UNIVERSIDAD
DE GRANADA

Trabajo Fin de Grado

Facultad de Ciencias

Grado en Matemáticas

Teorema de la función Implícita

María del Carmen Andrades Moya

Departamento de Análisis Matemático

Tutor: Antonio Cañada Villar

Curso Académico 2020/2021

Trabajo fin de Grado. Curso Académico 2020/2021.

Responsable de tutorización:

Antonio Cañada Villar.

Departamento de Análisis Matemático.

Índice general

Prólogo

Antes de empezar a usar términos matemáticos más complejos, deberíamos previamente definir lo más simple, y la base, al fin y al cabo, de este trabajo, el concepto de función. Puede parecer una tarea sencilla, no obstante, la definición de ésta ha ido formándose poco a poco a lo largo de la historia de las matemáticas, y como tal, ha sufrido cambios, o mejor dicho, ha ido mejorando en rigor y precisión con el curso de los años.

Los babilonios fueron los primeros en emplearlas, tenían tablas de números naturales expresando sus cuadrados o sus cubos, aunque el hecho de que las usaran, no significa que supieran lo que eran de manera rigurosa. De este modo, no podemos, ni debemos, afirmar que los babilonios sean los creadores del concepto de función. Por estas mismas razones, por muy presentes que las funciones estuviesen durante todos estos años, no era algo formalmente definido. No fue hasta el siglo XIV, en las Universidades de Oxford y París, cuando se empezó a materializar una definición, refiriéndose a las leyes de la naturaleza como la dependencia de algunas cantidades de otras.

Más tarde, con **Galileo**, el concepto empezó a ver algo más de luz. Sus estudios sobre el movimiento contenían un claro entendimiento sobre las relaciones entre variables. Casi al mismo tiempo, la misma idea se formaba en la cabeza de **Descartes**, el cuál, en su libro *La Géometrie*(1637) comentaba cómo una curva podía dibujarse dándole infinitos valores a las variables, que al fin y al cabo, es lo que generalmente hacemos hoy en día para la construcción de una curva, dada una función.

El uso por primera vez de la expresión "función", no se dio hasta 1673, con el matemático **Leibniz**. Sin embargo la palabra no fue usada en un lenguaje matemático, como pasa como con la mayoría del lenguaje usado en las ciencias. Tras otros cien años de espera, en 1748, la definición de función apareció de la mano de **Léonard Euler**, y era la siguiente,

Una función de una variable es una expresión analítica compuesta de alguna manera de la variable en cuestión y una cantidad numerable de números o constantes. [2]

No obstante, nos damos cuenta al momento de que el definir una función como una fórmula no nos trae más que problemas. Sea por ejemplo, el conjunto de puntos de $y^5 + 16y - 32x^2 + 32x = 0$, la figura 1 mostrada en $\mathbb{R}^2$ nos hace suponer que se trata del gráfico de y en función de x, pero no tenemos una fórmula existente que nos lo diga.

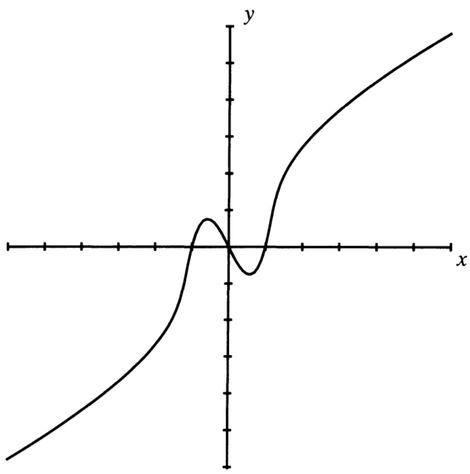

Figura 1: Gráfica de $y^5 + 16y - 32x^2 + 32x = 0$. [1]

La definición de Euler, por tanto al no ser una definición hecha con su respectiva expresión analítica, no nos aporta mucho, con ella solo podemos asumir que una función se trata de un conjunto de operaciones como sumar, multiplicar, raíces, etc. Si bien más tarde, en 1755, volvió a dar una nueva definición de función, esta última siendo algo más precisa.

Si algunas cantidades dependen de otras cantidades de forma que las últimas cambien cuando lo hagan las primeras, entonces podríamos decir que las últimas son funciones de las primeras. [1]

Durante las siguientes décadas diversos matemáticos fueron aportando sus propias definiciones de función, como **Lacroix**, **Cauchy**, **Fourier** o **Poincaré**, entre otros. Para ser más exactos muchos matemáticos de prestigio acabaron dando su propia versión. Cabe mencionar, que en la de Cauchy, se generaliza por primera vez una definición de función en la que se incluyen tanto las funciones explícitas como las implícitas, conceptos que definiré con precisión más adelante. Sin embargo, como pasaba con Euler, lo hace hablando de fórmulas, y como vimos previamente eso no nos da más que complicaciones. Fue en 1923, gracias a **Goursat**, con el que obtuvimos una definición similar a la moderna que tenemos actualmente, la cual doy a continuación.

Una *función* dada con *dominio* X y *codominio* o *rango* Y es un subconjunto, llamémosle f, del producto cartesiano

$$X \times Y = \{(x, y) : x \in X, y \in Y\}$$

con las siguientes propiedades, (i) para cada $x \in X$ hay un elemento $(x, y) \in f$, y (ii) si $(x, y) \in f$ y $(x, z) \in f$, entonces $y = z$.

Siempre que se cumplan estas dos condiciones, tenemos que dado un $x \in f$ cualquiera, tenemos un único $y \in Y$ tal que $(x, y) \in f$. Dado que la propia definición nos da la unicidad de dicha función, podemos escribir simplemente $y = f(x)$, para decir que $(x, y) \in f$.

Usando esta definición es fácil ver que, en efecto, el ejemplo que se puso anteriormente es una función. Podemos llegar a la conclusión de que para cada $x \in \mathbb{R}$, existe un único $y \in \mathbb{R}$ tal que (x, y) satisface la ecuación escrita más arriba. Además, tenemos que el conjunto de puntos que dibujamos antes, es el gráfico $y = f(x)$, luego ya tenemos la función f definida en el sentido moderno.

Nos bastaba con estudiar el comportamiento de $F(y) = y^5 + 16y - 32x^2 + 32x = 0$ para un x fijo, es decir, si para cada x, sabemos que existe un único y que verifica dicha ecuación, pero no somos capaces de calcularlo explícitamente, tenemos que la ecuación está definida de **forma implícita**.

Por tanto, no hace falta añadir, que también existen las **funciones explícitas**, las cuales no son más que las que vienen expresadas directamente como $y = f(x)$, como por ejemplo, $f(x) = x^4 + log(x) - 7$.

Ahora que tenemos más que definida una función, y hemos explicado las diferencias entre explícitas e implícitas, no nos queda otra que hablar sobre la razón por la que estamos aquí, el **Teorema de la Función Implícita**.

Uno espera que teniendo una ecuación de una variable, $F(x) = c$, siendo c una constante, debe ser suficiente para determinar el valor de x dado c. En el caso de tener dos variables, podríamos pensar de la misma manera, y suponer que dos ecuaciones serán sufucientes para encontrar los valores de esas dos variables, y así sucesivamente. De manera más general, lo que tenemos es que con un sistema de m ecuaciones con m variables,

$$\begin{aligned} F_1(x_1, x_2, \ldots, x_m) &= c_1, \\ F_2(x_1, x_2, \ldots, x_m) &= c_2, \\ &\vdots \\ F_m(x_1, x_2, \ldots, x_m) &= c_m, \end{aligned} \tag{1}$$

con $c_1, c_2, \ldots, c_m$ constantes, uno esperaría que es el número ideal para poder obtener los valores de las variables $x_1, x_2, \ldots, x_m$. Podemos incluso llegar a razonar que con un sistema de ecuaciones como el descrito anteriormente, pero ahora con más incógnitas (n) que ecuaciones (m), tratar las $n-m$ extra variables como parámetros, de forma que hay m variables que son funciones implícitas de $n - m$ parámetros.

Tenemos por tanto, que bajo unas condiciones especiales y variando $c_1, c_2, \ldots, c_m$, podremos definir estas m variables como funciones implícitas respecto de las otras. De forma que la finalidad del Teorema de la Función Implícita es de proporcionarnos un método lo suficientemente potente, o un conjunto de métodos, para poder asegurar que efectivamente estamos bajo esas condiciones especiales con las cuáles podremos dar el argumento anterior como válido.

El teorema de la función implícita se puede describir de diversas formas, a lo largo de este trabajo se expresarán varias, algunas más modernas, otras más complejas, pero al fin y al cabo todas siguen la misma base. A continuación, voy a enunciarlo de manera informal como se hace en [1],

Teorema de la Función Implícita (Informal): *Sean las ecuaciones siguientes, dadas por funciones continuamente diferenciables,*

$$\begin{aligned} F_1(x_1, x_2, \ldots, x_n) &= c_1, \\ F_2(x_1, x_2, \ldots, x_n) &= c_2, \\ &\vdots \\ F_m(x_1, x_2, \ldots, x_n) &= c_m, \end{aligned} \tag{2}$$

si (2) *se cumple para un punto* $a = (a_1, a_2, \ldots, a_n)$ *y si, cuando las funciones en* (2) *sean remplazadas por sus aproximaciones lineales, un conjunto en concreto de* m *variables puede ser expresado como funciones de las otras* $n-m$ *variables, entonces, para* (2)*, las mismas* m *variables pueden definirse como funciones implícitas de las otras* $n-m$ *variables en un entorno de* a*. Además, las funciones implícitas resultantes son también continuamente diferenciables y sus derivadas pueden obtenerse mediante derivación implícita.*

El teorema, actualmente, es lo que un alemán llamaría un *ansatz*, cuya traducción no es más que un *enfoque*, es decir, una manera de mirar a los problemas. Tenemos diversas versiones de él, como de otros muchos teoremas, y podemos generalizar cada uno de ellas a espacios euclídeos, de Banach o incluso conjuntos más complicados. En definitiva, el teorema lo que nos da es un **método para resolver ecuaciones**.

Para la realización de este prólogo se han utilizado las referencias [1] y [2].

Introducción

El presente Trabajo de fin de grado con título *Teorema de la función Implícita*, se ha realizado teniendo en mente afrontar los siguientes objetivos:

- Revisión histórica de sus orígenes (incluyendo las aportaciones de diversos matemáticos durante distintos momentos de la historia).
- Estudio comparativo de diferentes demostraciones que se han llevado a cabo sobre dicho teorema.
- Descripción de diferentes aplicaciones clásicas y modernas del teorema.

Estos objetivos se han alcanzado mediante la realización de los siguientes tres capítulos:

1. Origen histórico del teorema de la función implícita.
2. Estudio comparativo de diferentes demostraciones del T.F.I.
3. Generalizaciones del T.F.I y aplicaciones.

A continuación, se explica brevemente el contenido de cada uno de los capítulos.

Capítulo 1: *Origen histórico del teorema de la función implícita.*

Este capítulo trata sobre el del primer objetivo que hemos mencionado. Nos hemos centrado en cuatro matemáticos que tuvieron una gran importancia en el desarrollo del teorema de la función implícita.

El primero de ellos, Isaac Newton, lo hizo indirectamente mediante su *Polígono de Newton*. El segundo, Joseph Louis Lagrange, realizó lo que puede ser la primera demostración del teorema, aunque se asemejaba más a la demostración de un teorema de la función inversa, al cuál se le conoce como el *Teorema de Inversión de Lagrange*. A continuación tenemos a Augustin-Louis Cauchy, que es considerado el primer matemático que enunció y probó un teorema de la función implícita. Por último, Ulisse Dini. Los matemáticos anteriores a Dini habían estudiado versiones del teorema para funciones de una única variable, sin embargo, fue Dini, como se explica en el capítulo, el que describió de forma rigurosa por primera vez la versión que podemos encontrar en los libros de texto en la actualidad, que usa $n \geq 1$ variables.

Capítulo 2: *Estudio comparativo de diferentes demostraciones del T.F.I.*

El segundo capítulo, trata sobre el segundo objetivo como bien indica su título. A lo largo de éste, se van a enunciar tres versiones del teorema de la función implícita y sus respectivas demostraciones, además de los preliminares que hayan sido necesarios para su realización.

La primera demostración, es la realizada por Ulisse Dini. Se trata de una demostración inductiva. Primero se enuncia y demuestra la versión para para una sola variable dependiente y ecuación y un número arbitrario de variables independientes. Después mediante inducción, llegamos a la versión general que sirve para $n \geq 1$ variables. La segunda demostración es la demostración clásica, se trata de demostrar el teorema de la función implícita mediante el teorema de la función inversa. Ya que ambos son equivalentes, como veremos en este capítulo, bastará demostrar el segundo para obtener la prueba del primero. Para estas dos demostraciones se ha usado el libro de Krantz como referencia [1], no obstante, para la tercera y última demostración, que emplea el método de aproximaciones sucesivas he empleado un

Trabajo de Fin de Grado de 2017, de la Universidad de Sevilla [7]; la razón de esto ha sido mera preferencia respecto a la que se incluía el libro de Krantz.

Capítulo 3: *Generalizaciones del T.F.I y aplicaciones.*

En este último capítulo, se afronta el tercer objetivo de este trabajo. Se ha dividido en tres secciones, donde cada una de ellas engloba una aplicación distinta del teorema de la función implícita.

La primera de ellas se centra en las ecuaciones diferenciales ordinarias, ya que sabemos que existe una relación directa entre el teorema de la función implícita y el problema de Cauchy en teoría de ecuaciones diferenciales. Por esta razón, veremos que se puede demostrar, usando el teorema de existencia de ecuaciones diferenciales ordinarias (Picard), el teorema de la función implícita de forma iterativa. A continuación, se van a tratar algunos casos de sistemas de ecuaciones en los que la matriz jacobiana es degenerada, y veremos cuando se va a poder aplicar igualmente el teorema de la función implícita. Por último, se va a considerar el teorema de la función implícita en el ámbito del análisis complejo.

Previamente a estos capítulos, se ha realizado una pequeño prólogo para este trabajo, en el cuál se ha definido el concepto de función, en concreto el de función implícita, además de explicar brevemente su origen. En él, también se ha enunciado de manera informal una versión del teorema de la función implícita.

Finalmente, como se puede ver, mediante la realización de cada capítulo los objetivos iniciales se han cumplido satisfactoriamente, además del uso de la bibliografía que aparece al final y de la ayuda de mi tutor, Antonio Cañada.

Summary

The implicit function theorem is a classical theorem of mathematical analysis and it is used for equation systems with multiple variables. By using this theorem, we are able to obtain numerous related results, such as the well-known inverse function theorem. Its applications are countless (ordinary and partial differential equations, differential geometry, functional analysis, etc).

The present end degree work entitled *The Implicit Function Theorem*, is divided into three chapters and a brief preface. In the preface, we give diverse definitions of the mathematical term *function* throughout the history of mathematics, beginning with Euler:

A function of a variable quantity is an analytic expression composed in any way whatsoever of the variable quantity and numbers or constant quantities. [2]

As we all know, nowadays we use one slightly different definition, which was given by Goursat in 1923, and which can be found in details in the preface, as well as the meaning of the appreciated implicit function.

