Computational Physics I - Lecture 3, part 2

Adam Foster, Dorothea Golze, Patrick Rinke
Levi Keller, Yashasvi Ranawat, Ygor Morais Jaques

Aalto University
School of Science
Department of Applied Physics

Solutions of linear and non-linear equations

Linear equation: $ax + b = c$ **last week**

- sets of linear equations are very common in physics
- they can be solved with matrix algebra
- matrix algebra is one of the most important applications in computational physics

Non linear equation: $x = f(x)$ **today**

- non-linear equations are even more common than linear
- they are much harder to solve than linear equations
- numeric approaches for non-linear eqns are very important

Non linear equations - Relaxation method

non linear equation: $x = 2 - e^{-x}$

- We wish to know the value of x that solves the equation.

Non linear equations - Relaxation method

non linear equation: $x = 2 - e^{-x}$

- We wish to know the value of x that solves the equation.
- There is no known analytic solution.

Non linear equations - Relaxation method

non linear equation: $x = 2 - e^{-x}$

- We wish to know the value of x that solves the equation.
- There is no known analytic solution.
- If in doubt, try the simplest thing that comes to mind: start with a guess (e.g. $x=1$) and see what happens.

Non linear equations - Relaxation method

non linear equation: $x = 2 - e^{-x}$

- We wish to know the value of x that solves the equation.
- There is no known analytic solution.
- If in doubt, try the simplest thing that comes to mind: start with a guess (e.g. $x=1$) and see what happens.

$$x' = 2 - e^{-1} \approx 1.632$$

Non linear equations - Relaxation method

non linear equation: $x = 2 - e^{-x}$

- We wish to know the value of x that solves the equation.
- There is no known analytic solution.
- If in doubt, try the simplest thing that comes to mind: start with a guess (e.g. $x=1$) and see what happens.

$$x' = 2 - e^{-1} \approx 1.632$$

$$x'' = 2 - e^{-1.632} \approx 1.804$$

Non-linear equations - Exercise 1

Solve. **a)** $x = 2 - e^{-x}$ **b)** $x = e^{1-x^2}$

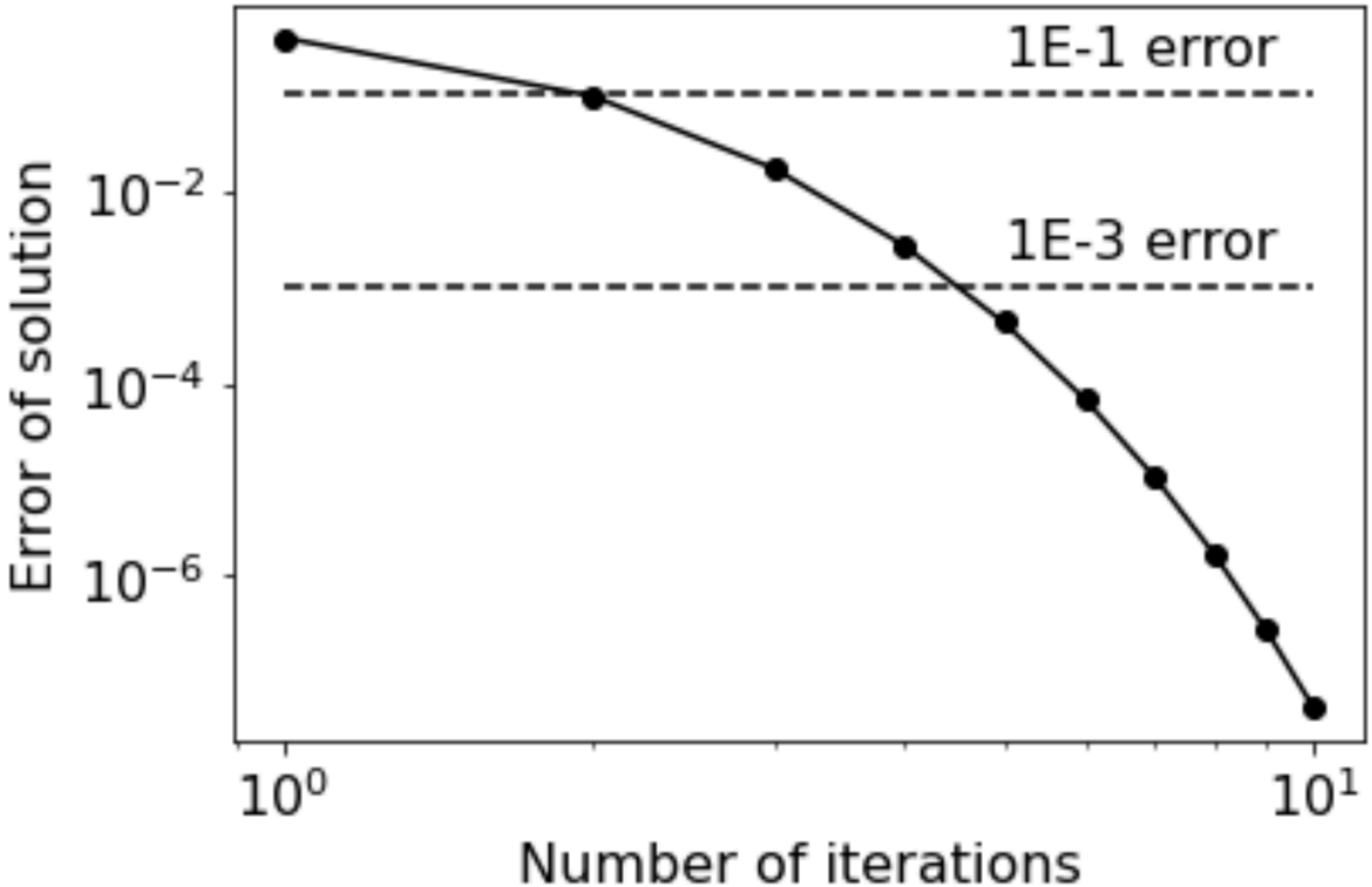
1. Write a short program that iterates equations a) and b).
2. Start iterating a) from $x=1.0$ and b) from $x=0.5$.
3. At every step, take the difference to the value at the previous step and plot this difference as a function of iteration number.

Talking points:

1. **What do you observe?**
2. **How many iterations do you need for 1E-3 accuracy?**
3. **What is happening in case b)?**

Non linear equations - Relaxation method

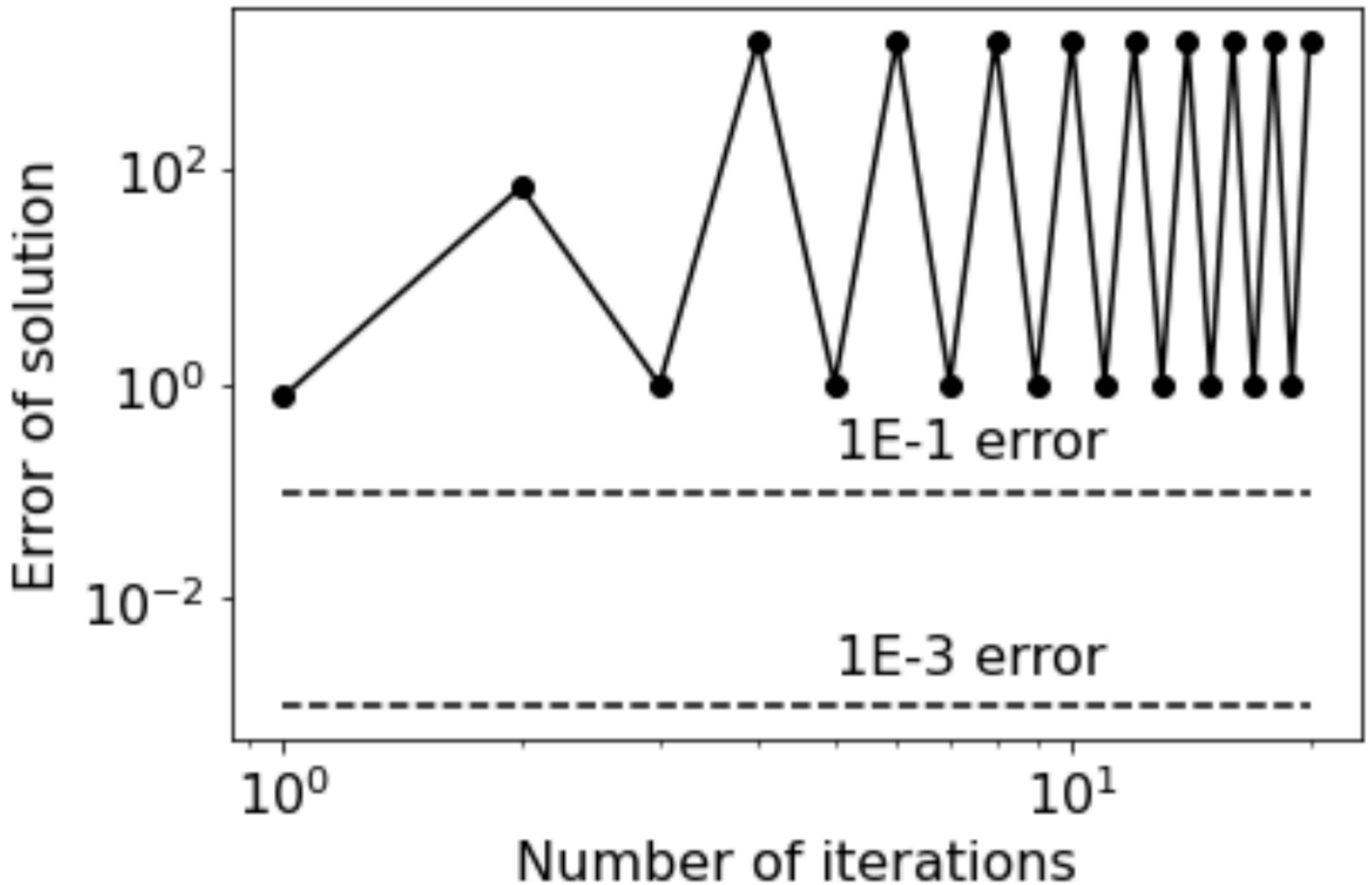
$$x = 2 - e^{-x}$$



- The solution converges to a *fixed point*.

Non linear equations - Relaxation method

$$x = e^{1-x^2}$$



- The solution does not converge to a *fixed point;* no matter where we start.

Non linear equations - error analysis

- To find out why iterating case b) did not work we assume we have an equation of type $x=f(x)$ with a solution x^* and then Taylor expand $f(x)$ around the solution x^*:

$$x' = f(x) = f(x^*) + (x - x^*)f'(x^*) + \ldots$$

Non linear equations - error analysis

- To find out why iterating case b) did not work we assume we have an equation of type $x=f(x)$ with a solution x^* and then Taylor expand $f(x)$ around the solution x^*:

$$x' = f(x) = f(x^*) + (x - x^*)f'(x^*) + \ldots$$

- Neglecting higher orders and rearranging gives:

$$x' - x^* = (x - x^*)f'(x^*)$$

distance to true solution

at every step the distance gets multiplied by the derivative at the solution

Non linear equations - error analysis

- Neglecting higher orders and rearranging gives:

$$x' - x^* = (x - x^*)f'(x^*)$$

distance to true solution

only if $|f'(x^*)|<1$ will the relaxation method converge

Non linear equations - error analysis

- Neglecting higher orders and rearranging gives:

$$x' - x^* = (x - x^*)f'(x^*)$$

distance to true solution

only if $|f'(x^*)|<1$ will the relaxation method converge

$$\textbf{a)} \quad |f'(x^*)| = \left| \left[e^{-x} \right]_{x=-1.8414} \right| = 0.1586$$ works

$$\textbf{b)} \quad |f'(x^*)| = \left| \left[-2xe^{1-x^2} \right]_{x=1} \right| = 2$$ doesn't

Non linear equations - error analysis

- Neglecting higher orders and rearranging gives:

$$x' - x^* = (x - x^*) f'(x^*)$$

distance to true solution

only if $|f'(x^*)|<1$ will the relaxation method converge

$$\textbf{a)} \quad |f'(x^*)| = \left| \left[e^{-x} \right]_{x=-1.8414} \right| = 0.1586 \quad \textbf{works}$$

$$\textbf{b)} \quad |f'(x^*)| = \left| \left[-2xe^{1-x^2} \right]_{x=1} \right| = 2 \quad \textbf{doesn't}$$

- We'll come back to better solution methods.

Non linear equations - error analysis

Key concept: iterating

Solving equations *iteratively* is a common technique in computational physics. Non linear equations are such an example.

$$\textbf{a)} \quad |f'(x^*)| = \left| \left[e^{-x} \right]_{x=-1.8414} \right| = 0.1586 \quad \textbf{works}$$

$$\textbf{b)} \quad |f'(x^*)| = \left| \left[-2xe^{1-x^2} \right]_{x=1} \right| = 2 \quad \textbf{doesn't}$$

Non linear equations - error analysis

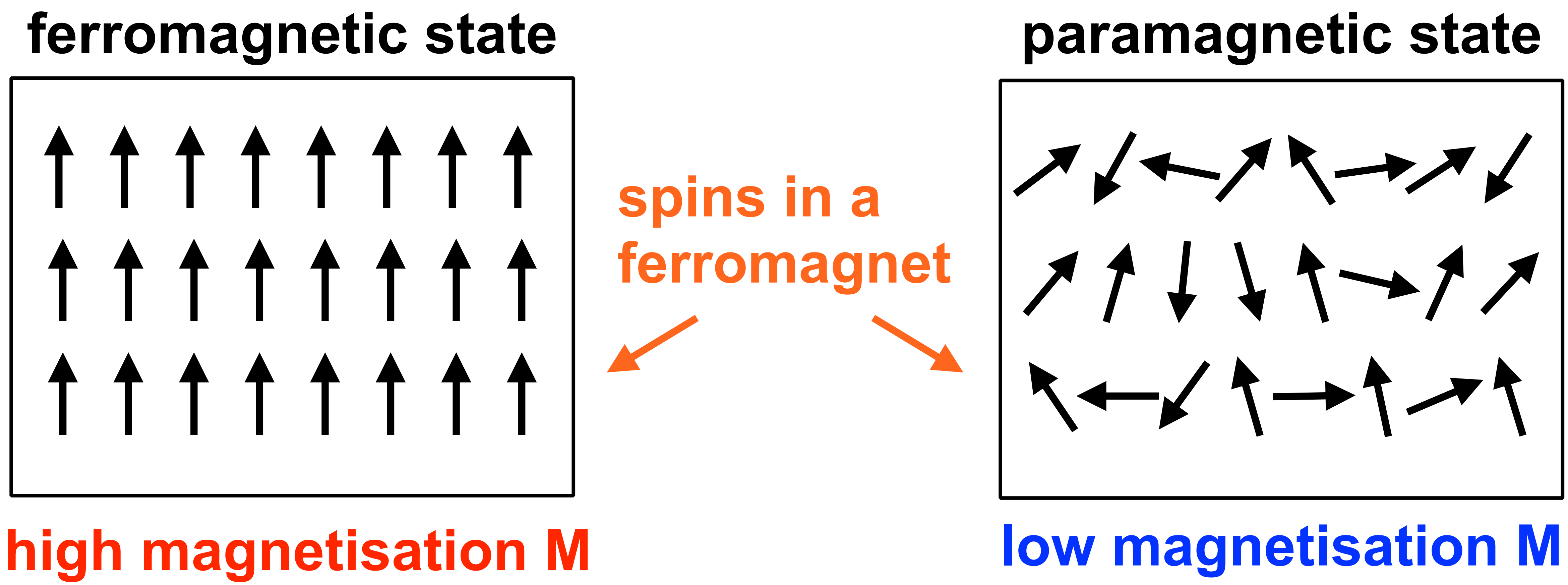
Key concept: iterating

Solving equations *iteratively* is a common technique in computational physics. Non linear equations are such an example.

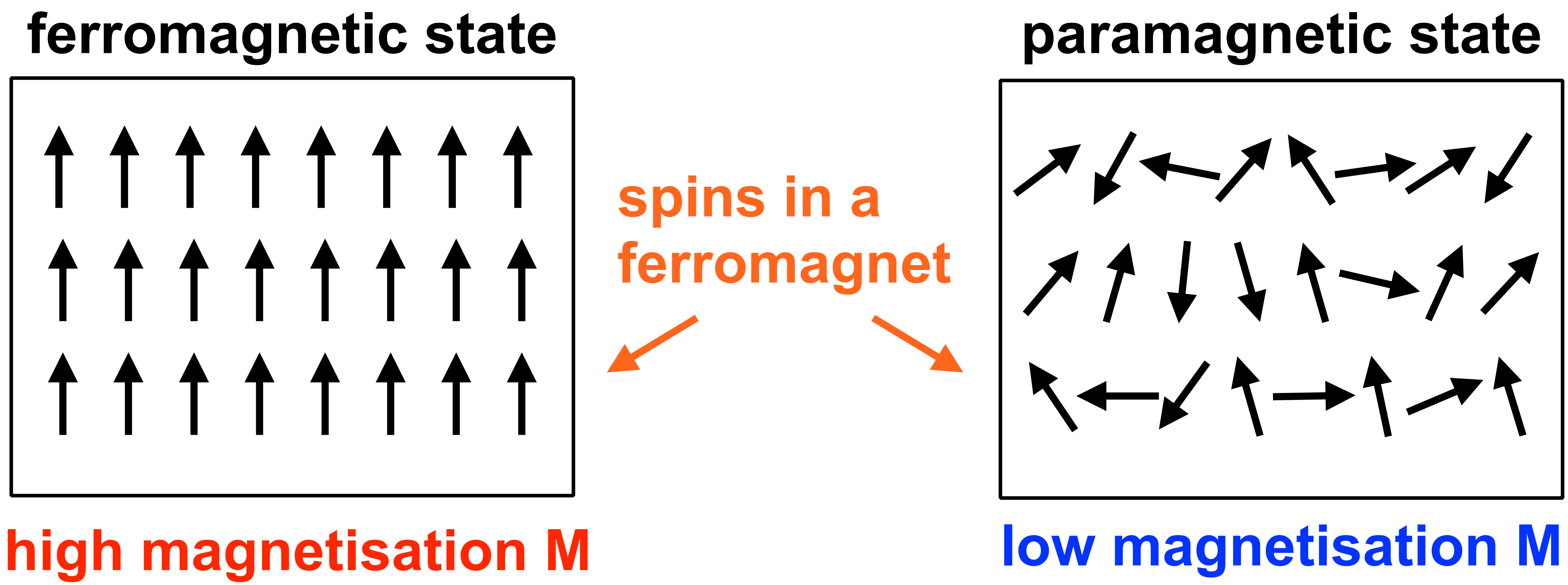
Key concept: stopping condition

Ensure that the iterative loop has a definite stopping condition. This could, for example, be the maximum number of iterations. Accuracy thresholds may not stop the iteration, if the iterative scheme does not converge.

Example - Ferromagnetism



Example - Ferromagnetism



- In the mean-field theory of ferromagnetism, the strength M of magnetization depends on temperature T according to

$$M = \mu \tanh \frac{JM}{k_B T}$$

coupling constant

magnetic moment

Example - Ferromagnetism

$$M = \mu \tanh \frac{JM}{k_B T}$$

- we make the following substitutions

$$m = M/\mu \quad \text{and} \quad C = \mu J/k_B$$

Example - Ferromagnetism

$$M = \mu \tanh \frac{JM}{k_B T}$$

- we make the following substitutions

$$m = M/\mu \quad \text{and} \quad C = \mu J/k_B$$

- to obtain

$$m = \tanh \frac{Cm}{T}$$

Example - Ferromagnetism

$$M = \mu \tanh \frac{JM}{k_B T}$$

- we make the following substitutions

$$m = M/\mu \quad \text{and} \quad C = \mu J/k_B$$

- to obtain

$$m = \tanh \frac{Cm}{T}$$

trivial solution for m=0

But we are interested in non-trivial solutions!

Example - Ferromagnetism

Example program

Relaxation method for two or more variables

non linear eqns: $x = f(x, y) \quad \text{and} \quad y = g(x, y)$

- The relaxation method can easily be applied to several variables.

Relaxation method for two or more variables

non linear eqns: $x = f(x, y) \quad \text{and} \quad y = g(x, y)$

- The relaxation method can easily be applied to several variables.
- We guess the initial values for x and y and then iterate.

