Lecture 14

CS 117, S26 © Katy Craig, 2026

Homework ~~6~~ 7 due Thursday, May 21st at 11:59pm

$(-\infty, +\infty)$ allowed

Def: Given an interval $I \subseteq \mathbb{R}$, a function $f: I \to \mathbb{R} \cup \{+\infty\}$ is convex if, $\forall x, y \in I, \alpha \in [0,1]$,

$$(*) \quad f((1-\alpha)x + \alpha y) \leq (1-\alpha) f(x) + \alpha f(y).$$

We say f is strictly convex if $\forall x \neq y, \alpha \in (0,1)$ $(*)$ holds with $<$.

Rmk: If $I \subseteq \mathbb{R}$ is an interval and $f: I \to \mathbb{R} \cup \{+\infty\}$ is convex, we may extend f to be a convex function on all of $\mathbb{R}$ by defining

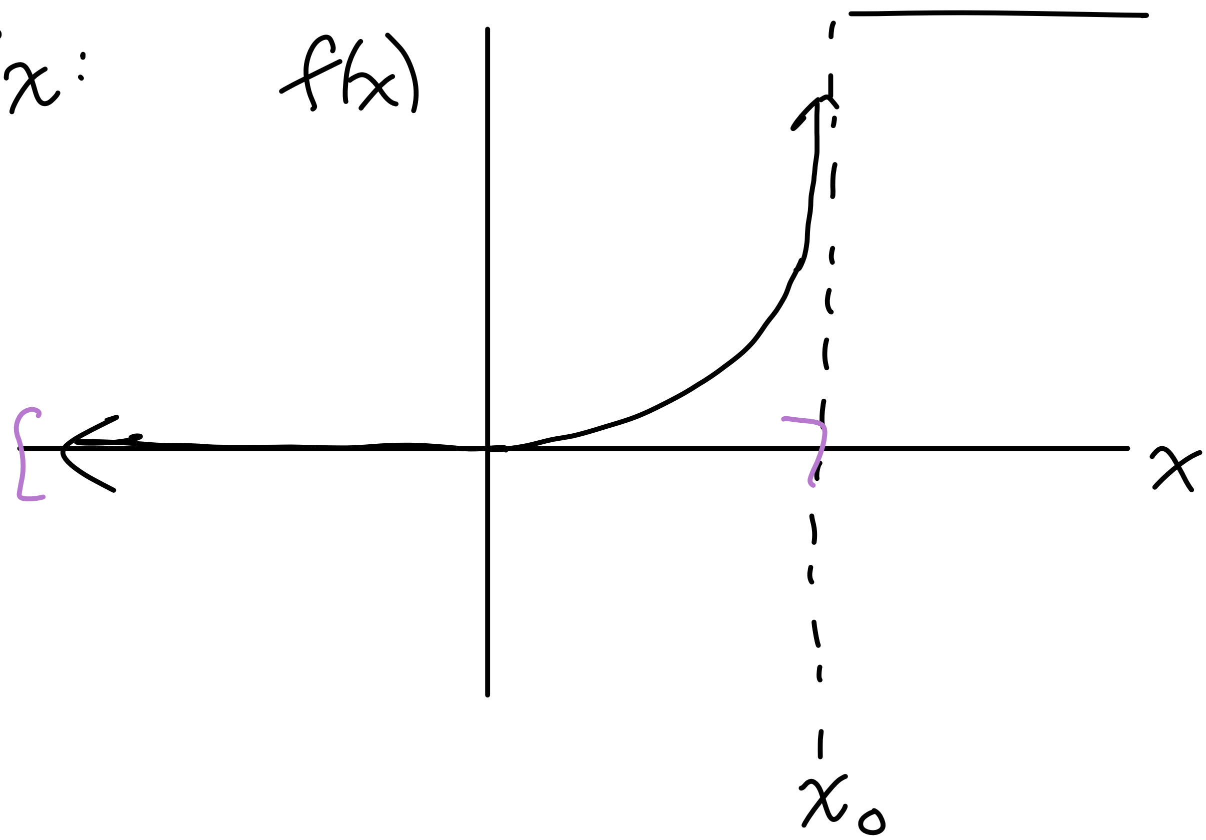
$$\tilde{f}(x) = \begin{cases} f(x) & x \in I \\ +\infty & x \notin I \end{cases}$$

Thm: If $f: \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is strictly convex, minimizers of f are unique.

