Asen L. Dontchev
R. Tyrrell Rockafellar

SPRINGER
MONOGRAPHS IN MATHEMATICS

Implicit Functions and Solution Mappings

A View from Variational Analysis

隐函数和解映射

Springer

世界图书出版公司
www.wpcbj.com.cn

Asen L. Dontchev • R. Tyrrell Rockafellar

Implicit Functions and Solution Mappings

A View from Variational Analysis

With 12 Illustrations

Springer

Asen L. Dontchev
Mathematical Reviews
416 Fourth Street
Ann Arbor, MI 48107-8604
USA
ald@ams.org

R. Tyrrell Rockafellar
University of Washington
Department of Mathematics
PO Box 354350
Seattle, WA 98195-4350
USA
rtr@math.washington.edu

Reprint from English language edition:
Implicit Functions and Solution Mappings
by Asen L.Dontchev, R.Tyrrell Rockafellar

ISSN 1439-7382 e-ISSN
ISBN 978-0-387-87820-1 e-ISBN 978-0-387-87821-8
DOI 10.1007/978-0-387-87821-8
Springer Dordrecht Heidelberg London New York

Library of Congress Control Number: 2009926485

Mathematics Subject Classification (2000): 26B10, 47J07, 58C15, 49J53, 49K40, 90C31, 93C70

图书在版编目（CIP）数据

隐函数和解映射 = Implicit functions and solution mappings：英文/（美）邓契夫（Dontchev，A. L.）著. —影印本. —北京：世界图书出版公司北京公司，2013. 3

ISBN 978 - 7 - 5100 - 5807 - 3

Ⅰ. ①隐…　Ⅱ. ①邓…　Ⅲ. ①隐函数定理—英文　Ⅳ. ①O177. 91

中国版本图书馆 CIP 数据核字（2013）第 035308 号

书　　名：	Implicit Functions and Solution Mappings：A View from Variational Analysis
作　　者：	Asen L. Dontchev，R. Tyrrell Rockafellar
中 译 名：	隐函数和解映射
责任编辑：	高蓉　刘慧
出 版 者：	世界图书出版公司北京公司
印 刷 者：	三河市国英印务有限公司
发　　行：	世界图书出版公司北京公司（北京朝内大街 137 号 100010）
联系电话：	010 - 64021602，010 - 64015659
电子信箱：	kjb@wpcbj.com.cn
开　　本：	24 开
印　　张：	16. 5
版　　次：	2013 年 3 月
版权登记：	图字：01 - 2012 - 8536
书　　号：	978 - 7 - 5100 - 5807 - 3　　定　　价：59. 00 元

Preface

Setting up equations and solving them has long been so important that, in popular imagination, it has virtually come to describe what mathematical analysis and its applications are all about. A central issue in the subject is whether the solution to an equation involving parameters may be viewed as a function of those parameters, and if so, what properties that function might have. This is addressed by the classical theory of implicit functions, which began with single real variables and progressed through multiple variables to equations in infinite dimensions, such as equations associated with integral and differential operators.

A major aim of the book is to lay out that celebrated theory in a broader way than usual, bringing to light many of its lesser known variants, for instance where standard assumptions of differentiability are relaxed. However, another major aim is to explain how the same constellation of ideas, when articulated in a suitably expanded framework, can deal successfully with many other problems than just solving equations.

These days, forms of modeling have evolved beyond equations, in terms, for example, of problems of minimizing or maximizing functions subject to constraints which may include systems of inequalities. The question comes up of whether the solution to such a problem may be expressed as a function of the problem's parameters, but differentiability no longer reigns. A function implicitly obtainable this manner may only have one-sided derivatives of some sort, or merely exhibit Lipschitz continuity or something weaker. Mathematical models resting on equations are replaced by "variational inequality" models, which are further subsumed by "generalized equation" models.

