Theorem 1 (Inverse Function Theorem) Suppose $U \subset \mathbb{R}^n$ is open, $f : U \to \mathbb{R}^n$ is C^1, $x_0 \in U$ and df_{x_0} is invertible. Then there exists a neighborhood V of x_0 in U and a neighborhood W of $f(x_0)$ in $\mathbb{R}^n$ such that f has a C^1 inverse $g = f^{-1} : W \to V$. (Thus $f(g(y)) = y$ for all $y \in W$ and $g(f(x)) = x$ for all $x \in V$.) Moreover

$$\left[\frac{\partial g_i}{\partial y_j}(y)\right]_{1\le i,j\le n} = dg_y = (df_{g(y)})^{-1} = \left[\frac{\partial f_i}{\partial x_j}(g(y))\right]^{-1}_{1\le i,j\le n} \quad \text{for all} \quad y \in W$$

and g is smooth whenever f is smooth.

Remark 1. The theorem says that a continuously differentiable function f between regions in $\mathbb{R}^n$ is locally invertible near points where its differential is invertible,
i.e. $W = \{y = (y_1(x), \ldots, y_n(x)) = (f_1(x), \ldots, f_n(x)) = f(x) \mid x \in V\}$,
$V = \{x = (x_1(y), \ldots, x_n(y)) = (g_1(y), \ldots, g_n(y)) = g(y) \mid y \in W\}$,
and each coordinate function is continuously differentiable.

Remark 2. Let $0 < a < 1$, define $f : \mathbb{R} \to \mathbb{R}$ by $f(x) = \begin{cases} ax + x^2 \sin\dfrac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$. Compute $f'(x)$ for all $x \in \mathbb{R}$. Show that $f'(0) > 0$, yet f is not one-to-one in any neighborhood of 0 [by using
(1) the fact that there exists a sequence of points $\{x_n \to 0\}$ at which $f'(x_n) = 0$, and $f''(x_n) \neq 0$,
(2) the observation that if $f'(p) = 0$, $f''(p) \neq 0$, then f cannot be one-to-one near p.].
This example shows that in the Inverse Function Theorem, the hypothesis that f is C^1 cannot be weaken to the hypothesis that f is differentiable.

Remark 3. Define $f : \mathbb{R}^2 \to \mathbb{R}^2$ by $f(x, y) = (e^x \cos y, e^x \sin y)$. Show that f is C^1 and $df_{(x,y)}$ is invertible for all $(x, y) \in \mathbb{R}^2$ and yet f is not one-to-one function globally. Why doesn't this contradict the Inverse Function Theorem?

Example 1. Use Inverse Function Theorem to determine whether the system

$$\begin{aligned} u(x, y, z) &= x + xyz \\ v(x, y, z) &= y + xy \\ w(x, y, z) &= z + 2x + 3z^2 \end{aligned}$$

can be solved for x, y, z in terms of u, v, w near $p = (0, 0, 0)$.

Solution: Set $F(x, y, z) = (u, v, w)$. Then

$$DF(p) = \begin{bmatrix} u_x & u_y & u_z \\ v_x & v_y & v_z \\ w_x & w_y & w_z \end{bmatrix}(p) = \begin{bmatrix} 1+yz & xz & xy \\ y & 1+x & 0 \\ 2 & 0 & 1+6z \end{bmatrix}(p) = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{bmatrix} \text{ and } \begin{vmatrix} 1 & 0 & 0 \\ 0 & 1 & 0 \\ 2 & 0 & 1 \end{vmatrix} = 1 \neq 0.$$

By the Inverse Function Theorem, the inverse $F^{-1}(u, v, w)$ exists near $p = (0, 0, 0)$, i.e. we can solve x, y, z in terms of u, v, w near $p = (0, 0, 0)$.

