TECHNISCHE UNIVERSITÄT DARMSTADT

Nichtparametrische Kurvenschätzung für latente Variablen

vom Fachbereich Mathematik der Technischen Universität Darmstadt

zur Erlangungn des Grades

Doktor rerum naturalium

(Dr.rer.nat.)

Dissertation

von Florian Müller

Erstgutachter: Prof. Dr. Michael Kohler

Zweitgutachter: Prof. Dr. Frank Aurzada

Darmstadt 2017

Müller, Florian: Nichtparametrische Kurvenschätzung für latente Variablen
Darmstadt, Technische Universität Darmstadt
Jahr der Veröffentlichung der Dissertation auf TUprints: 2018
Tag der mündlichen Prüfung: 21.6.2017

Danksagung

An dieser Stelle möchte ich mich bei allen bedanken, die mich während meiner Promotion und beim Erstellen dieser Dissertation begleitet und unterstützt haben.

Besonderer Dank gilt dabei Herrn Professor Dr. Michael Kohler für die Vergabe des äußerst interessanten Themas der Arbeit und der stets aufmerksamen und zielgerichteten Betreuung und Unterstützung beim Erstellen dieser Arbeit. Außerdem danke ich den Gutachtern und Prüfern Prof. Dr. Frank Aurzada, Prof. Dr. Ulrich Reif und PD Dr. Robert Haller-Dintelmann für die aufgebrachte Zeit.

Ich bedanke mich ebenfalls bei allen anderen Mitgliedern der Arbeitsgruppe Stochastik, die mir alle in den vergangenen Jahren mit ihren offenen Ohren und ihren Tipps und Ratschlägen eine große Hilfe waren. Die freundschafliche Atmosphäre in der Arbeitsgruppe hat mir die Arbeit erleichtert. Bei Reinhard Tent und Benedikt Bauer bedanke ich mich für das Korrekturlesen der Arbeit und das Finden der kleinen Fehler und versteckten Inkonsistenzen, für deren entdecken mir die nötige Distanz vom Text gefehlt hat. Außerdem bedanke ich mich bei Ann-Kathrin Bott für den produktiven fachlichen Austausch und die angenehme Arbeitsatmosphäre. Bei Fabian Werner und Moritz Egert bedanke ich mich für die vielen anregenden und aufschlussreichen Gespräche während unseres gemeinsamen Studiums. Fabians Werners interessiertes Nachfragen zum Dichteschätzer half mir beim Überwinden mancher Hürde in diesem Bereich der Arbeit.

Weiterer Dank gilt meiner Mutter, Petra Müller, und meinem Vater, Michael Busch, die mich während meines gesamten Studiums begleitet und ermutigt haben. Ich bedanke mich bei Johanna Sallanz und unserem Sohn Jonas für ihre Geduld und ihre liebevolle Unterstützung, gerade auch in den schwierigen Phasen meiner Promotion. Danke, dass ihr an mich geglaubt habt.

Inhaltsverzeichnis

1 Einleitung

In diesem Kapitel werden die grundlegenden Problemstellungen der Verteilungsschätzung sowie der Dichte- und Regressionsschätzung erläutert und die relevante Literatur vorgestellt. Abschließend werden im letzten Unterkapitel die in dieser Arbeit betrachteten Probleme und die neu erarbeiteten Resultate präsentiert.

1.1 Schätzung von Verteilungen

Im Folgenden sei X eine reellwertige Zufallsvariable mit Verteilung μ, also $\mathbf{P}[X \in B] = \mu(B)$ für alle Mengen B aus der Borelschen σ-Algebra $\mathcal{B}$ auf $\mathbb{R}$. Falls μ unbekannt ist, aber ein Datensatz $X_1, \ldots, X_n$ von unabhängig und identisch verteilten Kopien von X zur Verfügung steht, so stellt sich die Frage, ob sich μ aus den Daten rekonstruieren oder zumindest schätzen lässt. Dieses Problem ist von fundamentaler Bedeutung für die mathematische Statistik und entsprechend gründlich untersucht. Offensichtlich ist das empirische Maß $\mu_n : \mathcal{B} \to \mathbb{R}$,

$$\mu_n(B) := \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_B(X_j),$$

punktweise konsistent für alle $B \in \mathcal{B}$, da für unabhängig und identisch verteilte $X_1, X_2, \ldots, X_n$ für jedes $B \in \mathcal{B}$ auch $\mathbb{1}_B(X_1), \ldots, \mathbb{1}_B(X_n)$ unabhängig und identisch verteilt sind und damit nach dem starken Gesetz der großen Zahlen mit Wahrscheinlichkeit 1

$$\frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_B(X_j) \to \mathbf{E}\{\mathbb{1}_B(X)\} = \mu(B)$$

gilt. In der obigen Gleichung sowie im weiteren Verlauf der Arbeit ist jeder Grenzwert, sofern nicht anders angegeben, als Grenzwert für $n \to \infty$ zu verstehen. Nun stellt sich sofort die Frage, ob es auch Schätzer $\hat{\mu}_n$ gibt, die in einem geeigneten Sinne gleichmäßig konvergieren. Eine erste Antwort liefert das bekannte Theorem von Glivenko-Cantelli [1]: Sind $X, X_1, \ldots, X_n$ unabhängig und identisch verteilt so gilt für das empirische Maß $\mu_n : \mathcal{B} \to \mathbb{R}$ mit Wahrscheinlichkeit 1

$$\lim_{n\to\infty} \sup_{x\in\mathbb{R}} |\mu_n((-\infty, x]) - \mu((-\infty, x])| = 0.$$

Das empirische Maß erlaubt also eine gleichmäßig konsistente Schätzung von $\mu((\infty, x])$ für alle $x \in \mathbb{R}$ und damit auch eine konsistente Schätzung für $\mu(I)$ für alle Intervalle $I \subset \mathbb{R}$. Wie Devroye und Györfi in [2] gezeigt haben, gibt es jedoch keinen Schätzer $\hat{\mu}_n$, sodass der totale Variationsfehler

$$\sup_{B\in\mathcal{B}} |\hat{\mu}_n(B) - \mu(B)| \tag{1.1}$$

für alle Verteilungen fast sicher gegen 0 konvergiert. Besitzt μ jedoch eine Lebesgue-Dichte f, d.h. eine $\mathcal{B}$-messbare Funktion, so dass

$$\int_B f \, d\lambda = \mu(B)$$

für alle $B \in \mathcal{B}$ gilt, so lässt sich für jede Schätzung $\hat{f}_n = \hat{f}_n(\cdot, X_1, \ldots, X_n)$ der Dichte f durch

$$\hat{\mu}_n(B) := \int_B \hat{f}_n \, d\lambda \tag{1.2}$$

auch die Verteilung μ von X schätzen. Aus dem Lemma von Scheffé [3] folgt nun, dass für solche Schätzer, und alle Verteilungen μ die eine Dichte besitzen, der totale Variationsfehler (1.1) durch den halben L^1-Fehler

$$\int |\hat{f}_n - f| \, d\lambda$$

beschränkt ist. Aus universell konsistenten Dichteschätzern, also Schätzern für welche der L^1-Fehler für jede zu schätzende Dichte fast sicher gegen 0 konvergiert, erhalten wir also mittels (1.2) sofort einen über alle Borelmengen gleichmäßig konsistenten Schätzer. Das nächste Unterkapitel liefert daher einen Überblick über das Gebiet der Dichteschätzung.

1.2 Schätzung von Dichten

In Unterkapitel 1.1 haben wir gesehen, dass stark konsistente Dichteschätzer, also Dichteschätzer $\hat{f}_n(\cdot, X_1, \ldots, X_n)$, für die für jede Dichte f von $X_1, \ldots, X_n$ fast sicher

$$\lim_{n \to \infty} \int |\hat{f}_n - f| \, d\lambda = 0$$

gilt, zur Konstruktion von Verteilungsschätzern benutzt werden können. Die Konstruktion solcher Dichteschätzer stellt also ebenfalls eines der grundlegenden Probleme der mathematischen Statistik dar und wurde entsprechend ausgiebig untersucht. Als Ausgangspunkt dient hierbei oft der Kerndichteschätzer von Rosenblatt und Parzen ([4], [5]), der die Dichte f von X durch

$$\hat{f}_n(x) = \frac{1}{nh_n} \sum_{j=1}^{n} K\left(\frac{x - X_j}{h_n}\right)$$

schätzt. Hierbei ist K der Kern und h_n die Bandbreite des Schätzers. Ist die Funktion K eine Dichte, so ist auch $\hat{f}_n$ eine Dichte. In diesem Fall ist der Schätzer L^1 konsistent, falls die Bandbreite die Bedingungen

$$h_n \to 0 \quad \text{und} \quad nh_n \to \infty$$

erfüllt und der Kern K selbst ebenfalls eine Dichte ist (siehe [6] und [7]). Also ist die konsistente Schätzung von Dichten, und damit auch die konsistente Schätzung von Verteilungen bezüglich des totalen Variationsfehlers (1.1), ausgehend von unabhängig und identisch verteilten Daten $X_1, \ldots, X_n$, grundsätzlich möglich. Weiterführend ist für den Kerndichteschätzer auch die datenabhängige Wahl der Bandbreite, beispielsweise durch Kreuzvalidierung (siehe [8]) oder kombinatorische Methoden (siehe [9]) und Kerndichteschätzer, deren Kern keine Dichte ist ([10],[11]), auf Konsistenz, Konvergenzgeschwindigkeit und Optimalität der Konvergenzgeschwindigkeit untersucht worden. Für andere Methoden der Dichteschätzung, wie beispielsweise den Histogrammschätzer [12] existieren ebenfalls entsprechende Konsistenzaussagen. Auch hinsichtlich des mittleren lokalen

L^2-Fehlers, gegeben durch $\mathbf{E}|f - f_n|^2$, und des mittleren integrierten L^2-Fehlers, gegeben durch $\mathbf{E}\int|f - f_n|^2\, d\lambda$, welche für uns aber von geringerer Relevanz sind, existieren ebenfalls Resultate zur Konsistenz und Konvergenzgeschwindigkeit (siehe beispielsweise [13] oder [4]).

Wie steht nun die Lage, falls zur Konstruktion des Schätzers kein Datensatz von unabhängig und identisch verteilten Kopien $X_1, \ldots, X_n$ von X vorliegt, sondern die Verteilung der vorliegenden Daten $\bar{X}_1, \ldots, \bar{X}_n$ von der Verteilung der wahren $X_1, \ldots, X_n$ abweicht? Unter Annahmen über die Art des Messfehlers $\varepsilon_i = \bar{X}_i - X_i$ können auch für diese Problemstellung Konsistenzaussagen bewiesen werden, wie beispielsweise für Messfehler mit bekannter Dichte (sogenannte Deconvolution-Probleme) von Johannes [14] und Meister [15] gezeigt wurde. Bott, Kohler und Devroye [16] behandeln dieses Problem ohne Annahme an die Verteilung der Fehler, setzen aber voraus, dass der mittlere Messfehler fast sicher gegen 0 konvergiert. Wie wir in Kapitel 1.4 sehen werden, lässt sich auch die in dieser Arbeit behandelte Dichteschätzung als Dichteschätzung für fehlerbehaftete Daten interpretieren. Zuvor stellen wir im nächsten Unterkapitel aber noch ein weiteres Problem vor, mit dem sich diese Arbeit auseinandersetzt.

1.3 Nichtparametrische Regression

Bei der nichtparametrischen Regression handelt es sich um eines der grundlegenden Probleme der mathematischen Statistik. Sind X und Y reellwertige Zufallsvariablen, so versuchen wir den Zusammenhang zwischen X und Y durch eine Funktion f zu erklären, d.h. eine Funktion zu finden, sodass der Abstand zwischen $f(X)$ und Y in einem geeigneten Sinne klein ist. Im Gegensatz zur parametrischen Regression, bei der die prinzipielle Bauart von f als bekannt vorausgesetzt wird, untersucht man in der nichtparametrischen Regressionsschätzung das Problem unter möglichst wenigen Annahmen an die Verteilung von (X, Y). Verwendet man als Fehlerkriterium das L^2-Risiko

$$\mathbf{E}\{|m(X) - Y|^2\},$$

so ist bekannt, dass dieses durch die Regressionsfunktion $m := \mathbf{E}\{Y|X = \cdot\}$ minimiert wird (siehe beispielsweise [17, Section 1.1]). Da die Verteilung von (X, Y), welche die Regressionsfunktion eindeutig bestimmt, im Allgemeinen nicht bekannt ist, stellt sich also die Aufgabe, die Regressionsfunktion m ausgehend von Daten $\mathcal{D}_n = (X_1, Y_1), \ldots, (X_n, Y_n)$ zu schätzen. Da für jeden messbaren Schätzer $\hat{m}_n = \hat{m}_n(\cdot, (X_1, Y_1), \ldots, (X_n, Y_n))$ von m die Gleichung

$$\mathbf{E}\{|\hat{m}_n(X) - Y|^2|\mathcal{D}_n\} = \mathbf{E}\{|m(X) - Y|^2\} + \int |\hat{m}_n(x) - m(x)|^2 \mathbf{P}_X(dx) \qquad (1.3)$$

gilt, ist die Minimierung des L^2-Risikos des Schätzers äquivalent zur Minimierung des L^2-Fehlers $\int|\hat{m}_n(x) - m(x)|^2\mathbf{P}_X(dx)$. Auch dieses Problem wurde bereits ausgiebig untersucht und es sind mehrere Schätzer bekannt, für die

$$\lim_{n\to\infty} \int |\hat{m}_n(x) - m(x)|^2 \mathbf{P}_X(dx) = 0 \quad \text{f.s.}$$

beziehungsweise

$$\lim_{n\to\infty} \mathbf{E} \int |\hat{m}_n(x) - m(x)|^2 \mathbf{P}_X(dx) = 0$$

für alle Verteilungen von (X, Y) mit $\mathbf{E}\{Y^2\} < \infty$ gilt. Solche Regressionsschätzer bezeichnet man als stark bzw. schwach universell konsistent. Schwache und starke Konsistenz wurden beispielsweise bereits für den nächste-Nachbarschätzer (Stone [18], Devroye et al. [19]), den Partitionsschätzer (Györfi [20]) , den Kernschätzer (Nadaraya, Watson [21], [22]), den kleinste-Quadrate-Schätzer (Lugosi,Zeger [23]) und den Splineschätzer (Kohler [24]) nachgewiesen. Da wir später einen kleinste-Quadrate-Schätzer zur Schätzung der Regressionsfunktion verwenden werden, wollen wir diesen bereits hier etwas genauer vorstellen. Im Gegensatz zum Kerndichteschätzer und dem damit verwandten Kernschätzer $\hat{m}_{n,K,h_n}$ von Nadaraya und Watson,

$$\hat{m}_{n,K,h_n}(x) := \frac{\sum_{j=1}^{n} Y_j K\left(\frac{x-X_j}{h_n}\right)}{\sum_{j=1}^{n} K\left(\frac{x-X_j}{h_n}\right)},$$

welche versuchen, die Zielfunktion (also Dichte beziehungsweise Regressionsfunktion) durch lokales Mitteln der beobachteten Werte zu approximieren, handelt es sich beim kleinste-Quadrate-Schätzer um ein globales Schätzverfahren. Ist $\mathcal{F}_n$ eine Klasse von L^1-Funktionen, so ist der zu $\mathcal{F}_n$ und den Daten $(X_1, Y_1), \ldots, (X_n, Y_n)$ gehörige kleinste-Quadrate-Schätzer gegeben durch

$$\hat{f}_n(x) = \hat{f}_n(x, (X_1, Y_1), \ldots, (X_n, Y_n)) = \operatorname*{argmin}_{f \in \mathcal{F}_n} \sum_{j=1}^{n} |f(X_j) - Y_j|^2. \tag{1.4}$$

Der kleinste-Quadrate-Schätzer ist also diejenige Funktion aus $\mathcal{F}_n$, welche den empirischen L^2-Fehler minimiert. Also kann der kleinste-Quadrate- Schätzer nur konsistent sein, wenn die Folge $\mathcal{F}_n$ von Funktionenklassen geeignet gewählt ist. Tatsächlich gilt die Abschätzung (siehe [17, Lemma 10.1])

$$\begin{aligned} \int |\hat{m}_n(y) - m(y)|^2 \mathbf{P}_Y(dy) &\leq \inf_{f \in \mathcal{F}_n} \int |f(y) - m(y)|^2 \mathbf{P}_Y(dy) \\ &+ 2 \sup_{f \in \mathcal{F}_n} \left| \frac{1}{n} \sum_{j=1}^{n} |f(X_j) - Y_j|^2 - \mathbf{E}\{|f(X) - Y|^2\} \right|. \end{aligned} \tag{1.5}$$

Auch für diesen Schätzer wurden starke und schwache Konsistenz ausführlich untersucht (siehe beispielsweise [25], [26] und [27]).