The first chapter of this work is dedicated to the founders of the theorem, in other words, the prestigious mathematicians who were key in development of the implicit function theorem. We will describe how the British physicist Isaac Newton without realising it, was the pioneer, thanks to his Newton's Polygon. Later, Lagrange, gave us the first proof of what we would call now, an inverse function theorem, but it

was the French mathematician Cauchy who is considered to be the first ever to enunciate and prove our main theorem. Nevertheless, the implicit function theorem we learn now was described by the Italian Ulisse Dini, as he is the one who gave the general version, which applies for systems of equations of multiple variables. Next, we enunciate this version,

Implicit function theorem for one dependent variable: *Let be the open set $\mathcal{O} \subseteq \mathbb{R}^n$, and $f : \mathcal{O} \to \mathbb{R}$ a continuously differentiable function and $p = (p_1, p_2) \in \mathbb{R}^{n-1} \times \mathbb{R}$, where $p \in \mathcal{O}$ is a point for which:*

$$f(p) = 0 \text{ and } \frac{\partial f}{\partial x_n}(p) \neq 0.$$

Then there exists an open set $\mathcal{U} \subset \mathbb{R}^{n-1}$ with $p_1 \in \mathcal{U}$ and a unique continuously differentiable function $F : \mathcal{U} \to \mathbb{R}$ such that

$$p_2 = F(p_1) \text{ and } f[x, F(x)] = 0, \forall x \in \mathcal{U}.$$

The middle chapter consists in giving three different versions of the implicit function theorem and their respective demonstrations. We start by proving the Ulisse Dini's version, as described in [1]. We do this by the induction method on the number of dependent variables. That is, we first enunciate and prove the version for one dependent variable, one equation and any number of independent variables, which is quite simple, and then by induction, we finally give and demonstrate the general version, which is the next one:

Implicit function theorem: *There exists a neighborhood $\mathcal{O} \subset \mathbb{R}^m$ of p and a set of continuously differentiable functions $\delta_j : \mathcal{O} \to \mathbb{R}, j = 1, 2, \ldots, n$, such that $\delta_j(p) = q_j, j = 1, 2, \ldots, n$ and*

$$f_i[x; \delta_1(x), \delta_2(x), \ldots, \delta_n(x)] = 0, i = 1, 2, \ldots, n, \forall x \in \mathcal{O}.$$

In the proof, we use the classical approach. This one is relatively easier in comparison, that is the reason this is the preferred method to explain this theorem to students. On this section, we enunciate both the inverse and implicit theorem, and prove that they are equivalent to each other. In this way, we obtain our implicit function theorem demonstration by just proving the inverse function one. As stated earlier on the introduction of this work, the main source used for this two proofs has been the Krantz's book named *The Implicit Function Theorem. History, theory and applications* [1].

Finally in this chapter, we shall use the contraction mapping fixed point principle to obtain our last variant of the implicit function theorem. This section, unlike the previous ones, is based on functional analysis, therefore it is more abstract. However, for this particular reason, this version of the theorem can be applied in different and more general settings. For this one, we have relied on a previous end degree work from another university, also entitled *Teorema de la Función Implícita* [7].

On the final chapter of this work, we explain some applications and generalisations of the implicit function theorem, as stated on the introduction.

It is known about the direct correlation between this theorem and ordinary differential equations. The mathematician Goursat studied this connection and gave an iterative proof of the implicit function theorem. He was inspired by the Picard's iterative proof used on the existence theorem for ordinary differential equations. In this section, we will enunciate Picard's theorem and then explain how it can be used to obtain an implicit function theorem demonstration.

Secondly, we have been requiring through this work, the condition of a non degenerate jacobian matrix throughout this work. However, in this section, we will give some examples of cases where it does degenerate, and see whether we can still apply the implicit function theorem or not. After describing some necessary preliminary remarks, we will study and differentiate various cases where the jacobian has range one, and later on we will illustrate an example where its range is zero.

Last but not least, we will consider the implicit function theorem in both cases: the real analytic and the complex analytic frameworks. We also now that they are closely related as the problem in the real analytic category can be turned into a holomorhic theorem and viceversa. Consequently, an implicit function problem with complex analytic data will have a $\mathcal{C}^\infty$ solution by the classical $\mathcal{C}^\infty$ implicit function theorem, and hence, it also has a real analytic solution by the real analytic implicit function theorem. The key of the matter is to see that it has a holomorphic solution, as we describe bellow, from [1],

Implicit function theorem: *Suppose that the power series,*

$$F(x,y) = \sum_{\alpha,k} a_{\alpha,k} x^\alpha y^k,$$

is absolutely convergent for $|x| \leq R_1$*,* $|y| \leq R_2$*, if* $a_{0,0} = 0$ *and* $a_{0,1} \neq 0$*, then there exits* $r_0 > 0$ *and a power series*

$$f(x) = \sum_{|\alpha|>0} c_\alpha x^\alpha, \tag{3}$$

such that (3) is absolutely convergent for $|x| \leq r_0$ *and*

$$F(x, f(x)) = 0.$$

Capítulo 1

Origen histórico del Teorema de la función Implícita

El teorema de la función implícita, junto con el teorema de la función inversa, es uno de los más importantes y antiguos paradigmas que se han dado en las matemáticas. Se podría decir que la semilla del teorema apareció por primera vez entre los escritos del físico matemático **Isaac Newton** (1642-1727), y en los de **Gottfried Leibniz** (1646-1716), el cuál llego incluso a comentar la derivabilidad implícita.

Por otro lado, tenemos a **Joseph Louis Lagrange** (1736-1813),el cual llegó a encontrar un teorema que es prácticamente el teorema de la función inversa. Fue **Augustin-Louis Cauchy** (1789-1857) el que con más detalle se centró en el teorema de la función implícita. Debido a ello, es el matemático al que se le suele dar el título de creador del teorema. Esto es gracias al matemático Willian Fogg Osgood, el cual fue el que mencionó la *Memoria de Turin*, como la fuente del teorema, definiendo así a Cauchy como el creador de éste.

El teorema, como todo en la vida, fue evolucionando, hasta llegar a convertirse en el que se enseña en las aulas de matemáticas por todo el mundo. Fue formalmente definido por primera vez en el ambiente del análisis complejo y de las series de potencias complejas. El trabajo de Cauchy mencionado anteriormente estaba basado principalmente en resultados del análisis complejo, ya que no fue hasta el siglo XIX, cuando se pudo apreciar por primera vez diferencias importantes entre el análisis complejo y el real.

Apareció entonces una versión del teorema para variables reales. Éste tomó la versión más simple, es decir, la de dos variables reales, siendo la hipótesis principal que al menos una derivada parcial fuese no nula, hipótesis correspondiente a que la matriz jacobiana sea no singular. Ésta vino de la mano del matemático italiano **Ulisse Dini** (1845-1918) que generalizó la versión del teorema con variables reales para cualquier número de ecuaciones. Una vez que los matemáticos fueron estudiando el teorema más a fondo y, paralelamente, comprendiéndolo mejor, empezaron a salir nuevas demostraciones, que usaban técnicas algo más modernas, gracias a las cuáles se fueron desarrollando nuevas versiones del teorema.

Durante este capítulo, se va ir haciendo un recorrido por las distintas contribuciones que aportaron algunos de los matemáticos que contribuyeron al descubrimiento del teorema, enunciando detalladamente pero sin un rigor total, cada una de ellas.

1.1. Isaac Newton

Antes del siglo XVIII parece ser que ningún matemático llegó a tener un mínimo de interés en probar la existencia de funciones implícitas. Para ser más precisos, y citando las propias palabras de Léonard Euler [1],

"Tenemos que, efectivamante, un gran número de funciones algebraicas no es posible expresarlas de manera explícita. Sea por ejemplo, la función Z de z que vienen definida por la ecuación, $Z^5 = az^2Z^3 - bz^4Z^2 + cz^3Z - 1$. Aunque tengamos la certeza de que esta ecuación es irresoluble, no podemos negar el hecho de que sí podemos expresar Z como una expresión que dependa exclusivamente de la variable z y constantes, luego; podemos afirmar que Z es una función de z."

A los matemáticos de entonces les importaba más estudiar el comportamiento de tales funciones, que probar su propia existencia. El trabajo de Newton que comentaré más abajo, se puede considerar uno de los primeros en analizar el comportamiento de funciones definidas implícitamente.

Isaac Newton (1643-1727), fue un físico matemático nacido en Inglaterra. Es conocido fundamentalmente por sus estudios sobre las leyes de la mecánica clásica, a través de las leyes del movimiento y su ley de la gravitación universal. No obstante, también hizo descubrimientos fundamentales en el área de las matemáticas. Se le atribuye la creación del cálculo integral y diferencial, junto con Gottfried Leibniz. Pero esta no es la razón por la que estamos aquí.

Entre los muchos manuscritos que Newton escribió a lo largo de su vida, en uno precisamente escrito en latín, *De Analysi per AEquationes Infinitas* (1669), Newton se plantea el problema de cómo expresar la siguiente ecuación,

$$Y^3 + a^2Y - 2a^3 + axY - x^3 = 0 \tag{1.1}$$

respecto de la variable x, de forma que ésta sea válida en un entorno del punto $x = 0$ y tal que nos de una solución de la ecuación para $a \neq 0$ cuando $x = 0$. Este ejemplo en cuestión se encuentra dentro de un apartado del libro llamado *Exempla per Resolutionem AEquationum Affectarum*.

Este apartado empieza con lo que hoy en dia llamamos el **Método de Newton**, usado principalmente en la matemática aplicada como un método numérico para encontrar raíces. No obstante, se trata de la primera referencia en la historia de las matemáticas que tenemos de algo similar al teorema de la función implícita.

Newton tuvo la oportunidad de desarrollar más a fondo este procedimiento en otro manuscrito, *De Methodis Serierum et Fluxionum* (1670), en el cuál el método mejorado que construye es al que llamamos **Polígono de Newton** o **Diagrama de Newton**.

A continuación voy a explicar brevemente en qué consiste este método.

Construcción del polígono de Newton: El polígono de Newton se usa para determinar el comportamiento de un conjunto de puntos que siguen una función polinómica del tipo,

$$P(x,y) = \sum_{n=0}^{N} \sum_{i+j=n} a_{i,j} x^i y^j \tag{1.2}$$

en un entorno de uno de dichos puntos. Por comodidad, mediante el uso de traslaciones podemos hacer cambios en la función para que ese punto sea el $(0,0)$. También podemos suponer que no hay ningún factor común de x o y en el polinomio.

Con la ayuda del ejemplo que el mismo Isaac Newton usó en su momento, es decir la ecuación (1.1), me dispongo a realizar la construcción de su polígono de Newton. Además, de esta función implícita podremos obtener una solución $y = y(x)$, de forma que su gráfica pase por el origen y tal que esté definida al menos para valores suficientemente pequeños de x. En particular, tendremos que se verifica $y(0) = 0$. Empezamos realizando una pequeña traslación sobre la variable y, de forma que se cumpla la condición que dijimos antes y de paso, para quitarnos de en medio a la variable Y, hacemos el cambio $y = Y - a$, y sustituyendo en la ecuación (1.1) obtenemos,

$$y^3 + 3ay^2 + 4a^2y + axy + a^2x - x^3 = 0. \tag{1.3}$$

Empleando ahora la notación de (1.2), tenemos que,

$$a_{1,0} = a^2, a_{0,1} = 4a^2, a_{1,1} = a, a_{0,2} = 3a, a_{3,0} = -1, a_{0,3} = 1,$$

con el resto de coeficientes igual a 0.

El conjunto formado por todos los segmentos lineales creados al conectar los pares de puntos del conjunto

$$\{(i,j) : a_{i,j} \neq 0\} \tag{1.4}$$

contiene a un conjunto convexo especial, al que vamos a llamar C. De hecho, C es la envolvente conexa del conjunto dado en (1.4). La frontera, ∂C, es un polígono cerrado contenido en el primer cuadrante del eje de coordenadas, que además interseca a ambos ejes.

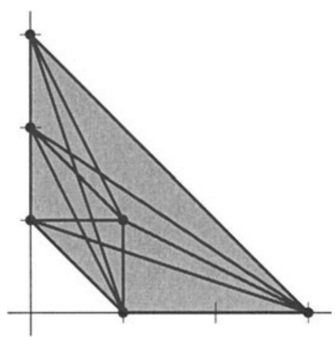

Figura 1.1: Polígono de Newton. [1]

Ya tenemos construido por tanto el polígono de Newton, y es el de arriba.

Tenemos que para cada segmento lineal formado se le corresponde un α, Siendo entonces $-1/\alpha$ la pendiente de dicho segmento de línea. Una de las ideas de Newton detrás del polinomio, es asumir que se puede escribir $y(x) = x^{\alpha}\tilde{y}(x)$, siendo $\tilde{y}(x)$ una función continua que no se anula en $x = 0$. Primero se obtenían los α que cumplían esta condición y después para saber con cuál quedarse se empleaba un método geométrico algo más elaborado.

Para la ecuación (1.3) en concreto, solo tenemos un segmento en el polígono de Newton, y este tiene pendiente -1, luego sustituyendo $y = x\tilde{y}$ y, eliminando los factores comunes de x, obtenemos la siguiente expresión,

$$x^2\tilde{y}^3 + 3ax\tilde{y}^2 + ax\tilde{y} - x^2 + 4a^2\tilde{y} + a^2 = 0. \tag{1.5}$$

Cuya solución en un entorno de $(0, 0)$ satisface $\tilde{y} \approx -1/4$, por lo que concluimos que $y \approx -1/4x$ y finalmente, deshaciendo la traslación que se hizo al principio,

$$Y \approx a - 1/4x.$$

Luego utilizando el método de Newton, mediante la construcción de su polígono, hemos llegado a la misma aproximación lineal que se hubiese obtenido con el teorema de la función implícita.

1.2. Joseph Louis Lagrange

En 1770, Lagrange probó lo que puede ser el primer teorema de la función implícita, aunque se asemejaba más al teorema de la función inversa. Al resultado como tal se le conoce como el **Teorema de Inversión de Lagrange**, y fue uno de los resultados más trascendentales que hizo a lo largo de su vida. El teorema se podría considerar como un caso especial del teorema de la función inversa para series de potencias, que nos permite obtener la expansión en serie de Taylor de la función inversa de una función analítica.