Def: A function $f: \mathbb{R} \to \underbrace{\{-\infty\} \cup \mathbb{R} \cup \{+\infty\}}_{\overline{\mathbb{R}}}$ is lower semicontinuous at x_0 if, for all $x_n \to x_0$,

$$f(x_0) \leq \liminf_{n \to \infty} f(x_n)$$

Ex:



f is convex on $(-\infty, x_0)$

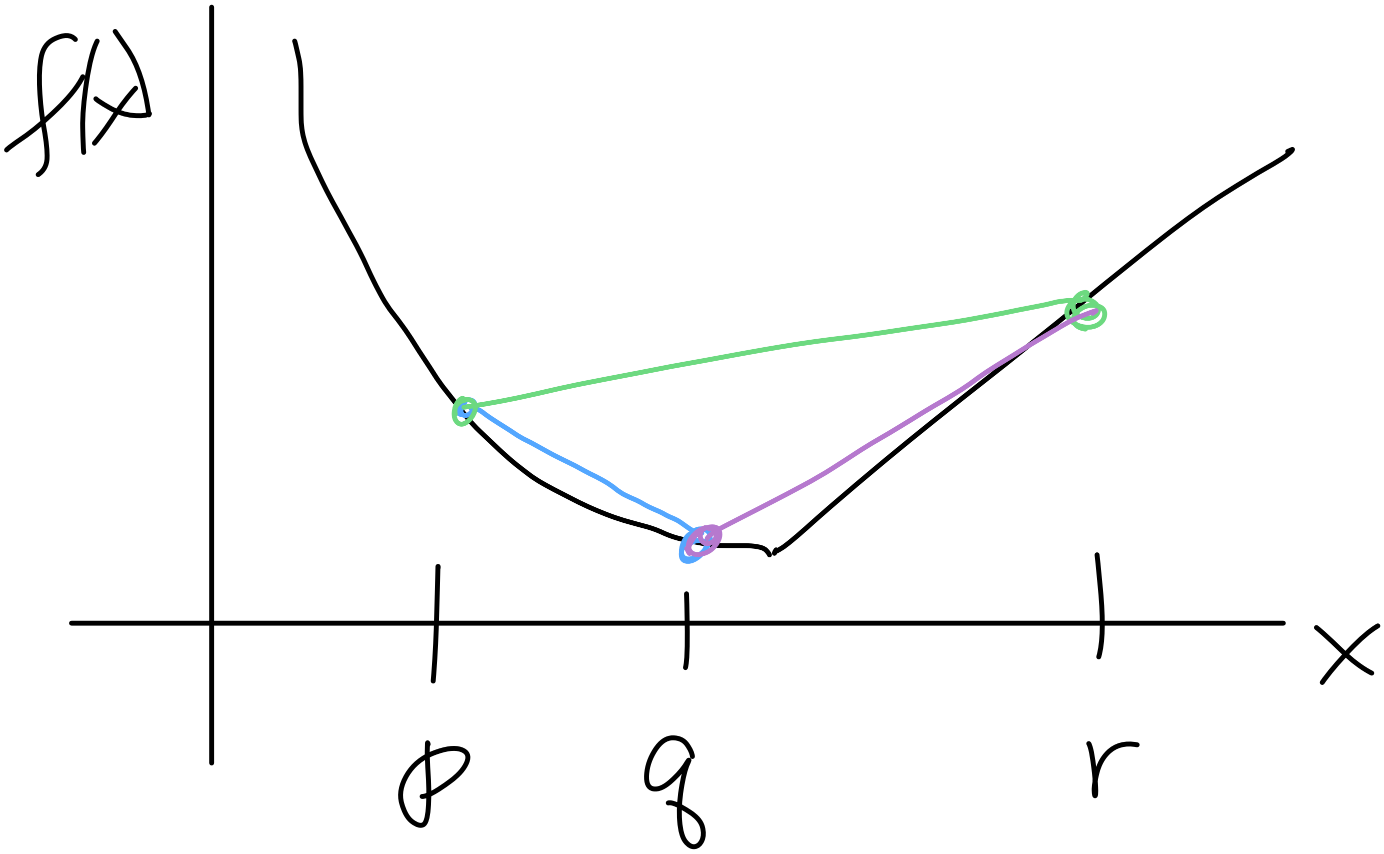
Thm: Given $f: \mathbb{R} \to \overline{\mathbb{R}}$ **lower** semicontinuous, f attains its minimum on any $[a,b]$.

