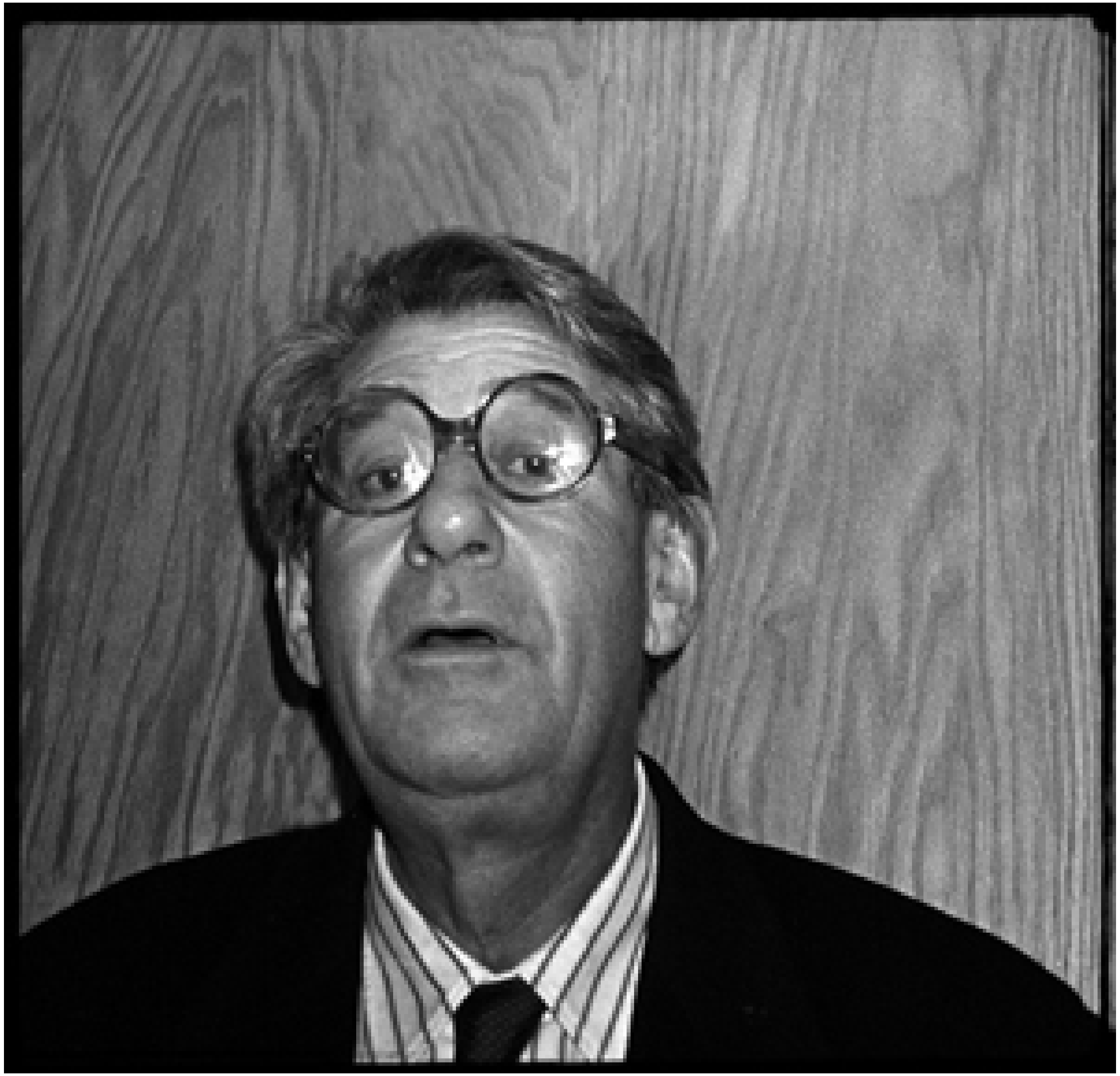
Newtonian Program Analysis

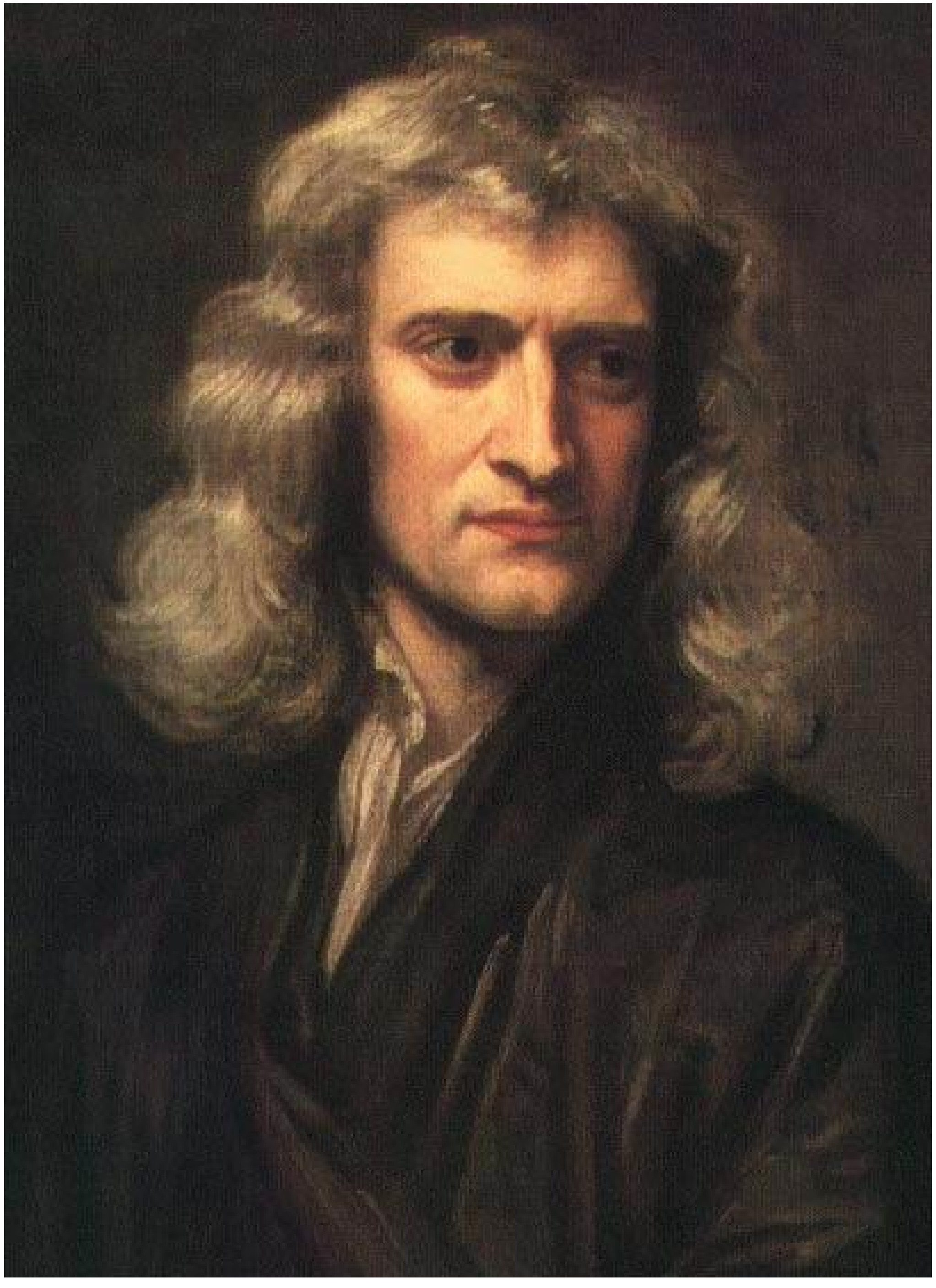
Javier Esparza

Technische Universität München

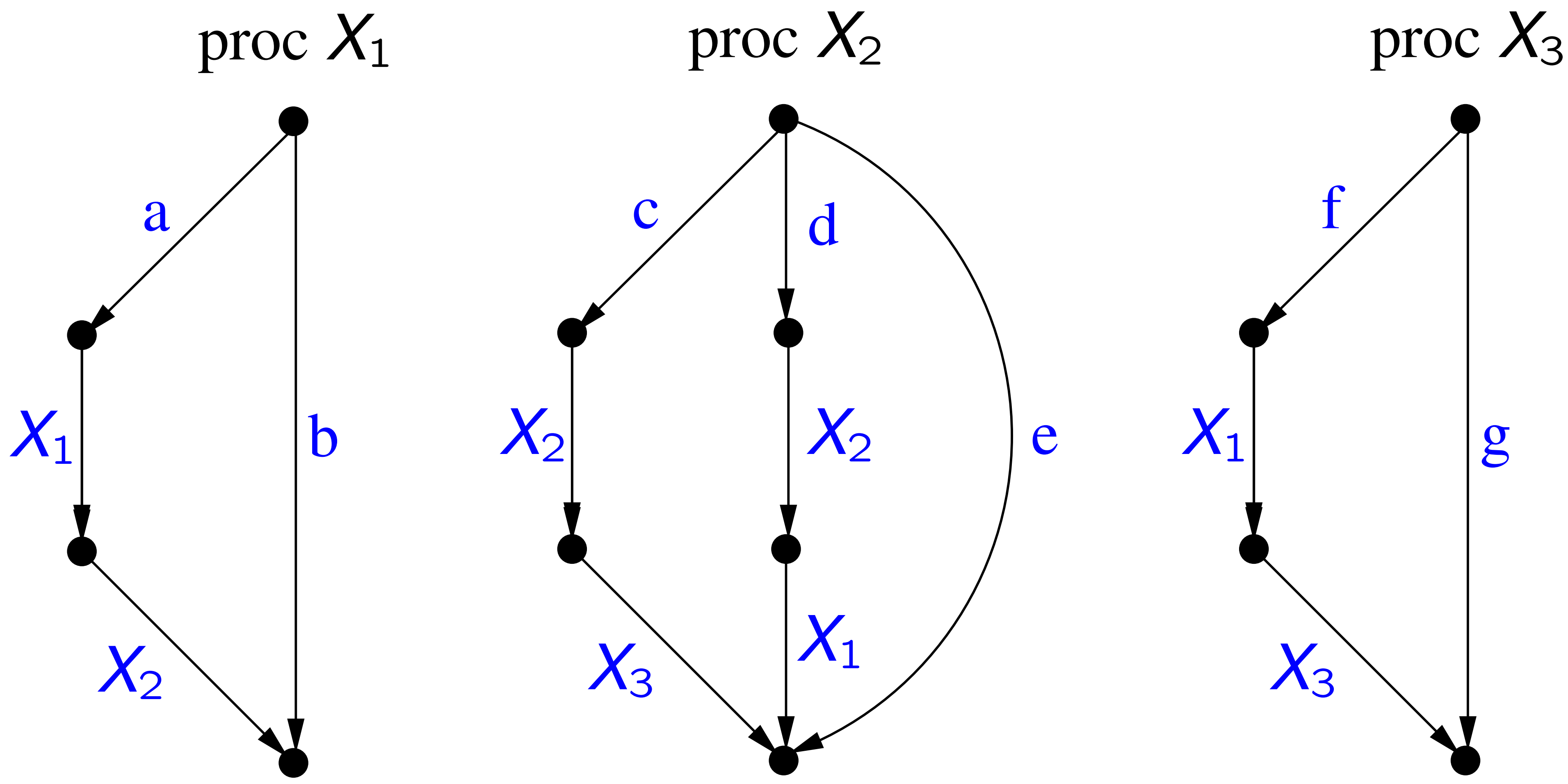
Joint work with

Stefan Kiefer and Michael Luttenberger





From programs to flowgraphs



From flowgraphs to equations

Again a syntactic transformation.

$$
\begin{aligned}
X_1 &= a \cdot X_1 \cdot X_2 + b \\
X_2 &= c \cdot X_2 \cdot X_3 + d \cdot X_2 \cdot X_1 + e \\
X_3 &= f \cdot X_1 \cdot X_3 + g
\end{aligned}
$$

But how should the equations be interpreted mathematically?

- What kind of objects are $a, \ldots, g$?
- What kind of operations are sum and product ?

From flowgraphs to equations

Again a syntactic transformation.

$$
\begin{aligned}
X_1 &= a \cdot X_1 \cdot X_2 + b \\
X_2 &= c \cdot X_2 \cdot X_3 + d \cdot X_2 \cdot X_1 + e \\
X_3 &= f \cdot X_1 \cdot X_3 + g
\end{aligned}
$$

But how should the equations be interpreted mathematically?

- What kind of objects are $a, \ldots, g$?
- What kind of operations are sum and product ?

It depends. Different interpretations lead to different semantics.

Input/output relational semantics

Interpret $a, \ldots, g$ as assignments or guards over a set of program variables V with set of valuations Val.

$R(X_i) = (v, v') \in Val \times Val$ such that X_i started at v, may terminate at v'.

Input/output relational semantics

Interpret $a, \ldots, g$ as assignments or guards over a set of program variables V with set of valuations Val.

$R(X_i) = (v, v') \in Val \times Val$ such that X_i started at v, may terminate at v'.

$(\; R(X_1), R(X_2), R(X_3) \;)$ is the least solution of the equations under the following interpretation:

- Universe: $2^{V \times V}$ (input/output relations)
- $a, \ldots, g$ are relations for assignment/guards
- sum is union of relations, product is join of relations:

$$R_1 \cdot R_2 = \{(a, b) \mid \exists c \; (a, c) \in R_1 \wedge (c, b) \in R_2\}$$

Language semantics

Interpret the atomic actions as letters of an alphabet A.

$L(X_i) =$ words $w \in A^*$ such that X_i can execute w and terminate.

Language semantics

Interpret the atomic actions as letters of an alphabet A.

$L(X_i) =$ words $w \in A^*$ such that X_i can execute w and terminate.

