

DAS EXPONENTIELLE EULER-VERFAHREN  
ZUR REGULARISIERUNG SCHLECHT  
GESTELLTER PROBLEME



INAUGURAL-DISSERTATION

zur  
Erlangung des Doktorgrades der  
Mathematisch-Naturwissenschaftlichen Fakultät  
der Heinrich-Heine-Universität Düsseldorf

vorgelegt von  
Michael Hönig  
aus Karlsruhe

November 2008

Mathematisches Institut  
der Heinrich-Heine-Universität Düsseldorf

Gedruckt mit der Genehmigung der  
Mathematisch-Naturwissenschaftlichen Fakultät der  
Heinrich-Heine-Universität Düsseldorf

Referent: Prof. Dr. Marlis Hochbruck  
Koreferent: Prof. Dr. Martin Hanke-Bourgeois

Tag der mündlichen Prüfung: 19.12.2008

# DAS EXPONENTIELLE EULER-VERFAHREN ZUR REGULARISIERUNG SCHLECHT GESTELLTER PROBLEME

## ZUSAMMENFASSUNG

In dieser Arbeit wird das exponentielle Euler-Verfahren zur Regularisierung linearer und nichtlinearer schlecht gestellter Probleme vorgestellt und untersucht.

Nach einer kurzen Einführung wird im ersten Teil, ausgehend vom Showalter-Anfangswertproblem, das exponentielle Euler-Verfahren zur Regularisierung linearer schlecht gestellter Probleme hergeleitet. Die Ordnungsoptimalität dieses Verfahrens wird durch einen neuen kurzen Beweis gezeigt. Ferner wird erläutert, wie das Verfahren mittels Krylov-Unterraum-Verfahren effizient implementiert werden kann. Den Abschluss des ersten Teils bilden dann numerische Beispiele, welche die theoretischen Ergebnisse validieren und die Konkurrenzfähigkeit dieses Regularisierungsverfahrens gegenüber einem Standardverfahren darlegen.

Im zweiten Teil der Arbeit wird das Verfahren aus Teil Eins verallgemeinert zu einem vereinfachten exponentiellen Euler-Verfahren für nichtlineare schlecht gestellte Probleme. Dabei bildet die nichtlineare Version des Showalter-Anfangswertproblems den Ausgangspunkt. Unter Voraussetzung der Tangentialkegelbedingung und der Verwendung eines geeigneten Diskrepanzprinzips zur Bestimmung der Schrittweiten, wird die Konvergenz dieses Verfahrens bei exakten und gestörten Daten nachgewiesen. Im Anschluss daran wird die Ordnungsoptimalität des Verfahrens unter den gleichen schwachen Voraussetzungen, unter denen die entsprechende Aussage für die asymptotische Regularisierung gezeigt wurde, bewiesen. Dabei wird nur eine Beschränkung des Schrittweitenwachstums gefordert. Dagegen sind variable Schrittweiten erlaubt. Das Verfahren wird in die bestehende Literatur eingeordnet und insbesondere wird die Verbindung zu inexakten Newton-Verfahren erläutert. Danach wird eine effiziente Implementierung durch Krylov-Unterraum-Verfahren vorgestellt. Auch der zweite Teil wird mit numerischen Beispielen abgeschlossen. Die untersuchten Parameteridentifikationsprobleme reflektieren die Konvergenz und Ordnungsoptimalität des Verfahrens und ein Vergleich mit Standardverfahren zeigt die Wettbewerbsfähigkeit des exponentiellen Euler-Verfahrens.



# DAS EXPONENTIELLE EULER-VERFAHREN ZUR REGULARISIERUNG SCHLECHT GESTELLTER PROBLEME

## ABSTRACT

This thesis considers the exponential Euler scheme as a regularization method for linear and nonlinear ill-posed problems.

After a short introduction, the exponential Euler scheme as regularization method for linear ill-posed problems is derived from the Showalter initial value problem. A new short proof of the orderoptimality of this method is presented. It is shown that this method can be efficiently implemented using Krylov subspace methods. The first part of this thesis is closed by numerical experiments which reflect the theoretical results and show the competitiveness of this regularization method.

The second part of this thesis considers nonlinear ill-posed problems. The exponential Euler method from part one is generalized for nonlinear ill-posed problems, based on the nonlinear version of the Showalter initial value problem. Under the tangential cone condition and a suitable discrepancy principle, the method is shown to be convergent for exact and noisy data. Furthermore it is proved that the method converges with optimal order under the same assumptions that are needed for the continuous case. The presented convergence analysis admits variable stepsizes under a moderate restriction to the growth of the stepsizes. The connection to inexact Newton iterations is discussed. We close the second part with numerical experiments. Two parameter identification problems reflect convergence and orderoptimality of the method and a comparison with standard regularization schemes shows the competitiveness of the exponential Euler method.

# INHALTSVERZEICHNIS

---

EINLEITUNG	1
------------	---

---

KAPITEL 1	LINEARE SCHLECHT GESTELLTE PROBLEME	4
1.1.	Theorie linearer schlecht gestellter Probleme . . . . .	5
1.2.	Diskrepanzprinzip . . . . .	11
1.3.	Asymptotische Regularisierung . . . . .	13
1.4.	Regularisierung durch numerische Integratoren . . . . .	14
1.5.	Implementierung . . . . .	16
1.6.	Numerische Beispiele . . . . .	19
1.6.1.	Differentiation . . . . .	19
1.6.2.	Geologische Prospektion . . . . .	24

---

KAPITEL 2	NICHTLINEARE SCHLECHT GESTELLTE PROBLEME	29
2.1.	Theorie nichtlinearer schlecht gestellter Probleme . . . . .	30
2.2.	Asymptotische Regularisierung . . . . .	34
2.3.	Das nichtlineare exponentielle Euler-Verfahren . . . . .	37
2.4.	Konvergenz des exponentiellen Euler-Verfahrens . . . . .	40
2.5.	Optimale Konvergenzraten . . . . .	48
2.6.	Implementierung . . . . .	62
2.7.	Numerische Beispiele . . . . .	65

2.7.1. Parameteridentifikation in einer elliptischen partiellen Differentialgleichung . . . . .	65
2.7.2. Ein inverses Problem aus der Grundwasserhydrologie . . . . .	74

---

FAZIT	79
-------	----

---

KAPITEL A	EXPONENTIELLE INTEGRATOREN	81
-----------	----------------------------	----

---

KAPITEL B	APPROXIMATION VON MATRIXFUNKTIONEN	83
-----------	------------------------------------	----



# EINLEITUNG

Die vorliegende Arbeit verknüpft zwei interessante Gebiete der Mathematik, die in den letzten Jahren aktiv erforscht wurden. Auf der einen Seite handelt es sich dabei um schlecht gestellte Probleme und auf der anderen Seite um exponentielle Integratoren.

In der Geophysik, der medizinischen Bildgebung, der Astronomie und vielen anderen Bereichen sind wir häufig mit dem Problem konfrontiert aus bestimmten Daten die Parameter des zugrundeliegenden Modells zu bestimmen. In einem solchen Fall sprechen wir dann von einem *inversen Problem*. Eine ganze Reihe von Zeitschriften („Inverse Problems“, „Inverse and Ill-Posed Problems“, „Inverse Problems in Science and Engineering“, „Inverse Problems and Imaging“) und Büchern [2, 20, 25, 29, 43, 51, 58, 73, 85] zeigen das große Interesse an dieser Problemstellung. Die große Schwierigkeit inverser Probleme liegt darin begründet, dass sie vielmals nicht die von Hadamard [26] formulierten Eigenschaften eines **gut gestellten** Problems besitzen:

1. Es existiert für beliebige Daten eine Lösung.
2. Die Lösung ist eindeutig.
3. Die Lösung hängt stetig von den Daten ab.

Ist eine dieser Eigenschaften nicht erfüllt, so nennen wir das Problem **schlecht gestellt**. Zwar ist der Raum, aus dem die Daten stammen, in der Regel der Lösungsraum des direkten Problems, sind die Daten aber verrauscht, so kann man oft die Existenz einer Lösung nicht sicherstellen. Um die Problematik der Eindeutigkeit zu umgehen, werden wir in dieser Arbeit den Raum der zulässigen Lösungen so restringieren, dass wir eine eindeutige Lösung definieren können. Den größten Aufwand müssen wir allerdings betreiben, wenn ein Problem die dritte Forderung von Hadamard nicht erfüllt. In diesem Fall ist die numerisch berechnete Lösung des Problems aufgrund der Verstärkung von Rundungsfehlern und des in der Regel in den Daten vorhandenen Rauschens völlig unbrauchbar. Es sind daher Methoden notwendig um diesem Verhalten entgegenzuwirken. Wir nennen diese Regularisierungsverfahren. Ein wesentlicher Bestandteil von Regularisierungsverfahren ist es, die Lösung nur

bis zu einer bestimmten, mit Bedacht gewählten Genauigkeit zu bestimmen. Wir werden so genau rechnen bis die Norm des Residuums in etwa dem Rauschpegel entspricht.

Die Basis des in dieser Arbeit verwendeten Regularisierungsverfahrens bildet das sogenannte Showalter-Anfangswertproblem, welches durch Einführung einer künstlichen Zeit entsteht. Zur numerischen Lösung von Anfangswertproblemen kommen Integrationsverfahren zum Einsatz. Diverse wohlbekanntes Regularisierungsverfahren können über diesen Ansatz motiviert werden. So entsprechen im Falle eines linearen schlecht gestellten Problems das Landweber-Verfahren und die Tikhonov-Philipps-Regularisierung dem expliziten bzw. impliziten Euler-Verfahren zur Lösung des Showalter-Anfangswertproblems. Im nichtlinearen Fall erhält man aus Varianten des Euler-Verfahrens das Landweber-Verfahren oder das Levenberg-Marquardt-Verfahren. Die Theorie linearer schlecht gestellter Probleme ist schon ziemlich komplett und beispielsweise in den Büchern [20, 43, 58, 73] zu finden. Ferner wurden in diesem Fall die regularisierende Eigenschaft von Runge-Kutta-Verfahren in [75] nachgewiesen. Im nichtlinearen Fall (siehe [51]), gibt es dagegen noch einige offene Fragen. So ist das Landweber-Verfahren nur für konstante Schrittweiten untersucht [32] und für das Levenberg-Marquardt-Verfahren konnte bisher nur die Konvergenz nachgewiesen werden (siehe [30]). Auch in den aktuellen Arbeiten [8, 57] konnte die Konvergenz der untersuchten Verfahren nur unter starken Schrittweitenbeschränkungen gezeigt werden. Ein Resultat eines Verfahrens mit variabler Schrittweite und optimalen Konvergenzraten fehlt bisher noch ganz. Für ein spezielles Verfahren, das exponentielle Euler-Verfahren, wird in der vorliegenden Arbeit erstmals ein solches Resultat gezeigt.

Das exponentielle Euler-Verfahren gehört zur Klasse der exponentiellen Integratoren. Diese Verfahren zur numerischen Lösung von Anfangswertproblemen sind schon recht lange bekannt (siehe zum Beispiel [56, 66, 70]). Sie eignen sich besonders für steife und oszillatorische Probleme und lösen lineare Anfangswertprobleme exakt. Allerdings galten exponentielle Integratoren lange Zeit als nicht praktikabel, da bei diesen Verfahren Funktionen von Operatoren oder, im diskretisierten Fall, Matrizen berechnet werden müssen und dies zunächst nur für kleine Matrizen möglich war. Mit den Arbeiten [16, 17, 22, 38, 62] zur Approximation von Matrixfunktionen änderte sich dies und exponentielle Integratoren konnten nun auch zur Lösung von großen Problemen erfolgreich verwendet werden. Seitdem sind diese Methoden intensiv untersucht (vgl. [11, 14, 40, 41, 42, 67]) und in Anwendungen eingesetzt (vgl. [6, 5, 78]) worden.

Die Arbeit teilt sich in zwei Kapitel, in denen zum einen lineare und zum anderen nichtlineare schlecht gestellte Probleme behandelt werden.

Nachdem wir das lineare schlecht gestellte Problem definiert haben, wiederholen wir kurz die wichtigsten bekannten Resultate zur Theorie linearer schlecht gestellter Probleme. Dabei gehen wir insbesondere genauer auf die asymptotische Regularisierung und die Regularisierung durch Integratoren ein. Ferner präsentieren wir einen neuen kurzen Beweis der in [47] gezeigten Regularisierungseigenschaft des exponentiellen Euler-Verfahrens. In einem

kurzen Abschnitt fassen wir dann die Eckpunkte der Implementierung zusammen. Mit numerischen Beispielen, welche die Regularisierungseigenschaft dieser Methode widerspiegeln und die Konkurrenzfähigkeit dieser Regularisierung darlegen, schließen wir den Teil über lineare Probleme ab.

Auch den zweiten Teil der Arbeit beginnen wir mit der Problemstellung und einem kurzen Überblick über die zugehörige Theorie und bekannte Regularisierungsverfahren. Dabei wiederholen wir insbesondere die Resultate Tautenhahns [80] zur asymptotischen Regularisierung nichtlinearer schlecht gestellter Probleme und motivieren daraus den Einsatz eines vereinfachten exponentiellen Euler-Verfahrens. Im darauf folgenden Abschnitt zeigen wir die Konvergenz dieser Methode bei exakten und gestörten Daten unter in diesem Zusammenhang üblichen Voraussetzungen. Danach zeigen wir, dass das Verfahren mit optimaler Ordnung konvergiert unter denselben schwachen Voraussetzungen unter denen die entsprechende Aussage für die asymptotische Regularisierung gezeigt wurde und einer leichten Einschränkung an das Wachstum der Zeitschrittweite. Wir ordnen das Resultat in die vorhandene Literatur ein. Nach kurzen Bemerkungen zur Implementierung des vereinfachten exponentiellen Euler-Verfahrens zur Regularisierung schlecht gestellter Probleme stellen wir die Ergebnisse zweier numerischer Experimente vor. Auch hier wird die zuvor präsentierte Theorie wieder reflektiert. Ein Vergleich mit inexakten Newton-Verfahren zeigt die Qualität des Verfahrens auf.

Abschließend blicken wir in einem Fazit auf die Ergebnisse dieser Arbeit zurück und stellen einige aus dieser Arbeit hervorgehende, noch ungeklärte Fragestellungen vor.

# KAPITEL 1

## LINEARE SCHLECHT GESTELLTE PROBLEME

In diesem Kapitel beschäftigen wir uns mit linearen schlecht gestellten Problemen. Dazu wiederholen wir zunächst die wohlbekanntete Theorie linearer schlecht gestellter Probleme und der Regularisierungsverfahren, wie sie beispielsweise in [20, 58, 73] zu finden ist. Insbesondere untersuchen wir dabei die auf dem Showalter-Anfangswertproblem basierende asymptotische Regularisierung. Danach geben wir einen kurzen Überblick über Verfahren die auf der Integration dieses Anfangswertproblems basieren und stellen das exponentielle Euler-Verfahren vor. In Abschnitt 1.4 präsentieren wir einen neuen kurzen Beweis der Ordnungsoptimalität des exponentiellen Euler-Verfahrens zur Regularisierung linearer schlecht gestellter Probleme. Den Abschluss dieses Kapitels bilden dann numerische Beispiele, die die Konkurrenzfähigkeit dieses Verfahrens belegen.

## 1.1 Theorie linearer schlecht gestellter Probleme

Im Folgenden untersuchen wir die lineare schlecht gestellte Operatorgleichung

$$(1.1) \quad Ax = y, \quad y \in \mathcal{R}(A),$$

wobei  $A$  ein linearer, beschränkter Operator zwischen den Hilberträumen  $\mathcal{X}$  und  $\mathcal{Y}$  ist, und das Bild  $\mathcal{R}(A)$  von  $A$  nicht abgeschlossen ist. Die Eigenschaft des nicht abgeschlossenen Bildes macht die Lösung der Gleichung (1.1) zu einem schlecht gestellten Problem. In diesem Fall ist nämlich die verallgemeinerte Inverse (Moore-Penrose-Inverse)  $A^\dagger$  von  $A$  unbeschränkt [20, Proposition 2.4]. Die verallgemeinerte Inverse ist der Operator, der  $y$  aus (1.1) auf die eindeutige Lösung in  $\mathcal{N}(A)^\perp$  der zugehörigen Normalengleichung

$$(1.2) \quad A^*Ax = A^*y$$

abbildet. Dies ist die eindeutige Lösung  $x^\dagger \in \mathcal{X}$  der Gleichung (1.1) mit minimaler Norm. Ferner nehmen wir an, dass uns nur verrauschte Daten  $y^\delta \in \mathcal{Y}$  mit

$$(1.3) \quad \|y - y^\delta\|_{\mathcal{Y}} \leq \delta$$

zur Verfügung stehen. Der nicht-negative Parameter  $\delta$  entspricht dann dem Rauschpegel und wir gehen davon aus ihn zu kennen. Da das Rauschen in der Regel zufällig ist, können wir uns im Allgemeinen nicht sicher sein, dass  $y^\delta \in \mathcal{R}(A)$  ist, dass also die gestörte Gleichung

$$(1.4) \quad Au = y^\delta$$

überhaupt eine Lösung hat.

Wir bezeichnen im Folgenden die zur gestörten Gleichung (1.4) gehörenden Lösungen und deren Approximationen mit  $u$ .

Sofern  $y^\delta$  im Definitionsbereich  $\mathcal{D}(A^\dagger) = \mathcal{R}(A) \oplus \mathcal{R}(A)^\perp$  der verallgemeinerten Inversen  $A^\dagger$  von  $A$  liegt, können wir  $x^\dagger = A^\dagger y^\delta$  schreiben. Da wir  $\mathcal{R}(A)$  als nicht abgeschlossen vorausgesetzt haben, ist die verallgemeinerte Inverse  $A^\dagger$  von  $A$  allerdings unstetig (siehe [73, Satz 2.1.8]). Falls wir es also mit einem gestörten Problem zu tun haben, das heißt  $\delta > 0$  ist, liefert uns  $A^\dagger y^\delta$ , sofern es existiert, im Allgemeinen keine gute Approximation der Lösung von (1.1). Um trotzdem eine stabile Näherung der Lösung  $x^\dagger$  zu bekommen, müssen wir daher  $A^\dagger$  durch einen stetigen Operator  $R : \mathcal{Y} \rightarrow \mathcal{X}$  approximieren.

Wir interessieren uns im weiteren Verlauf für (numerische) Methoden  $R$ , die die folgenden regularisierenden Eigenschaften haben:

**Definition 1.1.** Sei  $\{R_t\}_{t>0}$  eine Familie stetiger Operatoren zwischen den Hilberträumen  $\mathcal{Y}$  und  $\mathcal{X}$  mit  $R_t 0 = 0$ . Existiert eine Funktion  $\alpha : (0, \infty) \times \mathcal{Y} \rightarrow (0, \infty)$  so, dass für jedes  $y \in \mathcal{R}(A)$  gilt

$$(1.5) \quad \sup \{ \|A^\dagger y - R_{\alpha(\delta, y^\delta)} y^\delta\|_{\mathcal{X}} : y^\delta \in \mathcal{Y}, \|y - y^\delta\|_{\mathcal{Y}} \leq \delta \} \longrightarrow 0 \quad \text{für } \delta \rightarrow 0,$$

dann heißt  $(\{R_t\}_{t>0}, \alpha)$  **Regularisierungsverfahren** für  $A^\dagger$ . Ist  $R_t$  für jedes  $t > 0$  linear, so sprechen wir von einem **linearen Regularisierungsverfahren**. Die **Parameterwahl**  $\alpha$  sei so orientiert, dass für jedes  $y \in \mathcal{R}(F)$

$$\lim_{\delta \rightarrow 0} \sup \{ \alpha(\delta, y^\delta) : y^\delta \in \mathcal{Y}, \|y - y^\delta\|_{\mathcal{Y}} \leq \delta \} = 0.$$

gilt. Die Funktion  $\alpha$  heißt **a priori Parameterwahl**, falls sie nur von  $\delta$  abhängt, ansonsten heißt  $\alpha$  **a posteriori Parameterwahl**.

**Bemerkung.** Es genügt auch, die Operatoren  $R_t$  anstatt auf der positiven reellen Achse nur für eine diskrete Menge  $\mathcal{T}$  positiver Zahlen mit Häufungspunkt in 0 zu definieren. Dies wird insbesondere bei iterativen Verfahren der Fall sein.  $\diamond$

### Beispiel 1.2.

1. Der Regularisierungsoperator  $R_t$  der Tikhonov-Phillips-Regularisierung ordnet  $y^\delta$  die Lösung des Minimierungsproblems

$$\|Ax - y^\delta\|_{\mathcal{Y}}^2 + t \|x\|_{\mathcal{X}}^2 \stackrel{!}{=} \min$$

zu.

2. Beim Landweberverfahren überführt man die zur gestörten Gleichung (1.4) gehörende Normalengleichung in die Fixpunktform  $u = u + A^*(y^\delta - Au)$  und erhält ausgehend von einem Startwert  $u_0$  eine Folge von Iterierten via

$$u_{n+1} = u_n + A^*(y^\delta - Au_n), \quad \text{für } n = 0, 1, 2, \dots$$

Die Regularisierungsoperatoren sind dann auf der Menge  $\{\frac{1}{n} : n \in \mathbb{N}\}$  definiert und  $R_{1/n}y^\delta = u_n$ .

Ist  $(\{R_t\}_{t>0}, \alpha)$  ein Regularisierungsverfahren für  $A^\dagger$ , dann folgt aus (1.5) mit  $y^\delta = y$ , dass

$$\lim_{\delta \rightarrow 0} R_{\alpha(\delta, y)}y = A^\dagger y$$

für alle  $y \in \mathcal{D}(A^\dagger)$  gilt. Für eine im ersten Argument stetige Parameterwahl  $\alpha$  und  $\Gamma := \{\alpha(\delta, y) : \delta > 0, y \in \mathcal{R}(A)\}$  gilt dann auch

$$\lim_{\Gamma \ni \sigma \rightarrow 0} R_\sigma y = A^\dagger y.$$

Ein Regularisierungsverfahren ist daher eine punktweise Approximation der verallgemeinerten Inversen  $A^\dagger$ . Dies erzwingt allerdings im Fall der Unstetigkeit von  $A^\dagger$  die Unbeschränktheit jedes linearen Regularisierungsverfahrens für  $A^\dagger$ , siehe [73, Lemma 3.1.3].

Wir sind natürlich an der Genauigkeit der obigen Approximation interessiert. Wie verhält sich die Näherungslösung  $R_{\alpha(\delta, y^\delta)}y^\delta$  zu (1.1) bei gegebenen Daten  $y^\delta$  für  $\delta$  gegen Null? Leider ist die Konvergenz im Allgemeinen beliebig langsam.

**Lemma 1.3.** *Es sei  $(\{R_t\}_{t>0}, \alpha)$  ein Regularisierungsverfahren für  $A^\dagger$ . Dann gibt es keine Funktion  $h : [0, \infty) \rightarrow [0, \infty)$  mit  $\lim_{\delta \rightarrow 0} h(\delta) = 0$  so, dass*

$$\sup \{ \|A^\dagger y - R_{\alpha(\delta, y^\delta)} y^\delta\|_{\mathcal{X}} : y \in \mathcal{R}(A), \|y\|_{\mathcal{Y}} \leq 1, y^\delta \in \mathcal{Y}, \|y - y^\delta\|_{\mathcal{Y}} \leq \delta \} \leq h(\delta).$$

Einen Beweis findet man [20]. Es macht also nur Sinn, Konvergenzraten auf Teilmengen von  $\mathcal{D}(A)$  oder auch  $\mathcal{X}$  zu untersuchen. Dazu verallgemeinern wir zunächst die Klasse der Regularisierungsverfahren und definieren den zugehörigen schlimmsten Fehler.

**Definition 1.4.** *Eine stetige Abbildung  $R : \mathcal{Y} \rightarrow \mathcal{X}$  mit  $R0 = 0$  heißt ein **stabiles Rekonstruktionsverfahren** zur Lösung der Operatorgleichung (1.1).*

Den **schlimmsten Fehler** dieses Verfahrens auf  $\mathcal{M} \subset \mathcal{X}$  und bei einem Rauschpegel  $\delta$  bezeichnen wir mit

$$E(\delta, \mathcal{M}, R) := \sup \{ \|Ry^\delta - A^\dagger y\|_{\mathcal{Y}} : y \in \mathcal{R}(A), y^\delta \in \mathcal{Y}, \|y - y^\delta\|_{\mathcal{Y}} \leq \delta, A^\dagger y \in \mathcal{M} \}.$$

Ein optimales Rekonstruktionsverfahren  $R_{\text{opt}}$  aus der Menge  $\Upsilon$  aller Rekonstruktionsverfahren erfüllt dann

$$E(\delta, \mathcal{M}, R_{\text{opt}}) = \inf \{ E(\delta, \mathcal{M}, R) : R \in \Upsilon \}.$$

Wir definieren daher:

**Definition 1.5.** *Sei  $\mathcal{M} \subset \mathcal{X}$ . Wir bezeichnen den **besten schlimmsten Fehler** bei gegebenem Rauschpegel  $\delta$  mit*

$$E(\delta, \mathcal{M}) := \inf \{ E(\delta, \mathcal{M}, R) : R \in \Upsilon \}.$$

Man zeigt leicht [20, Proposition 3.10], dass der beste schlimmste Fehler nach unten abgeschätzt werden kann:

$$(1.6) \quad E(\delta, \mathcal{M}) \geq \sup \{ \|x\|_{\mathcal{X}} : x \in \mathcal{M}, \|Ax\|_{\mathcal{Y}} \leq \delta \} =: e(\delta, \mathcal{M}).$$

Wir beschränken uns bei unseren weiteren Untersuchungen auf Teilmengen von  $\mathcal{X}$ , die a priori Informationen an die exakte Lösung beinhalten. Dazu betrachten wir Teilräume  $\mathcal{X}_{\gamma, \varrho}$  und  $\mathcal{X}_\gamma$  von  $\mathcal{X}$ , definiert durch

$$(1.7) \quad \begin{aligned} \mathcal{X}_{\gamma, \varrho} &:= \{(A^*A)^\gamma z : \|z\|_{\mathcal{X}} \leq \varrho\}, \quad \gamma, \varrho > 0, \\ \mathcal{X}_\gamma &:= \bigcup_{\varrho > 0} \mathcal{X}_{\gamma, \varrho} = \mathcal{R}((A^*A)^\gamma), \quad \gamma > 0. \end{aligned}$$

Bei schlecht gestellten Problemen wirkt der Operator  $A$  häufig glättend, das heißt  $Ax$  hat weniger hochfrequente Anteile als  $x$ . Die Forderung  $x \in \mathcal{X}_{\gamma, \varrho}$  kann also als eine Glattheitsbedingung an  $x$  aufgefasst werden. Tautenhahn untersucht in [82] eine allgemeinere Klasse von Teilräumen von  $\mathcal{X}$ , siehe dazu auch den kürzlich erschienenen Artikel [44].

Kommen wir nun aber wieder zurück zu den Räumen  $\mathcal{X}_{\gamma, \varrho}$ . In diesen gilt sogar Gleichheit in (1.6), also

$$E(\delta, \mathcal{X}_{\gamma, \varrho}) = e(\delta, \mathcal{X}_{\gamma, \varrho}).$$

(siehe [73, Satz 3.2.4]). Nun sind wir in der Lage, die Größe des besten schlimmsten Fehlers anzugeben (siehe Lemma 1.6), wohlgermerkt ohne überhaupt ein Rekonstruktionsverfahren zu kennen. Da wir die Beweisidee später bei nichtlinearen schlecht gestellten Problemen wiederverwenden werden, präsentieren wir den Beweis hier in der Version von [20, Proposition 3.14]. Die zum Beweis nötige Interpolationsungleichung

$$(1.8) \quad \|(A^*A)^r x\|_{\mathcal{X}} \leq \|(A^*A)^q x\|_{\mathcal{X}}^{\frac{r}{q}} \|x\|_{\mathcal{X}}^{1-\frac{r}{q}}, \quad q > r \geq 0, \quad x \in \mathcal{X}$$

findet man ebenfalls in [20, Gleichung (2.49)].

**Lemma 1.6.** *Für  $\gamma, \varrho > 0$  sei  $\mathcal{X}_{\gamma, \varrho}$  durch (1.7) definiert. Dann gilt für alle  $\delta > 0$*

$$(1.9) \quad e(\delta, \mathcal{X}_{\gamma, \varrho}) \leq \delta^{\frac{2\gamma}{2\gamma+1}} \varrho^{\frac{1}{2\gamma+1}}.$$

*Beweis.* Sei  $x$  aus  $\mathcal{X}_{\gamma, \varrho}$ . Dann existiert ein  $z \in \mathcal{X}$  mit  $\|z\|_{\mathcal{X}} \leq \varrho$  und  $x = (A^*A)^{\gamma} z$ . Mit Hilfe der Interpolationsungleichung folgt dann

$$\begin{aligned} \|x\|_{\mathcal{X}} &= \|(A^*A)^{\gamma} z\|_{\mathcal{X}} \leq \|(A^*A)^{\gamma+1/2} z\|_{\mathcal{X}}^{\frac{2\gamma}{2\gamma+1}} \|z\|_{\mathcal{X}}^{\frac{1}{2\gamma+1}} \\ &= \|(A^*A)^{1/2} x\|_{\mathcal{X}}^{\frac{2\gamma}{2\gamma+1}} \|z\|_{\mathcal{X}}^{\frac{1}{2\gamma+1}} = \|Ax\|_{\mathcal{Y}}^{\frac{2\gamma}{2\gamma+1}} \|z\|_{\mathcal{X}}^{\frac{1}{2\gamma+1}}. \end{aligned}$$

Die letzte Gleichung folgt dabei aus

$$\|(A^*A)^{1/2} x\|_{\mathcal{X}} = \langle (A^*A)^{1/2} x, (A^*A)^{1/2} x \rangle_{\mathcal{X}} = \langle x, A^*Ax \rangle_{\mathcal{X}} = \langle Ax, Ax \rangle_{\mathcal{Y}} = \|Ax\|_{\mathcal{Y}}.$$

Ist nun zusätzlich  $\|Ax\|_{\mathcal{Y}} \leq \delta$ , so gilt  $\|x\|_{\mathcal{X}} \leq \delta^{\frac{2\gamma}{2\gamma+1}} \varrho^{\frac{1}{2\gamma+1}}$ . Die Behauptung folgt nun indem wir das Supremum über alle  $x \in \mathcal{X}_{\gamma, \varrho}$  mit  $\|Ax\|_{\mathcal{Y}} \leq \delta$  bilden.  $\square$

Der beste schlimmste Fehler geht also mindestens so schnell gegen Null wie  $O(\delta^{2\gamma/(2\gamma+1)})$ . Man kann auch zeigen, dass die obige Abschätzung (1.9) in gewisser Weise scharf ist, das heißt man findet zu  $\gamma, \varrho > 0$  eine positive Nullfolge  $\{\delta_k\}_{k \in \mathbb{N}}$  so, dass  $e(\delta_k, \mathcal{X}_{\gamma, \varrho})$  nicht schneller gegen Null konvergiert als  $\delta_k^{2\gamma/(2\gamma+1)}$ . Beweise für kompakte Operatoren findet man in [20, 73]; ein Beweis für lineare Operatoren folgt mit [19, Proposition 2.3].

Da wir nun wissen, was die beste zu erwartende asymptotische Konvergenzrate ist, können wir definieren was wir unter optimalen Rekonstruktionsverfahren verstehen.

**Definition 1.7.** Eine Familie von Rekonstruktionsverfahren  $\{R_\delta\}_{\delta>0}$  heißt **ordnungsoptimal** bezüglich  $X_{\gamma,\varrho}$  (und  $A$ ), wenn ein  $C_\gamma \geq 1$  existiert so, dass für alle  $\delta > 0$  hinreichend klein und alle  $\varrho \geq 0$  gilt

$$E_\nu(\delta, X_{\gamma,\varrho}, R_\delta) \leq C_\gamma \delta^{\frac{2\gamma}{2\gamma+1}} \varrho^{\frac{1}{2\gamma+1}}.$$

Ist diese Abschätzung mit  $C_\gamma = 1$  erfüllt, so heißt  $\{R_\delta\}_{\delta>0}$  ein **optimales Rekonstruktionsverfahren**.

Da Regularisierungsverfahren auch Rekonstruktionsverfahren sind, gilt diese Definition auch für Regularisierungsverfahren. Es besteht aber auch ein umgekehrter Zusammenhang zwischen den beiden Verfahrensklassen, wie Plato [69] zeigt.

**Definition 1.8.** Das Supremum  $\gamma_0$  über alle  $\gamma > 0$ , für das ein Regularisierungsverfahren  $R$  ordnungsoptimal ist, bezeichnen wir als die **Qualifikation** des Verfahrens.

Ist die Qualifikation endlich, so kann man mit diesem Verfahren keine bessere Konvergenzrate als  $O(\delta^{2\gamma_0/(2\gamma_0+1)})$  erzielen. Bei unendlicher Qualifikation ist die bestmögliche Konvergenzrate  $O(\delta)$  erreichbar. Diese gilt natürlich nur, sofern die exakten Lösungen genügend glatt sind.

Wie sehen nun aber Regularisierungsverfahren aus? Eine elegante Darstellung von Regularisierungsverfahren erhält man mit Hilfe von regularisierenden Filtern .

**Definition 1.9.** Eine Familie  $\{G_t\}_{t>0}$  von stückweise stetigen Funktionen  $G_t : [0, \|A\|^2] \rightarrow \mathbb{R}$  für die eine Konstante  $C > 0$  existiert so, dass

$$|\lambda G_t(\lambda)| \leq C \quad \text{und} \quad \lim_{t \rightarrow 0} G_t(\lambda) = \frac{1}{\lambda}$$

für alle  $\lambda \in (0, \|A\|^2]$  gilt, heißt **regularisierender Filter** (für  $A$ ).

