Implicit Function Theorem in Perturbation Theory

khonvanny

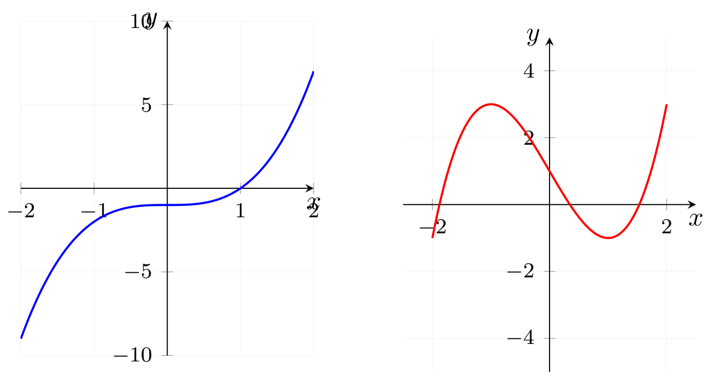
July 2025

1 Introduction

Consider the equation $x^3 - 1 = 0$. We know that this equation has a unique real solution, namely $x_0 = 1$. Now we want to know what happen to this unique solution if we add something small to this equation, for example if we add εx to the equation we obtain

$$x^3 + \varepsilon x - 1 = 0. \tag{E_1}$$

If ε is small, does equation (E_1) still has unique solution x_ε? And if it does, how does this solution x_ε relates to x_0. Note that if ε is large, say $\varepsilon = 3$, the equation has 3 real roots. See figures below. This is why we talk about ε small.



(a) Graph of $y = x^3 - 1$

(b) Graph of $y = x^3 - 3x + 1$

The equation $(E_1) : x^3 + \varepsilon x - 1$ is a perturbation of $x^3 - 1 = 0$ with ε is a perturbation parameter, and the equation $x^3 - 1 = 0$ is known as the unperturbed equation. In general, when we have equation $f(x) = 0$, a perturbed equation $f(x;\varepsilon) = f(x) + \varepsilon h(x,\varepsilon)$ is called an unfolding of f.

Before we proceed further, we state and prove the Banach Fixed Point Theorem, which is one of the most powerful tool in analysis.

Theorem 1 (Banach Fixed-Point Theorem)**.** *Let (X, d) be a complete metric space and let $T : X \to X$ be a contraction mapping, i.e., there exists a constant $L \in [0, 1)$ such that for all $x, y \in X$,*

$$d(T(x), T(y)) \leq L \cdot d(x, y).$$

Then T has a unique fixed point in X.

Proof. The proof consists of two parts: showing the existence of a fixed point and then its uniqueness.

First, for existence, we construct a sequence. Let $x_0 \in X$ be an arbitrary point. Define the sequence $(x_n)_{n\geq 0}$ by $x_{n+1} = T(x_n)$. We will show that this is a Cauchy sequence. By repeatedly applying the contraction property, we can bound the distance between consecutive terms:

$$d(x_{n+1}, x_n) = d(T(x_n), T(x_{n-1})) \leq L \cdot d(x_n, x_{n-1}) \leq \cdots \leq L^n \cdot d(x_1, x_0).$$

Now, for any integers $m > n \geq 0$, we use the triangle inequality and the geometric series formula to bound the distance $d(x_m, x_n)$:

$$\begin{aligned} d(x_m, x_n) &\leq \sum_{i=n}^{m-1} d(x_{i+1}, x_i) \\ &\leq \sum_{i=n}^{m-1} L^i d(x_1, x_0) \\ &= L^n(1 + L + \cdots + L^{m-n-1}) d(x_1, x_0) \\ &\leq L^n \left(\sum_{k=0}^{\infty} L^k \right) d(x_1, x_0) = \frac{L^n}{1 - L} d(x_1, x_0). \end{aligned}$$

Since $L \in [0, 1)$, we know that $L^n \to 0$ as $n \to \infty$. This implies that for any $\epsilon > 0$, we can find an integer N such that for all $m, n > N$, $d(x_m, x_n) < \epsilon$. This shows that (x_n) is a Cauchy sequence.

Since (X, d) is a complete metric space, the sequence (x_n) must converge to some limit $x^* \in X$. We now show that this limit is a fixed point. A contraction mapping is necessarily continuous, so we can pass the limit through the function:

$$x^* = \lim_{n\to\infty} x_n = \lim_{n\to\infty} x_{n+1} = \lim_{n\to\infty} T(x_n) = T(\lim_{n\to\infty} x_n) = T(x^*).$$

Thus, x^* is a fixed point of T.

For uniqueness, suppose there are two distinct fixed points, p and q, such that $T(p) = p$ and $T(q) = q$. Consider the distance between them and apply the contraction property:

$$d(p, q) = d(T(p), T(q)) \leq L \cdot d(p, q).$$

This inequality can be rewritten as $(1 - L)d(p, q) \leq 0$. Since $L < 1$, the term $(1 - L)$ is strictly positive. This forces $d(p, q) \leq 0$. As distance must be non-negative, we must have $d(p, q) = 0$, which implies $p = q$. Therefore, the fixed point is unique. $\square$

2 Implicit Function Theorem: First Encounter

Now we answer the above equation in the following theorem.

Theorem 2 (Implicit Function Theorem)**.** *Let's* $f : \mathbb{R}\times\mathbb{R} \to \mathbb{R}$ *be a* C^k *function in both variables, with* $k \geq 1$*. If* $f(x_0;\varepsilon_0) = 0$ *and* $D_x f(x_0, \varepsilon_0) \neq 0$*, then there exist* $\mu, \delta > 0$ *such that there is a* C^k *map* $g : (\varepsilon_0 - \mu, \varepsilon_) + \mu) \to (x_0 - \delta, x_0 + \delta)$ *such that* $g(\varepsilon_0) = x_0$ *and* $f(g(\varepsilon);\varepsilon) = 0$

Before we prove this theorem, we give some intuition of what the what it trying to say, and why it should be true. First, by definition if f is differentiable at x_0, then $f(x;\varepsilon) = f(x_0,\varepsilon)+D_x f(x_0,\varepsilon)(x-x_0)+O((x-x_0)^2)$. So $f(x,\varepsilon) = 0$ means we can find x that solve $f(x_0,\varepsilon)+D_x f(x_0,\varepsilon)(x-x_0)+O((x-x_0)^2)$. If we change variable form x to $x - x_0$, we have equation roughly look like $f(x_0,\varepsilon)+f'(x_0)x+O(x^2) = 0$, and the hypothesis of the theorem say that $f'(x_0,\varepsilon)$ is non-zero. If we at the theorem at linear level, we have $f(x_0,\varepsilon)+f'(x_0,\varepsilon)x = 0$. Since $f'(x_0,\varepsilon_0) \neq 0$, so when we change ε slightly from ε_0, we still have $f'(x_0,\varepsilon) \neq 0$ by continuity. Therefore, we can solve $x = -f(x_0,\varepsilon)/f'(x_0,\varepsilon)$. One can say that the implicit function assert that if we can solve the equation at linear level, then we can solve it in non-linear level and the solution depends smoothly on the parameter.