$(L(X_1), L(X_2), L(X_3))$ is the least solution of the equations under the following interpretation:

- Universe: 2^{A^*} (languages over A).
- $a, \ldots, g$ are the singleton languages $\{a\}, \ldots, \{g\}$.
- sum is union of languages, product is concatenation:

$$L_1 \cdot L_2 = \{w_1 w_2 \mid w_1 \in L_1 \wedge w_2 \in l_2\}$$

.

Counting semantics

Given a word w, denote by $\#(w)$ the vector saying how many times each of $a, \ldots, g$ occurs in w .

Define $Co(X_i) = \{\#(w) \mid w \in L(X_i)\}$.

Counting semantics

Given a word w, denote by $\#(w)$ the vector saying how many times each of a, . . . , g occurs in w.

Define $Co(X_i) = \{\#(w) \mid w \in L(X_i)\}$.

$(Co(X_1), Co(X_2), Co(X_3))$ is the least solution of the equations under the following interpretation:

- Universe: sets of vectors of naturals
- $a, \ldots, g$ are the singleton sets $\{(1, 0, \ldots, 0)\}, \ldots, \{(0, 0, \ldots, 1)\}$
- sum is union of sets, product is given by

$$S_1 \cdot S_2 = \{v_1 +_{\mathbb{R}} v_2 \mid v_1 \in S_1, v_2 \in S_2\}$$

Probabilistic termination semantics

Interpret $a, \ldots, g$ as probabilities.

$T(X_i) =$ probability that X_i terminates.

Probabilistic termination semantics

Interpret $a, \ldots, g$ as probabilities.

$T(X_i) =$ probability that X_i terminates.

$(\ T(X_1), T(X_2), T(X_3)\ )$ is the least solution of the equations under the following interpretation:

- Universe: $\mathbb{R}^+$
- $a, \ldots, g$ are the probabilities of taking the transitions
- sum and product are addition and multiplication of reals

Abstract interpretation [Cousot, Cousot 77] determines an interpretation given

- its universe, and
- its relation to a reference semantics (the concrete semantics).

ω-continuous semirings

Underlying mathematical structure: ω-continuous semirings

Algebra $(C, +, \cdot, 0, 1)$

– $(C, +, 0)$ is a commutative monoid

– $(C, \cdot, 1)$ is a monoid

– $a \sqsubseteq a + b$ is a partial order

– $\cdot$ distributes over $+$

– $0 \cdot a = a \cdot 0 = 0$

– $\sqsubseteq$-chains have limits

System of equations $X = f(X)$ where

- $X = (X_1, \ldots, X_n)$ vector of variables,
- $f(X) = (f_1(X), \ldots, f_n(X))$ vector of terms over $C \cup \{X_1, \ldots, X_n\}$.

Notice: the f_i are polynomials!!

Static program analysis

Static program analysis = computing the least solution of a system of polynomial equations over a suitable ω-continuous semiring

Program	$\Longrightarrow$	system of equations
Analysis problem	$\Longrightarrow$	concrete semiring
Algorithmic solution	$\Longrightarrow$	equation solver
Theory of static analysis	$\Longrightarrow$	generic solution techniques

In this talk: generic solution techniques and some consequences.

Kleenean program analysis

Theorem [Kleene]: The least solution μf is the supremum of $\{k_i\}_{i \geq 0}$, where

$$\begin{aligned} k_0 &= f(0) \\ k_{i+1} &= f(k_i) \end{aligned}$$

Basic algorithm: compute $k_0, k_1, k_2, \ldots$ until either $k_i = k_{i+1}$ or the approximation is considered adequate.

Current state-of-the-art:

- sufficient condition for termination: finite ascending chains
- if condition does not hold: widening and narrowing.

Kleenean program analysis is slow

Set interpretations: Kleene iteration never terminates if μf is an infinite set.

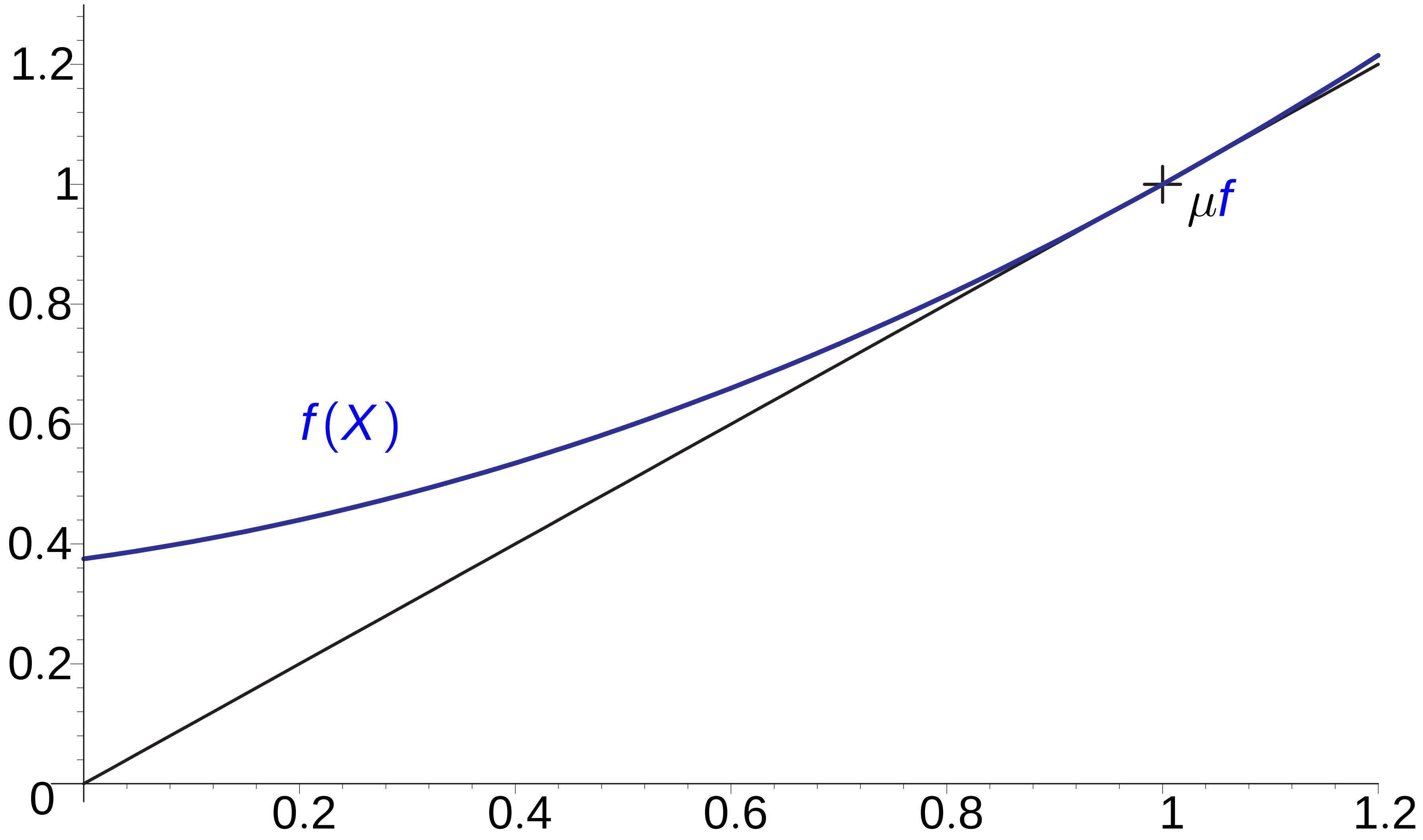
- $X = a \cdot X + b \qquad \mu f = a^* b$
- Kleene approximants are finite sets: $k_i = (\epsilon + a + \ldots + a^i)b$

Probabilistic interpretation: convergence can be very slow [EY STACS05].

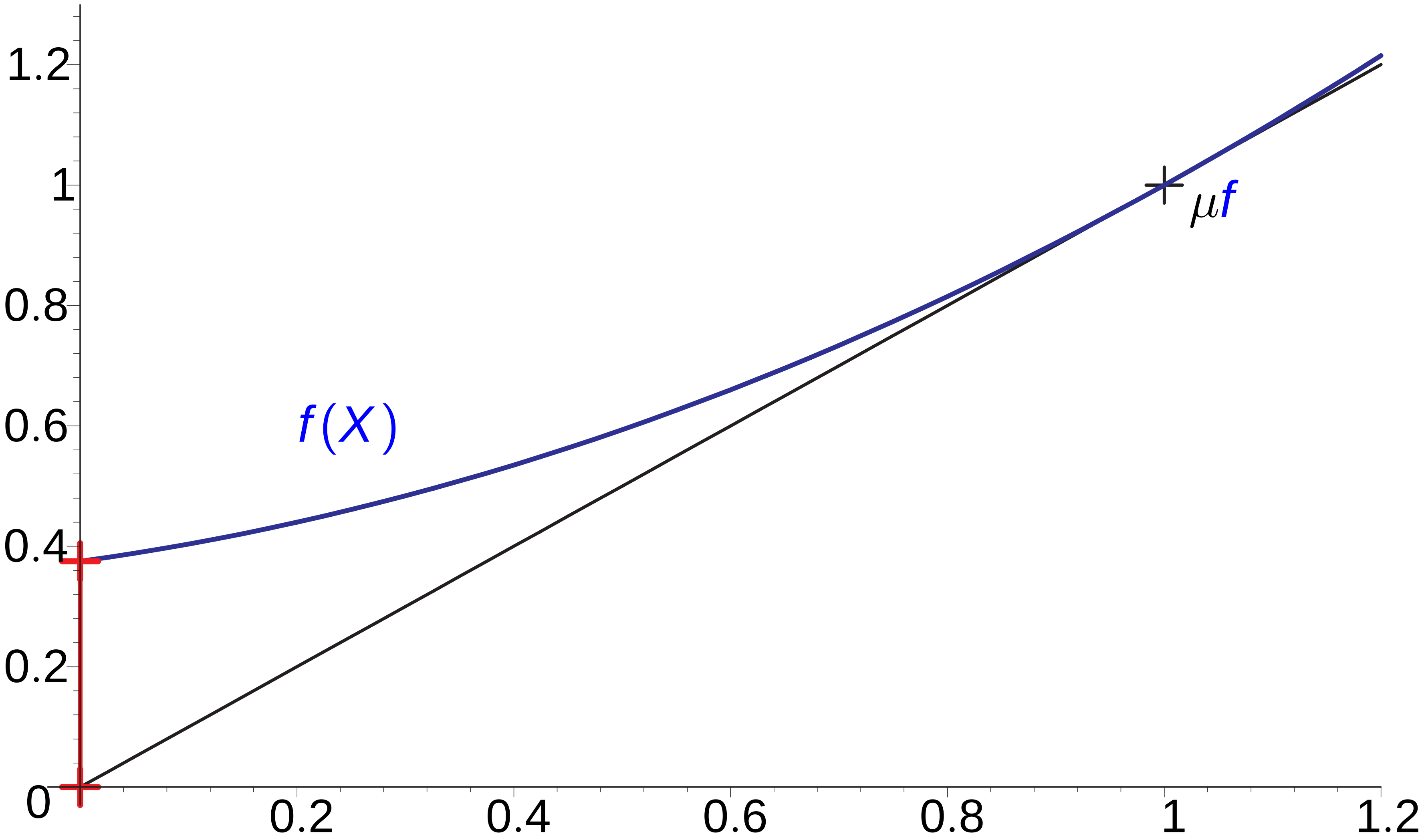
- $X = \frac{1}{2} X^2 + \frac{1}{2} \qquad \mu f = 1 = 0.99999\ldots$
- "Logarithmic convergence": k iterations to get $\log k$ bits of accuracy.

$$k_n \leq 1 - \frac{1}{n+1} \qquad k_{2000} = 0.9990$$

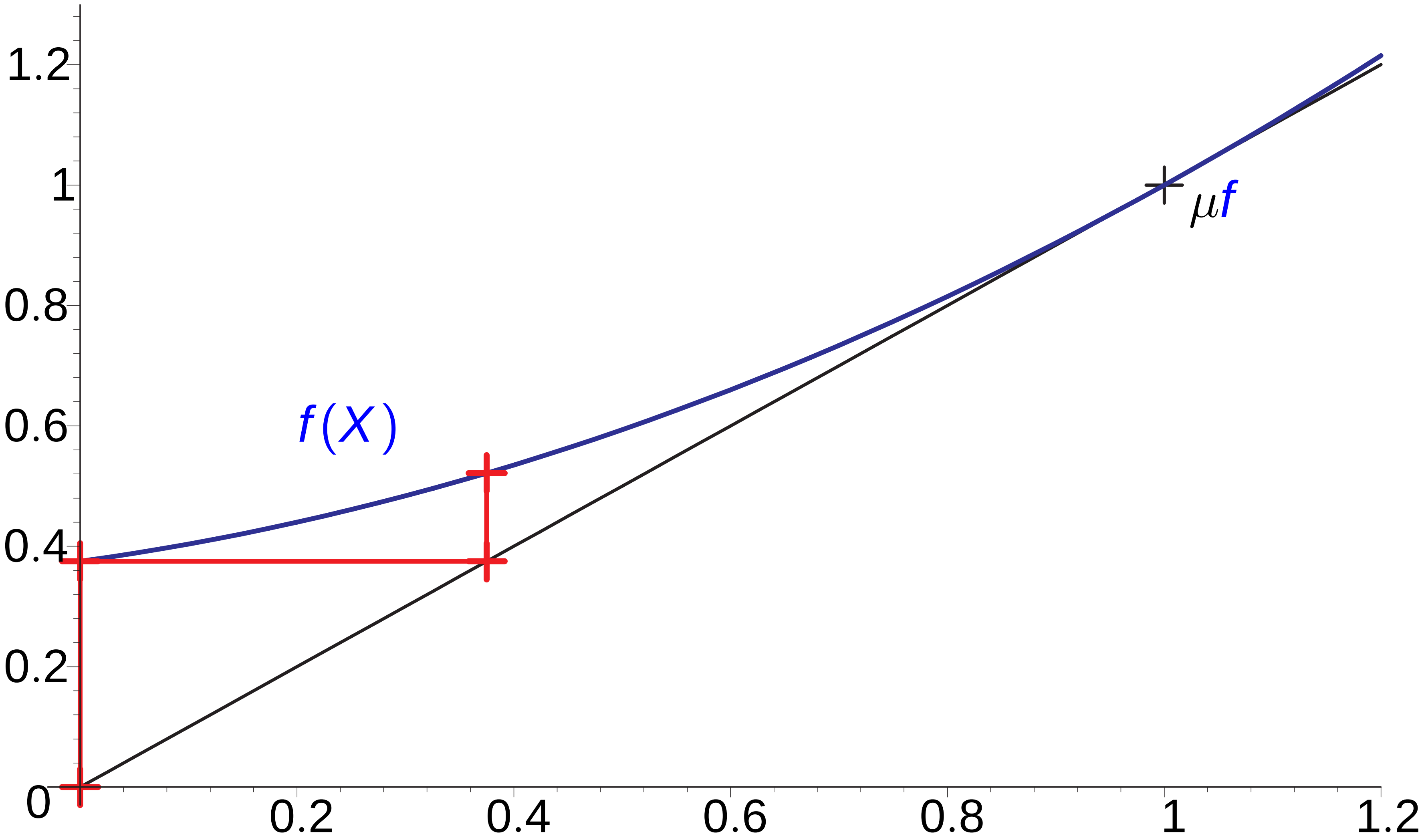
Kleene Iteration for $X = f(X)$ (univariate case)



