

Universität Duisburg-Essen
Campus Duisburg
Fakultät 4 / Institut für Mathematik

Diplomarbeit

Gradientenflussmethoden bei inversen Problemen in der Fernerkundung

Benjamin Berkels

18.4.2005

Inhaltsverzeichnis

1	Einleitung	5
1.1	Inverse und schlecht gestellte Probleme	5
1.2	Ein inverses Problem der Fernerkundung	5
2	Das kontinuierliche Problem	9
2.1	Das Problem als Minimierungsproblem	9
2.2	Gradientenfluss zur Energieminimierung	12
2.2.1	Motivation: Der endlich-dimensionale Fall	12
2.2.2	Gradientenflüsse auf Funktionenräumen	12
2.2.3	Der spezielle Gradientenfluss	13
2.3	Wahl und Eigenschaften der Metrik	13
2.3.1	Wahl der Metrik	13
2.3.2	Invertierbarkeit der Metrik	17
2.3.3	Stetigkeit der Inversen der Metrik	19
2.3.4	Stetigkeit der Metrik	23
2.4	Variation der Energie	26
2.5	Adjoint variable method	27
2.6	Stetigkeit der Variation der Energie	28
2.7	Kurzzeit-Existenz des Gradientenflusses	35
2.8	Regularisierung	37
2.8.1	Erster Teil der Regularisierungsenergie	37
2.8.2	Zweiter Teil der Regularisierungsenergie	38
2.8.3	Variation der Regularisierungsenergie	39
2.9	Stetigkeit der Variation der Regularisierungsenergie	42
3	Diskretisierung	45
3.1	Ortsdiskretisierung	45
3.2	Testfunktionen	46
3.2.1	Variation der Temperatur	47
3.2.2	Variation der Höhe	47
3.3	Zeitdiskretisierung	48
3.3.1	Diskreter Kurvenraum	48
3.3.2	Zeitdiskretisierung mittels Differenzenquotient	50
3.4	Die diskrete Metrik	51
3.5	Unterschiede der Testfunktionen im diskreten und kontinuierlichen Fall	53
3.6	Numerische Lösung der Differentialgleichungen	54
3.7	Auswertung der benutzten Funktionen	56
3.7.1	Auswertung von p	56

3.7.2	Auswertung von q	56
3.7.3	Auswertung von σ_a	56
3.8	Diskretisierung der Regularisierungsenergie	57
3.9	Variation der diskreten Regularisierungsenergie	58
3.10	Der Algorithmus	60
3.10.1	Wahl des Startprofils	60
3.10.2	Berechnung der Variation	61
3.10.3	Zeitschritt	62
3.10.4	Wahl von A	62
3.10.5	Zeitschrittweitensteuerung	62
3.10.6	Abbruchkriterium	64
3.11	Vorgehen bei unbekanntem Druck	64
3.12	Validierung	65
3.12.1	Validierung der Störungsrechnung	65
3.12.2	Validierung der Variationsberechnung	66
4	Ergebnisse	67
4.1	Auswertung des Blindtests	67
4.1.1	Auswertung für Inversionen	70
4.1.2	Ausgewählte Ergebnisse	74
4.2	Ausblick	75
5	Anhang	77
5.1	Wichtige Definitionen und Sätze	77
5.2	Programm-Dokumentation	81
5.2.1	Dateiformat der Eingabe Datensätze	81
5.2.2	Format der Parameterdatei <code>meteo.par</code>	82
5.2.3	Format der Ausgabedatei <code>output.dat</code>	82
	Literaturverzeichnis	85
	Index	87

1 Einleitung

In dieser Arbeit beschäftigen wir uns mit einem inversen und schlecht gestellten Problem aus der Fernerkundung. In Kapitel 2 führen wir eine Energie ein, deren Nullminima Lösungen dieses Problems sind, und minimieren diese Energie mit Hilfe eines Gradientenflusses. Anschließend wird in Kapitel 3 der Gradientenfluss diskretisiert und damit ein Algorithmus erarbeitet, der aus gegebenen Eingangsdaten lokale Minima der Energie berechnet. Die Qualität der konkreten Implementierung des Algorithmus wird abschließend in Kapitel 4 mit einer Fehleranalyse der Ergebnisse eines Blindtests untersucht.

Informationen zum Hintergrund und weitere Anwendungen der Grundidee dieses Ansatzes, also das Benutzen eines sog. regularisierten Gradientenflusses zur Lösung eines inversen Problems, finden sich in [24, 22, 7]. Zunächst klären wir jedoch, was sich hinter den Begriffen inverses und schlecht gestelltes Problem verbirgt, und skizzieren unser Problem aus der Fernerkundung.

1.1 Inverse und schlecht gestellte Probleme

Die folgenden Definitionen stammen aus [20], dort sind auch weitere, allgemeine Ausführungen zu inversen und schlecht gestellten Problemen zu finden.

Wir sprechen von einem **inversen Problem**, wenn wir Eigenschaften eines Objektes ermitteln wollen, dies aber nicht durch direkte Messung möglich ist, sondern nur indirekt, indem wir aus Beobachtungen anderer Eigenschaften Rückschlüsse auf die gesuchten Eigenschaften ziehen. Ein bekanntes Beispiel hierfür ist die Computer-Tomographie: Man möchte die Dichte des Körpergewebes eines Menschen bestimmen, kann dies jedoch nicht durch direkte Messung im Körper, sondern schickt Röntgenstrahlen durch den Körper, vergleicht die Intensität der Strahlen vor und nach dem Durchqueren des Körpers und ermittelt damit die gesuchte Dichte.

1.1.i Definition. Seien X, Y topologische Räume und $A : X \rightarrow Y$. Das Problem (A, X, Y) heißt **gut gestellt**, wenn

- (i) $Af = g$ für jedes $g \in Y$ eine Lösung f hat,
- (ii) diese Lösung eindeutig ist,
- (iii) die Lösung stetig von g abhängt.

Ist eine der Bedingungen nicht erfüllt, dann heißt das Problem **schlecht gestellt**.

1.2 Ein inverses Problem der Fernerkundung

Für die numerische Wettervorhersage sind kontinuierlich gemessene Profile bezüglich der Höhe von

- Temperatur (T) in K
- Feuchte (q) in kgm^{-3}
- Druck (p) in Pa

äußerst nützlich (vgl. Abbildung 1.1). Eine direkte Messung dieser Profile ist durch Radiosondenaufstiege zwar im Prinzip möglich und wird derzeit auch betrieben, aber die sehr hohen laufenden Kosten für Personal und Material bei Radiosondenaufstiegen (ca. 300.000 Euro pro Jahr und Radiosondenstation) sorgen dafür, dass nach günstigeren Alternativen gesucht wird.

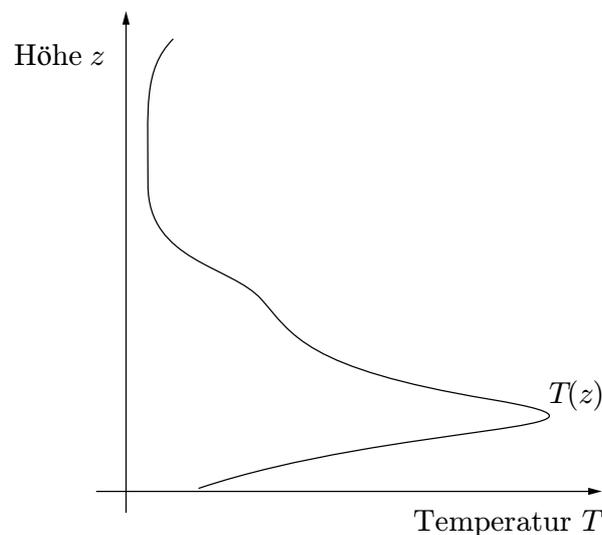


Abbildung 1.1: ein höhenabhängiges Temperaturprofil

Seit einiger Zeit gibt es so genannte Mikrowellen-Radiometer, mit denen man gleichzeitig Strahlungstemperaturen an mehreren Frequenzen messen kann. Genau geht es hier um die Mikrowellen-Frequenzen an der linken Flanke des Sauerstoff Absorptionskomplexes (eine Erläuterung dazu ist in [16] zu finden). Damit ergibt sich eine Möglichkeit, indirekt Rückschlüsse über die Profile zu ziehen, denn es gibt Modelle, mit denen aus den Atmosphärenprofilen die Strahlungstemperaturen errechnet werden können, und mit einer Invertierung dieser Modelle könnte man aus gemessenen Strahlungstemperaturen die Profile bestimmen. Einen Überblick über das Potenzial dieser Art Ansätze geben [3, 26]. Wir bezeichnen die Berechnung der Strahlungstemperaturen aus den Atmosphärenprofilen als **Vorwärtsrechnung** und die Umkehrung als **Rückwärtsrechnung**. In dieser Arbeit beschränken wir uns darauf, das Temperaturprofil (und daraus das Druckprofil) zu bestimmen, das Feuchteprofil nehmen wir immer als gegeben an. Ein mit einer Rückwärtsrechnung gewonnenes Temperaturprofil nennen wir **Retrieval**.

Existierende Ansätze dieser Art sind z.B.

- Mehrfachregression und Neuronale Netze in [13],
- Ansatz nach Bayes in [23] und [10].

Von diesen Ansätzen ist der erste rein statistisch, der zweite physikalisch mit statistischen Elementen, während wir hier einen neuartigen und schnellen numerischen Ansatz zur Bestimmung des Temperaturprofils vorstellen, der hauptsächlich auf dem physikalischen Vorwärtsmodell beruht und nur wenige (und in der Implementation abschaltbare) statistische Informationen benutzt. [4] ist eine erste Veröffentlichung zu unserem Ansatz.

2 Das kontinuierliche Problem

2.1 Das Problem als Minimierungsproblem

Wie schon erwähnt, benötigen wir ein Modell, um aus den Atmosphärenprofilen die Strahlungstemperaturen zu errechnen. Wir beschränken uns hier auf die Benutzung eines relativ einfachen Modells, es spricht jedoch nichts dagegen, ein anderes, komplizierteres Modell zu wählen, solange die Variation der Strahlungstemperaturen stetig von der Temperatur abhängt (vgl. 2.6). Mit Ausnahme der Berechnung der Strahlungstemperaturen und deren Variation bleiben die folgenden Ausführungen auch für andere Modelle gleich.

Für unser Modell müssen wir folgende Voraussetzungen machen:

- Messkonfiguration
 - Mehrkanal-Mikrowellen-Radiometer in Bodenhöhe z_0
 - Atmosphäre ist horizontal geschichtet, d.h. Temperatur, Druck, Feuchte sind in Höhe z konstant
 - Keine Wolken
- Voraussetzungen
 - Strahlungstemperaturen hängen nur von T, q, p , Anstiegswinkel Θ , Frequenz ν ab
 - Atmosphärenparameter $Q(z) := (q(z), p(z))$ bekannt für alle Höhen z
 - nm Strahlungstemperaturmessungen ($T_{bm11}, \dots, T_{bmnm}$) bei Frequenzen ν_1, \dots, ν_n Winkeln $\theta_1, \dots, \theta_m$ sind gegeben
- In unserem speziellen Fall sind die Frequenzen 50.8, 51.8, 52.8, 53.8, 54.8, 55.8, 56.8, 57.8, 58.8 GHz, entsprechend einem Band des 22-Kanal-Mikrowellen-Radiometers MICCY (MICrowave Radiometer for Cloud CarthographY) der Universität Bonn, die Winkel sind frei wählbar. Eine genauere Beschreibung des Radiometers ist in [9] zu finden.

Ist $T \in C^2[z_0, \infty[$ ein Temperaturprofil, so gilt nach [16] für die zugehörigen Strahlungstemperaturen $T_b \in C^2[z_0, \infty[$ die in der Meteorologie gebräuchliche Integralformulierung

$$T_b(\nu, \theta, z) = T_{b0} \exp \left(- \sec(\theta) \int_z^\infty \sigma_a(\nu, T(t), Q(t)) dt \right) \\ + \sec(\theta) \int_z^\infty \sigma_a(\nu, T(t), Q(t)) T(t) \\ \cdot \exp \left(- \sec(\theta) \int_z^t \sigma_a(\nu, T(s), Q(s)) ds \right) dt.$$

Hierbei ist

- T_{b0} kosmische Hintergrundstrahlung von 2.73 K
- σ_a Absorptions- und Emissionskoeffizient

Da wir numerisch Integrale nicht bis ∞ auswerten werden, beschränken wir auch den kontinuierlichen Fall darauf, dass die Strahlungstemperaturen nur von Parametern bis zu einer Höhe z_{\max} abhängen. Sei $\mathcal{T} := C^2([z_0, z_{\max}])$ der Raum der Temperaturprofile mit der Norm

$$\|T\|_{\mathcal{T}} := \|T\|_{\infty, [z_0, z_{\max}]} + \|T'\|_{\infty, [z_0, z_{\max}]} + \|T''\|_{\infty, [z_0, z_{\max}]},$$

wobei für $f \in C^0([a, b])$

$$\|f\|_{\infty, [a, b]} := \max\{|f(x)| \mid x \in [a, b]\}.$$

Ist $T \in \mathcal{T}$, so erfüllen die zugehörigen Strahlungstemperaturen $T_b \in \mathcal{T}$ nach dem Vorwärtsmodell folgende Gleichung:

$$\begin{aligned} T_b(\nu, \theta, z) = & T_{b0} \exp\left(-\sec(\theta) \int_z^{z_{\max}} \sigma_a(\nu, T(t), Q(t)) dt\right) \\ & + \sec(\theta) \int_z^{z_{\max}} \sigma_a(\nu, T(t), Q(t)) T(t) \\ & \cdot \exp\left(-\sec(\theta) \int_z^t \sigma_a(\nu, T(s), Q(s)) ds\right) dt. \end{aligned} \quad (2.1)$$

Aus Satz 5.1.i folgt, dass (2.1) die Integralformulierung der folgenden Differentialgleichung ist:

$$\begin{aligned} \frac{d}{dz} T_b(\nu, \theta, z) &= G(T_b(z, \nu, \theta), \nu, \theta, T(z), Q(z)), \\ T_b(\nu, \theta, z_{\max}) &= T_{b0}, \end{aligned} \quad (2.2)$$

wobei die rechte Seite G gegeben ist durch

$$\begin{aligned} G(T_b, \nu, \theta, T(z), Q(z)) &= \sec(\theta) \sigma_a(\nu, T(z), Q(z)) (T_b - T(z)) \\ &= A(z, \nu, \theta) T_b + B(z, \nu, \theta) \end{aligned} \quad (2.3)$$

mit

$$\begin{aligned} A(z, \nu, \theta) &:= \sec(\theta) \sigma_a(\nu, T(z), Q(z)), \\ B(z, \nu, \theta) &:= -\sec(\theta) \sigma_a(\nu, T(z), Q(z)) T(z). \end{aligned}$$

Sind $T \in \mathcal{T}$, ein Winkel θ und eine Frequenz ν gegeben, so ist dies eine lineare gewöhnliche Differentialgleichung erster Ordnung in z . Wir setzen ferner voraus, dass die Atmosphärenparameter $Q(z)$ und der Absorptions- und Emissionskoeffizient σ_a stetig differenzierbar sind. Wir setzen ebenfalls voraus, dass die Ableitung von σ_a nach T , also $\partial_T \sigma_a$, stetig differenzierbar ist, dies wird in Abschnitt 2.4 und 2.6 benötigt. Also sind A und B stetig und (2.2) besitzt eine eindeutige Lösung, nämlich (2.1). Sei $T_b(\nu, \theta, z)$ die Lösung von (2.2), dann schreiben wir

$$T_b(\nu, \theta, z)[T] := T_b(\nu, \theta, z),$$

um die Abhängigkeit von T klar zu machen, und als Kurzschreibweise für die Strahlungstemperaturen in Bodenhöhe

$$T_b(\nu, \theta)[T] := T_b(\nu, \theta, z_0)[T].$$

Da die einzigen gegebenen Messdaten die Strahlungstemperaturen $T_{bm11}, \dots, T_{bmnm}$ sind, müssen wir nun ein realistisches $T \in \mathcal{T}$ finden, dessen theoretische Strahlungstemperaturen zu den Messdaten passen, d.h. wir suchen ein $T \in \mathcal{T}$, für das gilt

$$\begin{array}{c} \text{Messdaten} \quad \text{Theoretische Daten} \\ \forall_{i \in \{1, \dots, n\}} \forall_{k \in \{1, \dots, m\}} \quad T_{bmik} = T_b(\nu_i, \theta_k)[T]. \end{array} \quad (2.4)$$

Um zu prüfen, wie gut ein bestimmtes $T \in \mathcal{T}$ diese Bedingung erfüllt, definieren wir als Energie das quadratische Kostenfunktional $E : \mathcal{T} \rightarrow \mathbb{R}$ für dieses Problem durch

$$E[T] = \frac{1}{2} \sum_{i=1}^n \sum_{k=1}^m (T_{bmik} - T_b(\nu_i, \theta_k)[T])^2. \quad (2.5)$$

Für $T \in \mathcal{T}$ gilt also

$$T \text{ erfüllt (2.4)} \Leftrightarrow E[T] = 0,$$

d.h. T ist ein Minimum der Energie und somit ist das physikalische Problem in ein Optimierungsproblem umgewandelt. Das Problem ist schlecht gestellt, denn zumindest eine der Bedingungen für gut gestellte Probleme ist verletzt: Die Lösung ist nicht eindeutig (vgl. Abbildung 2.1). Die beiden Profile in dieser Abbildung wurden numerisch bestimmt: Das grüne Profil (nicht gemessen, sondern generiert mit einer e -Funktion, um ein glattes Ausgangsprofil zu haben) wurde als Ausgangsdatensatz benutzt, d.h. es wurden zu diesem Profil die zugehörigen Strahlungstemperaturen mit der Vorwärtsrechnung ermittelt und auf diesem Datensatz das Programm ausgeführt, das das rote Profil als Minimierer bestimmt hat. Dies ist natürlich nicht der eigentliche Programm-Code, sondern eine Abwandlung, die speziell für die Generierung des Beispiels geschrieben wurde.

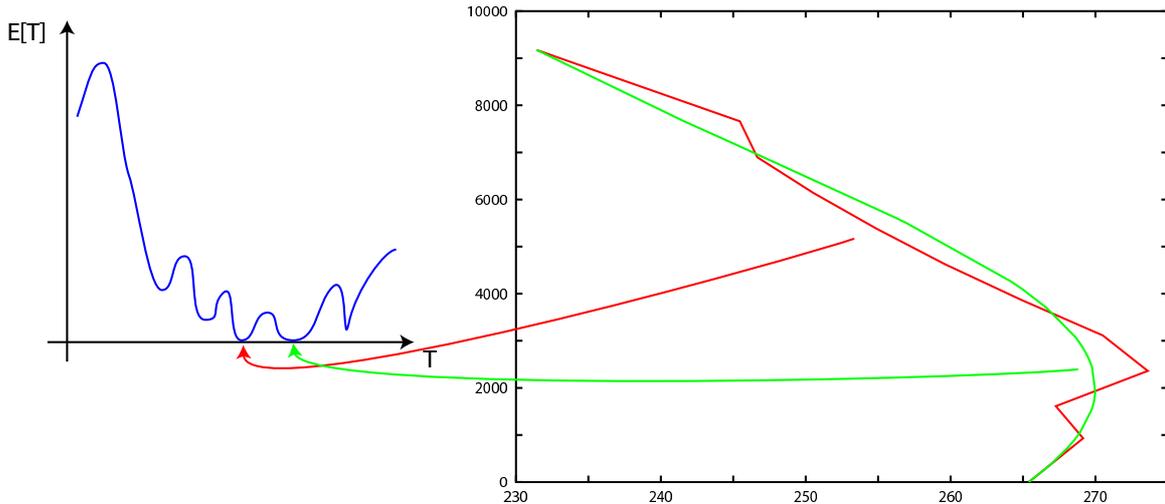


Abbildung 2.1: Das irreguläre, rote Profil und das glatte, grüne Profil sind beide Minima der Energie für einen bestimmten Datensatz

2.2 Gradientenfluss zur Energieminimierung

Ein bekanntes Verfahren zur Minimierung eines Funktionales ist der Gradientenfluss. Diesen wollen wir zur Lösung unseres Minimierungsproblems nutzen.

2.2.1 Motivation: Der endlich-dimensionale Fall

Sei $f \in C^1(\mathbb{R}^n, \mathbb{R})$. Um f zu minimieren, nutzen wir, dass $\text{grad } f(x) = \left(\frac{\partial}{\partial x_1} f(x), \dots, \frac{\partial}{\partial x_n} f(x) \right)$ lokal immer in die Richtung des steilsten Anstiegs zeigt, $-\text{grad } f(x)$ zeigt also in die Richtung des steilsten Abstiegs. Um f ausgehend von einem Startpunkt $x_0 \in \mathbb{R}^n$ zu minimieren, betrachten wir die Differentialgleichung

$$\begin{aligned} \frac{d}{dt} x(t) &= -\text{grad } f(x(t)), \\ x(0) &= x_0. \end{aligned}$$

Ist $x(t)$ eine Lösung, so gilt für $0 \leq t_1 < t_2$

$$f(x(t_1)) \geq f(x(t_2)),$$

denn:

$$\begin{aligned} f(x(t_2)) - f(x(t_1)) &= (f \circ x)(t_2) - (f \circ x)(t_1) = \int_{t_1}^{t_2} (f \circ x)'(t) dt \\ &= \int_{t_1}^{t_2} \text{grad } f(x(t)) \cdot \frac{d}{dt} x(t) dt = - \int_{t_1}^{t_2} \text{grad } f(x(t)) \cdot \text{grad } f(x(t)) dt \\ &= - \int_{t_1}^{t_2} \|\text{grad } f(x(t))\|_2^2 dt \leq 0. \end{aligned}$$

Da $\text{grad } f$ stetig ist, folgt hiermit sogar mehr: Ist $\text{grad } f(x(t_1)) \neq 0$ und $t_1 < t_2$, so folgt

$$f(x(t_1)) > f(x(t_2)).$$

Ist also x_0 kein kritischer Punkt von f (also $\text{grad } f(x_0) \neq 0$), so liefert die Differentialgleichung einen Punkt, an dem f echt kleiner ist als in x_0 . Ist allerdings x_0 ein kritischer Punkt von f , so ist die Lösung der Differentialgleichung die konstante Kurve $x(t) = x_0$ und das Verfahren minimiert f nicht weiter.

2.2.2 Gradientenflüsse auf Funktionenräumen

Sei X ein normierter, vollständiger reeller Funktionenraum und $E : X \rightarrow \mathbb{R}$ ein Funktional. Zu jedem $x \in X$ sei ein Skalarprodukt, bzw. eine Metrik, $g_x : X \times X \rightarrow \mathbb{R}$ gegeben. Dann ist der Gradient von E bzgl. der Metrik g_x definiert durch folgende Bedingung

$$\text{grad}_{g_x} E[x] = v \in X \Leftrightarrow \forall_{w \in X} g_x(v, w) = \langle E'[x], w \rangle.$$

Für einen Startwert $x_0 \in X$ ist der Gradientenfluss gegeben durch folgende Gleichung

$$\begin{aligned} \frac{\partial}{\partial t} x(t) &= -\text{grad}_{g_{x(t)}} E[x(t)], \\ x(0) &= x_0. \end{aligned}$$

Nach Definition des Gradienten muss also gelten

$$\forall_{w \in X} \quad g_{x(t)} \left(\frac{\partial}{\partial t} x(t), w \right) = - \langle E'[x(t)], w \rangle.$$

2.2.3 Der spezielle Gradientenfluss

Nun betrachten wir unseren Funktionenraum \mathcal{T} . Zu jedem $T \in \mathcal{T}$ sei eine Metrik g_T auf \mathcal{T} gegeben, derart, dass eine Darstellung

$$A_T : \mathcal{T} \rightarrow \mathbb{R}(A_T) := \{A_T v | v \in \mathcal{T}\} \subset \mathcal{T}'$$

von g_T existiert, d.h. es soll gelten

$$\forall_{T_1, T_2 \in \mathcal{T}} \quad g_T(T_1, T_2) = \langle A_T T_1, T_2 \rangle.$$

Der Gradientenfluss ist, wie eben gezeigt, definiert durch

$$\forall_{\vartheta \in \mathcal{T}} \quad g_T \left(\frac{\partial}{\partial s} T, \vartheta \right) = - \langle E'[T], \vartheta \rangle.$$

Setzen wir die Darstellung von g ein, so folgt

$$\forall_{\vartheta \in \mathcal{T}} \quad \left\langle A_T \frac{\partial}{\partial s} T, \vartheta \right\rangle = - \langle E'[T], \vartheta \rangle,$$

also

$$A_T \partial_s T = -E'[T].$$

Ist nun A_T auf $\mathbb{R}(A_T)$ invertierbar und ferner $E'[T] \in \mathbb{R}(A_T)$, so folgt

$$\partial_s T = -A_T^{-1} E'[T].$$

2.3 Wahl und Eigenschaften der Metrik

2.3.1 Wahl der Metrik

Da unsere Energie mehrere lokale und globale Minima haben kann, müssen wir dafür sorgen, dass der Gradientenfluss möglichst zu einem globalen und regulären Minimum führt. Daher benutzen wir die Metrik als regularisierendes Element, sie sorgt dafür, dass glatte Abstiegsrichtungen bevorzugt werden und vor allem eine reguläre Lösung aus der Menge der Minima gefunden wird. Zusammen mit der Zeitdiskretisierung, die wir in 3.3 benutzten, ist dieser Ansatz eng verwandt mit der iterativen Tikhonov Regularisierung (vgl. [14, 24, 8])

Der erste Ansatz für die Wahl von A ist es, die Metrik so zu wählen, dass A^{-1} ein Standardfilter ist (einfache Glättung, z. B. Konvexkombination). Mit einer geschickten Wahl der Metrik kann man dafür sorgen, dass Temperaturprofile, die gewisse Bedingungen erfüllen, wie z.B. Bedingungen, die für reale Temperaturprofile charakteristisch sind, begünstigt werden. Wir werden zwar auch dafür sorgen, dass unser Verfahren dies tut, aber nicht durch die Metrik, sondern durch einen zusätzlichen Regularisierungsterm in der Energie, vgl. 2.8, da es

relativ naheliegend ist, wie der Regularisierungsterm zu wählen ist. Wir entscheiden uns für eine Metrik, die für eine natürliche Glättung sorgt.

Zu gegebenem Temperaturprofil $T \in \mathcal{T}$ sei $\Gamma = \Gamma[T]$ die Kurve, die gegeben ist durch den Graphen von T :

$$\Gamma : [z_0, z_{\max}] \rightarrow \mathbb{R}^2, z \mapsto (z, T(z)).$$

Sind $v, w \in C^2(\Gamma, \mathbb{R})$, so wählen wir als Metrik

$$g_\Gamma(v, w) = g_{\Gamma,1}(v, w) + g_{\Gamma,2}(v, w)$$

mit

$$g_{\Gamma,1}(v, w) = \frac{1}{L(\Gamma)} \int_\Gamma v \cdot w ds,$$

$$g_{\Gamma,2}(v, w) = L(\Gamma) \int_\Gamma \nabla_\Gamma v \cdot \nabla_\Gamma w ds,$$

wobei für $x \in \Gamma$ der tangentielle Gradient von v in x mit $\nabla_\Gamma v(x)$ und die Länge der Kurve Γ mit $L(\Gamma)$ bezeichnet wird. Den Faktor $\frac{1}{L(\Gamma)}$ bzw. $L(\Gamma)$ wählen wir aus Normierungsgründen, um numerische Probleme zu verhindern, denn die Metrik verhält sich deswegen immer so, als wäre die Länge der Kurve auf 1 normiert. Dies zeigt folgendes Lemma:

2.3.i Lemma. *Sei $T \in \mathcal{T}$ und $\Gamma = \Gamma[T]$. Dann gilt für $c := L(\Gamma) > 0$ und $v, w \in C^2(\Gamma, \mathbb{R})$*

$$g_{\Gamma,1}(v, w) = g_{\frac{\Gamma}{c},1}(v \circ c, w \circ c),$$

$$g_{\Gamma,2}(v, w) = g_{\frac{\Gamma}{c},2}(v \circ c, w \circ c),$$

wobei $c \in \mathbb{R}$ auch als Abbildung $c : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto cx$ aufgefasst wird, $a\Gamma := \{ax | x \in \Gamma\}$ für $a \in \mathbb{R}$ und $\frac{\Gamma}{c} := \frac{1}{c}\Gamma$.

Beweis Seien $v, w \in C^2(\Gamma, \mathbb{R})$. Offensichtlich gilt dann $v \circ c, w \circ c \in C^2(\frac{\Gamma}{c}, \mathbb{R})$. Die erste Gleichung folgt direkt mit

$$g_{\Gamma,1}(v, w) = \frac{1}{L(\Gamma)} \int_\Gamma v(x) \cdot w(x) ds(x) \stackrel{(5.1.ix)}{=} \frac{1}{L(\Gamma)} \int_{\frac{\Gamma}{c}} v(cx) \cdot w(cx) cds(x)$$

$$= \frac{1}{L(\frac{\Gamma}{c})} \int_{\frac{\Gamma}{c}} v(cx) \cdot w(cx) ds(x) = g_{\frac{\Gamma}{c},1}(v \circ c, w \circ c).$$

Für die zweite Gleichung müssen wir zunächst einige Vorbereitungen treffen: $t_\Gamma(x)$ sei die Tangente im Punkt x an die Kurve Γ . Ist $a > 0$ und $\Gamma'(z) \neq 0$ (bei Graphen immer erfüllt) so gilt

$$t_\Gamma(\Gamma(z)) = \frac{\Gamma'(z)}{\|\Gamma'(z)\|} = \frac{a\Gamma'(z)}{\|a\Gamma'(z)\|} = t_{a\Gamma}(a\Gamma(z)).$$

Sei $x \in \frac{\Gamma}{c}$. Dann existiert ein $z \in [z_0, z_{\max}]$ mit $x = \frac{\Gamma(z)}{c}$ und es folgt

$$t_\Gamma(cx) = t_\Gamma(\Gamma(z)) = t_{\frac{\Gamma}{c}}\left(\frac{\Gamma(z)}{c}\right) = t_{\frac{\Gamma}{c}}(x).$$

Sind v, w nach $C^2([z_0, z_{\max}] \times \mathbb{R}, \mathbb{R})$ fortgesetzt (Da Γ ein Graph ist, ist dies auf jeden Fall möglich), so gilt

$$\nabla_{\Gamma} v(x) = (\mathbb{1} - n_{\Gamma}(x) \otimes n_{\Gamma}(x)) \nabla v(x),$$

wobei $n_{\Gamma}(x)$ die obere Normale an Γ im Punkt x bezeichnet. Für $x \in \frac{\Gamma}{c}$ gilt dann

$$n_{\Gamma}(cx) = t_{\Gamma}(cx)^{\perp} = t_{\frac{\Gamma}{c}}(x)^{\perp} = n_{\frac{\Gamma}{c}}(x)$$

und es folgt

$$\begin{aligned} \nabla_{\frac{\Gamma}{c}}(v \circ c)(x) &= (\mathbb{1} - n_{\frac{\Gamma}{c}}(x) \otimes n_{\frac{\Gamma}{c}}(x)) \nabla(v \circ c)(x) \\ &= (\mathbb{1} - n_{\Gamma}(cx) \otimes n_{\Gamma}(cx)) c \nabla v(cx) = c \nabla_{\Gamma} v(cx). \end{aligned}$$

Ebenso gilt

$$\nabla_{\frac{\Gamma}{c}}(w \circ c)(x) = c \nabla_{\Gamma} w(cx),$$

so dass insgesamt folgt

$$\begin{aligned} g_{\frac{\Gamma}{c}, 2}(v \circ c, w \circ c) &= L\left(\frac{\Gamma}{c}\right) \int_{\frac{\Gamma}{c}} \nabla_{\frac{\Gamma}{c}}(v \circ c)(x) \cdot \nabla_{\frac{\Gamma}{c}}(w \circ c)(x) ds = 1 \cdot \int_{\frac{\Gamma}{c}} c \nabla_{\Gamma} v(cx) \cdot c \nabla_{\Gamma} w(cx) ds \\ &= c \int_{\frac{\Gamma}{c}} \nabla_{\Gamma} v(cx) \cdot \nabla_{\Gamma} w(cx) c ds \stackrel{(5.1.ix)}{=} c \int_{\Gamma} \nabla_{\Gamma} v(x) \cdot \nabla_{\Gamma} w(x) ds \\ &= g_{\Gamma, 2}(v, w). \end{aligned}$$

□

Aus dem Lemma folgt also direkt

$$g_{\Gamma}(v, w) = g_{\frac{\Gamma}{c}}(v \circ c, w \circ c),$$

also die behauptete Normierungseigenschaft. Nun benötigen wir aber eine Metrik auf $\mathcal{T} \times \mathcal{T}$ und nicht auf $C^2(\Gamma, \mathbb{R}) \times C^2(\Gamma, \mathbb{R})$, also müssen wir einem $v \in \mathcal{T}$ ein Element aus $C^2(\Gamma, \mathbb{R})$ zuordnen. Dazu sei

$$p_1 : \mathbb{R}^2 \rightarrow \mathbb{R}, (x_1, x_2) \mapsto x_1$$

die kanonische Projektion auf die erste Koordinate. Ist $v \in \mathcal{T}$, so ist $v \circ p_1 \in C^2([z_0, z_{\max}] \times \mathbb{R}, \mathbb{R}) \supset C^2(\Gamma, \mathbb{R})$ und wir setzen für $v, w \in \mathcal{T}$:

$$g_{\mathcal{T}}(v, w) := g_{\Gamma}(v \circ p_1, w \circ p_1) = \underbrace{g_{\Gamma, 1}(v \circ p_1, w \circ p_1)}_{=: g_{\mathcal{T}, 1}(v, w)} + \underbrace{g_{\Gamma, 2}(v \circ p_1, w \circ p_1)}_{=: g_{\mathcal{T}, 2}(v, w)}.$$

2.3.ii Lemma. Für $v, w \in \mathcal{T}$ gilt

$$\begin{aligned} g_{\mathcal{T}, 1}(v, w) &= \frac{1}{L(\Gamma)} \int_{z_0}^{z_{\max}} v(z) \cdot w(z) \sqrt{1 + T'(z)^2} dz, \\ g_{\mathcal{T}, 2}(v, w) &= L(\Gamma) \int_{z_0}^{z_{\max}} \frac{v'(z) w'(z)}{\sqrt{1 + T'(z)^2}} dz. \end{aligned}$$

Beweis Seien $v \in \mathcal{T}$ und $z \in [z_0, z_{\max}]$. Dann gilt $(z, T(z)) = \Gamma(z) \in \Gamma$ und ferner

$$n(\Gamma(z)) := n_\Gamma(\Gamma(z)) = \frac{(-T'(z), 1)}{\|(-T'(z), 1)\|} = \frac{(-T'(z), 1)}{\sqrt{1 + T'(z)^2}}.$$

Also gilt

$$n(\Gamma(z)) \otimes n(\Gamma(z)) = \frac{1}{1 + T'(z)^2} \begin{pmatrix} T'(z)^2 & -T'(z) \\ -T'(z) & 1 \end{pmatrix}$$

und damit folgt

$$\begin{aligned} \nabla_\Gamma(v \circ p_1)(\Gamma(z)) &= (\mathbb{1} - n(\Gamma(z)) \otimes n(\Gamma(z))) \nabla(v \circ p_1)(\Gamma(z)) \\ &= (\mathbb{1} - n(\Gamma(z)) \otimes n(\Gamma(z))) v'(p_1(\Gamma(z))) \nabla(p_1)(\Gamma(z)) \\ &= (\mathbb{1} - n(\Gamma(z)) \otimes n(\Gamma(z))) v'(z)(1, 0) \\ &= v'(z) \left((1, 0) - \frac{1}{1 + T'(z)^2} (T'(z)^2, -T'(z)) \right) \\ &= \frac{v'(z)}{1 + T'(z)^2} ((1 + T'(z)^2, 0) - (T'(z)^2, -T'(z))) \\ &= \frac{v'(z)}{1 + T'(z)^2} (1, T'(z)). \end{aligned}$$

Also folgt für $v, w \in \mathcal{T}$

$$\begin{aligned} g_{T,2}(v, w) &= L(\Gamma) \int_\Gamma \nabla_\Gamma(v \circ p_1)(x) \cdot \nabla_\Gamma(w \circ p_1)(x) ds(x) \\ &= L(\Gamma) \int_{z_0}^{z_{\max}} \nabla_\Gamma(v \circ p_1)(\Gamma(z)) \cdot \nabla_\Gamma(w \circ p_1)(\Gamma(z)) |\Gamma'(z)| dz \\ &= L(\Gamma) \int_{z_0}^{z_{\max}} \frac{v'(z)}{1 + T'(z)^2} (1, T'(z)) \cdot \frac{w'(z)}{1 + T'(z)^2} (1, T'(z)) \sqrt{1 + T'(z)^2} dz \\ &= L(\Gamma) \int_{z_0}^{z_{\max}} \frac{v'(z)w'(z)}{(1 + T'(z)^2)^2} (1 + T'(z)^2) \sqrt{1 + T'(z)^2} dz \\ &= L(\Gamma) \int_{z_0}^{z_{\max}} \frac{v'(z)w'(z)}{\sqrt{1 + T'(z)^2}} dz, \end{aligned}$$

$$\begin{aligned} g_{T,1}(v, w) &= \frac{1}{L(\Gamma)} \int_\Gamma (v \circ p_1)(x) \cdot (w \circ p_1)(x) ds(x) \\ &= \frac{1}{L(\Gamma)} \int_{z_0}^{z_{\max}} (v \circ p_1)(\Gamma(z)) \cdot (w \circ p_1)(\Gamma(z)) |\Gamma'(z)| dz \\ &= \frac{1}{L(\Gamma)} \int_{z_0}^{z_{\max}} v(z) \cdot w(z) \sqrt{1 + T'(z)^2} dz. \end{aligned}$$

□

2.3.2 Invertierbarkeit der Metrik

Jetzt, da wir das Kurvenintegral in ein Integral einer Veränderlichen umgewandelt haben, kümmern wir uns um die Existenz einer invertierbaren Darstellung $A_T : \mathcal{T} \rightarrow \mathbb{R}(A_T)$ der Metrik, da wir eine solche benötigen, um den Gradientenfluss folgendermaßen darzustellen (falls $E'[T] \in \mathbb{R}(A_T)$):

$$\partial_s T = -A_T^{-1} E'[T].$$

Diese Darstellung des Gradientenflusses benötigen wir später, um die Kurzzeit-Existenz des Gradientenflusses zu zeigen.