The key concept for working at this level of generality, but with advantages even in the context of equations, is that of the set-valued *solution mapping* which assigns to each instance of the parameter element in the model *all* the corresponding solutions, if any. The central question is whether a solution mapping can be localized graphically in order to achieve single-valuedness and in that sense produce a function, the desired *implicit function*.

In modern variational analysis, set-valued mappings are an accepted workhorse in problem formulation and analysis, and many tools have been developed for handling them. There are helpful extensions of continuity, differentiability, and regularity of several types, together with powerful results about how they can be applied. A corresponding further aim of this book is to bring such ideas to wider attention by demonstrating their aptness for the fundamental topic at hand.

In line with classical themes, we concentrate primarily on local properties of solution mappings that can be captured metrically, rather than on results derived from topological considerations or involving exotic spaces. In particular, we only briefly discuss the Nash–Moser inverse function theorem. We keep to finite dimensions in Chapters 1 to 4, but in Chapters 5 and 6 provide bridges to infinite dimensions. Global implicit function theorems, including the classical Hadamard theorem, are not discussed in the book.

In Chapter 1 we consider the implicit function paradigm in the classical case of the solution mapping associated with a parameterized equation. We give two proofs of the classical inverse function theorem and then derive two equivalent forms of it: the implicit function theorem and the correction function theorem. Then we gradually relax the differentiability assumption in various ways and even completely exit from it, relying instead on the Lipschitz continuity. We also discuss situations in which an implicit function fails to exist as a graphical localization of the solution mapping, but there nevertheless exists a function with desirable properties serving locally as a selection of the set-valued solution mapping. This chapter does not demand of the reader more than calculus and some linear algebra, and it could therefore be used by both teachers and students in analysis courses.

Motivated by optimization problems and models of competitive equilibrium, Chapter 2 moves into wider territory. The questions are essentially the same as in the first chapter, namely, when a solution mapping can be localized to a function with some continuity properties. But it is no longer an equation that is being solved. Instead it is a condition called a generalized equation which captures a more complicated dependence and covers, as a special case, variational inequality conditions formulated in terms of the set-valued normal cone mapping associated with a convex set. Although our prime focus here is variational models, the presentation is self-contained and again could be handled by students and others without special background. It provides an introduction to a subject of great applicability which is hardly known to the mathematical community familiar with classical implicit functions, perhaps because of inadequate accessibility.

In Chapter 3 we depart from insisting on localizations that yield implicit functions and approach solution mappings from the angle of a "varying set." We identify continuity properties which support the paradigm of the implicit function theorem in a set-valued sense. This chapter may be read independently from the first two. Chapter 4 continues to view solution mappings from this angle but investigates substitutes for classical differentiability. By utilizing concepts of generalized derivatives, we are able to get implicit mapping theorems that reach far beyond the classical scope.

Chapter 5 takes a different direction. It presents extensions of the Banach open mapping theorem which are shown to fit infinite-dimensionally into the paradigm of the theory developed finite-dimensionally in Chapter 3. Some background in basic functional analysis is required. Chapter 6 goes further down that road and illustrates how some of the implicit function/mapping theorems from earlier in the book can be used in the study of problems in numerical analysis.

This book is targeted at a broad audience of researchers, teachers and graduate students, along with practitioners in mathematical sciences, engineering, economics and beyond. In summary, it concerns one of the chief topics in all of analysis, historically and now, an aid not only in theoretical developments but also in methods for solving specific problems. It crosses through several disciplines such as real and functional analysis, variational analysis, optimization, and numerical analysis, and can be used in part as a graduate text as well as a reference. It starts with elementary results and with each chapter, step by step, opens wider horizons by increasing the complexity of the problems and concepts that generate implicit function phenomena.

Many exercises are included, most of them supplied with detailed guides. These exercises complement and enrich the main results. The facts they encompass are sometimes invoked in the subsequent sections.

Each chapter ends with a short commentary which indicates sources in the literature for the results presented (but is not a survey of all the related literature). The commentaries to some of the chapters additionally provide historical overviews of past developments.

