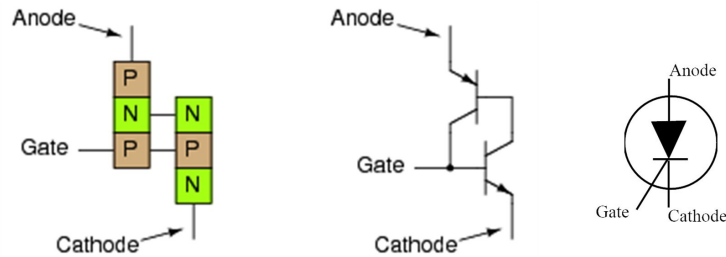


FIGURE 2.15: Circuit components: SCR. from left to right: physical diagram, equivalent schematic, and circuit symbol

SCRs are mainly used in devices where the control of high power, possibly coupled with high voltage, is demanded. Their operation makes them suitable for use in medium- to high-voltage AC power control applications, such as lamp dimming, regulators and motor control.

SCRs and similar devices are used for rectification of high-power AC in high-voltage direct-current power transmission

Transistor

A transistor is a device composed of semiconductor material usually with at least three terminals for connection to an external circuit. The transistor is the fundamental building block of modern electronic devices, and is ubiquitous in modern electronic systems; ranging from signal amplification to the design of digital logic and memory circuits.

The basic principle involved is the use of the voltage between two terminals to control the current flowing in the third terminal, however, the material and technology used to produce transistors are very different, and thus circuit symbols, abbreviations and details of current-voltage relations vary from type to type. In this thesis, we only introduce a Bipolar⁷ Junction Transistor (BJT).

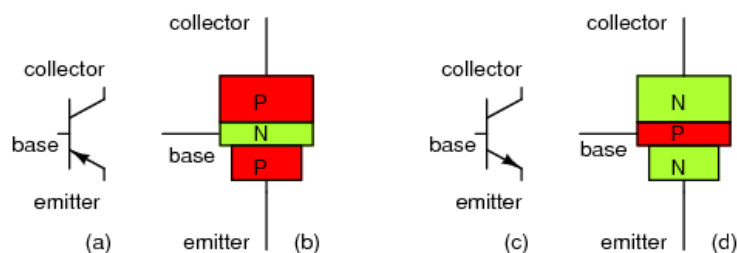


FIGURE 2.16: Circuit components: BJT transistor: (a) PNP schematic symbol, (b) physical layout (c) NPN symbol, (d) layout.

⁷ Bipolar transistors are called bipolar because the main flow of electrons through them takes place in two types of semiconductor material: P and N, as the main current goes from emitter to collector (or vice versa). In other words, two types of charge carriers - electrons and holes - comprise this main current through the transistor.

The BJT consists of two *pn* junctions, the *emitter-base* junction (**EBJ**) and the *collector-base* junction (**CBJ**). Depending on the bias condition (forward or reverse) of each of these junctions, four different modes of operation of the BJT are obtained. In order to understand and analysis the behaviour of a BJT in a circuit, one should study the voltage dropped over each junction and the current flow obtained. Figure 2.17 describes the current flow in an *npn* transistor biased to operate in the *active mode* (that is, when EBJ is in forward bias, and CBJ is in reverse bias).

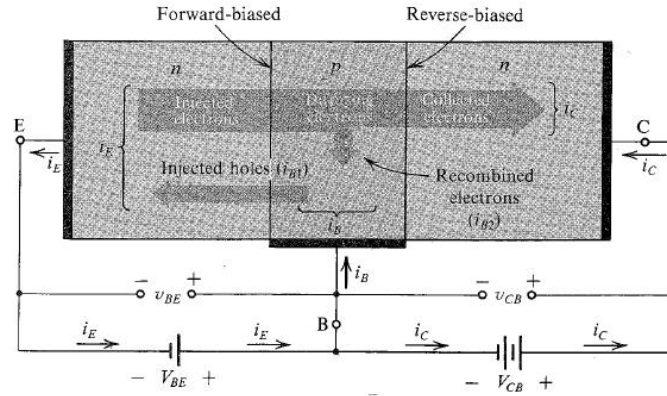


FIGURE 2.17: An *npn* BJT biased in active mode

Among different models that has been suggested for analysing BJT transistor circuits, Ebers and Moll, two early workers in the area, have shown that the following composite model (cf. Figure 2.18) can be used to predict the operation of the BJT in *all* of its possible modes.

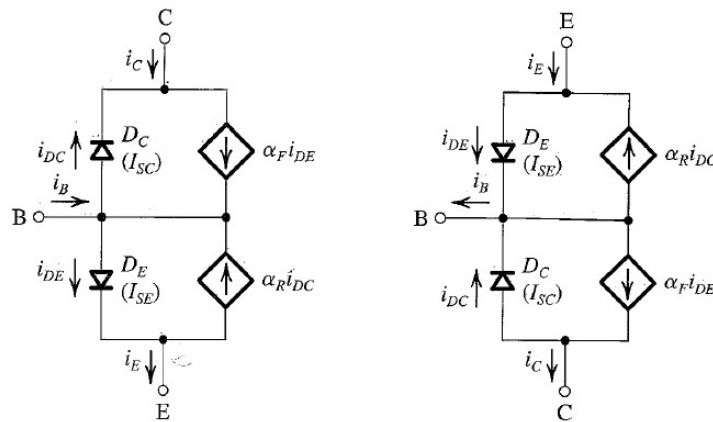


FIGURE 2.18: The Ebers-Moll (EM) model of the *npn* transistor (left), and the *pnp* transistor (right)

Thus, theoretically, one can replace a transistor with two diodes and two current-controlled current sources. The diodes in Ebers-Moll model are assumed to be ideal diodes.

Remark 2.1.1. Before proceeding to the next part of the circuit theory, we want to provide an interpretation of the set-valued map representing the diodes $i - v$ characteristic. If we compare the $i - v$ characteristic of a resistor with the characteristic of a practical diode (Figures 2.3 and 2.11), we can obtain two important facts:

(a) the characteristic of a diode is not enough to identify the exact value of the voltage over it or the current passing through it by knowing the other one; while for the resistor at any point of the $i - v$ characteristic, if you have in hand the value of v_R , you can obtain the value of i_R , and vice versa. However, the characteristic of the diode gives the idea of what could be the possible values of v_D or i_D , and what range of values are not allowed.

In other words, the set-valued characteristic talks about the possibility of the values a diode can take as v_D for a certain value i_D , and vice versa.

(b) in a specific circuit, if you want to change the current passing through a resistor R_1 , while the voltage over it should be fixed at a certain value v_1 , you have no option but to change the resistor; while for the diodes the possibility of accepting different voltages lets us use the same diode. In other words, the voltage and current of a resistor could be determined by its own characteristic, and for the diodes it depends also on the other components of the circuit. Thus, for a range of different components and different combinations, one can still use the same diode, that is, the diode could be matched with a “set” of other components.

Electrical components could be divided into two groups: active, and passive.

The *active* components produce the electrical energy (again described in terms of voltage and current), while the *passive* components are users of this energy. Resistors, and capacitors are examples of passive components, while voltage sources are considered as active elements. Transistors could be biased to work in an active mode, too, so that they would amplify the AC signal.

We need to consider this division while writing the Kirchhoff’s circuit laws.

Kirchhoff’s Laws To analyse a circuit, we need to know the $i - v$ characteristics of all the components in the circuit, but that is not enough. In addition, we need some rules to describe how the current passes through different components and branches, and how to measure the drop of potential differences over each component in a specific circuit. These rules (which are a reduced version of the Maxwell’s electromagnetic equations), are known as *Kirchhoff’s laws*:

- *Kirchhoff’s current law (KCL)*:

at any node (junction) in an electrical circuit, the sum of currents flowing into that

node is equal to the sum of currents flowing out of that node; equivalently, the algebraic sum of currents in a network of conductors meeting at a point is zero.

○ *Kirchhoff's voltage law (KVL):*

the directed sum of the electrical potential differences (voltage) around any closed network is zero; equivalently,

the sum of electrical potential drops in a closed loop is equal to the total electromotive force (emf) available in that loop.

The word “directed” in the statement of KVL, means that the active and passive components have different voltage signs in the sum. It is mostly optional to choose, but both can not be positive (or negative) at the same time. Another way to consider this fact is that the product of current and voltage of passive component should be positive, while the iv of an active component is negative, or vice versa.

For example, note that the conventional direction for voltages and currents in Figure 2.6 imply $i_s \times v_s < 0$, meaning that voltage sources are active components, while for a resistor as shown in Figure 2.3, we have $i_R \times v_R > 0$.

Solving or analysing a circuit means to use the the equations obtained from KVL, KCL, and $i-v$ characteristics of components, to find the current passing through and the voltage over each component of the circuit. This usually ends up to a system of n equations, and n variables.

Example 2.1.2. Figure 2.19 shows a simple circuit with a voltage source as the model for electrical power producer, a resistor $R_B = 1.4\text{ k}\Omega$ as the bias resistor, and a consumer modelled as a load resistor $R_L = 5.6\text{ k}\Omega$. Our goal is to obtain the percentage of the produced voltage that is delivered to the consumer.

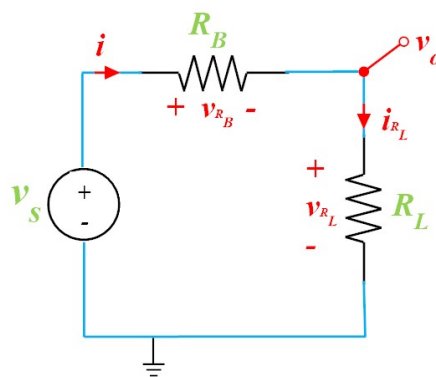


FIGURE 2.19: A simple circuit that divides voltage between a bias resistor and a load

Since there is only one loop in this circuit, KCL implies $i_{R_B} = i_{R_L} = i$. For using KVL, consider a clockwise closed path starting from the voltage source.

$$-v_s + v_{R_B} + v_{R_L} = 0.$$

Now by using the component relations $v_{R_B} = R_B i_{R_B}$, and $v_{R_L} = R_L i_{R_L}$, one can calculate the current value in the circuit as $i = \frac{v_s}{R_B + R_L}$. Using the $i - v$ relation of the resistor once more, yields

$$v_o = v_{R_L} = \frac{R_L}{R_B + R_L} v_s.$$

Thus, with the given values for R_B , and R_L , we would have a 20% loss in delivering the voltage produced by v_s to the load.

2.2 Formulating the Problem

Based on the components in the circuit, and the exactness of the solution required, there is a huge theory and lots of work done around it till now, in electrical engineering literature. We are not exactly interested in this topic, but we will use the setting of circuit theory and the rules mentioned in previous section to formulate our problem.

We start with a simple circuit as shown below (Figure 2.20) which involves an electronic component named *Zener Diode*. Our aim is to find a relation between the given input voltage (E , here) and the variable I , current, in the circuit.

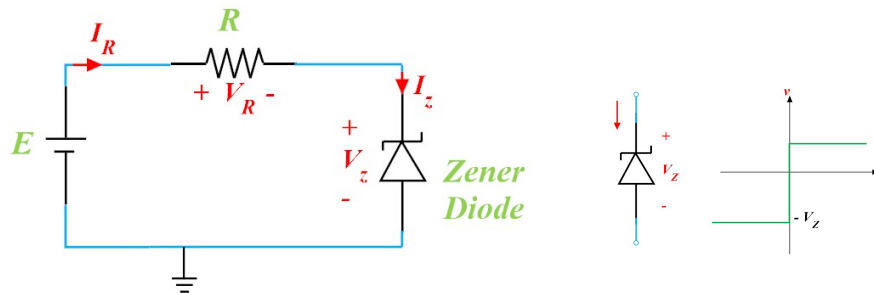


FIGURE 2.20: A simple electrical circuit, the schematic and $i - v$ characteristic of Zener Diode

$$\left. \begin{array}{l} \text{KVL : } -E + V_R + V_z = 0 \\ \text{KCL : } I_R = I_z = I \\ V_R = R I_R \\ V_z \in F(I_z) \end{array} \right\} \Rightarrow E \in R I + F(I), \quad (2.1)$$

in which $F(I)$ describes the relation between I_z and V_z as a set-valued map.

To match our mathematical setting, we change E to p , in order to indicate that it is a *parameter*; and I to z , to show that z is the variable, thus we get

$$p \in f(z) + F(z), \quad (2.2)$$

in which $f(z) = Rz$, in this particular example. For a given p , we are interested in the *solution mapping* defined as

$$S(p) = \{z \in \mathbb{R} \mid p \in f(z) + F(z)\}. \quad (2.3)$$

It is also useful to introduce another notation

$$\boxed{\Phi(z) := f(z) + F(z), \quad \text{so} \quad S(p) = \{z \in \mathbb{R} \mid p \in \Phi(z)\} = \Phi^{-1}(p)}. \quad (2.4)$$

Before proceeding our study with this model, we need to justify its appropriateness.

One may wonder why this form of generalized equation does not appear in electronic books. This depends on the way we try to model the behaviour of the diode. When one is using the ideal model for diodes, usually the two different working modes of the diode (that is, when the diode is on and acting like a wire, and when it is off and acting like a gap or open connection) are studied separately.

When one is only interested in the forward bias, the $i-v$ relation is usually approximated by an exponential term, or some approximations of this exponential map.

Since the $i-v$ characteristic of a Zener Diode (and diodes in general) fails from a function-form description only in one point of its domain, those books approximate the given set-valued map (and similar other maps for other type of diodes) with a single-valued map (look at Figure 2.21). So the later inclusion in the right hand side of (2.1) becomes equality, and they end up in an equation.

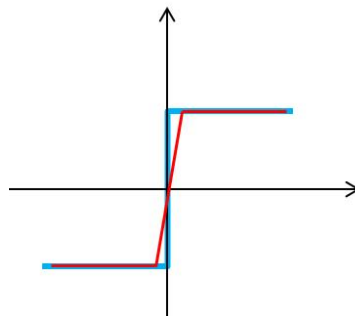


FIGURE 2.21: Approximation of the Diode's $i-v$ characteristic

Moreover, we are not going to solve the circuit equations, which is mostly the aim in electronics. We want to study the behaviour of the solution mapping in a neighborhood of a certain point (that is a local study) and it needs precise models concerning each point (p, z) in the input-output relation.

Regarding different electronic components in the circuit and various circuits, the obtained solution mapping (2.3) might change a bit, but we will come back to this difficulty in few pages, after stating the general purpose of introducing these type of problems.

We are going to answer this question that what will happen if the voltage of source

changes from \bar{p} to p' . There are three specific reasons to validate the importance of this question:

- **Failure in Precise Measurements**

When dealing with real world applications, one should always be aware of inexact measurements. Any producer provides an error percentage or tolerance range for his products. For example, when you have a resistor with the following colored bands : Brown, Green, Red, and Gold, then its resistance is $15 \times 10^2 \pm 5\%$.

Thus, either the signal source is a single component with a tolerance range, or it is a circuit itself, made of several components, you would have a deviation from the exact value \bar{p} .

Also when you are designing a circuit, you find the optimal value of your DC voltage source 10.53 *volts*, for example. But in the market, you cannot have such a battery or combination of batteries that provide this precise value. Thus, you are forced to use a physical component close to your calculations, but not exactly the same.

- **Process of Ageing**

Consider a simple chemical pile as the voltage source. As soon as you start to use it in a circuit, the chemical ionization process that provides the electrical energy runs inside the pile. After a certain time, you would notice that the battery is dead and you need to change it or charge it.

During this period, although the battery level was at the appropriate level, but it was decreasing slowly to get to the unacceptable level. In other words, there were small changes in its precise value.

- **Thermal Effect**

Especially when the signal source is an electronic circuit itself, one should consider that semiconductors are very sensitive to change of temperature.

For example, when an output of a circuit with diodes or transistors is used as the signal source for the next circuit (like the cascade structure in amplifiers), the behaviour of these components change with temperature and thus, the outcome would have a little tolerance around the precise expected value.

We would now try to provide an interpretation of the local stability properties of the solution mapping in terms of the circuit parameters. Let us note that based on the problem one may face during the design process, one of these properties would fit better to his/her demands.

Stability Formulation Suppose for a given \bar{p} , we know the previous current of *operating point*⁸, say \bar{z} . We also assume that the input change is small, that is, in mathematical terms, $p' \in \mathbb{B}_r(\bar{p}) =: V$ for some small $r > 0$.

In fact, we are interested in those circuits that keep the small input-change, small. More precisely, the distance between a $z \in S(p')$ and \bar{z} , is controlled by the distance $\|p' - \bar{p}\|$. Thus we wish (and search for) S having the following property

$$\|z - \bar{z}\| \leq \kappa \|p' - \bar{p}\| \quad \text{when} \quad z \in S(p') \cap U, \quad p' \in V \quad (2.5)$$

where U is a neighborhood of \bar{z} , V is a neighborhood of \bar{p} , and $\kappa \geq 0$ is a constant.

We defined this property as *isolated calmness* in Chapter 1 (cf. Definition 1.2.10).

The reason we considered the intersection $S(p') \cap U$ in the above formulation, is that while it is possible to have different values in $S(p')$, we are just interested in quarantining the existence of a z near enough to \bar{z} . S may be not single-valued at p' , but we do not care (in other words, the control is not about that); what matters is that it must have a value in U .

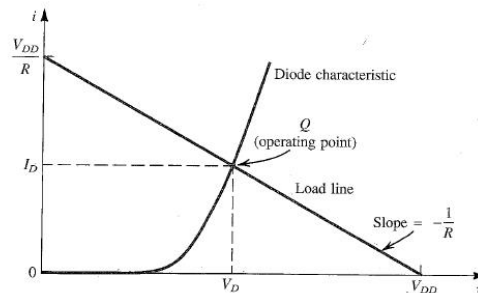
The very next question about equation (2.5) would be whether S is single-valued at \bar{p} or not. Actually, single-valuedness is implicitly assumed in that formulation, but one could be in a situation described in Figure 2.22 involving a Tunnel diode.

The answer is that it is not a necessary assumption in general, though in some particular cases it might be a request. Thus, we get a slight modification of the previous formulation and ask for the existence of $\kappa \geq 0$, neighborhoods U of \bar{z} , and V of \bar{p} such that

$$e(S(p') \cap U, S(\bar{p})) \leq \kappa \|p' - \bar{p}\| \quad \text{for all } p' \in V \quad (2.6)$$

where $\bar{z} \in S(\bar{p})$, and $e(A, B)$ is the *excess* of A beyond B . We defined this property as *calmness* in Chapter 1 (cf. Definition 1.2.8).

⁸ In the graphical analysis of the circuit, we plot two maps on the same $i - v$ plain: the $i - v$ characteristic of the diode, and the ordered (with respect to i) equation of the circuit gained from KVL. The solution can then be obtained as the equilibrium point, that is the coordinates of the intersection point of the two graphs. This point is called “operating point” (look at the figure below). It is easy to see that it depends both on the inner structure of the diode and the rest of the circuit.



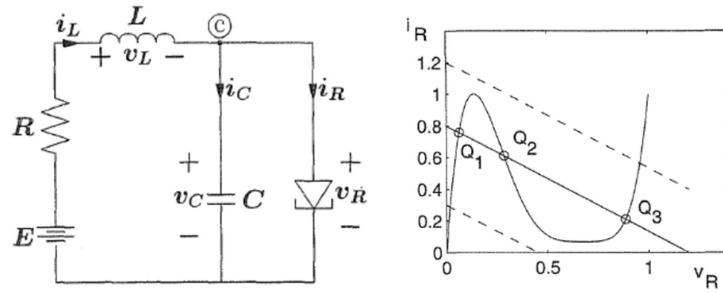


FIGURE 2.22: A circuit with Tunnel diode (left) and the graphical analysis of it (right)

If we are investigating a general local property of the solution mapping and the point $(\bar{p}, \bar{z}) \in \text{gph } S$ does not play a crucial role in our study, we would be interested in the “two-variable” version of the previous condition, that is

$$e(S(p') \cap U, S(p)) \leq \kappa \|p' - p\| \quad \text{for all } p', p \in V := \mathbb{B}_r(\bar{p}), \quad (2.7)$$

We defined this as **Aubin property** in Chapter 1 (cf. Definition 1.2.2).

Regularity Formulation Up to now, our construction were built under the assumption that the explicit form of the solution mapping $S(\cdot)$ is in hand and so we can easily calculate values like $S(p')$.

It is not true in general. All we are sure we can get from the circuit is $\Phi = S^{-1}$. It is not always simple (or even possible) to derive the formula of $S(\cdot)$, so we need to provide the proper formulation of the desired properties with Φ .

Two slightly different interpretations of the problem will result in the two different approaches described below.

Approach 1. One can take almost the same procedure. Starting from \bar{z} , and $\bar{p} \in \Phi(\bar{z})$, then consider a point $z \in U =: \mathbb{B}_r(\bar{z})$, calculate $\Phi(z)$, and ask for a control $\kappa \geq 0$ satisfying the following condition:

$$d(z, \bar{z}) \leq \kappa d(\bar{p}, \Phi(z) \cap V) \quad \text{for all } z \in U. \quad (2.8)$$

This property was introduced as **strong metric sub-regularity** in Chapter 1 (cf. Definition 1.3.16).

If one releases the single-valuedness condition $\Phi(\bar{z}) = \{\bar{p}\}$, gets

$$d(z, \Phi^{-1}(\bar{p})) \leq \kappa d(\bar{p}, \Phi(z) \cap V) \quad \text{for all } z \in U. \quad (2.9)$$

which was defined as **metric sub-regularity** in Chapter 1 (cf. Definition 1.3.13).

Then, we can do a general local study around an arbitrary point $(\bar{z}, \bar{p}) \in \text{gph } \Phi$ by using

a *two variable* condition

$$d(z, \Phi^{-1}(p)) \leq \kappa d(p, \Phi(z)) \quad \text{whenever } (z, p) \in U \times V. \quad (2.10)$$

which was defined as **metric regularity** in Chapter 1 (cf. Definition 1.3.1).

Approach 2. One can consider a different approach toward the problem, by starting from a point $(\bar{z}, \bar{p}) \in \text{gph } \Phi$, and taking an arbitrary $z' \in U =: \mathbb{B}_{r_1}(\bar{z})$, calculating $\Phi(z')$, taking a $p \in V =: \mathbb{B}_{r_2}(\bar{p})$, and then asking for the chance of having the output distance $d(z', \Phi^{-1}(p))$ being controlled by the input distance $d(p, \Phi(z'))$, that is,

$$d(z', \Phi^{-1}(p)) \leq \kappa d(p, \Phi(z')) \quad \text{whenever } (z', p) \in U \times V, \quad (2.11)$$

which is again **metric regularity**. Then make slight modifications by considering the additional condition of single-valuedness at \bar{z} (that is, $\Phi(\bar{z}) = \{\bar{p}\}$, called **strong metric regularity**, cf. Definition 1.3.9). Or considering the one-variable version of the above condition (introduced as **metric sub-regularity**).

Note 2.2.1. (a) The numerical advantage of regularity formulation was already discussed in Remark 1.3.4.

(b) One should pay attention to the fact that z' , and p , are chosen independently. Indeed, $p \in \Phi(z')$ ends up at an obvious situation $0 \leq 0$.

(c) We have already seen that there is some relations between the two mentioned formulations (cf. Section 1.3):

$$\begin{array}{ccccccc} \text{for } \Phi : & SMR & \subset & MR & \subset & MSR & \supset & SMSR \\ & & & \cong & & \cong & & \cong \\ \text{for } S : & & & \text{Aubin Property} & \subset & \text{Calmness} & \supset & \text{Isolated Calmness} \end{array}$$

Toward Complexity We want to introduce two techniques that play an important role in our future calculations. They would also annihilate the idea that our modelling is so much simplified and will not be able to cover the real cases in practical complicated circuits, though we may not refer to such circuits here.

- **Cascading**

In some circuits, one can “slice” the circuit to smaller ones, study each separately and sum them up at the end. This is a frequently used technique in designing Amplifiers, and other circuits with very special limitations. Here we used “cascading” as

a general term for objects connected serially that start a chain reaction⁹. For a better understanding of the situation look at the working Diagram 2.23, in which the network is divided into three subnets:

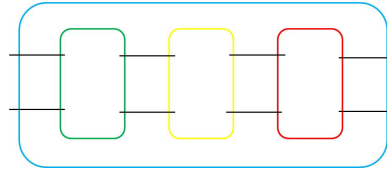


FIGURE 2.23: Cascading

The “slicing” process should be done by considering the working diagram and how components are connected to each other for their functionality. Thus, one should be careful to avoid cutting the feedback loops¹⁰, or separating a current controlled source from the control branch, ... and such mistakes.

The calculus rules provided in Subsection 3.1.1, especially the composition Propositions 3.1.26 and 3.1.27 will be the mathematical tools which give us the permission to use this technique while studying the stability properties.

Example 2.2.2. In the following circuit (Figure 2.24) we can study the stability of the two simpler circuits on the right.

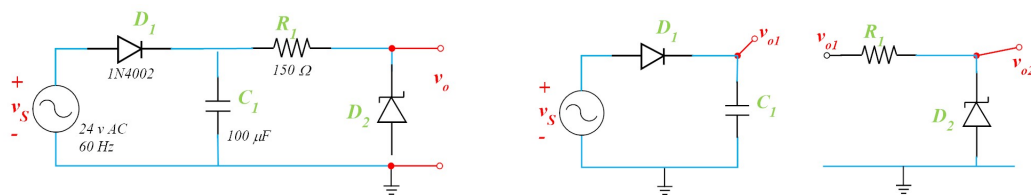


FIGURE 2.24: Simplifying the study of the circuit by splitting the layers

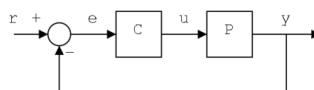
Indeed, in the right hand-side figure, we can study the stability of the *transfer functions* $\frac{v_{o1}}{v_s}$ and $\frac{v_{o2}}{v_{o1}}$ separately, considering the output of the first sub-circuit as an independent source v_{o1} of the second one. Then use a composition theorem to obtain the stability of the transfer function $\frac{v_o}{v_s}$.

• Multivalued Simplification

As we usually end up to a term like $f(z) + F(z)$, we try to make the set-valued map

⁹ In fact, the title “cascading” was drawn from *cascade amplifier circuits*, which are circuits made of different stages, each serves a special improvement of the current or voltage gain in a serial connection.

¹⁰ Feedback loop is a design method that returns part of the output of the system into input, in order to enable the system to adjust its performance to meet a desired output response. The operational diagram of a typical feedback loop could be seen in the following figure



as simple as possible. Such a simplification may not have a real physical meaning, but it would be a correct mathematical operation. To be more clear, consider the following figure that shows the use of this technique for a DIAC.

Although the pointwise sum of two graphs on the right would be equal to the graph on the left, the first graph (blue one, representing a single-valued relation) is not an $i-v$ of a component. One can think of the possibility of making a complicated circuit that has such an $i-v$ characteristic. Though it is possible in theory, and thus our technique is valid, in practice, it might be contrary to the idea of “simplification” we were trying to follow.

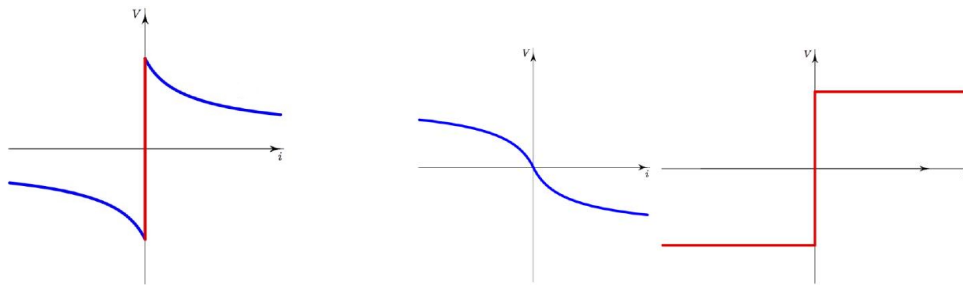


FIGURE 2.25: A simplified $i-v$ characteristic of a DIAC and the equivalent maps

This technique will be very useful, and essential somehow when calculating the tangent and normal cones, the tools we will provide for our study in next chapter.

Now we are ready to go back to the question we left before: the effect of changing components or their order on the final form of generalized equation that would be obtained. We try to provide the answer of this question with some examples.

Example 2.2.3 (Increasing Components and Loops). In Figure 2.26, we can see a circuit with two loops, each one containing a Diode. This will be a good example about how the setting will change in \mathbb{R}^n .

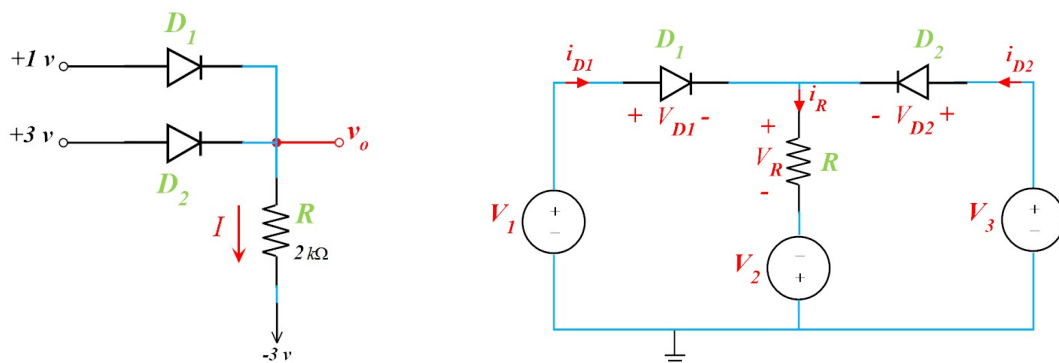


FIGURE 2.26: A circuit with two loops

Suppose the diodes do not have the same characteristic. One can write the following:

$$\left. \begin{array}{l} \text{KVL1 : } -V_1 + V_{D_1} + V_R - V_2 = 0 \\ \text{KVL2 : } -V_3 + V_{D_2} + V_R - V_2 = 0 \\ \text{KCL : } \quad i_{D_1} + i_{D_2} = i_R \\ \quad \quad V_R = R i_R \\ \quad \quad V_{D_1} \in F_1(i_{D_1}) \\ \quad \quad V_{D_2} \in F_2(i_{D_2}) \end{array} \right\} \Rightarrow \begin{array}{l} (V_1 + V_2) \in R(i_{D_1} + i_{D_2}) + F_1(i_{D_1}) \\ (V_3 + V_2) \in R(i_{D_1} + i_{D_2}) + F_2(i_{D_2}) \end{array} \quad (2.12)$$

So we get the following unified form

$$\begin{pmatrix} p_1 \\ p_2 \end{pmatrix} \in \begin{pmatrix} R & R \\ R & R \end{pmatrix} \begin{pmatrix} i_{D_1} \\ i_{D_2} \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} F \left(\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} i_{D_1} \\ i_{D_2} \end{pmatrix} \right) \quad (2.13)$$

in which $p_1 = V_1 + V_2$, $p_2 = V_3 + V_2$, and $F : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ is defined by

$$F \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} := \left\{ (x, y) \mid x \in F_1(z_1), y \in F_2(z_2) \right\} = \prod_{i=1}^2 F_i(z_i)$$

by letting $p := (p_1, p_2)^T$, and $z := (z_1, z_2)^T$, we get to a general form

$$p \in f(z) + BF(Cz) \quad (2.14)$$

In this example, $f(z) = Az = \begin{pmatrix} R & R \\ R & R \end{pmatrix} z$, $B = C = \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$.

Most of our results till now is about this form of generalized equation. We correct the introduced Φ and the solution mapping in the following way

$$\boxed{\begin{array}{l} \Phi(z) := f(z) + BF(Cz) \\ S(p) := \{z \mid p \in \Phi(z)\} \end{array}} \quad (2.15)$$

Thus, one can see increasing the components with the $i - v$ characteristic described as a set-valued map, in different loops will reformulate the problem in \mathbb{R}^n with some correction coefficients to indicate the presence or absence of those set-valued components in different branches. We will study this type of problem in Chapter 3.

Now, we consider the change of power sources from constant (that is, DC sources) to time-varying ones (that is, AC sources).

Example 2.2.4 (Effect of AC Sources). In the following figure, we have an alternating voltage source, i.e. a voltage source depending on time.

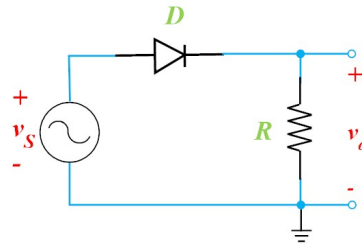


FIGURE 2.27: A circuit with AC voltage source

Mostly, such voltage sources would be of the form $v_s(t) = A \sin(t - t_0)$. To indicate the time-dependent nature of this circuit, not only we use (t) in front of the functions, but also use small letters to refer to a voltage or current.

We can easily obtain the generalized equation in this case as $v_s(t) \in Rz(t) + F(z(t))$.

In complicated circuits we may have

$$\begin{aligned} \Phi(z(t)) &:= f(z(t)) + BF(Cz(t)) \\ S(p) &:= \{z \mid p \in \Phi(z) \text{ for all } t \in [a, b]\}. \end{aligned} \quad (2.16)$$

So this case may reduce to solving the previous case (in which p was a fixed vector in \mathbb{R}^n) for any t in an interval $[a, b]$. We will study this type of problem in Chapter 4.

Example 2.2.5 (AC Source with a Capacitor; A Rectifier Circuit). Using AC voltage sources will allow us to use passive components apart from resistors. In Figure 2.28, we can see another simple circuit with a diode and a capacitor, and the voltage source is a sinusoid function.

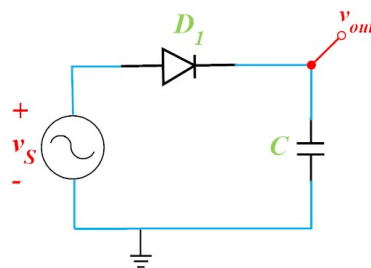


FIGURE 2.28: A simple AC to DC Rectifier circuit

Although we tried so much to avoid encountering capacitors and inductors till now, in almost any practical circuit one will find these components (mostly capacitors). So it is better to show in a simple circuit why we have ignored them up to now.

The $i - v$ relation of the capacitor is $i_c = C \frac{dv_c}{dt}$, in which $C > 0$ is a constant called capacitance, and the $i - v$ relation of the inductor is $v_L = L \frac{di_L}{dt}$, in which $L > 0$ is a

constant called inductance.

$$\left. \begin{array}{l} \text{KVL : } -v_s + v_{D_1} + v_c = 0 \\ \text{KCL : } i_{D_1} = i_c =: i \\ i_c = C \frac{dv_c}{dt} \\ V_{D_1} \in F(i_{D_1}) \end{array} \right\} \Rightarrow v_s(t) \in \frac{1}{C} \int_0^t i \, d\tau + F(i) \quad (2.17)$$

2.2.1 Alternative Formulation

In this subsection, we will shortly discuss another suggested formulation for components like diode, and transistor in the literature (see for example [2, 17, 35, 36]). We would explain why the setting of generalized equation (GE) is more appropriate for our study and how it covers more circuits.

Although one can follow the process used in the previous section, the special $i - v$ characteristic of the ideal diode give us a chance to gain another formulation. Let us consider the circuit in Figure 2.20 with ideal diode instead of a Zener diode and observe that at any point $(i_D, v_D) \in \text{gph } F$ (where F is the set-valued map relating the current passing through ideal diode to the voltage over it) one can write $v_D i_D = 0$.

So in this example one can write

$$\langle p - Rz, z \rangle = 0, \quad (2.18)$$

where $p = V_I$, $z = i_D$, and $\langle \cdot, \cdot \rangle$ indicates the scalar product in \mathbb{R}^n . Following this method we will end up at what is called *Variational Inequality*.

Not only for Ideal Diodes, but also when dealing with transistors as switches, we can formulate our problem as a VI.

2.3 Examples

This section will contain some real world examples. Each example will display part of the mathematical formulation difficulties, and will shed light on different aspects of the problem. We will provide more details of these examples in the following chapters, after providing the necessary mathematical tools for this study.

Example 2.3.1. (Driving an LED with AC, on the Importance of Understanding the Behaviour of the Circuit) [1]

Figure 2.29 depicts an LED paralalled backward with a rectifying diode. Since LEDs are made of different chemical substances than silicon diodes, their forward voltage drops will be different. Typically, LEDs have much larger forward voltage drops than rectifying

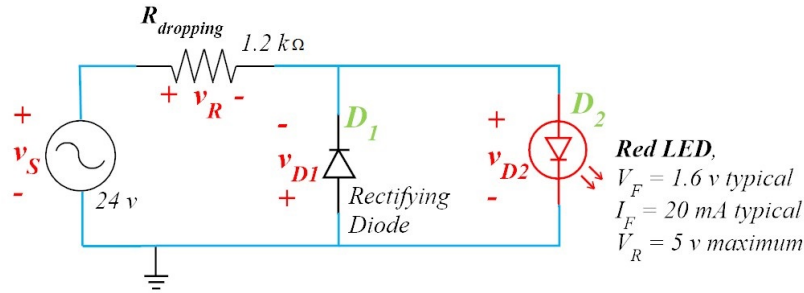


FIGURE 2.29: Driving an LED with AC

diodes, anywhere from about 1.6 *volts* to over 3 *volts*, depending on the color. Typical operating current for a standard-sized LED is around 20 *mA*. When operating an LED from a DC voltage source greater than the LED's forward voltage, a series-connected “dropping” resistor must be included to prevent full source voltage from damaging the LED.

Also because of their unique chemical makeup, LEDs have much, much lower peak-inverse voltage (PIV) ratings than ordinary rectifying diodes. A typical LED might only be rated at 5 volts in reverse-bias mode. Therefore, when using alternating current to power an LED, connect a protective rectifying diode anti-parallel with the LED to prevent reverse breakdown every other half-cycle as in Figure 2.29.

Now, we try to give the appropriate mathematical model for this circuit; which one can guess such a model could not have ideal diode model, or same set-valued map for both diodes. Let D_1 indicates the protective rectifying diode, and refer to LED with D_2 . KVL, KCL, and Diode Characteristics are as follows:

$$\begin{cases} -v_s + v_R - v_{D_1} = 0, \\ i_R = -i_{D_1} + i_{D_2} \end{cases}, \quad \begin{cases} v_{D_1} = -v_{D_2}, \\ v_{D_j} \in F(i_{D_j}), \quad j = 1, 2. \end{cases}$$

Let $x = \begin{pmatrix} v_{D_1} \\ v_{D_2} \end{pmatrix}$, and $u = \begin{pmatrix} i_{D_1} \\ i_{D_2} \end{pmatrix}$. Then from the above relations, we get

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \begin{pmatrix} -1 \\ +1 \end{pmatrix} v_s + \begin{pmatrix} -R & R \\ R & -R \end{pmatrix} \begin{pmatrix} u_1 \\ u_2 \end{pmatrix}$$

with $x_j \in F_j(u_j)$.

So, one can form the generalized equation as follows:

$$B + AU \in F(U), \quad (2.19)$$

where $B = \begin{pmatrix} -v_s \\ +v_s \end{pmatrix}$, $A = \begin{pmatrix} -R & R \\ R & -R \end{pmatrix}$, and $F : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ is defined as

$$F \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} := \begin{pmatrix} F_1(u_1) \\ F_2(u_2) \end{pmatrix}$$

In Chapter 3, we would see why this formulation cannot be written as a variational inequality (see Example 3.1.9).

Example 2.3.2. (A Simple Circuit with DIAC)

Let us consider the simple circuit in Figure 2.30 with a DIAC, whose $i - v$ characteristic for given $V > 0$, and $a > 0$ is as follows:

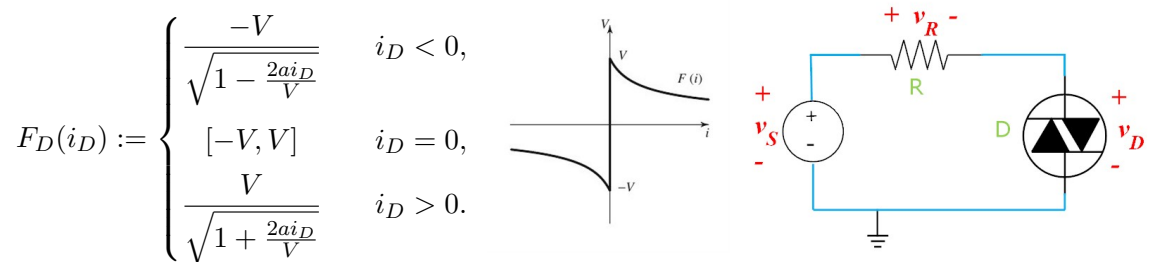


FIGURE 2.30: A simple circuit with DIAC

Using KVL, KCL, and the characteristic relations we obtain that

$$\begin{cases} -v_s + v_R + v_D = 0, \\ i_R = i_D = i. \end{cases} \quad \begin{cases} v_R = Ri_R \\ v_D \in F_D(i_D). \end{cases}$$

By replacing $p = v_s$, and $z = i$, we get the generalized equation

$$p \in f(z) + F(z), \tag{2.20}$$

where $f(z) = Rz$, and $F(z) = F_D(z)$ for every $z \in \mathbb{R}$. For each $p \in \mathbb{R}$, one can look at the solution mapping associated to (2.20) as the intersection of the line $\varphi(z) = p$ and the set-valued mapping $\Phi = f + F$. This interpretation would provide a general idea about the solution sets $S(p)$. Figure 2.31, shows the graph of Φ for different values of parameters R and a .

Note that for each $p \in \mathbb{R}$ there is a solution $z \in \mathbb{R}$ to the equation (2.20). Indeed, as $\Phi(z) = Rz + F(z)$ for $z \in \mathbb{R}$, we have $\text{rge } \Phi = \mathbb{R}$. Moreover, this solution is unique if

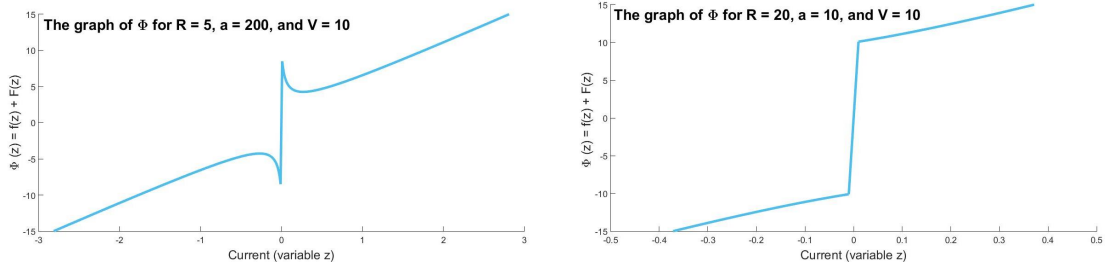


FIGURE 2.31: The graph of Φ for different values of parameters R , a and V (left: $a = 200, R = 5, V = 10$; right: $a = 10, R = 20, V = 10$)

$a \leq R$, since for a non-zero $z \in \mathbb{R}$ one infers that

$$\Phi'(z) = R - \frac{a}{\left(\frac{2a|z|}{V} + 1\right)^{\frac{3}{2}}} = \frac{R \left(\frac{2a|z|}{V} + 1\right)^{\frac{3}{2}} - a}{\left(\frac{2a|z|}{V} + 1\right)^{\frac{3}{2}}} > \frac{R - a}{\left(\frac{2a|z|}{V} + 1\right)^{\frac{3}{2}}} \geq 0.$$

In this case, the solution equals to zero if $p \in [-V, V]$; it is positive if $p > V$, and negative when $p < -V$.

Clearly, we can split Φ into different functions f and F which produce the same inclusion (2.20). Namely, from now on, we assume that

$$f(z) := \begin{cases} Rz - \frac{V}{\sqrt{1 - \frac{2az}{V}}} + V & z < 0, \\ Rz + \frac{V}{\sqrt{1 + \frac{2az}{V}}} - V & z \geq 0, \end{cases} \quad \text{and} \quad F(z) := \begin{cases} -V & z < 0, \\ [-V, V] & z = 0, \\ V & z > 0. \end{cases}$$

This is the *simplification idea* we discussed before in this chapter. Obviously, this new function f does not represent an electronic component but the sum $f + F$ and thus, the solution mapping remain unchanged. Hence, from the analytic point of view, we are allowed to do this change of functions in the sum.

$$f'(z) = R - \frac{a}{\left(\frac{2a|z|}{V} + 1\right)^{\frac{3}{2}}} \quad \text{whenever } z \in \mathbb{R}.$$

Indeed, $f'(z) = \Phi'(z)$ if $z \neq 0$, and $f'(0) = R - a$ could be obtained from the following simple calculations:

$$\lim_{h \rightarrow 0^+} \frac{Rh + \frac{V}{\sqrt{1 + \frac{2ah}{V}}} - V}{h} = \lim_{h \rightarrow 0^+} \left(R - \frac{a}{\left(\frac{2ah}{V} + 1\right)^{\frac{3}{2}}} \right) = R - a.$$

The case $a > R$ could be discussed similarly, but the solution is not necessarily unique

any more (see Figure 2.31). In Chapter 3, we will examine the Aubin property and the isolated calmness of the solution mapping at various reference points (refer to Examples 3.2.10 and 3.3.6).

Apart from the electrical importance, this examples describes a simple situation in which the solution set $S(p)$ is not a singleton for some $p \in \mathbb{R}$, even if we approximate the set-valued part of the sum $f + F$ with a line from $-\varepsilon$ to $+\varepsilon$ (for a small $\varepsilon > 0$, and with a steep slope $\frac{V}{\varepsilon}$).

We can also observe the importance and ease of dealing with the metric regularity of Φ instead of working with the Aubin property of S . Although they are describing the same thing, which is the behaviour of the system, it is much easier to obtain the graph of Φ .

Example 2.3.3. (A Circuit with SCR and Zener Diode)

Consider the circuit in Figure 2.32 involving a load resistance $R > 0$, two bias sources $E_2 > E_1 > 0$, an input-signal source u with corresponding instantaneous current i , and two non-smooth elements: the SCR and the Zener Diode.

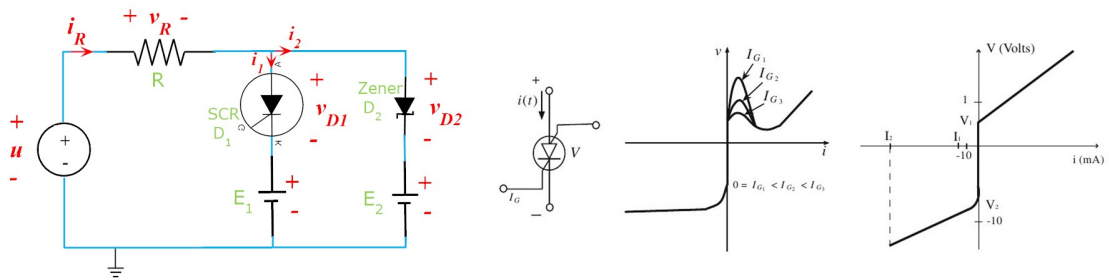


FIGURE 2.32: A circuit with SCR and Zener Diode

Suppose that $V > 0, V_1 < 0, \alpha > 0, a > 0$ and $b > 0$ are given numbers, and that $\varphi : \mathbb{R} \rightarrow (0, \infty)$ is a continuously differentiable function with $\varphi(\alpha) < \varphi(0), \varphi'(0) > 0$ and $\varphi'(\alpha) > 0$. Suppose that the $i - v$ characteristics F_1 of SCR and F_2 of the Zener diode are defined by

$$F_1(z) := \begin{cases} az + V_1 & z < 0, \\ [V_1, \varphi(0)] & z = 0, \\ \varphi(z) & z \in [0, \alpha], \\ a(z - \alpha) + \varphi(\alpha) & z > \alpha. \end{cases} \quad F_2(z) := \begin{cases} bz - V & z < 0, \\ [-V, V] & z = 0, \\ bz + V & z > 0. \end{cases}$$

Using KVL, KCL, and the characteristic relations we obtain that

$$\begin{cases} -u + v_R + v_{D_1} + E_1 = 0, \\ -u + v_R + v_{D_2} + E_2 = 0, \\ i_R = i_{D_1} + i_{D_2}. \end{cases} \quad \begin{cases} v_R = R i_R \\ v_{D_1} \in F_1(i_{D_1}), \\ v_{D_2} \in F_2(i_{D_2}). \end{cases}$$

By replacing $p_1 = u - E_1$, $p_2 = u - E_2$, $z_1 = i_{D_1}$, and $z_2 = i_{D_2}$ we would have

$$\begin{cases} p_1 \in R(z_1 + z_2) + F_1(z_1), \\ p_2 \in R(z_1 + z_2) + F_2(z_2), \end{cases} \quad (2.21)$$

which is in the form of the generalized equation

$$p \in f(z) + BF(Cz), \quad (2.22)$$

with $m = n = 2$, $B = C = I_2$, and $f(z) = Az$ for $z \in \mathbb{R}^2$, where $A = \begin{pmatrix} R & R \\ R & R \end{pmatrix}$, and $F : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ is defined as

$$F \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} := \begin{pmatrix} F_1(z_1) \\ F_2(z_2) \end{pmatrix}.$$

This example represents a more complicated situation with respect to the previous example. Not only there are two different set-valued maps in the model, but also any simplification for one of them, would destroy the symmetry of f . Also, despite the parallel connection of diodes, the role of resistor showed clearly in formulation (2.21) warns us not to study each diode (and so, each generalized equation) separately.

Later, in Example 3.2.11 we will discuss the Aubin property of the solution mapping and we will see the dependence of this property to the various parameters of the circuit.

Example 2.3.4. (Sampling Gate)

Consider a particular circuit of the *sample and hold circuits*¹¹ family in Figure 2.33, composed of four diodes D_1, D_2, D_3, D_4 which are controlled symmetrically by gate voltages $+V_c$ and $-V_c$, and the control resistors $R_c > 0$. The input signal is given by v_i and the output signal is defined by the voltage v_{out} over the load resistor $R_L > 0$.

During *sampling*, equal currents passing thorough resistors R_c , flow into the diode bridge, and v_{in} is copied over R_L . During *hold*, no current flows through the bridge. The reason we are interested in this circuit is that we have three independent loops (that is, $n = 3$), and four diodes (that is, $m = 4$), while in all the previous examples we had $m \leq n$. Assume that all the diodes have the same characteristics given for $V_{D1} < 0 < V_{D2}$ by the mentioned set-valued map.

¹¹ In electronics, a sample and hold (S&H) circuit is a device that captures (samples) the voltage of a continuously varying analog signal and freezes (holds) its value at a constant level for a specified minimum period of time. Sample and hold circuits and related peak detectors are the elementary analog memory devices. If the output is available during the sampling period, we would have a track and hold (T&H) circuit.

