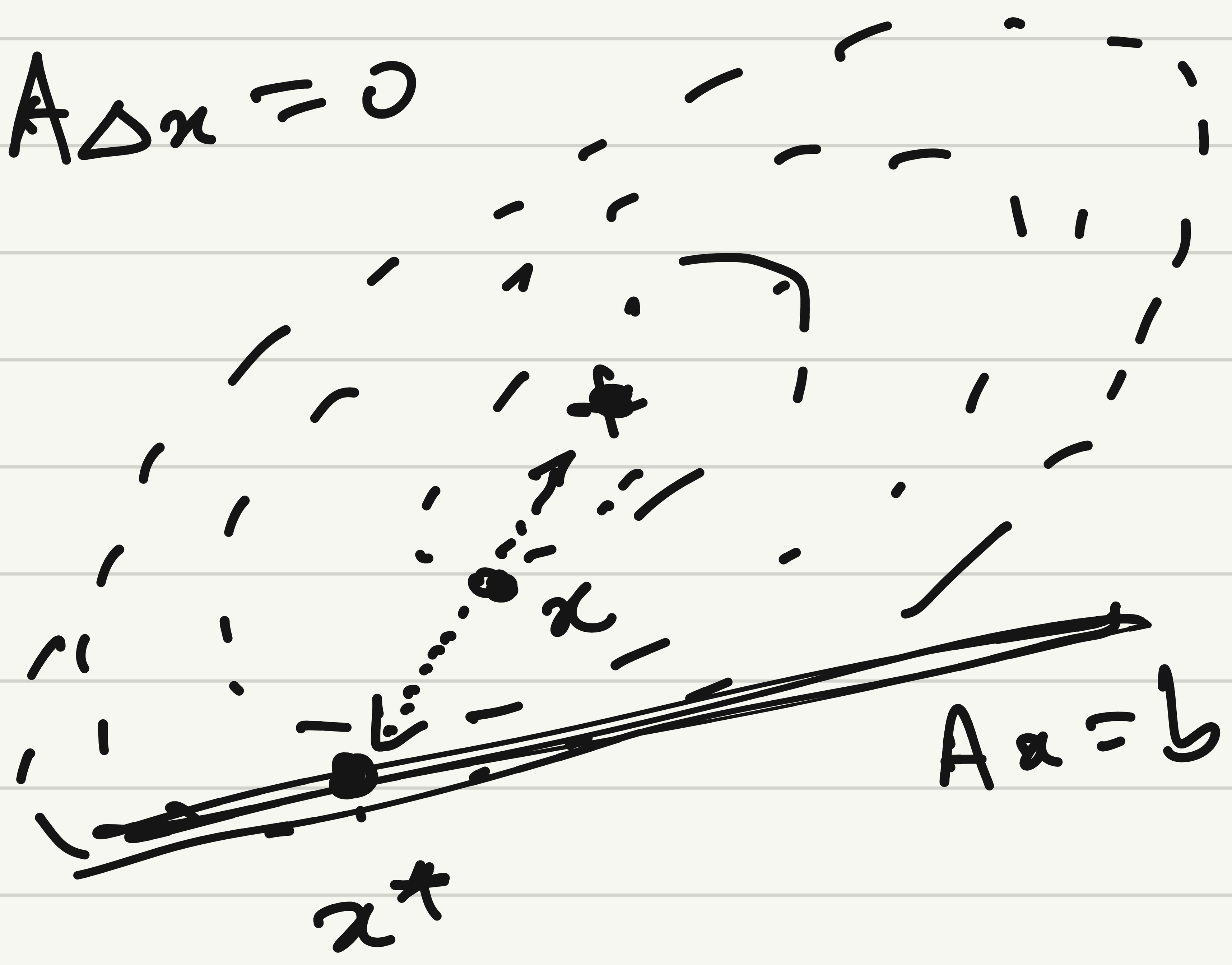
Equality constrained optimization

$$\begin{aligned} \min\ & f(x) \\ \text{s.t.}\ & Ax = b \end{aligned} \quad \longleftrightarrow \quad \left.\begin{aligned} r_{dual}(x,\nu) &= \nabla f(x) + A^T \nu = 0 \\ r_{pri}(x,\nu) &= Ax - b = 0 \end{aligned}\right\} \quad r(x,\nu) = \begin{bmatrix} r_{dual}(x,\nu) \\ r_{pri}(x,\nu) \end{bmatrix}$$

$$\begin{bmatrix} \nabla^2 f(x) & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \Delta \nu \end{bmatrix} = \begin{bmatrix} -\nabla f(x) - A^T \nu \\ b - Ax \end{bmatrix}$$

$A\Delta x = 0$

$$\begin{bmatrix} \nabla^2 f & A^T \\ A & 0 \end{bmatrix} \begin{bmatrix} \Delta x \\ \nu \end{bmatrix} = \begin{bmatrix} -\nabla f(x) \\ b - Ax \end{bmatrix}$$



current x : $Ax \neq b$

$x^+ = x + \Delta x$

$\boxed{\text{Infeasible start Newton's method}}$

Δx may not be descent dir for f

$\begin{bmatrix} \Delta x \\ \Delta \nu \end{bmatrix}$ <u>is</u> descent dir. for $\|r\|$

repeat:

Solve KKT sys for $\begin{bmatrix} \Delta x \\ \Delta \nu \end{bmatrix}$

Do backtracking LS on $\| r(x + t\Delta x, \nu + t\Delta \nu) \|$

update $(x, \nu) \mathrel{+}= t(\Delta x, \Delta \nu)$

until $Ax = b$ and $\|r\| < \varepsilon$

$\Rightarrow$ If $t = 1$, $x + \Delta x$ is exactly feasible: $A(x + \Delta x) = b$

$\rightarrow$ If x is feasible, equiv to previous Newton's method

Inequality constrained optimization

$$\min \; f_0(x)$$
$$\text{s.t.} \quad f_i(x) \le 0 \qquad i = 1, \dots, m$$
$$h_i(x) = 0 \qquad i = 1, \dots, p \xrightarrow{\text{convex}} Ax = b, \quad A \in \mathbb{R}^{p \times n}$$

$\longrightarrow$ $*$

Lagrangian: $L(x, \lambda, \nu) = f_0(x) + \sum \lambda_i f_i(x) + \sum \nu_i h_i(x)$

