1a

The constraints are affine and thus convex.

The domain of ϕ is $\mathbb{R}^m_{+}$, which is convex.

We have

$$\nabla\phi(y) = \begin{pmatrix} -\frac{1}{y_1} \\ \vdots \\ -\frac{1}{y_m} \end{pmatrix}$$

and

$$\nabla^2\phi(y) = \begin{pmatrix} \frac{1}{y_1^2} & & 0 \\ & \ddots & \\ 0 & & \frac{1}{y_m^2} \end{pmatrix} \succeq 0$$

So, ϕ is convex by the second-order condition for convexity.

Hence, the problem is convex.

1b $\nabla\phi(y) + r = 0 \Rightarrow y_i = \frac{1}{r_i}$

Lagrangian:

$$L(x, y, r) = \phi(y) + r^T(y - b + Ax)$$

Lagrange dual function:

$$g(r) = \inf_{x,y} L(x, y, r)$$

$$= \begin{cases} -\sum_i \log\left(\frac{1}{r_i}\right) + m - r^T b & \text{if } A^T r = 0,\ r \geq 0 \\ -\infty & \text{otherwise} \end{cases}$$

Dual problem:

$$\max \sum_i \log(r_i) + m - r^T b$$

s.t. $A^T r = 0$

$r \geq 0$

can remove this constraint because of the domain of the objective function

2a

The problem is convex and admits the strictly feasible solution $x = \frac{B}{2n} \mathbb{1}$, so by Slater's criterion strong duality holds.

Then optimality of x implies the existence of λ and ξ s.t. the KKT conditions hold. In particular, we get complementary slackness

$$\lambda \left(\sum_i x_i^* - B \right) = 0$$

$$\xi_i x_i = 0, \quad i = 1, \dots, n$$

and the stationarity condition

$$u_i'(x_i^*) + \lambda - \xi_i = 0, \quad i = 1, \dots, n$$

Then $\xi_i = 0$ for each i with $x_i^* \neq 0$, and thus $u_i'(x_i^*) = \lambda$ for each such i.

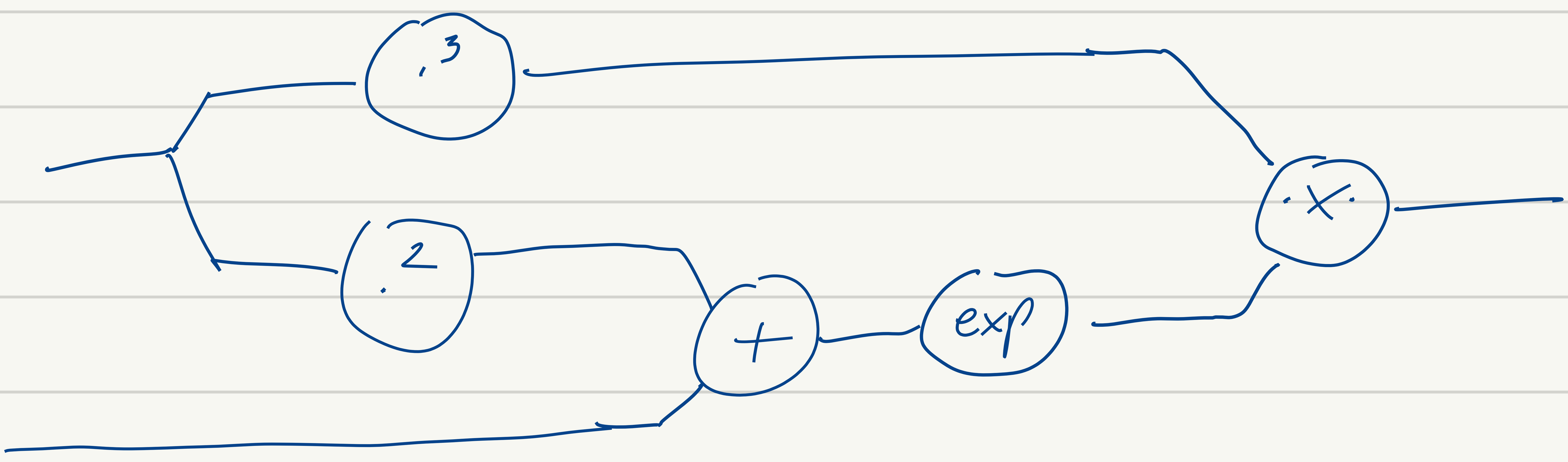
2b)

