Theorem 2.10.7 (The inverse function theorem). *Let $W \subset \mathbb{R}^m$ be an open neighborhood of $\mathbf{x}_0$, and let $\mathbf{f} : W \to \mathbb{R}^m$ be a continuously differentiable function. Set $\mathbf{y}_0 = \mathbf{f}(\mathbf{x}_0)$.*

If the derivative $[\mathbf{Df}(\mathbf{x}_0)]$ is invertible, then $\mathbf{f}$ is invertible in some neighborhood of $\mathbf{y}_0$, and the inverse is differentiable.

To quantify this statement, we will specify the radius R of a ball V centered at $\mathbf{y}_0$, in which the inverse function is defined. First simplify notation by setting $L = [\mathbf{Df}(\mathbf{x}_0)]$. Now find $R > 0$ satisfying the following conditions:

1. *The ball W_0 of radius $2R|L^{-1}|$ and centered at $\mathbf{x}_0$ is contained in W.*
2. *In W_0, the derivative of $\mathbf{f}$ satisfies the Lipschitz condition*

$$|[\mathbf{Df}(\mathbf{u})] - [\mathbf{Df}(\mathbf{v})]| \le \frac{1}{2R|L^{-1}|^2} |\mathbf{u} - \mathbf{v}| . \qquad 2.10.11$$

Set $V = B_R(\mathbf{y}_0)$. Then

1. *There exists a unique continuously differentiable mapping $\mathbf{g} : V \to W_0$ such that*

$$\mathbf{g}(\mathbf{y}_0) = \mathbf{x}_0 \quad \text{and} \quad \mathbf{f}\big(\mathbf{g}(\mathbf{y})\big) = \mathbf{y} \text{ for all } \mathbf{y} \in V. \qquad 2.10.12$$

   Since the derivative of the identity is the identity, by the chain rule, the derivative of $\mathbf{g}$ is $[\mathbf{Dg}(\mathbf{y})] = [\mathbf{Df}(\mathbf{g}(\mathbf{y}))]^{-1}$.
2. *The image of $\mathbf{g}$ contains the ball of radius R_1 around $\mathbf{x}_0$, where*

$$R_1 = R\left|L^{-1}\right|^2 \left(\sqrt{|L|^2 + \frac{2}{|L^{-1}|^2}} - |L| \right). \qquad 2.10.13$$

Theorem 2.10.7 is illustrated by Figure 2.10.6.

Note that we could replace the length of the derivatives by their norm (Definition 2.9.6) to get a slightly stronger theorem.

On first reading, skip the last sentence about the ball with radius R_1. It is a minor point.

The ball V gives a lower bound for the domain of $\mathbf{g}$; the actual domain may be bigger. But there are cases where the largest R satisfying the conditions of Theorem 2.10.7 is optimal. We invite you to check that if $f(x) = (x-1)^2$, with $x_0 = 0$, so that $y_0 = 1$, then the largest R satisfying equation 2.10.11 is $R = 1$. Thus the interval $V = (0, 2)$ is the largest interval centered at 1 on which an inverse can be defined. Indeed, since the function g is $g(y) = 1 - \sqrt{y}$, any value of y smaller than 0 is not in the domain of g.

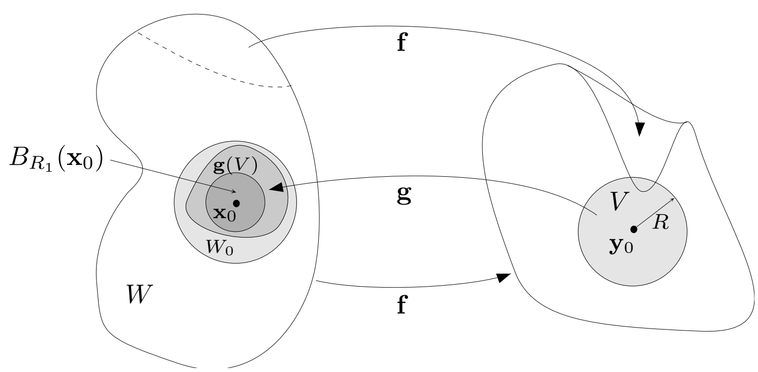

FIGURE 2.10.6. The function $\mathbf{f} : W \to \mathbb{R}^m$ maps every point in $\mathbf{g}(V)$ to a point in V; in particular, it sends $\mathbf{x}_0$ to $\mathbf{f}(\mathbf{x}_0) = \mathbf{y}_0$. Its inverse function $\mathbf{g} : V \to W_0$ undoes that mapping. The image $\mathbf{g}(V)$ of $\mathbf{g}$ may be smaller than its codomain W_0. The ball $B_{R_1}(\mathbf{x}_0)$ is guaranteed to be inside $\mathbf{g}(V)$; it quantifies "locally invertible".