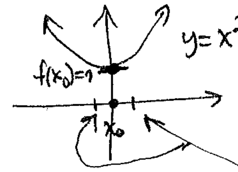
Q: Where did this proof fail for $f(x) = x^2+1$ having no roots $x \in \mathbb{R}$?



It *does* achieve a minimum value $f(x_0)=1$ at $x_0=0$.
But the "man" can't walk around the "flagpole" in <u>a full circle</u>, only at <u>2 points</u>

COR 1.6.14: A polynomial $p(z) = z^k + a_{k-1}z^{k-1} + \cdots + a_1 z^1 + a_0$ with $a_i \in \mathbb{C}$ and $k \geq 1$ has exactly k roots $r_1, \ldots, r_k$ in $\mathbb{C}$ (if you count <u>with</u> <u>multiplicity</u>), and factors as $p(z) = (z-r_1)(z-r_2)\cdots(z-r_k)$.

proof: Induct on k. The <u>base case</u> $k=1$ has $p(z) = z^1 + a_0 = z - r_1$ with $r_1 = a_0$.

In the <u>inductive step</u>, assume it for $k-1$, and given $p(z)$ of degree k, find some root r_1 using Fund'l Thm. Alg.

Use <u>long</u> <u>division algorithm</u> to write $p(z) = q(z)(z-r_1) + b$ for some $b \in \mathbb{C}$ and polynomial $q(z)$ of degree $k-1$, also <u>monic</u>, i.e. $q(z) = z^{k-1} + b_{k-2}z^{k-2} + \cdots$

$$q(z) = z^{k-1} + b_{k-2}z^{k-2} + \cdots$$
$$z - r_1 \,\big)\, z^k + a_{k-1}z^{k-1} + \cdots + a_1 z + a_0 = p(z)$$
$$\vdots$$
$$b$$

remainder b is of degree 0, i.e. $b \in \mathbb{C}$

However r_1 a root of $p(z)$ forces

$$0 = p(r_1) = q(r_1)\underbrace{(r_1 - r_1)}_{0} + b = b \quad \text{, i.e. } p(z) = q(z)(z-r_1)$$

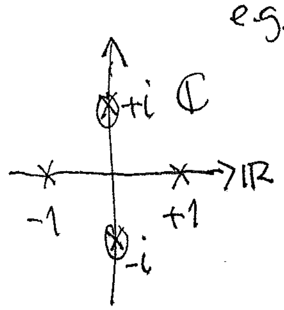
Now apply induction to $q(z)$ ∎

10/10/2016 > What about <u>irreducible</u> <u>factors</u> of $p(x) = x^k + a_{k-1}x^{k-1} + \cdots + a_0$ with $a_i \in \mathbb{R}$ if we only allow real coefficients in the factors?

e.g. $x^4 - 1 = (x^2-1)(x^2+1)$

$= (x-1)(x+1)\underbrace{(x^2+1)}$ irreducible over $\mathbb{R}$

$\left(= (x-1)(x+1)(x-i)(x+i) \text{ over } \mathbb{C} \right)$



(47) <u>COR 1.6.15</u>: If $p(x) = x^k + a_{k-1}x^{k-1} + \dots + a_1 x + a_0$ with $a_i \in \mathbb{R}$, $k \geq 1$

then it can be factored $p(x) = (x-r_1)\cdots(x-r_\ell)(x^2+c_1x+d_1)\cdots(x^2+c_mx+d_m)$

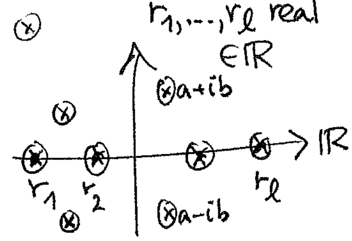
for some $r_1, \dots, r_\ell \in \mathbb{R}$

$c_1, \dots, c_m$

$d_1, \dots, d_m$

having $2m+\ell = k$ and each quadratic $x^2+c_ix+d_i$ irreducible over $\mathbb{R}$.