Proof. We prove this theorem by a contraction-mapping argument. We let $D = D_x f(x_0,\varepsilon_0) \neq 0$, and for each fixed ε we consider the map

$$\phi(x,\varepsilon) = x - D^{-1} f(x,\varepsilon).$$

A solution x_ε to the equation $f(x,\varepsilon) = 0$ is a fixed point of $\phi(x,\varepsilon)$, since $\phi(x,\varepsilon) = x$ implies $D^{-1}f(x,\varepsilon) = 0$, and because D is invertible, this means $f(x,\varepsilon) = 0$. Our task is then to show that for ε sufficiently close to ε_0, we can find a neighborhood $B_\delta = [x_0 - \delta, x_0 + \delta]$ of x_0 that makes the map $x \mapsto \phi(x,\varepsilon)$ a contraction that maps B_δ to itself.

First, to show that ϕ is a contraction, we analyze its derivative with respect to x:

$$D_x\phi(x,\varepsilon) = 1 - D^{-1}D_x f(x,\varepsilon).$$

At the point (x_0,ε_0), this derivative is zero:

$$D_x\phi(x_0,\varepsilon_0) = 1 - D^{-1}D_x f(x_0,\varepsilon_0) = 1 - D^{-1}D = 0.$$

Since f is C^k (with $k \geq 1$), $D_x f(x,\varepsilon)$ is continuous, and so is $D_x\phi(x,\varepsilon)$. By continuity, we can find $\delta_1 > 0$ and $\mu_1 > 0$ such that for all $x \in (x_0 - \delta_1, x_0 + \delta_1)$ and $\varepsilon \in (\varepsilon_0 - \mu_1, \varepsilon_0 + \mu_1)$, we have $|D_x\phi(x,\varepsilon)| \leq 1/2$. By the Mean Value Theorem, for any x_1, x_2 in this neighborhood of x_0,

$$|\phi(x_1,\varepsilon) - \phi(x_2,\varepsilon)| \leq \frac{1}{2}|x_1 - x_2|.$$

This shows ϕ is a contraction in x on this neighborhood.

Next, we must show that ϕ maps a ball $B_\delta = [x_0 - \delta, x_0 + \delta]$ into itself for a suitable choice of $\delta \le \delta_1$ and $\mu \le \mu_1$. For any $x \in B_\delta$, we estimate $|\phi(x, \varepsilon) - x_0|$ using the triangle inequality:

$$|\phi(x, \varepsilon) - x_0| \le |\phi(x, \varepsilon) - \phi(x_0, \varepsilon)| + |\phi(x_0, \varepsilon) - x_0|.$$

The first term is bounded by the contraction property: $|\phi(x, \varepsilon) - \phi(x_0, \varepsilon)| \le \frac{1}{2}|x - x_0| \le \frac{1}{2}\delta$. For the second term, we use the fact that $f(x_0, \varepsilon_0) = 0$:

$$|\phi(x_0, \varepsilon) - x_0| = |x_0 - D^{-1} f(x_0, \varepsilon) - x_0| = |D^{-1} f(x_0, \varepsilon)|.$$

Since f is continuous, we can choose $\mu_2 > 0$ such that for $|\varepsilon - \varepsilon_0| < \mu_2$, we have $|f(x_0, \varepsilon)| \le \frac{\delta}{2|D^{-1}|}$. This makes the second term $|\phi(x_0, \varepsilon) - x_0| \le \frac{\delta}{2}$. Combining these bounds, and choosing $\mu = \min(\mu_1, \mu_2)$, we get

$$|\phi(x, \varepsilon) - x_0| \le \frac{1}{2}\delta + \frac{1}{2}\delta = \delta.$$

Thus, $\phi(\cdot, \varepsilon)$ maps the closed interval B_δ to itself.

Since $\phi(\cdot, \varepsilon)$ is a contraction on the complete metric space B_δ, the Banach Fixed-Point Theorem guarantees that for each $\varepsilon \in (\varepsilon_0 - \mu, \varepsilon_0 + \mu)$, there is a unique fixed point, which we call $g(\varepsilon) \in B_\delta$. This function satisfies $f(g(\varepsilon), \varepsilon) = 0$ and by construction $g(\varepsilon_0) = x_0$.

Finally, we must show that the function $g(\varepsilon)$ is C^k. This is done by first proving continuity, then showing it is C^1, and finally using induction.

To prove the continuity of g, let $\varepsilon_n \to \varepsilon$ be a convergent sequence in the interval $(\varepsilon_0 - \mu, \varepsilon_0 + \mu)$. Let $x_n = g(\varepsilon_n)$ and $x = g(\varepsilon)$. We want to show that $x_n \to x$. Each x_n lies in the compact set $B_\delta = [x_0 - \delta, x_0 + \delta]$, so the sequence (x_n) must have a convergent subsequence, (x_{n_k}), which converges to some limit $x^* \in B_\delta$. By definition, x_{n_k} is the fixed point for ε_{n_k}:

$$x_{n_k} = \phi(x_{n_k}, \varepsilon_{n_k}).$$

Since $\phi(x, \varepsilon)$ is a continuous function of both its arguments, we can take the limit as $k \to \infty$:

$$\lim_{k\to\infty} x_{n_k} = \lim_{k\to\infty} \phi(x_{n_k}, \varepsilon_{n_k}) = \phi(\lim_{k\to\infty} x_{n_k}, \lim_{k\to\infty} \varepsilon_{n_k}).$$

This implies $x^* = \phi(x^*, \varepsilon)$. Thus, x^* is a fixed point of the map $\phi(\cdot, \varepsilon)$. However, the fixed point in B_δ is unique and is equal to $x = g(\varepsilon)$. Therefore, $x^* = x$. We have shown that any convergent subsequence of (x_n) converges to the same limit x. For a sequence in a compact set, this implies that the entire sequence (x_n) must converge to x. Thus, g is continuous.

