KERNFORSCHUNGSZENTRUM KARLSRUHE

Institut für Meteorologie und Klimaforschung

KfK 5421

Inversion schlecht gestellter Probleme am Beispiel der bodengebundenen, passiven Millimeterwellenradiometrie

Martin Kuntz

Kernforschungszentrum Karlsruhe GmbH, Karlsruhe

Als Manuskript gedruckt
Für diesen Bericht behalten wir uns alle Rechte vor

Kernforschungszentrum Karlsruhe GmbH
Postfach 3640, 76021 Karlsruhe

ISSN 0303-4003

Zusammenfassung

Inversion schlecht gestellter Probleme am Beispiel der bodengebundenen, passiven Millimeterwellenradiometrie

Inverse Probleme werden kurz charakterisiert und typische Schwierigkeiten bei der Auswertung aufgenommener Daten aufgezeigt. Zur Lösung inverser Probleme geeignete mathematische Verfahren werden vorgestellt, ihre Stärken und Schwächen diskutiert und auf ein einfaches Problem aus dem Bereich der Fernerkundung mittels bodengebundener, passiver Millimeterwellenradiometrie angewandt. Als besonders geeignete Inversionsverfahren kristallisieren sich die abgeschnittene Singulärwertzerlegung, die Tikhonov–Phillips–Regularisierung, das konjugierte Gradienten–Verfahren sowie die stochastischen Verfahren heraus.

Stable solutions of inverse problems using ground based passive millimeterwave radiometry

Inverse problems will be briefly characterized and typical difficulties associated with the evaluation of measured data will be dealt with. Some mathematics appropriate to the solution of inverse problems will be presented and their advantages and disadvantages will be discussed. Following this, the techniques will be applied to a simple problem occuring in ground based passive millimeterwave radiometry. The truncated singular value decomposition, the Tikhonov–Phillips regularization, the conjugate gradient method, and the statistical methods will prove to be the most suitable inversion techniques.

Vorwort

Vielen mag das numerische Lösen von Integralgleichungen, gemeinhin auch als *inverses Problem* bezeichnet, als ein ziemlich obskures Thema erscheinen. Dies hat seine Ursache darin, daß dieses Gebiet erst mit dem Aufkommen ausreichender Rechnerleistung in den Blickpunkt aktueller Forschung gelangt und daher noch relativ neu und unübersichtlich ist. Ein weiterer Grund für dieses Unbehagen dürfte in der Vielfalt der interessierenden Integralgleichungen liegen, eine jede gekennzeichnet durch tückische, für die betrachtete Gleichung geradezu typische Fallgruben. Überdies existieren zur *Inversion* einer einzigen Klasse von Integralgleichungen oftmals viele verschiedene Lösungsverfahren.

Ziel dieses Berichtes ist es, eine spezielle Art von Integralgleichungen, wie sie in der Fernerkundung atmosphärischer Parameter auftreten, die *Fredholmschen Integralgleichungen erster Art* systematisch auf ihre Lösungsmöglichkeiten hin zu untersuchen. Hierzu werden zunächst mehrere Lösungsverfahren erarbeitet, diskutiert und anschließend einer Reihe einfacher Tests unterzogen. Die am Beispiel der bodengebundenen, passiven Millimeterwellenradiometrie gewonnenen Erkenntnisse über die Einsatzmöglichkeiten dieser Lösungsverfahren in der Fernerkundung werden in diesem Bericht festgehalten.

Karlsruhe, den 3. August 1994 Martin Kuntz

Inhaltsverzeichnis

Zusammenfassung iii

Vorwort v

1 Theorie der Inversion 1
- 1.1 Inverse und schlecht gestellte Probleme . . . 1
- 1.2 Methode der kleinsten Quadrate . . . 2
- 1.3 Singulärwertzerlegung und Regularisierung . . . 3
- 1.4 Wahl des Regularisierungsparameters . . . 5
- 1.5 Klassifizierung schlecht gestellter Probleme . . . 6
- 1.6 Erreichbare Auflösung und Fehler . . . 7

2 Regularisierungsverfahren 10
- 2.1 Abgeschnittene Singulärwertzerlegung . . . 10
- 2.2 Tikhonov–Phillips–Regularisierung . . . 12
- 2.3 Iterative Regularisierungsverfahren . . . 14
  - 2.3.1 Landweber–Iteration . . . 15
  - 2.3.2 Erweiterung der Landweber–Iteration nach Strand . . . 17
  - 2.3.3 Verfahren der konjugierten Gradienten . . . 17
  - 2.3.4 Chahinesches Verfahren . . . 20
  - 2.3.5 Eigene Modifikation des Chahineschen Verfahrens . . . 21
- 2.4 Stochastische Verfahren . . . 22
  - 2.4.1 Erwartungswert und Kovarianzmatrix . . . 23
  - 2.4.2 Verfahren der maximalen Wahrscheinlichkeit . . . 24
  - 2.4.3 Praktische Inversionsgleichung . . . 26
- 2.5 Backus–Gilbert–Verfahren . . . 27

3 Anwendung der Regularisierungsverfahren 30
- 3.1 Abgeschnittene Singulärwertzerlegung . . . 33
- 3.2 Zur Definition der Auflösung . . . 37
- 3.3 Anwendung der Tikhonov–Phillips–Regularisierung . . . 41
- 3.4 Anwendung der Landweber–Iteration . . . 45
- 3.5 Anwendung des konjugierten Gradienten–Verfahrens . . . 45
- 3.6 Anwendung des Chahineschen Verfahrens . . . 50

3.7 Anwendung der stochastischen Verfahren 53
3.8 Anwendung des Backus–Gilbert–Verfahrens 58

4 Beurteilung der Regularisierungsverfahren **62**

A Nebenrechnungen **65**
A.1 Zur Berechnung der verallgemeinerten Inversen 65
A.2 Zur Tikhonov–Phillips Regularisierung 65
A.3 Zur Landweber–Iteration 66
A.4 Zum Verfahren der maximalen Wahrscheinlichkeit 67
A.5 Zur praktischen Inversionsgleichung 67
A.6 Zum Backus Gilbert Verfahren 68

B Symbole und Abkürzungen **71**

Literaturverzeichnis **74**

Abbildungsverzeichnis

1 Rekonstruktionsfehler 6
2 Auflösungsmatrix 8
3 Filterwirkung der abgeschnittenen Singulärwertzerlegung 11
4 Filterwirkung der Tikhonov–Phillips–Regularisierung 13
5 Filterwirkung der Landweber–Iteration 16
6 Verfahren der maximalen Wahrscheinlichkeit 25
7 Datenfehler kontra Auflösung 29
8 Druck, Temperatur und Volumenmischungsverhältnis von Ozon . . . 31
9 Intensitätsspektrum und Eigenwerte 31
10 Gewichtsfunktionen 32
11 Abgeschnittene Singulärwertzerlegung 1 34
12 Abgeschnittene Singulärwertzerlegung 2 35
13 Fehler der abgeschnittenen Singulärwertzerlegung 36
14 Auflösungsmatrix der abgeschnittenen Singulärwertzerlegung 38
15 Auflösung der abgeschnittenen Singulärwertzerlegung 39
16 Tikhonov–Phillips–Regularisierung 1 42
17 Tikhonov–Phillips–Regularisierung 2 43
18 Auflösung der Tikhonov–Phillips–Regularisierung 44
19 Landweber–Iteration 46
20 Verfahren der konjugierten Gradienten 48
21 Einfluß des Startprofils auf das Verfahren der konjugierten Gradienten 49
22 Modifiziertes Chahinesches Verfahren 51
23 Einfluß des Startprofils auf das modifizierte Chahinesche Verfahren . 52
24 A priori Information für die stochastischen Verfahren 54
25 Ergebnisse der stochastischen Verfahren 55
26 Auflösung der stochastischen Verfahren 57
27 Backus–Gilbert–Verfahren 59
28 Auflösung des Backus–Gilbert–Verfahrens 60

Kapitel 1

Theorie der Inversion schlecht gestellter Probleme

1.1 Inverse und schlecht gestellte Probleme

Können die Eigenschaften eines Objektes nicht direkt gemessen werden, sondern muß man von indirekten Beobachtungen auf diese Größen zurückschließen, so handelt es sich um ein *inverses Problem* oder auch um ein sogenanntes *Identifizierungsproblem.* Ein Beispiel aus der Meteorologie ist die Fernerkundung atmosphärischer Parameter mittels bodengebundener, passiver Millimeterwellenradiometrie [18]. Die Möglichkeit der Fernerkundung resultiert aus der spektral veränderlichen Absorption, Emission und Streuung der innerhalb der Atmosphäre emittierten elektromagnetischen Strahlung. Die dafür eingesetzten Meßgeräte erfassen die Strahldichte in Abhängigkeit von der Wellenlänge.

Der Zusammenhang zwischen Strahlungsfeld und momentanem Zustand der Atmosphäre wird durch die Strahlungsübertragungsgleichung (SÜG) beschrieben [8]. Die Berechnung dieses Strahlungsfeldes aus Atmosphärendaten, das sogenannte *direkte* oder *Vorwärts–Problem*, ist in mathematischer Hinsicht relativ einfach. Im Gegensatz dazu beinhaltet der umgekehrte Fall, das inverse Problem, das Auffinden einer bestimmten Lösung der SÜG. Das Problem ist nichtlinear, da die zu ermittelnden Parameter, wie die Höhenverteilungen der Volumenmischungsverhältnisse interessierender Spurengase oder der Temperatur, nichtlinear über die Transmissionsfunktion in die SÜG eingehen.

Bisher wurde noch keine Theorie der direkten Inversion nichtlinearer Probleme entwickelt und entsprechend wenig Literatur ist verfügbar. [1] Statt dessen wurden primär lineare Fragestellungen behandelt, so daß hierfür eine ganze Reihe wichtiger mathematischer Hilfsmittel existiert, die u. a. auch Aussagen über die erreichbare Höhenauflösung sowie den Fehler der gefundenen Lösung erlauben. Aufgrund dessen betrachten wir hier nur den linearen bzw. linearisierten Fall [45, 31]. Vom mathematischen Standpunkt reduziert sich das Problem der Inversion der SÜG damit auf das Lösen einer *homogenen Fredholmschen Integralgleichung erster Art*, [37, Bd. IV/1

[1] Erste Fortschritte in dieser Richtung erzielt z. B. Schock [35], indem er eine, der in Abschn. 1.3 näher beschriebenen, linearen Singulärwertzerlegung entsprechende, nichtlineare Theorie entwickelt.

S. 14]

$$g(y) = \int_a^b K(y,x) f(x) dx. \tag{1}$$

Hierbei ist $f(x)$ das gesuchte Höhenprofil, die Funktion $K(y,x)$ der sogenannte Kern der Integralgleichung, im wesentlichen gegeben durch die Transmission der Atmosphäre, und $g(y)$ das beobachtete Strahldichtespektrum.

Die Unterteilung in *gut* und *schlecht gestellte* Probleme stammt von Hadamard [17]. Seiner Definition zufolge bezeichnet man Abbildungen der Form, $\mathcal{K}(f) = g$, als gut gestellt, wenn

i) für alle möglichen Daten g eine Lösung existiert,

ii) diese eindeutig ist und

iii) überdies stetig von den Daten abhängt.

Unter der letzten Bedingung ist zu verstehen, daß kleine Meßfehler in der rechten Seite g nur zu kleinen Fehlern in der Lösung f führen sollten. Ist eine dieser drei Bedingungen nicht erfüllt, so bezeichnet man das Problem als schlecht gestellt. Kennzeichnend für inverse Probleme der Art (1) ist nun, daß sie im obigen Sinne schlecht gestellt sind. Schwierigkeiten unter numerischen Gesichtspunkten bereitet hierbei vor allen Dingen Punkt iii), die Instabilität der Lösung gegenüber kleinen Meßfehlern in den Daten.

1.2 Verallgemeinerte Inverse – Methode der kleinsten Quadrate

Die Theorie der Intergralgleichungen ist in mancher Hinsicht der linearen Algebra analog. Wie bereits angedeutet, besteht die grundlegende Idee der Fernerkundung der Atmosphäre in der indirekten Bestimmung des Höhenprofils $f(x)$ aus dem Strahldichtespektrum $g(y)$, das an mehreren spektralen Stützstellen y_i, $i = 1, \ldots, n$ aufgenommen wird. Führt man überdies eine *Quadratur* des Typs $\int_a^b h(x)dx = \sum_{j=1}^m w_j\, h(x_j)$ ein, so ergibt sich aus (1) die Matrixgleichung [2]

$$\boldsymbol{g} = \boldsymbol{A}\boldsymbol{f}, \tag{2}$$

mit $A_{ij} = w_j K(y_i, x_j)$, $i = 1, \ldots, n$; $j = 1, \ldots, m$ sowie $f_j = f(x_j)$, $j = 1, \ldots, m$ bzw. $g_i = g(y_i)$, $i = 1, \ldots, n$. Die Lösung dieses Gleichungssystems, d. h. die Rekonstruktion des Urbildes $\boldsymbol{f}$, ist sehr einfach, wenn $\boldsymbol{A}$ eine quadratische Matrix und

[2] Fette Großbuchstaben kennzeichnen im folgenden Matrizen, fette Kleinbuchstaben kennzeichnen Vektoren.

stabil, d. h. weder singulär noch nahezu singulär ist. In diesem Falle ist das Problem nach Hadamard gut gestellt und es ergibt sich

$$\boldsymbol{f} = \boldsymbol{A}^{-1}\boldsymbol{g}.$$

I. a. ist die Matrix $\boldsymbol{A}$ jedoch nicht quadratisch, so daß das Inversionsproblem bereits formal bzgl. der Existenz und Eindeutigkeit der Lösung schlecht gestellt ist. Wir erweitern deshalb den Lösungsbegriff, um auch diesen Fall mit einzuschließen und betrachten statt der Gleichung (2) das Problem, den Abstand [3]

$$||\boldsymbol{Af} - \boldsymbol{g}|| = \sqrt{< \boldsymbol{Af} - \boldsymbol{g} | \boldsymbol{Af} - \boldsymbol{g} >}$$

zu minimieren. Wir suchen also nach derjenigen Lösung, die nach erfolgter Vorwärtsrechnung die gemessenen Daten bzgl. der L_2-Norm bestmöglich approximiert, auch bekannt als die Methode der kleinsten Quadrate. Dieses Minimierungsproblem führt nach einigen Zwischenschritten (siehe Anhang A.1) auf die Gleichung [4]

$$\boldsymbol{A}^*\boldsymbol{Af} - \boldsymbol{A}^*\boldsymbol{g} = 0. \tag{3}$$

Hieraus folgt für die bestapproximierte oder *verallgemeinerte* [5] *Lösung*, oft auch als *Moore–Penrose–Lösung* [27, S. 85] oder *least squares solution* bezeichnet,

$$\boldsymbol{f} = \underbrace{(\boldsymbol{A}^*\boldsymbol{A})^{-1}\boldsymbol{A}^*}_{\boldsymbol{A}^{-A}}\boldsymbol{g}. \tag{4}$$

Die neu eingeführte Matrix $\boldsymbol{A}^{-A}$ bezeichnen wir als *verallgemeinerte Inverse* [6].

Charakteristisch für inverse Probleme ist nun, daß selbst die verallgemeinerte Lösung (4) nicht berechnet werden kann, da das inverse Problem schlecht gestellt ist bzgl. Punkt iii) der Hadamardschen Definition. Selbst kleinste Meßfehler in den Daten führen zu sehr großen Fehlern in der Lösung. Ursache hierfür ist die fast lineare Abhängigkeit der Gleichungen des zugrundeliegenden Gleichungssystems (2), bzw. die nahezu Singularität der Matrix $\boldsymbol{A}^*\boldsymbol{A}$.

1.3 Singulärwertzerlegung und Regularisierung

Um die Ursachen dieser Schwierigkeiten besser zu verstehen und um ein Mittel in die Hand zu bekommen, dieser extremen Verstärkung der Meßfehler entgegenzuwirken,

[3] Im folgenden wird für das Skalarprodukt die Diracsche Notation $< | >$ verwendet.

[4] Der Stern kennzeichnet die Adjungierte einer Matrix.

[5] Verallgemeinert in dem Sinne, daß ein überbestimmtes Gleichungssystem keine eindeutige Lösung besitzt, wir jedoch nach der Lösung mit der kleinsten L_2-Norm suchen.

[6] Die Schreibweise als $\boldsymbol{A}^{-A}$ soll die enge Verwandtschaft mit der nur im quadratischen Fall existierenden Inversen $\boldsymbol{A}^{-1}$ symbolisieren.

wollen wir zu einer anderen Darstellung der verallgemeinerten Lösung (4) gelangen. Hierzu bestimmen wir die Eigenwerte σ_n^2 und Eigenvektoren $\boldsymbol{v}_n$ der Matrix $\boldsymbol{A}^*\boldsymbol{A}$, [4], [23, S. 23], [27, S. 85 ff.], [28, S. 59 ff.]

$$\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{v}_n = \sigma_n^2 \boldsymbol{v}_n.$$

und stellen die verallgemeinerte Lösung (4) dar in der Basis der Eigenvektoren,

$$\begin{aligned} \boldsymbol{f} &= \sum_n < \boldsymbol{f}|\boldsymbol{v}_n > \boldsymbol{v}_n \\ &= \sum_n < (\boldsymbol{A}^*\boldsymbol{A})^{-1}\boldsymbol{A}^*\boldsymbol{g}|\boldsymbol{v}_n > \boldsymbol{v}_n \\ &= \sum_n \sigma_n^{-2} < (\boldsymbol{A}^*\boldsymbol{A})^{-1}\boldsymbol{A}^*\boldsymbol{g}|\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{v}_n > \boldsymbol{v}_n \\ &= \sum_n \sigma_n^{-2} < \boldsymbol{g}|\boldsymbol{A}\boldsymbol{v}_n > \boldsymbol{v}_n. \end{aligned} \tag{5}$$

Anhand dieser als Singulärwertzerlegung bezeichneten Darstellung der verallgemeinerten Lösung werden die bei Inversionsproblemen auftretenden Schwierigkeiten deutlich. Sind nämlich die Daten $\boldsymbol{g}$ mit Meßfehlern behaftet, so werden diese Fehler aufgrund kleiner Eigenwerte σ_n^2 enorm verstärkt und die verallgemeinerte Lösung unbrauchbar.

Um trotzdem zu einem vernünftigen Ergebnis zu gelangen, müssen alle diejenigen Beiträge, die aufgrund kleiner Eigenwerte eine extreme Verstärkung von Meßfehlern bewirken, unterdrückt werden. Z. B. durch den Einbau eines entsprechenden Filters $F_\gamma(\sigma_n)$, das alle Beiträge zu (5), kleiner als ein noch zu wählender *Regularisierungsparameter* γ, eliminiert. Statt von einer Filterung sprechen wir auch von einer *Regularisierung* der verallgemeinerten Lösung, im folgenden jeweils gekennzeichnet durch ein tiefgestelltes γ,

$$\boxed{\boldsymbol{f}_\gamma = \sum_n \sigma_n^{-2} F_\gamma(\sigma_n) < \boldsymbol{g}|\boldsymbol{A}\boldsymbol{v}_n > \boldsymbol{v}_n.} \tag{6}$$

Gl. 6 können wir auch schreiben als [7]

$$\begin{aligned} \boldsymbol{f}_\gamma &= \sum_n \sigma_n^{-2} F_\gamma(\sigma_n)\, \boldsymbol{v}_n < \boldsymbol{A}\boldsymbol{v}_n|\boldsymbol{g} > \\ &= \underbrace{\sum_n \sigma_n^{-2} F_\gamma(\sigma_n)\, \boldsymbol{v}_n \boldsymbol{v}_n^* \boldsymbol{A}^*}_{\boldsymbol{A}_\gamma^{-A}}\, \boldsymbol{g}, \end{aligned} \tag{7}$$

wobei wir die neu eingeführte *regularisierte verallgemeinerte Inverse* $\boldsymbol{A}_\gamma^{-A}$ zur Abgrenzung gegenüber der bereits in in Abschn. 1.2 eingeführten verallgemeinerten Inversen $\boldsymbol{A}^{-A}$ mit einem tiefgestellten Regularisierungsparameter γ versehen.

[7] Nachteil dieser zweiten Darstellung ist, daß sie auf dem Rechner wesentlich mehr Speicherplatz benötigt als die Darstellung (6).

Die eigentliche Kunst der Regularisierung besteht in der Wahl des optimalen Regularisierungsparameters γ, der die Stärke einer Regularisierung festlegt, wie auch in der Wahl der richten Art Filter, das abhängig vom gewählten Regularisierungsverfahren jeweils unterschiedlich beschaffen sein kann. In Kapitel 2 werden daher einige der bekanntesten Regularisierungsverfahren vorgestellt und speziell im Hinblick auf ihre Filterwirkung näher untersucht.

1.4 Wahl des Regularisierungsparameters

Stehen nur gestörte Daten, $g_\varepsilon = g + \varepsilon$, zur Verfügung, so erhalten wir nach erfolgter Regularisierung mit dem Regularisierungsparameter γ aufgrund von (7) die Näherungslösung

$$f_\gamma = A_\gamma^{-A} g_\varepsilon .$$

Der hierbei gemachte *Rekonstruktionsfehler*, d. h. der Fehler der Näherungslösung,

$$f_\gamma - f = A_\gamma^{-A} g_\varepsilon - A^{-A} g = \underbrace{A_\gamma^{-A} \varepsilon}_{Datenfehler} + \underbrace{(A_\gamma^{-A} - A^{-A}) g}_{Regularisierungsfehler} , \tag{8}$$

setzt sich aus zwei Anteilen zusammen [23, S. 10], vgl. auch [39, S. 804 Gl. (26)] Den ersten Teil nennen wir *Datenfehler* [8], da er in erster Linie auf Meßfehler in den Daten zurückzuführen ist, den zweiten Teil nennen wir *Regularisierungsfehler*, weil er sich als Folge der Regularisierung ergibt. Qualitativ erhalten wir nun folgendes Bild (siehe Abb. 1), das kennzeichnend für alle inversen und schlecht gestellten Probleme ist. Wird der Regularisierungsparameter γ zu 0 gewählt, d. h. wird überhaupt nicht regularisiert (wir berechnen lediglich die verallgemeinerte Lösung oder least squares solution), so ist der Datenfehler aufgrund der Beiträge kleiner Eigenwerte sehr groß. Regularisiert man infolgedessen stärker, so nimmt der Datenfehler zwar ab, doch wächst der Regularisierungsfehler umso stärker, je größer der Regularisierungsparameter γ gewählt wird. Der Rekonstruktionsfehler kann deshalb nicht beliebig klein gemacht werden, sondern es ist lediglich möglich, ihn bezüglich γ zu optimieren.

In der Praxis ist die exakte Lösung nicht bekannt, so daß das optimale γ gar nicht ermittelt werden kann. Statt dessen erfolgt die Wahl des Regularisierungsparameters entweder *a priori* aufgrund von Zusatzinformationen, die Erfahrungswerten oder anderen Messungen entstammen (z. B. bei den stochastischen Verfahren in Abschn. 2.4) oder *a posteriori* mittels *try and error* [23, S. 54]. Im letzteren Fall wird die Wahl abhängig gemacht vom Defekt zwischen gemessenen Daten g_ε und den aus

[8] Die Bezeichnung Datenfehler ist insofern irreführend, als unter dem Datenfehler nicht der Fehler in den Daten an sich, den wir im folgenden zur Unterscheidung als Defekt in den Daten bezeichnen wollen, sondern der aus diesem Defekt in den Daten resultierende Anteil am Gesamtfehler der Lösung zu verstehen ist. Der Begriff Datenfehler greift also in der „Lösungsebene“ an.

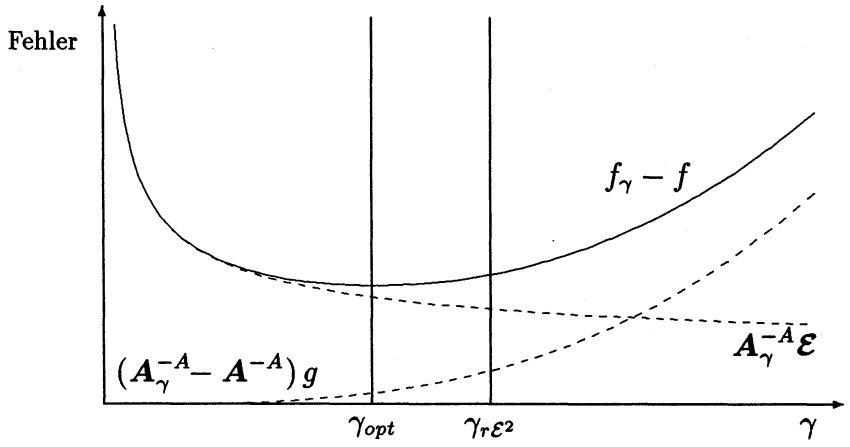

Abbildung 1: Regularisierungsfehler (monoton steigend), Datenfehler (monoton fallend) und Gesamt- oder Rekonstruktionsfehler in Abhängigkeit vom Regularisierungsparameter γ.

der Näherungslösung durch Vorwärtsrechnung ermittelten Daten $\boldsymbol{g}_\gamma$,

$$||\boldsymbol{g}_\mathcal{E} - \boldsymbol{g}_\gamma|| = ||\boldsymbol{g} + \mathcal{E} - \boldsymbol{g}_\gamma|| = ||\boldsymbol{A}(\boldsymbol{f} - \boldsymbol{f}_\gamma) + \mathcal{E}||. \tag{9}$$

(So z. B. bei der abgeschnittenen Singulärwertzerlegung, Abschn. 2.1, bei der Tikhonov–Phillips Regularisierung, Abschn. 2.2, oder bei den iterativen Verfahren, Abschn. 2.3). Man geht hierbei so vor, daß man ausgehend von einem hinreichend großen γ durch sukzessives Verkleinern des Regularisierungsparameters den Defekt in den Daten (9) unter eine vorgegebene Schranke $r\mathcal{E}^2 = r||\mathcal{E}||^2$ bringt [23, S. 68], [39, S. 805 Gl. (30)]. Hierbei muß der ansonsten frei wählbare Parameter $r \geq 1$ gesetzt werden. Mit anderen Worten, man bestimmt aus der Vielzahl der möglichen Lösungen, für die nach erneuter Vorwärtsrechnung die Abweichung von den gemessenen Daten kleiner ist als ein angenommener Fehler $r\mathcal{E}^2$, die am stärksten regularisierte Lösung. Auf diese Weise wird sichergestellt, daß einerseits die Näherungslösung nur geringfügig mit Datenfehlern behaftet und daß andererseits der Regularisierungsfehler nicht allzu groß ist. Wie aus Abb. 1 ersichtlich, liefert die a posteriori Parameterwahl mit dem aus ihr folgenden Wert für den Regularisierungsparameter, $\gamma_{r\mathcal{E}^2}$, keinesfalls das optimalen γ, sondern einen etwas zu großen Wert.

1.5 Klassifizierung schlecht gestellter Probleme – Regularisierung als Tiefpaß

Ordnet man die in Abschn. 1.3 berechneten Eigenwerte der Matrix $\boldsymbol{A}^*\boldsymbol{A}$ entsprechend ihrer Größe, $\sigma_1 > \sigma_2 > \ldots > \sigma_n$, so können anhand dieses Kriteriums schlecht gestellte Probleme klassifiziert werden [23]. Man unterscheidet hierbei zwischen Problemen,

die *schlecht gestellt* sind *von der Ordnung* α, wenn $\sigma_n = \mathcal{O}(n^{-\alpha})$ mit $\alpha > 0$, und *exponentiell schlecht gestellten* Problemen, sofern $|\ln \sigma_n| \geq cn^{\rho}$ und $\rho > 0$. Bei der Fernerkundung atmosphärischer Parameter mittels bodengebundener, passiver Millimeterwellenradiometrie handelt es sich um ein exponentiell schlecht gestelltes Problem, wie sich aufgrund der abgeschnittenen Singulärwertzerlegung (siehe Abschn. 2.1 bzw. Abschn. 3.1) noch zeigen wird.

Wie Louis [23, S. 49-50] zeigt, entsprechen kleinen Eigenwerten bei schlecht gestellten Operatoren stark oszillierende Eigenvektoren $\boldsymbol{v}_n$, so daß sich Meßfehler in den Daten in einem starken Oszillieren der Lösung bemerkbar machen (siehe auch die Abbildungen 11 und 12 in Abschn. 3.1). Der Einbau eines Filters F_γ wirkt daher wie ein Tiefpaß, der hochfrequente Anteile der Lösung abschneidet. Letztlich wird dies erreicht, indem über benachbarte Elemente der unregularisierten Lösung lokal gemittelt wird, so daß eine Reduktion der Anzahl der Unbekannten f_j (und damit der Auflösung (siehe Abschn. 1.6)) erfolgt. Man hätte sich also von vornherein darauf beschränken können – sofern dies überhaupt möglich ist [9] – weniger Unbekannte aus den Daten zu ermitteln. Bei einer weniger feinen Diskretisierung gleich zu Beginn hätte man dann unter Umständen ganz auf eine Regularisierung verzichten können.

Bei einer solchen Vorgehensweise ist es jedoch relativ schwierig, ein Maximum an Information aus den Daten zu extrahieren, da man im voraus leider nicht weiß, wie hoch der Informationsgehalt der Messungen tatsächlich ist. Statt dessen ist es „wesentlich günstiger, zunächst etwas feiner zu diskretisieren und anschließend die diskreten Probleme zu regularisieren“ [23]. Die im folgenden Kapitel 2 vorgestellten Regularisierungsverfahren geben hierzu das notwendige numerische Rüstzeug, um anhand weniger Steuer- bzw. Regularisierungsparameter eine vernünftige Inversion mit optimal angepaßter Auflösung durchzuführen.

1.6 Erreichbare Auflösung – Fehler der regularisierten Lösung

Aufschluß darüber, inwieweit Strukturen in der ursprünglichen Funktion $\boldsymbol{f}$ nach erfolgter Regularisierung noch aufgelöst werden können, gibt die Auflösungsmatrix $\boldsymbol{R}$ [24, S. 64], [34]. Man erhält sie, indem man ausgehend vom Vorwärtsproblem (2) dieses in die regularisierte Lösung $\boldsymbol{f}_\gamma$ (6) bzw. (7) einsetzt:

$$\boldsymbol{f}_\gamma = \boldsymbol{A}_\gamma^{-A}\boldsymbol{g} = \boldsymbol{A}_\gamma^{-A}\boldsymbol{A}\boldsymbol{f} = \boldsymbol{R}\boldsymbol{f}$$

Im Idealfall, wenn alle Elemente der Lösung eindeutig bestimmt sind und dem Urbild entsprechen, sollte sich für $\boldsymbol{R}$ die Einheitsmatrix ergeben. Ansonsten erhalten wir als regularisierte Näherungslösung lokale Mittelwerte des exakten Urbildes, d. h. ein mehr

[9] Z. B. ist bei den Horizontsondierungsverfahren die Zahl der Unbekannten bzw. die Zahl der vertikalen Stützstellen bereits durch die Beobachtungsgeometrie festgelegt [9].

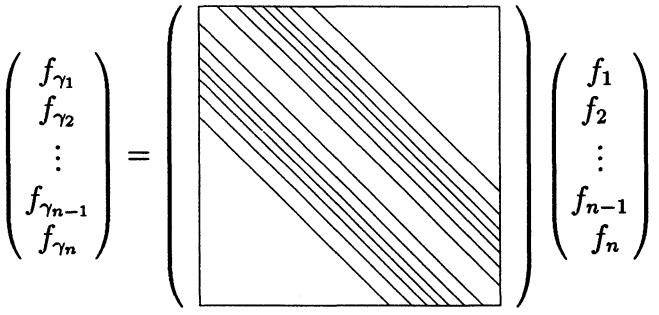

Abbildung 2: Die Zeilen der Auflösungsmatrix $\boldsymbol{R}$ geben Aufschluß darüber, inwieweit Strukturen im Urbild noch aufgelöst werden können.

oder weniger stark geglättetes Urbild. $\boldsymbol{R}$ ist dann je nach sich ergebender Auflösung eine mehr oder weniger breite Bandmatrix (siehe Abb. 2). Das Band muß dabei nicht überall gleich breit sein, sondern die Auflösung kann sich auch mit der Höhe ändern (vgl. Abb. 14 in Abschn. 3.1).

Überdies kann für die Fehlergrenzen bzw. die Streuung der regularisierten Lösung $\boldsymbol{f}_\gamma$, wenn auch nur bedingt, eine Abschätzung vorgenommen werden. Ausgangspunkt sind die Überlegungen zum Rekonstruktionsfehler (8) in Abschn. 1.4. Ist die Kovarianzmatrix [10] $\boldsymbol{\mathcal{S}}_\varepsilon$ und damit das stochastische Verhalten des Meßfehlers $\boldsymbol{\mathcal{E}}$ bekannt, so kann der Beitrag des Datenfehlers $\boldsymbol{A}_\gamma^{-A}\boldsymbol{\mathcal{E}}$, zum Gesamtfehler $\boldsymbol{f}_\gamma - \boldsymbol{f}$ berechnet werden. Wir wissen dann zumindest, inwieweit die Streuung des Meßfehlers sich auf eine Streuung der regularisierten Lösung auswirkt. Aufgrund der Rechenregeln für Kovarianzmatrizen [26, S. 82 ff.] folgt für die Kovarianzmatrix $\boldsymbol{\mathcal{S}}_\gamma$ der regularisierten Lösung $\boldsymbol{f}_\gamma$ unmittelbar

$$\boldsymbol{\mathcal{S}}_\gamma = \boldsymbol{A}_\gamma^{-A}\boldsymbol{\mathcal{S}}_\varepsilon\, \boldsymbol{A}_\gamma^{-A\,*}. \tag{10}$$

Die Kovarianzmatrix stellt die Erweiterung der vom eindimensionalen Fall bekannten Varianz auf Vektoren dar, wobei die Diagonalelemente der Kovarianzmatrix durch die Varianzen der Komponenten von $\boldsymbol{f}_\gamma$ gegeben sind. Da die Streuung oder Standardabweichung definiert ist als die Wurzel aus der Varianz, ist somit eine Aussage über die Streuung der einzelnen Komponenten der regularisierten Lösung möglich.

Sowohl die Auflösungsmatrix $\boldsymbol{R}$ als auch die Kovarianzmatrix $\boldsymbol{\mathcal{S}}_\gamma$ sind überaus wichtige Diagnostik-Werkzeuge bei der Durchführung von Inversionsrechnungen; dies insbesondere, da beide lediglich Funktionen des Kerns $\boldsymbol{A}$ sowie der regularisierten verallgemeinerten Inversen $\boldsymbol{A}_\gamma^{-A}$ sind, jedoch nicht von den aktuellen Daten $\boldsymbol{g}_\varepsilon$ abhängen. Einzige Ausnahme bilden die nichtlinearen iterativen Verfahren, wie z. B das konjugierte Gradienten–Verfahren oder das Chahinesche Verfahren (siehe Abschn. 2.3.3

[10] Im Abschn. 2.4 über die stochastischen Regularisierungsverfahren werden die Eigenschaften und die Definition der Kovarianzmatrix ausführlicher behandelt, so daß wir uns hier auf die zum Verständnis notwendigen Erläuterungen beschränken.

bzw. Abschn. 2.3.4), bei denen die Daten $\boldsymbol{g}_{\mathcal{E}}$ in die Berechnung des Filters $F_\gamma(\sigma_n^2)$ mit einfließen. Die Matrix $\boldsymbol{A}_\gamma^{-A}$ ist bei diesen Verfahren deshalb nicht explizit angebbar, so daß auch obige Betrachtungen zur Auflösung und zum Fehler nicht durchgeführt werden können.

