



UNIVERSITÄT ZU LÜBECK
INSTITUTE OF MATHEMATICS AND
IMAGE COMPUTING

Funktionalanalytische Untersuchung von Residualnetzwerken

A functional-analytic perspective on residual neural networks

Bachelorarbeit

im Rahmen des Studiengangs
Mathematik in Medizin und Lebenswissenschaften
der Universität zu Lübeck

Vorgelegt von

Christian Peter

Ausgegeben und betreut von

Prof. Dr. Jan Lellmann
Institute of Mathematics and Image Computing

Mit Unterstützung von

Maximilian Wessel, M.Sc.
Institute of Mathematics and Image Computing

9. März 2021

Eidesstattliche Erklärung

Ich versichere an Eides statt, die vorliegende Arbeit selbstständig und nur unter Benutzung der angegebenen Quellen und Hilfsmittel angefertigt zu haben.

Lübeck,
6. Mai 2021

Autor

Kurzfassung

Tiefe residuale neuronale Netzwerke haben sich als effektives Werkzeug unter anderem in der Bildverarbeitung und zum Lösen von Klassifizierungsproblemen etabliert. Durch eine hohe Anzahl an Schichten des Netzwerkes steigt die Anzahl der zu lernenden Parameter sowie der Rechenoperationen und somit die Komplexität, was in einem abnehmenden theoretischen Verständnis des Lernprozesses und einer Einschränkung der Interpretierbarkeit der gelernten Parameter resultiert. Dafür legen wir in Kapitel 1 und 2 die mathematischen Grundlagen, um das Lernproblem eines n -schichtigen residualen neuronalen Netzwerkes für allgemeine Trainingsdaten als Minimierungsproblem über diskreten Funktionenräumen zu formalisieren. Weiterhin zeigen wir für $n \rightarrow \infty$ die Γ -Konvergenz des im Lernprozess zu minimierenden Zielfunktional gegen ein von uns definiertes Grenzfunktional, welches über kontinuierlichen Funktionenräumen definiert ist. Mithilfe der Γ -Konvergenz können wir dann die Konvergenz der Minimierer des Zielfunktional gegen die Minimierer des Grenzfunktional zeigen. Zum Abschluss beschreiben wir unseren gescheiterten Versuch, die theoretischen Erkenntnisse in der Praxis zu beobachten und diskutieren mögliche Hindernisse bei der Untersuchung der Γ -Konvergenz in der Praxis.

Abstract

Deep residual networks have proven very effective in fields such as computer vision and classification problems. Due to a higher number of network layers, the amount of learnable parameters and arithmetic operations rises. This causes a decline in theoretical understanding of how the networks learn and limits the interpretability of the learned parameters. Therefore, in chapter 1 and 2, we lay the mathematical foundations to formalize the learning process of an n -layered residual neural network, for general training data, as a minimization problem of an objective functional over discrete L^p -spaces in chapter 3. Furthermore, we show that the *deep layer limit*, e.g. the limit as $n \rightarrow \infty$, of the objective functional coincides with a previously defined limiting objective functional. Using the properties of Γ -convergence, we show that the discrete minimizers of the n -layered objective functional converge to the continuous minimizers of the limiting objective functional, as $n \rightarrow \infty$. Finally we briefly conclude a failed experiment in chapter 4, which should have backed up the theory from chapter 3, and discuss possible reasons of failure.

Inhaltsverzeichnis

1	Γ-Konvergenz	1
1.1	Motivation	1
1.2	Γ -Konvergenz	1
2	Funktionenräume	5
2.1	L^p -Räume	5
2.2	Sobolev-Räume	7
3	Γ-Konvergenz des ResNet	9
3.1	Neuronale Netzwerke	9
3.2	Residuale neuronale Netzwerke	11
3.2.1	Das n -schichtige Modell	11
3.2.2	Das kontinuierliche Grenzfunktional	12
3.3	Parameterräume und Konvergenzbegriffe	13
3.4	Hauptresultat	17
3.4.1	Relative Kompaktheit von unter \mathcal{E}_n beschränkten Folgen in $\Theta^{(n)}$	18
3.4.2	Existenz von Minimierern von \mathcal{E}_n in $\Theta^{(n)}$	23
3.4.3	Γ -Konvergenz von \mathcal{E}_n gegen \mathcal{E}_∞	24
3.4.4	Beweis von Satz 3.3	39
4	Experiment	40
4.1	Versuchsaufbau und -durchführung	40
4.2	Resultate	41
5	Fazit	44

Kapitel 1: Γ -Konvergenz

1.1 Motivation

In diesem Kapitel wird die Γ -Konvergenz behandelt. Die Γ -Konvergenz beschreibt eine Konvergenz von Funktionalen, die unter anderem zur Folge hat, dass die Minima beziehungsweise Minimierer der Funktione konvergieren. Im Folgenden bezeichne X einen metrischen Raum mit Metrik d . Oftmals wird eine Lösung für das Problem

$$\min_{x \in X} f(x) \tag{1}$$

gesucht, wobei f ein Funktional auf X ist. Durch eine mögliche hohe Komplexität von f können Probleme wie (1) im Allgemeinen jedoch sehr schwierig zu lösen sein. Daher kann es sich anbieten, das Funktional f durch einfachere Funktionale $f_n : X \mapsto \overline{\mathbb{R}}$, $n \in \mathbb{N}$ zu approximieren und das Verhalten von Minimierern dieser, sofern sie denn existieren, zu untersuchen. Sei dafür x_n die Lösung von

$$\min_{x \in X} f_n(x). \tag{2}$$

Jetzt stellt sich die Frage, ob und wogegen die Folge der x_n konvergiert. Die Γ -Konvergenz ist ein sehr nützliches Werkzeug zur Beantwortung dieser Fragestellung.

1.2 Γ -Konvergenz

Definition 1.1 (Γ -Konvergenz)

Eine Folge $f_n : X \mapsto \overline{\mathbb{R}}$ Γ -konvergiert in X gegen eine Grenzfunktion $f_\infty : X \mapsto \overline{\mathbb{R}}$, wenn für alle $x \in X$ die folgenden zwei Bedingungen erfüllt sind:

- für alle Folgen $x_n \rightarrow x$ gilt:

$$f_\infty(x) \leq \liminf_{n \rightarrow \infty} f_n(x_n), \tag{3}$$

- es existiert eine Folge $x_n \rightarrow x$, für die gilt:

$$f_\infty(x) \geq \limsup_{n \rightarrow \infty} f_n(x_n). \tag{4}$$

Wobei (3) in der Literatur oft als Liminf-Ungleichung und (4) als Limsup-Ungleichung bezeichnet wird. In Kurzschreibweise: $f_\infty = \Gamma\text{-lim } f_n$.

Für eine Folge $(f_n)_n$ von Γ -konvergenten Funktionalen mit Grenzwert f_∞ lässt sich unter passenden Voraussetzungen zeigen, dass die Minimierer der f_n , bis auf Auswahl einer Teilfolge, gegen einen Minimierer von f_∞ konvergieren. Um dieses Resultat angeben zu können, werden einige technische Hilfsmittel benötigt. Wir charakterisieren zunächst die Kompaktheit von Mengen in metrischen Räumen.

Definition 1.2 (Kompaktheit)

Sei $K \subseteq X$. Dann heißt K

- kompakt in X , wenn jede Folge $(x_n)_{n \in \mathbb{N}} \in K^{\mathbb{N}}$ eine konvergente Teilfolge x_{n_j} besitzt, die gegen einen Punkt aus K konvergiert,
- relativ kompakt in X , wenn jede Folge $(x_n)_{n \in \mathbb{N}} \in K^{\mathbb{N}}$ eine konvergente Teilfolge x_{n_j} besitzt, die gegen einen Punkt aus X konvergiert.

Um die Frage zu untersuchen, wann ein Funktional $f : X \mapsto \overline{\mathbb{R}}$ überhaupt einen Minimierer besitzt, kann man sich der direkten Methode der Variationsrechnung bedienen. Dazu klären wir zunächst, wie der Limes inferior für Funktionale auf metrischen Räumen definiert ist. Anschließend führen wir den Begriff der Unterhalbstetigkeit ein.

Definition 1.3 (Limes inferior)

Sei $f : X \mapsto \overline{\mathbb{R}}$. Dann wird durch

$$\liminf_{y \rightarrow x} f(y) := \inf \left\{ \liminf_{n \rightarrow \infty} f(x_n) \mid x_n \in X, x_n \rightarrow x \right\} \quad (5)$$

der Limes inferior von f in x definiert.

Mithilfe des Limes inferior lässt sich die Unterhalbstetigkeit definieren.

Definition 1.4 (Unterhalbstetigkeit)

Sei $f : X \mapsto \overline{\mathbb{R}}$. Dann heißt f unterhalbstetig auf X , wenn für alle $x \in X$ gilt, dass

$$f(x) \leq \liminf_{y \rightarrow x} f(y). \quad (6)$$

Um einen Minimierer eines beliebigen Funktionals zu konstruieren, ist es vonnöten für eine gegebene Minimalfolge des Funktionals eine konvergente Teilfolge zu finden. Ein Kriterium, welches man für diesen Zweck an das Funktional stellen kann, wird durch die Koerzivität beschrieben.

Definition 1.5 (Koerzivität)

Eine Funktion $f : X \rightarrow \overline{\mathbb{R}}$ heißt koerziv, falls für alle $t \in \mathbb{R}$ die Menge

$$\{f \leq t\} := \{x \in X : f(x) \leq t\} \quad (7)$$

relativ kompakt in X ist.

Eine Funktion $f : X \rightarrow \overline{\mathbb{R}}$ heißt leicht koerziv, falls eine nichtleere kompakte Menge $K \subseteq X$ existiert, für die gilt:

$$\inf_{x \in X} f(x) = \inf_{x \in K} f(x). \quad (8)$$

Damit lässt sich nun ein Kriterium für die Existenz eines Minimierers charakterisieren.

Satz 1.6

Wenn $f : X \mapsto \overline{\mathbb{R}}$ unterhalbstetig und leicht koerziv ist, dann existiert ein $x^* \in X$ mit $f(x^*) = \min_{x \in X} f(x)$.

Beweis: Sei $f : X \mapsto \overline{\mathbb{R}}$ eine Funktion mit genannten Eigenschaften. Dann lässt sich durch die Koerzivitat von f und die Definition des Infimums eine relativ kompakte Menge K und eine Folge $(x_n)_{n \in \mathbb{N}} \in K^{\mathbb{N}}$ finden, fur die gilt:

$$\lim_{n \rightarrow \infty} f(x_n) = \inf_{x \in X} f(x). \quad (9)$$

Da K relativ kompakt ist, hat (x_n) also eine konvergente Teilfolge $x_{n_j} \rightarrow x^* \in X$. Durch die Eigenschaften konvergenter Folgen gilt zusatzlich, dass:

$$\lim_{j \rightarrow \infty} f(x_{n_j}) = \lim_{n \rightarrow \infty} f(x_n) = \inf_{x \in X} f(x). \quad (10)$$

Da f unterhalbstetig ist, lasst sich folgern, dass:

$$f(x^*) \leq \liminf_{y \rightarrow x^*} f(y) \leq \lim_{j \rightarrow \infty} f(x_{n_j}) = \inf_{x \in X} f(x). \quad (11)$$

Also ist x^* ein Minimierer von f . ■

Um nun fur eine Γ -konvergente Folge eine Aussage uber den Minimierer des Γ -Grenzwertes zu treffen, bietet es sich an die Koerzivitat auf Folgen von Funktionen zu verallgemeinern.

Definition 1.7 (gleichgradige leichte Koerzivitat)

Eine Folge von Funktionen $f_n : X \rightarrow \overline{\mathbb{R}}$ heit gleichgradig leicht koerziv, wenn eine nichtleere kompakte Menge $K \subseteq X$ existiert, fur die gilt:

$$\inf_{x \in X} f_n(x) = \inf_{x \in K} f_n(x) \quad \forall n \in \mathbb{N}. \quad (12)$$

Damit lasst sich der folgende Satz beschreiben.

Satz 1.8 ((DM12, Cor. 7.20))

Sei $(f_n)_{n \in \mathbb{N}}$ eine gleichgradig leicht koerzive Folge von Funktionalen auf X , fur die gilt: $\Gamma\text{-lim } f_n = f_\infty$.

Dann nimmt f_∞ ein Minimum auf X an und es gilt:

$$\min_{x \in X} f_\infty(x) = \lim_{n \rightarrow \infty} \inf_{x \in X} f_n(x). \quad (13)$$

Weiterhin gilt, dass wenn fur alle $n \in \mathbb{N}$ ein Minimierer von f_n existiert, dann ist jeder Haufungswert der Folge der Minimierer auch ein Minimierer von f_∞ .

Beweis: Der Beweis orientiert sich stark an (B⁺02)[Thm. 1.21].

Da f gleichgradig leicht koerziv ist, existiert eine relativ kompakte Menge $K \subseteq X$ mit

$$\inf_{x \in X} f_n(x) = \inf_{x \in K} f_n(x) \quad \forall n \in \mathbb{N}. \quad (14)$$

Daraus folgt, dass eine Folge $(x_n)_{n \in \mathbb{N}} \in K^{\mathbb{N}}$ existiert, für die gilt

$$\liminf_{n \rightarrow \infty} f_n(x_n) = \liminf_{n \rightarrow \infty} \inf_{x \in K} f_n(x) = \liminf_{n \rightarrow \infty} \inf_{x \in X} f_n(x). \quad (15)$$

Da K relativ kompakt ist, existiert eine konvergente Teilfolge $x_{n_k} \rightarrow x^* \in X$, für die gilt

$$\lim_{k \rightarrow \infty} f_{n_k}(x_{n_k}) = \liminf_{n \rightarrow \infty} \inf_{x \in K} f_n(x) = \liminf_{n \rightarrow \infty} \inf_{x \in X} f_n(x). \quad (16)$$

Wenn wir die Folge \tilde{x}_n definieren durch

$$\tilde{x}_n = \begin{cases} x_{n_k}, & n = n_k, \\ x^*, & n \neq n_k, \end{cases} \quad (17)$$

dann gilt $\tilde{x}_n \rightarrow x^*$ und somit folgt mit der Liminf-Ungleichung (3), dass:

$$f_\infty(x^*) \leq \liminf_{k \rightarrow \infty} f_n(\tilde{x}_n) \leq \liminf_{k \rightarrow \infty} f_{n_k}(\tilde{x}_{n_k}) = \lim_{k \rightarrow \infty} f_{n_k}(x_{n_k}). \quad (18)$$

Gleichung (16) und (18) kombiniert ergibt

$$f_\infty(x^*) \leq \limsup_{n \rightarrow \infty} \inf_{x \in X} f_n(x). \quad (19)$$

Mit (B⁺02, Prop 1.18(ii)) folgt schlussendlich

$$\limsup_{n \rightarrow \infty} \inf_{x \in X} f_n(x) \leq \inf_{x \in X} f_\infty(x), \quad (20)$$

was zusammen mit (19) die Behauptung impliziert. ■

Wenn also beispielsweise die Minimiererfolge der f_n existiert und relativ kompakt in X ist, dann folgt daraus direkt, dass f_n gleichgradig leicht koerziv ist, da bei Bildung des Abschlusses der Minimiererfolge offensichtlich eine nichtleere, kompakte Menge $K \subseteq X$ entsteht, für die gilt:

$$\inf_{x \in X} f_n(x) = \inf_{x \in K} f_n(x) \quad \forall n \in \mathbb{N}. \quad (21)$$

Außerdem gilt, dass durch die relative Kompaktheit mindestens ein Häufungswert der Minimiererfolge existiert und damit nach Satz 1.8 eine Folge konstruiert werden kann, die gegen den Minimierer von f_∞ konvergiert.

Zwei weitere interessante Eigenschaften der Γ -Konvergenz sind, dass Γ -Grenzwerte, sofern sie denn existieren, immer eindeutig und unterhalbstetig sind (B⁺02, Rem. 1.6, Prop. 1.28). Daraus folgt beispielsweise, für konstante Folgen von Funktionalen $f_n \equiv f$ gilt nur dann $f = \Gamma\text{-lim } f_n$, wenn f unterhalbstetig ist. Andernfalls gilt $\Gamma\text{-lim } f_n = \text{cl}f$, wobei $\text{cl}f = \sup\{g : g \leq f, \quad g \text{ unterhalbstetig}\}$ die unterhalbstetige Hülle von f ist.

Kapitel 2: Funktionenräume

In Kapitel 3 wird das Training eines neuronalen Netzes als Minimierungsproblem über passenden Funktionenräumen formuliert. Deshalb führen wir zunächst die relevanten Funktionenräume ein. Wir starten mit den L^p Räumen bezüglich allgemeiner Maße. Anschließend erklären wir den Begriff der schwachen Ableitung und stellen die sogenannten Sobolev-Räume vor. Zum Abschluss dieses Kapitels nennen wir noch zwei Einbettungssätze, die als technische Hilfsmittel in Beweisen aus dem dritten Kapitel genutzt werden.

2.1 L^p -Räume

Versucht man die p -Norm von endlichdimensionalen Räumen auf Funktionenräume zu erweitern, ergeben sich die L^p -Räume, welche wie folgt definiert sind.

Definition 2.1

Sei (X, \mathcal{A}, μ) ein Maßraum, $(Y, \|\cdot\|)$ ein Banachraum und $1 \leq p \leq \infty$. Dann ist

$$L^p(X, \mathcal{A}, \mu; Y, \|\cdot\|) := \left\{ f : X \mapsto Y \mid f \text{ ist messbar, } \int_X \|f\|^p d\mu < \infty \right\} \quad (22)$$

zusammen mit der Norm

$$\|f\|_{L^p} := \begin{cases} \int_X \|f\|^p d\mu, & 1 \leq p < \infty, \\ \text{ess sup}_{x \in X} \|f(x)\|, & p = \infty, \end{cases} \quad (23)$$

ein Banachraum, wobei μ -fast überall gleiche Funktionen miteinander identifiziert werden.

Im Folgenden, insbesondere in Kapitel 3, wird $Y = \mathbb{R}^\kappa$ gewählt, wobei κ ein Platzhalter für eine natürliche Zahl oder ein 2-Tupel natürlicher Zahlen ist, ausgestattet mit einer geeigneten Norm. Daher schreiben wir abkürzend $L^p(X, \mu) := L^p(X, \mathcal{A}, \mu; \mathbb{R}^\kappa, \|\cdot\|)$, beziehungsweise $L^p(U) := L^p(U, \lambda^n)$ für Teilmengen $U \subseteq \mathbb{R}^n$, wobei λ^n das n -dimensionale Lebesgue-Maß beschreibt. Aufgrund der Tatsache, dass alle Normen auf endlich dimensionalen Räumen, wie zum Beispiel dem \mathbb{R}^κ , äquivalent sind, wird nicht genau spezifiziert welche Norm betrachtet wird, da der tatsächliche Wert der Norm nicht von Interesse ist, sondern nur die durch die Norm induzierte Topologie. Einen besonderen Stellenwert nimmt der Raum $L^2(X, \mu)$ ein, da dieser sich als einziger mit einem Skalarprodukt versehen lässt und damit zu einem Hilbertraum wird. Seien dafür $f, g \in L^2(X, \mu)$. Dann wird durch

$$\langle f, g \rangle_{L^2} := \int_X \langle f(x), g(x) \rangle_{\mathbb{R}^\kappa} d\mu(x) \quad (24)$$

ein Skalarprodukt auf $L^2(X, \mu)$ definiert, wobei $\langle \cdot, \cdot \rangle_{\mathbb{R}^\kappa}$ das Standardskalarprodukt auf \mathbb{R}^κ beschreibt.

Definition 2.2 (Faltung)

Seien $f : \mathbb{R}^n \mapsto \mathbb{R}$, $g : \mathbb{R}^n \mapsto \mathbb{R}^\kappa$. Dann wird durch

$$(f * g)(t) := \int_{\mathbb{R}^n} f(s)g(t - s)ds \quad (25)$$

das Faltungsprodukt von f und g definiert, wobei die Integration der vektor- oder matrixwertigen Funktion $f(\cdot)g(t - \cdot)$ komponentenweise zu verstehen ist.

Durch die Faltung mit einem geeigneten Glättungskern kann man L^p -Funktionen durch C^∞ -Funktionen approximieren. Dafür sei $\eta_1 : \mathbb{R}^n \mapsto \mathbb{R}$ definiert durch

$$\eta_1(x) := \begin{cases} \alpha \exp(-\frac{1}{1-\|x\|^2}), & \|x\| < 1, \\ 0, & \text{sonst.} \end{cases} \quad (26)$$

Der Parameter $\alpha > 0$ sei dabei so gewählt, dass

$$\int_{\mathbb{R}^n} \eta_1(x)dx = 1 \quad (27)$$

gilt. Damit lässt sich der Standardglättungskern für jedes $\varepsilon > 0$ definieren durch

$$\eta_\varepsilon : \mathbb{R}^n \mapsto \mathbb{R}, \quad \eta_\varepsilon(x) = \frac{1}{\varepsilon^n} \eta_1\left(\frac{x}{\varepsilon}\right). \quad (28)$$

Dann gilt, dass (Eva98, Appendix C.4)

$$\eta_\varepsilon \in C^\infty(\mathbb{R}^n), \text{ spt}(\eta_\varepsilon) \subseteq B(0, \varepsilon) \text{ und } \int_{\mathbb{R}^n} \eta_\varepsilon(x)dx = 1. \quad (29)$$

Für eine gegebene integrierbare Funktion $f : \mathbb{R}^n \mapsto \mathbb{R}^\kappa$ sei die Glättung von f definiert durch

$$f^\varepsilon := \eta_\varepsilon * f. \quad (30)$$

Zwei wichtige Eigenschaften von geglätteten L^p -Funktionen werden in dem folgenden Satz charakterisiert.

Satz 2.3

Sei $1 \leq p < \infty$, $U \subseteq \mathbb{R}^n$ und $f \in L^p(U)$. Dann gilt

- $f^\varepsilon \in C^\infty(U)$,
- $f^\varepsilon \rightarrow f$ in $L^p(U)$ für $\varepsilon \rightarrow 0$.

