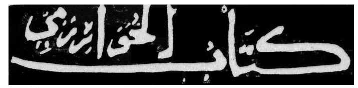
Résolution d'équations

- L'identification des points critiques revient à résoudre

$$\nabla f(x) = 0.$$

- Il s'agit d'un système de n équations à n inconnues.
- On verra que les conditions d'optimalité pour les problèmes avec contraintes se ramènent également à un système d'équations.
- Analysons d'abord les algorithmes permettant de résoudre ces systèmes d'équations.
- Ils seront ensuite adaptés pour les problèmes d'optimisation.

Algorithme



- Abu Ja’far Muhammad ibn Musa Al-Khwarizmi (780—840)
- Traité *al Kitab almukhtasar fi hisab al-jabr w’al muqabala*, qui est à l’origine de l’algèbre
- traduction latine de cet ouvrage, intitulée *Algoritmi de numero Indorum*

Méthode de Newton

Idée :

- Méthode itérative
- A chaque itération, simplification du problème
- Technique : linéarisation

Equation à une inconnue

$$f(x) = x^2 - 2, \quad \widehat{x} = 2$$

Théorème de Taylor

$$\begin{aligned} f(\widehat{x} + d) &= f(\widehat{x}) + df'(\widehat{x}) + o(|d|) \\ &= \widehat{x}^2 - 2 + 2\widehat{x}d + o(|d|) \\ &= 2 + 4d + o(|d|). \end{aligned}$$

Equation à une inconnue

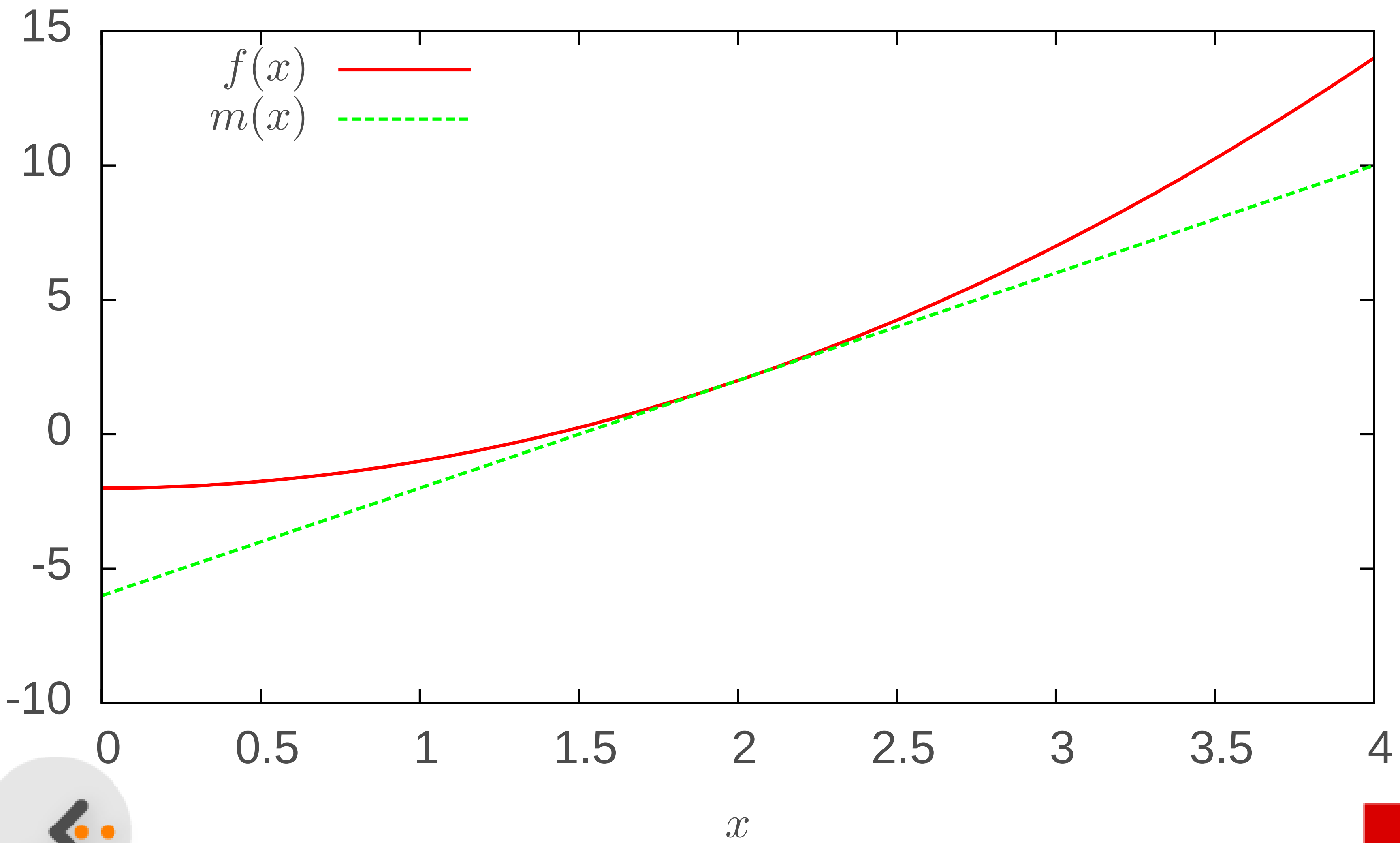
Ignorons le terme d'erreur pour obtenir un modèle :