Kapitel 2

Regularisierungsverfahren

In den letzten 10-20 Jahren hat die Zahl der Publikationen über inverse und schlecht gestellte Probleme stark zugenommen, so daß in diesem Kapitel lediglich die wichtigsten Regularisierungsverfahren kurz skizziert werden können. Soweit als möglich wird dabei jedes Verfahren interpretiert als Filterung der verallgemeinerten Inversen (6). Eine Übersicht über die verschiedenen Regularisierungsverfahren von der mathematischen Seite geben z. B. Baumeister [2], Natterer [27], Strand [39] sowie Louis [23], dem wir in diesem Abschnitt im wesentlichen folgen wollen. Eher anwendungsorientiert sind dagegen die Ausführungen in Westwater [45], Menke [24], Rodgers [31], Twomey [42], sowie Ulaby [43].

2.1 Abgeschnittene Singulärwertzerlegung

Das aufgrund der bisherigen Überlegungen sicherlich naheliegendste Regularisierungsverfahren ist die abgeschnittene Singulärwertzerlegung, [2, S. 49], [23, S. 79], [25] sowie [39]. Hierbei werden die in (6) gegenüber Datenfehlern stark empfindlichen Beiträge durch das spezielle Filter

$$F_\gamma(\sigma_n) = \begin{cases} 1 & \text{für} \quad \sigma_n \geq \gamma \\ 0 & \text{für} \quad \sigma_n < \gamma \end{cases}$$

einfach abgeschnitten (siehe Abb. 3), so daß aus (6) in Abschn. 1.3 folgt

$$\boxed{\boldsymbol{f}_\gamma = \sum_{\sigma_n > \gamma} \sigma_n^{-2} < \boldsymbol{g}_\mathcal{E} | \boldsymbol{A}\boldsymbol{v}_n > \boldsymbol{v}_n.} \tag{11}$$

Dies setzt zunächst die Durchführung einer Singulärwertzerlegung samt Berechnung der zugehörigen Eigenvektoren $\boldsymbol{v}_n$ (z. B. Jacobi- oder Householder–Verfahren [11, 28]) voraus. Die abgeschnittene Singulärwertzerlegung ist damit in numerischer Hinsicht ein recht aufwendiges Verfahren. Andererseits besticht sie durch die Tatsache, daß, sofern die Parameterwahl a posteriori erfolgt, keine a priori Information benötigt wird. Die abgeschnittene Singulärwertzerlegung ist also insbesondere dann sinnvoll, wenn keine oder nur sehr wenig a priori Information zur Verfügung steht. Überdies ist die in

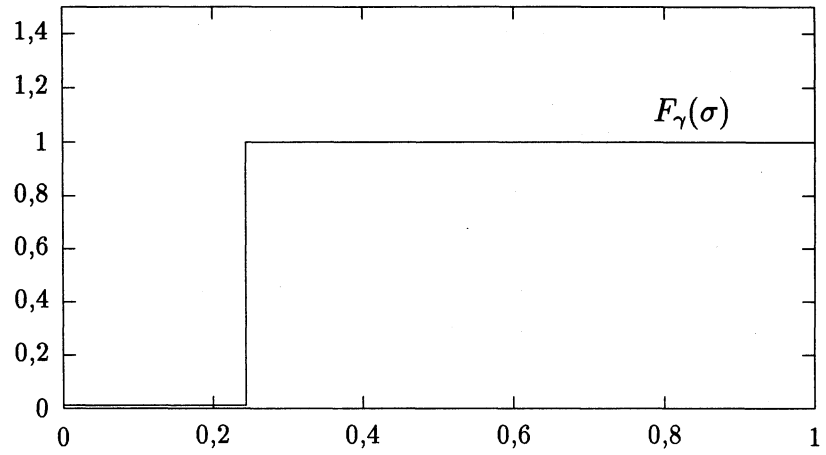

Abbildung 3: Filterwirkung der abgeschnittenen Singulärwertzerlegung für einen Regularisierungsparameter $\gamma = 0,25$.

Abschn. 1.4 Gl. (9) beschriebene a posteriori Parameterwahl bei der abgeschnittenen Singulärwertzerlegung besonders effizient durchzuführen [23, S. 82]. Da

$$\begin{aligned} ||\boldsymbol{g}_\varepsilon - \boldsymbol{g}_\gamma||^2 \leq ||\boldsymbol{g}_\varepsilon||^2 - ||\boldsymbol{g}_\gamma||^2 &= ||\boldsymbol{g}_\varepsilon||^2 - ||\boldsymbol{A}\boldsymbol{f}_\gamma||^2 \\ &= ||\boldsymbol{g}_\varepsilon||^2 - ||\sum_{\sigma_n > \gamma} \sigma_n^{-2} < \boldsymbol{g}_\varepsilon | \boldsymbol{A}\boldsymbol{v}_n > \boldsymbol{A}\boldsymbol{v}_n||^2 \\ &= ||\boldsymbol{g}_\varepsilon||^2 - \sum_{\sigma_n > \gamma} \sigma_n^{-2} < \boldsymbol{g}_\varepsilon | \boldsymbol{A}\boldsymbol{v}_n >^2, \end{aligned} \tag{12}$$

wobei von der Orthogonalität der Eigenvektoren $\boldsymbol{v}_n$ Gebrauch gemacht wird, kann der Defekt zwischen Meßdaten $\boldsymbol{g}_\varepsilon$ und den sich aus der Näherungslösung (9) ergebenden Daten $\boldsymbol{g}_\gamma$ unter minimalem numerischen Zusatzaufwand berechnet werden. Man beginnt mit einem $\gamma = \infty$ und subtrahiert vom Betragsquadrat $||\boldsymbol{g}_\varepsilon||^2$ so lange Terme $\sigma_n^{-2} < \boldsymbol{g}_\varepsilon | \boldsymbol{A}\boldsymbol{v}_n >^2$, bis der Defekt unter die Meßfehlergenauigkeit $r\mathcal{E}^2$ gelangt. Da sowohl die Eigenwerte σ_n^2 wie auch die Skalarprodukte $< \boldsymbol{g}_\varepsilon | \boldsymbol{A}\boldsymbol{v}_n >$ bei der Berechnung der Lösung (11) automatisch anfallen, ist der Zusatzaufwand entsprechend gering.

Ein weiterer Vorteil der abgeschnittenen Singulärwertzerlegung ist, daß aufgrund der Kenntnis der Eigenwerte die Schlechtgestelltheit des Problems leicht klassifiziert werden kann (siehe Abschn. 1.5 bzw. Abschn. 3.1 Abb. 9.b). Auch vom mathematischen Standpunkt aus betrachtet ist die abgeschnittene Singulärwertzerlegung eines der saubersten Verfahren, da im Verlauf des Verfahrens lediglich die stark fehlerbehafteten Beiträge abgeschnitten werden, andererseits jedoch keine (unter Umständen aufgrund falscher Annahmen ebenfalls mit Fehlern behaftete) a priori Information in die Lösung mit einfließt.

2.2 Tikhonov–Phillips–Regularisierung

Die Tikhonov–Phillips–Regularisierung [2, S. 33], [23, S. 87] ist eines der am häufigsten angewandten Regularisierungsverfahren und hat unter dementsprechend vielen Namen Eingang in die Literatur gefunden, z. B. als *smoothing method* [20, S. 504], als *damped least squares solution* [24, S. 52], als *regularization method* [28, S. 808], als *Twomey–Tikhonov method* [31] oder als *linear constraint solution* [42, S. 122]. Ausgangspunkt des Tikhonov–Phillips–Verfahrens ist die Beobachtung, daß der Informationsgehalt der fehlerbehafteten Daten $\boldsymbol{g_\varepsilon}$ sowie des Kerns $\boldsymbol{A}$ allein nicht ausreicht, um eine Lösung zu bestimmen. Deshalb werden dem Gleichungssystem a priori Informationen der Form

$$\boldsymbol{Bf} = \boldsymbol{r} \tag{13}$$

zugeführt. Dies können z. B. Informationen über Randbedingungen bzw. irgendwelche Glattheitsbedingungen an $\boldsymbol{f}$ sein. Die Aufgabenstellung lautet damit: Minimiere

$$||\boldsymbol{Af} - \boldsymbol{g_\varepsilon}||^2 \tag{14}$$

unter der Restriktion, daß die Zusatzbedingungen (13) ebenfalls erfüllt werden. Mit Hilfe eines *Lagrangeschen Multiplikators* γ^2 führen wir dieses restringierte Problem über in ein unrestringiertes, indem wir statt des Funktionals (14) das Funktional

$$||\boldsymbol{Af}_\gamma - \boldsymbol{g_\varepsilon}||^2 + \gamma^2||\boldsymbol{Bf}_\gamma - \boldsymbol{r}||^2 \tag{15}$$

minimieren. Nach einigen Zwischenschritten (siehe Anhang A.2) führt dies auf die Gleichung

$$(\boldsymbol{A}^*\boldsymbol{A} + \gamma^2\boldsymbol{B}^*\boldsymbol{B})\boldsymbol{f}_\gamma = \boldsymbol{A}^*\boldsymbol{g_\varepsilon} + \gamma^2\boldsymbol{B}^*\boldsymbol{r}$$

und damit auf die regularisierte Lösung

$$\boxed{\boldsymbol{f}_\gamma = (\boldsymbol{A}^*\boldsymbol{A} + \gamma^2\boldsymbol{B}^*\boldsymbol{B})^{-1}(\boldsymbol{A}^*\boldsymbol{g_\varepsilon} + \gamma^2\boldsymbol{B}^*\boldsymbol{r}).} \tag{16}$$

Wir wollen nun (16) als Filterung der verallgemeinerten Inversen (4) interpretieren. Unter der einschränkenden Annahme, daß $\boldsymbol{A}^*\boldsymbol{A}$ und $\boldsymbol{B}^*\boldsymbol{B}$ ein gemeinsames System von Eigenvektoren $\boldsymbol{v}_n$ besitzen,

$$\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{v}_n = \sigma_n^2\boldsymbol{v}_n \quad \text{und} \quad \boldsymbol{B}^*\boldsymbol{B}\boldsymbol{v}_n = \beta_n^2\boldsymbol{v}_n,$$

können wir für $\boldsymbol{f}_\gamma$ eine Singulärwertzerlegung durchführen

$$\begin{aligned}
\boldsymbol{f}_\gamma &= \sum_n < \boldsymbol{f}_\gamma|\boldsymbol{v}_n > \boldsymbol{v}_n \\
&= \sum_n < \boldsymbol{A}^*\boldsymbol{g_\varepsilon} + \gamma^2\boldsymbol{B}^*\boldsymbol{r}|(\boldsymbol{A}^*\boldsymbol{A} + \gamma^2\boldsymbol{B}^*\boldsymbol{B})^{-1}\boldsymbol{v}_n > \boldsymbol{v}_n \\
&= \sum_n (\sigma_n^2 + \gamma^2\beta_n^2)^{-1}(< \boldsymbol{g_\varepsilon}|\boldsymbol{A}\boldsymbol{v}_n > + \gamma^2 < \boldsymbol{r}|\boldsymbol{B}\boldsymbol{v}_n >)\boldsymbol{v}_n,
\end{aligned}$$

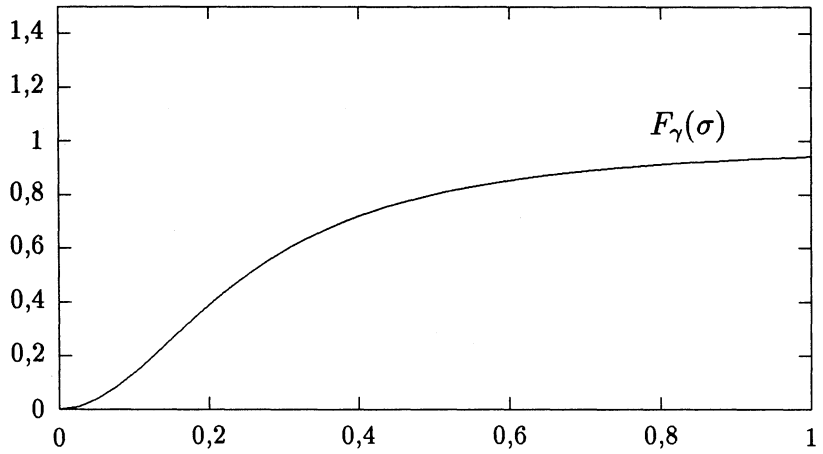

Abbildung 4: Filterwirkung der Tikhonov–Phillips–Regularisierung für einen Regularisierungsparameter $\gamma = 0,25$ und Parameter $\beta = 1$.

und erhalten für die spezielle Wahl $\boldsymbol{r} = 0$ die gewünschte Darstellung (vgl. (6))

$$\boldsymbol{f}_\gamma = \sum_n \sigma_n^{-2} F_\gamma(\sigma_n) < \boldsymbol{g}_\mathcal{E} | \boldsymbol{A}\boldsymbol{v}_n > \boldsymbol{v}_n$$

mit einem Filter

$$F_\gamma(\sigma_n) = \frac{\sigma_n^2}{\sigma_n^2 + \gamma^2 \beta_n^2}. \tag{17}$$

Die Wirkung der Tikhonov–Phillips–Regularisierung (16) läßt sich daher wie folgt beschreiben (siehe Abb. 4). Ist σ_n^2 groß im Vergleich zu $\gamma^2\beta_n^2$, werden die entsprechenden Anteile der Lösung nur wenig verfälscht. Ist σ_n^2 dagegen klein, so werden diejenigen Anteile, die einer starken Beeinflußung durch Datenfehler unterliegen, weggedämpft – im Gegensatz zur abgeschnittenen Singulärwertzerlegung jedoch nicht vollständig abgeschnitten.

In der ursprünglichen, auf Tikhonov zurückgehenden Formulierung mit der Zusatzbedingung $\boldsymbol{f} = 0$, d. h. $\boldsymbol{B} \equiv \boldsymbol{I}$ sowie $\boldsymbol{r} \equiv 0$ in (13) waren obige Annahmen natürlich erfüllt, so daß das Verfahren als Filterung der verallgemeinerten Inversen interpretiert werden konnte. Phillips erweiterte dann das Verfahren, indem er feststellte, daß Oszillationen besser durch die zusätzliche Bedingung $\boldsymbol{f}' = 0$, also das Verschwinden der ersten Ableitung, unterdrückt werden können, d. h.

$$\boldsymbol{B}\boldsymbol{f} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-1} \\ f_n \end{pmatrix} = 0.$$

Er fordert also explizit, daß die Lösung hinreichend glatt sein soll. Grundsätzlich kann jedoch jede Art von a priori Information in der Zusatzbedingung (13) mit berücksichtigt werden. Neben Randbedingungen können dies z. B. auch Kenntnisse über das Verhalten höherer Ableitungen von $\boldsymbol{f}$ usw. sein. Überdies ist es möglich, unterschiedliche Meßgenauigkeiten getrennt zu berücksichtigen, indem man einzelne Messungen unterschiedlich gewichtet. Am einfachsten geschieht dies durch Anwendung einer positiv definiten Diagonalmatrix $\boldsymbol{C}$, so daß statt des Funktionals (15) das Funktional

$$(\boldsymbol{A}\boldsymbol{f}_\gamma - \boldsymbol{g}_\varepsilon)^* \boldsymbol{C} (\boldsymbol{A}\boldsymbol{f}_\gamma - \boldsymbol{g}_\varepsilon) + \gamma^2 (\boldsymbol{B}\boldsymbol{f}_\gamma - \boldsymbol{r})^* (\boldsymbol{B}\boldsymbol{f}_\gamma - \boldsymbol{r}) \tag{18}$$

minimiert werden muß. Letztlich führt dies auf die gewichtete, regularisierte Lösung

$$\boxed{\boldsymbol{f}_\gamma = (\boldsymbol{A}^* \boldsymbol{C} \boldsymbol{A} + \gamma^2 \boldsymbol{B}^* \boldsymbol{B})^{-1} (\boldsymbol{A}^* \boldsymbol{C} \boldsymbol{g}_\varepsilon + \gamma^2 \boldsymbol{B} \boldsymbol{r}).} \tag{19}$$

Die Wahl des Regularisierungsparameters γ erfolgt wie auch bei der abgeschnittenen Singulärwertzerlegung entweder a priori aufgrund irgendwelcher Vergleichswerte oder a posteriori, wie in Abschn. 1.4 beschrieben, wobei der numerische Aufwand nun größer ist.

Ein wichtiger Vorteil der Tikhonov–Phillips–Regularisierung gegenüber der abgeschnittenen Singulärwertzerlegung ist, daß bei diesem Verfahren der Regularisierungsparameter kontinuierlich eingestellt werden kann. Es ist damit möglich, die Stärke der Regularisierung stufenlos zu regeln. Bei der abgeschnittenen Singulärwertzerlegung besteht dagegen nur die Möglichkeit, einen weiteren Eigenvektor in die Lösung mit einzubeziehen.

2.3 Iterative Regularisierungsverfahren

Je nachdem, ob das regularisierende Filter F_γ von den aktuellen Daten $\boldsymbol{g}_\varepsilon$ abhängt oder nicht, unterscheidet man zwischen nichtlinearen und linearen iterativen Verfahren. Als einfaches Beispiel eines linearen iterativen Verfahrens betrachten wir zunächst die *Landweber–Iteration* und eine Erweiterung dieses Verfahrens nach Strand [39]. Im Anschluß daran werden die *Methode der konjugierten Gradienten* wie auch das *Chahinesche Verfahren* als Beispiele nichtlinearer Regularisierungsverfahren behandelt, wobei für letzteres eine neue Modifikation vorgeschlagen wird, die eine Anwendung dieses Verfahrens auch auf nicht quadratische Matrizen ermöglicht.

Die regularisierende Wirkung der iterativen Verfahren wurde bereits früh beobachtet [21]. Pionierarbeit leistete auf diesem Gebiet vor allen Dingen Strand [39], indem er die enge Verbindung zwischen den iterativen Verfahren und anderen Regularisierungsverfahren aufzeigte (siehe hierzu auch die Ausführungen in [36]).

2.3.1 Landweber–Iteration

Ausgehend von der Normalgleichung

$$\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{f} = \boldsymbol{A}^*\boldsymbol{g}_\varepsilon$$

erzeugt man Iterationsverfahren ganz allgemein durch Aufspalten von

$$\boldsymbol{A}^*\boldsymbol{A} = \boldsymbol{S} - \boldsymbol{T}.$$

Die Matrix $\boldsymbol{S}$ ist dabei so zu wählen, daß sie leicht invertiert werden kann [23, S. 103], [28, S. 832]. Dies eingesetzt in die Ausgangsgleichung und von links multipliziert mit $\boldsymbol{S}^{-1}$ ergibt

$$\boldsymbol{f} - \boldsymbol{S}^{-1}\boldsymbol{T}\boldsymbol{f} = \boldsymbol{S}^{-1}\boldsymbol{A}^*\boldsymbol{g}_\varepsilon,$$

woraus die Iterationsgleichung

$$\boldsymbol{f}^{m+1} = \boldsymbol{S}^{-1}(\boldsymbol{T}\boldsymbol{f}^m + \boldsymbol{A}^*\boldsymbol{g}_\varepsilon)$$

ensteht. Mit der speziellen Wahl

$$\boldsymbol{S} = \beta^{-1}\boldsymbol{I}$$

ergibt sich hieraus das von Landweber vorgeschlagene Iterationsverfahren [21], [23, S. 107], [39, S. 806], [42, S. 176]

$$\boxed{\boldsymbol{f}^{m+1} = (\boldsymbol{I} - \beta\boldsymbol{A}^*\boldsymbol{A})\boldsymbol{f}^m + \beta\boldsymbol{A}^*\boldsymbol{g}_\varepsilon = \boldsymbol{f}^m + \beta\boldsymbol{A}^*(\boldsymbol{g}_\varepsilon - \boldsymbol{A}\boldsymbol{f}^m).} \tag{20}$$

Gleichung (20) können wir auch schreiben als

$$\boldsymbol{f}^{m+1} = \sum_{k=0}^{m}(\boldsymbol{I} - \beta\boldsymbol{A}^*\boldsymbol{A})^k\beta\boldsymbol{A}^*\boldsymbol{g}_\varepsilon + (\boldsymbol{I} - \beta\boldsymbol{A}^*\boldsymbol{A})^m\boldsymbol{f}^0,$$

so daß der lineare Charakter dieses iterativen Verfahrens offenbar wird. Mit der Anfangsschätzung $\boldsymbol{f}^0 = 0$ läßt sich die Landweber–Iteration wiederum darstellen als Filterung der verallgemeinerten Inversen. Letztlich führen diese Rechnungen (siehe Anhang A.3) auf die spektrale Darstellung

$$\boldsymbol{f}^m = \sum_n \sigma_n^{-2} F_m(\sigma_n, \beta) < \boldsymbol{g}_\varepsilon | \boldsymbol{A}\boldsymbol{v}_n > \boldsymbol{v}_n$$

mit dem Filter

$$F_m(\sigma_n, \beta) = 1 - (1 - \beta\sigma_n^2)^m, \tag{21}$$

(siehe Abb. 5). Die Iterationstiefe m spielt hierbei die Rolle des Regularisierungsparameters γ, indem wir $\gamma = 1/m$ setzen. Wird nur kurz iteriert, d. h. ist γ relativ groß, wird stark regularisiert. Entsprechend groß ist auch der entsprechende Regularisierungsfehler, vgl. Abb. 1 in Abschn. 1.4. Mit jedem zusätzlichen Iterationsschritt

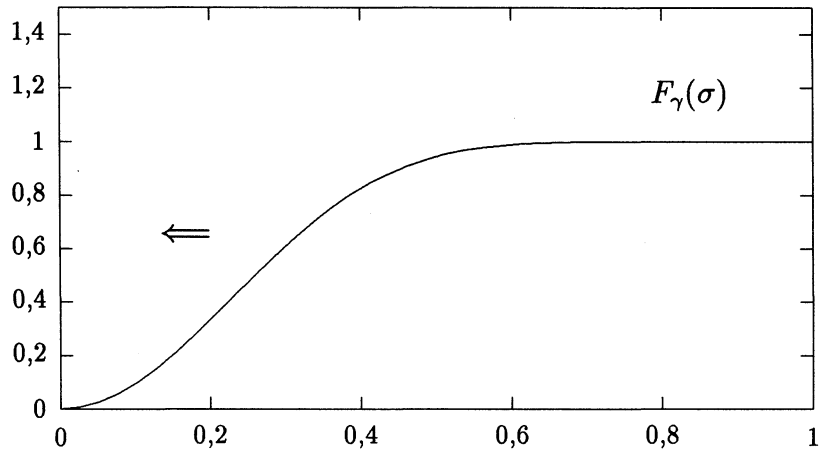

Abbildung 5: Filterwirkung der Landweber–Iteration mit einem Regularisierungsparameter $\gamma = \frac{1}{m} = \frac{1}{10}$ und einem Parameter $\beta = 1$. Mit zunehmender Iterationstiefe nimmt die regularisierende Wirkung des Verfahrens ab, angedeutet durch einen Pfeil.

wird jedoch der Regularisierungsfehler weiter reduziert, so daß mit zunehmender Iterationstiefe das Verfahren zu konvergieren scheint. Dann, für großes m bzw. kleines γ wächst der Datenfehler auf einmal stark an, so daß das Verfahren letztlich divergiert. Dieses *Semikonvergenzverhalten* ist typisch für alle iterativen Verfahren. Die Schwierigkeit der iterativen Verfahren besteht nun darin, die Iterationstiefe geeignet zu wählen, wobei die Wahl meist a posteriori, wie in Abschn. 1.4 beschrieben, erfolgt.

Der Parameter β in Gl. (20) ist ein weiterer Steuerparameter und dient der Festlegung der Iterationsschrittweite. Er darf nicht völlig frei gewählt werden, da ein Filter per Definition immer nur Werte im Bereich $0 \leq F_m(\sigma_n, \beta) \leq 1$ annehmen darf. Diese Bedingung muß für alle Eigenwerte, insbesondere für den größten Eigenwert, σ_1, erfüllt sein. Aus (21) folgt damit,

$$0 \leq 1 - (1 - \beta\sigma_1^2)^m \leq 1 \qquad \text{bzw.} \qquad 0 \leq \beta \leq \sigma_1^{-2}. \tag{22}$$

Hieraus ergibt sich die interessante Konsequenz, daß die Landweber–Iteration in einigen Fällen recht ineffizient arbeitet, und zwar dann, wenn einerseits nur schwach regularisiert werden soll, was einer großen Iterationstiefe entspricht, und wenn andererseits der größte auftretende Eigenwert σ_1 recht groß ist. In diesem Fall muß nämlich aufgrund von (22) die Iterationsschrittweite relativ klein gewählt werden, so daß im Anschluß sehr viele Iterationen notwendig sind, um aus dem Bereich der stark regularisierten Lösungen in den Bereich der schwach regularisierten Lösungen zu gelangen (siehe [39, S. 815]). Dies gilt insbesondere für die exponentiell schlecht gestellten Probleme.

2.3.2 Erweiterung der Landweber–Iteration nach Strand

Strand [39] schlägt deshalb vor, die Landweber–Iteration zu modifizieren und β in (20) zu ersetzen durch einen Operator $\boldsymbol{D}$, der auf dem gleichen System von Eigenvektoren basiert wie die Matrix $\boldsymbol{A}^*\boldsymbol{A}$, also

$$\boxed{\boldsymbol{f}^{m+1} = \boldsymbol{f}^m + \boldsymbol{D}\boldsymbol{A}^*(\boldsymbol{g}_\varepsilon - \boldsymbol{A}\boldsymbol{f}^m), \qquad \text{mit} \qquad \boldsymbol{D}\boldsymbol{v}_n = \delta_n^2 \boldsymbol{v}_n.}$$

Strand untersucht nun eine ganze Reihe von Operatoren $\boldsymbol{D}$, wobei er obige Bedingung an $\boldsymbol{D}$ trivialerweise erfüllt, indem er $\boldsymbol{D} = \mathcal{F}(\boldsymbol{A}^*\boldsymbol{A})$ wählt; hierbei ist $\mathcal{F}(\sigma_n^2)$ ein Polynom oder rationaler Bruch. Zu erwähnen ist besonders die spezielle Wahl

$$\boldsymbol{D} = (\boldsymbol{A}^*\boldsymbol{A} + \omega^2 \boldsymbol{I})^{-1}.$$

Hieraus folgt für die Eigenwerte $\delta_n^2 = (\sigma_n^2 + \omega^2)^{-1}$ und damit für das Filter

$$F_m(\sigma_n) = 1 - \left(1 - \frac{\sigma_n^2}{\sigma_n^2 + \omega^2}\right)^m = 1 - \left(\frac{\omega^2}{\sigma_n^2 + \omega^2}\right)^m.$$

In diesem Falle ergibt also die erste Iteration $m = 1$ mit einem Startwert $f_0 = 0$ exakt die gleiche Lösung wie die Tikhonov–Phillips–Regularisierung (vgl. Gl. (16) bzw. (17)) mit einem Regularisierungsparameter ω. Besser kann die enge Verbindung zwischen den iterativen und anderen Regularisierungsverfahren nicht demonstriert werden.

Unabhängig von der aktuellen Wahl des zusätzlichen Parameters ω, konvergiert die Iteration immer zur nicht regularisierten Lösung (4); aufgrund des Semikonvergenzverhaltens darf wiederum nicht zu tief iteriert werden. Für gegebenes ω und m kann der effektive Regularisierungsparameter γ abgeschätzt werden, indem man $F_m(\gamma) = \frac{1}{2}$ setzt und nach γ auflöst. Dies liefert

$$\gamma = (2^{1/m} - 1)^{1/2}\omega.$$

Für $m = 1$ folgt hieraus $\gamma = \omega$. Soll nun ein bestimmter Regularisierungsparameter γ eingestellt werden oder soll die Parameterwahl a posteriori, wie in Abschn. 1.4 beschrieben erfolgen, so ist es zweckmäßig, zunächst ein ω etwas größer als der gewünschte effektive Regularisierungsparameter γ zu wählen und anschließend solange zu iterieren bis dieser erreicht wird, bzw. der Defekt in den Daten den angenommenen Meßfehler unterschreitet. Dieses *iterative Tikhonov–Phillips–Verfahren* stellt also eine besonders elegante Möglichkeit dar, eine Tikhonov–Phillips–Regularisierung mit anschließender a posteriori Parameterwahl effizient durchzuführen.

2.3.3 Verfahren der konjugierten Gradienten

Ziel des konjugierten Gradienten–Verfahrens [23, S. 112] [28, S. 414-425] ist die Minimierung des Funktionals

$$J(\boldsymbol{f}) = ||\boldsymbol{A}\boldsymbol{f} - \boldsymbol{g}_\varepsilon||^2,$$

wobei von der Tatsache Gebrauch gemacht wird, daß der Gradient dieses Funktionals (vgl. Rechnung im Anhang A.1 für den Abschn. 1.2) bekannt ist,

$$\text{grad}\, J(\boldsymbol{f}) = 2\boldsymbol{A}^*(\boldsymbol{A}\boldsymbol{f} - \boldsymbol{g}_\varepsilon) \stackrel{\text{def}}{=} 2\boldsymbol{r}. \tag{23}$$

Es erscheint zweckmäßig, ausgehend von einem beliebigen Startwert $\boldsymbol{f}^0$, das Funktional $J(\boldsymbol{f})$ jeweils in Richtung des steilsten Abfalls, d. h. in Richtung $\boldsymbol{d} = -\boldsymbol{r}$ zu minimieren und auf diese Weise iterativ das richtige $\boldsymbol{f}$ zu ermitteln. Führen wir diese Minimierung durch, so folgt wegen [1]

$$J(\boldsymbol{f} + \alpha\boldsymbol{d}) = J(\boldsymbol{f}) + 2\alpha(\boldsymbol{A}\boldsymbol{f} - \boldsymbol{g}_\varepsilon)^*\boldsymbol{A}\boldsymbol{d} + \alpha^2||\boldsymbol{A}\boldsymbol{d}||^2$$

und damit

$$\frac{dJ(\boldsymbol{f} + \alpha\boldsymbol{d})}{d\alpha} = 2(\boldsymbol{A}\boldsymbol{f} - \boldsymbol{g}_\varepsilon)^*\boldsymbol{A}\boldsymbol{d} + 2\alpha||\boldsymbol{A}\boldsymbol{d}||^2$$

für das Minimum in Richtung $\boldsymbol{d}$

$$\alpha = -\frac{(\boldsymbol{A}^*(\boldsymbol{A}\boldsymbol{f} - \boldsymbol{g}_\varepsilon))^*\,\boldsymbol{d}}{||\boldsymbol{A}\boldsymbol{d}||^2} = -\frac{\boldsymbol{r}^*\boldsymbol{d}}{||\boldsymbol{A}\boldsymbol{d}||^2}. \tag{24}$$

Hieraus ergibt sich die Iterationsfolge

$$\boxed{\boldsymbol{f}^{m+1} = \boldsymbol{f}^m + \alpha^m\boldsymbol{d}^m,} \tag{25}$$

wobei immer gilt $\boldsymbol{d}^m = -\boldsymbol{r}^m$ (Richtung des steilsten Abfalls). Diese ad hoc Vorgehensweise ist jedoch wenig effizient, da die Lösung in einem aufwendigen Zickzackkurs angestrebt wird. Dies folgt unmittelbar aus der Tatsache, daß, nachdem das Funktional in Richtung des alten steilsten negativen Gradienten minimiert wurde, der neue Gradient jeweils senkrecht auf dem vorhergehenden steht.

Beim Verfahren der konjugierten Gradienten wird deshalb versucht diesen Zickzackkurs zu „glätten“, indem neben der Richtung des neuen steilsten Abstieges die alte Fortschreitungsrichtung nahezu beibehalten und nur etwas in Richtung des neuen steilsten Abstieges gebeugt wird (konjugieren = beugen). Lediglich im ersten Iterationsschritt entspricht die Ausbreitungsrichtung exakt der Richtung des steilsten Abstieges, $\boldsymbol{d}^0 = -\boldsymbol{r}^0$; bei allen nachfolgenden Schritten berechnet sich die neue Fortschreitungsrichtung hingegen zu

$$\boldsymbol{d}^{m+1} = -\boldsymbol{r}^{m+1} + \beta^m\boldsymbol{d}^m. \tag{26}$$

Die Größe β^m ist ein empirisch zu bestimmender Wichtungsfaktor, der angibt, wie stark die alte Fortschreitungsrichtung gegenüber der Richtung des neuen steilsten

[1] Entwicklung dieser mehrdimensionalen Funktion in eine Potenzreihe um $\boldsymbol{f}$. Glieder höherer als zweiter Ordnung in α sind allesamt identisch 0.

Abstieges gewichtet wird. Am weitesten verbreitet ist die Methode nach Fletcher-Reeves [28, S. 422],

$$\beta^m = \frac{||\boldsymbol{r}^{m+1}||^2}{||\boldsymbol{r}^m||^2}, \tag{27}$$

die sich vor allen Dingen durch ihren Erfolg rechtfertigt. In der endgültigen Version des Verfahrens wollen wir noch den Faktor α^m in Gl. (24) etwas weiter vereinfachen, indem wir ausnutzen daß $\boldsymbol{r}^m$ und $\boldsymbol{d}^{m-1}$ aufeinander senkrecht stehen,

$$\alpha^m = \frac{\boldsymbol{r}^{m*}\boldsymbol{d}^m}{||\boldsymbol{A}\boldsymbol{d}^m||^2} = \frac{\boldsymbol{r}^{m*}(-\boldsymbol{r}^m + \beta^{m-1}\boldsymbol{d}^{m-1})}{||\boldsymbol{A}\boldsymbol{d}^m||^2} = \frac{||\boldsymbol{r}^m||^2}{||\boldsymbol{A}\boldsymbol{d}^m||^2}, \tag{28}$$

so daß mit den Gln. (23), (25), (26), (27) sowie (28) der Algorithmus nun vollständig ist.

Auch das Verfahren der konjugierten Gradienten kann als Filterung der verallgemeinerten Inversen interpretiert werden. Nach m Iterationen läßt sich die Lösung darstellen als [23, S. 115 Gl. (4.3.5)]

$$\boldsymbol{f}^m = \boldsymbol{f}^0 - P_{m-1}(\boldsymbol{A}^*\boldsymbol{A})\boldsymbol{r}^0.$$

$P_{m-1}(\boldsymbol{A}^*\boldsymbol{A})$ ist hierbei ein Polynom $m-1$.ten Grades in $\boldsymbol{A}^*\boldsymbol{A}$. Mit der Singulärwertzerlegung des Operators $\boldsymbol{A}^*\boldsymbol{A}$ folgt daraus

$$\boldsymbol{f}^m = \sum_n < \boldsymbol{f}^m | \boldsymbol{v}_n > \boldsymbol{v}_n = \sum_n < \boldsymbol{f}^0 - P_{m-1}(\boldsymbol{A}^*\boldsymbol{A})\boldsymbol{r}^0 | \boldsymbol{v}_n > \boldsymbol{v}_n.$$

Für die spezielle Wahl $\boldsymbol{f}^0 = 0$, so daß $\boldsymbol{r}^0 = -\boldsymbol{A}^*\boldsymbol{g}_\mathcal{E}$ (siehe Gl. (23)) erhalten wir

$$\begin{aligned}
\boldsymbol{f}^m &= \sum_n < P_{m-1}(\boldsymbol{A}^*\boldsymbol{A})\boldsymbol{A}^*\boldsymbol{g}_\mathcal{E} | \boldsymbol{v}_n > \boldsymbol{v}_n \\
&= \sum_n < \boldsymbol{A}^*\boldsymbol{g}_\mathcal{E} | P_{m-1}(\sigma_n^2)\boldsymbol{v}_n > \boldsymbol{v}_n \\
&= \sum_n \sigma_n^{-2} F_m(\sigma_n) < \boldsymbol{g}_\mathcal{E} | \boldsymbol{A}\boldsymbol{v}_n > \boldsymbol{v}_n
\end{aligned}$$

mit einem Filter

$$F_m(\sigma_n) = \sigma_n^2 P_{m-1}(\sigma_n^2). \tag{29}$$

Sowohl das Filter als auch das Verfahren sind nichtlinear, da das Polynom P_{m-1} über die Faktoren α (27) und β (28) von $\boldsymbol{r}$ und damit von den Daten $\boldsymbol{g}_\mathcal{E}$ abhängt. Aufgrund dessen kann auch die in Abschn. 1.6 geführte Betrachtung hinsichtlich Auflösung und Fehler bei diesem Verfahren nicht durchgeführt werden. Der Regularisierungsparameter γ ist wiederum, wie bei der Landweber–Iteration, gegeben durch die Iterationstiefe, $\gamma = \frac{1}{m}$, so daß wir wieder nicht zu tief iterieren dürfen, sollen die Beiträge kleiner Eigenwerte unterdrückt bleiben. Als Abbruchkriterium eignet sich hierbei insbesondere die in Abschn. 1.4 beschriebene a posteriori Parameterwahl.

