Let see some examples before proving the IFT.

eg 3.8: Let $F : (0,\infty)\times(-\infty,\infty) \to \mathbb{R}^2$

$(r, \theta) \mapsto (r\cos\theta, r\sin\theta)$

Then $DF = \begin{pmatrix} \cos\theta & -r\sin\theta \\ \sin\theta & r\cos\theta \end{pmatrix}$ invertible $\forall (r,\theta)$

Then IFT $\Rightarrow$ F is locally invertible at every point

$(r,\theta) \in (0,\infty)\times(-\infty,\infty)$. But F is clearly not globally invertible as it is not one-to-one:

$$F(r, \theta+2\pi) = F(r,\theta).$$

eg 3.9 U = open interval (a,b) in $\mathbb{R}$ $(n=1)$,

is a special case : C^1 function $f:(a,b)\to\mathbb{R}$

with $f'\neq 0 \Rightarrow f$ strictly increasing or decreasing

$\Rightarrow$ global inverse exists.

($\therefore$ 1-dim has stronger result than high dimensions) eg 3.10: (i) $F: \mathbb{R}^2 \to \mathbb{R}^2 : (x,y) \mapsto (x^2, y)$.

$$DF = \begin{pmatrix} 2x & 0 \\ 0 & 1 \end{pmatrix} \text{ singular at } (x,y) = (0,0).$$

F doesn't satisfy the condition DF invertible in the IFT.

And clearly F is not invertible near $(x,y) = (0,0)$ as $F(\pm a, b) = (a^2, b)$ (2-to-1 near $(0,0)$).

$\therefore$ "DF invertible" condition can't be removed from IFT.

(ii) $H: \mathbb{R}^n \to \mathbb{R}^n : (x,y) \mapsto (x^3, y)$

is bijective & $H^{-1}(x,y) = (x^{1/3}, y)$ exists.

But $$DH = \begin{pmatrix} 3x^2 & 0 \\ 0 & 1 \end{pmatrix} \text{ singular at } (x,y) = (0,0).$$

The point is:

H^{-1} is not C^1 near $(x,y) = (0,0)$.

$\therefore$ "DF invertible" is only a "sufficient" condition for local invertibility.

Terminology: The condition in IFT that $DF(x_0)$ is invertible is called the nondegeneracy condition.

By eg 3.10, without nondegeneracy condition, the map may or may not be local invertible.

But Nondegeneracy condition is necessary for the differentiability of the local inverse:

Prop 3.8: Let $F: U \xrightarrow{\subset \mathbb{R}^n \text{ open}} \mathbb{R}^n$ be C^1, and $x_0 \in U$.

Suppose $\exists$ open V s.t. $x_0 \in V \subset U$, and $F|_V$ has a differentiable inverse.

Then $DF(x_0)$ is non-singular (ie. invertible).

Pf: Suppose the local inverse $(F|_V)^{-1}$ exists and is differentiable at the point $y_0 = F(x_0)$. Then

Chain rule $\Rightarrow$ $D(F^{-1})(y_0)\, DF(x_0) = \text{Identity}$

$\Rightarrow$ $DF(x_0)$ is invertible. #

Proof of IFT (Thm 3.7)

Special Case: $x_0 = 0$, $y_0 = F(x_0) = F(0) = 0$,

$DF(0) = I$ (the Identity)

Step 1 Let $\Psi(x) = -x + F(x)$.

Then $\exists r > 0$ s.t.

$$|\Psi(x_2) - \Psi(x_1)| \leq \frac{1}{2}|x_2 - x_1| \quad \text{on } \overline{B_r(0)}.$$

Pf of Step 1:

As $0 \in U$ and U is open, $\exists r_0 > 0$ s.t.

$$\overline{B_{r_0}(0)} \subset U.$$

Then $\Psi(x_1) - \Psi(x_2) = -x_1 + F(x_1) + x_2 - F(x_2)$

(prop. 3.6) $= \left(\int_0^1 DF(x_2 + t(x_1 - x_2))\,dt\right)(x_1 - x_2) - (x_1 - x_2)$

$= \left[\int_0^1 DF(x_2 + t(x_1 - x_2))\,dt - I\right](x_1 - x_2)$.