Now that we have established that g is continuous, we can show it is C^1. From the identity $f(g(\varepsilon), \varepsilon) = 0$, we differentiate with respect to ε using the chain rule:

$$D_x f(g(\varepsilon), \varepsilon) \cdot g'(\varepsilon) + D_\varepsilon f(g(\varepsilon), \varepsilon) = 0.$$

Since $D_x f(x_0, \varepsilon_0) \neq 0$ and all functions are continuous, $D_x f(g(\varepsilon), \varepsilon)$ is non-zero for ε near ε_0. We can therefore solve for $g'(\varepsilon)$:

$$g'(\varepsilon) = -\frac{D_\varepsilon f(g(\varepsilon), \varepsilon)}{D_x f(g(\varepsilon), \varepsilon)}.$$

If f is C^1, then its partial derivatives are continuous. Since g is also continuous, the right-hand side is a composition and quotient of continuous functions with a non-zero denominator, and is therefore continuous. This proves that $g'(\varepsilon)$ is continuous, so $g(\varepsilon)$ is a C^1 function.

Lastly, we use induction to show that if f is C^k, then g is C^k. The base case $k = 1$ has just been proven. Assume for some m with $1 \leq m < k$, that if f is C^m, then g is C^m. Now, let f be C^{m+1}. This means its partial derivatives, $D_x f$ and $D_\varepsilon f$, are C^m functions. By our inductive hypothesis, g is C^m. In the formula for $g'(\varepsilon)$, the right-hand side involves compositions and quotients of C^m functions, which means the result is a C^m function. So, $g'(\varepsilon)$ is C^m. If a function's derivative is C^m, the function itself must be C^{m+1}. Thus, $g(\varepsilon)$ is C^{m+1}. By the principle of induction, the statement holds for all $k \geq 1$. This completes the proof. $\square$

We can now apply this theorem to the example from the introduction, $x^3 + \varepsilon x - 1 = 0$, to help us understand the solution $x(\varepsilon)$ near $\varepsilon = 0$. Let $f(x, \varepsilon) = x^3 + \varepsilon x - 1$. For the unperturbed case $\varepsilon_0 = 0$, we have the solution $x_0 = 1$. At the point $(x_0, \varepsilon_0) = (1, 0)$, we have $f(1, 0) = 0$ and the partial derivative $D_x f(1, 0) = [3x^2 + \varepsilon]_{(1,0)} = 3 \neq 0$. The conditions of the theorem are satisfied, which guarantees that a unique, smooth solution $x(\varepsilon)$ exists in a neighborhood of $\varepsilon = 0$, with $x(0) = 1$.

To find the expansion of this solution, we use the fact that the identity $f(x(\varepsilon), \varepsilon) = 0$ holds for all ε in this neighborhood. We can therefore differentiate the expression with respect to ε:

$$\frac{d}{d\varepsilon}\big(x(\varepsilon)^3 + \varepsilon x(\varepsilon) - 1\big) = 0.$$

Using the chain rule and the product rule, we get

$$3x(\varepsilon)^2 x'(\varepsilon) + x(\varepsilon) + \varepsilon x'(\varepsilon) = 0.$$

We can find the first coefficient of the Taylor series, $x_1 = x'(0)$, by evaluating this equation at $\varepsilon = 0$. Since we know $x(0) = 1$:

$$3(1)^2 x'(0) + 1 + (0)x'(0) = 0 \implies 3x'(0) = -1 \implies x'(0) = -\frac{1}{3}.$$

To find the next term, we differentiate the expression again with respect to ε:

$$\frac{d}{d\varepsilon}\big(3x(\varepsilon)^2 x'(\varepsilon) + x(\varepsilon) + \varepsilon x'(\varepsilon)\big) = 0.$$

This gives:

$$\big(6x(\varepsilon)x'(\varepsilon)^2 + 3x(\varepsilon)^2x''(\varepsilon)\big) + x'(\varepsilon) + (x'(\varepsilon) + \varepsilon x''(\varepsilon)) = 0.$$

Again, we evaluate at $\varepsilon = 0$, using our known values $x(0) = 1$ and $x'(0) = -1/3$:

$$6(1)\left(-\frac{1}{3}\right)^2 + 3(1)^2x''(0) + 2\left(-\frac{1}{3}\right) + (0)x''(0) = 0.$$

$$6\left(\frac{1}{9}\right) + 3x''(0) - \frac{2}{3} = 0 \implies \frac{2}{3} + 3x''(0) - \frac{2}{3} = 0.$$

This simplifies to $3x''(0) = 0$, so $x''(0) = 0$. The second-order coefficient is $x_2 = x''(0)/2! = 0$. Therefore, the Taylor expansion of the solution up to second order is:

$$x(\varepsilon) = x(0) + x'(0)\varepsilon + \frac{x''(0)}{2!}\varepsilon^2 + O(\varepsilon^3) = 1 - \frac{1}{3}\varepsilon + O(\varepsilon^3).$$

The example we've just seen can be easily done without implicit function as the unfolding $x^3 + \varepsilon x - 1 = 0$ is still a cubic equation. Next we present another example that look slightly more complicated.

Example 1. *Consider the equation $x^3 + \varepsilon \sin(x) - 1 = 0$. We seek the second-order expansion for the solution $x(\varepsilon)$ that continues from the unperturbed solution $x_0 = 1$.*

Proof. Let $f(x, \varepsilon) = x^3 + \varepsilon \sin(x) - 1$. At the point $(x_0, \varepsilon_0) = (1, 0)$, we have $f(1, 0) = 0$. The partial derivative is $D_x f(x, \varepsilon) = 3x^2 + \varepsilon \cos(x)$, and at our point, $D_x f(1, 0) = 3 \neq 0$. The Implicit Function Theorem guarantees a unique smooth solution $x(\varepsilon)$ with $x(0) = 1$.

