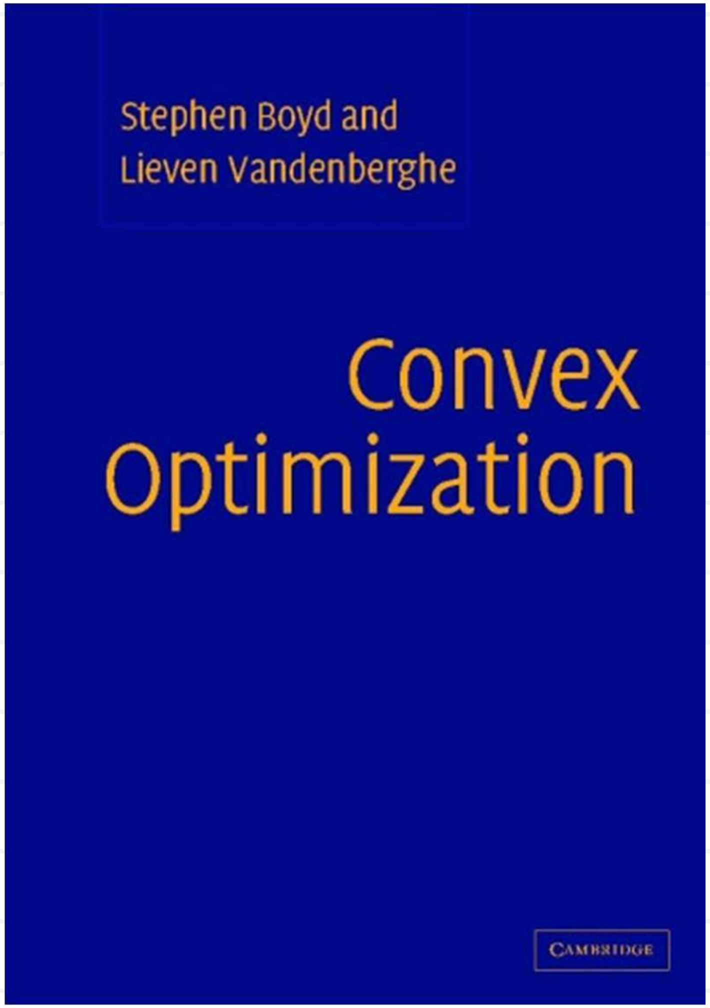
Lecture 10: Interior point methods

Reading



Section 11

Recap: Karush-Kuhn-Tucker conditions

$$\begin{cases} f_i(x^*) \le 0 \quad i = 1 \ldots m \qquad h_i(x^*) = 0 \quad i = 1 \ldots p \\ \lambda_i^* \ge 0 \quad i = 1 \ldots m \\ \nabla f_0(x^*) + \sum_{i=1}^{m} \lambda_i^* \nabla f_i(x^*) + \sum_{i=1}^{p} \nu_i^* \nabla h_i(x^*) = 0 \\ \lambda_i^* f_i(x^*) = 0 \quad i = 1 \ldots m \end{cases}$$

Optimization with inequality constraints

$$\min \; f_0(x)$$
$$\text{s.t.} \; f_i(x) \le 0$$

Option 1: active set methods (e.g. simplex)

- (tend to) stick to the boundary of *D*
- Can get slow for complex boundaries

Option 2: interior-point methods

- Do not stick to the boundary
- Converge to the optimum from inside
- Provably fast

