IMPLICIT FUNCTION THEOREM: ESTIMATES ON THE SIZE OF THE DOMAIN

ASHUTOSH JINDAL, RAVI BANAVAR, AND DEBASISH CHATTERJEE

Systems & Control Engineering

ABSTRACT. In this article we provide explicit estimates on the domain on which the Implicit Function Theorem (ImFT) and the Inverse Function Theorem (IFT) are valid. For maps which are continuously differentiable up to second order, the estimates depend upon the magnitude of the first order derivatives evaluated at the point of interest and a bound on the second order derivatives over a region of interest. One of the key contributions of this article is to come up with estimates that require minimal numerical computation. In particular, we do not perform any optimization to come up with these estimates. The derived bounds are then applied to compute the robustness margin for Quadratic Problems and then utilize these bounds to compute the allowable power variations to ensure stable operations of the Power System Networks. Another key application of these bounds is in estimating the domain of feedback linearization for discrete time control systems which is to be presented in a companion paper along with results on the feedback linearizability of the numerical integration techniques.

§ 1. Introduction

The Implicit Function Theorem (ImFT) and the Inverse Function Theorem (IFT) constitute a cornerstone of mathematical analysis and multivariate calculus. These theorems serve as the basis of several existential results in mathematics, and find applications in areas such as optimization [BHM99, Don99, Fle01], control theory, and numerical analysis [Mor94, NS96]. In a control-theoretic setting, the ImFT finds its applications in showing the existence of solutions of ordinary differential equations [Eld13, DK00] and thus asserts the existence of a unique trajectory for the dynamical system and also asserts the continuity properties of the solutions with respect to initial conditions and the control input. The implicit function theorem also finds its applications in the areas such as feedback linearization of numerically discretized systems [JBC21]. These qualitative existential results are of significant importance and serve as the basis of several algorithms and engineering applications.

However, as an engineer, one is also interested in knowing an estimate of the domain on which these methods are applicable. This serves as our motivation to arrive at an estimate of the domain on which the ImFT and IFT are valid. Despite the wide applicability of ImFT and IFT, very scarce attempts have been made on the quantitative side of the results asserted by these theorems. To the best of the authors’ knowledge, the first nontrivial estimate of the neighborhoods involved in ImFT is provided in [CHP03]. These estimates are based on the application of the Roche Theorem [Ash14] for the scalar case and then are applied inductively to generalize it for vector valued maps. The bounds provided in [CHP03] are based on the boundedness of the underlying complex map over a bounded domain. The accuracy of these estimates is limited by the accuracy with which one computes the bound of the underlying map over the domain of interest. Moreover, for the vector valued maps, the
bounds are calculated component wise using induction. Thus, estimates get poorer with an increase in the dimension. Furthermore, at each step, one needs to solve a set of nonlinear equations recursively to compute the bounds for that particular step. This makes the process computationally intensive and difficult to scale for very large dimensions. Another set of bounds based on bounded subgradients was provided by [Pap05] for the ImFT for the Lipschitz continuous functions. Although the bounds provided by [Pap05] cover a larger class of functions than those covered by [CHP03], to compute these bounds one needs to compute the subgradients over the domain of interest and needs to ensure that these gradients are invertible.

For IFT, [AMR07] provides a set of bounds based on the magnitude of the first-order derivatives evaluated at the point of interest and the boundedness of the magnitude of second-order derivatives over a domain. The boundedness of the second-order derivative restricts the application of these bounds to functions that are continuously differentiable at least up to second order. However, these bounds require minimal numerical computations, since one only needs to compute the bounds on the second-order derivatives, and thus are useful in scenarios where limited computation capacity is available.

§ 1.1. Contributions.
(i) As our first contribution, adopting a similar approach to [AMR07], we provide a set of bounds for the ImFT which require minimal numerical computation. Similar to [AMR07], these bounds are dependent on the first order derivatives evaluated at the point of interest and the second order derivatives evaluated on a bounded set containing the point of interest.
(ii) As an intermediate result, on the way to derive the bounds for the ImFT, we also improve the bounds provided by [AMR07] for IFT. However, this is achieved at the cost of a loss of the uniqueness of the inverse map.
(iii) We apply the obtained bounds on ImFT and IFT to investigate the robustness of the solutions of a system of nonlinear equations. We illustrate the method by applying it to Quadratically Constrained Quadratic Problems (QCQP) and its application in the power system networks.

§ 1.2. Notations. We use standard notations throughout the article. The set of real numbers is denoted by \( \mathbb{R} \) and the set of integers is denoted by \( \mathbb{Z} \). The set of positive integers is denoted by \( \mathbb{N} \), and the set of all positive integers less than or equal to \( n \) is denoted by \( \mathbb{N}_n \).

For \( U \subset \mathbb{R}^n \) a nonempty set, \( \text{cl} U \) denotes the smallest closed set containing \( U \) called the closure of \( U \); \( \text{int} U \) is the largest open set contained in \( U \) called the interior of \( U \); and \( \text{bd} U = (\text{cl} U) \setminus \text{int} U \) denotes the boundary of \( U \). For a given \( r > 0 \) and \( x_0 \in \mathbb{R}^n \),

\[
\mathcal{B} (x_0, r) = \{ x \in \mathbb{R} \mid \| x - x_0 \| < r \}
\]

denotes the open ball of radius \( r \) centered around \( x_0 \), for some fixed well-defined norm \( \| \cdot \| \) on \( \mathbb{R}^n \). The identity matrix of order \( n \) is denoted by \( I_n \); if the order is unambiguous from the expression, we may drop the subscript and simply write \( I \). For given \( n, m \in \mathbb{N} \), for \( A : \mathbb{R}^n \rightarrow \mathbb{R}^m \) a linear map, \( \| A \| \) is defined by

\[
\| A \| = \sup \{ \| Ax \| \mid x \in \mathcal{B} (0, 1) \subset \mathbb{R}^n \} .
\]

For a given \( \nu \in \mathbb{N} \) and nonempty set \( U \) and \( V \), \( C^\nu (U, V) \) denotes the class of \( \nu \)-times continuously differentiable maps with domain and co-domain as \( U \) and \( V \) respectively. The differential operator is denoted by \( \mathcal{D} \). For a given \( f \in C^\nu (U, \mathbb{R}^m) \), mapping \( \mathbb{R}^n \ni U \rightarrow f (x) \in \mathbb{R}^m, x_0 \in U, \mathcal{D} f (x_0) \) is the Jacobian matrix of \( f \) evaluated at \( x_0 \). The partial derivatives are denoted by adding subscript to differential operator i.e., \( \frac{\partial}{\partial x} : \mathcal{D} x \).
§2. Estimates for the Inverse Function Theorem

The Inverse Function Theorem (IFT) typically serves as a precursor to the Implicit Function Theorem (ImFT). In this light, we begin our discussion by stating the IFT. Due to a rich history of the theorem and its wide applications, it has been studied, with varying degree of generalizations, by several authors and in several sources [Spi95, Lan97, Cla76]. We shall refer to the following version from [AMR07] as a prototypical version of the IFT.

**Theorem 2.1.** (Inverse Function Theorem) [AMR07, Theorem 2.5.2]: Let $U \subset \mathbb{R}^n$ be a nonempty open set and $f: U \rightarrow \mathbb{R}^n$ be a $C^2$ map. Suppose $x_0 \in U$ is such that $Df(x_0): \mathbb{R}^n \rightarrow \mathbb{R}^n$ is an isomorphism. Then there exists a neighborhood $\mathcal{O}$ of $x_0$ such that $f|_{\mathcal{O}}$ is a $C^2$ diffeomorphism to its image. Moreover, for $(x, y) \in \mathcal{O} \times f(\mathcal{O})$, satisfying $y = f(x)$ one has

$$Df^{-1}(y) = (Df(x))^{-1}$$

The following proposition provides an initial set of estimates for the IFT in terms of (a) the magnitude of the first order derivatives evaluated at the point of interest, and (b) the bounds on the magnitude of second order derivatives over a region of interest.

**Proposition 2.2.** [AMR07, Proposition 2.5.6] Let $U \subset \mathbb{R}^n$, be a nonempty open set and $f: U \rightarrow \mathbb{R}^n$ be defined as in Theorem 2.1. Let $L_f := \|Df(x_0)\|$ and $M_f := \|Df^{-1}(x_0)\|$. For $R_f > 0$, set $K_f := \sup \left\{ \|D^2f(x)\| \mid x \in B(x_0, R_f) \right\}$, $N_f := 8M_f^3K_f$, and define

$$P_f := \min \left\{ \frac{2K_f M_f}{R_f}, R_f \right\}, \quad P'_f := \frac{P_f}{2M},$$

$$Q_f := \min \left\{ \frac{2M}{2N_f L_f}, \frac{P_f}{M_f}, P_f \right\}, \quad Q'_f := \frac{Q_f}{2L_f}.$$ 

Then there exist

(2.2-a) an open set $H_f \subset B(0, P_f)$ such that $f|_{H_f}$ is $C^2$ diffeomorphism to $B(0, P'_f)$, and

(2.2-b) an open set $H'_f \subset B(x_0, Q_f)$ such that $f^{-1}|_{H'_f}$ is $C^2$ diffeomorphism to $B(x_0, Q'_f)$. 

---

**Figure 1. Neighborhoods given in Proposition 2.2**
Remark 2.3. Here are three key observations on Proposition 2.2:

(2.3-a) $P_f$ is independent of $L_f$ (the magnitude of the derivative of the forward map), and is only dependent upon the magnitude of the derivative of the inverse map $M_f$, and the bound of the second order derivative $K_f$.

(2.3-b) $P_f$ is determined by the minimum of two quantities $\frac{1}{2K_f M_f}$ and $R_f$. A larger value of $R_f$ will lead to larger value of $K_f$ and hence $\frac{1}{2K_f M_f}$ will limit the choice of $P_f$. A smaller choice for $R_f$ will give larger $\frac{1}{2K_f M_f}$. However, since $R_f$ is smaller than $\frac{1}{2K_f M_f}$, $P_f$ is now limited by $R_f$.

(2.3-c) The choice of $P_f$ affects $Q_f$, this is because $f^{-1}$ is defined over $B\left(y_0, P_f'\right)$ and hence $Q_f$ is limited by $P_f'$.

A pictorial representation of these neighborhoods is given in Figure 1. Based on this result, let us now examine the ImFT and attempt to get similar bounds on the domain of its applicability in the following section.

§ 3. Estimating Neighborhoods for the Implicit Function Theorem

For the sake of clarity, we first state the Implicit Function Theorem (ImFT). As in the case of the IFT, the ImFT has been studied, with varying degrees of generalizations, by several authors in various sources [Zei93, KP02, Zei95b, Spi95, Lan97, AMR07]. In this article the following theorem from [AMR07] shall be referred to as the Implicit Function Theorem (ImFT).

**Theorem 3.1.** (Implicit Function Theorem) [AMR07, Theorem 2.5.7] Let $U \subseteq \mathbb{R}^n$, $V \subseteq \mathbb{R}^m$ be nonempty open sets. Let $f : U \times V \longrightarrow \mathbb{R}^n$ be a $C^2$ map. Suppose for some $x_0 \in U$ and $y_0 \in V$ the partial derivative in the second argument

$$D_y f(x_0, y_0) : \mathbb{R}^m \longrightarrow \mathbb{R}^n,$$

is an isomorphism. Then there exist neighborhoods $\mathcal{O}(x_0) \ni x_0$ and $\mathcal{O}(y_0) \ni y_0 = f(x_0, y_0)$ and a unique $C^2$ map

$$g : \mathcal{O}(x_0) \times \mathcal{O}(y_0) \longrightarrow V$$

such that

$$f(x, g(x, w)) = w$$

for all $(x, w) \in \mathcal{O}(x_0) \times \mathcal{O}(y_0)$.