Relaxation method for two or more variables

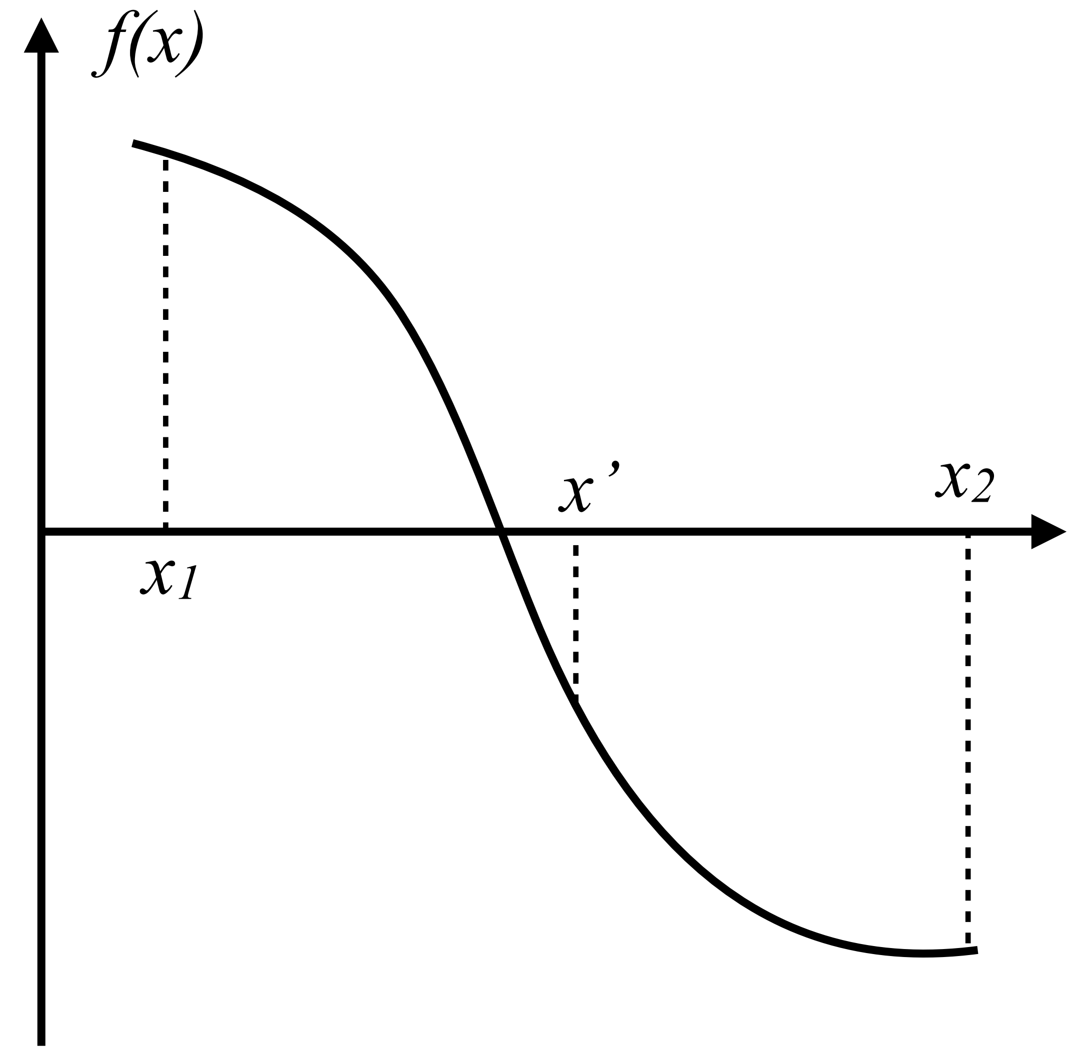
non linear eqns: $x = f(x, y) \quad \text{and} \quad y = g(x, y)$

- The relaxation method can easily be applied to several variables.
- We guess the initial values for x and y and then iterate.
- However, just like in the one-dimensional case it is not guaranteed that the solution converges.

Non linear equations - Binary search

non linear equation: $x = g(x)$

- *Binary search* (also called *bisection method*) is more robust than the relaxation method.

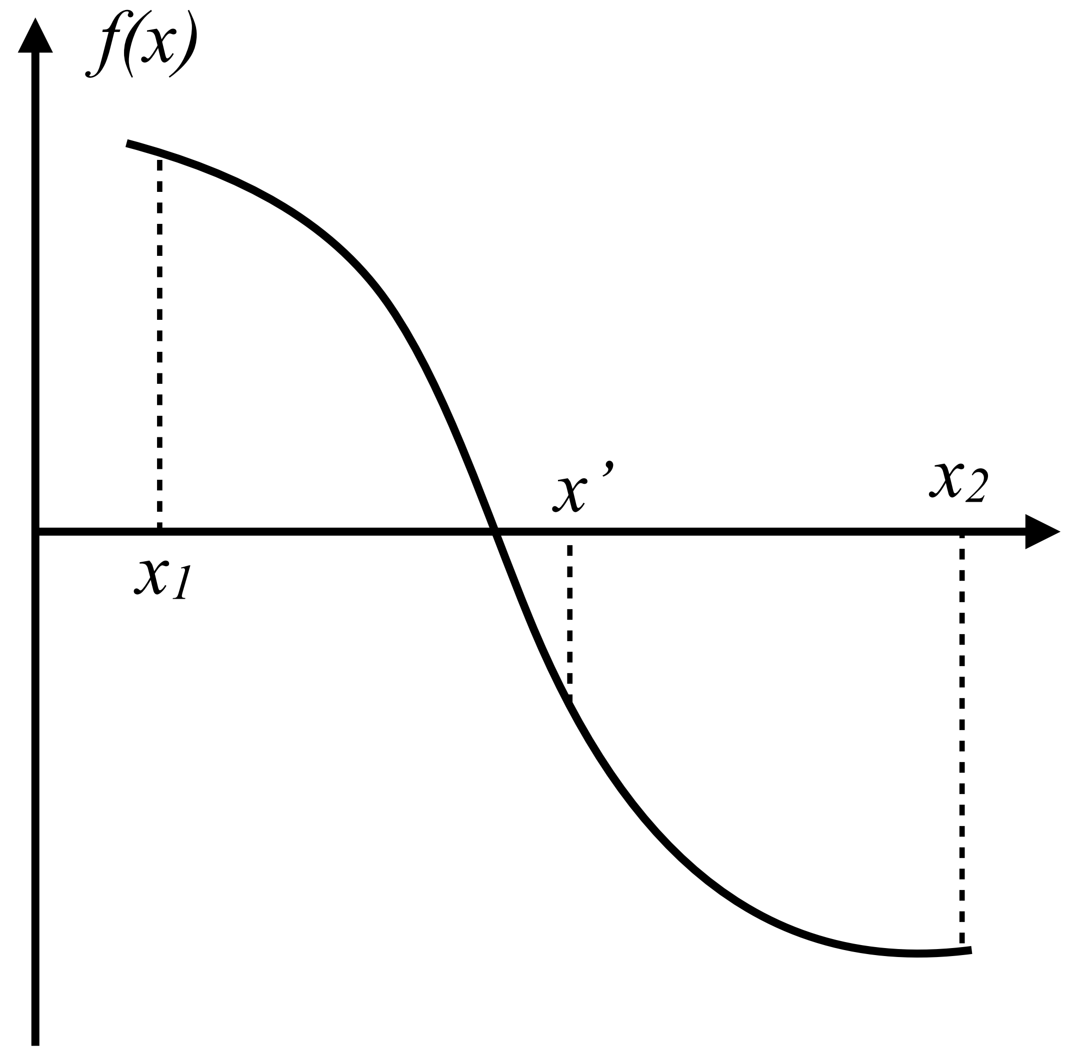


Non linear equations - Binary search

non linear equation: $x = g(x)$

- *Binary search* (also called *bisection method*) is more robust than the relaxation method.
- We recast the problem into one of finding zeros:

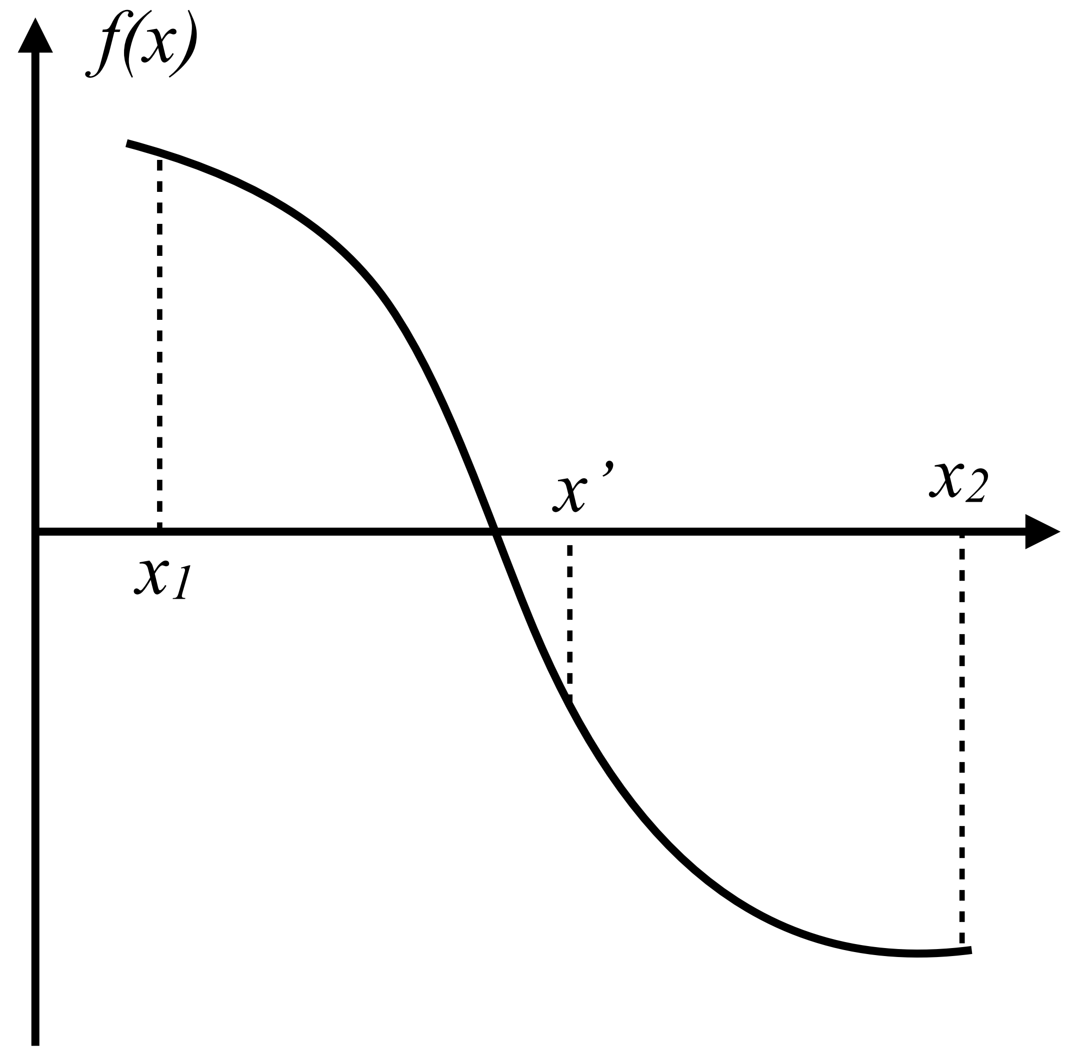
$$\underbrace{g(x) - x}_{f(x)} = 0$$



Non linear equations - Binary search

non linear equation: $f(x) = 0$

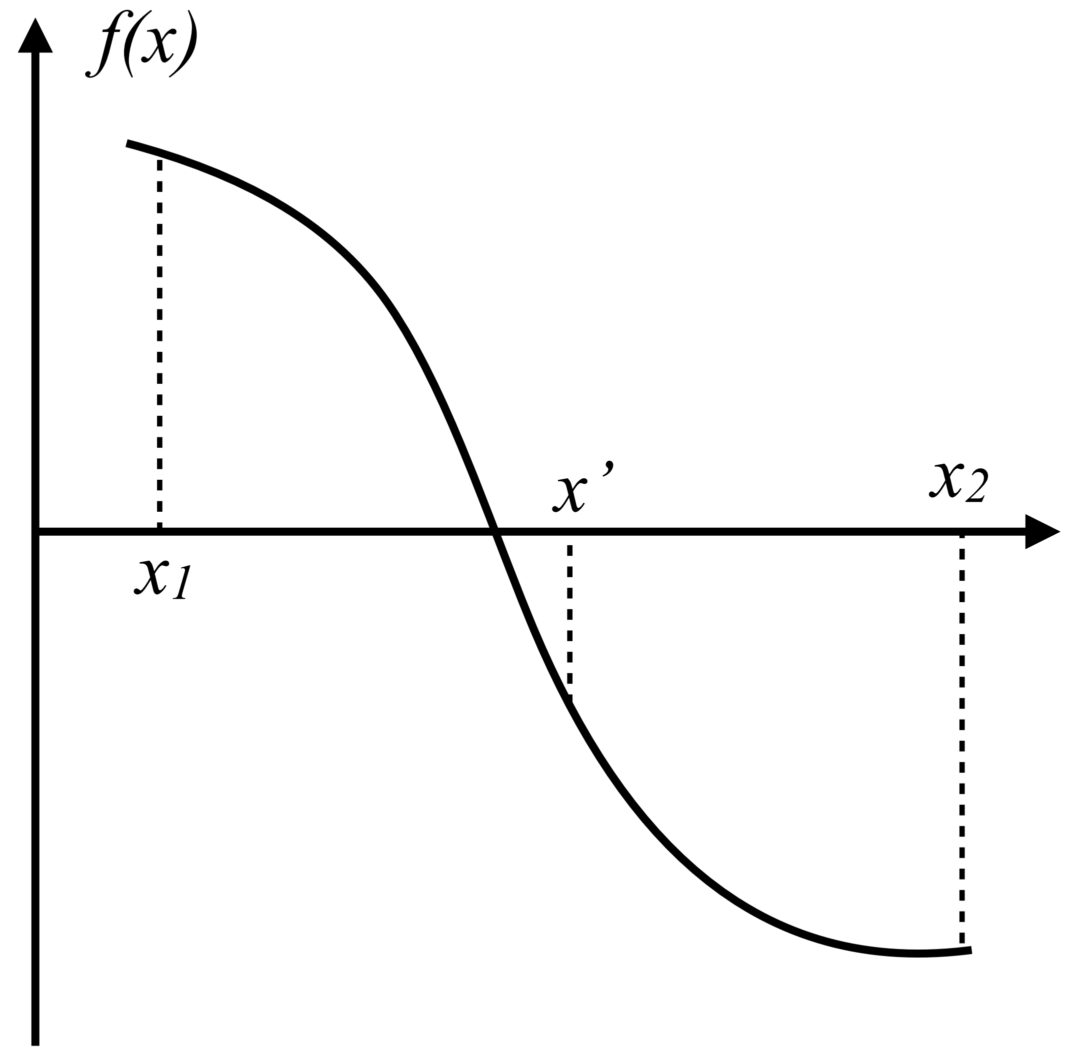
1. We start with an interval $[x_1,x_2]$.



Non linear equations - Binary search

non linear equation: $f(x) = 0$

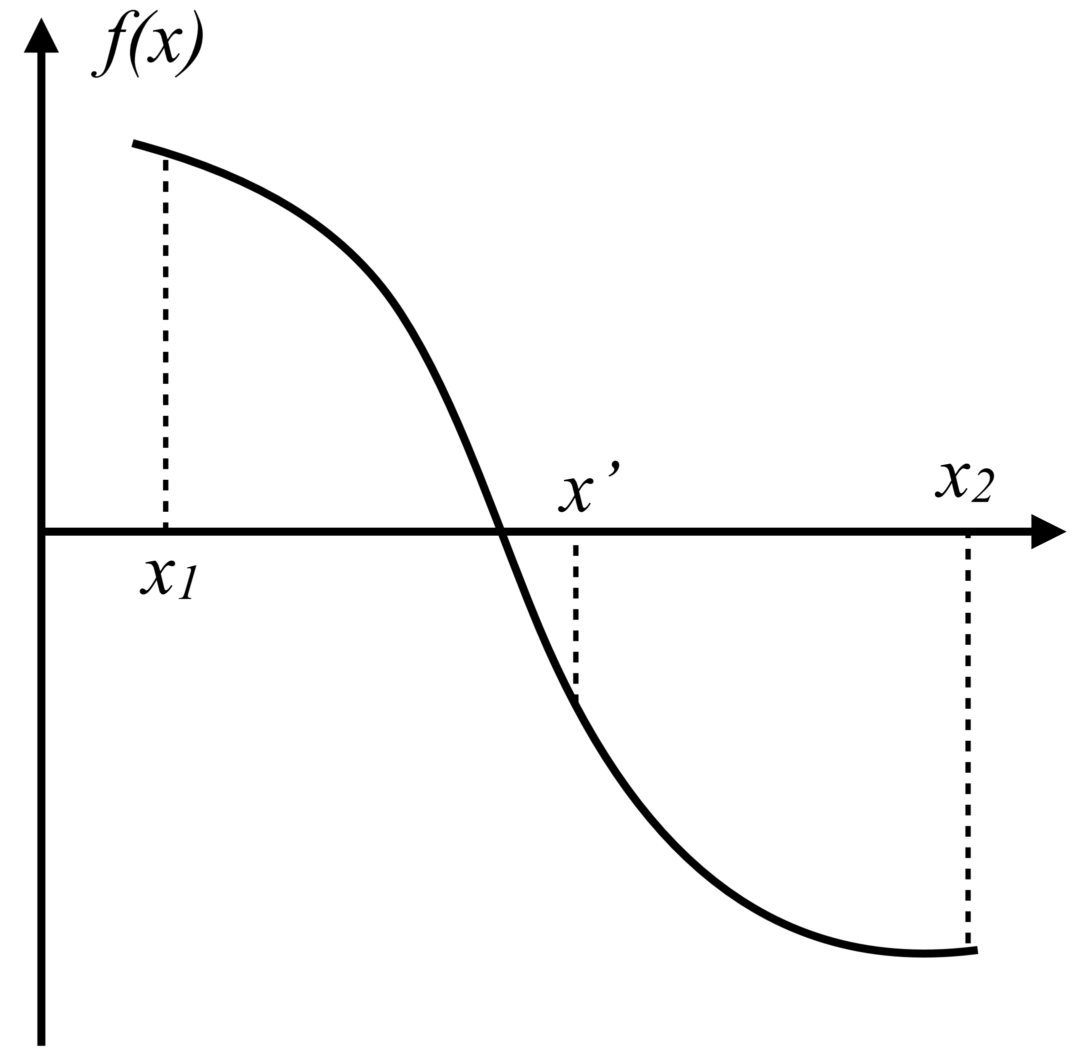
1. We start with an interval $[x_1, x_2]$.
2. Check that $f(x_1)$ and $f(x_2)$ have different signs.



Non linear equations - Binary search

non linear equation: $f(x) = 0$

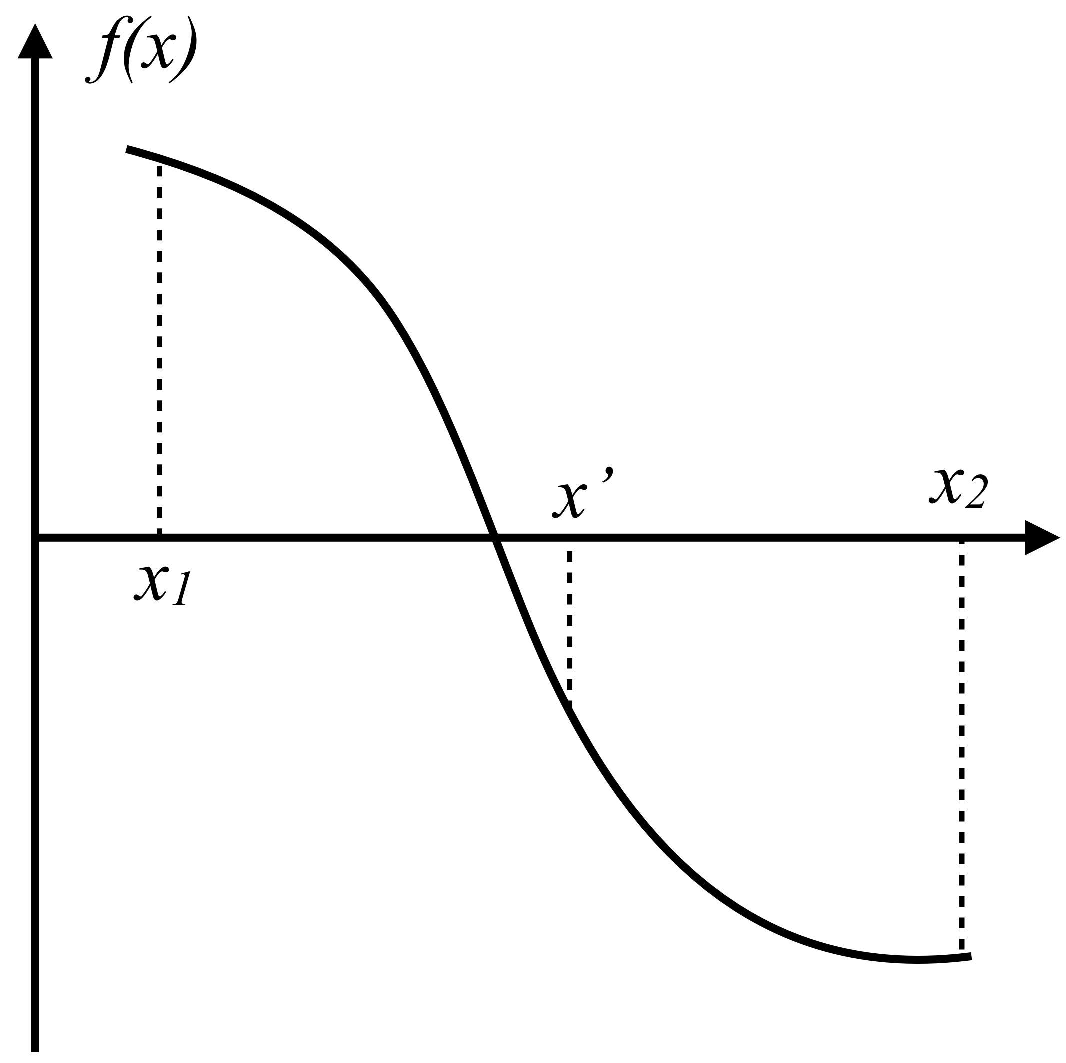
1. We start with an interval $[x_1, x_2]$.
2. Check that $f(x_1)$ and $f(x_2)$ have different signs.
3. Calculate midpoint $x' = 0.5(x_1 + x_2)$ and evaluate $f(x')$.



