

Artur Wachowicz

Methodical Analysis of the Implicit Function Theorem*

Abstract. This paper provides a systematic analysis of the Implicit Function Theorem, intended as a case study for learning mathematical analysis. The discussion covers classical formulations, key assumptions, generalizations, and applications, with an emphasis on opportunities to apply active learning methods at each stage. This analysis is particularly inspired by the works of Sierpińska (1994), Krantz and Parks (2013), and Freeman et al. (2014). As intended by the author, a scenario is presented illustrating how the concept of the implicit function can be effectively introduced to students.

1. Introduction

In Sierpińska (1994), it is argued that mathematical understanding is a multifaceted and complex process that extends far beyond the mere acquisition of definitions and procedural knowledge. She identifies three principal forms of understanding: **instrumental understanding** (the ability to apply rules and algorithms), **logical understanding** (comprehension of relationships among mathematical concepts), and **intuitive understanding** (a personal, internal sense of meaning associated with those concepts). Sierpińska draws attention to the difficulties learners often encounter, which she attributes to epistemological obstacles – the cognitive dissonance that arises when intuitive, everyday modes of thinking come into conflict with the abstract and formal nature of mathematical reasoning. She underscores the crucial role of the teacher in supporting students' conceptual development, particularly by guiding them from concrete operational thinking toward abstract conceptualization. Essential to this process is the creation of learning environments that foster reflection, encourage questioning, and enable students to actively construct their own mathematical meanings.

*2020 Mathematics Subject Classification: *Primary: 97I50; Secondary: 97I40*

Keywords and phrases: *implicit function theorem, active methods of education*

In the paper by Freeman et al. (2014), the authors present a comprehensive meta-analysis of 225 studies comparing traditional lecturing with active learning approaches in undergraduate STEM courses. The results strongly support the conclusion that active learning significantly enhances student performance. Across 158 studies that measured exam scores or concept inventories, the average effect size was 0.47 standard deviations, which translates into an improvement of roughly six percentage points on examinations. Equally important are the findings regarding failure rates: in the analysis of 67 studies, students in lecture-based courses were 1.95 times more likely to fail (receive a D/F or withdraw) compared to those in active learning courses. In absolute terms, failure rates decreased from an average of 33.8% under traditional lecturing to 21.8% when active methods were used, a reduction of about one and a half times. The consistency of these results across disciplines, class sizes, and methodological designs makes the evidence especially robust. Interestingly, the strongest effects appeared in smaller classes with fifty students or fewer, but benefits were still observable in larger cohorts. Taken together, these findings suggest that active learning not only improves academic achievement but also substantially reduces the risk of attrition in STEM courses. The effect is large enough to be pedagogically meaningful – roughly equivalent to raising performance by half a letter grade. Because of both the magnitude and the reliability of the observed outcomes, the authors argue that continuing to rely only on traditional lectures as a control condition in future research may be questionable, since the evidence indicates that students systematically learn less and fail more often in such environments. In what follows we focus on indication where and how one can apply selected active methods of education.

The Implicit Function Theorem is one of the fundamental results in mathematical analysis and plays a significant role in higher-level mathematics education. Its importance stems not only from the richness of its applications across various fields of mathematics, such as analysis, differential geometry, and the theory of differential equations, but also from the fact that it serves as an excellent example of a theorem that enables the development of students' competencies in three mentioned areas of mathematical understanding.

A methodical analysis of this theorem involves breaking down the proof process into logical and pedagogical steps that can be consciously conveyed to students in order to understand not only the theorem itself but also the way mathematical arguments are constructed. Due to the complexity and precision of the assumptions involved (continuity, differentiability, local conditions), the Implicit Function Theorem is an excellent teaching material for learning by doing, as well as for developing critical analysis skills and independent problem-solving abilities. Effective teaching of the Implicit Function Theorem can utilize the following methods:

- **Group discussions and brainstorming** — facilitate the exchange of ideas and arguments.
- **Practical exercises and open problems** — allow students to independently apply the theorem to specific functions.
- **Visual methods** — e.g., function plots using software such as GeoGebra, to intuitively demonstrate the theorem's action.

- **Research projects and presentations** — encourage exploration of generalizations and applications of the theorem.
- **Computer experiments** — enable simulations and testing of hypotheses related to the theorem's conditions.

We will discuss the circumstances and ways in which selected active teaching methods can be employed in the educational process, with the Implicit Function Theorem serving as an example. Usually, this theorem is not part of Calculus I; therefore, the proposed approaches are aimed at more advanced students. Nevertheless, the analysis of the basic formulation can be applied to second-year students in mathematics programs. Other aspects are rather intended for graduating students (thesis students) and can, for example, serve as part of a diploma seminar.

2. Analysis of the basic formulation of the Implicit Function Theorem

The basic formulation of the Implicit Function Theorem specifies the conditions under which the equation $f(x, y) = 0$ in real variables has a solution for y with respect to any fixed x from a certain interval. In this way, we obtain a function $y = y(x)$ that satisfies the equation identically on some interval. In other words, the general paradigm depends here on providing conditions on f under which a point $(x_0, y_0) \in f^{-1}(\{0\})$ can be extended to the graph of a function $y = y(x)$ defined locally on a non-degenerate interval I , such that $y(x_0) = y_0$ and the graph $\{(x, y(x)) : x \in I\}$ is contained in $f^{-1}(\{0\})$.

Of course, if y can be solved explicitly in terms of x from the equation $f(x, y) = 0$, the problem is trivial. It only becomes interesting in the context of more complex equations where y cannot be explicitly expressed as a function of x , especially those that define the graph of a function (sic!). An interesting and worthy example of such an equation is $y^5 + 16y - 32x^3 + 32x = 0$, which generates the graph of a certain function in \mathbb{R}^2 and for which an explicit form $y = y(x)$ cannot be provided (for more details, see Krantz and Parks (2013, Ch. 1.1.)). The earliest correct formulation and rigorous proof of the Implicit Function Theorem, meeting modern standards of mathematical rigor, is generally attributed to Dini (1878). Below, we provide this formulation.

THEOREM 2.1 (BASIC IMPLICIT FUNCTION THEOREM)

Let $f : \Omega \subseteq \mathbb{R}^2 \rightarrow \mathbb{R}$ be a C^1 function defined on an open set Ω , and let $(x_0, y_0) \in \Omega$ satisfy

$$f(x_0, y_0) = 0 \quad \text{and} \quad \frac{\partial f}{\partial y}(x_0, y_0) \neq 0.$$

Then there exists a neighborhood $U \subseteq \mathbb{R}$ of x_0 and a unique C^1 function $\varphi : U \rightarrow \mathbb{R}$ such that

$$\varphi(x_0) = y_0 \quad \text{and} \quad f(x, \varphi(x)) = 0 \quad \text{for all } x \in U.$$

Moreover, the derivative of φ is given by

$$\varphi'(x) = -\frac{\frac{\partial f}{\partial x}(x, \varphi(x))}{\frac{\partial f}{\partial y}(x, \varphi(x))}.$$

2.1. Preparation for understanding the formulation and its proof

Let us begin with a proposed analysis of how to organize students' work to successfully achieve the intended goal – understanding the necessity of the theorem, its proof, and its potential for further applications. First, as part of a flipped classroom approach, students should review and analyze the following concepts related to a function $f : \mathbb{R}^n \rightarrow \mathbb{R}$ (or, more generally, to $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$):

- partial derivatives (of any order), and the Jacobian matrix (in the case $f : \mathbb{R}^n \rightarrow \mathbb{R}^m$) – instrumental understanding,
- theorem on calculating partial derivatives of composite functions of several variables – instrumental understanding,
- continuity and differentiability – instrumental and logical understanding,
- continuous differentiability (or, more generally, functions of class C^k) – instrumental understanding,
- local sign preservation theorem – instrumental understanding.