**Proof.** Our proof traverses the usual route to prove the ImFT which involves augmenting the system with extra equations to extend dimension and then invoke the IFT. This procedure helps us to usher similar notions of bounds into the ImFT as well. Without loss of generality, assume $x_0 = 0$, $y_0 = 0$, and $w_0 := f(x_0, y_0) = 0$. Define

$$(3.1) \quad U \times V \ni (x, y) \longmapsto \Phi(x, y) := (A_1 x, A_2 f(x, y)) \in \mathbb{R}^n \times \mathbb{R}^m,$$

where $A_1$ and $A_2$ are isomorphisms between appropriate spaces. Clearly $D\Phi(0, 0)$ is invertible, given by

$$D\Phi(0, 0) = \begin{pmatrix} A_1 & 0 \\ A_2 D_x f(0, 0) & A_2 D_y f(0, 0) \end{pmatrix}.$$ 

Thus, from Theorem 2.1, there exist an open set $\mathcal{O}(x_0) \times \mathcal{O}(u_0) \ni (x, w)$, such that $\Phi$ has a unique $C^2$ inverse defined by

$$\mathcal{O}(x_0) \times \mathcal{O}(y_0) \ni (x, w) \longmapsto \Phi^{-1}(x, w) = (A_1^{-1} x, \psi(x, w)) \in \mathcal{O}(x_0) \times \mathcal{O}(y_0).$$
Indeed, it is clear that
\[ \Phi(A_1^{-1}, \psi(x, w)) = (x, A_2 f(x, \psi(x, w))) = (x, w), \]
which implies \( f(x, \psi(x, w)) = A_2^{-1} w. \) Replacing \( w \) with \( A_2 w \), and letting \( g(x, w) := \psi(x, A_2 w) \) we have
\[ f(x, g(x, w)) = w \] for all \( (x, w) \in \mathbb{O}((x_0) \times \mathbb{O}(w_0)) \)
concluding the proof. \( \square \)

**Remark.** The proof of the ImFT encountered in literature generally employs \( A_1, A_2 \) as the identity map. We introduce these two constructs here since they will turn out to be useful in computing novel bounds for the ImFT. The goal is now to come with estimates on \( \mathbb{O}(x_0) \) and \( \mathbb{O}(w_0) \); the latter correspond to a pair of sets on which the map \( g \) is well defined. We shall use the bounds derived for the Theorem 2.1 (IFT) for this purpose. Note that, the map \( g \) as given by Theorem 3.1 is well defined if and only if \( \Phi \) is invertible.

**Proposition 3.2.** Let \( U \subset \mathbb{R}^n \), and \( V \subset \mathbb{R}^m \) be nonempty open sets, and let \( f : U \times V \to \mathbb{R}^m \) be as defined in Theorem 3.1. Let \( A_1, A_2 \) be two isomorphisms on appropriate spaces and let
\[ (3.2) \quad U \times V \ni (x, y) \mapsto \Phi(x, y) := (A_1 x, A_2 f(x, y)) \in \mathbb{R}^n \times \mathbb{R}^m. \]
Define \( L_x \coloneqq \|D_x f(0, 0)\|, L_y \coloneqq \|D_y f(0, 0)\|, \) and \( M_y \coloneqq \|D_y f^{-1}(0, 0)\|. \) Moreover, let \( R > 0 \) and \( K_f \) be such that
\[ \|D^2 f(x_0, y_0)\| \leq K_f \text{ for all } (x, y) \in cl \mathcal{B}((x_0, y_0), R). \]
Set \( L := \|A_1\| + \|A_2\| (L_x + L_y), M := \|A_1^{-1}\| (1 + M_y L_x) + \|A_2^{-1}\| M_y, \) and \( K := \|A_2\| K_f \) to define
\[ P := \min \left\{ \frac{1}{2KM}, R \right\}, \quad P' := \frac{P(1 - MKP)}{2M}. \]
Further, set \( N = 8M^2 K \) and define
\[ Q := \min \left\{ \frac{1}{2NL}, \frac{P}{M}, P \right\}, \quad \text{and } Q' := \frac{Q(1 - NLQ)}{L}. \]
Then there exist
\[ (3.2-a) \quad \text{an open set } H \subset \mathcal{B}((x_0, y_0), P) \text{ such that } \Phi \text{ maps } H \text{ diffeomorphically onto } \mathcal{B}(\omega_0, P') \text{ with } \omega_0 := \Phi(x_0, y_0), \text{ and} \]
\[ (3.2-b) \quad \text{an open set } H' \subset \mathcal{B}(\omega_0, Q) \text{ such that } \Phi^{-1} \text{ maps } H' \text{ diffeomorphically onto } \mathcal{B}((x_0, y_0), Q'). \]

**Proof.** Without loss of generality, set \( x_0 = 0 \in \mathbb{R}^n, \ y_0 = 0 \in \mathbb{R}^m \) and \( f(0, 0) = 0 \in \mathbb{R}^m. \) Clearly \( \Phi(0, 0) = (0, 0). \) From the definition of \( \Phi, \) we see that its Jacobian at \( (0, 0) \) is given by
\[ D\Phi(0, 0) = \begin{pmatrix} A_1 & 0 \\ A_2 D_x f(0, 0) & A_2 D_y f(0, 0) \end{pmatrix}. \]
From the submultiplicative property of the induced norms\(^1\) and the triangle inequality we have
\[ \|D\Phi(0, 0)\| \leq \|A_1\| + \|A_2\| (L_x + L_y) =: L. \]
Similarly,
\[ D\Phi^{-1}(0, 0) = \begin{pmatrix} A_1^{-1} & 0 \\ -(D_y f(0, 0))^{-1} D_x f(0, 0) A_1^{-1} & (D_y f(0, 0))^{-1} A_1^{-1} \end{pmatrix}, \]
\(^1\)Let \( A : \mathbb{R}^n \to \mathbb{R}^m \) and \( B : \mathbb{R}^m \to \mathbb{R}^k \) be two linear operators. Then \( \|AB\| \leq \|A\| \|B\| \) where \( \|A\| \) is the induced operator norm.
Fix \( D \) in (3.3) and finds out the domain on which it is invertible. Let us examine the expression

\[ g \text{ is invertible. Hereinafter the proof proceeds by closely examining term in the parentheses in (3.3) and finds out the domain on which it is invertible. Let us examine the expression} \]

\[ D\Phi(x, y) = D\Phi(0, 0) + \int_{0}^{1} D^2\Phi(tx, ty) \cdot (x, y) \, dt \]

\[ (3.3) \]

From the above equation, \( D\Phi(x, y) \) is invertible if and only if

\[ \left(1 + D\Phi^{-1}(0, 0) \int_{0}^{1} D^2\Phi(tx, ty) \cdot (x, y) \, dt \right) \]

is invertible. Hereinafter the proof proceeds by closely examining term in the parentheses in (3.3) and finds out the domain on which it is invertible. Let us examine the expression \( D\Phi^{-1}(0, 0) \int_{0}^{1} D^2\Phi(tx, ty) \cdot (x, y) \, dt \). We have

\[ \left\| D\Phi^{-1}(0, 0) \int_{0}^{1} D^2\Phi(tx, ty) \cdot (x, y) \, dt \right\| \leq \left\| D\Phi^{-1}(0, 0) \int_{0}^{1} \left\| D^2\Phi(tx, ty) \right\| \| (x, y) \| \, dt \right\| \leq MK \| (x, y) \|. \]

Fix \( \lambda \in ]0, 1[ \), for all \( (x, y) \in B \left( (x_0, y_0), \frac{A}{MK} \right) \)

\[ \left\| D\Phi^{-1}(0, 0) \int_{0}^{1} D^2\Phi(tx, ty) \cdot (x, y) \, dt \right\| \leq \lambda < 1. \]

Using [AMR07, Lemma 2.5.4], \( D\Phi(x, y) \) is invertible for all \( (x, y) \in B \left( (x_0, y_0), \frac{A}{MK} \right) \) and

\[ \left\| (D\Phi(x, y))^{-1} \right\| = \left\| \left(1 + D\Phi^{-1}(0, 0) \int_{0}^{1} D^2\Phi(tx, ty) \cdot (x, y) \, dt \right)^{-1} D\Phi^{-1}(0, 0) \right\| \]

\[ \leq \left\| D\Phi^{-1}(0, 0) \right\| \left\| \left(1 + D\Phi^{-1}(0, 0) \int_{0}^{1} D^2\Phi(tx, ty) \cdot (x, y) \, dt \right)^{-1} \right\| \]

\[ \leq \frac{M}{1 - MK \| (x, y) \|} \quad \text{(from [AMR07, Lemma 2.5.4])} \]

\[ \leq \frac{M}{1 - \lambda} \quad \text{for all } (x, y) \in B \left( (0, 0), \frac{A}{MK} \right). \]

Our goal here is to find, for a given \( \omega \), an \( (x^*, y^*) \) such that \( \Phi(x^*, y^*) = \omega \). We do this by converting it into a fixed point problem.\(^2\) In this direction, set \( P_0 = \min \left\{ \frac{A}{MK}, R \right\} \), and for each fixed \( \omega \), define \( g_\omega : \text{cl} B ((0, 0), P_0) \rightarrow \text{cl} B ((0, 0), P_0) \) mapping

\[ (x, y) \mapsto g_\omega(x, y) := D\Phi^{-1}(0, 0) \left( \omega + D\Phi(0, 0) \cdot (x, y) - \Phi(x, y) \right). \]

Clearly \( g_\omega(x^*, y^*) = (x^*, y^*) \) implies \( \Phi(x^*, y^*) = \omega \). Thus for each fixed \( \omega \), \( g_\omega \) has a fixed point \( (x^*, y^*) \) if and only if \( \Phi(x^*, y^*) = \omega \). In this way, the problem of finding an inverse image is now transformed into a fixed point paradigm. To obtain the estimates on the neighborhoods let us look at the expression

\[ D\Phi(0, 0)(x, y) - \Phi(x, y) = \int_{0}^{1} \int_{0}^{1} D^2\Phi(stx, sty) \cdot ((sx, sy), (x, y)) \, ds \, dt. \]

Since \( \left\| D^2\Phi(x, y) \right\| \leq K \) for all \( (x, y) \in \text{cl} B ((0, 0), P_0) \) we have

\[ \left\| D\Phi(0, 0)(x, y) - \Phi(x, y) \right\| \leq K \| (x, y) \|^2 \] for all \( (x, y) \in B ((0, 0), P_0) \).