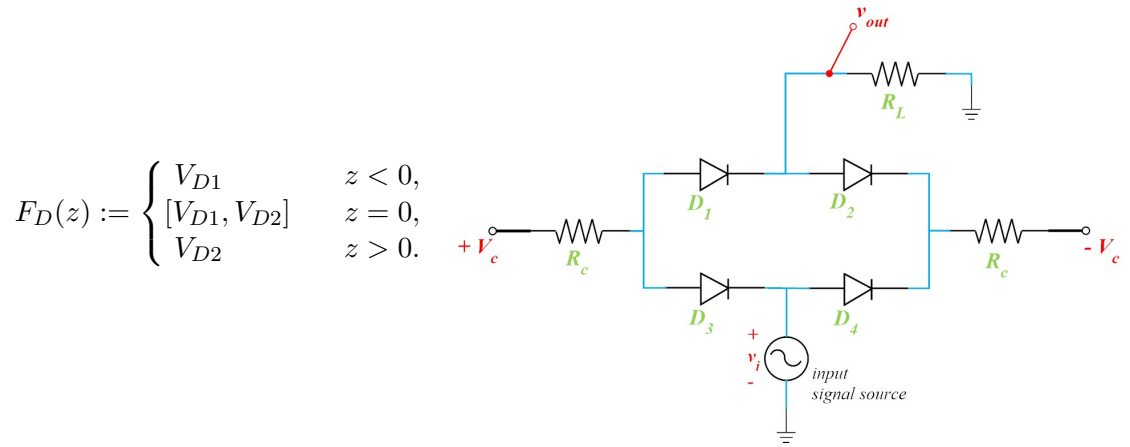


FIGURE 2.33: Loops and variables definition in sampling gate circuit

One can write the Kirchhoff's laws in the following form (Look at Figure 2.34)

$$\left\{ \begin{array}{l} \text{KVL1: } -v_i + v_4 - v_2 + R_L i_7 = 0, \\ \text{KVL2: } V_c - R_c i_6 - v_4 - v_3 - R_c i_5 + V_c = 0, \\ \text{KVL3: } v_1 + v_2 - v_4 - v_3 = 0, \end{array} \right. \quad \left\{ \begin{array}{l} \text{KCL1: } i_5 = i_1 + i_3 = i_2 + i_4 = i_6, \\ \text{KCL2: } i_1 - i_2 = i_4 - i_3 = i_7, \end{array} \right.$$

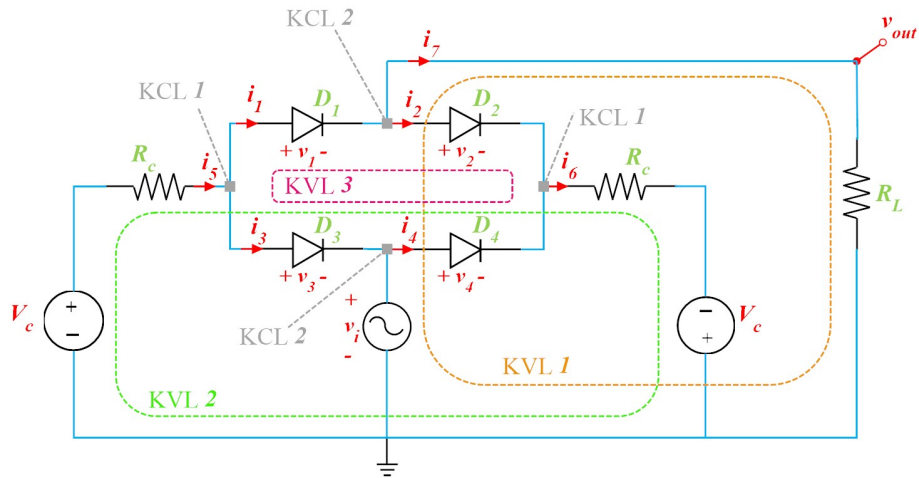


FIGURE 2.34: Sampling gate circuit

where v_j is the voltage over the diode D_j for $j \in \{1, 2, 3, 4\}$. Therefore, by letting $z = (i_7, i_6, i_1)^T$, $v = (v_1, v_2, v_3, v_4)^T$, and $u = (v_i, 2V_c)^T$, one can obtain the linear equation

$$Du = Az + Bv,$$

where

$$D = \begin{pmatrix} 1 & 0 \\ 0 & 1 \\ 0 & 0 \end{pmatrix}, \quad A = \begin{pmatrix} R_L & 0 & 0 \\ 0 & 2R_c & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad \text{and } B = \begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{pmatrix}$$

Now, from the diode characteristics we have that $v_1 \in F_D(i_1)$, $v_2 \in F_D(i_2) = F_D(i_1 - i_7)$, $v_3 \in F_D(i_3) = F_D(i_6 - i_1)$, and $v_4 \in F_D(i_4) = F_D(i_7 + i_6 - i_1)$. Hence, we obtain $v \in F(Cz)$ where

$$C = \begin{pmatrix} 0 & 0 & 1 \\ -1 & 0 & 1 \\ 0 & 1 & -1 \\ 1 & 1 & -1 \end{pmatrix} = B^T, \quad \text{and } F(y) = \begin{pmatrix} F_D(y_1) \\ F_D(y_2) \\ F_D(y_3) \\ F_D(y_4) \end{pmatrix} \text{ for } y \in \mathbb{R}^4.$$

So, we arrived at the generalized equation $p \in f(z) + BF(Cz)$ with $n = 3$, $m = 4$, $p := Du$, and $f(z) := Az$. In Chapter 3 we will investigate the isolated calmness property of the solution mapping at some reference points (ref. Example 3.3.12).

Chapter 3

Tools of Variational Analysis for Studying the Local Stability of the Solution Mapping in the Static Case

In mathematics the art of proposing a question must be held of higher value than solving it.

Georg Cantor

We start this chapter with introducing some tools from variational analysis. To be more precise, the concept of graphical derivative for set-valued maps is introduced in Section 3.1, and some calculus rules for these derivatives is provided in subsection 3.1.1. Then we provide two norm-like maps in Subsection 3.1.2, which allow us to characterize the stability-like properties of a set-valued map in terms of the *inner* and *outer norm* of their graphical derivatives. Subsection 3.1.3 will introduce *subdifferentials* and *generalized Jacobians* for functions that may not be differentiable. These concepts would be used later in this chapter to show a possible way to continue the stability study when dealing with a generalized equation with a non-smooth single-valued part.

The remaining sections of this chapter have a common structure. Each section will discuss one of the stability-like properties for the specific generalized equation $f(\cdot) + BF(C\cdot) \ni p$, with the following assumptions:

- (A1) B is injective;
- (A2) $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a continuously differentiable function;
- (A3) $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ is a set-valued map with closed graph;
- (A4) C is surjective; and

(A5) there are $F_i : \mathbb{R} \rightrightarrows \mathbb{R}$, $i \in \{1, \dots, m\}$ such that $F(x) = \prod_{i=1}^m F_i(x_i)$ whenever $x = (x_1, \dots, x_m)^T \in \mathbb{R}^m$.

We try to use the norm characterization obtained for an arbitrary set-valued mapping in Subsection 3.1.2 for each property, and calculus rules of Subsection 3.1.1 to go step by step toward a pointwise easy-to-check criteria for the local stability of the solution mapping.

The main theorem of each section uses only assumptions (A1) - (A3).

Among the first three assumptions, (A3) is not such a strong requirement and holds for the $i - v$ characteristics of semiconductors like diodes in our study. In Subsection 3.3.1 we try to consider the case where B is not injective, and use the following condition instead of (A1) to provide some similar statements for isolated calmness in this case.

($\widetilde{A1}$) Suppose that there is $\bar{v} \in F(C\bar{z})$ such that

$$\bar{p} = f(\bar{z}) + B\bar{v} \quad \text{and} \quad \bigcup_{t>0} \overline{\frac{\text{rge } F_C - \bar{v}}{t}} \cap \ker B = \{0_{\mathbb{R}^m}\}.$$

The process is almost the same as before.

In Subsection 3.5.1 we focus on functions f which are not smooth enough to satisfy (A2). A possible approach would be considering the generalized Jacobians and replace (A2) with

($\widetilde{A2}$) f is locally Lipschitz continuous on \mathbb{R}^n .

In Section 3.5, using the relations between the regularity terms and local stability properties of the inverse map expressed in Section 1.3, we restate the results of previous sections in terms of metric regularities in Theorems 3.5.6, and 3.5.13.

3.1 Variational Geometry

In this section we will first introduce some concepts that let us define derivatives of a set-valued map in a proper way. From elementary calculus, we know that the derivative of a function f at a certain point \bar{x} , could be interpreted as the slope of the tangent line to the graph of f at the point $(\bar{x}, f(\bar{x}))$. This tangent line could be then used to approximate the function in a vicinity of the reference point.

The idea is quite the same for set-valued maps. We will define approximating cones to the graph of a set-valued map at a certain point, and use this *graphical approximation* to introduce the derivatives.

Then, in Subsection 3.1.1 some simple calculus rules for the newly defined derivatives

will be given. Subsection 3.1.2, will bridge the local stability concepts of Chapter 1 and derivative concepts of this chapter.

Finally, we will look at the non differentiable (in the standard common sense) functions and introduce another graphical-based concept that could be considered as a replacement of the non-existing derivative, in Subsection 3.1.3.

Definition 3.1.1. (Tangent Cones)

[67, p. 132] [54, p. 69] [11, p. 162] Let Ω be an arbitrary non-empty subset of \mathbb{R}^d containing a point \bar{x} .

- (a) The *Bouligand-Severi tangent/contingent cone* $T(\bar{x}; \Omega)$ to Ω at \bar{x} contains those $u \in \mathbb{R}^d$ for which there are sequences $(t_k)_{k \in \mathbb{N}}$ in $(0, \infty)$ converging to 0, and $(u_k)_{k \in \mathbb{N}}$ in \mathbb{R}^d converging to u , such that $\bar{x} + t_k u_k \in \Omega$ whenever $k \in \mathbb{N}$;

$$T(\bar{x}; \Omega) := \left\{ u \in \mathbb{R}^d \mid \exists t_k \downarrow 0, u_k \rightarrow u \text{ with } \bar{x} + t_k u_k \in \Omega \right\}. \quad (3.1)$$

- (b) the *Bouligand paratingent cone* $\tilde{T}(\bar{x}; \Omega)$ to Ω at \bar{x} contains those $u \in \mathbb{R}^d$ for which there are sequences $(t_k)_{k \in \mathbb{N}}$ in $(0, \infty)$ converging to 0, $(u_k)_{k \in \mathbb{N}}$ in \mathbb{R}^d converging to u , and $(x_k)_{k \in \mathbb{N}}$ in Ω converging to \bar{x} , such that $x_k + t_k u_k \in \Omega$ whenever $k \in \mathbb{N}$; ¹

$$\tilde{T}(\bar{x}; \Omega) := \left\{ u \in \mathbb{R}^d \mid \exists t_k \downarrow 0, u_k \rightarrow u, x_k \xrightarrow{\Omega} \bar{x} \text{ with } x_k + t_k u_k \in \Omega \right\}. \quad (3.2)$$

After reviewing these definitions in the following graphical example, we will provide some general properties of tangent cones.

Example 3.1.2. We try to calculate the contingent and paratingent cones for two different sets in \mathbb{R}^2 at the reference point $(0, 0)$.

- (a) Consider the set $\Omega = \{(x, y) \in \mathbb{R}_+^2 \mid xy = 0\}$, as shown in Figure 3.1.

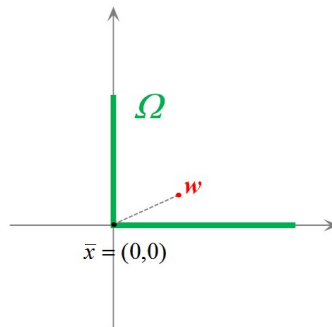


FIGURE 3.1: Calculating contingent and paratingent cones for a set

¹One should be aware that this definition is different from *Clarke tangent cone*, shown by $T_C(\bar{x}; \Omega)$ and defined as

$$T_C(\bar{x}; \Omega) := \left\{ u \in \mathbb{R}^d \mid \forall x_k \xrightarrow{\Omega} \bar{x}, \forall t_k \downarrow 0, \exists u_k \rightarrow u, \text{ with } x_k + t_k u_k \in \Omega \forall k \right\}.$$

To calculate the contingent cone, we must find points $u \in \mathbb{R}^2$, such that there exist sequences $t_k \downarrow 0$, and $u_k \rightarrow u$ with $t_k u_k \in \Omega$. Thus, any point $u \in \Omega$ would belong to the contingent cone (for example let $u_k = u$, and $t_k = \frac{1}{k}$). Any other point like w shown in the figure does not belong to $T((0,0);\Omega)$, since for any sequence $(w_k) \rightarrow w$, no matter how small the $t_k > 0$ would be, $t_k w_k$ would be out of Ω . Hence,

$$T((0,0);\Omega) = \mathbb{R}_+ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cup \mathbb{R}_+ \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

In order to compute $\tilde{T}((0,0);\Omega)$, first observe that by letting $x_k = x$, we get that $T((0,0);\Omega) \subset \tilde{T}((0,0);\Omega)$. We argue the set $\mathbb{R}_+ \begin{pmatrix} -1 \\ 0 \end{pmatrix}$ also belongs to the paratingent cone at $(0,0)$.

Let $u = (\lambda, 0)^T$ with $\lambda < 0$ be an arbitrary point. Then, by assuming $u_k = u$, $t_k = \frac{1}{-\lambda k}$, and $x_k = (\frac{2}{k}, 0)^T$ for every $k \in \mathbb{N}$, we get $x_k + t_k u_k = (\frac{1}{k}, 0)^T \in \Omega$.

A similar argument would prove that $\mathbb{R}_+ \begin{pmatrix} 0 \\ -1 \end{pmatrix} \subset \tilde{T}((0,0);\Omega)$. Hence,

$$\tilde{T}((0,0);\Omega) = \mathbb{R} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix}.$$

One may consider if Ω contains some points between the two lines, the situation would be different; in fact, it is true.

(b) This time we consider the convex set Ω in Figure 3.2 (left).

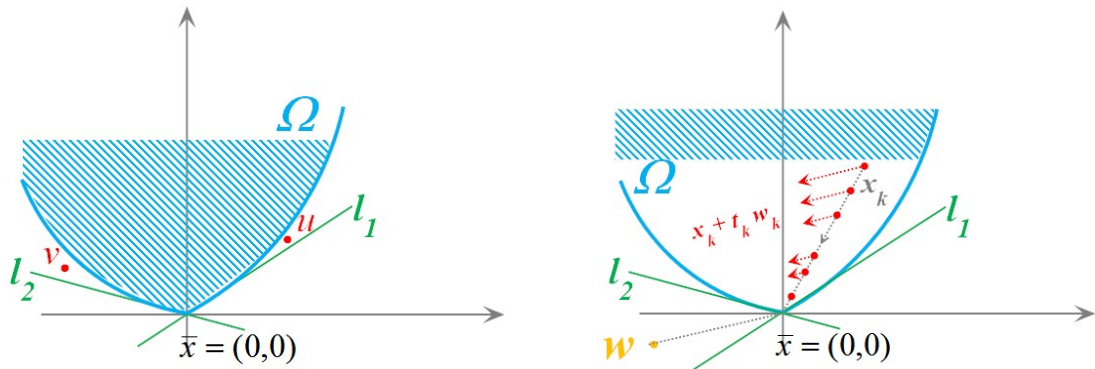


FIGURE 3.2: Calculating contingent and paratingent cones for a convex set

Consider l_1 , and l_2 to be the “tangent lines” to Ω at $\bar{x} = (0,0)$. Not only all points in Ω are inside the contingent cone, but also points like u , and v (shown in the figure) belong to $T(\bar{x};\Omega)$. The convexity of Ω , enables us to find a sequence $u_k \rightarrow u$, and a sequence of numbers t_k such that $t_k u_k$ is small enough to fit inside Ω , near \bar{x} .

Thus, the contingent cone to Ω at $(0,0)$ would be the whole area between line segments

l_1 , and l_2 . Assuming the slope of these lines to be m_1 , and m_2 , respectively, we can write

$$T((0, 0); \Omega) = \mathbb{R}_+ \begin{pmatrix} 1 \\ m_1 \end{pmatrix} \cup \mathbb{R}_+ \begin{pmatrix} 1 \\ m_2 \end{pmatrix} \cup \text{cone} \left\{ \begin{pmatrix} 1 \\ m_1 \end{pmatrix}, \begin{pmatrix} 1 \\ m_2 \end{pmatrix} \right\}.$$

It is expected to assume that the paratingent cone will include the negative part of this cone, too. What is more, the paratingent cone is even bigger.

For example, consider the point w (shown in the figure). Then, one can take a sequence $x_k \in \text{int } \Omega$ converging to $(0, 0)$ slower than the convergent of $w_k t_k$ to $(0, 0)$ (a possible choice is shown in Figure 3.2 (right)). Hence, $x_k + t_k w_k$ remains inside the Ω , and this reasoning reveals that $\tilde{T}((0, 0); \Omega) = \mathbb{R}^2$.

Proposition 3.1.3 (Some Properties of the Contingent Cone). [67, p. 232]

Let Ω be a non-empty subset of \mathbb{R}^d , and consider a point $\bar{x} \in \Omega$. Then the following properties for the contingent cone hold:

- (a) $T(\bar{x}; \Omega)$ is a cone, containing the zero element
- (b) $T(\bar{x}; \Omega)$ is closed;
- (c) if U is a neighborhood of \bar{x} , then $T(\bar{x}; \Omega) = T(\bar{x}; \Omega \cap U)$;
- (d) if Ω is convex, then $T(\bar{x}; \Omega) = \text{cl}(\mathbb{R}_+(\Omega - \bar{x}))$.

Statement (c) of the above proposition means that the contingent cone depends on the local properties of Ω near \bar{x} only.

It is time to define the “dual” concept for tangent cones.

Definition 3.1.4. (Normal Cones)

[54, p. 4][29, p. 229] Let Ω be a non-empty subset of \mathbb{R}^n and consider a point $\bar{x} \in \Omega$, at which Ω is locally closed.

- (a) The Fréchet/regular normal cone to Ω at \bar{x} is the set

$$\widehat{N}(\bar{x}; \Omega) := \left\{ v \in \mathbb{R}^n \mid \limsup_{x \xrightarrow{\Omega} \bar{x}} \frac{\langle v, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leq 0 \right\} \quad (3.3)$$

in which $x \xrightarrow{\Omega} \bar{x}$ means $x \rightarrow \bar{x}$ with $x \in \Omega$.

- (b) The Mordukhovich/general/limiting normal cone $N(\bar{x}; \Omega)$ to Ω at \bar{x} contains those $v \in \mathbb{R}^n$ for which there are sequences $(x_k)_{k \in \mathbb{N}}$ in Ω converging to \bar{x} , and $(v_k)_{k \in \mathbb{N}}$ in \mathbb{R}^n converging to v such that $v_k \in \widehat{N}(x_k; \Omega)$ for each $k \in \mathbb{N}$;

$$N(\bar{x}; \Omega) := \left\{ v \in \mathbb{R}^n \mid \exists v_k \rightarrow v, x_k \xrightarrow{\Omega} \bar{x} \text{ with } v_k \in \widehat{N}(x_k; \Omega) \right\}. \quad (3.4)$$

Remark 3.1.5. (a) We define $\widehat{N}(x; \Omega) = \emptyset$ for any $x \notin \Omega$.

(b) Very often, the limit process in the definition of the general normal cone $N(\bar{x}; \Omega)$ is

superfluous; no additional vectors v are produced in that manner, and one merely has $N(\bar{x}; \Omega) = \widehat{N}(\bar{x}; \Omega)$. This circumstance is termed as the *Clarke regularity*² of Ω at \bar{x} . Anyway, $N(\bar{x}; \Omega)$ is always a closed cone.

Example 3.1.6. Let us consider the set $\Omega = \{(x, y) \in \mathbb{R}_+^2 \mid xy = 0\}$. In Example 3.1.2 we calculated the contingent and paratingent cones at $\bar{x} = (0, 0)$. Now we take a look at the normal cones to Ω at \bar{x} .

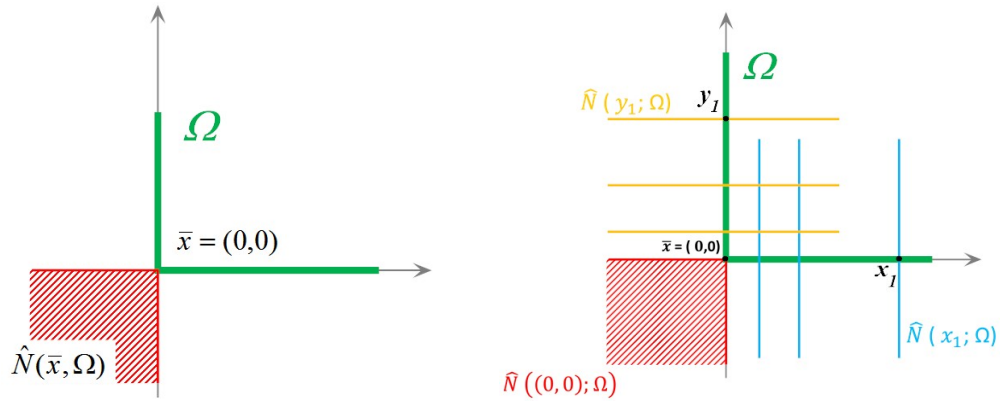


FIGURE 3.3: Regular and limiting normal cones to Ω at $\bar{x} = (0, 0)$

- (a) In the left image of Figure 3.3, all vectors $v \in \mathbb{R}^2$ for which $\langle v, x \rangle \leq 0$ for those $x \in \Omega$ that are near to $(0, 0)$, are shown with red. Thus,

$$\widehat{N}((0, 0); \Omega) = \mathbb{R}_+ \begin{pmatrix} 0 \\ -1 \end{pmatrix} \cup \mathbb{R}_+ \begin{pmatrix} -1 \\ 0 \end{pmatrix} \cup \text{cone} \left\{ \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\}.$$

- (b) In the right hand side of Figure 3.3, one can choose $x_k = \bar{x}$ ($\forall k$), to get $\widehat{N}(x_k; \Omega) = \widehat{N}((0, 0); \Omega)$. Another option is moving along the horizontal line in Ω , that will produce blue lines which converge to the vertical line passing through origin, as $k \rightarrow \infty$.

Moving along the vertical line in Ω , produces normals shown as orange lines which converge to the horizontal line passing through origin. Any other sequence $x_k \in \Omega$, that converges to \bar{x} , will produce a normal cone subset of the above ones. Hence,

$$N((0, 0); \Omega) = \mathbb{R} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cup \text{cone} \left\{ \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\}.$$

² **Definition (Clarke regularity of sets).** A set $C \subset \mathbb{R}^n$ is *regular* at one of its points \bar{x} in the sense of *Clarke* if it is locally closed at \bar{x} and every normal vector to C at \bar{x} is a regular normal vector, i.e., $N(\bar{x}; C) = \widehat{N}(\bar{x}; C)$ [65, p. 199].

Example 3.1.7. Given a differentiable convex function $h : \mathbb{R}^m \rightarrow \mathbb{R}$, consider the set $\Omega := \{u \in \mathbb{R}^m \mid h(u) \leq 0\}$. Then

$$N(u; \Omega) = \begin{cases} \{0\} & \text{if } h(u) < 0; \\ \{\lambda \nabla h(u) : \lambda \geq 0\} & \text{if } h(u) = 0; \\ \emptyset & \text{if } h(u) > 0. \end{cases}$$

Indeed, Ω is a convex set and one can easily deduce the direction of the normal vector at a point in the boundary of Ω from $\nabla h(\cdot)$ (since h is differentiable). The other two relations are obvious.

Proposition 3.1.8 (Some Properties of the Fréchet Normal Cone). [54, p. 17] [67, p. 35] Let $\Omega \subset \mathbb{R}^n$ be a non-empty set, and $\bar{x} \in \Omega$. Then,

- (a) the Fréchet normal cone $\widehat{N}(\bar{x}; \Omega)$ is a non-empty closed convex cone;
- (b) the Fréchet normal cone to Ω at \bar{x} is dual to the contingent cone to Ω at this point, that is,

$$\widehat{N}(\bar{x}; \Omega) = T(\bar{x}; \Omega)^\circ := \{x^* \in \mathbb{R}^n \mid \langle x^*, v \rangle \leq 0 \text{ whenever } v \in T(\bar{x}; \Omega)\}^3 \quad (3.5)$$

- (c) the Mordukhovich normal cone is not a convex set in general (like the case in Example 3.1.6), and it is not dual to any tangent cone.

Example 3.1.9. Consider the circuit in Figure 2.29 with an LED. In Example 2.3.1, we observed that the generalized equation has the following form:

$$B + AU \in F(U), \quad (3.6)$$

where $B = \begin{pmatrix} -v_s \\ +v_s \end{pmatrix}$, $A = \begin{pmatrix} -R & R \\ R & -R \end{pmatrix}$, and $F : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ is defined as

$$F \begin{pmatrix} u_1 \\ u_2 \end{pmatrix} := \begin{pmatrix} F_1(u_1) \\ F_2(u_2) \end{pmatrix} = \begin{pmatrix} F_1(u_1) \\ N(u_2; \mathbb{R}^+) \end{pmatrix}$$

Since the matrix A is not invertible and F_1 could not be expressed as a normal cone to a set, One can see it is not possible to write this generalized equation as a VI (Variational Inequality).

³ For a set A , subset of a normed vector space X , the (negative) *polar cone* of A is defined as

$$A^\circ := \{x^* \in X^* \mid \langle x^*, x \rangle \leq 0 \ \forall x \in A\},$$

while the *dual cone* to K , a non-empty closed convex cone in \mathbb{R}^m , is defined as

$$K^* := \{p \in \mathbb{R}^m \mid \langle p, v \rangle \geq 0 \text{ for all } v \in K\}.$$

For more information, look at [67, p. 35].

In other words, although one can change the role of voltage and current in the ideal model of diodes using the following fact

$$\boxed{v_{\text{diode}}^{\text{ideal}} \in N\left(i_{\text{diode}}^{\text{ideal}}; \mathbb{R}^+\right) \iff -i_{\text{diode}}^{\text{ideal}} \in N\left(-v_{\text{diode}}^{\text{ideal}}; \mathbb{R}^+\right)}, \quad (3.7)$$

in the practical model, or Zener diode case, one can only have the normal cone inclusion for the voltage, that is

$$\boxed{i_{\text{diode}}^{\text{non ideal}} \in N\left(v_{\text{diode}}^{\text{non ideal}}; [-\alpha, \beta]\right)}. \quad (3.8)$$

Thus, in order to use the VI setting, one needs to choose the variables in a way that everything reaches at $i_D = f(v_s, v_D, v_R, \dots)$ and $i_D \in F(v_D)$.

Before defining the “derivatives”, let us review few statements concerning product and sum operations with normal and tangent cones. Later, in Subsection 3.1.1, we would use these propositions to provide calculus rules.

Proposition 3.1.10 (Tangents and Normals to Product Sets). [65, p. 227]

With \mathbb{R}^n expressed as $\mathbb{R}^{n_1} \times \dots \times \mathbb{R}^{n_m}$, write $x \in \mathbb{R}^n$ as (x_1, \dots, x_m) with components $x_i \in \mathbb{R}^{n_i}$. If $C = C_1 \times \dots \times C_m$ for closed sets $C_i \in \mathbb{R}^{n_i}$, then at any $\bar{x} = (\bar{x}_1, \dots, \bar{x}_m)$ with $\bar{x}_i \in C_i$ one has

$$N(\bar{x}; C) = N(\bar{x}_1; C_1) \times \dots \times N(\bar{x}_m; C_m), \quad (3.9)$$

$$\widehat{N}(\bar{x}; C) = \widehat{N}(\bar{x}_1; C_1) \times \dots \times \widehat{N}(\bar{x}_m; C_m), \quad (3.10)$$

$$T(\bar{x}; C) \subset T(\bar{x}_1; C_1) \times \dots \times T(\bar{x}_m; C_m), \quad (3.11)$$

$$\widetilde{T}(\bar{x}; C) = \widetilde{T}(\bar{x}_1; C_1) \times \dots \times \widetilde{T}(\bar{x}_m; C_m). \quad (3.12)$$

Furthermore, if C is a regular set, then the inclusion for $T(\bar{x}; C)$ becomes an equality like the others.

Example 3.1.11. [65, p. 203] (Tangents and Normals to Convex Sets)

For a convex set $C \subset \mathbb{R}^n$ and any point $\bar{x} \in C$, one can observe that the generalized normal cone $N(\bar{x}; \Omega)$ agrees with the normal cone in the sense of convex analysis. That is,

$$N(\bar{x}; C) = \widehat{N}(\bar{x}; C) = \left\{ v \mid \langle v, x - \bar{x} \rangle \leq 0 \quad \text{for all } x \in C \right\}, \text{ and}$$

$$T(\bar{x}; C) = \text{cl} \left\{ w \mid \exists \lambda > 0 \text{ with } \bar{x} + \lambda w \in C \right\}.$$

Furthermore, C is regular at \bar{x} as long as C is locally closed at \bar{x} .

Example 3.1.12. [65, p. 204] (**Tangents and Normals to Boxes**)

Suppose $C = C_1 \times \cdots \times C_n$, where each C_j is a closed interval in \mathbb{R} (not necessarily bounded, perhaps just consisting of a single number). Then C is regular at every one of its points $\bar{x} = (\bar{x}_1, \dots, \bar{x}_n)$. Its tangent cone has the form

$$T(\bar{x}; C) = T(\bar{x}_1; C_1) \times \cdots \times T(\bar{x}_m; C_m), \text{ where}$$

$$T(\bar{x}_j; C_j) = \begin{cases} (-\infty, 0] & \text{if } \bar{x}_j \text{ is the right endpoint of } C_j, \\ [0, \infty) & \text{if } \bar{x}_j \text{ is the left endpoint of } C_j, \\ (-\infty, \infty) & \text{if } \bar{x}_j \text{ is an interior point of } C_j, \\ [0, 0] & \text{if } C_j \text{ is a singleton} \end{cases}$$

while its normal cones have the form

$$N(\bar{x}; C) = N(\bar{x}_1; C_1) \times \cdots \times N(\bar{x}_m; C_m), \text{ where}$$

$$N(\bar{x}_j; C_j) = \begin{cases} [0, \infty) & \text{if } \bar{x}_j \text{ is the right endpoint of } C_j, \\ (-\infty, 0] & \text{if } \bar{x}_j \text{ is the left endpoint of } C_j, \\ [0, 0] & \text{if } \bar{x}_j \text{ is an interior point of } C_j, \\ (-\infty, \infty) & \text{if } C_j \text{ is a singleton} \end{cases}$$

In particular, C is a closed convex set. The formulas in previous example relative to a tangent vector $w = (w_1, \dots, w_m)$ or a normal vector $v = (v_1, \dots, v_n)$ translate directly into the indicated requirements on the signs of the components w_j and v_j .

Proposition 3.1.13 (**Tangents and Normals under Set Addition**). [65, p. 230]

Let $C = C_1 + \cdots + C_m$ for closed sets $C_i \subset \mathbb{R}^n$. Then, at any point $\bar{x} \in C$, one has

$$T(\bar{x}; C) \supset \bigcup_{\substack{\bar{x}_1 + \cdots + \bar{x}_m = \bar{x} \\ \bar{x}_i \in C_i}} \left[T(\bar{x}_1; C_1) + \cdots + T(\bar{x}_m; C_m) \right],$$

$$\widehat{N}(\bar{x}; C) \subset \bigcap_{\substack{\bar{x}_1 + \cdots + \bar{x}_m = \bar{x} \\ \bar{x}_i \in C_i}} \left[\widehat{N}(\bar{x}_1; C_1) \cap \cdots \cap \widehat{N}(\bar{x}_m; C_m) \right].$$

Remark 3.1.14. The principal difference between tangential and normal approximations is that the former constructions (tangents) provide local approximations of sets in primal spaces, while the latter ones (normals) are defined in dual spaces carrying “dual” information for the study of the local behaviour.

Being applied to epigraphs of extended-real-valued functions and graphs of set-valued mappings, tangential approximations will generate corresponding *directional derivatives*/*-subderivatives* of functions and *graphical derivatives* of mappings, while normal approximations will relate to *subdifferentials* and *coderivatives*, respectively, as we would introduce them soon.

Definition 3.1.15. (Graphical Derivative)

[29, p. 215] For a mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a pair (x, y) with $y \in F(x)$, the *graphical derivative* of F at x for y is the mapping $DF(x|y) : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ whose graph is the tangent cone $T((x, y); \text{gph } F)$ to $\text{gph } F$ at (x, y) :

$$v \in DF(x|y)(u) \Leftrightarrow (u, v) \in T((x, y); \text{gph } F). \quad (3.13)$$

Thus, $v \in DF(x|y)(u)$ if and only if there exist sequences $u_k \rightarrow u$, $v_k \rightarrow v$ and $\tau_k \downarrow 0$ such that $y + \tau_k v_k \in F(x + \tau_k u_k)$ for all k .

Definition 3.1.16. (Coderivative)

[29, p. 229] For a mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a pair $(x, y) \in \text{gph } F$ at which $\text{gph } F$ is locally closed, the *coderivative* of F at x for y is the mapping $D^*F(x|y) : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ defined by

$$w \in D^*F(x|y)(z) \Leftrightarrow (w, -z) \in N((x, y); \text{gph } F). \quad (3.14)$$

Example 3.1.17. (Coderivatives of Differentiable Mappings)

[54, p. 45] Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be strictly differentiable at \bar{x} . Then

$$D^*f(\bar{x})(u) = \{\nabla f(\bar{x})^T u\} \quad \text{for all } u \in \mathbb{R}^m.$$

It is sufficient to show that for an arbitrary $u \in \mathbb{R}^m$ and any $w \in D^*f(\bar{x})(u)$, we have $w = \nabla f(\bar{x})^T u$. By Definitions 3.1.16 and 3.1.4, we have sequences $\epsilon_k \downarrow 0$, $x_k \rightarrow \bar{x}$, and $(w_k, u_k) \rightarrow (w, u)$ such that

$$\langle w_k, x - x_k \rangle - \langle u_k, f(x) - f(x_k) \rangle \leq \epsilon_k (\|x - x_k\| + \|f(x) - f(x_k)\|)$$

for all x close enough to x_k and all $k \in \mathbb{N}$. It follows from Definition 1.1.11 of strict differentiability that for any sequence $\tau_j \downarrow 0$ as $j \rightarrow \infty$ there is a sequence of neighborhoods U_j of \bar{x} with

$$\|f(x') - f(x) - \nabla f(\bar{x})(x' - x)\| \leq \tau_j \|x' - x\| \quad \text{for all } x, x' \in U_j, j \in \mathbb{N}.$$

This allows us to select a subsequence $\{k_j\}$ of \mathbb{N} such that for all $x, x_{k_j} \in U_{k_j}$, $j \in \mathbb{N}$,

$$\begin{aligned} \langle u_{k_j}, f(x) - f(x_{k_j}) \rangle &= \langle u_{k_j}, \nabla f(\bar{x})(x - x_{k_j}) + o(x - \bar{x}, x_{k_j} - \bar{x}) \rangle \\ &= \langle \nabla f(\bar{x})^T u_{k_j}, x - x_{k_j} \rangle + o(x - \bar{x}, x_{k_j} - \bar{x}) \langle u_{k_j}, \cdot \rangle \end{aligned}$$

where $o(\cdot, \cdot)$ is defined in a natural way, that is $\lim_{(x, x_{k_j}) \rightarrow (\bar{x}, \bar{x})} \frac{o(x - \bar{x}, x_{k_j} - \bar{x})}{\|(x - \bar{x}, x_{k_j} - \bar{x})\|} = 0$.

$$\begin{aligned} \langle w_{k_j} - \nabla f(\bar{x})^T u_{k_j}, x - x_{k_j} \rangle &\leq \epsilon_{k_j} (\|x - x_{k_j}\| + \|f(x) - f(x_{k_j})\|) + o(x - \bar{x}, x_{k_j} - \bar{x}) u_{k_j} \\ &\leq \epsilon_{k_j} (\|x - x_{k_j}\| + l \|x - x_{k_j}\|) + \tau_j \|u_{k_j}\| \|x - x_{k_j}\| \\ &\leq \tilde{\epsilon}_j \|x - x_{k_j}\| \end{aligned}$$

for all $x, x_{k_j} \in U_{k_j}$, $j \in \mathbb{N}$, where $\tilde{\epsilon}_j := (1 + l)(\epsilon_{k_j} + \tau_j \|u_{k_j}\|)$ and l denotes a Lipschitz constant of f around \bar{x} (existence of which is guaranteed by the strict differentiability of f around \bar{x}). The latter implies that

$$\|w_{k_j} - \nabla f(\bar{x})^T u_{k_j}\| \leq \tilde{\epsilon}_j \quad \text{for large } j \in \mathbb{N},$$

which gives $w = \nabla f(\bar{x})^T u$.

The inverse inclusion, which would be equivalent to show

$$(\nabla f(\bar{x})^T u, -u) \in N((\bar{x}, f(\bar{x})); \text{gph } f) \quad \text{for any } u \in \mathbb{R}^m,$$

is a direct consequence of the strict differentiability of f around \bar{x} and the definition of limiting normal cone.

Remark 3.1.18. One may wonder the competence of “*derivative*” in the last two definitions. Let us see what happens when F is single-valued, thus reducing to a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Suppose that f is strictly differentiable at x ; then, for $y = f(x)$, the graphical derivative $Df(x|y)$ is of course the linear mapping $Df(x)$ from \mathbb{R}^n to \mathbb{R}^m with matrix $\nabla f(x)$. In contrast, the coderivative $D^*f(x|y)$ comes out as the adjoint linear mapping $Df(x)^*$ from \mathbb{R}^m to \mathbb{R}^n with matrix $\nabla f(x)^T$.

Definition 3.1.19. (Strict Graphical Derivative)

[29, p. 238] For a mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ the *strict graphical derivative* mapping $\tilde{D}F(\bar{x}|\bar{y})$ at \bar{x} for \bar{y} , where $\bar{y} \in F(\bar{x})$, is defined as a mapping whose graph is the collection of vectors (u, v) for which there exist sequences $(x_k, y_k) \in \text{gph } F$, $(x_k, y_k) \rightarrow (\bar{x}, \bar{y})$, as well as $\tau_k \downarrow 0$ and $(u_k, v_k) \rightarrow (u, v)$ such that $(x_k + \tau_k u_k, y_k + \tau_k v_k) \in \text{gph } F$. Thus,

$$v \in \tilde{D}F(x|y)(u) \Leftrightarrow (u, v) \in \tilde{T}((x, y); \text{gph } F). \quad (3.15)$$

Note 3.1.20. There are many different ways to define the tangential and normal approximation cones. We only defined those cones that tend to be used in derivatives definitions. One can refer to [67, Chapter 11] to see other possibilities.

In a general Banach space setting, one has the possibility to define different versions

of coderivatives (cf. [54, Chapter 1]); but since they coincide in the case of finite dimensional spaces, we avoid introducing them. The one we defined here is referred to as *Normal Coderivative* in [54].

Remark 3.1.21. [29, p. 216] Note that the graphical differentiation comes from an operation on graphs, and the graph of a mapping F can be converted to the graph of its inverse F^{-1} just by interchanging variables, thus, we immediately have the rule that

$$D(F^{-1})(y|x) = DF(x|y)^{-1}.$$

3.1.1 Calculus Rules

The very next step would be to provide some calculus rules for the derivation operators we introduced recently. Being interested in generalized equations (of the form $f + F \ni 0$), we will focus only on the rules dealing with the sum of a function and a set-valued map. For possible general calculus rules one can refer to [14, 31, 33, 50, 71].

Proposition 3.1.22 (Sum Rule for Graphical Derivatives).

[29, p. 216] For a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ which is differentiable at x , a set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and any $y \in F(x)$, one has

$$D(f + F)(x|f(x) + y) = Df(x) + DF(x|y). \quad (3.16)$$

Proof. If $v \in D(f + F)(x|f(x) + y)(u)$ there exist sequences $\tau_k \downarrow 0$, $u_k \rightarrow u$ and $v_k \rightarrow v$ such that $(x, f(x) + y) + \tau_k(u_k, v_k) \in \text{gph}(f + F)$, that is,

$$f(x) + y + \tau_k v_k - f(x + \tau_k u_k) \in F(x + \tau_k u_k) \text{ for every } k.$$

By using the definition of the derivative for f , we get

$$y + \tau_k(v_k - Df(x)u_k) + o(\tau_k) \in F(x + \tau_k u_k).$$

Hence, by the definition of the graphical derivative, $v - Df(x)u \in DF(x|y)(u)$.

Conversely, if $v - Df(x)u \in DF(x|y)(u)$ then there exist sequences $\tau_k \downarrow 0$, $u_k \rightarrow u$ and $w_k \rightarrow v - Df(x)u$ such that $y + \tau_k w_k \in F(x + \tau_k u_k)$. By the differentiability of f , and letting $v_k = w_k + Df(x)u_k$, we get

$$y + f(x) + \tau_k v_k + o(\tau_k) \in (f + F)(x + \tau_k u_k)$$

which yields $v \in D(f + F)(x|f(x) + y)(u)$. □

Example 3.1.23. (Graphical Derivative for a Constraint System) [29, p. 217]

Consider a general constraint system of the form

$$f(x) - D \ni y, \quad (3.17)$$

for a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$, a set $D \subset \mathbb{R}^m$ and a parameter vector y ; and let \bar{x} be a solution of (3.17) for \bar{y} at which f is differentiable. Then for the (set-valued) mapping

$$G : x \mapsto f(x) - D, \quad \text{with } \bar{y} \in G(\bar{x}),$$

one has

$$DG(\bar{x}|\bar{y})(u) = Df(\bar{x})u - T(f(\bar{x}) - \bar{y}; D). \quad (3.18)$$

First observe that for the constant mapping $F := -D$, the definition of the graphical derivative gives

$$v \in DF(x|z)(u) \Leftrightarrow v \in T(z; -D) \Leftrightarrow -v \in T(-z; D).$$

Now, using the sum rule (3.16) at the reference point (\bar{x}, \bar{y}) gives the relation (3.18).

In the special case where $D = \mathbb{R}_-^s \times \{0\}^{m-s}$ with $f = (f_1, \dots, f_m)$, the constraint system (3.17) with respect to $y = (y_1, \dots, y_m)$ takes the form

$$\begin{cases} f_i(x) \leq y_i & \text{for } i = 1, \dots, s, \\ f_i(x) = y_i & \text{for } i = s + 1, \dots, m. \end{cases}$$

The graphical derivative formula (3.18) then says that a vector $v = (v_1, \dots, v_m)$ is in $DG(x|y)(u)$ if and only if

$$\begin{cases} \nabla f_i(x)u \leq v_i & \text{for } i = 1, \dots, s \text{ with } f_i(x) = y_i, \\ \nabla f_i(x)u = v_i & \text{for } i = s + 1, \dots, m. \end{cases}$$

For $i = 1, \dots, s$, one can easily calculate that $T(f_i(x) - y_i; \mathbb{R}_-) = \mathbb{R}_-$, and then use the product law (3.1.10) to obtain $T(f(x) - y; D)$.

Proposition 3.1.24 (Sum Rule for Strict Graphical Derivatives). *For a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ which is differentiable at x , a set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and any $y \in F(x)$, one has*

$$\tilde{D}(f + F)(x|f(x) + y) = Df(x) + \tilde{D}F(x|y).$$

Proof. The proof is very similar to the proof of Proposition 3.1.22. It is easy to check that for a differentiable function f at a point $x \in \text{dom } f$, the strict graphical derivative

$\tilde{D}f(x|f(x))$ is equal to its graphical derivative $Df(x|f(x))$, both equal to the linear mapping $Df(x)$ with matrix $\nabla f(x)$.

We first prove that $\tilde{D}(f+F)(x|f(x)+y) \subset Df(x) + \tilde{D}F(x|y)$.

Assume that $v \in \tilde{D}(f+F)(x|f(x)+y)(u)$. Then, $(u, v) \in \tilde{T}\left((x, f(x)+y); \text{gph}(f+F)\right)$, which in turn, means that there are sequences $(t_k)_{k \in \mathbb{N}}$ in $(0, \infty)$, $(v_k)_{k \in \mathbb{N}}$ in \mathbb{R}^d , $(u_k)_{k \in \mathbb{N}}$ in \mathbb{R}^d , and $(x_k, w_k)_{k \in \mathbb{N}}$ in $\text{gph}(f+F)$, converging to $0, v, u$, and $(x, f(x)+y)$ respectively, with $d = n + m$ such that

$$(x_k + t_k u_k, w_k + t_k v_k) \in \text{gph}(f+F).$$

For each k , choose $y_k \in F(x_k)$ such that $w_k = f(x_k) + y_k$. Then, $w_k + t_k v_k \in f(x_k + t_k u_k) + F(x_k + t_k u_k)$, yields $y_k + f(x_k) - f(x_k + t_k u_k) + t_k v_k \in F(x_k + t_k u_k)$. The differentiability of f implies $y_k + t_k(v_k - u_k Df(x_k)) + o(t_k) \in F(x_k + t_k u_k)$. By definition, this means $(u, v - Df(x)u) \in \tilde{T}\left((x, y); \text{gph} F\right)$. So, $v - Df(x)u \in \tilde{D}F(x|y)(u)$.

On the other hand, assume $v \in Df(x)u + \tilde{D}F(x|y)(u)$. Then, $(u, v - Df(x)u) \in \tilde{T}\left((x, y); \text{gph} F\right)$. This implies the existence of sequences $t_k \rightarrow 0$, $u_k \rightarrow u$, $w_k \rightarrow v - Df(x)u$, and $(x_k, y_k) \rightarrow (x, y)$ such that $y_k + t_k w_k \in F(x_k + t_k u_k)$. Let $v_k = w_k + Df(x_k)u_k$. Using the differentiability of f once more, we obtain

$$\begin{aligned} y_k + t_k v_k - f(x_k + t_k u_k) + f(x_k) + o(t_k) &\in F(x_k + t_k u_k), \\ y_k + f(x_k) + t_k v_k + o(t_k) &\in (f+F)(x_k + t_k u_k). \end{aligned}$$

By definition, this means $(u, v) \in \tilde{T}\left((x, y + f(x)); \text{gph}(f+F)\right)$. Thus, v belongs to $\tilde{D}(f+F)(x|y + f(x))(u)$, and the proof is complete. \square

Proposition 3.1.25 (Sum Rule for Coderivatives). [29, p. 232]

For a function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ which is strictly differentiable at \bar{x} and a mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ with $\bar{y} \in F(\bar{x})$, for all $u \in \mathbb{R}^m$ one obtains

$$D^*(f+F)(\bar{x} | f(\bar{x}) + \bar{y})(u) = Df(\bar{x})^* u + D^*F(\bar{x} | \bar{y})(u).^4 \quad (3.19)$$

Proof. Let $w \in D^*(f+F)(\bar{x} | f(\bar{x}) + \bar{y})(u)$. Then, $(w, -u) \in N((\bar{x}, \bar{z}); \text{gph} f+F)$ where $\bar{z} = f(\bar{x}) + \bar{y}$. By definition, there exist sequences $(w_k, u_k) \rightarrow (w, u)$, and $(x_k, z_k) \xrightarrow{\text{gph} f+F} (\bar{x}, \bar{z})$ with $z_k = f(x_k) + y_k$, such that for an arbitrary $\tau > 0$ one has

$$\langle (w_k, -u_k), (x - x_k, z - z_k) \rangle \leq \tau (\|x - x_k\| + \|z - z_k\|)$$

⁴ Once more, we want to draw the readers' attention to the distinction between Jacobian matrix and derivative linear mapping mentioned in 1.1.10. This should explain the difference of notations regarding this theorem with some references.

for x, z close enough to x_k, z_k , respectively. Thus,

$$\langle w_k, x - x_k \rangle - \langle u_k, z - z_k \rangle \leq \tau(\|x - x_k\| + \|z - z_k\|). \quad (3.20)$$

Now, we try to replace the terms involving z, \bar{z}, z_k with known variables of the theorem in the above equation and use strict differentiability of f (cf. Definition 1.1.11).

$$\begin{aligned} \langle u_k, z - z_k \rangle &= \langle u_k, f(x) + y - f(x_k) - y_k \rangle = \langle u_k, y - y_k \rangle + \langle u_k, f(x) - f(x_k) \rangle \\ &= \langle u_k, y - y_k \rangle + \langle u_k, Df(\bar{x})(x - x_k) + o((x, x_k) - (\bar{x}, \bar{x})) \rangle \\ &= \langle u_k, y - y_k \rangle + \langle u_k, Df(\bar{x})(x - x_k) \rangle + o((x - \bar{x}, x_k - \bar{x})) u_k \\ &= \langle u_k, y - y_k \rangle + \langle Df(\bar{x})^* u_k, x - x_k \rangle + o((x - \bar{x}, x_k - \bar{x})) u_k \end{aligned}$$

For the term $\|z - z_k\|$ on the left hand side of (3.20), again by strict differentiability of f (and maybe replacing $k \in \mathbb{N}$ with a subsequence k_j of \mathbb{N} such that for each $j \in \mathbb{N}$, the points x, x_{k_j} belong to a neighborhood U_j of \bar{x} ⁵), one gets

$$\begin{aligned} \|z - z_k\| &\leq \|y - y_k\| + \|f(x) - f(x_k)\| \\ &\leq \|y - y_k\| + \epsilon \|x - x_k\| + \|Df(\bar{x})\| \|x - x_k\| \end{aligned}$$

where $\epsilon > 0$ could be chosen arbitrarily small. Putting all calculations in (3.20), gives

$$\langle w_k - Df(\bar{x})^* u_k, x - x_k \rangle - \langle u_k, y - y_k \rangle \leq \tau'(\|x - x_k\| + \|y - y_k\|),$$

which is by definition of coderivative, equal to $w - Df(\bar{x})^* u \in D^*F(\bar{x} | \bar{y})(u)$ or $w \in Df(\bar{x})^* u + D^*F(\bar{x} | \bar{y})(u)$.

Applying the currently proved inclusion to the sum $(f + F) + (-f)$, we get

$$\begin{aligned} D^*(F)(\bar{x}|\bar{y})(u) &= D^*((f + F) + (-f))(\bar{x}|f(\bar{x}) + \bar{y} - f(\bar{x}))(u) \\ &\subset D^*(f + F)(\bar{x}|f(\bar{x})(u) + \bar{y}) + D^*(-f)(\bar{x}| - f(\bar{x}))(u) \\ &\subset D^*(f + F)(\bar{x}|f(\bar{x})(u) + \bar{y}) - D(f)^*(\bar{x}|f(\bar{x}))(u). \end{aligned}$$

which gives the opposite inclusion and thus establish the equality. \square

We will end this subsection by two chain rules, both for coderivatives of the composition of a single-valued function with a set-valued map. We will need these special chain rules when dealing with calculation of $D^*(BF(Cz))$, in which B and C are linear operators and F is the set-valued map representing the diodes $i - v$ characteristic.

In this way, instead of referring to the definition for calculating the coderivative in each special example, where B , and C could be different, we just calculate the coderivative of F once, and then use the tools provided here and those in Sections 3.2 and 3.5 to obtain

⁵ Details of this change is almost the same as Example 3.1.17, so we avoid repeating that.

the coderivative of the composition. The first chain rule would be about the coderivative of $f \circ G$.