En efecto, la mayor parte de la fama de Lagrange como matemático se debe a sus estudios en mecánica celeste, ya que sus mayores contribuciones estaban relacionadas con ella. Recibió numerosos premios por su trabajo, empezando en 1764, otorgado por la Academia de las Ciencias de París, gracias a su escrito sobre la libración de la luna. [1]

Las series de Lagrange aparecieron por primera vez durante sus investigaciones sobre métodos para encontrar soluciones numéricas de ecuaciones algebraicas. Usando este mismo método, fue como obtuvo la solución del problema de **Johannes Kepler** (1571-1630) de la teoría planetaria, convirtiéndose así en la aplicación más conocida de las series de Lagrange. [6]

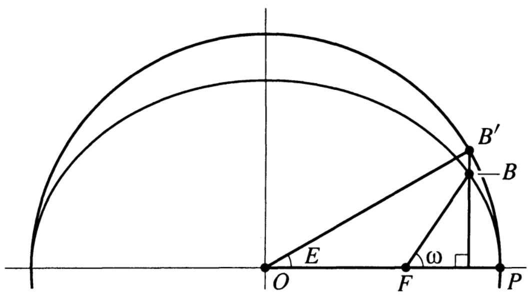

Figura 1.2: Kepler. [1]

La **ecuación de Kepler** es una de las fundamentales de la mecánica celeste,

$$E = M + e \cdot sen(E). \tag{1.6}$$

En astronomía, una anomalía se define como la cantidad angular usada para describir la posición de la órbita de un cuerpo celeste. En (1.6) M es la *anomalía media*, o ángulo que recorrería un planeta que se moviese uniformemente a lo largo de la circunferencia principal. E es la *anomalía excéntrica* y e es la *excentricidad* de la órbita. Tenemos que M y e son cantidades medibles, donde e es preferiblemente de un valor pequeño.

Gracias al teorema mencionado anteriormente, el Teorema de Inversión de Lagrange, se obtuvo una fórmula para la corrección que se debe hacer cuando para alguna función $\psi(.)$, $\psi(M)$ es remplazado por $\psi(E)$. La corrección toma la forma de una serie de potencias en e. Por tanto, se puede ajustar la diferencia entre la anomalía media y la excéntrica.

No voy a entrar demasiado en detalle en el problema de Kepler, pero si que me parece importante mencionar que antes de la intervención de Lagrange, las diversas soluciones que había eran erróneas, incluso la dada por el propio Kepler. Para determinados valores de la excentricidad, las soluciones fallaban, en concreto para valores demasiado altos. A continuación enuncio el teorema de Lagrange, con el lenguaje y rigor de la época. [1]

Teorema de Inversión de Lagrange: *Sea $f(z)$ y $g(z)$ funciones analíticas en el disco abierto $D(a,r) \subset \mathbb{C}$ y continuas en el cierre de dicho disco. Para un ϵ de módulo lo suficientemente pequeño tal que la desigualdad*

$$|\epsilon g(z)| < |z-a| \tag{1.7}$$

sea cierta para $z \in \partial D(a,r)$, entonces

$$\delta = a + \epsilon g(z) \tag{1.8}$$

tiene exactamente una raíz en $D(a,r)$ y, si esa raíz $\delta = \delta(\epsilon)$ es considerada como una función de ϵ, tenemos que

$$f(\delta) = f(a) + \sum_{n=1}^{\infty} \frac{t^n}{n!} \left(\frac{d^{n-1}}{dz^{n-1}} (f'(z) g(z)^n) \right) \Bigg|_{z=a} . \tag{1.9}$$

1.3. Augustin-Louis Cauchy

Cauchy definía una función de manera muy general como una variable que puede ser expresada mediante una u otras variables, a las que llamaba **variables independientes**. Si la función era multivaluada, Cauchy la denotaba como $f((x))$. Además, llegó a diferenciar varios tipos de funciones: explícitas, implícitas, simples, algebraicas y trigonométricas. Como ya dije al principio del capítulo, es considerado el primer matemático que enunció y probó un teorema de la función implícita, o al menos eso aseguraba Osgood.

Por razones políticas, tras la revolución de 1830, en 1831 Cauchy se auto exilió de Francia, yéndose a Italia, donde el rey de Cerdeña le ofreció la cátedra de profesor de física teórica en la Universidad de Turin, la cual fue creada expresamente para él. En 1831 publicó sus conocidas ***Memorias de Turin*** a la Real Academia de las Ciencias de Turin. Me voy a centrar en la primera, publicada el 11 de octubre, *Sur la mecanique celeste et sur un nouveau calcul qui s'applique a un grand nombre de questions diverses*, escrita en francés, ya que es la que contiene el teorema de la función implícita.

Consta de tres partes, de las cuáles solo nos interesa la Parte I, que fue publicada bajo algunas modificaciones en *Ejercicios de análisis y física matemática* de Cauchy. Como dato adicional, las partes de las memorias se encuentran entre los trabajos de Cauchy, es decir en diferentes libros o artículos, pero no fueron publicadas todas juntas como un solo tomo.

El tema fundamental en la que basa esta parte es la determinación de condiciones para la expansión de funciones explícitas e implícitas en series de potencias. A lo largo del trabajo, Cauchy desarrolló diversos métodos para la estimación de restos en esas expansiones tras un número finito de términos. A esta técnica Cauchy la nombró *calcul des limites*, que hoy se conoce como el **método de mayorantes**.

En los preliminares del libro, Cauchy insistió en la necesidad de probar la convergencia de las series infinitas más importantes de la mecánica celeste. Describió sus resultados sobre la convergencia de series de Taylor de una función y de las expansiones de series de potencias de funciones implícitas. También describe en términos generales las aplicaciones de dichos resultados en la mecánica celeste.

El siguiente teorema que voy a formular, bajo el contexto de funciones holomorfas, nos da la existencia de una función definida implícitamente bajo unas hipótesis generales, y también nos ofrece una representación integral de la función. [1]

Teorema de la Función Implícita (Versión Cauchy): *Sea una función $F(x, y)$ holomorfa en el bidisco $D(x_0, R_1) \times D(y_0, R_2) \subseteq \mathbb{C}^2$ con*

$$D_2F = \frac{\partial F}{\partial y}$$

Tenemos que si $F(x_0, y_0) = 0$ y $D_2F(x_0, y_0) \neq 0$, entonces existe un disco $D(x_0, r)$ y una única función holomorfa $f(x)$ definida en dicho disco con $f(x_0) = y_0$ tal que

$$F(x, f(x)) = 0$$

se cumple para $x \in D(x_0, r)$. Además, la función $f(x)$ es representada como

$$f(x) = \frac{1}{2\pi i} \oint_C \frac{D_2F(x, y)}{F(x, y)}\, dy$$

donde $C = \partial D(y_0, r_1)$ es un círculo elegido adecuadamente.

Cauchy demostró esta versión del teorema mediante técnicas de análisis complejo, luego no me voy a parar a demostrarlo. Sin embargo, también llegó a enunciar y demostrar el teorema de la función implícita mediante el método de mayorantes, que mencioné anteriormente. Lo enuncio a continuación, [1].

Teorema: *Sea la siguiente serie de potencias,*

$$F(x, y) = \sum_{j,k=0}^{\infty} a_{kj} x^j y^k$$

que suponemos absolutamente convergente para $|x| \leq R_1$, $|y| \leq R_2$. Si $a_{00} = 0$ y $a_{01} \neq 0$, entonces existe $r_0 > 0$ y una serie de potencias,

$$f(x) = \sum_{j=1}^{\infty} c_j x^j$$

tal que $f(x)$ *es absolutamente convergente para* $|x| \leq r_0$ *y*

$$F(x, f(x)) = 0.$$

Para esta sección se han usado los siguientes libros [1], [5] y [6].

1.4. Ulisse Dini

Hasta ahora, los teoremas de la función implícita que he descrito han sido para funciones de una única variable. Sin embargo, en 1878, el matemático italiano **Ulisse Dini** (1845-1918) describió de forma rigurosa por primera vez la versión que podemos encontrar en los libros de texto en la actualidad.

Ulisse Dini fue un matemático nacido en Pisa. Durante sus años de estudiante tuvo la oportunidad de tener a mentores como **Ottavio Mossotti** y **Enrico Betti**. El primero era físico matemático, muy influenciado por el trabajo de Lagrange, y enseñaba geodésica en la Universidad de Pisa. El segundo, profesor de física matemática, fue el tutor de la tesis de Dini.

En 1865, gracias a una beca, se fue a vivir a París, donde estudió bajo Charles Hermite y Joseph Bertrand, además de publicar diversos artículos. Volvió a Pisa un año más tarde, empezando allí su carrera como profesor en la Universidad Real de Pisa, como profesor de geodésica y análisis avanzado. Además de su carrera matemática, debido en parte a la problemática situación en la que se encontraba Italia durante aquella época, Dini se inició en la política en 1871, llegando incluso en 1890 a formar parte del parlamento italiano como senador. [3]

En 1877, Dini publicó el libro ***Lezioni di Analisi Infinitesimale***, el cuál contiene las lecciones dadas por él mismo durante el año académico anterior (1876-1877), en la Universidad de Pisa. Existen dos versiones de este libro, ya que fue publicado por dos editoriales distintas, pero en ambas se trata de un único volumen dividido en dos partes. La primera parte está basada en el cálculo diferencial y la segunda en el cálculo integral. La teoría de funciones implícitas se encuentra en el capítulo XIII de la parte 1ª, con título *Derivate e differenziali dei vari ordini di funzioni implicite di una o più variabili indipendenti.*

En este capítulo, Dini introduce la notación de funciones implícitas de dos o más variables, discutiendo sus propiedades de unicidad, existencia y regularidad $\mathcal{C}^1$, primero para una única función implícita con derivadas de orden igual o mayor que la primera y después, extendiendo sus resultados a sistemas de funciones implícitas.

El teorema de Dini nos asegura la existencia, regularidad y unicidad de funciones de $n \geq 1$ variables y con regularidad de clase $\mathcal{C}^r$ con $r \geq 1$, pero solo en campos escalares reales. Sin embargo posteriormente, Dini mencionó en uno de sus trabajos que el caso de campos complejos se podía deducir fácilmente, bastaba considerar la descomposición de una función compleja de la forma $f(x, y) = g(x, y) + ih(x, y)$, siendo g y h funciones reales. La versión de Dini [4], para el caso de una variable, se enuncia a continuación.

Teorema de la función implícita: *Sea Ω un conjunto abierto en $\mathbb{R}^2$ y sea $f : \Omega \longrightarrow \mathbb{R}$ una una función $\mathcal{C}^1$. Supongamos que existe un punto $(x_0, y_0) \in \Omega$ tal que:*

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

entonces debe existir un intervarlo real A con centro x_0, otro I con centro y_0 y una función $\varphi : A \longrightarrow \mathbb{R}$ de forma que:

1. *$A \times I \subset \Omega$,*

2. *para cualquier punto $(x, y) \in A \times I$ se tiene que $\frac{\partial f}{\partial y}(x, y) \neq 0$,*

3. *si $(x, y) \in A \times I$ entonces $f(x, y) = 0$, sí y solo sí $y = \varphi(x)$,*

4. *$y_0 = \varphi(x_0)$,*

5. *$\varphi \in \mathcal{C}^1(A)$ y para cualquier $x \in A$:*

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

Tras el trabajo realizado por Dini, otros autores han propuesto algunas mejoras de sus resultados. En primer lugar, mencionar a Elcia Sadun, que aún siendo poco conocido, fue el primero en destacar la importancia del trabajo que Dini había hecho sobre las funciones implícitas, además de realizar algunas mejoras y añadir generalizaciones, algunas de éstas con aplicaciones geométricas. Más tarde en 1899, E. Lindelöf hizo su propia contribución considerando de nuevo funciones analíticas de varias variables en el campo escalar complejo, probando el teorema mediante la expansión de series de potencias, que fue una restricción que el mismo Dini eliminó.

Otras muchas contribuciones se han hecho desde entonces, algunas por W.F. Osgood en 1901, É. Goursat en 1903 y E.H. Young en 1909, la versión del cual requería de unas hipótesis más débiles que las dadas por Dini, ya que se pedía un orden más bajo de diferenciabilidad [3].

Capítulo 2

Estudio comparativo de diferentes demostraciones del T.F.I.

Podemos decir que, a grandes rasgos, existen dos puntos de vista para demostrar el teorema de la función implícita clásico. El primero se basa en conceptos y resultados del cálculo de varias variables, mientras que el segundo usa nociones del análisis funcional consiguiendo llegar así a una demostración mucho más rápida y sencilla, en general.

En este capítulo voy a dar dos demostraciones basadas en el cálculo, ambas obtenidas del libro de Krantz [1], la primera de ellas es la que el propio Dini usó en su momento. También usaré el análisis funcional para dar una tercera demostración que, aún siendo algo más abstracta que las anteriores, nos va a permitir tener una versión del teorema más general, para esta he usado [7].

2.1. Demostración inductiva. Ulisse Dini

La demostración de en este capítulo, como bien dice el título es la dada por Ulisse Dini en 1870. Ésta se realiza mediante un procedimiento inductivo en el número de variables dependientes. Primero se demuestra el teorema de la función implícita para una sola variable dependiente, una ecuación y un número arbitrario de variables independientes.

Después mediante inducción, iremos tomando de nuevo una única ecuación con una variable dependiente a la que se le aplica la versión del teorema anterior. Así obtenemos una función definida de forma implícita que podemos sustituir en el resto de ecuaciones, de esta forma se acaba reduciendo el número de variables dependientes y ecuaciones a solo una.

A continuación se enuncia y demuestra el teorema de la función implícita para una variable dependiente:

Teorema de la función implícita con una variable dependiente : *Sea el abierto $\mathcal{O} \subseteq \mathbb{R}^n$, $f : \mathcal{O} \to \mathbb{R}$ función continuamente diferenciable y sea $p = (p_1, p_2) \in \mathbb{R}^{n-1} \times \mathbb{R}$, donde $p \in \mathcal{O}$ es un punto tal que:*

$$f(p) = 0 \text{ y } \frac{\partial f}{\partial x_n}(p) \neq 0.$$

Entonces tenemos que existe un conjunto abierto $\mathcal{U} \subset \mathbb{R}^{n-1}$ con $p_1 \in \mathcal{U}$ y una función continuamente diferenciable $F : \mathcal{U} \to \mathbb{R}$ tal que

$$p_2 = F(p_1) \text{ y } f[x, F(x)] = 0, \forall x \in \mathcal{U}.$$

Demostración:

Podemos suponer sin problema alguno que $\frac{\partial f}{\partial x_n}(p) > 0$. Por la continuidad de $\partial f / \partial x_n$, y utilizando un entorno más pequeño $\mathcal{O}$ de p en caso de necesidad, asumimos que:

$$\frac{\partial f}{\partial x_n}(q) > 0, \forall q \in \mathcal{O}. \tag{2.1}$$

Por esta razón tenemos que $f(p_1, .)$ es una función creciente en un cierto intervalo que contiene a r. Luego podemos encontrar $r_1 < r < r_2$ tal que,

$$f(p_1, r_1) < 0 < f(p_1, r_2) \tag{2.2}$$

se cumple. Usando ahora la continuidad de f, podemos encontrar un nuevo entorno $\mathcal{U}$ de p_1 tal que $\mathcal{U} \times [r_1, r_2] \subseteq \mathcal{O}$ y con

$$f(x, r_1) < 0 < f(x, r_2), \forall x \in \mathcal{U}. \tag{2.3}$$

Por el Teorema del Valor Intermedio, más concretamente el caso particular del Teorema de Bolzano, tenemos que para cada $x \in \mathcal{U}$ existe un único número y tal que $r_1 < y < r_2$ con $f(x, y) = 0$. Definimos F como $F(x) = y$ y debido a la unicidad de dicho y tenemos la continuidad de F. Además, se cumple,

$$\frac{\partial F}{\partial x_j} = -\frac{\partial f}{\partial x_j} / \frac{\partial f}{\partial x_n}, j = 1, 2 \ldots n-1. \tag{2.4}$$