\(^2\)Let \( U \subset \mathbb{R}^n \) be a nonempty set. For a given \( f : U \subset \mathbb{R}^n \rightarrow U, x^* \in U \) is called a fixed point of \( f \) if it satisfies \( f (x^*) = x^* \).
To prove (3.2-b), denote a diffeomorphism. From (3.5) we get
\[ \|Dg_\omega(x, y)\| = \|D\Phi^{-1}(0, 0)(\omega + D\Phi(0, 0)(x, y) - \Phi(x, y))\| \]
\[ \leq \|D\Phi^{-1}(0, 0)\| (\|\omega\| + \|D\Phi(0, 0)(x, y) - \Phi(x, y)\|) \]
\[ \leq M(\|\omega\| + K \|(x, y)\|^2). \]
Therefore, for all \( \omega \in B ((0, 0), P_0') \),
\[ clB ((0, 0), P_0) \ni (x, y) \mapsto g_\omega(x, y) \in clB ((0, 0), P_0). \]
Moreover, \( Dg_\omega(x, y) \) satisfies
\[ \|Dg_\omega(x, y)\| = \|D\Phi^{-1}(0, 0)\| \int_0^1 D^2\Phi(tx, ty) \cdot (x, y) \, dt \]
\[ \leq \|D\Phi^{-1}(0, 0)\| \int_0^1 D^2\Phi(tx, ty) \cdot (x, y) \, dt \]
\[ \leq MK \|(x, y)\| \]
\[ \leq MK \frac{\lambda}{MK} \leq \lambda < 1 \text{ for all } (x, y) \in B ((0, 0), P_0). \]
Thus for all \( \omega \in B ((0, 0), P_0') \), \( g_\omega \) is a contraction [Zei93, Definition 1.1]. Hence, from the Banach Fixed Point Theorem [Zei95a, Theorem 1.A], for each \( \omega \in B ((0, 0), P_0') \) there exists a unique \( (x^*, y^*) := (x^*(\omega), y^*(\omega)) \) such that
\[ g_\omega(x^*, y^*) = D\Phi^{-1}(0, 0) \cdot (\omega + D\Phi(0, 0)(x^*, y^*) - \Phi(x^*, y^*)) = (x^*, y^*), \]
which implies \( \omega = \Phi(x^*, y^*) \). Setting \( \omega \mapsto \Phi^{-1}(\omega) := (x^*(\omega), y^*(\omega)) \) establishes the existence of local inverse. From the expression of \( P_0' \), one can show that it achieves its maximum for \( P_0 = \frac{1}{MK} \). Setting
\[ P := \min \left\{ \frac{1}{2MK}, R \right\} \quad \text{and} \quad P' := \frac{P(1 - MKP)}{M} \]
proves (3.2-a) i.e., there exists an \( H \subset B ((x_0, y_0), P) \) such that \( \Phi|_H : H \rightarrow B (\omega_0, P') \) is a diffeomorphism.

To prove (3.2-b), denote \( \Phi^{-1} \) as \( \Psi \). From theorem 2.1 we have
\[ D\Psi(\omega_0) = D\Phi^{-1}(\omega_0) = (D\Phi(x_0, y_0))^{-1}. \]
Clearly \( D\Psi(\omega_0) \) is an isomorphism and therefore \( \Psi : B (\omega_0, P') \rightarrow H \) satisfies the conditions of Theorem 2.1. Moreover, we have
\[ D\Psi^{-1}(x_0, y_0) = D\Phi(x_0, y_0) \]
Consequently, \( \|D\Psi(\omega_0)\| = M \) and \( \|D\Psi^{-1}(x_0, y_0)\| = L. \) Since \( \Phi \circ \Psi(\omega) = \omega \) for all \( \omega \in B ((0, 0), P/2M) \), with simple mathematical manipulations one can show that
\[ \|D^2\Psi(\omega)\| \leq 8M^3K =: N \text{ for all } \omega \in B ((0, 0), P/2M). \]
Applying Proposition (3.2-a) on \( \Psi \) proves (3.2-b). \( \square \)

Remark. Although, in essence Proposition 3.2 is just an application of Proposition 2.2 on \( \Phi \). We give a detailed proof for it as it will help us develop the improved versions of these proposition which provide tighter estimates and or are applicable to a larger class of functions.

Remark. Proposition 3.2 gives the domain on which the inverse map \( \Phi^{-1} \) exists and hence relies on uniqueness of the solution of \( \Phi(x, y) = \omega \in \mathbb{R}^n \times \mathbb{R}^m \). However, if we relax the uniqueness requirement, and are just interested in finding out a domain \( \mathcal{O} (\omega_0) \) around \( \omega_0 = \Phi(x_0, y_0) \) and an open set \( \mathcal{O} (x_0, y_0) \ni (x_0, y_0), \) such that for all \( \omega \in \mathcal{O} (\omega_0), \) there
exists (not necessarily unique) \((x, y) \in \Theta(x_0, y_0)\) satisfying \(\Phi(x, y) = \omega\). Then we can get improved estimates on \(\Theta(\omega_0)\) and \(\Theta(x_0, y_0)\) given by the following proposition.

**Proposition 3.3.** Let \(U \subset \mathbb{R}^n\) and \(V \subset \mathbb{R}^m\) be nonempty open sets and \(U \times V \ni (x, y) \mapsto \Phi(x, y) \in \mathbb{R}^n \times \mathbb{R}^m\). Let \(\Phi : H(z_0) \subset B(z_0, P_0) \rightarrow B(\omega_0, P_0')\) be as defined in Proposition 3.2. Then, for any 

\[
0 < r < \min \left( \frac{1}{MK}, \frac{R}{A} \right) \quad \text{and} \quad \epsilon = \frac{r(2 - rMK)}{2M},
\]

for all \(\omega \in B(\omega_0, \epsilon)\) with \(\omega_0 = \Phi(x_0, y_0)\), there exists a (not necessarily unique) \((x, y) \in B((x_0, y_0), r)\) such that \(\Phi(x, y) = \omega\).

**Proof.** The proof involves repeated applications of Proposition 3.2 and we employ mathematical induction to prove the above result.

**Base Case:** For notational convenience set \((x, y) = z\). Since from the hypothesis \(D\Phi(z_0)\) is nonsingular, for a given \(\lambda \in [0, 1]\), Proposition 3.2 asserts that there exists an open set \(H(z_0) \ni z_0\), such that \(\Phi : H(z_0) \subset B(z_0, P_0) \rightarrow B(\omega_0, P_0')\) is a diffeomorphism. Further from (3.4) we have

\[
\left\| D\Phi(z)^{-1} \right\| < \frac{M}{1 - A} \quad \text{for all} \quad z \in B(z_0, P_0).
\]

For any \(z_1 \in \text{bd} H(z_0) \subset \text{cl} B(z_0, P_0)\), we have

\[
\left\| D\Phi(z_1)^{-1} \right\| \leq \frac{M}{1 - A} = M_1 < \infty,
\]

which implies that \(D\Phi(z_1)\) is nonsingular. Applying Proposition 3.2 on \(\Phi\) at \(z_1\), for

\[
P_1 := \frac{A(1 - A)}{MK} \quad \text{and} \quad P_1' := \frac{A(1 - A)^3}{MK}
\]

(for simplicity we have assumed \(R \gg 1\)) and \((z_1, \omega_1) := (z_1, \Phi(z_1))\), there exists an open set \(H(z_1) \subset B(z_1, P_1)\) such that \(\Phi|_{H(z_1)}\) is a diffeomorphism onto \(B(\omega_1, P_1')\) (refer Figure 2). Similarly, for any given \(z_2 \in \text{bd} H(z_1)\) one has
Therefore

\[ \left\| D\Phi(z_2)^{-1}\right\| \leq \frac{M/(1 - \lambda)}{1 - MK \left\| z_1 - z_2 \right\|/(1 - \lambda)} \]
\[ \leq \frac{M}{(1 - \lambda)^2} =: M_2. \]

Since \( D\Phi(z_2) \) is nonsingular, Proposition 3.2 is applicable at \( z_2 \).

**Setting up the induction:** For a given \( k \in \mathbb{N} \) define

\[ M_k := \frac{M}{(1 - \lambda)^k}, \quad P_k := \frac{\lambda(1 - \lambda)^k}{MK}, \quad \text{and} \quad P'_k := \frac{\lambda(1 - \lambda)^{2k+1}}{MK}. \]

Let \( z_k \in \text{bd} \, H(z_{k-1}), \omega_k = \Phi(z_k) \). Let \( H(z_k) \subset B(z_k, P_k) \) be an open set such that \( \Phi|_{H(z_k)} \) is a diffeomorphism onto \( B(\omega_k, P'_k) \).

**Induction Step:** Consider a \( z_{k+1} \in \text{bd} \, H(z_k) \). Then we have

\[ \left\| D\Phi(z_{k+1})^{-1}\right\| \leq \frac{M_k}{1 - M_kK \left\| z_k - z_{k+1}\right\|} \]
\[ \leq \frac{M}{(1 - \lambda)^{k+1}} =: M_{k+1}. \]

Therefore \( D\Phi(z_{k+1}) \) is nonsingular. From proposition 3.8, we have

\[ P_{k+1} := \frac{\lambda}{M_{k+1}K} = \frac{\lambda(1 - \lambda)^{k+1}}{MK}, \]

and

\[ P'_{k+1} := \frac{\lambda(1 - \lambda)P_{k+1}}{M_{k+1}K} = \frac{\lambda(1 - \lambda)^{2k+3}}{MK}, \]

and an open set \( H(z_{k+1}) \subset B(z_{k+1}, P_{k+1}) \) such that \( \Phi|_{H(z_{k+1})} \) is a diffeomorphism onto \( B(\omega_{k+1}, P'_{k+1}) \) with \( \omega_{k+1} = \Phi(z_{k+1}) \).

By induction we conclude that for all \( k \in \mathbb{N} \), for any \( \omega_k \in \text{bd} \, B(\omega_{k-1}, P'_{k-1}) \), there exists an open set \( H(z_k) \subset B(z_k, P_k) \) such that \( \Phi|_{H(z_k)} \) is diffeomorphism onto \( B(\omega_k, P'_k) \). Since \( z_k \) are chosen arbitrarily, one can take the union of the open sets \( H(z_k) \) and \( B(\omega_k, P_k) \) to define

\[ \mathcal{O}_k(\omega_0) := \mathcal{O}_{k-1}(\omega_0) \cup \bigcup_{\omega \in \text{bd} \, \mathcal{O}_{k-1}(\omega_0)} B(\omega, P'_k), \quad k > 0, \mathcal{O}_0(\omega_0) := B(\omega_0, P), \]

and

\[ \mathcal{O}_k(\omega_0) := \mathcal{O}_{k-1}(\omega_0) \cup \bigcup_{z \in \text{bd} \, \mathcal{O}_{k-1}(z)} H(z), \quad k > 0, \mathcal{O}_0(\omega_0) := H(z). \]

From the existence of local inverses, for any \( \omega \in \mathcal{O}_k(\omega_0) \), there exist some \( z \in \mathcal{O}_k(\omega) \) such that

\[ \Phi(z) = \omega. \]

At this point we have shown that for any \( \omega \in \mathcal{O}_k(\omega_0) \), there exist a \( z \in \mathcal{O}_k(\omega_0) \) satisfying \( \Phi(z) = \omega \). To come up with the expressions for \( r \) and \( \epsilon \), all that remains is to now compute bounds on \( \mathcal{O}_k(\omega_0) \) and \( \mathcal{O}_k(\omega_0) \). For any given \( r < \frac{1}{MK} \), there exists a \( k^* \) such that \( \mathcal{O}_{k^*}(\omega_0) \subset B(z_0, P_{k^*}) \), where

\[ P_{k^*} := P_0 + \sum_{i \in [k^*]} P_k, \]

and

\[ P_{k^*} > r. \]
From (3.7) substituting $P_k$ in (3.8) we get

$$P_{k^*} = \frac{1 - (1 - \lambda)^{k^*}}{MK}.$$  

(3.9)

Correspondingly,

$$P_{k^*}' := P_0' + \sum_{i \in [K']} P_{k^*}' = \frac{1}{M^2K} \left( \frac{(1 - \lambda)(1 - (1 - \lambda)^{2k^*})}{2 - \lambda} \right).$$

It is easy to show that $\mathcal{O}_k(\omega) = \mathcal{B}(\omega, P_{k^*})$. From (3.9) substitute $(1 - \lambda)^{k^*} = 1 - MKP_{k^*}$ to get

$$P_{k^*}' = \frac{1 - \lambda}{2 - \lambda} \left( \frac{1 - (1 - MKP_{k^*})^2}{M^2K} \right).$$