We differentiate the identity $x(\varepsilon)^3 + \varepsilon \sin(x(\varepsilon)) - 1 = 0$ with respect to ε:

$$3x(\varepsilon)^2x'(\varepsilon) + \sin(x(\varepsilon)) + \varepsilon \cos(x(\varepsilon))x'(\varepsilon) = 0.$$

Evaluating at $\varepsilon = 0$ with $x(0) = 1$ gives:

$$3(1)^2x'(0) + \sin(1) + 0 = 0 \implies x'(0) = -\frac{\sin(1)}{3}.$$

To find the second derivative, we differentiate the expression again:

$$\frac{d}{d\varepsilon}\big(3x^2x' + \sin(x) + \varepsilon x' \cos(x)\big) = 0.$$

This yields:

$$6x(x')^2 + 3x^2x'' + 2x' \cos(x) + \varepsilon(\dots)' = 0.$$

Evaluating at $\varepsilon = 0$ with $x(0) = 1$ and $x'(0) = -\sin(1)/3$:

$$6(1)\left(-\frac{\sin(1)}{3}\right)^2 + 3(1)^2x''(0) + 2\left(-\frac{\sin(1)}{3}\right)\cos(1) = 0.$$

$$\frac{2\sin^2(1)}{3} + 3x''(0) - \frac{2\sin(1)\cos(1)}{3} = 0.$$

Solving for $x''(0)$ gives:

$$3x''(0) = \frac{2\sin(1)\cos(1) - 2\sin^2(1)}{3} \implies x''(0) = \frac{2\sin(1)(\cos(1) - \sin(1))}{9}.$$

The Taylor expansion is therefore:

$$x(\varepsilon) = 1 - \frac{\sin(1)}{3}\varepsilon + \frac{\sin(1)(\cos(1) - \sin(1))}{9}\varepsilon^2 + O(\varepsilon^3).$$

□

To visualize the accuracy of these perturbation expansions, we can compare them to a numerical solution obtained using a root-finding algorithm. Figure 1 shows the numerical solution (obtained using MATLAB's `fzero` function) plotted against the first-order and second-order perturbation approximations for a range of ε values.

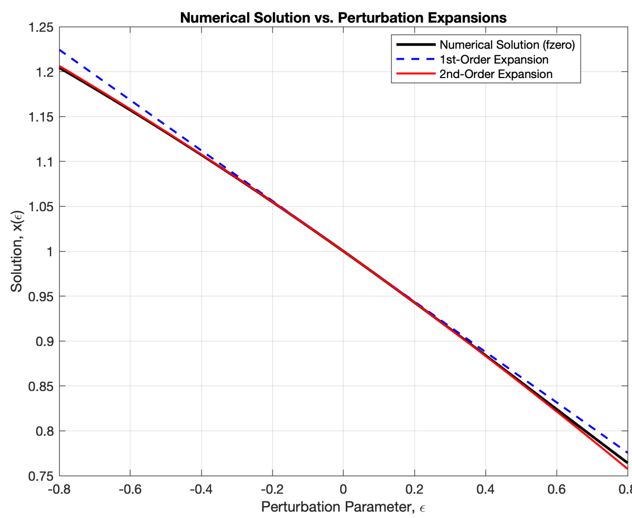


Figure 1: Comparison of numerical solution and perturbation expansions for $x^3 + \varepsilon\sin(x) - 1 = 0$.

3 Implicit Function Theorem: Second Look

The Implicit Function Theorem seen in previous section can be generalized higher dimensional setting, where we consider a system of n equations whose solution depends on d parameters. The core idea remains the same: if we have a solution and the system is non-singular with respect to its variables at that point, we can still describe the solutions as a smooth function of the parameters.

Theorem 3 (Implicit Function Theorem)**.** *Let $U \subseteq \mathbb{R}^n \times \mathbb{R}^d$ be an open set and let $f : U \to \mathbb{R}^n$ be a C^k function for $k \geq 1$. Suppose there is a point $(x_0, \varepsilon_0) \in U$ such that $f(x_0, \varepsilon_0) = \mathbf{0}$. If the $n \times n$ Jacobian matrix of f with respect to x, denoted $D_x f(x_0, \varepsilon_0)$, is invertible, then there exist open neighborhoods $W \subseteq \mathbb{R}^d$ of ε_0 and $V \subseteq \mathbb{R}^n$ of x_0 such that there is a unique C^k function $g : W \to V$ satisfying $g(\varepsilon_0) = x_0$ and $f(g(\varepsilon), \varepsilon) = \mathbf{0}$ for all $\varepsilon \in W$.*

Proof. The proof follows the same contraction-mapping strategy used in the one-dimensional case. Let $D = D_x f(x_0, \varepsilon_0)$ be the invertible $n \times n$ Jacobian matrix. For a fixed ε, we seek a fixed point for the operator $\phi : \mathbb{R}^n \to \mathbb{R}^n$ defined by

$$\phi(x, \varepsilon) = x - D^{-1} f(x, \varepsilon).$$

A point x is a fixed point of ϕ if and only if $f(x, \varepsilon) = \mathbf{0}$. Our goal is to find a closed ball $\bar{B}_\delta(x_0) = \{x \in \mathbb{R}^n : \|x - x_0\| \leq \delta\}$ on which $\phi(\cdot, \varepsilon)$ is a contraction that maps the ball to itself.

First, we show ϕ is a contraction. The Jacobian of ϕ with respect to x is given by

$$D_x \phi(x, \varepsilon) = I - D^{-1} D_x f(x, \varepsilon),$$

where I is the $n \times n$ identity matrix. At our point (x_0, ε_0), this Jacobian is the zero matrix:

$$D_x \phi(x_0, \varepsilon_0) = I - D^{-1} D_x f(x_0, \varepsilon_0) = I - D^{-1} D = \mathbf{0}.$$

Since f is C^k, its Jacobian $D_x f$ is a continuous function of (x, ε). Therefore, $D_x \phi$ is also continuous. By continuity, we can choose $\delta_1 > 0$ and $\mu_1 > 0$ such that for all x in the ball $B_{\delta_1}(x_0)$ and ε in $B_{\mu_1}(\varepsilon_0)$, the operator norm of the Jacobian is bounded: $\|D_x \phi(x, \varepsilon)\| \leq 1/2$. By the Mean Value Inequality, for any $x_1, x_2 \in B_{\delta_1}(x_0)$, we have

$$\|\phi(x_1, \varepsilon) - \phi(x_2, \varepsilon)\| \leq \sup_{c \in [x_1, x_2]} \|D_x \phi(c, \varepsilon)\| \cdot \|x_1 - x_2\| \leq \frac{1}{2} \|x_1 - x_2\|.$$

This shows that $\phi(\cdot, \varepsilon)$ is a contraction on $B_{\delta_1}(x_0)$.