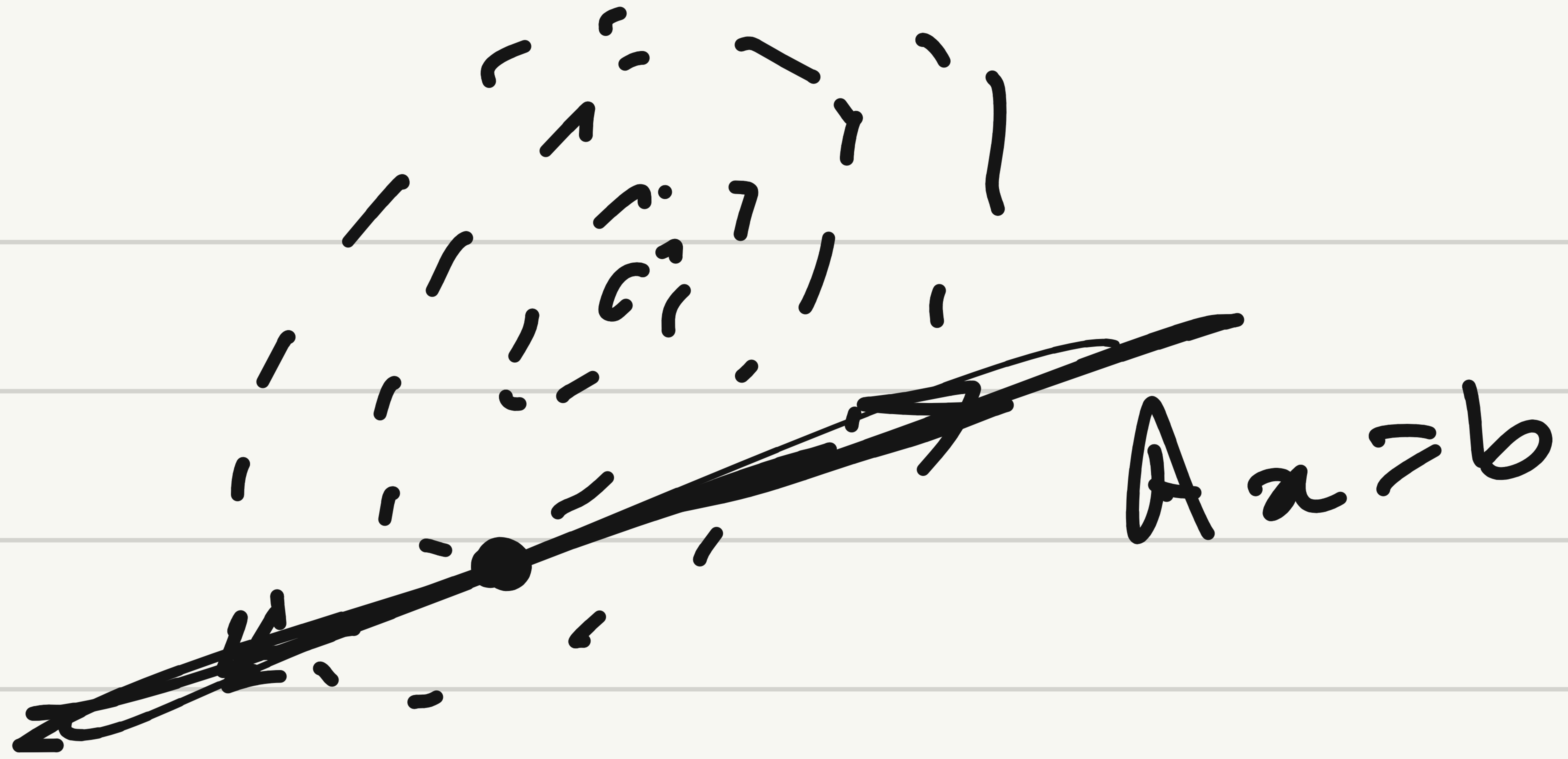
$A^T \nu$

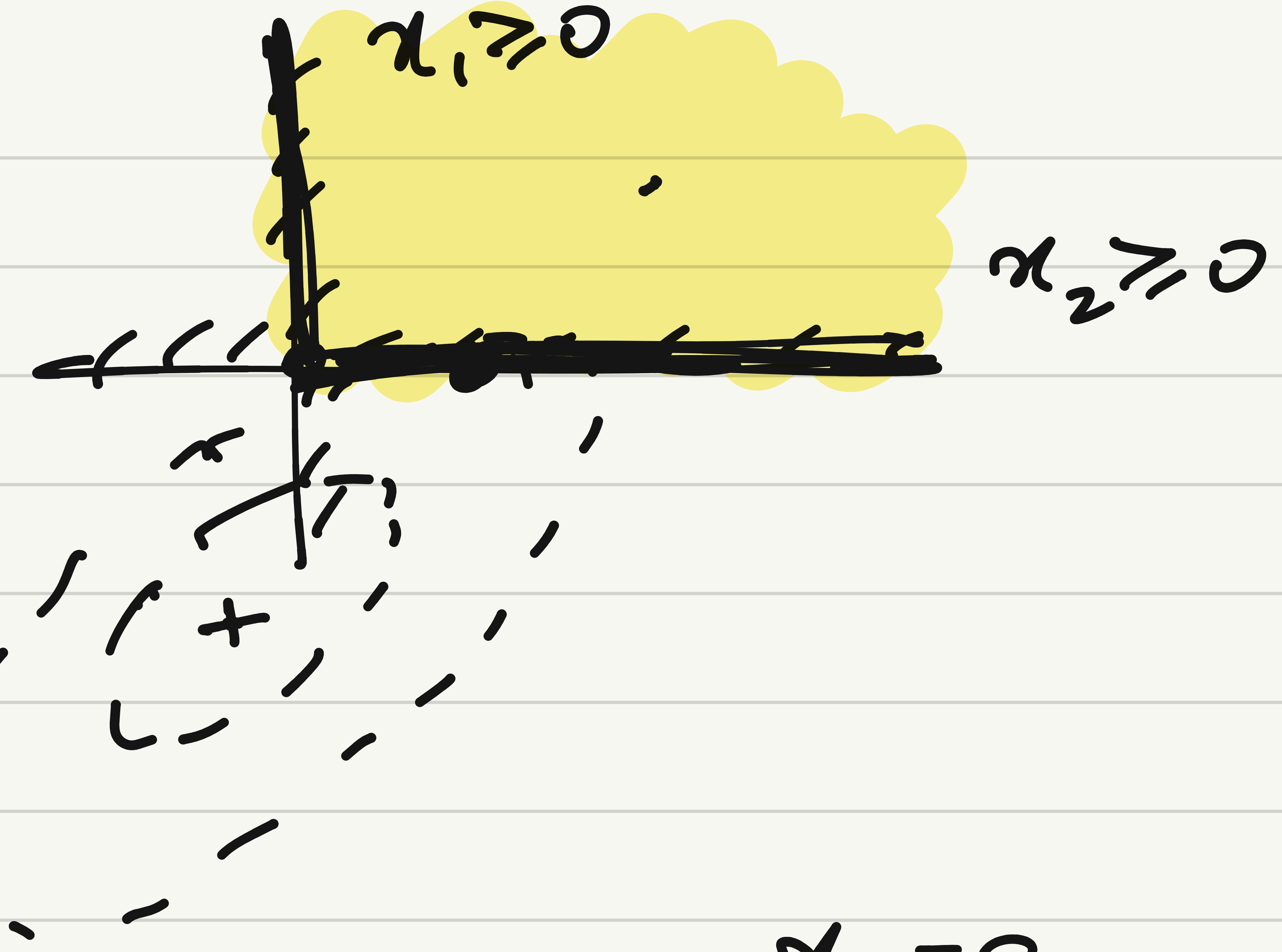
Dual fn: $g(\lambda, \nu) = \inf\limits_{x} L(x, \lambda, \nu)$