For any given $\varepsilon > 0$, there exists a $\lambda$ and a $k^*$ such that $r - P_{k^*} \leq \varepsilon$. Therefore,

$$\lim_{\lambda \to 0} P_{k^*}' = \frac{1 - (1 - MKr)^2}{2M^2K} = \frac{r(2 - MKr)}{2M} =: \varepsilon.$$  

(3.10)

This concludes the proof; for any given $r < 1/MK$, and $\varepsilon$ satisfying (3.10), for all $\omega \in \mathcal{B}(\omega_0, \varepsilon)$ there exists a $\zeta \in \mathcal{B}(\omega_0, r)$ satisfying $\Phi(\zeta) = \omega$. \hfill \square

**Example.** In order to understand the non-uniqueness of $z$ let us look at a common example. Let $\mathbb{R}^2 \ni (r, \theta) \mapsto f(r, \theta) := (r \cos \theta, r \sin \theta)$. Clearly,

$$Df(r, \theta) = \begin{pmatrix} \cos \theta & -r \sin \theta \\ \sin \theta & r \cos \theta \end{pmatrix}$$

is full rank for all $(r, \theta) \in \mathbb{R}^2 \setminus (0, \theta)$, and is therefore invertible. However, no global inverse exists for $f$. For a point $p = (r \cos \theta, r \sin \theta), \theta \in [0, 2\pi[$, the preimage of $p$ is

$$f^{-1}(p) = \{(r, \theta + 2n\pi), n \in \mathbb{Z}\}.$$  

However, for any given open set $\mathcal{O}(p)$ around $p$ there exists an open set $\mathcal{O}(x)$ around $(r, \theta + 2n\pi)$ for a particular $n$ such that there exists a solution to $f(x') = p'$ in $\mathcal{O}(x)$ for all $p' \in \mathcal{O}(p)$.

**Remark.** Although the bounds in Proposition 3.3 are presented in context of the ImFT, the same can be applied to the IFT as well. Indeed, replacing $\Phi$ with $f$ and accordingly replacing $L, M, K$ with $L_f, M_f, K_f$ we arrive at the following result:

**Proposition 3.4.** Let $U \subset \mathbb{R}^n$ be a nonempty open set and $f : U \to \mathbb{R}^n$, satisfy the assumptions of Theorem 2.1 and $L_f, M_f, K_f$ be as defined in Proposition 2.2. Then for any $0 \leq r \leq \min\{1/M_fK_f, R_f\}$, and $\varepsilon = r(2 - rM_fK_f)/2M$, for all $y \in \mathcal{B}(y_0, \varepsilon)$, there exist (not necessarily unique) $x \in \mathcal{B}(x_0, r)$ such that

$$f(x) = y.$$  

We omit the proof of Proposition 3.4 for brevity.

**Remark 3.5.** The bounds given in Proposition 3.4 are double than those given by Proposition 2.2. However, this comes at the cost of the loss of uniqueness of $x$.

Propositions 2.2 and 3.2 consider an unconstrained variation around $y_0$ and $\omega_0$ respectively. However, often we are interested to know how the preimage of $y$ under $f$ varies when $y$ is varied about $y_0$ along a particular direction. We define this formally as follows: Let $W \subset \mathbb{R}^n$ be a subspace, and $f : U \to \mathbb{R}^n$ be a $C^2$ map. For a given $r > 0$, find an $\varepsilon_W > 0$ such that for all $\tilde{y} \in W$ satisfying $||\tilde{y}|| < \varepsilon_W$, there exists an $x \in \mathcal{B}(x_0, r)$ satisfying $f(x) = y_0 + \tilde{y}$. 
PROPOSITION 3.6. Let $U \subset \mathbb{R}^n$ be a nonempty open set and $f: U \rightarrow \mathbb{R}^n$ satisfy the assumptions of Theorem 2.1 and $L_f, M_f, K_f, P_f$ and $R_f$ be as defined in Proposition 2.2. Define

$$B_W(x_0, r) := \{x \in \mathbb{R}^n \mid x - x_0 \in W \text{ and } \|x - x_0\| < r\}$$

and

$$M_f^W := \sup \{\|Df^{-1}(y)\| \mid y \in B_W(0, 1)\}.$$ 

Then (3.6-a) for any fixed $\lambda \in [0, 1]$ there exists a set $H_f^W \subset B \left( x_0, \frac{4}{M_f K_f} \right) \subset \mathbb{R}^n$ such that $f$ maps $H_f^W$ diffeomorphically onto $B_W \left( y_0, \frac{\lambda (1 - \lambda)}{M_f^p M_f K_f} \right)$.

(3.6-b) for any $0 \leq r \leq \min \left\{ \frac{1}{M_f K_f}, R_f \right\}$, and $\epsilon_W = \frac{r (2 - r M_f K_f)}{2 M_f^p}$, for all $y \in B_W(y_0, \epsilon_W)$, there exist (not necessarily unique) $x \in B(x_0, r) \subset \mathbb{R}^n$ such that $f(x) = y$.

Proof. Without loss of generality set $(x_0, y_0) = (0, 0)$ and define

$$x \mapsto g_y(x) := Df^{-1}(0)(y + Df(0)x - f(x)).$$

Then we have

$$\|g_y(x)\| \leq \|Df^{-1}(0)y\| + \|Df^{-1}(0) \cdot (Df(0)x - f(x))\|$$

$$\leq \|Df^{-1}(0)y\| + \|Df^{-1}(0)\| \|Df(0)x - f(x)\|$$

$$\leq M_f^W \|y\| + M_f K_f \|x\|^2$$

for all $y \in W$.

Therefore, for all $y \in B_W \left( y_0, \frac{\lambda (1 - \lambda)}{M_f^p M_f K_f} \right)$ we have

$$\text{cl} \, B \left( x_0, \frac{\lambda}{M_f K_f} \right) \ni x \mapsto g_y(x) \in \text{cl} \, B \left( x_0, \frac{\lambda}{M_f K_f} \right)$$

and

$$\|Dg_y(x)\| < 1$$

for all $x \in B \left( x_0, \frac{\lambda}{M_f K_f} \right)$.

Hence, from the Banach Fixed Point Theorem [Zei95a, Theorem 1.A], there exists a unique $x \in B \left( x_0, \frac{\lambda}{M_f K_f} \right)$ such that $g_y(x) = x$, which implies $f(x) = y$. This proves (3.6-a).

Proposition (3.6-b) follows from arguments similar to Proposition 3.3 and is omitted. □

The bounds derived in Propositions 2.2 and 3.2 are dependent on the second order derivatives of the underlying maps. This needs $f, \Phi \in \mathcal{C}^2$. In particular we use the bounds on $D^2 \Phi(x, y)$ to get an estimate of $D\Phi(x, y)$ and $(D\Phi(x, y))^{-1}$ in a neighborhood of $(x_0, y_0)$. We can relax this requirements by alternatively using bounds on the first order derivatives.

PROPOSITION 3.7. Let $\Phi$ be defined as Proposition 3.2 and $M := \|D\Phi(0, 0)^{-1}\|$. For a given $r > 0$, define

$$L(r) := \sup \{\|D\Phi(x_0, y_0) - D\Phi(x, y)\| \mid (x, y) \in B((x_0, y_0), r)\}.$$ 

Then for all $r$ such that $L(r) < \frac{1}{3M}$ and $\epsilon = \frac{r(1 - M L(r))}{3M}$, for all $\omega \in B((\omega_0, \epsilon), \omega_0 = \Phi(x_0, y_0)$, there exists $(x(\omega), y(\omega)) \in B((x_0, y_0), r)$ such that

$$\Phi(x(\omega), y(\omega)) = \omega$$

for all $\omega \in B((\omega_0, \epsilon)$. 


Proof. Without loss of generality set \((x_0, y_0) = (0, 0), \omega_0 = 0\). Define
\[
(3.11) \quad (x, y) \mapsto g_\omega(x, y) := \mathcal{D}\Phi^{-1}(0, 0) (\omega + \mathcal{D}\Phi(0, 0)(x, y) - \Phi(x, y)) .
\]
From (3.11)
\[
g_\omega(x^*, y^*) = (x^*, y^*) \implies \Phi(x^*, y^*) = \omega.
\]
Moreover,
\[
\|g_\omega(x, y)\| = \left\|\mathcal{D}\Phi^{-1}(0, 0) (\omega + \mathcal{D}\Phi(0, 0)(x, y) - \Phi(x, y))\right\|
\leq \left\|\mathcal{D}\Phi^{-1}(0, 0)\right\| (\|\omega\| + \|\mathcal{D}\Phi(0, 0)(x, y) - \Phi(x, y)\|)
\leq \left\|\mathcal{D}\Phi^{-1}(0, 0)\right\| \left(\|\omega\| + \int_0^1 \|\mathcal{D}\Phi(0, 0) - \mathcal{D}\Phi(tx, ty)\| \cdot (x, y) \|dt\right)
\leq M(\|\omega\| + rL(r)).
\]
For all \(\|\omega\| \leq \epsilon, g_\omega\) maps \(\epsilon\) \(\mathcal{B}(r, (0, 0))\) to \(\epsilon\) \(\mathcal{B}(r, (0, 0))\), and for all \((x, y) \in \mathcal{B}((x_0, y_0), r)\)
\[
\|Dg_\omega(x, y)\| \leq \left\|\mathcal{D}\Phi^{-1}(0, 0)\right\| \|\mathcal{D}\Phi(0, 0) - \mathcal{D}\Phi(x, y)\| < 1.
\]
Therefore, from the Banach Fixed Point Theorem [Zei95a, Theorem 1.4], for all \(\omega \in \mathcal{B}(\omega_0, \epsilon)\) there exist a unique \((x(\omega), y(\omega)) \in \mathcal{B}((x_0, y_0), r)\) such that
\[
g_\omega(x(w), x(w)) = (x(\omega), y(\omega)) \implies \Phi(x(\omega), y(\omega)) = \omega,
\]
thereby concluding the proof. \(\square\)

\S 3.1. Computing estimates for the ImFT.\) Proposition 3.2 serves as a precursor to the actual estimates for the ImFT in the following sense: We cast the ImFT in a larger space by adding extra equations to convert it to an IFT premise, and then obtain bounds on the relevant objects. These bounds are then converted back to those relevant for the ImFT. We can now use these bounds to get estimates on \(\mathcal{O}(x_0)\) and \(\mathcal{O}(x_0)\) as defined in Theorem 3.1.