2.3.4 Chahinesches Verfahren

Eine gewisse Sonderstellung unter den iterativen Verfahren nimmt das Chahinesche Verfahren ein, vor allem bedingt durch einen eher pragmatischen Zugang zur Iterationsgleichung [6, 7, 12, 13], [16, S. 266 ff.], [29, 31, 41]. Um dieses Verfahren anwenden zu können, muß die Matrix $\boldsymbol{A}$ zwei wichtige Voraussetzungen erfüllen. Sie muß a) quadratisch sein, eine Einschränkung, die auf die bisher besprochenen Regularisierungsverfahren nicht zutrifft, und sie muß b) eine Bandmatrix sein mit einem eindeutigen Maximum längs ihrer Diagonalen. [2] Gleichung (2) können wir dann unter Vernachlässigung etwaiger Meßfehler auch schreiben als

$$\begin{aligned} g_i^{mes} &= \sum_j A_{ij} f_j \qquad i = 1, \ldots, n;\ j = 1, \ldots, n; \\ &\approx c A_{ii} f_i, \end{aligned} \tag{30}$$

mit c als einer unbekannten Konstanten. Diese Approximation kann man über den Mittelwertsatz der Integralrechnung rechtfertigen oder einfach als physikalisch sinnvoll erachten aufgrund der Dominanz der Diagonalelemente in $\boldsymbol{A}$. Ausgehend von einer Anfangsschätzung f_i^0, erhalten wir dann

$$g_i^0 = \sum_j A_{ij} f_j^0 \approx c A_{ii} f_i^0. \tag{31}$$

Im allgemeinen ist $f_i^0 \neq f_i$, da g_i^0 nicht den gemessenen Daten g_i^{mes} entspricht. Wir können jedoch die Gleichungen (30) und (31) zueinander ins Verhältnis setzen und so eine verbesserte Lösung erhalten,

$$\frac{f_i^1}{f_i^0} = \frac{g_i^{mes}}{g_i^0},$$

woraus die Iterationsgleichung

$$f_i^{m+1} = \frac{g_i^{mes}}{g_i^m} f_i^m$$

[2] Bei der Vertikalsondierung der Atmosphäre vom Satelliten aus ist die Vorraussetzung b) automatisch erfüllt, da hier die *Gewichtsfunktionen*, gegeben durch die Zeilen der Matrix $\boldsymbol{A}$, aufgrund der Beschaffenheit der Atmosphäre bereits von vornherein ein solches Maximum aufweisen. Bei entsprechenden Messungen vom Boden aus besteht hingegen die Möglichkeit, die einzelnen Zeilen der Matrix $\boldsymbol{A}$ und entsprechend auch die einzelnen Elemente des Datenvektors $\boldsymbol{g}^{mes}$ voneinander zu subtrahieren, $A'_{ij} = A_{ij} - A_{i+1\,j}$ sowie $g_i^{mes\prime} = g_i^{mes} - g_{i+1}^{mes}$, wodurch das „gestrichene" Gleichungssystem

$$\boldsymbol{g}^{mes\prime} = \boldsymbol{A}'\boldsymbol{f}$$

entsteht. Auf dieses Gleichungssystem, das nun seinerseits quadratisch sein muß, kann das Chahinesche Verfahren angewendet werden [29]. Allerdings ist der Informationsgehalt dieses neuen Gleichungssystems wesentlich geringer, da durch die Subtraktion benachbarter Elemente des Datenvektors $\boldsymbol{g}^{mes}$ sich das "Signal zu Rausch Verhältnis" erheblich verschlechtert.

entsteht. In den meisten Fällen konvergiert dieses einfach zu programmierende Verfahren recht schnell, zeigt jedoch das für alle iterativen Verfahren typische Semikonvergenzverhalten, indem die regularisierende Wirkung des Verfahrens mit wachsender Iterationstiefe abnimmt.

Die regularisierende Wirkung des Chahineschen Verfahrens bei geringer Iterationstiefe ist unmittelbar einzusehen. Wie wir bereits im Abschn. 1.5 festgestellt hatten, zeichnen sich Regularisierungsverfahren dadurch aus, daß die aufgrund von Meßfehlern stark oszillierenden Anteile der unregularisierten Lösung in der regularisierten Lösung unterdrückt werden. Dies erfolgt dadurch, daß über benachbarte Elemente der unregularisierten Lösung lokal gemittelt wird. Betrachten wir die Approximation (30) bzw. (31) eingehender, so kann diese ebenfalls als lokale Mittelung aufgefaßt werden, wobei diese Mittelung sich über die zugehörige Bandbreite der Matrix $\boldsymbol{A}$, d. h. über die vertikale Breite der entsprechenden Gewichtsfunktionen erstreckt. Nach dem ersten Iterationsschritt ist die Auflösunglösungsmatrix $\boldsymbol{R}$ (siehe Abschn. 1.6) daher im wesentlichen gleich der Matrix $\boldsymbol{A}$. Mit zunehmender Iterationstiefe nimmt die regularisierende Wirkung des Chahineschen Verfahrens zusehends ab, so daß die Auflösung sich zwar erhöht, andererseits Meßfehler sich aber immer stärker auf die Lösung auswirken. Es hängt nun von der Größe des Meßfehlers ab, bei welcher Iterationstiefe m, bzw. bei welchem Regularisierungsparameter $\gamma = 1/m$ das Verfahren abgebrochen werden sollte, wobei die Parameterwahl i. d. Regel a posteriori erfolgt (siehe Abschn. 1.4).

2.3.5 Eigene Modifikation des Chahineschen Verfahrens

Neben seinen beiden größten Vorzügen, der leichten Programmierbarkeit und der schnellen Konvergenz, ist das Chahinesche Verfahrens auch mit Nachteilen behaftet. Hierzu gehört zum einen die Tatsache, daß das Chahinesche Verfahren als nichtlineares Regularisierungsverfahren keine unmittelbaren Aussagen über Fehler und erreichbare Auflösung erlaubt, und zum anderen, daß die Matrix $\boldsymbol{A}$ quadratisch und von bandartiger Struktur sein muß. Von den beiden letztgenannten Einschränkungen können wir zumindest die Forderung nach einer quadratischen Form der Matrix $\boldsymbol{A}$ fallenlassen, wenn wir in Analogie zu den bisher betrachteten Verfahren statt des Problems

$$\boldsymbol{g}^{mes} = \boldsymbol{A}\boldsymbol{f}$$

das verallgemeinerte Problem (vgl. Gl. (3) in Abschn. 1.2)

$$\boldsymbol{A}^*\boldsymbol{g}^{mes} = \boldsymbol{A}^*\boldsymbol{A}\boldsymbol{f}$$

betrachten und auf dieses das Chahinesche Verfahren anwenden. Dieser Schritt ist möglich, da, sofern die Matrix $\boldsymbol{A}$ von bandartiger Strukur ist, auch die Matrix $\boldsymbol{A}^*\boldsymbol{A}$ eine solche Bandstruktur aufweist.

Gegenüber dem Chahineschen Verfahren zeichnet sich dieses modifizierte Chahinesche Verfahren dadurch aus, daß es zwar etwas langsamer konvergiert, andererseits jedoch in den betrachteten Fällen deutlich bessere Inversionsergebnisse liefert, weshalb im sich anschließenden Kapitel 3 über die Anwendung der Regularisierungsverfahren nur das modifizierte Chahinesche Verfahren behandelt wird. Überdies kann aufgrund dieser Modifikation das Chahinesche Verfahren nun auch auf Probleme anwendet werden, bei denen die Matrix $\boldsymbol{A}$ nicht quadratisch ist, und ist insofern allgemeiner anwendbar.

2.4 Stochastische Verfahren

Bei allen bisher besprochenen Verfahren stellt sich jeweils die Frage nach der Wahl des richtigen Regularisierungsparameters γ. Diese erfolgt entweder a posteriori – indem von einer stark regularisierten Lösung ausgehend der Regularisierungsparameter immer kleiner gewählt wird, bis der Fehler in den Daten eine vorgegebene Schranke unterschreitet (siehe Abschn. 1.4) – oder a priori – aufgrund von Erfahrungswerten. Wie bereits in Abschn. 1.4 angedeutet, liefert die a posteriori Parameterwahl keinesfalls das optimale γ sondern immer etwas zu große Werte, so daß die Lösung zwar leicht überregularisiert, andererseits jedoch mit Sicherheit nur wenig mit Datenfehlern behaftet ist. Im Gegensatz dazu ist es bei der a priori Parameterwahl zwar möglich, genau das optimale γ zu treffen, andererseits besteht die Gefahr, daß der Regularisierungsparameter erheblich zu groß oder zu klein gewählt wird und infolgedessen die Lösung $\boldsymbol{f}_\gamma$ stark fehlerbehaftet ist.

Die stochastischen Verfahren stellen den Versuch dar, aufgrund zahlreicher, früherer Inversionsergebnisse, aufgrund der Meßergebnisse anderer Meßverfahren oder einfach aufgrund von Erfahrungswerten – Rodgers [31] spricht in diesem Zusammenhang gar von *virtuellen Messungen* – diesen optimalen Wert des Regularisierungsparameters anhand wahrscheinlichkeitstheoretischer Überlegungen zu ermitteln. Man verspricht sich hiervon ein Ergebnis, das dem Urbild näher kommt als bei den bisher besprochenen Verfahren, bei denen die Wahl des Regularisierungsparameters im Prinzip dem Anwender selbst obliegt. Doch auch die stochastisch motivierte a priori Parameterwahl schließt die Möglichkeit einer zu starken oder zu schwachen Regularisierung nicht aus, nämlich dann, wenn die aktuellen Daten stark vom bisherigen Mittelwert abweichen.

Für einige Verwirrung beim Studium der einschlägigen Literatur sorgt vor allem die Tatsache, daß es sich bei den stochastischen Verfahren nicht um verschiedene Regularisierungsverfahren an sich, sondern lediglich um verschiedene Ableitungen ein und desselben Regularisierungsverfahrens handelt [45]. Das im folgenden vorgestellte *Verfahren der maximalen Wahrscheinlichkeit* unterscheidet sich also im wesentlichen nicht von der sogenannten *Bayes-Schätzung* [23, S. 134 ff.], auch *optimum estimation method* [19, 30, 31, 32, 33] oder *Rodgers Verfahren* genannt; und auch das als *bester*

linearer Schätzer [23, S. 130 ff.] oder *minimum variance method* [14, 15, 20, 43] bekannte Verfahren führt, gaußsche Statistik vorausgesetzt, letztlich wieder auf ein und diesselbe Inversionsgleichung (37) bzw. (38). Es ist insofern wenig sinnvoll alle drei Zugänge zu dieser Inversionsgleichung im Detail zu behandeln, so daß wir auf die beiden letztgenannten stochastischen Verfahren nicht näher eingehen werden.

2.4.1 Erwartungswert und Kovarianzmatrix

Unabdingbare Voraussetzung für die Anwendung der stochastischen Verfahren ist zunächst, daß bereits ausreichend a priori Information zur Verfügung steht, um überhaupt Stochastik betreiben zu können. Bezeichnen wir mit $\overline{\boldsymbol{f}}$ den Mittel- oder Erwartungswert eines Ensembles von früheren Lösungen $\boldsymbol{f}$, $\overline{\boldsymbol{f}} = E(\boldsymbol{f})$, so berechnet sich die Kovarianzmatrix $\boldsymbol{\mathcal{S}}_f$ für diese Verteilung von Zufallsvektoren zu

$$\boldsymbol{\mathcal{S}}_f = E\left((\boldsymbol{f} - \overline{\boldsymbol{f}})(\boldsymbol{f} - \overline{\boldsymbol{f}})^*\right). \tag{32}$$

Des weiteren gehen wir davon aus, daß wir die stochastischen Kenngrößen des Meßfehlervektors $\boldsymbol{\mathcal{E}}$ der aktuellen Messung in Form seines Erwartungswertes [3] $\overline{\boldsymbol{\mathcal{E}}} = 0$ sowie in Form seiner Kovarianzmatrix $\boldsymbol{\mathcal{S}}_{\mathcal{E}} = E(\boldsymbol{\mathcal{E}}\boldsymbol{\mathcal{E}}^*)$ kennen, wenn auch der Meßfehlervektor $\boldsymbol{\mathcal{E}}$ selbst nicht bekannt ist.

Die Kovarianzmatrix stellt die Erweiterung der vom Fall einer eindimensionalen Verteilung her bekannten Varianz auf den Fall einer Verteilung von Zufallsvektoren dar und ergibt sich direkt aus der Definition der zweiten zentralen Momente (siehe [5, S. 671], [24, S. 21 ff.] und [26, S. 82]). In den Diagonalelementen der Kovarianzmatrix steckt jeweils die Varianz der entsprechenden Elemente von Zufallsvektoren; sie geben an, inwieweit die einzelnen Elemente der Zufallsvektoren streuen. Dagegen sind die Matrixelemente außerhalb der Diagonalen ein Maß für die Korrelation der verschiedenen Elemente von Zufallsvektoren untereinander. Sind z. B. die ersten und zweiten Elemente einer Verteilung von Zufallsvektoren stark positiv miteinander korreliert, d. h. ist das erste Element eines Zufallsvektors immer dann groß, wenn auch das zweite Element groß ist und umgekehrt, so sind die entsprechenden Matrixelemente $(1,2)$ und $(2,1)$ der zugehörigen Kovarianzmatrix positiv und relativ groß.

Liegt statt dessen keine Korrelation vor, so ist die Kovarianzmatrix eine reine Diagonalmatrix. Diese Annahme ist z. B. für die Kovarianzmatrix $\boldsymbol{\mathcal{S}}_{\mathcal{E}}$ im Falle des in [18] verwendeten Radiometers weitgehend gerechtfertigt. Geht man des weiteren davon aus, daß alle Elemente des Meßfehlervektors gleich stark streuen, so ist $\boldsymbol{\mathcal{S}}_{\mathcal{E}} = \sigma_{\mathcal{E}}^2 \boldsymbol{I}$ mit $\sigma_{\mathcal{E}}^2$ der Varianz und $\sigma_{\mathcal{E}}$ der Standardabweichung [4] eines jeden Elements des Meßfehlervektors.

[3] Wir gehen davon aus, daß es sich bei den Meßfehlern um rein statistische Fehler handelt (Rauschen), so daß der Erwartungswert für den Meßfehlervektor $\boldsymbol{\mathcal{E}}$ gleich 0 ist.

[4] Im wesentlichen hängt die Größe der Standardabweichung ab von der Integrationszeit, über die das aktuelle Spektrum aufgenommen wird. Es gilt die Radiometerformel, nach der die Standardabweichung der Wurzel aus der Integrationszeit umgekehrt proportional ist [18].

2.4.2 Verfahren der maximalen Wahrscheinlichkeit

Diesem im englischen Sprachgebrauch als *maximum likelihood method* bezeichneten Verfahren (siehe [45, S. 153], etwas ausführlicher beschrieben in [24, S. 79 ff.], sehr gut dargestellt in [26], sowie [38]) liegen die Annahmen zugrunde, daß

a) die Lösungen $\boldsymbol{f}$ und die Meßfehler $\boldsymbol{\mathcal{E}}$ durch eine Gaußverteilung beschrieben werden können [5] und daß

b) die Lösungen $\boldsymbol{f}$ und Meßfehler $\boldsymbol{\mathcal{E}}$ nicht miteinander korreliert, also unabhängig voneinander sind.

Aus der ersten Annahme folgt für die entsprechenden Wahrscheinlichkeitsdichten unmittelbar

$$p_f(\boldsymbol{f}) = \frac{1}{2\pi^{m/2}|\boldsymbol{S}_f|} e^{-\frac{1}{2}(\boldsymbol{f}-\overline{\boldsymbol{f}})^* \boldsymbol{S}_f^{-1}(\boldsymbol{f}-\overline{\boldsymbol{f}})} \tag{33}$$

bzw.

$$p_{\mathcal{E}}(\boldsymbol{\mathcal{E}}) = \frac{1}{2\pi^{n/2}|\boldsymbol{S}_{\mathcal{E}}|} e^{-\frac{1}{2}\boldsymbol{\mathcal{E}}^* \boldsymbol{S}_{\mathcal{E}}^{-1}\boldsymbol{\mathcal{E}}}. \tag{34}$$

Aus der Annahme b) folgt, daß die zusammengesetzte Wahrscheinlichkeitsdichte $p_{f,\mathcal{E}}(\boldsymbol{f},\boldsymbol{\mathcal{E}})$, d. h. die Wahrscheinlichkeitsdichte für das gleichzeitige Auftreten einer Lösung $\boldsymbol{f}$ mit einem zugehörigen Meßfehler $\boldsymbol{\mathcal{E}}$, sich einfach ergibt als das Produkt der beiden Einzelwahrscheinlichkeitsdichten $p_f(\boldsymbol{f})$ und $p_{\mathcal{E}}(\boldsymbol{\mathcal{E}})$,

$$p_{f,\mathcal{E}}(\boldsymbol{f},\boldsymbol{\mathcal{E}}) = p_f(\boldsymbol{f})\, p_{\mathcal{E}}(\boldsymbol{\mathcal{E}}). \tag{35}$$

In Abb. 6 ist der Verlauf der zusammengesetzten Wahrscheinlichkeitsdichte durch Höhenlinien skizziert.

Bisher haben wir noch keinen Gebrauch gemacht von unserem Modell, $\boldsymbol{A}\boldsymbol{f} = \boldsymbol{g}$, sowie von unserer aktuellen, wenn auch fehlerbehafteten Messung $\boldsymbol{g}_{\mathcal{E}}$, sondern lediglich Stochastik aufgrund der vorhandenen a priori Information betrieben. Wie bereits bekannt, liefert unser Modell innerhalb der Meßfehlergrenzen keine eindeutige Lösung (siehe Abschn. 1.4). Dies ist in in Abb. 6 angedeutet durch eine durchgezogene Linie, wobei jeder Punkt der Linie eine aufgrund des Modells mögliche Lösung darstellt. Aus dieser Schar der möglichen Lösungen bestimmen wir die aufgrund von Gl. (35) wahrscheinlichste Lösung $\boldsymbol{f}_{opt}$. Auf mathematischem Wege können wir dies erreichen, indem wir $\boldsymbol{\mathcal{E}}$ in (35) ersetzen durch

$$\boldsymbol{\mathcal{E}} = \boldsymbol{g}_{\mathcal{E}} - \boldsymbol{g} = \boldsymbol{g}_{\mathcal{E}} - \boldsymbol{A}\boldsymbol{f} = \boldsymbol{\mathcal{E}}(\boldsymbol{f}),$$

[5] Eine Annahme, die in unserem Fall, in dem die Lösungen $\boldsymbol{f}$ extrem variable klimatologische Verhältnisse repräsentieren nur unter bestimmten Umständen zu rechtfertigen ist.

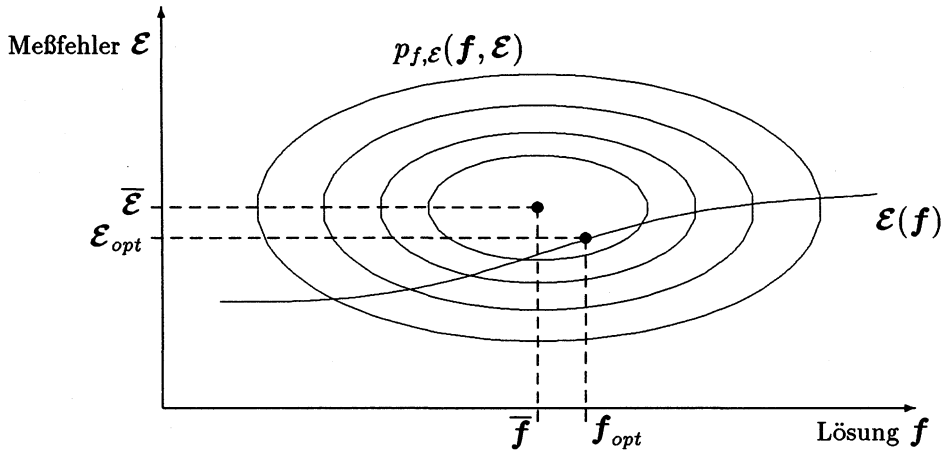

Abbildung 6: Das Zustandekommen der optimalen Näherungslösung f_{opt} durch Kombination von aktueller Messung $\mathcal{E}(f)$ und a priori Information $p_{f,\mathcal{E}}(f,\mathcal{E})$.

und anschließend das Maximum f_{opt} dieses dann nur noch von f abhängigen Ausdruckes ermitteln.

$$\frac{d\,p_{f,\mathcal{E}}(f, g_{\mathcal{E}} - Af)}{df_k} = 0.$$

Da f und $\mathcal{E}$ in (33) und (34) jeweils mit negativem Vorzeichen im Exponenten stehen, führt dieses Extremwertproblem auf die Minimierung des Ausdrucks

$$(g_{\mathcal{E}} - Af)^* \mathcal{S}_{\mathcal{E}}^{-1} (g_{\mathcal{E}} - Af) + (f - \overline{f})^* \mathcal{S}_f^{-1} (f - \overline{f}) \tag{36}$$

und damit auf die Lösung (siehe Anhang A.4)

$$\boxed{f_{opt} = \left(A^* \mathcal{S}_{\mathcal{E}}^{-1} A + \mathcal{S}_f^{-1}\right)^{-1} \left(A^* \mathcal{S}_{\mathcal{E}}^{-1} g_{\mathcal{E}} + \mathcal{S}_f^{-1} \overline{f}\right).} \tag{37}$$

Identische Gleichungen ergeben sich, gaußsche Statistik vorausgesetzt, auch für die beiden anderen, hier nicht betrachteten, stochastischen Verfahren.

Auffallend ist die enge Verwandtschaft der Lösung (37) mit der Lösung der Tikhonov–Phillips–Regularisierung, vgl. (16) bzw. (19) in Abschn. 2.2, doch sind die stochastischen Verfahren viel allgemeiner. Die Kovarianzmatrix $\mathcal{S}_{\mathcal{E}}$ spielt hierbei die Rolle des Regularisierungsparameters γ und die Matrix $\mathcal{S}_f$ die der Zusatzinformationen B, vgl. auch (36) mit den entsprechenden Ausdrücken (15) bzw. (18). Ist z. B. $\mathcal{S}_{\mathcal{E}}$ klein gegenüber $\mathcal{S}_f$, d. h. wurde die aktuelle Messung mit großer Präzision durchgeführt, so dominiert der erste Term im zu minimierenden Ausdruck (36) gegenüber dem zweiten und wir erhalten eine schwach regularisierte Lösung. Im Idealfall, wenn gar keine Meßfehler auftreten, wird $g_{\mathcal{E}} \to g$ und $\mathcal{S}_{\mathcal{E}} \to 0$, so daß (36) übergeht in $||Af - g||^2$. Als Lösung ergibt sich dann die verallgemeinerte oder Moore–Penrose–Lösung (4), die wir bereits in Abschn. 1.2 für den Fall nicht fehlerbehafteter Daten

kennengelernt hatten. Im umgekehrten Fall, wenn die Meßfehler sehr groß sind und der Informationsgehalt der Messung entsprechend gering ist, geht (36) über in $||\boldsymbol{f} - \overline{\boldsymbol{f}}||^2$. Als Lösung ergibt sich demnach der Erwartungswert $\overline{\boldsymbol{f}}$, d. h. wir erhalten lediglich unsere a priori Information wieder zurück, da durch die ungenaue Messung keine neue Information hinzugefügt wurde.

2.4.3 Praktische Inversionsgleichung

Für praktische Anwendungen ist die Inversionsgleichung (37) in ihrer obigen Form weniger geeignet, da zu ihrer Auswertung mehrere relativ große Matrizen invertiert werden müssen. Rodgers [31] schlägt deshalb vor, die Gleichung (37) identisch umzuformen (siehe Anhang A.5), so daß nur noch eine relativ kleine Matrix invertiert werden muß [6]

$$\boxed{\boldsymbol{f}_{opt} = \overline{\boldsymbol{f}} + \underbrace{\boldsymbol{S}_f \boldsymbol{A}^* (\boldsymbol{A}\boldsymbol{S}_f\boldsymbol{A}^* + \boldsymbol{S}_\varepsilon)^{-1}}_{\boldsymbol{A}_\gamma^{-A}} (\boldsymbol{g}_\varepsilon - \boldsymbol{A}\overline{\boldsymbol{f}}).} \tag{38}$$

Hierbei identifizieren wir die Matrix $\boldsymbol{A}_\gamma^{-A}$ als regularisierte verallgemeinerte Inverse (siehe Abschn. 1.3 ff.), die wir zur Berechnung der Auflösungsmatrix $\boldsymbol{R} = \boldsymbol{A}_\gamma^{-A}\boldsymbol{A}$ und der Kovarianzmatrix $\boldsymbol{S}_\gamma$ der regularisierten Lösung in Abschn. 1.6 benötigten. Die Lösung (38) läßt sich dann auch schreiben als

$$\begin{aligned} \boldsymbol{f}_{opt} &= \overline{\boldsymbol{f}} + \boldsymbol{A}_\gamma^{-A}(\boldsymbol{g}_\varepsilon - \boldsymbol{A}\overline{\boldsymbol{f}}) \\ &= \boldsymbol{A}_\gamma^{-A}\boldsymbol{g}_\varepsilon + (\boldsymbol{I} - \boldsymbol{R})\overline{\boldsymbol{f}}. \end{aligned}$$

Für die Kovarianzmatrix $\boldsymbol{S}_{\gamma\,stoch}$ der regularisierten Lösung, deren Diagonalelemente die Varianz der einzelnen Komponenten der Lösung enthalten, die also ein Maß für den Fehler der regularisierten Lösung ist, folgt analog zum Abschn. 1.6 (siehe [24, S. 93] bzw. [26, S. 82 ff.]),

$$\boldsymbol{S}_{\gamma\,stoch} = \underbrace{\boldsymbol{A}_\gamma^{-A}\boldsymbol{S}_\varepsilon\boldsymbol{A}_\gamma^{-A^*}}_{\boldsymbol{S}_\gamma} + (\boldsymbol{I} - \boldsymbol{R})\boldsymbol{S}_f(\boldsymbol{I} - \boldsymbol{R})^*. \tag{39}$$

Dieser Ausdruck unterscheidet sich von dem im Abschn. 1.6 hergeleiteten Ausdruck $\boldsymbol{S}_\gamma$ dadurch, daß ein zusätzlicher Term abhängig von der Kovarinzmatrix $\boldsymbol{S}_f$ der a priori Lösungen und der Auflösungsmatrix $\boldsymbol{R}$ auftritt. Dieser zweite Term ist ein Maß für den Regularisierungsfehler (siehe Abschn. 1.4), den zu berücksichtigen uns in Abschn. 1.6 nicht gelungen war. Die Kovarianzmatrix $\boldsymbol{S}_{\gamma\,stoch}$ liefert uns daher eine recht gute Abschätzung für den Gesamtfehler in $\boldsymbol{f}_{opt}$.

[6] Rodgers [31, S. 618 Gl. (69)] löst die verbleibende Gleichung (38) iterativ, so daß er im Prinzip überhaupt keine Matrixinversion durchführt.

2.5 Backus–Gilbert–Verfahren

Ein ebenfalls weit verbreitetes Regularisierungsverfahren ist das Verfahren nach Backus und Gilbert, [1, 3, 4, 10], [19, S. 140 ff.], [20, S. 509 ff.], [22], [24, S. 61 ff.], [28, S. 815 ff.], [31, 40], [42, S. 160 ff.], [43, S. 1297 ff.], [44]. Gegenüber den bisher betrachteten Regularisierungsverfahren zeichnet es sich dadurch aus, daß ihm eine völlig neue Lösungsstrategie zugrunde liegt, indem es das Inversionsproblem unter einem gänzlich neuen Blickwinkel zu lösen versucht.

Nach Backus und Gilbert ist eine Näherungslösung $\boldsymbol{f}_\gamma = \boldsymbol{A}_\gamma^{-A}\boldsymbol{g}_\mathcal{E}$ genau dann eine gute Näherungslösung, wenn a) ihr Datenfehler klein ist, wenn also Meßfehler sich möglichst schwach auf die Näherungslösung auswirken, und wenn b) die Auflösung möglichst hoch ist. Diese Auffassung stimmt mit obiger Fehlerabschätzung für die stochastischen Verfahren (vgl. Gl. (39)) überein, nach der der Gesamtfehler $\boldsymbol{\mathcal{S}}_{\gamma\,stoch}$ genau dann klein ist, wenn a) die den Datenfehler charakterisierende Kovarianzmatrix $\boldsymbol{\mathcal{S}}_\gamma$ klein ist und wenn b) die Auflösungsmatrix $\boldsymbol{R}$ eine schmale Bandmatrix ist und nahezu der Einheitsmatrix $\boldsymbol{I}$ entspricht.

Wir suchen nun nach einer verallgemeinerten Inversen $\boldsymbol{A}_\gamma^{-A}$, die obige zwei Bedingungen erfüllt. Zuvor müssen uns wir jedoch geeignete Kriterien überlegen, um ein Maß für die Güte der Kovarianzmatrix $\boldsymbol{\mathcal{S}}_\gamma$ sowie der Auflösungsmatrix $\boldsymbol{R}$ zu besitzen. Für die Kovarianzmatrix $\boldsymbol{S}_\gamma$, deren Diagonalelemente die Varianz der Näherungslösung widerspiegeln, ist ein solches geeignetes Kriterium die Summe aller Diagonalelemente,

$$\text{size}\,(\boldsymbol{\mathcal{S}}_\gamma) = \sum_{i=1}^{m} \mathcal{S}_{\gamma\,ii}. \tag{40}$$

Um die Güte der Auflösungsmatrix $\boldsymbol{R}$ beurteilen zu können, müssen wir wissen, wie breit das Band der Auflösungsmatrix ist, d. h. wie weit die Auflösungsmatrix $\boldsymbol{R}$ von ihrer Idealgestalt, der Einheitsmatrix $\boldsymbol{I}$, abweicht. Backus und Gilbert betrachten hierzu die Größe

$$\text{spread}\,(\boldsymbol{R}) = \sum_{i=1}^{m}\sum_{j=1}^{m} w(i,j)(R_{ij} - I_{ij})^2 = \sum_{i=1}^{m}\sum_{j=1}^{m} w(i,j)R_{ij}^2 \tag{41}$$

mit einem Wichtungsfaktor

$$w(i,j) = (i-j)^2.$$

Dieser Faktor wichtet Elemente der Auflösungsmatrix $\boldsymbol{R}$ entsprechend ihrer Entfernung von der Diagonalen, so daß vorzugsweise „schmalbandige“ Auflösungsmatrizen selektiert werden. Durch dieses Wichten geht jedoch Information über die Diagonalelemente von $\boldsymbol{R}$ verloren, da der Wichtungsfaktor $w(i,j)$ für $i = j$ identisch Null ist. Neben der Auflösungsmatrix $\boldsymbol{R} = \boldsymbol{I}$ erfüllt daher auch jede beliebige andere diagonale

Auflösungsmatrix $\boldsymbol{R}$ die Bedingung spread($\boldsymbol{R}$) = 0. An die Zeilen der Auflösungsmatrix $\boldsymbol{R}$ muß deshalb die Zusatzbedingung

$$\sum_{j=1}^{m} R_{ij} = 1$$

geknüpft werden, so daß die Diagonalelemente der Auflösungsmatrix endlich bleiben und die „Fläche" einer jeden Zeile der Auflösungsmatrix der Einheitsfläche entspricht. Ein geeignetes Maß für die Erfüllung dieser Nebenbedingung ist die Größe

$$\text{diff}(\boldsymbol{R}) = \sum_{i=1}^{m} \Big[(\sum_{j=1}^{m} R_{ij}) - 1 \Big]. \tag{42}$$

Die gesuchte verallgemeinerte Inverse $\boldsymbol{A}_\gamma^{-A}$ sollte nun die beiden Bedingungen, size($\boldsymbol{\mathcal{S}}_\gamma$) = 0 und spread($\boldsymbol{R}$) = 0 möglichst gut, sowie die Zusatzbedingung diff($\boldsymbol{R}$) = 0 exakt erfüllen, was wir, wie in Abschn. 2.2, mit Hilfe zweier Lagrangescher Multiplikatoren γ und β erreichen. Wir definieren hierzu eine Lagrangefunktion

$$\Phi = \text{size}(\boldsymbol{\mathcal{S}}_\gamma) + \gamma \ \text{spread}(\boldsymbol{R}) + \beta \ \text{diff}(\boldsymbol{R}) \tag{43}$$

und minimieren diese bzgl. der Elemente $A_{\gamma ij}^{-A}$. Im Anschluß eliminieren wir dann den zweiten Lagrangeschen Multiplikator β, indem wir dieses Zwischenergebnis in die Zusatzbedingung einsetzen und β so bestimmen, daß diese automatisch erfüllt wird (ausführliche Rechnung siehe Anhang A.6). Als Endergebnis erhalten wir für die Elemente der verallgemeinerten Inversen die Bestimmungsgleichung

$$\boxed{A_{\gamma rk}^{-A} = \frac{\sum_{s=1}^{n} \mathcal{H}_{\gamma ks}^{r\,-1} u_s}{\sum_{k=1}^{n} \sum_{s=1}^{n} \mathcal{H}_{\gamma ks}^{r\,-1} u_s u_k}} \tag{44}$$

und damit

$$\boxed{\boldsymbol{f}_\gamma = \boldsymbol{A}_\gamma^{-A} \boldsymbol{g}_\mathcal{E}.}$$

Wichtig ist zu beachten, daß die (44) neu auftretende Hilfsmatrix $\boldsymbol{\mathcal{H}}_\gamma^r$, definiert durch (siehe Anhang A.6)

$$\mathcal{H}_{\gamma\,kl}^{r} = \mathcal{S}_{\mathcal{E}kl} + \gamma \sum_{j=1}^{m} w(r,j) A_{kj} A_{lj},$$

für jede Zeile r der verallgemeinerten Inversen $\boldsymbol{A}_\gamma^A$ neu berechnet und invertiert werden muß. Nur einmal zu erfolgen braucht dagegen die Berechnung des ebenfalls neu eingeführten Hilfsvektors $\boldsymbol{u}$,

$$u_k = \sum_{j=1}^{m} A_{kj}.$$

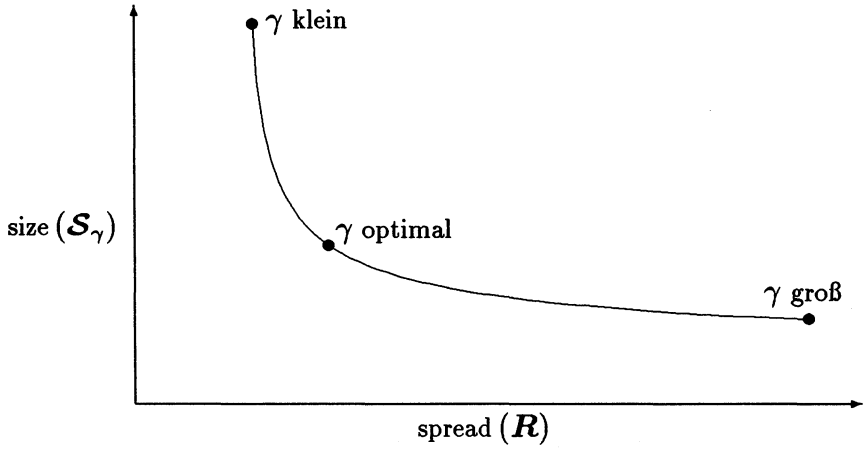

Abbildung 7: Zum Backus–Gilbert–Verfahren: Datenfehler kontra Auflösung.

