$f^{-1}: W \mapsto V$ is a C^1 function.

We have already proved the existence of open sets V, W such that $f: V \mapsto W$ is injective and surjective. Thus the inverse function $f^{-1}: W \mapsto V$ is well-defined. In the following we present two proofs of its differentiability.

Proof 1: Direct proof of total differentiability.

Let $y_0 \in W$ be arbitrary. Let $x_0 := f^{-1}(y_0)$. The goal is to show

$$\lim_{y \longrightarrow y_0, y \neq y_0} \frac{\left\| f^{-1}(y) - f^{-1}(y_0) - J_f^{-1}(x_0)\,(y - y_0) \right\|}{\|y - y_0\|} = 0. \tag{1}$$

Setting $x := f^{-1}(y)$ we can re-write this as

$$\lim_{y \longrightarrow y_0, y \neq y_0} \frac{\|x - x_0\|}{\|y - y_0\|} \frac{\left\| x - x_0 - J_f^{-1}(x_0)\,(f(x) - f(x_0)) \right\|}{\|x - x_0\|} = 0. \tag{2}$$

(2) clearly follows from the following.

$$\lim_{y \longrightarrow y_0, y \neq y_0} \left\| J_f^{-1}(x_0) \right\| \frac{\|x - x_0\|}{\|y - y_0\|} \frac{\| f(x) - f(x_0) - J_f(x_0)\,(x - x_0) \|}{\|x - x_0\|} = 0. \tag{3}$$

As $J_f^{-1}(x_0)$ is a constant matrix, independent of y, we only need to show

$$\lim_{y \longrightarrow y_0, y \neq y_0} \frac{\|x - x_0\|}{\|y - y_0\|} \frac{\| f(x) - f(x_0) - J_f(x_0)\,(x - x_0) \|}{\|x - x_0\|} = 0 \tag{4}$$

where keep in mind that $x := f^{-1}(y)$ and $x_0 := f^{-1}(y_0)$.

Since

$$\lim_{x \to x_0, x \neq x_0} \frac{\| f(x) - f(x_0) - J_f(x_0)\,(x - x_0) \|}{\|x - x_0\|} = 0. \tag{5}$$

it suffices to prove the following:

$$\exists \varepsilon > 0, C > 0 \qquad \frac{\|x - x_0\|}{\|y - y_0\|} = \frac{\|x - x_0\|}{\| f(x) - f(x_0) \|} \leqslant C \text{ for all } y \in B_\varepsilon(y_0), y \neq y_0. \tag{6}$$

We now prove (6). Notice that

$$\|x - x_0\| = \| J_f(x_0)^{-1} J_f(x_0)\,(x - x_0) \| \leqslant \| J_f(x_0)^{-1} \| \, \| J_f(x_0)\,(x - x_0) \|. \tag{7}$$

Setting $C := 2\,\| J_f(x_0)^{-1} \|$, we have

$$\|x - x_0\| \leqslant \frac{C}{2} \| J_f(x_0)\,(x - x_0) \|. \tag{8}$$

We note that

$$\begin{aligned} \| J_f(x_0)\,(x - x_0) \| &\leqslant \| f(x) - f(x_0) \| + \| f(x) - f(x_0) - J_f(x_0)\,(x - x_0) \| \\ &= \|y - y_0\| + \| f(x) - f(x_0) - J_f(x_0)\,(x - x_0) \|. \end{aligned} \tag{9}$$

Now thanks to (5), there is $r > 0$ such that whenever $\tilde{x} \in B_r(x_0)$, there holds

$$\| f(\tilde{x}) - f(x_0) - J_f(x_0)\,(\tilde{x} - x_0) \| < \frac{1}{C} \|\tilde{x} - x_0\|. \tag{10}$$

Recalling that $f(B_r(x_0))$ is open, there is $\varepsilon > 0$ such that $B_\varepsilon(y_0) \subseteq f(B_r(x_0))$. For every $y \in B_\varepsilon(y_0), y \neq y_0$, following (8–10), we conclude

$$\begin{aligned} \|x - x_0\| &\leqslant \frac{C}{2} \left[\|y - y_0\| + \frac{1}{C} \|x - x_0\| \right] \\ &= \frac{C}{2} \| f(x) - f(x_0) \| + \frac{1}{2} \|x - x_0\|, \end{aligned} \tag{11}$$

and (6) follows.

Proof 2: Detour proof through partial derivatives.

We have shown that there are bounded open sets V, W such that $f: V \mapsto W$ has an inverse function f^{-1}: $W \mapsto V$.

Now let $y_0 \in W$. Take $\varepsilon > 0$ such that $\overline{B_\varepsilon(y_0)} \subseteq W$. In the following we denote $x = f^{-1}(y)$ and $x_0 = f^{-1}(y_0)$. Then we have

$$f(x) - f(x_0) = f(f^{-1}(y)) - f(f^{-1}(y_0)) = y - y_0. \tag{12}$$

Application of mean value theorem to every f_i, we see that

$$\tilde{J} \cdot (f^{-1}(y) - f^{-1}(y_0)) = y - y_0 \tag{13}$$

where

$$\tilde{J} = \left(\frac{\partial f_i}{\partial x_j}(\xi_i) \right) \tag{14}$$

with every ξ_i lying on the line segment connecting $x = f^{-1}(y)$ and $x_0 = f^{-1}(y_0)$.

