→ 11/30/2016

(110)

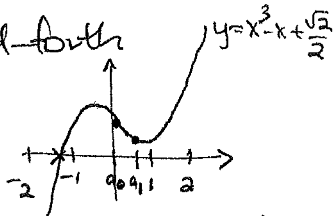
Does it always work?

- No, e.g. see book's EXAMPLE 2.8.3 solving $x^3 - x + \frac{\sqrt{2}}{2} = 0$

  where starting with a guess $a_0 = 0$ one bounces back-and-forth

  between $a_1 = \frac{\sqrt{2}}{2} = a_3 = a_5 = \cdots$

  and $a_0 = a_2 = a_4 = \cdots = 0$



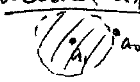
  never finding the real root

  (and it's still approximately true if $a_0 = \epsilon$ small, then $a_2 = c\epsilon^2$ for some c)

- Nevertheless, it often does work, and most of the issues are already illustrated in the 1-variable case $f(x) = 0$

  The book spends much of §2.8, 2.9 + Appendix A.5 developing a sufficient condition called Kantorovich's Theorem (THM 2.8.13)

  (guarantees convergence of $\bar{a}_0, \bar{a}_1, \ldots$ to a root in a ~~certain~~ this ball around $\bar{a}_1$:)



  then a consequence (PROP 2.8.14)

  (guarantees it in a certain ball around $\bar{a}_0$)

  and a superconvergence result

  (guarantees the distances $|\bar{a}_{n+1} - \bar{a}_n| =: \epsilon_n$ to eventually shrink incredibly fast, like $\epsilon_{n+m} \leq \left(\begin{smallmatrix}\text{some}\\\text{constant}\end{smallmatrix}\right)\left(\frac{1}{2}\right)^{2^m}$)

I found it a bit technical to even state them, so we'll skip them,

but roughly they require knowing that the product of 3 quantities are not large:

$$|\bar{f}(\bar{a}_0)| \cdot |Df(\bar{a}_0)^{-1}|^2 \cdot \underbrace{M}_{\sup\left\{\frac{|[Df(\bar{u}_1)] - [Df(\bar{u}_2)]|}{|\bar{u}_1 - \bar{u}_2|} \,:\, u_1 \neq u_2 \text{ in } B_{|\bar{a}_1 - \bar{a}_0|}(\bar{a}_1)\right\}} < \frac{1}{2}$$

see the "airplane crash" discussion on p. 244 to get a feeling for these 3 factors!

M is called a Lipschitz constant

(DEF'N 2.8.4: The derivative Df of $f: U \to \mathbb{R}^m$, $U \subset \mathbb{R}^n$, satisfies a Lipschitz condition with constant M on $V \subset U$ if $|[Df(\bar{x})] - [Df(\bar{y})]| \leq M\,|\bar{x} - \bar{y}|$ $\forall \bar{x} \neq \bar{y}$ in V)

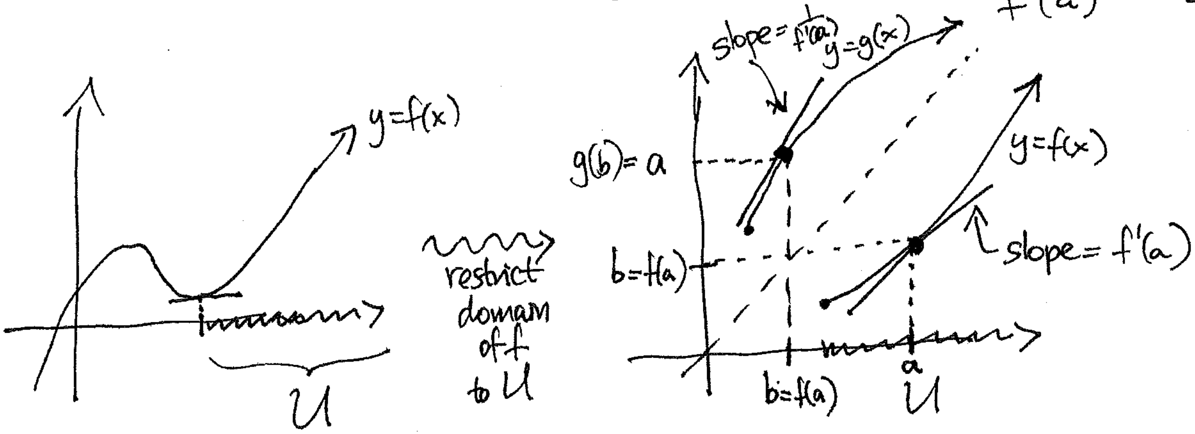
(111)

§2.10 Inverse, implicit function theorems

— tell us when we can invert nonlinear maps locally ($I_{nv.}$ F.T.)
and when we can use an equation system to write some variables in terms of others ($I_{mp.}$ F.T.)

