THE IMPLICIT FUNCTION THEOREM WHEN THE MATRIX $\frac{\partial F}{\partial y}(x,y)$ IS ONLY CONTINUOUS AT THE BASE POINT

Oswaldo Rio Branco de Oliveira

Abstract

This article presents a very elementary proof of the Implicit Function Theorem for differentiable maps $F(x,y)$, defined on a finite-dimensional Euclidean space, with $\frac{\partial F}{\partial y}(x,y)$ only continuous at the base point. In the case of a single scalar equation, this continuity hypothesis is not required. The inverse Function Theorem is also shown. The proofs are builded upon the mean-value theorem, the intermediate-value theorem, and Darboux's property (the intermediate-value property for derivatives). These proofs avoid compactness arguments and fixed-point theorems.

Mathematics Subject Classification: 26B10, 90C30
Key words and phrases: Implicit Function Theorems, Nonlinear Programming.

1 Introduction.

The objective of this article is to present a very elementary proof of a generally easy to apply Implicit Function Theorem. We prove this theorem for differentiable maps $F(x,y)$ defined on a finite-dimensional Euclidean space with the matrix $\frac{\partial F}{\partial y}(x,y)$ only continuous at the base point. In the case of a single scalar equation, we show that this continuity hypothesis is unnecessary. The Inverse Function Theorem is also shown. Besides following Dini's approach (see [3]), these proofs do not employ compactness arguments, the contraction principle, or any fixed-point theorem. Instead of such tools, the proofs in this article use the intermediate-value theorem, the mean-value theorem on the real line, and the intermediate-value property for derivatives on $\mathbb{R}$ (Darboux's property).

Throughout what follows, we shall freely assume that all the functions are defined on a subset of a finite-dimensional Euclidean space.

Some comments are worthwhile concerning proofs of the implicit and inverse function theorems. Most proofs of the classical versions (enunciated for maps of class C^1 on an open set) start with a demonstration of the Inverse Function Theorem and then prove the Implicit Function Theorem as a consequence of the former. Yet, in general these proofs employ either a compactness argument or the contraction mapping principle, see Krantz and Parks [9, pp. 41–52] and Dontchev and Rockafellar [4, pp. 9–20]. On the other hand, a proof of the classical Implicit Function Theorem that does not use either a compactness argument or any fixed-point theorem can be seen in de Oliveira [2].

Taking into account everywhere differentiable maps, a proof of the Implicit Function Theorem can be found in Hurwicz and Richter [5], whereas a proof of the Inverse Function Theorem can be seen in Saint Raymond [10]. The first proof employs Brower's fixed-point theorem while the second relies on a compactness argument. Instead of assuming the continuity of the first order partial derivatives, these proofs assume an appropriate nondegeneracy condition at all points inside some open set containing the base point. It is worth noting that this quite general condition can be difficult to verify.

Considering maps that are differentiable at the base point, but not necessarily on a neighborhood of it, one can find proofs of the implicit and inverse function theorems in Hurwicz and Richter [5] and Nijenhuis [8]. This second work employs Banach's fixed-point theorem.

Removing altogether the differentiability hypothesis, a proof of the Inverse Function Theorem for a map satisfying a Lipschitz condition can be seen in Clarke [1]. Yet, proofs of the Implicit Function Theorem for continuous maps can be found in Jittorntrum [6] and Kumagai [7].

In this article, the overall stategy of the proof of the Implicit Function Theorem is as follows. First, we prove it for a differentiable real function. Then, given a finite number of equations, we prove it supposing that the matrix $\frac{\partial F}{\partial y}(x,y)$ is continuous at the base point. In addition, we prove the Inverse Function Theorem for a map whose Jacobian matrix is continuous at the base point.