Genau wie bei der Dichteschätzung stellt sich für das Problem der Regressionsschätzung die Frage, ob sich die Zielfunktion ebenfalls schätzen lässt, falls keine exakten Beobachtungen der erklärenden Variable X zur Verfügung stehen, sondern der Regressionsschätzer ausgehend von $(\tilde{X}_1, Y_1), \ldots, (\tilde{X}_n, Y_n)$ konstruiert werden muss. Für dieses „error-in-variables“ Modell gibt es ebenfalls konsistente nicht-parametrische Regressionsschätzer (siehe beispielsweise [28]).

1.4 Latente Variablen

In den vorherigen Abschnitten haben wir bereits darauf verwiesen, dass sich die Probleme der Verteilungs-, Dichte- und Regressionsschätzung auch für Daten mit Messfehlern betrachten lassen. In diesem Abschnitt werden wir das Modell vorstellen, für das wir im

Folgenden die Probleme der Verteilungs-, Dichte-, und Regressionsschätzung behandeln werden. Wir gehen im Folgenden davon aus, dass wir die für uns interesanten Zufallsgrößen nicht direkt beobachten können. Man spricht dann von einem Modell mit sogenannten „latenten Variablen“. Statt der latenten Variablen beobachten wir Transformationen der latenten Variablen, die sogenannten expliziten oder manifesten Variablen. Solche Modelle finden in den Wirtschafts- und Geisteswissenschaften Anwendung und können verwendet werden, um abstrakte, schwer direkt messbare Größen wie beispielsweise Intelligenz, Kreativität, räumliches Vorstellungsvermögen oder Wohlstand zu untersuchen, falls für die zu untersuchenden latenten Größen entsprechende explizite Größen, etwa Bewertungen aus standardisierten Tests, zur Verfügung stehen. Eine Übersicht der behandelten Problemstellungen findet sich in der Survey von Bollen [29] und der Monographie von Skrondal und Rabe-Hesketh [30]. Ein konkretes Beispiel ist Intelligenz als abstrakte, latente Variable, für die als explizite Variablen die Testergebnisse unterschiedlicher Intelligenztests beobachtbar sind. Wir nehmen an, dass die Zusammenhänge zwischen latenten und expliziten Variablen linear sind. Wir normieren die erste dieser linearen Transformationen und erhalten für die Verteilungs- beziehungsweise Dichteschätzung das Modell

$$X = \begin{pmatrix} X^{(1)} \\ X^{(2)} \\ \vdots \\ X^{(d)} \end{pmatrix} = \begin{pmatrix} 1 \cdot Z \\ a_2 Z \\ \vdots \\ a_d Z \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_d \end{pmatrix}, \tag{1.6}$$

wobei X die expliziten Variablen enthält und Z die zu untersuchende latente Variable ist. Die Zufallsvariable $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_n)$ repräsentiert in disem Modell den additiven Messfehler, der die genaue Beobachtung von Z verhindert. Wir gehen außerdem wir davon aus, dass alle vorliegenden expliziten Variablen von genau einer latenten Variable beeinflusst werden, wir müssen uns also nicht mit dem Problem der Identifizierbarkeit, welches für die allgemeineren Strukturgleichungsmodelle etwa in [30] betrachtet wird, auseinandersetzen. Vielmehr ist unser Ziel, unter Verwendung von obigem Modell, die unbekannte Verteilung μ von Z anhand von unabhängig und identisch verteilten Kopien $X_1, \ldots, X_n$ von X zu schätzen. Ausgehend von diesen Beobachtungen der expliziten Variablen werden wir außerdem Schätzer für die Verteilungsfunktion von Z und, falls Z eine Dichte besitzt, auch der Dichte von Z konstruieren und die Konsistenz der Schätzer beweisen.

Für die Regressionsschätzung werden wir das Modell

$$\begin{pmatrix} X^{(1)} \\ X^{(2)} \\ \vdots \\ X^{(d)} \\ Y^{(1)} \\ Y^{(2)} \\ \vdots \\ Y^{(l)} \end{pmatrix} = \begin{pmatrix} 1 \cdot Z_1 \\ a_2 \cdot Z_1 \\ \vdots \\ a_d \cdot Z_1 \\ 1 \cdot Z_2 \\ b_2 \cdot Z_2 \\ \vdots \\ b_l \cdot Z_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_d \\ \delta_1 \\ \delta_2 \\ \vdots \\ \delta_l \end{pmatrix} \tag{1.7}$$

verwenden. $X = (X^{(1)}, \ldots, X^{(d)})$ und $Y = (Y^{(1)}, \ldots, Y^{(l)})$ sind hierbei die expliziten Variablen und Z_1 und Z_2 die latenten Variablen, deren Regressionsfunktion $\mathbf{E}\{Z_2 | Z_1 = \cdot\}$ geschätzt werden soll.

Das Problem der Regressionsschätzung für latente Variablen wurde bisher hauptsächlich im Kontext der parametrischen Regressionsschätzung behandelt, oft im Kontext von Strukturgleichungsmodellen (siehe [31], [30] oder [32]). In [33] wird die hochdimensionale lineare Regressionsfunktion der expliziten Variablen anhand eines latenten-Variablen-Modells untersucht, wobei davon ausgegangen wird, dass alle auftretenden Zufallsvariablen normalverteilt sind. Die zugrundeliegenden latenten Variablen werden mittels der Hauptkomponentenanalyse untersucht. Der einzige dem Autor bekannte nichtparametrische Ansatz findet sich in [34]. In [34] wird ein geschätztes Sample $(\hat{z}_1, \hat{z}_2)_1^n = ((\hat{z}_{1,1}, \hat{z}_{2,1}), \ldots, (\hat{z}_{1,n}, \hat{z}_{2,n})$ der latenten Variablen konstruiert. Sind im Model (1.7) die Zufallsvariablen $(Z_1, Z_2), \varepsilon_1, \ldots, \varepsilon_d, \delta_1, \ldots, \delta_n$ unabhängig und ist ν die gemeinsame Verteilung von (Z_1, Z_2),$\varepsilon_1, \ldots, \varepsilon_d$,$\delta_1, \ldots, \delta_n$, so gilt

$$\nu\left(\prod_{j=1}^{1+d+l} B_j\right) = \nu\left(B_1 \times \prod_{j=1}^{1+d+l} \mathbb{R}\right) \prod_{l=2}^{1+d+l} \nu\left(\mathbb{R}^2 \times \prod_{j=2}^{l-1} \mathbb{R} \times B_l \times \prod_{j=l+1}^{1+d+l} \mathbb{R}\right)$$

Für alle B_1 aus der borrelschen σ-Algebra auf $\mathbb{R}^2$ und alle B_j mit $j \geq 2$ aus der borrelschen σ-Algebra auf $\mathbb{R}$. Das geschätzte Sample $(\hat{z}_1, \hat{z}_2)_1^n$ wir in [34] derart konstruiert, dass die obige Beziehung für das zu $(\hat{z}_1, \hat{z}_2)_1^n$ gehörige empirische Maß auf geeigneten Intervalle approximativ gilt. In [34] wird gezeigt, dass das zu $(\hat{z}_1, \hat{z}_2)_1^n$ gehörige empirische Maß schwach gegen ν konvergiert, und dass es möglich ist ausgehend von $(\hat{z}_1, \hat{z}_2)_1^n$ einen stark konsistenten Splineschätzer m_n für die Regressionsfunktion $m(\cdot) := \mathbf{E}\{Z_2 | Z_1 = \cdot\}$ zu konstruieren. Wie wir im Folgenden Abschnitt erläutern werden, gehen die neuen Resultate in dieser Arbeit über die in [34] gezeigten hinaus.

1.5 Neue Resultate in dieser Arbeit

In diesem Abschnitt der Einleitung werden die neuen Resultate dieser Doktorarbeit vorgestellt. Dabei werden wir für die entsprechenden Theoreme nicht alle benötigten technischen Voraussetzungen auflisten, sondern verweisen für die vollständige Formulierung auf die folgenden Kapitel. Stattdessen geben wir für die hier präsentierten Resultate jeweils nur die entscheidenden Voraussetzungen an, welche das Theorem von den anderen Resultaten dieser Arbeit absetzen. Der Struktur der bisherigen Abschnitte der Einleitung folgend, stellen wir zunächst die grundlegenden Ideen hinter allen in dieser Arbeit neu gezeigten Konvergenzresultaten vor und präsentieren danach Resultate zur Verteilungsschätzung, zur Schätzung der Verteilungsfunktion, zur Dichteschätzung und zur Regressionsschätzung. Wir beginnen mit einer grundlegenden Beobachtung, welche alle hier vorgestellten Schätzverfahren motiviert. Sei

$$\varphi_Z(u) := \mathbf{E}\{e^{iuZ}\} = \int e^{iuz}\, \mathbf{P}_Z(dz)$$

die charakteristische Funktion von Z und sei

$$\varphi_X(u_1, \ldots, u_d) := \int \exp\left(i \sum_{l=1}^{d} u_l x_l\right) \mathbf{P}_X(dx)$$

die charakteristische Funktion von $X = (X^{(1)}, \dots, X^{(d)})$. Setzt man in Modell (1.6) voraus, dass die auftretenden Messfehler $\varepsilon_1, \dots, \varepsilon_d$ von Z unabhängig sind, so folgt

$$\begin{aligned}
\varphi_X(u_1, \dots, u_d) &= \mathbf{E}\left\{\exp\left(i\sum_{j=1}^{d} u_j X^{(j)}\right)\right\} \\
&= \mathbf{E}\left\{\exp\left(i\sum_{j=1}^{d} u_j (a_j Z + \varepsilon_j)\right)\right\} \\
&= \mathbf{E}\left\{\exp\left(i\sum_{j=1}^{d} u_j a_j Z\right) \cdot \exp\left(i\sum_{j=1}^{d} u_j \varepsilon_j\right)\right\} \\
&= \varphi_Z\left(\sum_{j=1}^{d} a_j u_j\right) \cdot \prod_{j=1}^{d} \varphi_{\varepsilon_j}(u_j).
\end{aligned}$$

Zusammen mit $\varphi_{\varepsilon_j}(0) = 1$ und $\frac{\partial}{\partial u_j}\varphi_{\varepsilon_j}(0) = \mathbf{E}\{\varepsilon_j\} = 0$ folgt

$$\varphi_X((u_1, 0, \dots, 0)) = \varphi_Z(u_1) \cdot \varphi_{\varepsilon_1}(u_1) \tag{1.8}$$

und

$$\frac{\partial}{\partial u_2}\varphi_X(u_1, 0, \dots, 0) = a_2 \frac{\partial}{\partial u_1}\varphi_Z(u_1)\varphi_{\varepsilon_1}(u_1). \tag{1.9}$$

In dieser Arbeit zeigen wir, dass die expliziten Variablen $(X_1, \dots, X_n)$ aus Modell (1.6) genutzt werden können, um ein Sample $\hat{z}_1, \dots, \hat{z}_n$ der latenten Variable Z zu schätzen, sodass (1.8) und (1.9) für die empirischen charakteristischen Funktionen

$$\hat{\varphi}_{X_1^n}(u_1, \dots, u_d) := \frac{1}{n}\sum_{j=1}^{n} \exp\left(i\sum_{l=1}^{d} u_l X_j^{(l)}\right) \tag{1.10}$$

approximativ gelten, also so, dass

$$\frac{1}{n}\sum_{j=1}^{n} e^{iuX_j^{(1)}} \approx \frac{1}{n}\sum_{j=1}^{n} e^{iu\hat{z}_j} \cdot \frac{1}{n}\sum_{j=1}^{n} e^{iu\hat{\varepsilon}_{1,j}}$$

und

$$\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(u, 0, \dots, 0) \approx \hat{a}_2 \frac{\partial}{\partial u}\frac{1}{n}\sum_{j=1}^{n} e^{iu\hat{z}_j} \cdot \frac{1}{n}\sum_{j=1}^{n} e^{iu\hat{\varepsilon}_{1,j}}$$

gelten, wobei $\hat{a}_2$ ein ebenfalls aus $X_1, \dots, X_n$ konstruierter Schätzer für a_2 ist und $\hat{\varepsilon}_{1,1}$ bis $\hat{\varepsilon}_{1,n}$ durch

$$\hat{\varepsilon}_{1,l} := X_l^{(1)} - \hat{a}_1 \hat{z}_l.$$

gegeben sind. In dieser Arbeit zeigen wir, dass die obigen Approximationsfehler

$$\left\| \frac{1}{n}\sum_{j=1}^{n} e^{iuX_j^{(1)}} - \frac{1}{n}\sum_{j=1}^{n} e^{iu\hat{z}_j} \cdot \frac{1}{n}\sum_{j=1}^{n} e^{iu\hat{\varepsilon}_{1,j}} \right\|$$

und

$$\left\| \frac{\partial}{\partial u_2} \hat{\varphi}_{X_1^n}(u, 0, \dots, 0) - \hat{a}_2 \frac{\partial}{\partial u} \frac{1}{n} \sum_{j=1}^{n} e^{iu\hat{z}_j} \cdot \frac{1}{n} \sum_{j=1}^{n} e^{iu\hat{\varepsilon}_{1,j}} \right\|$$

für $n \to \infty$ gegen 0 konvergieren, falls die vierten Momente von Z und von den ε_j endlich sind und die auftretenden charakteristischen Funktionen nirgends verschwinden. Aus diesem geschätzten Sample $\hat{z}_1, \dots, \hat{z}_n$ werden wir im Folgenden unsere Verteilungs-, Dichte- und Regressionsschätzer konstruieren.

Wir beginnen mit dem Verteilungsschätzer $\hat{\mu}_n$, den wir als empirisches Maß bezüglich der geschätzten Daten $\hat{z}_1, \dots, \hat{z}_n$, also für $B \in \mathcal{B}$ durch

$$\hat{\mu}_n(B) := \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_B(\hat{z}_j)$$

definieren. Dieser Schätzer ist konsistent für Intervalle. Zentrale Vorraussetzung für die folgenden Resultate ist die Nullstellenfreiheit der charakteristischen Funktionen von Z und der ε_j. Sind die auftretenden charakteristischen Funktionen nullstellenfrei, so gilt, unter weiteren technischen Regularitätsannahmen, die wir in Kapitel 1.2 genauer ausführen, das folgende Theorem.