$$m(\widehat{x}+d) = 2+4d.$$

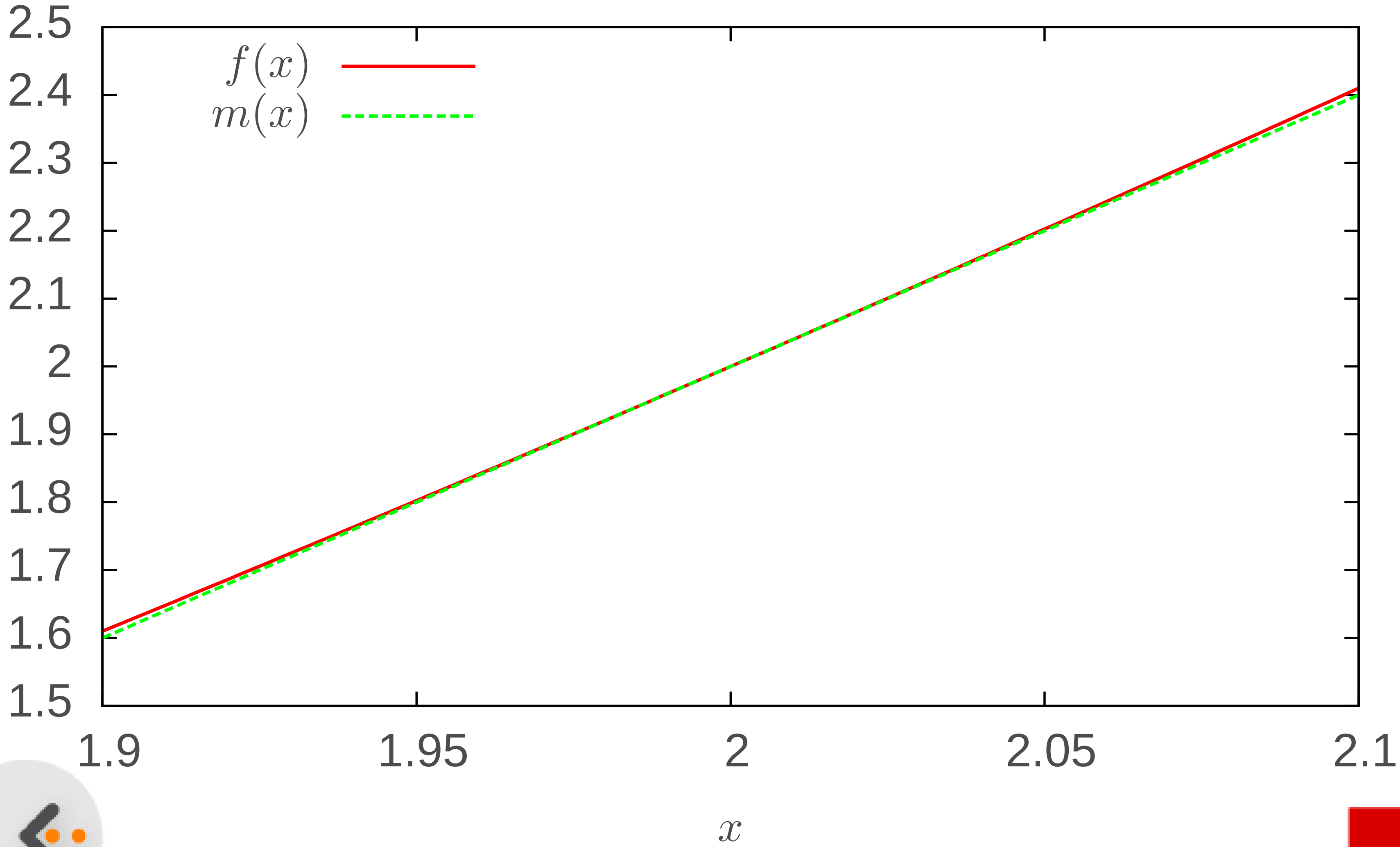
En posant $x = \widehat{x}+d$, nous obtenons

$$m(x) = 2+4(x-2) = 4x-6.$$

Equation à une inconnue



Equation à une inconnue



Equation à une inconnue

Modèle linéaire d'une fonction à une variable

Soit $f : \mathbb{R} \to \mathbb{R}$ *une fonction différentiable. Le* modèle linéaire *de* f *en* $\widehat{x}$ *est une fonction* $m_{\widehat{x}} : \mathbb{R} \to \mathbb{R}$ *définie par*

$$m_{\widehat{x}}(x) = f(\widehat{x}) + (x - \widehat{x})f'(\widehat{x}).$$

Equation à une inconnue

Algorithme :

1. Calculer le modèle linéaire en $\widehat{x}$:

$$f(\widehat{x}) + (x - \widehat{x}) f'(\widehat{x}) = 0,$$

2. Calculer sa racine x^+

$$x^+ = \widehat{x} - \frac{f(\widehat{x})}{f'(\widehat{x})},$$

3. Si x^+ n'est pas une racine du problème de départ, considérer x^+ comme nouvelle approximation, et recommencer.

Equation à une inconnue

Critère d'arrêt :

- En théorie $f(x^+) = 0$.
- En pratique, arithmétique finie.
- On définit une précision ε, et la condition est

$$|f(x^+)| \leq \varepsilon.$$

Algorithme : Méthode de Newton — une variable

Objectif

Trouver une approximation de la solution de l'équation

$$f(x) = 0.$$

Input

- La fonction $f : \mathbb{R} \to \mathbb{R}$;
- La dérivée de la fonction $f' : \mathbb{R} \to \mathbb{R}$;
- Une première approximation de la solution $x_0 \in \mathbb{R}$;
- La précision demandée $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$.

Algorithme : Méthode de Newton — une variable

Output

Une approximation de la solution $x^* \in \mathbb{R}$

Initialisation

$k = 0$

Itérations

1. $x_{k+1} = x_k - f(x_k)/f'(x_k)$,
2. $k = k + 1$.

Critère d'arrêt

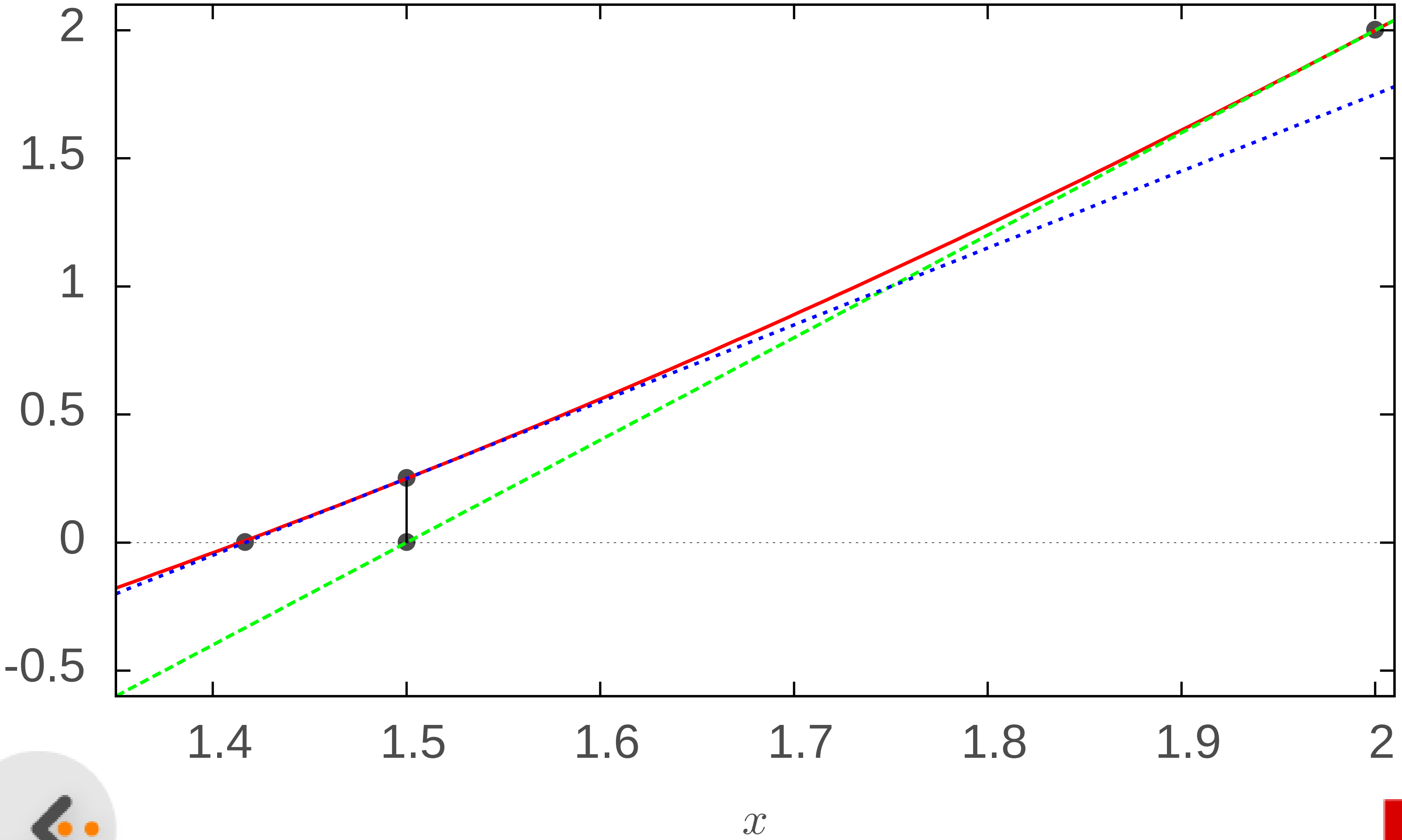
Si $|f(x_k)| \leq \varepsilon$, alors $x^* = x_k$.

