1. Introduction

We plan to introduce the calculus on $\mathbb{R}^n$, namely the concept of total derivatives of multivalued functions $f : \mathbb{R}^n \to \mathbb{R}^m$ in more than one variable. We are indeed familiar with the notion of partial derivatives $\partial_i f_j = \frac{\partial f_j}{\partial x_i}, 1 \leq i \leq n, 1 \leq j \leq m$. In the sequel, we see the importance of introducing the powerful concept of Total Derivative and its connection to the partial derivatives. The reader is expected to have the sound knowledge of basic linear algebra and the notions like: $\mathbb{R}^n$ as normed linear space, basis, dimension, linear operators and transformations and so on. We remark that the total derivative (known also as Frechét Derivative) can be extended to infinite dimensional normed linear spaces which allows one to solve more complicated problems especially arising from optimal control problems, calculus of variations, partial differential equations and so on.

Motivation: One of the fundamental problems in mathematics (and hence in applications as well) is follows: Let $f : \mathbb{R}^n \to \mathbb{R}^n$. Given $y \in \mathbb{R}^n$, solve for $x$ from the system of equations

\[(1.1) \quad f(x) = y\]

and represent $x = g(y)$ and if possible find good properties of $g$, namely smoothness. More generally, if $f : \mathbb{R}^{n+m} \to \mathbb{R}^n, x \in \mathbb{R}^n, y \in \mathbb{R}^m$, solve the implicit system of equations

\[(1.2) \quad f(x, y) = 0\]

and represent $x = g(y)$.

Note that $f : \mathbb{R}^n \to \mathbb{R}^n$ can be written as $f = (f_1, \cdots, f_n)^T$, where $f_i : \mathbb{R}^n \to \mathbb{R}$ are real valued functions.

Linear system: Let us look at the well-known linear system

\[(1.3) \quad Ax = y\]
where $A$ is a given $n \times n$ matrix. The system (1.3) can be rewritten as

(1.4) \[ \sum_{j=1}^{n} a_{ij} x_j = y_i, \quad 1 \leq i \leq n \]

The system (1.3) or (1.4) is uniquely solvable for $x$ in terms of $y$ if and only if $\det A \neq 0$. In this case

\[ x = A^{-1} y \]

and $A^{-1}$ is also an $n \times n$ matrix.

We would like to address the solvability of (1.1) and (1.2) giving appropriate conditions like $\det A \neq 0$ as in linear system.

**Example 1.1.** Define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = x^2$. Clearly $f(0) = 0$. For $y > 0$, the equation $x^2 = y$ has two solutions $x_1 = +\sqrt{y}$ and $x_2 = -\sqrt{y}$ (no uniqueness) and $y < 0$, the equation has no solution. Thus we sense a trouble around $y = 0$. Note that $\frac{\partial f}{\partial x}|_{x=0} = 2x|_{x=0} = 0$ is the cause for trouble which we will see later. If we take any $a \neq 0$, and $b = a^2 = f(a)$. Then for any $y \in (b-\epsilon, b+\epsilon)$, $\epsilon$ small $\exists! x \in (a-\delta, a+\delta)$ for some $\delta$ such that $f(x) = y$. That is the equation is solvable in a neighbourhood of the equilibrium point $(a, b)$. Here observe that $\frac{\partial f}{\partial x}|_{x=a} = 2x|_{x=a} = 2a \neq 0$.

**Example 1.2.** More generally, consider $f : \mathbb{R} \times \mathbb{R} \to \mathbb{R}$ defined by $f(x, y) = x^2 + y^2 - 1$. Indeed one knows that the solutions $(x, y)$ of the equation $f(x, y) = 0$ are points on the unit circle. Consider the solvability of $x$ in terms of $y$ near the solution $(0, 1)$ of $x^2 + y^2 - 1 = 0$, that $x^2 = 1 - y^2$. For $y$ near 1, we have two solutions $x_1 = +\sqrt{1-y^2}, x_2 = -\sqrt{1-y^2}$. Similarly the case near the point $(0, -1)$. Again observe that $\frac{\partial f}{\partial x}|_{(0,\pm1)} = 2x|_{x=a} = 2a \neq 0$.