$= \int_0^1 [DF(x_2 + t(x_1 - x_2)) - DF(0)]\,dt \cdot (x_1 - x_2)$

As F is C^1, we have

$\forall \varepsilon > 0$, $\exists r > 0$, $r \leq r_0$ such that

$$\|DF(x) - DF(0)\| < \varepsilon, \quad \forall x \in \overline{B_r(0)},$$

where $\|(b_{ij})\| = \sqrt{\sum_{i,j} b_{ij}^2}$ for any $n \times n$ matrix (b_{ij}).

Since $\overline{B_r(0)}$ is convex,

$$x_1, x_2 \in \overline{B_r(0)} \Rightarrow x_2 + t(x_1 - x_2) \in \overline{B_r(0)}.$$

Hence

$\forall \varepsilon > 0$, $\exists r > 0$ $(r \leq r_0)$ such that

$$\|DF(x_2 + t(x_1 - x_2)) - DF(0)\| < \varepsilon, \quad \forall x_1, x_2 \in \overline{B_r(0)},$$

and $t \in (0,1)$.

Therefore

$$|\Psi(x_1) - \Psi(x_2)| \leq \varepsilon |x_1 - x_2|.$$

Choosing $\varepsilon = \frac{1}{2} > 0$, then $\exists r > 0$, $r \leq r_0$ s.t.

$$|\Psi(x_1) - \Psi(x_2)| \leq \frac{1}{2}|x_1 - x_2|, \quad \forall x_1, x_2 \in \overline{B_r(0)}.$$

This completes the proof of Step 1 of the special case.

#

Step 2 $r>0$ as in Step 1. Then

$$\forall y \in B_{\frac{r}{2}}(0), \ \exists x \in B_r(0) \text{ such that}$$

$$F(x) = y.$$

And the local inverse G of F,

$$G: B_{\frac{r}{2}}(0) \to G(B_{\frac{r}{2}}(0)) \subset B_r(0)$$

satisfies $|G(y_1) - G(y_2)| \leq 2|y_1 - y_2|, \ \forall y_1, y_2 \in B_{\frac{r}{2}}(0).$

with $G(B_{\frac{r}{2}}(0))$ open in $B_r(0)$.

Pf of Step 2: By Step 1, one can apply Thm 3.4 (Perturbation of Identity) to show that

$$\forall y \in \overline{B_R(0)} \text{ with } R = (1 - \tfrac{1}{2}) \cdot r = \tfrac{r}{2}$$

$$\exists x \in \overline{B_r(0)} \quad \text{s.t.} \quad F(x) = y.$$

Then by remark (3) (after the proof of Thm 3.4) (which is the Remark 3.1 of Prof Chou's notes).

we have
$$|G(y_1) - G(y_2)| \leq \frac{1}{1 - \frac{1}{2}} |y_1 - y_2|$$
$$= 2|y_1 - y_2| \quad \forall y_1, y_2 \in \overline{B_{\frac{r}{2}}}(0).$$

Finally, remark (2) $\Rightarrow$ $G(B_{\frac{r}{2}}(0))$ is open in $B_r(0)$.

#

Step 3 G is differentiable on $B_{\frac{r}{2}}(0)$.

and $DG(y) = (DF)^{-1}(G(y))$, $\forall y \in B_{\frac{r}{2}}(0)$.

Pf of Step 3: As $DF(0) = I$, we may assume that $DF(x)$ is invertiable $\forall x \in B_r(0)$ for the $r>0$ given in Step 1 (Since we may always choose a smaller r in the proof of Step 1.).

Let $W = B_{\frac{r}{2}}(0) (\subseteq B_R(0))$ and

$V = G(B_{\frac{r}{2}}(0)) = G(W) \ni 0$.

Then $G: W \to V$ (and $F: V \to W$)

By Chain rule, if G is differentible, then

$$DF(G(y))\, DG(y) = I, \quad \forall y \in W$$

Hence

$$DG(y) = (DF)^{-1}(G(y)).$$ (So we target $(DF)^{-1}(G(y))$ as the linear map required.)