Def: The (convex) domain of $f: \mathbb{R} \to \overline{\mathbb{R}}$ is

$$D(f) := \{x : f(x) < +\infty\}$$

Prop: $f: \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is convex iff $D(f)$ is an interval and for all $p, q, r \in D(f)$, $p < q < r$, we have

$$\frac{f(q)-f(p)}{q-p} \underset{\star}{\leq} \frac{f(r)-f(p)}{r-p} \underset{\star}{\leq} \frac{f(r)-f(q)}{r-q}$$



Slopes of secant lines are increasing.

Pf: First, suppose f is convex.
Given $p,q,r \in D(f)$, $p<q<r$,
write

$$q = (1-t)p + tr \ , \quad t = \frac{q-p}{r-p} \in (0,1)$$

By convexity of f,

$$f(q) \le (1-t)\,f(p) + t\,f(r)$$

$t = \frac{q-p}{r-p}$

$1-t = \frac{(r-p)-(q-p)}{r-p} = \frac{r-q}{r-p}$

Rearranging...

$$\frac{1}{t}\left(f(q) - f(p)\right) \le f(r) - f(p)$$

This gives ☆, since $\frac{1}{t} = \frac{r-p}{q-p}$

Rearranging...

$$f(q) \le (1-t)f(p) - (1-t)f(r) + \overbrace{(1-t)f(r) + tf(r)}^{f(r)}$$

$$f(q) - f(r) \le -(1-t)\left[f(r) - f(p)\right]$$

$$\frac{f(q)-f(r)}{-(1-t)} \ge f(r) - f(p)$$

This gives ☆, since $\frac{1}{1-t} = \frac{r-p}{r-q}$.

Next, assume the second monotonicity property holds.

Fix $x, y \in \mathbb{R}$ and $t \in [0,1]$. We must show

$$f(\underbrace{(1-t)x + ty}_{q}) \le (1-t)f(x) + tf(y).$$

The inequality always holds if $x = y$ or $t = 0, 1$. WLOG, $x \neq y$, $t \in (0,1)$.

The inequality always holds if either x or $y \notin D(f)$. WLOG, x and $y \in D(f)$.

WLOG $x < y$. Then $x < q < y$. Since $D(f)$ is an interval, $q \in D(f)$.

By monotonicity of secant line,

$$(q-x)\cdot\frac{f(q)-f(x)}{q-x} \le \frac{f(y)-f(x)}{y-x}\cdot(q-x)$$

$$f(q)-f(x) \le \overbrace{\left(\frac{q-x}{y-x}\right)}^{t}\left(f(y)-f(x)\right)$$