Beweis: In (Eva98, Appendix C.4) wird der Beweis für Funktionen $f : \mathbb{R}^n \mapsto \mathbb{R}$ geführt. Durch eine komponentenweise Betrachtung der geglätteten Funktionen lassen sich die Eigenschaften aus dem Satz jedoch auch problemlos auf Funktionen $f : \mathbb{R}^n \mapsto \mathbb{R}^\kappa$ verallgemeinern. ■

2.2 Sobolev-Räume

Versucht man nun vollständige Funktionenräume mit L^p -Normen zu finden, die sich nicht nur auf die Funktionen selber, sondern auch auf ihre Ableitungen beziehen, ergeben sich die Sobolev-Räume. Dafür definieren wir zuerst ein Konzept der Ableitung für nicht klassisch differenzierbare Funktionen.

Definition 2.4 (schwache Ableitungen)

Sei $U \subseteq \mathbb{R}^n$ offen, $1 \leq p \leq \infty$, $f \in L^p(U)$ und $\alpha := (\alpha_1, \dots, \alpha_n) \in \mathbb{N}^n$ ein Multiindex. Dann heißt ein $g \in L^p$ die α -te schwache Ableitung von f , falls gilt, dass

$$\int_U f(x)\varphi(x)dx = (-1)^{|\alpha|} \int_U g(x)\partial^\alpha \varphi(x)dx, \quad \forall \varphi \in C_c^\infty(U). \quad (31)$$

$C_c^\infty(U)$ beschreibt hier die Menge aller beliebig oft stetig differenzierbaren Funktionen mit kompaktem Träger auf U .

Falls eine α -te schwache Ableitung von einer Funktion f existiert, dann ist sie bis auf einer Nullmenge eindeutig definiert (Eva98, Seite 243). Damit lassen sich die Sobolev-Räume definieren.

Definition 2.5 (Sobolev-Räume)

Sei $1 \leq p \leq \infty$, $k \in \mathbb{N}$ und $U \subseteq \mathbb{R}^n$ offen. Dann ist der Sobolev-Raum

$$W^{k,p}(U) := \{f \in L^p(U) : \partial^\alpha f \in L^p(U), \forall |\alpha| \leq k\} \quad (32)$$

zusammen mit der Norm

$$\|f\|_{W^{k,p}(U)} := \begin{cases} \left(\sum_{|\alpha| \leq k} \|f\|_{L^p(U)}^p \right)^{\frac{1}{p}}, & 1 \leq p < \infty, \\ \sum_{|\alpha| \leq k} \|f\|_{L^\infty(U)}, & p = \infty, \end{cases} \quad (33)$$

ein Banachraum.

Analog zu den L^p -Räumen nimmt auch hier der Raum $W^{k,2}(U)$ einen besonderen Stellenwert ein und wird daher mit $H^k(U) := W^{k,2}(U)$ abgekürzt. Auch hier lässt sich für alle $k \in \mathbb{N}$ ein Skalarprodukt auf $H^k(U)$ definieren. Seien dafür $f, g \in H^k(U)$. Dann ist das Skalarprodukt definiert durch

$$\langle f, g \rangle_{H^k(U)} := \sum_{|\alpha| \leq k} \langle \partial^\alpha f, \partial^\alpha g \rangle_{L^2(U)}. \quad (34)$$

Um die in Kapitel 3 benötigten Einbettungssätze zu formulieren, ist es notwendig zunächst den Begriff der Hölderstetigkeit zu definieren.

Definition 2.6 (Hölderstetigkeit)

Sei $U \subseteq \mathbb{R}^n$, $f : U \mapsto \mathbb{R}^k$ und $0 < \gamma \leq 1$. Dann heißt f hölderstetig mit Exponent γ , wenn eine Konstante $C \in \mathbb{R}$ existiert, sodass gilt

$$\sup_{\substack{x, y \in U \\ x \neq y}} \frac{\|f(x) - f(y)\|}{\|x - y\|^\gamma} \leq C. \quad (35)$$

Mit $(C^{0,\gamma}(U), \|\cdot\|_{C^{0,\gamma}(U)})$ wird der Raum der Funktionen bezeichnet, die beschränkt und hölderstetig mit Exponent γ sind, wobei

$$\|f\|_{C^{0,\gamma}(U)} := \sup_{x \in U} \|f(x)\| + \sup_{x \neq y} \frac{\|f(x) - f(y)\|}{\|x - y\|^\gamma}. \quad (36)$$

Bei $\gamma = 1$ spricht man von Lipschitzstetigkeit. Die Hölderräume lassen sich für abfallende Exponenten kompakt ineinander einbetten, wie der folgende Satz besagt.

Satz 2.7

Sei $U \subseteq \mathbb{R}^n$ beschränkt und $0 < \alpha < \beta \leq 1$. Dann ist der Identitätsoperator

$$\text{id} : C^{0,\beta}(U) \mapsto C^{0,\alpha}(U) \quad (37)$$

kompakt, bedeutet beschränkte Mengen in $C^{0,\beta}(U)$ werden auf relativ kompakte Mengen in $C^{0,\alpha}(U)$ abgebildet.

Beweis : (GT01, Lemma 6.33) ■

Die Hölderräume lassen sich außerdem unter bestimmten Bedingungen in Sobolev-Räume einbetten. Diese Eigenschaft wird durch den folgenden Satz charakterisiert.

Satz 2.8 (Morrey)

Sei $n < p \leq \infty$, $U \subseteq \mathbb{R}^n$ beschränkt und es gelte ∂U ist C^1 . Dann existiert eine von p und n abhängige Konstante $C \in \mathbb{R}$, mit

$$\|u\|_{C^{0,\gamma}(U)} \leq C \|u\|_{W^{1,p}(U)} \quad (38)$$

für alle $u \in C^1(U)$, wobei

$$\gamma := 1 - \frac{n}{p}. \quad (39)$$

Beweis : Der Beweis für $U = \mathbb{R}^n$ wird in (Eva98, S. 240, Thm. 4) geführt und lässt sich auf beschränkte Teilmengen U mit glattem Rand verallgemeinern, indem $W^{1,p}(U)$ normerhaltend auf $W^{1,p}(\mathbb{R}^n)$ fortgesetzt wird. ■

Kapitel 3: Γ -Konvergenz des ResNet

Im Folgenden werden wir ein allgemein gehaltenes Modell eines residualen neuronalen Netzwerkes (HZRS16) mit n Schichten definieren und das Training des Netzwerkes für einen gegebenen Trainingsdatensatz als Minimierungsproblem eines Zielfunktional über diskreten L^p -Räumen formulieren. Anschließend untersuchen wir das Verhalten des Modells für $n \rightarrow \infty$ und zeigen, dass das n -schichtige Zielfunktional unter bestimmten Bedingungen gegen ein kontinuierliches Grenzfunktional Γ -konvergiert, indem wir die Beweise aus (TG18) ausformulieren. Mithilfe dieser Tatsache zeigen wir dann, dass Minimierer des Zielfunktional gegen Minimierer des Grenzfunktional konvergieren.

3.1 Neuronale Netzwerke

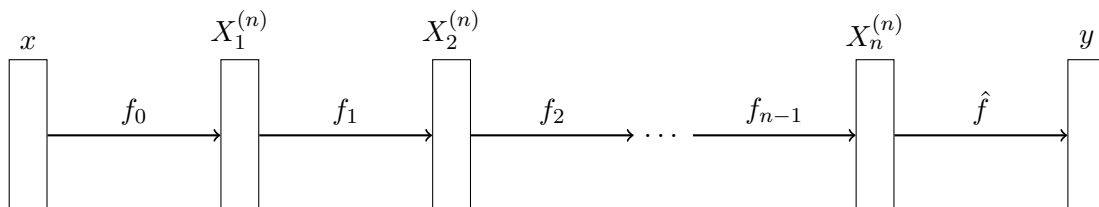
Ein Anwendungsfeld von neuronalen Netzwerken besteht aus Klassifizierungsproblemen, wobei ein Eingabewert x gegeben ist und sinnvoll einer Klasse aus einer Menge von m Klassen $\{y_i\}_{i=1}^m$ zugeordnet werden soll. Ein prominentes Beispiel ist die Klassifizierung von handgeschriebenen Ziffern (Den12). Dafür sei der Eingabewert des Netzwerkes $x \in \mathbb{R}^d$ ein vektorisiertes Bild einer handgeschriebenen Ziffer und die Menge der Klassen sei $\{0, 1, \dots, 9\}$. Dann ist also eine Funktion $F : \mathbb{R}^d \rightarrow \{0, 1, \dots, 9\}$ gesucht, welche die handgeschriebenen Ziffern möglichst korrekt den gegebenen Klassen zuordnet. Eine solche Funktion F kann beispielsweise durch das Training eines neuronalen Netzwerkes konstruiert werden. Dafür klären wir zunächst, wie ein solches Netzwerk aufgebaut ist. Ein n -schichtiges neuronales Netzwerk ist aus einem mathematischen Standpunkt eine Verknüpfung von n Funktionen f_0, \dots, f_{n-1} , sodass für einen Eingabewert $x \in \mathbb{R}^d$ der Zustand der $(i + 1)$ -ten Schicht definiert ist durch

$$X_{i+1}^{(n)} := f_i(f_{i-1}(\dots f_0(x) \dots)) \in \mathbb{R}^d, \quad i = 0, \dots, n - 1. \quad (40)$$

Dabei wird die Annahme getroffen, dass jede Schicht aus d Neuronen besteht und die 0-te Schicht $X_0^{(n)} := x$ als der Eingabewert des Netzwerkes definiert ist. Der Ausgabewert des neuronalen Netzwerkes ergibt sich durch die Anwendung von einer Klassifizierungsfunktion $\hat{f} : \mathbb{R}^d \rightarrow \mathbb{R}^m$ auf $X_n^{(n)}$, die die Werte in $X_n^{(n)}$ den jeweiligen Klassen zuordnet. Zusammengefasst lässt sich der Ausgabewert also darstellen durch

$$y = \hat{f}(X_n^{(n)}) = \hat{f}(f_n(f_{n-1}(\dots f_0(x) \dots))) =: F(x), \quad (41)$$

wie in der folgenden Grafik visualisiert wird.



Für gegebene Trainingsdaten $\{x_s, y_s\}_{s=1}^S$, wobei x_s ein Eingabewert mit zugehöriger Klasse y_s ist, besteht das Ziel also darin, Funktionen $f_0, \dots, f_{n-1}, \hat{f}$ zu finden, sodass die Ausgabewerte des neuronalen Netzwerkes $F(x_s)$ möglichst nah an den zugehörigen Klassen y_s liegen, in der Hoffnung mithilfe des Netzwerkes auch neue Eingabewerte sinnvoll klassifizieren zu können. Um den Fehler von $F(x_s)$ und y_s quantifizieren zu können, wird eine geeignete Kostenfunktion $\mathcal{L} : \mathbb{R}^m \times \mathbb{R}^m \rightarrow \mathbb{R}_{\geq 0}$ definiert. Also werden Funktionen $f_0, \dots, f_{n-1}, \hat{f}$ gesucht, sodass der Ausdruck

$$\sum_{s=1}^S \mathcal{L}(F(x_s), y_s) = \sum_{s=1}^S \mathcal{L}(\hat{f}(f_n(f_{n-1}(\dots f_0(x_s) \dots))), y_s) \quad (42)$$

minimal wird. Die Wahl der Kostenfunktion ist offensichtlich ausschlaggebend für die Funktionalität des Netzwerkes. Daher werden wir, um einen hohen Grad an Allgemeingültigkeit zu erhalten, nur leichte Anforderungen an die Kostenfunktion unseres Modells stellen, doch dazu später mehr. Üblicherweise wird in neuronalen Netzwerken die Annahme getroffen, dass die Übergangsfunktionen der Form

$$f_i(x) = \sigma_i(K_i^{(n)}x + b_i^{(n)}) \quad (43)$$

sind, wobei $K_i^{(n)} \in \mathbb{R}^{d \times d}$ die Gewichtsmatrix und $b_i^{(n)} \in \mathbb{R}^d$ der Schwellenwert ist. In dem hier untersuchten Modell wird angenommen, dass alle Übergangsfunktionen die gleiche Form haben, also $\sigma_i = \sigma$, für alle $i = 1, \dots, n-1$. Eine beliebige Wahl für die Übergangsfunktionen sind Aktivierungsfunktionen, wie zum Beispiel die *rectified linear unit*, kurz *ReLU* (NH10). Diese ist definiert durch

$$\text{relu}(x) := \max(0, x), \quad (44)$$

wobei die Funktion für vektorwertige Argumente komponentenweise agiert. Aber auch andere Aktivierungsfunktionen wie zum Beispiel die auch komponentenweise agierende, stetig differenzierbare Sigmoidfunktion

$$\text{sig}(x) := \frac{1}{1 + e^{-x}} \quad (45)$$

sind oft das Mittel der Wahl. In unserem Modell stellen wir auch hier nur leichte Anforderungen an die Übergangsfunktion σ , welche nur in einer Lipschitz-Bedingung und einem Verlauf durch den Ursprung besteht. Die Klassifizierungsfunktion hat in unserem Fall die Form

$$\hat{f}(x) := h(Wx + c), \quad (46)$$

wobei $h : \mathbb{R}^m \mapsto \mathbb{R}^m$ eine gegebene Funktion ist, $W \in \mathbb{R}^{m \times d}$ und $c \in \mathbb{R}^m$.

Für gegebene Trainingsdaten $\{x_s, y_s\}_{s=1}^S$ lässt sich das Problem des Findens von optimalen Funktionen $f_0, \dots, f_{n-1}, \hat{f}$ also reduzieren auf das Finden von optimalen Parametern $\{K_i^{(n)}\}, \{b_i^{(n)}\}, \{W\}, \{c\}$ mit $i = 0, \dots, n-1$, durch welche die Funktionen $f_0, \dots, f_{n-1}, \hat{f}$ eindeutig definiert sind.

3.2 Residuale neuronale Netzwerke

Tiefe neuronale Netzwerke, also Netzwerke mit einer großen Anzahl von Schichten, sind sehr effektiv in Bereichen wie der *Computer Vision* und der Spracherkennung. Das Trainieren eines solchen tiefen Netzwerkes stellt eine große Herausforderung dar, welche in (ZLSZ20) diskutiert wird. In der Praxis hat sich gezeigt, dass eine residuale Netzwerkarchitektur das Training von tiefen, neuronalen Netzen stabilisieren kann (HZRS16). In einem residualen neuronalen Netzwerk, welches als *ResNet* bezeichnet wird, addiert man in der Übergangsfunktion von einer Schicht zur Nächsten zu der üblichen affin-linearen Transformation gefolgt von einer nichtlinearen Aktivierungsfunktion den Zustandsvektor der vorherigen Schicht hinzu.

$$f_i(x) = x + \sigma(K_i x + b_i). \quad (47)$$

Im Folgenden werden wir zunächst ein Modell eines n -schichtigen residualen Netzwerkes definieren und die Parameter des Netzwerkes mit diskreten Funktionen identifizieren. Im Anschluss werden wir die Anzahl n der Schichten gegen unendlich gehen lassen und dadurch für die diskreten Komponenten des n -schichten Netzwerkes kontinuierliche Grenzwerte erhalten. Dies liefert uns eine kontinuierliche Formulierung des Modells.

3.2.1 Das n -schichtige Modell

In unserem Modell wird in (47) noch der Faktor $\frac{1}{n}$ eingeführt. Der Übergang der i -ten zur $(i+1)$ -ten Schicht sei also gegeben durch

$$X_{i+1}^{(n)} = X_i^{(n)} + \frac{1}{n} \sigma \left(K_i^{(n)} X_i^{(n)} + b_i^{(n)} \right), \quad i = 0, \dots, n-1. \quad (48)$$

Im Folgenden schreiben wir abkürzend $\mathbf{K}^{(n)} = \{K_i^{(n)}\}_{i=0}^{n-1}$ und $\mathbf{b}^{(n)} = \{b_i^{(n)}\}_{i=0}^{n-1}$. Wir können so den Ausgabewert des Netzwerkes mit Eingabewert $X_0^{(n)} = x$ für gegebene Parameter $\mathbf{K}^{(n)}, \mathbf{b}^{(n)}, W, c$ kompakt schreiben als

$$h(WX_n^{(n)}[x; \mathbf{K}^{(n)}, \mathbf{b}^{(n)}] + c), \quad (49)$$

wobei $X_i^{(n)}[x; \mathbf{K}^{(n)}, \mathbf{b}^{(n)}]$ für alle $i = 1, \dots, n$ rekursiv durch (48) gegeben ist. Damit können wir also für gegebene Trainingsdaten $\{x_s, y_s\}_{s=1}^S$ den Trainingsprozess des Netzwerkes durch das Minimierungsproblem

$$\min_{\mathbf{K}^{(n)}, \mathbf{b}^{(n)}, W, c} \sum_{s=1}^S E_n(\mathbf{K}^{(n)}, \mathbf{b}^{(n)}, W, c; x_s, y_s) \quad (50)$$

beschreiben, wobei die Funktion E_n den Fehler, also die Abweichung der vorhergesagten Klassen $f(x_s)$ von den wahren Klassen y_s , in Abhängigkeit der Parameter des Netzwerkes beschreibt und definiert ist durch

$$E_n(\mathbf{K}^{(n)}, \mathbf{b}^{(n)}, W, c; x, y) := \mathcal{L}(h(WX_n^{(n)}[x; \mathbf{K}^{(n)}, \mathbf{b}^{(n)}] + c), y). \quad (51)$$

Von der Kostenfunktion \mathcal{L} fordern wir lediglich die Stetigkeit im ersten Argument und dass sie nichtnegativ ist.

Lässt man nun die Anzahl der Schichten beliebig groß werden, so übersteigt die Anzahl der Netzparameter $\mathbf{K}^{(n)}, \mathbf{b}^{(n)}, W, c$ schnell die Anzahl der Trainingsdaten. Dies führt in der Praxis zur Überanpassung des Netzwerkes, wodurch eine nahezu perfekte Vorhersage der Klassen auf den Trainingsdaten erreicht wird, auf ungesehen Daten schlägt eine sinnvolle Klassifizierung jedoch fehl. Um dieses Problem zu umgehen, führen wir noch Regularisierer für die Parameter $\mathbf{K}^{(n)}, \mathbf{b}^{(n)}, W, c$ ein, die zu große Parameter und ein zu starkes Variieren der einzelnen Parameter zwischen den Schichten bestrafen. Die Regularisierer seien wie folgt definiert

$$R_n^{(1)}(\mathbf{K}^{(n)}) := n \sum_{i=0}^{n-1} \|K_i^{(n)} - K_{i-1}^{(n)}\|^2 + \tau_1 \|K_0^{(n)}\|^2, \quad (52)$$

$$R_n^{(2)}(\mathbf{b}^{(n)}) := n \sum_{i=0}^{n-1} \|b_i^{(n)} - b_{i-1}^{(n)}\|^2 + \tau_2 \|b_0^{(n)}\|^2, \quad (53)$$

$$R^{(3)}(W) := \|W\|^2, \quad (54)$$

$$R^{(4)}(c) := \|c\|^2, \quad (55)$$

wobei $\tau_1, \tau_2 > 0$ gegebene Gewichtungsfaktoren sind. Da die Parameter alle in endlich-dimensionalen Räumen liegen spezifizieren wir analog zu Abschnitt 2.1 auch hier die zu betrachtende Norm nicht. Zusammen ergibt sich dann für gegebene Trainingsdaten $\{x_s, y_s\}_{s=1}^S$ und Regularisierungsparameter $\alpha_i, i = 1, \dots, 4$ das finale zu minimierende Zielfunktional

$$\begin{aligned} \mathcal{E}_n(\mathbf{K}^{(n)}, \mathbf{b}^{(n)}, W, c) &:= \sum_{s=1}^S E_n(\mathbf{K}^{(n)}, \mathbf{b}^{(n)}, W, c; x_s, y_s) + \alpha_1 R_n^{(1)}(\mathbf{K}^{(n)}) \\ &+ \alpha_2 R_n^{(2)}(\mathbf{b}^{(n)}) + \alpha_3 R^3(W) + \alpha_4 R^4(c). \end{aligned} \quad (56)$$

3.2.2 Das kontinuierliche Grenzfunktional

Um den Übergang von diskreten zu kontinuierlichen Parametern zu betrachten, identifizieren wir $\{K_i^{(n)}\}_{i=0}^{n-1}, \{b_i^{(n)}\}_{i=0}^{n-1}$ und $\{X_i^{(n)}\}_{i=0}^{n-1}$ mit Funktionen $K : [0, 1] \mapsto \mathbb{R}^{d \times d}, b : [0, 1] \mapsto \mathbb{R}^d$ und $X : [0, 1] \mapsto \mathbb{R}^d$, sodass

$$K_i^{(n)} = K\left(\frac{i}{n}\right), \quad b_i^{(n)} = b\left(\frac{i}{n}\right), \quad X_i^{(n)} = X\left(\frac{i}{n}\right) \quad (57)$$

gilt. Durch umstellen von (48) ergibt sich als Übergangsvorschrift

$$\frac{X\left(\frac{i+1}{n}\right) - X\left(\frac{i}{n}\right)}{\frac{i}{n}} = \sigma\left(K\left(\frac{i}{n}\right) X\left(\frac{i}{n}\right) + b\left(\frac{i}{n}\right)\right), \quad (58)$$

was eine Diskretisierung der gewöhnlichen Differentialgleichung

$$\dot{X}(t) = \sigma(K(t)X(t) + b(t)) \quad (59)$$

mit Schrittweite $\frac{1}{n}$ beschreibt. Analog zu (49) beschreibe $X(t; x, K, b)$ die Lösung von (59) mit Anfangswertbedingung $X(0) = x$ und Parametern K, b . Im Folgenden nehmen wir an, dass die Funktion σ komponentenweise lipschitzstetig ist. Daraus folgt mit dem Satz von Picard-Lindelöf (For17, §12 Satz 4), dass eine eindeutige Lösung der Differentialgleichung (59) existiert, sofern wir \dot{X} in den Randpunkten $\{0\}$ und $\{1\}$ als rechtsseitige, beziehungsweise linksseitige Ableitung von X verstehen.

