Iterative Solvers for Large Linear Systems

Part Ib: Consistency and Convergence

Andreas Meister

University of Kassel, Department of Analysis and Applied Mathematics

Outline

- Basics of Iterative Methods
- Splitting-schemes
  - Jacobi- u. Gauß-Seidel-scheme
  - Relaxation methods
- Methods for symmetric, positive definite Matrices
  - Method of steepest descent
  - Method of conjugate directions
  - CG-scheme

Outline

- Multigrid Method
  - Smoother, Prolongation, Restriction
  - Twogrid Method and Extension
- Methods for non-singular Matrices
  - GMRES
  - BiCG, CGS and BiCGSTAB
- Preconditioning
  - ILU, IC, GS, SGS, ...

A trivial scheme

- Choose $B = I$

$$
\begin{aligned}
\Longrightarrow \quad M &= I^{-1}(I - A) = I - A \\
N &= I \\
\Longrightarrow \quad x_{m+1} &= (I - A)x_m + b
\end{aligned}
$$

"+ "	:	no assumptions on A
"+ "	:	$I^{-1}x$ is easy to calculate
"- "	:	bad approximation of A in general

Model problem:

$$
\underbrace{\begin{pmatrix} 0.7 & -0.4 \\ -0.2 & 0.5 \end{pmatrix}}_{A:=} \underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{x:=} = \underbrace{\begin{pmatrix} 0.3 \\ 0.3 \end{pmatrix}}_{b:=}
$$

- A is non-singular ($\det A = 0.27$) and $x^\star = A^{-1}b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$

Fundamental questions

Aim: Find an answer to each of the following questions

1. When does a Splitting scheme converge?
2. Which are the ingredients that determine the rate of convergence?

A trivial scheme

Trivial scheme				
m	$x_{m,1}$	$x_{m,2}$	$\varepsilon_m := \lVert x_m - x^* \rVert_\infty$	$\varepsilon_m / \varepsilon_{m-1}$
0	2.100000e+01	-1.900000e+01	2.000000e+01	
1	-1.000000e+00	-5.000000e+00	6.000000e+00	3.000000e-01
2	-2.000000e+00	-2.400000e+00	3.400000e+00	5.666667e-01
3	-1.260000e+00	-1.300000e+00	2.300000e+00	6.764706e-01
4	-5.980000e-01	-6.020000e-01	1.602000e+00	6.965217e-01
5	-1.202000e-01	-1.206000e-01	1.120600e+00	6.995006e-01
6	2.157000e-01	2.156600e-01	7.843400e-01	6.999286e-01
7	4.509740e-01	4.509700e-01	5.490300e-01	6.999898e-01
8	6.156802e-01	6.156798e-01	3.843202e-01	6.999985e-01
9	7.309760e-01	7.309759e-01	2.690241e-01	6.999998e-01
10	8.116832e-01	8.116832e-01	1.883168e-01	7.000000e-01
11	8.681782e-01	8.681782e-01	1.318218e-01	7.000000e-01
12	9.077248e-01	9.077248e-01	9.227525e-02	7.000000e-01
13	9.354073e-01	9.354073e-01	6.459267e-02	7.000000e-01
14	9.547851e-01	9.547851e-01	4.521487e-02	7.000000e-01
15	9.683496e-01	9.683496e-01	3.165041e-02	7.000000e-01
20	9.946805e-01	9.946805e-01	5.319484e-03	7.000000e-01
25	9.991060e-01	9.991060e-01	8.940457e-04	7.000000e-01
30	9.998497e-01	9.998497e-01	1.502623e-04	7.000000e-01
40	9.999958e-01	9.999958e-01	4.244537e-06	7.000000e-01
55	1.000000e-00	1.000000e-00	2.015120e-08	7.000000e-01
70	1.000000e-00	1.000000e-00	9.566903e-11	7.000002e-01
85	1.000000e-00	1.000000e-00	4.540812e-13	6.998631e-01
96	1.000000e-00	1.000000e-00	8.881784e-15	6.956522e-01

A trivial scheme

Model problem:

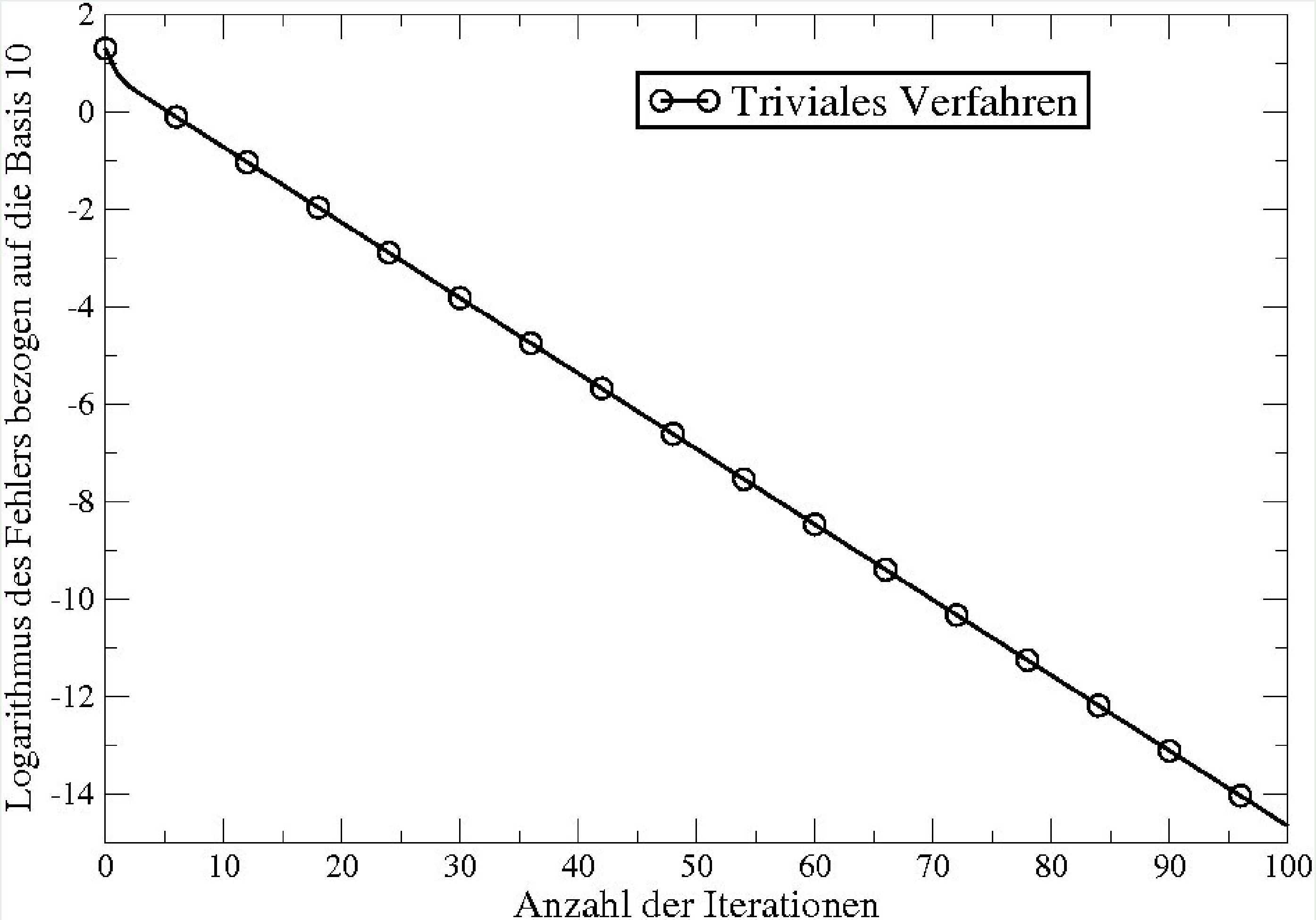


Abbildung: Convergence history $\log_{10} \varepsilon_m$

A trivial scheme

Definition: Spectral radius

A number $\lambda \in \mathbb{C}$ is called eigenvalue of A, if a vector $x \neq 0$ exists, such that $Ax = \lambda x$. The number

$$\rho(A) := \max\{|\lambda| : \lambda \text{ is eigenvalue of } A\}$$

is called spectral radius of A.