Proposition 3.1.26 (Special Chain Rules for Coderivatives). [54, p. 146]

For mappings $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^d$ and $f : \mathbb{R}^n \times \mathbb{R}^d \rightarrow \mathbb{R}^m$, define

$$(f \circ G)(x) := f(x, G(x)) = \bigcup \{f(x, y) \mid y \in G(x)\}. \quad (3.21)$$

Given $\bar{x} \in \text{dom } G$, we assume that:

- (a) $f(x, \cdot) \in L(\mathbb{R}^d, \mathbb{R}^m)$ around \bar{x} , i.e., it is a linear operator from \mathbb{R}^d into \mathbb{R}^m . Moreover, $f(\bar{x}, \cdot)$ is injective.
- (b) The mapping $x \rightarrow f(x, \cdot)$ from \mathbb{R}^n into the operator space $L(\mathbb{R}^d, \mathbb{R}^m)$ is strictly differentiable at \bar{x} .

Take any $\bar{y} \in G(\bar{x})$ and denote $\bar{z} := f(\bar{x}, \bar{y})$. Then one has

$$D^*(f \circ G)(\bar{x} | \bar{z})(z^*) = \nabla_x f(\bar{x}, \bar{y})^* z^* + D^*G(\bar{x} | \bar{y})(f(\bar{x}, \cdot)^* z^*) \quad \text{for all } z^* \in \mathbb{R}^m. \quad (3.22)$$

As a counterpart of the previous proposition, the next chain rule would be about the coderivative of $G \circ f$. Thanks to the special structure of \mathbb{R}^m , both propositions find a simplified presentation here, comparing to the a general Banach space setting in [54, 56], where they were first stated and proved.

Proposition 3.1.27 (Special Chain Rules for Coderivatives).

[56, Theorem 3.10, p. 2669] Assume that $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is strictly differentiable at $\bar{x} \in \text{dom } f$ and Restrictive Metrically Regular (RMR)⁶ around this point. Let $G : \mathbb{R}^m \rightrightarrows \mathbb{R}^d$ be a set-valued mapping, such that $\bar{z} \in G(f(\bar{x}))$. Then

$$D^*(G \circ f)(\bar{x} | \bar{z}) = Df(\bar{x})^* \circ D^*G|_f(f(\bar{x}) | \bar{z}),$$

where $G|_f$ is the restriction of G on f (cf. Notation 1.2).

3.1.2 Inner and Outer Norms

This subsection starts with the introduction of positively homogeneous set-valued mappings; in particular, we will see that the derivatives defined at the beginning of this

⁶ **Definition (RMR).** A function $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$ is said to have the *RMR property* around \bar{x} , or f is RMR around this point, if the restrictive mapping $f : \mathbb{R}^n \rightarrow f(\mathbb{R}^n)$ between \mathbb{R}^n and the metric space $f(\mathbb{R}^n) \subset \mathbb{R}^m$, whose metric is induced by the norm on \mathbb{R}^m , is metrically regular around \bar{x} in the sense of Definition 1.3.1.

One can easily see, by the classical open mapping theorem, that for linear mappings f , the RMR property always holds since the subspace $f(\mathbb{R}^n)$ is closed in \mathbb{R}^m . However, the situation is much more complicated for nonlinear mappings when the RMR property may be violated even in the simplest cases as, for example, for $f(x) = x^2$ around $\bar{x} = 0 \in \mathbb{R}$ (for more details, refer to [56]).

section, are positively homogeneous. Then, we provide two *norm-like* operators which let us characterize, quantitatively, the behaviour of these maps (Proposition 3.1.35).

The rest of this section would be devoted to provide general statements relating the local stability or (metric) regularity of a set valued map to the (inner or outer) norm of its derivative.

Definition 3.1.28. (Positively Homogeneous Mappings)

[29, p. 216] A mapping $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ is called *positively homogeneous* when $\text{gph } H$ is a cone, which is equivalent to H satisfying

$$0 \in H(0) \quad \text{and} \quad H(\lambda x) = \lambda H(x) \quad \text{for } \lambda > 0.$$

Remark 3.1.29. The inverse of a positively homogeneous mapping is another positively homogeneous mapping. Linear mappings are positively homogeneous as a special case, their graphs being not just cones but linear subspaces.

Because graphical derivative mappings are positively homogeneous, the general properties of positively homogeneous mappings can be applied to them. Norm concepts are available in particular for capturing quantitative characteristics.

Definition 3.1.30. (Outer and Inner Norms)

[29, p. 218] For any positively homogeneous mapping $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, the *outer norm* and the *inner norm* are defined, respectively, by

$$|H|^+ := \sup_{\|x\| \leq 1} \sup_{y \in H(x)} \|y\|, \quad \text{and} \quad |H|^- := \sup_{\|x\| \leq 1} \inf_{y \in H(x)} \|y\| \quad (3.23)$$

with the convention $\inf_{y \in \emptyset} \|y\| = \infty$ and $\sup_{y \in \emptyset} \|y\| = -\infty$.

Note 3.1.31. When H is a linear mapping, both $|H|^+$ and $|H|^-$ reduce to the operator (matrix) norm $\|H\|$ associated with the Euclidean norm. However, it must be noted that neither $|H|^+$ nor $|H|^-$ satisfies the conditions in the definition of a true “norm”, inasmuch as in general, the positively homogeneous mappings $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ do not even form a vector space under addition and scalar multiplication.

Consider the “inverse elements of addition” axiom. One may define $-H$ in a natural way, that is for any $x \in \text{dom } H$, let $(-H)(x) := \{-y \in \mathbb{R}^m \mid y \in H(x)\}$. Then, of course

$$0_{\mathbb{R}^m} \in (H + (-H))(x) \quad \text{for all } x \in \text{dom}(H \cap -H),$$

but $H + (-H) \neq O$, the zero mapping, which is needed to be defined as

$$O(x) = \{0\} \quad \text{for all } x \in \mathbb{R}^n.$$

Nevertheless, some elementary rules are valid, like the following ones:

$$\begin{aligned} |\lambda H|^+ &= \|\lambda\| |H|^+, & |H_1 + H_2|^+ &\leq |H_1|^+ + |H_2|^+, & |H_2 \circ H_1|^+ &\leq |H_2|^+ |H_1|^+, \\ |\lambda H|^- &= \|\lambda\| |H|^-, & |H_1 + H_2|^- &\leq |H_1|^- + |H_2|^-, & |H_2 \circ H_1|^- &\leq |H_2|^- |H_1|^- . \end{aligned}$$

The following example shows that in general, some properties of the norm definition may not be satisfied as well.

Example 3.1.32. Define the set-valued mapping $H : \mathbb{R} \rightrightarrows \mathbb{R}$ with $H(x) = \{0, x\}$. We show that H is a positively homogeneous map with $|H|^- = 0$, and so the inner norm is not a true norm, since H is not the *zero mapping*.

First, observe that $H(0) = \{0\}$. Letting $\lambda > 0$, we obtain

$$H(\lambda x) = \{0, \lambda x\} = \lambda \{0, x\} = \lambda H(x).$$

Thus, H is positively homogeneous; but

$$|H|^- = \sup_{\|x\| \leq 1} \inf_{y \in H(x)} \|y\| = \sup_{\|x\| \leq 1} \inf \{0, \|x\|\} = 0$$

Note 3.1.33. For a positively homogeneous mapping $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, when $\text{dom } H = \mathbb{R}^n$, it would be a simple and immediate observation from the defining relation (3.23) and its conventions concerning the empty set, that

$$|H|^+ \geq |H|^-.$$

However, when $\text{dom } H \neq \mathbb{R}^n$, things are a bit tricky. For example, define $H : \mathbb{R} \rightrightarrows \mathbb{R}$ as

$$H(x) = \begin{cases} 0 & \text{for } x \geq 0, \\ \emptyset & \text{otherwise.} \end{cases}$$

Then, for $x \in [-1, 0)$, we get $\sup_{y \in H(x)} \|y\| = \sup_{y \in \emptyset} \|y\| = -\infty$. Thus,

$$|H|^+ = \sup_{\|x\| \leq 1} \sup_{y \in H(x)} \|y\| = \sup \{0, -\infty\} = 0.$$

While $|H|^- = \sup_{\|x\| \leq 1} \inf_{y \in H(x)} \|y\| = \sup \{0, \infty\} = \infty$ (for more details refer to [9]).

Example 3.1.34. It would be interesting to see the interpretations of the inner and outer norms, when $H = A^{-1}$ for a linear function $A : \mathbb{R}^n \rightarrow \mathbb{R}^m$. Let the $m \times n$ matrix for this linear function be denoted likewise by A , for simplicity.

- If $m < n$, we have A surjective (the associated matrix being of rank m) if and only if $|A^{-1}|^-$ is finite, this expression being the norm of the right inverse of A , that is,

$|A^{-1}|^- = \|A^T(AA^T)^{-1}\|$. Then $|A^{-1}|^+ = \infty$.

This fact is a result of the convention and the definition of inner and outer norms. To be more clear, when A is surjective, for any $y \in \mathbb{R}^m$, there exists $x \in \mathbb{R}^n$ such that $Ax = y$. That is, A^{-1} as a set-valued mapping has domain equal to \mathbb{R}^m

$$|H|^- = \sup_{\|y\| \leq 1} \inf_{x \in A^{-1}(y)} \|x\| < \infty, \text{ since it is not the } \sup_{x \in \emptyset} \inf_{x \in \emptyset} \|x\|.$$

And vice versa, when $|H|^- < \infty$, it means that $A^{-1}(y) \neq \emptyset$ for any $y \in \mathbb{R}^m$, that is, A is surjective.

One can easily observe that H is not single-valued, by reminding that the system of linear equations $Ax = y$, with number of equations less than the variables (components of x). Then there are infinite solutions corresponding to free variables. Thus

$$|H|^+ = \sup_{\|y\| \leq 1} \sup_{x \in A^{-1}(y)} \|x\| = \infty.$$

From linear algebra we know that since $m < n$, and A is surjective; it has a right inverse, that is, there exist a linear function $A_{\text{right}}^{-1} := A^T(AA^T)^{-1}$ such that $AA_{\text{right}}^{-1} = I_m$. Hence, $|A^{-1}|^- = \|A_{\text{right}}^{-1}\|$.

• On the other hand, if $m > n$, we have $|A^{-1}|^+ < \infty$ if and only if A is injective (the associated matrix has rank n), and then $|A^{-1}|^+ = |(A^T A)^{-1} A^T|$ but $|A^{-1}|^- = \infty$. Since $\text{dom } H \subsetneq \mathbb{R}^m$, then for some $y \in \mathbb{R}^m$, $H(y) = \emptyset$, and so, by convention

$$|H|^- = \sup_{\|y\| \leq 1} \inf_{x \in \emptyset} \|x\| = \infty.$$

But these points are not important for the outer norm calculations (since $\sup_{x \in \emptyset} \|x\| = -\infty$).

For the other points we have a single-valued map which is the left inverse of A , given by $A_{\text{left}}^{-1} := (A^T A)^{-1} A^T$ and thus the assertion yields.

• For $m = n$, of course, both norms agree with the usual matrix norm $\|A^{-1}\|$, and the finiteness of this quantity is equivalent to non-singularity of A .

The inner and outer norms have interesting properties specially when dealing with sub-linear mappings. We only mention the most needed ones here, however, the enthusiast reader can refer to [9] for more information.

Proposition 3.1.35 (Norm Characterizations). [29, p. 219]

The inner norm of a positively homogeneous mapping $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ satisfies

$$|H|^- = \inf \{ \kappa > 0 \mid H(x) \cap \kappa \mathbb{B} \neq \emptyset \quad \text{for all } x \in \mathbb{B} \}. \quad (3.24)$$

In parallel, the outer norm satisfies

$$|H|^+ = \inf \{ \kappa > 0 \mid y \in H(x) \Rightarrow |y| \leq \kappa|x| \} = \sup_{|y|=1} \frac{1}{d(0, H^{-1}(y))}. \quad (3.25)$$

If H has closed graph, then furthermore

$$|H|^+ < \infty \iff H(0) = \{0\}. \quad (3.26)$$

If H has closed and convex graph, then

$$|H|^- < \infty \iff \text{dom } H = \mathbb{R}^n, \quad (3.27)$$

and in that case $|H|^- < \infty$ if and only if H is surjective.

In order to prove the next theorem, we need a proposition which is a simple case of [29, Theorem 4B.3].

Proposition 3.1.36. *For a set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a point $(\bar{x}, \bar{y}) \in \text{gph } F$ such that $\text{gph } F$ is locally closed at it, and for any $c \in (0, \infty)$ satisfying*

$$\limsup_{\substack{(x,y) \rightarrow (\bar{x}, \bar{y}) \\ (x,y) \in \text{gph } F}} |DF(x|y)^{-1}|^- < c, \quad (3.28)$$

there are neighborhoods V of \bar{y} and U of \bar{x} such that

$$d(x, F^{-1}(y)) \leq c d(y, F(x)) \quad \text{for } x \in U, \text{ and } y \in V. \quad (3.29)$$

Proof. As mentioned above, this is a special case of [29, Theorem 4B.3, p. 222] where $G(p, x) = F(x) - p$, and y takes the place of p ; we have $S(y) = F^{-1}(y)$ and $d(0, G(y, x)) = d(y, F(x))$. The upper semicontinuity of $y \mapsto d(y, F(\bar{x}))$ is granted and the rest are simple modifications. \square

We start derivative description statements with derivative criteria for metric regularity, in terms of graphical derivative, in the following theorem and coderivative, in the next theorem.

Theorem 3.1.37 (Graphical Derivative Criterion for Metric Regularity).

[29, p. 221] *For a mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a point $(\bar{x}, \bar{y}) \in \text{gph } F$ at which the $\text{gph } F$ is locally closed, one has*

$$\text{reg}(F; \bar{x}|\bar{y}) = \limsup_{\substack{(x,y) \rightarrow (\bar{x}, \bar{y}) \\ (x,y) \in \text{gph } F}} |DF(x|y)^{-1}|^- \quad (3.30)$$

Thus, F is metrically regular at \bar{x} for \bar{y} if and only if the right side of (3.30) is finite.

Proof. For short, let d_{DF} denote the right side of (3.30).

We will start by showing that $\text{reg}(F; \bar{x}|\bar{y}) \leq d_{DF}$. If $d_{DF} = \infty$, there is nothing to prove.

Let $d_{DF} < c < \infty$. Applying the previous proposition, Condition (3.29), becomes the definition of metric regularity of F at \bar{x} for \bar{y} , and therefore $\text{reg}(F; \bar{x}|\bar{y}) \leq c$. Since c can be taken arbitrarily close to d_{DF} we conclude that $\text{reg}(F; \bar{x}|\bar{y}) \leq d_{DF}$.

Now, we turn to demonstrating the opposite inequality,

$$\text{reg}(F; \bar{x}|\bar{y}) \geq d_{DF}. \quad (3.31)$$

If $\text{reg}(F; \bar{x}|\bar{y}) = \infty$ we are done. Suppose therefore that F is metrically regular at \bar{x} for \bar{y} with respect to a constant κ and neighborhoods U for \bar{x} and V for \bar{y} . Then

$$d(x', F^{-1}(y)) \leq \kappa d(y, F(x')) \leq \kappa \|y - y'\|, \quad (3.32)$$

whenever $(x', y') \in \text{gph } F$, $x' \in U$, $y \in V$.

We know from Proposition 1.2.7 that V can be chosen so small that $F^{-1}(y) \cap U \neq \emptyset$ for every $y \in V$. Pick any $y' \in V$ and $x' \in F^{-1}(y') \cap U$, and let $v \in \mathbb{B}$. Take a sequence $\tau_k \downarrow 0$ such that $y_k := y' + \tau_k v \in V$ for all k . By (3.32) and the local closedness of $\text{gph } F$ at (\bar{x}, \bar{y}) there exists $x_k \in F^{-1}(y' + \tau_k v)$ such that

$$\|x' - x_k\| = d(x', F^{-1}(y_k)) \leq \kappa \|y_k - y'\| = \kappa \tau_k \|v\|.$$

For $u_k := (x_k - x')/\tau_k$ we obtain

$$\|u_k\| \leq \kappa \|v\|. \quad (3.33)$$

Thus, u_k is bounded, so $u_{k_i} \rightarrow u$ for a subsequence $k_i \rightarrow \infty$. Since $(x_{k_i}, y' + \tau_{k_i} v) \in \text{gph } F$, we obtain $(u, v) \in T((x', y'); \text{gph } F)$. Hence, by the definition of the graphical derivative, we have $u \in DF^{-1}(y'|x')(v) = DF(x'|y')^{-1}(v)$. The bound (3.33) guarantees that

$$|DF(x|y)^{-1}|^- \leq \kappa.$$

Since $(x, y) \in \text{gph } F$ is arbitrarily chosen near (\bar{x}, \bar{y}) , and κ is independent of this choice, we conclude that (3.31) holds and hence we have (3.30). \square

Theorem 3.1.38 (Coderivative Criterion for Metric Regularity). [29, p. 232]

For a mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a pair $(\bar{x}, \bar{y}) \in \text{gph } F$ at which $\text{gph } F$ is locally closed, one has

$$\text{reg}(F; \bar{x}|\bar{y}) = |D^*F(\bar{x}|\bar{y})^{-1}|^+. \quad (3.34)$$

Thus, F is metrically regular at \bar{x} for \bar{y} if and only if the right side of (3.34) is finite, which is equivalent to

$$D^*F(\bar{x}|\bar{y})(u) \ni 0 \implies u = 0.$$

Proof. Indeed, the proof is a consequence of the following equality and the previous theorem (Theorem 3.1.37).

(Basic Equality) Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a set-valued map, let $\bar{y} \in F(\bar{x})$, and assume that $\text{gph } F$ is locally closed at (\bar{x}, \bar{y}) . Then

$$\limsup_{\substack{(x,y) \rightarrow (\bar{x}, \bar{y}) \\ (x,y) \in \text{gph } F}} |DF(x|y)^{-1}|^- = |D^*F(\bar{x}|\bar{y})^{-1}|^+. \quad (3.35)$$

For the proof of this equality one can refer to [29, Theorem 4C.3, p. 233]. \square

Considering Theorem 1.3.5, describing the equality of metric regularity of F and Aubin property of F^{-1} at the proper reference points, the following theorem would be another way of expressing the previous criteria.

Theorem 3.1.39 (Graphical Derivative Criterion for Aubin Property).

[29, p. 222] For a mapping $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ and a point $(\bar{y}, \bar{x}) \in \text{gph } S$ at which $\text{gph } S$ is locally closed, one has

$$\text{lip}(S; \bar{x}|\bar{y}) = \limsup_{\substack{(y,x) \rightarrow (\bar{y}, \bar{x}) \\ (y,x) \in \text{gph } S}} |DS(y|x)|^-. \quad (3.36)$$

Thus, S has the Aubin property at \bar{y} for \bar{x} if and only if the right side of (3.36) is finite.

Next, we provide a criterion for strong metric sub-regularity in terms of graphical derivative, and as a consequence of Theorem 1.3.17, a corollary about isolated calmness.

Theorem 3.1.40 (Graphical Derivative Criterion for Strong Metric Sub-regularity). [29, p. 246]

A mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ whose graph is locally closed at $(\bar{x}, \bar{y}) \in \text{gph } F$ is strongly metrically sub-regular at \bar{x} for \bar{y} if and only if

$$DF(\bar{x}|\bar{y})^{-1}(0) = \{0\}, \quad (3.37)$$

this being equivalent to

$$|DF(\bar{x}|\bar{y})^{-1}|^+ < \infty, \quad (3.38)$$

and in that case

$$\text{subreg}(F; \bar{x}|\bar{y}) = |DF(\bar{x}|\bar{y})^{-1}|^+. \quad (3.39)$$

Corollary 3.1.41 (Graphical Derivative Criterion for Isolated Calmness).

[29, p. 246] For a mapping $S : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ and a point $(\bar{y}, \bar{x}) \in \text{gph } S$ at which $\text{gph } S$ is

locally closed, one has

$$\text{clm}(S; \bar{y} | \bar{x}) = |DS(\bar{y} | \bar{x})|^+. \quad (3.40)$$

The last statement of this subsection would be a criterion for strong metric regularity in terms of strict graphical derivative.

Theorem 3.1.42 (Strict Graphical Derivative Criterion for Strong Metric Regularity). [29, p. 238]

Consider a set-valued mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and $(\bar{x}, \bar{y}) \in \text{gph } F$. If F is strongly metrically regular at \bar{x} for \bar{y} , then

$$|\tilde{D}F(\bar{x} | \bar{y})^{-1}|^+ < \infty. \quad (3.41)$$

On the other hand, if the graph of F is locally closed at (\bar{x}, \bar{y}) and

$$\bar{x} \in \liminf_{y \rightarrow \bar{y}} F^{-1}(y), \quad (3.42)$$

then condition (3.41) is also sufficient for strong metric regularity of F at \bar{x} for \bar{y} . In this case the quantity on the left side of (3.41) equals to $\text{reg}(F; \bar{x} | \bar{y})$.

Proof. Proposition 1.3.11 says that a mapping F is strongly metrically regular at \bar{x} for \bar{y} if and only if it is metrically regular there and F^{-1} has a localization around \bar{y} for \bar{x} which is nowhere multivalued. Furthermore, in this case for every $c > \text{reg}(F; \bar{x} | \bar{y})$ there exists a neighborhood V of \bar{y} such that F^{-1} has a localization around \bar{y} for \bar{x} which is a Lipschitz continuous function on V with constant c .

Let F be strongly metrically regular at \bar{x} for \bar{y} , let $c > \text{reg}(F; \bar{x} | \bar{y})$ and let U and V be open neighborhoods of \bar{x} and \bar{y} , respectively, such that the localization $V \ni y \mapsto \varphi(y) := F^{-1}(y) \cap U$ is a Lipschitz continuous function on V with a Lipschitz constant c .

We will show first that for every $v \in \mathbb{R}^m$ the set $\tilde{D}F(\bar{x} | \bar{y})^{-1}(v)$ is non-empty. Let $v \in \mathbb{R}^m$. Since $\text{dom } \varphi \supset V$, we can choose sequences $\tau_k \downarrow 0$ and u_k such that $\bar{x} + \tau_k u_k = \varphi(\bar{y} + \tau_k v)$ for large k . Then, from the Lipschitz continuity of φ with Lipschitz constant c we conclude that $\|u_k\| \leq c \|v\|$, hence u_k has a cluster point u which, by definition, is from $\tilde{D}F(\bar{x} | \bar{y})^{-1}(v)$.

Now, choose any $v \in \mathbb{R}^m$ and $u \in \tilde{D}F(\bar{x} | \bar{y})^{-1}(v)$; then, there exist sequences $(x_k, y_k) \in \text{gph } F$, $(x_k, y_k) \rightarrow (\bar{x}, \bar{y})$, $\tau_k \downarrow 0$, $u_k \rightarrow u$ and $v_k \rightarrow v$ such that $y_k + \tau_k v_k \in V$, $x_k = \varphi(y_k)$ and $x_k + \tau_k u_k = \varphi(y_k + \tau_k v_k)$ for k sufficiently large. But then, again from the Lipschitz continuity of φ with Lipschitz constant c , we obtain that $\|u_k\| \leq c \|v_k\|$. Passing to the limit we conclude that $\|u_k\| \leq c \|v\|$, which implies that

$$|\tilde{D}F(\bar{x} | \bar{y})^{-1}|^+ \leq c.$$

Hence (3.41) is satisfied; moreover, the quantity on the left side of (3.41) is less than or equal to $\text{reg}(F; \bar{x} | \bar{y})$.

To prove the second statement, we first show that F^{-1} has a single-valued bounded localization, that is there exist a bounded neighborhood U of \bar{x} and a neighborhood V of \bar{y} such that $V \ni y \mapsto F^{-1}(y) \cap U$ is single valued.

On the contrary, assume that for any bounded neighborhood U of \bar{x} and any neighborhood V of \bar{y} the intersection $\text{gph } F^{-1} \cap (V \times U)$ is the graph of a multivalued mapping. This means that there exist sequences $\epsilon_k \downarrow 0, x_k \rightarrow \bar{x}, x'_k \rightarrow \bar{x}, x_k \neq x'_k$ for all k such that

$$F(x_k) \cap F(x'_k) \cap \mathbb{B}_{\epsilon_k}(\bar{y}) \neq \emptyset \quad \text{for all } k.$$

Let $t_k = \|x_k - x'_k\|$ and let $u_k = \frac{x_k - x'_k}{t_k}$. Then $t_k \downarrow 0$ and $\|u_k\| = 1$ for all k . Hence $\{u_k\}$ has a cluster point $u \neq 0$. Consider any $y_k \in F(x_k) \cap F(x'_k) \cap \mathbb{B}_{\epsilon_k}(\bar{y})$. Then, $y_k + t_k 0 \in F(x'_k + t_k u_k)$ for all k . By the definition of the strict graphical derivative, $0 \in \tilde{D}F(\bar{x} | \bar{y})(u)$. Hence $|\tilde{D}F(\bar{x} | \bar{y})^{-1}|^+ = \infty$, which contradicts (3.41). Thus, there exist neighborhoods U of \bar{x} and V of \bar{y} such that $\varphi(y) := F^{-1}(y) \cap U$ is at most single-valued on V , and U is bounded. By assumption (3.42), there exists a neighborhood $V' \subset V$ of \bar{y} such that $F^{-1}(y) \cap U \neq \emptyset$ for any $y \in V'$, hence $V' \subset \text{dom } \varphi$. Further, since $\text{gph } F$ is locally closed at (\bar{x}, \bar{y}) and φ is bounded, there exists an open neighborhood $V'' \subset V'$ of \bar{y} such that φ is a continuous function on V'' .

From the definition of the strict graphical derivative we obtain that the set-valued mapping $(x, y) \mapsto \tilde{D}F(x | y)$ has closed graph. We claim that condition (3.41) implies that

$$\limsup_{\substack{(x,y) \rightarrow (\bar{x}, \bar{y}) \\ (x,y) \in \text{gph } F}} |\tilde{D}F(x | y)^{-1}|^+ < \infty. \quad (3.43)$$

On the contrary, assume that there exist sequences $(x_k, y_k) \in \text{gph } F$ converging to (\bar{x}, \bar{y}) , $v_k \in \mathbb{B}$ and $u_k \in \tilde{D}F(x_k | y_k)^{-1}(v_k)$ such that $\|u_k\| > k \|v_k\|$.

Case 1. There exists a subsequence $v_{k_i} = 0$ for all k_i . Since $\text{gph } \tilde{D}F(x_{k_i} | y_{k_i})^{-1}$ is a cone, we may assume that $\|u_{k_i}\| = 1$. Let u be a cluster point of u_{k_i} . Then, passing to the limit we get $0 \neq u \in \tilde{D}F(\bar{x} | \bar{y})^{-1}(0)$ which, combined with formula (3.25) in Proposition 3.1.35, contradicts (3.41).

Case 2. For all large k , $v_k \neq 0$. Since $\text{gph } \tilde{D}F(x_k | y_k)^{-1}$ is a cone, we may assume that $\|v_k\| = 1$. Then, $\lim_{k \rightarrow \infty} \|u_k\| = \infty$. Define

$$w_k := \frac{u_k}{\|u_k\|} \in \tilde{D}F(x_k | y_k)^{-1} \left(\frac{v_k}{\|u_k\|} \right)$$

and let w be a cluster point of w_k . Then, passing to the limit we obtain $0 \neq w \in \tilde{D}F(\bar{x} | \bar{y})^{-1}(0)$ which, combined with formula (3.25), again contradicts (3.41).

Hence (3.43) is satisfied. Therefore, there exists an open neighborhood $\tilde{V} \subset V''$ of \bar{y} such that $|\tilde{D}F(\varphi(y)|y)^{-1}|^+ < \infty$ for all $y \in \tilde{V}$. We will now prove that for every $(x, y) \in \text{gph } F$ near (\bar{x}, \bar{y}) and every $v \in \mathbb{R}^m$ we have that $DF(x|y)^{-1}(v) \neq \emptyset$. Fix $(x, y) \in \text{gph } F \cap (U \times \tilde{V})$ and $v \in \mathbb{R}^m$, and let $t_k \downarrow 0$; then there exist $u_k \in \mathbb{R}^n$ such that $x + t_k u_k = F^{-1}(y + t_k v) \cap U = \varphi(y + t_k v)$ for all large k and we also have that $t_k u_k \rightarrow 0$ by the continuity of φ . Assume that $\|u_k\| \rightarrow \infty$ for some subsequence (which is denoted in the same way without loss of generality). Set $\tau_k = t_k \|u_k\|$ and $w_k = \frac{u_k}{\|u_k\|}$. Then $\tau_k \downarrow 0$ and, for a further subsequence, $w_k \rightarrow w$ for some w with $\|w\| = 1$. Since $(y + \frac{\tau_k}{\|u_k\|}v, x + \tau_k w_k) \in \text{gph } F^{-1}$, we obtain that $w \in DF(x|y)^{-1}(0) \subset \tilde{D}F(x|y)^{-1}(0)$ for some $w \neq 0$.

Thus, $|\tilde{D}F(x|y)^{-1}|^+ = \infty$ contradicting the choice of \tilde{V} . Hence the sequence $\{u_k\}$ cannot be unbounded and since $y + t_k v \in F(x + t_k u_k)$ for all k , any cluster point u of $\{u_k\}$ satisfies $u \in DF(x|y)^{-1}(v)$. Hence, $DF(x|y)^{-1}$ is non-empty-valued. From this, Note 3.1.31, and the inclusion $DF(x|y)^{-1}(v) \subset \tilde{D}F(x|y)^{-1}(v)$ we obtain

$$|DF(x|y)^{-1}|^- \leq |\tilde{D}F(x|y)^{-1}|^+. \quad (3.44)$$

Putting together (3.43) and (3.44), and utilizing the derivative criterion for metric regularity in Theorem 3.1.37, we obtain that F is metrically regular at \bar{x} for \bar{y} with $\text{reg}(F; \bar{x}|\bar{y})$ bounded by the quantity on the left side of (3.43). But since F^{-1} has a single-valued localization at \bar{y} for \bar{x} , we conclude that F is strongly metrically regular at \bar{x} for \bar{y} . Moreover, $\text{reg}(F; \bar{x}|\bar{y}) = |\tilde{D}F(x|y)^{-1}|^+$. The proof is complete. \square

3.1.3 Subdifferentials

In this subsection we will have a very short look into the theory of subdifferentials. We have two reasons for that; first of all, though it is undesired, it is still probable that we encounter problems in which the single-valued function f is not smooth enough. Secondly, the subdifferential of a function is (in general) a set-valued map itself and it has the same idea as graphical derivatives we already introduced in this section.

To be more clear, as we already saw in Chapter 2, we need to deal with generalized equations of the form $f + F \ni 0$, that is, sum of a function and a set-valued map. In the previous subsections of this chapter, we tried to provide a tool (based on derivatives of f , and F) for studying the local stability properties of the solution of the GE.

We would see in the next sections of this chapter that continuous differentiability or twice differentiability of f is part of the assumptions of most of our theorems. Thus, one may get worried, or at least curious to know what will happen when f does not satisfy such conditions.

Let us recall that the single valued function in our model not only represents the combination of all components of circuit apart from diodes and transistors, but also by using some simplification ideas we extract some parts of the graph of the set-valued map and add it to f (refer to Figure 2.25 and explanations there).

The concept of subdifferential (for functionals, and subgradient for functions to \mathbb{R}^n) is an appropriate substitute to generalize the derivative concept to functions which are not differentiable.

According to [67], the theory of the subdifferential and the conjugate of convex functionals as well as its various applications originated in the work of Moreau and Rockafellar in the early 1960s. The Rockafellar's *Convex Analysis* [63] is one of the classic texts on the subject in finite-dimensional spaces.

Definition 3.1.43. (Subdifferentials)

[67, p. 167, p. 288] Given a lower semicontinuous⁷ function $f : \mathbb{R}^d \rightarrow \overline{\mathbb{R}} := \mathbb{R} \cup \{\pm\infty\}$ and a point $\bar{x} \in \mathbb{R}^d$ such that $f(\bar{x}) \in \mathbb{R}$,

- (a) the functional f is said to be *Fréchet subdifferentiable* (*F-subdifferentiable*) at \bar{x} if there exists $\xi \in \mathbb{R}^d$, the *F-subderivative* of f at \bar{x} , such that

$$\liminf_{h \rightarrow 0} \frac{f(\bar{x} + h) - f(\bar{x}) - \langle \xi, h \rangle}{\|h\|} \geq 0. \quad (3.45)$$

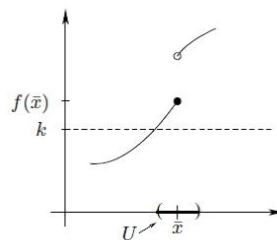
The set of all F-subderivatives of f at \bar{x} is called *Fréchet subdifferential* (*F-subdifferential*) of f at \bar{x} , and is shown with $\partial_F f(\bar{x})$.

- (b) the *limiting (Mordukhovich) subdifferential* of f at \bar{x} is the set $\partial f(\bar{x})$ containing all $\xi \in \mathbb{R}^d$ such that there are sequences $(x_k)_{k \in \mathbb{N}}$ and $(\xi_k)_{k \in \mathbb{N}}$ converging to \bar{x} and ξ , respectively, with

$$f(x_k) \rightarrow f(\bar{x}) \text{ as } k \rightarrow \infty, \text{ and } \xi_k \in \partial_F f(x_k) \text{ for each } k \in \mathbb{N}.$$

⁷ **Definition (lower semicontinuity).** The functional $f : \mathbb{R}^d \rightarrow \overline{\mathbb{R}}$ is called *lower semicontinuous* (*l.s.c.*) at $\bar{x} \in \text{dom } f$ if either $f(\bar{x}) = -\infty$ or for every $k < f(\bar{x})$ there exists a neighborhood U of \bar{x} such that

$$k < f(x) \text{ for all } x \in \text{dom } f \cap U.$$



The figure gives an idea of lower semicontinuity property [67, p. 22].

- (c) the *outer subdifferential* of f at \bar{x} is the set $\partial_{>}f(\bar{x})$, containing those $\xi \in \mathbb{R}^d$ for which, there are sequences $(x_k)_{k \in \mathbb{N}}$ and $(\xi_k)_{k \in \mathbb{N}}$ converging to \bar{x} and ξ , respectively, with

$$f(x_k) \downarrow f(\bar{x}) \text{ as } k \rightarrow \infty, \text{ and } \xi_k \in \partial_F f(x_k) \text{ for each } k \in \mathbb{N}.$$

Note 3.1.44. If $f(\bar{x})$ is infinite, then all the above subdifferentials of f at \bar{x} are defined to be empty sets.

Note that nothing will change if for each $k \in \mathbb{N}$ one will take ξ_k from $\partial f(x_k)$ instead of $\partial_F f(x_k)$ in the definition of the outer subdifferential.

Furthermore, note that the choice of any other (equivalent) norm on \mathbb{R}^d (instead of the usual one induced by the scalar product) does not affect the above subdifferential constructions.

When the function f is in \mathbb{R}^n we can follow the similar procedure and define generalized Jacobians.

Definition 3.1.45. (Generalized Jacobians) [22, p. 70]

- (a) The *Bouligand's limiting Jacobian* of a locally Lipschitz continuous function $h : \mathbb{R}^l \rightarrow \mathbb{R}^d$ at a point $\bar{u} \in \mathbb{R}^l$, is the (non-empty compact) set $\partial_B h(\bar{u})$, consisting of all matrices $A \in \mathbb{R}^{d \times l}$ for which there is a sequence $(u_n)_{n \in \mathbb{N}}$ converging to \bar{u} such that h is differentiable at each u_n and $\nabla h(u_n) \rightarrow A$ as $n \rightarrow \infty$.
- (b) The *Clarke's generalized Jacobian* of h at \bar{u} , denoted by $\partial h(\bar{u})$, is the convex hull of $\partial_B h(\bar{u})$.

3.2 Results about Aubin Property

In this section and the rest of this chapter we would try to focus on the setting we find convenient for the study of electronic circuits in the static case in Chapter 2, that is

$$\begin{array}{l} \Phi(z) := f(z) + BF(Cz) \\ S(p) := \{z \mid p \in \Phi(z)\} \end{array} \quad (3.46)$$

where $p \in \mathbb{R}^n$ is a fixed vector, $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ is a function, and $F : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$ is a set-valued map (with certain assumptions), and $B \in \mathbb{R}^{n \times m}$, $C \in \mathbb{R}^{m \times n}$ are given matrices with $m \leq n$, unless otherwise is stated.

The problem is how we can provide some handy theorems in order to check the local stability properties of the of solution mapping in (3.46). We will use the derivative criteria of Subsection 3.1.2, and the calculus rules of Subsection 3.1.1.

Let us provide the general description of the process for Aubin property here. The starting point would be a point-based description of the Aubin property.

Theorem 3.2.1 (Point-based Characterizations of Aubin Property).⁸

Let $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be a set-valued mapping with closed-graph around $(\bar{x}, \bar{y}) \in \text{gph } F$. Then the following properties are equivalent:

- (a) F has the Aubin property around (\bar{x}, \bar{y}) ;
- (b) $\|D^*F(\bar{x}, \bar{y})\| < \infty$;
- (c) $D^*F(\bar{x}, \bar{y})(0) = \{0\}$.

Moreover, in this case for the exact Lipschitzian bound of F around (\bar{x}, \bar{y}) , one has

$$\text{lip}F(\bar{x}, \bar{y}) = \|D^*F(\bar{x}, \bar{y})\|. \quad (3.47)$$

Proof. One only needs to combine Theorem 1.3.5 (concerning equivalence relation between metric regularity of F^{-1} and the Aubin property of F) with the outer norm characterization in Proposition 3.1.35 and the coderivative criterion for metric regularity in Theorem 3.1.38. \square

Note 3.2.2. Considering the property (c) in the above theorem, we try to go step by step from D^*S to D^*F using some calculus rules and adapting them to our situation. Almost the same procedure would be done to obtain criteria for other local stability properties or for the metric regularity counterparts.

For the setting (3.46), let us define the mappings $Q : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, and $F_C : \mathbb{R}^m \rightrightarrows \mathbb{R}^m$, by

$$\begin{aligned} Q(z) &:= BF(Cz), \\ F_C(u) &:= \begin{cases} F(u) & \text{if } u = Cz \text{ for some } z \in \mathbb{R}^n, \\ \emptyset & \text{otherwise.} \end{cases} \end{aligned} \quad (3.48)$$

We suppose that we have in hand a point $(\bar{z}, \bar{p}) \in \text{gph } \Phi$. Put $\bar{v} := (B^T B)^{-1} B^T (\bar{p} - f(\bar{z}))$. We may also need to refer to the following assumptions:

⁸ This theorem has been proved by Mordukhovich (hence known as Mordukhovich criterion in the literature) in a more general setting, that is when $F : X \rightrightarrows Y$ is a set-valued mapping between Asplund spaces. In that case, one needs to add another condition to (b), and (c) which is: F is partially sequentially normally compact (PSNC) at (\bar{x}, \bar{y}) [54, Section 4.2]. Since we work on \mathbb{R}^m , F is automatically PSNC, and we avoid expressing such definitions and results. Another important situation when the conditions of Theorem 3.2.1 can be essentially simplified and efficiently specified, concerns set-valued mappings with closed and convex graphs; see, for example, the Aubin property of convex-graph multifunctions [54, p. 389].

General Assumptions:

(A1) B is injective;

(A2) f is continuously differentiable in \mathbb{R}^n ;

(A3) F has closed graph;

(A4) C is surjective; and

(A5) there are $F_i : \mathbb{R} \rightrightarrows \mathbb{R}$, $i \in \{1, \dots, m\}$ such that $F(x) = \prod_{i=1}^m F_i(x_i)$ whenever $x = (x_1, \dots, x_m)^T \in \mathbb{R}^m$.

A short review of components descriptions (especially Diodes (jump to 2.1)) of the last chapter, and the examples there (especially, Example 2.2.3), reveals that conditions (A3), and (A5) hold true automatically. Assumption (A4) helps us to replace F_C with F , without any further concern. A discussion about loosing the assumption (A1) will be given in Subsection 3.3.1.

The first proposition would provide a description for the coderivative of Φ in terms of ∇f , and the coderivative of F .

Proposition 3.2.3 (Coderivative of Sum for GE). [3, Proposition 3.1, p. 94]

Under the assumptions (A1) - (A3), for any $\xi \in \mathbb{R}^n$ one has

$$D^*\Phi(\bar{z}|\bar{p})(\xi) = \nabla f(\bar{z})^T \xi + C^T D^*F_C(C\bar{z}|\bar{v})(B^T \xi).$$

Proof. Fix any $\xi \in \mathbb{R}^n$. As (A2) ensures the strict differentiability of f at \bar{z} (see Remark 1.1.13), Theorem 3.1.25 implies that

$$D^*\Phi(\bar{z}|\bar{p})(\xi) = \nabla f(\bar{z})^T \xi + D^*Q(\bar{z}|\bar{p} - f(\bar{z}))(\xi).$$

Define $h : \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ by $h(z, v) := Bv$, for each $z \in \mathbb{R}^n$, $v \in \mathbb{R}^m$ and $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, by $G(z) := F(Cz)$, $z \in \mathbb{R}^n$. Thus, defining $h \circ G$ as introduced in Proposition 3.1.26, gives

$$h(z, G(z)) = \bigcup \{Bv : v \in G(z)\} = BG(z) = BF(Cz) = Q(z), \quad \text{whenever } z \in \mathbb{R}^n.$$

As $\bar{p} - f(\bar{z}) \in BF(C\bar{z})$, it is a simple observation that $\bar{v} = (B^T B)^{-1} B^T (\bar{p} - f(\bar{z})) \in G(\bar{z})$ and $h(\bar{z}, \bar{v}) = B\bar{v} = \bar{p} - f(\bar{z})$.

Obviously, $z \rightarrow h(z, \cdot)$ is strictly differentiable at \bar{z} . Since B is injective (A1), $\text{rge } h(\bar{z}, \cdot) = \mathbb{R}^n$; thus, using Proposition 3.1.26 reveals that

$$\begin{aligned} D^*Q(\bar{z}|\bar{p} - f(\bar{z}))(\xi) &= D^*(h \circ G)(\bar{z}|\bar{p} - f(\bar{z}))(\xi) \\ &= \nabla_z h(\bar{z}|\bar{v})^*(\xi) + D^*G(\bar{z}|\bar{v})(h(\bar{z}, \cdot)^*\xi) \\ &= D^*G(\bar{z}|\bar{v})(B^T \xi). \end{aligned}$$

Finally, notice that F_C is actually the restriction of F to the linear mapping C (that is, the function $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ defined by $g(z) := Cz$, cf. Notation 1.2), which easily satisfies the conditions of strict differentiability at \bar{z} and RMR around this point. Thus, Proposition 3.1.27 yields that

$$\begin{aligned} D^*G(\bar{z} | \bar{v})(B^T \xi) &= D^*(F \circ g)(\bar{z} | \bar{v})(B^T \xi) \\ &= \nabla g(\bar{z})^* \circ D^*(F|_g)(g(\bar{z}) | \bar{v})(B^T \xi) \\ &= C^T D^*F_C(C\bar{z} | \bar{v})(B^T \xi). \end{aligned}$$

Combining the above expressions ends the proof. \square

Theorem 3.2.4 (Aubin Property Criterion for the Solution Mapping of GE). [3, Theorem 3.1, p. 95] *Under the assumptions (A1)–(A3), S has the Aubin property at (\bar{p}, \bar{z}) if and only if*

$$0 \in \nabla f(\bar{z})^T \xi + C^T D^*F_C(C\bar{z} | \bar{v})(B^T \xi) \quad \Rightarrow \quad \xi = 0 \quad (3.49)$$

Moreover, its Lipschitz modulus is given by

$$\text{lip}(S; (\bar{p}, \bar{z})) = \sup \{ \|\xi\| : (\nabla f(\bar{z})^T \xi + C^T D^*F_C(C\bar{z} | \bar{v})(B^T \xi)) \cap \mathbb{B} \neq \emptyset \}. \quad (3.50)$$

Proof. First observe that since $S = \Phi^{-1}$, one has $(\eta, -\xi) \in N((\bar{z}, \bar{p}); \text{gph } \Phi)$ if and only if $(-\xi, \eta) \in N((\bar{p}, \bar{z}); \text{gph } S)$ or, equivalently,

$$\eta \in D^*\Phi(\bar{z} | \bar{p})(\xi) \quad \text{if and only if} \quad -\xi \in D^*S(\bar{p} | \bar{z})(-\eta).$$

Having Proposition 3.2.3 in hand, one applies the well-known Mordukhovich criterion in the finite-dimensional setting (Theorem 3.2.1) to conclude the proof. \square

Employing more assumptions, we may get the following corollaries.

Corollary 3.2.5. [3, Corollary 3.1, p. 95] *Suppose that the assumptions (A1)–(A4) hold true. Then S has the Aubin property at (\bar{p}, \bar{z}) if and only if*

$$\left. \begin{array}{l} ((CC^T)^{-1}C\nabla f(\bar{z})^T \xi, B^T \xi) \in -N((C\bar{z}, \bar{v}); \text{gph } F) \\ \nabla f(\bar{z})^T \xi \in \text{rge } C^T \end{array} \right\} \quad \Rightarrow \quad \xi = 0. \quad (3.51)$$

Proof. Since C is surjective, then $F_C = F$ and $CC^T \in \mathbb{R}^{m \times m}$ is non-singular.

First, let $\xi \in \mathbb{R}^n$ be such that $0 \in \nabla f(\bar{z})^T \xi + C^T D^*F(C\bar{z} | \bar{v})(B^T \xi)$.

Find $w \in D^*F(C\bar{z} | \bar{v})(B^T \xi)$ with $\nabla f(\bar{z})^T \xi + C^T w = 0$. Thus $-(CC^T)^{-1}C\nabla f(\bar{z})^T \xi$ is in $D^*F(C\bar{z} | \bar{v})(B^T \xi)$ or equivalently, $(-(CC^T)^{-1}C\nabla f(\bar{z})^T \xi, -B^T \xi) \in N((C\bar{z}, \bar{v}); \text{gph } F)$.

Clearly, we have $\nabla f(\bar{z})^T \xi \in \text{rge } C^T$ (for each ξ , one can define $u \in \mathbb{R}^n$ as $u =$

$(CC^T)^{-1}C\nabla f(\bar{z})^T\xi$, thanks to surjectivity of C) and Theorem 3.2.4 yields the rest.

On the other hand, pick any $\xi \in \mathbb{R}^n$ with $((CC^T)^{-1}C\nabla f(\bar{z})^T\xi, B^T\xi) \in -N((C\bar{z}, \bar{v}); \text{gph } F)$ and $\nabla f(\bar{z})^T\xi \in \text{rge } C^T$. The definition of the coderivative of F implies that $w := -(CC^T)^{-1}C\nabla f(\bar{z})^T\xi \in D^*F(C\bar{z}|\bar{v})(B^T\xi)$. Thus $CC^T w = -C\nabla f(\bar{z})^T\xi$. This implies that $C^T w + \nabla f(\bar{z})^T\xi \in \ker C \cap \text{rge } C^T = \{0\}$. Therefore $0 \in \nabla f(\bar{z})^T\xi + C^T D^*F(C\bar{z}|\bar{v})(B^T\xi)$ and Theorem 3.2.4 ends the proof. \square

Remark 3.2.6. If (A5) also holds, then $\text{gph } F = \prod_{j=1}^m \text{gph } F_j$, and Proposition 3.1.10 (about tangents and normals to product sets) implies that

$$N((C\bar{z}, \bar{v}); \text{gph } F) = \prod_{j=1}^m N(((C\bar{z})_j, \bar{v}_j); \text{gph } F_j).$$

Thus, the first condition in (3.51) can be checked coordinate-wise.

In order to express the next corollary, we need to remind the following definition from linear algebra.

Definition 3.2.7. (P-matrix) [58, p. 145] A matrix $A \in \mathbb{R}^{n \times n}$ is called a *P-matrix* if all its k -by- k principal minor determinants are positive whenever $k \in \{1, \dots, n\}$.

Using the well known Sylvester's criterion for positive definite matrices⁹, one can obtain an easy to verify rule for P-matrices:

A is a P-matrix if and only if for any non-zero $x \in \mathbb{R}^n$,
there is $j \in \{1, \dots, n\}$ such that $x_j(Ax)_j > 0$.

Corollary 3.2.8. [4, Corollary 1, p. 336] In addition to (A1) - (A5), assume that $n = m$, that $B = C = I_n$, that $\nabla f(\bar{z})^T$ is a P-matrix, and that for each $j \in \{1, 2, \dots, n\}$, we have

$$N((\bar{z}_j, \bar{v}_j); \text{gph } F_j) \subset \{(a, b)^T \in \mathbb{R}^2 : ab \leq 0\},$$

Then S has the Aubin property at $(\bar{p}, \bar{z}) \in \text{gph } S$.

Proof. In view of Remark 3.2.6, the condition (3.51) says that S has the Aubin property at (\bar{p}, \bar{z}) provided that $\xi = 0_{\mathbb{R}^n}$ is the only point which satisfies

$$(\nabla f(\bar{z})^T\xi, \xi) \in -\prod_{j=1}^n N((\bar{z}_j, \bar{v}_j); \text{gph } F_j) \subset \left[\{(a, b)^T \in \mathbb{R}^2 : ab \leq 0\} \right]^n.$$

⁹ In linear algebra, a symmetric $n \times n$ real matrix M is said to be *positive definite* if the scalar $z^T M z$ is positive for every non-zero column vector $z \in \mathbb{R}^n$.

To see a proof of Sylvester's necessary and sufficient criterion, refer to [40].

Suppose on the contrary that ξ is non-zero. Then for each $j \in \{1, 2, \dots, n\}$, we must have $\xi_j (\nabla f(\bar{z})^T \xi)_j \leq 0$; which is a contradiction, since $\nabla f(\bar{z})^T$ is a P-matrix. \square

Remark 3.2.9. Note that the inclusion in Corollary 3.2.8 is satisfied when F_j is a maximal monotone operator (ref. Definition 1.1.22).

For simplicity, let $F : \mathbb{R} \rightrightarrows \mathbb{R}$ be a maximal monotone operator, and $(a, b)^T \in N((\bar{z}, \bar{v}); \text{gph } F)$. Then, by definition of normal cones there exist sequences $(a_k, b_k) \rightarrow (a, b)$, and $(z_k, v_k) \in \text{gph } F$ with $(z_k, v_k) \rightarrow (\bar{z}, \bar{v})$, such that

$$\limsup_{(z,v) \xrightarrow{\text{gph } F} (z_k, v_k)} \frac{\langle (a_k, b_k), (z, v) - (z_k, v_k) \rangle}{\|(z, v) - (z_k, v_k)\|} \leq 0.$$

Thus, we must have $a_k(z - z_k) + b_k(v - v_k) \leq 0$.