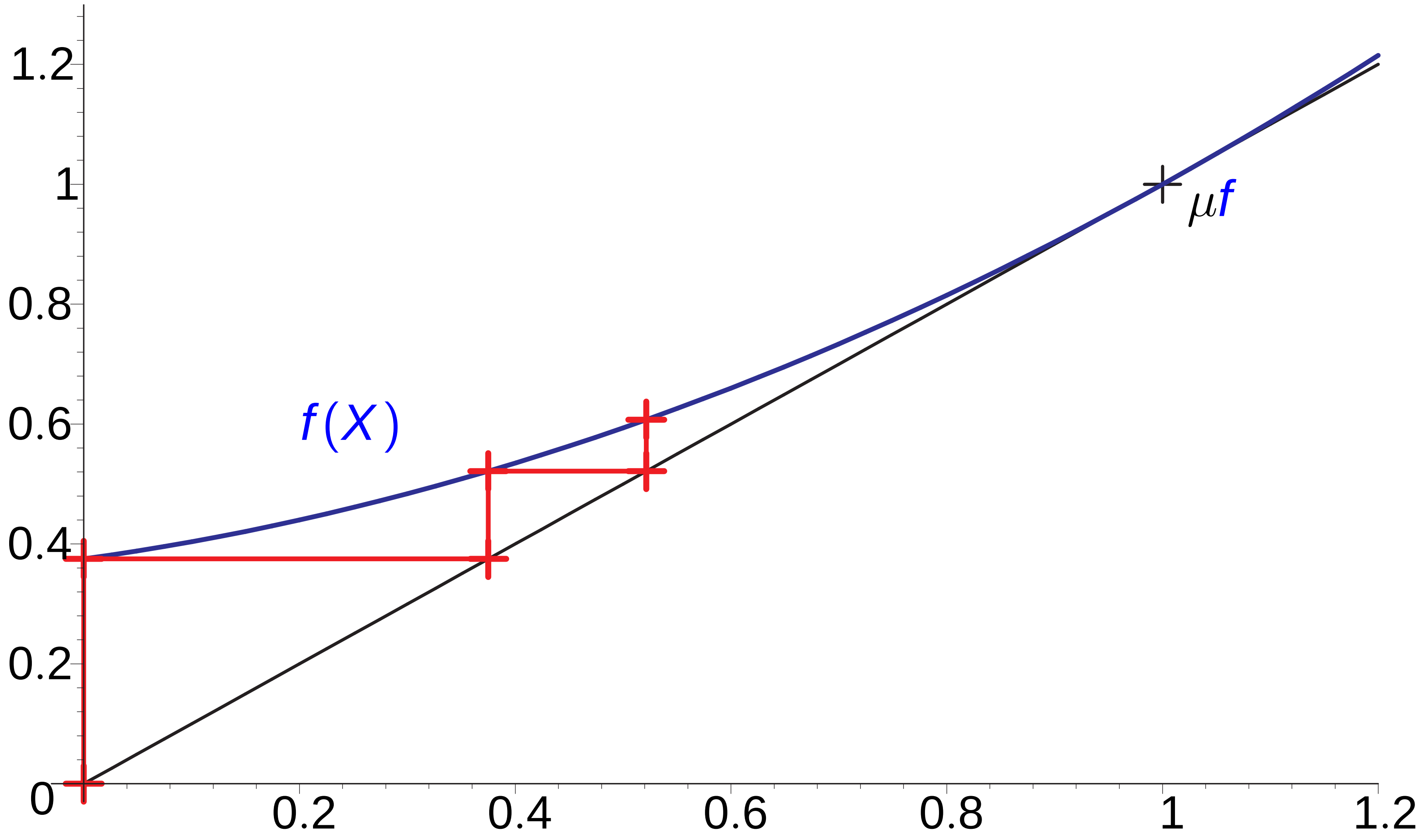
Kleene Iteration for $X = f(X)$ (univariate case)



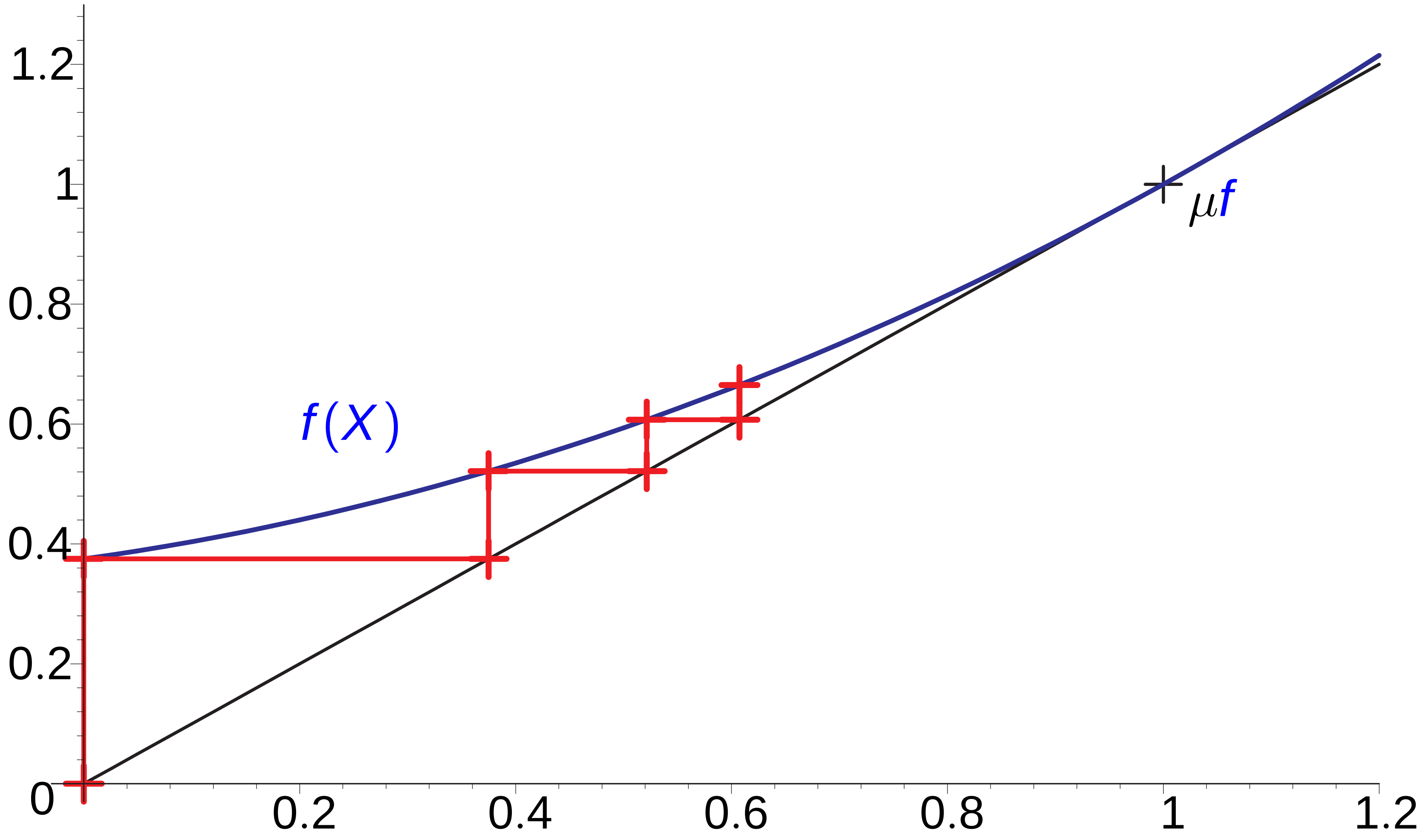
Kleene Iteration for $X = f(X)$ (univariate case)



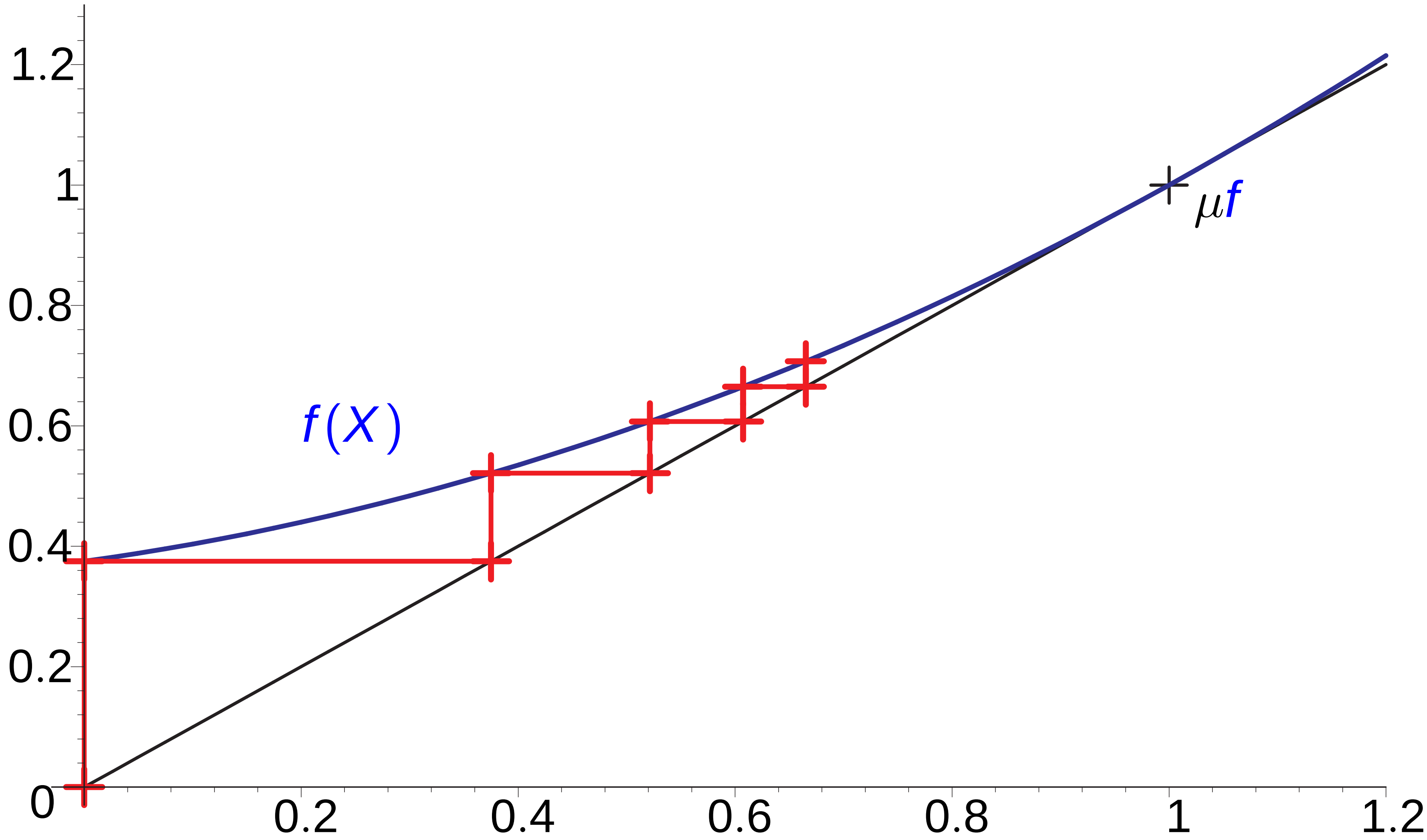
Kleene Iteration for $X = f(X)$ (univariate case)