As part of the project-based learning (PBL) component, students can be assigned the following topics:

- geometric interpretation of the equation $f(x, y) = 0$, $(x, y) \in \Omega$, possibly supported by examples using IT tools (e.g., GeoGebra) to visualize the corresponding curves, with an emphasis on cases where the equation cannot be explicitly solved by a function $y = y(x)$; special attention should be paid to equations of the form $f(x, y) = 0$ that either do not generate differentiable functions $y = y(x)$ satisfying the equation, or allow for the existence of more than one such function passing through a given point — this is a suitable moment to engage students' intuition (e.g., through brainstorming).
- implications of the C^1 assumption on f in the context of the condition $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$ – with a focus on building the intuitive and logical understanding of this condition mentioned earlier,
- as an element of instrumental understanding, carrying out calculations of partial derivatives of composite functions of several variables using examples, including derivatives of the form $\frac{\partial f}{\partial x}(x, \varphi(x))$.

After these preparations, we are ready to present the proof for students. Of course, there are many elegant descriptions of this proof, so it is not our intention to reproduce any of them; instead, we will indicate only the most important steps/remarks taking into account all didactic components considered earlier:

- drawing attention to the openness of the domain of the function f and explaining how this property is used when considering one of the possible conditions, e.g., $\frac{\partial f}{\partial y}(x_0, y_0) > 0$ – with reference to the theorem on the local preservation of sign in a neighbourhood in \mathbb{R}^2 , and its relation to generating an appropriate open rectangle $P := (x_0 - \alpha, x_0 + \alpha) \times (y_0 - \beta, y_0 + \beta)$ such that $\frac{\partial f}{\partial y}(x, y) > 0$ for all $(x, y) \in P$,
- the above gives that the function

$$y \mapsto f(x_0, y)$$

is strictly increasing on the interval $(y_0 - \beta, y_0 + \beta)$, which, in view of the assumption $f(x_0, y_0) = 0$, yields, for any $0 < \varepsilon < \beta$, the inequalities

$$f(x_0, y_0 - \varepsilon) < 0 \quad \text{and} \quad f(x_0, y_0 + \varepsilon) > 0.$$

- taking into account the continuity of f in a neighbourhood of the point (x_0, y_0) and the property of local preservation of sign, this in turn implies the existence of a number $0 < \delta < \alpha$ such that

$$f(x, y_0 - \varepsilon) < 0, \quad f(x, y_0 + \varepsilon) > 0$$

for all $x \in (x_0 - \delta, x_0 + \delta)$;

- from the above inequalities, it follows that if $x^* \in (x_0 - \delta, x_0 + \delta)$, then, by continuity and the strict monotonicity of the function $y \mapsto f(x^*, y)$, there exists exactly one $y^* \in (y_0 - \varepsilon, y_0 + \varepsilon)$ such that $f(x^*, y^*) = 0$ (in fact, the Darboux property is used here). Thus, the most important part of the theorem, namely the existence and uniqueness of the implicit function, has been established;
- let us observe that, so far, the assumption that $\frac{\partial f}{\partial x}$ is continuous has not been used; thus, we see that this condition is not necessary to prove the existence of an implicit function. This assumption becomes important in the second part of the proof, in particular for proving that an implicit function $y = \varphi(x)$ is continuously differentiable. In that case, the formula

$$\varphi'(x) = - \frac{\frac{\partial f}{\partial x}(x, \varphi(x))}{\frac{\partial f}{\partial y}(x, \varphi(x))} \quad (1)$$

can be obtained simply by differentiating, with respect to x , the equality $f(x, \varphi(x)) = 0$;

- we should also be aware that the assumption $\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$ is not used merely to justify equation (1), but its main purpose is to guarantee the existence of an implicit function; the possibility of computing this derivative is rather a byproduct of this assumption.

Remark It is worth noting that the assumption $\partial_y f(x_0, y_0) \neq 0$ in the Implicit Function Theorem is *sufficient* but not *necessary*. This may be viewed as part of the brainstorming during discussions with students; for the reader's convenience, we independently provide a few illustrative examples. When at a solution point (x_0, y_0) we have $\partial_y f(x_0, y_0) = 0$, different behaviors of so called level set $f^{-1}(\{0\})$ may occur:

(A) **No locally defined solutions (isolated point).**

Example: $f(x, y) = x^2 + y^2$. The point $(0, 0)$ satisfies $f(0, 0) = 0$ and $\partial_y f(0, 0) = 0$. There is only the trivial solution $(0, 0)$ in the neighborhood, so no function $y = y(x)$ exists locally.

(B) **Multiple solutions (no uniqueness).**

Example: $f(x, y) = y^2 - x^2$. At $(0, 0)$, we have $f = 0$ and $\partial_y f = 0$. For $x \in \mathbb{R}$, two differentiable functions exist: $y = \pm x$ (but 4 continuous solutions, and infinitely many non-continuous).

(C) **Unique solution exists but is not smooth.**

Example: $f(x, y) = y^3 - x$. At $(0, 0)$, $\partial_y f(0, 0) = 3y^2|_{y=0} = 0$. The solution is unique and given by $y(x) = \sqrt[3]{x}$, but this function is not differentiable at $x = 0$.

(D) **Solutions exist only on one side.**

Example: $f(x, y) = y^2 + x$. At $(0, 0)$, $f = 0$ and $\partial_y f = 0$. For $x < 0$, there are real solutions $y = \pm\sqrt{-x}$, while for $x > 0$ there are no real solutions.

(E) **Unique smooth function.**

There are also situations where, despite the failure of the condition $\partial_y f(x_0, y_0) \neq 0$, one still obtains a *unique and differentiable* solution $y = y(x)$ in a neighborhood of the point.

Example: $f(x, y) = y^3 + x^3$. We have $f(0, 0) = 0$ and $\partial_y f(0, 0) = 3y^2|_{y=0} = 0$. The equation $f(x, y) = 0$ is equivalent to $y^3 = -x^3$, hence (over \mathbb{R}) uniquely

$$y(x) = -x.$$

This branch is smooth and satisfies $y'(0) = -1$, even though the condition $\partial_y f(x_0, y_0) \neq 0$ is violated.

General comment. The condition $\partial_y f \neq 0$ in the Implicit Function Theorem is a *sufficient* but not a *necessary* condition. It may fail while a unique and even smooth solution branch $y = y(x)$ still exists (as in the example above). One of the fundamental applications of the Implicit Function Theorem is that it allows for a *partial analysis of the behavior of an implicit function without having its explicit formula*. In particular, this includes the possibility of investigating local extrema, inflection points, and monotonicity properties of implicitly defined functions.

Illustration. Using earlier prepared methods of differentiation of a composite function of many variables, the following task can be given as an exercise. Suppose

that $f(x, y) = 0$ defines locally a C^2 function $y = y(x)$. Without knowing $y(x)$ explicitly, check that

$$y'(x) = -\frac{\partial f}{\partial x}(x, y(x)) / \frac{\partial f}{\partial y}(x, y(x)),$$

whenever $\frac{\partial f}{\partial y}(x, y(x)) \neq 0$. Similarly, prove that the second derivative can be expressed as:

$$y''(x) = -\frac{\frac{\partial^2 f}{\partial x^2}(x, y(x)) + 2\frac{\partial^2 f}{\partial x \partial y}(x, y(x))y'(x) + \frac{\partial^2 f}{\partial y^2}(x, y(x))(y'(x))^2}{\frac{\partial f}{\partial y}(x, y(x))}.$$

Thus, one can analyze the local extrema of $y = y(x)$ without solving explicitly for y . By computing higher-order derivatives, one can also identify inflection points of $y = y(x)$. Working on these problems allows students to integrate both logical and instrumental understanding in their mathematical development. The teacher should provide a list of interesting examples illustrating the determination of local extrema or inflection points for formulas $f(x, y) = 0$ that cannot be explicitly solved for y ; the case teaching method can be applied here.