Damit können wir ein kontinuierliches Gegenstück von (51) konstruieren durch

$$E_\infty(K, b, W, c; x, y) := \mathcal{L}(h(WX(t; x, K, b) + c), y). \quad (60)$$

Die Summen in der Definition der Regularisierer $R_n^{(1)}$, und $R_n^{(2)}$ sind Diskretisierungen der H^1 -Norm. Daher definieren wir Regularisierer für die kontinuierlichen Funktionen K und b durch

$$R_\infty^{(1)}(K) := \|\dot{K}\|_{L^2([0,1])}^2 + \tau_1 \|K(0)\|^2 \quad (61)$$

und

$$R_\infty^{(2)}(b) := \|\dot{b}\|_{L^2([0,1])}^2 + \tau_2 \|b(0)\|^2. \quad (62)$$

Da $R^{(3)}$ und $R^{(4)}$ unabhängig von n sind, bleiben diese unverändert. Damit lässt sich also eine kontinuierliche Version von \mathcal{E}_n formulieren durch

$$\begin{aligned} \mathcal{E}_\infty(K, b, W, c) := & \sum_{s=1}^S E_\infty(K, b, W, c; x_s, y_s) + \alpha_1 R_\infty^{(1)}(K) \\ & \alpha_2 R_\infty^{(2)}(b) + \alpha_3 R^{(3)}(W) + \alpha_4 R^4(c). \end{aligned} \quad (63)$$

Das Hauptresultat der Arbeit besteht darin zu zeigen, dass das Minimum sowie die Minimierer von \mathcal{E}_n für $n \rightarrow \infty$ gegen das Minimum, sowie die Minimierer von \mathcal{E}_∞ konvergieren. Dafür muss jedoch erst geklärt werden, in welchen Räumen die Minimierer der Funktionale \mathcal{E}_n und \mathcal{E}_∞ gesucht werden und wie die Konvergenz zwischen diesen Räumen charakterisiert wird. Diese Fragestellung wird im nächsten Abschnitt behandelt.

3.3 Parameterräume und Konvergenzbegriffe

In dem n -schichtigen Modell des Netzwerkes wurden die Parameter $\mathbf{K}^{(n)}, \mathbf{b}^{(n)}, W, c$ definiert, wobei $\mathbf{K}^{(n)}$ und $\mathbf{b}^{(n)}$ Mengen von Matrizen beziehungsweise Vektoren sind. In Hinsicht darauf, dass wir einen Konvergenzbegriff zwischen den diskreten Parametern $\mathbf{K}^{(n)}$ und $\mathbf{b}^{(n)}$ und den kontinuierlichen Parametern K und b , welche Funktionen auf $[0, 1]$ sind, finden wollen, ist es sinnvoll die Parameter $\mathbf{K}^{(n)} = \{K_i^{(n)}\}_{i=0}^{n-1}$ und $\mathbf{b}^{(n)} = \{b_i^{(n)}\}_{i=0}^{n-1}$ als Funktionen bezüglich eines diskreten Maßes, welches auf der Menge $\{\frac{i}{n}\}_{i=0}^{n-1}$ konzentriert ist, zu verstehen. Dafür sei im Folgenden $t_i^{(n)}$ definiert durch $t_i^{(n)} := \frac{i}{n}$. Damit definieren wir das Maß

$$\mu_n := \frac{1}{n} \sum_{i=0}^{n-1} \delta_{\frac{i}{n}} = \frac{1}{n} \sum_{i=0}^{n-1} \delta_{t_i^{(n)}} \quad (64)$$

auf $[0, 1]$. In Bemerkung 3.1 werden zwei wichtige Eigenschaften von μ_n festgehalten.

Bemerkung 3.1

Das Maß μ_n ist ein Borel-Wahrscheinlichkeitsmaß auf $[0, 1]$ und konvergiert schwach gegen das Lebesgue-Maß λ auf $[0, 1]$.

Beweis: Dass μ_n ein Borel-Wahrscheinlichkeitsmaß ist, folgt direkt daraus, dass es als eine Linearkombination von Borelmaßen mit positiven Koeffizienten definiert ist und

$$0 \leq \mu_n(A) \leq 1, \quad \forall A \subseteq [0, 1] \quad (65)$$

und

$$\mu_n([0, 1]) = \frac{1}{n} \sum_{i=0}^{n-1} 1 = 1. \quad (66)$$

Für die schwache Konvergenz muss gezeigt werden, dass für alle beschränkten, stetigen Funktionen f auf $[0, 1]$ gilt, dass

$$\lim_{n \rightarrow \infty} \int_{[0,1]} f d\mu_n = \int_{[0,1]} f d\lambda. \quad (67)$$

Sei f dafür eine beliebige beschränkte, stetige Funktion auf $[0, 1]$. Dann gilt

$$\lim_{n \rightarrow \infty} \int_{[0,1]} f d\mu_n = \lim_{n \rightarrow \infty} \frac{1}{n} \sum_{i=0}^{n-1} f\left(\frac{i}{n}\right). \quad (68)$$

Die rechte Seite von Gleichung (68) beschreibt genau das Riemann-Integral von f über $[0, 1]$ und ist durch die Stetigkeit von f gleich dem Lebesgue-Integral von f über $[0, 1]$, woraus die Behauptung folgt. ■

Damit können wir also die Parameter $\mathbf{K}^{(n)}$ und $\mathbf{b}^{(n)}$ als Elemente des diskreten L^p Raumes $L^2\left(\left\{t_i^{(n)}\right\}_{i=0}^{n-1}, \mu_n\right)$ auffassen. Dafür führen wir die Schreibweise $K^{(n)}$, beziehungsweise $b^{(n)}$ für die diskreten Funktionen ein und schreiben weiterhin $K_i^{(n)} = K^{(n)}\left(\frac{i}{n}\right) \in \mathbb{R}^{d \times d}$, beziehungsweise $b_i^{(n)} = b^{(n)}\left(\frac{i}{n}\right) \in \mathbb{R}^d$ für die Matrizen beziehungsweise die Vektoren. Also definieren wir den Parameterraum für das n -schichtige Netzwerk durch

$$\Theta^{(n)} := L^2\left(\left\{t_i^{(n)}\right\}_{i=0}^{n-1}, \mu_n; \mathbb{R}^{d \times d}\right) \times L^2\left(\left\{t_i^{(n)}\right\}_{i=0}^{n-1}, \mu_n; \mathbb{R}^d\right) \times \mathbb{R}^{m \times d} \times \mathbb{R}^m. \quad (69)$$

Da die Regularisierer für K und b des Grenzfunktional \mathcal{E}_∞ als H^1 -Normen konstruiert wurden, wählen wir dementsprechend den Parameterraum des Grenzfunktional als

$$\Theta := H^1([0, 1]; \mathbb{R}^{d \times d}) \times H^1([0, 1]; \mathbb{R}^d) \times \mathbb{R}^{m \times d} \times \mathbb{R}^m. \quad (70)$$

Hierbei ist zu beachten, dass in Kapitel 2 die Sobolevräume $H^1(U)$ nur für offene Mengen U definiert wurden. Da aber aus Satz 2.8 folgt, dass Elemente von $H^1((0, 1))$ hölderstetig mit Exponent $\gamma = \frac{1}{2}$ und somit stetig sind, beziehungsweise zumindest einen stetigen

Vertreter besitzen, lassen sich die Elemente aus $H^1((0, 1))$ problemlos stetig auf $[0, 1]$ fortsetzen, was uns die Betrachtung von $H^1([0, 1])$ und die Anwendung der Sätze aus Kapitel 2 auf diesen ermöglicht. Das Problem, einen Konvergenzbegriff zwischen Elementen aus $\Theta^{(n)}$ und Θ zu konstruieren, beschränkt sich darauf, einen sinnvollen Abstandsbe-
griff zwischen Elementen aus $L^2\left(\left\{t_i^{(n)}\right\}_{i=0}^{n-1}, \mu_n; \mathbb{R}^\kappa\right)$ und $H^1([0, 1]; \mathbb{R}^\kappa)$ zu finden. Seien dafür $f_n \in L^2\left(\left\{t_i^{(n)}\right\}_{i=0}^{n-1}, \mu_n; \mathbb{R}^\kappa\right)$ und $f \in H^1([0, 1]; \mathbb{R}^\kappa)$. Dann setzen wir zunächst die diskrete Funktion f_n von $\left\{t_i^{(n)}\right\}_{i=0}^{n-1}$ stückweise konstant zu einer kontinuierlichen Funktion \tilde{f}_n auf $[0, 1]$ fort durch

$$\tilde{f}_n(t) := f_n(t_i^{(n)}), \text{ für } t \in [t_i^{(n)}, t_{i+1}^{(n)}), \quad i = 0, \dots, n-1. \quad (71)$$

Damit lässt sich die Funktion \tilde{f}_n bezüglich der L^2 -Norm mit f vergleichen. So können wir einen Abstandsbe-
griff zwischen f_n und f definieren durch

$$\begin{aligned} \hat{d} : L^2\left(\left\{t_i^{(n)}\right\}_{i=0}^{n-1}, \mu_n; \mathbb{R}^\kappa\right) \times H^1([0, 1]; \mathbb{R}^\kappa) &\mapsto [0, \infty) \\ \hat{d}(f_n, f) &:= \|\tilde{f}_n - f\|_{L^2[0,1]}. \end{aligned} \quad (72)$$

Hierbei ist jedoch zu beachten, dass \hat{d} keine Metrik ist, da die Symmetrieeigenschaft nicht gegeben ist. Der Ausdruck $\hat{d}(f, f_n)$ ist nicht definiert. Um trotzdem die Eigenschaften von Metriken ausnutzen zu können, was insbesondere für die Anwendung der Theorie aus Kapitel 1 vonnöten ist, führen wir den TL^2 -Abstand ein. Dieser ist für eine offene Teilmenge $U \subseteq \mathbb{R}^n$ definiert durch

$$d_{TL^2}((\mu, f), (\nu, g)) := \inf_{\pi \in \Gamma(\mu, \nu)} \left(\iint_{U \times U} \|x - y\|^2 + \|f(x) - g(y)\|^2 d\pi(x, y) \right)^{\frac{1}{2}} \quad (73)$$

und ist eine Metrik auf dem Raum

$$TL^2(U) := \left\{ (\mu, f) : \mu \in \mathcal{P}(U), f \in L^2(U, \mu) \right\}, \quad (74)$$

wie in (TS16) untersucht wird. Hierbei beschreibt $\mathcal{P}(U)$ die Menge aller Borel-Wahr-
scheinlichkeitsmaße auf U und $\Gamma(\mu, \nu)$ die Menge aller Borel-Wahrscheinlichkeitsmaße auf $U \times U$, deren Randverteilungen im ersten Faktor μ und im zweiten Faktor ν entsprechen. Wählen wir $\mu = \mu_n$ und $\nu = \lambda$, dann sind die Abstandsbe-
griffe \hat{d} und d_{TL^2} in gewisser Weise äquivalent, wie der folgende Satz besagt.

Satz 3.2

$$\begin{array}{|l} \text{Sei } f_n \in L^2\left(\left\{t_i^{(n)}\right\}_{i=0}^{n-1}, \mu_n; \mathbb{R}^\kappa\right), f \in H^1([0, 1]). \text{ Dann gilt} \\ \\ d_{TL^2((0,1))}((f_n, \mu_n), (f, \lambda)) \xrightarrow{n \rightarrow \infty} 0 \iff \hat{d}(f_n, f) \xrightarrow{n \rightarrow \infty} 0. \end{array} \quad (75)$$

Beweis: Da λ das Lebesgue-Maß beschreibt und somit trivialerweise absolut stetig bezüglich des Lebesgue-Maßes ist, lässt sich (TS16, Proposition 3.12 4.) anwenden. Dafür definieren wir eine Folge von Abbildungen $\{T_n\}_{n \in \mathbb{N}}$ durch

$$\begin{aligned} T_n &: (0, 1) \mapsto \{t_i^{(n)}\}_{i=0}^{n-1} \\ T_n(x) &= t_i^{(n)}, \quad x \in [t_i^{(n)}, t_{i+1}^{(n)}) \text{ für } i = 1, \dots, n-1 \\ &\text{und } T_n(x) = t_0, \quad x \in (t_0, t_1). \end{aligned} \quad (76)$$

und zeigen, dass $\{T_n\}_{n \in \mathbb{N}}$ eine stagnierende Folge von Transportabbildungen zwischen μ_n und λ im Sinne von (TS16) ist. Dafür müssen die folgenden zwei Aussagen gezeigt werden.

1. Für festes $n \in \mathbb{N}$ ist das Bildmaß von λ unter T_n gegeben durch μ_n .
2. $\int_{(0,1)} |x - T_n(x)| d\lambda(x) \rightarrow 0$ für $n \rightarrow \infty$

Wir starten mit Aussage 1. Das Bildmaß von λ unter T_n ist definiert durch

$$T_{n*}\lambda(A) := \lambda(T_n^{-1}(A)), \quad \forall A \subseteq \mathcal{B}((0, 1)) \quad (77)$$

Sei $A \subseteq \mathcal{B}((0, 1))$ beliebig. Wir definieren $I \subseteq \{0, \dots, n-1\}$ als die Menge aller $i = 0, \dots, n-1$, für die gilt, dass $\frac{i}{n} \in A$. Falls $I = \emptyset$, ist $T^{-1}(A) = \emptyset$ und somit

$$T_{n*}\lambda(A) = \lambda(\emptyset) = 0 = \mu_n(A). \quad (78)$$

Falls $I \neq \emptyset$, ist $T^{-1}(A) = \bigcup_{i \in I} [\frac{i}{n}, \frac{i+1}{n})$. Damit folgt, dass

$$T_{n*}\lambda(A) = \lambda\left(\bigcup_{i \in I} \left[\frac{i}{n}, \frac{i+1}{n}\right)\right) = \sum_{i \in I} \lambda\left(\left[\frac{i}{n}, \frac{i+1}{n}\right)\right) = \sum_{i \in I} \frac{1}{n} = \frac{1}{n} \sum_{i=0}^{n-1} \delta_{\frac{i}{n}}(A) = \mu_n(A). \quad (79)$$

Für die 2. Aussage betrachten wir

$$\int_{(0,1)} |x - T_n(x)| d\lambda(x) = \sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} |x - t_i| dx. \quad (80)$$

Mit der Substitution $u = x - t_i$ folgt

$$\sum_{i=0}^{n-1} \int_{t_i}^{t_{i+1}} |x - t_i| dx = \sum_{i=0}^{n-1} \int_0^{\frac{1}{n}} u du = \sum_{i=0}^{n-1} \frac{1}{2n^2} = \frac{1}{2n} \rightarrow 0, \text{ für } n \rightarrow \infty. \quad (81)$$

Also ist $\{T_n\}_{n \in \mathbb{N}}$ eine stagnierende Folge von Transportabbildungen zwischen μ_n und λ . Daher folgt mit (TS16, Proposition 3.12 4.), dass

$$d_{TL^2((0,1))}((f_n, \mu_n), (f, \lambda)) \xrightarrow{n \rightarrow \infty} 0 \iff \|f - f_n \circ T_n\|_{L^2((0,1))} \xrightarrow{n \rightarrow \infty} 0. \quad (82)$$

Die Behauptung von Satz 3.2 folgt aus der Tatsache, dass

$$\tilde{f} = f_n \circ T_n. \quad (83) \quad \blacksquare$$

Satz 3.2 hat die Konsequenz, dass wir die Konvergenz zwischen Elementen aus $L^2\left(\left\{t_i^{(n)}\right\}_{i=0}^{n-1}, \mu_n; \mathbb{R}^\kappa\right)$ und $H^1([0, 1])$ bezüglich des Abstandsbegriffes \hat{d} beschreiben und gleichzeitig die Eigenschaften einer Metrik ausnutzen können, was es uns ermöglicht die Theorie aus Kapitel 1 anzuwenden. Der hier betrachtete Abstandsbegriff zwischen $\Theta^{(n)}$ und Θ sei also für $\theta^{(n)} = (K^{(n)}, b^{(n)}, W^{(n)}, c^{(n)}) \in \Theta^{(n)}$ und $\theta = (K, b, W, c) \in \Theta$ definiert durch

$$d : \Theta^{(n)} \times \Theta \mapsto [0, \infty)$$

$$d(\theta^{(n)}, \theta) := \hat{d}(K^{(n)}, K) + \hat{d}(b^{(n)}, b) + \|W^{(n)} - W\| + \|c^{(n)} - c\|. \quad (84)$$

Damit lässt sich das Hauptresultat formulieren.

3.4 Hauptresultat

Das Hauptresultat wird in Satz 3.3 festgehalten und charakterisiert die Konvergenz der Minimierer von \mathcal{E}_n gegen Minimierer von \mathcal{E}_∞ .

Satz 3.3 ((TG18, Theorem 2.1))

Seien $\Theta^{(n)}$ und Θ durch (69) beziehungsweise (70) gegeben und seien $E_n, R_n^{(i)}, R^{(j)}, \mathcal{E}_n$ und $E_\infty, R_\infty^{(i)}, \mathcal{E}_\infty$ für $i = 1, 2, j = 2, 3$ wie in Abschnitt 3.2.1 beziehungsweise 3.2.2 definiert. Dann gilt für gegebene Trainingsdaten $\{x_s, y_s\}_{s=1}^S$, unter den Annahmen

1. $\alpha_i > 0$ für $i = 1, 2, 3, 4$ und $\tau_j > 0$ für $j = 1, 2$,
2. h ist stetig,
3. σ ist komponentenweise lipschitzstetig,
4. $\sigma(0) = 0$,
5. $\mathcal{L} \geq 0$ und \mathcal{L} ist stetig im ersten Argument,

dass Minimierer von \mathcal{E}_n und \mathcal{E}_∞ in $\Theta^{(n)}$ beziehungsweise Θ existieren. Weiterhin gilt, dass wenn $\theta^{(n)} \subseteq \Theta^{(n)}$ eine Folge von Minimierern von \mathcal{E}_n ist, dann ist $\{\theta^{(n)}\}_{n \in \mathbb{N}}$ relativ kompakt in $\Theta^{(n)}$ bezüglich d ,

$$\min_{\Theta^{(n)}} \mathcal{E}_n = \mathcal{E}_n(\theta^{(n)}) \rightarrow \min_{\Theta} \mathcal{E}_\infty \quad (85)$$

und jeder Häufungswert von $\theta^{(n)}$ ist ein Minimierer von \mathcal{E}_∞ .

Hierbei ist die relative Kompaktheit von $\{\theta^{(n)}\}_{n \in \mathbb{N}}$ in $\Theta^{(n)}$ bezüglich d so zu verstehen, dass im Sinne von Definition 1.2 jede Folge in $\{\theta^{(n)}\}_{n \in \mathbb{N}}$ eine Teilfolge besitzt, welche bezüglich d gegen ein Element aus Θ konvergiert. Der Beweis von Satz 3.3 wird mithilfe von Satz 1.8 geführt und ist in vier Teile unterteilt. In Abschnitt 3.4.1 zeigen wir die relative Kompaktheit von unter \mathcal{E}_n beschränkten Folgen in $\Theta^{(n)}$, in Abschnitt 3.4.2 zeigen wir die Existenz von Minimierern von \mathcal{E}_n in $\Theta^{(n)}$ und in Abschnitt 3.4.3 zeigen wir die

Γ -Konvergenz von \mathcal{E}_n gegen \mathcal{E}_∞ , was eine Voraussetzung von Satz 1.8 ist. In Abschnitt 3.4.4 werden dann mit Satz 1.8 dann endgültig die Behauptung von Satz 3.3 beweisen.

3.4.1 Relative Kompaktheit von unter \mathcal{E}_n beschränkten Folgen in $\Theta^{(n)}$

Sei $\{(K^{(n)}, b^{(n)}, W^{(n)}, c^{(n)})\}_{n \in \mathbb{N}} \subseteq \Theta^{(n)}$ eine unter \mathcal{E}_n beschränkte Folge. Also eine Folge, für die gilt

$$\sup_{n \in \mathbb{N}} \mathcal{E}_n(K^{(n)}, b^{(n)}, W^{(n)}, c^{(n)}) < \infty. \quad (86)$$

Für die Parameter $\{W^{(n)}\}_{n \in \mathbb{N}}$ und $\{c^{(n)}\}_{n \in \mathbb{N}}$ folgt die Beschränktheit in $\mathbb{R}^{m \times d}$, beziehungsweise \mathbb{R}^m direkt aus der Konstruktion der Regularisierer $R^{(3)}$ und $R^{(4)}$ zusammen mit (86). Da der \mathbb{R}^k endlichdimensional ist, ist Beschränktheit äquivalent zu relativer Kompaktheit. Damit muss nur noch die relative Kompaktheit von $\{K^{(n)}\}_{n \in \mathbb{N}}$ und $\{b^{(n)}\}_{n \in \mathbb{N}}$ gezeigt werden. Wir beginnen mit einem Lemma, welches besagt, dass für die Parameter $K^{(n)}$ und $b^{(n)}$ Beschränktheit unter den Regularisierern $R_n^{(1)}$ beziehungsweise $R_n^{(2)}$ in \mathbb{R} Beschränktheit von $K^{(n)}$ und $b^{(n)}$ in $L^\infty(\{t_i^{(n)}\}_{i=0}^{n-1}, \mu_n)$ impliziert.