Méthode de Newton — une variable

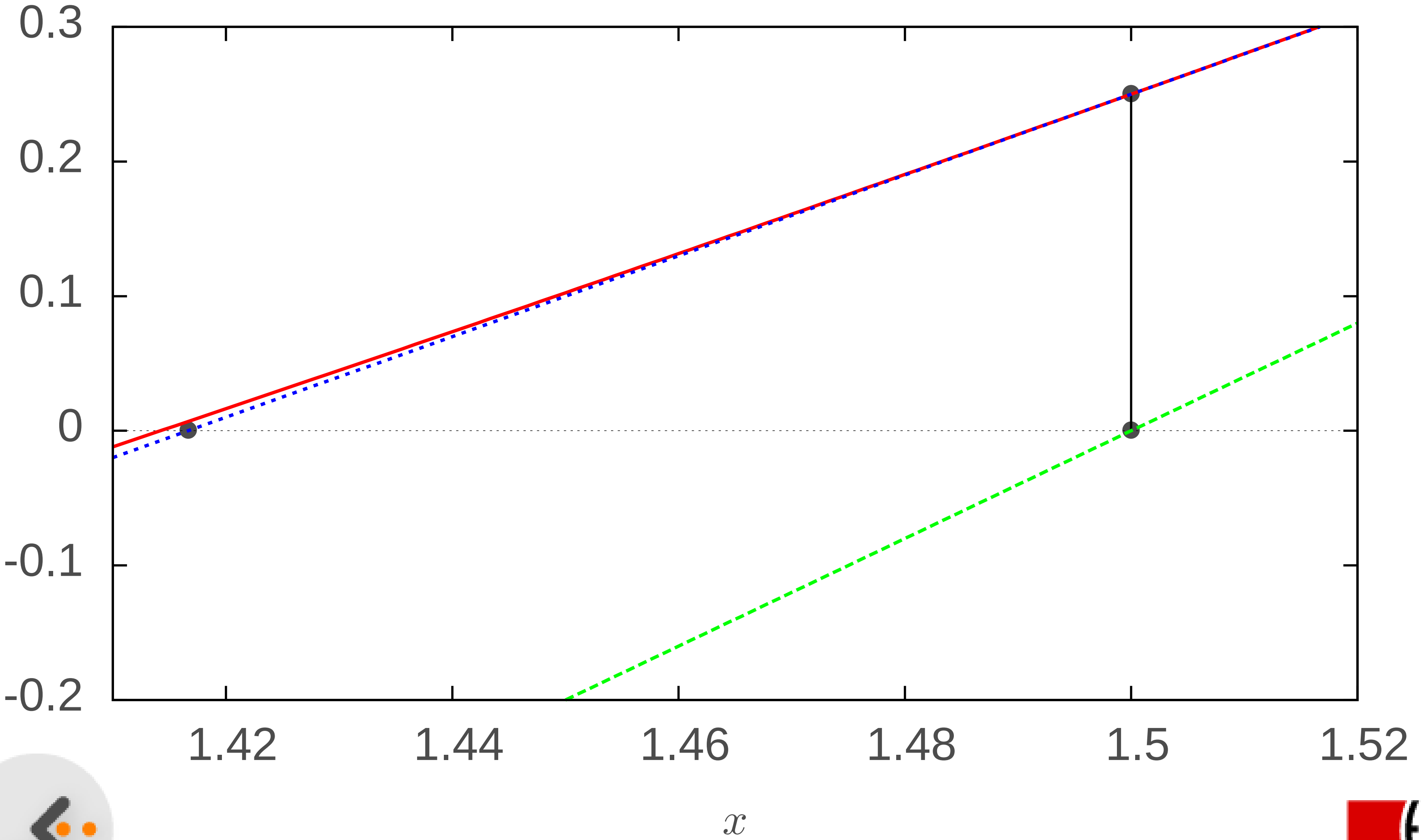
$$f(x) = x^2 - 2 = 0, \quad x_0 = 2, \quad \varepsilon = 10^{-15}$$

k	x_k	$f(x_k)$	$f'(x_k)$
0	+2.00000000E+00	+2.00000000E+00	+4.00000000E+00
1	+1.50000000E+00	+2.50000000E-01	+3.00000000E+00
2	+1.41666667E+00	+6.94444444E-03	+2.83333333E+00
3	+1.41421569E+00	+6.00730488E-06	+2.82843137E+00
4	+1.41421356E+00	+4.51061410E-12	+2.82842712E+00
5	+1.41421356E+00	+4.44089210E-16	+2.82842712E+00