Non linear equations - Binary search

non linear equation: $f(x) = 0$

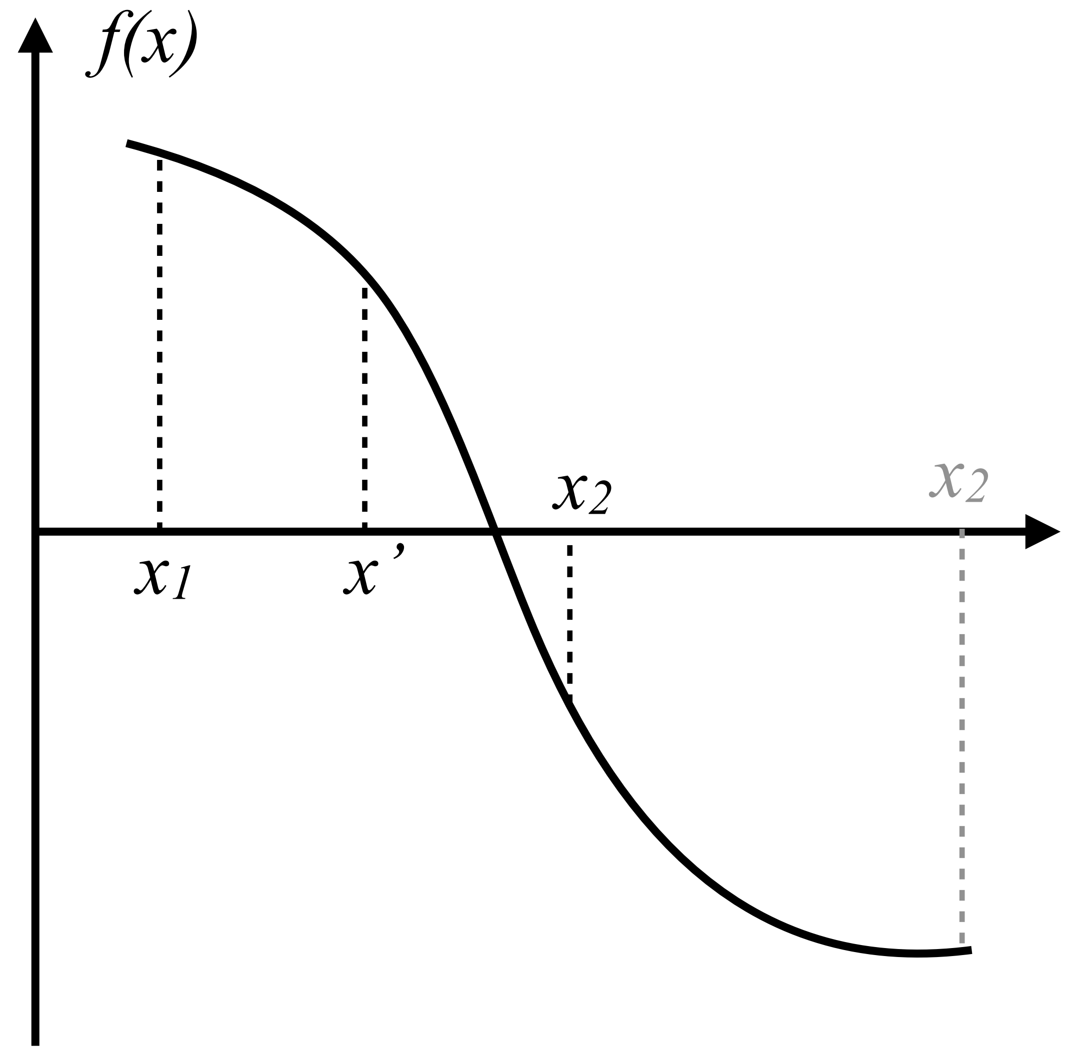
1. We start with an interval $[x_1,x_2]$.
2. Check that $f(x_1)$ and $f(x_2)$ have different signs.
3. Calculate midpoint $x' = 0.5(x_1+x_2)$ and evaluate $f(x')$.
4. If $f(x')$ has the same sign as $f(x_1)$ set $x_1=x'$, otherwise $x_2=x'$.



Non linear equations - Binary search

non linear equation: $f(x) = 0$

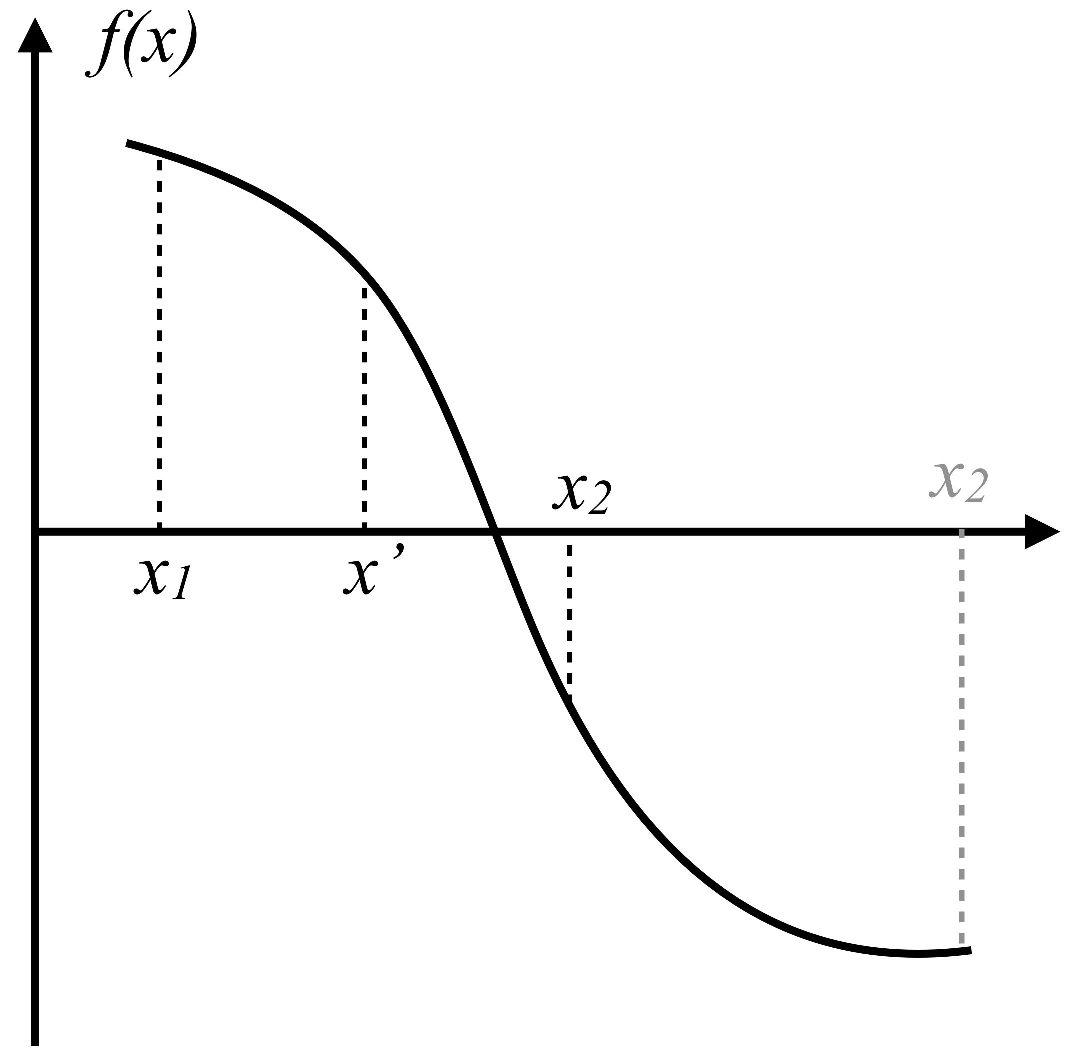
1. We start with an interval $[x_1,x_2]$.
2. Check that $f(x_1)$ and $f(x_2)$ have different signs.
3. Calculate midpoint $x' = 0.5(x_1+x_2)$ and evaluate $f(x')$.
4. If $f(x')$ has the same sign as $f(x_1)$ set $x_1=x'$, otherwise $x_2=x'$.



Non linear equations - Binary search

non linear equation: $f(x) = 0$

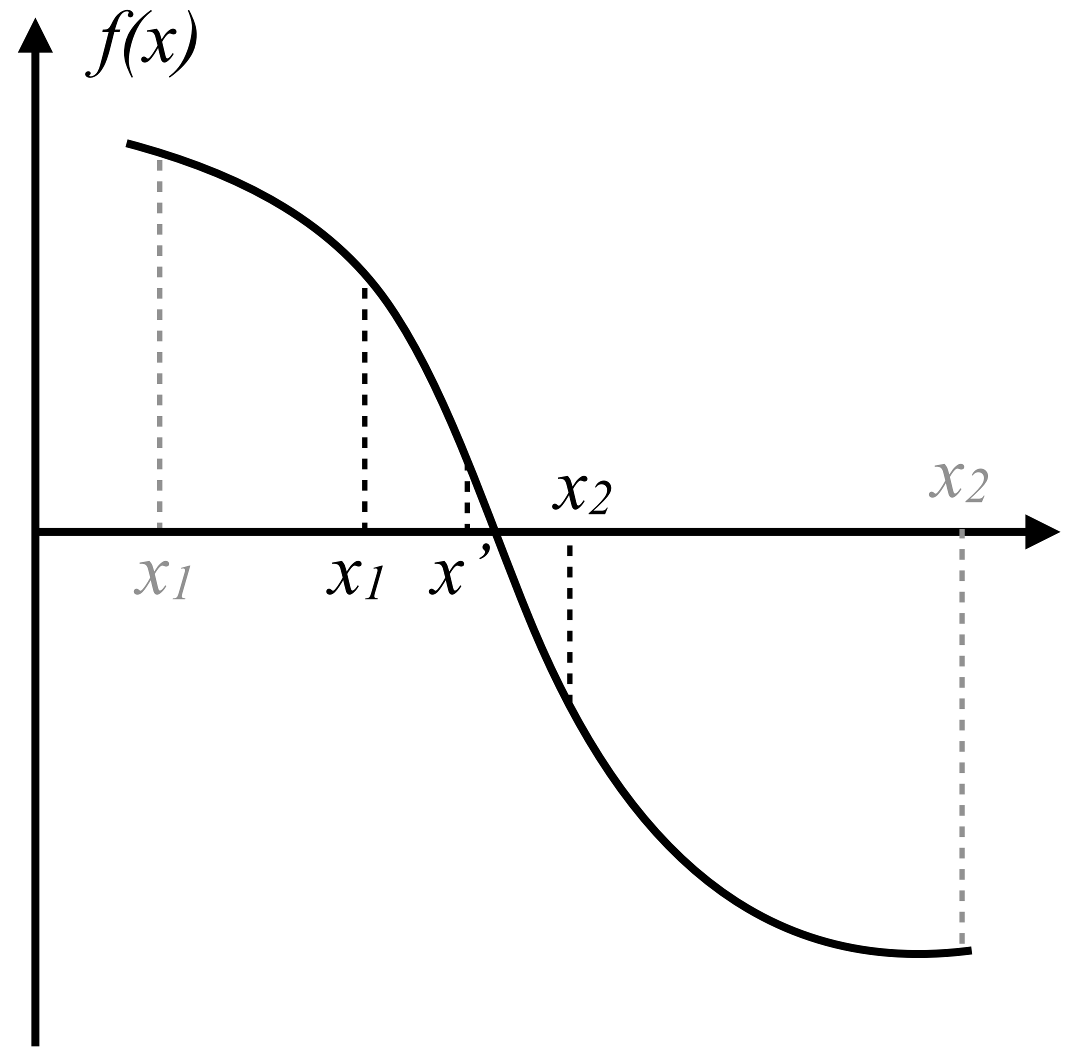
1. We start with an interval $[x_1,x_2]$.
2. Check that $f(x_1)$ and $f(x_2)$ have different signs.
3. Calculate midpoint $x' = 0.5(x_1+x_2)$ and evaluate $f(x')$.
4. If $f(x')$ has the same sign as $f(x_1)$ set $x_1=x'$, otherwise $x_2=x'$.
5. If $|x_1-x_2|$ is greater than the specified accuracy, go to 2.



Non linear equations - Binary search

non linear equation: $f(x) = 0$

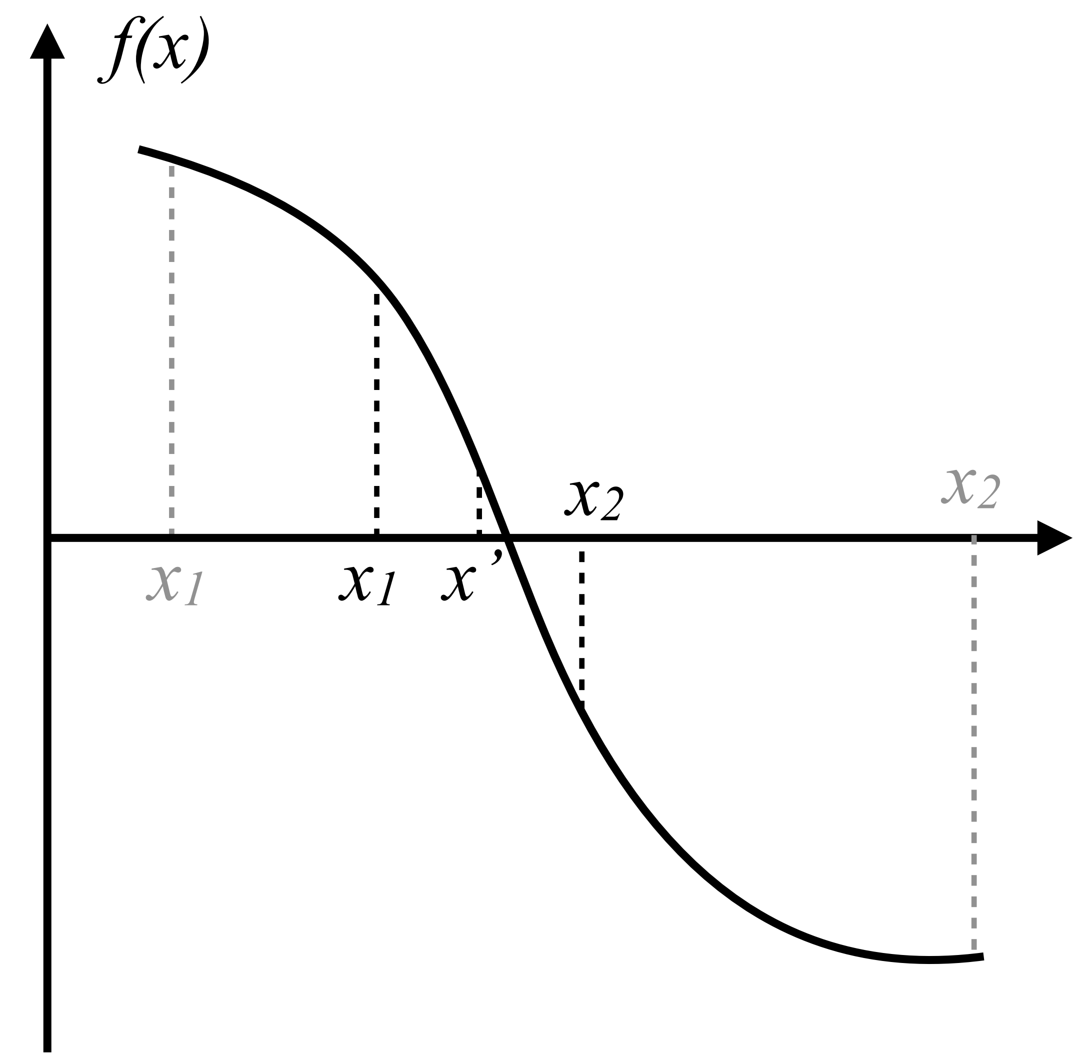
1. We start with an interval $[x_1, x_2]$.
2. Check that $f(x_1)$ and $f(x_2)$ have different signs.
3. Calculate midpoint $x' = 0.5(x_1 + x_2)$ and evaluate $f(x')$.
4. If $f(x')$ has the same sign as $f(x_1)$ set $x_1 = x'$, otherwise $x_2 = x'$.
5. If $|x_1 - x_2|$ is greater than the specified accuracy, go to 2.



Non linear equations - Binary search

non linear equation: $f(x) = 0$

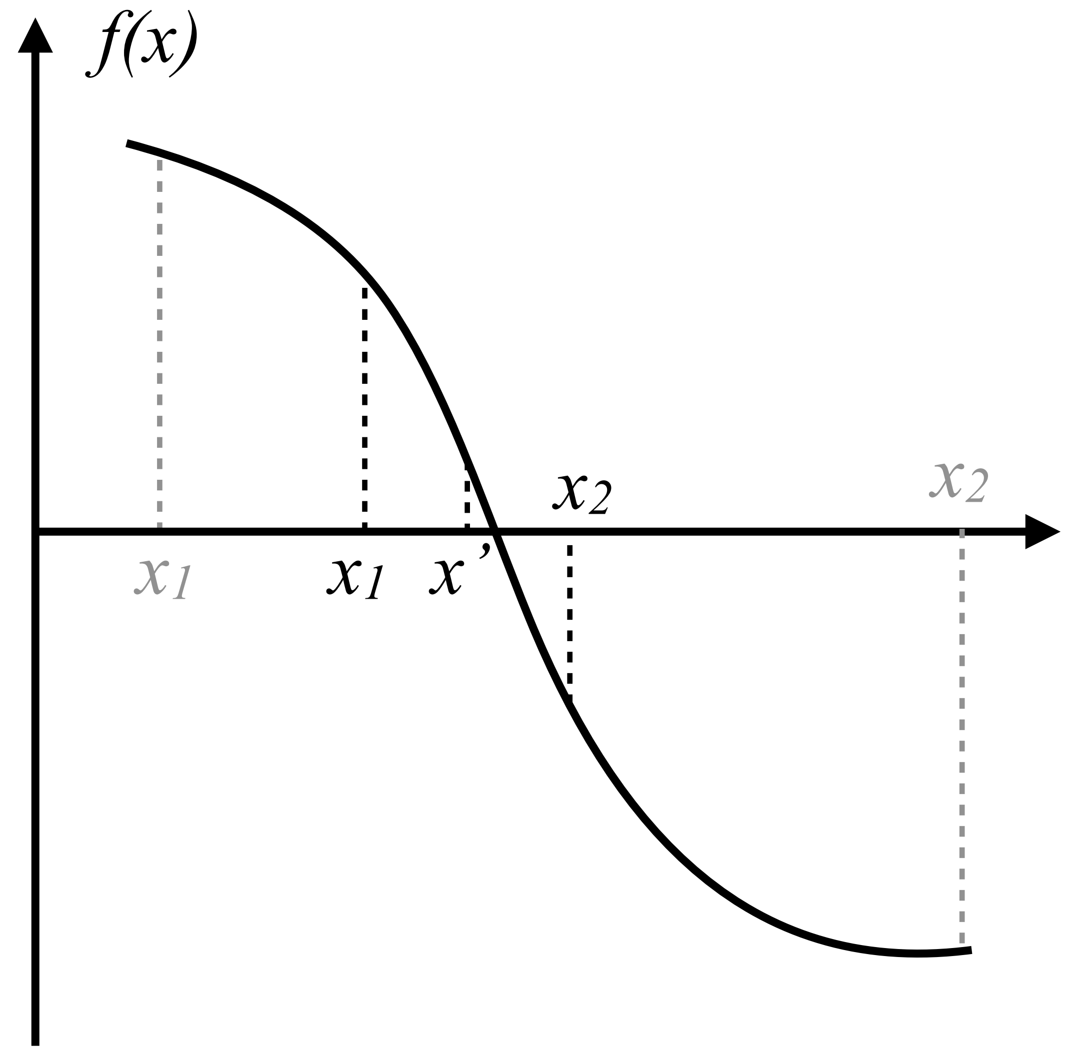
- Binary search is easy and fast.