Die Wahl des Regularisierungsparameters γ, der über die Hilfsmatrix $\mathcal{H}_\gamma^r$ in die Bestimmungsgleichung (44) eingeht, erfolgt ähnlich wie bei den bisher betrachteten Regularisierungsverfahren. Wiederum treten Datenfehler und erreichbare Auflösung miteinander in Konkurrenz, so daß eine Erhöhung der Auflösung (der Regularisierungsparameter γ wird größer gewählt, entsprechend verringert sich spread$(\boldsymbol{R})$) nur auf Kosten einer Vergrößerung des Datenfehlers erreicht werden kann und umgekehrt (siehe Abb. 7). Das Optimum ist erreicht, wenn size$(\mathcal{S}_\gamma)$ und spread $(\boldsymbol{R})$ sich in etwa entsprechen. Für den Anwender bedeutet dies, er muß im Prinzip für jedes spezielle Problem obigen Verlauf der Kurve bestimmen und hat dann die Möglichkeit, den Regularisierungsparameter γ optimal zu wählen. Es ist jedoch durchaus denkbar, daß man auch eine a posterori Parameterwahl, wie in Abschn. 1.4 beschrieben, durchführt.

Der große Vorteil des Backus–Gilbert–Verfahrens ist, daß es einen unmittelbaren Zusammenhang zwischen der Verstärkung von Meßfehlern sowie der erreichbaren Auflösung herstellt und somit von sicherlich hohem explanativem Wert ist. Überdies liefert es mit der spread-Funktion eine Definition der Auflösung, auf die wir kommenden Kapitel, in dem wir die bisher besprochenen Verfahren auf ein konkretes Beispiel anwenden wollen, noch einmal zurückkommen werden. In der Meteorologie spielt das Backus–Gilbert–Verfahren jedoch eine weniger bedeutende Rolle, was aufgrund des hohen numerischen Aufwandes [7] unmittelbar einzusehen ist. Es kommt daher insbesondere dann zum Einsatz, wenn nur sehr wenige Daten zur Verfügung stehen (n sehr klein), wie z. B. in der Seismik.

[7] Es müssen m (Anzahl der Höhenstützstellen) Matrixinversionen durchgeführt werden, wobei die Dimension der zu invertierenden Matrizen $\mathcal{H}_\gamma^r$ quadratisch mit der Anzahl n der spektralen Stützstellen anwächst.

Kapitel 3

Anwendung der Regularisierungsverfahren

Um die speziellen Eigenschaften der im letzten Kapitel vorgestellten Regularisierungsverfahren besser kennenzulernen, sollen die verschiedenen Verfahren im folgenden auf ein konkretes Beispiel angewendet werden. Hierzu betrachten wir die 142,175 GHz Spektrallinie von Ozon, einem starken Absorber und Emitter innerhalb der Atmosphäre, mit einem normalerweise ausgeprägtem Maximum im Volumenmischungsverhältnis von etwa 8–10 ppmv in 30–40 km Höhe. Unser Ziel ist es, anhand des am Boden zu beobachtenden Intensitätsspektrums auf den Höhenverlauf des Volumenmischungsverhältnisses von Ozon zurückzuschließen. Die Einflüsse anderer Atmosphärengase, insbesondere der von Wasserdampf und Sauerstoff bleiben in dieser einfachen Betrachtung unberücksichtigt. Für den Druckverlauf innerhalb der Atmosphäre wird die barometrische Höhenformel angenommen und der reale Temperaturverlauf wird durch eine Kosinusfunktion approximiert (siehe Abb. 8.a).

Ausgangspunkt für die nachfolgenden Betrachtungen ist das willkürlich angenommene Ozon–Höhenprofil der Abb. 8.b, mit einem ausgeprägten Maximum im Volumenmischungsverhältnis in 35 km Höhe. Aus diesem Ozon–Höhenprofil wird zunächst unter Berücksichtigung obiger Höhenprofile von Druck und Temperatur das am Boden zu beobachtende Intensitätsspektrum (siehe Abb. 9.a) über Vorwärtsrechnung ermittelt. Anschließend wird aus diesem Intensitätsspektrum mit Hilfe der besprochenen Regularisierungsverfahren auf das Ausgangsprofil zurückgeschlossen und das Inversionsergebnis mit dem Ausgangsprofil verglichen.

Das berechnete Intensitätsspektrum der Abb. 9.a besteht aus 399 diskreten spektralen Stützstellen. Doch wird, aufgrund der Symmetrie des Intensitätsspektrums bzgl. der Linienmitte, zur Wiedergewinnung des Ozon–Höhenprofils im folgenden nur die rechte, hochfrequente Flanke des Intensitätsspektrums berücksichtigt. Die in der Betrachtung verbleibenden 200 spektralen Stützstellen werden hierbei von der Linienmitte ausgehend von 1 bis 200 durchnummeriert. Des weiteren werden, entsprechend dem zu erwartenden höheren Informationsgehalt des Intensitätsspektrums im Bereich der Linienmitte, die spektralen Stützstellen in diesem Bereich etwas dichter angeordnet. Der minimale Abstand zwischen den Stützstellen beträgt hierbei 0.8 MHz. Nach außen hin dagegen nimmt die Dichte der spektralen Stützstellen exponentiell ab, so

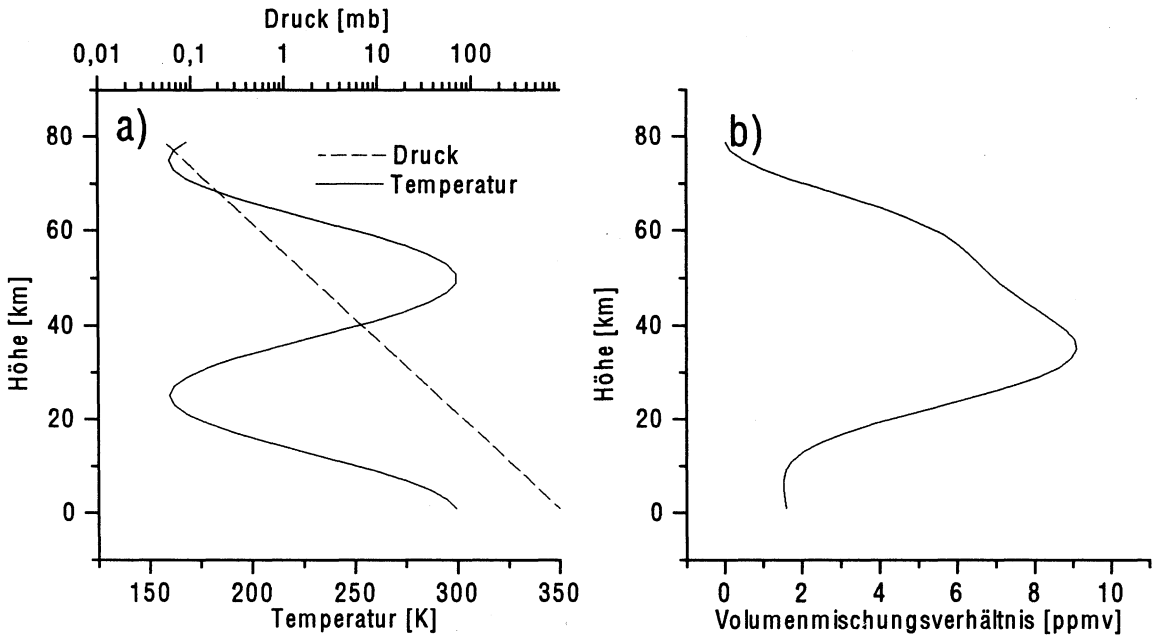

Abbildung 8: Willkürlich angenommene Höhenprofile: In a) Druck (barometrische Höhenformel) und Temperatur; in b) Volumenmischungsverhältnis von Ozon.

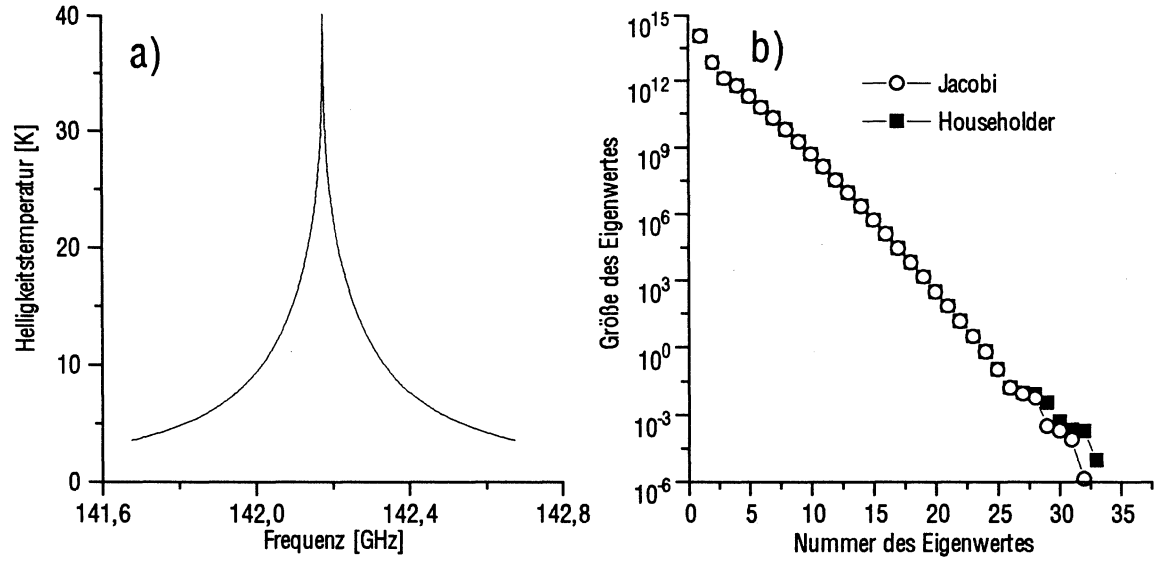

Abbildung 9: In a) das sich unter Berücksichtigung von Druck und Temperatur aus dem Ozon–Höhenprofil der Abb. 8b) ergebende Intensitätsspektrum (ausgedrückt über die Helligkeitstemperatur). In b) die Eigenwerte der Matrix $\boldsymbol{A}^*\boldsymbol{A}$ berechnet nach dem Jacobi- und dem Householder-Verfahren [28] (logarithmische Auftragung).

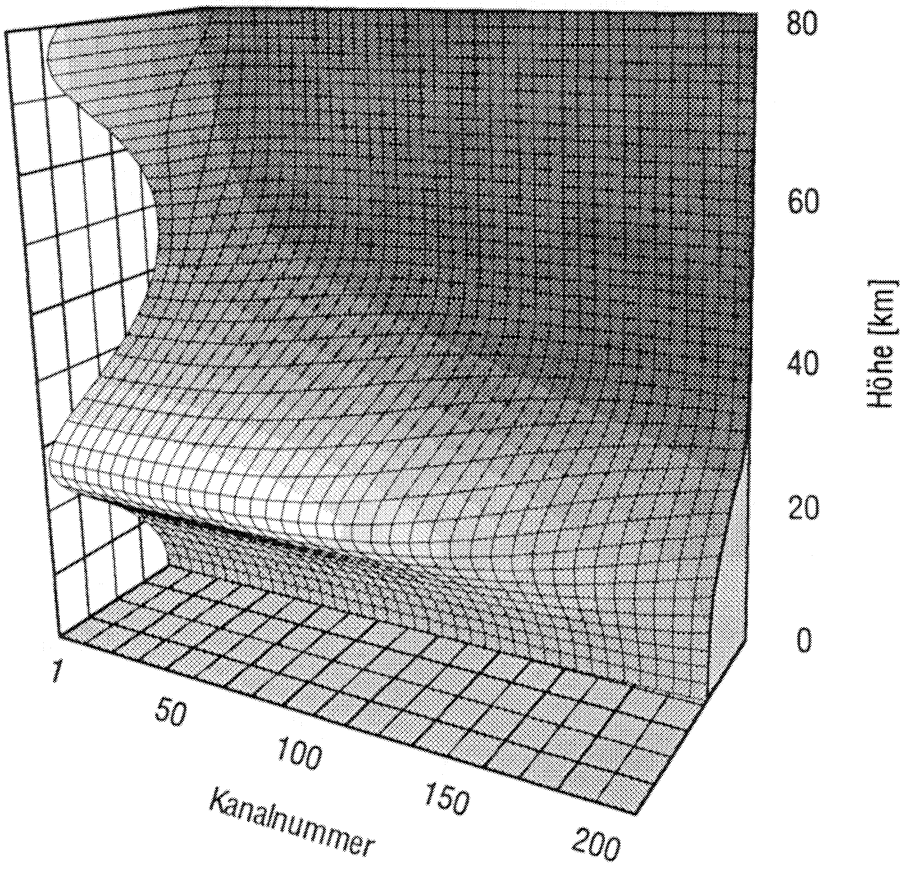

Abbildung 10: Die transponierte Matrix $\boldsymbol{A}^*$, deren Zeilen angeben, mit welchem Gewicht das Volumenmischungsverhältnis einer jeden Höhenschicht zum am Boden meßbaren Intensitätsspektrum beiträgt. Die Spalten der Matrix $\boldsymbol{A}^*$ bezeichnet man auch als Gewichtsfunktionen. Die Darstellung erfolgt dreidimensional, wobei bzgl. des Gewichts willkürliche Einheiten gewählt wurden.

daß am äußeren rechten Rand des Intensitätsspektrums ein Abstand von 5,4 MHz zwischen den Stützstellen erreicht wird. [1]

In Abb. 10 ist die aus der Strahlungstransfergleichung unter Zugrundelegung obiger Höhenprofile sich ergebende Matrix $\boldsymbol{A}$ bzw. deren Transponierte [2] $\boldsymbol{A}^*$ dargestellt,

[1] Bei dem in [18] verwendeten Millimeterwellen-Radiometer sind die verfügbaren 1200 spektralen Stützstellen äquidistant in einem Abstand von 0.8 MHz angeordnet, so daß die gewählte spektrale Auflösung im Bereich der Linienmitte exakt der maximalen Geräteauflösung entspricht. Am äußeren Rand hingegen entspricht eine spektrale Stützstelle im Modell dem Mittelwert von 6-7 Kanälen des Millimeterwellen-Radiometers.

[2] Korrekt wäre Adjungierte, doch ist im vorliegenden Fall die Matrix $\boldsymbol{A}$ reell.

wobei zur Berechnung die Atmosphäre in 40 Höhenschichten der Dicke 2 km unterteilt wurde – getreu dem bereits Ende des Abschn. 1.5 zitierten Grundsatz, daß es wesentlich günstiger sei, zunächst etwas feiner zu diskretisieren und anschließend zu regularisieren. Überdies erweist es sich auch bei der Berechnung der Auflösungsmatrix $\boldsymbol{R}$ in Abschn. 3.2 als von Vorteil, die Dicke der Schichten deutlich kleiner als die tatsächlich erreichbare Auflösung zu wählen. Die Matrix $\boldsymbol{A}$ ist damit eine 200×40-Matrix. Auffallend ist in Abb. 10, daß, obwohl die spektralen Stützstellen im Bereich der Linienmitte sehr dicht liegen, lediglich die ersten 20 Spalten der Matrix $\boldsymbol{A}^*$ Aufschluß über den Höhenbereich oberhalb von 50 km geben. Dagegen überstreichen die restlichen 180 Spalten nahezu alle den selben Höhenbereich unterhalb von 50 km. Aufgrund dieser Redundanz ist es daher wenig verwunderlich, daß die Matrix $\boldsymbol{A}^*\boldsymbol{A}$ nahezu singulär ist.

3.1 Abgeschnittene Singulärwertzerlegung

Die Abb. 9.b zeigt das bereits in Abschn. 1.5 erwähnte exponentielle Fallen der Eigenwerte der Matrix $\boldsymbol{A}^*\boldsymbol{A}$ (logarithmische Auftragung), aufgrund dessen die bodengebundene passive Millimeterwellenradiometrie zur Klasse der exponentiell schlecht gestellten Probleme gehört [23]. Die Durchführung der abgeschnittenen Singulärwertzerlegung erfolgt wie in Abschn. 2.1 beschrieben, d. h. wir summieren entsprechend obiger Reihenfolge der Eigenwerte über die Beiträge zur regularisierten Lösung auf. Um den realen Verhältnissen möglichst nahe zu kommen und um überdies eine realistische Aussage über die erreichbare Höhenauflösung zu erhalten, wird dem über Vorwärtsrechnung ermittelten Intensitätsspektrum der Abb. 9.a zusätzlich ein entsprechendes Rauschen überlagert. Im aktuellen Fall wird hierbei ein weißes Rauschen mit Gaußscher Häufigkeitsverteilung und einer Standardabweichung von 0,01 K angenommen. [3]

In den Abbildungen 11 und 12 sind die Ergebnisse der ersten 16 Summationen, die zur Näherungs- bzw. regularisierten Lösung führen, dargestellt. Sie sollten vor dem Hintergrund der in Abschn. 1.4 beschriebenen Wahl des Regularisierungsparameters und der dortigen Abb. 1 betrachtet werden. Zu Beginn ist der Fehler der Näherungslösung aufgrund der allzu starken Regularisierung (erste wenige Summationen) noch sehr groß, doch bereits nach 6, 7 Summationen gibt die regularisierte Lösung das Ausgangsprofil recht gut wieder. Mit jeder weiteren Summation verringert sich die Stärke der Regularisierung, so daß nach und nach auch die Feinstruktur des Ausgangsprofils immer besser herausgearbeitet wird. Nach etwa 11, 12 Summationen kommt im vorliegenden Fall die regularisierte Lösung dem Urbild am nächsten. Hier sollte das Verfahren abgebrochen werden. Wird statt dessen weiter aufsummiert,

[3] Dieser Wert sollte mit dem in [18] verwendeten Radiometer bei Integrationszeiten von 1-2 Stunden unter optimalen Voraussetzungen erreicht werden können.

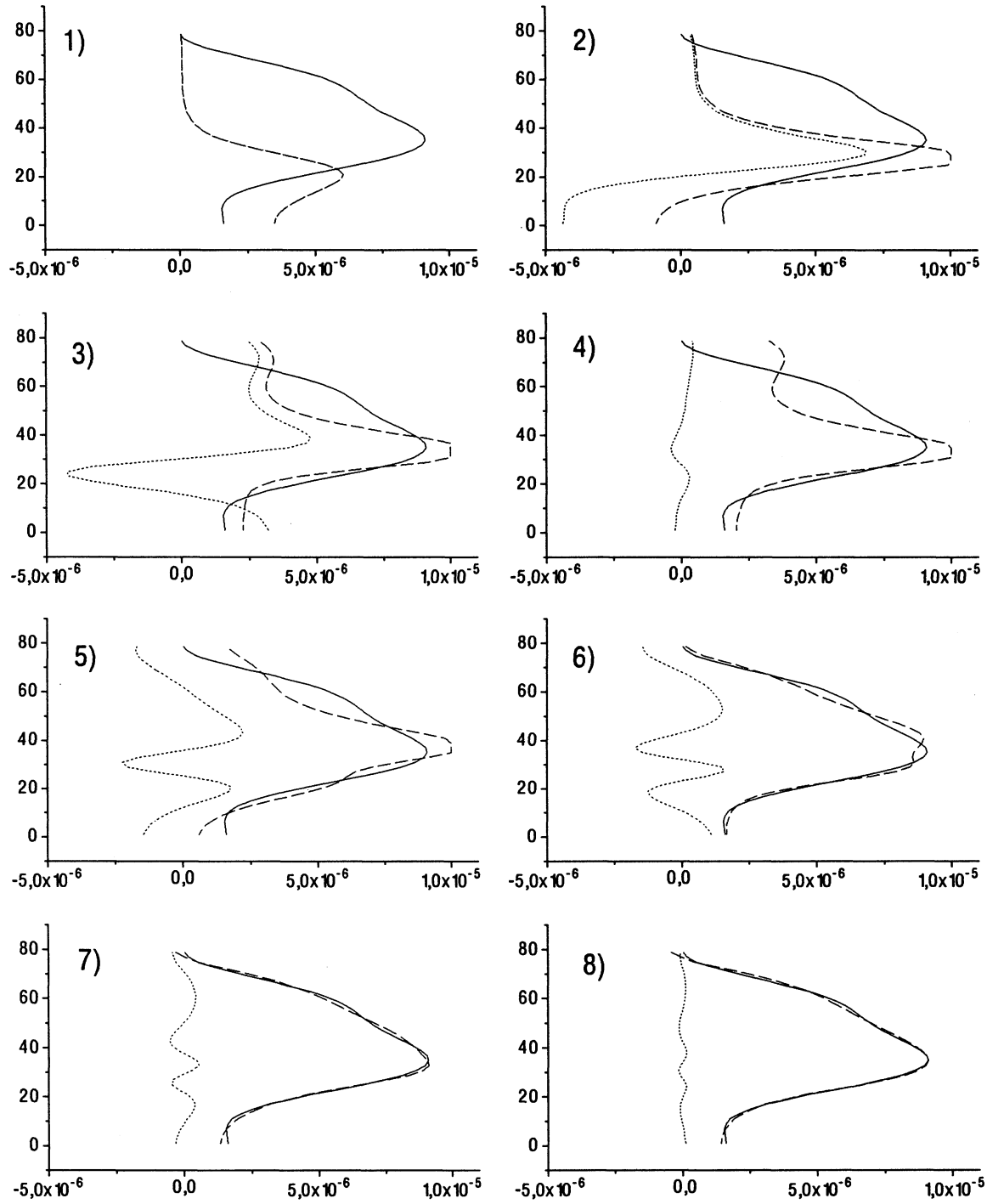

Abbildung 11: Abgeschnittene Singulärwertzerlegung. Das Entstehen der Näherungslösung durch Summation über die Eigenvektoren. Dargestellt ist jeweils das Ausgangsprofil (durchgezogen), die momentane Näherungslösung (gestrichelt) sowie der aktuelle Beitrag zur Näherungslösung (gepunktet). Im wesentlichen entspricht dieser aktuelle Beitrag dem Eigenvektor multipliziert mit einem vom aktuellen Eigenwert und den meßfehlerbehafteten Daten abhängigen Entwicklungskoeffizienten.

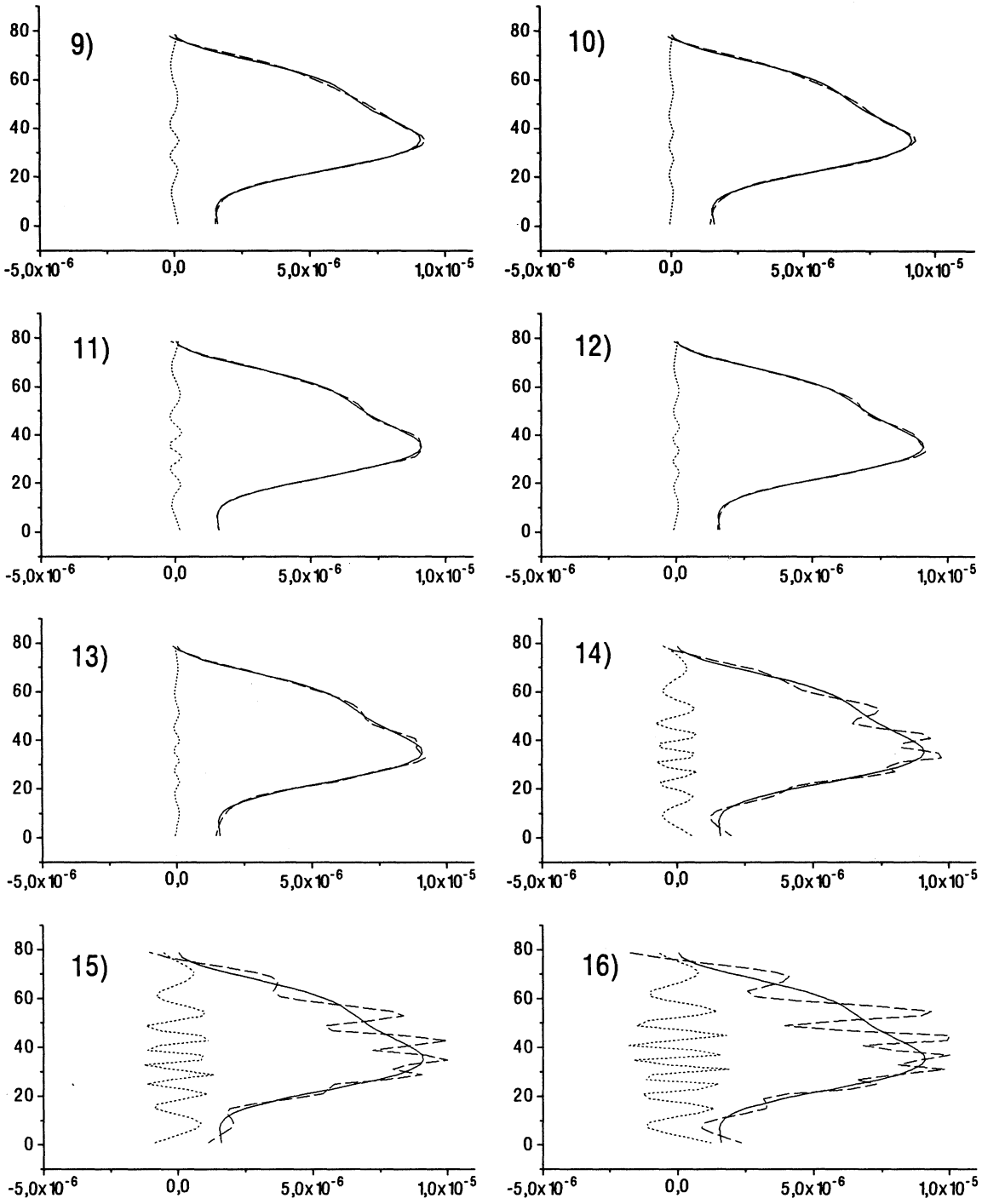

Abbildung 12: Abgeschnittene Singulärwertzerlegung. Zur Erläuterung beachte die Bildunterschrift der Abb. 11. Interessant ist zu verfolgen, wie die Eigenvektoren mit jedem Schritt „schneller oszillieren“ und dadurch die Höhenauflösung immer weiter verbessern. Die größte Höhenauflösung wird in etwa 30 km Höhe erreicht, da in diesem Bereich die „Frequenz“ der oszillierenden Eigenvektoren am höchsten ist. Wird die Höhenauflösung zu hoch getrieben, d. h. wird versucht, mehr Information aus den Meßdaten herauszuholen als vorhanden, wird die Näherungslösung instabil.

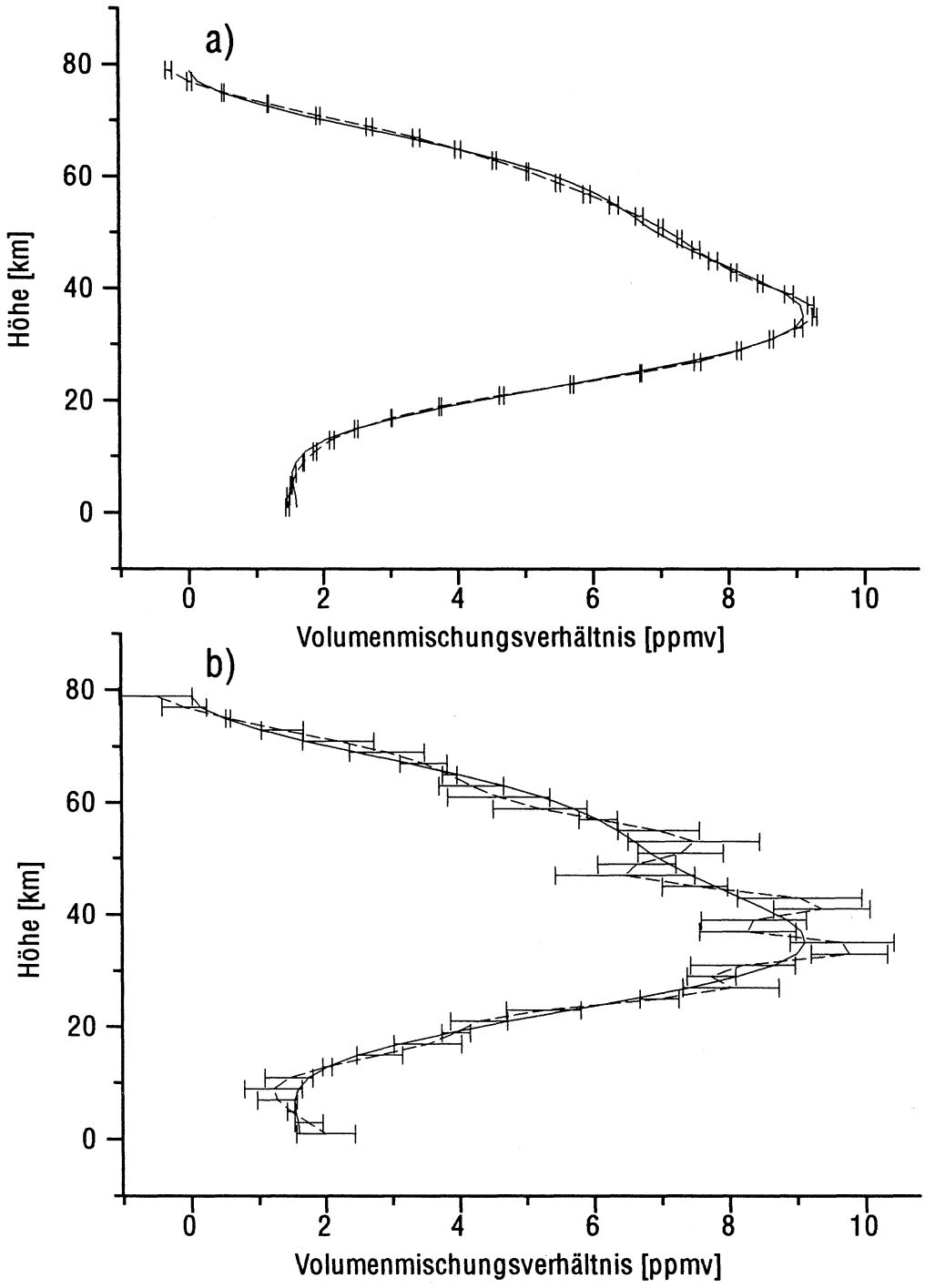

Abbildung 13: Fehler der abgeschnittenen Singulärwertzerlegung. Dargestellt sind die aus der Fehlerabschätzung in Abschn. 1.6 Gl. (10) sich ergebenden Fehlerbalken, in a) nach 10 sowie in b) nach 14 Summationen. Besonders nach 14 Summationen, wenn der Datenfehler gegenüber dem Regularisierungsfehler eindeutig dominiert, geben die Fehlerbalken den Fehler der Näherungslösung (gestrichelt) verglichen mit der exakten Lösung (durchgezogen) recht gut wieder.

so machen sich mit jedem weiteren Schritt die dem Intensitätsspektrum überlagerten Meßfehler stärker bemerkbar. In der Folge beginnt die Näherungslösung stark zu oszillieren und wird bereits nach wenigen zusätzlichen Schritten vollkommen unbrauchbar (siehe Abb. 12.16).

In der Realität ist das Ausgangsprofil natürlich nicht bekannt, so daß der Regularisierungsparameter γ bzw. der Eigenwert, bis zu dem über die Beiträge zur regularisierten Lösung aufsummiert wird, nicht optimal gewählt werden kann. Statt dessen erfolgt die Parameterwahl meist a posteriori, wie in Abschn. 1.4 letzter Absatz und Abschn. 2.1 Gl. (12) beschrieben; vor allen Dingen dann, wenn das Inversionsverfahren operationell betrieben werden soll. In diesem Fall wird das Abbruchkriterium (mit r gleich 1) bereits nach 9 Summationen erreicht. Doch auch diese Näherungslösung gibt das Ausgangsprofil gut wieder. Allerdings ist die Höhenauflösung noch deutlich niedriger als z. B. nach 13 Summationen, wie aus der Struktur des letzten aufsummierten Eigenvektors ersichtlich.

Es bleibt wiederum dem erfahrenen Anwender selbst überlassen, den Regularisierungsparameter nach Gutdünken besser zu wählen. Ein wichtiges Hilfsmittel ist ihm hierbei die in Abschn. 1.6 Gl. (10) geführte Betrachtung über den Fehler der regularisierten Lösung. Leider liefert diese nur eine Abschätzung für den Datenfehler, so daß besonders zu Beginn der Summation, wenn der Regularisierungsfehler groß ist gegenüber dem Datenfehler, eine Aussage über den Gesamtfehler nicht möglich ist. Da in der Folge jedoch der Regularisierungsfehler mit jeder weiteren Summation abnimmt, so daß irgendwann der stark anwachsende Datenfehler gegenüber dem kleiner werdenden Regularisierungsfehler dominiert (im vorliegenden Fall nach etwa 13 Summationen), kann irgendwann der Gesamtfehler mit dem Datenfehler gleichgesetzt werden. So liefern die aufgrund obiger Fehlerabschätzung in die Abbildungen 13.a und vor allen Dingen 13.b eingetragenen Datenfehlerbalken eine ganz vorzügliche Abschätzung auch für den Gesamtfehler der Näherungslösung. Jedoch sollte das Verfahren abgebrochen werden, bevor der Datenfehler zu groß zu werden droht und die Näherungslösung zu oszillieren beginnt, im vorliegenden Fall also nach 11 oder 12 Summationen.

3.2 Zur Definition der Auflösung

Um die erreichbare Höhenauflösung quantitativ zu bestimmen, berechnen wir die in Abschn. 1.6 eingeführte Auflösungsmatrix $\boldsymbol{R}$ (siehe Abb. 14). Wie aufgrund der Abb. 12 nicht anders zu erwarten, erhalten wir die höchste Auflösung im Bereich zwischen 25 und 35 km Höhe (hier ist das Band am schmalsten). Oberhalb und unterhalb dagegen ist die Auflösung deutlich niedriger.