Lemma 3.4

Sei $n \in \mathbb{N}$, $t_i := t_i^{(n)}$ und $f_n : \{t_i\}_{i=0}^{n-1}$. Dann gilt

$$\|f_n\|_{L^\infty(\{t_i\}_{i=0}^{n-1}, \mu_n)}^2 \leq 2 \left(\|f_n(t_0)\|^2 + n \sum_{j=1}^{n-1} \|f_n(t_j) - f_n(t_{j-1})\|^2 \right). \quad (87)$$

Beweis: Mit der Dreiecksungleichung folgt, dass für beliebiges $i = 1, \dots, n-1$

$$\|f_n(t_i) - f_n(t_0)\|^2 \leq \left(\sum_{j=1}^i \|f_n(t_j) - f_n(t_{j-1})\| \right)^2 \leq \left(\sum_{j=1}^{n-1} \|f_n(t_j) - f_n(t_{j-1})\| \right)^2 \quad (88)$$

gilt. Aus der Jensenschen Ungleichung (Rud87, Chapter 3, Theorem 3.2) folgt, dass für konvexe Funktionen g und für nichtnegative λ_j mit $\sum_{j=1}^{n-1} \lambda_j = 1$ gilt, dass

$$g \left(\sum_{j=1}^{n-1} \lambda_j x_j \right) \leq \sum_{j=1}^{n-1} \lambda_j g(x_j) \quad (89)$$

Wählen wir $\lambda_j = \frac{1}{n-1}$, $x_j = \|f_n(t_j) - f_n(t_{j-1})\|$ und $g(x) = x^2$, dann ergibt sich

$$\begin{aligned} \left(\frac{1}{n-1} \sum_{j=1}^{n-1} \|f_n(t_j) - f_n(t_{j-1})\| \right)^2 &\leq \frac{1}{n-1} \sum_{j=1}^{n-1} \|f_n(t_j) - f_n(t_{j-1})\|^2 \\ \Leftrightarrow \left(\sum_{j=1}^{n-1} \|f_n(t_j) - f_n(t_{j-1})\| \right)^2 &\leq (n-1) \sum_{j=1}^{n-1} \|f_n(t_j) - f_n(t_{j-1})\|^2. \end{aligned}$$

Also folgt für die rechte Seite von (88), dass

$$\left(\sum_{j=1}^{n-1} \|f_n(t_j) - f_n(t_{j-1})\| \right)^2 \leq (n-1) \sum_{j=1}^{n-1} \|f_n(t_j) - f_n(t_{j-1})\|^2 \leq n \sum_{j=1}^{n-1} \|f_n(t_j) - f_n(t_{j-1})\|^2. \quad (90)$$

gilt. Damit können wir die folgende Abschätzung für beliebiges $i = 1, \dots, n-1$ durchführen:

$$\begin{aligned} \|f_n(t_i)\|^2 &= \|f_n(t_i) - f_n(t_0) + f_n(t_0)\|^2 \\ &\leq (\|f_n(t_i) - f_n(t_0)\| + \|f_n(t_0)\|)^2 \\ &= \|f_n(t_i) - f_n(t_0)\|^2 + 2\|f_n(t_i) - f_n(t_0)\|\|f_n(t_0)\| + \|f_n(t_0)\|^2 \\ &\leq \|f_n(t_i) - f_n(t_0)\|^2 + \|f_n(t_i) - f_n(t_0)\|^2 + \|f_n(t_0)\|^2 + \|f_n(t_0)\|^2 \\ &= 2(\|f_n(t_i) - f_n(t_0)\|^2 + \|f_n(t_0)\|^2) \\ &\stackrel{(90)}{\leq} 2 \left(n \sum_{j=1}^{n-1} \|f_n(t_j) - f_n(t_{j-1})\|^2 + \|f_n(t_0)\|^2 \right), \end{aligned} \quad (91)$$

wobei wir in (91) ausgenutzt haben, dass $2ab \leq a^2 + b^2$ gilt. Bilden des Supremums liefert

$$\|f_n\|_{L^\infty(\{t_i\}_{i=0}^{n-1}, \mu_n)}^2 = \sup_{i=1, \dots, n} \|f_n(t_i)\|^2 \leq 2 \left(n \sum_{j=1}^{n-1} \|f_n(t_j) - f_n(t_{j-1})\|^2 + \|f_n(t_0)\|^2 \right), \quad (92)$$

woraus die Behauptung folgt. ■

Mithilfe von Lemma 3.4 lässt sich nun der folgende Satz beweisen.

Satz 3.5

Sei $n \in \mathbb{N}$, μ_n definiert wie in (64) und $f_n : \{t_i^{(n)}\}_{i=0}^{n-1} \mapsto \mathbb{R}^\kappa$. Wenn

$$\sup_{n \in \mathbb{N}} \left(\|f_n(0)\|^2 + n \sum_{j=0}^{n-1} \|f_n(t_j^{(n)}) - f_n(t_{j-1}^{(n)})\|^2 \right) < \infty \quad (93)$$

gilt, dann ist $\{(\mu_n, f_n)\}_{n \in \mathbb{N}}$ relativ kompakt in TL^2 und für jeden Häufungswert (μ, f) gilt, dass $\mu = \lambda_{[0,1]}$ und $f \in C^{0,\gamma}(\{t_i^{(n)}\}_{i=0}^{n-1}, \mathbb{R}^\kappa)$. Hierbei beschreibt $\lambda_{[0,1]}$ die Beschränkung des Lebesgue-Maßes auf $[0, 1]$ und ist definiert durch

$$\lambda_{[0,1]}(A) := \lambda(A \cap [0, 1]), \quad \forall A \subseteq \mathbb{R}. \quad (94)$$

Weiterhin existiert für jede konvergente Teilfolge $\{(\mu_m, f_m)\}_{m \in \mathbb{N}}$ von $\{(\mu_n, f_n)\}_{n \in \mathbb{N}}$ eine weitere Teilfolge $\{(\mu_l, f_l)\}_{l \in \mathbb{N}}$ und ein $f \in C^{0,\gamma}(\{t_i^{(n)}\}_{i=0}^{n-1}, \mathbb{R}^\kappa)$, sodass

$$\max_{i \in \{0, \dots, l-1\}} \|f_l(t_i^{(l)}) - f(t_i^{(l)})\| \rightarrow 0, \quad \text{für } l \rightarrow \infty. \quad (95)$$

gilt.

Beweis: Mit Lemma 3.4 und der Voraussetzung (93) folgt, dass ein $M \in \mathbb{R}$ existiert mit

$$\|f_n(0)\|^2 + n \sum_{j=0}^{n-1} \|f_n(t_j^{(n)}) - f_n(t_{j-1}^{(n)})\|^2 \leq M \text{ und } \|f_n\|_{L^\infty(\{t_i\}_{i=0}^{n-1}, \mu_n)} \leq M. \quad (96)$$

Sei \tilde{f}_n die stückweise konstante, kontinuierliche Fortsetzung von f_n auf \mathbb{R} , definiert durch

$$\tilde{f}_n(t) := \begin{cases} f_n(0), & t < 0, \\ f_n(t_i^{(n)}), & t \in [t_i^{(n)}, t_{i+1}^{(n)}), \text{ für } i = 0, \dots, n-1, \\ f_n(t_{n-1}^{(n)}), & t \geq 1. \end{cases} \quad (97)$$

Dann definieren wir g_n als die Glättung von f_n

$$g_n = \eta_{\varepsilon_n} * \tilde{f}_n, \quad (98)$$

wobei η_{ε_n} durch (28) mit $\varepsilon_n = \frac{1}{2n}$ definiert ist. Zunächst zeigen wir, dass g_n beschränkt in $H^1([0, 1]; \mathbb{R}^\kappa)$ ist. Dafür müssen wir zeigen, dass

$$\sup_{n \in \mathbb{N}} \|g_n\|_{H^1([0,1])} = \|\dot{g}_n\|_{L^2([0,1])} + \|g_n\|_{L^2([0,1])} < \infty \quad (99)$$

gilt. Aus (29) und Satz 2.3 folgt, dass g_n beschränkt in $L^\infty([0, 1]; \mathbb{R}^\kappa)$ ist. Da $[0, 1]$ kompakt ist, folgt daraus die Beschränktheit von g_n in $L^2([0, 1]; \mathbb{R}^\kappa)$. Es bleibt also zu zeigen, dass

$$\sup_{n \in \mathbb{N}} \|\dot{g}_n\|_{L^2([0,1])} < \infty \quad (100)$$

gilt. Sei dafür $t \in [t_i^{(n)}, t_i^{(n)} + \varepsilon_n]$ für $i \geq 1$. Unter Anbetracht dessen, dass

$$\begin{aligned} \int_{\mathbb{R}} \frac{d}{dt} \eta_{\varepsilon_n}(s-t) ds &= \int_{|s-t| < \varepsilon_n} \frac{-2(s-t)}{((s-t)^2 - \varepsilon_n^2)^2} \alpha \varepsilon_n e^{\frac{-1}{1-\frac{s-t}{\varepsilon_n}}} ds \\ &= \int_{|u| < \varepsilon_n} \frac{-2u}{(u^2 - \varepsilon_n^2)^2} \alpha \varepsilon_n e^{\frac{-1}{1-\frac{u}{\varepsilon_n}}} du \\ &= 0 \end{aligned}$$

gilt, wobei die letzte Gleichheit daraus folgt, dass eine ungerade Funktion über ein symmetrisch um 0 verteiltes Intervall integriert wird, können wir $\|\dot{g}(t)\|$ auch schreiben als

$$\begin{aligned} \|\dot{g}_n(t)\| &= \left\| \int_{\mathbb{R}} \frac{d}{dt} \eta_{\varepsilon_n}(s-t) \tilde{f}_n(s) ds - \tilde{f}_n(t) \int_{\mathbb{R}} \frac{d}{dt} \eta_{\varepsilon_n}(s-t) ds \right\| \\ &= \left\| \int_{\mathbb{R}} \frac{d}{dt} \eta_{\varepsilon_n}(s-t) (\tilde{f}_n(s) - \tilde{f}_n(t)) ds \right\| \\ &\leq \int_{\mathbb{R}} \left\| \frac{d}{dt} \eta_{\varepsilon_n}(s-t) \right\| \|\tilde{f}_n(s) - \tilde{f}_n(t)\| ds \\ &\leq \int_{t_{i-1}^{(n)}}^{t_{i+1}^{(n)}} \left\| \frac{d}{dt} \eta_{\varepsilon_n}(s-t) \right\| \|\tilde{f}_n(s) - \tilde{f}_n(t)\| ds, \end{aligned} \quad (101)$$

wobei die letzte Ungleichung aus der Wahl von $t \in [t_i^{(n)}, t_i^{(n)} + \varepsilon_n]$ und dem Träger von $\frac{d}{dt} \eta_{\varepsilon_n}$ folgt. Da $\frac{d}{dt} \eta_1(t)$ stetig und $[0, 1]$ kompakt ist, lässt sich eine Konstante $\beta \in \mathbb{R}$ finden, sodass

$\|\frac{d}{dt}\eta_1\|_{L^\infty([0,1])} = \|\frac{d}{dt}\eta_1\|_{L^\infty(\mathbb{R})} \leq \beta$. Da $\frac{d}{dt}\eta_{\varepsilon_n}(t) = \frac{1}{\varepsilon_n^2}\eta_1(\frac{t}{\varepsilon_n})$ gilt, können wir die Norm von $\frac{d}{dt}\eta_{\varepsilon_n}$ also abschätzen durch

$$\left\| \frac{d}{dt}\eta_{\varepsilon_n} \right\|_{L^\infty(\mathbb{R})} \leq \frac{\beta}{\varepsilon_n^2}. \quad (102)$$

Mit (101) folgt also für $t \in [t_i^{(n)}, t_i^{(n)} + \varepsilon_n]$ mit $i \geq 1$, dass

$$\begin{aligned} \|\dot{g}_n(t)\| &\leq \frac{\beta}{\varepsilon_n^2} \int_{t_{i-1}^{(n)}}^{t_{i+1}^{(n)}} \|\tilde{f}_n(s) - \tilde{f}_n(t)\| ds \\ &= 4n^2\beta \left(\int_{t_{i-1}^{(n)}}^{t_i^{(n)}} \|\tilde{f}_n(s) - \tilde{f}_n(t)\| ds + \int_{t_i^{(n)}}^{t_{i+1}^{(n)}} \|\tilde{f}_n(s) - \tilde{f}_n(t)\| ds \right) \\ &= 4n^2\beta \left(\frac{1}{n} \|f_n(t_i^{(n)}) - f_n(t_{i-1}^{(n)})\| + 0 \right) \\ &= 4n\beta \|f_n(t_i^{(n)}) - f_n(t_{i-1}^{(n)})\| \end{aligned} \quad (103)$$

gilt. Analog folgt für $t \in [t_i^{(n)} + \varepsilon_n, t_{i+1}^{(n)}]$ und $i \leq n-2$, dass

$$\|\dot{g}_n(t)\| \leq 4n\beta \|f_n(t_{i+1}^{(n)}) - f_n(t_i^{(n)})\| \quad (104)$$

gilt. Da $\dot{g}_n(t) = 0$ für $t \leq \varepsilon_n$ oder $t \geq 1 - \varepsilon_n$, ist

$$\begin{aligned} \|\dot{g}_n\|_{L^2([0,1])}^2 &= \int_{\varepsilon_n}^{1-\varepsilon_n} \|\dot{g}_n(t)\|^2 dt \\ &= \sum_{i=1}^{n-1} \left(\int_{t_{i-1}^{(n)} + \varepsilon_n}^{t_i^{(n)}} \|\dot{g}_n(t)\|^2 dt + \int_{t_i^{(n)}}^{t_i^{(n)} + \varepsilon_n} \|\dot{g}_n(t)\|^2 dt \right) \\ &\leq 16n^2\beta^2 \sum_{i=1}^{n-1} \|f_n(t_i^{(n)}) - f_n(t_{i-1}^{(n)})\|^2 \frac{1}{2n} + \|f_n(t_i^{(n)}) - f_n(t_{i-1}^{(n)})\|^2 \frac{1}{2n} \end{aligned} \quad (105)$$

$$= 16n\beta^2 \sum_{i=1}^{n-1} \|f_n(t_i^{(n)}) - f_n(t_{i-1}^{(n)})\|^2, \quad (106)$$

wobei die Abschätzung (105) aus (103) und (104) folgt. Zusammen mit der Annahme (93) folgt also, dass $\{\dot{g}_n\}_{n \in \mathbb{N}}$ beschränkt in $L^2([0,1])$, und somit $\{g_n\}_{n \in \mathbb{N}}$ beschränkt in $H^1([0,1])$ ist. Weiterhin folgt mit Satz 2.8 die Beschränktheit von $\{g_n\}_{n \in \mathbb{N}}$ in $C^{0, \frac{1}{2}}([0,1])$. Mit Satz 2.7 ergibt sich also die relative Kompaktheit von $\{g_n\}_{n \in \mathbb{N}}$ in $C^{0, \gamma}([0,1])$ für alle $\gamma \in (0, \frac{1}{2})$. Da die Topologie in $C^{0, \gamma}([0,1])$ stärker als die in $L^\infty([0,1])$ ist, ist $\{g_n\}_{n \in \mathbb{N}}$ also insbesondere auch relativ kompakt in $L^\infty([0,1])$. Das heißt, es existiert eine Teilfolge, die hier auch mit $(g_n)_{n \in \mathbb{N}}$ bezeichnet wird und ein $g \in C^{0, \gamma}([0,1])$, sodass $g_n \rightarrow g$ in $L^\infty([0,1])$. Jetzt muss nur noch gezeigt werden, dass $\tilde{f}_n \rightarrow g$ in $L^\infty([0,1])$. Dafür zeigen wir $\|\tilde{f}_n - g_n\|_{L^\infty([0,1])} \rightarrow 0$.

Sei $t \in [t_i, t_i + \varepsilon_n]$. Dann gilt, wenn wir ε_n und β wie bisher definieren, dass

$$\begin{aligned}
 \|\tilde{f}_n(t) - g_n(t)\| &= \left\| \int_{\mathbb{R}} \eta_{\varepsilon_n}(s-t) \tilde{f}_n(s) ds - f(t) \right\| \\
 &= \left\| \int_{\mathbb{R}} \eta_{\varepsilon_n}(s-t) \tilde{f}_n(s) ds - f(t) \int_{\mathbb{R}} \eta_{\varepsilon_n}(s-t) ds \right\| \\
 &= \left\| \int_{\mathbb{R}} \eta_{\varepsilon_n}(s-t) (\tilde{f}_n(s) - \tilde{f}_n(t)) ds \right\| \\
 &\leq \int_{\mathbb{R}} \|\eta_{\varepsilon_n}(s-t)\| \|(\tilde{f}_n(s) - \tilde{f}_n(t))\| ds \\
 &\leq \frac{\beta}{\varepsilon_n} \int_{t_{i-1}^{(n)}}^{t_{i+1}^{(n)}} \|\tilde{f}_n(s) - \tilde{f}_n(t)\| ds \\
 &= \begin{cases} 2\beta \|f_n(t_i^{(n)}) - f_n(t_{i-1}^{(n)})\|, & i \geq 1, \\ 0, & i = 0, \end{cases} \tag{108}
 \end{aligned}$$

wobei Gleichung (107) aus dem Fakt folgt, dass $\int_{\mathbb{R}} \eta_{\varepsilon_n}(t) dt = 1$. Analog bekommen wir für $t \in [t_i^{(n)} + \varepsilon_n, t_{i+1}^{(n)}]$ die Abschätzung

$$\|\tilde{f}_n(t) - g_n(t)\| \leq \begin{cases} 2\beta \|f_n(t_{i+1}^{(n)}) - f_n(t_i^{(n)})\|, & i \leq n-2, \\ 0, & i = n-1. \end{cases} \tag{109}$$

Also ist

$$\begin{aligned}
 \|\tilde{f}_n - g_n\|_{L^\infty([0,1])}^2 &\leq 4\beta^2 \sup_{i \in \{1, \dots, n-1\}} \|f_n(t_i^{(n)}) - f_n(t_{i-1}^{(n)})\|^2 \\
 &\leq 4\beta^2 \sum_{i=1}^{n-1} \|f_n(t_i^{(n)}) - f_n(t_{i-1}^{(n)})\|^2 \\
 &= \mathcal{O}\left(\frac{1}{n}\right), \tag{110}
 \end{aligned}$$

wobei (110) aus der Annahme (93) folgt. Somit gilt, dass $\tilde{f}_n \rightarrow g$ in $L^\infty([0,1])$ und durch die Kompaktheit von $[0,1]$ auch in $L^2([0,1])$. Da $\tilde{f}_n|_{[0,1]} = f_n \circ T_n$, folgt mit Satz 3.2 die Konvergenz in TL^2 und somit die Behauptung. ■

Damit können wir in folgendem Korollar das Ergebnis dieses Abschnittes festhalten.

Korollar 3.6

Seien $\mathcal{E}_n, \mathcal{E}_\infty, E_n, E_\infty, R_n^{(i)}, R_\infty^{(i)}$ und $R^{(j)}$, für $i = 1, 2, j = 3, 4$, definiert wie in den Abschnitten 3.2.1 und 3.2.2. Weiterhin seien $\alpha_i > 0$ für $i = 1, 2, 3, 4$, $\tau_j > 0$ für $j = 1, 2$, $h(x) \in \mathbb{R}$ für alle $x \in \mathbb{R}$, $\mathcal{L}(z, y) \in \mathbb{R}_{\geq 0}$ für alle $z, y \in \mathbb{R}^d$ und σ sei Lipschitzstetig mit $\sigma(0) = 0$. Sei $\theta^{(n)} = (K^{(n)}, b^{(n)}, W^{(n)}, c^{(n)}) \in (\Theta^{(n)})^{\mathbb{N}}$ mit

$$\sup_{n \in \mathbb{N}} \mathcal{E}_n(\theta^{(n)}) < \infty. \quad (111)$$

Dann existiert eine Teilfolge $\theta^{(n_m)}$ und $\theta = (K, b, W, c) \in \Theta$, sodass

$$d(\theta^{(n_m)}, \theta) \rightarrow 0, \quad \text{für } m \rightarrow \infty. \quad (112)$$

Weiterhin gilt, dass $\mathcal{E}_\infty(\theta) < \infty$.

Beweis: Die relative Kompaktheit von $\{W^{(n)}\}_{n \in \mathbb{N}}$ und $\{c^{(n)}\}_{n \in \mathbb{N}}$ folgt direkt aus der Konstruktion der Regularisierer. Die relative Kompaktheit von $\{K^{(n)}\}_{n \in \mathbb{N}}$ und $\{b^{(n)}\}_{n \in \mathbb{N}}$ folgt aus Satz 3.5 zusammen mit der Konstruktion der Regularisierer. Aus der Lipschitzstetigkeit von σ und dem Fakt, dass $\sigma(0) = 0$ ist, folgt, dass $X(1; x, K, b) < \infty$ für alle $x \in \mathbb{R}^d$. Also ist insbesondere auch $E_\infty(K, b, W, c; x, y)$ für alle (x, y) endlich, woraus zusammen mit der Konstruktion von Θ folgt, dass $\mathcal{E}_\infty(K, b, W, c) < \infty$. ■

3.4.2 Existenz von Minimierern von \mathcal{E}_n in $\Theta^{(n)}$

Die Existenz der Minimierer wird in dem folgenden Satz formalisiert.