From the maximal monotonicity of F we have $\langle v - v_k, z - z_k \rangle \geq 0$ (ref. Remark 1.1.24), which means $z - z_k$ and $v - v_k$ must have the same sign. There are three possibilities:

- I. $z - z_k = 0$, then $v - v_k$ could be an arbitrary real number. Thus $b_k = 0$ and for any $a_k \in \mathbb{R}$, the desired relation $a_k b_k \leq 0$ holds.
- II. $z - z_k > 0$, then $v - v_k \geq 0$. So one can obtain $a_k + b_k \frac{v - v_k}{z - z_k} \leq 0$. Thus,

$$a_k b_k \leq \frac{-b_k(v - v_k)}{z - z_k} b_k \leq \frac{-b_k^2(v - v_k)}{z - z_k} \leq 0$$

- III. $z - z_k < 0$, then $v - v_k \leq 0$. So one can obtain $a_k \geq -b_k \frac{v - v_k}{z - z_k}$.

One can easily observe that if $b_k \leq 0$, then we must have $a_k \geq 0$. On the other hand, if $a_k \leq 0$, then $b_k \geq 0$ (omitting the case when $v - v_k = 0$, and we must have $a_k = 0$).

Thus we have proved $(a, b)^T \in N((\bar{z}, \bar{v}); \text{gph } F)$ implies $ab \leq 0$.

Example 3.2.10. (A Simple Circuit with DIAC) [4, p. 339]

Consider the circuit in Figure 3.4 with a DIAC. Suppose that $V > 0$ and $a > 0$ are given, and the $i - v$ characteristic of the DIAC is given by:

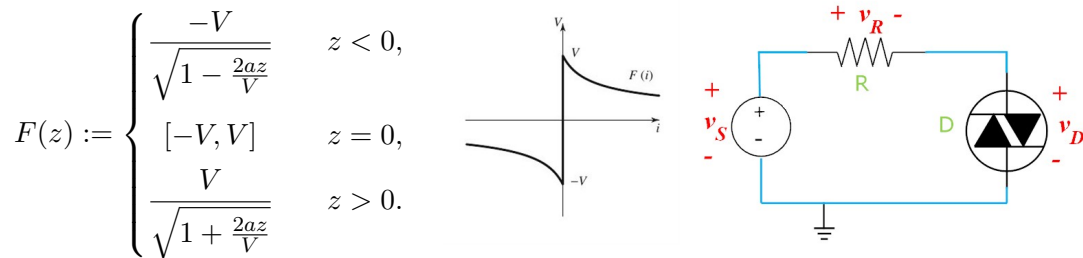


FIGURE 3.4: Calculating the Aubin property in a circuit with DIAC

In Example 2.3.2, we discussed the effect of parameters R , and a on the solution set at each point and changed the maps to

$$f(z) := \begin{cases} Rz - \frac{V}{\sqrt{1 - \frac{2az}{V}}} + V & z < 0, \\ Rz + \frac{V}{\sqrt{1 + \frac{2az}{V}}} - V & z \geq 0, \end{cases} \quad \text{and} \quad F(z) := \begin{cases} -V & z < 0, \\ [-V, V] & z = 0, \\ V & z > 0. \end{cases}$$

In this setting, $B = C = I_1$ and the assumptions (A1) - (A4) hold true. For a fixed $\bar{p} \in \mathbb{R}$, let \bar{z} be the corresponding solution to (3.46), and $\bar{v} = \bar{p} - f(\bar{z})$. Now, condition (3.51) in Corollary 3.2.5 reads as

$$(f'(\bar{z})\xi, \xi) \in -N((\bar{z}, \bar{v}); \text{gph } F) \implies \xi = 0,$$

where $-N((x, y); \text{gph } F)$ is given by

$$\left\{ \begin{array}{ll} \mathbb{R} \begin{pmatrix} 0 \\ 1 \end{pmatrix} & x > 0, y = V, \text{ or } x < 0, y = -V, \\ \mathbb{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix} & x = 0, y \in (-V, V), \\ \mathbb{R} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cup \text{cone} \left\{ \begin{pmatrix} 0 \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} & x = 0, y = V, \\ \mathbb{R} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cup \text{cone} \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\} & x = 0, y = -V, \\ \emptyset & \text{otherwise.} \end{array} \right.$$

Let us check the first rule of the normal cone, for example. The condition $(f'(\bar{z})\xi, \xi) = (0, \lambda)$ for some $\lambda \in \mathbb{R}$, when $\bar{z} > 0$, and $\bar{p} - f(\bar{z}) = V$ or $\bar{z} < 0$, and $\bar{p} - f(\bar{z}) = -V$, should result in $\xi = 0$. The condition will fail if for $\lambda \neq 0$, we can obtain $f'(\bar{z}) = 0$.

If $a < R$, one knows that $f'(z) > 0 \forall z \in \mathbb{R}$. If $a = R$, then $f'(z) = 0$ only happens at $z = 0$ which is out of the available range of \bar{z} . Thus, $\lambda = 0$ and the condition is satisfied.

In the third and fourth rule of the normal cone, $\bar{z} = 0$ is allowed and so, in case of $a = R$, for $\bar{p} \in \{V, -V\}$, the condition (3.51) will be violated.

In the fourth rule, let us check the criteria $(f'(\bar{z})\xi, \xi) \in \text{cone} \left\{ \begin{pmatrix} 0 \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\}$ when $\bar{z} = 0$, and $\bar{v} = \bar{p} - f(0) = -V$. For $a < R$, since $f'(0) = R - a > 0$, the points $(f'(\bar{z})\xi, \xi)$ define a line passing through origin in the first and third quadrant. Thus, the only possible intersection with the cone would be at $(0, 0)$ which implies $\xi = 0$. To sum up these observations,

if $a < R$, then S has the Aubin property at any point $(\bar{p}, \bar{z}) \in \text{gph } S$;

if $a = R$, then S has the Aubin property at (\bar{p}, \bar{z}) provided that $\bar{p} \notin \{-V, V\}$.

The case $a > R$ could be discussed similarly, but the solution is not unique any more. For example, if $\bar{z} = 0$, then S has the Aubin property at (\bar{p}, \bar{z}) when $\bar{p} \in (-V, V)$, since the ‘‘cone’’ terms in the third and fourth pieces of the normal cone calculation permit the ξ to be non-zero.

Example 3.2.11. (A Circuit with SCR and Zener Diode) [4, p. 344]

Consider the circuit in Figure 3.5 involving two non-smooth elements in parallel: the SCR and the Zener Diode.

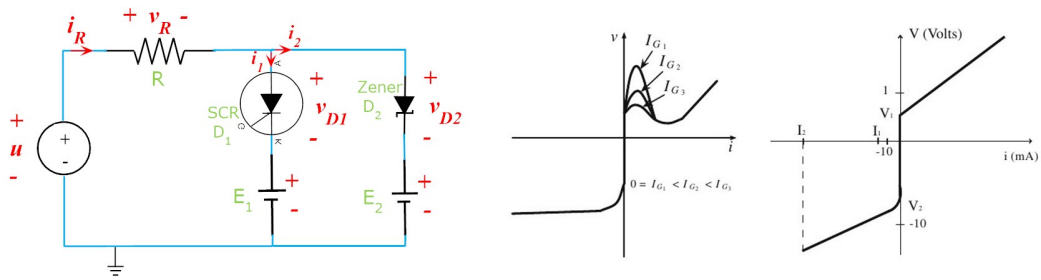


FIGURE 3.5: Calculating the Aubin property in a circuit with SCR and Zener Diode.

In Example 2.3.3, we obtained the mathematical model for this circuit in the form of (3.46) with $m = n = 2$, $B = C = I_2$, $f(z) = Az$ for $z \in \mathbb{R}^2$, where $A = \begin{pmatrix} R & R \\ R & R \end{pmatrix}$, and

$$F : \mathbb{R}^2 \rightrightarrows \mathbb{R}^2 \text{ is defined as } F \begin{pmatrix} z_1 \\ z_2 \end{pmatrix} := \begin{pmatrix} F_1(z_1) \\ F_2(z_2) \end{pmatrix}.$$

The $i - v$ characteristics F_1 of SCR and F_2 of the Zener diode are defined as

$$F_1(z) := \begin{cases} az + V_1 & z < 0, \\ [V_1, \varphi(0)] & z = 0, \\ \varphi(z) & z \in [0, \alpha], \\ a(z - \alpha) + \varphi(\alpha) & z > \alpha, \end{cases} \quad F_2(z) := \begin{cases} bz - V & z < 0, \\ [-V, V] & z = 0, \\ bz + V & z > 0, \end{cases}$$

where $V, -V_1, \alpha, a$, and b are positive constants, and $\varphi : \mathbb{R} \rightarrow (0, \infty)$ is a continuously differentiable function with $\varphi(\alpha) < \varphi(0)$, $\varphi'(0) > 0$, and $\varphi'(\alpha) > 0$.

Given $\bar{p} = (\bar{p}_1, \bar{p}_2) \in \mathbb{R}^2$, denote by $\bar{z} = (\bar{z}_1, \bar{z}_2)^T$ the corresponding solution to (3.46). Condition (3.51) of Corollary 3.2.5 reads as

$$(I_2 A \xi, I_2 \xi) \in -N\left((I_2 \bar{z}, \bar{v}); \text{gph } F\right) \implies \xi = (0, 0),$$

where $\bar{v} = \bar{p} - f(\bar{z})$. In view of Remark 3.2.6, the necessary and sufficient condition for the Aubin property of S at the reference point is

$$\left. \begin{aligned} (R(\xi_1 + \xi_2), \xi_1) &\in -N\left(\left(\bar{z}_1, \bar{p}_1 - R(\bar{z}_1 + \bar{z}_2)\right); \text{gph } F_1\right) \\ (R(\xi_1 + \xi_2), \xi_2) &\in -N\left(\left(\bar{z}_2, \bar{p}_2 - R(\bar{z}_1 + \bar{z}_2)\right); \text{gph } F_2\right) \end{aligned} \right\} \implies \xi_1 = \xi_2 = 0, \quad (3.52)$$

where the first normal cone, $N((x, y); \text{gph } F_1)$, can be calculated as

$$\left\{ \begin{array}{ll} \mathbb{R} \begin{pmatrix} a \\ -1 \end{pmatrix} & x < 0, y = ax + V_1, \text{ or} \\ & x > \alpha, y = a(x - \alpha) + \varphi(\alpha), \\ \mathbb{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix} & x = 0, y \in (V_1, \varphi(0)), \\ \mathbb{R} \begin{pmatrix} \varphi'(x) \\ -1 \end{pmatrix} & x \in (0, \alpha), y = \varphi(x), \\ \mathbb{R} \begin{pmatrix} a \\ -1 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cup \text{cone} \left\{ \begin{pmatrix} a \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} & x = 0, y = V_1, \\ \mathbb{R} \begin{pmatrix} -\varphi'(0) \\ 1 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} -1 \\ 0 \end{pmatrix} \cup \text{cone} \left\{ \begin{pmatrix} -\varphi'(0) \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\} & x = 0, y = \varphi(0), \\ \mathbb{R} \begin{pmatrix} a \\ -1 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} \varphi'(\alpha) \\ -1 \end{pmatrix} \cup \text{cone} \left\{ \begin{pmatrix} a \\ -1 \end{pmatrix}, \begin{pmatrix} \varphi'(\alpha) \\ -1 \end{pmatrix} \right\} & x = \alpha, y = \varphi(\alpha), \\ \emptyset & \text{otherwise.} \end{array} \right.$$

The second normal cone, $N((x, y); \text{gph } F_2)$, could be obtained similarly

$$\left\{ \begin{array}{ll} \mathbb{R} \begin{pmatrix} b \\ -1 \end{pmatrix} & x < 0, y = bx - V, \text{ or} \\ & x > 0, y = bx + V \\ \mathbb{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix} & x = 0, y \in (-V, V), \\ \mathbb{R} \begin{pmatrix} b \\ -1 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cup \text{cone} \left\{ \begin{pmatrix} b \\ -1 \end{pmatrix}, \begin{pmatrix} 1 \\ 0 \end{pmatrix} \right\} & x = 0, y = -V, \\ \mathbb{R} \begin{pmatrix} -b \\ 1 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} -1 \\ 0 \end{pmatrix} \cup \text{cone} \left\{ \begin{pmatrix} -b \\ 1 \end{pmatrix}, \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\} & x = 0, y = V, \\ \emptyset & \text{otherwise.} \end{array} \right.$$

In order to simplify the process of checking condition (3.52), we multiply the inclusions in (3.52) by the matrices $M_1 := \begin{pmatrix} 0 & 1 \\ \frac{1}{R} & -1 \end{pmatrix}$, and $M_2 := \begin{pmatrix} \frac{1}{R} & -1 \\ 0 & 1 \end{pmatrix}$, respectively, to get the following equivalent condition

$$M_1\left(N((\bar{z}_1, \bar{v}_1); \text{gph } F_1)\right) \cap M_2\left(N((\bar{z}_2, \bar{v}_2); \text{gph } F_2)\right) = (0, 0)^{10}. \quad (3.53)$$

Now, one infers that $M_1\left(N((\bar{z}_1, \bar{v}_1); \text{gph } F_1)\right)$ is given by

$$\left\{ \begin{array}{l} \mathbb{R} \begin{pmatrix} -1 \\ \frac{a}{R} + 1 \end{pmatrix} \\ \mathbb{R} \begin{pmatrix} 0 \\ \frac{1}{R} \end{pmatrix} \\ \mathbb{R} \begin{pmatrix} -1 \\ \frac{\varphi'(x)}{R} + 1 \end{pmatrix} \\ \mathbb{R} \begin{pmatrix} -1 \\ \frac{a}{R} + 1 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} 0 \\ \frac{1}{R} \end{pmatrix} \cup \text{cone} \left\{ \begin{pmatrix} -1 \\ \frac{a}{R} + 1 \end{pmatrix}, \begin{pmatrix} 0 \\ \frac{1}{R} \end{pmatrix} \right\} \\ \mathbb{R} \begin{pmatrix} 1 \\ -\frac{\varphi'(0)}{R} + 1 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} 0 \\ -\frac{1}{R} \end{pmatrix} \cup \text{cone} \left\{ \begin{pmatrix} 1 \\ -\frac{\varphi'(0)}{R} + 1 \end{pmatrix}, \begin{pmatrix} 0 \\ -\frac{1}{R} \end{pmatrix} \right\} \\ \mathbb{R} \begin{pmatrix} -1 \\ \frac{a}{R} + 1 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} -1 \\ \frac{\varphi'(\alpha)}{R} + 1 \end{pmatrix} \cup \text{cone} \left\{ \begin{pmatrix} -1 \\ \frac{a}{R} + 1 \end{pmatrix}, \begin{pmatrix} -1 \\ \frac{\varphi'(\alpha)}{R} + 1 \end{pmatrix} \right\} \\ \emptyset \end{array} \right. \quad \begin{array}{l} \bar{z}_1 < 0, \bar{v}_1 = a\bar{z}_1 + V_1, \text{ or} \\ \bar{z}_1 > \alpha, \bar{v}_1 = a(\bar{z}_1 - \alpha) + \varphi(\alpha), \\ \bar{z}_1 = 0, \bar{v}_1 \in (V_1, \varphi(0)), \\ \bar{z}_1 \in (0, \alpha), \bar{v}_1 = \varphi(\bar{z}_1), \\ \bar{z}_1 = 0, \bar{v}_1 = V_1, \\ \bar{z}_1 = 0, \bar{v}_1 = \varphi(0), \\ \bar{z}_1 = \alpha, \bar{v}_1 = \varphi(\alpha), \\ \text{otherwise.} \end{array}$$

And $M_2\left(N((\bar{z}_2, \bar{v}_2); \text{gph } F_2)\right)$ is given by

¹⁰ The reason is that

$$M_1 \begin{pmatrix} R\xi_1 + R\xi_2 \\ R\xi_1 \end{pmatrix} = \begin{pmatrix} \xi_1 \\ \xi_2 \end{pmatrix} = M_2 \begin{pmatrix} R\xi_1 + R\xi_2 \\ R\xi_2 \end{pmatrix}.$$

Therefore, the condition will hold true if and only if $(0, 0)$ is the only point in the intersection of $M_1\left(N((\bar{z}_1, \bar{v}_1); \text{gph } F_1)\right)$ and $M_2\left(N((\bar{z}_2, \bar{v}_2); \text{gph } F_2)\right)$.

$$\left\{ \begin{array}{ll} \mathbb{R} \begin{pmatrix} \frac{b}{R} + 1 \\ -1 \end{pmatrix} & \bar{z}_2 < 0, \bar{v}_2 = b\bar{z}_2 - V, \text{ or} \\ & \bar{z}_2 > 0, \bar{v}_2 = b\bar{z}_2 + V \end{array} \right. , \\ \left\{ \begin{array}{ll} \mathbb{R} \begin{pmatrix} \frac{1}{R} \\ 0 \end{pmatrix} & \bar{z}_2 = 0, \bar{v}_2 \in (-V, V), \\ \mathbb{R} \begin{pmatrix} \frac{b}{R} + 1 \\ -1 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} \frac{1}{R} \\ 0 \end{pmatrix} \cup \text{cone} \left\{ \begin{pmatrix} \frac{b}{R} + 1 \\ -1 \end{pmatrix}, \begin{pmatrix} \frac{1}{R} \\ 0 \end{pmatrix} \right\} & \bar{z}_2 = 0, \bar{v}_2 = -V, \\ \mathbb{R} \begin{pmatrix} \frac{-b}{R} - 1 \\ 1 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} \frac{-1}{R} \\ 0 \end{pmatrix} \cup \text{cone} \left\{ \begin{pmatrix} \frac{-b}{R} - 1 \\ 1 \end{pmatrix}, \begin{pmatrix} \frac{-1}{R} \\ 0 \end{pmatrix} \right\} & \bar{z}_2 = 0, \bar{v}_2 = V, \\ \emptyset & \text{otherwise.} \end{array} \right.$$

Hence, one can easily observe that the Aubin property of the solution map depends explicitly on the exact values of the parameters of the circuit and the reference point. For example, if (\bar{z}_1, \bar{v}_1) , and (\bar{z}_2, \bar{v}_2) reside in the first sub-domain of $M_1(N((\bar{z}_1, \bar{v}_1); \text{gph } F_1))$ and $M_2(N((\bar{z}_2, \bar{v}_2); \text{gph } F_2))$, respectively, then any point of the form

$$\lambda \begin{pmatrix} -1 \\ \frac{a}{R} + 1 \end{pmatrix} = \lambda' \begin{pmatrix} \frac{b}{R} + 1 \\ -1 \end{pmatrix},$$

would belong to the intersection in (3.53) for $\lambda, \lambda' \in \mathbb{R}$. The equality holds if and only if $\lambda = \lambda' = 0$. Thus, the solution mapping has the Aubin property.

When $\bar{z} = (0, 0)$, condition (3.53) holds if $\bar{p}_1 \in (V_1, \varphi(0))$, and $\bar{p}_2 \in [-V, V]$. Thus, the solution mapping has the Aubin property for these reference points.

For $\bar{z}_1 \in (0, \alpha)$, $\bar{p}_1 = \varphi(\bar{z}_1)$, and $\bar{z}_2 = 0, \bar{p}_2 \in \{-V, V\}$; one should check the intersection of the line $\mathbb{R} \left(-1, \frac{\varphi'(\bar{z}_1)}{R} + 1 \right)^T$ with the third piece in $M_2(N((\bar{z}_2, \bar{v}_2); \text{gph } F_2))$, which is shown in Figure 3.6.

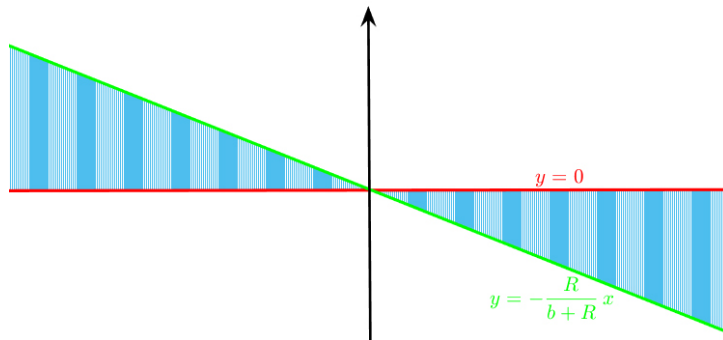


FIGURE 3.6: A part of the set $M_2(N((\bar{z}_2, \bar{v}_2); \text{gph } F_2))$

Clearly, the condition (3.53) is violated when $0 \leq \frac{\varphi'(\bar{z}_1)}{R} + 1 \leq \frac{R}{b+R}$. Therefore, the solution mapping does not have the Aubin property if $-R \leq \varphi'(\bar{z}_1) \leq \frac{-Rb}{b+R}$.

3.3 Results about Isolated Calmness

The starting point here would be a characterization of strong metric sub-regularity with graphical derivatives (see Theorem 3.1.40 and Corollary 3.1.41). First we assume that the linear operator B in the setting (3.46) is injective, and in the subsection afterwards, we change this assumption with a slightly different one.

Consider the mappings Q and F_C as defined in (3.48). In order to pass from the graphical derivative of Q to the graphical derivative of F_C (which in fact, needs a special chain rule for the graphical derivative of the composition mapping like the idea in Propositions 3.1.26, and 3.1.27 for coderivatives), we need the following lemma which provides a rule on how a linear operator alters a tangent cone relation.

Lemma 3.3.1 (Matrix Operation over Tangent Cone Relations). [3, Lemma 4.1, p. 96] *Let $E \in \mathbb{R}^{k \times d}$ be any matrix, let $G \in \mathbb{R}^{l \times d}$ be injective, and let Γ be a subset of $\text{rge } E$. Put $\Xi := E^{-1}(\Gamma)$ and $\Lambda := G(\Xi)$. For $\bar{x} \in \Lambda$, denote by \bar{y} the (unique) point in Ξ with $G\bar{y} = \bar{x}$. Then*

$$T(\bar{x}; \Lambda) = \{u \in \mathbb{R}^l : \exists w \in \mathbb{R}^d \text{ such that } u = Gw \text{ and } Ew \in T(E\bar{y}; \Gamma)\}.$$

Proof. We claim that

$$T(\bar{y}; \Xi) = \{w \in \mathbb{R}^d : Ew \in T(E\bar{y}; \Gamma)\}. \quad (3.54)$$

First, take any $w \in T(\bar{y}; \Xi)$. Find $(t_n)_{n \in \mathbb{N}}$ in $(0, \infty)$ converging to 0, and $(w_n)_{n \in \mathbb{N}}$ in \mathbb{R}^d converging w , such that $\bar{y} + t_n w_n \in \Xi$ whenever $n \in \mathbb{N}$. Then we have that

$$E\bar{y} + t_n Ew_n = E(\bar{y} + t_n w_n) \in \Gamma \text{ for each } n \in \mathbb{N}.$$

Hence $Ew \in T(E\bar{y}; \Gamma)$. On the other hand, let $w \in \mathbb{R}^d$ be such that $Ew \in T(E\bar{y}; \Gamma)$. Pick $(t_n)_{n \in \mathbb{N}}$ in $(0, \infty)$ and $(v_n)_{n \in \mathbb{N}}$ in \mathbb{R}^k converging to 0 and Ew , respectively, such that $E\bar{y} + t_n v_n \in \Gamma$ whenever $n \in \mathbb{N}$. As $\Gamma \subset \text{rge } E$, and $\text{rge } E$ is a closed subspace of \mathbb{R}^k (which is a direct result of the continuity of E as a bounded linear operator on \mathbb{R}^k), one infers that $v_n \in \text{rge } E$ for each $n \in \mathbb{N}$.

Therefore, by Banach open mapping theorem there is $(w_n)_{n \in \mathbb{N}}$ in \mathbb{R}^d converging to w such

that $Ew_n = v_n$ for each $n \in \mathbb{N}$ ¹¹. Thus, for an arbitrary index n , we have $E(\bar{y} + t_n w_n) \in \Gamma$, hence $\bar{y} + t_n w_n \in E^{-1}(\Gamma) = \Xi$. So $w \in T(\bar{y}; \Xi)$. The claim is proved.

Now we prove that

$$T(\bar{x}; \Lambda) = \{Gw : w \in T(\bar{y}; \Xi)\}, \quad (3.55)$$

using exactly the same steps as in the proof of Lemma 4.1 in [7]¹².

Let us first prove the inclusion $\{Gw : w \in T(\bar{y}; \Xi)\} \subset T(\bar{x}; \Lambda)$.

Consider $w \in T(\bar{y}; \Xi)$. By definition of the tangent cone, there are sequences $t_k \downarrow 0$, and $w_k \rightarrow w$ such that $\bar{y} + t_k w_k \in \Xi$. Clearly, we have

$$\bar{x} + t_k Gw_k = G\bar{y} + t_k Gw_k \in G(\Xi) = \Lambda$$

Thus, $Gw \in T(\bar{x}; \Lambda)$.

It remains to prove that $T(\bar{x}; \Lambda) \subset \{Gw : w \in T(\bar{y}; \Xi)\}$, which can be conducted by a similar reasoning. Let $x \in T(\bar{x}; \Lambda)$. By definition, there are sequences $x_k \rightarrow x$, and $\lambda_k \downarrow 0$ such that $\bar{x} + \lambda_k x_k \in \Lambda$. Hence, there is a sequence w_k in Ξ such that for each k , one has $\bar{x} + \lambda_k x_k = Gw_k$. Thus, $x_k = G \frac{w_k - \bar{y}}{\lambda_k}$.

Consider the sequence (h_k) defined as

$$h_k = \frac{w_k - \bar{y}}{\lambda_k}.$$

For each k , let $\lambda'_k = \lambda_k$ and observe that $\lambda'_k h_k + \bar{y} = w_k \in \Xi$, and $\lambda'_k \downarrow 0$. In order to show that h_k has a convergent subsequence, having Bolzano-Weierstrass Theorem in mind, it only suffices to prove the boundedness of h_k .

Assume by contradiction that h_k is unbounded. Hence there exists a subsequence (still denoted by h_k) such that $\|h_k\| \rightarrow +\infty$ as $k \rightarrow +\infty$. By passing to a subsequence if necessary, we have

$$\frac{h_k}{\|h_k\|} \rightarrow s \quad \text{with} \quad \|s\| = 1.$$

It follows that $Gs = \lim_{k \rightarrow \infty} \frac{x_k}{\|h_k\|} = 0$, which contradicts the injectivity of G . Thus, the sequence h_k possesses a convergent subsequence with the limit h_0 in $T(\bar{y}; \Xi)$. We can also conclude that $x = Gh_0 \in GT(\bar{y}; \Xi)$. Hence, (3.55) is proved.

Combining (3.54) and (3.55) yield the assertion. \square

¹¹ Indeed, the restricted map $\tilde{E} : (\mathbb{R}^d / \ker E) \rightarrow \text{rge } E$, induced by E , is a bijective continuous linear operator, hence from the open mapping theorem, \tilde{E}^{-1} is continuous, too. Then, $Ew_n = v_n \rightarrow Ew$ implies that $w_n \rightarrow w$.

¹² Actually the Lemma 4.1 of [7] is a similar statement with two more conditions: E is assumed to be surjective and Γ is a closed subset of \mathbb{R}^n .

But these two conditions are only used in the first part of their proof (referring to [65, Exercise 6.7]) which is equal to the claim above, proved without these assumptions. Thus, this lemma could be considered as a generalization of that result.

The above Lemma could be written in the following symbolic representation now, which helps us remember what is happening there.

$$\begin{aligned}
Gw \in T(\bar{x}; G(E^{-1}(\Gamma))) &\iff G^{-1}Gw \in T(G^{-1}\bar{x}; G^{-1}G(E^{-1}(\Gamma))) \\
&\iff w \in T(\bar{y}; E^{-1}(\Gamma)) \\
&\iff Ew \in T(E\bar{y}; \Gamma).
\end{aligned} \tag{3.56}$$

Proposition 3.3.2 (Graphical Derivative of Sum for GE). [3, Proposition 4.1, p. 96] Under the assumptions (A1) - (A2), for any $b \in \mathbb{R}^n$ one has

$$D\Phi(\bar{z} | \bar{p})(b) = \nabla f(\bar{z})b + BDF_C(C\bar{z} | \bar{v})(Cb).$$

Proof. Fix any $b \in \mathbb{R}^n$. By Proposition 3.1.22 and Remark 3.1.18, we have

$$D\Phi(\bar{z} | \bar{p})(b) = \nabla f(\bar{z})b + DQ(\bar{z} | \bar{p} - f(\bar{z}))(b).$$

Observe that by the definition of Q as $Q(z) := BF(Cz)$,

$$\text{gph } Q = \left\{ \begin{array}{l} \begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^{2n} : \exists \begin{pmatrix} b \\ c \end{pmatrix} \in \mathbb{R}^n \times \mathbb{R}^m \text{ such that} \\ \begin{pmatrix} u \\ v \end{pmatrix} = G \begin{pmatrix} b \\ c \end{pmatrix} \text{ and } E \begin{pmatrix} b \\ c \end{pmatrix} \in \text{gph } F_C \end{array} \right\},$$

with

$$G := \begin{pmatrix} I_n & 0 \\ 0 & B \end{pmatrix}, \quad E := \begin{pmatrix} C & 0 \\ 0 & I_m \end{pmatrix}.$$

As B is injective, so is G . Using Lemma 3.3.1 with $k := 2m, l := 2n, d := n + m, \Gamma := \text{gph } F_C, \bar{x} := (\bar{z}, \bar{p} - f(\bar{z}))^T$, and $\bar{y} := (\bar{z}, \bar{v})^T$ reveals that

$$T((\bar{z}, \bar{p} - f(\bar{z})); \text{gph } Q) = \left\{ \begin{pmatrix} b \\ Bc \end{pmatrix} : \begin{pmatrix} Cb \\ c \end{pmatrix} \in T((C\bar{z}, \bar{v}); \text{gph } F_C) \right\}.$$

This means that $DQ(\bar{z} | \bar{p} - f(\bar{z}))(b) = BDF_C(C\bar{z} | \bar{v})(Cb)$. The assertion is proved. \square

Theorem 3.3.3 (Isolated Calmness Criterion for the Solution Mapping of GE). [3, Theorem 4.1, p. 97] Under the assumptions (A1) - (A3), S has the isolated calmness property at (\bar{p}, \bar{z}) if and only if

$$0 \in \nabla f(\bar{z})b + BDF_C(C\bar{z} | \bar{v})(Cb) \implies b = 0. \tag{3.57}$$

Moreover, its calmness modulus is given by

$$\text{clm}(S; (\bar{p}, \bar{z})) = \sup \{ \|b\| : (\nabla f(\bar{z})b + B DF_C(C\bar{z}|\bar{v})(Cb)) \cap \mathbb{B} \neq \emptyset \}. \quad (3.58)$$

Proof. Note that $x \in D\Phi(\bar{z}|\bar{p})(y)$ if and only if $y \in DS(\bar{p}|\bar{z})(x)$.

Indeed, $(x, y) \in T((\bar{p}, \bar{z}); \text{gph } S)$ implies the existence of sequences $t_k \downarrow 0, (u_k, v_k) \rightarrow (x, y)$ such that $\bar{z} + t_k v_k \in S(\bar{p} + t_k u_k)$. Then, $\bar{p} + t_k u_k \in \Phi(\bar{z} + t_k v_k)$, by definition. Thus, $(y, x) \in T((\bar{z}, \bar{p}); \text{gph } \Phi)$ and vice versa.

It would be enough to combine Corollary 3.1.41, Theorem 3.1.40, and Proposition 3.3.2, to conclude the proof. We just want to explain a bit, the modulus formula here.

$$\begin{aligned} \text{clm}(S; (\bar{p}, \bar{z})) &= |DS(\bar{p}|\bar{z})|^+ := \sup_{\|x\| \leq 1} \sup_{y \in DS(\bar{p}|\bar{z})(x)} \|y\| \\ &= \sup_{\|x\| \leq 1} \sup_{x \in D\Phi(\bar{z}|\bar{p})(y)} \|y\| \\ &= \sup_{\|x\| \leq 1} \sup \{ \|y\| : x \in \nabla f(\bar{z})y + B DF_C(C\bar{z}|\bar{v})(Cy) \} \\ &= \sup \{ \|y\| : \|\nabla f(\bar{z})y + B DF_C(C\bar{z}|\bar{v})(Cy)\| \leq 1 \} \\ &= \sup \{ \|y\| : (\nabla f(\bar{z})y + B DF_C(C\bar{z}|\bar{v})(Cy)) \cap \mathbb{B} \neq \emptyset \}. \end{aligned}$$

The equations show the step by step process of combining the mentioned results and definition of outer norm to obtain the claimed formula. \square

Again, imposing the additional assumption that C is surjective we can go one step further and get the following statement.

Corollary 3.3.4. [3, Corollary 4.1, p. 97] Suppose that the assumptions (A1) - (A4) (cf. Note 3.2.2) hold true. Then S has the isolated calmness property at (\bar{p}, \bar{z}) if and only if

$$\left. \begin{aligned} (Cb, -(B^T B)^{-1} B^T \nabla f(\bar{z})b) \in T((C\bar{z}, \bar{v}); \text{gph } F) \\ \nabla f(\bar{z})b \in \text{rge } B \end{aligned} \right\} \implies b = 0. \quad (3.59)$$

Proof. Indeed, if C is surjective, then $F_C = F$. Note that (A1) ensures that $B^T B \in \mathbb{R}^{m \times m}$ is non-singular.

First, let $b \in \mathbb{R}^n$ be such that $0 \in \nabla f(\bar{z})b + B DF_C(C\bar{z}, \bar{v})(Cb)$. Then, find a point $w \in DF_C(C\bar{z}, \bar{v})(Cb)$ with $\nabla f(\bar{z})b + Bw = 0$. Thus, $-(B^T B)^{-1} B^T \nabla f(\bar{z})b$ is in $DF_C(C\bar{z}, \bar{v})(Cb)$. Clearly, we have $\nabla f(\bar{z})b \in \text{rge } B$ and the definition of the contingent derivative of F yields the rest.

On the other hand, pick any $b \in \mathbb{R}^n$ with $(Cb, -(B^T B)^{-1} B^T \nabla f(\bar{z})b) \in T((C\bar{z}, \bar{v}); \text{gph } F)$

and $\nabla f(\bar{z})b \in \text{rge } B$. The definition of the contingent derivative yields that

$$w := -(B^T B)^{-1} B^T \nabla f(\bar{z})b \in DF(C\bar{z}, \bar{v})(Cb).$$

Thus, we have $B^T Bw = -B^T \nabla f(\bar{z})b$. So $Bw + \nabla f(\bar{z})b \in \ker B^T \cap \text{rge } B = \{0\}$. Therefore, $0 \in \nabla f(\bar{z})b + BDF(C\bar{z}, \bar{v})(Cb)$. Using Theorem 3.3.3 ends the proof. \square

Remark 3.3.5. If (A5) is also satisfied and $\text{gph } F_j$ is *Clarke regular* at $((C\bar{z})_j, \bar{v}_j)$ for each $j \in \{1, \dots, m\}$ (cf. Remark 3.1.5), then Proposition 3.1.10 implies that

$$T((C\bar{z}, \bar{v}); \text{gph } F) = \prod_{j=1}^m T(((C\bar{z})_j, \bar{v}_j); \text{gph } F_j).$$

Hence the first condition in (3.59) can be checked coordinate-wise.

Example 3.3.6. (A Simple Circuit with DIAC) [4, p. 339]

Consider the circuit in Figure 3.4 with a DIAC. We discussed the Aubin property of the solution mapping at different points of its graph in Example 3.2.10. Here we want to investigate the isolated calmness property.

If $a < R$, the uniqueness of the solution (ref. Example 2.3.2) implies that S has also the isolated calmness property at any $(\bar{p}, \bar{z}) \in \text{gph } S$. In order to use Corollary 3.3.4, we observe that the assumptions (A1) - (A4) hold true and the condition (3.59) has the form

$$(b, -f'(\bar{z})b) \in T((\bar{z}, \bar{p} - f(\bar{z})); \text{gph } F) \implies b = 0, \quad (3.60)$$

with $T((x, y); \text{gph } F)$ given by

$$\left\{ \begin{array}{ll} \mathbb{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix} & x > 0, y = V, \text{ or } x < 0, y = -V, \\ \mathbb{R} \begin{pmatrix} 0 \\ 1 \end{pmatrix} & x = 0, y \in (-V, V), \\ \mathbb{R}_+ \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cup \mathbb{R}_+ \begin{pmatrix} 0 \\ -1 \end{pmatrix} & x = 0, y = V, \\ \mathbb{R}_+ \begin{pmatrix} -1 \\ 0 \end{pmatrix} \cup \mathbb{R}_+ \begin{pmatrix} 0 \\ 1 \end{pmatrix} & x = 0, y = -V, \\ \emptyset & \text{otherwise.} \end{array} \right.$$

From the third and fourth pieces one can not guarantee that $b = 0$ when $a = R$. Hence, if $a = R$, S has the isolated calmness property at (\bar{p}, \bar{z}) provided that $\bar{p} \notin \{-V, V\}$.

The case $a > R$ is more interesting, since the solution is no more unique. Simple calculations show that it is possible to have $f'(\bar{z}) = 0$ for some $\bar{z} \neq 0$. So, the solution mapping is not isolated calm when $\bar{z} < 0$, or $\bar{z} > 0$.

Let $\bar{z} = 0$. Then, $f'(0) = R - a \neq 0$, and (3.60) reveals that S has the isolated calmness property at (\bar{p}, \bar{z}) when $\bar{p} \in [-V, V]$ (while we have already observed that it does not have the Aubin property if $\bar{p} \in \{-V, V\}$). Let us compute the calmness modulus provided by Theorem 3.3.3 in this case, with simplifying assumptions $R = V = 1$. Formula (3.58) reads as

$$\text{clm}(S; (-1, 0)) = \sup \left\{ \|b\| : \left((1-a)b + DF(0| -1)(b) \right) \cap \mathbb{B} \neq \emptyset \right\}.$$

Since $DF(0| -1)(b) = \left\{ v \in \mathbb{R} \mid \begin{pmatrix} b \\ v \end{pmatrix} \in \begin{pmatrix} -\lambda \\ 0 \end{pmatrix} \cup \begin{pmatrix} 0 \\ \lambda \end{pmatrix} \text{ for some } \lambda \in \mathbb{R}_+ \right\}$, one obtains that the supremum of b should be taken over the constraint $-1 \leq (1-a)b + v \leq 1$, where the points $(b, v)^T$ belong to the previously mentioned sets. Hence,

$$\text{clm}(S; (-1, 0)) = \begin{cases} \frac{1}{|1-a|} & a \neq 1, \\ \infty & \text{otherwise.} \end{cases}$$

Example 3.3.7. (A Circuit with SCR and Zener Diode) [4, p. 342]

Consider the circuit in Figure 3.5 with a SCR and Zener diode. We want to investigate the isolated calmness property of the solution mapping at different points of its graph (in Example 3.2.11, we discussed the Aubin property). The condition (3.59) in view of Remark 3.3.5 reads as

$$\left. \begin{aligned} (\beta_1, -R(\beta_1 + \beta_2)) &\in T\left(\left(\bar{z}_1, \bar{p}_1 - R(\bar{z}_1 + \bar{z}_2)\right); \text{gph } F_1\right) \\ (\beta_2, -R(\beta_1 + \beta_2)) &\in T\left(\left(\bar{z}_2, \bar{p}_2 - R(\bar{z}_1 + \bar{z}_2)\right); \text{gph } F_2\right) \end{aligned} \right\} \implies \beta_1 = \beta_2 = 0, \quad (3.61)$$

where the first tangent cone, $T((x, y); \text{gph } F_1)$, can be calculated as

$$\left\{ \begin{array}{l} \mathbb{R} \begin{pmatrix} 1 \\ a \end{pmatrix} \\ \mathbb{R} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \mathbb{R} \begin{pmatrix} 1 \\ \varphi'(x) \end{pmatrix} \\ \mathbb{R}_+ \begin{pmatrix} -1 \\ -a \end{pmatrix} \cup \mathbb{R}_+ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \mathbb{R}_+ \begin{pmatrix} 1 \\ \varphi'(0) \end{pmatrix} \cup \mathbb{R}_+ \begin{pmatrix} 0 \\ -1 \end{pmatrix} \\ \mathbb{R}_+ \begin{pmatrix} 1 \\ a \end{pmatrix} \cup \mathbb{R}_+ \begin{pmatrix} -1 \\ -\varphi'(\alpha) \end{pmatrix} \\ \emptyset \end{array} \right. \begin{array}{l} x < 0, y = ax + V_1, \text{ or} \\ x > \alpha, y = a(x - \alpha) + \varphi(\alpha), \\ x = 0, y \in (V_1, \varphi(0)), \\ x \in (0, \alpha), y = \varphi(x), \\ x = 0, y = V_1, \\ x = 0, y = \varphi(0), \\ x = \alpha, y = \varphi(\alpha), \\ \text{otherwise.} \end{array}$$

and the second tangent cone, $T((x, y); \text{gph } F_2)$, can be computed similarly to obtain

$$\left\{ \begin{array}{l} \mathbb{R} \begin{pmatrix} 1 \\ b \end{pmatrix} \\ \mathbb{R} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \mathbb{R}_+ \begin{pmatrix} -1 \\ -b \end{pmatrix} \cup \mathbb{R}_+ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \mathbb{R}_+ \begin{pmatrix} 1 \\ b \end{pmatrix} \cup \mathbb{R}_+ \begin{pmatrix} 0 \\ -1 \end{pmatrix} \\ \emptyset \end{array} \right. \begin{array}{l} x < 0, y = bx - V, \text{ or} \\ x > 0, y = bx + V \\ x = 0, y \in (-V, V), \\ x = 0, y = -V, \\ x = 0, y = V, \\ \text{otherwise.} \end{array}$$

In order to facilitate the process of checking the inclusions in (3.61), we multiply them by the matrices $M_3 := \begin{pmatrix} 1 & 0 \\ -1 & \frac{-1}{R} \end{pmatrix}$ and $M_4 := \begin{pmatrix} -1 & \frac{-1}{R} \\ 1 & 0 \end{pmatrix}$, respectively, to get the following equivalent condition

$$M_3 \left(T((\bar{z}_1, \bar{v}_1); \text{gph } F_1) \right) \cap M_4 \left(T((\bar{z}_2, \bar{v}_2); \text{gph } F_2) \right) = \emptyset.$$

Hence, $M_3 \left(T((\bar{z}_1, \bar{v}_1); \text{gph } F_1) \right)$ is given by

$$\left\{ \begin{array}{l}
\mathbb{R} \begin{pmatrix} 1 \\ \frac{-a}{R} - 1 \end{pmatrix} \\
\mathbb{R} \begin{pmatrix} 0 \\ \frac{-1}{R} \end{pmatrix} \\
\mathbb{R} \begin{pmatrix} 1 \\ \frac{-\varphi'(\bar{z}_1)}{R} - 1 \end{pmatrix} \\
\mathbb{R}_+ \begin{pmatrix} -1 \\ \frac{a}{R} + 1 \end{pmatrix} \cup \mathbb{R}_+ \begin{pmatrix} 0 \\ \frac{-1}{R} \end{pmatrix} \\
\mathbb{R}_+ \begin{pmatrix} 1 \\ \frac{-\varphi'(0)}{R} - 1 \end{pmatrix} \cup \mathbb{R}_+ \begin{pmatrix} 0 \\ \frac{1}{R} \end{pmatrix} \\
\mathbb{R}_+ \begin{pmatrix} 1 \\ \frac{-a}{R} - 1 \end{pmatrix} \cup \mathbb{R}_+ \begin{pmatrix} -1 \\ \frac{\varphi'(\alpha)}{R} + 1 \end{pmatrix} \\
\emptyset
\end{array} \right. \begin{array}{l}
\bar{z}_1 < 0, \bar{v}_1 = a\bar{z}_1 + V_1, \text{ or} \\
\bar{z}_1 > \alpha, \bar{v}_1 = a(\bar{z}_1 - \alpha) + \varphi(\alpha), \\
\bar{z}_1 = 0, \bar{v}_1 \in (V_1, \varphi(0)), \\
\bar{z}_1 \in (0, \alpha), \bar{v}_1 = \varphi(\bar{z}_1), \\
\bar{z}_1 = 0, \bar{v}_1 = V_1, \\
\bar{z}_1 = 0, \bar{v}_1 = \varphi(0), \\
\bar{z}_1 = \alpha, \bar{v}_1 = \varphi(\alpha), \\
\text{otherwise.}
\end{array}$$

And $M_4(T((\bar{z}_2, \bar{v}_2); \text{gph } F_2))$ is given by

$$\left\{ \begin{array}{l}
\mathbb{R} \begin{pmatrix} \frac{-b}{R} - 1 \\ 1 \end{pmatrix} \\
\mathbb{R} \begin{pmatrix} \frac{-1}{R} \\ 0 \end{pmatrix} \\
\mathbb{R}_+ \begin{pmatrix} \frac{b}{R} + 1 \\ -1 \end{pmatrix} \cup \mathbb{R}_+ \begin{pmatrix} \frac{-1}{R} \\ 0 \end{pmatrix} \\
\mathbb{R}_+ \begin{pmatrix} \frac{-b}{R} - 1 \\ 1 \end{pmatrix} \cup \mathbb{R}_+ \begin{pmatrix} \frac{1}{R} \\ 0 \end{pmatrix} \\
\emptyset
\end{array} \right. \begin{array}{l}
\bar{z}_2 < 0, \bar{v}_2 = b\bar{z}_2 - V, \text{ or} \\
\bar{z}_2 > 0, \bar{v}_2 = b\bar{z}_2 + V \\
\bar{z}_2 = 0, \bar{v}_2 \in (-V, V), \\
\bar{z}_2 = 0, \bar{v}_2 = -V, \\
\bar{z}_2 = 0, \bar{v}_2 = V, \\
\text{otherwise.}
\end{array}$$

One can easily check each sub-domain of $M_3(T((\bar{z}_1, \bar{v}_1); \text{gph } F_1))$ with sub-domains of $M_4(T((\bar{z}_2, \bar{v}_2); \text{gph } F_2))$ to see if they have a non-zero intersection. The result is that the isolated calmness of the solution map depends on the parameters of a particular circuit, and the reference point.

For example, in the case of $\bar{z}_1 < 0$, $\bar{v}_1 = a\bar{z}_1 + V_1$ or $\bar{z}_1 > \alpha$, $\bar{v}_1 = a(\bar{z}_1 - \alpha) + \varphi(\alpha)$, and for any $(\bar{z}_2, \bar{v}_2) \in \text{gph } F_2$, the solution map has the isolated calmness property. While for $\bar{z}_1 \in (0, \alpha)$, $\bar{v}_1 = \varphi(\bar{z}_1)$, the condition (3.61) will be violated for $\bar{z}_2 \neq 0$ if $\varphi'(\bar{z}_1) = \frac{-Rb}{R+b}$; and for $\bar{z}_2 = 0$, $\bar{v}_2 \in (-V, V)$ if $\varphi'(\bar{z}_1) = -R$; and for $\bar{z}_2 = 0$, $\bar{v}_2 \in \{-V, V\}$ if $\varphi'(\bar{z}_1) \in \left\{ -R, \frac{-Rb}{R+b} \right\}$.

3.3.1 Isolated Calmness Without Injectivity Assumption

Up to now, we have assumed that $m \leq n$. However, one can have $m > n$ in some applications. In such a case, one cannot expect the assumptions (A1) and (A4) (cf. Note 3.2.2) to hold true. So till the end of this section, $m, n \in \mathbb{N}$ are not related to each other by an inequality. Also, in order to avoid confusion during the calculations of this subsection, we refer to the zero vector of different spaces with the space indicated as a subscript (like $0_{\mathbb{R}^d}$). We assume the following assumption (instead of (A1)):

($\widetilde{\text{A1}}$) Suppose that there is $\bar{v} \in F(C\bar{z})$ such that

$$\bar{p} = f(\bar{z}) + B\bar{v} \quad \text{and} \quad \overline{\bigcup_{t>0} \frac{\text{rge } F_C - \bar{v}}{t}} \cap \ker B = \{0_{\mathbb{R}^m}\}.$$

Of course, if \bar{v} is an interior point of $\text{rge } F_C$ then ($\widetilde{\text{A1}}$) reduces to (A1). However, when \bar{v} is in the boundary of $\text{rge } F_C$, this is not true. For example, consider $\text{rge } F_C = (-3, -1] \subset \mathbb{R}$. Then, for $\bar{v} = -2$, we have $\text{rge } F_C - \bar{v} = (-1, 1]$, and

$$\overline{\bigcup_{t>0} \frac{\text{rge } F_C - \bar{v}}{t}} \cap \ker B = \mathbb{R} \cap \ker B,$$

which implies ($\widetilde{\text{A1}}$) being equivalent with (A1). While for $\bar{v} = -1$, we have $\text{rge } F_C - \bar{v} = (-2, 0]$, and

$$\overline{\bigcup_{t>0} \frac{\text{rge } F_C - \bar{v}}{t}} \cap \ker B = \mathbb{R}_- \cap \ker B,$$

which is not necessarily equal to $\ker B$.

Lemma 3.3.8. [4, Lemma 2, p. 353] Let $G \in \mathbb{R}^{l \times d}$, let $\Xi \subset \mathbb{R}^d$, and set $\Lambda = G(\Xi)$. Suppose that $\bar{x} \in \Lambda$ and $\bar{y} \in \Xi$ are such that

$$G(\bar{y}) = \bar{x} \quad \text{and} \quad \overline{\bigcup_{t>0} \frac{\Xi - \bar{y}}{t}} \cap \ker G = \{0_{\mathbb{R}^d}\}. \quad (3.62)$$

Then

$$T(\bar{x}; \Lambda) = G(T(\bar{y}; \Xi)).$$

Proof. To prove that $G(T(\bar{y}; \Xi)) \subset T(\bar{x}; \Lambda)$, pick any $w \in G(T(\bar{y}; \Xi))$. Find $v \in T(\bar{y}; \Xi)$ with $G(v) = w$. Thus there is $(t_n)_{n \in \mathbb{N}}$ in $(0, \infty)$ converging to 0 and $(v_n)_{n \in \mathbb{N}}$ in \mathbb{R}^d converging to v such that $\bar{y} + t_n v_n \in \Xi$ whenever $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, put $w_n := G(v_n)$. Clearly, $(w_n)_{n \in \mathbb{N}}$ converges to w . Moreover,

$$x + t_n w_n = G(\bar{y} + t_n v_n) \in G(\Xi) = \Lambda \quad \text{whenever } n \in \mathbb{N}.$$

So $w \in T(\bar{x}; \Lambda)$.

To see the opposite inclusion, pick any $w \in T(\bar{x}; \Lambda)$. Find $(t_n)_{n \in \mathbb{N}}$ in $(0, \infty)$ converging to 0 and $(w_n)_{n \in \mathbb{N}}$ in \mathbb{R}^l converging to w such that

$$\bar{x} + t_n w_n \in \Lambda \text{ or } G(\bar{y}) + t_n w_n \in G(\Xi) \quad \text{for each } n \in \mathbb{N}.$$

For each $n \in \mathbb{N}$, find $v_n \in \Xi$ such that $w_n = G((v_n - \bar{y})/t_n)$, and put $u_n = (v_n - \bar{y})/t_n$. We claim that $(u_n)_{n \in \mathbb{N}}$ is bounded.