Etwas überraschen mag zunächst das Auftreten von negativen Werten innerhalb der Auflösungsmatrix $\boldsymbol{R}$, die eine Größe von bis zu 15 % der Werte entlang der Hauptdiagonalen erreichen. Aufgrund der Ausführungen in Abschn. 1.5 und Abschn. 1.6,

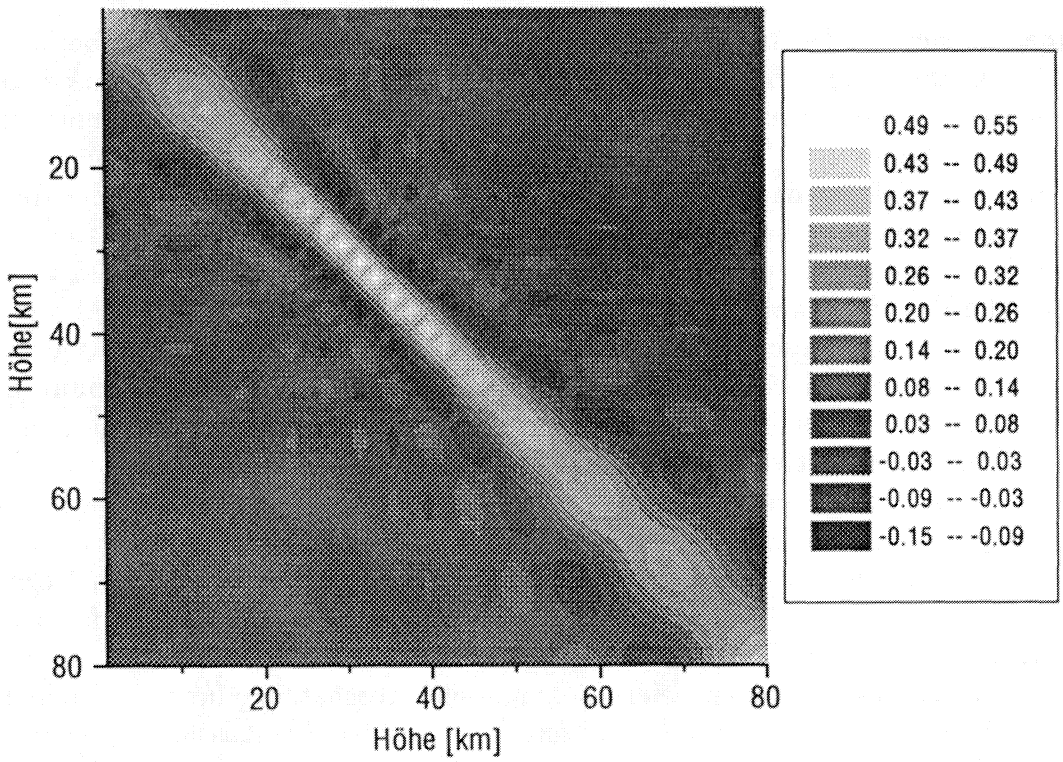

Abbildung 14: Die Auflösungsmatrix $\boldsymbol{R}$ der regularisierten Lösung, berechnet für eine nach der 12. Summation abgeschnittene Singulärwertzerlegung. Die Zeilen der Auflösungsmatrix geben Aufschluß darüber, über welchen Höhenbereich (Abszisse) die unbekannte exakte Lösung „gemittelt" werden müßte, um den Wert der regularisierten Lösung in einer bestimmten Höhe (Ordinate) zu erhalten.

in denen wir die Regularsierung als eine Art lokale Mittelung interpretiert hatten, hätten wir statt dessen das Auftreten von ausschließlich positiven Werten erwartet. Um das Zustandekommen von negativen Werten plausibel zu machen und um die erreichbare Auflösung in konkrete Zahlenwerte zu fassen, stellen wir im folgenden einige ausgewählte Zeilen der Auflösungsmatrix $\boldsymbol{R}$ dar in Form von Balkendiagrammen (siehe Abb. 15).

Wäre die Auflösung beliebig hoch, d. h. wäre die Auflösungsmatrix $\boldsymbol{R}$ gleich der Einheitsmatrix $\boldsymbol{I}$, so würde der Wert der Näherungslösung in einer bestimmten Höhenschicht sich allein aus dem Wert des Originalprofils in dieser Höhenschicht ergeben (In jedem Diagramm würde jeweils nur ein Balken auftreten, was im kontinuierlichen Fall einer Deltafunktion entspräche). Hierzu hätten wir über alle Eigenvektoren aufsummieren müssen. Aufgrund von Meßfehlern in den Daten mußte die

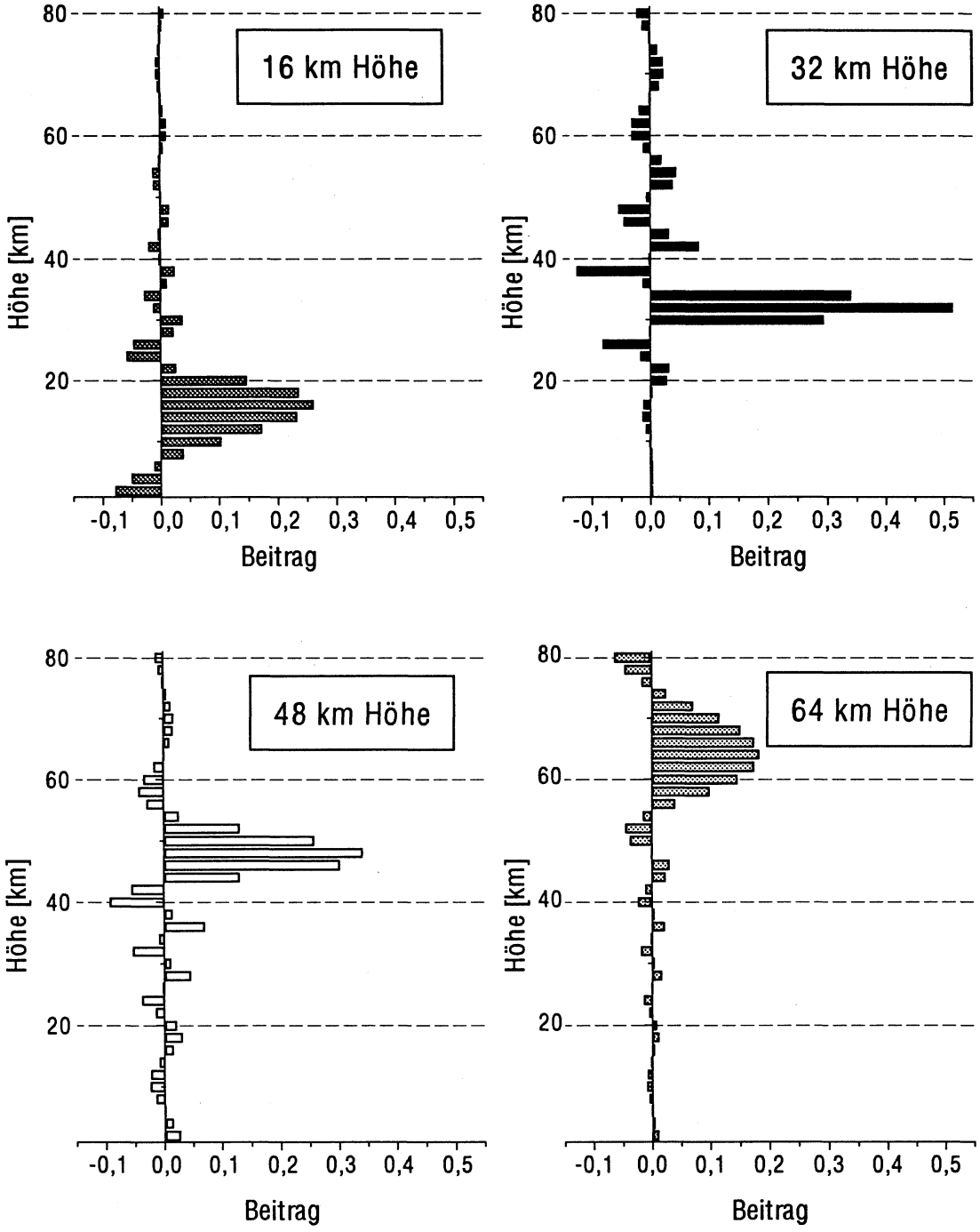

Abbildung 15: Ausgewählte Zeilen der Auflösungsmatrix zur abgeschnittenen Singulärwertzerlegung. Ein jedes Balkendiagramm gibt an, wie stark der Wert des Originalprofils in einer bestimmten Höhenschicht zu dem von uns ermittelten Wert der Näherungslösung in der angegebenen Höhe beiträgt. Z. B. ergibt sich der Wert der Näherungslösung in 32 km Höhe im wesentlichen aus den Beiträgen dreier Schichten, eine jede der Dicke 2 km, so daß die dortige Höhenauflösung 6-7 km beträgt. In 64 km Höhe ist die Höhenauflösung mit etwa 20 km dagegen deutlich niedriger.

Summation jedoch bereits vorher abgebrochen werden, so daß insbesondere „hochfrequente“ Eigenvektoren, die aufgrund ihrer zugehörigen kleinen Eigenwerte zu einer enormen Verstärkung des Datenfehlers geführt hätten, abgeschnitten wurden. Da aber die hochfrequenten Anteile bei der spektralen Darstellung der Deltafunktion fehlen, beobachten wir ein Gibbssches Phämomen, nämlich das Auftreten von Überschwingern und damit das Auftreten von negativen Werten. Für unsere weiteren Betrachtungen hat dies zur Folge, daß wir zunächst unser Verständnis für den Begriff lokale Mittelung (besser wäre vielleicht lokale Glättung) dahingehend erweitern müssen, daß Höhenschichten zu dieser Mittelung auch mit einem negativen Gewicht beitragen können. Weitaus schwieriger ist hingegen die Frage zu beantworten, was wir unter der Auflösung konkret verstehen wollen.

Die in der Fernerkundung atmosphärischer Parameter gebräuchlichste Definition der Auflösung ist die als Wellenlänge eines sinusförmigen Ausgangsprofils, das mittels Inversion gerade noch aufgelöst werden kann. Doch ändert sich, wie wir in Abb. 15 gesehen haben, die Auflösung mit der Höhe, wohingegen die Wellenlänge einer sinusförmigen Störung über der Höhe konstant bleibt. Rodgers schlägt deshalb vor [34], statt dessen die von Backus und Gilbert eingeführte spread-Funktion zu verwenden. Da in der Physik der Begriff der Auflösung eng mit dem Auftreten kontinuierlicher Funktionen verbunden ist, geht Rodgers von deren ursprünglicher, kontinuierlicher Definition der spread-Funktion aus und definiert die erreichbare Auflösung in einer bestimmten Höhe z als

$$\text{spread'}(z) = 12\frac{\int (z-z')^2 R(z,z')^2 dz}{\left[\int R(z,z')dz\right]^2}$$

bzw. in diskretisierter Form

$$\text{spread'}(i) = 4\,\Delta z\frac{\sum_{j=1}^{m}\left[(i-j+1/2)^3-(i-j-1/2)^3\right]R_{ij}^2}{\left[\sum_{j=1}^{m} R_{ij}\right]^2}. \tag{45}$$

Die Normierungsfaktoren 12 bzw. 4 wurden hierbei so gewählt, daß sich für eine Rechteckfunktion der Breite b und der Höhe $1/b$, gerade die Breite b dieser Rechteckfunktion als Auflösung ergibt; Δz ist die Dicke einer jeden Höhenschicht. Für das Backus–Gilbert–Verfahren erweist sich die Gl. (45) als die am besten geeignete Definition der Auflösung, da in der Auflösungsmatrix dieses Verfahrens keine Überschwinger auftreten (siehe Abb. 28). Für die anderen, linearen Verfahren ist diese Definition jedoch weniger geeignet, wie entsprechende Untersuchungen in [3] sowie eigene Erfahrungen zeigen. Obwohl die mit diesen Verfahren erreichbare Auflösung nahezu doppelt so hoch ist als die des Backus–Gilbert–Verfahrens [3], schneiden sie aufgrund von Überschwingern innerhalb der Auflösungsmatrix $\boldsymbol{R}$ bei obiger Definition der Auflösung deutlich schlechter ab.

Wir verwenden statt dessen zur Definition der Auflösung bei diesen Verfahren das Rayleigh Kriterium und verstehen unter der maximal erreichbaren Auflösung den kleinsten Abstand zweier Deltafunktionen im Originalprofil, die mittels Inversion gerade noch aufgelöst werden können. Nach Rayleigh entspricht dann die Auflösung gerade dem Abstand zwischen Hauptmaximum und erstem Minimum bzw. der Hälfte des Abstandes zwischen den beiden ersten Minima. Für die abgeschnittene Singulärwertzerlegung (siehe Abb. 15) beträgt die maximal mögliche Auflösung für ein Rauschen von 0,01 K deshalb in 16 km Höhe etwa 12 km, in 32 km ungefähr 6-7 km, in 48 km Höhe 9-10 km und in 64 km Höhe 14-15 km.

3.3 Anwendung der Tikhonov–Phillips–Regularisierung

Bei der Betrachtung der Tikhonov–Phillips–Regularisierung beschränken wir uns auf die von Phillips vorgeschlagene Modifikation des Verfahrens, indem wir als Nebenbedingung fordern, daß die erste Ableitung des gesuchten Profils verschwindet, $\boldsymbol{f}' = 0$. Das gesuchte Profil soll also möglichst glatt sein (siehe Abschn. 2.2). Diese Annahme ist im vorliegenden Fall, in dem das Originalprofil relativ glatt ist, sicherlich gerechtfertigt. Doch wird man auch dann, wenn das Originalprofil nicht bekannt ist, i. d. Regel auf obige Nebenbedingung zurückgreifen. Gegenüber der ursprünglichen Forderung von Tikhonov, $\boldsymbol{f} = 0$, besitzt diese den Vorteil, daß selbst bei extrem starker Regularisierung zumindest eine Aussage über das mittlere Volumenmischungsverhältnis bzw. den Säulengehalt des betrachteten Gases möglich ist (siehe Abb. 16.1)

Die Wahl des bei der Tikhonov–Phillips–Regularisierung im Gegensatz zur abgeschnittenen Singulärwertzerlegung kontinuierlich einstellbaren Regularisierungsparameters γ erfolgt derart, daß beide Verfahren direkt miteinander verglichen werden können (vgl. Abb. 11 und 12 mit Abb. 16 bzw. 17). Hierzu wählen wir den Regularisierungsparameter zunächst relativ groß, $\gamma = 8\,000$ in Abb. 16.1, und multiplizieren im Anschluß den Regularisierungsparameter jeweils mit einem Faktor 0,2. Wie bei der abgeschnittenen Singulärwertzerlegung kommt mit dieser Wahl des Regularisierungsparameters die Näherungslösung auch hier dem Originalprofil nach 11, 12 Schritten am nächsten und beginnt nach 13 Schritten zu oszillieren.

Im Vergleich zur abgeschnittenen Singulärwertzerlegung liefert die Tikhonov–Phillips–Regularisierung geringfügig bessere Inversionsergebnisse. Dies liegt in erster Linie daran, daß das Originalprofil relativ glatt ist, so daß die Nebenbedingung $\boldsymbol{f}' = 0$ zusätzliche (wahre) Information über das zu ermittelnde Höhenprofil liefert. Wird die Wahl des Regularisierungsparamters a posteriori mit einem Paramter $r = 1$ durchgeführt (siehe Abschn. 1.4) so erhalten wir als Näherungslösung die Abb. 17.10. Dieses Ergebnis ist besser als das entsprechende Resultat der abgeschnittenen Singulärwertzerlegung, wenn auch noch nicht optimal. Hinsichtlich der Fehlerbetrachtung

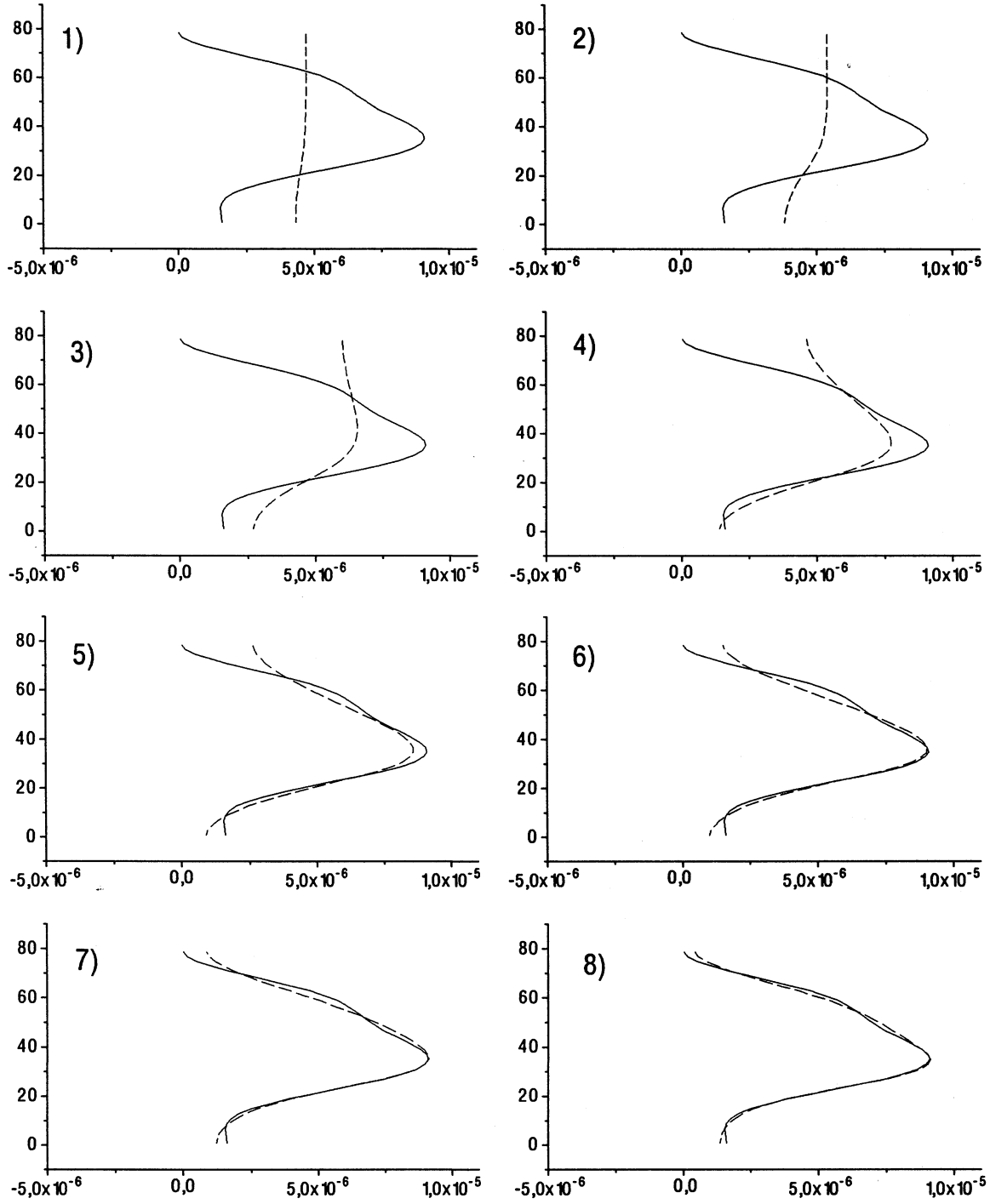

Abbildung 16: Tikhonov–Phillips–Regularisierung mit der Nebenbedingung $f' = 0$, d. h. das gesuchte Höhenprofil soll möglichst glatt sein. Dargestellt ist jeweils das Ausgangsprofil (durchgezogen) sowie die momentane Näherungslösung (gestrichelt). Interessant ist zu beobachten, wie sich bei extrem starker Regularisierung ein über der Höhe konstantes Höhenprofil einstellt (Bild 1), aus dem auf den Säulengehalt des betrachteten Gases geschlossen werden kann.

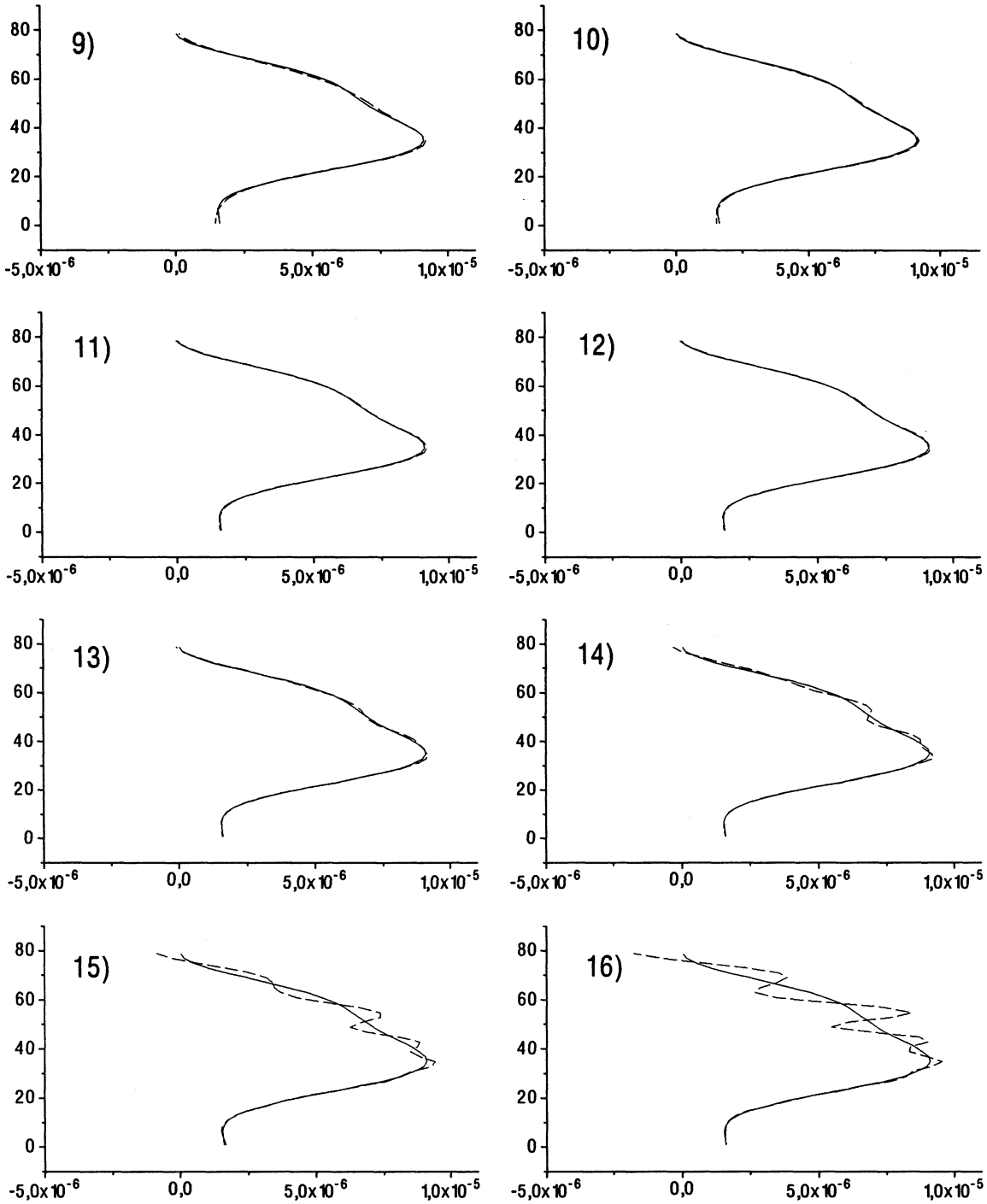

Abbildung 17: Tikhonov–Phillips–Regularisierung. Zur Erläuterung beachte die Bildunterschrift der Abb. 16. Wie auch bei der abgeschnittenen Singulärwertzerlegung (siehe Abb. 12) beginnt die Lösung bei zu schwacher Regularisierung zu oszillieren, so daß das Verfahren nach 12 oder 13 Schritten abgebrochen werden sollte.

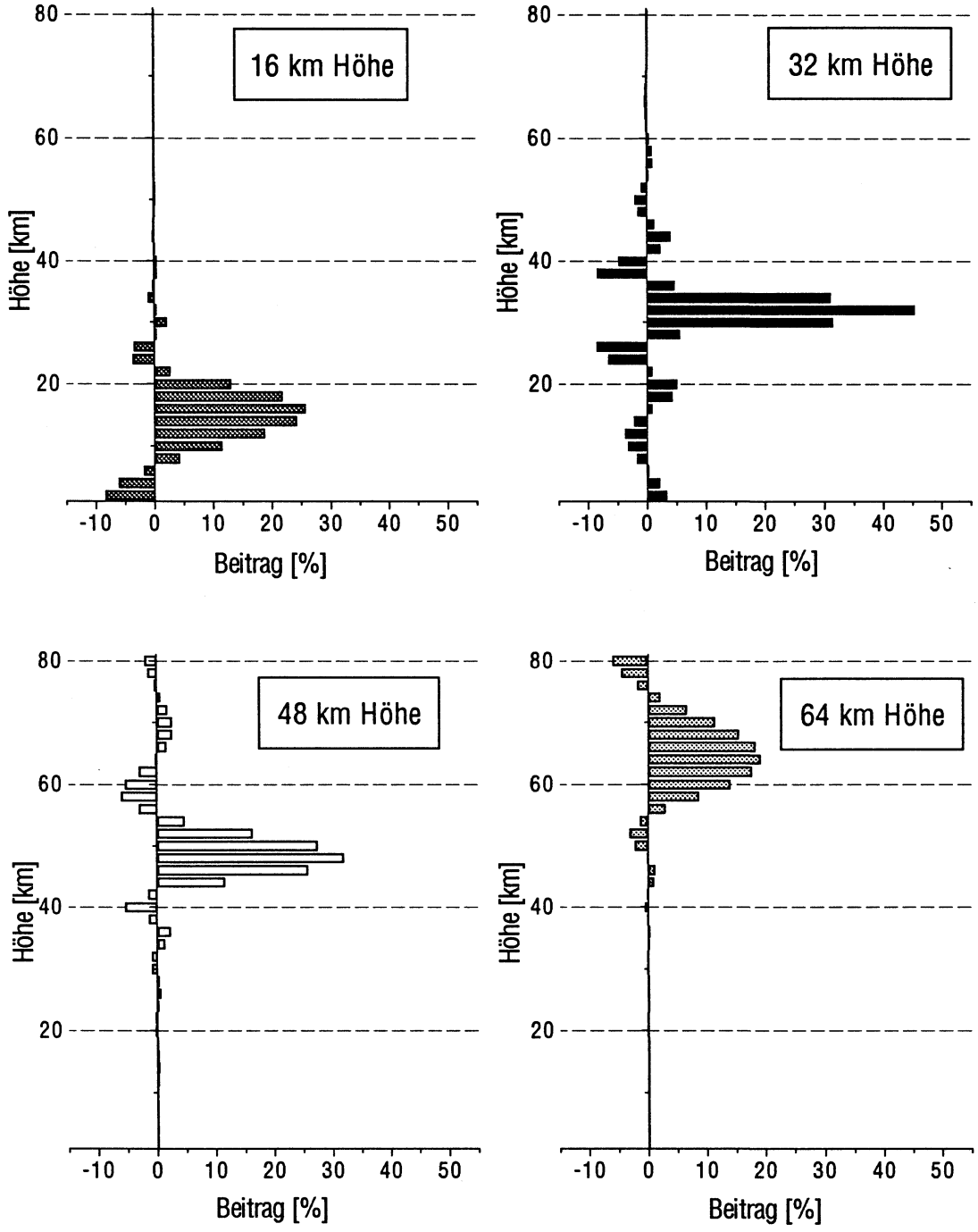

Abbildung 18: Ausgewählte Zeilen der Auflösungsmatrix zur Tikhonov–Phillips–Regularisierung. Ein jedes Balkendiagramm gibt an, wie stark der Wert des Originalprofils in einer bestimmten Höhenschicht zu dem von uns ermittelten Wert der Näherungslösung in der angegebenen Höhe beiträgt. Die maximal erreichbare Auflösung ist bei abgeschnittener Singulärwertzerlegung (vgl. Abb. 15) und Tikhonov–Phillips–Regularisierung mit jeweils 6-7 km in 32 km Höhe bei einem Rauschen von 0,01 K etwa gleich hoch.

in Abschn. 1.6, die eine Abschätzung für den Datenfehler liefert, unterscheiden sich abgeschnittene Singulärwertzerlegung und Tikhonov–Phillips–Regularisierung praktisch nicht, weshalb auf eine Darstellung der aus dieser Abschätzung sich ergebenden Fehlerbalken verzichtet und nochmals auf die Abb. 13 verwiesen wird. Bzgl. der Auflösungsmatrix $\boldsymbol{R}$ bestehen zwischen abgeschnittener Singulärwertzerlegung und Tikhonov–Phillips–Regularisierung nur geringe Unterschiede (vgl. die Abbildungen 15 und 18). Zwar sind die Überschwinger bei der Tikhonov–Phillips–Regularisierung etwas kleiner, doch ist die erreichbare Auflösung bei beiden Verfahren letztlich gleich.

3.4 Anwendung der Landweber–Iteration

Bereits bei der Betrachtung des Verfahrens in Abschn. 2.3.1 hatten wir festgestellt, daß die Landweber–Iteration ein nur wenig effizientes Regularisierungsverfahren sein kann – insbesondere dann, wenn wie im vorliegenden Fall das Problem exponentiell schlecht gestellt ist. Bei der Anwendung dieses Verfahrens auf die Inversion der Spektrallinie von Ozon finden wir diese Vermutung bestätigt, indem erstmals nach 8000 Iterationen die Lösung dem ursprünglichen Ozonprofil näherungsweise entspricht (siehe Abb.19). Nach weiteren 72 000 Iterationen werden die Ergebnisse der Landweber–Iteration vergleichbar mit denen von abgeschnittener Singulärwertzerlegung und Tikhonov–Phillips–Regularisierung nach 9 oder 10 Schritten. Um gar die optimale Näherungslösung zu erhalten müßten sicherlich weitere 50 000 - 100 000 Iterationsschritte durchgeführt werden, worauf aus Gründen des hohen numerischen Aufwandes jedoch verzichtet wird.

Zwar ist bei der Landweber–Iteration im Vergleich zu den anderen iterativen Verfahren (z. B. Verfahren der konjugierten Gradienten) der numerische Aufwand für einen einzelnen Iterationsschritt relativ gering, doch reicht dies bei weitem nicht aus, um die hohe Zahl der notwendigen Iterationsschritte auch nur annähernd zu kompensieren. Wir verzichten deshalb auf eine eingehendere Anwendung dieses numerisch so unattraktiven Verfahrens und wenden uns statt dessen dem weitaus effizienteren Verfahren der konjugierten Gradienten zu.

3.5 Anwendung des konjugierten Gradienten–Verfahrens

Wie bereits Ende des Abschnitts 1.6 erwähnt, ist bei den nichtlinearen iterativen Regularisierungsverfahren eine Berechnung der Auflösungsmatrix $\boldsymbol{R}$ sowie eine Fehlerabschätzung entsprechend Gl. (10) nicht möglich. Statt dessen beschränken wir uns bei der Anwendung dieses Verfahrens wie im vorigen Abschnitt auf die Darstellung der Iterationsfolge und untersuchen im Anschluß den Einfluß verschiedener Startprofile auf die Konvergenz des Verfahrens.

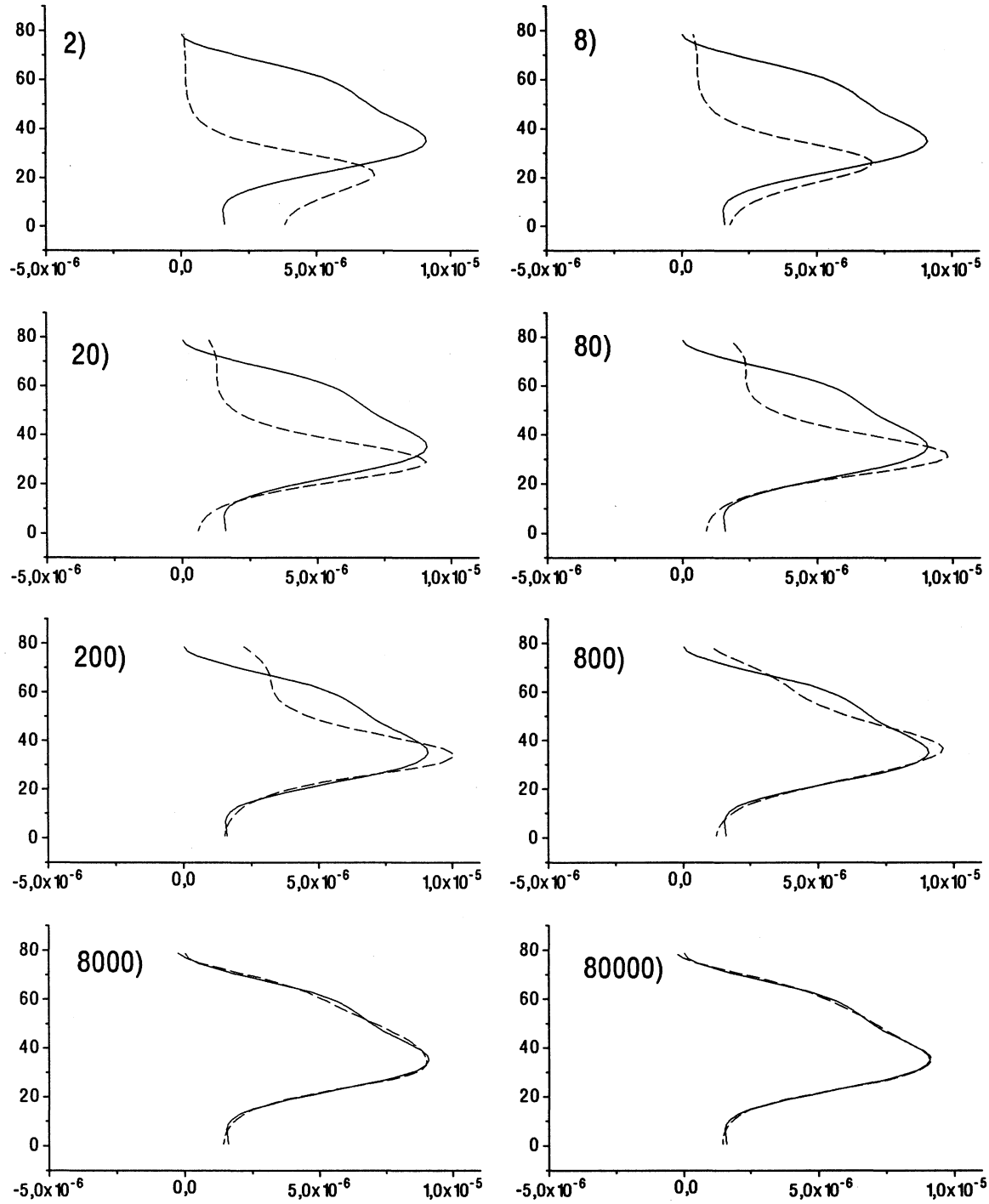

Abbildung 19: Landweber–Iteration. Dargestellt sind jeweils Originalprofil (durchgezogen) sowie momentane Näherungslösung (gestrichelt). Die Nummern kennzeichnen die Anzahl der bereits durchgeführten Iterationsschritte, wobei das Verfahren aufgrund des hohen numerischen Aufwandes bereits nach 80 000 Iterationsschritten, also weit vor Erreichen des instabilen Bereichs, abgebrochen wird.