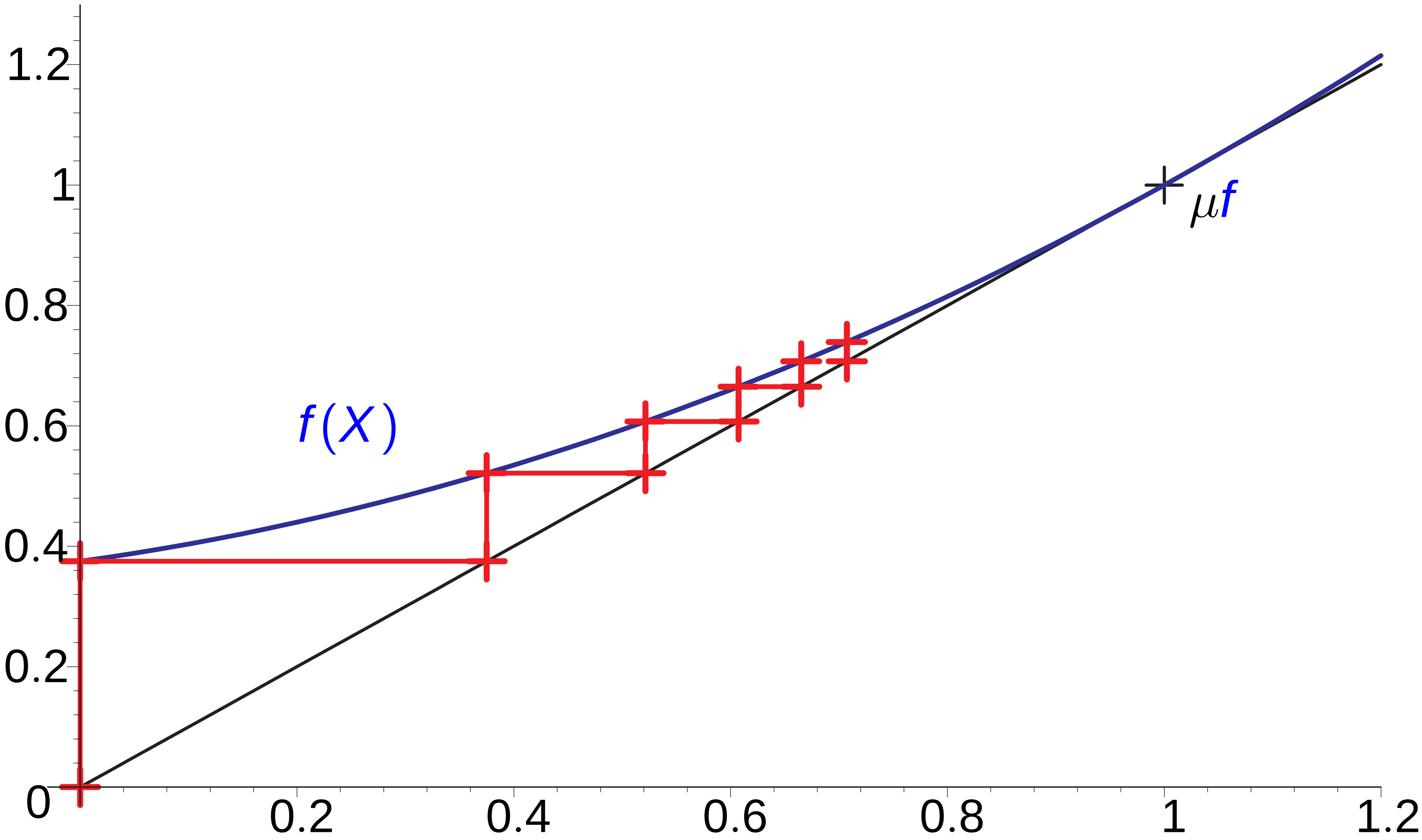
Kleene Iteration for $X = f(X)$ (univariate case)



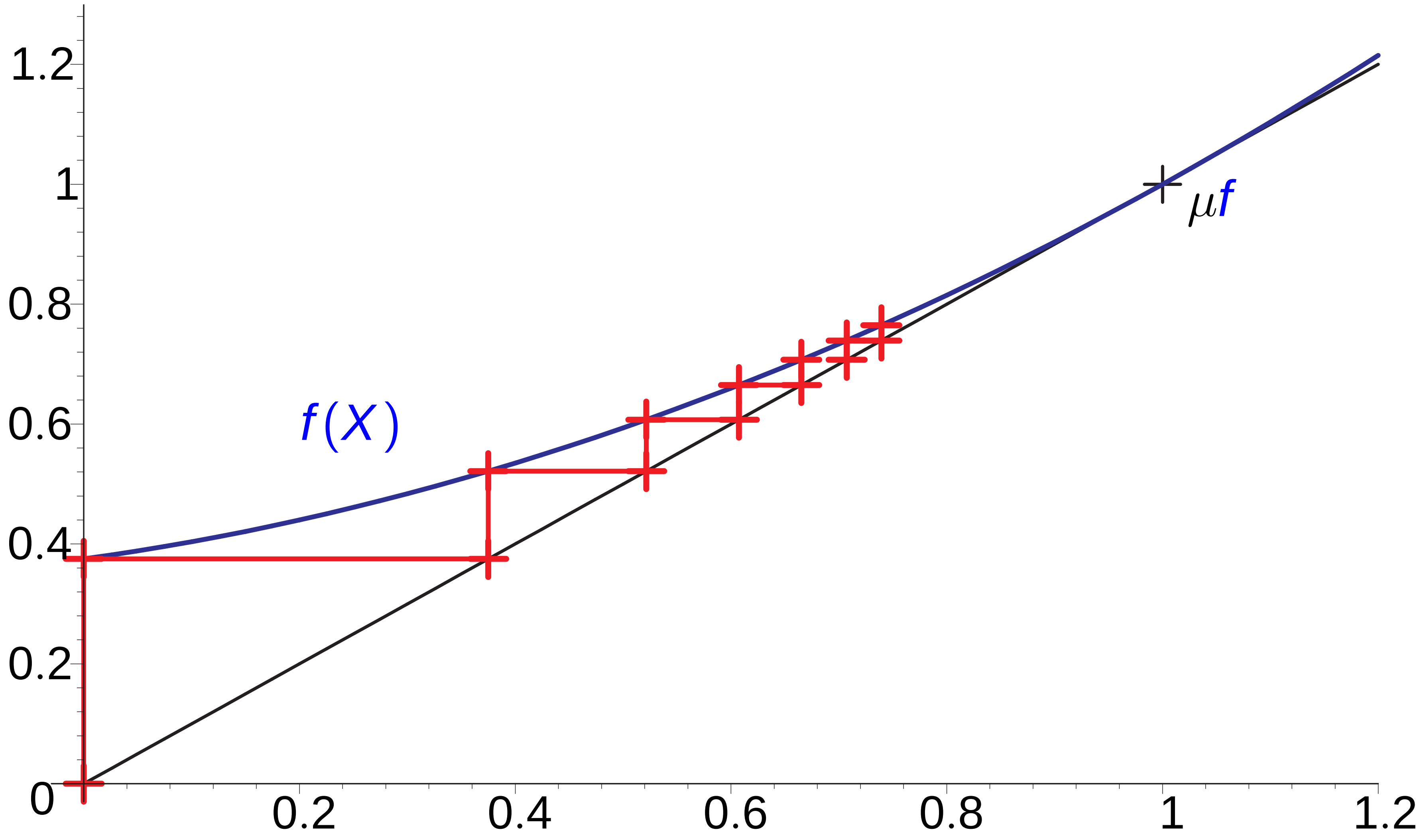
Kleene Iteration for $X = f(X)$ (univariate case)



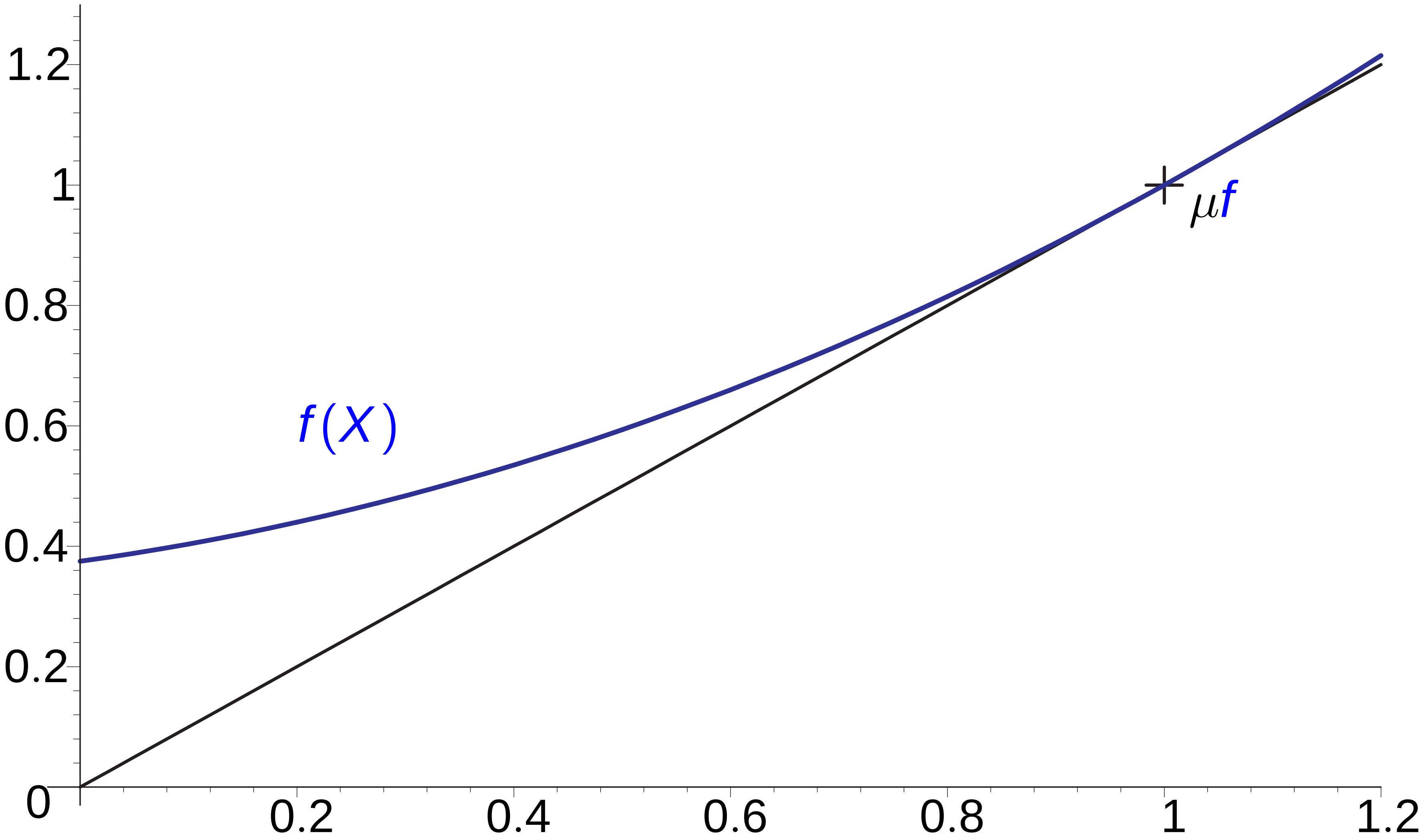
Kleene Iteration for $X = f(X)$ (univariate case)



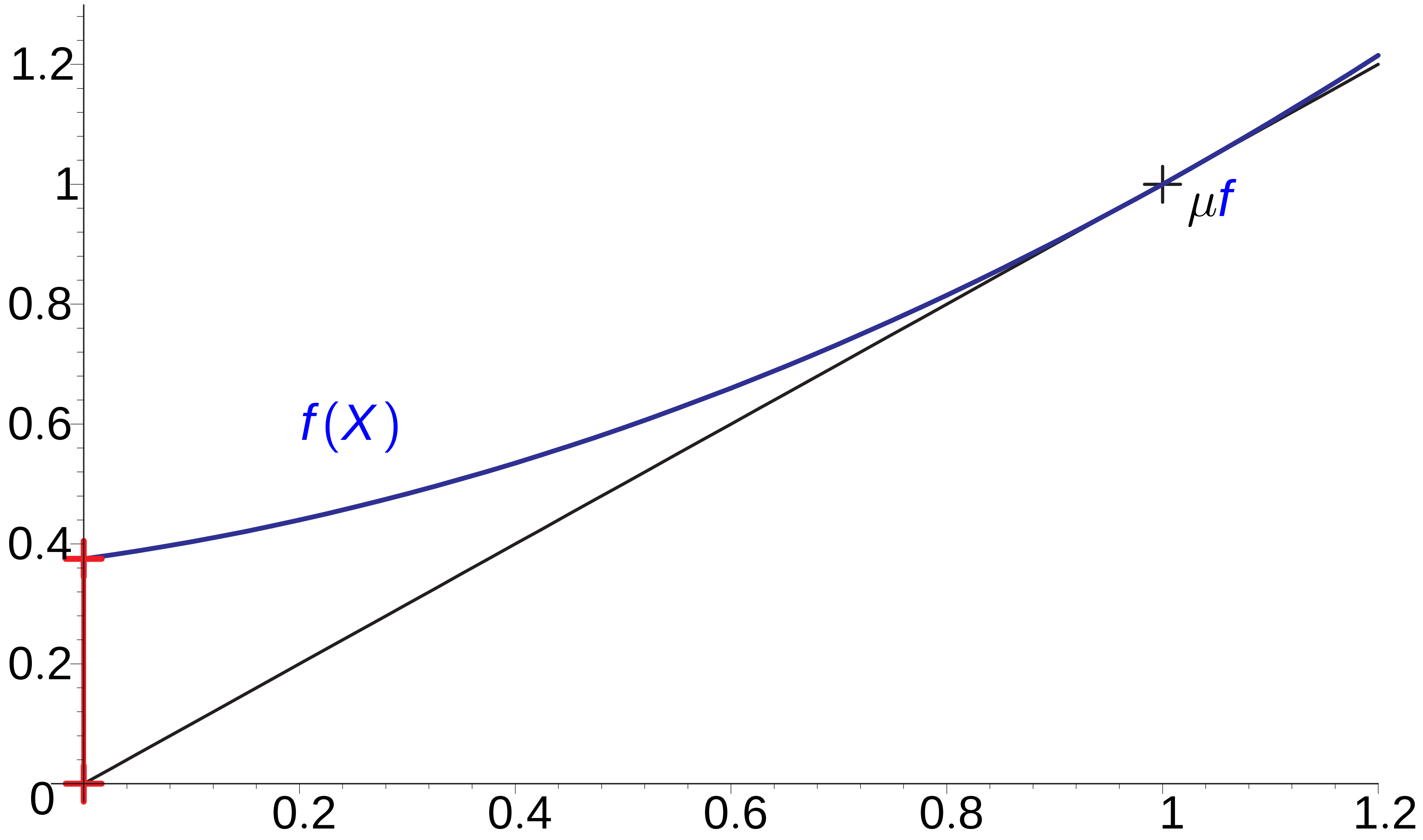
Kleene Iteration for $X = f(X)$ (univariate case)