Next, we show the self-mapping property on a possibly smaller ball $\bar{B}_\delta(x_0)$ with $\delta \leq \delta_1$. For any $x \in \bar{B}_\delta(x_0)$, we have

$$\|\phi(x, \varepsilon) - x_0\| \leq \|\phi(x, \varepsilon) - \phi(x_0, \varepsilon)\| + \|\phi(x_0, \varepsilon) - x_0\|.$$

The first term is bounded by the contraction property: $\|\phi(x, \varepsilon) - \phi(x_0, \varepsilon)\| \leq \frac{1}{2}\|x - x_0\| \leq \frac{1}{2}\delta$. The second term is $\|\phi(x_0, \varepsilon) - x_0\| = \| - D^{-1} f(x_0, \varepsilon)\|$. Since f is continuous and $f(x_0, \varepsilon_0) = \mathbf{0}$, we can choose $\mu \leq \mu_1$ small enough such that for $\varepsilon \in B_\mu(\varepsilon_0)$, we have $\|f(x_0, \varepsilon)\| \leq \frac{\delta}{2\|D^{-1}\|}$. This ensures $\|\phi(x_0, \varepsilon) - x_0\| \leq \|D^{-1}\| \frac{\delta}{2\|D^{-1}\|} = \frac{\delta}{2}$. Combining the bounds, we get

$$\|\phi(x, \varepsilon) - x_0\| \leq \frac{1}{2}\delta + \frac{1}{2}\delta = \delta.$$

This confirms that ϕ maps the closed ball $\bar{B}_\delta(x_0)$ into itself.

Since $\bar{B}_\delta(x_0)$ is a closed subset of $\mathbb{R}^n$, it is a complete metric space. The Banach Fixed-Point Theorem thus guarantees the existence of a unique fixed point $g(\varepsilon) \in \bar{B}_\delta(x_0)$ for each $\varepsilon \in B_\mu(\varepsilon_0)$. The proof of the C^k smoothness of g follows by differentiating the identity $f(g(\varepsilon), \varepsilon) = \mathbf{0}$ with respect to ε, which yields the expression for the Jacobian of g:

$$Dg(\varepsilon) = -[D_x f(g(\varepsilon), \varepsilon)]^{-1} D_\varepsilon f(g(\varepsilon), \varepsilon).$$

The same continuity and induction argument used in the one-dimensional case shows that if f is C^k, then g must also be C^k. □

Now we give an important application of the implicit function theorem, which the inverse function theorem, that is a common tools in geometry and topology. The theorem simply say that if at linear level, a function is invertible, then it must be invertible at non-linear level.

Example 2 (The Inverse Function Theorem)**.** *Let $F : \mathbb{R}^n \to \mathbb{R}^n$ be a C^k function for $k \geq 1$. If at a point $x_0 \in \mathbb{R}^n$, the derivative (Jacobian matrix) $DF(x_0)$ is invertible, then F is locally invertible near x_0. Specifically, there exist open neighborhoods V of x_0 and W of $y_0 = F(x_0)$ such that there is a unique C^k inverse function $g : W \to V$ satisfying $F(g(y)) = y$ for all $y \in W$.*

Proof. The statement is about finding a function g that inverts F. We can rephrase this as a root-finding problem. We want to find a function $x = g(y)$ that solves the equation $F(x) = y$.

To fit this into the framework of the Implicit Function Theorem, we define a new function $f : \mathbb{R}^n \times \mathbb{R}^n \to \mathbb{R}^n$ by

$$f(x, y) = F(x) - y.$$

Here, x plays the role of the variable we wish to solve for, and y plays the role of the parameter ε. Our goal is to solve the equation $f(x, y) = \mathbf{0}$ for x as a function of y.

Let $y_0 = F(x_0)$. We have a point (x_0, y_0) where $f(x_0, y_0) = F(x_0) - y_0 = \mathbf{0}$. We now check the main condition of the Implicit Function Theorem. We need to compute the Jacobian of f with respect to the variable x and check if it is invertible at (x_0, y_0).

$$D_x f(x, y) = D_x(F(x) - y) = DF(x).$$

The condition of the Inverse Function Theorem is precisely that this matrix, $DF(x_0)$, is invertible at the point x_0.

Since f is C^k and its Jacobian $D_x f(x_0, y_0) = DF(x_0)$ is invertible, all conditions of the Implicit Function Theorem are met. The theorem guarantees the existence of open neighborhoods W of y_0 and V of x_0, and a unique C^k function $g : W \to V$ such that $f(g(y), y) = \mathbf{0}$ for all $y \in W$.

Substituting the definition of f back in, this means

$$F(g(y)) - y = \mathbf{0},$$

or $F(g(y)) = y$. This function g is precisely the local inverse of F. Thus, the Inverse Function Theorem is established as a direct consequence of the Implicit Function Theorem. $\square$

The inverse function is very important in the study of differential topology and geometry. However, we do not spend time on these topics here, and instead we will proceed with the implicit function theorem which has more connection to perturbation theory, and to fully get there, we need to talk about a more general form of the theorem which is on a Banach Space setting where we take up in the next sectio.

4 Implicit Function Theorem: Third Look

To do implicit function theorem in Banach spaces, we need to discuss some notion involving Banach space. First, recall that a Banach space $\mathcal{B}$ is a normed vector space over $\mathbb{R}$ or $\mathbb{C}$ that is complete. The spaces $\mathbb{R}^d$, $\mathbb{C}^d$ are examples of Banach spaces. The space of linear map $\mathcal{L}(\mathbb{C}^m, \mathbb{C}^n)$, which can be identified with $\mathbb{C}^{mn}$, are also Banach spaces. The space of function $C(S^1) = \{f : S^1 \to \mathbb{C}\}$ is a Banach space with respect to the sup norm $\|.\|_\infty$. The space p integrable function $L^p(\mathbb{R})$, equipped with L^p norm, are also Banach Spaces. The Sobolev space $H^k(\mathbb{R})$ of function with weak derivative up to order k and each of those derivative is square integrable are also Banach spaces.

To property state the implicit function theorem on Banach Space, we need to talk about continuity and derivative first. There is not mystery for continuity since a Banach Space is just a norm space and therefore we can use the norm to define continuity whenever we need to. For differentiability, the idea is the same as in finite dimensional case, if a function $f : X \to Y$ is differentiable at x_0 if it can be approximated by linear map near x_0, and f is differentiable on $U \subset X$ if f is differentiable for each $x \in U$. We will make this more precise later, but for now we want to state this idea to highly that we need to talk about linear map on Banach spaces and this is where finite dimensional spaces and infinite dimensional spaces are different. The key different is that in finite dimensional space, linearity implies continuity. However, this is not true in infinite dimensional space. Let's give an example of this phenomenon.