Abb. 20 zeigt die für verschiedene Iterationstiefen erhaltenen Inversionsergebnisse, ausgehend von einem Startprofil $\boldsymbol{f}^0 = 0$, für das wir das konjugierte Gradienten-Verfahren als Filterung der verallgemeinerten Inversen hatten darstellen können (siehe Ende Abschn. 2.3.3). Auffallend ist, wie gut zumindest während der ersten drei bis vier Iterationsschritte die Ergebnisse des konjugierten Gradienten-Verfahrens und der abgeschnittenen Singulärwertzerlegung miteinander übereinstimmen (vgl. Abb. 11). Beim konjugierte Gradienten-Verfahren erreichen wir die optimale Lösung nach etwa 14 Iterationsschritten, doch kann der optimale Zeitpunkt zum Abbruch des Verfahrens nur dann bestimmt werden, wenn auch das Originalprofil bekannt ist. Führen wir statt dessen eine a posteriori Parameterwahl durch, wie in Abschn. 1.4 beschrieben, so erhalten wir als Näherungslösung die Abb. 20.11. Die mit dem konjugierten Gradienten-Verfahren maximal erreichbare Höhenauflösung entspricht in etwa der von abgeschnittener Singulärwertzerlegung und Tikhonov-Phillips-Regularisierung. Doch kann aufgrund der Nichtlinearität des Verfahrens die Auflösung nicht direkt, sondern nur indirekt z. B. durch Vergleich mit einem anderen Verfahren bekannter Auflösung ermittelt werden.

Die Wahl eines beliebigen Startprofiles $\boldsymbol{f}^0$ für das konjugierte Gradienten-Verfahren ist bei schlecht gestellten Problemen nicht ohne Einfluß auf das Inversionsergebnis. [4] Dies folgt unmittelbar aus der Tatsache, daß wir die Lösung nach m Iterationen (vgl. Abschn. 2.3.3 Gl. 29) darstellen können als

$$\boldsymbol{f}^m = \boldsymbol{f}^0 - P_{m-1}(\boldsymbol{A}^*\boldsymbol{A})\boldsymbol{r}^0.$$

$P_{m-1}(\boldsymbol{A}^*\boldsymbol{A})$ ist hierbei ein Polynom $m-1$.ten Grades in $\boldsymbol{A}^*\boldsymbol{A}$ und $\boldsymbol{r}^0$ der Gradient an der Stelle $\boldsymbol{f}^0$. Treten nun im Startprofil $\boldsymbol{f}^0$ detaillierte Strukturen, wie z. B. eine scharfe Kante auf (siehe Abb. 21), so muß im Prinzip über alle Eigenvektoren der Matrix $\boldsymbol{A}^*\boldsymbol{A}$ aufsummiert, d.h. unendlich tief ($m \to \infty$) iteriert werden, um eine solche, im Originalprofil nicht vorhande Kante wieder „auszubügeln". Da andererseits das konjugierte Gradienten-Verfahren aufgrund von Meßfehlern in den Daten das für alle iterativen Verfahren typische Semikonvergenzverhalten zeigt, darf eine bestimmte Iterationstiefe nicht überschreiten werden. Dies hat zur Folge, daß detaillierte Strukturen im Startprofil zwangsläufig auch im Inversionsergebnis auftreten müssen – gleichsam auf das Inversionsergebnis abgebildet werden (siehe Abb. 21). Als geeignete Startprofile kommen insofern nur hinreichend glatte Höhenprofile in Frage, da ansonsten im Inversionsergebnis ein Informationsgehalt bzw. eine Auflösung suggeriert wird, die im Ergebnis nicht enthalten sein kann.

Neben seiner Glattheit ist eine weitere Voraussetzung für ein gutes Startprofil, daß es nicht allzu weit vom Mittelwert des Originalprofils abweicht. Z. B. werden im vorliegenden Fall die besten Inversionsergebnisse erzielt mit einem Startprofil $\boldsymbol{f}^0 = 3,5$ ppmv, wohingegen die Startprofile $\boldsymbol{f}^0 = 100$ ppmv oder das unphysikalische Startprofil $\boldsymbol{f}^0 = -10$ ppmv zu relativ großen Abweichungen zwischen Inversionsergebnis

[4] Dies gilt im Prinzip für alle iterativen Verfahren.

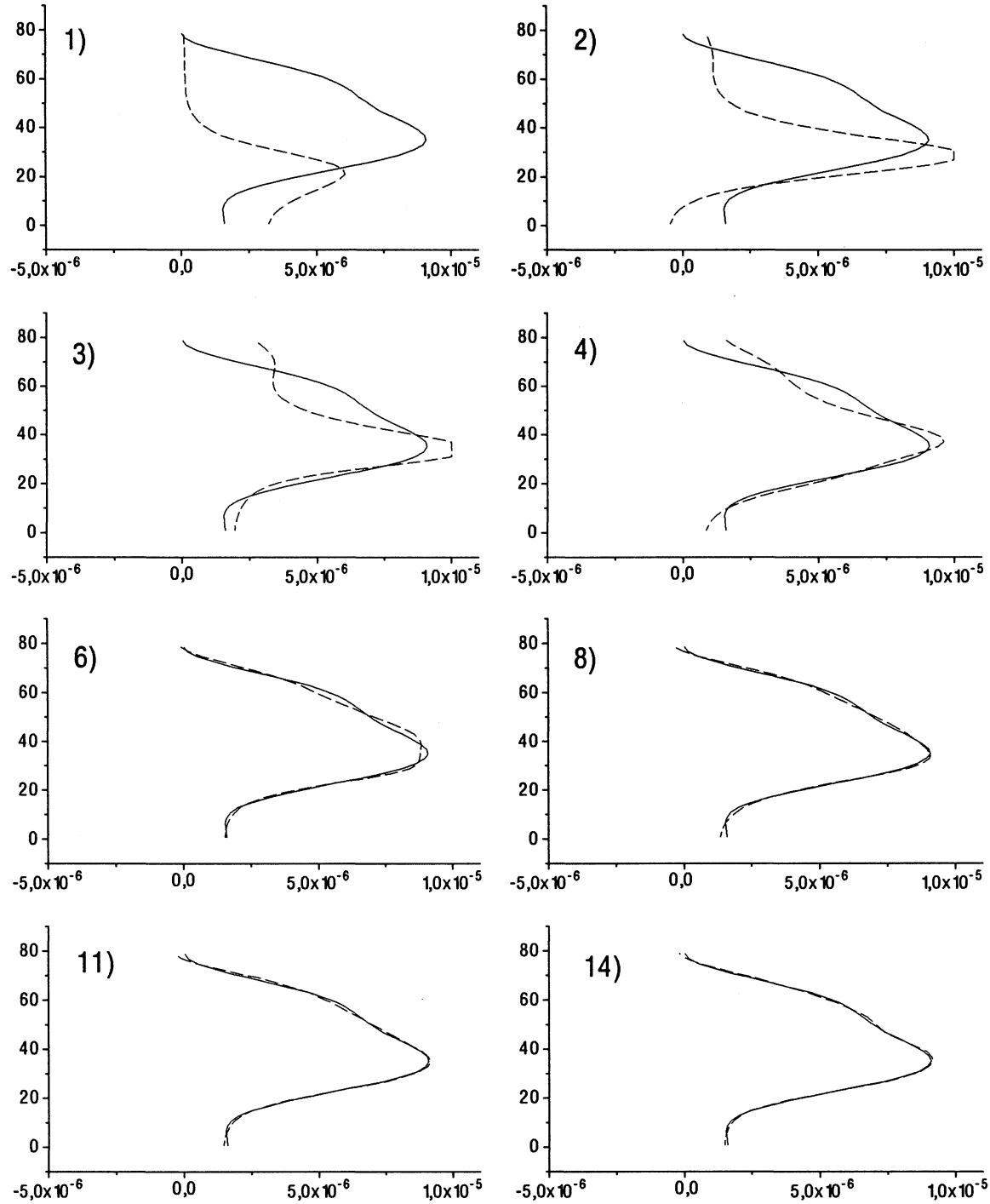

Abbildung 20: Verfahren der konjugierten Gradienten. Dargestellt sind jeweils Originalprofil (durchgezogen) sowie aktuelle Näherungslösung (gestrichelt). Die Nummern kennzeichnen die Anzahl der durchgeführten Iterationsschritte. Das Verfahren wird nach 14 Schritten abgebrochen, da hier die optimale Lösung erreicht ist. Wird statt dessen weiter iteriert, wird das Verfahren instabil.

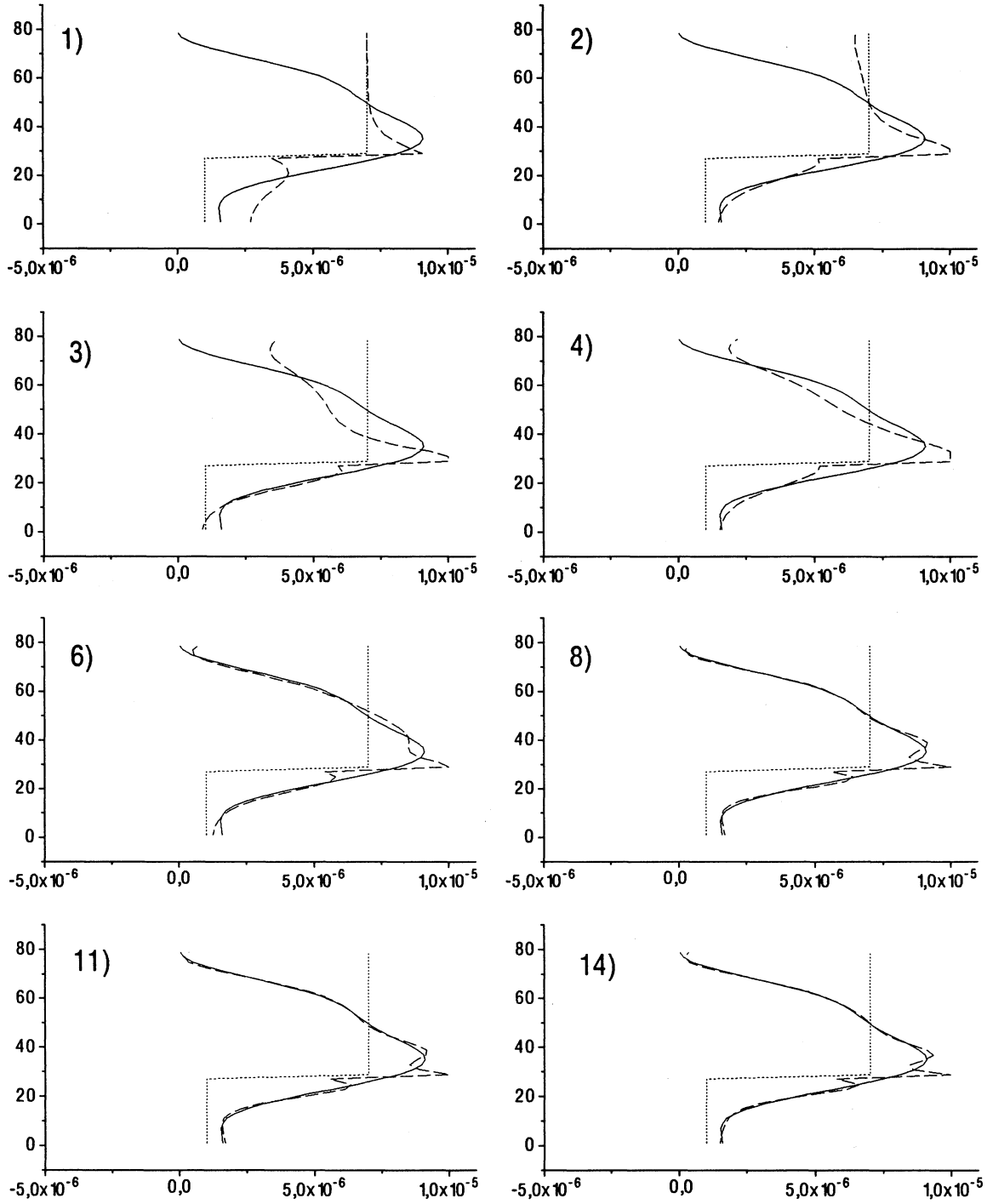

Abbildung 21: Einfluß des Startprofils auf das konjugierte Gradienten–Verfahren. Dargestellt sind jeweils Originalprofil (durchgezogen), momentane Näherungslösung (gestrichelt) sowie Startprofil (gepunktet). Insbesondere im Startprofil angenommene detaillierte Strukturen wie z. B. scharfe „Kanten“, die im Originalprofil nicht vorhanden sind, können im Verlauf des Verfahren nicht vollständig korrigiert werden und werden auf die Näherungslösung abgebildet.

und Originalprofil führen.

3.6 Anwendung des Chahineschen Verfahrens

Die in diesem Abschnitt gezeigten Abbildungen beschränken sich auf eine Anwendung des modifizierten Chahineschen Verfahrens nach Abschn. 2.3.5, da nur diese Verallgemeinerung des Verfahren auf die konkrete Problemstellung mit $\boldsymbol{A}$ als einer 200×40-Matrix angewendet werden kann. Doch wurden nebenher etliche andere Rechnungen mit einer entsprechend quadratischen Matrix $\boldsymbol{A}'$ (siehe Abschn. 2.3.4) durchgeführt, in denen das modifizierte Chahinesche Verfahren sich dem ursprünglichen Verfahren bzgl. der Inversionsergebnisse in allen betrachteten Fällen überlegen zeigte – auch wenn es deutlich langsamer konvergierte. [5]

In Abb. 22 ist das Zustandekommen des invertierten Profils mit dem modifizierten Chahineschen Verfahren und einer Anfangsschätzung $\boldsymbol{f}^0 = 1$ ppmv dargestellt, wobei das Optimum nach etwa 60 Iterationen erreicht wird (bei a posteriori-Parameterwahl mit $r = 1$ nach etwa 55 Iterationen). Danach beginnt die Lösung – wenn auch sehr langsam – zu oszillieren.

Im Vergleich zu den bisher betrachteten Verfahren schneiden das modifizierte Chahinesche Verfahren und das ursprüngliche Chahinesche Verfahren deutlich schlechter ab. So ist es z. B. nicht möglich, das Originalprofil bis in eine Höhe von 80 km wie bei den vorangegangenen Verfahren zu detektieren. Oberhalb von 70 km wird die Lösung instabil, weshalb sich die Berechnungen zur Abb. 22 auf einen Höhenbereich von 0-70 km beschränken. [6] Doch selbst in diesem eingeschränkten Höhenbereich sind die Ergebnisse von abgeschnittener Singulärwertzerlegung (vgl. Abb. 12.11), Tikhonov–Phillips–Regularisierung (vgl. Abb. 17.11) sowie konjugiertem Gradienten–Verfahren (vgl. Abb. 20.14) deutlich besser.

Ein weiterer Nachteil der beiden Chahineschen Verfahren ist ihre im Vergleich zum Verfahren der konjugierten Gradienten stärkere Empfindlichkeit gegenüber detaillierten Strukturen im Startprofil (vgl. hierzu Abb. 23 mit Abb. 21). Beiden Iterationsfolgen liegt dasselbe Startprofil mit einer scharfen Kante in jeweils 30 km Höhe zugrunde. Während jedoch beim Verfahren der konjugierten Gradienten diese im Originalprofil nicht vorhandene Kante im Laufe der Iteration nahezu „ausgebügelt“ wird, wird sie bei den Chahineschen Verfahren nahezu vollständig auf das Endergebnis abgebildet.

Zur „Ehrenrettung“ der Chahineschen Verfahren sei jedoch darauf hingewiesen, daß diese ausgeprägte Sensibilität der Verfahren gegenüber der Wahl des Startprofils eine ausgesprochene Besonderheit der bodengebundenen Fernerkundung darstellt

[5] Allerdings ist diese Abnahme der Konvergenzgeschwindigkeit nicht so stark, daß deswegen das Vefahren ineffizient wäre.

[6] Zur Ursache dieser Instabilität siehe die Ausführungen zur Anwendungen des Chahineschen Verfahrens auf bodengebundene Mesungen in Abschn. 2.3.4.

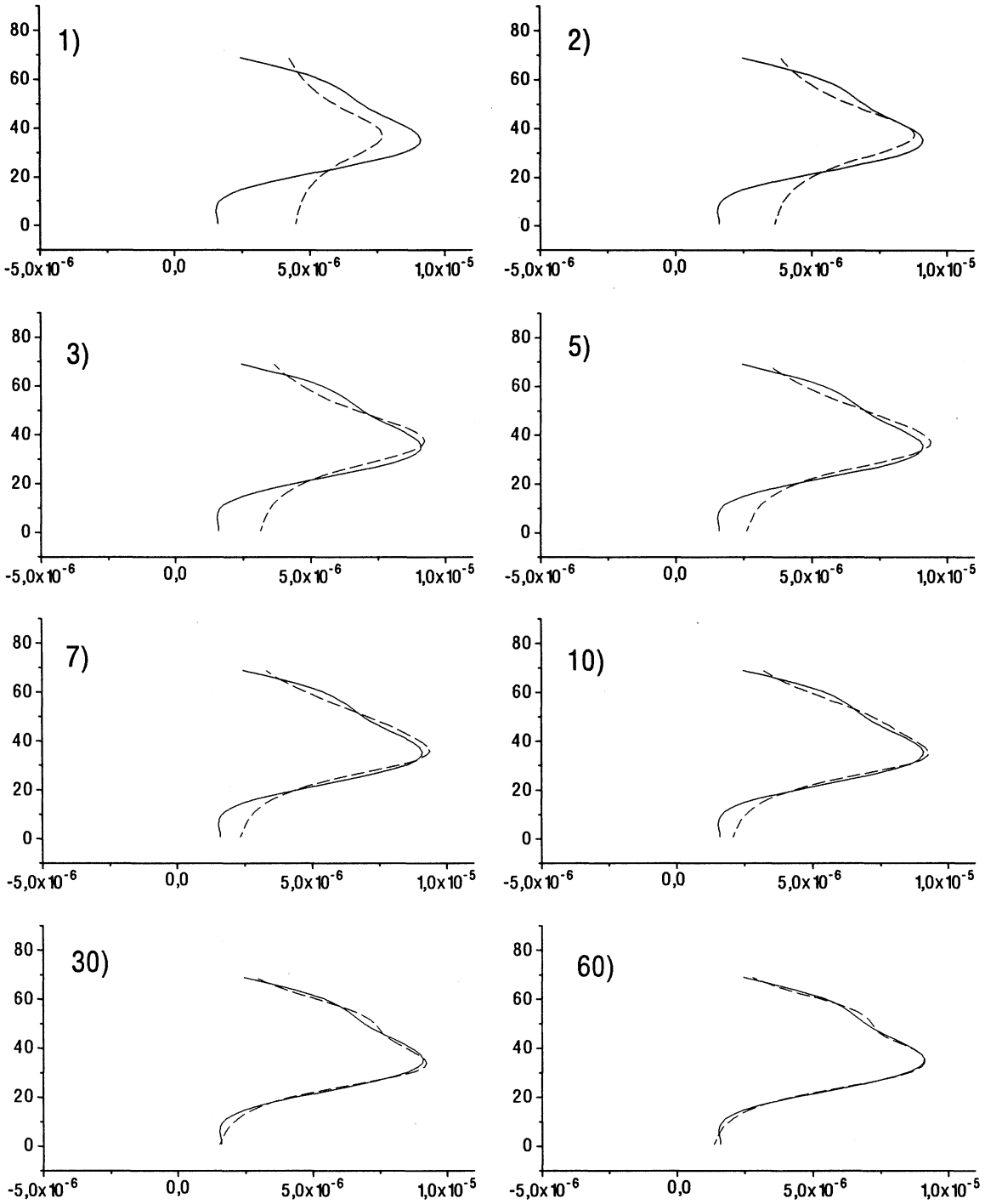

Abbildung 22: Modifiziertes Chahinesches Verfahren mit einer Anfangsschätzung $f^0 = 1$. Dargestellt sind jeweils Originalprofil (durchgezogen) sowie die aktuelle Näherungslösung (gestrichelt). Die Nummern kennzeichnen die Anzahl der durchgeführten Iterationsschritte. Nach etwa 60 Iterationen ist die optimale Lösung erreicht, weshalb das Verfahren abgebrochen wird.

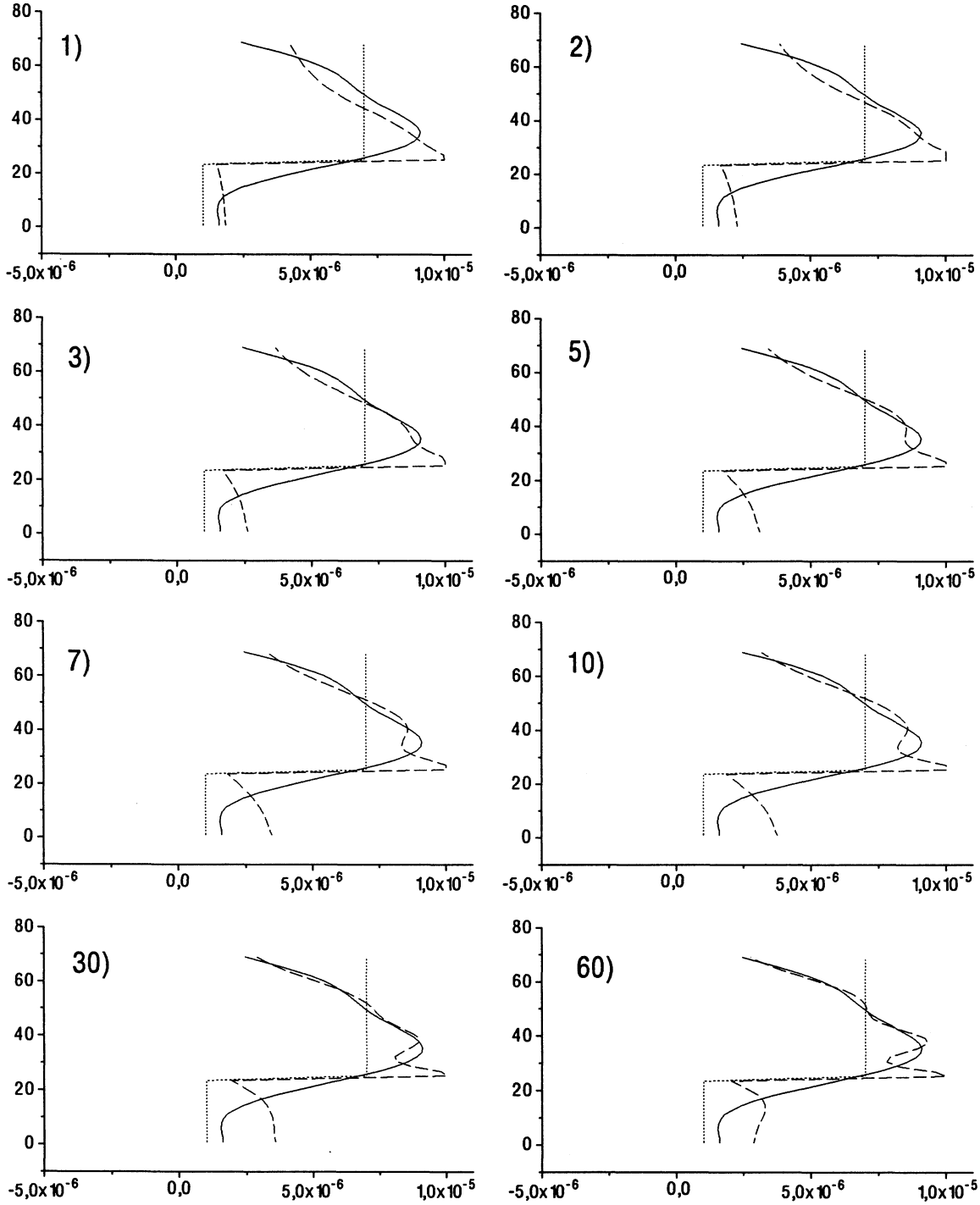

Abbildung 23: Einfluß des Startprofils auf das modifizierte Chahinesche Verfahren. Dargestellt sind jeweils Originalprofil (durchgezogen), aktuelle Näherungslösung (gestrichelt) sowie Startprofil (gepunktet). Gegenüber dem konjugierten Gradienten–Verfahren reagiert das modifizierte Chahinesche Verfahren weitaus sensibler auf eine falsche Wahl des Startprofils, wie aus dem Vergleich mit der Abb. 21 deutlich wird.

(Blickrichtung nach oben). Denn die Chahineschen Verfahren dürfen nur auf Bandmatrizen angewandt werden, die ein ausgeprägtes Maximum längs ihrer Diagonalen aufweisen – eine Bedingung, die im Falle bodengebundener Messungen nicht erfüllt ist. Erst durch den in Abschn. 2.3.4 erwähnten Trick, nämlich die Schaffung eines obige Bedingungen erfüllenden Ersatzproblems, was jedoch andererseits eine erhebliche Verschlechterung des „Signal–zu–Rausch–Verhältnisses" bewirkt, wird eine Anwendung der Chahineschen Verfahren überhaupt möglich.

3.7 Anwendung der stochastischen Verfahren

In der Praxis bereitet die Prämisse der stochastischen Verfahren nach ausreichend a priori Information und die damit verbundene Initialisierung der Kovarianzmatrix $\boldsymbol{S}_f$ Schwierigkeiten. Zumindest bei der Fernerkundung meteorologischer Parameter stehen a priori meist nur sehr wenige Höhenprofile $\boldsymbol{f}$ zur Verfügung, so daß aufgrund dieser wenigen Daten kaum Stochastik betrieben werden kann. Überdies ist die Annahme, daß die Wahrscheinlichkeit für das Auftreten einer bestimmten Lösung bereits vor Durchführung der aktuellen Messung aus der Verteilung dieser a priori Höhenprofile ermittelt werden kann, oft nicht zutreffend. Z. B. können die vorhandenen a priori Höhenprofile eine völlig andere Jahreszeit, eine andere Tageszeit, einen anderen Breitengrad oder vollkommen andere Wetterverhältnisse als die bei Durchführung der Messung aktuellen repräsentieren. Zwangsläufig ist deshalb die Annahme einer Gaußschen Häufigkeitsverteilung meteorologischer Profile um einen vorgegebenen Mittelwert mit den zum Zeitpunkt der Messung herrschenden meteorologischen Bedingungen nur schwer in Einklang zu bringen. Letztlich führt dies dazu, daß wir statt des tatsächlichen Höhenprofils eine stark in Richtung a priori Profile verzerrte Näherungslösung erhalten.

Um dies zu verdeutlichen, betrachten wir die in Abb. 24.a dargestellten, willkürlich angenommenen a priori Volumenmischungsverhältnisprofile, die alle mehr oder weniger glatt sind und stark um unser Originalprofil streuen. Aus diesen berechnen wir Erwartungswert $\overline{\boldsymbol{f}}$ (siehe Abb. 24.b) und Kovarianzmatrix $\boldsymbol{S}_f$ gemäß Gl. (32). Für die Initialisierung der Fehlerkovarianzmatrix $\boldsymbol{S}_\varepsilon$ nehmen wir an, daß die Meßfehler unterschiedlicher spektraler Kanäle nicht miteinander korreliert sind und wie bisher eine Gaußsche Häufigkeitsverteilung mit einer Standardabweichung von 0,01 K besitzen. $\boldsymbol{S}_\varepsilon$ ist damit eine reine Diagonalmatrix, deren Diagonalelemente die quadrierte Standarabweichung der zugehörigen Komponente des Meßfehlervektors $\boldsymbol{\varepsilon}$ enthalten.

Die Abb. 25.a zeigt die aufgrund obiger Verteilung der a priori Profile ermittelte Näherungslösung $\boldsymbol{f}_{opt}$ samt zugehörigen Fehlerbalken, wobei die Berechnung mit Hilfe der praktischen Inversionsgleichung (38) bzw. der Gl. (39) erfolgt. Im Vergleich zu den bisherigen Verfahren ist dieses Inversionsergebnis weniger befriedigend, trotzdem das Originalprofil nicht allzu stark vom bisherigen Mittelwert abweicht (Abb. 24.b).

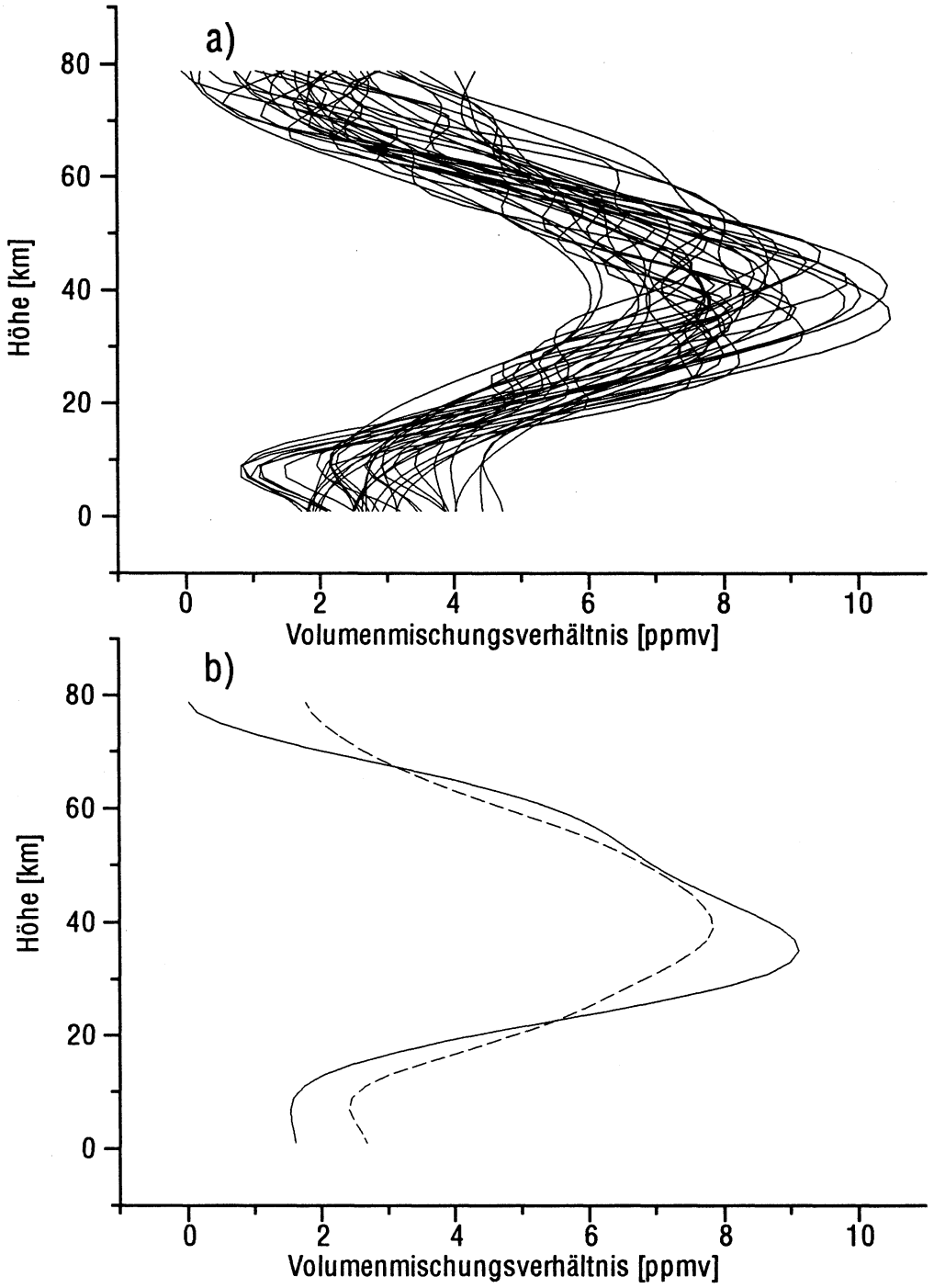

Abbildung 24: A priori Information für die Anwendung der stochastischen Verfahren. In a) 40 willkürlich angenommene Volumenmischungsverhältnisprofile von Ozon, in b) der sich aus obiger Verteilung ergebende Mittel- oder Erwartungswert (gestrichelt). Zum Vergleich ist neben dem Mittelwert auch das zu invertierende Originalprofil (durchgezogen) mit dargestellt.

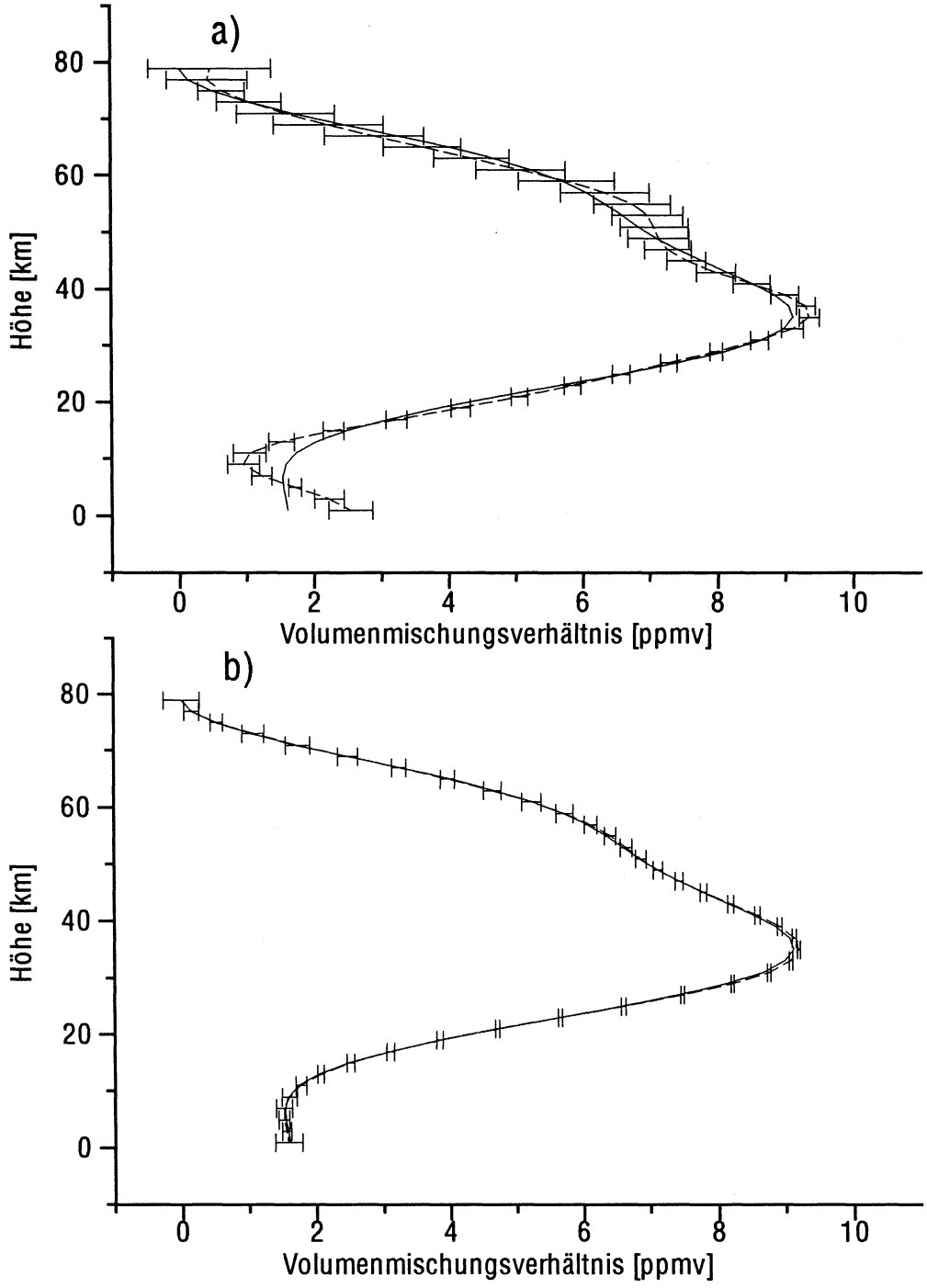

Abbildung 25: Ergebnisse der stochastischen Verfahren. In a) auf Grund der in Abb. 24 willkürlich angenommenen a priori Information berechnete Näherungslösung. In b) durch Initialisierung der Kovarianzmatrix $\boldsymbol{S}_f$ von Hand erzieltes Ergebnis. Beiden Lösungen liegt der gleiche Erwartungswert $\overline{\boldsymbol{f}}$ (Abb. 24.b) zugrunde. Im Gegensatz zur Abb. 13, in der die Fehlerbalken nur den Datenfehler bezeichnen, liegt den Fehlerbalken hier der Gesamtfehler zugrunde.