Non linear equations - Binary search

non linear equation: $f(x) = 0$

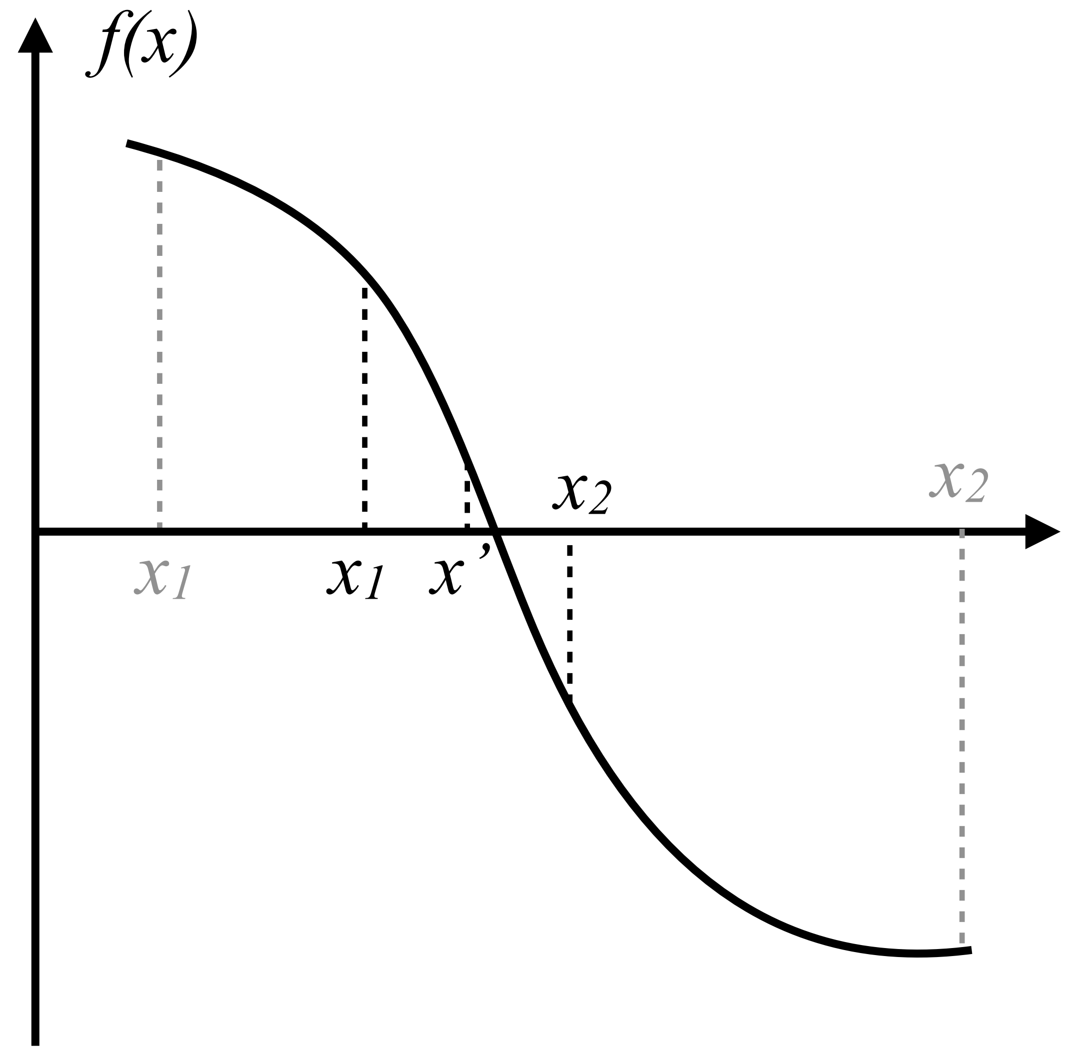
- Binary search is easy and fast.
- But it depends very strongly on the initial points x_1 and x_2.



Non linear equations - Binary search

non linear equation: $f(x) = 0$

- Binary search is easy and fast.
- But it depends very strongly on the initial points x_1 and x_2.
- It is not guaranteed to find all roots, if the function has more than one zero.



Non linear equations - Binary search

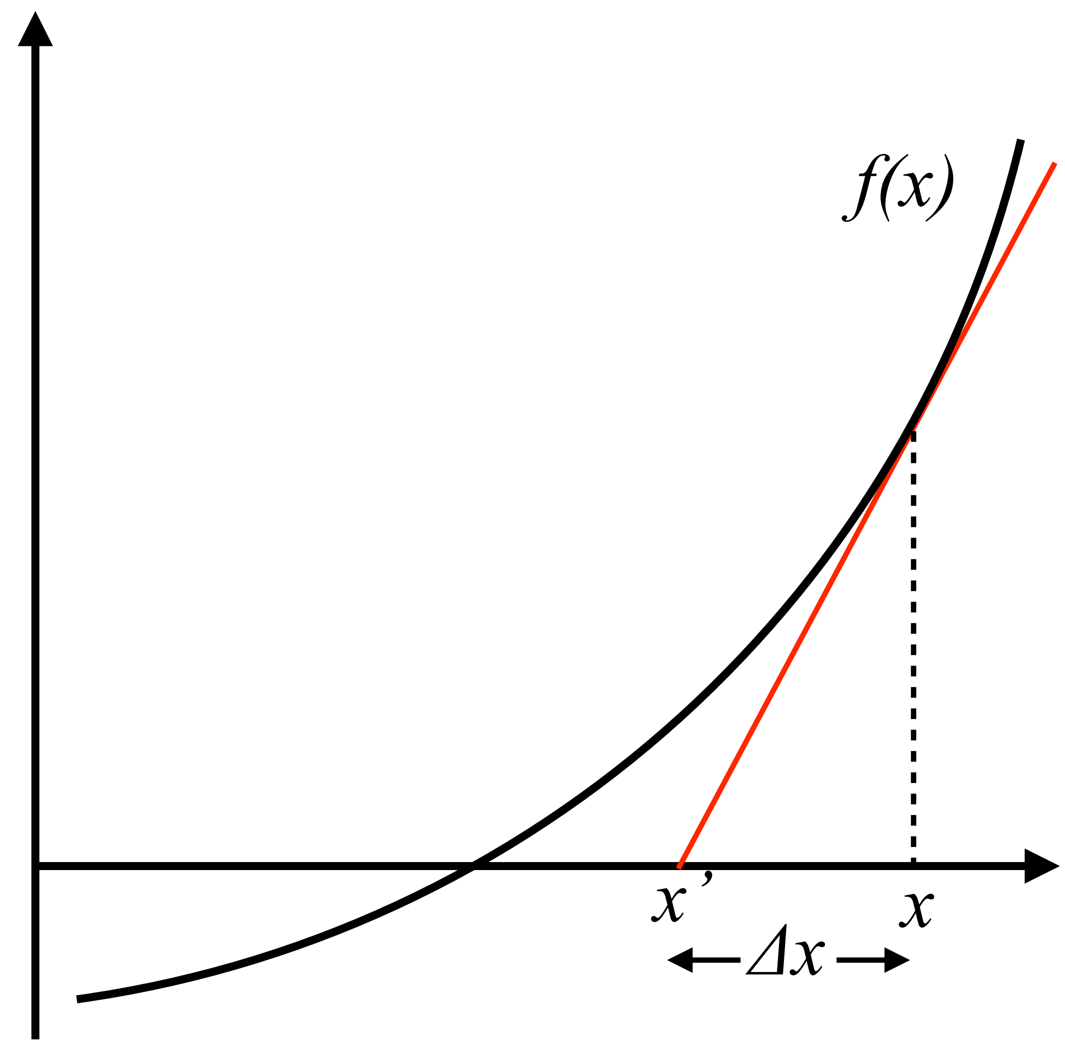
Key concept: bisection method

The bisection method is a general bracketing method. It can be used as an iterative method for non-linear equations, at which it is more robust than the relaxation method, but not without fail.

Non linear equations - Newton's method

non linear equation: $f(x) = 0$

- In Newton's method we make use of the first derivative.

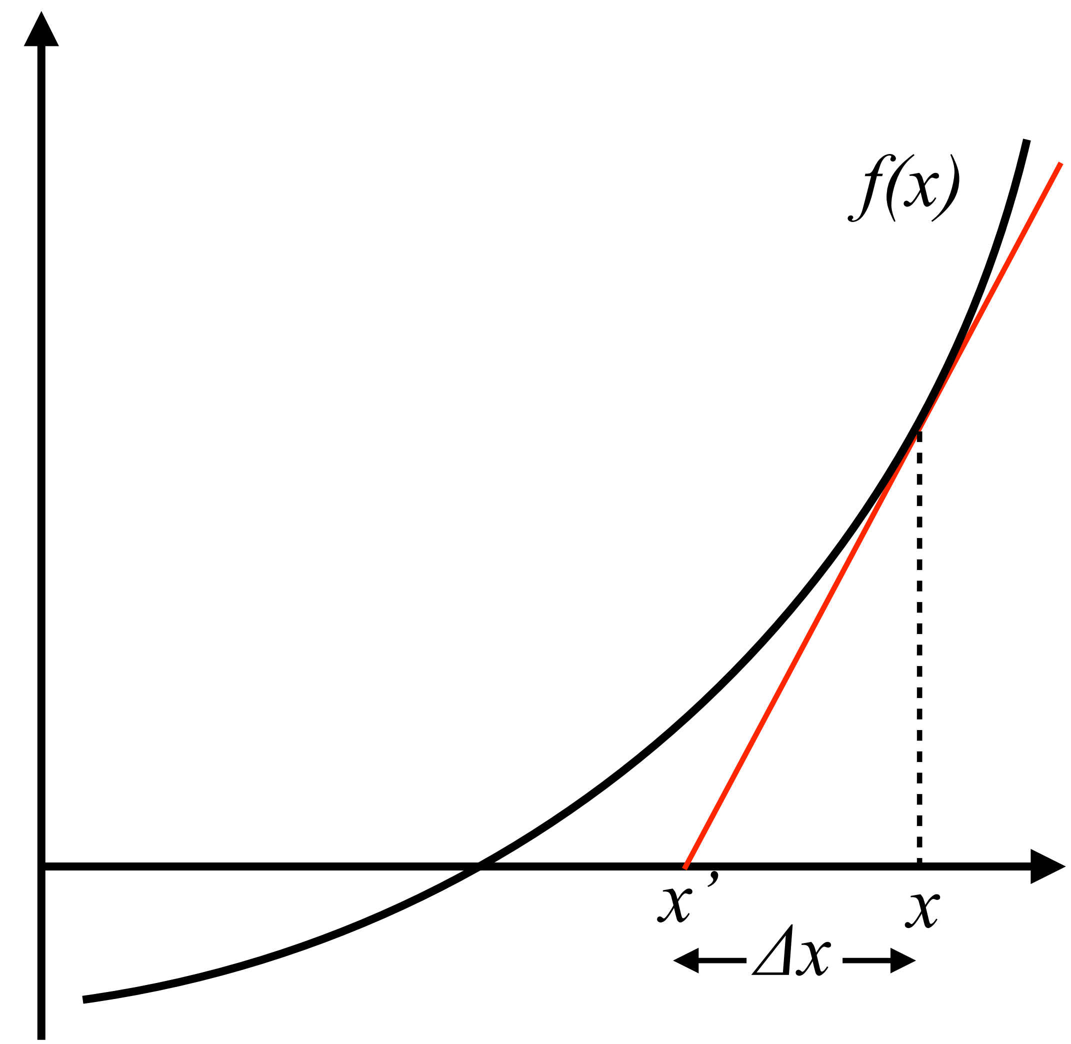


Non linear equations - Newton's method

non linear equation: $f(x) = 0$

- In Newton's method we make use of the first derivative.
- We see in the graph that:

$$f'(x) = \frac{f(x)}{\Delta x}$$



Non linear equations - Newton's method

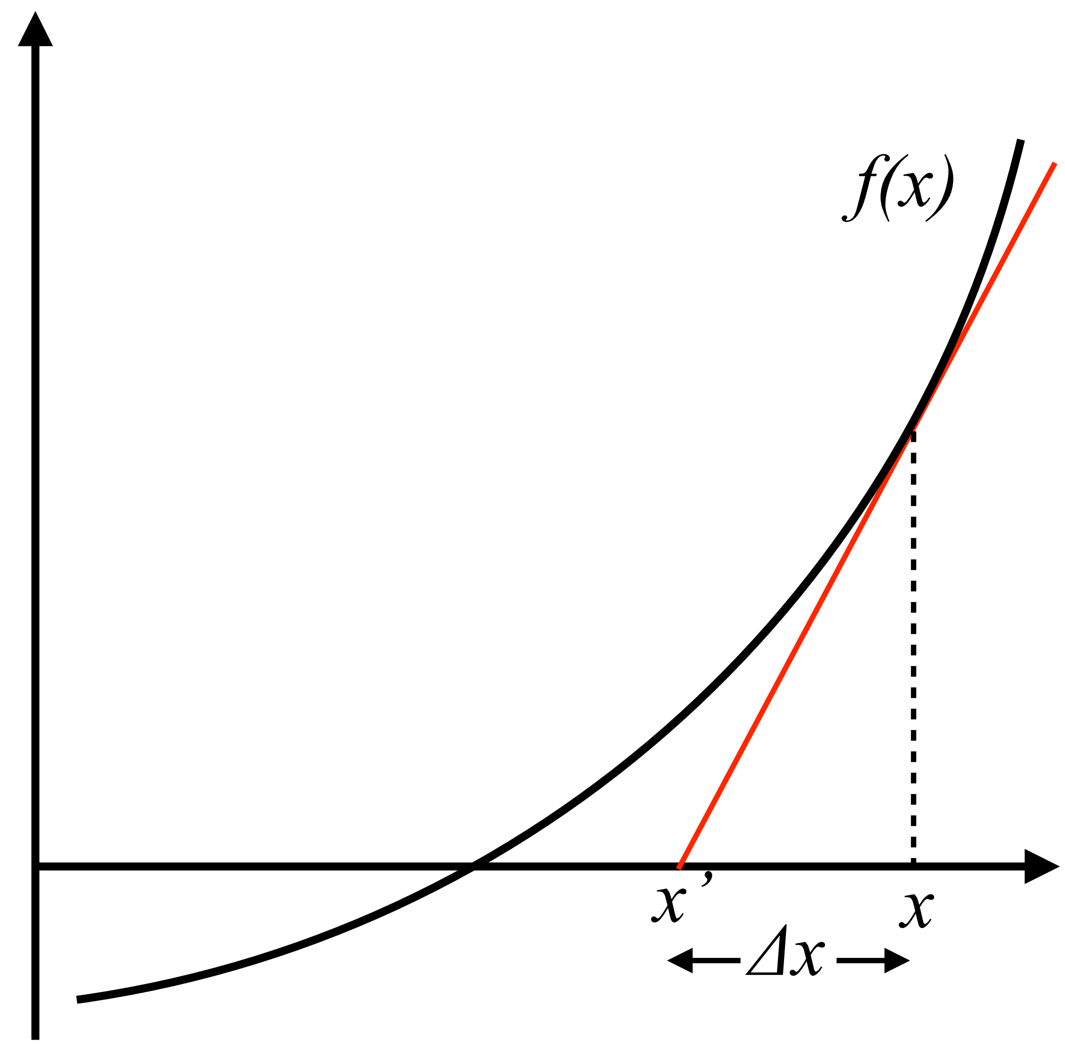
non linear equation: $f(x) = 0$

- In Newton's method we make use of the first derivative.
- We see in the graph that:

$$f'(x) = \frac{f(x)}{\Delta x}$$

- Our new guess x' is then:

$$x' = x - \Delta x = x - \frac{f(x)}{f'(x)}$$

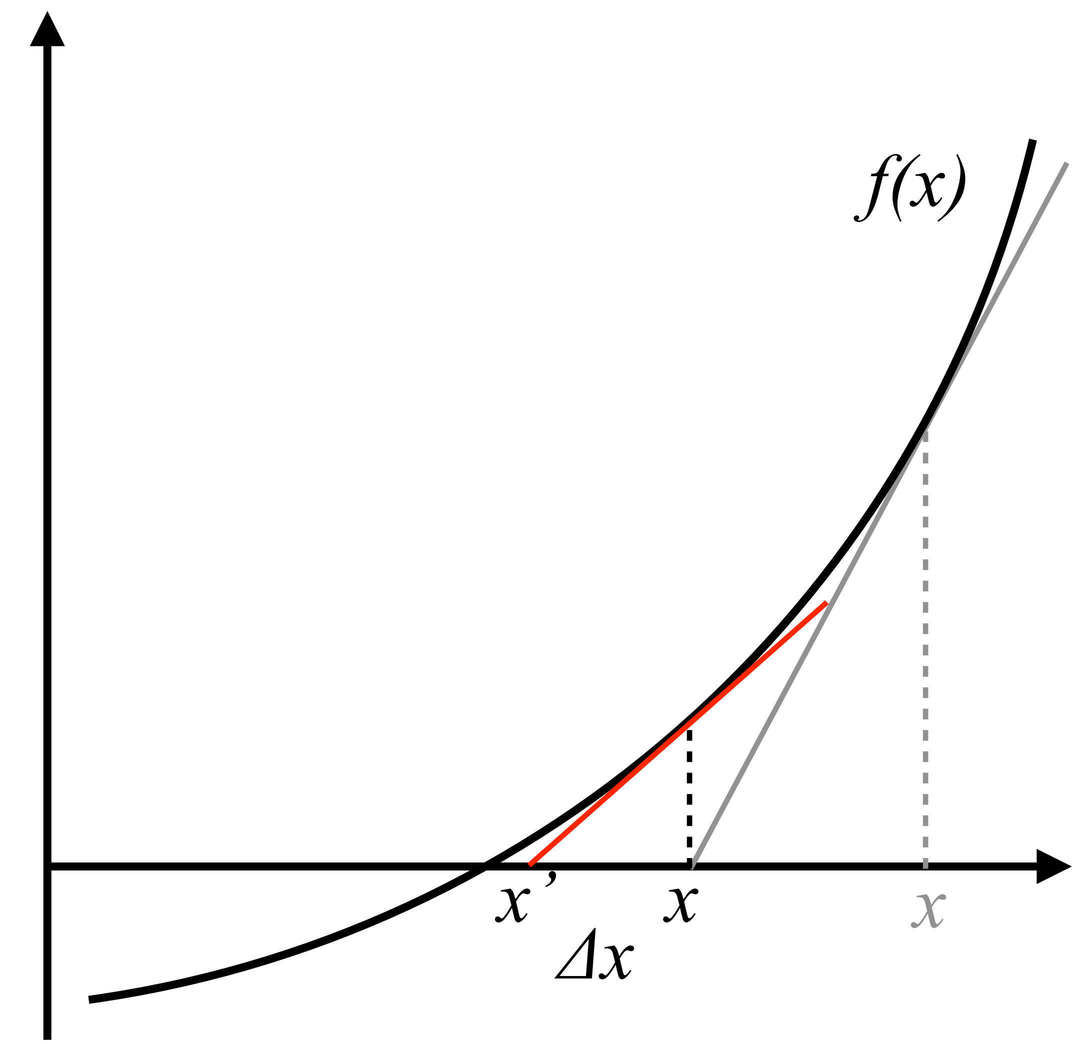


Non linear equations - Newton's method

non linear equation: $f(x) = 0$

- Then we iterate:

$$x' = x - \Delta x = x - \frac{f(x)}{f'(x)}$$



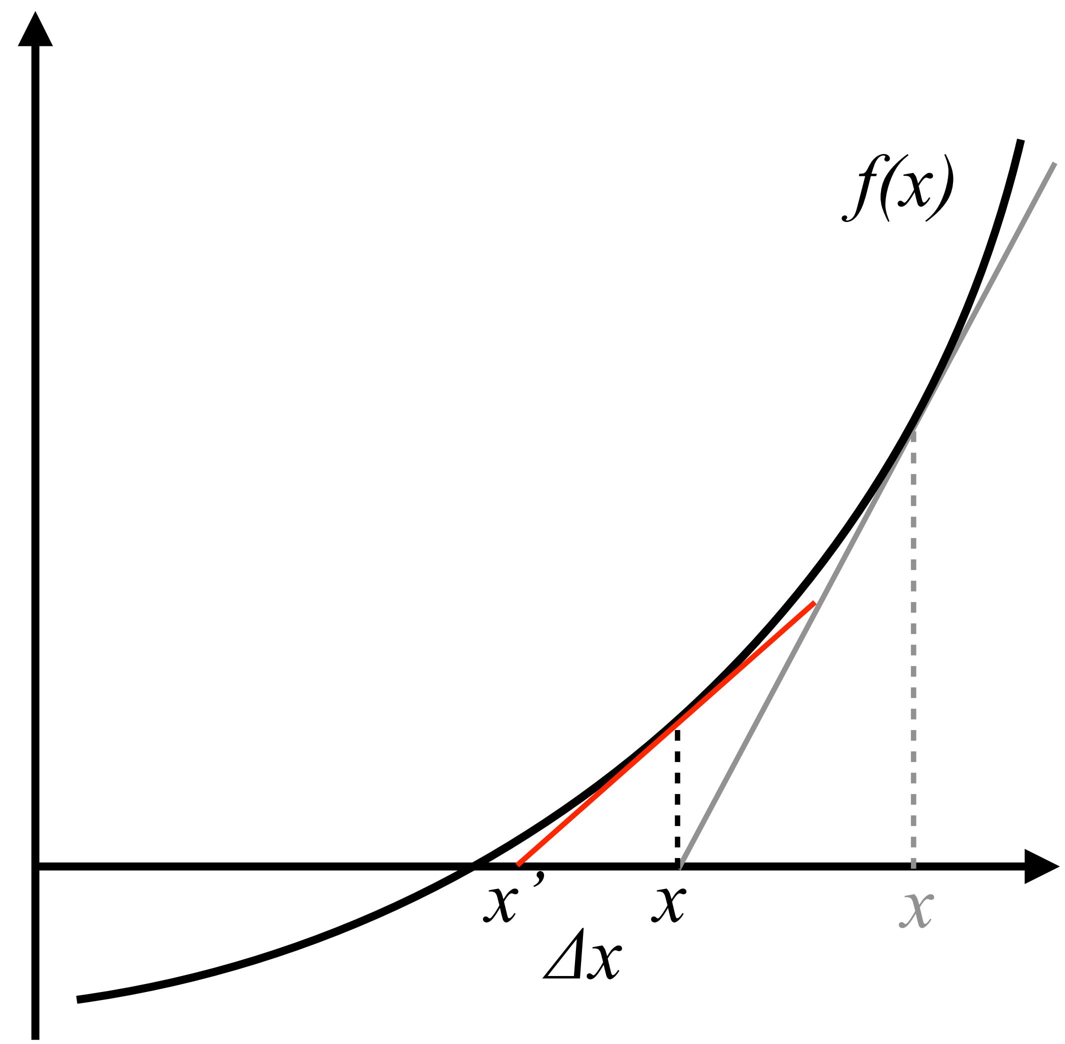
Non linear equations - Newton's method

non linear equation: $f(x) = 0$

- Then we iterate:

$$x' = x - \Delta x = x - \frac{f(x)}{f'(x)}$$

- Newton's method requires access to the first derivative.



