The inverse and implicit function theorems

Motivations.

Consider the 1-dimensional situation.

Q1. Suppose $f:(a,b) \to \mathbb{R}$ is differentiable and $f'(x) > 0$ for every $x \in (a,b)$.

(a) What can you say about f? f is strictly increasing on (a,b).

(b) Is f injective? Yes.

(c) Is f surjective? Not necessarily; eg. $f(x) = x$, $a=0$, $b=1$.

(d) If $(c,d) = f(a,b)$ is the image of (a,b) under f, can we define $g:(c,d) \to (a,b)$ such that

$$\begin{cases} g(f(x)) = x & \forall x \in (a,b) \\ f(g(y)) = y & \forall y \in (c,d) \end{cases} ?$$

Yes.

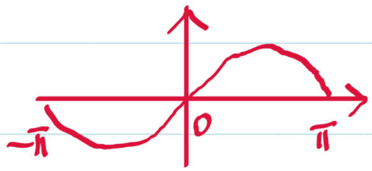
(e) Is the function g defined in (d) differentiable? Yes

What is $g'(y)$? $g'(y) = \dfrac{1}{f'(g(y))}$ by chain rule.

Q2. Suppose now $f:(a,b)\to\mathbb{R}$ is still differentiable $0\in(a,b)$, and $f'(0)>0$.

(a) Must f be strictly increasing on (a,b)?

No, eg. $f(x)=\sin x$, $(a,b)=(-\pi,\pi)$



(b) Must there be some (possibly tiny but non-empty) open interval I containing 0 so that f is strictly increasing on I?

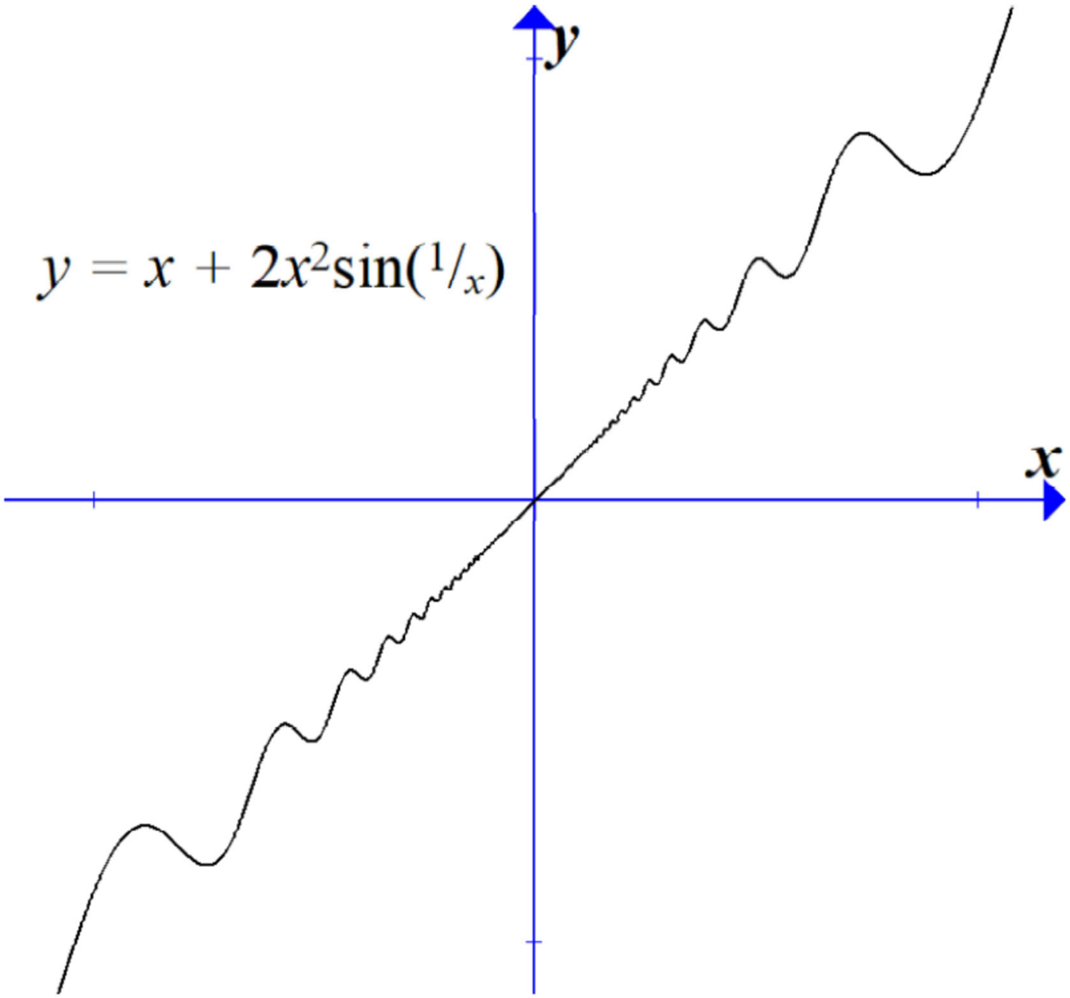
Again no!