Fijando el punto $x \in \mathcal{U}$ y $F(x) = y$, por la diferenciabilidad de la función f tenemos que:

$$f(x + se_j, y + t) - f(x, y) = s\frac{\partial F}{\partial x_i}(x, y) + t\frac{\partial F}{\partial x_n}(x, y) + \epsilon\sqrt{s^2 + t^2} \tag{2.5}$$

donde e_j es el vector j-ésimo de la base canónica y $\epsilon = \epsilon(s, t) \to 0$ cuando $(s, t) \to 0$. Tomando por tanto $t = F(x + se_j) - F(x)$ y recordando que $y = F(x)$ sustituyendo en (2.5) tenemos

$$f(x + se_j, F(x) + F(x + se_j) - F(x)) - f(x, F(x)) = 0$$
$$s\frac{\partial F}{\partial x_i}(x, y) + t\frac{\partial F}{\partial x_n}(x, y) + \epsilon\sqrt{s^2 + t^2} = 0$$

luego llegamos a:

$$|t| \left| \frac{\partial F}{\partial x_n}(x, y) \right| \leq |s| \left| \frac{\partial F}{\partial x_j}(x, y) \right| + |\epsilon||s| + |\epsilon||t| \tag{2.6}$$

donde tomando un valor de s suficientemente pequeño obtenemos que

$$|\epsilon| \leq \frac{1}{2} |\frac{\partial F}{\partial x_n}(x, y)| \ y \ |\epsilon| \leq 2|\frac{\partial F}{\partial x_j}(x, y)|$$

tal que

$$|t| \leq 6|s| \left| \frac{\frac{\partial F}{\partial x_j}(x, y)}{\frac{\partial F}{\partial x_n}(x, y)} \right| \tag{2.7}$$

se cumple, y por tanto tenemos que,

$$\left| \frac{F(x+sej)-F(x)}{s} + \frac{\frac{\partial F}{\partial x_j}(x,y)}{\frac{\partial F}{\partial x_n}(x,y)} \right| = \left| \frac{t}{s} + \frac{\frac{\partial F}{\partial x_j}(x,y)}{\frac{\partial F}{\partial x_n}(x,y)} \right| \leq |\epsilon| \left(1 + 6\frac{\sqrt{s^2+t^2}}{\frac{\partial F}{\partial x_n}(x,y)} \right) \tag{2.8}$$

Tomando el límite cuando $s \to 0$ en la expresión anterior, tenemos que $\frac{\partial F}{\partial x_j}$ existe en x y es el dado en la fórmula (2.4), luego ya tenemos que F es continuamente diferenciable como queríamos demostrar. □

Para poder demostrar la versión general del teorema, vamos a enunciar el siguiente lema, además de algunos conceptos y notación nuevos.

Lema: *Sea M una matriz real $m \times m$. Entonces tenemos que existe una matriz invertible real A tal que AM es triangular superior.*

Notación: Supongamos que tenemos un conjunto dado de ecuaciones:

$$f_i(x_1, \dots, x_j; y_1, \dots, y_n) = 0, i = 1, \dots n, \tag{2.9}$$

siendo las funciones f_i todas continuamente diferenciables en $\mathbb{R}$. Vamos a asumir que $(p;q) = (p_1, \dots, p_j; q_1, \dots q_n)$ es un punto el que (2.9) se cumple y además tenemos que

$$det \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \cdots & \frac{\partial f_1}{\partial y_n} \\ \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial y_1} & \cdots & \frac{\partial f_n}{\partial y_n} \end{pmatrix} \neq 0. \tag{2.10}$$

Tomamos dichas funciones $f_i(p; \cdot)$ como una aplicación $F : \mathbb{R}^n \to \mathbb{R}^n$ definida como sigue:

$$y \mapsto F(y) = (f_1(p;y), \dots, f_n(p;y)). \tag{2.11}$$

Luego aplicando el lema enunciado más arriba a la matriz de derivadas parciales de F podemos obtener mediante transformaciones lineales de F una nueva función:

$y \mapsto \tilde{F}(y) = (\tilde{f}_1(p;y), \ldots, \tilde{f}_n(p;y))$, de forma que, $\frac{\partial \tilde{f}_i}{\partial y_j} = 0$, cuando $i > j$.

No es restrictivo suponer que,

$$\frac{\partial f_i}{\partial y_j} = 0, i > j \tag{2.12}$$

De esta forma la matriz de derivadas parciales se nos ha quedado de la siguiente forma,

$$\begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} & \cdots & \frac{\partial f_1}{\partial y_n} \\ \frac{\partial f_2}{\partial y_1} & \frac{\partial f_2}{\partial y_2} & \cdots & \frac{\partial f_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_n}{\partial y_1} & \frac{\partial f_n}{\partial y_2} & \cdots & \frac{\partial f_n}{\partial y_n} \end{pmatrix} = \begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} & \cdots & \frac{\partial f_1}{\partial y_n} \\ 0 & \frac{\partial f_2}{\partial y_2} & \cdots & \frac{\partial f_2}{\partial y_n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \frac{\partial f_n}{\partial y_n} \end{pmatrix} \tag{2.13}$$

con esta transformación tenemos que el determinante de dicha matriz es simplemente la multiplicación de la diagonal, luego la ecuación (2.10) se escribe como

$$\prod_{i=1}^{n} \frac{\partial f_i}{\partial y_i} \neq 0 \tag{2.14}$$

Podemos enunciar ya, la citada versión general.

Teorema de la función implícita: *Existe un entorno $\mathcal{O} \subset \mathbb{R}^m$ del punto p y un conjunto de funciones $\delta_j : \mathcal{O} \to \mathbb{R}, j = 1, 2, \ldots, n$, tal que $\delta_j(p) = q_j, j = 1, 2, \ldots, n$ y con*

$$f_i[x; \delta_1(x), \delta_2(x), \ldots, \delta_n(x)] = 0, i = 1, 2, \ldots, n, \forall x \in \mathcal{O}. \tag{2.15}$$

Demostración:

La vamos a hacer mediante inducción como se comentó al principio de este apartado.

El caso $n = 1$, es el teorema de la función implícita para una sola variable dependiente (*TFI1* a partir de ahora), que ya tenemos probado. Supongamos que es cierto para el caso $n - 1$ y vamos a probarlo para n. Usando la nueva notación que definimos antes, por la expresión (2.14), tenemos que,

$$\frac{\partial f_n}{\partial y_n}(p; q) \neq 0 \tag{2.16}$$

Introducimos ahora las nuevas variables $y` = (y_1, y_2, \ldots, y_{n-1})$ y a $q' = (q_1, \ldots, q_{n-1})$.Por el *TFI1* aplicado a la ecuación,

$$f_n(x, y'; y_n) = 0 \tag{2.17}$$

para el punto $(p; q'; q_n)$, existe un entorno $\mathcal{U} \subset \mathbb{R}^{m+n-1}$ de (p, q') y una función continuamente diferenciable $F : \mathcal{U} \to \mathbb{R}$ tal que $F(p; q') = q_n$ y

$$f_n(x; y'; F(x; y')] = 0, \forall (x; y') \in \mathcal{U}. \tag{2.18}$$

Derivando f_n respecto de la variable $y_j, 1 \leq j \leq n-1$ obtenemos

$$\frac{\partial f_n}{\partial y_j} + \frac{\partial f_n}{\partial y_n}\frac{\partial F}{\partial y_j} = 0. \tag{2.19}$$

Evaluando ahora en el punto $x = p, y' = q'$, y usando (2.12) y (2.16), vemos que

$$\frac{\partial F}{\partial y_j}(p; q') = 0 \tag{2.20}$$

se cumple para $j = 1, \ldots, n-1$. Ahora para cada $i = 1, 2, \ldots, n-1$ definimos la función h_i como sigue,

$$h_i(x_1, \ldots, x_m, y_1, \ldots, y_n - 1) = f_i(x; y', F(x; y')). \tag{2.21}$$

Considerando el sistema de ecuaciones, $h_i(x_1, \ldots, x_m, y_1, \ldots, y_{n-1}) = 0$, para $i = 1, \ldots, n-1$, y para cada $j = 1, \ldots, n-1$, por (2.20), tenemos que,

$$\frac{\partial h_i}{\partial y_j}(p; q') = \frac{\partial f_i}{\partial y_j}(p; q'; q_n) + \frac{\partial f_i}{\partial y_n}(p; q'; q_n)\frac{\partial F}{\partial y_j}(p; q') = \frac{\partial f_i}{\partial y_j}(p; q') \tag{2.22}$$

y por la expresión (2.14),

$$det\begin{pmatrix} \frac{\partial h_1}{\partial y_1} & \frac{\partial h_1}{\partial y_2} & \cdots & \frac{\partial h_1}{\partial y_{n-1}} \\ \frac{\partial h_2}{\partial y_1} & \frac{\partial h_2}{\partial y_2} & \cdots & \frac{\partial h_2}{\partial y_{n-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial h_{n-1}}{\partial y_1} & \frac{\partial h_{n-1}}{\partial y_2} & \cdots & \frac{\partial h_{n-1}}{\partial y_{n-1}} \end{pmatrix} = det\begin{pmatrix} \frac{\partial f_1}{\partial y_1} & \frac{\partial f_1}{\partial y_2} & \cdots & \frac{\partial f_1}{\partial y_{n-1}} \\ \frac{\partial f_2}{\partial y_1} & \frac{\partial f_2}{\partial y_2} & \cdots & \frac{\partial f_2}{\partial y_{n-1}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_{n-1}}{\partial y_1} & \frac{\partial f_{n-1}}{\partial y_2} & \cdots & \frac{\partial f_{n-1}}{\partial y_{n-1}} \end{pmatrix} \neq 0. \tag{2.23}$$

Tenemos que las parciales de la primera matriz están evaluadas en $(p; q')$ mientras que las de la segunda, en $(p; q)$. Por inducción, tenemos que existe un entorno $\mathcal{V} \subset \mathbb{R}^m$

de p y una aplicación continuamente diferenciable $\phi_j : \mathcal{V} \to \mathbb{R}$ tal que $\phi_j(p) = q_j$ y con

$$h_i[x; \phi_1(x), \dots, \phi_{n-1}(x)] = 0, i = 1, 2, \dots, n-1, \forall x \in \mathcal{V}$$

Tomando ahora $\Phi(x) = (x.\phi_1(x), \dots, \phi_{n-1}(x))$ y $\mathcal{O} = \mathcal{V} \cap \Phi^{-1}(\mathcal{V})$ y definimos $\phi_n : \mathcal{O} \to \mathbb{R}$ como $\phi_n(x) = F[x; \phi_1, \dots, \phi_{n-1}(x)]$. Y por la definición que hicimos de h_i tenemos que

$$h_i(x_1, \dots, x_m, y_1, \dots, y_{n-1}) = f_i(x; y', F(x; y'))$$

luego vemos que se cumple $f_i(x; \phi_1(x), \dots, \phi_n(x)) = 0, i = 1, \dots, n, \forall x \in \mathcal{O}$. □

2.2. Demostración clásica

En este apartado se va a demostrar el teorema de la función implícita mediante el teorema de la función inversa. Ya que ambos son equivalentes, como veremos a continuación, bastará demostrar el segundo para tener la prueba del primero. Antes sin embargo, nos va a hacer falta describir algunas definiciones previas.

Empezaremos por la matriz jacobiana y su determinante, además de comentar el papel que desempeñan en el cálculo. Sean dos entornos abiertos $\mathcal{U}, \mathcal{V} \subset \mathbb{R}^N$ y sea la función $F : \mathcal{U} \to \mathcal{V} \in \mathcal{C}^1$, que la describimos como $F(x) = (f_1(x), \dots, f_N(x))$. Tenemos por tanto, que para un punto $p \in \mathcal{U}$, la *matriz jacobiana* de F en p es, simplemente $F'(p)$,

$$JacF(p) \equiv \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \frac{\partial f_1}{\partial x_2}(p) & \cdots & \frac{\partial f_1}{\partial x_N}(p) \\ \frac{\partial f_2}{\partial x_1}(p) & \frac{\partial f_2}{\partial x_2}(p) & \cdots & \frac{\partial f_2}{\partial x_N}(p) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_N}{\partial x_1}(p) & \frac{\partial f_N}{\partial x_2}(p) & \cdots & \frac{\partial f_N}{\partial x_N}(p) \end{pmatrix} \tag{2.24}$$

La matriz jacobiana juega a veces el mismo papel en el cálculo de varias variables que la primera derivada (una matriz 1x1) en el cálculo de una variable. Para el teorema de la función inversa también nos va a hacer falta el *jacobiano*, siendo éste el determinante de la matriz jacobiana,

$$det(JacF(p)) = det \begin{pmatrix} \frac{\partial f_1}{\partial x_1}(p) & \frac{\partial f_1}{\partial x_2}(p) & \cdots & \frac{\partial f_1}{\partial x_N}(p) \\ \frac{\partial f_2}{\partial x_1}(p) & \frac{\partial f_2}{\partial x_2}(p) & \cdots & \frac{\partial f_2}{\partial x_N}(p) \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_N}{\partial x_1}(p) & \frac{\partial f_N}{\partial x_2}(p) & \cdots & \frac{\partial f_N}{\partial x_N}(p) \end{pmatrix} \tag{2.25}$$

Veremos que la condición de que dicho determinante $|JacF(p)| = det(JacF(p))$ sea no nulo será la condición suficiente para que exista inversa de F en un entorno de p. Una notación que vamos usar respecto a las componentes de una función es la siguiente:

$$F(x) = F(x_1, \dots x_N) \equiv (f_1(x_1, \dots, x_N), \dots, f_M(x_1, \dots, x_N)), \tag{2.26}$$

usando la elección de $M \leq N$ argumentos posibles, que expresamos como $x_{i_1}, \dots, x_{i_M}$, y siendo el jacobiano,

$$det(JacF(p)) = \frac{\partial(f_1, \dots, f_M)}{\partial(x_{i_1}, \dots, x_{i_M})} = det \begin{pmatrix} \frac{\partial f_1}{\partial x_{i_1}} & \frac{\partial f_1}{\partial x_{i_2}} & \cdots & \frac{\partial f_1}{\partial x_{i_M}} \\ \frac{\partial f_2}{\partial x_{i_1}} & \frac{\partial f_2}{\partial x_{i_2}} & \cdots & \frac{\partial f_2}{\partial x_{i_M}} \\ \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_M}{\partial x_{i_1}} & \frac{\partial f_M}{\partial x_{i_2}} & \cdots & \frac{\partial f_M}{\partial x_{i_M}} \end{pmatrix} \tag{2.27}$$

Teorema de la función inversa: *Sea* $\mathcal{U} \subseteq \mathbb{R}^n$ *un abierto y sea* $G : \mathcal{U} \to \mathbb{R}^n \in \mathcal{C}^k$, *con* $k \geq 1$. *Sea* x_0 *un punto fijo de* $\mathcal{U}$, *suponemos que* $|JacG(x_0)| \neq 0$. *Entonces existe un entorno abierto* $\mathcal{V} \subseteq \mathcal{U}$ *de* x_0 *tal que:*

- *La restricción* $G|_{\mathcal{V}}$ *es inyectiva.*
- *El conjunto* $G(\mathcal{V}) = \mathcal{W}$ *es abierto.*
- *La inversa* G^{-1} *de* $G|_{\mathcal{U}}$ *es de clase* $\mathcal{C}^k$.

