Lemma 0.1. Let $f\colon \mathbf{R}^n \to \mathbf{R}^n$ be continuously differentiable with $Df(a)$ non-singular for some point $a \in \mathbf{R}^n$. For every $\epsilon > 0$ there exists $r > 0$ such that $B_r(f(a)) \subseteq f(B_\epsilon(a))$.

Proof: Choose $\delta > 0$ such that $\delta < \epsilon$ and

$$\|x - a\| < \delta \quad \text{implies} \quad \|Df(a)^{-1}\|\|Df(x) - Df(a)\| < 1/2.$$

Set $r = \frac{1}{2}\delta/\|Df(a)^{-1}\|$. For $y \in B_r(f(a))$ we solve for x such that $f(x) = y$ using a pseudo Newton method. Let $g(x) = f(x) - y$ and define

$$x_0 = a \quad \text{and} \quad x_{k+1} = x_k - Dg(a)^{-1} g(x_k).$$

Claim that $\|x_k - a\| \le \delta$ and $\|x_{k+1} - x_k\| \le \delta/2^k$ for all $k \in \mathbf{N}$. We prove this claim by induction. For $k = 0$ we have $\|x_0 - a\| = 0 < \delta$ and

$$\|x_1 - x_0\| = \|x_0 - Dg(a)^{-1} g(x_0) - x_0\| \le \|Dg(a)^{-1}\|\|f(a) - y\| \le \|Dg(a)^{-1}\| r = \delta/2.$$

For the induction hypothesis suppose that the claim holds for $k \le K$. Then

$$\|x_{K+1} - a\| \le \sum_{k=0}^{K} \|x_{k+1} - x_k\| \le \delta \sum_{k=0}^{K} \frac{1}{2^{k+1}} = \delta(1 - 1/2^{k+1}) < \delta$$

Let $v_K = g(x_{K+1}) - g(x_K) - Dg(a)(x_{K+1} - x_K)$. By the intermediate value theorem there is c_K lying on the line segment between x_{K+1} and x_K such that

$$(g(x_{K+1}) - g(x_K)) \cdot v_K = Dg(c_K)(x_{K+1} - x_K) \cdot v_K.$$

Therefore,

$$\begin{aligned}\|v_K\|^2 = v_K \cdot v_K &= (g(x_{K+1}) - g(x_K) - Dg(a)(x_{K+1} - x_K)) \cdot v_K \\ &= (Dg(c_K) - Dg(a))(x_{K+1} - x_K) \cdot v_K\end{aligned}$$

implies by the Cauchy inequality that

$$\|g(x_{K+1}) - g(x_K) - Dg(a)(x_{K+1} - x_K)\| \le \|Dg(c_K) - Dg(a)\|\|x_{K+1} - x_K\|.$$

Since $\|c_K - a\| \le \|x_{K+1} - a\| < \delta$ then

$$\begin{aligned}\|x_{K+2} - x_{K+1}\| &\le \|x_{K+1} - Dg(a)^{-1} g(x_{K+1}) - x_K + Dg(a)^{-1} g(x_K)\| \\ &\le \|Dg(a)^{-1}\|\|g(x_{K+1}) - g(x_K) - Dg(a)(x_{K+1} - x_K)\| \\ &\le \|Dg(a)^{-1}\|\|Dg(c_K) - Dg(a)\|\|x_{K+1} - x_K\| \\ &\le \frac{1}{2}\|x_{K+1} - x_K\| \le \delta/2^{K+1}.\end{aligned}$$

This completes the induction and proves the claim. We now show that x_k is a Cauchy sequence. Let $p, q \in \mathbf{N}$ with $p > q$. Then

$$\|x_p - x_q\| \le \sum_{k=q}^{p-1} \|x_{k+1} - x_k\| \le \delta \sum_{k=q}^{p-1} \frac{1}{2^k} = 2\delta\Big(\frac{1}{2^q} - \frac{1}{2^p}\Big) \to 0 \text{ as } p, q \to \infty.$$

Thus, there is $x \in \mathbf{R}^n$ such that $x_k \to x$ as $k \to \infty$. Therefore

$$x = \lim_{k\to\infty} x_{k+1} = \lim_{k\to\infty} x_k - Dg(a)^{-1} g(x_k) = x - Dg(a)^{-1} g(x)$$

implies $g(x) = 0$ or that $f(x) = y$. □

Lemma 0.2. Let $f\colon \mathbf{R}^n \to \mathbf{R}^n$ be continuously differentiable with $Df(a)$ non-singular for some point $a \in \mathbf{R}^n$. Then there exists $\epsilon > 0$ such that f is one-to-one on $B_\epsilon(a)$. Moreover there is $\lambda > 0$ such that

$$\|x_2 - x_1\| < \lambda \|f(x_2) - f(x_1)\| \quad \text{for all} \quad x_1, x_2 \in B_\epsilon(a).$$

Proof: Choose $\epsilon > 0$ so that $x \in B_\epsilon(a)$ implies $\|Df(a)^{-1}\|\|Df(x) - Df(a)\| < 1/2$. Let $x_1, x_2 \in B_\epsilon(a)$, $v_0 = Df(a)(x_2 - x_1)$ and $w_t = Df(x_1 + t(x_2 - x_1))(x_2 - x_1)$. Then

$$\begin{aligned}\|w_t - v_0\| &\le \|Df(x_1 + t(x_2 - x_1)) - Df(a)\|\|x_2 - x_1\| \\ &\le \|Df(x_1 + t(x_2 - x_1)) - Df(a)\|\|Df(a)^{-1}\|\|v_0\| < \tfrac{1}{2}\|v_0\|.\end{aligned}$$