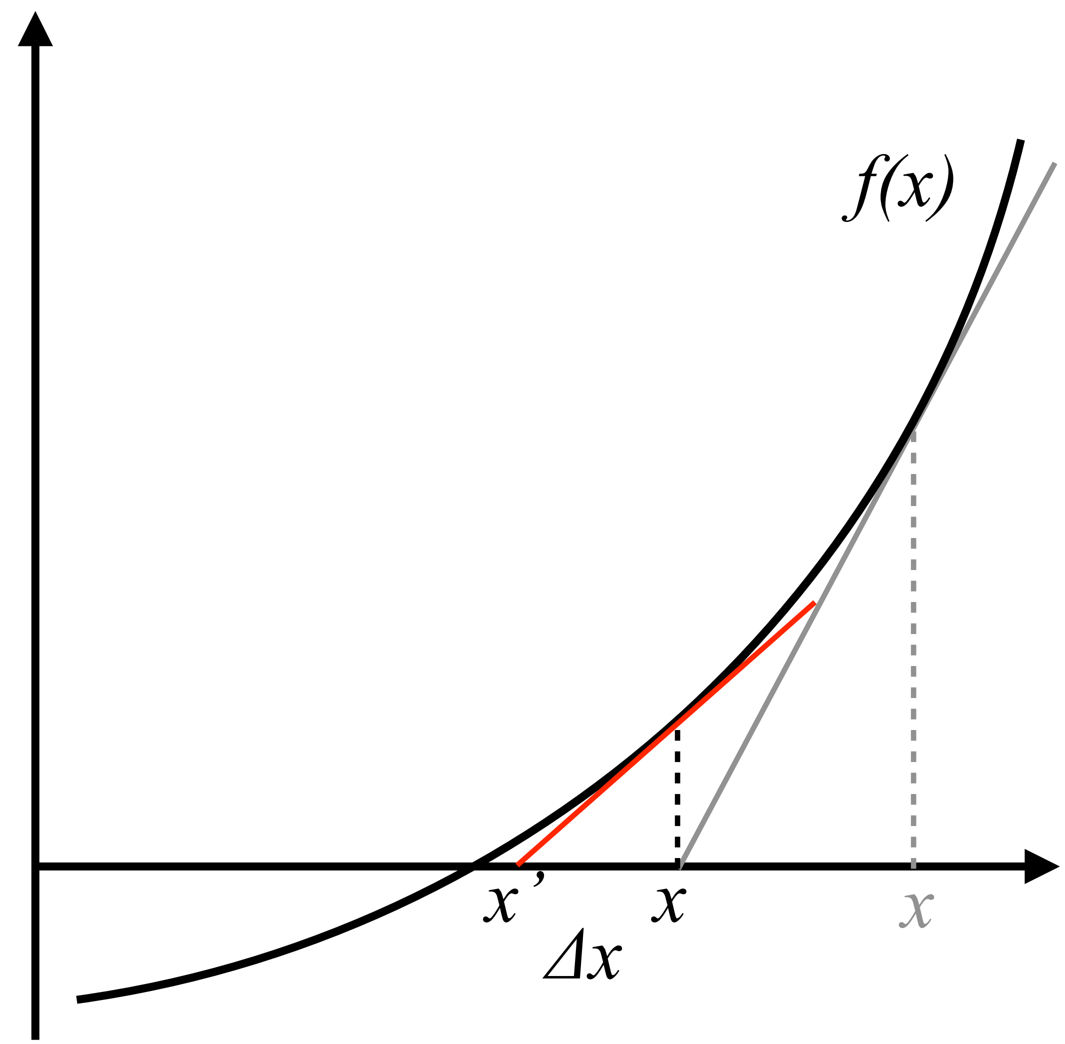
Non linear equations - Newton's method

non linear equation: $f(x) = 0$

- Then we iterate:

$$x' = x - \Delta x = x - \frac{f(x)}{f'(x)}$$

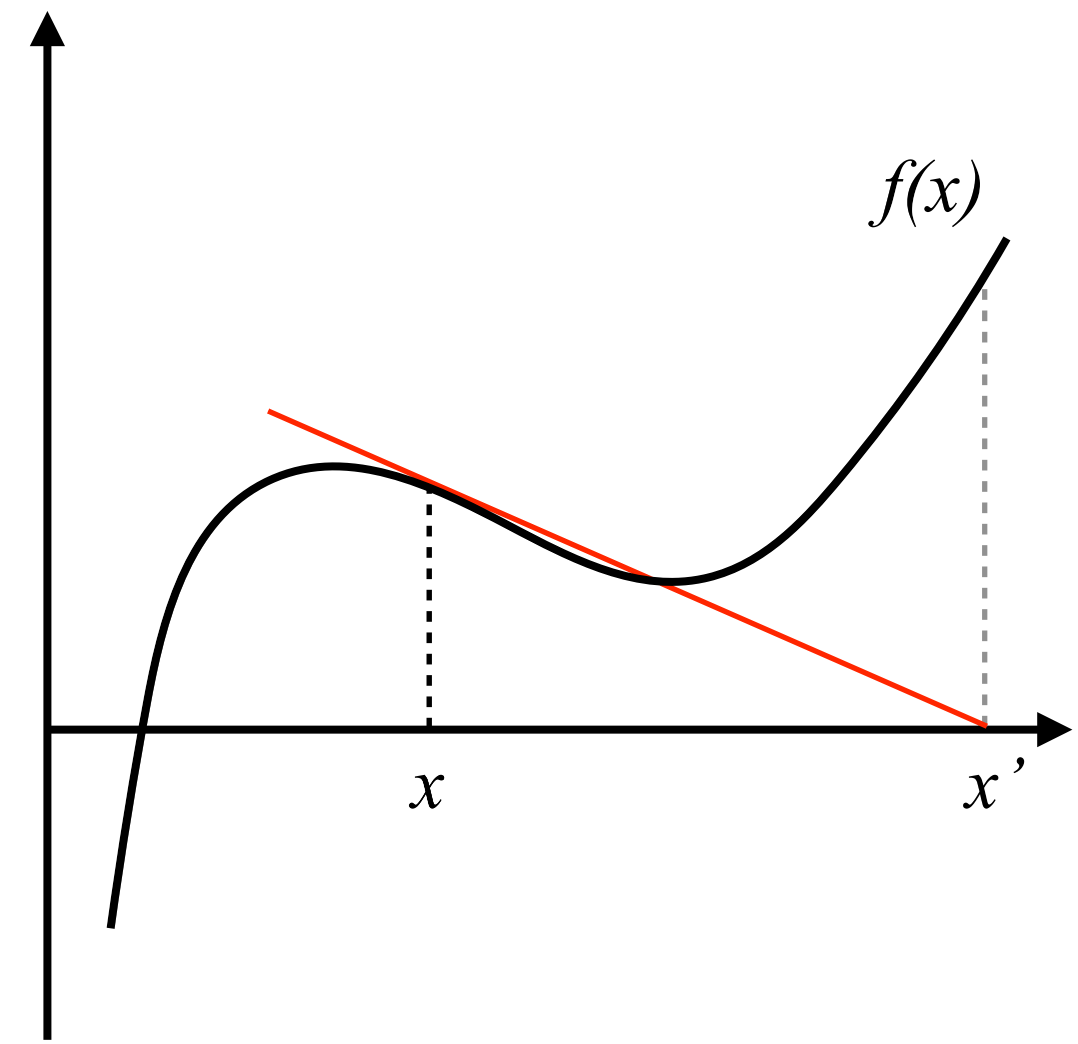
- Newton's method requires access to the first derivative.
- If we do not have it analytically, we now know how to compute the derivative.



Non linear equations - Newton's method

non linear equation: $f(x) = 0$

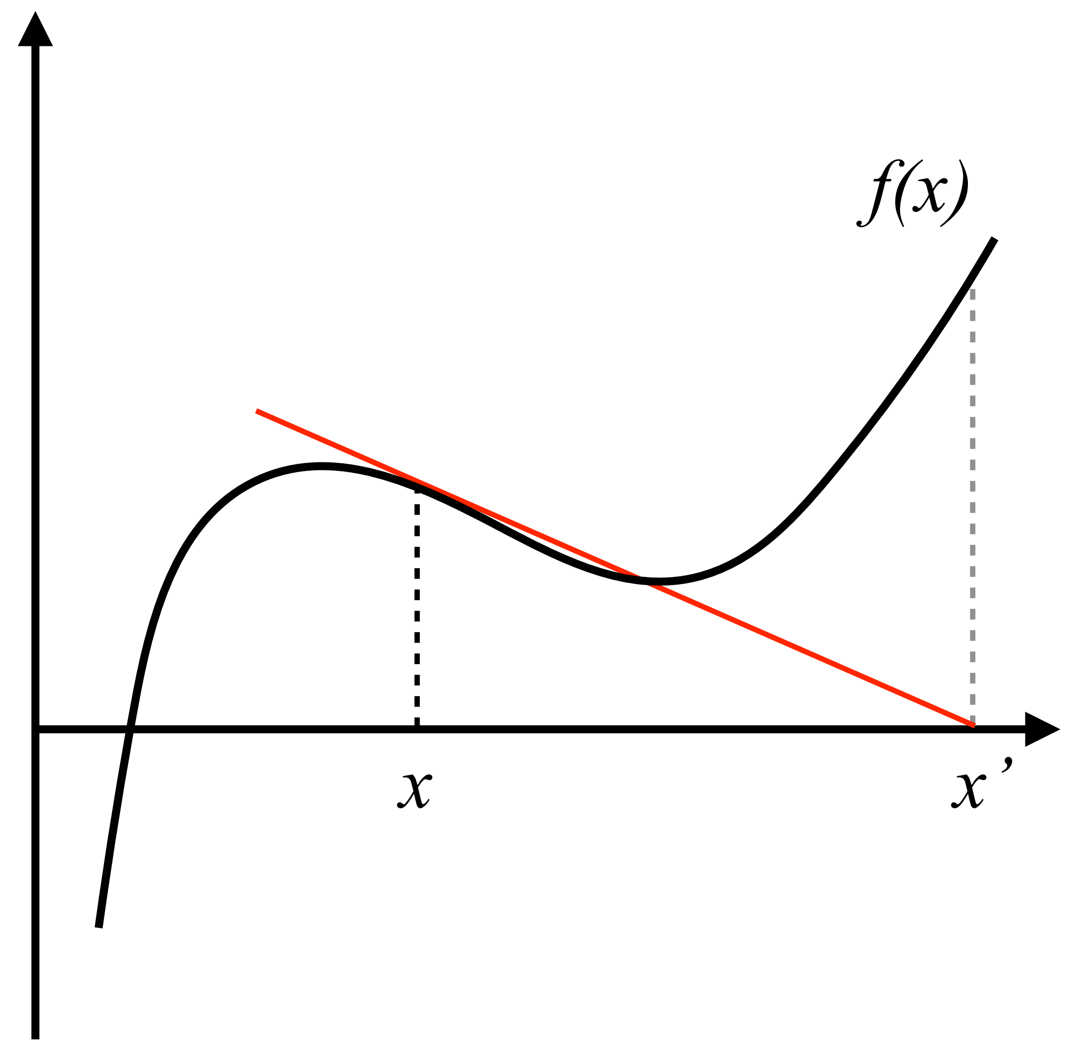
- Newton's method is more robust than the relaxation or bisection method.



Non linear equations - Newton's method

non linear equation: $f(x) = 0$

- Newton's method is more robust than the relaxation or bisection method.
- But it can also fail.



Non-linear equations - Exercise 2

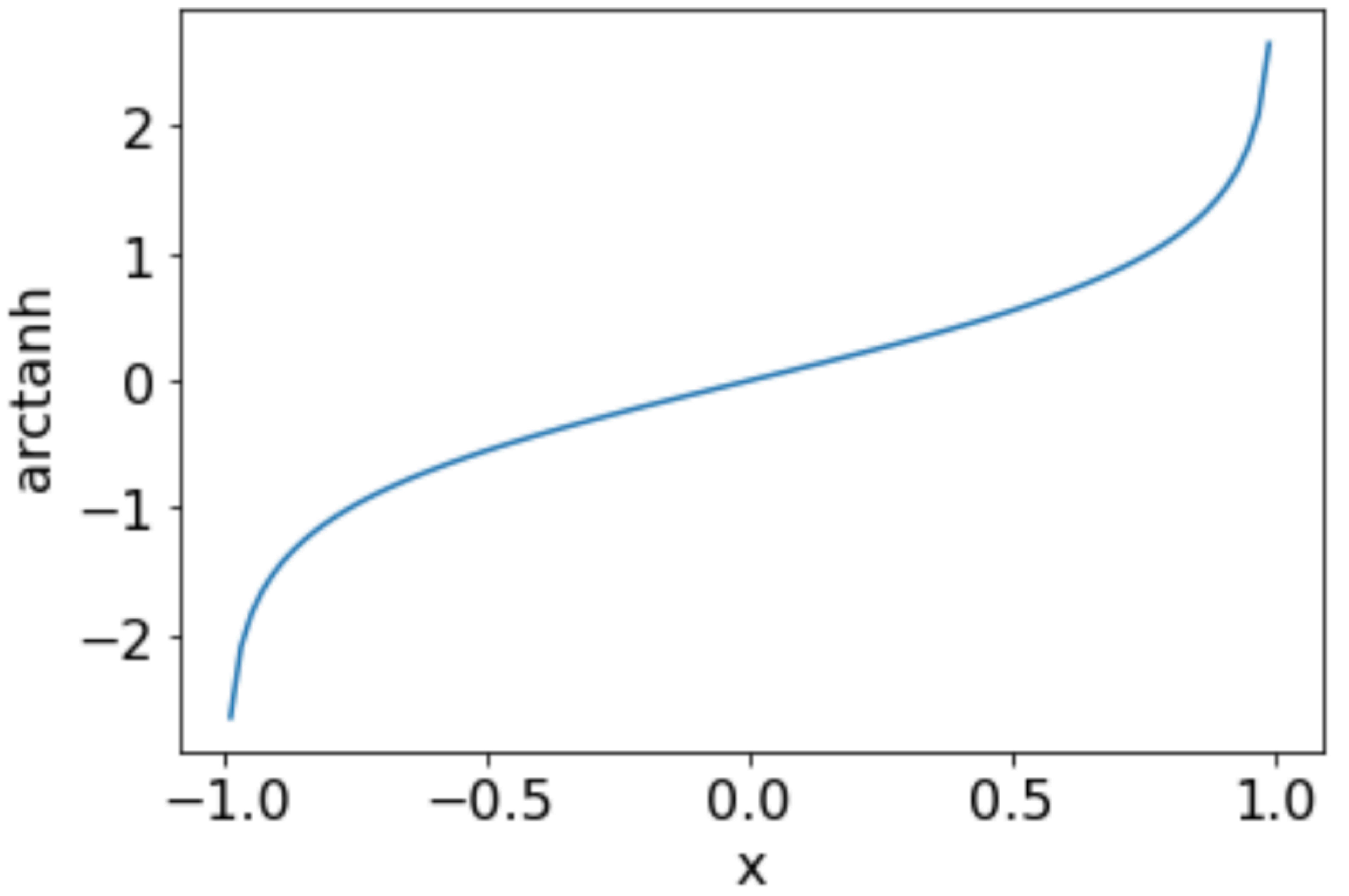
Solve: **a)** $x = e^{1-x^2}$ **b)** $\tanh^{-1}(u)$

1. For a), adapt your relaxation program to Newton's method.
2. For b), consult the in-class exercise sheet. Then write a function that calculates $\tanh^{-1}(u)$.
3. Plot $\tanh^{-1}(u)$ from -1 to 1.

