

FORSCHUNGSZENTRUM KARLSRUHE

Technik und Umwelt

Wissenschaftliche Berichte

FZKA 5719

**Bestimmung der Höhenverteilung
stratosphärischer Spurengase aus
Emissionsspektren eines bodengebundenen
Millimeterwellen–Radiometers***

Martin Kuntz

Institut für Meteorologie und Klimaforschung

* Von der Fakultät für Physik der Universität (T. H.) Karlsruhe genehmigte
Dissertation

Forschungszentrum Karlsruhe GmbH, Karlsruhe
1996

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

Forschungszentrum Karlsruhe GmbH
Postfach 3640, 76021 Karlsruhe

ISSN 0947-862

Kurzfassung

Bestimmung der Höhenverteilung stratosphärischer Spurengase aus Emissionsspektren eines bodengebundenen Millimeterwellen–Radiometers

Zahlreiche natürliche Bestandteile der Erdatmosphäre sowie auch anthropogene Spurenstoffe zeigen charakteristische Spektrallinien im Millimeterwellenbereich, die durch Rotationsübergänge der Moleküle verursacht werden. Mit den am Institut für Meteorologie und Klimaforschung des Forschungszentrums Karlsruhe und der Universität Karlsruhe entwickelten Millimeterwellen–Radiometern kann die von der Atmosphäre emittierte schwache Strahlung spektral soweit aufgelöst werden, daß über eine eindeutige Stoffzuordnung hinaus eine Auswertung der Druckverbreiterung der Spektrallinien möglich ist. Vom Boden aus durchgeführte Messungen der Strahldichtespektren entsprechender Konstituenten erlauben damit die Bestimmung von Temperatur und Volumenmischungsverhältnis als Funktion der Höhe.

Der Schwerpunkt der Arbeit liegt auf der Entwicklung geeigneter Algorithmen für die Auswertung der aufgenommenen Spektren; dabei wird besonderes Augenmerk auf die Effizienz der Verfahren gelegt. Es wird zunächst eingegangen auf die Berechnung von Strahldichtespektren aus einem vorgegebenen Zustand der Atmosphäre, allgemein als direktes oder Vorwärts–Problem bezeichnet. Im Gegensatz dazu beinhaltet der umgekehrte Fall, das inverse Problem, den Schluß von den gemessenen Spektren zurück auf den Zustand der Atmosphäre. Nach einer kurzen Einführung in die Theorie der inversen Probleme und anschließender Diskussion geeigneter Inversionsverfahren wird eine Erweiterung dieser Verfahren vorgeschlagen, mit der es möglich ist, Höhenprofile mehrerer Spurengase simultan aus einem Spektrum abzuleiten. Außerdem wird ein neue Methode entwickelt, mit der stehende Wellen, die den Spektren häufig überlagert sind, bei der Auswertung gleich mitberücksichtigt werden können.

In einer Reihe numerisch modellierter Experimente konnte die Leistungsfähigkeit dieser Erweiterungen demonstriert werden. Schließlich wurden die Verfahren angewandt auf eine Anzahl von im Frühjahr '94 gemessenen Spektren von Ozon; dabei zeigen die invertierten Höhenprofile eine gute Übereinstimmung mit entsprechenden in situ–Messungen.

Abstract

Retrieval of stratospheric trace gas profiles from emission spectra of a ground based radiometer in the millimeter wave range

Several natural components of the atmosphere as well as anthropogenic trace gases show characteristic spectral lines in the millimeter wave region due to rotational transitions of the molecules. The millimeter wave radiometers at the institute of meteorology and climate research of the Research Center Karlsruhe and the University of Karlsruhe were developed for measuring this weak radiation emitted by the atmosphere with high spectral resolution. Besides a clear attachment of trace gases the evaluation of the pressure broadening allows the calculation of profiles of temperature and volume mixing ratio of appropriate trace gases from the measured spectra.

The present work is focused on the development of adequate retrieval algorithms, while other topics only touched briefly. First the direct problem, the numerical simulation of the radiative transfer in the atmosphere, is discussed; special attention is paid to the efficiency of the forward-algorithm. Second the inverse problem (finding a solution of the complicated equation of radiative transfer) is described. After a short introduction into the theory of inverse problems and a discussion of appropriate retrieval algorithms a new extension of two already existing inversion techniques is given, which allows for the simultaneous evaluation of several profiles of trace gases from one spectrum. Furthermore a new method is suggested for inverting spectra, which are frequently blotted out by device-caused standing waves.

Finally several numerical experiments are performed. Derived results from spectra of ozone, measured in spring '94 are in good agreement with corresponding in situ measurements taken at high altitudes.

Inhaltsverzeichnis

Kurzfassung	iii
1 Einleitung	1
2 Grundbegriffe der Strahlungsübertragung	4
2.1 Die Strahlungsübertragungsgleichung (SÜG)	4
2.1.1 Die differentielle Form der SÜG	4
2.1.2 Die integrale Form der SÜG	6
2.1.3 Implementierung der SÜG	7
2.2 Sensitivitäts- und Gewichtsfunktionen	9
2.2.1 Zur Unterscheidung zwischen beiden Funktionen	9
2.2.2 Berechnung der Sensitivitätsfunktionen	10
2.3 Linearisierung der SÜG	15
2.3.1 Gauß-Newton-Verfahren	15
2.3.2 Reduktion auf ein lineares Gleichungssystem	16
3 Inversion der linearisierten SÜG	19
3.1 Theorie der Inversion	19
3.1.1 Schlecht gestellte Probleme	20
3.1.2 Methode der kleinsten Quadrate	21
3.1.3 Singulärwertzerlegung und Regularisierung	22
3.1.4 Wahl des Regularisierungsparameters	23
3.1.5 Erreichbare vertikale Auflösung im Profil	26
3.1.6 Abschätzung des Fehlers	28
3.1.7 Zusammenfassung	29
3.2 Regularisierungsverfahren	30
3.2.1 Abgeschnittene Singulärwertzerlegung	32
3.2.2 Tikhonov-Phillips-Regularisierung	34
3.2.3 Linear statistisches Verfahren	38
3.3 Mögliche Fehlerquellen	43
4 Erweiterung der Inversionsverfahren	46
4.1 Ableitung mehrerer Spurengase simultan	46
4.2 Berücksichtigung von stehenden Wellen	49

5	Anwendung der Inversionsverfahren	55
5.1	Inversion synthetischer Spektren	56
5.1.1	Unverrauschte synthetische Spektren	56
5.1.2	Verrauschte synthetische Spektren	61
5.1.3	Unverrauschte synthetische Spektren mit stehenden Wellen	66
5.1.4	Verrauschte synthetische Spektren mit stehenden Wellen	74
5.2	Anwendung auf gemessene Spektren	78
5.2.1	Berücksichtigung des troposphärischen Wasserdampfs	78
5.2.2	Die 142,175 GHz-Linie von Ozon	79
5.2.3	Fehlerbetrachtung	89
5.3	Inversion von Chlormonoxid bei 278,631 GHz	98
6	Zusammenfassung	105
A	Vergleichsrechnungen	107
A.1	Rechnungen zum Vorwärtsproblem	107
A.2	Rechnungen zum inversen Problem	110
B	Ergänzungen	112
B.1	Berechnungen	112
B.1.1	Zur Berechnung der verallgemeinerten Inversen	112
B.1.2	Zur Minimierung des Tikhonov–Phillips–Funktionalis	112
B.1.3	Zur Tikhonov–Phillips Regularisierung	113
B.1.4	Zum linear statistischen Verfahren	113
B.1.5	Zur praktischen Inversionsgleichung	114
B.2	Spektralbereiche	115
C	Symbole und Abkürzungen	121
	Literaturverzeichnis	127

Abbildungsverzeichnis

1	Differentielle Form der SÜG	4
2	Verhältnis von Streuquerschnitt zu Absorptionsquerschnitt	5
3	Strahlungsübertragung innerhalb der Atmosphäre.	6
4	Linearisierung der SÜG	11
5	Rekonstruktionsfehler	25
6	Auflösungsmatrix	28
7	Filterwirkung der abgeschnittenen Singulärwertzerlegung	33
8	Filterwirkung der Tikhonov–Phillips–Regularisierung	36
9	Verfahren der maximalen Wahrscheinlichkeit	41
10	Unverraushtes synthetisches Spektrum	57
11	Inversion von O_3 und HNO_3 aus einem unverrauschten Spektrum	58
12	Auflösungsmatrix des unverrauschten Spektrums	60
13	Verrauschte synthetische Spektren	62
14	Inversion von O_3 und HNO_3 aus verrauschten Spektren	63
15	Fehler der Inversionsergebnisse	64
16	Auflösungsmatrix des verrauschten Spektrums	65
17	Mit mehreren stehenden Wellen überlagertes, unverraushtes Spektrum	67
18	Mit einer stehenden Welle kurzer Periode überlagertes Spektrum	70
19	Inversion von O_3 und HNO_3 aus einem mit einer stehenden Welle kurzer Periode überlagerten Spektrum	71
20	Mit einer stehenden Welle langer Periode überlagertes Spektrum	72
21	Inversion von O_3 und HNO_3 aus mit einer stehenden Wellen langer Periode überlagerten Spektrum	73
22	Mit mehreren stehenden Wellen überlagertes, verraushtes Spektrum	76
23	Inversion von O_3 und HNO_3 aus einem mit stehenden Wellen überlagerten, verrauschten Spektrum	77
24	gemessene Spektren von 28. März 1994	81
25	Temperaturprofile vom 28. März 1994	83
26	Anpassung an ein gemessenes Spektrum	85
27	Ozon–Höhenprofile vom März und April 1994	86
28	Fehler der invertierten Ozon–Profile	87
29	Höhenauflösung der Messungen vom 28. März 1994	88
30	Einfluß der Troposphäre sowie der Frequenzumsetzung auf das Spektrum	90
31	Relativer Fehler der Inversionsergebnisse	93

32	Verrauschtes Spektrum von ClO bei 278 GHz	100
33	Inversion von O ₃ und ClO aus einem verrauschten Spektrum	101
34	Auflösungsmatrix des verrauschten Spektrums	102
35	Internationaler Vergleich von Vorwärtsrechnungen	108
36	Vergleichsrechnungen mit FASCOD2	109
37	Internationaler Vergleich von Inversionsrechnungen	110
38	Millimeterwellenspektrum bei 142 GHz	115
39	Millimeterwellenspektrum bei 278 GHz	117
40	Millimeterwellenspektrum bei 204 GHz	119

Tabellenverzeichnis

1	Anpassung von stehenden Wellen aus einem unverrauschten Spektrum	68
2	Anpassung von stehenden Wellen aus einem verrauschten Spektrum .	75
3	Systematische Fehler	98

Kapitel 1

Einleitung

Es besteht kein Zweifel: Der massive Einsatz von Naturwissenschaft und Technik hat einen materiellen Wohlstand für alle Gruppen der Gesellschaft mit sich gebracht, wie er noch vor hundert Jahren unvorstellbar gewesen wäre. Die Kehrseite dieser Entwicklung ist ebenso wohlbekannt. Ein Teilaspekt ist die anthropogene Emission von Aerosolen und Spurengasen, in deren Folge die Konzentration verschiedener Spurenstoffe in der Atmosphäre seit Jahrzehnten proportional zu Wirtschaftswachstum und Bevölkerungsexplosion ansteigt. Bedingt durch die chemische Trägheit und die damit verbundene lange Lebensdauer vieler Spurengase sowie großräumige Transportvorgänge innerhalb der Atmosphäre ist die Zunahme der Konzentration dieser Gase nicht nur auf deren Entstehungsort und dessen Umgebung begrenzt. Vielmehr hat die Belastung der Atmosphäre inzwischen globale Bedeutung erlangt. Konkrete Beispiele für anthropogene Klimaänderungen sind der immer deutlicher werdende Abbau der Ozonschicht [81] sowie die Änderung der Strahlungsbilanz durch die Konzentrationszunahme verschiedener Spurenstoffe [90].

Seit der Entdeckung des antarktischen Ozonlochs [20] haben große nationale und internationale Anstrengungen eine rapide Entwicklung der Ozonforschung eingeleitet. Die Ursachen des antarktischen Ozonabbaus im Frühling sind jetzt weitgehend bekannt [23, 24, 79, 80], wenn auch noch viele Details geklärt werden müssen. In diesem Zusammenhang liefert die Fernerkundung wichtige Zustandsparameter der Atmosphäre, wobei die aus der Atmosphäre stammende elektromagnetische Strahlung als Informationsträger dient. Für globale und vergleichende Beobachtungen der mittleren Atmosphäre liefern Fernerkundungsinstrumente auf hochfliegenden Flugzeugen und Satelliten hervorragende Voraussetzungen. Vom Boden aus bestehen in Form der Millimeterwellen-Radiometrie günstige Möglichkeiten für eine zeitlich weitgehend kontinuierliche Messung von Höhenprofilen, aufgrund der Eigenschaft der Millimeterwellen, Wolken und in gewissen Grenzen auch Regen zu durchdringen (siehe Abb. 2).

Die Möglichkeit der Fernerkundung resultiert aus der spektral veränderlichen Absorption, Emission und Streuung der innerhalb der Atmosphäre übertragenen elektromagnetischen Strahlung. Die sich aus dieser Wechselwirkung zwischen elektromagnetischer Strahlung und Spurengasmolekülen ergebenden Bandenspektren stellen geradezu „Fingerabdrücke“ eines spektralen Gases dar. Mit den im IMK entwickelten

Millimeterwellen–Radiometern kann diese von der Atmosphäre emittierte schwache Strahlung spektral soweit aufgelöst werden, daß über eine eindeutige Stoffzuordnung hinaus eine Auswertung der Druckverbreiterung der Spektrallinien möglich ist. Vom Boden aus durchgeführte Messungen entsprechender Konstituenten erlauben damit die Bestimmung von Temperatur und Volumenmischungsverhältnis als Funktion der Höhe [32].

Der Zusammenhang zwischen dem Strahlungsfeld und dem Zustand der Atmosphäre wird durch die Strahlungsübertragungsgleichung (SÜG) beschrieben [10]. 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, nämlich das Auffinden einer bestimmten Lösung der komplizierten SÜG. Beide Probleme sind nichtlinear, da die zu ermittelnden Parameter, wie die Höhenverteilungen der Volumenmischungsverhältnisse interessierender Spurengase oder der Temperatur, nichtlinear über die Transmissionsfunktion in die SÜG eingehen. Erschwerend kommt hinzu, daß das inverse Problem *schlecht gestellt* ist im Sinne von Hadamard [30]: Kleine Meßfehler in den Strahldichtespektren führen demnach – ohne besondere Vorkehrungen – zu großen Fehlern im Lösungsprofil. In der Folge bedarf die Lösung einer sehr sorgfältigen numerischen Behandlung [19].

Entsprechend aufwendig sind die hierfür bereitzustellenden Auswertalgorithmen, so daß Entwicklungsarbeit geleistet werden muß, bevor aus den gewonnenen Spektren Konzentrationsprofile interessierender Gasprofile berechnet werden können. Ein entsprechendes Computerprogramm wurde im Rahmen dieser Arbeit für den Millimeterwellenbereich entwickelt und getestet. Dies schließt die Entwicklung eines effizienten Vorwärts–Programmes für die Lösung des direkten Problems mit ein.

Die vorliegende Arbeit gliedert sich in zwei Teile: Einen theoretischen Teil, dieser umfaßt die Kapitel 2 bis 4, in welchen die Problematik der Inversion der SÜG diskutiert und die Entwicklung eines geeigneten Algorithmus beschrieben wird. Sowie einen praktischen Teil, im wesentlichen bestehend aus Kapitel 5. Dieser widmet sich der praktischen Anwendung dieser Methode und demonstriert ihre Funktionstüchtigkeit. Im Detail gliedert sich die Arbeit wie folgt:

In Kapitel 2 wird zunächst die Implementierung der SÜG in ein Vorwärtsprogramm beschrieben und anschließend das Problem des Auffindens einer Lösung der komplizierten SÜG reduziert auf das Lösen eines linearen Gleichungssystems, welches jedoch *schlecht gestellt* ist.

Das Kapitel 3 ist ganz der Lösung dieses linearen Gleichungssystems gewidmet. Hierzu wird zunächst eine Einführung in die Theorie der Inversion *schlecht gestellter* Probleme gegeben und anschließend werden drei *Regularisierungsverfahren*, die abgeschnittene Singulärwertzerlegung, die Tikhonov–Phillips–Regularisierung sowie das linear statistische Verfahren, vorgestellt.

In Kapitel 4 erfolgt die Erweiterung der beiden letztgenannten Regularisierungsverfahren auf den Fall der simultanen Inversion mehrerer Spurengase aus einem Spektrum. Überdies wird eine Erweiterung der Regularisierungsverfahren vorgeschlagen, mit der von stehenden Wellen überlagerte Spektren direkt invertiert werden können, ohne diese zuvor einer aufwendigen Elimination von stehenden Wellen unterziehen zu müssen. Beide Erweiterungen sind in diesem Kontext Neuentwicklungen.

Im Anschluß an diesen theoretischen Teil folgt in Kapitel 5 eine Reihe von Inversionsrechnungen auf der Basis synthetisch erzeugter Spektren. Hieran schließt sich die Auswertung einer Anzahl gemessener Spektren des Spurengases Ozon, aufgenommen im März und April 1994 über der Hornisgrinde (Nordschwarzwald, 1165 m N.N.) an. Für diese Messungen wird eine detaillierte Fehlerbetrachtung durchgeführt. Schließlich wird am Ende des Kapitels noch eine Untersuchung zur Meßbarkeit des Spurengases ClO für den Standort Hornisgrinde angestellt.

Den Abschluß bildet eine kurze Zusammenfassung der Arbeit in Kapitel 6.

Hinweis:

Um dem Leser die Übersicht über die teilweise relativ komplizierten Formeln der folgenden Abschnitte zu erleichtern, wurden die runden Klammern () der Kennzeichnung funktionaler Abhängigkeiten vorbehalten. Zur Kennzeichnung einer Klammerung wurden statt dessen ausschließlich die eckigen Klammern [] verwendet.

Kapitel 2

Grundbegriffe der Strahlungsübertragung

2.1 Die Strahlungsübertragungsgleichung (SÜG)

2.1.1 Die differentielle Form der SÜG

Die klassische Arbeit auf dem Gebiet der Strahlungsübertragung ist die von Chandrasekhar [10]. Ausgangspunkt seiner Betrachtungen ist die Beschreibung des Strahlungsfeldes über die spektrale Strahldichte L_ν . Hierunter versteht man die aus einer vorgegebenen Richtung bei gegebener Frequenz einfallende Strahlungsintensität, bezogen auf die Einheitsfläche, Einheitsfrequenzintervall und Raumwinkel. Die Einheit der spektralen Strahldichte ist $\text{W} / (\text{sr m}^2 \text{s}^{-1})$. Wie in Abb. 1 skizziert, ändert sich die

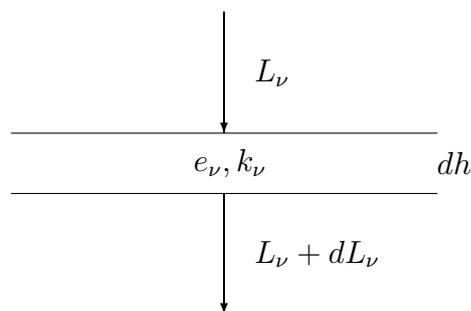


Abbildung 1: Änderung der spektralen Strahldichte L_ν beim Durchgang durch eine Schicht der Dicke dh .

spektrale Strahldichte beim Durchgang durch eine atmosphärische Schicht der Dicke dh aufgrund der innerhalb der Schicht emittierten Strahlung $e_\nu dh$ sowie der innerhalb der Schicht erfolgenden Absorption $k_\nu L_\nu dh$. Dabei ist e_ν der Quellterm (Emissivität)

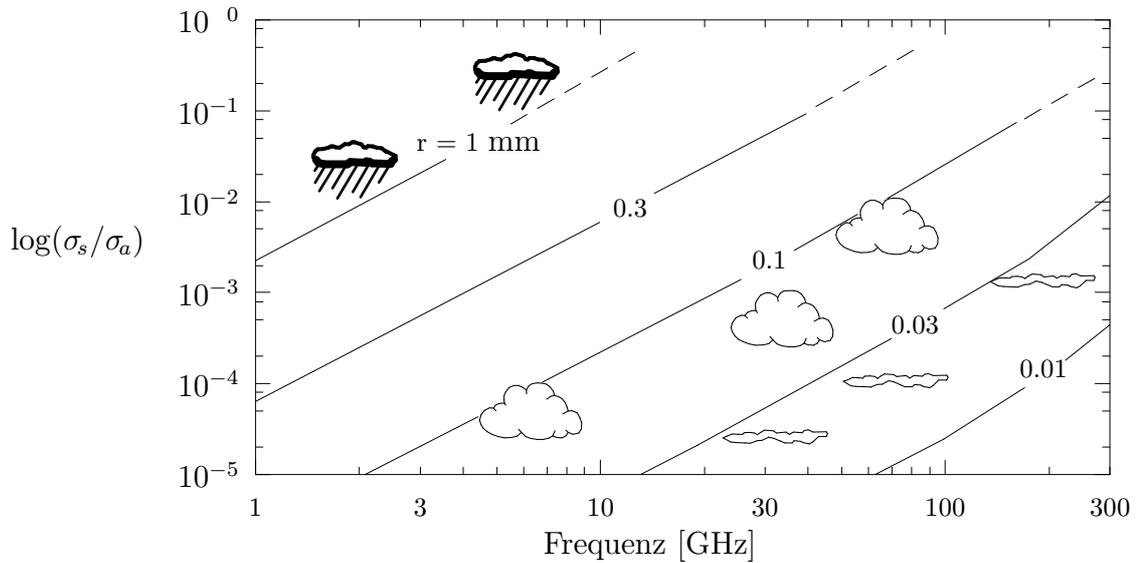


Abbildung 2: Verhältnis von Streuquerschnitt σ_s zu Absorptionsquerschnitt σ_a für Wassertröpfchen mit Radius r nach [35].² Die für die verschiedene Wolken typischen Tröpfchengrößen sind jeweils mit angegeben.

und k_ν der Absorptionskoeffizient. Dies führt auf die *differentielle Form der Strahlungsübertragungsgleichung (SÜG)*,¹

$$dL_\nu = e_\nu dh - k_\nu L_\nu dh = k_\nu [B_\nu(T) - L_\nu] dh. \quad (1)$$

Streuungseffekte, die ebenfalls zu einer Zu- oder Abnahme der spektralen Strahlendichte entlang der Ausbreitungsrichtung führen, werden vernachlässigt. Dies ist im Millimeterwellenbereich gerechtfertigt, da Aerosol und typische Wolkenpartikel nur schwach streuen (siehe Abb. 2). Bei starker Bewölkung und Regen hingegen darf die Streuung nicht mehr vernachlässigt werden, was die Lösung der SÜG erheblich kompliziert [10, 28, 34, 35, 87]. Liebe [42]–[47] stellt für diesen Fall empirische Korrekturmöglichkeiten bereit. Andererseits kann man bei solch widrigen Witterungsverhältnissen jedoch davon ausgehen, daß die untere Troposphäre optisch so dicht ist, daß selbst mit der im Vergleich zum infraroten Bereich gegenüber Witterungseinflüssen weniger empfindlichen Millimeterwellenradiometrie auch stark emittierende stratosphärische Spurengase, wie z. B. Ozon, vom Boden aus nicht mehr detektiert werden können. Von daher bedeutet die Beschränkung auf eine nicht streuende Atmosphäre für die Auswertung von Meßdaten keine wesentliche Einschränkung.

¹In Gl. (1) wird vorausgesetzt, daß die Bedingungen (1) des lokalen thermodynamischen Gleichgewichts erfüllt sind, daß also der Quellterm e_ν gleich dem Produkt aus Absorptionskoeffizient k_ν und Planckfunktion $B_\nu(T)$ ist.

²Der Absorptions- bzw. Streukoeffizient ergibt sich aus dem Absorptions- bzw. Streuquerschnitt durch Multiplikation mit der Anzahldichte der absorbierenden bzw. streuenden Teilchen.

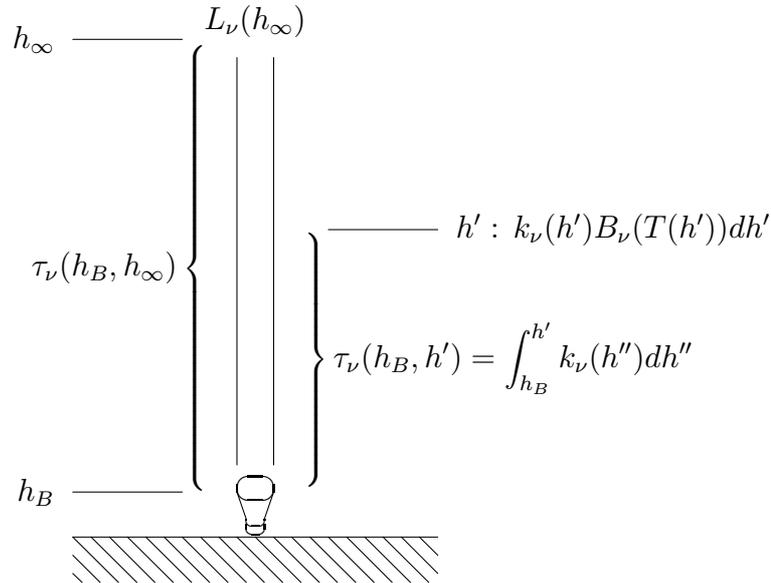


Abbildung 3: Strahlungsübertragung innerhalb der Atmosphäre.

2.1.2 Die integrale Form der SÜG

Maßgeblich für die Berechnung der zu empfangenden Strahlendichte bei bekanntem Zustand der Atmosphäre (Druck, Temperatur und Zusammensetzung) ist die SÜG in ihrer integralen Form. Sie ergibt sich durch einfache Integration aus der differentiellen Form (1). Für eine nichtstreuende Atmosphäre im lokalen thermodynamischen Gleichgewicht mit Blickrichtung senkrecht nach oben lautet sie [10, 89]³

$$L_\nu(h_B) = e^{-\tau_\nu(h_B, h_\infty)} L_\nu(h_\infty) + \int_{h_B}^{h_\infty} \underbrace{k_\nu(h') B_\nu(T(h'))}_{e_\nu(h)} e^{-\tau_\nu(h_B, h')} dh'. \quad (2)$$

Dabei ist $L_\nu(h_B)$ die in der Beobachterhöhe h_B bei der Frequenz ν meßbare spektrale Strahlendichte, $B_\nu(T)$ die Planckfunktion, $k_\nu(h')$ der höhenabhängige, monochromatische Volumenabsorptionskoeffizient und

$$\tau_\nu(h_B, h') = \int_{h_B}^{h'} k_\nu(h'') dh'' \quad (3)$$

die optische Dicke der Atmosphäre zwischen Beobachterhöhe h_B und h' .

³Im folgenden wird der besseren Übersichtlichkeit willen nur der Fall einer Blickrichtung senkrecht nach oben behandelt. Die im Rahmen dieser Arbeit entwickelten Algorithmen unterstützen selbstverständlich auch die Anwendung auf frei wählbare Elevationswinkel.

Die Beiträge auf der rechten Seite der SÜG (2) können physikalisch leicht interpretiert werden (siehe Abb. 3). Der erste Term ist der Beitrag der kosmischen Hintergrundstrahlung, $L_\nu(h_\infty)$, die vom Weltraum kommend in die Atmosphäre eintritt, vermindert um die beim Durchgang durch die gesamte Atmosphäre erfolgende Absorption (Faktor $e^{-\tau_\nu(h_B, h_\infty)}$). Der zweite Term setzt sich zusammen aus den Beiträgen des gesamten Höhenbereichs (Integral von h_B bis h_∞), wobei die in der betrachteten Höhe h' emittierte Strahlung $k_\nu(h')B_\nu(T(h'))dh'$, veringert wird um die auf dem Weg durch die Restatmosphäre erfolgende Absorption, $e^{-\tau_\nu(h_B, h')}$.

2.1.3 Implementierung der SÜG

Die SÜG in ihrer integralen Form (2) ist Ausdruck des direkten oder Vorwärts-Problems. Zur numerischen Berechnung der in der Beobachterhöhe h_B zu empfangenden spektralen Strahldichte $L_\nu(h_B)$ wird die reale, inhomogene Atmosphäre approximiert durch ein Atmosphärenmodell mit homogenen Schichten bzgl. des Druckes, der Temperatur und des Volumenmischungsverhältnisses. Man spricht auch von einer Quadratur der SÜG [10, Kap. 2]. Dabei darf

1. die Schichtdicke nicht zu groß gewählt werden und
2. außerdem dürfen die Schichtmittelwerte für Druck, Temperatur und Volumenmischungsverhältnis nicht einfach als arithmetisches Mittel der Werte an den Schichtgrenzen berechnet werden, sondern müssen der sich mit der Höhe ändernden Dichte der absorbierenden Gase Rechnung tragen.

Es ist üblich, die Inhomogenität der Atmosphäre durch die sogenannten *Curtis–Godson–Mittelwerte* zu repräsentieren [11, 22, 66]:⁴

$$p_j = \frac{\int_{h_j}^{h_{j+1}} p(h')n_g(h')dh'}{\int_{h_j}^{h_{j+1}} n_g(h')dh'}, \quad T_j = \frac{\int_{h_j}^{h_{j+1}} T(h')n_g(h')dh'}{\int_{h_j}^{h_{j+1}} n_g(h')dh'} \quad \text{und} \quad c_{vj} = \frac{\int_{h_j}^{h_{j+1}} c_v(h')n_g(h')dh'}{\int_{h_j}^{h_{j+1}} n_g(h')dh'}. \quad (4)$$

Dabei bedeuten p_j , T_j und c_{vj} die gewichteten Schichtmittelwerte für Druck, Temperatur und Volumenmischungsverhältnis, n_g ist die Anzahldichte der absorbierenden Gasmoleküle und j die Nummer der unteren Schichtgrenze. Entsprechend der Gleichung (4) erhält man für verschiedene Spurengasprofile auch unterschiedliche Schichtmittelwerte.

⁴Auch die Curtis–Godson–Mittelwerte sind nur für den Fall einer Blickrichtung senkrecht nach oben angegeben. Im allgemeinen Fall muß die Änderung des Wegelementes mit berücksichtigt werden, wie dies im Vorwärtsprogramm geschieht.

Für die numerische Implementierung der SÜG (2) in ein Vorwärts-Programm ist es zweckmäßig, zunächst die Wirkung einer einzelnen Schicht auf den Verlauf der Strahldichte näher zu untersuchen. Sei $\Delta h_j = h_{j+1} - h_j$ die Dicke der j -ten Schicht sowie p_j , T_j und $c_{\nu j}$ die über (4) berechneten Schichtmittelwerte von Druck, Temperatur und Volumenmischungsverhältnis. Weiterhin sei $k_{\nu j}$ der aus diesen Werten resultierende Absorptionskoeffizient und $L_\nu(h_{j+1})$ die spektrale Strahldichte an der oberen Schichtgrenze h_{j+1} . Für die Strahldichte $L_\nu(h_j)$ nach Durchgang durch die Schicht folgt dann mit Hilfe der Gln. (2) und (3)

$$L_\nu(h_j) = e^{-k_{\nu j} \Delta h_j} L_\nu(h_{j+1}) + k_{\nu j} B_\nu(T_j) \int_{h_j}^{h_{j+1}} e^{-k_{\nu j} [h' - h_j]} dh',$$

so daß

$$\boxed{L_\nu(h_j) = e^{-k_{\nu j} \Delta h_j} L_\nu(h_{j+1}) + [1 - e^{-k_{\nu j} \Delta h_j}] B_\nu(T_j).} \quad (5)$$

Die resultierende spektrale Strahldichte $L_\nu(h_j)$ ergibt sich also einfach als Summe der einfallenden Strahldichte $L_\nu(h_{j+1})$ und der Planckfunktion $B_\nu(T_j)$, gewichtet mit den jeweiligen, von der Dicke der Schicht und dem Absorptionkoeffizienten abhängigen Vorfaktoren.

Es ist lehrreich, sich die Wirkungsweise der Gl. (5) anhand dreier Beispiele zu verdeutlichen [89]. Zunächst der Fall eines transparenten Mediums, $k_{\nu j} = 0$, für das sich (5) vereinfacht zu $L_\nu(h_j) = L_\nu(h_{j+1})$; d. h. die Strahldichte an der unteren Schichtgrenze h_j ist gleich der einfallenden Strahldichte, wie für ein transparentes Medium nicht anders zu erwarten. Zweitens der Fall eines undurchlässigen Mediums, $\Delta h_j k_{\nu j} \gg 1$, so daß $L_\nu(h_j) = B_\nu(T_j)$; die resultierende Strahldichte ist gleich der innerhalb der Schicht emittierten Strahlung. Und drittens der Fall eines Mediums, das so temperiert ist, daß $L_\nu(h_{j+1}) = B_\nu(T_j)$; das Medium emittiert genau so viel wie es absorbiert. Die Strahldichte ändert sich längs der Schicht nicht (unabhängig vom Absorptionskoeffizienten $k_{\nu j}$).

Ist die Wirkung einer einzelnen Schicht auf die spektrale Strahldichte einmal bekannt, kann die integrale Wirkung eines ganzen Schichtenmodells relativ einfach berechnet werden. Am effizientesten erhält man die spektrale Strahldichte in Beobachterhöhe $L_\nu(h_B)$, indem man ausgehend vom oberen Rand der Atmosphäre, also dem Beitrag der kosmischen Hintergrundstrahlung $L_\nu(h_\infty)$, die Gleichung (5) sukzessive auf die jeweils darunterliegende Schicht anwendet. Dabei sollte die Schichtdicke mit abnehmender Höhe immer feiner gewählt werden, um der nach unten hin zunehmenden Dichte der Atmosphäre Rechnung zu tragen. Ist die Beobachterhöhe h_B schließlich erreicht, wird das Verfahren abgebrochen. (In Anhang A.1 wird ein Vergleich des im Rahmen dieser Arbeit entwickelten Vorwärtsprogrammes mit anderen Programmen durchgeführt).

2.2 Sensitivitäts- und Gewichtsfunktionen

2.2.1 Zur Unterscheidung zwischen beiden Funktionen

Von großer Bedeutung im Hinblick auf das übergeordnete Ziel – die Ableitung von Höhenprofilen aus den gemessenen Strahldichtespektren – ist die Frage, wie stark sich eine Änderung des Volumenmischungsverhältnisses c_v oder eine Änderung der Temperatur T in einer bestimmten Höhe auf die unten zu empfangende spektrale Strahldichte $L_\nu(h_B)$ auswirkt.⁵ Die Antwort auf diese Frage führt auf den Begriff der *Sensitivitätsfunktion*. Wohl unterschieden werden hiervon muß die Frage, mit welchem Gewicht das Volumenmischungsverhältnis bzw. die Temperatur einer bestimmten Höhe zu der unten meßbaren spektralen Strahldichte beitragen; das Ergebnis dieser Betrachtungsweise führt auf den Begriff der *Gewichtsfunktion*.⁶

Anhand des obigen Atmosphärenmodells sei der Unterschied zwischen beiden Funktionen näher erläutert. Hierzu ist anzumerken, daß im folgenden die Sensitivitäts- und Gewichtsfunktionen $S_\nu(h)$ bzw. $G_\nu(h)$ – entsprechend der diskreten Form des Atmosphärenmodells – zunächst nur in ihrer diskretisierten Form auftreten werden, gekennzeichnet durch eine runde Klammer mit tiefgestelltem Index für die jeweilige Höhenschicht, z. B. $(S_\nu)_j$ oder $(G_\nu)_j$. Erst gegen Ende des nächsten Abschnitts wird auch ein geschlossener analytischer Ausdruck für die Sensitivitätsfunktionen abgeleitet werden. Bis dahin jedoch soll sich der Begriff der Funktion – wenn im folgenden von Sensitivitäts- und Gewichtsfunktionen die Rede sein wird – auf die diskretisierte Form beziehen. Bezeichne nun $L_\nu(h_B)_j$ den Beitrag der j -ten Höhenschicht zu der in Beobachterhöhe meßbaren spektralen Strahldichte $L_\nu(h_B)$ (Dämpfung auf dem Weg durch die Restatmosphäre inbegriffen), so daß

$$L_\nu(h_B) = \sum_j L_\nu(h_B)_j.$$

Weiterhin seien c_{vj} und T_j die Mittelwerte von Volumenmischungsverhältnis und Temperatur der Schicht. Für die Sensitivität $(S_{\nu c_v})_j$, die ein Maß ist für die marginale Änderung der spektralen Strahldichte aufgrund einer marginalen Änderung des Volumenmischungsverhältnisses der j -ten Höhenschicht, folgt hieraus

$$(S_{\nu c_v})_j \propto \frac{\partial L_\nu(h_B)_j}{\partial c_{vj}} = \frac{\partial L_\nu(h_B)}{\partial c_{vj}}.$$

⁵Oder anders formuliert, wie sensibel die spektrale Strahldichte auf Änderungen des Volumenmischungsverhältnisses bzw. der Temperatur in einer bestimmten Höhe reagiert.

⁶Im infraroten Spektralbereich wird zusätzlich noch zwischen Gewichts- und Beitragsfunktionen unterschieden. Im Millimeterwellenbereich ist diese Unterscheidung obsolet, da die Planckfunktion der Rayleigh-Jeans-Näherung genügt. Die Planckfunktion ist also im wesentlichen proportional zur Temperatur, so daß die Gewichts- und die Beitragsfunktionen sich nur um einen konstanten Faktor unterscheiden.

Entsprechend ergibt sich bzgl. der Temperatur

$$(S_{\nu T})_j \propto \frac{\partial L_{\nu}(h_B)_j}{\partial T_j} = \frac{\partial L_{\nu}(h_B)}{\partial T_j}.$$

Hingegen gilt für das Gewicht, mit dem das Volumenmischungsverhältnis bzw. die Temperatur der j -ten Schicht zu der unten meßbaren spektralen Strahldichte beiträgt

$$(G_{\nu c_v})_j \propto \frac{L_{\nu}(h_B)_j}{c_{vj}} \quad \text{bzw.} \quad (G_{\nu T})_j \propto \frac{L_{\nu}(h_B)_j}{T_j}.$$

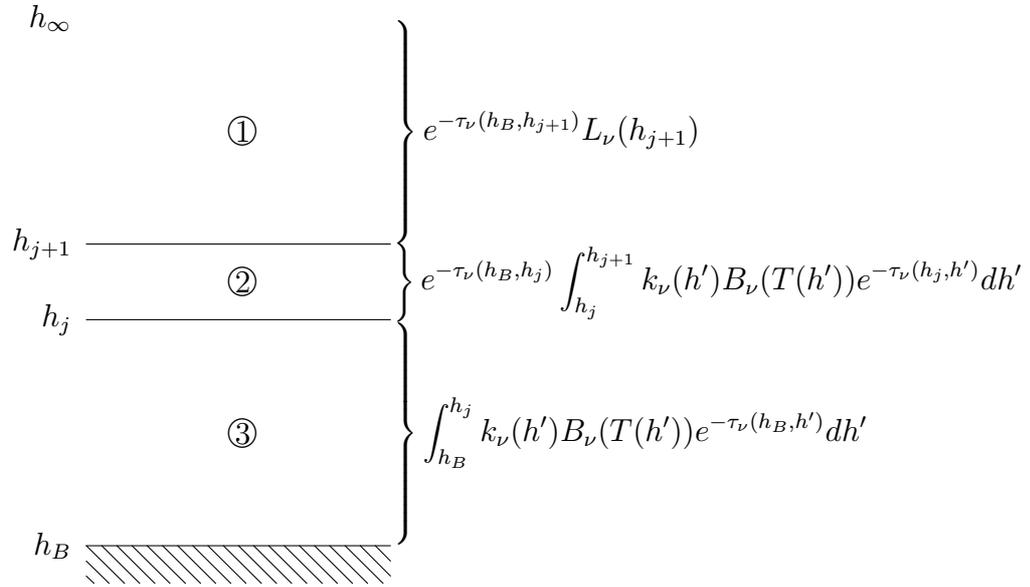
Die Sensitivitäts- und Gewichtsfunktionen sind also nur dann gleich, wenn zwischen der spektralen Strahldichte und dem Volumenmischungsverhältnis – bzw. zwischen der spektralen Strahldichte und der Temperatur – ein linearer Zusammenhang besteht. Die SÜG (2) an sich ist jedoch nichtlinear, da sowohl das Volumenmischungsverhältnis als auch die Temperatur nichtlinear über den Absorptionskoeffizienten k_{ν} in die SÜG eingehen (siehe Gl. (3)). Zwischen Sensitivitäts- und Gewichtsfunktionen muß daher streng unterschieden werden. Leider wird diesem Umstand in der Millimeterwellenradiometrie oftmals nicht Rechnung getragen; statt dessen werden die Sensitivitätsfunktionen fälschlicherweise als Gewichtsfunktionen bezeichnet oder wird gar bei der Inversion der SÜG statt mit den Sensitivitätsfunktionen mit Gewichtsfunktionen gerechnet.

2.2.2 Berechnung der Sensitivitätsfunktionen

Für die Inversion der SÜG werden, wie in Abschn. 2.3 noch deutlich wird, lediglich die Sensitivitätsfunktionen $(S_{\nu c_v})_j$ und $(S_{\nu T})_j$ bzw. die entsprechenden partiellen Ableitungen $\partial L(h_B)/\partial c_{vj}$ und $\partial L(h_B)/\partial T_j$ benötigt. Sie sollen deshalb im folgenden soweit als möglich auf analytischem Wege berechnet werden. Die Gewichtsfunktionen $(G_{\nu c_v})_j$ und $(G_{\nu T})_j$ hingegen spielen bei der Inversion der SÜG eine eher untergeordnete Rolle, fallen jedoch bei der analytischen Berechnung der Sensitivitätsfunktionen als Nebenprodukt quasi automatisch mit an.

Grundsätzlich besteht die Möglichkeit, die Sensitivitätsfunktionen auch auf rein numerischem Wege zu erhalten. Das Vorwärtsmodell stellt in diesem Falle eine Art „black box“ dar, welche für ein gegebenes Volumenmischungsverhältnis- und Temperaturprofil die zugehörige spektrale Strahldichte $L_{\nu}(h_B)$ als Output liefert. Statt der obigen partiellen Ableitungen können dann die Differenzenquotienten $\Delta L(h_B)/\Delta c_{vj}$ bzw. $\Delta L(h_B)/\Delta T_j$ berechnet werden, indem man für ein gegebenes Höhenprofil sowie ein zweites, gegenüber dem ersten in einer bestimmten Höhe leicht variiertes Höhenprofil die zugehörigen spektralen Strahldichten ermittelt. Diese Vorgehensweise ist sehr rechenzeitaufwendig, da für jeden Iterationsschritt ⁷ – abhängig von der Anzahl

⁷Warum bei der Inversion der SÜG iteriert werden muß, wird deutlich werden in Abschn. 2.3.1 über das Gauß-Newton-Verfahren. Es ist jedoch unmittelbar einleuchtend, daß ein nichtlineares Problem, wie die SÜG, immer nur iterativ gelöst werden kann, sofern überhaupt eine Lösung existiert.



$$L_\nu(h_B) = \textcircled{1} + \textcircled{2} + \textcircled{3}$$

Abbildung 4: In die Linearisierung der SÜG involvierte Geometrie nach [93]. Die in den Beiträgen ① und ② auftretenden Exponentialfunktionen beschreiben die auf dem Weg durch die Restatmosphäre zum Beobachter erfolgende Dämpfung.

der zu invertierenden Spurengase und der Anzahl der Höhenschichten – oft weit über 100 Vorwärtsrechnungen durchgeführt werden müssen, welche im Falle einer analytischen Berechnung auf eine einzige Vorwärtsrechnung je Iterationsschritt hätten reduziert werden können.

Für die analytische Berechnung der Sensitivitätsfunktionen ist es zweckmäßig, zunächst eine beliebige, dünne Schicht (die j -te) aus dem Atmosphärenmodell herauszugreifen, und die integrale SÜG (2) aufzuspalten in Beiträge, die von oberhalb dieser Schicht stammen, die innerhalb dieser Schicht emittiert werden bzw. die von unterhalb der Schicht kommen [93] (siehe Abb. 4).⁸

$$L_\nu(h_B) = e^{-\tau_\nu(h_B, h_{j+1})} L_\nu(h_{j+1}) \quad \textcircled{1}$$

$$+ e^{-\tau_\nu(h_B, h_j)} \int_{h_j}^{h_{j+1}} k_\nu(h') B_\nu(T(h')) e^{-\tau_\nu(h_j, h')} dh' \quad \textcircled{2}$$

$$+ \int_{h_B}^{h_j} k_\nu(h') B_\nu(T(h')) e^{-\tau_\nu(h_B, h')} dh' \quad \textcircled{3}$$

⁸Eine andere Möglichkeit zur Berechnung der Sensitivitätsfunktionen existiert über die Variationsrechnung, siehe z. B. [62, S. 76].

Da der Absorptionskoeffizient und die Temperatur innerhalb der dünnen Schicht als konstant angenommen werden können, $k_{\nu j}, T_j$, vereinfacht sich die obige Gleichung in Anlehnung an Gl. (5) zu

$$\begin{aligned} L_{\nu}(h_B) &= e^{-\tau_{\nu}(h_B, h_j)} e^{-k_{\nu j} \Delta h_j} L_{\nu}(h_{j+1}) \\ &+ e^{-\tau_{\nu}(h_B, h_j)} [1 - e^{-k_{\nu j} \Delta h_j}] B_{\nu}(T_j) \\ &+ \int_{h_B}^{h_j} k_{\nu}(h') B_{\nu}(T(h')) e^{-\tau_{\nu}(h_B, h')} dh'. \end{aligned} \quad (6)$$

Sensitivitätsfunktionen für das Volumenmischungsverhältnis

Gesucht werden die partiellen Ableitungen der zu empfangenden spektralen Strahl-dichte $L_{\nu}(h_B)$ nach dem Volumenmischungsverhältnis $c_{\nu j}$ bzw. nach der Temperatur T_j der betrachteten Schicht. Erstere sind einfacher zu berechnen, da das Volumenmischungsverhältnis – im Gegensatz zur Temperatur – lediglich über den Absorptionskoeffizienten k_{ν} und nicht zusätzlich über die Planckfunktion in die SÜG eingeht. Aus (6) folgt unmittelbar

$$\frac{\partial L_{\nu}(h_B)}{\partial k_{\nu j}} = \Delta h_j e^{-\tau_{\nu}(h_B, h_{j+1})} [B_{\nu}(T_j) - L_{\nu}(h_{j+1})], \quad (7)$$

so daß

$$\begin{aligned} \frac{\partial L_{\nu}(h_B)}{\partial c_{\nu j}} &= \frac{\partial L_{\nu}(h_B)}{\partial k_{\nu j}} \frac{\partial k_{\nu j}}{\partial c_{\nu j}} \\ &= \Delta h_j e^{-\tau_{\nu}(h_B, h_j)} \underbrace{\frac{\partial k_{\nu j}}{\partial c_{\nu j}} [B_{\nu}(T_j) - L_{\nu}(h_{j+1})]}_{(S_{\nu c_{\nu}})_j}. \end{aligned} \quad (8)$$

Gl. (8) liefert bereits den gesuchten Zusammenhang zwischen einer marginalen Änderung des Volumenmischungsverhältnisses $c_{\nu j}$ und der Reaktion der spektralen Strahl-dichte $L_{\nu}(h_B)$ auf diese Variation. In ihrer diskretisierten Form (8) ist sie besonders für eine Implementierung im Programm geeignet, wobei die Größe $L_{\nu}(h_{j+1})$ leicht mit Hilfe von (5) ermittelt werden kann. Die Implementierung erfolgt ähnlich wie bei der Integration der Gl. (5), indem wiederum ausgehend vom oberen Rand der Atmosphäre, das Atmosphärenmodell Schicht für Schicht abgearbeitet wird. Probleme bereitet jedoch die Transmission $e^{-\tau_{\nu}(h_B, h_j)}$, da die optische Dicke $\tau_{\nu}(h_B, h_j)$ zwischen der Höhe des Beobachters h_B und der Untergrenze der j -ten Schicht h_j nicht von vornherein bekannt ist, sondern erst im weiteren Verlauf der Bearbeitung anfällt. Doch kann man die Transmission der Atmosphäre unterhalb der aktuellen Schicht auch darstellen als Produkt der Einzeltransmissionen aller darunterliegenden Schichten $e^{-\tau_{\nu}(h_B, h_j)} = e^{-\tau_{\nu}(h_B, h_2)} \dots e^{-\tau_{\nu}(h_{j-2}, h_{j-1})} e^{-\tau_{\nu}(h_{j-1}, h_j)}$. Von daher besteht die

Möglichkeit, die Gleichung (8) zunächst bis auf die Transmission zu berechnen und dann im nachhinein bei Bearbeitung der darunterliegenden Schichten, wenn die optischen Dicken $\tau_\nu(h_B, h_2), \dots, \tau_\nu(h_{j-1}, h_j), \tau_\nu(h_{j-2}, h_{j-1})$ etc. bekannt sind, mit den entsprechenden Transmissionen zu multiplizieren.

Aus Gl. (8) kann man auch einen geschlossenen analytischen Ausdruck für die Antwort der spektralen Strahldichte $L_\nu(h_B)$ auf eine Variation des Volumenmischungsverhältnisses c_ν in einer bestimmten Höhe h herleiten, indem man $L_\nu(h_{j+1})$ darstellt wie in Gl. (2) und zu infinitesimal dünnen Schichten übergeht, $\Delta h_j \rightarrow \delta h$. Es folgt

$$\left(\frac{\partial L_\nu(h_B)}{\partial c_\nu(h)} \right)_h = \delta h S_{\nu c_\nu}(h), \quad (9)$$

wobei

$$S_{\nu c_\nu}(h) = \overbrace{e^{-\tau_\nu(h_B, h)} \frac{\partial k_\nu(h)}{\partial c_\nu(h)}}^{G_{\nu c_\nu}(h)} \left[B_\nu(T(h)) - \left\{ e^{-\tau(h, h_\infty)} L_\nu(h_\infty) + \int_h^{h_\infty} B_\nu(T(h')) k_\nu(h') e^{-\tau_\nu(h, h')} dh' \right\} \right].$$

Die $S_{\nu c_\nu}(h)$ sind die analytischen Sensitivitätsfunktionen für das Volumenmischungsverhältnis. Sie enthalten die analytischen Gewichtsfunktionen $G_{\nu c_\nu}(h)$ als Teilfunktionen. Anhand der obigen Gleichung wird deutlich, daß die Sensitivitäts- und Gewichtsfunktionen einander nur dann entsprechen, wenn der in der geschweiften Klammer stehende Ausdruck klein ist gegen $B_\nu(T(h))$, d. h. nur dann, wenn die optische Dicke der Atmosphäre klein ist und keine Selbstabsorption auftritt. Im Falle der bodengebundenen Millimeterwellenradiometrie ist diese Bedingung jedoch zumindest für die stärker absorbierenden Spurengase wie Ozon nicht mehr erfüllt. Dies gilt insbesondere dann, wenn das Radiometer nicht senkrecht nach oben, sondern unter einem relativ flachen Elevationswinkel auf die Atmosphäre gerichtet ist.⁹ In diesem Fall sollte unbedingt mit den Sensitivitätsfunktionen gerechnet werden. Das bedeutet jedoch nicht,

⁹Typische Einstellungen sind Elevationswinkel um 20°. Im Falle der 142,175 GHz-Signatur von Ozon (vom Einfluß des troposphärischen Wasserdampfes werde im Moment abgesehen) ergibt sich bei diesem Winkel bereits eine maximale Helligkeitstemperatur von 50–60 K in der Liniennitte (dies entspricht dem maximalen Wert, den die geschweifte Klammer in obiger Gleichung in etwa annehmen kann). Dem gegenüber ist die mittlere Temperatur der Atmosphäre von etwa 250 K zu sehen (dieser Wert müßte im Rahmen der Rayleigh–Jeans–Näherung für die Helligkeitstemperatur der Planckfunktion angesetzt werden). Wird nun anstelle der Sensitivitätsfunktionen mit den Gewichtsfunktionen gerechnet, ergibt sich für die eckige Klammer in obiger Gleichung ein Wert von 250 K gegenüber dem wahren Wert von 250 K – 55 K = 195 K. Hieraus resultiert ein relativer Fehler von bis zu

$$\frac{250 \text{ K} - 195 \text{ K}}{195 \text{ K}} \approx 28\%,$$

um den die Gewichtsfunktionen im unteren Bereich der Atmosphäre gegenüber den Sensitivitätsfunktionen zu groß sind.

daß eine Inversion der SÜG (2) allein auf der Basis der Gewichtsfunktionen grundsätzlich unmöglich ist.¹⁰ Allerdings dürften – sofern das Verfahren überhaupt konvergiert – i. d. Regel mehr Iterationsschritte benötigt werden als bei einer Verwendung der Sensitivitätsfunktionen, da in diesem Fall die in jeden Iterationsschritt einfließenden partiellen Ableitungen (8) nicht exakt berechnet werden.

Sensitivitätsfunktionen für die Temperatur

Für die partielle Ableitung der spektralen Strahldichte $L_\nu(h_B)$ nach der sowohl in der Planckfunktion als auch im Absorptionskoeffizienten enthaltenen Temperatur folgt im Grenzfall dünner Schichten aus (6) zunächst

$$\frac{\partial L_\nu(h_B)}{\partial B_\nu(T_j)} = e^{-\tau_\nu(h_B, h_j)} \left[1 - e^{-k_{\nu j} \Delta h_j} \right] \approx \Delta h_j k_{\nu j} e^{-\tau_\nu(h_B, h_j)}$$

und daraus mit Hilfe von (7)

$$\boxed{\begin{aligned} \frac{\partial L_\nu(h_B)}{\partial T_j} &= \frac{\partial L_\nu(h_B)}{\partial B_\nu(T_j)} \frac{\partial B_\nu(T_j)}{\partial T_j} + \frac{\partial L_\nu(h_B)}{\partial k_{\nu j}} \frac{\partial k_{\nu j}}{\partial T_j} \\ &= \Delta h_j e^{-\tau_\nu(h_B, h_j)} \underbrace{\left[\frac{\partial B_\nu(T_j)}{\partial T_j} k_{\nu j} + \frac{\partial k_{\nu j}}{\partial T_j} \{B_\nu(T_j) - L_\nu(h_{j+1})\} \right]}_{(S_{\nu T})_j}. \end{aligned}} \quad (10)$$

Gl. (10) stellt das Analogon zur Gl. (8) bzgl. der Temperatur dar. Sie beschreibt die Antwort der zu empfangenden spektralen Strahldichte auf eine Variation der Temperatur der j -ten Höhengschicht. Auch sie kann leicht in ein Programm implementiert werden.

Für den analytischen Ausdruck folgt in Analogie zu (9)

$$\left(\frac{\partial L_\nu(h_B)}{\partial T(h)} \right)_h = \delta h S_{\nu T}(h),$$

wobei

$$\begin{aligned} S_{\nu T}(h) &= \overbrace{e^{-\tau_\nu(h_B, h)} \left[\frac{\partial B_\nu(T(h))}{\partial T(h)} k_\nu(h) + \frac{\partial k_\nu(h)}{\partial T(h)} \right]}^{G_{\nu T}(h)} \times \\ &\quad \left\{ B_\nu(T(h)) - e^{-\tau(h, h_\infty)} L_\nu(h_\infty) - \int_h^{h_\infty} B_\nu(T(h')) k_\nu(h') e^{-\tau_\nu(h, h')} dh' \right\}. \end{aligned} \quad (11)$$

¹⁰Z. B. wird dies am Institut für Fernerkundung der Universität Bremen so gehandhabt und dennoch werden gute Inversionsergebnisse erzielt [36, 91].

Die $S_{\nu T}(h)$ sind die analytischen Temperatur-Sensitivitätsfunktionen, welche aufgrund der meist schwachen Abhängigkeit des Absorptionskoeffizienten von der Temperatur vom ersten Term in Gl. (11), d. h. den Gewichtsfunktionen $G_{\nu T}(h)$ dominiert werden. Die Unterscheidung in Sensitivitäts- und Gewichtsfunktionen ist deshalb bzgl. der Temperatur weniger wichtig als für das Volumenmischungsverhältnis, wenngleich dies auch – streng genommen – nur für den Fall $\partial k_{\nu}/\partial T = 0$ gilt.

2.3 Linearisierung der SÜG

Im folgenden Abschnitt sollen die mit der Berechnung der Sensitivitätsfunktionen begonnenen Vorarbeiten für die Ableitung von Höhenprofilen aus den aufgenommenen Strahldichtespektren weiter fortgeführt werden. Die SÜG (2) ist – wie bereits in Abschn. 2.2.1 deutlich wurde – nichtlinear, da das Volumenmischungsverhältnis und die Temperatur nichtlinear über den Absorptionskoeffizienten k_{ν} in die SÜG eingehen. Nun ist vom mathematischen Standpunkt aus betrachtet der Umgang mit nichtlinearen Integralgleichungen äußerst kompliziert. Wesentlich einfacher hingegen sind die linearen Integralgleichungen zu behandeln, wenngleich auch sie gekennzeichnet sind durch zahlreiche, oft geradezu tückische Fallgruben [61, Kap. 18]. Wünschenswert wäre es deshalb, eine linearisierte Form der SÜG zu finden, in die sowohl das Volumenmischungsverhältnis als auch die Temperatur linear eingehen. Die Gewinnung einer solchen linearisierten Form der SÜG verfolgt das im nächsten Abschnitt vorgestellte Gauß-Newton-Verfahren, worunter eine einfache Taylorreihenentwicklung der am Boden meßbaren spektralen Strahldichte um Anfangsschätzwerte für die Höhenprofile von Temperatur und Volumenmischungsverhältnis zu verstehen ist.