Now, we will conduct a detailed analysis of how to effectively engage students in understanding the necessity of the Implicit Function Theorem. This includes exploring the historical and mathematical motivations behind its development. Such an approach not only fosters deeper conceptual understanding but also promotes critical thinking and appreciation for the structure of mathematical reasoning.

2.2. The Kepler equation as an example of the necessity of the Implicit Function Theorem

There are many examples of problems involving the study of implicit functions. Let us focus on one of the most historically significant implicit equations that cannot be solved explicitly – the Kepler equation, arising in celestial mechanics (see Spivak (1965), Krantz and Parks (2013)). It can be given to students as part of flipped education and also includes a physical component that requires additional analysis and understanding (so, first, the student must be familiar with the description of considered physical phenomenon and quantities).

Consider the equation

$$M = E - e \sin E, \tag{2}$$

where:

- $M \in [0, 2\pi]$ is the *mean anomaly* (a measure of time),
- E is the *eccentric anomaly* (a geometric angle),
- $e \in [0, 1)$ is the *eccentricity* of the orbit.

Equation (2) cannot be solved explicitly for E in terms of M . Nevertheless, astronomers and physicists need to understand how E depends on M . This provides a perfect opportunity to apply the Implicit Function Theorem.

Implicit formulation and analysis

Let us define the function:

$$f(M, E) := E - e \sin E - M.$$

We want to solve the equation $f(M, E) = 0$ for E , given M .

Regularity. The function f is smooth: $f \in C^\infty(\mathbb{R}^2)$. Its partial derivative with respect to E is:

$$\frac{\partial f}{\partial E} = 1 - e \cos E.$$

Since $\cos E \in [-1, 1]$ and $0 \leq e < 1$, we have $1 - e \cos E > 0$ for all E .

Conclusion. By the Implicit Function Theorem, for every $M \in \mathbb{R}$ there exists a unique differentiable function $E = \varphi(M)$ such that $f(M, \varphi(M)) = 0$. Moreover, this function is differentiable, and:

$$\frac{dE}{dM} = \frac{1}{1 - e \cos E}.$$

This derivative gives valuable insight into how the eccentric anomaly E changes with respect to time M , even though E itself cannot be expressed in terms of elementary functions.

Historical and educational significance

Before the advent of modern analysis, equations like (2) posed serious challenges. Astronomers used geometric constructions and tables, but could not rigorously justify the differentiability or even the existence of a function $E(M)$.

The Implicit Function Theorem resolved this, providing a firm theoretical foundation for problems involving hidden dependencies between variables. The Kepler equation thus serves as a powerful educational example:

- it illustrates the need for the Implicit Function Theorem in real-world scientific problems;
- it offers an opportunity to analyze differentiability without having a closed-form solution;
- it connects historical context with modern mathematical tools.

This example is particularly suitable for active learning activities, such as numerical approximation of $E(M)$, visualization of the function's graph, or investigating the behavior of its derivative.

Numerical analysis of Newton's method for the Kepler equation

The Kepler equation expresses implicitly as

$$f(E) = E - e \sin E - M = 0.$$

Newton iteration Starting from an initial guess E_0 , the Newton–Raphson iteration is

$$E_{n+1} = E_n - \frac{f(E_n)}{f'(E_n)} = E_n - \frac{E_n - e \sin E_n - M}{1 - e \cos E_n}.$$

Smoothness and uniqueness The function f is infinitely differentiable (C^∞) in \mathbb{R} , and

$$f'(E) = 1 - e \cos E \geq 1 - e > 0,$$

which guarantees that for each fixed M , the equation $f(E) = 0$ has a unique solution $E = \varphi(M)$.

Local quadratic convergence Newton’s method enjoys *local quadratic convergence*, meaning it converges quadratically if the initial guess E_0 is sufficiently close to the true root E^* .

Formally, if x^* is the true solution and x_n the n -th iterate, then there exists a neighborhood around x^* and a constant $C > 0$ such that for all n sufficiently large,

$$|x_{n+1} - x^*| \leq C|x_n - x^*|^2.$$

This means that the number of correct digits roughly doubles at each step once close enough to the root, resulting in very rapid convergence.

Kantorovich condition A sufficient condition for convergence can be derived from the Kantorovich theorem: let

$$L = \sup_E |f''(E)| = \sup_E |e \sin E| = e,$$

and

$$m = \inf_E |f'(E)| = 1 - e.$$

Define

$$\alpha := \frac{|f(E_0)|}{m} = \frac{|E_0 - e \sin E_0 - M|}{1 - e}.$$

If

$$L\alpha = e \frac{|f(E_0)|}{1 - e} \leq \frac{1}{2},$$

then the Newton iteration starting at E_0 is guaranteed to converge quadratically to the unique root E^* .

Interpretation for the common initial guess Choosing $E_0 = M$, we have

$$f(M) = M - e \sin M - M = -e \sin M,$$

so

$$|f(M)| \leq e.$$

Hence, the Kantorovich criterion is satisfied if

$$e \frac{e}{1-e} = \frac{e^2}{1-e} \leq \frac{1}{2}.$$

This inequality holds for all eccentricities e such that

$$0 \leq e \leq \frac{1}{2}.$$

Therefore, for $e \leq 0.5$, the Newton iteration with initial guess $E_0 = M$ converges quadratically.

Behavior for larger eccentricities For eccentricities $e > 0.5$, the initial guess $E_0 = M$ may fail the Kantorovich condition, and convergence is no longer guaranteed theoretically. In practice, Newton's method may still converge, but it is advisable to use improved initial guesses (e.g., Danby or Mikkola approximations) or globalized Newton variants to ensure robust convergence.

Numerical illustration For example, with $e = 0.5$, $M = 1.0$, and $E_0 = M$, the Newton iterations produce:

$$E_0 = 1.0000000,$$

$$E_1 = 1.5764694,$$

$$E_2 = 1.5002083,$$

$$E_3 = 1.4987017,$$

$$E_4 = 1.4987011,$$

demonstrating rapid quadratic convergence.

Summary:

- The Implicit Function Theorem guarantees existence and uniqueness of E ;
- Newton's method is an effective tool to compute E numerically;
- Convergence is local; Kantorovich's theorem provides a sufficient condition;
- For moderate eccentricities ($e \leq 0.5$), the natural initial guess $E_0 = M$ suffices;
- For higher eccentricities, better starters or safeguarded methods are recommended.

Pedagogical suggestions Students can implement Newton's method for various e , verify Kantorovich conditions numerically, and compare convergence behavior with different initial guesses.

3. Selected generalizations of the basic version of the Implicit Function Theorem

Let us begin with the most natural generalization, namely by increasing the number of independent variables (see Krantz and Parks (2013, Thm. 3.2.1)).

THEOREM 3.1

Let $m \in \mathbb{N}$, $m \geq 2$. If $W \subset \mathbb{R}^m$ is open, $F : W \rightarrow \mathbb{R}$ is continuously differentiable, and $p = (p', q) \in \mathbb{R}^{m-1} \times \mathbb{R}$ with $p \in W$ is a point for which

$$F(p) = 0 \quad \text{and} \quad \frac{\partial F}{\partial x_m}(p) \neq 0,$$

then there exists an open set $W' \subset \mathbb{R}^{m-1}$ with $p' \in W'$ and a unique continuously differentiable function $\psi : W' \rightarrow \mathbb{R}$ such that $q = \psi(p')$ and

$$F(x', \psi(x')) = 0 \quad \text{for all } x' \in W'.$$

Interestingly, this can be given as a simple exercise for students (as part of flipped education), asking them to verify that all proof steps can be repeated exactly as in Theorem 2.1. This is a straightforward instance of comparative analysis, where the inclusion of additional dependent variables does not affect the method (and intuition) of proof. Similarly, students can verify (with some additional calculations), using the differentiation formulas for composite functions, that:

$$\frac{\partial \psi}{\partial x_j} = - \frac{\frac{\partial F}{\partial x_j}}{\frac{\partial F}{\partial x_m}}, \quad j = 1, 2, \dots, m-1.$$

The next generalization we consider is somewhat different, as the values of the function now lie in \mathbb{R}^2 . We highlight this particular case to provide an elementary explanation (and some intuition) for the general situation of this type (see Krantz and Parks (2013, Thm. 1.3.4)). This step seems to be the most important to catch the general paradigm of the Implicit Function Theorem.