Méthode de Newton — une variable



Méthode de Newton — une variable



Méthode de Newton — une variable

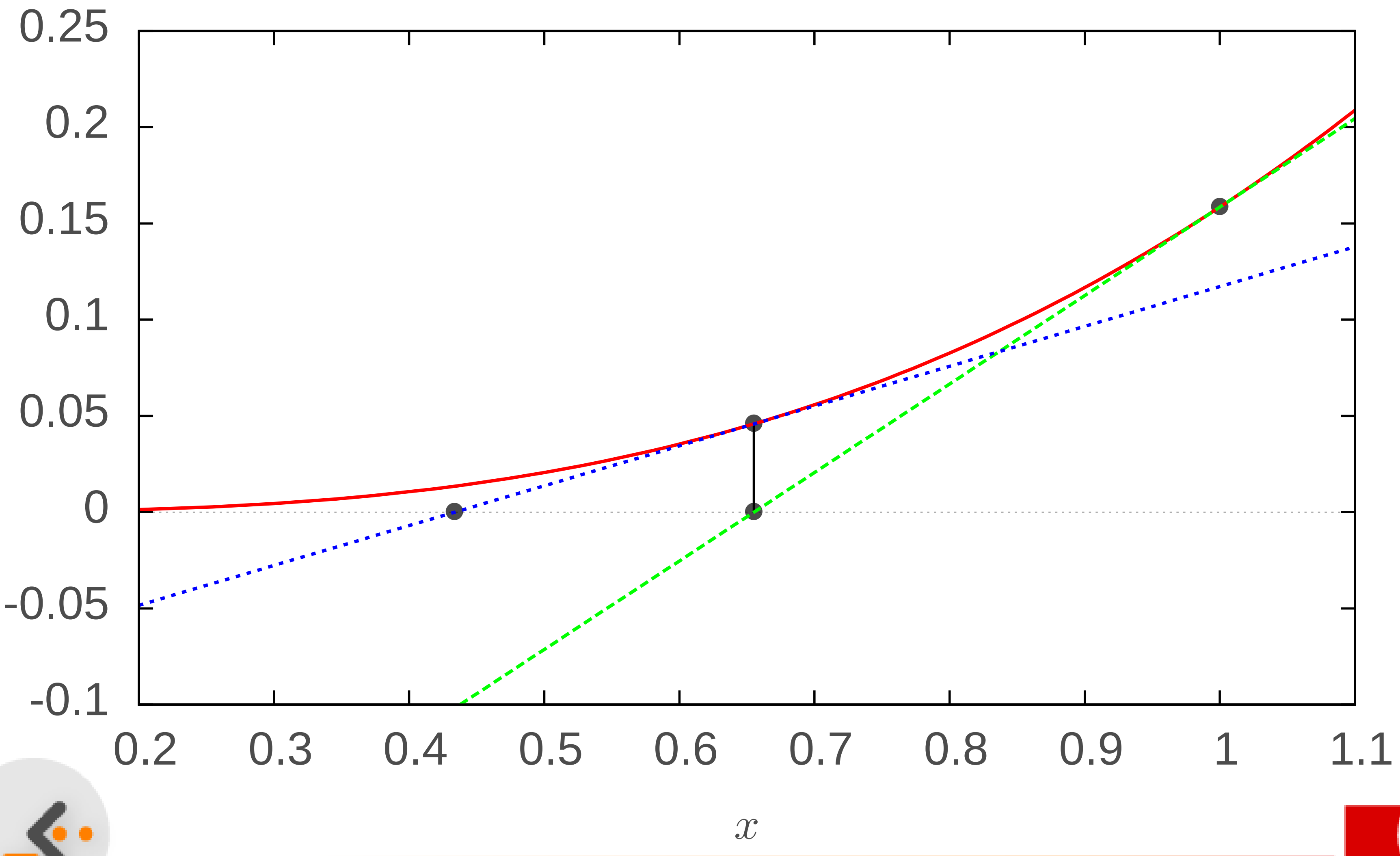
- La méthode semble très rapide
- Malheureusement, elle ne fonctionne pas toujours bien

Méthode de Newton — une variable

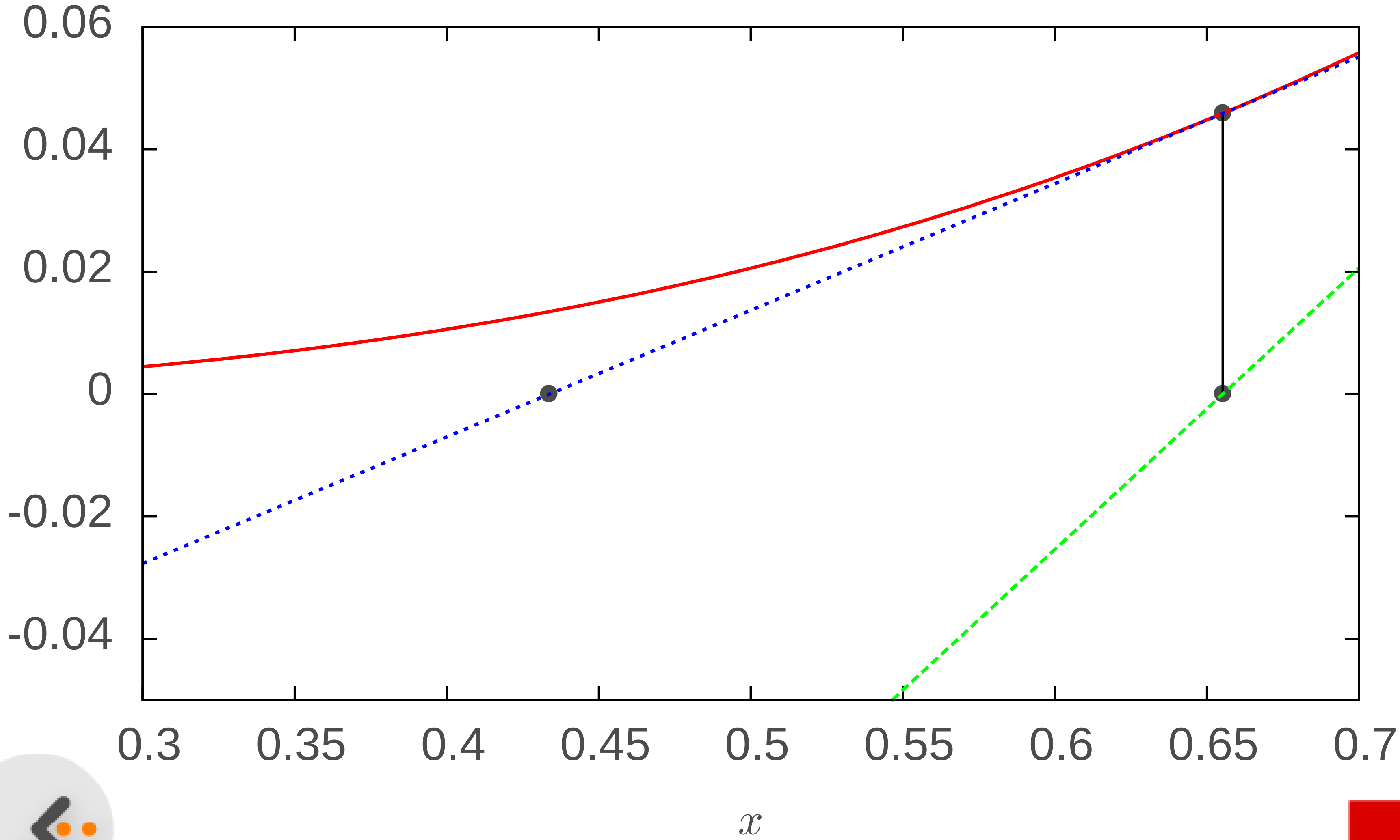
$$f(x) = x - \sin x = 0, \quad x_0 = 1, \quad \varepsilon = 10^{-15}$$

k	x_k	$f(x_k)$	$f'(x_k)$
0	+1.00000000E+00	+1.58529015E-01	+4.59697694E-01
1	+6.55145072E-01	+4.58707860E-02	+2.07040452E-01
2	+4.33590368E-01	+1.34587380E-02	+9.25368255E-02
3	+2.88148401E-01	+3.97094846E-03	+4.12282985E-02
4	+1.91832312E-01	+1.17439692E-03	+1.83434616E-02
$\vdots$			
23	+8.64386893E-05	+1.07639890E-13	+3.73582354E-09
24	+5.76257947E-05	+3.18933045E-14	+1.66036607E-09
25	+3.84171966E-05	+9.44986548E-15	+7.37940486E-10
26	+2.56114682E-05	+2.79996227E-15	+3.27973648E-10
27	+1.70743119E-05	+8.29617950E-16	+1.45766066E-10