Theorem 2 (Implicit Function Theorem) Let $U \subset \mathbb{R}^{m+n} \equiv \mathbb{R}^m \times \mathbb{R}^n$ be an open set, $f = (f_1, \ldots, f_n) : U \to \mathbb{R}^n$ a C^1 function, $(a, b) \in U$ a point such that $f(a, b) = 0$ and the $n \times n$ matrix $d_y f|_{(a,b)} = \left[\dfrac{\partial f_i}{\partial y_j}(a, b)\right]_{1\le i,j\le n}$ is invertible. Then there exists a neighborhood V of (a, b) in U a neighborhood W of a in $\mathbb{R}^m$ and a C^1 function $g : W \to \mathbb{R}^n$ such that

$$\{(x, y) \in V \subset \mathbb{R}^m \times \mathbb{R}^n \mid f(x, y) = 0\} = \{(x, g(x)) \mid x \in W\} = \text{the graph of } g \text{ over } W$$

i.e. Letting $S = \{(x, y) \in \mathbb{R}^m \times \mathbb{R}^n \mid f(x, y) = 0\}$ denote the solution set, then $S \cap V$ is of the dimension m, and it is equal to graph(g), the graph of a C^1 function g, over W. Moreover

$$dg_x = -(d_y f)^{-1}|_{(x,g(x))} d_x f|_{(x,g(x))}$$

and g is smooth whenever f is smooth.
Example 2. Let $F:\mathbb{R}^4\to\mathbb{R}^2$ be given by $F(x,y,z,w)=(G(x,y,z,w),H(x,y,z,w))$ $=(y^2+w^2-2xz,y^3+w^3+x^3-z^3)$, and let $p=(1,-1,1,1)$.

(a) Show that we can solve $F(x,y,z,w)=(0,0)$ for (x,z) in terms of (y,w) near $(-1,1)$.

Solution: Since $DF(p)=\begin{bmatrix}G_x & G_y & G_z & G_w\\ H_x & H_y & H_z & H_w\end{bmatrix}(p)=\begin{bmatrix}-2 & -2 & -2 & 2\\ 3 & 3 & -3 & 3\end{bmatrix}$ and $\begin{vmatrix}G_x & G_z\\ H_x & H_z\end{vmatrix}(p)=\begin{vmatrix}-2 & -2\\ 3 & -3\end{vmatrix}=12\neq 0$, we can write (x,z) in terms of (y,w) near $(-1,1)$ by Implicit Function Theorem.

(b) If $(x,z)=\Phi(y,w)$ is the solution in part (a), show that $D\Phi(-1,1)$ is given by the matrix

$$-\begin{bmatrix}-2 & -2\\ 3 & -3\end{bmatrix}^{-1}\begin{bmatrix}-2 & 2\\ 3 & 3\end{bmatrix}=\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}$$

Solution: The Implicit Function Theorem implies that, near p,the solution set $\{(x,y,z,w)\mid F(x,y,z,w)=(0,0)\}$ is the graph of $(x,z)=\Phi(y,w)$ near $(-1,1)$.
Hence, we have $\dfrac{\partial F}{\partial y}=(0,0)$, and $\dfrac{\partial F}{\partial w}=(0,0)$ near $(-1,1)$.
Therefore, $0=G_x\dfrac{\partial x}{\partial y}+G_y+G_z\dfrac{\partial z}{\partial y}$, and $0=G_x\dfrac{\partial x}{\partial w}+G_z\dfrac{\partial z}{\partial w}+G_w$,
which implies that $-[G_y,G_w]=[G_x,G_z]\begin{bmatrix}\dfrac{\partial x}{\partial y} & \dfrac{\partial x}{\partial w}\\ \dfrac{\partial z}{\partial y} & \dfrac{\partial z}{\partial w}\end{bmatrix}$.
Similarly, we have $-[H_y,H_w]=[H_x,H_z]\begin{bmatrix}\dfrac{\partial x}{\partial y} & \dfrac{\partial x}{\partial w}\\ \dfrac{\partial z}{\partial y} & \dfrac{\partial z}{\partial w}\end{bmatrix}$.
Thus, we have $-\begin{bmatrix}G_y & G_w\\ H_y & H_w\end{bmatrix}=\begin{bmatrix}G_x & G_z\\ H_x & H_z\end{bmatrix}\begin{bmatrix}\dfrac{\partial x}{\partial y} & \dfrac{\partial x}{\partial w}\\ \dfrac{\partial z}{\partial y} & \dfrac{\partial z}{\partial w}\end{bmatrix}$
or $D\Phi=\begin{bmatrix}\dfrac{\partial x}{\partial y} & \dfrac{\partial x}{\partial w}\\ \dfrac{\partial z}{\partial y} & \dfrac{\partial z}{\partial w}\end{bmatrix}=-\begin{bmatrix}G_x & G_z\\ H_x & H_z\end{bmatrix}^{-1}\begin{bmatrix}G_y & G_w\\ H_y & H_w\end{bmatrix}$
Hence, $D\Phi(-1,1)$ is given by the matrix