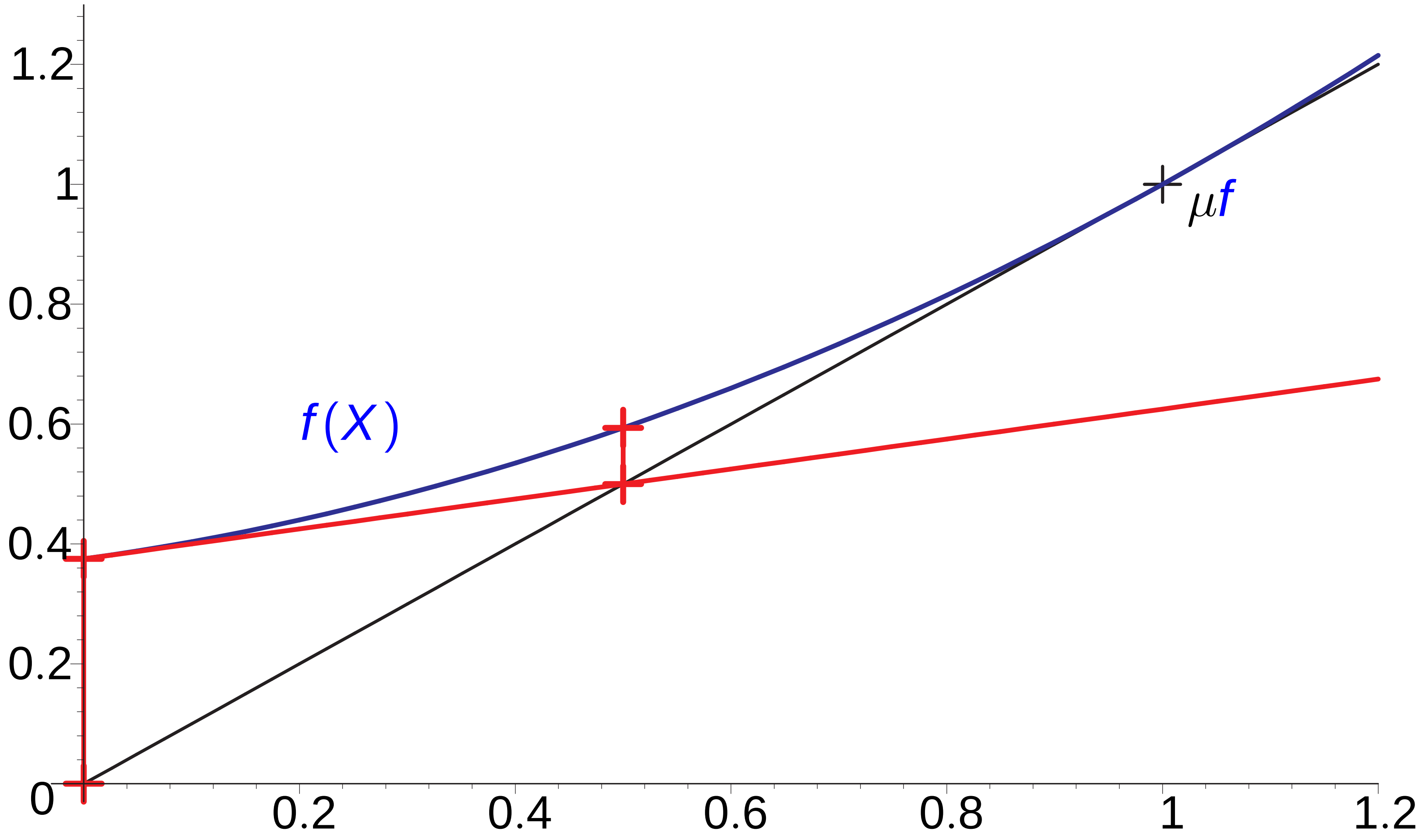
Newton's Method for $X = f(X)$ (univariate case)



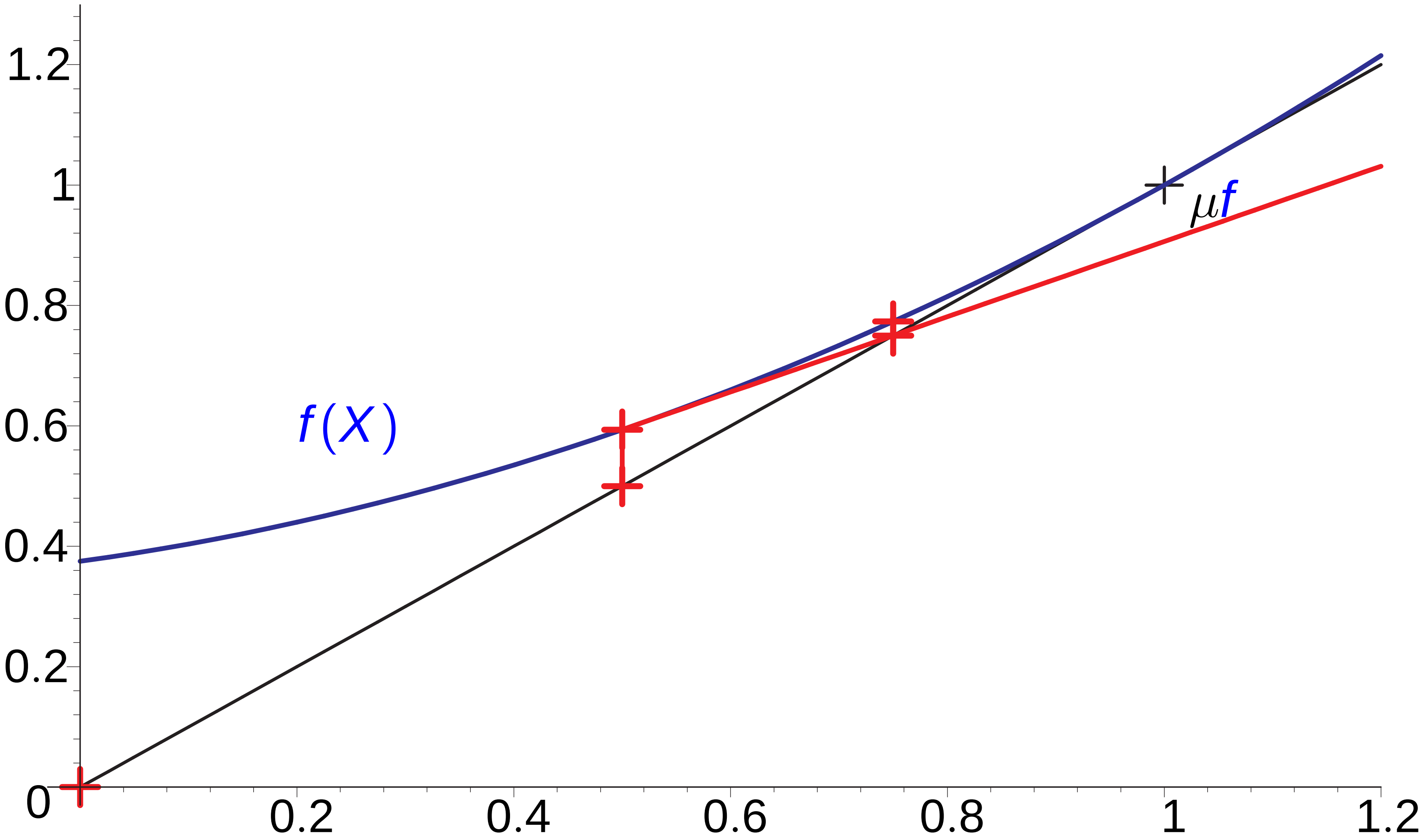
Newton's Method for $X = f(X)$ (univariate case)



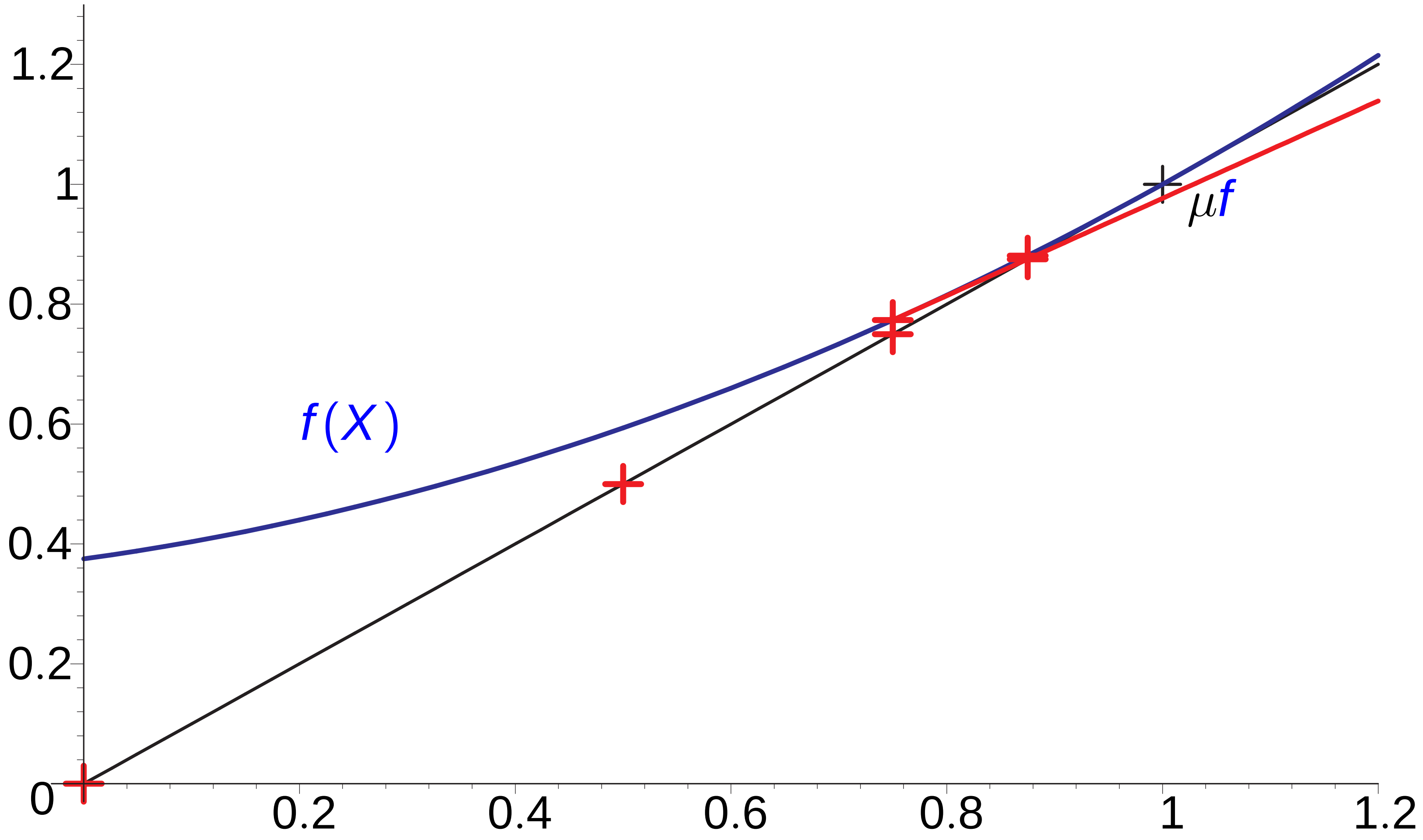
Newton's Method for $X = f(X)$ (univariate case)



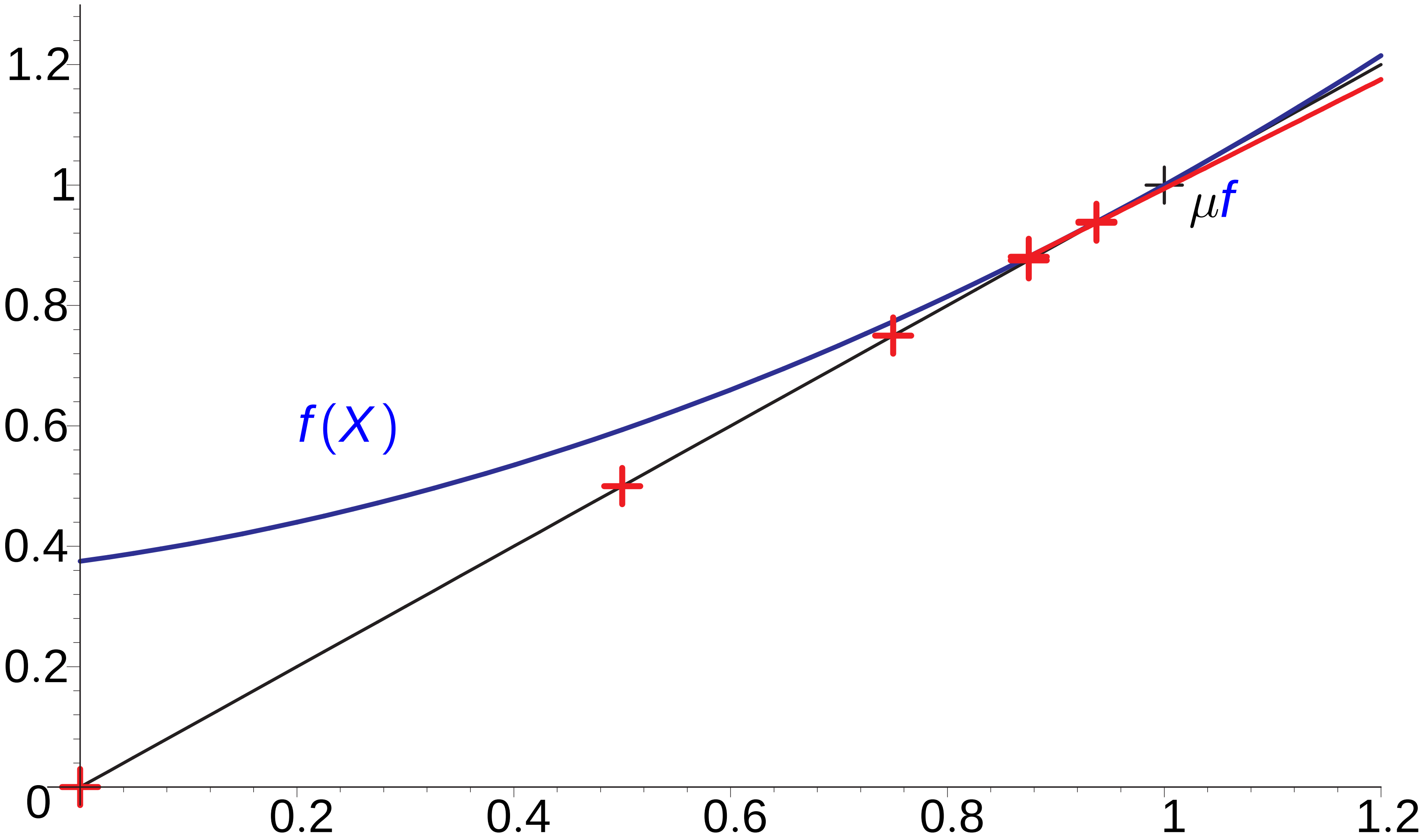
Newton's Method for $X = f(X)$ (univariate case)