Satz 3.7

Sei $n \in \mathbb{N}$. Seien weiterhin $\Theta^{(n)}, \mathcal{E}_n, E_n, R_n^{(i)}, R^{(j)}$ für $i = 1, 2, j = 3, 4$ definiert wie in den Abschnitten 3.2.1, 3.2.2 und 3.3. Dann existiert unter den Annahmen, dass σ und h stetig sind, $\sigma(0) = 0$ sowie \mathcal{L} nichtnegativ und stetig im ersten Argument ist, das Minimum von \mathcal{E}_n über $\Theta^{(n)}$ und ein $\theta^{(n)} \in \Theta^{(n)}$ mit

$$\min_{\theta \in \Theta^{(n)}} \mathcal{E}_n(\theta) = \mathcal{E}_n(\theta^{(n)}). \quad (113)$$

Beweis: Sei $n \in \mathbb{N}$ beliebig. Sei $(\theta_m^{(n)})_{m \in \mathbb{N}} = (K_m^{(n)}, b_m^{(n)}, W_m, c_m)_{m \in \mathbb{N}} \in (\Theta^{(n)})^{\mathbb{N}}$ eine Minimalfolge von \mathcal{E}_n , also eine Folge für die gilt

$$\mathcal{E}_n(\theta_m^{(n)}) \rightarrow \inf_{\theta \in \Theta^{(n)}} \mathcal{E}_n(\theta), \quad \text{für } m \rightarrow \infty. \quad (114)$$

Da $\sigma(0) = 0$, gilt $\mathcal{E}_n(0, 0, 0, 0) := \mathcal{E}_n(\underline{0}) = \sum_{s=1}^S \mathcal{L}(h(\underline{0}, y_i)) =: C < \infty$. Also existiert eine Teilfolge $(\theta_k^{(n)})_{k \in \mathbb{N}}$ von $\theta_m^{(n)}$, für die gilt

$$\sup_{k \in \mathbb{N}} \mathcal{E}_n(\theta_k^{(n)}) = \sup_{k \in \mathbb{N}} \mathcal{E}_n(K_k^{(n)}, b_k^{(n)}, W_k^{(n)}, c_k^{(n)}) \leq C. \quad (115)$$

Also ist insbesondere auch

$$\sup_{k \in \mathbb{N}} \max\{R_n^{(1)}(K_k^{(n)}), R_n^{(2)}(b_k^{(n)}), R^{(3)}(W_k), R^{(4)}(c_k)\} \leq C. \quad (116)$$

Aus der Konstruktion der Regularisierer und mit Lemma 3.4 folgt die Beschränktheit und dadurch die relative Kompaktheit von $\{K_k^{(n)}\}_{k \in \mathbb{N}}$, $\{b_k^{(n)}\}_{k \in \mathbb{N}}$, $\{W_k\}_{k \in \mathbb{N}}$ und $\{c_k\}_{k \in \mathbb{N}}$ in \mathbb{R}^κ . Somit folgt, dass \mathcal{E}_n leicht koerziv ist. Durch die Lipschitzstetigkeit von σ folgt induktiv die Stetigkeit von $X_i^{(n)}[x; \cdot, \cdot]$ für alle $i = 1, \dots, n-1$, in dem Sinne als dass für konvergente Folgen $K_l \rightarrow K$ und $b_l \rightarrow b$

$$\lim_{l \rightarrow \infty} X_i^{(n)}[x; K_l, b_l] \rightarrow X_i^{(n)}[x; K, b] \quad (117)$$

gilt. Weiterhin folgt durch die Stetigkeit von h und $\mathcal{L}(\cdot, y_i)$, dass \mathcal{E}_n stetig und somit auch unterhalbstetig ist. Mit Satz 1.6 folgt dann die Existenz eines Minimierers von \mathcal{E}_n in $\Theta^{(n)}$. ■

3.4.3 Γ-Konvergenz von \mathcal{E}_n gegen \mathcal{E}_∞

Der Beweis der Γ-Konvergenz ist in drei Teile unterteilt. Im ersten Teil werden wir einige vorbereitende Resultate zeigen, die für den zweiten und dritten Teil, in welchen wir die Liminf-Ungleichung beziehungsweise die Limsup-Ungleichung aus Definition 1.1 zeigen, als technische Hilfsmittel dienen. Dafür beginnen wir mit einem Lemma, welches die gleichmäßige Konvergenz des diskreten Modells $X_i^{(n)}$ gegen das kontinuierliche Modell X beschreibt.

Lemma 3.8

Seien $(K^{(n)})_{n \in \mathbb{N}} \in (L^2(\mu_n, \mathbb{R}^{d \times d}))^{\mathbb{N}}$ und $(b^{(n)})_{n \in \mathbb{N}} \in (L^2(\mu_n, \mathbb{R}^d))^{\mathbb{N}}$, mit $K^{(n)} \rightarrow K \in H^1([0, 1])$ und $b^{(n)} \rightarrow b \in H^1([0, 1])$. Weiterhin seien $R_n^{(i)}$ für $i = 1, 2$ definiert wie in Abschnitt 3.2.1. Unter der Annahme, dass σ Lipschitzstetig mit Lipschitzkonstante L_σ , $\sigma(0) = 0$, $\max\left\{\sup_{n \in \mathbb{N}} R_n^{(1)}(K^{(n)}), \sup_{n \in \mathbb{N}} R_n^{(2)}(b^{(n)})\right\} < \infty$ und $x \in \mathbb{R}$ ist, existiert eine von L_σ , $\|x\|$, $\|K\|_{L^\infty([0, 1])}$ und $\|b\|_{L^\infty([0, 1])}$ abhängige Konstante $C \in \mathbb{R}$, sodass $\|X(\cdot, x, K, b)\|_{L^\infty([0, 1])} \leq C$. Weiterhin gilt

$$\sup_{i \in \{0, 1, \dots, n-1\}} \sup_{t \in [t_i^{(n)}, t_{i+1}^{(n)}]} \|X(t; x, K, b) - X_i^{(n)}[x; K^{(n)}, b^{(n)}]\| \rightarrow 0, \quad \text{für } n \rightarrow \infty. \quad (118)$$

Beweis: Seien $X_i^{(n)} := X_i^{(n)}[x; K^{(n)}, b^{(n)}]$ und $X(t) = X(t; x, K, b)$. Dann ist

$$\begin{aligned} \left\| X\left(\frac{i}{n}\right) - X_i^{(n)} \right\| &= \left\| X\left(\frac{i-1}{n}\right) + \left(X\left(\frac{i}{n}\right) - X\left(\frac{i-1}{n}\right) \right) - X_{i-1}^{(n)} - \left(X_i^{(n)} - X_{i-1}^{(n)} \right) \right\| \\ &= \left\| X\left(\frac{i-1}{n}\right) + \int_{\frac{i-1}{n}}^{\frac{i}{n}} \dot{X}(t) dt - X_{i-1}^{(n)} - \left(X_i^{(n)} - X_{i-1}^{(n)} \right) \right\| \\ &\leq \left\| X\left(\frac{i-1}{n}\right) - X_{i-1}^{(n)} \right\| + \left\| \int_{\frac{i-1}{n}}^{\frac{i}{n}} \dot{X}(t) dt - \left(X_i^{(n)} - X_{i-1}^{(n)} \right) \right\|. \end{aligned} \quad (119)$$

Setzen wir die Definitionen von \dot{X} und $X_i^{(n)}$ ein, können wir den rechten Teil von (119) folgendermaßen abschätzen

$$\begin{aligned}
 & \left\| \int_{\frac{i-1}{n}}^{\frac{i}{n}} \dot{X}(t) dt - \left(X_i^{(n)} - X_{i-1}^{(n)} \right) \right\| \\
 &= \left\| \int_{\frac{i-1}{n}}^{\frac{i}{n}} \sigma(K(t)X(t) + b(t)) dt - \left(X_{i-1}^{(n)} + \left(\frac{i}{n} - \frac{i-1}{n} \right) \sigma \left(K_{i-1}^{(n)} X_{i-1}^{(n)} + b_{i-1}^{(n)} \right) - X_{i-1}^{(n)} \right) \right\| \\
 &= \left\| \int_{\frac{i-1}{n}}^{\frac{i}{n}} \sigma(K(t)X(t) + b(t)) - \sigma \left(K_{i-1}^{(n)} X_{i-1}^{(n)} + b_{i-1}^{(n)} \right) dt \right\| \\
 &\leq L_\sigma \int_{\frac{i-1}{n}}^{\frac{i}{n}} \left\| K(t)X(t) + b(t) - \left(K_{i-1}^{(n)} X_{i-1}^{(n)} + b_{i-1}^{(n)} \right) \right\| dt \\
 &\leq L_\sigma \int_{\frac{i-1}{n}}^{\frac{i}{n}} \left\| b(t) - b_{i-1}^{(n)} \right\| + \left\| K(t)X(t) - K_{i-1}^{(n)} X_{i-1}^{(n)} \right\| dt. \tag{120}
 \end{aligned}$$

Aus den Annahmen $\max \left\{ \sup_{n \in \mathbb{N}} R_n^{(1)}(K^{(n)}), \sup_{n \in \mathbb{N}} R_n^{(2)}(b^{(n)}) \right\} < \infty$, $K^{(n)} \rightarrow K$ und $b^{(n)} \rightarrow b$ folgt mit Lemma 3.4

$$\|K\|_{L^\infty([0,1])}, \|b\|_{L^\infty([0,1])} < \infty. \tag{121}$$

Im Folgenden zeigen wir zunächst

$$\sup_{n \in \mathbb{N}} \|X^{(n)}\|_{L^\infty(\mu_n)} < \infty. \tag{122}$$

Dafür führen wir die folgende Abschätzung durch

$$\begin{aligned}
 \|X_{i+1}^{(n)} - X_i^{(n)}\| &= \left\| \frac{1}{n} \sigma \left(K_i^{(n)} X_i^{(n)} + b_i^{(n)} \right) \right\| \\
 &\leq \frac{L_\sigma}{n} \left\| K_i^{(n)} X_i^{(n)} + b_i^{(n)} \right\| \\
 &\leq \frac{L_\sigma}{n} \left(\left\| K_i^{(n)} X_i^{(n)} \right\| + \left\| b_i^{(n)} \right\| \right) \\
 &\leq \frac{L_\sigma}{n} \left(\left\| K_i^{(n)} \right\| \left\| X_i^{(n)} \right\| + \left\| b_i^{(n)} \right\| \right) \\
 &\leq \frac{L_\sigma}{n} M_1 \left(\left\| X_i^{(n)} \right\| + 1 \right), \tag{123}
 \end{aligned}$$

wobei $M_1 := \max \left\{ \sup_{n \in \mathbb{N}} R_n^{(1)}(K^{(n)}), \sup_{n \in \mathbb{N}} R_n^{(2)}(b^{(n)}) \right\} < \infty$. Also gilt

$$\begin{aligned}
 \|X_{i+1}^{(n)}\| &= \left\| X_i^{(n)} + \frac{1}{n} \sigma \left(K_i^{(n)} X_i^{(n)} + b_i^{(n)} \right) \right\| \\
 &\leq \left\| X_i^{(n)} \right\| + \frac{1}{n} \left\| \sigma \left(K_i^{(n)} X_i^{(n)} + b_i^{(n)} \right) \right\| \\
 &\leq \left\| X_i^{(n)} \right\| + \frac{L_\sigma}{n} M_1 \left(\left\| X_i^{(n)} \right\| + 1 \right) \\
 &= \left\| X_i^{(n)} \right\| \left(1 + \frac{L_\sigma}{n} M_1 \right) + \frac{L_\sigma}{n} M_1. \tag{124}
 \end{aligned}$$

Induktion mittels der rekursiven Vorschrift (124) ergibt für alle $j = 1, \dots, n$

$$\begin{aligned}
 \|X_j^{(n)}\| &\leq \|x\| \left(1 + \frac{L_\sigma}{n} M_1\right)^j + \frac{L_\sigma}{n} M_1 \sum_{i=1}^j \left(1 + \frac{L_\sigma}{n} M_1\right)^{i-1} \\
 &\leq \|x\| \left(1 + \frac{L_\sigma}{n} M_1\right)^n + \frac{L_\sigma}{n} M_1 \sum_{i=1}^n \left(1 + \frac{L_\sigma}{n} M_1\right)^n \\
 &= \|x\| \left(1 + \frac{L_\sigma}{n} M_1\right)^n + L_\sigma M_1 \left(1 + \frac{L_\sigma}{n} M_1\right)^n \\
 &= (\|x\| + L_\sigma M_1) \left(1 + \frac{L_\sigma}{n} M_1\right)^n \\
 &\rightarrow (\|x\| + L_\sigma M_1) e^{L_\sigma M_1}, \quad \text{für } n \rightarrow \infty,
 \end{aligned}$$

woraus folgt

$$\sup_{n \in \mathbb{N}} \|X^{(n)}\|_{L^\infty(\mu_n)} < \infty.$$

Im Folgenden zeigen wir die Lipschitzstetigkeit von X . Dafür betrachten wir für $0 \leq t \leq 1$ den Ausdruck

$$\begin{aligned}
 \|X\|_{L^\infty([0,t])} &= \sup_{s \in [0,t]} \|X(s)\| \\
 &= \sup_{s \in [0,t]} \left\| \int_0^s \dot{X}(\xi) d\xi + x \right\| \\
 &= \sup_{s \in [0,t]} \left\| \int_0^s \sigma(K(\xi)X(\xi) + b(\xi)) d\xi + x \right\| \\
 &\leq \sup_{s \in [0,t]} L_\sigma \int_0^s \|K(\xi)X(\xi) + b(\xi)\| d\xi + \|x\| \\
 &\leq \sup_{s \in [0,t]} L_\sigma s (\|K\|_{L^\infty([0,s])} \|X\|_{L^\infty([0,s])} + \|b\|_{L^\infty([0,s])}) + \|x\| \\
 &\leq \sup_{s \in [0,t]} L_\sigma s (\|K\|_{L^\infty([0,1])} \|X\|_{L^\infty([0,s])} + \|b\|_{L^\infty([0,1])}) + \|x\| \\
 &\leq L_\sigma t \|X\|_{L^\infty([0,t])} \|K\|_{L^\infty([0,1])} + L_\sigma \|b\|_{L^\infty([0,1])} + \|x\| \tag{125}
 \end{aligned}$$

Mit $t = \frac{1}{2L_\sigma \|K\|_{L^\infty([0,1])}}$ und anschließendem Umstellen ergibt sich

$$\|X\|_{L^\infty([0,t])} \leq 2L_\sigma \|b\|_{L^\infty([0,1])} + 2\|x\| \tag{126}$$

Analog zu (125) gilt unter Berücksichtigung von (126)

$$\begin{aligned}
 \|X\|_{L^\infty([t,2t])} &\leq 2L_\sigma \|b\|_{L^\infty([0,1])} + 2\|X(t)\| \\
 &\leq 2L_\sigma \|b\|_{L^\infty([0,1])} + 2\|X\|_{L^\infty([0,t])} \\
 &\leq 2L_\sigma \|b\|_{L^\infty([0,1])} + 2(2L_\sigma \|b\|_{L^\infty([0,1])} + 2\|x\|) \\
 &= 2(2^2 - 1)L_\sigma \|b\|_{L^\infty([0,1])} + 2^2\|x\|.
 \end{aligned}$$

Induktiv folgt dementsprechend für $n \in \mathbb{N}$

$$\|X\|_{L^\infty([(n-1)t, nt])} \leq 2(2^n - 1)L_\sigma \|b\|_{L^\infty([0,1])} + 2^n\|x\|.$$

Somit ist für $N = \lceil 2L_\sigma \|K\|_{L^\infty([0,1])} \rceil$

$$\begin{aligned} \|X\|_{L^\infty([0,1])} &\leq \|X\|_{L^\infty([0,Nt])} \\ &\leq \max_{k \in \{1, \dots, N\}} \|X\|_{L^\infty([(k-1)t, kt])} \\ &\leq 2(2^N - 1)L_\sigma \|b\|_{L^\infty([0,1])} + 2^N \|x\|, \end{aligned}$$

wobei wir in der letzten Ungleichung ausgenutzt haben, dass der Ausdruck $\|X\|_{L^\infty([(k-1)t, kt])}$ monoton wachsend in k ist. Damit können wir $\|\dot{X}\|_{L^\infty([0,1])}$ abschätzen durch

$$\|\dot{X}\|_{L^\infty([0,1])} = \|\sigma(KX + b)\|_{L^\infty([0,1])} \leq L_\sigma (\|K\|_{L^\infty([0,1])} + \|X\|_{L^\infty([0,1])} + \|b\|_{L^\infty([0,1])})$$

Also ist \dot{X} beschränkt auf $[0, 1]$ und somit ist X Lipschitzstetig auf $[0, 1]$. Sei L_X die Lipschitzkonstante von X . Damit können wir den rechten Teil von (120) weiter abschätzen durch

$$\begin{aligned} \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|K(t)X(t) - K_{i-1}^{(n)}X_{i-1}^{(n)}\| dt &= \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|K(t)X(t) - K(t)X_{i-1}^{(n)} + K(t)X_{i-1}^{(n)} - K_{i-1}^{(n)}X_{i-1}^{(n)}\| dt \\ &\leq \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|X(t) - X_{i-1}^{(n)}\| \|K(t)\| dt + \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|K(t) - K_{i-1}^{(n)}\| \|X(t)\| dt \\ &\leq M_2 \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \|X(t) - X_{i-1}^{(n)}\| dt + \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|K(t) - K_{i-1}^{(n)}\| dt \right), \end{aligned} \tag{127}$$

wobei $M_2 := \max\{M_1, \|X^{(n)}\|_{L^\infty(\mu_n)}\}$. Im Weiteren schätzen wir den linken Teil von (127) ab durch

$$\begin{aligned} \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|X(t) - X_{i-1}^{(n)}\| dt &\leq \int_{\frac{i-1}{n}}^{\frac{i}{n}} \left\| X_{i-1}^{(n)} - X\left(\frac{i-1}{n}\right) \right\| + \left\| X\left(\frac{i-1}{n}\right) - X(t) \right\| dt \\ &= \frac{1}{n} \left\| X_{i-1}^{(n)} - X\left(\frac{i-1}{n}\right) \right\| + \int_{\frac{i-1}{n}}^{\frac{i}{n}} \left\| X\left(\frac{i-1}{n}\right) - X(t) \right\| dt \\ &\leq \frac{1}{n} \left\| X_{i-1}^{(n)} - X\left(\frac{i-1}{n}\right) \right\| + L_X \int_{\frac{i-1}{n}}^{\frac{i}{n}} \left| t - \frac{i-1}{n} \right| dt \\ &= \frac{1}{n} \left\| X_{i-1}^{(n)} - X\left(\frac{i-1}{n}\right) \right\| + L_X \int_0^{\frac{1}{n}} u du \\ &= \frac{1}{n} \left\| X_{i-1}^{(n)} - X\left(\frac{i-1}{n}\right) \right\| + \frac{L_X}{2n^2}. \end{aligned} \tag{128}$$

Das Kombinieren von (119), (120), (127) und (128) ergibt

$$\begin{aligned}
\left\| X\left(\frac{i}{n}\right) - X_i^{(n)} \right\| &\stackrel{(119)}{\leq} \left\| X\left(\frac{i-1}{n}\right) - X_{i-1}^{(n)} \right\| + \left\| \int_{\frac{i-1}{n}}^{\frac{i}{n}} \dot{X}(t) dt - \left(X_i^{(n)} - X_{i-1}^{(n)} \right) \right\| \\
&\stackrel{(120)}{\leq} \left\| X\left(\frac{i-1}{n}\right) - X_{i-1}^{(n)} \right\| \\
&\quad + L_\sigma \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|b(t) - b_i^{(n)}\| dt + \left\| K(t)X(t) - K_{i-1}^{(n)}X_{i-1}^{(n)} \right\| dt \\
&\stackrel{(127)}{\leq} \left\| X\left(\frac{i-1}{n}\right) - X_{i-1}^{(n)} \right\| + L_\sigma \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|b(t) - b_i^{(n)}\| dt \\
&\quad + L_\sigma M_2 \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \|X(t) - X_{i-1}^{(n)}\| dt + \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|K(t) - K_{i-1}^{(n)}\| dt \right) \\
&\stackrel{(128)}{\leq} \left\| X\left(\frac{i-1}{n}\right) - X_{i-1}^{(n)} \right\| + L_\sigma \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|b(t) - b_i^{(n)}\| dt \\
&\quad + L_\sigma M_2 \left(\frac{1}{n} \left\| X_{i-1}^{(n)} - X\left(\frac{i-1}{n}\right) \right\| + \frac{L_X}{2n^2} + \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|K(t) - K_{i-1}^{(n)}\| dt \right) \\
&= \left(1 + \frac{L_\sigma M_2}{n} \right) \left\| X\left(\frac{i-1}{n}\right) - X_{i-1}^{(n)} \right\| + L_\sigma \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|b(t) - b_i^{(n)}\| dt \\
&\quad + L_\sigma M_2 \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|K(t) - K_{i-1}^{(n)}\| dt + \frac{L_X L_\sigma M_2}{2n^2}. \tag{129}
\end{aligned}$$

Für beliebiges $k \in \{0, 1, \dots, n\}$ folgt dementsprechend induktiv

$$\begin{aligned}
\left\| X\left(\frac{k}{n}\right) - X_k^{(n)} \right\| &\leq L_\sigma M_2 \sum_{i=1}^n \left(1 + \frac{L_\sigma M_2}{n} \right)^{n-i} \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|K(t) - K_{i-1}^{(n)}\| dt \\
&+ L_\sigma \sum_{i=1}^n \left(1 + \frac{L_\sigma M_2}{n} \right)^{n-i} \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|b(t) - b_{i-1}^{(n)}\| dt + \frac{L_X L_\sigma M_2}{2n^2} \sum_{i=1}^n \left(1 + \frac{L_\sigma M_2}{n} \right)^{n-i}. \tag{130}
\end{aligned}$$

Aus der youngschen Ungleichung für Produkte (HJG34, S. 111, 4.8) folgt, dass für nichtnegative reelle Zahlen a, b und für $p, q > 1$ mit $\frac{1}{p} + \frac{1}{q} = 1$

$$ab \leq \frac{a^p}{p} + \frac{b^q}{q} \tag{131}$$

gilt. Für beliebiges $\varepsilon > 0$, $n > 0$ und $p = q = 2$, $a' = \frac{a\sqrt{n}}{\sqrt{\varepsilon}}$, $b' = \frac{\sqrt{\varepsilon}b'}{\sqrt{n}}$ folgt die Ungleichung

$$a'b' \leq \frac{a'^2 n}{2\varepsilon} + \frac{\varepsilon b'^2}{2n}. \tag{132}$$

(132) auf die Produkte in den Summen der ersten zwei Summanden aus (130) angewandt

ergibt für beliebiges $\varepsilon > 0$

$$\begin{aligned} \left\| X\left(\frac{k}{n}\right) - X_k^{(n)} \right\| &\leq \frac{\varepsilon L_\sigma M_2}{2n} \sum_{i=1}^n \left(1 + \frac{L_\sigma M_2}{n}\right)^{2(n-i)} + \frac{L_\sigma M_2 n}{2\varepsilon} \sum_{i=1}^n \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \|K(t) - K_{i-1}^{(n)}\| dt \right)^2 \\ &\quad + \frac{\varepsilon L_\sigma}{2n} \sum_{i=1}^n \left(1 + \frac{L_\sigma M_2}{n}\right)^{2(n-i)} + \frac{L_\sigma n}{2\varepsilon} \sum_{i=1}^n \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \|b(t) - b_{i-1}^{(n)}\| dt \right)^2 \\ &\quad \frac{L_X L_\sigma M_2}{2n^2} \sum_{i=0}^{n-1} \left(1 + \frac{L_\sigma M_2}{n}\right). \end{aligned} \quad (133)$$

Die Jensen'sche Ungleichung für Integrale (Rud87, Ch 3, Thm 3.2) besagt, dass für Lebesgue-integrierbare Funktionen $f : [a, b] \rightarrow \mathbb{R}$ und für konvexe Funktionen $\varphi : \mathbb{R} \rightarrow \mathbb{R}$

$$\varphi\left(\frac{1}{b-a} \int_a^b f(t) dt\right) \leq \frac{1}{b-a} \int_a^b \varphi(f(t)) dt \quad (134)$$

gilt. Mit $a = \frac{i-1}{n}$, $b = \frac{i}{n}$, $\varphi(x) = x^2$ und $f(t) = \|K(t) - K_{i-1}^{(n)}\|^2$ folgt

$$n^2 \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \|K(t) - K_{i-1}^{(n)}\| dt \right)^2 \leq n \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|K(t) - K_{i-1}^{(n)}\|^2 dt \quad (135)$$

und somit

$$n \sum_{i=1}^n \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \|K(t) - K_{i-1}^{(n)}\| dt \right)^2 \leq \sum_{i=1}^n \int_{\frac{i-1}{n}}^{\frac{i}{n}} \|K(t) - K_{i-1}^{(n)}\|^2 dt \rightarrow 0, \quad \text{für } n \rightarrow \infty, \quad (136)$$

wobei die Konvergenz aus (136) durch die Annahme $K^{(n)} \rightarrow K$ folgt. Für $b^{(n)}$ folgt analog

$$n \sum_{i=1}^n \left(\int_{\frac{i-1}{n}}^{\frac{i}{n}} \|b(t) - b_{i-1}^{(n)}\| dt \right)^2 \rightarrow 0, \quad \text{für } n \rightarrow \infty. \quad (137)$$

Um die Behauptung

$$\sup_{k \in \{0, 1, \dots, n\}} \left\| X\left(\frac{k}{n}\right) - X_k^{(n)} \right\| \rightarrow 0, \quad \text{für } n \rightarrow \infty$$

zu beweisen, reicht es also die folgenden zwei Aussagen zu zeigen.