$$-\begin{bmatrix}-2 & -2\\ 3 & -3\end{bmatrix}^{-1}\begin{bmatrix}-2 & 2\\ 3 & 3\end{bmatrix}=\begin{bmatrix}-1 & 0\\ 0 & 1\end{bmatrix}$$

Remark 4. The theorem says that if $f:U\subset\mathbb{R}^{m+n}\to\mathbb{R}^n$ is a map of the class $C^1(U)$, (a,b) is a point in U, and $d_yf|_{(a,b)}=\left[\dfrac{\partial f_i}{\partial y_j}(a,b)\right]_{1\le i,j\le n}$ is invertible, then, **locally, the solution set** $S=\{(x,y)\in U\mid f(x,y)=f(a,b)\}$ **is a** C^1 **"manifold" of dimension** m.

Remark 5. Let $f:U\subset\mathbb{R}^{m+n}\to\mathbb{R}^n$ **is a map of the class** C^1 **on an open subset** U **of** $\mathbb{R}^{(m+n)}$**, and assume that the differential,** $df_p=\left[\dfrac{\partial f_i}{\partial x_j}\right]_{1\le i\le n;1\le j\le (m+n)}$**, is of the constant rank** k **at**

each $p \in U$. By relabeling, if necessary, we may assume that $\hat{f} = (f_1, \ldots, f_k) : U \subset \mathbb{R}^{m+n} \to \mathbb{R}^k$ is a map with its differential $d\hat{f}_p = \left[\dfrac{\partial f_i}{\partial x_j}\right]_{1\le i\le k; 1\le j\le (m+n)}$,of rank k. The Implicit Function Theorem says that locally the solution set $S = \{x \in U \mid \hat{f}(x) = 0\}$ is a C^1 "manifold" of dimension $(m+n) - k$. Identify S with an open neighborhood $W \subset \mathbb{R}^{(m+n)-k}$ of $0 \in \mathbb{R}^{(m+n)-k}$ and identify U with $V \times W$, such that we write each $x \in V \times W \subset \mathbb{R}^{(m+n)}$ as $x = (\bar{x}, \hat{x})$, where $\bar{x} = (\bar{x}_1, \ldots, \bar{x}_k) \in V$, and $\hat{x} = (\hat{x}_{k+1}, \ldots, \hat{x}_{m+n}) \in W$. This implies $\hat{f}(0, \hat{x}) = 0$, and $\left[\dfrac{\partial \hat{f}_i}{\partial \bar{x}_j}\right]_{1\le i,j\le k}$ is invertible on $V \times W$. **Using the Inverse Function Theorem, we may show that the range $\hat{f}(V \times W) = f(U)$ locally is a $k-$dimensional "manifold".**

Remark 6. Two mappings are said to be locally equivalent if under suitable choices of local coordinate systems in the domain and range spaces (with origin at 0) they can be written by the same formulas. The Implicit Function Theorem says that if $f, g : U \subset \mathbb{R}^{m+n} \to \mathbb{R}^n$ is a map of the class $C^1(U)$, p is a point in U, and $\mathrm{rk}[df(p)] = \mathrm{rk}[dg(p)] = n$, then f and g are equivalent, i.e. there exists diffeomorphisms $h : \mathbb{R}^{m+n} \to \mathbb{R}^{m+n}$, $k : \mathbb{R}^n \to \mathbb{R}^n$, such that $f \circ h = k \circ g$.