A trivial scheme

Model problem:

$$\underbrace{\begin{pmatrix} 0.7 & -0.4 \\ -0.2 & 0.5 \end{pmatrix}}_{A:=} \underbrace{\begin{pmatrix} x_1 \\ x_2 \end{pmatrix}}_{x:=} = \underbrace{\begin{pmatrix} 0.3 \\ 0.3 \end{pmatrix}}_{b:=}$$

- A is non-singular ($\det A = 0.27$)
- $x^\star = A^{-1} b = \begin{pmatrix} 1 \\ 1 \end{pmatrix}$
- Spectral radius of the iteration matrix:

$$\rho(M) = \rho(I - A) = \rho \begin{pmatrix} 0.3 & 0.4 \\ 0.2 & 0.5 \end{pmatrix} = 0.7$$

Consistency, convergence and rate of convergence

Aim: Find an answer to each of the following questions

1. When does a Splitting scheme converge?
2. Which are the ingredients that determine the rate of convergence?

Consistency, convergence and rate of convergence

Consistency:

An iterative solution method $x_{m+1} = \phi(x_m, b)$ is called consistent w.r.t. the matrix A, if the solution $x^\star = A^{-1}b$ represents a fixpoint of ϕ, that means

$$x^\star = \phi(x^\star, b)$$

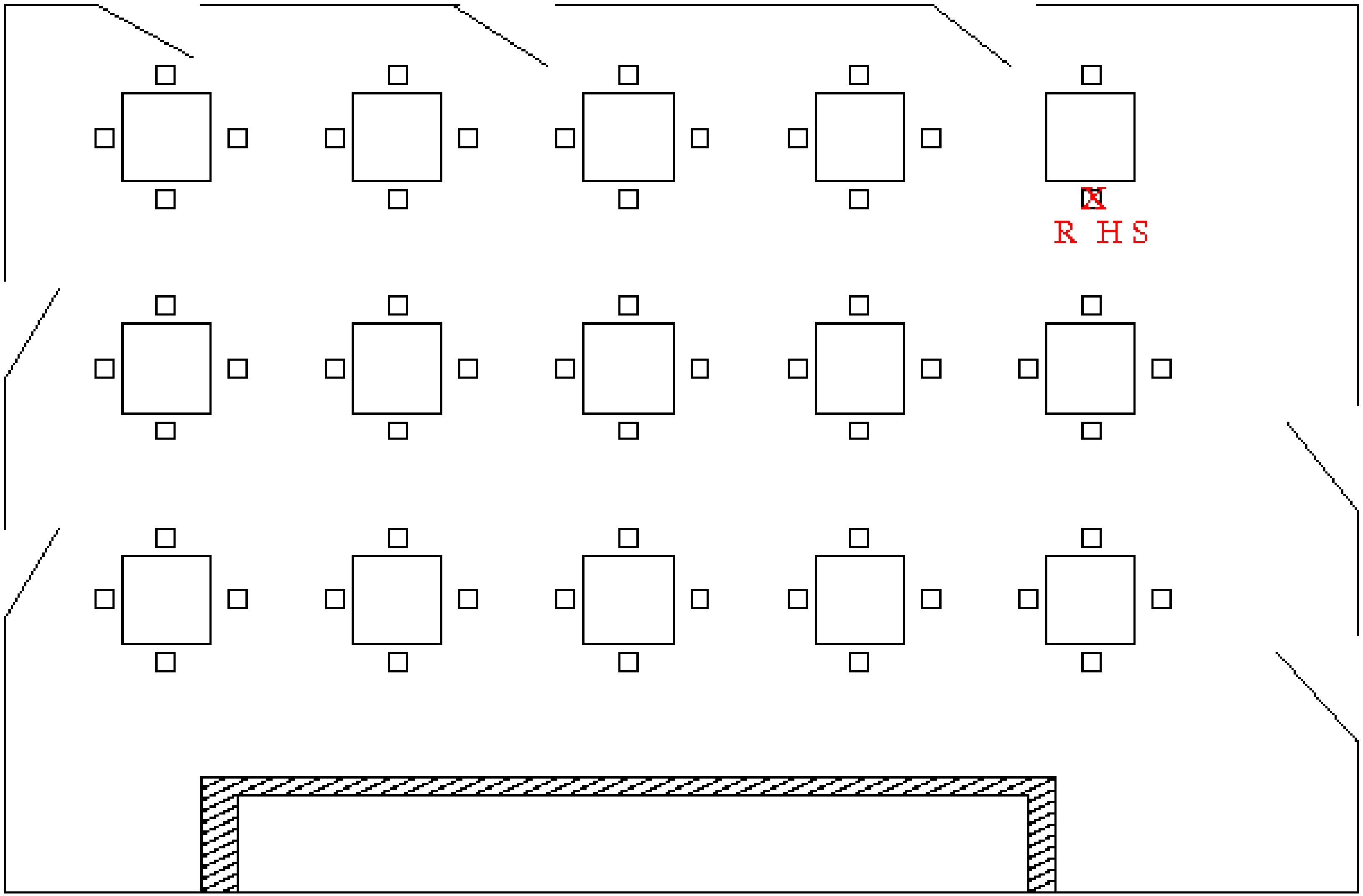
for each right hand side $b \in \mathbb{C}^n$.