On the other hand if we consider the equilibrium point $(+1, 0)$. For $y$ near 0, $\exists! x = +\sqrt{1-y^2}$ and for $(-1, 0)$ and $y$ near 0, $\exists! x = -\sqrt{1-y^2}$. In fact, For any $(a, b)$ with $a^2 + b^2 - 1 = 0$ and $a \neq 0$, we get $\frac{\partial f}{\partial x}|_{(a,b)} \neq 0$ and the system is uniquely solvable for $y$ in a neighborhood of $b$. The situation is reversed if one looks at the possibility of solving $y$ in terms of $x$.

**Remark 1.3.** Thus we see the impact of non vanishing of the derivative on the solvability similar to $\det A \neq 0$ in the linear systems. In higher dimensional case, we have many derivatives and we need a systematic procedure to deal with such complicated case. In other words, we would like to understand the solvability of a system of non-linear equations in many unknowns. This is given via inverse and implicit function theorems. We also remark that we will only get a local theorem not a global theorem like in linear systems.
2. Partial, Directional and Frechét Derivatives

Let \( f : \mathbb{R} \to \mathbb{R} \) and \( x_0 \in \mathbb{R} \). Then \( f'(x_0) \) is normally defined as

\[
f'(x_0) = \lim_{h \to 0} \frac{f(x_0 + h) - f(x_0)}{h}
\]

(2.1)

We also aware of the fact that \( f'(x_0) \) is the slope of the tangent to the curve \( y = f(x) \) at the point \((x_0, f(x_0))\). We will soon give another interpretation of the derivatives via linear transformation which is at the heart of the concept Frechét derivative.

Let \( U \) be a open subset of \( \mathbb{R}^n \) and \( f : U \to \mathbb{R}^m \) be a multi-valued map represented by \( f = (f_1, \ldots, f_m)^T \), where \( f_i : U \to \mathbb{R} \) are real valued maps. The limit definition can easily be used to define the directional derivatives in any direction and in particular partial derivatives which are nothing but the directional derivatives along the co-ordinate axes.

**Directional and Partial Derivatives:** Recall that the derivative in (2.1) is the instantaneous rate of change of the output \( f(x) \) with respect to the input \( x \). Thus, if we consider \( f(x) \) at \( x_0 \in \mathbb{R}^n \), there are infinitely many radial directions emanating from \( x_0 \). In particular, if we have two variable function \( f(x, y) \), then \( \frac{\partial f}{\partial x}(x_0, y_0) \) is the instantaneous rate of change of \( f \) along the x-axis (keeping y-fixed) and is given by

\[
\frac{\partial f}{\partial x}(x_0, y_0) = \lim_{h \to 0} \frac{f(x_0 + h, y) - f(x_0, y)}{h}.
\]

(2.2)

Similarly \( \frac{\partial f}{\partial y}(x_0, y_0) \). Any given vector \( v \in \mathbb{R}^n \) determines a direction given by its position vector. Thus for \( x_0 \in \mathbb{R}^n \), \( f(x_0 + hv) - f(x_0), h \in \mathbb{R} \) is the change in \( f \) in the direction \( v \). This motivates us to define the directional derivative of \( f \) at \( x_0 \in \mathbb{R}^n \) in the direction \( v \), denoted by \( D_v f(x_0) \), is defined as

\[
D_v f(x_0) = \lim_{h \to 0} \frac{f(x_0 + hv) - f(x_0)}{h}
\]

whenever the limit exists. Note that if \( f = (f_1, \ldots, f_m)^T \), then

\[
D_v f(x_0) = (D_v f_1(x_0), \ldots, D_v f_m(x_0))^T.
\]

**Corollary 2.1.** If \( v = e_i = (0, \ldots, 0, 1, 0, \ldots, 0) \) is the co-ordinate axis vector, then clearly

\[
D_{e_i} f(x_0) = \frac{\partial f}{\partial x_i}(x_0) = \left( \frac{\partial f_1}{\partial x_i}(x_0), \ldots, \frac{\partial f_m}{\partial x_i}(x_0) \right)^T.
\]