Newton's Method for $X = f(X)$ (univariate case)



Newton's Method for $X = f(X)$ (univariate case)



Evaluation of Newton's method

Newton's Method is usually very efficient

- often exponential convergence

...but not robust:

- may not converge, or
- may converge only locally (in some neighborhood of the least fixed-point), or
- may converge very slowly.

A puzzling mismatch

Program analysis:

- General domain: arbitrary ω-continuous semirings
- Kleene Iteration is robust and generally applicable
- . . . but converges slowly.

Numerical mathematics:

- Particular domain: the real field
- Newton's Method converges fast
- . . . but is not robust

Two questions

- Can Newton's Method be generalized to arbitrary ω-continuous semirings?

- Is Newton's method robust when restricted to the real semiring?

Mathematical formulation of Newton's Method

Let ν be some approximation of μf. (We start with $\nu = f(0)$.)

- Compute the function $T_\nu(X)$ describing the tangent to $f(X)$ at ν
- Solve $X = T_\nu(X)$ (instead of $X = f(X)$), and take the solution as the new approximation

Elementary analysis: $T_\nu(X) = Df_\nu(X) + f(\nu) - \nu$

where $Df_{x_0}(X)$ is the differential of f at x_0

So:
$$\begin{aligned} \nu_0 &= 0 \\ \nu_{i+1} &= \nu_i + \Delta_i \qquad \Delta_i \text{ solution of } X = Df_{\nu_i}(X) + f(\nu_i) - \nu_i \end{aligned}$$

Generalizing Newton's method

Key point: generalize $X = Df_\nu(X) + f(\nu) - \nu$

In an arbitrary ω-continuous semiring

- neither the differential $Df_\nu(X)$, nor
- the difference $f(\nu) - \nu$

are defined.

Differentials in semirings

Standard solution: take the algebraic definition

$$Df(X) = \begin{cases} 0 & \text{if } f(X) = c \\ X & \text{if } f(X) = X \\ Dg(X) + Dh(X) & \text{if } f(X) = g(X) + h(X) \\ Dg(X) \cdot h(X) + g(X) \cdot Dh(X) & \text{if } f(X) = g(X) \cdot h(X) \\ \sum_{i \in I} Df(X) & \text{if } f(X) = \sum_{i \in I} f_i(X). \end{cases}$$

The difference $f(\nu_i) - \nu_i$

Solution: Replace $f(\nu_i) - \nu_i$ by any δ_i such that $f(\nu_i) = \nu_i + \delta_i$

$$\nu_{i+1} = \nu_i + \Delta_i \quad \text{where } \Delta_i \text{ solution of } \quad X = Df_{\nu_i}(X) + \delta_i$$

The difference $f(\nu_i) - \nu_i$

Solution: Replace $f(\nu_i) - \nu_i$ by any δ_i such that $f(\nu_i) = \nu_i + \delta_i$

$$\nu_{i+1} = \nu_i + \Delta_i \quad \text{where } \Delta_i \text{ solution of } \quad X = Df_{\nu_i}(X) + \delta_i$$

But does δ_i always exist?

The difference $f(\nu_i) - \nu_i$

Solution: Replace $f(\nu_i) - \nu_i$ by any δ_i such that $f(\nu_i) = \nu_i + \delta_i$

$$\nu_{i+1} = \nu_i + \Delta_i \quad \text{where } \Delta_i \text{ solution of } \quad X = Df_{\nu_i}(X) + \delta_i$$

But does δ_i always exist? Proposition: Yes

The difference $f(\nu_i) - \nu_i$

Solution: Replace $f(\nu_i) - \nu_i$ by any δ_i such that $f(\nu_i) = \nu_i + \delta_i$

$$\nu_{i+1} = \nu_i + \Delta_i \quad \text{where } \Delta_i \text{ solution of } \quad X = Df_{\nu_i}(X) + \delta_i$$

But does δ_i always exist? Proposition: Yes

But ν_{i+i} depends on your choice of δ_i !

The difference $f(\nu_i) - \nu_i$

Solution: Replace $f(\nu_i) - \nu_i$ by any δ_i such that $f(\nu_i) = \nu_i + \delta_i$

$\nu_{i+1} = \nu_i + \Delta_i$ where Δ_i solution of $X = Df_{\nu_i}(X) + \delta_i$

But does δ_i always exist? Proposition: Yes

But ν_{i+i} depends on your choice of δ_i ! Theorem: No, it doesn't

The difference $f(\nu_i) - \nu_i$

Solution: Replace $f(\nu_i) - \nu_i$ by any δ_i such that $f(\nu_i) = \nu_i + \delta_i$

$\nu_{i+1} = \nu_i + \Delta_i$ where Δ_i solution of $X = Df_{\nu_i}(X) + \delta_i$

But does δ_i always exist? Proposition: Yes

But ν_{i+i} depends on your choice of δ_i ! Theorem: No, it doesn't

Can't you give a closed form for ν_{i+1} ?

The difference $f(\nu_i) - \nu_i$

Solution: Replace $f(\nu_i) - \nu_i$ by any δ_i such that $f(\nu_i) = \nu_i + \delta_i$

$$\nu_{i+1} = \nu_i + \Delta_i \quad \text{where } \Delta_i \text{ solution of } \quad X = Df_{\nu_i}(X) + \delta_i$$

But does δ_i always exist? Proposition: Yes

But ν_{i+i} depends on your choice of δ_i ! Theorem: No, it doesn't

Can't you give a closed form for ν_{i+1} ? Proposition: Yes

The difference $f(\nu_i) - \nu_i$

Solution: Replace $f(\nu_i) - \nu_i$ by any δ_i such that $f(\nu_i) = \nu_i + \delta_i$

$$\nu_{i+1} = \nu_i + \Delta_i \quad \text{where } \Delta_i \text{ solution of } \quad X = Df_{\nu_i}(X) + \delta_i$$

But does δ_i always exist? Proposition: Yes

But ν_{i+i} depends on your choice of δ_i ! Theorem: No, it doesn't

Can't you give a closed form for ν_{i+1} ? Proposition: Yes

The least solution of $X = Df_{\nu_i}(X) + \delta_i$ is $Df^*_{\nu_i}(\delta_i) := \sum_{j=0}^{\infty} Df^j_{\nu_i}(\delta_i)$

and so: $\nu_{i+1} = \nu_i + Df^*_{\nu_i}(\delta_i)$

Theorem [EKL DLT07]: Let $X = f(X)$ be an equation over an arbitrary ω-continuous semiring. The sequence

$$\begin{aligned} \nu_0 &= f(0) \\ \nu_{i+1} &= \nu_i + Df^*_{\nu_i}(\delta_i) \end{aligned}$$

where δ_i satisfies $f(\nu_i) = \nu_i + \delta_i$ exists, is unique and satisfies

$$k_i \sqsubseteq \nu_i \sqsubseteq \mu f$$

for every $i \geq 0$.

Extensions and simplifications

Systems of equations:

- $\nu_i, \ \Delta_i, \ \delta_i$ become vectors (elements of S^n)
- The differential becomes a function $S^n \to S^n$
  Geometric intuition: $Df_{\nu_i}(X_1, \ldots, X_n)$ is the hyperplane tangent to f at the (n-dimensional) point ν_i

Commutative semirings (and left-linear equations):

- One variable: $Df_\nu(X) = f'(\nu) \cdot X$, and so $\nu_{i+1} = \nu_i + f'^*(\nu_i) \cdot \delta_i$
- Many variables: $Df_\nu(X) = J(\nu) \cdot X$, where $J(\nu)$ is the Jacobi matrix of partial derivatives evaluated at ν, and so $\nu_{i+1} = \nu_i + J^*(\nu_i) \cdot \delta_i$

Newton's method for language equations

Language semiring: Universe is 2^{A^*}, $+$ is union, $\cdot$ is concatenation.

For left-linear systems of equations, Newton's method terminates after 1 iteration:

$$\begin{aligned} X_1 &= a \cdot X_1 + b \cdot X_2 \\ X_2 &= a \cdot X_1 + b \cdot X_2 + 1 \end{aligned}$$

$$\begin{aligned} \nu_0 &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} \\ \nu_1 &= \begin{pmatrix} 0 \\ 1 \end{pmatrix} + \begin{pmatrix} a & b \\ a & b \end{pmatrix}^* \cdot \begin{pmatrix} 0 \\ 1 \end{pmatrix} = \begin{pmatrix} (a + bb^*a)^* & (a + bba^*)^* b \\ (b + aa^*b)^* a & (b + aa^*b)^* \end{pmatrix} \\ &= \begin{pmatrix} (a^* + bba^*)^* b \\ (b^* + aa^*b)^* \end{pmatrix} \end{aligned}$$

A nonlinear equation

$$\begin{aligned} X &= a \cdot X \cdot X + b \\ f(X) &= a \cdot X \cdot X + b \\ Df_\nu(X) &= a \cdot \nu \cdot X + a \cdot X \cdot \nu \end{aligned}$$

$$\begin{aligned} \nu_0 &= b \qquad \nu_0 + \delta_0 = f(\nu_0) \Longrightarrow \delta_0 := abb \\ \nu_1 &= \nu_0 + Df_b^*(\delta_0) = b + Df_b^*(abb) \\ &= b + (X + abX + aXb + abaXb + \ldots)(abb) \\ &= b + abb + ababb + aabbb + abaabbb + \ldots \\ \nu_2 &= \cdots \end{aligned}$$

The method does not terminate. Can we characterize the approximants?

Finite-index approximations

System $X = f(X)$ induces context-free grammar $G \stackrel{def}{=} X \rightarrow f(X)$.

[Ginsburg, Spanier, Salomaa, Gruska, Yntema 67-71]:
A word $w \in L(G)$ has index k if there is a derivation

$$S \Rightarrow w_1 \Rightarrow w_2 \ldots \Rightarrow w_n \Rightarrow w$$

such that each of $S, w_1, \ldots, w_n$ contains at most k occurrences of non-terminals (and one of them contains k non-terminals).

Example: $X = a \cdot X \cdot X + b$

b	has index 1	$X \Rightarrow b$
$(ab)^i b$	has index 2	$X \Rightarrow aXX \Rightarrow abX \stackrel{*}{\Rightarrow} (ab)^i X \Rightarrow (ab)^i b$
$aabbabb$	has index 3	$X \stackrel{*}{\Rightarrow} aaXXX \stackrel{*}{\Rightarrow} aabbabb$