2.3.1 Gauß-Newton-Verfahren

Sei $L_{\nu}(h_B)_j$ wie in Abschn. 2.2.1 der Beitrag der j -ten Höhengschicht zu der in Beobachterhöhe meßbaren spektralen Strahldichte $L_{\nu}(h_B)$, so läßt sich die nichtlineare Abhängigkeit dieses Beitrages von der Temperatur T_j sowie dem Volumenmischungsverhältnis c_{vj} der j -ten Schicht darstellen als

$$L_{\nu}(h_B)_j = L_{\nu}(h_B)_j(T_j, c_{vj}). \quad (12)$$

Unter der Linearisierung der SÜG ist – wie bereits oben angedeutet – eine einfache Taylorreihenentwicklung des direkten Modells (12) um eine vorgenommene Anfangsschätzung T_j^0, c_{vj}^0 zu verstehen,

$$\begin{aligned} L_{\nu}(h_B)_j(T_j, c_{vj}) &= L_{\nu}(h_B)_j(T_j^0, c_{vj}^0) \\ &+ \left. \frac{\partial L_{\nu}(h_B)_j}{\partial T_j} \right|_{T_j^0} [T_j - T_j^0] + \left. \frac{\partial L_{\nu}(h_B)_j}{\partial c_{vj}} \right|_{c_{vj}^0} [c_{vj} - c_{vj}^0] \end{aligned}$$

$$+ \mathcal{O}\left([T_j - T_j^0][c_{vj} - c_{vj}^0]\right).$$

Alle Terme höher als erster Ordnung werden vernachlässigt. Mit Hilfe der Gln. (8) und (10) läßt sich dies auch schreiben als

$$L_\nu(h_B)_j(T_j, c_{vj}) - L_\nu(h_B)_j(T_j^0, c_{vj}^0) = \Delta h_j(S_{\nu T})_j \Big|_{T_j^0} [T_j - T_j^0] + \Delta h_j(S_{\nu c_v})_j \Big|_{c_{vj}^0} [c_{vj} - c_{vj}^0].$$

Aufsummiert bzw. aufintegriert über alle Höhengschichten folgt schließlich

$$\begin{aligned} & L_\nu(h_B)(T(h), c_v(h)) - L_\nu(h_B)(T^0(h), c_v^0(h)) \\ &= \sum_j \Delta h_j(S_{\nu T})_j \Big|_{T_j^0} [T_j - T_j^0] + \sum_j \Delta h_j(S_{\nu c_v})_j \Big|_{c_{vj}^0} [c_{vj} - c_{vj}^0] \quad (13) \\ &= \int_{h_B}^{h_\infty} S_{\nu T}(h') \Big|_{T^0} [T(h') - T^0(h')] dh' + \int_{h_B}^{h_\infty} S_{\nu c_v}(h') \Big|_{c_v^0} [c_v(h') - c_v^0(h')] dh'. \end{aligned}$$

Gleichung (13) ist die gesuchte linearisierte Darstellung der SÜG und wird als die *linearisierte Strahlungsübertragungsgleichung* bezeichnet. Sie wird im folgenden Abschnitt noch weiter vereinfacht und auf ein lineares Gleichungssystem reduziert werden.

Zuvor soll jedoch bemerkt werden, daß bzgl. des Volumennmischungsverhältnisses c_v bisher noch nicht näher spezifiziert wurde, auf welches Gas sich diese Angabe denn eigentlich bezieht. Statt dessen war bisher nur ganz allgemein vom Volumennmischungsverhältnis eines beliebigen Gases die Rede. Da sich die reale Atmosphäre aus vielen verschiedenen Gasen zusammensetzt, müßten in der linearisierten SÜG (13) im Prinzip die Beiträge vieler verschiedener (relevanter) Gase berücksichtigt werden, ein jedes mit seinem Volumennmischungsverhältnis und zugehörigen Sensitivitätsfunktionen. Aus Gründen der besseren Übersichtlichkeit wurde hierauf bisher verzichtet. Künftig soll die mögliche Erweiterung der obigen Gleichung auf mehrere Gase jedoch zumindest durch Punkte angedeutet werden.

2.3.2 Reduktion auf ein lineares Gleichungssystem

Die linearisierte SÜG (13) ist eine homogene Integralgleichung – eine sogenannte *Fredholmsche Integralgleichung erster Art* [76, Bd. IV/1 S. 14]. Nun ist die Theorie der homogenen Integralgleichungen in mancher Hinsicht der linearen Algebra analog, wie allein schon anhand der diskretisierten Form der linearisierten SÜG (13) deutlich wird. Berücksichtigt man überdies, daß das Strahldichtespektrum nur an endlich vielen spektralen Stützstellen ν_i aufgenommen wird, $i = 1 \dots m$, so läßt sich die linearisierte SÜG in Matrixschreibweise darstellen als ¹¹

$$\mathbf{b} - \mathbf{b}^0 = \mathbf{A}_T^0 [\mathbf{f}_T - \mathbf{f}_T^0] + \mathbf{A}_{c_v}^0 [\mathbf{f}_{c_v} - \mathbf{f}_{c_v}^0] + \dots \quad (14)$$

¹¹Fette Großbuchstaben kennzeichnen im folgenden Matrizen, fette Kleinbuchstaben Vektoren.

Dabei bezeichnet der Vektor \mathbf{b} das an m diskreten spektralen Stützstellen gemessene Strahldichtespektrum. Die Vektoren $\mathbf{f}_T, \mathbf{f}_{c_v}$ usw. repräsentieren die diesem Spektrum zugrundeliegenden wahren Höhenprofile von Temperatur und Volumenmischungsverhältnissen relevanter Spurengase, welche im Zuge der Inversionsrechnungen aus dem Strahldichtespektrum berechnet werden sollen.¹² Die Vektoren $\mathbf{f}_T^0, \mathbf{f}_{c_v}^0, \dots$ kennzeichnen die Anfangsschätzwerte für die Temperatur- und Volumenmischungsverhältnishöhenprofile und \mathbf{b}^0 ist das aus dieser Anfangsschätzung über Vorwärtsrechnung ermittelte, synthetische Strahldichtespektrum. Die $(m \times n)$ Matrizen $\mathbf{A}_T^0, \mathbf{A}_{c_v}^0$ usw. schließlich – oder vielmehr deren Elemente – sind im wesentlichen gegeben durch die den spektralen Stützstellen i und Höhenschichten j entsprechenden partiellen Ableitungen (8) und (10) bzw. durch die zugehörigen diskretisierten Sensitivitätsfunktionen (siehe hierzu auch Gl. 13)),

$$A_{T\ i,j}^0 = \left. \frac{\partial L_\nu(h_B)}{\partial T_j} \right|_{T_j^0} = \Delta h_j(S_{\nu_i T})_j \Big|_{T_j^0}$$

sowie

$$A_{c_v\ i,j}^0 = \left. \frac{\partial L_\nu(h_B)}{\partial c_{v_j}} \right|_{c_{v_j}^0} = \Delta h_j(S_{\nu_i c_v})_j \Big|_{c_{v_j}^0}.$$

Durch geeignetes Zusammenfassen dieser sogenannten *Jakobimatrizen* $\mathbf{A}_T^0, \mathbf{A}_{c_v}^0$ usw. zu einer einzigen, großen Matrix \mathbf{A}^0 und entsprechendes Untereinanderschreiben der Höhenprofile $\mathbf{f}_T, \mathbf{f}_{c_v}, \dots$ sowie $\mathbf{f}_T^0, \mathbf{f}_{c_v}^0, \dots$ vereinfacht sich die Gl. (14) weiter zu¹³

$$\mathbf{b} - \mathbf{b}^0 = \underbrace{(\mathbf{A}_T^0, \mathbf{A}_{c_v}^0, \dots)}_{\mathbf{A}^0} \left[\underbrace{\begin{pmatrix} \mathbf{f}_T \\ \mathbf{f}_{c_v} \\ \vdots \end{pmatrix}}_{\mathbf{f}} - \underbrace{\begin{pmatrix} \mathbf{f}_T^0 \\ \mathbf{f}_{c_v}^0 \\ \vdots \end{pmatrix}}_{\mathbf{f}^0} \right], \quad (15)$$

so daß das lineare Gleichungssystem

$$\boxed{\mathbf{A}^0 \mathbf{f} = \mathbf{b} - \mathbf{b}^0 + \mathbf{A}^0 \mathbf{f}^0} \quad (16)$$

entsteht. Wichtig ist, nicht zu schreiben

$$\mathbf{A}^0[\mathbf{f} - \mathbf{f}^0] = \mathbf{b} - \mathbf{b}^0, \quad (17)$$

wobei die Gründe hierfür erst Ende des nächsten Kapitels deutlich werden.

¹²Wie bereits oben angedeutet repräsentieren die Ellipsen künftig die mögliche Erweiterung der linearisierten SÜG auf mehrere Gase.

¹³Zur Beschreibung zusammengefaßter Matrizen und Vektoren wird die Konvention, daß runde Klammern ausschließlich der Kennzeichnung funktionaler Abhängigkeiten vorbehalten sein sollten, durchbrochen.

Mit der Herleitung des linearen Gleichungssystem (16) ist das vorläufig gesteckte Ziel, die Reduktion der nichtlinearen, integralen SÜG (2) auf ein lineares Gleichungssystem, erreicht. Im Prinzip wurde das Problem der Inversion der SÜG dadurch reduziert auf das Lösen eines linearen Gleichungssystems, eine unter gewöhnlichen Umständen relativ einfach zu bewältigende Aufgabe. Leider erweist sich das Problem als schlecht gestellt im Sinne von Hadamard [30], so daß dieser vermeintlich einfachen Aufgabe im Anschluß gleich zwei ganze Kapitel gewidmet werden.

Ist die Lösung des obigen linearen Gleichungssystems erst einmal bekannt, ist die weitere Vorgehensweise klar. Anhand der ermittelten Höhenprofile von Temperatur und Volumenmischungsverhältnissen relevanter Spurengase können neue, verbesserte Jakobimatrizen \mathbf{A}_T^0 , $\mathbf{A}_{c_v}^0$, ... und ein neues, verbessertes synthetisches Spektrum \mathbf{b}^0 berechnet werden. Diese führen auf ein neues, verbessertes lineares Gleichungssystem, welches es erneut zu lösen gilt. Auf diese Weise kann iterativ die nichtlineare SÜG (2) invertiert werden, bzw. können aus gegebenen Strahldichtespektren Höhenprofile von Temperatur und Volumenmischungsverhältnisse interessierender Spurengase abgeleitet werden. – Voraussetzung ist natürlich, daß der Informationsgehalt des betrachteten Spektralbereichs und die Qualität der Messung entsprechend hoch sind. In der Regel werden für eine Inversion 2–3 solcher Iterationsschritte benötigt, wobei häufig nach dem ersten Iterationsschritt die Lösung die wahren Höhenprofile bereits sehr gut wiedergibt. Offen ist bisher noch die Frage nach einem geeigneten Abbruchkriterium für das beschriebene Iterationsverfahren, doch kann diese erst im Anschluß an die Ausführungen des folgenden Kapitels beantwortet werden.

Kapitel 3

Inversion der linearisierten Strahlungsübertragungsgleichung

3.1 Theorie der Inversion

Das zentrale Thema des folgenden Kapitels ist die Lösung des oben abgeleiteten linearen Gleichungssystems (16). Man spricht auch von einer *Inversion* oder von einem *inversen Problem*, da es im Prinzip die in Gl. (16) auftretende Matrix \mathbf{A}^0 zu invertieren gilt. Wie sich im weiteren Verlauf des Kapitels herausstellen wird, ist das Problem *schlecht gestellt* im Sinne von Hadamard [30] und bedarf deshalb einer besonders sorgfältigen numerischen Behandlung. Die Lösung muß *regularisiert*, d.h. speziellen, sie stabilisierenden Verfahren unterzogen werden. Wichtig hierbei ist zu beachten, daß die Auswirkungen einer solchen Regularisierung auf die Lösung stark davon abhängen, wie das der Lösung zugrundeliegende Gleichungssystem ursprünglich formuliert ist. Insbesondere bestehen zwischen den beiden linearen Gleichungssystemen (16) und (17), welche sich in mathematischer Hinsicht ja lediglich bzgl. ihrer Schreibweise unterscheiden, unter numerischen Gesichtspunkten beträchtliche Unterschiede. Dies wird in Abschn. 3.3 näher diskutiert. Statt dessen sei – zunächst noch repräsentativ für beide Gleichungssysteme – das lineare Gleichungssystem

$$\mathbf{A}\mathbf{f} = \mathbf{g} \tag{18}$$

betrachtet. Dabei steht die Matrix \mathbf{A} für die in den beiden Gleichungssystemen (16) und (17) auftretende $m \times n$ -Matrix \mathbf{A}^0 , \mathbf{f} für den n -komponentigen Vektor \mathbf{f} bzw. die Differenz $\mathbf{f} - \mathbf{f}^0$ sowie der m -komponentige Vektor \mathbf{g} für die rechte Seite $\mathbf{b} - \mathbf{b}^0 + \mathbf{A}^0 \mathbf{f}^0$ bzw. $\mathbf{b} - \mathbf{b}^0$. Vereinfachend sei weiterhin angenommen, daß zunächst lediglich ein einzelnes Höhenprofil aus den Meßdaten abgeleitet werden soll. Dabei ist es für die nachfolgenden Betrachtungen unerheblich, ob es sich um ein Temperaturhöhenprofil oder um ein Höhenprofil des Volumenmischungsverhältnisses handelt: wichtig ist zunächst nur, daß der Vektor \mathbf{f} nur einzelnes Höhenprofil repräsentiert. In diesem Fall treffen die nachfolgenden Ausführungen, welche den allgemeinen, interdisziplinären Wissensstand bzgl. der Inversion schlecht gestellter Probleme widerspiegeln, in unveränderter Form auch auf die bodengebundene, passive Millimeterwellenradiometrie zu. Sollen hingegen mehrere Höhenprofile gleichzeitig aus einem Spektrum ermittelt werden,

müssen einige zusätzliche Punkte beachtet werden, deren Diskussion in Kapitel 4 folgt.

3.1.1 Schlecht gestellte Probleme

Die Frage nach der Lösbarkeit eines linearen Gleichungssystems – oder allgemeiner – die Frage nach der Umkehrbarkeit einer Abbildung A von der Menge F der *Parameter* nach der Menge G der *Resultate*

$$A : F \rightarrow G$$

stellte bereits 1923 Hadamard [30]. Seiner Definition zufolge bezeichnet man ein solches Problem als gut gestellt (und damit lösbar), wenn

- i) für alle möglichen Resultate eine Lösung existiert (*Existenz*),
- ii) diese eindeutig ist (*Eindeutigkeit*)
- iii) und überdies stetig von den Resultaten abhängt (*Stabilität*).

Ist eine dieser Bedingungen nicht erfüllt, bezeichnet man das Problem als inkorrekt oder schlecht gestellt.¹ Dabei wird vor allen Dingen dem letzten Punkt, der Stabilität der Lösung gegenüber Meßfehlern in den Resultaten, das Hauptaugenmerk bei der Behandlung inkorrekt gestellter Fragestellungen gelten, denn die Nichtexistenz läßt sich durch eine Ausweitung des Lösungsbegriffs und die Nichteindeutigkeit durch das Vorschreiben von Normierungsbedingungen meist erzwingen.

Bezogen auf das obige, lineare Gleichungssystem bedeutet dies, daß das Problem (18) von vornherein schlecht gestellt ist, da die Matrix \mathbf{A} nicht quadratisch ist. Statt dessen ist \mathbf{A} eine $m \times n$ -Matrix, wobei $m > n$, d. h. das Gleichungssystem ist überbestimmt. Lediglich für den Fall einer quadratischen Matrix \mathbf{A} , die überdies nicht singulär sein darf (ihre Zeilen oder Spalten müssen linear unabhängig sein), ist das Problem gut gestellt. Das Gleichungssystem kann unter diesen Umständen sehr leicht gelöst werden zu

$$\mathbf{f} = \mathbf{A}^{-1}\mathbf{g}. \quad (19)$$

¹Die Bezeichnung „schlecht“ oder „inkorrekt“ gestellte Probleme suggeriert einen Makel des Falschen oder Unrichtigen. Hadamard war der Meinung, daß jede korrekte mathematische Formulierung einer sinnvollen physikalischen Fragestellung zu einer mathematischen Gleichung führen muß, welche den Bedingungen i) – iii) genügt. Daß dem nicht so ist, sondern daß sehr wohl ein korrektes Modell eines sinnvollen technischen Problems beschrieben werden kann, welches eine der Bedingungen i) – iii) verletzt, beweist nicht zuletzt die bodengebundene, passive Millimeterwellenradiometrie. Nicht das Modell ist schlecht abgefaßt oder das Problem falsch formuliert, sondern die sie beschreibende Gleichung erfordert eine besonders sorgfältige (numerische) Behandlung.

3.1.2 Methode der kleinsten Quadrate

Man behilft sich, indem man den Lösungsbegriff erweitert und statt des ursprünglichen, schlecht gestellten Problems (18) ein Ersatzproblem löst. An dieser Stelle sei auf die besonders lehrreichen Monographien von Louis [50], Menke [52] und Twomey [85] sowie auf die Kapitel 2, 15 und 18 in den „Numerical Recipes“ [61] verwiesen, in welchen die folgenden Schritte, die im Rahmen dieser Arbeit nur grob skizziert werden können, im Detail behandelt werden. Anstatt nach der im Falle eines überbestimmten Gleichungssystems nicht existierenden exakten Lösung, sucht man nach derjenigen Lösung, welche alle Gleichungen des überbestimmten Gleichungssystems simultan möglichst gut erfüllt. Man bestimmt die Lösung mit der kleinsten Norm (L_2), auch bekannt als die Methode der kleinsten Quadrate. Minimiert werden muß dann das Betragsquadrat der Abweichung ²

$$|\mathbf{A}\mathbf{f} - \mathbf{g}|^2 = \langle \mathbf{A}\mathbf{f} - \mathbf{g} | \mathbf{A}\mathbf{f} - \mathbf{g} \rangle .$$

Dieses Minimierungsproblem führt nach einigen Zwischenschritten (siehe Anhang B.1.1) auf das *normale* oder *verallgemeinerte* Gleichungssystem

$$\mathbf{A}^T \mathbf{A} \mathbf{f} = \mathbf{A}^T \mathbf{g}, \quad (20)$$

welches nun an Stelle der ursprünglichen Gleichung tritt. Im Gegensatz zu (18) kann das Gleichungssystem (20) auch für den Fall einer nicht quadratischen Matrix \mathbf{A} gelöst werden – vorausgesetzt das Problem verstößt nicht noch gegen Punkt iii) der Hadamardschen Bedingungen –, da das Produkt $\mathbf{A}^T \mathbf{A}$ quadratisch ist, so daß gilt

$$\mathbf{f} = \underbrace{[\mathbf{A}^T \mathbf{A}]^{-1} \mathbf{A}^T}_{\mathbf{A}^{-A}} \mathbf{g}. \quad (21)$$

Man spricht von Gl. (21) als der *bestapproximierten* oder *verallgemeinerten Lösung*, oft auch als *Moore–Penrose–Lösung* [55, S. 85] bezeichnet. ³ Die Matrix \mathbf{A}^{-A} betitelt man dazu passend als *verallgemeinerte Inverse* oder *Pseudoinverse*, wobei die Schreibweise mit einem hochgestellten $^{-A}$ die enge Verwandtschaft mit der nur im Falle einer quadratischen Matrix \mathbf{A} existierenden Inversen \mathbf{A}^{-1} symbolisieren soll. Penrose [60] zeigte 1955, daß für jede endliche Matrix \mathbf{A} , reell oder komplex, eine solche eindeutige, verallgemeinerte Inverse \mathbf{A}^{-A} existiert.

Charakteristisch für inverse Probleme ist nun, daß selbst die verallgemeinerte Lösung (21) nicht berechnet werden kann, da das Problem schlecht gestellt ist bzgl.

²Das hochgestellte T kennzeichnet im folgenden die Transponierte eines Vektors oder Matrix und für das Skalarprodukt wird die Diracsche Notation $\langle | \rangle$ verwendet.

³Verallgemeinert in dem Sinne, daß ein überbestimmtes Gleichungssystem keine eindeutige Lösung besitzt, jedoch durch das Vorschreiben einer Normierungsbedingung (L_2 -Norm) eindeutig gemacht wurde.

Punkt iii) der Hadamardschen Definition. Die Matrix $\mathbf{A}^T\mathbf{A}$ ist singulär bzw. – was dem unter numerischen Gesichtspunkten im Prinzip gleichkommt – nahezu singulär. (Der Rang der Matrix $\mathbf{A}^T\mathbf{A}$ ist kleiner als die Anzahl ihrer Zeilen bzw. Spalten.) Ursache dieser Nahezu-Singularität ist der geringe Informationsgehalt des zugrundeliegenden Gleichungssystems, d. h. sehr viele Gleichungen sind linear abhängig. Press et al. [61] sprechen diesbezüglich gar von einer besonderen Ironie der Mathematik, daß nämlich solche, auf Integralgleichungen zurückgehende Gleichungssysteme gleichzeitig überbestimmt (was die Anzahl der Gleichungen und die Anzahl der Unbekannten angeht) und unterbestimmt sind (was den Informationsgehalt des Gleichungssystems betrifft).

3.1.3 Singulärwertzerlegung und Regularisierung

Immer dann, wenn die herkömmlichen Verfahren, wie z. B. Gauß-Elimination oder LU-Zerlegung ⁴ [61], bei der Lösung eines linearen Gleichungssystems versagen, empfiehlt sich die Durchführung einer sogenannten Singulärwertzerlegung (SVD ⁵), um die Ursache der auftretenden Schwierigkeiten zu diagnostizieren und gegebenenfalls zu eliminieren. Unter diesem Verfahren versteht man die Berechnung der Eigenwerte σ_k^2 und Eigenvektoren \mathbf{v}_k der Matrix $\mathbf{A}^T\mathbf{A}$ (z. B. mit dem Jakobi- oder Householder-Verfahren [61]),

$$\mathbf{A}^T\mathbf{A}\mathbf{v}_k = \sigma_k^2\mathbf{v}_k,$$

und die anschließende Darstellung der verallgemeinerten Lösung (21) in der Basis der Eigenvektoren \mathbf{v}_k [6], [50, S. 23], [55, S. 85 ff.], [61, S. 59 ff.].

$$\begin{aligned} \mathbf{f} &= \sum_{k=1}^n \langle \mathbf{f} | \mathbf{v}_k \rangle \mathbf{v}_k \\ &= \sum_{k=1}^n \langle [\mathbf{A}^T\mathbf{A}]^{-1}\mathbf{A}^T\mathbf{g} | \mathbf{v}_k \rangle \mathbf{v}_k \\ &= \sum_{k=1}^n \sigma_k^{-2} \langle [\mathbf{A}^T\mathbf{A}]^{-1}\mathbf{A}^T\mathbf{g} | \mathbf{A}^T\mathbf{A}\mathbf{v}_k \rangle \mathbf{v}_k, \end{aligned}$$

so daß

$$\mathbf{f} = \sum_{k=1}^n \sigma_k^{-2} \langle \mathbf{g} | \mathbf{A}\mathbf{v}_k \rangle \mathbf{v}_k. \quad (22)$$

Anhand der Darstellung (22) werden die für Inversionsprobleme typischen Schwierigkeiten deutlich. Sind nämlich einige Eigenwerte σ_n^2 der Matrix $\mathbf{A}^T\mathbf{A}$ sehr klein und die

⁴Zerlegung der zu invertierenden Matrix $\mathbf{A}^T\mathbf{A}$ in das Produkt zweier Dreiecksmatrizen \mathbf{L} und \mathbf{U} , \mathbf{L} = lower triangular matrix, \mathbf{U} = upper triangular matrix.

⁵SVD = singular value decomposition [61]

Daten \mathbf{g} mit Meßfehlern behaftet (dabei genügen häufig bereits Fehler in der Größenordnung des numerischen Rauschens), so werden diese Fehler aufgrund der kleinen Eigenwerte enorm verstärkt und die verallgemeinerte Lösung unbrauchbar.

Um trotzdem zu einem vernünftigen Ergebnis zu gelangen, müssen alle diejenigen Beiträge zu (22), die aufgrund kleiner Eigenwerte eine extreme Verstärkung von Meßfehlern bewirken würden, unterdrückt werden. Dahinter verbirgt sich die Philosophie, daß es unter Umständen wesentlich günstiger sein kann von der Lösung etwas abzuschneiden, von dem man von vornherein weiß, daß es stark mit Fehlern behaftet ist, als blindlings über alle Beiträge aufzusummieren. Mathematisch kann dieses Unterdrücken erreicht werden durch den Einbau eines geeigneten Filters $F_\gamma(\sigma_n)$, welches alle Beiträge zu (22), deren Eigenwerte kleiner als ein noch zu wählender *Regularisierungsparameter* γ sind, eliminiert. Statt von einer Filterung spricht man auch von einer *Regularisierung* der verallgemeinerten Lösung, im folgenden jeweils gekennzeichnet durch tiefgestelltes γ ,

$$\mathbf{f}_\gamma = \sum_n \sigma_n^{-2} F_\gamma(\sigma_n) \langle \mathbf{g} | \mathbf{A} \mathbf{v}_n \rangle \mathbf{v}_n. \quad (23)$$

Gl. (23) läßt sich auch schreiben in der Form ⁶

$$\begin{aligned} \mathbf{f}_\gamma &= \sum_n \sigma_n^{-2} F_\gamma(\sigma_n) \mathbf{v}_n \langle \mathbf{A} \mathbf{v}_n | \mathbf{g} \rangle \\ &= \underbrace{\sum_n \sigma_n^{-2} F_\gamma(\sigma_n) \mathbf{v}_n \mathbf{v}_n^T \mathbf{A}^T}_{\mathbf{A}_\gamma^{-A}} \mathbf{g}. \end{aligned} \quad (24)$$

D. h. selbst bei einer Regularisierung der verallgemeinerten Lösung wird im Prinzip eine Art inverse Matrix berechnet, die *regularisierte verallgemeinerte Inverse* \mathbf{A}_γ^{-A} . Zur Abgrenzung gegenüber der bereits in Abschn. 3.1.2 eingeführten verallgemeinerten Inversen \mathbf{A}^{-A} wird sie im folgenden durch einen tiefgestellten Regularisierungsparameter γ gekennzeichnet, \mathbf{A}_γ^{-A} .

3.1.4 Wahl des Regularisierungsparameters

Die eigentliche Kunst der Regularisierung besteht nun in der Wahl des richtigen Regularisierungsparameters γ , welcher festlegt, wie stark gefiltert werden soll, bzw. ab welcher Schwelle zu kleinen Eigenwerten gehörende Beiträge abgeschnitten werden sollen. Je nach Art des gewählten Filters bzw. Regularisierungsverfahrens können die richtigen Werte für den Regularisierungsparameter unterschiedlich groß sein. D.h. ein bestimmtes γ , welches sich für ein spezielles Regularisierungsverfahren als günstig

⁶Ein erheblicher Nachteil dieser zweiten Darstellung ist, daß sie auf dem Rechner wesentlich mehr Speicherplatz benötigt als die Darstellung (23).

herausgestellt hat, muß nicht notwendigerweise auch für ein anderes Regularisierungsverfahren geeignet sein. Ein jedes Regularisierungsverfahren erfordert deshalb eine eigene Strategie, um den Regularisierungsparameter γ geeignet zu wählen.

Im folgenden sollen zunächst einige grundsätzliche Überlegungen zu dieser Problematik angestellt und anschließend zwei Strategien zur Wahl des Regularisierungsparameters, die a priori- und die a posteriori- Parameterwahl, vorgestellt werden. Weitere Strategien zur Wahl des Regularisierungsparameters finden sich in [1], [31] und [83, S. 804 ff].

Ausgangspunkt der nachfolgenden Überlegungen sind die gestörten Daten $\mathbf{g}_\varepsilon = \mathbf{g} + \mathbf{E}$, mit \mathbf{E} als dem Meßfehlervektor. Aus diesen gestörten Daten ergibt sich nach erfolgter Regularisierung mit dem Regularisierungsparameter γ aufgrund von (24) die Näherungslösung

$$\mathbf{f}_\gamma = \mathbf{A}_\gamma^{-A} \mathbf{g}_\varepsilon.$$

Der hierbei gemachte *Gesamt-* oder *Rekonstruktionsfehler*, d. h. die Abweichung der Näherungslösung \mathbf{f}_γ vom wahren Wert \mathbf{f} (siehe die Gleichungen (21) und (24)),

$$\mathbf{f}_\gamma - \mathbf{f} = \mathbf{A}_\gamma^{-A} \mathbf{g}_\varepsilon - \mathbf{A}^{-A} \mathbf{g} = \underbrace{\mathbf{A}_\gamma^{-A} \mathbf{E}}_{\text{Datenfehler}} + \underbrace{[\mathbf{A}_\gamma^{-A} - \mathbf{A}^{-A}] \mathbf{g}}_{\text{Regularisierungsfehler}}, \quad (25)$$

setzt sich aus zwei Anteilen zusammen [50, S. 10], vgl. auch [83, S. 804 Gl. (26)]. Der erste Teil ist der sogenannte *Datenfehler*⁷, da er in erster Linie auf Meßfehler in den Daten zurückzuführen ist; der zweite Teil ist der sogenannte *Regularisierungsfehler*, weil er sich als Folge der Regularisierung (des Abschneidens) ergibt. Qualitativ ergibt sich nun Abb. 5, welche kennzeichnend für alle inversen und schlecht gestellten Probleme ist.

Wird nicht regularisiert, $\gamma = 0$ (wird also lediglich die verallgemeinerte Lösung oder least squares solution (21) berechnet), so ist der Datenfehler aufgrund der stark fehlerbehafteten Beiträge zu kleinen Eigenwerten zunächst sehr groß. Regularisiert man infolgedessen stärker, nimmt der Datenfehler ab, doch wächst der Regularisierungsfehler – und zwar umso stärker, je größer der Regularisierungsparameter γ gewählt wird. Der Gesamt- oder Rekonstruktionsfehler kann deshalb nicht beliebig klein gemacht werden, sondern es ist lediglich möglich, ihn bezüglich γ zu optimieren – am naheliegendsten durch einen Vergleich der Näherungslösung mit der exakten Lösung.

In der Praxis ist die exakte Lösung meist nicht bekannt, so daß das optimale γ nicht ermittelt werden kann. Statt dessen erfolgt die Wahl des Regularisierungsparameters γ

⁷Die Bezeichnung Datenfehler ist insofern irreführend, als unter dem Datenfehler nicht der Meßfehler \mathbf{E} in den fehlerbehafteten Daten \mathbf{g}_ε zu verstehen ist, sondern der aus diesem Meßfehler in den Daten resultierende Anteil am Gesamtfehler der Lösung. Der Begriff Datenfehler greift also in der „Lösungsebene“ an.

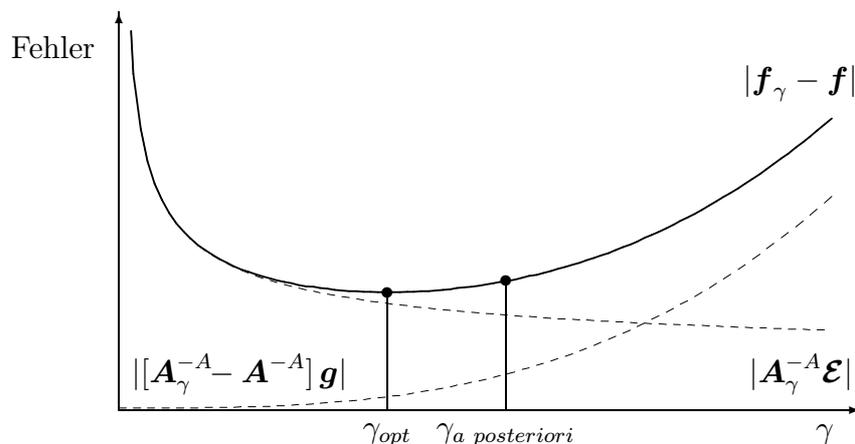


Abbildung 5: Regularisierungs- oder Abschneidefehler (gestrichelt, monoton steigend), Datenfehler (gestrichelt, monoton fallend) und Gesamt- oder Rekonstruktionsfehler (durchgezogen) in Abhängigkeit vom Regularisierungsparameter γ (siehe Gl. (25)).

entweder *a priori*, basierend auf Zusatzinformationen, die irgendwelchen Erfahrungswerten, früheren Messungen oder sonstigen Kenntnissen entstammen, oder *a posteriori* über *trial and error* [50, S. 54]. Die *a priori*-Strategie, nämlich einen Wert für den Regularisierungsparameter γ von vornherein fest vorzugeben, verfolgen die sogenannten statistischen Regularisierungsverfahren, wie z. B. das Rodgers-Verfahren [68] oder das linear statistische Verfahren, welches zusammen mit der *a priori*-Parameterwahl in Abschn. 3.2.3 besprochen wird.

Bei der *a posteriori*-Parameterwahl hingegen wird die Wahl abhängig gemacht vom Defekt zwischen den gemessenen Daten \mathbf{g}_ε und den aus der Näherungslösung \mathbf{f}_γ über Vorwärtsrechnung ermittelten Daten $\mathbf{g}_\gamma = \mathbf{A}\mathbf{f}_\gamma$,

$$|\mathbf{g}_\varepsilon - \mathbf{g}_\gamma| = |\mathbf{g} + \boldsymbol{\varepsilon} - \mathbf{g}_\gamma| = |\boldsymbol{\varepsilon} + \mathbf{A}[\mathbf{f} - \mathbf{f}_\gamma]|. \quad (26)$$

Dieses gilt für die abgeschnittene Singulärwertzerlegung (siehe Abschn. 3.2.1), für die Tikhonov-Phillips Regularisierung (siehe Abschn. 3.2.2) oder für die iterativen Regularisierungsverfahren [40]. Aus der Gl. (26) geht hervor, daß selbst im Idealfall, wenn die Näherungslösung \mathbf{f}_γ der exakten Lösung \mathbf{f} entspricht, der Defekt $|\mathbf{g}_\varepsilon - \mathbf{g}_\gamma|$ nicht identisch Null wird, sondern den Wert $\boldsymbol{\varepsilon} = |\boldsymbol{\varepsilon}|$ annimmt. Es erscheint daher naheliegend, eine Näherungslösung dann als optimal (im Rahmen der Meßgenauigkeit) zu erachten, sobald der Defekt in den Daten den Wert $\boldsymbol{\varepsilon}$ annimmt und nicht etwa zu versuchen – in dem Bestreben die Anpassung verbessern zu wollen – den Defekt weiter zu verringern. Eine Voraussetzung für die Umsetzung dieser Beobachtung in ein geeignetes Kriterium zur Wahl des Regularisierungsparameters ist allerdings, daß der Betrag $\boldsymbol{\varepsilon}$ des Meßfehlervektors zumindest ungefähr bekannt ist, wovon im folgenden

ausgangen werden soll.⁸

Die a posteriori-Strategie zur Wahl des Regularisierungsparameters stellt sich nun wie folgt dar [50, S. 68], [83, S. 805 Gl. (30)]: Ausgehend von einem hinreichend großen γ wird versucht, den Defekt (26) durch sukzessives Verkleinern des Regularisierungsparameters weiter zu verringern. Das Verfahren wird schließlich abgebrochen, sobald dieser die untere Schranke \mathcal{E} erreicht. Häufig erscheint es sinnvoll, das Verfahren bereits etwas früher abzubrechen, z. B. bei einem Wert $r\mathcal{E}$, wobei $r \geq 1$. (Dies ist aus Sicherheitsgründen dann erforderlich, wenn nicht erwartet werden kann, daß die Näherungslösung der exakten Lösung beliebig nahe kommt). Typische Werte für r liegen im Falle der bodengebundenen, passiven Millimeterwellenradiometrie im Bereich zwischen 1,0 und 1,5.

Als Anwender sollte man sich bei einer Verfolgung der a posteriori-Strategie zur Wahl des Regularisierungsparameters der Tatsache bewußt sein, daß sie keinesfalls das optimale γ liefert, sondern i. d. Regel etwas zu große Werte (angedeutet in Abb. 5). D. h. die Näherungslösung kann meist noch etwas verbessert werden (wenn häufig auch nur geringfügig). Andererseits ist man sich als Anwender bei einer Verwendung der a posteriori-Parameterwahl jedoch sicher, daß die Näherungslösung nur geringfügig mit Datenfehlern behaftet ist (der Datenfehler kann ja, wie aus Abb. 5 ersichtlich, für kleine Werte des Regularisierungsparameters γ nahezu beliebig groß werden). Dies ist vor allen Dingen dann von Vorteil, wenn der Inversionsalgorithmus operationell betrieben werden soll. Wird hingegen „von Hand“ ausgewertet, bleibt es letztlich dem Ermessen des Anwenders selbst überlassen, den Regularisierungsparameter γ nach eigenem Gutdünken noch etwas kleiner zu wählen, um die Lösung zu optimieren.

3.1.5 Erreichbare vertikale Auflösung im Profil

Wie Louis [50, S. 49-50] zeigt, entsprechen kleinen Eigenwerten bei schlecht gestellten Operatoren stark oszillierende Eigenvektoren \mathbf{v}_n , so daß sich der Meßfehler \mathcal{E} in einem starken Oszillieren der unregularisierten Lösung bemerkbar macht (siehe auch [40, S. 35]). Der Einbau eines Filters $F_\gamma(\sigma_n)$ wirkt daher wie ein Tiefpaß, welcher die hochfrequenten Anteile der unregularisierten Lösung abschneidet. Im Prinzip wird dieses erreicht, indem über benachbarte Elemente der unregularisierten Lösung lokal gemittelt oder verschmiert wird. Von daher kann man die Regularisierung auch

⁸Im wesentlichen hängt die Größe dieses Wertes ab von der Integrationszeit, über die das aktuelle Spektrum aufgenommen wurde. Geht man davon aus, daß es sich bei den Meßfehlern um rein statistische Fehler handelt (weißes Rauschen), so ist der Erwartungswert für den Meßfehler gleich 0. In diesem Falle aber ist der Betrag $\mathcal{E} = |\mathcal{E}|$ des Meßfehlervektors \mathcal{E} gleich der Summe der Standardabweichungen eines jeden Elements des Meßfehlervektors. Des weiteren gilt die Radiometerformel, nach der die Standardabweichung der Wurzel aus der Integrationszeit umgekehrt proportional ist [86, 88].

auffassen als eine Art Faltung der unregularisierten Lösung mit einer „Fenster“-funktion variabler Breite: man denke z. B. an eine Art Gauß- oder Sinc-Funktion (siehe [50, S. 86]). Letztlich führt diese Mittelung bzw. Faltung zu einer Reduktion der Anzahl der Unbekannten und damit zu einer Reduktion der Höhen-Auflösung.

Man hätte sich also gleich von vornherein – sofern dies technisch realisiert werden kann – auf eine etwas geringere (Höhen-)Auflösung beschränken können.⁹ Bei einer weniger feinen Diskretisierung gleich zu Beginn hätte man dann unter Umständen ganz auf eine Regularisierung verzichten können. Doch ist es in einem solchen Fall relativ schwierig, ein Maximum an Information aus den Daten zu extrahieren, da man im voraus leider nicht weiß, wie hoch der Informationsgehalt der Messungen nun tatsächlich ist. Statt dessen ist es nach Louis [50] „wesentlich günstiger, zunächst etwas feiner zu diskretisieren und anschließend die diskreten Probleme zu regularisieren“. Hierzu geben die im Abschn. 3.2 vorgestellten Regularisierungsverfahren das notwendige numerische Rüstzeug, so daß anhand weniger Steuer- bzw. Regularisierungsparameter eine vernünftige Inversion mit optimal angepaßter Auflösung durchgeführt werden kann.

Aufschluß darüber, inwieweit Strukturen in der exakten Lösung \mathbf{f} nach erfolgter Regularisierung schließlich aufgelöst werden können, gibt die *Auflösungsmatrix* oder *Averaging Kernel Matrix* \mathbf{R} [52, S. 64], [71]. Man erhält sie, indem man das Vorwärtsproblem (18) in die regularisierte Lösung \mathbf{f}_γ (23) bzw. (24) einsetzt,

$$\mathbf{f}_\gamma = \mathbf{A}_\gamma^{-A} \mathbf{g} = \mathbf{A}_\gamma^{-A} \mathbf{A} \mathbf{f} \stackrel{!}{=} \mathbf{R} \mathbf{f}.$$

Die Auflösungsmatrix \mathbf{R} ist von der regularisierten verallgemeinerten Inversen \mathbf{A}_γ^{-A} sowie der Matrix \mathbf{A} abhängig.¹⁰ Eine Beurteilung der Güte einer Inversion ist mit ihrer Hilfe also möglich, ohne eine Messung durchführen oder mit synthetischen Daten arbeiten zu müssen.

Im Idealfall, wenn die Näherungslösung \mathbf{f}_γ der exakten Lösung entspricht, sollte sich für \mathbf{R} die Einheitsmatrix \mathbf{I} ergeben. Ansonsten zeigt die regularisierte Lösung \mathbf{f}_γ ein mehr oder weniger stark geglättetes Abbild der exakten Lösung und die Matrix \mathbf{R} ist je nach erreichter Auflösung eine mehr oder weniger breite Bandmatrix (siehe Abb. 6). Das Band muß dabei nicht überall gleich breit sein, sondern die Auflösung kann sich auch mit der Höhe ändern (vgl. die Abbildungen 12 und 16 in Abschn. 5.1).

Die Verschmierung der Höhenschichten der exakten Lösung ist durch die Zeilenvektoren der Matrix \mathbf{R} gegeben. Mit Hilfe dieser „Verschmierungsfunktionen“ läßt

⁹Z. B. ist bei den Horizontsondierungsverfahren die Zahl und Position der Höhenstützstellen und damit die Zahl der Unbekannten im allgemeinen bereits durch die Beobachtungsgeometrie festgelegt [12].

¹⁰Dies gilt nur für die linearen Regularisierungsverfahren (siehe Abschn. 3.2). Bei den nichtlinearen, iterativen Regularisierungsverfahren, wie z. B. beim Chahineschen Verfahren oder beim Verfahren der konjugierten Gradienten, kann die regularisierte verallgemeinerte Inverse nicht explizit angegeben werden, so daß die folgenden Angaben über die erreichbare Auflösung und den Fehler der regularisierten Lösung bei den nichtlinearen Regularisierungsverfahren nicht möglich sind.

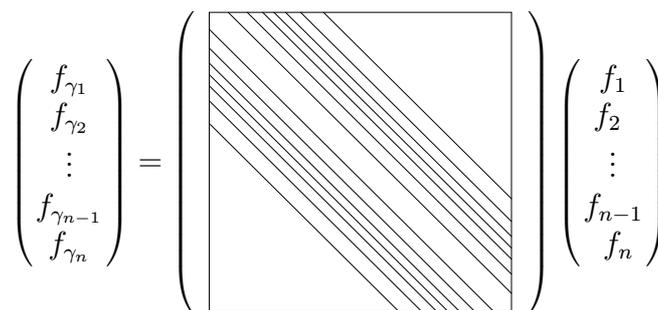


Abbildung 6: Die Zeilen der Auflösungsmatrix \mathbf{R} geben Aufschluß darüber, inwieweit in der exakten Lösung eventuell vorhandene Strukturen in der Näherungslösung noch aufgelöst werden können. Die Matrix selbst ist dargestellt als Höhenlinienplot, wobei das Maximum entlang der Diagonalen verläuft.

sich die mögliche (Höhen-)Auflösung der Inversion abschätzen, etwa durch die Halbwertsbreite einer solchen Verschmierungsfunktion (mehr hierzu in [40, S. 37 ff.]).

3.1.6 Abschätzung des Fehlers

Leider liefern die Regularisierungsverfahren – mit Ausnahme der statistischen Verfahren – keine Aussage über den Gesamt- oder Rekonstruktionsfehler der regularisierten Lösung; und selbst deren Angaben über den Gesamtfehler sind meist mit einer gewissen Vorsicht zur Kenntnis zu nehmen. Detaillierte Angaben sind hingegen bei allen linearen Regularisierungsverfahren möglich über die Größe des Datenfehlers $\mathbf{A}_\gamma^{-A} \boldsymbol{\varepsilon}$, welcher – wie aus Gl. (25) hervorgeht – entscheidend zum Gesamtfehler mit beiträgt. Eine Voraussetzung hierfür ist allerdings, daß das statistische Verhalten des Meßfehlervektors in Form seiner Kovarianzmatrix $\mathbf{S}_\boldsymbol{\varepsilon}$ bekannt ist, wenn man auch den Meßfehlervektor $\boldsymbol{\varepsilon}$ selbst nicht kennt.¹¹ Die Kovarianzmatrix stellt die Erweiterung der vom eindimensionalen Fall her bekannten Varianz auf ein Ensemble von Zufallsvektoren dar, wobei die Diagonalelemente der Kovarianzmatrix gegeben sind durch die Varianzen der Elemente der Zufallsvektoren. Die Matrixelemente außerhalb der Diagonalen hingegen sind ein Maß für die Korrelation der Elemente der Zufallsvektoren untereinander [7, S. 671] [54, S. 82 ff.]. Sind die Elemente der Zufallsvektoren nicht korreliert, ist die Kovarianzmatrix eine reine Diagonalmatrix.¹²

Wie bereits in Abschn. 3.1.3 Gl. (24) gesehen, ergibt sich die regularisierte Lösung

¹¹In Abschn. 3.2.3 über das linear statistische Verfahren werden die Eigenschaften und die Definition der Kovarianzmatrix ausführlicher behandelt, so daß hier lediglich die zum Verständnis notwendigen Erläuterungen gegeben werden.

¹²Diese Annahme ist z. B. für die Meßfehlerkovarianzmatrix $\mathbf{S}_\boldsymbol{\varepsilon}$ im Falle der bodengebundenen, passiven Millimeterwellenradiometrie weitgehend gerechtfertigt, da es sich bei den Meßfehlern im wesentlichen um statistische Fehler, also „Gaußsches Rauschen“ handelt.

\mathbf{f}_γ durch Multiplikation der regularisierten verallgemeinerten Inversen \mathbf{A}_γ^{-A} mit der rechten Seite \mathbf{g}_ε ,

$$\mathbf{f}_\gamma = \mathbf{A}_\gamma^{-A} \mathbf{g}_\varepsilon.$$

Dies impliziert aufgrund der Rechenregeln für Kovarianzmatrizen (siehe z. B. [54, S. 82 ff.]) für die Kovarianzmatrix \mathbf{S}_γ der regularisierten Lösung \mathbf{f}_γ unmittelbar

$$\mathbf{S}_\gamma = \mathbf{A}_\gamma^{-A} \mathbf{S}_\varepsilon [\mathbf{A}_\gamma^{-A}]^T. \quad (27)$$

In der Physik gibt man für den (statistischen) Fehler einer Meßgröße üblicherweise deren Standardabweichung an, welche definiert ist als die Wurzel aus der Varianz. Ist daher die Kovarianzmatrix \mathbf{S}_γ bekannt, deren Diagonalelemente ja gegeben sind durch die Varianzen der Elemente der regularisierten Lösung \mathbf{f}_γ , kann man mit Hilfe von (27) den (Daten-)Fehler der regularisierten Lösung leicht abschätzen.

Der Vollständigkeit halber sei an dieser Stelle auch gleich der entsprechende Ausdruck für die statistischen Regularisierungsverfahren mit angegeben, ohne ihn hier jedoch herleiten zu wollen (Näheres siehe Abschn. 3.2.3). Wie bereits eingangs erwähnt, erlauben die statistischen Verfahren eine Aussage über den Gesamtfehler der regularisierten Lösung, weshalb im Vergleich zu (27) ein zusätzlicher, den Abschneide- oder Regularisierungsfehler charakterisierender Term, abhängig von der Auflösungsmatrix \mathbf{R} und einer weiteren Kovarianzmatrix, der a priori-Kovarianzmatrix \mathbf{S}_f , auftritt,

$$\mathbf{S}_\gamma = \mathbf{A}_\gamma^{-A} \mathbf{S}_\varepsilon [\mathbf{A}_\gamma^{-A}]^T + [\mathbf{I} - \mathbf{R}] \mathbf{S}_f [\mathbf{I} - \mathbf{R}]^T. \quad (28)$$

Die a priori-Kovarianzmatrix \mathbf{S}_f enthält a priori-Informationen über die Lösung, z. B. über deren zu erwartende Form und muß bereits vor der eigentlichen Inversion bekannt sein. D. h. im wesentlichen erhält man an zusätzlicher Information über den Regularisierungsfehler bei den statistischen Verfahren wieder das zurück, was von vornherein an a priori-Information über die Lösung in den Algorithmus hineingesteckt wurde.

3.1.7 Zusammenfassung

Bevor im folgenden Unterkapitel auf einige spezielle Regularisierungsverfahren näher eingegangen wird, sei im folgenden Abschnitt ein kurzer Überblick über das in diesem Kapitel bisher Gesagte gegeben. Zusammenfassend läßt sich feststellen, daß je nach Beschaffenheit des linearen Gleichungssystems (18) jeweils verschiedene Wege zu dessen Lösung eingeschlagen werden müssen:

1. Ist das Problem gut gestellt, d. h. ist das Gleichungssystem von vornherein quadratisch und sind alle Gleichungen voneinander linear unabhängig, so kann die inverse Matrix \mathbf{A}^{-1} berechnet werden, so daß

$$\mathbf{f} = \mathbf{A}^{-1} \mathbf{g}_\varepsilon.$$

Dabei wird die volle (Höhen-)Auflösung erreicht, d.h. die Auflösungsmatrix entspricht der Einheitsmatrix, $\mathbf{R} = \mathbf{I}$.

2. Ist das Gleichungssystem „formal“ schlecht gestellt, was im vorliegenden Fall soviel heißt, wie das Gleichungssystem ist überbestimmt, so muß statt dessen die verallgemeinerte Inverse \mathbf{A}^{-A} aus Gl. (21) berechnet werden. Als Lösung ergibt sich die verallgemeinerte oder bestapproximierte Lösung

$$\mathbf{f} = \mathbf{A}^{-A} \mathbf{g}_\varepsilon.$$

Auch hier wird die volle Auflösung erreicht, $\mathbf{R} = \mathbf{I}$.

3. Ist das Gleichungssystem hingegen überbestimmt und überdies schlecht gestellt bzgl. der Stabilität, sind also sehr viele Gleichungen des überbestimmten Gleichungssystems voneinander linear abhängig bzw. nahezu linear abhängig (dies ist für die bodengebundene, passive Millimeterwellenradiometrie praktisch immer der Fall), muß regularisiert werden. Zu berechnen ist dann die regularisierte verallgemeinerte Inverse \mathbf{A}_γ^{-A} (siehe die Gln. (23) und (24) oder die im folgenden Abschnitt besprochenen Regularisierungsverfahren) und als Näherungslösung ergibt sich ein geglättetes bzw. mehr oder weniger verschmiertes Abbild der exakten Lösung,

$$\mathbf{f}_\gamma = \mathbf{A}_\gamma^{-A} \mathbf{g}_\varepsilon.$$

Leider ist die Regularisierung der verallgemeinerten Lösung nur möglich auf Kosten der erreichbaren (Höhen-)Auflösung, so daß nicht mehr – wie bisher – die Auflösungsmatrix der Einheitsmatrix entspricht, d. h. $\mathbf{R} \neq \mathbf{I}$. Statt dessen ist die Auflösungsmatrix \mathbf{R} eine mehr oder weniger breite Bandmatrix, wobei die Breite des Bandes angibt, inwieweit Strukturen der exakten Lösung noch aufgelöst werden können.

3.2 Regularisierungsverfahren

In der Literatur sind eine Vielzahl von Regularisierungsverfahren vorgeschlagen worden. Eine Übersicht über die verschiedenen Regularisierungsverfahren von der mathematischen Seite geben z. B. Baumeister [3], Louis [50], Natterer [55] sowie Strand [83]. Eher anwendungsorientiert sind dagegen die Ausführungen in Ishimaru [34], Menke [52], Rodgers [68], Twomey [85], Ulaby [87], Westwater [93] sowie in den Numerical Recipes [61].

In einer früheren Arbeit des Autors [40] wurden bereits mehrere Regularisierungsverfahren im Hinblick auf ihre Eignung zur Lösung von inversen Problemen aus dem Bereich der Fernerkundung mit Hilfe der bodengebundenen, passiven Millimeterradiometrie getestet. Neben der Güte der Inversionsergebnisse waren sowohl der Umfang der für eine Anwendung der Verfahren benötigten a priori-Information als auch die

Effizienz der Verfahren Gegenstand der Untersuchung. Verglichen wurden neben einigen direkten (linearen) Verfahren etliche iterative Regularisierungsverfahren (sowohl lineare als auch nichtlineare).

Aus dem Bereich der direkten Verfahren kristallisierten sich dabei insbesondere die *abgeschnittene Singulärwertzerlegung*, die *Tikhonov–Phillips–Regularisierung* sowie die *statistischen Verfahren* als geeignet heraus. Das ebenfalls weit verbreitete *Backus–Gilbert–Verfahren* [2, 16, 49, 92] hingegen erwies sich als den anderen direkten Verfahren gegenüber deutlich unterlegen, sowohl was die Effizienz des Verfahrens als auch die Qualität der Inversionsergebnisse betrifft. Aus der Klasse der linearen iterativen Regularisierungsverfahren wurden die *Landweber–Iteration* [41, 50, 83, 85] sowie eine Erweiterung dieses Verfahrens nach Strand [83] untersucht. Es zeigte sich, daß die Landweber–Iteration äußerst ineffizient arbeitete. Die Erweiterung des Verfahrens nach Strand hingegen lieferte gute Inversionsergebnisse bei gleichzeitig hoher Effizienz des Algorithmus. Allerdings war das Verfahren durch die Modifikation von Strand der Tikhonov–Phillips–Regularisierung so ähnlich geworden, daß es als Spezialfall der Tikhonov–Phillips–Regularisierung angesehen werden konnte und deshalb im folgenden nicht weiter betrachtet wird. Aus der Gruppe der nichtlinearen iterativen Verfahren wurden das *Verfahren der konjugierten Gradienten* [50, 61], das *Chahinesche Verfahren* [8, 9, 29] sowie eine eigene Modifikation des Chahineschen Verfahrens betrachtet. Es zeigte sich, daß das Verfahren der konjugierten Gradienten von allen betrachteten Regularisierungsverfahren am effizientesten arbeitete bei gleichzeitig hoher Qualität der erzielten Inversionsergebnisse. Die beiden Chahineschen Verfahren hingegen erwiesen sich zwar als ähnlich effizient, doch war die Qualität der Inversionsergebnisse in beiden Fällen allenfalls befriedigend. Dies ist darauf zurückzuführen, daß die Chahineschen Verfahren ursprünglich für die Nadirsondierung konzipiert wurden und deshalb nicht unmittelbar auch auf die bodengebundenen Messungen angewendet werden können. Statt dessen muß das ursprüngliche Problem zunächst in ein anderes, für eine Anwendung der Chahineschen Verfahren geeignetes Problem überführt werden [40, S. 20], [65].¹³ Leider ist dieses Vorgehen nur möglich auf Kosten einer erheblichen Verschlechterung des Signal–zu–Rausch–Verhältnisses, was sich wiederum negativ auf die Qualität der Inversionsergebnisse auswirkt.

¹³Die wichtigste Voraussetzung für eine Anwendung der Chahineschen Verfahren ist die Diagonaldominanz der Matrix \mathbf{A} . Bei der Nadirsondierung der Atmosphäre vom Satelliten aus ist diese Bedingung automatisch erfüllt, da die Sensitivitätsfunktionen aufgrund der Beschaffenheit der Atmosphäre von vornherein ein solches Maximum aufweisen. Dies gilt nicht für die bodengebundenen Messungen. Allerdings besteht hier die Möglichkeit, durch paarweise Subtraktion benachbarter Zeilen der Matrix \mathbf{A} und benachbarter Elemente des Datenvektors \mathbf{g} , also, $A'_{ij} = A_{ij} - A_{i+1j}$ sowie $g'_i = g_i - g_{i+1}$, das Problem in ein diagonaldominantes, „gestrichenes“ Gleichungssystem

$$\mathbf{g}' = \mathbf{A}' \mathbf{f}$$

zu überführen. Auf dieses können die Chahineschen Verfahren nun angewendet werden [65]. Allerdings ist der Informationsgehalt des neuen Gleichungssystems geringer als der des ursprünglichen, da

Von den für die bodengebundene passive Millimeterwellenradiometrie prinzipiell geeigneten Regularisierungsverfahren werden im folgenden nur die abgeschnittene Singulärwertzerlegung, das Tikhonov–Phillips–Verfahren und als Vertreter der statistischen Regularisierungsverfahren das linear statistische Verfahren näher betrachtet. Nicht eingegangen wird hingegen auf das Verfahren der konjugierten Gradienten,¹⁴ da es als nichtlineares Regularisierungsverfahren keine solch umfassenden Aussagen über die Auflösung oder den Fehler der regularisierten Lösung erlaubt, wie dies in Abschn. 3.1.5 beschrieben wurde und im Prinzip für alle linearen Verfahren möglich ist [40].

3.2.1 Abgeschnittene Singulärwertzerlegung

Das aufgrund der Überlegungen in Abschn. 3.1.3 über die Regularisierung schlecht gestellter Probleme sicherlich naheliegendste Regularisierungsverfahren ist die abgeschnittene Singulärwertzerlegung, (siehe [3, S. 49], [50, S. 79], [53, 83]). Dabei werden die in Gl. (23) gegenüber Datenfehlern stark empfindlichen Beiträge durch das spezielle Filter

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

einfach abgeschnitten, so daß für die regularisierte Lösung folgt¹⁵

$$\begin{aligned} \mathbf{f}_\gamma &= \sum_{\sigma_n > \gamma} \sigma_n^{-2} \langle \mathbf{g}_\varepsilon | \mathbf{A} \mathbf{v}_n \rangle \mathbf{v}_n \\ &= \underbrace{\sum_{\sigma_n > \gamma} \sigma_n^{-2} \mathbf{v}_n \mathbf{v}_n^T \mathbf{A}^T}_{\mathbf{A}_\gamma^{-A}} \mathbf{g}_\varepsilon. \end{aligned} \quad (29)$$

Die Filterwirkung dieses Verfahrens ist in Abb. 7 skizziert. Bevor die Gl. (29) jedoch zur Berechnung der regularisierten Lösung verwendet werden kann, müssen noch die durch die Subtraktion benachbarter Elemente des Datenvektors \mathbf{g} das Signal–zu–Rausch–Verhältnis erheblich verschlechtert wurde.