Wir zeigen, dass A_T aufgefasst werden kann als Abbildung

$$A_T : \mathcal{T} = C^2([z_0, z_{\max}]) \rightarrow C^0([z_0, z_{\max}]) \times \mathbb{R}^2$$

und dann bijektiv ist. Dabei gilt $C^0([z_0, z_{\max}]) \times \mathbb{R}^2 \subset \mathcal{T}'$, indem wir ein Element $(f, a, b) \in C^0([z_0, z_{\max}]) \times \mathbb{R}^2$ mit einem Element von \mathcal{T}' identifizieren, mittels

$$\forall_{w \in \mathcal{T}} \quad \langle (f, a, b), w \rangle = \int_{z_0}^{z_{\max}} f(z)w(z) + bw(z_{\max}) - aw(z_0). \quad (2.6)$$

Es gilt

$$\begin{aligned} \langle A_T v, w \rangle &= \frac{1}{L(\Gamma)} \int_{z_0}^{z_{\max}} v(z) \cdot w(z) \sqrt{1 + T'(z)^2} dz + L(\Gamma) \int_{z_0}^{z_{\max}} \frac{v'(z)w'(z)}{\sqrt{1 + T'(z)^2}} dz \\ &= \frac{1}{L(\Gamma)} \int_{z_0}^{z_{\max}} v(z) \cdot w(z) \sqrt{1 + T'(z)^2} dz - L(\Gamma) \int_{z_0}^{z_{\max}} \left(\frac{v'(z)}{\sqrt{1 + T'(z)^2}} \right)' w(z) dz \\ &\quad + \frac{L(\Gamma)v'(z_{\max})}{\sqrt{1 + T'(z_{\max})^2}} w(z_{\max}) - \frac{L(\Gamma)v'(z_0)}{\sqrt{1 + T'(z_0)^2}} w(z_0) \\ &= \int_{z_0}^{z_{\max}} \left(\frac{\sqrt{1 + T'(z)^2}}{L(\Gamma)} v(z) - \left(\frac{L(\Gamma)}{\sqrt{1 + T'(z)^2}} v'(z) \right)' \right) w(z) dz \\ &\quad + \frac{L(\Gamma)v'(z_{\max})}{\sqrt{1 + T'(z_{\max})^2}} w(z_{\max}) - \frac{L(\Gamma)v'(z_0)}{\sqrt{1 + T'(z_0)^2}} w(z_0), \end{aligned} \quad (2.7)$$

für beliebige $v, w \in \mathcal{T}$. Es gilt $A_T v =_{\mathcal{T}'} (f, a, b)$, genau dann wenn

$$\forall_{w \in \mathcal{T}} \quad \langle A_T v, w \rangle = \langle (f, a, b), w \rangle, \quad (2.8)$$

wobei $a =_{\mathcal{T}'} b$ bedeutet, dass a und b als Element von \mathcal{T}' gleich sind. Wegen (2.7) gilt offenbar

$$\begin{aligned} A_T v =_{\mathcal{T}'} \left(\left(\frac{\sqrt{1 + T'(z)^2}}{L(\Gamma)} v(z) - \left(\frac{L(\Gamma)}{\sqrt{1 + T'(z)^2}} v'(z) \right)' \right), \right. \\ \left. \frac{L(\Gamma)v'(z_0)}{\sqrt{1 + T'(z_0)^2}}, \frac{L(\Gamma)v'(z_{\max})}{\sqrt{1 + T'(z_{\max})^2}} \right). \end{aligned} \quad (2.9)$$

Daher ist $A_T v \in C^0([z_0, z_{\max}]) \times \mathbb{R}^2$ und somit ist $A_T : C^2([z_0, z_{\max}]) \rightarrow C^0([z_0, z_{\max}]) \times \mathbb{R}^2$ wohldefiniert.

Gilt nun $A_T v =_{\mathcal{T}'} (f, a, b)$, so folgt aus (2.8), eingeschränkt auf $w \in C_0^\infty$, und (2.9)

$$f(z) = \left(\frac{\sqrt{1+T'(z)^2}}{L(\Gamma)} v(z) - \left(\frac{L(\Gamma)}{\sqrt{1+T'(z)^2}} v'(z) \right)' \right).$$

Für $w_1 \in \mathcal{T}$ mit $w_1(z_0) \neq 0$ und $w_1(z_{\max}) = 0$, bzw. $w_2 \in \mathcal{T}$ mit $w_2(z_0) = 0$ und $w_2(z_{\max}) \neq 0$ folgt nun mit (2.8) und (2.9)

$$a = \frac{L(\Gamma)v'(z_0)}{\sqrt{1+T'(z_0)^2}}, \text{ bzw. } b = \frac{L(\Gamma)v'(z_{\max})}{\sqrt{1+T'(z_{\max})^2}}.$$

Wir haben somit gezeigt: $A_T v =_{\mathcal{T}'} (f, a, b)$ gilt genau dann, wenn v Lösung des Randwertproblems

$$\begin{aligned} f(z) &= q_T(z)v(z) - (p_T(z)v'(z))' =: (L_T v)(z), \\ R_{1,T} v &:= p_T(z_0)v'(z_0) = a, \quad R_{2,T} v := p_T(z_{\max})v'(z_{\max}) = b \end{aligned} \quad (2.10)$$

ist, wobei

$$p_T(z) := \frac{L(\Gamma)}{\sqrt{1+T'(z)^2}}, \quad q_T(z) := \frac{\sqrt{1+T'(z)^2}}{L(\Gamma)}.$$

2.3.iii Satz. Seien $a < b$, $p \in C^1([a, b])$ mit $p > 0$, $q \in C^0([a, b])$ mit $q > 0$, $\alpha_1, \alpha_2, \beta_1, \beta_2 \in \mathbb{R}$ mit $\alpha_1 \alpha_2 \leq 0$ und $\beta_1 \beta_2 \geq 0$ und $(\alpha_1, \alpha_2), (\beta_1, \beta_2) \neq (0, 0)$. Dann ex. zu allen $f \in C^0([a, b])$ und $\varrho_1, \varrho_2 \in \mathbb{R}$ genau ein $v \in C^2([a, b])$ mit

$$\begin{aligned} (p(x)v'(x))' - q(x)v(x) &= f(x), \\ R_1 u &:= \alpha_1 u(a) + \alpha_2 u'(a) = \varrho_1, \quad R_2 u := \beta_1 u(b) + \beta_2 u'(b) = \varrho_2. \end{aligned}$$

Beweis Nach [15] (S. 389, Aufgabe 9) sind bei obigen Voraussetzungen alle Eigenwerte λ der Sturm-Liouvilleschen Eigenwertaufgabe

$$(p(x)v'(x))' + \lambda q(x)v(x) = 0, \quad R_1 u = 0, \quad R_2 u = 0$$

nichtnegativ. Insbesondere ist -1 kein Eigenwert, d.h. das homogene Randwertproblem

$$Lv := (p(x)v'(x))' - q(x)v(x) = 0, \quad R_1 u = 0, \quad R_2 u = 0$$

hat nur die triviale Lösung. Dann folgt mit [15] (Satz 35.2), da dieser (mit analogem Beweis) auch für Sturmische Randwertaufgaben gilt, die eindeutige Lösbarkeit der Randwertaufgabe aus der Behauptung. \square

Sei $(f, a, b) \in C^0([z_0, z_{\max}]) \times \mathbb{R}^2$. Nach 2.3.iii gibt es genau ein $v \in C^2([z_0, z_{\max}])$ mit

$$\begin{aligned} (p_T(z)v'(z))' - q_T(z)v(z) &= -f(z), \\ p_T(z_0)v'(z_0) &= a, \quad p_T(z_{\max})v'(z_{\max}) = b, \end{aligned}$$

und somit hat (2.10) genau eine Lösung. Damit ist $A_T : C^2([z_0, z_{\max}]) \rightarrow C^0([z_0, z_{\max}]) \times \mathbb{R}^2$ bijektiv. Ferner haben wir gezeigt, dass gilt

$$R(A_T) =_{\mathcal{T}'} C^0([z_0, z_{\max}]) \times \mathbb{R}^2 =: \mathcal{S}. \quad (2.11)$$

Anders als im mehrdimensionalen Fall bei partiellen Differentialgleichungen, ist es hier nicht nötig zusätzlich Hölder-Stetigkeit zu fordern, um die Invertierbarkeit zu erhalten (vgl. [12]).

2.3.3 Stetigkeit der Inversen der Metrik

Für die Existenz des Gradientenflusses brauchen wir auch die Stetigkeit der Inversen der Metrik. Dazu benötigen wir die stetige Abhängigkeit der Lösung einer halbhomogenen Sturmischen Randwertaufgabe von den Daten und den Koeffizienten.

Im Folgenden ist mit C^0 immer $C^0([z_0, z_{\max}])$ gemeint, analog für C^1 und C^2 .

2.3.iv Satz. (*Stetige Abhängigkeit von den Daten und den Koeffizienten*) Sei $T \in \mathcal{T}$ beliebig. Für $S \in \mathcal{T}, g \in C^0$ sei $v_{(S,g)}$ die Lösung von

$$L_S v_{(S,g)} = g, \quad R_{1,S} v_{(S,g)} = 0, \quad R_{2,S} v_{(S,g)} = 0.$$

Dann gilt

$$\forall_{T \in \mathcal{T}} \quad \forall_{f \in C^0} \quad \forall_{\epsilon > 0} \quad \exists_{\delta > 0} \quad \forall_{\substack{S \in \mathcal{T} \\ \|T-S\|_{\mathcal{T}} < \delta}} \quad \forall_{\substack{g \in C^0 \\ \|f-g\|_{C^0} < \delta}} \quad \|v_{(T,f)} - v_{(S,g)}\|_{\mathcal{T}} < \epsilon.$$

Beweis Seien $T \in \mathcal{T}$ und $f \in C^0$ beliebig, aber fest. Angenommen die Behauptung ist falsch, dann gilt

$$\exists_{c > 0} \quad \forall_{n \in \mathbb{N}} \quad \exists_{\substack{T_n \in \mathcal{T} \\ \|T-T_n\|_{\mathcal{T}} < \frac{1}{n}}} \quad \exists_{\substack{f_n \in C^0 \\ \|f-f_n\|_{C^0} < \frac{1}{n}}} \quad \|v_{(T,f)} - v_{(T_n,f_n)}\|_{\mathcal{T}} \geq c.$$

Seien $v := v_{(T,f)}$ und $v_n := v_{(T_n,f_n)}$. Offenbar konvergiert T_n gegen T in der \mathcal{T} -Norm und f_n konvergiert gleichmäßig gegen f . Es folgt die gleichmäßige Konvergenz von q_{T_n} gegen q_T und die Konvergenz von p_{T_n} gegen p_T in der C^1 -Norm (folgt mit Hilfe der Aussagen über h im Beweis von 2.3.vii). Insbesondere ist q_{T_n} in der C^0 -Norm beschränkt. Mit den Sätzen 21.2, 21.3 und 36.1 aus [15] kann man zeigen, dass nun v_n gleichmäßig gegen v konvergiert. Aus

$$q_{T_n} v_n - (p_{T_n} v_n')' = L_{T_n} v_n = f_n \quad \text{und} \quad q_T v - (p_T v')' = L_T v = f$$

folgt für $g_n := p_{T_n} v_n' - p_T v'$

$$\begin{aligned} \|g_n'\|_{C^0} &= \left\| (p_{T_n} v_n' - p_T v')' \right\|_{C^0} \leq \|f_n - f\|_{C^0} + \|q_{T_n} v_n - q_T v\|_{C^0} \\ &\leq \|f_n - f\|_{C^0} + \|q_{T_n} v_n - q_{T_n} v\|_{C^0} + \|q_{T_n} v - q_T v\|_{C^0} \\ &\leq \underbrace{\|f_n - f\|_{C^0}}_{\rightarrow 0 \text{ für } n \rightarrow \infty} + \underbrace{\|q_{T_n}\|_{C^0}}_{\text{beschr.}} \underbrace{\|v_n - v\|_{C^0}}_{\rightarrow 0 \text{ für } n \rightarrow \infty} + \underbrace{\|q_{T_n} - q_T\|_{C^0}}_{\rightarrow 0 \text{ für } n \rightarrow \infty} \|v\|_{C^0}. \end{aligned}$$

Somit konvergiert g_n' gleichmäßig gegen 0. Ferner gilt

$$g_n(z_0) = p_{T_n}(z_0) v_n'(z_0) - p_T(z_0) v'(z_0) = R_{1,T_n} v_n - R_{1,T} v = 0 - 0 = 0,$$

also konvergiert insbesondere $g_n(z_0)$ gegen 0. Nach 5.1.ii konvergiert g_n gleichmäßig gegen ein $g \in C^1$ und ist insbesondere in der C^0 -Norm beschränkt.

Wegen $p_T(z) > 0$ für alle $z \in [z_0, z_{\max}]$ und der Stetigkeit von p_T , existiert ein $d > 0$ mit $|p_T(z)| = p_T(z) \geq d$ und es folgt $\left\| \frac{1}{p_T} \right\|_{C^0} \leq \frac{1}{d}$. Da p_{T_n} gleichmäßig gegen p_T konvergiert, existiert ein $N \in \mathbb{N}$, so dass für alle $n \geq N$ gilt $\|p_{T_n} - p_T\|_{C^0} \leq \frac{d}{2}$. Für $z \in [z_0, z_{\max}]$ folgt

$$|p_{T_n}(z)| \geq |p_T(z)| - |p_{T_n}(z) - p_T(z)| \geq d - \frac{d}{2} = \frac{d}{2} \quad \Rightarrow \quad \left| \frac{1}{p_{T_n}(z)} \right| \leq \frac{2}{d},$$

also gilt $\left\| \frac{1}{p_{T_n}} \right\|_{C^0} \leq \frac{2}{d}$. Nun folgt

$$\begin{aligned} \|v'_n\|_{C^0} &\leq \left\| \frac{1}{p_{T_n}} \right\|_{C^0} \|p_{T_n} v'_n\|_{C^0} \leq \frac{2}{d} (\|p_T v'\|_{C^0} + \|p_{T_n} v'_n - p_T v'\|_{C^0}) \\ &= \frac{2}{d} \|p_T v'\|_{C^0} + \frac{2}{d} \underbrace{\|g_n\|_{C^0}}_{\text{beschr.}}, \end{aligned}$$

d.h. v'_n ist in der C^0 -Norm beschränkt. Ferner gilt

$$\begin{aligned} \left\| v'_n - \frac{1}{p_T} (p_T v' + g) \right\|_{C^0} &\leq \left\| \frac{1}{p_T} \right\|_{C^0} \|p_T v'_n - p_T v' - g\|_{C^0} \\ &\leq \frac{1}{d} (\|p_T v'_n - p_{T_n} v'_n\|_{C^0} + \|p_{T_n} v'_n - p_T v' - g\|_{C^0}) \\ &\leq \frac{1}{d} (\|p_T - p_{T_n}\|_{C^0} \|v'_n\|_{C^0} + \|p_{T_n} v'_n - p_T v' - g\|_{C^0}) \\ &= \frac{1}{d} \underbrace{\|p_T - p_{T_n}\|_{C^0}}_{\rightarrow 0 \text{ für } n \rightarrow \infty} \underbrace{\|v'_n\|_{C^0}}_{\text{beschr.}} + \frac{1}{d} \underbrace{\|g_n - g\|_{C^0}}_{\rightarrow 0 \text{ für } n \rightarrow \infty} \end{aligned}$$

und somit konvergiert v'_n gleichmäßig. Da auch v_n gleichmäßig gegen v konvergiert folgt nun mit Satz 5.1.ii und aus der Eindeutigkeit des Grenzwertes, dass

$$\lim_{n \rightarrow \infty} v'_n(z) = v'(z).$$

Insgesamt konvergiert also v'_n gegen v' . Aus

$$q_{T_n} v_n - p'_{T_n} v'_n - p_{T_n} v''_n = L_{T_n} v_n = f_n \text{ und } q_T v - p'_T v' - p_T v'' = L_T v = f$$

folgt

$$\begin{aligned} \|p_{T_n} v''_n - p_T v''\|_{C^0} &\leq \|f_n - f\|_{C^0} + \|q_{T_n} v_n - q_T v\|_{C^0} + \|p'_{T_n} v'_n - p'_T v'\|_{C^0} \\ &\leq \|f_n - f\|_{C^0} + \|q_{T_n} v_n - q_T v_n\|_{C^0} + \|q_T v_n - q_T v\|_{C^0} \\ &\quad + \|p'_{T_n} v'_n - p'_T v'_n\|_{C^0} + \|p'_T v'_n - p'_T v'\|_{C^0} \\ &\leq \underbrace{\|f_n - f\|_{C^0}}_{\rightarrow 0 \text{ für } n \rightarrow \infty} + \underbrace{\|q_{T_n} - q_T\|_{C^0}}_{\rightarrow 0 \text{ für } n \rightarrow \infty} \underbrace{\|v_n\|_{C^0}}_{\text{beschr.}} + \|q_T\|_{C^0} \underbrace{\|v_n - v\|_{C^0}}_{\rightarrow 0 \text{ für } n \rightarrow \infty} \\ &\quad + \underbrace{\|p'_{T_n} - p'_T\|_{C^0}}_{\rightarrow 0 \text{ für } n \rightarrow \infty} \underbrace{\|v'_n\|_{C^0}}_{\text{beschr.}} + \|p'_T\|_{C^0} \underbrace{\|v'_n - v'\|_{C^0}}_{\rightarrow 0 \text{ für } n \rightarrow \infty}. \end{aligned}$$

Damit ergibt sich

$$\|v''_n\|_{C^0} \leq \left\| \frac{1}{p_{T_n}} \right\|_{C^0} \|p_{T_n} v''_n\|_{C^0} \leq \frac{2}{d} \|p_T v''\|_{C^0} + \frac{2}{d} \underbrace{\|p_{T_n} v''_n - p_T v''\|_{C^0}}_{\rightarrow 0 \text{ für } n \rightarrow \infty},$$

d.h. v''_n ist in der C^0 -Norm beschränkt. Ferner gilt

$$\begin{aligned} \|v''_n - v''\|_{C^0} &\leq \left\| \frac{1}{p_T} \right\|_{C^0} \|p_T v''_n - p_T v''\|_{C^0} \\ &\leq \frac{1}{d} (\|p_T v''_n - p_{T_n} v''_n\|_{C^0} + \|p_{T_n} v''_n - p_T v''\|_{C^0}) \\ &\leq \frac{1}{d} \underbrace{\|p_T - p_{T_n}\|_{C^0}}_{\rightarrow 0 \text{ für } n \rightarrow \infty} \underbrace{\|v''_n\|_{C^0}}_{\text{beschr.}} + \frac{1}{d} \underbrace{\|p_{T_n} v''_n - p_T v''\|_{C^0}}_{\rightarrow 0 \text{ für } n \rightarrow \infty}, \end{aligned}$$

d.h. v_n'' konvergiert gleichmäßig gegen v'' . Insgesamt konvergiert v_n in der \mathcal{T} -Norm gegen v . Dies ist aber ein Widerspruch zu $\|v - v_n\|_{\mathcal{T}} = \|v_{(T,f)} - v_{(T_n,f_n)}\|_{\mathcal{T}} \geq c$ für alle $n \in \mathbb{N}$. Somit ist die Annahme falsch und die Behauptung richtig. \square

Ein weiterer Baustein, der für die Stetigkeit der Inversen der Metrik benötigt wird, ist das folgende Lemma:

2.3.v Lemma. *Zu jedem $T \in \mathcal{T}$ ex. ein $c > 0$, so dass für alle $v \in \mathcal{T}$ mit $v'(z_0) = v'(z_{\max}) = 0$ gilt*

$$\|L_T v\|_{C^0} \geq c \|v\|_{\mathcal{T}}.$$

Beweis Sei $T \in \mathcal{T}$ beliebig. Für $v = 0$ ist die Aussage trivial. Für $v \neq 0$ gilt

$$\|L_T v\|_{C^0} \geq c \|v\|_{\mathcal{T}} \Leftrightarrow \left\| L_T \frac{v}{\|v\|_{\mathcal{T}}} \right\|_{C^0} \geq c,$$

es genügt also die Aussage für $v \in X := \{w \in \mathcal{T} | w'(z_0) = w'(z_{\max}) = 0, \|w\|_{\mathcal{T}} = 1\}$ zu beweisen. Angenommen die Aussage ist falsch, dann gilt insbesondere

$$\forall_{n \in \mathbb{N}} \quad \exists_{v_n \in X} \quad \|L_T v_n\|_{C^0} < \frac{1}{n}.$$

v_n ist Lösung des Randwertproblems

$$L_T v_n =: f_n, \quad R_{1,T} v_n = p_T(z_0) v_n'(z_0) = 0, \quad R_{2,T} v_n = p_T(z_{\max}) v_n'(z_{\max}) = 0.$$

Für $g \in C^0$ sei $v_g = v_{(T,g)}$, wobei $v_{(T,g)}$ wie in 2.3.iv. Dann existiert nach 2.3.iv ein $\delta > 0$, so dass gilt

$$\forall_{\substack{g \in C^0 \\ \|0-g\|_{C^0} < \delta}} \quad \|v_0 - v_g\|_{\mathcal{T}} < 1.$$

0 und v_0 lösen

$$L_T v = 0, \quad R_{1,T} v = 0, \quad R_{2,T} v = 0,$$

also folgt aus der eindeutigen Lösbarkeit $v_0 = 0$. Für n mit $\frac{1}{n} < \delta$ gilt

$$\|f_n\|_{C^0} = \|L_T v_n\|_{C^0} < \frac{1}{n} < \delta$$

und es folgt

$$1 = \|v_n\|_{\mathcal{T}} = \|v_n - 0\|_{\mathcal{T}} = \|v_{f_n} - v_0\|_{\mathcal{T}} < 1.$$

Widerspruch. Also war die Annahme falsch und es folgt die Behauptung. \square

2.3.vi Bemerkung. Die Aussage von 2.3.v ist falsch, wenn man auf Nullrandwerte der Ableitung verzichtet: Das Randwertproblem

$$\begin{aligned} L_T v &= 0, \\ R_{1,T} v &= 1, \quad R_{2,T} v = 0, \end{aligned}$$

hat eine Lösung, die offenbar von Null verschieden ist. Sei v diese Lösung. Wegen $v \neq 0$ gilt $\|v\|_{\mathcal{T}} \neq 0$ und es folgt für beliebiges $c > 0$

$$\|L_T v\|_{C^0([z_0, z_{\max}])} = 0 < c \|v\|_{\mathcal{T}}.$$

2.3.vii Satz. Die Inverse der Metrik ist in folgendem Sinne stetig:

$$\forall_{T \in \mathcal{T}} \quad \forall_{f \in C^0} \quad \forall_{\epsilon > 0} \quad \exists_{\delta > 0} \quad \forall_{\substack{S \in \mathcal{T} \\ \|T-S\|_{\mathcal{T}} < \delta}} \quad \forall_{\substack{g \in C^0 \\ \|f-g\|_{C^0} < \delta}} \quad \|A_T^{-1}(f, 0, 0) - A_S^{-1}(g, 0, 0)\|_{\mathcal{T}} < \epsilon.$$

Beweis Seien $T \in \mathcal{T}$, $f \in C^0$ und $\epsilon > 0$ beliebig, aber fest. Nach 2.3.v existiert ein $c > 0$, so dass für alle $v \in \mathcal{T}$ mit $v'(z_0) = v'(z_{\max}) = 0$ gilt

$$\|L_T v\|_{C^0} \geq c \|v\|_{\mathcal{T}}.$$

Für $S \in \mathcal{T}$, $g \in C^0$ sei $v_{(S,g)} := A_S^{-1}(g, 0, 0) \in \mathcal{T}$. Dann gilt

$$L_S v_{(S,g)} = g, \quad R_{1,S} v_{(S,g)} = 0, \quad R_{2,S} v_{(S,g)} = 0.$$

Daraus folgt $v'_{(S,g)}(z_0) = v'_{(S,g)}(z_{\max}) = 0$ und

$$\begin{aligned} & L_T v_{(T,f)} - L_S v_{(S,g)} = f - g \\ \Rightarrow & q_T v_{(T,f)} - \left(p_T v'_{(T,f)}\right)' - q_S v_{(S,g)} + \left(p_S v'_{(S,g)}\right)' = f - g \\ \Rightarrow & q_T v_{(T,f)} - p'_T v'_{(T,f)} - p_T v''_{(T,f)} - q_S v_{(S,g)} + p'_S v'_{(S,g)} + p_S v''_{(S,g)} = f - g \\ \Rightarrow & q_T (v_{(T,f)} - v_{(S,g)}) - p'_T (v'_{(T,f)} - v'_{(S,g)}) - p_T (v''_{(T,f)} - v''_{(S,g)}) \\ \Rightarrow & = (q_S - q_T) v_{(S,g)} - (p'_S - p'_T) v'_{(S,g)} - (p_S - p_T) v''_{(S,g)} + f - g \\ \Rightarrow & \|L_T (v_{(T,f)} - v_{(S,g)})\|_{C^0} \leq \|v_{(S,g)}\|_{\mathcal{T}} (\|q_S - q_T\|_{C^0} + \|p_S - p_T\|_{C^1}) + \|f - g\|_{C^0}. \end{aligned}$$

Wegen $(v_{(T,f)}(z_0) - v_{(S,g)}(z_0))' = (v_{(T,f)}(z_{\max}) - v_{(S,g)}(z_{\max}))' = 0$ folgt

$$\begin{aligned} & \|A_T^{-1}(f, 0, 0) - A_S^{-1}(g, 0, 0)\|_{\mathcal{T}} = \|v_{(T,f)} - v_{(S,g)}\|_{\mathcal{T}} \leq \frac{1}{c} \|L_T (v_{(T,f)} - v_{(S,g)})\|_{C^0} \\ & \leq \frac{1}{c} \|v_{(S,g)}\|_{\mathcal{T}} (\|q_S - q_T\|_{C^0} + \|p_S - p_T\|_{C^1}) + \frac{1}{c} \|f - g\|_{C^0}. \end{aligned}$$

Wegen der stetigen Abhängigkeit der Lösung der halbhomogenen Sturmischen Randwertaufgabe von den Daten und den Koeffizienten (vgl. 2.3.iv), folgt die Existenz eines $\delta_1 > 0$, so dass für alle $S \in \mathcal{T}$ und alle $g \in C^0$ gilt

$$\begin{aligned} & \|T - S\|_{\mathcal{T}} < \delta_1, \|f - g\|_{C^0} < \delta_1 \Rightarrow \|v_{(S,g)} - v_{(T,f)}\|_{\mathcal{T}} < 1 \\ & \Rightarrow \|v_{(S,g)}\|_{\mathcal{T}} \leq \|v_{(S,g)} - v_{(T,f)}\|_{\mathcal{T}} + \|v_{(T,f)}\|_{\mathcal{T}} < 1 + \|v_{(T,f)}\|_{\mathcal{T}} =: C. \end{aligned}$$

Ist $h \in C^0(\mathbb{R})$, so existiert zu jedem $\tilde{\epsilon}$ ein $\delta > 0$, so dass für alle $S \in \mathcal{T}$ gilt

$$\|T - S\|_{\mathcal{T}} < \delta \Rightarrow \|h(T'(\cdot)) - h(S'(\cdot))\|_{C^0} < \tilde{\epsilon},$$

denn: Sei $I := [-\|T'\|_{C^0} - 1, \|T'\|_{C^0} + 1]$. Da h auf I gleichmäßig stetig ist, existiert ein $\delta_h > 0$ so dass gilt

$$|x - y| < \delta_h \Rightarrow |h(x) - h(y)| < \frac{\tilde{\epsilon}}{2}.$$

Für $S \in \mathcal{T}$ mit $\|S - T\|_{\mathcal{T}} < \min(\delta_h, 1)$ gilt dann offenbar für alle $z \in ([z_0, z_{\max}])$.

$$\begin{aligned} & T'(z), S'(z) \in I, \quad |S'(z) - T'(z)| < \delta_h \\ \Rightarrow & |h(S'(z)) - h(T'(z))| < \frac{\tilde{\epsilon}}{2} \\ \Rightarrow & \|h(T'(\cdot)) - h(S'(\cdot))\|_{C^0} \leq \frac{\tilde{\epsilon}}{2} < \tilde{\epsilon}. \end{aligned}$$

Ferner ist die Abbildung $L(\Gamma[\cdot]) : \mathcal{T} \rightarrow \mathbb{R}$ stetig (vgl. Beweis von 2.3.ix), also existiert wegen der Gestalt von p_T und q_T ein $\delta_2 > 0$, so dass für alle $S \in \mathcal{T}$ gilt

$$\|S - T\|_{\mathcal{T}} < \delta_2 \Rightarrow \|q_S - q_T\|_{C^0}, \|p_S - p_T\|_{C^1} < \frac{c\epsilon}{3C}.$$

Sei $\delta := \min\{\delta_1, \delta_2, \frac{c\epsilon}{3}\} > 0$. Dann gilt für $S \in \mathcal{T}, g \in C^0$ mit $\|S - T\|_{\mathcal{T}} < \delta$ und $\|f - g\|_{C^0} < \delta$

$$\begin{aligned} & \|A_T^{-1}(f, 0, 0) - A_S^{-1}(g, 0, 0)\|_{\mathcal{T}} \\ & \leq \frac{1}{c} \|v_{(S,g)}\|_{\mathcal{T}} (\|q_S - q_T\|_{C^0} + \|p_S - p_T\|_{C^1}) + \frac{1}{c} \|f - g\|_{C^0} \\ & < \frac{1}{c} C \left(\frac{c\epsilon}{3C} + \frac{c\epsilon}{3C} \right) + \frac{1}{c} \frac{c\epsilon}{3} = \epsilon. \end{aligned}$$

□

2.3.4 Stetigkeit der Metrik

Zunächst zeigen wir einen allgemeineren Satz, der uns die Stetigkeit von A_T in T liefern wird.

2.3.viii Satz. Seien $\mathcal{V} \subset X := C^2[a, b], a < b$ und $F : \mathcal{V} \rightarrow L(X, X')$ so, dass ein $f : \mathbb{R} \rightarrow \mathbb{R}$ und ein $g : \mathcal{V} \rightarrow \mathbb{R}$ existieren mit entweder

$$\forall_{x \in \mathcal{V}} \quad \forall_{v, w \in X} \quad \langle F(x)v, w \rangle = g(x) \int_a^b f(x'(z))v(z)w(z)dz$$

oder

$$\forall_{x \in \mathcal{V}} \quad \forall_{v, w \in X} \quad \langle F(x)v, w \rangle = g(x) \int_a^b f(x'(z))v'(z)w'(z)dz.$$

Sind f und g stetig, so ist auch F stetig.

Beweis Wir beweisen die Aussage nur für den Fall, dass die zweite Bedingung für F erfüllt ist. Der andere Fall folgt analog, da im Beweis nur $|v'(z)| \leq \|v\|_X$ benutzt wird und auch $|v(z)| \leq \|v\|_X$ (analog für w) gilt.

Sei $x \in \mathcal{V}$ und $\epsilon > 0$ beliebig. Dann existiert wegen der Stetigkeit von g ein $\delta_g > 0$ und ein $C_g > 0$, so dass für alle $y \in \mathcal{V}$ gilt

$$\|y - x\|_X < \delta_g \Rightarrow |g(y)| \leq C_g.$$

Analog zum Beweis von 2.9.i existiert wegen $|v(x)w(x)| \leq \|v\|_X \|w\|_X$ ein $\delta_f > 0$, so dass für alle $y \in \mathcal{V}$ mit $\|y - x\|_X < \delta_f$ und alle $v, w \in X$ mit $v, w \neq 0$ gilt

$$\left| \int_a^b f(x'(z))v'(z)w'(z)dz - \int_a^b f(y'(z))v'(z)w'(z)dz \right| < \frac{\epsilon}{2C_g} \|v\|_X \|w\|_X.$$

Ferner gilt für alle $y \in \mathcal{V}$ mit $\|y - x\|_X < \delta_f$

$$\begin{aligned} \left| \int_a^b f(y'(z))v'(z)w'(z)dz \right| &\leq \left| \int_a^b f(y'(z))v'(z)w'(z)dz - \int_a^b f(x'(z))v'(z)w'(z)dz \right| \\ &\quad + \left| \int_a^b f(x'(z))v'(z)w'(z)dz \right| \\ &\leq \underbrace{\left(\frac{\epsilon}{2C_g} + \int_a^b |f(x'(z))| dz \right)}_{=: C_f > 0} \|v\|_X \|w\|_X. \end{aligned}$$

Wegen der Stetigkeit von g existiert ein $\tilde{\delta}_g > 0$, so dass für alle $y \in \mathcal{V}$ gilt

$$\|y - x\|_X < \tilde{\delta}_g \Rightarrow |g(x) - g(y)| < \frac{\epsilon}{2C_f}.$$

Seien $\delta := \min(\delta_f, \delta_g, \tilde{\delta}_g)$, $y \in X$ mit $\|y - x\|_X < \delta_g$ und $v, w \in X$ mit $v, w \neq 0$. Dann folgt

$$\begin{aligned} |\langle (F(x) - F(y))v, w \rangle| &= \left| g(x) \int_a^b f(x'(z))v(z)w(z)dz - g(y) \int_a^b f(y'(z))v(z)w(z)dz \right| \\ &\leq \underbrace{\left| g(x) \int_a^b f(x'(z))v(z)w(z)dz - g(x) \int_a^b f(y'(z))v(z)w(z)dz \right|}_{< C_g \frac{\epsilon}{2C_g} \|v\|_X \|w\|_X} \\ &\quad + \underbrace{\left| g(x) \int_a^b f(y'(z))v(z)w(z)dz - g(y) \int_a^b f(y'(z))v(z)w(z)dz \right|}_{< \frac{\epsilon}{2C_f} C_f \|v\|_X \|w\|_X} \\ &< \epsilon \|v\|_X \|w\|_X. \end{aligned}$$

Die Stetigkeit von F folgt nun aus 5.1.iii. □

2.3.ix Satz. Die Abbildung $A[\cdot] : \mathcal{T} \rightarrow L(\mathcal{T}, \mathcal{S}), T \mapsto [v \mapsto (w \mapsto g_T(v, w))]$ ist stetig.