Méthode de Newton — une variable



Méthode de Newton — une variable

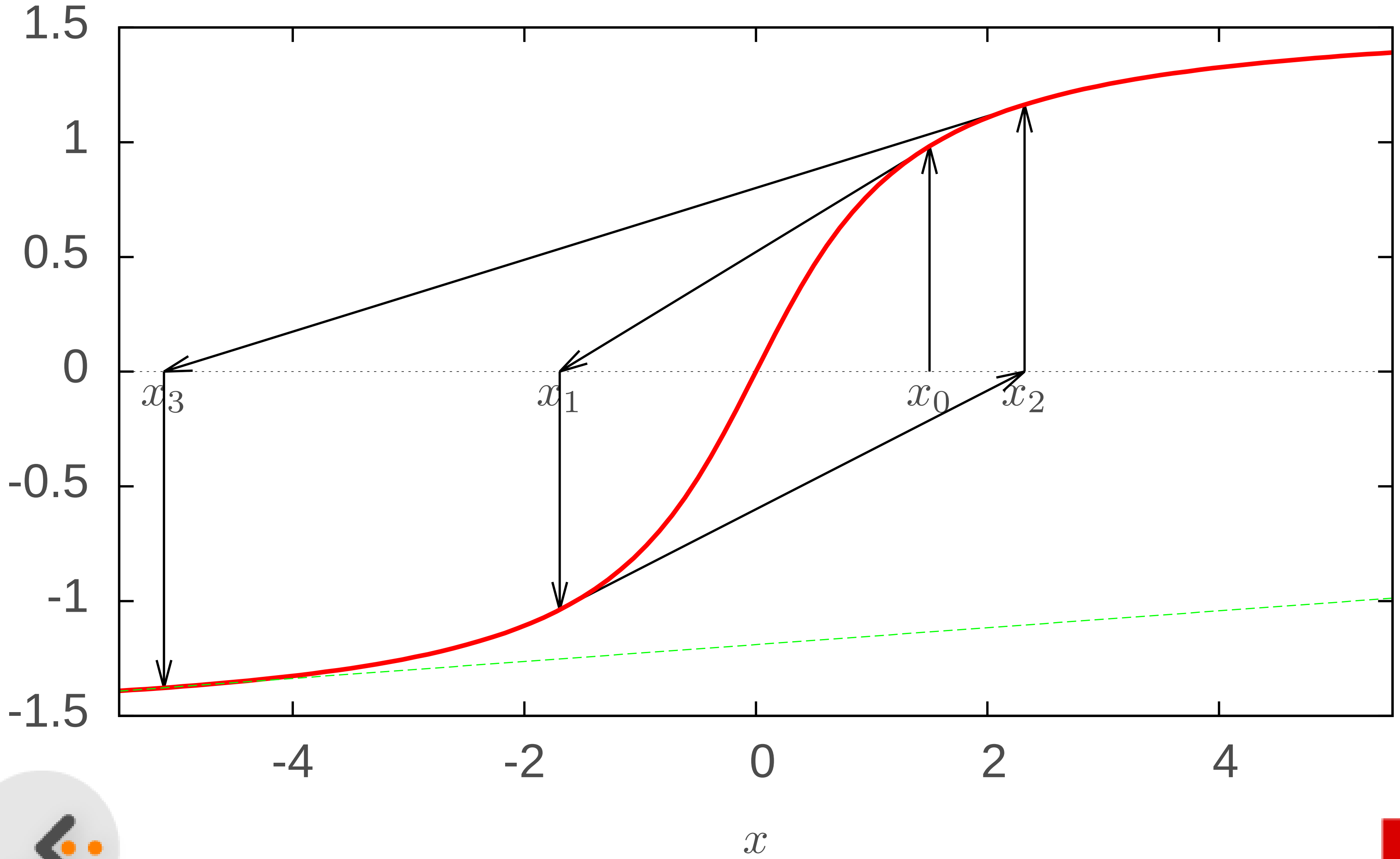


Méthode de Newton — une variable

$$f(x) = \arctan x = 0, \quad x_0 = 1.5, \quad \varepsilon = 10^{-15}$$

k	x_k	$f(x_k)$	$f'(x_k)$
0	+1.50000000E+00	+9.82793723E-01	+3.07692308E-01
1	-1.69407960E+00	-1.03754636E+00	+2.58404230E-01
2	+2.32112696E+00	+1.16400204E+00	+1.56552578E-01
3	-5.11408784E+00	-1.37769453E+00	+3.68271300E-02
4	+3.22956839E+01	+1.53984233E+00	+9.57844131E-04
5	-1.57531695E+03	-1.57016153E+00	+4.02961851E-07
6	+3.89497601E+06	+1.57079607E+00	+6.59159364E-14
7	-2.38302890E+13	-1.57079633E+00	+1.76092712E-27
8	+8.92028016E+26	+1.57079633E+00	+1.25673298E-54
9	-1.24990460E+54	-1.57079633E+00	+6.40097701E-109
10	+2.45399464E+108	+1.57079633E+00	+1.66055315E-217

Méthode de Newton — une variable



Méthode de Newton — une variable

Performance de la méthode de Newton

- Si la fonction n'est pas trop non-linéaire;
- Si la dérivée de f à la solution n'est pas trop proche de 0;
- Si x_0 n'est pas trop éloigné de la racine;
- Alors la méthode de Newton converge très vite vers la solution.

Méthode de Newton — une variable

Erreur du modèle linéaire — une variable *Soit un intervalle ouvert $X \subseteq \mathbb{R}$, et une fonction f dont la dérivée est continue au sens de Lipschitz sur X, la constante de Lipschitz étant M. Alors, pour tout $\widehat{x}$, $x^+ \in X$,*

$$|f(x^+) - m_{\widehat{x}}(x^+)| \leq M \frac{(x^+ - \widehat{x})^2}{2}.$$

(p. 196)

Rappel

Borne pour l'intégrale *Soit $f : X \subset \mathbb{R}^n \longrightarrow \mathbb{R}^{m\times n}$, où X est un ensemble ouvert convexe, et soient x et $x+d$ dans X. Alors, si f est intégrable sur $[x, x+d]$,*

$$\begin{aligned}\left\|\int_0^1 f(x+td)d\,dt\right\| &\leq \int_0^1 \|f(x+td)d\|\,dt \\ &\leq \int_0^1 \|f(x+td)\|\,\|d\|\,dt\end{aligned}$$

où $\|\cdot\|$ est une norme sur $\mathbb{R}^{m\times n}$.