THEOREM 3.2

Let $U \subset \mathbb{R}^3$ be an open set, and assume that

$$F = (F_1, F_2) : U \rightarrow \mathbb{R}^2$$

is continuously differentiable. Further, assume that at a point $a = (a_1, a_2, a_3) \in U$, it holds that $F(a) = 0$ and

$$\det \begin{pmatrix} \frac{\partial F_1}{\partial x_2}(a) & \frac{\partial F_1}{\partial x_3}(a) \\ \frac{\partial F_2}{\partial x_2}(a) & \frac{\partial F_2}{\partial x_3}(a) \end{pmatrix} \neq 0.$$

Then there exists a product neighborhood $V \times W \subset U$, with $a_1 \in V \subset \mathbb{R}$ and $(a_2, a_3) \in W \subset \mathbb{R}^2$, and a unique continuously differentiable mapping

$$\phi = (\phi_1, \phi_2) : V \rightarrow W$$

such that $(a_2, a_3) = \phi(a_1)$ and, for each $x_1 \in V$, it holds that

$$F[x_1, \phi_1(x_1), \phi_2(x_1)] = 0.$$

In simple terms, Theorem 3.2 can be thought of like this: suppose we have two equations with three variables x_1, x_2, x_3 :

$$F_1(x_1, x_2, x_3) = 0, \quad F_2(x_1, x_2, x_3) = 0.$$

We want to solve for x_2 and x_3 in terms of the remaining variable x_1 . Ideally, x_2 and x_3 would be smooth functions of x_1 .

The condition that guarantees this is that the derivative with respect to the variables we want to solve for—here, x_2 and x_3 —is invertible. In this case, the derivative is a linear map from \mathbb{R}^2 to \mathbb{R}^2 , which is invertible if and only if its determinant is nonzero. This is exactly the higher-dimensional analogue of the condition

$$\frac{\partial f}{\partial y}(x_0, y_0) \neq 0$$

in Theorem 2.1. Before we present the key steps of the proof (building in particular an intuitive understanding), let us recall that the main idea relies on the appropriate use of the Inverse Function Theorem, which we state here for the reader's convenience (in this particular case).

THEOREM 3.3 (INVERSE FUNCTION THEOREM, $\mathbb{R}^2 \rightarrow \mathbb{R}^2$)

Let $G : U \subset \mathbb{R}^2 \rightarrow \mathbb{R}^2$ be continuously differentiable function on an open set U , and let $b \in U$ be such that

$$\det \begin{pmatrix} \frac{\partial G_1}{\partial y_1}(b) & \frac{\partial G_1}{\partial y_2}(b) \\ \frac{\partial G_2}{\partial y_1}(b) & \frac{\partial G_2}{\partial y_2}(b) \end{pmatrix} \neq 0.$$

Then there exist a neighborhood V of the point b and a neighborhood W of the point $G(b)$, and a **unique differentiable function**

$$G^{-1} : W \rightarrow V$$

such that for each $z \in W$ there exists exactly one $y \in V$ satisfying $G(y) = z$, and

$$G^{-1}(G(y)) = y, \quad G(G^{-1}(z)) = z.$$

So, first, as part of group work (PBL), students should recall the formulation and analyze the precise proof of the above theorem, and possibly as part of flipped class the notion and properties of a diffeomorphism. Now let us focus on the key steps in the proof of Theorem 3.2, in order to provide a clear picture corresponding to the proof of Theorem 2.1.

Let $a = (a_1, a_2, a_3) \in U$ and define

$$J = \begin{pmatrix} \frac{\partial F_1}{\partial x_2}(a) & \frac{\partial F_1}{\partial x_3}(a) \\ \frac{\partial F_2}{\partial x_2}(a) & \frac{\partial F_2}{\partial x_3}(a) \end{pmatrix}, \quad \det(J) \neq 0.$$

1. For the fixed $x_1 = a_1$, consider the function

$$G(x_2, x_3) := F(a_1, x_2, x_3).$$

2. Since $\det(J) \neq 0$, the Inverse Function Theorem guarantees the existence of a neighborhood $W \subset \mathbb{R}^2$ of (a_2, a_3) such that

$$G : W \rightarrow G(W) \subset \mathbb{R}^2$$

is a diffeomorphism. In particular, there exists a unique continuously differentiable mapping

$$G^{-1} : G(W) \rightarrow W$$

satisfying $G(G^{-1}(y)) = y$ for all $y \in G(W)$.

3. Now, for each x_1 sufficiently close to a_1 , consider the function

$$G_{x_1}(x_2, x_3) := F(x_1, x_2, x_3), \quad (x_2, x_3) \in \mathbb{R}^2.$$

By continuity of the partial derivatives, the Jacobian of G_{x_1} with respect to (x_2, x_3) at (a_2, a_3) remains invertible for x_1 near a_1 . Therefore, by the Inverse Function Theorem, there exists a unique continuously differentiable mapping

$$\phi(x_1) = (\phi_1(x_1), \phi_2(x_1))$$

such that

$$G_{x_1}(\phi(x_1)) = F(x_1, \phi_1(x_1), \phi_2(x_1)) = 0$$

for all x_1 sufficiently close to a_1 . In other words, we define the implicit function ϕ by

$$\phi(x_1) := G_{x_1}^{-1}((0, 0)),$$

where $G_{x_1}^{-1}$ is the local inverse guaranteed by the Inverse Function Theorem.

4. By construction, $\phi(a_1) = (a_2, a_3)$ and

$$F[x_1, \phi_1(x_1), \phi_2(x_1)] = 0$$

for all x_1 in a small neighborhood of a_1 . Continuity and differentiability of ϕ follow from the differentiability of F and the local inverse $G_{x_1}^{-1}$.

5. Uniqueness of ϕ comes from the uniqueness provided by the Inverse Function Theorem: for each x_1 sufficiently close to a_1 , there is exactly one $(x_2, x_3) \in W$ such that $F[x_1, x_2, x_3] = 0$.

All the above steps should be discussed with students, including possible completions; in particular, it should be emphasized at the beginning that the function G is defined in terms of the dependent variables x_2 and x_3 , and that the setting \mathbb{R}^3 is the minimal nontrivial case where two equations define two dependent variables as functions of a single independent variable.

Now it is time for the natural generalization of Theorem 3.2, allowing for any finite number of independent and dependent variables, followed by a comparative analysis of its proof with that of Theorem 3.2.

THEOREM 3.4

Let $F : U \subset \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ be continuously differentiable on an open set U , and write the variables as $(x, y) \in \mathbb{R}^n \times \mathbb{R}^m$. Suppose that at a point $(a, b) \in U$, we have

$$F(a, b) = 0,$$

and that the Jacobian matrix, with respect to the y -variables,

$$D_y F(a, b) = \begin{pmatrix} \frac{\partial F_1}{\partial y_1}(a, b) & \cdots & \frac{\partial F_1}{\partial y_m}(a, b) \\ \vdots & \ddots & \vdots \\ \frac{\partial F_m}{\partial y_1}(a, b) & \cdots & \frac{\partial F_m}{\partial y_m}(a, b) \end{pmatrix},$$

is invertible (i.e., $\det D_y F(a, b) \neq 0$).