Suppose on the contrary that this is not the case. Passing to a subsequence if necessary, we may assume that $(u_n/\|u_n\|)_{n \in \mathbb{N}}$ converges, to $\bar{u} \in S_{\mathbb{R}^d}$, the unit sphere in \mathbb{R}^d . For each $n \in \mathbb{N}$, one has that

$$\frac{u_n}{\|u_n\|} = \frac{v_n - \bar{y}}{\|v_n - \bar{y}\|} \quad \text{and} \quad \frac{w_n}{\|u_n\|} = G\left(\frac{u_n}{\|u_n\|}\right).$$

Passing to the limit for $n \rightarrow \infty$ one gets, that

$$\bar{u} \in \bigcup_{t>0} \frac{\Xi - \bar{y}}{t} \quad \text{and} \quad G(\bar{u}) = \lim_{n \rightarrow \infty} \frac{w_n}{\|u_n\|} = 0_{\mathbb{R}^l},$$

which, by condition (3.62) implies $\bar{u} = 0_{\mathbb{R}^d}$, a contradiction since $S_{\mathbb{R}^d} \ni \bar{u} \neq 0_{\mathbb{R}^d}$.

Having the claim in hand, one may assume without loss of generality that $(u_n)_{n \in \mathbb{N}}$ converges to some $u \in \mathbb{R}^d$. For each $n \in \mathbb{N}$, we have $\bar{y} + t_n u_n = v_n \in \Xi$, therefore $u \in T(\bar{y}; \Xi)$. Moreover, $w = G(u)$. The proof is finished. \square

Proposition 3.3.9. [4, Proposition 1, p. 354] Assume that $(\widetilde{A1})$ and (A2) hold true. Then

$$D\Phi(\bar{z} | \bar{p})(b) = \nabla f(\bar{z})b + B DF_C(C\bar{z} | \bar{v})(Cb) \quad \text{whenever } b \in \mathbb{R}^n.$$

Proof. Fix any $b \in \mathbb{R}^n$. By Proposition 3.1.22 and Remark 3.1.18, we have

$$D\Phi(\bar{z} | \bar{p})(b) = \nabla f(\bar{z})b + DQ(\bar{z} | \bar{p} - f(\bar{z}))(b).$$

Define $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ by $H(z) := F(Cz)$, for $z \in \mathbb{R}^n$. Then,

$$\text{gph } Q = G(\text{gph } H) \quad \text{with} \quad G := \begin{pmatrix} I_n & 0 \\ 0 & B \end{pmatrix}$$

Then $\ker G = \{0_{\mathbb{R}^n}\} \times \ker B$ and $\text{gph } H \subset \mathbb{R}^n \times \text{rge } F_C$. Therefore

$$\bigcup_{t>0} \frac{\overline{\text{gph } H - (\bar{z}, \bar{v})}}{t} \cap \ker G \subset \{0_{\mathbb{R}^n}\} \times \left(\bigcup_{t>0} \frac{\overline{\text{rge } F_C - \bar{v}}}{t} \cap \ker B \right) = \{0_{\mathbb{R}^{n+m}}\}.$$

Applying Lemma 3.3.8 with $l := 2n$, $d := n + m$, $\Xi := \text{gph } H$, $\bar{x} = (\bar{z}, \bar{p} - f(\bar{z}))$, and $\bar{y} = (\bar{z}, \bar{v})$ reveals that

$$T\left((\bar{z}, \bar{p} - f(\bar{z})); \text{gph } Q\right) = \left\{ \begin{pmatrix} b \\ Bc \end{pmatrix} : \begin{pmatrix} b \\ c \end{pmatrix} \in T((\bar{z}, \bar{v}); \text{gph } H) \right\}.$$

This means that $DQ(\bar{z}, \bar{p} - f(\bar{z}))(b) = BDH(\bar{z}, \bar{v})(b)$. Moreover,

$$\text{gph } H = \left\{ \begin{pmatrix} z \\ v \end{pmatrix} \in \mathbb{R}^n \times \mathbb{R}^m : E \begin{pmatrix} z \\ v \end{pmatrix} \in \text{gph } F_C \right\} \quad \text{with} \quad E := \begin{pmatrix} C & 0 \\ 0 & I_m \end{pmatrix}.$$

Finally, using the claim in the first part of the proof of Lemma 3.3.1 with $k := 2m$, $d := n + m$, $\Gamma := \text{gph } F_C$, and $\bar{y} = (\bar{z}, \bar{v})$ reveals that

$$T\left((\bar{z}, \bar{v}); \text{gph } H\right) = \left\{ \begin{pmatrix} b \\ c \end{pmatrix} \in \mathbb{R}^n \times \mathbb{R}^m : E \begin{pmatrix} b \\ c \end{pmatrix} \in T((C\bar{z}, \bar{v}); \text{gph } F_C) \right\}.$$

Hence $DH(\bar{z} | \bar{v})(b) = DF_C(C\bar{z} | \bar{v})(Cb)$. Combining the above facts ends the proof. \square

Theorem 3.3.10. [4, Theorem 3, p. 355] Suppose that $(\widetilde{A1})$, (A2), and (A3) hold true. Then S has the isolated calmness property at (\bar{p}, \bar{z}) if and only if

$$0 \in \nabla f(\bar{z})b + B DF_C(C\bar{z} | \bar{v})(Cb) \quad \Rightarrow \quad b = 0_{\mathbb{R}^n}.$$

Proof. The proof is similar to the proof of Theorem 3.3.3, and could be concluded by combining Corollary 3.1.41, Theorem 3.1.40, and Proposition 3.3.9 above; noticing that $x \in D\Phi(\bar{z} | \bar{p})(y)$ if and only if $y \in DS(\bar{p} | \bar{z})(x)$. \square

To use the above statement, one has to know the range of the matrix C . Sometimes the following sufficient condition may be useful.

Corollary 3.3.11. [4, Corollary 2, p. 355] Under the assumptions of Theorem 3.3.10, S has the isolated calmness property at (\bar{p}, \bar{z}) , if $b = 0_{\mathbb{R}^n}$ is the only point that

$$\nabla f(\bar{z})b + Bw = 0 \quad \text{and} \quad (Cb, w) \in T\left((C\bar{z}, \bar{v}); \text{gph } F\right) \quad \text{for some} \quad w \in \mathbb{R}^m. \quad (3.63)$$

Proof. Clearly, $\text{gph } F_C \subset \text{gph } F$, so $T\left((C\bar{z}, \bar{v}); \text{gph } F_C\right) \subset T\left((C\bar{z}, \bar{v}); \text{gph } F\right)$. Take any $b \in \mathbb{R}^n$ such that $0 \in \nabla f(\bar{z})b + BDF_C(C\bar{z}, \bar{v})(Cb)$. Find $w \in \mathbb{R}^m$ such that $0 = \nabla f(\bar{z})b + Bw$ and $(Cb, w) \in T\left((C\bar{z}, \bar{v}); \text{gph } F_C\right)$. Thus $b = 0_{\mathbb{R}^n}$. \square

Let us finish this subsection with an example of a circuit that exhibits the failure of the assumption (A1), yet with the tools provided here under the assumption $(\widetilde{\text{A1}})$, we can investigate the isolated calmness property.

Example 3.3.12. (Sampling Gate) [4, page 355]

Consider the circuit in Figure 3.7, composed of four diodes D_1, D_2, D_3, D_4 which are controlled symmetrically by gate voltages $+V_c$ and $-V_c$, and the control resistors $R_c > 0$. Assume that all the diodes have the same characteristics given for $V_{D1} < 0 < V_{D2}$ by

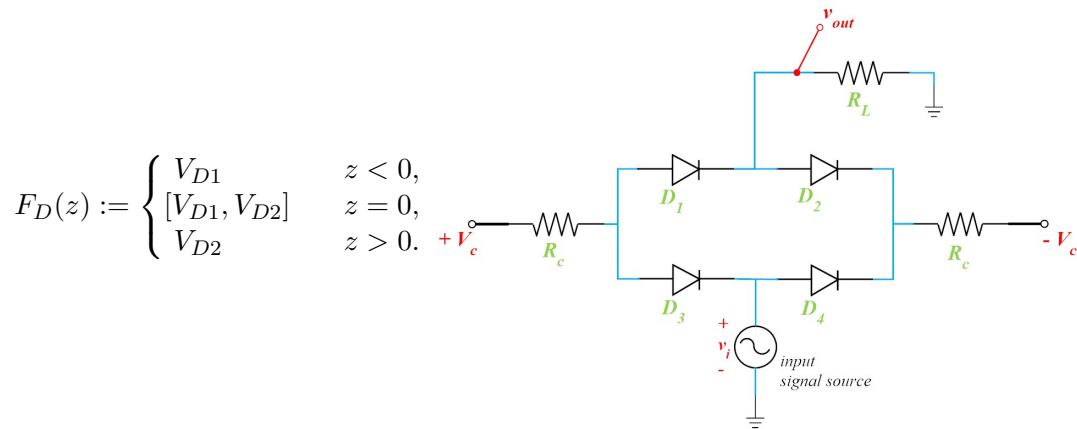


FIGURE 3.7: Calculating isolated calmness in a circuit without the (A1) assumption

In Example 2.3.4, we obtained the mathematical model for this circuit in the form of (3.46) with $m = 4$, $n = 3$, $p = (v_i, 2V_c, 0)^T$, $C = B^T$, $f(z) = Az$ for $z \in \mathbb{R}^3$, where

$$A = \begin{pmatrix} R_L & 0 & 0 \\ 0 & 2R_c & 0 \\ 0 & 0 & 0 \end{pmatrix}, \quad B = \begin{pmatrix} 0 & -1 & 0 & 1 \\ 0 & 0 & 1 & 1 \\ 1 & 1 & -1 & -1 \end{pmatrix}, \quad \text{and} \quad F(y) = \begin{pmatrix} F_D(y_1) \\ F_D(y_2) \\ F_D(y_3) \\ F_D(y_4) \end{pmatrix} \quad \text{for } y \in \mathbb{R}^4.$$

Since B is not invertible (in fact, $\ker B = \{\lambda(-1, 1, -1, 1)^T \mid \lambda \in \mathbb{R}\}$), we can not use Theorem 3.3.3 or Corollary 3.3.4. However, for the points $(\bar{z}, \bar{v}) \in \text{gph } F_C$ with \bar{v} in the boundary of $\text{rge } F_C$, we can check whether the assumptions of Theorem 3.3.10 are

satisfied and then, use Corollary 3.3.11. Let $\bar{p} = (0, 2V_{D1}, 0)^T$ and $\bar{z} = (0, 0, 0)^T$. Put $\bar{v} = (V_{D1}, V_{D1}, V_{D1}, V_{D1})^T$. Then

$$(\bar{z}, \bar{p}) \in \text{gph } \Phi, \quad \bar{p} = A\bar{z} + B\bar{v} \quad \text{and} \quad \bar{v} \in F(C\bar{z}).$$

Moreover, $\text{rge } F_C \subset \text{rge } F = [V_{D1}, V_{D2}]^4$. Therefore

$$\bigcup_{t>0} \overline{\frac{\text{rge } F_C - \bar{v}}{t}} \cap \ker B \subset \mathbb{R}_+^4 \cap \mathbb{R} \{(-1, 1, -1, 1)^T\} = \{0_{\mathbb{R}^4}\}.$$

Hence $(\widetilde{A1})$ holds. In order to check criteria (3.63), let $b \in \mathbb{R}^3$ be such that

$$Ab + Bw = 0_{\mathbb{R}^3} \quad \text{and} \quad (Cb, w) \in T((0, \bar{v}); \text{gph } F) \quad \text{for some } w \in \mathbb{R}^4.$$

Clearly, $T((0, \bar{v}); \text{gph } F)$ equals to

$$\prod_{j=1}^4 T((0, V_{D1}); \text{gph } F_D) = \left[\left\{ \begin{pmatrix} x \\ y \end{pmatrix} \in \mathbb{R}^2 \mid x \leq 0, y \geq 0, xy = 0 \right\} \right]^4.$$

Therefore, one should have $\langle Cb, w \rangle = 0$. But

$$\begin{aligned} \langle Cb, w \rangle &= \langle B^T b, w \rangle = \langle b, Bw \rangle = \langle b, -Ab \rangle = \left\langle \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix}, \begin{pmatrix} -R_L b_1 \\ -2R_c b_2 \\ 0 \end{pmatrix} \right\rangle \\ &= -R_L b_1^2 - 2R_c b_2^2, \end{aligned}$$

which implies that $b_1 = b_2 = 0$. Thus, $Cb = (b_3, b_3, -b_3, -b_3)^T$. From the tangent cone relation, we know that all the coordinates of this vector should be non-positive, so $b_3 = 0$. Applying Corollary 3.3.11 one concludes that S has the isolated calmness property at the reference point.

3.4 Results about Calmness

Although we know that Aubin property or isolated calmness implies calmness, up to the results obtained so far, the criteria guaranteeing the calmness are more complicated than those for the isolated calmness or Aubin property. Hence, one should always consider the latter property first, and employ more sophisticated tools of this section only in case of their failure.

One reason for this difficulty is the lack of stability under perturbation for calmness. To be more clear, suppose that S has the Aubin property at (\bar{p}, \bar{z}) . Then, there exists $\gamma > 0$

such that for every function $\tilde{f} : \mathbb{R}^n \rightarrow \mathbb{R}^n$, every point $(\tilde{p}, \tilde{z}) \in \mathbb{B}_\gamma(\tilde{p}) \times \mathbb{B}_\gamma(\tilde{z})$ with

$$\begin{aligned} \tilde{p} \in \tilde{f}(\tilde{z}) + BF(C\tilde{z}), \quad \|\tilde{f}(\tilde{z}) - f(\tilde{z})\| \leq \gamma, \text{ and} \\ \left\| \left(\tilde{f}(z') - f(z') \right) - \left(\tilde{f}(z'') - f(z'') \right) \right\| \leq \gamma \|z' - z''\| \quad \text{whenever } z', z'' \in \mathbb{B}_\gamma(\tilde{z}), \end{aligned}$$

the mapping $\tilde{S} := (\tilde{f} + BF(C\cdot))^{-1}$ has the Aubin property at (\tilde{p}, \tilde{z}) (cf. [30, Theorem 1]). A similar statement can be said about isolated calmness (cf. [24]), but not for calmness. In fact, in Example 1.3.15 we have observed that we cannot expect to find an equivalent characterization of metric sub-regularity by means of a derivative-like object which is really computable in general, because any such derivative would have the property that the derivatives of the multifunctions G and \tilde{G} of that example are different. But it is possible that G and \tilde{G} differ only by a C^∞ -function, where all derivatives vanish at \bar{x} and so the usual calculus rules cannot be valid.

We would provide a sufficient condition (Proposition 3.4.3) for calmness and an estimate for calmness modulus in this section. In order to prove that proposition, we need some tools connecting the calmness of the solution mapping in the setting (3.46) to the outer subdifferential of an auxiliary function.

Lemma 3.4.1. [41, Lemma 1, p. 441] *Let X, U, Y be normed spaces. Consider multifunctions $M : X \rightrightarrows Y$ defined by*

$$M(y) := \{x \in X : g(x) + y \in \Lambda\},$$

where $g : X \rightarrow Y$ and $\Lambda \subseteq Y$ is a closed subset; and $M^* : U \rightrightarrows X$ defined on the basis of some locally Lipschitzian (with respect to the product topology) function $h : U \times X \rightarrow Y$ by means of

$$M^*(u) := \{x \in X \mid h(u, x) \in \Lambda\}.$$

Assume that $h(\bar{u}, \bar{x}) \in \Lambda$ for some $\bar{u} \in U$ and $\bar{x} \in X$. Then, M^* is calm at (\bar{u}, \bar{x}) provided that M is calm at $(0, \bar{x})$ with $g(x) := h(\bar{u}, x)$ for all $x \in X$.

Proof. The local Lipschitz continuity of h and the calmness of M yield constants $K, L, \epsilon > 0$ such that

$$\begin{aligned} \|h(u', x) - h(u'', x)\| &\leq K \|u' - u''\| \quad \text{for all } u', u'' \in \mathbb{B}_\epsilon(\bar{u}), \quad \text{for all } x \in \mathbb{B}_\epsilon(\bar{x}), \\ d(x, M(0)) &\leq L \|y\| \quad \text{for all } y \in \mathbb{B}_\epsilon(0), \quad \text{for all } x \in \mathbb{B}_\epsilon(\bar{x}) \cap M(y). \end{aligned}$$

Choose ϵ' such that $0 < \epsilon' \leq \epsilon$ and

$$\|h(u, x) - h(\bar{u}, \bar{x})\| \leq \epsilon \quad \text{for all } (u, x) \in \mathbb{B}_{\epsilon'}(\bar{u}) \times \mathbb{B}_{\epsilon'}(\bar{x}).$$

Let $x \in M^*(u) \cap \mathbb{B}_{\epsilon'}(\bar{x})$, and $u \in \mathbb{B}_{\epsilon'}(\bar{u})$, be arbitrary. Then, $x \in M(h(u, x) - g(x)) \cap \mathbb{B}_{\epsilon'}(\bar{x})$, by definition of M and M^* . It follows the calmness of M^* at (\bar{u}, \bar{x}) :

$$d(x, M^*(\bar{u})) = d(x, M(0)) \leq L \|h(u, x) - g(x)\| \leq LK \|u - \bar{u}\|.$$

□

Theorem 3.4.2. [45, Theorem 2.1, p. 203]

Let $f : \mathbb{R}^n \rightarrow \bar{\mathbb{R}}$ be lower semicontinuous in a neighborhood of $\bar{x} \in \mathbb{R}^n$ and $f(\bar{x}) = 0$.

Fix a $\bar{\gamma} > 0$ and consider the following properties:

- (a) For any $\gamma < \bar{\gamma}$ there is a $\delta > 0$ such that $d(x, [f \leq 0]) \leq \gamma^{-1} f^+(x)$ if $\|x - \bar{x}\| < \delta$;
- (b) For any $\gamma < \bar{\gamma}$ there is a $\delta > 0$ such that $\liminf \|h_k\|^{-1} f(x_k + h_k) \geq \gamma$ whenever $f(x_k) \leq 0$, $\|x_k - \bar{x}\| \leq \delta$ and $h_k \rightarrow 0, h_k \in N_F([f \leq 0], x_k) \setminus \{0\}$;
- (c) For any $\gamma < \bar{\gamma}$ there is a $\delta > 0$ such that $\|x^*\| \geq \gamma$ if $x^* \in \partial f(x)$ for some x satisfying $\|x - \bar{x}\| < 2\delta$ and $0 < f(x) < \delta\gamma$;
- (d) $\|x^*\| \geq \bar{\gamma}$ if $x^* \in \partial_{>} f(\bar{x})$;
- (e) For any $\gamma < \bar{\gamma}$ there is a $\delta > 0$ such that $|\nabla f|(x) \geq \gamma$ if $\|x - \bar{x}\| < 2\delta$ and $0 < f(x) < \delta\gamma$.

Then, (e) \Rightarrow (d) \Leftrightarrow (c) \Rightarrow (a) \Leftarrow (b).

Proposition 3.4.3 (Calmness Sufficient Criterion for the Solution Mapping of GE). [3, Proposition 5.1, p. 98]

Suppose that the assumptions (A1) - (A3) are satisfied. Put $\Lambda = \text{gph } F \times \text{rge } B$ and define the functions

$$g : \mathbb{R}^n \rightarrow \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^n \text{ by } g(z) := (Cz, (B^T B)^{-1} B^T (\bar{p} - f(z)), \bar{p} - f(z)), \text{ and}$$

$$h : \mathbb{R}^n \rightarrow \mathbb{R}^+ \text{ by } h(z) := d(g(z), \Lambda) \text{ for each } z \in \mathbb{R}^n.$$

Then S has the calmness property at (\bar{p}, \bar{z}) , provided that

$$0 \notin \partial_{>} h(\bar{z}). \quad (3.64)$$

Proof. The inclusion (3.46) says that, for each $p \in \mathbb{R}^n$, one has

$$S(p) = \left\{ z \in \mathbb{R}^n : \left(Cz, (B^T B)^{-1} B^T (p - f(z)), p - f(z) \right) \in \Lambda \right\}.$$

Indeed, fix any $p \in \mathbb{R}^n$. First, pick any $z \in \mathbb{R}^n$ with $(z, p) \in \text{gph } \Phi$. Inclusion (3.46) reveals that $p - f(z) \in \text{rge } B$. We have already mentioned that $B^T B \in \mathbb{R}^{m \times m}$ is non-singular. Hence, (3.46) implies that $\left(Cz, (B^T B)^{-1} B^T (p - f(z)) \right) \in \text{gph } F$. On the other hand, fix an arbitrary $z \in \mathbb{R}^n$ with

$$\left(Cz, (B^T B)^{-1} B^T (p - f(z)), p - f(z) \right) \in \Lambda.$$

Then $B^T(p - f(z)) = B^T Bw$ for some $w \in F(Cz)$. As $p - f(z) \in \text{rge } B$, we have $p - f(z) - Bw \in \ker B^T \cap \text{rge } B = \{0\}$. Therefore $(z, p) \in \text{gph } \Phi$.

Define the mapping $M : \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ as follows:

$$M(y) := \{z \in \mathbb{R}^n : g(z) + y \in \Lambda\}, \quad y \in \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^n.$$

As Λ is closed and the mapping

$$\mathbb{R}^n \times \mathbb{R}^n \ni (p, z) \mapsto \left(Cz, (B^T B)^{-1} B^T(p - f(z)), p - f(z) \right) \in \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^n \quad (3.65)$$

is continuously differentiable (hence strictly differentiable and therefore locally Lipschitz continuous), Lemma 3.4.1 says that S is calm at (\bar{p}, \bar{z}) , provided that so is M at $(0, \bar{z})$. As observed in [41, p. 438], M is calm at $(0, \bar{z})$ if and only if there are $L > 0$ and $\epsilon > 0$ such that

$$d(z, g^{-1}(\Lambda)) \leq L d(g(z), \Lambda) \quad \text{whenever } z \in \mathbb{B}(\bar{z}, \epsilon).$$

Since g is locally Lipschitz continuous, so is h . Moreover, $[h \leq 0] := \{u \in \mathbb{R}^n : h(u) \leq 0\} = g^{-1}(\Lambda)$. Theorem 3.4.2 reveals that M is calm at $(0, \bar{z})$, provided that (3.64) holds.

The proposition is proved. \square

Remark 3.4.4. (Estimate for Calmness Modulus)

Clearly, if the matrix B is surjective, it suffices to consider $\Lambda = \text{gph } F$, and $g(z) = (Cz, B^{-1}(\bar{p} - f(z)))$, $z \in \mathbb{R}^n$. Furthermore, Theorem 3.4.2 also gives an upper estimate of the corresponding calmness modulus.

Namely, if there is $\gamma > 0$ such that $\|\xi\| \geq \gamma$ for each $\xi \in \partial_{>} h(\bar{z})$, then S is calm at (\bar{p}, \bar{z}) with modulus not exceeding K/γ , where $K > 0$ is the Lipschitz constant of the map φ at (\bar{p}, \bar{z}) . φ is defined in the proof of the Proposition 3.4.3 as

$$\mathbb{R}^n \times \mathbb{R}^n \ni (p, z) \mapsto \left(Cz, (B^T B)^{-1} B^T(p - f(z)), p - f(z) \right) \in \mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^n$$

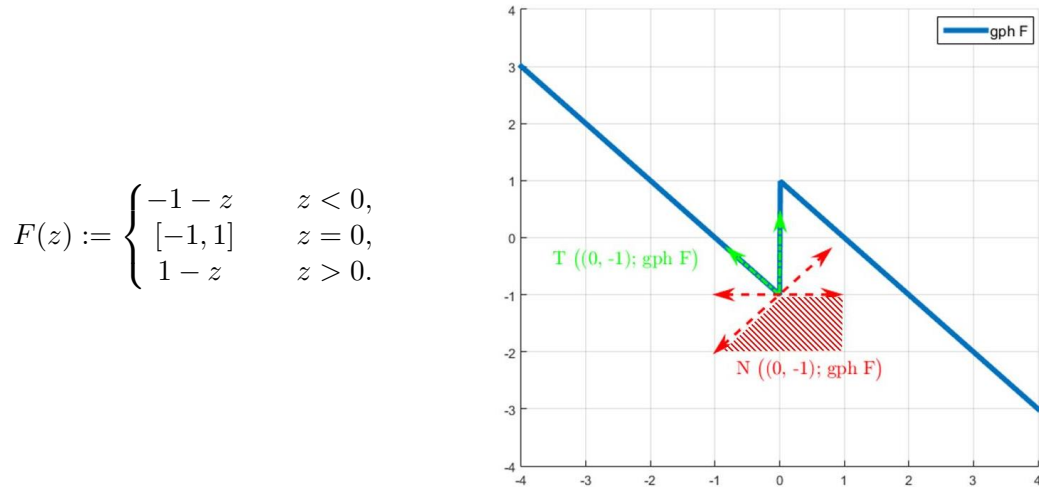
To be more precise, there is $\epsilon > 0$ such that

$$\|\varphi(p', z) - \varphi(p'', z)\| \leq K \|p' - p''\| \quad \text{whenever } p', p'' \in \mathbb{B}_\epsilon(\bar{p}), z \in \mathbb{B}_\epsilon(\bar{z}).$$

Without loss of generality, we may assume that the norm on $\mathbb{R}^m \times \mathbb{R}^m \times \mathbb{R}^n$ for each point (u, v, w) is given by $\|(u, v, w)\| = \|u\|_m + \|v\|_m + \|w\|_n$, where $\|\cdot\|_k$ is the Euclidean norm on \mathbb{R}^k .

Example 3.4.5. [3, page 99]

Suppose that $f : \mathbb{R} \rightarrow \mathbb{R}$ is given by $f(z) := z$ for every $z \in \mathbb{R}$, and $F : \mathbb{R} \rightrightarrows \mathbb{R}$ is defined as: Let $\bar{z} := 0$ and $\bar{p} := -1$. It is easy to verify that both criteria (3.51) and

FIGURE 3.8: Graph of F with normal and tangent cones at $(0, -1)$.

(3.59) are violated. In fact, $\bar{v} = \bar{p} - f(\bar{z}) = -1$, $\nabla f(\bar{z}) = 1$, and (3.51) reads as

$$(\xi, \xi) \in -N((0, -1); \text{gph } F) \implies \xi = 0,$$

which is not true, since points other than $(0, 0)$ from the identity line belong to the normal cone (see Figure 3.8). Indeed, $N((0, -1); \text{gph } F) = \mathbb{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} -1 \\ -1 \end{pmatrix} \cup \text{cone} \left\{ \begin{pmatrix} 1 \\ 0 \end{pmatrix}, \begin{pmatrix} -1 \\ -1 \end{pmatrix} \right\}$. The condition (3.59) reads as

$$(b, -b) \in T((0, -1); \text{gph } F) \implies b = 0,$$

which is not true, since points other than $(0, 0)$ from the bisector of second quadrant belong to the tangent cone (see Figure 3.8). Indeed, $T((0, -1); \text{gph } F) = \mathbb{R}_+ \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cup \mathbb{R}_+ \begin{pmatrix} -1 \\ 1 \end{pmatrix}$. So, S has neither the Aubin nor the isolated calmness property at the reference point. Now we want to check the sufficient criteria of Proposition 3.4.3 for the calmness property.

Note that, $g(z) = (z, -1 - z, -1 - z)$ in this example and $\Lambda = \text{gph } F \times \mathbb{R}$. Thus,

$$h(z) := d(g(z), \Lambda) = d((z, -1 - z, -1 - z), \text{gph } F \times \mathbb{R}) = d((z, -1 - z), \text{gph } F).$$

Let us compute the outer subdifferential of h at $z = 0$. Recall that, the set $\partial_{>}h(0)$ contains those points $\xi \in \mathbb{R}$ such that there are sequences $(z_k)_{k \in \mathbb{N}}$ and $(\xi_k)_{k \in \mathbb{N}}$ converging to 0 and ξ , respectively, with $h(z_k) \downarrow h(0) = 0$ as $k \rightarrow \infty$ and $\xi_k \in \partial_F h(z_k)$ for each $k \in \mathbb{N}$ (ref. Definition 3.1.43).

If $z < 0$, then the point $(z, -1 - z) \in \text{gph } F$, thus $h(z) = 0$, hence we just need to consider $(z_k)_{k \in \mathbb{N}}$ in $(0, \infty)$. Moreover, there is $\delta > 0$ such that, for each $z \in (0, \delta)$, the

unique nearest point in $\text{gph } F$ to the point $(z, -1 - z)$ is $(0, -1)$. Therefore,

$$h(z) = d((z, -1 - z), \text{gph } F) = d((z, -1 - z), (0, -1)) = \sqrt{2}z \quad \text{for } z \in (0, \delta).$$

Hence,

$$\partial_F h(z) := \left\{ \xi \in \mathbb{R} \mid \liminf_{0 \neq r \rightarrow 0} \frac{h(z+r) - h(z) - \langle \xi, r \rangle}{\|r\|} \geq 0 \right\} = \{\sqrt{2}\},$$

for any point $z \in (0, \delta)$. Let $(z_k)_{k \in \mathbb{N}}$ be any sequence converging to 0 such that $h(z_k) \downarrow h(0) = 0$ as $k \rightarrow \infty$. We may assume without any loss of generality that $z_k \in (0, \delta)$ for each $k \in \mathbb{N}$. Thus $\partial_F h(z_k) = \{\sqrt{2}\}$, hence $\partial_{>} h(0) = \{\sqrt{2}\}$. Therefore, in view of Remark 3.4.4, S is calm at the reference point with modulus not exceeding $\frac{2}{\sqrt{2}}$.

Example 3.4.6. (A Simple Circuit with DIAC) [3, page 100]

Let us go back to Examples 3.2.10 and 3.3.6, and this time, check the calmness property in the circuit of Figure 3.4 with a DIAC, at the reference point $(\bar{p}, \bar{z}) = (-1, 0)$.

In order to avoid non-necessary details and to ease the computations, let us assume that $R = 1$, $V = 1$, and the $i - v$ characteristic of the DIAC is given by a set-valued mapping $\Psi : \mathbb{R} \rightrightarrows \mathbb{R}$, where $\Psi(0) := [-1, 1]$, and Ψ is single-valued and continuously differentiable on $\mathbb{R} \setminus \{0\}$; its graph is symmetric with respect to the origin, and $-1 < \Psi(z) < -1 - z$ whenever $z \in (-\infty, 0)$; and finally $\Psi(0_-) = -1$ and $\Psi'(0_-) = -a$ for some $a > 0$.

Putting $F = \partial|\cdot|$, one infers that there is a continuously differentiable odd function $\psi : \mathbb{R} \rightarrow \mathbb{R}$ such that $\Psi = \psi + F$. Moreover, $\psi(0) = 0$, $\psi'(0) = -a$, and $0 < \psi(z) < -z$ whenever $z < 0$ (see Figure 2.25). Therefore, we arrive at (3.46) with $m = n = 1$, $B = C = I$, and f defined for each $z \in \mathbb{R}$ by $f(z) := z + \psi(z)$.

We have already seen that if $a = 1$, then S has neither the Aubin nor the isolated calmness property at $(-1, 0)$. To see whether Proposition 3.4.3 is helpful or not, note that assumptions (A1) - (A3) are satisfied, $\Lambda = \text{gph } F \times \mathbb{R}$, $g(z) = (z, -1 - f(z), -1 - f(z))$, and $h(z) = d((z, -1 - z - \psi(z)), \text{gph } F)$.

When $z < 0$, $\psi(z) > 0$ and thus, $-1 - z - \psi(z) < -1 - z$ and by assumption $\psi(z) < -z$. So, the nearest point in $\text{gph } F$ to $-1 - z - \psi(z)$ for $z < 0$ is $(z, -1)$. Therefore,

$$h(z) = \|(z, -1 - z - \psi(z)), (z, -1)\| = -z - \psi(z).$$

Hence, $\partial_F h(z) = \{-1 - \psi'(z)\}$ whenever $z < 0$. For each $k \in \mathbb{N}$, put $z_k := \frac{-1}{k}$ and $\xi_k := -1 - \psi'(z_k)$. Then both $(z_k)_{k \in \mathbb{N}}$ and $(\xi_k)_{k \in \mathbb{N}}$ converge to 0 (in fact, $\xi_k \rightarrow -1 - \psi'(0) = -1 + a = 0$). For any $k \in \mathbb{N}$, we have $\xi_k \in \partial_F h(z_k)$, $h(z_k) = -z_k - \psi(z_k) > 0$, and $h(z_k) \downarrow h(0) = 0$. Thus $0 \in \partial_{>} h(0)$. Hence, the sufficient condition (3.64) is not satisfied in this case, and Proposition 3.4.3 does not provide any information about the calmness property of S at the reference point $(-1, 0)$.

However, S is not calm at $(-1, 0)$. Indeed, suppose on the contrary that it is. As $S(-1) = \{0\}$, the point 0 is an isolated point of $S(-1)$, thus S would have the isolated calmness property at $(-1, 0)$, which is not the case as seen before.

3.5 Results in terms of Metric Regularity

Considering the relation between local stability properties and regularity definitions, one may be able to restate the last three sections' results, in terms of (strong) metric (sub-) regularity. All the new things here are about strong metric regularity, for which we need some calculus rules regarding strict graphical derivatives.

We will first provide such rules and then sum up all the results in this area (translated into regularity paradigm) in a single theorem to ease the access (cf. Theorem 3.5.6).

The last subsection of this chapter, will discuss the case of a non-differentiable single-valued part (in the sum $f + F$). We believe that would be the proper time to consider this change in our general assumptions (cf. Note 3.2.2), since all those results we wanted to cover about the static GE has been stated by then.

Let us start with a lemma that provides a chain rule for strict graphical derivatives.

Lemma 3.5.1 (Matrix Operation over Paratingent Cone Relations). [20, Lemma 1, p. 4] *Let $E \in \mathbb{R}^{k \times d}$ be any matrix, let $G \in \mathbb{R}^{l \times d}$ be injective, and let Γ be a subset of $\text{rge } E$. Put $\Xi := E^{-1}(\Gamma)$ and $\Lambda := G(\Xi)$. For $\bar{x} \in \Lambda$ denote by \bar{y} the (unique) point in Ξ with $G(\bar{y}) = \bar{x}$. Then*

$$\tilde{T}(\bar{x}; \Lambda) = \{u \in \mathbb{R}^l : \exists w \in \mathbb{R}^d \text{ such that } u = Gw \text{ and } Ew \in \tilde{T}(E\bar{y}; \Gamma)\}.$$

Proof. The proof would be very similar to the proof of Lemma 3.3.1. One only needs to consider the small necessary changes when replacing T with \tilde{T} .

First, we claim that $\tilde{T}(\bar{y}; \Xi) = \{w \in \mathbb{R}^d : Ew \in \tilde{T}(E\bar{y}; \Gamma)\}$.

Take any $w \in \tilde{T}(\bar{y}; \Xi)$, Find $(t_n)_{n \in \mathbb{N}}$ in $(0, \infty)$, $(y_n)_{n \in \mathbb{N}}$ in Ξ , and $(w_n)_{n \in \mathbb{N}}$ in \mathbb{R}^d converging to 0, \bar{y} , and w , respectively, such that $y_n + t_n w_n \in \Xi$ whenever $n \in \mathbb{N}$. Then we have that

$$Ey_n + t_n Ew_n = E(y_n + t_n w_n) \in \Gamma \text{ for each } n \in \mathbb{N}.$$

Hence $Ew \in \tilde{T}(E\bar{y}; \Gamma)$.

On the other hand, let $w \in \mathbb{R}^d$ be such that $Ew \in \tilde{T}(E\bar{y}; \Gamma)$. By definition, pick $(t_n)_{n \in \mathbb{N}}$ in $(0, \infty)$, and $(u_n)_{n \in \mathbb{N}}$ in Γ , and $(v_n)_{n \in \mathbb{N}}$ in \mathbb{R}^k converging to 0, $E\bar{y}$, and Ew , respectively, such that $u_n + t_n v_n \in \Gamma$ whenever $n \in \mathbb{N}$ (see explanations in Footnote 11). $\Gamma \subset \text{rge } E$, and since $\text{rge } E$ is a closed subspace of \mathbb{R}^k ; one infers that $v_n \in \text{rge } E$ for each $n \in \mathbb{N}$. Therefore, by Banach open mapping theorem there are sequences $(y_n)_{n \in \mathbb{N}}$ converging to

\bar{y} and $(w_n)_{n \in \mathbb{N}}$ converging to w , both in \mathbb{R}^d , such that $Ey_n = u_n$ and $Ew_n = v_n$ for each $n \in \mathbb{N}$.

Thus, for an arbitrary index n , we have

$$\begin{aligned} u_n &= Ey_n \in \Gamma & \text{so } y_n &\in E^{-1}(\Gamma) = \Xi \\ u_n + t_nv_n &= E(y_n + tnw_n) \in \Gamma & \text{so } y_n + tnw_n &\in \Xi \end{aligned}$$

Thus, $w \in \tilde{T}(\bar{y}; \Xi)$, and the claim is proved.

Second, we show that $\tilde{T}(\bar{x}; \Lambda) = \{Gw : w \in \tilde{T}(\bar{y}; \Xi)\}$.

To prove that $G(\tilde{T}(\bar{y}; \Xi)) \subset \tilde{T}(\bar{x}; \Lambda)$, pick any $w \in G(\tilde{T}(\bar{y}; \Xi))$, and find $v \in G(\tilde{T}(\bar{y}; \Xi))$ with $Gv = w$. Thus there is $(t_n)_{n \in \mathbb{N}}$ in $(0, \infty)$ converging to 0, and $(y_n)_{n \in \mathbb{N}}$ in Ξ converging to \bar{y} , and $(v_n)_{n \in \mathbb{N}}$ in \mathbb{R}^d converging to v , such that $y_n + t_nv_n \in \Xi$ whenever $n \in \mathbb{N}$. For each $n \in \mathbb{N}$, put $u_n = Gy_n$ and $w_n = Gv_n$. Clearly, $(u_n)_{n \in \mathbb{N}}$ converges to $G\bar{y} = \bar{x}$ and $(w_n)_{n \in \mathbb{N}}$ converges to $w = Gv$. Moreover,

$$u_n + t_nw_n = G(y_n + t_nv_n) \in G(\Xi) = \Lambda \text{ whenever } n \in \mathbb{N}.$$

Thus, $w \in \tilde{T}(\bar{x}; \Lambda)$.

To see the opposite inclusion, pick any $w \in \tilde{T}(\bar{x}; \Lambda)$. Find $(t_n)_{n \in \mathbb{N}}$ in $(0, \infty)$ converging to 0, $(x_n)_{n \in \mathbb{N}}$ in Λ converging to \bar{x} , and $(w_n)_{n \in \mathbb{N}}$ in \mathbb{R}^l converging to w , such that $x_n + t_nw_n \in \Lambda = G(\Xi)$ for each $n \in \mathbb{N}$.

For each $n \in \mathbb{N}$, find $y_n \in \Xi$ such that $x_n = Gy_n$. Then, $(y_n)_{n \in \mathbb{N}}$ is bounded.

Indeed, if this is not the case, find a cluster point \bar{h} of $(y_n/\|y_n\|)_{n \in \mathbb{N}}$. Let N be an infinite subset of \mathbb{N} such that $\lim_{N \ni n \rightarrow \infty} \frac{y_n}{\|y_n\|} = \bar{h}$. Then

$$0 = \lim_{N \ni n \rightarrow \infty} \frac{x_n}{\|y_n\|} = \lim_{N \ni n \rightarrow \infty} G\left(\frac{y_n}{\|y_n\|}\right) = G(\bar{h}).$$

This contradicts the injectivity of G because $\|\bar{h}\| = 1$. Therefore there is an infinite subset N of \mathbb{N} such that $(y_n)_{n \in N}$ converges to $\tilde{y} \in \mathbb{R}^d$, say. Then

$$G\tilde{y} = \bar{x} = \lim_{N \ni n \rightarrow \infty} x_n = \lim_{N \ni n \rightarrow \infty} Gy_n = G\tilde{y}.$$

Employing, the injectivity once more, we get $\bar{y} = \tilde{y}$.

For each $n \in N$, find v_n in Ξ such that $w_n = G\left(\frac{v_n - y_n}{t_n}\right)$, and put $u_n = \frac{v_n - y_n}{t_n}$. Similar argument as in the case of (y_n) shows that $(u_n)_{n \in N}$ is bounded. Therefore there is an infinite subset N' , of N such that $(u_n)_{n \in N'}$ converges to some $u \in \mathbb{R}^d$. For each $n \in N'$, we have $y_n + t_nu_n = v_n \in \Xi$, therefore $u \in \tilde{T}(\bar{y}; \Xi)$; Moreover, $w = G(u)$ and thus, $w \in G(\tilde{T}(\bar{y}; \Xi))$. The proof of second part is finished.

Combining the two parts yields the assertion. \square

Proposition 3.5.2 (Strict Graphical Derivative of Sum for GE). [20, Proposition 1, p. 5] Under the assumptions (A1) - (A2), for any $b \in \mathbb{R}^n$ one has

$$\tilde{D}\Phi(\bar{z}|\bar{p})(b) = \nabla f(\bar{z})b + B\tilde{D}F_C(C\bar{z}|\bar{v})(Cb).$$

Proof. From Proposition 3.1.24, one obtains

$$\tilde{D}\Phi(\bar{z}|\bar{p})(b) = \nabla f(\bar{z})b + B\tilde{D}Q(\bar{z}|\bar{p} - f(\bar{z}))(b) \quad \text{for each } b \in \mathbb{R}^n.$$

Moreover, observe that

$$\text{gph } Q = \left\{ \begin{array}{l} \begin{pmatrix} u \\ v \end{pmatrix} \in \mathbb{R}^{2n} : \text{there exists } \begin{pmatrix} b \\ c \end{pmatrix} \in \mathbb{R}^m \times \mathbb{R}^n \text{ such that} \\ \begin{pmatrix} u \\ v \end{pmatrix} = G \begin{pmatrix} b \\ c \end{pmatrix} \text{ and } E \begin{pmatrix} b \\ c \end{pmatrix} \in \text{gph } F_C \end{array} \right\},$$

with

$$G := \begin{pmatrix} I_n & 0 \\ 0 & B \end{pmatrix}, \quad E := \begin{pmatrix} C & 0 \\ 0 & I_m \end{pmatrix}.$$

As B is injective, so is G . Using Lemma 3.5.1 with $k := 2m, l := 2n, d := n + m, \Gamma := \text{gph } F_C, \bar{x} := (\bar{z}, \bar{p} - f(\bar{z}))^T$, and $\bar{y} := (\bar{z}, \bar{v})^T$ reveals that

$$\tilde{T}\left((\bar{z}, \bar{p} - f(\bar{z})); \text{gph } Q\right) = \left\{ \begin{pmatrix} b \\ Bc \end{pmatrix} : \begin{pmatrix} Cb \\ c \end{pmatrix} \in \tilde{T}((C\bar{z}, \bar{v}); \text{gph } F_C) \right\}.$$

This means that $\tilde{D}Q(\bar{z}, \bar{p} - f(\bar{z}))(b) = B\tilde{D}F_C(C\bar{z}, \bar{v})(Cb)$. The assertion is proved. \square

The following proposition is the strict graphical version of Theorem 3.1.38, and relates the strong metric regularity of a set-valued map at a reference point to a property of its strict graphical derivative. Remember that the necessary condition was already obtained in Theorem 3.1.42.

Proposition 3.5.3 (Strong Metric Regularity in terms of Strict Graphical Derivative). [20, Proposition 2, p.5]

Consider a set-valued mapping $H : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ and a point $(\bar{x}, \bar{y}) \in \text{gph } H$. Then H is strongly metrically regular at \bar{x} for \bar{y} if and only if it satisfies the following three conditions:

- (a) for each neighborhood U of \bar{x} there is a neighborhood V of \bar{y} such that $H^{-1}(y) \cap U \neq \emptyset$ whenever $y \in V$;
- (b) the set $\text{gph } H \cap (\mathbb{B}_r(\bar{x}) \times \mathbb{B}_r(\bar{y}))$ is closed for some $r > 0$;
- (c) $0 \in \tilde{D}H(\bar{x}|\bar{y})(u) \implies u = 0$.

Proof. Suppose that H is strongly metrically regular at \bar{x} for \bar{y} . Then (c) holds by Theorem 3.1.42. Observe also that H has necessarily locally closed graph at the reference point. Finally, (a) is satisfied since H is open at (\bar{x}, \bar{y}) , that is, for any neighborhood U of \bar{x} , the set $V := H(U)$ is a neighborhood of \bar{y} .

The converse implication is proved in Theorem 3.1.42. \square

Theorem 3.5.4 (Strong Metric Regularity Criterion for GE).

[20, Theorem 1, p. 6] Assume that (A1) - (A3) hold true. Then Φ is strongly metrically regular at \bar{z} for \bar{p} if and only if

(a) for each neighborhood U of \bar{z} there is a neighborhood V of \bar{p} such that

$$\Phi^{-1}(p) \cap U \neq \emptyset \text{ whenever } p \in V;$$

(b) $0 \in \nabla f(\bar{z})b + B\tilde{D}F_C(C\bar{z}|\bar{v})(Cb) \implies b = 0$.

Moreover, its regularity modulus is given by

$$\text{reg}(\Phi; \bar{z}|\bar{p}) = \sup \left\{ \|b\| : \left(\nabla f(\bar{z})b + B\tilde{D}F_C(C\bar{z}|\bar{v})(Cb) \right) \cap \mathbb{B} \neq \emptyset \right\}.$$

Proof. Having in mind that Φ has closed graph, one only needs to combine Proposition 3.5.2 and Proposition 3.5.3 to get the desired equivalence. The formula for the regularity modulus follows from Theorem 3.1.42. \square

Example 3.5.5. (A Simple Circuit with DIAC)

Consider the circuit in Figure 3.4 with a DIAC. In previous sections we have studied the Aubin property, isolated calmness, and calmness of the solution mapping at different reference points. Let us now, investigate the strong metric regularity of the map Φ in view of Theorem 3.5.4. Note that $\Phi : \mathbb{R} \rightrightarrows \mathbb{R}$ is given by $\Phi(z) = f(z) + F(z)$, where

$$f(z) := \begin{cases} Rz - \frac{V}{\sqrt{1 - \frac{2az}{V}}} + V & z < 0, \\ Rz + \frac{V}{\sqrt{1 + \frac{2az}{V}}} - V & z \geq 0, \end{cases} \quad \text{and} \quad F(z) := \begin{cases} -V & z < 0, \\ [-V, V] & z = 0, \\ V & z > 0, \end{cases}$$

in which R , a , and V are positive constants (ref. Example 2.3.2). Assumptions (A1) - (A3) hold true. We will consider two reference points.

For $(\bar{z}, \bar{p}) = (0, 0)$, we would have $f'(0) = R - a$, $\bar{v} = 0$. A close look at graph of Φ (see Figure 2.31) reveals that condition (a) is satisfied. Condition (b) reads as

$$0 \in (R - a)\beta + \tilde{D}F(0|0)(\beta) \implies \beta = 0.$$

Since $\tilde{T}\left((0, 0); \text{gph } F\right) = \mathbb{R} \begin{pmatrix} 0 \\ 1 \end{pmatrix}$, also this condition is satisfied, and Φ is strongly metrically regular at 0 for 0.

For $(\bar{z}, \bar{p}) = (0, -V)$, parameters R , and a play an important role, as we might have expected. If $a > R$, the point $(0, -V)$ is a local minimum of $\text{gph } \Phi$, and hence condition (a) is not satisfied. When $a \leq R$, there would be no more problems for this condition (see Figure 2.31). To check condition (b), notice that

$$\begin{aligned} \tilde{D}F(0 | -V)(\beta) &= \left\{ v \in \mathbb{R} \mid \begin{pmatrix} \beta \\ v \end{pmatrix} \in \tilde{T}\left((0, -V); \text{gph } F\right) \right\} \\ &= \left\{ v \mid \begin{pmatrix} \beta \\ v \end{pmatrix} \in \mathbb{R} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} -1 \\ 0 \end{pmatrix} \right\}. \end{aligned}$$

Hence, $(\beta, (a - R)\beta)^T$ should belong to the latter union. If $a < R$, the only possibility would be $\beta = 0$, and thus, the condition holds true. But in case of $a = R$, it is possible to have $\beta \neq 0$.

Therefore, Φ will not have the strong metric regularity property at 0 for $-V$ if $a \geq R$.

Theorem 3.5.6 (Summary of Metric Regularity Criteria for GE).

[20, Theorem 2, p. 6] Suppose that the assumptions (A1) - (A4) hold true. Then

(i) Φ is metrically regular at \bar{z} for \bar{p} if and only if

$$\left. \begin{aligned} ((CC^T)^{-1}C\nabla f(\bar{z})^T\xi, B^T\xi) &\in -N((C\bar{z}, \bar{v}); \text{gph } F) \\ \nabla f(\bar{z})^T\xi &\in \text{rge } C^T \end{aligned} \right\} \Rightarrow \xi = 0;$$

(ii) Φ is strongly metrically sub-regular at \bar{z} for \bar{p} if and only if

$$\left. \begin{aligned} (Cb, -(B^TB)^{-1}B^T\nabla f(\bar{z})b) &\in T((C\bar{z}, \bar{v}); \text{gph } F) \\ \nabla f(\bar{z})b &\in \text{rge } B \end{aligned} \right\} \Rightarrow b = 0;$$

(iii) Φ is strongly metrically regular at \bar{z} for \bar{p} if and only if

(a) for each neighborhood U of \bar{z} there is a neighborhood V of \bar{p} such that

$$\Phi^{-1}(p) \cap U \neq \emptyset \text{ whenever } p \in V;$$

(b)

$$\left. \begin{aligned} (Cb, -(B^TB)^{-1}B^T\nabla f(\bar{z})b) &\in \tilde{T}((C\bar{z}, \bar{v}); \text{gph } F) \\ \nabla f(\bar{z})b &\in \text{rge } B \end{aligned} \right\} \Rightarrow b = 0.$$

Proof. The statement (i) is Corollary 3.2.5, the proof of which was based on Mordukhovich coderivative criterion. Whereas (ii) is Corollary 3.3.4. To see the last one, note that if C is surjective, then $F_C = F$. Moreover, (A1) ensures that $B^TB \in \mathbb{R}^{m \times m}$ is non-singular. It suffices to show that (b) is equivalent to (b) in Theorem 3.5.4.

First, let $b \in \mathbb{R}^n$ be such that $0 \in \nabla f(\bar{z})b + B\tilde{D}F(C\bar{z}|\bar{v})(Cb)$. Find a point $w \in \tilde{D}F(C\bar{z}|\bar{v})(Cb)$ with $\nabla f(\bar{z})b + Bw = 0$. Thus, $-(B^TB)^{-1}B^T\nabla f(\bar{z})b$ is in $\tilde{D}F(C\bar{z}|\bar{v})(Cb)$. Clearly, we have $\nabla f(\bar{z})b \in \text{rge } B$ and the definition of the strict graphical derivative of F yields the rest.