(sans preuve)

Méthode de Newton — une variable

Convergence de la méthode de Newton — une variable *Soit un intervalle ouvert $X \subseteq \mathbb{R}$, et une fonction f dont la dérivée est continue au sens de Lipschitz sur X, la constante de Lipschitz étant M. Supposons qu'il existe $\rho > 0$ tel que $|f'(x)| \geq \rho \quad \forall x \in X$. Supposons qu'il existe $x^* \in X$ tel que $f(x_*) = 0$.*

()

Méthode de Newton — une variable

Convergence de la méthode de Newton — une variable (suite)

Alors, il existe $\eta > 0$ *tel que, si* $|x_0 - x^*| < \eta$ *avec* $x_0 \in X$*, la suite* $(x_k)_k$ *définie par*

$$x_{k+1} = x_k - \frac{f(x_k)}{f'(x_k)} \quad k = 0, 1, \ldots$$

est bien définie et converge vers x^**. De plus,* $|x_{k+1} - x^*| \leq \frac{M}{2\rho}|x_k - x^*|^2$*.*

(p. 197)

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Méthode de Newton — une variable

- Fonction pas trop non-linéaire M petit
- Dérivée pas trop proche de 0 $|f'(x)| \geq \rho \quad \forall x \in X.$
- x_0 n'est pas trop éloigné de la racine $|x_0 - x^*| < \eta$, avec

$$\eta = \min(r, \tau \frac{2\rho}{M}).$$

M petit $\Rightarrow \eta$ grand

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Méthode de Newton — une variable

- Convergence quadratique

$$|x_{k+1} - x^*| \leq \frac{M}{2\rho} |x_k - x^*|^2.$$

Méthode de Newton — une variable

Convergence q-quadratique

Soit une suite $(x_k)_k$ dans $\mathbb{R}^n$ convergeant vers x^*. On dit que la suite converge q-quadratiquement vers x^* s'il existe $c \geq 0$ et $\widehat{k} \in \mathbb{N}$ tels que

$$\|x_{k+1} - x^*\| \leq c\|x_k - x^*\|^2 \quad \forall k \geq \widehat{k}.$$

Méthode de Newton — n variables

Modèle linéaire d'une fonction à n variables

Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction différentiable. Le modèle linéaire *de f en $\widehat{x}$ est une fonction $m_{\widehat{x}} : \mathbb{R}^n \to \mathbb{R}^m$ définie par*

$$m_{\widehat{x}}(x) = f(\widehat{x}) + \nabla f(\widehat{x})^T(x - \widehat{x}) = f(\widehat{x}) + J(\widehat{x})(x - \widehat{x})$$

où $\nabla f(\widehat{x})$ est la matrice gradient de f en $\widehat{x}$ et $J(\widehat{x}) = \nabla f(\widehat{x})^T$ la matrice Jacobienne.

Méthode de Newton — n variables

Erreur du modèle linéaire — n variables *Soit $f : \mathbb{R}^n \to \mathbb{R}^m$ une fonction différentiable sur un ensemble ouvert convexe $X \subset \mathbb{R}^n$, dont la matrice jacobienne est continue au sens de Lipschitz sur X, où M est la constante de Lipschitz. Alors, pour tout $\widehat{x}$, $x^+ \in X$,*

$$\|f(x^+) - m_{\widehat{x}}(x^+)\| \leq M \frac{\|x^+ - \widehat{x}\|^2}{2}.$$

(p. 200)

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Algorithme : Méthode de Newton — n variables

Objectif

Trouver une approximation de la solution du système d'équations

$$f(x) = 0. \tag{1}$$

Input

- La fonction $f : \mathbb{R}^n \to \mathbb{R}^n$;
- La matrice jacobienne de la fonction $J : \mathbb{R}^n \to \mathbb{R}^{n \times n}$;
- Une première approximation de la solution $x_0 \in \mathbb{R}^n$;
- La précision demandée $\varepsilon \in \mathbb{R}$, $\varepsilon > 0$.

Output

Une approximation de la solution $x^* \in \mathbb{R}^n$

Algorithme : Méthode de Newton — n variables

Initialisation

$k = 0$

Itérations

1. Calculer d_{k+1} solution de $J(x_k)d_{k+1} = -f(x_k)$.
2. $x_{k+1} = x_k + d_{k+1}$.
3. $k = k + 1$.

Critère d'arrêt

Si $\|f(x_k)\| \leq \varepsilon$, alors $x^* = x_k$.

Méthode de Newton — n variables

$$
\begin{array}{rcl}
(x_1+1)^2 + x_2^2 & = & 2 \\
e^{x_1} + x_2^3 & = & 2.
\end{array}
$$

$$
f(x) = \begin{pmatrix} (x_1+1)^2 + x_2^2 - 2 \\ e^{x_1} + x_2^3 - 2 \end{pmatrix} \; J(x) = \begin{pmatrix} 2(x_1+1) & 2x_2 \\ e^{x_1} & 3x_2^2 \end{pmatrix}
$$

Si $x_0 = (1\ 1)^T$, nous avons

$$
f(x_0) = \begin{pmatrix} 3 \\ e-1 \end{pmatrix} \approx \begin{pmatrix} 3 \\ 1.7182 \end{pmatrix}, \quad J(x_0) = \begin{pmatrix} 4 & 2 \\ e & 3 \end{pmatrix}
$$

Méthode de Newton — n variables

k	x_k	$f(x_k)$	$\|f(x_k)\|$
0	1.00000000e+00	3.00000000e+00	3.45723768e+00
	1.00000000e+00	1.71828182e+00	
1	1.52359213e-01	7.56629795e-01	1.15470870e+00
	1.19528157e+00	8.72274931e-01	
2	-1.08376809e-02	5.19684443e-02	1.14042557e-01
	1.03611116e+00	1.01513475e-01	
3	-8.89664601e-04	1.29445248e-03	3.94232975e-03
	1.00153531e+00	3.72375572e-03	
4	-1.37008875e-06	3.13724882e-06	8.07998556e-06
	1.00000293e+00	7.44606181e-06	
5	-5.53838974e-12	1.05133679e-11	2.88316980e-11
	1.00000000e+00	2.68465250e-11	
6	-1.53209346e-16	-2.22044604e-16	2.22044604e-16
	1.00000000e+00	0.00000000e+00	

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

Méthode de Newton — n variables

Convergence de la méthode de Newton — n variables *Soit un ensemble convexe ouvert $X \subseteq \mathbb{R}^n$, et une fonction $f : X \to \mathbb{R}^n$. Supposons qu'il existe $x^* \in X$, une boule $B(x^*, r)$ centrée en x^* de rayon r, et une constante $\rho > 0$ tels que $f(x^*) = 0$, $B(x^*, r) \subset X$, $J(x^*)$ est inversible,*

$$\|J(x^*)^{-1}\| \leq \frac{1}{\rho},$$

et J est continue au sens de Lipschitz sur $B(x^*, r)$, la constante de Lipschitz étant M.