¹⁴Zur regularisierenden Wirkung des gemeinhin als Minimierungsverfahren bekannten Verfahrens der konjugierten Gradienten nur soviel: Als iteratives Verfahren benötigt das Verfahren eine geeignete Anfangsschätzung, von der es sich im Laufe der Iteration immer weiter entfernt und hin zur wahren Lösung iteriert. Normalerweise hat die Anfangsschätzung daher nur seinen sehr geringen Einfluß auf das Endergebnis. Ist das Problem hingegen schlecht gestellt im Sinne von Hadamard, erweist es sich als vorteilhaft, nicht allzu tief zu iterieren, um der Anfangsschätzung (a priori–Information) einen etwas stärkeren Einfluß auf das Inversionsergebnis einzuräumen. Das Inversionsergebnis erbt auf diese Weise quasi einen Teil der Eigenschaften der Anfangsschätzung (z. B. deren Stabilität) und wird dadurch regularisiert. Die Rolle des Regularisierungsparameters γ übernimmt bei diesem Verfahren also die Iterationstiefe.

¹⁵Die obere Gleichung ist die numerisch günstigere Alternative, wenngleich nur die zweite Schreibweise die zur Berechnung der Auflösungsmatrix \mathbf{R} benötigte regularisierte verallgemeinerte Inverse \mathbf{A}_γ^{-A} liefert, siehe Abschn. 3.1.5.

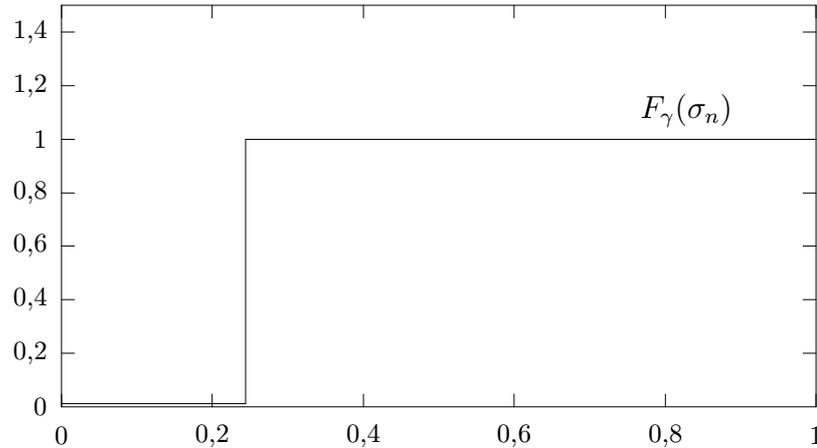


Abbildung 7: Filterwirkung der abgeschnittenen Singulärwertzerlegung für einen Regularisierungsparameter $\gamma = 0,25$.

Eigenwerte σ_n^2 der Matrix $\mathbf{A}^T\mathbf{A}$ samt den zugehörigen Eigenvektoren \mathbf{v}_n berechnet werden (z. B. mit dem Jacobi- oder dem Householder-Verfahren [17, 61]). Die abgeschnittene Singulärwertzerlegung ist damit in numerischer Hinsicht ein ziemlich aufwendiges Verfahren. Andererseits besteht sie durch die Tatsache, daß sie praktisch keine a priori Information benötigt, sofern die Parameterwahl a posteriori erfolgt. Die abgeschnittene Singulärwertzerlegung ist also insbesondere dann sinnvoll, wenn keine oder nur sehr wenig a priori Information zur Verfügung steht. Überdies kann die in Abschn. 3.1.4 Gl. (26) beschriebene a posteriori Parameterwahl bei der abgeschnittenen Singulärwertzerlegung besonders effizient durchgeführt werden [50, S. 82]. Da

$$\begin{aligned}
 |\mathbf{g}_\varepsilon - \mathbf{g}_\gamma|^2 &\leq |\mathbf{g}_\varepsilon|^2 - |\mathbf{g}_\gamma|^2 = |\mathbf{g}_\varepsilon|^2 - |\mathbf{A}\mathbf{f}_\gamma|^2 \\
 &= |\mathbf{g}_\varepsilon|^2 - \left| \sum_{\sigma_n > \gamma} \sigma_n^{-2} \langle \mathbf{g}_\varepsilon | \mathbf{A}\mathbf{v}_n \rangle \mathbf{A}\mathbf{v}_n \right|^2 \\
 &= |\mathbf{g}_\varepsilon|^2 - \sum_{\sigma_n > \gamma} \sigma_n^{-2} \langle \mathbf{g}_\varepsilon | \mathbf{A}\mathbf{v}_n \rangle^2, \quad (30)
 \end{aligned}$$

(hierbei wurde von der Orthogonalität der Eigenvektoren \mathbf{v}_n Gebrauch gemacht), kann der Defekt (26) zwischen den Meßdaten \mathbf{g}_ε und den aus der Näherungslösung über Vorwärtsrechnung ermittelten Daten \mathbf{g}_γ unter minimalem Zusatzaufwand berechnet werden. Man beginnt mit einem hinreichend großen Regularisierungsparameter, $\gamma \approx \infty$, und subtrahiert vom Betragsquadrat $|\mathbf{g}_\varepsilon|^2$ so lange Terme $\sigma_n^{-2} \langle \mathbf{g}_\varepsilon | \mathbf{A}\mathbf{v}_n \rangle^2$, bis das Betragsquadrat des Defekts (30) die untere Schranke $r^2\mathcal{E}^2$ erreicht. Da sowohl die Eigenwerte σ_n^2 wie auch die Skalarprodukte $\langle \mathbf{g}_\varepsilon | \mathbf{A}\mathbf{v}_n \rangle$ bei der Berechnung der Lösung (29) automatisch anfallen, ist der Zusatzaufwand entsprechend gering.

Ein weiterer Vorteil der abgeschnittenen Singulärwertzerlegung ist, daß aufgrund der Kenntnis der Eigenwerte die Schlechtgestellttheit des Problems klassifiziert werden kann. Ordnet man nämlich die Eigenwerte der Matrix $\mathbf{A}^T\mathbf{A}$ entsprechend ihrer Größe, $\sigma_1^2 > \sigma_2^2 > \dots$, so können anhand des Fallens der Eigenwerte schlecht gestellte Probleme unterschieden werden [50]. Man differenziert hierbei zwischen Problemen, die *schlecht gestellt* sind von der Ordnung α , wenn $\sigma_n^2 = \mathcal{O}(n^{-\alpha})$ mit $\alpha > 0$, und *exponentiell schlecht gestellten* Problemen, sofern $|\ln \sigma_n^2| \geq cn^\rho$ und $\rho > 0$. Bei der bodengebundenen, passiven Millimeterwellenradiometrie handelt es sich insofern um ein exponentiell schlecht gestelltes Problem, wie anhand der in [40, S. 30 ff.] durchgeführten abgeschnittenen Singulärwertzerlegung deutlich wird.

Vom mathematischen Standpunkt aus betrachtet ist die abgeschnittene Singulärwertzerlegung sicherlich eines der „saubersten“ Regularisierungsverfahren, da im Verlaufe des Verfahrens lediglich die stark fehlerbehafteten Beiträge abgeschnitten werden, andererseits jedoch keine (unter Umständen aufgrund falscher Annahmen ebenso falsche) a priori Information in die Lösung mit einfließt.

3.2.2 Tikhonov–Phillips–Regularisierung

Die Tikhonov–Phillips–Regularisierung [3, S. 33], [50, 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* [34, S. 504], als *damped least squares solution* [52, S. 52], als *regularization method* [61, S. 808], [84, S. 45], als *Twomey–Tikhonov method* [68] oder als *linear constraint solution* [85, S. 122]. Im Gegensatz zur abgeschnittenen Singulärwertzerlegung wird bei diesem Verfahren von der verallgemeinerten Lösung (21) nicht etwas abgeschnitten, sondern zur Lösung etwas hinzugefügt. Das Verfahren benötigt von daher deutlich mehr a priori–Information als die abgeschnittene Singulärwertzerlegung, bei der ja allenfalls die Wahl des Regularisierungsparameters γ a priori erfolgte.

Ausgangspunkt der Tikhonov–Phillips–Regularisierung ist die Beobachtung, daß die durch den Kern \mathbf{A} und die Inhomogenität \mathbf{g}_ε gegebene Information allein nicht ausreicht, um eine Lösung des (formal überbestimmten) linearen Gleichungssystems

$$\mathbf{A}\mathbf{f} = \mathbf{g}_\varepsilon \quad (31)$$

zu bestimmen. Deshalb sollen Zusatzinformationen über die Lösung in Form eines weiteren linearen Gleichungssystems,

$$\mathbf{B}\mathbf{f} = \mathbf{r}, \quad (32)$$

eingeführt werden. Dies können z. B. Kenntnisse über das Verhalten der Lösung am Rand sein. Möglich wäre es auch, bestimmte Nebenbedingungen an die Lösung zu stellen. Z. B. die Forderung, die gesuchte Lösung solle möglichst glatt sein, oder

die Elemente des Lösungsvektors sollten betragsmäßig möglichst klein sein. Letztere könnte realisiert werden durch die spezielle Wahl $\mathbf{B} = \mathbf{I}$ und $\mathbf{r} = 0$.

Gesucht wird wieder – wie in Abschn. 3.1.2 – die bestapproximierte Lösung des Gleichungssystems (31) (der Grund hierfür ist dessen formale Überbestimmtheit), doch diesmal unter einer Restriktion, nämlich der Zusatzbedingung (32). Es handelt sich dabei um ein schwierig zu lösendes restringiertes Minimierungsproblem. Mit Hilfe eines *Lagrangeschen Multiplikators*, γ^2 , kann dieses jedoch überführt werden in ein nicht restringiertes für

$$J_\gamma(\mathbf{f}_\gamma) := |\mathbf{A}\mathbf{f}_\gamma - \mathbf{g}_\varepsilon|^2 + \gamma^2 |\mathbf{B}\mathbf{f}_\gamma - \mathbf{r}|^2. \quad (33)$$

Das Funktional $J_\gamma(\mathbf{f}_\gamma)$ bezeichnet man hierbei als *Tikhonov–Phillips–Funktional* und $\gamma^2 |\mathbf{B}\mathbf{f}_\gamma - \mathbf{r}|^2$ als *Strafterm* [50, S. 87]. Wie anhand von Gl. (33) deutlich wird, geben große Werte von γ dem Strafterm mehr Gewicht, die Lösungen werden immer „glatter“. ¹⁶ Bei kleinen Werten von γ hat statt dessen der Defekt $|\mathbf{A}\mathbf{f}_\gamma - \mathbf{g}_\varepsilon|^2$ den größeren Einfluß auf die Lösung \mathbf{f}_γ . Sind die Daten nur wenig gestört, dann wird man immer ein kleines γ wählen. Hat man hingegen relativ ungenaue Daten vorliegen, wird man den Einfluß der Zusatzinformation durch die Wahl eines größeren γ erhöhen.

Aus dem Minimierungsproblem (33) folgt nach einigen Zwischenschritten (siehe Anhang B.1.2) das regularisierte verallgemeinerte Gleichungssystem

$$[\mathbf{A}^T\mathbf{A} + \gamma^2\mathbf{B}^T\mathbf{B}]\mathbf{f}_\gamma - \mathbf{A}^T\mathbf{g}_\varepsilon - \gamma^2\mathbf{B}^T\mathbf{r} = 0,$$

welches schließlich auf die regularisierte Lösung

$$\mathbf{f}_\gamma = \underbrace{[\mathbf{A}^T\mathbf{A} + \gamma^2\mathbf{B}^T\mathbf{B}]^{-1}}_{\mathbf{A}_\gamma^{-A}} [\mathbf{A}^T\mathbf{g}_\varepsilon + \gamma^2\mathbf{B}^T\mathbf{r}] \quad (34)$$

führt. ¹⁷ Diese kann unmittelbar für die numerische Berechnung verwendet werden.

Etwas verwundern mag zunächst die Tatsache, daß das erklärte Ziel aller Regularisierungsverfahren, die Stabilisierung der verallgemeinerten Lösung, bei der abgeschnittenen Singulärwertzerlegung einerseits erreicht wird durch ein Abschneiden der gegenüber Meßfehlern stark empfindlichen Beiträge zur verallgemeinerten Lösung (vgl. Gl. (29) im vorigen Abschnitt), während bei der Tikhonov–Phillips–Regularisierung andererseits zur Lösung etwas hinzugefügt wird – nämlich die vorhandene a priori–Information. Dieser nur scheinbare Widerspruch läßt sich jedoch beseitigen unter der einschränkenden Annahme, daß $\mathbf{A}^T\mathbf{A}$ und $\mathbf{B}^T\mathbf{B}$ ein gemeinsames System von Eigenvektoren \mathbf{v}_n besitzen und $\mathbf{r} = 0$ gilt. In diesem Fall kann die Tikhonov–Phillips–Regularisierung ebenfalls dargestellt werden als Filterung der verallgemeinerten Inversen (siehe Anhang B.1.3) mit einem Filter

¹⁶Sofern man die Glattheit als Nebenbedingung an die Lösung stellt.

¹⁷Die Matrix \mathbf{A}_γ^{-A} ist die regularisierte verallgemeinerte Inverse der Tikhonov–Phillips–Regularisierung, siehe Abschn. 3.1.3.

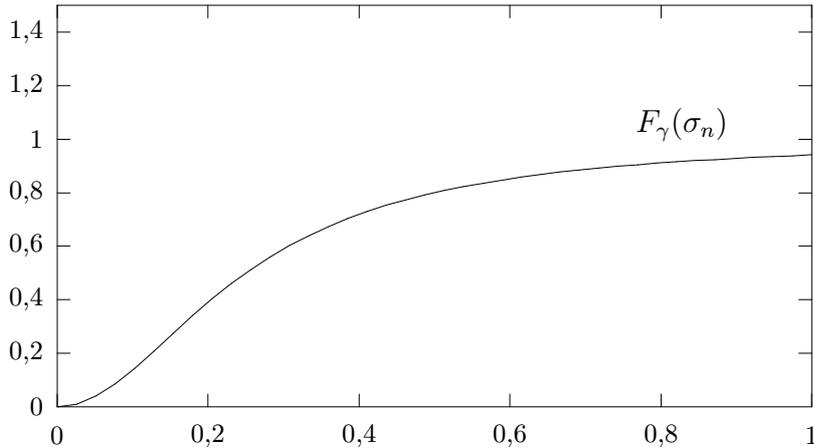


Abbildung 8: Filterwirkung der Tikhonov–Phillips–Regularisierung für einen Regularisierungsparameter $\gamma = 0,25$ und einen Parameter $\beta = 1$.

$$F_\gamma(\sigma_n) = \frac{\sigma_n^2}{\sigma_n^2 + \gamma^2 \beta_n^2}. \quad (35)$$

Dabei sind die σ_n^2 und β_n^2 die Eigenwerte der Matrizen $\mathbf{A}^T \mathbf{A}$ und $\mathbf{B}^T \mathbf{B}$. Aus Gl. (35) geht hervor, daß auch bei der Tikhonov–Phillips–Regularisierung letztlich wieder die einer starken Beeinflußung durch Meßfehler unterliegenden Anteile unterdrückt werden, daß also auch hier ein „Wegdämpfen“ unerwünschter Beiträge erfolgt (siehe Abb. 8). Die Wirkungsweise der Tikhonov–Phillips–Regularisierung (34) läßt sich daher wie folgt beschreiben: Ist ein Eigenwert σ_n^2 groß im Vergleich zum korrespondierenden $\gamma^2 \beta_n^2$, dann ist $F_\gamma(\sigma_n) \approx 1$, d. h. die entsprechenden Anteile der Lösung werden nur wenig verfälscht. Ist σ_n^2 hingegen klein, so ist wegen der Bedingung an die β_n auch das Filter $F_\gamma(\sigma_n)$ klein – die gegenüber Meßfehlern stark empfindlichen Anteile werden weggedämpft. Im Gegensatz zur abgeschnittenen Singulärwertzerlegung werden diese Terme somit nicht weggelassen, sondern durch etwas „Falsches“ ersetzt, das aber keinen Schaden anrichtet [50, S. 91].¹⁸

In der ursprünglichen, auf Tikhonov zurückgehenden Formulierung des Verfahrens mit der Restriktion $\mathbf{f} \approx 0$, d. h. $\mathbf{B} = \mathbf{I}$ und $\mathbf{r} = 0$ waren obige Annahmen bzgl. $\mathbf{A}^T \mathbf{A}$ und $\mathbf{B}^T \mathbf{B}$ 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 Bedingung $\mathbf{f}' = 0$ bzw. $\mathbf{f}'' = 0$, das

¹⁸In jüngerer Zeit werden auch das Verfahren der abgeschnittenen Singulärwertzerlegung und die Tikhonov–Phillips–Regularisierung miteinander kombiniert. Hieraus resultiert ein Regularisierungsverfahren, dessen Wirkungsweise für große σ_n^2 der Tikhonov–Phillips–Regularisierung entspricht und für kleine σ_n^2 in die der abgeschnittenen Singulärwertzerlegung übergeht [1, 31].

Verschwinden der ersten oder zweiten Ableitung, unterdrückt werden können, d. h.

$$\mathbf{B}\mathbf{f} = \begin{pmatrix} 1 & -1 & 0 & \cdots & 0 & 0 \\ 0 & 1 & -1 & \cdots & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -1 \\ 0 & 0 & 0 & \cdots & 0 & 0 \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-1} \\ f_n \end{pmatrix} = 0, \quad (36)$$

bzw.

$$\mathbf{B}\mathbf{f} = \begin{pmatrix} 1 & -2 & 1 & \cdots & 0 & 0 & 0 \\ 0 & 1 & -2 & \cdots & 0 & 0 & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & \cdots & 1 & -2 & 1 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \\ 0 & 0 & 0 & \cdots & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} f_1 \\ f_2 \\ \vdots \\ f_{n-1} \\ f_n \end{pmatrix} = 0.$$

Phillips forderte also explizit, die Lösung solle hinreichend glatt sein. Grundsätzlich kann jedoch jede Art von a priori Information in der Zusatzbedingung (32) mit berücksichtigt werden, wenngleich auch im eigentlichen Sinne regularisierend nur die Bedingungen $\mathbf{f}^{(k)} = 0$ mit $k = 0, 1, \text{etc.}$ wirken.

Überdies ist es bei der Tikhonov–Phillips–Regularisierung möglich, unterschiedliche Meßgenauigkeiten der Elemente des Datenvektors \mathbf{g}_ε zu berücksichtigen, indem man die Einzelmessungen verschieden stark gewichtet. Am einfachsten geschieht dies durch Anwendung einer positiv definiten Diagonalmatrix \mathbf{C} , so daß statt des Funktionals (33) nun das Funktional

$$[\mathbf{A}\mathbf{f}_\gamma - \mathbf{g}_\varepsilon]^T \mathbf{C} [\mathbf{A}\mathbf{f}_\gamma - \mathbf{g}_\varepsilon] + \gamma^2 [\mathbf{B}\mathbf{f}_\gamma - \mathbf{r}]^T [\mathbf{B}\mathbf{f}_\gamma - \mathbf{r}] \quad (37)$$

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

$$\mathbf{f}_\gamma = \underbrace{[\mathbf{A}^T \mathbf{C} \mathbf{A} + \gamma^2 \mathbf{B}^T \mathbf{B}]^{-1}}_{\mathbf{A}_\gamma^{-A}} [\mathbf{A}^T \mathbf{C} \mathbf{g}_\varepsilon + \gamma^2 \mathbf{B} \mathbf{r}]. \quad (38)$$

Die Wahl des Regularisierungsparameters γ erfolgt wie bei der abgeschnittenen Singulärwertzerlegung entweder a priori aufgrund irgendwelcher Erfahrungswerte oder a posteriori, wie in Abschn. 3.1.4 beschrieben. Der numerische Aufwand ist nun allerdings größer als bei der abgeschnittenen Singulärwertzerlegung, da für jeden neuen Wert des Regularisierungsparameters γ die regularisierte Lösung (34) bzw. (38) neu berechnet werden, d. h. jeweils eine zusätzliche Matrixinversion durchgeführt werden muß.

Ein entscheidender Vorteil der Tikhonov–Phillips–Regularisierung gegenüber der abgeschnittenen Singulärwertzerlegung ist jedoch, daß bei diesem Verfahren der Regularisierungsparameter kontinuierlich eingestellt werden kann. Es ist somit möglich,

die Stärke der Regularisierung stufenlos zu regeln. Bei der abgeschnittenen Singulärwertzerlegung bestand dagegen nur die Möglichkeit, jeweils einen weiteren Eigenvektor in die Lösung mit einzubeziehen.

3.2.3 Linear statistisches Verfahren

Bei den beiden bisher besprochenen Regularisierungsverfahren stellte sich jeweils die Frage nach der Wahl des richtigen Regularisierungsparameters γ . Diese erfolgte entweder a posteriori, indem ausgehend von einer stark regularisierten Lösung der Regularisierungsparameter sukzessive immer kleiner gewählt wurde, bis der Fehler in den Daten eine vorgegebene Schranke unterschritt (vgl. Abschn. 3.1.4), oder a priori, aufgrund irgendwelcher Erfahrungswerte. Wie bereits in Abschn. 3.1.4 deutlich wurde, liefert die a posteriori-Parameterwahl keinesfalls das optimale γ , sondern i. d. Regel etwas zu große Werte, so daß die Lösung zwar leicht überregularisiert, andererseits jedoch mit Sicherheit nur geringfügig mit Datenfehlern behaftet ist; diese können ja – wie aus Abb. 5 hervorgeht – im Falle einer zu schwachen Regularisierung nahezu beliebig groß werden. Im Gegensatz dazu ist es bei der a priori-Parameterwahl sehr wohl möglich, das γ optimal zu treffen, doch besteht die Gefahr, daß der voreingestellte Regularisierungsparameter unter Umständen erheblich zu groß oder zu klein gewählt wird und die Lösung infolgedessen stark fehlerbehaftet ist.

Die statischen Regularisierungsverfahren, für die stellvertretend hier das *linear statistische Verfahren* betrachtet werden soll, stellen den Versuch dar, diesen optimalen Wert des Regularisierungsparameters aufgrund wahrscheinlichkeitstheoretischer Überlegungen zu ermitteln, basierend auf den Ergebnissen früherer Inversionsrechnungen, auf den Ergebnissen anderer Meßverfahren oder sonstigen Erkenntnissen. Doch auch die statistisch motivierte a priori-Parameterwahl schließt die Gefahr einer zu starken oder zu schwachen Regularisierung nicht aus, nämlich dann, wenn die aktuelle Messung stark vom bisherigen Mittelwert abweicht. Für einige Verwirrung beim Studium der einschlägigen Literatur sorgt nun vor allen Dingen die Tatsache, daß ein ganze Reihe statistischer Regularisierungsverfahren existiert, welche sich im individuellen Zugang zur Problematik stark unterscheiden, Gaußsche Statistik vorausgesetzt jedoch alle auf ein und diesselbe Inversionsgleichung führen [93]. Erwähnt werden sollten in diesem Zusammenhang neben dem linear statistischen Verfahren die *Bayes-Schätzung* [27], [50, S. 134 ff.], [67], die *optimum estimation method* [33], [68]–[70] (auch *Rodgers-Verfahren* genannt) sowie die *minimum variance method* [25, 26, 34, 87].

Eine unabdingbare Voraussetzung für die Anwendung der statistischen Verfahren ist zunächst die Tatsache, daß ausreichend a priori-Information zur Verfügung stehen muß, um überhaupt Statistik betreiben zu können. Rodgers [68] spricht in diesem Zusammenhang gar von *virtuellen Messungen*, welche neben der eigentlichen Messung

in die Ermittlung der Lösung mit einfließen.¹⁹ Sei $\bar{\mathbf{f}}$ der Mittel- oder Erwartungswert eines Ensembles von a priori bekannten Lösungen \mathbf{f}^i , $i = 1, \dots, N$,

$$\bar{\mathbf{f}} = E(\mathbf{f}) = \frac{1}{N} \sum_{i=1}^N \mathbf{f}^i,$$

so ist die Kovarianzmatrix \mathbf{S}_f dieser Verteilung von Zufallsvektoren gegeben durch

$$\mathbf{S}_f = E([\mathbf{f} - \bar{\mathbf{f}}][\mathbf{f} - \bar{\mathbf{f}}]^T) \approx \frac{1}{N-1} \sum_{i=1}^N [\mathbf{f}^i - \bar{\mathbf{f}}][\mathbf{f}^i - \bar{\mathbf{f}}]^T. \quad (39)$$

Des weiteren seien die statistischen Kenngrößen des Meßfehlervektors \mathbf{E} der aktuellen Messung in Form seines Erwartungswertes $\bar{\mathbf{E}} = 0$ sowie in Form seiner Meßfehlerkovarianzmatrix $\mathbf{S}_E = E(\mathbf{E}\mathbf{E}^T)$ bekannt, wenn auch der Meßfehlervektor \mathbf{E} selbst nicht gegeben ist.²⁰ Wie in Abschn. 3.1.6 angedeutet, stellt die Kovarianzmatrix die Erweiterung der vom Fall einer eindimensionalen Verteilung her bekannten Varianz auf eine Verteilung von Zufallsvektoren dar. Sie ergibt sich direkt aus der Definition der zweiten zentralen Momente (siehe [7, S. 671], [52, S. 21 ff.] und [54, S. 82]). In den Diagonalelementen der Kovarianzmatrix steht die Varianz der Elemente der Zufallsvektoren. Sie gibt an, inwieweit die Elemente der Zufallsvektoren streuen. Die Matrixelemente außerhalb der Diagonalen hingegen sind ein Maß für die Stärke der Korrelation der Zufallsvektoren untereinander. Besteht die Korrelation beispielsweise darin, daß die ersten und zweiten Elemente der Zufallsvektoren jeweils in einem festen Verhältnis zueinander stehen, so sind die entsprechenden Matrixelemente (1, 2) und (2, 1) ungleich Null und relativ groß. Sind die Zufallsvektoren hingegen nicht miteinander korreliert, so ist die Kovarianzmatrix eine reine Diagonalmatrix. Diese Annahme ist für die Meßfehlerkovarianzmatrix \mathbf{S}_E des hier am IMK des FZK entwickelten Radiometers gerechtfertigt. Nimmt man überdies an, daß alle Elemente der Zufallsvektoren gleich stark streuen, so kann die Kovarianzmatrix \mathbf{S}_E geschrieben werden als $\mathbf{S}_E = \sigma_E^2 \mathbf{I}$. Dabei ist σ_E^2 die Varianz und σ_E die Standardabweichung²¹ eines jeden Elements des Meßfehlervektors.

Die bisherigen Angaben bezogen sich auf die statistischen Verfahren ganz allgemein. Dem linear statistischen Verfahren, auch als *bester linearer Schätzer* [50,

¹⁹Diese virtuellen Messungen sind allerdings auch nichts anderes als klimatologische Daten. Inwieweit bei diesen jedoch – wie später gefordert werden wird – a) immer von einer Gaußverteilung ausgegangen werden kann und b) eine Anbindung der Lösung nicht nur in ihrer Struktur sondern auch in ihrem Betrag an die a priori-Information sinnvoll ist, um unbekannte Phänomene zu erforschen sei dahingestellt.

²⁰Es wird angenommen, daß es sich bei den Meßfehlern um rein statistische Fehler handelt (Gaußsches Rauschen), so daß der Erwartungswert $\bar{\mathbf{E}}$ für den Meßfehlervektor \mathbf{E} gleich 0 ist.

²¹Im 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 [88].

S. 130 ff.] oder *maximum likelihood method* [52, S. 79 ff.], [54, 82], [93, S. 153] bezeichnet, liegen nun die folgenden Annahmen zugrunde:

- a) Die Lösungen \mathbf{f} sowie die Meßfehler \mathcal{E} genügen der Gaußschen Statistik ²² und
- b) die Lösungen \mathbf{f} und die Meßfehler \mathcal{E} sind nicht miteinander korreliert.

Aus der ersten Annahme ergibt sich die Wahrscheinlichkeitsdichte für das Auftreten eines bestimmten Höhenprofils \mathbf{f} bzw. eines Meßfehlers \mathcal{E} unmittelbar zu

$$p_f(\mathbf{f}) = \frac{1}{2\pi^{m/2}|\mathcal{S}_f|} e^{-\frac{1}{2}[\mathbf{f}-\bar{\mathbf{f}}]^T \mathcal{S}_f^{-1}[\mathbf{f}-\bar{\mathbf{f}}]} \quad (40)$$

bzw.

$$p_{\mathcal{E}}(\mathcal{E}) = \frac{1}{2\pi^{n/2}|\mathcal{S}_{\mathcal{E}}|} e^{-\frac{1}{2}\mathcal{E}^T \mathcal{S}_{\mathcal{E}}^{-1}\mathcal{E}}. \quad (41)$$

Aus der Annahme b) folgt weiterhin, daß die zusammengesetzte Wahrscheinlichkeitsdichte $p_{f,\mathcal{E}}(\mathbf{f}, \mathcal{E})$ für das gleichzeitige Auftreten einer Lösung \mathbf{f} mit dem Meßfehler \mathcal{E} sich ergibt als das Produkt der beiden Einzelwahrscheinlichkeitsdichten $p_f(\mathbf{f})$ und $p_{\mathcal{E}}(\mathcal{E})$,

$$p_{f,\mathcal{E}}(\mathbf{f}, \mathcal{E}) = p_f(\mathbf{f}) p_{\mathcal{E}}(\mathcal{E}). \quad (42)$$

In Abb. 9 ist der Verlauf dieser zusammengesetzten Wahrscheinlichkeitsdichte, welche bereits vor der eigentlichen Auswertung aus der vorhandenen a priori-Information ermittelt werden kann, durch Höhenlinien skizziert.

Ausgangspunkt der weiteren Überlegungen ist die aktuelle Messung $\mathbf{g}_{\mathcal{E}}$, welche sich zusammensetzt aus dem wahren Meßwert \mathbf{g} und einem (noch) unbekanntem Meßfehler \mathcal{E} , $\mathbf{g}_{\mathcal{E}} = \mathbf{g} + \mathcal{E}$. Die Schwierigkeit besteht nun darin, abzuschätzen, welcher Anteil der aktuellen Messung $\mathbf{g}_{\mathcal{E}}$ dem wahren Meßwert und welcher Anteil dem Meßfehler entspricht. Da weder \mathbf{g} noch \mathcal{E} bekannt sind, läßt sich diese Frage zunächst nur schwer beantworten, doch kann mit Hilfe des mathematischen Modells $\mathbf{A}\mathbf{f} = \mathbf{g}$ zumindest ein Zusammenhang hergestellt werden zwischen dem Meßfehler \mathcal{E} und der Lösung \mathbf{f} ,

$$\mathcal{E} = \mathbf{g}_{\mathcal{E}} - \mathbf{g} = \mathbf{g}_{\mathcal{E}} - \mathbf{A}\mathbf{f} = \mathcal{E}(\mathbf{f}). \quad (43)$$

Gl. (43) ist Ausdruck der Tatsache, daß das Modell innerhalb der Meßfehlergrenzen keine eindeutige Lösung liefert (siehe z. B. Abschn. 3.1.4). Je nachdem, welche Lösung \mathbf{f} aus der Schar der möglichen Lösungen gerade betrachtet wird, ergibt sich ein anderer Meßfehler \mathcal{E} . In Abb. 9 ist dieser Zusammenhang durch eine durchgezogene Linie angedeutet, wobei jeder Punkt der Linie eine aufgrund des Modells mögliche Kombinationen Meßfehler-Lösung repräsentiert. Aus dieser Schar der möglichen Lösungen wird die aufgrund der Gl. (42) wahrscheinlichste Lösung \mathbf{f}_{opt} ermittelt, indem man

²²Eine Annahme, die im vorliegenden Fall, in welchem die Lösungen \mathbf{f} extrem variable klimatische Verhältnisse repräsentieren nur bedingt gerechtfertigt ist.

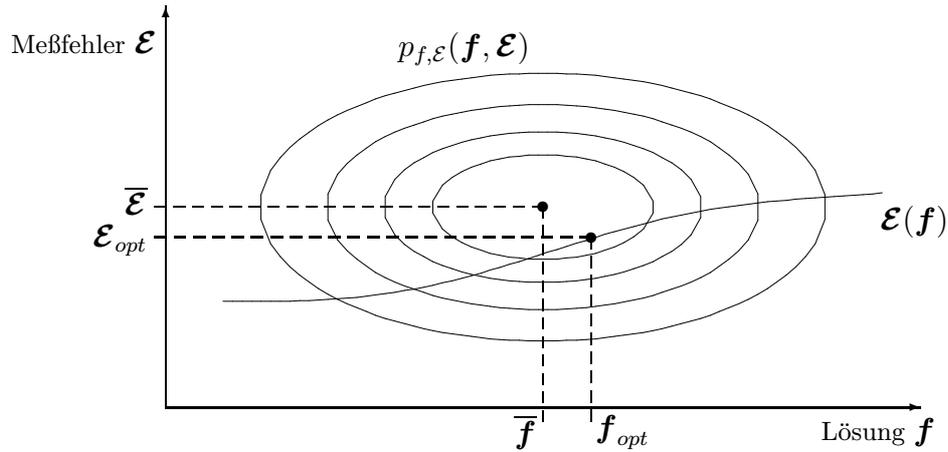


Abbildung 9: Das Zustandekommen der optimalen Näherungslösung \mathbf{f}_{opt} durch Kombination der aktuellen Messung samt Meßfehler $\mathcal{E}(\mathbf{f})$ mit der zur Verfügung stehenden a priori Information $p_{f,\mathcal{E}}(\mathbf{f}, \mathcal{E})$.

zunächst den Meßfehler \mathcal{E} in (42) ersetzt durch Gl. (43) und anschließend das Maximum \mathbf{f}_{opt} dieses dann nur noch von \mathbf{f} abhängigen Ausdrucks bestimmt,

$$\frac{d p_{f,\mathcal{E}}(\mathbf{f}, \mathbf{g}_{\mathcal{E}} - \mathbf{A}\mathbf{f})}{d f_k} = 0.$$

Da \mathbf{f} und \mathcal{E} in (40) und (41) jeweils mit negativem Vorzeichen im Exponenten stehen, führt dieses Extremwertproblem auf die Minimierung des Ausdrucks

$$[\mathbf{g}_{\mathcal{E}} - \mathbf{A}\mathbf{f}]^T \mathbf{S}_{\mathcal{E}}^{-1} [\mathbf{g}_{\mathcal{E}} - \mathbf{A}\mathbf{f}] + [\mathbf{f} - \bar{\mathbf{f}}]^T \mathbf{S}_f^{-1} [\mathbf{f} - \bar{\mathbf{f}}] \quad (44)$$

und schließlich auf die Lösung (siehe Anhang B.1.4)

$$\boxed{\mathbf{f}_{opt} = \underbrace{[\mathbf{A}^T \mathbf{S}_{\mathcal{E}}^{-1} \mathbf{A} + \mathbf{S}_f^{-1}]^{-1}}_{\mathbf{A}_{\gamma}^{-A}} [\mathbf{A}^T \mathbf{S}_{\mathcal{E}}^{-1} \mathbf{g}_{\mathcal{E}} + \mathbf{S}_f^{-1} \bar{\mathbf{f}}]}. \quad (45)$$

Identische Gleichungen ergeben sich, Gaußsche Statistik vorausgesetzt, auch für die anderen, hier nicht betrachteten, statistischen Verfahren.

Auffallend ist die enge Verwandtschaft der regularisierten Lösung (45) mit der Lösung der Tikhonov–Phillips–Regularisierung (vgl. Gl. (45) mit den entsprechenden Gln. (34) bzw. (38) in Abschn. 3.2.2). Die inverse Meßfehlerkovarianzmatrix $\mathbf{S}_{\mathcal{E}}^{-1}$ spielt hierbei die Rolle der Einzelmessungen gewichtenden Matrix \mathbf{C} und die inverse Kovarianzmatrix Matrix \mathbf{S}_f^{-1} die der Zusatzinformationen \mathbf{B} sowie des Regularisierungsparameters γ (vgl. Gl. (44) mit den entsprechenden Ausdrücken (33) bzw. (37)).

Ist $\mathcal{S}_\varepsilon^{-1}$ groß gegenüber \mathcal{S}_f^{-1} , d. h. wurde die Messung mit großer Präzision durchgeführt, so dominiert der erste Term im zu minimierenden Ausdruck (44) gegenüber dem zweiten – die Lösung ist entsprechend schwach regularisiert. Im Idealfall, wenn gar keine Meßfehler auftreten, $\mathbf{g}_\varepsilon \rightarrow \mathbf{g}$ und $\mathcal{S}_\varepsilon^{-1} \rightarrow \infty$, geht (44) über in $|\mathbf{A}\mathbf{f} - \mathbf{g}|^2$. Als Lösung ergibt sich die verallgemeinerte oder Moore–Penrose–Lösung (21), die bereits in Abschn. 3.1.2 für den Fall nicht fehlerbehafteter Daten vorgestellt wurde. Im umgekehrten Fall, wenn die Meßfehler sehr groß sind und der Informationsgehalt der Messung entsprechend gering, wird (44) zu $|\mathbf{f} - \bar{\mathbf{f}}|^2$ und als Lösung ergibt sich der Erwartungswert $\bar{\mathbf{f}}$. D. h. man erhält als Resultat lediglich wieder das zurück, was man bereits zu Beginn an a priori–Information über die Lösung hineinsteckte, da durch die ungenaue Messung keine neue Information hinzukam.

Laut Rodgers [68] ist die Inversionsgleichung (45) in ihrer obigen Form für praktische Anwendungen weniger geeignet, da zu ihrer Auswertung mehrere relativ große Matrizen invertiert werden müssen. Er schlägt deshalb vor, die Gleichung (45) identisch umzuformen (siehe Anhang B.1.5), so daß nur noch eine relativ kleine Matrix invertiert werden muß ²³

$$\mathbf{f}_{opt} = \bar{\mathbf{f}} + \underbrace{\mathcal{S}_f \mathbf{A}^T (\mathbf{A} \mathcal{S}_f \mathbf{A}^T + \mathcal{S}_\varepsilon)^{-1}}_{\mathbf{A}_\gamma^{-A}} (\mathbf{g}_\varepsilon - \mathbf{A}\bar{\mathbf{f}}). \quad (46)$$

Aus heutiger Sicht ist die Verwendung der praktischen Inversionsgleichung (46) jedoch nicht mehr zwingend erforderlich, da die Rechnerleistung in den letzten 20 Jahren um ein Vielfaches verbessert wurde. In der vorliegenden Arbeit wird deshalb die Gleichung (45) auch weiterhin bevorzugt werden, in erster Linie ihrer engen Verwandtschaft zur Inversionsgleichung der Tikhonov–Phillips–Regularisierung wegen. Auf diese Weise ist es möglich, beide Regularisierungsverfahren in einem Programm miteinander zu kombinieren, so daß während einer Inversion per Tastendruck zwischen den beiden Regularisierungsverfahren hin– und hergeschaltet werden kann. ²⁴

Über das im aktuellen Fall rein akademische Interesse hinaus liefert die Gl. (46) jedoch auch wichtige Aufschlüsse über das Fehlerverhalten der Lösung des linear statistischen Verfahrens. Z. B. kann an ihr leicht die der Vollständigkeit halber bereits in Abschn. 3.1.6 präsentierte Gl. (28) zur Berechnung der Kovarianzmatrix \mathcal{S}_γ hergeleitet werden, welche Aussagen über den Gesamtfehler der regularisierten Lösung erlaubt. Mit Hilfe der regularisierten verallgemeinerten Inversen \mathbf{A}_γ^{-A} läßt sich die Gl. (46) nämlich schreiben als

$$\mathbf{f}_{opt} = \bar{\mathbf{f}} + \mathbf{A}_\gamma^{-A} [\mathbf{g}_\varepsilon - \mathbf{A}\bar{\mathbf{f}}]$$

²³Rodgers [68, S. 618 Gl. (69)] verwendet hierzu ein iteratives Verfahren, so daß er letztlich überhaupt keine direkte Matrixinversion mehr durchführt.

²⁴Alle Berechnungen dieser Arbeit wurden auf einem Pentium PC mit 90 MHz und einem Arbeitsspeicher von 16 MB durchgeführt. Zeitlich sehr viel aufwendiger als die eigentlichen Matrixinversionen sind ohnehin die Vorwärtsrechnungen.

$$= \mathbf{A}_\gamma^{-A} \mathbf{g}_\varepsilon + [\mathbf{I} - \mathbf{R}] \bar{\mathbf{f}},$$

dabei ist $\mathbf{R} = \mathbf{A}_\gamma^{-A} \mathbf{A}$ die in Abschn. 3.1.5 eingeführte Auflösungsmatrix. Aus dieser Darstellung der Lösung folgt aufgrund der Rechenregeln für Kovarianzmatrizen (siehe z. B. [52, S. 93] oder [54, S. 82 ff.]) unmittelbar die für alle statistischen Verfahren charakteristische Formel (28) zur Berechnung der Lösungskovarianzmatrix \mathbf{S}_γ , welche im Gegensatz zur Gl. (27) auch eine Aussage über den Regularisierungsfehler liefert (siehe Abschn. 3.1.6).

3.3 Mögliche Fehlerquellen

Am Ende dieses Kapitels sei noch eingegangen auf die bereits zu Beginn des Kapitels aufgeworfene Frage nach der Abhängigkeit der Wirkungsweise der Regularisierungsverfahren von der Formulierung des zugrundeliegenden linearen Gleichungssystems, warum zwischen den beiden Gleichungssystemen (16) und (17) aus Abschn. 2.3.2, welche sich in mathematischer Hinsicht ja lediglich bzgl. ihrer Schreibweise unterscheiden, unter numerischen Gesichtspunkten wesentliche Unterschiede bestehen.

Im Laufe des Kapitels wurde bereits mehrfach deutlich, daß die Regularisierungsverfahren – und insbesondere die Tikhonov–Phillips–Regularisierung – eine Stabilisierung der verallgemeinerten Lösung dadurch erreichen, daß sie a priori–Informationen einführen, wie z. B. die Forderung nach Glattheit der Lösung. Diese werden in Form von Nebenbedingungen an den Lösungsvektor \mathbf{f} gestellt. Nun entspricht der Lösungsvektor \mathbf{f} lediglich im Falle des Gleichungssystems (16) dem eigentlich gesuchten Höhenprofil, im Falle des Gleichungssystems (17) hingegen der Differenz aus Höhenprofil und Anfangsschätzung (siehe die Bemerkungen zur Gl. (18)). Die Nebenbedingungen werden also einmal an das Höhenprofil selbst und einmal an den Differenzvektor aus Höhenprofil und Anfangsschätzung gestellt, wobei letzteres zu sicherlich unvorhergesehenen Inversionsergebnissen führt; denn nicht die Differenz aus Lösung und Anfangsschätzung, sondern die Lösung selbst, das zu invertierende Höhenprofil, soll ja möglichst glatt sein. Im Einzelfall mag es zwar durchaus sinnvoll sein, auch einmal auf die Formulierung (17) zurückgreifen zu können, in der Mehrzahl der Fälle werden jedoch ausschließlich a priori–Informationen über die Höhenprofile selbst und nicht über deren Abweichungen von einer Anfangsschätzung zur Verfügung stehen.

Überdies birgt das lineare Gleichungssystem (17) eine weitere, noch viel größere Gefahr. Rodgers [68, S. 621] spricht in diesem Zusammenhang gar von einer tückischen Fallgrube, in die viele Autoren sicherlich ungewollt stürzen.²⁵ Wie in Abschn. 2.3 über die Linearisierung der SÜG deutlich wurde, muß zur Lösung der nicht-linearen SÜG iteriert werden. D. h. nachdem das lineare Gleichungssystem (17) in einem ersten Schritt erfolgreich gelöst wurde, dient die Lösung als neue, verbesserte

²⁵Auch der Autor dieser Arbeit lag während der Implementierungsphase seines Auswerteprogrammes gleich mehrmals darin.

Anfangsschätzung für den jeweils nachfolgenden Iterationsschritt. Zu diesem Zweck werden basierend auf der neuen Anfangsschätzung ein neues synthetisches Spektrum \mathbf{b}^0 sowie eine verbesserte Jakobimatrix \mathbf{A}^0 berechnet, welche ein neues lineares Gleichungssystem definieren. Die große Gefahr des Gleichungssystems (17) besteht nun darin, daß die ursprünglich über \mathbf{f}^0 in den Inversionsalgorithmus hineingesteckte (und zur Regularisierung der Lösung natürlich unbedingt erforderliche) a priori-Information im Laufe der Iteration vom Nutzer unbemerkt immer mehr verloren geht, weil er auf der linken Seite von (17) für \mathbf{f}^0 die jeweils neueste Anfangsschätzung einsetzte, anstatt weiterhin die ursprüngliche a priori-Information zu verwenden.

Ein Beispiel möge dies verdeutlichen: Angenommen als Anfangsschätzung diene das repräsentative Höhenprofil $\bar{\mathbf{f}}$ und die zur Regularisierung verwendete a priori-Information laute, die Abweichung der Lösung von diesem Höhenprofil sei möglichst klein. Nach dem ersten Iterationsschritt ergibt sich aus (17) – wie erwünscht – eine Näherungslösung, welche von dem repräsentativen Höhenprofil $\bar{\mathbf{f}}$ nur geringfügig abweicht. Wird jedoch in den nachfolgenden Iterationsschritten auf der linken Seite der Gl. (17) für \mathbf{f}^0 die Lösung des jeweils vorangegangenen Iterationsschritts eingesetzt (wie die Notation der Gl. (17) dies nahelegt), so lautet die Nebenbedingung nicht mehr – wie ursprünglich beabsichtigt – die Abweichung der Lösung von dem repräsentativen Höhenprofil $\bar{\mathbf{f}}$ sei möglichst klein, sondern die Abweichung der Lösung von der Lösung des jeweils vorangegangenen Iterationsschritts sei möglichst klein. Im Laufe der Iteration wird sich die Lösung deshalb immer weiter von dem repräsentativen Höhenprofil $\bar{\mathbf{f}}$ entfernen, so daß letztlich das zur Lösung des linearen Gleichungssystems verwendete Regularisierungsverfahren nicht mehr regularisierend wirkt. Statt dessen wird die nicht erwünschte exakte, i. a. aber instabile Lösung erreicht.

Diese Gefahr besteht bei dem linearen Gleichungssystem (16) nicht, da in diesem der Vektor \mathbf{f}^0 mit auf der rechten Seite der Gleichung steht. Vielmehr muß in (16) für \mathbf{f}^0 sogar die jeweils aktuellste Anfangsschätzung eingesetzt werden, wie man dies aufgrund der Notation der Gl. (16) auch erwarten würde. Aus den besagten Gründen sollte das Gleichungssystem (16) daher unbedingt bevorzugt werden.

Zum Schluß noch einige Bemerkungen zu einem geeigneten Abbruchkriterium für das iterative Gauß-Newton-Verfahren. Wie aus der Diskussion der a posteriori-Parameterwahl in Abschn. 3.1.4 hervorgeht, sollte die Differenz zwischen den gemessenen Daten \mathbf{g}_ε und den aus der Lösung \mathbf{f}_γ über Vorwärtsrechnung ermittelten Daten $\mathbf{g}_\gamma = \mathbf{A}\mathbf{f}_\gamma$ keinesfalls kleiner werden als das Meßfehlerauschen \mathcal{E} . Zur Formulierung eines geeigneten Abbruchkriteriums ist der Defekt in den Daten infolgedessen wenig geeignet; vielmehr dient er primär der Wahl des Regularisierungsparameters γ . Statt dessen sollte ein geeignetes Abbruchkriterium bevorzugt in der Lösungsebene angreifen und das iterative Gauß-Newton-Verfahren dann abgebrochen werden, wenn die Differenz der Lösungen zweier aufeinanderfolgende Iterationsschritte klein wird und einen noch vorzugebenden Schwellenwert unterschreitet. In dieser Arbeit war dies nach 2–3 Iterationsschritten praktisch immer der Fall (abhängig natürlich vom Grad

der Nichtlinearität der SÜG).

Damit stehen die wichtigsten, der für eine Inversion der nichtlinearen SÜG benötigten, numerischen Werkzeuge nun zur Verfügung, so daß einer ersten Auswertung von Meßdaten im Prinzip nichts mehr im Wege stehen sollte. Bevor dies konkret in Angriff genommen wird, soll im nächsten Kapitel zunächst noch eingegangen werden auf die Erweiterung der besprochenen Regularisierungsverfahren auf die Inversion mehrerer Höhenprofile simultan aus einem Spektrum; dies ist insbesondere dann sinnvoll, wenn ein Spektrum Signaturen mehrerer verschiedener Gase enthält, sowie auf die Frage der Berücksichtigung von stehenden Wellen, welche den Spektren in der Millimeterwellenradiometrie häufig überlagert sind. Überdies werden in Kapitel 5 über die Auswertung von Meßdaten noch einige gerätespezifische Anpassungen des Auswertalgorithmus notwendig sein.

Kapitel 4

Erweiterung der Inversionsverfahren

4.1 Ableitung mehrerer Spurengasprofile simultan aus Messungen vom Boden aus

Wie bereits mehrfach erwähnt, waren die Ausführungen des vorigen Kapitels über die Theorie der Inversion und über die Regularisierungsverfahren eher allgemeiner Natur und spiegelten den aktuellen Wissensstand über die Inversion schlecht gestellter Probleme wider – losgelöst von dem speziellen Problem der bodengebundenen, passiven Millimeterwellenradiometrie. Hierzu war bisher vereinfachend angenommen worden, daß lediglich versucht werden sollte, ein Höhenprofil aus dem gemessenen Strahldichtespektrum zu ermitteln, um dem interdisziplinären Charakter des bisher Gesagten Rechnung zu tragen. Diese Einschränkung soll nun fallengelassen werden, d. h. es soll versucht werden, aus einem Strahldichtespektrum Höhenprofile verschiedener Spurengase simultan abzuleiten.¹ Dieser Ansatz ist neu. Wie aus der Gl. (15) in Abschn. 2.3.2 hervorgeht, setzt sich der Lösungsvektor \mathbf{f} in diesem Fall gleich aus mehreren Höhenprofilen zusammen und die Matrix \mathbf{A} entsteht durch die Hintereinanderreihung mehrerer quadratischer Teilmatrizen,

$$\mathbf{f} = \begin{pmatrix} \mathbf{f}_T \\ \mathbf{f}_{c_v} \\ \vdots \end{pmatrix} \quad \text{bzw.} \quad \mathbf{A} = (\mathbf{A}_T, \mathbf{A}_{c_v}, \dots).$$

Es leuchtet unmittelbar ein, daß die Regularisierungsverfahren nicht unverändert auch auf diesen neuen Sachverhalt übertragen werden können, sondern entsprechend modifiziert werden müssen.

Beim Tikhonov–Phillips–Verfahren erfolgte die Regularisierung der verallgemeinerten Lösung bisher dadurch, daß Zusatzinformationen in Form von Nebenbedingungen an die Lösung eingeführt wurden; die Lösung sollte beispielsweise möglichst glatt sein. Diese Bedingungen dürfen nun nicht mehr an den Lösungsvektor \mathbf{f} gestellt werden, sondern müssen an den Höhenprofilen \mathbf{f}_T , \mathbf{f}_{c_v} , usw. angreifen. Nicht der

¹Dies wird natürlich nur unter bestimmten Voraussetzungen möglich sein, deren wichtigste sicherlich ist, daß das Strahldichtespektrum überhaupt Signaturen all jener Spurengase aufweist, welche aus dem Spektrum invertiert werden sollen.

zusammengesetzte Lösungsvektor \mathbf{f} , sondern die Höhenprofile von Temperatur und Volumenmischungsverhältnissen interessierender Gase sollen möglichst glatt sein, wobei zwischen den einzelnen Lösungsprofilen keine Kopplung erwünscht ist. Des Weiteren stellt sich bei dem Versuch der Ermittlung mehrerer Höhenprofile simultan aus einem Spektrum die Frage nach der Stärke der Regularisierung. Es darf sicherlich nicht erwartet werden, daß zur Stabilisierung aller Höhenprofile gleich stark reguliert werden muß. Vielmehr ist wohl davon auszugehen, daß ein jedes Höhenprofil eines eigenen Regularisierungsparameters bedarf, welcher individuell auf das jeweilige Höhenprofil eingestellt werden muß. Dies folgt allein schon aus der Tatsache, daß die Signaturen verschiedener Gase innerhalb eines Spektrum in ihrer Stärke i. d. Regel enorm variieren.

Diese Überlegungen führen dazu, daß – statt bisher lediglich einer – nun gleich mehrere Nebenbedingungen an die verallgemeinerte Lösung gestellt werden, eine jede für ein bestimmtes Höhenprofil. Anstelle des Tikhonov–Phillips–Funktional (33) muß deshalb das modifizierte Funktional

$$J_\gamma(\mathbf{f}_\gamma) := |\mathbf{A}\mathbf{f}_\gamma - \mathbf{g}_\mathcal{E}|^2 + \gamma_T^2 |\widetilde{\mathbf{B}}_T \mathbf{f}_\gamma - \widetilde{\mathbf{r}}_T|^2 + \gamma_{c_v}^2 |\widetilde{\mathbf{B}}_{c_v} \mathbf{f}_\gamma - \widetilde{\mathbf{r}}_{c_v}|^2 + \dots \quad (47)$$

minimiert werden. Dabei sind die $\gamma_T^2, \gamma_{c_v}^2, \dots$ Langrangesche Multiplikatoren, bzw. – mehr in Bezug auf die beabsichtigte Regularisierung der verallgemeinerten Lösung – die einem jeden Höhenprofil zugeordneten Regularisierungsparameter. Da eine jede dieser Nebenbedingungen sich auf genau ein Höhenprofil bezieht, sind die Matrizen $\widetilde{\mathbf{B}}_T, \widetilde{\mathbf{B}}_{c_v}, \dots$ von der besonders einfachen Form

$$\widetilde{\mathbf{B}}_T = \begin{pmatrix} \mathbf{B}_T & \mathbf{O} & \dots \\ \mathbf{O} & \mathbf{O} & \\ \vdots & & \ddots \end{pmatrix}, \quad \widetilde{\mathbf{B}}_{c_v} = \begin{pmatrix} \mathbf{O} & \mathbf{O} & \dots \\ \mathbf{O} & \mathbf{B}_{c_v} & \\ \vdots & & \ddots \end{pmatrix} \quad \text{usw.},$$

wobei die Teilmatrizen $\mathbf{B}_T, \mathbf{B}_{c_v}, \dots$ im wesentlichen der Matrix \mathbf{B} aus Abschn. 3.2.2 entsprechen. Die Matrix \mathbf{O} ist die Nullmatrix. Die Vektoren $\widetilde{\mathbf{r}}_T, \widetilde{\mathbf{r}}_{c_v}, \dots$ schließlich sind gegeben durch

$$\widetilde{\mathbf{r}}_T = \begin{pmatrix} \mathbf{r}_T \\ \mathbf{o} \\ \vdots \end{pmatrix}, \quad \widetilde{\mathbf{r}}_{c_v} = \begin{pmatrix} \mathbf{o} \\ \mathbf{r}_{c_v} \\ \vdots \end{pmatrix} \quad \text{etc.},$$

dabei stehen die Teilvektoren $\mathbf{r}_T, \mathbf{r}_{c_v}, \dots$ für die einer jeden Nebenbedingung zugeordnete rechte Seite und entsprechen im wesentlichen der rechten Seite \mathbf{r} aus Gl. (32) in Abschn. 3.2.2. Der Vektor \mathbf{o} ist der Nullvektor.

Die Minimierung des obigen Funktional (47) erfolgt analog zum Abschn. 3.2.2 über die Tikhonov–Phillips–Regularisierung (vgl. hierzu auch Anhang B.1.2) und

föhrt auf die regularisierte verallgemeinerte Lösung

$$\begin{aligned} \mathbf{f}_\gamma &= \left[\mathbf{A}^T \mathbf{A} + \gamma_T^2 \widetilde{\mathbf{B}}_T^T \widetilde{\mathbf{B}}_T + \gamma_{c_v}^2 \widetilde{\mathbf{B}}_{c_v}^T \widetilde{\mathbf{B}}_{c_v} + \dots \right]^{-1} \\ &\times \left[\mathbf{A}^T \mathbf{g}_\varepsilon + \gamma_T^2 \widetilde{\mathbf{B}}_T^T \widetilde{\mathbf{r}}_T + \gamma_{c_v}^2 \widetilde{\mathbf{B}}_{c_v}^T \widetilde{\mathbf{r}}_{c_v} + \dots \right]. \end{aligned} \quad (48)$$

Für die Implementierung in ein Programm ist diese Gleichung allerdings weniger geeignet, da sie eine enorme Verschwendung des grundsätzlich knappen Speicherplatzes zur Folge hätte. Schließlich bestehen die Matrizen $\widetilde{\mathbf{B}}_T, \widetilde{\mathbf{B}}_{c_v}, \dots$ und die rechten Seiten $\widetilde{\mathbf{r}}_T, \widetilde{\mathbf{r}}_{c_v}, \dots$ im wesentlichen aus Nullen. Bezieht man dieses Wissen in die Berechnung der Lösung mit ein, so kann die regularisierte Lösung nach wenigen Zwischenschritten auch geschrieben werden als

$$\mathbf{f}_\gamma = \left[\mathbf{A}^T \mathbf{A} + \begin{pmatrix} \gamma_T^2 \mathbf{B}_T^T \mathbf{B}_T & \mathbf{O} & \dots \\ \mathbf{O} & \gamma_{c_v}^2 \mathbf{B}_{c_v}^T \mathbf{B}_{c_v} & \\ \vdots & & \ddots \end{pmatrix} \right]^{-1} \left[\mathbf{A}^T \mathbf{g}_\varepsilon + \begin{pmatrix} \gamma_T^2 \mathbf{B}_T^T \mathbf{r}_T \\ \gamma_{c_v}^2 \mathbf{B}_{c_v}^T \mathbf{r}_{c_v} \\ \vdots \end{pmatrix} \right]. \quad (49)$$

Gegenüber der Gl. (48) bedeutet diese Schreibweise einen erheblich geringeren Speicherplatzbedarf. Implementiert man die Gl. (49) überdies noch geschickt, indem man zunächst die Matrixprodukte $\mathbf{B}_T^T \mathbf{B}_T, \mathbf{B}_{c_v}^T \mathbf{B}_{c_v}$, usw. berechnet und anschließend gewichtet mit dem zugehörigen Regularisierungsparameter geeignet zur Matrix $\mathbf{A}^T \mathbf{A}$ hinzuaddiert, wächst der Speicherplatzbedarf nur noch proportional zu $(n^2 + n)$ gegenüber $(n^3 + n^2)$ in Gl. (48). Dabei ist n die Anzahl der gleichzeitig zu invertierenden Höhenprofile.