**Proposition 3.8.** Let \(U \subset \mathbb{R}^n, V \subset \mathbb{R}^m\) be nonempty open sets and \(f : U \times V \rightarrow \mathbb{R}^m\) be as defined in Theorem 3.1. Suppose that \(L, M, N, K, P\) and \(P'\) are as defined in Theorem 3.2. Let \(A_1 : \mathbb{R}^n \rightarrow \mathbb{R}^n, A_2 : \mathbb{R}^m \rightarrow \mathbb{R}^m\) be given isomorphisms. Let \((x_0, y_0) \in U \times V\) and \(f(x_0, y_0) = w_0\), such that \(Df(x_0, y_0)\) is an isomorphism, then whenever
\[
0 < r \leq \frac{P}{2M \|A_1\|} \quad \text{and} \quad \epsilon = \frac{1}{\|A_2\|} \left(\frac{P'}{2M} - \|A_1\| r\right),
\]
there exists a map \(g : \mathcal{B}(x_0, r) \times \mathcal{B}(w_0, \epsilon) \rightarrow \mathbb{R}^m\) such that
\[
f(x, g(x, w)) = w \quad \text{for all } (x, w) \in \mathcal{B}(x_0, r) \times \mathcal{B}(w_0, \epsilon).
\]

**Proof.** Define \(\Phi : U \times V \rightarrow \mathbb{R}^n \times \mathbb{R}^m\) as defined in Proposition 3.2, which asserts that there exists an open set \(H \subset \mathcal{B}((x_0, y_0), P)\) such that \(\Phi|_H\) is a diffeomorphism onto \(\mathcal{B}(w_0, P')\), and \(\omega_0 = \Phi(x_0, y_0)\). Therefore, for all \((u, v) \in \mathcal{B}(\omega_0, P')\) there exists a unique \((x, y) \in H\) satisfying
\[
\Phi(x, y) = (A_1x, A_2f(x, y)) = (u, v).
\]

From the above equation, we have \(u = A_1x\) and \(v = A_2f(x, y) = A_2w\). For any given
\[
(3.12) \quad 0 < r \leq \frac{P'}{\|A_1\|} \quad \text{and} \quad \epsilon = \frac{P'}{\|A_2\|} - \|A_1\| r,
\]
for all \(x \in \mathcal{B}(x_0, r)\) and \(\omega \in \mathcal{B}(w_0, \epsilon)\), \((A_1x, A_2w) = (u, v) \in \mathcal{B}(\omega_0, P')\). In other words there exists an \((x, y) \in \mathcal{B}((x_0, y_0), P)\) such that
\[
\Phi(x, y) = (A_1x, A_2f(x, y)) = (A_1x, A_2w) = (u, v).
\]

Therefore, for a given \(r, \epsilon\) satisfying (3.12), for all \(x \in \mathcal{B}(x_0, r)\) and \(w \in \mathcal{B}(w_0, \epsilon)\), there exists a unique \(y\) satisfying \((x, y) \in \mathcal{B}((x_0, y_0), P)\) such that \(f(x, y) = w\). Setting \(y = g(x, w)\) completes the proof. \(\square\)
REMARK. Proposition 3.8 gives bounds on neighborhoods around $x_0$ and $u_0$ on which Theorem 3.1 holds. The neighborhood around $u_0$ is dependent upon the neighborhood chosen around $x_0$ as seen from (3.12). As to Proposition 2.2, the bounds are dependent on the first order derivatives $D_x f$ and $D_y f$ evaluated at the operating point $(x_0, y_0)$ and bounds on the second order derivative $D^2 f(x, y)$ over domain of interest (determined by $R$).

§ 3.2. Improving estimates by a careful choice of $A_1$ and $A_2$. The estimates given in Proposition 3.8 are dependent upon the choice of $A_1$ and $A_2$. As seen from (3.12), the bounds on $r$ and $\epsilon$ are explicitly dependent on $A_1$ and $A_2$. Moreover, $L$, $M$, $K$ as defined in Proposition 3.2 are also dependent upon $A_1$ and $A_2$ as follows

$$L = \|A_1\| + \|A_2\| (L_x + L_y), \quad M = \|A_1^{-1}\| (1 + M_y L_x) + \|A_2^{-1}\| M_y, \quad K = \|A_2\| K_f.$$ 

Substituting these in (3.6) we write

$$P = \min \left\{ \frac{1}{2 \|A_2\| K_f \left( \|A_1^{-1}\| (1 + M_y L_x) + \|A_2^{-1}\| M_y \right),} R \right\}. \tag{3.13}$$

REMARK. Note that, both $P$ and $K_f$ are dependent on $R$, making it difficult to optimize $P$ with respect to $A_1$ and $A_2$. However, if $R$ is chosen sufficiently large, the minimum will be decided by the first term in (3.13). This is also justifiable: choosing a smaller $R$ limits our estimates to a ball of radius $R$. Further, $K_f$ is dependent on $R$ and increases monotonically with $R$ i.e., for $R_1 \geq R_2$ we have $K_f (R_1) \geq K_f (R_2)$ because

$$K_f (R) = \sup \{ \|D^2 f (x, y)\| \mid (x, y) \in \mathcal{B}((x_0, y_0), R) \}.$$ 

Therefore the term

$$\frac{1}{2 \|A_2\| K_f \left( \|A_1^{-1}\| (1 + M_y L_x) + \|A_2^{-1}\| M_y \right)}$$

decreases with $R$. Moreover, from continuity of $D^2 f$ there exists an $R'$ such that for all $R > R'$

$$\frac{1}{2 \|A_2\| K_f \left( \|A_1^{-1}\| (1 + M_y L_x) + \|A_2^{-1}\| M_y \right)} \leq R.$$

Hence, for further discussion we will assume $R \gg 1$ so as to be ignored from the analysis. An upper bound on $r$ in Proposition 3.8 is given by

$$\frac{P'}{\|A_1\|} = \frac{1}{4K_f \|A_2\| \|A_1\| \left( \|A_1^{-1}\| (1 + M_y L_x) + \|A_2^{-1}\| M_y \right)^2}. \tag{3.14}$$

We now analyze the effect of the terms in the denominator to arrive at the further simplifications.

Defining $\kappa_{A_1} := \|A_1\| \|A_1^{-1}\|$, $\kappa_{A_2} := \|A_2\| \|A_2^{-1}\|$ as the condition number of $A_1$ and $A_2$ respectively, and $\Lambda := (\|A_2\| / \|A_1\|)$, we rewrite (3.14) as

$$\frac{P'}{\|A_1\|} = \frac{1}{4K_f} \left( \kappa_{A_1} (1 + M_y L_x) \Lambda^2 + \frac{\kappa_{A_2} M_y}{\Lambda^2} \right) \quad \text{for } R \gg 1.$$ 

Since $\kappa_{A_1}$, $\kappa_{A_2}$, $\Lambda$ can be chosen independently with $\kappa_{A_1} \geq 1$ and $\kappa_{A_2} \geq 1$, the optimal choice for $\kappa_{A_1}$, $\kappa_{A_2}$, $\Lambda$ that maximizes $\frac{P'}{\|A_1\|}$ is

$$\kappa_{A_1} = \kappa_{A_2} = 1.$$
and

\[ \Lambda^* = \left( \frac{M_y}{1 + M_y L_x} \right). \]

Substituting these quantities back in various expressions, we have

\[ P^* := \frac{1}{4K_f M_y^2}, \quad \frac{P^*}{2M^* \|\Lambda^*\|} = \frac{1}{16K_f M_y(1 + M_y L_x)}, \]

and

\[ \frac{P^*}{2M^* \|\Lambda^1\|} = \frac{1}{16K_f M_y^2}. \]

Collecting these quantities, the optimized estimates are then given by the following proposition.

**Proposition 3.9.** Let \( U \subset \mathbb{R}^n, V \subset \mathbb{R}^m \) be nonempty open sets. Let \( f : U \times V \rightarrow \mathbb{R}^n \) satisfy the assumptions of Theorem 3.1. Define

\[ L_x := \|D_x f(x_0, y_0)\|, \quad M_y := \|\left(D_x f(x_0, y_0)\right)^{-1}\|, \]

\[ K_f := \sup\{\|D^2 f(x, y)\| : (x, y) \in U \times V\}. \]

Define \( P^* := \frac{1}{4K_f M_y^2}. \) For given

\[ 0 \leq r \leq \frac{1}{16K_f M_y(1 + M_y L_x)} \]

and

\[ \epsilon = \frac{1}{16K_f M_y^2} \frac{(1 + M_y L_x)r}{M_y}, \]

there exists a unique map \( g : B(x_0, r) \times B(u_0, \epsilon) \rightarrow \mathbb{R}^m \) with \( (x, g(x, w)) \in B((x_0, y_0), P^*) \) and satisfying

\[ f(x, g(x, w)) = w \quad \text{for all } x \in B(x_0, r), w \in B(u_0, \epsilon). \]

These bounds can be represented graphically as shown in Figure 3.

**Remark.** Our objective is to come up with bounds that are easier to compute rather than tighter bounds. In this light, we relax several inequalities and choose rather conservative versions of them. For instance at several places submultiplicativity of matrix norms is employed to bound \( \|AB\| \) by \( \|A\| \|B\| \). If exact expressions of \( A \) and \( B \) are known, then one can come up with better estimates of \( \|AB\| \). Likewise, the full potential of the matrices \( A_1 \) and \( A_2 \) are not utilized in the general case since that will involve solving several non-trivial optimization routines on the set of linear isomorphisms.
Remark. As in Proposition 3.3 if our objective is to show the existence of \( u \) satisfying \( f(x, u) = w \) in a neighborhood of \((x_0, u_0)\)and \( w_0 \), then one can relax the uniqueness requirement and come up with the improved estimates, as the following proposition demonstrates.

**Proposition 3.10.** Let \( U, V \) and \( W \) be as defined in Theorem 3.1, and let \( f : U \times V \rightarrow W \) satisfy the assumptions of Theorem 3.1. Let \( L_x, M_y, K_f \) be as defined in Proposition 3.2. Define

\[
P^* := \frac{1}{2K_f M_y}.
\]

Then for a given

\[
0 \leq r \leq \frac{1}{8K_f M_y (1 + M_y L_x)}
\]

and

\[
e^*(r) = \frac{1}{8K_f M_y^2} - \frac{(1 + M_y L_x) r}{M_y},
\]

for each \( x \in B(x_0, r) \) and \( w \in B(w_0, e^*(r)) \), there exists (not necessarily unique) \( u \), satisfying \((x, u) \in B((x_0, u_0), P^*)\) and

\[
f(x, u) = w.
\]

The proof follows from arguments similar to Proposition 3.8 and hence omitted.

§ 3.3. An alternate set of estimates for the Implicit Function Theorem. The proof of Theorem 3.1 involved extending the function to an invertible map by adding an extra equation. This leads to an increase in the dimension of the problem and hence provides coarser bounds. An alternate proof method that does not involve invoking inverse function theorem, intuitively speaking, shall give better bounds. Using this as the motivation, we now look at a more commonly used definition of the Implicit Function Theorem as follows.

**Theorem 3.11.** Implicit Function Theorem (classical) [Zei93, Theorem 4.B]: Let \( X \subset \mathbb{R}^n \), \( Y \subset \mathbb{R}^m \) be nonempty open sets. Let \( X \times Y \ni (x, y) \mapsto f(x, y) \in \mathbb{R}^m \) be a \( \mathcal{C}^1 \) map. Suppose \((x_0, y_0) \in X \times Y\) is such that \( D_y f(x_0, y_0) \) is invertible. Then for any neighborhood \( O(y_0) \) of \( y_0 \), there is a neighborhood \( O(x_0) \) of \( x_0 \) and a unique continuous map \( g : O(x_0) \rightarrow O(y_0) \) such that

1. \((3.11-a) \ g(x_0) = y_0.\)
2. \((3.11-b) \ f(x, g(x)) = f(x_0, y_0) \) for all \( x \in U,\)
3. \((3.11-c) \ g \) is continuous at point \( x_0.\)

Remark. The classical version of the implicit function theorem can be retrieved from Theorem 3.1 by keeping \( w = w_0 \) fixed and denoting \( g(x, u_0) \) as \( g(x) \).