for any feasible x, any $\lambda \ge 0$, any ν,

$$\underbrace{f_0(x)} \ge p^* \ge d^* \ge \underbrace{g(\lambda, \nu)}$$

Duality gap: $f_0(x) - g(\lambda, \nu)$





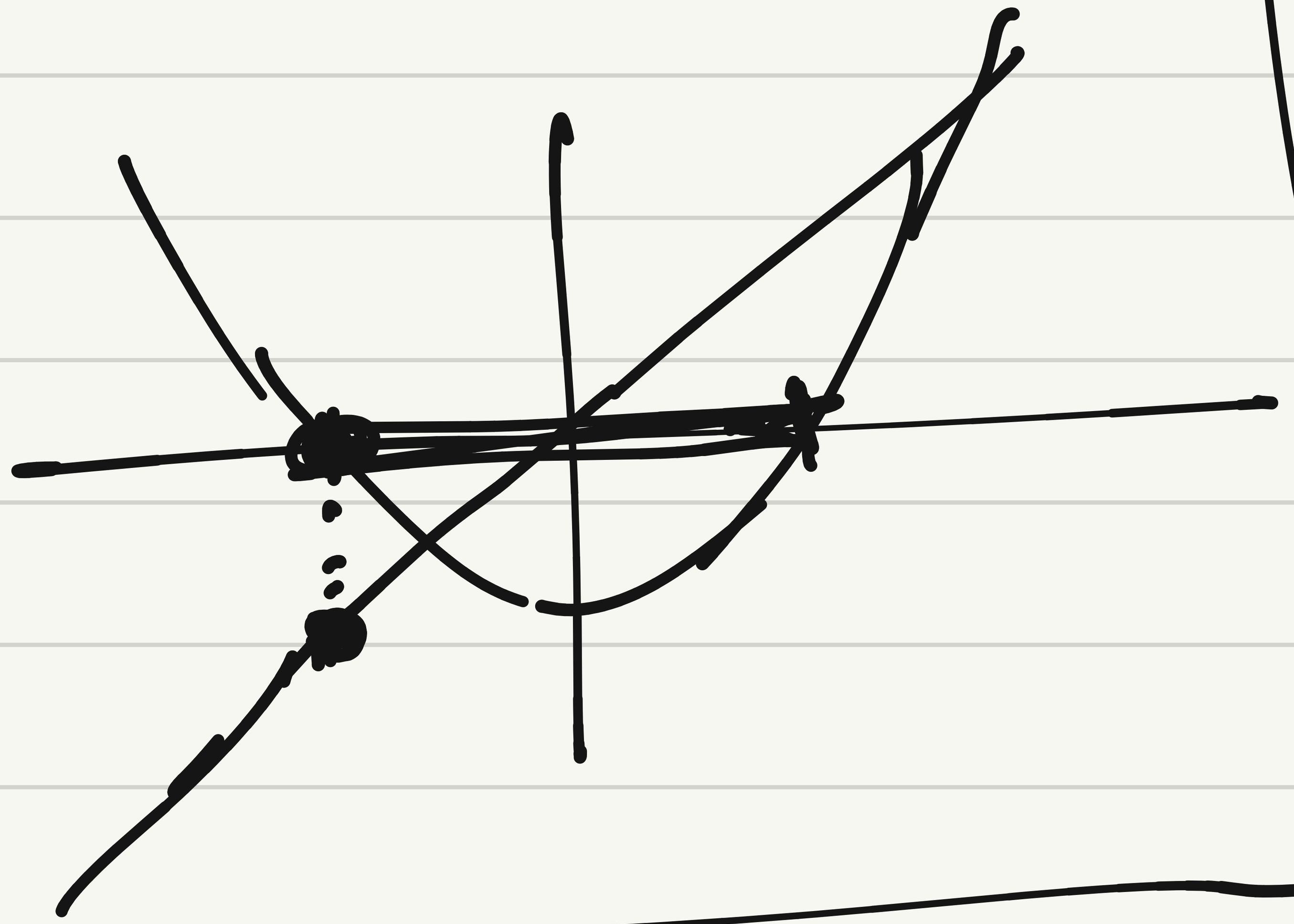
active : $f_i(x) = 0$

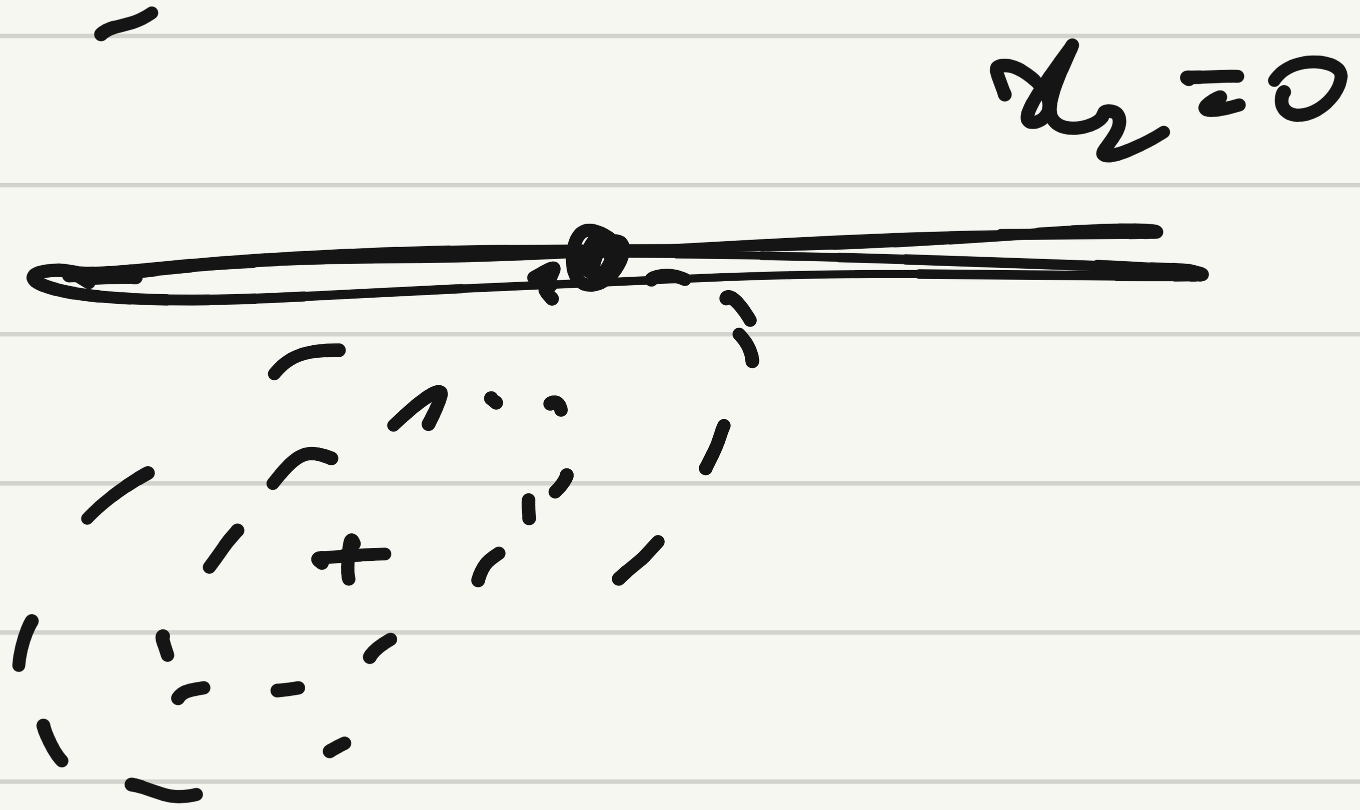
inactive : $f_i(x) < 0$

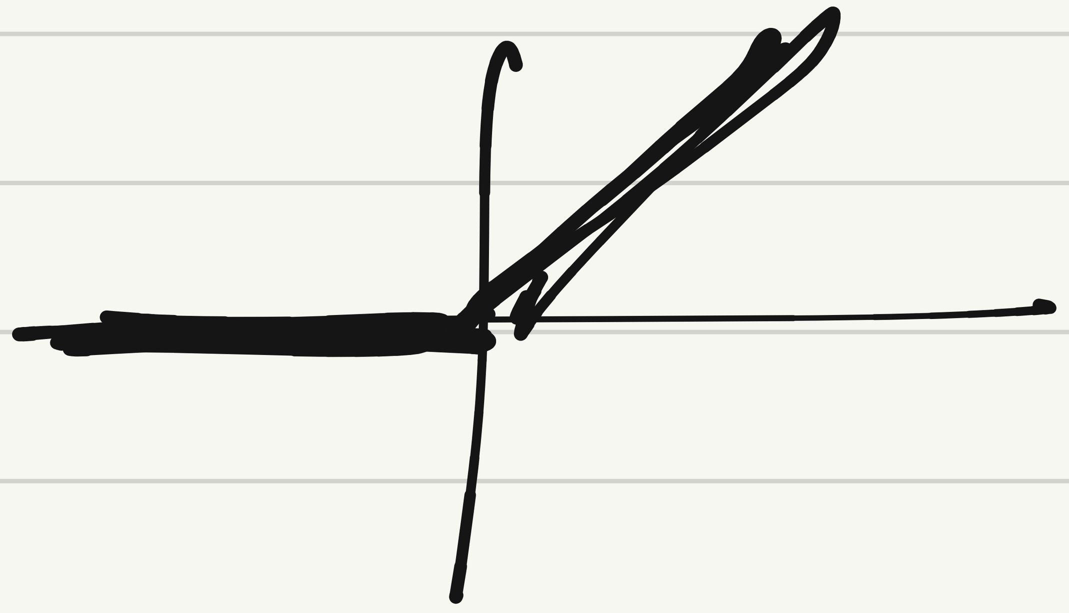
m ineq cons $\Rightarrow 2^m$ possibilities