It follows from

$$\|w\|^2 - 2w_t \cdot v_0 + \|v_0\|^2 = \|w_t - v_0\|^2 \le \frac{1}{4}\|v_0\|^2$$

that $2w_t \cdot v_0 \ge \frac{3}{8}\|v_0\|^2$. Let $g(t) = f(x_1 + t(x_2 - x_1)) \cdot v_0$. Then $g'(t) = w_t \cdot v_0$ and

$$\begin{aligned}\|f(x_2) - f(x_1)\|\|v_0\| \ge (f(x_2) - f(x_1)) \cdot v_0 = g(1) - g(0) = \int_0^1 g'(t)dt = \int_0^1 w_t \cdot w_0 \\ \ge \frac{3}{8}\int_0^1 \|v_0\|^2 = \frac{3}{8}\|v_0\|^2.\end{aligned}$$

Therefore

$$\|f(x_2) - f(x_1)\| \ge \frac{3}{8}\|Df(a)(x_2 - x_1)\| \ge \frac{3}{8\|Df(a)^{-1}\|}\|x_2 - x_1\|$$

shows f is one-to-one. Taking $\lambda = (8/3)\|Df(a)^{-1}\|$ finishes the proof. □

Lemma 0.3. Let $V \subseteq \mathbf{R}^n$ be open and $a \in V$. Let $f\colon V \to \mathbf{R}^n$ be one-to-one and continuous and let $b = f(a)$. If f has continuous inverse $f^{-1}\colon B_r(b) \to \mathbf{R}^n$ for some $r > 0$, then

$$\lim_{y \to b} g(y) \quad \text{exists if only if} \quad \lim_{x \to a} g \circ f(x) \quad \text{exists.}$$

Moreover, if the limits exists they are equal.

Proof: "$\longrightarrow$" Suppose $\lim_{y \to b} g(y) = L$ exists. Then for every $\epsilon > 0$ there exists $\delta_1 > 0$ such that $y \in \mathcal{D}(g)$ and $0 < \|y - b\| < \delta_1$ implies $\|g(y) - L\| < \epsilon$. Since f is continuous there is $\delta > 0$ such that $x \in V$ and $\|x - a\| < \delta$ implies $\|f(x) - b\| < \delta_1$. Moreover, since f is one-to-one $0 < \|x - a\|$ implies $0 < \|f(x) - b\|$. It follows that $0 < \|x - a\| < \delta$ and $x \in \mathcal{D}(g \circ f)$ imply $f(x) \in \mathcal{D}(g)$ and therefore $\|g \circ f(x) - L\| < \epsilon$.

"$\longleftarrow$" Suppose $\lim_{x \to a} g \circ f(x) = L$ exists. By the previous part $\lim_{y \to b} g \circ f \circ f^{-1}(y) = L$. Therefore for $\epsilon > 0$ there is $\delta_1 > 0$ such that $y \in \mathcal{D}(g \circ f \circ f^{-1})$ and $0 < \|y - b\| < \delta_1$ implies $\|g \circ f \circ f^{-1}(y) - L\| < \epsilon$. Let $\delta = \min(\delta_1, r)$. Then $y \in \mathcal{D}(g)$ and $0 < \|y - b\| < \delta$ implies $y \in \mathcal{D}(f^{-1})$ and therefore $\|g(y) - L\| = \|g \circ f \circ f^{-1}(y) - L\| < \epsilon$. □

Corollary 0.4. Let $f\colon \mathbf{R}^n \to \mathbf{R}^n$ be continuously differentiable with $Df(a)$ non-singular for some point $a \in \mathbf{R}^n$ and let $b = f(a)$. Then there is $r > 0$ such that f has an inverse function $f^{-1}\colon B_r(b) \to \mathbf{R}^n$ and f^{-1} is differentiable with derivative $Df^{-1}(b) = Df(a)^{-1}$.

Proof: Let $\epsilon > 0$ be chosen as in Lemma 0.2 and $r > 0$ be chosen as in Lemma 0.1. Then f is one-to-one on $B_\epsilon(a)$ and $B_r(b) \subseteq f(B_\epsilon(a))$. It follows $f^{-1}\colon B_r(b) \to B_\epsilon(a)$ exists.

Let $y \in B_r(b)$ and $x = f^{-1}(y)$. Then $x \in B_\epsilon(a)$ and Lemma 0.2 implies

$$\|f^{-1}(y) - f^{-1}(b)\| = \|x - a\| \le \lambda \|f(x) - f(a)\| = \lambda \|y - b\|.$$

Now since $f^{-1}\colon B_r(b) \to \mathbf{R}^n$ is one-to-one and continuous then Lemma 0.3 implies

$$\lim_{y \to b} \frac{\|f^{-1}(y) - f^{-1}(b) - Df(a)^{-1}(y - b)\|}{\|y - b\|}$$

exists if and only if

$$\lim_{x \to a} \frac{\|x - a - Df(a)^{-1}(f(x) - f(a))\|}{\|f(x) - f(a)\|}$$

exists. Estimating gives

$$\begin{aligned} &\frac{\|x - a - Df(a)^{-1}(f(x) - f(a))\|}{\|f(x) - f(a)\|} \\ &\qquad\qquad \le \|Df(a)^{-1}\| \frac{\|f(x) - f(a) - Df(a)(x - a)\|}{\|f(x) - f(a)\|} \\ &\qquad\qquad \le \lambda \|Df(a)^{-1}\| \frac{\|f(x) - f(a) - Df(a)(x - a)\|}{\|x - a\|} \to 0 \end{aligned}$$

as $x \to a$, which implies f^{-1} is differentiable. □