Theorem 1.1. *Ist $[a, b]$ ein Intervall mit $\mathbf{P}_Z(a) = \mathbf{P}_Z(b) = 0$, so gilt*

$$\hat{\mu}_n([a, b]) \to \mathbf{P}_Z([a, b])$$

fast sicher.

Bemerkung. In Kapitel 2 befindet sich die exakte Formulierung dieses Resultats und der benötigten Annahmen als Theorem 2.6.

Analog zur Situation der klassischen Dichtschätzung, benötigen wir weitere Annahmen um die gleichmäßige Konvergenz des Schätzers zu beweisen. Besitzt $\mathbf{P}_Z$ eine beschränkte Dichte, so erhalten wir auf einer fast sicher die reelle Achse ausschöpfenden Folge von Intervallen auch gleichmäßige Konvergenz. Unter Regularitätsannahmen gilt das folgende Theorem.

Theorem 1.2. *Besitzt $\mathbf{P}_Z$ eine beschränkte Dichte, so existiert eine Folge $(\hat{K}_n)_{n \in \mathbb{N}}$ von reellwertigen, positiven Zufallsvariablen mit $\hat{K}_n \to \infty$ fast sicher, so dass mit Wahrscheinlichkeit 1*

$$\sup_{a, b \in [-\hat{K}_n, \hat{K}_n]} |\hat{\mu}_n([a, b]) - \mathbf{P}_Z([a, b])| \to 0$$

gilt.

Bemerkung. In Kapitel 2 befindet sich die exakte Formulierung diese Resultats und der benötigten Annahmen als Theorem 2.7.

Für beide Resultate werden nur Momentenbedingungen und die Nullstellenfreiheit der auftretenden charakteristischen Funktionen benötigt. Insbesondere benötigen wir keinerlei Information über die Verteilung der Messfehler $\varepsilon_1, \dots, \varepsilon_k$ oder über die Glattheit der Dichte. Die Zufallsvariablen $\hat{K}_n$, deren Divergenzverhalten als Indikator für die Konvergenzgüte verstanden werden kann, hängen dabei nur von den Daten $(X_1, \ldots, X_n)$ und der charakteristischen Funktion der latenten Variable Z ab und können explizit aus $(X_1, \ldots, X_n)$ berechnet werden. Wie wir in Theorem 2.7 zeigen werden, lässt sich außerdem die Konvergenzgeschwindigkeit des Verteilungsschätzers durch $\hat{K}_n$ beschreiben.

Ausgehend von der Verteilungsschätzung können wir auch die Verteilungsfunktion $F_Z(t) := \mathbf{P}_Z((-\infty, t])$ schätzen. Ist s_n eine Folge von Zufallsvariablen so definieren wir den Schätzer $\hat{F}_{\hat{z}_1^n}(t)$ der Verteilungsfunktion als

$$\hat{F}_{\hat{z}_1^n}(t) := \begin{cases} 0 & \text{für } t \leq -s_n \\ \hat{\mu}_n([-\hat{K}_n, t]) & \text{für } t \in (-s_n, s_n) \\ 1 & \text{für } t \geq s_n. \end{cases} \tag{1.11}$$

Dann gilt das folgende Theorem, welches analog zum Satz von Glivenko-Cantelli die gleichmäßige Konvergenz des Verteilunktionsfunktionsschätzers gewährleistet.

Theorem 1.3. *Sei $\hat{F}_n$ der Verteilungsfunktionsschätzer aus (1.11). Sei $\hat{K}_n$ die Folge von Zufallsvariablen aus dem vorherigen Theorem. Dann existiert eine Funktion f, so dass für $s_n = f(\hat{K}_n)$ mit Wahrscheinlichkeit 1*

$$\sup_{t \in \mathbb{R}} |\hat{F}_n(t) - F_Z(t)| \to 0$$

gilt.

Bemerkung. In Kapitel 2 befindet sich die exakte Formulierung dieses Resultats und der benötigten Annahmen als Theorem 2.10.

Außerdem lässt sich die zu den geschätzten Daten gehörige empirische Verteilung $\hat{\mu}_n$ verwenden, um die Dichte f_Z von Z zu schätzten. Wegen

$$\hat{\mu}_n([a, b]) = \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_{[a,b]}(\hat{z}_j)$$

verwenden wir den durch

$$\hat{f}_n(x) = \frac{1}{n} \sum_{j=1}^{n} K\left(\frac{x - \hat{z}_j}{h_n}\right) \tag{1.12}$$

gegebenen Kerndichteschätzer mit naivem Kern $K(x) = \mathbb{1}_{[-\frac{1}{2}, \frac{1}{2}]}(x)$ und Bandbreite h_n. Für diesen Schätzer gilt, unter entsprechenden Regularitätsannahmen, das folgende Theorem.

Theorem 1.4. *Sei $\hat{f}_n$ definiert wie in (1.12). Dann existiert eine Folge von zufälligen Bandbreiten $\hat{h}_n$, so dass*

$$\int |f(t) - \hat{f}_n(t)| \, dt \to 0$$

mit Wahrscheinlichkeit 1 gilt.

Bemerkung. Die zufälligen Bandbreiten $\hat{h}_n$ im obigen Theorem hängen, genau wie die Folge der $\hat{K}_n$ nur von den Daten $X_1, \ldots, X_n$ und der Verteilung von Z ab. Außerdem gilt $\hat{h}_n \to 0$ fast sicher.

Bemerkung. In Kapitel 2 befindet sich die exakte Formulierung dieses Resultats und der benötigten Annahmen als Theorem 2.12.

Wie schon beim Verteilungsschätzer werden auch beim Dichteschätzer unter stärkeren Annahmen, insbesondere an die Glattheit der Dichte, Aussagen zur Konvergenzgeschwindigkeit bewiesen. In Theorem 2.16 zeigen wir außerdem, dass für normalverteilte Zufallsvariablen eine deterministische untere Schranke für die Konvergenzgeschwindigkeit hergeleitet werden kann.

Ausgehend von den beobachtbaren expliziten Variablen wird in der vorliegenden Arbeit also eine geschätzte Stichprobe der unbeobachtbaren latenten Variable konstruiert. Diese geschätzte Stichprobe wird dann verwendet, um konsistente Schätzer für die Verteilung, für die Verteilungsfunktion und für die Dichte der zugrundliegenden latenten Variable zu konstruieren. Während die parametrische Dichteschätzung für latente Variablen bereits ausgiebig untersucht wurde (siehe beispielsweise [35]), stellen diese Resultate, soweit dem Autor bekannt, die ersten im Bereich der nichtparametrischen Dichteschätzung für latente Variablen dar.

Zur Schätzung der Regressionsfunktion der latenten Variablen

$$m(\cdot) = \mathbf{E}\{Z_2 | Z_1 = \cdot\}$$

gehen wir ähnlich zur Dichteschätzung vor. Sind im Modell (1.7) die Fehler $\varepsilon_1, \dots, \varepsilon_d$ und $\delta_1, \dots, \delta_n$ von Z_1 und Z_2 unabhängig, so erhalten wir die zu den Identitäten (1.8) und (1.9) analogen Gleichungen

$$\varphi_{(X,Y)}(u_1, 0, \dots, 0, v_1, 0, \dots, 0) = \varphi_{(Z_1,Z_2)}(u_1, v_1) \cdot \varphi_{\varepsilon_1}(u_1) \cdot \varphi_{\delta_1}(v_1),$$

$$\begin{aligned}&\frac{\partial}{\partial u_2}\varphi_{(X,Y)}(u_1, 0, \dots, 0, v_1, 0, \dots, 0)\\ &\qquad = a_2 \cdot \frac{\partial}{\partial u_1}\varphi_{(Z_1,Z_2)}(u_1, v_1) \cdot \varphi_{\varepsilon_1}(u_1) \cdot \varphi_{\delta_1}(v_1)\end{aligned}$$

und

$$\begin{aligned}&\frac{\partial}{\partial v_2}\varphi_{(X,Y)}(u_1, 0, \dots, 0, v_1, 0, \dots, 0)\\ &\qquad = b_2 \cdot \frac{\partial}{\partial v_1}\varphi_{(Z_1,Z_2)}(u_1, v_1) \cdot \varphi_{\varepsilon_1}(u_1) \cdot \varphi_{\delta_1}(v_1).\end{aligned}$$

Ausgehend von den expliziten Variablen $(X_1, Y_1), \dots, (X_n, Y_n)$ konstruieren wir ein geschätztes Sample $\hat{z}_{1,1}, \dots, \hat{z}_{1,n}, \hat{z}_{2,1}, \dots, \hat{z}_{2,n}$ der latenten Variablen, welches diese Gleichungen approximativ erfüllt. Ausgehend von dieser geschätzten Stichprobe konstruieren wir den Regressionsschätzer $\hat{m}_n$ als kleinste-Quadrate-Schätzer zu einer geeignet gewählten Menge $\mathcal{F}_n$ von trigonometrischen Polynomen, es gilt also

$$\hat{m}_n := \operatorname{argmin}_{f \in \mathcal{F}_n} \sum_{j=1}^{n} \left(f(\hat{z}_{1,j}) - \hat{z}_{2,j}\right)^2. \tag{1.13}$$

Für den so konstruierten Schätzer $\hat{m}_n$ gilt das folgenden Theorem.

Theorem. *Gilt in* (1.7) $d \geq 3$ *und* $l \geq 3$*, sind* $(Z_1, Z_2), \varepsilon_1, \ldots, \varepsilon_d, \delta_1, \ldots, \delta_l$ *unabhängig, sind die* 8*-ten Momente von* Z_1 *und* Z_2 *sowie* $\varepsilon_1, \ldots, \delta_l$ *endlich, und ist die charakteristische Funktion* $\varphi_{(X,Y)}(u, v) = \int \exp(i(\langle u, X\rangle + \langle v, Y\rangle)) \, d\mathbf{P}$ *von* (X, Y) *nullstellenfrei, so existiert eine, ausgehend von den Daten* (X_1, Y_1)*, ...,* (X_n, Y_n) *explizit konstruierbare, Folge* $\mathcal{F}_n$ *von zufälligen Funktionenräumen, so dass für* $\hat{m}_n$*, definiert wie in* (1.13)*, mit Wahrscheinlichkeit* 1

$$\int |\hat{m}_n(x) - m(x)|^2 \mathbf{P}_{Z_1}(dx) \to 0$$

gilt.

Auch hierbei benötigen wir keine Annahmen über die Verteilung der Messfehler ε_j, δ_j außer Zentriertheit. Es werden ebenfalls keinerlei Annahmen über die Glattheit der Regressionsfunktion benötigt. Es genügen Momentenbedingungen für Z_1, Z_2, sowie die Nullstellenfreiheit der auftretenden charakteristischen Funktionen.

Bisherige Ansätze zur Bestimmung der Regressionsfunktion für latente Variablen waren größtenteils parametrischer Natur. Die einzige dem Autor bekannte Arbeit, welche für dieses Problem einen nichtparametrischen Ansatz wählt, ist [34]. Die in der vorliegenden Arbeit bewiesenen Resultate gehen über die dortigen Resultate hinaus, da in [34] zur Konstruktion des Splineschätzers für die Regressionsfunktion benötigt wird, dass die Regressionsfunktion Lipschitz-stetig ist und dass eine obere Schranke für die Lipschitz-Konstante bekannt ist. Der in der vorliegenden Arbeit vorgestellte Regressionsschätzer benötigt keinerlei solche Annahmen an die Regressionsfunktion. Die in der vorliegenden Arbeit benötigten Voraussetzungen an die charakteristische Funktion der latenten Variablen werden in ähnlicher Form in [34] ebenfallls benötigt.

2 Schätzung der Verteilung bei latenten Variablen

In diesem Abschnitt der Arbeit werden wir uns mit der Schätzung der Verteilung und der Dichtefunktion von latenten Variablen beschäftigen. Da wir die latenten Variablen nicht direkt beobachten können, konstruieren wir, nachdem das zugrundeliegende Modell in Abschnitt 2.1 vorgestellt wurde, aus den beobachteten Variablen geschätzte Daten unserer latenten Variable. Mit diesen geschätzten Daten definieren wir in Abschnitt 2.2 den Schätzer für die Verteilung, in Abschnitt 2.4 den Schätzer für die Verteilungsfunktion und in Abschnitt 2.5 den Schätzer für die Dichte der latenten Variable. Für diese Schätzer beweisen wir jeweils punktweise Konsistenz, d.h. Konvergenz gegen die zu schätzende Größe im Grenzwert |„Menge der Daten "| $= n \to \infty$. Unter stärkeren Voraussetzungen werden auch Aussagen zur gleichmäßigen Konsistenz und zur Konvergenzgeschwindigkeit, im jeweils geeigneten Sinne, hergeleitet. Wichtigstes Hilfsresultat hierzu ist die Konvergenz der empirischen charakteristischen Funktion der geschätzten Daten, die im nächsten Kapitel in Abschnitt 3.1 vorgestellt und bewiesen wird.

2.1 Modell

Unser Ziel ist die Schätzung der Verteilung und der Dichte einer latenten Variablen Z, d.h. einer Zufallsgröße, die von uns nicht direkt gemessen werden kann. Wir haben jedoch ein Sample der sogenannten manifesten Variable

$$X = (X^{(1)}, \dots, X^{(d)})$$

deren Komponenten *linear* von Z abhängen.

Definition 2.1 (Vorläufige Definition)**.** Das lineare latente-Variablen-Modell in einer latenten Variablen und d manifesten Variablen ist gegeben durch

$$X = \begin{pmatrix} X^{(1)} \\ X^{(2)} \\ \vdots \\ X^{(d)} \end{pmatrix} = \begin{pmatrix} a_1 Z \\ a_2 Z \\ \vdots \\ a_d Z \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_d \end{pmatrix}, \tag{2.1}$$

wobei $a_1, \dots, a_d \in \mathbb{R}$ sind, Z eine reellwertige Zufallsvariable ist, und die $\varepsilon_1, \dots, \varepsilon_d$ von Z unabhängige Zufallsvariablen mit $\mathbf{E}\{\varepsilon_1\} = \dots = \mathbf{E}\{\varepsilon_d\} = 0$ sind.

Gibt es ein $a_j \in \{a_1, \dots, a_n\}$, welches nicht Null ist, so können wir durch Umsortierung der Vektoren und Renormierung von Z und der ε_j im Folgenden o.B.d.A. annehmen, dass in (2.1) $a_1 = 1$ gilt. Ist dies nicht der Fall, sind also alle Koeffizienten $a_j = 0$, so ist es offensichtlich unmöglich, irgendwelche Aussagen über Z oder die Verteilung von Z aus unabhängigen und zu $X = (X^{(1)}, \dots, X^{(d)})$ identisch verteilten manifesten Variablen $X_1, \dots, X_n$ herzuleiten. Unter welchen Voraussetzungen die Verteilung von Z durch die Verteilung von X eindeutig bestimmt ist, wurde in der Literatur bereits untersucht. In [34] wird beispielsweise für das durch Gleichung (1.7) beschriebene Modell für die Regressionschätzung untersucht, unter welchen Voraussetzungen die Verteilung von (Z_1, Z_2) durch die Verteilung von (X, Y) eindeutig bestimmt ist. Durch eine kleine Modifikation der in [34] vorgestellten Resultate erhält man das folgende Lemma, welches wir in Kapitel 3.1 beweisen werden.