Whidbey Island, Washington
August, 2008

Asen L. Dontchev
R. Tyrrell Rockafellar

Acknowledgements

Special thanks are owed to our readers Marius Durea, Shu Lu, Yoshiyuki Sekiguchi and Hristo Sendov, who gave us valuable feedback on the entire manuscript, and to Francisco J. Aragón Artacho, who besides reviewing most of the book helped us masterfully with all the figures. During various stages of the writing we also benefited from discussions with Aris Daniilidis, Darinka Dentcheva, Hélène Frankowska, Michel Geoffroy, Alexander Ioffe, Stephen Robinson, Vladimir Veliov, and Constantin Zălinescu. We are also grateful to Mary Anglin for her help with the final copy-editing of the book.

The authors

Contents

Chapter 1
Functions Defined Implicitly by Equations

The idea of solving an equation $f(p,x) = 0$ for x as a function of p, say $x = s(p)$, plays a huge role in classical analysis and its applications. The function obtained in this way is said to be defined *implicitly* by the equation. The closely related idea of solving an equation $f(x) = y$ for x as a function of y concerns the *inversion* of f. The circumstances in which an implicit function or an inverse function exists and has properties like differentiability have long been studied. Still, there are features which are not widely appreciated and variants which are essential to seeing how the subject might be extended beyond solving only equations. For one thing, properties other than differentiability, such as Lipschitz continuity, can come in. But fundamental expansions in concept, away from thinking just about functions, can serve in interesting ways as well.

As a starter, consider for real variables x and y the extent to which the equation $x^2 = y$ can be solved for x as a function of y. This concerns the inversion of the function $f(x) = x^2$ in Figure 1.1 below, as depicted through the reflection that interchanges the x and y axes. The reflection of the graph is not the graph of a function, but some parts of it may have that character. For instance, a function is obtained from a neighborhood of the point B, but not from one of the point A, no matter how small.

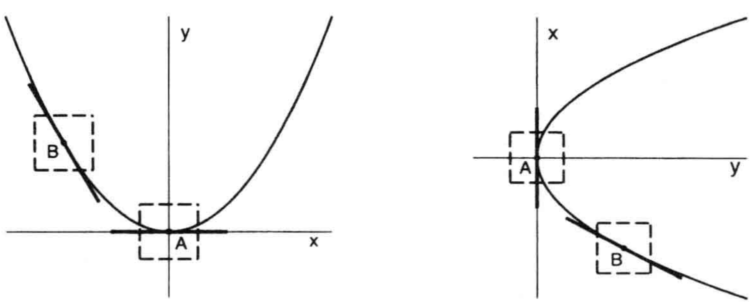

Fig. 1.1 Graphical localizations of the function $y = x^2$ and its inverse.

A.L. Dontchev and R.T. Rockafellar, *Implicit Functions and Solution Mappings: A View from Variational Analysis*, Springer Monographs in Mathematics,
DOI 10.1007/978-0-387-87821-8_1, © Springer Science+Business Media, LLC 2009

Although the reflected graph in this figure is not, as a whole, the graph of a function, it can be regarded as the graph of something more general, a "set-valued mapping" in terminology which will be formalized shortly. The question revolves then around the extent to which a "graphical localization" of a set-valued mapping might be a function, and if so, what properties that function would possess. In the case at hand, the reflected graph assigns two different x's to y when $y > 0$, but no x when $y < 0$, and just $x = 0$ when $y = 0$.