2 Notations and Preliminaries.

Apart from the intermediate-value and the mean-value theorems, both on the real line, we assume the intermediate-value theorem for derivatives on $\mathbb{R}$ (Darboux's property): *Given a differentiable function $f : [a,b] \to \mathbb{R}$, the image of the derivative function is an interval.*

Let us consider n and m, both in $\mathbb{N}$, and fix the canonical bases $\{e_1, \ldots, e_n\}$ and $\{f_1, \ldots, f_m\}$, of $\mathbb{R}^n$ and $\mathbb{R}^m$, respectively. Given $x = (x_1, \ldots, x_n)$ and $y = (y_1, \ldots, y_n)$, both in $\mathbb{R}^n$, we put $\langle x, y\rangle = x_1y_1 + \cdots + x_ny_n$ and $|x| = \sqrt{\langle x, x\rangle}$. Given $r > 0$, let us write $B(x;r) = \{y \text{ in } \mathbb{R}^n : \ |y-x| < r\}$. We identify a linear map $T : \mathbb{R}^n \to \mathbb{R}^m$ with the $m \times n$ matrix $M = (a_{ij})$, where $T(e_j) = a_{1j}f_1 + \cdots + a_{mj}f_m$ for $j = 1, \ldots, n$. We also write Tv for $T(v)$.

In this section, Ω denotes a nonempty open subset of $\mathbb{R}^n$, where $n \geq 1$. Given a map $F : \Omega \to \mathbb{R}^m$ and a point p in Ω, we write $F(p) = \big(F_1(p), \ldots, F_m(p)\big)$.

Let us suppose that F is differentiable at p. The Jacobian matrix of F at p is

$$JF(p) = \left(\frac{\partial F_i}{\partial x_j}(p)\right)_{\substack{1\leq i\leq m\\ 1\leq j\leq n}} = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(p) & \cdots & \frac{\partial F_1}{\partial x_n}(p) \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial x_1}(p) & \cdots & \frac{\partial F_m}{\partial x_n}(p) \end{pmatrix}.$$

If F is a real function, then we have $JF(p) = \nabla F(p)$, the gradient of F at p.

The following lemma (a particular case of the chain rule but sufficient for our purposes) is a local result. For practical reasons we state it for $\Omega = \mathbb{R}^n$. We omit the proof of the lemma.

Lemma 1 *Let $F : \mathbb{R}^n \to \mathbb{R}^m$ be differentiable, $T : \mathbb{R}^k \to \mathbb{R}^n$ be the linear function associated to a $n \times k$ real matrix M, and y be a fixed point in $\mathbb{R}^n$. Then, the function $G(x) = F(y + Tx)$, where x is in $\mathbb{R}^k$, is differentiable and satisfies $JG(x) = JF(y + Tx)M$, for all x in $\mathbb{R}^k$.*

Given a and b, both in $\mathbb{R}^n$, we put $\overline{ab} = \{a + t(b - a) : 0 \leq t \leq 1\}$. The following mean-value theorem (in several variables) is a trivial consequence of the mean-value theorem on the real line and thus we omit the proof.

Lemma 2 *Let us consider a differentiable real function $F : \Omega \to \mathbb{R}$, with Ω open in $\mathbb{R}^n$. Let a and b be points in Ω such that the segment $\overline{ab}$ is within Ω. Then, there exists c in $\overline{ab}$ satisfying*

$$F(b) - F(a) = \langle \nabla F(c), b - a \rangle \,.$$

We denote the determinant of a real square matrix M by $\det M$.

Lemma 3 *Let us consider a differentiable map $F : \Omega \to \mathbb{R}^n$, with Ω open within $\mathbb{R}^n$, and p a point in Ω satisfying $\det JF(p) \neq 0$. Let us suppose that the real function $\det\left(\frac{\partial F_i}{\partial x_j}(\xi_{ij})\right)$ in the n^2 variables ξ_{ij}, with $1 \leq i, j \leq n$ and ξ_{ij} running in Ω, is continuous at the point defined by $\xi_{ij} = p$, for all $1 \leq i, j \leq n$. Then, the restriction of F to some non-degenerate open ball $B(p; r)$ is injective.*