min x

s.t. $x^2 - 1 \leq 0$



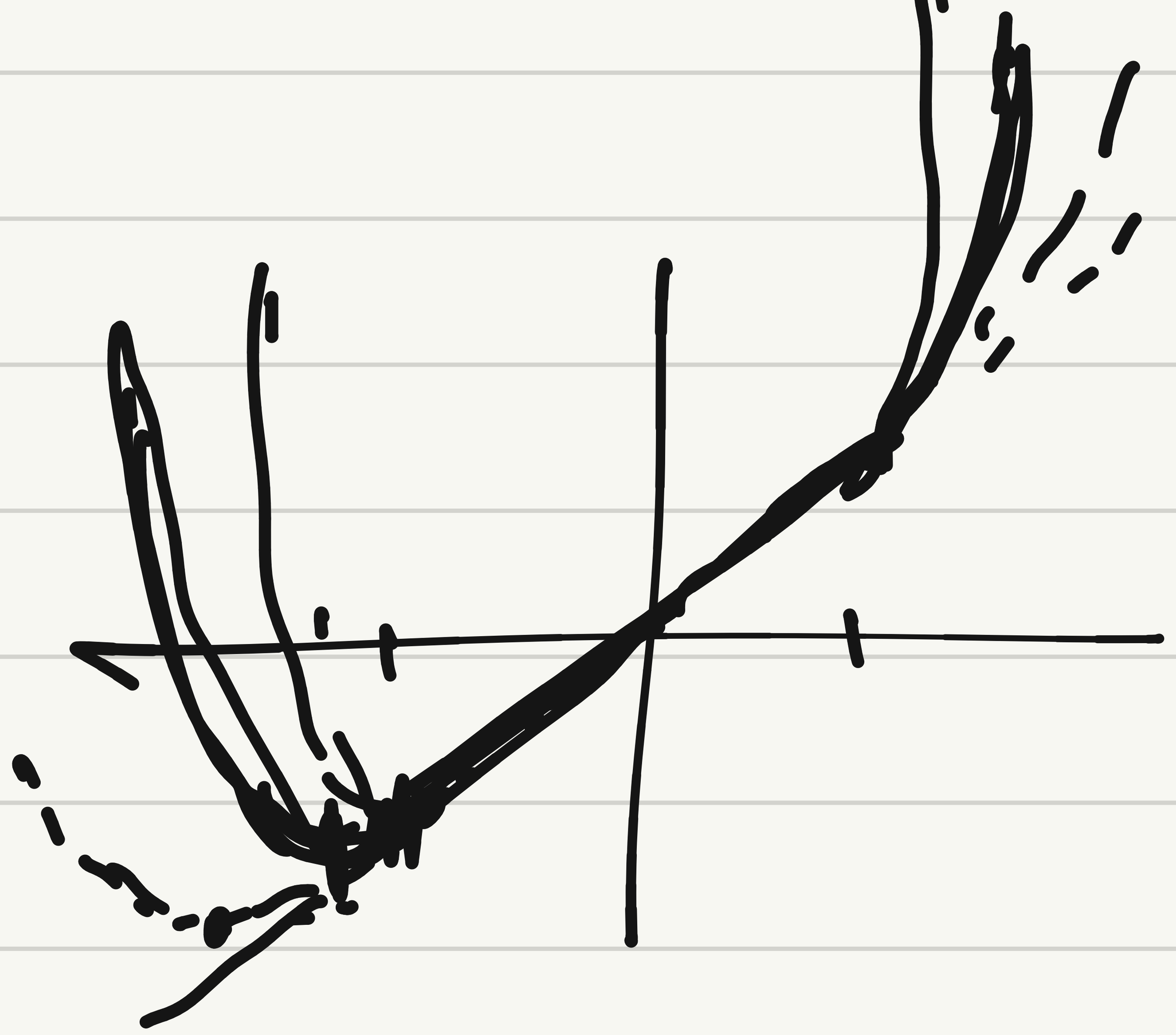




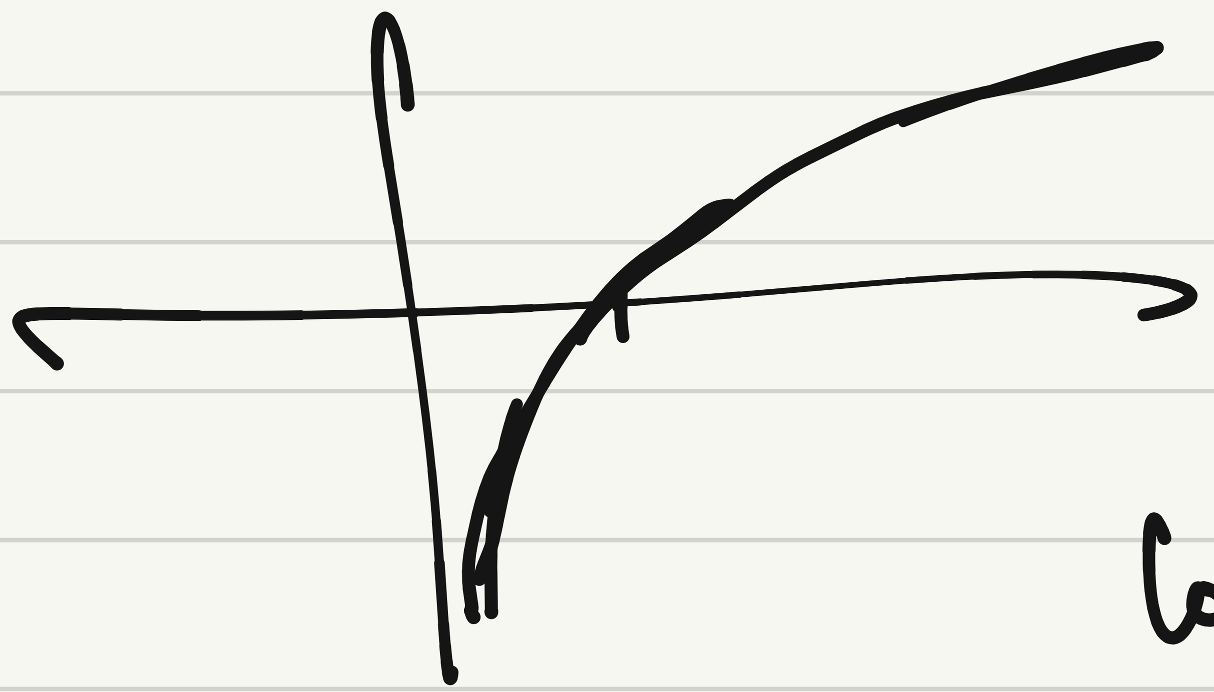
$[u]^+ = \max(u, 0)$

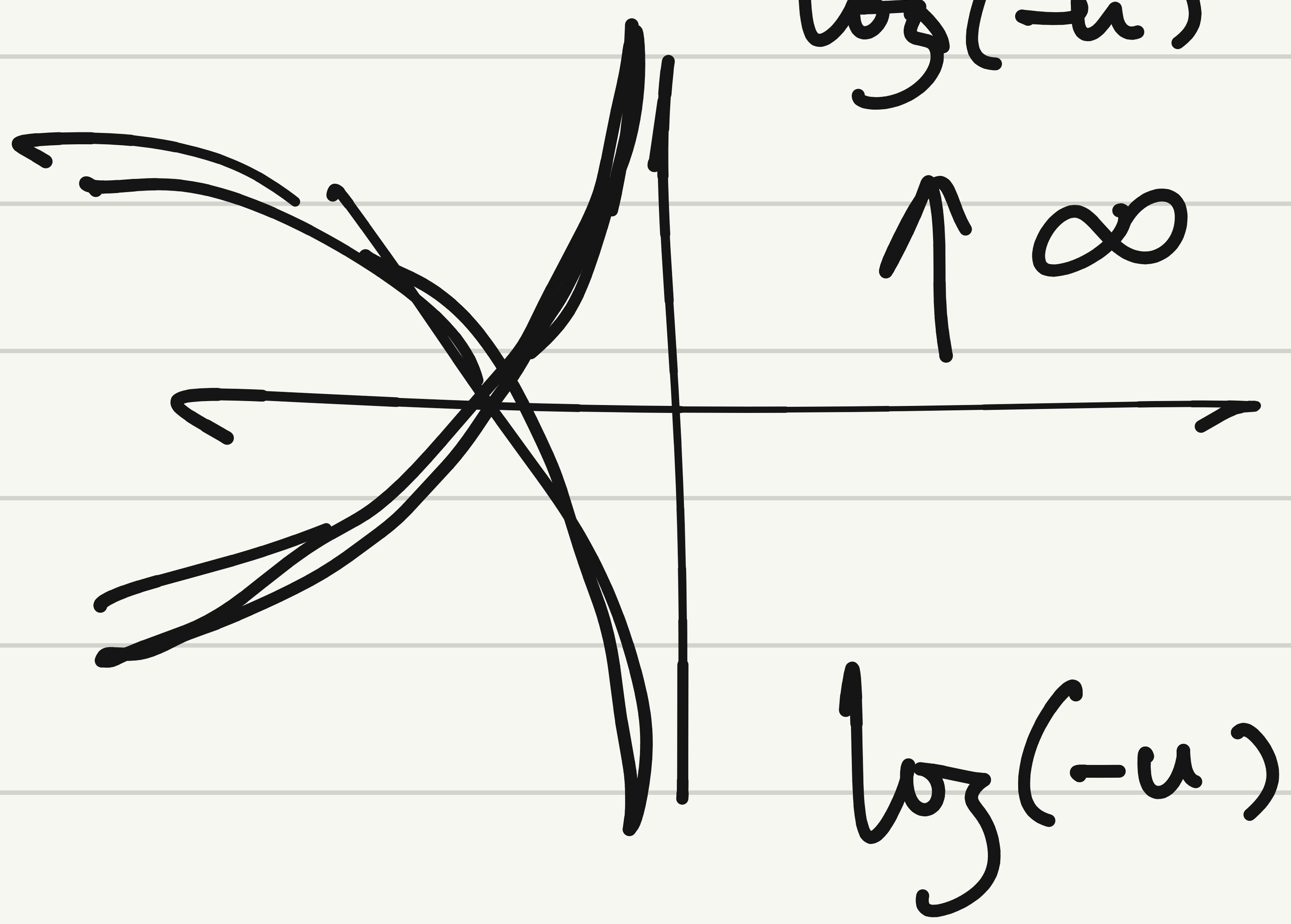
$\boxed{\text{Penalty method}}$: eg. $\min f_0(x) + \frac{k}{2} \sum \underbrace{\left([f_i(x)]^+\right)^2}$