Definition 1. A map $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ is called a Lipschitz map on U if there exists a constant $C \ge 0$ such that

$$\|f(x) - f(y)\| \le C\|x - y\| \quad \text{for all } x, y \in U.$$

If one can choose a (Lipschitz) constant $C < 1$ such that the above Lipschitz condition hold on U, then f is called a contraction map.

Example 3. Let U be a convex subset of $\mathbb{R}^m$, and $f : U \subset \mathbb{R}^m \to \mathbb{R}^n$ be a map with bounded $\|df\| = \sup\{\dfrac{\|df_p(x)\|_{\mathbb{R}^n}}{\|x\|_{\mathbb{R}^m}} \mid p \in U, x \ne 0\}$ (which implies that df_p is an $n \times m$ matrix, so if $\|df\| \le M$ then $\|\nabla f_i\| \le M$ for $i = 1, \ldots, m$.). Then f is Lipscitz on U.

Definition 2. Let U be a subset of $\mathbb{R}^m$, and f be a map that maps U into U, i.e. $f : U \to U$. A point $p \in U$ is said to be a fixed point of f if $f(p) = p$.

Theorem 3. (Fixed point theorem for contractions) Let $f : \mathbb{R}^m \to \mathbb{R}^m$ be a contraction map. Then f has a fixed point.
Note that $\mathbb{R}^m$ is complete which implies that any Cauchy sequence (is bounded and has a limiting point by Bolzano-Weierstrass theorem, and Cauchy condition implies that it) converges.

Definition 3. (Uniform Boundedness) Let K be a compact subset of $\mathbb{R}^p$, and $C_{pq}(K) = BC_{pq}(K) = \{f : K \subset \mathbb{R}^p \to \mathbb{R}^q\}$ denotes the set of all continuous (and bounded) functions from K into $\mathbb{R}^q$. We say that a set $\mathscr{F} \subset C_{pq}(K)$ is bounded (or uniformly bounded) on K if there exists a constant M such that $\|f\|_K = \sup\{f(x) \mid x \in K\} \le M$, for all $f \in \mathscr{F}$.

Definition 4. (Uniform Equicontinuity) A set $\mathscr{F}$ of functions on K to $\mathbb{R}^q$ is said to be uniformly equicontinuous on K if, for each real number $\epsilon > 0$ there is a number $\delta(\epsilon) > 0$ such that if x, y belong to K and $\|x - y\| < \delta(\epsilon)$ and f is a function in $\mathscr{F}$, then $\|f(x) - f(y)\| < \epsilon$

Definition 5. (Uniform Convergence) A sequence (f_n) of functions on $U \subset \mathbb{R}^p$ to $\mathbb{R}^q$ converges uniformly on a subset $D \subset U$ to a function f if for each $\epsilon > 0$ there is a natural number $L(\epsilon)$ such that for all $n \geq L(\epsilon)$ and $x \in D$, then

$$\|f_n(x) - f(x)\| < \epsilon.$$

In this case, we say that the sequence is uniformly convergent on D. It is immediate from the definition that a sequence $f_n \in B_{pq}(D) = \{$the set of bounded functions from $D \subset \mathbb{R}^p$ to $\mathbb{R}^q\}$ converges uniformly to $f \in B_{pq}(D)$ on D if and only if $\lim_{n\to\infty} \|f_n - f\|_D = \lim_{n\to\infty} \sup\{\|f_n(x) - f(x)\| \mid x \in D\} = 0$.