Lemma 2.2. *Seien X, Z, $\varepsilon_0, \ldots, \varepsilon_d$ und $a_1, \ldots, a_d$ wie in der vorläufigen Definition 2.1. Das heißt, seien X, Z, $\varepsilon_0, \ldots, \varepsilon_d$ wie in Gleichung (2.1), seien $\varepsilon_0, \ldots, \varepsilon_d$ von Z unabhängig. Es gelte $\mathbf{E}\{\varepsilon_j\} < \infty$ für $j \in \{1, \ldots, d\}$ und $\mathbf{E}\{\varepsilon_1\} = \cdots = \mathbf{E}\{\varepsilon_d\} = 0$. Sei Z zudem quadratisch integriebar mit $\mathbf{E}\{Z^2\} < \infty$ sowie $\mathbf{E}\{Z^2\} \neq 0$ und es gelte $\varphi_X(t) := \mathbf{E}\{\exp(i\langle X, t\rangle)\} \neq 0$ für alle $t \in \mathbb{R}^d$ sowie $\varphi_{\varepsilon_j}(t) \neq 0$ für alle $j \in \{1, \ldots, d\}$ und alle $t \in \mathbb{R}$. Es gelte außerdem $d \geq 3$ und $a_{j_i} \neq 0$ für mindestens 3 verschiedene $j_1, j_2, j_3 \in \{1, \ldots, d\}$. Dann ist die Verteilung von Z eindeutig durch die Verteilung von X bestimmt.*

Wie schon zuvor bemerkt, können wir o.B.d.A. davon ausgehen, dass $a_1 = 1$ gilt, falls nicht $a_j = 0$ für alle $j \in \{1, \ldots, d\}$ gilt. Sind insgesamt d' der a_j von 0 verschieden, so können wir durch Umsortierung der Einträge von X ebenfalls o.B.d.A. annehmen, das die von 0 verschiedenen Koeffizienten den Vektor $a = (a_1, \ldots, a_d)$ anführen, also dass $a_j \neq 0$ für alle $j \leq d'$ gilt. Wir erhalten die endgültige Definition unseres Modells, indem wir diese, die Allgemeinheit nicht beschränkenden Annahmen sowie entsprechende, zu den Voraussetzungen von Lemma 2.2 äquivalente, Voraussetzungen mit in die Definition aufnehmen.

Definition 2.3. Ist Z eine reellwertige Zufallsvariable mit $\mathbf{E}\{Z^2\} \neq 0$ und $\varphi_Z(u) \neq 0$ für alle $u \in \mathbb{R}$, sind $a_1, \ldots, a_d \in \mathbb{R}$ mit $a_1 = 1$,$a_2 \neq 0$ und $a_3 \neq 0$ und sind $\varepsilon_1, \ldots, \varepsilon_d$ von Z unabhängige reellwertige Zufallsvariablen mit $\mathbf{E}\{\varepsilon_1\} = \cdots = \mathbf{E}\{\varepsilon_d\} = 0$ und $\varphi_{\varepsilon_j}(u) \neq 0$ für alle $u \in \mathbb{R}$ und alle $j \in \{1, \ldots, d\}$, so ist das Lineare-latente-Variablen-Modell in einer latenten Variable gegeben durch

$$X = \begin{pmatrix} X^{(1)} \\ X^{(2)} \\ \vdots \\ X^{(d)} \end{pmatrix} = \begin{pmatrix} a_1 Z \\ a_2 Z \\ \vdots \\ a_d Z \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_d \end{pmatrix} \tag{2.2}$$

2.2 Definition des Verteilungsschätzers

Im Folgenden gehen wir davon aus, dass ein Sample von X vorliegt, welches wir verwenden werden, um die Verteilung, die Verteilungsfunktion und die Dichte von Z zu schätzen. Seien also im Folgenden $Z, Z_1, \ldots, Z_n$ reellwertig und identisch verteilt und ebenso die $\mathbb{R}^d$-wertigen $\varepsilon, \varepsilon_1, \ldots, \varepsilon_n$ identisch verteilt. Außerdem seien $Z, Z_1, \ldots, Z_n, \varepsilon, \varepsilon_1, \ldots, \varepsilon_n$ unabhahängig. Für $j \in \{1, \ldots, d\}$ und $l \in \{1, \ldots, n\}$ sei $\varepsilon_{j,l}$ die j-te Komponente von ε_l. Nach Definition unseres Modells sind auch Z und ε unabhängig, was impliziert das auch Z_l und ε_l für alle $l \in \{1, \ldots, n\}$ unabhängig sind. Nun definieren wir

$$X_l^{(j)} := a_j Z_l + \varepsilon_{j,l}$$

und

$$X_l = \begin{pmatrix} X_l^{(1)} \\ \vdots \\ X_l^{(d)} \end{pmatrix}.$$

Damit sind auch $X_1, \ldots, X_n$ unabhängig und identisch verteilt. Ausgehend von $X_1^n = X_1, \ldots, X_n$ werden wir in diesem Kapitel nun konsistente Schätzer für die Verteilung, die Verteilungsfunktion und die Dichte von Z konstruieren. Grundlegendes Werkzeug für die Konstruktion der Schätzer in diesem Kapitel ist der in diesem Abschnitt vorgestellte Verteilungsschätzer $\hat{\mu}_n$ für die Verteilung μ von Z. Liegt ein unabhängig und identisch verteiltes Sample $Y_1, \ldots, Y_n$ einer Zufallsvariablen Y vor, so lässt sich die Verteilung ν von Y klassischerweise durch das bereits in der Einleitung in Abschnitt 1.1 vorgestellte empirische Maß

$$\nu_n(A) = \nu_n(A, Y_1, \ldots, Y_n) = \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_{\{A(Y_j)\}}$$

schätzen. Da in unserem Fall kein unabhängig und identisch verteiltes Sample der Zufallsvariable Z, deren Verteilung wir schätzen wollen, vorliegt, halten wir uns an das Motto „Hast du keins, dann mach dir eins“ und konstruieren im Folgenden ein geschätztes Sample $\hat{z}_1, \ldots, \hat{z}_n$ aus unseren unabhängig und identisch verteilten $X_1, \ldots, X_n$.

Wir beginnen mit der Schätzung der Koeffizienten $a_1, \ldots, a_d$. Nach Voraussetzung gilt $a_1 = 1$. Da die ε_j von Z und untereinander unabhängig sind und da $\mathbf{E}\{\varepsilon_j\} = 0$ für alle $j \in \{1, \ldots, d\}$ gilt, folgt

$$\begin{aligned}
&\mathbf{E}\{X^{(j)} X^{(i)}\} = \mathbf{E}\{(a_j Z + \varepsilon_j)(a_i Z + \varepsilon_i)\} \\
&\quad = a_j a_i \mathbf{E}\{Z^2\} + \mathbf{E}\{a_j Z + \varepsilon_j\}\mathbf{E}\{\varepsilon_i\} + \mathbf{E}\{a_i Z + \varepsilon_i\}\mathbf{E}\{\varepsilon_j\} + \mathbf{E}\{\varepsilon_j\}\mathbf{E}\{\varepsilon_i\} \\
&\quad = a_j a_i \mathbf{E}\{Z^2\}.
\end{aligned}$$

Nach Vorausestzung gilt $a_1 = 1$, $a_2 \neq 0$ und $a_3 \neq 0$. Zusammen mit $0 < \mathbf{E}\{Z^2\} < \infty$ folgt

$$a_2 = \frac{a_2 a_3 \mathbf{E}\{Z^2\}}{a_1 a_3 \mathbf{E}\{Z^2\}} = \frac{\mathbf{E}\{X^{(2)} X^{(3)}\}}{\mathbf{E}\{X^{(1)} X^{(3)}\}} \quad \text{und} \quad a_j = \frac{\mathbf{E}\{X^{(2)} X^{(j)}\}}{\mathbf{E}\{X^{(1)} X^{(2)}\}}$$

für $j \in \{3, \ldots, d\}$. Ersetzen der Erwartungswerte durch empirische Mittel liefert die Schätzer

$$\hat{a}_2 = \frac{\sum_{l=1}^{n} X_l^{(2)} X_l^{(3)}}{\sum_{l=1}^{n} X_l^{(1)} X_l^{(3)}} \quad \text{und} \quad \hat{a}_j = \frac{\sum_{l=1}^{n} X_l^{(2)} X_l^{(j)}}{\sum_{l=1}^{n} X_l^{(1)} X_l^{(2)}}$$

für $j \in \{3, \ldots, d\}$, sowie $\hat{a}_1 := 1$. Wegen $X_l^{(j)} - a_j Z_l = \varepsilon_{j,l}$ für $j \in \{1, \ldots, d\}$ und $l \in \{1, \ldots, n\}$ erhalten wir aus jeder Schätzung $\hat{z}_1, \ldots \hat{z}_n$ von $Z_1, \ldots Z_n$ eine Schätzung $\hat{\varepsilon}_{j,l}$ der Fehler durch

$$\hat{\varepsilon}_{j,l} := X_l^{(j)} - \hat{a}_j \hat{z}_l.$$

Wir können also die $\hat{\varepsilon}_{j,l}$ als Funktion der $\hat{z}_1, \ldots \hat{z}_n$ auffassen und benötigen damit nur noch diese Schätzung der Z_j. In Kapitel 1.5 haben wir gezeigt, dass

$$\varphi_X(u_1, 0, \ldots, 0) = \varphi_Z(u_1)\varphi_{\varepsilon_1}(u_1)$$

und

$$\frac{\partial}{\partial u_2}\varphi_X(u_1, 0, \ldots, 0) = a_2 \frac{\partial}{\partial u_1}\varphi_Z(u_1)\varphi_{\varepsilon_1}(u_1)$$

gilt. Diese Beziehung soll approximativ auch für die empirische charakteristische Funktion von X und die empirischen charakteristischen Funktionen unserer geschätzten Daten und Fehler gelten. Da die charakteristische Funktion φ_X nullstellenfrei ist, kann die Approximation durch

$$\hat{\varphi}_{X_1^n}(u) = \frac{1}{n} \sum_{j=1}^{n} \exp\left(i(\langle u, X_j \rangle\right)$$

nur dort gelingen, wo auch $\hat{\varphi}_{X_1^n}$ vom Ursprung wegbeschränkt ist.

Definition 2.4 (Definition der geschätzten Daten)**.** Seien Δ_n und u_n zwei reellwertige Folgen. Wir definieren $\hat{K}_n$ in Abhängigkeit von diesen Folgen als

$$\hat{K}_n := \min\{u_n, \max\{x \in \mathbb{R} : \inf_{u\in[0,x]} |\hat{\varphi}_{X_1^n}(u, 0, \dots, 0)| \geq \Delta_n\}\}. \tag{2.3}$$

Weiterhin definieren wir den empirischen Approximationsfehler der Daten im Frequenzbild, $T_n : \mathbb{R}^n \to \mathbb{R}$, durch

$$\begin{aligned} T_n(z) = \sup_{u\in[0,\hat{K}_n]} \Big[&\big|\hat{\varphi}_{X_1^n}(u, 0, \dots) - \hat{\varphi}_z(u) \cdot \hat{\varphi}_{\hat{\varepsilon}_1^n}(u)\big| \\ &+ \left|\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(u, 0, \dots) - \hat{a}_2 \frac{\partial}{\partial u}\hat{\varphi}_z(u) \cdot \hat{\varphi}_{\hat{\varepsilon}_1^n}(u)\right| \Big]. \end{aligned} \tag{2.4}$$

Dann erhalten wir $\hat{z}_1^n = (\hat{z}_1, \dots, \hat{z}_n)$ als Minimum von $T_n(z)$ unter der Nebenbedingung

$$\frac{1}{n}\sum_{j=1}^n |\hat{z}_j|^2 \leq \frac{1}{n}\sum_{j=1}^n |X_j^{(1)}|^2 + 1. \tag{2.5}$$

Bemerkung. In Definition 2.4 werden die geschätzten Daten als Minimum der Funktion $T_n(z)$ unter der Nebenbedingung (2.5) definiert. Wir sollten sicherstellen, dass dieses Minimum tatsächlich existiert. Offensichtlich sind die Funktionen

$$(u, z) \mapsto \hat{\varphi}_{X_1^n}(u, 0, \dots) = \frac{1}{n}\sum_{j=1}^n \exp\left(iuX_j^{(1)}\right),$$

$$(u, z) \mapsto \hat{\varphi}_z(u) \cdot \hat{\varphi}_{\hat{\varepsilon}_1^n}(u) = \Big(\frac{1}{n}\sum_{j=1}^n \exp\left(iuz_j\right)\Big) \cdot \Big(\frac{1}{n}\sum_{j=1}^n \exp\left(iu(X_j^{(1)} - z_j)\right)\Big),$$

$$(u, z) \mapsto \frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(u, 0, \dots) = \frac{1}{n}\sum_{j=1}^n iX_j^{(2)} \exp\left(iuX_j^{(1)}\right),$$

und

$$\begin{aligned} (u, z) \mapsto \hat{a}_2 &\frac{\partial}{\partial u}\hat{\varphi}_z(u) \cdot \hat{\varphi}_{\hat{\varepsilon}_1^n}(u) \\ &= \Big(\frac{\sum_{l=1}^n X_l^{(2)}X_l^{(3)}}{\sum_{l=1}^n X_l^{(1)}X_l^{(3)}} \frac{1}{n}\sum_{j=1}^n iz_j \exp\left(iuz_j\right)\Big) \cdot \Big(\frac{1}{n}\sum_{j=1}^n \exp\left(iu(X_j^{(1)} - z_j)\right)\Big) \end{aligned}$$

stetige Funktionen in u und z. Also ist auch

$$\begin{aligned} F(u, z) := &\big|\hat{\varphi}_{X_1^n}(u, 0, \dots) - \hat{\varphi}_z(u) \cdot \hat{\varphi}_{\hat{\varepsilon}_1^n}(u)\big| \\ &+ \left|\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(u, 0, \dots) - \hat{a}_2 \frac{\partial}{\partial u}\hat{\varphi}_z(u) \cdot \hat{\varphi}_{\hat{\varepsilon}_1^n}(u)\right| \end{aligned}$$

von $\mathbb{R}^{1+n}$ nach $\mathbb{R}$ stetig. Da sowohl $[0, \hat{K}_n]$ als auch

$$\hat{M} = \left\{ z \in \mathbb{R}^n : \frac{1}{n}\sum_{j=1}^n |z_j|^2 \leq \frac{1}{n}\sum_{j=1}^n |X_j^{(1)}|^2 + 1 \right\}$$

kompakt sind, ist die Einschränkung von F auf $[0,\hat{K}_n] \times \hat{M}$ sogar gleichmäßig stetig. Sind nun $z, h \in \mathbb{R}^n$ mit $z \in \hat{M}$ und $z+h \in \hat{M}$. so gilt, da $[0,\hat{K}_n]$ kompakt ist,

$$\begin{aligned}\left|\sup_{u\in[0,\hat{K}_n]} F(u,z) - \sup_{u\in[0,\hat{K}_n]} F(u,z+h)\right| &\leq \sup_{u\in[0,\hat{K}_n]} |F(u,z) - F(u,z+h)| \\ &= |F(u^*,z) - F(u^*,z+h)|\end{aligned}$$

für ein geeignetes $u^* \in [0,\hat{K}_n]$, welches nur von z und h abhängt. Da F auf $[0,\hat{K}_n] \times \hat{M}$ gleichmäßig stetig ist, folgt

$$\begin{aligned}&\lim_{h\to 0}\left|\sup_{u\in[0,\hat{K}_n]} F(u,z) - \sup_{u\in[0,\hat{K}_n]} F(u,z+h)\right| \\ &\quad\leq \lim_{h\to 0} |F(u^*,z) - F(u^*,z+h)| = 0.\end{aligned}$$

Also ist die Funktion $z \mapsto T_n(z)$ mit

$$\begin{aligned}T_n(z) = \sup_{u\in[0,\hat{K}_n]} &\Big[\left|\hat{\varphi}_{X_1^n}(u,0,\dots) - \hat{\varphi}_z(u)\cdot\hat{\varphi}_{\hat{\varepsilon}_1^n}(u)\right| \\ &+ \left|\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(u,0,\dots) - \hat{a}_2\frac{\partial}{\partial u}\hat{\varphi}_z(u)\cdot\hat{\varphi}_{\hat{\varepsilon}_1^n}(u)\right|\Big].\end{aligned}$$

auf $\hat{M}$ stetig und nimmt auf $\hat{M}$ auch ihr Minimum an, da $\hat{M}$ kompakt ist.