To formalize that framework for the general purposes of this chapter, we focus on set-valued mappings F from $\mathbb{R}^n$ and $\mathbb{R}^m$, signaled by the notation

$$F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m,$$

by which we mean correspondences which assign to each $x \in \mathbb{R}^n$ one or more elements of $\mathbb{R}^m$, or possibly none. The set of elements $y \in \mathbb{R}^m$ assigned by F to x is denoted by $F(x)$. However, instead of regarding F as going from $\mathbb{R}^n$ to a space of subsets of $\mathbb{R}^m$ we identify as the *graph* of F the set

$$\operatorname{gph} F = \{ (x,y) \in \mathbb{R}^n \times \mathbb{R}^m \mid y \in F(x) \}.$$

Every subset of $\mathbb{R}^n \times \mathbb{R}^m$ serves as $\operatorname{gph} F$ for a uniquely determined $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$, so this concept is very broad indeed, but it opens up many possibilities.

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 $F(x) = \{y\}$ and thereby build a bridge to handling functions as special cases of set-valued mappings.

Domains and ranges get flexible treatment in this way. For $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ the *domain* is the set

$$\operatorname{dom} F = \{ x \mid F(x) \neq \emptyset \},$$

while the *range* is

$$\operatorname{rge} F = \{ y \mid y \in F(x) \text{ for some } x \},$$

so that $\operatorname{dom} F$ and $\operatorname{rge} F$ are the projections of $\operatorname{gph} F$ on $\mathbb{R}^n$ and $\mathbb{R}^m$ respectively. Any subset of $\operatorname{gph} F$ can freely be regarded then as itself the graph of a set-valued submapping which likewise projects to some domain in $\mathbb{R}^n$ and range in $\mathbb{R}^m$.

The *functions* from $\mathbb{R}^n$ to $\mathbb{R}^m$ are identified in this context with the set-valued mappings $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ such that F is single-valued at every point of $\operatorname{dom} F$. When F is a function, we can emphasize this by writing $F : \mathbb{R}^n \to \mathbb{R}^m$, but the notation $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ doesn't preclude F from actually being a function. Usually, though, we use lower case letters for functions: $f : \mathbb{R}^n \to \mathbb{R}^m$. Note that in this notation f can still be empty-valued in places; it's single-valued only on the subset $\operatorname{dom} f$ of $\mathbb{R}^n$. Note also that, although we employ "mapping" in a sense allowing for potential multi-valuedness (as in a "set-valued mapping"), no multi-valuedness is ever involved when we speak of a "function."

A clear advantage of the framework of set-valued mappings over that of only functions is that *every set-valued mapping* $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ *has an inverse*, namely the set-valued mapping $F^{-1} : \mathbb{R}^m \rightrightarrows \mathbb{R}^n$ defined by

$$F^{-1}(y) = \{ x \,|\, y \in F(x) \}.$$

The graph of F^{-1} is generated from the graph of F simply by reversing (x,y) to (y,x), which in the case of $m = n = 1$ corresponds to the reflection in Figure 1.1. 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}. For that, the notion of "localization" must come into play, as we are about to explain after a bit more background. Traditionally, a function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *surjective* when $\operatorname{rge} f = \mathbb{R}^m$ and *injective* when $\operatorname{dom} f = \mathbb{R}^n$ and f^{-1} is a function; full *invertibility* of f corresponds to the juxtaposition of these two properties.

In working with $\mathbb{R}^n$ we will, for now, keep to the Euclidean norm $|x|$ associated with the canonical inner product

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

namely

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

The closed ball around $\bar{x}$ with radius r is then

$$\mathbb{B}_r(\bar{x}) = \{ x \,|\, |x - \bar{x}| \le r \}.$$

We denote the closed unit ball $\mathbb{B}_1(0)$ by $\mathbb{B}$. A *neighborhood* of $\bar{x}$ is any set U such that $\mathbb{B}_r(\bar{x}) \subset U$ for some $r > 0$. We recall for future needs that the interior of a set $C \subset \mathbb{R}^n$ consists of all points x such that C is a neighborhood of x, whereas the closure of C consists of all points x such that the *complement* of C is *not* a neighborhood of x; C is *open* if it coincides with its interior and *closed* if it coincides with its closure. The interior and closure of C will be denoted by $\operatorname{int} C$ and $\operatorname{cl} C$.