Proof. Since $\det\left(\frac{\partial F_i}{\partial x_j}(p)\right) \neq 0$, the continuity hypothesis yields a $r > 0$ such that $\det\left(\frac{\partial F_i}{\partial x_j}(\xi_{ij})\right)$ does not vanish, for all $\xi_{ij} \in B(p; r)$ and $1 \leq i, j \leq n$. Now, let a and b be distinct in $B(p; r)$. By employing the mean-value theorem in several variables to each component F_i of F, we find c_i in the segment $\overline{ab}$, within $B(p; r)$, such that $F_i(b) - F_i(a) = \langle \nabla F_i(c_i), b - a \rangle$. Hence,

$$\begin{pmatrix} F_1(b) - F_1(a) \\ \vdots \\ F_n(b) - F_n(a) \end{pmatrix} = \begin{pmatrix} \frac{\partial F_1}{\partial x_1}(c_1) & \cdots & \frac{\partial F_1}{\partial x_n}(c_1) \\ \vdots & & \vdots \\ \frac{\partial F_n}{\partial x_1}(c_n) & \cdots & \frac{\partial F_n}{\partial x_n}(c_n) \end{pmatrix} \begin{pmatrix} b_1 - a_1 \\ \vdots \\ b_n - a_n \end{pmatrix}.$$

Thus, since $\det\left(\frac{\partial F_i}{\partial x_j}(c_i)\right) \neq 0$ and $b - a \neq 0$, we conclude that $F(b) \neq F(a)$. $\square$

Given a real function $F : \Omega \to \mathbb{R}$, a short computation shows that the following definition of differentiability is equivalent to that which is most commonly used. We say that F is differentiable at p in Ω if there are a ball $B(p;r)$ within Ω, with $r > 0$, a v in $\mathbb{R}^n$, and a vector-valued map $E : B(0;r) \to \mathbb{R}^n$ satisfying

$$\begin{cases} F(p+h) = F(p) + \langle v, h\rangle + \langle E(h), h\rangle\,, \text{ for all } |h| < r, \\ \text{where } E(h) = 0 \text{ and } E(h) \to 0 \text{ as } h \to 0. \end{cases}$$

3 The Implicit and Inverse Function Theorems.

The first implicit function result we prove concerns one equation, several variables and a differentiable real function whose partial derivatives need not be continuous at any point. In its proof, we denote the variable in $\mathbb{R}^{n+1} = \mathbb{R}^n \times \mathbb{R}$ by (x, y), where $x = (x_1, \ldots, x_n)$ is in $\mathbb{R}^n$ and y is in $\mathbb{R}$. Given a subset X of $\mathbb{R}^n$ and a subset Y of $\mathbb{R}$, let us use the notation $X \times Y = \{(x, y) : x \in X \text{ and } y \in Y\}$. It is well-known that $X \times Y$ is open in $\mathbb{R}^n \times \mathbb{R}$ if and only if X and Y are open.

In this section, Ω denotes a nonempty open set within $\mathbb{R}^n \times \mathbb{R}$.

Theorem 1 *Let $F : \Omega \to \mathbb{R}$ be differentiable, with $\frac{\partial F}{\partial y}$ nowhere vanishing, and (a, b) a point in Ω such that $F(a, b) = 0$. Then, there exists an open set $X \times Y$, within Ω and containing the point (a, b), that satisfies the following.*

- *For each x in X there is a unique $y = g(x)$ in Y such that $F\big(x, g(x)\big) = 0$.*
- *We have $g(a) = b$. Moreover, $g : X \to Y$ is differentiable and satisfies*

$$\frac{\partial g}{\partial x_j}(x) = -\frac{\frac{\partial F}{\partial x_j}(x, g(x))}{\frac{\partial F}{\partial y}(x, g(x))}, \text{ for all } x \text{ in } X, \text{ where } j = 1, \ldots, n.$$

Moreover, if $\nabla F(x, y)$ is continuous at (a, b) then $\nabla g(x)$ is continuous at $x = a$.