Talking points:

1. **What do you observe?**
2. **How quickly does Newton's method find the right solution?**
3. **Does your function $\tanh^{-1}(u)$ give the right solution?**

Non linear equations - Exercise 2

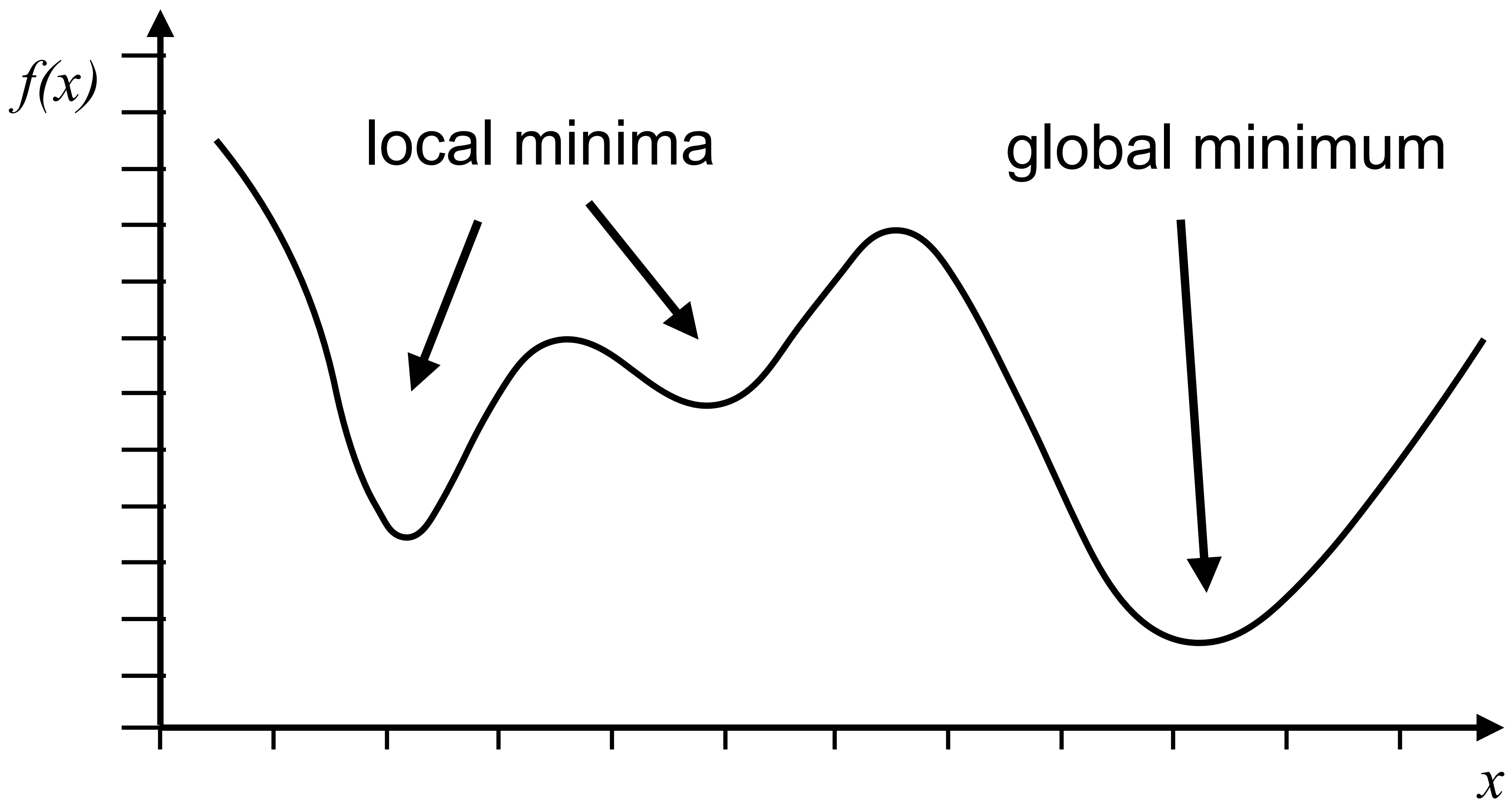


Non linear equations - Newton’s method

Key concept: Newton’s method

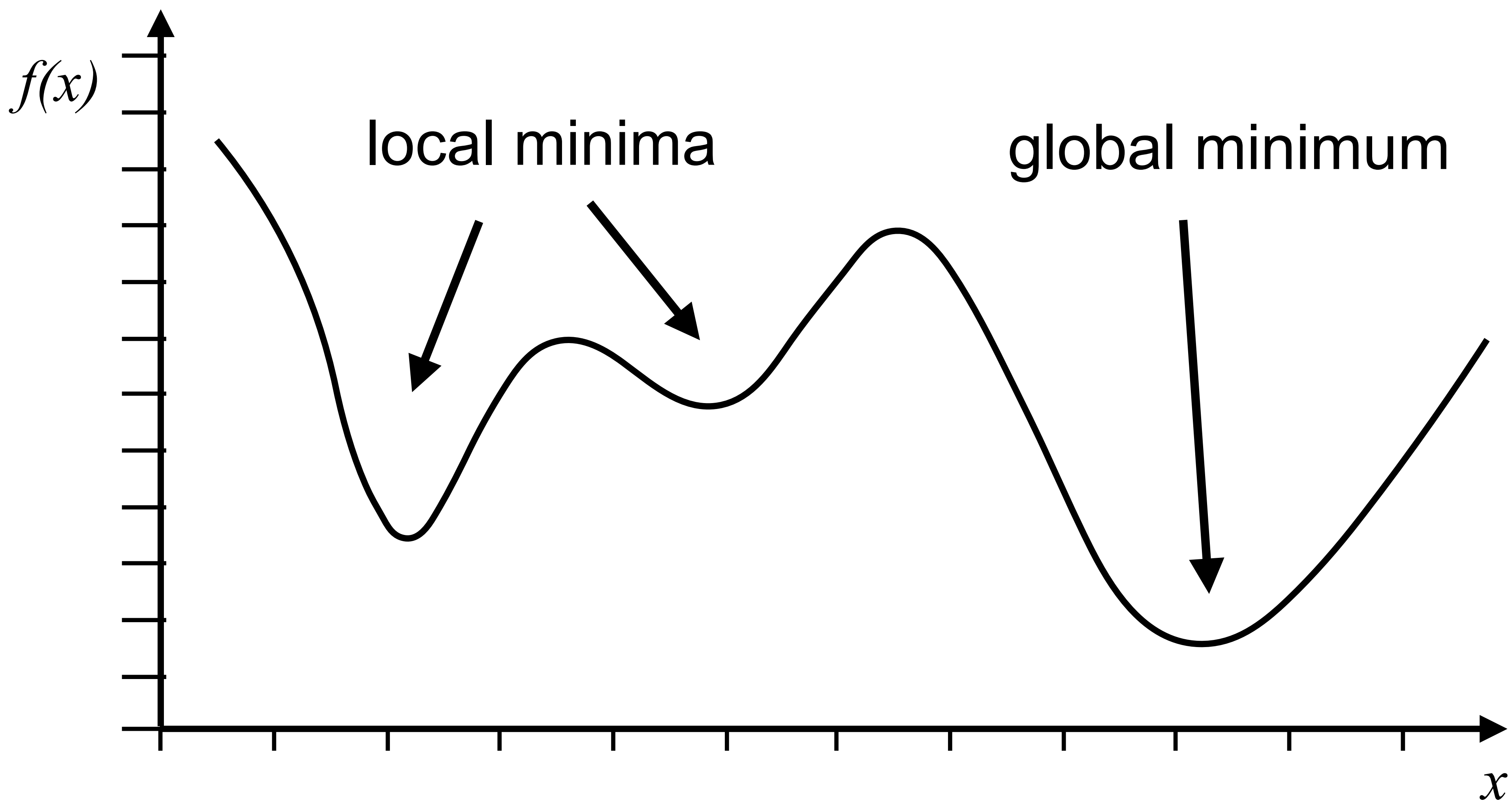
Newton’s method is another simple, iterative method for non-linear equations. It is more robust than the relaxation and the bisection method, but not without fail either.

Maxima and minima of functions



Maxima and minima of functions

At extrema: $$\frac{\partial f(x_1, x_2, \ldots)}{\partial x_i} = 0 \quad \text{for all } i$$



Maxima and minima of functions

At extrema: $$\frac{\partial f(x_1, x_2, \ldots)}{\partial x_i} = 0 \quad \text{for all } i$$

- In principle, we could apply the root finding techniques we just learned directly to find the roots of the first derivatives.

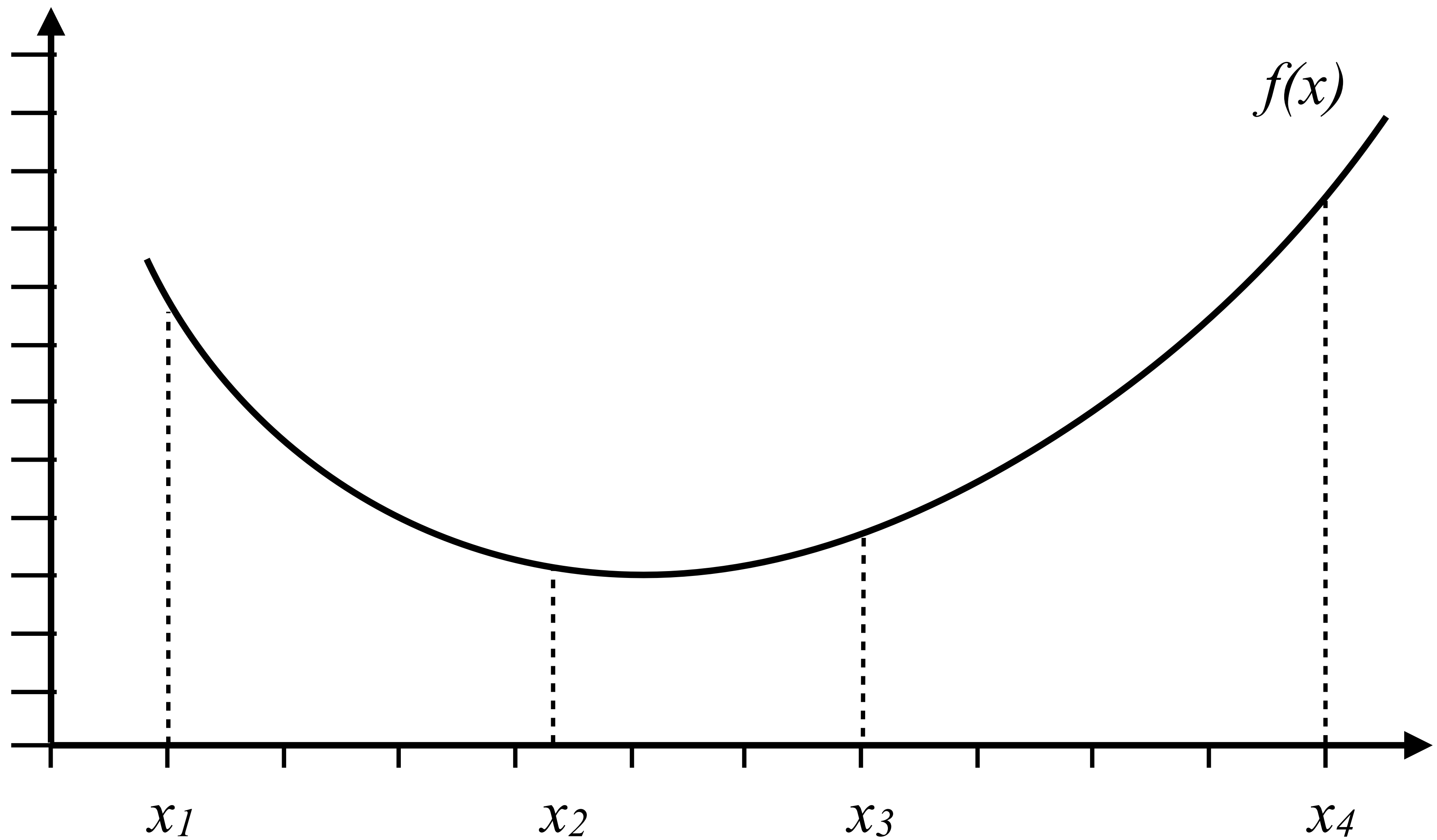
Maxima and minima of functions

At extrema: $$\frac{\partial f(x_1, x_2, \ldots)}{\partial x_i} = 0 \quad \text{for all } i$$

- In principle, we could apply the root finding techniques we just learned directly to find the roots of the first derivatives.
- However, we do not always have access to analytic first derivatives. For this reason, we consider methods that also work without derivatives.

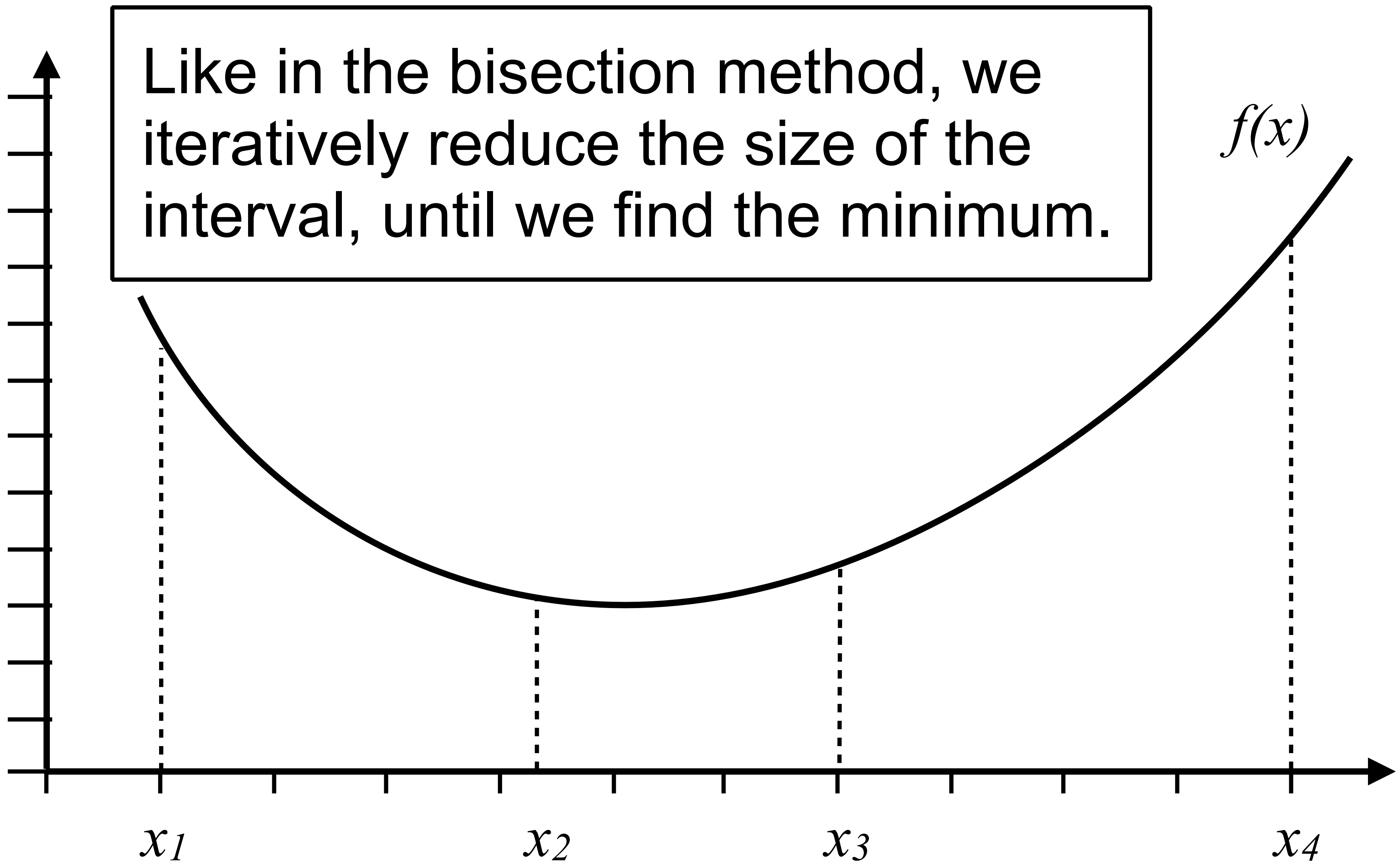
Maxima and minima - Golden ratio search

- Let us assume we are minimising $f(x)$. For maxima, we can always minimise $-f(x)$.



Maxima and minima - Golden ratio search

- Let us assume we are minimising *f(x)*. For maxima, we can always minimise *-f(x)*.



Maxima and minima - Golden ratio search

- Note, we did not half the interval like in the bisection method, but introduced two new points x_3 and x_4.

Maxima and minima - Golden ratio search

- Note, we did not half the interval like in the bisection method, but introduced two new points x_3 and x_4.

- Now we wish to place x_3 and x_4 optimally. Our two conditions are:

  1. The interval in which the minimum falls should decrease by the largest amount possible.
  2. x_3 and x_4 should be positioned symmetrically, since we do not know in which interval the minimum will be.

Maxima and minima - Golden ratio search

- Note, we did not half the interval like in the bisection method, but introduced two new points x_3 and x_4.

- Now we wish to place x_3 and x_4 optimally. Our two conditions are:

    1. The interval in which the minimum falls should decrease by the largest amount possible.
    2. x_3 and x_4 should be positioned symmetrically, since we do not know in which interval the minimum will be.

- To satisfy 2, we choose:

$$x_2 - x_1 = x_4 - x_3$$

Maxima and minima - Golden ratio search

- For 1, we impose that the ratio of the search interval before and after the next partition should be as large as possible.

Maxima and minima - Golden ratio search

- For 1, we impose that the ratio of the search interval before and after the next partition should be as large as possible.

- In our example the minimum lies in the interval $[x_1,x_3]$.

Maxima and minima - Golden ratio search

- For 1, we impose that the ratio of the search interval before and after the next partition should be as large as possible.

- In our example the minimum lies in the interval $[x_1, x_3]$.

- We define the following ratio:

$$z = \frac{x_4 - x_1}{x_3 - x_1} = \frac{x_2 - x_1 + x_3 - x_1}{x_3 - x_1} = \frac{x_2 - x_1}{x_3 - x_1} + 1$$

Here we have used the 2nd condition to eliminate x_4.

Maxima and minima - Golden ratio search

- We define the following ratio:

$$z = \frac{x_4 - x_1}{x_3 - x_1} = \frac{x_2 - x_1 + x_3 - x_1}{x_3 - x_1} = \frac{x_2 - x_1}{x_3 - x_1} + 1$$

- In the next step, the minimum lies in the interval $[x_1, x_2]$, thus:

$$z = \frac{x_3 - x_1}{x_2 - x_1}$$

Maxima and minima - Golden ratio search

- We define the following ratio:

$$z = \frac{x_4 - x_1}{x_3 - x_1} = \frac{x_2 - x_1 + x_3 - x_1}{x_3 - x_1} = \frac{x_2 - x_1}{x_3 - x_1} + 1$$

- In the next step, the minimum lies in the interval $[x_1, x_2]$, thus:

$$z = \frac{x_3 - x_1}{x_2 - x_1}$$

- If we now want z to be the same (condition 1), we get:

$$z = 1/z + 1 \quad \text{or equivalently} \quad z^2 - z - 1 = 0$$

Maxima and minima - Golden ratio search

- If we now want z to be the same (condition 1), we get:

$$z = 1/z + 1 \quad \text{or equivalently} \quad z^2 - z - 1 = 0$$

- z assumes the Golden ratio:

$$z = \frac{1+\sqrt{5}}{2} = 1.618\ldots$$

Golden ratio search - algorithm

1. Choose two initial outside points x_1 and x_4, then calculate the interior points x_2 and x_3 according to the golden ration rule.

Golden ratio search - algorithm

1. Choose two initial outside points x_1 and x_4, then calculate the interior points x_2 and x_3 according to the golden ration rule.
2. Evaluate $f(x)$ at each of the four points and check that at least one of the points x_2 and x_3 gives a function value less than at both x_1 and x_4. Also choose a target accuracy for the position of the minimum.

Golden ratio search - algorithm

1. Choose two initial outside points x_1 and x_4, then calculate the interior points x_2 and x_3 according to the golden ration rule.
2. Evaluate $f(x)$ at each of the four points and check that at least one of the points x_2 and x_3 gives a function value less than at both x_1 and x_4. Also choose a target accuracy for the position of the minimum.
3. If $f(x_2)<f(x_3)$ then the minimum lies between x_1 and x_3. In this case, x_3 becomes the new x_4, x_2 becomes the new x_3 and there will be a new value for x_2, chosen once again according to the golden ratio rule. Evaluate $f(x)$ at this new point.

Golden ratio search - algorithm

4. Otherwise, the minimum lies between x_2 and x_4. Then x_2 becomes the new x_1, x_3 becomes the new x_2, and there will be a new value for x_3. Evaluate $f(x)$ at this new point.

Golden ratio search - algorithm

4. Otherwise, the minimum lies between x_2 and x_4. Then x_2 becomes the new x_1, x_3 becomes the new x_2, and there will be a new value for x_3. Evaluate $f(x)$ at this new point.
5. If $|x_4 - x_1|$ is greater than the target accuracy, repeat from step 3. Otherwise, calculate $0.5(x_2 + x_3)$ and this the final estimate of the position of the minimum.

Golden ratio search - algorithm

4. Otherwise, the minimum lies between x_2 and x_4. Then x_2 becomes the new x_1, x_3 becomes the new x_2, and there will be a new value for x_3. Evaluate $f(x)$ at this new point.
5. If $|x_4 - x_1|$ is greater than the target accuracy, repeat from step 3. Otherwise, calculate $0.5(x_2 + x_3)$ and this the final estimate of the position of the minimum.

- Golden ratio search usually converges fast, but it has the same problem as the bisection method:

  If the minimum does not lie within the initial interval, it cannot be found.

Non-linear equations - Exercise 3

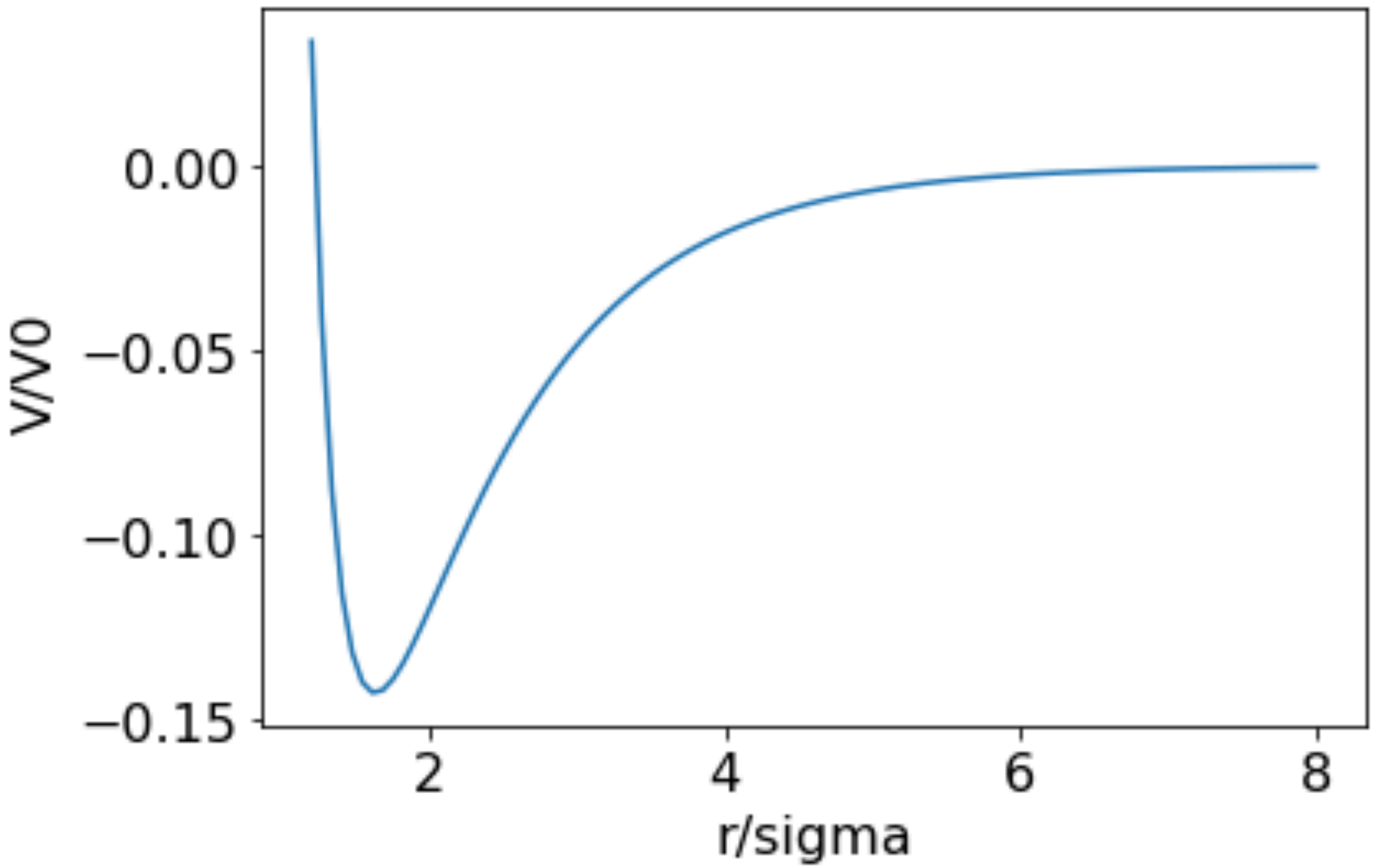
Buckingham potential: $V(r) = V_0 \left[\left(\frac{\sigma}{r}\right)^6 - e^{-r/\sigma} \right]$

1. Plot the Buckingham potential for $\sigma=1$.
2. Complete the golden ratio example program to find the minimum of the potential.
3. Check your computational against the analytic solution.