**Example 2.2.** Define \( f : \mathbb{R}^n \to \mathbb{R} \) by \( f(x) = |x|^2 = \sum_{i=1}^n x_i^2 = \sum_{i=1}^n x_i x_i = \langle x, x \rangle \). Then \( \frac{\partial f}{\partial x_i}(x_0) = 2x_i \).
Now for $v \in \mathbb{R}^n$,

$$f(x_0 + hv) = \sum_{i=1}^{n} (x_{0i} + hv_i)^2$$

$$= \sum_{i=1}^{n} x_{0i}^2 + 2h \sum_{i=1}^{n} x_{0i}v_i + h^2 \sum_{i=1}^{n} v_i^2$$

$$= f(x_0) + 2h(x_0, v) + h^2 |v|^2$$

It follows that

$$D_v f(x_0) = 2(x_0, v).$$

Remark 2.3. As seen earlier the existence of all directional derivatives implies the existence of partial derivatives. But, the converse is not true.

Exercise 2.4. Let $f : \mathbb{R}^2 \to \mathbb{R}$ defined by

$$f(x, y) = \begin{cases} x + y & \text{if } x = 0 \text{ or } y = 0 \\ 1 & \text{otherwise} \end{cases}$$

Then $D_{(1,0)} f(0, 0) = D_{(0,1)} f(0, 0) = 1$, but $D_{(a,b)} f(0, 0), a \neq 0, b \neq 0$ does not exists.

Remark 2.5. Normally, we expect differentiable functions to be continuous. But the existence of all directional derivatives at a point does not imply the continuity at that point. This is serious drawback and prompts us to look for a stronger concept of derivative, namely the notion total derivative.

Exercise 2.6. Consider

$$f(x, y) = \begin{cases} \frac{xy^2}{x^2 + y^4}, & \text{if } x \neq 0 \\ 0 & \text{if } x = 0. \end{cases}$$

Show that $D_v f(0, 0)$ exists for all $v \in \mathbb{R}^2$, but $f$ is not continuous.

Exercise 2.7. Consider

$$f(x, y) = \begin{cases} \frac{xy(x^2 - y^2)}{x^2 + y^2}, & \text{if } (x, y) \neq (0, 0) \\ 0 & \text{if } (x, y) = (0, 0). \end{cases}$$

Then show that $\frac{\partial^2 f}{\partial x \partial y}(0, 0) = 1$ and $\frac{\partial^2 f}{\partial y \partial x}(0, 0) = -1$ which shows that, in general, the order of partial derivatives cannot be interchanged.
Total (Fréchet) Derivative: Recall that if \( f : \mathbb{R} \to \mathbb{R} \), then \( f'(x_0) = \alpha \) represents the slope of the tangent to the curve \( y = f(x) \) at the point \((x_0, f(x_0))\). In this case, one can associate the linear equations, namely the line \( y = \alpha x = f'(x_0)x \). In other words, the derivatives can be viewed as the linear mapping, \( T_\alpha : \mathbb{R} \to \mathbb{R} \) defined by

\[
T_\alpha x = \alpha x = f'(x_0)x.
\]

Thus interpreting any differentiation concept as a linearization is the crux of the matter not only in finite dimension, but in infinite dimensions as well. Once again recall \( f'(x_0) \) in dimension one which can be re-casted as

\[
f(x_0 + h) = f(x_0) + f'(x_0)h + r(h)
\]

where the reminder (or error) terms satisfies

\[
\lim_{h \to 0} \frac{r(h)}{h} = 0
\]

That is

\[
f(x_0 + h) = \text{value of } f \text{ at } x_0 + \text{linearized term} + \text{reminder } o(h).
\]

This can be easily extended to vector valued functions with one variable, namely \( f : \mathbb{R} \to \mathbb{R}^m \). Here \( f(x) = (f_1(x), \ldots, f_m(x))^T \) with \( x \in \mathbb{R} \) and \( f'(x_0) = (f'_1(x_0), \ldots, f'_m(x_0))^T \). If we define \( \alpha_i = f'_i(x_0) \), then \( \alpha = (\alpha_1, \ldots, \alpha_m) \in \mathbb{R}^m \). Correspondingly, we can associate a linear operator \( T_\alpha : \mathbb{R} \to \mathbb{R}^m \) defined by

\[
T_\alpha x = \alpha x = (\alpha_1 x, \ldots, \alpha_m x)^T = (f'_1(x_0)x, \ldots, f'_m(x_0)x)^T, x \in \mathbb{R}
\]