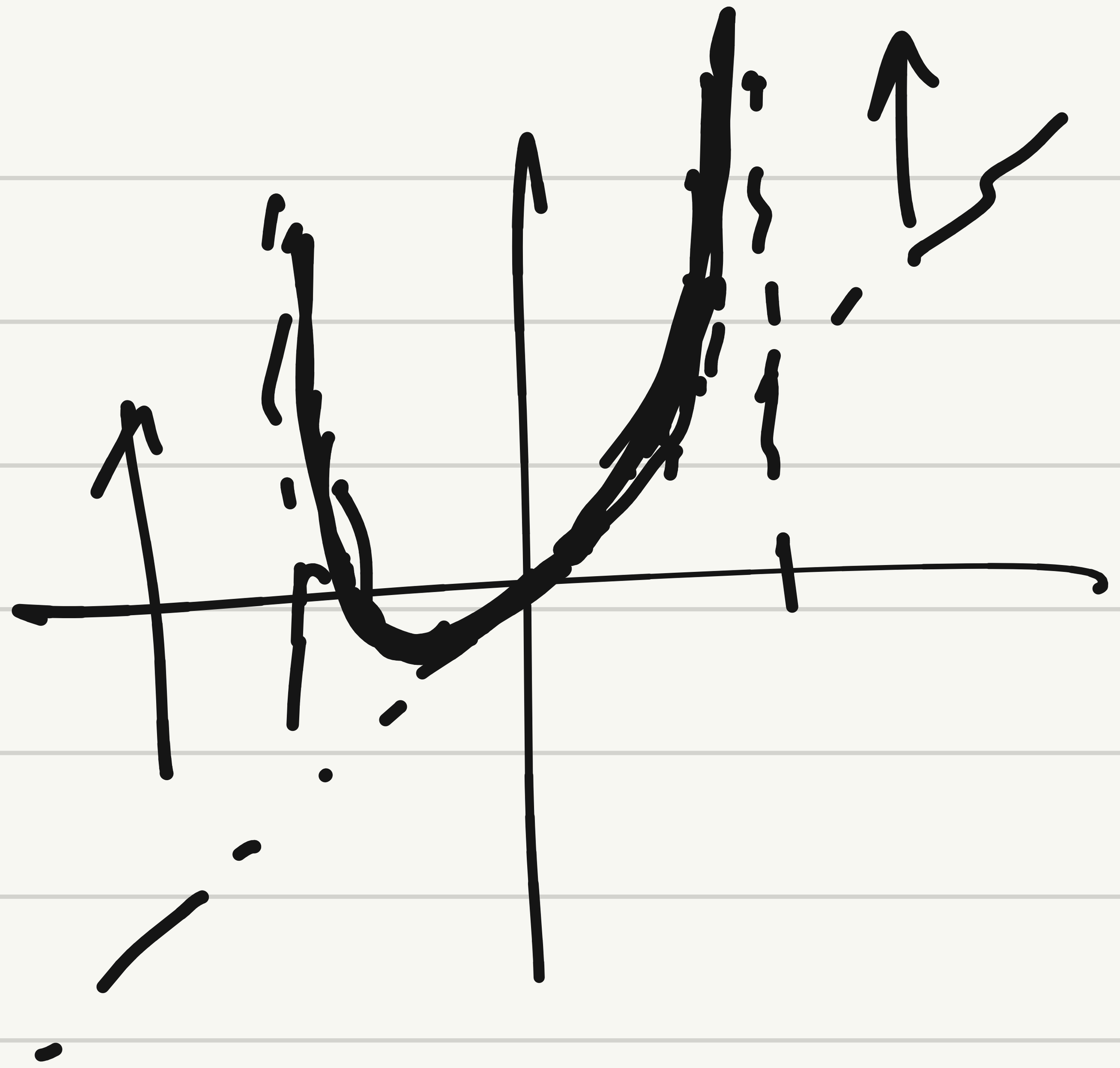
As $k \to \infty$, uncon. opt $\to$ cons. opt