<u>proof</u>: Factor $p(z) = \underbrace{(z-r_1)\cdots(z-r_\ell)}_{r_1,\dots,r_\ell \text{ real } \in \mathbb{R}}\underbrace{(z-r_{\ell+1})\cdots(z-r_k)}_{r_{\ell+1},\dots,r_k \in \mathbb{C}\setminus\mathbb{R}}$ with $r_i \in \mathbb{C}$



Note that any root $r = a+ib \in \mathbb{C}$ of $p(z)$

has <u>conjugate</u> $\bar{r} = a-ib \in \mathbb{C}$ an<u>other</u> <u>root</u> of $p(z)$

since $\quad 0 = p(r) = r^k + \sum_{j=0}^{k-1} a_j r^j$

$$\Downarrow$$

$$0 = \bar{0} = \overline{p(r)} = \overline{r^k + \sum_{j=0}^{k-1} a_j r^j} = \overline{r^k} + \sum_{j=0}^{k-1} \overline{a_j r^j}$$

$$= (\bar{r})^k + \sum_{j=0}^{k-1} \bar{a}_j \cdot (\bar{r})^j$$

$$= (\bar{r})^k + \sum_{j=0}^{k-1} a_j (\bar{r})^j = p(\bar{r})$$

$\boxed{\text{since } a_j \in \mathbb{R}\ \forall j \Rightarrow \bar{a}_j = a_j}$

Thus the roots $r_{\ell+1}, \dots, r_k \in \mathbb{C}\setminus\mathbb{R}$ must come in <u>conjugate</u> <u>pairs</u> $\{a+ib, a-ib\}$ with $b \neq 0$ and $(z-(a+ib))(z-(a-ib))$

$$= z^2 - ((a+ib)+(a-ib))z + (a+ib)(a-ib)$$

$$= \underbrace{z^2 - 2az + (a^2+b^2)}_{\text{irreducible over } \mathbb{R} \text{ when } b \neq 0}$$

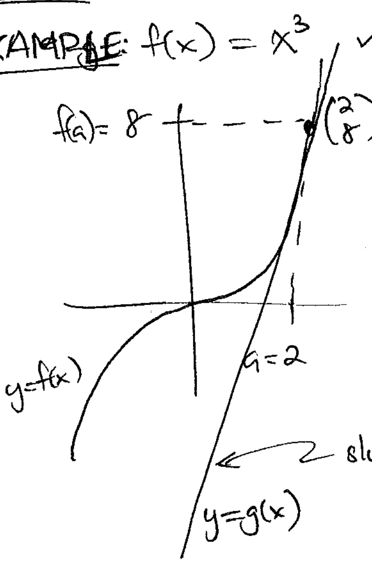
since $\underbrace{(-2a)^2 - 4\cdot 1\cdot(a^2+b^2)}_{\text{discriminant}} = -4b^2 < 0$ ∎

§1.7 Multivariate derivatives

– answering the question: given $f: \mathbb{R}^n \to \mathbb{R}^m$, is there a linear function $\mathbb{R}^n \to \mathbb{R}^m$ that best approximates f near some point $\bar{x} = \bar{a} \in \mathbb{R}^n$, and how to compute it?

We have a good idea already for $n = m = 1$, i.e. $f: \mathbb{R}^1 \to \mathbb{R}^1$

EXAMPLE: $f(x) = x^3$, near $x = 2$



$f'(x) = 3x^2$ $\left(= \lim_{h\to 0} \frac{f(x+h)-f(x)}{h}\right)$ exists for all $x \in \mathbb{R}^1$

$f'(a) = f'(2) = 3 \cdot 2^2 = 12$

$g(x) - \cancel{f(a)} = f'(a)(x-a)$

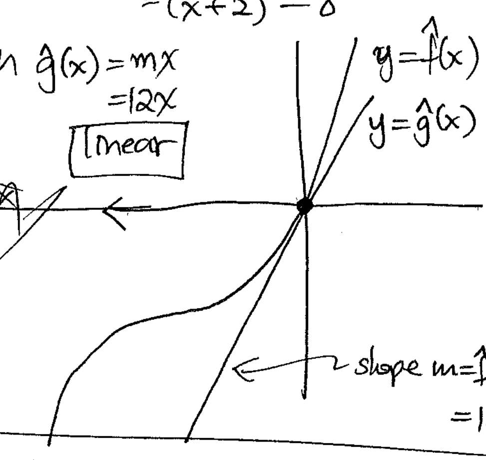
$g(x) - 8 = 12(x-2)$

$\boxed{g(x) = 12x - 16}$ ← not linear $\mathbb{R}^1 \to \mathbb{R}^1$ ("affine-linear")

← slope $m = f'(2) = 12$

But we could have considered just as well $\hat{f}(x) = f(x+a) - f(a)$
$= (x+2)^3 - 8$