Dass diese Familie von Funktionen ihren Namen „regularisierend“ auch verdient, zeigt das folgende Lemma [73, Korollar 3.3.4].

**Lemma 1.10.** Seien  $\{G_t\}$  ein regularisierender Filter (für  $A$ ) und

$$M(t) := \sup\{|G_t(\lambda)| : \lambda \in [0, \|A\|^2]\}.$$

Gilt ferner für die (a priori) Parameterwahl  $\alpha : (0, \infty) \rightarrow (0, \infty)$

$$\alpha(\delta) \rightarrow 0 \quad \text{und} \quad \delta \sqrt{M(\alpha(\delta))} \quad \text{für} \quad \delta \rightarrow 0,$$

dann ist  $(\{R_t\}_{t>0}, \alpha)$  mit  $R_t := G_t(A^*A)A^*$  ein Regularisierungsverfahren für  $A^\dagger$ .

Wir können nun die Frage nach der Existenz von Regularisierungsverfahren recht leicht beantworten. In der Tat lassen sich viele bekannte Regularisierungsverfahren mit Hilfe von regularisierenden Filtern schreiben.

**Beispiel 1.11.**

1. Zum wohl bekanntesten Regularisierungsverfahren, der Tikhonov-Phillips-Regularisierung, gehört der regularisierende Filter

$$G_t(\lambda) = \frac{1}{\lambda + t}.$$

Es hat die Qualifikation  $\gamma_0 = 2$ .

2. Das relaxierte Landweber-Verfahren (siehe [55]) können wir mit dem Filter

$$G_{1/m}(\lambda) = \frac{1 - (1 - \omega\lambda)^m}{\lambda}$$

beschreiben. Dabei ist  $\omega < 2/\|A\|^2$  der Relaxationsparameter. Da dies kein kontinuierliches, sondern ein iteratives Verfahren ist, können wir den Filter natürlich auch nur für  $\frac{1}{m}$  mit  $m \in \mathbb{N}$  definieren. Das Landweber-Verfahren hat unendliche Qualifikation  $\gamma_0 = \infty$ .

3. Der Filter zur abgeschnittenen Singulärwertzerlegung lautet

$$G_t(\lambda) = \begin{cases} 1/\lambda, & \lambda \geq t, \\ 0, & \lambda < t. \end{cases}$$

Auch die abgeschnittene Singulärwertzerlegung hat unendliche Qualifikation  $\gamma_0 = \infty$ .

Regularisierungsverfahren bestehen nicht nur aus einer Familie von Operatoren, sondern auch aus einer Parameterwahl. Im folgenden Abschnitt beschäftigen wir uns mit dieser.

## 1.2 Diskrepanzprinzip

Wir können den Rekonstruktionsfehler  $\|A^\dagger g - R_t y^\delta\|_X$  in Approximationsfehler und Datenfehler aufspalten, also

$$\|A^\dagger y - R_t y^\delta\|_X \leq \|A^\dagger y - R_t y\|_X + \|R_t(y - y^\delta)\|_X.$$

Die einzelnen Fehler zeigen dann typischerweise das in Abbildung 1.1 dargestellte Verhalten. Da der Rekonstruktionsfehler sowohl für  $t \rightarrow 0$  als auch für  $t \rightarrow \infty$  explodiert, stellt sich die Frage nach der Wahl des optimalen Regularisierungsparameters  $t_{opt}$ , der den Rekonstruktionsfehler minimieren soll. Diesen Parameter zu finden ist die Aufgabe der Parameterwahl.

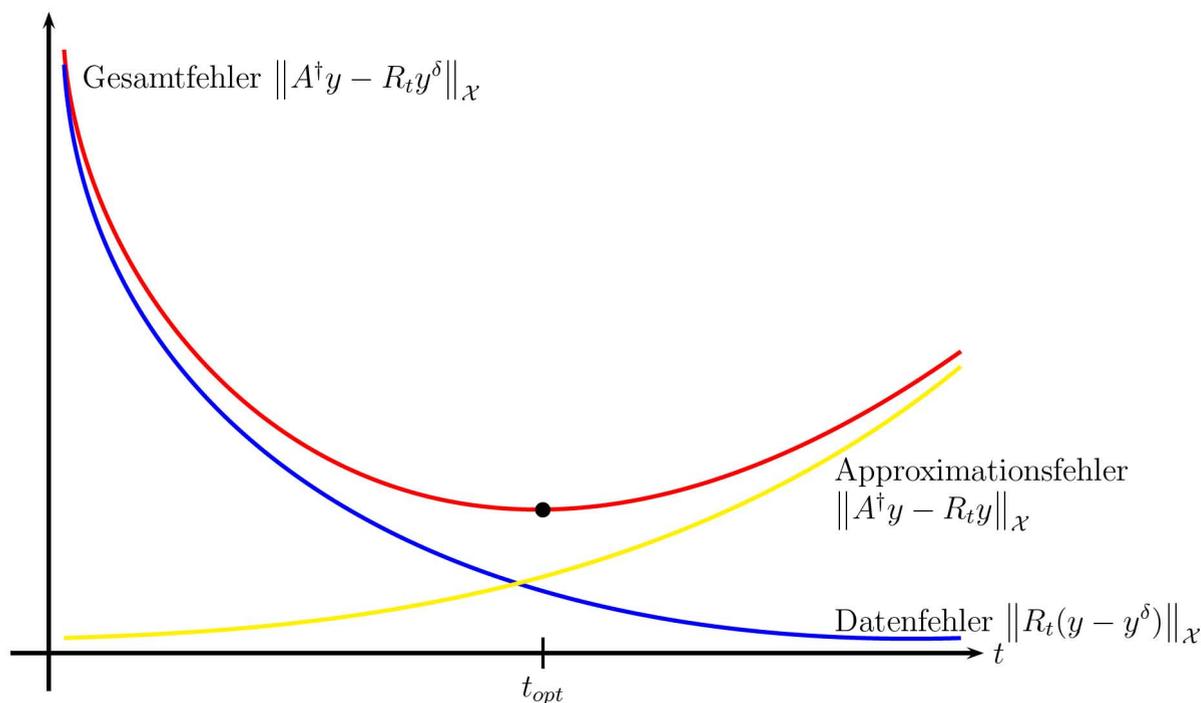


Abbildung 1.1: Zusammensetzung des Rekonstruktionsfehlers aus Approximationsfehler und Datenfehler.

Es gibt eine ganze Reihe von etablierten Parameterwahlen. Im Allgemeinen ist  $\gamma$  mit  $A^\dagger x \in \mathcal{X}_\gamma$  nicht bekannt. Man kann daher keine a priori Parameterwahl so konstruieren, dass das Regularisierungsverfahren optimale Ordnung hat. Es gibt in der Literatur eine

ganze Reihe von Vorschlägen zur Wahl des (a posteriori) Regularisierungsparameters (siehe [4, 23, 24, 28, 33, 59, 63, 86]). Die Ansätze von [24, 33, 86] kommen dabei sogar ohne die Kenntnis von  $\delta$  aus, wodurch jedoch der Nachweis der regularisierenden Eigenschaften deutlich erschwert wird. Wir konzentrieren uns auf die a posteriori Parameterwahl wie sie Morozov [63] vorgeschlagen hat.

Da uns zur Lösung von (1.1) anstatt der exakten rechten Seite  $y$  nur verrauschte Daten  $y^\delta$  mit (1.3) zur Verfügung stehen, macht es keinen Sinn, von einer approximativen Lösung eine höhere Genauigkeit zu verlangen als dieses  $\delta$ . Ziel ist es also, den Regularisierungsparameter so zu wählen, dass

$$\|AR_{\alpha(\delta, y^\delta)}y^\delta - y^\delta\|_{\mathcal{Y}} \approx \delta$$

gilt. Da der Defekt  $\lim_{t \searrow 0} \|Af_t^\delta - y^\delta\|_{\mathcal{Y}} \leq \delta$  (vgl. [73, Abschnitt 3.4]) erfüllt, existiert zu jedem  $\tau > 1$  ein  $t_0 > 0$  mit  $\|AR_t y^\delta - y^\delta\|_{\mathcal{Y}} \leq \tau\delta$  für  $0 < t < t_0$ . Das folgende Diskrepanzprinzip von Morozov ist also wohldefiniert.

**Definition 1.12.** Seien  $\{t_k\}_{k \in \mathbb{N}}$  eine monoton fallende Nullfolge und  $\tau > 1$  fest gewählt. Bestimme  $k^*$  so, dass

$$\|AR_{t_{k^*}}y^\delta - y^\delta\|_{\mathcal{Y}} \leq \tau\delta < \|AR_{t_i}y^\delta - y^\delta\|_{\mathcal{Y}} \quad \text{für } i = 1, \dots, k^* - 1.$$

Die Parameterwahl

$$\alpha(\delta, y^\delta) = \alpha(\delta, y^\delta, \tau) := t_{k^*}$$

heißt **Diskrepanzprinzip** (von Morozov).

Das Diskrepanzprinzip liefert zusammen mit geeigneten Regularisierungsoperatoren ordnungsoptimale Regularisierungsverfahren (vgl. [73, Satz 3.4.1] oder [20, Theorem 4.17]).

**Satz 1.13.** Der Filter  $\{G_t\}_{t>0}$  sei regularisierend mit Qualifikation  $\mu_0 > 1$  und es sei  $tM(t) \leq C_M$  für  $t \rightarrow 0$ . Die Parameterwahl erfolge nach dem Diskrepanzprinzip (Definition 1.12). Außerdem sei

$$\tau > S_p := \sup \{1 - \lambda G_t(\lambda) : t > 0, 0 \leq \lambda \leq \|A\|^2\} \geq 1.$$

Dann ist  $(\{R_t\}, \alpha)$  mit  $R_t = G_t(A^*A)A^*$  ein ordnungsoptimales Regularisierungsverfahren für  $A^\dagger$  bzgl.  $X_{\mu, \varrho}$  für alle  $\mu \in ]0, \mu_0 - 1/2]$ .

**Bemerkung.** Der Satz 1.13 gilt auch für diskrete Filter  $\{G_{t_n}\}_{n \in \mathbb{N}}$ , sofern die Folge  $\{t_n\}_{n \in \mathbb{N}}$  die Eigenschaft

$$(1.10) \quad t_n = Q_n t_{n-1} \quad \text{für } n \geq 1 \text{ mit } 0 < \theta \leq Q_n < 1$$

besitzt. ◇

### 1.3 Asymptotische Regularisierung

Showalter hat in [79] ein Verfahren zur Approximation der verallgemeinerten Inversen einer linearen Abbildung vorgestellt. Er zeigt in [79, Theorem I], dass

$$(1.11) \quad \lim_{t \rightarrow \infty} \int_0^t \exp(-(t-s)A^*A)A^* ds = \lim_{t \rightarrow \infty} t\varphi(-tA^*A)A^* = A^\dagger$$

sofern  $A^\dagger$  beschränkt ist. Hierbei ist die analytische  $\varphi$ -Funktion gegeben durch

$$(1.12) \quad \varphi(\lambda) = \frac{e^\lambda - 1}{\lambda}.$$

Für  $\lambda \rightarrow 0$  erfüllt diese

$$\varphi(\lambda) = 1 + \frac{\lambda}{2!} + \frac{\lambda^2}{3!} + \dots$$

Bei schlecht gestellten Problemen, wie wir sie betrachten, haben wir es allerdings mit einer unbeschränkten verallgemeinerten Inversen zu tun, weshalb uns das Resultat von Showalter nicht weiterhilft. Wir können den Showalter Ansatz aber auch mit Hilfe eines Anfangswertproblems definieren, denn das Integral in (1.11) ist im Wesentlichen der Lösungsoperator des Anfangswertproblems

$$(1.13) \quad \begin{aligned} \frac{d}{dt}u(t) &= A^*(y^\delta - Au(t)), \quad \text{für } t > 0, \\ u(0) &= 0. \end{aligned}$$

Aufgrund der Linearität und Beschränktheit von  $A^*A$  ist (1.13) eindeutig lösbar (siehe [68, Chapter 4, Corollary 1.5]). Ferner konvergiert die Lösung  $u(t)$  von (1.13) für  $t \rightarrow \infty$  gegen einen stationären Zustand  $u_\infty$ , welcher offensichtlich die Normalgleichung  $A^*Au_\infty = A^*y^\delta$  erfüllt. Als differenzierbare Funktion ist die Lösung  $u(t)$  stetig und wegen  $u(0) = 0$  erhalten wir eine Familie von Rekonstruktionsverfahren  $\{R_t\}$  aus

$$R_t y^\delta := u\left(\frac{1}{t}\right).$$

Mit Hilfe der Variation-der-Konstanten-Formel [35] können wir  $u$  berechnen als

$$u(t) = t\varphi(-tA^*A)A^*y^\delta$$

mit der  $\varphi$ -Funktion aus (1.12). Daran können wir auch direkt den zugehörigen Filter ablesen

$$G_t(\lambda) = \frac{1}{t}\varphi\left(-\frac{\lambda}{t}\right),$$

welchen wir nach genauem Betrachten als regularisierend erkennen. Die unendliche Qualifikation und dass  $tM(t) \leq 1$  gilt, haben wir in [47] nachgewiesen. Zusammen mit dem Diskrepanzprinzip erhalten wir daher mit Hilfe von Satz 1.13 ein ordnungsoptimales Regularisierungsverfahren (für  $A^\dagger$ ) mit unendlicher Qualifikation. Dieser als *asymptotische Regularisierung* bezeichnete Ansatz taucht unseres Wissens nach zum ersten mal bei Vainikko [83] in der englischsprachigen Literatur auf (zuvor ist er schon im russischsprachigen Raum bekannt gewesen; siehe dazu die in [83] zitierte Literatur).

## 1.4 Regularisierung durch numerische Integratoren

Zur numerischen Umsetzung der asymptotischen Regularisierung, also zur Lösung des Anfangswertproblems (1.13), setzt man Integrationsverfahren ein. Zu den Bekanntesten unter diesen Verfahren zählen die Runge-Kutta-Verfahren. Rieder untersucht in [75] die regularisierende Wirkung dieser Methode.

Ein  $s$ -stufiges Runge-Kutta-Verfahren zur Lösung von (1.13) ist gegeben durch

$$(1.14) \quad \begin{aligned} u_1 &= h_0 \sum_{i=1}^s b_i U'_i, \\ U'_i &= A^*(y^\delta - AU_i), \\ U_i &= h_0 \sum_{j=1}^s a_{ij} U'_j. \end{aligned}$$

Ein Runge-Kutta-Verfahren ist explizit, sofern die Matrix  $\mathcal{A} = (a_{ij})_{i,j=1}^s \in \mathbb{R}^{s,s}$  nur Einträge unterhalb der Diagonalen hat. Sei  $b = (b_i)_{i=1}^s$  der aus den Koeffizienten  $b_i$  des Integrators bestehende Vektor. Bezeichnen wir dann mit

$$S(z) = 1 + zb^*(I_s - z\mathcal{A})\mathbb{1}, \quad \mathbb{1} = (1, \dots, 1)^* \in \mathbb{R}^s$$

die zum obigen Verfahren gehörende Stabilitätsfunktion (siehe [27, Definition 2.1]), so ist

$$(1.15) \quad G_{1/t_n}(\lambda) = \frac{1 - \prod_{k=1}^n S(-h_k \lambda)}{\lambda}$$

der Filter des resultierenden Regularisierungsverfahrens [75]. Dabei ist  $h_k$  die Zeitschrittweite im  $k$ -ten Schritt des Integrators und  $t_n = \sum_{k=1}^n h_k$  die Integrationszeit. Rieder konnte nun in [75, Theorem 3.2] zeigen, dass zu jedem Integrator eine maximale Schrittweite  $h_{\max}$  existiert so, dass die nach dem Diskrepanzprinzip gestoppte Integrationsmethode ein ordnungsoptimales Regularisierungsverfahren mit unendlicher Qualifikation ist, sofern die Schrittweiten von Null weg beschränkt bleiben. Die Beschränkung der maximalen Schrittweite rührt dabei aus der Stabilität des Verfahrens. Bei A-stabilen Verfahren kann die Bedingung der maximalen Schrittweite wegfallen; man darf daher in diesem Fall unabhängig vom Integrator große Zeitschrittweiten wählen.

**Satz 1.14.** [75, Theorem 3.3] *Sei  $S$  die Stabilitätsfunktion eines A-stabilen Runge-Kutta-Verfahrens, welches*

$$|S(-z)| < 1 \quad \text{für alle } z > 0$$

*erfüllt. Sind ferner die Zeitschrittweiten  $h_k$  nach unten von der Null weg und nach oben beschränkt, so ist das Regularisierungsverfahren mit dem Filter (1.15) ordnungsoptimal und von unendlicher Qualifikation, wenn es nach dem Diskrepanzprinzip gestoppt wird.*

**Bemerkung.** Die Aussage von Rieder lässt uns einige Regularisierungsverfahren für lineare Probleme in neuem Licht erscheinen. Verwendet man beispielsweise das explizite oder implizite Euler-Verfahren zur Lösung des Showalter-Anfangswertproblems (1.13), so erhält man das schon bekannte relaxierte Landweber-Verfahren bzw. die Tikhonov-Philips-Regularisierung. Unter diesem Aspekt erkennen wir, dass die Beschränkung des Relaxationsparameters  $\omega < 2/\|A\|^2$  im Landweber-Verfahren dem beschränkten Stabilitätsbereich  $B_1(-1)$  des expliziten Euler-Verfahrens geschuldet ist.  $\diamond$

In [47] konnten wir diese Aussage auf exponentielle Integratoren, siehe Anhang A, erweitern. Wir betrachten hier nur den Spezialfall des exponentiellen Euler-Verfahrens angewandt auf (1.13):

$$(1.16) \quad u_{n+1} = u_n + h_n \varphi(-h_n A^* A) A^* (y^\delta - Au_n).$$

In [9] und [10] wird dieser Ansatz für große diskrete schlecht gestellte Probleme untersucht und die Konvergenz gezeigt. Eine Aussage über die Konvergenzrate ist dort aber nicht zu finden.

**Satz 1.15.** *Sind die Schrittweiten  $h_n$  durch eine positive Konstante  $c$  nach unten beschränkt, gilt also  $h_n \geq c > 0$  für alle  $n \in \mathbb{N}_0$ , so liefert das exponentielle Euler-Verfahren angewandt auf das Anfangswertproblem (1.3) zusammen mit dem Diskrepanzprinzip ein ordnungsoptimales Regularisierungsverfahren mit unendlicher Qualifikation.*

*Beweis.* Da das exponentielle Euler-Verfahren lineare Anfangswertprobleme exakt löst, entspricht das zugehörige Regularisierungsverfahren der asymptotischen Regularisierung, deren Eigenschaften oben schon besprochen wurden. Durch die Beschränkung der Zeitschrittweite nach unten erfüllt das Verfahren auch (1.10) und der Satz ist bewiesen.  $\square$

## 1.5 Implementierung

In diesem Abschnitt beschäftigen wir uns mit der Umsetzung des exponentiellen Euler-Verfahrens auf dem Rechner. Wir gehen davon aus eine diskrete Version von (1.4) zu haben. Auf die Art der Diskretisierung gehen wir bei den einzelnen numerischen Beispielen ein. Die diskreten Variablen schreiben wir im Folgenden fettgedruckt. Sei nun

$$(1.17) \quad \begin{aligned} \frac{d}{dt} \mathbf{u}(t) &= \mathbf{A}^* (\mathbf{y}^\delta + \mathbf{A} \mathbf{u}(t)), \quad \text{für } t > 0, \\ \mathbf{u}(0) &= \mathbf{0}. \end{aligned}$$

die diskrete Version von (1.13).

Der Algorithmus zur Lösung dieses diskreten Anfangswertproblems ist in Algorithmus 1 schematisch dargestellt. Als Eingabeparameter benötigt dieser Algorithmus  $\mathbf{y}^\delta$ , eine Startschrittweite  $h_c$ , den Rauschpegel  $\delta$  und  $\tau$  aus dem Diskrepanzprinzip (Definition 1.12). Ferner muss eine Funktion bereitgestellt werden, die die Matrix-Vektor-Produkte  $\mathbf{A} \mathbf{v}$  bzw.  $\mathbf{A}^* \mathbf{v}$  berechnet. Es ist dabei nicht nötig, die Matrizen  $\mathbf{A}$  und  $\mathbf{A}^*$  explizit zu kennen. Es ist daher eine matrixfreie Implementierung möglich.

Um die Matrix-Vektor-Multiplikation im Diskrepanzprinzip einzusparen nutzen wir

$$\|\mathbf{A} \mathbf{u}(t) - \mathbf{y}^\delta\| = \|t \mathbf{A} \mathbf{A}^* \varphi(-t \mathbf{A} \mathbf{A}^*) \mathbf{y}^\delta - \mathbf{y}^\delta\| = \|\exp(-t \mathbf{A} \mathbf{A}^*) \mathbf{y}^\delta\|$$

und approximieren daher zunächst  $\|\exp(-t \mathbf{A} \mathbf{A}^*) \mathbf{y}^\delta\|$ . Dies erledigen wir mittels Krylov-Verfahren. Eine kurze Erläuterung dieser Methoden ist in Anhang B zu finden. Die Genauigkeit der Krylov-Approximation

$$(1.18) \quad \mathbf{V}_m \exp(-t_1 \mathbf{H}_m) \beta \mathbf{e}_1 \approx \exp(-t_1 \mathbf{A} \mathbf{A}^*) \mathbf{y}^\delta$$

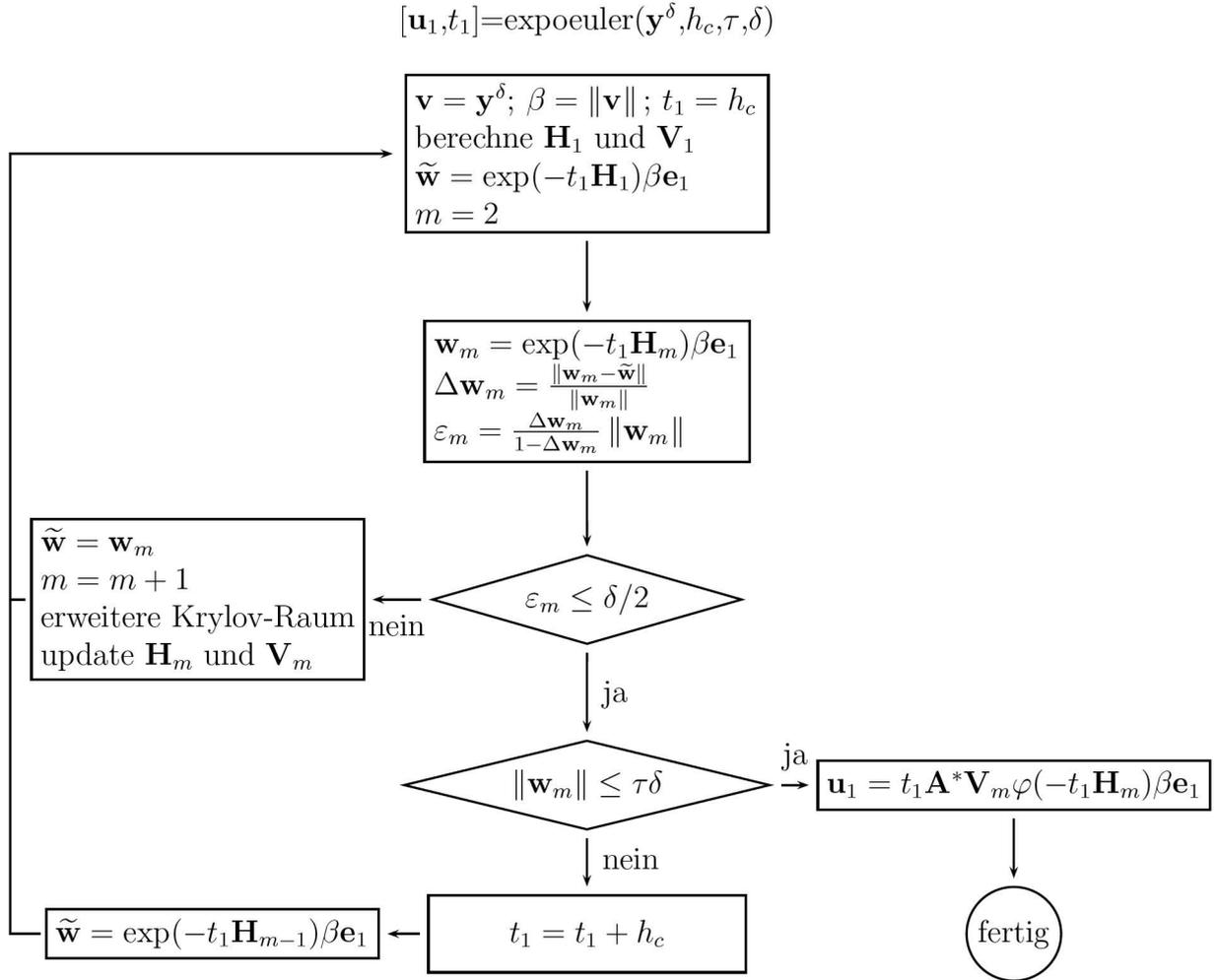
mit  $\beta = \|\mathbf{y}^\delta\|$  schätzen wir dabei mit Hilfe des Kriteriums von Hochbruck und van den Eshof [84]. Wir approximieren den Fehler  $\varepsilon_m$  durch

$$\varepsilon_m \lesssim \frac{\Delta_m}{1 - \Delta_m} \|\mathbf{w}_m\|,$$

wobei  $\mathbf{w}_m := \exp(-h \mathbf{H}_m) \beta \mathbf{e}_1$  und

$$\Delta_m = \frac{\|\mathbf{w}_m - \mathbf{w}_{m-k}\|}{\|\mathbf{w}_m\|} \quad \text{für } k \geq 1,$$

eine Näherung des relativen Fehlers im  $m$ -ten Krylov-Schritt ist. In numerischen Experimenten hat sich herausgestellt, dass es genügt, mit einer Genauigkeit in der Größenordnung von  $\delta$  zu rechnen. Ferner hat es sich in unseren numerischen Experimenten als sinnvoll erwiesen die Approximation für jedes  $m$  zu berechnen und nicht nur für bestimmte, wie es

**Algorithmus 1** Exponentielles Euler-Verfahren für lineare Probleme

in [39] vorgeschlagen wurde. Hat die Approximation (1.18) die gewünschte Genauigkeit, so überprüfen wir das Diskrepanzprinzip.

Ist das Diskrepanzprinzip schließlich erfüllt, so berechnen wir

$$\mathbf{u}_1 = t_1 \mathbf{A}^* \mathbf{V}_m \varphi(-t_1 \mathbf{H}_m) \beta \mathbf{e}_1.$$

Falls das Diskrepanzprinzip nicht erfüllt ist, erhöhen wir die Schrittweite. Wir beschränken die Dimension des Krylov-Raums auf  $m_{\max} = 100$ . Ist das Diskrepanzprinzip in diesem Krylov-Raum nicht erfüllbar, so starten wir den Prozess mit  $\mathbf{y}^\delta - \mathbf{A} \mathbf{u}_1$  statt  $\mathbf{y}^\delta$  neu.

Wurde die Zeitschrittweite im letzten Schritt zu groß gewählt, so kann es passieren, dass wir den Stoppzeitpunkt nicht optimal treffen. In diesem Fall liegt die Diskrepanz

$\|\mathbf{A}\mathbf{u}_1 - \mathbf{y}^\delta\| > \tau\delta$  deutlich unter  $\tau\delta$ . Um diesen Fehler zu korrigieren, führen wir am Ende des Algorithmus eine Monitoringstrategie durch (vgl. [75, Abschnitt 4.2]). Dabei reduzieren wir die Schrittweite im letzten Zeitschritt so lange, bis die Diskrepanz größer als  $0.9\tau\delta$  ist. Eine Pseudocodeformulierung ist in Algorithmus 2 zu finden.

Zum Vergleich des Verfahrens haben wir das von Hanke in [29] untersuchte **cgnr**-Verfahren, welches dem cg-Verfahren angewandt auf die Normalengleichung (1.2) entspricht, implementiert (in [29] heißt das Verfahren **cgne**). Wir stellen es in Algorithmus 3 dar.

---

**Algorithmus 2** Monitoringstrategie für das exponentielle Euler-Verfahren
 

---

```

while  $\|\mathbf{A}\mathbf{u}_1 - \mathbf{y}^\delta\| \leq 0.9\tau\delta$  do
   $t_1 = 0.9t_1$ 
   $\mathbf{d} = \mathbf{V}_m\varphi(-t_1\mathbf{H}_m)\beta\mathbf{e}_1$ 
   $\mathbf{u}_1 = \mathbf{u}_0 + t_1\mathbf{d}$ 
end while

```

---



---

**Algorithmus 3** **cgnr**


---

```

 $[\mathbf{u}] = \text{cgnr}(\mathbf{y}^\delta, \tau\delta)$ 
 $m = 0; \mathbf{u} = \mathbf{0}; \mathbf{r} = \mathbf{y}^\delta; \mathbf{d} = \tilde{\mathbf{r}}_0 = \mathbf{A}^*\mathbf{r}$ 
while  $\|\mathbf{A}\mathbf{u} - \mathbf{y}^\delta\| > \tau\delta$  do
   $\tilde{\mathbf{d}} = \mathbf{A}\mathbf{d}$ 
   $\alpha = \|\tilde{\mathbf{r}}_m\|^2 / \|\tilde{\mathbf{d}}\|^2$ 
   $\mathbf{u} = \mathbf{u} + \alpha\mathbf{d}$ 
   $\mathbf{r}_m = \mathbf{r}_m + \alpha\tilde{\mathbf{d}}$ 
   $\tilde{\mathbf{r}}_{m+1} = \mathbf{A}^*\mathbf{r}$ 
   $\beta = \|\tilde{\mathbf{r}}_{m+1}\|^2 / \|\tilde{\mathbf{r}}_m\|^2$ 
   $\mathbf{d} = \tilde{\mathbf{r}}_{m+1} + \beta\mathbf{d}$ 
   $m = m + 1$ 
end while

```

---

## 1.6 Numerische Beispiele

Nachdem wir nun die Details der Implementierung besprochen haben, möchten wir das exponentielle Euler-Verfahren anhand zweier Verfahren testen.

Alle numerischen Experimente dieser Arbeit wurden mit MATLAB Release 2006a auf einem Intel® Core™ 2 Duo 6400 PC mit 2.13 GHz und 2 Gigabyte Hauptspeicher gerechnet.

### 1.6.1 Differentiation

Das erste Beispiel ist [73, Abschnitt 6.1.3.1] entnommen. Wir betrachten die Fredholm-Integralgleichung erster Art

$$(1.19) \quad Ax = y,$$

wobei  $A : \mathcal{X} \rightarrow \mathcal{X}$ ,  $\mathcal{X} = L^2([0, 1])$ , der Integraloperator

$$Ax(\xi) := \int_0^1 k(\xi, \eta)x(\eta) \, d\eta$$

mit Kern

$$k(\xi, \eta) = \begin{cases} \xi - \eta, & \xi \geq \eta, \\ 0, & \xi < \eta, \end{cases}$$

ist. Da  $k \in L^2([0, 1]^2)$  ist, ist  $A$  ein kompakter Operator [87, Beispiel II.3(c)]. Ferner gilt  $Ax = y$  genau dann, wenn  $x = y''$  und  $y(0) = y'(0) = 0$  ist. Deshalb besteht der Kern von  $A$  nur aus der 0.

Speziell sei nun die rechte Seite  $y$  durch

$$y(\xi) = \frac{\xi^2(\xi^2 - 4\xi + 9)}{12} + \frac{1 - \cos(3\pi\xi)}{18\pi^2} \quad \text{für } \xi \in [0, 1]$$

gegeben. Die zugehörige Lösung  $x^\dagger$  lautet

$$x^\dagger(\xi) = (1 - \xi)^2 + \cos^2\left(\frac{3\pi\xi}{2}\right) \quad \text{für } \xi \in [0, 1].$$

Da  $x^\dagger = A^*w$  mit  $w(\xi) = 2 - 9\pi^2 \cos(3\pi\xi)/2$  ist, haben wir  $x^\dagger \in \mathcal{R}(A^*) = \mathcal{R}((A^*A)^{\frac{1}{2}}) = \mathcal{X}_1$ . Damit sollten wir eine Konvergenzrate von  $O(\sqrt{\delta})$  erwarten können.

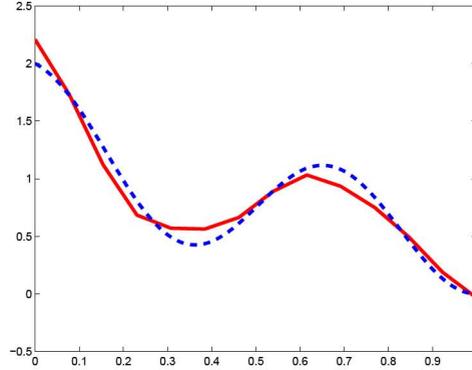


Abbildung 1.2: Exakte Lösung  $x^\dagger$  (gestrichelte Linie) und die Rekonstruktion  $u_{14,t}$  mittels des exponentiellen Euler-Verfahrens bei  $\delta = 0.0116$  und  $t = 10^4$  (durchgezogene Linie).