Proof. By considering the function $F(x + a, \frac{y}{c} + b)$, with $c = \frac{\partial F}{\partial y}(a, b)$, we may assume that $(a, b) = (0, 0)$ and $\frac{\partial F}{\partial y}(0, 0) = 1$. Next, we split the proof into three parts: existence and uniqueness, continuity at the origin, and differentiability.

$\diamond$ Existence and Uniqueness. Let us choose a non-degenerate $(n+1)$-dimensional parallelepiped $X \times [-r, r]$, centered at $(0, 0)$ and within Ω, whose edges are parallel to the coordinate axes and X is open. Then, the function $\varphi(y) = F(0, y)$, where y runs over $[-r, r]$, is differentiable with φ' nowhere vanishing and $\varphi'(0) = 1$. Thus, by Darboux's property we have $\varphi' > 0$ everywhere and we conclude that φ is strictly increasing. Hence, by the continuity of F and shrinking X (if necessary) we may assume that

$$F\Big|_{X\times\{-r\}} < 0 \quad \text{and} \quad F\Big|_{X\times\{r\}} > 0.$$

As a consequence, fixing an arbitrary x in X, the function

$$\psi(y) = F(x,y), \text{ where } y \in [-r,r],$$

satisfies $\psi(-r) < 0 < \psi(r)$. Hence, by the mean-value theorem there exists a point η in the open interval $Y = (-r,r)$ such that $\psi'(\eta) = \frac{\partial F}{\partial y}(x,\eta) > 0$. Therefore, by Darboux's property we have $\psi'(y) > 0$ at every y in Y. Thus, ψ is strictly increasing and the intermediate-value theorem yields the existence of a unique $y = g(x)$ in the open interval Y such that $F(x, g(x)) = 0$.

- ◇ Continuity at the origin. Let δ satisfy $0 < \delta < r$. From above, there exists an open set $\mathcal{X}$, contained in X and containing 0, such that $g(x)$ is in the interval $(-\delta, \delta)$, for all x in $\mathcal{X}$. Thus, g is continuous at $x = 0$.

- ◇ Differentiability. From the differentiability of the real function F at $(0,0)$, and writing $\nabla F(0,0) = (v,1) \in \mathbb{R}^n \times \mathbb{R}$ for the gradient of F at $(0,0)$, it follows that there are functions $E_1 : \Omega \to \mathbb{R}^n$ and $E_2 : \Omega \to \mathbb{R}$ satisfying

$$\begin{cases} F(h,k) = \langle v,h\rangle + k + \langle E_1(h,k), h\rangle + E_2(h,k)k, \\ \\ \text{where } \lim\limits_{(h,k)\to(0,0)} E_j(h,k) = 0 = E_j(0,0), \text{ for } j = 1,2. \end{cases}$$

Hence, substituting [we already proved that $g(h) \xrightarrow{h\to 0} g(0) = 0$]

$$\begin{cases} k = g(h), \\ E_j\big(h, g(h)\big) = \epsilon_j(h), \text{ with } \lim\limits_{h\to 0} \epsilon_j(h) = \epsilon_j(0) = 0 \text{ for } j = 1,2, \end{cases}$$

and noticing that we have $F\big(h, g(h)\big) = 0$, for all possible h, we obtain

$$\langle v,h\rangle + g(h) \ + \langle \epsilon_1(h), h\rangle + \epsilon_2(h)g(h) = 0.$$

Thus,

$$[1 + \epsilon_2(h)]g(h) = -\langle v,h\rangle - \langle \epsilon_1(h), h\rangle .$$

If $|h|$ is small enough, then we have $1 + \epsilon_2(h) \neq 0$ and we may write

$$g(h) = \langle -v, h\rangle + \langle \epsilon_3(h), h\rangle ,$$

where

$$\epsilon_3(h) = \frac{\epsilon_2(h)}{1+\epsilon_2(h)}v - \frac{\epsilon_1(h)}{1+\epsilon_2(h)} \text{ and } \lim_{h\to 0} \epsilon_3(h) = 0.$$

Therefore, g is differentiable at 0 and $\nabla g(0) = -v$.