For $y_1 \in W = B_{\frac{r}{2}}(0)$ & $y_1 + y \in W = B_{\frac{r}{2}}(0)$,

we have
$$\begin{aligned} y &= (y_1 + y) - y_1 \\ &= F(G(y_1+y)) - F(G(y_1)) \end{aligned}$$

Denote $x_1 = G(y_1+y)$ and $x_2 = G(y_1)$

Then $y = F(x_1) - F(x_2)$

$$= \left[\int_0^1 DF(x_2 + t(x_1 - x_2))\, dt\right](x_1 - x_2) \quad \text{(Prop 3.6)}$$

$$= \int_0^1 \left[DF(x_2 + t(x_1 - x_2)) - DF(x_2)\right] dt \cdot (x_1 - x_2)$$

$$+ DF(x_2)(x_1 - x_2)$$

Hence

$$(DF)^{-1}(x_2)y = (DF)^{-1}(x_2)\int_0^1 \left[DF(x_2 + t(x_1 - x_2)) - DF(x_2)\right] dt\, (x_1 - x_2)$$
$$+ (x_1 - x_2)$$

i.e. $\quad G(y_1 + y) - G(y_1) = (DF)^{-1}(G(y_1))y + R, \qquad (x_2 = G(y_1))$

where

$$R = (DF)^{-1}(x_2)\int_0^1 \left[DF(x_2) - DF(x_2 + t(x_1 - x_2))\right] dt\, (x_1 - x_2)$$

Observes that

$$|x_1 - x_2| = |G(y_1 + y) - G(y_1)| \le 2\, |(y_1 + y) - y_1| \quad \text{(step 2)}$$
$$= 2|y|,$$

we have, $|x_1 - x_2| \to 0$ as $|y| \to 0$ and

$$\frac{|R|}{|y|} \le 2\|DF^{-1}(x_2)\| \int_0^1 \|DF(x_2) - DF(x_2 + t(x_1 - x_2))\| dt$$

By assumption F is C^1 ($x_1, x_2 \in \overline{B_r(0)}$), we have

$$\lim_{|y| \to 0} \frac{|R|}{|y|} = 0 .$$

Therefore $G(y_1 + y) - G(y_1) = (DF)^{-1}(G(y_1))\, y + o(|y|)$,

which implies G is differentiable at $y_1 \in B_{\frac{r}{2}}(0) = W$

and $\quad DG(y_1) = (DF)^{-1}(G(y_1))$. #

<u>Final Step for special case</u>: G is $C^1(B_{\frac{r}{2}}(0))$ and furthermore, if F is C^k ($k \ge 1$), then G is C^k.

<u>Pf of final step</u>: By assumption, DF is continuous and invertible on $B_r(0)$. Linear Algebra $\Rightarrow$ $(DF)^{-1}$ is continuous. Therefore, by Step 3 (and step 2)

$DG(y) = (DF)^{-1}(G(y))$ is continuous. Hence G is C^1.

The fact that F is C^k $(k \geq 1) \Rightarrow G$ is C^k is by differentiating the identity $DG(y) = (DF)^{-1}(G(y))$ and induction. ##

General Case:

Consider $\tilde{F}(x) = (DF)^{-1}(x_0)\left[F(x+x_0) - y_0\right]$

Then $\tilde{F}(0) = 0$

$\tilde{F}$ defined on an open set $\tilde{U} = U - x_0$
$= \{x : x + x_0 \in U\}$
with $0 \in \tilde{U}$.

and
$D\tilde{F}(0) = (DF)^{-1}(x_0)\, DF(x_0) = I$.

By the special case, $\exists\, r > 0$ such that

$$\exists : \tilde{G} : B_{\frac{r}{2}}(0) \longrightarrow \tilde{G}(B_{\frac{r}{2}}(0)) \subset B_r(0) \subset \tilde{U}$$

s.t. $\tilde{G}$ is the local inverse of $\tilde{F}$.

Let $W = DF(x_0)(B_{\frac{r}{2}}(0)) + y_0$

$V = \tilde{G}(B_{\frac{r}{2}}(0)) + x_0$, (then $V \subset U$ & $x_0 \in V$)

and

$$G : W \to V \text{ by}$$

$$G(y) = \tilde{G}\left((DF)^{-1}(x_0)(y - y_0)\right) + x_0, \quad \forall y \in W.$$

Clearly G maps W bijective onto V.