Bevor wir das Problem auf dem Rechner umsetzen können, müssen es wir noch geeignet diskretisieren. Dazu zerlegen wir das Intervall  $[0, 1]$  äquidistant in  $l - 1$  Teilintervalle der Länge  $h_l = \frac{1}{l-1}$ . Die zugehörigen Knoten bezeichnen wir mit

$$\xi_{l,j} = (j - 1)h_l, \quad \text{für } j = 1, \dots, l.$$

Als Ansatzraum  $\mathcal{X}^{(l)} \subset \mathcal{X}$  wählen wir dann den von den Hutfunktionen  $\phi_{l,i}$ , ( $i = 1, \dots, l$ ), gegeben durch

$$\phi_{l,i}(\xi_{l,j}) = \delta_{i,j},$$

aufgespannten  $l$ -dimensionalen Raum der stetigen, stückweise linearen Funktionen. Durch den Index  $l$  kennzeichnen wir im Folgenden, in welchem Raum  $\mathcal{X}^{(l)}$  wir uns befinden. Die Gramschen Matrizen sind

$$\mathbf{G}_l = (\langle \phi_{l,i}, \phi_{l,j} \rangle_{\mathcal{X}})_{i,j=1}^l = \mathbf{G}_l = \frac{h_l}{6} \begin{pmatrix} 2 & 1 & 0 & \cdots & 0 \\ 1 & 4 & \ddots & \ddots & \vdots \\ 0 & \ddots & \ddots & \ddots & 0 \\ \vdots & \ddots & \ddots & 4 & 1 \\ 0 & \cdots & 0 & 1 & 2 \end{pmatrix} \in \mathbb{R}^{l \times l}.$$

Wir gehen von der realistischen Situation aus, die verrauschten Daten  $y^\delta$  nur an den Knoten  $\xi_{l,j}$ ,  $j = 1, \dots, l$ , zu kennen und interpolieren diese daher linear, um die kontinuierliche Theorie anwenden zu können. Wir haben also

$$g_l^\delta = \sum_{j=1}^l (g(\xi_{l,j}) + \varepsilon_l(\xi_{l,j})) \phi_{l,j}$$

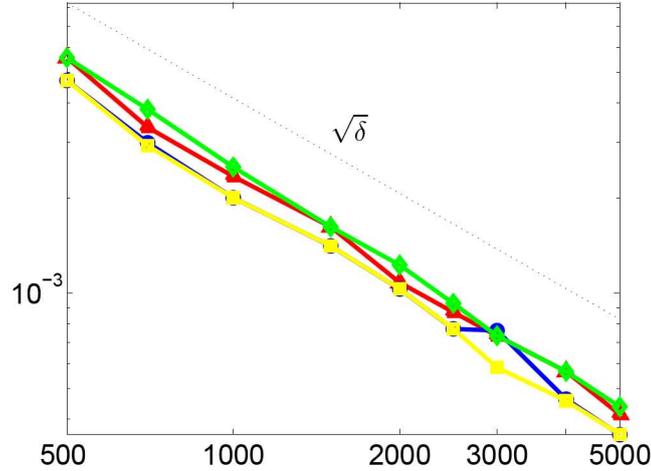


Abbildung 1.3: Relativer  $L^2$ -Regularisierungsfehler und erwartetes Abklingverhalten  $\sqrt{\delta_l}$  (schwarz gepunktet) als Funktion der Dimension  $l$ . Exponentieller Euler mit/ohne Monitoring (rot/grün), `cgnr` mit/ohne Reorthogonalisierung (blau/gelb).

mit einer in  $[-\varepsilon_l, \varepsilon_l]$ ,  $\varepsilon_l > 0$  gleichverteilten Zufallsvariablen  $\varepsilon_l : [0, 1] \rightarrow [-\varepsilon_l, \varepsilon_l]$ . Die Störung in den Daten besteht daher nicht nur aus dem Messrauschen, sondern auch aus dem Interpolationsfehler und es gilt die folgende Abschätzung

$$\begin{aligned} \|y - y_l^\delta\|_{L^2(0,1)} &\leq \left\| y - \sum_{j=1}^l y(\xi_{l,j}) \varphi_{l,j} \right\|_{L^2(0,1)} + \left\| \sum_{j=1}^l \varepsilon_l(\xi_{l,j}) \varphi_{l,j} \right\|_{L^2(0,1)} \\ &\leq h_l^2 \|y''\|_{L^2(0,1)} + \varepsilon_l =: \delta_l. \end{aligned}$$

Mit  $\varepsilon_l = h_l^2$  müssten wir also trotz der Diskretisierung

$$\|x^\dagger - x_{l,t_k^*}^\delta\|_{L^2(0,1)} = O(h_l) = O(\sqrt{\delta_l}) \quad \text{für } l \rightarrow \infty$$

beobachten können. Den relativen Rekonstruktionsfehler der verwendeten Verfahren berechnen wir nicht exakt, sondern mit Hilfe von Trapezsummen. Da diese Approximation eine Genauigkeit von  $O(h_l^2)$  hat, fällt dieser Fehler bei der Beobachtung des Rekonstruktionsfehlers nicht ins Gewicht.

In numerischen Experimenten haben wir das exponentielle Euler-Verfahren mit dem vorzeitig gestoppten `cgnr`-Verfahren aus [29] verglichen. Dabei haben wir jeweils  $\tau = 1.01$  und als Startzeitschrittweite  $t = 10 \cdot l^{3/2}$  gewählt. Die Ergebnisse spiegeln das erwartete Abklingverhalten des relativen Rekonstruktionsfehlers wieder. Abbildung 1.3 zeigt, dass der Fehler des

**cgnr**-Verfahrens minimal geringer ist als der Fehler des exponentiellen Euler-Verfahrens. Dabei macht es kaum einen Unterschied ob wir beim exponentiellen Euler-Verfahren die Monitoringstrategie (siehe Algorithmus 2) einsetzen oder nicht.

Ein Blick auf die Dimension der jeweils benötigten Krylov-Räume (siehe Abbildung 1.4) zeigt, dass in fast allen Fällen die beiden exponentiellen Euler-Verfahren mit Krylov-Räumen geringerer Dimension auskommen als das **cgnr**-Verfahren ohne Reorthogonalisierung. Dies steht eigentlich im Widerspruch dazu, dass das **cgnr**-Verfahren das Residuum im Krylov-Raum minimiert und daher als erstes das Diskrepanzprinzip erfüllen müsste [29, Seite 16]. Durch den Verlust der Orthogonalität im **cgnr**-Verfahren (siehe [60] und die dort zitierte Literatur) geht diese Eigenschaft allerdings verloren. Wir haben daher auch das **cgnr**-Verfahren mit Reorthogonalisierung implementiert. In Abbildung 1.4 (oben) erkennt man, dass das so modifizierte **cgnr**-Verfahren mit geringfügig kleineren Krylov-Räumen auskommt als die exponentiellen Euler-Verfahren.

Die kleineren Krylov-Räume spiegeln sich auch im Vergleich der Rechenzeiten wieder (Abbildung 1.4). So ist das exponentielle Euler-Verfahren bei kleinem  $\delta$  und den damit einhergehenden großen räumlichen Diskretisierungen um bis zu 40% schneller als **cgnr** ohne Reorthogonalisierung. Der Vergleich mit dem **cgnr**-Verfahren mit Reorthogonalisierung zeigt, dass die exponentiellen Euler-Verfahren hinsichtlich der Rechenzeit mithalten können.

Werfen wir zum Schluss noch einen Blick auf unsere Monitoringstrategie aus Algorithmus 2. In Abbildung 1.4 (unten) ist deutlich zu erkennen, dass wir durch diese Strategie den Stoppzeitpunkt der Zeitintegration so steuern können, dass er nahe am gewünschten Stoppzeitpunkt  $t_{\text{opt}}$  mit  $\|Ay(t_{\text{opt}}) - y^\delta\| = \tau\delta$  liegt. Ferner liefert das exponentielle Euler-Verfahren in der Regel einen besseren Stoppzeitpunkt als die **cgnr**-Verfahren.

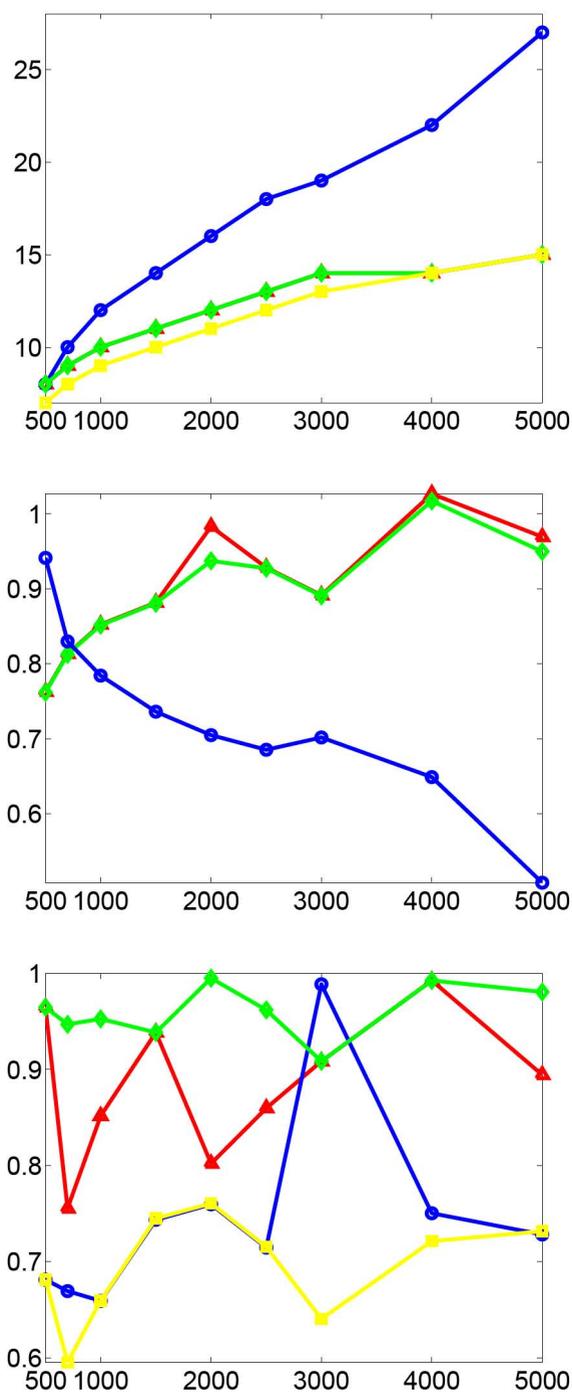


Abbildung 1.4: Dimension der Krylov-Räume (oben), Speedup gegenüber `cgnr` mit Reorthogonalisierung (mittig) und  $\|AR_{l,t_k} y^\delta - y^\delta\| / (\tau\delta)$  (unten) jeweils als Funktion der Dimension. Exponentieller Euler mit/ohne Monitoring (rot/grün), `cgnr` mit/ohne Reorthogonalisierung (blau/gelb).

### 1.6.2 Geologische Prospektion

Das zweite numerische Beispiel ist eine Vereinfachung eines Problems aus der Geologie (vergleiche [25, 34]). Dabei möchte man von Messungen auf der Erdoberfläche auf den Ort, die Form und die Zusammensetzung unterirdischer Körper schließen. Wir beschränken uns auf ein stark vereinfachtes eindimensionales Problem auf dem Intervall  $\Omega = [0, 1]$ . Wir gehen davon aus, dass auf einer Parallelen zur horizontalen (Mess-)Linie eine zu bestimmende Masseverteilung  $x$  vorliegt. Auf der Messlinie wird die vertikale Komponente der Gravitationskraft  $y$  gemessen. Der Zusammenhang zwischen dieser Messung und der tatsächlichen Masseverteilung  $x$  wird, wie schon im Beispiel aus Abschnitt 1.6.1, durch eine Fredholm-Integralgleichung erster Art (1.19) beschrieben. Der Kern  $k$  des Integraloperators  $A : L^2([0, 1]) \rightarrow L^2([0, 1])$  ist dieses Mal gegeben durch die Funktion

$$k(\xi, \eta) = d(d^2 + (\xi - \eta)^2)^{-3/2} \quad \text{für } (\xi, \eta) \in [0, 1]^2.$$

Der Parameter  $d$  beschreibt dabei den Abstand zwischen der Messlinie und der Parallelen mit der Masseverteilung  $x$ . In unseren Experimenten setzen wir ihn auf  $d = 0.25$ . Als exakte Lösung wählen wir

$$x^\dagger(\xi) = \sin(\pi\xi) + \frac{\sin(2\pi\xi)}{2}$$

und berechnen die zugehörige rechte Seite numerisch. Dazu diskretisieren wir das Problem wieder wie in Abschnitt 1.6.1.

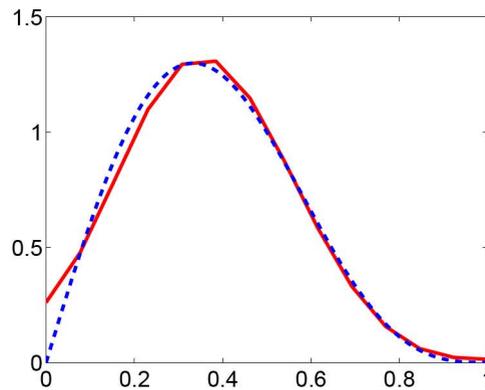


Abbildung 1.5: Exakte Lösung  $x^\dagger$  (gestrichelte Linie) und die Rekonstruktion  $u_{14,t}$  mittels des exponentiellen Euler-Verfahrens bei  $\delta = 0.2972$  und  $t = 1.96$  (durchgezogene Linie).

Wie zuvor vergleichen wir das exponentielle Euler-Verfahren mit dem vorzeitig gestoppten **cgnr**-Verfahren und setzen  $\tau = 1.01$  und die Startzeitschrittweite  $t = l^2/100$ . In Abbildung

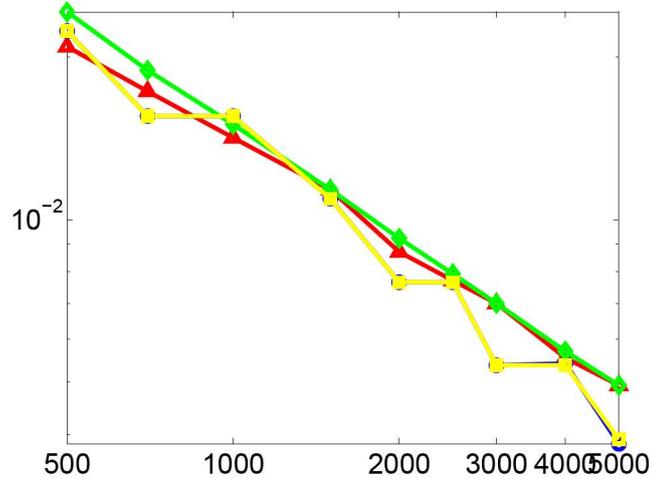


Abbildung 1.6: Relativer  $L^2$ -Regularisierungsfehler als Funktion der Dimension. Exponentieller Euler mit/ohne Monitoring (rot/grün), `cgnr` mit/ohne Reorthogonalisierung (blau/gelb).

1.5 ist die Funktion  $x^\dagger$  und deren Rekonstruktion durch das exponentielle Euler-Verfahren mit  $l = 14$  und einem relativen Rekonstruktionsfehler von 7.9% dargestellt. Die weiteren Tests der Verfahren an diesem Beispiel bestätigen die Ergebnisse des Experiments aus Abschnitt 1.6.1. Da wir nicht wissen, ob die gewählte Funktion  $x^\dagger$  in einem der Teilräume  $\mathcal{X}_{\gamma,\ell}$  (vergleiche (1.7)) liegt, liefert uns die Theorie keine Aussage über das Verhalten des Fehlers  $\delta \rightarrow 0$ . An Abbildung 1.6 sehen wir jedoch, dass der relative Fehler bei allen vier Verfahren mit der gleichen Rate abnimmt.

Wie schon im ersten numerischen Beispiel kommen die exponentiellen Euler-Verfahren bei fast allen Dimensionen  $l$  (und damit größerem Rauschpegel  $\delta \approx h_l^2$ ) mit deutlich kleineren Krylov-Räumen aus als `cgnr` ohne Reorthogonalisierung (siehe Abbildung 1.7 (oben)). Dies spiegelt sich natürlich auch in der Effizienz der Verfahren wider. In Abbildung 1.7 (mittig) erkennt man, dass das exponentielle Euler-Verfahren ohne die Monitoringstrategie um bis zu 25% schneller ist wie das `cgnr`-Verfahren ohne Reorthogonalisierung. Setzt man die Monitoringstrategie ein, so reduziert sich der Vorteil nur minimal und die Methode ist somit noch deutlich schneller als das `cgnr`-Verfahren ohne Reorthogonalisierung. Zum Vergleich haben wir in Abbildung 1.7 auch die Ergebnisse von `cgnr` mit Reorthogonalisierung dargestellt. Man erkennt sehr gut, dass durch die erhöhte Stabilität des Verfahrens mit Reorthogonalisierung der numerische Mehraufwand mehr als wett gemacht wird. Die exponentiellen Euler-Verfahren sind bei den Beispielen großer Dimension nur um etwa 10% langsamer als `cgnr` mit Reorthogonalisierung also durchaus konkurrenzfähig.

Zum Test der Monitoringstrategie ist in Abbildung 1.7 (unten) das Verhältnis der Norm der Diskrepanz zu  $\tau\delta$  graphisch dargestellt. Offensichtlich erfüllt die Monitoringstrategie wieder ihre Aufgabe. Das so modifizierte Verfahren trifft den Stoppzeitpunkt wesentlich genauer als das Verfahren ohne Verwendung dieser Strategie.

Aufgrund des symmetrischen Kerns des Fredholm-Integraloperators  $A$  ist dieser selbstadjungiert. Man kann sich daher den Weg über die Normalengleichung 1.2 sparen und direkt das gestörte Problem 1.4 lösen. Die Theorie lässt sich auf diesen Fall übertragen. Wie Hanke in [29, Kapitel 4] zeigt, eignet sich das **cg**-Verfahren kombiniert mit dem Diskrepanzprinzip nicht für das symmetrische Problem. Wir vergleichen die exponentiellen Euler-Verfahren daher, wie in [29] vorgeschlagen, mit dem **mr**-Verfahren (**mr**=*minimal residual*, siehe [29, Abschnitt 2.2]). Die Umsetzung auf dem Rechner zeigt, dass sich am Konvergenzverhalten nichts Wesentliches ändert (siehe Abbildung 1.8). Die relativen Fehler aller vier Verfahren zeigen wieder ein gleiches Verhalten. Die Effizienz der Verfahren steigt allerdings, da in jedem Krylov-Schritt nur noch eine Matrix-Vektor-Multiplikation durchgeführt werden muss. Dabei profitiert das **mr**-Verfahren ohne Reorthogonalisierung am stärksten vom besser konditionierten System und macht den Rückstand auf die anderen Verfahren zum Teil wett. Die drei Verfahren, welche Reorthogonalisierung einsetzen, gewinnen durch diesen Ansatz nicht so stark. Die vier Verfahren haben daher eine vergleichbare Effizienz.

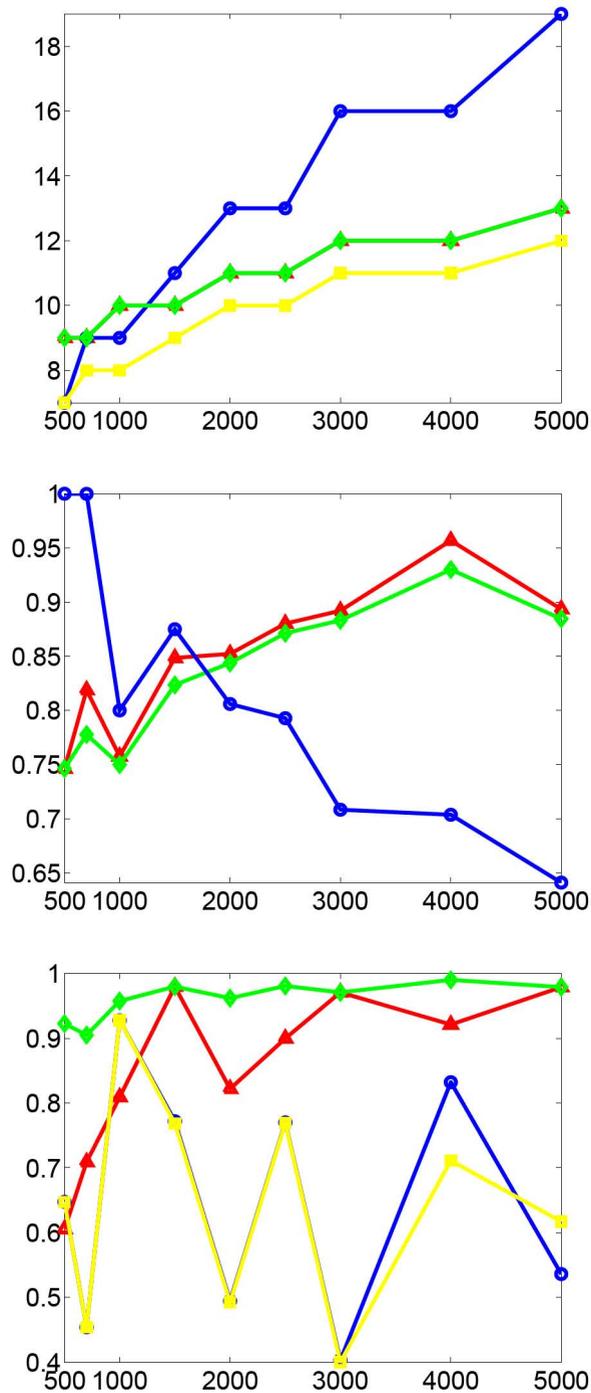


Abbildung 1.7: Dimension der Krylov-Räume (oben), Speedup gegenüber `cgnr` mit Reorthogonalisierung (mittig) und  $\|AR_{l,t_k^*}y^\delta - y^\delta\| / (\tau\delta)$  (unten) jeweils als Funktion der Dimension. Exponentieller Euler mit/ohne Monitoring (rot/grün), `cgnr` mit/ohne Reorthogonalisierung (blau/gelb).

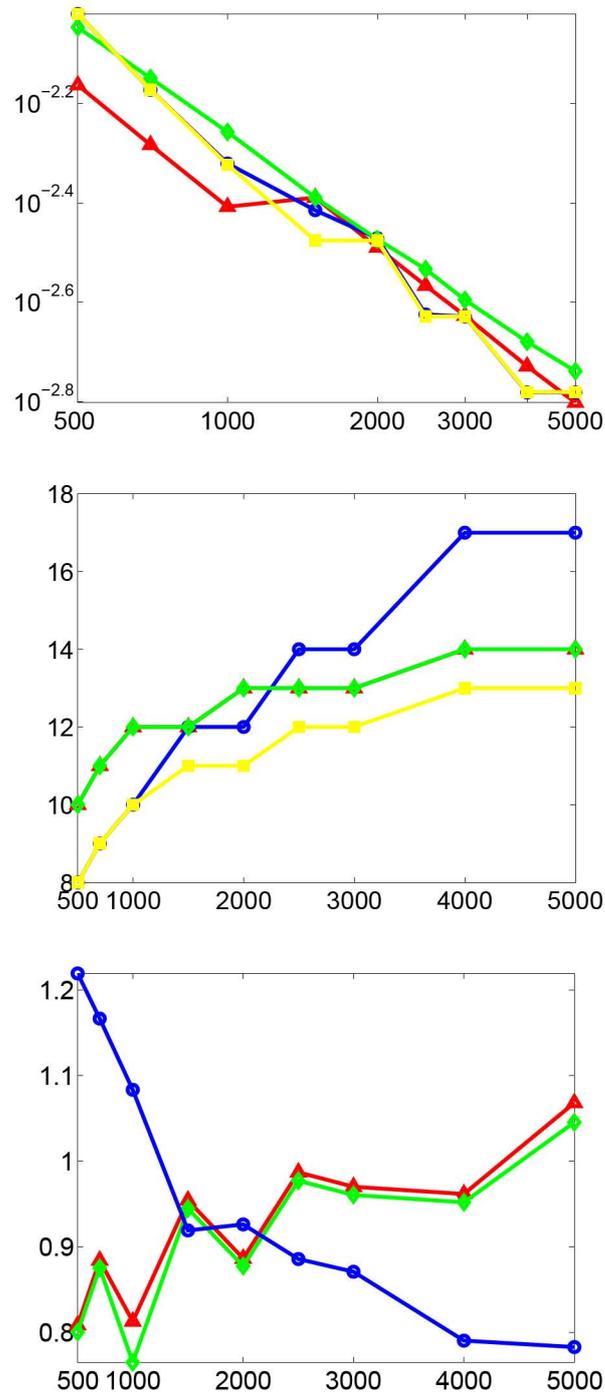


Abbildung 1.8: Relativer  $L^2$ -Regularisierungsfehler (oben), Dimension der Krylov-Räume (mittig) und Speedup gegenüber  $\text{cg}$  mit Reorthogonalisierung (unten) jeweils als Funktion der Dimension. Exponentieller Euler mit/ohne Monitoring (rot/grün),  $\text{cg}$  mit/ohne Reorthogonalisierung (blau/gelb).

## KAPITEL 2

# NICHTLINEARE SCHLECHT GESTELLTE PROBLEME

Im zweiten Kapitel dieser Arbeit beschäftigen wir uns mit nichtlinearen schlecht gestellten Problemen. Ziel ist es, festzustellen ob und wie sich das für die linearen Probleme entwickelte Verfahren auf diese Klasse von Problemen erweitern lässt und welche Eigenschaften es hat.

Zunächst verschaffen wir uns einen Überblick über die Theorie nichtlinearer schlecht gestellter Probleme und fassen bekannte Regularisierungsverfahren für diese Problemklasse zusammen. Im Anschluss daran präsentieren wir die Ergebnisse von Tautenhahn [80] zur asymptotischen Regularisierung nichtlinearer schlecht gestellter Probleme, welche die Basis sowohl für das hier entwickelte Verfahren als auch für den Beweis der regularisierenden Eigenschaft bilden. Nach diesem Rückblick stellen wir das vereinfachte exponentielle Euler-Verfahren für nichtlineare schlecht gestellte Probleme vor und weisen in Abschnitt 2.4 nach, dass dieses Verfahren gegen die gesuchte Lösung konvergiert. Im Anschluss daran zeigen wir, dass diese Konvergenz auch von optimaler Ordnung ist. An zwei numerischen Beispielen verifizieren wir schließlich, dass die theoretischen Ergebnisse auch numerisch beobachtbar sind und dass das exponentielle Euler-Verfahren zur Regularisierung schlecht gestellter Probleme konkurrenzfähig ist. Die Ergebnisse zur Konvergenz sind in dem mit Hochbruck und Ostermann verfassten Artikel [36] zu finden. Vorgestellt wurde das Verfahren schon in [37].

Beginnen wir nun mit einem kurzen Überblick über die Theorie.

## 2.1 Theorie nichtlinearer schlecht gestellter Probleme

Wir betrachten im Folgenden die Gleichung

$$(2.1) \quad F(x) = y$$

wobei  $F : \mathcal{D}(F) \subset \mathcal{X} \rightarrow \mathcal{Y}$  ein nichtlinearer Operator zwischen den Hilberträumen  $\mathcal{X}$  und  $\mathcal{Y}$  mit Definitionsbereich  $\mathcal{D}(F)$  ist. Wir gehen wieder davon aus, dass uns nur gestörte Daten  $y^\delta$  mit

$$(2.2) \quad \|y - y^\delta\|_{\mathcal{Y}} \leq \delta$$

zur Verfügung stehen. Die gestörte Gleichung lautet dann

$$(2.3) \quad F(u) = y^\delta.$$

Für lineare Probleme  $Ax = y$  kennen wir „einfache“ Kriterien, um die Schlechtgestellttheit zu überprüfen. So sind lineare Probleme genau dann schlecht gestellt, wenn das Bild von  $A$  nicht abgeschlossen ist, oder  $A$  kompakt und das Bild von  $A$  unendlichdimensional ist. Bei nichtlinearen Problemen gestaltet sich solch eine Charakterisierung schwieriger. Wir formulieren daher eine lokale Schlechtgestellttheit.

**Definition 2.1.** Die nichtlineare Operatorgleichung (2.1) auf den Hilberträumen  $\mathcal{X}$  und  $\mathcal{Y}$  heißt **lokal schlecht gestellt** in  $x_0 \in \mathcal{D}(F)$ , falls es zu jedem  $r > 0$  eine Folge  $\{x_n\}_{n \in \mathbb{N}} \subset B_r(x_0) \cap \mathcal{D}(F)$  gibt, die nicht gegen  $x_0$  konvergiert, deren Bildfolge  $\{F(x_n)\}_{n \in \mathbb{N}} \subset \mathcal{Y}$  aber gegen  $F(x_0)$  konvergiert, d.h.

$$\lim_{n \rightarrow \infty} \|F(x_n) - F(x_0)\|_{\mathcal{Y}} = 0, \quad \text{aber } x_n \not\rightarrow x_0 \quad \text{für } n \rightarrow \infty.$$

Andernfalls heißt (2.1) **lokal gut gestellt** in  $x_0$ .

Im linearen Fall liefern kompakte Operatoren schlecht gestellte Probleme. Dies verhält sich bei nichtlinearen Problemen ähnlich. Ein Kriterium für die lokale Schlechtgestellttheit eines nichtlinearen Problems erhält man aus dem folgenden Lemma (vgl. [21, Proposition A.3]).

**Lemma 2.2.** Sei  $F : \mathcal{D}(F) \subset \mathcal{X} \rightarrow \mathcal{Y}$  ein stetiger, kompakter und schwach folgenabgeschlossener nichtlinearer Operator zwischen den Hilberträumen  $\mathcal{X}$  und  $\mathcal{Y}$ . Ist  $\mathcal{X}$  separabel und unendlichdimensional, so ist (2.1) lokal schlecht gestellt in jedem  $x_0 \in \text{int}(\mathcal{D}(F))$ .

Viele Verfahren zur Lösung nichtlinearer Probleme linearisieren diese. Der Zusammenhang der Schlechtgestellttheit des nichtlinearen Problems und einer Linearisierung ist allerdings

nicht so einfach wie man vermuten könnte. Man kann im Allgemeinen nicht von der Schlechtgestelltheit von Linearisierungen von (2.1) auf die Schlechtgestelltheit von (2.1) selbst schließen. Genausowenig ist der umgekehrte Schluss von der Schlechtgestelltheit von (2.1) auf die Schlechtgestelltheit der zugehörigen Linearisierungen korrekt. Gegenbeispiele findet man bei Engl, Kunisch und Neubauer [21] und bei Schock [77].

Unter zusätzlichen Voraussetzungen an die Funktion  $F$ , zum Beispiel einer Lipschitz-Bedingung an die Fréchet-Ableitung  $F'$  (die natürlich existieren sollte), überträgt sich die Schlechtgestelltheit jedoch von (2.1) auf die Linearisierung

$$F'(x^\dagger)x = y.$$

Näheres hierzu findet man beispielsweise in [45].

Schon bei linearen Problemen haben wir die Eindeutigkeit der Lösung erzielt, indem wir die Lösung mit minimaler Norm ausgezeichnet haben. Da wir uns im nichtlinearen Fall lokal in einer Kugel um  $x_0$  bewegen, suchen wir nun die  $x_0$ -Minimum-Norm-Lösung.

**Definition 2.3.** Seien  $x_0 \in \mathcal{X}$  und  $y \in \mathcal{R}(F)$ . Der Vektor  $x^\dagger$  heißt  $x_0$ -**Minimum-Norm-Lösung** von (2.1) bezüglich  $y$ , falls

$$\|x^\dagger - x_0\|_{\mathcal{X}} = \min \{ \|x - x_0\|_{\mathcal{X}} : x \in \mathcal{D}(F), F(x) = y \}$$

gilt.

Die  $x_0$ -Minimum-Norm-Lösung ist unter gewissen Voraussetzungen lokal eindeutig (siehe Lemma 2.10). Verfahren, um die  $x_0$ -Minimum-Norm-Lösungen zu finden, wollen wir wieder als Regularisierungsverfahren bezeichnen.

**Definition 2.4.** Seien  $F : \mathcal{D}(F) \subset \mathcal{X} \rightarrow \mathcal{Y}$  stetig und  $\{R_t\}_{t>0}$  eine Familie stetiger Operatoren  $R_t : \mathcal{X} \times \mathcal{Y} \rightarrow \mathcal{Y}$ . Existiert eine Funktion  $\alpha : (0, \infty) \times \mathcal{Y} \rightarrow (0, \infty)$  so, dass für eine  $x_0$ -Minimum-Norm-Lösung  $x^\dagger$  von (2.1) mit  $y \in \mathcal{R}(F)$

$$\sup \{ \|x^\dagger - R_{\alpha(\delta, y^\delta)}(x_0, y^\delta)\|_{\mathcal{X}} : y^\delta \in \mathcal{Y}, \|y - y^\delta\|_{\mathcal{Y}} \leq \delta \} \longrightarrow 0 \quad \text{für } \delta \rightarrow 0$$

gilt, dann heißt  $(\{R_t\}_{t>0}, \alpha)$  **Regularisierungsverfahren** für  $F$ . Die **Parameterwahl**  $\alpha$  sei so orientiert, dass für jedes  $y \in \mathcal{R}(F)$ :

$$\lim_{\delta \rightarrow 0} \sup \{ \alpha(\delta, y^\delta) : y^\delta \in \mathcal{Y}, \|y - y^\delta\|_{\mathcal{Y}} \leq \delta \} = 0.$$

Es gibt auch für nichtlineare Probleme eine ganze Reihe von Regularisierungsverfahren. Das vermutlich einfachste Verfahren ist eine Verallgemeinerung des Landweber-Verfahrens für lineare Probleme. Ausgehend von  $x_0$  wird dabei via

$$(2.4) \quad u_{k+1} = u_k + F'(u_k)^*(y^\delta - F(u_k)) \quad \text{für } k = 0, 1, 2, \dots$$

iterativ eine Folge von Approximationen an die gesuchte Lösung berechnet. Eine Analyse dieser Methode findet man in [32]. Das vermutlich bekannteste Verfahren ist die nichtlineare Tikhonov-Phillips-Regularisierung. Hier wird analog zum linearen Fall (siehe Beispiel 1.2) das Minimierungsproblem

$$\|F(u) - y^\delta\|_Y^2 + t \|u - x_0\|_X^2 \stackrel{!}{=} \min$$

gelöst. Der Vektor  $x_0$  ist dabei eine „gute“ Startapproximation an die Lösung. Untersucht wird die regularisierende Eigenschaft der nichtlinearen Tikhonov-Phillips-Regularisierung in [21, 64]. Eine ganze Klasse von Regularisierungsmethoden sind Regularisierungsverfahren vom Newton-Typ. Wir haben zuvor schon festgestellt, dass sich die Schlechtgestellttheit eines Problems unter gewissen Voraussetzungen auf die Linearisierung überträgt. Bei regularisierten Newton-Verfahren wird daher in jedem Schritt ein lineares Regularisierungsverfahren auf das linearisierte Problem

$$(2.5) \quad F'(u_n)\Delta u_n = y^\delta - F(u_n)$$

angewandt (siehe [3, 46, 48, 49, 50, 72, 71, 81]). Eine Übersicht über derartige Methoden findet man im kürzlich erschienenen Buch von Kaltenbacher, Neubauer und Scherzer [51]. Jin und Tautenhahn zeigen in [48] die Konvergenz und Ordnungsoptimalität für eine ganze Klasse solcher Verfahren. Sie untersuchen iterative Verfahren der Form

$$(2.6) \quad u_{n+1} = u_0 + G_t(F'(u_k)^* F'(u_k)) F'(u_k)^* (y^\delta - F(u_k) - F'(u_k)(u_0 - u_k)),$$

wobei  $G_t$  die zu einem linearen Regularisierungsverfahren gehörende Filterfunktion (siehe Definition 1.9) ist.