Proof. Without loss of generality assume \( x_0 = 0, y_0 = 0 \) and \( w_0 := f(x_0, y_0) = 0 \). Let \( T := D_y f(x_0, y_0) \). Define \( X \times Y \ni (x, y) \mapsto \varphi(x, y) := y - T^{-1}(f(x, y) - w_0) \in \mathbb{R}^m \), then \( \varphi(x, y) = y \) if and only if \( f(x, y) = w_0 \). The proof revolves around showing that for each \( x \in U, \varphi(x, \cdot) \) has a fixed point. From expression of \( \varphi \) we have

\[
D_y \varphi(x, y) = I - T^{-1} D_y f(x, y).
\]

Since \( D_y \varphi(x_0, y_0) = 0 \), there exists an \( r > 0 \) such that

\[
\frac{||\varphi(x_0, y) - \varphi(x_0, y_0)||}{||y - y_0||} \leq \frac{1}{2} \text{ for all } y \in \text{cl } B(y_0, r).
\]
On rearranging,

\[(3.15) \quad \| \varphi(x_0, y) - y_0 \| \leq \frac{r}{2} \quad \text{for all } y \in \text{cl} B(y_0, r). \]

For each \( x \) set \( m(x) := \max \{ \| \varphi(x, y) - \varphi(x_0, y) \| \mid y \in \text{cl} B(y_0, r) \} \). Since \( m(x) \) is continuous and \( m(x_0) = 0 \), there exists an \( \epsilon \) such that

\[(3.16) \quad m(x) < \frac{r}{2} \quad \text{for all } x \in B(x_0, \epsilon). \]

For each \( x \in B(x_0, \epsilon) \) we have

\[
\| \varphi(x, y) - y_0 \| \leq \| \varphi(x, y) - \varphi(x, y_0) \| + \| \varphi(x, y_0) - y_0 \| \\
< \frac{r}{2} + \frac{r}{2} = r.
\]

Thus, for each \( x \in B(x_0, \epsilon) \), \( \varphi(x, \cdot) \) maps

\[\text{cl} B(y_0, r) \ni y \mapsto \varphi(x, y) \in \text{cl} B(y_0, r).\]

Hence by Brouwer’s Fixed Point Theorem [Zei95a, Theorem 1.B], for each \( x \in B(x_0, \epsilon) \), there exists a \( g(x) \in \text{cl} B(y_0, r) \) such that \( \varphi(x, g(x)) = g(x) \) or consequently

\[f(x, g(x)) = u_0.\]

**Remark.** Although Theorem 3.11 requires the map \( f \) to be simply continuous on \( X \times Y \), and differentiable at \((x_0, y_0)\). For estimating the bounds on the neighborhood we require the bounds on the second order derivatives. Hence, we additionally assume \( f \in \mathcal{C}^2 \) and that all partial derivatives up to second order exist and are continuous.

Inspecting the proof of Theorem 3.11, one can see that (3.15) and (3.16) plays a crucial role in showing existence of the implicit map \( g \). Based on these two inequalities, the estimates on neighborhoods for Theorem 3.11 are given by the following proposition.

**Proposition 3.12.** Let \( X \subset \mathbb{R}^n, Y \subset \mathbb{R}^m \) be nonempty open sets. Let \( f : X \times Y \rightarrow \mathbb{R}^m \) be a \( \mathcal{C}^2 \) map and satisfies the conditions of Theorem 3.1. Define

\[L_x := \| D_x f(x_0, y_0) \|, \quad M_y := \left\| (D_y f(x_0, y_0))^{-1} \right\|\]

and for given \( R' > 0 \), \( R > 0 \) define

\[K_{y} := \sup \{ \left\| D_y^2 f(x_0, y) \right\| \mid y \in B(y_0, R) \}, \]

\[K_{xy} := \sup \{ \left\| D_y D_x f(x, y) \right\| \mid (x, y) \in B((x_0, y_0), R') \}, \]

and \( K_{xy} := \sup \{ \left\| D_x D_y f(x, y) \right\| \mid (x, y) \in B((x_0, y_0), R') \}. \)

Then for all \( r < P, P : = \min \left\{ \frac{1}{2M_x K_{xy} R}, R \right\}, \) and \( \epsilon \) satisfying

\[M_y (L_x + K_{xy} r + K_{xx} \epsilon) \epsilon < r/2,\]

for all \( x \in B(x_0, \epsilon) \) there exists a \( g(x) \in B(y_0, r) \) satisfying \( f(x, g(x)) = 0. \)

**Proof.** From the proof of Theorem 3.11, we can see (3.15) and (3.16) serve as a starting point for estimating the bounds. From (3.15) we have

\[
\| \varphi(x_0, y) - y_0 \| \leq \frac{r}{2}.
\]

From the Fundamental Theorem of Calculus [Rud76, Theorem 6.24],

\[
\varphi(x_0, y) = \varphi(x_0, y_0) + \int_0^1 D_y \varphi(x_0, sy) y \, ds \\
= \int_0^1 D_y \varphi(x_0, sy) y \, ds.
\]
Similarly,
\[ D_y \varphi(x_0, y) = \int_0^1 D_y^2 \varphi(x_0, sy) y \, ds \]
\[ = -T^{-1} \int_0^1 D_y^2 f(x_0, sy) y \, ds \]
\[ \implies \left\| D_y \varphi(x_0, y) \right\| \leq M_y K_{yy}, \]
where
\[ M_y := \| T^{-1} \| \quad \text{and} \quad K_{yy} := \sup \{ \| D_y^2 f(x_0, y) \| \mid B(y_0, R) \}. \]
Substituting these in expression of \( D_y \varphi(x_0, y) \), we get
\[ \frac{\| \varphi(x_0, y) - y_0 \|}{\| y - y_0 \|} \leq \left\| D_y \varphi(x_0, y) \right\| \leq \frac{1}{2} \quad \text{for all} \quad y \in B(y_0, P), \]
with
\[ P := \min \left\{ \frac{1}{2M_y K_{yy}}, R \right\}. \]
We use (3.16) to estimate bounds on \( \epsilon \). From (3.16), for a given \( r < P \), we have
\[ \max \{ \| \varphi(x, y) - \varphi(x_0, y) \| \mid y \in B(y_0, r) \} < r/2 \quad \text{for all} \quad x \in B(x_0, \epsilon). \]
Using the Fundamental Theorem of Calculus we write
\[ D_x \varphi(x, y) = D_x \varphi(x_0, y_0) + \int_0^1 \int_0^1 \left( D_x^2 \varphi(sx, ty) - D_x D_y \varphi(sx, ty) \right) \, ds \, dt, \]
and from triangle inequality and submultiplicative property of induced norms we have
\[ \| D_x \varphi(x, y) \| \leq \| D_x \varphi(x_0, y_0) \| + \int_0^1 \int_0^1 \left( \| D_x^2 \varphi(sx, ty) \| \| x \| + \| D_x D_y \varphi(sx, ty) \| \| y \| \right) \, ds \, dt. \]
Substituting for \( L_x, M_y, K_{xx}, K_{yy}, \) and \( K_{xy} \) in above equation we get
\[ \| D_x \varphi(x, y) \| \leq M_y \left( L_x + K_{xy} \| y \| + K_{xx} \| x \| \right), \]
and
\[ \| \varphi(x, y) - \varphi(x_0, y_0) \| \leq M_y \left( L_x + K_{xy} \| y \| + K_{xx} \| x \| \right) \| x \|. \]
From (3.16), we have
\[ M_y (L_x + K_{xy} r + K_{xx} \epsilon) \leq r/2. \]
For a given \( r < P \), one can solve inequality (3.17) to get a respective \( \epsilon \). Then for every \( x \in B(x_0, \epsilon) \) there exists a \( g(x) \in B(y_0, r) \) such that \( f(x, g(x)) = 0 \), which completes the proof.

§ 3.4. Comparison between the bounds obtained from Theorem 3.1 and 3.11. In this article we have presented two different set of bounds on the neighborhoods in Implicit Function Theorem. The first set of the bounds obtained in Proposition 3.9 involves adding an extra equation and thus increases the dimension while the Proposition 3.12 relies on a direct proof to come up with estimates. From Proposition 3.12, we have for a given \( r \leq P \), with \( P = \min \{ 1/(2K_{yy}M_y), R \} \) there exists an \( \epsilon \) satisfying
\[ M_y (L_x + K_{xy} r + K_{xx} \epsilon) \leq r/2, \]
such that for all \( x \in B(x_0, \epsilon) \) there exists a \( g(x) \in B(y_0, r) \) such that
\[ f(x, g(x)) = 0. \]
Now, set $r = 1/(2K_{yy}M_y)$ and $\epsilon = 1/(16K_f M_y(1 + M_yL_x))$ and substitute in (3.18), then we have
\begin{align*}
M_y(L_x + K_{xy}r + K_{xx}\epsilon)\epsilon
&= \left( L_x + \frac{K_{xy}}{2K_{yy}M_y} + \frac{K_{xx}}{16K_f M_y(1 + M_yL_x)} \right) \frac{M_y}{16K_f M_y(1 + M_yL_x)} \\
&\leq \left( L_x + \frac{K_{xy}}{2K_{yy}M_y} + \frac{1}{16M_y} \right) \frac{M_y}{16K_f M_y(1 + M_yL_x)} \\
&= \frac{1}{4K_{yy}M_y} \left( \frac{M_yL_x + 1/16 + K_{xy}/2K_{yy}}{4(1 + M_yL_x)} \right).
\end{align*}
(3.19)

Now, if $M_y, L_x, K_{xy}, K_{yy}$ are such that
\begin{align*}
\frac{M_yL_x + 1/16 + K_{xy}/2K_{yy}}{4(1 + M_yL_x)} < 1,
\end{align*}
(3.20)
then $\epsilon$ satisfies (3.18). Since right hand side of (3.18) is an increasing function, and for $\epsilon = 1/(16K_f M_y(1 + M_yL_x))$, it is strictly less than $r/2$. Therefore there exists an $\epsilon^* > \epsilon$ satisfying (3.18) for $r = P$. One can further relax (3.20) as follows
\begin{align*}
\frac{M_yL_x + 1/16 + K_{xy}/2K_{yy}}{4(1 + M_yL_x)} \leq \frac{1}{4} + \frac{K_{xy}}{8K_{yy}},
\end{align*}
to get following condition on $K_{xy}$ and $K_{yy}$:
\begin{align*}
K_{xy} \leq 6K_{yy}.
\end{align*}
(3.21)

Hence, for functions that satisfying (3.20) and (3.21), the bounds given by Proposition 3.12 are larger than those given by Proposition 3.9.

§4. Applications

Solving system of equations is ubiquitous in computational mathematics. The problem may often be challenging due to uncertainty in the parameters and coefficients associated with the set of equations. Power Flow studies and Optimal Power Flow studies require us to calculate operating voltages and currents by balancing the load demand and the power generated. Integration of renewable sources like solar and wind energy in the power system network leads to uncertainty in the generated supply. Due to this uncertain supply of renewable sources, it is difficult to ensure that there will be sufficient power generation to meet the power demand while adhering to the stable operation limits of the power system network. As described by [DNT17], the AC Power Flow equations are written as follows
\begin{align*}
\text{Re}\left( V_i Y_{i0} V_0 + \sum_{k=1}^{n} V_i Y_{ik} V_k \right) &= p_i, \quad \text{for all } i \in \text{PQ}, \\
\text{Im}\left( V_i Y_{i0} V_0 + \sum_{k=1}^{n} V_i Y_{ik} V_k \right) &= q_i, \quad \text{for all } i \in \text{PQ}, \\
\text{Re}\left( V_i Y_{i0} V_0 + \sum_{k=1}^{n} V_i Y_{ik} V_k \right) &= p_i, \quad \text{for all } i \in \text{PV}, \\
|V_i|^2 &= v_i^2, \quad \text{for all } i \in \text{PV},
\end{align*}
(4.1)
where $V_i$ denotes the complex voltage phasor, $p_i$, $q_i$ denotes the active and reactive power injection at the node $i$, and $Y$ denotes the admittance matrix. PV denotes the set of generator buses, and PQ denotes the set of load bus. $v_i$ denotes the root mean square (rms) voltage at the $i^{th}$ bus. $V_0$ denotes the reference slack bus and is kept fixed at 1 + j0 per unit magnitude, where $j = \sqrt{-1}$. Power Flow as given by (4.1) is quadratic. Further, for stable
operation of the power system network, one needs to maintain the bus voltage magnitude within prescribed limits. This imposes a quadratic constraint on (4.1). Thus, Power Flow equations involve solving a Quadratically Constrained Quadratic Programming (QCQP) which is, in general, an NP-Hard problem. This serves as the motivation to come up with apriori estimates on the maximum allowable variations in the power generated and expected demand at bus nodes, for which a stable operation of the power system (i.e., maintaining the operating bus voltages within the allowable limits) is feasible. In this direction, we first study the robustness of the solutions of the generalized nonlinear equation over a given constraint set.