Now, given any a' in X, we put $b' = g(a')$. Then, $g : X \to Y$ solves the problem $F\big(x, h(x)\big) = 0$, for all x in X, with the condition $h(a') = b'$. From what we have just done it follows that g is differentiable at a'. $\square$

Next, we prove the implicit function theorem for a finite number of equations. Some notation is appropriate. We denote the variable in $\mathbb{R}^n \times \mathbb{R}^m = \mathbb{R}^{n+m}$ by $(x; y)$, where $x = (x_1, \ldots, x_n)$ is in $\mathbb{R}^n$ and $y = (y_1, \ldots, y_m)$ in $\mathbb{R}^m$. Given Ω an open subset of $\mathbb{R}^n \times \mathbb{R}^m$ and a differentiable map $F : \Omega \to \mathbb{R}^m$, we write $F = (F_1, \ldots, F_m)$ with F_i the ith component of F and $i = 1, \ldots, m$, and

$$\frac{\partial F}{\partial y} = \left(\frac{\partial F_i}{\partial y_j}\right)_{\substack{1\leq i\leq m\\ 1\leq j\leq m}} = \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \ldots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \vdots \\ \frac{\partial F_m}{\partial y_1} & \ldots & \frac{\partial F_m}{\partial y_m} \end{pmatrix}.$$

Analogously, we define the matrix $\frac{\partial F}{\partial x} = \left(\frac{\partial F_i}{\partial x_k}\right)$, where $1 \leq i \leq m$ and $1 \leq k \leq n$.

Theorem 2 (The Implicit Function Theorem). *Let $F : \Omega \to \mathbb{R}^m$ be differentiable, where Ω is an open set in $\mathbb{R}^n \times \mathbb{R}^m$. Let us suppose that (a, b) is a point in Ω satisfying $F(a, b) = 0$ and $\det \frac{\partial F}{\partial y}(a, b) \neq 0$, with $\frac{\partial F}{\partial y}(x, y)$ continuous at (a, b). Then, there exist an open set $X \times Y$, within Ω and containing (a, b), satisfying the following conditions.*

- *Given x in X, there is a unique $y = g(x)$ in Y such that $F(x, g(x)) = 0$.*
- *We have $g(a) = b$. Moreover, the map $g : X \to Y$ is differentiable and*

$$Jg(x) = -\left[\frac{\partial F}{\partial y}(x, g(x))\right]^{-1}_{m\times m} \left[\frac{\partial F}{\partial x}(x, g(x))\right]_{m\times n}, \text{ for all } x \text{ in } X.$$

In addition, if $JF(x, y)$ is continuous at (a, b) then $Jg(x)$ is continuous at $x = a$.

Proof. Let us consider the invertible matrix $\frac{\partial F}{\partial y}(a, b) = M$ and the associated bijective linear function $\mathcal{M} : \mathbb{R}^m \to \mathbb{R}^m$. By employing Lemma 1 we conclude that the map $G(x; z) = F[x; b + \mathcal{M}^{-1}(z - b)]$, defined on a small enough neighborhood of (a, b), satisfy $\frac{\partial G}{\partial z}(a; b) = MM^{-1}$ and the condition $G(a; b) = 0$. Thus, we may suppose that M is the identity matrix of order m.

Next, we split the proof into four parts: finding Y, existence and differentiability, differentiation formula, and uniqueness.