1. $\frac{1}{n^2} \sum_{i=0}^{n-1} \left(1 + \frac{L_\sigma M_2}{n}\right)^i \rightarrow 0$ für $n \rightarrow \infty$,
2. $\sup_{n \in \mathbb{N}} \frac{1}{n} \sum_{i=1}^n \left(1 + \frac{L_\sigma M_2}{n}\right)^{2(n-i)} < \infty$.

Aussage 1 folgt aus

$$\begin{aligned} 0 &\leq \frac{1}{n^2} \sum_{i=0}^{n-1} \left(1 + \frac{L_\sigma M_2}{n}\right)^i \\ &\leq \frac{1}{n^2} \sum_{i=0}^{n-1} \left(1 + \frac{L_\sigma M_2}{n}\right)^n \\ &= \frac{1}{n} \left(1 + \frac{L_\sigma M_2}{n}\right)^n \rightarrow 0 \cdot e^{L_\sigma M_2} = 0, \quad \text{für } n \rightarrow \infty. \end{aligned}$$

Aussage 2 folgt ähnlich:

$$\frac{1}{n} \sum_{i=1}^n \left(1 + \frac{L_\sigma M_2}{n}\right)^{2(n-i)} \leq \left(1 + \frac{L_\sigma M_2}{n}\right)^{2n} \rightarrow e^{2L_\sigma M_2}, \quad \text{für } n \rightarrow \infty.$$

Damit haben wir also alle Terme aus (133) ausreichend abgeschätzt, sodass für eine Folge $\varepsilon := \varepsilon_n \rightarrow 0^+$, welche langsam genug konvergiert,

$$\sup_{k \in \{0, 1, \dots, n\}} \left\| X\left(\frac{k}{n}\right) - X_k^{(n)} \right\| \rightarrow 0, \quad \text{für } n \rightarrow \infty \quad (138)$$

gilt. Mit

$$\begin{aligned} \sup_{t \in [t_k^{(n)}, t_{k+1}^{(n)}]} \left\| X(t) - X_k^{(n)} \right\| &\leq \sup_{t \in [t_k^{(n)}, t_{k+1}^{(n)}]} \left(\|X(t) - X(t_k)\| + \|X(t_k) - X_k^{(n)}\| \right) \\ &\leq \sup_{t \in [t_k^{(n)}, t_{k+1}^{(n)}]} \left(L_x |t - t_k| + \|X(t_k) - X_k^{(n)}\| \right) \\ &= \frac{L_x}{n} + \|X(t_k) - X_k^{(n)}\| \rightarrow 0, \quad \text{für } n \rightarrow \infty, \end{aligned}$$

wobei aus (138) folgt, dass die Konvergenz gleichmäßig über $k \in \{0, 1, \dots, k\}$ ist, folgt die Behauptung. ■

Mithilfe des Lemmas 3.8 können wir den folgenden Satz beweisen, welcher besagt, dass die Fehlerfunktion des diskreten Modells gegen die Fehlerfunktion des kontinuierlichen Modells konvergiert.

Satz 3.9

Sei $(\Theta^{(n)})^{\mathbb{N}} \ni (K^{(n)}, b^{(n)}, W^{(n)}, c^{(n)})_{n \in \mathbb{N}} =: (\theta^{(n)})_{n \in \mathbb{N}} \rightarrow \theta := (K, b, W, c) \in \Theta$.

Wenn zusätzlich zu den Annahmen aus Satz 3.3 angenommen wird

$$\max\left\{\sup_{n \in \mathbb{N}} R_n^{(1)}(K^{(n)}), \sup_{n \in \mathbb{N}} R_n^{(2)}(b^{(n)})\right\} < \infty, \quad (139)$$

dann gilt

$$\lim_{n \rightarrow \infty} E_n(\theta^{(n)}; x, y) = E_\infty(\theta, x, y). \quad (140)$$

Beweis : Mit den Annahmen des zu beweisenden Satzes gilt

$$\begin{aligned} \lim_{n \rightarrow \infty} E_n(\theta^{(n)}; x, y) &= \lim_{n \rightarrow \infty} \mathcal{L} \left(h(W^{(n)} X_n^{(n)}[x; K^{(n)}, b^{(n)}] + c^{(n)}), y \right) \\ &= \mathcal{L} \left(\lim_{n \rightarrow \infty} h(W^{(n)} X_n^{(n)}[x; K^{(n)}, b^{(n)}] + c), y \right) \\ &= \mathcal{L} \left(h \left(\lim_{n \rightarrow \infty} W^{(n)} X_n^{(n)}[x; K^{(n)}, b^{(n)}] + c^{(n)} \right), y \right) \\ &= \mathcal{L}(h(WX(1; x, K, b) + c), y) = E_\infty(\theta; x, y), \end{aligned}$$

woraus die Behauptung folgt. Dabei wurde bei der zweiten Gleichheit die Stetigkeit von \mathcal{L} im ersten Argument, bei der dritten Gleichheit die Stetigkeit von h und bei der vierten Gleichheit Lemma 3.8 ausgenutzt. \blacksquare

Im Folgenden ein Lemma, in dem zwei Ungleichungen bewiesen werden, welche wir brauchen um den darauffolgenden Satz zu beweisen.

Lemma 3.10

Sei $\varepsilon_n \rightarrow 0^+$, $\delta, \delta' > 0$. Dann gelten die folgenden zwei Aussagen.

1. $\int_{\delta'}^{1-\delta'} \|\eta_\delta * g(t) - \eta_\delta * g(t - \varepsilon_n)\|^2 dt \leq \int_{\varepsilon_n}^1 \|g(t) - g(t - \varepsilon_n)\|^2 dt$,
für beliebiges $g \in L^2([0, 1], \mathbb{R}^\kappa)$ und $\varepsilon_n + \delta \leq \delta'$, wobei η_δ der Standardglättungskern aus 28 ist.
2. $\int_{2\delta'}^{1-2\delta'} \|\dot{g}(t)\|^2 dt \leq \liminf_{n \rightarrow \infty} \frac{1}{\varepsilon_n^2} \int_{2\delta'}^{1-2\delta'} \|g_n(t) - g_n(t - \varepsilon_n)\|^2 dt$,
für beliebige $g, g_n \in C^\infty([\delta', 1 - \delta']; \mathbb{R}^\kappa)$ mit $\dot{g}_n \rightarrow \dot{g}$ in $L^\infty([\delta', 1 - \delta']; \mathbb{R}^\kappa)$ und $\sup_{n \in \mathbb{N}} \|\ddot{g}_n\|_{L^\infty([\delta', 1 - \delta'])} < \infty$.

Beweis : Seien die Bedingungen von 1 erfüllt. Dann gilt

$$\begin{aligned} &\left(\int_{\delta'}^{1-\delta'} \|\eta_\delta * g(t) - \eta_\delta * g(t - \varepsilon_n)\|^2 dt \right)^{\frac{1}{2}} \\ &= \left(\int_{\delta'}^{1-\delta'} \left\| \int_{-\delta}^{\delta} \eta_\delta(s) (g(t-s) - g(t - \varepsilon_n - s)) ds \right\|^2 dt \right)^{\frac{1}{2}} \\ &\leq \left(\int_{\delta'}^{1-\delta'} \left(\int_{-\delta}^{\delta} \|\eta_\delta(s) (g(t-s) - g(t - \varepsilon_n - s))\| ds \right)^2 dt \right)^{\frac{1}{2}}. \quad (141) \end{aligned}$$

Die Minkowski-Ungleichung für Integrale (HJG34, Thm. 200-202) besagt, dass für $p < \infty$, Maßräume (S_1, μ_1) , (S_2, μ_2) und messbare Funktionen $F : S_1 \times S_2 \rightarrow \mathbb{K}$

$$\left[\int_{S_2} \left(\int_{S_1} |F(s, t)| d\mu(s) \right)^p d\mu_2(t) \right]^{\frac{1}{p}} \leq \int_{S_1} \left(\int_{S_2} |F(s, t)|^p d\mu_2(t) \right)^{\frac{1}{p}} d\mu_1(s) \quad (142)$$

gilt. Mit $p = 2$, $(S_1, \mu) = ((\delta', 1 - \delta'), \lambda)$, $(S_2, \mu) = ((-\delta', \delta'), \lambda)$ und $F(s, t) = \|\eta_\delta(s)(g(t - s) - g(t - \varepsilon_n - s))\|$ folgt für (141)

$$\begin{aligned} & \left(\int_{\delta'}^{1-\delta'} \left(\int_{-\delta}^{\delta} \|\eta_\delta(s)(g(t - s) - g(t - \varepsilon_n - s))\| ds \right)^2 dt \right)^{\frac{1}{2}} \\ & \leq \int_{-\delta}^{\delta} \left(\int_{\delta'}^{1-\delta'} \eta_\delta(s)^2 \|g(t - s) - g(t - \varepsilon_n - s)\|^2 dt \right)^{\frac{1}{2}} ds \\ & = \int_{-\delta}^{\delta} \eta_\delta(s) \left(\int_{\delta'}^{1-\delta'} \|g(t - s) - g(t - \varepsilon_n - s)\|^2 dt \right)^{\frac{1}{2}} ds \\ & = \int_{-\delta}^{\delta} \eta_\delta(s) \left(\int_{\delta'-s}^{1-\delta'-s} \|g(u) - g(u - \varepsilon_n)\|^2 du \right)^{\frac{1}{2}} ds \end{aligned} \quad (143)$$

$$\leq \int_{-\delta}^{\delta} \eta_\delta(s) \left(\int_{\varepsilon_n}^1 \|g(u) - g(u - \varepsilon_n)\|^2 du \right)^{\frac{1}{2}} ds \quad (144)$$

$$\begin{aligned} & = \int_{\varepsilon_n}^1 \|g(t) - g(t - \varepsilon_n)\|^2 dt \int_{-\delta}^{\delta} \eta_\delta(s) ds \\ & = \int_{\varepsilon_n}^1 \|g(t) - g(t - \varepsilon_n)\|^2 dt, \end{aligned} \quad (145)$$

wobei in (143) $u = t - s$ substituiert, in (144) die Voraussetzung $\varepsilon_n + \delta \leq \delta'$ und in (145) die Normiertheit des Glättungskerns ausgenutzt wurde. Damit folgt die Ungleichung 1. Seien nun die Bedingungen von 2 erfüllt. Mithilfe der Taylor-Formel mit Lagrangeschem Restglied (For07, §22, Satz 2) lässt sich g_n entwickelt im Entwicklungspunkt t und ausgewertet in $t - \varepsilon_n$ schreiben als

$$\begin{aligned} g_n(t - \varepsilon_n) &= T_1 g_n(t - \varepsilon_n; t) + R_1 g_n(t - \varepsilon_n; t) \\ &= g_n(t) + \dot{g}_n(t)(t - \varepsilon_n - t) + R_1 g_n(t - \varepsilon_n; t) \\ &= g_n(t) - \dot{g}_n(t)\varepsilon_n + \ddot{g}_n(z)\varepsilon_n^2, \end{aligned}$$

für ein $z \in [t - \varepsilon_n, t]$. Umstellen ergibt

$$g_n(t) - g_n(t - \varepsilon_n) = \dot{g}_n(t)\varepsilon_n - \ddot{g}_n(z)\varepsilon_n^2.$$

Durch Bildung der Norm, Division durch $\varepsilon_n \neq 0$ und Anwendung der umgekehrten Dreiecksungleichung bekommen wir

$$\frac{\|g_n(t) - g_n(t - \varepsilon_n)\|}{\varepsilon_n} = \|\dot{g}_n(t) - \ddot{g}_n(z)\varepsilon_n\| \geq \|\dot{g}_n(t)\| - \varepsilon_n \|\ddot{g}_n(z)\|.$$

Da z aus dem Intervall $[t - \varepsilon_n, t]$ kommt, können wir für $t \in [2\delta', 1 - 2\delta']$ und $\varepsilon_n < \delta'$ die Norm von $\ddot{g}_n(z)$ durch die $L^\infty([\delta', 1 - \delta'])$ -Norm von \ddot{g}_n abschätzen, welche nach Voraussetzung für alle $n \in \mathbb{N}$ endlich ist. Also erhalten wir

$$\begin{aligned} \frac{\|g_n(t) - g_n(t - \varepsilon_n)\|}{\varepsilon_n} &\geq \|\dot{g}_n(t)\| - \varepsilon_n \|\ddot{g}_n\|_{L^\infty([\delta', 1 - \delta'])} \\ \iff \|\dot{g}_n(t)\| &\leq \frac{\|g_n(t) - g_n(t - \varepsilon_n)\|}{\varepsilon_n} + \varepsilon_n \|\ddot{g}_n\|_{L^\infty([\delta', 1 - \delta'])}. \end{aligned} \quad (146)$$

Sei $\xi > 0$ beliebig. Dann folgt aus der youngschen Ungleichung für beliebige $a, b \geq 0$

$$2ab \leq \xi a^2 + \frac{1}{\xi} b^2$$

und somit

$$\begin{aligned} (a+b)^2 &= a^2 + 2ab + b^2 \\ &\leq a^2 + \xi a^2 + \frac{1}{\xi} b^2 + b^2 = (1+\xi)a^2 + \left(1 + \frac{1}{\xi}\right)b^2. \end{aligned}$$

Wir setzen $C_\xi := \left(1 + \frac{1}{\xi}\right)$. Damit folgt mit $a = \frac{\|g_n(t) - g_n(t - \varepsilon_n)\|}{\varepsilon_n}$ und $b = \varepsilon_n \|\ddot{g}\|_{L^\infty([\delta', 1 - \delta'])}$ in (146)

$$\|\dot{g}_n(t)\|^2 \leq (1+\xi) \frac{\|g_n(t) - g_n(t - \varepsilon_n)\|^2}{\varepsilon_n^2} + C_\xi \varepsilon_n^2 \|\ddot{g}_n\|_{L^\infty([\delta', 1 - \delta'])}^2.$$

Durch die gleichmäßige Konvergenz von \dot{g}_n gegen \dot{g} folgt

$$\begin{aligned} \int_{2\delta'}^{1-2\delta'} \|\dot{g}(t)\|^2 dt &= \lim_{n \rightarrow \infty} \int_{2\delta'}^{1-2\delta'} \|\dot{g}_n(t)\|^2 dt \\ &\leq \liminf_{n \rightarrow \infty} \int_{2\delta'}^{1-2\delta'} (1+\xi) \frac{\|g_n(t) - g_n(t - \varepsilon_n)\|^2}{\varepsilon_n^2} + C_\xi \varepsilon_n^2 \|\ddot{g}_n\|_{L^\infty([\delta', 1 - \delta'])}^2 dt \\ &= (1+\xi) \liminf_{n \rightarrow \infty} \int_{2\delta'}^{1-2\delta'} \frac{\|g_n(t) - g_n(t - \varepsilon_n)\|^2}{\varepsilon_n^2} dt \\ &\quad + \lim_{n \rightarrow \infty} C_\xi \varepsilon_n^2 \|\ddot{g}_n\|_{L^\infty([\delta', 1 - \delta'])}^2 (1 - 4\delta') \\ &= (1+\xi) \liminf_{n \rightarrow \infty} \int_{2\delta'}^{1-2\delta'} \frac{\|g_n(t) - g_n(t - \varepsilon_n)\|^2}{\varepsilon_n^2} dt. \end{aligned}$$

Da $\xi > 0$ beliebig war, können wir guten Gewissens $\xi \rightarrow 0$ gehen lassen und somit folgt die Behauptung. ■

Damit können wir den folgenden Satz beweisen, welcher uns als technisches Hilfsmittel im Beweis von der Liminf-Ungleichung dienen wird.

Satz 3.11

Sei $(f_n)_{n \in \mathbb{N}} \in (L^2([0, 1]; \mathbb{R}^\kappa))^{\mathbb{N}}$, $f \in L^2([0, 1]; \mathbb{R}^\kappa)$, mit $f_n \rightarrow f$ in $L^2([0, 1]; \mathbb{R}^\kappa)$ und $\varepsilon_n \rightarrow 0^+$. Wenn

$$\liminf_{n \rightarrow \infty} \frac{1}{\varepsilon_n} \int_{\varepsilon_n}^1 \|f_n(t) - f_n(t - \varepsilon_n)\|^2 dt < \infty \quad (147)$$

gilt, dann ist $f \in H^1([0, 1]; \mathbb{R}^\kappa)$ und

$$\liminf_{n \rightarrow \infty} \frac{1}{\varepsilon_n} \int_{\varepsilon_n}^1 \|f_n(t) - f_n(t - \varepsilon_n)\|^2 dt \geq \int_0^1 \|\dot{f}(t)\|^2 dt. \quad (148)$$

Beweis: Seien $\delta, \delta' > 0$, mit $\varepsilon_n + \delta < \delta'$. Dann folgt mit Ungleichung 1 aus Lemma 3.10 und der Voraussetzung (147), dass ein $M \in \mathbb{R}$ existiert, sodass

$$\liminf_{n \rightarrow \infty} \int_{\delta'}^{1-\delta'} \|\eta_\delta * f_n(t) - \eta_\delta * f_n(t - \varepsilon_n)\|^2 dt \leq M \quad (149)$$

gilt, wobei η_δ den Standardglättungskern beschreibt. Zunächst halten wir mithilfe der Faltungsungleichung von Young (Wer11, Satz II.4.4) die folgenden zwei Aussagen fest.