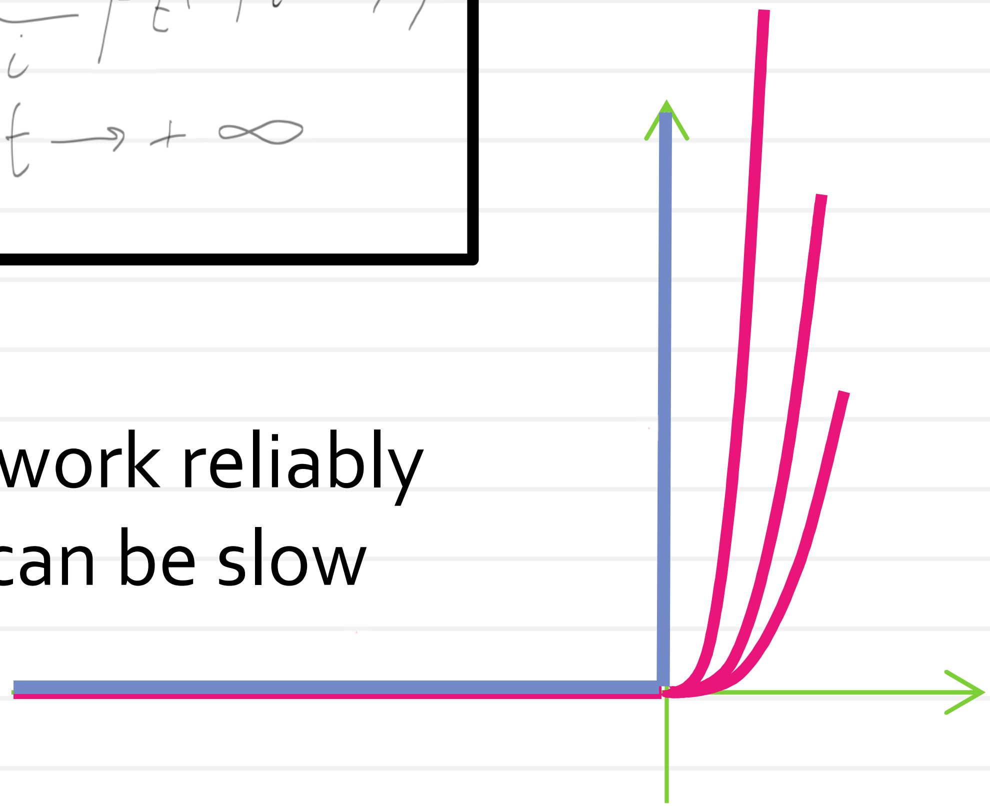
"Naïve" interior-point: penalty method

$$\min f_0(x) \quad \text{s.t.} \quad f_i(x) \le 0$$

$$\min_x f_0(x) + \sum_i \begin{cases} 0, & f_i(x) \le 0 \\ +\infty, & f_i(x) > 0 \end{cases}$$

$$\boxed{\min f_0(x) + \sum_i p_t(f_i(x)) \qquad t \to +\infty}$$

- Hard to make work reliably
- Convergence can be slow



"Smart" interior-point: barrier method

$$\min \; f_0(x) \quad \text{s.t.} \; f_i(x) \le 0$$

$$\min_x \; f_0(x) + \sum_i \begin{cases} 0, & f_i(x) \le 0 \\ +\infty, & f_i(x) > 0 \end{cases}$$