Der iterative Ansatz bei Newton-Verfahren ermöglicht es, die Fülle der Regularisierungsoperatoren für lineare Probleme auch im nichtlinearen Fall einzusetzen.

### Beispiel 2.5.

1. Lösen wir das (eventuell schlecht gestellte) lineare Problem (2.5) mit Hilfe der Tikhonov-Phillips-Regularisierung, so erhalten wir das von Hanke [30] analysierte Levenberg-Marquardt-Verfahren.
2. Für Probleme mit großer Dimension sind natürlich iterative Verfahren zur Regularisierung des linearisierten Problems die Methoden der Wahl. Eine Analyse des Newton-Ansatzes mit dem **cgnr**-Verfahren als innerer Iteration findet man beispielsweise bei Hanke [31] oder Rieder [74].

Es wäre vermessen, bei nichtlinearen Problemen eine bessere optimale Konvergenzrate zu erwarten als bei linearen Problemen. Wie übertragen daher die Definition der Ordnungsoptimalität auf den nichtlinearen Fall.

**Definition 2.6.** Seien  $F : \mathcal{D}(F) \subset \mathcal{X} \rightarrow \mathcal{Y}$  stetig,  $x^\dagger$  eine  $x_0$ -Minimum-Norm-Lösung von (2.1) mit  $y \in \mathcal{R}(F)$ ,  $F$  Fréchet-differenzierbar und

$$(2.7) \quad x^\dagger - x_0 = (F'(x^\dagger)^* F'(x^\dagger))^\gamma w \quad \text{für ein } w \in X \text{ mit } \|w\|_{\mathcal{X}} \leq \varrho \text{ und ein } \gamma > 0.$$

Das Regularisierungsverfahren  $(\{R_t\}_{t>0}, \alpha)$  heißt **ordnungsoptimal**, wenn für  $\delta$  hinreichend klein eine Konstante  $C_\gamma$  existiert so, dass

$$\sup \{ \|x^\dagger - R_{\alpha(\delta, y^\delta)}(x_0, y^\delta)\|_{\mathcal{X}} : y^\delta \in \mathcal{Y}, \|y - y^\delta\|_{\mathcal{Y}} \leq \delta \} \leq C_\gamma \varrho^{\frac{1}{2\gamma+1}} \delta^{\frac{2\gamma}{2\gamma+1}}.$$

Die Quellbedingung (2.7) ist das Analogon zu den im linearen Kapitel definierten Räumen  $\mathcal{X}_{\gamma, \varrho}$  (siehe (1.7)).

**Beispiel 2.7.** Das Landweber-Verfahren für nichtlineare schlecht gestellt Probleme, definiert durch (2.4), konvergiert unter geeigneten Voraussetzungen (vgl. [32], unter anderem  $\gamma \in (0, 1/2]$ ) mit optimaler Ordnung.

## 2.2 Asymptotische Regularisierung

Da das von uns untersuchte Verfahren als diskrete Version der asymptotischen Regularisierung aufgefasst werden kann, geben wir in diesem Abschnitt einen kurzen Überblick über dieses Verfahren und präsentieren die für uns interessanten Eigenschaften aus der Arbeit von Tautenhahn [80]. Er untersucht dabei die nichtlineare Variante der asymptotischen Regularisierung (vgl. Abschnitt 1.3)

$$(2.8) \quad \begin{aligned} u'(t) &= F'(u(t))^* (y^\delta - F(u(t))) & t \geq 0, \\ u(0) &= x_0. \end{aligned}$$

Der Regularisierungsparameter  $t_*$  wird dabei wieder durch das Diskrepanzprinzip

$$(2.9) \quad \|F(u(t_*)) - y^\delta\| \leq \tau\delta < \|F(u(t)) - y^\delta\|, \quad 0 \leq t < t_*$$

bestimmt. Tautenhahn konnte nun unter der Voraussetzung der Tangentialkegelbedingung

**Annahme 1.**  $F : \mathcal{D}(F) \subset \mathcal{X} \rightarrow \mathcal{Y}$  ist Fréchet differenzierbar und erfüllt

$$(2.10) \quad \|F(x) - F(\tilde{x}) - F'(\tilde{x})(x - \tilde{x})\| \leq \eta \|F(x) - F(\tilde{x})\|,$$

für ein  $\eta < 1$ ,  $x, \tilde{x} \in B_r(x_0) \subset \mathcal{D}(F)$ .

zeigen, dass die durch das Diskrepanzprinzip (2.9) bestimmte Stoppzeit  $t_*$  endlich ist und dass das Verfahren sowohl bei exakten als auch bei gestörten Daten gegen eine Lösung von (2.1) konvergiert.

**Satz 2.8.** [80, Theorem 3] Es sei Annahme 1 erfüllt und (2.1) in  $B_r(x_0)$  lösbar. Falls  $u(t)$  eine Lösung von (2.8) mit  $y^\delta = y$  ist, konvergiert  $u(t)$  für  $t \rightarrow \infty$  gegen eine Lösung  $x_* \in B_r(x_0)$  von (2.1). Gilt zusätzlich  $\mathcal{N}(F'(x^\dagger)) \subset \mathcal{N}(F'(x))$  für alle  $x \in B_r(x_0)$ , so konvergiert  $u(t)$  sogar gegen  $x^\dagger$ .

**Satz 2.9.** [80, Theorem 4] Falls Annahme 1, (2.2) und  $\|F(x_0) - y^\delta\| > \tau\delta > 0$  erfüllt sind,  $u(t)$  eine Lösung von (2.8) ist,  $t_*$  nach dem Diskrepanzprinzip (2.9) mit

$$\tau > \frac{1 + \eta}{1 - \eta}$$

gewählt wird und (2.1) in  $B_r(x_0)$  lösbar ist, konvergiert  $u(t_*)$  für  $\delta \rightarrow 0$  gegen eine Lösung  $x_*$  von (2.1). Gilt zusätzlich  $\mathcal{N}(F'(x^\dagger)) \subset \mathcal{N}(F'(x))$  für alle  $x \in B_r(x_0)$ , so konvergiert  $u(t)$  sogar gegen  $x^\dagger$ .

**Bemerkung.** Die Bedingung (2.10) bedeutet, dass der Fehler der Linearisierung von (2.1) durch das Residuum beschränkt ist und liefert mit der Dreiecksungleichung die folgende Abschätzung:

$$(2.11) \quad \frac{1}{1+\eta} \|F'(x)(x - \tilde{x})\| \leq \|F(x) - F(\tilde{x})\| \leq \frac{1}{1-\eta} \|F'(x)(x - \tilde{x})\|$$

für alle  $x, \tilde{x} \in B_r(x_0)$ . ◇

Eine direkte Konsequenz hiervon ist

**Lemma 2.10.** *Unter der Annahme (2.11) gilt: sind  $x_*, \tilde{x}_*$  zwei Lösungen von (2.1) in  $B_r(x_0)$ , dann gilt*

$$x_* - \tilde{x}_* \in \mathcal{N}(F'(x_*)).$$

Dieses Lemma charakterisiert die Lösungen von (2.1) in  $B_r(x_0)$  und ermöglicht später die Konvergenz gegen  $x^\dagger$  zu zeigen und nicht nur die Konvergenz gegen irgendeine Lösung des nichtlinearen Problems (2.1).

Zum Beweis der Konvergenzrate muss man zunächst die Quellbedingung aus Definition 2.6 fordern.

**Annahme 2.** *Es existieren  $w \in \mathcal{X}$ , sowie Konstanten  $\gamma \in (0, 1/2]$  und  $\varrho \geq 0$  so, dass*

$$e_0 = x_0 - x^\dagger = (F'(x^\dagger)^* F'(x^\dagger))^\gamma w \quad \text{und} \quad \|w\| \leq \varrho.$$

Ferner setzt Tautenhahn die folgende Annahme voraus:

**Annahme 3.** *Für jedes  $x \in B_r(x_0)$  existiert ein beschränkter linearer Operator  $R_x : \mathcal{Y} \rightarrow \mathcal{Y}$  und eine Konstante  $C_+ \geq 0$  so, dass*

1.  $F'(x) = R_x F'(x^\dagger)$  und
2.  $\|R_x - I\| \leq C_+ \|x - x^\dagger\|$

gelten.

**Bemerkung.** In der Literatur [71, 72, 74] ist als Alternative zu Annahme 3 auch die Version

$$(2.12) \quad \begin{aligned} F'(x) &= R_{(x, \tilde{x})} F'(\tilde{x}) \\ \|R_{(x, \tilde{x})} - I\| &\leq C_R \|x - \tilde{x}\|_{\mathcal{X}} \end{aligned}$$

für  $x, \tilde{x} \in B_r(u_0)$ , oder eine Variante hiervon [7, 49, 51] zu finden. ◇

Im Folgenden Lemma 2.11 zeigen wir, dass (2.12) schärfer als Annahme 3 ist.

**Lemma 2.11.** *Ist  $C_{Rr} < \frac{1}{2}$  so folgt aus (2.12) die Annahme 3 mit  $\eta = \frac{C_{Rr}}{1-C_{Rr}} < 1$ .*

*Beweis.* Aus

$$\begin{aligned} \|(F'(x) - F'(\tilde{x}))(x - \tilde{x})\|_y &= \|(R_{(x,\tilde{x})-I}F'(\tilde{x}))(x - \tilde{x})\|_y \\ &\leq C_R \|x - \tilde{x}\|_X \|F'(\tilde{x})(x - \tilde{x})\|_y \end{aligned}$$

folgt mit  $x_t = \tilde{x} + t(x - \tilde{x})$  und dem Mittelwertsatz

$$\begin{aligned} \|F(x) - F(\tilde{x}) - F'(\tilde{x})(x - \tilde{x})\|_y &= \left\| \int_0^1 (F'(x_t) - F'(\tilde{x}))(x - \tilde{x}) dt \right\|_y \\ &\leq \int_0^1 C_R \|x_t - \tilde{x}\|_X \|F'(\tilde{x})(x - \tilde{x})\|_y dt \\ &\leq C_R \frac{1}{2} \|x - \tilde{x}\|_X \|F'(\tilde{x})(x - \tilde{x})\|_y. \end{aligned}$$

Mit der Dreiecksungleichung und  $\|x - \tilde{x}\|_X < 2r$  ergibt sich daraus

$$(1 - C_{Rr}) \|F'(\tilde{x})(x - \tilde{x})\|_y \leq \|F(x) - F(\tilde{x})\|_y$$

und schließlich wegen  $C_{Rr} < 1$

$$\|F(x) - F(\tilde{x}) - F'(\tilde{x})(x - \tilde{x})\|_y \leq \frac{C_{Rr}}{1 - C_{Rr}} \|F(x) - F(\tilde{x})\|_y.$$

Für  $C_{Rr} < \frac{1}{2}$  ist  $\frac{C_{Rr}}{1-C_{Rr}} < 1$  und die Behauptung ist gezeigt.  $\square$

Ohne Beschränkung der Allgemeinheit (man kann (2.1) entsprechend skalieren) nimmt Tautenhahn ferner an:

**Annahme 4.** *Für  $x \in B_r(x_0)$  gilt*

$$(2.13) \quad \|F'(x)\| \leq 1.$$

Unter all diesen Annahmen beweist Tautenhahn nun in [80, Theorem 6] die Ordnungsoptimalität der nichtlinearen asymptotischen Regularisierung.

**Satz 2.12.** *Falls zusätzlich zu den Voraussetzungen von Satz 2.9 die Annahmen 2 bis 4 und*

$$(2.14) \quad \tau > \frac{2 - \eta}{1 - \eta},$$

*erfüllt sind, existiert eine Konstante  $c_* = c_*(\gamma, \tau, \eta)$  so, dass für die Lösung  $u$  von (2.8) und für genügend kleines  $\varrho$*

$$\|u(t) - x^\dagger\| \leq c_* \varrho^{1/(2\gamma+1)} \delta^{2\gamma/(2\gamma+1)}$$

*gilt.*

## 2.3 Das nichtlineare exponentielle Euler-Verfahren

Eine numerische Umsetzung der asymptotischen Regularisierung ist natürlich nicht direkt möglich, sondern erfordert den Einsatz eines Zeitintegrators. Böckmann und Pornsawad beweisen in [8] die regularisierende Eigenschaft vereinfachter Runge-Kutta-Verfahren angewandt auf (2.8)

$$\begin{aligned} u_{n+1} &= u_n + h_n \sum_{i=1}^s b_i F'(U_{n,i})^* (y^\delta - F(U_{n,i})), \\ U'_{n,i} &= h_n \sum_{j=1}^s a_{ij} F'(u_n)^* (y^\delta - F(u_n) - F'(u_n)U'_{n,j}) \\ U_{n,i} &= u_n + U'_{n,i}. \end{aligned}$$

Diese Methode hat allerdings mehrere schwerwiegende Nachteile. Zwar werden durch die dort vorgeschlagene Vereinfachung nichtlineare Probleme auf lineare Probleme reduziert, allerdings muss man in jedem Integrationsschritt  $(s + 1)$ -mal die Funktion  $F$  und ihre Ableitung  $F'$  auswerten, was beides in der Regel teuer ist ( $s$  ist dabei die Anzahl der Stufen des Verfahrens). Zum Beweis der Konvergenz ist ferner eine Abschätzung

$$(2.15) \quad \frac{c_1}{\sqrt{2}} \leq b^*(I_s + h_n \lambda \mathcal{A})^{-1} \mathbb{1} \leq c_1 \quad \text{für } \lambda \in [0, \|F'(u_n)\|^2]$$

nötig. Der Koeffizientenvektor  $b$  und die Koeffizientenmatrix  $\mathcal{A}$  sind dabei wieder wie im linearen Fall (1.14) definiert. Die Bedingung (2.15) zieht eine drastische Beschränkung der Schrittweite nach sich und erhöht damit die Anzahl der Integrationsschritte. Setzen wir wie in [8] die Annahme 4 voraus, so ergibt sich beispielsweise für das implizite Euler-Verfahren die Schrittweitenbeschränkung  $h_n \leq \sqrt{2} - 1$ .

Diese Beschränkung scheint allerdings nur beweistechnisch notwendig zu sein, da die in [8] beschriebenen numerischen Experimente auch mit größeren Schrittweiten gerechnet werden können.

Auch Li et al. [57] verfolgen den Ansatz, das nichtlineare Showalter-Anfangswertproblem mit einem Runge-Kutta-Verfahren zu lösen. Sie untersuchen dabei insbesondere das explizite zweistufige Verfahren von Runge gegeben durch das Butcher-Tableau

$$\begin{array}{c|cc} 0 & 0 & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 \\ \hline & 0 & 1 \end{array}.$$

Für konstante Schrittweite  $h_n = 1$  wird in [57] unter den Annahmen 1 – 3 gezeigt, dass dieses Regularisierungsverfahren für  $\gamma \in (0, \frac{1}{2}]$  mit optimaler Ordnung konvergiert. Die

numerischen Beispiele legen die zu erwartenden Vorteile dieses Verfahrens gegenüber dem Landweber-Verfahren dar.

In dieser Arbeit betrachten wir eine vereinfachte Version des exponentiellen Euler-Verfahrens (siehe Anhang A) angewandt auf (2.8). Dazu approximieren wir die Jacobi-Matrix

$$(2.16) \quad \frac{d}{du} F'(u)(y^\delta - F(u)) \approx -F'(u)^* F'(u) =: -J(u),$$

indem wir die Ableitung zweiter Ordnung wegfällen lassen. Die zugehörige Iteration ist dann

$$(2.17) \quad u_{n+1} = u_n + h_n \varphi(-h_n J(u_n)) F'(u_n)^*(y^\delta - F(u_n)) \quad \text{für } n \in \mathbb{N}_0$$

mit der  $\varphi$ -Funktion aus (1.12). Man kann das Verfahren (2.17) auch als ein regularisiertes inexaktes Newton-Verfahren [71, 72] mit dem exponentiellen Euler-Verfahren als innerer Regularisierung interpretieren. Verwendet man nämlich das exponentielle Euler-Verfahren um die im Newton-Verfahren auftretenden linearisierten Gleichungen (2.5) zu regularisieren, so erhält man

$$\Delta u_n = h_n \varphi(-h_n J(u_n)) F'(u_n)^*(y^\delta - F(u_n)).$$

Zusammen mit  $u_{n+1} = u_n + \Delta u_n$  ergibt dies genau die Iterationsvorschrift (2.17).

Wie schon in Kapitel 1 werden wir die zu den exakten Daten  $y$  gehörenden Iterierten mit  $x_n$  bezeichnen.

### Bemerkungen.

1. Wendet man das Verfahren (2.17) auf eine lineares Problem an, so erhält man das in Kapitel 1 untersuchte Verfahren (1.16). Somit ist das vereinfachte exponentielle Euler-Verfahren zur Regularisierung nichtlinearer schlecht gestellter Probleme (2.17) eine Verallgemeinerung des Verfahrens (1.16) für lineare Probleme.
2. Auch im nichtlinearen Fall liefern verschiedene, teilweise vereinfachte, Integratoren zur Lösung des Anfangswertproblems (2.8) schon bekannte Regularisierungsverfahren. So entspricht das explizite Euler-Verfahren der nichtlinearen Landweber-Iteration und das linear implizite Euler-Verfahren, vereinfacht durch die Approximation (2.16), führt auf das Verfahren von Levenberg-Marquardt.

◇

Wie schon bei den linearen Problemen müssen wir die Iteration wieder nach einer geeigneten Anzahl von Schritten stoppen. Diesen Stoppindex wählen wir wieder mit Hilfe des Diskrepanzprinzips

$$(2.18) \quad \|y^\delta - F(u_{n^*})\| \leq \tau \delta < \|y^\delta - F(u_n)\|, \quad \text{für } n < n^*.$$

Zur Vereinfachung der Notation führen wir noch einige abkürzende Schreibweisen ein:

$$\begin{aligned}
A_n &= F'(u_n), & J_n &= A_n^* A_n, & K_n &= A_n A_n^*, \\
A_+ &= F'(x^\dagger), & J_+ &= A_+^* A_+, & K_+ &= A_+ A_+^*, \\
\Delta u_n &= u_{n+1} - u_n, & \Delta F_n &= y^\delta - F(u_n), & \Delta F_n^\delta &= y^\delta - F(u_n), \\
\Phi_n &= \varphi(-h_n J_n), & \tilde{\Phi}_n &= \varphi(-h_n K_n), \\
\Phi_{n,+} &= \varphi(-h_n J_+), & \tilde{\Phi}_{n,+} &= \varphi(-h_n K_+), \\
R_n &= R_{u_n}, \\
e_n &= u_n - x^\dagger.
\end{aligned}$$

Um die Lesbarkeit zu erhalten, haben wir die Abhängigkeit von  $\delta$ , beispielsweise bei  $e_n$ , nicht immer gekennzeichnet.

## 2.4 Konvergenz des exponentiellen Euler-Verfahrens

Ziel dieses Abschnitts ist es, die Konvergenz der Iterierten  $u_n$  des exponentiellen Euler-Verfahrens gegen die  $x_0$ -Minimum-Norm-Lösung  $x^\dagger$  zu zeigen. Dazu ist jedoch noch eine Strategie nötig, mittels derer wir die Zeitschrittweiten  $h_n$  wählen. In der Literatur gibt es dazu verschiedene Vorschläge. Einen sehr einfachen Ansatz verfolgen Jin und Tautenhahn [48]. Sie fordern, dass die Schrittweiten monoton, aber auch nicht zu schnell wachsend sind, genauer verlangen sie, dass eine Konstante  $\Delta h_{\max}$  existiert, sodass

$$h_n \leq h_{n+1} \leq h_n + \Delta h_{\max} \quad \text{für } n > 0$$

gilt. Rieder [72] schlägt vor, die Schrittweite so zu wählen, dass ein Diskrepanzprinzip für das um  $u_n$  linearisierte Problem zum ersten Mal erfüllt ist, also

$$\|A_n^* \Delta u_n - \Delta F_n^\delta\|_{\mathcal{Y}} < \mu_n \|\Delta F_n^\delta\|_{\mathcal{Y}} \quad \text{für } n > 0,$$

wobei die Parameter  $\mu_n$  aus dem Intervall  $(\eta + (1 + \eta)\delta / \|\Delta F_n^\delta\|_{\mathcal{Y}}, 1]$  gewählt werden sollen.

Unsere Analyse beruht auf dem von Hanke [30] gewählten Weg. Der Parameter  $\mu$  erfülle dazu  $0 < \mu < 1$  und die Funktion  $\psi$  sei durch

$$(2.19) \quad \psi(z) = \varphi(z)(1 - z)$$

definiert. Wir wählen dann die Zeitschrittweite  $h_n$  nach dem folgenden Diskrepanzprinzip:

$$(2.20) \quad p_n^\delta(h_n) := \|\psi(-h_n K_n) \Delta F_n^\delta - A_n \Delta u_n\|_{\mathcal{Y}} = \mu \|\Delta F_n^\delta\|_{\mathcal{Y}}.$$

Dies können wir auch äquivalent umformen zu

$$(2.21) \quad p_n^\delta(h_n) = \|\varphi(-h_n K_n) \Delta F_n^\delta\|_{\mathcal{Y}}.$$

Um die Wohldefiniertheit der durch 2.20 definierten Schrittweite  $h_n$  zu zeigen nehmen wir an, dass nach  $n$  Zeitschritten sei die Ungleichung

$$(2.22) \quad \|\psi(-h_n K_n) \Delta F_n^\delta + A_n e_n\|_{\mathcal{Y}} \leq \frac{\mu}{\nu} \|\Delta F_n^\delta\|_{\mathcal{Y}}$$

erfüllt ist. Die Gültigkeit Annahme wird später in den Sätzen 2.15 und 2.16 verifiziert werden.

**Bemerkung.** Für das in [30] untersuchte Levenberg-Marquardt-Verfahren vereinfacht sich die Situation wegen  $\psi \equiv 1$ . In diesem Fall ergeben das Diskrepanzprinzip (2.20) und die Annahme (2.22) die gleichen Bedingungen wie in [30].  $\diamond$

Im folgenden Lemma stellen wir sicher  $h_n$  durch (2.20) bzw. (2.21) eindeutig bestimmt ist.

**Lemma 2.13.** *Es gelte (2.22), dann hat (2.20) eine eindeutige Lösung  $h_n > 0$ .*

*Beweis.* Die Funktion  $p_n^\delta(h)$  ist stetig und streng monoton fallend in  $h$ , sofern  $\Delta F_n^\delta \neq 0$  gilt. Aus (2.21) und  $\varphi(0) = 1$  folgt unmittelbar

$$\lim_{h \rightarrow 0} p_n^\delta(h) = \|\Delta F_n^\delta\|_{\mathcal{Y}}.$$

Wegen  $\varphi(-z) \rightarrow 0$  für  $z \rightarrow \infty$  gilt auch

$$\lim_{h \rightarrow \infty} p_n^\delta(h) = \|\Pi_n \Delta F_n^\delta\|_{\mathcal{Y}},$$

wobei  $\Pi_n$  die Orthogonalprojektion auf  $\mathcal{R}(A_n)^\perp$  ist. Nach Definition von  $\psi$  gilt  $\psi(0) = 1$ . Aus der Potenzreihendarstellung von  $\psi$  folgt daher unmittelbar

$$\Pi_n \Delta F_n^\delta = \Pi_n \psi(-hK_n) \Delta F_n^\delta.$$

Somit gilt für alle  $u \in \mathcal{D}(F)$

$$\Pi_n \Delta F_n^\delta = \Pi_n \left( \psi(-hK_n) \Delta F_n^\delta + A_n(u_n - u) \right).$$

Dies liefert für alle  $u \in \mathcal{D}(F)$  und  $h > 0$

$$\|\Pi_n \Delta F_n^\delta\|_{\mathcal{Y}} \leq \|\psi(-hK_n) \Delta F_n^\delta + A_n(u_n - u)\|_{\mathcal{Y}}.$$

Die Existenz der (2.20) erfüllenden Schrittweite  $h_n$  folgt nun mittels (2.22) aus

$$\|\Pi_n \Delta F_n^\delta\|_{\mathcal{Y}} \leq \|\psi(-h_n K_n) \Delta F_n^\delta + A_n e_n\|_{\mathcal{Y}} \leq \frac{\mu}{\nu} \|\Delta F_n^\delta\|_{\mathcal{Y}} \leq \mu \|\Delta F_n^\delta\|_{\mathcal{Y}} \leq \|\Delta F_n^\delta\|_{\mathcal{Y}}.$$

□

Dieses Lemmas stellt, vorausgesetzt (2.22) ist erfüllt, sicher, dass die Iterierten  $u_n$  auch wohldefiniert sind. Der Nachweis der Eigenschaft (2.22) wird wie schon erwähnt später in den Beweisen zur Konvergenz bei exakten (Satz 2.15) und gestörten Daten (Satz 2.16) erbracht. Zunächst zeigen wir aber im folgenden Lemma, dass  $u_{n+1}$  eine bessere Approximation an  $x^\dagger$  als  $u_n$  liefert.

**Lemma 2.14.** *Sei (2.22) erfüllt und  $0 < \mu < 1 < \nu$ , so dass wir  $h_n$  nach (2.20) wählen können. Dann gelten für  $\Delta u_n \neq 0$  die Ungleichungen*

$$(2.23) \quad \|e_n\|_{\mathcal{X}}^2 - \|e_{n+1}\|_{\mathcal{X}}^2 > \frac{2h_n(\nu-1)\mu}{\nu} \|\Delta F_n^\delta\|_{\mathcal{Y}} \cdot \|\tilde{\Phi}_n \Delta F_n^\delta\|_{\mathcal{Y}},$$

$$(2.24) \quad \|e_n\|_{\mathcal{X}}^2 - \|e_{n+1}\|_{\mathcal{X}}^2 > -\frac{2\varphi^{-1}(\mu)(\nu-1)\mu^2}{\nu \|A_n\|^2} \|\Delta F_n^\delta\|_{\mathcal{Y}}^2.$$

*Beweis.* Für den Fehler  $e_{n+1} = u_{n+1} - x^\dagger$  der  $(n+1)$ -ten Iterierten  $u_{n+1}$  gilt

$$\begin{aligned}
\|e_{n+1}\|_{\mathcal{X}}^2 &= \|\Delta u_n + e_n\|_{\mathcal{X}}^2 \\
&= \|\Delta u_n\|_{\mathcal{X}}^2 + 2\langle \Delta u_n, e_n \rangle_{\mathcal{X}} + \|e_n\|_{\mathcal{X}}^2 \\
&= h_n^2 \langle \Phi_n A_n^* \Delta F_n^\delta, \Phi_n A_n^* \Delta F_n^\delta \rangle_{\mathcal{X}} \\
&\quad + 2h_n \langle \Phi_n A_n^* \Delta F_n^\delta, e_n \rangle_{\mathcal{X}} + \|e_n\|_{\mathcal{X}}^2 \\
&= h_n^2 \left\langle \Delta F_n^\delta, K_n \tilde{\Phi}_n^2 \Delta F_n^\delta \right\rangle_{\mathcal{Y}} \\
&\quad - 2h_n \left\langle \tilde{\Phi}_n \Delta F_n^\delta, \psi(-h_n K_n) \Delta F_n^\delta \right\rangle_{\mathcal{Y}} \\
&\quad + 2h_n \left\langle \tilde{\Phi}_n \Delta F_n^\delta, \psi(-h_n K_n) \Delta F_n^\delta + A_n e_n \right\rangle_{\mathcal{Y}} + \|e_n\|_{\mathcal{X}}^2 \\
&= -h_n^2 \left\langle \Delta F_n^\delta, K_n \tilde{\Phi}_n^2 \Delta F_n^\delta \right\rangle_{\mathcal{Y}} \\
&\quad + 2h_n \left\langle \tilde{\Phi}_n \Delta F_n^\delta, h_n K_n \tilde{\Phi}_n \Delta F_n^\delta - \psi(-h_n K_n) \Delta F_n^\delta \right\rangle_{\mathcal{Y}} \\
&\quad + 2h_n \left\langle \tilde{\Phi}_n \Delta F_n^\delta, \psi(-h_n K_n) \Delta F_n^\delta + A_n e_n \right\rangle_{\mathcal{Y}} + \|e_n\|_{\mathcal{X}}^2.
\end{aligned}$$

Die Wahl  $\psi(z) = \varphi(z)(1-z)$  in (2.19) führt auf die Identität

$$z\varphi(-z) - \psi(-z) = -\varphi(-z).$$

Damit formen wir den obigen Ausdruck nun weiter um zu

$$\begin{aligned}
\|e_{n+1}\|_{\mathcal{X}}^2 - \|e_n\|_{\mathcal{X}}^2 &= -h_n^2 \left\langle \Delta F_n^\delta, K_n \tilde{\Phi}_n^2 \Delta F_n^\delta \right\rangle_{\mathcal{Y}} - 2h_n \left\langle \tilde{\Phi}_n \Delta F_n^\delta, \tilde{\Phi}_n \Delta F_n^\delta \right\rangle_{\mathcal{Y}} \\
&\quad + 2h_n \left\langle \tilde{\Phi}_n \Delta F_n^\delta, \psi(-h_n K_n) \Delta F_n^\delta + A_n e_n \right\rangle_{\mathcal{Y}} \\
&= -\|\Delta u_n\|_{\mathcal{Y}}^2 - 2h_n \left\| \tilde{\Phi}_n \Delta F_n^\delta \right\|_{\mathcal{Y}}^2 \\
&\quad + 2h_n \left\langle \tilde{\Phi}_n \Delta F_n^\delta, \psi(-h_n K_n) \Delta F_n^\delta + A_n e_n \right\rangle_{\mathcal{Y}} \\
&< -2h_n \left\| \tilde{\Phi}_n \Delta F_n^\delta \right\|_{\mathcal{Y}}^2 + 2h_n \left\langle \tilde{\Phi}_n \Delta F_n^\delta, \psi(-h_n K_n) \Delta F_n^\delta + A_n e_n \right\rangle_{\mathcal{Y}} \\
&\leq 2h_n \left\| \tilde{\Phi}_n \Delta F_n^\delta \right\|_{\mathcal{Y}} \left( \left\| \psi(-h_n K_n) \Delta F_n^\delta + A_n e_n \right\|_{\mathcal{Y}} - \left\| \tilde{\Phi}_n \Delta F_n^\delta \right\|_{\mathcal{Y}} \right) \\
&\leq 2h_n \mu \frac{1-\nu}{\nu} \left\| \tilde{\Phi}_n \Delta F_n^\delta \right\|_{\mathcal{Y}} \left\| \Delta F_n^\delta \right\|_{\mathcal{Y}}.
\end{aligned}$$

Damit haben wir (2.23) nachgewiesen.

Zum Beweis der zweiten Ungleichung (2.24) verwenden wir

$$\mu \left\| \Delta F_n^\delta \right\|_{\mathcal{Y}} = \left\| \tilde{\Phi}_n \Delta F_n^\delta \right\|_{\mathcal{Y}} \geq \varphi(-h_n \|K\|) \left\| \Delta F_n^\delta \right\|_{\mathcal{Y}},$$

woraus wir

$$h_n \geq -\frac{\varphi^{-1}(\mu)}{\|A_n\|^2}$$

schließen. Einsetzen in (2.23) liefert nun das gewünschte Resultat.  $\square$

Wir weisen an dieser Stelle darauf hin, dass  $\varphi^{-1}(\mu)$  für  $\mu \in (0, 1)$  negativ ist und daher die rechte Seite von (2.24) positiv.

Nach all diesen Vorbereitungen sind wir nun in der Lage die Konvergenz des exponentiellen Euler-Verfahrens bei exakten Daten nachzuweisen. Dieser Beweis orientiert sich an [31, Theorem 4.2].

**Satz 2.15.** *Sei  $0.3 + \eta < \mu < 1$  und seien Annahme 1 und 4 erfüllt. Falls  $y^\delta = y = F(x^\dagger)$  gilt, konvergiert  $u_n$  für  $n \rightarrow \infty$  gegen eine Lösung von (2.1). Gilt zusätzlich*

$$(2.25) \quad \mathcal{N}(A_+) \subset \mathcal{N}(F'(x)) \quad \text{für alle } x \in B_r(x_0),$$

so konvergiert  $u_n$  für  $n \rightarrow \infty$  gegen  $x^\dagger$ .