Consider $\partial_x : C^1(S^1) \to C^0(S^1)$ is a linear map. Now a map is continuous if for any sequence $f_n \to f$ in C^0, we must have $\partial_x f_n \to \partial_x f$ in C^0. However, if we look at the sequence $f_n(x) = 1/n \sin(n^2 x)$. This sequence converge to 0 function, since $|f_n(x) - 0| = \sup_{x\in[0,2\pi]} 1/n|\sin(n^2 x)| = 1/n \to 0$. However, $\partial_x f_n(x) = n \sin(n^2 x)$, and the norm $\|\partial_x f_n - 0\| = \sup_{x\in[0,2\pi]} |n \sin(n^2 x)| = n \to +\infty$.

It turn out that for linear map, in any dimension, the condition that ensures continuity is being continuous at a point.

Theorem 4. *Let $L : X \to Y$ be a linear map, where X and Y are Banach spaces. The map L is continuous if and only if one of the following statement is true*

1. *There exist $x_0 \in X$ such that L is continuous at x_0.*
2. *There exist $M > 0$ such that $|L(x)| \leq M|x|$ for all $x \in X$.*

Remark 1. *Because of condition 2, we also call continuous linear map as bounded map.*

We will not prove this theorem here and instead we encourage the reader to prove it as it will help understand continuity and linearity. However, we remark that the second condition allow us to generalize the notion of size of matrices, which are linear map in finite dimensional spaces, to general linear map.

Definition 1. *Let $L : X \to Y$ be a continuous linear map between Banach spaces X and Y. The norm of L is defined as*

$$\|L\|_\infty = \inf_{M>0}\{M > 0 \text{ such that } |L(x)| \leq M|x| \text{ for all } x \in X\}$$

This norm give the space of all continuous linear map between two Banach spaces a Banach space structure. First, the space of continuous linear map from X to Y is denoted by $\mathcal{L}(X, Y)$.

Theorem 5. *The norm $\|.\|_\infty$ on linear map from X to Y make $\mathcal{L}(X, Y)$ a Banach space.*

Definition 2 (Invertible Operator)**.** *A continuous linear map $L \in \mathcal{L}(X, Y)$ is called an* ***isomorphism*** *if it is a bijection (one-to-one and onto) and its inverse L^{-1} is also a continuous linear map, i.e., $L^{-1} \in \mathcal{L}(Y, X)$.*

Remark 2. *A celebrated result called the Bounded Inverse Theorem states that if a bounded linear operator between two Banach spaces is a bijection, its inverse is automatically bounded (and thus continuous). Therefore, for a map in $\mathcal{L}(X, Y)$ between Banach spaces, being invertible is equivalent to being an isomorphism.*

Now we can precisely define the derivative of a map between Banach spaces. This is often called the Fréchet derivative.

Definition 3 (Differentiability)**.** *Let X and Y be Banach spaces and $U \subseteq X$ be an open set. A function $f : U \to Y$ is said to be* ***differentiable*** *at a point $x_0 \in U$ if there exists a continuous linear map $L \in \mathcal{L}(X, Y)$ such that*

$$\lim_{\|h\|_X \to 0} \frac{\|f(x_0 + h) - f(x_0) - L(h)\|_Y}{\|h\|_X} = 0.$$

The unique linear map L is called the derivative of f at x_0 and is denoted by $Df(x_0)$.

Example 3 (Derivative of a Linear Operator)**.** *Let $L : X \to Y$ be a continuous (i.e., bounded) linear operator between Banach spaces. The derivative of L at any point $x_0 \in X$ is the operator L itself.*

Proof. Let $f(x) = L(x)$. To find the derivative at x_0, we must find a linear operator, let's call it A, that satisfies the limit definition. Let's propose that the derivative is $A = L$. We must check if the remainder term goes to zero appropriately:

$$f(x_0 + h) - f(x_0) - L(h).$$

By the linearity of $f = L$, we have $f(x_0 + h) = L(x_0 + h) = L(x_0) + L(h)$. Substituting this in gives:

$$(L(x_0) + L(h)) - L(x_0) - L(h) = 0.$$

The limit in the definition of the derivative is therefore

$$\lim_{\|h\| \to 0} \frac{\|0\|_Y}{\|h\|_X} = 0.$$

This holds trivially. Thus, the Fréchet derivative of a bounded linear map is the map itself at every point: $D_x L = L$. □

A more substantial example, which is fundamental in matrix analysis and operator theory, is finding the derivative of the inversion map.

Example 4 (Derivative of a Non-linear Integral Operator)**.** *Let $X = C([0, 1])$ be the Banach space of continuous functions on the interval $[0, 1]$ with the sup norm. Consider the non-linear map $F : X \to X$ defined by*

$$F(u)(x) = \int_0^1 e^{x-y} \sin(u(y)) dy.$$

We will compute its Fréchet derivative at a point $u_0 \in X$.

Proof. To find the derivative, we analyze the expression $F(u_0 + h)$ for a small function $h \in X$.

$$F(u_0 + h)(x) = \int_0^1 e^{x-y} \sin(u_0(y) + h(y)) dy.$$

We use the Taylor expansion for the sine function around the point $u_0(y)$:

$$\sin(u_0(y) + h(y)) = \sin(u_0(y)) + \cos(u_0(y))h(y) + O(h(y)^2).$$

Substituting this into the integral gives

$$F(u_0 + h)(x) = \int_0^1 e^{x-y} \left(\sin(u_0(y)) + \cos(u_0(y))h(y) + O(h(y)^2)\right) dy.$$

We can split this into three parts by linearity of the integral:

$$F(u_0+h)(x) = \underbrace{\int_0^1 e^{x-y}\sin(u_0(y))dy}_{F(u_0)(x)} + \underbrace{\int_0^1 e^{x-y}\cos(u_0(y))h(y)dy}_{\text{Linear part in } h} + \underbrace{\int_0^1 e^{x-y}O(h(y)^2)dy}_{\text{Remainder}}.$$

The derivative $DF(u_0)$ is the linear operator that maps the perturbation h to the linear part of the response. So, the derivative is the operator $L = DF(u_0)$ defined by:

$$[L(h)](x) = [DF(u_0)(h)](x) = \int_0^1 e^{x-y}\cos(u_0(y))h(y)dy.$$

To confirm this, we must show that the norm of the remainder term is $o(\|h\|)$. The remainder term $R(h)$ satisfies

$$\|R(h)\|_\infty = \sup_{x\in[0,1]} \left| \int_0^1 e^{x-y}O(h(y)^2)dy \right| \le \sup_{x\in[0,1]} \int_0^1 |e^{x-y}||O(h(y)^2)|dy.$$

Since $|O(h(y)^2)| \le C\|h\|_\infty^2$ for some constant C, and the kernel e^{x-y} is bounded on the domain, the entire expression is bounded by a constant times $\|h\|_\infty^2$. An expression that is $O(\|h\|_\infty^2)$ is also $o(\|h\|_\infty)$, so the limit condition for the Fréchet derivative is satisfied. □

With these tools, we are ready to state and prove the Implicit Function Theorem in its full generality. The statement and proof are remarkably similar to the finite-dimensional case, which highlights the power of the abstract framework.