In other words: Consistency means

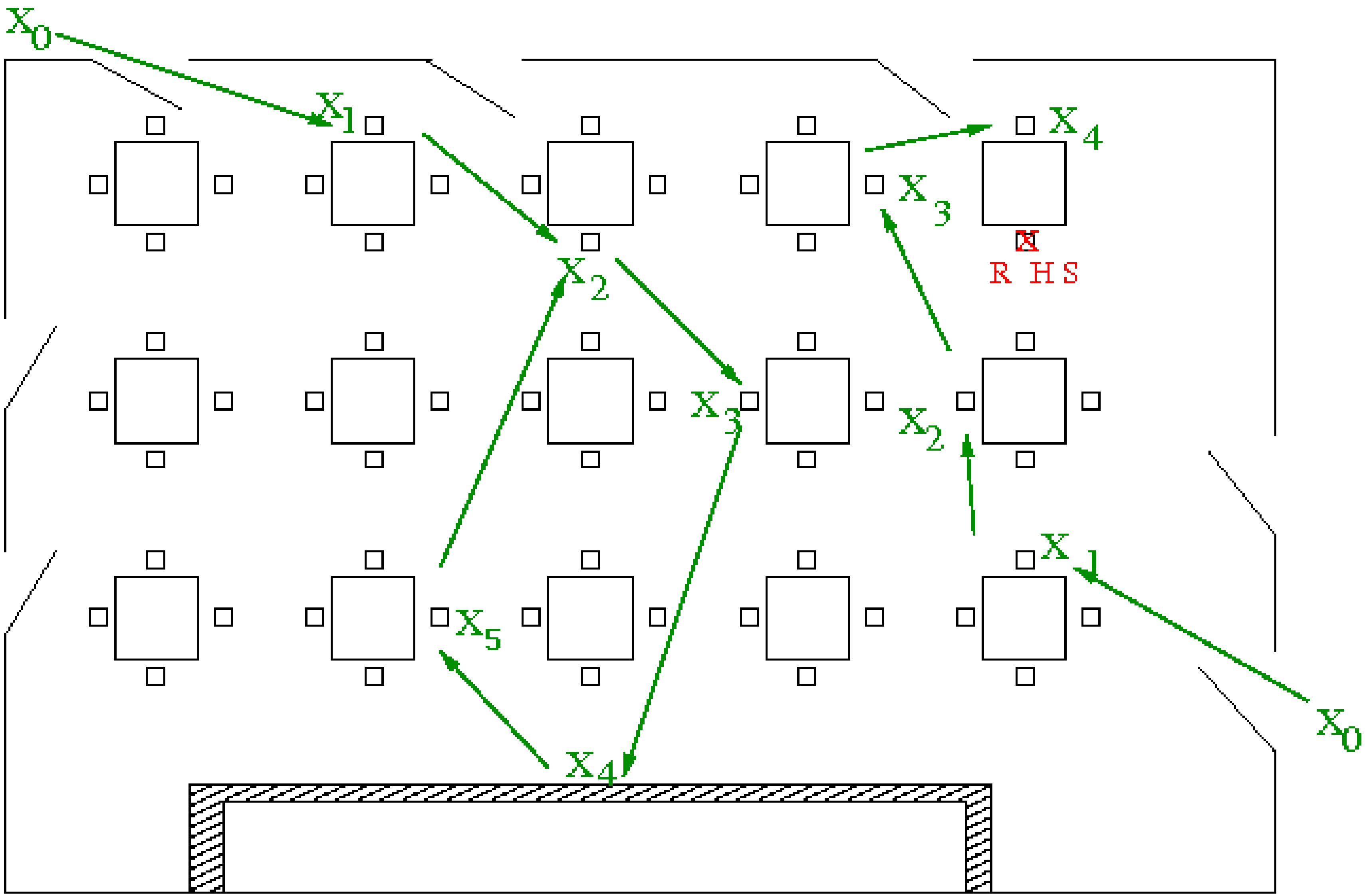
If the iterative solution method yields $\quad x_m = A^{-1}b,$

then $\quad x_k = A^{-1}b \quad$ for all $k \geq m$.