*Beweis.* Wir definieren  $\nu = \frac{\mu}{0.3 + \eta} > 1$ . Dann gilt

$$\begin{aligned} & \|\psi(-h_n K_n) \Delta F_n + A_n e_n\|_{\mathcal{Y}} \\ &= \left\| (\tilde{\Phi}_n - e^{-h_n K_n}) \Delta F_n + \Delta F_n + A_n e_n \right\|_{\mathcal{Y}} \\ &\leq \left\| (\tilde{\Phi}_n - e^{-h_n K_n}) \right\| \|\Delta F_n\|_{\mathcal{Y}} + \|\Delta F_n + A_n e_n\|_{\mathcal{Y}} \\ &\leq \sup_{z \in [-\infty, 0]} |\varphi(z) - e^z| \|\Delta F_n\|_{\mathcal{Y}} + \eta \|\Delta F_n\|_{\mathcal{Y}} \\ &\leq (0.3 + \eta) \|\Delta F_n\|_{\mathcal{Y}} \\ &\leq \frac{\mu}{\nu} \|\Delta F_n\|_{\mathcal{Y}}. \end{aligned}$$

Damit ist die Ungleichung (2.22) für  $n = 0$  erfüllt und wir erhalten mit Hilfe von Lemma 2.14

$$\|x_1 - x^\dagger\|_{\mathcal{X}} < \|x_0 - x^\dagger\|_{\mathcal{X}}.$$

Induktion liefert nun, dass die Norm der Fehler streng monoton fällt.

Es bleibt zu zeigen, dass  $u_n$  gegen eine Lösung von (2.1) konvergiert. Dazu zeigen wir zunächst, dass die Fehler  $e_n$  eine Cauchy-Folge bilden. Seien dazu  $m, n \in \mathbb{N}$  und  $m > n$ . Der Index  $l \in \{n, \dots, m\}$  sei so gewählt, dass das Residuum minimiert wird, d.h.

$$\|\Delta F_l\|_{\mathcal{Y}} \leq \|\Delta F_i\|_{\mathcal{Y}} \quad \text{für } i = n, \dots, m.$$

Wir betrachten nun

$$(2.26) \quad \|e_l - e_n\|_{\mathcal{X}}^2 = \|e_n\|_{\mathcal{X}}^2 + 2 \langle e_l - e_n, e_l \rangle_{\mathcal{X}} - \|e_l\|_{\mathcal{X}}^2.$$

Mit Hilfe von Lemma 2.14 erhalten wir die folgende Abschätzung für das Skalarprodukt aus (2.26):

$$\begin{aligned} |\langle e_l - e_n, e_l \rangle_{\mathcal{X}}| &= \left| \sum_{i=n}^{l-1} \langle h_i A_i^* \tilde{\Phi}_i \Delta F_i, e_l \rangle_{\mathcal{X}} \right| \\ &= \left| \sum_{i=n}^{l-1} \langle h_i \tilde{\Phi}_i \Delta F_i, A_i e_l \rangle_{\mathcal{Y}} \right| \\ &\leq \sum_{i=n}^{l-1} \|h_i \tilde{\Phi}_i \Delta F_i\|_{\mathcal{Y}} \|A_i e_l\|_{\mathcal{Y}}. \end{aligned}$$

Für den letzten Faktor dieser Ungleichung erhalten wir weiter

$$\begin{aligned} \|A_i e_l\|_{\mathcal{Y}} &= \|A_i e_i - A_i(e_i - e_l)\|_{\mathcal{Y}} \\ &\leq \|\Delta F_i + A_i e_i\|_{\mathcal{Y}} + \|F_i - F_l - A_i(e_i - e_l)\|_{\mathcal{Y}} + \|\Delta F_l\|_{\mathcal{Y}} \\ &\leq (3\eta + 1) \|\Delta F_i\|_{\mathcal{Y}}. \end{aligned}$$

Daraus folgt mit Hilfe von (2.23)

$$\begin{aligned} |\langle e_l - e_n, e_l \rangle_{\mathcal{X}}| &\leq (3\eta + 1) \sum_{i=n}^{l-1} \|h_i \tilde{\Phi}_i \Delta F_i\|_{\mathcal{Y}} \cdot \|\Delta F_i\|_{\mathcal{Y}} \\ &\leq \frac{(3\eta + 1)\nu}{2(\nu - 1)\mu} \sum_{i=n}^{l-1} (\|e_i\|_{\mathcal{X}}^2 - \|e_{i+1}\|_{\mathcal{X}}^2) \\ &= \frac{(3\eta + 1)\nu}{2(\nu - 1)\mu} (\|e_n\|_{\mathcal{X}}^2 - \|e_l\|_{\mathcal{X}}^2) \end{aligned}$$

und damit schließlich

$$\|e_l - e_n\|_{\mathcal{X}}^2 \leq \left( \frac{(3\eta + 1)\nu}{(\nu - 1)\mu} + 1 \right) (\|e_n\|_{\mathcal{X}}^2 - \|e_l\|_{\mathcal{X}}^2).$$

Analog zeigt man nun auch

$$\|e_m - e_l\|_{\mathcal{X}}^2 \leq \left( \frac{(3\eta + 1)\nu}{(\nu - 1)\mu} + 1 \right) (\|e_l\|_{\mathcal{X}}^2 - \|e_m\|_{\mathcal{X}}^2).$$

Zusammen ergibt dies

$$\begin{aligned} \|u_m - u_n\|_{\mathcal{X}}^2 &= \|e_m - e_n\|_{\mathcal{X}}^2 \leq 2 \|e_m - e_l\|_{\mathcal{X}}^2 + 2 \|e_l - e_m\|_{\mathcal{X}}^2 \\ &\leq 2 \left( \frac{(3\eta + 1)\nu}{(\nu - 1)\mu} + 1 \right) (\|e_n\|_{\mathcal{X}}^2 - \|e_m\|_{\mathcal{X}}^2). \end{aligned}$$

Da, wie wir oben gezeigt haben, die Norm des Fehlers monoton fällt, konvergiert die rechte für  $m, n \rightarrow \infty$  gegen 0. Damit bilden die Iterierten  $u_n$  eine Cauchy-Folge. Da  $\|A_n\|$  beschränkt ist (wir haben ja  $\|A_n\| \leq 1$  angenommen), folgt nun aus (2.24), dass die Reihe  $\sum_{n=0}^{\infty} \|y - F(u_n)\|^2$  konvergiert und damit  $F(u_n) \rightarrow y$  für  $n \rightarrow \infty$  gelten muss. Damit ist  $u_{\infty} = \lim_{n \rightarrow \infty} u_n$  eine Lösung von (2.1).

Aus (2.25) folgt unmittelbar  $N(A_+) \subset N(A_k)$  für alle  $k \in \mathbb{N}_0$ . Dann gilt aber auch

$$u_k - x_0 \in N(A_+)^{\perp}, \quad k \in \mathbb{N},$$

und damit

$$x^{\dagger} - u_{\infty} = x^{\dagger} - x_0 + x_0 - u_{\infty} \in N(A_+)^{\perp}.$$

Zusammen mit Lemma 2.10 folgt  $u_{\infty} = x^{\dagger}$ . □

Wir können nun im folgenden Satz die Konvergenz bei gestörten Daten nachweisen.

**Satz 2.16.** *Zusätzlich zu den Voraussetzungen in Satz 2.15 seien (2.2) erfüllt und*

$$\mu > \frac{1 + \eta + \tau(0.3 + \eta)}{\tau}.$$

*Dann konvergiert das exponentielle Euler-Verfahren, gestoppt nach dem Diskrepanzprinzip (2.18), nach  $n_* = n_*(\delta) < \infty$  Iterationen und die zugehörigen Approximationen  $u_{n_*}$  konvergieren für  $\delta \rightarrow 0$  gegen eine Lösung von (2.1). Gilt zusätzlich*

$$(2.27) \quad \mathcal{N}(F'(x^{\dagger})) \subset \mathcal{N}(F'(x)) \quad \text{für } x \in B_r(x_0),$$

*so konvergiert  $u_n$  für  $n \rightarrow \infty$  gegen  $x^{\dagger}$ .*

*Beweis.* Wir zeigen zunächst, dass die Norm des Fehlers streng monoton fällt, solange das Diskrepanzprinzip (2.18) noch nicht erfüllt ist, d.h. wir verifizieren

$$(2.28) \quad \|e_n\|_{\mathcal{X}} < \|e_{n-1}\|_{\mathcal{X}} \quad \text{für } n = 1, \dots, n_*.$$

In diesem Fall gilt

$$\begin{aligned} &\|\psi(-h_n K_n) \Delta F_n^{\delta} + A_n e_n\| \\ &\leq \left\| (\tilde{\Phi}_n - e^{-h_n K_n}) \right\| \|\Delta F_n^{\delta}\|_{\mathcal{Y}} + \|\Delta F_n^{\delta} + A_n e_n\|_{\mathcal{Y}} \\ &\leq \sup_{z \in [-\infty, 0]} |\varphi(z) - e^z| \|\Delta F_n^{\delta}\|_{\mathcal{Y}} + \delta + \eta \|\Delta F_n\|_{\mathcal{Y}} \\ &\leq (0.3 + \eta) \|\Delta F_n^{\delta}\|_{\mathcal{Y}} + \delta(1 + \eta). \end{aligned}$$

Wegen  $\delta < \|\Delta F_n^\delta\|/\tau$  folgt weiter:

$$\|\psi(-h_n K_n) \Delta F_n^\delta + A_n e_n\|_{\mathcal{Y}} \leq \frac{1 + \eta + \tau(0.3 + \eta)}{\tau} \|\Delta F_n^\delta\|_{\mathcal{Y}}.$$

Somit gilt (2.22) für  $n = 0$  und  $\nu = \mu\tau/(1 + \eta + \tau(0.3 + \eta))$ , was nach Voraussetzung größer als 1 ist. Wir können daher Lemma 2.14 anwenden und wie im Beweis von 2.15 folgt die Monotonie (2.28) per Induktion.

Wir fahren nun fort wie im Beweis von Hanke [30, Theorem 2.3], der der Vollständigkeit halber wiederholt wird. Analog zum Beweis von Satz 2.15 summieren wir (2.24), dieses mal aber nur für  $n = 0, \dots, n_* - 1$ , auf und erhalten mit Hilfe des Diskrepanzprinzips (2.18) und von Annahme 4

$$\begin{aligned} n_* \tau^2 \delta^2 &\leq \sum_{n=0}^{n_*-1} \|\Delta F_n^\delta\|_{\mathcal{Y}}^2 \\ &\leq \frac{\nu}{2\varphi^{-1}(\mu)(\nu-1)\mu^2} (\|e_{n_*}\|_{\mathcal{X}}^2 - \|e_0\|_{\mathcal{X}}^2) \\ &\leq \frac{\nu}{2\varphi^{-1}(\mu)(1-\nu)\mu^2} \|e_0\|_{\mathcal{X}}^2 \\ &< \infty \end{aligned}$$

und damit die Endlichkeit des Stoppindex  $n_*$ .

Wir betrachten des Weiteren das Verhalten von  $u_{n_*}$  für  $\delta \rightarrow 0$ . Dazu unterscheiden wir zwei Fälle. Sei zunächst  $n_* = n_*(\delta_m)$  für eine Nullfolge  $\{\delta_m\}_{m \in \mathbb{N}}$  und zugehörigen verrauschten Daten  $y^{\delta_m}$  beschränkt. Ferner sei  $n$  ein (endlicher) Häufungspunkt von  $\{n_*(\delta_m) : m \in \mathbb{N}\}$ . Wir nehmen ohne Beschränkung der Allgemeinheit an, dass  $n_*(\delta_m) = n$  für alle  $m \in \mathbb{N}$  (ansonsten betrachten wir eine Teilfolge von  $\{\delta_m\}_{m \in \mathbb{N}}$ ). Dann gelten mit den Iterierten  $x_n$  des exponentiellen Euler-Verfahrens bei exakten Daten

$$u_n = x_n^{\delta_m} \rightarrow x_n \quad \text{und} \quad F(x_n^{\delta_m}) \rightarrow F(x_n) \quad \text{für} \quad m \rightarrow \infty,$$

da die Iterierten  $x_n^\delta$  stetig von den Daten  $y^\delta$  abhängen. Wegen

$$\left\| y^{\delta_m} - F(x_{n_*}^{\delta_m}) \right\|_{\mathcal{Y}} \leq \tau \delta_m$$

folgt außerdem  $F(x_n) = y$  und mit Satz 2.15 schließlich

$$x_n^{\delta_m} \rightarrow x_* \quad \text{für} \quad m \rightarrow \infty,$$

wobei  $x_*$  wieder eine Lösung von (2.1) in  $B_r(x_0)$  ist. Es gelte nun  $n_*(\delta) \rightarrow \infty$  für  $\delta \rightarrow 0$ . Wegen (2.28) existiert zu  $\varepsilon > 0$  ein  $m(\varepsilon)$  so, dass  $\|x^\dagger - x_m\| \leq \varepsilon/2$  für  $m > m(\varepsilon)$  gilt. Ferner sei  $\delta(\varepsilon)$  so klein, dass  $n_*(\delta) > m(\varepsilon)$  für  $\delta < \delta(\varepsilon)$  gilt. Dann folgt aus (2.28)

$$\|e_{n_*(\delta)}\|_{\mathcal{X}} < \|e_m\|_{\mathcal{X}} \leq \|x_* - x_m\|_{\mathcal{X}} + \|x_m - u_m\|_{\mathcal{X}} \leq \varepsilon/2 + \|x_m - u_m\|_{\mathcal{X}}$$

für  $m = m(\varepsilon)$  und alle  $\delta < \delta(\varepsilon)$ . Wie eben folgt aus einem Stetigkeitsargument, dass  $\|x_m - u_m\|_{\mathcal{X}} < \varepsilon/2$  und damit

$$\|e_{n_*(\delta)}\|_{\mathcal{X}} < \varepsilon$$

für genügend kleines  $\delta$  gilt. Damit konvergiert auch im zweiten Fall  $u_{n_*(\delta)}$  gegen  $x_*$  für  $\delta \rightarrow 0$ .

Die Konvergenz gegen  $x^\dagger$  folgt nun, da unter der Voraussetzung (2.27) die Folge  $\{x_n\}_{n \in \mathbb{N}}$  gegen  $x^\dagger$  konvergiert.  $\square$

**Bemerkung.** Die Schrittweitenbedingung (2.20) bei der Umsetzung auf dem Rechner exakt zu erfüllen ist in der Regel nicht möglich. Dies ist jedoch nicht problematisch, da man die Parameter  $\mu$  und  $\nu$  in (2.20) durch variable Parameter  $\mu_n$  und  $\nu_n$  ersetzen kann. Um dann trotzdem noch die Gültigkeit von Satz 2.16 und Satz 2.15 zu gewährleisten, muss man sicherstellen, dass  $0 < \mu_{\min} \leq \mu_n \leq \mu_{\max} < 1$  für alle  $n \in \mathbb{N}_0$  (bzw.  $\leq n_*$ ) gilt. Durch eine flexible Steuerung der  $\mu_n$  lässt sich außerdem die Konvergenzgeschwindigkeit des Verfahren deutlich erhöhen (siehe Abschnitt 2.7).  $\diamond$

## 2.5 Optimale Konvergenzraten

Im vorherigen Abschnitt haben wir die Konvergenz des exponentiellen Euler-Verfahrens zur Regularisierung nichtlinearer schlecht gestellter Probleme nachgewiesen. Ziel dieses Abschnitts ist es nun, diese Aussage zu verschärfen, indem wir auch die Konvergenzgeschwindigkeit angeben. Es wird sich herausstellen, dass das Verfahren unter geeigneten Voraussetzungen mit optimaler Ordnung konvergiert. Hierfür benötigen wir allerdings zusätzlich zu Annahme 1, wie im kontinuierlichen Fall [80], auch die Annahmen 2 und 3.

Eine direkte Folgerung aus Annahme 3 erlaubt uns eine Kontrolle des bei der Linearisierung von  $F$  entstehenden Fehlers.

**Lemma 2.17.** *Unter Annahme 3 gilt für  $x \in B_r(x_0)$*

$$\|F(x) - F(x^\dagger) - F'(x^\dagger)(x - x^\dagger)\|_{\mathcal{Y}} \leq \frac{1}{2}C_+ \|x - x^\dagger\|_{\mathcal{X}} \|F'(x^\dagger)(x - x^\dagger)\|_{\mathcal{Y}}.$$

*Beweis.* Die Aussage folgt wie im Beweis von Lemma 2.11 mit  $\tilde{x} = x^\dagger$ .  $\square$

Als weitere Vorbereitung zeigen wir noch eine Konsequenz aus dem Diskrepanzprinzip.

**Lemma 2.18.** *Es gelten (2.2),  $\|F(x_0) - y^\delta\|_{\mathcal{Y}} > \tau\delta$  mit  $\tau > \frac{1+\eta}{1-\eta}$  und die Annahme 1. Wird das exponentielle Euler-Verfahren nach dem Diskrepanzprinzip (2.18) gestoppt, so gilt für  $n < n^*$*

$$\|\Delta F_n^\delta\|_{\mathcal{Y}} \leq \frac{\tau}{(\tau - 1)(1 - \eta)} \|A_+ e_n\|_{\mathcal{Y}}.$$

*Beweis.* Aus (2.11) und (2.2) ergibt sich

$$\|\Delta F_n^\delta\|_{\mathcal{Y}} \leq \|y^\delta - y\|_{\mathcal{Y}} + \|F(x^\dagger) - F(u_n)\|_{\mathcal{Y}} \leq \delta + \frac{1}{1 - \eta} \|A_+ e_n\|_{\mathcal{Y}}.$$

Mit Hilfe des Diskrepanzprinzips (2.18) und  $n < n^*$  folgt  $\delta < \frac{1}{\tau} \|\Delta F_n^\delta\|_{\mathcal{Y}}$  und damit die gewünschte Abschätzung.  $\square$

Nun können wir mit der eigentlichen Untersuchung der Konvergenzraten beginnen. Dazu betrachten wir zunächst die Fehlerrekursion.

**Lemma 2.19.** *Unter Annahme 3 erfüllt der Fehler  $e_n$  des exponentiellen Euler-Verfahrens die Rekursion*

$$(2.29) \quad e_{n+1} = \exp(-h_n J_+) e_n + h_n A_+^* (r_n + \tilde{\Phi}_{n,+}(y^\delta - y)).$$

wobei  $r_n = r_n^{(1)} - r_n^{(2)}$  in

$$\begin{aligned} r_n^{(1)} &= \tilde{\Phi}_{n,+}(F(x^\dagger) - F(u_n) + A_+e_n) - (I - R_n^*)\tilde{\Phi}_n\Delta F_n^\delta \\ r_n^{(2)} &= (\tilde{\Phi}_{n,+} - \tilde{\Phi}_n)\Delta F_n^\delta \end{aligned}$$

zerlegbar ist.

*Beweis.* Den Fehler im  $(n+1)$ -ten Schritt können wir mittels (2.17) umformen zu

$$\begin{aligned} e_{n+1} &= e_n + h_n\Phi_n A_n^* \Delta F_n^\delta \\ &= e_n + h_n\Phi_{n,+}\Delta F_n^\delta - h_n[\Phi_{n,+}A_+^* - \Phi_n A_n^*] \Delta F_n^\delta \\ &= e_n + h_n\Phi_{n,+}A_+^*(F(x^\dagger) - F(u_n)) - h_n[\Phi_{n,+}A_+^* - \Phi_n A_n^*] \Delta F_n^\delta \\ &\quad + h_n\Phi_{n,+}A_+^*(y^\delta - y). \end{aligned}$$

Mit Hilfe der Darstellung

$$e_n = \exp(-h_n J_+)e_n + h_n\Phi_{n,+}A_+^*A_+e_n$$

erhalten wir somit

$$\begin{aligned} e_{n+1} &= \exp(-h_n J_+)e_n + h_n A_+^* \tilde{\Phi}_{n,+}(F(x^\dagger) - F(u_n) + A_+e_n) \\ &\quad - h_n[\Phi_{n,+}A_+^* - \Phi_n A_n^*] \Delta F_n^\delta + h_n A_+^* \tilde{\Phi}_{n,+}(y^\delta - y). \end{aligned}$$

Aus Annahme 3 leiten wir nun die Gleichung

$$\begin{aligned} \Phi_{n,+}A_+^* - \Phi_n A_n^* &= A_+^* \tilde{\Phi}_{n,+} - A_n^* \tilde{\Phi}_n \\ &= A_+^* (\tilde{\Phi}_{n,+} - R_n^* \tilde{\Phi}_n) \\ &= A_+^* (\tilde{\Phi}_{n,+} - \tilde{\Phi}_n + (I - R_n^*)\tilde{\Phi}_n) \end{aligned}$$

ab und erhalten schließlich die gewünschte Darstellung

$$e_{n+1} = \exp(-h_n J_+)e_n + h_n A_+^* (r_n + \tilde{\Phi}_{n,+}(y^\delta - y)).$$

□

Für die weitere Analyse der Konvergenzraten setzen wir noch eine Beschränkung der  $j$ -ten Zeitschrittweite  $h_j$  durch die Gesamtzeitschrittweite  $t_j$  nach dem  $j$ -ten Schritt wie folgt voraus. Dazu existiere ein  $\varepsilon > 0$  so, dass gilt

$$(2.30) \quad h_j \leq c_h(1 + t_j)^{\gamma-\varepsilon} \quad \text{für } j \in \mathbb{N}_0.$$

Die maximale Schrittweite notieren wir mit

$$h_{\max}^{(n)} = \max_{0 \leq j < n} h_j \leq c_h(1 + t_n)^{\gamma-\varepsilon}.$$

Unter dieser weiteren Voraussetzung sind wir nun in der Lage, die Restterme  $r_n^{(i)}$ ,  $i = 1, 2$ , aus Lemma 2.19 abzuschätzen.

**Lemma 2.20.** *Es gelten (2.2) und die Annahmen 1 und 3 seien erfüllt. Unter Voraussetzung der Zeitschrittweitenbeschränkung (2.30) existieren Konstanten  $C_1 = C_1(C_+\tau, \eta)$ ,  $C_2 = C_2(C_+, \tau, \eta, r, c_h)$  und  $C_3 = C_3(C_+, \tau, \eta, r)$  so, dass*

$$\begin{aligned}\|r_n^{(1)}\|_{\mathcal{Y}} &\leq C_1 \|e_n\|_{\mathcal{X}} \|A_+e_n\|_{\mathcal{Y}}, \\ \|r_n^{(2)}\|_{\mathcal{Y}} &\leq C_2(1+t_n)^{\gamma-\varepsilon} \|e_n\|_{\mathcal{X}} \|A_+e_n\|_{\mathcal{Y}}\end{aligned}$$

und

$$\|K_+r_n^{(2)}\|_{\mathcal{Y}} \leq C_3 \|e_n\|_{\mathcal{X}} \|A_+e_n\|_{\mathcal{Y}}$$

für  $n < n_*$  gelten.

*Beweis.* Den ersten Summanden von  $r_n^{(1)}$  können wir wegen  $\|\tilde{\Phi}_{n,+}\| \leq 1$  mit Hilfe von Lemma 2.17 abschätzen durch  $\frac{C_+}{2} \|e_n\|_{\mathcal{X}} \|A_+e_n\|_{\mathcal{Y}}$ . Als obere Schranke für die Norm des zweiten Summanden erhalten wir mit Hilfe von Annahme 3 und Lemma 2.18  $C_+ \frac{\tau}{(\tau-1)(1-\eta)} \|e_n\|_{\mathcal{X}} \|A_+e_n\|_{\mathcal{Y}}$ . Die erste Ungleichung gilt also mit der Konstanten

$$C_1 = C_+ \left( \frac{1}{2} + \frac{\tau}{(\tau-1)(1-\eta)} \right).$$

Zur Beschränkung von  $\|r_n^{(2)}\|_{\mathcal{Y}}$  liefert uns der Mittelwertsatz zunächst

$$\tilde{\Phi}_{n,+} - \tilde{\Phi}_n = h_n \int_0^1 \varphi'(-h_n K_+ + s h_n (K_+ - K_n))(K_n - K_+) ds.$$

Die darin auftretende Differenz von  $K_n$  und  $K_+$  können wir wegen

$$K_n = A_n A_n^* = R_n A_+ A_+^* R_n^* = R_n K_+ R_n^*$$

umformen zu

$$\begin{aligned}(2.31) \quad K_n - K_+ &= (R_n - I)K_+ R_n^* + K_+ R_n^* - K_+ \\ &= (R_n - I)K_+ R_n^* + K_+(R_n^* - I).\end{aligned}$$

Die zweite gewünschte Schranke folgt nun wegen

$$\max_{z \geq 0} \varphi'(-z) = 1/2$$

und mit Lemma 2.18 sowie Annahme 3 und (2.30) aus

$$\begin{aligned}
\|r_n^{(2)}\|_{\mathcal{Y}} &\leq \left\| (\tilde{\Phi}_{n,+} - \tilde{\Phi}_n) \Delta F_n^\delta \right\|_{\mathcal{Y}} \\
&\leq h_n \left\| \int_0^1 \varphi'(-h_n K_+ + sh_n(K_+ - K_n))(K_n - K_+) ds \right\| \cdot \|\Delta F_n^\delta\|_{\mathcal{Y}} \\
&\leq c_h (1 + t_n)^{\gamma-\varepsilon} \left\| \int_0^1 \varphi'(-h_n K_+ + sh_n(K_+ - K_n)) ds \right\| \cdot \|K_n - K_+\| \\
&\quad \cdot \frac{\tau}{(\tau-1)(1-\eta)} \|A_+ e_n\|_{\mathcal{Y}} \\
&\leq c_h (1 + t_n)^{\gamma-\varepsilon} \frac{1}{2} C_+ \|e_n\|_{\mathcal{X}} (1 + \|R_n\|) \frac{\tau}{(\tau-1)(1-\eta)} \|A_+ e_n\|_{\mathcal{Y}} \\
&\leq C_+ c_h \frac{\tau}{2(\tau-1)(1-\eta)} (2 + rC_+) (1 + t_n)^{\gamma-\varepsilon} \|e_n\|_{\mathcal{X}} \|A_+ e_n\|_{\mathcal{Y}}.
\end{aligned}$$

Damit ist

$$C_2 = C_+ c_h \frac{\tau}{2(\tau-1)(1-\eta)} (2 + rC_+),$$

wobei  $r$  Radius der Kugel  $B_r(x_0)$  aus den Annahmen 1 und 3 ist. Zum Beweis der letzten Ungleichung nutzen wir

$$\begin{aligned}
K_+(\tilde{\Phi}_{n,+} - \tilde{\Phi}_n) &= K_+ \tilde{\Phi}_{n,+} - K_n \tilde{\Phi}_n + (K_n - K_+) \tilde{\Phi}_n \\
&= \frac{1}{h_n} (I - \exp(-h_n K_+)) - \frac{1}{h_n} (I - \exp(-h_n K_n)) + (K_n - K_+) \tilde{\Phi}_n \\
&= \frac{1}{h_n} (\exp(-h_n K_n) - \exp(-h_n K_+)) + (K_n - K_+) \tilde{\Phi}_n.
\end{aligned}$$

Für den ersten Summanden nutzen wir wie zuvor die Darstellung mittels des Mittelwertsatzes

$$\exp(-h_n K_n) - \exp(-h_n K_+) = h_n \int_0^1 \exp(-h_n K_+ + sh_n(K_+ - K_n))(K_n - K_+) dt$$

und verwenden zur Abschätzung des zweiten Summanden wiederum die Gleichung (2.31). Die gewünschte Ungleichung folgt dann mit

$$C_3 = C_+ \frac{2\tau}{(\tau-1)(1-\eta)} (2 + rC_+).$$

□

Dieses Lemma ermöglicht es nun einen großen Schritt in Richtung Ordnungsoptimalität zu gehen. Inspiriert von [32, Theorem 3.1] und [81, Theorem 5 & Proposition 7] werden wir eine Monotonie des Fehlers nachweisen. Im Vergleich zum kontinuierlichen Verfahren, der asymptotischen Regularisierung [80], fordern wir dabei zusätzlich die Einschränkung an die Zeitschrittweiten  $h_j$  (2.30).

**Satz 2.21.** *Es seien die Annahmen 1 bis 4 erfüllt und es gelten die Ungleichungen (2.2), (2.14) sowie  $\|F(x_0) - y^\delta\|_y > \tau\delta$ . Ferner sei die Folge der Zeitschrittweiten  $(h_j)_{j \geq 0}$  monoton wachsend und erfülle die Ungleichung (2.30) mit  $\varepsilon > 0$ . Dann existiert für genügend kleines  $\varrho > 0$  eine Konstante  $C_* = C_*(\tau, \eta, C_+, c_h, r, \gamma, \varepsilon)$  so, dass die Ungleichungen*

$$(2.32) \quad \|e_n\|_{\mathcal{X}} \leq C_* \frac{\varrho}{(1+t_n)^\gamma},$$

$$(2.33) \quad \|A_+ e_n\|_y \leq C_* \frac{\varrho}{(1+t_n)^{\gamma+1/2}}$$

für  $n < n_*$  gelten.

*Beweis.* Durch Auflösen der Fehlerrekursion (2.29) erhalten wir die für beliebiges  $n \in \mathbb{N}_0$  gültigen diskreten Versionen der Variation-der-Konstanten-Formel

$$(2.34) \quad e_n = \exp(-t_n J_+) e_0 + \sum_{j=0}^{n-1} h_j \exp(-(t_n - t_{j+1}) J_+) A_+^* (r_j + \tilde{\Phi}_{j,+}(y^\delta - y))$$

und

$$(2.35) \quad A_+ e_n = \exp(-t_n K_+) A_+ e_0 + \sum_{j=0}^{n-1} h_j \exp(-(t_n - t_{j+1}) K_+) K_+ (r_j + \tilde{\Phi}_{j,+}(y^\delta - y)).$$

Sei nun  $n < n_*$ . Mit Annahme 2 und

$$(2.36) \quad \|\exp(-t J_+) J_+^\alpha\| \leq \frac{1}{(1+t)^\alpha} \quad \text{für } \alpha \in [0, 1] \quad \text{und } t \geq 0$$

können wir die Norm des ersten Summanden von (2.34) abschätzen. Zur Abschätzung des zweiten Summanden aus (2.34) verwenden wir einerseits

$$(2.37) \quad \begin{aligned} & \left\| \sum_{j=0}^{n-1} h_j \exp(-(t_n - t_{j+1}) J_+) A_+^* \tilde{\Phi}_{j,+}(y^\delta - y) \right\|_y \\ & \leq \left\| \sum_{j=0}^{n-1} h_j \exp(-(t_n - t_{j+1}) J_+) J_+^{1/2} \Phi_{j,+} \right\| \|y^\delta - y\|_y \\ & \leq \left( \sup_{\lambda \in [0,1]} \int_0^{t_n} e^{-(t_n-s)\lambda^2} \lambda \, ds + h_{\max}^{(n)} \right) \delta \\ & \leq \left( \sqrt{t_n} \sup_{z \in [0, \infty]} \varphi(-z) \sqrt{z} + h_{\max}^{(n)} \right) \delta \\ & \leq \left( \sqrt{\frac{t_n}{2}} + h_{\max}^{(n)} \right) \delta \end{aligned}$$

und andererseits (2.36) mit  $\alpha = \frac{1}{2}$  sowie die ersten beiden Abschätzungen aus Lemma 2.20. Wir erhalten dann

$$\begin{aligned} \|e_n\|_{\mathcal{X}} &\leq \frac{\varrho}{(1+t_n)^\gamma} + \left( \sqrt{\frac{t_n}{2}} + h_{\max}^{(n)} \right) \delta \\ &\quad + C_1 \sum_{j=0}^{n-1} h_j \frac{1}{\sqrt{1+t_n-t_{j+1}}} \|e_j\|_{\mathcal{X}} \|A_+ e_j\|_{\mathcal{Y}} \\ &\quad + C_2 \sum_{j=0}^{n-1} h_j \frac{(1+t_j)^{\gamma-\varepsilon}}{\sqrt{1+t_n-t_{j+1}}} \|e_j\|_{\mathcal{X}} \|A_+ e_j\|_{\mathcal{Y}}. \end{aligned}$$

Auch in (2.35) schätzen wir den ersten Summanden wieder mit Hilfe von Annahme 2 und (2.36) ab. Ferner nutzen wir die Teleskopsumme

$$(2.38) \quad \sum_{j=0}^{n-1} h_j \exp(-(t_n - t_{j+1})K_+) K_+ \tilde{\Phi}_{j,+} = I - \exp(-t_n K_+)$$

für den zu  $y^\delta - y$  gehörenden Term. Des Weiteren verfahren wir wie zuvor mit dem Unterschied, dass wir den letzten Term  $j = n - 1$  getrennt behandeln. Dann erhalten wir

$$\begin{aligned} \|A_+ e_n\|_{\mathcal{Y}} &\leq \frac{\varrho}{(1+t_n)^{\gamma+1/2}} + \delta \\ &\quad + C_1 \sum_{j=0}^{n-1} h_j \frac{1}{1+t_n-t_{j+1}} \|e_j\|_{\mathcal{X}} \|A_+ e_j\|_{\mathcal{Y}} \\ &\quad + C_2 \sum_{j=0}^{n-2} h_j \frac{(1+t_j)^{\gamma-\varepsilon}}{1+t_n-t_{j+1}} \|e_j\|_{\mathcal{X}} \|A_+ e_j\|_{\mathcal{Y}} \\ &\quad + C_3 h_{n-1} \|e_{n-1}\|_{\mathcal{X}} \|A_+ e_{n-1}\|_{\mathcal{Y}}. \end{aligned}$$

Der Rest des Beweises läuft nun via Induktion über  $n$ .