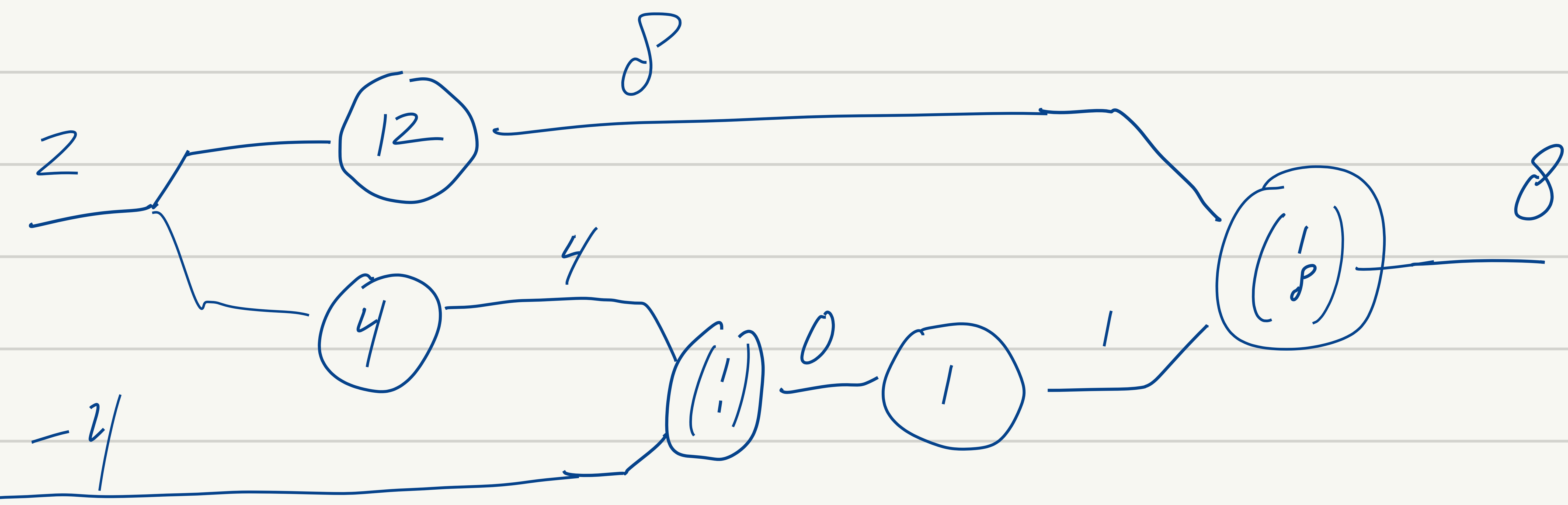
Can easily construct a solution that lies far away from the boundary.