Die hier abgeleitete Gleichung (49) ist die Erweiterung der einfachen Inversionsgleichung (34) der Tikhonov–Phillips–Regularisierung auf den Fall gleich mehrerer Höhenprofile. Es ist dabei möglich, unterschiedliche Meßgenauigkeiten der rechten Seite \mathbf{g}_ε zu berücksichtigen, indem man analog zur Gl. (37) eine positiv definite Diagonalmatrix \mathbf{C} eingeföhrt, welche die Einzelmessungen entsprechend gewichtet. Es resultiert die allgemeinere Inversionsgleichung

$$\mathbf{f}_\gamma = \left[\mathbf{A}^T \mathbf{C} \mathbf{A} + \begin{pmatrix} \gamma_T^2 \mathbf{B}_T^T \mathbf{B}_T & \mathbf{O} & \dots \\ \mathbf{O} & \gamma_{c_v}^2 \mathbf{B}_{c_v}^T \mathbf{B}_{c_v} & \\ \vdots & & \ddots \end{pmatrix} \right]^{-1} \left[\mathbf{A}^T \mathbf{C} \mathbf{g}_\varepsilon + \begin{pmatrix} \gamma_T^2 \mathbf{B}_T^T \mathbf{r}_T \\ \gamma_{c_v}^2 \mathbf{B}_{c_v}^T \mathbf{r}_{c_v} \\ \vdots \end{pmatrix} \right]. \quad (50)$$

Ähnlich der Tikhonov–Phillips–Regularisierung können auch die statistischen Regularisierungsverfahren auf den Fall gleich mehrerer Höhenprofile erweitert werden. Die hieraus resultierende Inversionsgleichung soll jedoch nicht abgeleitet, sondern einfach durch Vergleich der Inversionsgleichung der Tikhonov–Phillips–Regularisierung

(siehe Abschn. 3.2.2) mit der entsprechenden Inversionsgleichung des linear statistischen Verfahrens (siehe Abschn. 3.2.3) aus obiger Gleichung ermittelt werden. Wie aus dem Vergleich der beiden Gln. (45) und (38) hervorgeht, müssen hierzu in obiger Gleichung lediglich die Ausdrücke $\gamma^2 \mathbf{B}^T \mathbf{B}$ ersetzt werden durch die Inversen der entsprechenden a priori-Kovarianzmatrizen $\mathbf{S}_\varepsilon^{-1}$, die Vektoren \mathbf{r} durch die entsprechenden a priori-Profile $\bar{\mathbf{f}}$ sowie die Matrix \mathbf{C} durch die Inverse der Fehlerkovarianzmatrix $\mathbf{S}_\varepsilon^{-1}$. Es resultiert die Erweiterung des linear statistischen Verfahrens auf den Fall mehrerer Höhenprofile,

$$\mathbf{f}_\gamma = \left[\mathbf{A}^T \mathbf{S}_\varepsilon^{-1} \mathbf{A} + \begin{pmatrix} \mathbf{S}_{f_T}^{-1} & \mathbf{O} & \cdots \\ \mathbf{O} & \mathbf{S}_{f_{c_v}}^{-1} & \\ \vdots & & \ddots \end{pmatrix} \right]^{-1} \left[\mathbf{A}^T \mathbf{S}_\varepsilon^{-1} \mathbf{g}_\varepsilon + \begin{pmatrix} \mathbf{S}_{f_T}^{-1} \bar{\mathbf{f}}_T \\ \mathbf{S}_{f_{c_v}}^{-1} \bar{\mathbf{f}}_{c_v} \\ \vdots \end{pmatrix} \right]. \quad (51)$$

Problematisch ist hingegen die Erweiterung der abgeschnittenen Singulärwertzerlegung auf den Fall der Ableitung mehrerer Höhenprofile simultan aus einem Spektrum, da bei diesem Regularisierungsverfahren keine Nebenbedingungen an die Lösung gestellt werden, sondern die Regularisierung durch Aufsummieren über die Eigenvektoren der Matrix $\mathbf{A}^T \mathbf{A}$ erfolgt (siehe Abschn. 3.2.1). Insbesondere ist es bei diesem Verfahren schwierig, die Stärke der Regularisierung für jedes Höhenprofil individuell zu regeln, da ausgehend von einer für alle Höhenprofile gleich schwachen Regularisierung immer nur über einen weiteren Eigenvektor aufsummiert werden kann. Letzten Endes führt dies auf eine insgesamt stärkere, jedoch erneut für alle Höhenprofile gleiche Regularisierung. Unter anderem wurde deshalb versucht, nur die einem jeden Höhenprofil zugeordneten Teile der besagten Eigenvektoren aufzuaddieren, um vielleicht auf diese Weise eine individuelle Regularisierung der Höhenprofile zu erreichen. Zwar war es somit möglich mehrere Spurengase simultan zu invertieren, doch blieb die Qualität der erzielten Inversionsergebnisse deutlich gegenüber derjenigen obiger beider Verfahren zurück. Über die Ursachen kann bisher nur spekuliert werden.

Als Abbruchkriterium dient auch im Falle der Ableitung mehrerer Höhenprofile simultan aus einem Spektrum weiterhin die Tatsache, daß die Differenz der Lösungen zweier aufeinanderfolgender Iterationsschritte möglichst klein sein muß und einen noch vorzugebenden Schwellenwert unterschreiten sollte.

4.2 Berücksichtigung von stehenden Wellen

Ein in der Mikro- oder Millimeterwellenradiometrie häufig auftretendes Problem sind stehende Wellen, welche von Mehrfachreflexionen innerhalb des Meßgerätes herrühren (siehe z. B. Abb. 24 auf Seite 81). Zwar ist ein teilweises Unterdrücken dieser unerwünschten Effekte durch periodisches Einfügen eines geeigneten Dielektrikums oder

durch mechanische Weglängenmodulation möglich, so daß die stehende Wellen sich weitgehend herausmitteln – doch können sie im allgemeinen nicht vollständig eliminiert werden. Dem von der Atmosphäre emittierten Spektrum sind deshalb stehende Wellen der Form

$$A \sin(k\nu + \phi) \quad (52)$$

überlagert. Es ist nun die Aufgabe der Auswertung, diese stehenden Wellen entweder im nachhinein zu eliminieren (jedoch noch vor der eigentlichen Inversion der Spektren) oder besser noch – falls dies irgendwie realisiert werden könnte – bei der Inversion der gemessenen Spektren gleich mit zu berücksichtigen.

Bisher gehen alle in der Atmosphärenchemie tätigen Gruppen in Ermangelung eines geeigneten Verfahrens zur Realisation der Möglichkeit zwei nach der ersten Methode vor. Das heißt sie unterziehen ihre gemessenen „Rohspektren“ zunächst einer relativ aufwendigen Elimination von stehenden Wellen, bevor sie die von stehenden Wellen bereinigten Spektren dann invertieren. Es ist leicht einzusehen, daß diese Aufbereitung der gemessenen Spektren nicht gerade einfach zu bewerkstelligen ist, denn schließlich sind die stehenden Wellen einem im Prinzip unbekanntem Spektrum überlagert. Die Gruppe des Instituts für Fernerkundung der Universität Bremen [36, S. 92 ff.], [64, S. 58 ff.] geht deshalb so vor, daß sie in einem Fitprogramm (least-squares-Verfahren) gemeinsam mit den stehenden Wellen ein synthetisches Spektrum an das gemessene Spektrum anpaßt. Zu diesem Zweck wird ein typisches, mittels Vorwärtsrechnung synthetisch erzeugtes Spektrum mit drei Fitparametern (für Untergrund, Steigung des Untergrundes und passende Skalierung) versehen, welche zusammen mit den Amplituden und Phasen der stehenden Wellen entsprechend variiert werden können. Im Anschluß an die bestmögliche Anpassung wird das gemessene Spektrum schließlich um die gefitteten stehenden Wellen korrigiert.

Die große Gefahr der obigen Vorgehensweise besteht in einer falschen Anpassung der stehenden Wellen an das gemessene Spektrum, da das mit lediglich drei Fitparametern versehene synthetische Spektrum das der Messung zugrundeliegende Spektrum nur unzulänglich beschreibt. Dies gilt insbesondere für stehende Wellen relativ großer Periode ([36, S. 92 unten]). So läßt sich beispielsweise eine Differenz zwischen dem angepaßten synthetischen Spektrum und dem gemessenen Spektrum in der Linienamplitude leicht durch Anpassen eines zur Linienmitte symmetrischen Kosinus-Terms ausgleichen. In der Folge sind alle korrigierten Spektren sehr ähnlich dem für die komplette Anpassung verwendeten synthetischen Spektrum und auch die letztlich invertierten Höhenprofile weichen in ihrer Form nur geringfügig von dem für die Berechnung des synthetischen Spektrums verwendeten Höhenprofil ab [36, S. 106].

Im folgenden soll deshalb ein neues Verfahren vorgeschlagen werden, mit dem die stehende Wellen im Zuge der Inversion des gemessenen Spektrums gleich mit berücksichtigt werden können. Da in diesem Falle eine exakte Anpassung der Vorwärtsrechnung an das wahre Spektrum möglich ist, besteht insbesondere die Gefahr einer Fehlanpassung von stehenden Wellen an das gemessene Spektrum nicht mehr bzw.

nur noch in weitaus geringerem Umfang. Überdies spielt auch der zeitliche Aspekt eine nicht unwesentliche Rolle, da die sehr aufwendige Korrektur von stehenden Wellen nun nicht mehr durchgeführt werden muß.

Aus der Periode $P = 2\pi/k$ der stehenden Wellen im Spektrum kann über die Fabry–Pérot–Formel

$$l = \frac{c}{2P}$$

unmittelbar auf die Länge l der Resonatoren geschlossen werden; dabei ist c die Lichtgeschwindigkeit. Da sich die Geometrie des Meßgerätes zumindest während eines Meßtages nicht ändert, sind die Längen der Resonatoren und folglich auch die Perioden P der stehenden Wellen im Spektrum konstant, von äußerst geringfügigen Schwankungen, z. B. bedingt durch die geringe thermische Ausdehnung des Meßgerätes, einmal abgesehen. Die Ermittlung der Periode der stehenden Wellen im Spektrum bedeutet von daher kein großes Problem, zumal sie anhand der Längen der Resonatoren, nachdem diese erst einmal identifiziert wurden, relativ genau bestimmt werden können.

Andererseits wirken sich geringfügige Schwankungen der Resonatorlänge stark aus auf die Phase ϕ des obigen Musters der stehenden Wellen (52);² überdies ändern sich auch die Amplituden der stehenden Wellen im Spektrum von Messung zu Messung relativ stark. Neben dem Problem der Inversion des linearen Gleichungssystems (18) besteht deshalb das Problem der gleichzeitigen Anpassung von Amplituden und Phasen etwaiger stehenden Wellen an das gemessene Spektrum. In der Form der Gl. (52) ist das Problem der Anpassung der Phase einer stehenden Welle an ein Spektrum nichtlinear; es kann jedoch in ein lineares überführt werden, indem man Gl. (52) schreibt als

$$a \sin(k\nu) + b \cos(k\nu),$$

wobei $A = \sqrt{a^2 + b^2}$ und $\tan \phi = b/a$. Dieses nun lineare Problem der Anpassung der beiden Parameter a und b (für eine jede, im Spektrum auftretende stehende Welle) kann in das lineare Problem der Inversion des Gleichungssystems (18) integriert werden; dabei werden die Perioden $P = 2\pi/k$ der stehenden Wellen als bekannt

²Z. B. trat im Falle des am IMK des FZK entwickelten Ozonradiometers folgender Effekt auf: Eine im Spektrum deutlich sichtbare stehende Welle der Periode $P = 0,113$ GHz (dies entspricht einer Resonatorlänge von $l = 1,33$ m) konnte eindeutig identifiziert werden als Mehrfachreflexionen zwischen dem Antennen–Hohlleiter–Übergang und dem kalten Absorber. Letzterer bestand zum damaligen Zeitpunkt im wesentlichen aus einem mit flüssigem Stickstoff gefüllten Dewar, welches innen mit einem geeigneten Absorbermaterial verkleidet war. Der „Sehstrahl“ wurde von oben mittels Spiegelschaltung auf diese Anordnung gerichtet. Die eigentliche Ursache der stehenden Welle waren Reflexionen an der Stickstoffoberfläche, wobei der Stickstoffpegel im Laufe der Messungen aufgrund der Verdampfung von flüssigem Stickstoff kontinuierlich sank. Der Resonator wurde deshalb immer länger (im Laufe einer Stunde um etwa 0,5 cm, was einer relativen Änderung der Periode P um weniger als 0,5 % entspricht; insofern konnte die Periode als konstant angenommen werden). Hingegen verschob sich die Phase ϕ während dieser Zeit um etwa 10π . Dies war auch in den integrierten Spektren zu beobachten, indem sich die „stehende Welle“ im Laufe einer Messung langsam verschob und schließlich herausmittelte.

vorausgesetzt. Beachtet werden muß jedoch, daß das inverse Problem schlecht gestellt ist, daß zu seiner Lösung also regularisiert werden muß, während die beiden Parameter a und b eindeutig aus der obigen Gleichung hervorgehen.

Formal läßt sich die Überlagerung des gemessenen Spektrums bzw. der rechten Seite \mathbf{g}_ε des linearen Gleichungssystems (18) mit stehenden Wellen darstellen als ³

$$\begin{pmatrix} a_{11} & \cdots & a_{1m} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nm} \end{pmatrix} \begin{pmatrix} f_1 \\ \vdots \\ f_m \end{pmatrix} + a \begin{pmatrix} \sin(k\nu_1) \\ \vdots \\ \sin(k\nu_n) \end{pmatrix} + b \begin{pmatrix} \cos(k\nu_1) \\ \vdots \\ \cos(k\nu_n) \end{pmatrix} = \begin{pmatrix} g_{\varepsilon 1} \\ \vdots \\ g_{\varepsilon n} \end{pmatrix}.$$

Geschickter ist es jedoch, den Lösungsvektor \mathbf{f} sowie die beiden Parameter a und b zu einem neuen Vektor \mathbf{f}' zusammenzufassen, woraus das gestrichene Gleichungssystem (53) entsteht,

$$\underbrace{\begin{pmatrix} a_{11} & \cdots & a_{1m} & \sin(k\nu_1) & \cos(k\nu_1) \\ \vdots & \ddots & \vdots & \vdots & \vdots \\ a_{n1} & \cdots & a_{nm} & \sin(k\nu_n) & \cos(k\nu_n) \end{pmatrix}}_{\mathbf{A}'} \underbrace{\begin{pmatrix} f_1 \\ \vdots \\ f_m \\ a \\ b \end{pmatrix}}_{\mathbf{f}'} = \underbrace{\begin{pmatrix} g_{\varepsilon 1} \\ \vdots \\ g_{\varepsilon n} \end{pmatrix}}_{\mathbf{g}_\varepsilon}. \quad (53)$$

Wie das ursprüngliche Gleichungssystem (18) ist natürlich auch dieses Gleichungssystem schlecht gestellt bzgl. des „ungestrichenen Anteils“ \mathbf{f} der Lösung, so daß zu seiner Berechnung auf jeden Fall regularisiert werden muß. Andererseits ist das Gleichungssystem gut gestellt bzgl. der beiden Parameter a und b , welche eindeutig aus (53) hervorgehen und deshalb nicht regularisiert werden dürfen. Diese Nebenbedingung einer nur teilweisen Regularisierung des Lösungsvektors \mathbf{f}' wird realisiert in Gl. (54), welche sich in Anlehnung an die Inversionsgleichung (50) der Tikhonov-Phillips-Regularisierung unmittelbar als Lösung des obigen Gleichungssystems (53)

³Im folgenden wird der besseren Übersichtlichkeit willen nur von einer dem Spektrum überlagerten stehenden Welle ausgegangen. Die Erweiterung des Algorithmusses auf mehrere stehende Wellen ist trivial.

ergibt:

$$\mathbf{f}'_{\gamma} = \begin{pmatrix} f_{\gamma 1} \\ \vdots \\ f_{\gamma m} \\ a \\ b \end{pmatrix} = \left[\mathbf{A}'^T \mathbf{C} \mathbf{A}' + \begin{pmatrix} \left(\begin{matrix} \gamma_T^2 \mathbf{B}_T^T \mathbf{B}_T & \mathbf{O} & \cdots \\ \mathbf{O} & \gamma_{c_v}^2 \mathbf{B}_{c_v}^T \mathbf{B}_{c_v} & \cdots \\ \vdots & \vdots & \ddots \\ 0 & \cdots & \\ 0 & \cdots & \end{matrix} \right) & \begin{matrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 0 & 0 \end{matrix} \right]^{-1} \times \left[\mathbf{A}'^T \mathbf{C} \mathbf{g}_{\varepsilon} + \begin{pmatrix} \left(\begin{matrix} \gamma_T^2 \mathbf{B}_T^T \mathbf{r}_T \\ \gamma_{c_v}^2 \mathbf{B}_{c_v}^T \mathbf{r}_{c_v} \\ \vdots \\ 0 \\ 0 \end{matrix} \right) \right] \quad (54)$$

Im Vergleich zu (50) wurden in (54) lediglich die für die Regularisierung verantwortlichen Matrizen und Vektoren an den entsprechenden Stellen um Nullen erweitert, so daß eine Nicht-Regularisierung der beiden Parameter a und b gewährleistet ist.

Gleichung (54) ist der bisher allgemeinste Fall der Lösung der Tikhonov-Phillips-Regularisierung. Neben der Erweiterung des Verfahrens auf mehrere Höhenprofile beinhaltet sie die Möglichkeit der Berücksichtigung gleich mehrerer, dem Spektrum überlagerter stehender Wellen bekannter Periode.⁴ Analog ergibt sich aus Gl. (51) für die statistischen Regularisierungsverfahren die Inversionsgleichung

$$\mathbf{f}'_{\gamma} = \begin{pmatrix} f_{\gamma 1} \\ \vdots \\ f_{\gamma m} \\ a \\ b \end{pmatrix} = \left[\mathbf{A}'^T \mathbf{S}_{\varepsilon}^{-1} \mathbf{A}' + \begin{pmatrix} \left(\begin{matrix} \mathbf{S}_{f_T}^{-1} & \mathbf{O} & \cdots \\ \mathbf{O} & \mathbf{S}_{f_{c_v}}^{-1} & \cdots \\ \vdots & \vdots & \ddots \\ 0 & \cdots & \\ 0 & \cdots & \end{matrix} \right) & \begin{matrix} 0 & 0 \\ \vdots & \vdots \\ 0 & 0 \\ 0 & 0 \end{matrix} \right]^{-1} \times \left[\mathbf{A}'^T \mathbf{S}_{\varepsilon}^{-1} \mathbf{g}_{\varepsilon} + \begin{pmatrix} \left(\begin{matrix} \mathbf{S}_{f_T}^{-1} \bar{\mathbf{f}}_T \\ \mathbf{S}_{f_{c_v}}^{-1} \bar{\mathbf{f}}_{c_v} \\ \vdots \\ 0 \\ 0 \end{matrix} \right) \right] \quad (55)$$

Die beiden Inversionsgleichungen (54) und (55) wurden im folgenden auch für die Auswertung der Meßdaten verwendet. Nicht realisiert werden konnte hingegen die

⁴Natürlich müssen hierzu der Vektor \mathbf{f}' , die Matrix \mathbf{A}' sowie die regularisierenden Matrizen bzw. Vektoren in (54) auf offensichtliche Art und Weise auf den Fall gleich mehrerer stehender Wellen erweitert werden.

Erweiterung des abgeschnittenen Singulärwertzerlegung auf den Fall der Berücksichtigung von stehenden Wellen aus den im Prinzip gleichen Gründen wie in Abschn. 4.1. Es war einfach nicht klar, wie bei diesem Verfahren eine nur teilweise Regularisierung der gestrichenen Lösung \mathbf{f}' realisiert werden könnte. Doch stellt bereits die Erweiterung der Tikhonov–Phillips–Regularisierung sowie der statistischen Regularisierungsverfahren auf die Inversion gleich mehrerer Spurengase und die Berücksichtigung von stehenden Wellen einen wichtigen Fortschritt dar.

Kapitel 5

Anwendung der erweiterten Inversionsverfahren

Wie schon Niple 1980 allgemeingültig feststellte, besteht das Standardverfahren zum Testen der Inversionsprogramme aus drei Schritten: der Inversion von unverrauschten synthetischen Daten, der Inversion von synthetischen Daten, denen ein Rauschen überlagert wurde und schließlich der Inversion von Feldmessungen [57]. Im Falle der Millimeterwellenradiometrie kommt häufig noch ein weiterer Aspekt hinzu, nämlich die Überlagerung der Spektren mit stehenden Wellen, welche von Mehrfachreflexionen innerhalb des Meßgerätes herrühren (siehe voriger Abschnitt). Das obige Testverfahren sollte deshalb um die beiden folgenden Schritte erweitert werden: die Inversion von unverrauschten synthetischen Spektren, zu denen stehende Wellen addiert wurden sowie die Inversion von synthetischen Daten, denen Rauschen und stehende Wellen überlagert wurden.

Eine Überprüfung der Inversionsverfahren für das gesamte Millimeterwellenspektrum wäre numerisch sehr aufwendig und würde den Rahmen dieser Arbeit sprengen. Statt dessen sollen nur einige wenige, repräsentative Bereiche aus dem Millimeterwellenspektrum herausgegriffen werden. Ohnehin interessieren primär diejenigen Spektralbereiche, welche mit den im IMK entwickelten bzw. sich gerade im Bau befindlichen Millimeterwellenradiometern gemessen werden können, da nur für diese Meßdaten verfügbar oder zu erwarten sind. Bei systematischen Messungen im Einsatz befand sich bisher nur das Ozon-Radiometer bei 142 GHz (siehe Abschn. 5.2.2). Der mit diesem Radiometer durch Verstimmen des lokalen Oszillators (LO) maximal erreichbare Spektralbereich erstreckt sich von 136 – 146,8 GHz.¹ In einer Einzelmessung erfaßt werden kann hiervon nur ein schmaler Ausschnitt der Bandbreite 1,2 GHz.² Inzwischen fertiggestellt wurde das ClO-Radiometer bei 278 GHz; es befindet sich

¹ Ob über diesen gesamten angegebenen Spektralbereich auch tatsächlich gemessen werden kann, hängt unter anderem ab vom spektralen Verlauf der Rauschtemperatur des Systems, da eine hohe Rauschtemperatur mit hohen Konversionsverlusten des Mischers einhergeht. Je höher diese sind, desto schlechter ist das resultierende Signal-zu-Rausch-Verhältnis. Zumindest liefert die angegebene Spanne jedoch einen Anhaltspunkt dafür, welcher Bereich des Spektrums mit dem Radiometer maximal erreichbar ist.

² Bei einer spektralen Auflösung des verwendeten akusto-optischen Spektrometers (AOS) von $\approx 0,7$ MHz und etwa 1700 Kanälen entspricht dies gerade der maximalen Empfangsbandbreite der

momentan in der Testphase. Mit ihm kann prinzipiell der Frequenzbereich zwischen 268,6 und 279,5 GHz eingestellt werden. Noch im Bau hingegen befindet sich das CIO-Radiometer bei 204 GHz. Mit diesem Radiometer sind die Spektralbereiche zwischen 200 und 208 GHz sowie zwischen 216 – 224 GHz erfaßbar. In den Abbildungen 38 – 40 in Anhang B.2 sind die für die drei Radiometer relevanten Spektralbereiche dargestellt. Sie wurden gerechnet für eine Blickrichtung senkrecht nach oben und eine Beobachterhöhe von 10 km. Diese wurde gewählt, um den Einfluß des troposphärischen Wasserdampfes auf die Spektren möglichst gering zu halten. Als Folge wären die Spektren sonst zu stark verzerrt worden und hätten ihren eigentlichen Zweck, nämlich einen Überblick über die relevanten Spektralbereiche zu gewähren, nicht mehr erfüllen können.

Im folgenden sollen anhand der Abbildungen 38 – 40 in Anhang B.2 die jeweils interessantesten Teilspektren der Bandbreite 1,2 GHz (dies entspricht, wie bereits erwähnt, der maximalen Empfangsbandbreite der verwendeten Radiometer) aus den relevanten Spektralbereichen ausgewählt und die Inversionsverfahren auf diese Konstellationen angewendet werden. Dabei wird davon ausgegangen, daß die Höhenprofile der nicht interessierenden Spurengase, soweit diese mit zum Spektrum beitragen, sowie die Höhenprofile von Druck und Temperatur bekannt sind. Dies gilt insbesondere für den troposphärischen Wasserdampf, welcher sich in Form eines breitbandigen Untergrundes durch das gesamte Millimeterwellenspektrum zieht. Diese Annahme ist höchst unrealistisch, da gerade der troposphärische Wasserdampf starken räumlichen und zeitlichen Schwankungen unterliegt (dies wird anhand der gemessenen Spektren der Abb. 24 in Abschn. 5.2.2 besonders deutlich). Doch soll im folgenden zunächst die Funktionstüchtigkeit der Inversionsalgorithmen überprüft werden, so daß eine Berücksichtigung des stark variierenden Wasserdampf-Untergrundes erst in Abschn. 5.2.1 erfolgt.

5.1 Inversion synthetischer Spektren

5.1.1 Unverrauschte synthetische Spektren

Unverrauschte, synthetisch erzeugte Spektren, welche mit Hilfe des Vorwärtsmodells aus vorgegebenen Volumenmischungsverhältnisprofilen sowie Druck- und Temperaturdaten generiert werden können, erlauben eine Überprüfung der Inversionsverfahren auf Selbstkonsistenz, da das Inversionsergebnis und das für die Berechnung der synthetischen Spektren verwendete Originalprofil übereinstimmen müssen. Wie bereits oben erwähnt, werden hierzu die Höhenprofile von Druck, Temperatur, sowie der

im IMK verwendeten Radiometer.

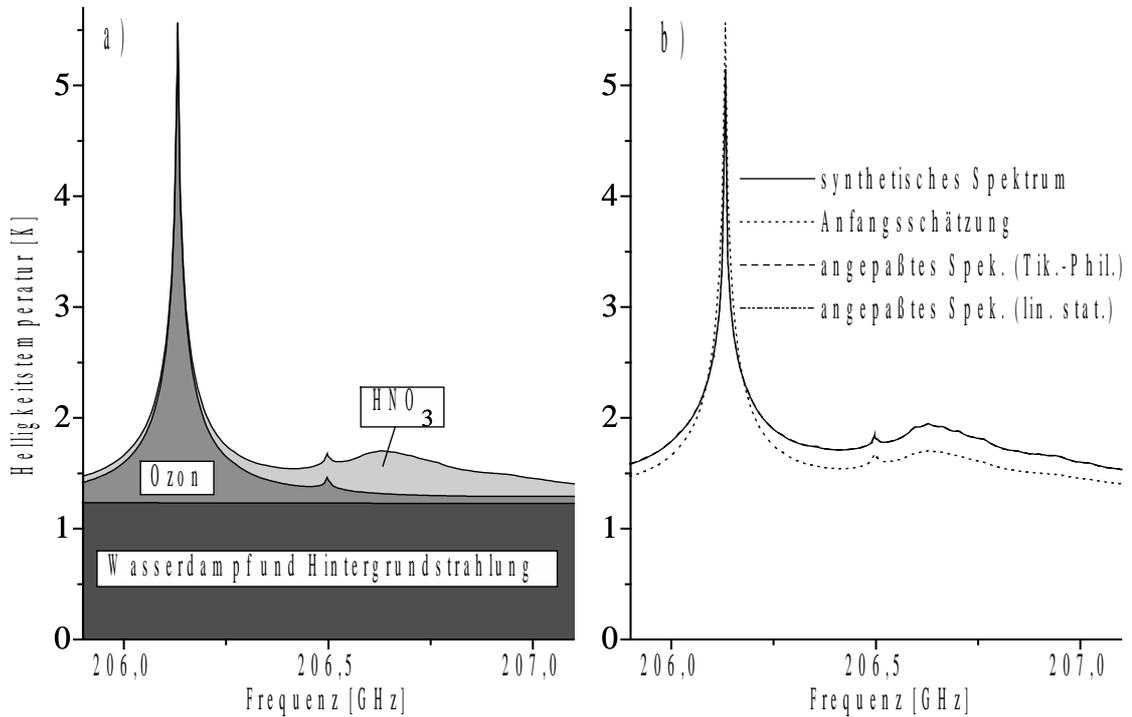


Abbildung 10: In a) Ausschnitt des Millimeterwellenspektrums in 10 km Höhe mit Blickrichtung senkrecht nach oben, gerechnet für eine Standardatmosphäre für mittlere Breiten.⁴ In b) Anpassung an ein unverraushtes, synthetisches Spektrum (durchgezogene, gestrichelte und gestrichpunktete Linie decken sich).

nicht interessierenden Spurengase, soweit sie zum Spektrum mit beitragen, als bekannt vorausgesetzt. Abweichungen des invertierten vom exakten Volumenmischungsverhältnisprofil sind deshalb in erster Linie auf numerisches Rauschen (endliche Rechengenauigkeit!) und geringe Interpolations-Inkonsistenzen zurückzuführen. Letztere treten auf, da das Vorwärts- und das Inversionsprogramm eine unterschiedliche Höhenschichteneinteilung verwenden.³

³Es hat sich gezeigt, daß für eine Anwendung der Tikhonov-Phillips-Regularisierung eine äquidistante Schichteneinteilung vorteilhaft ist. Natürlich sollten die Schichten dann entsprechend fein gewählt (≤ 2 km) und durch ihre Curtis-Godson-Mittelwerte repräsentiert werden (siehe Abschn. 2.1.3).

⁴Im Millimeterwellenbereich kann für die in der Atmosphäre herrschenden Temperaturen die Planckfunktion nach Rayleigh-Jeans angenähert werden durch

$$B_\nu(T) = 2k_B \frac{\nu^2}{c^2} T,$$

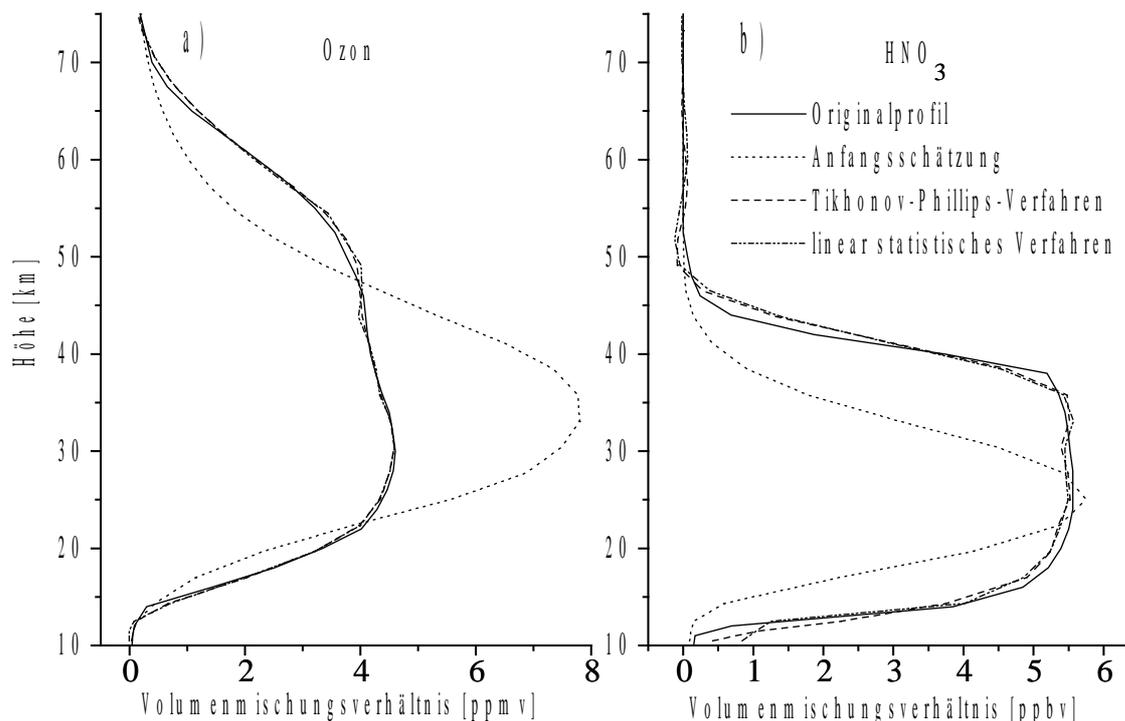


Abbildung 11: Gleichzeitige Inversion von a) Ozon und b) HNO_3 aus einem unverrauschten Spektrum.

Zum Testen der Inversionsverfahren wurde ein schmaler Ausschnitt des Millimeterwellenspektrums zwischen 205,9 und 207,1 GHz ausgewählt. Dieser ist Teil des mit dem ClO-Radiometer bei 204 GHz insgesamt erreichbaren Spektralbereichs (siehe Abb. 40) und zeichnet sich aus durch zwei relativ starke, eng benachbarte Signaturen von Ozon und HNO_3 . Beide Signaturen liegen innerhalb der Empfangsbandbreite des Radiometers von 1,2 GHz, so daß dieser Ausschnitt sich für eine gleichzeitige Inversion von Ozon- und HNO_3 -Volumenmischungsverhältnisprofilen geradezu anbietet. In Abb. 10 a ist dieser Bereich des Spektrums noch einmal vergrößert dargestellt.

Die Abb. 11 zeigt die unter diesen idealen Bedingungen (nur numerisches Rauschen bedingt durch die endliche Rechnergenauigkeit) erhaltenen Inversionsergebnisse. Den Ausgangspunkt der Simulation bildeten die beiden willkürlich angenommenen Höhenprofile von Ozon und HNO_3 (durchgezogene Linien) – im folgenden Original- oder

d. h. die Planckfunktion ist im wesentlichen proportional der Temperatur. Es bietet sich deshalb an, die SÜG auf beiden Seiten durch diesen Proportionalitätsfaktor zu dividieren. Die Strahldichte erhält damit die Dimension einer Temperatur, der sogenannten Helligkeitstemperatur. Im folgenden wird die Strahldichte deshalb jeweils in Kelvin (K) angegeben.

Referenzprofile genannt –, mit deren Hilfe zunächst ein synthetisches Strahldichtespektrum berechnet wurde. Dieses diente dann – stellvertretend für eine “ideale“ Messung – als Eingabe für die nachfolgenden Inversionsrechnungen und ist in Abb. 10 b dargestellt (durchgezogene Linie). Als Anfangsschätzung für die Inversionrechnungen wurden die beiden gepunkteten Linien in Abb. 11 verwendet, welche der Standardatmosphäre für mittlere Breiten entstammen und die zu dem gepunkteten Strahldichtespektrum in Abb. 10 b führen [13, 14]. Die gestrichelten und gestrichpunkteten Linien in Abb. 11 schließlich geben die mit dem Tikhonov–Phillips–Verfahren sowie dem linear statistischen Verfahren invertierten Höhenprofile wieder. Beide stimmen vorzüglich mit den Referenzprofilen überein. Dabei ist die Übereinstimmung bzgl. des Spurengases Ozon noch besser als für HNO_3 , was in erster Linie auf die deutlich schwächere Signatur von HNO_3 gegenüber derjenigen von Ozon zurückzuführen ist (siehe Abb. 10 a). Auch die Anpassung der Spektren an das vorgegebene synthetische Spektrum ist für beide Inversionsverfahren ganz ausgezeichnet (gestrichelte, gestrichpunktete sowie durchgezogene Linie decken sich in Abb. 10 b).⁵

An a priori–Information wurde für die Tikhonov–Phillips–Regularisierung verwendet, daß die Höhenprofile möglichst glatt sein sollten (siehe Gl. (36) in Abschn. 3.2.2), wobei die Wahl des Regularisierungsparameters γ a posteriori erfolgte (siehe Abschn. 3.1.4). Im Falle des linear statistischen Verfahrens dienten die beiden Anfangsschätzungen als a priori–Volumenmischungsverhältnisprofile und die a priori–Kovarianzmatrizen $\mathcal{S}_{f\text{O}_3}$, $\mathcal{S}_{f\text{HNO}_3}$ sowie die Fehlerkovarianzmatrix $\mathcal{S}_\mathcal{E}$ wurden als Diagonalmatrizen initialisiert.⁶ Die Diagonalelemente selbst wurden für die a priori–Kovarianzmatrizen mit den Werten 0 , $25 \cdot 10^{-12}$ und $1 \cdot 10^{-18}$ angenommen (was einer Standardabweichung der a priori–Höhenprofile für Ozon sowie HNO_3 von $\pm 0,5$ ppmv bzw. ± 1 ppbv entspricht) und für die Fehlerkovarianzmatrix $\mathcal{S}_\mathcal{E}$ mit dem sehr kleinen Wert $1 \cdot 10^{-16}$ K^2 initialisiert (dies entspricht einem in der Praxis sicherlich nie

⁵Dies war nicht anders zu erwarten, denn aus der Schlechtgestellttheit des Problems geht hervor, daß viele, wenn auch falsche Lösungen im Parameterraum das gemessene Spektrum mehr oder weniger exakt befriedigen. Andernfalls wäre keine Regularisierung erforderlich, da bereits alle nötige Information im Spektrum enthalten wäre. In diesem Falle lieferte die unrestringierte verallgemeinerte Inverse die richtige Lösung.

⁶Dies impliziert, daß die Volumenmischungsverhältnisse benachbarter Höhenschichten der a priori–Höhenprofile nicht korreliert sind. Wehr [91, S. 57 ff.] hat in seiner Diplomarbeit den Einfluß verschiedener a priori–Kovarianzmatrizen auf das Inversionsergebnis untersucht und festgestellt, daß bei einer zu starken Höhenkorrelation die a priori–Profile zu „steif“ sind und in ihrer Form nahezu unverändert auf das Inversionsergebnis abgebildet werden. Statt dessen empfiehlt er, keine Korrelation der Höhenschichten anzunehmen. Noch günstiger wäre natürlich die Verwendung einer a priori–Kovarianzmatrix, welche gemäß der Gl. (39) in Abschn. 3.2.3 aus einer großen Anzahl typischer, a priori bekannter Volumenmischungsverhältnisprofile berechnet werden könnte; doch stand eine solche Verteilung von a priori–Profilen leider nicht zur Verfügung. Was die Fehlerkovarianzmatrix angeht, so muß diese i. d. Regel sogar als eine reine Diagonalmatrix angenommen werden, da es sich in der Millimeterwellenradiometrie bei den Meßfehlern meist um rein statistische Fehler (numerisches Rauschen) handelt, welche nicht korreliert sind.

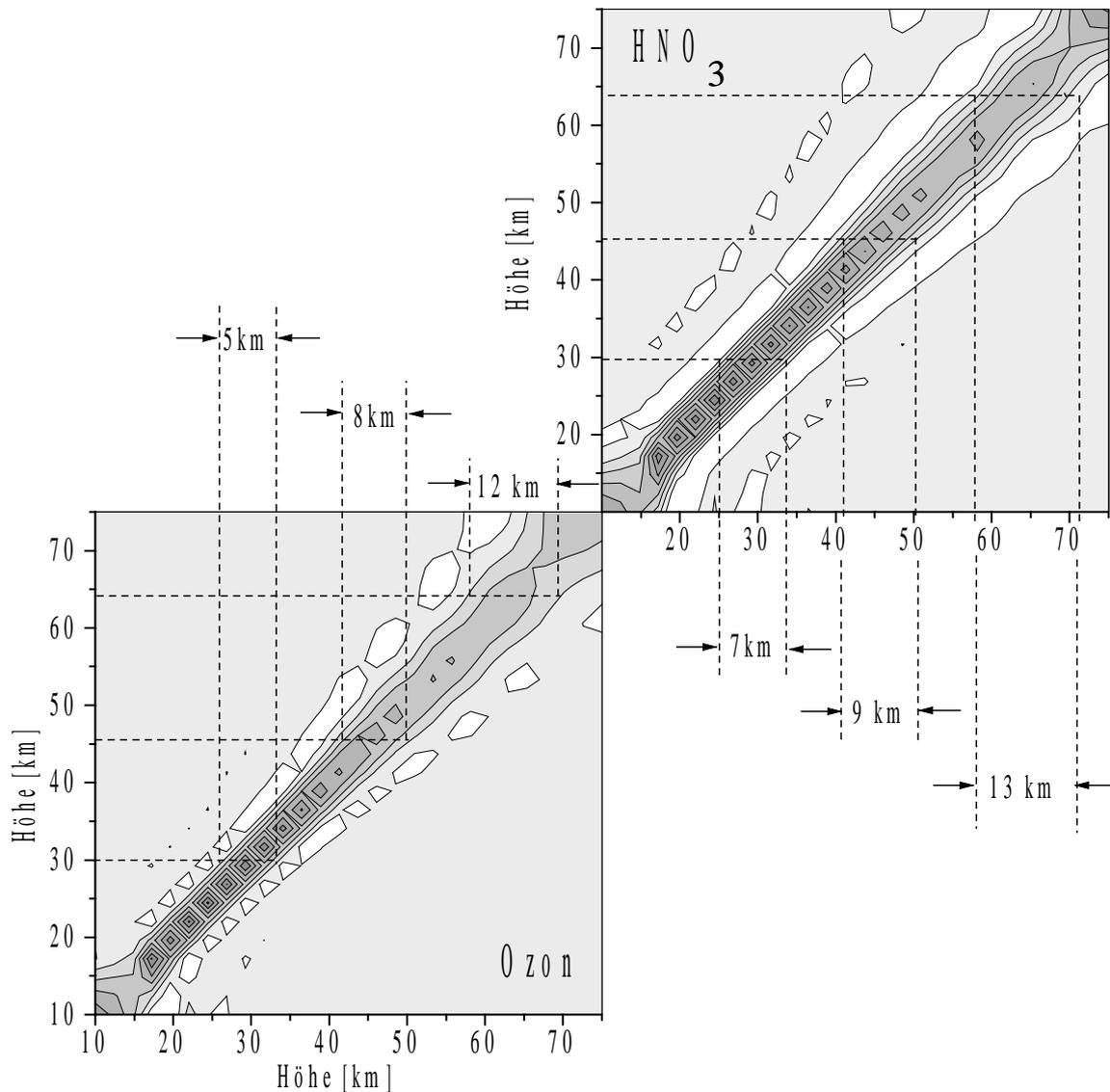


Abbildung 12: Auflösungsmatrix des Tikhonov-Phillips-Verfahrens bei einer gleichzeitigen Inversion der beiden Spurengase Ozon und HNO_3 aus einem unverrauschten synthetischen Spektrum (siehe hierzu die Ausführungen in Abschn. 3.1.5 auf Seite 26 ff.). Zu Darstellungszwecken wurden die Zeilen der Matrix in umgekehrter Reihenfolge angeordnet und die uninteressanten Bereiche der Matrix ausgespart. Offensichtlich kommt die Auflösungsmatrix der Einheitsmatrix sehr nahe. Eingezeichnet ist das Höhenauflösungsvermögen für beide Spurengase in 30, 46 und 63 km Höhe, dabei wurde die Halbwertsbreite des Bandes der Auflösungsmatrix als Kriterium verwendet. Begrenzt wird die Höhenauflösung in beiden Fällen letztlich nur durch das numerische Rauschen (endliche Rechnergenauigkeit) und geringe Interpolationsinkonsistenzen.

erreichbaren Meßfehlerausch von $\pm 1 \cdot 10^{-8}$ K und war notwendig, da diese Klasse von Inversionsverfahren für rauschfreie Messungen, wie sie hier betrachtet werden, formal nicht funktioniert).

Auf die Berechnung der Ergebniskovarianzmatrix \mathcal{S}_γ , die – wie in Abschn. 3.1.6 erwähnt – eine Aussage über den Fehler der Inversion liefert, wurde an dieser Stelle verzichtet, da sich im Falle unverrauschter Spektren ohnehin nur unrealistische Fehlerbalken ergeben hätten. Wesentlich aufschlußreicher ist hingegen die Auflösungsmatrix \mathbf{R} , welche bereits in Abschn. 3.1.5 als ein weiteres Kriterium für die Abschätzung der Güte einer Inversion vorgestellt wurde und die eine Art obere Schranke für die unter realistischen Bedingungen (endliches Meßfehlerausch) erreichbare Höhengauflösung liefert. Wie aus der Abb. 12 hervorgeht, ist diese für den betrachteten Spektralbereich am größten im Höhenbereich zwischen 25 und 35 km mit Werten von etwa 5 km für Ozon und 7 km für HNO_3 und nimmt nach oben und unten hin deutlich ab. Als Kriterium für die Ermittlung der Höhengauflösung wurde dabei die Halbwertsbreite des Bandes einer jeden Zeile der Auflösungsmatrix verwendet (näheres hierzu in einer früheren Arbeit [40, S. 37 ff.]).

Durch Testrechnungen ohne Rauschen wurden in allen Fällen gute Ergebnisse erzielt, d. h. die Originalprofile wurden auch bei deutlich abweichenden Anfangsschätzungen nahezu exakt reproduziert. Das Ziel, die Implementierung der Algorithmen zu validieren wurde damit erreicht. Grobe Fehler können ausgeschlossen werden.

5.1.2 Verrauschte synthetische Spektren

Verrauschte synthetische Spektren zeichnen ein bereits deutlich realistischeres Bild der bei der Auswertung von gemessenen Spektren anzutreffenden Verhältnisse. Wie bisher soll hierzu der Spektralbereich zwischen 205,9 und 207,1 GHz betrachtet werden, wobei den Spektren nun ein Gaußsches Rauschen der Standardabweichung 1/100 K bzw. 1/10 K überlagert wird (siehe Abb. 13). Da bereits bei einem Rauschen von 1/100 K Feinheiten der HNO_3 -Signatur nicht mehr zu erkennen sind, ist zu befürchten, daß sich die Qualität des zu invertierenden HNO_3 -Profils gegenüber derjenigen des unverrauschten Falles stark verringern wird. Bei einem Rauschen von 1/10 K geht die HNO_3 -Signatur gar vollständig im Rauschen unter, so daß fraglich ist, ob anhand dieses Spektrums das Spurengas HNO_3 überhaupt noch invertiert werden kann. Bzgl. Ozon sind in dieser Hinsicht sicherlich weniger Schwierigkeiten zu erwarten, da dessen Signatur sich doch deutlich gegen den Rauschuntergrund abhebt.

In Abb. 14 sind die Inversionsergebnisse der Tikhonov-Phillips-Regularisierung sowie des linear statistischen Verfahrens dargestellt, welche die obigen Erwartungen im wesentlichen bestätigen. Während Ozon in beiden Fällen (1/100 und 1/10 K Rauschen) relativ gut invertiert werden kann (Abb. 14 a und c) ergeben sich für HNO_3 doch mehr oder weniger große Abweichungen von dem Referenzprofil (siehe Abb. 14 c

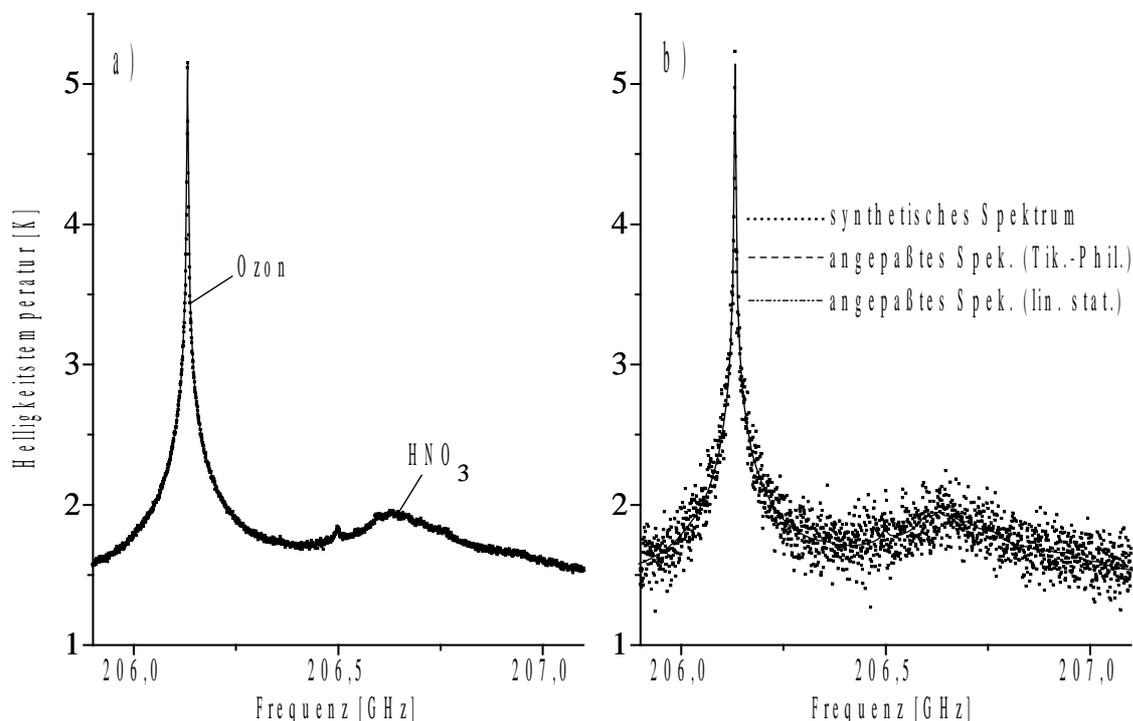


Abbildung 13: In a) mit einer Standardabweichung von $1/100$ K je Kanal der Bandbreite $0,7$ MHz verrauschtes Szenario, in b) mit einer Standardabweichung von $1/10$ K verrauschtes Spektrum. Die kompletten Anpassungen der Tikhonov–Phillips–Regularisierung sowie des linear statistischen Verfahrens sind jeweils mit dargestellt. Beide sind so genau, daß sie im Bild nicht zu unterscheiden sind.

und d). Insbesondere bei einem Rauschen von $1/10$ K (dies entspricht einem Signalzu-Rausch-Verhältnis ⁷ für HNO_3 von $5:1$ ⁸) lassen die Inversionsergebnisse sowohl der Tikhonov–Phillips–Regularisierung als auch des linear statistischen Verfahrens doch sehr zu wünschen übrig. Offensichtlich ist der Informationsgehalt des Spektrums für eine höhenauflösende Inversion von HNO_3 zu gering. (Natürlich ist auch ein schlecht aufgelöstes Höhenprofil wertvoll, sofern mit diesem eine relativ genaue integrierte Säulendichte verbunden ist).

⁷Unter Signalzu-Rausch-Verhältnis wird in diesem Zusammenhang das Verhältnis der maximalen Linienamplitude in der Mitte einer Signatur zu dem jeweils angegebenen Rauschen verstanden. Dieses wiederum ist bezogen auf eine Bandbreite von $0,7$ MHz je Frequenzkanal.

⁸Dies gilt, wie eingangs erwähnt, für eine Beobachterhöhe von 10 km. Am Boden würde das Signalzu-Rauschverhältnis bei unverändertem Rauschen noch ungünstiger ausfallen aufgrund der in der Troposphäre erfolgenden Absorption.

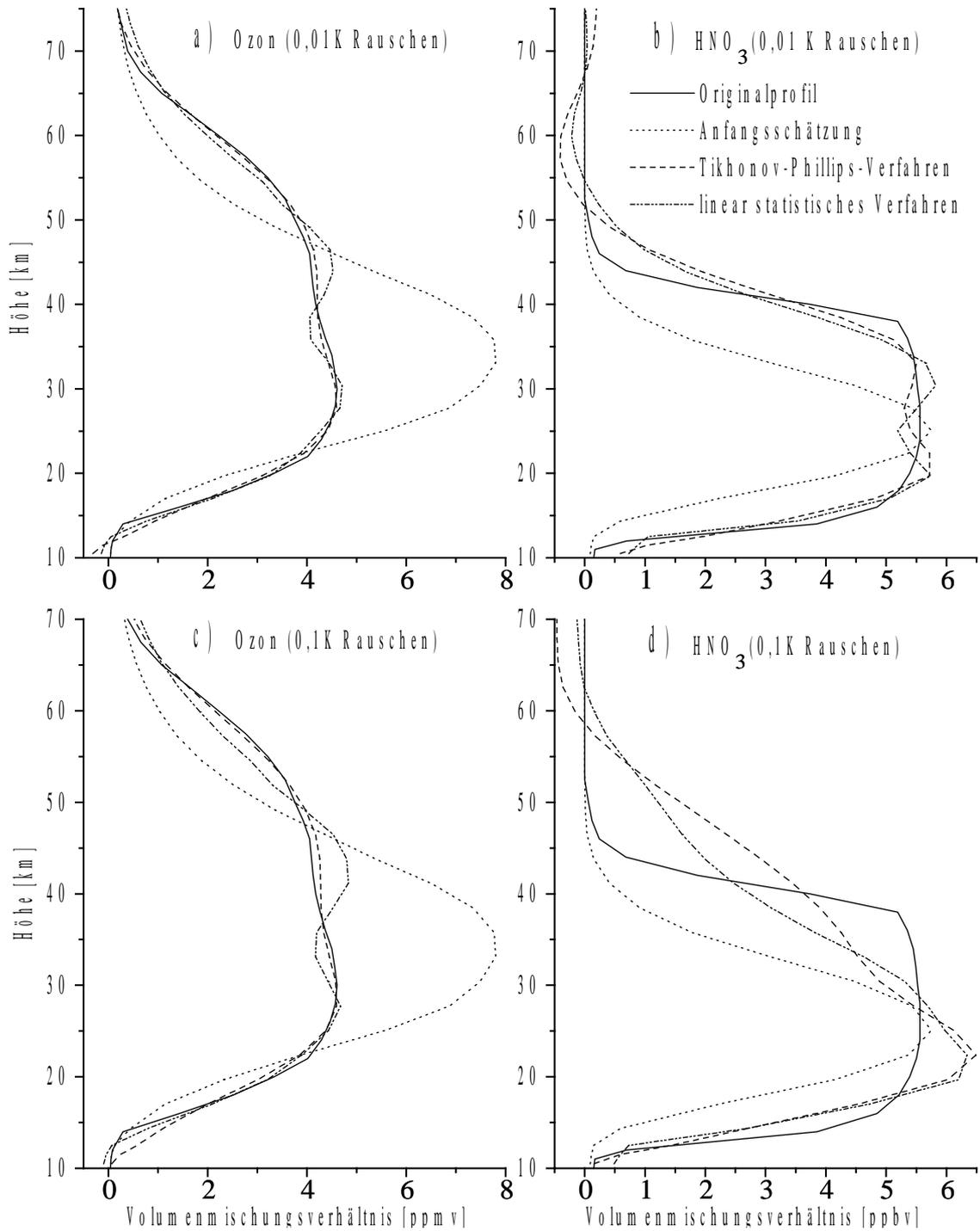


Abbildung 14: Gleichzeitige Inversion von Ozon und HNO₃ aus einem verrauschten Spektrum. In a) und b) 1/100 K Rauschen, in c) und d) 1/10 K Rauschen.