Further, for each \( i \),

\[
f_i(x_0 + h) = f_i(x_0) + f'_i(x_0)h + r_i(h), h \in \mathbb{R}
\]

where \( \frac{r_i(h)}{h} \to 0 \) as \( h \to 0 \). In vector notation

\[
f(x_0 + h) = f(x_0) + f'(x_0)h + r(h)
\]

\[
\lim_{h \to 0} \frac{r(h)}{h} = 0
\]

If \( f : \mathbb{R}^n \to \mathbb{R}^m \), then \( x_0, h \) are vectors and one has to interpret the meaning of product \( f'(x_0)h \) and division \( \frac{r(h)}{h} \).

The equation (2.7) equivalently can be rewritten as

\[
\lim_{h \to 0} \frac{|f(x_0 + h) - f(x_0) - f'(x_0)h|}{|h|} = 0
\]
where \( f'(x_0)h \) is the action of the linear operator \( f'(x_0) \) at \( h \).

**Definition 2.8.** *(Frechét Derivative):* Let \( U \subset \mathbb{R}^n \) be open and \( f : U \to \mathbb{R}^m, x_0 \in E \). We say \( f \) is differentiable (\( F \)-differentiable) at \( x_0 \) if there exists a linear operator \( T : \mathbb{R}^n \to \mathbb{R}^m \) such that

\[
(2.8) \quad \lim_{h \to 0} \frac{|f(x_0 + h) - f(x_0) - Th|}{|h|} = 0
\]

If such a linear transformation \( T \) exists, we denote \( T = f'(x_0) \). If \( f \) is differentiable at all points in \( E \), we say \( f \) is differentiable in \( E \) and \( f'(x_0) \) is also known as the total derivative or Frechét derivative of \( f \) at \( x_0 \).

**Example 2.9.** Suppose \( A \in L(\mathbb{R}^n, \mathbb{R}^m) \) be an \( m \times n \) matrix. Define \( f : \mathbb{R}^n \to \mathbb{R}^m \) by \( f(x) = Ax \). Then clearly \( f(x_0 + h) - f(x_0) = A(x_0 + h) - Ax_0 = Ah \) by linearity. Therefore \( r(h) = 0 \) and \( f'(x_0) = A \).

**Example 2.10.** Let \( f : \mathbb{R}^n \to \mathbb{R} \) by \( f(x) = |x|^2 = (x,x) \). Then \( f'(x_0)h = 2(x_0,h) \) or \( f'(x_0) = 2x_0 \)

**Remark 2.11.** The definition can be extended to infinite dimensional normed linear spaces without much difficulty and hence the enormous applications.

The following results can be verified by the reader.

**Proposition 2.12.** If \( f : \mathbb{R}^n \to \mathbb{R}^n \) and differentiable at \( x_0 \), then \( F \) is continuous at \( x_0 \).

**Proposition 2.13.** *(Chain rule):* Let \( f : U \subset \mathbb{R}^n \to \mathbb{R}^m, U \) open be differentiable at \( x_0 \) and \( g : \mathbb{R}^m \to \mathbb{R}^k \) be differentiable at \( y_0 = f(x_0) \). Then the composite function \( F(x) = g \circ f(x) = g(f(x)) \), \( F : \mathbb{R}^n \to \mathbb{R}^k \) is differentiable at \( x_0 \) and

\[
(2.9) \quad F'(x_0) = g'(f(x_0))f'(x_0).
\]

We have to interpret the product \( g'(f(x_0))f'(x_0) \) as they are linear operators.