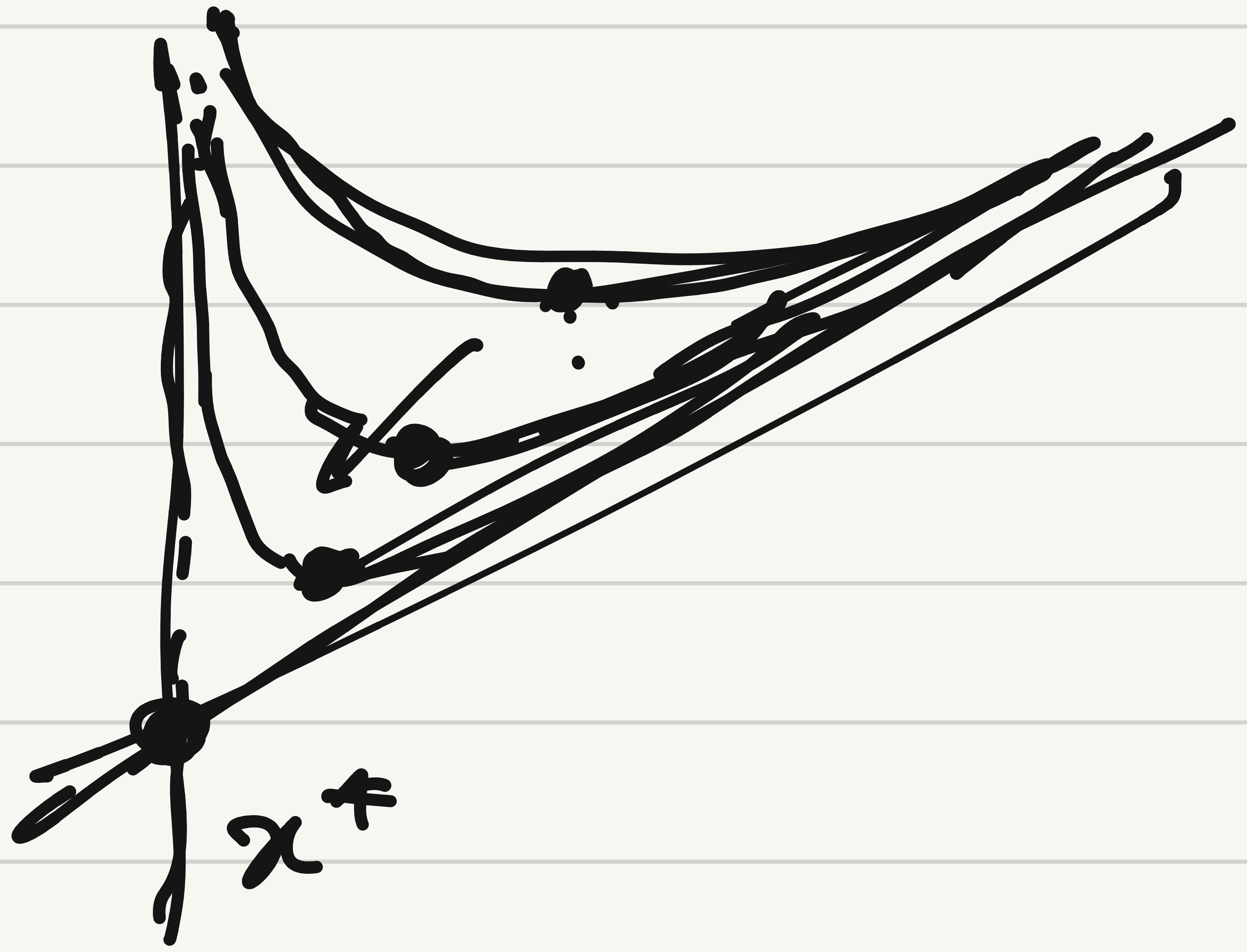


$\boxed{\text{Barrier method}}$ eg. $f_0(x) + k \sum -\log(-f_i(x))$









As $k \to 0$, sol $\to$ con. opt. sol x^*

$\boxed{\text{Augmented Lagrangian method}}$: penalty term

+ Lagrangian term $\sum \lambda_i f_i(x)$

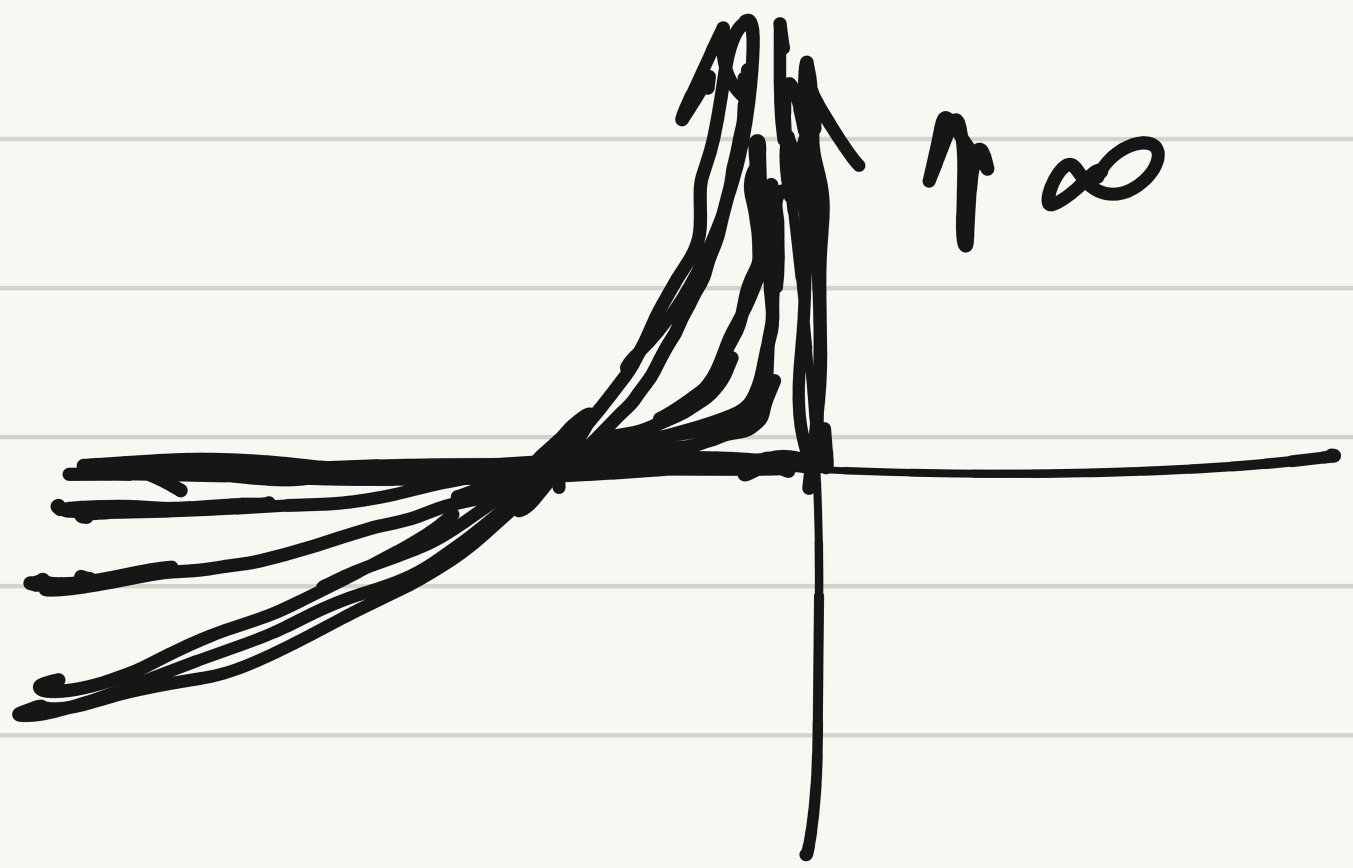
$k = 1/t, \quad t \to \infty$

Barrier method

$$\min f_0(x) \quad \text{s.t.} \quad f_i(x) \le 0, \quad i = 1, \dots, m, \quad Ax = b \qquad (*)$$

$$\longrightarrow \quad \min f_0(x) + \frac{1}{t} \underbrace{\sum -\log(-f_i(x))} \quad \text{s.t.} \quad Ax = b \qquad (**)$$

As $t \to \infty$, $-\frac{1}{t}\log(-u) \to \begin{cases} 0 & \text{if } u < 0 \\ \infty & \text{if } u \geq 0 \end{cases}$

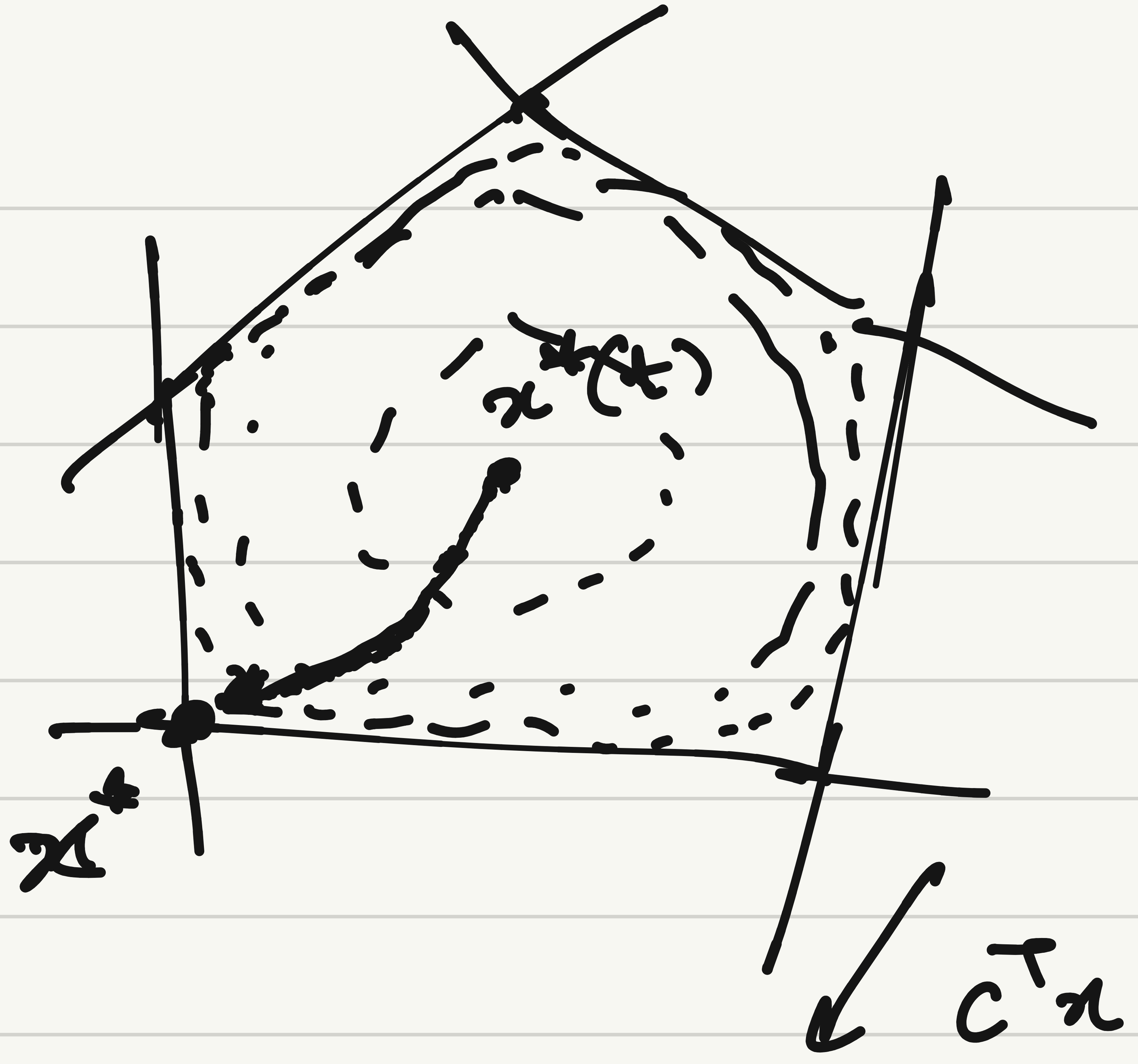


also $-\log(-u)$ is convex & increasing $\Rightarrow$ problem remains convex.

$$\phi(x) = \sum_{i=1}^{m} -\log(-f_i(x)) \quad : \quad \boxed{\text{log barrier function}}$$

$$\Rightarrow \min f_0(x) + \frac{1}{t}\phi(x)$$

for diff values of t, opt point for $(**)$ $= x^*(t)$: $\boxed{\text{central path}}$



As $t \to \infty$, $x^*(t) \to x^*$, $f_0(x^*(t)) \to p^*$

$x^*(t)$ is always strictly feasible : $f_i(x^*(t)) < 0$

$Ax^*(t) = b$

$\underline{x^*(t)}$ minimizes $f_0(x) + \frac{1}{t} \sum -\log(-f_i(x))$

s.t. $Ax = b$

$\to \nabla f_0(x) + \sum \left(\frac{1}{t f_i(x)} \nabla f_i(x) \right) + A^T \underline{\nu} = 0$

Compare $\nabla L(x, \lambda, \nu) = \nabla f_0(x) + \sum \lambda_i \nabla f_i(x) + A^T \nu$