1. $\left\| \frac{d}{dt} \eta_\delta * f - \frac{d}{dt} \eta_\delta * f_n \right\|_{L^\infty([\delta', 1-\delta'])} = \left\| \frac{d}{dt} \eta_\delta * (f - f_n) \right\|_{L^\infty([\delta', 1-\delta'])}$
 $\leq \left\| \frac{d}{dt} \eta_\delta \right\|_{L^2(\mathbb{R})} \|f - f_n\|_{L^2([0,1])}$
2. $\left\| \frac{d^2}{dt^2} \eta_\delta * f_n \right\|_{L^\infty([\delta', 1-\delta'])} \leq \left\| \frac{d^2}{dt^2} \eta_\delta \right\|_{L^2(\mathbb{R})} \|f_n\|_{L^2([0,1])}$

Damit können wir Ungleichung 2 aus Lemma 3.10 auf $g := \eta_\delta * f$ und $g_n := \eta_\delta * f_n$ anwenden. Also gilt

$$\begin{aligned} \int_{2\delta'}^{1-2\delta'} \left\| \frac{d}{dt} \eta_\delta * f(t) \right\|^2 dt &\leq \liminf_{n \rightarrow \infty} \frac{1}{\varepsilon_n^2} \int_{2\delta'}^{1-2\delta'} \|\eta_\delta * f_n(t) - \eta_\delta * f_n(t - \varepsilon_n)\|^2 dt \\ &\leq \liminf_{n \rightarrow \infty} \frac{1}{\varepsilon_n^2} \int_{\delta'}^{1-\delta'} \|\eta_\delta * f_n(t) - \eta_\delta * f_n(t - \varepsilon_n)\|^2 dt \\ &\leq M, \end{aligned}$$

wobei die letzte Ungleichung aus 149 folgt. Also ist $\{\frac{d}{dt} \eta_\delta * f : \delta > 0\}$ beschränkt in $L^2([2\delta', 1 - 2\delta'])$. Da $L^2([2\delta', 1 - 2\delta'])$ reflexiv ist (Wer11, S. 105) und jede beschränkte Folge in einem reflexiven Raum eine schwach konvergente Teilfolge besitzt (Wer11, Theorem III.3.7), existiert also ein $h \in L^2([2\delta', 1 - 2\delta'])$ und eine Folge $\delta_m \rightarrow 0^+$ mit $\frac{d}{dt} \eta_{\delta_m} * f \xrightarrow{m \rightarrow \infty} h$. Somit folgt mittels der Definition der schwachen Konvergenz in Kombination mit der Selbstdualität von L^2 , welche eine Konsequenz aus dem Darstellungssatz von Frechet-Riesz (Wer11, Theorem V.3.6) ist, dass für alle $\varphi \in L^2([2\delta', 1 - 2\delta'])$ und somit insbesondere für alle $\varphi \in C_c^1([2\delta', 1 - 2\delta'])$ gilt, dass

$$\int_{2\delta'}^{1-2\delta'} \varphi(t) \frac{d}{dt} (\eta_{\delta_m} * f)(t) dt \rightarrow \int_{2\delta'}^{1-2\delta^*} \varphi(t) h(t) dt, \quad \text{für } m \rightarrow \infty. \quad (150)$$

Mittels partieller Integration unter Ausnutzung des kompakten Trägers von φ können wir weiter folgern

$$\int_{2\delta'}^{1-2\delta'} \varphi(t) \frac{d}{dt} (\eta_{\delta_m} * f)(t) dt = - \int_{2\delta'}^{1-2\delta'} (\eta_{\delta_m} * f)(t) \frac{d}{dt} \varphi(t) dt. \quad (151)$$

Weiterhin folgt mit Satz 2.3

$$- \int_{2\delta'}^{1-2\delta'} (\eta_{\delta_m} * f)(t) \frac{d}{dt} \varphi(t) dt \rightarrow - \int_{2\delta'}^{1-2\delta'} f(t) \frac{d}{dt} \varphi(t) dt, \quad \text{für } m \rightarrow \infty. \quad (152)$$

Kombination von (150), (151) und (152) ergibt

$$\int_{2\delta'}^{1-2\delta^*} \varphi(t) h(t) dt = - \int_{2\delta'}^{1-2\delta'} f(t) \frac{d}{dt} \varphi(t) dt$$

und somit folgt mit Definition 2.4, dass h die schwache Ableitung von f ist. Wir schreiben daher $h = \dot{f}$. Damit bleibt nur noch Aussage (148) zu zeigen. Da $\frac{d}{dt}(\eta_{\delta_m} * f) = \eta_{\delta_m} * \frac{d}{dt}f$ gilt, folgt also mit Satz 2.3: $\frac{d}{dt}(\eta_{\delta_m} * f) \rightarrow \dot{f}$ in $L^2([2\delta, 1 - 2\delta])$, für $m \rightarrow \infty$. Mit Ungleichung 1 und 2 aus Lemma 3.10 bekommen wir

$$\begin{aligned} \int_{2\delta'}^{1-2\delta'} \left\| \frac{d}{dt}(\eta_{\delta_m} * f)(t) \right\| dt &\leq \liminf_{n \rightarrow \infty} \int_{2\delta'}^{1-2\delta'} \left\| \frac{\eta_{\delta_m} * f_n(t) - \eta_{\delta_m} * f_n(t - \varepsilon_n)}{\varepsilon_n} \right\| dt \\ &\leq \liminf_{n \rightarrow \infty} \int_{\delta'}^{1-\delta'} \left\| \frac{\eta_{\delta_m} * f_n(t) - \eta_{\delta_m} * f_n(t - \varepsilon_n)}{\varepsilon_n} \right\| dt \\ &\leq \liminf_{n \rightarrow \infty} \int_{\varepsilon_n}^1 \left\| \frac{f_n(t) - f_n(t - \varepsilon_n)}{\varepsilon_n} \right\|^2 dt. \end{aligned}$$

Aus der $L^2([2\delta, 1 - 2\delta])$ Konvergenz von $\frac{d}{dt}(\eta_{\delta_m} * f)$ folgt

$$\int_{2\delta'}^{1-2\delta'} \|\dot{f}(t)\|^2 dt \leq \liminf_{n \rightarrow \infty} \int_{\varepsilon_n}^1 \left\| \frac{f_n(t) - f_n(t - \varepsilon_n)}{\varepsilon_n} \right\|^2 dt. \quad (153)$$

Mit $\delta' \rightarrow 0^+$ folgt die Behauptung. ■

Damit können wir die Liminf-Ungleichung aus Definition 1.1 für \mathcal{E}_n zeigen.

Satz 3.12

Seien die Bedingungen aus Satz 3.3 erfüllt. Sei weiterhin $(\Theta^{(n)})_{n \in \mathbb{N}} \ni (K^{(n)}, b^{(n)}, W^{(n)}, c^{(n)})_{n \in \mathbb{N}} =: (\theta^{(n)})_{n \in \mathbb{N}} \rightarrow \theta := (K, b, W, c) \in \Theta$. Dann gilt

$$\liminf_{n \rightarrow \infty} \mathcal{E}_n(\theta^{(n)}) \geq \mathcal{E}_\infty(\theta). \quad (154)$$

Beweis: Sei $\theta \in \Theta$ beliebig. Sei weiterhin $\theta^{(n)} \rightarrow \theta$ beliebig. Falls $\liminf_{n \rightarrow \infty} \mathcal{E}_n(\theta^{(n)}) = \infty$, folgt die Behauptung trivialerweise. Somit reicht es, unter Betrachtung einer Teilfolge, die Aussage (154) nur für Folgen $\theta^{(n)} \rightarrow \theta$ zu zeigen, für welche $\mathcal{E}_n(\theta^{(n)})$ beschränkt ist. Sei $\theta^{(n)}$ eine solche Folge. Dann folgt aus Satz 3.9, dass $\lim_{n \rightarrow \infty} E_n(\theta^{(n)}; x, y) = E_\infty(\theta; x, y)$. Da $R^{(3)}$ und $R^{(4)}$ unabhängig von n und stetig sind, folgt $\lim_{n \rightarrow \infty} R^{(3)}(W^{(n)}) = R^{(3)}(W)$ und $\lim_{n \rightarrow \infty} R^{(4)}(c^{(n)}) = R^{(4)}(c)$. Für die Behauptung müssen wir also nur noch die folgenden zwei Aussagen zeigen:

1. $\liminf_{n \rightarrow \infty} R_n^{(1)}(K^{(n)}) \geq R_\infty^{(1)}(K)$,
2. $\liminf_{n \rightarrow \infty} R_n^{(2)}(b^{(n)}) \geq R_\infty^{(2)}(b)$.

Der Beweis der Aussagen verläuft analog, deshalb zeigen wir nur Aussage 1.

Sei $\tilde{K}^{(n)}(t) = K_i^{(n)}$ für $t \in (t_i^{(n)}, t_{i+1}^{(n)}]$ und $i = 0, \dots, n-1$. Dann ist

$$\begin{aligned} \liminf_{n \rightarrow \infty} R_n^{(1)}(K^{(n)}) &= \liminf_{n \rightarrow \infty} n \sum_{i=1}^{n-1} \|K_i^{(n)} - K_{i-1}^{(n)}\|^2 + \tau_1 \|K_0^{(n)}\|^2 \\ &\geq \liminf_{n \rightarrow \infty} \sum_{i=1}^{n-1} \|K_i^{(n)} - K_{i-1}^{(n)}\|^2 + \liminf_{n \rightarrow \infty} \tau_1 \|K_0^{(n)}\|^2 \\ &= \liminf_{n \rightarrow \infty} n^2 \int_{\frac{1}{n}}^1 \left\| \tilde{K}^{(n)}(t) - \tilde{K}^{(n)}\left(t - \frac{1}{n}\right) \right\|^2 dt + \tau_1 \|K_0^{(n)}\|^2 \\ &\geq \int_0^1 \|\dot{K}(t)\|^2 dt + \tau_1 \|K(0)\|^2 = R_\infty^{(1)}(K), \end{aligned}$$

wobei die letzte Ungleichung aus Satz 3.11 mit $\varepsilon_n = \frac{1}{n}$ folgt. Zusammengefasst bekommen wir also

$$\begin{aligned} \liminf_{n \rightarrow \infty} \mathcal{E}_n(\theta^{(n)}) &= \liminf_{n \rightarrow \infty} \sum_{s=1}^S E_n(\theta^{(n)}; x_s, y_s) + R_n^{(1)}(K^{(n)}) + R_n^{(2)}(b^{(n)}) + R^{(3)}(W^{(n)}) + R^{(4)}(c^{(n)}) \\ &= \sum_{s=1}^S E_\infty(\theta, x_s, y_s) + \liminf_{n \rightarrow \infty} R_n^{(1)}(K^{(n)}) + R_n^{(2)}(b^{(n)}) + R^{(3)}(W) + R^{(4)}(c) \\ &\geq \sum_{s=1}^S E_\infty(\theta, x_s, y_s) + R_\infty^{(1)}(K) + R_\infty^{(2)}(b) + R^{(3)}(W) + R^{(4)}(c) \\ &= \mathcal{E}_\infty(\theta). \end{aligned}$$

woraus die Behauptung folgt. ■

Um die Limsup-Ungleichung zu zeigen, definieren wir eine Folge $(\theta^{(n)}) \in (\Theta^{(n)})^{\mathbb{N}}$ durch $\theta^{(n)} = (K^{(n)}, b^{(n)}, W^{(n)}, c^{(n)})$, mit

$$K_i^{(n)} = n \int_{\frac{i}{n}}^{\frac{i+1}{n}} K(t) dt, \quad \text{für } i = 0, \dots, n-1, \quad (155)$$

$$b_i^{(n)} = n \int_{\frac{i}{n}}^{\frac{i+1}{n}} b(t) dt, \quad \text{für } i = 0, \dots, n-1, \quad (156)$$

$$W^{(n)} = W, \quad (157)$$

$$c^{(n)} = c. \quad (158)$$

Im Folgenden zeigen wir die Konvergenz von $\theta^{(n)} \rightarrow \theta$ in der TL^2 -Topologie und im Anschluss die Limsup-Ungleichung.

Lemma 3.13

Seien die Bedingungen aus Satz 3.3 erfüllt, sei $\theta = (K, b, W, c) \in \Theta$ und sei $(\theta^{(n)})_{n \in \mathbb{N}} \in (\Theta^{(n)})^{\mathbb{N}}$ definiert durch (155-158). Dann konvergiert $\theta^{(n)} \rightarrow \theta$ bezüglich der TL^2 -Topologie.

Beweis: Da $W^{(n)}$ und $c^{(n)}$ konstante Folgen sind, folgt dessen Konvergenz trivialerweise. Die Beweise für $K^{(n)} \rightarrow K$ und $b^{(n)} \rightarrow b$ laufen analog ab, daher beschränken wir uns auf den Beweis von $K^{(n)} \rightarrow K$. Sei $\tilde{K}^{(n)}(t) = K_i^{(n)}$ für $t \in [\frac{i}{n}, \frac{i+1}{n})$, für $i = 0, \dots, n-1$ und $K^{(n)}(1) = K_{n-1}^{(n)}$. Da $K \in H^1([0, 1])$ folgt mit Satz 2.8, dass $K \in C^{0, \frac{1}{2}}([0, 1]; \mathbb{R}^\kappa)$. Also existiert nach Definition 2.6 eine Konstante $C_K \in \mathbb{R}$, sodass $\|K(s) - K(t)\| \leq C_K \sqrt{|s-t|}$ für alle $s, t \in [0, 1]$. Damit können wir folgendermaßen abschätzen:

$$\begin{aligned} \|\tilde{K}^{(n)} - K\|_{L^2([0,1])}^2 &= \int_0^1 \|\tilde{K}^{(n)}(t) - K(t)\|^2 dt \\ &= \sum_{i=0}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \|K_i^{(n)} - K(t)\|^2 dt \\ &= \sum_{i=0}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left\| n \int_{\frac{i}{n}}^{\frac{i+1}{n}} K_i^{(n)} - K(t) ds \right\|^2 dt \end{aligned} \quad (159)$$

$$\begin{aligned} &\leq n^2 \sum_{i=0}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left(\int_{\frac{i}{n}}^{\frac{i+1}{n}} \|K(s) - K(t)\| ds \right)^2 dt \\ &\leq n \sum_{i=0}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \|K(s) - K(t)\|^2 ds dt \end{aligned} \quad (160)$$

$$\leq C_K^2 n \sum_{i=0}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \int_{\frac{i}{n}}^{\frac{i+1}{n}} |s-t| ds dt, \quad (161)$$

wobei in (159) die Definition von $K_i^{(n)}$ eingesetzt und in (160) die Jensensche Ungleichung benutzt wurde. Zunächst widmen wir uns dem Summanden in (161):

$$\begin{aligned} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \int_{\frac{i}{n}}^{\frac{i+1}{n}} |s-t| ds dt &= \int_{\frac{i}{n}}^{\frac{i+1}{n}} \int_{\frac{i}{n}}^t t-s ds + \int_{\frac{i}{n}}^{\frac{i+1}{n}} s-t ds dt \\ &= \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left(t \int_{\frac{i}{n}}^t ds - \int_{\frac{i}{n}}^t s ds \right) + \left(\int_t^{\frac{i+1}{n}} s ds - t \int_t^{\frac{i+1}{n}} ds \right) dt \\ &= \int_{\frac{i}{n}}^{\frac{i+1}{n}} \frac{1}{2} \left(t - \frac{i}{n} \right)^2 + \frac{1}{2} \left(t - \frac{i+1}{n} \right)^2 dt \\ &= \frac{1}{2} \left(\int_0^{\frac{1}{n}} u^2 du + \int_{-\frac{1}{n}}^0 v^2 dv \right) \\ &= \frac{1}{2} \int_{-\frac{1}{n}}^{\frac{1}{n}} u^2 du \\ &= \frac{1}{3n^3}. \end{aligned}$$

Einsetzen in (161) ergibt

$$\begin{aligned} \|\tilde{K}^{(n)} - K\|_{L^2([0,1])}^2 &\leq C_K^2 n \sum_{i=0}^{n-1} \frac{1}{3n^3} \\ &= \frac{C_K}{3n} \rightarrow 0, \quad \text{für } n \rightarrow \infty. \end{aligned}$$

Mit Satz 3.2 folgt die Behauptung. ■

Wir zeigen jetzt, dass für die durch (155-158) definierte Folge $(\theta^{(n)})_{n \in \mathbb{N}}$ die Limsup-Ungleichung erfüllt ist.

Satz 3.14

Seien die Bedingungen aus Satz 3.3 erfüllt. Weiterhin definieren wir für beliebiges $\theta \in \Theta$ die Folge $(\theta^{(n)})_{n \in \mathbb{N}} \in (\Theta^{(n)})^{\mathbb{N}}$ wie in Lemma 3.13. Dann gilt $\theta^{(n)} \rightarrow \theta$ und

$$\limsup_{n \rightarrow \infty} \mathcal{E}_n(\theta^{(n)}) \leq \mathcal{E}_\infty.$$

Beweis: Sei $\theta = (K, b, W, c) \in \Theta$. Falls $\mathcal{E}_\infty = \infty$ folgt die Behauptung trivialerweise. Daher betrachten wir nur den Fall $\mathcal{E}_\infty < \infty$. $\theta^{(n)} \rightarrow \theta$ haben wir bereits in Lemma 3.13 gezeigt. Aus Satz 3.9 folgt somit, dass $\lim_{n \rightarrow \infty} E_n(\theta^{(n)}) = E_\infty(\theta)$. Daher bleiben für die Behauptung, ähnlich wie im Beweis von Satz 3.12, nur noch die folgenden zwei Aussagen zu zeigen:

1. $\limsup_{n \rightarrow \infty} R_n^{(1)}(K^{(n)}) \leq R_\infty^{(1)}(K)$
2. $\limsup_{n \rightarrow \infty} R_n^{(2)}(b^{(n)}) \leq R_\infty^{(2)}(b)$

Da der Beweis der Aussagen 1 und 2 analog verläuft, zeigen wir nur Aussage 1. Wir schätzen ab:

$$R_n^{(1)}(K^{(n)}) = n \sum_{i=1}^{n-1} \|K_i^{(n)} - K_{i-1}^{(n)}\|^2 + \tau_1 \|K_0^{(n)}\|^2 \tag{162}$$

$$= n \sum_{i=1}^{n-1} \left\| n \int_{\frac{i}{n}}^{\frac{i+1}{n}} K(t) - K\left(t - \frac{1}{n}\right) dt \right\|^2 + \tau_1 \left\| n \int_0^{\frac{1}{n}} K(t) dt \right\|^2 \tag{163}$$

$$\leq n^2 \sum_{i=1}^{n-1} \int_{\frac{i}{n}}^{\frac{i+1}{n}} \left\| K(t) - K\left(t - \frac{1}{n}\right) \right\|^2 dt + \tau_1 \left\| n \int_0^{\frac{1}{n}} K(t) - K(0) + K(0) dt \right\|^2 \tag{164}$$

$$\leq n^2 \int_{\frac{1}{n}}^1 \left\| K(t) - K\left(t - \frac{1}{n}\right) \right\|^2 dt + \tau_1 n \int_0^{\frac{1}{n}} \|K(t) - K(0) + K(0)\|^2 dt \tag{165}$$

Sei $\varepsilon > 0$ beliebig. Mit der youngschen Ungleichung für Produkte folgt, dass eine Konstante $C_\varepsilon = (1 + \frac{1}{\varepsilon}) \in \mathbb{R}$ existiert, sodass für beliebige $a, b \in \mathbb{R}$

$$\|a + b\|^2 \leq (1 + \varepsilon) \|a\|^2 + C_\varepsilon \|b\|^2$$

gilt. Damit können wir den rechten Teil von (165) abschätzen durch

$$\begin{aligned} \tau_1 n \int_0^{\frac{1}{n}} \|K(t) - K(0) + K(0)\|^2 dt &\leq \tau_1 n \int_0^{\frac{1}{n}} C_\varepsilon \|K(t) - K(0)\|^2 + (1 + \varepsilon) \|K(0)\|^2 dt \\ &\leq (1 + \varepsilon) \tau_1 \|K(0)\|^2 + \tau_1 C_\varepsilon n C_K^2 \int_0^{\frac{1}{n}} t dt \\ &= (1 + \varepsilon) \tau_1 \|K(0)\|^2 + \frac{\tau_1 C_\varepsilon C_K^2}{2n}, \end{aligned}$$

wobei C_K wie in dem Beweis von Lemma 3.13 definiert ist. Eingesetzt in (165) bekommen wir

$$\begin{aligned} R_n^{(1)}(K^{(n)}) &\leq n^2 \int_{\frac{1}{n}}^1 \left\| K(t) - K\left(t - \frac{1}{n}\right) \right\|^2 dt + (1 + \varepsilon) \tau_1 \|K(0)\|^2 + \frac{\tau_1 C_\varepsilon C_K^2}{2n} \\ &\leq \int_0^1 \|\dot{K}\|^2 dt + (1 + \varepsilon) \tau_1 \|K(0)\|^2 + \frac{\tau_1 C_\varepsilon C_K^2}{2n} \\ &= \|\dot{K}\|_{L^2(0,1)}^2 + (1 + \varepsilon) \tau_1 \|K(0)\|^2 + \frac{\tau_1 C_\varepsilon C_K^2}{2n}, \end{aligned}$$

wobei die letzte Ungleichung aus Satz 3.11 folgt. Grenzübergang $n \rightarrow \infty$ ergibt

$$\limsup_{n \rightarrow \infty} R_n^{(1)}(K^{(n)}) \leq \int_0^1 \|\dot{K}(t)\|^2 dt + (1 + \varepsilon) \tau_1 \|K(0)\|^2 \leq (1 + \varepsilon) R_\infty^{(1)}(K).$$

Die Behauptung folgt für $\varepsilon \rightarrow 0$. ■

3.4.4 Beweis von Satz 3.3

Indem wir die Ergebnisse aus den vorherigen drei Abschnitten zusammentragen, können wir nun das Hauptresultat beweisen. Seien dafür die Annahmen aus Satz 3.3 erfüllt. Aus Satz 3.7 folgt, dass eine Minimiererfolge von $\theta^{(n)}$ von \mathcal{E}_n über $\Theta^{(n)}$ existiert. Weiterhin resultiert aus Satz 3.12 und Satz 3.14

$$\mathcal{E}_\infty = \Gamma\text{-lim } \mathcal{E}_n.$$

Daher folgt mit (B⁺02, Prop. 1.18(ii)), dass $\mathcal{E}_n(\theta^{(n)})$ beschränkt ist. Also können wir durch Anwendung von Korollar 3.6 folgern, dass $\theta^{(n)}$ kompakt in $\Theta^{(n)}$ im Sinne von Abschnitt 3.4 ist. Durch Bildung des Abschlusses von $\{\theta^{(n)} : n \in \mathbb{N}\}$ ergibt sich die gleichgradige Koerzivität von \mathcal{E}_n . Somit folgt aus Satz 1.8 die Behauptung von Satz 3.3.

Kapitel 4: Experiment

Mit dem Ziel der Verifikation der Theorie aus Kapitel 3 haben wir ein Experiment durchgeführt, in welchem wir analog zu dem in Kapitel 3 beschriebenen Modell ein residuales neuronales Netzwerk mithilfe der Programmbibliothek *PyTorch* (PGM⁺19) implementiert haben. Dabei wollten wir die Konvergenz der diskreten Netzparameter gegen die kontinuierlichen Parameter beobachten. Im Folgenden werden der Versuchsaufbau, die Versuchsdurchführung und die Resultate des Versuches präsentiert.