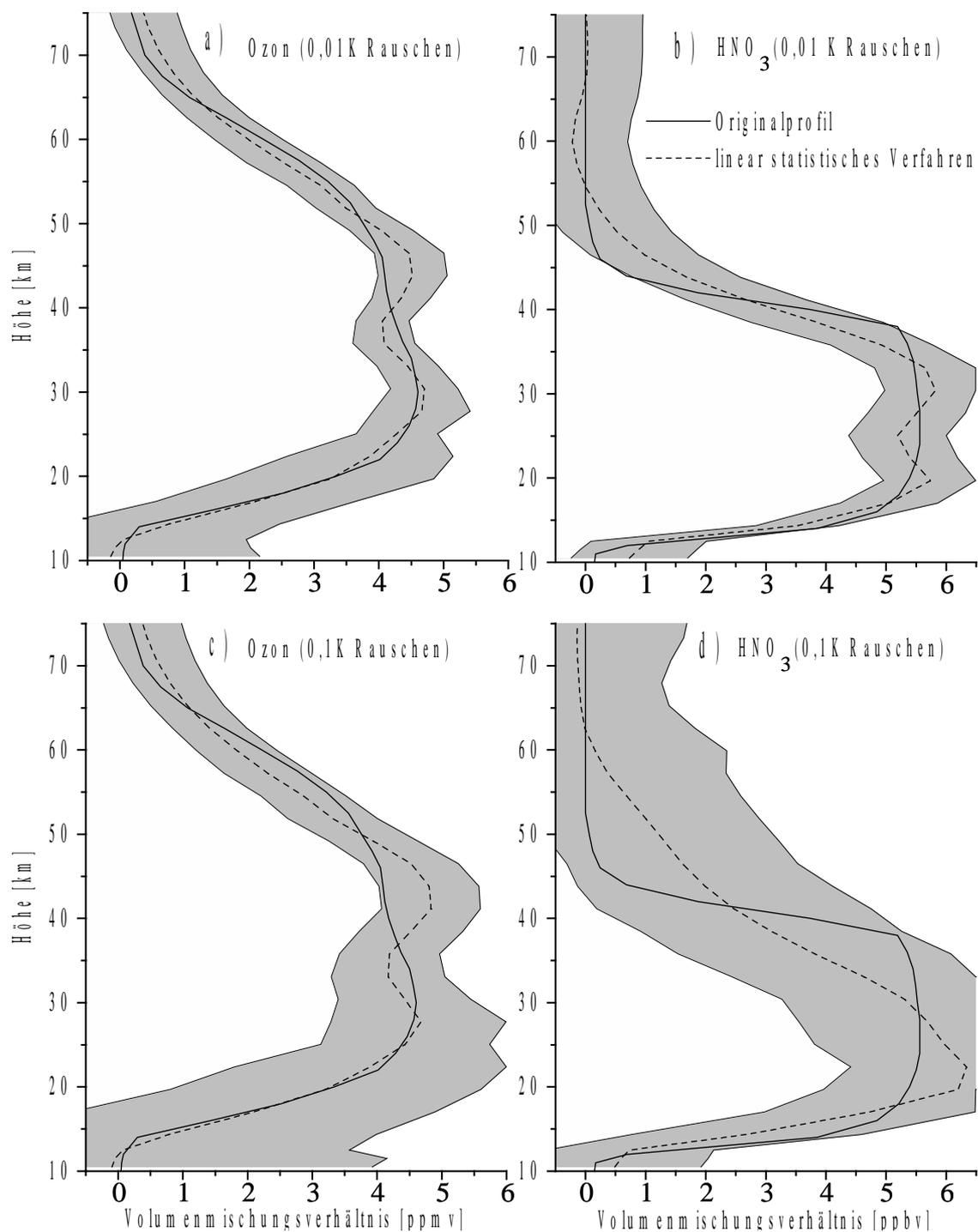


Abbildung 15: Fehlerbalken bei einer gleichzeitigen Inversion von Ozon und HNO₃ mit linear statistischen Verfahren. In a) und b) 1/100 K, in c) und d) 1/10 K Rauschen.

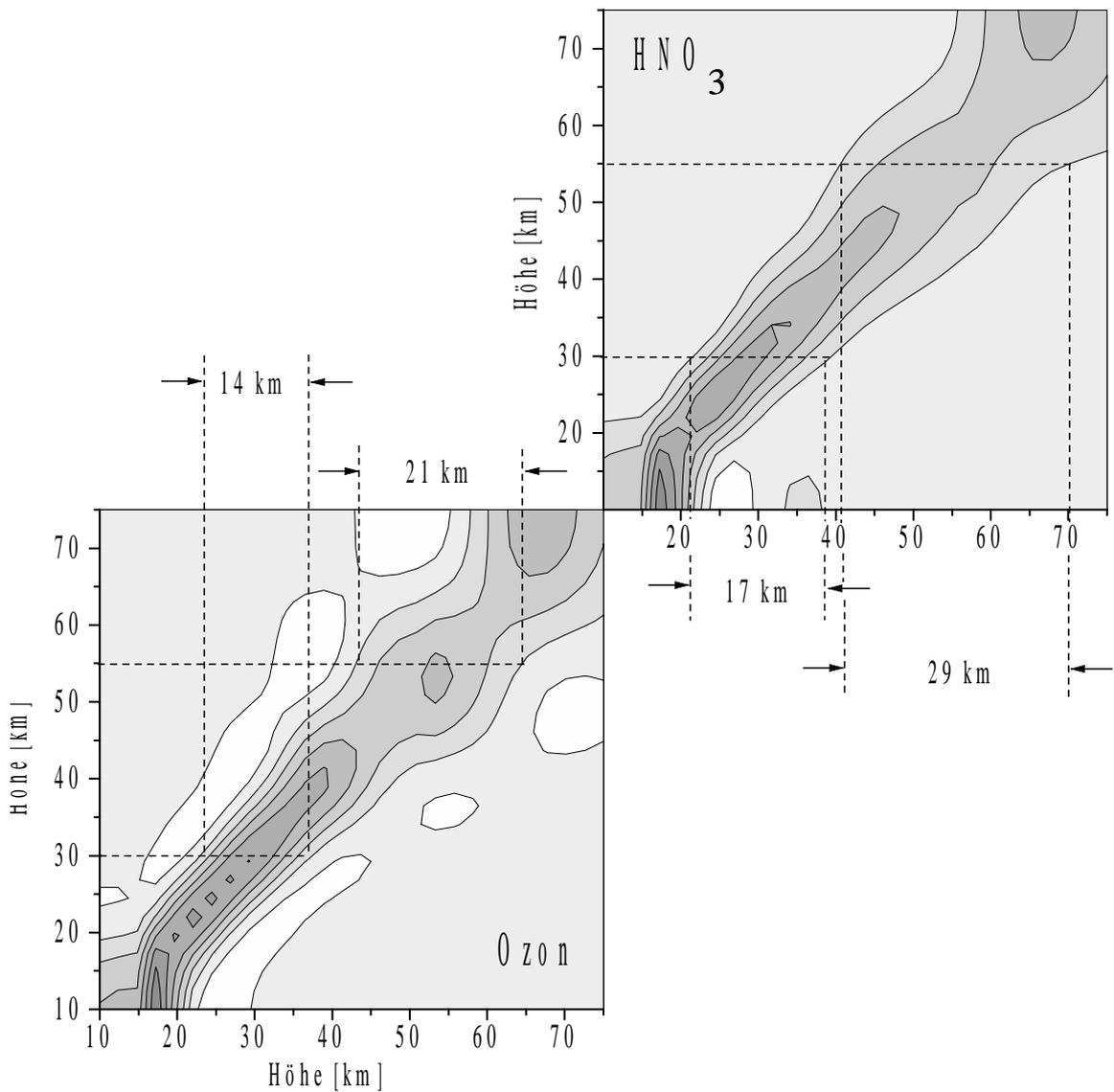


Abbildung 16: Auflösungsmatrix des Tikhonov-Phillips-Verfahrens bei einer gleichzeitigen Inversion von Ozon und HNO_3 aus einem mit $1/10$ K verrauschten synthetischen Spektrum. Eingezeichnet ist das Höhenauflösungsvermögen in einer Höhe von 30 und 55 km.

An a priori-Information wurde im Prinzip die gleiche wie im Falle des unverrauschten Spektrums verwendet; lediglich die Diagonalelemente der Fehlerkovarianzmatrix \mathcal{S}_ε wurden mit den Werten $(0,01 \text{ K})^2$ bzw. $(0,1 \text{ K})^2$ neu initialisiert. Die Abb. 15 zeigt die mit diesen Eingabedaten für das lineare statistischen Verfahren über eine Berechnung der Ergebniskovarianzmatrix \mathcal{S}_γ erhaltenen Fehlerbalken (siehe Gl. (28))

in Abschn. 3.1.6).⁹ Wie aus der Abbildung hervorgeht, wachsen diese mit zunehmendem Rauschen stark an und sind für HNO_3 erheblich größer als für Ozon aufgrund des ungünstigeren Signal-zu-Rausch-Verhältnisses. Dabei bewegt man sich mit dem Versuch der Ableitung des HNO_3 -Profils aus dem mit 1/10 K verrauschten Spektrum offensichtlich in einer Art Grenzbereich, unterhalb dessen eine Anwendung von höhenauflösenden Inversionsverfahren kaum mehr sinnvoll erscheint: zum einen gibt nämlich das invertierte HNO_3 -Profil das Referenzprofil nur unzulänglich wieder und zum anderen wäre bei einer weiteren Verschlechterung des ohnehin ungünstigen Signal-zu-Rausch-Verhältnisses aufgrund der stark anwachsenden Fehlerbalken im Prinzip überhaupt keine Aussage über den Höhenverlauf von HNO_3 mehr möglich (siehe Abb. 14 d). Kleiner als 5:1 darf das Signal-zu-Rauschverhältnis daher keinesfalls werden, soll überhaupt noch eine Höhenauflösung von Spurengasen möglich sein. Diese Auffassung wird durch die zugehörige Auflösungsmatrix im wesentlichen bestätigt (siehe Abb. 16). Kaum mehr als zwei Höhenstufen können für HNO_3 noch aufgelöst werden. Unter diesen Umständen wäre es vielleicht sogar sinnvoll gewesen, für HNO_3 von vornherein mit einem Zwei-Schichten-Modell zu arbeiten und dieses bei der Initialisierung der a priori-Kovarianzmatrix in geeigneter Form zu berücksichtigen; eventuell hätte man auch die Form des Höhenprofils gleich ganz vorschreiben und dieses anschließend für die Anpassung an das gemessene Spektrum entsprechend skalieren können. Doch sind dies alles Feinheiten, welche hier nicht allgemeingültig diskutiert werden können, sondern jeweils von Fall zu Fall entschieden werden müssen. Als Quintessenz der Untersuchungen in diesem Abschnitt soll jedoch festgehalten werden, daß für eine einigermaßen höhenauflösende Inversionrechnung (d. h. Unterscheidung wenigstens zweier Höhenschichten) das Signal-zu-Rausch-Verhältnis mindestens 5:1 betragen muß.

5.1.3 Unverrauschte synthetische Spektren mit stehenden Wellen

Im folgenden soll der Einfluß von stehenden Wellen auf die Qualität der zu invertierenden Höhenprofile näher untersucht werden. Zu diesem Zweck werden dem bisher betrachteten Spektrum drei stehende Wellen unterschiedlicher Periode, Amplitude und Phase überlagert (siehe Abb. 17 a). In Abschn. 4.2 über die Grundlagen zur Berücksichtigung von stehenden Wellen wurden die Perioden der überlagerten stehenden Wellen als bekannt vorausgesetzt. Dies bereitete insofern keine Probleme, als die Perioden der stehenden Wellen leicht über die Fabry-Pérot-Formel aus der Länge der zugehörigen Resonatoren ermittelt werden können, nachdem diese innerhalb des

⁹Im Gegensatz zu allen anderen Regularisierungsverfahren, welche nur Aussagen über den Datenfehler der regularisierten Lösung erlauben, ermöglichen die statistischen Verfahren eine Abschätzung des Gesamtfehlers der Lösung – wenngleich diese Angaben über den Fehler mit einer gewissen Skepsis zur Kenntnis zu nehmen sind (siehe die Ausführungen in Abschn. 3.1.6).

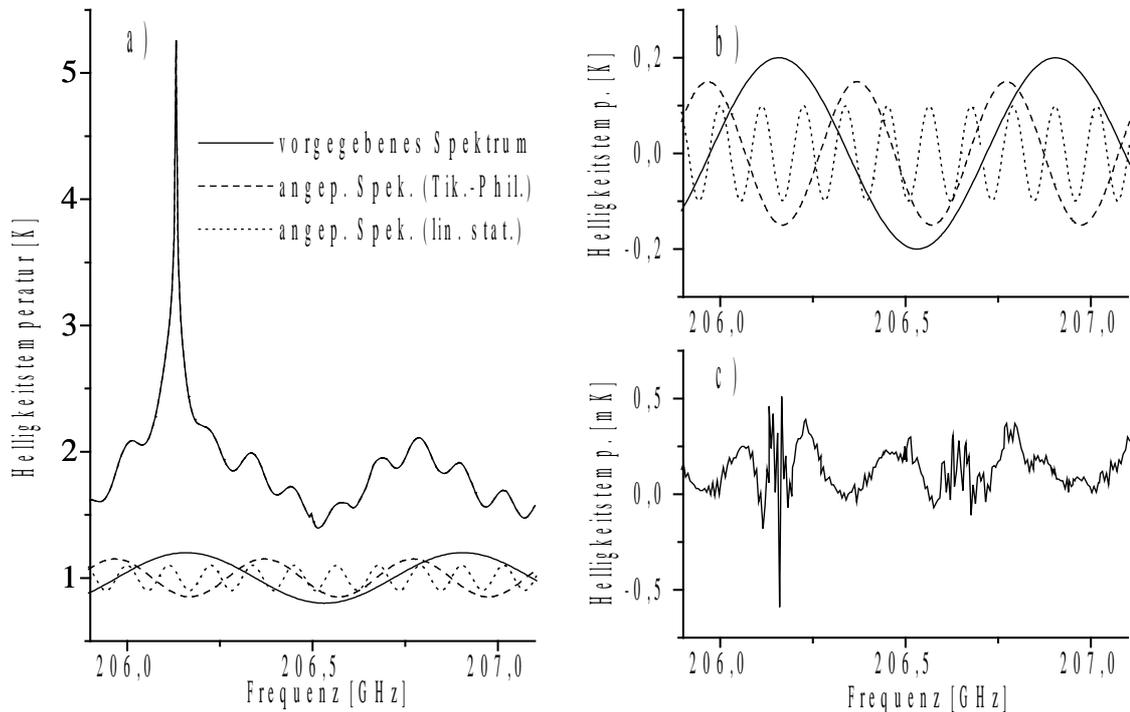


Abbildung 17: In a) unverraushtes synthetisches, mit stehenden Wellen überlagertes Spektrum zusammen mit den kompletten Anpassungen der Tikhonov–Phillips–Regularisierung sowie des linear statistischen Verfahrens (Linien überdecken sich). Die überlagerten stehenden Wellen sind unten maßstabsgetreu (zu Darstellungszwecken um ein Kelvin nach oben versetzt) mit dargestellt. Abb. b zeigt die stehenden Wellen noch einmal etwas vergrößert und Abb. c die Differenz zwischen dem vorgegebenen Spektrum und der kompletten Anpassung der Tikhonov–Phillips–Regularisierung an das vorgegebene Spektrum (man beachte die Skalierung der Ordinate in mK).

Radiometers erst einmal identifiziert wurden. Entsprechend sollen auch hier in einem ersten Test die Perioden der stehenden Wellen bei der Inversion als bekannt vorausgesetzt werden. In weiteren Versuchen soll dann überprüft werden, inwieweit sich geringfügige Abweichungen der bei der Inversion verwendeten Perioden von den Perioden der dem Spektrum überlagerten stehenden Wellen auf die Qualität der Inversionsergebnisse auswirken.

In Tabelle 1 sind die Perioden, Amplituden und Phasen der dem synthetischen Spektrum überlagerten stehenden Wellen zusammengefaßt. Sie wurden in Anlehnung an die bei realen Messungen zu beobachtenden stehenden Wellen (vgl. Abschn. 5.2.2) willkürlich vorgegeben. Die Abb. 17 a zeigt die kompletten Anpassungen

Nr.	Periode GHz	Ampl. K	Phase Grad	Ampl.* K	Phase* Grad
1.	0,7465	0,2	30	0,200	29,94
2.	0,403	0,15	60	0,150	59,93
3.	0,113	0,1	90	0,100	90,01

Tabelle 1: Liste der dem synthetischen Spektrum zusätzlich überlagerten stehenden Wellen. Die beiden mit einem Stern gekennzeichneten Spalten geben die über eine Inversion des Spektrums ermittelten Werte von Amplitude und Phase wieder und stimmen sehr gut mit den vorgegebenen Werten überein.

der um die Berücksichtigung von stehenden Wellen erweiterten Tikhonov–Phillips-Regularisierung, sowie des entsprechend erweiterten linear statistischen Verfahrens an das vorgegebene Spektrum; dabei wurde an a priori-Information wieder die des ungestörten Spektrums aus Abschn. 5.1.1 verwendet. Wie bereits im Falle des gänzlich ungestörten Spektrums ist auch hier die Übereinstimmung zwischen Anpassung und Spektrum ausgezeichnet (durchgezogene, gestrichelte und gepunktete Linie decken sich in Abb. 17 a). In Abb. 17 c ist deshalb zusätzlich die Differenz zwischen dem vorgegebenen Spektrum und der Anpassung der Tikhonov–Phillips-Regularisierung mit aufgetragen. Da letztere praktisch keine stehenden Wellen mehr enthält (man beachte die Skalierung der Ordinate in mK), konnten diese offensichtlich vollständig eliminiert werden, wobei – wie bereits mehrfach erwähnt – die Perioden der dem Spektrum überlagerten stehenden Wellen bei der Inversion als bekannt vorausgesetzt wurden. Von hoher Qualität sind auch die invertierten Höhenprofile (vergleichbar mit der des ungestörten Spektrums, siehe Abb. 11), weshalb auf deren Darstellung verzichtet wurde; dasselbe gilt für die Auflösungsmatrix \mathbf{R} . Angegeben sind jedoch die im Zuge der Inversion des Spektrums ermittelten Amplituden und Phasen der drei stehenden Wellen (siehe die mit einem Stern versehenen Spalten in Tab. 1), welche hervorragend mit den Amplituden und Phasen der dem synthetischen Spektrum überlagerten stehenden Wellen übereinstimmen.

Um den Einfluß einer Abweichung der bei der Inversion des Spektrums verwendeten Perioden von den Perioden der dem Spektrum überlagerten stehenden Wellen näher untersuchen zu können, sollen die stehenden Wellen künftig isoliert betrachtet werden; d. h. dem synthetischen Spektrum wird im folgenden nur noch jeweils eine stehende Welle überlagert werden. Interessant sind hierbei vor allen Dingen die beiden Spezialfälle der Überlagerung des Spektrums mit einer stehenden Welle relativ kurzer Periode (Periode klein gegen die Linienbreite) sowie der Fall der Überlagerung des Spektrums mit einer stehenden Welle relativ langer Periode (Periode größer oder vergleichbar der Linienbreite). Zunächst zum Fall der Überlagerung des synthetischen Spektrums mit einer stehenden Welle kurzer Periode – 0,113 GHz – dies entspricht

der stehenden Welle 3 in Tab. 1 und ist dargestellt in Abb. 18 a (durchgezogene Linie). Für die Inversion dieses Spektrums wurde eine stehende Welle der Periode 0,115 GHz angenommen, was einer relativen Abweichung von 2% entspricht.¹⁰

Das Ergebnis der mit Hilfe der Tikhonov–Phillips–Regularisierung erreichten Anpassung an das vorgegebene Spektrum ist in Abb. 18 a gezeigt (gestrichelte Linie). Auf den ersten Blick stimmen beide gut überein, doch treten die Abweichungen bereits deutlich stärker hervor, sobald man Spektrum und Anpassung um die gefittete stehende Welle der Periode 0,115 GHz korrigiert (Abb. 18 b); offensichtlich konnte die dem synthetischen Spektrum überlagerte stehende Welle nur teilweise eliminiert werden, da insbesondere an den beiden Rändern noch eine leichte Restwelligkeit zu erkennen ist. In Abb. 18 c ist deshalb zusätzlich die Differenz zwischen dem vorgegebenen Spektrum und der kompletten Anpassung der Tikhonov–Phillips–Regularisierung aufgetragen. Letztere ist in der Mitte des Spektrums relativ klein ($\approx 0,02$ K) – ein Hinweis darauf, daß in diesem Bereich die Anpassung relativ gut funktioniert – und wird nach außen hin deutlich schlechter mit Abweichungen bis zu 0,06 K. Die Amplitude der dem synthetischen Spektrum überlagerten stehenden Welle betrug zum Vergleich 0,1 K.¹¹ In Abb. 18 d ist weiterhin der Fehler aufgetragen, welcher sich theoretisch aufgrund einer Abweichung der Periode der gefitteten stehenden Welle von der Periode der vorgegebenen stehenden Welle ergeben sollte: dieser entspricht einfach der Differenz der beiden stehenden Wellen. Bis auf ein paar kleine Abweichungen im Bereich der Ozon–Signatur stimmt dessen spektraler Verlauf sehr gut mit dem Differenzspektrum der Abb. 18 c überein, so daß man zumindest für Ozon eine einigermaßen gute Übereinstimmung des invertierten Höhenprofils mit dem Referenzprofil erwarten würde – bzgl. HNO_3 ist in dieser Hinsicht sicherlich keine Aussage möglich, da dessen Signatur zu schwach ist.

Tatsächlich geben die Inversionsergebnisse der Tikhonov–Phillips–Regularisierung sowie des linear statistischen Verfahrens den Verlauf des Referenzprofils von Ozon relativ gut wieder (siehe Abb. 19 a). Dagegen weichen die invertierten HNO_3 –Profile stark von dem Referenzprofil der Abb. 19 b ab. Alles in allem sind die Inversionsergebnisse jedoch sehr ähnlich denen des mit 0,1 K verrauschten synthetischen Spektrums (vgl. Abb. 19 mit den entsprechenden Abbildungen 14 c und d), was auch plausibel erscheint, da eine nicht eliminierte stehende Welle relativ kurzer Periode sich im Prinzip so auswirken sollte, als wäre das Spektrum stark verrauscht. Dementsprechend beschaffen ist auch die Auflösungsmatrix \mathbf{R} , weshalb auf deren Darstellung an dieser Stelle verzichtet wurde (siehe statt dessen die Abb. 16).

¹⁰Als a priori–Information dient weiterhin die des ungestörten Spektrums aus Abschn. 5.1.1.

¹¹Normalerweise würde man aus diesem Verlauf der Differenzspektrums schließen, daß die für die Inversion verwendete Periode nicht optimal war und würde in einem zweiten Versuch die Periode der anzupassenden stehenden Welle entsprechend variieren. Im vorliegenden Fall sollen jedoch gerade die Auswirkungen einer falschen Annahme der Periode der stehenden Welle auf die Inversionsergebnisse näher untersucht werden.

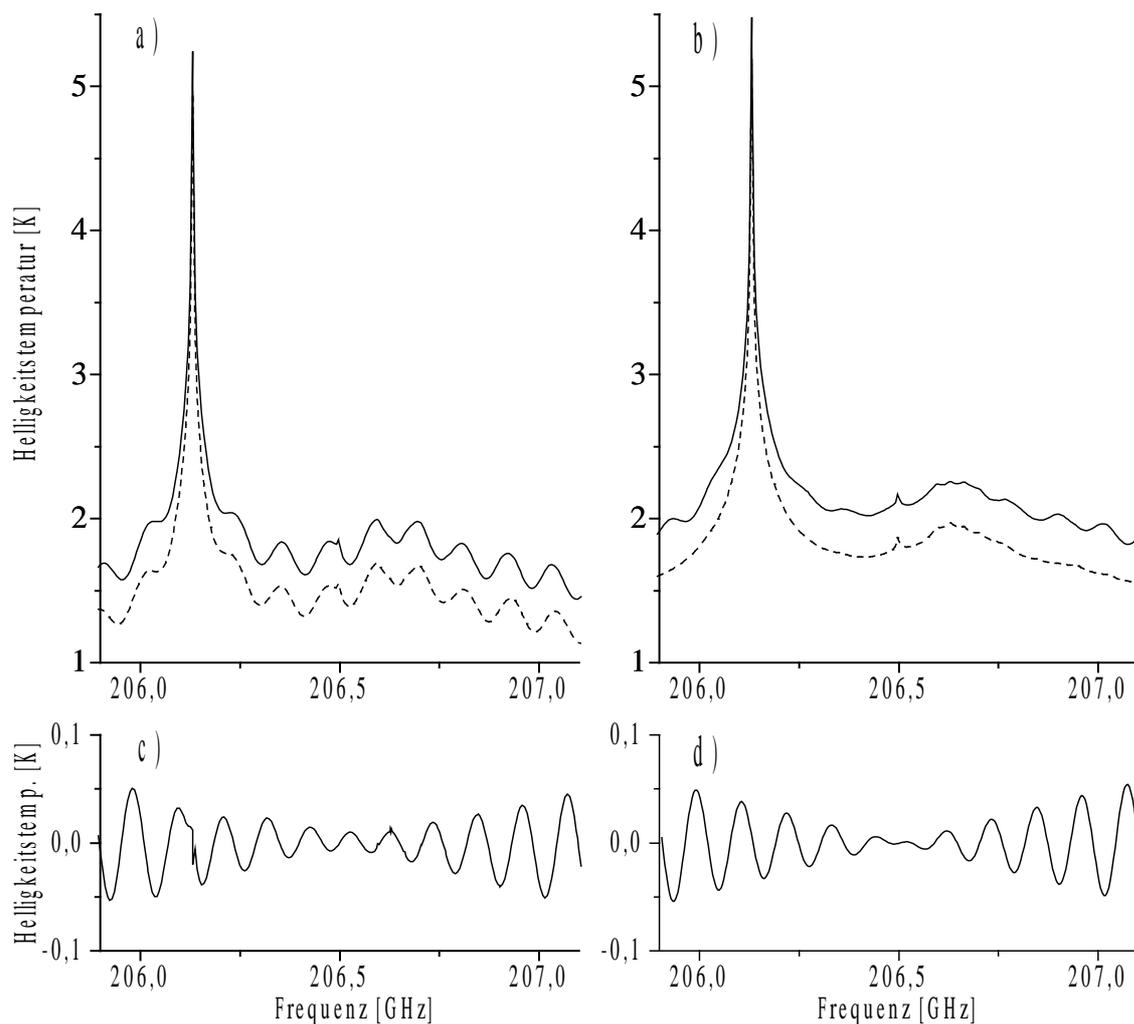


Abbildung 18: In a) vorgegebenes, mit einer stehenden Welle der Periode 0,113 GHz und der Amplitude 0,1 K überlagertes Spektrum (durchgezogen) sowie die komplette Anpassung der Tikhonov–Phillips–Regularisierung bei Verwendung einer stehenden Welle der Periode 0,115 GHz (gestrichelt und zu Darstellungszwecken um 0,3 K nach unten versetzt). In b) um die gefittete stehende Welle der Periode 0,115 GHz korrigierte Spektren. In c) Differenz zwischen dem vorgegebenen Spektrum und der kompletten Anpassung sowie in d) aufgrund der Abweichung der Perioden der stehenden Wellen theoretisch zu erwartende Abweichung zwischen Spektrum und kompletter Anpassung (entspricht einfach der Differenz obiger zweier stehenden Wellen).

Von den Inversionsergebnissen des verrauschten Spektrums deutlich unterscheiden

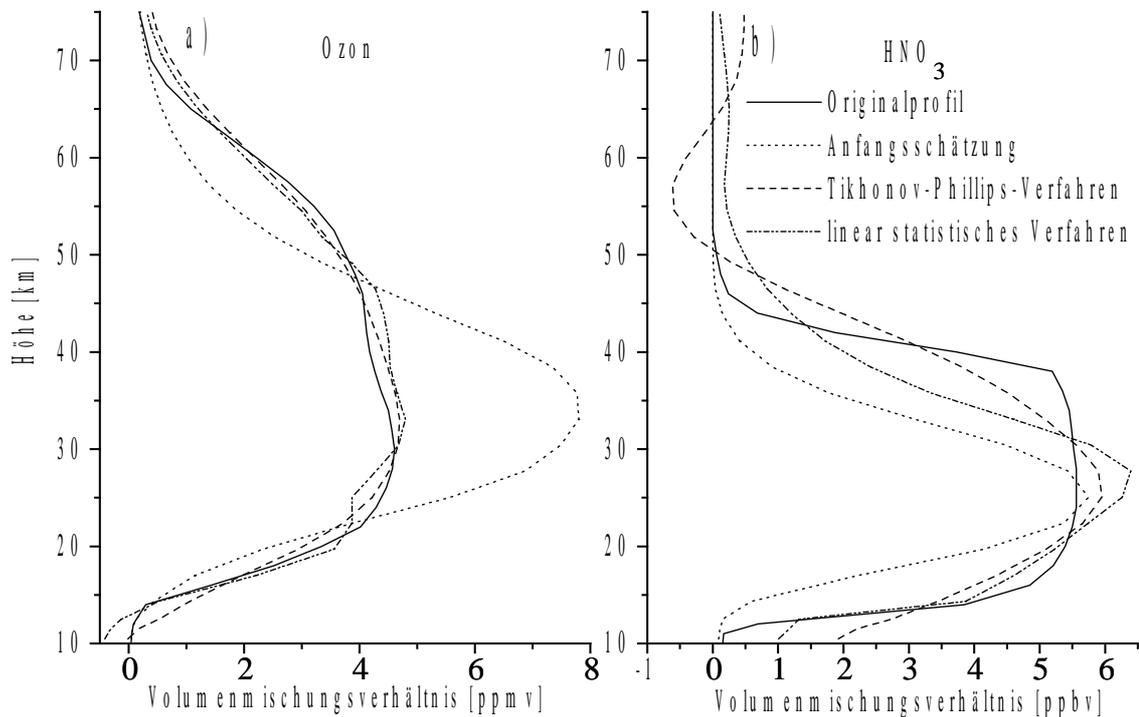


Abbildung 19: Gleichzeitige Inversion von a) Ozon und b) HNO₃ aus einem unverrauschten, mit einer stehenden Welle der Periode 0,113 GHz überlagerten Spektrum. Zur Inversion wurde eine stehende Welle der Periode von 0,115 GHz verwendet.

sollten sich die Resultate eines mit einer stehenden Welle relativ langer Periode überlagerten Spektrums. Um dieses zu überprüfen, wird dem bisher betrachteten Spektrum eine stehende Welle der Periode 0,7465 GHz und Amplitude 0,2 K überlagert (dies entspricht der stehenden Welle 1 aus Tabelle 1). Ansonsten bleiben die Rahmenbedingungen die gleichen wie bisher, lediglich die Inversion wird mit einer stehenden Welle veränderter Periode – 0,76 GHz – durchgeführt, was einer relativen Abweichung von 2% entspricht. In Abb. 20 a ist das sich hieraus ergebende Spektrum zusammen mit der kompletten Anpassung der Tikhonov–Phillips–Regularisierung dargestellt. Beide stimmen gut überein, besser noch als im Falle des mit einer stehenden Welle relativ kurzer Periode überlagerten Spektrums, bei dem nach der Korrektur des vorgegebenen Spektrums um die gefittete stehende Welle noch eine gewisse Restwelligkeit zu erkennen war (vgl. die Abb. 20 b mit der Abb. 18 b). In Abb. 20 c ist deshalb zusätzlich die Differenz zwischen Spektrum und Anpassung aufgetragen. Letztere ist erheblich kleiner als die entsprechende Differenz bei dem mit einer stehenden Welle kurzer Periode überlagerten Spektrum (vgl. mit der Abb. 18 c), obwohl die Amplitude der langen stehenden Welle mit 0,2 K doppelt so groß war als die der kurzen mit 0,1 K

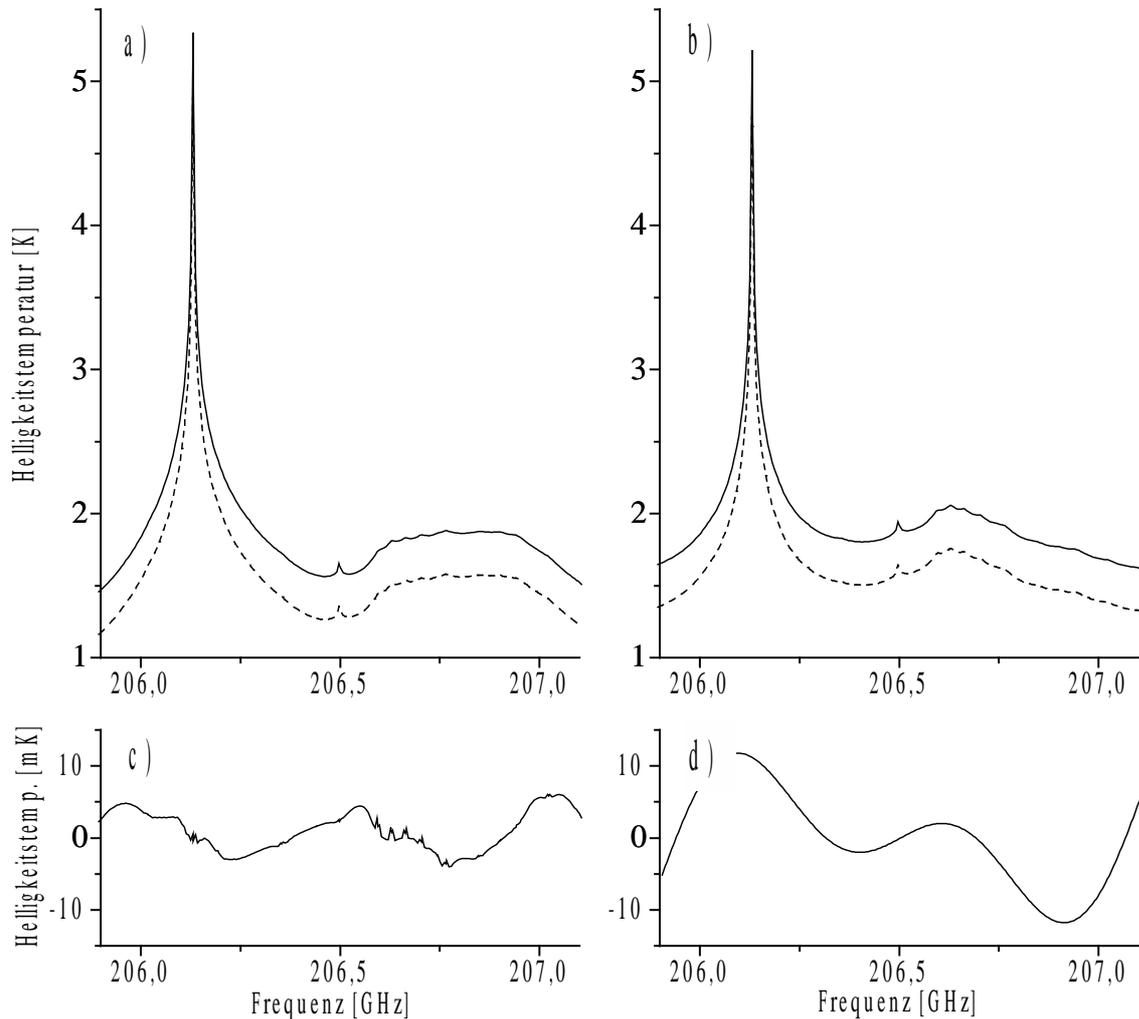


Abbildung 20: In a) vorgegebenes, mit einer stehenden Welle der Periode 0,7465 GHz und der Amplitude 0,2 K überlagertes Spektrum (durchgezogen) sowie die komplette Anpassung der Tikhonov-Phillips-Regularisierung bei Verwendung einer stehenden Welle der Periode 0,76 GHz (gestrichelt und zu Darstellungszwecken um 0,3 K nach unten versetzt). In b) um die gefittete stehende Welle der Periode 0,76 GHz korrigierte Spektren. In c) Differenz zwischen dem vorgegebenen Spektrum und der kompletten Anpassung (man beachte die Auftragung der Ordinate in mK) sowie in d) aufgrund der Abweichung der Perioden der stehenden Wellen im Grunde zu erwartende Differenz zwischen dem Spektrum und der kompletten Anpassung (Differenz obiger zweier stehenden Wellen).

(man beachte die unterschiedliche Skalierung der beiden Ordinaten). Vergleicht man die Differenz weiterhin mit dem Fehler, welcher sich theoretisch aufgrund der Abweichung der Periode der gefitteten stehenden Welle von der Periode der vorgegebenen

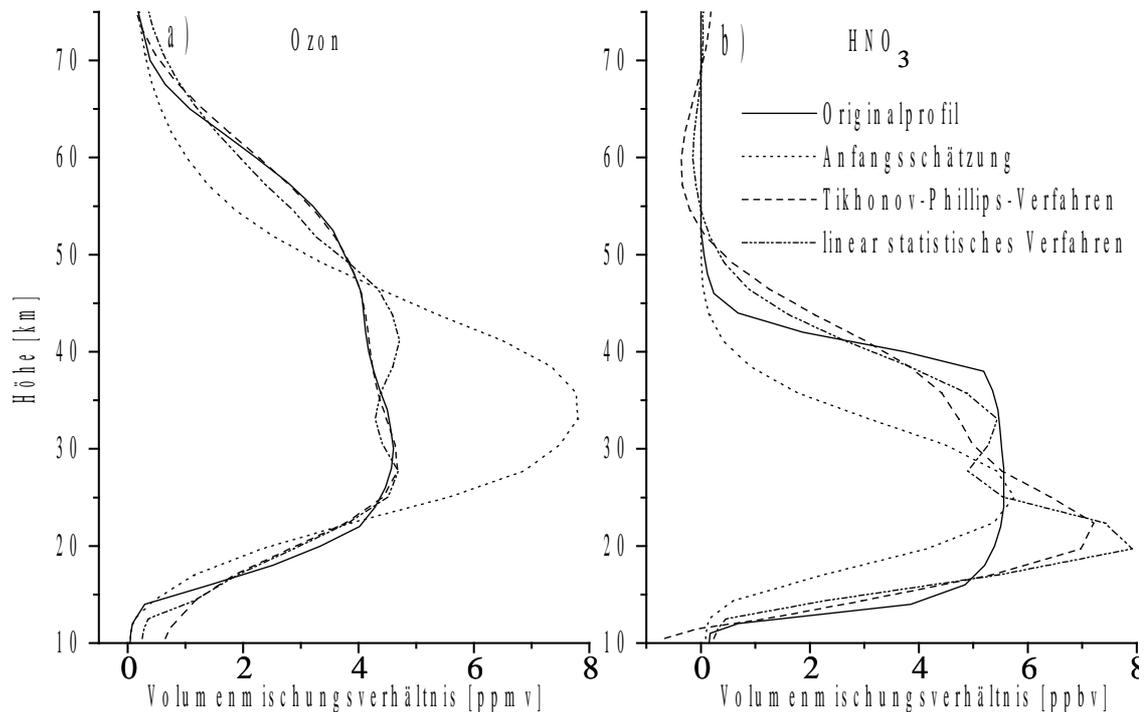


Abbildung 21: Gleichzeitige Inversion von a) Ozon und b) HNO_3 aus einem unverrauschten, mit einer stehenden Welle der Periode $0,7465 \text{ GHz}$ überlagerten Spektrum. Zur Inversion wurde eine stehende Welle der Periode $0,76 \text{ GHz}$ verwendet.

stehenden Welle ergeben sollte (dieser entspricht wieder der Differenz der beiden stehenden Wellen, siehe Abb. 20 d), so ist dieser ebenfalls größer. Die dem Spektrum überlagerte stehende Welle muß im Laufe der Inversion also falsch angepaßt worden sein, was nur auf Kosten der Qualität der erzielten Inversionsergebnisse erfolgt sein kann (siehe Abb. 21).

Vergleicht man die Abb. 20 d mit der Abb. 20 b, so wird deutlich, wie diese falsche Anpassung zustande kam. Die Signaturen von O_3 und HNO_3 liegen nämlich genau über den Maxima der Abb. 20 d, so daß durch eine Skalierung der O_3 - und HNO_3 -Profile ein auf die Abweichung der Perioden der beiden stehenden Wellen zurückzuführender Fehler leicht kompensiert werden kann.

In Abb. 21 sind die Inversionsergebnisse der Tikhonov-Phillips-Regularisierung sowie des linear statistischen Verfahrens dargestellt. Beide Verfahren geben das Referenzprofil von Ozon relativ gut wieder mit geringen Vorteilen für die Tikhonov-Phillips-Regularisierung. Für HNO_3 hingegen ergeben sich bei beiden Verfahren deutliche Abweichungen gegenüber dem Referenzprofil mit einem ausgeprägten Peak in 20 km Höhe. In diesem Höhenbereich erreicht die Linienbreite Werte zwischen $0,2$ und

0,3 GHz, liegt also gerade in der Größenordnung einer halben Wellenlänge der dem Spektrum überlagerten stehenden Welle. Von daher ist zu vermuten, daß dieser Peak ursächlich zurückzuführen ist auf die erwähnte falsche Anpassung der überlagerten stehenden Welle. Andererseits sollte man diesen Peak natürlich auch nicht überbewerten, zumal die Abweichungen im Differenzspektrum mit nur wenigen mK ja doch sehr klein sind (siehe Abb. 20 c). Im Falle von Ozon ist ein solcher Peak nicht zu beobachten, da dessen Signatur gegenüber derjenigen von HNO_3 ungleich stärker ist.

Damit sollen die Untersuchungen über stehende Wellen in unverrauschten Spektren nun abgeschlossen werden, da alle wesentlichen Effekte bereits besprochen wurden. Zwar hätte man noch überprüfen können, inwieweit sich die Inversionsergebnisse verschlechtern, wenn dem synthetischen Spektrum gleich mehrere stehende Wellen unterschiedlicher Periode überlagert werden und die Perioden der für die Inversion verwendeten stehenden Wellen von Perioden der obigen stehenden Wellen abweichen, doch führt dies am Ende nur zu einer Überlagerung der besprochenen Effekte. Festzuhalten bleibt jedenfalls, daß die stehende Wellen im Spektrum die Qualität der Inversionsergebnisse erheblich beeinträchtigen können, insbesondere dann, wenn die Perioden der für die Inversion verwendeten stehenden Wellen nicht mit den Perioden der dem Spektrum überlagerten stehenden Wellen übereinstimmen. Dabei wirken sich stehende Wellen kurzer Periode so aus, als wäre das Spektrum stark verrauscht. Bei stehenden Wellen langer Periode besteht insbesondere die Gefahr einer falschen Anpassung der überlagerten stehenden Welle mit der Folge einer Meßwertverfälschung.

5.1.4 Verrauschte synthetische Spektren mit stehenden Wellen

Wurden das Rauschen und die Überlagerung der Spektren mit stehenden Wellen bisher getrennt behandelt, so sollen die mit stehenden Wellen überlagerten Spektren nun zusätzlich verrauscht werden. Dies entspricht dem realistischen, bei einer Messung anzutreffenden Fall. Betrachtet werden soll hierzu das mit drei stehenden Wellen überlagerte Spektrum der Abb. 17 a, auf das zusätzlich ein Rauschen der Standardabweichung $1/100$ K bzw. $1/10$ K (bezogen auf die Kanalbreite von 0,7 MHz) addiert wird (siehe Abb. 22 a und b). Dies hat zur Folge dazu, daß die HNO_3 -Signatur in beiden Spektren praktisch nicht mehr zu erkennen ist; insbesondere im Falle des mit $1/10$ K verrauschten Spektrums ist mit einer vernünftigen Inversion des HNO_3 -Profils nicht mehr zu rechnen.

An a priori-Information wurde im wesentlichen wieder die des Abschnitts 5.1.2 verwendet, wobei die Perioden der dem Spektrum überlagerten stehenden Wellen bei der Inversion als bekannt vorausgesetzt wurden (siehe Tabelle 2).¹² In Abb. 22 sind

¹²Noch realistischer wäre sicherlich der Fall, daß die Perioden der stehenden Wellen nicht genau bekannt sind, doch entsprechen die Auswirkungen abweichender Perioden im Falle verrauschter Spektren im wesentlichen denen des vorhergehenden Abschnitts.

Nr.	Periode GHz	Ampl. K	Phase Grad	Ampl.* K	Phase* Grad	Ampl.** K	Phase** Grad
1.	0,7465	0,2	30	0,201	30,34	0,206	27,50
2.	0,403	0,15	60	0,150	60,02	0,159	59,48
3.	0,113	0,1	90	0,100	90,06	0,102	91,69

Tabelle 2: Liste der dem synthetischen Spektrum zusätzlich überlagerten stehenden Wellen. Die mit einem und zwei Sternen gekennzeichneten Spalten geben die aus dem mit 1/100 bzw. 1/10 K verrauschten Spektrum ermittelten Werte von Amplitude und Phase wieder.

die kompletten Anpassungen der Tikhonov–Phillips–Regularisierung sowie die Differenzen zwischen dem jeweils vorgegebenem Spektrum und der kompletten Anpassung dargestellt. Da letztere praktisch frei von stehenden Wellen sind, ist anzunehmen, daß diese im Laufe der Inversion vollständig angepaßt werden konnten. Wie ein Blick auf die im Zuge der Inversion ermittelten Werte von Amplitude und Phase jedoch zeigt (siehe Tabelle 2), betragen die Abweichungen im Falle des mit 1/10 K verrauschten Spektrums immerhin noch 1/100 K in der Amplitude und bis zu 2,5 Grad in der Phase, was sich wiederum negativ auf die Qualität der Inversionsergebnisse auswirkt (siehe Abb. 22). Letztere ist deutlich geringer als im Falle des ebenfalls verrauschten, jedoch nicht mit stehenden Wellen behafteten Spektrums (vgl. hierzu die Abb. 14). Vor allem für das Höhenprofil von HNO_3 ist die Beeinträchtigung aufgrund der Überlagerung von stehenden Wellen beträchtlich. Konnte zuvor zumindest das mit 1/100 K verrauschte Spektrum relativ gut invertiert werden, so gibt im vorliegenden Fall das ermittelte HNO_3 -Profil den ursprünglichen Verlauf nur unzureichend wieder. Im Falle des mit 1/10 K verrauschten Spektrums gar kann aus der Lösung der Tikhonov–Phillips–Regularisierung allenfalls geschlossen werden, daß sich in großer Höhe weniger HNO_3 befindet als in der unteren Stratosphäre; für detailliertere Aussagen ist der Verlauf des Höhenprofils ansonsten zu glatt.¹³ Auf den ersten Blick geringfügig besser schneidet das linear statistische Verfahren ab, doch ist dieses in erster Linie auf die unterschiedliche a priori–Information zurückzuführen.¹⁴ Verzichtet wurde an dieser Stelle auf eine Darstellung der Fehlerbalken und der Auflösungsmatrix \mathbf{R} , da im Prinzip von vornherein klar ist, daß die Fehler erheblich größer sein werden und die Auflösung entsprechend geringer sein wird als in den vergleichbaren Fällen der

¹³Die zur Stabilisierung der Lösung verwendete a priori–Information war ja im Falle der Tikhonov–Phillips–Regularisierung gerade die Forderung, das Höhenprofil solle möglichst glatt sein. Offensichtlich ist der Einfluß des Spektrums auf das Inversionsergebnis im vorliegenden Fall also äußerst gering und die Lösung wird durch die Regularisierung dominiert.

¹⁴Im Gegensatz zur Tikhonov–Phillips–Regularisierung, bei der die invertierten Höhenprofile möglichst glatt sein sollten, wurden bei dem linear statistischen Verfahren die Volumenmischungsverhältnisprofile der Standardatmosphäre für mittlere Breiten [13, 14] als a priori–Information verwendet.

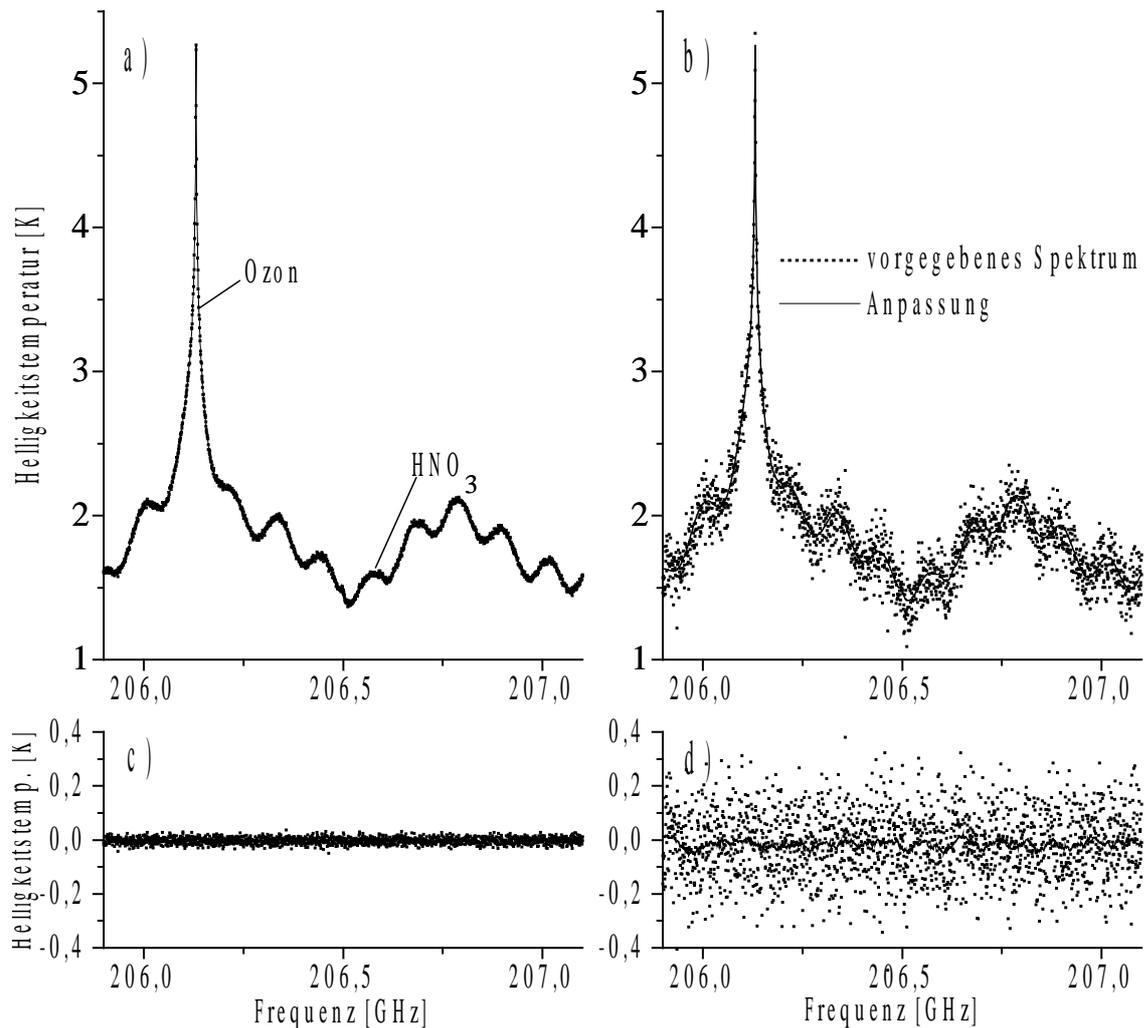


Abbildung 22: In a) mit drei stehenden Wellen und einem Rauschen der Standardabweichung 1/100 K je Kanal der Bandbreite 0,7 MHz überlagertes Spektrum, in b) entsprechendes, mit einer Standardabweichung von 1/10 K verrauschtes Spektrum. Zusätzlich sind in a) und b) die kompletten Anpassungen der Tikhonov–Phillips–Regularisierung mit dargestellt (durchgezogene Linien). In c) und d) ist die Differenz zwischen dem jeweils vorgegebenen Spektrum und der kompletten Anpassung aufgetragen; in d) wurde zusätzlich das über jeweils 30 benachbarte Kanäle gemittelte Differenzspektrum mit dargestellt.

beiden Abbildungen 15 und 16.

Zusammenfassend läßt sich feststellen, daß mit der neuen Methode zur Berücksichtigung von stehenden Wellen und der simultanen Inversion mehrerer verschiedener Spurengase aus einem Spektrum, die Tikhonov–Phillips–Regularisierung sowie die

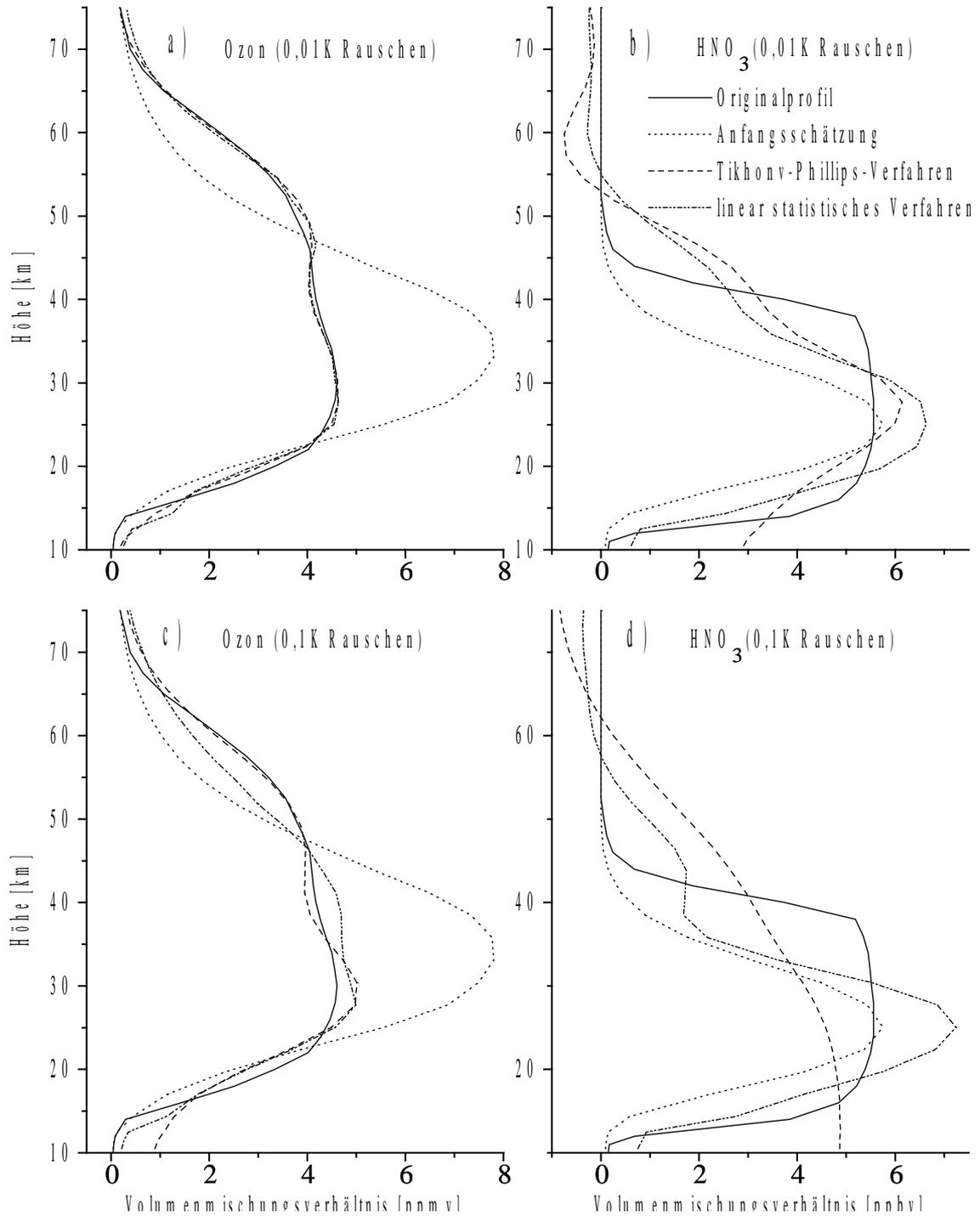


Abbildung 23: Gleichzeitige Inversion von Ozon und HNO₃ aus einem mit drei stehenden Wellen überlagerten, verrauschten Spektrum. In a) und b) 1/100 K Rauschen, in c) und d) 1/10 K Rauschen.

statistischen Regularisierungsverfahren als leistungsfähige Inversionswerkzeuge bezeichnet werden können. Mit ihrer Hilfe können die gemessenen Spektren nun direkt, d. h. ohne die Spektren vorher einer aufwendigen Aufbereitung zur Elimination von stehenden Wellen unterziehen zu müssen, in den Inversionsalgorithmus eingespeist und invertiert werden.¹⁵

5.2 Anwendung auf gemessene Spektren

5.2.1 Berücksichtigung des troposphärischen Wasserdampfs

Bisher wurden die Volumenmischungsverhältnisse aller nicht relevanten Spurengase als bekannt vorausgesetzt. Dies galt insbesondere für den troposphärischen Wasserdampf, welcher starken zeitlichen und räumlichen Schwankungen unterliegt (siehe Abb. 24). Dessen starke Variation soll bei der Inversion künftig mit berücksichtigt werden.

Nachdem die Tikhonov–Phillips–Regularisierung und das linear statistische Verfahren erfolgreich auf den Fall der Ableitung gleich mehrerer Spurengase aus einem Spektrum erweitert werden konnten, liegt es natürlich nahe, zunächst zu versuchen, den troposphärischen Wasserdampf aus einem gegebenen Spektrum gleich mit zu invertieren. Generell bereitet jedoch die Ableitung von Höhenprofilen all jener Spurengase Probleme, deren Volumenmischungsverhältnis mit der Höhe stark abnimmt, im Gegensatz zu Höhenprofilen stratosphärischer Spurengase, wie z. B. Ozon, HNO_3 oder ClO , welche ein ausgeprägtes Maximum in großer Höhe besitzen. Erschwerend kommt hinzu, daß der troposphärische Wasserdampf sich in der Mehrzahl der Spektren lediglich in Form eines breitbandigen Untergrundes auswirkt, da die Linienzentren weit außerhalb der für die vorliegende Aufgabe relevanten Spektralbereiche liegen (siehe die Abbildungen 38 – 40 in Anhang B.2). Deshalb ist der Informationsgehalt für eine höhenauflösende Inversion von Wasserdampfprofilen zu gering.

Viele Autoren [48, 36, 59, 97] approximieren die Troposphäre deshalb durch ein geeignetes Einschichtenmodell und korrigieren ihre Spektren um den Einfluß des troposphärischen Wasserdampfs vor der eigentlichen Inversion.¹⁶ Eine weitere Möglichkeit der Berücksichtigung besteht in der Verwendung eines in seiner Form fest vorgegebenen, für die aktuellen Wetterverhältnisse charakteristischen Wasserdampfprofils, welches zur Anpassung an das gegebene Spektrum entsprechend skaliert wird. Diese Möglichkeit wird im folgenden bevorzugt werden. Natürlich erfolgt auch bei diesem Verfahren die Skalierung des Wasserdampfprofils noch vor der eigentlichen Inversion,

¹⁵Selbstverständlich müssen die Einzelspektren zuvor immer noch integriert und etwaige Ausreißer eliminiert werden.