Definition 2.5 (Definition des Verteilungsschätzers)**.** Sind Δ_n und u_n reelle Folgen, definieren wir den Verteilungsschätzer $\hat{\mu}_n$ als empirisches Maß bezüglich der geschätzten Daten aus Definition 2.4. Wir schätzen die Verteilung von Z also durch

$$\hat{\mu}_n(B) := \frac{1}{n}\sum_{j=1}^{n} \mathbb{1}_B(\hat{z}_j). \quad (B \in \mathcal{B}_1)$$

2.3 Konsistenz des Verteilungsschätzers

Nachdem wir im vorherigen Abschnitt den Verteilungsschätzer definiert haben, stellt sich die Frage, ob unser Schätzer die Verteilung der latenten Variable Z in geeignetem Sinne approximiert. Mindestanforderung an einen solchen Schätzer ist dabei die, bei ungestörten Daten aus dem starken Gesetz der großen Zahlen folgende, punktweise Konvergenz von $\hat{\mu_n}$ gegen die wahre Verteilung μ. Das folgende Theorem zeigt, dass unser Schätzer diese Anforderung bei geigneter Wahl der Parameter und unter schwachen zusätzlichen Momentannahmen an Z und ε erfüllt.

Theorem 2.6 (Punktweise Konvergenz von $\hat{\mu}_n$)**.** *Sei Δ_n eine reelle Folge mit $\frac{2\log(n)}{\sqrt{n}} \leq \Delta_n$ für alle n in $\mathbb{N}$ und $\lim_{n\to\infty}\Delta_n = 0$. Sei u_n eine reelle Folge mit $\lim_{n\to\infty} u_n = \infty$ und $u_n \leq n^\alpha$ für ein $\alpha \in \mathbb{R}$ mit $0 < \alpha \leq \frac{1}{20}$ und alle $n \in \mathbb{N}$. Seien u_n und Δ_n außerdem so, dass $u_n \leq \Delta_n^2 n^{\frac{1}{10}}$ für alle $n \in \mathbb{N}$ gilt. Seien X, Z und ε wie in Definition 2.3 und seien außerdem die 4-ten Momente von Z und $\varepsilon_1,\dots,\varepsilon_d$ endlich. Seien die geschätzten Daten $\hat{z}_1^n$ ausgehend von $X_1,\dots,X_n$ sowie Δ_n und u_n definiert wie in Definition 2.4 und sei $\hat{\mu}_n$ definiert wie in Definition 2.5. Ist dann $a < b \in \mathbb{R}$ mit $\mathbf{P}_Z(\{a\}) = \mathbf{P}_Z(\{b\}) = 0$, so folgt*

$$\hat{\mu}_n([a,b]) \to \mathbf{P}_Z([a,b]), \quad \text{f.s.}$$

Nehmen wir zudem an, dass Z eine beschänkte Dichte besitzt, so erhalten wir das folgene Resultat vom Glivenko-Cantelli-Typ, welches die gleichmäßige Konvergenz des Schätzers über alle Borelmengen sicherstellt.

Theorem 2.7 (Gleichmäßige Konvergenz von $\hat{\mu}_n$)**.** *Sei $\beta \in \mathbb{R}$ mit $0 \leq \beta < \frac{1}{2}$ und sei $n_0 \in \mathbb{N}$. Sei Δ_n eine reelle Folge mit $\frac{2\log(n)}{\sqrt{n}} \leq \Delta_n$ für alle n in $\mathbb{N}$ mit $n \geq n_0$ und $\lim_{n\to\infty} \Delta_n = 0$. Sei u_n eine reelle Folge mit $\lim_{n\to\infty} u_n = \infty$ und $u_n \leq n^\alpha$ für ein $\alpha \in \mathbb{R}$ und alle $n \in \mathbb{N}$ mit $n \geq n_0$. Seien u_n und Δ_n außerdem so, dass*

$$4u_n^{\frac{5}{2}+\beta} \frac{1}{\Delta_n^2} \leq \frac{\sqrt[4]{n}}{\log(n)} \tag{2.6}$$

für alle $n \in \mathbb{N}$ mit $n \geq n_0$ gilt. Seien X, Z und ε wie in Definition 2.3 und seien außerdem die 4-ten Momente von Z und $\varepsilon_1, \ldots, \varepsilon_d$ endlich und Z besitze eine beschränkte Dichte f. Seien die geschätzten Daten $\hat{z}_1^n$ und der Parameter $\hat{K}_n$ ausgehend von $X_1, \ldots, X_n$ sowie Δ_n und u_n definiert wie in Definition 2.4 und sei $\hat{\mu}_n$ definiert wie in Definition 2.5. Sei $s_n = \frac{1}{\sqrt{\lfloor \hat{K}_n \rfloor}}$. Dann gilt

$$\hat{K}_n^\beta \sup_{a,b\in[\frac{-\pi}{s_n}+1, \frac{\pi}{s_n}-1]} \left| \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{[a,b]}(\hat{z}_j) - \int \mathbb{1}_{[a,b]} \, \mathbf{P}_Z \right| \to 0 \quad \textit{f.s.}$$

Bemerkung. Die Voraussetzungen von Theorem 2.7 enthalten mehrere Bedingungen an die Folgen u_n und Δ_n. Es lohnt sich also kurz sicherzustellen, dass sich alle diese Bedingungen gleichzeitig erfüllen lassen, um zu zeigen, dass die Voraussetzungen des Theorems nicht wiedersprüchlich sind. Die Bedingung (2.6) ist für $\beta = \frac{4}{10}$ erfüllt, falls es ein $n_0 \in \mathbb{N}$ gibt, so dass

$$u_n^{\frac{29}{10}} \Delta_n^{-2} \leq n^{\frac{1}{5}}$$

für alle $n \geq n_0$ gilt. Dies ist beispielsweise für $u_n = n^{\frac{1}{29}}$ und $\Delta_n = n^{-\frac{1}{20}}$ der Fall. Diese Wahl von u_n und Δ_n erfüllt auch alle anderen Bedingungen an u_n und Δ_n aus den Voraussetzungen von Theorem 2.7.

2.4 Schätzung der Verteilungsfunktion

Ausgehend von unserer Schätzung der empirischen Verteilung können wir auch einen Schätzer für die Verteilungsfunktion F von Z herleiten.

Definition 2.8 (Schätzer der Verteilungsfunktion)**.** Sei $\hat{\mu}_n$ definiert wie in Definition 2.5 und sei (s_n) eine reelle Folge. Wir definieren den Schätzer $\hat{F}_n$ der Verteilungsfunktion in Abhängigkeit von s_n über

$$\hat{F}_n(t) := \begin{cases} 0 & \text{für } t \leq \frac{-\pi}{s_n} + 1 \\ \hat{\mu}_n([-\frac{\pi}{s_n} + 1, t]) & \text{für } t \in \left(\frac{-\pi}{s_n} + 1, \frac{\pi}{s_n} - 1\right) \\ 1 & \text{für } t \geq \frac{\pi}{s_n} - 1. \end{cases} \tag{2.7}$$

Das folgende Theorem zeigt, dass der in Definition 2.8 definierte Schätzer für jede Nullfolge s_n punktweise gegen die wahre Verteilungsfunktion konvergiert.

Theorem 2.9. *Sei Δ_n eine reelle Folge mit $\frac{2\log(n)}{\sqrt{n}} \leq \Delta_n$ für alle n in $\mathbb{N}$ und $\lim_{n\to\infty} \Delta_n = 0$. Sei u_n eine reelle Folge mit $\lim_{n\to\infty} u_n = \infty$ und $u_n \leq n^\alpha$ für ein $\alpha \in \mathbb{R}$ mit $0 < \alpha \leq \frac{1}{20}$ und alle $n \in \mathbb{N}$. Seien u_n und Δ_n außerdem so, dass $u_n \leq \Delta_n^2 n^{\frac{1}{10}}$ für alle $n \in \mathbb{N}$ gilt. Seien X, Z und ε wie in Definition 2.3 und seien außerdem die 4-ten Momente von Z und $\varepsilon_1, \ldots, \varepsilon_d$ endlich. Außerdem gelte $\mathbf{P}_Z(\{a\}) = 0$ für alle $a \in \mathbb{R}$. Seien die geschätzten Daten $\hat{z}_1^n$ ausgehend von $X_1, \ldots, X_n$ sowie Δ_n und u_n definiert wie in Definition 2.4 und sei $\hat{\mu}_n$ definiert wie in Definition 2.5. Sei s_n eine Nullfolge und sei $\hat{F}_n$ definiert wie in Definition 2.8. Dann gilt*

$$|F_n(t) - F(t)| \to 0 \quad \text{f.s.}$$

für alle $t \in \mathbb{R}$.

Unter etwas stärkeren Voraussetzungen erhalten wir ein Resultat vom Glivenko-Cantelli-Typ, d.h., wir erhalten die gleichmäßige Konvergenz von $\hat{F}_n$ gegen F.

Theorem 2.10 (Ein Glivenko-Cantelli Resultat). *Sei Δ_n eine reelle Folge mit $\frac{2\log(n)}{\sqrt{n}} \leq \Delta_n$ für alle n in $\mathbb{N}$ mit $n \geq n_0$ und $\lim_{n\to\infty} \Delta_n = 0$. Sei u_n eine reelle Folge mit $\lim_{n\to\infty} u_n = \infty$ und $u_n \leq n^\alpha$ für ein $\alpha \in \mathbb{R}$ und alle $n \in \mathbb{N}$ mit $n \geq n_0$. Seien u_n und Δ_n außerdem so, dass*

$$4u_n^{\frac{5}{2}} \frac{1}{\Delta_n^2} \leq \frac{\sqrt[4]{n}}{\log(n)}$$

für alle $n \in \mathbb{N}$ mit $n \geq n_0$ gilt. Seien X, Z und ε wie in Definition 2.3 und seien außerdem die 4-ten Momente von Z und $\varepsilon_1, \ldots, \varepsilon_d$ endlich und besitze Z eine beschränkte Dichte. Seien die geschätzten Daten $\hat{z}_1^n$ und der Parameter $\hat{K}_n$ ausgehend von $X_1, \ldots, X_n$ sowie Δ_n und u_n definiert wie in Definition 2.4 und sei $\hat{\mu}_n$ definiert wie in Definition 2.5. Sei dann $\hat{F}_n$ definiert wie in Definition 2.8 mit $s_n = \frac{1}{\sqrt{\lfloor \hat{K}_n \rfloor}}$. Dann gilt

$$\sup_{t\in\mathbb{R}} |\hat{F}_n(t) - F(t)| \to 0 \quad \text{f.s.}$$

Bemerkung. Verstärkt man die Annahmen von Theorem 2.10 weiter, so ist es ebenfalls möglich eine Aussage zur Konvergenzgeschwindigkeit von der Form

$$\hat{K}_n^\beta \sup_{t\in\mathbb{R}} |\hat{F}_n(t) - F(t)| \to 0 \quad \text{f.s.}$$

herzuleiten.

2.5 Schätzung der Dichtefunktion

Um die gleichmäßige Konvergenz des Verteilungsschätzers und des Schätzers der Verteilungsfunktion zur beweisen, benötigten wir die Annahme, dass Z eine beschränkte Dichte f besitzt. Diese werden wir als Nächstes schätzen. Wir verwenden einen Kerndichteschätzer, den wir auf die geschätzten Daten $\hat{z}_j$ anwenden.

Definition 2.11. Sei h_n eine Folge von Bandbreiten mit $h_n \to 0$. Wir definieren $\hat{f}_n = \hat{f}_n(h_n)$ als

$$\hat{f}_n(t) := \frac{1}{2nh_n} \sum_{j=1}^{n} \mathbb{1}_{[-1,1]} \left(\frac{\hat{z}_j - t}{h_n} \right) = \frac{\hat{\mu}_n (B_{h_n}(t))}{\lambda (B_{h_n}(t))}, \tag{2.8}$$

wobei $B_{h_n}(t)$ die Kugel um t mit Radius h_n bezeichnet.

Wie wir bereits in der Einleitung gesehn haben, ist bei bei der Dichteschätzung der L^1-Fehler

$$\int_R |\hat{f}_n(t) - f(t)| \, dt$$

von besonderer Relevanz. Nach dem Lemma von Scheffé (siehe [3] oder [36, Theorem 1]) gilt nämlich

$$\sup_{B \in \mathcal{B}} \left| \int_B \hat{f}_n(t) \, dt - \int_B f(t) \, dt \right| = \frac{1}{2} \int_{\mathbb{R}} |\hat{f}_n(t) - f(t)| \, dt.$$

Ist unser Dichteschätzer also L_1-konsistent, so können wir

$$P[Z \in B] = \int_B f(t) \, dt$$

gleichmäßig über alle Borelmengen $B \in \mathcal{B}$ durch $\int_B \hat{f}_n(t)$ annähern. Im Gegensatz zu dem in Theorem 2.7 vorgestellten Schätzer lassen wir dabei zu, dass unsere Schätzung von $\mathbf{P}[Z \in B]$ auch von (geschätzten) Datenpunkten beeinflusst wird, welche nicht selbst in B liegen.

Theorem 2.12. *Sei $\beta \in \mathbb{R}$ mit $0 \leq \beta < \frac{1}{2}$ und sei $n_0 \in \mathbb{N}$. Sei Δ_n eine reelle Folge mit $\frac{2\log(n)}{\sqrt{n}} \leq \Delta_n$ für alle n in $\mathbb{N}$ mit $n \geq n_0$ und $\lim_{n\to\infty} \Delta_n = 0$. Sei u_n eine reelle Folge mit $\lim_{n\to\infty} u_n = \infty$ und $u_n \leq n^\alpha$ für ein $\alpha \in \mathbb{R}$ und alle $n \in \mathbb{N}$ mit $n \geq n_0$. Seien u_n und Δ_n außerdem so, dass*

$$4u_n^{\frac{5}{2}+\beta} \frac{1}{\Delta_n^2} \leq \frac{\sqrt[4]{n}}{\log(n)} \tag{2.9}$$

für alle $n \in \mathbb{N}$ mit $n \geq n_0$ gilt. Seien X, Z und ε wie in Definition 2.3 und seien außerdem die 4-ten Momente von Z und $\varepsilon_1, \ldots, \varepsilon_d$ endlich und Z besitze eine beschränkte Dichte f. Seien die geschätzten Daten $\hat{z}_1^n$ und der Parameter $\hat{K}_n$ ausgehend von $X_1, \ldots, X_n$ sowie Δ_n und u_n definiert wie in Definition 2.4 und sei $\hat{\mu}_n$ definiert wie in Definition 2.5. Sei h_n eine Folge von Bandbreiten mit $h_n \geq \hat{K}_n^{-\beta}$ und sei $\hat{f}_n$ definiert wie in (2.8). Dann gilt

$$\int |f(t) - \hat{f}_n(t)| \, dt \to 0 \quad \textit{f.s.}$$

Um für den Dichteschätzer eine Aussage zur Konvergenzgeschwindigkeit zu erhalten, benötigen wir weitere Annahman an f. Im folgenden benötigen wir die Hölderstetigkeit von f oder die Hölderstetigkeit der Ableitung von f. Eine Vereinheitlichung dieser Eigenschaften liefert der Begriff der (p, C)-Glattheit, den wir im Folgenden definieren.