Nach Annahme 2 gilt der Induktionsanfang  $n = 0$  für beliebiges  $C_* \geq 1$ . Wir nehmen nun an, dass die gewünschten Ungleichungen für alle Indizes kleiner gleich  $n - 1$  gelten. In diesem Fall gelten dann

$$\begin{aligned} \|e_n\|_{\mathcal{X}} &\leq \frac{\varrho}{(1+t_n)^\gamma} + \left( \sqrt{\frac{t_n}{2}} + h_{\max}^{(n)} \right) \delta \\ &\quad + C_*^2 \varrho^2 \left( C_1 S_n \left( \frac{1}{2}, 2\gamma + \frac{1}{2} \right) + C_2 S_n \left( \frac{1}{2}, \gamma + \frac{1}{2} + \varepsilon \right) \right), \end{aligned}$$

wie auch

$$(2.39) \quad \begin{aligned} \|A_+ e_n\|_Y &\leq \frac{\varrho}{(1+t_n)^{\gamma+1/2}} + \delta \\ &+ C_*^2 \varrho^2 \left( C_1 S_n \left( 1, 2\gamma + \frac{1}{2} \right) + C_2 S'_n \left( 1, \gamma + \frac{1}{2} + \varepsilon \right) \right) \\ &+ C_3 C_*^2 \varrho^2 c_h \frac{1}{(1+t_{n-1})^{\gamma+1/2+\varepsilon}}, \end{aligned}$$

wobei

$$(2.40) \quad S_n(\alpha, \beta) = \sum_{j=0}^{n-1} \frac{h_j}{(1+t_n - t_{j+1})^\alpha (1+t_j)^\beta},$$

$$(2.41) \quad S'_n(\alpha, \beta) = \sum_{j=0}^{n-2} \frac{h_j}{(1+t_n - t_{j+1})^\alpha (1+t_j)^\beta}.$$

Die Abschätzungen für die Summen  $S_n$  und  $S'_n$  aus dem an späterer Stelle vorgestellten Lemma 2.24 führen auf

$$(2.42) \quad \|e_n\|_X \leq \frac{\varrho}{(1+t_n)^\gamma} \left( 1 + C_*^2 \varrho C_4 \right) + \left( \sqrt{\frac{t_n}{2}} + h_{\max}^{(n)} \right) \delta,$$

$$(2.43) \quad \|A_+ e_n\|_Y \leq \frac{\varrho}{(1+t_n)^{\gamma+1/2}} \left( 1 + C_*^2 \varrho C_5 \right) + \delta$$

mit

$$\begin{aligned} C_4 &= C_1 C_9 + C_2 C_7, \\ C_5 &= C_1 C_9 + C_2 C_8 + C_3 c_h (1+c_h)^{\gamma+1/2+\varepsilon}. \end{aligned}$$

Zur Abschätzung des letzten Summanden in (2.39) verwendeten wir, dass mit (2.30) aus der Ungleichung

$$1 + t_{j+1} = 1 + t_j + h_j \leq (1 + c_h)(1 + t_j)$$

unmittelbar

$$(2.44) \quad \frac{1}{(1+t_j)^\beta} \leq (1+c_h)^\beta \frac{1}{(1+t_{j+1})^\beta},$$

folgt.

Das Diskrepanzprinzip (2.18) zusammen mit Lemma 2.18 liefert dann

$$\begin{aligned} \delta &\leq \frac{1}{(\tau-1)(1-\eta)} \|A_+ e_n\|_Y \\ &\leq \frac{1}{(\tau-1)(1-\eta)} \left( \frac{\varrho}{(1+t_n)^{\gamma+1/2}} \left( 1 + C_*^2 \varrho C_5 \right) + \delta \right). \end{aligned}$$

Damit haben wir mit (2.14) für alle  $n < n_*$

$$(2.45) \quad \delta \leq C_6 \frac{\varrho}{(1+t_n)^{\gamma+1/2}}$$

mit der Konstanten

$$C_6 = \frac{1}{(\tau-1)(1-\eta)-1} (1 + C_*^2 \varrho C_5).$$

Wir setzen nun diese Schranke für  $\delta$  in (2.42) ein und erhalten mit (2.30)

$$\begin{aligned} \|e_n\|_{\mathcal{X}} &\leq \frac{\varrho}{(1+t_n)^\gamma} (1 + C_*^2 \varrho C_4) + \left( \sqrt{\frac{t_n}{2}} + h_{\max}^{(n)} \right) C_6 \frac{\varrho}{(1+t_n)^{\gamma+1/2}} \\ &\leq \frac{\varrho}{(1+t_n)^\gamma} \left( 1 + C_*^2 \varrho C_4 + C_6 \frac{\sqrt{\frac{t_n}{2}} + h_{\max}^{(n)}}{\sqrt{1+t_n}} \right) \\ &\leq \frac{\varrho}{(1+t_n)^\gamma} \left( 1 + C_*^2 \varrho C_4 + C_6 (2^{-1/2} + c_h) \right). \end{aligned}$$

Die Ungleichung (2.32) gilt also, falls

$$(2.46) \quad 1 + C_*^2 \varrho C_4 + C_6 (2^{-1/2} + c_h) \leq C_*$$

erfüllt ist. Zum Beweis der zweiten Schranke setzen wir (2.45) in (2.43) ein und erhalten

$$\begin{aligned} \|A_+ e_n\|_{\mathcal{Y}} &\leq \frac{\varrho}{(1+t_n)^{\gamma+1/2}} (1 + C_*^2 \varrho C_5) + C_6 \frac{\varrho}{(1+t_n)^{\gamma+1/2}} \\ &\leq \frac{\varrho}{(1+t_n)^{\gamma+1/2}} \left( 1 + C_*^2 \varrho C_5 + C_6 \right). \end{aligned}$$

Somit gilt (2.33), falls

$$(2.47) \quad 1 + C_*^2 \varrho C_5 + C_6 \leq C_*$$

erfüllt ist. Die beiden Bedingungen (2.46) und (2.47) sind für genügend kleines  $\varrho$  erfüllt, womit der Satz bewiesen ist.  $\square$

**Lemma 2.22.** *Seien  $S_n, S'_n$  gegeben durch (2.40) bzw. (2.41),  $0 < \alpha \leq 1$  und  $\beta > 0$ . Dann gilt*

$$S_n(\alpha, \beta) \leq (1 + c_h)^\beta \left( I_{t_n}(\alpha, \beta) + \frac{h_{\max}^{(n)}}{(1+t_n)^\beta} \right),$$

wobei  $I_t(\alpha, \beta)$  durch

$$(2.48) \quad I_t(\alpha, \beta) = \int_0^t (1+t-s)^{-\alpha} (1+s)^{-\beta} ds$$

definiert ist. Ist zusätzlich die Schrittweitenfolge monoton wachsend, so gilt

$$S'_n(\alpha, \beta) \leq (1 + c_h)^\beta I_{t_n}(\alpha, \beta).$$

*Beweis.* Mit Hilfe von (2.44) können wir in beiden Summen wie folgt abschätzen

$$S_n(\alpha, \beta) \leq (1 + c_h)^\beta \sum_{j=0}^{n-1} \frac{h_j}{(1 + t_n - t_{j+1})^\alpha (1 + t_{j+1})^\beta} =: \tilde{S}_n(\alpha, \beta),$$

$$S'_n(\alpha, \beta) \leq (1 + c_h)^\beta \sum_{j=0}^{n-2} \frac{h_j}{(1 + t_n - t_{j+1})^\alpha (1 + t_{j+1})^\beta} =: \tilde{S}'_n(\alpha, \beta).$$

Die so modifizierten Summen  $\tilde{S}_n$  und  $\tilde{S}'_n$  entsprechen Riemann-Summen wobei das Integral jeweils am rechten Randpunkt der Teilintervalle ausgewertet wird. Da der Integrand konvex

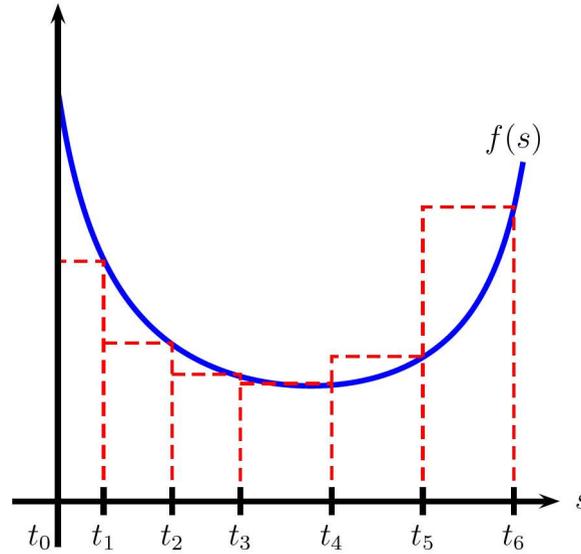


Abbildung 2.1: Skizze der Funktion  $f(s) = \frac{1}{\sqrt{1+t_n-s}(1+s)^{\gamma+1/2}}$  und einer zugehörigen Riemann-Summe.

ist (siehe Abbildung 2.1), ist  $\tilde{S}_n$  auf dem ersten Teil des Integrationsintervalls eine Riemann-Untersumme und auf dem restlichen Teil eine Riemann-Obersumme. Um diese durch das Integral abschätzen zu können verschieben wir die zur Obersumme gehörenden Kästchen (siehe Abbildung 2.1) um  $h_{\max}^{(n)}$  nach rechts. Da das letzte Kästchen dann über das Intervall  $[0, t_n]$  hinausragt erhalten wir den Zusatzterm

$$(1 + c_h)^\beta \frac{h_{\max}^{(n)}}{(1+t_n)^\beta}.$$

Die restlichen Kästchen liegen nun unter dem Integranden und wir können die Summe durch das Integral abschätzen.

Die Abschätzung für  $S'_n$  folgt mit der gleichen Idee. Der Zusatzterm entfällt dabei. Da die Summe nur bis  $n-2$  läuft und die Schrittweiten monoton wachsend sind, also  $h_{\max}^{(n)} = h_{n-1}$  ist liegen alle verschobenen Kästchen nicht nur unter dem Integranden sondern auch innerhalb des Intervalls  $[0, t_n]$ .  $\square$

Das folgende Lemma liefert uns Abschätzungen für die in Lemma 2.22 auftretenden Integrale.

**Lemma 2.23.** *Es existieren Konstanten  $c_{\alpha,\beta}$  so, dass*

$$I_t(\alpha, \beta) \leq \begin{cases} c_{\alpha,\beta} (1+t)^{1-\alpha-\beta}, & 0 \leq \alpha, \beta < 1, \\ c_{\alpha,\beta} \frac{1 + \log(2+t)}{(1+t)^\beta}, & \alpha = 1, 0 \leq \beta < 1, \\ c_{\alpha,\beta} (1+t)^{-\alpha}, & 0 \leq \alpha \leq 1, \beta > 1. \end{cases}$$

*Beweis.* Für  $0 \leq \alpha, \beta < 1$  ergibt sich das gewünschte Resultat aus

$$\int_0^t (1+t-s)^{-\alpha} (1+s)^{-\beta} ds \leq \int_0^t (t-s)^{-\alpha} s^{-\beta} ds = \frac{\Gamma(1-\alpha)\Gamma(1-\beta)}{\Gamma(2-\alpha-\beta)} t^{1-\alpha-\beta}.$$

Ist  $\alpha = 0$  und  $\beta > 1$ , so folgt die Behauptung mit

$$I_t(0, \beta) = \int_0^t (1+s)^{-\beta} ds \leq c_{0,\beta}.$$

Für  $0 < \alpha < 1$  und  $\beta > 1$  haben wir

$$\begin{aligned} I_t(\alpha, \beta) &= \int_0^{t/2} (1+t-s)^{-1} (1+s)^{-\beta} ds + \int_{t/2}^t (1+t-s)^{-1} (1+s)^{-\beta} ds \\ &\leq \left(\frac{2}{2+t}\right)^\alpha \int_0^{t/2} (1+s)^{-\beta} ds + \left(\frac{2}{2+t}\right)^\beta \int_{t/2}^t (1+t-s)^{-\alpha} ds \\ &\leq \tilde{c}_{\alpha,\beta} [(1+t)^{-\alpha} + (1+t)^{-\alpha-\beta+1} + (1+t)^{-\beta}] \\ &\leq c_{\alpha,\beta} (1+t)^{-\alpha}. \end{aligned}$$

Ist  $\alpha = 1$  und  $\beta > 1$  so folgt wie eben

$$\begin{aligned} I_t(1, \beta) &\leq \tilde{c}_{1,\beta} [(1+t)^{-\beta} + (1+t)^{-1} + \log(2+t)(1+t)^{-\beta}] \\ &\leq c_{1,\beta} (1+t)^{-1}. \end{aligned}$$

Auch im Fall  $\alpha = 1$  und  $\beta < 1$  erhalten wir wie zuvor schon

$$\begin{aligned} I_t(1, \beta) &\leq \tilde{c}_{1,\beta} [(1+t)^{-\beta} + (1+t)^{-1} + \log(2+t)(1+t)^{-\beta}] \\ &\leq c_{1,\beta}(1 + \log(2+t))(1+t)^{-\beta}. \end{aligned}$$

Das Integral für  $\alpha = \beta = 1$  können wir mit der Partialbruchzerlegung

$$\frac{1}{(1+t-s)(1+s)} = \frac{1}{(2+t)(1+t-s)} + \frac{1}{(2+t)(1+s)}$$

exakt berechnen und erhalten schließlich die letzte noch fehlende Schranke aus

$$I_t(1, 1) = 2\log(1+t)(2+t)^{-1}.$$

□

Nun zeigen wir das im Beweis von Satz 2.21 benötigte Abklingverhalten der Summen  $S_n$  und  $S'_n$ .

**Lemma 2.24.** *Sei  $0 < \gamma \leq 1/2$  und  $S_n, S'_n$  seien gegeben durch (2.40) bzw. (2.41). Ferner gelte die Schrittweitenbeschränkung (2.30).*

1. Für  $0 \leq \alpha \leq 1/2$  existiert eine Konstante  $C_7 = C_7(\alpha, \gamma, \varepsilon, c_h)$  so, dass

$$S_n \left( \alpha, \gamma + \frac{1}{2} + \varepsilon \right) \leq C_7 \frac{1}{(1+t_n)^{\alpha+\gamma-1/2}}.$$

2. Die Schrittweiten  $h_j$  seien zusätzlich monoton wachsend. Dann existiert eine Konstante  $C_8 = C_8(\gamma, \varepsilon, c_h)$  so, dass

$$S'_n \left( 1, \gamma + \frac{1}{2} + \varepsilon \right) \leq C_8 \frac{1}{(1+t_n)^{\gamma+1/2}}.$$

3. Für  $\alpha = \frac{1}{2}, 1, \frac{1}{2} - \gamma$  existiert eine Konstante  $C_9 = C_9(\alpha, \gamma, c_h)$  so, dass

$$S_n \left( \alpha, 2\gamma + \frac{1}{2} \right) \leq C_9 \frac{1}{(1+t_n)^{\alpha+\gamma-1/2}}.$$

*Beweis.* Die zweite Abschätzung folgt unmittelbar aus den Lemmas 2.22 und 2.23. Im ersten Fall verwenden wir zusätzlich zu den eben verwendeten Lemmas, dass durch (2.30) die maximale Schrittweite durch  $h_{\max}^{(n)} \leq c_h(1+t_n)^\gamma$  beschränkt ist. Die letzte Abschätzung folgt wie die Erste. Allerdings benutzen wir anstatt Lemma 2.23 die Ungleichungen

$$I_t \left( \alpha, 2\gamma + \frac{1}{2} \right) \leq \tilde{c}_{\alpha,\gamma} \frac{1}{(1+t)^{\alpha+\gamma-1/2}}, \quad \alpha = \frac{1}{2}, 1, \frac{1}{2} - \gamma,$$

welche [80, Proposition 6] entnommen sind. □

Nach all diesen Vorbereitungen können wir nun den zentralen Satz dieses Kapitels, die Ordnungsoptimalität des exponentiellen Euler-Verfahrens beweisen.

**Satz 2.25.** *Seien die Annahmen 1 bis 4 und die Ungleichungen (2.2) und (2.14) erfüllt. Ferner seien die Schrittweiten monoton wachsend und erfüllen (2.30) für alle  $j \leq n_*$ . Der Stoppindex  $n_*$  sei dabei durch das Diskrepanzprinzip (2.18) gewählt. Dann existiert für genügend kleines  $\varrho$  eine Konstante  $C = C(\tau, \eta, C_+, c_h, r, \gamma, \varepsilon) > 0$  so, dass*

$$\|e_{n_*}\|_{\mathcal{X}} \leq C \varrho^{1/(2\gamma+1)} \delta^{2\gamma/(2\gamma+1)}$$

*gilt.*

*Beweis.* Den Fehler  $e_{n_*}$  können wir mittels (2.34) für  $n = n_*$  in der Form

$$e_{n_*} = J_+^\gamma v_* + \sum_{j=0}^{n_*-1} h_j \exp(-(t_{n_*} - t_{j+1})J_+) A_+^* \tilde{\Phi}_{j,+} (y^\delta - y)$$

schreiben, wobei

$$v_* = \exp(-t_{n_*} J_+) w + \sum_{j=0}^{n_*-1} h_j \exp(-(t_{n_*} - t_{j+1})J_+) J_+^{-\gamma} A_+^* r_j$$

ist. Wir bemerken dazu, dass wegen  $\gamma \leq \frac{1}{2}$  der Operator

$$J_+^{-\gamma} A_+^* : \mathcal{N}(A_+^*)^\perp \rightarrow \mathcal{X}$$

beschränkt ist.

Als nächstes beschränken wir  $v_*$  mit Hilfe von (2.36) und der Lemmas 2.20 und 2.24 durch

$$\begin{aligned} \|v_*\|_{\mathcal{X}} &= \|\exp(-t_{n_*} J_+) w\|_{\mathcal{X}} + \left\| \sum_{j=0}^{n_*-1} h_j \exp(-(t_{n_*} - t_{j+1})J_+) J_+^{-\gamma} A_+^* r_j \right\|_{\mathcal{X}} \\ &\leq \varrho + C_*^2 \varrho^2 \left( C_1 S_{n_*} \left( \frac{1}{2} - \gamma, 2\gamma + \frac{1}{2} \right) + C_2 S_{n_*} \left( \frac{1}{2} - \gamma, \gamma + \frac{1}{2} + \varepsilon \right) \right) \\ &\leq C_{10} \varrho \end{aligned}$$

mit

$$C_{10} = 1 + C_*^2 \varrho (C_1 C_9 + C_2 C_7).$$

Wegen

$$\begin{aligned} A_+ J_+^\gamma v_* &= A_+ e_{n_*} - \sum_{j=0}^{n_*-1} h_j \exp(-(t_{n_*} - t_{j+1})K_+) K_+^* \tilde{\Phi}_{j,+} (y^\delta - y) \\ &= A_+ e_{n_*} + (\exp(-t_{n_*} K_+) - I)(y^\delta - y) \end{aligned}$$

folgt außerdem

$$\begin{aligned}
\left\| J_+^{\gamma+1/2} v_* \right\|_{\mathcal{X}} &= \|A_+ J_+^\gamma v_*\|_{\mathcal{Y}} \\
&\leq \|A_+ e_{n_*}\|_{\mathcal{Y}} + \delta \\
&\leq (1 + \eta) \|F(u_{n_*}) - F(x^\dagger)\|_{\mathcal{Y}} + \delta \\
&\leq (1 + \eta) \left( \|\Delta F_{n_*}^\delta\|_{\mathcal{Y}} + \delta \right) + \delta \\
&\leq ((1 + \eta)(1 + \tau) + 1) \delta \\
&= C_{11} \delta
\end{aligned}$$

mit

$$C_{11} = (1 + \eta)(1 + \tau) + 1$$

aus (2.11) und dem Diskrepanzprinzip (2.18).

Mit Hilfe der Interpolationsungleichung (1.8) erhalten wir wie bei der Abschätzung des besten schlimmsten Fehlers für lineare Regularisierungsverfahren (siehe Beweis von Lemma 1.6)

$$\|J_+^\gamma v_*\|_{\mathcal{X}} \leq \left\| J_+^{\gamma+1/2} v_* \right\|_{\mathcal{X}}^{2\gamma/(2\gamma+1)} \|v_*\|_{\mathcal{X}}^{1/(2\gamma+1)} \leq (C_{11} \delta)^{2\gamma/(2\gamma+1)} (C_{10} \varrho)^{1/(2\gamma+1)}.$$

Zusammen mit der für alle  $n \in \mathbb{N}$  gültigen Ungleichung (2.37) erhalten wir unter Verwendung von (2.30)

$$\begin{aligned}
\|e_{n_*}\|_{\mathcal{X}} &\leq \|J_+^\gamma v_*\|_{\mathcal{X}} + \left\| \sum_{j=0}^{n_*-1} h_j \exp(-(t_{n_*} - t_{j+1})J_+) A_+^* \tilde{\Phi}_{j,+} (y^\delta - y) \right\|_{\mathcal{X}} \\
(2.49) \quad &\leq (C_{11} \delta)^{2\gamma/(2\gamma+1)} (C_{10} \varrho)^{1/(2\gamma+1)} + (2^{-1/2} + c_h) \sqrt{t_{n_*} + 1} \delta.
\end{aligned}$$

Die Abschätzung (2.45) für  $n = n_* - 1$  kombiniert mit (2.44) liefert ferner

$$\sqrt{1 + t_{n_*}} \delta^{1/(2\gamma+1)} \leq C_6^{1/(2\gamma+1)} (1 + c_h)^{1/2} \varrho^{1/(2\gamma+1)}.$$

Setzen wir dies nun in (2.49) ein, so erhalten wir das gewünschte Resultat mit der Konstanten

$$C = C_{11}^{2\gamma/(2\gamma+1)} C_{10}^{1/(2\gamma+1)} + C_6^{1/(2\gamma+1)} (1 + c_h)^{1/2} (2^{-1/2} + c_h).$$

□

**Bemerkung.** Die Bedingung der Monotonie der Schrittweiten in Satz 2.25 ist nicht zwingend notwendig. Man muss lediglich sicherstellen, dass die Ungleichungen aus Lemma 2.23 erfüllt sind. ◊

Zum Abschluss dieses Abschnitts über die Konvergenzraten wollen wir die hier vorgestellten Ergebnisse noch in die vorhandene Literatur einordnen. Da wir das von uns untersuchte nichtlineare exponentielle Euler-Verfahren auch als Newton-Verfahren mit dem linearen exponentiellen-Euler Verfahren als innerer Regularisierung interpretieren können (siehe Abschnitt 2.3), ist natürlich auch die Theorie von Rieder [71, 72] anwendbar. Dort wird allerdings unter der Verschärfung (2.12) von Annahme 3 nur eine Konvergenzrate  $O(\delta^{(2\gamma-\alpha)/(1+2\gamma)})$  für  $\gamma \in (\alpha, \frac{1}{2}]$  mit einem  $\alpha > 0$  bewiesen, also nicht die hier gezeigte Ordnungsoptimalität.

Jin und Tautenhahn zeigen in [48, Example 3], dass der Ansatz 2.6, versehen mit dem Filter des linearen exponentiellen Euler-Verfahrens, unter geeigneten Voraussetzungen auch ein ordnungsoptimales Regularisierungsverfahren für beliebiges  $\gamma > 0$  liefert. Allerdings sind die dort verwendeten Voraussetzungen etwas restriktiver als die in dieser Arbeit benutzten. So ist wie bei Rieder [71, 72] die Voraussetzung (2.12) nötig und ferner muss die Abschätzung

$$\|(F'(x) - F'(\tilde{x}))z\|_{\mathcal{Y}} \leq c_1 \|x - \tilde{x}\|_{\mathcal{X}} \|F'(\tilde{x})z\|_{\mathcal{Y}} + c_2 \|F'(\tilde{x})(x - \tilde{x})\|_{\mathcal{Y}} \|z\|_{\mathcal{X}}$$

für alle  $x, \tilde{x} \in B_r(x^\dagger)$  und  $z \in \mathcal{X}$  erfüllt sein. Ferner wird eine Monotonie und maximal lineare Wachstumsrate der Schrittweiten vorausgesetzt. Jin und Tautenhahn wiesen zusätzlich noch Konvergenzraten unter logarithmischen Quellbedingungen

$$x_0 - x^\dagger = (-\log(J_+))^{-\gamma} w$$

mit  $\gamma > 0$  und  $w \in \mathcal{X}$  oder unter einer Lipschitz-Bedingung an  $F'$  nach. Wie die Autoren in [48, Remark 2] bemerken sind die in [48] verwendeten Beweistechniken nicht auf den in dieser Arbeit betrachteten Ansatz übertragbar.

## 2.6 Implementierung

Bevor wir das in den vorangegangenen Abschnitten untersuchte Verfahren am Rechner testen können, müssen wir uns noch ein paar Gedanken über die Implementierung machen. Natürlich muss jeder Umsetzung auf dem Rechner eine Diskretisierung des Problems (2.3) vorausgehen. Auf diese gehen wir, da sie nicht immer gleich sein wird, in den einzelnen Beispielen gesondert ein. Wir schreiben die diskreten Variablen wieder fettgedruckt.

Wie zuvor schon bemerkt können wir unsere Methode auch als ein Newton-artiges Verfahren interpretieren, bei dem als innere Regularisierung das exponentielle Euler-Verfahren für lineare Probleme angewandt wird. Dies machen wir uns bei der Implementierung zu nutze, indem wir das lineare Verfahren (siehe Abschnitt 1.5) übernehmen und dort nur das Diskrepanzprinzip auf das in (2.20) definierte innere Diskrepanzprinzip anpassen. Insbesondere verwenden wir auch wieder Krylov-Verfahren zur Berechnung der Matrixfunktionen. Im Unterschied zur Implementierung im linearen Fall hat es sich bei nichtlinearen Problemen als sinnvoll erwiesen, wie in [39] vorgeschlagen, die Näherungslösung nicht in jedem Krylov-Raum, sondern nur, wenn

$$m \in \{1, 2, 3, 4, 6, 8, 11, 15, 20, 27, 36, 46, 57, 70, 85, 100\} =: \mathcal{I}$$

ist zu berechnen.

Im Gegensatz zum Verfahren (2.17) müssen wir im inneren Diskrepanzprinzip (2.20) bzw. der Vereinfachung (2.21) eine Funktion der Matrix  $\mathbf{K}_n$  mit  $\Delta \mathbf{F}_n^\delta$  multiplizieren und nicht eine Funktion mit der Matrix  $\mathbf{J}_n$ . Um trotzdem nur einen Krylov-Raum aufbauen zu müssen, schreiben wir das Verfahren (2.17) um zu

$$\mathbf{u}_{n+1} = \mathbf{u}_n + h_n \mathbf{A}_n^* \Delta \tilde{\mathbf{u}}_n = \mathbf{u}_n + h_n \mathbf{A}_n^* \varphi(-h_n \mathbf{K}_n) \Delta \mathbf{F}_n^\delta.$$

Dadurch müssen wir nur  $\varphi(-h_n \mathbf{K}_n) \Delta \mathbf{F}_n^\delta$  approximieren. Da wir es ferner im diskretisierten Fall mit einer skalierten euklidischen Norm zu tun haben, können wir uns die Multiplikation mit der Matrix  $\mathbf{V}_m$  (siehe Anhang B) zur Überprüfung des inneren Abbruchkriteriums sparen. Das Verfahren haben wir in den Algorithmen 5 und 4 schematisch dargestellt.

Von uns durchgeführte numerische Experimente zeigten, dass es günstiger ist, den Parameter  $\mu$  aus dem inneren Diskrepanzprinzip (2.20) nicht konstant zu wählen, sondern flexibel zu steuern. Dazu verwenden wir eine Variante der von Rieder [73, Kapitel 7.5.3.4] vorgeschlagenen dynamischen Steuerung der  $\mu_n$ . Ausgehend von einem Startwert  $\mu_{\text{start}} \in (0, 1)$  wählen wir  $\mu_{\text{max}} \in (\mu_{\text{start}}, 1)$  und berechnen dann in jedem Schritt

$$\hat{\mu}_n = \mu_{\text{max}} \max \left\{ \frac{\tau \delta}{\|\Delta \mathbf{F}_n^\delta\|}, \tilde{\mu}_n \right\},$$

wobei

$$\tilde{\mu}_n = \begin{cases} 1 - \left(\frac{r_{n-2}}{r_{n-1}}\right)^{3/2} (1 - \mu_{n-1}), & r_{n-1} \geq r_{n-2} \quad \text{und} \quad n > 1, \\ \mu_{n-1}, & r_{n-1} < r_{n-2} \quad \text{und} \quad n > 1, \\ \mu_{\text{start}}, & n \leq 1, \end{cases}$$

Dabei sind die  $r_n$  die Dimensionen des tatsächlich benötigten Krylov-Raums im  $n$ -ten Schritt. Diese adaptive Steuerung erfüllt  $0 < \mu_{\text{start}} \leq \hat{\mu}_n \leq \mu_{\text{max}} < 1$  und damit die Voraussetzungen zur Konvergenz des Verfahrens.

In jedem Fall ist es nützlich wie bei inexakten Newton-Verfahren für gut gestellte Probleme eine sogenannte *safeguarding*-Strategie einzusetzen (siehe [18]). Dies verhindert, dass man, falls der nichtlineare Defekt  $\|\Delta \mathbf{F}_{n-1}^\delta\|$  schon nahe bei  $\tau\delta$  liegt, in der inneren Iteration zu genau approximiert, also  $\mu_n$  zu klein wählt. Wir schlagen daher in Analogie zu [18] vor,

$$(2.50) \quad \mu_n = \max\left\{\hat{\mu}_n, \mu_{\text{max}} \frac{\tau\delta}{\|\Delta \mathbf{F}_n^\delta\|}\right\}$$

zu wählen.

---

#### Algorithmus 4 Exponentielles Euler-Verfahren für nichtlineare Probleme

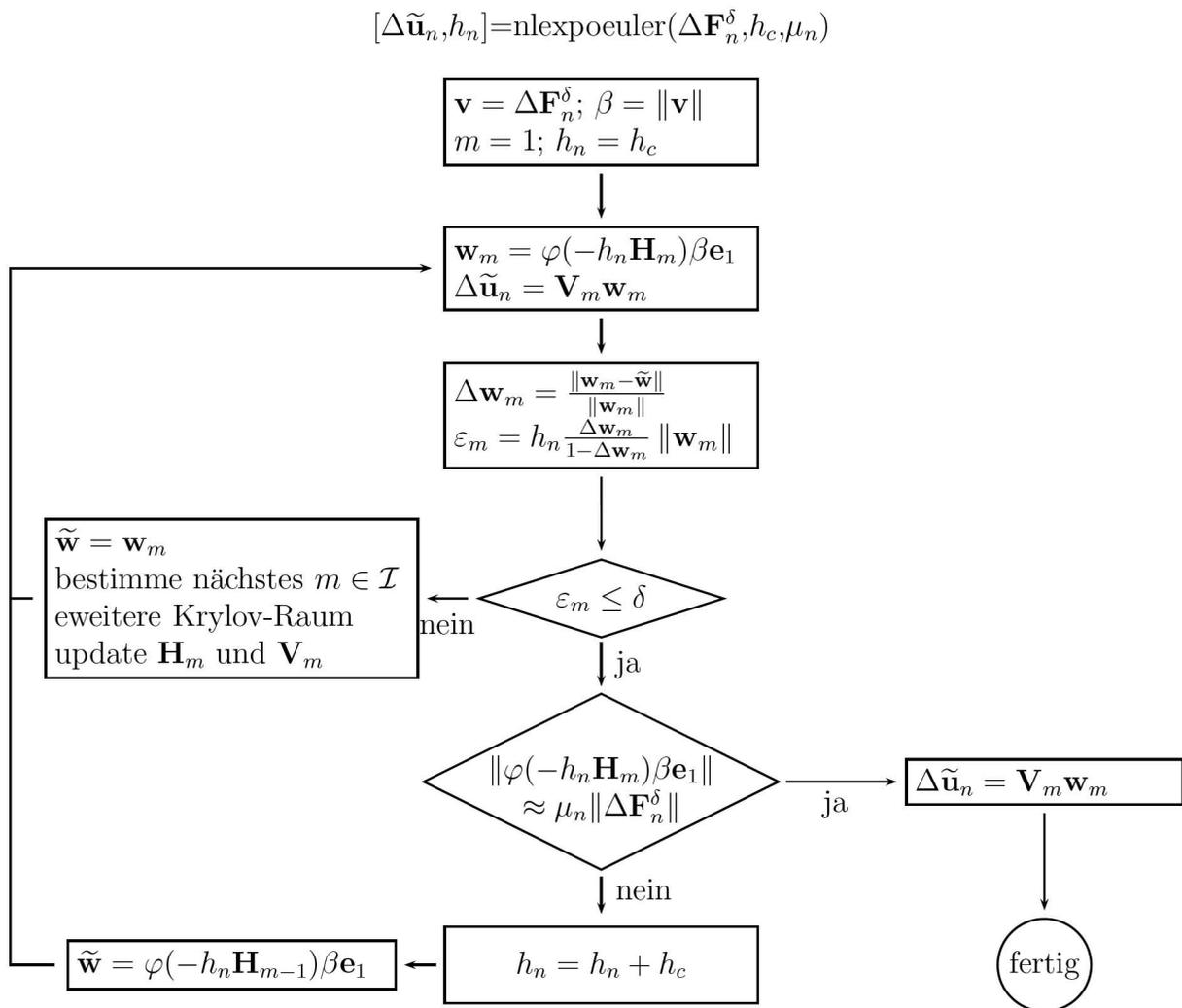
---

```

[ $\mathbf{u}_{n_*}, t$ ]=nichtlin-expoeuler( $\mathbf{x}_0, \tau, \delta$ )
 $t = 0$ ;  $n = 0$ ;  $\mu_0 = \mu_{\text{start}}$ 
 $\mathbf{u}_0 = \mathbf{x}_0$ 
while  $t \leq t_{\text{max}}$  &  $\|\Delta \mathbf{F}_n^\delta\| > \tau\delta$  do {Zeitschleife}
  [ $\Delta \tilde{\mathbf{u}}_n, h_n$ ]=nlexpoeuler( $\Delta \mathbf{F}_n^\delta, h_c, \mu_n$ )
   $\mathbf{u}_{n+1} = \mathbf{u}_n + h_n \mathbf{A}_n^* \Delta \tilde{\mathbf{u}}_n$ 
   $t = t + h_n$ 
   $n = n + 1$ 
  berechne neues  $\Delta \mathbf{F}_n^\delta$ 
  bestimme neues  $\mu_n$  nach (2.50)
end while
 $\mathbf{u}_{n_*} = \mathbf{u}_n$ 

```

---

**Algorithmus 5** Innere Iteration des nichtlinearen exponentiellen Euler-Verfahrens

## 2.7 Numerische Beispiele

Wir haben bisher die Konvergenz und Ordnungsoptimalität des exponentiellen Euler-Verfahrens zur Regularisierung nichtlinearer schlecht gestellter Probleme nachgewiesen, und eine effiziente Implementierung angegeben. Nun wollen wir noch einige numerische Experimente präsentieren, die Regularisierungseigenschaften widerspiegeln und die Wettbewerbsfähigkeit dieses Verfahrens gegenüber Standardverfahren darlegen.