Examples 4. **(a)** For each $n \in \mathbb{N}$, let $f_n : \mathbb{R} \to \mathbb{R}$ be defined by $f_n(x) = \dfrac{x}{n}$ for each $x \in \mathbb{R}$.
(b) For each $n \in \mathbb{N}$, let $f_n : I = [0,1] \subset \mathbb{R} \to \mathbb{R}$ be defined by $f_n(x) = x^n$ for each $x \in I$.
(c) For each $n \in \mathbb{N}$, let $f_n : \mathbb{R} \to \mathbb{R}$ be defined by $f_n(x) = \dfrac{x^2 + nx}{n}$ for each $x \in \mathbb{R}$.
(d) For each $n \in \mathbb{N}$, let $f_n : \mathbb{R} \to \mathbb{R}$ be defined by $f_n(x) = \dfrac{1}{n}\sin(nx + n)$ for each $x \in \mathbb{R}$.
One can easily see that the limiting functions are $f \equiv 0$, $f(x) = \begin{cases} 0, & 0 \leq x < 1 \\ 1, & x = 1 \end{cases}$, $f(x) = x$, and $f \equiv 0$, for (a), (b), (c), and (d), respectively, and only the convergence in (d) is uniform.

Theorem 4. (Arzela-Ascoli Theorem). Let K be a compact subset of $\mathbb{R}^p$ and let $\mathscr{F}$ be a collection of functions which are continuous on K and have values in $\mathbb{R}^q$. The following properties are equivalent:
(a) The family $\mathscr{F}$ is bounded and uniformly equicontinuous on K.
(b) Every sequence from $\mathscr{F}$ has a subsequence which is uniformly convergent on K.

Example 5. Consider the following sequences of functions which show that **Arzela-Ascoli Theorem** may fail if the various hypothesis are dropped.
(a) $f_n(x) = x + n \quad \text{for} \quad x \in [0,1]$
(b) $f_n(x) = x^n \quad \text{for} \quad x \in [0,1]$
(c) $f_n(x) = \dfrac{1}{1 + (x-n)^2} \quad \text{for} \quad x \in [0,\infty)$

Example 6. Let $f_n : [0,1] \to \mathbb{R}$ be continuous and be such that $|f_n(x)| \leq 100$ for every n and for all $x \in [0,1]$ and the derivatives $f_n'(x)$ exist and are uniformly bounded on $(0,1)$.
(a) Show that there is a constant M such that $|\, f_n(x) - f_n(y) \,| \leq M \,|\, x - y \,|$ for any $x, y \in [0,1]$ and any $n \in \mathbb{N}$.
Solution: Let M be a constant such that $|f_n'(x)| \leq M$ for all $x \in (0,1)$. By the mean value theorem, we get $|\, f_n(x) - f_n(y) \,| \leq M \,|\, x - y \,|$ for any $x, y \in [0,1]$ and any $n \in \mathbb{N}$.

(b) Prove that f_n has a uniformly convergent subsequence.
Solution: We apply the Arzela-Ascoli Theorem by verifying that $\{f_n\}$ is equicontinuous and bounded. Given ϵ, we can choose $\delta = \epsilon/M$, independent of x, y, and n. Thus $\{f_n\}$ is equicontinuous. It is bounded because $\|f_n\| = \sup_{x\in[0,1]} |f_n(x)| \leq 100$.

Example7. Let the functions $f_n : [a,b] \to \mathbb{R}$ be uniformly bounded continuous functions. Set $F_n(x) = \displaystyle\int_a^x f_n(t)\,dt$, for $x \in [a,b]$. Prove that F_n has a uniformly convergent subsequence.
Solution: Since $\|F_n\| \leq \|f_n\|(b-a)$, F_n is uniformly bounded. Also, since $|F_n'(x)| \leq \|f_n\|$, F_n is equicontinuous by the preceding result. Therefore, F_n has a uniformly convergent subsequence by Arzela-Ascoli Theorem.