Theorem [EKL DLT'07]: Let $X = f(X)$ be a system of language equations, and let G be the derived context-free grammar. For every $i \geq 0$:

$$\nu_i = L_{i+1}(G)$$

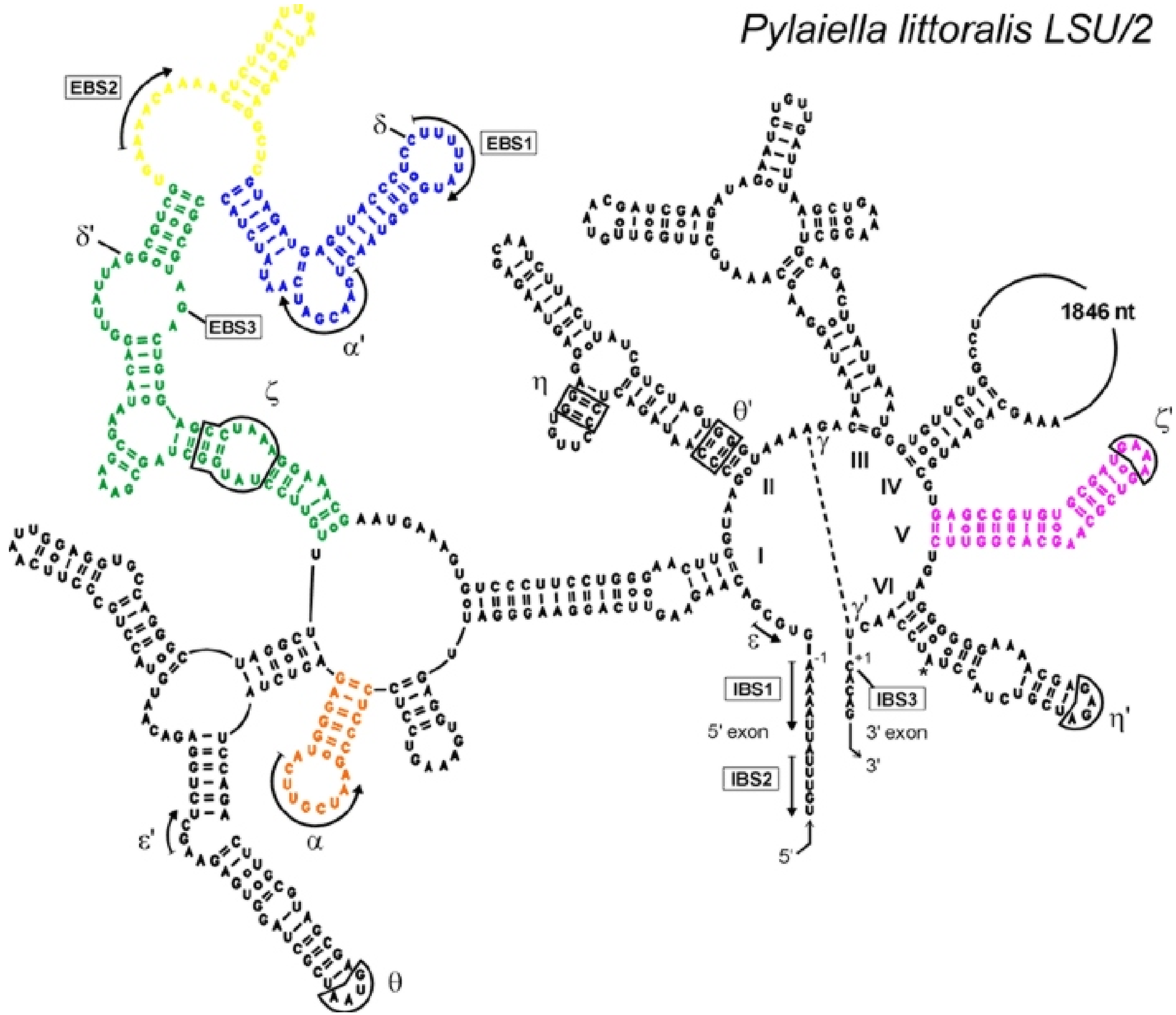
We can easily construct grammars G_i such that $L(G_{i+1}) = \nu_i$

$$X \;=\; a \cdot X \cdot X + b \qquad G \;=\; \{X \to aXX \mid b\}$$

$$\begin{array}{rcll} G_0 & = & & \{X_0 \to b\} \\ G_1 & = & G_0 & \cup \; \{X_1 \to aX_1X_0 \mid aX_0X_1 \mid aX_0X_0 + b\} \\ G_{i+1} & = & G_i & \cup \; \{X_{i+1} \to aX_{i+1}X_i \mid aX_iX_{i+1} \mid aX_iX_i + b\} \end{array}$$

Newton's method approximates a context-free grammar by context-free grammars of finite index.

Visualizing finite index: Secondary structure of RNA



(image by Bassi, Costa, Michel; www.cgm.cnrs-gif.fr/michel/)

An stochastic context-free grammar

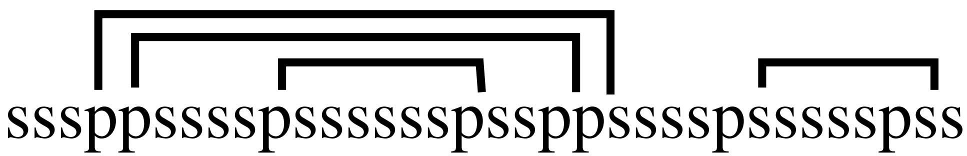
[ ]: Model the distribution of secondary structures as the derivation trees of the following stochastic context-free grammar:

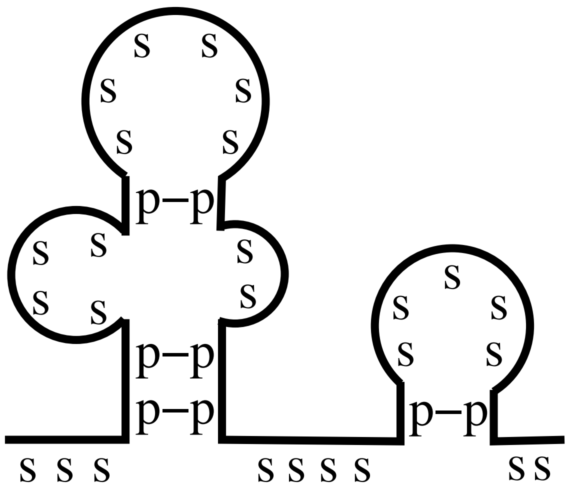
$$L \xrightarrow{0.869} CL \qquad L \xrightarrow{0.131} C$$

$$S \xrightarrow{0.788} pSp \qquad S \xrightarrow{0.212} CL$$

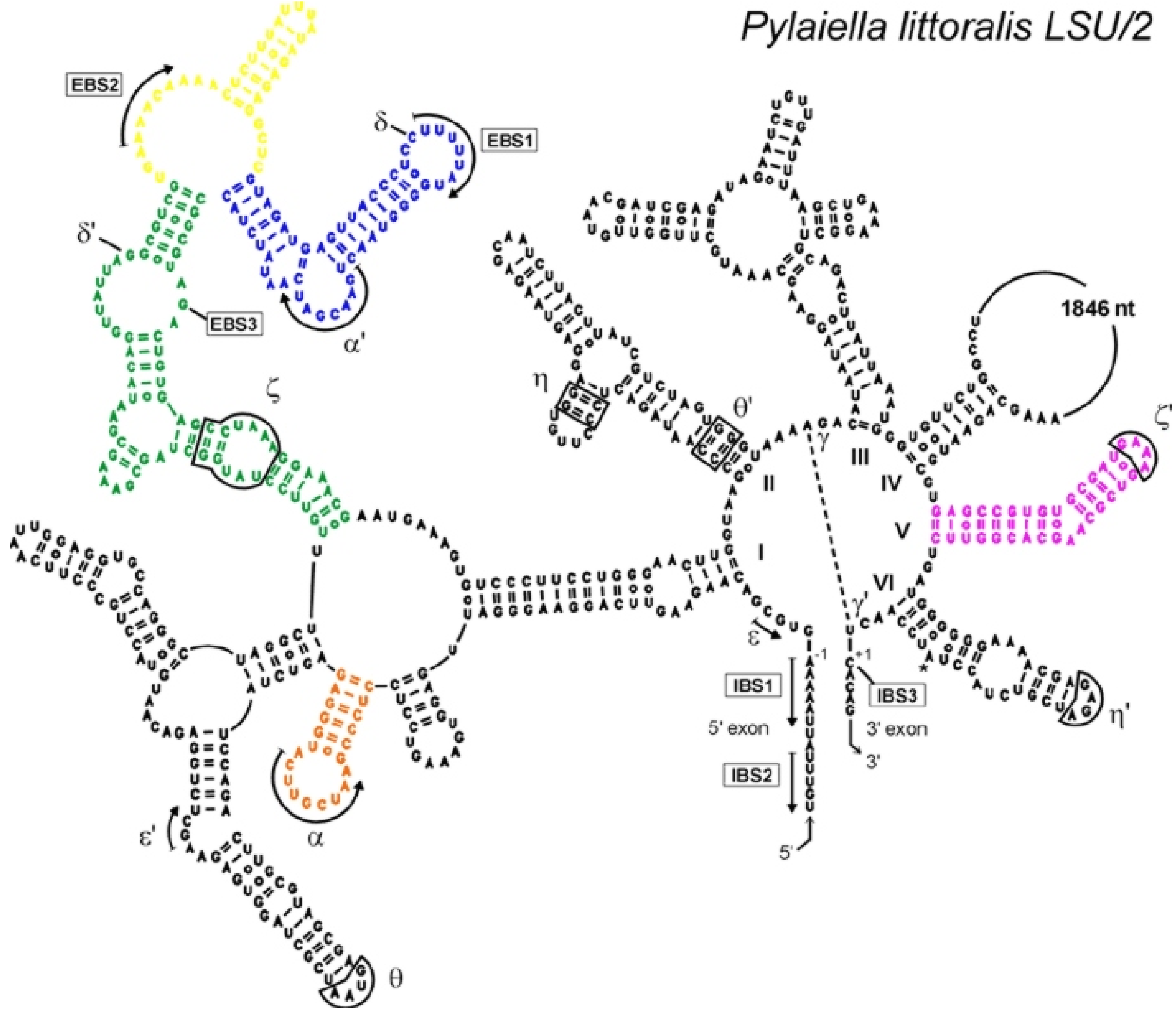
$$C \xrightarrow{0.895} s \qquad C \xrightarrow{0.105} pSp$$

Graphical interpretation:

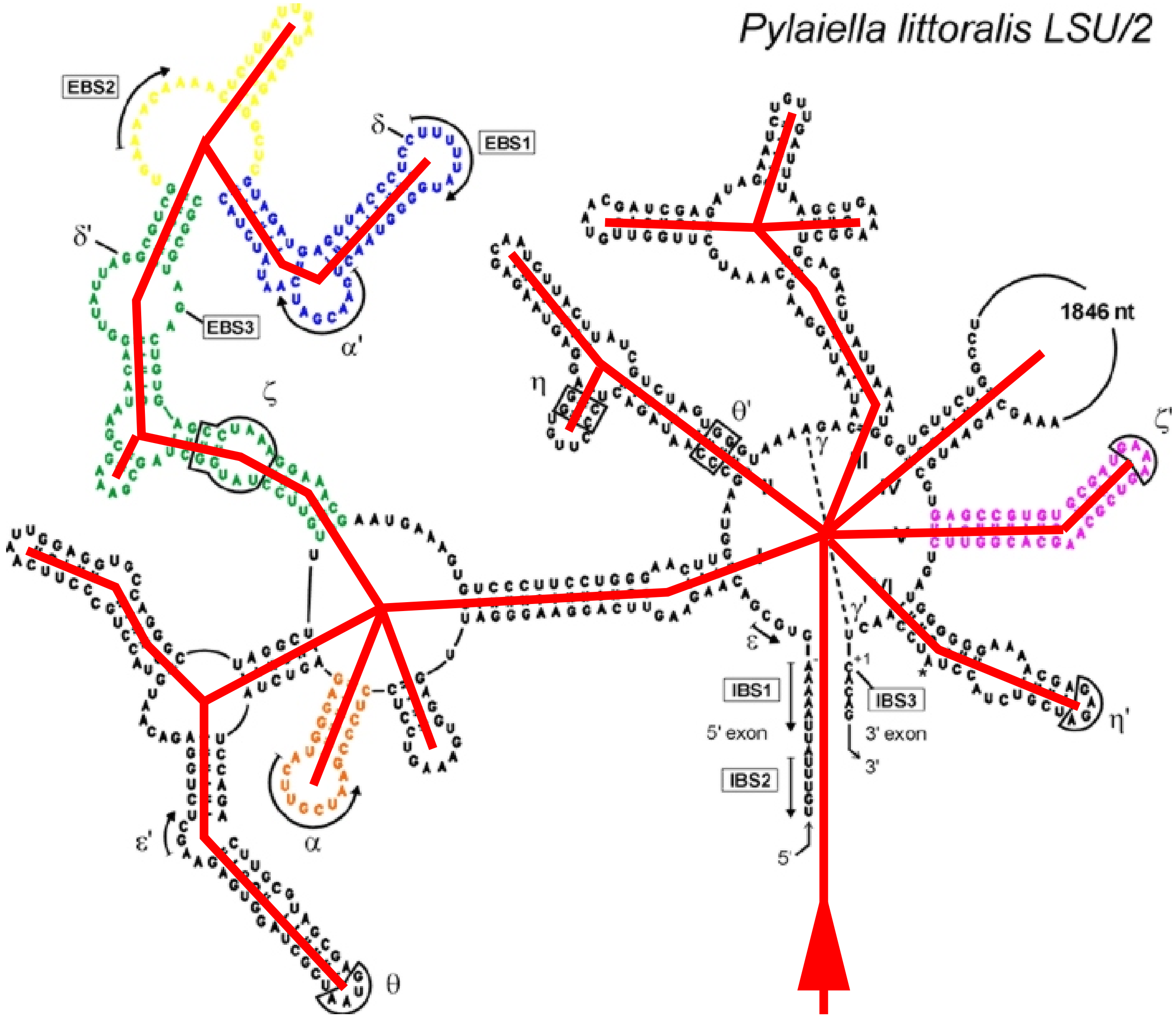




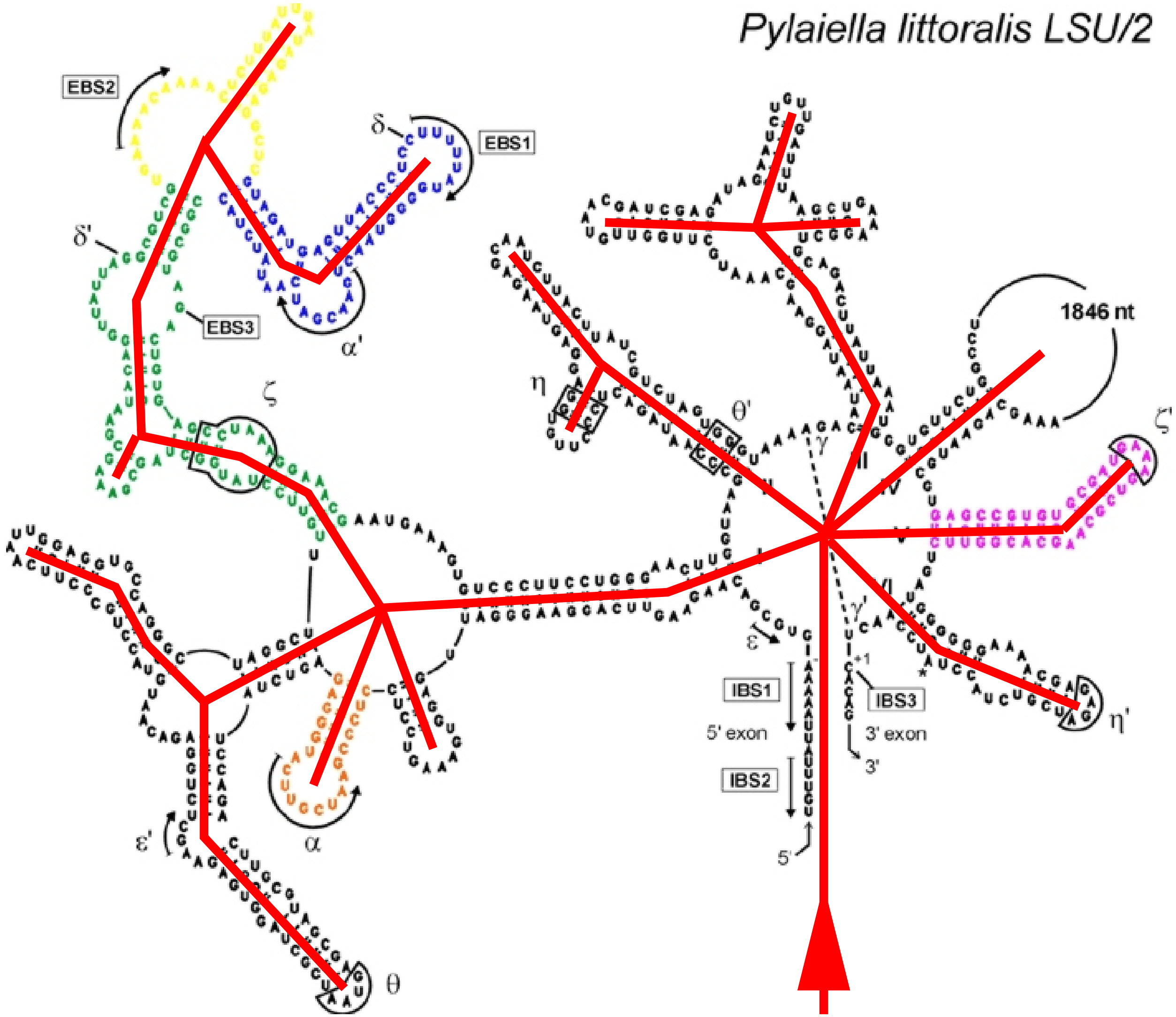
Visualizing the index of a derivation



Visualizing the index of a derivation



Visualizing the index of a derivation



Index = maximal number of branching points from root to leaf + 1

Grammar leads to two equation systems:

$$
\begin{aligned}
L &= C \cdot L + C & \qquad \hat{L} &= 0.869 \cdot \hat{C} \cdot \hat{L} + 0.131 \cdot \hat{C} \\
S &= p \cdot S \cdot p + C \cdot L & \qquad \hat{S} &= 0.788 \cdot \hat{S} + 0.212 \cdot \hat{C} \cdot \hat{L} \\
C &= s + p \cdot S \cdot p & \qquad \hat{C} &= 0.895 + 0.105 \cdot \hat{S}
\end{aligned}
$$

$$
\begin{aligned}
\nu_0(L) &= \text{der. of index} \leq 1 & \qquad \hat{\nu}_0(L) &= 0.5585 \\
\nu_1(L) &= \text{der. of index} \leq 2 & \qquad \hat{\nu}_1(L) &= 0.8050 \\
\nu_2(L) &= \text{der. of index} \leq 3 & \qquad \hat{\nu}_2(L) &= 0.9250 \\
\nu_3(L) &= \text{der. of index} \leq 4 & \qquad \hat{\nu}_3(L) &= 0.9789 \\
\nu_4(L) &= \text{der. of index} \leq 5 & \qquad \hat{\nu}_4(L) &= 0.9972 \\
\nu_5(L) &= \text{der. of index} \leq 6 & \qquad \hat{\nu}_5(L) &= 0.9999
\end{aligned}
$$

Idempotent and commutative semirings

Theorem [Hopkins-Kozen LICS '99]: The least fixed point of a system $X = f(X)$ of n equations over an ω-continuous idempotent and commutative semiring is reached by the sequence

$$
\begin{aligned}
\nu_0 &= f(0) \\
\nu_{i+1} &= J(\nu_i)^* \cdot f(\nu_i)
\end{aligned}
$$

after at most $O(3^n)$ iterations.

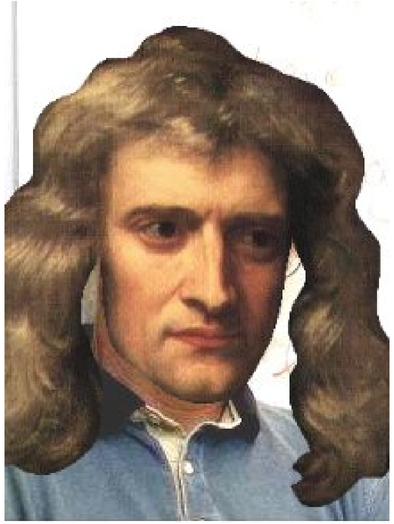
.

Idempotent and commutative semirings

Theorem [Hopkins-Kozen LICS '99]: The least fixed point of a system $X = f(X)$ of n equations over an ω-continuous idempotent and commutative semiring is reached by the sequence

$$\begin{aligned} \nu_0 &= f(0) \\ \nu_{i+1} &= J(\nu_i)^* \cdot f(\nu_i) \end{aligned}$$

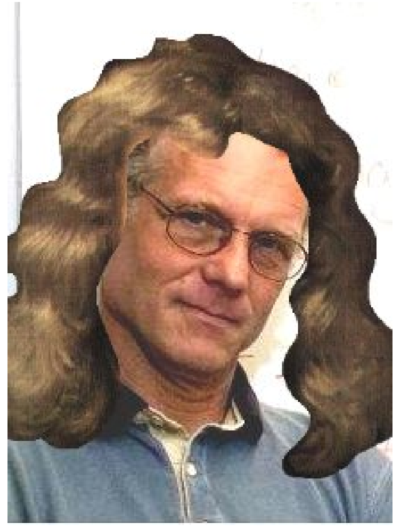
after at most $O(3^n)$ iterations.



.

Idempotent and commutative semirings

Theorem [Hopkins-Kozen LICS '99]: The least fixed point of a system $X = f(X)$ of n equations over an ω-continuous idempotent and commutative semiring is reached by the sequence



$$\begin{aligned} \nu_0 &= f(0) \\ \nu_{i+1} &= J(\nu_i)^* \cdot f(\nu_i) \end{aligned}$$

after at most $O(3^n)$ iterations.

.

Idempotent and commutative semirings

Theorem [Hopkins-Kozen LICS '99]: The least fixed point of a system $X = f(X)$ of n equations over an ω-continuous idempotent and commutative semiring is reached by the sequence

$$\begin{aligned} \nu_0 &= f(0) \\ \nu_{i+1} &= J(\nu_i)^* \cdot f(\nu_i) \end{aligned}$$

after at most $O(3^n)$ iterations.

Theorem [EKL STACS'07]: <u>This is exactly Newton's sequence</u>.

.

Idempotent and commutative semirings

Theorem [Hopkins-Kozen LICS '99]: The least fixed point of a system $X = f(X)$ of n equations over an ω-continuous idempotent and commutative semiring is reached by the sequence

$$\begin{aligned} \nu_0 &= f(0) \\ \nu_{i+1} &= J(\nu_i)^* \cdot f(\nu_i) \end{aligned}$$

after at most $O(3^n)$ iterations.

Theorem [EKL STACS'07]: This is exactly Newton's sequence.

Moreover, the fixed point is reached after at most n iterations.

An example

The Newton sequence terminates for all idempotent and commutative analyses, the Kleene sequence does not.

$$\begin{aligned} X &= a \cdot X \cdot X + b \\ f'(X) &= a \cdot X + a \cdot X = a \cdot X \end{aligned}$$

For one equation: $\mu f = \nu_1 = f'(\nu_0)^* \cdot \nu_0$

We obtain:
$$\begin{aligned} \nu_0 &= b \\ \nu_1 &= (ab)^* b \end{aligned}$$

This result provides a computational version of Parikh's theorem:
[Hopkins, Kozen LICS 99], [Aceto, Esik, Ingólfsdottir ITA 02]

The regular language

$$(a \cdot b)^* \cdot b$$

has the same Parikh image ("counting semantics") as the context-free language generated by the grammar

$$X \to aXX \mid b$$

Our two questions

Can Newton's Method be generalized to arbitrary ω-continuous semirings?

Is Newton's method robust when restricted to the real semiring?

Newton’s method on the real semiring

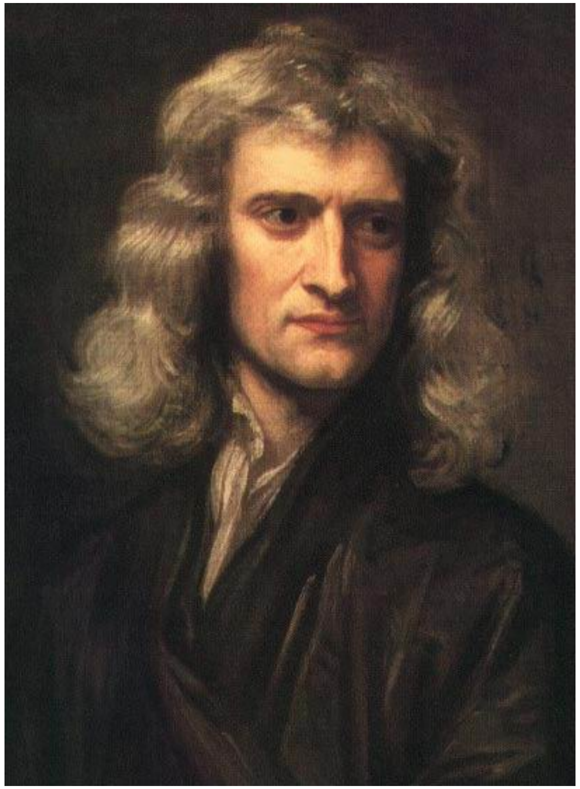
On the real field Newton’s method may not converge, or converge only locally

On the real semiring these problems disappear [EKL TCS 08]:

- Newton’s method always converges [EY STACS 05]
- It always exhibits linear or exponential convergence [EKL STOC 07]
- For strongly connected systems there is a threshold *k* such that after *k* iterations each subsequent iteration gains at least one bit of accuracy [EKL STACS 08]
- For important classes the threshold is linear in the size of the system [EKL STACS 08].

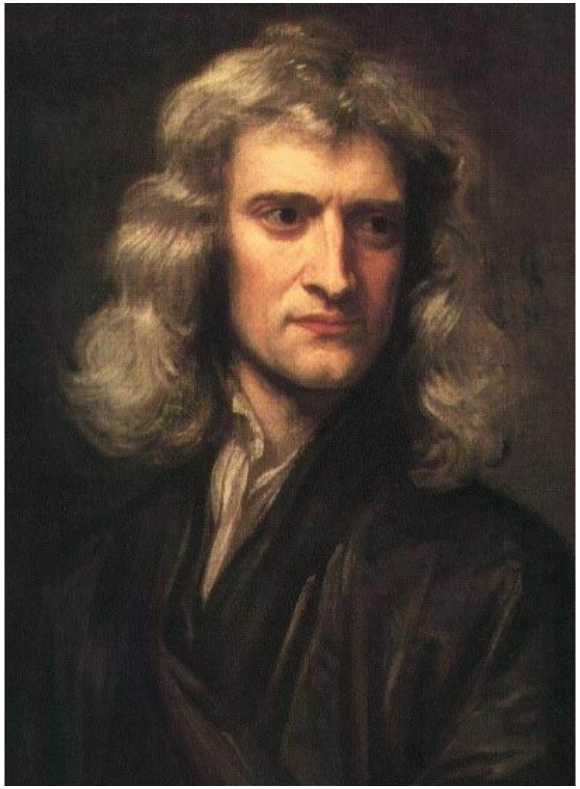
Conclusions

Newton did it all



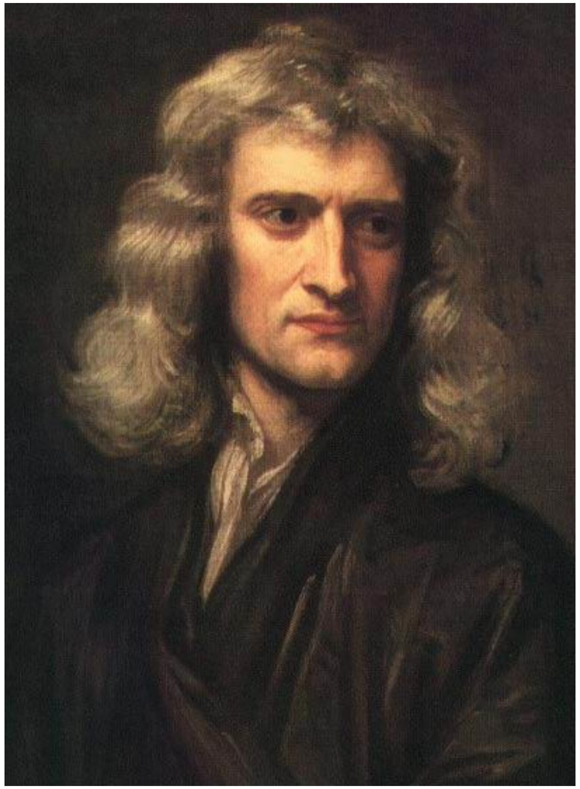
Conclusions

Newton did it all　　　but never saw Iceland



Conclusions

Newton did it all



but never saw Iceland

. . . and I did!