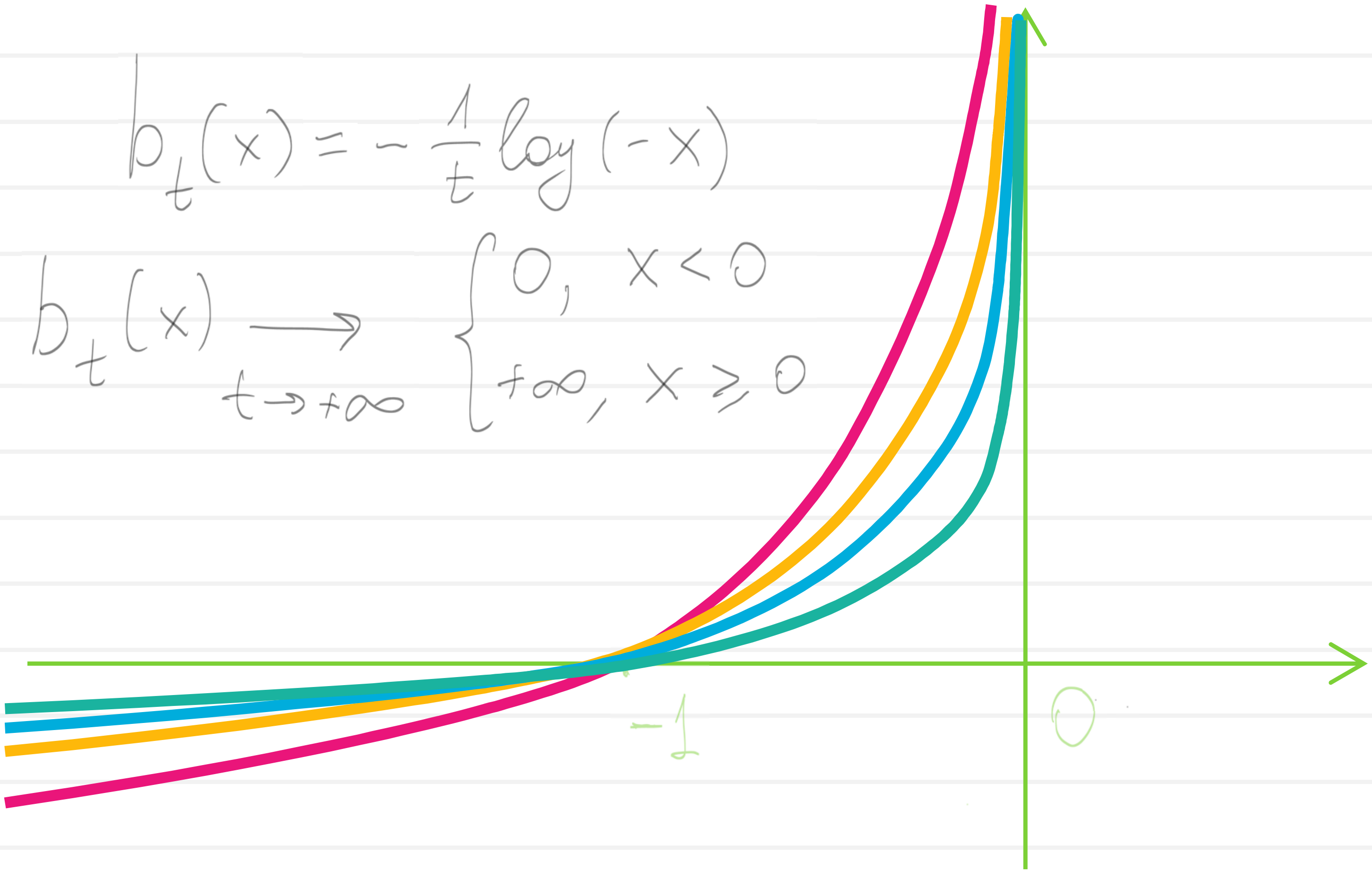
$$\boxed{\min \; f_0(x) + \sum_i b_t(f_i(x)) \qquad t \to +\infty}$$

- Polynomial time convergence for LP [Karmakar84]
- Highly efficient in practice

Logarithmic barrier

$$b_t(x) = -\frac{1}{t}\log(-x)$$

$$b_t(x) \xrightarrow[t \to +\infty]{} \begin{cases} 0, & x < 0 \\ +\infty, & x \geq 0 \end{cases}$$