◇ Finding Y. Defining $\Phi(x, y) = \big(x, F(x, y)\big)$, where (x, y) is in Ω, we have

$$J\Phi(x, y) = \left(\begin{array}{c|c} I & 0 \\ \hline \frac{\partial F}{\partial x} & \frac{\partial F}{\partial y} \end{array}\right) \text{ and } \det J\Phi(x, y) = \det \frac{\partial F}{\partial y}(x, y),$$

with I the identity matrix of order n and 0 the $n \times m$ zero matrix. Thus, $\det J\Phi(a, b) \neq 0$. By hypothesis, the matrix $\frac{\partial F}{\partial y}(x, y)$ is continuous at (a, b). Next, in order to apply Lemma 3 we introduce the variables ξ_{lk} in Ω , where l and k run in $\{1, \ldots, m + n\}$, and the notation $(z_1, \ldots, z_n; z_{n+1}, \ldots, z_{n+m}) = (x_1, \ldots, x_n; y_1, \ldots, y_m)$. Then, the real function $\det\left(\frac{\partial \Phi_l}{\partial z_k}(\xi_{lk})\right) = \det\left(\frac{\partial F_i}{\partial y_j}(\xi_{i+n,j+n})\right)$ is continuous at the point defined by $\xi_{lk} = (a, b)$, for all $l, k = 1, \ldots, m + n$. Therefore, by Lemma 3 and shrinking Ω if necessary, we may assume that Φ is injective. We may also assume that Ω is an open non-degenerate parallelepiped $\mathcal{X}_1 \times Y$ centered at (a, b) whose edges are parallel to the coordinate axes.

◇ Existence and differentiability. We claim that the system

$$\begin{cases} F_1(x; y_1, \ldots, y_m) = 0, \\ F_2(x; y_1, \ldots, y_m) = 0, \\ \qquad\vdots \\ F_m(x; y_1, \ldots, y_m) = 0, \end{cases} \quad \text{with the conditions} \quad \begin{cases} y_1(a) = b_1 \\ y_2(a) = b_2 \\ \quad\vdots \\ y_m(a) = b_m, \end{cases}$$

has a differentiable solution $g(x) = \big(g_1(x), \ldots, g_m(x)\big)$ on some open set X containing a [i.e., we have $F\big(x, g(x)\big) = 0$ for all x in X and $g(a) = b$].

Let us prove it by induction on m. The case $m = 1$ follows from Theorem 1 since $\frac{\partial F}{\partial y}(a; b) = 1$ and, by continuity, we can assume $\frac{\partial F}{\partial y} \neq 0$ everywhere.

Assuming that the claim holds for $m-1$, let us examine the case m. Then, given a pair $(x; y) = (x; y_1, \ldots, y_m)$ we introduce the helpful notations $y' = (y_2, \ldots, y_m)$, $y = (y_1; y')$, and $(x; y) = (x; y_1; y')$.

Next, let us consider the equation $F_1(x; y_1; y') = 0$, where x and y' are independent variables and y_1 is the dependent variable, with the condition $y_1(a; b') = b_1$. Since $\frac{\partial F_1}{\partial y_1}(a; b_1; b') = 1$, by continuity we may assume that the function $\frac{\partial F_1}{\partial y_1}(x; y_1; y')$ does not vanish. Hence, by Theorem 1 there exists a differentiable function $\varphi(x; y')$ on some open set [let us say, $\mathcal{X}_2 \times \mathcal{Y}'$] containing $(a; b')$ that satisfies

$$F_1[x; \varphi(x; y'); y'] = 0 \text{ (on } \mathcal{X}_2 \times \mathcal{Y}') \text{ and the condition } \varphi(a; b') = b_1.$$

As a consequence, $\varphi(x; y')$ also satisfies the $m - 1$ equations

$$\frac{\partial F_1}{\partial y_1}[x; \varphi(x; y'); y'] \frac{\partial \varphi}{\partial y_j}(x; y') + \frac{\partial F_1}{\partial y_j}[x; \varphi(x; y'); y'] = 0, \text{ for } j = 2, \ldots, m.$$

Thus, since $\frac{\partial F_1}{\partial y} = \big(\frac{\partial F_1}{\partial y_1}, \ldots, \frac{\partial F_1}{\partial y_m}\big)$ is continuous at $(a; b_1; b')$, with $\frac{\partial F_1}{\partial y_1}$ nowhere vanishing, and φ is continuous, with $\varphi(a; b') = b_1$, we conclude that $\frac{\partial \varphi}{\partial y'} = \big(\frac{\partial \varphi}{\partial y_2}, \ldots, \frac{\partial \varphi}{\partial y_m}\big)$ is continuous at $(a; b')$.