Part I: The cafeteria



Consistency:



Consistency

Statement for consistency

An iterative solution method is consistent if and only if

$$M = I - NA.$$

Justification: Let $x^\star = A^{-1}b$

"$\Longleftarrow$ "Let $M = I - NA$, then we obtain

$$x^\star = Mx^\star + N\underbrace{Ax^\star}_{=b} = Mx^\star + Nb = \phi(x^\star, b).$$

"$\Longrightarrow$ "Let ϕ be consistent, then

$$\begin{aligned} x^\star &= \phi(x^\star, b) = Mx^\star + Nb = Mx^\star + NAx^\star \\ &= (M + NA)x^\star \end{aligned}$$

$$\stackrel{b=Ax^\star}{\Longrightarrow} M = I - NA.$$

Consistency

General form of a Splitting method

$$x_{m+1} = \underbrace{B^{-1}(B-A)}_{M:=} x_m + \underbrace{B^{-1}}_{N:=} b, \quad m = 0, 1, \ldots \,.$$

For each Splitting method, one gets:

$$M = B^{-1}(B-A) = I - B^{-1}A = I - NA$$

Hence:

Each Splitting method is linear and consistent.

Convergence

Convergence:

An iterative solution method $x_{m+1} = \phi(x_m, b)$ is called convergent, if there exists a limit

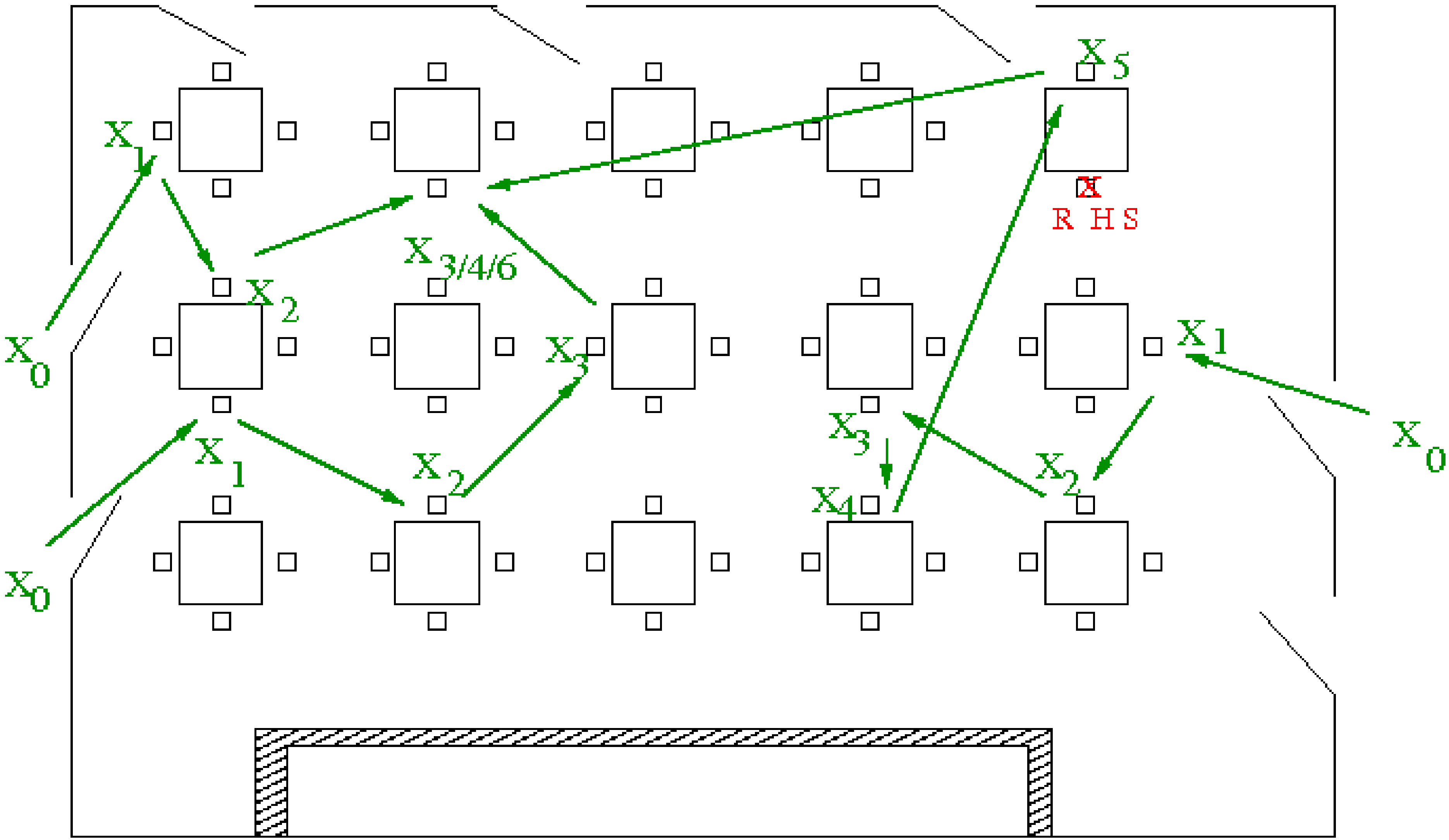
$$x = \lim_{m \to \infty} x_m = \lim_{m \to \infty} \phi(x_{m-1}, b)$$