¹⁶Bei der Fehlerbetrachtung in Abschn. 5.2.3 wird diese Strategie ebenfalls verfolgt, so daß sich eine Behandlung der Methode im Detail an dieser Stelle erübrigt.

im Gegensatz zum Einschichtenmodell fließt die Information über den troposphärischen Wasserdampfgehalt jedoch in die Berechnung der Sensitivitätsfunktionen mit ein. Auf diese Weise ist es möglich, den Höhenverlauf des troposphärischen Wasserdampfs zumindest näherungsweise mit zu berücksichtigen; und falls ein Höhenprofil aus Radiosondenaufstiegen oder sonstigen Meßverfahren zur Verfügung steht, so kann dieses ohne großen Aufwand in die Auswertung mit einbezogen werden. Natürlich ist der numerische Aufwand der zweiten Methode entsprechend größer; allerdings besteht die Möglichkeit, da die Absorption und Emission des troposphärischen Wasserdampfs sehr breitbandig erfolgt, in jeder Höhengschicht den Einfluß des Wasserdampfs nur am linken und rechten Rand des Spektrums exakt zu berechnen und die Werte dazwischen linear zu interpolieren. In diesem Falle arbeitet auch die zweite Methode sehr effizient.

Im Großen und Ganzen funktionierte das angegebene Verfahren zur Anpassung des Wasserdampfs so gut, so daß auf eine eingehendere Untersuchung der Auswirkungen eines die aktuellen Wetterverhältnisse nicht exakt widerspiegelnden Wasserdampfprofils an dieser Stelle verzichtet werden soll. Ohnehin sollen bei der Fehlerdiskussion in Abschn. 5.2.3 die Auswirkungen des troposphärischen Wasserdampfprofils auf das Inversionsergebnis noch näher untersucht werden.

5.2.2 Die 142,175 GHz-Linie von Ozon

Am 28. und 30. März sowie am 26. und 28. April 1994 konnten auf der Hornisgrinde (Nordschwarzwald, 1165 m N.N.) erstmals Messungen mit dem Ozon-Radiometer bei 142 GHz in einer Höhe von 1120 m N.N. durchgeführt werden. Bereits zuvor waren vom FZK aus (100 m N.N.) erste Testmessungen unternommen worden, doch war die Qualität der damals aufgenommenen Spektren für eine Auswertung nicht ausreichend.

Charakterisierung der Messungen

Das Interesse dieser Messungen galt dabei in erster Linie der starken Signatur von Ozon bei 142,17512 GHz. Diese wurde gewählt, da sie neben ihrer relativen Stärke¹⁷ von den Signaturen anderer Spurengase gut separiert ist (siehe Abb. 38). Eine Ausnahme bildet in dieser Hinsicht wieder der troposphärische Wasserdampf, dessen Einfluß sich in Form eines breitbandigen Untergrundes über das gesamte Millimeterwellenspektrum zieht. Anders als im Spektrum der Abb. 38, welches für eine Höhe von 10 km über Meereshöhe gerechnet wurde, ist der Wasserdampfuntergrund bei den Messungen von der Hornisgrinde jedoch deutlich stärker ausgeprägt, bedingt durch die geringe Höhe von lediglich 1120 m. Als weitere Komplikation empfängt das auf der

¹⁷Die Linienamplitude erreicht in der trockenen Atmosphäre Werte um die 20 K bei einem Elevationswinkel von 90°.

Hornisgrinde eingesetzte Ozon-Radiometer beide Seitenbänder, im Gegensatz zu den neu konzipierten CIO-Radiometern bei 204 und 278 GHz.¹⁸

Dies hat zur Folge, daß dem eigentlich interessierenden Spektrum der Bandbreite 1,2 GHz bei 142 GHz das zweite Seitenband bei 138 GHz überlagert ist. In einer Messung ist die O₃-Signatur deshalb von vornherein nur halb so hoch. (Hinzu kommt die Dämpfung der Signatur auf dem Weg durch die Troposphäre zum Boden, siehe hierzu die Abbildungen 30 und 38). Dies muß bei der Auswertung entsprechend berücksichtigt werden; abgesehen vom doppelten numerischen Aufwand bei der Berechnung der Sensitivitätsfunktionen, welche nun für beide Seitenbänder berechnet werden müssen, bedeutet dies jedoch kein prinzipielles Problem.

In der Abb. 24 sind die Ergebnisse sechs zeitlich aufeinanderfolgender Einzelmessungen dargestellt. Sie wurden aufgenommen am 28. März 1994 unter einem Elevationswinkel von 29,5° bei Integrationszeiten zwischen 150 und 435 s. Auffallend an dieser Meßreihe ist die starke Variation des Wasserdampfuntergrundes, welcher von einer mittleren Helligkeitstemperatur um die 65 K zu Beginn (10:53 Uhr) innerhalb von eineinhalb Stunden auf einen Wert von 95 K anstieg (12:18 Uhr). Um 14:00 Uhr erreichte der Wasserdampfuntergrund gar einen Wert von 160 K (in Abb. 24 nicht mit dargestellt). Die Einzelspektren konnten deshalb nicht einfach gemittelt werden, sondern mußten zuvor nach dem herkömmlichen Verfahren um den Einfluß der Troposphäre korrigiert werden. An den anderen Meßtagen waren die Wetterverhältnisse hingegen deutlich stabiler.

Weiterhin fällt die Überlagerung der Spektren mit stehenden Wellen auf. Deutlich zu erkennen ist eine stehende Welle relativ kurzer Periode (0,113 GHz) in den Spektren um 10:53, 11:17 und 11:32 Uhr. Wie sich bereits während der Messungen herausstellte, wurde diese hervorgerufen durch Mehrfachreflexionen zwischen dem Antennen-Hohlleiter-Übergang und dem kalten Absorber (siehe hierzu die Ausführungen der

¹⁸Alle drei im FZK verwendeten Radiometer arbeiten nach dem Heterodyn-Prinzip. Dabei wird das zu messende Signal der Frequenz ν_{SF} mit dem Signal eines lokalen Oszillators (LO) der Frequenz ν_{LO} überlagert. In einem Mischer, welcher ein nichtlineares Halbleiter-Bauteil enthält, gemeinhin eine Schottky-Diode, wird die zu messende Signalfrequenz anschließend auf eine Zwischenfrequenz ν_{RF} umgesetzt. Nach [38] ergeben sich unter der Voraussetzung, daß die Leistung des LO groß gegenüber der Signalleistung ist, die beiden Empfangsfrequenzen

$$\nu_{SF_{1,2}} = \nu_{LO} \pm \nu_{ZF}$$

Der Heterodynempfänger empfängt also, sofern keine besonderen Vorkehrungen getroffen werden, beide Seitenbänder (Doppelseitenbandbetrieb, DSB). Wird hingegen – wie im Falle der beiden CIO-Radiometer – vor dem Mischen ein Seitenband unterdrückt, spricht man vom Einseitenbandbetrieb (SSB).

Bei den obigen Messungen von Ozon war der LO i. d. Regel eingestellt auf die Frequenz 139,9531 GHz. Dem eigentlich interessierenden Seitenband zwischen 141,575 und 142,775 GHz (dies entspricht gerade der maximalen Bandbreite von 1,2 GHz) ist deshalb das „gespiegelte“ zweite Seitenband zwischen 137,131 und 138,331 GHz überlagert.

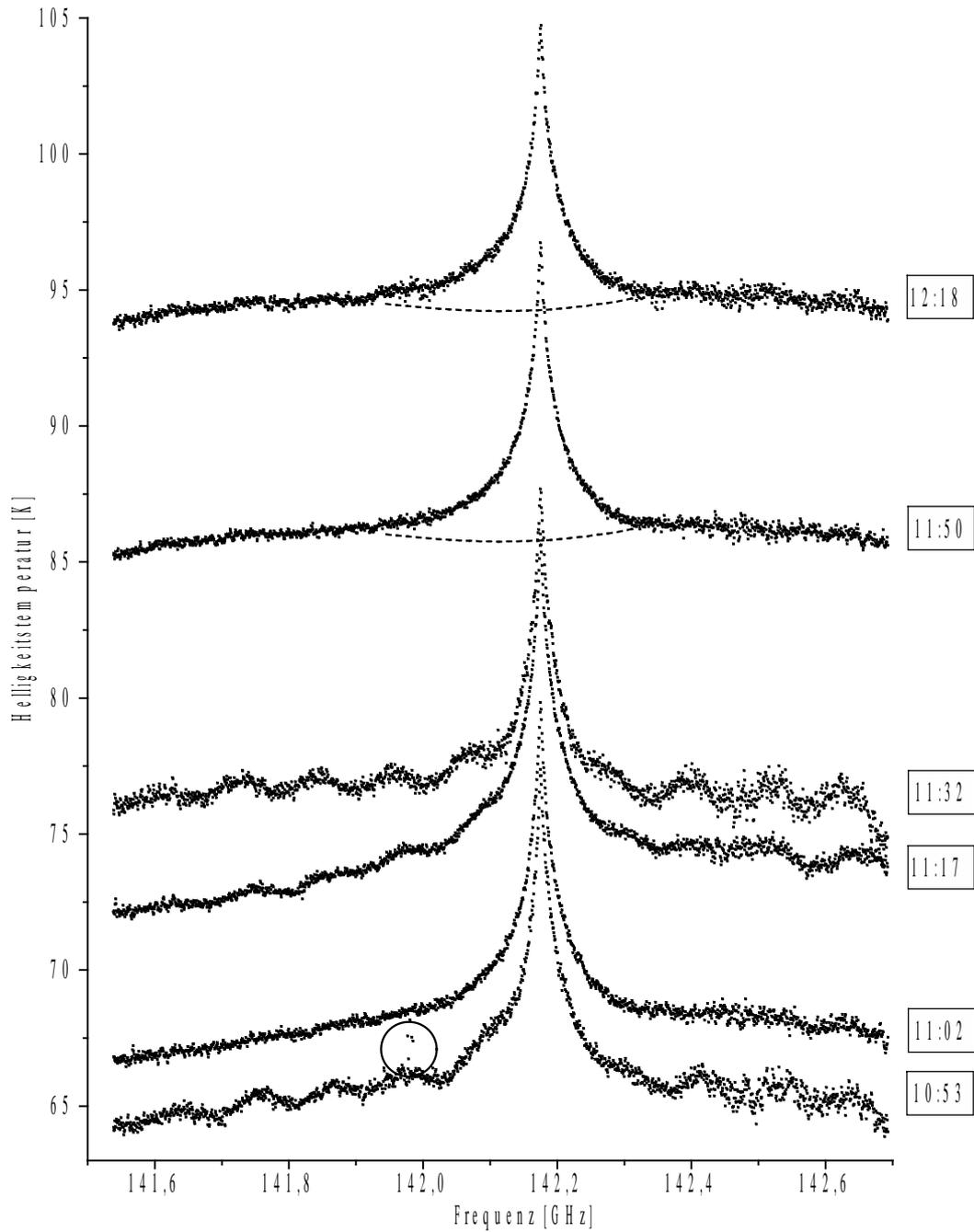


Abbildung 24: Sechs zeitlich aufeinanderfolgende Messungen von 28. März 1994. Die Spektren wurden aufgenommen unter einem Elevationswinkel von 29,5 Grad bei Integrationszeiten zwischen 150 und 435 s.

Fußnote 2 auf Seite 51). In den beiden Spektren um 11:50 und 12:18 Uhr ist überdies eine stehende Welle relativ großer Periode (0,746 GHz) zu erkennen (in Abb. 24 hervorgehoben durch zwei zusätzlich mit eingezeichnete, gestrichelte Linien). Diese stehende Welle war auch allen anderen Spektren der Meßkampagne überlagert (in den zusätzlich mit einer stehenden Welle kurzer Periode überlagerten Spektren nur schwer zu erkennen) und ist zurückzuführen auf Mehrfachreflexionen zwischen der Linse und dem Hohlleiter-Übergang der Antenne. In einigen weiteren Spektren (nicht denen der Abb. 24) trat überdies eine stehende Welle mittlerer Periode (0,403 GHz) auf, hervorgerufen durch Mehrfachreflexionen zwischen dem Antennen-Hohlleiter-Übergang und dem warmen Absorber.

Ansonsten lief das Radiometer sehr stabil (lediglich drei Ausreißer im Spektrum vom 28. März um 10:53 Uhr, siehe die eingekreisten Werte in Abb. 24), so daß nahezu alle während der vier Meßtage aufgenommenen Spektren zur Auswertung herangezogen werden konnten.

Präprozessierung der Messungen

Die eigentliche Auswertung lief nun wie folgt ab: Zunächst wurden – sofern der Wasserdampfuntergrund zwischen den einzelnen Messungen nicht allzu stark variierte – über mehrere Einzelspektren gemittelt, um so eine Verringerung der Rauschamplitude zu erreichen.¹⁹ Dies wirkte sich unter anderem auch vorteilhaft aus auf die Überlagerung der Spektren mit stehenden Wellen, da insbesondere die stehenden Wellen kurzer Periode in ihrer Phase von Messung zu Messung stark variieren und sich im Laufe der Integration mit herausmitteln.²⁰ Anschließend wurde für jedes Spektrum ein geeignetes Temperaturprofil generiert.

Eine der Voraussetzungen für die Auswertung von gemessenen Spektren ist ja die Kenntnis der Höhenprofile von Druck und Temperatur, welche im aktuellen Fall (Hornisgrinde) – von den Bodenwerten einmal abgesehen – vor Ort nicht bekannt waren. Stattdessen wurden Radiosondenprofile aus der näheren Umgebung (Stuttgart) [18] zur Auswertung herangezogen und nach unten hin um die aktuellen Bodenwerte, sowie nach oben hin um das Temperaturprofil der Standardatmosphäre für mittlere Breiten [13, 14] ergänzt.²¹ Schließlich wurde das resultierende Temperaturprofil leicht geglättet, da nicht erwartet werden konnte, daß Feinheiten des Temperaturprofils aus Stuttgart auch auf die Hornisgrinde zutreffen würden. Ein typisches Resultat dieses

¹⁹Nach der Radiometerformel [86, 88] ist die Rauschamplitude der Wurzel aus der Integrationszeit umgekehrt proportional.

²⁰Leider war diese Mittelung bei den sich schnell ändernden Witterungsverhältnissen vom 28. März nicht immer möglich, so daß die Einzelspektren entweder einzeln ausgewertet werden mußten (trotz des relativ starken Rauschens), oder vor der Integration um den troposphärischen Wasserdampf korrigiert werden mußten.

²¹Radiosonden erreichen eine Höhe von maximal 35 km, so daß oberhalb leider keine besseren Temperaturdaten zur Verfügung standen.

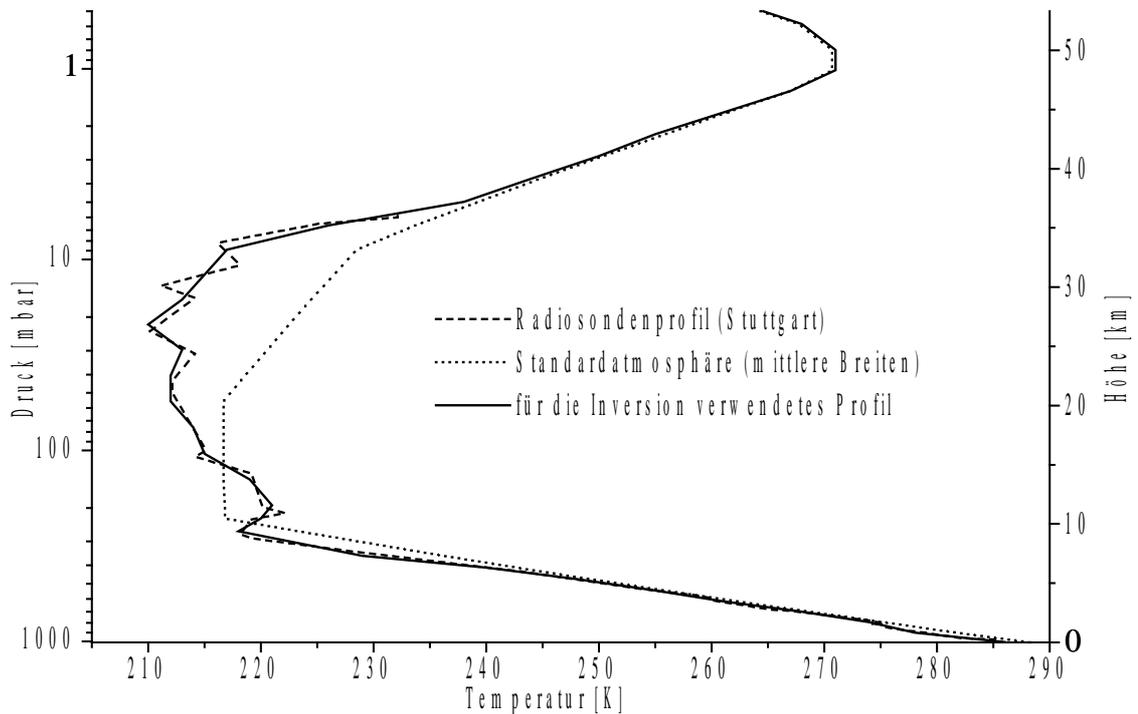


Abbildung 25: Temperaturprofile vom 28. März 1994, 12:00 Uhr. Für die Eingabe in die Inversionsrechnungen wurde das Radiosondenprofil leicht geglättet.

Verfahrens ist in Abb. 25 dargestellt (durchgezogene Linie).

Der nächste Schritt war die Anpassung des troposphärischen Wasserdampfuntergrundes an die aufintegrierten Spektren (im wesentlichen wie in Abschn. 5.2.1 beschrieben). Als Eingabe diente dabei ein Wasserdampfprofil, welches typisch war für die während der Messungen auf der Hornisgrinde herrschenden Wetterverhältnisse und früheren Radiosondenaufstiegen im Rahmen des REKLIP-Projektes 1992 entstammte [21, 37]. Dieses wurde zur Anpassung entsprechend skaliert.

Inversion

Erst im Anschluß an diese Vorarbeiten erfolgte die eigentliche Inversion. Hierzu wurde bevorzugt die Tikhonov-Phillips-Regularisierung angewendet, mit der Nebenbedingung, das zu invertierende Höhenprofil solle möglichst glatt sein. Im Gegensatz zu den statistischen Inversionsverfahren, welche neben der a priori-Kovarianzmatrix als weitere Eingabe ein typisches Volumenmischungsverhältnisprofil benötigen, war es somit möglich, eine Regularisierung der Lösung zu erreichen, ohne bereits in den Inversionsalgorithmus a priori-Informationen über die Form der Lösung hineinstecken zu müssen. Lediglich im Höhenbereich unterhalb von 15 km wurde in einigen Fällen die

Form des zu invertierenden Profils in Anlehnung an die Standardatmosphäre für mittlere Breiten vorgegeben; bzgl. des Volumenmischenverhältnisses war das Höhenprofil jedoch auch in diesem Bereich frei skalierbar. Ziel dieser Maßnahme war es, selbst für diejenigen Spektren, für die die Anpassung des troposphärischen Wasserdampfuntergrundes nicht optimal arbeitete, die Konvergenz des iterativen Gauß–Newton–Verfahrens zu erreichen.

Im Zuge der Inversion eines jeden Spektrums wurden die erwähnten drei stehenden Wellen der Perioden 0,7465, 0,403 sowie 0,133 GHz gleich mit angepaßt, wobei die Wahl des Regularisierungsparameters γ i. d. Regel a posteriori erfolgte (siehe Abschn. 3.1.4). Die Fehlerkovarianzmatrix \mathcal{S}_ε eines jeden Spektrums wurde als reine Diagonalmatrix angenommen und die Diagonalelemente selbst wurden gemäß der Radiometerformel initialisiert, nach der die Standardabweichung der Rauschamplitude der Wurzel aus der Integrationszeit umgekehrt proportional ist [86, 88].²² Zusätzlich wurde auf die Diagonalelemente der Fehlerkovarianzmatrix ein Wert von 0,05 K addiert, um eventuell im Zuge der Inversion der Spektren nicht vollständig angepaßte stehende Wellen bei der Fehlerrechnung näherungsweise mit zu berücksichtigen.

Anhand der Messung vom 28. März 1994, 12:18 Uhr mit einer Integrationszeit von 435 s (siehe Abb. 24) ist in der Abb. 26 exemplarisch aufgezeigt, wie im Laufe der Iteration das gemessene Spektrum immer besser angepaßt wird. Dabei ist nach drei Iterationen die Übereinstimmung bereits so gut, daß das Verfahren abgebrochen werden kann. Insbesondere sind in dem Differenzspektrum der Abb. 26 unten links keinerlei Reste von stehenden Wellen mehr zu erkennen.

Ergebnisse

In Abb. 27 sind die ermittelten Ozon–Volumenmischungsverhältnisprofile zusammen mit Radiosondenprofilen des jeweiligen Tages aus Lindenbergr (Potsdam) und vom Hohenpeißenbergr dargestellt. Letztere wurden im Rahmen der SESAME–Kampagne 1994 gemessen. Dabei ergibt sich im Vergleich mit den Radiosondenprofilen eine relativ gute Übereinstimmung, wobei natürlich nicht erwartet werden kann, daß Feinstrukturen der Radiosondenprofile in den invertierten Profilen noch aufgelöst werden können. Ohnehin war das Höhengauflösungsvermögen der '94er Messungen auf der Hornisgrinde mit 5–6 Höhengschichten und einer maximalen Auflösung von 7,5 km in 30 km Höhe nicht allzu hoch (siehe Abb. 29). Ursache hierfür war die geringe Anzahl der insgesamt aufgenommenen Spektren und das hieraus resultierende relativ starke Meßfehlerrauschen selbst nach der Integration.²³ Geringfügig besser, mit etwa sechs

²²Trotz der Überlagerung der Spektren mit stehenden Wellen wurden diese innerhalb der Fehlerkovarianzmatrix nicht mit berücksichtigt, da diese im Laufe der Inversion ja mit invertiert werden sollten.

²³Häufig wurde zu Testzwecken auch die Frequenz des LO verstimmt oder der Elevationswinkel verändert, so daß die Spektren nicht aufintegriert werden konnten.

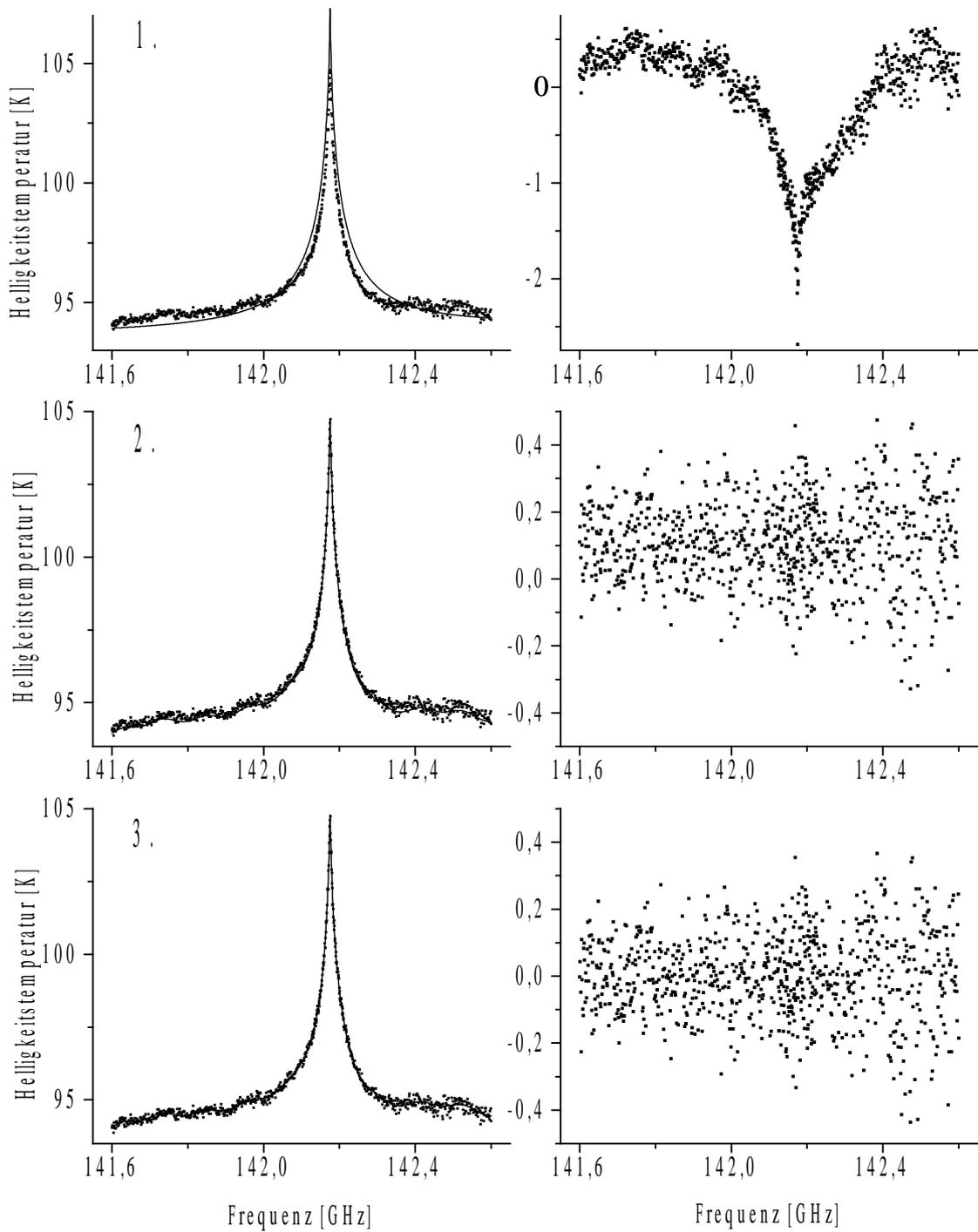


Abbildung 26: Komplette Anpassung der Tikhonov–Phillips–Regularisierung an das am 28. März 1994 um 12:18 Uhr gemessene Spektrum links; Differenz zwischen Spektrum und kompletter Anpassung rechts. Nach drei Iterationsschritten stimmen Spektrum und Anpassung gut überein.

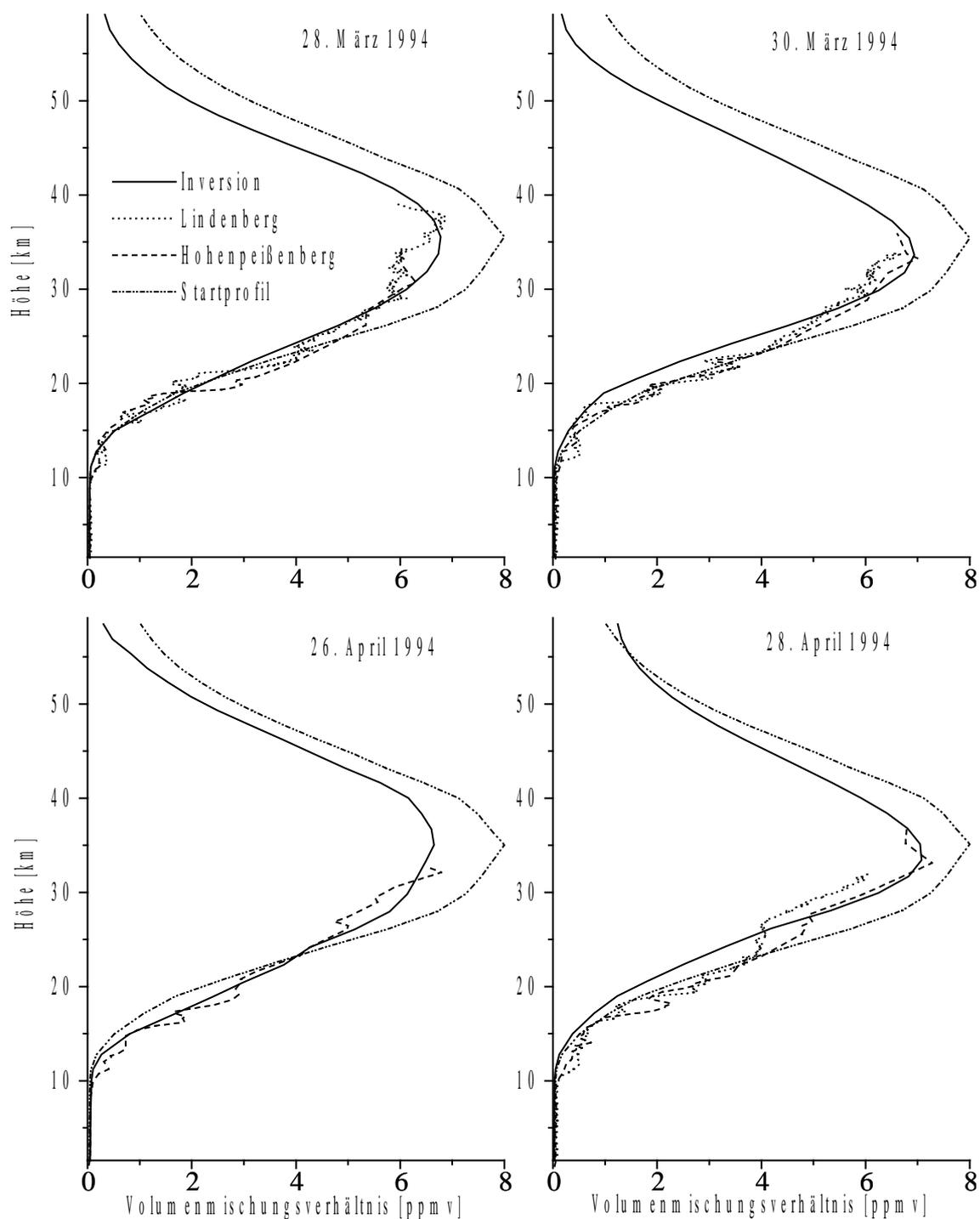


Abbildung 27: Invertierte Ozon-Volumenmischungsverhältnisprofile vom März und April 1994 aus Messungen auf der Hornisgrinde. Zum Vergleich sind Radiosondenprofile aus Lindenberg (Potsdam) und vom Hohenpeißenberg mit dargestellt. Als Startprofil diente die Standardatmosphäre für mittlere Breiten [13, 14].

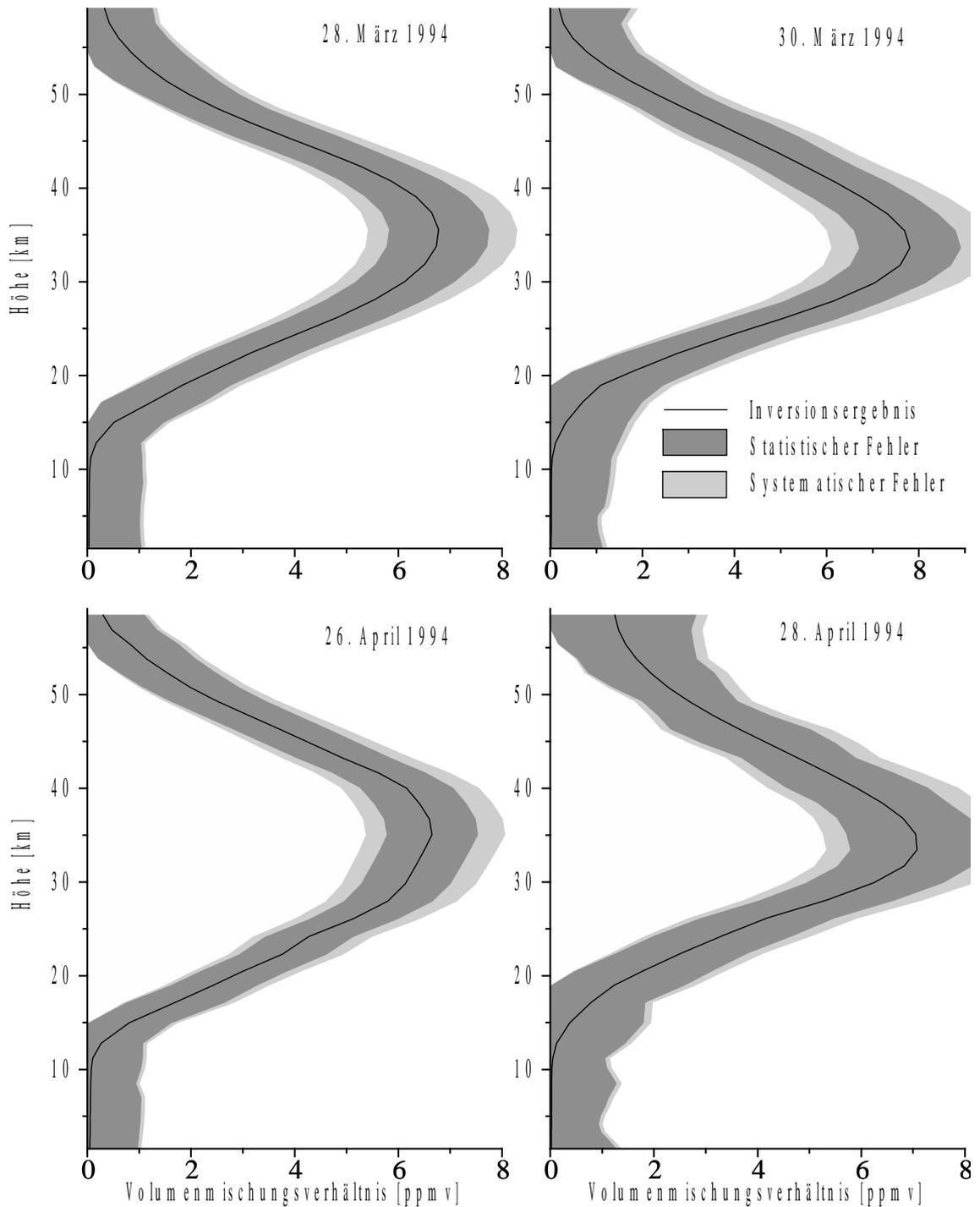


Abbildung 28: Fehlerbalken der invertierten Ozon-Volumenmischungsverhältnisprofile vom März und April 1994. Berücksichtigt wurden sowohl der statistische Fehler (linear statistisches Verfahren) sowie der systematische Fehler (siehe Abschn. 5.2.3).

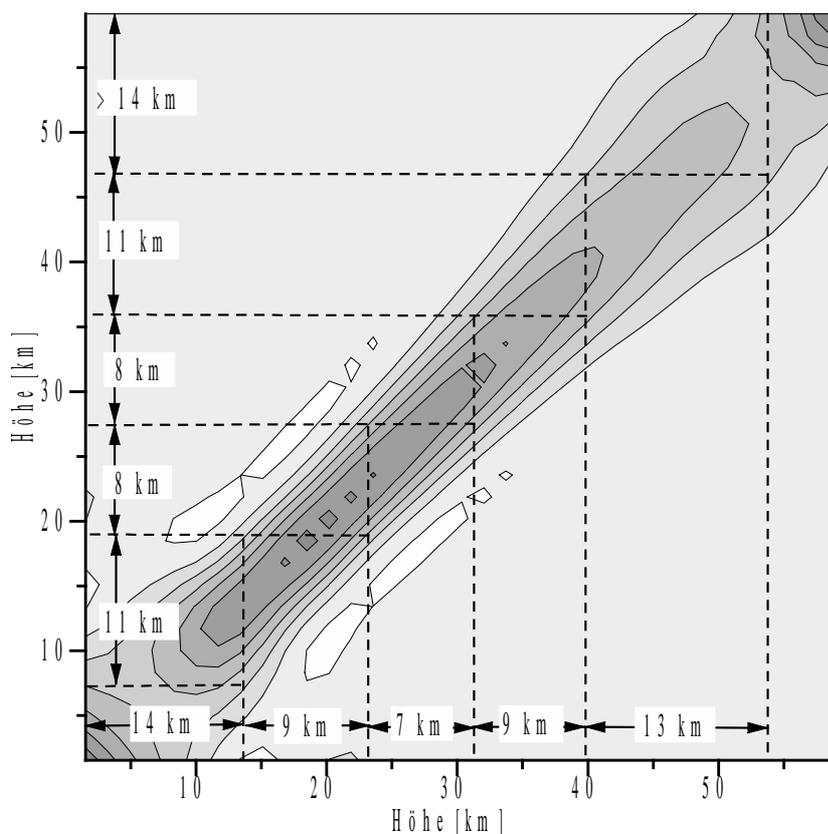


Abbildung 29: Auflösungsmatrix des Tikhonov-Phillips-Verfahrens für die Messungen vom 28. März 1994.

Höhenschichten, war die Höhenauflösung lediglich am 28. April 1994 aufgrund der etwas längeren Integrationszeit von 3850 s.

Des weiteren sind in Abb. 28 die Meßfehler aufgetragen, welche sich aus einer Berechnung der Ergebniskovarianzmatrix \mathbf{S}_γ des linear statistischen Verfahrens ergeben (siehe Gl. (28) in Abschn. 3.1.6, dunkelgrau dargestellt in Abb. 28). Um diese zu ermitteln wurde die a priori-Kovarianzmatrix \mathbf{S}_f als reine Diagonalmatrix mit einem Wert von $1 \cdot 10^{-12}$ entlang der Diagonalen initialisiert, was einer Standardabweichung des a priori-Volumenmischungsverhältnisprofils für Ozon von ± 1 ppmv entspricht.²⁴ Gleich mit dargestellt sind die Auswirkungen der systematischen Fehler auf die Inversionsergebnisse (hellgrau dargestellt in Abb. 28), wie sie durch Unsicherheit der für die Inversion verwendeten Liniendaten (Linienstärke, Lorentzhalbwidthsbreite, unteres Anregungsniveau, Mittenfrequenz), durch eine ungenaue Kalibration des Meßgerätes

²⁴Wie sich herausstellt, ist diese Annahme beinahe etwas zu optimistisch, da die invertierten Höhenprofile gerade noch innerhalb dieses Bereiches liegen.

oder durch Fehleinstellungen des Elevationswinkels zustande kommen können. Eine ausführliche Diskussion dieser Fehlerquellen und ihrer Auswirkungen erfolgt im nächsten Abschnitt.

Im Mittel beträgt die Unsicherheit der invertierten Höhenprofile aller vier Meßtage 1,5 ppmv, primär bedingt durch den relativ großen statistischen Fehler (dunkelgrau). Für die Zukunft ist zu erwarten, daß dieser aufgrund längerer Integrationszeiten und rauschärmerer Empfänger deutlich reduziert werden kann. Die systematischen Fehler (hellgrau) sind erheblich kleiner mit Werten zwischen 7–10 % (siehe hierzu die Ausführungen des folgenden Abschnitts). Doch werden auch sie in Zukunft um etwa 2–3 % reduziert werden können.

Als Ergebnis der Messungen läßt sich festhalten, daß an den betrachteten Tagen keine ungewöhnlichen Veränderungen der Ozonschicht über der Hornisgrinde auftreten (dies wäre sicherlich auch nicht zu erwarten gewesen). Statt dessen deuten die Inversionsergebnisse auf ein gegenüber der Standardatmosphäre für mittlere Breiten [13, 14] um etwa 15 % verringertes Volumenmischungsverhältnis von Ozon über den gesamten Höhenbereich hin (siehe Abb. 27). Dies ist guter Übereinstimmung mit den Radiosondenprofilen aus Lindenberg (Potsdam) und vom Hohenpeißenberg.

5.2.3 Fehlerbetrachtung

Eine vollständige Fehlerbetrachtung ist für die Inversionsrechnungen relativ schwierig durchzuführen, da eine analytische Formulierung des Auswerteproblems nicht existiert. Glücklicherweise liefern die in Abschn. 3.1.6 über die Theorie der Inversion vorgestellten Fehlerabschätzungen jedoch zumindest eine Aussage über den Datenfehler, welcher aus dem Meßfehlerauschen resultiert (siehe Gl. (27)). Im Falle der statistischen Regularisierungsverfahren ist darüber hinaus eine Aussage über den Regularisierungsfehler möglich, (siehe Gl. (28) auf Seite 29); insofern wurden die statistischen Fehler bereits erschöpfend behandelt und werden im folgenden nicht weiter berücksichtigt. Nicht in die obige Fehlerabschätzung mit eingeschlossen sind jedoch die systematischen Fehler, wie sie durch die Verwendung eines falschen Temperaturprofils, durch die Verwendung falscher Liniendaten (Linienstärke, Lorentzhalbwertbreite, unteres Anregungsniveau, Mittenfrequenz) oder durch eine ungenaue Kalibration des Meßgeräts entstehen können. Ihre Auswirkungen auf die Qualität der Inversionsergebnisse sollen im folgenden näher untersucht werden. Angegeben wird dabei jeweils die Standardabweichung (1σ -Fehler).

Um dieses zu erleichtern, wird die Troposphäre zunächst durch ein einfaches Einschichtenmodell approximiert. Dies hat den Vorteil, daß anhand dieses einfachen Modells die relevanten Zusammenhänge besser aufgezeigt werden können. Weiterhin wird angenommen, daß die Spektren vor der eigentlichen Inversion jeweils um den Einfluß

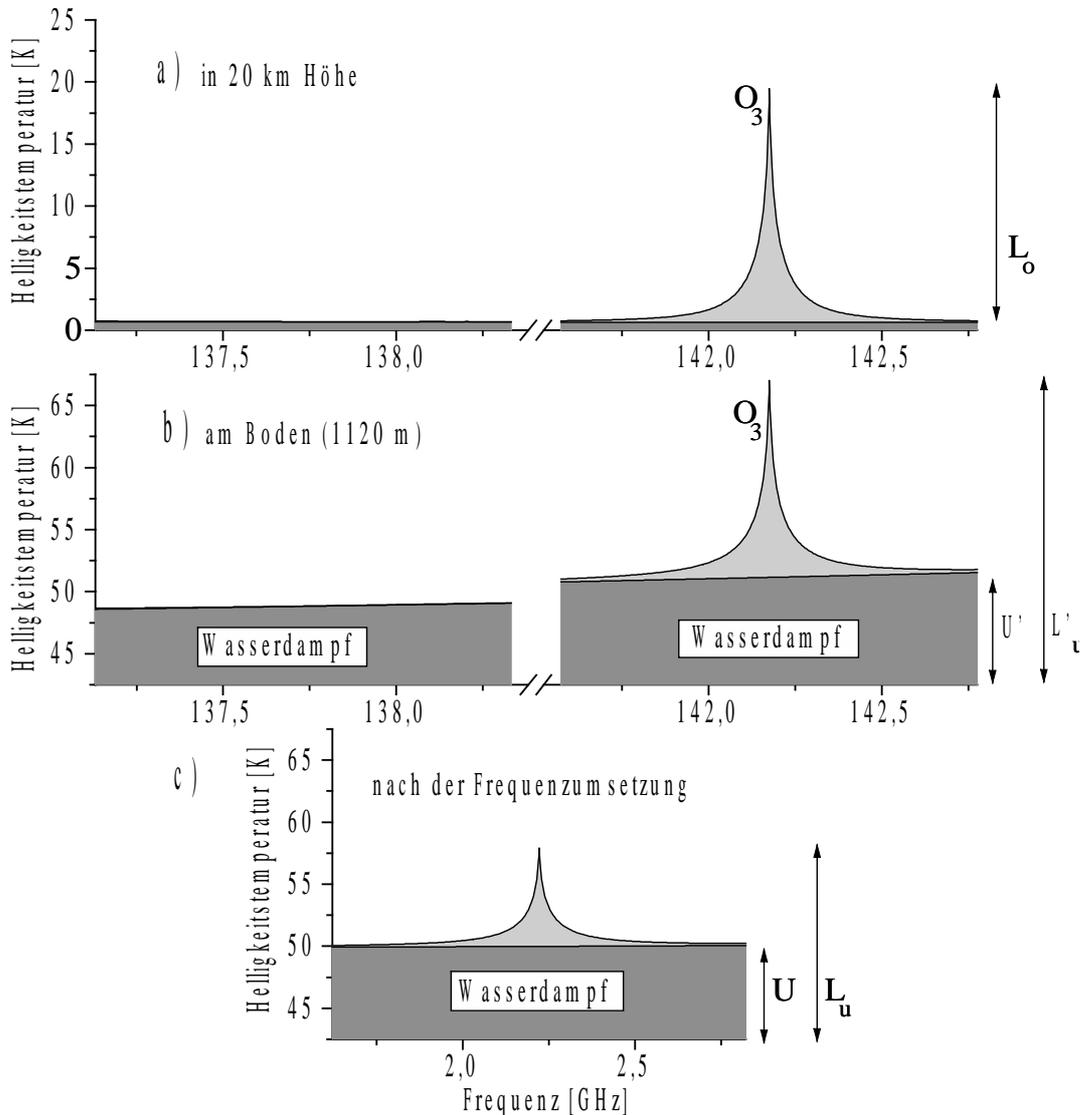


Abbildung 30: Einfluß der Troposphäre sowie der Frequenzumsetzung auf das Spektrum am Beispiel von Ozon bei 142,175 GHz, gerechnet für die Standardatmosphäre für mittlere Breiten [13, 14]. In a) O₃-Signatur und zweites Seitenband in 20 km Höhe, in b) Strahldichtespektrum nach Durchgang durch die Troposphäre und in c) resultierendes Spektrum nach der Frequenzumsetzung.

der Troposphäre korrigiert wurden.²⁵ Beide Annahmen werden dadurch gerechtfertigt, daß die Auswirkungen der systematischen Fehler auf die Inversionsergebnisse ja

²⁵In Abschn. 5.2.1 war ja gerade in Anlehnung an dieses gängige Verfahren ein verbessertes Verfahren vorgeschlagen worden, bei dem die Spektren nicht mehr troposphärenkorrigiert zu werden brauchten, sondern der Einfluß der Troposphäre bei der Berechnung der Sensitivitätsfunktionen gleich mit berücksichtigt wurde. Insofern handelt es sich um eine Abschätzung des ungünstigsten

lediglich abgeschätzt werden sollen. In Abb. 30 ist für den Fall des Ozon-Radiometers bei 142 GHz (vgl. Abb. 38) noch einmal rekapituliert, wie das am Boden gemessene Signal letztlich zustande kommt.

Sei L_o die oberhalb der Troposphäre unter einem Elevationswinkel von 90° (Blickrichtung senkrecht nach oben) zu beobachtende stratosphärische O_3 -Signatur, so folgt unter der Annahme eines Einschichtenmodells für die Troposphäre für das unter einem Elevationswinkel α am Boden meßbare Signal L'_u in Anlehnung an die Gl. (5) aus Abschn. 2.1.3 ²⁶

$$L'_u = e^{-k \Delta h / \sin \alpha} \sin^{-1} \alpha L_o + \underbrace{[1 - e^{-k \Delta h / \sin \alpha}] B(T_{trop})}_{U'}. \quad (56)$$

Dabei ist k der mittlere Absorptionskoeffizient und Δh die Dicke der Troposphäre. Die Größe $B(T_{trop})$ ist die Planckfunktion – abhängig von der mittleren Troposphären-temperatur T_{trop} – und U' ist der am Boden im rechten Seitenband zu beobachtende Strahldichteuntergrund, hervorgerufen durch die starke Emission des troposphärischen Wasserdampfs (siehe Abb. 30 a und b). Überdies wurde in Erweiterung der Gl. (5) die Abhängigkeit der Strahldichte vom Elevationswinkel mit berücksichtigt. ²⁷

Dieses am Boden zu beobachtende Spektrum L'_u soll nun um den Einfluß der Troposphäre korrigiert werden. Wegen

$$U' = [1 - e^{-k \Delta h / \sin \alpha}] B(T_{trop})$$

läßt sich die Transmission der Troposphäre auch schreiben als

$$e^{-k \Delta h / \sin \alpha} = \frac{B(T_{trop}) - U'}{B(T_{trop})},$$

so daß

$$L'_u = \frac{B(T_{trop}) - U'}{B(T_{trop})} \sin^{-1} \alpha L_o + U'.$$

Hieraus folgt für die troposphärenkorrigierte O_3 -Signatur (Blickrichtung senkrecht nach oben)

$$L_o = \frac{B(T_{trop})}{B(T_{trop}) - U'} \sin \alpha [L'_u - U'].$$

Falles.

²⁶Der besseren Übersichtlichkeit wegen wird die Frequenzabhängigkeit der spektralen Strahldichte in der Schreibweise im folgenden vernachlässigt. Ohnehin absorbiert die Troposphäre in diesem Frequenzbereich sehr breitbandig.

²⁷Dabei wurde für die Fehlerabschätzung von dem besonders einfachen Fall einer planaren Atmosphäre ausgegangen. Diese Vereinfachung ist zulässig, da der Elevationswinkel bei Messungen vom Boden aus immer größer als 15° gewählt wird, so daß die Krümmung der Erdatmosphäre vernachlässigt werden kann.

Man kann sich leicht überlegen, daß für den Fall, daß dem eigentlich interessierenden Spektrum ein zweites Seitenband überlagert ist (siehe Abb. 30), die beiden obigen Gleichungen erweitert werden müssen zu

$$\begin{aligned} L_u &= w \frac{B(T_{trop}) - U}{B(T_{trop})} \sin^{-1} \alpha L_o + U \\ L_o &= \frac{1}{w} \frac{B(T_{trop})}{B(T_{trop}) - U} \sin \alpha [L_u - U] \end{aligned} \quad 0 < w < 1. \quad (57)$$

Dabei ist w das Gewicht, mit dem das interessierende rechte Seitenband zu dem im Doppelseitenbandbetrieb gemessenen Spektrum der Abb. 30 c beiträgt. Sind die Konversionsverluste bei der Frequenzumsetzung in beiden Seitenbändern gleich hoch, nimmt w den Wert 0,5 an. Weiterhin wurde in (57) vereinfachend angenommen, daß der Strahldichteuntergrund in beiden Seitenbändern etwa gleich groß ist.

Anhand der Gl. (57) können die Einflüsse eines falsch eingestellten Elevationswinkels, einer falschen Gewichtung der beiden Seitenbänder, einer ungenauen Kalibration des Meßgeräts (gemeint ist die absolute Amplitudenkalibration) sowie einer falsch angenommenen mittleren Troposphärentemperatur auf die stratosphärische O_3 -Signatur L_o leicht untersucht werden. Es handelt sich hierbei sämtlich um Fehler, die sich multiplikativ auf das Inversionsergebnis auswirken, dieses also lediglich skalieren, jedoch nicht in seiner Form verändern.

- **Auswirkungen eines falsch eingestellten Elevationswinkels**

Zunächst soll untersucht werden, wie sich ein bei der Auswertung als zu groß angenommener Elevationswinkel auf das Inversionsergebnis auswirkt. D. h. das Radiometer war bei der Messung unter einem flacheren Winkel auf die Atmosphäre gerichtet als in der Inversion berücksichtigt. Die Unsicherheit der Winkeleinstellung beträgt etwa 1/2 Grad, wobei dieser Wert in Zukunft weiter verbessert werden dürfte. Hieraus ergibt sich für die troposphären- und winkelkorrigierte, stratosphärische Signatur L_o und damit für das Inversionsergebnis ein relativer Fehler in Abhängigkeit vom eingestellten Elevationswinkel von

$$\frac{L_o(\alpha + 0,5^\circ) - L_o(\alpha)}{L_o(\alpha)} = \frac{\sin(\alpha + 0,5^\circ)}{\sin \alpha} - 1.$$

Der Verlauf dieser Kurve ist in Abb. 31 a dargestellt. Demzufolge ist der relative Fehler bei einem Elevationswinkel von 90° (Blickrichtung senkrecht nach oben) und einer Abweichung von einem halben Grad vernachlässigbar klein; hingegen beträgt der relative Fehler bei einem Elevationswinkel von 20° bereits 2 %.

- **Auswirkungen einer falschen Gewichtung der Seitenbänder**

Wie aus dem Datenblatt des Herstellers hervorgeht, sind die Konversionsverluste des Mischers für das linke Seitenband um etwa 1 dB höher als im eigentlich

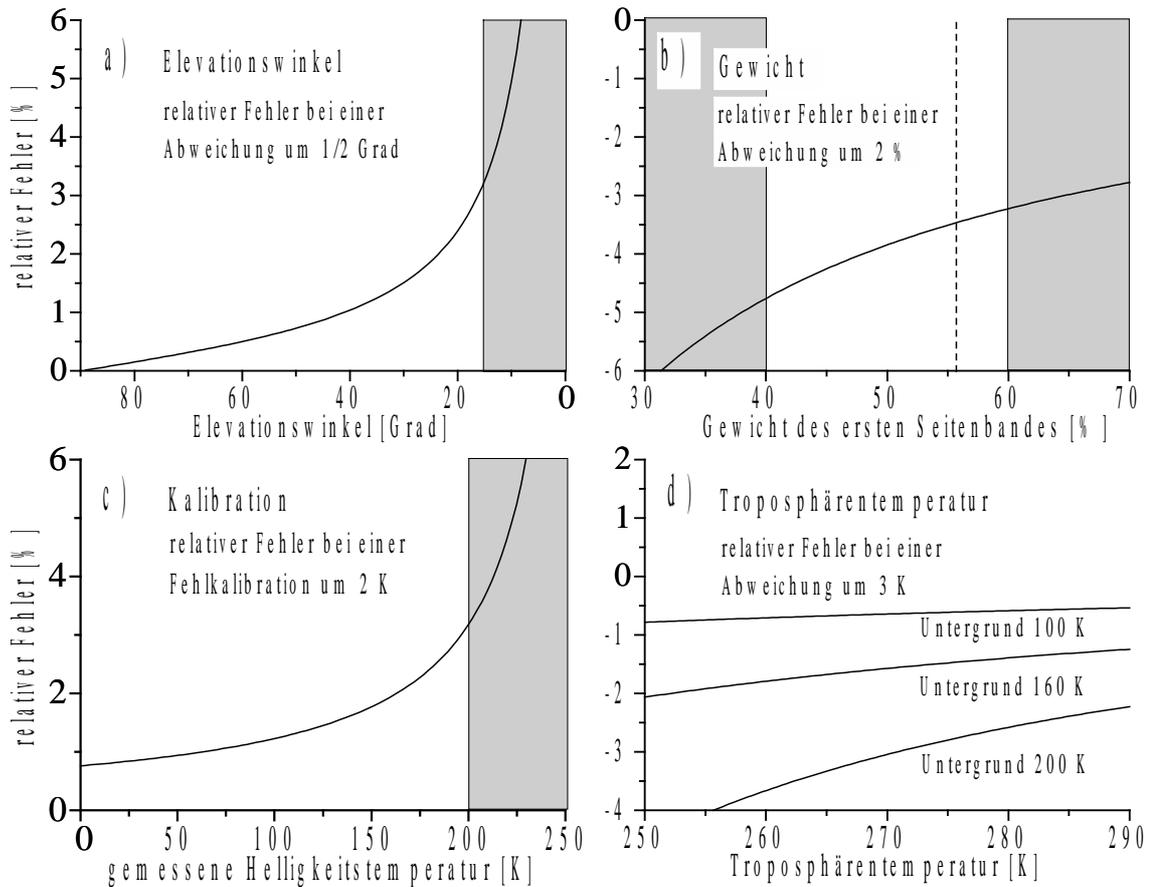


Abbildung 31: Relativer Fehler der invertierten Volumenmischungsverhältnisprofile. In a) relativer Fehler infolge eines um 1/2 Grad zu groß angenommenen Elevationswinkels; in b) relativer Fehler infolge eines um 2 % zu groß angenommenen Gewichts des ersten Seitenbandes; in c) relativer Fehler infolge einer um 2 K zu hoch bestimmten Helligkeitstemperatur; in d) relativer Fehler infolge einer um 3 K zu hoch angenommenen Troposphärentemperatur.

interessierenden rechten Seitenband bei 142 GHz [5]. Dies entspricht einem relativen Gewicht des rechten Seitenbandes von $w = 0,56$. Leider sind die Unsicherheiten der Messungen relativ hoch mit $\Delta w \approx 2/100$. Hieraus ergibt sich für das Inversionsergebnis ein relativer Fehler in Abhängigkeit vom Gewicht w des interessierenden Seitenbandes von

$$\frac{L_o(w + 2/100) - L_o(w)}{L_o(w)} = \frac{w}{w + 2/100} - 1.$$

Der Verlauf dieser Kurve ist in Abb. 31 b dargestellt. Offensichtlich wirkt sich eine falsch angenommene Gewichtung der beiden Seitenbänder relativ stark

aus auf die Qualität der Inversionsergebnisse. Z. B führt eine Abweichung von 2/100 nach oben bei einem Gewicht $w = 0,56$ auf ein um 3,5 % zu niedriges Inversionsergebnis.²⁸

- **Auswirkungen einer ungenauen Kalibration**

Die Kalibration der Meßdaten erfolgt durch Vergleich der zu messenden Strahldichte mit der Strahldichte zweier Eichlasten bekannter Temperatur und unter der Annahme einer linearen Gesamtverstärkung. Als Eichlasten dienen dabei zwei Schwarzkörper unterschiedlicher physikalischer Temperatur, nämlich 77 K (flüssiger Stickstoff) und ca. 300 K (Umgebungstemperatur). Die Unsicherheit des Kalibrationsverfahrens beträgt etwa 2 K und wirkt sich gleichermaßen aus auf das Signal L_u sowie den Untergrund U . Für den relativen Fehler der Inversionsergebnisse in Abhängigkeit von der Stärke des Untergrundes U folgt damit

$$\frac{L_o(L_u + 2\text{K}, U + 2\text{K}) - L_o(L_u, U)}{L_o(L_u, U)} = \frac{B(T_{trop}) - U}{B(T_{trop}) - U - 2\text{K}} - 1.$$

Der Verlauf dieser Kurve ist in Abb. 31 c dargestellt. Eine ungenaue Kalibration wirkt sich demnach umso stärker aus, je größer der Einfluß der Troposphäre auf das Spektrum, d. h. je größer der Untergrund U ist. Bei einer gegenüber dem wahren Wert um 2 K zu hoch angenommenen Helligkeitstemperatur und einem Untergrund U von 175 K führt dies auf eine Abweichung von etwa 2 %.²⁹

- **Auswirkungen einer falsch angenommenen Troposphärentemperatur**

Wie aus der Gl. (57) hervorgeht, muß für eine Troposphärenkorrektur der Spektren die mittlere Temperatur T_{trop} der Troposphäre bekannt sein.³⁰ Einige Autoren [59, 64] gehen davon aus, daß die mittlere Temperatur der Troposphäre relativ gut durch die Bodentemperatur abzüglich 7 K angenähert werden kann. Im aktuellen Fall wurden neben den aktuellen Bodenwerten Radiosondenprofile aus der näheren Umgebung (Stuttgart) zur Auswertung herangezogen. Infolgedessen kann der Temperaturfehler relativ klein angenommen werden mit 3 K (Standardabweichung). Dies führt auf einen relativen Fehler in Abhängigkeit

²⁸Bei den beiden CIO-Radiometern stellt sich dieses Problem nicht mehr, da beide Radiometer im Einseitenbandbetrieb arbeiten.