### 2.7.1 Parameteridentifikation in einer elliptischen partiellen Differentialgleichung

Das erste numerische Beispiel für nichtlineare Probleme ist Rieder [72] entnommen. Ziel ist es, den nichtnegativen Parameter  $x$  des zweidimensionalen elliptischen Randwertproblems

$$(2.51) \quad \begin{aligned} -\Delta y + xy &= f && \text{in } \Omega \\ y &= g && \text{auf } \partial\Omega \end{aligned}$$

auf  $\Omega = [0, 1]^2$  aus der gegebenen Lösung  $y$  und den bekannten rechten Seiten  $f \in L^2(\Omega)$  und  $g \in H^2(\partial\Omega)$  zu rekonstruieren. Hat  $y$  keine Nullstellen, so können wir die Lösung dieses Problems explizit darstellen durch

$$x^\dagger = \frac{f + \Delta y}{y}.$$

Wir erkennen daran, dass die Lösung zwar eindeutig ist, aber nicht stetig von den Daten  $y$  abhängt. Da wir aber von gestörten Daten  $y^\delta$  ausgehen müssen, ist eine Regularisierung notwendig. Dazu definieren wir den nichtlinearen Operator  $F : \mathcal{D}(F) \subset L^2(\Omega) \rightarrow L^2(\Omega)$  als die Abbildung, die den Parameter  $x$  auf die Lösung des Randwertproblems (2.51) abbildet. Die Wohldefiniertheit dieser Abbildung ist für genügend kleines  $\kappa > 0$  in

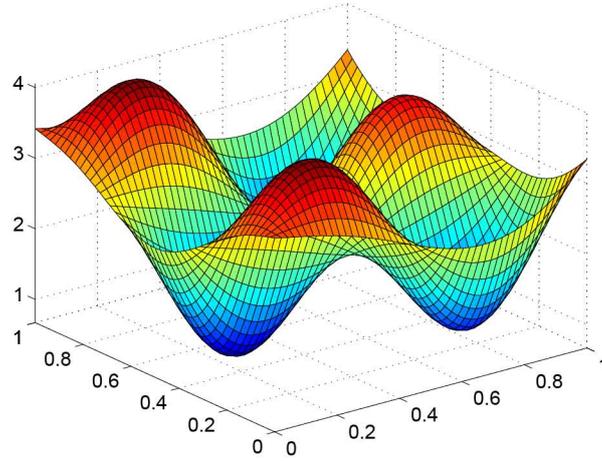
$$\mathcal{D}(F) := \{x \in L^2(\Omega) : \|x - \tilde{x}\|_{L^2} \leq \kappa \text{ für ein } \tilde{x} \in L^2(\Omega) \text{ mit } \tilde{x} \geq 0 \text{ f. ü.}\}$$

sichergestellt. Das entsprechende Resultat ist beispielsweise in [13] zu finden. Die Fréchet Ableitung des Operators  $F$  und seine Adjungierte sind dann gegeben durch

$$\begin{aligned} F'(\mathbf{c})w &= -A(x)^{-1}(wF(c)), \\ F'(\mathbf{c})^*w &= -F(c)A(c)^{-1}w, \end{aligned}$$

wobei der Operator  $A(c) : H^2(\Omega) \cap H_0^1(\Omega) \rightarrow L^2(\Omega)$  durch

$$(2.52) \quad A(c)y = -\Delta y + cy$$

Abbildung 2.2: Exakte Lösung  $x^\dagger$ .

definiert ist (vgl. [13]). Ist  $F(x) \geq \theta > 0$  für alle  $x \in B_r(x_0)$  mit  $r \leq \kappa$ , so kann man zeigen (siehe [32] oder [73, Beispiel 7.3.10]), dass  $F$  die verallgemeinerte Version von Annahme 3 (siehe Bemerkung nach Annahme 3) erfüllt. Der Operator  $R_{x,\tilde{x}}$  ist dabei gegeben durch seine Adjungierte

$$R_{x,\tilde{x}}^* w = A(x) \left( \frac{F(\tilde{x})}{F(x)} A(\tilde{x})^{-1} w \right).$$

Insbesondere sind somit in einer genügend kleinen Kugel um  $x_0$  die Tangentialkegelbedingung Annahme 1 und natürlich auch Annahme 3 erfüllt. Ferner erfüllt  $F$  auch die Skalierungsannahme (2.13).

Die rechten Seiten

$$f(\xi, \eta) = 32(\xi(\xi - 1) + \eta(\eta - 1)) + x^\dagger(\xi, \eta)(16\xi(\xi - 1)\eta(\eta - 1) + 1)$$

und  $g$  von (2.51) wählen wir so, dass die zum Parameter

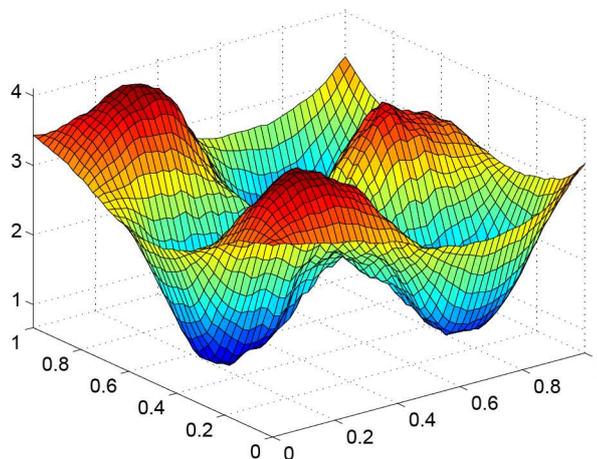
$$x^\dagger(\xi, \eta) = 1.5 \sin(2\pi\xi) \sin(3\pi\eta) + 3((\xi - 0.5)^2 + (\eta - 0.5)^2) + 2,$$

(siehe Abbildung 2.2) gehörende Lösung  $y$  von (2.51) durch

$$y(\xi, \eta) = 16\xi(\xi - 1)\eta(1 - \eta) + 1$$

gegeben ist. Der Startwert

$$x_0(\xi, \eta) = 3((\xi - 0.5)^2 + (\eta - 0.5)^2) + 2 + 128\xi(\xi - 1)\eta(1 - \eta)$$

Abbildung 2.3: Rekonstruktion  $u_{n*}$ .

erfüllt dann wegen  $\mathcal{R}(F'(x^\dagger)^*) = \mathcal{R}((F'(x^\dagger)^*F'(x^\dagger))^{1/2})$  und

$$(2.53) \quad (x^\dagger - x_0)/F(x^\dagger) \in H^2(\Omega) \cap H_0^1(\Omega)$$

die Glattheitsbedingung Annahme 2 mit  $\gamma = 1/2$ . Unsere numerischen Experimente sollten daher die in Satz 2.25 bewiesene Abklingrate  $O(\delta^{1/2})$  reflektieren.

Zur numerischen Umsetzung auf dem Rechner müssen wir zunächst noch eine Diskretisierung des Randwertproblems (2.51) durchführen. Dazu verwenden wir finite Differenzen mit Gitterweite  $\Delta\xi = \Delta\eta = \frac{1}{m+1}$  mit  $m \in \mathbb{N}$  in beiden Raumdimensionen. Wir verwenden die skalierte euklidische Norm

$$\|\cdot\|_{\Delta\xi} = \frac{1}{(\Delta\xi)^2} \|\cdot\|_2$$

und bezeichnen die Gitterpunkte mit  $(\xi_i, \eta_j) = (i\Delta\xi, j\Delta\eta)$ ,  $0 \leq i, j \leq m+1$ . Den Laplaceoperator  $-\Delta$  in (2.52) approximieren wir durch zentrale Differenzen 2. Ordnung

$$-\Delta \approx -\Delta_{\Delta\xi} = \frac{1}{(\Delta\xi)^2} \begin{bmatrix} 0 & -1 & 0 \\ -1 & 4 & -1 \\ 0 & -1 & 0 \end{bmatrix}.$$

Eine lexikographische Anordnung der Gitterpunkte ergibt nun das  $(m^2 \times m^2)$ -Gleichungssystem

$$(\mathbf{B} + \text{diag}(\mathbf{x}))\mathbf{y} = \mathbf{f}.$$

wobei  $\mathbf{B}$  die zu  $\Delta_{\Delta\xi}$  gehörende Blocktridiagonalmatrix ist und  $\mathbf{x}$ ,  $\mathbf{y}$  und  $\mathbf{f}$  die entsprechenden lexikographisch angeordneten diskretisierten Funktionen bezeichnen. Die gestörten Daten

$\mathbf{y}^\delta$  erzeugten wir durch

$$\mathbf{y}^\delta = \mathbf{y} + \delta \mathbf{z} / \|\mathbf{z}\|_{\Delta\xi},$$

wobei die Einträge von  $\mathbf{z}$  gleichverteilte Zufallsvariablen in  $[-1, 1]$  sind. In Abbildung 2.3 ist eine Rekonstruktion durch das exponentielle Euler-Verfahren dargestellt. Die Norm des Rauschens war  $\delta = 0.001$ , die Diskretisierungsschrittweite betrug  $\Delta\xi = 0.02$  und als Parameter des Verfahrens wurden  $\mu = 0.9$  und  $\tau = 1.5$  gewählt. Der Fehler der Regularisierung war dann, gemessen in der  $\Delta\xi$ -Norm, bei 3%. Für die weiteren Beispiele dieses Abschnitts verwenden wir eine Diskretisierung mit  $\Delta\xi = 0.01$ ,  $\tau = 1.5$  und benutzen die dynamische Steuerung der  $\mu_n$  (2.6) mit  $\mu_{\text{start}} = 0.1$  und  $\mu_{\text{max}} = 0.999$ . Zunächst haben wir die regularisierende Eigenschaft des exponentiellen Euler-Verfahrens untersucht. Abbildung 2.4 zeigt den relativen Rekonstruktionsfehler  $\|\mathbf{e}_{n_*}\|_{\Delta\xi} / \|\mathbf{x}^\dagger\|_{\Delta\xi}$  des exponentiellen Euler-Verfahrens für verschiedene Rauschpegel  $\delta$ . Das wegen  $\gamma = \frac{1}{2}$  von Satz 2.25 vorhergesagte Verhalten  $O(\sqrt{\delta})$  wird deutlich widergespiegelt.

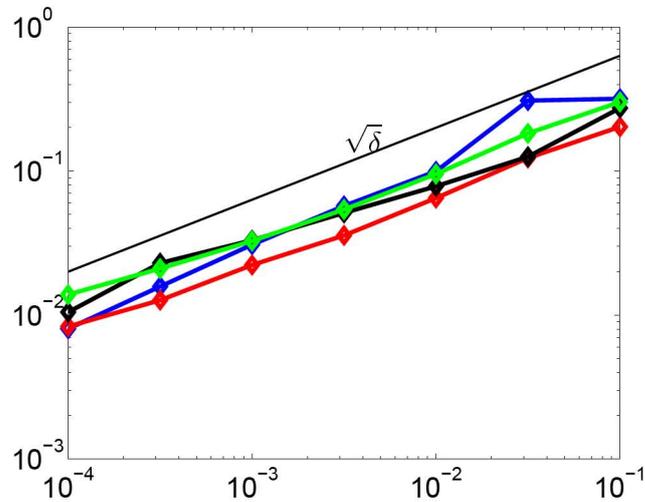


Abbildung 2.4: Relativer Rekonstruktionsfehler  $\|\mathbf{e}_{n_*}\|_{\Delta\xi} / \|\mathbf{x}^\dagger\|_{\Delta\xi}$  des exponentiellen Euler-Verfahrens (rot), Newton-cgnr (blau), (2.6) mit exponentiellem Euler (schwarz) und des iterativ regularisierten Gauss-Newton-Verfahrens (grün) jeweils als Funktion von  $\delta$ .

Zum Vergleich haben wir die Fehlerkurven eines regularisierten Newton-Verfahrens mit einem cgnr-Verfahren als innerer Iteration (siehe [31, 74]), des Verfahrens von Jin und Tautenhahn (siehe (2.6) und [48]) mit dem linearen exponentiellen Euler-Verfahren als innerer Regularisierung und des iterativ regularisierten Gauss-Newton-Verfahrens (siehe zum Beispiel [3, 7, 48]) graphisch dargestellt. Auch die drei Vergleichsverfahren spiegeln die ordnungsoptimale Konvergenzrate wider. Am Vergleich der Rechenzeit (siehe Abbildung

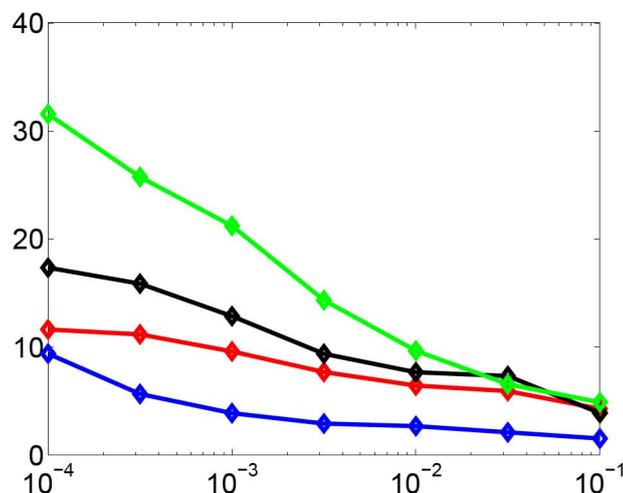


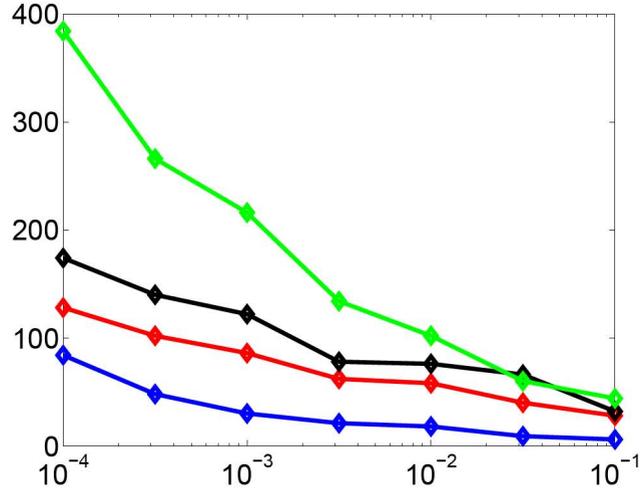
Abbildung 2.5: Rechenzeit des exponentiellen Euler-Verfahrens (rot), Newton-cgnr (blau), (2.6) mit exponentiellem Euler (schwarz) und des iterativ regularisierten Gauss-Newton-Verfahrens (grün) jeweils als Funktion von  $\delta$ .

2.5) erkennt man, dass das Newton-cgnr-Verfahren die anderen drei Verfahren schlägt. Das zweitschnellste ist das von uns untersuchte exponentielle Euler-Verfahren, welches für kleines  $\delta$  nur etwa 20% langsamer ist.

$\delta$	$10^{-1}$	$10^{-1.5}$	$10^{-2}$	$10^{-2.5}$	$10^{-3}$	$10^{-3.5}$	$10^{-4}$
Newton-cgnr	3	4	4	4	5	6	8
expo. Euler	3	4	4	4	4	5	5
expo. Euler (Jin)	3	4	4	4	5	5	5
Tikhonov (Jin)	3	3	4	4	5	5	6

Tabelle 2.1: Anzahl der äußeren Iterationen

Betrachtet man die Effizienz der Verfahren genauer, so erkennt man an Tabelle 2.1, dass die drei langsamen Verfahren im Wesentlichen die gleiche Anzahl an äußeren Iterationen benötigen. Beim Newton-cgnr-Verfahren ist dagegen der Anstieg der Anzahl der äußeren Iterationen für kleines  $\delta$  auffallend. Betrachtet man die Gesamtzahl der Matrix-Vektor-Produkte (siehe Abbildung 2.6), so erkennt man, dass sich die Kurven für das Newton-cgnr-Verfahren und das exponentielle Euler-Verfahren für  $\delta$  gegen Null leicht aufeinander zulaufen. Zusammen mit dem Verhalten der äußeren Iterationen ergibt sich, wie schon anhand des Vergleichs der Rechenzeit erkannt, dass für kleines  $\delta$  das in dieser Arbeit vorgestellte Verfahren mit dem Newton-cgnr-Verfahren mithalten kann.



Abbildungung 2.6: Anzahl der Matrix-Vektor-Produkte des exponentiellen Euler-Verfahrens (rot), Newton-cgnr (blau), (2.6) mit exponentiellem Euler (schwarz) und des iterativ regulierten Gauss-Newton-Verfahrens (grün) jeweils als Funktion von  $\delta$ .

Im Folgenden wollen wir noch das Verhalten der Regularisierungsverfahren bei verschiedenen glatten Anfangswerten untersuchen. Dazu seien

$$\tilde{x}_0(\xi, \eta) = 3((\xi - 0.5)^2 + (\eta - 0.5)^2) + 2 + 15\tilde{z}(\xi)\tilde{z}(\eta)$$

mit der Hutfunktion auf dem Intervall  $[0, 1]$

$$\tilde{z}(\xi, \eta) = \begin{cases} t, & 0 \leq t \leq 1/2, \\ 1 - t, & 1/2 < t \leq 1 \end{cases}$$

und

$$\hat{x}_0(\xi, \eta) = 3((\xi - 0.5)^2 + (\eta - 0.5)^2) + 2 + \sqrt{90}\hat{z}(\xi)\hat{z}(\eta)$$

mit

$$\hat{z}(\xi, \eta) = \begin{cases} t, & 0 \leq t \leq 1/2, \\ 2(1 - t), & 1/2 < t \leq 1. \end{cases}$$

Die drei Startwerte sind in Abbildung 2.8 dargestellt.

Wegen  $\tilde{x}_0 \in H^s(\Omega)$  und  $\hat{x}_0 \in H^{s-1}(\Omega)$  für jedes  $s < 3/2$  sind beide Anfangswerte nicht so glatt wie  $x_0$ . Es erfüllt auch keiner der beiden die Quellbedingung (2.53). Es gilt jedoch weiterhin  $(\tilde{x}_0 - x^\dagger)|_{\partial\Omega} = 0$  bzw.  $(\hat{x}_0 - x^\dagger)|_{\partial\Omega} = 0$ , so dass die Randterme der zugehörigen

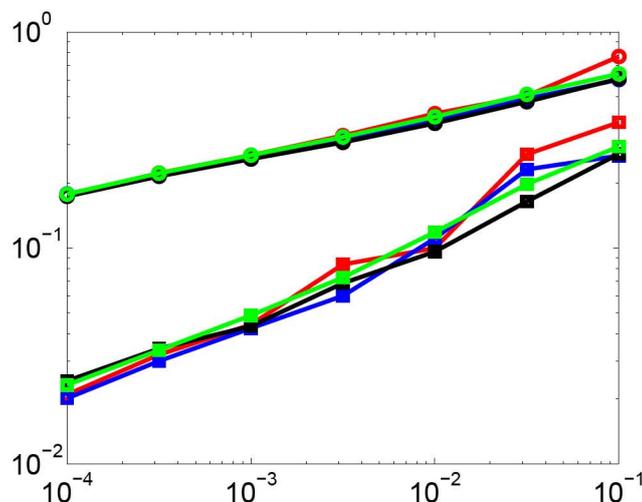


Abbildung 2.7: Relativer Rekonstruktionsfehler  $\|e_{n_*}\|_{\Delta\xi} / \|\mathbf{x}^\dagger\|_{\Delta\xi}$  von exponentiellem Euler (blau), Newton-cgmr (rot), (2.6) mit exponentiellem Euler (schwarz), und iterativ reguliertem Gauss-Newton-Verfahren (grün) bei verschiedenen glatten Anfangsdaten  $\tilde{x}_0$  ( $\circ$ ) und  $\hat{x}_0$  ( $\square$ ) jeweils als Funktion von  $\delta$ .

Lösungen passen. Aufgrund der geringeren Glattheit ist zu erwarten, dass der Rekonstruktionsfehler mit einer geringeren Rate, also  $O(\delta^\kappa)$  mit  $\kappa < \frac{1}{2}$ , fällt. Wie man an Abbildung 2.7 erkennt, spiegeln die numerischen Ergebnisse dieses Verhalten wider und zwar unabhängig vom gewählten Regularisierungsverfahren. Die zum Startwert  $\tilde{x}_0$  gehörenden Kurven fallen deutlich schneller als die zu  $\hat{x}_0$  gehörenden. Ein Vergleich mit Abbildung 2.4 zeigt auch, dass der die Quellbedingung (2.53) erfüllende Startwert  $x_0$  die schnellste Konvergenz liefert.

Zuletzt wollen wir noch den Einfluss der dynamischen Schrittweitensteuerung untersuchen. Dazu starten wir das Verfahren mit konstanter Wahl von  $\mu = 0.7$  und der flexiblen Wahl  $\mu_n$  nach (2.6) mit den Parametern  $\mu_{\text{start}} = 0.1$  und  $\mu_{\text{max}} = 0.999$ . Ferner setzen wir den Parameter des Diskrepanzprinzips  $\tau = 2$ . Die in Abbildung 2.9 dargestellten Ergebnisse zeigen, dass die adaptive Wahl der  $\mu_n$  die Effektivität des Verfahrens deutlich erhöhen.

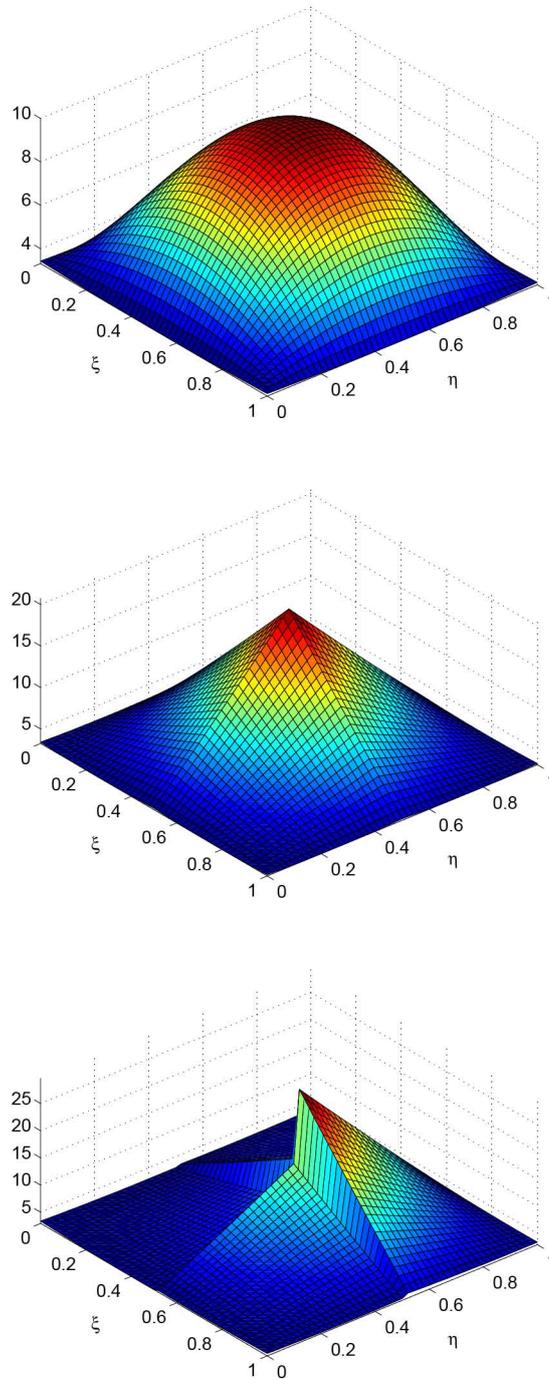


Abbildung 2.8: Startwerte  $x_0$  (oben),  $\tilde{x}_0$  (mitte) und  $\hat{x}_0$  (unten).

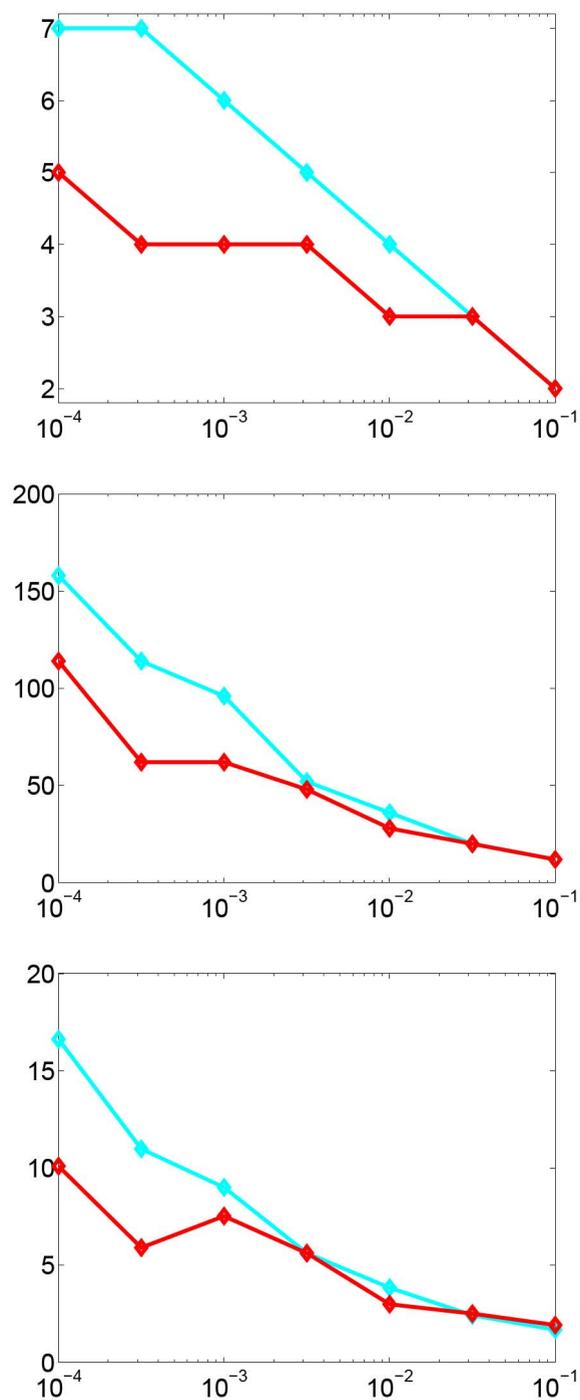


Abbildung 2.9: Anzahl der äußeren Iterationen (oben), Anzahl der Matrix-Vektor-Produkte und Rechenzeit des exponentiellen Euler-Verfahrens mit (rot) und ohne (cyan) adaptive Wahl der  $\mu_n$  jeweils als Funktion von  $\delta$ .

### 2.7.2 Ein inverses Problem aus der Grundwasserhydrologie

Das zweite numerische Beispiel beschreibt eine vereinfachte Form eines inversen Problems aus der Grundwasserhydrologie (siehe [12, 30]). Durch das Randwertproblem

$$(2.54) \quad \begin{aligned} -\operatorname{div}(x \operatorname{grad} y) &= f && \text{in } \Omega \\ y &= g && \text{auf } \partial\Omega \end{aligned}$$

wird eine quadratische ( $\Omega = [0, 6]^2$ ) grundwasserführende Schicht simuliert. Die gemischten Dirichlet-Neumann-Randwerte

$$y(\xi, 0) = 100, \quad y_\xi(6, \eta) = 0, \quad (xy_\xi)(0, \eta) = -500, \quad y_\eta(\xi, 6) = 0$$

beschreiben beispielsweise eine Verwerfung (kein Wasserfluss, Nullrandbedingungen) oder einen Fluss (Wasserzufluss und Wasserabfluss). Die rechte Seite

$$f(\xi, \eta) = \begin{cases} 0, & 0 < \eta < 4, \\ 137, & 4 < \eta < 5, \\ 274, & 5 < \eta < 6 \end{cases}$$

reflektiert die unterschiedliche Grundwasseranreicherung beispielsweise durch Vegetation oder Niederschlag.

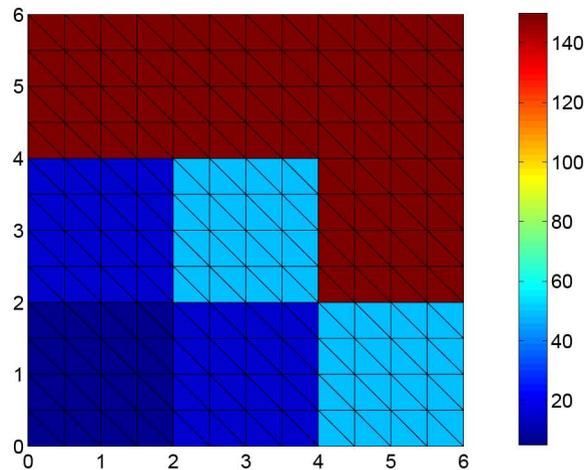


Abbildung 2.10: Exakter Koeffizient  $x^\dagger$  bei einer Diskretisierung mit 169 Ecken und 288 Dreiecken.

Die Lösung  $y$  des direkten Problems (2.54) gibt den Grundwasserspiegel an, der gesuchte Parameter des Randwertproblems  $x^\dagger$  beschreibt die Durchlässigkeit der grundwasserführenden Schicht. Wir wählen den Parameter  $x^\dagger$  als eine stückweise konstante Funktion, wie in Abbildung 2.10 dargestellt. Die Variation der Durchlässigkeit in  $x^\dagger$  könnte dann die Verdichtung des Sediments simulieren. Weitere Details zur Plausibilität dieses synthetischen Modells findet man in [12, Teil 1].

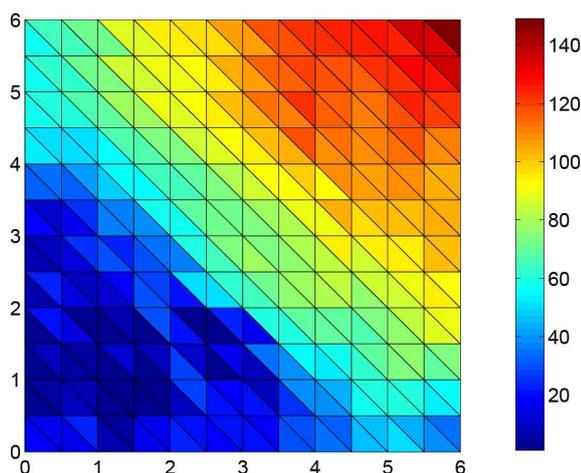


Abbildung 2.11: Rekonstruktion  $u_{n*}$  durch das exponentielle Euler-Verfahren.

Zur Diskretisierung des Problems setzen wir stückweise konstante finite Elemente ein. Eine Triangulierung mit 288 Dreiecken und 169 Gitterpunkten ist in Abbildung 2.10 und 2.11 dargestellt.

Obwohl dieses Problem die zur theoretischen Untersuchung in Kapitel 2 gemachten Voraussetzungen nicht erfüllt (Details dazu findet man bei Hanke [30]), zeigen die numerischen Experimente, dass das exponentielle Euler-Verfahren auch hier regularisierend wirkt. In Abbildung 2.11 ist die Rekonstruktion  $u_{n*}$  von  $x^\dagger$  mit um  $\delta = 10^{-3/2}$  gestörten Daten zu sehen. Der relative Regularisierungsfehler der Rekonstruktion beträgt 39%.

Für die weiteren numerischen Experimente in diesem Abschnitt haben wir allerdings eine Verfeinerung dieser Triangulierung mit 1152 Dreiecken und 625 Gitterpunkten verwendet. Als Vergleichsverfahren wählen wir wieder das regularisierte Newton-Verfahren mit `cgmr` als innerer Iteration.