As $f \in C^1$, by taking ε small enough we can guarantee the existence of a constant $C > 0$ such that $\|\tilde{J}^{-1}\| < C$ independent of the points $\xi_1, \ldots, \xi_N$. Consequently

$$y \longrightarrow y_0 \Longrightarrow f^{-1}(y) \longrightarrow f^{-1}(y_0) \Longrightarrow \tilde{J} \longrightarrow J_f(x_0). \tag{15}$$

Now take $y = y_0 + \delta e_i$ and let $\delta \longrightarrow 0$, we see that all the partial derivatives exist for f^{-1} and furthermore

$$J_{f^{-1}}(y_0) = J_f(x_0)^{-1}. \tag{16}$$

By arbitrariness of y_0 we have

$$J_{f^{-1}}(y) = J_f(f^{-1}(y))^{-1} \tag{17}$$

for all $y \in W$. As J_f and f^{-1} are both continuous, $J_g(y)$ is continuous. Consequently g is differentiable.

Exercise 1. Prove the following.

a) Let $\varnothing \neq U \subseteq \mathbb{R}^N$ be open, let $M \geqslant N$, and let $f \in C^1(U, \mathbb{R}^M)$ be such that $\operatorname{rank} J_f(x) = N$ for all $x \in U$. Then f is locally injective on U.

b) If "$\operatorname{rank} J_f(x) = N$ for all $x \in U$" is replaced by $\operatorname{rank} J_f(x_0) = N$, then there is $V \supset x_0$ such that f is injective on V.

c) Let $\varnothing \neq U \subset \mathbb{R}^N$ be open, let $M \leqslant N$, and let $f \in C^1(U, \mathbb{R}^M)$ with $\operatorname{rank} J_f(x) = M$ for all $x \in U$. Then $f(U)$ is open.

Exercise 2. Prove that if $f \in C^k$ for some $k > 1$, then $f^{-1} \in C^k$ for the same k.

Proof of the implicit function theorem.

Let $\varnothing \neq U \subseteq \mathbb{R}^{M+N}$ be open, let $f \in C^1(U, \mathbb{R}^N)$, and let $(x_0, y_0) \in U$ be such that $f(x_0, y_0) = 0$ and $\det \frac{\partial f}{\partial y}(x_0, y_0) \neq 0$.

We define

$$F: \mathbb{R}^M \times U \mapsto \mathbb{R}^{M+N}, \qquad F(x, y) := \begin{pmatrix} x \\ f(x, y) \end{pmatrix}. \tag{18}$$

Simple calculation gives

$$J_F(x_0, y_0) = \begin{pmatrix} I_M & 0 \\ \frac{\partial f}{\partial x}(x_0, y_0) & \frac{\partial f}{\partial y}(x_0, y_0) \end{pmatrix} \tag{19}$$

which gives

$$\det J_F(x_0, y_0) = \det \frac{\partial f}{\partial y}(x_0, y_0). \tag{20}$$

It is clear that $F \in C^1(\mathbb{R}^M \times U, \mathbb{R}^{M+N})$.

Thus by Theorem ? there is an open set $S \subseteq \mathbb{R}^M \times U$ such that $(x_0, y_0) \in S$, $T := F(S)$ is open, and F has a C^1 inverse function $G: T \mapsto S$. Checking the proof of Theorem ? we see that S can be taken as $V \times W$ where $V \subseteq \mathbb{R}^M, W \subseteq \mathbb{R}^N$ are open sets with $x_0 \in V, y_0 \in W$.

Now denote

$$G(z) = \begin{pmatrix} g_1(z) \\ \vdots \\ g_{M+N}(z) \end{pmatrix} \text{ where } z \in \mathbb{R}^{M+N} \tag{21}$$

and define

$$\phi(x) = \begin{pmatrix} g_{M+1}(x,0) \\ \vdots \\ g_{M+N}(x,0) \end{pmatrix}. \tag{22}$$

Note that as $(x_0, 0) = F(x_0, y_0) \in T$, the set $\{x \in V \mid (x,0) \in T\}$ is open. We now replace V by this set to make ϕ well-defined on V.

It is clear that $\phi \in C^1(V, W)$ and furthermore $J_\phi(x) = -\left(\frac{\partial f}{\partial y}\right)^{-1}\left(\frac{\partial f}{\partial x}\right)$. All we need to show now is $f(x, \phi(x)) = 0$.

We have

$$\begin{pmatrix} x \\ f(x, \phi(x)) \end{pmatrix} = F(x, \phi(x)) = F(x_1, ..., x_M, g_{M+1}(x,0), ..., g_{M+N}(x,0)). \tag{23}$$

All we need to show now is that $g_i(x,0) = x_i$, $i = 1, 2, ..., M$. Let $G(x,0) = \begin{pmatrix} x' \\ y' \end{pmatrix}$. Then we have

$$\begin{pmatrix} x' \\ f(x', y') \end{pmatrix} = F(x', y') = F(G(x,0)) = \begin{pmatrix} x \\ 0 \end{pmatrix}. \tag{24}$$

We see that necessarily $x' = x$. Consequently

$$F(x_1, ..., x_M, g_{M+1}(x,0), ..., g_{M+N}(x,0)) = F(G(x,0)) = \begin{pmatrix} x \\ 0 \end{pmatrix} \tag{25}$$

and it follows from (23) that $f(x, \phi(x)) = 0$.