Beweis Für $T \in \mathcal{T}$, sowie $v, w \in \mathcal{T}$ gilt

$$\begin{aligned} |g_{T,1}(v, w)| &= \frac{1}{L(\Gamma)} \left| \int_{z_0}^{z_{\max}} v(z) \cdot w(z) \sqrt{1 + T'(z)^2} dz \right| \\ &\leq \frac{\sqrt{1 + \|T\|_{\mathcal{T}}^2}}{L(\Gamma)} \int_{z_0}^{z_{\max}} |v(z) \cdot w(z)| dz \\ &\leq \frac{\sqrt{1 + \|T\|_{\mathcal{T}}^2}}{L(\Gamma)} |z_{\max} - z_0| \|v\|_{C^0} \|w\|_{C^0} \\ |g_{T,2}(v, w)| &= L(\Gamma) \left| \int_{z_0}^{z_{\max}} \frac{v'(z)w'(z)}{\sqrt{1 + T'(z)^2}} dz \right| \leq L(\Gamma) \int_{z_0}^{z_{\max}} |v'(z)w'(z)| dz \\ &\leq L(\Gamma) |z_{\max} - z_0| \|v'\|_{C^0} \|w'\|_{C^0} \end{aligned}$$

und wegen $\|v\|_{C^0} + \|v'\|_{C^0} \leq \|v\|_{\mathcal{T}}$, sowie $g_T(v, w) = g_{T,1}(v, w) + g_{T,2}(v, w)$, folgt

$$|g_T(v, w)| \leq C_T \|v\|_{\mathcal{T}} \|w\|_{\mathcal{T}} \quad (2.12)$$

wobei $C_T > 0$ nur von T abhängt. Für $T \in \mathcal{T}$ gilt also

$$\|A[T]\|_{L(\mathcal{T}, \mathcal{T}')} = \sup_{0 \neq v \in \mathcal{T}} \frac{\|A[T]v\|_{\mathcal{T}'}}{\|v\|_{\mathcal{T}}} = \sup_{0 \neq v \in \mathcal{T}} \sup_{0 \neq w \in \mathcal{T}} \frac{|\langle A[T]v, w \rangle|}{\|v\|_{\mathcal{T}} \|w\|_{\mathcal{T}}} \leq C,$$

wegen

$$|\langle A[T]v, w \rangle| = |g_T(v, w)| \leq C \|v\|_{\mathcal{T}} \|w\|_{\mathcal{T}}.$$

Somit ist $A[T] \in L(\mathcal{T}, \mathcal{T}')$ und da wir bereits gezeigt haben, dass $A[T] : \mathcal{T} \rightarrow \mathcal{S}$ und $\mathcal{S} \subset \mathcal{T}'$, gilt $A[T] \in L(\mathcal{T}, \mathcal{S})$, d.h. $A[\cdot]$ ist wohldefiniert. Die Abbildung

$$\mathcal{T} \rightarrow \mathbb{R}, T \mapsto \int_{z_0}^{z_{\max}} \sqrt{1 + T'(z)^2} dt (= L(\Gamma[T]))$$

ist bekanntlich stetig, also ist $L(\Gamma[\cdot]) : \mathcal{T} \rightarrow \mathbb{R}$ stetig. Die Stetigkeit von $A[\cdot]$ folgt nun durch Anwenden von 2.3.viii auf beide Teile von $A[\cdot]$, wobei die Voraussetzungen wegen 2.3.ii und der Stetigkeit von $L(\Gamma[\cdot])$ erfüllt sind. \square

2.3.x Satz. Seien X, Y, Z normierte Vektorräume und $x \in X$. Ferner sei $A[\cdot] : X \rightarrow L(Y, Z)$, $x \mapsto A[x]$ stetig in X und es existiere ein $\delta > 0$ und ein $C > 0$, so dass für alle $y \in X$ mit $\|x - y\|_X < \delta$ gilt

$$A[y] \text{ ist invertierbar und } \|A^{-1}[y]\|_{L(Z, Y)} < C.$$

Dann ist $A^{-1}[\cdot] : X \rightarrow L(Z, Y)$, $x \mapsto A^{-1}[x]$ stetig in x .

Beweis Sei $\epsilon > 0$. Nach Voraussetzung existieren ein $\delta_1 > 0$ und ein $C > 0$, so dass für alle $y \in X$ gilt

$$\|x - y\|_X < \delta_1 \Rightarrow \|A^{-1}[y]\|_{L(Z, Y)} < C.$$

Ferner existiert wegen der Stetigkeit von $A[\cdot]$ in x ein $\delta_2 > 0$, so dass für alle $y \in X$ gilt

$$\|x - y\|_X < \delta_2 \Rightarrow \|A[x] - A[y]\|_{L(Y, Z)} < \frac{\epsilon}{C^2}.$$

Sei nun $\delta := \min(\delta_1, \delta_2)$ und $y \in X$ mit $\|x - y\|_X < \delta$. Dann gilt

$$\begin{aligned} \|A^{-1}[x] - A^{-1}[y]\|_{L(Z, Y)} &= \|A^{-1}[x] (\mathbb{1} - A[x]A^{-1}[y])\|_{L(Z, Y)} \\ &= \|A^{-1}[x] (A[y] - A[x]) A^{-1}[y]\|_{L(Z, Y)} \\ &\leq \|A^{-1}[x]\|_{L(Z, Y)} \|A[y] - A[x]\|_{L(Y, Z)} \|A^{-1}[y]\|_{L(Z, Y)} \\ &< C \frac{\epsilon}{C^2} C = \epsilon. \end{aligned}$$

\square

2.4 Variation der Energie

Wie wir nun gesehen haben, ist es für die Anwendung des Gradientenflusses nötig, die Variation $E'[T]$ der Energie auszurechnen. Sei

$$J_{ik}[T] = \frac{1}{2} |T_b(\nu_i, \theta_k, z_0)[T] - T_{bmik}|^2,$$

dann gilt

$$E[T] = \sum_{i=1}^n \sum_{k=1}^m J_{ik}[T].$$

Die Variation von T_b ist eine Funktion, abhängig von z :

$$\langle T_b'(\nu_i, \theta_k, z)[T], \vartheta \rangle = \left. \frac{d}{d\epsilon} T_b(\nu_i, \theta_k, z)[T + \epsilon\vartheta] \right|_{\epsilon=0}.$$

Es gilt dann nach der Kettenregel und der Definition der ersten Variation

$$\langle J'_{ik}[T], \vartheta \rangle = (T_b(\nu_i, \theta_k, z_0)[T] - T_{bmik}) \langle T_b'(\nu_i, \theta_k, z_0)[T], \vartheta \rangle.$$

Wegen der Linearität der ersten Variation gilt

$$\begin{aligned} \langle E'[T], \vartheta \rangle &= \sum_{i=1}^n \sum_{k=1}^m \langle J'_{ik}[T], \vartheta \rangle \\ &= \sum_{i=1}^n \sum_{k=1}^m (T_b(\nu_i, \theta_k, z_0)[T] - T_{bik}) \langle T_b'(\nu_i, \theta_k, z_0)[T], \vartheta \rangle \end{aligned}$$

und unser Problem reduziert sich auf das Auswerten von

$$\langle T_b'(\nu_i, \theta_k, z_0)[T], \vartheta \rangle.$$

Um die Notation übersichtlicher zu machen, seien $i \in \{1, \dots, n\}$ und $k \in \{1, \dots, m\}$ fest. Für die Frequenz ν_i und den Winkel θ_k benutzen wir im Folgenden für die Gleichung (2.2) die Kurzschreibweise

$$\begin{aligned} \frac{d}{dz} T_b &= G(T_b, T) \\ T_b(z_{\max}) &= T_{b0}. \end{aligned}$$

und $T_b(z)[T] := T_b(z)$, um die Abhängigkeit von T hervorzuheben. Ferner benutzen wir

$$\langle T_b'[T], \vartheta \rangle$$

als Kurzschreibweise für die von z abhängige Funktion. Bei der Auswertung von T_b' hilft uns das folgende Lemma:

2.4.i Lemma. *Die Variation der Strahlungstemperaturen erfüllt die Differentialgleichung*

$$\begin{aligned} \frac{d}{dz} \langle T_b'[T], \vartheta \rangle &= \frac{\partial}{\partial T_b} G(T_b, T) \cdot \langle T_b'[T], \vartheta \rangle + \frac{\partial}{\partial T} G(T_b, T) \cdot \vartheta, \\ \langle T_b'(z_{\max})[T], \vartheta \rangle &= 0. \end{aligned} \tag{2.13}$$

Beweis Aus (2.2) folgt

$$\begin{aligned}
\underbrace{\frac{d}{dz} \langle T_b'[T], \vartheta \rangle}_{\Delta'} &= \frac{d}{dz} \left(\frac{d}{d\epsilon} T_b[T + \epsilon\vartheta] \Big|_{\epsilon=0} \right) = \frac{d}{d\epsilon} \left(\frac{d}{dz} T_b[T + \epsilon\vartheta] \right) \Big|_{\epsilon=0} \\
&= \frac{d}{d\epsilon} (G(T_b[T + \epsilon\vartheta], T + \epsilon\vartheta)) \Big|_{\epsilon=0} \\
&= \frac{\partial}{\partial T_b} G(T_b, T) \cdot \underbrace{\langle T_b'[T], \vartheta \rangle}_{\Delta} + \frac{\partial}{\partial T} G(T_b, T) \cdot \vartheta, \\
\langle T_b'(z_{\max})[T], \vartheta \rangle &= \frac{d}{d\epsilon} T_b(z_{\max})[T + \epsilon\vartheta] \Big|_{\epsilon=0} = \frac{d}{d\epsilon} T_{b0} \Big|_{\epsilon=0} = 0
\end{aligned} \tag{2.14}$$

□

Nun schreiben wir die Abhängigkeit von ν_i und θ_k wieder aus. Aus der eindeutigen Lösbarkeit von Gleichungen obigen Typs folgt ferner: Ist Δ eine Lösung von

$$\begin{aligned}
\frac{d}{dz} \Delta(\nu_i, \theta_k, z) &= (\partial_T G(T_b(z, \nu_i, \theta_k), \nu_i, \theta_k, T(z), Q(z))) \vartheta(z) + \\
&\quad (\partial_{T_b} G(T_b(z, \nu_i, \theta_k), \nu_i, \theta_k, T(z), Q(z))) \Delta(z), \\
\Delta(\nu_i, \theta_k, z_{\max}) &= 0,
\end{aligned} \tag{2.15}$$

so gilt

$$\langle T_b'(\nu_i, \theta_k, z_0)[T], \vartheta \rangle = \Delta(\nu_i, \theta_k, z_0),$$

und das Auswerten der Variation der Strahlungstemperaturen ist auf das Lösen obiger Differentialgleichung zurückgeführt. Im Folgenden sei immer

$$\langle T_b'(\nu_i, \theta_k)[T], \vartheta \rangle := \langle T_b'(\nu_i, \theta_k, z_0)[T], \vartheta \rangle.$$

Aus (2.3) folgt

$$\begin{aligned}
\partial_T G(T_b, \nu, \theta, T, Q) &= \sec(\theta) (\partial_T \sigma_a(\nu, T, Q) (T_b - T) - \sigma_a(\nu, T, Q)), \\
\partial_{T_b} G(T_b, \nu, \theta, T, Q) &= \sec(\theta) \sigma_a(\nu, T, Q).
\end{aligned}$$

Mit Satz 5.1.i erhält man nun die Integralform für die Variation der Strahlungstemperaturen:

$$\begin{aligned}
\langle T_b'(\nu, \theta, z)[T], \vartheta \rangle &= \\
&= -\sec(\theta) \int_z^{z_{\max}} \left(\partial_T \sigma_a(\nu, T(t), Q(t)) (T_b(\nu, \theta, t)[T] - T(t)) - \sigma_a(\nu, T(t), Q(t)) \right) \vartheta(t) \\
&\quad \exp \left(-\sec(\theta) \int_z^t \sigma_a(\nu, T(s), Q(s)) ds \right) dt.
\end{aligned} \tag{2.16}$$

2.5 Adjoint variable method

Bei inversen Problemen wird häufig die adjoint variable method angewendet. Informationen zum Hintergrund und Anwendungen dieser Methode gibt es in [25, 21]. Daher prüfen wir,

ob es sinnvoll ist, dieses Verfahren auf unser Problem anzuwenden. Seien dazu $i \in \{1, \dots, n\}$ und $k \in \{1, \dots, m\}$ fest. Für die Frequenz ν_i und den Winkel θ_k benutzen wir wieder die Kurzschreibweise aus 2.4.

Die zu (2.13) gehörige adjungierte Gleichung lautet

$$S'_{ik} = -\frac{\partial}{\partial T_b} G(T_b, T) S_{ik} + (T_b(z_0) - T_{bmik}) \delta_{z_0},$$

$$S_{ik}(z_0) = 0,$$

wobei δ_{z_0} die Deltadistribution ist. Ist S_{ik} eine Lösung dieser Gleichung, so folgt durch Umstellen, Multiplikation mit $\langle T_b'[T], \vartheta \rangle$ und Integration über $[z_0, z_{\max}]$

$$\begin{aligned} (T_b(z_0)[T] - T_{bmik}) \langle T_b'(z_0)[T], \vartheta \rangle &= \left\langle S'_{ik} + \frac{\partial}{\partial T_b} G(T_b, T) S_{ik}, \langle T_b'[T], \vartheta \rangle \right\rangle_{L_2[z_0, z_{\max}]} \\ &= \langle S'_{ik}, \langle T_b'[T], \vartheta \rangle \rangle_{L_2[z_0, z_{\max}]} + \left\langle \frac{\partial}{\partial T_b} G(T_b, T) S_{ik}, \langle T_b'[T], \vartheta \rangle \right\rangle_{L_2[z_0, z_{\max}]} \\ &\stackrel{\text{part. Int}}{=}_{S_{ik}(z_0)=0} - \left\langle S_{ik}, \frac{d}{dz} \langle T_b'[T], \vartheta \rangle \right\rangle_{L_2[z_0, z_{\max}]} + \left\langle S_{ik}, \frac{\partial}{\partial T_b} G(T_b, T) \langle T_b'[T], \vartheta \rangle \right\rangle_{L_2[z_0, z_{\max}]} \\ &\stackrel{(2.13)}{=} - \left\langle S_{ik}, \frac{\partial}{\partial T} G(T_b, T) \cdot \vartheta \right\rangle_{L_2[z_0, z_{\max}]} . \end{aligned}$$

Es gilt also

$$\langle E'[T], \vartheta \rangle = \sum_{i=1}^n \sum_{k=1}^m - \left\langle S_{ik}, \frac{\partial}{\partial T} G(T_b, T) \cdot \vartheta \right\rangle_{L_2[z_0, z_{\max}]} .$$

Wollen wir zu Testfunktionen $\vartheta_1, \dots, \vartheta_l$ jeweils die Variation der Energie ausrechnen, so müssen wir $m \cdot n \cdot l$ Differentialgleichungen lösen, wenn wir die Berechnung wie in 2.4 durchführen. Da S_{ik} nicht von der Testfunktion abhängig ist, müssen nur $m \cdot n$ Differentialgleichungen gelöst werden. Allerdings benötigt man noch $m \cdot n \cdot l$ L_2 -Skalarprodukte, also Integrationen, die bei entsprechender Genauigkeit genauso teuer sind wie das Lösen einer Differentialgleichung. Daher kann bei unserem Problem der Rechenaufwand durch die adjoint variable method nicht verringert werden und wir benutzen zur Berechnung der Variation der Energie die Methode aus 2.4.

2.6 Stetigkeit der Variation der Energie

Seien ν und θ fest. Wir benutzen wieder die Kurzschreibweise aus Abschnitt 2.4. Ferner sei

$$\sigma : \mathbb{R}^2 \rightarrow \mathbb{R}, (z, T) \mapsto \sec(\theta) \sigma_a(\nu, T, Q(z)).$$

Wegen der Voraussetzung an die stetige Differenzierbarkeit von σ_a , $\partial_T \sigma_a$ und Q gilt $\sigma, \partial_T \sigma \in C^1([z_0, z_{\max}] \times \mathbb{R})$. Wir setzen weiter für $z, t \in [z_0, z_{\max}]$ und $T \in \mathcal{T}$

$$f(t, z)[T] = \exp \left(- \int_z^t \sigma(s, T(s)) \right) ds.$$

2.6.i Hilfssatz.

$$\begin{aligned} \forall_{v \in \mathcal{T}} \quad \forall_{\epsilon > 0} \quad \exists_{\delta > 0} \quad \exists_{C > 0} \quad \forall_{\substack{w \in \mathcal{T} \\ \|v-w\|_{\mathcal{T}} < \delta}} \quad \forall_{z, t \in [z_0, z_{\max}]} \quad & |f(t, z)[v] - f(t, z)[w]| < \epsilon, \\ & |f(t, z)[w]| < C. \end{aligned}$$

Beweis Sei $v \in \mathcal{T}$ und $\epsilon > 0$. Für $z, t \in [z_0, z_{\max}]$ gilt

$$\left| \int_z^t \sigma(s, v(s)) ds \right| \leq \left| \int_z^t |\sigma(s, v(s))| ds \right| \leq \int_{z_0}^{z_{\max}} |\sigma(s, v(s))| ds := M.$$

Da \exp stetig und somit auf $[-M-1, M+1]$ gleichmäßig stetig, existiert ein $\delta_1 > 0$, so dass für alle $a, b \in [-M-1, M+1]$ gilt

$$|a - b| < \delta_1 \Rightarrow |\exp(a) - \exp(b)| < \epsilon.$$

Ferner gilt o.E. $\delta_1 \leq 1$. Sei

$$I(v) := \left[\min_{z \in [z_0, z_{\max}]} v(z) - 1, \max_{z \in [z_0, z_{\max}]} v(z) + 1 \right].$$

Da σ stetig, ist σ auf $[z_0, z_{\max}] \times I(v)$ gleichmäßig stetig und es existiert ein $\delta_2 > 0$, so dass für alle $s \in [z_0, z_{\max}]$, $x, y \in I(v)$ gilt

$$|x - y| < \delta_2 \Rightarrow |\sigma(s, x) - \sigma(s, y)| < \frac{\delta_1}{z_{\max} - z_0}.$$

Sei $w \in \mathcal{T}$ mit $\|v - w\|_{\mathcal{T}} < \delta := \min(\delta_2, 1)$. Dann gilt für alle $s \in [z_0, z_{\max}]$

$$\min_{z \in [z_0, z_{\max}]} v(z) - 1 \leq v(s) - 1 < w(s) < v(s) + 1 \leq \max_{z \in [z_0, z_{\max}]} v(z) + 1,$$

d.h. $w(s) \in I(v)$ für alle $s \in [z_0, z_{\max}]$ ($v(s) \in I(v)$ gilt offensichtlich). Wegen $\|v - w\|_{\mathcal{T}} < \delta$ gilt auch $|v(s) - w(s)| < \delta_2$ für alle $s \in [z_0, z_{\max}]$. Es folgt für $z, t \in [z_0, z_{\max}]$

$$\left| \int_z^t \sigma(s, w(s)) ds - \int_z^t \sigma(s, v(s)) ds \right| \leq \left| \int_z^t \underbrace{|\sigma(s, w(s)) - \sigma(s, v(s))|}_{< \frac{\delta_1}{z_{\max} - z_0}} ds \right| < \frac{\delta_1 |t - z|}{z_{\max} - z_0} \leq \delta_1$$

und

$$\left| \int_z^t \sigma(s, w(s)) ds \right| \leq \underbrace{\left| \int_z^t \sigma(s, w(s)) ds - \int_z^t \sigma(s, v(s)) ds \right|}_{< \delta_1 \leq 1} + \underbrace{\left| \int_z^t \sigma(s, v(s)) ds \right|}_{\leq M} < M + 1.$$

Seien

$$a := - \int_z^t \sigma(s, v(s)) ds, b := - \int_z^t \sigma(s, w(s)) ds,$$

wir haben gezeigt $a, b \in [-M-1, M+1]$ und $|a - b| < \delta_1$, also folgt

$$|f(t, z)[v] - f(t, z)[w]| = |\exp(a) - \exp(b)| < \epsilon.$$

Ferner gilt, da \exp wachsend,

$$|f(t, z)[w]| \leq \exp\left(\left|\int_z^t \sigma(s, w(s))ds\right|\right) < \exp(M + 1) =: C.$$

□

2.6.ii Satz. Für d gelte die Aussage aus 2.6.i und sei $g \in C^0([z_0, z_{\max}] \times \mathbb{R})$. Ferner sei für $z \in [z_0, z_{\max}]$, $v \in \mathcal{T}$ und $\vartheta \in \mathcal{T}$

$$\langle h(z)[v], \vartheta \rangle = \int_z^{z_{\max}} \underbrace{g(t, v(t))d(t, z)[v]}_{=: D(t, z)[v]} \vartheta(t) dt.$$

Dann gilt

$$\begin{aligned} \forall_{v \in \mathcal{T}} \quad \forall_{\epsilon > 0} \quad \exists_{\substack{\delta > 0 \\ C > 0}} \quad \forall_{\substack{w \in \mathcal{T} \\ \|v - w\|_{\mathcal{T}} < \delta}} \quad \forall_{t, z \in [z_0, z_{\max}]} \quad \forall_{\substack{\vartheta \in \mathcal{T} \\ \vartheta \neq 0}} \quad & |\langle h(z)[v], \vartheta \rangle - \langle h(z)[w], \vartheta \rangle| < \epsilon \|\vartheta\|_{\mathcal{T}} \\ \wedge |\langle h(z)[w], \vartheta \rangle| < C \|\vartheta\|_{\mathcal{T}} \wedge |D(t, z)[v] - D(t, z)[w]| < \epsilon \wedge |D(t, z)[w]| < C. \end{aligned}$$

Beweis Sei $v \in \mathcal{T}$ und $\epsilon > 0$. Sei $M := \|g\|_{\infty, [z_0, z_{\max}] \times I(v)} + 1$. Wegen 2.6.i existiert $\delta_2 > 0$ und ein $C > 0$, so dass für alle $w \in \mathcal{T}$ und alle $z, t \in [z_0, z_{\max}]$ gilt

$$\begin{aligned} \|v - w\|_{\mathcal{T}} < \delta_2 &\Rightarrow |d(t, z)[v] - d(t, z)[w]| < \frac{\epsilon}{2M(z_{\max} - z_0)}, \\ |d(t, z)[w]| &< C. \end{aligned}$$

g ist in $[z_0, z_{\max}] \times I(v)$ gleichmäßig stetig, also existiert ein $\delta_1 > 0$, so dass für alle $t \in [z_0, z_{\max}]$, $x, y \in I(v)$ gilt

$$|x - y| < \delta_2 \Rightarrow |g(t, x) - g(t, y)| < \frac{\epsilon}{2C(z_{\max} - z_0)}.$$

Sei $w \in \mathcal{T}$ mit $\|v - w\|_{\mathcal{T}} < \delta := \min(\delta_1, \delta_2, 1)$ und $\vartheta \in \mathcal{T}$. Wie im Beweis von 2.6.i folgt

$$\forall_{s \in [z_0, z_{\max}]} \quad v(t), w(t) \in I(v) \wedge |v(t) - w(t)| < \delta_1, \delta_2.$$

Es folgt

$$\begin{aligned} |\langle h(z)[v], \vartheta \rangle - \langle h(z)[w], \vartheta \rangle| &\leq \int_z^{z_{\max}} |g(t, v(t))d(t, z)[v] - g(t, w(t))d(t, z)[w]| \underbrace{|\vartheta(t)|}_{\leq \|\vartheta\|_{\mathcal{T}}} dt \\ &\leq \|\vartheta\|_{\mathcal{T}} \int_z^{z_{\max}} \underbrace{|g(t, v(t))d(t, z)[v] - g(t, v(t))d(t, z)[w]|}_{=|g(t, v(t)) \cdot |d(t, z)[v] - d(t, z)[w]| < M \frac{\epsilon}{2M(z_{\max} - z_0)}} dt \\ &\quad + \|\vartheta\|_{\mathcal{T}} \int_z^{z_{\max}} \underbrace{|g(t, v(t))d(t, z)[w] - g(t, w(t))d(t, z)[w]|}_{=|g(t, v(t)) - g(t, w(t))| \cdot |d(t, z)[w]| < \frac{\epsilon}{2C(z_{\max} - z_0)} C} dt \\ &< \|\vartheta\|_{\mathcal{T}} \epsilon \frac{z_{\max} - z}{z_{\max} - z_0} \leq \epsilon \|\vartheta\|_{\mathcal{T}} \end{aligned}$$

und

$$\begin{aligned} |\langle h(z)[w], \vartheta \rangle| &\leq \int_z^{z_{\max}} \underbrace{|g(t, w(t))|}_{<M} \underbrace{|d(t, z)[w]|}_{<C} \underbrace{|\vartheta(t)|}_{\leq \|\vartheta\|_{\mathcal{T}}} dt < MC(z_{\max} - z) \|\vartheta\|_{\mathcal{T}} \\ &\leq MC(z_{\max} - z_0) \|\vartheta\|_{\mathcal{T}} = \tilde{C} \|\vartheta\|_{\mathcal{T}}. \end{aligned}$$

Analog zu den Abschätzungen für den Integranden, zeigt man die beiden Ungleichungen für D . \square

2.6.iii Satz. *Die Strahlungstemperaturen sind im folgenden Sinne stetig von der Temperatur abhängig:*

$$\begin{aligned} \forall v \in \mathcal{T} \quad \forall \epsilon > 0 \quad \exists \delta > 0 \quad \exists C > 0 \quad \forall \substack{w \in \mathcal{T} \\ \|v-w\|_{\mathcal{T}} < \delta} \quad \forall z \in [z_0, z_{\max}] \quad & |T_b(z)[v] - T_b(z)[w]| < \epsilon, \\ & |T_b(z)[w]| < C. \end{aligned}$$

Beweis Wegen (2.1) gilt

$$T_b(z)[v] = \underbrace{T_{b0} f(z_{\max}, z)[v]}_{=:I} + \underbrace{\int_z^{z_{\max}} \sigma(t, v(t))v(t)f(t, z)[v]dt}_{=:II}.$$

Mit $g(t, s) := \sigma(t, s)s$ gilt $g \in C^0([z_0, z_{\max}] \times \mathbb{R})$ und die Behauptung folgt durch Anwenden von 2.6.i auf I und durch Anwenden von 2.6.ii mit $\vartheta \equiv 1$ (Voraussetzungen sind unter anderem wegen 2.6.i erfüllt) auf II. \square

2.6.iv Satz. *Die Variation der Strahlungstemperaturen ist im folgenden Sinne stetig von der Temperatur abhängig:*

$$\begin{aligned} \forall v \in \mathcal{T} \quad \forall \epsilon > 0 \quad \exists \substack{\delta > 0 \\ C > 0} \quad \forall \substack{w \in \mathcal{T} \\ \|v-w\|_{\mathcal{T}} < \delta} \quad \forall z \in [z_0, z_{\max}] \quad \forall \substack{\vartheta \in \mathcal{T} \\ \vartheta \neq 0} \quad & |\langle T_b'(z)[v], \vartheta \rangle - \langle T_b'(z)[w], \vartheta \rangle| < \epsilon \|\vartheta\|_{\mathcal{T}}, \\ & |\langle T_b'(z)[w], \vartheta \rangle| < C \|\vartheta\|_{\mathcal{T}}. \end{aligned}$$

Ferner existiert ein $F : [z_0, z_{\max}]^2 \times \mathcal{T}, (t, z, v) \mapsto F(t, z)[v]$, so dass für alle $v \in \mathcal{T}$ gilt $F(\cdot, \cdot)[v] \in C^1([z_0, z_{\max}]^2)$ und

$$\forall \vartheta \in \mathcal{T} \quad \forall z \in [z_0, z_{\max}] \quad \langle T_b'(z)[v], \vartheta \rangle = \int_z^{z_{\max}} F(t, z)[v]\vartheta(t)dt.$$

Weiter gilt

$$\begin{aligned} \forall v \in \mathcal{T} \quad \forall \epsilon > 0 \quad \exists \substack{\delta > 0 \\ C > 0} \quad \forall \substack{w \in \mathcal{T} \\ \|v-w\|_{\mathcal{T}} < \delta} \quad \forall t, z \in [z_0, z_{\max}] \quad & |F(t, z)[v] - F(t, z)[w]| < \epsilon, \\ & |F(t, z)[w]| < C. \end{aligned}$$

also gilt für d die Aussage aus 2.6.i und somit folgt aus 2.6.ii, dass ein $\delta_I > 0$ und ein $C_I > 0$ existieren, so dass gilt

$$\begin{aligned} \forall_{\substack{w \in \mathcal{T} \\ \|v-w\|_{\mathcal{T}} < \delta_I}} \quad \forall_{z \in [z_0, z_{\max}]} \quad \forall_{\substack{\vartheta \in \mathcal{T} \\ \vartheta \neq 0}} \quad & |I(v, z, \vartheta) - I(w, z, \vartheta)| < \frac{\epsilon}{2} \|\vartheta\|_{\mathcal{T}}, \\ & |I(w, z, \vartheta)| < C_I \|\vartheta\|_{\mathcal{T}}. \end{aligned}$$

Mit $g_{II}(t, s) := \partial_T \sigma(t, s)s + \sigma(t, s)$ gilt $g_{II} \in C^0([z_0, z_{\max}] \times \mathbb{R})$ und somit folgt aus 2.6.ii (Voraussetzungen sind unter anderem wegen 2.6.i erfüllt), dass ein $\delta_{II} > 0$ und ein $C_{II} > 0$ existieren, so dass gilt

$$\begin{aligned} \forall_{\substack{w \in \mathcal{T} \\ \|v-w\|_{\mathcal{T}} < \delta_{II}}} \quad \forall_{z \in [z_0, z_{\max}]} \quad \forall_{\substack{\vartheta \in \mathcal{T} \\ \vartheta \neq 0}} \quad & |\mathbb{II}(v, z, \vartheta) - \mathbb{II}(w, z, \vartheta)| < \frac{\epsilon}{2} \|\vartheta\|_{\mathcal{T}}, \\ & |\mathbb{II}(w, z, \vartheta)| < C_{II} \|\vartheta\|_{\mathcal{T}}. \end{aligned}$$

Mit $\delta := \min(\delta_I, \delta_{II})$ und $C := C_I + C_{II}$ folgt wegen

$$|\langle T_b'(z)[v], \vartheta \rangle - \langle T_b'(z)[w], \vartheta \rangle| \leq |I(v, z, \vartheta) - I(w, z, \vartheta)| + |\mathbb{II}(v, z, \vartheta) - \mathbb{II}(w, z, \vartheta)|$$

und

$$|\langle T_b'(z)[w], \vartheta \rangle| \leq |I(w, z, \vartheta)| + |\mathbb{II}(w, z, \vartheta)|$$

der erste Teil der Behauptung. Für $v \in \mathcal{T}$ gilt offenbar $F(\cdot, \cdot)[v] \in C^1([z_0, z_{\max}]^2)$, da F aus stetig differenzierbaren Funktionen zusammen gesetzt ist. Analog zu den Abschätzungen des ersten Teils der Behauptung folgen die Abschätzungen von F , wenn man die Abschätzungen von D aus 2.6.ii anwendet, um die nötigen Abschätzungen für F_1 und F_2 zu erhalten (so wie man die Abschätzungen von h aus 2.6.ii für I und II verwendet hat). \square

Nun haben wir genug Vorbereitungen getroffen, um die Stetigkeit der Variation der Energie zu beweisen:

2.6.v Satz.

$$E'[\cdot] : \mathcal{T} \rightarrow \mathcal{S}, T \mapsto E'[T]$$

ist stetig.

Beweis Hier zeigen wir zunächst nur, dass $E'[\cdot] : \mathcal{T} \rightarrow \mathcal{T}'$ stetig ist. Die Behauptung folgt daraus mit $R(E'[\cdot]) \subset \mathcal{S}$, was wir im Beweis von 2.7.i zeigen werden.

Seien $v \in \mathcal{T}$ und $\epsilon > 0$. Ferner seien $i \in \{1, \dots, n\}$ und $k \in \{1, \dots, m\}$ fest. Seien $M := mn > 0$ und $N := |T_{b_{ik}}| + 1 > 0$. Wir benutzen wieder die Kurzschreibweise und lassen ν_i und θ_k weg. Nach 2.6.iii existiert ein $\delta_1 > 0$ und ein $C_1 > 0$ mit

$$\forall_{\substack{w \in \mathcal{T} \\ \|v-w\|_{\mathcal{T}} < \delta_1}} \quad |T_b(z_0)[w]| < C_1.$$

Nach 2.6.iv existiert ein $\delta_2 > 0$ und ein $C_2 > 0$ mit

$$\begin{aligned} \forall_{\substack{w \in \mathcal{T} \\ \|v-w\|_{\mathcal{T}} < \delta_2}} \quad \forall_{\substack{\vartheta \in \mathcal{T} \\ \vartheta \neq 0}} \quad & |\langle T_b'(z_0)[v], \vartheta \rangle - \langle T_b'(z_0)[w], \vartheta \rangle| < \min\left(\frac{\epsilon}{3C_1M}, \frac{\epsilon}{3NM}\right) \|\vartheta\|_{\mathcal{T}}, \\ & |\langle T_b'(z_0)[w], \vartheta \rangle| < C_2 \|\vartheta\|_{\mathcal{T}}. \end{aligned}$$

Nach 2.6.iii existiert ein $\delta_3 > 0$ mit

$$\forall_{\substack{w \in \mathcal{T} \\ \|v-w\|_{\mathcal{T}} < \delta_3}} |T_b(z_0)[v] - T_b(z_0)[w]| < \frac{\epsilon}{3C_2M}.$$

Sei $\delta_{ik} := \min(\delta_1, \delta_2, \delta_3)$. Seien $w \in \mathcal{T}$ mit $\|v-w\|_{\mathcal{T}} < \delta_{ik}$ und $\vartheta \in \mathcal{T}$ mit $\vartheta \neq 0$, dann gilt

$$\begin{aligned} & |\langle J'_{ik}[v], \vartheta \rangle - \langle J'_{ik}[w], \vartheta \rangle| \\ &= |(T_b(z_0)[v] - T_{b_{ik}}) \langle T_b'(z_0)[v], \vartheta \rangle - (T_b(z_0)[w] - T_{b_{ik}}) \langle T_b'(z_0)[w], \vartheta \rangle| \\ &\leq |T_b(z_0)[v] \langle T_b'(z_0)[v], \vartheta \rangle - T_b(z_0)[w] \langle T_b'(z_0)[w], \vartheta \rangle| \\ &\quad + \underbrace{|T_{b_{ik}} \langle T_b'(z_0)[v], \vartheta \rangle - T_{b_{ik}} \langle T_b'(z_0)[w], \vartheta \rangle|}_{< N \frac{\epsilon}{3NM} \|\vartheta\|_{\mathcal{T}}} \\ &< \underbrace{|T_b(z_0)[v] \langle T_b'(z_0)[v], \vartheta \rangle - T_b(z_0)[v] \langle T_b'(z_0)[w], \vartheta \rangle|}_{< C_1 \frac{\epsilon}{3C_1M} \|\vartheta\|_{\mathcal{T}}} \\ &\quad + \underbrace{|T_b(z_0)[v] \langle T_b'(z_0)[w], \vartheta \rangle - T_b(z_0)[w] \langle T_b'(z_0)[w], \vartheta \rangle|}_{< \frac{\epsilon}{3M} \|\vartheta\|_{\mathcal{T}}} \\ &\quad < \frac{\epsilon}{3C_2M} C_2 \|\vartheta\|_{\mathcal{T}} \\ &< \frac{\epsilon}{M} \|\vartheta\|_{\mathcal{T}}. \end{aligned}$$

Analog findet man zu jedem $i \in \{1, \dots, n\}$ und $k \in \{1, \dots, m\}$ ein δ_{ik} und setzt $\delta := \min\{\delta_{ik} | i \in \{1, \dots, n\}, k \in \{1, \dots, m\}\}$. Seien nun $w \in \mathcal{T}$ mit $\|v-w\|_{\mathcal{T}} < \delta$ und $\vartheta \in \mathcal{T}$ mit $\vartheta \neq 0$, dann gilt

$$\begin{aligned} |\langle E'[v], \vartheta \rangle - \langle E'[w], \vartheta \rangle| &\leq \sum_{i=1}^n \sum_{k=1}^m |\langle J'_{ik}[v], \vartheta \rangle - \langle J'_{ik}[w], \vartheta \rangle| \\ &< \sum_{i=1}^n \sum_{k=1}^m \frac{\epsilon}{M} \|\vartheta\|_{\mathcal{T}} = \epsilon \|\vartheta\|_{\mathcal{T}}. \end{aligned}$$

Die Stetigkeit von $E'[\cdot]$ folgt nun mit 5.1.iv. □

2.6.vi Satz. *Es existiert ein $H : [z_0, z_{\max}] \times \mathcal{T}, (t, v) \mapsto H(t)[v]$, so dass für alle $v \in \mathcal{T}$ gilt $H(\cdot)[v] \in C^1([z_0, z_{\max}])$ und*

$$\forall_{\vartheta \in \mathcal{T}} \langle E'[v], \vartheta \rangle = \int_z^{z_{\max}} H(t)[v] \vartheta(t) dt.$$

Ferner gilt

$$\forall_{v \in \mathcal{T}} \quad \forall_{\epsilon > 0} \quad \exists_{\delta > 0} \quad \forall_{\substack{w \in \mathcal{T} \\ \|v-w\|_{\mathcal{T}} < \delta}} \quad \forall_{t \in [z_0, z_{\max}]} |H(t)[v] - H(t)[w]| < \epsilon.$$

Beweis 2.6.iv gilt für jede Frequenz ν_i und für jeden Winkel θ_k , daher existiert zu allen $i \in \{1, \dots, n\}$ und $k \in \{1, \dots, m\}$ ein F_{ik} , so dass für beliebiges $v \in \mathcal{T}$ gilt

$$\forall_{\vartheta \in \mathcal{T}} \langle T_b'(\nu_i, \theta_k, z_0)[v], \vartheta \rangle = \int_z^{z_{\max}} F_{ik}(t, z_0)[v] \vartheta(t) dt.$$

Für $v, \vartheta \in \mathcal{T}$ folgt daraus

$$\begin{aligned} \langle E'[v], \vartheta \rangle &= \sum_{i=1}^n \sum_{k=1}^m (T_b(\nu_i, \theta_k, z_0)[v] - T_{b_{ik}}) \langle T_b'(\nu_i, \theta_k, z_0)[v], \vartheta \rangle \\ &= \sum_{i=1}^n \sum_{k=1}^m (T_b(\nu_i, \theta_k, z_0)[v] - T_{b_{ik}}) \int_z^{z_{\max}} F_{ik}(t, z_0)[v] \vartheta(t) dt \\ &= \int_z^{z_{\max}} \underbrace{\sum_{i=1}^n \sum_{k=1}^m (T_b(\nu_i, \theta_k, z_0)[v] - T_{b_{ik}}) F_{ik}(t, z_0)[v]}_{=: H(t)[v]} \vartheta(t) dt. \end{aligned}$$

Da $F_{ik}(\cdot, \cdot)[v] \in C^1([z_0, z_{\max}]^2)$ (vgl. 2.6.iv), gilt offenbar $H(\cdot)[v] \in C^1([z_0, z_{\max}])$ und die behauptete Abschätzung kann man analog zum Beweis von 2.6.v zeigen, indem man statt des ersten Teils von 2.6.iv den zweiten anwendet. \square

2.7 Kurzzeit-Existenz des Gradientenflusses

2.7.i Satz. *Die Abbildung*

$$\mathcal{T} \rightarrow \mathcal{T}, T \mapsto A_T^{-1} E'[T]$$

ist stetig.