Now, we look at solving the system with $m - 1$ equations

$$\begin{cases} F_2[x; \varphi(x; y'); y'] = 0 \\ \qquad\vdots \\ F_m[x; \varphi(x; y'); y'] = 0 \end{cases}, \text{ with the condition } \; y'(a) = b'.$$

Let us define $\mathcal{F}_i(x; y') = F_i[x; \varphi(x; y'); y']$, with $i = 2, \ldots, m$, and write $\mathcal{F} = (\mathcal{F}_2, \ldots, \mathcal{F}_m)$. Then, since the entries of the matrices $\frac{\partial \varphi}{\partial y'}(x; y')$ and

$\frac{\partial F}{\partial y}(x;y)$ are continuous at $(a;b')$ and $(a;b)$, respectively, with $\varphi(a;b') = b_1$, we conclude that the entries of $\frac{\partial \mathcal{F}}{\partial y'}(x;y')$ are continuous at $(a;b')$. Yet, by hypothesis $\frac{\partial F}{\partial y}(a;b)$ is the identity matrix of order m and thus we find

$$\frac{\partial \mathcal{F}_i}{\partial y_j}(a;b') = \frac{\partial F_i}{\partial y_1}(a;b)\frac{\partial \varphi}{\partial y_j}(a;b') + \frac{\partial F_i}{\partial y_j}(a;b) = 0 + \frac{\partial F_i}{\partial y_j}(a;b), \text{ for } 2 \leq i,j \leq m.$$

This shows that $\frac{\partial \mathcal{F}}{\partial y'}(a;b')$ is the identity matrix of order $m-1$. Therefore, by induction hypothesis there exists a differentiable function ψ on an open set X containing a [with $\psi(X)$ contained in $\mathcal{Y}'$] that satisfies

$$\begin{cases} F_i[x;\varphi\big(x;\psi(x)\big);\psi(x)] = 0, \text{ for all } x \text{ in } X, \text{ for all } i = 2,\ldots,m, \\ \text{and the condition } \psi(a) = b'. \end{cases}$$

Clearly, we also have $F_1\big[x;\varphi\big(x;\psi(x)\big);\psi(x)\big] = 0$, for all x in X. Defining $g(x) = \big(\varphi(x;\psi(x));\psi(x)\big)$, with x in X, we obtain $F[x;g(x)] = 0$, for all x in X, and $g(a) = \big(\varphi(a;b');b'\big) = (b_1;b') = b$, with g differentiable on X.

◇ Differentiation formula. Differentiating $F[x;g(x)] = 0$ we find

$$\frac{\partial F_i}{\partial x_k} + \sum_{j=1}^{m} \frac{\partial F_i}{\partial y_j}\frac{\partial g_j}{\partial x_k} = 0, \text{ with } 1 \leq i \leq m \text{ and } 1 \leq k \leq n.$$

In matricial form, we write $\frac{\partial F}{\partial x}\big(x,g(x)\big) + \frac{\partial F}{\partial y}\big(x,g(x)\big)Jg(x) = 0$.