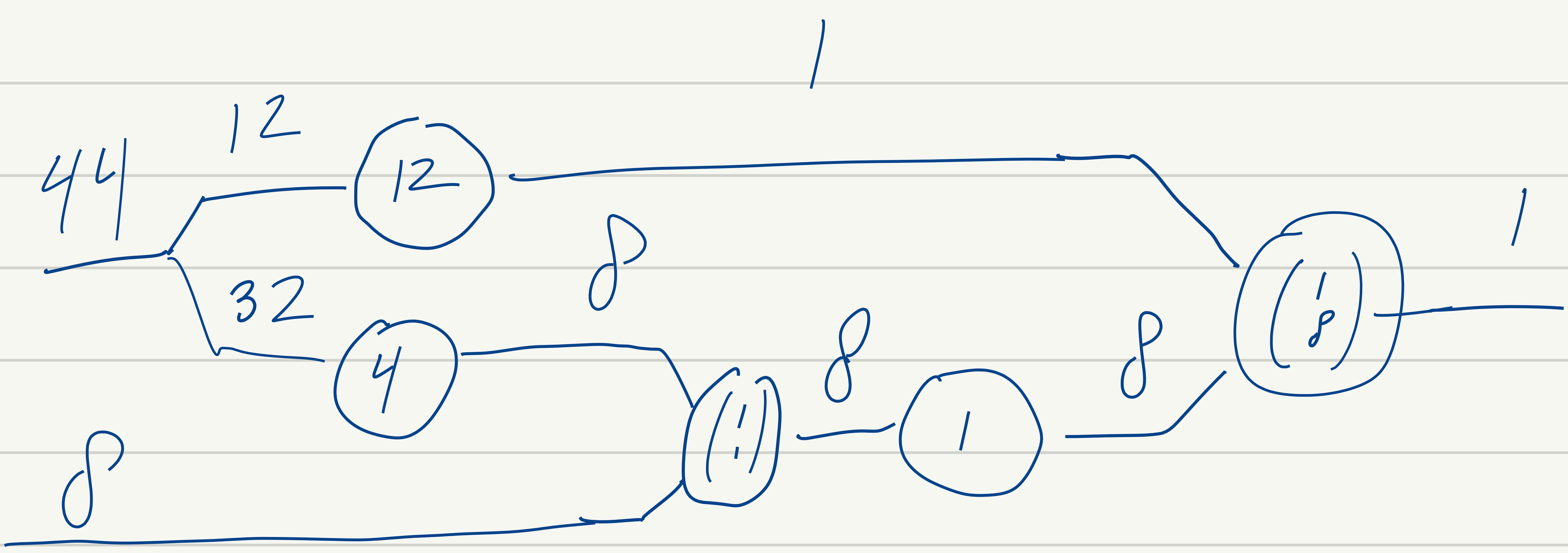
3

$$F(x_1, x_2) = x_1^3 e^{x_1^2 + x_2}$$

$$\nabla F(2, -4)$$

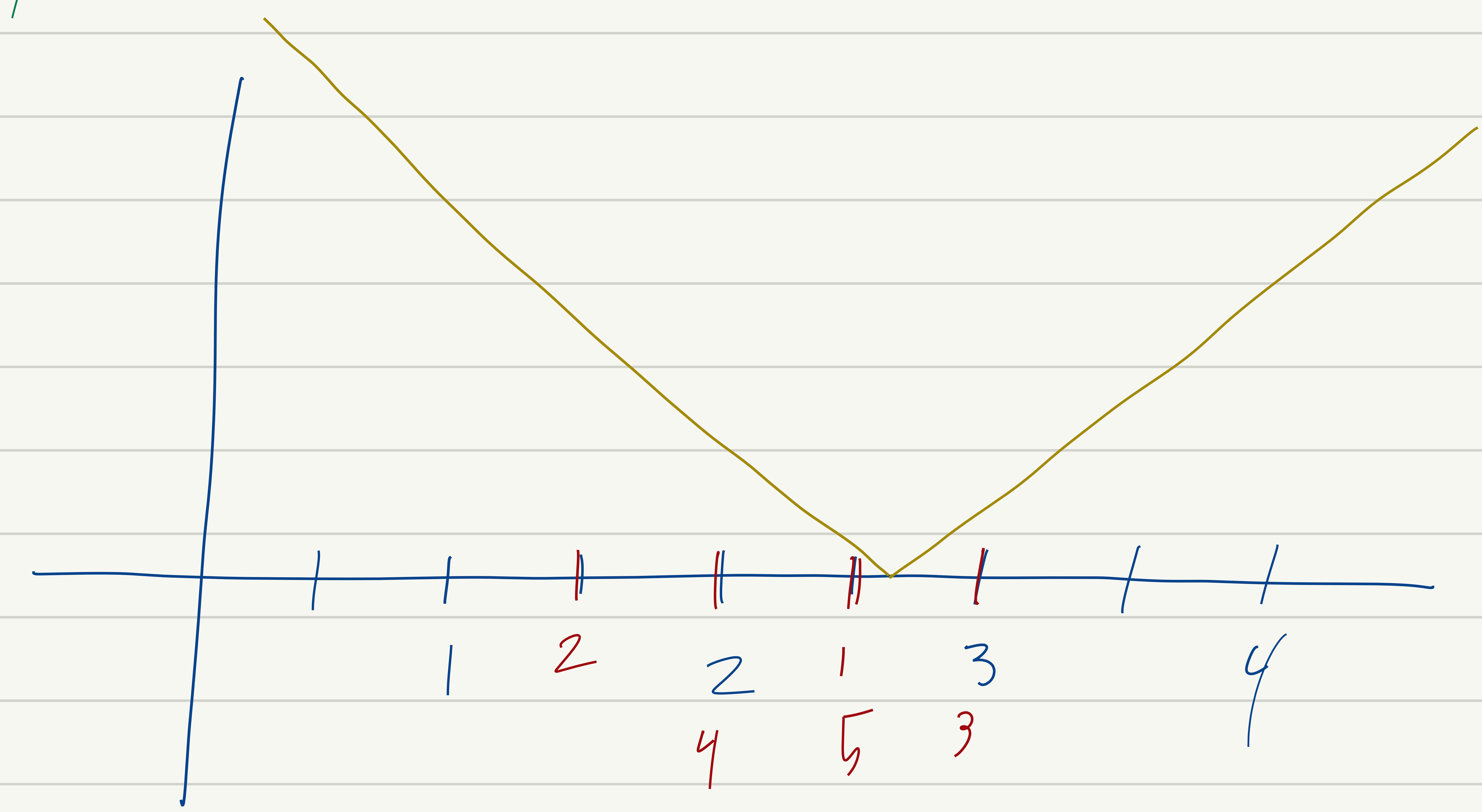






$$\nabla F(2, -4) = \begin{pmatrix} 44 \\ 8 \end{pmatrix}$$

4a



F_0 F_1 F_2 F_3 F_4 F_5

1 1 2 3 5 8

Final bracket : [2.5, 3]

[0, 4]

[1.5, 4]

[1.5, 3]

[2, 3]

[2.5+ε, 3]

4b

The interval size is $\frac{1}{\varphi^k}$ after k iterations.

$$\frac{1}{\varphi^k} \approx \frac{1}{10^{16}}$$

$$\varphi^k \approx 10^{16}$$

$$k \approx \log_\varphi 10^{16}$$

$$= 16 \log_\varphi 10$$

$$\approx 16 \cdot 5$$

$$= 80$$

Also OK:

$$\frac{-16 \ln 10}{\ln \frac{1}{\varphi}} \quad \text{or} \quad \log_{\frac{1}{\varphi}} 10^{-16}$$

5a

Since Δx is a descent direction, we have $\nabla f(x)^T \Delta x < 0$.

This means that if $t>0$ is obtained with backtracking line search, then

$$f(x+t\Delta x) \leq f(x) + \alpha \nabla f(x)^T (t\Delta x)$$
$$\leq f(x).$$

By Taylor's theorem we have

$$f(x+t\Delta x) = f(x) + \nabla f(x)^T (t\Delta x) + o(\|t\Delta x\|)$$