$$\boxed{\min_x f_0(x) - \frac{1}{t}\sum \log(-f_i(x)) \quad t \to +\infty}$$

$$\boxed{\min_x t f_0(x) - \sum \log(-f_i(x)) \quad t \to +\infty}$$

Barrier method

Input: strictly feasible point

t = 1

Loop

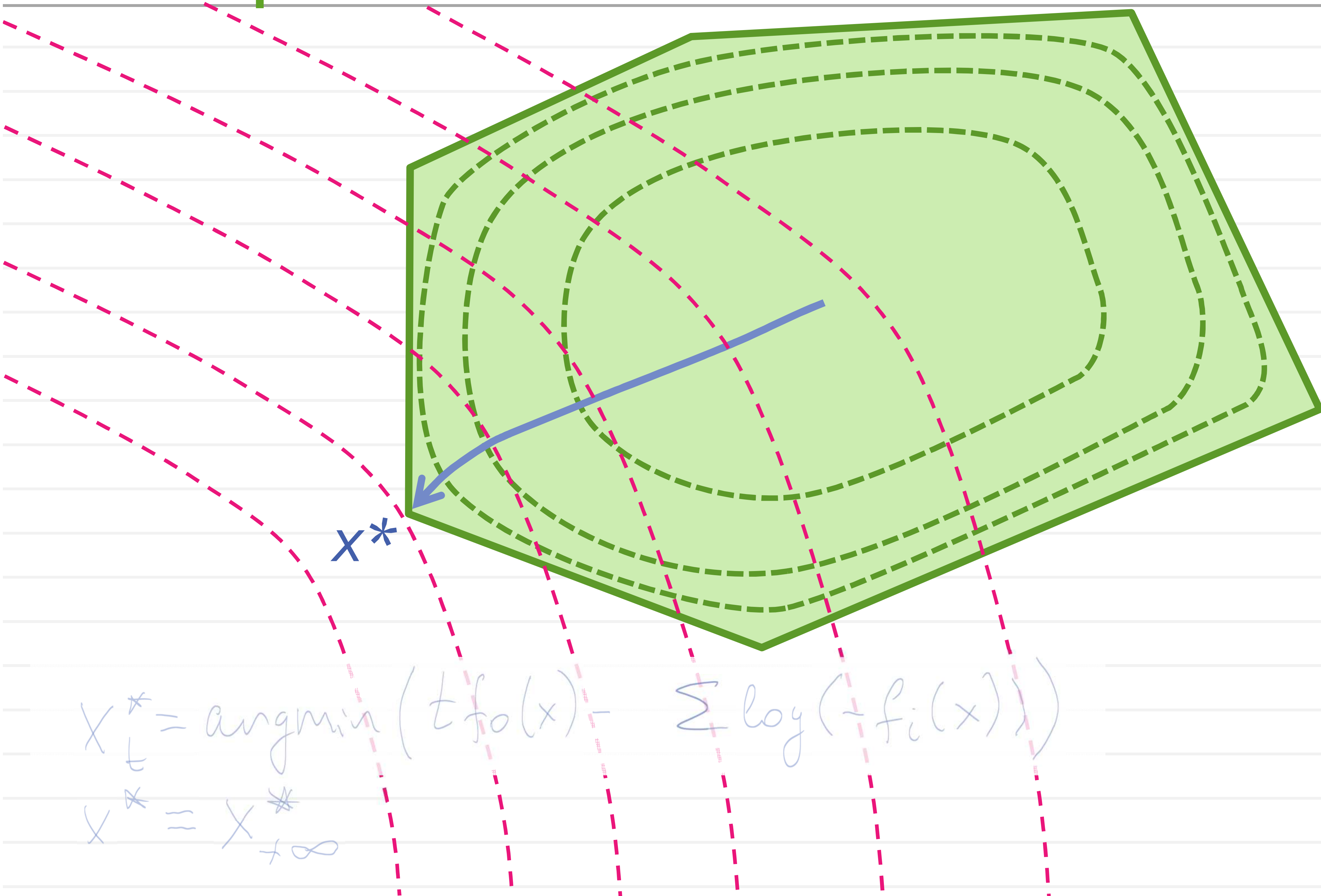
Newton($t f_0(x) - \sum \log(-f_i(x))$)

Increase t := μt

End

- Key performance factor: Newton method is very efficient with good initialization.