◇ Uniqueness. Let X, Y, and g be as described in Theorem 2. Given arbitraries $h : X \to Y$ and x in X satisfying $F\big(x,h(x)\big) = 0$, we have $\Phi(x,h(x)) = (x,0) = \Phi(x,g(x)\big)$. Thus, since Φ is injective, $h(x) = g(x)$. $\square$

Theorem 3 (The Inverse Function Theorem). *Let $F : \Omega \to \mathbb{R}^n$ be differentiable, where Ω is an open set in $\mathbb{R}^n$. Let us suppose that x_0 is a point in Ω such that $JF(x_0)$ is invertible, with $JF(x)$ continuous at x_0. Then, there exist an open set X containing x_0, an open set Y containing $y_0 = F(x_0)$, and a differentiable function $G : Y \to X$ that satisfies $F\big(G(y)\big) = y$, for all y in Y, and $G\big(F(x)\big) = x$, for all x in X. In addition,*

$$JG(y) = JF\big(G(y)\big)^{-1}, \textit{ for all } y \textit{ in } Y,$$

and $JG(y)$ is continuous at $y = y_0$.

Proof. By Lemma 3 we may assume that F is injective. The map $\Phi(y,x) = F(x) - y$, where (y,x) runs over $\mathbb{R}^n \times \Omega$, is differentiable and $\Phi(y_0,x_0) = 0$. Yet, $\frac{\partial \Phi}{\partial x}\big(y_0,x_0\big) = JF(x_0)$ is invertible and $J\Phi(y,x)$ is continuous at (y_0,x_0). The Implicit Function Theorem guarantees an open set Y containing y_0 and a differentiable map $G : Y \to \Omega$, with $JG(y)$ continuous at $y = y_0$, satisfying

$$F\big(G(y)\big) = y, \text{ for all } y \text{ in } Y.$$

Thus, G is bijective from Y to $X = G(Y)$ and F is bijective from X to Y. We also have $X = F^{-1}(Y)$. Since F is continuous, X is open (and contains x_0).

Putting $F(x) = \big(F_1(x), \ldots, F_n(x)\big)$ and $G(y) = \big(G_1(y), \ldots, G_n(y)\big)$ and differentiating $\big(F_1(G(y)), \ldots, F_n(G(y))\big)$ we find

$$\sum_{k=1}^{n} \frac{\partial F_i}{\partial x_k} \frac{\partial G_k}{\partial y_j} = \frac{\partial y_i}{\partial y_j} = \begin{cases} 1, & \text{if } i = j, \\ 0, & \text{if } i \neq j. \end{cases}$$

$\square$

Acknowledgments. The author is greatly indebted to Professors Robert B. Burckel and James V. Ralston for their very valuable comments and suggestions.

References

[1] F. H. Clarke, On the inverse function theorem, *Pacific. J. Math.*, **64**(1) (1976) 97–102.

[2] O. R. B. de Oliveira, The implicit and the inverse function theorems: easy proofs, *Real Anal. Exchange*, to appear. Available at arXiv preprint arXiv:1212.2066, 2012.

[3] U. Dini, *Lezione di Analisi Infinitesimale*, volume 1, Pisa, 1907, 197–241.

[4] A. L. Dontchev and R. T. Rockafellar, *Implicit Functions and Solution Mappings*, Springer, New York, 2009.

[5] L. Hurwicz and M. K. Richter, Implicit functions and diffeomorphisms without C^1, *Adv. Math. Econ.*, **5** (2006) 65–96.

[6] K. Jittorntrum, An implicit function theorem, *J. Optim. Theory Appl.* **25**(4) (1978) 575–577.

[7] S. Kumagai, An implicit function theorem: comment, *J. Optim. Theory Appl.*, **31**(2) (1980) 285–288.

[8] A. Nijenhuis, Strong derivatives and inverse mappings, *Amer. Math. Monthly*, **81** (9) (1974) 969–980.

[9] S. G. Krantz and H. R. Parks, *The Implicit Function Theorem - History, Theory, and Applications*, Birkhaüser, Boston, 2002.

[10] J. Saint Raymond, Local inversion for differentiable functions and the Darboux property, *Mathematika*, **49** (2002), 141–158.

Departamento de Matemática, Universidade de São Paulo
Rua do Matão 1010 - CEP 05508-090
São Paulo, SP - Brasil
oliveira@ime.usp.br