A continuación procedo a enunciar y demostrar el teorema de la función implícita mediante el teorema de la función inversa. Tras esta demostración, haré la prueba del teorema de la función inversa dando así por finalizada la demostración del primero.

Teorema de la función implícita: *Sea*

$$F(x) = F(x_1, \dots, x_n) \equiv (f_1(x_1, \dots, x_n), \dots, f_m(x_1, \dots, x_n))$$

una función de clase $\mathcal{C}^k$, con $k \geq 1$, definida en un conjunto abierto $\mathcal{U} \subseteq \mathbb{R}^n$, y tal que tome valores en $\mathbb{R}^m$, suponiendo que $1 \leq m < n$, y definimos $q = n - m$. Sea $x' = (x'_1, \dots, x'_n)$ un punto fijo de $\mathcal{U}$ y sea $x = (x_1, \dots, x_n)$ un punto cualquiera también de $\mathcal{U}$. Tomamos por tanto:

$$x_a = (x_1, \dots, x_q) \text{ las primeras } q \text{ coordenadas de } x,$$
$$x'_a = (x'_1, \dots, x'_q) \text{ las primeras } q \text{ coordenadas de } x'.$$

Suponemos que,

$$\frac{\partial(f_1, \dots, f_m)}{\partial(x_{q+1}, \dots, x_n)}(x') \neq 0 \tag{2.28}$$

Entonces existe un entorno $\mathcal{V}$ de x', un conjunto abierto $\mathcal{W} \subseteq \mathcal{R}^q$ que contiene a x'_a y funciones $g_1, \dots, g_m \in \mathcal{C}^k$ en $\mathcal{W}$ tales que:

$$F(x_1, \dots, x_q; g_1(x_a), \dots, g_m(x_a)) = 0, \forall x_a \in \mathcal{W}.$$

Además, $g_1, \dots, g_m$ son las únicas funciones que cumplen,

$$\{x \in \mathcal{V} : F(x) = 0\} = \{x \in \mathcal{V} : x_a \in \mathcal{W}, x_{q+i} = g_i(x_a), i = 1, \dots, m\}.$$

Demostración:

Tenemos que la función que hemos definido F es al menos, $\mathcal{C}^1$, luego podemos definir su jacobiano, el cuál es continuo y en él existe un entorno $\mathcal{U}$ de x', en el que éste no se anula.

Sea la transformación $G : \mathcal{U} \to \mathbb{R}^n$ dada por:

$$G(x) = (x_1, \dots, x_n, f_1(x), \dots, f_m(x)),$$

con lo que tenemos que $G \in \mathcal{C}^k$, al igual que G. Su matriz jacobiana de F es la siguiente:

$$Jac_G = \begin{pmatrix} 1 & 0 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 1 & \cdots & 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 1 & 0 & \cdots & 0 \\ \frac{\partial f_1}{\partial x_1} & \frac{\partial f_1}{\partial x_2} & \cdots & \frac{\partial f_1}{\partial x_q} & \frac{\partial f_1}{\partial x_{q+1}} & \cdots & \frac{\partial f_1}{\partial x_n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \ddots & \vdots \\ \frac{\partial f_m}{\partial x_1} & \frac{\partial f_m}{\partial x_2} & \cdots & \frac{\partial f_m}{\partial x_q} & \frac{\partial f_m}{\partial x_{q+1}} & \cdots & \frac{\partial f_m}{\partial x_n} \end{pmatrix}$$

Tenemos por tanto que el determinante de dicha matriz, debido a su peculiar forma, es simplemente del bloque inferior derecho $m \times m$, que se trata del jacobiano (2.28), por tanto $det(Jac_G(x)) \neq 0$, $\forall x \in \mathcal{U}$.

Ahora, usando el teorema de la función inversa que hemos enunciado más arriba, concluimos que existe un entorno $\mathcal{V}$ de x' de forma que $G(\mathcal{V})$ es un conjunto abierto y tal que la restricción $G|_{\mathcal{V}}$ tiene una función inversa $G^{-1} \in \mathcal{C}^k$.

Por último, si definimos $(x_a, 0) = (x_1, \ldots, x_q, 0, \ldots, 0)$ y el conjunto,

$$R = \{x_a : (x_a, 0) \in G(\mathcal{V})\}.$$

Como el conjunto $G(\mathcal{V})$ es un abierto de $\mathbb{R}^n$, tenemos que R es también un abierto. Ahora, para cada $x_a \in R$, sea

$$g_i(x_a) = f_{q+1}(x_a, 0), \forall i = 1, \ldots, m.$$

Para $x \in \mathcal{V}$, $G(x) = 0$ si y solo si, $x_a \in R$ y $G(x) = (x_a, 0)$, ya que $G\_\mathcal{V}$ y G^{-1} son inversas, $G(x) = (x_a, 0)$, si y solo si $x = G^{-1}(x_a, 0)$. $\square$

Demostración (Teorema función Inversa):

PASO 1: La función F es localmente inyectiva:

Sea $x_0 \in \mathcal{U}$ un punto fijo pero arbitrario y sea h la matriz inversa de Jac_G, y notamos por $\|h\|$ la norma de h, la cual es un operador lineal. Se define

$$c = \frac{1}{\|h(x_0)\|} \tag{2.29}$$

Sea L la matriz jacobiana de G en el punto x_0, y tomamos $\tilde{G}(t) = G(t) - L$. Luego para $s, t \in \mathcal{U}$:

$$G(s) - G(t) = L(s) - L(t) + [\tilde{G}(s) - \tilde{G}(t)], \ pero$$

$$\frac{|\tilde{G}(s) - \tilde{G}(t)|}{|s-t|} \to 0, \ cuando \ s, t \to 0.$$

Luego para un $\epsilon > 0$, tenemos que cuando el s esté cerca de t,

$$|G(s) - G(t)| \geq |L(s) - L(t)| - \epsilon \dot{|} s - t|$$

Sin embargo,

$$|L(s) - L(t)| \geq c \dot{|} s - t|. \tag{2.30}$$

por tanto tenemos que, $|G(s) - G(t)| \geq (c - \epsilon)\dot{|} s - t|$, donde tomando $\epsilon = \frac{c}{2}$ obtenemos:

$$|G(s) - G(t)| \geq \frac{c}{2} \dot{|} s - t|.$$

Con estos cálculos llegamos a la conclusión de que $G(s) = G(t)$ implica necesariamente que $s = t$, con lo que hemos demostrado que G es localmente inyectiva en el conjunto $\mathcal{U}$.

PASO 2: El conjunto $G(\mathcal{U})$ es abierto:

Sea $W = G(\mathcal{U})$. Sea x^* un punto cualquiera de W, vamos a ver que x^* tiene un entorno en W. Por lo visto en el Paso 1, podemos tomar $t* \in \mathcal{V}$, de forma que $G(t^*) = x^*$, siendo $\mathcal{V}$ un entorno de t^* en el que G es inyectiva. Sea ahora B una bola abierta con centro t^*, y tal que su clausura esté dentro de $\mathcal{V}$ y con $\partial B = S$, su frontera.

La inyectividad implica que $x^* \notin G(S)$. Además por ser G continua y S un conjunto compacto, tenemos que $G(S)$ es también compacto. Definimos

$$\sigma^* = \frac{1}{2} dist(x*, G(S)).$$

Sea ahora $W^* = B(x^*, \sigma^*)$. Fijando el punto arbitrario $x \in W^*$, y con $t \in S$,

$$2\sigma^* \leq |x^* - G(t)| \leq |x^* - x| + |x - G(t)|.$$

y como tenemos que $|x^* - x| < \sigma^*$, $\sigma^* < |x - G(t)|$, $\forall t \in S$. Sea ahora un $t \in \mathcal{U}$, definimos,

$$F(t) = |x - G(t)|^2 = \sum_{j=1}^{Q} [x_j - g_j(t)]^2,$$

donde g_j son las componentes de G. Por tanto, $F \subset \mathcal{C}^k$ y debe existir un mínimo en la bola compacta $\bar{B}$, pero

$$F(t^*) = |x - x^*|^2 < [\sigma^*]^2, \ y$$

$$F(t) > [\sigma^*]^2 \forall t \in S.$$

Por tanto, el valor mínimo de F en $\bar{B}$ es menor que $[\sigma^*]^2$, luego debe alcanzarse en algún punto interior $\tilde{t} \in W^*$, y en ese caso la derivada parcial de F en $\tilde{t}$ debe ser 0,

$$\frac{\partial F(t)}{\partial t_k} = -2 \sum_{j=1}^{Q} (x_j - g_j(t)) \frac{\partial f_j(t)}{\partial t_k}(t),$$

donde tomando $c_j = x_j - g_j(\tilde{t})$, obtenemos,

$$\sum_{j=1}^{Q} c_j \frac{\partial f_j(\tilde{t})}{\partial t_k}, \ para\ cada\ k.$$

Como además $det(Jac_G(\tilde{t})) \neq 0$, tenemos que los vectores columna

$$\frac{\partial g_1}{\partial t_j}(\tilde{t}), \dots, \frac{\partial g_Q}{\partial t_j}(\tilde{t}), j = 1, \dots, Q,$$

son linealmente independientes. Con esto concluimos que $c_j = 0$, $j = 1, \dots Q$, y por tanto, $x = G(\tilde{t})$.

Hemos demostrado así que si $x \in B$, entonces $x = G(\tilde{t})$ para algún $\tilde{t} \in B$, o $x \in G(B)$. Luego $W^* \subseteq G(B) \subseteq W$, luego $W = G(\mathcal{V})$ es un conjunto abierto.

Notación: Durante el resto de demostración, se va a dejar fijo el entorno $\mathcal{V}$ de x_0, donde G es inyectiva y por tanto $det(Jac_G) \neq 0$.

PASO 3: La función $(G|_{\mathcal{V}})^{-1} \in \mathcal{C}^1$:

Tenemos que $(G|_{\mathcal{V}})^{-1} \subset \mathcal{C}^1$ existe, por la inyectividad local que se probó en el Paso 1. Sea $x^* \in W$, $t^* = (G|_{\mathcal{V}}^{-1}(x^*))$ y $L^* = Jac_G(t^*)$. Vamos a ver que G_{-1} es diferenciable en x^* y que $Jac_G^{-1}(x^*) = (L^*)^{-1}$. Sea $\tilde{c} = \frac{1}{\|(L^*)^{-1}\|}$. Para algún $\epsilon > 0$, existe una bola $\tilde{B} = B(t^*, r^*) \subseteq \mathcal{V}$ tal que:

$$|G(t) - G(t^*) - L^*(t - t^*)| \leq \frac{\epsilon \cdot \tilde{c} \cdot c}{2}|t - t^*|, \forall t \in \tilde{B}. \tag{2.31}$$

La constante c es la misma que la definida en el Paso 1, donde,

$$|L(s) - L(t)| \geq c|s - t|.$$

Por lo demostrado en el Paso 2, tenemos que existe un entorno $B^* = B(x^*, s)$ tal que $B^* \subseteq G(\tilde{B})$. Sea $x \in B^*$, entonces $x = G(t)$ para algún $t \in \tilde{B}$. Como además, $x^* = G(t^*)$, por el Paso 1 tenemos que,

$$\frac{c}{2} \cdot |t - t^*| \leq |x - x^*|. \tag{2.32}$$

Es más, como $t = G^{-1}(x^*)$, vemos que,

$$L^*[G^{-1}(x) - G^{-1}(x^*) - (L^*)^{-1}(x - x^*)] \leq |G(t) - G(t^*) - L^*(x - x^*)|.$$

Como tenemos que $\tilde{c}|\omega| \leq L^*(\omega)$ para todo $\omega \in \mathbb{R}^Q$, tenemos que,

$$\tilde{c}|G^{-1}(x) - G^{-1}(x^*) - (L^*)^{-1}(x - x^*)| \leq |G(t) - G(t^*) - L^*(x - x^*)|.$$

Usando ahora (2.31) y (2.32), para cada $x \in \tilde{B}$,

$$|G^{-1}(x) - G^{x^*} - (L^*)^{-1}(x - x^*)| \leq \epsilon|x - x^*|$$

Por tanto, hemos visto que G^{-1} es diferenciable en x^* y $Jac_G^{-1}(x^*) = L^{-1}$. Resumiendo, tenemos que G^{-1} es una función diferenciable y,

$$Jac_G^{-1}(x^*) = Jac_G(G^{-1}(x)), \forall x \in W, \tag{2.33}$$

de donde deducimos que G^{-1} es continua.Y como además, cada $\frac{\partial g_i}{\partial x_j}$ es continua, la composición $\frac{\partial g_i}{\partial x_j} \circ G$, también lo es. Luego por (2.33) y por la regla de Cramer, tenemos que todas las derivadas parciales,

$$\frac{\partial (G^{-1})_i}{\partial x_j} \textit{ son continuas.}$$

Siendo ahora G de clase $\mathcal{C}^m$, entonces cada $\frac{\partial g_i}{\partial x_j}$ es de clase $\mathcal{C}^{m-1}$, y por tanto la composición $\frac{\partial g_i}{\partial x_j} \circ G^{-1} \in \mathcal{C}^{m-1}$. Por tanto, $\frac{\partial g_i}{\partial x_j} \circ G$ es de clase $\mathcal{C}^{m-1}$ y a la vez $G^{-1} \in \mathcal{C}^m$. Así, de forma inductiva tenemos que si $G \in \mathcal{C}^k$, entonces también lo será G^{-1}, con lo que se termina la demostración de este último paso, y por consiguiente, la del teorema de la función inversa. $\square$

2.3. Demostración con el método de aproximaciones sucesivas

La tercera y última demostración que se va a dar es algo más abstracta en comparación con las anteriores, ya que emplea el análisis funcional, pero nos va a permitir tener una versión del teorema más general. Como bien dice el nombre de esta sección, se va a usar el método de aproximaciones sucesivas, es decir, mediante iteraciones vamos a ir mejorando el resultado partiendo de un punto inicial. Con esto se conseguirá una sucesión de funciones cuyo límite va a acabar siendo la función implícita que necesitábamos.

Antes de demostrar nada, nos es necesario enunciar algunas definiciones que se van a emplear durante esta sección.

Sea X un espacio métrico completo y sea δ su métrica. Tenemos que la función $F : X \to X$ se define como una *contracción* si existe una constante $0 < c < 1$ tal que

$$\delta(F(x), F(y)) \leq c \cdot \delta(x, y), \forall x, y \in X.$$

El hecho de que $c < 1$ nos asegura que la imagen de un conjunto bajo F es *contraída*, así los puntos en $F(x)$ van a estar más pegados entre ellos que los puntos originales en X.

El siguiente teorema que voy a enunciar se considera el teorema básico de funciones contractivas.