²⁹Spektren mit einer Untergrund-Helligkeitstemperatur von mehr als 200 K sollten nicht mehr ausgewertet werden, da die Troposphäre in diesen Fällen wenig durchlässig und der Informationsgehalt des Spektrums entsprechend gering ist. Die Temperatur der Troposphäre wurde für die obige Darstellung mit 275 K angenommen.

³⁰Im wesentlichen wird die Stärke der Absorption der Troposphäre bestimmt aus der Stärke des Strahldichteuntergrundes U , welcher ein Maß ist für die innerhalb der Troposphäre emittierte Strahlung und in etwa dem Produkt aus Absorptionskoeffizient k und Planckfunktion $B(T_{trop})$ entspricht.

von der mittleren Troposphärentemperatur von

$$\frac{L_o(T_{trop} + 3\text{K}) - L_o(T_{trop})}{L_o(T_{trop})} = \frac{B(T_{trop} + 3\text{K})[B(T_{trop}) - U]}{[B(T_{trop} + 3\text{K}) - U]B(T_{trop})} - 1.$$

Der Verlauf dieser Kurvenschar ist in Abb. 31 d für verschiedene Werte des Untergrundes U dargestellt. Z. B. führt bei einem troposphärischen Untergrund U von 165 K eine um 3 K zu hoch angenommene mittlere Troposphärentemperatur auf ein um 2 % zu niedriges Inversionsergebnis. Dabei sind die Auswirkungen umso kleiner, je geringer der Einfluß der Troposphäre, d. h. U , ist.³¹

Bisher noch nicht diskutiert wurde die Frage einer Abweichung der bei der Inversion verwendeten Liniendaten von den wahren Werten sowie die Frage der Verwendung eines falschen stratosphärischen Temperaturprofils. Auf Seiten der Liniendaten wird das primäre Interesse dabei den Auswirkungen einer falschen Linienstärke S sowie eines falschen Druckverbreiterungsparameter γ gelten, welche beide katalogisierten Spektraldaten [73]–[75] entstammen. Wie der Vergleich mit anderen Quellen zeigt, [63, 89], liegen die Unsicherheiten bzgl. der Linienstärke in der Größenordnung von 5 %. Für den Druckverbreiterungsparameter γ gibt Rothman [75] einen Fehler von ebenfalls 5 % an. Während ersterer sich jedoch multiplikativ auf das Inversionsergebnis auswirkt, bewirkt ein falscher Druckverbreiterungsparameter eine Höhenverschiebung des invertierten Profils. Für die Unsicherheit des stratosphärischen Temperaturprofils wird im folgenden ein Wert von 5 K angenommen und zwar für das gesamte Profil.

- **Auswirkungen eines falschen Druckverbreiterungsparameters γ**

Um den Einfluß eines falschen Druckverbreiterungsparameters besser untersuchen zu können, wird im folgenden von einer isothermen Atmosphäre ausgegangen, da in dieser der Luftdruck mit zunehmender Höhe nach einer einfachen Exponentialfunktion abnimmt [39, S. 126],

$$p = p_0 e^{-ch},$$

wobei $c \approx 0,125/\text{km}$ für $T = 273 \text{ K}$. Überdies ist die Linienbreite b in diesem Falle dem Druck direkt proportional, $b = \gamma p$, so daß

$$h(b) = -\frac{\ln(b/\gamma p_0)}{c} \quad (58)$$

bzw.

$$b(h) = \gamma p_0 e^{-ch}. \quad (59)$$

³¹Natürlich gilt auch in diesem Falle, daß Spektren mit einer Untergrund-Helligkeitstemperatur von mehr als 200 K nicht mehr invertiert werden sollten, da die Troposphäre zu undurchlässig ist.

Für den Fall, daß der Druckverbreiterungsparameter γ um 5% zu groß angenommen wird, ergibt sich aus (58) die neue Höhenabhängigkeit

$$h'(b) = -\frac{\ln(b/1,05\gamma p_0)}{c},$$

woraus mit Hilfe von (59) folgt

$$h'(b) = -\frac{\ln(\gamma p_0 e^{-ch}/1,05\gamma p_0)}{c} = h(b) + \frac{\ln(1,05)}{c}.$$

Die neue Höhenabhängigkeit ist gegenüber der alten um $\ln(1,05)/c \approx 0,39$ km nach oben verschoben. Entsprechend führt ein um 5% zu klein angenommener Druckverbreiterungsparameter auf eine Verschiebung des invertierten Höhenprofils um 0,41 km nach unten. Beide können gegenüber der maximal erreichbaren Höhengauflösung von 7–10 km (siehe Abb. 29) vernachlässigt werden.

- **Auswirkungen einer falsch angenommenen Linienstärke**

Wie aus einer einfachen Plausibilitätsbetrachtung hervorgeht, ist die Stärke der stratosphärischen O_3 -Signatur L_o im Grenzfall schwacher Linien im wesentlichen proportional dem Produkt aus Volumenmischungsverhältnis, Linienstärke und Planckfunktion, $L_o \propto c_v S(T)B(T)$, welche ihrerseits von der Temperatur T abhängen, so daß³²

$$c_v \propto \frac{L_o}{S(T)B(T)} \quad (60)$$

Wird die Linienstärke S bei der Inversion um 5% zu groß angenommen, folgt hieraus für das ermittelte Volumenmischungsverhältnis ein relativer Fehler von

$$\frac{c_v(1,05 S(T)) - c_v(S(T))}{c_v(S(T))} = \frac{1}{1,05} - 1 \approx -5\%.$$

Das Inversionsverfahren reagiert auf Fehler in der Linienstärke also äußerst empfindlich, viel stärker als auf eine falsch angenommene Stratosphärentemperatur, wie die folgende Betrachtung zeigt.

- **Auswirkungen eines falschen stratosphärischen Temperaturprofils**

Für jede Temperatur T läßt sich die relativ komplexe Temperaturabhängigkeit der Linienstärke $S(T)$ (siehe z. B. [14, 89]) vereinfacht darstellen als³³

$$S(T) = aT^m \quad \text{wobei} \quad \begin{cases} m = E_n/k_B T - 2 & \text{für symmetrische Moleküle} \\ m = E_n/k_B T - 2,5 & \text{für asymmetrische Moleküle.} \end{cases}$$

³²Dabei wurde die schwache Nichtlinearität der SÜG (Eigenabsorption von Ozon in der Stratosphäre) vernachlässigt. Diese Vereinfachung sei für eine Fehlerbetrachtung gestattet.

³³Approximation der Funktion $S(T)$ durch das besonders einfache Polynom aT^m . Aus den beiden Gleichungen $S(T) = aT^m$ sowie $\partial S/\partial T = maT^{m-1}$ folgt schließlich $m = T/S \partial S/\partial T$ und hieraus nach etwas Rechnung der obige Zusammenhang.

Hierbei ist a eine Konstante, E_n die Energie des unteren Anregungsniveaus und k_B die Boltzmannkonstante. Den obigen Ausdruck eingesetzt in die Gl. (60) führt auf eine Temperaturabhängigkeit des Volumenmischungsverhältnisses von³⁴

$$c_v \propto T^{-m-1}.$$

Hieraus ergibt sich für den relativen Fehler aufgrund eines um 5 K zu hoch angenommenen stratosphärischen Temperaturprofils

$$\frac{c_v(T + 5 \text{ K}) - c_v(T)}{c_v(T)} = \left[\frac{T + 5 \text{ K}}{T} \right]^{-m-1}.$$

Im Falle der 142,175 GHz-Linie von Ozon, einem asymmetrischen Molekül, folgt bei einer mittleren Temperatur der Stratosphäre von 250 K für m ein Wert von $-2,23$. Die Linienstärke ist also relativ schwach temperaturabhängig. Entsprechend gilt für das invertierte Volumenmischungsverhältnis $c_v \propto T^{-1,23}$, woraus sich ein relativer Fehler von etwa 2,5 % ergibt.

In Tabelle 3 sind die relevanten Fehlerquellen samt ihren Auswirkungen auf das Inversionsergebnis noch einmal zusammengefaßt. Wie bereits erwähnt, zieht die Unsicherheit in der Linienstärke dabei den größten Fehler nach sich. An zweiter Stelle folgt die Unsicherheit in der Gewichtung der beiden Seitenbänder (dieses Problem wird sich in Zukunft bei den beiden ClO-Radiometern nicht mehr stellen, da diese im Einseitenbandbetrieb arbeiten) und erst an dritter Stelle die Auswirkungen eines falschen stratosphärischen Temperaturprofils. Sicherlich werden in naher Zukunft erhebliche Verbesserungen bzgl. der Einstellung des Elevationswinkels sowie der Kalibration erreicht werden, doch sind von diesen Verbesserungen in Anbetracht der großen Unsicherheit der Linienstärke nur geringfügige Verbesserungen des Gesamtfehlers zu erwarten.

Indem man die Gl. (57) in die Gl. (60) einsetzt, kann man auch einen Ausdruck herleiten, an dem alle oben diskutierten Fehlerquellen mit Ausnahme des Druckverbreiterungsparameters γ gemeinsam studiert werden können.³⁵ Insbesondere kann auf diesen Ausdruck das Gaußsche Fehlerfortpflanzungsgesetz angewendet werden, was aus Gründen der besseren Übersichtlichkeit hier jedoch nicht geschehen soll. Statt dessen wird lediglich das Endergebnis angegeben.

Hiernach ergibt sich für die in Tabelle 3 angegebenen Parameter ein systematischer Fehler von 7 %. Dieser Wert ist typisch für die während der Meßtage im März und

³⁴Dabei wurde im wesentlichen Gebrauch gemacht von der Tatsache, daß für die Planckfunktion im Millimeterwellenbereich die Rayleigh-Jeans-Näherung gilt, daß also die Planckfunktion im wesentlichen der Temperatur proportional ist.

³⁵Der Ausschluß des Druckverbreiterungsparameters bedeutet keine wesentliche Einschränkung, da dessen Auswirkungen auf die invertierten Profile gegenüber der maximal erreichbaren Höhenauflösung vernachlässigt werden können.

Eingabeparameter	Wert	Unsicherheit	rel. Fehler
Elevationswinkel	29,5°	0,5°	1,5 %
Gewicht	0,56	0,02	3,6 %
Troposphärentemperatur	270 K	3 K	0,4 %
Stratosphärentemperatur	250 K	5 K	2,5 %
Untergrund	75 K		
Linienstärke		5 %	5,0 %
Kalibration		2 K	1,0 %
gesamt ³⁶			6,9 %

Tabelle 3: Liste der systematischen Fehler.

April 1994 auf der Hornisgrinde herrschenden Wetterverhältnisse (der systematische Fehler bewegte sich zwischen 6,7 und 10 %) und muß zu dem statistischen Fehler, der bereits aus dem Meßfehlerauschen resultiert, quadratisch addiert werden. In der Abb. 28 des vorhergehenden Abschnitts ist dieser systematische Fehler gleich mit berücksichtigt.

5.3 Inversion von Chlormonoxid bei 278,631 GHz

Im folgenden sollen noch einige Untersuchungen zum ClO-Radiometer bei 278 GHz durchgeführt werden. Letzteres befindet sich momentan in der Testphase, so daß für die Auswertung bisher keine gemessenen ClO-Spektren zu Verfügung stehen. ³⁷ Stattdessen erfolgen die Untersuchungen auf der Basis synthetisch erzeugter Spektren. Insbesondere soll in diesem Zusammenhang die Frage geklärt werden, inwieweit das Spurengas ClO überhaupt von der Hornisgrinde aus detektiert werden kann, oder ob für künftige Testmessungen von vornherein auf höhergelegene Standorte ausgewichen werden sollte. ³⁸

Wie sich aus der geringen Intensität der ClO-Signatur bei 278,631 GHz unmittelbar schließen läßt (siehe Abb. 39 in Anhang B.2), stellt die Messung dieses Spurengases höchste Anforderungen an die zur Verfügung stehende Meßtechnik. Für die Auswertung resultiert aus dieser geringen Intensität der Linie bei gleicher Integrationszeit wie bisher ein wesentlich ungünstigeres Signal-zu-Rausch-Verhältnis. Gerade

³⁶Dieser Wert entspricht der Quadratwurzel aus der Summe der Quadrate der Einzelfehler.

³⁷In ersten Messungen mit dem neuen ClO-Radiometer konnten zumindest die beiden starken Linien von Ozon bei 273,051 GHz und 274,478 – bei allerdings sehr ungünstigen Witterungsverhältnissen vom Forschungszentrum (100 m N.N.) aus – gut detektiert werden.

³⁸Seit März 1994 befindet sich zu Testzwecken ein Meßcontainer auf der Hornisgrinde (Nordschwarzwald, 1165 m N.N.). Dieser Standort wurde gewählt aufgrund seiner geringen räumlichen Entfernung zum Forschungszentrum Karlsruhe. Sobald das Meßgerät jedoch genügend ausgereift ist, sollen Messungen an hochalpinen Standorten oder in der Nähe des Polarkreises erfolgen.

ein günstiges Signal–zu–Rausch–Verhältnis ist jedoch mit die wichtigste Voraussetzung für eine erfolgreiche Inversion der Spektren (siehe die Ausführungen der Abschnitte 5.1.2 und 5.1.4). Ziel der Messungen muß es deshalb sein, Spektren mit einem möglichst günstigen Signal–zu–Rausch–Verhältnis zu liefern.

Im Prinzip kann eine Verbesserung des Signal–zu–Rausch–Verhältnisses auf zwei-erlei Wegen erreicht werden, zum einen durch eine Vergrößerung des Signals bei unverändertem Rauschen und zum anderen durch eine Reduktion des Rauschens bei unverändertem Signal. Der ersten Möglichkeit, der Realisierung einer stärkeren ClO–Signatur, sind mit der Auswahl geeigneter, möglichst hochgelegener Standorte (im aktuellen Fall Hornisgrinde), der Ausnutzung besonders günstiger Witterungsverhältnisse sowie der Einstellung eines optimalen Elevationswinkels relativ enge Grenzen gesetzt.³⁹ Eine deutliche Verbesserung des Signal–zu–Rausch–Verhältnisses ist deshalb in erster Linie über eine Verringerung des Rauschens aufgrund rauschärmerer Empfänger oder aufgrund längerer Integrationszeiten zu erreichen. Um letztere nicht allzu groß werden zu lassen (de Zafra [95] und Solomon [78] berichten von Integrationszeiten bis zu 36 bzw. bis zu 55 Stunden), sei im folgenden von besonders günstigen Witterungsverhältnissen ausgegangen, wie sie an kalten, trockenen Wintertagen auftreten. Zu diesem Zweck wurde als Wasserdampfprofil das Radiosondenprofil eines relativ kalten Januartages verwendet, aufgenommen am 30. Januar 1992 über der Hornisgrinde im Rahmen der REKLIP–Kampagne [21, 37], und nach oben hin um die Standardatmosphäre für mittlere Breiten ergänzt. Insgesamt betrug der Wasserdampfgehalt über der Hornisgrinde dabei 3 mm (ausfällbare Wassermenge) gegenüber 9 mm unter den Bedingungen der Standardatmosphäre für mittlere Breiten. Hieraus ergibt sich ein optimaler Elevationswinkel von etwa 15°.

Für die nachfolgenden Betrachtungen wurde der Spektralbereich zwischen 277,9 und 279,1 GHz ausgewählt (siehe Abb. 39 in Anhang B.2). Dieser zeichnet sich durch eine verhältnismäßig kräftige Signatur von ClO aus.⁴⁰ Für die Auswertung kommt erschwerend hinzu, daß sich diese ClO–Signatur auf der Flanke einer Ozon–Linie bei 278,522 GHz befindet, welche unter den Bedingungen der Standardatmosphäre um etwa einen Faktor 7 stärker ist. Es wurde deshalb versucht, Ozon und ClO simultan

³⁹Die Existenz eines optimalen Elevationswinkels ist unmittelbar einleuchtend. Wird das Radiometer unter einem großen Elevationswinkel auf die Atmosphäre gerichtet (z. B. Blickrichtung senkrecht nach oben), ist das stratosphärische Signal aufgrund der geringen absorbierenden Masse entlang des Sehstrahls relativ klein. Richtet man das Radiometer deshalb unter einem flacheren Winkel auf die Atmosphäre, nimmt die absorbierende Masse der Stratosphäre zwar zu, andererseits wird die Troposphäre zunehmend undurchlässiger und im Grenzfall eines sehr flachen Winkels praktisch vollkommen dicht. Dazwischen existiert ein optimaler Elevationswinkel (bei welcher Einstellung dieser im konkreten Einzelfall liegt, hängt unter anderem ab von den herrschenden Witterungsverhältnissen d. h. dem Wasserdampfgehalt der Troposphäre), bei dem die zum Boden gelangende stratosphärische Signatur maximal wird.

⁴⁰Kräftig, gemessen an den Linien zweier benachbarter Übergänge bei 204 und 241 GHz, die beide um einen Faktor 2–3 schwächer sind. Allerdings wird dieser Nachteil größtenteils kompensiert durch die höhere Durchlässigkeit der Troposphäre in diesen Frequenzbereichen.

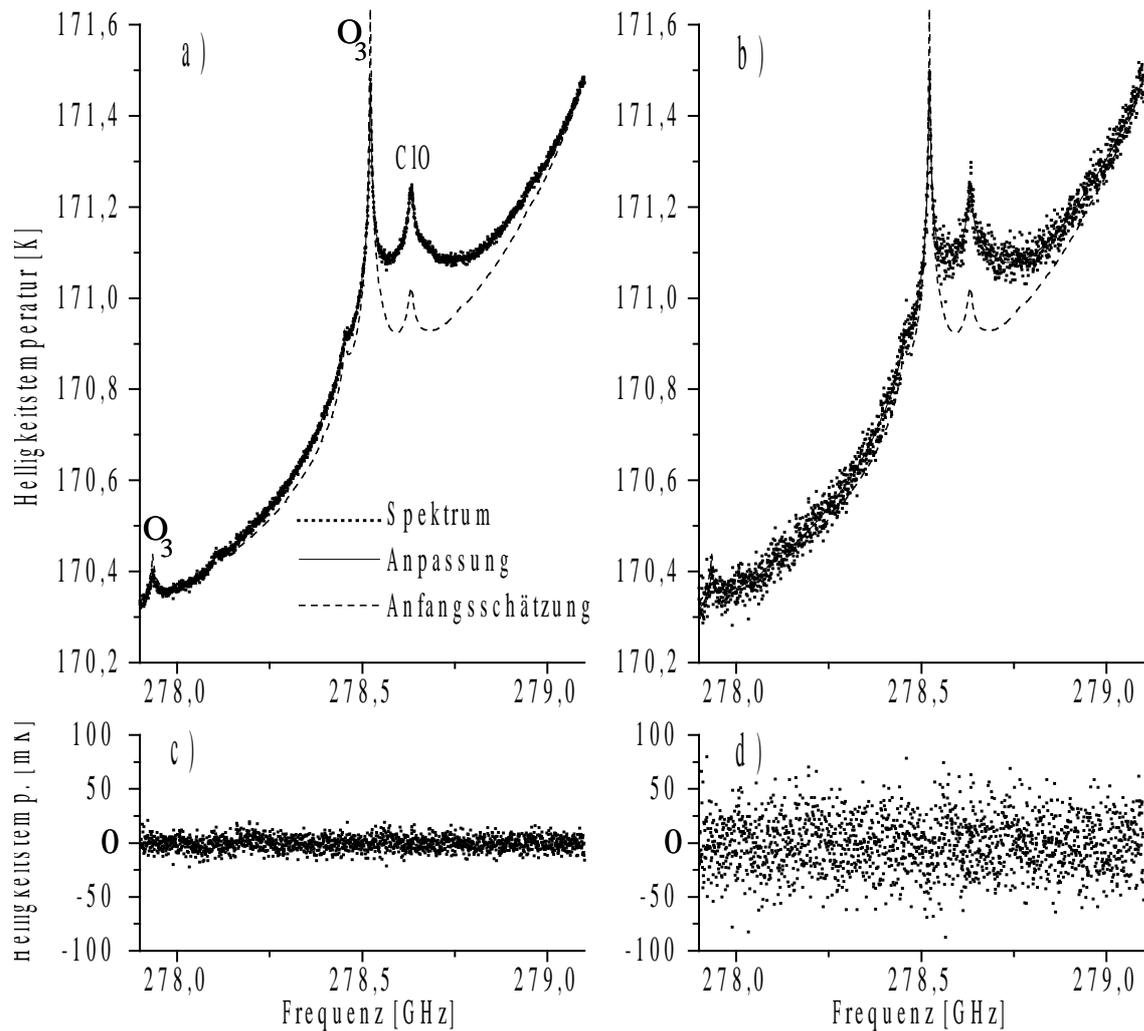


Abbildung 32: In a) mit einem Rauschen der Standardabweichung $5/1000$ K je Kanal der Bandbreite $0,7$ MHz überlagertes Szenario, in b) entsprechendes, mit einer Standardabweichung von $2/100$ K verrauschtes Spektrum. Zusätzlich sind in a) und b) die kompletten Anpassungen der Tikhonov–Phillips–Regularisierung mit dargestellt (durchgezogene Linien). In c) und d) ist die Differenz zwischen dem jeweils vorgegebenen Spektrum und der kompletten Anpassung aufgetragen.

zu invertieren.

Den Ausgangspunkt der Simulation bildeten die beiden – in Anlehnung an Messungen von Connor [15] und de Zafrá [96] in der Antarktis – willkürlich vorgegebenen Höhenprofile von Ozon und ClO (durchgezogene Linien in Abb. 33). Gegenüber der Standardatmosphäre für mittlere Breiten (gepunktete Linien) lassen diese auf eine

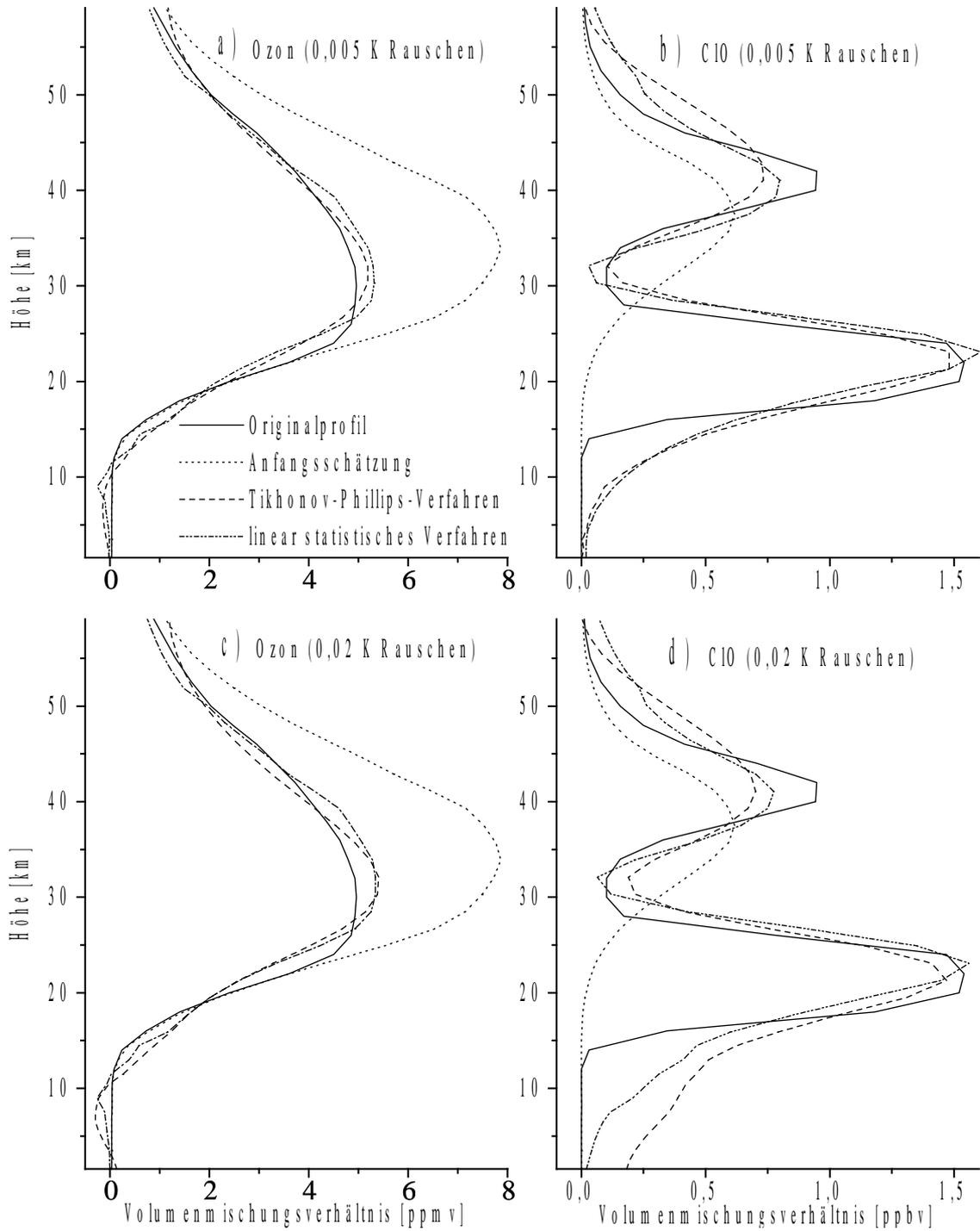


Abbildung 33: Gleichzeitige Inversion von Ozon und ClO aus einem verrauschten Spektrum. In a) und b) 5/1000 K Rauschen, in c) und d) 2/100 K Rauschen.

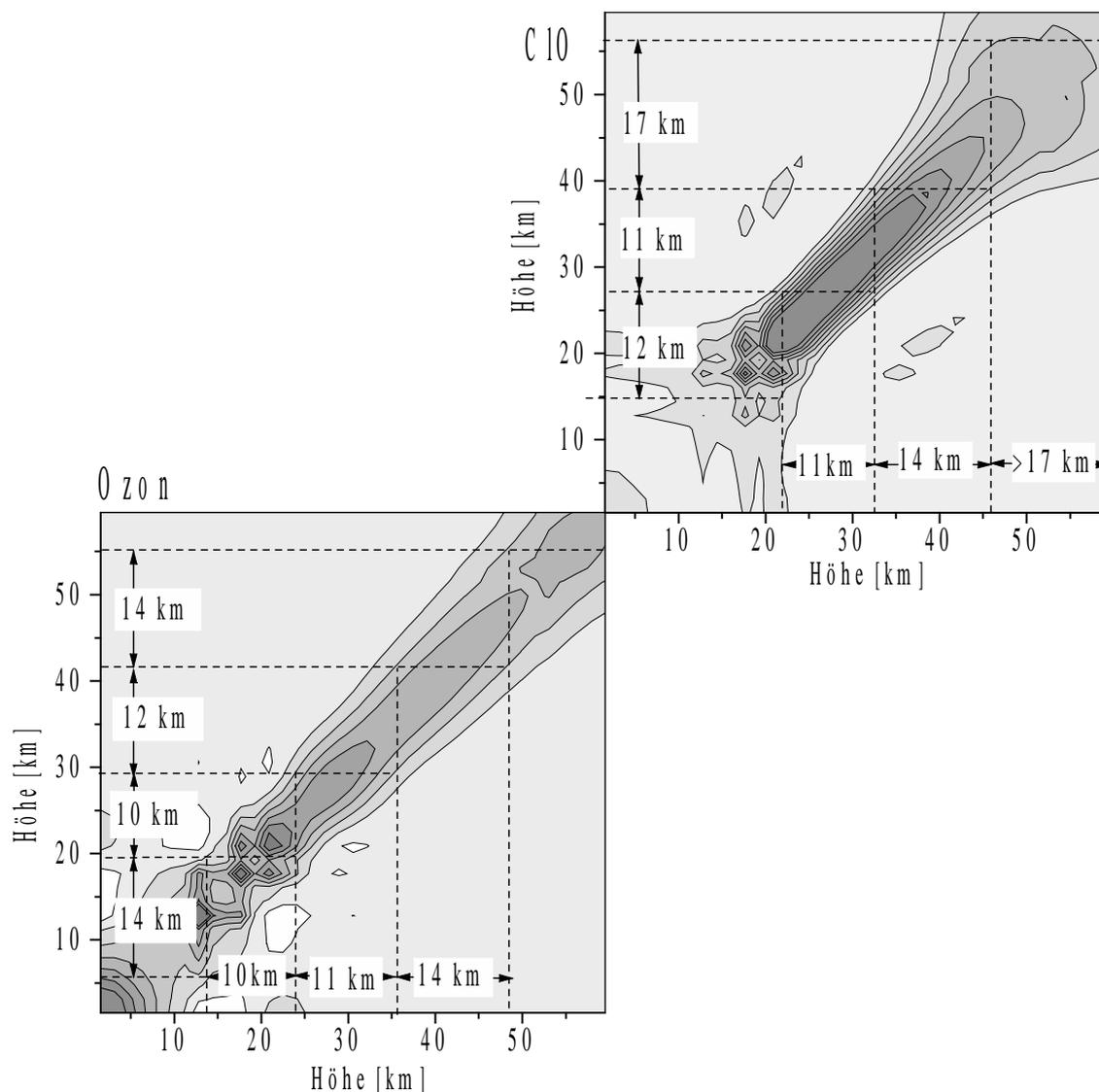


Abbildung 34: Auflösungsmatrix des Tikhonov–Phillips–Verfahrens bei einer gleichzeitigen Inversion von Ozon und CIO aus einem mit 2/100 K verrauschten synthetischen Spektrum.

deutlich gestörte Atmosphärenchemie schließen. Insbesondere weist das CIO–Profil jene für den antarktischen Ozonabbau im Frühling typische Zweischichtenstruktur auf. Aus diesen beiden Profilen wurde unter Berücksichtigung des obigen Wasserdampfprofils ein synthetisches Strahldichtespektrum berechnet, welches dann – zusätzlich überlagert mit einem Rauschen der Amplitude 5/1000 bzw. 2/100 K – als Eingabe für die nachfolgenden Inversionsrechnungen diente (siehe Abb. 32 a und b).

Als Anfangsschätzungen wurden die gepunkteten Linien der Abb. 33 verwendet, welche der Standardatmosphäre für mittlere Breiten entstammen und auf das gepunktete Strahldichtespektrum der Abb. 32 a und b führten. Verglichen mit dem Spektrum der Standardatmosphäre für mittlere Breiten war die ClO-Signatur in den zwei zu invertierenden Spektren also erheblich stärker ausgeprägt. Die gestrichelten und gestrichpunkteten Linien der Abb. 33 schießlich geben die mit der Tikhonov-Phillips-Regularisierung sowie dem linear statistischen Verfahren invertierten Höhenprofile wieder. Während Ozon in beiden Fällen (5/1000 und 2/100 K) relativ gut invertiert werden konnte (Abb. 33 a und c), ergeben sich für ClO doch mehr oder weniger große Abweichungen von dem wahren Profil, vor allem im Höhenbereich unterhalb von 15 km.

Erfreulicherweise kann jedoch selbst bei einem Rauschen von 2/100 K – dies entspricht einem Signal-zu-Rausch-Verhältnis von 25:1 (Peak:Rauschen) bezogen auf eine Kanalbreite von 0,7 MHz – noch die Zweischichtenstruktur des ClO-Profiles aufgelöst werden. Das Auflösungsvermögen beträgt demnach mindestens drei Höhenschichten, was durch die Auflösungsmatrix der Abb. 34 im wesentlichen bestätigt wird. Hiernach ergibt sich für Ozon ein maximales Höhenauflösungsvermögen von 10–11 km in 30 km Höhe, entsprechend vier Höhenschichten und für ClO ein maximales Höhenauflösungsvermögen von 11 km in 30 km Höhe, entsprechend etwas mehr als drei Höhenschichten. Interessant ist auch die Aussage der Auflösungsmatrix, daß das Spurengas ClO unterhalb von 15 km nicht mehr invertiert werden kann (das Band der ClO-Auflösungsmatrix hört bei etwa 15 km auf bzw. wird so breit, daß von einer Höhenauflösung nicht mehr gesprochen werden kann). Dies steht im Einklang mit den Inversionsergebnissen der Abb. 33 d, welche unterhalb von 15 km deutlich von dem wahren Profil abweichen. Die Anpassungen der Tikhonov-Phillips-Regularisierung sowie des linear statistischen Verfahrens stimmen in beiden Fällen sehr gut mit den vorgegebenen Spektren überein (siehe Abb. 32).

An a priori Information wurde im Falle der Tikhonov-Phillips-Regularisierung verwendet, daß die Höhenprofile möglichst glatt sein sollten, wobei die Wahl des Regularisierungsparameters γ a posteriori erfolgte. Im Falle des linear statistischen Verfahrens hingegen dienten die beiden Anfangsschätzungen als a priori-Volumenmischungsverhältnisprofile und die a priori-Kovarianzmatrizen \mathbf{S}_{fO_3} und \mathbf{S}_{fClO} wurden als Diagonalmatrizen mit den Werten $0,25 \cdot 10^{-12}$ und $9 \cdot 10^{-20}$ initialisiert (dies entspricht einer Standardabweichung der a priori-Höhenprofile von O_3 und ClO von $\pm 0,5$ ppmv bzw. $\pm 0,3$ ppbv). Die Fehlerkovarianzmatrix $\mathbf{S}_{\mathcal{E}}$ wurde als reine Diagonalmatrix mit den Werten $2,5 \cdot 10^{-5} K^2$ und $4 \cdot 10^{-4} K^2$ angenommen, entsprechend einem Rauschen von 5/1000 bzw. 2/100 K.

Zusammenfassend läßt sich feststellen, daß das Spurengas ClO von der Hornisgrinde aus nur unter extrem günstigen Witterungsverhältnissen detektiert werden kann, d. h. nur an sehr kalten, trockenen Wintertagen, wenn der Wasserdampfgehalt der

Atmosphäre über der Hornisgrinde nicht mehr als 3 mm beträgt (ausfällbare Wassermenge). Überdies sollte bei diesen Messungen ein Signal-zu-Rausch-Verhältnis von mindestens 25:1 erreicht werden, damit wenigstens drei Höhengschichten aufgelöst werden können. Im gerade betrachteten Fall (gestörte Atmosphäre, d. h. entsprechend erhöhte ClO-Konzentration) würde dies einem Rauschen von 2/100 K bzw. einer Integrationszeit von knapp drei Stunden entsprechen. Hingegen würde unter den Bedingungen der Standardatmosphäre (bei unverändert geringem Wasserdampfgehalt) die ClO-Signatur von vornherein um einen Faktor 2–4 schwächer ausfallen (siehe Abb. 32). In diesem Fall müßte ein Rauschen von 5/1000 K realisiert werden, um ein Signal-zu-Rausch-Verhältnis 25:1 zu erreichen. Dies würde einer Integrationszeit von 45 Stunden, d. h. zwei Tagen entsprechen.

Kapitel 6

Zusammenfassung

Mit den entwickelten Spurengas-Radiometern im Millimeterwellenbereich stehen im Forschungszentrum günstige Meßmöglichkeiten für eine zeitlich weitgehend kontinuierliche Messung von Höhenprofilen zur Verfügung. Ziel der vorliegenden Arbeit war es, die für die Berechnung von Konzentrationsprofilen aus den gewonnenen Strahllichtespektren benötigten Auswerteprogramme bereitzustellen.

Hierzu wurde aufbauend auf einem empirischen Rechenmodell von H. J. Liebe [42]–[47] und unter Einbeziehung der HITRAN- und AFGL-Daten zunächst ein effizientes Linie-für-Linie Vorwärtsprogramm entwickelt, mit dem die entsprechenden Intensitätsspektren berechnet werden konnten. In internationalen Vergleichsrechnungen erwies sich dieses Programm als sehr leistungsfähig (siehe Anhang A.1).

Sodann wurde das Problem des Auffindens einer bestimmten Lösung der nichtlinearen SÜG reduziert auf das Lösen eines linearen Gleichungssystems. Dieses wurde erreicht über eine Taylorreihenentwicklung der nichtlinearen SÜG um ein angenommenes Startprofil, wobei nur die erste Ordnung mit berücksichtigt wurde. Durch sukzessives Lösen dieses linearisierten Problems konnte iterativ die Lösung der nichtlinearen SÜG ermittelt werden (Gauß-Newton-Iteration). Eine gewisse Besonderheit stellt in diesem Zusammenhang die Art und Weise dar, auf die die Sensitivitätsfunktionen im Programm implementiert wurden. Deren Berechnung erfolgte auf weitgehend analytischem Wege, so daß die Sensitivitätsfunktionen bei der Vorwärtsrechnung gleich mit bestimmt werden konnten. Dies hatte eine enorme Beschleunigung des Auswerteprogrammes zur Folge.

Einige Schwierigkeiten bei der Lösung des obigen linearen Gleichungssystems ergaben sich aus der Tatsache, daß das Gleichungssystem schlecht gestellt war im Sinne von Hadamard [30] und einer besonders sorgfältigen numerischen Behandlung bedurfte. Zunächst wurde deshalb ein Einblick in die Theorie der Inversion gegeben und anschließend drei Regularisierungsverfahren vorgestellt, welche sich bereits in einer früheren Arbeit als für die Auswertung von bodengebundenen Messungen prinzipiell geeignet herauskristallisiert hatten [40]: die abgeschnittene Singulärwertzerlegung, die Tikhonov-Phillips-Regularisierung sowie das linear statistische Verfahren.

Neu an dieser Arbeit war die Erweiterung der beiden letztgenannten Regularisierungsverfahren auf den Fall der simultanen Inversion mehrerer Spurengase aus einem Spektrum. Dies ist insbesondere dann von Bedeutung, wenn in den Spektren

– wie im Falle der 278 GHz–Linie von ClO – der Signatur des primär interessierenden Gases die Signatur eines weiteren Gases – in diesem Falle Ozon – überlagert ist. Desweiteren wurde eine Erweiterung der Regularisierungsverfahren durchgeführt, die eine Einbeziehung der gerätebedingten stehenden Wellen in den Inversionsalgorithmus ermöglicht. Mit stehenden Wellen überlagerte Spektren konnten auf diese Weise direkt invertiert werden, ohne sie vorher einer aufwendigen Elimination von stehenden Wellen unterziehen zu müssen.

Die Implementierung der beiden erweiterten Regularisierungsverfahren in ein Programm erfolgte so, daß während eines jedes Iterationsschritts wahlweise zwischen der Tikhonov–Phillips–Regularisierung und dem linear statistischen Verfahren zur Lösung des schlecht gestellten linearen Gleichungssystems hin– und hergeschaltet werden konnte. Dies hatte den Vorteil, daß die Inversionsergebnisse der beiden Verfahren unmittelbar miteinander verglichen werden konnten. Auch die abgeschnittene Singulärwertzerlegung wurde mit implementiert; allerdings wurden für dieses Inversionsverfahren nur wenige Rechnungen durchgeführt, da es nicht gelang, die abgeschnittene Singulärwertzerlegung entsprechend den beiden obigen Verfahren zu erweitern.

In einer Reihe von Inversionrechnungen konnte auf der Basis synthetisch erzeugter Spektren die Leistungsfähigkeit der erweiterten Regularisierungsverfahren eindrucksvoll demonstriert werden. Beide erwiesen sich hierbei als überaus nützliche Inversionswerkzeuge, mit denen es selbst stark mit stehenden Wellen gestörte Spektren zu invertieren gelang. Probleme bei der Auswertung ergaben sich allenfalls dann, wenn die Spektren entweder zu stark verrauscht waren oder die in den Inversionsalgorithmus hineingesteckten Perioden von den Perioden der den Spektren überlagerten stehenden Wellen zu sehr abwichen. Unter anderem wurde mit dem entwickelten Auswertprogramm auch an einem Vergleich von Inversionsprogrammen teilgenommen, wobei dieses vergleichsweise sehr gut abschnitt (siehe Anhang A.2).

Schließlich wurde eine Anzahl gemessener Spektren von Ozon invertiert. Hierbei ergab sich eine gute Übereinstimmung der invertierten Höhenprofile mit den entsprechenden Radiosondenprofilen. Für die Auswertung wurde überdies eine detaillierte Fehlerrechnung durchgeführt; dabei erwies sich die nicht genau bekannte Linienstärke als der größte Unsicherheitsfaktor.

Den Abschluß bildete eine Untersuchung auf der Basis synthetisch erzeugter Spektren zur Meßbarkeit des Spurengases ClO vom Standort Hornisgrinde. Es stellte sich heraus, daß das Spurengas ClO prinzipiell auch von der Hornisgrinde aus beobachtet werden kann, jedoch nur bei extrem günstigen Witterungsverhältnissen. Um eine akzeptable Höhenauflösung zu erreichen (wenigstens drei Höhenschichten) ist dabei mit Integrationszeiten zwischen drei Stunden und zwei Tagen zu rechnen.

Anhang A

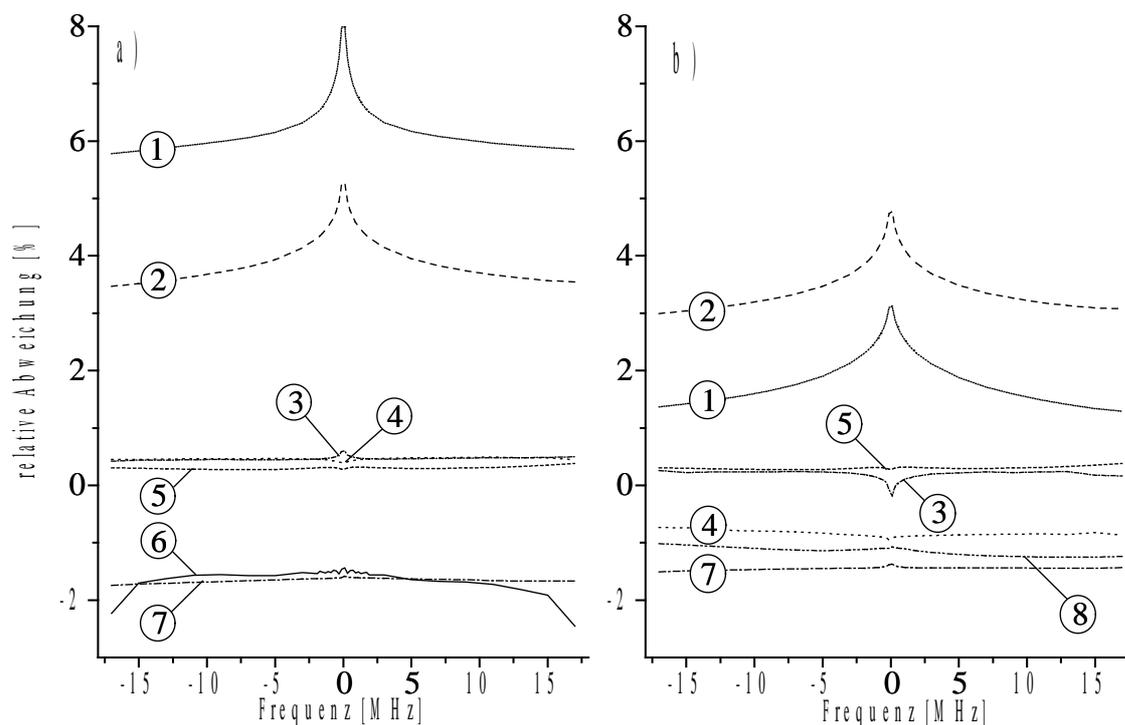
Vergleichsrechnungen

A.1 Rechnungen zum Vorwärtsproblem

Leider liefern die Untersuchungen des Kapitels 5 keinen Aufschluß darüber, wie gut das entwickelte Vorwärtsprogramm die Strahlungsübertragung innerhalb der Atmosphäre modelliert. Z. B. bestehen Unsicherheiten bzgl. der für die Vorwärtsrechnung verwendeten Molekülparameter (Linienstärken, Druckverarbeitungsparameter etc.), welche katalogisierten Spektraldaten [73]–[75] entstammen. Diese Frage kann im Prinzip nur durch einen direkten Vergleich der theoretisch berechneten mit gemessenen Strahldichtespektren beantwortet werden, wenngleich große Unsicherheiten bestehen bzgl. des Zustands der Atmosphäre zum Zeitpunkt der Messung, welcher für einen solchen Vergleich bekannt sein muß.

Große Bedeutung kommt deshalb internationalen Vergleichsrechnungen zu, in welchen ausgehend von einem wohldefinierten Zustand der Atmosphäre Transmissions- und Strahldichteberechnungen mit den verschiedenen Programmen durchgeführt werden. Diese liefern wichtige Erkenntnisse über Vorzüge, Schwächen und Charakteristika der verwendeten Modelle. Überdies kann ein erfolgreiches Abschneiden bei einem solchen Vergleich als Validierung des eigenen Programmes angesehen werden, da in diesen das Wissen gleich mehrerer Arbeitsgruppen mit einfließt. Diesem Zweck diene auch die Teilnahme an einem internationalen Vergleich von Vorwärtsprogrammen, initiiert von G. E. Nedoluha und R. M. Bevilacqua vom Naval Research Laboratory in Washington D. C., an welchem sich insgesamt acht Gruppen beteiligten.

Zu berechnen war das Strahldichtespektrum der 22,236 GHz-Linie von Wasserdampf für eine Beobachterhöhe von 20 km und einen Elevationswinkel von 20° (siehe die Abb. 36 links). Hierzu wurden den Teilnehmern in einem ersten Test die Absorptionskoeffizienten für jede Höhenschicht und jede Frequenz zur Verfügung gestellt. Im Prinzip brauchte also nur noch über die Strahlungsübertragungsgleichung integriert zu werden, natürlich unter Berücksichtigung des vorgegebenen Elevationswinkels bei der Berechnung der Schichtmittelwerte (siehe Abschn. 2.1.3). Ziel dieses Schrittes war es, zunächst die Aufintegration der verschiedenen Programme untereinander zu vergleichen, frei von den bei der Berechnung der Absorptionskoeffizienten entstehenden Abweichungen aufgrund unterschiedlicher Liniendaten. In einem



- ① Ricaud, de la Noe, Observatoire Bordeaux
- ② Schrader, Olivero, State University of Pennsylvania
- ③ Kuntz, Hochschild, Forschungszentrum Karlsruhe
- ④ Peter, Kämpfer, Universität Bern
- ⑤ Nedoluha, Bevilacqua, Naval Research Lab. Washington
- ⑥ de Zafra, Parrish, University New York
- ⑦ Raffalski, Künzi, Universität Bremen
- ⑧ O'Connor, USA, Institut nicht bekannt

Abbildung 35: Internationaler Vergleich von Vorwärtsrechnungen am Beispiel der 22,235 GHz-Linie von Wasserdampf. Dargestellt sind jeweils die relativen Abweichungen vom gemeinsamen Mittelwert aller plausibel erscheinenden Vorwärtsrechnungen. In a) relative Abweichungen bei Verwendung von für jede Höhengschicht und Frequenz vorgegebenen Absorptionskoeffizienten. In b) relative Abweichungen bei Verwendung eigener Absorptionskoeffizienten. Beide Fälle wurden gerechnet für eine Höhe von 20 km und einen Elevationswinkel von 20°.

zweiten Schritt sollten die Teilnehmer die Absorptionskoeffizienten aus einem vorgegebenen Wasserdampf-Volumenmischungsverhältnisprofil dann selbst berechnen. In

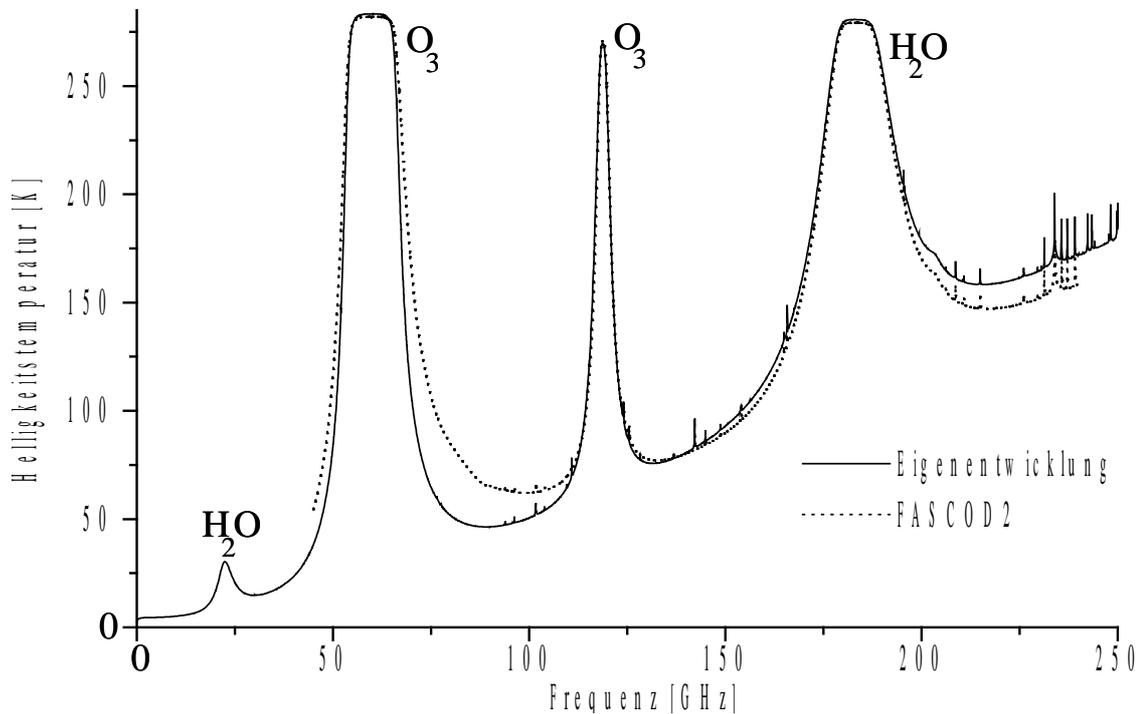


Abbildung 36: Vergleich der Eigenentwicklung mit FASCOD2. Gerechnet wurde das Millimeterwellenspektrum in Meereshöhe unter Bedingungen der Standardatmosphäre (mittlere Breiten) und einen Elevationswinkel von 90° .

beiden Fällen erwies sich das neu entwickelte Vorwärtsprogramm als außerordentlich leistungsfähig (siehe Abb. 35). Die relativen Abweichungen vom gemeinsamen Mittelwert aller plausibel erscheinenden Vorwärtsrechnungen betragen jeweils weniger als 0,5 %.

Darüberhinaus wurden Vergleiche mit dem international verwendeten Programm FASCOD2 angestellt [13]. Gerechnet wurde der Bereich des Millimeterwellenspektrums zwischen 50 und 250 GHz. Hierbei ergab sich eine gute Übereinstimmung bzgl. der Feinstruktur des Spektrums. In ihrer Grobstruktur wichen die Helligkeitstemperaturverläufe beider Programme geringfügig voneinander ab. Dieses war nicht anders zu erwarten, da die Eigenentwicklung und FASCOD2 eine unterschiedliche Linienformfunktion und unterschiedliche Formulierungen für das Wasserdampf-Kontinuum verwenden. In Anlehnung an Liebe [42]–[47] wurde in der Eigenentwicklung ein an Messungen empirisch angepaßtes Wasserdampf-Kontinuum und eine nach Rosenkranz modifizierte Van-Vleck-Weißkopf-Formfunktion [72] verwendet. Manabe et alii [51] und Smith [77] bescheinigen diesem Modell eine gute Übereinstimmung mit dem Experiment im Millimeterwellenbereich .

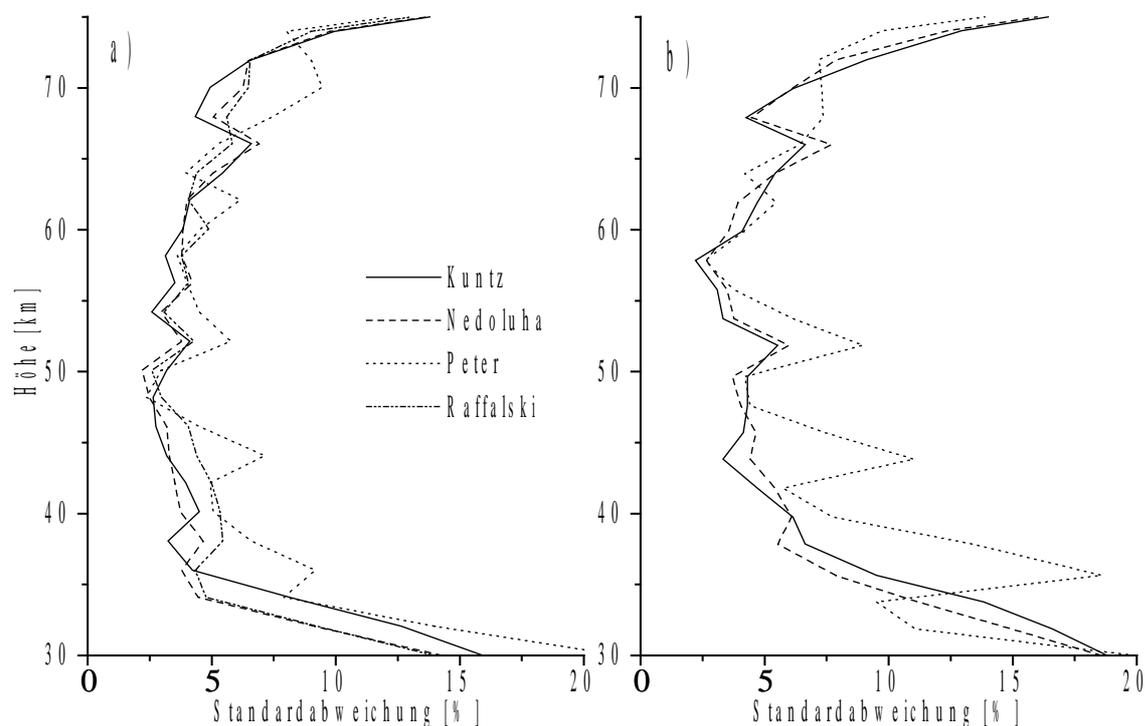


Abbildung 37: Internationaler Vergleich von Inversionsrechnungen. Aufgetragen ist die Streuung (Standardabweichung) der invertierten Profile um die Originalprofile, $\sigma(\mathbf{f}_{\text{inv}} - \mathbf{f}_{\text{wahr}})$. In a) Streuung der Inversionsergebnisse der lediglich verrauschten Spektren, in b) Streuung der Inversionsergebnisse der zusätzlich mit einer stehenden Welle überlagerten Spektren.¹

A.2 Rechnungen zum inversen Problem

Im Anschluß an die erfolgreiche Teilnahme an dem Vergleich der Vorwärtsrechnungen bestand die Möglichkeit zu einem Vergleich der Inversionsprogramme, an welchem sich insgesamt vier Gruppen beteiligten. Zu diesem Zweck wurden einem jeden Teilnehmer vierzig synthetische berechnete, leicht verrauschte Spektren übersandt (als Beispiel diente wieder die 22,235 GHz-Linie von Wasserdampf), welche es zu invertieren galt.

¹Weder G. Nedoluha, R. Peter noch U. Raffalski verfügten über ein Programm zur gleichzeitigen Anpassung von stehenden Wellen, so daß letzterer sich an der Inversion der mit einer stehenden Welle überlagerten Spektren überhaupt nicht mehr beteiligte. R. Peter invertierte diese Spektren, jedoch ohne die Überlagerung der stehenden Wellen bei der Auswertung in irgendeiner Form mit zu berücksichtigen; dies erklärt denn auch dessen vergleichsweise schlechtes Abschneiden in der Abb. 37 b). G. Nedoluha berücksichtigte die Überlagerung der Spektren mit einer stehenden Welle hingegen in bei der Initialisierung seiner Fehler-Kovarianzmatrix – er arbeitete nach dem Rodgers-Verfahren – und erzielte damit gute Ergebnisse [56].

Von diesen vierzig Spektren waren – um die Auswertung noch etwas zu verkomplizieren – zwanzig zusätzlich mit einer stehenden Welle der Periode 104,72 MHz überlagert; allerdings war die Amplitude der stehenden Welle sehr klein, ≈ 15 mK, so daß die Spektren näherungsweise auch ohne die Berücksichtigung der stehenden Welle invertiert werden konnten.

In Abb. 37 sind die Ergebnisse dieser Vergleichrechnungen dargestellt (Vorsicht, aufgetragen sind keine Höhenprofile, sondern die Streuungen der invertierten Höhenprofile um die wahren Werte). Dabei schnitt das entwickelte Inversionsprogramm in beiden Fällen (mit und ohne überlagerte stehende Welle) sehr gut ab. Die invertierten Profile streuten in beiden Fällen im interessierenden Höhenbereich zwischen 40 und 65 km um weniger als 5 % um die wahren Profile, im Einklang mit den Inversionsergebnissen der anderen Teilnehmer.²

²Praktisch mit Abschluß dieser Arbeit wurde von G. Nedoluha im Internet eine Homepage unter der Adresse <http://wvms.nrl.navy.mil/NDSC/ndsc.html> eingerichtet, unter der die Ergebnisse der Vergleichsrechnungen eingesehen werden können.