Beweis Seien $\epsilon > 0$ und $T \in \mathcal{T}$ beliebig, aber fest, und sei H wie in 2.6.vi. Dann gilt wegen (2.6) und 2.6.vi

$$\forall_{v \in \mathcal{T}} \quad E'[v] = (H(\cdot)[v], 0, 0).$$

Damit folgt mit (2.11) insbesondere $E'[T] \in C^1([z_0, z_{\max}]) \times \{(0, 0)\} \subset \mathcal{R}(A_T) = \mathcal{S}$, die Abbildung $\mathcal{T} \rightarrow \mathcal{T}, T \mapsto A_T^{-1} E'[T]$ ist also wohldefiniert. Sei $C^0 := C^0([z_0, z_{\max}])$. Wegen 2.3.vii und $H(\cdot)[T] \in C^0$ existiert ein $\delta_1 > 0$, so dass gilt

$$\forall_{\substack{S \in \mathcal{T} \\ \|T-S\|_{\mathcal{T}} < \delta_1}} \quad \forall_{\substack{g \in C^0 \\ \|H(\cdot)[T]-g\|_{C^0} < \delta_1}} \quad \|A_T^{-1}(H(\cdot)[T], 0, 0) - A_S^{-1}(g, 0, 0)\|_{\mathcal{T}} < \epsilon.$$

Nach 2.6.vi existiert ein $\delta_2 > 0$, so dass

$$\forall_{\substack{S \in \mathcal{T} \\ \|T-S\|_{\mathcal{T}} < \delta_2}} \quad \forall_{t \in [z_0, z_{\max}]} \quad |H(t)[T] - H(t)[S]| < \frac{\delta_1}{2}.$$

Daraus folgt

$$\forall_{\substack{S \in \mathcal{T} \\ \|T-S\|_{\mathcal{T}} < \delta_2}} \quad \|H(\cdot)[T] - H(\cdot)[S]\|_{C^0} \leq \frac{\delta_1}{2} < \delta_1.$$

Für $\delta := \min(\delta_1, \delta_2)$ und $S \in \mathcal{T}$ mit $\|T - S\|_{\mathcal{T}} < \delta$ folgt

$$\|A_T^{-1} E'[T] - A_S^{-1} E'[S]\|_{\mathcal{T}} = \|A_T^{-1}(H(\cdot)[T], 0, 0) - A_S^{-1}(H(\cdot)[S], 0, 0)\|_{\mathcal{T}} < \epsilon.$$

\square

2.7.ii Satz. (*Existenz des Gradientenflusses*) Zu jedem $T_0 \in \mathcal{T}$ existiert ein $\epsilon > 0$ und ein $T : [0, \epsilon] \rightarrow \mathcal{T}$, so dass gilt

$$\begin{aligned} \frac{\partial}{\partial t} T(t) &= -A_{T_0}^{-1} E'[T(t)] \text{ für } t \in]0, \epsilon[, \\ T(0) &= T_0. \end{aligned}$$

Beweis Sei $T_0 \in \mathcal{T}$. Analog zum Beweis von 2.7.i folgt, dass die Abbildung

$$F : \mathcal{T} \rightarrow \mathcal{T}, F[T] = -A_{T_0}^{-1} E'[T]$$

stetig ist. Es existiert also ein $\delta_1 > 0$, so dass für alle $T \in \mathcal{T}$ mit $\|T - T_0\| < \delta_1$ gilt

$$\|F[T] - F[T_0]\|_{\mathcal{T}} < 1 \Rightarrow \|F[T]\|_{\mathcal{T}} \leq \|F[T] - F[T_0]\|_{\mathcal{T}} + \|F[T_0]\|_{\mathcal{T}} < 1 + \|F[T_0]\|_{\mathcal{T}} =: c.$$

Sei H wie in 2.6.vi. Dann ex. ein $\delta_2 > 0$, so dass

$$\bigvee_{\substack{T \in \mathcal{T} \\ \|T_0 - T\|_{\mathcal{T}} < \delta_2}} \bigvee_{t \in [z_0, z_{\max}]} |H(t)[T_0] - H(t)[T]| < 1.$$

Seien $\delta := \min(\delta_1, \delta_2)$ und $D := B_{\frac{\delta}{2}}(T_0) := \{T \in \mathcal{T} \mid \|T - T_0\|_{\mathcal{T}} \leq \frac{\delta}{2}\}$. Für $T \in D$ folgt

$$\|H(\cdot)[T_0] - H(\cdot)[T]\|_{C^0} \leq 1,$$

also

$$\|H(\cdot)[T]\|_{C^0} \leq \|H(\cdot)[T] - H(\cdot)[T_0]\|_{C^0} + \|H(\cdot)[T_0]\|_{C^0} \leq 1 + \|H(\cdot)[T_0]\|_{C^0} =: C.$$

Sei $X := C^0([z_0, z_{\max}]) \times \{(0, 0)\}$ mit $\|(f, 0, 0)\|_X = \|f\|_{C^0}$. Wie im Beweis von 2.7.i folgt für $T \in \mathcal{T}$

$$E'[T] \in Y := C^1([z_0, z_{\max}]) \times \{(0, 0)\} \subset X \text{ und } E'[T] = (H(\cdot)[T], 0, 0).$$

Für $T \in D$ folgt

$$\|E'[T]\|_X = \|(H(\cdot)[T], 0, 0)\|_X = \|H(\cdot)[T]\|_{C^0} \leq C,$$

also ist $E'[D]$ eine bezüglich der Norm von X beschränkte Teilmenge von Y . Nach 8.6 aus [1] ist die Einbettung $C^1([z_0, z_{\max}]) \rightarrow C^0([z_0, z_{\max}])$ kompakt, also ist auch die Einbettung

$$J : Y = C^1([z_0, z_{\max}]) \times \{(0, 0)\} \rightarrow X = C^0([z_0, z_{\max}]) \times \{(0, 0)\}$$

kompakt. Da $E'[D]$ eine beschränkte Teilmenge von Y ist, ist $J(E'[D]) = E'[D]$ relativ kompakt in X . Wegen 2.3.vii ist

$$-A_{T_0}^{-1} : X \rightarrow \mathcal{T}$$

stetig, also ist $-A_{T_0}^{-1} E'[D] = F[D]$ relativ kompakt in \mathcal{T} . Für $B \subset D$ gilt $F[B] \subset F[D]$, also ist auch $F[B]$ relativ kompakt. Für $T \in D$ gilt $\|T - T_0\| \leq \frac{\delta}{2} < \delta_1$ und somit $\|F[T]\|_{\mathcal{T}} \leq c$. Nun folgt mit Theorem 2.2 und Lemma 2.2(i) aus [1] die Behauptung. \square

2.8 Regularisierung

Tests mit einer früheren Programmversion haben gezeigt, dass die damit berechneten Retrievals zwar Minima der Energie aus 2.1 sind, aber teilweise deutlich von den gesuchten Profilen abweichen und insbesondere Eigenschaften aufweisen, die in der Natur normalerweise nicht auftreten. Bei Beratungen mit Meteorologen aus Bonn hat uns Professor Simmer empfohlen, besonders auf Folgendes zu achten, das bei unseren Retrievals nicht erfüllt war: Natürliche Temperaturprofile oberhalb der Inversion haben in der Regel die Eigenschaften:

- (i) Das Profil ist konkav
- (ii) Der Erwartungswert des Gradienten liegt bei $-6.4 \frac{\text{K}}{\text{km}}$.

Wir erweitern nun die Energie, die bisher durch das quadratische Kostenfunktional gegeben ist, um einen Regularisierungsterm, der Profile begünstigen soll, die obige Eigenschaften haben. Damit verhindern wir, dass Profile, die deutlich von diesen Eigenschaften abweichen, Minima der Energie sind. Dieser Ansatz ist auch bekannt als Tikhonov Regularisierung, eine detaillierte Beschreibung findet sich in [20], typische Anwendungen in [2].

Zuerst führen wir einen Term für die Konkavität ein:

2.8.1 Erster Teil der Regularisierungsenergie

Zu gegebenem Temperaturprofil $T \in \mathcal{T}$ sei

$$\text{inv}[T] := \{z \in [200 \text{ m}, 8000 \text{ m}] | T'(z) = 0 \wedge T''(z) < 0\}$$

und

$$h_1 = h_1[T] = \sup(\text{inv}[T], 200 \text{ m}),$$

d.h. $\text{inv}[T]$ ist die Menge der Höhen aller Inversionen zwischen 200 m und 8000 m von T und $h_1[T]$ ist die Höhe der höchsten dieser Inversionen (falls vorhanden, ansonsten $h_1 = 200 \text{ m}$) und $h_2 = 8000 \text{ m}$, d.h. h_1 ist abhängig von T , h_2 jedoch nicht. Die Kurve $\Gamma = \Gamma[T]$ des Temperaturprofils sei gegeben durch den Graphen des Temperaturprofils:

$$\Gamma : [h_1, h_2] \rightarrow \mathbb{R}^2, z \mapsto (z, T(z)).$$

Für $x \in \Gamma$ sei $H(x)$ die Krümmung im Punkte x und $N(x)$ die (oberhalb des Graphen liegende) Normale im Punkte x . Es gilt

$$\Gamma \text{ ist konvex in } x \Leftrightarrow H(x) \cdot N(x) \geq 0,$$

$$\Gamma \text{ ist konkav in } x \Leftrightarrow H(x) \cdot N(x) \leq 0.$$

Um Konvexität stark zu bestrafen und gleichzeitig die Energie nach unten zu beschränken, wenden wir die Exponentialfunktion auf den Term an und integrieren ihn über die Kurve. Zusätzlich führen wir die Parameter $\alpha, \gamma > 0$ ein und erhalten folgende Energie:

$$E_c[T] = \gamma \int_{\Gamma[T]} e^{\alpha H(x) \cdot N(x)} ds(x). \quad (2.17)$$

2.8.2 Zweiter Teil der Regularisierungsenergie

Jetzt kümmern wir uns um den Erwartungswert des Gradienten. Seien h_1, h_2 wieder wie oben. Ferner nehmen wir an, dass der Gradient der Temperaturprofile in jedem Punkt normalverteilt ist mit Erwartungswert μ und Varianz σ . Für $z \in [h_1, h_2]$ und $\iota \in \mathbb{R}$ nehmen wir als Maß für die Wahrscheinlichkeit, dass der Gradient den Wert $T'(z)$ annimmt,

$$p_\iota[T](z) = 1 - P\left[(\mu - |T'(z) - \mu|_\iota) \leq X \leq (\mu + |T'(z) - \mu|_\iota)\right],$$

wobei $|x|_\iota := \sqrt{x^2 + \iota^2}$ und $P[X \leq a]$ die Wahrscheinlichkeit, dass eine normalverteilte Zufallsvariable X mit Erwartungswert μ und Varianz σ einen Wert kleiner gleich a annimmt. Dabei wählen wir $|\cdot|_\iota$ anstatt von $|\cdot|$, damit die Variation der Energie stetig ist (vgl. Abschnitt 2.9). Offensichtlich liegt $p_\iota[T](z)$ zwischen 0 und 1 und je näher der Wert bei 0 liegt, desto weiter ist $T'(z)$ vom Erwartungswert μ entfernt. Da die Energie Profile bevorzugen soll, bei denen $T'(z)$ nah beim Erwartungswert liegt, muss sie dafür sorgen, dass $p_\iota[T](z)$ nahe bei 1 liegt. Daher betrachten wir

$$\left(\frac{1}{p_\iota[T](z)} - 1\right) \in [0, \infty[$$

und definieren als Energie mit dem Parameter $\beta > 0$

$$E_P[T] = \beta \int_{h_1[T]}^{h_2} \left(\frac{1}{p_\iota[T](z)} - 1\right) dz. \quad (2.18)$$

Insgesamt erhalten wir die Regularisierungsenergie

$$\begin{aligned} E_{\text{reg}}[T] &= E_c[T] + E_P[T] \\ &= \gamma \int_{\Gamma[T]} e^{\alpha H(x) \cdot N(x)} ds(x) + \beta \int_{h_1[T]}^{h_2} \left(\frac{1}{p_\iota[T](z)} - 1\right) dz. \end{aligned}$$

Statt E zu minimieren, minimieren wir nun $E_{\text{ges}}[\cdot] := E[\cdot] + E_{\text{reg}}[\cdot]$.

Zur Implementierung von $E_P[T]$ ist es nützlich, die Auswertung von $p_\iota[T](z)$ auf die sog. Errorfunction erf zurückzuführen, die bereits in der C++ Bibliothek `<math.h>` vorhanden ist. Dies liefert folgendes Lemma

2.8.i Lemma. *Es gilt*

$$p_\iota[T](z) = 1 - \text{erf}\left(\frac{|T'(z) - \mu|_\iota}{\sqrt{2}\sigma}\right).$$

Beweis Sei Φ die Gauß'sche Verteilungsfunktion, dann gilt

$$p_\iota[T](z) = 2 - 2\Phi\left(\frac{|T'(z) - \mu|_\iota}{\sigma}\right),$$

denn aus

$$P[a \leq X \leq b] = \Phi\left(\frac{b - \mu}{\sigma}\right) - \Phi\left(\frac{a - \mu}{\sigma}\right)$$

folgt

$$\begin{aligned}
 p_\iota[T](z) &= 1 - \left(\Phi \left(\frac{|T'(z) - \mu|_\iota}{\sigma} \right) - \Phi \left(-\frac{|T'(z) - \mu|_\iota}{\sigma} \right) \right) \\
 &= 1 - \left(2\Phi \left(\frac{|T'(z) - \mu|_\iota}{\sigma} \right) - 1 \right) \\
 &= 2 - 2\Phi \left(\frac{|T'(z) - \mu|_\iota}{\sigma} \right).
 \end{aligned} \tag{2.19}$$

Um Φ auszuwerten, nutzen wir nun erf:

$$\begin{aligned}
 \int_0^x e^{-\frac{t^2}{2}} dt &\stackrel{\phi(t)=\frac{t}{\sqrt{2}}}{=} \sqrt{2} \int_0^{\frac{x}{\sqrt{2}}} e^{-\phi(t)^2} \phi'(t) dt = \sqrt{2} \int_0^{\frac{x}{\sqrt{2}}} e^{-t^2} dt \\
 &= \sqrt{2} \frac{\sqrt{\pi}}{2} \frac{2}{\sqrt{\pi}} \int_0^{\frac{x}{\sqrt{2}}} e^{-t^2} dt = \frac{\sqrt{\pi}}{\sqrt{2}} \operatorname{erf} \left(\frac{x}{\sqrt{2}} \right).
 \end{aligned}$$

Damit folgt

$$\begin{aligned}
 \Phi(x) &= \frac{1}{\sqrt{2\pi}} \int_{-\infty}^x e^{-\frac{t^2}{2}} dt = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^0 e^{-\frac{t^2}{2}} dt + \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{t^2}{2}} dt \\
 &= \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \frac{\sqrt{\pi}}{\sqrt{2}} \operatorname{erf} \left(\frac{x}{\sqrt{2}} \right) = \frac{1}{2} + \frac{1}{2} \operatorname{erf} \left(\frac{x}{\sqrt{2}} \right),
 \end{aligned}$$

insgesamt gilt also

$$\begin{aligned}
 p_\iota[T](z) &= 2 - 2\Phi \left(\frac{|T'(z) - \mu|_\iota}{\sigma} \right) = 2 - 2 \left(\frac{1}{2} + \frac{1}{2} \operatorname{erf} \left(\frac{|T'(z) - \mu|_\iota}{\sqrt{2}\sigma} \right) \right) \\
 &= 1 - \operatorname{erf} \left(\frac{|T'(z) - \mu|_\iota}{\sqrt{2}\sigma} \right).
 \end{aligned}$$

□

2.8.3 Variation der Regularisierungsenergie

Um die Variation von $E_{\text{reg}}[T]$ ausrechnen zu können, müssen wir zuvor

$$\left. \frac{d}{d\epsilon} h_1[T + \epsilon\vartheta] \right|_{\epsilon=0}$$

berechnen. Dazu sei

$$\mathcal{T}_{\text{inv}} := \{T \in \mathcal{T} \mid \operatorname{inv}[T] \neq \emptyset\}.$$

2.8.ii Lemma. Für $T \in \mathcal{T}_{\text{inv}}$ gilt

$$\left. \frac{d}{d\epsilon} h_1[T + \epsilon\vartheta] \right|_{\epsilon=0} = -\frac{\vartheta'(h_1[T])}{T''(h_1[T])}, \tag{2.20}$$

insbesondere ist h_1 stetig in T .

Beweis Sei $T \in \mathcal{T}_{\text{inv}}$, also $\text{inv}[T] \neq \emptyset$, und sei $f(\epsilon, z) := T'(z) + \epsilon\vartheta'(z)$ und $h(\epsilon) := h_1[T + \epsilon\vartheta]$. Für kleine ϵ gilt dann

$$f(\epsilon, h(\epsilon)) = T'(h(\epsilon)) + \epsilon\vartheta'(h(\epsilon)) = (T' + \epsilon\vartheta')(h_1[T + \epsilon\vartheta]) = 0, \quad (2.21)$$

da die Inversion ein lokales Maximum des Temperaturprofils $T + \epsilon\vartheta$ ist. h ist also durch die Gleichung (2.21) implizit definiert und falls gilt $\partial_2 f(0, h(0)) \neq 0$, so folgt aus dem Satz über implizit definierte Funktionen:

$$h'(\epsilon) = -\frac{\partial_1 f(\epsilon, h(\epsilon))}{\partial_2 f(\epsilon, h(\epsilon))} = -\frac{\vartheta'(h(\epsilon))}{(T + \epsilon\vartheta)''(h(\epsilon))},$$

und es folgt

$$\left. \frac{d}{d\epsilon} h_1[T + \epsilon\vartheta] \right|_{\epsilon=0} = h'(0) = -\frac{\vartheta'(h(0))}{T''(h(0))} = -\frac{\vartheta'(h_1[T])}{T''(h_1[T])}.$$

□

Sind $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ stetig und $b : \mathbb{R} \rightarrow \mathbb{R}$ differenzierbar, so gilt mit $F(x, y) := \int_a^x f(z, y) dz$

$$\begin{aligned} \frac{d}{d\epsilon} \left(\int_a^{b(\epsilon)} f(z, \epsilon) dz \right) &= \frac{d}{d\epsilon} (F(b(\epsilon), \epsilon)) = \partial_1 F(b(\epsilon)) b'(\epsilon) + \partial_2 F(b(\epsilon)) \cdot 1 \\ &= f(b(\epsilon), \epsilon) b'(\epsilon) + \int_a^{b(\epsilon)} \partial_2 f(z, \epsilon) dz. \end{aligned}$$

Daraus folgt

$$\frac{d}{d\epsilon} \left(\int_{b(\epsilon)}^a f(z, \epsilon) dz \right) = \int_{b(\epsilon)}^a \partial_2 f(z, \epsilon) dz - f(b(\epsilon), \epsilon) b'(\epsilon). \quad (2.22)$$

Für $z \in [h_1[T], h_2]$ und $x = \Gamma(z) = (z, T(z))$ gilt

$$\begin{aligned} H(x) &= H(\Gamma(z)) = \Gamma''(z) = (0, T''(z)) \\ N(x) &= N(\Gamma(z)) = \frac{(-\Gamma_2'(z), \Gamma_1'(z))}{\|(-\Gamma_2'(z), \Gamma_1'(z))\|} = \frac{(-T'(z), 1)}{\sqrt{1 + T'(z)^2}} \\ \Rightarrow H(x) \cdot N(x) &= \frac{T''(z)}{\sqrt{1 + T'(z)^2}}. \end{aligned}$$

Im Folgenden sei $w(z) := \sqrt{1 + T'(z)^2}$, dann gilt

$$\begin{aligned} \left. \frac{d}{d\epsilon} \sqrt{1 + (T + \epsilon\vartheta)'(z)^2} \right|_{\epsilon=0} &= \frac{1}{2w(z)} 2T'(z)\vartheta'(z) = \frac{T'(z)\vartheta'(z)}{w(z)} \\ \left. \frac{d}{d\epsilon} \left(\frac{(T + \epsilon\vartheta)''(z)}{\sqrt{1 + (T + \epsilon\vartheta)'(z)^2}} \right) \right|_{\epsilon=0} &= \frac{\vartheta''(z)w(z) - \frac{T'(z)\vartheta'(z)}{w(z)} T''(z)}{w(z)^2} \\ &= \frac{\vartheta''(z)}{w(z)} - \frac{T'(z)T''(z)\vartheta'(z)}{w(z)^3} \end{aligned}$$

$$\begin{aligned}
& \frac{d}{d\epsilon} \left(\sqrt{1 + (T + \epsilon\vartheta)'(z)^2} \cdot e^{\alpha \frac{(T + \epsilon\vartheta)''(z)}{\sqrt{1 + (T + \epsilon\vartheta)'(z)^2}}} \right) \Big|_{\epsilon=0} \\
&= \frac{T'(z)\vartheta'(z)}{w(z)} e^{\alpha \frac{T''(z)}{w(z)}} + w(z) e^{\alpha \frac{T''(z)}{w(z)}} \alpha \left(\frac{\vartheta''(z)}{w(z)} - \frac{T'(z)T''(z)\vartheta'(z)}{w(z)^3} \right) \\
&= e^{\alpha \frac{T''(z)}{w(z)}} \left(\frac{T'(z)\vartheta'(z)}{w(z)} - \alpha \frac{T'(z)T''(z)\vartheta'(z)}{w(z)^2} + \alpha \vartheta''(z) \right) \\
&= e^{\alpha \frac{T''(z)}{w(z)}} \left(\frac{T'(z)}{w(z)} - \alpha \frac{T'(z)T''(z)}{w(z)^2} \right) \vartheta'(z) + \alpha e^{\alpha \frac{T''(z)}{w(z)}} \vartheta''(z).
\end{aligned}$$

Wegen (2.17) und obiger Rechnung gilt

$$\begin{aligned}
E_c[T] &= \gamma \int_{\Gamma[T]} e^{\alpha H(x) \cdot N(x)} ds(x) = \gamma \int_{h_1[T]}^{h_2} e^{\alpha H(\Gamma(z)) \cdot N(\Gamma(z))} |\Gamma'(z)| dz \\
&= \gamma \int_{h_1[T]}^{h_2} e^{\alpha \frac{T''(z)}{\sqrt{1+T'(z)^2}} \sqrt{1+T'(z)^2}} dz
\end{aligned}$$

und es folgt daraus mit (2.20) und (2.22), dass

$$\begin{aligned}
\langle E'_c[T], \vartheta \rangle &= \gamma \int_{h_1[T]}^{h_2} e^{\alpha \frac{T''(z)}{w(z)}} \left(\frac{T'(z)}{w(z)} - \alpha \frac{T'(z)T''(z)}{w(z)^2} \right) \vartheta'(z) + \alpha e^{\alpha \frac{T''(z)}{w(z)}} \vartheta''(z) dz \\
&\quad + \gamma e^{\alpha \frac{T''(h_1[T])}{w(h_1[T])}} w(h_1[T]) \frac{\vartheta'(h_1[T])}{T''(h_1[T])}.
\end{aligned} \tag{2.23}$$

Ferner folgt mit Hilfe von (2.19)

$$\begin{aligned}
\frac{d}{d\epsilon} p_l[T + \epsilon\vartheta](z) \Big|_{\epsilon=0} &= \frac{d}{d\epsilon} \left(2 - 2\Phi \left(\frac{|(T + \epsilon\vartheta)'(z) - \mu|_l}{\sigma} \right) \right) \Big|_{\epsilon=0} \\
&= -2\varphi \left(\frac{|T'(z) - \mu|_l}{\sigma} \right) \frac{1}{\sigma} \frac{T'(z) - \mu}{|T'(z) - \mu|_l} \vartheta'(z),
\end{aligned}$$

wobei $\varphi(t) = e^{-\frac{t^2}{2}}$ die Dichte der Gauß'schen Verteilung ist. Somit ist

$$\frac{d}{d\epsilon} \left(\frac{1}{p_l[T + \epsilon\vartheta](z)} - 1 \right) \Big|_{\epsilon=0} = \frac{2\varphi \left(\frac{|T'(z) - \mu|_l}{\sigma} \right) \frac{1}{\sigma} \frac{T'(z) - \mu}{|T'(z) - \mu|_l} \vartheta'(z)}{p_l[T](z)^2}$$

und es folgt mit (2.18), (2.20) und (2.22)

$$\begin{aligned}
\langle E'_P[T], \vartheta \rangle &= \beta \int_{h_1[T]}^{h_2} \frac{2\varphi \left(\frac{|T'(z) - \mu|_l}{\sigma} \right) \frac{1}{\sigma} \frac{T'(z) - \mu}{|T'(z) - \mu|_l} \vartheta'(z)}{p_l[T](z)^2} dz \\
&\quad + \beta \left(\frac{1}{p_l[T](h_1[T])} - 1 \right) \frac{\vartheta'(h_1[T])}{T''(h_1[T])}.
\end{aligned} \tag{2.24}$$

(2.23) und (2.24) liefern eine Darstellung von $\langle E'_{\text{reg}}[T], \vartheta \rangle$.

2.9 Stetigkeit der Variation der Regularisierungsenergie

Aus (2.23) und (2.24) sieht man, dass $E'_c[\cdot]$ und $E'_P[\cdot]$ jeweils aus zwei Teilen bestehen und sich von der Struktur her sehr ähneln. Wir beweisen daher, um die Stetigkeit von $E'_{\text{reg}}[\cdot]$ zu zeigen, zunächst zwei allgemeinere Sätze, die dann zusammen die Stetigkeit von $E'_c[\cdot]$ und $E'_P[\cdot]$ liefern und somit auch die von $E'_{\text{reg}}[\cdot]$, allerdings gilt die Stetigkeit nicht auf ganz \mathcal{T} .

2.9.i Satz. *Seien $\mathcal{V} \subset X := C^2[a, b]$, $a < h_2 \leq b$, $h_1 : \mathcal{V} \rightarrow]a, h_2]$ und $F : \mathcal{V} \rightarrow X'$ so, dass ein $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ existiert mit entweder*

$$\forall v \in \mathcal{V} \quad \forall \vartheta \in X \quad \langle F(v), \vartheta \rangle = \int_{h_1(v)}^{h_2} f(v'(z), v''(z)) \vartheta'(z) dz$$

oder

$$\forall v \in \mathcal{V} \quad \forall \vartheta \in X \quad \langle F(v), \vartheta \rangle = \int_{h_1(v)}^{h_2} f(v'(z), v''(z)) \vartheta''(z) dz.$$

Ist h_1 stetig in v und existiert ein $c > 0$, so dass $a < h_1(v) - c$ und f stetig in $I := I_c(v') \times I_c(v'')$ ist, wobei

$$I_c(g) := \left[\min_{z \in [h_1(v) - c, h_2]} g(z) - c, \max_{z \in [h_1(v) - c, h_2]} g(z) + c \right] \quad \text{für } g \in C^0[a, b],$$

so ist F stetig in v .

Beweis Wir beweisen die Aussage nur für den Fall, dass die obere Bedingung für F erfüllt ist. Der andere Fall folgt analog, da im Beweis nur $|\vartheta'(z)| \leq \|\vartheta\|_{2,\infty}$ benutzt wird und auch $|\vartheta''(z)| \leq \|\vartheta\|_{2,\infty}$ gilt.

Sei $\epsilon > 0$ beliebig und sei $M := \|f\|_{\infty, I} + 1 > 0$. Dann existiert wegen der Stetigkeit von h_1 in v ein $\delta_1 > 0$, so dass für alle $w \in \mathcal{V}$ gilt

$$\|v - w\|_{2,\infty} < \delta_1 \Rightarrow |h_1(v) - h_1(w)| < \min\left(\frac{\epsilon}{2M}, c\right). \quad (2.25)$$

Da I kompakt ist, ist f auf I sogar gleichmäßig stetig, also existiert ein $\delta_2 > 0$, so dass für alle $x_1, y_1 \in I_c(v')$ und alle $x_2, y_2 \in I_c(v'')$ gilt

$$|x_1 - y_1| < \delta_2, |x_2 - y_2| < \delta_2 \Rightarrow |f(x_1, y_1) - f(x_2, y_2)| < \frac{\epsilon}{2(b-a)}. \quad (2.26)$$

Sei nun $\delta := \min(\delta_1, \delta_2, c)$ und $w \in \mathcal{V}$ mit $\|v - w\|_{2,\infty} < \delta$. Wegen $\|v - w\|_{2,\infty} < \delta \leq \delta_2$ folgt

$$|v'(z) - w'(z)| < \delta_2, |v''(z) - w''(z)| < \delta_2 \quad \text{für } z \in [a, b].$$

Wegen $\|v - w\|_{2,\infty} < \delta \leq \delta_1$ folgt aus (2.25)

$$h_1(v) - c < h_1(w) < h_1(v) + c,$$

also insbesondere

$$[h_1(w), h_2] \subset [h_1(v) - c, h_2], \text{ sowie offensichtlich } [h_1(v), h_2] \subset [h_1(v) - c, h_2].$$

Wegen $\|v - w\|_{2,\infty} < \delta \leq c$ gilt für $z \in [h_1(v) - c, h_2] \subset [a, b]$

$$\min_{z \in [h_1(v) - c, h_2]} v'(z) - c \leq v'(z) - c \leq w'(z) \leq v'(z) + c \leq \max_{z \in [h_1(v) - c, h_2]} v'(z) + c,$$

also $w'(z) \in I_c(v')$. Analog folgt $w''(z) \in I_c(v'')$ für $z \in [h_1(v) - c, h_2]$. Ferner gilt offensichtlich $(v'(z), v''(z)) \in I_c(v') \times I_c(v'')$ für $z \in [h_1(v) - c, h_2]$, es gilt also insgesamt

$$\begin{aligned} (v'(z), v''(z)) \in I, (w'(z), w''(z)) \in I \\ |v'(z) - w'(z)| < \delta_2, |v''(z) - w''(z)| < \delta_2 \end{aligned} \quad \text{für } z \in [h_1(v), h_2] \cup [h_1(w), h_2]. \quad (2.27)$$

Sei nun $\vartheta \in X, \vartheta \neq 0$ beliebig. Im Falle $h_1(v) \leq h_1(w)$ gilt

$$\begin{aligned} | \langle (F(v) - F(w)), \vartheta \rangle | &= \left| \int_{h_1(v)}^{h_2} f(v'(z), v''(z)) \vartheta'(z) dz - \int_{h_1(w)}^{h_2} f(w'(z), w''(z)) \vartheta'(z) dz \right| \\ &\leq \int_{h_1(w)}^{h_2} \underbrace{|f(v'(z), v''(z)) - f(w'(z), w''(z))|}_{< \frac{\epsilon}{2(b-a)} \text{ nach (2.26), (2.27)}} \underbrace{|\vartheta'(z)|}_{\leq \|\vartheta\|_{2,\infty}} dz \\ &\quad + \int_{h_1(v)}^{h_1(w)} \underbrace{|f(v'(z), v''(z))|}_{< M \text{ wegen (2.27)}} \underbrace{|\vartheta'(z)|}_{\leq \|\vartheta\|_{2,\infty}} dz \\ &< \underbrace{(h_2 - h_1(w))}_{\leq b-a} \frac{\epsilon}{2(b-a)} \|\vartheta\|_{2,\infty} + \underbrace{(h_1(w) - h_1(v))}_{< \frac{\epsilon}{2M} \text{ nach (2.25)}} M \|\vartheta\|_{2,\infty} \\ &< \epsilon \|\vartheta\|_{2,\infty}. \end{aligned}$$

Für den Fall $h_1(v) > h_1(w)$ folgt analog

$$| \langle (F(v) - F(w)), \vartheta \rangle | < \epsilon \|\vartheta\|_{2,\infty}.$$

Mit Lemma 5.1.v folgt nun die Stetigkeit von F in v . □

2.9.ii Satz. Seien $\mathcal{V} \subset X := C^2[a, b], a < h_2 \leq b, h_1 : \mathcal{V} \rightarrow]a, h_2]$ und $F : \mathcal{V} \rightarrow X'$ so, dass ein $f : \mathbb{R}^2 \rightarrow \mathbb{R}$ existiert mit

$$\forall_{v \in \mathcal{V}} \quad \forall_{\vartheta \in X} \quad \langle F(v), \vartheta \rangle = f(v'(h_1(v)), v''(h_1(v))) \vartheta'(h_1(v)).$$

Ist h_1 stetig in v und ist f stetig in $(v'(h_1(v)), v''(h_1(v)))$, so ist F stetig in v .