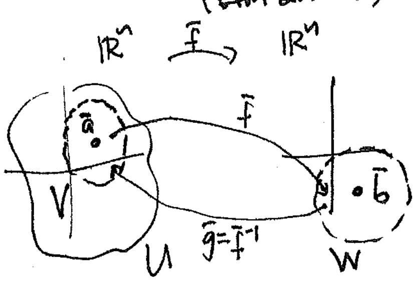
Recall in 1 variable, if $f(x)$ was differentiable on some $U \subset \mathbb{R}$ and $f'(x)$ was nonvanishing on U (say always $f'(x) > 0$ on U) then f was monotone on U ($x<y \Rightarrow f(x) < f(y)$), and had a well-defined inverse $g = f^{-1}$ on $f(U)$; with related derivative: if $a \in U$ has $b = f(a) \in f(U)$

then $$g'(b) = g'(f(a)) = \frac{1}{f'(a)} = \frac{1}{f'(g(b))}$$



In many variables, $f'(a) \neq 0$ is replaced by $D\bar{f}(\bar{a})$ invertible.

Inverse Function Thm: (more than THM 2.10.4, less than THM 2.10.7) If $\bar{f}: \underset{\mathbb{R}^n}{\overset{\text{open}}{U}} \to \mathbb{R}^n$ is in $C^1(U)$ (i.e. continuous partial derivatives on U)



and $\bar{a} \in U$ has $D\bar{f}(\bar{a})$ invertible (i.e. $\det[J\bar{f}(\bar{a})] \neq 0$), then $\exists$ open sets $\begin{cases} V \subset \mathbb{R}^n, \bar{a} \in V \subset U \\ W \subset \mathbb{R}^n \text{ containing } \bar{b} = \bar{f}(\bar{a}) \end{cases}$ such that

(i) one has an inverse $\bar{g}: W \to V$ such that $\bar{f} \circ \bar{g} = 1_W$, $\bar{g} \circ \bar{f} = 1_V$

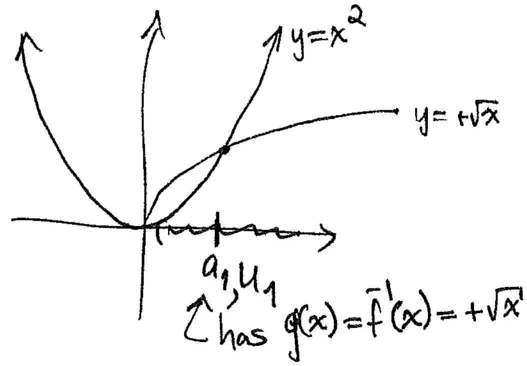
(ii) $\bar{g}$ is differentiable on W

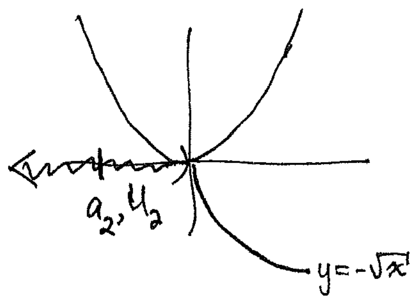
($\underset{\text{Chain rule}}{\Rightarrow}$ (iii) $[D\bar{f}(\underset{\bar{g}(\bar{b})}{\bar{a}})][D\bar{g}(\bar{b})] = 1_{\mathbb{R}^n}$ i.e. $[J\bar{g}(\bar{b})] = [J\bar{f}(\bar{a})]^{-1}$)

(112)

Note: 1. For a given function $\bar{f}(\bar{x})$, the W and $\bar{g}=\bar{f}^{-1}$ may depend on $\bar{a}$ (and U)
2. We may have no nice formula for $\bar{g}=\bar{f}^{-1}$

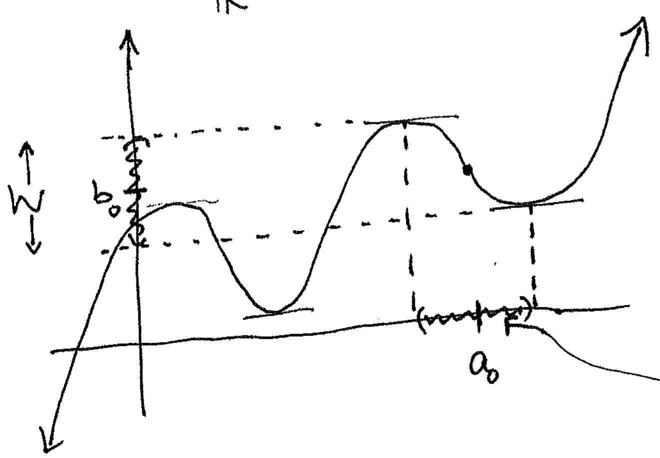
EXAMPLES ① $f(x)=x^2$
$\mathbb{R}^1 \to \mathbb{R}^1$





② A quintic polynomial

$$f(x)= x^5+ax^4+bx^3+\cancel{c}x^2+dx+\cancel{e}$$
$$\mathbb{R}^1 \longrightarrow \mathbb{R}^1$$



. has no simple radical formula for its 5 roots to $f(x)=a_0$

Nevertheless, for a_0 in here, we can solve (approximately) for $\cancel{x}=f^{-1}(\cancel{y})$ with y lying in the region W shown, via Newton's method.

12/1/2016 >

③ (EXAMPLE 2.10.6) One can use the fact that $D\bar{f}(\bar{a})$ not invertible gives a clue to where $\bar{f}: U \to \mathbb{R}^n$, $U \subset \mathbb{R}^n$ may have done some "folding" near $\bar{a}$, to guess where the boundary of $\mathrm{img}(\bar{f})$ lies:
If C_1, C_2 are the circles shown here, where do the midpoints of line segments between their points actually trace out in $\mathbb{R}^2$?