Dies ist in erster Linie auf die aus obiger Verteilung der a priori Profile folgende Initialisierung der Kovarianzmatrix $\boldsymbol{S}_f$ zurückzuführen. Z. B. weisen alle a priori Profile der Abb. 24.a, welche sich in Bodennähe durch ein besonders niedriges Volumenmischungsverhältnis auszeichnen, in etwa 10 km Höhe einen „Knick" auf. Diese auch in der Kovarianzmatrix $\boldsymbol{S}_f$ enthaltene Information bewirkt, daß dieser Knick in seiner From nahezu unverändert auf die Näherungslösung abgebildet wird, für die wir aufgrund der aktuellen Daten $\boldsymbol{g}_\mathcal{E}$ in Bodennähe ein ebenfalls niedriges Volumenmischungsverhältnis erwarten (siehe Abb. 25.a). Wie dieses Beispiel zeigt, muß deshalb besondere Sorgfalt auf die Auswahl der a priori Profile verwendet werden, die für die aktuellen klimatologischen Verhältnisse repräsentativ sein sollten.

Existieren keine repräsentativen a priori Profile, so besteht immerhin noch die Möglichkeit, Mittelwert $\overline{\boldsymbol{f}}$ und Kovarianzmatrix $\boldsymbol{S}_f$ von Hand zu initialisieren. Als besonders zweckmäßig erweist sich hierbei die Initialisierung der Matrix $\boldsymbol{S}_f$ als Bandmatrix mit einem Band von der Form einer Gaußschen Glockenkurve. Für das Maximum längs der Diagonalen wählen wir einen Wert von $(2\ \text{ppmv})^2$; dies bedeutet, daß die Volumenmischungsverhältnisse fiktiv angenommener a priori Profile um $\pm$ 2 ppmv um den vorgegebenen Mittelwert $\overline{\boldsymbol{f}}$ streuen. Für die Breite des Bandes (gleich der zweifachen Standardabweichung der Gaußfunktion) hingegen wählen wir einen Wert von 6 (Höhenschichten). Die Breite des Bandes gibt an, wieviele benachbarte Höhenschichten jeweils miteinander korreliert sind; sie definiert eine Art "Steifigkeit" des vorgegebenen Mittelwerts $\overline{\boldsymbol{f}}$ gegenüber Verformungen.

Das Ergebnis dieser Initialisierung zeigt die Abb. 25.b., wobei als Mittelwert das gemittelte Profil von oben (Abb. 24.b) verwendet wird. Im Gegensatz zur Abb. 25.a ist dieses Inversionsergebnis nahezu optimal. Doch dürfen wir nicht verkennen, daß wir durch die Initialisierung der Matrix $\boldsymbol{S}_f$ den eigentlichen Zweck der stochastischen Verfahren, nämlich die automatische Bestimmung des optimalen Regularisierungsparameters auf der Basis der vorhandenen a priori verfehlt haben. Schließlich müssen wir Maximalwert und Breite der zur Initialisierung der Kovarianzmatrix $\boldsymbol{S}_f$ verwandten Gaußfunktion vorgeben, haben also wiederum zwei Regularisierungsparameter, mit denen wir spielen können. Im Prinzip betreiben wir deshalb lediglich eine spezielle Art von Tikhonov–Phillips–Regularisierung (vgl. Ausführungen Ende des Abschnitts 2.4.2).

Für die Wahl des Mittelwerts $\overline{\boldsymbol{f}}$ gilt im wesentlichen dasselbe wie für die Wahl des Startprofils bei den iterativen Verfahren (vgl. Abschn. 3.5). Sofern das vorgegebene Mittelwertprofil $\overline{\boldsymbol{f}}$ hinreichend glatt ist und nicht um Größenordnungen vom aktuellen Volumenmischungsverhältnisprofil abweicht, ist das Ergebnis der Inversion $\boldsymbol{f}_{opt}$ von diesem weitgehend unabhängig. Z. B. ergibt sich auch für die Wahl $\overline{\boldsymbol{f}} = 0$ ein nahezu identisches Inversionsergebnis.

Bzgl. der erreichbaren Auflösung schneiden die stochastischen Verfahren im Vergleich zu abgeschnittener Singulärwertzerlegung oder Tikhonov–Phillips–Regularisierung geringfügig schlechter ab (vgl. Abb. 26 mit Abb. 15 bzw. Abb. 18),

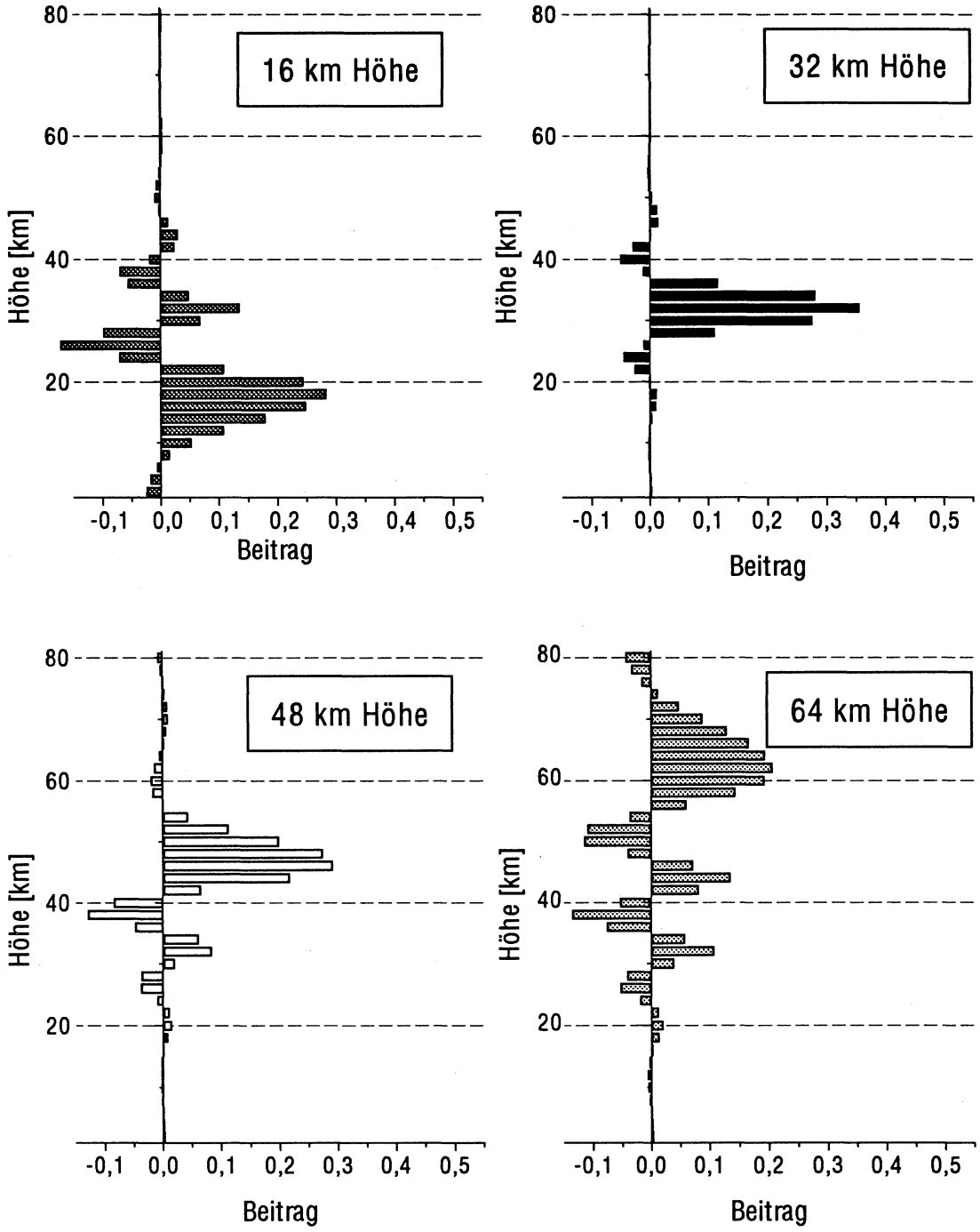

Abbildung 26: Ausgewählte Zeilen der Auflösungsmatrix für die stochastischen Verfahren. Ein jedes Balkendiagramm gibt an, wie stark der Wert des Originalprofils in einer bestimmten Höhenschicht zu dem von uns ermittelten Wert der Näherungslösung in der abgegebenen Höhe beiträgt. In 16 km Höhe beträgt die Auflösung 12 km, in 32 km Höhe 32 km Höhe ungefähr 8 km, in 48 km Höhe 10 km sowie in 64 km Höhe 14-15 km.

wobei vermehrt Überschwinger auftreten. Vermutlich hängt dies damit zusammen, daß die Initialisierung der Kovarianzmatrix von Hand doch relativ willkürlich ist, so daß z. T. auch Fehlinformation über das zu ermittelnde Höhenprofil mit in die Inversion einfließt.

Zusammenfassend läßt sich deshalb über die stochastischen Verfahren feststellen, daß sie zwar von der Konzeption her das allgemeinste und vielseitigste aller Regularisierungsverfahren darstellen, andererseits in der (meteorologischen) Praxis jedoch gerade wegen ihrer Vielseitigkeit und der damit verbundenen Vielzahl der einzustellenden Parameter [7] oft Schwierigkeiten bereiten. Gerade hier ist deshalb die Gefahr, durch eine falsche Anwendung des Verfahrens das Inversionsergebnis zu verfälschen, am größten.

3.8 Anwendung des Backus–Gilbert–Verfahrens

Wie bereits Ende des Abschnitts 2.5 bei der Vorstellung des Backus–Gilbert–Verfahrens festgestellt, wächst der numerische Aufwand dieses Verfahrens linear mit der Anzahl m der Höhenstützstellen f_j und quadratisch mit der Anzahl n der spektralen Stützstellen $g_{\varepsilon i}$. Aufgrund dessen beschränken wir uns bei der Anwendung des Verfahrens auf das Invertieren einer 20×20 Matrix. [8] In Abb. 27 sind die Ergebnisse dieser Inversionsrechnungen für verschiedene Regularisierungsparameter γ dargestellt. Ausgehend von einem relativ kleinen Regularisierungsparameter von 1×10^{-8} in Abb. 27.1 multiplizieren wir γ in jedem weiteren Schritt mit einem Faktor 8 und erhalten so nach 7 Schritten mit einem Regularisierungsparameter $\gamma = 0,168$ die optimale Lösung. [9] Nach 8 Schritten wird die Lösung instabil.

Im Vergleich zu den bisherigen Verfahren schneidet das Backus–Gilbert–Verfahren deutlich schlechter ab, was teilweise auf den geringeren Informationsgehalt des 20×20 Gleichungssystems gegenüber dem bisher betrachteten 200×40 Gleichungssystem zurückzuführen, zum Großteil aber durch das Verfahren selbst bedingt ist. Selbst für die optimale Lösung bestehen noch beträchtliche Abweichungen zwischen Näherungslösung und Originalprofil.

Im Gegensatz zu den bisher betrachteten Verfahren, treten in der Auflösungsmatrix des Backus–Gilbert–Verfahrens nahezu keine Überschwinger auf (siehe Abb. 28). Das Rayleigh Kriterium kann deshalb zur Definition der erreichbaren Auflösung nicht angewendet werden. Wir verwenden statt dessen die in Gl. (45) definierte spread'-Funktion und erhalten in 16 km eine Auflösung von ungefähr 12 km, in 32 km Höhe eine Auflösung von 6-7 km, in 48 km Höhe eine Auflösung von 9 km sowie in 64 km

[7] Neben dem Mittelwert $\overline{f}$ muß jedes Element der Kovarianzmatrix S_f initialisiert werden, sei es aufgrund der vorhandenen a priori Profile oder von Hand.

[8] Hierzu werden Kanäle entsprechend zusammengefaßt.

[9] Auf eine Auftragung von Auflösung gegenüber Datenfehler entsprechend der Abb. 7 wird verzichtet.

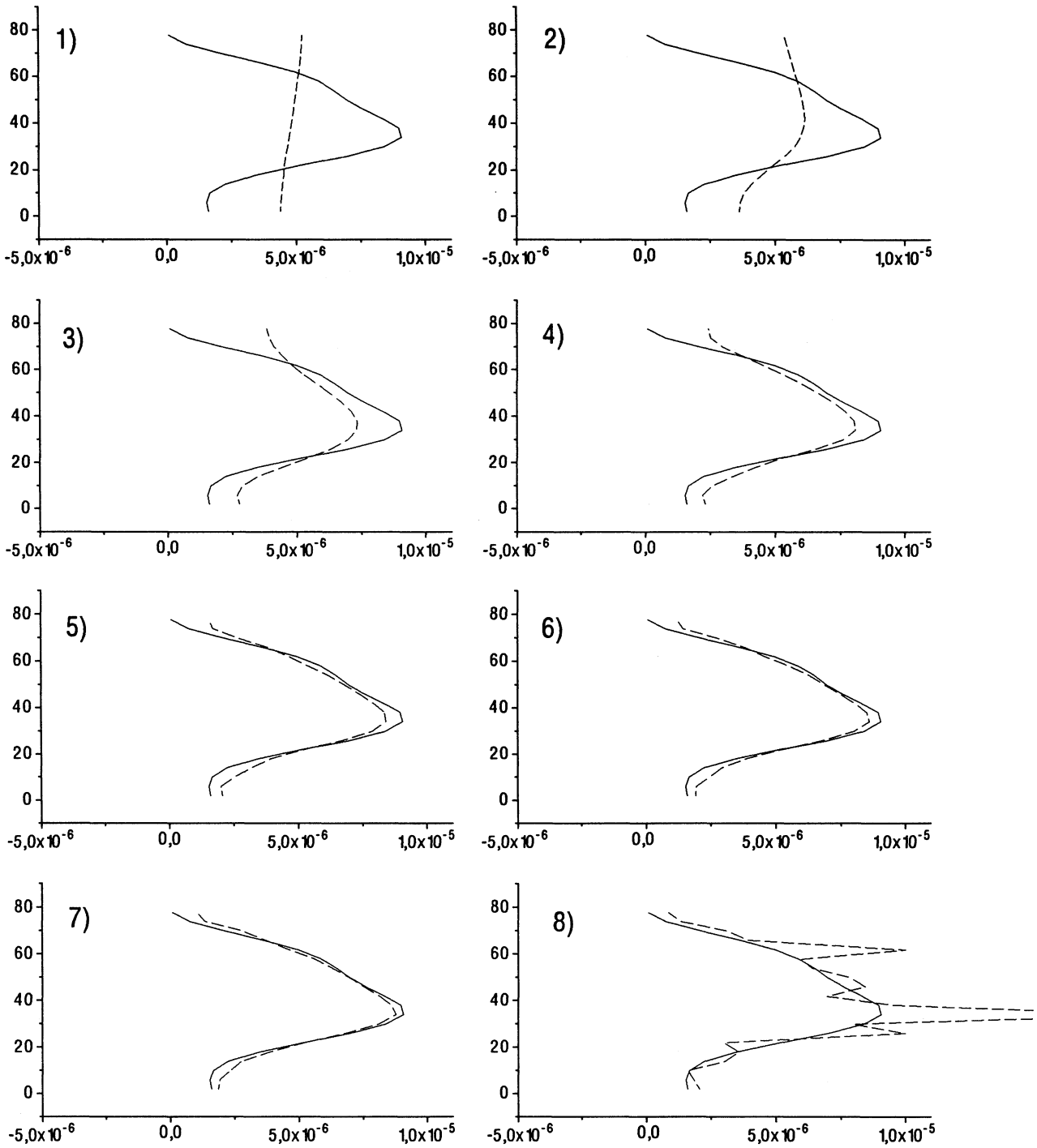

Abbildung 27: Backus–Gilbert–Verfahren. Dargestellt sind jeweils Originalprofil (durchgezogen) sowie die aktuelle Näherungslösung (gestrichelt). Im Vergleich zu den bisher betrachteten Verfahren sind die Ergebnisse des Backus–Gilbert–Verfahrens unbefriedigend.

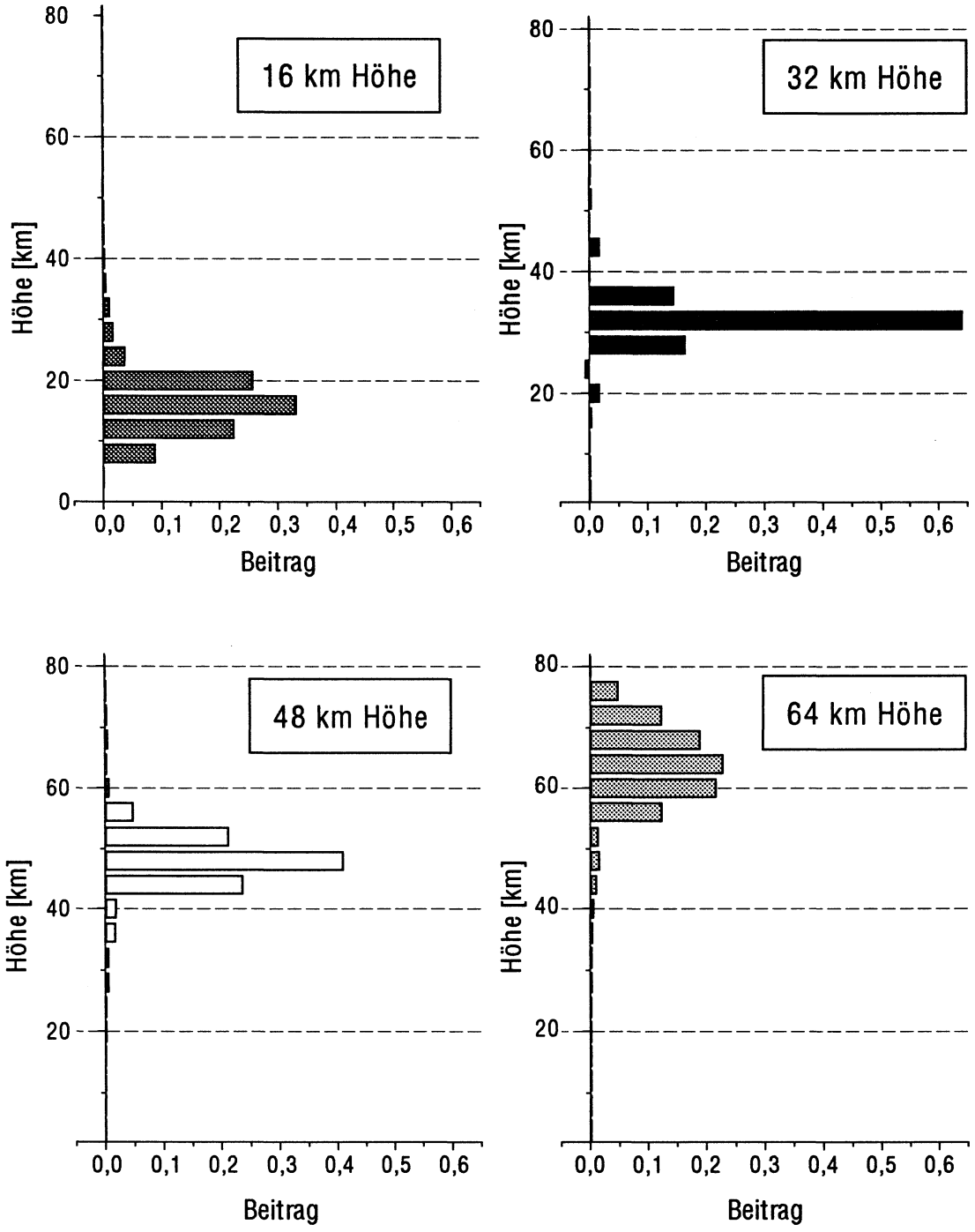

Abbildung 28: Auflösung des Backus–Gilbert–Verfahrens. Ein jedes Balkendiagramm gibt an, wie stark der Wert des Originalprofils in einer bestimmten Höhenschicht (Dicke 4km) zu dem von uns ermittelten Wert der Näherungslösung in der angegebenen Höhe beiträgt. Im Gegensatz zu den bisher betrachteten Balkendiagrammen treten beim Backus–Gilbert–Verfahren nahezu keine Überschwinger auf.

Höhe eine Auflösung von 14 km. Diese Werte entsprechen in etwa denen von abgeschnittener Singulärwertzerlegung und Tikhonov–Phillips–Regularisierung. Vergleicht man die angegebenen Werte mit den in Abb. 28 dargestellten Balkendiagrammen, so entspricht die erreichbare Auflösung der Halbwertsbreite der „Beitragsfunktion".

Kapitel 4

Beurteilung der Regularisierungsverfahren

Bei der Beurteilung der betrachteten Regularisierungsverfahren bilden sowohl die Effizienz des Verfahrens als auch die Güte der erhaltenen Inversionsergebnisse die Schwerpunkte. Als weiteres, wichtiges Kriterium verwenden wir den Umfang der für die Anwendung des Regularisierungsverfahrens benötigten a priori Information. Hierbei werten wir es als positiv, wenn ein Regularisierungsverfahren mit möglichst wenig a priori Information auskommt. Zwar bedeutet dies eine wesentliche Einschränkung der Handlungsfreiheit für den Anwender des Verfahrens[1], andererseits befindet er sich jedoch immer auf der sicheren Seite und läuft nicht Gefahr, durch eine irrtümlich falsche Initialisierung das Inversionsergebnis zu verfälschen.

Ein Überblick über die betrachteten Regularisierungsverfahren ist in Tabelle 4 zusammengestellt, wobei zusätzlich mit angegeben wird, ob es sich um ein lineares oder nichtlineares bzw. um ein iteratives Regularisierungsverfahren handelt. Dies ist insofern von Bedeutung, als die linearen Verfahren eine umfassende Diagnostik des Fehlers und der erreichten Auflösung eines Inversionsergebnisses erlauben (siehe Ausführungen in Kapitel 1), die nichtlinearen Verfahren hingegen nicht.

Vergleicht man die Verfahren bzgl. ihrer Inversionsergebnisse, so schneidet neben dem Backus–Gilbert–Verfahren das Chahinesche Verfahren besonders schlecht ab; auch das modifizierte Chahinesche Verfahren liefert allenfalls befriedigende Inversionsergebnisse. Bei den Chahineschen Verfahren ist dieses schlechte Abschneiden darauf zurückzuführen, daß die Verfahren auf die Inversion bodengebundener Messungen nicht direkt angewendet werden können, sondern daß statt dessen ein Ersatzproblem gelöst werden muß (siehe die Ausführungen in Abschn. 2.3.4). Die restlichen Regularisierungsverfahren liefern alle gleich gute Resultate, wobei die Tikhonov–Phillips–Regularisierung im betrachteten Fall besonders gut abschneidet.

Bezieht man die Effizienz der Regularisierungsverfahren in die Betrachtung mit ein, so fallen Landweber–Iteration und Backus–Gilbert–Verfahren negativ auf; auch

[1] Z. B. kann der Anwender bei der abgeschnittenen Singulärwertzerlegung lediglich den Regularisierungsparameter γ variieren, wohingegen er bei den iterativen Iterativen Verfahren die Möglichkeit hat, neben der Iterationstiefe auch das Anfangsprofil nach seinem Ermessen festzulegen.

Regularisierungs-verfahren	*linear*	*iterativ*	*wenig a priori Information*	*Effizienz*	*Inversions-ergebnisse*
abgeschnittene Singulärwertzerlegung	×		+++	++	++
Tikhonov–Phillips–Regularisierung	×		++	++	+++
Landweber–Iteration	×	×	+	−−−	++
Landweber–Iteration nach Strand	×	×	+	++	++
Konjugierte Gradienten		×	+	+++	++
Chahinesches Verfahren		×	−[2]	+++	−−
Modifiziertes Chahinesches Verfahren		×	−[2]	+	+
Stochastische Verfahren	×		−−−	++	++
Backus–Gilbert–Verfahren	×		+++	−−	−−−

Tabelle 1: Übersicht über die Regularisierungsverfahren. In dieser Übersicht wurden die Regularisierungsverfahren bzgl. ihrer Eigenschaften bewertet. Die Skala reicht hierbei von „+ + +“ bis „− − −“. Es bleibt anzumerken, daß diese Beurteilung der Regularisierungsverfahren sich einzig auf die untersuchte Anwendung bezieht, also nicht notwendigerweise auch auf andere Anwendungsbereiche so zutreffen muß.

das modifizierte Chahinesche Verfahren konvergiert nur mäßig schnell. Besonders effizient hingegen sind das Verfahren der konjugierten Gradienten und das Chahinesche Verfahren.

Zusammenfassend läßt sich feststellen, daß bei hoher Effizienz des betrachteten Regularisierungsverfahrens mit der abgeschnittenen Singulärwertzerlegung, mit der Tikhonov–Phillips–Regularisierung, mit dem Verfahren der konjugierten Gradienten oder mit den stochastischen Verfahren qualitativ hochwertige Inversionsergebnisse erzielt werden können. Die Erweiterung der Landweber–Iteration nach Strand

[2]Normalerweise benötigten die Chahineschen Verfahren keine oder nur sehr wenig a priori Information (nämlich allenfalls Informationen über die optimale Iterationstiefe bei der a priori Parameterwahl). Im Falle der bodengebundenen Millimeterwellenradiometrie hängt das Inversionsergebnis jedoch zusätzlich so stark ab von dem gewählten Startprofil (siehe Abschn. 3.6), daß wir die Chahineschen Verfahren in diesem Punkt mit einem (−) bewerten.

wird nicht gesondert aufgeführt, da sie als eine Art *iterative Tikhonov–Phillips–Regularisierung* in der Tikhonov–Phillips–Regularisierung mit eingeschlossen ist (siehe Abschn. 2.3.2).

Welches dieser nahezu gleichwertigen Regularisierungsverfahren im Einzelfall bevorzugt werden sollte, hängt ab von der konkreten Aufgabenstellung und dem Umfang der vorhandenen a priori Information. Doch sollte zumindest einmal eine abgeschnittene Singulärwertzerlegung durchgeführt werden, da sie aus mathematischer Sicht das sauberste Regularisierungsverfahren darstellt und überdies eine Klassifizierung des betrachteten Problems erlaubt.

Für das konjugierte Gradienten–Verfahren spricht die hohe Effizienz des Verfahrens, die es für eine Anwendung auf nichtlineare Probleme, welche durch sukzessives Betrachten des linearisierten Problems (Newton–Iteration) gelöst werden können, geradezu prädestinieren. Gegen das konjugierte Gradienten–Verfahren wiederum spricht, daß es als nichtlineares Verfahren keine Aussage über Fehler und erreichte Auflösung erlaubt. Doch besteht die Möglichkeit, im Anschluß an das Lösen des nichtlinearen Problems mit dem konjugierten Gradienten–Verfahren ein lineares Regularisierungsverfahren anzuwenden und so eine umfassende Fehlerdiagnostik samt Aussage über die erreichbare Auflösung zu erhalten.

Die Ergebnisse der stochastischen Verfahren sollten stets mit einer gewissen Skepsis zur Kenntnis genommen werden, da immer erst geprüft werden muß, inwieweit die verwendete a priori Information für die klimatologischen Verhältnisse zum Zeitpunkt der Messung repräsentativ ist. Ist dies der Fall, so liefern die stochastischen Verfahren zuverlässige Inversionsergebnisse samt umfassender Fehlerdiagnostik.

Diese Beurteilung der Regularisierungsverfahren bezieht sich zunächst nur auf die untersuchte Anwendung, nämlich die bodengebundene Millimeterwellenradiometrie. Sicherlich sind einige der Aussagen auch allgemeingültig, jedoch kann die Gesamtbewertung der Verfahren bei anderen Fernerkundungsverfahren anders ausfallen; dies gilt insbesondere für die Chahineschen Verfahren bei einer Sondierung der Atmosphäre vom Satelliten aus.

Anhang A

Nebenrechnungen

A.1 Zur Berechnung der verallgemeinerten Inversen

In Abschn. 1.2 ist der Ausdruck

$$\begin{aligned} ||\boldsymbol{A}\boldsymbol{f}-\boldsymbol{g}||^2 &= (\boldsymbol{A}\boldsymbol{f}-\boldsymbol{g})^*(\boldsymbol{A}\boldsymbol{f}-\boldsymbol{g}) \\ &= (\boldsymbol{f}^*\boldsymbol{A}^*-\boldsymbol{g}^*)(\boldsymbol{A}\boldsymbol{f}-\boldsymbol{g}) \\ &= \boldsymbol{f}^*\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{f}-\boldsymbol{f}^*\boldsymbol{A}^*\boldsymbol{g}-\boldsymbol{g}^*\boldsymbol{A}\boldsymbol{f}+\boldsymbol{g}^*\boldsymbol{g} \end{aligned}$$

bzgl. $\boldsymbol{f}$ zu minimieren. Einfaches komponentenweises Differenzieren[1]

$$\begin{aligned} \partial||\boldsymbol{A}\boldsymbol{f}-\boldsymbol{g}||^2/\partial f_k &= \boldsymbol{e}_k^*\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{f}+\boldsymbol{f}^*\boldsymbol{A}^*\boldsymbol{A}\,\boldsymbol{e}_k-\boldsymbol{e}_k^*\boldsymbol{A}^*\boldsymbol{g}-\boldsymbol{g}^*\boldsymbol{A}\boldsymbol{e}_k \\ &= \boldsymbol{e}_k^*(\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{f}-\boldsymbol{A}^*\boldsymbol{g})+(\boldsymbol{f}^*\boldsymbol{A}^*\boldsymbol{A}-\boldsymbol{g}^*\boldsymbol{A})\boldsymbol{e}_k \\ &= \boldsymbol{e}_k^*(\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{f}-\boldsymbol{A}^*\boldsymbol{g})+(\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{f}-\boldsymbol{A}^*\boldsymbol{g})^*\boldsymbol{e}_k \\ &= 2\boldsymbol{e}_k^*(\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{f}-\boldsymbol{A}^*\boldsymbol{g}) \end{aligned}$$

und anschließendes zu Null setzen liefert

$$\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{f}-\boldsymbol{A}^*\boldsymbol{g}=0.$$

A.2 Zur Tikhonov–Phillips Regularisierung

In Abschn. 2.2 ist das Funktional (15)

$$\begin{aligned} J_\gamma(\boldsymbol{f}_\gamma) &= ||\boldsymbol{A}\boldsymbol{f}_\gamma-\boldsymbol{g}_\mathcal{E}||^2+\gamma^2||\boldsymbol{B}\boldsymbol{f}_\gamma-\boldsymbol{r}||^2 \\ &= (\boldsymbol{A}\boldsymbol{f}_\gamma-\boldsymbol{g}_\mathcal{E})^*(\boldsymbol{A}\boldsymbol{f}_\gamma-\boldsymbol{g}_\mathcal{E})+\gamma^2(\boldsymbol{B}\boldsymbol{f}_\gamma-\boldsymbol{r})^*(\boldsymbol{B}\boldsymbol{f}_\gamma-\boldsymbol{r}) \\ &= \boldsymbol{f}_\gamma^*\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{f}_\gamma-\boldsymbol{f}_\gamma^*\boldsymbol{A}^*\boldsymbol{g}_\mathcal{E}-\boldsymbol{g}_\mathcal{E}^*\boldsymbol{A}\boldsymbol{f}_\gamma+\boldsymbol{g}_\mathcal{E}^*\boldsymbol{g}_\mathcal{E}+ \\ &\quad \gamma^2\left[\boldsymbol{f}_\gamma^*\boldsymbol{B}^*\boldsymbol{B}\boldsymbol{f}_\gamma-\boldsymbol{f}_\gamma^*\boldsymbol{B}^*\boldsymbol{r}-\boldsymbol{r}^*\boldsymbol{B}\boldsymbol{f}_\gamma+\boldsymbol{r}^*\boldsymbol{r}\right] \end{aligned}$$

[1]Der Vektor $\boldsymbol{e}_k$ ist der k-te Einheitsvektor.

bezüglich $\boldsymbol{f}_\gamma$ zu minimieren. Komponentenweises Differenzieren ergibt

$$\begin{aligned}
\partial J_\gamma(\boldsymbol{f}_\gamma)/\partial f_{\gamma k} &= \boldsymbol{e}_k^*\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{f}_\gamma + \boldsymbol{f}_\gamma^*\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{e}_k - \boldsymbol{e}_k^*\boldsymbol{A}^*\boldsymbol{g}_\mathcal{E} - \boldsymbol{g}_\mathcal{E}^*\boldsymbol{A}\boldsymbol{e}_k + \\
&\quad \gamma^2\left[\boldsymbol{e}_k^*\boldsymbol{B}^*\boldsymbol{B}\boldsymbol{f}_\gamma + \boldsymbol{f}_\gamma^*\boldsymbol{B}^*\boldsymbol{B}\boldsymbol{e}_k - \boldsymbol{e}_k^*\boldsymbol{B}^*\boldsymbol{r} - \boldsymbol{r}^*\boldsymbol{B}\boldsymbol{e}_k\right] \\
&= \boldsymbol{e}_k^*\left[\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{f}_\gamma - \boldsymbol{A}^*\boldsymbol{g}_\mathcal{E} + \gamma^2\boldsymbol{B}^*\boldsymbol{B}\boldsymbol{f}_\gamma - \gamma^2\boldsymbol{B}^*\boldsymbol{r}\right] + \\
&\quad \left[\boldsymbol{f}_\gamma^*\boldsymbol{A}^*\boldsymbol{A} - \boldsymbol{g}_\mathcal{E}^*\boldsymbol{A} + \gamma^2\boldsymbol{f}_\gamma^*\boldsymbol{B}^*\boldsymbol{B} - \gamma^2\boldsymbol{r}^*\boldsymbol{B}\right]\boldsymbol{e}_k \\
&= 2\boldsymbol{e}_k^*\left[\boldsymbol{A}^*\boldsymbol{A}\boldsymbol{f}_\gamma - \boldsymbol{A}^*\boldsymbol{g}_\mathcal{E} + \gamma^2\boldsymbol{B}^*\boldsymbol{B}\boldsymbol{f}_\gamma - \gamma^2\boldsymbol{B}^*\boldsymbol{r}\right],
\end{aligned}$$

so daß hieraus folgt:

$$(\boldsymbol{A}^*\boldsymbol{A} + \gamma^2\boldsymbol{B}^*\boldsymbol{B})\boldsymbol{f}_\gamma = \boldsymbol{A}^*\boldsymbol{g}_\mathcal{E} + \gamma^2\boldsymbol{B}^*\boldsymbol{r}.$$

A.3 Zur Landweber–Iteration

In Abschn. 2.3.1 soll die Iterationsfolge

$$\boldsymbol{f}^m = (\boldsymbol{I} - \beta\boldsymbol{A}^*\boldsymbol{A})\boldsymbol{f}^{m-1} + \beta\boldsymbol{A}^*\boldsymbol{g}_\mathcal{E}$$

mit der Anfangsschätzung $\boldsymbol{f}^0 = 0$ dargestellt werden als Filterung der verallgemeinerten Inversen. Hierzu stellen wir $\boldsymbol{f}^m$ dar in der Basis der Eigenvektoren der Matrix $\boldsymbol{A}^*\boldsymbol{A}$:

$$\begin{aligned}
\boldsymbol{f}^m &= \sum_n < \boldsymbol{f}^m|\boldsymbol{v}_n > \boldsymbol{v}_n \\
&= \sum_n < \sum_{k=0}^{m-1}(\boldsymbol{I} - \beta\boldsymbol{A}^*\boldsymbol{A})^k\beta\boldsymbol{A}^*\boldsymbol{g}_\mathcal{E}|\boldsymbol{v}_n > \boldsymbol{v}_n \\
&= \sum_n \Big\{ < \beta\boldsymbol{A}^*\boldsymbol{g}_\mathcal{E}|\boldsymbol{v}_n > + \\
&\qquad < (\boldsymbol{I} - \beta\boldsymbol{A}^*\boldsymbol{A})\beta\boldsymbol{A}^*\boldsymbol{g}_\mathcal{E}|\boldsymbol{v}_n > + \cdots \\
&\qquad < (\boldsymbol{I} - \beta\boldsymbol{A}^*\boldsymbol{A})^{m-1}\beta\boldsymbol{A}^*\boldsymbol{g}_\mathcal{E}|\boldsymbol{v}_n > \Big\} \\
&= \sum_n \Big\{\beta < \boldsymbol{g}_\mathcal{E}|\boldsymbol{A}\boldsymbol{v}_n > + \\
&\qquad \beta < (1 - \beta\sigma_n^2)\boldsymbol{g}_\mathcal{E}|\boldsymbol{A}\boldsymbol{v}_n > + \cdots \\
&\qquad \beta < (1 - \beta\sigma_n^2)^{m-1}\boldsymbol{g}_\mathcal{E}|\boldsymbol{A}\boldsymbol{v}_n > \Big\}\boldsymbol{v}_n \\
&= \sum_n \beta\sum_{k=0}^{m-1}(1 - \beta\sigma_n^2)^k < \boldsymbol{g}_\mathcal{E}|\boldsymbol{A}\boldsymbol{v}_n > \boldsymbol{v}_n \\
&= \sum_n \beta(1 - \beta\sigma_n^2)^m\sum_{k=1}^{m}(1 - \beta\sigma_n^2)^{-k} < \boldsymbol{g}_\mathcal{E}|\boldsymbol{A}\boldsymbol{v}_n > \boldsymbol{v}_n
\end{aligned}$$

Die zweite Summe in obiger Formel ist eine geometrische Reihe in $(1-\beta\sigma_n^2)$, so daß wir $\boldsymbol{f}^m$ auch schreiben können als

$$\begin{aligned}\boldsymbol{f}^m &= \sum_n \beta(1-\beta\sigma_n^2)^m \frac{1-(1-\beta\sigma_n^2)^{-m}}{(1-\beta\sigma_n^2)-1} < \boldsymbol{g}_\mathcal{E}|\boldsymbol{A}\boldsymbol{v}_n > \boldsymbol{v}_n \\ &= \sum_n \sigma_n^{-2} \underbrace{[1-(1-\beta\sigma_n^2)^m]}_{F_m(\sigma_n,\beta)} < \boldsymbol{g}_\mathcal{E}|\boldsymbol{A}\boldsymbol{v}_n > \boldsymbol{v}_n\end{aligned}$$

mit einem Filter $F_m(\sigma_n,\beta) = 1-(1-\beta\sigma_n^2)^m$.