Beweis Sei $\epsilon > 0$ beliebig. Sei $M := |f(v'(h_1(v)), v''(h_1(v)))| + 1 > 0$. Dann existiert wegen der Stetigkeit von f in $(v'(h_1(v)), v''(h_1(v)))$ ein $\delta_1 > 0$, so dass für alle $x, y \in \mathbb{R}$ gilt

$$|v'(h_1(v)) - x| < \delta_1, |v''(h_1(v)) - y| < \delta_1 \Rightarrow |f(v'(h_1(v)), v''(h_1(v))) - f(x, y)| < \frac{\epsilon}{2}.$$

Ferner existiert wegen der Stetigkeit von v' in $h_1(v)$, ein $\delta_2 > 0$, so dass für alle $x \in [a, b]$ gilt

$$|h_1(v) - x| < \delta_2 \Rightarrow |v'(h_1(v)) - v'(x)| < \frac{\delta_1}{2},$$

und wegen der Stetigkeit von h_1 in v existiert ein $\delta_3 > 0$, so dass für alle $w \in \mathcal{V}$ gilt

$$\|v - w\|_{2,\infty} < \delta_3 \Rightarrow |h_1(v) - h_1(w)| < \min\left(\delta_2, \frac{\epsilon}{2M}\right).$$

Sei nun $w \in \mathcal{V}$ mit $\|v - w\|_{2,\infty} < \delta' := \min\left(\frac{\delta_1}{2}, \delta_3\right)$. Dann gilt insbesondere

$$|v'(h_1(w)) - w'(h_1(w))| < \frac{\delta_1}{2}.$$

Es folgt nun

$$|v'(h_1(v)) - w'(h_1(w))| \leq \underbrace{|v'(h_1(v)) - v'(h_1(w))|}_{< \frac{\delta_1}{2}} + \underbrace{|v'(h_1(w)) - w'(h_1(w))|}_{< \frac{\delta_1}{2}} < \delta_1.$$

Analog findet man ein δ'' , so dass für $w \in \mathcal{V}$ mit $\|v - w\|_{2,\infty} < \delta''$ gilt

$$|v''(h_1(v)) - w''(h_1(w))| < \delta_1.$$

Sei nun $\delta := \min(\delta', \delta'')$ und $w \in \mathcal{V}$ mit $\|v - w\|_{2,\infty} < \delta$. Dann gilt

$$|v'(h_1(v)) - w'(h_1(w))| < \delta_1, |v''(h_1(v)) - w''(h_1(w))| < \delta_1.$$

Sei nun $\vartheta \in X, \vartheta \neq 0$ beliebig. Nach dem Mittelwertsatz existiert ein ζ , das zwischen $h_1(v)$ und $h_1(w)$ liegt, mit

$$\vartheta'(h_1(v)) - \vartheta'(h_1(w)) = \vartheta''(\zeta)(h_1(v) - h_1(w)),$$

also gilt

$$|\vartheta'(h_1(v)) - \vartheta'(h_1(w))| \leq \|\vartheta\|_{2,\infty} |h_1(v) - h_1(w)| < \frac{\epsilon}{2M} \|\vartheta\|_{2,\infty}.$$

Insgesamt folgt

$$\begin{aligned} & | \langle (F(v) - F(w)), \vartheta \rangle | \\ &= | f(v'(h_1(v)), v''(h_1(v)))\vartheta'(h_1(v)) - f(w'(h_1(w)), w''(h_1(w)))\vartheta'(h_1(w)) | \\ &\leq \underbrace{| f(v'(h_1(v)), v''(h_1(v)))\vartheta'(h_1(v)) - f(v'(h_1(v)), v''(h_1(v)))\vartheta'(h_1(w)) |}_{< M|\vartheta'(h_1(v)) - \vartheta'(h_1(w))| < M\frac{\epsilon}{2M}\|\vartheta\|_{2,\infty}} \\ &\quad + \underbrace{| f(v'(h_1(v)), v''(h_1(v)))\vartheta'(h_1(w)) - f(w'(h_1(w)), w''(h_1(w)))\vartheta'(h_1(w)) |}_{< \frac{\epsilon}{2}|\vartheta'(h_1(w))| \leq \frac{\epsilon}{2}\|\vartheta\|_{2,\infty}} \\ &< \epsilon \|\vartheta\|_{2,\infty}. \end{aligned}$$

Mit Lemma 5.1.v folgt nun die Stetigkeit von F in v . □

2.9.iii Satz.

$$E'_c[\cdot] : \mathcal{T}_{\text{inv}} \rightarrow \mathcal{T}', T \mapsto E'[T] \text{ und } E'_P[\cdot] : \mathcal{T}_{\text{inv}} \rightarrow \mathcal{T}', T \mapsto E'[T]$$

sind stetig.

Beweis Anwendung von 2.9.i und 2.9.ii auf $E'_c[\cdot]$ bzw. $E'_P[\cdot]$, wobei die Voraussetzungen wegen 2.8.ii, (2.23) und (2.24) erfüllt sind. □

3 Diskretisierung

3.1 Ortsdiskretisierung

Um den Ort sinnvoll zu diskretisieren, ist es nötig zu wissen, wie die Datensätze, auf denen in der Praxis gerechnet werden soll, aufgebaut sind und dies bei der Diskretisierung zu berücksichtigen. Ein typischer Datensatz besteht aus

- einem Gitter, das sind l Höhen, $z_1 < z_2 < \dots < z_l$,
- dem Profil der Atmosphärenparameter in Form der Werte $Q(z_1), \dots, Q(z_l)$,
- den Strahlungstemperaturen $T_{bm11}, \dots, T_{bmnm}$ des Referenztemperaturprofils T_{ref} , das wir bestimmen wollen,
- den dazu gehörenden Frequenzen ν_1, \dots, ν_n und Winkeln $\theta_1, \dots, \theta_m$ und
- den Werten des Referenztemperaturprofils auf den ersten n_1 und den letzten n_2 Gitterpunkten ($n_1, n_2 \in \mathbb{N}$)

$$T_{\text{ref}}(z_1), \dots, T_{\text{ref}}(z_{n_1}), T_{\text{ref}}(z_{l-n_2+1}), \dots, T_{\text{ref}}(z_l).$$

In unserem Fall sind dies meist die Temperaturen bis 200m, da es bei dem Radiometer in Cabauw einen 200m hohen Turm gibt, an dem die Temperatur gemessen werden kann, und die Temperaturen über 8000m, da diese hinreichend bekannt sind (vgl. Abschnitt 4.1).

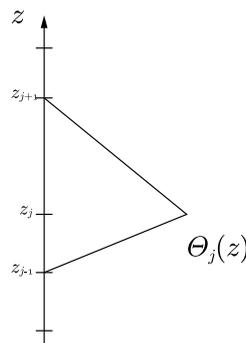


Abbildung 3.1: Θ_j , die FE-Basisfunktion zum Knoten z_j

Die naheliegende FE-Diskretisierung sieht folgendermaßen aus (eine Einführung in die Theorie der Finiten Element wird in [5] gegeben): Für $j \in \{1, \dots, l\}$ sei Θ_j die FE-Basisfunktion zum Knoten z_j (vgl. Abbildung 3.1). Ein diskretes Profil T ist dann gegeben durch

$$T(z) = \sum_{i=1}^l T_i \Theta_i(z),$$

wobei

$$T_i = T_{\text{ref}}(z_i) \text{ für } i \in \{1, \dots, n_1, l - n_2 + 1, \dots, l\},$$

um die Vorgabe der Werte an ersten n_1 und den letzten n_2 Gitterpunkten zu erfüllen, und

$$T_i \in [0, \infty[\text{ für } i \in \{n_1 + 1, \dots, l - n_2\}.$$

Allerdings werden wir diese Diskretisierung hier nicht verwenden: Die natürlichen Temperaturprofile sind in der Regel näherungsweise stückweise linear. Bei bestimmten Wetterbedingungen gibt es sog. **Inversionen**, das sind Zonen, in denen die Temperatur bei steigender Höhe zunimmt (normalerweise ist genau das Gegenteil der Fall). Die Zonen sind von besonderem Interesse und sollen daher so genau wie möglich bestimmt werden können. Da die Höhen, an denen die eventuell vorhandene Inversion liegt, ebenfalls von den Wetterbedingungen abhängen, gibt es mit obiger Diskretisierung unter Umständen keine Gitterpunkte in der Nähe der Inversionshöhen, wenn nicht sehr viele Freiheitsgrade eingeführt werden. Da wir nur wenige Messdaten haben, sollte man sich auf möglichst wenige Freiheitsgrade beschränken. Anstatt also die Diskretisierung durch ein feineres Gitter zu verbessern, soll auch das Gitter variabel sein. Für festes $M \in \mathbb{N}$ wählen wir folgenden Ansatzraum

$$\begin{aligned} \mathcal{T}_M := \{T \in \mathcal{T} \mid & \exists z_{n_1} < Z_1 < Z_2 < \dots < Z_M < z_{l-n_2+1} \\ & T|_{[z_i, z_{i+1}]} \text{ ist affin für } i \in \{1, \dots, n_1 - 1, l - n_2 + 1, \dots, l - 1\}, \\ & T|_{[Z_i, Z_{i+1}]} \text{ ist affin für } i \in \{1, \dots, M - 1\}, \\ & T|_{[z_{n_1}, Z_1]}, T|_{[Z_M, z_{l-n_2+1}]} \text{ sind affin,} \\ & T(z_i) = T_{\text{ref}}(z_i) \text{ für } i \in \{1, \dots, n_1, l - n_2 + 1, \dots, l\}\}. \end{aligned}$$

\mathcal{T}_M ist mit der üblichen Addition kein Vektorraum! Ein Element $T \in \mathcal{T}_M$ können wir mit den zwei Vektoren

$$\bar{T} = (T_1, \dots, T_M) \text{ und } \bar{Z} = (Z_1, \dots, Z_M)$$

identifizieren, wobei

$$T_i := T(Z_i) \text{ für } i \in \{1, \dots, M\}.$$

\mathcal{T}_M enthält auch physikalisch verbotene Profile, da wir nicht $T_i \geq 0$ für $i \in \{1, \dots, M\}$ fordern. Dies ist in der Praxis jedoch kein Problem, da solche Profile nicht in der Nähe von Minima der Energie liegen, wenn die Strahlungstemperatur-Messwerte von einem Profil mit positiver Temperatur kommen. Zur Energieminimierung sollen die Einträge dieser Vektoren frei bewegt werden können, mit der Einschränkung

$$z_{n_1} < Z_1 < Z_2 < \dots < Z_M < z_{l-n_2+1},$$

daher müssen wir geeignete Testfunktionen finden, die diesen Variationen entsprechen, während die Zeitschrittweitensteuerung dafür sorgt, dass die Profile diese Einschränkung immer erfüllen (vgl. 3.10.5).

3.2 Testfunktionen

Sei $z_1 < \dots < z_l$ ein Gitter und $T_1 < \dots < T_l$ Werte an den Gitterpunkten. Dann ist dadurch eindeutig eine stückweise affine Funktion $T : [z_1, z_l] \rightarrow \mathbb{R}$ bestimmt, für die gilt $T(z_i) = T_i$ für $1 \leq i \leq l$.

3.2.1 Variation der Temperatur

Sei Θ_j die FE-Basisfunktion zum Gitter z_1, \dots, z_l und Knoten z_j . Für die Funktion $T_\epsilon(z) := T(z) + \epsilon\Theta_j(z)$ gilt dann für $i \neq j$

$$T_\epsilon(z_i) = T(z_i) + \underbrace{\epsilon\Theta_j(z_i)}_{=0} = T_i,$$

und

$$T_\epsilon(z_j) = T(z_j) + \underbrace{\epsilon\Theta_j(z_j)}_{=1} = T_j + \epsilon.$$

T_ϵ entspricht also der Funktion T , nur dass der Wert am Knoten z_j um ϵ erhöht ist. Es gilt also

$$\frac{d}{dT_j} T(z) = \lim_{\epsilon \rightarrow 0} \frac{T_\epsilon(z) - T(z)}{\epsilon} = \lim_{\epsilon \rightarrow 0} \frac{T_\epsilon(z) - T_0(z)}{\epsilon} = \left. \frac{d}{d\epsilon} (T_\epsilon(z)) \right|_{\epsilon=0} = \Theta_j(z),$$

und somit ist Θ_j die Testfunktion zur Variation der Temperatur am Knoten z_j .

3.2.2 Variation der Höhe

Sei Θ_j^ϵ die FE-Basisfunktion zum Gitter $\hat{z}_1, \dots, \hat{z}_l$ und Knoten \hat{z}_j , wobei

$$\hat{z}_i = \begin{cases} z_i, & i \neq j \\ z_j + \epsilon, & i = j, \end{cases}$$

vgl. Abbildung 3.2. Für die Funktion $\hat{T}_\epsilon(z) := T(z - \epsilon\Theta_j^\epsilon(z))$ gilt dann für $i \neq j$

$$\hat{T}_\epsilon(\hat{z}_i) = \hat{T}_\epsilon(z_i) = T(z_i - \underbrace{\epsilon\Theta_j^\epsilon(z_i)}_{=0}) = T_i$$

und

$$\hat{T}_\epsilon(\hat{z}_j) = \hat{T}_\epsilon(z_j + \epsilon) = T(z_j + \epsilon - \underbrace{\epsilon\Theta_j^\epsilon(z_j + \epsilon)}_{=1}) = T_j.$$

\hat{T}_ϵ entspricht also der Funktion T , nur dass der Knoten z_j um ϵ erhöht ist (vgl. Abbildung 3.2).

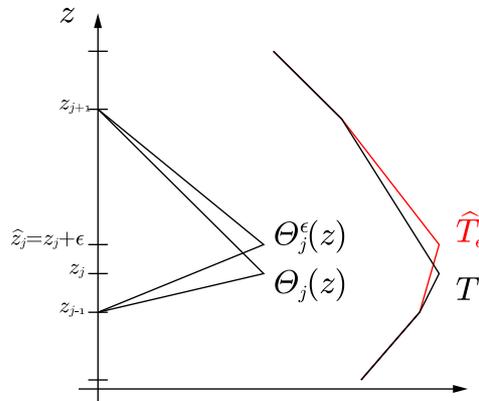


Abbildung 3.2: Θ_j^ϵ und \hat{T}_ϵ im Vergleich zu Θ_j und T

Es gilt also

$$\begin{aligned} \frac{d}{dz_j} T(z) &= \lim_{\epsilon \rightarrow 0} \frac{\widehat{T}_\epsilon(z) - T(z)}{\epsilon} = \lim_{\epsilon \rightarrow 0} \frac{\widehat{T}_\epsilon(z) - \widehat{T}_0(z)}{\epsilon} \\ &= \frac{d}{d\epsilon} \left(\widehat{T}_\epsilon(z) \right) \Big|_{\epsilon=0} = \frac{d}{d\epsilon} \left(T(z - \epsilon \Theta_j^\epsilon(z)) \right) \Big|_{\epsilon=0} \\ &= T'(z) \cdot \left(-\Theta_j^\epsilon(z) - \epsilon \frac{d}{d\epsilon} (\Theta_j^\epsilon(z)) \right) \Big|_{\epsilon=0} = -T'(z) \Theta_j(z), \end{aligned}$$

also ist $-T' \Theta_j$ die Testfunktion zur Variation der Höhe am Knoten z_j .

3.3 Zeitdiskretisierung

Für die Zeitdiskretisierung benötigen wir eine Darstellung von \mathcal{T} , in der \mathcal{T} ein Vektorraum ist, also insbesondere eine Basis hat (vgl. Abschnitt 3.3.2). Passend zu der Idee unserer geometrischen Metrik werden wir die Temperaturprofile nicht wie bisher als Abbildungen von einem Intervall nach \mathbb{R} auffassen, sondern als Kurven, die den Graphen dieser Abbildungen entsprechen.

3.3.1 Diskreter Kurvenraum

Sei $n \geq 2$ die Anzahl der Punkte, durch die die Kurven bestimmt werden sollen. Für $1 \leq i \leq n$ sei

$$h_i := \frac{i-1}{n-1}$$

und $\tilde{\Theta}_j$ die FE Basisfunktion zum Gitter h_1, \dots, h_n und Knoten h_j . Ferner sei für $1 \leq j \leq n$ (wegen $h_1 = 0$ und $h_n = 1$ wohldefiniert)

$$\tilde{\Theta}_j^z : [0, 1] \rightarrow \mathbb{R}^2, x \mapsto \left(\tilde{\Theta}_j(x), 0 \right),$$

$$\tilde{\Theta}_j^T : [0, 1] \rightarrow \mathbb{R}^2, x \mapsto \left(0, \tilde{\Theta}_j(x) \right).$$

Für $T_i, z_i \in \mathbb{R}$ ($1 \leq i \leq n$) mit $z_i < z_{i+1}$ ($1 \leq i \leq n-1$) sei

$$\tilde{T} := \sum_{i=1}^n \left(z_i \tilde{\Theta}_i^z + T_i \tilde{\Theta}_i^T \right).$$

Dann gilt

$$\tilde{T}(h_j) = \sum_{i=1}^n (z_i(\delta_{ij}, 0) + T_i(0, \delta_{ij})) = (z_j, T_j),$$

d.h. \tilde{T} ist der Polygonzug (eine stückweise gerade Kurve) durch die Punkte $(z_1, T_1), \dots, (z_n, T_n)$ (in dieser Reihenfolge). Als Entsprechung von \mathcal{T}_M setzen wir $n = n_1 + n_2 + M$ und definieren

$$\tilde{\mathcal{T}}_M := \left\{ \sum_{i=1}^n \left(\tilde{z}_i \tilde{\Theta}_i^z + \tilde{T}_i \tilde{\Theta}_i^T \right) \left| \begin{array}{l} \exists z_{n_1} < Z_1 < Z_2 < \dots < Z_M < z_{l-n_2+1} \\ \tilde{z}_i = z_i \text{ für } i \in \{1, \dots, n_1\}, \\ \tilde{T}_i = T_{\text{ref}}(z_i) \text{ für } i \in \{1, \dots, n_1\} \\ \tilde{T}_{i+n_1} \in \mathbb{R}, \tilde{z}_{i+n_1} = Z_i \text{ für } i \in \{1, \dots, M\}, \\ \tilde{z}_{i+n_1+M} = z_{l-n_2+i} \text{ für } i \in \{1, \dots, n_2\}, \\ \tilde{T}_{i+n_1+M} = T_{\text{ref}}(z_{l-n_2+i}) \text{ für } i \in \{1, \dots, n_2\} \end{array} \right. \right\}.$$

Wie bei \mathcal{T}_M können wir ein Element $\tilde{T} \in \tilde{\mathcal{T}}_M$ mit den zwei Vektoren

$$\bar{T} = (T_1, \dots, T_M) \text{ und } \bar{Z} = (Z_1, \dots, Z_M)$$

identifizieren, wobei

$$T_i := \tilde{T}_{i+n_1} \text{ für } i \in \{1, \dots, M\},$$

$$Z_i := \tilde{z}_{i+n_1} \text{ für } i \in \{1, \dots, M\}.$$

Offensichtlich gibt es eine bijektive Abbildung φ von \mathcal{T}_M nach $\tilde{\mathcal{T}}_M$, die einem Element $T \in \mathcal{T}_M$ die Kurve $\tilde{T} \in \tilde{\mathcal{T}}_M$ zuordnet, die dem Graph der Abbildung von T entspricht. So können wir auch auf $\tilde{\mathcal{T}}_M$ unsere Energie definieren:

$$\tilde{E}_{\text{ges}} : \tilde{\mathcal{T}}_M \rightarrow \mathbb{R}, \tilde{T} \mapsto E_{\text{ges}}[\varphi^{-1}(\tilde{T})].$$

Ferner seien

$$\Theta_i^z := \tilde{\Theta}_{i+n_1}^z \text{ und } \Theta_i^T := \tilde{\Theta}_{i+n_1}^T$$

für $1 \leq i \leq M$. Mit diesen Definitionen gilt für $1 \leq j \leq M$

$$\begin{aligned} \langle E'_{\text{ges}}[T], \Theta_j \rangle &= \frac{\partial}{\partial T_j} (E_{\text{ges}}[T]) = \frac{\partial}{\partial T_j} \left(\tilde{E}_{\text{ges}}[\tilde{T}] \right) = \langle \tilde{E}'_{\text{ges}}[\tilde{T}], \Theta_j^T \rangle, \\ \langle E'_{\text{ges}}[T], -T' \Theta_j \rangle &= \frac{\partial}{\partial z_j} (E_{\text{ges}}[T]) = \frac{\partial}{\partial z_j} \left(\tilde{E}_{\text{ges}}[\tilde{T}] \right) = \langle \tilde{E}'_{\text{ges}}[\tilde{T}], \Theta_j^z \rangle. \end{aligned} \quad (3.1)$$

Der Raum der Testfunktionen ist somit gegeben durch

$$\tilde{\mathcal{T}}_0 := \left\{ \sum_{i=1}^M (z_i \Theta_i^z + T_i \Theta_i^T) \left| z_i, T_i \in \mathbb{R} (1 \leq i \leq M) \right. \right\},$$

bzw. durch $\mathcal{T}_0 := \varphi^{-1}(\tilde{\mathcal{T}}_0)$.

3.3.2 Zeitdiskretisierung mittels Differenzenquotient

Sei $\tilde{T}_0 \in \tilde{T}_M$ ein Startprofil und $\tau > 0$ die Zeitschrittweite. Ist \tilde{T} die Lösung des Gradientenflusses aus 2.2 auf dem diskreten Raum \tilde{T}_M mit Startwert \tilde{T}_0 und der Energie \tilde{E}_{ges} , so wollen wir diese Lösung an den Stellen $n\tau$ für $n \in \mathbb{N}_0$ durch $\tilde{T}^{(n)}$ über folgende Beziehung approximieren:

$$\frac{\partial}{\partial s} \tilde{T}(n\tau) \approx \frac{\tilde{T}((n+1)\tau) - \tilde{T}(n\tau)}{\tau}.$$

Seien

$$\bar{T}^n = (T_1^n, \dots, T_M^n) \text{ und } \bar{Z}^n = (Z_1^n, \dots, Z_M^n)$$

die $\tilde{T}^{(n)}$ entsprechenden Vektoren und $T^{(n)} := \varphi^{-1}(\tilde{T}^{(n)})$. Im Folgenden schreiben wir kurz g_n statt $g_{\tilde{T}^{(n)}}$. Ist ϑ eine Testfunktion, so gilt

$$\begin{aligned} g_n \left(\frac{\tilde{T}^{(n+1)} - \tilde{T}^{(n)}}{\tau}, \vartheta \right) &= - \langle \tilde{E}'_{\text{ges}}[\tilde{T}^{(n)}], \vartheta \rangle \\ \Leftrightarrow g_n \left(\sum_{i=1}^M ((Z_i^{n+1} - Z_i^n) \Theta_i^z + (T_i^{n+1} - T_i^n) \Theta_i^T), \vartheta \right) &= -\tau \langle \tilde{E}'_{\text{ges}}[\tilde{T}^{(n)}], \vartheta \rangle \\ g_n \left(\sum_{i=1}^M (Z_i^{n+1} \Theta_i^z + T_i^{n+1} \Theta_i^T), \vartheta \right) \\ \Leftrightarrow &= g_n \left(\sum_{i=1}^M (Z_i^n \Theta_i^z + T_i^n \Theta_i^T), \vartheta \right) - \tau \langle \tilde{E}'_{\text{ges}}[\tilde{T}^{(n)}], \vartheta \rangle \\ &= \sum_{i=1}^M (Z_i^{n+1} g_n(\Theta_i^z, \vartheta) + T_i^{n+1} g_n(\Theta_i^T, \vartheta)) \\ \Leftrightarrow &= \sum_{i=1}^M (Z_i^n g_n(\Theta_i^z, \vartheta) + T_i^n g_n(\Theta_i^T, \vartheta)) - \tau \langle \tilde{E}'_{\text{ges}}[\tilde{T}^{(n)}], \vartheta \rangle. \end{aligned}$$

Der n -te Zeitschritt muss also folgende Bedingung erfüllen:

$$A^n \begin{pmatrix} \bar{T}^{n+1} \\ \bar{Z}^{n+1} \end{pmatrix} = A^n \begin{pmatrix} \bar{T}^n \\ \bar{Z}^n \end{pmatrix} - \tau E,$$

wobei

$$E := \begin{pmatrix} \bar{E}^1 \\ \bar{E}^2 \end{pmatrix}, \quad A^n := \begin{pmatrix} A_1^n & A_2^n \\ A_2^n & A_3^n \end{pmatrix}$$

und (wegen (3.1))

$$\begin{aligned} \bar{E}^1 &:= \left(\langle E'_{\text{ges}}[T^{(n)}], \Theta_i \rangle \right)_{i=1}^M, & \bar{E}^2 &:= \left(\langle E'_{\text{ges}}[T^{(n)}], -(T^{(n)})' \Theta_i \rangle \right)_{i=1}^M, \\ A_1^n &= (g_n(\Theta_i^T, \Theta_j^T))_{i,j=1}^M, & A_2^n &= (g_n(\Theta_i^T, \Theta_j^z))_{i,j=1}^M, & A_3^n &= (g_n(\Theta_i^z, \Theta_j^z))_{i,j=1}^M. \end{aligned}$$

Es gilt also

$$\begin{pmatrix} \overline{T}^{n+1} \\ \overline{Z}^{n+1} \end{pmatrix} = \begin{pmatrix} \overline{T}^n \\ \overline{Z}^n \end{pmatrix} - \tau (A^n)^{-1} E$$

und ist die Metrik g_n so gewählt, dass $g_n(\Theta_i^T, \Theta_j^z) = 0$ für $1 \leq i, j \leq M$, so gilt $A_2 = 0$ und somit

$$\overline{T}^{n+1} = \overline{T}^n - \tau (A_1^n)^{-1} \overline{E}^1 \text{ und } \overline{Z}^{n+1} = \overline{Z}^n - \tau (A_3^n)^{-1} \overline{E}^2.$$

3.4 Die diskrete Metrik

Da wir im diskreten Ansatzraum Randwerte vorgeben, wird bei der Metrik nur noch der Teil der Kurven betrachtet, der variabel ist. Wegen 2.3.ii setzen also für $v, w \in C^2([z_0, z_{\max}], \mathbb{R})$

$$g_{T^{(n)},1}(v, w) = \frac{1}{L(\Gamma^{(n)})} \int_{z_{n_1}}^{z_{l-n_2+1}} v(z) \cdot w(z) \sqrt{1 + (T^{(n)})'(z)^2} dz,$$

$$g_{T^{(n)},2}(v, w) = L(\Gamma^{(n)}) \int_{z_{n_1}}^{z_{l-n_2+1}} \frac{v'(z)w'(z)}{\sqrt{1 + (T^{(n)})'(z)^2}} dz,$$

wobei

$$\Gamma^{(n)} = \tilde{T}^{(n)} \cap [z_{n_1}, z_{l-n_2+1}] \times \mathbb{R}.$$

Zu beachten ist auch, dass die Metrik im n -ten Zeitschritt abhängig von $\tilde{T}^{(n)}$ ist. Nun benötigen wir aber eine Metrik auf $\tilde{\mathcal{T}}_0 \subset C([0, 1], \mathbb{R}^2)$ und nicht auf $C^2([z_0, z_{\max}], \mathbb{R})$. Sei $n = n_1 + n_2 + M$ und sei $p : [z_0, z_{\max}] \rightarrow [0, 1]$ die Abbildung, die gegeben ist durch

$$p(z_i) = \frac{i-1}{n-1} \text{ für } i \in \{1, \dots, n_1\},$$

$$p(Z_i) = \frac{i-1+n_1}{n-1} \text{ für } i \in \{1, \dots, M\},$$

$$p(z_{l-n_2+i}) = \frac{i-1+n_1+M}{n-1} \text{ für } i \in \{1, \dots, n_2\},$$

$$p|_{[z_i, z_{i+1}]} \text{ ist affin für } i \in \{1, \dots, n_1-1, l-n_2+1, \dots, l-1\},$$

$$p|_{[Z_i^n, Z_{i+1}^n]} \text{ ist affin für } i \in \{1, \dots, M-1\},$$

$$p|_{[z_{n_1}, Z_1^n]}, p|_{[Z_M^n, z_{l-n_2+1}]} \text{ sind affin.}$$

Dann gilt $\tilde{\Theta}_{j+n_1} \circ p = \Theta_j$ für $j \in \{1, \dots, M\}$, wobei $\tilde{\Theta}_{j+n_1}$ die Basisfunktion aus Abschnitt 3.3.1 und Θ_j die FE-Basisfunktion zum Gitter $z_1 < \dots < z_{n_1} < Z_1^n < Z_2^n < \dots < Z_M^n < z_{l-n_2+1} < \dots < z_l$ und Knoten Z_j^n ist. Für $v = (v_1, v_2) \in \tilde{\mathcal{T}}_0$ sind dann $v_1 \circ p, v_2 \circ p : [z_0, z_{\max}] \rightarrow \mathbb{R}$ und wir setzen

$$g_{n,i}(v, w) := g_{T^{(n)},i}(v_1 \circ p, w_1 \circ p) + g_{T^{(n)},i}(v_2 \circ p, w_2 \circ p), \quad i = 1, 2.$$

(Dies ist möglich, da die Elemente aus $\tilde{\mathcal{T}}_0$ fast überall differenzierbar sind). Sei ferner

$$x_i = (x_{1,i}, x_{2,i}) = \begin{cases} (z_{n_1}, T_{\text{ref}}(z_{n_1})) & i = 0 \\ (Z_i^n, T_i^n) & 1 \leq i \leq M \\ (z_{l-n_2+1}, T_{\text{ref}}(z_{l-n_2+1})) & i = M+1 \end{cases}$$

Dann gilt mit

$$L_n := \sum_{j=0}^M |x_{j+1} - x_j| = L(\Gamma^{(n)})$$

für $1 \leq j \leq M$

$$\begin{aligned} g_{n,1}(\Theta_j^T, \Theta_j^T) &= g_{T^{(n)},1}(0, 0) + g_{T^{(n)},1}(\tilde{\Theta}_{j+n_1} \circ p, \tilde{\Theta}_{j+n_1} \circ p) \\ &= 0 + L_n \int_{z_{n_1}}^{z_{l-n_2+1}} (\tilde{\Theta}_{j+n_1} \circ p)(z)^2 \sqrt{1 + (T^{(n)})'(z)^2} dz \\ &= \frac{1}{L_n} \int_{z_{n_1}}^{z_{l-n_2+1}} \Theta_j(z)^2 \sqrt{1 + (T^{(n)})'(z)^2} dz \\ &= \frac{1}{L_n} \sum_{i=0}^M \int_{x_{1,i}}^{x_{1,i+1}} \Theta_j(z)^2 \sqrt{1 + (T^{(n)})'(z)^2} dz \\ &= \frac{1}{L_n} \sum_{i=0}^M \sqrt{1 + \left(\frac{x_{2,i+1} - x_{2,i}}{x_{1,i+1} - x_{1,i}} \right)^2} \int_{x_{1,i}}^{x_{1,i+1}} \Theta_j(z)^2 dz \\ &= \frac{1}{L_n} \sum_{i=j-1}^j \sqrt{1 + \left(\frac{x_{2,i+1} - x_{2,i}}{x_{1,i+1} - x_{1,i}} \right)^2} \frac{x_{1,i+1} - x_{1,i}}{3} \quad (\text{gilt nach 5.1.viii}) \\ &= \frac{1}{3L_n} \sum_{i=j-1}^j \sqrt{(x_{1,i+1} - x_{1,i})^2 + (x_{2,i+1} - x_{2,i})^2} \\ &= \frac{|x_j - x_{j-1}|}{3L_n} + \frac{|x_{j+1} - x_j|}{3L_n}, \end{aligned}$$

und analog zeigt man

$$g_{n,1}(\Theta_j^T, \Theta_{j+1}^T) = \frac{|x_{j+1} - x_j|}{6L_n}.$$

Ferner gilt offenbar für $1 \leq i, j \leq M$

$$g_{n,1}(\Theta_i^T, \Theta_j^z) = 0,$$

$$g_{n,1}(\Theta_i^T, \Theta_j^T) = g_{n,1}(\Theta_i^z, \Theta_j^z),$$

$$g_{n,1}(\Theta_i^T, \Theta_j^T) = 0 \text{ falls } j \notin \{i-1, i, i+1\}.$$

Für $1 \leq j \leq M$ ist

$$\begin{aligned} g_{n,2}(\Theta_j^T, \Theta_j^T) &= g_{T^{(n)},2}(0, 0) + g_{T^{(n)},2}(\tilde{\Theta}_{j+n_1} \circ p, \tilde{\Theta}_{j+n_1} \circ p) \\ &= 0 + L_n \int_{z_{n_1}}^{z_{l-n_2+1}} \frac{(\tilde{\Theta}_{j+n_1} \circ p)'(z)^2}{\sqrt{1 + (T^{(n)})'(z)^2}} dz = L_n \int_{z_{n_1}}^{z_{l-n_2+1}} \frac{\Theta_j'(z)^2}{\sqrt{1 + (T^{(n)})'(z)^2}} dz \\ &= L_n \sum_{i=0}^M \int_{x_{1,i}}^{x_{1,i+1}} \frac{\Theta_j'(z)^2}{\sqrt{1 + (T^{(n)})'(z)^2}} dz \end{aligned}$$

$$\begin{aligned}
 &= L_n \sum_{i=0}^M \left(\frac{(\delta_{j,i+1} - \delta_{j,i})^2}{(x_{1,i+1} - x_{1,i})^2} \Big/ \sqrt{1 + \left(\frac{x_{2,i+1} - x_{2,i}}{x_{1,i+1} - x_{1,i}} \right)^2} \right) \int_{x_{1,i}}^{x_{1,i+1}} dz \\
 &= L_n \sum_{i=j-1}^j \left(\frac{1}{(x_{1,i+1} - x_{1,i})^2} \Big/ \sqrt{1 + \left(\frac{x_{2,i+1} - x_{2,i}}{x_{1,i+1} - x_{1,i}} \right)^2} \right) (x_{1,i+1} - x_{1,i}) \\
 &= L_n \sum_{i=j-1}^j \frac{1}{\sqrt{(x_{1,i+1} - x_{1,i})^2 + (x_{2,i+1} - x_{2,i})^2}} = \frac{L_n}{|x_j - x_{j-1}|} + \frac{L_n}{|x_{j+1} - x_j|},
 \end{aligned}$$

und analog zeigt man

$$g_{n,2}(\Theta_j^T, \Theta_{j+1}^T) = -\frac{L_n}{|x_{j+1} - x_j|}.$$

Ferner gilt offenbar für $1 \leq i, j \leq M$

$$g_{n,2}(\Theta_i^T, \Theta_j^z) = 0,$$

$$g_{n,2}(\Theta_i^T, \Theta_j^T) = g_{n,2}(\Theta_i^z, \Theta_j^z),$$

$$g_{n,2}(\Theta_i^T, \Theta_j^T) = 0 \text{ falls } j \notin \{i-1, i, i+1\}.$$

Seien $\lambda_1, \lambda_2 > 0$. Dann definieren wir für $v, w \in \tilde{\mathcal{T}}_0$

$$g_n(v, w) = \lambda_1 g_{n,1}(v, w) + \lambda_2 g_{n,2}(v, w),$$

und erhalten so eine Metrik auf $\tilde{\mathcal{T}}_0$. Es gilt also $A_1 = A_3$, $A_2 = 0$ und A_1 hat Tridiagonalgestalt und ist schwach diagonaldominant, also invertierbar.

3.5 Unterschiede der Testfunktionen im diskreten und kontinuierlichen Fall

Im kontinuierlichen Fall werden die Testfunktionen als Elemente von $C^2([z_0, z_{\max}], \mathbb{R})$ aufgefasst, während im diskreten Fall Testfunktionen aus $C^2(\Gamma, \mathbb{R}^2)$ (bzw. aus $C^2([0, 1], \mathbb{R}^2)$) betrachtet werden. Dies ist jedoch in gewissem Sinne äquivalent: Ist $T \in \mathcal{T}$, $\Gamma = \Gamma[T]$ und $v \in C^2(\Gamma, \mathbb{R}^2)$, so dass $\Gamma + v$ ein Graph über $[z_0, z_{\max}]$ ist, so entspricht v einem Element aus $C^2([z_0, z_{\max}], \mathbb{R})$: Da $\Gamma + v := \{x + v(x) | x \in \Gamma\}$ Graph ist, ex. zu $z \in [z_0, z_{\max}]$ genau ein $x_z \in \Gamma$ mit $z = p_1(x_z + v(x_z))$ und v entspricht

$$w : [z_0, z_{\max}] \rightarrow \mathbb{R}, z \mapsto p_2(x_z + v(x_z)) - T(z),$$

denn:

3.5.i Lemma. Für $v \in C^2(\Gamma, \mathbb{R}^2)$ und w wie oben gilt

$$\Gamma[T] + v = \Gamma[T + w].$$

Beweis

„ \Rightarrow “: Sei $y \in \Gamma[T] + v$, dann existiert ein $x \in \Gamma$ mit $y = x + v(x)$. Ferner sei $z := p_1(y)$, dann gilt $z = p_1(x + v(x))$, d.h. $x = x_z$.

$$\begin{aligned} (z, T(z) + w(z)) &= (z, T(z) + p_2(x_z + v(x_z)) - T(z)) \\ &= (p_1(x + v(x)), p_2(x + v(x))) = x + v(x) = y, \end{aligned}$$

also $y = (z, T(z) + w(z)) \in \Gamma[T + w]$.