Graphical localization. *For $F : \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ and a pair $(\bar{x}, \bar{y}) \in \operatorname{gph} F$, a graphical localization of F at $\bar{x}$ for $\bar{y}$ is a set-valued mapping $\tilde{F}$ such that*

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

so that

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

The inverse of $\tilde{F}$ then has

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

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

Often the neighborhoods U and V can conveniently be taken to be closed balls $\mathbb{B}_a(\bar{x})$ and $\mathbb{B}_b(\bar{y})$. Observe, however, that the domain of a graphical localization $\tilde{F}$ of F with respect to U and V may differ from $U \cap \operatorname{dom} F$ and may well depend on the choice of V.

Single-valuedness in localizations. *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 a single-valued localization of* F around $\bar{x}$ *for* $\bar{y}$ *instead of just* at $\bar{x}$ *for* $\bar{y}$.

For the function $f(x) = x^2$ from $\mathbb{R}$ to $\mathbb{R}$ we started with, the set-valued inverse mapping f^{-1}, which is single-valued only at 0 with $f^{-1}(0) = 0$, fails to have a single-valued localization at 0 for 0. But as observed in Figure 1.1, it has a single-valued localization around $\bar{y} = 1$ for $\bar{x} = -1$.

In passing from inverse functions to implicit functions more generally, we need to pass from an equation $f(x) = y$ to one of the form

$$(1) \qquad f(p,x) = 0 \quad \text{for a function } f : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^m$$

in which p acts as a parameter. The question is no longer that of inverting f, but the framework of set-valuedness is valuable nonetheless because it allows us to immediately introduce the *solution mapping*

$$(2) \qquad S : \mathbb{R}^d \rightrightarrows \mathbb{R}^n \text{ with } S(p) = \{ x \mid f(p,x) = 0 \}.$$

We can then look at pairs $(\bar{p},\bar{x})$ in $\operatorname{gph} S$ 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 equation. The classical implicit function theorem deduces the existence from certain assumptions on f. A review of the form of this theorem will help in setting the stage for later developments because of the pattern it provides. Again, some basic background needs to be recalled, and this is also an opportunity to fix some additional notation and terminology for subsequent use.

A function $f : \mathbb{R}^n \to \mathbb{R}$ is *upper semicontinuous* at a point $\bar{x}$ when $\bar{x} \in \operatorname{int} \operatorname{dom} f$ and for every $\varepsilon > 0$ there exists $\delta > 0$ for which

$$f(x) - f(\bar{x}) < \varepsilon \quad \text{whenever } x \in \operatorname{dom} f \text{ with } |x - \bar{x}| < \delta.$$

If instead we have

$$-\varepsilon < f(x) - f(\bar{x}) \quad \text{whenever } x \in \operatorname{dom} f \text{ with } |x - \bar{x}| < \delta,$$

then f is said to be *lower semicontinuous* at $\bar{x}$. Such upper and lower semicontinuity combine to *continuity*, meaning the existence for every $\varepsilon > 0$ of a $\delta > 0$ for which

$$|f(x) - f(\bar{x})| < \varepsilon \quad \text{whenever } x \in \operatorname{dom} f \text{ with } |x - \bar{x}| < \delta.$$

This condition, in our norm notation, carries over to defining the continuity of a vector-valued function $f : \mathbb{R}^n \to \mathbb{R}^m$ at a point $\bar{x} \in \operatorname{int} \operatorname{dom} f$. However, we also speak more generally then of f being *continuous at* $\bar{x}$ *relative to a set* D when $\bar{x} \in D \subset \operatorname{dom} f$ and this last estimate holds for $x \in D$; in that case $\bar{x}$ need not belong to $\operatorname{int} \operatorname{dom} f$. When f is continuous relative to D at every point of D, we say it is continuous on D. The graph $\operatorname{gph} f$ of a function $f : \mathbb{R}^n \to \mathbb{R}^m$ with closed domain $\operatorname{dom} f$ that is continuous on $D = \operatorname{dom} f$ is a closed set in $\mathbb{R}^n \times \mathbb{R}^m$.