for each right hand side $b \in \mathbb{C}^n$, which is independent of the initial guess $x_0 \in \mathbb{C}^n$

In other words: Convergence means:

The method has a unique destination.

Convergence:



Convergence and Consistency

We obtain:

For a consistent and convergent linear iterative solution method $x_{m+1} = \phi(x_m, b)$ one gets

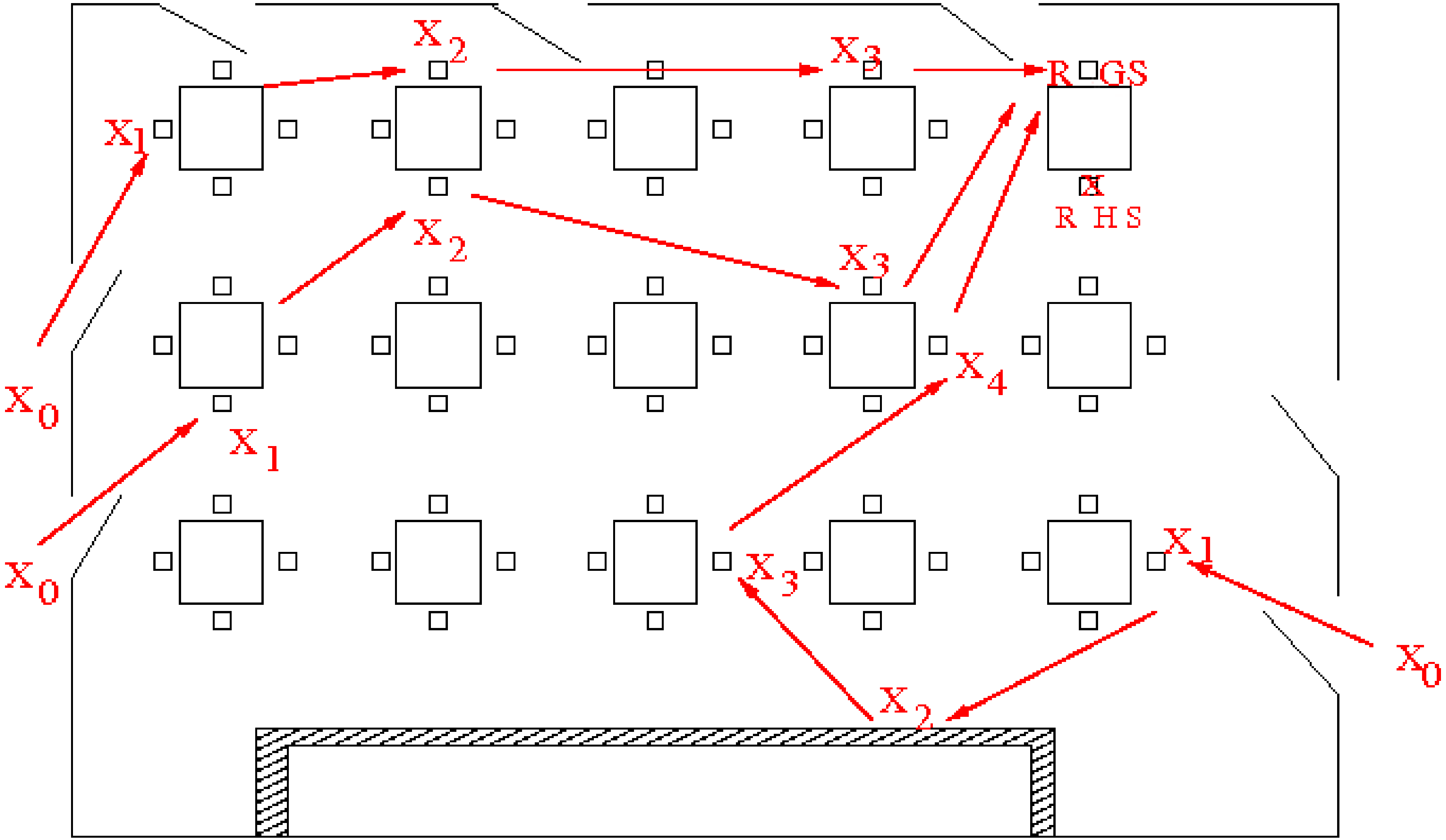
$$x^\star = A^{-1}b = \lim_{m\to\infty} \phi(x_m, b)$$

for all $x_0 \in \mathbb{C}^n$.