$$f(q) \le (1-t)f(x) + t f(y) \qquad \square$$

<u>Ex</u>: Statistical divergences

Density function $\rho(x),\ x\in\mathbb{R}^d$

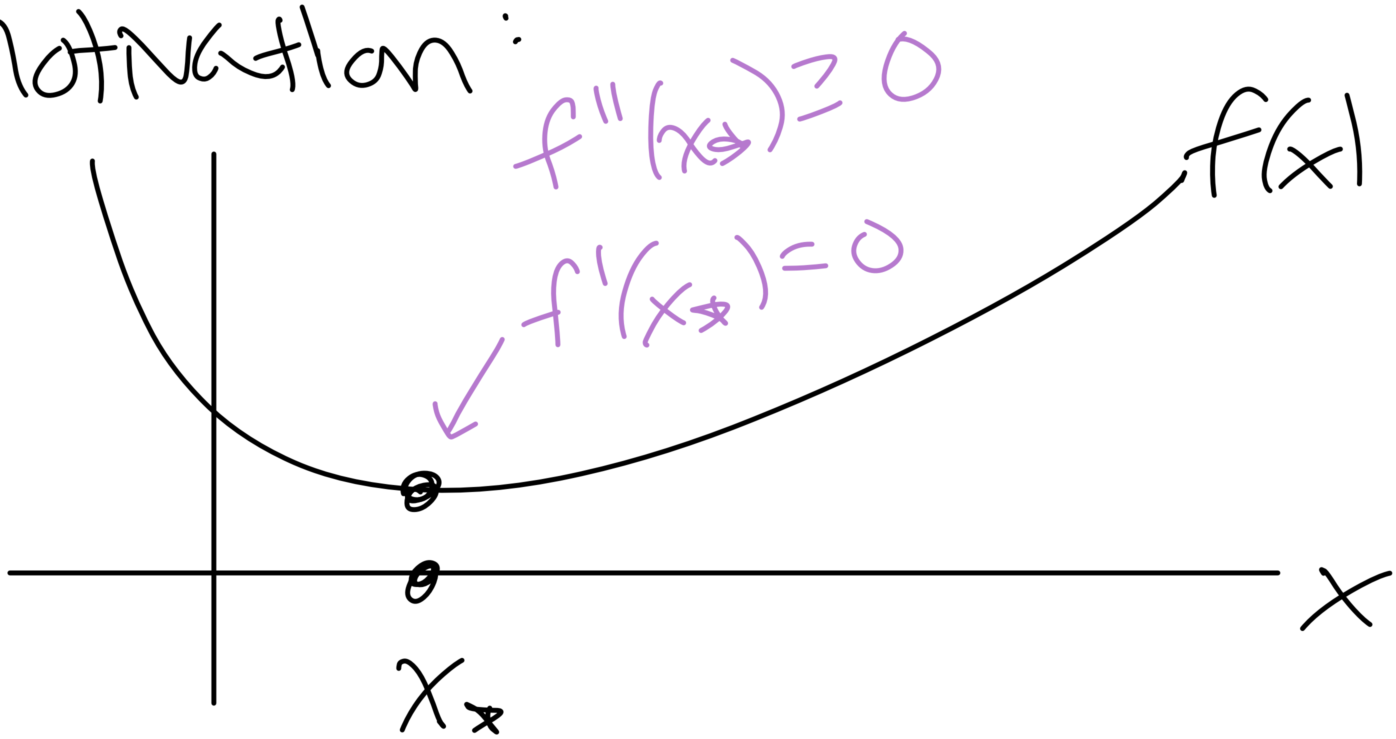
$$\mathcal{F}(\rho) := \int_{\mathbb{R}^d} f(\rho)\,dx$$

$$f(s)=\begin{cases} s\log s & s\ge 0 \\ +\infty & \text{otherwise}\end{cases}$$

$\mathcal{F}$ is the entropy.

To prove convergence of proximal point method we will use a generalization of the notion of derivative to characterize optimizers.

Motivation:

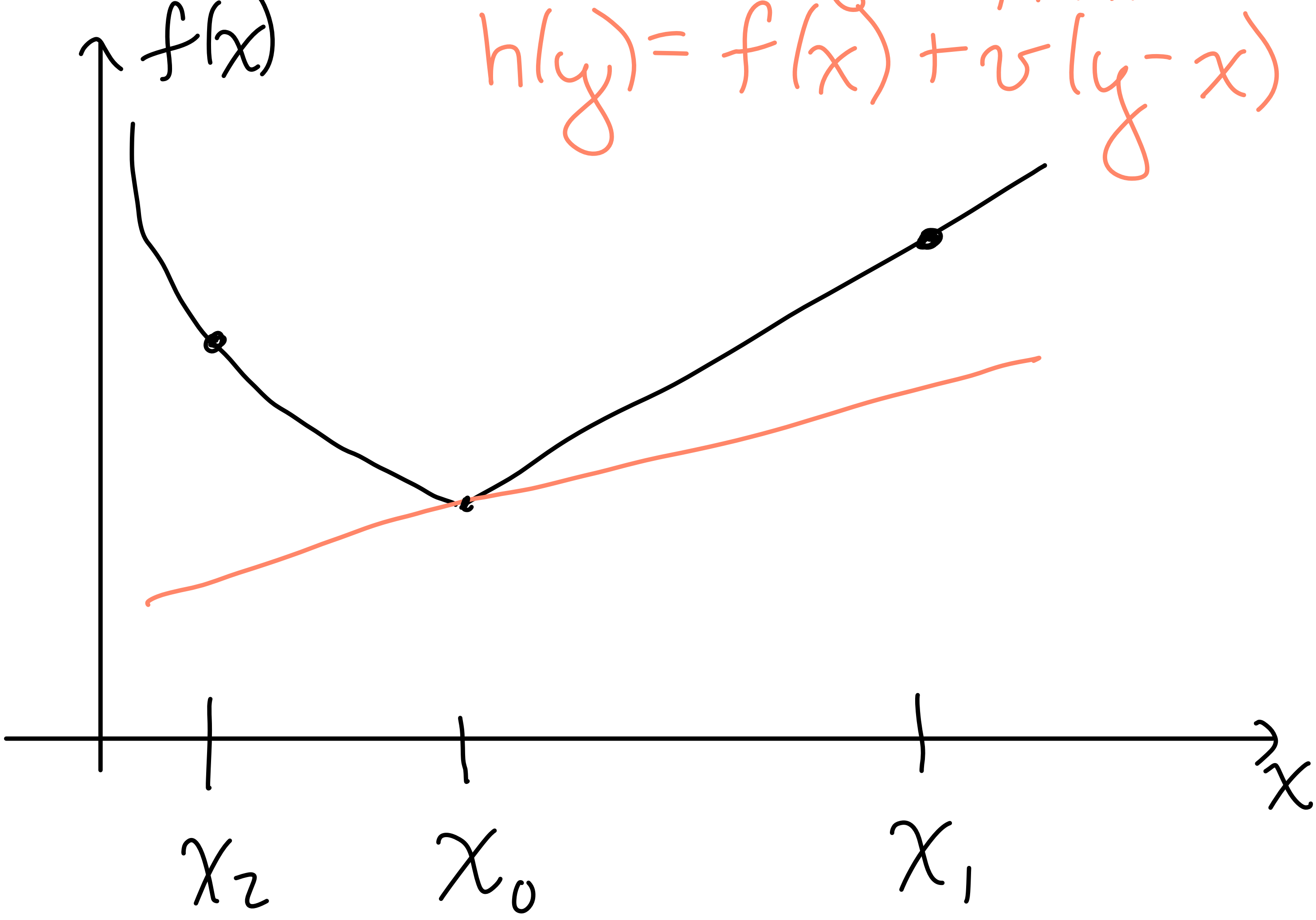


Def: Given $f: \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ convex, the <u>subdifferential</u> of f at $x \in \mathbb{R}$ is

$$\partial f(x) := \{v \in \mathbb{R} : f(y) \geq f(x) + v(y-x) \ \forall y \in \mathbb{R}\}$$

line with slope v through $(x, f(x))$

$h(y) = f(x) + v(y-x)$



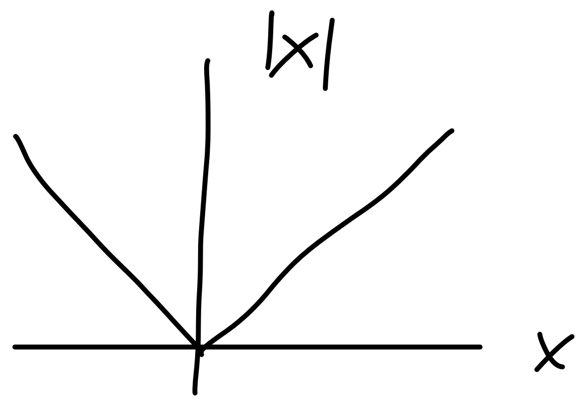
Def: $f: \mathbb{R} \to \mathbb{R} \cup \{+\infty\}$ is proper if $\bigcup D(f) \neq \emptyset$.