On the other hand, pick any $b \in \mathbb{R}^n$ with $(Cb, -(B^T B)^{-1} B^T \nabla f(\bar{z})b)$ in $\tilde{T}((C\bar{z}, \bar{v}); \text{gph } F)$ and $\nabla f(\bar{z})b \in \text{rge } B$. The definition of the strict graphical derivative says that

$$w := -(B^T B)^{-1} B^T \nabla f(\bar{z})b \in \tilde{D}F(C\bar{z} | \bar{v})(Cb).$$

Thus we have $B^T Bw = -B^T \nabla f(\bar{z})b$. So, $Bw + \nabla f(\bar{z})b \in \ker B^T \cap \text{rge } B = \{0\}$. Therefore, $0 \in \nabla f(\bar{z})b + B \tilde{D}F(C\bar{z} | \bar{v})(Cb)$. \square

Corollary 3.5.7. [20, Corollary 1, p. 7]

In addition to (A1) - (A5), assume that $n = m$, $B = C = I_n$, $\nabla f(\bar{z})$ is a P-matrix, and for each $i \in \{1, 2, \dots, n\}$, the mapping $F_i : \mathbb{R} \rightrightarrows \mathbb{R}$ is maximal monotone. Then Φ is strongly metrically regular at \bar{z} for \bar{v} .

Proof. For any $x = (x_1, \dots, x_n), y = (y_1, \dots, y_n) \in \mathbb{R}^n$, define the function $\varphi : \mathbb{R}^n \times \mathbb{R}^n \rightarrow (\mathbb{R}^2)^n$ as $\varphi(x, y) := ((x_1, y_1), \dots, (x_n, y_n))$.

Clearly, φ is linear and one-to-one. Also, by (A5), we have $\prod_{i=1}^n \text{gph } F_i = \varphi(\text{gph } F)$. The definition of the paratingent cone, Proposition 3.1.10, and Lemma 3.5.1 (with $E = I$, and G be the representative matrix of φ) imply that

$$\begin{aligned} \tilde{T}(\tilde{T}((\bar{z}, \bar{v}); \text{gph } F)) &= \tilde{T}(\varphi(\bar{z}, \bar{v}); \varphi(\text{gph } F)) \\ &= \tilde{T}(\varphi(\bar{z}, \bar{v}); \prod_{i=1}^n \text{gph } F_i) \subset \prod_{i=1}^n \tilde{T}((\bar{z}_i, \bar{v}_i); \text{gph } F_i). \end{aligned}$$

Also, it is well-known that

$$\prod_{i=1}^n N((\bar{z}_i, \bar{v}_i); \text{gph } F_i) = N(\varphi(\bar{z}, \bar{v}); \prod_{i=1}^n \text{gph } F_i) = \varphi(N((\bar{z}, \bar{v}); \text{gph } F)).$$

As all F_i 's are maximal monotone, we have

$$\begin{aligned} N((\bar{z}_i, \bar{v}_i); \text{gph } F_i) &\subset \{(a, b) \in \mathbb{R}^2 : ab \leq 0\}, \\ \tilde{T}((\bar{z}_i, \bar{v}_i); \text{gph } F_i) &\subset \{(a, b) \in \mathbb{R}^2 : ab \geq 0\}, \text{ for each } i \in \{1, \dots, n\}. \end{aligned}$$

Fix any non-zero $\eta \in \mathbb{R}^n$. Since $\nabla f(\bar{z})$ is a P-matrix, so is $\nabla f(\bar{z})^T$. There are $k, l \in \{1, \dots, n\}$ such that $\eta_k(\nabla f(\bar{z})\eta)_k > 0$ and $\eta_l(\nabla f(\bar{z})^T \eta)_l > 0$, which means that $(\eta_k, -(\nabla f(\bar{z})\eta)_k) \notin \tilde{T}((\bar{z}_k, \bar{v}_k); \text{gph } F_k)$ and $((\nabla f(\bar{z})^T \eta)_l, \eta_l) \notin -N((\bar{z}_l, \bar{v}_l); \text{gph } F_l)$. The above relations for the normal and paratingent cone and the fact that φ is one-to-one, imply that conditions (i) and (iii - b) in Theorem 3.5.6 hold. Then, (iii - a) is a result of (i)¹³, and the proof is complete. \square

¹³ This fact is indeed, the very first lines of the proof of Theorem 1.3.5, (equivalence of metric regularity and inverse Aubin property). Here is a brief review.

Example 3.5.8. Let us have a closer look at Example 3.5.5. We have $n = m = 1$, $B = C = I_1$, and not only assumptions (A1)-(A3), but also assumptions (A4) and (A5) hold true. It is also easy to verify that F is maximal monotone. Hence, in view of Corollary 3.5.7, the strong metric regularity of Φ could be concluded if $f'(\bar{z})$ is a P-matrix.

From Definition 3.2.7, and calculations in Example 2.3.2, we obtain that everything reduces to check whether

$$f'(\bar{z}) = R - \frac{a}{\left(\frac{2a|\bar{z}|}{V} + 1\right)^{3/2}}$$

is positive or not. For $\bar{z} = 0$, $f'(0) = R - a > 0$ whenever $a < R$, which confirms the results of Example 3.5.5.

Note that the corollary is a sufficient condition and thus, does not say anything about the strong metric regularity when $f'(\bar{z}) \leq 0$, as we observed that for example, at the reference point $(0,0)$, the strong metric regularity always holds no matter what the relation of R and a is.

3.5.1 Strong Metric (Sub-) Regularity with a Nonsmooth Single-valued Part

The rest of this section discusses the case of a non-smooth f in the setting (3.46), that is when f is not continuously differentiable on \mathbb{R}^n as it was assumed by condition (A2) in Note 3.2.2. Since there is no differentiability of f any more to use, one can think of generalized Jacobians as a possible alternative.

We shall recall that from the circuit interpretation point of view, in general, F represents the $i - v$ characteristic of diodes, transistors and such components and f indicates all the other components and how they are connected to each other. However, as we have explained before in Chapter 2, one can use a simplification technique in order to simplify the graph of set-valued part (which has a computational importance) and thus, add some points to f . Hence, it is very probable to obtain a single-valued part which is not smooth enough (see Example 3.5.14).

We will discuss the results obtained by Ismailov [46] for strong metric regularity and the similar results considering his method for strong metric sub-regularity [19].

Let $U = \mathbb{B}_a(\bar{z})$, $V = \mathbb{B}_b(\bar{p})$, $\kappa > \text{reg}(\Phi; \bar{z} | \bar{p})$, where a and b are positive constants such that $b < \frac{a}{\kappa}$. Choose an arbitrary $p \in \mathbb{B}_b(\bar{p})$. If $\Phi^{-1}(p) \cap \mathbb{B}_a(\bar{z}) = \emptyset$, then $d(\bar{z}, \Phi^{-1}(p)) \geq a$. Thus, by definition of metric regularity one obtains

$$a \leq d(\bar{z}, \Phi^{-1}(p)) \leq \kappa d(p, \Phi(\bar{z})) \leq \kappa \|p - \bar{p}\| \leq \kappa b < a,$$

which is a contradiction. So, (iii - a) holds.

It is worth mentioning that as Ismailov claimed, the setting

$$f(x) + F(x) \ni 0, \quad (3.66)$$

with a non-differentiable function $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ and a set-valued map $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$, could be considered as a way of unifying two classical theorems of variational analysis. The first theorem is due to Robinson [60] and is based on the assumptions we were working with till now (see also Subsection 4.1.1). Thus, Theorem 3.5.12 extends Theorem 3.5.6 to the case of a non-smooth function f .

The second result is Clarke's inverse function theorem [21], which is concerned with the case of a usual nonlinear equation

$$f(x) = 0,$$

corresponds to (3.66) with $F(\cdot) \equiv \{0\}$; but assumes local Lipschitz continuity of f rather than smoothness. Thus, Theorem 3.5.12 at the same time, extends Clarke's theorem (see also [22, Theorem 7.1.1, p. 253]) from usual equations to GEs.

Thus, we assume the following condition instead of (A2):

($\widetilde{A2}$) f is locally Lipschitz continuous on \mathbb{R}^n .

Let us first provide a perturbation results like Theorem 1.4.14, with a strongly metrically sub-regular set-valued map and a calm single-valued term.

Lemma 3.5.9 (Stability of Strong Metric Sub-regularity Under Single-valued Calm Perturbations). [20, Lemma 2, p. 8]

Let $(\bar{x}, \bar{y}) \in \mathbb{R}^n \times \mathbb{R}^m$ and $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ be such that $\bar{y} \in G(\bar{x})$. Suppose that G is strongly metrically sub-regular at \bar{x} for \bar{y} (that is, there is $\kappa > 0$ along with a neighborhood U of \bar{x} such that $\|x - \bar{x}\| \leq \kappa d(\bar{y}, G(x))$ whenever $x \in U$).

Then, for any function $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ which is calm at \bar{x} relative to $U \subset \text{dom } g$ with the constant $\mu < \frac{1}{\kappa}$, one has

$$\|x - \bar{x}\| \leq \frac{\kappa}{1 - \kappa\mu} d(\bar{y} + g(\bar{x}), g(x) + G(x)) \quad \text{for each } x \in U.$$

That is, $g + G$ is strongly metrically sub-regular at \bar{x} for $\bar{y} + g(\bar{x})$.

Proof. Fix any $x \in U$. The calmness of g means that $\|g(x) - g(\bar{x})\| \leq \mu \|x - \bar{x}\|$. Therefore,

$$\begin{aligned} \|x - \bar{x}\| &\leq \kappa d(\bar{y}, G(x)) \leq \kappa \left[d(\bar{y}, \bar{y} - g(x) + g(\bar{x})) + d(\bar{y} - g(x) + g(\bar{x}), G(x)) \right] \\ &\leq \kappa \|g(x) - g(\bar{x})\| + \kappa d(\bar{y} + g(\bar{x}), g(x) + G(x)) \\ &\leq \kappa\mu \|x - \bar{x}\| + \kappa d(\bar{y} + g(\bar{x}), g(x) + G(x)). \end{aligned}$$

Performing a small rearrangement and dividing by $1 - \kappa\mu > 0$, we obtain the desired inequality. \square

Now, we go one step ahead with dropping the calmness condition on g .

Theorem 3.5.10 (Stability of SMSR Under Single-valued Perturbations).

[19, Theorem 3.5, page 11] Let $(\bar{x}, \bar{y}) \in \mathbb{R}^n \times \mathbb{R}^l$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^l$, and $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^l$ be such that $\bar{y} \in g(\bar{x}) + G(\bar{x})$. Suppose that there exists a compact subset \mathcal{A} of $\mathbb{R}^{l \times n}$ such that

- (a) there are $\varepsilon > 0$ and $r > 0$ such that for each $u \in \text{int } \mathbb{B}_r(\bar{x})$, one can find $A \in \mathcal{A}$ such that $\|g(u) - g(\bar{x}) - A(u - \bar{x})\| \leq \varepsilon \|u - \bar{x}\|$;
- (b) for every $A \in \mathcal{A}$ the mapping $G_A : \mathbb{R}^n \ni x \mapsto g(\bar{x}) + A(x - \bar{x}) + G(x) \subset \mathbb{R}^l$ is strongly metrically sub-regular at \bar{x} for \bar{y} and let $m := \sup_{A \in \mathcal{A}} \text{subreg}(G_A; \bar{x} | \bar{y}) < \frac{1}{2\varepsilon}$.

Then $g + G$ is strongly metrically sub-regular at \bar{x} for \bar{y} ; and

$$\text{subreg}(g + G; \bar{x} | \bar{y}) \leq \frac{m}{1 - 2m\varepsilon}.$$

Proof. Without loss of generality assume that $\bar{y} = 0$ and note that $m < \infty$. Fix any $\kappa > m$ such that $2\kappa\varepsilon < 1$. Let $r > 0$ be as in (a). First, we show that there exists $a \in (0, r]$ such that

$$\|x - \bar{x}\| \leq \frac{\kappa}{1 - \kappa\varepsilon} d(0, G_A(x)) \quad \text{whenever } x \in \text{int } \mathbb{B}_a(\bar{x}), \text{ and } A \in \mathcal{A}. \quad (3.67)$$

As \mathcal{A} is compact, there is a finite set $\mathcal{A}_F \subset \mathcal{A}$ such that

$$\mathcal{A} \subset \mathcal{A}_F + \varepsilon\mathbb{B}. \quad (3.68)$$

Pick any $\tilde{A} \in \mathcal{A}_F$. Then, by assumption (b), there exists $\alpha_{\tilde{A}} > 0$ such that

$$\|x - \bar{x}\| \leq \kappa d(0, G_{\tilde{A}}(x)) \quad \text{whenever } x \in \text{int } \mathbb{B}_{\alpha_{\tilde{A}}}(\bar{x}).$$

Fix any $A' \in \varepsilon\mathbb{B}$. As $G_{\tilde{A}+A'} = G_{\tilde{A}} + A'(x - \bar{x})$, Lemma 3.5.9 with $g(\cdot) = A'(\cdot - \bar{x})$ reveals that

$$\|x - \bar{x}\| \leq \frac{\kappa}{1 - \kappa\varepsilon} d(0, G_{\tilde{A}+A'}(x)) \quad \text{for any } x \in \text{int } \mathbb{B}_{\alpha_{\tilde{A}}}(\bar{x}).$$

Thus for any $\tilde{A} \in \mathcal{A}_F$, there is $\alpha_{\tilde{A}} > 0$ such that for each $A' \in \varepsilon\mathbb{B}$ the above inequality holds. Let $a = \min \left\{ r, \min_{\tilde{A} \in \mathcal{A}_F} \alpha_{\tilde{A}} \right\}$. Taking into account (3.68), we obtain (3.67).

Fix any $x \in \text{int } \mathbb{B}_a(\bar{x})$. Use (a) to find $A \in \mathcal{A}$ such that $\|g(x) - g(\bar{x}) - A(x - \bar{x})\| \leq$

$\varepsilon \|x - \bar{x}\|$. This and (3.67) imply that

$$\begin{aligned} \|x - \bar{x}\| &\leq \frac{\kappa}{1 - \kappa\varepsilon} d(0, G_A(x)) = \frac{\kappa}{1 - \kappa\varepsilon} d(-g(\bar{x}) - A(x - \bar{x}), G(x)) \\ &\leq \frac{\kappa}{1 - \kappa\varepsilon} \left[d(-g(x), G(x)) + \|g(x) - g(\bar{x}) - A(x - \bar{x})\| \right] \\ &\leq \frac{\kappa}{1 - \kappa\varepsilon} d(0, g(x) + G(x)) + \frac{\kappa}{1 - \kappa\varepsilon} \varepsilon \|x - \bar{x}\| \end{aligned}$$

Since $2\kappa\varepsilon < 1$, and thus, $\frac{\kappa\varepsilon}{1 - \kappa\varepsilon} < 1$, we get that

$$\|x - \bar{x}\| \leq \frac{\kappa}{1 - 2\kappa\varepsilon} d(0, g(x) + G(x)).$$

Thus, $g + G$ is strongly metrically sub-regular at \bar{x} for 0. As $\kappa \in (m, \frac{1}{2\varepsilon})$ was arbitrary, we get the desired estimate on the sub-regularity modulus. \square

Unfortunately, Theorem 3.5.10 does not say anything about the possible choices of subsets \mathcal{A} . The following corollary suggests a family of subsets \mathcal{A} , for which the condition (a) in Theorem 3.5.10 is satisfied.

Corollary 3.5.11. [20, Corollary 2, p. 10] *Let $(\bar{x}, \bar{y}) \in \mathbb{R}^n \times \mathbb{R}^l$, $g : \mathbb{R}^n \rightarrow \mathbb{R}^l$, and $G : \mathbb{R}^n \rightrightarrows \mathbb{R}^l$ be such that $\bar{y} \in g(\bar{x}) + G(\bar{x})$. Suppose that g is locally Lipschitz continuous at \bar{x} and that for every $\varepsilon > 0$ there exists $r > 0$ along with a selection h for $\partial_B g$ such that*

$$\|g(u) - g(\bar{x}) - h(u)(u - \bar{x})\| \leq \varepsilon \|u - \bar{x}\| \quad \text{whenever } u \in \text{int } \mathbb{B}_r(\bar{x}). \quad (3.69)$$

Assume that the assumption (b) in Theorem 3.5.10 is satisfied with $\mathcal{A} := \partial_B g(\bar{x})$. Then $g + G$ is strongly metrically sub-regular at \bar{x} for \bar{y} ; and $\text{subreg}(g + G; \bar{x} | \bar{y}) \leq m$.

Proof. Let $\gamma \in (0, 1)$ be such that $\gamma(m + \gamma) < 1$. Set $\mathcal{A} = \partial_B g(\bar{x}) + \gamma\mathbb{B}$. Then \mathcal{A} is compact (as the sum of two compact sets). By Lemma 3.5.9, for any $\tilde{A} \in \partial_B g(\bar{x})$ and any $A' \in \gamma\mathbb{B}$, the mapping $G_{\tilde{A}+A'} = G_{\tilde{A}} + A'(x - \bar{x})$ is strongly metrically sub-regular at \bar{x} for \bar{y} with the modulus at most $\frac{m+\gamma}{1-(m+\gamma)\gamma}$. Thus, for every $A \in \mathcal{A}$ the mapping G_A is strongly metrically sub-regular at \bar{x} for \bar{y} ; and

$$m' = \sup_{A \in \mathcal{A}} \text{subreg}(G_A; \bar{x} | \bar{y}) \leq \frac{m + \gamma}{1 - (m + \gamma)\gamma}$$

Let $\varepsilon \in (0, \gamma)$ be arbitrary. By the outer semi-continuity of $\partial_B g$ and (3.69), there is $r > 0$ and a selection h for $\partial_B g$ such that, for each $u \in \text{int } \mathbb{B}_r(\bar{x})$, one has

$$h(u) \in \partial_B g(u) \subset \partial_B g(\bar{x}) + \varepsilon\mathbb{B} \subset \mathcal{A} \quad \text{and} \quad \|g(u) - g(\bar{x}) - h(u)(u - \bar{x})\| \leq \varepsilon \|u - \bar{x}\|.$$

Thus, for each $u \in \text{int } \mathbb{B}_r(\bar{x})$ one can find $A \in \mathcal{A}$ such that $\|g(u) - g(\bar{x}) - A(u - \bar{x})\| \leq \varepsilon \|u - \bar{x}\|$. Theorem 3.5.10 implies that the mapping $g + G$ is strongly metrically sub-regular at \bar{x} for \bar{y} ; and $\text{subreg}(g + G; \bar{x} | \bar{y}) \leq m'$. As $\gamma > 0$ can be arbitrarily small, the proof is finished. \square

Now, we will express the Izmailov's theorem [46] about the strong metric regularity in case of a non-smooth single-valued part, and then summarize the results of this subsection in one theorem.

Theorem 3.5.12 (SMR of Perturbed GE with a Nonsmooth Single-valued Part). [46, Theorem 3, p. 583] *Let $f : \mathbb{R}^n \rightarrow \mathbb{R}^n$ be a Lipschitz continuous function in a neighborhood of $\bar{x} \in \mathbb{R}^n$, and $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ be a set-valued map such that $0 \in f(\bar{x}) + F(\bar{x})$, and for each $A \in \partial f(\bar{z})$ the mapping*

$$\mathbb{R}^n \ni x \mapsto J_A(x) := f(\bar{x}) + A(x - \bar{x}) + F(x)$$

is strongly metrically regular at \bar{x} for 0.

Then there exist neighborhoods U of \bar{x} and V of 0 such that for every $y \in V$, there exists a unique $x(y) \in U$ satisfying the perturbed generalized equation $y \in f(x) + F(x)$, and the mapping $y \mapsto x(y)$ is Lipschitz continuous on V .

Theorem 3.5.13 (Summary of MR Criteria for GE in Case of a Nonsmooth Single-valued Part). [20, Theorem 4, p. 10]

Under the assumptions (A1), $\widetilde{\text{(A2)}}$, (A3), and (A4), for any $A \in \partial f(\bar{z})$, define the mapping

$$J_A : \mathbb{R}^n \ni z \mapsto f(\bar{z}) + A(z - \bar{z}) + BF(Cz).$$

(a) *The mapping Φ is strongly metrically sub-regular at \bar{z} for \bar{p} provided that for each $A \in \partial f(\bar{z})$, one has that*

$$\left. \begin{array}{l} (Cb, -(B^T B)^{-1} B^T Ab) \in T((C\bar{z}, \bar{v}); \text{gph } F) \\ Ab \in \text{rge } B \end{array} \right\} \Rightarrow b = 0;$$

(b) *The mapping Φ is strongly metrically regular at \bar{z} for \bar{p} provided that for each $A \in \partial f(\bar{z})$, one has that*

(i) *for each neighborhood U of \bar{z} there is a neighborhood V of \bar{p} such that*

$$J_A^{-1}(p) \cap U \neq \emptyset \text{ whenever } p \in V;$$

(ii)

$$\left. \begin{array}{l} (Cb, -(B^T B)^{-1} B^T Ab) \in \widetilde{T}((C\bar{z}, \bar{v}); \text{gph } F) \\ Ab \in \text{rge } B \end{array} \right\} \Rightarrow b = 0.$$

Proof. (a) For each $A \in \partial f(\bar{z})$, the function $f^*(\cdot) := f(\bar{z}) + A(\cdot - \bar{z})$ is continuously differentiable on \mathbb{R}^n ; so, condition (A2) is satisfied and the mapping J_A is strongly metrically sub-regular at \bar{z} for \bar{p} by Theorem 3.5.6 (ii) with $\Phi := J_A = f^* + BF(C \cdot)$. Now, apply Theorem 3.5.10 with $g = f$, $G = F$, and $\mathcal{A} = \partial f(\bar{z})$, to get the conclusion. (b) The conditions (i) and (ii) guarantee that, for each $A \in \partial f(\bar{z})$, the mapping J_A is strongly metrically regular at \bar{z} for \bar{p} (using Theorem 3.5.6 (iii)). By Theorem 3.5.12, Φ is strongly metrically regular at \bar{z} for \bar{p} . \square

Example 3.5.14. (A simple circuit with SCR) [20]

Consider the circuit in Figure 3.9, with an SCR, whose $i - v$ characteristic is given by

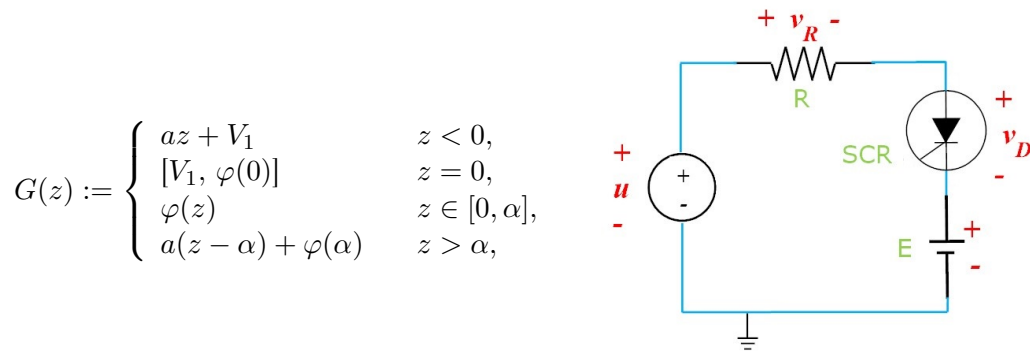


FIGURE 3.9: A simple circuit with SCR

where $-V_1$, α , and a are positive constants, and $\varphi : \mathbb{R} \rightarrow (0, \infty)$ is a continuously differentiable function with $V_1 < \varphi(0)$, $\varphi(\alpha) < \varphi(0)$, $\varphi'(0) > 0$, and $a > \varphi'(\alpha) > 0$. Note that $G = g + F$ with

$$F(z) := \begin{cases} V_1 & z < 0, \\ [V_1, \varphi(0)] & z = 0, \\ \varphi(0) & z > 0, \end{cases} \quad \text{and} \quad g(z) := \begin{cases} az & z < 0, \\ \varphi(z) - \varphi(0) & z \in [0, \alpha], \\ a(z - \alpha) + \varphi(\alpha) - \varphi(0) & z > \alpha. \end{cases}$$

A review of Example 2.3.2 reveals that by setting $p = u - E$ and $z = i$, we get (3.46) with $m = n = 1$, $B = C = I_1$, and $f(z) = Rz + g(z)$, for $z \in \mathbb{R}$.

Then f is locally Lipschitz continuous on \mathbb{R} with

$$\partial f(z) = \begin{cases} R + a & z < 0, \\ [R + a, \varphi'(0)] & z = 0, \\ R + \varphi'(z) & z \in (0, \alpha), \\ [R + \varphi'(\alpha), R + a] & z = \alpha, \\ R + a & z > \alpha. \end{cases}$$

Suppose that $\varphi'(z) > -R$ for each $z \in (0, \alpha)$. Then, for any $z \in \mathbb{R}$, all the elements of $\partial f(z)$ are positive. Assumptions (A1), $(\widetilde{A2})$, (A3), and (A4), hold true. Given $(\bar{z}, \bar{p}) \in \text{gph } \Phi$, we get that the assumptions of Theorem 3.5.13 (b) are satisfied. Thus, Φ is

strongly metrically regular at any reference point.

Let us do the computations for two reference points. For $\bar{z} = 0$, $\bar{p} = V_1$, we have $f(\bar{z}) = 0$, $\bar{v} = V_1$, $J_A(z) = Az + F(z)$ where $A \in [R + a, \varphi'(0)]$, and condition (i) is satisfied (see Figure 3.10). Condition (ii) reads as

$$(b, -Ab) \in \tilde{T}\left((0, V_1); \text{gph } F\right) \implies b = 0, \quad \text{for all } A \in [R + a, \varphi'(0)].$$

Since $\tilde{T}\left((0, V_1); \text{gph } F\right) = \mathbb{R} \begin{pmatrix} 0 \\ 1 \end{pmatrix} \cup \mathbb{R} \begin{pmatrix} -1 \\ 0 \end{pmatrix}$, and $A > 0$, the condition is satisfied.

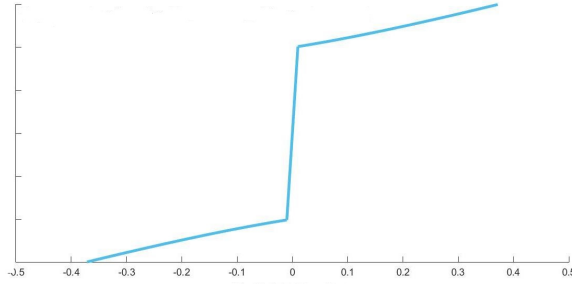


FIGURE 3.10: General form of the map J_A in Example 3.5.14

For $\bar{z} = \alpha$, $\bar{p} = R\alpha + \varphi(\alpha)$, we have $\bar{v} = \varphi(0)$, and $f(\alpha) = R\alpha + \varphi(\alpha) - \varphi(0) > 0$. Then $J_A(z) = f(\alpha) + A(z - \alpha) + F(z)$ where $A \in [R + \varphi'(\alpha), R + a]$, which is again of the form shown in Figure 3.10. Hence, condition (i) is satisfied.

Moreover, $\tilde{T}\left((\alpha, \varphi(0)); \text{gph } F\right) = \mathbb{R} \begin{pmatrix} 1 \\ 0 \end{pmatrix}$, and thus, condition (ii) also holds.

Chapter 4

Metric Regularity in the Case of Time-Varying Sources

Some mathematician, I believe, has said that true pleasure lies not in the discovery of truth, but in the search for it.

Leo Tolstoy

In this chapter, we try to answer the question of perturbation effect when the vector p becomes a function of time, which corresponds to the electronic circuits working with AC sources. We start with obtaining a proper model for the electrical circuit in this case, and explain the shift of our goal from studying the solution mapping to solution trajectories in Section 4.1. In Subsection 4.1.1, we express an important theorem (cf. Theorem 4.1.2) to build our structure.

Section 4.2 starts with a proposition that has a simple proof, but contains a novel idea which considerably eases the study of the problem in the time varying case. In Subsection 4.2.1, we provide results that highlight smoothness dependence of trajectories on the input signal.

In Subsection 4.2.2, we deviate from the study about solution trajectories for a while to present a uniform strong metric regularity result along a trajectory, with assumptions weaker than Theorem 4.1.2. It is necessary for obtaining the results about perturbation effect in Subsection 4.2.3, where we provide a theorem that guarantees the existence of a solution trajectory for the perturbed problem which is continuous, and whose distance from the solution trajectory of the (non-perturbed) generalized equation is controlled by the distance between the input signal and its perturbed function.

4.1 Obtaining a Proper Model

In this section, we would reconsider the situation described in Example 2.2.4, in order to introduce and study in depth the problem in the case of time-varying sources. We would see that there is a need to change the formulation and thus, the point of view towards the perturbation study. A primary result that guarantees the existence of continuous solution trajectories will be given in Subsection 4.1.1, Theorem 4.1.2. This theorem will be a starting point for the detailed study of next sections.

Let us start with an example. In Figure 4.1, a simple regulator circuit with a practical model for the diode is shown. The voltage source is made of n batteries connected to each other in a serial scheme, that provides $n + 1$ different levels between $0 V_s$, and $1 V_s$.

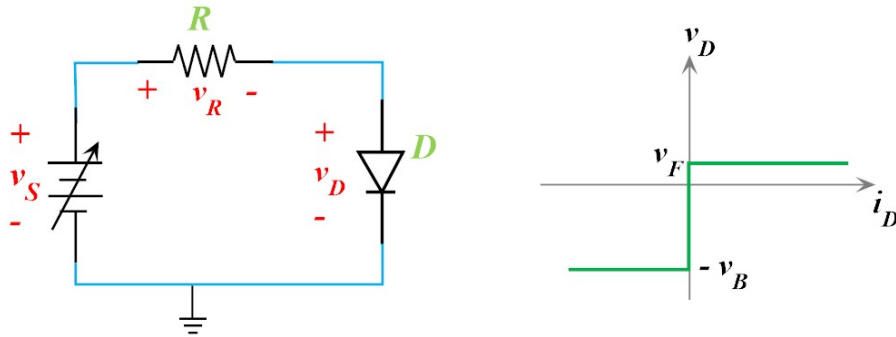


FIGURE 4.1: A regulator circuit with multiple DC sources

Using Kirchhoff's laws and $i - v$ characteristics of diode and resistor, we obtain that:

$$\left. \begin{array}{l} \text{KVL : } -\frac{m}{n}V_S + V_R + V_D = 0 \\ \text{KCL : } I_D = I_R =: I \\ V_R = R I_R \\ V_D \in F(I_D) \end{array} \right\} \Rightarrow 0 \in -p + Rz + F(z), \quad (4.1)$$

where $p = \frac{m}{n}V_S$, $z = I$, and $m = 0, 1, \dots, n$, indicates the number of turned-on batteries in the circuit. Then, the solution mapping would be

$$S(p) = \{ z \in \mathbb{R} \mid -p + Rz + F(z) \ni 0 \}.$$

In order to find $S(p)$, we can consider the three parts of $\text{gph } F$ separately to solve the generalized equation, fortunately, analytically this time.

1. For $z > 0$, with $F(z) = \{v_F\}$.

Then, we would have an equation, $-p + Rz + v_F = 0$. Thus, $z = \frac{p - v_F}{R}$ which is only valid for $z > 0$, that is, when $p > v_F$.

2. For $z = 0$, with $F(z) = [-v_B, v_F]$.

Then, $-p + 0 + [-v_B, v_F] \ni 0$. That is, $z = 0$ for $p \in [-v_B, v_F]$.

3. For $z < 0$, with $F(z) = \{-v_B\}$.

Then, again we would have an equation, $-p + Rz - v_B = 0$. Thus, $z = \frac{p + v_B}{R}$, as long as $p < -v_B$.

Therefore, S is a single-valued map in this problem, with the graph shown in Figure 4.2 (left), and the rule given as:

$$S(p) = \begin{cases} \left\{ \frac{p + v_B}{R} \right\} & p < -v_B \\ \{0\} & p \in [-v_B, v_F] \\ \left\{ \frac{p - v_F}{R} \right\} & p > v_F \end{cases} \quad (4.2)$$

When we deal with an AC voltage source, theoretically we can follow the same procedure. For any $t \in [0, 1]$, use the specific value $p(t)$ and the transformation graph to find the value of z at that time, that is $z(t)$. Then, we can obtain the graph of $z(\cdot)$ with respect to time, similar to the one shown in Figure 4.2 (right) for a sinusoid signal.

There are two interesting facts to highlight here:

- (a) Very naturally, instead of asking for the graph of the solution mapping with respect to the input signal, we focused on the graph of the solution mapping with respect to the time. Of course, when the solution mapping is not a function like this problem, the latter expression needs a clarification.
- (b) Dealing with a function as the input signal, we searched for a function as the output signal. In order to keep the notations consistent, yet without ambiguity, we will refer to these functions as $p(\cdot)$, $p: \mathbb{R} \rightarrow \mathbb{R}^n$, $z(\cdot)$, and so on.

Therefore, in the case of time varying sources, we can assume that t as a parameter, belongs to a set like $[0, 1]$, and it would be more appropriate to consider the solution mapping as the (generally set-valued) map that associates to every $t \in [0, 1]$, the set of all possible vectors z in \mathbb{R}^n that fits the generalized equation

$$f(z) - p(t) + F(z) \ni 0. \quad (4.3)$$

The *solution mapping* is therefore given by

$$S: t \mapsto S(t) = \{z \in \mathbb{R}^n \mid f(z) - p(t) + F(z) \ni 0\}, \quad (4.4)$$

and a *solution trajectory* over $[0, 1]$ is, in this case, a function $\bar{z}(\cdot)$ such that $\bar{z}(t) \in S(t)$ for all $t \in [0, 1]$, that is, $\bar{z}(\cdot)$ is a selection for S over $[0, 1]$. Every solution trajectory

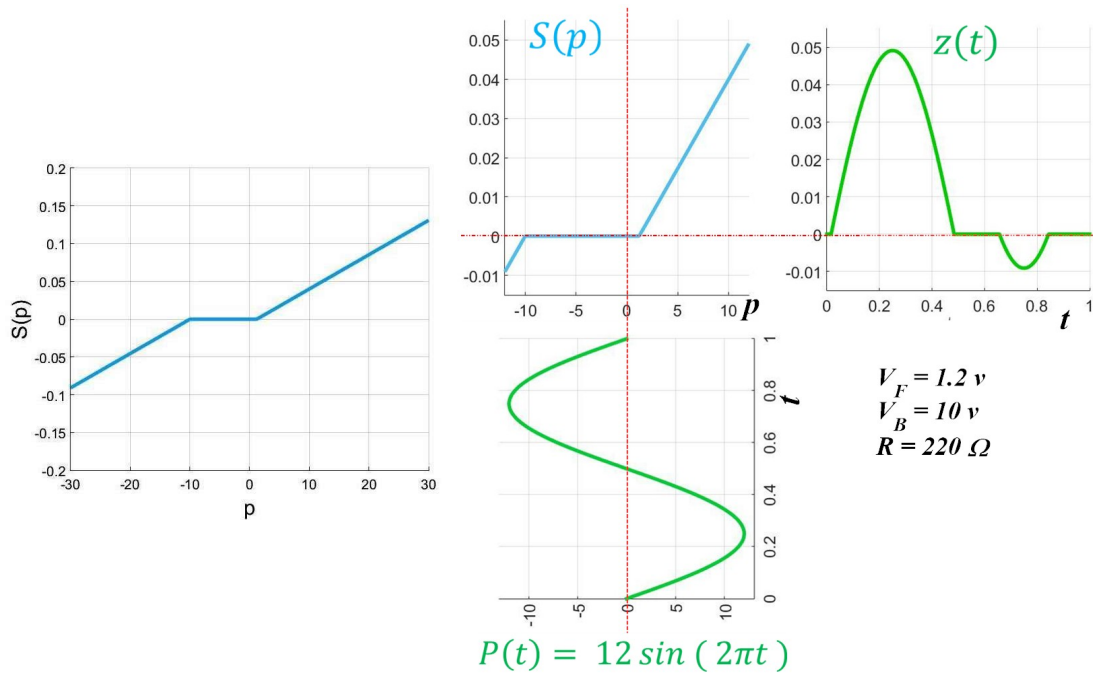


FIGURE 4.2: Solution mapping for the circuit in Figure 4.1 (left), graphical method to find the output for a typical input (right)

defined in this way, would fit in Equation (2.16) and fulfils our intuition.

Defining the solution mapping in terms of the parameter t , and not directly of the input signal p , will cause some difficulties to study the perturbation problem.

Comparing to the situation in Chapter 3, although for each t one needs to solve a generalized equation of the type discussed in depth in Chapter 3, the Aubin property¹ of the solution mapping at a certain point is not sufficient any more to guarantee the stability of the output with respect to the perturbations of the input signal. In other words, the relation between S and $p(\cdot)$ is not explicitly expressed now.

So the first question one needs to answer is whether some pointwise local stability property of the new solution mapping can be related to a general statement about it or not. We will tackle this problem in Subsections 4.1.1 and 4.2.2.

Next, we will focus on the solution trajectories. In particular, we ask questions related to the existence of selections which are smooth functions, their relationship to the input signal, and their reaction to the small perturbations of the input signal. Section 4.2 is devoted to provide answers to these questions.

Let us note that under the general assumptions:

- (A2) f is continuously differentiable in \mathbb{R}^n ;
- (A3) F has closed graph;

¹ or any other local stability property like calmness or isolated calmness of S or equivalently, metric regularities (all four different definitions) of S^{-1} .

when p is a continuous function, the map S has closed graph.

Indeed, consider $\{(t_n), (z_n)\}_{n \in \mathbb{N}}$ such that $(t_n, z_n) \in \text{gph } S$ for every n ; and $(t_n, z_n) \rightarrow (t_0, z_0)$. We show that $(t_0, z_0) \in \text{gph } S$.

For each n , consider $w_n := -f(z_n) + p(t_n)$. Since, by assumption, $z_n \in S(t_n)$, we get that $w_n \in F(z_n)$. Now, by the continuity of p and f , we obtain that $w_n \rightarrow w_0 = -f(z_0) + p(t_0)$.

Since F has closed graph we get that $(w_0, z_0) \in \text{gph } F$, i.e. $w_0 \in F(z_0)$. Therefore, $F(z_0) + f(z_0) - p(t_0) \ni 0$, that is, $z_0 \in S(t_0)$.

Consider a function $h : \mathbb{R} \times \mathbb{R}^n \rightarrow \mathbb{R}^n$, defined as $h(t, v) = -p(t) + f(v)$. Then, the generalized equation (4.3) can be written as

$$h(t, v) + F(v) \ni 0, \quad (4.5)$$

which is called a *parametric generalized equation*, since the function h now depends on a scalar parameter $t \in [0, 1]$. For any given $(t, z) \in \text{gph } S$, define the mapping

$$v \mapsto G_{t,z}(v) := h(t, z) + \nabla_z h(t, z)(v - z) + F(v). \quad (4.6)$$

A point $(t, z) \in \mathbb{R}^{1+n}$ is said to be a *strongly regular point*² for the generalized equation (4.3, or equivalently, 4.5) when $(t, z) \in \text{gph } S$ and the mapping $G_{t,z}$ is strongly metrically regular at z for 0. That is, there exist constants $a_t, b_t, \lambda_t > 0$ such that the mapping

$$\mathbb{B}_{b_t}(0) \ni y \mapsto G_{t,z}^{-1}(y) \cap \mathbb{B}_{a_t}(z(t)) \quad (4.7)$$

is a Lipschitz continuous function with a Lipschitz constant λ_t .

The following theorem will show that the *strong regularity condition* ensures a good behaviour of the solutions of the nonlinear problem. Although it is not explicitly mentioned, a closer look at the definition of estimators (Definition 1.4.3) and the note after it (Note 1.4.4), reveals its role behind the scene.

² **Historical note.** This term was first introduced at 1980 by Robinson [60], and then became popular in the literature. In his work, the set-valued mapping F is always a normal cone to a non-empty closed convex subset C of a normed linear space, and he found this setting a convenient tool for formulating some problems in complementarity and in mathematical programming, as well as variational inequalities.

As he mentions, the idea behind this condition is that it is analogous to the non-singularity condition imposed in the usual implicit function theorem for nonlinear equations, and indeed, it reduces to that condition (in his setting) when C is the whole space (so that the generalized equation reduces to the equation $f(x) = 0$).

Note that applying a condition on the “linearization” of (4.5) will allow the numerical methods to enter the scene more effectively.

Theorem 4.1.1 (Implicit Function Theorem for Generalized Equations).

[29, Theorem 2B.7, p. 89] Consider a function $f : \mathbb{R}^d \times \mathbb{R}^n \rightarrow \mathbb{R}^n$ and a mapping $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ with $(\bar{p}, \bar{x}) \in \text{int dom } f$ and $f(\bar{p}, \bar{x}) + F(\bar{x}) \ni 0$, and suppose that $\widehat{\text{lip}}_p(f; (\bar{p}, \bar{x})) \leq \gamma < \infty$. Let h be a strict estimator of f with respect to x uniformly in p at (\bar{p}, \bar{x}) with constant μ . Suppose that $(h + F)^{-1}$ has a Lipschitz continuous single-valued localization σ around 0 for \bar{x} with $\text{lip}(\sigma; 0) \leq \kappa$ for a constant κ such that $\kappa\mu < 1$. Then the solution mapping

$$S : p \mapsto \{x \in \mathbb{R}^n \mid f(p, x) + F(x) \ni 0\} \quad \text{for } p \in \mathbb{R}^d$$

has a Lipschitz continuous single-valued localization s around \bar{p} for \bar{x} with

$$\text{lip}(s; \bar{p}) \leq \frac{\kappa\gamma}{1 - \kappa\mu}.$$

4.1.1 Review on an Existence Theorem

From Theorem 4.1.1, one obtains that when (\bar{t}, \bar{z}) is a strongly regular point for (4.3), there are open neighborhoods T of \bar{t} and U of \bar{z} such that the mapping

$$T \cap [0, 1] \ni \tau \mapsto S(\tau) \cap U \tag{4.8}$$

is single-valued and Lipschitz continuous on $T \cap [0, 1]$.

The theorem which follows shows that if each point in $\text{gph } S$ is strongly regular, then there are finitely many Lipschitz continuous solution trajectories defined on $[0, 1]$ whose graphs never intersect each other. In addition, along any such trajectory $\bar{u}(\cdot)$ the mapping $G_{t, \bar{u}(t)}$ is strongly regular uniformly in $t \in [0, 1]$, meaning that the neighborhoods and the constants involved in the definition do not depend on t .

Since we will work on the idea and assumptions of this theorem, and its proof needs some clarification, we add the proof here for the sake of completeness. We preferred to insert the added details as footnotes as long as it was possible.³

Theorem 4.1.2 (Uniform Strong Metric Regularity). [29, Theorem 6G.1, p. 426]

Suppose that there exists a bounded set $C \subset \mathbb{R}^n$ such that, for each $t \in [0, 1]$, the set $S(t)$ is non-empty and contained in C for all $t \in [0, 1]$. Also, suppose that each point in $\text{gph } S$ is strongly regular. Then there are finitely many Lipschitz continuous functions $\bar{u}_j : [0, 1] \rightarrow \mathbb{R}^n$, $j = 1, 2, \dots, M$ such that, for each $t \in [0, 1]$, one has $S(t) = \bigcup_{1 \leq j \leq M} \{\bar{u}_j(t)\}$.

Moreover, the graphs of the functions \bar{u}_j are isolated from each other, in the sense that

³ Check also [this web page](#) for Errata and Addenda.

there exists $\delta > 0$ such that

$$\|\bar{u}_{j'}(t) - \bar{u}_j(t)\| \geq \delta \quad \text{for every } j' \neq j \quad \text{and every } t \in [0, 1].$$

Furthermore, there exist positive constants a, b and λ such that, for each such function \bar{u}_i , and for each $t \in [0, 1]$ the mapping

$$\mathbb{B}_b(0) \ni w \longmapsto G_{t, \bar{u}_i(t)}^{-1}(w) \cap \mathbb{B}_a(\bar{u}_i(t))$$

is a Lipschitz continuous function with a Lipschitz constant λ .

Proof. From the assumed uniform boundedness of the solution mapping S and the continuity of f and its derivatives, we get the existence of a constant $K > 0$ such that

$$\sup_{t \in [0, 1], v \in C} \left(\|\nabla_t f(t, v)\| + \|\nabla_u f(t, v)\| + \|\nabla_{uu}^2 f(t, v)\| + \|\nabla_{ut}^2 f(t, v)\| \right) \leq K. \quad (4.9)$$

Let $(t, v) \in \text{gph } S$. Then, according to Theorem 4.1.1 there exists a neighborhood $T_{t,v}$ of t which is open relative to $[0, 1]$ and an open neighborhood $U_{t,v}$ of v such that the mapping $T_{t,v} \ni \tau \mapsto S(\tau) \cap U_{t,v}$ is a function, denoted by $u_{t,v}(\cdot)$, which is Lipschitz continuous on $T_{t,v}$ with Lipschitz constant $L_{t,v}$. From the open covering $\{T_{t,v} \times U_{t,v}\}_{(t,v) \in \text{gph } S}$ of the graph of S , which is a compact set in \mathbb{R}^{1+n} (due to the boundedness assumption on C), we can extract a finite subcovering $\{T_{t_j, v_j} \times U_{t_j, v_j}\}_{j=1}^M$. Let $L = \max_{1 \leq j \leq M} L_{t_j, v_j}$.

We will prove the theorem in three steps corresponding to the following claims:

1. existence of finitely many Lipschitz continuous trajectories;
2. isolation of the graphs of the trajectories;
3. existence of uniform bounds.

STEP 1. Let $\tau \in [0, 1]$ and choose any $\bar{u} \in S(\tau)$. Now we will prove that there exists a Lipschitz continuous function $\bar{u}(\cdot)$ with Lipschitz constant L such that $\bar{u}(t) \in S(t)$ for all $t \in [0, 1]$ and $\bar{u}(\tau) = \bar{u}$.

Assume $\tau < 1$. Then there exists $j \in \{1, \dots, M\}$ such that $(\tau, \bar{u}) \in T_{t_j, v_j} \times U_{t_j, v_j}$. Define $\bar{u}(t) = u_{t_j, v_j}(t)$ for all $t \in (t'_j, t''_j) := T_{t_j, v_j}$. Then $\bar{u}(\tau) = u_{t_j, v_j}(\tau) = \bar{u}$ [first part of the claim] and $\bar{u}(\cdot)$ is Lipschitz continuous on $[t'_j, t''_j]$.⁴ If $t''_j < 1$ then there exists some $i \in \{1, \dots, M\}$ such that $(t''_j, \bar{u}(t''_j)) \in T_{t_i, v_i} \times U_{t_i, v_i} := (t'_i, t''_i) \times U_{t_i, v_i}$. Then of course $u_{t_i, v_i}(t''_j) = \bar{u}(t''_j)$ [single-valuedness of $S \cap U_{t_i, v_i}$]; and we can extend $\bar{u}(\cdot)$ to $[t_j, t''_i]$ as $\bar{u}(t) = u_{t_i, v_i}(t)$ for $t \in [t_j, t''_i]$. After at most M such steps we extend $\bar{u}(\cdot)$ to $[t_j, 1]$. By

⁴ Let $t \in (t'_j, t''_j)$, and consider a sequence $(t_j^n) \rightarrow t''_j$ such that $t_j^n < t''_j$ for each n . Then, by the Lipschitz continuity of \bar{u} inside the interval, and the continuity of the norm we get

$$\begin{aligned} \|\bar{u}(t''_j) - \bar{u}(t)\| &= \left\| \bar{u}\left(\lim_{n \rightarrow \infty} t_j^n\right) - \bar{u}(t) \right\| = \left\| \lim_{n \rightarrow \infty} \bar{u}(t_j^n) - \bar{u}(t) \right\| = \lim_{n \rightarrow \infty} \|\bar{u}(t_j^n) - \bar{u}(t)\| \\ &\leq \lim_{n \rightarrow \infty} L \|t_j^n - t\| = L \|t''_j - t\|. \end{aligned}$$

repeating the same argument on the interval $[0, \tau]$ we extend $\bar{u}(\cdot)$ on the entire interval $[0, 1]$ thus obtaining a Lipschitz continuous selection for S . If $\tau = 1$ then we repeat the same argument on $[0, 1]$ starting from 1 and going to the left.

To finish the proof of the first claim, we assume that (τ, \bar{u}) and (θ, \tilde{u}) are two points in $\text{gph } S$ and let $\bar{u}(\cdot)$ and $u(\cdot)$ be the functions determined by the above procedure such that $\bar{u}(\tau) = \bar{u}$ and $\tilde{u}(\theta) = \tilde{u}$.

Assume that $\bar{u}(0) \neq \tilde{u}(0)$ and the set $\Delta := \{t \in [0, 1] \mid \bar{u}(t) = \tilde{u}(t)\}$ is non-empty. Since Δ is closed⁵, $\inf \Delta := v > 0$ is attained and then we have that $\bar{u}(v) = \tilde{u}(v)$ and $\bar{u}(t) \neq \tilde{u}(t)$ for $t \in [0, v)$. But then $(v, \bar{u}(v)) \in \text{gph } S$ cannot be a strongly regular point of S , a contradiction⁶. Thus, the number of different Lipschitz continuous functions $\bar{u}(\cdot)$ constructed from points $(\tau, \bar{u}) \in \text{gph } S$ is not more than the number of points in $S(0)$ ⁷. Hence there are finitely many Lipschitz continuous functions $\bar{u}(\cdot)$ such that for every $t \in [0, 1]$ one has $S(t) = \bigcup_{1 \leq j \leq M} \{\bar{u}_j(t)\}$. This proves the first part of the theorem.

STEP 2. The fact that the solutions are isolated from each other, is implicitly shown in the contradiction above and follows from the fact that there are “finitely many” $\bar{u}_j(\cdot)$. The argument is that if there are two different functions $\bar{u}(\cdot)$, and $\tilde{u}(\cdot)$ with an intersection point, say $\bar{u}(\tau) = \tilde{u}(\tau) = \bar{u}$ for a point $(\tau, \bar{u}) \in \text{gph } S$, then $\Delta \neq \emptyset$ and we must have $\bar{u}(0) = \tilde{u}(0)$, and so on for every t . Thus, the two functions have either one point of intersection, and are totally the same, or they do not have any intersection at all.

STEP 3. Choose a Lipschitz continuous function $\bar{u}(\cdot)$ whose graph is in the graph of S , that is, $\bar{u}(\cdot)$ is one of the functions $\bar{u}_j(\cdot)$ and its Lipschitz constant is L . Let $t \in (0, 1)$ and let $G_t = G_{t, \bar{u}(t)}$ ⁸, for simplicity. Let a_t, b_t and λ_t be positive constants such that the

So $\bar{u}(\cdot)$ is Lipschitz continuous at t'_j (and also at t''_j with a similar reasoning). It remains to show that $(t'_j, \bar{u}(t'_j)) \in \text{gph } S$. Since $\text{gph } S$ is a compact set and for each n , we have $(t'_j, \bar{u}(t'_j)) \in \text{gph } S$;

$$t'_j \rightarrow t''_j, \text{ and } \bar{u}(t'_j) \rightarrow \bar{u}(t''_j) \text{ as } n \rightarrow \infty$$

we get the assertion.

⁵ In fact, Δ is the zero level set of the continuous function $\bar{u} - \tilde{u}$.