Definition 2.13. Sei $0 < p \leq 1$. Eine Funktion $f : \mathbb{R} \to \mathbb{R}$ heißt (p, C)-glatt, falls für alle $x \in \mathbb{R}$ und alle $h \in \mathbb{R}$ die Abschätzung

$$|f(x) - f(x+h)| \leq C|h|^p$$

gilt.

Ist $1 < p = k + \delta$ für ein $k \in \mathbb{N}$ und ein $\delta \in (0, 1)$, so heißt $f : \mathbb{R} \to \mathbb{R}$ (p, C)-glatt, falls f k-mal stetig differenzierbar ist und $f^{(k)}$ eine (δ, C)-glatte Funktion ist.

Ist f nun (p, C)-glatt, so erhalten wir durch Wahl einer geeigneten Bandbreite h_n eine Aussage über die Konvergenzgeschwindigkeit des L^1-Fehlers.

Theorem 2.14. *Sei $\beta \in \mathbb{R}$ mit $0 \leq \beta < \frac{1}{2}$ und sei $n_0 \in \mathbb{N}$. Sei Δ_n eine reelle Folge mit $\frac{2\log(n)}{\sqrt{n}} \leq \Delta_n$ für alle n in $\mathbb{N}$ mit $n \geq n_0$ und $\lim_{n\to\infty} \Delta_n = 0$. Sei u_n eine reelle Folge mit $\lim_{n\to\infty} u_n = \infty$ und $u_n \leq n^\alpha$ für ein $\alpha \in \mathbb{R}$ und alle $n \in \mathbb{N}$ mit $n \geq n_0$. Seien u_n und Δ_n außerdem so, dass*

$$4u_n^{\frac{5}{2}+\beta} \frac{1}{\Delta_n^2} \leq \frac{\sqrt[4]{n}}{\log(n)} \tag{2.10}$$

für alle $n \in \mathbb{N}$ mit $n \geq n_0$ gilt. Seien X, Z und ε wie in Definition 2.3 und seien außerdem die 4-ten Momente von Z und $\varepsilon_1, \ldots, \varepsilon_d$ endlich und Z besitze eine beschränkte Dichte f. Sei f außerdem (p, C)-glatt für ein $p \leq 2$. Seien die geschätzten Daten $\hat{z}_1^n$ und der Parameter $\hat{K}_n$ ausgehend von $X_1, \ldots, X_n$ sowie Δ_n und u_n definiert wie in Definition 2.4 und sei $\hat{\mu}_n$ definiert wie in Definition 2.5. Sei $\gamma < \frac{p}{1+p}\beta$ und sei $h_n = \hat{K}_n^{\gamma-\beta}$. Sei $\hat{f}_n$ definiert wie in (2.8). Dann gilt

$$\hat{K}_n^\gamma \int |f(t) - \hat{f}_n(t)|\, dt \to 0 \quad \textit{f.s.}$$

2.6 Deterministische untere Schranke der Konvergenzgeschwindigkeit

In den vorausgehenden Abschnitten wurden die Resultate zur Konvergenzgeschwindigkeit der vorliegenden Schätzer unter Zuhilfenahme der Folge $\hat{K}_n$ aus Definition 2.4 formuliert. Da die Voraussetzungen von Lemma 3.8 unter den Annahmen der obigen Theoreme erfüllt sind, konvergiert $\hat{K}_n$ fast sicher gegen ∞. Da $\hat{K}_n$ von den Daten $X_1, \ldots, X_n$ abhängt, kann diese Konvergenz zumindest a-priori beliebig langsam stattfinden. Unter Annahmen an das Abfallverhalten der charakteristischen Funktionen von Z und ε ist es aber möglich eine deterministische untere Schranke für $\hat{K}_n$ herzuleiten und damit auch eine deterministische Konvergenzrate in den vorhergehenden Theoremen zu erhalten.

Proposition 2.15. *Sei Δ_n eine Nullfolge, so dass*

$$\frac{\sqrt{n}}{\log(n)}\Delta_n \to \infty$$

für $n \to \infty$ gilt. Sei X eine reelle Zufallsvariable mit $\mathbf{E}\{|X|\} < \infty$ und seien $X, X_1, \ldots, X_n$ unabhängig und identisch verteilt. Sei

$$\hat{x}_n := \sup\left\{x \in \mathbb{R}^+ : \inf_{u\in[0,x]} |\hat{\varphi}_{X_1^n}(u, 0, \ldots, 0)| \geq \Delta_n\right\}$$

und sei A_n eine reelle Folge, so dass ein $n_0 \in \mathbb{N}$ existiert für welche

$$\inf_{u\in[0,A_n]} |\varphi_{X^{(1)}}(u)| \geq (1+\varepsilon)\Delta_n$$

für ein $\varepsilon > 0$ und alle $n \geq n_0$ gilt, und so, dass für ein $k \in \mathbb{N}$ die Folge $\frac{A_n}{n^k}$ beschänkt ist. Dann gilt mit Wahrscheinlichkeit 1

$$\hat{x}_n \geq A_n$$

für n genügend groß.

Bemerkung. In Proposition 2.15 haben wir vorausgesetzt, dass der Betrag der charakteristischen Funktion $\varphi_{X^{(1)}}(t)$ von $X^{(1)}$ für $|t| \to \infty$ hinreichend langsam abfällt. Das folgende Theorem von Pólya zeigt, dass es in der Tat charakteristische Funktionen von reellen Zufallsvariablen gibt, so dass $|\varphi_{X^{(1)}}(t)|$ für $t \to \infty$ beliebig langsam gegen 0 konvergiert. Wir zitieren das Theorem an dieser Stelle nur und verweisen für den Beweis in die Literatur (beispielsweise [37] oder [38, Satz 15.24]).

Theorem (Pólya)**.** *Sei $f : \mathbb{R} \to \mathbb{R}$ mit den folgenden Eigenschaften:*

1. *f ist stetig.*
2. *Es gilt $f(0) = 1$.*
3. *Es gilt $\lim_{t\to\infty} f(t) = 0$.*
4. *f ist gerade.*
5. *Die Einschränkung von f auf $[0, \infty)$ ist konvex.*

Dann ist f die charakteristische Funktion einer reellen Zufallsvariable X und X besitzt eine Dichte, welche auf $\mathbb{R} \setminus \{0\}$ stetig ist.

Wir beenden diesen Abschnitt mit dem folgendem Theorem, in welchem die Ergebnisse aus Theorem 2.14 und Proposition 2.15 anwenden um für den Fall, dass die Zufallsvariablen Z und ε_1 jeweils Normalverteilt mit unbekannter Varianz und Erwartungswert 0 sind, eine deterministische Konvergenzgeschwindigkeit für unseren Dichteschätzer herzuleiten.

Theorem 2.16. *Sei $\Delta_n = n^{-\frac{1}{20}}$ udn sei $u_n = n^{\frac{1}{29}}$. Seien X, Z und ε wie in Definition 2.3. Sei außerdem Z Normalverteilt mit Erwartungswert 0 und Varianz $\sigma_1^2 > 0$ und sei ε_1 Normalverteilt mit Erwartungswert 0 und Varianz $\sigma_2^2 > 0$. Zudem seien die 4-ten Momente von $\varepsilon_2, \dots, \varepsilon_d$ endlich und die charakteristischen Funktionen von $\varepsilon_2, \dots, \varepsilon_d$ seien nullstellenfrei. Seien die geschätzten Daten $\hat{z}_1^n$ und der Parameter $\hat{K}_n$ ausgehend von $X_1, \dots, X_n$ sowie Δ_n und u_n definiert wie in Definition 2.4 und sei $\hat{\mu}_n$ definiert wie in Definition 2.5. Sei $\hat{h}_n = \hat{K}_n^{\frac{8}{31} - \frac{4}{10}}$. Ist dann $\hat{f}_n$ ausgehend von $\hat{\mu}_n$ und $\hat{h}_n$ definiert wie in (2.8), so gilt mit Wahrscheinlichkeit 1*

$$\Big(\frac{1}{\sigma_1^2 + \sigma_2^2} \log\big(\frac{n^{\frac{1}{10}}}{4}\big)\Big)^{\frac{8}{62}} \int |f(t) - \hat{f}_n(t)|\, dt \to 0$$

wobei f die Dichte von Z bezeichnet.

Bemerkung. Bevor wir im nächsten Kapitel mit den Beweisen der hier vorgestellten Resultate beginnen, noch zwei Anmerkungen zu Theorem 2.16. Wie wir im späteren Beweis sehen werden, hängt die Konvergenzrate maßgeblich vom Abfallverhalten der charakteristischen Funktion von $X^{(1)} = Z + \varepsilon_1$ ab. Während es also durchaus möglich erscheint, die in der Rate auftretenden Konstanten und Exponenten durch geschickte Wahl von u_n und Δ_n zu verbessern, können die in diesem Kapitel vorgestellten Methoden für normalverteilte Zufallsvariablen nur Raten der Form $C(\log(n))^\alpha$ liefern. Durch die Beziehung

$$\varphi_{X^{(1)}} = \varphi_Z \cdot \varphi_{\varepsilon_1}$$

wird anderseits nahegelegt, dass die Konvergenzgeschwindigkeit des Schätzers im wesentlichen vom dem asymptotisch dominanten Faktor auf der rechten Seite bestimmt wird. Selbst im Falle $\varepsilon_1 = 0$, was $\varphi_{\varepsilon_1} = 1$ impliziert, lässt sich also für normalverteiltes Z keine substanzielle Verbesserung der untere Schranke für die Konvergenzgeschwindigkeit herleiten. Damit kommen wir zum zweiten Teil dieser abschließenden Bemerkung: Wie gut ist die von dem hier vorgestellten Schätzer ereichte Rate im Vergleich zu anderen Schätzverfahren? Eine genaue Einordnung des Resultats erweist sich als schwierig, da dem Autor, wie in der Einleitung bereits erwähnt, in dem hier vorgestellten Modell keine anderen nichtparametrischen Dichteschätzer bekannt sind. Wir vergleichen daher mit der (optimalen) Rate aus der Dekonvolutions-Theorie, welche sich ebenfalls mit der nichtparametrischen Schätzung der Dichte f von Z ausgehend von Daten X mit einem unabhängigen additiven Fehler, also

$$X = Z + \varepsilon$$

beschäftigt. Wir werden in dieser Bemerkung nicht im Detail auf die benötigten Konzepte und Definitionen eingehen, sondern nur das für uns entscheidende Resultat grob skizzieren. Die entsprechenden Ausführungen und Beweise finden sich beispielsweise in [15].

Auch im Dekonvolutions-Modell wird zur konstruktion des Schätzers die Annahme benötigt, dass die charakteristische Funktion von ε nullstellenfrei ist. Zusätzlich wird allerdings davon ausgegangen, dass die Dichte, und damit auch die charakteristische Funktion, des Fehlers ε bekannt ist. Ist dann

$$\mathcal{F}_{\beta,C,L_2} = \left\{ f \in L_2 : \int |Ft(f)(t)|^2 |t|^{2\beta} \, dt \leq C \right\},$$

wobei $Ft(f)$ die Fouriertransformierte von f bezeichnet, so gilt für Z Normalverteilt auch $f_Z \in \mathcal{F}_{\beta,C,L_2}$ für geeignetes $\beta > 0$ und $C > 0$. Ist nun auch der Fehler ε normalverteilt, so gilt [15, Theorem 2.9]

$$\sup_{f \in \mathcal{F}_{\beta,C,L_2}} \mathbf{E} \int |\hat{f} - f|^2 \, dt = O\left((\log(n))^{-\beta} \right)$$

hierbei bezeichnet $\hat{f}$ den Dekonvolutions-Kerndichteschätzer bezüglich einem geeigneten Kern. Obwohl in der vorliegenden Arbeit ein anderes Modell als auch ein anderes Fehlerkriterium für den Dichteschätzer betrachtet haben, erreichen wir zumindest im Fall von normalverteilten Zufallsvariablen eine Konvergenzgeschwindigkeit in der gleichen Größenordnung wie der Dekonvolutions-Kerndichteschätzer.

3 Beweise zu Kapitel 2

3.1 Beweis der Konvergenz der geschätzten Daten

Beweis von Lemma 2.2. Bei diesem Lemma handelt es sich um eine Abwandlung des analogen Lemmas 4.1 für das multivariate Modell aus Kapitel 4, welches sich in [34] findet. Wir präsentieren hier der Vollständigkeit halber eine Variation des Beweises aus [34]. Für den Beweis benötigen wir zwei Hilfsresultate, die wir auch zur Definition unseres Schätzers und im Beweis von Lemma 3.3 noch brauchen werden. Die erste Hilfsaussage beschäftigt sich mit den Koeffizienten. Gilt sowohl (2.2) als auch

$$X = \begin{pmatrix} X^{(1)} \\ X^{(2)} \\ \vdots \\ X^{(d)} \end{pmatrix} = \begin{pmatrix} \tilde{a}_1 \tilde{Z} \\ \tilde{a}_2 \tilde{Z} \\ \vdots \\ \tilde{a}_d \tilde{Z} \end{pmatrix} + \begin{pmatrix} \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \\ \vdots \\ \tilde{\varepsilon}_d \end{pmatrix}$$

und erfüllen $\tilde{Z}$, $\tilde{a}_1, \ldots, \tilde{a}_d$ und $\tilde{\varepsilon}_1, \ldots, \tilde{\varepsilon}_d$ die Bedingungen aus Definition 2.3, so folgt aus der Unabhängigkeit von $\varepsilon_1, \ldots, \varepsilon_d, Z$ und der Unabhängigkeit von $\tilde{\varepsilon}_1, \ldots, \tilde{\varepsilon}_d, \tilde{Z}$, das sowohl

$$\begin{aligned}
&\mathbf{E}\{X^{(j)}X^{(i)}\} = \mathbf{E}\{(a_j Z + \varepsilon_j)(a_i Z + \varepsilon_i)\} \\
&\quad = a_j a_i \mathbf{E}\{Z^2\} + \mathbf{E}\{a_j Z + \varepsilon_j\}\mathbf{E}\{\varepsilon_i\} + \mathbf{E}\{a_i Z + \varepsilon_i\}\mathbf{E}\{\varepsilon_j\} + \mathbf{E}\{\varepsilon_j\}\mathbf{E}\{\varepsilon_i\} \\
&\quad = a_j a_i \mathbf{E}\{Z^2\}
\end{aligned}$$

als auch, völlig analog,

$$\mathbf{E}\{X^{(j)}X^{(i)}\} = \tilde{a}_j \tilde{a}_i \mathbf{E}\{\tilde{Z}^2\}$$