Teorema (Punto fijo de Banach): *Sea $F : X \to X$ una contracción con X un espacio métrico completo. Entonces F va a tener un único punto fijo, $x_0 \in X$ que cumple $F(x_0) = x_0$.*

Demostración: Para un punto $p \in X$ cualquiera, vamos a definir una sucesión de forma inductiva de la siguiente forma,

$$x_1 = p, x_2 = F(x_1), \dots, x_i = F(x_{i-1}).$$

A continuación se prueba que dicha sucesión $\{x_i\}$ es de Cauchy en X. Sea $i \geq 1$:

$$\begin{aligned} \delta(x_i, x_{i+1}) = \delta(F(x_{i-1}), F(i_{i-2})) \leq c \cdot \delta(x_{i-1}, x_{i-2}) = \\ = c \cdot \delta(F(x_{i-2}), F(x_{i-3})) \leq c^2 \cdot \delta(x_{i-2}, x_{i-3}) \leq \dots \leq \\ \leq c^{i-1} \cdot \delta(x_1, x_0) \end{aligned}$$

De aquí obtenemos,

$$\begin{aligned} \delta(x_{i+k}, x_i) \leq \delta(x_{i+k}, x_{i+k-1}) + \delta(x_{i+k-1}, x_{i+k-2}) + \delta(x_{i+1}, x_i) \leq \\ \leq [c^{i+k-1} + c^{i+k-2} + \dots c^j] \cdot \delta(x_1, x_0) \leq c^i \cdot \frac{1}{1-c} \cdot \delta(x_1, x - 0). \end{aligned}$$

Luego para un $\epsilon > 0$ y para un i lo suficientemente grande, tenemos que cualquier sea el valor de $k \geq 1$, $\delta(x_{i+k}, x_i) < c$, luego la sucesión $\{x_i\}$ es de Cauchy. Ahora, como el espacio X es completo, tenemos que la sucesión tiene límite, al que vamos a denotar x_0, y además,

$$F(x_0) = F(\lim_{i \to \infty}(x_i) = \lim_{i \to \infty}(F(x_i)) = \lim_{i \to \infty}(x_{i+1}) = x_0.$$

Luego hemos probado que efectivamente, x_0 es un punto fijo de F, ahora nos falta probar la unicidad de este. Para ello, vamos a suponer que existe otro punto fijo, para llegar a la conclusión de que tienen que ser el mismo. Sea $\tilde{x}_0$ el otro punto fijo, entonces como F es una contracción tenemos que existe una métrica tal que,

$$\delta(x_0, \tilde{x}_0) = \delta(F(x_0), F(\tilde{x}_0)) \leq c \cdot \delta(x_0, \tilde{x}_0).$$

Y ya que tenemos la condición de $0 < c < 1$, las únicas posibilidades son $\delta(\tilde{x}, \tilde{y}) = 0$ o que $\tilde{x} = \tilde{y}$, con lo probamos la unicidad, y por consiguiente el teorema. $\square$

Para la demostración de la siguiente versión del teorema de la función implícita he usado el trabajo de Castaño Muñoz [7]. Enuncio antes el siguiente lema que va a ser usado en la demostración.

Lema: *Sea f un homeomorfismo y diferenciable en a. Entonces tenemos que la función $g = f^{-1}$ es diferenciable en $b = f(a)$ si y solo si $f'(a)$ es un isomorfismo.*

Teorema de la función implícita: *Sean $\mathcal{U} \subseteq \mathbb{R}^n$ y $\mathcal{V} \subseteq \mathbb{R}^m$, dos conjuntos abiertos, y sea la función $F : \mathcal{U} \times \mathcal{V} \to \mathbb{R}^m$ diferenciable de forma que existan $a \in \mathcal{U}$ y $b \in \mathcal{V}$, tal que $F(a, b) = 0$. Sea M la matriz jacobiana de $F(a, \cdot)$ en b invertible. Entonces existen entornos abiertos de a y b respectivamente, $\mathcal{U}_0 \subset \mathcal{U}$ y $\mathcal{V}_0 \subset \mathcal{V}$, tales que para todo $x \in \mathcal{U}_0$ existe un único $y \in \mathcal{V}_0$ que cumple $F(x, y) = 0$, y además, existe una aplicación diferenciable $u : \mathcal{U}_0 \to \mathcal{V}_0$ tal que:*

- $u(a) = b$, $(x, u(x)) \in \mathcal{U} \times \mathcal{V}$,
- $F(x, u(x)) = 0$, $\forall x \in \mathcal{U}$.

Demostración: Sea la sucesión de funciones definida como sigue:

$$\begin{cases} u_0(x) = y_0, \\ u_n(x) = f(x, u_{n-1}(x)) \in \mathcal{C}(\mathcal{U}_0, \mathbb{R}^m). \end{cases}$$

Donde $f(x, y) = y - M^{-1}(F(x, y))$, que verifica la condición $f(x, y) = y$ si y solo si $F(x, y) = 0$. Consideremos ahora la aplicación, con x fijo:

$$\begin{array}{rccc} H\colon & \mathcal{C}^0(\mathcal{U}_0, \mathbb{R}^m) & \longrightarrow & \mathcal{C}^0(\mathcal{U}_0, \mathbb{R}^m) \\ & v & \longmapsto & f(x, v(x)) \end{array}$$

Para un punto fijo de esta aplicación se va a tener que:

$$H(v) = v \Leftrightarrow f(x, v(x)) = v(x) \Leftrightarrow F(x, v(x)) = 0.$$

Veamos ahora que para un $\mathcal{U}_0$ lo suficientemente pequeño la aplicación H es contractiva, es decir,

$$\|H(v_1) - H(v_2)\|_{\infty,\mathcal{U}_0} \leq c \cdot \|v_1 - v_2\|_{\infty,\mathcal{U}_0}, \text{ para } 0 < c < 1, \text{ luego,}$$

$$\begin{aligned}
\|H(v_1)-H(v_2)\|_{\infty,\mathcal{U}_0} &= \sup_{x\in\mathcal{U}_0}|H(v_1)(x)-H(v_2)(x)|_{\mathbb{R}^m} = \\
&= \sup_{x\in\mathcal{U}_0}|f(x,v_1(x))-f(x,v_2(x))| \leq (*) \\
&\leq \sup_{x\in\mathcal{U}_0}\frac{1}{2}\cdot|v_1(x)-v_2(x)|_{\mathbb{R}^m} = \\
&= \frac{1}{2}\cdot\|v_1-v_2\|_{\infty,\mathcal{U}_0}.
\end{aligned}$$

En $(*)$ hemos usado la igualdad, $d_2f(a,b) = Id_{\mathbb{R}^m} - M^{-1}(d_2F(a,b)) = 0$, y empleando el teorema de los incrementos finitos, de forma que,

$$\|d_2f(x,y)\| \leq \frac{1}{2}, \forall(x,y)\in\mathcal{U}_0\times\mathcal{V}_0,$$

concluimos que existe un punto fijo en nuestra aplicación.

Para finalizar, nos queda ver que como $F(x,u(x))=0$, tenemos que,

$$\left\{\begin{array}{c}
\frac{\partial F}{\partial x_1}+\frac{\partial F}{\partial x_{n+1}}\cdot\frac{\partial u_1}{\partial x_1}+\dots\frac{\partial F}{\partial x_{n+m}}\cdot\frac{\partial u_m}{\partial x_1}=0 \\
\frac{\partial F}{\partial x_2}+\frac{\partial F}{\partial x_{n+1}}\cdot\frac{\partial u_1}{\partial x_2}+\dots\frac{\partial F}{\partial x_{n+m}}\cdot\frac{\partial u_m}{\partial x_2}=0 \\
\dots \\
\frac{\partial F}{\partial x_n}+\frac{\partial F}{\partial x_{n+1}}\cdot\frac{\partial u_1}{\partial x_n}+\dots\frac{\partial F}{\partial x_{n+m}}\cdot\frac{\partial u_m}{\partial x_n}=0
\end{array}\right.$$

con lo que podemos concluir por la regla de Cramer, que las parciales, $\frac{\partial u}{\partial x_1},\dots,\frac{\partial u}{\partial x_n}$, pueden ser expresadas como funciones continuas, y efectivamente $u\in\mathcal{C}^1$. Solo nos queda por demostrar que esas parciales existen, para ello vamos a emplear el lema enunciado antes.

Como teníamos que $F(x,u(x))=0$, $\frac{\partial F}{\partial y}(x,y)$ se trata de un difeomorfismo de $\mathbb{R}^m$ en un entorno de (a,b). Por tanto, $G(x,u(x)) = (x,F(x,u(x)))$ es una aplicación inyectiva y su diferencial es un isomorfismo en un entorno de (a,b).

Luego, G y G^{-1} son difeomorfismos de $\mathcal{C}^1$, por lo que el lema nos asegura la existencia de las parciales de u, con lo que queda por terminada la demostración. $\square$

Capítulo 3

Generalizaciones del T.F.I. y aplicaciones

Las siguientes generalizaciones y aplicaciones del teorema de la función implícita han sido obtenidas del libro de Krantz [1].

3.1. Ecuaciones diferenciales ordinarias

Sabemos que existe una relación directa entre el teorema de la función implícita y la teoría de ecuaciones diferenciales. Picard, al demostrar el teorema de existencia de ecuaciones diferenciales ordinarias, dio lugar a que Goursat demostrase de la misma manera, mediante una prueba iterativa, el teorema de la función implícita.

Más tarde, en el siglo XX, John Nash fue de los primeros en emplear el teorema de la función implícita para el estudio de ecuaciones en derivadas parciales. Sin embargo, para esta sección me voy a centrar, por simplicidad, solo en las ecuaciones diferenciales ordinarias. Se va a demostrar cómo el teorema de existencia de soluciones de ecuaciones diferenciales ordinarias puede ser usado para probar nuestro teorema de la función implícita.

Teorema (Cauchy-Peano): *Si* $F(t,x), (t,x) \in \mathbb{R}\times\mathbb{R}^n$, *es una función continua en la región* $(t_0-a, t_0+a)\times\mathbb{B}(x_0,r)$, *con* $a, r > 0$, *entonces existe una solución* $x(t)$ *de*

$$\frac{dx}{dt} = F(t,x), \quad x(t_0) = x_0,$$

en un cierto intervalo (t_0-h, t_0+h).

Observación: Recalcar que el teorema anterior no nos asegura la unicidad de la solución, y es que para funciones F solo continuas no tiene por qué ser la única solución del problema. Sea por ejemplo el problema $x' = x^{2/3}, x(0) = 0$. Tenemos que existen dos soluciones $x(t)$ verificando el problema, $x \equiv 0$ y $x \equiv (t/3)^3$.

Para obtener la unicidad nos bastará con añadir la condición de que F sea lipschitziana respecto a x, pasando a convertirse el teorema anterior en el Teorema de Existencia y Unicidad de Picard.

Ahora podemos dar una demostración del teorema de la función implícita como un corolario del teorema de Picard.

Teorema: *Supongamos que* $\mathcal{U} \subset \mathbb{R}^{n+1}$, *y sea* $G : \mathcal{U} \to \mathbb{R}^n$ *una función* $\mathcal{C}^1$. *Si* $G(t_0, x_0) = 0, (t_0, x_0) \in \mathcal{R}\times\mathcal{R}^n$, *y tal que la matriz* $(n\times n)$, *siendo* G_i *la componente i-ésima de* G,

$$\left(\frac{\partial G_i}{\partial x_j}(t_0, x_0)\right)_{i,j=1,2,\ldots,n}$$

sea no singular, entonces existe un intervalo abierto (t_0-h, t_0+h) *y una función continuamente diferenciable* $g : (t_0-h, t_0+h) \to \mathbb{R}^n$ *tal que* $g(t_0) = x_0$ *y* $G(t, g(t)) = 0$, $\forall t \in \mathbb{R}$.

Demostración:

Consideremos en primer lugar el caso para $n = 1$. Elegimos un $a, r > 0$ de forma que $(t_0-a, t_0+a)\times(x_0-r, x_0+r) \subseteq \mathcal{U}$ y tal que $(\partial G/\partial x)(t,x)$ sea distinto de cero en $(t_0-a, t_0+a)\times(x_0-r, x_0+r)$. Definimos ahora $F : (t_0-a, t_0+a)\times(x_0-r, x_0+r) \to \mathbb{R}$ como,

$$F(t,x) = -\frac{\partial G}{\partial t}(t,x)/\frac{\partial G}{\partial x}(t,x). \tag{3.1}$$

Como F es continua, podemos aplicar el teorema de existencia de ecuaciones diferenciales para concluir que existe una solución del problema

$$\frac{dx}{dt} = F(t,x), \quad x(t_0) = x_0,$$

definida en un intervalo $(t_0 - h, t_0 + h)$. Sea por tanto $g : (t_0 - h, t_0 + h) \to \mathbb{R}^n$ tal que $g(t) = x(t)$. Nótese además que $g(t_0) = x_0$ y que,

$$\begin{aligned} g'(t) &= \frac{dx}{dt}(t) = F(t, x(t)), \\ &= -\frac{\partial G}{\partial t}(t, g(t)) / \frac{\partial G}{\partial x}(t, g(t)). \end{aligned}$$

Como $g(t_0) = x_0$, tenemos que $G(t_0, g(t_0)) = G(t_0, x_0)$, y por lo anterior,

$$\frac{d}{dt} G(t, g(t)) = \frac{\partial G}{\partial t}(t, g(t)) + g'(t) \frac{\partial G}{\partial x}(t, g(t)) = 0.$$

Luego tenemos que $G(t, g(t)) = 0$ en el intervalo $(t_0 - h, t_0 + h)$.

Para el caso de $n > 1$, basta elegir $a, r > 0$ tal que $(t_0 - a, t_0 + a) \times \mathbb{B}(x_0, r) \subseteq \mathcal{U}$ de forma que la matriz $(n \times n)$,

$$D_x G = \left(\frac{\partial G_i}{\partial x_j}(t_0, x_0) \right)_{i,j=1,2,\ldots,n}$$

tenga determinante distinto de 0 en $(t_0 - a, t_0 + a) \times \mathbb{B}(x_0, r)$. Y por último sustituimos en (3.1), para obtener,

$$F(t,x) = -\left[D_x G(t + t_0, x + x_0) \right]^{-1} \left(\frac{\partial G}{\partial t}(t + t_0, x + x_0) \right),$$

con lo que la demostración es similar al caso $n = 1$. $\square$

Observación: La demostración del teorema de la función implícita que acabo de dar solo es válida para el caso de una variable independiente. Luego para el caso de tener una sola variable dependiente y varias independientes se tendría que cambiar la función $F(t,x)$ que es usada en la prueba, por un sistema de ecuaciones en derivadas parciales de primer orden. Para el caso general, es decir varias variables dependientes e independientes, se podrá usar el mismo método que usa Ulisse Dini en su demostración (dada en la sección 2.1 de este trabajo).

3.2. Algunos casos singulares del T.F.I.

Tanto el teorema de la función implícita como el de la inversa requieren que la función con la que se trabaje sea continuamente diferenciable, además de la condición de que una cierta matriz jacobiana tenga un determinante distinto de cero.

En esta sección se van a tratar algunos casos en los que esto no ocurre, es decir, en el que la matriz jacobiana es degenerada, y ver cuando se va a poder aplicar dichos teoremas igualmente. Pero antes es necesario exponer algunos preliminares.