4.1 Versuchsaufbau und -durchführung

Bei dem Versuch wurde das Modell aus Kapitel 3 etwas vereinfacht, indem wir auf die Klassifizierungsfunktion h verzichtet haben und somit auch die Parameter W und c wegfallen. Die übrigbleibenden Komponenten des Modells haben wir mit $d = 1$ gewählt durch

$$\begin{aligned}\sigma(x) &= \text{relu}(x), \\ \mathcal{L}(x, y) &= |x - y|^2, \\ \alpha_i &= 1 \text{ für } i = 1, 2 \\ \tau_j &> 0 \text{ für } j = 1, 2.\end{aligned}$$

Hierbei ist zu beachten, dass die Voraussetzungen aus Satz 3.3 erfüllt sind. Auf die genauen Werte von τ_j gehen wir in Abschnitt 4.2 ein. Weiterhin haben wir versucht Trainingsdaten $\{x_s, y_s\}_{s=1}^S$ so zu konstruieren, dass wir die Minimierer von \mathcal{E}_∞ , welche für gegebene Trainingsdaten analytisch nur sehr schwer zu bestimmen sind, vorhersagen können. Dafür haben wir Funktionen $K, b : [0, 1] \rightarrow \mathbb{R}$ definiert durch

$$K(t) = b(t) = e^{-5(t-0.5)^2}.$$

Im Anschluss haben wir die gewöhnliche Differentialgleichung

$$\dot{X}(t) = \sigma(K(t)X(t) + b(t)) \quad (166)$$

über $[0, 1]$ für $S = 50000$ verschiedene Anfangswerte numerisch mittels der *odeint* Funktion aus der *Python*-Bibliothek *SciPy* (VGO⁺20) gelöst. Die Anfangswerte $\{x_s\}_{s=1}^{50000}$ sind dabei zufällige, gleichverteilte Werte aus $[0, 1]$. Damit bekommen wir also die folgende Menge an Lösungen der Differentialgleichung:

$$\{X_s : [0, 1] \rightarrow \mathbb{R}\}_{s=1}^S,$$

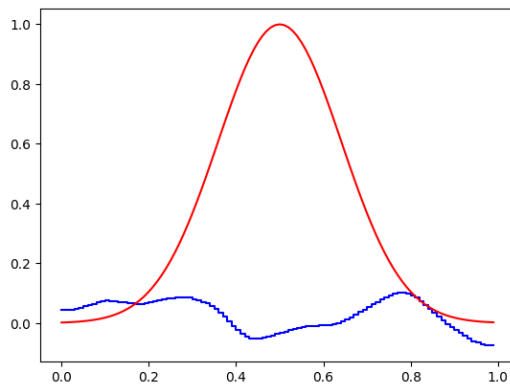
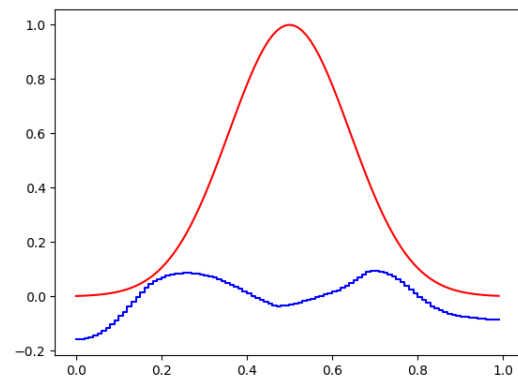
wobei X_s die Lösung von (166) unter der Bedingung $X(0) = x_s$ beschreibt. Die Trainingsdaten seien dann definiert durch

$$\{x_s, y_s\}_{s=1}^S = \{x_s, X_s(1)\}_{s=1}^S = \{X_s(0), X_s(1)\}_{s=1}^S. \quad (167)$$

Wir haben dann die Vermutung angestellt, dass die Minimierer $(K^*, b^*) \in H^1([0, 1], \mathbb{R}) \times H^1([0, 1], \mathbb{R})$ von \mathcal{E}_∞ durch $K^*(t) = K(t)$ und $b^*(t) = b(t)$ gegeben sind. Daraufhin haben wir das Netzwerk mit $n \in \{100, 1000\}$ Schichten mithilfe der Trainingsdaten aus (167) über 500 Epochen trainiert und anschließend die Parameter $\mathbf{K}^{(n)}$ und $\mathbf{b}^{(n)}$ ausgelesen.

4.2 Resultate

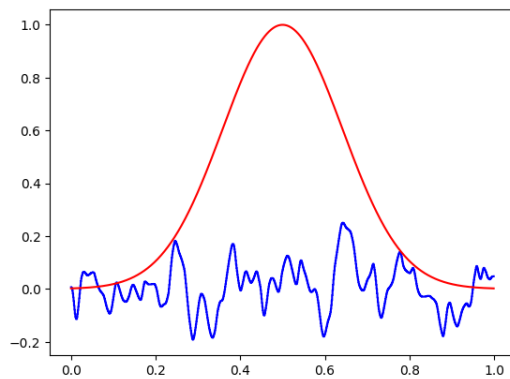
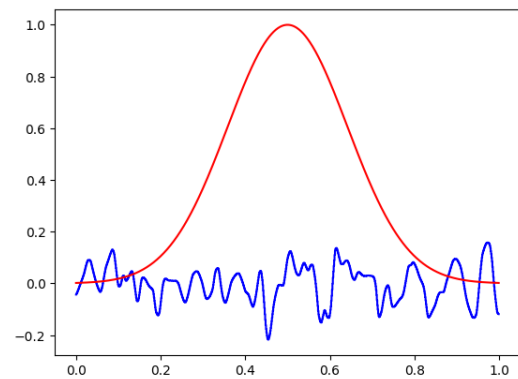
Gemäß der Theorie aus Kapitel 3 haben wir erwartet, dass die Parameter $\mathbf{K}^{(n)}$ und $\mathbf{b}^{(n)}$ gegen die mutmaßlichen Minimierer K^* und b^* bezüglich des Abstandes d aus (84) konvergieren. Um das zu überprüfen haben wir die Funktionen $\tilde{K}, \tilde{b} : [0, 1] \rightarrow \mathbb{R}$, analog zu der Vorgehensweise in der Konstruktion von d , als stückweise konstante Fortsetzung der als diskreten Funktionen verstandenen gelernten Parameter $K^{(n)}$ und $b^{(n)}$ definiert. Im Folgenden eine Visualisierung der Funktionen K^* und \tilde{K} , beziehungsweise b^* und \tilde{b} für verschiedene Werte von τ_j und $n \in \{100, 1000\}$.

(a) Rot: $K(t)$, Blau: $\tilde{K}(t)$ (b) Rot: $b(t)$, Blau: $\tilde{b}(t)$

$$n = 100,$$

$$\tau_1 = 0.1, \tau_2 = 0.1,$$

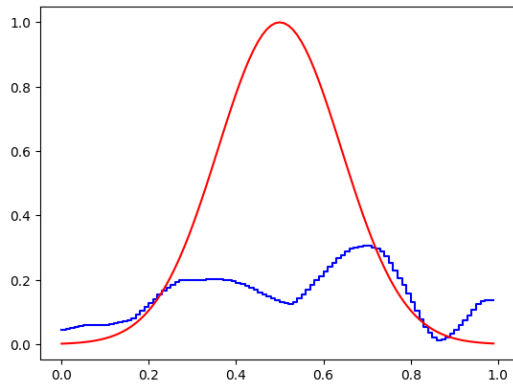
Trainingsfehler: 3.2077, Validierungsfehler: 3.2361

(c) Rot: $K(t)$, Blau: $\tilde{K}(t)$ (d) Rot: $b(t)$, Blau: $\tilde{b}(t)$

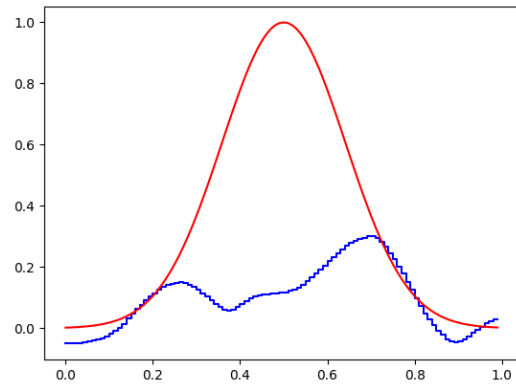
$$n = 1000,$$

$$\tau_1 = 0.1, \tau_2 = 0.1,$$

Trainingsfehler: 143.6049, Validierungsfehler: 144.3028



(e) Rot: $K(t)$, Blau: $\tilde{K}(t)$

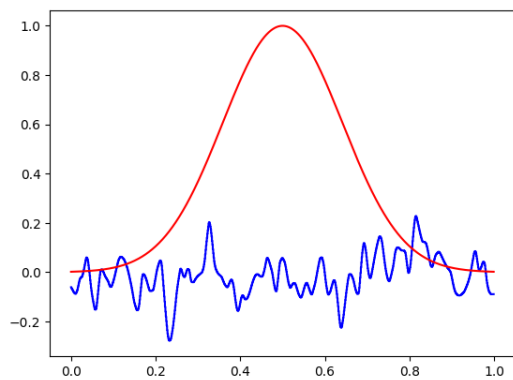


(f) Rot: $b(t)$, Blau: $\tilde{b}(t)$

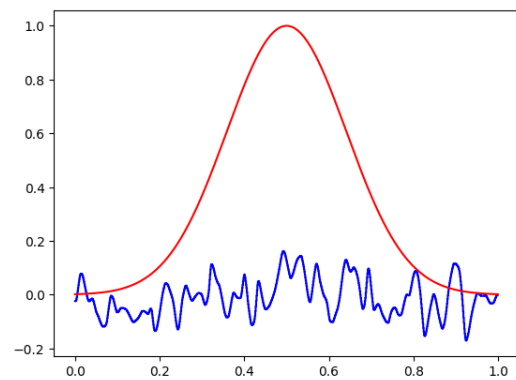
$n = 100$

$\tau_1 = 1, \tau_2 = 1,$

Trainingsfehler: 4.2109, Validierungsfehler: 4.5058



(g) Rot: $K(t)$, Blau: $\tilde{K}(t)$

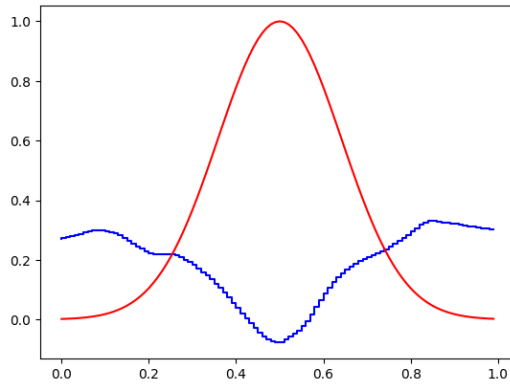
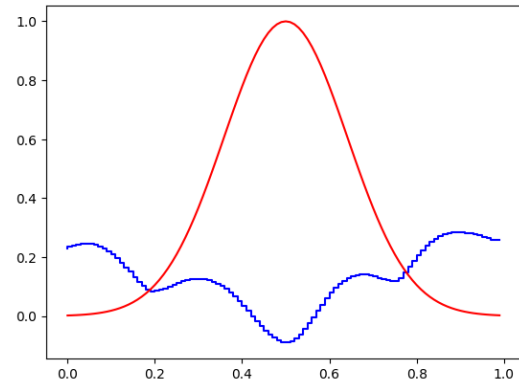


(h) Rot: $b(t)$, Blau: $\tilde{b}(t)$

$n = 1000,$

$\tau_1 = 1, \tau_2 = 1,$

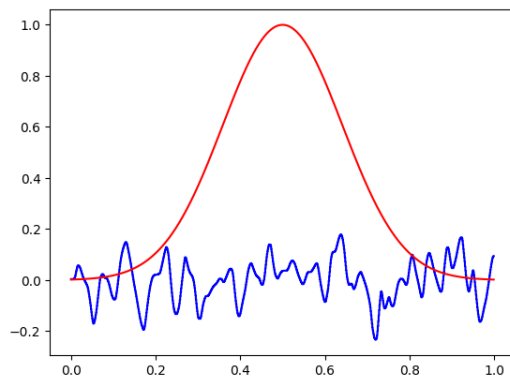
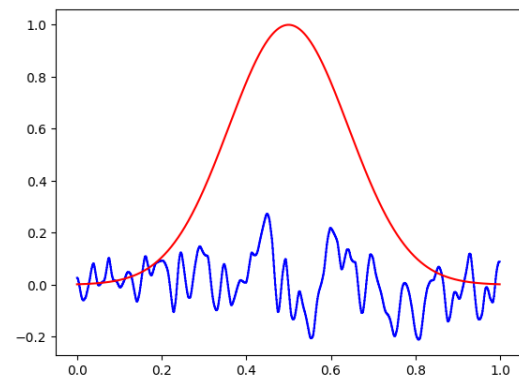
Trainingsfehler: 143.6049, Validierungsfehler: 144.3028

(a) Rot: $K(t)$, Blau: $\tilde{K}(t)$ (b) Rot: $b(t)$, Blau: $\tilde{b}(t)$

$$n = 100$$

$$\tau_1 = 0.5, \tau_2 = 3,$$

Trainingsfehler: 3.9370, Validierungsfehler: 4.5943

(c) Rot: $K(t)$, Blau: $\tilde{K}(t)$ (d) Rot: $b(t)$, Blau: $\tilde{b}(t)$

$$n = 1000,$$

$$\tau_1 = 0.5, \tau_2 = 3,$$

Trainingsfehler: 143.6049, Validierungsfehler: 144.3028

Wie in den Grafiken zu erkennen ist, konnten wir keine L^2 -Konvergenz der stückweise konstant fortgesetzten Parameter gegen die mutmaßlichen Minimierer beobachten. Die wohl offensichtlichste Fehlerquelle ist, dass K^* und b^* nicht notwendigerweise Minimierer des Grenzfunktional \mathcal{E}_∞ sein müssen. Aber auch die Möglichkeit, dass die Schichtanzahl n nicht groß genug gewählt ist, um die Konvergenz zu zeigen, ist nicht auszuschließen.

Kapitel 5: Fazit

Wir haben in Kapitel 1 das Konzept und einige Eigenschaften der Γ -Konvergenz erarbeitet, insbesondere Satz 1.8, welcher die Konvergenz der Minimierer einer Γ -konvergenten Folge von Funktionalen gegen die Minimierer des Γ -Grenzwertes beschreibt. Weiterhin haben wir in Kapitel 2 die notwendigen funktionalanalytischen Grundlagen gelegt, mit welchen es uns möglich war in Kapitel 3 das Lernproblem eines n -schichtigen residualen neuronalen Netzwerkes für gegebene Trainingsdaten als Minimierungsproblem eines Zielfunktional \mathcal{E}_n über diskreten Funktionenräumen zu formulieren. Zusätzlich haben wir in Kapitel 3 für $n \rightarrow \infty$ und allgemeine Trainingsdaten die Γ -Konvergenz des zu minimierenden Zielfunktional gegen ein von uns definiertes, kontinuierliches Grenzfunktional sowie die Kompaktheit von Minimiererfolgen von \mathcal{E}_n gezeigt, indem wir die Beweise aus (TG18) detailliert ausformuliert haben. Somit konnten wir unter Anwendung von Satz 1.8 die Konvergenz der Minimierer des Zielfunktional gegen die Minimierer des Grenzfunktional zeigen, was das Hauptresultat dieser Arbeit darstellt.

In Kapitel 4 haben wir versucht die zuvor erarbeitete Theorie experimentell zu beobachten. Dafür haben wir das in Kapitel 3 definierte Modell implementiert und Trainingsdaten aus Randwerten von Lösungen von Differentialgleichungen erstellt, in der Hoffnung die Minimierer von \mathcal{E}_∞ vorhersagen und somit für eine hohe Anzahl von Schichten n die Konvergenz der Parameter beobachten zu können. In unseren Versuchen war es uns nicht möglich die Γ -Konvergenz der Netzwerk Parameter zu beobachten. Der wahrscheinlichste Grund dafür ist eine falsche Vorhersage der Minimierer von \mathcal{E}_∞ .

Daher wäre ein weiterer erfolgversprechender Ansatz, die Minimierer von \mathcal{E}_∞ analytisch zu bestimmen, um so passende Trainingsdaten zu konstruieren. Eine analytische Berechnung der Minimierer von \mathcal{E}_∞ erscheint jedoch äußerst herausfordernd und konnte im Rahmen dieser Arbeit nicht durchgeführt werden. Für weitere Arbeiten bietet sich deshalb eine nähere Untersuchung dieses Ansatzes an. Ein aktueller Ansatz zur Untersuchung des Trainings von neuronalen Netzen ist die Interpretation des Minimierungsproblem als Problem der optimalen Steuerung (CSZ19). Durch solch eine Interpretation ist es möglich Minimierer von \mathcal{E}_∞ für gegebene Trainingsdaten zu berechnen. Durch die Berechnung dieser Minimierer stehen für die Untersuchung der Γ -Konvergenz eine Grundwahrheit für die erwarteten Grenzwerte zur Verfügung.

Literatur

- [B⁺02] BRAIDES, Andrea u. a.: *Gamma-convergence for Beginners*. Bd. 22. Clarendon Press, 2002
- [CSZ19] CHEN, Yize ; SHI, Yuanyuan ; ZHANG, Baosen: *Optimal Control Via Neural Networks: A Convex Approach*. 2019
- [Den12] DENG, Li: The mnist database of handwritten digit images for machine learning research [best of the web]. In: *IEEE Signal Processing Magazine* 29 (2012), Nr. 6, S. 141–142
- [DM12] DAL MASO, Gianni: *An introduction to Γ -convergence*. Bd. 8. Springer Science & Business Media, 2012
- [Eva98] EVANS, Lawrence C.: Partial differential equations. In: *Graduate studies in mathematics* 19 (1998), Nr. 2
- [For07] FORSTER, Otto: *Analysis*. Springer, 2007
- [For17] FORSTER, Otto: *Analysis 2: Differentialrechnung im \mathbb{R}^n , gewöhnliche Differentialgleichungen*. Springer-Verlag, 2017
- [GT01] GILBARG, David ; TRUDINGER, Neil S.: *Elliptic partial differential equations of second order*. Bd. 224. springer, 2001
- [HJG34] H., Hardy G. ; J.E., Littlewood ; G., Polya: *Inequalities*. Cambridge University Press, 1934
- [HZRS16] HE, Kaiming ; ZHANG, Xiangyu ; REN, Shaoqing ; SUN, Jian: Deep residual learning for image recognition. In: *Proceedings of the IEEE conference on computer vision and pattern recognition*, 2016, S. 770–778
- [Lel11] LELLMANN, Jan: *Nonsmooth convex variational approaches to image analysis*, Diss., 2011
- [Leo09] LEONI, Giovanni: *A First Course in Sobolev Spaces*. Bd. 105. American Mathematical Soc., 2009
- [NH10] NAIR, Vinod ; HINTON, Geoffrey E.: Rectified linear units improve restricted boltzmann machines. In: *Icml*, 2010
- [PGM⁺19] PASZKE, Adam ; GROSS, Sam ; MASSA, Francisco ; LERER, Adam ; BRADBURY, James ; CHANAN, Gregory ; KILLEEN, Trevor ; LIN,

- Zeming ; GIMELSHEIN, Natalia ; ANTIGA, Luca ; DESMAISON, Alban ; KOPF, Andreas ; YANG, Edward ; DEVITO, Zachary ; RAISON, Martin ; TEJANI, Alykhan ; CHILAMKURTHY, Sasank ; STEINER, Benoit ; FANG, Lu ; BAI, Junjie ; CHINTALA, Soumith: PyTorch: An Imperative Style, High-Performance Deep Learning Library. Version:2019. <http://papers.neurips.cc/paper/9015-pytorch-an-imperative-style-high-performance-deep-learning-library.pdf>. In: WALLACH, H. (Hrsg.) ; LAROCHELLE, H. (Hrsg.) ; BEYGELZIMER, A. (Hrsg.) ; ALCHÉ-BUC, F. d'(Hrsg.) ; FOX, E. (Hrsg.) ; GARNETT, R. (Hrsg.): *Advances in Neural Information Processing Systems 32*. Curran Associates, Inc., 2019, 8024–8035
- [Rud87] RUDIN, Walter: *Real and Complex Analysis, 3rd Ed.* USA : McGraw-Hill, Inc., 1987. – ISBN 0070542341
- [TG18] THORPE, Matthew ; GENNIP, Yves van: Deep limits of residual neural networks. In: *arXiv preprint arXiv:1810.11741* (2018)
- [TS16] TRILLOS, Nicolás G. ; SLEPČEV, Dejan: Continuum limit of total variation on point clouds. In: *Archive for rational mechanics and analysis* 220 (2016), Nr. 1, S. 193–241
- [VGO⁺20] VIRTANEN, Pauli ; GOMMERS, Ralf ; OLIPHANT, Travis E. ; HABERLAND, Matt ; REDDY, Tyler ; COURNAPEAU, David ; BUROVSKI, Evgeni ; PETERSON, Pearu ; WECKESSER, Warren ; BRIGHT, Jonathan ; VAN DER WALT, Stéfan J. ; BRETT, Matthew ; WILSON, Joshua ; MILLMAN, K. J. ; MAYOROV, Nikolay ; NELSON, Andrew R. J. ; JONES, Eric ; KERN, Robert ; LARSON, Eric ; CAREY, C J. ; POLAT, İlhan ; FENG, Yu ; MOORE, Eric W. ; VANDERPLAS, Jake ; LAXALDE, Denis ; PERKTOLD, Josef ; CIMRMAN, Robert ; HENRIKSEN, Ian ; QUINTERO, E. A. ; HARRIS, Charles R. ; ARCHIBALD, Anne M. ; RIBEIRO, Antônio H. ; PEDREGOSA, Fabian ; VAN MULBREGT, Paul ; SCIPY 1.0 CONTRIBUTORS: SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python. In: *Nature Methods* 17 (2020), S. 261–272. <http://dx.doi.org/10.1038/s41592-019-0686-2>. – DOI 10.1038/s41592-019-0686-2
- [Wer11] WERNER, Dirk: *Funktionalanalysis*. Springer, 2011
- [ZLSZ20] ZHANG, Gege ; LI, Gangwei ; SHEN, Weining ; ZHANG, Weidong: The expressivity and training of deep neural networks: Toward the edge of chaos? In: *Neurocomputing* 386 (2020), S. 8–17