für $i, j \in \{1, \ldots, d\}$ mit $i \neq j$ gilt. Nach Voraussetzung gilt $\mathbf{E}\{\tilde{Z}^2\} \neq 0$ und $\mathbf{E}\{Z^2\} \neq 0$ sowie $a_1 = \tilde{a}_1 = 1$, $a_2 \neq 0$, $\tilde{a}_2 \neq 0$, $a_3 \neq 0$ und $\tilde{a}_3 \neq 0$. Also gilt

$$a_2 = \frac{a_2 a_3 \mathbf{E}\{Z^2\}}{a_1 a_3 \mathbf{E}\{Z^2\}} = \frac{\mathbf{E}\{X^{(2)}X^{(3)}\}}{\mathbf{E}\{X^{(1)}X^{(3)}\}} = \frac{\tilde{a}_2 \tilde{a}_3 \mathbf{E}\{\tilde{Z}^2\}}{\tilde{a}_1 \tilde{a}_3 \mathbf{E}\{\tilde{Z}^2\}} = \tilde{a}_2$$

und, analog,

$$a_j = \frac{\mathbf{E}\{X^{(2)}X^{(j)}\}}{\mathbf{E}\{X^{(1)}X^{(2)}\}} = \tilde{a}_j$$

für $j \in \{3, \ldots, d\}$. Also stimmen die Koeffizienten überein. Die zweite Hilfsaussage liefert einen Zusammenhang zwischen den charakteristischen Funktionen von X, Z und ε. Es gilt

$$\begin{aligned}
\varphi_X(u_1, \ldots, u_d) &= \mathbf{E}\left\{\exp\left(i\sum_{j=1}^d u_j X^{(j)}\right)\right\} \\
&= \mathbf{E}\left\{\exp\left(i\sum_{j=1}^d u_j (a_j Z + \varepsilon_j)\right)\right\} \\
&= \mathbf{E}\left\{\exp\left(i\sum_{j=1}^d u_j a_j Z\right) \cdot \exp\left(i\sum_{j=1}^d u_j \varepsilon_j\right)\right\}
\end{aligned}$$

$$= \varphi_Z \left(\sum_{j=1}^{d} a_j u_j \right) \cdot \prod_{j=1}^{d} \varphi_{\varepsilon_j}(u_j),$$

da $Z, \varepsilon_1, \ldots, \varepsilon_d$ nach Voraussetzung voneinander unabhängig sind. Zusammen mit $\varphi_{\varepsilon_j}(0) = 1$ und $\frac{\partial}{\partial u}\varphi_{\varepsilon_j}(0) = i \cdot \mathbf{E}\{\varepsilon_j\} = 0$ folgt

$$\varphi_X(u_1, 0, \ldots, 0) = \varphi_Z(u_1)\varphi_{\varepsilon_1}(u_1). \tag{3.1}$$

Da die auftretenden charakteristischen Funktionen wegen $\mathbf{E}\{Z\} < \infty$ und $\mathbf{E}\{\varepsilon_j\} < \infty$ differenzierbar sind, gilt auch

$$\begin{aligned}
\frac{\partial}{\partial u_2}\varphi_X(u_1, \ldots, u_d) &= \frac{\partial}{\partial u_2}\Big(\varphi_Z \left(\sum_{j=1}^{d} a_j u_j \right) \cdot \prod_{j=1}^{d} \varphi_{\varepsilon_j}(u_j) \Big) \\
&= \frac{\partial}{\partial u_2}\varphi_Z \left(\sum_{j=1}^{d} a_j u_j \right) \cdot \prod_{j=1}^{d} \varphi_{\varepsilon_j}(u_j) + \varphi_Z \left(\sum_{j=1}^{d} a_j u_j \right) \cdot \frac{\partial}{\partial u_2} \prod_{j=1}^{d} \varphi_{\varepsilon_j}(u_j) \\
&= a_2 \frac{\partial}{\partial u}\varphi_Z \left(\sum_{j=1}^{d} a_j u_j \right) \cdot \prod_{j=1}^{d} \varphi_{\varepsilon_j}(u_j) \\
&\quad + \varphi_Z \left(\sum_{j=1}^{d} a_j u_j \right) \cdot \frac{\partial}{\partial u}\varphi_{\varepsilon_j}(u_j) \prod_{j=1, j\neq 2}^{d} \varphi_{\varepsilon_j}(u_j).
\end{aligned}$$

Zusammen mit $\varphi_{\varepsilon_j}(0) = 1$ und $\frac{\partial}{\partial u}\varphi_{\varepsilon_j}(0) = 0$ folgt

$$\frac{\partial}{\partial u_2}\varphi_X(u_1, 0, \ldots, 0) = a_2 \frac{\partial}{\partial u_1}\varphi_Z(u_1)\varphi_{\varepsilon_1}(u_1). \tag{3.2}$$

Da die auftretenden charakteristischen Funktionen differenzierbar sind und keine Nullstellen besitzen, ist die Funktion L_{φ_Z} gegeben durch

$$L_{\varphi_Z}(v) := \log(\varphi_Z(0)) + \int_0^v \frac{\frac{\partial}{\partial u_1}\varphi_Z(u)}{\varphi_Z(u)}\, du = \int_0^v \frac{\frac{\partial}{\partial u_1}\varphi_Z(u)}{\varphi_Z(u)}\, du$$

wohldefiniert, und es gilt $\exp(L_{\varphi_Z}(v)) = \varphi_Z(v)$. (siehe beispielsweise [39, Seite 123ff.]) Damit folgt aber, zusammen mit $a_2 = \tilde{a}_2$ und (3.1) sowie (3.2),

$$\begin{aligned}
\varphi_Z(v) &= \exp\left(\int_0^v \frac{\frac{\partial}{\partial u_1}\varphi_Z(u)}{\varphi_Z(u)}\, du \right) \\
&= \exp\left(\int_0^v \frac{1}{a_2} \frac{a_2 \frac{\partial}{\partial u_1}\varphi_Z(u) \cdot \varphi_{\varepsilon_1}(u_1)}{\varphi_Z(u) \cdot \varphi_{\varepsilon_1}(u_1)}\, du \right) \\
&= \exp\left(\int_0^v \frac{1}{a_2} \frac{\frac{\partial}{\partial u_2}\varphi_X(u, 0, \ldots, 0)}{\varphi_X(u, 0, \ldots, 0)}\, du \right) \\
&= \exp\left(\int_0^v \frac{1}{\tilde{a}_2} \frac{\frac{\partial}{\partial u_2}\varphi_X(u, 0, \ldots, 0)}{\varphi_X(u, 0, \ldots, 0)}\, du \right)
\end{aligned}$$

$$= \varphi_{\tilde{Z}}(v).$$

Nach dem Eindeutigkeitssatz für charakteristische Funktionen (siehe [40, Theorem 23.4]) stimmt die Verteilung von Z also mit der Verteilung von $\tilde{Z}$ überein. Für $j \in \{1, \dots, d\}$ gilt, analog zu (3.1),

$$\varphi_{\tilde{Z}}(\tilde{a}_j u_j)\varphi_{\tilde{\varepsilon}_j}(u_j) = \varphi_X(0, \dots, u_j, \dots, 0) = \varphi_Z(\tilde{a}_j u_j)\varphi_{\varepsilon_j}(u_j).$$

Zusammen mit $\varphi_{\tilde{Z}} = \varphi_Z$ sowie $\varphi_Z(u) \neq 0$ und $\varphi_{\tilde{Z}}(u) \neq 0$ für alle $u \in \mathbb{R}$ folgt daraus, dass auch die Verteilungen der ε_j mit den Verteilungen der $\tilde{\varepsilon}_j$ identisch sind. Damit ist der Beweis beendet. □

Bevor wir die Theoreme aus Kapitel 2 über die Konsistenz des Schätzers für die Verteilungsfunktion und die Konsistenz des Dichteschätzers beweisen können, benötigen wir einige Lemmata und Theoreme zur Konvergenz der empirischen charakteristischen Funktionen und der Konvergenz des Verteilungsschätzers $\hat{\mu}_n$ aus Definition 2.5. Die Verbindung zwischen diesen beiden Themen liefert das folgende Lemma, welches erlaubt, die Konvergenz der empirischen charakteristischen Funktion der geschätzten Daten auf Konvergenz des empirischen Mittels von trigonometrischen Polynomen zu übertragen. Ein wichtiges Werkzeug für die späteren Konsistenzbeweise ist die Approximation der zu schätzenden Funktion durch trigonometrische Polynome. Dabei ist darauf zu achten, dass Grad und Frequenz der approximierenden Funktionen an die geschätzten Daten angepasst sind.

Definition 3.1. Seien α_n und s_n reelle Folgen mit $\alpha_n \geq 0$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} \alpha_n = \infty$ sowie mit $s_n \geq 0$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} s_n = 0$. Außerdem sei $d_n \to \infty$ eine Folge von ganzen Zahlen. Dann ist $\mathcal{G}_n = \mathcal{G}_n(\alpha_n, d_n, s_n)$ definiert als die Menge der trigonometrischen Polynome

$$g(u) = a_0 + \sum_{j=1}^{d_n} a_j \cos(j \cdot s_n \cdot u) + \sum_{j=1}^{d_n} b_j \sin(j \cdot s_n \cdot u)$$

mit $|a_0|, |a_j|, |b_j| \leq \alpha_n$.

Lemma 3.2. *Seien r_n und I_n reelle Folgen, so dass mit Wahrscheinlichkeit 1*

$$r_n \sup_{u\in[0,I_n]} |\hat{\varphi}_{\hat{z}_1^n}(u) - \varphi_Z(u)| \to 0$$

für $n \to \infty$ gilt. Seien α_n und s_n reelle Folgen mit $\alpha_n \geq 0$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} \alpha_n = \infty$ sowie mit $s_n \geq 0$ für alle $n \in \mathbb{N}$ und $\lim_{n\to\infty} s_n = 0$. Außerdem sei $d_n \to \infty$ eine Folge von ganzen Zahlen, so dass $s_n d_n \leq I_n$ für alle $n \in \mathbb{N}$ gilt. Sei $\mathcal{G}_n = \mathcal{G}_n(\alpha_n, d_n, s_n)$, definiert wie in 3.1, die Menge der trigonometrischen Polynome

$$g(u) = a_0 + \sum_{j=1}^{d_n} a_j \cos(j \cdot s_n \cdot u) + \sum_{j=1}^{d_n} b_j \sin(j \cdot s_n \cdot u)$$

mit $|a_0| \leq \alpha_n$, $|a_j| \leq \alpha_n$ und $|b_j| \leq \alpha_n$. Ist dann r'_n eine reelle Folge, so dass $2d_n\alpha_n r'_n \leq r_n$ für alle $n \in \mathbb{N}$ gilt, so gilt auch

$$r'_n \sup_{g\in\mathcal{G}_n} \left| \frac{1}{n} \sum_{i=1}^{n} g(\hat{z}_i) - \mathbf{E}\{g(Z)\} \right| \to 0$$

fast sicher.

Bemerkung. Die Aussage von Lemma 3.2 gilt auch, wenn wir die deterministischen Folgen r_n, s_n, d_n und α_n durch datenabhängige $\hat{r}_n$, $\hat{s}_n$, $\hat{d}_n$ und $\hat{\alpha}_n$ ersetzen, die mit Wahrscheinlichkeit 1 das gewünschte Konvergenzverhalten besitzen.

Beweis von Lemma 3.2. Sei $g \in \mathcal{G}_n$ beliebig, dann folgt

$$\begin{aligned}
&\Big|\frac{1}{n}\sum_{i=1}^{n} g(\hat{z}_i) - \mathbf{E}\{g(Z)\}\Big| = \\
&\Big|\frac{1}{n}\sum_{i=1}^{n}\left(a_0 + \sum_{j=1}^{d_n}(a_j\cos(\hat{z}_i \cdot s_n \cdot j) + b_j \sin(\hat{z}_i \cdot s_n \cdot i))\right) \\
&\quad - \mathbf{E}\left\{a_0 + \sum_{j=1}^{d_n}(a_j\cos(Z \cdot s_n \cdot j) + b_j \sin(Z \cdot s_n \cdot i))\right\}\Big| \\
&\le \sum_{j=1}^{d_n}|a_j|\left|\frac{1}{n}\sum_{i=1}^{n}\Big[\cos(\hat{z}_i \cdot s_n \cdot j) - \mathbf{E}\big\{\cos(Z \cdot s_n \cdot j)\big\}\Big]\right| \\
&\quad + |b_j|\left|\frac{1}{n}\sum_{i=1}^{n}\Big[\sin(\hat{z}_i \cdot s_n \cdot j) - \mathbf{E}\big\{\sin(Z \cdot s_n \cdot j)\big\}\Big]\right|.
\end{aligned}$$

Wegen $|a_j| \le \alpha_n$, $|b_j| \le \alpha_n$, $2d_n\alpha_n r'_n \le r_n$ und $s_n \cdot j \le s_n \cdot d_n \le I_n$ folgt

$$\begin{aligned}
r'_n\Big|\frac{1}{n}\sum_{i=1}^{n} g(z_i) - \mathbf{E}\{g(Z)\}\Big| \\
&\le 2d_n\alpha_n r'_n \sup_{u\in[0,I_n]}|\varphi_{\hat{z}_i^n}(u) - \varphi_Z(u)| \\
&\le r_n \sup_{u\in[0,I_n]}|\varphi_{\hat{z}_i^n}(u) - \varphi_Z(u)| \to 0 \quad \text{f.s.}
\end{aligned}$$

Da $g \in \mathcal{G}_n$ beliebig war, impliziert dies die Behauptung. □

Um die Vorraussetzung von Lemma 3.2 zu verifizieren, benötigen wir einige Hilfsresultate über das Konvergenzverhalten empirischer charakteristischer Funktionen. Unser Ziel ist der Beweis von Lemma 3.3, welches angibt, unter welchen Umständen und in welchem Sinne die empirische charakteristische Funktion der wie in Definition 2.4 geschätzten Daten gegen die charakteristische Funktion der latenten Variable Z konvergiert. Dazu benötigen wir zunächst einige Lemmata über die Konvergenz von empirischen charakteristischen Funktionen.

Wir beginnen mit einer Vorüberlegung. Sind $X, X_1, X_2, \ldots$ unabhängig und identisch verteilt, so ist klar, dass die empirische charakteristische Funktion

$$\hat{\varphi}_1^n(X_1, \ldots, X_n, u) = \varphi_{X_1^n}(u) = \frac{1}{n}\sum_{j=1}^{n}\exp(i \cdot X_j \cdot u)$$

im Allgemeinen nicht gleichmäßig gegen die charakteristische Funktion

$$\varphi_X(u) = \mathbf{E}\left\{\exp(iXu)\right\}$$

konvergieren kann. Denn besitzt X eine stetige Dichte, so verschwindet φ nach dem Lemma von Riemann-Lebesgue (siehe [41, Satz 3.9]) im Unendlichen. Es gilt also

$$\lim_{|u|\to\infty} |\varphi_X(u)| \to 0.$$

Die Funktion $\hat{\varphi}_1^n$ hingegen ist periodisch, falls alle X_j rationale Vielfache voneinander sind. Denn ist $x \in \mathbb{R}$ und sind $q_1, \ldots, q_n, p_1, \ldots, p_n \in \mathbb{Z}$ mit $X_j = \frac{q_j}{p_j} x$, und ist

$$h = \frac{2\pi}{x} \prod_{j=1}^{n} p_j,$$

so gilt $\frac{q_j}{p_j} xh \in \{2k\pi : k \in \mathbb{Z}\}$ für jedes $j \in \{1, \ldots, n\}$ und es folgt

$$\hat{\varphi}_{X_1^n}(u) = \frac{1}{n} \sum_{j=1}^{n} \exp(i \cdot \frac{q_j}{p_j} x \cdot u) = \frac{1}{n} \sum_{j=1}^{n} \exp(i \cdot \frac{q_j}{p_j} x \cdot (u + h)).$$

Sind die X_j keine rationalen Vielfachen, so ist $\varphi_{X_1^n}$ zumindest „fast-periodisch“[1]. Das nächste Lemma zeigt, dass die erwünschte Konvergenz von $\hat{\varphi}_1^n$ gegen φ zumindest auf einer die reelle Halbachse ausschöpfenden Folge von Intervallen gleichmäßig ist.