A.4 Zum Verfahren der maximalen Wahrscheinlichkeit

Sei

$$\begin{aligned}J(\boldsymbol{f}) &= (\boldsymbol{g}_\mathcal{E}-\boldsymbol{A}\boldsymbol{f})^*\mathcal{S}_\mathcal{E}^{-1}(\boldsymbol{g}_\mathcal{E}-\boldsymbol{A}\boldsymbol{f}) + (\boldsymbol{f}-\overline{\boldsymbol{f}})^*\mathcal{S}_f^{-1}(\boldsymbol{f}-\overline{\boldsymbol{f}}) \\ &= \boldsymbol{g}_\mathcal{E}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{g}_\mathcal{E} - \boldsymbol{g}_\mathcal{E}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{A}\boldsymbol{f} - \boldsymbol{f}^*\boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{g}_\mathcal{E} + \boldsymbol{f}^*\boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{A}\boldsymbol{f} + \\ &\quad \boldsymbol{f}^*\mathcal{S}_f^{-1}\boldsymbol{f} - \boldsymbol{f}^*\mathcal{S}_f^{-1}\overline{\boldsymbol{f}} - \overline{\boldsymbol{f}}^*\mathcal{S}_f^{-1}\boldsymbol{f} + \overline{\boldsymbol{f}}^*\mathcal{S}_f^{-1}\overline{\boldsymbol{f}}\end{aligned}$$

das in Abschn. 2.4.2 Gl. (36) bzgl. $\boldsymbol{f}$ zu minimierende Funktional. Komponentenweises Differenzieren

$$\begin{aligned}\partial J(\boldsymbol{f})/\partial f_k &= -\boldsymbol{g}_\mathcal{E}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{A}\boldsymbol{e}_k - \boldsymbol{e}_k^*\boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{g}_\mathcal{E} + \boldsymbol{e}_k^*\boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{A}\boldsymbol{f} + \boldsymbol{f}^*\boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{A}\boldsymbol{e}_k \\ &\quad + \boldsymbol{e}_k^*\mathcal{S}_f^{-1}\boldsymbol{f} + \boldsymbol{f}^*\mathcal{S}_f^{-1}\boldsymbol{e}_k - \boldsymbol{e}_k^*\mathcal{S}_f^{-1}\overline{\boldsymbol{f}} - \overline{\boldsymbol{f}}^*\mathcal{S}_f^{-1}\boldsymbol{e}_k \\ &= \boldsymbol{e}_k^*\left[\boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{A}\boldsymbol{f} + \mathcal{S}_f^{-1}\boldsymbol{f} - \boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{g}_\mathcal{E} - \mathcal{S}_f^{-1}\overline{\boldsymbol{f}}\right] + \\ &\quad + \left[\boldsymbol{f}^*\boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{A} + \boldsymbol{f}^*\mathcal{S}_f^{-1} - \boldsymbol{g}_\mathcal{E}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{A} - \overline{\boldsymbol{f}}^*\mathcal{S}_f^{-1}\right]\boldsymbol{e}_k \\ &= 2\boldsymbol{e}_k^*\left[\boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{A}\boldsymbol{f} + \mathcal{S}_f^{-1}\boldsymbol{f} - \boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{g}_\mathcal{E} - \mathcal{S}_f^{-1}\overline{\boldsymbol{f}}\right]\end{aligned}$$

und anschließendes zu 0 setzen liefert

$$(\boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{A} + \mathcal{S}_f^{-1})\boldsymbol{f} = \boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{g}_\mathcal{E} + \mathcal{S}_f^{-1}\overline{\boldsymbol{f}}.$$

A.5 Zur praktischen Inversionsgleichung

$$\begin{aligned}\boldsymbol{f}_{opt} &= (\boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{A} + \mathcal{S}_f^{-1})^{-1}(\boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{g}_\mathcal{E} + \mathcal{S}_f^{-1}\overline{\boldsymbol{f}}) \\ &= \overline{\boldsymbol{f}} + (\boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{A} + \mathcal{S}_f^{-1})^{-1}(\boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{g}_\mathcal{E} + \mathcal{S}_f^{-1}\overline{\boldsymbol{f}}) - \overline{\boldsymbol{f}} \\ &= \overline{\boldsymbol{f}} + (\boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}\boldsymbol{A} + \mathcal{S}_f^{-1})^{-1}\boldsymbol{A}^*\mathcal{S}_\mathcal{E}^{-1}(\boldsymbol{g}_\mathcal{E} - \boldsymbol{A}\overline{\boldsymbol{f}})\end{aligned}$$

$$
\begin{aligned}
&= \overline{\boldsymbol{f}} + \boldsymbol{\mathcal{S}}_f(\boldsymbol{A}^*\boldsymbol{\mathcal{S}}_\varepsilon^{-1}\boldsymbol{A}\boldsymbol{\mathcal{S}}_f + \boldsymbol{I})^{-1}\boldsymbol{A}^*\boldsymbol{\mathcal{S}}_\varepsilon^{-1}(\boldsymbol{g}_\varepsilon - \boldsymbol{A}\overline{\boldsymbol{f}}) \\
&= \overline{\boldsymbol{f}} + \boldsymbol{\mathcal{S}}_f(\boldsymbol{\mathcal{S}}_\varepsilon\boldsymbol{A}^{*-1}(\boldsymbol{A}^*\boldsymbol{\mathcal{S}}_\varepsilon^{-1}\boldsymbol{A}\boldsymbol{\mathcal{S}}_f + \boldsymbol{I}))^{-1}(\boldsymbol{g}_\varepsilon - \boldsymbol{A}\overline{\boldsymbol{f}}) \\
&= \overline{\boldsymbol{f}} + \boldsymbol{\mathcal{S}}_f(\boldsymbol{A}\boldsymbol{\mathcal{S}}_f + \boldsymbol{\mathcal{S}}_\varepsilon\boldsymbol{A}^{*-1})^{-1}(\boldsymbol{g}_\varepsilon - \boldsymbol{A}\overline{\boldsymbol{f}}) \\
&= \overline{\boldsymbol{f}} + \boldsymbol{\mathcal{S}}_f\boldsymbol{A}^*(\boldsymbol{A}\boldsymbol{\mathcal{S}}_f\boldsymbol{A}^* + \boldsymbol{\mathcal{S}}_\varepsilon)^{-1}(\boldsymbol{g}_\varepsilon - \boldsymbol{A}\overline{\boldsymbol{f}})
\end{aligned}
$$

A.6 Zum Backus Gilbert Verfahren

In Abschn. 2.5 soll aus der Lagrangefunktion (43)

$$\Phi = \text{size}(\boldsymbol{\mathcal{S}}_\gamma) + \gamma\ \text{spread}(\boldsymbol{R}) + \beta\ \text{diff}(\boldsymbol{R})$$

die verallgemeinerte Inverse $\boldsymbol{A}_\gamma^{-A}$ ermittelt werden. Hieraus folgt mit Hilfe der Definitionen (40), (41) und (42) unmittelbar

$$\Phi = \sum_{i=1}^{m}\left\{\mathcal{S}_{\gamma ii} + \gamma\sum_{j=1}^{m} w(i,j)R_{ij}R_{ij} + \beta\sum_{j=1}^{m} R_{ij} - \beta\right\}$$

bzw. die Definitionen $\boldsymbol{\mathcal{S}}_\gamma = \boldsymbol{A}_\gamma^{-A}\boldsymbol{\mathcal{S}}_\varepsilon\boldsymbol{A}_\gamma^{-A*}$ sowie $\boldsymbol{R} = \boldsymbol{A}_\gamma^{-A}\boldsymbol{A}$ eingesetzt

$$
\begin{aligned}
\Phi &= \sum_{i=1}^{m}\left\{\sum_{k=1}^{n}\sum_{l=1}^{n} A_{\gamma ik}^{-A}\mathcal{S}_{\varepsilon kl}A_{\gamma il}^{-A} + \gamma\sum_{j=1}^{m} w(i,j)\Big[\sum_{k=1}^{n} A_{\gamma ik}^{-A}A_{kj}\Big]\Big[\sum_{l=1}^{n} A_{\gamma il}^{-A}A_{lj}\Big]\right. \\
&\qquad \left. + \beta\sum_{j=1}^{m}\sum_{k=1}^{n} A_{\gamma ik}^{-A}A_{kj} - \beta\right\} \\
&= \sum_{i=1}^{m}\left\{\sum_{k=1}^{n}\sum_{l=1}^{n} A_{\gamma ik}^{-A}A_{\gamma il}^{-A}\underbrace{\Big[\mathcal{S}_{\varepsilon kl} + \gamma\sum_{j=1}^{m} w(i,j)A_{kj}A_{lj}\Big]}_{\mathcal{H}_{\gamma kl}^{i}} + \beta\sum_{k=1}^{n} A_{\gamma ik}^{-A}\underbrace{\sum_{j=1}^{m} A_{kj}}_{u_k} - \beta\right\}.
\end{aligned}
$$

Neu eingeführt wurden hierbei der Vektor $\boldsymbol{u}$, dessen n Komponenten u_k sich aus der Summation über die m Elemente der k-ten Zeile der Matrix $\boldsymbol{A}$ ergeben,

$$u_k = \sum_{j=1}^{m} A_{kj}, \tag{46}$$

sowie eine $n \times n$- Matrix $\boldsymbol{\mathcal{H}}_\gamma^i$, definiert durch

$$\mathcal{H}_{\gamma kl}^{i} = \mathcal{S}_{\varepsilon kl} + \gamma\sum_{j=1}^{m} w(i,j)A_{kj}A_{lj}. \tag{47}$$

Der hochgestellte Index i soll andeuten, daß die Matrix $\boldsymbol{\mathcal{H}}_\gamma^i$ zusätzlich noch über den Wichtungsfaktor $w(i,j)$ vom Laufindex i abhängt. Ableiten der Lagrangefunktion Φ

nach den Elementen der verallgemeinerten Inversen $A^{-A}_{\gamma rs}$ und anschließendes zu Null setzen ergibt

$$\begin{aligned}\frac{\partial \Phi}{\partial A^{-A}_{\gamma rs}} &= \sum_{i=1}^{m}\left\{\sum_{k=1}^{n}\sum_{l=1}^{n}\left[\delta_{ri}\delta_{sk}A^{-A}_{\gamma il}+\delta_{ri}\delta_{sl}A^{-A}_{\gamma ik}\right]\mathcal{H}^{i}_{\gamma kl}+\beta\sum_{k=1}^{n}\delta_{ri}\delta_{sk}u_k\right\}\\ &= \sum_{l=1}^{n}A^{-A}_{\gamma rl}\mathcal{H}^{r}_{\gamma sl}+\sum_{k=1}^{n}A^{-A}_{\gamma rk}\mathcal{H}^{r}_{\gamma ks}+\beta u_s\\ &= 2\sum_{k=1}^{n}A^{-A}_{\gamma rk}\mathcal{H}^{r}_{\gamma ks}+\beta u_s=0.\end{aligned}$$

Die Lösung dieses Gleichungssystems ist

$$A^{-A}_{\gamma rk}=-\frac{1}{2}\beta\sum_{s=1}^{n}\mathcal{H}^{r}_{\gamma ks}u_s.$$

Setzen wir diesen Ausdruck in die Bedingung $\mathrm{diff}(\boldsymbol{R})=0$ ein, um nachzuprüfen inwieweit sie erfüllt ist,

$$\begin{aligned}\mathrm{diff}(\boldsymbol{R}) &= \sum_{i=1}^{m}\Big[\sum_{j=1}^{m}R_{ij}-1\Big]\\ &= \sum_{i=1}^{m}\Big[\sum_{k=1}^{n}A^{-A}_{\gamma ik}u_k-1\Big]\\ &= \sum_{i=1}^{m}\Big[-\frac{1}{2}\beta\sum_{k=1}^{n}\sum_{s=1}^{n}\mathcal{H}^{i}_{\gamma ks}u_su_k-1\Big]=0,\end{aligned}$$

so stellen wir fest, daß die Zusatzbedingung nur erfüllt werden kann, wenn der Klammerausdruck in obiger Gleichung für alle i, und insbesondere für $i=r$, identisch Null ist. Hieraus folgt unmittelbar

$$\beta=\frac{-2}{\sum_{k=1}^{n}\sum_{s=1}^{n}\mathcal{H}^{r}_{\gamma ks}u_su_k}$$

und damit als Endresultat

$$A^{-A}_{\gamma rk}=\frac{\sum_{s=1}^{n}\mathcal{H}^{r\,-1}_{\gamma ks}u_s}{\sum_{k=1}^{n}\sum_{s=1}^{n}\mathcal{H}^{r}_{\gamma ks}u_su_k}. \tag{48}$$

Wichtig bei der Berechnung der Elemente von $\boldsymbol{A}^{-A}_{\gamma}$ ist, zu beachten, daß die Matrix $\mathcal{H}^{r}_{\gamma}$ vom Zeilenindex r der verallgemeinerten Inversen abhängt und für jede Zeile der verallgemeinerten Inversen $\boldsymbol{A}^{-A}_{\gamma}$ neu berechnet werden muß.

In der Literatur findet man gewöhnlich eine etwas kompaktere Darstellung unseres Endergebnisses (48). Um zu dieser zu gelangen, müssen wir die r-te Zeile der verallgemeinerten Inversen $\boldsymbol{A}_\gamma^{-A}$ als einen Vektor interpretieren und erhalten für jeden dieser „Zeilenvektoren" $\boldsymbol{a}_r$ die dann etwas übersichtlichere Bestimmungsgleichung

$$\boldsymbol{a}_r = \frac{\boldsymbol{\mathcal{H}}_\gamma^{r\,-1}\boldsymbol{u}}{\boldsymbol{u}^*\boldsymbol{\mathcal{H}}_\gamma^r\boldsymbol{u}}$$

mit $\boldsymbol{u}$ und $\boldsymbol{\mathcal{H}}_\gamma^r$ aus (46) bzw. (47).

Anhang B

Symbole und Abkürzungen

$\boldsymbol{A}$	Eine $n \times m$ Matrix, im wesentlichen der Kern der Integralgleichung
$\boldsymbol{A}^*$	Adjungierte (bzw. Transponierte) der Matrix $\boldsymbol{A}$
$\boldsymbol{A}^{-1}$	Inverse der Matrix $\boldsymbol{A}$, sofern diese berechnet werden kann
$\boldsymbol{A}^{-A}$	Verallgemeinerte Inverse der Matrix $\boldsymbol{A}$
$\boldsymbol{A}_\gamma^{-A}$	Regularisierte verallgemeinerte Inverse der Matrix $\boldsymbol{A}$
$\boldsymbol{B}$	Eine $n \times n$ Matrix zur Formulierung von Nebenbedingungen bei der Tikhonov–Phillips–Regularisierung
$\boldsymbol{C}$	Eine positiv definite $m \times m$ Diagonalmatrix zur unterschiedlichen Gewichtung von Messungen beim Tikhonov–Phillips–Verfahren
$\boldsymbol{D}$	Eine $n \times n$ Matrix bei der Erweiterung der Landweber–Iteration nach Strand
$\boldsymbol{d}$	Fortschreitungsrichtung beim konjugierten Gradienten–Verfahren
$\boldsymbol{\mathcal{E}}$	Meßfehlervektor (Vektor der Länge m)
$\overline{\boldsymbol{\mathcal{E}}}$	Erwartungswert des Meßfehlervektors (Vektor der Länge m)
$\boldsymbol{e}_k$	k-ter Einheitsvektor
$\boldsymbol{f}$	Exakte Lösung bzw. das gesuchte Höhenprofil (Vektor der Länge n)
$\overline{\boldsymbol{f}}$	Erwartungswert des Lösungsvektors (Vektor der Länge n)
$\boldsymbol{f}_\gamma$	Regularisierte Lösung (Vektor der Länge n)
$\boldsymbol{f}_{opt}$	Regularisierte oder optimale Lösung bei den stochastischen Verfahren (Vektor der Länge n)
$F_\gamma()$	Regularisierendes Filter
$\boldsymbol{g}$	Die exakten Daten, ohne Meßfehler (Vektor der Länge m)

$\boldsymbol{g}^{mes}$	Die gemessenen Daten beim Chahineschen Verfahren (entspricht $\boldsymbol{g}_\varepsilon$)
$\boldsymbol{g}_\varepsilon$	Die gemessenen Daten, mit Meßfehler (Vektor der Länge m)
$\boldsymbol{g}_\gamma$	Aus der regularisierten Lösung $\boldsymbol{f}_\gamma$ über Vorwärtsrechnung ermittelte Daten (Vektor der Länge m)
$\mathcal{H}_\gamma^r$	Vom Zeilenindex r der verallgemeinerten Inversen abhängige Hilfsmatrizen beim Backus–Gilbert–Verfahren ($n \times n$)
$\boldsymbol{I}$	$n \times n$ Einheitsmatrix
$J()$	Funktional beim konjugierten Gradienten–Verfahren
m	Iterationstiefe bei den iterativen Verfahren Dimension des Datenvektors $\boldsymbol{g}$ (Anzahl der spektralen Stützstellen)
n	Die Dimension des Lösungsvektors $\boldsymbol{f}$ (Anzahl der Höhenschichten)
$\mathcal{O}()$	Landau–Symbol
$P_{m-1}()$	Polynom vom Grad $m-1$ im Argument
$p_f()$	Wahrscheinlichkeitsdichte für das Auftreten einer Lösung $\boldsymbol{f}$
$p_\varepsilon()$	Wahrscheinlichkeitsdichte für das Auftreten eines Meßfehlers $\boldsymbol{\varepsilon}$
$p_{f,\varepsilon}()$	Wahrscheinlichkeitsdichte für das Auftreten einer Lösung $\boldsymbol{f}$ mit Meßfehlervektor $\boldsymbol{\varepsilon}$
$\boldsymbol{R}$	Auflösungsmatrix ($n \times n$)
$\boldsymbol{r}$	Hilfsvektor zur Formulierung von Nebenbedingungen bei der Tikhonov–Phillips–Regularisirung (Vektor der Länge m) Richtung des steilsten Anstieges beim konjugierte Gradienten–Verfahren
r	Parameter bei der a posteriori Parameterwahl
$\boldsymbol{S}$	Eine $n \times n$ Matrix bei der Landweber–Iteration
$\boldsymbol{\mathcal{S}}_\varepsilon$	Fehlerkovarianzmatrix ($m \times m$)
$\boldsymbol{\mathcal{S}}_\gamma$	Kovarianzmatrix der regularisierten Lösung $\boldsymbol{f}_\gamma$ ($n \times n$)
$\boldsymbol{\mathcal{S}}_{\gamma\,stoch}$	Kovarianzmatrix der sogenannten optimalen Lösung $\boldsymbol{f}_{opt}$ bei den stochastischen Verfahren ($n \times n$)

size()	Kriterium zur Beurteilung der Größe des Datenfehlers beim Backus–Gilbert Verfahren
spread()	Kriterium zur Beurteilung der Güte der Auflösung beim Backus–Gilbert Verfahren
$\boldsymbol{T}$	Eine $n \times n$ Matrix bei der Landweber–Iteration
$\boldsymbol{u}$	Hilfsvektor beim Backus–Gilbert–Verfahren (Vektor der Länge n)
$\boldsymbol{v}_n$	Eigenvektor der Matrix $\boldsymbol{A}^*\boldsymbol{A}$ zum Eigenwert σ_n^2 (Vektor der Länge n)
$w()$	Wichtungsfaktor beim Backus–Gilbert–Verfahren
α	Hilfsgröße beim konjugierten Gradienten–Verfahren
β	Steuerparameter bei der Landweber–Iteration Hilfsgröße beim konjugierten Gradienten–Verfahren Hilfsgröße beim Backus–Gilbert–Verfahren
β_n^2	Eigenwert der Matrix $\boldsymbol{B}^*\boldsymbol{B}$
δ_n^2	Eigenwert der Matrix $\boldsymbol{D}$
γ	Regularisierungsparameter
ω	Steuerparameter bei der Erweiterung der Landweber–Iteration nach Strand
σ_ε^2	Varianz eines Elements des Meßfehlervektors
σ_n^2	Eigenwert der Matrix $\boldsymbol{A}^*\boldsymbol{A}$
$\|\ \ \|$	Betrag eines Vektors bezüglich der L_2-Norm
$< \mid >$	Skalarprodukt in Diracscher Notation
SÜG	Strahlungsübertragungsgleichung

Literaturverzeichnis

[1] Backus, G., Gilbert, F., Uniqueness in the Inversion of Inaccurate Gross Earth Data, Philosophical Transactions of the Royal Society of London, A, Vol. 266, pp. 123-192, Mar. 1970.

[2] Baumeister, J., Stable Solutions of Inverse Problems, Advanced Lectures in Mathematics, Vieweg Verlag, Braunschweig Wiesbaden, 1987.

[3] Bevilacqua, R. M., Olivero, J. J., Vertical Resolution of Middle Atmospheric Measurements by Ground-Based Microwave Radiometry, Journal of Geophysical Research, Vol. 93, No. D8, pp. 9463-9475, Aug. 1988.

[4] Brillet, J., A Theoretical Study of Ozone Measurements Made with Ground-Based Microwave Sensors, Journal of Geophysical Research, Vol. 94, No. D10, pp. 12,833-12,850, Sep. 1989.

[5] Bronstein, I. N., Semendjajev, K. A., Taschenbuch der Mathematik, Verlag Harri Deutsch, 1985.

[6] Chahine, M. T., A General Method for Inverse Solution of the Full Radiative Transfer Equation, Journal of the Atmospheric Sciences, Vol. 29, pp. 741-747, May 1972.

[7] Chahine, M. T., Deepak, A. (ed.), Generalization of the Relaxation Method for the Inverse Solution of Nonlinear and Linear Transfer Equations, Inversion Methods in Atmospheric Remote Sounding (A workshop held at Langley Research Center, Hampton Virginia), Academic Press, NewYork, pp. 117-138, 1977.

[8] Chandrasekhar, S., Radiative Transfer, Dover Publications Inc., NewYork, 1960.

[9] Von Clarmann, T., Die Bestimmung von Mischungsverhältnissen stratosphärischer Spurengase aus Emissionsspektren im infraroten Spektralbereich, Dissertation, genehmigt von der Fakultät für Physik der Universität Karlsruhe, März 1990.

[10] Conrath, B. J., Deepak, A. (ed.), Backus-Gilbert Theory and its Application to Retrieval of ozone and Temperature Profiles, Inversion Methods in Atmospheric Remote Sounding, (A workshop held at Langley Research Center, Hampton, Virginia), Academic Press, New York, pp. 155-193, 1977.

[11] Davies, R. B., Documentation for NEWMAT07, an Experimental Matrix Package in C++, public domain software, 1993.

[12] Fischer, H., Fergg, F., Oelhaf H., Rabus D., Völker W., Burkert P., Simultaneous Detection of Trace Constituents in the Middle Atmosphere with a Small He-Cooled, High Resolution Michelson Interferometer (MIPAS), Beiträge zur Physik der Atmosphäre, Vol. 56, No. 2, pp. 260-275, May 1983.

[13] Fischer, H., Beiträge zur passiven Fernerkundung stratosphärischer Spurengase im infraroten Spektralbereich, Habilitationsschrift der Fakultät für Physik der Ludwigs-Maximilians-Universität München, Okt. 1985.

[14] Foster, M., An Application of the Wiener-Kolmogorov Smoothing Theory to Matrix Inversion, Journal of the Society for Industrial and Applied Mathematics, Vol. 9, No. 3, Sep. 1961.

[15] Franklin, J. N., Well-Posed Stochastic Extensions of Ill-Posed Linear Problems, Journal of Mathematical Analysis and Applications, Vol. 31, pp. 682-716, 1970.

[16] Goody, R. M., Yung, Y. L., Atmospheric Radiation, Theoretical Basis, Second Edition, Oxford University Press, 1989.

[17] Hadamard, J., Lectures on the Cauchy Problem in Linear Partial Differential Equations, Yale University Press, NewHaven, 1923.

[18] Hochschild, G., Krupa, R., Kuntz, M., Millimeterwellen-Radiometrie von Spurengasen in der Stratosphäre, KfK-Nachrichten, Jahrgang 25, Nr. 4, pp. 182-192, Dez. 1993.

[19] Houghton, J. T., Taylor, F. W., Rodgers, C. D., Remote Sounding of Atmospheres, Cambridge Planetary Sciences Series 5, Cambridge University Press, 1986.

[20] Ishimaru, A., Wave Propagation and Scattering in Random Media, Section 22, Vol. II, Academic Press NewYork SanFrancisco London, 1978.

[21] Landweber, L., An Iteration Formula for Fredholm Integral Equations of the First Kind, Americal Journal of Mathematics, Vol. 92, No. 191, pp. 615-624, 1951.

[22] Loredo, T. J., Epstein, R. I., Analyzing Gamma-Ray Burst Spectral Data, The Astrophysical Journal, Vol. 336, pp. 896-919, Jan. 1989.

[23] Louis, A. K., Inverse und schlecht gestellte Probleme, Teubner Studienbücher, 1989.

[24] Menke, W., Geophysical Data Analysis: Discrete Inverse Theory, Academic Press Inc., SanDiego, 1989.

[25] Miller, K., Least Squares Methods for Ill-Posed Problems with a Prescribed Bound, SIAM, Journal on Numerical Analysis, Vol. 16, pp. 52-74, 1970.

[26] Morrison, D. F., Multivariate Statistical Methods, McGraw-Hill, Second Edition, NewYork, 1967.

[27] Natterer, F., The Mathematics of Computerized Tomography, B. G. Teubner, Stuttgart, John Wiley & Sons, Chichester NewYork Brisbane Toronto Singapore, 1986.

[28] Press, W. H., Vetterling, W. T., Teukolsky, S. A., Flannery, B. P., Numerical Recipes in C, The Art of Scientific Computing, Cambridge University Press, 1992.

[29] Randegger, A. K., On the Determination of the Atmospheric Ozone Profile for Ground Based Microwave Measurements, Pageoph, Vol. 118, pp. 1052-1065, Birkhäuser Verlag, Basel, Feb. 1980.

[30] Rodgers, C. D., Remote Sounding of the Atmospheric Temperature Profile in the Presence of Cloud, Quarterly Journal of the Royal Meteorological Society, Vol. 96, pp. 654-666, 1970.

[31] Rodgers, C. D., Retrieval of Atmospheric Temperature and Composition From Remote Measurements of Thermal Radiation, Reviews of Geophysics and Space Physics, Vol. 14, No. 4, pp. 609-624, Nov. 1976.

[32] Rodgers, C. D., Deepak, A. (ed.), Statistical Principles of Inversion Theory, Inversion Methods in Atmospheric Remote Sounding (A workshop held at Langley Research Center, Hampton, Virginia), Academic Press, NewYork, pp. 117-138, 1977.

[33] Rodgers, C. D., Fleming, H. E. (ed.), Chahine, M. T. (ed.), A Strategy for Optimal Profile Retrieval from Limb Sounders, Advances in Remote Sensing Retrieval Methods, A. Deepak Publishing, 1985.

[34] Rodgers, C. D., Characterization and Error Analysis of Profiles Retrieved from Remote Sounding Measurements, Journal of Geophysical Research, Vol. 95, No. D5, pp. 5587-5595, Apr. 1990.

[35] Schock, E., Nonlinear Ill-Posed Equations: Singular Value Decomposition and the Picard Criterion, Journal of Mathematical Analysis and Applications, No. 116., pp. 200-208, 1986.

[36] Schock, E., Semi-Iterative Methods for the Approximate Solution of Ill-Posed Problems, Numerische Mathematik, Vol. 50, pp. 263-271, 1987.

[37] Smirnow, W. I., Lehrgang der Höheren Mathematik, VEB Deutscher Verlag der Wissenschaften, Berlin, 1988.

[38] Strand, O. N., Westwater, E. R., Statistical Estimation of the Numerical Solution of a Fredholm Intergral Equation of the First Kind, Journal of the Association for Computing Machinery, Vol. 15, No. 1, pp. 100-114, Jan. 1968.

[39] Strand, O. N., Theory and Methods Related to the Singular-Function Expansion and Landweber's Iteration for Integral Equations of the First Kind, SIAM, Journal on Numerical Analysis, Society for Industrial and Applied Mathematics, Vol. 11, pp. 798-825, Sep. 1974.

[40] Tarantola, A., Valette, B., Generalized Nonlinear Inverse Problems Using the Least Squares Criterion, Reviews of Geophysics and Space Physics, Vol. 20, No. 2, pp. 219-232, May 1982.

[41] Twomey, S., Herman, B., Rabinoff, R., An Extension to the Chahine Method of Inverting the Radiative Transfer Equation, Journal of the Atmospheric Sciences, Vol. 34, pp. 1085-1090, Jul. 1977.

[42] Twomey, S., Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements, Developments in Geomathematics 3, Elsevier Scientific Publishing Company, Amsterdam, Oxford, NewYork, 1977.

[43] Ulaby, F. T., Moore, R. K., Fung, A. K., Passive Microwave Sensing of the Atmosphere, Microwave Remote Sensing, Active and Passive, Artech House, Vol. 3, chap. 17, pp. 1283-1308, 1990.

[44] Westwater, E. R., Cohen, A., Application of Backus-Gilbert Inversion Technique to Determination of Aerosol Size Distributions from Optical Scattering Measurements, Applied Optics, Vol. 12, No. 6, pp. 1340-1348, June 1973.

[45] Westwater, E. R., Janssen, M. A. (ed.), Atmospheric Remote Sensing by Microwave Radiometry, Wiley Series in Remote Sensing, John Wiley & Sons Inc., NewYork Cichester Brisbane Toronto Signapore, 1993.

Index

Auflösung, 27, 45, 58
    Definition der, 37
Auflösungsmatrix, 7, 27, 37
a posteriori Parameterwahl, 5, 10, 22
a priori Information, 12, 22, 53
a priori Parameterwahl, 5, 22

Backus–Gilbert–Verfahren, 27, 58
Bayes–Schätzung, 22
bester linearer Schätzer, 22

Chahinesches Verfahren, 20, 50
    modifiziertes, 21, 50

Datenfehler, 5, 27, 37
Defekt in den Daten, 5
direktes Problem, 1

Einfluß des Startprofils, 47, 50
Erwartungswert, 23, 53

Fehler, 7, 26, 37
Filter, 4, 10, 12, 15, 17, 19
Fredholmsche Integralgleichung, 2

Gewichtsfunktionen, 20, 32
Gibbssches Phänomen, 40
gut gestelltes Problem, 2

Hadamard, 2

Identifizierungsproblem, 1
Intensitätsspektrum, 30
inverses Problem, 1
Inversion
    Theorie der, 1
    Verfahren, *siehe* Regularisierungsverfahren

Kern der Integralgleichung, 2
konjug. Gradienten–Verfahren, 17, 45
Korrelation, 23
Kovarianzmatrix
    a priori, 23, 53
      Initialisierung von Hand, 56
    Fehler–, 23, 53
    Lösungs–, 8, 26

Lagrangescher Multiplikator, 12, 28
Landweber–Iteration, 15, 45
    modifizierte, 17
linear constraint solution, 12
lineares Problem, 1

maximum likelihood method, 24
Meßfehlervektor, 23
Methode der kleinsten Quadrate, 2
minimum variance method, 22
Moore–Penrose–Lösung, 3

nichtlineares Problem, 1

optimum estimation method, 22
oszillierende Eigenvektoren, 6, 35

praktische Inversionsgleichung, 26

Quadratur, 2

Rayleigh Kriterium, 41
Regularisierung, 3
Regularisierungsparameter, 4
    Wahl des, 5
Regularisierungsverfahren, 10
    Beurteilung der, 62
    Effizienz, 62
    iteratives, 15, 17, 20, 21
    lineares, 10, 12, 15, 17, 22, 27
    nichtlineares, 17, 20, 21

Übersicht, 63
regularization method, 12
Rekonstruktionsfehler, 5
Rodgers Verfahren, 22

schlecht gestelltes Problem, 2
exponentiell, 6, 33
Klassifizierung, 6
von der Ordnung α, 6
Semikonvergenzverhalten, 16
Signal zu Rausch Verhältnis, 20
Singulärwertzerlegung, 3
abgeschnittene, 10, 33
smoothing method, 12
spektrale Stützstellen, 30
spread–Funktion, 27, 40
stochastische Verfahren, 22, 53
Strahlungsübertragungsgleichung, 1
Streuung, 23

Tiefpaß, 6
Tikhonov–Phillips–Regularisierung, 12, 41
iterative, 17
try and error Parameterwahl, 5

Überschwinger, 40

verallgemeinerte Inverse, 2
regularisierte, 4
verallgemeinerte Lösung, 3
Verfahren der maximalen Wahrscheinlichkeit, 24
virtuelle Messungen, 22
Vorwärts–Problem, *siehe* direktes Problem

Wahrscheinlichkeitsdichte, 24

Zufallsvektor, 23