Theorem 6 (Implicit Function Theorem on Banach Spaces)**.** *Let X, Y, Z be Banach spaces, let $U \subseteq X \times Y$ be an open set, and let $f : U \to Z$ be a C^k map for $k \ge 1$. Suppose there is a point $(x_0, y_0) \in U$ such that $f(x_0, y_0) = 0$. If the partial Fréchet derivative with respect to x, $D_x f(x_0, y_0) \in \mathcal{L}(X, Z)$, is an isomorphism, then there exist open neighborhoods $W \subseteq Y$ of y_0 and $V \subseteq X$ of x_0 such that there is a unique C^k function $g : W \to V$ satisfying $g(y_0) = x_0$ and $f(g(y), y) = 0$ for all $y \in W$.*

Proof. The proof is a direct application of the Banach Fixed-Point Theorem, mirroring the finite-dimensional case. Let $L = D_x f(x_0, y_0)$. By hypothesis, L is an isomorphism from X to Z. We want to find a solution x to $f(x, y) = 0$ for y near y_0. This is equivalent to finding a fixed point of the operator $\phi : X \to X$ defined by

$$\phi(x, y) = x - L^{-1}(f(x, y)).$$

Our goal is to find a closed ball $\bar{B}_\delta(x_0) = \{x \in X : \|x - x_0\| \le \delta\}$ on which $\phi(\cdot, y)$ is a contraction that maps the ball to itself.

First, we show ϕ is a contraction in the x variable. The partial Fréchet derivative of ϕ with respect to x is

$$D_x\phi(x, y) = I - L^{-1}D_x f(x, y),$$

where I is the identity operator on X. At the point (x_0, y_0), this derivative is the zero operator:

$$D_x\phi(x_0, y_0) = I - L^{-1}D_x f(x_0, y_0) = I - L^{-1}L = 0.$$

Since f is C^1, the map $(x, y) \mapsto D_x f(x, y)$ is continuous. Thus, we can find neighborhoods of x_0 and y_0 such that the operator norm is bounded: $\|D_x\phi(x, y)\| \leq 1/2$. By the Mean Value Inequality, for any x_1, x_2 in this neighborhood of x_0, we have

$$\|\phi(x_1, y) - \phi(x_2, y)\| \leq \frac{1}{2}\|x_1 - x_2\|.$$

Next, to show the self-mapping property, consider $x \in \bar{B}_\delta(x_0)$.

$$\|\phi(x, y) - x_0\| \leq \|\phi(x, y) - \phi(x_0, y)\| + \|\phi(x_0, y) - x_0\|.$$

The first term is bounded by the contraction property: $\|\phi(x, y) - \phi(x_0, y)\| \leq \frac{1}{2}\|x - x_0\| \leq \frac{1}{2}\delta$. The second term is $\|\phi(x_0, y) - x_0\| = \| - L^{-1}(f(x_0, y))\| \leq \|L^{-1}\|\|f(x_0, y)\|$. Since f is continuous and $f(x_0, y_0) = 0$, we can make $\|f(x_0, y)\|$ small by choosing y close to y_0. Specifically, we choose the neighborhood of y_0 such that $\|f(x_0, y)\| \leq \frac{\delta}{2\|L^{-1}\|}$. This ensures the second term is also bounded by $\delta/2$. Combining the bounds, we get $\|\phi(x, y) - x_0\| \leq \delta$.

The closed ball $\bar{B}_\delta(x_0)$ is a complete metric space. Since $\phi(\cdot, y)$ is a contraction mapping $\bar{B}_\delta(x_0)$ to itself, the Banach Fixed-Point Theorem guarantees a unique fixed point $g(y)$. The proof of the C^k smoothness of g follows from the same inductive argument as in the finite-dimensional case, using the chain rule for Fréchet derivatives. □

Example 5 (Non-linear Integral Equations)**.** *Consider the non-linear integral equation for an unknown function $u(x)$ on the interval $[0, 1]$:*

$$u(x) - \varepsilon \int_0^1 e^{x-y} \sin(u(y))dy = g(x)$$

where $g(x)$ is a given continuous function and ε is a small parameter. We can use the Implicit Function Theorem to show that for any given g, a unique solution $u(x)$ exists for all sufficiently small ε.

Proof. We formulate this as a root-finding problem in a Banach space. Let $X = C([0, 1])$ be the Banach space of continuous functions on $[0, 1]$ equipped with the supremum norm, $\|u\|_\infty = \sup_{x\in[0,1]} |u(x)|$. Our variable is the function $u \in X$ and the parameter is $\varepsilon \in \mathbb{R}$.

Let's define a kernel $K(x, y) = e^{x-y}$ and a non-linear operator $S(u)(y) = \sin(u(y))$. We can define an operator $F : X \times \mathbb{R} \to X$ by

$$F(u, \varepsilon) = u - \varepsilon \int_0^1 K(x, y)S(u)(y)dy - g.$$

We are looking for a solution (u, ε) to the equation $F(u, \varepsilon) = 0$.

For the unperturbed case $\varepsilon_0 = 0$, the equation becomes $F(u, 0) = u - g = 0$, which has the obvious and unique solution $u_0 = g$. Our starting point is thus $(u_0, \varepsilon_0) = (g, 0)$.

To apply the Implicit Function Theorem, we must compute the Fréchet derivative of F with respect to the variable u at the point $(g, 0)$ and check if it is invertible. The derivative $D_u F(g, 0)$ is a linear operator $L : X \to X$ that satisfies

$$F(g + h, 0) = F(g, 0) + L(h) + o(\|h\|).$$

Let's compute the left-hand side:

$$F(g + h, 0) = (g + h) - (0) \int_0^1 \ldots dy - g = h.$$

Since $F(g, 0) = 0$, we have $h = L(h) + o(\|h\|)$. The only linear operator that satisfies this for all h is the identity operator, $L(h) = h$. Thus,

$$D_u F(g, 0) = I.$$

The identity operator I on any Banach space is always invertible (its inverse is itself), and it is a bounded linear operator. Therefore, the central condition of the Implicit Function Theorem is satisfied.