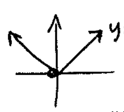
with $\hat{g}(x) = mx$
$= 12x$
$\boxed{\text{linear}}$



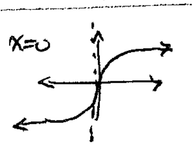
← slope $m = \hat{f}'(0) = 12$

NON-EXAMPLES:

① $f(x) = |x|$ near $x = 0$



② $f(x) = x^{1/3}$ near $x = 0$



(but we'll deal with it as an implicit function $x = y^3$ later in Chap. 2)

In fact, here's another equivalent definition of differentiability at a for $f: \mathbb{R}^1 \to \mathbb{R}^1$, that generalizes better to $\mathbb{R}^n \to \mathbb{R}^m$:

DEF'N 1.7.5: For U open in $\mathbb{R}^1$ and $f: U \to \mathbb{R}$, f is differentiable at some $a \in U$, with $f'(a) = m$,

if $$\lim_{h\to 0} \frac{1}{h}\Big(\underbrace{f(a+h) - f(a)}_{\hat{f}(h)} - \underbrace{mh}_{\text{a linear function } \mathbb{R}^1 \xrightarrow{\hat{g}(h)} \mathbb{R}^1,\ h \mapsto mh}\Big) = 0$$ (i.e. the error between $\hat{f}(h)$ & $\hat{g}(h)$ is smaller than h as $h \to 0$)

(49)

Why equivalent to the usual?

$$\lim_{h\to 0}\frac{1}{h}\left(f(a+h)-f(a) - mh\right) = 0 \quad \text{(and exists, in particular)}$$

$\Updownarrow$ since $\lim_{h\to 0}\frac{1}{h}(mh) = m$ exists, can add it to both sides

$$\lim_{h\to 0}\frac{1}{h}\left(f(a+h)-f(a)-mh\right) + \lim_{h\to 0}\frac{1}{h}(mh) = m$$

$\Updownarrow$ limit laws

$$\lim_{h\to 0}\frac{1}{h}\left(f(a+h)-f(a)\right) = m \quad \longleftarrow \boxed{\text{usual definition}}$$

Rather than doing something naive (and wrong) for $\bar{f}: \mathbb{R}^n \to \mathbb{R}^m$, like defining $f'(\bar{a}) = \lim_{\bar{h}\to\bar{0}} \frac{\bar{f}(\bar{a}+\bar{h})-\bar{f}(\bar{a})}{|\bar{h}|}$ (wrong even for $n=m=1$, since $|h|$ is always positive)

we ask for a linear function $L: \mathbb{R}^n \to \mathbb{R}^m$ that plays the above role of $\overset{h\mapsto}{mh}$ ...

DEF'N (1.7.10, essentially) For $\bar{f}: \overset{\text{open}}{U} \to \mathbb{R}^m$, $U \subset \mathbb{R}^n$, and $\bar{a} \in U$,

say that $\bar{f}$ is differentiable at $\bar{a}$

if $\exists$ some linear transformation $L: \mathbb{R}^n \to \mathbb{R}^m$, $\bar{x} \mapsto L(\bar{x})$

with $\lim_{\bar{h}\to\bar{0}} \frac{1}{|\bar{h}|}\left(\left(\bar{f}(\bar{a}+\bar{h})-\bar{f}(\bar{a})\right) - L(\bar{h})\right) = 0.$

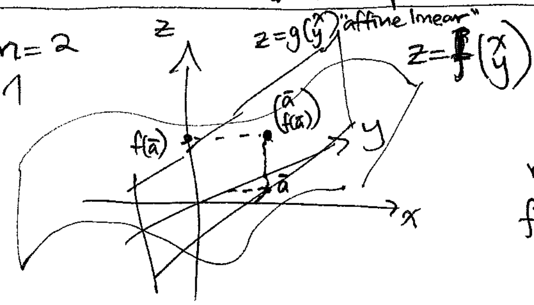
In this case we write $D\bar{f}(\bar{a}) = L$

(and we'll see shortly how to compute the matrix $[D\bar{f}(\bar{a})] = [L]$ that represents $L: \mathbb{R}^n \to \mathbb{R}^m$, via partial derivatives & Jacobian matrix)

10/12/2016

---

e.g. $n=2$, $m=1$



replace f with $\hat{f}\binom{x}{y} = f(\bar{x}+\bar{a}) - f(\bar{a})$