§ 4.1. Robustness of Solutions of Nonlinear Equations. Let \( X \subset \mathbb{R}^n \) and \( U \subset \mathbb{R}^m \) be nonempty open sets and \( f : X \times U \rightarrow \mathbb{R}^m \) be a \( \mathcal{C}^2 \) map. A system of nonlinear equations in \((x, u)\) on \( X \times U \) is defined as

\[
(4.2) \quad f(x, u) = 0
\]

(4.2) is solvable if there exists an \((x_0, u_0) \in X \times U\) such that \(f(x_0, u_0) = 0\).

**Definition.** Let \( r > 0 \) such that \( B(x_0, r) \subset X \). (4.2) is said to be robustly solvable on \( B(x_0, r) \), if there exists an \( \epsilon > 0 \), such that for all \( u' \in B(u_0, \epsilon) \subset U \), there exists an \( x' \in B(x_0, r) \) satisfying \( f(x', u') = 0 \). Moreover, for a given \( r, \epsilon \) is called the robustness margin for (4.2).

**Theorem 4.1.** Let \( X \subset \mathbb{R}^n \), \( U \subset \mathbb{R}^m \) be nonempty set and \( f : X \times U \rightarrow \mathbb{R}^m \) be a \( \mathcal{C}^2 \) map. Let \((x_0, u_0) \in X \times U\) be such that \(f(x_0, u_0) = 0\) and \( D_x f(x_0, u_0) \) is nonsingular. Define

\[
L_u := \|D_u f(x_0, u_0)\| \quad M_x := \|D_x f(x_0, u_0)\|^{-1}
\]

and for a given \( R > 0 \), \( R' > 0 \) set

\[
K_{xs} := \sup \left\{ \|D_x^2 f(x, u)\| : x \in B(x_0, R) \right\},
\]

\[
K_{us} := \sup \left\{ \|D_u^2 f(x, u)\| : (x, u) \in B((x_0, u_0), R') \right\}, \quad \text{and}
\]

\[
K_{su} := \sup \left\{ \|D_x D_u f(x, u)\| : (x, u) \in B((x_0, u_0), R') \right\}.
\]

Set \( P := \min \left\{ \frac{1}{2n \epsilon}, R \right\} \). Then for all \( r \leq P \), and \( \epsilon \) satisfying

\[
(4.3) \quad M_x (L_u + K_{xs} r + K_{us} \epsilon) \leq r/2,
\]

(4.2) is robustly solvable on \( B(x_0, r) \) and robustness margin is lower bounded by \( \epsilon \).

**Proof.** Since \( D_x f(x_0, u_0) \) is nonsingular, \( f \) satisfies the assumptions of Theorem 3.11. Thus, for any neighborhood \( \mathcal{O}(x_0) \ni x_0 \), there exists an open set \( \mathcal{O}(u_0) \ni u_0 \) and a continuous map \( g : \mathcal{O}(u_0) \rightarrow \mathcal{O}(x_0) \) such that \( f(g(u), u) = 0 \) for all \( u \in \mathcal{O}(u_0) \). This proves the solvability of (4.2) in a neighborhood of \((x_0, u_0)\). Applying Proposition 3.12 on \( f \) as defined in (4.2), for all \( r < P \), there exists an \( \epsilon \) satisfying (4.3) such that

\[
f(g(u), u) = 0 \quad \text{for all} \quad u \in B(u_0, \epsilon).
\]

with \( g(u) \in B(x_0, r) \). For any given \( u' \in B(u_0, \epsilon) \), equating \( x' := g(u') \) ensures existence of an \( x' \in B(x_0, r) \) satisfying \( f(x', u') = 0 \) which proves the second half of the above theorem.}

With the general nonlinear setting defined, we now apply this result on Quadratic Equations as follows:
§4.2. Robustness of Solutions of Quadratic Equations. Let $x \in \mathbb{R}^n$ and $u \in \mathbb{R}^n$, define a function

\begin{equation}
\mathbb{R}^n \ni x \mapsto F(x) := Q(x) + Lx \in \mathbb{R}^n,
\end{equation}

where $L \in \mathbb{R}^{n \times n}$ and $Q(x)$ is a quadratic function with its $i^{th}$ component defined as

\begin{equation}
(Q(x))_i := x^T Q_i x \text{ for all } i \in [n].
\end{equation}

where for all $i \in [n]$, $Q_i \in \mathbb{R}^{n \times n}$ is a symmetric matrix. Then for a given $u$, the system of quadratic equations is then defined as following

\begin{equation}
F(x) = Q(x) + Lx = u,
\end{equation}

subjected to constraint

\begin{equation}
x \in \Omega := \{ x \in \mathbb{R}^n \mid [Ax]_i \leq b_i \text{ for all } i \in [n] \}.
\end{equation}

where $[Ax]_i$ denotes the $i^{th}$ component of the vector $Ax$ and $b_i$ is a scalar. The constraint set $\Omega$ defines a polyhedra as shown in Figure 4. The dashed lines represent the equalities $[Ax]_i = b_i$ and the shaded interior along with the boundary defines the constraint set. It is assumed that the constraints are not redundant and $x$ and $u$ have the same dimension. We are interested in the following problem.

**Problem 4.2.** For a given $u_0$, let there be an $x_0 \in \Omega$ such that $(x_0, u_0)$ solves (4.6). Find the robustness margin i.e., the largest $\epsilon \geq 0$ such that for all $u$ satisfying $\|u_0\|_i - \|u\|_i \leq \epsilon$ for all $i \in [n]$, there exists an $x \in \Omega$ such that $(x, u)$ solves (4.6).

First we utilize the IFT and ImFT to show that a nonzero robustness margin exists and then use the bounds derived for IFT and ImFT to come up with a lower bound on $\epsilon$. Casting (4.6) as (4.2), define a map $f : \mathbb{R}^n \times \mathbb{R}^n \longrightarrow \mathbb{R}^n$ mapping

\begin{equation}
f(x, u) = Q(x) + Lx - u.
\end{equation}

The Jacobian for $f$ at $(x_0, u_0)$ is given by

\begin{equation}
\text{D} f(x_0, u_0) = (\text{D}_x f(x_0, u_0) \quad \text{D}_u f(x_0, u_0)) = (2[x_0^T Q_i] + L \quad I_n)
\end{equation}

where $[x_0^T Q_i]$ denotes the $n \times n$ matrix with its $i^{th}$ row as $x_0^T Q_i$ and $I_n$ denotes an identity matrix of order $n$. Since the inequalities are placed component wise, i.e., we are looking for an $n$-cube around $u_0$, and $\Omega$ describes the volume covered by a polyhedron $[Ax]_i = b_i$, we shall chose the infinity-norm to work with, defined as : for a vector $v \in \mathbb{R}^n$ the infinity norm is defined as

\[ \|v\|_\infty := \max \{ |v_i| \mid i \in [n] \} \]

The corresponding induced norm is defined as follows: let $A : \mathbb{R}^n \times \mathbb{R}^m$ be a linear operator, then $\|A\|_\infty$ is defined as

\[ \|A\|_\infty := \max \{ \|Ax\|_\infty \mid \|x\|_\infty \leq 1 \} \]
Similarly, we can define a norm on a vector valued bilinear maps. Let \( \mathbb{R}^n \times \mathbb{R}^n \ni (u, v) \mapsto B(u, v) \in \mathbb{R}^m \) be a bilinear map. The induced norm on \( B \) is defined as
\[
\|B\|_\infty := \max \{ \|B(u, v)\|_\infty \mid \|u\|_\infty \leq 1, \|v\|_\infty \leq 1 \}.
\]
With these premise set, we give our first result. Let \( D_x f(x_0, u_0) = 2[x_0^T Q_1] + L \) be nonsingular at \((x_0, u_0)\) and define
\[
L_x = \|D_x f(x_0, u_0)\|_\infty \quad \text{and} \quad M_x = \|D_x f(x_0, u_0)^{-1}\|_\infty.
\]
The second order partials for \( f \) are defined as follows
\[
\mathbb{R}^n \times \mathbb{R}^n \ni (v_1, v_2) \mapsto D_x^2 f(x, u)(v_1, v_2) := 2[v_1^T Q_1] \in \mathbb{R}^n
\]
and
\[
D_x D_u f(x, u) = D_x^2 f(x, u) = 0.
\]
Consequently, we have
\[
K_{xx} = \max \{ 2 \|Q_i\|_\infty \mid i \in [n] \} \quad \text{and} \quad K_{xy} = K_{ux} = 0.
\]
**Theorem 4.3.** Let \( \Omega \) be nondegenerate i.e., has a nonempty interior and \( Q, L \) be such that \( 2[x_0^T Q_1] + L \) is non-singular. Then (4.6) is robustly solvable on \( \Omega \). Moreover, let \( M_x \) and \( K_{xx} \) be defined by (4.10) and (4.11). Then for a given \( r > 0 \) such that \( B(x_0, r) \subset \Omega \), the robustness margin denoted by \( \epsilon \) is lower bounded by \( \min \left\{ \frac{r}{2M_x}, \frac{1}{4M_x^2 K_{xx}} \right\} \).

**Proof.** Since \( D_x f(x_0, u_0) = 2[x_0^T Q_1] + L \) is nonsingular, from Theorem 3.11 there exists open set \( \bigcup (u_0) \ni u_0 \) and \( \bigcup (x_0) \ni x_0 \) such that for all \( u \in \bigcup (u_0) \), there exists a unique \( x \in \bigcup (x_0) \) such that \( f(x, u) = Q(x) + Lx - u = 0 \) which implies \( Q(x) + Lx = u \), i.e., \( x \) solves (4.6).

To show \( x \in \Omega \), we use Proposition 3.12. From Proposition 3.12, for all \( r \leq \frac{1}{2M_x K_{xx}} \), there exists an \( \epsilon \) satisfying
\[
M_x (L_x + K_{xx} r + K_{uu} \epsilon) \leq r/2,
\]
such that for all \( u \in B(u_0, \epsilon) \) there exists a \( x \in B(x_0, r) \) satisfying \( F(x) = u \). Substituting for \( L_x, K_{xx}, \) and \( K_{uu} \) we get
\[
\epsilon = \min \left\{ \frac{r}{2M_x}, \frac{1}{4M_x^2 K_{xx}} \right\},
\]
completing the proof. \( \square \)

Problem 4.2 can be also approached from the Inverse Function Theorem. Define
\[
x \mapsto F(x) := Q(x) + Lx.
\]
Clearly, \( DF(x_0) = 2[x_0^T Q_1 + L] \) is nonsingular and
\[
\mathbb{R}^n \times \mathbb{R}^n \ni (v_1, v_2) \mapsto D^2 F(x)(v_1, v_2) := 2[v_1^T Q_1] \in \mathbb{R}^n.
\]
Computing the quantities involved in Proposition 2.2, we have
\[
L_F = \|DF(x_0)\|_\infty = L_x, \quad M_F = \|DF^{-1}(u_0)\|_\infty = M_x,
\]
and
\[
K_F = \max \left\{ \|D^2 F(x)\|_\infty \mid x \in B(x_0, R) \right\} = K_{xx}.
\]
This gives us the opportunity to improve our estimates on \( \epsilon \). Relaxing the requirement of unique solution on \( \Omega \) we can now invoke Proposition 3.3 to come up with the extended robustness margin as follows
Robustness Margin and Extended Robust Margin for Example (4.12)

Let $M_x - K_{xx}$ be defined by (4.10) and (4.11). For a given $r > 0$ such that $B(x_0, r) \subseteq \Omega$, for (4.6) the extended robustness margin denoted by $\epsilon^*$ is lower bounded by

$$\min \left\{ \frac{r(2 - M_x - K_{xx})}{2M_x} \right\},$$

i.e., for all $u \in B(u_0, \epsilon^*)$ there exist (not necessarily unique) $x \in B(x_0, r)$, satisfying $F(x) = u$.