Then there exist neighborhoods $V \subset \mathbb{R}^n$ of a and $W \subset \mathbb{R}^m$ of b , and a unique continuously differentiable mapping

$$\varphi : V \rightarrow W$$

such that

$$\varphi(a) = b, \quad \text{and} \quad F(x, \varphi(x)) = 0 \quad \text{for all } x \in V.$$

Moreover, φ is of class C^k if F is C^k for $k \geq 1$.

REMARK 3.5 (COMPARATIVE ANALYSIS OF PROOFS) *The proofs of the Implicit Function Theorem in the cases $F : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ and the general $F : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ follow the same underlying strategy, which is based on the Inverse Function Theorem. The following steps are hints for students on how to organize the general version of the proof; students can employ either technique: a flipped classroom or PBL. In the case of PBL, we may require a final presentation with all details and discuss possible differences in calculations in the multidimensional case (checking logical and instrumental understanding).*

Step 1: Choice of auxiliary map. *In both cases, we isolate the dependent variables and define an auxiliary mapping that combines the independent variables with the original function F . For $F : \mathbb{R}^3 \rightarrow \mathbb{R}^2$, this map is $G(x_2, x_3) := F(a_1, x_2, x_3)$ with x_1 fixed, whereas in the general case, we set $G(x, y) := (x, F(x, y))$. The purpose of G is to encode the system in a form suitable for applying the Inverse Function Theorem.*

Step 2: Verification of Jacobian invertibility. *Both proofs rely crucially on the invertibility of the Jacobian matrix with respect to the dependent variables. In the $\mathbb{R}^3 \rightarrow \mathbb{R}^2$ case, this is the 2×2 matrix of derivatives with respect to (x_2, x_3) . In the general case, it is the $m \times m$ block $D_y F(a, b)$. In both situations, a nonzero determinant ensures local invertibility of G .*

Step 3: Application of the Inverse Function Theorem. *Once invertibility is established, the Inverse Function Theorem yields a locally defined C^1 inverse G^{-1} , which is then used to construct the implicit function ϕ or φ . The structure of the argument is identical: for each value of the independent variables near the given point, there is a unique choice of dependent variables solving $F = 0$.*

Step 4: Conclusion. Both proofs conclude by showing that the resulting implicit function is C^1 (or C^k if F is C^k), satisfies $F(x, \phi(x)) = 0$, and is unique in a neighborhood. The general proof is a natural extension of the $\mathbb{R}^3 \rightarrow \mathbb{R}^2$ case, with the same logic applied to n independent and m dependent variables.

Comparison of implicit differentiation

Let us compare the computation of the derivative of the implicit function in the two cases discussed above.

1. Special case: $F : \mathbb{R}^3 \rightarrow \mathbb{R}^2$ We have $F(x_1, x_2, x_3) = (F_1, F_2)$ and the implicit functions ϕ_1, ϕ_2 such that

$$x_2 = \phi_1(x_1), \quad x_3 = \phi_2(x_1), \quad F(x_1, \phi_1(x_1), \phi_2(x_1)) = 0.$$

Differentiating with respect to x_1 and applying the chain rule gives

$$\begin{pmatrix} \frac{\partial F_1}{\partial x_2} & \frac{\partial F_1}{\partial x_3} \\ \frac{\partial F_2}{\partial x_2} & \frac{\partial F_2}{\partial x_3} \end{pmatrix} \begin{pmatrix} \phi_1'(x_1) \\ \phi_2'(x_1) \end{pmatrix} = - \begin{pmatrix} \frac{\partial F_1}{\partial x_1} \\ \frac{\partial F_2}{\partial x_1} \end{pmatrix},$$

so that

$$\begin{pmatrix} \phi_1'(x_1) \\ \phi_2'(x_1) \end{pmatrix} = - \begin{pmatrix} \frac{\partial F_1}{\partial x_2} & \frac{\partial F_1}{\partial x_3} \\ \frac{\partial F_2}{\partial x_2} & \frac{\partial F_2}{\partial x_3} \end{pmatrix}^{-1} \begin{pmatrix} \frac{\partial F_1}{\partial x_1} \\ \frac{\partial F_2}{\partial x_1} \end{pmatrix}.$$

2. General case: $F : \mathbb{R}^n \rightarrow \mathbb{R}^m$ Suppose $F(x, y) = 0$ with $x \in \mathbb{R}^n$, $y \in \mathbb{R}^m$, and let $\varphi : \mathbb{R}^n \rightarrow \mathbb{R}^m$ be the implicit function such that $y = \varphi(x)$ and $F(x, \varphi(x)) = 0$. Differentiating with respect to x gives the matrix equation

$$\frac{\partial F}{\partial y}(x, \varphi(x)) D\varphi(x) + \frac{\partial F}{\partial x}(x, \varphi(x)) = 0,$$

so that

$$D\varphi(x) = - \left(\frac{\partial F}{\partial y}(x, \varphi(x)) \right)^{-1} \frac{\partial F}{\partial x}(x, \varphi(x)).$$

One can see that the formulas coincide in structure. It is worth examining students using concrete examples (in particular for $n, m \geq 2$) such as

$$F(x_1, x_2, y_1, y_2, y_3) = \begin{cases} x_1 + y_1 - y_2^2 + y_3 = 0, \\ x_2^2 + y_2 - y_1 y_3 = 0, \\ x_1 x_2 + y_3 - y_1^2 = 0, \end{cases}$$

where one has to check that on some neighbourhood of $(0, 0, 0, 0, 0) \in \mathbb{R}^5$, there exists an implicit function $\varphi = (\varphi_1(x_1, x_2), \varphi_2(x_1, x_2), \varphi_3(x_1, x_2))$, and find its derivative at $(0, 0)$.

As the last generalization, we give the counterpart of Theorem 3.4 in Banach spaces (which also admits infinite dimensional case). Compare the formulation below to Krantz and Parks (2013, Thm. 3.4.10):

THEOREM 3.6 (IMPLICIT FUNCTION THEOREM IN BANACH SPACES)

Let X , Y , and Z be Banach spaces, and let $F : U \subset X \times Y \rightarrow Z$ be a continuously Fréchet differentiable mapping on an open set U . Suppose that $(x_0, y_0) \in U$ satisfies

$$F(x_0, y_0) = 0,$$

and that the partial derivative with respect to y , $D_y F(x_0, y_0) : Y \rightarrow Z$, is a continuous linear operator with a continuous inverse. Then there exist neighborhoods $V \subset X$ of x_0 and $W \subset Y$ of y_0 , and a unique continuously Fréchet differentiable function

$$\varphi : V \rightarrow W$$

such that

$$\varphi(x_0) = y_0 \quad \text{and} \quad F(x, \varphi(x)) = 0 \quad \text{for all } x \in V.$$

Moreover, the derivative of φ is given by

$$D\varphi(x) = -(D_y F(x, \varphi(x)))^{-1} D_x F(x, \varphi(x)).$$

To analyze this result, we must first ensure that students have the appropriate background in the fundamentals of functional analysis, as outlined below (including a review of all related concepts that appear – this can generally be organized as a group project with a final presentation during the seminar):

- (i) **Bounded inverse / Open mapping theorem:** If $T : Y \rightarrow Z$ is a bounded linear bijection between Banach spaces, then T^{-1} exists and is bounded. (This relies on completeness of Y, Z);
- (ii) **Inverse Function Theorem (Banach version):** If G is C^1 on a Banach space and $DG(p)$ is a bounded linear isomorphism, then G is a C^1 local diffeomorphism near p ;
- (iii) Existence of Fréchet derivatives and usual rules (chain rule, block derivative) hold in Banach spaces.

Now, let us focus on a proposed sketch of the proof, emphasizing where and how the above facts are applied. It is also valuable for students to compare these steps with the corresponding ones in Theorem 3.4.

Step 1 — Define an auxiliary map.

Define

$$G : U \rightarrow X \times Z, \quad G(x, y) := (x, F(x, y)).$$

Then G is C^1 and $G(x_0, y_0) = (x_0, 0)$.