**Linear operators and matrices:** Let \( A \) be an \( m \times n \) matrix, then define \( f : \mathbb{R}^n \to \mathbb{R}^m \) by \( f(x) = Ax \) and \( f \) is a linear operator. Conversely, if \( f \) is a linear operator, then \( \exists \) an \( m \times n \) matrix \( A \) such that \( f(x) = Ax \). Of course the matrix representation depends on the basis of \( \mathbb{R}^n \) and \( \mathbb{R}^m \). Thus every linear operator \( f \in L(\mathbb{R}^n, \mathbb{R}^m) \) can be identified with a matrix \( A \in M(m,n) \).

In the above proposition, let \( A \in M(m,n), B \in M(k,m) \) and \( C \in M(k,n) \) be, respectively, the matrices corresponding to \( f'(x_0), g'(f(x_0)) \) and \( F'(x_0) \). Then the equality
(2.8) can be read as

\[ C = BA \]

Notice that the orders of the matrices are such that the product \( BA \) is well defined.

**Exercise 2.14.** Show that in Exercises 2.4, 2.6 and 2.7, \( f'(0,0) \) does not exist.

**Remark 2.15.** This indicates that the existence of all directional derivatives are not enough to guarantee the existence of total derivative. But if the total derivative exists, then all the directional derivatives exist and in fact one can compute the total derivative using the partial derivatives.

Let \( \{e_1, \ldots, e_n\} \) and \( \{\tilde{e}_1, \ldots, \tilde{e}_m\} \) be the standard basis of \( \mathbb{R}^n \) and \( \mathbb{R}^m \) respectively. If \( f'(x_0) \) exists, then for \( j \) fixed, by definition and linearity of \( f'(x_0) \), we get

\[
\begin{align*}
    f(x_0 + he_i) - f(x_0) &= f'(x_0)(he_i) + r(he_i), \\
    &= hf'(x_0)e_i + r(he_i)
\end{align*}
\]

where \( h \in \mathbb{R} \) on \( \frac{|r(he_i)|}{h} \to 0 \) as \( h \to 0 \).

Dividing by \( h \) and taking the limit as \( h \to 0 \), we get

\[
\frac{\partial f}{\partial x_i} = f'(x_0)e_i = \left( \frac{\partial f_1}{\partial x_i}, \ldots, \frac{\partial f_m}{\partial x_i} \right)^T
\]

(2.10)

More generally, if \( v \in \mathbb{R}^n \), then \( v = \sum v_i e_i \) and

\[
D_v f(x_0) = f'(x_0)v = \sum v_i f'(x_0)e_i \text{ by linearity}
\]

\[
= \sum v_i \frac{\partial f}{\partial x_i}
\]

Thus, the matrix representation of \( f'(x_0) \) is given by

\[
f'(x_0) = \left[ \frac{\partial f_j}{\partial x_i} \right]_{1 \leq j \leq m}^{1 \leq i \leq n}
\]

(2.11)

The above results can be consolidated in

**Theorem 2.16.** Let \( f : U \subset \mathbb{R}^n \to \mathbb{R}^m \) be differentiable at \( x_0 \in U \). Then \( \frac{\partial f_i}{\partial x_i} \) exists for all \( 1 \leq i \leq n, 1 \leq j \leq m \) and \( f'(x_0) \) is given as in (2.10). That is

\[
f'(x_0)e_i = \sum_{j=1}^m \frac{\partial f_j}{\partial x_i} \tilde{e}_j
\]
3. Inverse Function Theorem

In this section, we address the solvability of the non linear equation in explicit form:

(3.1) \[ f(x) = y \]

where \( f : E \subset \mathbb{R}^n \to \mathbb{R}^m \) and \( y \in \mathbb{R}^m \) is given and \( E \) is open. These are a set of \( m \) non-linear equation in \( n \) unknowns:

(3.2)
\[
\begin{align*}
  f_1(x_1, \ldots, x_n) &= y_1 \\
  &\vdots \\
  f_m(x_1, \ldots, x_n) &= y_n
\end{align*}
\]

Given \( a \in E \), let \( b = f(a) \), then \((a, b)\) is a solution to (3.1). We want to give conditions under which one can solve for \( x \) for all \( y \) in a neighborhood of \( b \).