„ \Leftarrow “: Sei $y \in \Gamma[T + w]$. Dann existiert ein $z \in [z_0, z_{\max}]$ mit

$$\begin{aligned} y &= (z, T(z) + w(z)) = (z, p_2(x_z + v(x_z))) = (p_1(x_z + v(x_z)), p_2(x_z + v(x_z))) \\ &= x_z + v(x_z) \in \Gamma[T] + v. \end{aligned}$$

□

3.6 Numerische Lösung der Differentialgleichungen

Um Differentialgleichungen (wie (2.2) und (2.15)) zu lösen, benötigen wir ein geeignetes numerisches Verfahren. Eine Vielzahl solcher Verfahren findet man in [27]. Wir werden hier ein sog. Einschrittverfahren verwenden, um gewöhnliche Differentialgleichungen der Form

$$y'(x) = f(x, y(x)),$$

$$y(a) = y_0$$

auf einem Intervall $[a, b]$ zu lösen. Dazu benutzen wir

$$y(t+h) = y(t) + \int_0^1 \frac{\partial}{\partial x} (y(t+xh)) dx = y(t) + h \int_0^1 y'(t+xh) dx.$$

Um das Integral numerisch auszuwerten, benutzen wir eine Quadraturformel, die Keplersche Fassregel (die dadurch entsteht, dass der Integrand an den Stellen $0, \frac{1}{2}$ und 1 durch ein quadratisches Polynom interpoliert und dieses dann exakt integriert wird). Damit erhalten wir

$$\begin{aligned} \int_0^1 y'(t+xh) dx &\approx \frac{1}{6} (y'(t) + 4y'(t + \frac{h}{2}) + y'(t+h)) \\ &= \frac{1}{6} (f(t, y(t)) + 4f(t + \frac{h}{2}, y(t + \frac{h}{2})) + f(t+h, y(t+h))). \end{aligned}$$

Ferner gilt

$$y(t + \frac{h}{2}) = y(t) + \frac{h}{2}y'(t) + O(h^2)$$

und

$$\begin{aligned} y(t + \frac{h}{2}) &= y(t) + \frac{h}{2}y'(\xi) \quad (\text{für ein } \xi \in [t, t + \frac{h}{2}]) \\ &= y(t) + \frac{h}{2}y'(t + \frac{h}{2}) + O(h^2), \end{aligned}$$

daher folgt

$$\begin{aligned} 4f\left(t + \frac{h}{2}, y\left(t + \frac{h}{2}\right)\right) &= 2f\left(t + \frac{h}{2}, y\left(t + \frac{h}{2}\right)\right) + 2f\left(t + \frac{h}{2}, y\left(t + \frac{h}{2}\right)\right) \\ &\approx 2f\left(t + \frac{h}{2}, y(t) + \frac{h}{2}y'(t)\right) + 2f\left(t + \frac{h}{2}, y(t) + \frac{h}{2}y'(t + \frac{h}{2})\right). \end{aligned}$$

Deswegen setzen wir

$$\begin{aligned} k_1 &= f(t, y(t)) \\ k_2 &= f\left(t + \frac{h}{2}, y(t) + \frac{h}{2}k_1\right) \\ k_3 &= f\left(t + \frac{h}{2}, y(t) + \frac{h}{2}k_2\right) \\ k_4 &= f(t + h, y(t) + hk_3) \\ y(t + h) &= y(t) + \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4). \end{aligned}$$

Das so erhaltenen Verfahren heißt Runge-Kutta-Verfahren. Für den etwas spezielleren Typ von gewöhnliche Differentialgleichungen der Form

$$y'(x) = b(x) + c(x) \cdot y(x),$$

$$y(z_l) = y_0$$

auf $[z_1, z_l]$ für $b, c : [z_1, z_l] \rightarrow \mathbb{R}$, ist es implementiert. Gegenüber obiger Überlegung ist noch zu beachten, dass in unseren Differentialgleichungen der Anfangswert nicht am linken, sondern am rechten Intervall-Endpunkt gegeben ist und die Schrittweite durch unser Gitter schon in gewissem Sinne vorbestimmt ist.

Dies führt zu folgendem Algorithmus: Sei

$$f(x, y) := b(x) + c(x) \cdot y.$$

Die numerische Lösung $(y(z_1), \dots, y(z_l))$ wird berechnet durch

$$y(z_l) := y_0$$

und für $k < l$ iterativ absteigend mit

$$\begin{aligned} h &= z_{k+1} - z_k, \\ k_1 &= f(z_{k+1}, y(z_{k+1})), \\ k_2 &= f\left(z_{k+1} - \frac{h}{2}, y(z_{k+1}) - \frac{h}{2}k_1\right), \\ k_3 &= f\left(z_{k+1} - \frac{h}{2}, y(z_{k+1}) - \frac{h}{2}k_2\right), \\ k_4 &= f(z_{k+1} - h, y(z_{k+1}) - hk_3), \\ y(z_k) &= y(z_{k+1}) - \frac{h}{6}(k_1 + 2k_2 + 2k_3 + k_4). \end{aligned}$$

Der Löser ist so implementiert, dass er für $N \in \mathbb{N}$ auch die Lösung an den Stellen

$$y\left(z_k + \frac{i}{N}(z_{k+1} - z_k)\right), 1 < i < N - 1$$

berechnet. In der Praxis muss in $N \geq 3$ gewählt werden, um auf einem typischen, gemessenen Datensatz die Strahlungstemperaturen mit (2.2) hinreichend genau zu berechnen, da bei einem solchen Datensatz nur 20-40 Knoten für einen Höhenbereich von 0 – 30km vorhanden sind.

3.7 Auswertung der benutzten Funktionen

Da der Runge-Kutta-Löser die Funktion f an beliebigen Stellen auswerten können muss, die Funktionen zum Teil aber nur als Punktauswertungen gegeben sind, müssen sie interpoliert werden.

3.7.1 Auswertung von p

Aus der Natur ist bekannt (siehe dazu z.B. [19]), dass der Druck mit der Höhe exponentiell abnimmt. Daher wird der Druck exponentiell interpoliert: Seien $z_1 < z_2$ und p_1, p_2 die zugehörigen Druckwerte. Gesucht ist nun eine Funktion $p : [z_1, z_2] \rightarrow \mathbb{R}$ mit $p(z_1) = p_1$ und $p(z_2) = p_2$ der Form

$$p(z) = a \exp(c(z - z_1)).$$

Es folgt sofort $a = p_1$ und

$$\begin{aligned} p_1 \exp(c(z_2 - z_1)) &= p_2 \\ \Rightarrow c(z_2 - z_1) &= \ln\left(\frac{p_2}{p_1}\right) \\ \Rightarrow c &= \frac{1}{z_2 - z_1} \ln\left(\frac{p_2}{p_1}\right). \end{aligned}$$

Es gilt also

$$p(z) = p_1 \exp\left(\frac{1}{z_2 - z_1} \ln\left(\frac{p_2}{p_1}\right) (z - z_1)\right).$$

3.7.2 Auswertung von q

q wird linear interpoliert, d.h. sind $z_1 < \dots < z_l$ die Gitterknoten und q_1, \dots, q_l die zugehörigen Feuchtwerte, so gilt

$$q(z) = \sum_{i=1}^l q_i \Theta_i(z),$$

wobei für $j \in \{1, \dots, l\}$ Θ_j die FE-Basisfunktion zum Knoten z_j sei.

3.7.3 Auswertung von σ_a

Für gegebene Profile T, q, p ist der Absorptions- und Emissionskoeffizient σ_a eine Funktion, die in unserem Modell nur von der Höhe und der Frequenz abhängt:

$$\sigma_a : [z_1, z_l] \times \mathbb{R} \rightarrow \mathbb{R}, (z, \nu) \mapsto \sigma_a(\nu, T(z), q(z), p(z)).$$

Unsere Approximation basiert auf dem Rosenkranz-Modell, das in [16] beschrieben wird. Für $z < 8\text{km}$ kann σ_a hinreichend genau durch ein Polynom 3. Grades mit von ν abhängigen Frequenzen approximiert werden:

$$\sigma_a(\nu, T, q, p) = a_0(\nu) + \sum_{i=1}^3 (a_{1i}(\nu)T^i + a_{2i}(\nu)q^i + a_{3i}(\nu)p^i).$$

Sei

$$k := \min\{i \in \{1, \dots, l\} | z_i \geq 8\text{km}\}.$$

Für $i \geq k$ sind die Werte von σ_a an den Gitterknoten und den verwendeten Frequenzen im Datensatz enthalten, $\sigma_{a,i}(\nu)$. Für $z \geq z_k$ wird σ_a linear aus diesen Werten interpoliert. Insgesamt gilt also

$$\sigma_a(\nu, z) = \begin{cases} a_0(\nu) + \sum_{i=1}^3 (a_{1i}(\nu)T^i(z) + a_{2i}(\nu)q^i(z) + a_{3i}(\nu)p^i(z)) & z < z_k \\ \sum_{i=k}^l \sigma_{a,i}(\nu)\Theta_i(z) & z \geq z_k. \end{cases}$$

Die eventuell vorhandene Unstetigkeitsstelle bei z_k macht numerisch keine Probleme.

3.8 Diskretisierung der Regularisierungsenergie

Das gesuchte Profil kann als Polygonzug Γ durch die paarweise verschiedenen Punkte x_1, \dots, x_n aufgefasst werden, wobei $n = n_1 + n_2 + M$ und

$$x_i = (u_i, v_i) = \begin{cases} (z_i, T_{\text{ref}}(z_i)) & 1 \leq i \leq n_1 \\ (Z_{i-n_1}, T_{i-n_1}) & n_1 < i \leq n_1 + M \\ (z_{i-(n_1+M)+(l-n_2)}, T_{\text{ref}}(z_{i-(n_1+M)+(l-n_2)})) & n_1 + M < i \leq n_1 + n_2 + M \end{cases}.$$

Die diskrete Krümmung H_i des Polygonzuges im Punkte x_i für $1 < i < n$ ist gegeben durch

$$H_i = \frac{\frac{x_{i+1}-x_i}{\|x_{i+1}-x_i\|} - \frac{x_i-x_{i-1}}{\|x_i-x_{i-1}\|}}{\frac{1}{2}(\|x_{i+1}-x_i\| + \|x_i-x_{i-1}\|)} =: H(x_{i-1}, x_i, x_{i+1}),$$

die diskrete Tangente T_i an den Polygonzug im Punkte x_i für $1 < i < n$ ist gegeben durch

$$T_i = \frac{x_{i+1} - x_{i-1}}{\|x_{i+1} - x_{i-1}\|},$$

und die diskrete Normale N_i an den Polygonzug im Punkte x_i für $1 < i < n$ ist gegeben durch

$$N_i = D^{90}T_i = \begin{pmatrix} 0 & -1 \\ 1 & 0 \end{pmatrix} \frac{x_{i+1} - x_{i-1}}{\|x_{i+1} - x_{i-1}\|} =: N(x_{i-1}, x_{i+1}).$$

Sei $k := \max(\{k \in \{1, \dots, M\} | T_k > T_{k-1}\}, 1) + n_1$, dann ist x_k der Knoten der oberen Inversion (falls eine Inversion im variablen Teil des Profils vorhanden ist, ansonsten ist x_k der unterste der variable Knoten. Deswegen ist die 1 im Maximum nötig). Ferner sei $j := M + n_1 + 1$, dann ist x_j der erste Knoten oberhalb von 8km Höhe. Dann entsprechen u_k , bzw. u_j h_1 , bzw. h_2 aus dem Abschnitt über die Regularisierung und wir erhalten als ersten Teil der diskreten reg. Energie

$$E_c[T] = \gamma \sum_{i=k+1}^{j-1} \frac{1}{2} (\|x_{i+1} - x_i\| + \|x_i - x_{i-1}\|) e^{\alpha H_i \cdot N_i}.$$

Die Diskretisierung des zweiten Teils der Regularisierungsenergie ist naheliegend: Da $T(z)$ stückweise affin ist, ist $T'(z)$ stückweise konstant. Es gilt also für $\hat{x}_i := \frac{u_{i+1}+u_i}{2}$

$$\begin{aligned} E_P[T] &= \beta \int_{h_1}^{h_2} \left(\frac{1}{p_l[T](z)} - 1 \right) dz = \beta \int_{u_k}^{u_j} \left(\frac{1}{1 - \operatorname{erf} \left(\frac{|T'(z) - \mu|_l}{\sqrt{2}\sigma} \right)} - 1 \right) dz \\ &= \sum_{i=k}^{j-1} (u_{i+1} - u_i) \left(\frac{1}{1 - \operatorname{erf} \left(\frac{|T'(\hat{x}_i) - \mu|_l}{\sqrt{2}\sigma} \right)} - 1 \right) \\ &= \sum_{i=k}^{j-1} (u_{i+1} - u_i) \left(\frac{1}{1 - \operatorname{erf} \left(\frac{1}{\sqrt{2}\sigma} \left| \frac{v_{i+1} - v_i}{u_{i+1} - u_i} - \mu \right|_l \right)} - 1 \right). \end{aligned}$$

3.9 Variation der diskreten Regularisierungsenergie

Um den Gradientenfluss anwenden zu können, muss noch die Variation der Regularisierungsenergie in Richtung der Testfunktionen Θ_i und $-T'\Theta_i$ ausgerechnet werden. Da die Regularisierungsenergie bereits voll diskret ist, gilt für $n_1 < i \leq n_1 + M$ (vgl. 3.2)

$$\langle E'_{\text{reg}}[T], \Theta_{i-n_1} \rangle = \frac{\partial}{\partial v_i} (E_{\text{reg}}[T]) \quad \text{und} \quad \langle E'_{\text{reg}}[T], -T'\Theta_{i-n_1} \rangle = \frac{\partial}{\partial u_i} (E_{\text{reg}}[T]).$$

Um die Ableitungen von $E_c[T]$ auszurechnen, führen wir folgende Hilfsfunktion ein:

$$\begin{aligned} f_c : (\mathbb{R}^2)^3 &\rightarrow \mathbb{R}, \\ \underbrace{((u_1, v_1), (u_2, v_2), (u_3, v_3))}_{=(x_1, x_2, x_3)} &\mapsto \frac{1}{2} (\|x_3 - x_2\| + \|x_2 - x_1\|) e^{\alpha H(x_1, x_2, x_3) \cdot N(x_1, x_3)}. \end{aligned}$$

Damit gilt

$$E_c[T] = \gamma \sum_{i=k+1}^{j-1} f_c(x_{i-1}, x_i, x_{i+1}).$$

Ferner seien

$$\begin{aligned} g_c(i) &:= f_c(x_{i-1}, x_i, x_{i+1}), \\ g_{c,0}(i) &:= \frac{\partial}{\partial u_1} f_c(x_{i-1}, x_i, x_{i+1}), \\ g_{c,1}(i) &:= \frac{\partial}{\partial u_2} f_c(x_{i-1}, x_i, x_{i+1}), \\ g_{c,2}(i) &:= \frac{\partial}{\partial u_3} f_c(x_{i-1}, x_i, x_{i+1}). \end{aligned}$$

Mit dieser Notation folgt

$$\frac{\partial}{\partial u_k} (E_c[T]) = g_{c,0}(k+1) \quad \text{und} \quad \frac{\partial}{\partial u_j} (E_c[T]) = g_{c,2}(j-1).$$

Für die restlichen Ableitungen muss eine Fallunterscheidung gemacht werden:

(i) Fall $k + 1 = j - 1$:

$$\frac{\partial}{\partial u_{k+1}} (E_c[T]) = g_{c,1}(k + 1),$$

(ii) Fall $k + 1 = j - 2$:

$$\begin{aligned} \frac{\partial}{\partial u_{k+1}} (E_c[T]) &= g_{c,1}(k + 1) + g_{c,0}(k + 2), \\ \frac{\partial}{\partial u_{j-1}} (E_c[T]) &= g_{c,2}(j - 2) + g_{c,1}(j - 1), \end{aligned}$$

(iii) Fall $k + 1 \leq j - 3$, $k + 1 < i < j - 1$:

$$\frac{\partial}{\partial u_i} (E_c[T]) = g_{c,2}(i - 1) + g_{c,1}(i) + g_{c,0}(i + 1).$$

Für die Ableitungen nach v_i kann man völlig analog verfahren. Im Programm ist Auswertung der Ableitungen exakt auf diese Weise implementiert, wobei die 6 Ableitungen der Hilfsfunktion f_c von Maple berechnet und als C-Code exportiert worden sind.

Ähnlich verfährt man, um die Ableitungen von $E_P[T]$ auszurechnen. Sei

$$\begin{aligned} f_P : (\mathbb{R}^2)^2 &\rightarrow \mathbb{R}, \\ \underbrace{((u_1, v_1), (u_2, v_2))}_{=(x_1, x_2)} &\mapsto (u_2 - u_1) \left(\frac{1}{1 - \operatorname{erf}\left(\frac{1}{\sqrt{2}\sigma} \left| \frac{v_2 - v_1}{u_2 - u_1} - \mu \right| \right)} - 1 \right). \end{aligned}$$

Damit gilt

$$E_P[T] = \beta \sum_{i=k}^{j-1} f_P(x_i, x_{i+1}).$$

Ferner seien

$$\begin{aligned} g_P(i) &:= f_P(x_i, x_{i+1}), \\ g_{P,0}(i) &:= \frac{\partial}{\partial u_1} f_P(x_i, x_{i+1}), \\ g_{P,1}(i) &:= \frac{\partial}{\partial u_2} f_P(x_i, x_{i+1}). \end{aligned}$$

Mit dieser Notation folgt

$$\frac{\partial}{\partial u_k} (E_P[T]) = g_{P,0}(k) \quad \text{und} \quad \frac{\partial}{\partial u_j} (E_P[T]) = g_{P,1}(j - 1).$$

Für $k < i < j - 1$ gilt

$$\frac{\partial}{\partial u_i} (E_P[T]) = g_{P,1}(i - 1) + g_{P,0}(i),$$

und für $i < k$ oder $i > j$

$$\frac{\partial}{\partial u_i} (E_P[T]) = 0.$$

Für die Ableitungen nach v_i kann man völlig analog verfahren. Im Programm ist Auswertung der Ableitungen exakt auf diese Weise implementiert, wobei die 4 Ableitungen der Hilfsfunktion f_P von Maple berechnet und als C-Code exportiert worden sind.

3.10 Der Algorithmus

$\Theta_j : [z_1, z_l] \rightarrow \mathbb{R}$ sei die FE Basisfunktion zum Gitter $z_1, z_{n_1}, Z_1, \dots, Z_M, z_{l-n_2}, z_l$ und Knoten Z_j , d.h. stückweise affin mit

$$\Theta_j(Z_j) = 1,$$

$$\Theta_j(z) = 0 \text{ für } z \in \{z_1, z_{n_1}, Z_1, \dots, Z_{j-1}, Z_{j+1}, \dots, Z_M, z_{l-n_2}, z_l\}.$$

3.10.1 Wahl des Startprofils

Es sind zwei verschiedene Startprofile implementiert. Das erste ist das Startprofil $T^{0,\text{lin}}$, das bestimmt ist durch

$$\bar{T}^0 = (T_1^0, \dots, T_M^0) \text{ und } \bar{Z}^0 = (Z_1^0, \dots, Z_M^0),$$

wobei

$$T_i^0 = \frac{i}{M+1} T_{\text{ref}}(z_{l-n_2}) + \frac{M+1-i}{M+1} T_{\text{ref}}(z_{n_1}),$$

$$Z_i^0 = \frac{i}{M+1} z_{l-n_2} + \frac{M+1-i}{M+1} z_{n_1},$$

d.h. das Startprofil wird linear aus dem untersten der oberen und dem obersten der unteren vorgegebenen Werte interpoliert. Das zweite ist das Startprofil $T^{0,\text{inv}}$ das bestimmt ist durch

$$\bar{T}^0 = (T_1^0, \dots, T_M^0) \text{ und } \bar{Z}^0 = (Z_1^0, \dots, Z_M^0),$$

wobei

$$T_i = \frac{i-2}{M} T_{\text{ref}}(z_{l-n_2}) + \frac{M+2-i}{M} T_{\text{ref}}(z_1) \quad \text{für } i = 2, \dots, M,$$

$$Z_i = \frac{i-2}{M} z_{l-n_2} + \frac{M+2-i}{M} 900\text{m}$$

$$T_1 = \frac{1}{M+2} T_{\text{ref}}(z_{l-n_2}) + \frac{M+1}{M+2} T_{\text{ref}}(z_{n_1}),$$

$$Z_1 = \frac{1}{4} Z_2 + \frac{3}{4} z_{n_1}.$$

Damit hat $T^{0,\text{inv}}$ eine Inversion in einer Höhe von 900m, wobei die Temperatur an der Inversionshöhe gleich der Bodentemperatur ist. Der folgende Algorithmus wird für beide Startprofile durchgeführt und als Ergebnis wird das Profil mit der geringeren Energie ausgewählt, wobei nur der Teil der Energie, der gegeben ist durch das quadratische Kostenfunktional, betrachtet wird. Dieses Vorgehen ist sinnvoll, da die Energie viele lokale Minima besitzt, und daher der Gradientenfluss in einem unerwünschten lokalen Minimum enden kann, wenn das Startprofil zu weit vom gesuchten Minimum entfernt ist. Dabei bieten sich gerade diese beiden Startprofile an, da es in der Natur im wesentlichen näherungsweise lineare Temperaturprofile und solche mit Inversion gibt.

3.10.2 Berechnung der Variation

Für alle $i \in \{1, \dots, n\}, j \in \{1, \dots, M\}, k \in \{1, \dots, m\}$

(i) Löse

$$\begin{aligned} \frac{d}{dz} T_b(\nu_i, \theta_k, z) &= G(T_b(\nu_i, \theta_k, z), \nu_i, \theta_k, T^0(z), Q(z)), \\ T_b(\nu_i, \theta_k, z_l) &= T_{b0}, \end{aligned}$$

dann ist $T_b(\nu_i, \theta_k)[T^0] = T_b(\nu_i, \theta_k, z_1)$.

(ii) Löse (Variation für die Temperaturänderung)

$$\begin{aligned} \frac{d}{dz} \Delta(\nu_i, \theta_k, z) &= (\partial_T G(T_b(\nu_i, \theta_k, z), \nu_i, \theta_k, T^0(z), Q(z))) \Theta_j(z) + \\ &\quad (\partial_{T_b} G(T_b(\nu_i, \theta_k, z), \nu_i, \theta_k, T^0(z), Q(z))) \Delta(z), \\ \Delta(\nu_i, \theta_k, z_l) &= 0, \end{aligned}$$

dann ist $\langle T_b'(\nu_i, \theta_k)[T^0], \Theta_j(z) \rangle = \Delta(\nu_i, \theta_k, z_1)$.

(iii) Löse (Variation für die Höhenänderung)

$$\begin{aligned} \frac{d}{dz} \Delta(\nu_i, \theta_k, z) &= (\partial_T G(T_b(\nu_i, \theta_k, z), \nu_i, \theta_k, T^0(z), Q(z))) [-(T^0)'(z) \Theta_j(z)] + \\ &\quad (\partial_{T_b} G(T_b(\nu_i, \theta_k, z), \nu_i, \theta_k, T^0(z), Q(z))) \Delta(z), \\ \Delta(\nu_i, \theta_k, z_l) &= 0, \end{aligned}$$

dann ist $\langle T_b'(\nu_i, \theta_k)[T^0], -(T^0)'(z) \Theta_j(z) \rangle = \Delta(\nu_i, \theta_k, z_1)$.

Dann berechne für alle $j \in \{1, \dots, M\}$

$$\begin{aligned} \bar{E}_j^1 &= \langle E'[T^0], \Theta_i \rangle \\ &= - \sum_{i=1}^n \sum_{k=1}^m (T_{bmik} - T_b(\nu_i, \theta_k)[T^0]) \cdot \langle T_b'(\nu_i, \theta_k)[T^0], \Theta_j(z) \rangle, \\ \bar{E}_j^2 &= \langle E'[T^0], (T^0)'(z) \Theta_i \rangle \\ &= - \sum_{i=1}^n \sum_{k=1}^m (T_{bmik} - T_b(\nu_i, \theta_k)[T^0]) \cdot \langle T_b'(\nu_i, \theta_k)[T^0], -(T^0)'(z) \Theta_j(z) \rangle, \end{aligned}$$

$$\bar{E}_{\text{reg},j}^1 = \langle E'_{\text{reg}}[T], \Theta_i \rangle = \frac{\partial}{\partial \nu_{i+n_1}} (E_{\text{reg}}[T]),$$

$$\bar{E}_{\text{reg},j}^2 = \langle E'_{\text{reg}}[T], -T' \Theta_i \rangle = \frac{\partial}{\partial u_{i+n_1}} (E_{\text{reg}}[T]).$$

Sei $\bar{E}^i = (\bar{E}_1^i + \bar{E}_{\text{reg},1}^i, \dots, \bar{E}_M^i + \bar{E}_{\text{reg},M}^i)$.

3.10.3 Zeitschritt

Nun wird T^1 bestimmt durch

$$\bar{T}^1 = \bar{T}^0 - \tau_1 A^{-1} \bar{E}^1 \quad \text{und} \quad \bar{Z}^1 = \bar{Z}^0 - \tau_2 A^{-1} \bar{E}^2.$$

Hierbei ist τ_i , $i = 1, 2$ und A^{-1} wie im nächsten Punkt beschrieben.

3.10.4 Wahl von A

Zur Zeit implementiert ist die in 3.4 beschriebene geometrische Metrik. Wahlweise kann aber auch die folgende einfache Glättung gewählt werden: $A^{-1}(V_1, \dots, V_M) = B^s(V_1, \dots, V_M)$ mit

$$(B(V_1, \dots, V_M))_j = \begin{cases} \frac{1}{2\lambda} V_{j-1} + \frac{\lambda-1}{\lambda} V_j + \frac{1}{2\lambda} V_{j+1} & 1 < j < M \\ \frac{\lambda-1}{\lambda} V_1 + \frac{1}{\lambda} V_2 & j = 1 \\ \frac{1}{\lambda} V_{M-1} + \frac{\lambda-1}{\lambda} V_M & j = M \end{cases}$$

für $\lambda > 0$ und $s \in \mathbb{N}$.

3.10.5 Zeitschrittweitensteuerung

Dass die Zeitschrittweite nicht beliebig gewählt werden kann, sieht man am folgenden leichten Beispiel: Sei $E : \mathbb{R} \rightarrow \mathbb{R}, x \mapsto x^2$. Die Diskretisierung des Gradientenflusses

$$\begin{aligned} \frac{d}{dt} x(t) &= -\text{grad } E(x(t)), \\ x(0) &= x_0. \end{aligned}$$

mit Schrittweite $\tau = 1$ analog zu (3.3) führt zu

$$x_{k+1} = x_k - \tau \text{grad } E(x_k) = x_k - 2x_k = -x_k \quad \text{für } k \in \mathbb{N}.$$

Also gilt insgesamt

$$x_k = (-1)^k x_0,$$

diese Folge konvergiert für keinen Startwert $x_0 \neq 0$ gegen das Minimum $x = 0$, sie konvergiert sogar überhaupt nicht! Daher ist es notwendig, eine Zeitschrittweitensteuerung zu benutzen. Wir verwenden einen bewährten Ansatz, der in [18] zu finden ist: Seien $\sigma \in]0, \frac{1}{2}[$ und für $t \in \mathbb{R}$ sei T_t^0 das Profil, das bestimmt ist durch

$$\bar{T}^0 - t A^{-1} \bar{E}^1 \quad \text{und} \quad \bar{Z}^0.$$

Ferner sei $f : \mathbb{R} \rightarrow \mathbb{R}, t \mapsto E_{\text{ges}}[T_t^0]$. Gesucht ist nun ein τ_1 , so dass das Verhältnis der Steigung der Sekante durch $f(0)$ und $f(\tau_1)$ und der Tangentensteigung von f in 0 gilt:

$$\frac{\text{Sekantensteigung}}{\text{Tangentensteigung}} \geq \sigma,$$

siehe hierzu Abbildung 3.3. Die Tangentensteigung ist offensichtlich gegeben durch

$$f'(0) = ((-A^{-1}\bar{E}^1, 0), (\bar{E}^1, \bar{E}^2)) = -(A^{-1}\bar{E}^1, \bar{E}^1)$$

und ungleich Null, sofern f in Null kein Minimum hat, die Sekantensteigung durch

$$\frac{f(\tau_1) - f(0)}{\tau_1} = \frac{E_{\text{ges}}[T_{\tau_1}^0] - E_{\text{ges}}[T^0]}{\tau_1}.$$

Es muss also gelten

$$\frac{f(\tau_1) - f(0)}{\tau_1 f'(0)} = -\frac{E_{\text{ges}}[T_{\tau_1}^0] - E_{\text{ges}}[T^0]}{\tau_1 (A^{-1}\bar{E}^1, \bar{E}^1)} \geq \sigma.$$

Ist f genügend glatt, so gilt nach Taylor

$$f(\tau_1) = f(0) + \tau_1 f'(0) + O(\tau_1^2)$$

und es folgt

$$\frac{f(\tau_1) - f(0)}{\tau_1 f'(0)} = \frac{f(0) + \tau_1 f'(0) + O(\tau_1^2) - f(0)}{\tau_1 f'(0)} = 1 + O(\tau_1),$$

also ex. zu jedem $\sigma < 1$ ein $\tau > 0$, so dass für alle $0 < \tau_1 < \tau$ gilt

$$\frac{f(\tau_1) - f(0)}{\tau_1 f'(0)} \geq \sigma.$$

Die Existenz von τ_1 ist also (zumindest für glattes f) gesichert. Sei $\beta \in]0, 1[$. Die sog. **Armijo-Regel mit Aufweitung** bestimmt nun das kleinste $m \in \mathbb{Z}$, so dass gilt

$$\frac{f(\beta^m) - f(0)}{\beta^m f'(0)} \geq \sigma.$$

β^m ist also die größtmögliche Schrittweite (die eine Potenz von β ist), die die Bedingung erfüllt. Dies ist für $\sigma = \beta = \frac{1}{2}$ durch folgenden Algorithmus implementiert:

- Sei $\tau_1^0 = 1$ (oder $= \tau_1$ aus dem letzten Zeitschritt)

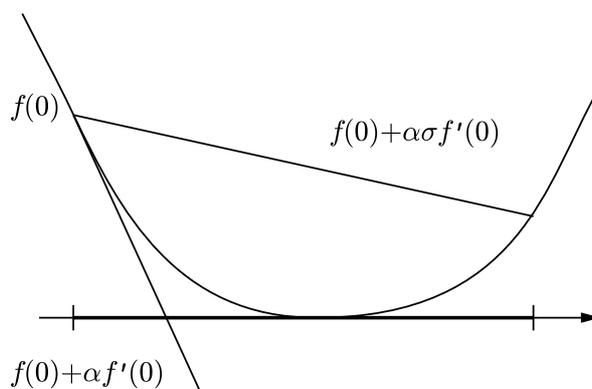


Abbildung 3.3: Man kann nicht erwarten, dass das Verhältnis der Steigungen größer oder gleich Eins ist, aber schon, dass es größer oder gleich $\sigma \in]0, 1[$ ist.

- Im Folgenden ist $T^{1,i}$ bestimmt durch

$$\overline{T}^{1,i} = \overline{T}^0 - \tau_1^i A^{-1} \overline{E}^1 \quad \text{und} \quad \overline{Z}^{1,i} = \overline{Z}^0.$$

- Check:

$$-\frac{E_{\text{ges}}[T^{1,i}] - E_{\text{ges}}[T^0]}{(A^{-1}\overline{E}^1, \overline{E}^1)} \geq \frac{1}{2}\tau_1^i$$

falls ja: $\tau_1^{i+1} = 2\tau_1^i$, solange bis nein, dann $\tau_1 = \tau_1^i$

falls nein: $\tau_1^{i+1} = \frac{1}{2}\tau_1^i$, solange bis ja, dann $\tau_1 = \tau_1^{i+1}$

- Bestimmung von τ_2 analog
- Nach der Bestimmung von τ_1 wird zusätzlich noch überprüft, ob $T_j^1 \geq 0$ für $1 \leq j \leq M$ gilt. Ist dies nicht der Fall, so wird τ_1 solange halbiert, bis dieses erfüllt es. Bei der Bestimmung von τ_2 wird analog dafür gesorgt, dass $z_{n_1} < Z_1^1 < Z_2^1 < \dots < Z_M^1 < z_{l-n_2+1}$ erfüllt ist. Dies ist nötig, um sicher zu stellen, dass T^1 ein Element von \mathcal{T}_M ist.

In der Praxis hat sich gezeigt, dass die Größenordnungen der möglichen Schrittweiten τ_1 und τ_2 meist verschieden sind, daher werden diese getrennt bestimmt und nicht nur eine Schrittweite benutzt.

3.10.6 Abbruchkriterium

Auf T^1 wendet man nun wieder den gerade beschriebenen Algorithmus an, um T^2 zu erhalten, und so weiter. So erhält man nun T^n für $n \in \mathbb{N}$. Da jedoch nur endlich viel Rechenzeit zur Verfügung steht, verwenden wir das folgende Abbruchkriterium: Falls im n -ten Zeitschritt

$$\begin{aligned} & (E_{\text{ges}}[T^n] - E_{\text{ges}}[T^{n+1}] > \frac{1}{1000} \vee n < 150) \\ & \wedge (E_{\text{ges}}[T^n] - E_{\text{ges}}[T^{n+1}] > \frac{1}{10000}) \\ & \wedge (E[T^{n+1}] > \frac{1}{10}) \end{aligned}$$

erfüllt ist, so wird ein $(n+1)$ -ter Zeitschritt durchgeführt, andernfalls wird das Verfahren abgebrochen. Dieses verschachtelte Verfahren hat sich in der Praxis bewährt, denn damit wird zum einen erreicht, dass das Verfahren abbricht, sobald die Energie nicht mehr sinkt (also ein lokales Minimum der Energie erreicht ist), zum anderen kommt es zum Abbruch, sobald die Strahlungstemperaturen des Retrievals so nah an den Messdaten sind, dass die Differenz im Bereich der Messungenauigkeit eines Mikrowellen-Radiometers liegt (dies ist der Fall, falls $E[T^{n+1}] \leq 0.1$). Ferner wird erreicht, dass nur mehr als 150 Zeitschritte gemacht werden, wenn die Energie noch merklich sinkt. Dadurch wird vermieden, dass zu viele Zeitschritte gemacht werden, falls sich das Retrieval schon in der Nähe eines lokalen Minimums befindet.

3.11 Vorgehen bei unbekanntem Druck

Ist der Druck an der Höhe z_1 und die Temperatur T_i an der Höhe z_i bekannt für $i = 1, \dots, l$ und Höhen $z_1 < z_2 < \dots < z_l$, so ist aus der Meteorologie bekannt, dass der Druck p_i an der Höhe z_i näherungsweise iterativ durch

$$p_i = p_{i-1} \exp\left(\frac{-2g(z_i - z_{i-1})}{R(t_i + t_{i-1})}\right) \quad \text{für } i = 2, \dots, l \quad (3.2)$$

bestimmt werden kann, wobei

$$R = 287.0 \frac{\text{J}}{\text{kgK}} \text{ (Gaskonstante für trockene Luft)}$$

$$g = 9.81 \frac{\text{m}}{\text{s}^2} \text{ (Fallbeschleunigung).}$$

Diesen Zusammenhang benutzen wir, um auf die Kenntnis des Druckprofils verzichten zu können, d.h. es muss nur noch der Druck am Boden bekannt sein. Dies ist in der Praxis kein Problem, da man ihn am Standort des Radiometers leicht messen kann. Dazu ändern wir im Algorithmus aus 3.10 nur einen Zeitschritt leicht ab: Im n -ten Zeitschritt bestimmen wir zu T^n mittels (3.2) das zugehörige Druckprofil an den Höhen $z_1, z_{n_1}, Z_1, \dots, Z_M, z_{l-n_2}, z_l$ und erhalten durch Interpolation wie in 3.7.1 ein komplettes Druckprofil p^n . Nun wird der Zeitschritt wie in 3.10 beschrieben mit $Q = (q, p^n)$ durchgeführt.

Im Programm ist dies realisiert und mittels der Option `useTheoreticPressure` aktivierbar (vgl. 5.2.2).

3.12 Validierung

Um die Richtigkeit bestimmter Programmteile zu prüfen, sind Testroutinen implementiert. Ruft man das Programm in einem dieser Testmodi auf, so wird der entsprechende Test aufgerufen und ausgegeben, ob der Test erfolgreich war oder nicht. Die Tests sind so angelegt, dass damit nicht nur Implementierungsfehler erkannt werden können, sondern auch Fehler in der theoretischen Herleitung der Variation der Energie (vgl. (2.4)).

3.12.1 Validierung der Störungsrechnung

Es gelte die Notation aus 3.2. Sei $\Delta[\vartheta]$ die Lösung von (2.15) zur Störung ϑ . Der Test basiert auf der Gleichung

$$\lim_{\epsilon \rightarrow 0} \frac{T_b(\nu_i, \theta_k)[T + \epsilon \vartheta] - T_b(\nu_i, \theta_k)[T]}{\epsilon} = \Delta[\vartheta](\nu_i, \theta_k, z_1).$$

Für kleines $\epsilon > 0$ und für alle $j \in \{1, \dots, M\}$ wird

$$\frac{T_b(\nu_i, \theta_k)[T + \epsilon \Theta_j] - T_b(\nu_i, \theta_k)[T]}{\epsilon}$$

und

$$\frac{T_b(\nu_i, \theta_k)[T \circ \Theta_j^\epsilon] - T_b(\nu_i, \theta_k)[T]}{\epsilon}$$

berechnet und mit der numerischen Lösung von

$$\Delta[\Theta_j](\nu_i, \theta_k, z_1), \text{ bzw. } \Delta[-T' \Theta_j](\nu_i, \theta_k, z_1)$$

verglichen. Ist die Differenz klein genug, so gilt der Test als bestanden.