Justification:

- Convergence
  - $x = \lim_{m\to\infty} x_m$ represents a fixpoint of the linear mapping ϕ.
  - There exists exactly one fixpoint.
- Consistency
  - $x^\star = A^{-1}b$ is a fixpoint.

Consistency and Convergence



Banach fixed point theorem

When does a Splitting scheme converge?

Banach fixed point theorem:

Let D be a complete subset of a normed space X and let $f : D \longrightarrow D$ be a contracting mapping on X, then the sequence

$$x_{m+1} = f(x_m) \quad , m = 0, 1, \ldots$$

is convergent independent of the initial guess $x_0 \in D$. Furthermore the unique limit satisfies the equation $x = f(x) \in D$ and thus represents the unique fixpoint of f. Thereby, two inequalities describe the rate of convergence:

a priori: $$\|x_m - x\| \leq \frac{q^m}{1-q} \|x_1 - x_0\|$$

a posteriori: $$\|x_m - x\| \leq \frac{q}{1-q} \|x_m - x_{m-1}\|$$

where $0 \leq q < 1$ represents the Lipschitz constant of f.

Banach fixed point theorem

Definition

Contractivity means:
We have

$$\|f(x) - f(y)\| \leq q\|x - y\| \qquad \text{with} \quad 0 \leq q < 1.$$

for all x, y.

Banach fixed point theorem

Example:

We are looking for an $x \in D = [0,1]$ which satisfies $x = \cos x$.
$\Longrightarrow$ Consequently, we are looking for a fixpoint of

$$f(x) = \cos x \quad \text{in } [0,1]$$

Properties:

1. $f : [0,1] \longrightarrow [0,1]$
2. $[0,1]$ represents a complete subset of $\mathbb{R}$ w.r.t. $\|x\| = |x|$.
3. $f'(x) = -\sin x$
   $\Longrightarrow q := \max_{x\in[0,1]} |f'(x)| < 1$
   $\Longrightarrow |f(x) - f(y)| \leq q \cdot |x - y| \quad \text{with} \quad 0 \leq q < 1$

$\longrightarrow$ The sequence $x_{m+1} = f(x_m)$ will converge to $x = f(x)$ independet of the initial value $x_0 \in [0,1]$.