**Theorem 3.1** (Inverse Function Theorem). Let \( f \) be as above satisfies

1. \( f \) is a \( C^1 \) map, that is \( f'(x) \) exists for all \( x \in E \) and the mapping \( x \mapsto f'(x), E \to L(\mathbb{R}^n, \mathbb{R}^m) \) is continuous.
2. The matrix \( f'(a) \) is invertible, that is \( \det f'(a) \neq 0 \). Then \( \exists \) open sets \( U \) and \( V \) in \( \mathbb{R}^n \) and \( \mathbb{R}^m \), containing \( a \) and \( b \), respectively such that
   - (i) \( f : U \to V \) is 1-1 and onto
   - (ii) \( g = f^{-1} : V \to U \) given by \( g(f(x)) = x, \forall x \in U \) is a \( C^1 \) map.

The above theorem tells us that \( y = f(x) \) can be uniquely solved for \( x \) for \( y \) in a neighborhood of \( b \). Further the inverse map is also smooth.

We will not present a proof of the above theorem, but it is based on contraction mapping theorem from functional analysis.

**Theorem 3.2.** Let \( \phi : \mathbb{R}^n \to \mathbb{R}^n \) be a contraction map, that is \( \exists \ 0 \leq \alpha < 1 \) such that \( |\phi(x) - \phi(y)| \leq \alpha |x - y| \) for all \( x, y \in \mathbb{R}^n \). Then \( \exists \) a unique solution to the problem

\( \phi(x) = x \)

**Corollary 3.3.** The above theorem is true for any complete metric space \( X \) in place of \( \mathbb{R}^n \).

**Remark 3.4.** The proof is beautiful and constructive. Take any arbitrary point \( x_0 \in \mathbb{R}^n \). Construct inductively \( x_{n+1} = \phi(x_n), n = 0, 1, 2, \ldots \). Then one can prove that \( x_n \to x \) and \( \phi(x) = x \).
Proof. (Inverse Function Theorem; sketch): Given that $A = f'(a)$ is invertible. Since $f'$ is continuous, given $\epsilon > 0$, $\exists U \subset E$ such that $\| f'(x) - A \| \leq \epsilon \forall x \in U$. Now define for $y \in \mathbb{R}^n$, $\phi(x) = x + A^{-1}(y - f(x))$. Then $f(x) = y$ is solvable if and only if $\phi(x) = x$ has a solution.

Step 1: Show that $\phi$ is a contraction to get the solution.

Step 2: Prove the inverse map then obtained is $C^1$.

Corollary 3.5. Let $f : E \subset \mathbb{R}^n \rightarrow \mathbb{R}^m$ be $C^1$ and $\det f'(x) \neq 0, \forall x \in E$, then $f$ is an open map. The matrix of $f'(x)$ is also known as Jacobian matrix.

4. Implicit Function Theorem

Quite often, we do not expect to get equations in explicit form $y = f(x)$ like in $x^2 + y^2 - 1 = 0$, we may get a relation connecting the variables $x$ and $y$. Let $f : E \subset \mathbb{R}^n \times \mathbb{R}^m \rightarrow \mathbb{R}^n$ be a $C^1$ map. We wish to solve for $x$ in terms of $y$ of the system of equations

$$(4.1) \quad f(x, y) = 0$$

This is a system of $n$ equations in $n + m$ variables:

$$(4.2) \quad \begin{cases} f_1(x_1, \cdots x_n, y_1, \cdots y_m) = 0 \\ \vdots \\ f_n(x_1, \cdots x_n, y_1, \cdots y_m) = 0 \end{cases}$$

**Linear System:** If $f_i$’s are linear, then $\exists n \times n$ matrix $A$ and $n \times m$ matrix $B$ so that (4.2) reduces to

$$(4.3) \quad Ax + By = 0$$

If $A$ is invertible, then $x$ can be solve as

$$x = -A^{-1}By$$

Let $T \in L(\mathbb{R}^{n+m}, \mathbb{R}^n)$ be a linear operator on $\mathbb{R}^{n+m}$. Indeed $T$ can be represented as an $n \times (n+m)$ matrix like $[A, B]$. For $(h, k) \in \mathbb{R}^{n+m}$, we write $(h, k) = (h, 0) + (0, k), h \in \mathbb{R}^n, k \in \mathbb{R}^m$ and by linearity of $T$, we get

$$T(h, k) = T(h, 0) + T(0, k)$$

Thus, define $T_x : \mathbb{R}^n \rightarrow \mathbb{R}^n, T_y : \mathbb{R}^m \rightarrow \mathbb{R}^n$ as

$$T_x h = T(h, 0), T_y k = T(0, k)$$
and
\[ T(h, k) = T_x h + T_y k \]