Central path



$x_t^* = \arg\min \left(t f_0(x) - \sum \log(-f_i(x)) \right)$

$x^* = x_{+\infty}^*$

Central path and dual variables

$$\frac{d}{dx}\left(t f_0(x) - \sum \log(-f_i(x))\right) =$$

$$= t \nabla f_0(x) + \sum \frac{1}{-f_i(x)} \nabla f_i(x) \Big|_{x_t^*} = 0$$

$$\lambda_{t,i}^* = \frac{1}{-t f_i(x_t^*)}$$

Conditions satisfied by $(x^*_t \ \lambda^*_t)$

$$\begin{cases} \nabla f_0(x_t^*) + \sum \lambda_{t,i}^* \nabla f_i(x_t^*) = 0 \\ f_i(x_t^*) \le 0 \\ \lambda_t^* \ge 0 \\ \lambda_{t,i}^* f_i(x_t^*) = -\frac{1}{t} \end{cases}$$

Slackness and duality gap

Assume that ($\tilde{x}$, $\tilde{\lambda}$, $\tilde{\nu}$) meet all KKT except complementary slackness

$$f_0(\tilde{x}) - g(\tilde{\lambda}, \tilde{\nu}) = f_0(\tilde{x}) - \min_x L(x, \tilde{\lambda}, \tilde{\nu}) =$$

$$= f_0(\tilde{x}) - L(\tilde{x}, \tilde{\lambda}, \tilde{\nu}) = -\sum \tilde{\lambda}_i f_i(\tilde{x}) - \sum \tilde{\nu}_i h_i(\tilde{x}) =$$

$$= -\sum \tilde{\lambda}_i f_i(\tilde{x})$$

Thus, the total slackness gives the duality gap.

Central path and duality gap

$$\lambda^*_{t,i} = \frac{1}{-t f_i(x^*_t)}$$

Conditions satisfied by $(x^*_t\ \ \lambda^*_t)$

$$\begin{cases} \nabla f_0(x^*_t) + \sum \lambda^*_{t,i} \nabla f_i(x^*_t) = 0 \\ f_i(x^*_t) \leq 0 \\ \lambda^*_t \geq 0 \\ \lambda^*_{t,i} f_i(x^*_t) = -\frac{1}{t} \end{cases}$$