Banach fixed point theorem

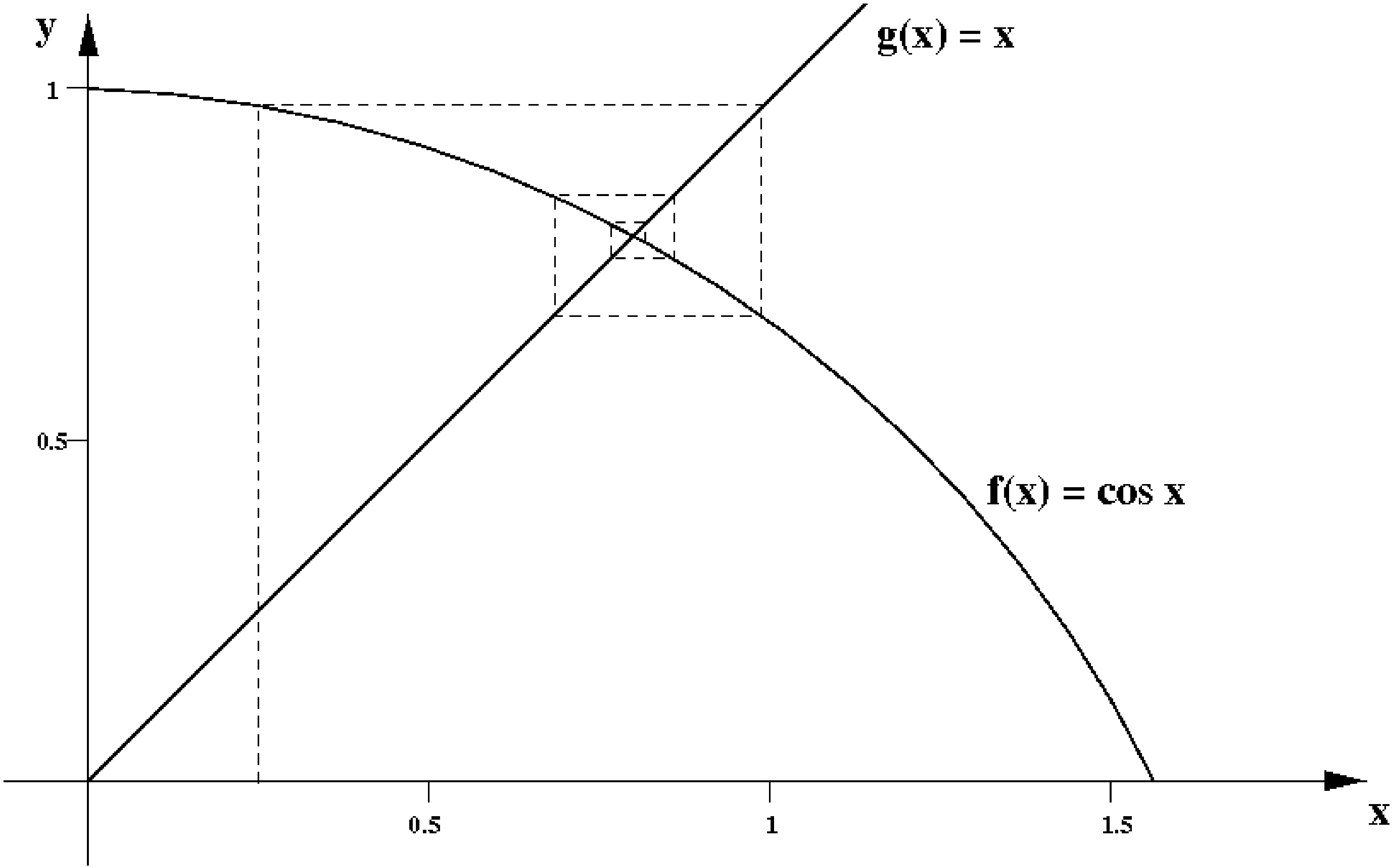


Fig.:Convergence history concerning $x_0 = 0.25$

Convergence

In the context of a Splitting scheme we have:

$$\|\phi(x,b)-\phi(y,b)\| = \|Mx+Nb-(My+Nb)\| = \|M(x-y)\| \leq \|M\|\|x-y\|$$

Thus our fixpoint theorem reads

Let $\|M\| < 1$, then the Splitting method

$$\phi(x,b) = Mx + Nb$$

convergent.

A-priori error estimate:

$$\|x_m - x^\star\| \leq \frac{\|M\|^m}{1 - \|M\|}\|x_1 - x_0\|$$

Conjuction between norm und spectral radius

There hold:

- $\rho(M) \leq \|M\|$ for each matrix norm $\|\cdot\|$.
- For each matrix M and each $\epsilon > 0$ there exists a norm such that

$$\|M\| \leq \rho(M) + \epsilon.$$

Thus, for each M we can write:

- If there exists a norm such that $\|M\| < 1$, then $\rho(M) < 1$
- if $\rho(M) < 1$, then there exists a norm such that $\|M\| < 1$.

Convergence

We obtain:

A Splitting method $\phi(x, b) = Mx + Nb$ is convergent if and only if

$$\rho(M) < 1$$

holds.

Definition: Rate of convergence

$\rho(M)$ is called rate of convergence.

Consistency, convergence and rate of convergence

Aim: Find an answer to each of the following questions

1. When does a Splitting scheme converge?

   Method is convergent if and only if $\iff \rho(M) < 1$

2. Which are the ingredients that determine the rate of convergence?

   The rate convergence directly depends on $\rho(M)$

   $\Longrightarrow$ The smaller the merrier

Summary

- Splitting methods are always linear.
- Splitting methods are always consistent.
- Splitting methods converge to $x^\star = A^{-1}b$ for each initial guess $x_0 \in \mathbb{C}^n$ to $x^\star = A^{-1}b$ if and only if $\rho(M) < 1$.
- Usually splitting methods are converging

  faster if the spectral radius $\rho(M)$ is smaller.
- **Rule of thumb** for convergent schemes:

  Squaring down the spectral radius leads to an iterative solution method, which requires only half of the iteration to reach the same error bound.