Anhang B

Ergänzungen

B.1 Berechnungen

B.1.1 Zur Berechnung der verallgemeinerten Inversen

In Abschn. 3.1.2 ist der Ausdruck

$$\begin{aligned}
 |\mathbf{A}\mathbf{f} - \mathbf{g}|^2 &= [\mathbf{A}\mathbf{f} - \mathbf{g}]^T [\mathbf{A}\mathbf{f} - \mathbf{g}] \\
 &= [\mathbf{f}^T \mathbf{A}^T - \mathbf{g}^T] [\mathbf{A}\mathbf{f} - \mathbf{g}] \\
 &= \mathbf{f}^T \mathbf{A}^T \mathbf{A} \mathbf{f} - \mathbf{f}^T \mathbf{A}^T \mathbf{g} - \mathbf{g}^T \mathbf{A} \mathbf{f} + \mathbf{g}^T \mathbf{g}
 \end{aligned}$$

bzgl. \mathbf{f} zu minimieren. Einfaches komponentenweises Differenzieren ¹

$$\begin{aligned}
 \partial |\mathbf{A}\mathbf{f} - \mathbf{g}|^2 / \partial f_k &= \mathbf{e}_k^T \mathbf{A}^T \mathbf{A} \mathbf{f} + \mathbf{f}^T \mathbf{A}^T \mathbf{A} \mathbf{e}_k - \mathbf{e}_k^T \mathbf{A}^T \mathbf{g} - \mathbf{g}^T \mathbf{A} \mathbf{e}_k \\
 &= \mathbf{e}_k^T [\mathbf{A}^T \mathbf{A} \mathbf{f} - \mathbf{A}^T \mathbf{g}] + [\mathbf{f}^T \mathbf{A}^T \mathbf{A} - \mathbf{g}^T \mathbf{A}] \mathbf{e}_k \\
 &= \mathbf{e}_k^T [\mathbf{A}^T \mathbf{A} \mathbf{f} - \mathbf{A}^T \mathbf{g}] + [\mathbf{A}^T \mathbf{A} \mathbf{f} - \mathbf{A}^T \mathbf{g}]^T \mathbf{e}_k \\
 &= 2\mathbf{e}_k^T [\mathbf{A}^T \mathbf{A} \mathbf{f} - \mathbf{A}^T \mathbf{g}]
 \end{aligned}$$

und anschließendes zu Null setzen liefert

$$\mathbf{A}^T \mathbf{A} \mathbf{f} - \mathbf{A}^T \mathbf{g} = 0.$$

B.1.2 Zur Minimierung des Tikhonov–Phillips–Funktional

In Abschn. 3.2.2 ist das Tikhonov–Phillips–Funktional (33)

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

¹Der Vektor \mathbf{e}_k ist der k -te Einheitsvektor.

bezüglich \mathbf{f}_γ zu minimieren. Komponentenweises Differenzieren ergibt

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

so daß hieraus folgt:

$$[\mathbf{A}^T \mathbf{A} + \gamma^2 \mathbf{B}^T \mathbf{B}] \mathbf{f}_\gamma - \mathbf{A}^T \mathbf{g}_\varepsilon - \gamma^2 \mathbf{B}^T \mathbf{r} = 0.$$

B.1.3 Zur Tikhonov–Phillips Regularisierung

Unter der einschränkenden Annahme, daß $\mathbf{A}^T \mathbf{A}$ und $\mathbf{B}^T \mathbf{B}$ ein gemeinsames System von Eigenvektoren \mathbf{v}_n besitzen,

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

kann auch für die Lösung der Tikhonov–Phillips–Regularisierung (34) eine Singulärwertzerlegung durchgeführt werden

$$\begin{aligned} \mathbf{f}_\gamma &= \sum_n \langle \mathbf{f}_\gamma | \mathbf{v}_n \rangle \mathbf{v}_n \\ &= \sum_n \langle \mathbf{A}^T \mathbf{g}_\varepsilon + \gamma^2 \mathbf{B}^T \mathbf{r} | (\mathbf{A}^T \mathbf{A} + \gamma^2 \mathbf{B}^T \mathbf{B})^{-1} \mathbf{v}_n \rangle \mathbf{v}_n \\ &= \sum_n (\sigma_n^2 + \gamma^2 \beta_n^2)^{-1} (\langle \mathbf{g}_\varepsilon | \mathbf{A} \mathbf{v}_n \rangle + \gamma^2 \langle \mathbf{r} | \mathbf{B} \mathbf{v}_n \rangle) \mathbf{v}_n. \end{aligned}$$

Für die spezielle Wahl $\mathbf{r} = 0$ folgt hieraus (siehe Gl. (23) in Abschn. 3.1.3)

$$\mathbf{f}_\gamma = \sum_n \sigma_n^{-2} F_\gamma(\sigma_n) \langle \mathbf{g}_\varepsilon | \mathbf{A} \mathbf{v}_n \rangle \mathbf{v}_n$$

mit einem Filter

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

B.1.4 Zum linear statistischen Verfahren

Sei

$$\begin{aligned} J(\mathbf{f}) &:= [\mathbf{g}_\varepsilon - \mathbf{A} \mathbf{f}]^T \mathbf{S}_\varepsilon^{-1} [\mathbf{g}_\varepsilon - \mathbf{A} \mathbf{f}] + [\mathbf{f} - \bar{\mathbf{f}}]^T \mathbf{S}_f^{-1} [\mathbf{f} - \bar{\mathbf{f}}] \\ &= \mathbf{g}_\varepsilon^T \mathbf{S}_\varepsilon^{-1} \mathbf{g}_\varepsilon - \mathbf{g}_\varepsilon^T \mathbf{S}_\varepsilon^{-1} \mathbf{A} \mathbf{f} - \mathbf{f}^T \mathbf{A}^T \mathbf{S}_\varepsilon^{-1} \mathbf{g}_\varepsilon + \mathbf{f}^T \mathbf{A}^T \mathbf{S}_\varepsilon^{-1} \mathbf{A} \mathbf{f} + \\ &\quad \mathbf{f}^T \mathbf{S}_f^{-1} \mathbf{f} - \mathbf{f}^T \mathbf{S}_f^{-1} \bar{\mathbf{f}} - \bar{\mathbf{f}}^T \mathbf{S}_f^{-1} \mathbf{f} + \bar{\mathbf{f}}^T \mathbf{S}_f^{-1} \bar{\mathbf{f}} \end{aligned}$$

das in Abschn. 3.2.3 Gl. (44) bzgl. \mathbf{f} zu minimierende Funktional. Komponentenweises Differenzieren

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

und anschließendes zu Null setzen liefert

$$[\mathbf{A}^T \mathcal{S}_\varepsilon^{-1} \mathbf{A} + \mathcal{S}_f^{-1}] \mathbf{f} = \mathbf{A}^T \mathcal{S}_\varepsilon^{-1} \mathbf{g}_\varepsilon + \mathcal{S}_f^{-1} \bar{\mathbf{f}}.$$

B.1.5 Zur praktischen Inversionsgleichung

$$\begin{aligned}
\mathbf{f}_{opt} &= [\mathbf{A}^T \mathcal{S}_\varepsilon^{-1} \mathbf{A} + \mathcal{S}_f^{-1}]^{-1} [\mathbf{A}^T \mathcal{S}_\varepsilon^{-1} \mathbf{g}_\varepsilon + \mathcal{S}_f^{-1} \bar{\mathbf{f}}] \\
&= \bar{\mathbf{f}} + [\mathbf{A}^T \mathcal{S}_\varepsilon^{-1} \mathbf{A} + \mathcal{S}_f^{-1}]^{-1} [\mathbf{A}^T \mathcal{S}_\varepsilon^{-1} \mathbf{g}_\varepsilon + \mathcal{S}_f^{-1} \bar{\mathbf{f}}] - \bar{\mathbf{f}} \\
&= \bar{\mathbf{f}} + [\mathbf{A}^T \mathcal{S}_\varepsilon^{-1} \mathbf{A} + \mathcal{S}_f^{-1}]^{-1} \mathbf{A}^T \mathcal{S}_\varepsilon^{-1} [\mathbf{g}_\varepsilon - \mathbf{A} \bar{\mathbf{f}}] \\
&= \bar{\mathbf{f}} + \mathcal{S}_f [\mathbf{A}^T \mathcal{S}_\varepsilon^{-1} \mathbf{A} \mathcal{S}_f + \mathbf{I}]^{-1} \mathbf{A}^T \mathcal{S}_\varepsilon^{-1} [\mathbf{g}_\varepsilon - \mathbf{A} \bar{\mathbf{f}}] \\
&= \bar{\mathbf{f}} + \mathcal{S}_f [\mathcal{S}_\varepsilon \mathbf{A}^{T-1} [\mathbf{A}^T \mathcal{S}_\varepsilon^{-1} \mathbf{A} \mathcal{S}_f + \mathbf{I}]^{-1} [\mathbf{g}_\varepsilon - \mathbf{A} \bar{\mathbf{f}}] \\
&= \bar{\mathbf{f}} + \mathcal{S}_f [\mathbf{A} \mathcal{S}_f + \mathcal{S}_\varepsilon \mathbf{A}^{T-1}]^{-1} [\mathbf{g}_\varepsilon - \mathbf{A} \bar{\mathbf{f}}] \\
&= \bar{\mathbf{f}} + \mathcal{S}_f \mathbf{A}^T [\mathbf{A} \mathcal{S}_f \mathbf{A}^T + \mathcal{S}_\varepsilon]^{-1} [\mathbf{g}_\varepsilon - \mathbf{A} \bar{\mathbf{f}}]
\end{aligned}$$

B.2 Spektralbereiche

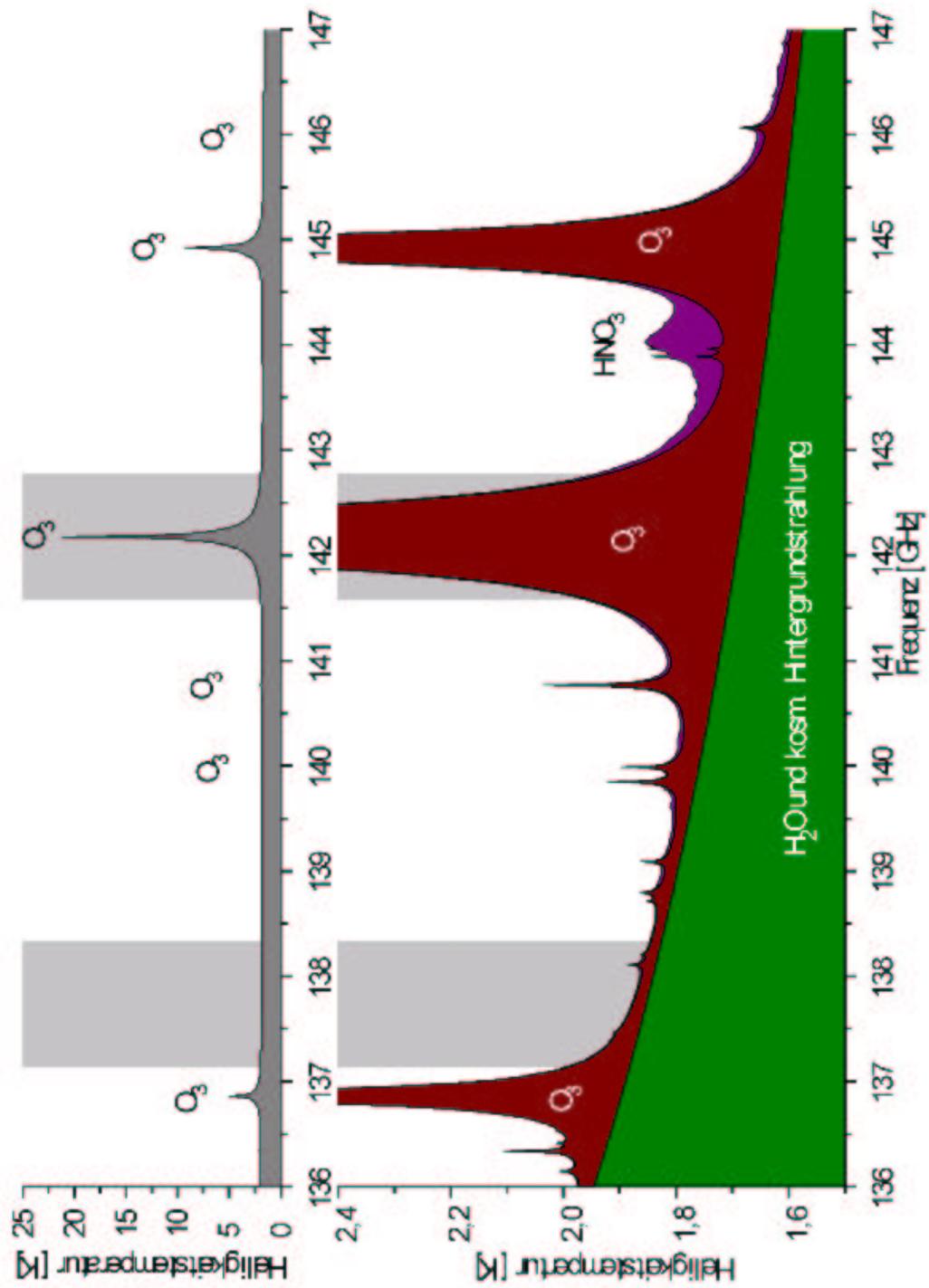


Abbildung 38: Ozon bei 142GHz, Erläuterungen siehe Rückseite

Abbildung 38: Mit dem Ozon-Radiometer bei 142 GHz maximal erreichbarer Spektralbereich, gerechnet für einen Elevationswinkel von 90° und eine Höhe von 10 km. Diese wurde gewählt, um den Einfluß des troposphärischen Wasserdampfs möglichst gering zu halten, ansonsten wäre das Spektrum für eine Übersichtsdarstellung zu stark verzerrt worden. Das Teilbild oben gewährt zunächst einen Überblick über die Intensität der Spektrallinien im besagten Spektralbereich und das Teilbild unten zeigt das Strahldichtespektrum, unterschieden nach den Beiträgen der verschiedenen Spurengase. Zusätzlich mit eingezeichnet sind die beiden Seitenbänder des Ozon-Radiometers für eine Frequenz des lokalen Oszillators (LO) von 139,9531 GHz.

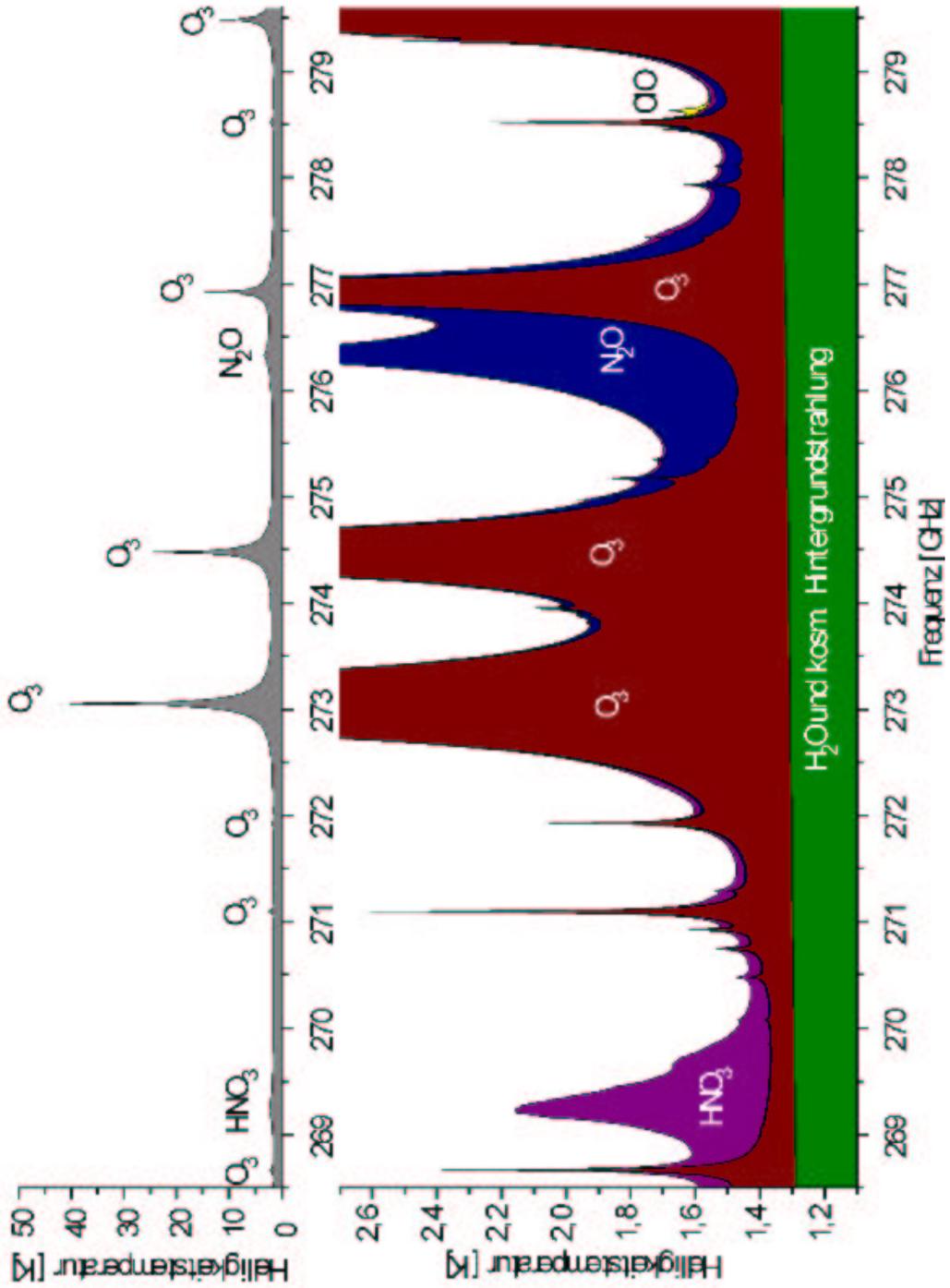


Abbildung 39: ClO bei 278GHz, Erläuterungen siehe Rückseite

Abbildung 39: Mit dem ClO-Radiometer bei 278 GHz überstreichbarer Spektralbereich, gerechnet für einen Elevationswinkel von 90° und eine Höhe von 10 km. Wie auch in der Abb. 38 wurde diese gewählt, um den Einfluß des troposphärischen Wasserdampfs auf das Spektrum möglichst gering zu halten. Am Boden hätte sich dieser ansonsten in Form eines sehr breitbandigen Untergrundes ausgewirkt. Überdies würde die Signatur des troposphärischen Spurengases N_2O am Boden sehr viel größer und breiter ausfallen, so daß insbesondere die ohnehin schwache Signatur des Spurengas ClO bei 278,631 GHz in dieser als Übersichtsdarstellung konzipierten Abbildung nicht mehr zu erkennen gewesen wäre. Das Teilbild oben gewährt zunächst einen Überblick über die Intensität der Spektrallinien. Das Teilbild unten schließlich zeigt das Strahl-dichtespektrum, unterschieden nach den Beiträgen der verschiedenen Spurengase.

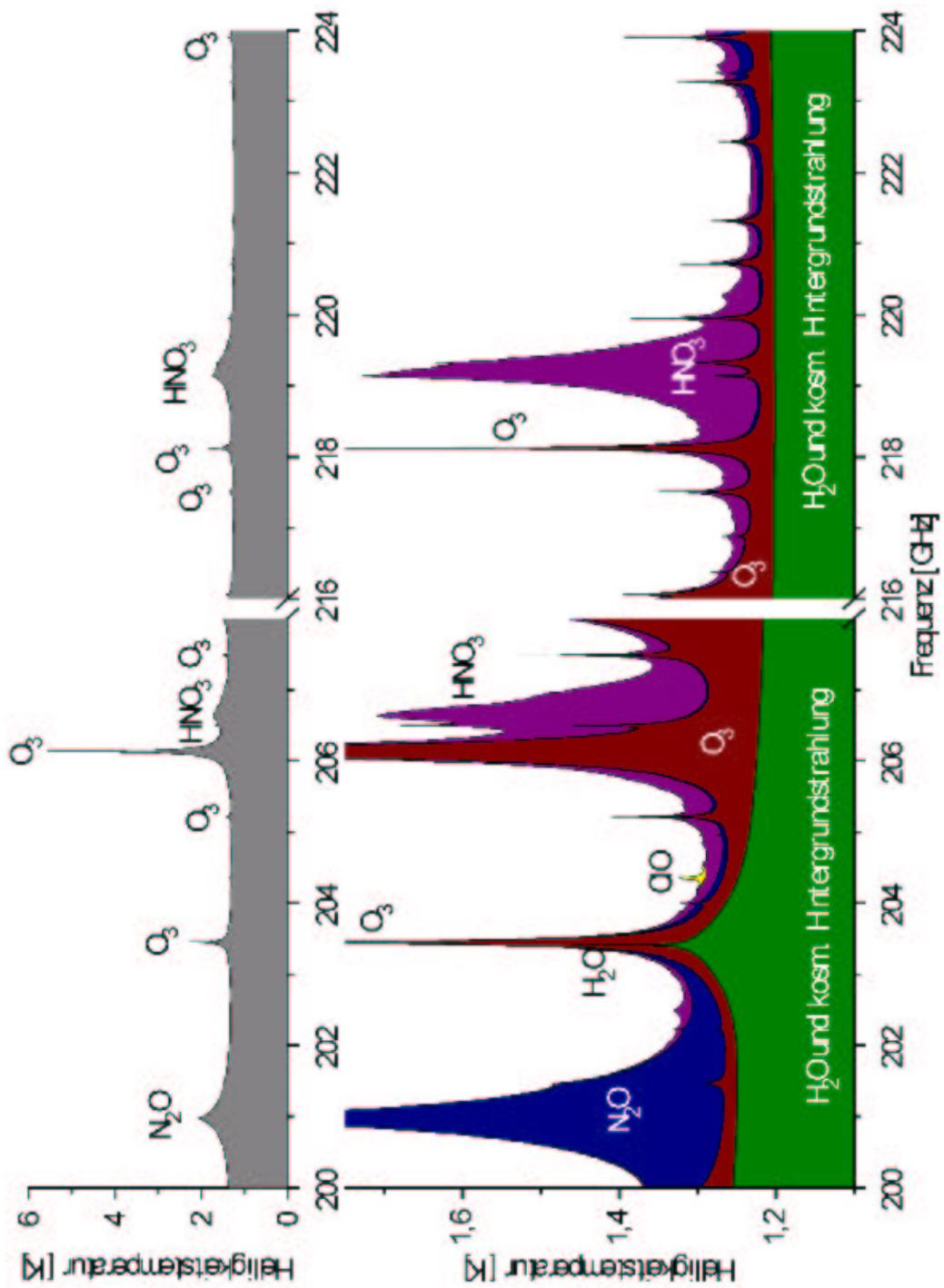


Abbildung 40: ClO bei 204GHz, Erläuterungen siehe Rückseite

Abbildung 40: Mit dem ClO-Radiometer bei 204 GHz überstreichbarer Spektralbereich, gerechnet für einen Elevationswinkel von 90° und eine Höhe von 10 km. Diese wurde diese gewählt, um den Einfluß des troposphärischen Wasserdampfs auf das Spektrum möglichst gering zu halten, welcher sich ansonsten in Form eines sehr breitbandigen Untergrundes ausgewirkt hätte. Überdies wäre die Signatur des troposphärischen Spurengases N_2O am Boden sehr viel größer und breiter ausgefallen, so daß die Abbildung ihren eigentlichen Zweck, einen Überblick über den interessierenden Spektralbereich zu gewähren nicht mehr hätte erfüllen können.

Das Teilbild oben gewährt zunächst einen Überblick über die Intensität der Spektrallinien. Die das Teilbild unten schließlich zeigt das Strahldichtespektrum, unterschieden nach den Beiträgen der verschiedenen Spurengase. Im Gegensatz zu den beiden Radiometern bei 142 und 278 GHz ist das ClO-Radiometer bei 204 GHz von vornherein für den routinemäßigen Einsatz bei verschiedenen Frequenzen vorgesehen, die dann rechnergesteuert angefahren werden können sollen.

Anhang C

Symbole und Abkürzungen

\mathbf{A}	Eine Matrix der Dimension $m \times l$, im wesentlichen der Kern der Integralgleichung, entspricht der Matrix \mathbf{A}^0
\mathbf{A}^T	Transponierte der Matrix \mathbf{A}
\mathbf{A}^0	durch Hintereinanderschreiben der beiden Jakobimatrizen \mathbf{A}_T^0 und $\mathbf{A}_{c_v}^0$ entstandene große Jakobimatrix
\mathbf{A}'	um die Berücksichtigung von stehenden Wellen erweiterte Matrix \mathbf{A}
\mathbf{A}^{-1}	Inverse der Matrix \mathbf{A} , sofern diese berechnet werden kann
\mathbf{A}^{-A}	Verallgemeinerte Inverse der Matrix \mathbf{A} , entspricht $[\mathbf{A}^T \mathbf{A}]^{-1} \mathbf{A}^T$
\mathbf{A}_γ^{-A}	Regularisierte (gefilterte) verallgemeinerte Inverse der Matrix \mathbf{A} , die Berechnungsvorschrift hängt ab vom gewählten Regularisierungsverfahren
$\mathbf{A}_{c_v}^0$	Jakobimatrix für das Volumenmischungsverhältnis, im wesentlichen gegeben durch die diskreten Sensitivitätsfunktionen $(S_{\nu c_v})_j$, von der Dimension $m \times n$
\mathbf{A}_T^0	Jakobimatrix für die Temperatur, im wesentlichen gegeben durch die diskreten Sensitivitätsfunktionen $(S_{\nu T})_j$, von der Dimension $m \times n$
A	Eine Abbildung von der Menge F der Parameter, nach der Menge G der Resultate
A	Amplitude einer dem Spektrum überlagerten stehenden Welle
a	ein Parameter
\mathbf{B}	Eine Matrix der Dimension $n \times n$, dient der Formulierung von Nebenbedingungen bei der Tikhonov–Phillips–Regularisierung
\mathbf{B}^T	Transponierte der Matrix \mathbf{B}
\mathbf{B}_{c_v}	Matrix mit Nebenbedingungen für das Volumenmischungsverhältnisprofil, von der Dimension $n \times n$
$\widetilde{\mathbf{B}}_{c_v}$	Matrix mit Nebenbedingungen für das Volumenmischungsverhältnisprofil, von der Dimension $l \times l$
\mathbf{B}_T	Matrix mit Nebenbedingungen für das Temperaturprofil, von der Dimension $n \times n$

$\widetilde{\mathbf{B}}_T$	Matrix mit Nebenbedingungen für das Temperaturprofil, von der Dimension $l \times l$
\mathbf{b}	Vektor der Dimension m , repräsentiert das an m spektralen Stützstellen gemessene Strahldichtespektrum
\mathbf{b}^0	Anfangsschätzung für den Vektor \mathbf{b}
$B_\nu(T)$	die Planckfunktion
b	ein Parameter
b	Linienbreite
\mathbf{C}	Eine positiv definite Diagonalmatrix der Dimension $m \times m$, dient der Gewichtung von Einzelmessungen beim Tikhonov–Phillips–Verfahren
c	Lichtgeschwindigkeit
c_ν	Volumenmischungsverhältnis eines nicht näher spezifizierten Gases
c_ν^0	Anfangsschätzung für das Volumenmischungsverhältnis
$c_{\nu j}$	Volumenmischungsverhältnis der j -ten Höhengschicht
$c_{\nu j}^0$	Anfangsschätzung für das Volumenmischungsverhältnis der j -ten Höhengschicht
$E()$	Erwartungswert einer Verteilung
\mathbf{e}_k	der k -te Einheitsvektor
e_ν	Quellterm, im Falle des thermodynamischen Gleichgewichts gleich dem Produkt aus Absorptionskoeffizient und Planckfunktion $k_\nu B_\nu(T)$
\mathcal{E}	Vektor der Dimension m , repräsentiert den Meßfehlervektor
\mathcal{E}^T	Transponierte des Meßfehlervektors
$\overline{\mathcal{E}}$	Erwartungswert des Meßfehlervektors, handelt es sich bei den Meßfehlern um rein statistische Fehler, entspricht dieser dem Nullvektor
\mathcal{E}_{opt}	Meßfehlervektor zur optimalen Lösung \mathbf{f}_{opt} der statistischen Regularisierungsverfahren
\mathcal{E}	Betrag des Meßfehlervektors, $\mathcal{E} = \mathcal{E} $
\mathbf{f}	Vektor der Dimension l , repräsentiert das (die) gesuchte(n) Höhenprofil(e)
\mathbf{f}^i	Vertreter eines ganzes Ensembles von Lösungsvektoren \mathbf{f}
$\overline{\mathbf{f}}$	Erwartungswert eines Ensembles von Lösungsvektoren, entspricht dem a priori Höhenprofil
\mathbf{f}^0	Anfangsschätzung für den Vektor \mathbf{f}
\mathbf{f}'	Um die Berücksichtigung von stehenden Wellen erweiterter Vektor \mathbf{f}
\mathbf{f}_{c_ν}	Vektor der Dimension n , repräsentiert ein gesuchtes Volumenmischungsverhältnisprofil

$\mathbf{f}_{c_v}^0$	Anfangsschätzung für den Vektor \mathbf{f}_{c_v}
\mathbf{f}_{opt}	regularisiertes oder optimales Höhenprofil bei den statistischen Regularisierungsverfahren, entspricht \mathbf{f}_γ
\mathbf{f}_T	Vektor der Dimension n , repräsentiert das gesuchte Temperaturprofil
\mathbf{f}_T^0	Anfangsschätzung für den Vektor \mathbf{f}_T
\mathbf{f}_γ	regularisiertes Höhenprofil, $\mathbf{f}_\gamma = \mathbf{A}_\gamma^{-A} \mathbf{g}$
F	Menge der Parameter
$F_\gamma(\sigma_n)$	regularisierendes Filter, die Filterwirkung hängt ab vom gewählten Regularisierungsverfahren
f_k	k -te Komponente des Vektors \mathbf{f}
\mathbf{g}	Vektor der Dimension m , repräsentiert die rechte Seite des entstandenen linearen Gleichungssystems
$\mathbf{g}_\mathcal{E}$	gestörte rechte Seite, $\mathbf{g}_\mathcal{E} = \mathbf{g} + \mathcal{E}$
\mathbf{g}_γ	aus der regularisierten Lösung über erneute Vorwärtsrechnung erhaltene rechte Seite, $\mathbf{g}_\gamma = \mathbf{A} \mathbf{f}_\gamma$
G	Menge der Resultate
$G_{\nu c_v}(h)$	Gewichtsfunktion für das Volumenmischungsverhältnis: gibt an, mit welchem Gewicht das Volumenmischungsverhältnis in der Höhe h beiträgt zu der in Beobachterhöhe meßbaren spektralen Strahldichte $L_\nu(h_B)$
$G_{\nu T}(h)$	Gewichtsfunktion für die Temperatur: gibt an, mit welchem Gewicht die Temperatur in der Höhe h beiträgt zu der in Beobachterhöhe meßbaren spektralen Strahldichte $L_\nu(h_B)$
$(G_{\nu c_v})_j$	Gewicht, mit dem das Volumenmischungsverhältnis der j -ten Höhenschicht zu der in Beobachterhöhe meßbaren spektralen Strahldichte beiträgt
$(G_{\nu T})_j$	Gewicht, mit dem die Temperatur der j -ten Höhenschicht zu der in Beobachterhöhe meßbaren spektralen Strahldichte beiträgt
h	Höhe
h_∞	obere Höhe der Atmosphäre, im Programm 100 km
h_B	Beobachterhöhe, Höhe des Standortes des Radiometers
\mathbf{I}	Einheitsmatrix von der Dimension $n \times n$
$J()$	Funktional
k	eine Konstante
k_ν	Absorptionskoeffizient
$k_{\nu j}$	Absorptionskoeffizient der j -ten Schicht
L_o	Maximale Amplitude einer stratosphärischen Signatur, Elevation 90°

L_u	Maximale Amplitude einer am Boden unter einem Elevationswinkel α im Zweiseitenbandbetrieb des Radiometers zu beobachtenden Signatur
L'_u	Maximale Amplitude einer am Boden unter einem Elevationswinkel α im Einseitenbandbetrieb des Radiometers zu beobachtenden Signatur
L_ν	spektrale Strahldichte, $[L_\nu] = \text{W/sr m}^2 \text{s}^{-1}$
$L_\nu(h_B)$	in Beobachterhöhe meßbare spektrale Strahldichte
$L_\nu(h_B)_j$	Beitrag der j -ten Höhenschicht zu der in Beobachterhöhe meßbaren spektralen Strahldichte
l	Anzahl der zu invertierenden Höhenprofile multipliziert mit n
l	Länge eines Resonators
m	Anzahl der Frequenzstützstellen
N	Anzahl der Lösungsvektoren im Ensemble
n	Anzahl der Höhenschichten
n_g	Anzahldichte der absorbierenden Gase
\mathbf{O}	Nullmatrix
\mathbf{o}	Nullvektor
$\mathcal{O}()$	Landau-Symbol, kennzeichnet das Restglied einer Reihenentwicklung
P	Periode einer dem Spektrum überlagerten stehenden Welle
p	Druck
$p_f(\mathbf{f})$	Wahrscheinlichkeitsdichte für das Auftreten eines bestimmten Höhenprofils \mathbf{f}
$p_\mathcal{E}(\mathcal{E})$	Wahrscheinlichkeitsdichte für das Auftreten eines bestimmten Meßfehlervektors \mathcal{E}
$p_{f,\mathcal{E}}(\mathbf{f}, \mathcal{E})$	Wahrscheinlichkeitsdichte für das Auftreten eines bestimmten Höhenprofils \mathbf{f} und eines bestimmten Meßfehlervektors \mathcal{E}
p_j	Druck der j -ten Höhenschicht
\mathbf{r}	Vektor der Dimension n , wird benötigt als Hilfsvektor bei der Formulierung der Nebenbedingungen für die Tikhonov-Phillips-Regularisierung
\mathbf{r}_{c_v}	Hilfsvektor für das Volumenmischungsverhältnis, von der Dimension n
$\widetilde{\mathbf{r}}_{c_v}$	Hilfsvektor für das Volumenmischungsverhältnis, von der Dimension l
\mathbf{r}_T	Hilfsvektor für das Temperaturprofil, von der Dimension n
$\widetilde{\mathbf{r}}_T$	Hilfsvektor für das Temperaturprofil, von der Dimension l
\mathbf{R}	Auflösungsmatrix oder Averaging Kernel Matrix, von der Dimension $l \times l$, $\mathbf{R} = \mathbf{A}_\gamma^{-A} \mathbf{A}$

r	Parameter bei der a posteriori-Parameterwahl
\mathcal{S}_f	a priori-Kovarianzmatrix, von der Dimension $l \times l$
$\mathcal{S}_{f c_v}$	a priori-Kovarianzmatrix für das Volumenmischungsverhältnis eines Gases, von der Dimension $n \times n$
$\mathcal{S}_{f T}$	a priori-Kovarianzmatrix für die Temperatur, von der Dimension $n \times n$
\mathcal{S}_f^{-1}	Inverse der a priori-Kovarianzmatrix
\mathcal{S}_ε	Fehlerkovarianzmatrix, von der Dimension $m \times m$
$\mathcal{S}_\varepsilon^{-1}$	Inverse der Fehlerkovarianzmatrix
\mathcal{S}_γ	Ergebniskovarianzmatrix der regularisierten Lösung, von der Dimension $l \times l$
$S(T)$	Linienstärke
$S_{\nu c_v}(h)$	Sensitivitätsfunktion für das Volumenmischungsverhältnis: gibt an, wie stark sich eine marginale Änderung des Volumenmischungsverhältnisses in der Höhe h auswirkt auf die in Beobachterhöhe meßbare spektrale Strahldichte $L_\nu(h_B)$
$S_{\nu T}(h)$	Sensitivitätsfunktion für die Temperatur: gibt an, wie stark sich eine marginale Änderung der Temperatur in der Höhe h auswirkt auf die in Beobachterhöhe meßbare spektrale Strahldichte $L_\nu(h_B)$
$(S_{\nu c_v})_j$	Sensitivität für das Volumenmischungsverhältnis der j -ten Höenschicht: gibt an, wie stark sich eine marginale Änderung des Volumenmischungsverhältnisses der j -ten Höenschicht auswirkt auf die in Beobachterhöhe meßbare spektrale Strahldichte $L_\nu(h_B)$
$(S_{\nu T})_j$	Sensitivität für die Temperatur der j -ten Höenschicht: gibt an, wie stark sich eine marginale Änderung der Temperatur der j -ten Höenschicht auswirkt auf die in Beobachterhöhe meßbare spektrale Strahldichte $L_\nu(h_B)$
T	Temperatur
T^0	Anfangsschätzung für die Temperatur
T_j	Temperatur der j -ten Höenschicht
T_j^0	Anfangsschätzung für die Temperatur der j -ten Höenschicht
T_{Trop}	Mittlere Temperatur der Troposphäre
U	breitbandiger Wasserdampfuntergrund eines unter einem Elevationswinkel α im Zweiseitenbandbetrieb arbeitenden Radiometers
U'	breitbandiger Wasserdampfuntergrund eines unter einem Elevationswinkel α im Einseitenbandbetrieb arbeitenden Radiometers
\mathbf{v}_k	Eigenvektor der Matrix $\mathbf{A}^T \mathbf{A}$ zum Eigenwert σ_k^2
\mathbf{v}_n	Eigenvektor der Matrix $\mathbf{A}^T \mathbf{A}$ zum Eigenwert σ_n^2

w	Gewicht des interessierenden Seitenbandes
α	Elevationswinkel
β_n^2	Eigenwert der Matrix $\mathbf{B}^T \mathbf{B}$
Δh_j	Dicke der j -ten Höenschicht, $h_{j+1} - h_j$
ϕ	Phase einer dem Spektrum überlagerten stehende Welle
γ	Regularisierungsparameter
γ	Druckverbreiterungsparameter
γ_{c_v}	Regularisierungsparameter für das Volumenmischungsverhältnis
γ_T	Regularisierungsparameter für die Temperatur
ν	Frequenz
σ_a	Absorptionsquerschnitt, entspricht dem Absorptionskoeffizienten dividiert durch die Anzahldichte der absorbierenden Teilchen
$\sigma_{\mathcal{E}}$	Standardabweichung eines jeden Elements des Meßfehlervektors
σ_k^2	Eigenwert der Matrix $\mathbf{A}^T \mathbf{A}$
σ_n^2	Eigenwert der Matrix $\mathbf{A}^T \mathbf{A}$
σ_s	Streuquerschnitt, entspricht dem Streukoeffizient dividiert durch die Anzahldichte der streuenden Teilchen
$\tau_{\nu}(h', h'')$	Optische Dicke der Atmosphäre zwischen der Höhe h' und der Höhe h''
[]	Klammerung
()	Kennzeichnung einer funktionalen Abhängigkeit
	Betrag eines Vektors bzgl. der L_2 -Norm
²	Betragsquadrat eines Vektors bzgl. der L_2 -Norm
< >	Skalarprodukt in Diracscher Notation
AFGL	Air Force Geophysics Laboratory
AOS	Akusto–Optisches Spektrometer
FASCOD2	Fast Atmospheric Signature Code 2
FZK	Forschungszentrum Karlsruhe
HITRAN	High Transmission
IMK	Institut für Meteorologie und Klimaforschung
REKLIP	Regionales Klima Projekt
SÜG	Strahlungsübertragungsgleichung
LO	Lokaler Oszillator
SVD	Singular Value Decomposition

Literaturverzeichnis

- [1] Amato, U., De Feis, I., Serio, C., Regularization for the Temperature Inverse Problem, Rapp. Tecnico n. 109/94, Dipartimento di Ingegneria e Fisica Ambientale, Univerisita della Basilicata, Potenza, Italia, 1994.
- [2] 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.
- [3] Baumeister, J., Stable Solutions of Inverse Problems, Advanced Lectures in Mathematics, Vieweg Verlag, Braunschweig Wiesbaden, 1987.
- [4] 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.
- [5] Boheim, Dr., Datenblatt des Mischers BM7 8636R, AEG, Feb. 1992.
- [6] 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.
- [7] Bronstein, I. N., Semendjajev, K. A., Taschenbuch der Mathematik, Verlag Harri Deutsch, 1985.
- [8] 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.
- [9] 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.
- [10] Chandrasekhar, S., Radiative Transfer, Dover Publications Inc. NewYork, 1960.
- [11] Von Clarmann, T., Untersuchungen zur Strahldichteberechnung mit Linie-für-Linie-Computerprogrammen, Diplomarbeit für Meteorologie, Ludwig-Maximilians-Universität München, Juli 1986.

- [12] Von Clarmann, T., Die Bestimmung von Mischungsverhältnissen stratosphärischer Spurengase aus Emissionsspektren im infraroten Spektralbereich, Dissertation, genehmigt von der Fakultät für Physik der Universität Karlsruhe, KfK-Bericht Nr. 4698, März 1990.
- [13] Clough, S. A., Kneizys, F. X., Shettle E. P., Anderson, G. P., Atmospheric Radiance and Transmittance: FASCOD2, Sixth Conference on Atmospheric Radiation, Williamsburg Va., May 1986.
- [14] Clough S. A., Kneizys, F. X., Anderson, G. P., Shettle, E. P., Chetwynd, J. H., Abreu, L. W., Hall, L. A., Worsham, R. D., FASCOD3: Spectral Simulation, IRS '88: Current Problems in Atmospheric Radiation, Proceedings of the International Radiation Symposium, Lille, France, Aug. 1988.
- [15] Connor, J. B., Barrett, J. W., Parrish, A., Solomon, P. M., De Zafra, R. L., Jaramillo, M., Ozone over McMurdo Station, Antarctica, Austral Spring 1986: Altitude Profiles for the Middle and Upper Stratosphere, Journal of Geophysical Research, Vol. 92, No. D11, pp. 13,221-13,230, Nov. 1987.
- [16] 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.
- [17] Davies, R. B., Documentation for NEWMAT07, an Experimental Matrix Package in C++, public domain software, 1993.
- [18] Deutscher Wetterdienst, Europäischer Wetterbericht, Amtsblatt des Deutschen Wetterdienstes, Zentralamt, Frankfurter Str. 135, 63067 Offenbach a. M., Postfach 100465, 63004 Offenbach, ISSN 0341-2970.
- [19] Dobner, H.-J., Was ist schlecht an schlecht gestellten Problemen?, KfK-Seminarreihe: Aktuelle Forschungsgebiete in der Mathematik, Seminarbeiträge 1993, KfK 5276, pp. 7-26, März 1994.
- [20] Farman, J. C., Gardiner, B. G., Shanklin, J. D., Large Losses of Total Ozone in Antarctica Reveal Seasonal ClO/NO Interaction, Nature, Vol. 315, pp. 207-210, May 1985.
- [21] Fiedler, F., Das Regio-Klima Projekt: Wie regeln die natürlichen Energieumsetzungen das Klima in einer Region? KfK-Nachrichten, Jahrgang 24, Nr. 3, pp. 125-131, Sep. 1993.

- [22] Fischer, H., Beiträge zur passiven Fernerkundung stratosphärischer Spurengase im infraroten Spektralbereich, Habilitationsschrift der Fakultät für Physik der Ludwigs-Maximilians-Universität München, Okt. 1985.
- [23] Fischer, H., Stratosphärisches Ozon: Dynamische und Chemische Prozesse, aus Atmosphärisches Ozon, Prozesse und Wirkungen, Vortragsveranstaltung der Arbeitsgemeinschaft der Großforschungseinrichtungen (AGF), Ahr-Straße 45, 53175 Bonn-Bad-Godesberg, pp. 33-36, Dez. 1993.
- [24] Fischer, H., Ozonveränderungen in der Stratosphäre: Dynamische und Chemische Prozesse KfK-Nachrichten, Jahrgang 26, Nr. 2, pp. 61-66, Juli 1994.
- [25] 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.
- [26] Franklin, J. N., Well-Posed Stochastic Extensions of Ill-Posed Linear Problems, Journal of Mathematical Analysis and Applications, Vol. 31, pp. 682-716, 1970.
- [27] Fröhner, F.,H., Ronen Y. (ed.), Elias E. (ed.), Assignment of Uncertainties to Scientific Data, Proc. Internat. Conf. on Reactor Physics and Reactor Computations, Tel Aviv, Ben-Gurion of the Negev Press, Beer-Sheva, p. 287, 1994.
- [28] Gasiewski, A. J., Janssen, M. A. (ed.), Microwave Radiative Transfer in Hydrometeors, Atmospheric Remote Sensing by Microwave Radiometry, Wiley Series in Remote Sensing, John Wiley & Sons Inc., New York, chap. 3, pp. 91-144, 1993.
- [29] Goody, R. M., Yung, Y. L., Atmospheric Radiation, Theoretical Basis, Second Edition, Oxford University Press, 1989.
- [30] Hadamard, J., Lectures on the Cauchy Problem in Linear Partial Differential Equations, Yale University Press, NewHaven, 1923.
- [31] Hansen, P. C., Regularization, GSVD, and Truncated GSVD, BIT, Vol. 29, pp. 491-504, 1989.
- [32] Hochschild, G., Krupa, R., Kuntz, M., Millimeterwellen-Radiometrie von Spurengasen in der Stratosphäre, KfK-Nachrichten, Jahrgang 25, Nr. 4, pp. 182-192, Dez. 1993.
- [33] Houghton, J. T., Taylor, F. W., Rodgers, C. D., Remote Sounding of Atmospheres, Cambridge Planetary Sciences Series 5, Cambridge University Press, 1986.
- [34] Ishimaru, A., Wave Propagation and Scattering in Random Media, Section 22, Vol. II, Academic Press NewYork SanFrancisco London, 1978.

- [35] Janssen, M. A., Microwave Radiative Transfer in Hydrometeors, Atmospheric Remote Sensing by Microwave Radiometry, Wiley Series in Remote Sensing, John Wiley & Sons Inc., New York, chap. 1, pp. 1-35, 1993.
- [36] Klein, U., Aufbau und Betrieb eines breitbandigen, bodengestützten Millimeterwellen-Radiometers zur Messung atmosphärischer Spurengase, Dissertation, genehmigt von der Fakultät für Physik der Universität Bremen, Juli 1993.
- [37] Kolle, O., private Mitteilungen, 1995.
- [38] Kraus, J. D., Radioastronomy, 2nd ed., Cygnus-Quasar Books, Powell, Ohio, 1986.
- [39] Kuchling, H., Physik, Formeln und Gesetze, Buch- und Zeit-Verlagsgesellschaft mbH Köln, 1969.
- [40] Kuntz, M., Inversion schlecht gestellter Probleme am Beispiel der bodengebundenen, passiven Millimeterwellenradiometrie, KfK-Bericht Nr. 5421, Institut für Meteorologie und Klimaforschung, Kernforschungszentrum Karlsruhe GmbH, Okt. 1994.
- [41] Landweber, L., An Iteration Formula for Fredholm Integral Equations of the First Kind, American Journal of Mathematics, Vol. 92, No. 191, pp. 615-624, 1951.
- [42] Liebe, H. J., Deepak, A. (ed.), Wilkerson, T. D. (ed.), Ruhnke, L. H. (ed.), Atmospheric Water Vapor: A Nemesis for Millimeter Wave Propagation, Atmospheric Water Vapor, Academic Press, New York London Toronto Sydney San Francisco, pp. 143-201, 1980.
- [43] Liebe, H. J., Temkin, R. J. (ed.), The Atmospheric Water Vapor Continuum Below 300 GHz, 8th International Conference on Infrared and Millimeter Waves, Miami Beach, Florida, IEEE Catalog No. 83CH1917-4, pp. TH6.3, Dec. 1983.
- [44] Liebe, H. J., The Atmospheric Water Vapor Continuum Below 300 GHz, International Journal of Infrared and Millimeter Waves, Vol. 5, No. 2, 1984.
- [45] Liebe, H. J., An Updated Model for Millimeter Wave Propagation in Moist Air, Radio Science, Vol. 20, No. 5, pp. 1069-1089, Oct. 1985.
- [46] Liebe, H. J., MPM - An Atmospheric Millimeter-Wave Propagation Model, International Journal of Infrared and Millimeter Waves, Vol. 10, No. 6, pp. 631-650, Feb. 1989.

- [47] Liebe, H. J., Hufford, G. A., Modeling Millimeter-Wave Propagation Effects in The Atmosphere, Atmospheric Propagation in the UV, Visible, IR, and MM-Wave Region and Related Systems, ARGARD-CP-454, pp. 18,1-18,13, Mar. 1990.
- [48] Lobsiger, E., Ground-Based Microwave Radiometry to Determine Stratospheric and Mesosheric Ozone Profiles, Journal of Atmospheric and Terrestrial Physics, Vol. 49, No. 5, pp. 493-501, 1987.
- [49] Loredo, T. J., Epstein, R. I., Analyzing Gamma-Ray Burst Spectral Data, The Astrophysical Journal, Vol. 336, pp. 896-919, Jan. 1989.
- [50] Louis, A. K., Inverse und schlecht gestellte Probleme, Teubner Studienbücher, 1989.
- [51] Manabe, T., Furuham, Y., Ihara, T., Saito, S., Tanaka, H., Ono, A., Measurements of Attenuation and Refractive Dispersion due to Atmospheric Water Vapor at 80 and 240 GHz, International Journal of Infrared and Millimeter Waves, Vol. 6, No. 4, pp. 313-322, 1985.
- [52] Menke, W., Geophysical Data Analysis: Discrete Inverse Theory, Academic Press Inc., SanDiego, 1989.
- [53] Miller, K., Least Squares Methods for Ill-Posed Problems with a Prescribed Bound, SIAM, Journal on Numerical Analysis, Vol. 16, pp. 52-74, 1970.
- [54] Morrison, D. F., Multivariate Statistical Methods, McGraw-Hill, Second Edition, NewYork, 1967.
- [55] Natterer, F., The Mathematics of Computerized Tomography, B. G. Teubner, Stuttgart, John Wiley & Sons, Chichester NewYork Brisbane Toronto Singapore, 1986.
- [56] Nedoluha, G. E., Bevilacqua R. M., Gomez R. M., Thacker, D. L., Waltman, W. B., Pauls, T. A., Ground-based Measurements of Water Vapor in the Middle Atmosphere, Journal of Geophysical Research, Vol. 100, No. D2, pp. 2927-2939, Feb. 1995.
- [57] Niple, E., Nonlinear least squares analysis of atmospheric absorption spectra, Applied Optics, Vol. 19, No. 20, pp. 3481-3490, Oct. 1980.
- [58] Oelhaf, H., Fischer, H., Lenoble, J. (ed.), Geleyns, J. F. (ed.), Relevance of Upper Atmosphere non-LTE Effects to Limb Emission of Stratospheric Constituents, IRS '88, Current Problems in Atmospheric Radiation, Proceedings of the International Radiation Symposium, Lille, France, A. Deepak Publishing, pp. 460-463, Aug. 1989.

- [59] Parrish, A., De Zafra, R. L., Solomon, P. M., Barrett, J. W., A Ground-Based Technique for Millimeter Wave Spectroscopic Observations of Stratospheric Trace Constituents, *Radio Science*, Vol. 23, pp. 106-118, 1988.
- [60] Penrose, R., A Generalized Inverse for Matrices, *Proc. Cambridge Philos. Soc.*, Vol. 51, pp. 406-413, 1955.
- [61] 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.
- [62] Pfister, A., Untersuchung der horizontalen und vertikalen Auflösung eines operationellen Inversionsverfahrens zur Bestimmung von mesoskaligen Temperaturfeldern aus Satellitendaten, Diplomarbeit im Fach Meteorologie, Institut für Meteorologie und Klimaforschung, Okt. 1988.
- [63] Poynter, R. L., Pickett, H. M., *Submillimeter, Millimeter, and Microwave Spectral Line Catalog*, Allied Optics, Vol. 24, No. 14, pp. 2235-2240, 1984.
- [64] Raffalski, U., Mikrowellen-Radiometrie in der Atmosphärenforschung unter Verwendung der Referenzstrahl-Methode, Diplomarbeit im Studiengang Physik, Institut für Fernerkundung der Universität Bremen, Sep. 1993.
- [65] 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.
- [66] Redemann, E., Ein FORTRAN-Rechenprogramm zur Berechnung der atmosphärischen Transmission und Strahldichte, Report des Meteorol. Inst. der LM-Universität in München, Mai 1984.
- [67] 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.
- [68] 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.
- [69] 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.

- [70] 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.
- [71] 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.
- [72] Rosenkranz, P. W., Janssen, M. A. (ed.), Absorption of Microwaves by Atmospheric Gases, in *Atmospheric Remote Sensing by Microwave Radiometry*, Wiley Series in Remote Sensing, John Wiley & Sons Inc., New York, chap. 2, pp.37-90, 1993.
- [73] Rothman, L. S., Gamache, R. R., Goldman, A., Toth, R. A., Pickett, H. M., Poynter, R. L., Flaud, J.-M., Camy-Peyret, C., Barbe, A., Husson N. et al., The HITRAN Database: 1986 Edition, *Applied Optics*, Vol. 26, No. 19, pp.4058-4097, Oct. 1987.
- [74] Rothman, L. S. (ed.), The Spectroscopic Database for High Resolution Atmospheric Modeling, SPIE, Vol. 928, *Modeling of the Atmosphere*, pp. 79-93, Apr. 1988.
- [75] Rothman, L. S., Gamache R. R., Tipping R. H., Rinsland, C. P., Smith, M. A. H., Benner, D. C., Devi, V. M., Flaud, J.-M., Camy-Peyret, C., et al., The HITRAN Molecular Database: Editions of 1991 and 1992, *Journal of Quantum Spectroscopy and Radiative Transfer*, Vol. 48, pp. 469-507, June 1992.
- [76] Smirnow, W. I., *Lehrgang der Höheren Mathematik*, VEB Deutscher Verlag der Wissenschaften, Berlin, 1988.
- [77] Smith, L. L., Chedin, A. (ed.), Atmospheric Transmittance Modeling in the Submillimeter Region, *Microwave Remote Sensing of the Earth System*, A. Deepak Publishing, Hampton, Virginia, pp. 1-10, 1989.
- [78] Solomon, P. M., De Zafra, R. L., Parrish, A., Barrett, J. W., Observations of Diurnal Variation of Stratospheric Chlorine Monoxide: A Critical Test of Chlorine Chemistry in the Ozone Layer, *Science*, Vol. 224, pp. 1210-1214, Mar. 1984.
- [79] Solomon, S., The Mystery of the Antarctic Ozone Hole, *Reviews of Geophysics*, Vol. 26, No. 1, pp. 131-148, Feb. 1988.
- [80] Solomon, S., Progress towards a Quantitative Understanding of Antarctic Ozone Depletion, *Nature*, Vol. 347, pp. 347-354, Sep. 1990.

- [81] Stolarski, R. S., Bojkov, R., Bishop, L., Zerefos, C., Staehelin, J., Zawodny, J., Measured Trends in Stratospheric Ozone, *Science*, Vol. 256, pp. 342-349, 1992.
- [82] 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.
- [83] 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.
- [84] Tikhonov, A., N., Arsenin, V., Y., Solutions of Ill-Posed Problems, John Wiley & Sons, Scripta Series in Mathematics, New York, 1977.
- [85] Twomey, S., Introduction to the Mathematics of Inversion in Remote Sensing and Indirect Measurements, *Developments in Geomathematics 3*, Elsevier Scientific Publishing Company, Amsterdam, Oxford, New York, 1977.
- [86] Ulaby, F. T., Moore, R. K., Funk, A. K., Volume Scattering and Emission Theory, *Microwave Remote Sensing, Active and Passive*, Artech House, Vol. 1, chap. 6, pp. 345-431, 1981.
- [87] Ulaby, F. T., Moore, R. K., Funk, A. K., Volume Scattering and Emission Theory, *Microwave Remote Sensing, Active and Passive*, Artech House, Vol. 3, chap. 13, pp. 1065-1205, 1990.
- [88] Vowinkel, B., *Passive Mikrowellenradiometrie*, Vieweg Verlag, Braunschweig, ISBN 3-528-08959-8, 1988.
- [89] Waters, J. W., Absorption and Emission by Atmospheric Gases, *Methods of Experimental Physics: Astrophysics*, 12B, chap. 2.3, Academic Press, San Diego, California, pp. 142-176, 1976.
- [90] Wang, W.-C., Molnar, G., A model Study of the Greenhouse Effects Due to Increasing Atmospheric CH₄, N₂O, CF₂Cl₂, and CFCl₃, *Journal of Geophysical Research*, Vol. 90, No. D7, pp. 12,971-12,980, Dec. 1985.
- [91] Wehr, T., Bestimmung atmosphärischer Spurengasgehalte aus Spektren eines flugzeuggestützten Submillimeter-Radiometers, Diplomarbeit im Studiengang Physik, Institut für Fernerkundung der Universität Bremen, Apr. 1993.
- [92] 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.

- [93] Westwater, E. R., Janssen, M. A. (ed.), Ground-Based Microwave Remote Sensing of Meteorological Variables, Atmospheric Remote Sensing by Microwave Radiometry, Wiley Series in Remote Sensing, John Wiley & Sons Inc., NewYork, chap. 4, pp. 145-213, 1993.
- [94] Wetzel, G., Eignung der Infrarotspektroskopie zur Fernerkundung troposphärischer Spurengase, Dissertation, genehmigt von der Fakultät für Physik der Universität Karlsruhe, März 1993.
- [95] De Zafra, R. L., Parrish, A., Solomon, P. M., Barrett, W., A Measurement of Stratospheric HO₂ by Ground-Based Millimeter-Wave Spectroscopy, Journal of Geophysical Research, Vol. 89, No. D1, pp. 1321-1326, Feb. 1984.
- [96] De Zafra, R. L., Jaramillo, M., Barrett, J., Emmons, L. K., Solomon, P. M., Parrish, A., New Observations of a Large Concentration of ClO in the Springtime Lower Stratosphere Over Antarctica and Its Implications for Ozone-Depleting Chem., Journal of Geophysical Research, Vol. 94, No. D9, pp. 11,423-11,428, Aug. 1989.
- [97] Zommerfelds, W. C., Künzi, K. F., Summers, M. E., Bevilacqua, R. M., Strobel, D. F., Allen, A., Sawchuck, W. J., Diurnal Variations of Mesospheric Ozone Obtained by Ground-Based Microwave Radiometry, Journal of Geophysical Research, Vol. 94, No. D10, pp. 12,819-12,832, Sep. 1989.