Define dual varible estimates $\begin{cases} \lambda_i^*(t) = \dfrac{1}{t f_i(x^*(t))} \\ \nu_i^*(t) = \nu_i^*(t) \text{ coming from} \end{cases}$

then $\underline{\nabla L(x^*, \lambda^*, \nu^*) = 0}$

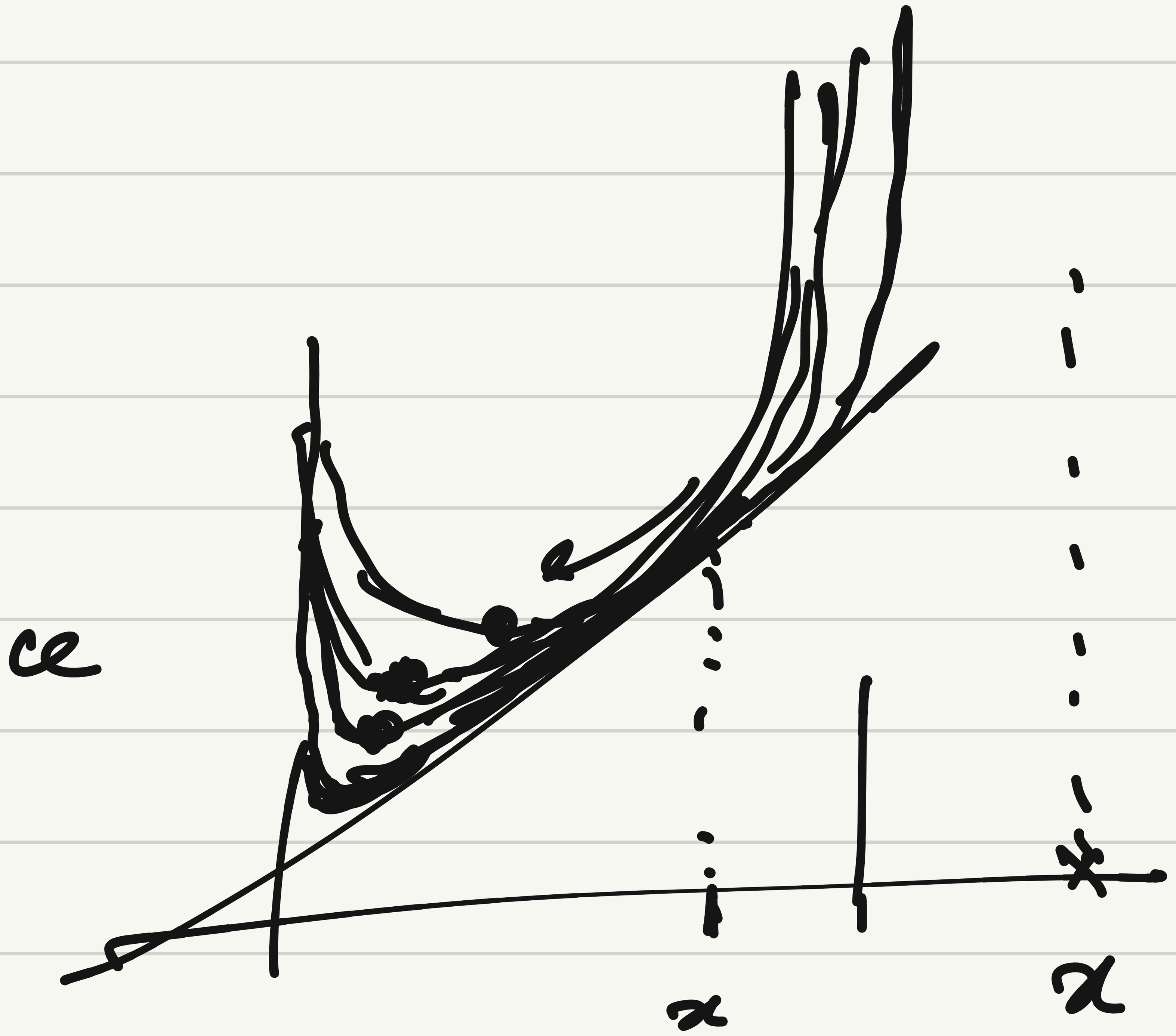
$\lambda^*(t), \nu^*(t)$ dual feasible : $\lambda_i^*(t) \geq 0$

$$g(\lambda^*(t), \nu^*(t)) = L(x^*(t), \lambda^*(t), \nu^*(t)) = f_0(x^*(t)) - m/t$$

duality gap $\boxed{f_0(x^*(t)) - g(\lambda^*(t), \nu^*(t)) = m/t}$

As $t \to \infty$, duality gap $\to 0$, $f_0(x^*(t)) \to p^*$

If accuracy $\varepsilon \Rightarrow t \geq m/\varepsilon$



Algorithm: increase t gradually

start with strictly feas. x, $t > 0$, $\underset{\text{param}}{\mu > 1}$, $\underset{\text{tolerance}}{\varepsilon > 0}$

Repeat:

1. Solve (**) to get $x^*(t)$ centering step
2. if $m/t < \varepsilon$: return $x^*(t)$
3. $t = \mu t$

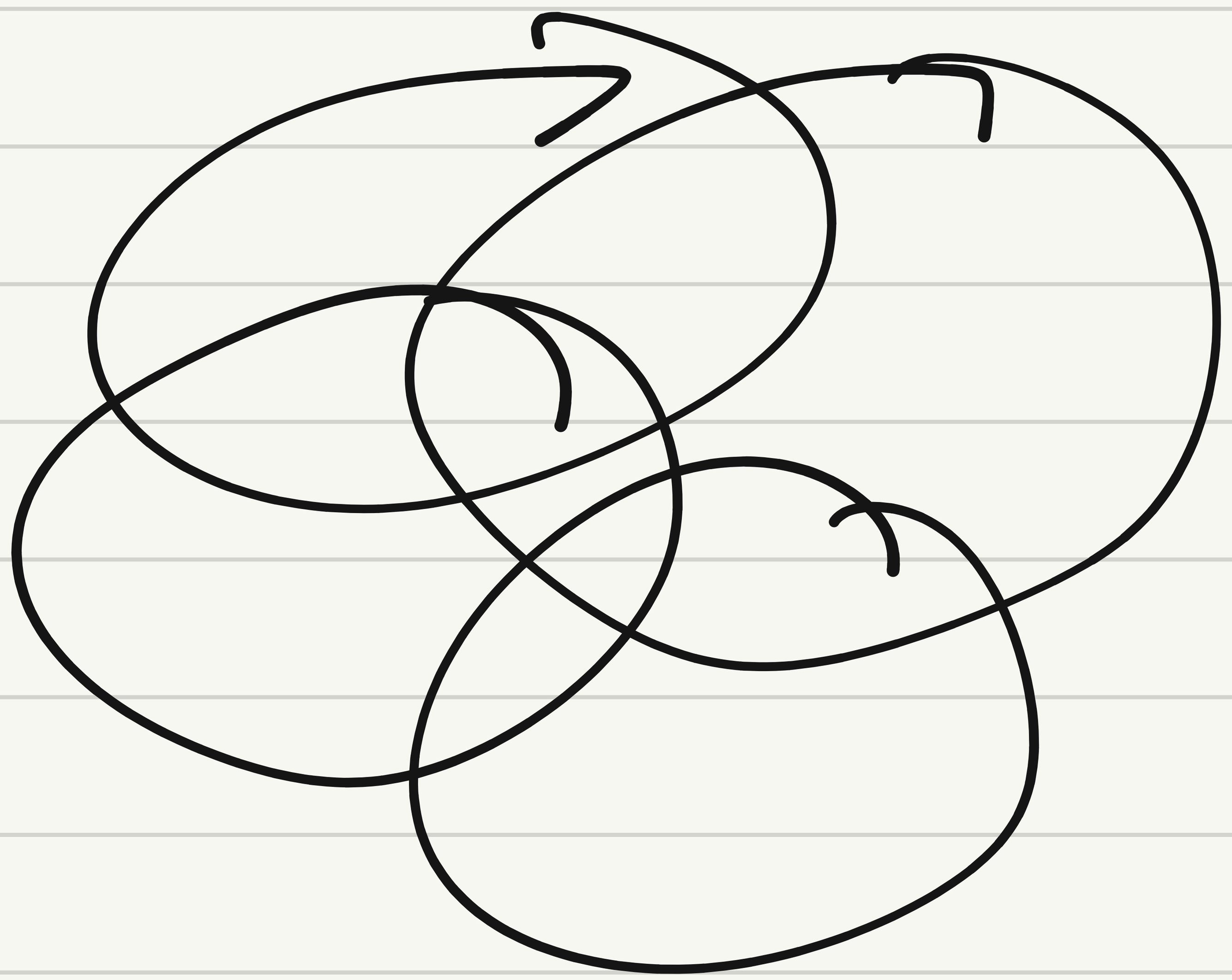
Choice of μ: $\left.\begin{array}{l}\text{small} \Rightarrow \text{many outer iters, few inner iters} \\ \text{large} \Rightarrow \text{few outer iters, many inner iters}\end{array}\right\}$ $\mu \approx 10 \sim 20$

<u>feasibility</u> How to find initial x s.t. $f_i(x) < 0$ for all $i = 1, \dots m$,

and $Ax = b$?