(suite...)

Méthode de Newton — n variables

Convergence de la méthode de Newton — n variables (suite)

Alors, il existe $\eta > 0$ tel que si

$$x_0 \in B(x^*, \eta), \tag{2}$$

alors la suite $(x_k)_k$ définie par

$$x_{k+1} = x_k - J(x_k)^{-1} f(x_k) \quad k = 0, 1, \ldots \tag{3}$$

est bien définie et converge vers x^*. De plus,

$$\|x_{k+1} - x^*\| \leq \frac{M}{\rho} \|x_k - x^*\|^2. \tag{4}$$

(p. 204)

TRANSP-OR

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

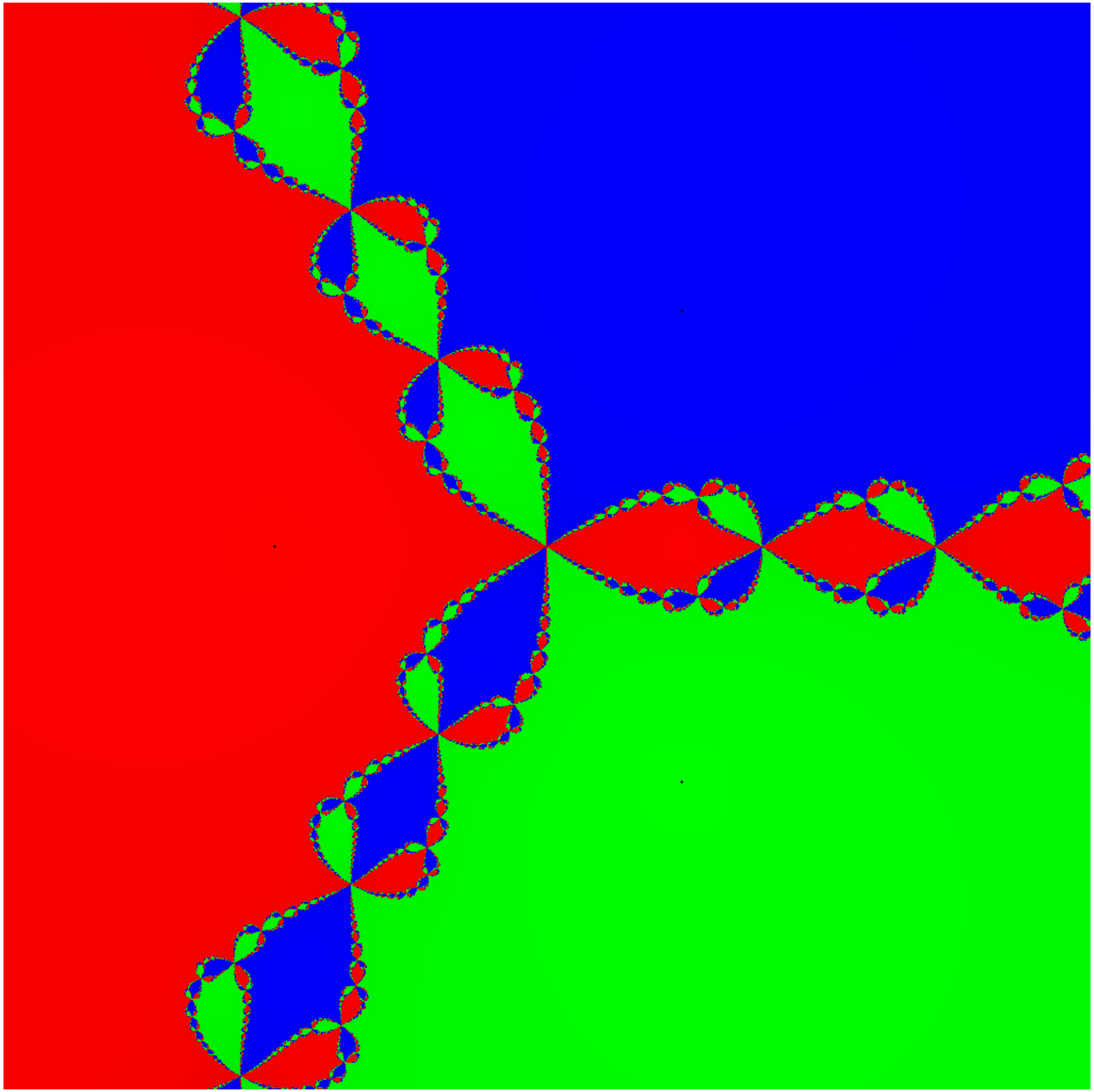
Exemple

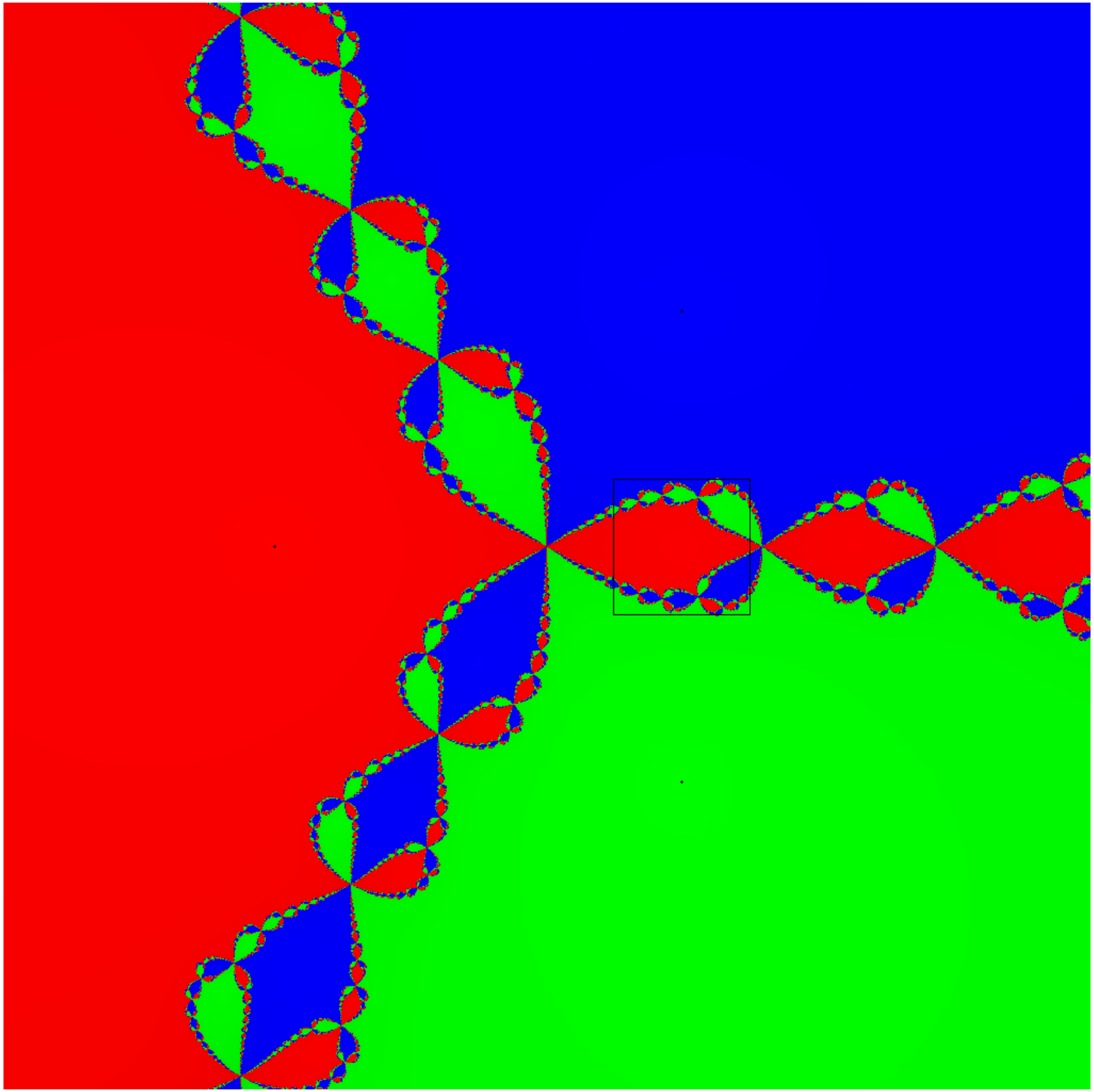
$$f(x,y) = \begin{pmatrix} x^3 - 3xy^2 - 1 \\ y^3 - 3x^2y \end{pmatrix} = 0$$