Step 2 — Derivative of G at (x_0, y_0) .

In block form (with I_X the identity on X),

$$DG(x_0, y_0) = \begin{pmatrix} I_X & 0 \\ D_x F(x_0, y_0) & D_y F(x_0, y_0) \end{pmatrix} = \begin{pmatrix} I_X & 0 \\ B & A \end{pmatrix},$$

where $B := D_x F(x_0, y_0) : X \rightarrow Z$ and $A := D_y F(x_0, y_0) : Y \rightarrow Z$.

In operator terms,

$$DG(x_0, y_0)(h, k) = (h, Bh + Ak), \quad (h, k) \in X \times Y.$$

Since A is invertible, the triangular operator $DG(x_0, y_0)$ is also invertible: given $(u, v) \in X \times Z$, one solves

$$h = u, \quad k = A^{-1}(v - Bu).$$

Thus, $DG(x_0, y_0)$ is a bounded isomorphism, and by (i), its inverse is bounded.

Step 3 — Apply the Inverse Function Theorem.

By (ii), there exist neighborhoods $V \times W \subset X \times Y$ of (x_0, y_0) and $V' \times W' \subset X \times Z$ of $(x_0, 0)$ such that

$$G : V \times W \rightarrow V' \times W'$$

is a C^1 diffeomorphism with inverse G^{-1} .

Step 4 — Define the implicit function.

For each $x \in V'$ close to x_0 , consider $(x, 0) \in V' \times W'$. Put

$$G^{-1}(x, 0) = (x, \varphi(x)).$$

Then $F(x, \varphi(x)) = 0$, and $\varphi : V' \rightarrow W$ is C^1 with $\varphi(x_0) = y_0$. Uniqueness follows from local invertibility of G .

Step 5 — Formula for the derivative.

Differentiating $F(x, \varphi(x)) \equiv 0$ and applying the chain rule gives

$$D_x F(x, \varphi(x)) + D_y F(x, \varphi(x)) D\varphi(x) = 0.$$

Thus

$$D\varphi(x) = -(D_y F(x, \varphi(x)))^{-1} D_x F(x, \varphi(x)).$$

Remark on the role of completeness.

In the general Banach space version of the Implicit Function Theorem, the assumption that the spaces are complete is crucial. Completeness is used in the Bounded Inverse / Open Mapping Theorem to guarantee that a bounded linear bijection has a bounded inverse, and in the Inverse Function Theorem (or contraction mapping argument) to ensure local solvability and convergence. Without completeness, these key functional-analytic results may fail in infinite-dimensional spaces, and the standard proof may not go through. However, in the finite-dimensional case, this assumption is unnecessary. Every finite-dimensional normed space is automatically complete, so the theorems hold without explicitly requiring completeness.

For the reader's convenience, we provide below an example illustrating the use of Theorem 3.6. It can be analyzed in detail as part of a case study demonstrating how this theorem works. Moreover, we also provide some additional examples for students' self-study.

Example of use the Implicit Function Theorem in an infinite-dimensional space

We consider the classical infinite-dimensional Banach space of continuous functions

$$X = C[0.1, 1]$$

with the supremum norm

$$\|f\|_\infty = \sup_{x \in [0.1, 1]} |f(x)|.$$

Example

Define the mapping

$$F : C[0.1, 1] \times \mathbb{R} \rightarrow C[0.1, 1], \quad F(f, \lambda)(x) = f(x)^2 - \lambda x, \quad x \in [0.1, 1].$$

We want to find an implicit function $f = f(\lambda)$ such that

$$F(f(\lambda), \lambda) = 0.$$

Step 1: Check the hypotheses of the Implicit Function Theorem

Choose a point

$$\lambda_0 = 1, \quad f_0(x) = \sqrt{x}.$$

Then

$$F(f_0, \lambda_0)(x) = f_0(x)^2 - \lambda_0 x = x - x = 0 \quad \text{for all } x \in [0.1, 1].$$

The Fréchet derivative with respect to f is

$$\begin{aligned} D_f F(f_0, \lambda_0)[h](x) &= \lim_{\epsilon \rightarrow 0} \frac{F(f_0 + \epsilon h, \lambda_0)(x) - F(f_0, \lambda_0)(x)}{\epsilon} \\ &= 2f_0(x)h(x) = 2\sqrt{x}h(x). \end{aligned}$$

This operator

$$D_f F(f_0, \lambda_0) : C[0.1, 1] \rightarrow C[0.1, 1]$$

is linear and bounded.

Step 2: Check invertibility of the operator

We need $D_f F(f_0, \lambda_0)$ to be invertible. It acts as

$$h \mapsto 2\sqrt{x}h(x),$$

so its inverse is

$$(D_f F(f_0, \lambda_0))^{-1}[g](x) = \frac{1}{2\sqrt{x}}g(x), \quad x \in [0.1, 1].$$

Since $x \in [0.1, 1]$, the operator is continuous and invertible.

Step 3: Compute the derivative of the implicit function

The Implicit Function Theorem guarantees that there exists $f(\lambda)$ near λ_0 such that

$$F(f(\lambda), \lambda) = 0.$$

Its Fréchet derivative with respect to λ is given by

$$f'(\lambda_0) = -(D_f F(f_0, \lambda_0))^{-1} \circ D_\lambda F(f_0, \lambda_0),$$

where

$$D_\lambda F(f_0, \lambda_0)[\mu](x) = \left. \frac{d}{d\epsilon} F(f_0, \lambda_0 + \epsilon\mu)(x) \right|_{\epsilon=0} = -\mu x.$$

Hence,

$$f'(\lambda_0)[\mu](x) = -\frac{1}{2\sqrt{x}}(-\mu x) = \frac{\mu x}{2\sqrt{x}} = \frac{\mu\sqrt{x}}{2}.$$

For $\mu = 1$, we have

$$f'(\lambda_0)(x) = \frac{\sqrt{x}}{2}.$$

Conclusion

The implicit function $f(\lambda)$ exists near $\lambda_0 = 1$, and its Fréchet derivative is equal to

$$f'(\lambda_0)(x) = \frac{\sqrt{x}}{2}.$$

Let us note that $f'(\lambda_0)$ is a function in this case (not a number).

Observe that each step in the above guide can be effectively supported by active learning methods. For example, step 1 can involve brainstorming to find the initial point, while practical competencies, such as computing the derivative independently, can be reinforced through flipped classroom activities or a small group project, emphasizing that the expression used here corresponds to a partial derivative with respect to the function variable f and indeed represents a proper Fréchet derivative, as required by its definition. Step 3 can be implemented as a form of group work with a discussion on the proper interpretation of the respective derivatives.

Examples for independent practice with hints. The following examples are proposed for independent work with the Implicit Function Theorem in infinite-dimensional Banach spaces. Students are encouraged to find the implicit function $f(\lambda)$ and compute its Fréchet derivative with respect to λ . Short hints are provided to guide the solution process.

1. Consider the functional

$$F : C[0.1, 1] \times \mathbb{R} \rightarrow C[0.1, 1], \quad F(f, \lambda)(x) = f(x)^2 - \lambda x, \quad x \in [0.1, 1].$$

Find the implicit function $f(\lambda)$ satisfying $F(f(\lambda), \lambda) = 0$ and compute its derivative at $\lambda_0 = 1$.

- **Hint:** Choose $f_0(x) = \sqrt{x}$ as the initial function. Compute the Fréchet derivative with respect to f , check invertibility of the corresponding operator, then apply the Implicit Function Theorem to find $f'(\lambda_0)$.

2. Consider the functional

$$F : C[0.1, 1] \times \mathbb{R} \rightarrow C[0.1, 1], \quad F(f, \lambda)(x) = f(x)^3 - \lambda \sin(x), \quad x \in [0.1, 1].$$

Find the implicit function $f(\lambda)$ satisfying $F(f(\lambda), \lambda) = 0$ and compute its derivative at $\lambda_0 = 1$.