<u>Phase 1</u>: find strictly feasible point <u>or</u> report infeasible

<u>Phase 2</u>: Barrier method



Define some fn $\psi(x)$ s.t. $\psi(x) \leq 0$ if x feasible

$\psi(x) > 0$ if infeasible

Then minimize $\psi(x)$

1. $\psi(x) = \max f_i(x)$ (maximum infeasibility)

$\min \psi(x)$ s.t. $Ax = b$

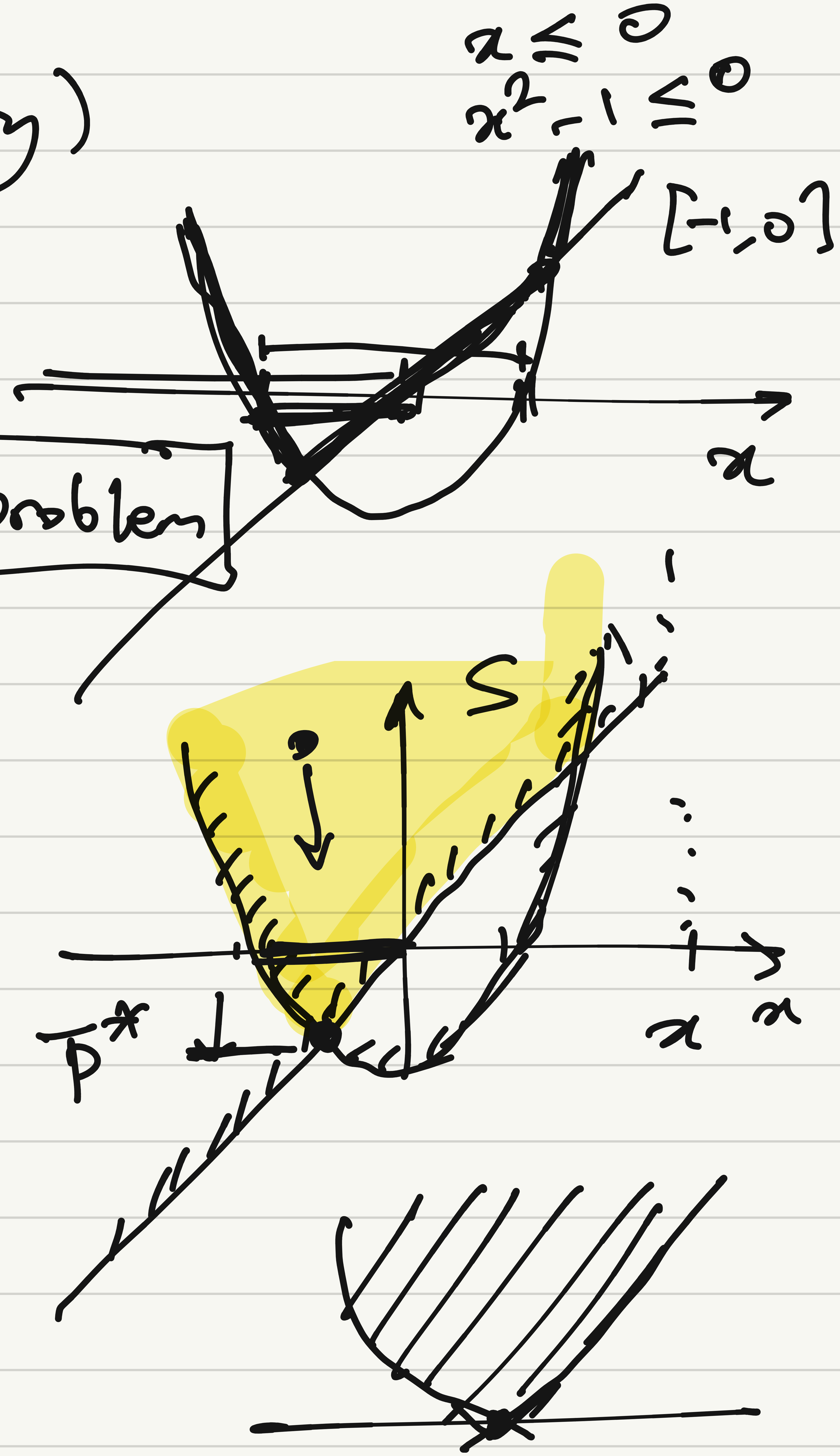
$\Leftrightarrow$ $\min \underset{\smile}{s}$ s.t. $\underline{s \geq f_i(x)}$, $Ax = b$: $\boxed{\text{Phase 1 opt problem}}$

feasible point: choose any x
choose $s \geq \max f_i(x)$

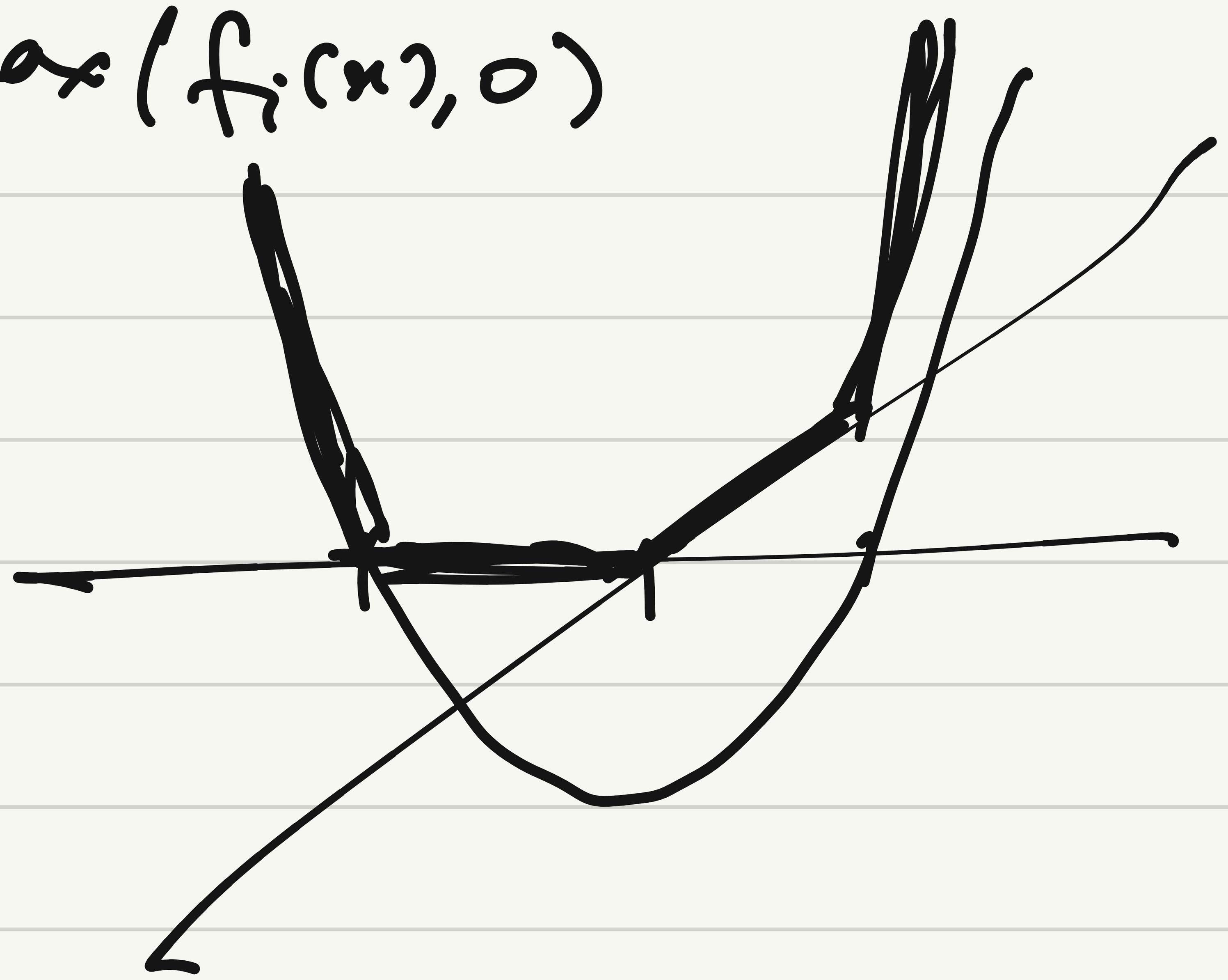
Solve Phase 1 problem w/ barrier method

<u>early termination</u>:

- if any $s < 0$ $\Rightarrow$ x is strictly feasible, return it
- if $g(\lambda, \nu) > 0$ $\Rightarrow$ problem is infeasible, return



2. $\Psi(x) =$ sum of infeasibilities $\sum_{i=1}^{m} [f_i(x)]^+$ $\max(f_i(x), 0)$



$\rightarrow$ $\min\ s_1 + \cdots + s_m$

s.t. $s_i \geq f_i(x)$, $Ax = b$, $s_i \geq 0$

if x feasible $\Rightarrow$ $\Psi(x) = 0$

If problem is infeasible $\Rightarrow$ $\Psi(x) > 0$ for all x