⁶ To be more clear, in order to have $(v, \bar{u}(v))$ as a strongly regular point of $\text{gph } S$, we should have that the mapping $\tau \mapsto S(\tau) \cap U$ is single-valued in a vicinity of $(v, \bar{u}(v))$, which is not.

⁷ Suppose $(0, u_i) \in \text{gph } S$. Since every point in $\text{gph } S$ is a strongly regular point, one has that the mapping $T_{0, u_i} \ni \tau \mapsto S(\tau) \cap U_{0, u_i}$ is single-valued, especially, $S(0) \cap U_{0, u_i}$ is a singleton. Thus, $S(0) = \bigcup_{i \in I} (S(0) \cap U_{0, u_i})$, for an arbitrary index set I . But since graph of S is a compact subset of \mathbb{R}^{n+1} , there exists a finite set $J \subset I$ such that $S(0) \subset \bigcup_{i \in J} (S(0) \cap U_{0, u_i}) = \bigcup_{i \in J} \{u_i\}$.

⁸ One should be careful not to confuse this change of notation, which is only valid for the rest of this proof, with the auxiliary map G_t that will be defined by (4.19) in Section 4.2 and will be used till the end of this chapter.

mapping

$$\mathbb{B}_{b_t}(0) \ni w \mapsto G_t^{-1}(w) \cap \mathbb{B}_{a_t}(\bar{u}(t)) \quad (4.10)$$

is a Lipschitz continuous function with Lipschitz constant λ_t . Make $b_t > 0$ smaller if necessary so that

$$2b_t\lambda_t < a_t. \quad (4.11)$$

Let $\rho_t > 0$ be such that $L\rho_t < \frac{a_t}{2}$. Then, from the Lipschitz continuity of \bar{u} around t we have that $\mathbb{B}_{\frac{a_t}{2}}(\bar{u}(\tau)) \subset \mathbb{B}_{a_t}(\bar{u}(t))$ for all $\tau \in (t - \rho_t, t + \rho_t)$ ⁹. Make $\rho_t > 0$ smaller if necessary so that

$$K(L+1)\rho_t < \frac{1}{\lambda_t}, \quad \text{and} \quad \lambda_t\rho_t < \frac{a_t}{16K} \quad (4.12)$$

Our aim is to apply Theorem 1.4.14 (the strong regularity part) in order to show that there exist a neighborhood O_t of t and positive constants α_t and β_t such that for each $\tau \in O_t \cap [0, 1]$ the mapping

$$\mathbb{B}_{\beta_t}(0) \ni w \mapsto G_\tau^{-1}(w) \cap \mathbb{B}_{\alpha_t}(\bar{u}(t)) \quad (4.13)$$

is a Lipschitz continuous function.

Consider the function $g_{t,\tau} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ defined as

$$\begin{aligned} g_{t,\tau}(v) &= f(\tau, \bar{u}(\tau)) - f(t, \bar{u}(t)) \\ &\quad + \left(\nabla_u f(\tau, \bar{u}(\tau)) - \nabla_u f(t, \bar{u}(t)) \right) v \\ &\quad + \nabla_u f(t, \bar{u}(t)) \bar{u}(t) - \nabla_u f(\tau, \bar{u}(\tau)) \bar{u}(\tau). \end{aligned} \quad (4.14)$$

For each v we have $G_\tau(v) = G_t(v) + g_{t,\tau}(v)$. Since $g_{t,\tau}$ is an affine map, it is Lipschitz continuous. Let us show that the Lipschitz constant is bounded by the expression on the left of (4.12)¹⁰, by using a mean-value theorem in \mathbb{R}^n ¹¹. For simplicity, let $p_\tau :=$

⁹ Let $z \in \mathbb{B}_{\frac{a_t}{2}}(\bar{u}(\tau))$ be an arbitrary point. Then

$$\|z - \bar{u}(t)\| \leq \|z - \bar{u}(\tau)\| + \|\bar{u}(\tau) - \bar{u}(t)\| \leq \frac{a_t}{2} + L\|\tau - t\| < \frac{a_t}{2} + L\rho_t < \frac{a_t}{2} + \frac{a_t}{2}$$

Thus, $z \in \mathbb{B}_{a_t}(\bar{u}(t))$.

¹⁰ Here, we have used the fact that f is a twice continuously differentiable function. The following proof requires only Lipschitz continuity of $\nabla_u f$.

$$\begin{aligned} \|g_{t,\tau}(v) - g_{t,\tau}(v')\| &\leq \|\nabla_u f(\tau, \bar{u}(\tau)) - \nabla_u f(t, \bar{u}(t))\| \|v - v'\| \\ &\leq L_{\nabla_u f} (\|\tau - t\| + \|\bar{u}(\tau) - \bar{u}(t)\|) \|v - v'\| \\ &< K'(\rho_t + L\rho_t) \|v - v'\| = K'(L+1)\rho_t \|v - v'\| \end{aligned}$$

In which $L_{\nabla_u f}$ is the Lipschitz constant of $\nabla_u f$, and K' could be chosen in accordance with ρ_t , in a way that $K'(L+1)\rho_t < \frac{1}{\lambda_t}$ holds true.

¹¹ For more details on mean value theorems in \mathbb{R}^n , refer to [58, Section 3.2], specially Theorem 3.2.3, page 69. We used in the above proof the main idea of that theorem (instead of using the end result

$(\tau, \bar{u}(\tau))$, and $p_t := (t, \bar{u}(t))$.

$$\begin{aligned}
\|g_{t,\tau}(v) - g_{t,\tau}(v')\| &\leq \|\nabla_u f(p_\tau) - \nabla_u f(p_t)\| \|v - v'\| \\
&\leq \left\| \int_0^1 \frac{d}{ds} [\nabla_u f(p_t + s(p_\tau - p_t))] ds \right\| \|v - v'\| \\
&\leq \int_0^1 \left\| \frac{d}{ds} [\nabla_u f(t + s(\tau - t), \bar{u}(t) + s(\bar{u}(\tau) - \bar{u}(t)))] \right\| ds \|v - v'\| \\
&\leq \int_0^1 \left\| \nabla_{ut}^2 f(p_t + s(p_\tau - p_t))(\tau - t) \right. \\
&\quad \left. + \nabla_{uu}^2 f(p_t + s(p_\tau - p_t))(\bar{u}(\tau) - \bar{u}(t)) \right\| ds \|v - v'\| \\
&\leq \left(\sup_{0 \leq s \leq 1} \left\| \nabla_{ut}^2 f(p_t + s(p_\tau - p_t)) \right\| \|\tau - t\| \right. \\
&\quad \left. + \sup_{0 \leq s \leq 1} \left\| \nabla_{uu}^2 f(p_t + s(p_\tau - p_t)) \right\| \|\bar{u}(\tau) - \bar{u}(t)\| \right) \|v - v'\| \\
&\leq K (\|\tau - t\| + \|\bar{u}(\tau) - \bar{u}(t)\|) \|v - v'\| \\
&\leq K(\rho_t + L\rho_t) \|v - v'\| \\
&\leq K(L + 1)\rho_t \|v - v'\|.
\end{aligned}$$

To show the boundedness of $\|g_{t,\tau}(\bar{u}(t))\|$, one can use elementary calculus and the mean value theorem mentioned in Footnote (11), again, to argue as follows

$$\begin{aligned}
\|g_{t,\tau}(\bar{u}(t))\| &= \|f(\tau, \bar{u}(\tau)) - f(t, \bar{u}(t)) + \nabla_u f(\tau, \bar{u}(\tau))(\bar{u}(t) - \bar{u}(\tau))\| \\
&= \left\| \int_0^1 \frac{d}{ds} f(t + s(\tau - t), \bar{u}(t) + s(\bar{u}(\tau) - \bar{u}(t))) ds + \nabla_u f(p_\tau)(\bar{u}(t) - \bar{u}(\tau)) \right\| \\
&= \left\| \int_0^1 (\tau - t) \nabla_t f(t + s(\tau - t), \bar{u}(t) + s(\bar{u}(\tau) - \bar{u}(t))) ds \right. \\
&\quad \left. + \int_0^1 (\bar{u}(\tau) - \bar{u}(t)) \nabla_u f(t + s(\tau - t), \bar{u}(t) + s(\bar{u}(\tau) - \bar{u}(t))) ds \right. \\
&\quad \left. - \int_0^1 \nabla_u f(\tau, \bar{u}(\tau))(\bar{u}(\tau) - \bar{u}(t)) ds \right\| \\
&\leq \|\tau - t\| \sup_{0 \leq s \leq 1} \left\| \nabla_t f(p_t + s(p_\tau - p_t)) \right\| \\
&\quad + \|\bar{u}(t) - \bar{u}(\tau)\| \left\| \int_0^1 \nabla_u f(p_t + s(p_\tau - p_t)) - \nabla_u f(p_\tau) ds \right\| \\
&\leq K\rho_t + L\rho_t \int_0^1 K \|(s-1)(p_\tau - p_t)\| ds \\
&\leq K\rho_t \left(1 + \frac{1}{2}L\rho_t + \frac{1}{2}L^2\rho_t\right).
\end{aligned}$$

of it, to give the reader a scheme of the proof). To be more precise, for a G-differentiable function $g : \mathbb{R}^n \rightarrow \mathbb{R}^m$ defined on a convex subset D_0 of \mathbb{R}^n , one has

$$\|g(x) - g(y)\| \leq \sup_{0 \leq s \leq 1} \|g'(x + s(y-x))\| \|x - y\| \quad \text{for all } x, y \in D_0.$$

We apply Theorem 1.4.14 with $F = G_t, (\bar{x}, \bar{y}) = (\bar{u}(t), 0), g = g_{t,\tau}, a = a_t, b = b_t, \kappa = \lambda_t$, and defining $\mu = \mu_t := K(L+1)\rho_t$, and

$$\kappa' = \lambda'_t := \frac{3\lambda_t}{2(1 - K(L+1)\rho_t\lambda_t)} > \frac{\lambda_t}{1 - \mu_t\lambda_t}. \quad (4.15)$$

For that purpose we need to show that there exist constants α_t and β_t that satisfy the inequalities

$$\alpha_t \leq \frac{a_t}{2}, \quad 2\mu_t\alpha_t + 2\beta_t \leq b_t, \quad 2\lambda'_t\beta_t \leq \alpha_t, \quad \|g_{t,\tau}(\bar{u}(t))\| \leq \beta_t. \quad (4.16)$$

Choose ρ_t smaller if necessary such that $\frac{1}{2}L\rho_t + \frac{1}{2}L^2\rho_t < 1$, then the above calculations reveal that $\|g_{t,\tau}(\bar{u}(t))\| \leq 2K\rho_t$.

Denoting $A := K(1+L)$ and $B := 2K$ we have

$$\mu_t = A\rho_t \quad \text{and} \quad \|g_{t,\tau}(\bar{u}(t))\| \leq B\rho_t.$$

Set $\beta_t := B\rho_t$. We will now show that there exists a positive α_t which satisfies all inequalities in (4.16).

Substituting the already chosen μ_t and β_t in (4.16), we obtain that α_t should satisfy

$$\begin{cases} \alpha_t \leq \frac{a_t}{2} \\ 2A\rho_t\alpha_t + 2B\rho_t \leq b_t \\ 3B\lambda_t\rho_t \leq \alpha_t(1 - A\rho_t\lambda_t). \end{cases} \quad (4.17)$$

The system (4.17) has a solution $\alpha_t > 0$ provided that

$$\frac{3B\lambda_t\rho_t}{1 - A\lambda_t\rho_t} \leq \frac{b_t - 2B\rho_t}{2A\rho_t} \quad \text{and} \quad \frac{b_t - 2B\rho_t}{2A\rho_t} \leq \frac{a_t}{2}.$$

Thus, everything comes down to checking whether this system of inequalities is consistent. But this system is consistent whenever

$$\rho_t(A\lambda_t b_t + 2B + 4AB\lambda_t\rho_t) \leq b_t \leq \rho_t(2B + Aa_t),$$

which holds when $A\lambda_t b_t + 4AB\lambda_t\rho_t \leq Aa_t$; which in turn always holds because of the assumptions (4.11) and (4.12).

We are now ready to apply Theorem 1.4.14 from which we conclude that the mapping in (4.13), which is

$$\mathbb{B}_{\beta_t}(0) \ni w \mapsto G_\tau^{-1}(w) \cap \mathbb{B}_{\alpha_t}(\bar{u}(t))$$

is a Lipschitz continuous function with Lipschitz constant λ'_t . The next step is to move the center of the second ball in the above intersection to $\bar{u}(\tau)$. We claim that the mapping

$$\mathbb{B}_{\frac{\beta_t}{2}}(0) \ni w \longmapsto G_\tau^{-1}(w) \cap \mathbb{B}_{\frac{\alpha_t}{2}}(\bar{u}(\tau)) \quad (4.18)$$

is a Lipschitz continuous function with Lipschitz constant λ'_t (or maybe less).

First we observe that $G_\tau^{-1}(w) \neq \emptyset$, for $w \in \mathbb{B}_{\frac{\beta_t}{2}}(0)$.

Let $u_1, u_2 \in G_\tau^{-1}(w) \cap \mathbb{B}_{\frac{\alpha_t}{2}}(\bar{u}(\tau))$. From the Lipschitz continuity of $\bar{u}(\cdot)$ around t we have that $\mathbb{B}_{\frac{\alpha_t}{2}}(\bar{u}(\tau)) \subset \mathbb{B}_{\alpha_t}(\bar{u}(t))$ (choosing ρ_t smaller if necessary such that $L\rho_t < \alpha_t/2$). Hence, we get $u_1, u_2 \in G_\tau^{-1}(w) \cap \mathbb{B}_{\alpha_t}(\bar{u}(t))$, which is a single-valued map and thus, we get $u_1 = u_2$.

The last part is to show $G_\tau^{-1}(w) \cap \mathbb{B}_{\frac{\alpha_t}{2}}(\bar{u}(\tau)) \neq \emptyset$.

For any $w_0 \in \mathbb{B}_{\frac{\beta_t}{2}}(0)$, we know that there exists a unique $u_0 := G_\tau^{-1}(w_0) \cap \mathbb{B}_{\alpha_t}(\bar{u}(t))$.

We also know that, by definition, $\bar{u}(\tau) \in G_\tau^{-1}(0)$. Since $\|\bar{u}(\tau) - \bar{u}(t)\| < L\rho_t < \frac{\alpha_t}{2}$, we get $\bar{u}(\tau) \in \mathbb{B}_{\alpha_t}(\bar{u}(t))$. So, $\bar{u}(\tau) = G_\tau^{-1}(0) \cap \mathbb{B}_{\alpha_t}(\bar{u}(t))$.

$$\begin{aligned} \|u_0 - \bar{u}(\tau)\| &= \|[G_\tau^{-1}(w_0) \cap \mathbb{B}_{\alpha_t}(\bar{u}(t))] - [G_\tau^{-1}(0) \cap \mathbb{B}_{\alpha_t}(\bar{u}(t))]\| \\ &\leq \lambda'_t \|w_0 - 0\| \\ &\leq \lambda'_t \frac{\beta_t}{2} \leq \frac{\alpha_t}{2}. \end{aligned}$$

Thus, $u_0 \in \mathbb{B}_{\frac{\alpha_t}{2}}(\bar{u}(\tau))$. In fact, we get $u_0 = G_\tau^{-1}(w_0) \cap \mathbb{B}_{\frac{\alpha_t}{2}}(\bar{u}(\tau))$.

A similar reasoning provides the Lipschitz property of the map in that neighborhood¹².

Now we are ready to unify the constants. From the open covering $\bigcup_{t \in [0,1]} (t - \rho_t, t + \rho_t)$ of $[0, 1]$ choose a finite subcovering of open intervals $(t_i - \rho_{t_i}, t_i + \rho_{t_i})$, $i = 1, 2, \dots, m$.

Let

$$\begin{aligned} a &:= \min \left\{ \frac{\alpha_{t_i}}{2} \mid i = 1, \dots, m \right\} \\ \lambda &:= \max \{ \lambda'_{t_i} \mid i = 1, \dots, m \} \\ \text{and } b &:= \min \left\{ \frac{a}{\lambda}, \min \left\{ \frac{\beta_{t_i}}{2} \mid i = 1, \dots, m \right\} \right\}. \end{aligned}$$

Since $b \leq a/\lambda$, the observation in Lemma 4.2.8 (possibility to reduce “proportionally” the radii of balls in the definition of a strongly metrically regular map) applies; hence, for each $\tau \in (t_i - \rho_{t_i}, t_i + \rho_{t_i}) \cap [0, 1]$ the mapping

$$\mathbb{B}_b(0) \ni w \longmapsto G_\tau^{-1}(w) \cap \mathbb{B}_a(\bar{u}(\tau))$$

¹² In fact, one may argue that in both (4.13) and (4.18), we deal with the graph of G_τ^{-1} , only the localization (i.e. the cutting area) changes. Since $(\mathbb{B}_{\frac{\beta_t}{2}}(0) \times \mathbb{B}_{\frac{\alpha_t}{2}}(\bar{u}(\tau))) \subset (\mathbb{B}_{\beta_t}(0) \times \mathbb{B}_{\alpha_t}(\bar{u}(t)))$, if there exists a point in (4.18), it is the same point of (4.13); and the only thing that needs to be checked is whether the new localization is empty-valued or not.

is a Lipschitz continuous function with Lipschitz constant λ . Let $t \in [0, 1]$; then $t \in (t_i - \rho_{t_i}, t_i + \rho_{t_i})$ for some $i \in \{1, \dots, m\}$. Hence the mapping

$$\mathbb{B}_b(0) \ni w \longmapsto G_t^{-1}(w) \cap \mathbb{B}_a(\bar{u}(t))$$

is a Lipschitz continuous function with Lipschitz constant λ . The proof is complete. \square

Although Theorem 4.1.2 is an important result in our study, there is an unpleasant assumption in that theorem which is: $S(t)$ is uniformly bounded. Even if this condition is fulfilled, it is hard to be checked since it requires the whole set $S(t)$ to be clarified and available for any $t \in [0, 1]$.

In the next section it will be shown that the uniform bound could be obtained without this extra assumption. Moreover, we will provide some results about the nature of the solution trajectories related to specific classes of input functions and their perturbations.

4.2 A Study about the Solution Trajectories

As a completion to the shift of interest from studying the solution mapping of a generalized equation to searching for certain properties of the solution trajectories, we described in the previous section; in this section, we will present some results about the behaviour of the solution trajectories, first for the generalized equation (4.3), and then for the perturbed generalized equation.

Though we are inspired by Robinson's idea of strongly regular points in defining the auxiliary map (4.6), and the techniques in Theorems 4.1.1 and 4.1.2, we find it more convenient to do some modifications in the setting in order to adapt it to our problem. Since our aim in this section is the study of the solution trajectories with respect to variations of the input function, $p(t)$, and since working with the function $f(z)$ or with its first order approximation does not play an important role in our case (the proof of this statement will follow soon), we assume to deal with $f(z)$ itself and so to consider the auxiliary mapping

$$v \mapsto G_t(v) := f(v) - p(t) + F(v). \quad (4.19)$$

For more details on different possible choices of auxiliary maps and how the strong metric regularity would be affected, we state the following proposition:

Proposition 4.2.1 (Different Auxiliary Maps).

Given the generalized equation $f(t, u) + F(u) \ni 0$, where $f : \mathbb{R} \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ is a function, and $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ is a set-valued map with closed graph, let us denote by h a strict estimator of f with respect to u uniformly in t , at (\bar{t}, \bar{u}) with a constant μ (cf. Definition

1.4.3). Consider the following auxiliary maps:

$$G_{\bar{t}, \bar{u}}(v) = h(v) + F(v), \quad (4.20)$$

$$G_{\bar{t}}(v) = f(\bar{t}, v) + F(v). \quad (4.21)$$

Then, $G_{\bar{t}, \bar{u}}$ is SMR at \bar{u} for 0, if and only if $G_{\bar{t}}$ is SMR at \bar{u} for 0, provided that the regularity modulus of each map times μ is less than 1.

Proof. First observe that, by definition of a strict estimator, $h(\bar{u}) = f(\bar{t}, \bar{u})$ and so, $0 \in G_{\bar{t}, \bar{u}}(\bar{u})$ is equivalent to $0 \in G_{\bar{t}}(\bar{u})$. Now taking into account the pointwise relation

$$G_{\bar{t}}(v) = f(\bar{t}, v) + F(v) = G_{\bar{t}, \bar{u}}(v) + f(\bar{t}, v) - h(v),$$

one can define a map $g_{\bar{t}, \bar{u}} : \mathbb{R}^n \rightarrow \mathbb{R}^n$ with $g_{\bar{t}, \bar{u}}(v) = f(\bar{t}, v) - h(v)$. For any $v_1, v_2 \in U$ (a neighborhood of \bar{u}), we get

$$\begin{aligned} \|g_{\bar{t}, \bar{u}}(v_1) - g_{\bar{t}, \bar{u}}(v_2)\| &= \|f(\bar{t}, v_1) - h(v_1) - f(\bar{t}, v_2) + h(v_2)\| \\ &= \|e(\bar{t}, v_1) - e(\bar{t}, v_2)\| \\ &\leq \mu \|v_1 - v_2\| \end{aligned}$$

where the last inequality is obtained by definition of strict estimator. Thus, $g_{\bar{t}, \bar{u}}$ is Lipschitz continuous around \bar{u} .

Now one can use Theorem 1.4.12 with $G = G_{\bar{t}, \bar{u}}$ and $g = g_{\bar{t}, \bar{u}}$ and, by assuming that $\mu \operatorname{reg}(G_{\bar{t}, \bar{u}}; \bar{u}) < 1$, to conclude that $g + G = G_{\bar{t}}$ has a Lipschitz continuous single-valued localization around $0 + g(\bar{u})$ for \bar{u} . Since $g_{\bar{t}, \bar{u}}(\bar{u}) = h(\bar{u}) - f(\bar{t}, \bar{u}) = 0$, the latter could be expressed as the SMR of $G_{\bar{t}}$ at \bar{u} for 0.

The converse implication is satisfied in a similar way by letting $G = G_{\bar{t}}$, $g = -g_{\bar{t}, \bar{u}}$ and assuming $\mu \operatorname{reg}(G_{\bar{t}}; \bar{u}) < 1$. \square

Remark 4.2.2. (a) A closer look at the proof reveals that if h is a strict estimator, then the regularity modulus of $G_{\bar{t}}$ and $G_{\bar{t}, \bar{u}}$ are related to each other with $\kappa' = \frac{\kappa}{1 - \kappa\mu}$. Considering a partially first order approximation of f like $h(v) = f(\bar{t}, \bar{u}) + \nabla f_u(\bar{t}, \bar{u})(v - \bar{u})$, will result in the same modulus for auxiliary maps (since $\mu = 0$ in this case).

(b) One should note that in general, $h_1(\cdot) = f(\bar{t}, \cdot)$ is not a strict estimator of f at the reference point. To guarantee this, one needs an extra assumption like the following:

$$f(t, \cdot) \text{ is Lipschitz continuous, for any } t \text{ in a neighborhood of } \bar{t}.$$

However, this is not a necessary condition. For example, in the specific case we are interested in, that is $h(t, u) = f(u) - p(t)$, h_1 is automatically a strict estimator with $\mu = 0$

(in fact, a partial first order approximation), without requiring a Lipschitz continuity assumption on $p(\cdot)$.

4.2.1 Continuity of Solution Trajectories

In this subsection we will discuss the smoothness relation between the input signal and solution trajectories under the strong metric regularity assumption of the auxiliary map (4.19). For future reference we remind that the *solution mapping*, in our setting, is defined as

$$S : t \mapsto S(t) = \{z \in \mathbb{R}^n \mid G_t(z) \ni 0\}, \quad (4.22)$$

and a function $z : [0, 1] \rightarrow \mathbb{R}^n$ is called a *solution trajectory* if

$$z(t) \in S(t), \quad \text{for all } t \in [0, 1]. \quad (4.23)$$

Throughout the whole subsection we will assume that, given a function $p(\cdot)$, a solution trajectory $z(\cdot)$ exists. The reason we mentioned the existence result in Subsection 4.1.1 is to guarantee that this assumption is not nonsense.

Let us start with a simple observation that will be used several times in this chapter. The following lemma will provide a rule for moving from one auxiliary map to another. This simple yet handy result is a consequence of our choice of auxiliary map and our setting.

Lemma 4.2.3. *Consider the generalized equation (4.3), and the auxiliary map (4.19). For arbitrary points $t, t' \in [0, 1]$, the following equalities hold*

$$\boxed{G_t(v) = G_{t'}(v) + p(t') - p(t)} \quad (4.24)$$

$$\boxed{G_t^{-1}(w) = G_{t'}^{-1}(w + p(t) - p(t'))} \quad (4.25)$$

Proof. The first equality is trivial. One only needs to write down the definition of the auxiliary map G_t . For the second one, consider $v \in G_t^{-1}(w)$. Then, $w \in G_t(v) = G_{t'}(v) + p(t') - p(t)$, implies $v \in G_{t'}^{-1}(w + p(t) - p(t'))$.

Since there is nothing special about t and t' , the inverse inclusion also holds true, and hence the equality is proved. \square

In the following proposition, we will prove a continuity result for a given solution trajectory under suitable assumptions. One of the assumptions is that “different” trajectories, that is, trajectories without intersections, may not get arbitrary close to each other. In

Theorem 4.1.2, we have already seen this assumption as a result, which comes after continuity of the trajectories. Here, as a somehow inverse statement, we start from “isolated trajectories” and prove their continuity.

Since continuity of the trajectories is not assumed any more, one should be careful about how to formulate this property. We use an expression based on the graphs of trajectories (see [18]). Example 4.2.7 after the proposition will provide some ideas for the difficulties that may arise by “bad” formulations.

Proposition 4.2.4 (Smoothness Dependence of Trajectories on Input Signal).

For the generalized equation (4.3), and the solution mapping (4.22), assume that

- (i) $z(\cdot)$ is a given solution trajectory which is **isolated** from other trajectories; that is, there is an open set $\mathcal{O} \in \mathbb{R}^{n+1}$ such that

$$\{(t, v) \mid t \in [0, 1] \text{ and } 0 \in G_t(v)\} \cap \mathcal{O} = \text{gph } z. \quad (4.26)$$

- (ii) $p(\cdot)$ is a continuous function;

- (iii) G_t is pointwise strongly metrically regular; i.e. for any $t \in [0, 1]$, G_t is strongly metrically regular at $z(t)$ for 0, with constants a_t, b_t , and $\kappa_t > 0$ defined as (4.7).

Then $z(\cdot)$ is a continuous function.

Proof. Fix $t \in [0, 1]$. We know that $(t, z(t)) \in \text{gph } S$, so $0 \in G_t(z(t))$ or $z(t) \in G_t^{-1}(0)$. For any $\epsilon > 0$, let $\epsilon_1 := \min\{\frac{\epsilon}{\kappa_t}, b_t\}$, in which b_t is the radius of the neighborhood around 0 in the assumption (iii). By the uniform continuity of $p(\cdot)$, there exists $\delta_p > 0$ such that

$$\|p(t) - p(\tau)\| < \epsilon_1 \quad \text{whenever} \quad \|\tau - t\| < \delta_p.$$

Let $\delta < \delta_p$ and consider $\tau \in [0, 1]$ such that $\|\tau - t\| < \delta$. By definition, $z(\tau) \in G_\tau^{-1}(0)$. Using Lemma 4.2.3 we obtain $z(\tau) \in G_t^{-1}(p(\tau) - p(t))$.

On the other hand, by assumption (i) we also know that $z(\tau) \in \mathbb{B}_{a_t}(z(t))$.

Indeed, assuming $z(\tau) \notin \mathbb{B}_{a_t}(z(t))$, allows us to define a Lipschitz continuous function \tilde{z} as

$$\tilde{z}(\tau) := G_t^{-1}(p(\tau) - p(t)) \cap \mathbb{B}_{a_t}(z(t))$$

on $\mathbb{B}_\delta(t)$. By Lemma 4.2.3, $\tilde{z}(\tau) \in G_\tau^{-1}(0)$ and thus, is (part of) a solution trajectory. Now, consider a sequence (t_n) in $\mathbb{B}_\delta(t)$ converging to t , and recall that, by definition, $z(t) = G_t^{-1}(0) \cap \mathbb{B}_{a_t}(z(t))$. Thus, $\tilde{z}(t_n) \rightarrow z(t)$.

This means that \tilde{z} is a solution trajectory that could get arbitrarily close to $z(\cdot)$ at $(t, z(t))$, which contradicts assumption (i).

So $z(\tau) \in G_t^{-1}(y) \cap \mathbb{B}_{a_t}(z(t))$ where $y \in \mathbb{B}_{b_t}(0)$.

Now, by assumption (iii), the mapping $G_t^{-1}(\cdot) \cap \mathbb{B}_{a_t}(z(t))$ is single-valued and Lipschitz

continuous on $\mathbb{B}_{b_t}(0)$ with Lipschitz constant κ_t . So

$$\|z(t) - z(\tau)\| \leq \kappa_t \|p(t) - p(\tau)\| < \epsilon.$$

Since t was an arbitrary point in $[0, 1]$, the proof is complete. \square

Remark 4.2.5. If we assume that $p(\cdot)$ is a Lipschitz continuous function, then following the previous proof by considering $\tau_1, \tau_2 \in \mathbb{B}_\delta(t)$, we can obtain

$$\|z(\tau_1) - z(\tau_2)\| \leq \kappa_t \| [p(\tau_1) - p(t)] - [p(\tau_2) - p(t)] \| \leq \kappa_t L_p \|\tau_1 - \tau_2\|.$$

This means that $z(\cdot)$ is locally Lipschitz on $[0, 1]$ which is a compact set; so it is globally Lipschitz and we can restate the proposition as the following corollary.

Corollary 4.2.6. *Assume that*

- (i) $z(\cdot)$ is a given continuous solution trajectory;
- (ii) $p(\cdot)$ is a Lipschitz continuous function;
- (iii) G_t is pointwise strongly metrically regular at $z(t)$ for 0.

Then $z(\cdot)$ is a Lipschitz continuous function.

Example 4.2.7. Having a look at Theorem 4.1.2, one can think of defining isolation of trajectories in this way: \bar{u}_j is isolated from other trajectories, in the sense that there exists $\delta > 0$ such that

$$\|\bar{u}_{j'}(t) - \bar{u}_j(t)\| \geq \delta \text{ for every } j' \neq j \text{ and every } t \in [0, 1]. \quad (4.27)$$

Since we have not proved the continuity of trajectories and a selection could be made as a function with “jump”, this condition could not be valid for any trajectory. Consider the functions z_1, z_2 , and \bar{z} in Figure 4.3.

$$\begin{cases} z_1 : [0, 1] \rightarrow \mathbb{R} \\ z_1(t) = 1 \end{cases}, \quad \begin{cases} z_2 : [0, 1] \rightarrow \mathbb{R} \\ z_2(t) = 0 \end{cases}, \quad \bar{z}(t) = \begin{cases} z_1(t) & t \in [0, t_0] \\ z_2(t) & t \in [t_0, 1] \end{cases}$$

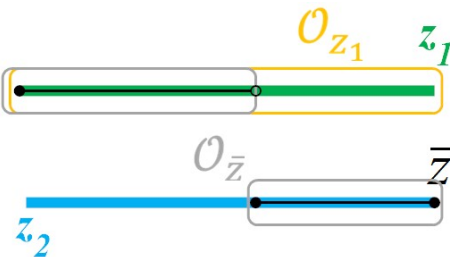


FIGURE 4.3: Difficulty in defining isolated trajectories

In this example, solution \bar{z} is not isolated from z_1 , nor z_2 in the sense of (4.27). In fact, there is no solution trajectory which is isolated from all other trajectories (infinitely many others!). Thus, criteria (4.27) is nonsense.

However, using the formulation of (4.26), one can see that z_1 is isolated from z_2 , and \bar{z} ; the same holds for z_2 . While for the trajectory \bar{z} , the situation is a bit different. Any open set $\mathcal{O} \in \mathbb{R}^2$ around the $\text{gph } \bar{z}$ (like the grey rectangles in the figure), may include points of $\text{gph } z_1$, or $\text{gph } z_2$ which does not belong to $\text{gph } \bar{z}$. Hence, \bar{z} (and any other solution trajectory that could be made in this way) is not isolated from z_1 , nor z_2 , and automatically drops out of the discussion of Proposition 4.2.4.

4.2.2 Uniform Strong Metric Regularity

In this subsection we focus our attention on the uniform strong metric regularity of G_t . One can consider this subsection as a quite independent section, but since we need the results we obtain here in the following subsection where some properties of the solution trajectories will be discussed, we prefer to keep logical hierarchy, and put this subsection under the general name of “a study about the solution trajectories”.

We have already seen Theorem 4.1.2 about uniform strong metric regularity. Our aim is to provide statements under simpler conditions, adapted to our particular setting.

In order to clarify the next statement, we remind that pointwise strong metric regularity of G_t for all $t \in [0, 1]$, guarantees for each $t \in [0, 1]$ the existence of constants $a_t, b_t, \kappa_t > 0$ such that the mapping

$$\mathbb{B}_{b_t}(0) \ni y \mapsto G_t^{-1}(y) \cap \mathbb{B}_{a_t}(z(t))$$

is single valued and Lipschitz continuous with a Lipschitz constant κ_t .

It is worthwhile noting that the radii a_t, b_t can be decreased provided that a suitable proportion is kept. The details are expressed in following lemma.

Lemma 4.2.8 (Proportional Reduction of Radii). *Let H be a strongly metrically regular map at \bar{x} for \bar{y} with a Lipschitz constant $\kappa > 0$ and neighborhoods $\mathbb{B}_a(\bar{x})$ and $\mathbb{B}_b(\bar{y})$. Then for every positive constants*

$$a' \leq a \text{ and } b' \leq b \text{ such that } \kappa b' \leq a',$$

the mapping H is strongly metrically regular with the corresponding Lipschitz constant κ and neighborhoods $\mathbb{B}_{a'}(\bar{x})$ and $\mathbb{B}_{b'}(\bar{y})$.

Proof. Since $B_{b'}(\bar{y}) \subset B_b(\bar{y})$ by assumption, $H^{-1}(y) \cap \mathbb{B}_a(\bar{x}) =: x$ for every $y \in B_{b'}(\bar{y})$. Taking into account that $H^{-1}(\cdot) \cap \mathbb{B}_a(\bar{x})$ is a Lipschitz continuous function on $B_b(\bar{y})$,

and by definition, $\bar{x} := H^{-1}(\bar{y}) \cap \mathbb{B}_a(\bar{x})$, we get:

$$\|x - \bar{x}\| = \|(H^{-1}(y) \cap \mathbb{B}_a(\bar{x})) - (H^{-1}(\bar{y}) \cap \mathbb{B}_a(\bar{x}))\| \leq \kappa \|y - \bar{y}\| \leq \kappa b'.$$

So, under the condition $\kappa b' \leq a'$, we get $x \in \mathbb{B}_{a'}(\bar{x})$.

Indeed, in this case any $y \in B_{b'}(\bar{y})$ will be in the domain of $H^{-1}(\cdot) \cap \mathbb{B}_{a'}(\bar{x})$. Then, the single-valuedness and Lipschitz continuity are the consequences of dealing with the same map (that is, $\text{gph } H$). \square

Remark 4.2.9. Fixing a point t , and the corresponding radii a_t , and b_t , we want to study the effect that increasing or decreasing these radii can have on the properties of the graphical localization.

Choosing a bigger a (letting $a > a_t$) may cause the graphical localization loose single valuedness by adding extra points to the localized map; while decreasing a may omit some points from the localized map and therefore, cause emptiness.

Increasing b , enlarges the domain of the localized map and may cause both unpleasant situations; multi-valuedness and emptiness. But even if the new points remain properly (i.e. in a single-valued manner) inside the second neighbourhood, the function may act not enough smoothly and cause losing the Lipschitz property of the localized map. On the other hand, decreasing b may not harm anything as it only drops some well-behaved points out of our scope.

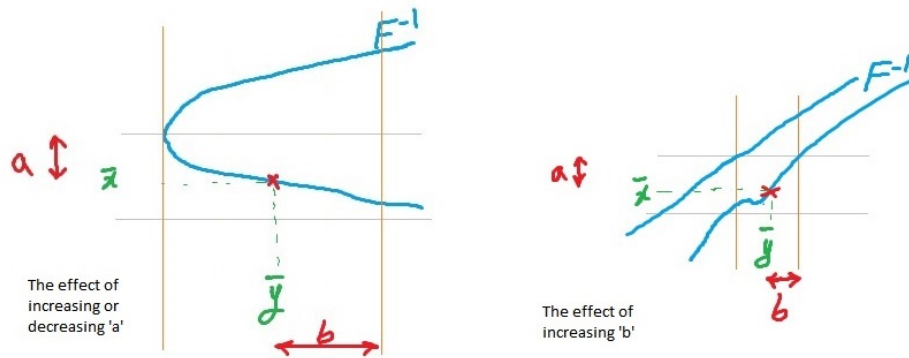


FIGURE 4.4: The side effects of changing radii independently

The previous lemma suggests that working with both radii a , and b at the same time is allowed (with a careful control, of course) in order to keep the desired properties of the localized map.

Theorem 4.2.10 (Uniform Strong Metric Regularity).

For the generalized equation (4.3), and the solution mapping (4.22), assume that

- (i) $z(\cdot)$ is a given continuous solution trajectory;

- (ii) $p(\cdot)$ is a continuous function;
 (iii) G_t is pointwise strongly metrically regular at $z(t)$ for 0.

Then there exist constants $a, b, \kappa > 0$ such that for any $t \in [0, 1]$, the mapping

$$\mathbb{B}_b(0) \ni y \longmapsto G_t^{-1}(y) \cap \mathbb{B}_a(z(t))$$

is single valued and Lipschitz continuous with a Lipschitz constant κ .

Proof. We prove the statement in two steps. First, by showing the mentioned map must be single-valued without caring about the Lipschitz regularity, and then by proving it is a Lipschitz continuous function.

STEP 1. Single-valuedness:

We show that there exist $a, b > 0$ such that for any $t \in [0, 1]$, the map

$$\mathbb{B}_b(0) \ni y \longmapsto G_t^{-1}(y) \cap \mathbb{B}_a(z(t)) \quad (4.28)$$

is single-valued. We argue by contradiction, by assuming that for any $a, b > 0$, there exists $t_{a,b} \in [0, 1]$ such that (4.28) does not hold. In particular, take $a_n = \frac{1}{n}, b_n = \frac{1}{n^3}$; then, for every $n \in \mathbb{N}$, there exists $t_n (:= t_{a_n, b_n}) \in [0, 1]$ such that

$$\mathbb{B}_{b_n}(0) \ni y \longmapsto G_{t_n}^{-1}(y) \cap \mathbb{B}_{a_n}(z(t_n)) \quad (4.29)$$

is not single-valued, which is equivalent to

Case 1. there exists $y_n \in \mathbb{B}_{b_n}(0)$ such that the cardinality of the set $G_{t_n}^{-1}(y_n) \cap \mathbb{B}_{a_n}(z(t_n))$ is grater than one, or

Case 2. there exists $y'_n \in \mathbb{B}_{b_n}(0)$ such that the set $G_{t_n}^{-1}(y'_n) \cap \mathbb{B}_{a_n}(z(t_n))$ is empty¹³.

By replacing (t_n) with a subsequence (if necessary), from the compactness of $[0, 1]$ in \mathbb{R} , we can assume that $t_n \rightarrow t_0 \in [0, 1]$. We will try to reach a contradiction in each case.

Case 1. Multi-valuedness

Suppose that, for any $n \in \mathbb{N}$, there exist $t_n \in [0, 1]$ and at least a $y_n \in \mathbb{B}_{b_n}(0)$ such that $z_n^1, z_n^2 \in G_{t_n}^{-1}(y_n) \cap \mathbb{B}_{a_n}(z(t_n))$ with $z_n^1 \neq z_n^2$.

By assumption (iii), there exist constants $a_{t_0}, b_{t_0}, \kappa_{t_0} > 0$ such that the mapping

$$\mathbb{B}_{b_{t_0}}(0) \ni w \longmapsto G_{t_0}^{-1}(w) \cap \mathbb{B}_{a_{t_0}}(z(t_0))$$

¹³ In other words, the mapping

$$y \longmapsto G_{t_n}^{-1}(y) \cap \mathbb{B}_{a_n}(z(t_n))$$

for at least a point $y \in \mathbb{B}_{b_n}(0)$, is empty, or multivalued, that is, it has at least two values.

is single valued and Lipschitz continuous with Lipschitz constant κ_{t_0} .
Make $b_{t_0} > 0$ smaller if necessary so that

$$\kappa_{t_0} b_{t_0} \leq a_{t_0}. \quad (4.30)$$

For n large enough (i.e. $n > N_0 \in \mathbb{N}$), one can have the following:

$$b_n < \frac{b_{t_0}}{2}, \quad \|p(t_n) - p(t_0)\| < \frac{b_{t_0}}{2}, \quad \|z(t_n) - z(t_0)\| < \frac{a_{t_0}}{2}, \quad a_n < \frac{a_{t_0}}{2}, \quad \kappa_{t_0} < n^2, \quad (4.31)$$

in which the second and third inequalities are the results of continuity assumptions of $p(\cdot)$ and $z(\cdot)$, respectively. Then,

$$\|z_n^1 - z(t_0)\| \leq \|z_n^1 - z(t_n)\| + \|z(t_n) - z(t_0)\| \leq a_n + \frac{a_{t_0}}{2} < a_{t_0}.$$

The same holds for z_n^2 ; thus, $z_n^1, z_n^2 \in \mathbb{B}_{a_{t_0}}(z(t_0))$. On the other hand, $z_n^i \in G_{t_n}^{-1}(y_n)$ for $i = 1, 2$, and Lemma 4.2.3 implies that $z_n^i \in G_{t_0}^{-1}(y_n + p(t_n) - p(t_0))$. But

$$\|y_n + p(t_n) - p(t_0)\| \leq \|y_n - 0\| + \|p(t_n) - p(t_0)\| \leq b_n + \frac{b_{t_0}}{2} < b_{t_0}.$$

Thus, $(y_n + p(t_n) - p(t_0)) \in \mathbb{B}_{b_{t_0}}(0)$, which is a contradiction since, in that neighborhood, $G_{t_0}^{-1}(\cdot) \cap \mathbb{B}_{a_{t_0}}(z(t_0))$ is single-valued.

Case 2. Emptiness

Let us now suppose that, for any $n \in \mathbb{N}$, there exist $t_n \in [0, 1]$ and at least a point $y'_n \in \mathbb{B}_{b_n}(0)$ such that $G_{t_n}^{-1}(y'_n) \cap \mathbb{B}_{a_n}(z(t_n))$ is empty.

For n large enough, the inequalities in (4.30) and (4.31) hold, and we have already proved that $y \in \mathbb{B}_{b_n}(0)$ implies $y + p(t_n) - p(t_0) \in \mathbb{B}_{b_{t_0}}(0)$. Therefore, since $y'_n \in \mathbb{B}_{b_n}(0)$, the mapping $G_{t_0}^{-1}(y'_n + p(t_n) - p(t_0)) \cap \mathbb{B}_{a_{t_0}}(z(t_0))$ is single-valued. In particular, it implies that $G_{t_0}^{-1}(y'_n + p(t_n) - p(t_0)) \neq \emptyset$. Let z be a point in $G_{t_0}^{-1}(y'_n + p(t_n) - p(t_0))$. Then, by using Lemma 4.2.3, we obtain $y'_n \in G_{t_n}(z)$, in particular, $G_{t_n}^{-1}(y'_n)$ is not empty. The contradiction assumption implies that

$$\|z - z(t_n)\| > a_n. \quad (4.32)$$

We will show the inconsistency between the contradiction assumption and the assumptions of the theorem with this inequality. In order to proceed, let us first prove that the mapping

$$\mathbb{B}_{\frac{b_{t_0}}{2}}(0) \ni y \longmapsto G_{t_n}^{-1}(y) \cap \mathbb{B}_{\frac{a_{t_0}}{2}}(z(t_n)) \quad (4.33)$$

is single-valued and Lipschitz continuous with Lipschitz constant κ_{t_0} .

As a matter of fact, we have already seen that $G_{t_n}^{-1}(y) \neq \emptyset$ for every $y \in \mathbb{B}_{b_n}(0)$, and $G_{t_n}^{-1}(y) \cap \mathbb{B}_{a_{t_0}/2}(z(t_n))$ is not multi-valued. Thus, it only remains to show that $G_{t_n}^{-1}(y) \cap \mathbb{B}_{a_{t_0}/2}(z(t_n)) \neq \emptyset$ for every $y \in \mathbb{B}_{b_{t_0}/2}(0)$.

Denote by z_y the point $z_y := G_{t_0}^{-1}(y + p(t_n) - p(t_0)) \cap \mathbb{B}_{a_{t_0}}(z(t_0))$. First observe that, by Lemma 4.2.3, $z_y \in G_{t_n}^{-1}(y)$.

On the other hand, by definition, $z(t_n) \in G_{t_n}^{-1}(0)$ and by using Lemma 4.2.3, we get $z(t_n) \in G_{t_0}^{-1}(p(t_n) - p(t_0))$. We also know that $z(t_n) \in \mathbb{B}_{a_{t_0}}(z(t_0))$ (from the inequalities in (4.31)). The single-valuedness of $G_{t_0}^{-1}(\cdot) \cap \mathbb{B}_{a_{t_0}}(z(t_0))$ over $\mathbb{B}_{b_{t_0}}(0)$ allows us to write $z(t_n) = G_{t_0}^{-1}(p(t_n) - p(t_0)) \cap \mathbb{B}_{a_{t_0}}(z(t_0))$ without ambiguity. Thus, we have:

$$\begin{aligned} \|z_y - z(t_n)\| &= \|[G_{t_0}^{-1}(y + p(t_n) - p(t_0)) \cap \mathbb{B}_{a_{t_0}}(z(t_0))] - [G_{t_0}^{-1}(p(t_n) - p(t_0)) \cap \mathbb{B}_{a_{t_0}}(z(t_0))]\| \\ &\leq \kappa_{t_0} \|y + p(t_n) - p(t_0) - (p(t_n) - p(t_0))\| \\ &\leq \kappa_{t_0} \|y - 0\| \\ &\leq \kappa_{t_0} \frac{b_{t_0}}{2} \leq \frac{1}{2} a_{t_0}. \end{aligned}$$

Which means $z_y \in G_{t_n}^{-1}(y'_n) \cap \mathbb{B}_{\frac{a_{t_0}}{2}}(z(t_n))$.

A similar reasoning reveals the Lipschitz continuity of the map $G_{t_n}^{-1}(\cdot) \cap \mathbb{B}_{\frac{a_{t_0}}{2}}(z(t_n))$.

Indeed, taking any $y_i \in \mathbb{B}_{b_{t_0}/2}(0)$, one can define $z_i := G_{t_n}^{-1}(y_i) \cap \mathbb{B}_{a_{t_0}/2}(z(t_n))$ for $i = 1, 2$ without ambiguity. Using the second and third inequalities in (4.31), we have

$$\begin{aligned} z_i \in G_{t_n}^{-1}(y_i) &= G_{t_0}^{-1}(y_i + p(t_n) - p(t_0)) \quad \text{and} \quad y_i + p(t_n) - p(t_0) \in \mathbb{B}_{b_{t_0}}(0) \\ z_i \in \mathbb{B}_{a_{t_0}/2}(z(t_n)), \quad \text{and} \quad \|z(t_n) - z(t_0)\| &< a_{t_0}/2, \quad \text{so} \quad z_i \in \mathbb{B}_{a_{t_0}}(z(t_0)). \end{aligned}$$

Thus we are allowed to use the single-valuedness and Lipschitz property of $G_{t_0}^{-1}$ to obtain

$$\begin{aligned} \|z_1 - z_2\| &= \|[G_{t_n}^{-1}(y_1) \cap \mathbb{B}_{a_{t_0}/2}(z(t_n))] - [G_{t_n}^{-1}(y_2) \cap \mathbb{B}_{a_{t_0}/2}(z(t_n))]\| \\ &= \|[G_{t_0}^{-1}(y_1 + p(t_n) - p(t_0)) \cap \mathbb{B}_{a_{t_0}}(z(t_0))] - [G_{t_0}^{-1}(y_2 + p(t_n) - p(t_0)) \cap \mathbb{B}_{a_{t_0}}(z(t_0))]\| \\ &\leq \kappa_{t_0} \|y_1 - y_2\|, \end{aligned}$$

which reveals the Lipschitz property of the map in (4.33).

Now, having the strong metric regularity of $G_{t_n}(\cdot)$ with constants $\frac{a_{t_0}}{2}, \frac{b_{t_0}}{2}, \kappa_{t_0}$ in hand, we use Lemma 4.2.8 with $a' = a_n = \frac{1}{n} \leq \frac{a_{t_0}}{2}, b' = b_n = \frac{1}{n^3} \leq \frac{b_{t_0}}{2}$, to obtain the strong metric regularity of $G_{t_n}(\cdot)$ with constants a_n, b_n, κ_{t_0} (reminding that the last inequality of (4.31) guarantees $\kappa_{t_0} b' \leq a'$). Now for the specific $y'_n \in \mathbb{B}_{b_n}(0)$ claimed before, there exists $z \in G_{t_n}^{-1}(y'_n) \cap \mathbb{B}_{a_n}(z(t_n))$ which contradicts (4.32).

Therefore, till now we have proved that there exist $a^*, b^* > 0$ such that the mapping

$$\mathbb{B}_{b^*}(0) \ni y \longmapsto G_t^{-1}(y) \cap \mathbb{B}_{a^*}(z(t))$$

is single-valued for any $t \in [0, 1]$.

STEP 2. Lipschitz Continuity:

Being sure that we deal with a function, now we proceed by claiming that there exist $b \leq b^*$, and $\kappa > 0$ such that the mapping

$$\mathbb{B}_b(0) \ni y \longmapsto G_t^{-1}(y) \cap \mathbb{B}_{a^*}(z(t))$$

is Lipschitz continuous with Lipschitz constant κ for all $t \in [0, 1]$.

We will prove the assertion by contradiction. Suppose the claim is false; then, for any $b \leq b^*$, and any $\kappa > 0$, there exists $t_{b,\kappa} \in [0, 1]$ such that the mapping

$$\mathbb{B}_b(0) \ni y \longmapsto G_{t_{b,\kappa}}^{-1}(y) \cap \mathbb{B}_{a^*}(z(t_{b,\kappa}))$$

is not Lipschitz with constant κ . Since we already know that this map is single-valued, due to the explanations in Remark 4.2.9, the contradiction assumption yields that for every $\kappa > 0$, there exist $y_1, y_2 \in \mathbb{B}_b(0)$, with $y_1 \neq y_2$ such that

$$\left\| \left[G_{t_{b,\kappa}}^{-1}(y_1) \cap \mathbb{B}_{a^*}(z(t_{b,\kappa})) \right] - \left[G_{t_{b,\kappa}}^{-1}(y_2) \cap \mathbb{B}_{a^*}(z(t_{b,\kappa})) \right] \right\| > \kappa \|y_1 - y_2\|.$$

For any $n \in \mathbb{N}$, let $b_n := \min\{\frac{1}{n^3}, b^*\}$, $\kappa_n := n$ and set $t_n := t_{b_n, \kappa_n}$. Then there exist at least $y_n^1, y_n^2 \in \mathbb{B}_{b_n}(0)$, with $y_n^1 \neq y_n^2$ such that

$$\|y_n^1 - y_n^2\| n < \left\| \left[G_{t_n}^{-1}(y_n^1) \cap \mathbb{B}_{a^*}(z(t_n)) \right] - \left[G_{t_n}^{-1}(y_n^2) \cap \mathbb{B}_{a^*}(z(t_n)) \right] \right\|$$

Let $g_n^i := G_{t_n}^{-1}(y_n^i) \cap \mathbb{B}_{a^*}(z(t_n))$ for $i = 1, 2$, and assume that t_n converges to a point, say t_0 .