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *Lipschitz continuous* relative to a set D, or on a set D, if $D \subset \operatorname{dom} f$ and there is a constant $\kappa \geq 0$ such that

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

If f is Lipschitz continuous relative to a neighborhood of a point $\bar{x} \in \operatorname{int} \operatorname{dom} f$, f is said to be Lipschitz continuous *around* $\bar{x}$. A function $f : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^m$ is Lipschitz continuous with respect to x uniformly in p near $(\bar{p}, \bar{x}) \in \operatorname{int} \operatorname{dom} f$ if there is a constant $\kappa \geq 0$ along with neighborhoods U of $\bar{x}$ and Q of $\bar{p}$ such that

$$|f(p,x') - f(p,x)| \leq \kappa |x' - x| \quad \text{for all } x', x \in U \text{ and } p \in Q.$$

Differentiability entails consideration of linear mappings. Although we generally allow for multi-valuedness and even empty-valuedness when speaking of "mappings," single-valuedness everywhere is required of a linear mapping, for which we typically use a letter like A. A linear mapping from $\mathbb{R}^n$ to $\mathbb{R}^m$ is thus a function $A : \mathbb{R}^n \to \mathbb{R}^m$ with $\operatorname{dom} A = \mathbb{R}^n$ which obeys the usual rule for linearity:

$$A(\alpha x + \beta y) = \alpha A x + \beta A y \text{ for all } x, y \in \mathbb{R}^n \text{ and all scalars } \alpha, \beta \in \mathbb{R}.$$

The *kernel* of A is

$$\ker A = \{ x \mid Ax = 0 \}.$$

In the finite-dimensional setting, we carefully distinguish between a linear mapping and its matrix, but often use the same notation for both. A linear mapping $A : \mathbb{R}^n \to \mathbb{R}^m$ is represented then by a matrix A with m rows, n columns, and components $a_{i,j}$:

$$A = \begin{pmatrix} a_{11} & a_{12} & \cdots & a_{1n} \\ a_{21} & a_{22} & \cdots & a_{2n} \\ \vdots & & & \vdots \\ a_{m1} & a_{m2} & \cdots & a_{mn} \end{pmatrix}.$$

The inverse A^{-1} of a linear mapping $A : \mathbb{R}^n \to \mathbb{R}^m$ always exists in the set-valued sense, but it isn't a linear mapping unless it is actually a function with all of $\mathbb{R}^m$ as its domain, in which case A is said to be *invertible*. From linear algebra, of course, that requires $m = n$ and corresponds to the matrix A being nonsingular. More generally, if $m \leq n$ and the rows of the matrix A are linearly independent, then the rank of the matrix A is m and the mapping A is surjective. In terms of the transpose of A, denoted by A^{T}, the matrix AA^{T} is in this case nonsingular. On the other hand, if $m \geq n$ and the columns of A are linearly independent then $A^{\mathsf{T}}A$ is nonsingular.

Both the identity mapping and its matrix will be denoted by I, regardless of dimensionality. By default, $|A|$ is the operator norm of A induced by the Euclidean norm,

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

A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is *differentiable* at a point $\bar{x}$ when $\bar{x} \in \operatorname{int} \operatorname{dom} f$ and there is a linear mapping $A : \mathbb{R}^n \to \mathbb{R}^m$ with the property that for every $\varepsilon > 0$ there exists $\delta > 0$ with

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

If such a mapping A exists at all, it is unique; it is denoted by $Df(\bar{x})$ and called the *derivative* of f at $\bar{x}$. A function $f : \mathbb{R}^n \to \mathbb{R}^m$ is said to be *twice differentiable* at a point $\bar{x} \in \operatorname{int} \operatorname{dom} f$ when there is a bilinear mapping $N : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^m$ with the property that for every $\varepsilon > 0$ there exists $\delta > 0$ with

$$|f(\bar{x}+h) - f(\bar{x}) - Df(\bar{x})h - N(h,h)| \leq \varepsilon |h|^2 \quad \text{for every } h \in \mathbb{R}^n \text{ with } |h| < \delta.$$

If such a mapping N exists it is unique and is called the *second derivative* of f at $\bar{x}$, denoted by $D^2 f(\bar{x})$. Higher-order derivatives can be defined accordingly.