3.12.2 Validierung der Variationsberechnung

Genau wie der vorige Test, basiert auch der nächste Test auf einem Differenzenquotienten, nämlich

$$\lim_{\epsilon \rightarrow 0} \frac{E_{\text{ges}}[T + \epsilon \vartheta] - E_{\text{ges}}[T]}{\epsilon} = \langle E'_{\text{ges}}[T], \vartheta \rangle.$$

Für kleines $\epsilon > 0$ und für alle $j \in \{1, \dots, M\}$ wird

$$\frac{E_{\text{ges}}[T + \epsilon \Theta_j] - E_{\text{ges}}[T]}{\epsilon}$$

und

$$\frac{E_{\text{ges}}[T \circ \Theta_j^\epsilon] - E_{\text{ges}}[T]}{\epsilon}$$

berechnet und mit der numerischen Lösung von

$$\langle E'_{\text{ges}}[T], \Theta_j \rangle, \text{ bzw. } \langle E'_{\text{ges}}[T], -T' \Theta_j \rangle$$

verglichen. Ist die Differenz klein genug, so gilt der Test als bestanden.

4 Ergebnisse

4.1 Auswertung des Blindtests

Um die Qualität des Verfahrens zu überprüfen, ist ein Blindtest auf zwei verschiedenen Datensätzen gemacht worden. Jeder der Datensätze enthält die in 5.2.1 angegebenen Informationen, wobei die Temperatur-, Druck- und Feuchteprofile aus Radiosondenaufstiegen stammen, und die Strahlungstemperaturen aus diesen Profilen mittels (2.2) an den in 2.1 genannten Frequenzen und einem Winkel von 0° simuliert wurden. Ferner ist der Bereich des Temperaturprofils, der gefunden werden soll, durch Platzhalterwerte unkenntlich gemacht worden.

Der erste Datensatz stammt aus Messungen in Cabauw (insgesamt 1494 Profile), der zweite aus Messungen in Essen (insgesamt 1407 Profile). Die Station in Cabauw (51.97° Nord, 4.93° Ost, Höhe über NN: -2m) ist der Stützpunkt des 4D-cloud Projekts (vgl. <http://www.meteo.uni-bonn.de/projekte/4d-clouds/bbc/>), die Station in Essen (51.40° Nord, 6.97° Ost, Höhe über NN: 153m) ist ein Radiosondenstützpunkt des Deutschen Wetterdienstes. Der entscheidende Unterschied zwischen den beiden Datensätzen ist, wie viel des Temperaturprofils am Boden bekannt ist: Bei den Daten aus Cabauw sind es die ersten 200m des Temperaturprofils (in der Regel zwei Messpunkte), während es in den Daten aus Essen nur die Bodentemperatur ist. Alle Rechnungen wurden mit den Parametern $\alpha = 1, \beta = 0.0001, \gamma = 0.0001, \lambda_1 = 0, \lambda_2 = 1$ durchgeführt (vgl. 3.4 und 3.8).

Zu gegebenem Datensatz seien n die Anzahl der Radiosondenaufstiege darin, T_{ref}^j das Referenzprofil und T_{ret}^j das Retrievalprofil zum j -ten Radiosondenaufstieg. Hierbei werden die Profile wie zuvor linear aus den gegebenen diskreten Werten interpoliert. Als systematischen Fehler s zur Höhe z bezeichnen wir das arithmetische Mittel der Differenzen der Retrievals zu den Referenzprofilen in der Höhe z , d.h.

$$s(z) := \frac{1}{n} \sum_{j=1}^n \left(T_{\text{ret}}^j(z) - T_{\text{ref}}^j(z) \right).$$

Als Maß für die Qualität der Retrievals wollen wir den RMS error (root mean square) der Differenzen der Retrievals zu den Referenzprofilen in der Höhe z benutzen, d.h.

$$e(z) := \sqrt{\frac{1}{n} \sum_{j=1}^n \left(T_{\text{ret}}^j(z) - T_{\text{ref}}^j(z) \right)^2}.$$

Zu den Höhen $100\text{m}, 200\text{m}, \dots, 8000\text{m}$ wurden e und s berechnet und sind auf den nächsten Seiten als Plot (linear interpoliert) zu sehen. Links immer Auswertungen der Daten aus Cabauw, rechts aus Essen.

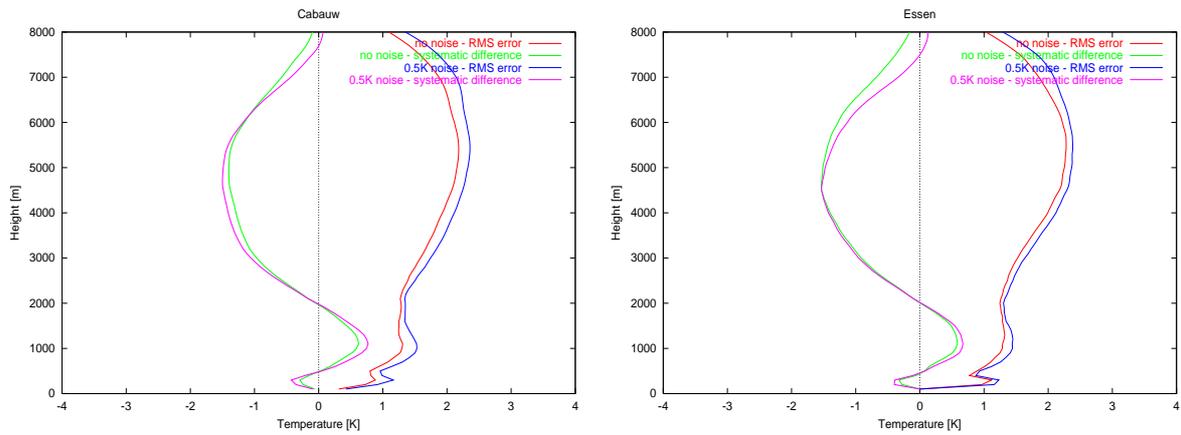


Abbildung 4.1: Einfluss von Instrumentenrauschen: Zur Simulation von Instrumentenrauschen wurde zu den $T_{bmi,k}$ eine in $[-0,5K, 0,5K]$ gleichverteilte Zufallsvariable addiert.

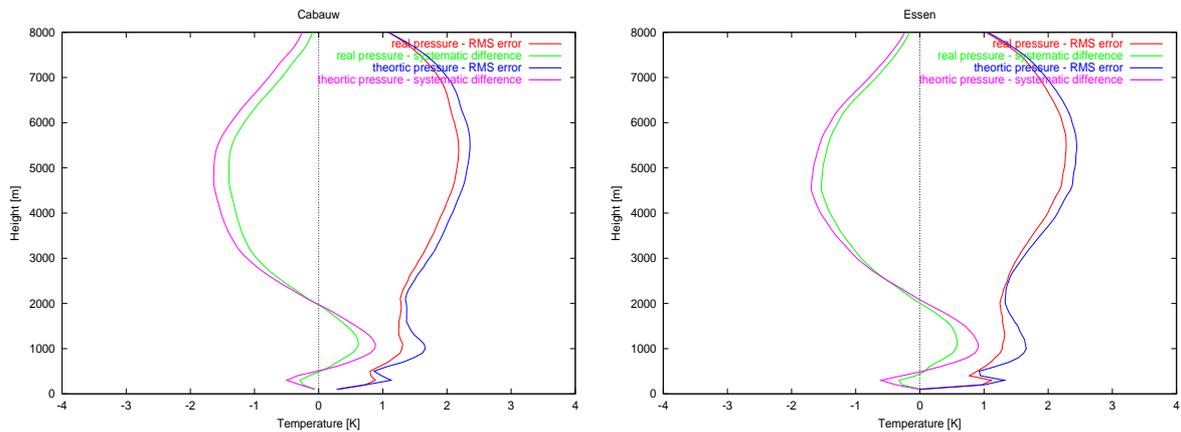


Abbildung 4.2: Einfluss von theoretischem Druck: Für diese Rechnung wurde statt dem realen Druck, der theoretische Druck, wie in 3.11 beschrieben, benutzt.

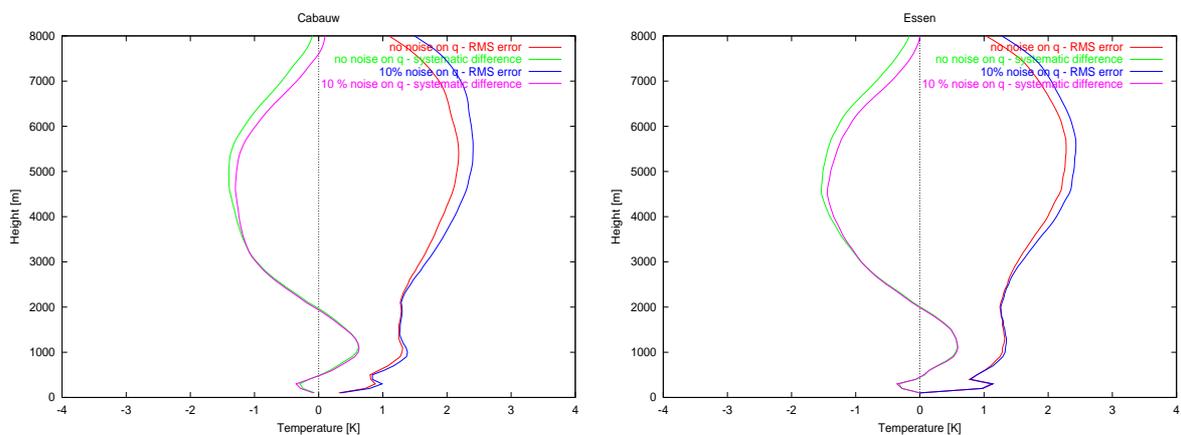


Abbildung 4.3: Einfluss der Feuchte: Das Feuchteprofile wurde mit 10% Gauß-Rauschen versehen.

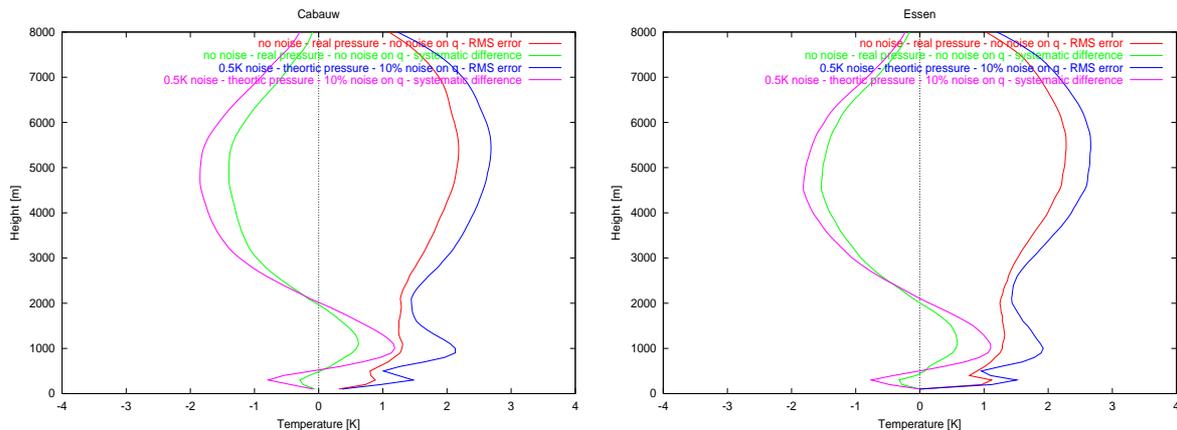


Abbildung 4.4: Einfluss von Instrumentenrauschen, theoretischem Druck und Feuchte

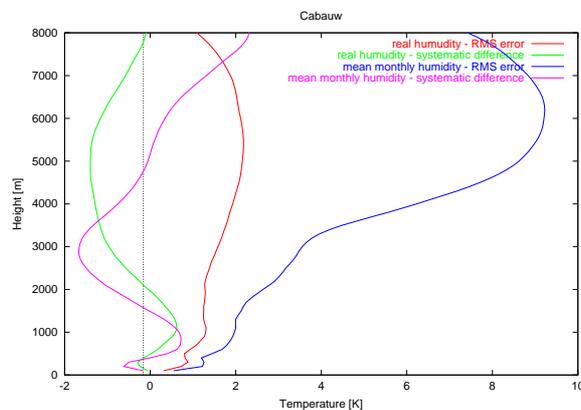


Abbildung 4.5: Einfluss der Feuchte: Das Retrieval ist nicht soweit von der Feuchte unabhängig, dass die Kenntnis der mittleren monatlichen Feuchte für ein akzeptables Retrieval ausreicht. Das aktuelle Feuchteprofil muss also (bis auf Rauschen) bekannt sein.

In Abbildung 4.1 ist zu sehen, dass Instrumentenrauschen von 0.5K, wie es in der Praxis mit einem gut kalibrierten Mikrowellen-Radiometer zu erreichen ist, keinen großen Einfluss auf die Qualität der Retrievals hat. Das Vorgeben des theoretischen Druckprofils statt des realen hingegen verändert die Retrievals stärker, wobei sich die Verschlechterungen auf die Höhen im Bereich von etwa 1000m konzentrieren (vgl. Abbildung 4.2). Die Genauigkeit bleibt zwar noch brauchbar, es lohnt sich jedoch sicherlich noch, nach Alternativen zu suchen, wie auf die Kenntnis des Druckprofils verzichtet werden kann. Eine Möglichkeit wäre z.B., ein möglichst aktuelles, bekanntest Druckprofil zu nutzen, etwa von der nächstgelegenen Radiosondenstation.

Abbildung 4.3 zeigt, dass geringes Rauschen im Feuchteprofile nahezu keinen Einfluss auf die Retrievals in einer Höhe bis etwa 2500m hat und auch in Höhen darüber die Retrievals nur geringfügig verschlechtert. Allerdings sind die Retrievals nicht unabhängig von der Feuchte: Gibt man statt des richtigen Feuchteprofils etwa das mittlere monatliche Feuchteprofil vor, so sind die Retrievals praktisch unbrauchbar (vgl. Abbildung 4.5). Damit wird deutlich, dass unser Verfahren von einem Einsatz „auf dem Feld“, d.h. vom Einsatz in der Praxis mit Messdaten von einem Mikrowellen-Radiometer, der Bodentemperatur und dem

Bodendruck, noch mindestens einen Schritt entfernt ist. Um dieses Problem zu lösen, könnte z.B. gleichzeitig zum Temperatur-Retrieval auf ähnliche Weise ein Retrieval für die Feuchte berechnet werden.

Die Kombination aller Störungen verschlechtert die Ergebnisse deutlich stärker als die Störungen es jeweils einzeln tun. Insbesondere bei einer Höhe von etwa 1000m nimmt die Genauigkeit ab (vgl. Abbildung 4.4). Abbildung 4.2 legt nahe, dass die Benutzung des theoretischen Drucks den größten Einfluss auf diese Ungenauigkeit hat, und dass speziell diese Höhe betroffen ist, liegt daran, dass in diesem Bereich die meisten Inversionen liegen, und diese (ohnehin schon schwierig zu findenden Temperaturphänomene) mit den Störungen schwieriger zu finden sind. Damit ergibt sich ein weiterer Grund, nach einer Alternative zur Nutzung des theoretischen Drucks zu suchen. Der Fehler bleibt allerdings durchaus vergleichbar mit rein statistischen Retrieval-Verfahren, wenn man für die Größe des Fehler bei statistischen Verfahren z.B. [13] heranzieht.

Ferner ist zu sehen, dass der Fehler generell ab einer Höhe von 3000m zu wachsen beginnt. Das kommt daher, dass der Einfluss der Temperatur auf die Strahlungstemperaturen am Boden mit der Höhe abnimmt. Dem zu widersprechen scheint, dass der Fehler ab einer Höhe von etwa 6000m wieder abzunehmen beginnt. Dieser Effekt entsteht, weil die richtige Temperatur an den Knoten, die über 8000m liegen, vorgegeben wird (bei den meisten Datensätzen liegt der erste Knoten, an dem die Temperatur wieder vorgegeben wird, bei ca. 8500m – 9000m).

Eine weitere Auffälligkeit ist, dass es kaum Unterschiede zwischen den Ergebnissen zu den Daten aus Cabauw und Essen gibt, d.h. das Vorgeben der Temperatur auf den ersten 200m ist nicht nötig. Ein Einsatz des Verfahrens ist (von der Feuchte-Abhängigkeit abgesehen) daher auch mit einem mobilen Mikrowellen-Radiometer denkbar, wenn kein 200m hoher Messmast verfügbar ist.

4.1.1 Auswertung für Inversionen

Wie schon erwähnt ist die Qualität der Retrievals mit den rein statistischen Ansätzen vergleichbar. Die Stärke unseres Verfahren liegt jedoch darin, auch im Falle von Inversionen gute Ergebnisse zu liefern. Daher entfernen wir aus den beiden Datensätzen diejenigen Profile, die keine Grenzschichtinversionen sind, und berechnen ansonsten auf die gleiche Weise den systematischen Fehler und den RMS error. Dabei wird ein Profil als Grenzschichtinversion gewertet, wenn es im Referenztemperaturprofil Knoten unterhalb vom 2000m gibt, bei denen die Temperatur mindestens 2K niedriger ist, als beim nächsthöheren Knoten. (Die 2K sollen dafür sorgen, dass sehr schwache Inversionen nicht beachtet werden, um die Auswertung nicht zu Gunsten unseres Verfahrens zu beeinflussen).

Um zu bestimmen, bei welchen Inversionen unser Verfahren besonders gute Ergebnisse liefert, benutzen wir noch ein alternatives Auswahlkriterium für Inversionen: Ein Profil wird als alternative Inversion gewertet, wenn es im Referenztemperaturprofil Knoten unterhalb vom 2000m gibt, bei denen die Temperatur größer ist als die Bodentemperatur (Der Begriff alternative Inversion wird hier nur als Abkürzung benutzt, es handelt sich dabei nicht um einen Begriff aus der Meteorologie).

Die Abbildungen 4.6, 4.7, 4.8 und 4.9 zeigen den Vergleich der Qualität der Retrievals bei allen Datensätzen und bei Grenzschichtinversionen mit unterschiedlichen Störungen, während die Abbildungen 4.10, 4.11, 4.12 und 4.13 den Vergleich der Qualität der Retrievals bei allen Datensätzen und bei alternativen Inversionen mit unterschiedlichen Störungen zeigen. Es ist zu erkennen, dass die Genauigkeit bei Grenzschichtinversionen zwar deutlich nachlässt, aber

dennoch in einer akzeptablen Größenordnung bleibt. Bei alternativen Inversionen bleibt die Genauigkeit sogar im Wesentlichen gleich, also hat unser Verfahren keinerlei Probleme damit, solche Inversionen aufzuspüren. In der Literatur scheint es bisher keine Untersuchung von statistischen Verfahren zu geben, die sich explizit mit der Qualität der Retrievals im Falle von Inversionsprofilen beschäftigen. Wir vermuten jedoch, dass die Genauigkeit dieser Verfahren stärker nachlässt, als bei unserem Verfahren, und hier ein wesentlicher Vorteil unseres Ansatzes liegt. Es ist also ratsam, noch eine solche Untersuchung durchzuführen, um in diesem Punkt Gewissheit zu haben.

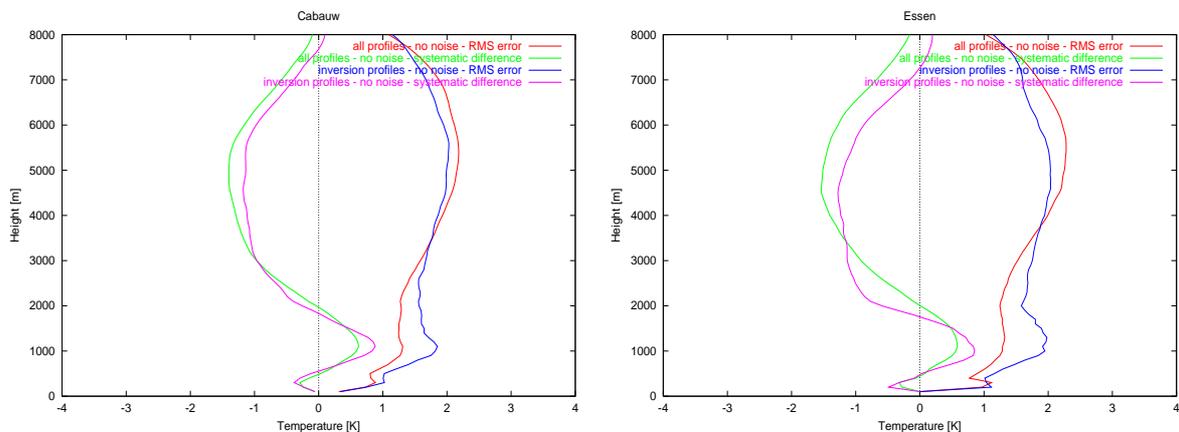


Abbildung 4.6: Vergleich der Qualität der Retrievals bei allen Datensätzen und bei Grenzschichtinversionen (Rechnungen ohne Rauschen)

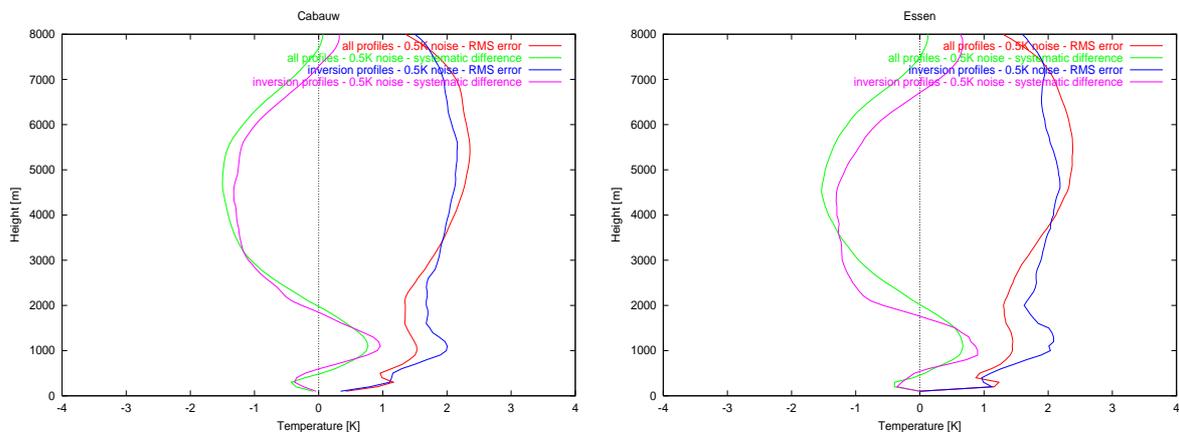


Abbildung 4.7: Vergleich der Qualität der Retrievals bei allen Datensätzen und bei Grenzschichtinversionen (0.5K Rauschen auf den Strahlungstemperaturen)

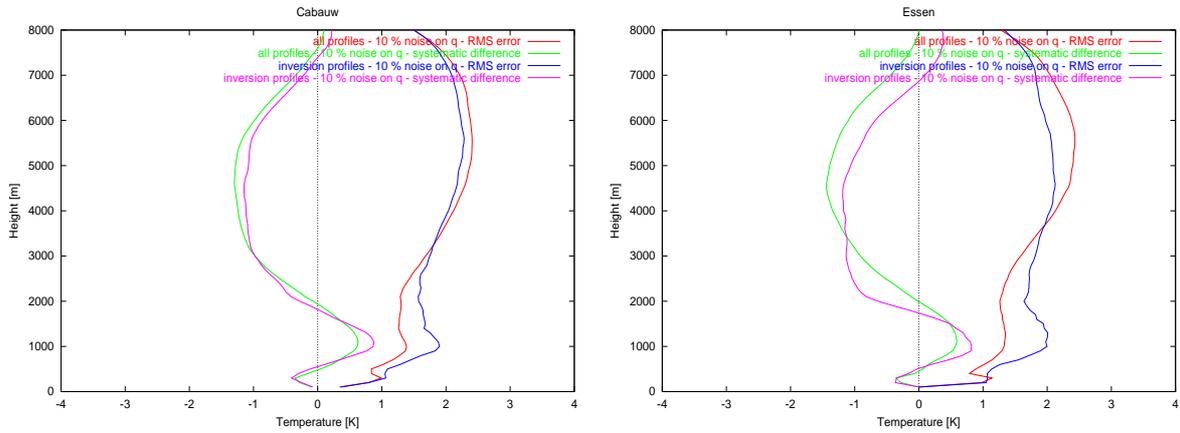


Abbildung 4.8: Vergleich der Qualität der Retrievals bei allen Datensätzen und bei Grenzschichtinversionen (10% Gauß-Rauschen auf der Feuchte)

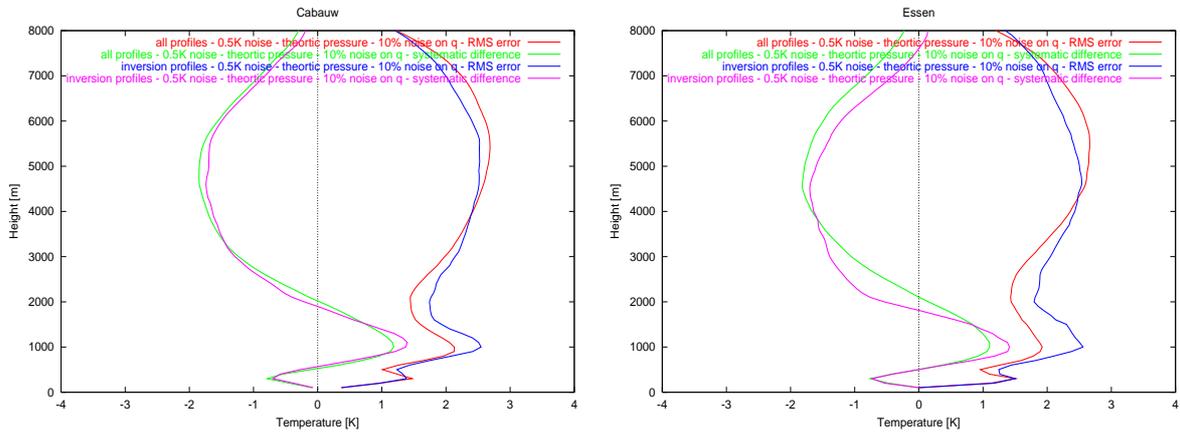


Abbildung 4.9: Vergleich der Qualität der Retrievals bei allen Datensätzen und bei Grenzschichtinversionen (Alle Störungen)

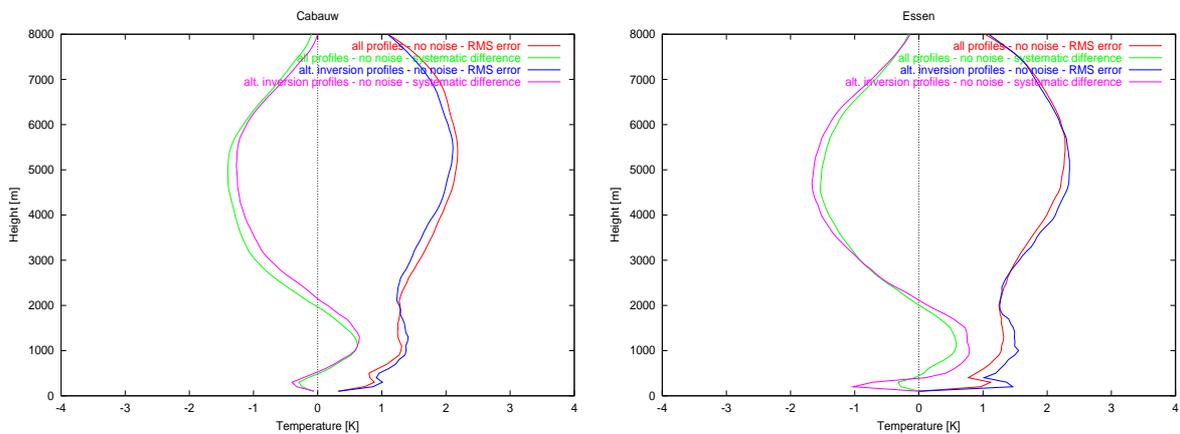


Abbildung 4.10: Vergleich der Qualität der Retrievals bei allen Datensätzen und bei alternativen Inversionen (Rechnungen ohne Rauschen)

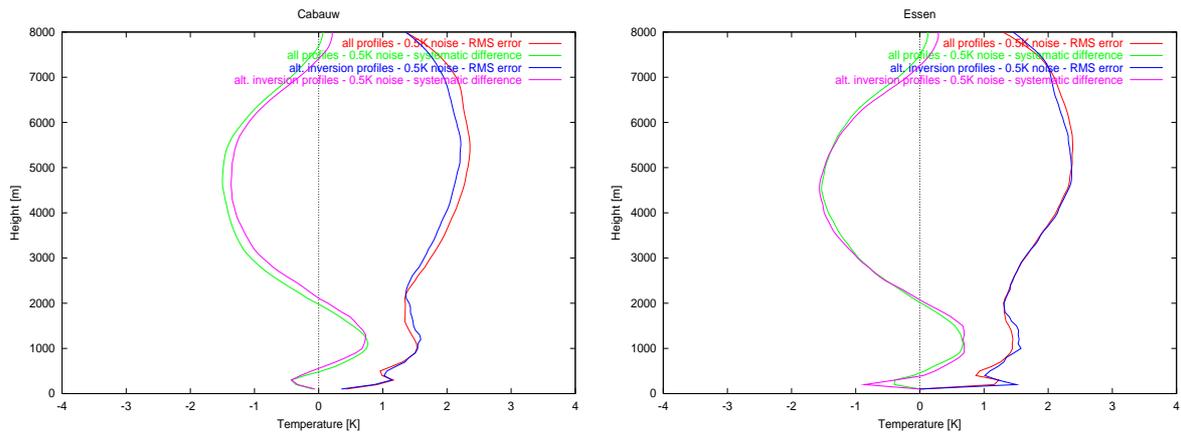


Abbildung 4.11: Vergleich der Qualität der Retrievals bei allen Datensätzen und bei alternativen Inversionen (0.5K Rauschen auf den Strahlungstemperaturen)

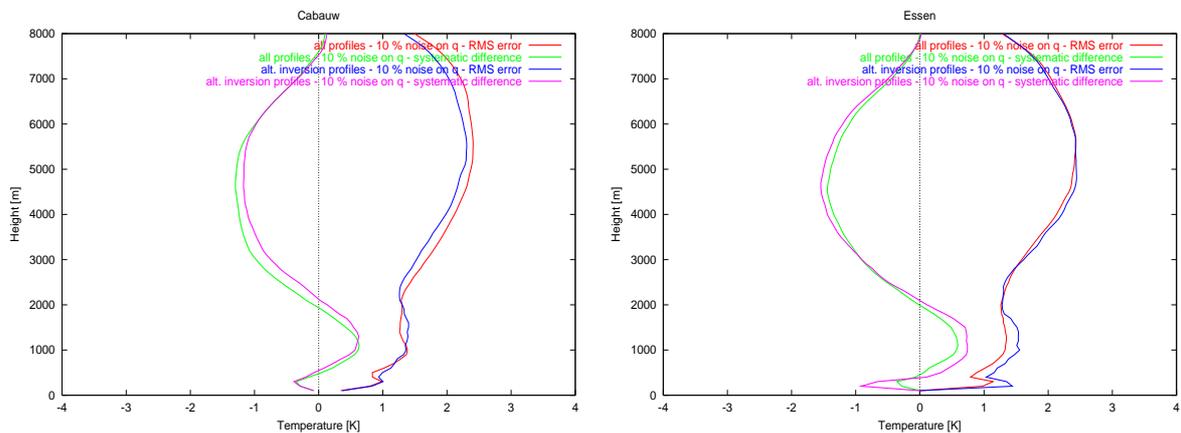


Abbildung 4.12: Vergleich der Qualität der Retrievals bei allen Datensätzen und bei alternativen Inversionen (10% Gauß-Rauschen auf der Feuchte)

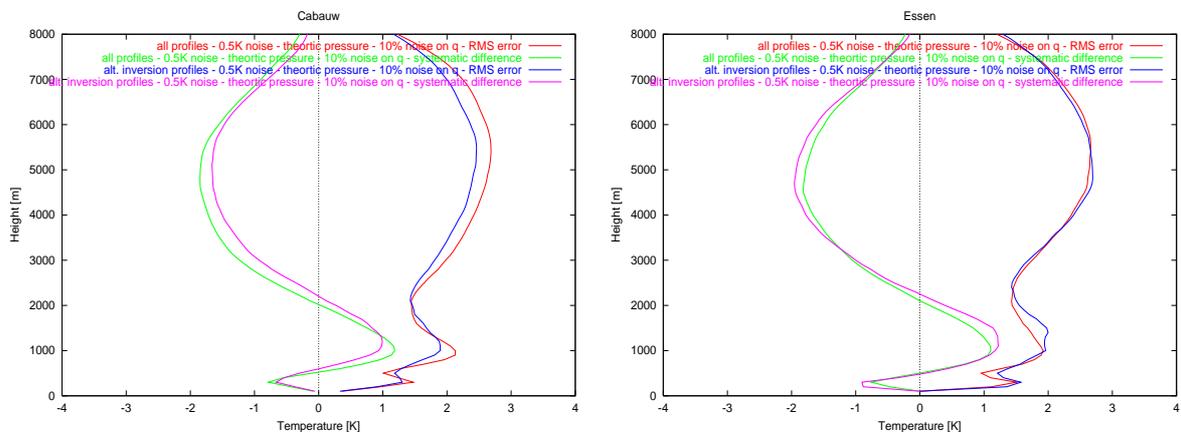


Abbildung 4.13: Vergleich der Qualität der Retrievals bei allen Datensätzen und bei alternativen Inversionen (Alle Störungen)

4.1.2 Ausgewählte Ergebnisse

Im Folgenden finden sich einige ausgewählte Ergebnisse des Blindtests auf den beiden Datensätzen, wobei die Ergebnisse von Rechnungen ohne Rauschen und mit realem Druckprofil stammen. Sie stehen beispielhaft für die Stärken und die Schwächen unseres Verfahrens. Zusammengefasst lässt sich folgendes beobachten:

- Profile ohne Inversion werden meist recht gut erkannt.
- Inversionen, deren Temperatur höher als die Bodentemperatur ist, werden oft sehr gut erkannt (vgl. Abschnitt 4.1.1.).
- Inversionen, deren Temperatur niedriger als die Bodentemperatur ist, werden meist nicht erkannt (vgl. Abschnitt 4.1.1.).
- Gelegentlich wird das falsche Startprofil gewählt.