This means that for $\alpha \in (0,1)$ we have

$$f(x+t\Delta x) \leq f(x) + \alpha \nabla f(x)^T (t\Delta x)$$

for $t>0$ small enough

5b

Let $x^+ = x + t\Delta x$, where t is computed with backtracking line search. Then

$$f(x^+) \leq f(x) - \alpha \|\nabla f(x)\|_2^2$$

or

$$f(x^+) \leq f(x) - \frac{\alpha\beta}{M} \|\nabla f(x)\|_2^2.$$

Hence

$$f(x^+) - p^* \leq f(x) - p^* - \min\{\alpha, \frac{\alpha\beta}{M}\} \|\nabla f(x)\|_2^2$$

Since

$$f(y) \geq f(x) - \frac{1}{2m} \|\nabla f(x)\|_2^2$$

for all y, we have

$$f(x) - p^* \leq \frac{1}{2m} \|\nabla f(x)\|_2^2.$$

Putting this together gives

$$f(x^+) - p^* \leq \left(1 - 2m \min\{\alpha, \frac{\alpha\beta}{M}\}\right)\left(f(x) - p^*\right)$$

This implies

$$C(1) = \limsup_{k \to \infty} \frac{|f(x_{k+1}) - p^*|}{|f(x_k) - p^*|}$$

$$\leq \frac{|f(x^+) - p^*|}{|f(x) - p^*|}$$

$$= 1 - 2m \min\{\alpha, \frac{\alpha\beta}{M}\},$$

so convergence is at least linear with rate

$$1 - \min\{2m\alpha, 2m\alpha\beta/M\}$$