Talking points:

1. **What do you observe?**
2. **What can you say about the Buckingham potential?**
3. **How does the number of iterations depend on the specified accuracy?**

Exercise 3 - The Buckingham potential



Non linear equations - Golden ratio

Key concept: Golden ratio

Two quantities are in the *golden ratio*, if their ratio is the same as the ratio of their sum to the larger of the two quantities. In minima search, the golden ratio gives the optimal distance for reducing the search interval.

The Gauss-Newton method and gradient descent

- The golden ratio method is robust and reliable, but it cannot be generalized to functions of more than one variable.

The Gauss-Newton method and gradient descent

- The golden ratio method is robust and reliable, but it cannot be generalized to functions of more than one variable.

- It also depends sensitively on the initial search interval. If the minimum does not fall into the search interval, it cannot be found.

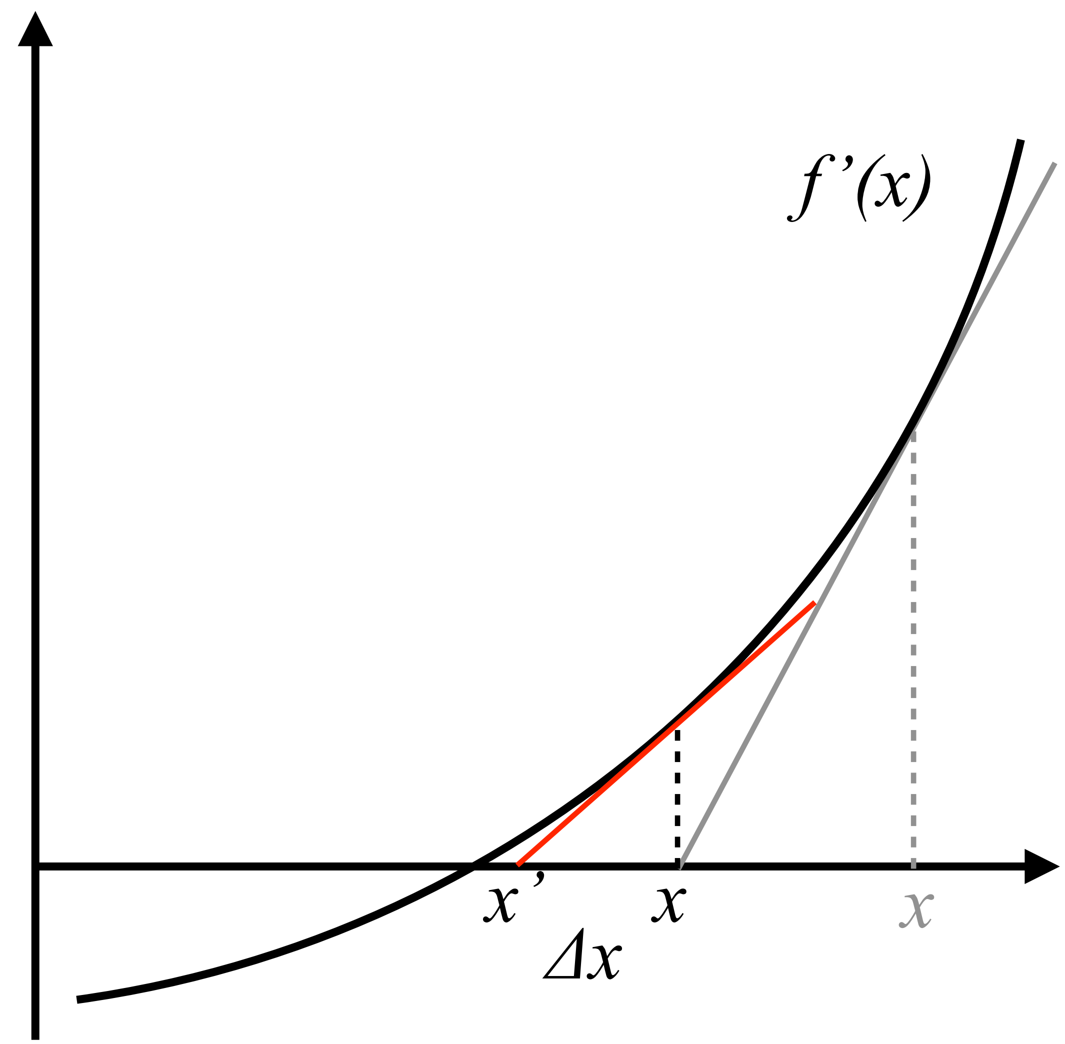
The Gauss-Newton method and gradient descent

- The golden ratio method is robust and reliable, but it cannot be generalized to functions of more than one variable.
- It also depends sensitively on the initial search interval. If the minimum does not fall into the search interval, it cannot be found.
- Let's try something better.

The Gauss-Newton method and gradient descent

at extrema solve: $f'(x) = 0$

- Minima or maxima are the roots of the first derivative.



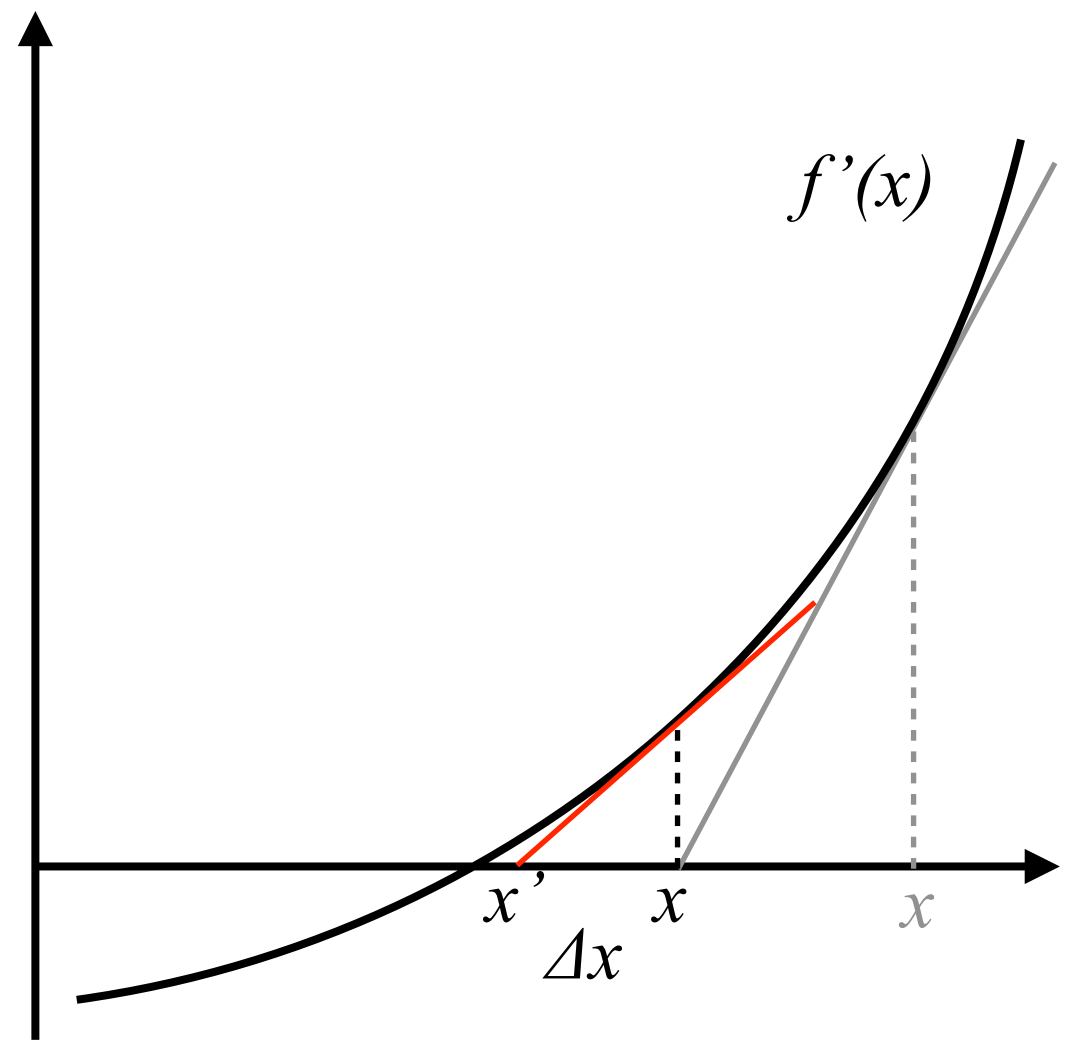
The Gauss-Newton method and gradient descent

at extrema solve: $f'(x) = 0$

- Minima or maxima are the roots of the first derivative.
- Newton's method was good for finding roots. Let's apply it:

$$x' = x - \Delta x = x - \frac{f'(x)}{f''(x)}$$

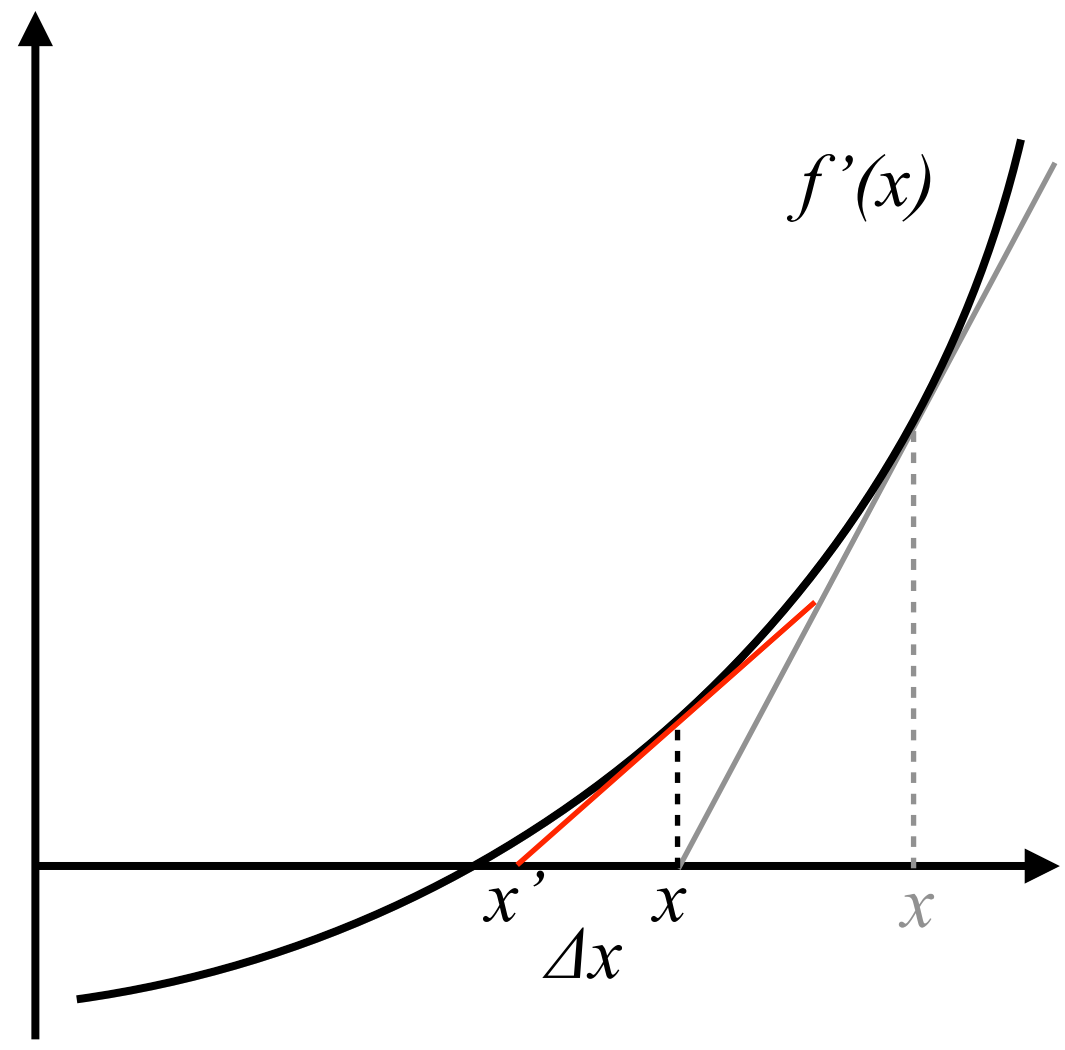
- This is the Gauss-Newton method.



The Gauss-Newton method and gradient descent

at extrema solve: $f'(x) = 0$

- If we have access to the 2nd derivative, Gauss-Newton's method works great.

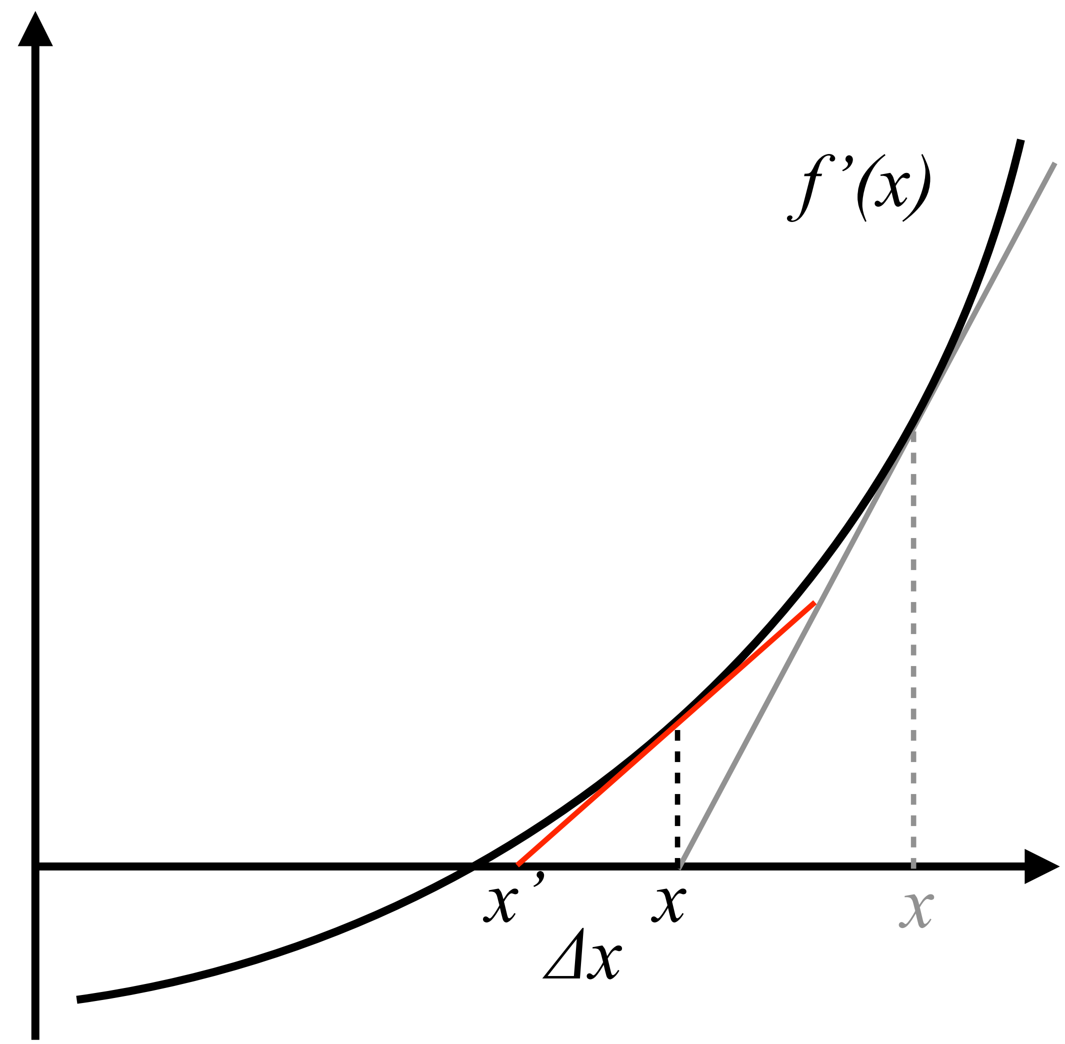


The Gauss-Newton method and gradient descent

at extrema solve: $f'(x) = 0$

- If we have access to the 2nd derivative, Gauss-Newton's method works great.
- If we don't, then we make a rough guess for it:

$$\gamma \approx \frac{1}{f''(x)}$$

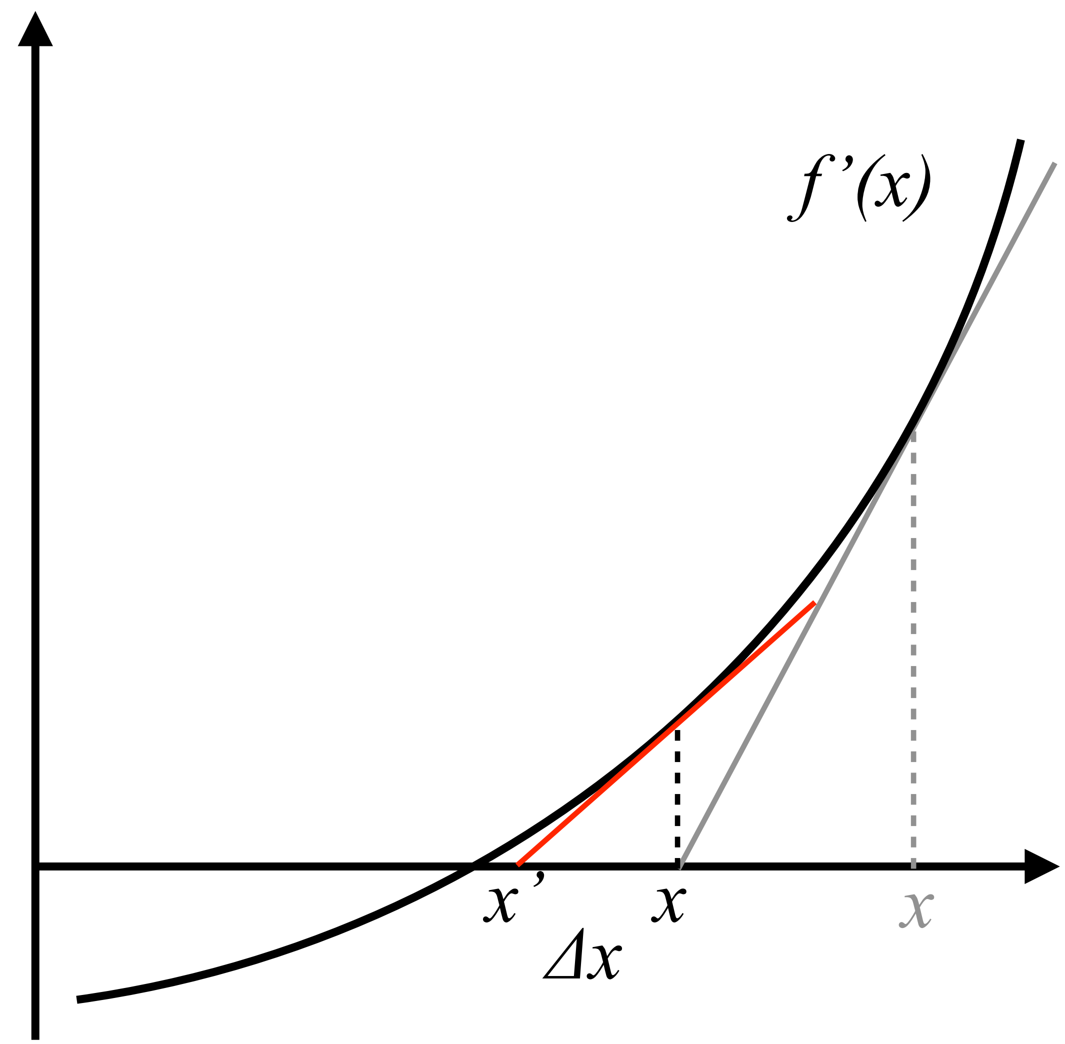


The Gauss-Newton method and gradient descent

at extrema solve: $f'(x) = 0$

- We obtain:

$$x' = x - \gamma f'(x)$$



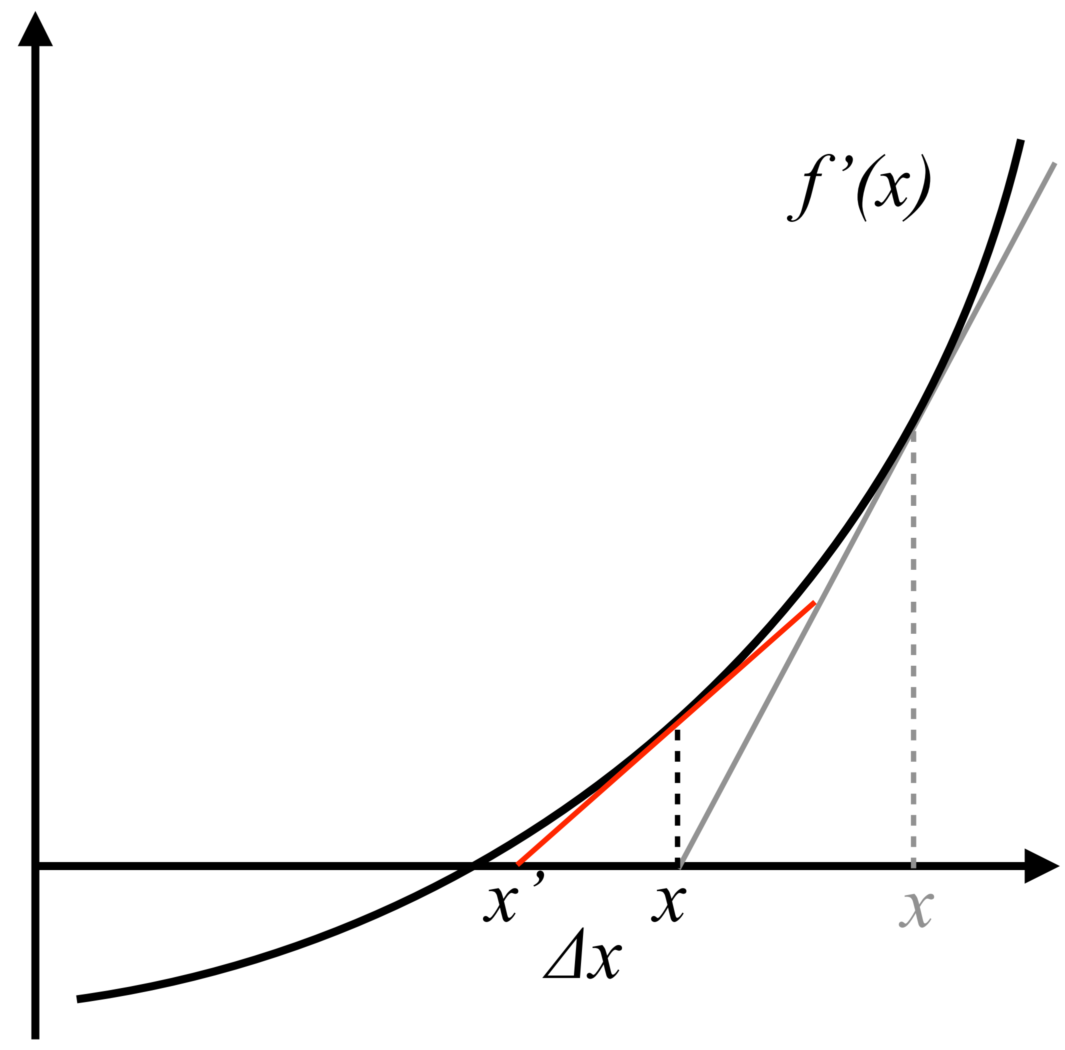
The Gauss-Newton method and gradient descent

at extrema solve: $f'(x) = 0$

- We obtain:

$$x' = x - \gamma f'(x)$$

- This method is called *gradient descent*.



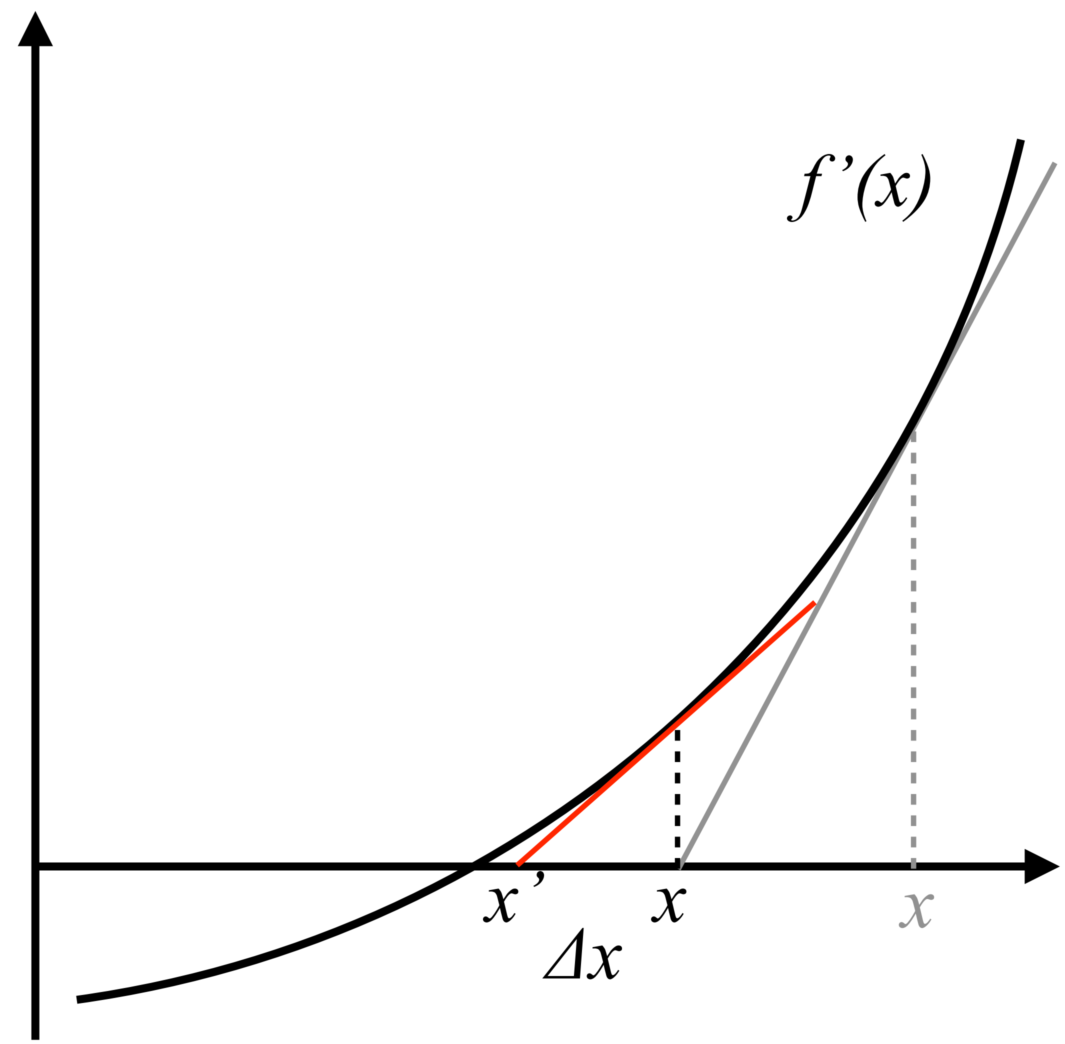
The Gauss-Newton method and gradient descent

at extrema solve: $f'(x) = 0$

- We obtain:

$$x' = x - \gamma f'(x)$$

- This method is called *gradient descent*.
- For reasonable values of γ, it will give you an answer with a reasonable number of steps.



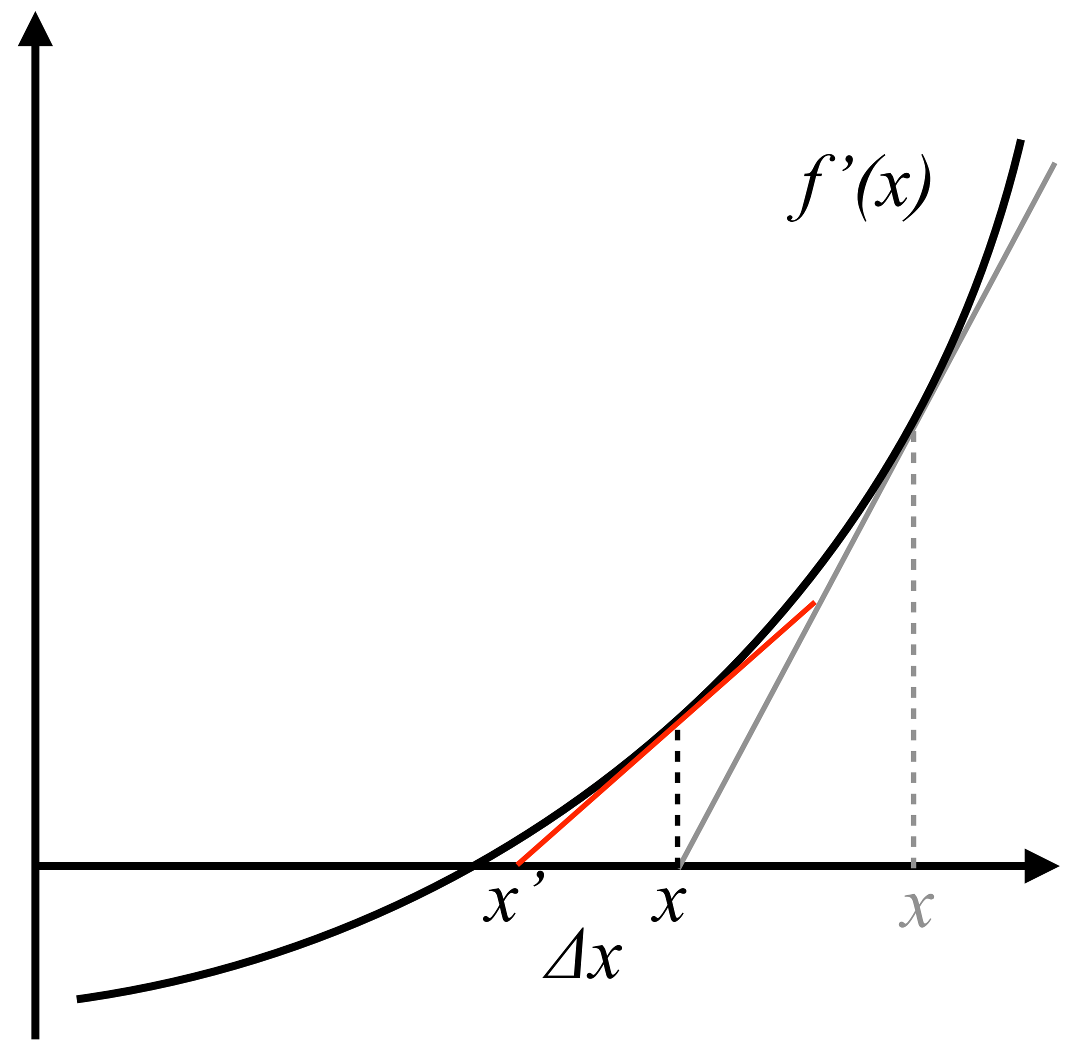
The Gauss-Newton method and gradient descent

at extrema solve: $f'(x) = 0$

- We obtain:

$$x' = x - \gamma f'(x)$$

for $\gamma > 0$ we find minima
for $\gamma < 0$ we find maxima

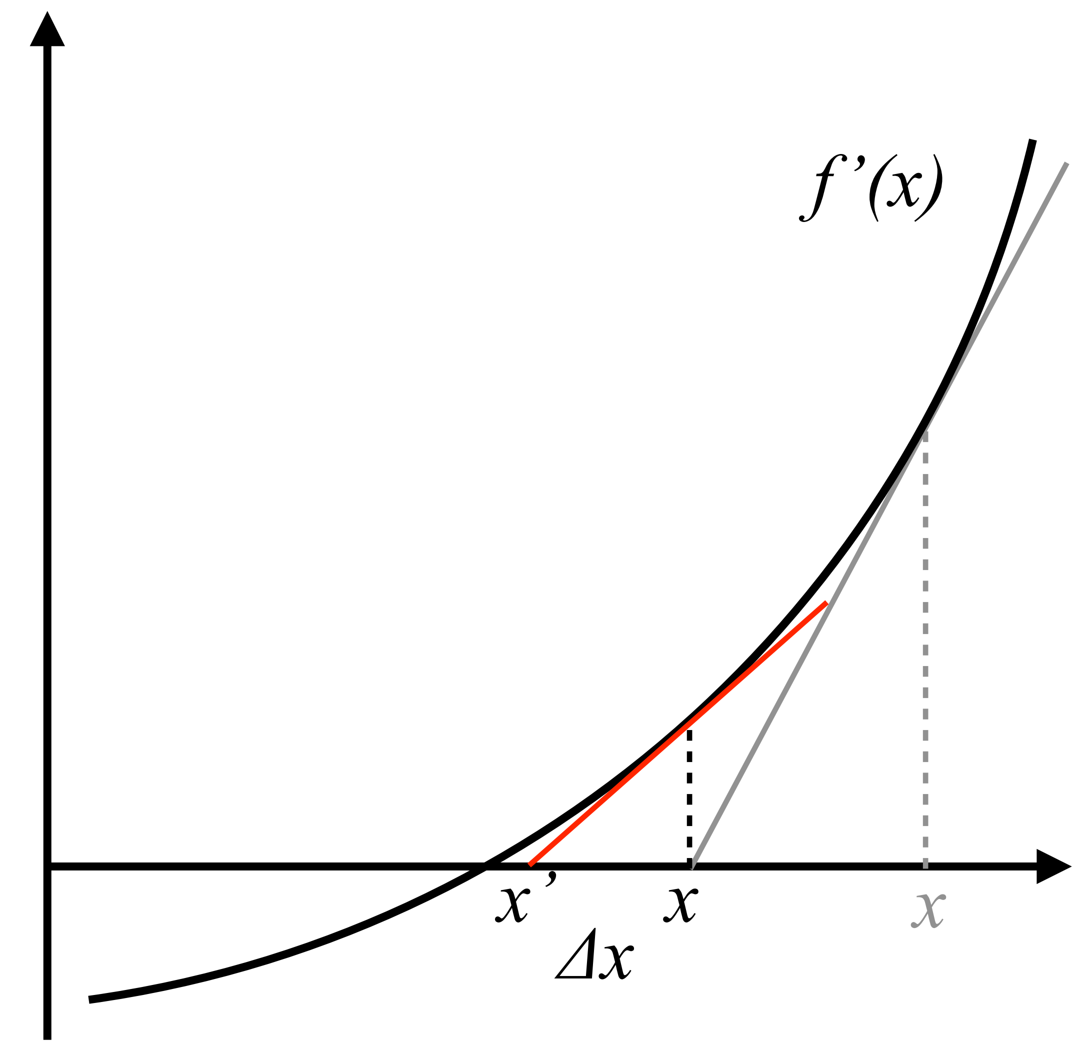


The Gauss-Newton method and gradient descent

at extrema solve: $f'(x) = 0$

- If we also don't have access to the first derivative, we have to approximate that, too:

$$x_3 = x_2 - \gamma \frac{f(x_2) - f(x_1)}{x_2 - x_1}$$



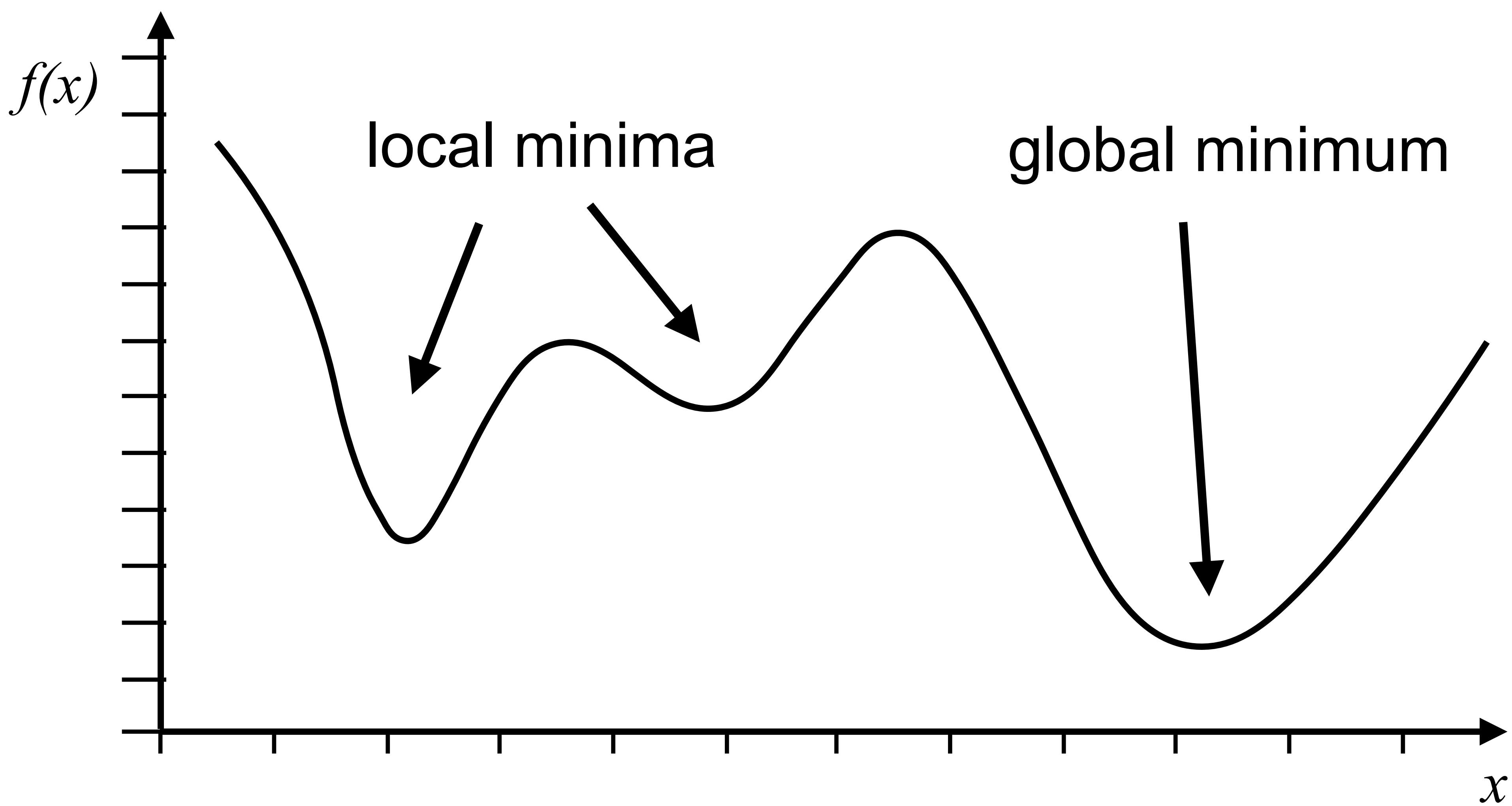
The Gauss-Newton method and gradient descent

Key concept: gradient descent

Gradient descent is a very common optimisation (meaning minima/maxima finding) method. It requires knowledge of the derivative of the optimisation objective (i.e. here our function $f(x)$). If the derivative is not available, it needs to be approximated numerically.

Maxima and minima of functions

Note that we have only discussed *local* and not *global* optimisation schemes!



Non-linear equations - Exercise 4

Buckingham potential: $V(r) = V_0 \left[\left(\frac{\sigma}{r} \right)^6 - e^{-r/\sigma} \right]$

Find the minimum of the Buckingham potential for σ=1:

1. For the Gauss-Newton method. Start from r=σ.
2. For gradient descent.
3. For gradient descent with numeric 1st derivative.

Talking points:

1. **What do you observe?**
2. **What happens when you start from r=4σ and why?**
3. **What is a good value for γ in gradient descent?**