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 the notation $x = (x_1, \ldots, x_n)$ and $f = (f_1, \ldots, f_m)$, the components of $\nabla f(\bar{x})$ are the partial derivatives of the component functions f_i:

$$\nabla f(\bar{x}) = \left(\frac{\partial f_i}{\partial x_j}(\bar{x}) \right)_{i,j=1}^{m,n}.$$

In distinguishing between $Df(\bar{x})$ as a linear mapping and $\nabla f(\bar{x})$ as its matrix, we can guard better against ambiguities which may arise in some situations. When the Jacobian $\nabla f(x)$ exists and is continuous (with respect to the matrix norms associated with the Euclidean norm) on a set $D \subset \mathbb{R}^n$, then we say that the function f is *continuously differentiable* on D; we also call such a function *smooth* or $\mathscr{C}^1$ on D. Accordingly, we define k times continuously differentiable ($\mathscr{C}^k$) functions.

For a function $f : \mathbb{R}^d \times \mathbb{R}^n \to \mathbb{R}^m$ and a pair $(\bar{p}, \bar{x}) \in \operatorname{int} \operatorname{dom} f$, the *partial derivative* mapping $D_x f(\bar{p}, \bar{x})$ of f with respect to x at $(\bar{p}, \bar{x})$ is the derivative of the function $g(x) = f(\bar{p}, x)$ at $\bar{x}$. If the partial derivative mapping is continuous as a function of the pair (p, x) in a neighborhood of $(\bar{p}, \bar{x})$, then f is said to be continuously differentiable with respect to x around $(\bar{p}, \bar{x})$. The partial derivative $D_x f(\bar{p}, \bar{x})$ is represented by an $m \times n$ matrix, denoted $\nabla_x f(\bar{p}, \bar{x})$ and called the partial Jacobian. Respectively, $D_p f(\bar{p}, \bar{x})$ is represented by the $m \times d$ partial Jacobian $\nabla_p f(\bar{p}, \bar{x})$. It's a standard fact from calculus that if f is differentiable with respect to both p and x around $(\bar{p}, \bar{x})$ and the partial Jacobians $\nabla_x f(p, x)$ and $\nabla_p f(p, x)$ depend continuously on p and x, then f is continuously differentiable around $(\bar{p}, \bar{x})$.


ANALISI
Infinitesimale
LEZIONI
DETTATE NELLA R. UNIVERSITÀ
PISA
dal Prof: Cav. Ulisse Dini
Anno Accademico 1877-78.
Parte Iª - CALCOLO DIFFERENZIALE
(L'Autore si riserva i diritti di Proprietà Letteraria)
Litografia Gozani
Pisa

Fig. 1.2 The front page of Dini's manuscript from 1877/78.

With this notation and terminology in hand, let us return to the setting of implicit functions in equation (1), as traditionally addressed with tools of differentiability. Most calculus books present a result going back to Dini[1], who formulated and proved it in lecture notes of 1877/78; the cover of Dini's manuscript is displayed above. The version typically seen in advanced texts is what we will refer to as the *classical implicit function theorem* or *Dini's theorem*. In those texts the set-valued solution mapping S in (2) never enters the picture directly, but a brief statement in that mode will help to show where we are headed in this book.

[1] Ulisse Dini (1845–1918). Many thanks to Danielle Ritelli from the University of Bologna for a copy of Dini's manuscript.