- **Hint:** Take $f_0(x) = (\sin x)^{1/3}$. Compute the Fréchet derivative with respect to f as an operator of multiplication, check invertibility, then determine $f'(\lambda_0)$ using the formula from the Implicit Function Theorem.

3. Consider the functional

$$F : C[0.1, 1] \times \mathbb{R} \rightarrow C[0.1, 1], \quad F(f, \lambda)(x) = e^{f(x)} - \lambda x^2, \quad x \in [0.1, 1].$$

Find the implicit function $f(\lambda)$ satisfying $F(f(\lambda), \lambda) = 0$ and compute its derivative at $\lambda_0 = 1$.

- **Hint:** Take $f_0(x) = \ln(x^2)$. Compute the Fréchet derivative with respect to f , verify invertibility of the multiplication operator, then calculate $f'(\lambda_0)$ using the Implicit Function Theorem.

4. Consider the functional

$$F : \ell^2 \times \mathbb{R} \rightarrow \ell^2, \quad F(x, \lambda) = (x_n^2 - \lambda/n)_{n \in \mathbb{N}}.$$

Find the implicit function $x(\lambda)$ satisfying $F(x(\lambda), \lambda) = 0$ and compute its Fréchet derivative at $\lambda_0 = 1$.

- **Hint:** Choose $x_0 = (1/\sqrt{n})_{n \in \mathbb{N}}$. Compute the Fréchet derivative with respect to x , which acts componentwise: $h \mapsto (2x_{0,n}h_n)_{n \in \mathbb{N}}$. Verify that it is invertible and then use the Implicit Function Theorem to find $x'(\lambda_0)$.

5. Consider the Banach space $X = \ell^2$ of square-summable sequences with the standard norm

$$\|x\|_{\ell^2} = \left(\sum_{n=1}^{\infty} |x_n|^2 \right)^{1/2}.$$

Define the operator

$$F : \ell^2 \times \mathbb{R} \rightarrow \ell^2, \quad (F(x, \lambda))_n = \left(1 + \frac{1}{n^2} \right) x_n - \frac{\lambda}{n^2}, \quad n \geq 1.$$

We are looking for a function $x = x(\lambda) \in \ell^2$ such that

$$F(x(\lambda), \lambda) = 0.$$

Hint for the solution:

- Compute the derivative of F with respect to x : $(D_x F(x_0, \lambda_0)[h])_n = (1 + \frac{1}{n^2})h_n$.
- Compute the derivative of F with respect to λ : $(D_\lambda F(x_0, \lambda_0)[\mu])_n = -\frac{\mu}{n^2}$.
- Use the Implicit Function Theorem formula

$$x'(\lambda_0) = -(D_x F(x_0, \lambda_0))^{-1} \circ D_\lambda F(x_0, \lambda_0)$$

to find the derivative of the implicit function $x(\lambda)$.

- Verify that $x'(\lambda_0) \in \ell^2$.

The above examples are the culmination of the presented variants of the Implicit Function Theorem. In particular, they illustrate the high level of mathematical abstraction that can be achieved using the outlined layered approach, starting from the “ $\mathbb{R}^2 \rightarrow \mathbb{R}$ ” case and extending to the generalization “ $X \times Y \rightarrow Z$ ” for Banach spaces. More examples of this kind can be found in Lang (1993) and Zeidler (1995).

4. Selected applications in higher mathematics

So far, we have considered direct applications of the Implicit Function Theorem according to its various formulations. It is important, especially for more advanced students, to review some significant results in higher mathematics where this theorem plays a crucial role, in order to emphasize its central importance. At present, there exists a vast number of generalizations and variations of the theorem, including many results concerning the global solvability of equations of the form $f(x, y) = 0$ in abstract spaces (see, for instance, Krantz and Parks (2013, Ch. 5)). In this paper, however, we will focus on certain classical theorems that are usually presented during undergraduate or graduate studies. This approach seems more useful, since students often fail to appreciate how fundamental the Implicit Function Theorem is in classical mathematical analysis. Long-term observations suggest that, in general, students are familiar with only one typical application of the theorem.

One of the first applications of the Implicit Function Theorem is the proof of the existence and uniqueness of solutions to ordinary differential equations (the Cauchy problem). Consider the equation

$$y'(t) = f(t, y(t)), \quad y(t_0) = y_0.$$

The fundamental Picard–Lindelöf theorem ensures existence and uniqueness of a solution when f is Lipschitz continuous. However, earlier classical approaches (including those of Cauchy and Peano) made use of the geometry of implicit equations. Indeed, if we attempt to describe the solution of a differential equation as a function satisfying the system

$$F(t, y, y') = y' - f(t, y) = 0,$$

the natural question arises about the existence of a function $y(t)$ that is “implicit” in this relation. To apply the Implicit Function Theorem in Banach spaces (Theorem 3.6), we first identify appropriate Banach spaces:

- $X = \mathbb{R} \times \mathbb{R}^n$, the space of initial parameters (t_0, y_0) ,
- $Y = C([t_0 - \delta, t_0 + \delta], \mathbb{R}^n)$, the space of candidate solution functions,
- $Z = C([t_0 - \delta, t_0 + \delta], \mathbb{R}^n)$, the space of values for the operator.

Next, we define a mapping $F : X \times Y \rightarrow Z$ by

$$F((t_0, y_0), y)(t) = y(t) - y_0 - \int_{t_0}^t f(s, y(s)) ds.$$

Then $F((t_0, y_0), y) = 0$ if and only if y satisfies the integral equation corresponding to the Cauchy problem.

The Fréchet derivative of F with respect to y is the linear operator

$$D_y F((t_0, y_0), y)h(t) = h(t) - \int_{t_0}^t D_y f(s, y(s)) h(s) ds,$$

which is invertible for sufficiently small $\delta > 0$, corresponding to a local existence neighborhood.

By Theorem 3.6, there exist neighborhoods $V \subset X$ and $W \subset Y$ and a unique continuously Fréchet differentiable function

$$\varphi : V \rightarrow W$$

such that

$$\varphi(t_0, y_0) = y \quad \text{and} \quad F((t_0, y_0), \varphi(t_0, y_0)) = 0 \quad \text{for all } (t_0, y_0) \in V.$$

Consequently, this guarantees the **local existence and uniqueness** of the solution to the Cauchy problem. Thanks to this, the implicit function appears in the background of many more advanced results in the theory of differential equations, for instance, when proving the dependence of solutions on parameters. If a parameter occurs in the differential equation, then the application of the theorem guarantees that the solution depends smoothly on that parameter, which is of fundamental importance, among other things, in bifurcation analysis. From the students' perspective, it is worthwhile, as part of a thesis seminar, to conduct a discussion on filling in the details of the above reasoning, especially since the matter concerns the last considered variant of the Implicit Function Theorem.

A similar idea appears in optimization theory. Consider the classical constrained optimization problem

$$\text{maximize/minimize } f(x, y) \quad \text{subject to } g(x, y) = 0,$$

where $f : \mathbb{R}^{n+m} \rightarrow \mathbb{R}$ and $g : \mathbb{R}^{n+m} \rightarrow \mathbb{R}^m$ are continuously differentiable functions. We split the variables as $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^m$.

Step 1: Identification of the function F .

We define

$$F(x, y) := g(x, y), \quad (x, y) \in \mathbb{R}^n \times \mathbb{R}^m.$$

Let $(a, b) \in \mathbb{R}^{n+m}$ be a point satisfying the constraints, i.e.,

$$F(a, b) = g(a, b) = 0.$$

Assume that the Jacobian matrix with respect to the y -variables,

$$D_y F(a, b) = \begin{pmatrix} \frac{\partial g_1}{\partial y_1} & \cdots & \frac{\partial g_1}{\partial y_m} \\ \vdots & \ddots & \vdots \\ \frac{\partial g_m}{\partial y_1} & \cdots & \frac{\partial g_m}{\partial y_m} \end{pmatrix} (a, b),$$

is invertible, i.e., $\det D_y F(a, b) \neq 0$.