Preliminares:

Sean $F(x, y, z)$ y $G(x, y, x)$ funciones $\mathcal{C}^1$ en un entorno del origen y tal que cumplan $F(0, 0, 0) = G(0, 0, 0) = 0$. Tenemos que las siguientes ecuaciones,

$$F(x, y, z) = 0, \quad G(x, y, z) = 0 \tag{3.2}$$

son de tres variables, luego el teorema de la función implícita usual nos dice que dos de dichas variables, pongamos por ejemplo la x y la y, van a poder ser expresadas como funciones de la tercera variable z. Como ya sabemos, la hipótesis para poder aplicar el teorema es que la matriz jacobiana sea no degenerada en el origen,

$$\mathcal{J} = det \begin{pmatrix} F_x & F_y \\ G_x & G_y \end{pmatrix} \tag{3.3}$$

teniendo en cuenta que las funciones F y G son una aplicación tal que,

$$(x, y) \quad \longmapsto \quad (F(x, y, z), G(x, y, z)).$$

Se puede ver que hemos tomado a la variable z como un parámetro. Además, como se ha hecho en el (3.3), a lo largo de este apartado los subíndices se van a emplear para referirse a las derivadas parciales de las funciones, y el superíndice del 0 para cuando estas estén evaluadas en el origen, es decir,

$$G_y^0 = \frac{\partial G}{\partial y}(0, 0, 0).$$

En esta sección, incluso sin tener la hipótesis de $\mathcal{J} \neq 0$, se van a obtener las expresiones de x e y respecto de z en un entorno el origen $(0,0,0)$. Por tanto, de la misma forma que en el teorema de la función implícita, si las funciones $x(z)$ e $y(z)$ existen y son diferenciables, entonces podremos encontrar (mediante la regla de la cadena) sus respectivas derivadas dx/dz y dy/dz, que para $z = 0$ satisfacen,

$$F_x \frac{dx}{dz} + F_y \frac{dy}{dz} + F_z = 0,$$
$$G_x \frac{dx}{dz} + G_y \frac{dy}{dz} + G_z = 0.$$

Luego es completamente necesario (para tener su existencia) que la matriz,

$$\begin{pmatrix} F_x^0 & F_y^0 & F_z^0 \\ G_x^0 & G_y^0 & G_z^0 \end{pmatrix} \tag{3.4}$$

tenga el mismo rango que la matriz jacobiana en el origen,

$$\mathcal{J}^0 = \begin{pmatrix} F_x^0 & F_y^0 \\ G_x^0 & G_y^0 \end{pmatrix}. \tag{3.5}$$

Para el caso en que la matriz (3.4) tenga un rango mayor que la de (3.5), entonces tendremos que la solución diferenciable que estamos buscando no existe.

3.2.1. Matriz Jacobiana de rango 1

Para esta sección vamos a asumir que el rango de (3.4) y de (3.5) es 1. Luego al menos una de las componentes de (3.5) es distinta de 0. Por simplicidad, vamos a suponer que,

$$\frac{\partial F}{\partial x}(0,0,0) = F_x^0 \neq 0.$$

Como (3.4) tiene rango 1 y por haber asumido que $F_x^0 \neq 0$, tenemos que,

$$G_y^0 = \frac{G_x^0}{F_x^0} F_y^0, G_z^0 = \frac{G_x^0}{F_x^0} F_z^0. \tag{3.6}$$

Vamos a cambiar el sistema de ecuaciones de antes $F(x,y,z) = 0$, $G(x,y,z) = 0$, por uno equivalente pero más simple.

$$H(x,y,z) = G(z,y,z) - \frac{G_x^0}{F_x^0} F(x,y,z).$$

Es obvio que $H_x^0 = 0$, y además podemos aplicar (3.6), y tenemos que,

$$H_x^0 = H_y^0 = H_z^0 = 0. \tag{3.7}$$

Por tanto, el siguiente sistema,

$$\begin{aligned} F(x,y,z) &= 0 \\ H(x,y,z) &= 0 \end{aligned} \tag{3.8}$$

es equivalente al original (3.2).

Como hemos supuesto que $F_x^0 \neq 0$, podemos aplicar el teorema de la función implícita para resolver la primera ecuación $F(x,y,z) = 0$ de (3.8), siendo x una función respecto de y y z cerca de $y = z = 0$. Por tanto, si expresamos $x = f(y,z)$, entonces $f(0,0) = 0$ y las derivadas parciales de $f(y,z)$ pueden ser expresadas respecto las de $F(x,y,z)$. Luego para un entorno de $(0,0)$, tenemos que,

$$f_y = -\frac{F_y(f(y,z),y,z)}{F_x(f(y,z),y,z)}, f_z = -\frac{F_z(f(y,z),y,z)}{F_x(f(y,z),y,z)}. \tag{3.9}$$

Sustituyendo además $x = f(y,z)$ dentro de la ecuación $H(x,y,z)$ en (3.8), ya hemos eliminado la variable x de nuestro sistema. Consideremos ahora la ecuación resultante de sustituir la x,

$$J(x,y) \equiv H(f(y,z),y,z) = 0. \tag{3.10}$$

Si la ecuación anterior se resolviese para y como una función de z, tal que $y = y(z)$, con dy/dz finita en $z = 0$, entonces esta función $y(z)$ junto con $x(z) = f(y(z),z)$, nos daría la solución que buscamos para el sistema $F(x,y,z) = 0$, $H(x,y,z) = 0$. Es más, la derivada dx/dz podría expresarse como,

$$\frac{dx}{dz} = -\frac{F_y \cdot (dy/dz) + F_z}{F_x}. \tag{3.11}$$

Solo nos queda ver cómo obtener nuestra función y respecto de la variable z. Si asumimos que todas las derivadas de F y de H que necesitamos existen y son continuas, entonces las derivadas de $J(y,z)$ pueden ser expresadas en términos de F y H usando (3.10). En concreto, tendríamos que,

$$\begin{aligned} J_y &= \frac{-H_x F_y + H_y F_x}{F_x}, \\ J_z &= \frac{-H_x F_z + H_z F_x}{F_x}. \end{aligned}$$

Entonces por (3.7), tenemos también que $J_y^0 = J_z^0 = 0$.

La segunda y tercera derivada de J son expresiones más complicadas. Por ejemplo, tenemos,

$$\begin{aligned} J_{yy} &= \frac{H_{xx}F_y^2 - 2H_{xy}F_yF_x + H_{yy}F_x^2}{F_x^2} \\ &= -H_x \cdot \frac{F_y^2 F_{xx} - 2F_xF_yF_{xy} + F_x^2 F_{yy}}{F_x^3}. \end{aligned}$$

Con lo que tenemos que por la última igualdad de la expresión anterior, $J_{yy} = 0$ en $(0,0)$ por ser $H_x^0 = 0$, sin embargo, no tenemos razón para pensar que la igualdad de la primera línea sea 0 en el origen.

Teniendo en cuenta que las primeras derivadas parciales de $J(y,z)$ son nulas en $(0,0)$, podemos escribir lo siguiente,

$$\begin{aligned} J(y,z) = {} & \frac{1}{2}J_{yy}^0 y^2 + J_{yz}^0 yz + \frac{1}{2}J_{zz}^0 z^2 + \frac{1}{6}J_{yyy}^0 y^3 - \frac{1}{2}J_{yyz}^0 y^2 z \\ & + \frac{1}{2}J_{yzz}^0 yz^2 + \frac{1}{6}J_{zzz}^0 z^3 + \text{otros términos de orden superior.} \end{aligned} \tag{3.12}$$

Como queremos encontrar una solución a la ecuación $J(y,z) = 0$ con $y = 0$ y $z = 0$ y para la que z tome valores no nulos. Podemos sustituir $y = \mu z$ en la expresión (3.12), siendo μ una función real evaluada en z. Con lo que tenemos que,

$$\begin{aligned} J(y,z) = {} & z^2 \left(\frac{1}{2}J_{yy}^0 \mu^2 + J_{yz}^0 \mu + \frac{1}{2}J_{zz}^0 \right) \\ & + z^3 \left(\frac{1}{6}J_{yyy}^0 \mu^3 + \frac{1}{2}J_{yyz}^0 \mu^2 + \frac{1}{2}J_{yzz}^0 \mu + \frac{1}{6}J_{zzz}^0 \right) \\ & + \text{otros términos de orden superior,} \end{aligned} \tag{3.13}$$

luego para $z \neq 0$, podemos simplificar y tenemos que $J(y,z) = 0$ es equivalente a escribir,

$$\begin{aligned} J(\mu, z) = &\left(\frac{1}{2}J^0_{yy}\mu^2 + J^0_{yz}\mu + \frac{1}{2}J^0_{zz}\right) \\ &+ z\left(\frac{1}{6}J^0_{yyy}\mu^3 + \frac{1}{2}J^0_{yyz}\mu^2 + \frac{1}{2}J^0_{yzz}\mu + \frac{1}{6}J^0_{zzz}\right) \\ &+ \text{otros términos de orden superior} = 0. \end{aligned} \tag{3.14}$$

Ahora podemos resolver la ecuación (3.14) para μ como función de z. Por tanto, la solución de $J(y,z) = 0$ para la variable y vendrá dada como $y(z) = z\mu(z)$. Si cogemos la primera expresión de la ecuación,

$$\frac{1}{2}J^0_{yy}\mu^2 + J^0_{yz}\mu + \frac{1}{2}J^0_{zz} = 0, \tag{3.15}$$

vemos que es el límite de (3.14) cuando $z \to 0$. Si $\mu = \mu(z)$ es solución de (3.14), entonces $\mu(0)$ debe cumplir (3.15). Tenemos por tanto cuatro casos dependiendo del signo del discriminante $(J^0_{yz})^2 - J^0_{yy}J^0_{zz}$ y del signo de las derivadas parciales segundas J^0_{yz}, J^0_{yy} y J^0_{zz}. Lo habitual es que el se pueda aplicar el *Caso I* o el *Caso II* que se van a describir a continuación.

Caso I

$$(J^0_{yz})^2 - J^0_{yy}J^0_{zz} < 0$$
$$\text{o } J^0_{yy} = J^0_{yz} = 0 \text{ y } J^0_{zz} \neq 0$$

En este caso, la ecuación (3.15) no tiene raíces reales. Luego no puede existir un valor real para $\mu(0)$, y por tanto tampoco va a existir una solución del tipo que estamos buscando.

Caso II

$$(J^0_{yz})^2 - J^0_{yy}J^0_{zz} > 0$$

Para este caso, (3.15) tiene una o dos raíces reales simples, dependiendo de si J^0_{yy} se anula o no. Igualmente, para ambos casos se puede aplicar el teorema de la función implícita, el cuál nos asegura la existencia de una solución $\mu = \mu(z)$ de la ecuación (3.14) para cada raíz real simple de (3.15). Luego, cuando μ_0 sea una raíz real, tendremos que,

$$\begin{aligned}
\mu(z) &= \mu_0 + o(1), \\
y(z) &= z\mu(z) = \mu_0(z) + o(z) \\
\frac{dy}{dz} &= \mu_0 \text{ para } z = 0.
\end{aligned}$$

A partir de ahora, cuando se use la notación $o(z)$, nos referiremos a que dicho infinitésimo está acotado. De (3.11) obtenemos además que para $z = 0$ tenemos,

$$\frac{dx}{dz} = -\frac{F_y\mu_0 + F_z}{F_x}. \tag{3.16}$$

Caso III

$$(J_{yz}^0)^2 - J_{yy}^0 J_{zz}^0 = 0 \ y \ J_{yy}^0 \neq 0$$

En esta ocasión, se puede resolver la $\mu(z)$ tanto para z positivo como en negativo. Sin embargo, hará falta utilizar potencias fraccionarias de z. Sea μ_0 una raíz doble de (3.15). Tenemos que entonces la ecuación (3.14) pasa a convertirse en,

$$\begin{aligned}
&\frac{1}{2}J_{yy}^0(\mu-\mu_0)^2 + z\left(\frac{1}{6}J_{yyy}^0(\mu-\mu_0)^3 + J_1(\mu-\mu_0)^2 + J_2(\mu-\mu_0) + J_3\right) \\
&\qquad + \text{ otros términos de orden superior} = 0.
\end{aligned}$$

Siendo

$$\begin{aligned}
J_1 &= \frac{1}{2}J_{yyy}^0\mu_0 + \frac{1}{2}J_{yyz}^0, \\
J_2 &= \frac{1}{2}J_{yyy}^0\mu_0^2 + J_{yyz}^0\mu_0 + \frac{1}{2}J_{yzz}^0, \\
J_3 &= \frac{1}{6}J_{yyy}^0\mu_0^3 + \frac{1}{2}J_{yyz}^0\mu_0^2 + \frac{1}{2}J_{yzz}^0\mu_0 + \frac{1}{6}J_{zzz}^0.
\end{aligned}$$

Caso III (a)

$$J_3 = 0$$

Este subcaso no puede resolverse sin usar derivadas de un orden todavía más superior, luego no lo vamos a ver en detalle.

Caso III (b)

$$J_3 \neq 0$$

Reemplazamos la z por u^2 si J_3 y J^0_{yy} tienen signos opuestos, y por $-u^2$ si son del mismo signo. Entonces, haciendo la raíz cuadrada, obtenemos que,

$$\mu - \mu_0 = \pm\sqrt{(-2J_3/J^0_{yy})u} + \text{otros términos de orden superior,}$$
$$\mu - \mu_0 = \pm\sqrt{(2J_3/J^0_{yy})u} + \text{otros términos de orden superior,}$$

Como resultado, tenemos que estas soluciones:

1. Si J_3 y J^0_{yy} tienen signos opuestos, entonces existen dos soluciones de $\mu(z)$ para z positiva y ninguna para z negativa.

2. Si J_3 y J^0_{yy} tienen el mismo signo, entonces existen dos soluciones de $\mu(z)$ para z negativa y ninguna para z positiva.

Para ambos casos, tenemos que $y = \mu_0(z) + O(|z|^{3/2})$, luego para $z = 0$, $dy/dz = \mu_0$ y dx/dz viene dado por (3.16).

Caso IV

$$J^0_{yy} = J^0_{yz} = J^0_{zz} = 0$$

Para este apartado, podemos remplazar (3.13) por,

$$\frac{1}{6}J^0_{yyy}\mu^3 + \frac{1}{2}J^0_{yyz}\mu^2 + \frac{1}{2}J^0_{yzz}\mu + \frac{1}{6}J^0_{zzz} + \text{otros términos de orden superior} = 0. \quad (3.17)$$

Y tomando ahora la ecuación,

$$\frac{1}{6}J^0_{yyy}\mu^3 + \frac{1}{2}J^0_{yyz}\mu^2 + \frac{1}{2}J^0_{yzz}\mu + \frac{1}{6}J^0_{zzz} = 0, \quad (3.18)$$

que es el límite de (3.17) cuando $z \to 0$. Si (3.18) no tuviese raíces reales, entonces no existirían soluciones del tipo que buscamos. Si μ_0 fuese una raíz real múltiple de (3.18), nos harían falta derivadas de un orden superior para poder encontrar una solución, y no nos vamos a meter en eso. Por último, en el caso de que μ_0 fuese una raíz real simple, entonces el teorema de la función implícita nos dice que (3.18) tiene una solución, $\mu = \mu(z)$ con $\mu(z) = \mu_0 + o(1)$. Luego, como antes, $y(z) = \mu_0(z) + o(z)$, y en $z = 0$, tendremos que,

$$\frac{dy}{dz} = \mu_0, \frac{dx}{dz} = -\frac{F_y\mu_0 + F_z}{F_x}.$$

3.2.2. Matriz Jacobiana de rango 0

Explicar este caso es algo más complicado que los mencionados anteriormente. Sin embargo, a continuación voy a hacer un ejemplo para poder tener una idea general.