**Proof.** The above result follows from applying Proposition 3.3 on $F$. \qed

**Example.** In order to illustrate the above calculated bounds we present a simple example. The example is taken from [DKLR19] so as to establish comparisons. The data is defined as follows:

$$A = \begin{bmatrix} -1 & 0 \\ -1 & 0 \\ 0 & 1 \\ 0 & 1 \end{bmatrix}, \quad B = \begin{bmatrix} -0.5 & 0 \\ 3 & -0.5 \\ 3 & 3 \end{bmatrix}, \quad Q_1 = \begin{bmatrix} 1 & 0 \\ 0 & 0 \end{bmatrix}, \quad Q_2 = \begin{bmatrix} 0 & 0 \\ 0 & 1 \end{bmatrix}, \quad L = \begin{bmatrix} 1 & -3 \\ 2 & -1 \end{bmatrix}, \quad \text{and} \quad u_0 = \begin{bmatrix} -2 \\ 4 \end{bmatrix}$$

The underlying expression is then given by

$$F(x) = \begin{bmatrix} x_1^2 + 3x_1 - 3x_2 \\ x_2^2 + 2x_1 - x_2 \end{bmatrix} = \begin{bmatrix} -2 \\ 4 \end{bmatrix} \quad \text{with} \quad \Omega = \left\{ x : \begin{bmatrix} 0.5 \\ 0.5 \end{bmatrix} \leq \begin{bmatrix} x_1 \\ x_2 \end{bmatrix} \leq \begin{bmatrix} 3 \\ 3 \end{bmatrix} \right\}.$$  

The unique solution is given by $x_0 \approx [1.36 \quad 1.74]^T$. One can check that $DF(x_0)$ is nonsingular. From $\Omega$, we can see that $B(x_0, 0.8) \subset \Omega$. The corresponding quantities are $L_x = 6.7204, M_x = 0.3763, L_u = 1,$ and $K_{xx} = 2$. The robustness margin for (4.12) comes out to be, $\epsilon = 0.8829$, while, the extended robustness margin $\epsilon^* = 1.5461$. Figure 5 shows simulation results for the same. The constraint set $\Omega$ is shown in Figure 5 (a), and the corresponding Robust Margin (RM) and Extended Robust Margin (ERM) are shown in (b). For verification, we solve the (4.6) for each $u$ on the boundary of RM and ERM and, show that the solution $x$ is contained inside $\Omega$.

§ 4.3. Applications to Power Systems. With general nonlinear equations and quadratic equation case explained, we can now apply these results on the powerflow equations (4.1) and come up with a robustness margin. Setting

$$x = \begin{bmatrix} \text{Re}(V_1) & \ldots & \text{Re}(V_n) & \text{Im}(V_1) & \ldots & \text{Im}(V_n) \end{bmatrix}$$

and

$$u = \begin{bmatrix} p_1 & \ldots & p_n & q_1 & \ldots & q_n & v_1^2 & \ldots & v_n^2 \end{bmatrix}$$

Figure 5. Robustness Margin and Extended Robust Margin for Example (4.12)
From Table 1, one can see that the robustness margin decreases with the increase in the 
and \( K_1 \), where consider the variation in the first five dimension of are recorded in P.U. magnitude. In order to simulate real life scenarios, we only con-

...lem is simulated for test cases ‘5’, ‘9’, ‘14’ and ‘57’. The maximum allowable analysis. The package is open source and is available online \( ZMS19 \). The prob-

solving problems like, Power Flow analysis, Optimal Load Flow, and DC Power Flow 
system analysis. The package contains well defined libraries and set of routines for 
MATLAB software \( ZMST10 \). The MatPower package serves an excellent tool for power 

...test cases are obtained from the dataset given in the MatPower package found in 
MATLAB software \( ZMST10 \). The MatPower package serves an excellent tool for power 

...is expected as with the increase in \( n \) the problem dimension increases and norm based inequalities such as triangular inequality and submultiplicative inequality become poorer. Our estimates are conservative when compared to those provided in \( DKL19 \). However, the computation cost required to compute these bounds is much simpler than that required by \( DKL19 \). The only computationally intensive step in the proposed method is to compute the inverse of the Jacobian matrix \( DF \) at \( x_0 \), which can be tackled efficiently. Another possible limitation of this method is that it can only provide us with a lower estimate on the robustness margin.

### Table 1. Simulation data for cases ‘5’, ‘9’, ‘14’, and ‘57’

| Case | \( M_F \) | \( M'_F \) | \( K_F \) | Max \( r \) | Max \( \epsilon^* \) |
|------|----------|----------|--------|--------|--------|
| 5    | 0.5154   | 0.5053   | 507.617| 0.0038 | 0.0038 |
| 9    | 1.3802   | 0.7968   | 142.2375| 0.0051 | 0.0032 |
| 14   | 2.4795   | 0.5291   | 154.3248| 0.0026 | 0.0025 |
| 57   | 12.153   | 0.3697   | 304.5159| 0.0002 | 0.0003 |

§ 5. Conclusion

In this article, we have provided a lower estimate of the size of the neighborhoods involved in ImFT. The bounds given in \( AMR07 \) for the IFT are extended to the ImFT. On the way, as an addendum, we have also improved the estimates of neighborhoods for
IFT. The extended bounds as derived in Proposition 3.4 are twice as large as those given in Proposition 2.2 given in [AMR07]. However, this improvement comes at the cost of losing the uniqueness of the solution. This improvement is significant for the problems that involve finding solutions to the equations of type \( f(x) = u \), as they provide the guarantee of the existence of a solution in the local neighborhood of the operating point \((x_0, u_0)\), \( f(x_0) = u_0 \). The conventional proofs of ImFT involved extending the dimension of the underlying problem and this motivated us to look for an alternate proof of the classical ImFT. Based on which we have come up with another set of bounds derived in Proposition 3.12. If the underlying function \( f \) satisfied certain conditions on the second order partials calculated along different directions, given in (3.21), then the bounds calculated from the classical definition of the ImFT are better than those calculated by extending the ImFT to the IFT paradigm. As an application of these bounds, we have looked at the robustness of the solutions of the QCQP under uncertainties. This type of equation has a critical application in the power system, where the uncertainty arises from the intermittent and unpredictable nature of the renewable source. We validate our results on the benchmark systems provided in MatPower package of MATLAB software. Another very important application of these bounds is in the feedback linearization methods, where we use these bounds to come up with an estimate on the domain of feedback linearizability. This will be presented in a companion paper, along with other results on the feedback linearization of the discrete time systems. The bounds given in this paper are by no means the optimum estimates and the problem is wide open on both theoretical front: improving these estimates by looking at the structures of the underlying mapping \( f \), as well as on the application front: coming up with different applications of these bounds in the engineering problems.

REFERENCES

[AMR07] R. Abraham, J. E. Marsden, and T. Ratiu, Manifolds, Tensor Analysis, and Applications, Applied Mathematical Sciences, Springer Science & Business Media, 2007.

[Ash14] R. B. Ash, Complex Variables, Academic Press, 2014.

[BHM99] D. P. Bertsekas, W. Hager, and O. Mangasarian, Nonlinear Programming, Athena Scientific Belmont, 1999.

[CHP03] H. Chang, W. He, and N. Prabha, The analytic domain in the implicit function theorem, JIPAM. Journal of Inequalities in Pure and Applied Mathematics 4 (2003), no. 1.

[Cla76] F. Clarke, On the inverse function theorem, Pacific Journal of Mathematics 64 (1976), no. 1, 97–102.

[DK00] J. J. Duistermaat and J. AC Kolk, Lie Groups, Universitext, Springer, Berlin, Heidelberg, 2000.

[DKLR19] K. Dvijotham, B. Krishnamoorthi, Y. Luo, and B. Rapone, Robust feasibility of systems of quadratic equations using topological degree theory, arXiv preprint arXiv:1907.12206 (2019).

[DNT17] K. Dvijotham, H. Nguyen, and K. Turitsyn, Solvability regions of affinely parameterized quadratic equations, IEEE Control Systems Letters 2 (2017), no. 1, 25–30.

[Don99] A. L. Dontchev, Lipschitzian stability of Newton’s method for variational inclusions, IFIP Conference on System Modeling and Optimization, Springer, 1999, pp. 119–147.

[Eld13] J. Elderling, Normally Hyperbolic Invariant Manifolds: The Noncompact Case, Atlantis Series in Dynamical Systems, vol. 2, Atlantis Press, Paris, 2013.

[Fle01] R. Fletcher, Practical methods of optimization, 2nd edn. Wiley, Chichester, 2000, Numerical Algorithms 26 (2001), no. 2.

[JBC21] A. Jindal, R. Banavar, and D. Chatterjee, Feedback linearization of implicit discrete systems, 2021.

[KP02] S. G. Krantz and H. R. Parks, The Implicit Function Theorem: History, Theory, and Applications, Modern Birkhäuser Classics, Springer Science & Business Media, 2002.

[Lan97] S. Lang, Undergraduate Analysis, Undergraduate Texts in Mathematics, Springer Science & Business Media, 1997.

[Mor94] B. Mordukhovich, Lipschitzian stability of constraint systems and generalized equations, Nonlinear Analysis: Theory, Methods & Applications 22 (1994), no. 2, 173–206.

[NS96] S. G. Nash and A. Sofer, Linear and Nonlinear Programming, McGraw-Hill Science, Engineering & Mathematics, 1996.

[Pap05] M. Pap, On the domain of the implicit function and applications, Journal of Inequalities and Applications 2005 (2005), no. 3, 1–14.

[Rad76] W. Rudin, Principles of Mathematical Analysis, 3 ed., McGraw-hill New York, 1976.
[Spi95] M. Spivak, *Calculus on Manifolds: A Modern Approach to Classical Theorems of Advanced Calculus*, CRC press, 1995.

[Zei93] E. Zeidler, *Nonlinear Functional Analysis and its Applications Vol. I: Fixed-point Theorems/Transl. by Peter R. Wadsack*, Springer-Verlag, 1993.

[Zei95a] ———, *Applied Functional Analysis: Applications to Mathematical Physics*, Applied Mathematical Sciences, vol. 108, Springer Science & Business Media, 1995.

[Zei95b] ———, *Applied Functional Analysis: Main Principles and Their Applications*, Applied Mathematical Sciences, vol. 109, Springer Science & Business Media, 1995.

[ZMS19] R. D. Zimmerman and C. E. Murillo-Sánchez, *Matpower (version 7.0)/software*, 2019.

[ZMST10] R. D. Zimmerman, C. E. Murillo-Sánchez, and R. J. Thomas, *Matpower: Steady-state operations, planning, and analysis tools for power systems research and education*, IEEE Transactions on power systems 26 (2010), no. 1, 12–19.