Step 2: Application of Theorem 3.4.

By Theorem 3.4, there exist neighborhoods $V \subset \mathbb{R}^n$ of a and $W \subset \mathbb{R}^m$ of b , and a unique continuously differentiable mapping

$$\varphi : V \rightarrow W$$

such that

$$\varphi(a) = b \quad \text{and} \quad F(x, \varphi(x)) = g(x, \varphi(x)) = 0 \quad \text{for all } x \in V.$$

In other words, locally the dependent variables y can be expressed as functions of the independent variables x in a way that satisfies the constraints.

Step 3: Reduction to an unconstrained problem.

The constrained optimization problem reduces to the unconstrained problem

$$\tilde{f}(x) := f(x, \varphi(x)), \quad x \in V.$$

The classical first-order necessary conditions then yield

$$\nabla_x \tilde{f}(x) = \nabla_x f(x, \varphi(x)) + \nabla_y f(x, \varphi(x)) D\varphi(x) = 0.$$

Substituting

$$D\varphi(x) = -(D_y F(x, \varphi(x)))^{-1} D_x F(x, \varphi(x)),$$

as guaranteed by Theorem 3.4, leads directly to the familiar Lagrange multiplier conditions

$$\nabla f(x, y) = \sum_{i=1}^m \lambda_i \nabla g_i(x, y).$$

Thus, Theorem 3.4 is precisely the tool that justifies expressing the dependent variables y locally in terms of x , which is the key step in rigorously deriving the Lagrange multiplier method. Let us note that this is only a sketch of the formal reasoning; similarly to the Cauchy problem, all steps should be considered in

detail, which in particular requires integrating the various demanding aspects of understanding emphasized in Sierpińska (1994).

Yet another perspective is offered by differential geometry. Consider a subset $M \subset \mathbb{R}^n$ defined as the level set of a smooth mapping

$$F : \mathbb{R}^n \rightarrow \mathbb{R}^m, \quad M := \{x \in \mathbb{R}^n : F(x) = 0\}.$$

Assume that $n > m$ and the Jacobian matrix $DF(x) \in \mathbb{R}^{m \times n}$ has full rank m at every point $x \in M$.

1. Local representation as a graph. Fix a point $a \in M$. Without loss of generality, we can relabel the coordinates as $x = (u, v) \in \mathbb{R}^{n-m} \times \mathbb{R}^m$ so that the $m \times m$ submatrix $D_v F(a)$ is invertible. Then, by the classical Implicit Function Theorem (Theorem 3.4), there exists a neighborhood $U \subset \mathbb{R}^{n-m}$ of u_0 and a unique smooth function

$$\varphi : U \rightarrow \mathbb{R}^m$$

such that

$$F(u, \varphi(u)) = 0 \quad \text{for all } u \in U.$$

Consequently, locally near a , the set M can be expressed as

$$M \cap (U \times \mathbb{R}^m) = \{(u, \varphi(u)) : u \in U\},$$

i.e., as the graph of a smooth function.

2. Manifold structure. The above local graph representation implies that M is locally diffeomorphic to an open subset of \mathbb{R}^{n-m} , so it is a smooth manifold of dimension $n - m$. The Implicit Function Theorem ensures not only the existence of φ , but also its smoothness, which allows the construction of local coordinate charts $(u, \varphi(u))$ for M .

3. Consequences. With these local coordinates, one can study the dimension of M , define tangent spaces at each point via

$$T_a M = \ker DF(a),$$

and introduce additional geometric structures such as Riemannian metrics, differential forms, and vector fields. Hence, the Implicit Function Theorem provides a rigorous foundation for treating level sets of smooth functions as smooth manifolds, making it a cornerstone of modern differential geometry and topology.

This last example requires students to have a basic knowledge of differential geometry. There are many excellent books covering the fundamental terminology; from our perspective, one of the most suitable references is Spivak (1965).

The presented examples demonstrate that the Implicit Function Theorem is far from being a merely technical result of analysis. Rather, it is a universal tool that forms the analytic foundation of several major areas of mathematics, including the theory of differential equations, optimization, and geometry.

5. Conclusion

What emerges most clearly from the preceding discussion is not only the mathematical depth of the Implicit Function Theorem, but also its exceptional value as a pedagogical tool. For students, the theorem can serve as a guiding thread that connects apparently distant topics of the curriculum. When one sees how the same principle explains the solvability of differential equations, the rigor behind constrained optimization, and the structure of smooth manifolds, mathematics begins to appear as a tightly interwoven system rather than a collection of unrelated facts. In this sense, the theorem plays the role of a conceptual bridge, fostering a holistic understanding of analysis.

The systematic analysis of different versions of the theorem makes it clear that its strength lies not only in local solvability of equations but also in the flexibility of its formulations. The transition from the classical finite-dimensional statement to the Banach space version mirrors the natural growth of mathematical analysis itself – from concrete computations to abstract functional-analytic frameworks. For students, this journey illustrates how the same idea can survive and flourish in increasingly general settings, thereby becoming a model of mathematical abstraction. It also shows how technical details – such as invertibility of derivatives or continuity of mappings – are not mere formalities but the precise conditions that guarantee the power of the theorem.

From the didactical standpoint, the Implicit Function Theorem provides fertile ground for active learning. Long-term observations indicate that many students remember it only as a tool for differentiating implicitly defined curves. By deliberately confronting them with applications in differential equations, optimization, and geometry, one can expand their horizons and demonstrate how a single theorem underlies an impressive variety of classical results. This methodical approach promotes a form of understanding that is both conceptual and procedural: students not only know *how* to apply the theorem, but also *why* it is central to analysis. In this way, the classroom becomes a space where theory, practice, and reflection meet.

In conclusion, the Implicit Function Theorem deserves to be taught not merely as a piece of analysis, but as a cornerstone of mathematical education. It embodies the essence of mathematics as a discipline: the search for general principles that illuminate diverse problems, the balance between rigor and intuition, and the interplay between abstraction and application. Future educational practice may benefit from designing seminar activities in which students reconstruct classical proofs step by step using this theorem, thereby experiencing first-hand how mathematics builds itself upon fundamental ideas. Such an approach aligns with the spirit of active learning emphasized in Sierpińska (1994) and Freeman et al. (2014), and confirms that genuine mathematical understanding is achieved when students see the unity of theory across seemingly distant fields.

References

- Dini, U. (1878). *Fondamenti per la teorica delle funzioni di variabili reali*. Pisa, Italy: Tipografia Nistri.
- Freeman, S., Eddy, S. L., McDonough, M., Smith, M. K., Okoroafor, N., Jordt, H., and Wenderoth, M. P. (2014). Active learning increases student performance in science, engineering, and mathematics. *Proceedings of the National Academy of Sciences of the United States of America*, **111**(23), pp. 8410–8415. <https://doi.org/10.1073/pnas.1319030111>
- Krantz, S. G., and Parks, H. R. (2013). *The Implicit Function Theorem: History, Theory, and Applications*. Reprint ed. Cham, Switzerland: Birkhäuser.
- Lang, S. (1993). *Real and Functional Analysis*. New York, NY: Springer.
- Sierpińska, A. (1994). *Understanding in Mathematics*. London, UK: Falmer Press.
- Spivak, M. (1965). *Calculus on Manifolds: A Modern Approach to Classical Theorems of Advanced Calculus*. New York, NY: W. A. Benjamin.
- Zeidler, E. (1995). *Applied Functional Analysis: Applications to Mathematical Physics*. New York, NY: Springer.

Artur Wachowicz
Institute of Mathematics
Lodz University of Technology
Poland
e-mail: artur.wachowicz@p.lodz.pl