Total duality gap: $f_0(x^*_t) - g(\lambda^*_t) = \frac{M}{t}$

$$g(\lambda^*_t) \leq d^* \leq p^* \leq f_0(x^*_t)$$

Barrier method: termination

Input: strictly feasible point

t = 1

Loop while $f_0(x_t^*) - g(\lambda_t^*) > \varepsilon$

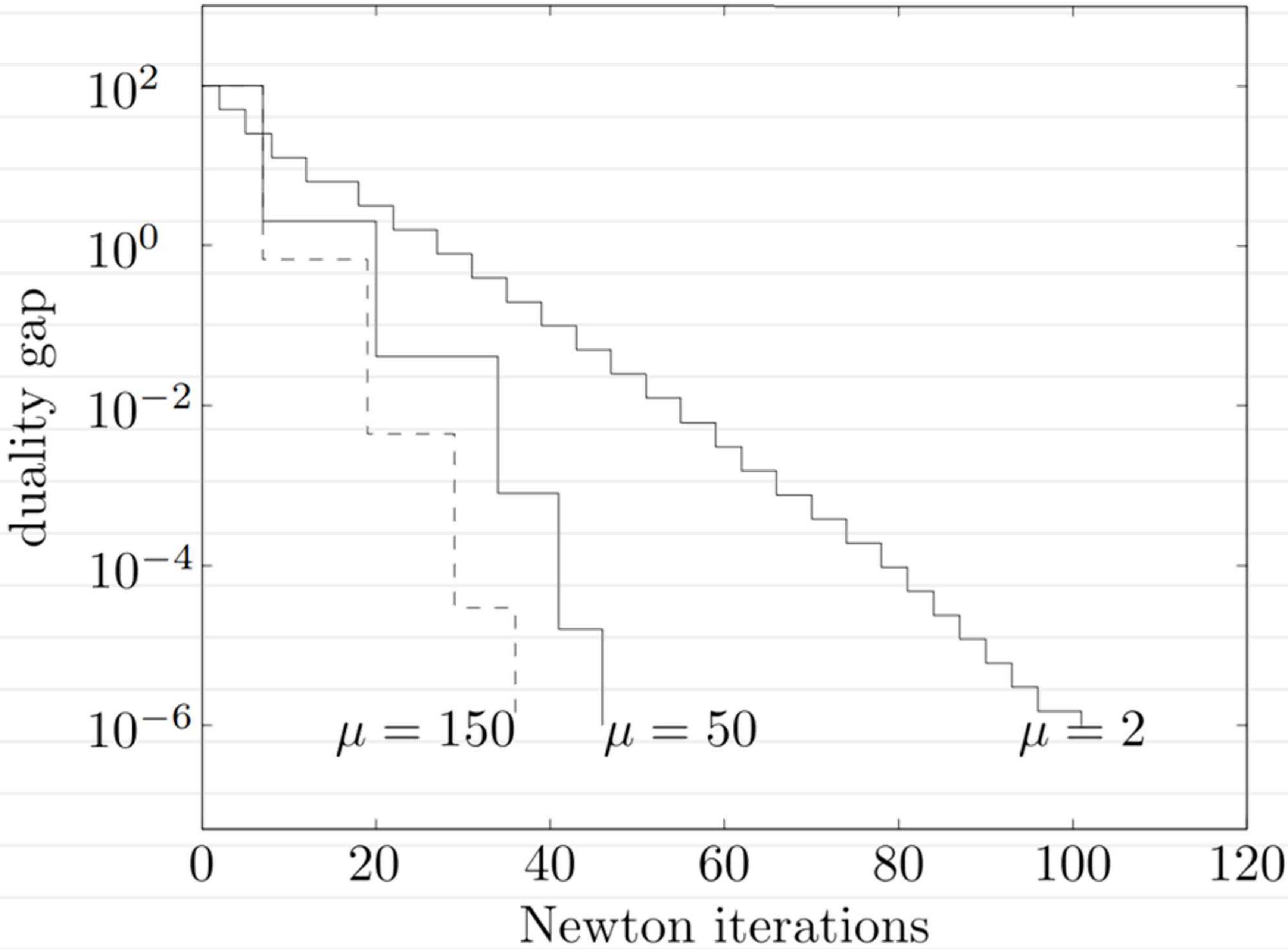
Newton($t f_0(x) - \sum \log(-f_i(x))$)

Increase t := μt

End

- Main loop terminates when duality gap is small enough
- Newton iterations can be terminated when the gradient is close enough (compared to the duality gap) to zero

Example: small ineq LP (N = 50, M = 100)



Phase 1

$$\min \; f_0(x)$$

$$\text{s.t.:} \quad f_i(x) \le 0 \quad i = 1..M$$

$$h_i(x) = 0 \quad i = 1..N$$

We need a strictly feasible point to start

Phase-1 program:

$$\min_{x,s} \; s$$

$$\text{s.t.:} \quad f_i(x) \le s \qquad s \ge 0$$

$$h_i(x) \le s$$

$$-h_i(x) \le s$$

Adding equality constraints

$$\min f_0(x)$$
$$\text{s.t. } f_i(x) \le 0$$
$$Cx = d$$

Input: feasible point (relative interior)

t = 1

Loop

Constrained Newton

$\left(t f_0(x) - \sum \log(-f_i(x)),\ Cx = d \right)$

Increase t := μt

End

Equality-constrained Newton method

$$\min \quad f_0(x)$$
$$\text{s.t.} \quad Cx = d$$

Constrained Newton steps:

$$\min_{\Delta x} \ \frac{1}{2}(x_t + \Delta x)^T \nabla^2 f_0(x_t)(x_t + \Delta x) + \nabla f_0(x_t)^T (x_t + \Delta x)$$
$$\text{s.t.}: \ C(x_t + \Delta x) = d$$