For n large enough, one has the following:

$$b_n < \frac{b_{t_0}}{2}, \quad \|p(t_n) - p(t_0)\| < \frac{b_{t_0}}{2}, \quad \|z(t_n) - z(t_0)\| < \frac{a^*}{2}, \quad \kappa_{t_0} < n^2.$$

On the one hand, $g_n^i \in G_{t_n}^{-1}(y_n^i) = G_{t_0}^{-1}(y_n^i + p(t_n) - p(t_0))$ and the above conditions imply that $(y_n^i + p(t_n) - p(t_0)) \in \mathbb{B}_{b_{t_0}}(0)$.

On the other hand, $g_n^i \in \mathbb{B}_{a^*}(z(t_n))$. We will show that $g_n^i \in B_{a_{t_0}}(z(t_0))$.

Indeed, let $w_n^i := G_{t_0}^{-1}(y_n^i + p(t_n) - p(t_0)) \cap B_{a_{t_0}}(z(t_0))$ for $i = 1, 2$.

Since $(y_n^i + p(t_n) - p(t_0)) \rightarrow 0$, by the continuity of $G_{t_0}^{-1}(\cdot) \cap B_{a_{t_0}}(z(t_0))$ around 0, we get

$$w_n^i \longrightarrow z(t_0) = G_{t_0}^{-1}(0) \cap B_{a_{t_0}}(z(t_0)).$$

Thus, for any $\epsilon > 0$, there exists $N_\epsilon \in \mathbb{N}$ such that, for $n > N_\epsilon$, one has $\|w_n^i - z(t_0)\| < \epsilon$. Let $\epsilon = a^*/2$. Then,

$$\begin{aligned} \|w_n^i - z(t_n)\| &\leq \|w_n^i - z(t_0)\| + \|z(t_0) - z(t_n)\| \\ &< a^*/2 + a^*/2 \\ &< a^*, \end{aligned}$$

which means that $w_n^i \in \mathbb{B}_{a^*}(z(t_n))$. Combining with $w_n^i \in G_{t_n}^{-1}(y_n^i)$ (obtained by using Lemma 4.2.3), we get that $w_n^i \in G_{t_n}^{-1}(y_n^i) \cap \mathbb{B}_{a^*}(z(t_n))$. Hence, by the single-valuedness of $G_{t_n}^{-1}(\cdot) \cap \mathbb{B}_{a^*}(z(t_n))$, we can conclude that $w_n^i = g_n^i \in B_{a_{t_0}}(z(t_0))$.

Then, the assumption (iii) of the theorem results in $\|g_n^1 - g_n^2\| \leq \kappa_{t_0} \|y_n^1 - y_n^2\|$. So

$$\|y_n^1 - y_n^2\| n < \|g_n^1 - g_n^2\| \leq \kappa_{t_0} \|y_n^1 - y_n^2\|,$$

which is a contradiction, since it implies boundedness of the sequence $(\kappa_n) := (n)$. Combining the two steps ends the proof. \square

Remark 4.2.11. A slightly different version of this Theorem could be stated and proved as follows. Under stronger assumptions on $p(\cdot)$ and $z(\cdot)$, a simpler and more direct proof can be provided. The proof is in the direction of Theorem 4.1.2, except that the special structure of the single-valued part here allows us to bypass the use of Theorem 1.4.14.

Theorem 4.2.12 (Uniform Strong Metric Regularity).

For the generalized equation (4.3), and the solution mapping (4.22), assume that

- (i) $z(\cdot)$ is a Lipschitz continuous solution trajectory with Lipschitz constant L_z ;
- (ii) $p(\cdot)$ is a Lipschitz continuous function with Lipschitz constant L_p ;
- (iii) G_t is pointwise strongly metrically regular at $z(t)$ for 0.

Then there exist constants $a, b, \kappa > 0$ such that the mapping

$$\mathbb{B}_b(0) \ni y \mapsto G_t^{-1}(y) \cap \mathbb{B}_a(z(t))$$

is single valued and Lipschitz continuous with a Lipschitz constant κ for any $t \in [0, 1]$.

Proof. For an arbitrary point $t_0 \in [0, 1]$, by assumption (iii), there exist constants $a_{t_0}, b_{t_0}, \kappa_{t_0} > 0$ such that the mapping

$$\mathbb{B}_{b_{t_0}}(0) \ni y \mapsto G_{t_0}^{-1}(y) \cap \mathbb{B}_{a_{t_0}}(z(t_0))$$

is single-valued and Lipschitz continuous with Lipschitz constant κ_{t_0} . Choose b_{t_0} smaller if necessary, so that $b_{t_0} \kappa_{t_0} < a_{t_0}$.

Fix $\rho_{t_0} > 0$ such that the following conditions are satisfied:

$$L_z \rho_{t_0} < a_{t_0}/2, \quad L_p \rho_{t_0} < b_{t_0}/2 \quad (4.34)$$

Now, for any $\tau \in (t_0 - \rho_{t_0}, t_0 + \rho_{t_0}) \cap [0, 1]$, we claim that the mapping

$$\mathbb{B}_{b_{t_0}/2}(0) \ni y \longmapsto G_\tau^{-1}(y) \cap \mathbb{B}_{a_{t_0}/2}(z(\tau)) \quad (4.35)$$

is single-valued and Lipschitz continuous with Lipschitz constant κ_{t_0} .

The proof will include the following simple steps:

1. the set $G_\tau^{-1}(y)$ is not empty for any $y \in \mathbb{B}_{b_{t_0}/2}(0)$;
 2. the mapping (4.35) is not multivalued;
 3. the set $G_\tau^{-1}(y) \cap \mathbb{B}_{a_{t_0}/2}(z(\tau))$ is a singleton for any $y \in \mathbb{B}_{b_{t_0}/2}(0)$;
 4. the mapping (4.35) is a Lipschitz continuous function (with constant κ_{t_0}).
1. Since $\|p(\tau) - p(t_0)\| \leq L_p \|\tau - t_0\| < L_p \rho_{t_0} < b_{t_0}/2$, for any $y \in \mathbb{B}_{b_{t_0}/2}(0)$, one gets $y + p(\tau) - p(t_0) \in \mathbb{B}_{b_{t_0}}(0)$; so $G_{t_0}^{-1}(y + p(\tau) - p(t_0)) \neq \emptyset$. Thus, from Lemma 4.2.3 one concludes that $G_\tau^{-1}(y) \neq \emptyset$.
 2. Let $u_1, u_2 \in G_\tau^{-1}(y) \cap \mathbb{B}_{a_{t_0}/2}(z(\tau))$. Observe that $u_i \in G_{t_0}^{-1}(y + p(\tau) - p(t_0))$ for $i = 1, 2$, by Lemma 4.2.3. Since Lipschitz continuity of $z(\cdot)$ implies that $\mathbb{B}_{a_{t_0}/2}(z(\tau))$ is a subset of $\mathbb{B}_{a_{t_0}}(z(t_0))$, one obtains that $u_i \in G_{t_0}^{-1}(y + p(\tau) - p(t_0)) \cap \mathbb{B}_{a_{t_0}/2}(z(t_0))$ for $i = 1, 2$. Thus, assumption (iii) implies $u_1 = u_2$.
 3. For any $w_0 \in \mathbb{B}_{b_{t_0}/2}(0)$, let $u_0 := G_{t_0}^{-1}(w_0 + p(\tau) - p(t_0)) \cap \mathbb{B}_{a_{t_0}}(z(t_0))$. By definition, $z(\tau) \in G_\tau^{-1}(0)$, so from (4.34) and Lemma 4.2.3 one can obtain that $z(\tau) \in G_{t_0}^{-1}(0 + p(\tau) - p(t_0)) \cap \mathbb{B}_{a_{t_0}}(z(t_0))$. Then,

$$\begin{aligned} \|u_0 - z(\tau)\| &= \|[G_{t_0}^{-1}(w_0 + p(\tau) - p(t_0)) \cap \mathbb{B}_{a_{t_0}}(z(t_0))] - [G_{t_0}^{-1}(0 + p(\tau) - p(t_0)) \cap \mathbb{B}_{a_{t_0}}(z(t_0))]\| \\ &\leq \kappa_{t_0} \|w_0 + p(\tau) - p(t_0) - (0 + p(\tau) - p(t_0))\| \\ &< \kappa_{t_0} \frac{b_{t_0}}{2} < \frac{a_{t_0}}{2}. \end{aligned}$$

Thus, $u_0 \in G_\tau^{-1}(w_0) \cap \mathbb{B}_{a_{t_0}/2}(z(\tau))$. In fact, by considering the result of the previous part of the proof we obtain $u_0 = G_\tau^{-1}(w_0) \cap \mathbb{B}_{a_{t_0}/2}(z(\tau))$.

4. Consider any $w_1, w_2 \in \mathbb{B}_{b_{t_0}/2}(0)$, and let $u_i = G_\tau^{-1}(w_i) \cap \mathbb{B}_{a_{t_0}/2}(z(\tau))$ for $i = 1, 2$. Then, by using Lemma 4.2.3 we can conclude that

$$\begin{aligned} \|u_1 - u_2\| &= \left\| [G_\tau^{-1}(w_1) \cap \mathbb{B}_{a_{t_0}/2}(z(\tau))] - [G_\tau^{-1}(w_2) \cap \mathbb{B}_{a_{t_0}/2}(z(\tau))] \right\| \\ &= \left\| [G_{t_0}^{-1}(w_1 + p(\tau) - p(t_0)) \cap \mathbb{B}_{a_{t_0}/2}(z(\tau))] - [G_{t_0}^{-1}(w_2 + p(\tau) - p(t_0)) \cap \mathbb{B}_{a_{t_0}/2}(z(\tau))] \right\| \\ &\leq \kappa_{t_0} \|w_1 + p(\tau) - p(t_0) - (w_2 + p(\tau) - p(t_0))\| \\ &\leq \kappa_{t_0} \|w_1 - w_2\|. \end{aligned}$$

Now that we have local uniformity, choose a finite subcover $\bigcup_{i=1}^m (t_i - \rho_{t_i}, t_i + \rho_{t_i})$ from the open covering of the compact interval $[0, 1]$. Let

$$\begin{aligned} a &:= \min\left\{\frac{a_{t_i}}{2} \mid i = 1, \dots, m\right\}, \\ \kappa &:= \max\{\kappa_{t_i} \mid i = 1, \dots, m\}, \\ \text{and } b &:= \min\left\{\frac{a}{\kappa}, \min\left\{\frac{b_{t_i}}{2} \mid i = 1, \dots, m\right\}\right\}. \end{aligned} \tag{4.36}$$

For any interval $(t_i - \rho_{t_i}, t_i + \rho_{t_i})$, we have the strong metric regularity with constants $\frac{a_{t_i}}{2}$, $\frac{b_{t_i}}{2}$, and κ . From (4.36) and Lemma 4.2.8, we obtain the strong metric regularity with constants a, b , and κ , which does not depend on t_i any more. Thus, the proof is complete. \square

4.2.3 Perturbations of the Input Signal

In this subsection we try to take into account the small variations of the function $p(\cdot)$. More precisely, we consider a continuous function $\tilde{p}(\cdot)$ such that $\|\tilde{p}(t) - p(t)\| < \epsilon$ for any $t \in [0, 1]$, and for a suitably small $\epsilon > 0$. The reason of such perturbations and the importance of this study in the case of electronic circuits was already discussed in Chapter 2 (cf. Section 2.2).

Unlike the static case, where p was a fixed vector in \mathbb{R}^n , since we changed the notation of solution mapping, the problem does not reduce to the study of the stability properties of $S(t)$; however, we can take advantage of those results by considering the fact that in the newly defined set-valued map G_t the variable t reduces to a parameter.

In this subsection we deal with the perturbed form of problem (4.3). To be more specific, we consider the generalized equation

$$f(z) - \tilde{p}(t) + F(z) \ni 0, \tag{4.37}$$

denote the corresponding solution mapping with \tilde{S} ,

$$\tilde{S} : t \mapsto \tilde{S}(t) = \{z \in \mathbb{R}^n \mid f(z) - \tilde{p}(t) + F(z) \ni 0\}, \quad (4.38)$$

and define the auxiliary map $\tilde{G}_t : \mathbb{R}^n \rightrightarrows \mathbb{R}^n$ as $\tilde{G}_t(v) = f(v) + F(v) - \tilde{p}(t)$.

The easy-to-check equalities

$$\boxed{\tilde{G}_t(v) = G_t(v) + p(t) - \tilde{p}(t)} \quad (4.39)$$

$$\boxed{\tilde{G}_t^{-1}(w) = G_t^{-1}(w + \tilde{p}(t) - p(t))} \quad (4.40)$$

for each $t \in [0, 1]$, will be useful for connecting the strong metric regularity properties of \tilde{G}_t to those of G_t as described in the following lemma.

Once more, we want to indicate that the straightforward equalities (4.39) and (4.40) are a consequence of our choice of the auxiliary maps and the special form of the single-valued part of the generalized equation (4.37).

Lemma 4.2.13 (Perturbation Effect on the Auxiliary Map).

Assume that $p(\cdot)$ and $\tilde{p}(\cdot)$ are continuous functions from $[0, 1]$ to \mathbb{R}^n with $\|\tilde{p}(t) - p(t)\| < \epsilon$ for any $t \in [0, 1]$. If G_t is strongly metrically regular at \bar{u} for 0 (i.e. $(\bar{u}, 0) \in \text{gph } G_t$ and there exist constants $a_t, b_t, \kappa_t > 0$ such that the mapping

$$\mathbb{B}_{b_t}(0) \ni y \mapsto G_t^{-1}(y) \cap \mathbb{B}_{a_t}(\bar{u})$$

is single valued and Lipschitz continuous with Lipschitz constant κ_t), then for any positive $\epsilon < b_t$ the mapping

$$w \mapsto \tilde{G}_t^{-1}(w) \cap \mathbb{B}_{a_t}(\bar{u}) \quad (4.41)$$

is a Lipschitz continuous function on $\mathbb{B}_{b_t - \epsilon}(0)$ with Lipschitz constant κ_t .

Proof. The process of proof is very similar to those steps we provided for the mapping (4.35) in the proof of the previous theorem, except that here, the intersecting ball $\mathbb{B}_{a_t}(\bar{u})$ is the same for both maps G_t^{-1} and \tilde{G}_t^{-1} and this makes the proof much easier. The proof should include the following steps:

1. the sets $\tilde{G}_t^{-1}(w)$ and also $\tilde{G}_t^{-1}(w) \cap \mathbb{B}_{a_t}(\bar{u})$ are not empty for any $w \in \mathbb{B}_{b_t - \epsilon}(0)$;
2. the mapping (4.41) is not multivalued;
3. the mapping (4.41) is a Lipschitz continuous function (with constant κ_t).

Choose any $w_1, w_2 \in \mathbb{B}_{b_t - \epsilon}(0)$. From assumption we get $(w_i + \tilde{p}(t) - p(t)) \in \mathbb{B}_{b_t}(0)$ for $i = 1, 2$. Then, the pointwise strong metric regularity of G_t , lets us define $u_i := G_t^{-1}(w_i + \tilde{p}(t) - p(t)) \cap \mathbb{B}_{a_t}(\bar{u})$ for $i = 1, 2$. By using (4.40), one obtains $u_i \in \tilde{G}_t^{-1}(w_i)$. In fact, $u_i = \tilde{G}_t^{-1}(w_i) \cap \mathbb{B}_{a_t}(\bar{u})$. Thus, steps 1. and 2. are proved.

But pointwise strong metric regularity of G_t provides more information, that is

$$\|u_1 - u_2\| \leq \kappa_t \|w_1 - w_2\|.$$

Therefore, step 3. is also proved. \square

Remark 4.2.14. (a) A careful look at the proof reveals that the lemma could be also expressed in the following way:

If G_t is SMR at \bar{u} for 0, then \widetilde{G}_t is SMR at \bar{u} for $p(t) - \tilde{p}(t)$.

In this case, ϵ could be as big as b_t .

In fact, in this case one can consider Theorem 1.4.9 with $F = G_t$, $(\bar{x}, \bar{y}) = (\bar{u}, 0)$, $\kappa = \kappa_t$, and $g(\cdot) = p(t) - \tilde{p}(t)$ which is a constant function with respect to u , so is obviously Lipschitz with any $\mu < \kappa^{-1}$, and immediately obtain the SMR at \bar{u} for $p(t) - \tilde{p}(t)$ of the map $(G_t + p(t) - \tilde{p}(t))$ which is exactly \widetilde{G}_t .

(b) Under the assumptions of Theorem 4.2.10, we would have uniform strong metric regularity for G_t at $z(t)$ for 0 and the proof shows that we obtain uniform strong metric regularity for \widetilde{G}_t at $z(t)$ for $p(t) - \tilde{p}(t)$, too.

Finally, we have provided enough instruments to declare the main result of this section, that is the existence of a solution trajectory $\tilde{z}(\cdot)$ close to $z(\cdot)$ that inherits its continuity properties.

We may recall that, under the assumptions of Theorem 4.2.10, existence of positive constants a, b , and κ not depending on t is guaranteed for uniform strong metric regularity. Since the following theorem satisfies those assumptions, we will use the uniform constants without ambiguity.

Theorem 4.2.15 (Existence of a Continuous Trajectory for the Perturbed Problem). *For the generalized equations (4.3), and (4.37) and the corresponding solution mappings (4.22), and (4.38), assume that*

- (i) $z(\cdot)$ is a given continuous solution trajectory (for S);
- (ii) $p(\cdot)$ and $\tilde{p}(\cdot)$ are continuous functions such that for every $t \in [0, 1]$, $\|\tilde{p}(t) - p(t)\| < \epsilon$ (with $\epsilon < b/4$);
- (iii) G_t is pointwise strongly metrically regular at $z(t)$ for 0.

Then there exists a continuous solution trajectory $\tilde{z}(\cdot)$ for \widetilde{S} such that, for every $t \in [0, 1]$, we have $\|\tilde{z}(t) - z(t)\| < \frac{4a\epsilon}{b}$.

Proof. We will present two proofs for this theorem, both are constructional methods, yet with different approaches. Remark 4.2.16 after the proof, will provide a comparison between the methods.

Method 1. Pointwise construction:

Consider an arbitrary $t_0 \in [0, 1]$. Since $(t_0, z(t_0)) \in \text{gph } S$ and G_{t_0} is strongly metrically regular at $z(t_0)$ for 0, by using Theorem 4.2.10, we obtain that the mapping

$$\mathbb{B}_b(0) \ni y \longmapsto G_{t_0}^{-1}(y) \cap \mathbb{B}_a(z(t_0))$$

is single-valued and Lipschitz continuous with constant κ . Let $y_0 = \tilde{p}(t_0) - p(t_0)$. For ϵ small enough (i.e. $\epsilon < b/4$), we have $y_0 \in \mathbb{B}_b(0)$. Let

$$\tilde{z}(t_0) := G_{t_0}^{-1}(y_0) \cap \mathbb{B}_a(z(t_0)). \quad (4.42)$$

Note that the right-hand side of this expression is a singleton and so $\tilde{z}(t_0)$ is exactly determined without ambiguity. Let us check if $(t_0, \tilde{z}(t_0)) \in \text{gph } \tilde{S}$ or, equivalently, $0 \in \widetilde{G_{t_0}}(\tilde{z}(t_0))$.

From the definition of $\tilde{z}(t_0)$ we have $y_0 \in G_{t_0}(\tilde{z}(t_0)) = f(\tilde{z}(t_0)) + F(\tilde{z}(t_0)) - p(t_0)$. Then, from $y_0 = \tilde{p}(t_0) - p(t_0)$, one gets $\tilde{p}(t_0) \in f(\tilde{z}(t_0)) + F(\tilde{z}(t_0))$ or $0 \in \widetilde{G_{t_0}}(\tilde{z}(t_0))$. Since t_0 is an arbitrary point in $[0, 1]$,

$$t \longmapsto G_t^{-1}(\tilde{p}(t) - p(t)) \cap \mathbb{B}_a(z(t))$$

defines a single-valued map $\tilde{z}(\cdot)$.

To prove the continuity, consider a sequence $(t_n) \in [0, 1]$ converging to t_0 . By continuity of $\tilde{p}(\cdot)$ and $p(\cdot)$ we know that $y_n := \tilde{p}(t_n) - p(t_n) \longrightarrow y_0 = \tilde{p}(t_0) - p(t_0)$.

By definition, $\tilde{z}(t_n) := G_{t_n}^{-1}(y_n) \cap \mathbb{B}_a(z(t_n))$. Lemma 4.2.3 yields that

$$\tilde{z}(t_n) \in G_{t_0}^{-1}(y_n + p(t_n) - p(t_0)).$$

On the other hand, $\tilde{z}(t_n) \in \mathbb{B}_a(z(t_n))$. We claim that $\tilde{z}(t_n) \in \mathbb{B}_a(z(t_0))$.

Indeed, for n large enough, one can have the following:

$$\|p(t_n) - p(t_0)\| < \frac{b}{4}, \quad \|z(t_n) - z(t_0)\| < \frac{a}{2}. \quad (4.43)$$

Considering Lemma 4.2.8, with $b' = b/2$, and $a' = a/2$, we obtain that the mapping

$$\mathbb{B}_{b/2}(0) \ni y \longmapsto G_{t_0}^{-1}(y) \cap \mathbb{B}_{a/2}(z(t_0))$$

is single-valued and Lipschitz continuous with Lipschitz constant κ . Now observing that

$$\begin{aligned} \|y_n + p(t_n) - p(t_0)\| &\leq \|\tilde{p}(t_n) - p(t_n)\| + \|p(t_n) - p(t_0)\| \\ &< \epsilon + \frac{b}{4} < \frac{b}{2}, \end{aligned}$$

we can define $w_n := G_{t_0}^{-1}(y_n + p(t_n) - p(t_0)) \cap \mathbb{B}_{a/2}(z(t_0))$ without ambiguity.

On the one hand, Lemma 4.2.3 implies that $w_n \in G_{t_n}^{-1}(y_n)$.

On the other hand, $w_n \in \mathbb{B}_{a/2}(z(t_0)) \subset \mathbb{B}_a(z(t_n))$. Thus, $w_n \in G_{t_n}^{-1}(y_n) \cap \mathbb{B}_a(z(t_n))$.

Since $G_{t_n}^{-1}$ is single-valued and Lipschitz continuous when restricted to $\mathbb{B}_b(0) \times \mathbb{B}_a(z(t_n))$, we obtain the equality $w_n = G_{t_n}^{-1}(y_n) \cap \mathbb{B}_a(z(t_n))$, and thus, $w_n = \tilde{z}(t_n)$. Therefore, $\tilde{z}(t_n) \in \mathbb{B}_{a/2}(z(t_0)) \subset \mathbb{B}_a(z(t_0))$.

The strong metric regularity of G_{t_0} implies that

$$\|\tilde{z}(t_n) - \tilde{z}(t_0)\| \leq \kappa \|(y_n + p(t_n) - p(t_0)) - y_0\|.$$

Hence, $\|\tilde{z}(t_n) - \tilde{z}(t_0)\|$ converges to zero as $n \rightarrow \infty$.

It only remains to remind that the estimate for the difference $\|\tilde{z}(t) - z(t)\|$ is a straightforward consequence of the way we constructed \tilde{z} . Indeed, let $r(\epsilon) := \sup_{t \in [0,1]} \|\tilde{z}(t) - z(t)\|$.

Starting from $\epsilon < \frac{b}{4}$, we obtained $r(\epsilon) < a$. If we let $\epsilon < \frac{b}{8}$, a deeper look into the proof reveals that we can obtain $r(\epsilon) < \frac{a}{2}$, and so on. Thus, the distance $\|\tilde{z}(t) - z(t)\|$ (for every $t \in [0, 1]$) is controlled linearly by ϵ and the proof is complete.

Method 2. Construction over an interval:

Fix $t \in [0, 1]$, and let b smaller if necessary such that $\kappa b < a$. By Remark 4.2.9, this will not affect the uniform strong metric regularity of G_t guaranteed by the assumptions of this theorem and Theorem 4.2.10. The uniform continuity of $\tilde{p}(\cdot)$, and $z(\cdot)$ allows us to choose $\rho > 0$ sufficiently small and independent of t , such that for any $\tau \in (t - \rho, t + \rho)$, the following hold:

$$\|\tilde{p}(\tau) - \tilde{p}(t)\| < \frac{b}{4}, \quad \|z(\tau) - z(t)\| < \frac{a}{2}. \quad (4.44)$$

Then, for any $\tau \in (t - \rho, t + \rho)$ the continuity of $\tilde{p}(\cdot)$ and its closeness to $p(\cdot)$ implies that

$$\|\tilde{p}(\tau) - p(t)\| \leq \|\tilde{p}(\tau) - \tilde{p}(t)\| + \|\tilde{p}(t) - p(t)\| < \frac{b}{2},$$

and therefore, by using Lemma 4.2.8 with $b' = \frac{b}{2}$, and $a' = \frac{a}{2}$, we obtain that the set $G_t^{-1}(\tilde{p}(\tau) - p(t)) \cap \mathbb{B}_{\frac{a}{2}}(z(t))$ is a singleton. Thus, we can define

$$\tilde{z}(\tau) := G_t^{-1}(\tilde{p}(\tau) - p(t)) \cap \mathbb{B}_{\frac{a}{2}}(z(t)) \quad \text{for any } \tau \in (t - \rho, t + \rho), \quad (4.45)$$

without ambiguity. In order to prove the continuity of this function, consider a sequence (τ_n) in $(t - \rho, t + \rho)$ such that $\tau_n \rightarrow \tau$. Then, from the Lipschitz continuity of $G_t^{-1}(\cdot) \cap$

$\mathbb{B}_a(z(t))$ over $\mathbb{B}_b(0)$ we obtain that

$$\begin{aligned} \|\tilde{z}(\tau_n) - \tilde{z}(\tau)\| &= \left\| \left[G_t^{-1}(\tilde{p}(\tau_n) - p(t)) \cap \mathbb{B}_a(z(t)) \right] - \left[G_t^{-1}(\tilde{p}(\tau) - p(t)) \cap \mathbb{B}_a(z(t)) \right] \right\| \\ &\leq \kappa \|\tilde{p}(\tau_n) - p(t) - (\tilde{p}(\tau) - p(t))\| \\ &\leq \kappa \|\tilde{p}(\tau_n) - \tilde{p}(\tau)\|. \end{aligned}$$

The continuity of \tilde{p} implies that $\|\tilde{z}(\tau_n) - \tilde{z}(\tau)\| \rightarrow 0$ as $\tau_n \rightarrow \tau$.

It remains to show that \tilde{z} is (part of) a solution trajectory, that is, $(\tau, \tilde{z}(\tau)) \in \text{gph } \tilde{S}$.

Since $\tilde{z}(\tau) \in G_t^{-1}(\tilde{p}(\tau) - p(t))$, from Lemma 4.2.3 we get $\tilde{z}(\tau) \in G_\tau^{-1}(\tilde{p}(\tau) - p(\tau))$. Then, Equation (4.40) implies that $\tilde{z}(\tau) \in \widetilde{G}_\tau^{-1}(0)$ for any $\tau \in (t - \rho, t + \rho)$, or equivalently, $\tilde{z}(\tau) \in \tilde{S}(\tau)$.

Up to now, we have proved that for each $t \in [0, 1]$, we can find a solution trajectory in the interval $(t - \rho, t + \rho)$. It remains to show that this construction over different intervals remains consistent. To be more clear, let us consider two points t_1 , and t_2 , with corresponding trajectory pieces \tilde{z}_1 , and \tilde{z}_2 . Suppose that $t_1 < t_2$ and let us consider the situation where $t_2 - \rho < \tau < t_1 + \rho$. We should prove that $\tilde{z}_1(\tau) = \tilde{z}_2(\tau)$.

By definition, $\tilde{z}_i(\tau) = G_{t_i}^{-1}(\tilde{p}(\tau) - p(t_i)) \cap \mathbb{B}_a(z(t_i))$ for $i = 1, 2$, and as already shown, Lemma 4.2.3, and Equality (4.40) imply that $\tilde{z}_i(\tau) \in \widetilde{G}_\tau^{-1}(0)$ for $i = 1, 2$.

On the other hand, the continuity of $z(\cdot)$, and inequalities in (4.44) reveal that

$$\|z(\tau) - \tilde{z}_i(\tau)\| \leq \|z(\tau) - z(t_i)\| + \|z(t_i) - \tilde{z}_i(\tau)\| < \frac{a}{2} + \frac{a}{2}.$$

Thus, $\tilde{z}_i(\tau) \in \mathbb{B}_a(z(\tau))$ for $i = 1, 2$. Using Lemma 4.2.13 and Remark 4.2.14 for \widetilde{G}_τ , we obtain that the mapping $\widetilde{G}_\tau^{-1}(\cdot) \cap \mathbb{B}_a(z(\tau))$ is single-valued and Lipschitz continuous over $\mathbb{B}_{\frac{3b}{4}}(0)$. So, $\tilde{z}_1(\tau) = \tilde{z}_2(\tau)$, and the proof is complete. \square

Remark 4.2.16. (a) A through observation reveals that in fact, the two methods produce the same function mainly because of the single-valuedness of the mapping

$$\mathbb{B}_b(0) \ni y \mapsto G_t^{-1}(y) \cap \mathbb{B}_a(z(t)).$$

To be more precise, let us denote the trajectory obtained from *Method 1.* by \tilde{z}_{M1} , and the other one by \tilde{z}_{M2} . Consider an arbitrary point $t \in [0, 1]$ and a neighborhood $(t - \rho, t + \rho)$ with $\rho > 0$ defined in such a way that (4.44) holds. First observe that from Equations (4.42), and (4.45) we obtain immediately that $\tilde{z}_{M1}(t) = \tilde{z}_{M2}(t)$.

Now for any $\tau \in (t - \rho, t + \rho)$, we have $\tilde{z}_{M1}(\tau) = G_\tau^{-1}(\tilde{p}(\tau) - p(\tau)) \cap \mathbb{B}_a(z(\tau))$.

We have already seen in proof *Method 1.* of the previous theorem that when inequalities in (4.43) are satisfied (which is the case, by Condition 4.44 on ρ), it is possible to conclude that $\tilde{z}_{M1}(\tau) \in \mathbb{B}_a(z(t))$.

On the other hand, Lemma 4.2.3 implies that $\tilde{z}_{M1}(\tau) \in G_t^{-1}(\tilde{p}(\tau) - p(t))$. Thus,

$\tilde{z}_{M_1}(\tau) \in G_t^{-1}(\tilde{p}(\tau) - p(t)) \cap B_a(z(t))$, and by strong metric regularity of G_t we can obtain the desired equality $\tilde{z}_{M_1}(\tau) = \tilde{z}_{M_2}(\tau)$.

(b) It is worth mentioning that the method of construction over intervals shows explicitly that Lipschitz continuity of $\tilde{z}(\cdot)$ could be easily obtained from Lipschitz continuity of $\tilde{p}(\cdot)$. But this is not something new or more than what we can obtain from the method of pointwise construction, as it was implicitly mentioned there, too. Indeed, in view of Lemma 4.2.13, Proposition 4.2.4, and Corollary 4.2.6, we get the same result.

References

- [1] Variational analysis and generalized differentiation. II. Applications, *Online book for special - purpose diodes*, <http://www.allaboutcircuits.com/textbook/semiconductors/chpt-3/special-purpose-diodes/> and <http://www.allaboutcircuits.com/textbook/semiconductors/chpt-7>, Last time accessed: 2017-01-10.
- [2] V. Acary, O. Bonnefon, and B. Brogliato, *Nonsmooth modeling and simulation for switched circuits*, Lecture Notes in Electrical Engineering, vol. 69, Springer, Dordrecht, 2011.
- [3] S. Adly and R. Cibulka, *Quantitative stability of a generalized equation*, J. Optim. Theory Appl. **160** (2014), no. 1, 90–110.
- [4] S. Adly, R. Cibulka, and H. Massias, *Variational analysis and generalized equations in electronics*, Set-Valued Var. Anal. **21** (2013), no. 2, 333–358.
- [5] S. Adly, R. Cibulka, and H. V. Ngai, *Newton's method for solving inclusions using set-valued approximations*, SIAM J. Optim. **25** (2015), no. 1, 159–184.
- [6] S. Adly, H. V. Ngai, and V. V. Nguyen, *Newton's method for solving generalized equations: Kantorovich's and Smale's approaches*, J. Math. Anal. Appl. **439** (2016), no. 1, 396–418.
- [7] S. Adly and J. V. Outrata, *Qualitative stability of a class of non-monotone variational inclusions. Application in electronics*, J. Convex Anal. **20** (2013), no. 1, 43–66.
- [8] F. J. Aragón Artacho, *On metric regularity of mappings in optimization*, Ph.D. thesis, University of Murcia, 2007.
- [9] F. J. Aragón Artacho and A. L. Dontchev, *On the inner and outer norms of sublinear mappings*, Set-Valued Anal. **15** (2007), no. 1, 61–65.

-
- [10] F. J. Aragón Artacho and B. S. Mordukhovich, *Metric regularity and Lipschitzian stability of parametric variational systems*, *Nonlinear Anal.* **72** (2010), no. 3-4, 1149–1170.
- [11] J-P. Aubin and H. Frankowska, *Set-valued analysis*, *Systems & Control: Foundations & Applications*, vol. 2, Birkhäuser Boston, Inc., Boston, MA, 1990.
- [12] C. Berge, *Topological spaces: including a treatment of multi-valued functions, vector spaces and convexity*, Dover Publications, Inc., Mineola, NY, 1997, Translated from the French original by E. M. Patterson, Reprint of the 1963 translation.
- [13] M. Bianchi, G. Kassay, and R. Pini, *An inverse map result and some applications to sensitivity of generalized equations*, *J. Math. Anal. Appl.* **399** (2013), no. 1, 279–290.
- [14] M. Bianchi, G. Kassay, and R. Pini, *Linear Openness of the Composition of Set-Valued Maps and Applications to Variational Systems*, *Set-Valued Var. Anal.* **24** (2016), no. 4, 581–595.
- [15] J. M. Borwein, *Stability and regular points of inequality systems*, *J. Optim. Theory Appl.* **48** (1986), no. 1, 9–52.
- [16] T. C. Carusone, D. A. Johns, and K. Martin, *Analog integrated circuit design*, John Wiley & Sons, 2nd edition, 2012.
- [17] R. Cibulka, *Differential variational inequalities: a gentle invitation*, proceedings of XXIX Seminar in differential equations, Moníec, April 14-18, 2014. Volume I, University of West Bohemia in Pilsen, 2016.
- [18] R. Cibulka, A. L. Dontchev, M. Krastanov, and V. M. Veliov, *Metrically regular differential generalized equations*, Tech. report, Institute of Statistics and Mathematical Methods in Economics, Vienna University of Technology, 09 2016.
- [19] R. Cibulka, A. L. Dontchev, and A. Y. Kruger, *Strong metric subregularity of mappings in variational analysis and optimization*, *J. Math. Anal. Appl.* (2016).
- [20] R. Cibulka and T. Roubal, *Solution stability and path-following for a class of generalized equations*, Private Communication.
- [21] F. H. Clarke, *On the inverse function theorem*, *Pacific J. Math.* **64** (1976), no. 1, 97–102.
- [22] F. H. Clarke, *Optimization and nonsmooth analysis*, second ed., *Classics in Applied Mathematics*, vol. 5, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 1990. MR 1058436

-
- [23] C. A. Desoer and E. S. Kuh, *Basic circuit theory*, McGraw Hill international editions: Electrical and electronic engineering series, McGraw-Hill, 1969.
- [24] A. L. Dontchev, *Characterizations of Lipschitz stability in optimization*, Recent developments in well-posed variational problems, Math. Appl., vol. 331, Kluwer Acad. Publ., Dordrecht, 1995, pp. 95–115.
- [25] A. L. Dontchev, M. I. Krastanov, R. T. Rockafellar, and V. M. Veliov, *An Euler-Newton continuation method for tracking solution trajectories of parametric variational inequalities*, SIAM J. Control Optim. **51** (2013), no. 3, 1823–1840.
- [26] A. L. Dontchev, A. S. Lewis, and R. T. Rockafellar, *The radius of metric regularity*, Trans. Amer. Math. Soc. **355** (2003), no. 2, 493–517.
- [27] A. L. Dontchev and R. T. Rockafellar, *Regularity and conditioning of solution mappings in variational analysis*, Set-Valued Anal. **12** (2004), no. 1-2, 79–109.
- [28] A. L. Dontchev and R. T. Rockafellar, *Newton’s method for generalized equations: a sequential implicit function theorem*, Math. Program. **123** (2010), no. 1, Ser. B, 139–159.
- [29] A. L. Dontchev and R. T. Rockafellar, *Implicit functions and solution mappings*, second ed., Springer Series in Operations Research and Financial Engineering, Springer, New York, 2014, A view from variational analysis.
- [30] A. L. Dontchev and V. M. Veliov, *Metric regularity under approximations*, Control Cybernet. **38** (2009), no. 4B, 1283–1303. MR 2760907
- [31] M. Durea and R. Strugariu, *Chain rules for linear openness in general Banach spaces*, SIAM J. Optim. **22** (2012), no. 3, 899–913.
- [32] M. Durea and R. Strugariu, *Openness stability and implicit multifunction theorems: applications to variational systems*, Nonlinear Anal. **75** (2012), no. 3, 1246–1259.
- [33] M. Durea and R. Strugariu, *Chain rules for linear openness in metric spaces and applications*, Math. Program. **143** (2014), no. 1-2, Ser. A, 147–176.
- [34] M. Durea and R. Strugariu, *An introduction to nonlinear optimization theory*, De Gruyter Open, Berlin, 2014, Edited by Aleksandra Nowacka-Leverton, Vicentiu Radulescu and Nick Rogers.
- [35] F. Facchinei and J-S. Pang, *Finite-dimensional variational inequalities and complementarity problems. Vol. I*, Springer Series in Operations Research, Springer-Verlag, New York, 2003.

-
- [36] F. Facchinei and J-S. Pang, *Finite-dimensional variational inequalities and complementarity problems. Vol. II*, Springer Series in Operations Research, Springer-Verlag, New York, 2003.
- [37] O. P. Ferreira and G. N. Silva, *Kantorovich's theorem on Newton's method for solving strongly regular generalized equation*, arXiv preprint arXiv:1604.04569 (2016).
- [38] O. P. Ferreira and G. N. Silva, *Unifying the local convergence analysis of Newton's method for strongly regular generalized equations*, arXiv preprint arXiv:1604.04568 (2016).
- [39] H. Gfrerer, *First order and second order characterizations of metric subregularity and calmness of constraint set mappings*, SIAM J. Optim. **21** (2011), no. 4, 1439–1474.
- [40] G. T. Gilbert, *Positive definite matrices and Sylvester's criterion*, Amer. Math. Monthly **98** (1991), no. 1, 44–46.
- [41] R. Henrion and J. V. Outrata, *Calmness of constraint systems with applications*, Math. Program. **104** (2005), no. 2-3, Ser. B, 437–464.
- [42] A. D. Ioffe, *Regular points of Lipschitz functions*, Trans. Amer. Math. Soc. **251** (1979), 61–69.
- [43] A. D. Ioffe, *Metric regularity—A survey part I. Theory*, J. Aust. Math. Soc. **101** (2016), no. 2, 188–243.
- [44] A. D. Ioffe, *Metric regularity—A survey part II. Applications*, J. Aust. Math. Soc. **101** (2016), no. 3, 376–417.
- [45] A. D. Ioffe and J. V. Outrata, *On metric and calmness qualification conditions in subdifferential calculus*, Set-Valued Anal. **16** (2008), no. 2-3, 199–227.
- [46] A. F. Izmailov, *Strongly regular nonsmooth generalized equations*, Math. Program. **147** (2014), no. 1-2, Ser. A, 581–590.
- [47] B. Kummer, *Generalized equations: solvability and regularity*, Math. Programming Stud. (1984), no. 21, 199–212, Sensitivity, stability and parametric analysis.
- [48] T. R. Kuphaldt, *Lessons in electric circuits. volume III-semiconductors*, Design Science License, Open Book Project, 5th edition, 2009.
- [49] J. Kyparisis, *Sensitivity analysis framework for variational inequalities*, Math. Programming **38** (1987), no. 2, 203–213.

- [50] A. B. Levy and B. S. Mordukhovich, *Coderivatives in parametric optimization*, Math. Program. **99** (2004), no. 2, Ser. A, 311–327.
- [51] B. S. Mordukhovich, *Generalized differential calculus for nonsmooth and set-valued mappings*, J. Math. Anal. Appl. **183** (1994), no. 1, 250–288.
- [52] B. S. Mordukhovich, *Stability theory for parametric generalized equations and variational inequalities via nonsmooth analysis*, Trans. Amer. Math. Soc. **343** (1994), no. 2, 609–657.
- [53] B. S. Mordukhovich, *Coderivatives of set-valued mappings: calculus and applications*, Proceedings of the Second World Congress of Nonlinear Analysts, Part 5 (Athens, 1996), vol. 30, 1997, pp. 3059–3070.
- [54] B. S. Mordukhovich, *Variational analysis and generalized differentiation. I. Basic theory*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 330, Springer-Verlag, Berlin, 2006.
- [55] B. S. Mordukhovich and Y. H. Shao, *Differential characterizations of covering, metric regularity, and Lipschitzian properties of multifunctions between Banach spaces*, Nonlinear Anal. **25** (1995), no. 12, 1401–1424.
- [56] B. S. Mordukhovich and B. Wang, *Restrictive metric regularity and generalized differential calculus in Banach spaces*, Int. J. Math. Math. Sci. (2004), no. 49-52, 2653–2680.
- [57] B. S. Mordukhovich, *Variational analysis and generalized differentiation. II. Applications*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 331, Springer-Verlag, Berlin, 2006.
- [58] J. M. Ortega and W. C. Rheinboldt, *Iterative solution of nonlinear equations in several variables*, Classics in Applied Mathematics, vol. 30, Society for Industrial and Applied Mathematics (SIAM), Philadelphia, PA, 2000, Reprint of the 1970 original.
- [59] S. M. Robinson, *Generalized equations and their solutions. I. Basic theory*, Math. Programming Stud. (1979), no. 10, 128–141, Point-to-set maps and mathematical programming.
- [60] S. M. Robinson, *Strongly regular generalized equations*, Math. Oper. Res. **5** (1980), no. 1, 43–62.
- [61] S. M. Robinson, *Generalized equations and their solutions. II. Applications to nonlinear programming*, Math. Programming Stud. (1982), no. 19, 200–221, Optimality and stability in mathematical programming.

-
- [62] S. M. Robinson, *Generalized equations*, Mathematical programming: the state of the art (Bonn, 1982), Springer, Berlin, 1983, pp. 346–367.
- [63] R. T. Rockafellar, *Convex analysis*, Princeton Mathematical Series, No. 28, Princeton University Press, Princeton, N.J., 1970.
- [64] R. T. Rockafellar, *Proto-differentiability of set-valued mappings and its applications in optimization*, Ann. Inst. H. Poincaré Anal. Non Linéaire **6** (1989), no. suppl., 449–482, Analyse non linéaire (Perpignan, 1987).
- [65] R. T. Rockafellar and R. J-B. Wets, *Variational analysis*, Grundlehren der Mathematischen Wissenschaften [Fundamental Principles of Mathematical Sciences], vol. 317, Springer-Verlag, Berlin, 1998.
- [66] T. Roubal, *Non-smooth analysis in engineering and economics*, Master’s thesis, University of West Bohemia in Pilsen, Czech Republic, 2015.
- [67] W. Schirotzek, *Nonsmooth analysis*, Universitext, Springer, Berlin, 2007.
- [68] A. S. Sedra and K. C. Smith, *Microelectronic circuits*, Oxford University Press, Inc., 5th edition, 2004.
- [69] A. Shapiro, *Sensitivity analysis of parametrized programs via generalized equations*, SIAM J. Control Optim. **32** (1994), no. 2, 553–571.
- [70] A. Uderzo, *On some regularity properties in variational analysis*, Set-Valued Var. Anal. **17** (2009), no. 4, 409–430.
- [71] A. Uderzo, *A strong metric subregularity analysis of nonsmooth mappings via steepest displacement rate*, J. Optim. Theory Appl. **171** (2016), no. 2, 573–599.
- [72] M. D. Voisei, *The sum and chain rules for maximal monotone operators*, Set-Valued Anal. **16** (2008), no. 4, 461–476.

Notation

\mathbb{R}	the real numbers
\mathbb{R}_+	the non-negative real numbers
\mathbb{R}^m	the m -dimensional Euclidean space
\mathbb{N}	the natural numbers: $1, 2, 3, \dots$
$\mathbb{B}_r(\bar{x})$	closed ball of radius $r > 0$ centred at \bar{x}
\mathbb{B}	closed unit ball, $\mathbb{B}_1(0)$
$\text{int } C$	interior of the set C
$\text{cl } C$	closure of the set C
$x_k \longrightarrow x$	the sequence (x_k) is convergent to x
$t_k \downarrow 0$	a sequence of positive numbers t_k tending to 0
$d(x, C)$	distance from x to the set C
$e(C, D)$	excess of the set C beyond the set D
$ x $	absolute value of $x \in \mathbb{R}$
$\ x\ $	norm of x
$\langle x, y \rangle$	canonical inner product, bilinear form
$ H ^+$	outer norm
$ H ^-$	inner norm
$T(\bar{x}; C)$	Bouligand-Severi tangent cone (contingent cone) to the set C at \bar{x}
$\tilde{T}(\bar{x}; C)$	Bouligand paratingent cone to the set C at \bar{x}
$\hat{N}(\bar{x}; C)$	Fréchet normal cone (regular normal cone) to the set C at \bar{x}
$N(\bar{x}; C)$	Mordukhovich normal cone (limiting normal cone) to the set C at \bar{x}
$F : X \rightrightarrows Y$	set-valued mapping from X into the subsets of Y
$f : X \rightarrow Y$	function f from X into Y
A^T	transpose of the matrix A
$\ker A$	kernel of the the linear operator A
$\det A$	determinant of the matrix A
$\text{gph } F$	graph of the mapping F
$\text{dom } F$	domain of the mapping F

$\text{rge } F$	range of the mapping F
$\partial_F f(\bar{x})$	Fréchet subdifferential (F-subdifferential) of the function f at \bar{x}
$\partial f(\bar{x})$	Mordukhovich (limiting) subdifferential of the function f at \bar{x}
$\partial_{>} f(\bar{x})$	outer subdifferential of the function f at \bar{x}
$\partial_B h(\bar{u})$	Bouligand's limiting Jacobian of the function h at \bar{u}
$\partial h(\bar{u})$	Clarke's generalized Jacobian of h at \bar{u}
$\nabla f(\bar{x})$	Jacobian matrix of the function f at \bar{x}
$Df(\bar{x})$	derivative of the function f at \bar{x}
$DF(\bar{x} \bar{y})$	graphical derivative of the mapping F at \bar{x} for \bar{y}
$D^*F(\bar{x} \bar{y})$	coderivative of the mapping F at \bar{x} for \bar{y}
$\tilde{D}F(\bar{x} \bar{y})$	strict graphical derivative of the mapping F at \bar{x} for \bar{y}
$\text{clm}(f; \bar{x})$	calmness modulus of the function f at \bar{x}
$\text{clm}_x(f; (\bar{p}, \bar{x}))$	partial calmness modulus of f with respect to x at (\bar{p}, \bar{x})
$\widehat{\text{clm}}_x(f; (\bar{p}, \bar{x}))$	uniform partial calmness modulus of f with respect to x
$\text{clm}(S; \bar{y} \bar{x})$	calmness modulus of S at \bar{y} for \bar{x}
$\text{lip}(f; \bar{x})$	Lipschitz modulus of the function f at \bar{x}
$\widehat{\text{lip}}_x(f; (\bar{p}, \bar{x}))$	uniform partial Lipschitz modulus of f with respect to x
$\text{lip}(S; \bar{y} \bar{x})$	Lipschitz modulus of S at \bar{y} for \bar{x}
$\text{reg}(F; \bar{x} \bar{y})$	regularity modulus of F at \bar{x} for \bar{y}
$\text{subreg}(F; \bar{x} \bar{y})$	modulus of metric sub-regularity of F at \bar{x} for \bar{y}

Index

- Aubin property, 19
- calmness
 - for functions, 12
 - for set-valued maps, 22
 - isolated calmness, 22
 - partial calmness, 34
- circuit components
 - capacitor, 46
 - current source, 47
 - DIAC, 51
 - diode, 48
 - inductor, 46
 - LED, 50
 - resistor, 45
 - SCR, 51
 - Shockley diode, 50
 - transistor, 52
 - voltage source, 47
 - Zener diode, 49
- Clarke
 - Clarke regularity of sets, 79
 - Clarke tangent cone, 76
 - Clarke's generalized Jacobian, 100
- coderivative, 83
- cone
 - Clarke tangent cone, 76
 - contingent cone, 76
 - limiting normal cone, 78
 - paratingent cone, 76
 - regular normal cone, 78
- excess, 12
- Fréchet subdifferentiable, 99
- function
 - differentiable, 13
 - estimator, 33
 - first-order approximation, 33
 - implicit, 16
 - inverse, 10
 - Jacobian, 13
 - Lipschitz continuous, 13
 - lower semicontinuous, 99
 - partial estimator, 34
 - strict differentiable, 14
 - subdifferential, 99
- generalized Jacobian, 100
- graphical derivative, 83
- Haudorff distance, 11
- inner norm, 90
- isolated trajectories, 157
- KCL, Kirchhoff's current law, 54
- KVL, Kirchhoff's voltage law, 55
- linear openness, 23
- localization
 - graphical localization, 15
 - single-valued localization, 16
- locally closed set, 19
- metric regularity
 - metric regularity, 24
 - metric sub-regularity, 29

- restrictive metrically regular (RMR), 89
- strong metric regularity, 28
- strong metric sub-regularity, 31
- uniform strong metric regularity, 147
- monotone
 - locally monotone, 17
 - maximal monotone, 17
- monotone map, 17
- Mordukhovich subdifferential, 99

- open at linear rate, 23
- outer norm, 90
- outer subdifferential, 100

- P-matrix, 104
- positively homogeneous map, 90
- problem
 - dynamic case, 3
 - static case, 3

- restriction, 12

- selection, 14
- solution mapping, 57
- solution trajectory, 144
- stability-like, 5
- strict graphical derivative, 84
- strongly regular point, 146