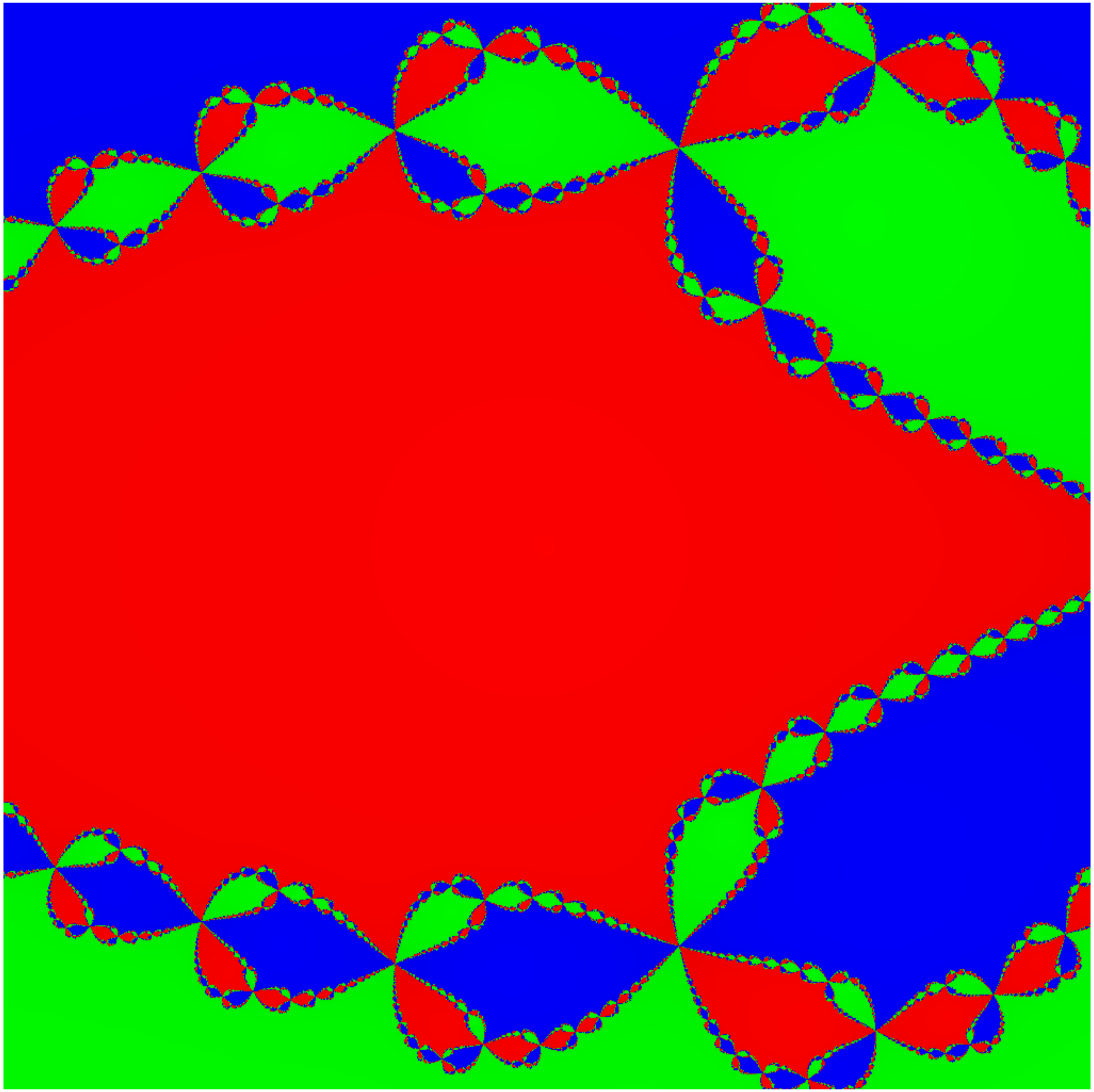
Solutions :

$$\begin{pmatrix} 1 \\ 0 \end{pmatrix} \quad \begin{pmatrix} -\frac{1}{2} \\ \frac{\sqrt{3}}{2} \end{pmatrix} \quad \begin{pmatrix} -\frac{1}{2} \\ -\frac{\sqrt{3}}{2} \end{pmatrix}$$

En fonction du point de départ, la méthode de Newton converge vers une racine ou l’autre







Méthode de Newton — n variables

Inconvénients de la méthode de Newton :

- Importance du point de départ. Solution : techniques de globalisation (seront vues dans le cadre de l'optimisation)
- Nécessite le calcul de la matrice des dérivées. Solution : méthodes quasi-Newton (pas vues dans ce cours).