Since $\tilde{F}(x) = (DF)^{-1}(x_0)\left[F(x + x_0) - y_0\right]$,

we have $F(x + x_0) = (DF)(x_0)\tilde{F}(x) + y_0, \quad \forall x \in B_r(0)$

$$\Rightarrow \quad F(x) = (DF)(x_0)\tilde{F}(x - x_0) + y_0, \quad x \in V$$

Hence $\forall y \in W$

$$F(G(y)) = y_0 + (DF)(x_0)\tilde{F}(G(y) - x_0)$$

$$= y_0 + (DF)(x_0)\tilde{F}\left[\tilde{G}\left((DF)^{-1}(x_0)(y - y_0)\right)\right]$$

by definition of G

$$= y_0 + (DF)(x_0)\left((DF)^{-1}(x_0)(y - y_0)\right)$$ since $\tilde{F} \circ \tilde{G} = I$

$$= y_0 + y - y_0$$

$$= y$$

$\therefore$ G is the local inverse of F

The remaining facts that $F \in C^k$ $(k \geq 1) \Rightarrow G \in C^k$ is clear from the definition of G, and the results on $\tilde{G}$ ($\& \tilde{F}$) in the special case. ##

Def: A C^k-map $F : V \to W$ (V, W open in $\mathbb{R}^n$) is a <u>C^k-diffeomorphism</u> if F^{-1} exists and is also C^k.

<u>Note</u>: (i) The IFT can be rephased as:

If $F : U \overset{\subset \mathbb{R}^n}{\longrightarrow} \mathbb{R}^n \in C^k$, and DF is nonsingular at a point $x_0 \in U$, then F is a C^k-diffeomorphism between some nbds V and W of x_0 & $F(x_0)$ respectively.

(ii) If $F: V \to W$ is a C^k-diffeomorphism, then $\forall$ function $\varphi: W \to \mathbb{R}$, there corresponds a function $\psi = \varphi \circ F : V \to \mathbb{R}$. Conversely, $\forall$ function $\psi: V \to \mathbb{R}$, there corresponds a function $\varphi = \psi \circ F^{-1} : W \to \mathbb{R}$. Moreover, φ is $C^k \iff \psi$ is C^k. Thus every C^k-diffeomorphism gives rise to a "<u>local</u> C^k-<u>change of coordinates</u>".

Thm 3.5 (Implicit Function Theorem)

Let U be an open set in $\mathbb{R}^n \times \mathbb{R}^m$

$F : U \to \mathbb{R}^m$ is a C^1-map.

Suppose that $(x_0, y_0) \in U$ satisfies $F(x_0, y_0) = 0$,

and $D_y F(x_0, y_0)$ is invertible in $\mathbb{R}^m$.

Then (1) $\exists$ an open set of the form $V_1 \times V_2 \subset U$ containing (x_0, y_0) and a C^1-map

$$\varphi : \overset{\subset \mathbb{R}^n}{V_1} \longrightarrow \overset{\subset \mathbb{R}^m}{V_2} \quad \text{with } \varphi(x_0) = y_0$$

such that $\quad F(x, \varphi(x)) = 0, \quad \forall\, x \in V_1.$

(2) $\varphi : V_1 \to V_2$ is C^k when F is C^k, $1 \le k \le \infty$.

(3) Moreover, assume further that DF_y is invertible in $V_1 \times V_2$.

Then, if $\psi : V_1 \to V_2$ is another C^1-map satisfying $\quad F(x, \psi(x)) = 0,$

we have $\psi \equiv \varphi$.

Note: If $F = \begin{pmatrix} F_1(x_1, \dots, x_n, y_1, \dots, y_m) \\ \vdots \\ F_m(x_1, \dots, x_n, y_1, \dots, y_m) \end{pmatrix}$, then

$$D_y F = \begin{pmatrix} \frac{\partial F_1}{\partial y_1} & \cdots & \frac{\partial F_1}{\partial y_m} \\ \vdots & & \\ \frac{\partial F_m}{\partial y_1} & \cdots & \frac{\partial F_m}{\partial y_m} \end{pmatrix}$$

is $m \times m$ & can be regarded as a linear transformation from $\mathbb{R}^m$ to $\mathbb{R}^m$.

In general, for a map F such that $DF(x_0, y_0)$ has rank m, then one can rearrange the independent variables to make the $m \times m$ submatrix corresponding to the last m columns of the Jocabian matrix invertible, i.e. in the situation of the theorem. Hence the condition $DF_y(x_0, y_0)$ is invertible in the Implicit Function Theorem can be generalized to $\operatorname{rank} DF(x_0, y_0) = m$.