Recap: equality-constrained QP

$$\min \frac{1}{2} x^T Q x + p^T x \qquad x \in \mathbb{R}^m \qquad Q \text{ s.p.d.}$$

$$\text{s.t.:} \quad Cx = d \qquad C \in \mathbb{R}^{n \times m}$$

$$L(x, \nu) = \frac{1}{2} x^T Q x + p^T x + \nu^T (Cx - d)$$

$$\text{KKT:} \quad \begin{cases} Qx + p + C^T \nu = \vec{0} \\ Cx = d \end{cases}$$

$$\left(\begin{array}{c|c} Q & C^T \\ \hline C & 0 \end{array}\right) \begin{pmatrix} x \\ \nu \end{pmatrix} = \begin{pmatrix} -p \\ d \end{pmatrix}$$

m+n equations on m+n variables

Updated central path conditions

$$\min\ t f_0(x) - \sum \log(-f_i(x))$$
$$\text{s.t.: } Cx = d$$

from the constrained Newton

$$\lambda^*_{t,i} = \frac{1}{-t f_i(x^*_t)} \qquad \nu^*_{t,i} = \frac{\nu}{t}$$

$$\begin{cases} \nabla f_0(x^*_t) + \sum \lambda^*_{t,i} \nabla f_i(x^*_t) + C^T \nu^*_{t,i} = 0 \\ f_i(x^*_t) \le 0 \qquad Cx^*_t = d \\ \lambda^*_t \ge 0 \\ \lambda^*_{t,i} f_i(x^*_t) = -\frac{1}{t} \end{cases}$$

Duality gap = total slackness = $\frac{M}{t}$

Logarithmic barrier path following method

$$\min \; f_0(x)$$
$$\text{s.t.} \quad f_i(x) \le 0$$
$$Cx = d$$

Input: feasible point (relative interior) from Phase-1

t = 1

Loop while $f(x_t^*) - g(\lambda_t^*, \nu_t^*) > \varepsilon$

Constrained Newton

$\left(t f_0(x) - \sum \log(-f_i(x)), \; Cx = d \right)$

Increase t := μt

End

Final recap for the method

$$\min f_0(x)$$
$$\text{s.t. } f_i(x) \leq 0$$
$$Cx = d$$

$$\min t f_0(x) - \sum \log(-f_i(x))$$
$$\text{s.t.: } Cx = d$$

$$\begin{cases} \nabla f_0(x_t^*) + \sum \lambda_{t,i}^* \nabla f_i(x_t^*) + C^T \nu_{t,i}^* = 0 \\ f_i(x_t^*) \leq 0 \qquad Cx_t^* = d \\ \lambda_t^* \geq 0 \\ \lambda_{t,i}^* f_i(x_t^*) = -\frac{1}{t} \end{cases}$$

Primal-Dual method idea

$$\min f_0(x) \qquad x \in \mathbb{R}^m$$
$$\text{s.t.} \quad f_i(x) \le 0 \qquad i = 1 \ldots M$$
$$Cx = d \qquad i = 1 \ldots N$$

dual residual

$$r_t(x, \lambda, \nu) = \begin{pmatrix} \nabla f_0(x) + \sum \lambda_i \nabla f_i(x) + \sum \nu_i \nabla h_i(x) \\ Cx - d \\ \operatorname{diag}(\lambda) \cdot f(x) - \mathbb{1} \cdot t \end{pmatrix}$$

primal residual

centering residual

Solving at each iteration:

$$\boxed{\begin{aligned} & r_t(x, \lambda, \nu) = 0 \\ & \text{s.t.} \quad f_i(x) \le 0 \\ & \qquad\qquad \lambda \ge 0 \end{aligned}}$$

Primal-Dual steps

$$\min f_0(x) \qquad x \in \mathbb{R}^m$$
$$\text{s.t. } f_i(x) \le 0 \qquad i = 1 \ldots M$$
$$Cx = d \qquad i = 1 \ldots N$$

Solving at each iteration:

$$\boxed{\begin{array}{l} r_t(x, \lambda, \nu) = 0 \\ \text{s.t. } f_i(x) \le 0 \\ \qquad \lambda \ge 0 \end{array}}$$

PD step:

- Get a Newton direction from $r_t(x, \lambda, \nu) = 0$
- Line-search with the cap (to stay within $f_i(x) < 0$, $\lambda > 0$)