The theorem guarantees that there exist a neighborhood of $\varepsilon = 0$ and a neighborhood of $u_0 = g$ in $C([0, 1])$, and a unique, smooth map $\varepsilon \mapsto u(\varepsilon)$ such that $F(u(\varepsilon), \varepsilon) = 0$. This proves the existence and uniqueness of a continuous solution $u(x)$ for all sufficiently small values of the parameter ε. $\square$

Consider the space $l^\infty(\mathbb{Z})$ of all bounded sequences $X = (x_j)_{j \in \mathbb{Z}}$, which is a Banach space under the supremum norm. For each sequence $X \in l^\infty(\mathbb{Z})$ and a real number $D \geq 0$, we consider the function $G : l^\infty(\mathbb{Z}) \times \mathbb{R} \to l^\infty(\mathbb{Z})$ defined by:

$$(G(X, D))_j = 2x_j - x_j^3 + D(x_{j-1} - 2x_j + x_{j+1})$$

This function can be viewed as defining a dynamical system. Starting with an initial sequence $X^{(0)}$, the state of the system evolves according to the rule $X^{(n+1)} = G(X^{(n)}, D)$.

The operator G models a reaction-diffusion process on a discrete lattice.

1. The **reaction** term, $2x_j - x_j^3$, describes the local dynamics at each site, representing the interaction of the state x_j with itself.

2. The **diffusion** term, $D(x_{j-1} - 2x_j + x_{j+1})$, models the interaction of the state at site j with its nearest neighbors. This term is a discrete version of the second derivative, which governs diffusion processes.

A central question for such systems is to understand the long-term behavior of an initial state $X^{(0)}$ by studying the limit of the sequence of iterates $X^{(n)} = G^n(X^{(0)}, D)$ as $n \to \infty$. The most fundamental long-term outcomes are **stationary states**, or fixed points, where the system ceases to evolve, ie. $G(X,0) = X$. The existence of such states is therefore a primary question of interest.

In the uncoupled case when $D = 0$, we can explicitly construct a stationary solution representing a front that connects the system's two stable states (-1 and $+1$). This solution is given by:

$$X_j^{(0)} = \begin{cases} -1 & \text{if } j \leq 0 \\ 1 & \text{if } j > 0 \end{cases}$$

The precise problem is to determine if this non-trivial spatial structure can persist when the coupling is introduced. Mathematically, we ask: for a small dispersal rate $D > 0$, does there exist a stationary solution $X^{(D)}$ that is a small perturbation of the unperturbed front $X^{(0)}$?

We formulate this as a root-finding problem by defining the function $f : l^\infty(\mathbb{Z}) \times \mathbb{R} \to l^\infty(\mathbb{Z})$ as

$$f(X,D) = G(X,D) - X$$

We seek a solution to $f(X,D) = 0$ that continues from the known solution $(X^{(0)}, D^{(0)}) = (X^{(0)}, 0)$. We apply the Implicit Function Theorem on Banach Spaces (Theorem 6). To do so, we must verify its central hypothesis: that the partial Fréchet derivative of f with respect to X, evaluated at $(X^{(0)}, 0)$, is an isomorphism.

To find the Fréchet derivative of the function $f(X,D) = G(X,D) - X$ with respect to X at the point $(X^{(0)}, 0)$, we use the limit definition. We must find a bounded linear operator L such that the remainder term satisfies the required limit condition. For the fixed parameter value $D = 0$, the function is $f(X,0) = X - X^3$, where the cube operation is component-wise.

We first evaluate the function at the perturbed point $X^{(0)} + H$, where $H \in l^\infty(\mathbb{Z})$ is a perturbation with small norm. The j-th component is given by

$$[f(X^{(0)} + H, 0)]_j = (x_j^{(0)} + h_j) - (x_j^{(0)} + h_j)^3$$

Using the binomial expansion and the facts that $(x_j^{(0)})^3 = x_j^{(0)}$ and $(x_j^{(0)})^2 = 1$ for all $j \in \mathbb{Z}$, the expression simplifies to

$$[f(X^{(0)}+H,0)]_j = (x_j^{(0)}+h_j)-(x_j^{(0)}+3x_j^{(0)2}h_j+3x_j^{(0)}h_j^2+h_j^3) = -2h_j-3x_j^{(0)}h_j^2-h_j^3$$

From this expansion, we identify the part that is linear in H, which is $-2H$. We therefore propose that the Fréchet derivative is the linear operator L defined by $L(H) = -2H$.

We now verify this by showing that the remainder term, $R(H) = f(X^{(0)} + H, 0) - f(X^{(0)}, 0) - L(H)$, is $o(\|H\|_\infty)$. Since $f(X^{(0)}, 0) = 0$ and $L(H) = -2H$, the remainder is

$$R(H) = (-2H - 3M_{X^{(0)}}H^2 - H^3) - (-2H) = -3M_{X^{(0)}}H^2 - H^3$$

where $M_{X^{(0)}}$ is multiplication by the sequence $X^{(0)}$ and H^k denotes component-wise powers. We bound its norm using the triangle inequality:

$$\|R(H)\|_\infty = \sup_j |-3x_j^{(0)}h_j^2 - h_j^3| \le 3\|H\|_\infty^2 + \|H\|_\infty^3$$

To check the limit condition for the derivative, we evaluate

$$\lim_{\|H\|_\infty \to 0} \frac{\|R(H)\|_\infty}{\|H\|_\infty} \le \lim_{\|H\|_\infty \to 0} \frac{3\|H\|_\infty^2 + \|H\|_\infty^3}{\|H\|_\infty} = \lim_{\|H\|_\infty \to 0} (3\|H\|_\infty + \|H\|_\infty^2) = 0$$

Since the limit is zero, our proposed linear operator $L(H) = -2H$ is indeed the Fréchet derivative. The derivative $D_X f(X^{(0)}, 0)$ is the operator $L = -2I$.

Since the derivative $D_X f(X^{(0)}, 0)$ is an isomorphism, the conditions of the theorem are satisfied. The Implicit Function Theorem guarantees the existence of a unique, continuous function $X(D)$ for D in a neighborhood of 0, such that $X(0) = X^{(0)}$ and $f(X(D), D) = 0$. This function $X(D)$ is the stationary front solution that persists from the unperturbed case.