e.g. $f(x)=\begin{cases} x+2x^2\sin\frac{1}{x} & \text{if } x\neq 0 \\ 0 & \text{if } x=0 \end{cases}$

Then $f'(0)=\lim\limits_{x\to 0}\dfrac{f(x)-f(0)}{x-0}=\lim\limits_{x\to 0}\dfrac{x+2x^2\sin\frac{1}{x}-0}{x}$

$=\lim\limits_{x\to 0}\left(1+2x\sin\frac{1}{x}\right)=1>0$, but for $x\neq 0$,

$f'(x)=1+4x\sin\frac{1}{x}-2\cos\frac{1}{x}$ which can be negative on intervals arbitrarily close to 0.



The function $f(x) = x + 2x^2\sin(\frac{1}{x})$ is bounded inside a quadratic envelope near the line $y = x$, so $f'(0) = 1$. Nevertheless, it has local max/min points accumulating at $x = 0$, so it is not one-to-one on any surrounding interval.

Source: Wikipedia

$$f'(x) = \begin{cases} 1 - 2\cos\frac{1}{x} + 4x\sin\frac{1}{x} & \text{if } x \neq 0 \\ 0 & \text{if } x = 0 \end{cases}$$

Problem: $f'(x)$ is not continuous at $x=0$!

Q3. Suppose $f : (a,b) \to \mathbb{R}$ is not only differentiable on (a,b), but f' is continuous on (a,b). Suppose also $f'(0) \neq 0$. Must there be some non-empty open interval I near 0 such that f is monotone on I?

Yes! Just choose some open interval I containing 0 so that $f'(x) \neq 0$ $\forall x \in I$.

This guarantees that f is injective on I.

This generalizes to higher dimension.

Theorem (Inverse function theorem on $\mathbb{R}^n$, $n \geq 1$)

Suppose $\Omega \subseteq \mathbb{R}^n$ is an open set and $f: \Omega \to \mathbb{R}^n$ is continuously differentiable (i.e. differentiable with a continuous derivative). If $Df(x_0)$ is invertible (on $\mathbb{R}^n$) for some $x_0 \in \Omega$, then $\exists$ open set $V \subseteq \Omega$ containing x_0 such that $f: V \to f(V)$ is bijective, and its inverse map $g: f(V) \to V$ is continuously differentiable. Furthermore,

$$Dg(y) = [Df(g(y))]^{-1} \quad \forall y \in f(V).$$

If $f: \Omega \to \mathbb{R}^n$ is C^∞, then so is $g: f(V) \to V$.

$Df(x_0)$ is an $n \times n$ matrix!

"$Df(x_0)$ invertible" is the same as saying "$\det(Df(x_0)) \neq 0$"

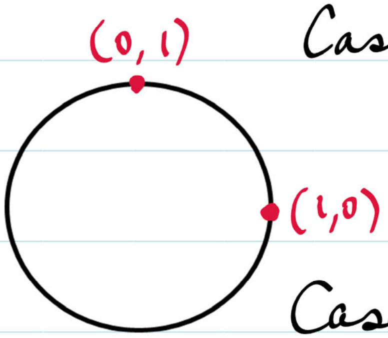
(The most difficult part of the proof is in establishing that f is injective on some open set V containing x_0.
This is done using e.g. the contraction mapping principle.)

We write $f \in C^1$ if f is continuously differentiable.

Closely related to the inverse function theorem is the implicit function theorem.

Question: If (x_0, y_0) is on a hypersurface $F(x,y)=0$, can we parametrize $\{F(x,y)=0\}$ as the graph of a function $y=\phi(x)$ near the point (x_0, y_0)?

eg. $F(x,y) = x^2+y^2-1$ so that "$F(x,y)=0$" is the unit circle centered at 0, (x_0,y_0) lies on the unit circle



Case 1: $(x_0, y_0) = (0,1)$.

Then near $(0,1)$, the circle can be described by $y = \sqrt{1-x^2}$, $x \in (-1,1)$.