Sea el siguiente sistema de ecuaciones,

$$\begin{aligned} xz + yz + xy - z^3 + z^2x &= 0, \\ xz + 2yz - xy - 3z^3 - z^2y &= 0. \end{aligned} \tag{3.19}$$

Como no hay ningún término lineal el rango del jacobiano del sistema es cero en el origen. Sin embargo, nuestra intención es resolverlo para las variables x e y en función de z. Esto se podría hacer de manera algebraica eliminando variables, pero vamos a emplear el teorema de la función implícita para poder mostrar su utilidad en un caso más general.

Teniendo en cuenta que no hay ningún término z^2 en (3.19), si sustituimos por las variables $x = uz$ $y = vz$, tendremos el siguiente sistema una vez que eliminamos los factores comunes de z^2,

$$\begin{aligned} u + v - z + zu + uv &= 0, \\ u + 2v - 3z - zv - uv &= 0. \end{aligned} \tag{3.20}$$

y tomando las siguientes ecuaciones,

$$\begin{aligned} F(u, v, z) &= u + v - z + zu, \\ G(u, v, z) &= u + 2v - 3z - zv, \end{aligned}$$

tenemos que cumplen,

$$\begin{pmatrix} F_u^0 & F_v^0 \\ G_u^0 & G_v^0 \end{pmatrix} . = \begin{pmatrix} 1 & 1 \\ 1 & 2 \end{pmatrix} . \tag{3.21}$$

Además, tenemos que $(u, v, z) = (0, 0, 0)$ es un punto que cumple el sistema (3.20). Luego como la matriz anterior tiene rango 2, por el teorema de la función implícita tenemos garantizada la existencia de las soluciones $u(z)$ y $v(z)$ para (3.20) con $u(0) = v(0) = 0$. Además, el teorema también nos permite obtener las derivadas de $u(z)$ y $v(z)$ de forma implícita. Tenemos por tanto que, $du/dz(0) = -1$ y que $dv/dz(0) = 2$. Y por tanto que,

$$x(z) = -z^2 + o(z^2),$$
$$y(z) = 2z^2 + o(z^2).$$

Por último, de forma adicional, también podemos ver que todos los puntos del conjunto,

$$\{(x, 0, 0) : x \in \mathbb{R}\} \cup \{(0, y, 0) : y \in \mathbb{R}\}$$

también resuelven el sistema (3.19), dándonos por tanto otras dos soluciones adicionales que nos resuelven el sistema en el origen.

3.3. El T.F.I para funciones analíticas

En esta sección vamos a considerar el teorema de la función implícita tanto en el análisis real como en el complejo. Estas categorías están relacionadas ya que podemos pasar el problema de la función implícita del análisis real al complejo, y al revés también.

En otras palabras, un problema de funciones implícitas con datos analíticos reales tiene una solución en $\mathcal{C}^\infty$ gracias al teorema de la función implícita clásico en $\mathcal{C}^\infty$, la cuestión es ver que tiene una solución analítica real. De la misma forma, el mismo problema pero con funciones holomorfas (análisis complejo) también va a tener una solución en $\mathcal{C}^\infty$, y además, también tendrá automáticamente una solución analítica real. Pero el caso ahora es ver que tiene una solución analítica compleja.

Notación: Durante este apartado vamos a usar la notación de *multiíndice*. En

$\mathbb{R}^N$, un multiíndice es una N-tupla $\alpha = (\alpha_1, \ldots, \alpha_N)$, donde cada índice α_j es un entero no negativo. Para un $x \in \mathbb{R}^N$, definimos,

$$x^\alpha \equiv x_1^{\alpha_1} \cdot x_2^{\alpha_2} \cdots x_N^{\alpha_N}.$$

Además, tendremos que $|\alpha| = \alpha_1 + \cdots + \alpha_N$, y que $\alpha! = \alpha_1! \cdot \alpha_2! \cdots \alpha_N!$. Por último, si tenemos otro multiíndice $\beta = (\beta_1, \ldots, \beta_N)$, entonces la suma será, $\alpha + \beta = (\alpha_1 + \beta_1, \ldots, \alpha_N + \beta_N)$. Se va a emplear también la notación e_j para referirse al multiíndice $(0, 0, \ldots, 0, 1, \ldots, 0)$, todo ceros menos el índice de la posición j que es un 1.

También se van a usar series de potencias para una función $\delta(x, y)$, siendo $x \in \mathbb{R}^N$ e $y \in \mathbb{R}$. Luego tendremos potencias de x^α, siendo α un multiíndice y también, y^k, con k un entero no negativo. La escribiremos como,

$$\delta(x, y) = \sum_{\alpha,k} a_{\alpha,k} x^\alpha y^k.$$

donde α se moverá dentro del rango de sus índices y k irá de 0 hasta ∞. Cuando se escriba $a_{0,0}$ se supondrá que α es la N-tupla de solo ceros y $K = 0$.

Enunciamos a continuación el siguiente lema que usa las estimaciones de Hadamard para el tamaño de los coeficientes de una serie de potencias convergente.

Lema: *Sea*

$$F(x) = \sum_{\alpha,k} a_\alpha x^\alpha y^k$$

una función definida por una serie de potencias en $x = (x_1, \ldots, x_N)$ e $y \in \mathbb{R}$ que es convergente para $|x| < R_1$, $|y| < R_2$. Asumiendo que F está acotada por K, entonces,

$$|a_{\alpha,k}| \leq K \cdot R_1^{-|\alpha|} \cdot R_2^{-k}.$$

Teorema: *Supongamos que la serie de potencias,*

$$F(x, y) = \sum_{\alpha,k} a_{\alpha,k} x^\alpha y^k. \tag{3.22}$$

es absolutamente convergente para $|x| \leq R_1$, $|y| \leq R_2$, si

$$a_{0,0} = 0 \text{ y } a_{0,1} \neq 0, \tag{3.23}$$

entonces existe un $r_0 > 0$ *y una serie de potencias,*

$$f(x) = \sum_{|\alpha|>0} c_\alpha x^\alpha \tag{3.24}$$

tal que (3.24) es absolutamente convergente para $|x| \leq r_0$ *y*

$$F(x, f(x)) = 0. \tag{3.25}$$

Demostración:

Tenemos que sin pérdida de generalidad podemos suponer que $a_{0,1} = 1$, luego (3.22) se expresa como,

$$F(x,y) = y + \sum_{|\alpha|>0} (a_{\alpha,0} + a_{\alpha,1} y) x^\alpha + \sum_{|\alpha|\geq 0} \sum_{k=2}^{\infty} a_{\alpha,k} x^\alpha y^k. \tag{3.26}$$

Usando ahora la notación $b_{\alpha,k} = -a_{\alpha,k}$, podemos reescribir la ecuación $F(x,y) = 0$ como,

$$y = \sum_{|\alpha|>0} (b_{\alpha,0} + b_{\alpha,1} y) x^\alpha + \sum_{|\alpha|\geq 0} \sum_{k=2}^{\infty} b_{\alpha,k} x^\alpha y^k, \tag{3.27}$$

y tomando también $y = B(x,y)$, entonces,

$$B(x,y) = \sum_{|\alpha|>0} (b_{\alpha,0} + b_{\alpha,1} y) x^\alpha + \sum_{|\alpha|\geq 0} \sum_{k=2}^{\infty} b_{\alpha,k} x^\alpha y^k. \tag{3.28}$$

Sustituyendo ahora $y = f(x)$ en (3.27) con la $f(x)$ dada en (3.24), obtenemos,

$$\begin{aligned} \sum_{|\alpha|>0} c_\alpha x^\alpha = \sum_{|\alpha|>0} b_{\alpha,0} x^\alpha + \sum_{|\alpha|>0} \sum_{|\beta|>0} b_{\alpha,1} c_\beta x^{\alpha+\beta} \\ + \sum_{|\alpha|\geq 0} \sum_{k=2}^{\infty} b_{\alpha,k} x^\alpha \left(\sum_{|\beta|>0} c_\beta x^\beta \right)^k . \end{aligned} \tag{3.29}$$

Si todas las series que aparecen en (3.29) fueran al final absolutamente convergentes, entonces el orden de los sumatorios podría ser reordenado sin problemas. Asumiendo por tanto que lo son, podemos igualar las potencias de x de la parte izquierda y derecha de la ecuación, obteniendo así las siguientes relaciones de recurrencia. En primer lugar, tenemos

$$c_{e_j} = b_{e_j,0}, \tag{3.30}$$

que nos permite resolverla para cada c_{e_j}. A continuación indicamos como para cada coeficiente c_α con un índice de orden mayor puede ser expresado en términos de $b_{\gamma,j}$ y otros c_β de índice menor. De hecho, vamos a asumir de forma inductiva que hemos resuelto para c_α para todos los multiíndices α con $|\alpha| \leq p$. Fijando entonces un multiíndice $|\alpha| = p$ e identificando las potencias de x, tenemos que,

$$\begin{aligned} c_{\alpha+e_j} &= b_{\alpha+e_j,0} + \sum_{\substack{|\beta|>0,|\gamma|>0,\\ \beta+\gamma=\alpha+e_j}} b_{\gamma,1} c_\beta \\ &+ \sum_{\substack{|\gamma|>0,k\geq 0,\\ |\beta^i|>0,\beta^1+\cdots+\beta^k+\gamma=\alpha+e_j}} M(\beta^1,\ldots,\beta^k)\cdot b_{\gamma,k}\cdot c_{\beta^1}\cdot c_{\beta^2}\cdots c_{\beta^k}. \end{aligned} \tag{3.31}$$

Tenemos que $M(\beta^1,\ldots,\beta^k)$ es un posible coeficiente multinomial (y los superíndices son etiquetas no potencias), y cada M es positivo. Se puede comprobar que todos los multiíndices del lado derecho son menores o iguales que p, justo como queríamos.

Las relaciones de recurrencia obtenidas por ahora, (3.30) y (3.31) únicamente determinan los coeficientes c_β en la serie de potencias para la función implícita, pero aún nos hace falta ver que (3.24) es convergente. La manera más sencilla para obtener las estimaciones necesarias es usando el método de mayorantes, que describo a continuación.

Sean dos series de potencias con el mismo número de variables,

$$\phi(x_1,x_2,\ldots,x_\rho) = \sum_{j_1,j_2,\ldots,j_\rho=0}^{\infty} \phi_{j_1,j_2,\ldots,j_\rho} x_1^{j^1} x_2^{j^2}\ldots x_\rho^{j^\rho}, \tag{3.32}$$

$$\psi(x_1,x_2,\ldots,x_\rho) = \sum_{j_1,j_2,\ldots,j_\rho=0}^{\infty} \psi_{j_1,j_2,\ldots,j_\rho} x_1^{j^1} x_2^{j^2}\ldots x_\rho^{j^\rho} \tag{3.33}$$

Diremos que $\psi(x_1,x_2,\ldots,x_\rho)$ es un mayorante de $\phi(x_1,x_2,\ldots,x_\rho)$ si,

$$|\phi_{j_1,j_2,\ldots,j_\rho}| \leq \psi_{j_1,j_2,\ldots,j_\rho} \tag{3.34}$$

para todo $j_1,j_2,\ldots,j_\rho$.

Como todos los coeficientes de $M(\beta^1,\ldots,\beta^k)$ en (3.31) son positivos, tenemos que si

$$G(x,y) = \sum_{|\alpha|\geq 0,k\geq 0} g_{\alpha,k} x^\alpha y^k,$$

(con $g_{0,0} = g_{0,1} = 0$) es un mayorante de

$$B(x,y) = \sum_{|\alpha|>0} (b_{\alpha,0} + b_{\alpha,1} y) x^\alpha + \sum_{|\alpha|\geq 0, k\geq 2} b_{\alpha,k} x^\alpha y^k$$

y si

$$h(x) = \sum_{|\alpha|} h_\alpha x_\alpha \tag{3.35}$$

resuelve

$$h(x) = G[x, h(x)], \tag{3.36}$$

entonces $h(x)$ será un mayorante de f(x). Como consecuencia, si la serie (3.34) para $h(x)$ es convergente, entonces la serie de (3.24) es convergente y su radio de convergencia es al menos tan grande como el radio de convergencia de (3.34). Tenemos

$$G(x,y) = \frac{Kr}{r - (x_1 + \cdots + x_N) - y},$$

donde $K = sup\{|B(x,y)| : x \in \bar{D}(o, R_1), y \in \bar{D}(0, r_2)\}$ y r son lo suficientemente pequeños (dependiendo de R_1 y R_2). Por el lema enunciando previamente, tenemos que G es un mayorante de B. Con esta elección de mayorante (3.35) es cuadrática y puede ser resuelta explícitamente. La solución además es holomorfa en $x = 0$. De hecho, $y = h(x)$ es la solución de dicha ecuación cuadrática

$$y^2 + (x_1 + \cdots + x_N - r)y + Kr = 0. \tag{3.37}$$

□

El teorema que acabamos de demostrar solo considera el caso de una variable dependiente y un número arbitrario de independientes, mediante el método inductivo de Dini se podría obtener una versión general.

Tenemos además, que el teorema de la función implícita para el análisis complejo se saca de la versión de $\mathcal{C}^1$. Una vez que sabemos que la solución $y = \psi(x)$ para un sistema holomorfo,

$$F(x,y) = 0$$

es continuamente diferenciable, basta aplicar $\partial/\partial\bar{z}_j$ a ambos lados y aplicar la regla de la cadena para determinar que ψ es holomorfa.

Bibliografía

[1] Steven G. Krantz and Harold R. Parks, The Implicit Function Theorem. History, theory and applications, Springer Science and Business Media, 2012.

[2] J.J. O'Connor and E.F. Robertson, The Function Concept, MacTutor History of Mathematics Archive, 2005. https://mathshistory.st-andrews.ac.uk/HistTopics/Functions/.

[3] Iurato, Giuseppe (2012) On the role played by the work of Ulisse Dini on implicit function theory in the modern differential geometry foundations: the case of the structure of a differentiable manifold, 1. http://philsci-archive.pitt.edu/9485/

[4] Mingari Scarpello, Giovanni and Ritelli, Daniele. (2002). A Historical Outline of the Theorem of Implicit Functions. Divulgaciones Matemáticas. 10.

[5] Bruno Belhoste, Augustin-Louis Cauchy: A Biography, Springer-Verlag New York, 1991.

[6] Frank Smithies, Cauchy and the Creation of Complex Function Theory, University of Cambridge, 1997.

[7] José Carlos Castaño Muñoz, Trabajo de Fin de Grado, El teorema de la función implícita, Universidad de Sevilla, curso académico 2017-2018. https://idus.us.es/handle/11441/77506