Remark:

f is proper and $x \notin D(f) \Rightarrow \partial f(x) = \emptyset$.

x is a minimizer of $f \Leftrightarrow 0 \in \partial f(x)$

Ex:



$f(x) = |x|$

$$\partial f(x) = \begin{cases} [-1, 1] & x = 0 \\ 1 & x > 0 \\ -1 & x < 0 \end{cases}$$

There is a nice relationship between subdifferential of a function and its derivatives (when they exist...)

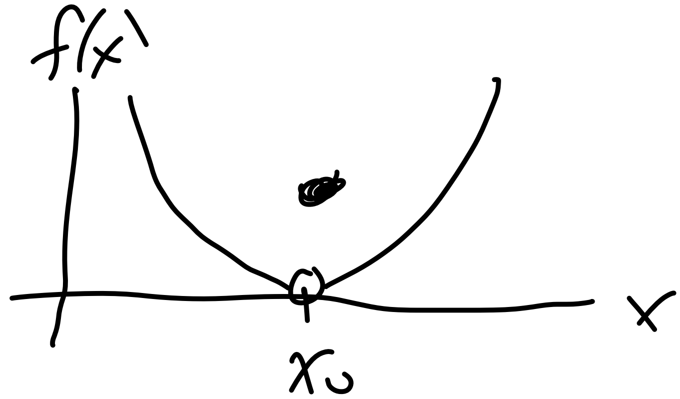
To talk about derivatives. _

Def: Given a function $f:\mathbb{R}\to\overline{\mathbb{R}}$, $a\in\overline{\mathbb{R}}$, $L\in\overline{\mathbb{R}}$, we say

$$\lim_{x\to a} f(x) = L$$

if, for all sequences $x_n\in\mathbb{R}\setminus\{a\}$ s.t. $x_n\to a$, $\lim_{n\to\infty} f(x_n) = L$.

Ex:



$$\lim_{x\to x_0} f(x) = 0$$

The previous def is sometimes known as "two sided limit".

The following is "one sided limit":

Def: Given a function $f: \overline{\mathbb{R}} \to \overline{\mathbb{R}}$, $a \in \mathbb{R}$, $L \in \overline{\mathbb{R}}$, we say

$$\lim_{x \to a^+} f(x) = L$$

"limit as x approaches a from above"

if, for all sequences $x_n \in (a, +\infty)$ s.t. $x_n \to a$ $\lim_{n \to \infty} f(x_n) = L$.

Likewise, we define $\lim_{x \to a^-} f(x) = L$.

Prop: Given a function $f:\mathbb{R}\to\overline{\mathbb{R}}$, $a\in\mathbb{R}$, $L\in\overline{\mathbb{R}}$,

$$\lim_{x\to a} f(x) = L \text{ iff } \lim_{x\to a^+} f(x) = \lim_{x\to a^-} f(x) = L.$$

Pf:
"$\Rightarrow$" is immediate.

Now, we show "$\Leftarrow$".

Fix $x_n : \mathbb{N}\to\mathbb{R}\setminus\{a\}$ s.t. $x_n\to a$.
We must show $\lim_{n\to\infty} f(x_n) = L$.

It suffices to show that every subsequence $f(x_{n_k})$ has a further subsequence

$f(x_{n_{k_\ell}})$ s.t. $\lim_{\ell \to \infty} f(x_{n_{k_\ell}}) = L$.

Fix $f(x_{n_k})$. Either $x_{n_k} > a$ for ∞-many k or $x_{n_k} < a$ for ∞-many k. WLOG suppose the former. Let $x_{n_{k_\ell}}$ be the subsequence of elements s.t. $x_{n_{k_\ell}} > a$.

By one sided limits,

$$\lim_{\ell \to \infty} f(x_{n_{k_\ell}}) = L. \qquad \square$$