Case 2: $(x_0, y_0) = (1,0)$ Problem: $\partial_2 F(1,0) = 0$!

Then near $(1,0)$, the circle cannot be described as a graph $y=\phi(x)$ [Two values of y for each $x \in (0,1)$].

We use the inverse function theorem to prove the following

<u>Theorem</u> (Implicit function theorem on $\mathbb{R}^n$, $n>1$)

Suppose $\Omega \subseteq \mathbb{R}^n$ is an open set and $F: \Omega \to \mathbb{R}$ is continuously differentiable. If $(x_0, y_0) \in \Omega$ satisfies

$$\begin{cases} F(x_0, y_0) = 0 \\ \partial_y F(x_0, y_0) \neq 0 \end{cases}$$

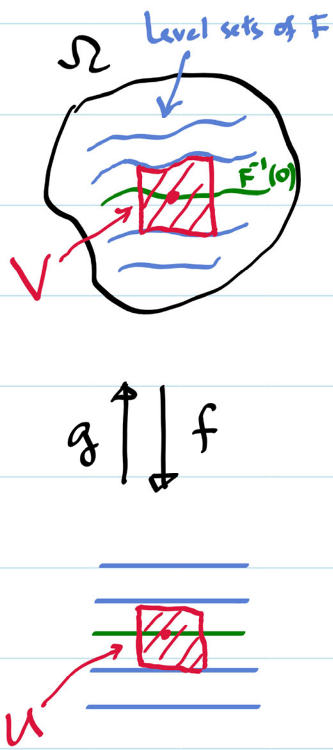
where y_0 is the last variable of (x_0, y_0) and $\partial_y F$ is partial derivative with respect to the last variable of $\mathbb{R}^n$, then

$\exists$ open set $V \subseteq \mathbb{R}^n$ containing (x_0, y_0), and a C^1 map $\phi: \pi(V) \to \mathbb{R}$ where $\pi: \mathbb{R}^n \to \mathbb{R}^{n-1}$ is the coordinate projection onto the first $(n-1)$ coordinates, such that

$$\begin{cases} (x,y) \in V \\ F(x,y) = 0 \end{cases} \text{ if and only if } \begin{cases} x \in \pi(V) \\ y = \phi(x) \end{cases}.$$

If F is C^∞ on Ω, then $\phi \in C^\infty$ on $\pi(V)$.

Proof. Given F and $(x_0, y_0) \in \Omega$ as in the implicit function theorem, let $f : \Omega \to \mathbb{R}^n$ be given by $f(x,y) = (x, F(x,y))$ $\forall x, y \in \Omega$. Then f is C^1 on Ω, and $\det(Df(x_0,y_0)) = \det \left(\begin{array}{c:c} Id_{n-1} & \partial_x F(x_0,y_0) \\ \hdashline 0 & \partial_y F(x_0,y_0) \end{array} \right) = \partial_y F(x_0,y_0) \neq 0.$



Hence the inverse function theorem applies, and $\exists$ open sets $V \subseteq \Omega$, $U := f(V) \subseteq \mathbb{R}^n$ and a C^1 map $g : U \to V$ such that

① $f(g(x,y)) = (x,y)$ $\forall (x,y) \in U$ and ② $g(f(x,y)) = (x,y)$ $\forall (x,y) \in V$.

The first $(n-1)$ components of ① shows that $g(x,y) = (x, G(x,y))$ for some C^1 function G on U. The last components of ① and ② then shows

③ $F(x, G(x,y)) = y$ $\forall (x,y) \in U$ and ④ $G(x, F(x,y)) = y$ $\forall (x,y) \in V$.

① and ② also ensures that $\pi(U) = \pi(V)$. By shrinking U and V if necessary, we assume $U = \pi(V) \times J$ for some open interval $J \ni 0$. Let $\phi : \pi(V) \to \mathbb{R}$ be given by $\phi(x) = G(x,0)$ $\forall x \in \pi(V)$. If $x \in \pi(V)$ and $y = \phi(x)$, then $(x,0) \in U$, so by ③, $F(x,y) = F(x,\phi(x)) = F(x, G(x,0)) = 0$; we also have $(x,y) = g(x,0) \in V$. On the other hand, if $(x,y) \in V$ and $F(x,y) = 0$, then by ④, $y = G(x, F(x,y)) = G(x,0) = \phi(x)$ and clearly $x \in \pi(V)$. Finally, if $F \in C^\infty$, then so is f, g, G and hence ϕ.