That is, we have
\[ T = T_x + T_y \]
with \( T_x \in L(\mathbb{R}^n, \mathbb{R}^n) \) and \( T_y \in L(\mathbb{R}^m, \mathbb{R}^n) \)

**Theorem 4.1 (Implicit Function Theorem (Linear version)).** Assume \( T \in L(\mathbb{R}^{n+m}, \mathbb{R}^n) \) and \( T_x \) is invertible, then for any \( k \in \mathbb{R}^m, \exists! h \in \mathbb{R}^n \) such that \( T(h, k) = 0 \) and the solution is given by \( h = -T_x^{-1}T_y(k) \).

**Theorem 4.2 (Implicit Function Theorem (Non Linear version)).** Let \( f : E \subset \mathbb{R}^{n+m} \rightarrow \mathbb{R}^n \) be a \( C^1 \) map such that \( f(a, b) = 0 \) for some \((a, b) \in E \). Put \( T = f'(a, b) \in L(\mathbb{R}^{n+m}, \mathbb{R}^n) \) and \( T = T_x + T_y \) as above and assume \( T_x \) is invertible. Then \( \exists \) open sets \( U \subset E \subset \mathbb{R}^{n+m}, W \subset \mathbb{R}^m \) with \( b \in W, (a, b) \in U \) satisfying

(i) for every \( y \in W, \exists! x \) such that \((x, y) \in U \) and \( f(x, y) = 0 \).

(ii) define \( g : W \rightarrow \mathbb{R}^n \) by \( g(y) = x \), then \( g \) is a \( C^1 \) map such that \( g(b) = a \), and \( f(g(y), y) = 0 \). Further
\[ g'(b) = -T_x^{-1}T_y. \]

**Proof.** (Idea): The proof is based as an application of Inverse function theorem applied to \( F : E \subset \mathbb{R}^{n+m} \rightarrow \mathbb{R}^{n+m} \) defined by
\[ F(x, y) = (f(x, y), y). \]

We will not get into the details. \( \square \)

**Remark 4.3.** For the system (4.2), we have \( T = [T_x, T_y] \), where
\[ T_x = \begin{bmatrix} D_{x_1}f_1, \cdots D_{x_n}f_1 \\ D_{x_1}f_n, \cdots D_{x_n}f_n \end{bmatrix} \quad \text{and} \quad T_y = \begin{bmatrix} D_{y_1}f_1, \cdots D_{y_m}f_1 \\ D_{y_1}f_n, \cdots D_{y_m}f_n \end{bmatrix} \]

**Example 4.4.** Define \( f : \mathbb{R}^{2+3} \rightarrow \mathbb{R}^2, n = 2, m = 3, \) by
\[ f_1(x_1, x_2, y_1, y_2, y_3) = 2e^{x_1} + x_2y_1 - 4y_2 + 3 \]
\[ f_2(x_1, x_2, y_1, y_2, y_3) = x_2 \cos x_1 - 6x_1 + 2y_1 - y_3 \]

Take \( a = (1, 1), b = (3, 2, 7), \) then \( f(a, b) = 0 \). Compute \( T = (T_x, T_y) \) where
\[ T_x = \begin{bmatrix} 2 & 3 \\ -6 & 1 \end{bmatrix} \quad \text{and} \quad T_y = \begin{bmatrix} 1 & -4 & 0 \\ 2 & 0 & 1 \end{bmatrix} \]
Clearly $T_x$ is invertible and

$$T_x^{-1} = \frac{1}{20} \begin{bmatrix} 1 & -3 \\ 6 & 2 \end{bmatrix}. $$

Hence one can solve for $x = g(y)$ in a neighborhood of $(a, b)$.

**REFERENCES**