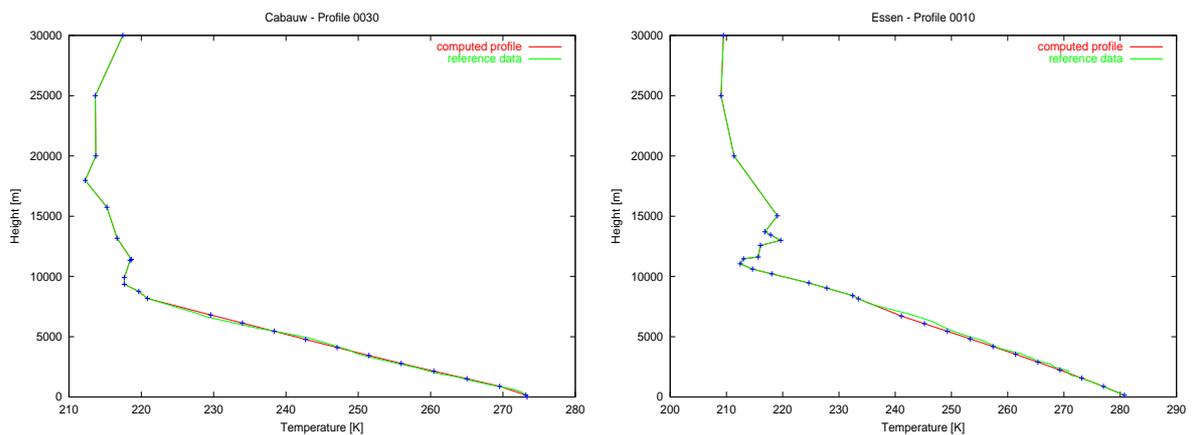


Abbildung 4.14: *Profile ohne Inversion*

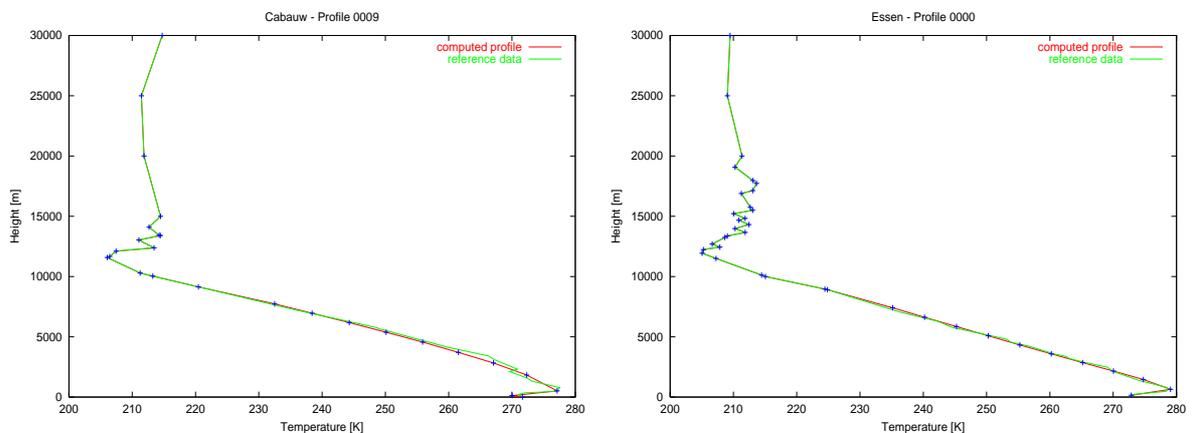


Abbildung 4.15: *Profile, mit einer Inversion, deren Temperatur höher als die Bodentemperatur ist*

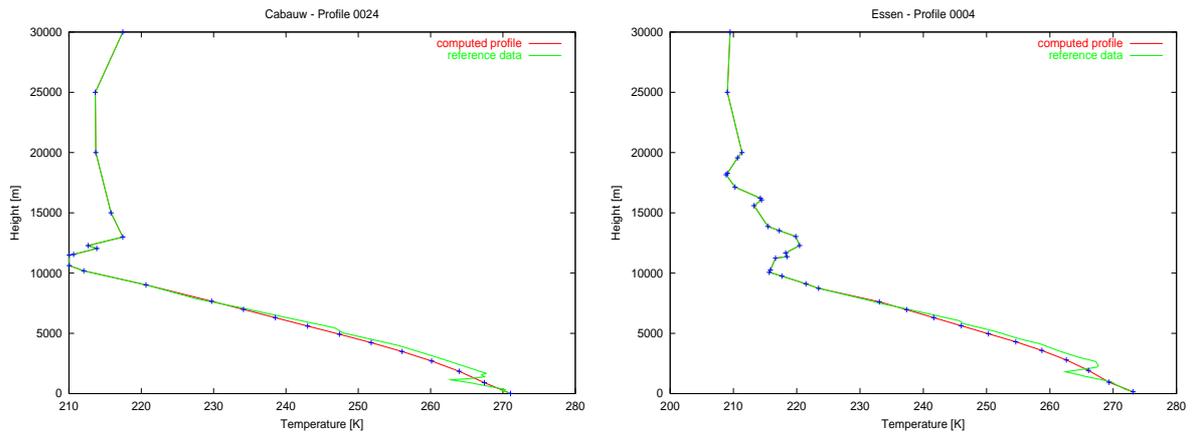


Abbildung 4.16: Profile, mit einer Inversion, deren Temperatur niedriger als die Bodentemperatur ist

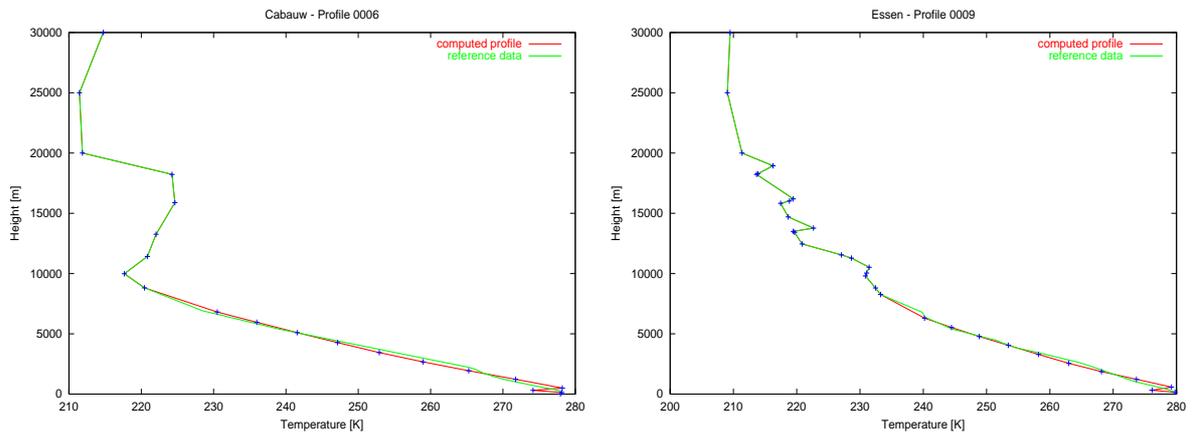


Abbildung 4.17: Profile, bei denen das falsche Startprofil gewählt wurde. Hier daran zu erkennen, dass das Retrieval eine Inversion hat, das Referenzprofil aber nicht.

4.2 Ausblick

Der Blindtest hat gezeigt, dass die Ergebnisse unseres Verfahrens gut genug sind, um in der Praxis Verwendung zu finden. Das Schwerwiegendste, das die praktische Nutzung unseres Verfahrens noch verhindert, ist die Abhängigkeit von der Kenntnis des Feuchteprofils. Dieses Problem könnte man lösen, indem man zuerst ein statistisches Feuchte-Retrieval-Verfahren benutzt und das dadurch erhaltene Feuchteprofil in unserem Verfahren verwendet. Eine andere Möglichkeit ist, wie schon erwähnt, gleichzeitig zum Temperatur-Retrieval auf ähnliche Weise ein Retrieval für die Feuchte zu berechnen. Für das Feuchte-Retrieval liegt es dabei nahe, als Messdaten nicht die Strahlungstemperaturen, die für das Temperatur-Retrieval verwendet werden, zu benutzen, sondern solche, die an der rechten Flanke der Wasserdampfabsorptionslinie liegen (vgl. [16]). Das zu Anfang schon erwähnte Radiometer MICCY kann in diesem Bereich an den Frequenzen von 22.235 bis 28.235GHz in 1GHz Schritten messen.

Ferner könnte man das benutzte Vorwärtsmodell gegen eines mit geringeren Einschränkungen ersetzen, da z.B. die Voraussetzung, dass der Himmel wolkenfrei ist, sehr stark und in

der Realität oft nicht erfüllt ist.

Außerdem ist der systematische Fehler der Ergebnisse noch relativ groß. Die Ursachen dafür sollte genauer untersucht und wenn möglich das Verfahren so verändert werden, dass der systematische Fehler reduziert wird.

Als letztes ist noch zu erwähnen, dass der Blindtest nur für Messdaten eines Winkels gemacht worden ist, obwohl der Algorithmus aus 3.10 die Benutzung mehrerer Winkel zulässt und auch vollständig implementiert ist. Die bisher gemachten Testrechnungen mit mehreren Winkeln zeigten keine sichtbaren Verbesserungen gegenüber den entsprechenden Rechnungen mit nur einem Winkel, daher wurde dies nicht im Blindtest verwendet. Aus [17] und [6] ist jedoch zu sehen, dass Messdaten an mehreren Winkeln mehr Informationen über das Temperaturprofil enthalten sollten. Somit lohnt es sich sicher, zu untersuchen, inwiefern es möglich ist, diese Informationen in unserem Ansatz zu nutzen.

5 Anhang

5.1 Wichtige Definitionen und Sätze

5.1.i Satz.

$$w(z) = w_0 \exp\left(-\int_z^b A(t)dt\right) - \int_z^b B(t) \exp\left(-\int_z^t A(s)ds\right) dt$$

ist Lösung des Anfangswertproblems

$$w'(z) = A(z)w(z) + B(z),$$

$$w(b) = w_0.$$

Beweis $w(b) = w_0$ ist offensichtlich. Wegen

$$\exp\left(-\int_z^t A(s)ds\right) = \exp\left(-\int_z^b A(s)ds\right) \cdot \exp\left(-\int_b^t A(s)ds\right)$$

folgt ferner

$$w(z) = \exp\left(-\int_z^b A(t)dt\right) \left(w_0 - \int_z^b B(t) \exp\left(\int_t^b A(s)ds\right) dt\right),$$

also gilt

$$\begin{aligned} w'(z) &= A(z) \exp\left(-\int_z^b A(t)dt\right) \left(w_0 - \int_z^b B(t) \exp\left(\int_t^b A(s)ds\right) dt\right) \\ &\quad + \exp\left(-\int_z^b A(t)dt\right) B(z) \exp\left(\int_z^b A(s)ds\right) \\ &= A(z)w(z) + B(z). \end{aligned}$$

□

5.1.ii Satz. Für $n \in \mathbb{N}$ sei $f_n \in C^1([a, b])$, so dass $(f_n(c)) \in \mathbb{R}^{\mathbb{N}}$ für ein $c \in [a, b]$ konvergiert und f'_n gleichmäßig auf $[a, b]$ konvergiert. Dann ex. $f \in C^1([a, b])$, so dass f_n gleichmäßig auf $[a, b]$ gegen f konvergiert und es gilt

$$f'(x) = \lim_{n \rightarrow \infty} f'_n(x) \text{ für alle } x \in [a, b].$$

Beweis Der Beweis ist in fast jedem Analysis Buch zu finden.

□

5.1.iii Lemma. Seien X, Y normierte \mathbb{R} -Vektorräume. Ferner seien $\mathcal{V} \subset X$ und $F : \mathcal{V} \rightarrow L(Y, Y')$. Dann ist F genau dann stetig, wenn

$$\forall v \in \mathcal{V} \quad \forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall \substack{w \in \mathcal{V} \\ \|v-w\|_X < \delta} \quad \forall \substack{y, z \in Y \\ y, z \neq 0} \quad |\langle (F(v) - F(w))y, z \rangle| < \epsilon \|y\|_Y \|z\|_Y.$$

Beweis

„ \Rightarrow “: Nach Definition der Stetigkeit gilt

$$\forall v \in \mathcal{V} \quad \forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall \substack{w \in \mathcal{V} \\ \|v-w\|_X < \delta} \quad \|F(v) - F(w)\|_{L(Y, Y')} < \epsilon.$$

Seien $y_0, z_0 \in Y, y_0, z_0 \neq 0$ beliebig. Es folgt

$$\begin{aligned} \epsilon > \|F(v) - F(w)\|_{L(Y, Y')} &= \sup_{\substack{y \in Y \\ y \neq 0}} \frac{\|(F(v) - F(w))y\|_{Y'}}{\|y\|_Y} \\ &= \sup_{\substack{y \in Y \\ y \neq 0}} \sup_{\substack{z \in Y \\ z \neq 0}} \frac{|\langle (F(v) - F(w))y, z \rangle|}{\|y\|_Y \|z\|_Y} \geq \frac{|\langle (F(v) - F(w))y_0, z_0 \rangle|}{\|y_0\|_Y \|z_0\|_Y} \end{aligned}$$

und damit die Behauptung.

„ \Leftarrow “: Aus der Voraussetzung folgt

$$\forall v \in \mathcal{V} \quad \forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall \substack{w \in \mathcal{V} \\ \|v-w\|_X < \delta} \quad \forall \substack{y, z \in Y \\ y, z \neq 0} \quad \frac{|\langle (F(v) - F(w))y, z \rangle|}{\|y\|_Y \|z\|_Y} < \frac{\epsilon}{2}.$$

Durch Bildung des Supremums über y und z folgt

$$\|F(v) - F(w)\|_{L(Y, Y')} = \sup_{\substack{y \in Y \\ y \neq 0}} \sup_{\substack{z \in Y \\ z \neq 0}} \frac{|\langle (F(v) - F(w))y, z \rangle|}{\|y\|_Y \|z\|_Y} \leq \frac{\epsilon}{2} < \epsilon.$$

□

5.1.iv Lemma. Seien X, Y normierte \mathbb{R} -Vektorräume. Ferner seien $\mathcal{V} \subset X$ und $F : \mathcal{V} \rightarrow Y'$. Dann ist F genau dann stetig, wenn

$$\forall v \in \mathcal{V} \quad \forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall \substack{w \in \mathcal{V} \\ \|v-w\|_X < \delta} \quad \forall \substack{\vartheta \in Y' \\ \vartheta \neq 0} \quad |\langle (F(v) - F(w)), \vartheta \rangle| < \epsilon \|\vartheta\|_{Y'}.$$

Beweis

„ \Rightarrow “: Nach Definition der Stetigkeit gilt

$$\forall v \in \mathcal{V} \quad \forall \epsilon > 0 \quad \exists \delta > 0 \quad \forall \substack{w \in \mathcal{V} \\ \|v-w\|_X < \delta} \quad \|F(v) - F(w)\|_{Y'} < \epsilon.$$

Sei $\vartheta_0 \in Y, \vartheta_0 \neq 0$ beliebig. Es folgt

$$\epsilon > \|F(v) - F(w)\|_{Y'} = \sup_{\substack{\vartheta \in Y' \\ \vartheta \neq 0}} \frac{|\langle F(v) - F(w), \vartheta \rangle|}{\|\vartheta\|_{Y'}} \geq \frac{|\langle F(v) - F(w), \vartheta_0 \rangle|}{\|\vartheta_0\|_{Y'}}$$

und damit die Behauptung.

„ \Leftarrow “: Aus der Voraussetzung folgt

$$\forall_{v \in \mathcal{V}} \quad \forall_{\epsilon > 0} \quad \exists_{\delta > 0} \quad \forall_{\substack{w \in \mathcal{V} \\ \|v-w\|_X < \delta}} \quad \forall_{\substack{\vartheta \in Y \\ \vartheta \neq 0}} \quad \frac{|\langle (F(v) - F(w)), \vartheta \rangle|}{\|\vartheta\|_Y} < \frac{\epsilon}{2}.$$

Durch Bildung des Supremums folgt

$$\|(F(v) - F(w))\|_{Y'} = \sup_{\substack{\vartheta \in Y \\ \vartheta \neq 0}} \frac{|\langle (F(v) - F(w)), \vartheta \rangle|}{\|\vartheta\|_Y} \leq \frac{\epsilon}{2} < \epsilon.$$

□

5.1.v Lemma. Seien X, Y normierte \mathbb{R} -Vektorräume. Ferner seien $\mathcal{V} \subset X$ und $F : \mathcal{V} \rightarrow Y'$. Dann ist F genau dann stetig in $v \in \mathcal{V}$, wenn

$$\forall_{\epsilon > 0} \quad \exists_{\delta > 0} \quad \forall_{\substack{w \in \mathcal{V} \\ \|v-w\|_X < \delta}} \quad \forall_{\substack{\vartheta \in Y \\ \vartheta \neq 0}} \quad |\langle (F(v) - F(w)), \vartheta \rangle| < \epsilon \|\vartheta\|_Y.$$

Beweis Analog zu vorigem Lemma. □

5.1.vi Definition. (erste Variation) Sei X ein \mathbb{R} -Vektorraum und $E : X \rightarrow \mathbb{R}$ ein Funktional. Dann ist für $x, v \in X$

$$\langle E'[x], v \rangle := \left. \frac{d}{d\epsilon} E[x + \epsilon v] \right|_{\epsilon=0}$$

die erste Variation von E in x in Richtung v . Damit gilt

$$E' : X' \rightarrow \mathbb{R}, x \mapsto (v \mapsto \langle E'[x], v \rangle).$$

5.1.vii Definition. Seien $z_1 < \dots < z_l$ ein Gitter und $j \in \{1, \dots, l\}$. Die Funktion $\Theta_j : [z_1, z_l] \rightarrow \mathbb{R}$, die bestimmt ist durch

$$\Theta_j(z_i) = \delta_{ij} \text{ für } 1 \leq i \leq l$$

und

$$\Theta_j|_{[z_i, z_{i+1}]} \text{ ist affin für } 1 \leq i \leq l-1,$$

heißt FE-Basisfunktion zum Gitter z_1, \dots, z_l und Knoten z_j .

5.1.viii Lemma. Mit den Voraussetzungen aus 5.1.vii gilt

$$\Theta_j(z) = \begin{cases} \frac{z_{j+1}-z}{z_{j+1}-z_j} & z \in [z_j, z_{j+1}] \\ \frac{z-z_{j-1}}{z_j-z_{j-1}} & z \in [z_{j-1}, z_j] \\ 0 & \text{sonst} \end{cases}$$

und

$$\int_{z_j}^{z_{j+1}} \Theta_j(z)^2 dz = \frac{z_{j+1} - z_j}{3},$$

$$\int_{z_{j-1}}^{z_j} \Theta_j(z)^2 dz = \frac{z_j - z_{j-1}}{3},$$

$$\int_{z_{j-1}}^{z_j} \Theta_j(z) \Theta_{j+1}(z) dz = \frac{z_{j+1} - z_j}{6}.$$

Beweis Durch die Bedingungen an Θ_j aus 5.1.vii ist Θ_j bekanntlich eindeutig bestimmt. Da die oben angegebene Funktion diese Bedingungen offensichtlich erfüllt, gilt die Gleichheit. Ferner gilt mit $h_j := z_{j+1} - z_j$

$$\begin{aligned} \int_{z_j}^{z_{j+1}} \Theta_j(z)^2 dz &= \int_{z_j}^{z_{j+1}} \left(\frac{z_{j+1} - z}{z_{j+1} - z_j} \right)^2 dz = \frac{1}{h_j^2} \int_{z_j}^{z_{j+1}} z_{j+1}^2 - 2z_{j+1}z + z^2 dz \\ &= \frac{1}{h_j^2} \left[z_{j+1}^2 z - z_{j+1} z^2 + \frac{1}{3} z^3 \right]_{z_j}^{z_{j+1}} \\ &= \frac{1}{h_j^2} \left[z_{j+1}^3 - z_{j+1}^3 + \frac{1}{3} z_{j+1}^3 - z_{j+1}^2 z_j + z_{j+1} z_j^2 - \frac{1}{3} z_j^3 \right] \\ &= \frac{1}{3h_j^2} [z_{j+1}^3 - 3z_{j+1}^2 z_j + 3z_{j+1} z_j^2 - z_j^3] = \frac{1}{3h_j^2} h_j^3 \\ &= \frac{z_{j+1} - z_j}{3}, \end{aligned}$$

$$\int_{z_{j-1}}^{z_j} \Theta_j(z)^2 dz = \int_{z_{j-1}}^{z_j} \left(\frac{z - z_{j-1}}{z_j - z_{j-1}} \right)^2 dz = - \int_{z_j}^{z_{j-1}} \left(\frac{z_{j-1} - z}{z_{j-1} - z_j} \right)^2 dz \stackrel{\text{s.o.}}{=} \frac{z_j - z_{j-1}}{3},$$

$$\begin{aligned} \int_{z_{j-1}}^{z_j} \Theta_j(z) \Theta_{j+1}(z) dz &= \int_{z_{j-1}}^{z_j} \left(\frac{z_{j+1} - z}{z_{j+1} - z_j} \right) \left(\frac{z - z_{j-1}}{z_{j+1} - z_j} \right) dz \\ &= \frac{1}{h_j^2} \int_{z_j}^{z_{j+1}} z_{j+1} z - z_{j+1} z_j - z^2 + z_j z dz \\ &= \frac{1}{h_j^2} \left[\frac{1}{2} z_{j+1} z^2 - z_{j+1} z_j z - \frac{1}{3} z^3 + \frac{1}{2} z_j z^2 \right]_{z_j}^{z_{j+1}} \\ &= \frac{1}{h_j^2} \left[\frac{1}{2} z_{j+1}^3 - z_{j+1}^2 z_j - \frac{1}{3} z_{j+1}^3 + \frac{1}{2} z_{j+1}^2 z_j - \left(\frac{1}{2} z_{j+1} z_j^2 - z_{j+1} z_j^2 - \frac{1}{3} z_j^3 + \frac{1}{2} z_j^3 \right) \right] \\ &= \frac{1}{h_j^2} \left[\frac{1}{6} z_{j+1}^3 - \frac{1}{2} z_{j+1}^2 z_j - \left(-\frac{1}{2} z_{j+1} z_j^2 + \frac{1}{6} z_j^3 \right) \right] \\ &= \frac{1}{6h_j^2} [z_{j+1}^3 - 3z_{j+1}^2 z_j + 3z_{j+1} z_j^2 - z_j^3] = \frac{1}{6h_j^2} h_j^3 = \frac{z_{j+1} - z_j}{6}. \end{aligned}$$

□

5.1.ix Satz. Sei $\Gamma \subset \mathbb{R}^2$ eine Kurve mit $c := L(\Gamma) > 0$ und $f : \Gamma \rightarrow \mathbb{R}$ integrierbar. Dann gilt

$$\int_{\Gamma} f(x) ds(x) = \int_{\frac{\Gamma}{c}} f(cx) c ds(x).$$

Beweis Sei $\phi : [0, 1] \rightarrow \Gamma$ eine Karte von Γ . Dann ist $\frac{\phi}{c}$ eine Karte von $\frac{\Gamma}{c}$ und es gilt

$$\begin{aligned} \int_{\Gamma} f(x) ds(x) &= \int_0^1 f(\phi(y)) |\phi'(y)| dy = \int_0^1 f\left(c \frac{\phi(y)}{c}\right) c \left| \frac{\phi'(y)}{c} \right| dy \\ &= \int_{\frac{\Gamma}{c}} f(cx) c ds(x). \end{aligned}$$

□

5.2 Programm-Dokumentation

Das Programm führt nach dem Start folgende Schritte durch:

- (i) Lesen der Parameter aus der Datei `meteo.par` (siehe 5.2.2).
- (ii) Anzeigen der Parameter in einer Liste, um sicherzustellen, dass alle Parameter richtig gelesen wurden.
- (iii) Einlesen des ersten noch nicht gelesenen Datensatzes aus der in der Parameter Datei angegebenen Eingabe-Datei (siehe 5.2.1).
- (iv) Berechnung des Retrievals entsprechend dem Algorithmus aus Abschnitt 3.10.
- (v) Schreiben des Retrievals in die `output.dat` (siehe 5.2.3).
- (vi) Wiederholen der letzten zwei Schritte, für jedes Feuchteprofil, das im Datensatz angegeben ist und auf dem gerechnet werden soll.
- (vii) Wiederholen der letzten vier Schritte, bis alle Datensätze abgearbeitet worden sind.

5.2.1 Dateiformat der Eingabe Datensätze

Das Programm liest Datensätze aus Text Dateien im folgendem Format, dabei ist die Notation genau wie in Kapitel 3:

- Integer: Jahr
- Integer: Monat
- Integer: Tag
- Integer: Stunde (diese Daten geben an, wann der Datensatz gemessen wurde)
- Integer: l = Anzahl der Höhenniveaus (n_2 bezeichne die Anzahl der Höhen, die größer oder gleich 8000m sind, soll nicht im Datensatz stehen, sondern wird berechnet)

- l double: z_1, \dots, z_l Höhenniveaus in m
- l double: Temperatur an den obigen Höhen, wobei die Werte an den Höhen, die echt größer 200m und echt kleiner 8000m sind, durch beliebige Zahlen ersetzt werden können, in K
- l double: Druck an den obigen Höhen in Pa
- l double: Absolute Feuchte an den obigen Höhen in $\frac{\text{kg}}{\text{m}^3}$ (Angabe von mehreren Profilen möglich)
- n_2 double: Wert von σ_a an obigen Höhen, die größer oder gleich 8000m sind, zur Frequenzen 50.8
- n_2 double: Wert von σ_a an obigen Höhen, die größer oder gleich 8000m sind, zur Frequenzen 51.8
- \vdots \vdots
- n_2 double: Wert von σ_a an obigen Höhen, die größer oder gleich 8000m sind, zur Frequenzen 58.8
- 9 double: Strahlungstemperaturen zu den Frequenzen 50.8, 51.8, 52.8, 53.8, 54.8, 55.8, 56.8, 57.8, 58.8 GHz

Eine Datei darf beliebig viele Datensätze enthalten, sie müssen nur in obigem Format hintereinander in der Datei stehen.

5.2.2 Format der Parameterdatei `meteo.par`

In der Parameterdatei `meteo.par` sind alle wichtigen Parameter gespeichert, die das Programm benötigt. Beim Start werden diese aus der Datei eingelesen, so dass die Parameter geändert werden können, ohne das Programm neu zu kompilieren. Dabei wird folgendes Format verwendet: In einer Zeile steht `Parameter-Name Parameter-Wert` z.B. `beta 0.0001`. Die einzelnen Parameter sind in Tabelle 5.1 aufgelistet.

5.2.3 Format der Ausgabedatei `output.dat`

Die berechneten Retrievals werden in die Textdatei `output.dat` geschrieben. Die einzelnen Retrievals stehen in der Datei hintereinander im folgendem Format:

- Integer: Jahr
- Integer: Monat
- Integer: Tag
- Integer: Stunde (diese Daten geben an, wann der Datensatz gemessen wurde)
- $9 \cdot \text{numangles}$ double: Strahlungstemperaturen zu denen ein Profil gefunden werden soll:

- 9 double: Strahlungstemperaturen zum Winkel 1 und den Frequenzen 50.8, 51.8, 52.8, 53.8, 54.8, 55.8, 56.8, 57.8, 58.8 GHz

⋮ ⋮

- 9 double: Strahlungstemperaturen zum Winkel `numangles` und den Frequenzen 50.8, 51.8, 52.8, 53.8, 54.8, 55.8, 56.8, 57.8, 58.8 GHz

- `numHumidityProfiles` Retrievals, jeweils bestehend aus

- Integer: Nummer des Startprofils, das für das Retrieval ausgewählt wurde (vgl. 3.11)

0: lineares Startprofil

1: Startprofil mit Inversion

- Integer: l = Anzahl der Höhengiveaus des Retrievals

- l double: z_1, \dots, z_l Höhengiveaus in m

- l double: ermittelte Temperatur an den obigen Höhen in K

- $9 \cdot \text{numangles}$ double: Strahlungstemperaturen, die das ermittelte Profil hat

- * 9 double: Strahlungstemperaturen zum Winkel 1 und den Frequenzen 50.8, 51.8, 52.8, 53.8, 54.8, 55.8, 56.8, 57.8, 58.8 GHz

⋮ ⋮

- * 9 double: Strahlungstemperaturen zum Winkel `numangles` und den Frequenzen 50.8, 51.8, 52.8, 53.8, 54.8, 55.8, 56.8, 57.8, 58.8 GHz

Name	Beschreibung
<code>M</code>	Anzahl der Freiheitsgrade/2
<code>numangles</code>	Anzahl der Messwinkel. Werte größer 1 funktionieren (derzeit) nur mit <code>calculateTB=1</code>
<code>startdataset</code>	Nummer des ersten Datensatzes, auf dem gerechnet werden soll (nummeriert ab 0). Alle vorigen werden ignoriert.
<code>TBNoise</code>	Aus dem Datensatz gelesene Strahlungstemperaturen werden um eine auf $[-\text{TBNoise}, \text{TBNoise}]$ gleichverteilte Zufallsvariable gestört
<code>alpha</code>	entspricht α aus 3.8
<code>beta</code>	entspricht β aus 3.8
<code>gamma</code>	entspricht γ aus 3.8
<code>lambda1</code>	entspricht λ_1 aus 3.4
<code>lambda2</code>	entspricht λ_2 aus 3.4
<code>calculateTB</code>	Ist <code>calculateTB = 1</code> , so werden die TBs aus dem Datensatz ignoriert und neu berechnet (dazu muss das komplette Temperaturprofil im Datensatz vorhanden sein!). Ist <code>calculateTB = 0</code> , so werden die TBs aus dem Datensatz für das Retrieval benutzt.
<code>useTheoreticPressure</code>	Ist <code>useTheoreticPressure = 0</code> , so wird der normale Algorithmus aus 3.10 benutzt. Ist <code>useTheoreticPressure = 1</code> , so wird der Algorithmus wie in 3.11 abgeändert.
<code>keepTempFiles</code>	Ist <code>keepTempFiles = 1</code> , so werden die Dateien, die zum plotten mit <code>gnuplot</code> benötigt werden, nicht gelöscht.
<code>plotGIF</code>	Ist <code>plotGIF = 1</code> so wird von jedem Retrieval ein <code>.GIF</code> mit <code>gnuplot</code> generiert.
<code>plotPS</code>	wie oben, mit <code>.PS</code> statt <code>.GIF</code>
<code>plotEPS</code>	wie oben, mit <code>.EPS</code> statt <code>.GIF</code>
<code>inputfile</code>	Name der Datei, die die Datensätze enthält, für die Retrievals berechnet werden sollen
<code>numdatasets</code>	echte Anzahl der Datensätze (nummeriert ab 1). Ist 0 angegeben, so wird versucht die Anzahl aus der Datei <code>inputfile</code> selbst zu berechnen.
<code>numHumidityProfiles</code>	Anzahl der Feuchteprofile, die pro Datensatz vorhanden sind
<code>numUsedHumidityProfiles</code>	Anzahl der Feuchteprofile, auf denen gerechnet werden soll
<code>numTBProfiles</code>	reservierte Größe, muss auf 1 gesetzt werden

Tabelle 5.1: mögliche Parameter der Datei `meteo.par`

Literaturverzeichnis

- [1] H. W. Alt. *Lineare Funktionalanalysis*. Springer, Berlin-Heidelberg-New York, 1985.
- [2] L. Alvarez, J. Weickert, and J. Sánchez. A scale-space approach to nonlocal optical flow calculations. In M. Nielsen, P. Johansen, O. F. Olsen, and J. Weickert, editors, *Scale-Space Theories in Computer Vision. Second International Conference, Scale-Space '99, Corfu, Greece, September 1999*, Lecture Notes in Computer Science; 1682, pages 235–246. Springer, 1999.
- [3] J. Askne and E. Westwater. A review of ground-based remote sensing of temperature and moisture by passive microwave radiometers. *IEEE Trans. Geosci. Rem. Sens.*, GE-24:340–352, 1986.
- [4] B. Berkels, U. Clarenz, S. Crewell, U. Löhnert, M. Rumpf, and C. Simmer. A physical temperature profiling method using gradient flows. In *Microwave Radiometry and Remote Sensing Applications, Feb. 2004, Rome, Italy*, 2004.
- [5] D. Braess. *Finite Elemente*. Springer, Berlin, 2003.
- [6] M. P. Cadeddu, G. E. Peckham, and C. Gaffard. The vertical resolution of ground-based microwave radiometers analyzed through a multiresolution wavelet technique. *IEEE Transactions on Geoscience and Remote Sensing*, 40(3), 2002.
- [7] U. Clarenz, M. Droske, and M. Rumpf. Towards fast non-rigid registration. In *Inverse Problems, Image Analysis and Medical Imaging, AMS Special Session Interaction of Inverse Problems and Image Analysis*, volume 313, pages 67–84. AMS, 2002.
- [8] U. Clarenz, S. Henn, M. Rumpf, and K. Witsch. Relations between optimization and gradient flow methods with applications to image registration. In W. Hackbusch and M. Griebel, editors, *Multigrid and related methods for optimization problems*, 2002.
- [9] S. Crewell, H. Czekala, U. Löhnert, and C. Simmer et.al. MICCY - a 22 channel ground-based microwave radiometer for atmospheric research. *Radio Science*, 36:621–638, 2001.
- [10] S. Crewell, U. Löhnert, and C. Simmer. An integrated approach towards retrieving physically consistent profiles of temperature, humidity, and cloud liquid water. *Journal of Applied Meteorology*, 43(9), 2004.
- [11] K. Deimling. *Ordinary differential equations in Banach spaces*, volume 96 of *Lecture Notes in Mathematics*. Springer, Berlin-Heidelberg-New York, 1977.
- [12] D. Gilbarg and N. S. Trudinger. *Elliptic partial differential equations of second order*. Springer, Berlin-Heidelberg-New York, 1992.

-
- [13] J. Güldner and D. Spänkuch. Remote sensing of the thermodynamic state of the atmospheric boundary layer by ground-based microwave radiometry. *J. Atmos. Oceanic Technol.*, 18:925–933, 2001.
- [14] M. Hanke and C. W. Groetsch. Nonstationary iterated tikhonov regularization. *J. Optim. Theory and Applications*, 98:37–53, 1998.
- [15] H. Heuser. *Gewöhnliche Differentialgleichungen*. B.G. Teubner, Stuttgart, 1995.
- [16] M. A. Janssen. *Atmospheric Remote Sensing by Microwave Radiometry*. John Wiley and Sons, Inc., New York, 1993.
- [17] E. N. Kadygrov and D. R. Pick. The potential for temperature retrieval from an angular-scanning single-channel microwave radiometer and some comparisons with in situ observations. *Meteorol. Appl.*, 5:393–404, 1998.
- [18] P. Kosmol. *Optimierung und Approximation*. de Gruyter Lehrbuch, 1991.
- [19] H. Kraus. *Die Atmosphäre der Erde: Eine Einführung in die Meteorologie*. Springer, Berlin, 2004.
- [20] A. K. Louis. *Inverse und schlecht gestellte Probleme*. B.G. Teubner, Stuttgart, 1989.
- [21] B. Mohammadi and O. Pironneau. *Applied Shape Optimization for Fluids*. Clarendon Press, Oxford, 2001.
- [22] F. Otto. The geometry of dissipative evolution equations: the porous medium equation. *Comm. Partial Differential Equations*, 26(1-2):101–174, 2001.
- [23] C. D. Rodgers. *Inverse Methods for Atmospheric Sounding - Theory and Practice*. World Scientific, Singapore, 2000.
- [24] O. Scherzer and J. Weickert. Relations between regularization and diffusion filtering. *Journal of Mathematical Imaging and Vision*, 12:43–63, 2000.
- [25] J. Sokolowski and J. P. Zolesio. *Introduction to Shape Optimization*. Springer, Berlin-Heidelberg-New York, 1992.
- [26] F. Solheim, J. Godwin, E. R. Westwater, Y. Han, S. Keihm, K. Marsh, and R. Ware. Radiometric profiling of temperature, water vapor and cloud liquid water using various inversion methods. *Radio Science*, 33:393–404, 1998.
- [27] J. Stör and R. Bulirsch. *Numerische Mathematik II*. Springer, Berlin, 1973.

Index

- A_T , 13, 17
 - $E[T]$, 11
 - $E_P[T]$, 38
 - $E_c[T]$, 37
 - E_{reg} , 38
 - E_{ges} , 38
 - $L(\Gamma)$, 14
 - \mathcal{S} , 18
 - \mathcal{T} , 10
 - T_b , 10
 - \mathcal{T}_0 , 49
 - \mathcal{T}_M , 46
 - \mathcal{T}_{inv} , 39
 - $\tilde{\mathcal{T}}_0$, 49
 - $\tilde{\mathcal{T}}_M$, 49
 - $R(\cdot)$, 13
 - \tilde{E}_{ges} , 49
 - $h_1[T]$, 37

 - Abbruchkriterium, 64
 - Absorptions- und Emissionskoeffizient, 10
 - Auswertung, 56
 - adjoint variable method, 27
 - Algorithmus, 60
 - Armijo-Regel mit Aufweitung, 63

 - Blindtest, 67

 - Differentialgleichung
 - numerische Lösung, 54
 - Differenzenquotient, 50
 - Diskretisierung
 - Ortsdiskretisierung, 45
 - Zeitdiskretisierung, 48

 - Einschrittverfahren, 54
 - Energie
 - Definition, 11
 - Variation
 - Stetigkeit, 33
 - Errorfunction, 38

 - FE-Basisfunktion
 - Definition, 79
 - Fehler
 - RMS, 67
 - systematischer, 67

 - Gradientenfluss, 13
 - Existenz, 36

 - Interpolation
 - exponentiell, 56
 - Inversion, 46

 - Kostenfunktional
 - quadratisches, 11

 - Metrik
 - Darstellung, 15
 - Definition, 14
 - diskret, 51
 - Inverse
 - Stetigkeit, 22
 - Invertierbarkeit, 18
 - Stetigkeit, 24

 - Optimierungsproblem, 11

 - Rückwärtsrechnung, 6
 - Regularisierungsenergie
 - Diskretisierung, 57
 - Variation, 58
 - Erster Teil, 37
 - Variation, 39
 - Stetigkeit, 44
 - Zweiter Teil, 38
- Retrieval, 6
- Runge-Kutta-Verfahren, 55
-
- schlecht gestellt, 5
- Startprofil, 60
- stetige Abhängigkeit, 19

Strahlungstemperaturen, 6

Stetigkeit, 31

Variation, 26

Integralform, 27

Stetigkeit, 31

Vorwärtsmodell, 10

Temperaturprofil

Graph, 14, 37

Testfunktionen, 46

Validierung, 65

Variation

erste, 79

Vorwärtsrechnung, 6

Zeitschritt, 62

Zeitschrittweite

Steuerung, 62

Erklärung

Hiermit versichere ich, dass ich diese Diplomarbeit selbständig und nur unter Gebrauch der angegebenen Hilfsmittel und Quellen angefertigt habe.

Duisburg, April 2005

(Benjamin Berkels)