In Abbildung 2.12 haben wir den relativen Fehler der beiden Verfahren bei unterschiedlich starken Störungen  $\delta$  dargestellt. Beide Verfahren zeigen dabei wieder ein ähnliches Verhalten. Einen Überblick über den Aufwand der untersuchten Verfahren geben die in Abbildung

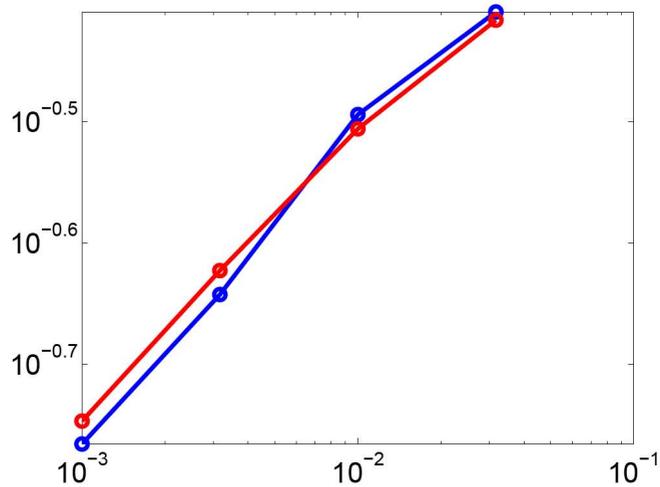


Abbildung 2.12: Relativer Regularisierungsfehler des exponentiellen Euler-Verfahrens (rot) und des Newton-cgnr-Verfahrens (blau) als Funktionen von  $\delta$ .

2.13 geplotteten Kurven. Wir erkennen, dass beide Verfahren mit etwa der gleichen Anzahl an Matrix-Vektor-Multiplikationen auskommen. Da sich die Anzahl der äußeren Iterationen aber um bis zu einem Faktor drei unterscheidet, arbeitet das exponentielle Euler-Verfahren letztendlich um bis zu einem Faktor zwei bis drei schneller als das Vergleichsverfahren.

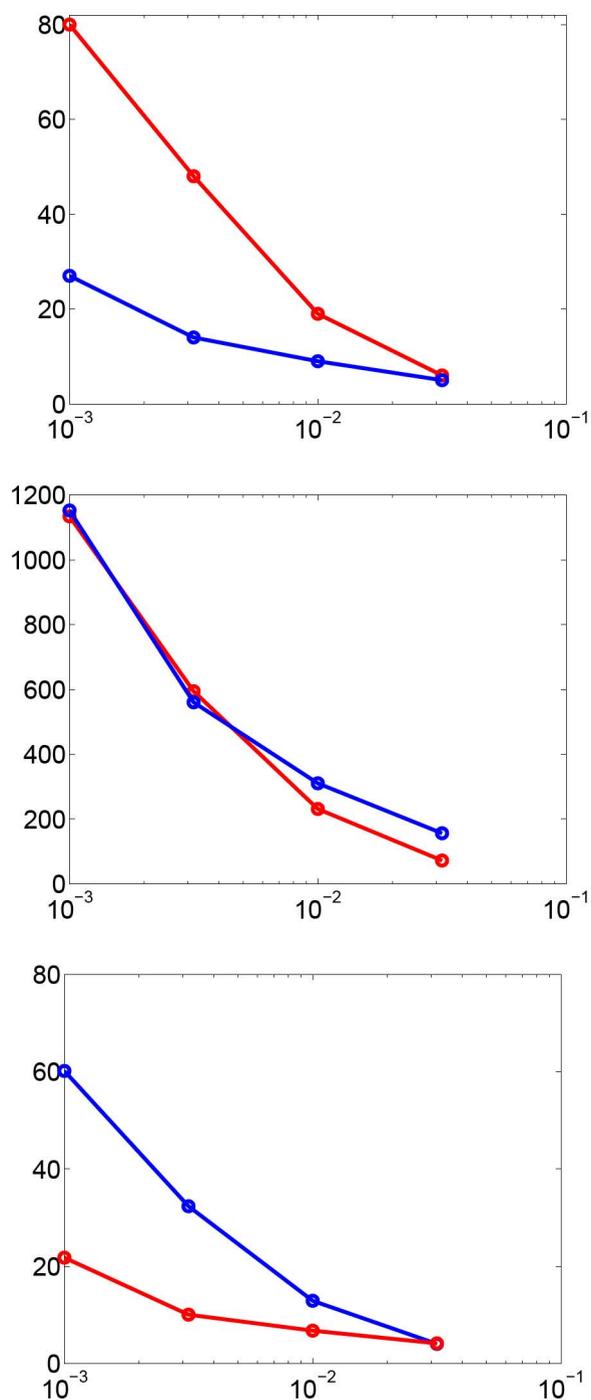


Abbildung 2.13: Anzahl der äußeren Iterationen (oben), Anzahl der Matrix-Vektor-Produkte und Rechenzeit des exponentiellen Euler-Verfahrens (rot) und des Newton-cgnr-Verfahrens (blau) jeweils als Funktion von  $\delta$ .



# FAZIT

Schlecht gestellte Probleme werden aktuell intensiv erforscht, unter anderem da sie in vielen Anwendungen auftreten. In dieser Arbeit wurde das exponentielle Euler-Verfahren zur Lösung linearer und erstmals auch nichtlinearer schlecht gestellter Probleme untersucht. Wir konnten zeigen, dass das Verfahren unter den in diesem Kontext üblichen Annahmen gegen die gewünschte Lösung des Problems konvergiert. Ferner haben wir bewiesen, dass die Konvergenz ordnungsoptimal ist. Die numerischen Beispiele reflektieren dieses Verhalten deutlich und legen dar, dass diese Methode konkurrenzfähig ist. Mit Blick auf das letzte numerische Beispiel scheint dies insbesondere bei Problemen, bei denen jede äußere Iteration teuer ist, der Fall zu sein.

Wie so oft wirft auch diese wissenschaftliche Arbeit mehr Fragen auf, als sie beantwortet. Im linearen Fall konnte Rieder [75] die regularisierende Eigenschaft von Runge-Kutta-Verfahren nachweisen. Die Situation im nichtlinearen Fall ist hingegen wesentlich unklarer. So konnte in [8] nur unter starken Schrittweitenbeschränkungen die Konvergenz für vereinfachte Runge-Kutta-Verfahren nachgewiesen werden. Eine Aussage über Konvergenzraten fehlt bisher noch ganz. Es scheint möglich, die in dieser Arbeit benutzten Techniken zum Nachweis der Konvergenz auf andere Integratoren zu übertragen. Allerdings ist auch hier zu befürchten, dass dies je nach Integrator eine Schrittweitenbeschränkung nach sich ziehen würde. Die Techniken, welche wir zum Beweis der Konvergenzraten benutzt haben, auf andere Integratoren anzuwenden scheint sich dagegen als schwierig zu gestalten, da sie sich wesentlich auf die Teleskopsumme (2.38) stützt. Da der Einsatz anderer Integratoren zu einer anderen  $\varphi$ -Funktion führt würde sich diese Summe nicht so einfach auflösen lassen und es ist unklar, wie man dann fortfahren könnte.

Die numerischen Beispiele lassen uns ferner vermuten, dass die Schrittweitenbeschränkung (2.30) zu scharf ist. Mit adaptiver Wahl der  $\mu_n$  sind die Schrittweitzuwächse der letzten Schritte oft zu groß, um dieses Kriterium erfüllen zu können. Es wäre daher interessant zu untersuchen, ob man die Ordnungsoptimalität auch unter einer schwächeren Schrittweitenbeschränkung zeigen kann.

Das letzte numerische Beispiel zeigte uns ferner, dass das exponentielle Euler-Verfahren auch zur Regularisierung schlecht gestellter Probleme, die die Annahmen 1 bis 3 nicht

erfüllen eingesetzt werden kann. Kann man auch in diesen Fall noch eine regularisierende Eigenschaft nachweisen? Ein erster Schritt wäre hier, Annahme 2 auf allgemeinere Quellbedingungen zu erweitern. Insbesondere sind dabei logarithmische Quellbedingungen, wie sie bei exponentiell schlecht gestellten Problemen auftreten, zu untersuchen.

Letztendlich ist die interessanteste offen geblieben Frage, ob man diese Arbeit als Beitrag zur *angewandten* Mathematik bezeichnen darf. Ist es möglich, das hier vorgestellte und untersuchte Verfahren in den vielfältigen schlecht gestellten Problemen der Natur- und Ingenieurwissenschaften einzusetzen? Die theoretischen Erkenntnisse und numerischen Ergebnisse lassen darauf hoffen.

Der Ball rollt!

# ANHANG A

## EXPONENTIELLE INTEGRATOREN

Der in dieser Arbeit verwendete Integrator gehört zur Klasse der exponentiellen Integratoren. Obwohl exponentielle Integratoren schon länger bekannt sind, werden sie erst in den letzten Jahren intensiv untersucht (siehe z.B. [11, 14, 40, 41, 42, 67]) und zur Lösung großer Systeme steifer und oszillatorischer Differentialgleichungen (siehe z.B. [6, 5, 78])

$$(A.1) \quad u'(t) = f(u(t)), \quad u(t_0) = 0$$

eingesetzt. Das in dieser Arbeit eingesetzte numerische Verfahren basiert auf einer Linearisierung der autonomen Gleichung (A.1). Ausgehend von einer Näherung  $u_n \approx u(t_n)$  ist diese Linearisierung gegeben durch

$$(A.2) \quad u'(t) = J_n u(t) + g_n(u(t))$$

mit

$$J_n = \frac{df}{du}(u_n), \quad g_n(u(t)) = f(u(t)) - J_n u(t).$$

**Bemerkung.** In Kapitel 1 liegt eine lineare Differentialgleichung vor. In diesem Fall sind  $J_n = A^* A$  und  $g_n = A^* y^\delta$  unabhängig von  $n$ . In Kapitel 2 haben wir eine vereinfachte Version des exponentiellen Integrators eingesetzt. Dort verwendeten wir statt der exakten Jacobi-Matrix nur die Approximation durch  $F'(u_n)^* F'(u_n)$ .  $\diamond$

Die Anwendung eines  $s$ -stufigen expliziten exponentiellen Runge-Kutta-Verfahrens auf (A.2) führt auf die Klasse der exponentiellen Rosenbrock-Verfahren

$$U_{ni} = e^{c_i h_n J_n} u_n + h_n \sum_{j=1}^{i-1} a_{ij}(h_n J_n) g_n(U_{nj}), \quad 1 \leq i \leq s,$$
$$u_{n+1} = e^{h_n J_n} u_n + h_n \sum_{j=i}^s b_i(h_n J_n) g_n(U_{ni}),$$

wobei  $h_n$  die Zeitschrittweite und  $u_{n+1}$  die numerische Approximation an die kontinuierliche Lösung  $u(t_{n+1})$  zur Zeit  $t_{n+1} = t_n + h_n$  ist (siehe [42]). Die Gewichte  $b_i(z)$  sind dabei Linearkombinationen der ganzen Funktionen

$$\varphi_k(z) = \int_0^1 e^{(1-s)z} \frac{s^{k-1}}{(k-1)!} ds, \quad k \geq 1.$$

Sie erfüllen die Rekursion

$$\varphi_k(z) = \frac{\varphi_{k-1}(z) - \varphi_{k-1}(0)}{z}, \quad \varphi_0(z) = e^z.$$

Das exponentielle Euler-Verfahren ist ein einstufiges ( $s = 1$ ) Verfahren und hat die Koeffizienten

$$b_1(z) = \varphi_1(z) \quad \text{und} \quad c_1 = 0.$$

Da somit nur die Funktion  $\varphi_1$  auftritt, lassen wir im Rest dieser Arbeit den Index weg und schreiben nur  $\varphi$ .

# ANHANG B

## APPROXIMATION VON MATRIXFUNKTIONEN

Zur numerischen Umsetzung exponentieller Integratoren, müssen wir Produkte von Funktionen von Matrizen und Vektoren berechnen. In dieser Arbeit approximieren wir für eine (symmetrische, negativ semidefinite) Matrix  $\mathbf{A} \in \mathbb{R}^{n,n}$  eine Zeitschrittweite  $h > 0$  und einen Vektor  $\mathbf{b} \in \mathbb{R}^n$  das Produkt

$$(B.1) \quad \varphi(\mathbf{A})\mathbf{b} \quad \text{mit} \quad \varphi(z) = \frac{e^z - 1}{z}$$

mittels Krylov-Methoden. Dabei projizieren wir das jeweils auf den  $m$ -ten Krylov-Raum

$$\mathcal{K}_m(\mathbf{A}, \mathbf{b}) = \text{span} \{ \mathbf{b}, \mathbf{A}\mathbf{b}, \dots, \mathbf{A}^{m-1}\mathbf{b} \}.$$

Verfahren dieser Art wurden in [15, 16, 17, 38, 52, 53, 65, 76] für die mit  $\varphi$  verwandte Exponentialfunktion untersucht.

In den numerischen Beispielen dieser Arbeit haben wir sowohl das Arnoldi-Verfahren [1] als auch das Lanczos-Verfahren (mit Reorthogonalisierung, [54]) verwendet. Diese liefern im  $m$ -ten Schritt eine orthogonale Matrix  $\mathbf{V}_m = [\mathbf{v}_1, \dots, \mathbf{v}_m]$  und eine obere Hessenberg- bzw. Tridiagonalmatrix  $\mathbf{H}_m$  so, dass

$$\mathbf{A}\mathbf{V}_m = \mathbf{V}_m\mathbf{H}_m + h_{m+1,m}\mathbf{v}_{m+1}\mathbf{e}_1^*$$

gilt. Die Approximation von (B.1) erhalten wir dann mittels

$$\varphi(\mathbf{A})\mathbf{b} \approx \mathbf{V}_m\varphi(\mathbf{H}_m)\beta\mathbf{e}_1.$$

Da typischerweise  $m \ll n$  gilt, siehe Hochbruck und Lubich [38] sowie Gallopoulos und Saad [22], kann man  $\varphi(\mathbf{H}_m)$  schnell durch ein direktes Verfahren, zum Beispiel Diagonalisierung, berechnen [61, 62].

**Algorithmus 6** Arnoldi-Algorithmus

---

Zu gegebenen  $\mathbf{A} \in \mathbb{R}^{n,n}$ ,  $\mathbf{b} \in \mathbb{R}^n$ , berechne  $\beta = \|\mathbf{b}\| > 0$   
 $\mathbf{v}_1 = \mathbf{b}/\beta$ ;  
**for**  $m = 1, 2, \dots$  **do**  
  **for**  $j = 1, \dots, m$  **do**  
     $h_{j,m} = \mathbf{v}_j^T \mathbf{A} \mathbf{v}_m$   
  **end for**  
   $\mathbf{v}_{m+1} = \mathbf{A} \mathbf{v}_m - \sum_{j=1}^m h_{j,m} \mathbf{v}_j$   
   $h_{m+1,m} = \|\mathbf{v}_{m+1}\|$   
   $\mathbf{v}_{m+1} = \mathbf{v}_{m+1}/h_{m+1,m}$   
**end for**

---

**Algorithmus 7** Lanczos-Algorithmus

---

Zu gegebenen  $\mathbf{A} \in \mathbb{R}^{n,n}$ ,  $\mathbf{b} \in \mathbb{R}^n$ , berechne  $\beta = \gamma_0 = \|\mathbf{b}\| > 0$   
 $\mathbf{r}_0 = \mathbf{b}$   
 $\mathbf{v}_1 = \mathbf{r}_0/\gamma_0$ ;  
**for**  $m = 1, 2, \dots$  **do**  
   $\mathbf{r}_m = \mathbf{A} \mathbf{v}_m - \gamma_{m-1} \mathbf{v}_{m-1}$   
   $\alpha_m = \tilde{\mathbf{v}}_m^T \mathbf{r}_m$   
   $\mathbf{r}_m = \mathbf{r}_m - \alpha_m \mathbf{q}_m$   
   $\gamma_m = \|\mathbf{r}_m\|$   
   $\mathbf{v}_{m+1} = \mathbf{r}_m/\gamma_{m+1}$ ;  
**end for**  
 $\mathbf{H} = \text{diag}([\gamma_1, \dots, \gamma_{m-1}], [\alpha_1, \dots, \alpha_m], [\gamma_1, \dots, \gamma_{m-1}])$

---

Wie in [16, 17, 38] gezeigt wird konvergiert der Fehler dieser Approximation für symmetrische, negativ definite Matrizen superlinear, sobald  $m > \sqrt{\|\mathbf{A}\|_2}$  ist. Im zweiten Teil unserer Arbeit ist die Matrix  $\mathbf{A}$  eine Diskretisierung von  $-h_n J_n$ . Da  $\|J_n\| \leq 1$  ist, beginnt die superlineare Konvergenz also bei  $\sqrt{h_n}$ .

# LITERATURVERZEICHNIS

- [1] W. E. ARNOLDI, *The principle of minimized iterations in the solution of the matrix eigenvalue problem*, Quart. Appl. Math., 9 (1951), Seiten 17–29.
- [2] A. BAKUSHINSKII UND M. KOKURIN, *Iterative Methods for Approximate Solution of Inverse Problems*, Springer, Dordrecht, 2004.
- [3] A. B. BAKUSHINSKII, *The problem of convergence of the iteratively regularized Gauss–newton method*, Comp. Math. Math. Phys., 32 (1992), Seiten 1353–1359.
- [4] F. BAUER UND T. HOHAGE, *A Lepskij-type stopping rule for regularized Newton methods*, Inverse Problems, 21 (2005), Seiten 1975–1991.
- [5] L. BERGAMASCHI, M. CALIARI, UND M. VIANELLO, *The ReLPM Exponential Integrator for FE Discretizations of Advection-Diffusion Equations*, Lecture Notes in Computer Science, 176 (2004), Seiten 434–442.
- [6] H. BERLAND, B. OWREN, UND B. SKAFLESTAD, *Solving the nonlinear Schrödinger equation using exponential integrators*, Modeling, Identification and Control, 27 (2006), Seiten 201–217.
- [7] B. BLASCHKE, A. NEUBAUER, UND O. SCHERZER, *On convergence rates for the iteratively regularized Gauss-Newton method*, IMA J. Num. Anal., 17 (1997), Seiten 421–436.
- [8] C. BÖCKMANN UND P. PORNSAWAD, *Iterative Runge–Kutta-type methods for nonlinear ill-posed problems*, Inverse Problems, 24 (2008), Artikel 025002.
- [9] D. CALVETTI UND L. REICHEL, *Lanczos-based exponential filtering for discrete ill-posed problems*, Numer. Alg., 29 (2002), Seiten 45–65.
- [10] D. CALVETTI, L. REICHEL, UND Q. ZHANG, *Iterative exponential filtering for large discrete ill-posed problems*, Numer. Math., 83 (1999), Seiten 535–556.

- 
- [11] M. P. CALVO UND C. PALENCIA, *A class of explicit multistep exponential integrators for semilinear problems*, Numer. Math., 102 (2005), Seiten 367–381.
- [12] J. CARRERA UND S. P. NEUMAN, *Estimation of aquifer parameters under transient and steady state conditions: 1-3*, Water Resour. Res., 22 (1986), Seiten 199–242.
- [13] F. COLONIUS UND K. KUNISCH, *Stability for parameter estimation in two point boundary volume problems*, J. Reine Angew. Math., 370 (1986), Seiten 1–29.
- [14] S. M. COX UND P. C. MATTHEWS, *Exponential time differencing for stiff systems*, J. Comp. Phys., 176 (2002), Seiten 430–455.
- [15] V. DRUSKIN, A. GREENBAUM, UND L. KNIZHNERMAN, *Using nonorthogonal Lanczos vectors in the computation of matrix functions*, SIAM J. Sci. Comp., 19 (1998), Seiten 38–54.
- [16] V. L. DRUSKIN UND L. A. KNIZHNERMAN, *Two polynomial methods of calculating functions of symmetric matrices*, U.S.S.R Comp. Math. Math. Phys., 29 (1989), Seiten 112–121.
- [17] ———, *Krylov subspace approximations of eigenpairs and matrix functions in exact and computer arithmetic*, Numer. Lin. Alg. Appl., 2 (1995), Seiten 205–217.
- [18] S. EISENSTAT UND H. WALKER, *Choosing the forcing terms in an inexact Newton method*, SIAM J. Sci. Comp., 17 (1996), Seiten 16–32.
- [19] H. W. ENGL, *Necessary and sufficient conditions for convergence of regularization methods for solving linear operator equations of the first kind*, Numer. Func. Anal. Optimiz., 3 (1981), Seiten 201–222.
- [20] H. W. ENGL, M. HANKE, UND A. NEUBAUER, *Regularization of Inverse Problems*, Kluwer Academic Publishers, Dordrecht, Boston, London, 1996.
- [21] H. W. ENGL, K. KUNISCH, UND A. NEUBAUER, *Convergence rates for Tikhonov regularization of nonlinear ill-posed problems*, Inverse Problems, 5 (1989), Seiten 523–540.
- [22] E. GALLOPOULOS UND Y. SAAD, *Efficient solution of parabolic equations by Krylov approximation methods*, SIAM J. Sci. Statist. Comp., 13 (1992), Seiten 1236–1264.
- [23] H. GFRERER, *An a posteriori parameter choice for ordinary and iterated Tikhonov regularization of ill-posed problems leading to optimal convergence rates*, Math. Comp., 49 (1987), Seiten 507–522.

- 
- [24] V. GLASKO UND A. TIKHONOV, *Use of the regularization method in non-linear problems*, USSR Comp. Math. Math. Phys., 5 (1965), Seiten 93–107.
- [25] C. W. GROETSCH, *Inverse Problems in the Mathematical Sciences*, Vieweg, Braunschweig, Wiesbaden, 1993.
- [26] J. HADAMARD, *Lectures on the Cauchy Problem in Linear Partial Differential Equations*, Dover Publ., New York, 1923.
- [27] E. HAIRER UND G. WANNER, *Solving ordinary differential equations II. Stiff and Differential-Algebraic problems*, Springer, Berlin, Heidelberg, New York, 1996.
- [28] U. HÄMARIK UND U. TAUTENHAHN, *The use of monotonicity for choosing the regularization parameter in ill-posed problems*, Inverse Problems, 15 (1999), Seiten 1487–1505.
- [29] M. HANKE, *Conjugate gradient type methods for ill-posed problems*, Longman Scientific & Technical, Harlow, Essex, 1995.
- [30] ———, *A regularizing Levenberg–Marquardt scheme, with applications to inverse groundwater filtration problems*, Inverse Problems, 13 (1997), Seiten 79–95.
- [31] ———, *Regularizing properties of a truncated newton-cg algorithm for nonlinear inverse problems*, Numer. Funct. Anal. Optim., 18 (1997), Seiten 971–993.
- [32] M. HANKE, A. NEUBAUER, UND O. SCHERZER, *A convergence analysis of the Landweber iteration for nonlinear ill-posed problems*, Num. Math., 72 (1995), Seiten 21–37.
- [33] P. HANSEN, *Analysis of discrete ill-posed problems by means of the L-curve*, SIAM Rev., 34 (1992), Seiten 561–580.
- [34] ———, *Regularization tools version 4.0 for MATLAB 7.3*, Numer. Alg., 46 (2007), Seiten 189–194.
- [35] D. HENRY, *Geometric Theory of Semilinear Parabolic Equations*, Springer, Berlin, Heidelberg, New York, 1981.
- [36] M. HOCHBRUCK, M. HÖNIG, UND A. OSTERMANN, *A convergence analysis of the exponential Euler iteration for nonlinear ill-posed problems*, Inverse Problems, (eingereicht).
- [37] ———, *Regularization of nonlinear ill-posed problems by exponential integrators*, Math. Mod. Numer. Anal., (eingereicht).
- [38] M. HOCHBRUCK UND C. LUBICH, *On Krylov subspace approximations to the matrix exponential operator*, SIAM J. Numer. Anal., 34 (1997), Seiten 1911–1925.

- 
- [39] M. HOCHBRUCK, C. LUBICH, UND H. SELHOFER, *Exponential integrators for large systems of differential equations*, SIAM J. Sci. Comp., 19 (1998), Seiten 1552–1574.
- [40] M. HOCHBRUCK UND A. OSTERMANN, *Explicit exponential runge–kutta methods for semilinear parabolic problems*, SIAM J. Numer. Anal., 43 (2005), Seiten 1069–1090.
- [41] ———, *Exponential Runge–Kutta methods for parabolic problems*, Appl. Numer. Math., 53 (2005), Seiten 323–339.
- [42] M. HOCHBRUCK, A. OSTERMANN, UND J. SCHWEITZER, *Exponential Rosenbrock-type methods*, SIAM J. Numer. Anal., (noch nicht erschienen).
- [43] B. HOFMANN, *Mathematik inverser Probleme*, B.G. Teubner, Stuttgart, Leipzig, 1999.
- [44] B. HOFMANN UND P. MATHÉ, *How general are general source conditions?*, Inverse Problems, 24 (2008), Artikel 015009.
- [45] B. HOFMANN UND O. SCHERZER, *Local ill-posedness and source conditions of operator equation in Hilbert spaces*, Inverse Problems, 14 (1998), Seiten 1189–1206.
- [46] T. HOHAGE UND S. LANGER, *Convergence analysis of an inexact iteratively regularized Gauss–Newton method under general source conditions*, Journal of Inverse and Ill-Posed Problems, 15 (2007), Seiten 19–35.
- [47] M. HÖNIG, *Asymptotische Regularisierung schlecht gestellter Probleme mittels steifer Integratoren*, Diplomarbeit, Universität Karlsruhe, 2004.
- [48] Q. JIN UND U. TAUTENHAHN, *On the discrepancy principles for some Newton type methods for solving nonlinear inverse problems*, Numer. Math., (noch nicht erschienen).
- [49] B. KALTENBACHER, *Some newton-type methods for the regularization of nonlinear ill-posed problems*, Inverse Problems, 13 (1997), Seiten 792–753.
- [50] ———, *A posteriori parameter choice strategies for some Newton type methods for the regularization of nonlinear ill-posed problems*, Numer. Math., 79 (1998), Seiten 501–528.
- [51] B. KALTENBACHER, A. NEUBAUER, UND O. SCHERZER, *Iterative Regularization Methods for Nonlinear Ill-Posed Problems*, de Gruyter, Berlin, New York, 2008.
- [52] L. A. KNIZHNERMAN, *Calculation of functions of unsymmetric matrices using Arnoldi’s method*, Comp. Math. Math. Phys., 31 (1992), Seiten 1–9.
- [53] ———, *Error bounds in Arnoldi’s method: The case of a normal matrix*, Comp. Math. Math. Phys., 32 (1992), Seiten 1199–1211.

- 
- [54] C. LANCZOS, *An iteration method for the solution of the eigenvalue problem of linear differential and integral operators*, Journal of Research of the National Bureau of Standards, 45 (1950), Seiten 255–282.
- [55] L. LANDWEBER, *An iteration formula for Fredholm integral equations of the first kind*, Amer. J. Math, 73 (1951), Seiten 614–624.
- [56] J. D. LAWSON, *Generalized Runge–Kutta processes for stable systems with large Lipschitz constants*, SIAM J. Numer. Anal., 4 (1967), Seiten 372–380.
- [57] L. LI, B. HAN, UND W. WANG, *R-K type Landweber method for nonlinear ill-posed problems*, J. Comp. Appl. Math., 206 (2007), Seiten 341–357.
- [58] A. K. LOUIS, *Inverse und schlecht gestellte Probleme*, B.G. Teubner, Stuttgart, 1989.
- [59] P. MATHÉ UND S. PEREVERZEV, *Geometry of linear ill-posed problems in Hilbert scales*, Inverse Problems, 19 (2003), Seiten 789–803.
- [60] G. MEURANT UND Z. STRAKOŠ, *The Lanczos and conjugate gradient algorithms in finite precision arithmetic*, Acta Numerica, 15 (2006), Seiten 471–542.
- [61] C. MOLER UND C. V. LOAN, *Nineteen dubious ways to compute the exponential of a matrix, twenty-five years later*, SIAM Review, 45 (2003), Seiten 3–49.
- [62] C. MOLER UND C. VAN LOAN, *Nineteen dubious ways to compute the exponential of a matrix*, SIAM Review, 20 (1978), Seiten 801–836.
- [63] V. A. MOROZOV, *The error principle in the solution of operational equations by the regularization method*, USSR Comp. Math. Math. Phys., 8 (1968), Seiten 63–87.
- [64] A. NEUBAUER, *Tikhonov regularization for non-linear ill-posed problems: optimal convergence rates and finite-dimensional approximation*, Inverse Problems, 5 (1989), Seiten 541–557.
- [65] J. NIEHOFF, *Projektionsverfahren zur Approximation von Matrixfunktionen mit Anwendungen auf die Implementierung exponentieller Integratoren*, Dissertatio, Heinrich-Heine-Universität Düsseldorf, 2006.
- [66] S. P. NØRSETT, *An A-stable modification of the Adams–Bashforth methods*, in Conference on the Numerical Solution of Differential Equations, J. Morris (Herausgeber), vol. 109 of Lecture Notes in Math., Academic Press, 1969, Seiten 214–219.
- [67] A. OSTERMANN, M. THALHAMMER, UND W. M. WRIGHT, *A class of explicit exponential general linear methods*, BIT, 46 (2006), Seiten 409–432.

- 
- [68] A. PAZY, *Semigroups of linear operators and applications to partial differential equations*, Springer, Berlin, Heidelberg, New York, 1983.
- [69] R. PLATO, *Optimal algorithms for linear ill-posed problems yield regularization methods*, Numer. Func. Anal. and Optimiz., 11 (1990), Seiten 111–118.
- [70] D. A. POPE, *An exponential method of numerical integration of ordinary differential equations*, Comm. ACM, 6 (1963), Seiten 491–493.
- [71] A. RIEDER, *On convergence rates of inexact Newton regularizations*, Inverse Problems, 15 (1999), Seiten 309–327.
- [72] —, *On the regularization of nonlinear ill-posed problems via inexact Newton iterations*, Inverse Problems, 15 (1999), Seiten 309–327.
- [73] —, *Keine Probleme mit inversen Problemen*, Vieweg, Wiesbaden, 2003.
- [74] —, *Inexact Newton regularization using conjugate gradients as inner iteration*, SIAM J. Numer. Anal., 43 (2005), Seiten 604–622.
- [75] —, *Runge–Kutta integrators yield optimal regularization schemes*, Inverse Problems, 21 (2005), Seiten 453–471.
- [76] Y. SAAD, *Analysis of some Krylov subspace approximations to the matrix exponential operator*, SIAM J. Numer. Anal., 29 (1992), Seiten 209–228.
- [77] E. SCHOCK, *Non-linear ill-posed equations: counter-examples*, Inverse Problems, 18 (2002), Seiten 715–717.
- [78] J. SCHWEITZER, *Numerical Simulation of Relativistic Laser-Plasma Interaction*, Dissertatio, Heinrich-Heine-Universität Düsseldorf, 2008.
- [79] D. SHOWALTER, *Representation and computation of the pseudoinverse*, Proceedings of the American Mathematical Society, 18 (1967), Seiten 584–586.
- [80] U. TAUTENHAHN, *On the asymptotical regularization of nonlinear ill-posed problems*, Inverse Problems, 10 (1994), Seiten 1405–1418.
- [81] —, *On a general regularization scheme for nonlinear ill-posed problems*, Inverse Problems, 13 (1997), Seiten 1427–1437.
- [82] —, *Optimality for ill-posed problems under general source conditions*, Numer. Funct. Anal. Optim., 19 (1998), Seiten 377–398.

- 
- [83] G. M. VAINIKKO, *On the optimality of regularization methods*, in Proceedings of the Alpine-U.S. Seminar on Inverse and Ill-Posed Problems, H. W. Engl und C. W. Groetsch (Herausgeber), Academic Press, 1986, Seiten 77–95.
- [84] J. VAN DEN ESHOF UND M. HOCHBRUCK, *Preconditioning lanczos approximations to the matrix exponential*, SIAM J. Sci. Comp., 27 (2006), Seiten 1438–1457.
- [85] C. VOGEL, *Computational methods for inverse problems*, SIAM, Philadelphia, 2002.
- [86] G. WAHBA, *Practical approximate solutions to linear operator equations when the data are noisy*, SIAM J. Numer. Anal., 14 (1977), Seiten 651–667.
- [87] D. WERNER, *Funktionalanalysis*, Springer, Berlin, Heidelberg, New York, 2000. 3. Auflage.



# DANKSAGUNG

Ohne die Unterstützung vieler lieber Menschen würde es die Arbeit nicht oder zumindest nicht so geben. Ich möchte mich daher an dieser Stelle bei ihnen allen herzlich bedanken.

Mein besonderer Dank gilt dabei meiner Doktormutter Prof. Dr. Marlis Hochbruck, die mir jederzeit in den vergangenen vier Jahren zur Seite stand. Durch die zahlreichen Gespräche mit und Anregungen von ihr, wurde diese Arbeit erst möglich gemacht.

Des Weiteren möchte ich mich bei Prof. Dr. Alexander Ostermann für die fruchtbare Zusammenarbeit der letzten Monate bedanken.

Bei Prof. Dr. Martin Hanke-Bourgeois bedanke ich mich für die Übernahme des Koreferats und die Hinweise, die zum Gelingen des zweiten Teils dieser Arbeit beigetragen haben.

Prof. Dr. Andreas Rieder danke ich dafür, dass er mich durch die Betreuung meiner Diplomarbeit auf dieses spannende Thema aufmerksam gemacht hat.

Meine aktuellen und besonders auch ehemaligen Kollegen der Lehrstühle für Angewandte Mathematik und Mathematische Optimierung haben dafür gesorgt, dass mir die Zeit an der Universität Düsseldorf immer in angenehmer Erinnerung bleiben wird. Hierfür herzlichen Dank.

Ich möchte mich bei all denen entschuldigen, die diese Arbeit beeinflusst haben, die ich aber vergessen habe namentlich zu erwähnen.

## Was wäre ich ohne meine Familie?

Ohne den Rückhalt und die Motivation durch meine lieben Eltern Rosemarie und Bruno, meine Geschwister Markus, Christina und Thomas und meine Freundin Veronika wäre diese Arbeit nicht entstanden.



# ERKLÄRUNG

Die hier vorgelegte Dissertation habe ich eigenständig und ohne unerlaubte Hilfe angefertigt. Die Dissertation wurde in der vorgelegten oder in ähnlicher Form noch bei keiner anderen Institution eingereicht. Ich habe bisher keine erfolglosen Promotionsversuche unternommen.

Düsseldorf, den 25.11.2008

(Michael Hönig)