Lemma 3.3. *Seien X, Z und ε Zufallsvariablen wie in Definition 2.3 und es gelte $\mathbf{E}\{Z^4\} < \infty$ und $\mathbf{E}\{\varepsilon_j^4\} < \infty$ für alle $j \in \{1, \ldots, n\}$. Sei u_n eine Folge mit $\lim_{n\to\infty} u_n = \infty$ und mit $u_n \leq n^\alpha$ für ein $\alpha > 0$ und alle $n \in \mathbb{N}$. Sei Δ_n eine Nullfolge mit $2\frac{\log(n)}{\sqrt{n}} \leq \Delta_n$ für alle $n \geq n_0$. Sei $\hat{K}_n$ definiert wie in Gleichung (2.3) und sei $\hat{z}_1^n$ ausgehend von u_n, Δ_n und $\hat{K}_n$ definiert wie in Definition 2.4. Dann gilt*

$$\frac{\Delta_n^2}{\hat{K}_n} \frac{\sqrt[4]{n}}{\log(n)} \sup_{u\in[0,\hat{K}_n]} |\hat{\varphi}_{\hat{z}_1^n}(u) - \varphi_Z(u)| \to 0 \quad \textit{f.s.}$$

Zum Beweis von Lemma 3.3 benötigen wir eine Reihe von Hilfsresultaten, die wir im Folgenden erarbeiten werden. Zunächst zeigen wir jedoch, dass sich durch geeignete Wahl von u_n und Δ_n in der Definition von $\hat{K}_n$ sicherstellen lässt, dass auch $\sup_{u\in[0,\hat{K}_n]} |\hat{\varphi}_{\hat{z}_1^n}(u) - \varphi_Z(u)|$ für $n \to \infty$ gegen 0 konvergiert. Obwohl die Rate in Lemma 3.3 den datenabhängigen Faktor $\hat{K}_n$ enthält, lässt sich also die Konvergenz durch geschickte, deterministische Wahl der Parameter u_n und Δ_n garantieren.

Korollar 3.4. *Seien X, Z und ε Zufallsvariablen wie in Definition 2.3 und es gelte $\mathbf{E}\{Z^4\} < \infty$ und $\mathbf{E}\{\varepsilon_j^4\} < \infty$ für alle $f \in \{1, \ldots, n\}$. Sei $\hat{z}_1^n$ definiert wie in Definition 2.4 und sei $\hat{K}_n$ definiert wie in Gleichung (2.3) mit $u_n \leq n^\alpha$ für ein $\alpha > 0$ und alle $n \in \mathbb{N}$ und $\frac{2\log(n)}{\sqrt{n}} \leq \Delta_n$ für n groß genug. Sei r_n eine Folge, so dass $r_n u_n \leq \Delta_n^2 n^{\frac{1}{5}}$ für n groß genug gilt. Dann gilt*

$$r_n \sup_{u\in[0,\hat{K}_n]} |\hat{\varphi}_{\hat{z}_1^n}(u) - \varphi_Z(u)| \to 0$$

fast sicher.

[1] Fast-periodische Funktionen wurden in der Literatur bereits ausgiebieg untersucht, hauptsächlich im Zusammenhang mit partiellen Diffenrentialgleichungen (siehe [42], [43] oder [44]). Die exakte Definition des Begriffes und die den Begriff umgebende Theorie sind für die vorliegende Arbeit allerdings von geringer Relevanz. Es sei jedoch ergänzend erwähnt, dass fast-periodische Funktionen auch als charakteristische Funktionen von diskreten Verteilungen eine Rolle spielen (siehe beispielsweise [45, Kapitel 2.2]).

Beweis von Korollar 3.4. Da es ein $n_0 \in \mathbb{N}$ gibt, so dass

$$2\frac{\log(n)}{\sqrt{n}} \leq \Delta_n$$

für alle $n \geq n_0$ gilt, und da $u_n \leq n^\alpha$ für ein $\alpha > 0$ gilt, sind die Vorraussetzungen von Lemma 3.3 erfüllt. Außerdem gilt

$$r_n \frac{\hat{K}_n}{\Delta_n^2} \leq r_n \frac{u_n}{\Delta_n^2} \leq n^{\frac{1}{5}} < \frac{\sqrt[4]{n}}{\log n}$$

für n groß genug. Für solche n folgt

$$r_n \leq \frac{\Delta_n^2}{\hat{K}_n} \frac{\sqrt[4]{n}}{\log(n)}.$$

Unter Verwendung von Lemma 3.3 erhalten wir also

$$r_n \sup_{u \in [0, \hat{K}_n]} |\hat{\varphi}_{\hat{z}_1^n}(u) - \varphi_Z(u)| \leq \frac{\Delta_n^2}{\hat{K}_n} \frac{\sqrt[4]{n}}{\log(n)} \sup_{u \in [0, \hat{K}_n]} |\hat{\varphi}_{\hat{z}_1^n}(u) - \varphi_Z(u)| \to 0 \quad \text{f.s.}$$

Damit ist der Beweis beendet. $\square$

Bemerkung. Es ist möglich r_n, u_n und Δ_n so zu wählen, dass die Bedingungen von Korollar 3.4 erfüllt sind. Beispielsweise durch $r_n = 1$, $u_n = n^{\frac{1}{10}}$ und $\Delta_n = \frac{1}{n^{\frac{1}{20}}}$.

Nun beginnen wir mit der Vorarbeit für den Beweis von Lemma 3.3. Als Erstes beweisen wir eine Variante von Lemma 3.3, bei dem statt der geschätzten Daten unabhängig und identisch verteilte Kopien $X_1, \dots, X_n$ einer Zufallsvariablen X vorliegen.

Lemma 3.5. *Sei X mit $\mathbf{E}\{|X|\} < \infty$ und seien $X, X_1, \dots$ unabhängig und identisch verteilte Zuvallsvariablen. Sind r_n und I_n Folgen in $\mathbb{R}$, so dass die Folgen $r_n \frac{\log(n)}{\sqrt{n}}$ und $\frac{I_n}{(n^k)}$ beschränkt sind, so folgt*

$$r_n \sup_{u \in [0, I_n]} |\varphi_X(u) - \hat{\varphi}_{X_1^n}(u)| \to 0 \quad \textit{f.s.} \tag{3.3}$$

Beweis von Lemma 3.5. Sei $M_n = \sqrt{n}$. Dann gibt es ein $C > 0$, so dass $\frac{r_n}{M_n} \leq \frac{C}{\log(n)} \to 0$ gilt. Außerdem gilt $\frac{M_n}{n} \to 0$. Sei $\operatorname{sign} : \mathbb{R} \to \{-1, 1\}$ mit$\operatorname{sign}(x) = 1$ für $x \geq 0$ und $\operatorname{sign}(x) = -1$ sonst. Wir setzen

$$X_{M_n} = \begin{cases} X, & \text{falls } |X| \leq M_n \\ \operatorname{sign}(X) M_n & \text{sonst} \end{cases}$$

und definieren $X_{M_n}{}_1^n = (X_{1,M_n}, \dots, X_{n,M_n}$. Dann folgt

$$\begin{aligned} T_n &:= \sup_{u \in [0, I_n]} |\varphi_X(u) - \hat{\varphi}_{X_1^n}(u)| \\ &\leq \sup_{u \in [0, I_n]} |\varphi_X(u) - \varphi_{X_{M_n}}(u)| + \sup_{u \in [0, I_n]} |\varphi_{X_{M_n}}(u) - \hat{\varphi}_{X_{M_n}{}_1^n}(u)| \end{aligned}$$

$$
\begin{aligned}
&+ \sup_{u\in[0,I_n]} |\hat{\varphi}_{X_{M_n}{}_1^n}(u) - \hat{\varphi}_{X_1^n}(u)| \\
&= T_{1,n} + T_{2,n} + T_{3,n}.
\end{aligned}
$$

Wir zeigen, dass alle drei Terme $T_{1,n}, T_{2.n}$ und $T_{3,n}$ fast sicher gegen null konvergieren. Wir beginnen mit $T_{1,n}$. Aus der Eulerschen Identität und der Beschränktheit von $\sin$ und $\cos$ folgt

$$
\begin{aligned}
T_{1,n} &= \sup_{u\in[0,I_n]} |\varphi_X(u) - \varphi_{X_{M_n}}(u)| \\
&\leq \sup_{u\in[0,I_n]} |\mathbf{E}\{\cos(Xu) - \cos(X_{M_n}u)\}| + |\mathbf{E}\{\sin(Xu) - \sin(X_{M_n}u)\}| \\
&\leq \sup_{u\in 0,I_n]} \big|\mathbf{E}\big\{\big(\cos(Xu) - \cos(X_{M_n}u)\big)\mathbb{1}_{\{|X|\geq M_n\}}\big\}\big| \\
&\qquad + \big|\mathbf{E}\big\{\big(\sin(Xu) - \sin(X_{M_n}u)\big)\mathbb{1}_{\{|X|\geq M_n\}}\big\}\big| \\
&\leq \sup_{u\in[0,I_n]} 2\mathbf{P}[|X| > M_n] + 2\mathbf{P}[|X| > M_n] \qquad (3.4) \\
&\leq 4\int \mathbb{1}_{\{|x|>M_n\}}\mathbf{P}_X \\
&\leq 4\int \frac{|x|}{M_n}\mathbf{P}_X \\
&\leq 4\mathbf{E}\{|X|\}\frac{1}{M_n}.
\end{aligned}
$$

Zusammen mit $\frac{r_n}{M_n} \to 0$ impliziert dies $r_n T_{1,n} \to 0$ fast sicher. Nun zur Konvergenz von $T_{3,n}$. Wir verwenden analog zur Herleitung von Gleichung (3.4) die Eulersche Identität sowie die Beschränktheit von $\sin$ und $\cos$ und erhalten

$$
\begin{aligned}
T_{3,n} &\leq 4\frac{1}{n}\sum_{j=1}^{n} \mathbb{1}_{\{|X_j|>M_n\}} \\
&\leq \frac{1}{M_n}\frac{1}{n}\sum_{j=1}^{n}|X_j|.
\end{aligned}
$$

Da die X_j unabhängig und identisch verteilt sind, konvergiert $\sum_{j=1}^{n}|X_j|$ nach dem starken Gesetz der großen Zahlen [40, Theorem 12.1] fast sicher gegen $\mathbf{E}|X| < \infty$. Wegen $\lim_{n\to\infty} \frac{r_n}{M_n} = 0$ folgt

$$
r_n T_{3,n} \leq \frac{r_n}{M_n}\frac{1}{n}\sum_{j=1}^{n}|X_j| \to 0 \quad \text{f.s.}
$$

Um $T_{2,n}$ abzuschätzen, diskretisieren wir. Es gilt

$$
\begin{aligned}
r_n T_{2,n} &= r_n \sup_{u\in[0,I_n]} |\varphi_{X_{M_n}}(u) - \hat{\varphi}_{X_{1,M_n}^n}(u)| \\
&\leq r_n \sup_{u\in[0,I_n]} |\mathbf{E}\{\cos(uX_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n}\cos(uX_{j,M_n})| \qquad (3.5) \\
&\quad + r_n \sup_{u\in[0,I_n]} |\mathbf{E}\{\sin(uX_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n}\sin(uX_{j,M_n})|.
\end{aligned}
$$

Wir wählen $u_1 = 0$ und $u_j = u_{j-1}+\frac{1}{n}$ für $j \in \{2, \ldots, I_n \cdot n\}$. Zu jedem $u \in [0, I_n]$ existiert also ein $\tilde{u} \in \{u_1, \ldots, u_{I_n \cdot n}\}$ mit $|u - \tilde{u}| \leq \frac{1}{n}$. Da Cosinus und Sinus Lipschitz-stetig sind, folgt

$$\begin{aligned}
&\Big| |\mathbf{E}\{\cos(uX_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n}\cos(uX_{j,M_n})| \\
&\quad - |\mathbf{E}\{\cos(\tilde{u}X_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n}\cos(\tilde{u}X_{j,M_n})|\Big| \\
&\leq |\mathbf{E}\{\cos(uX_{M_n}) - \cos(\tilde{u}X_{M_n})| + |\frac{1}{n}\sum_{j=1}^{n}\cos(uX_{j,M_n}) - cos(\tilde{u}X_{j,M_n})| \\
&\leq |X_{M_n}| \cdot |u - \tilde{u}| + \frac{1}{n}\sum_{j=1}^{n}|X_{j,M_n}| \cdot |u - \tilde{u}| \leq \frac{2M_n}{n}
\end{aligned} \tag{3.6}$$

aus der umgekehrten Dreiecksungleichung. Für den Sinus gilt die analoge Abschätzung. Nun verwenden wir das Lemma von Borell-Cantelli (siehe Lemma 11.1 in [40]), um die fast sichere Konvergenz von $r_n T_{2,n}$ zu zeigen. Da für ein geeignetes $C > 0$

$$\frac{n}{M_n r_n} = \sqrt{n}\frac{\frac{\log(n)}{\sqrt{n}}}{r_n \frac{\log(n)}{\sqrt{n}}} \geq C\log(n) \to \infty$$

gilt, existiert für jedes $\varepsilon > 0$ ein $n_0 \in \mathbb{N}$, so dass

$$\frac{2M_n}{n} < \frac{\varepsilon}{4r_n}$$

für alle $n \geq n_0$ gilt. Zusammen mit (3.5), (3.6) und der Ungleichung von Hoeffding [17, Appendix] erhalten wir für alle $n \geq n_0$

$$\begin{aligned}
&\mathbf{P}\Big[r_n T_{2,n} > \varepsilon\Big] \\
&\leq \mathbf{P}\Big[r_n\Big(\sup_{u\in[0,I_n]}|\mathbf{E}\{\cos(uX_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n}\cos(uX_{j,M_n})| \\
&\quad + \sup_{u\in[0,I_n]}|\mathbf{E}\{\sin(uX_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n}\sin(uX_{j,M_n})|\Big) > \varepsilon\Big] \\
&\leq \mathbf{P}\Big[r_n \sup_{u\in[0,I_n]}|\mathbf{E}\{\cos(uX_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n}\cos(uX_{j,M_n})| > \frac{\varepsilon}{2}\Big] \\
&\quad + \mathbf{P}\Big[r_n \sup_{u\in[0,I_n]}|\mathbf{E}\{\sin(uX_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n}\sin(uX_{j,M_n})| > \frac{\varepsilon}{2}\Big] \\
&\leq \mathbf{P}\Big[\exists\tilde{u} \in \{u_1, \ldots, u_{I_n \cdot n}\} : r_n|\mathbf{E}\{\cos(\tilde{u}X_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n}\cos(\tilde{u}X_{j,M_n})| > \frac{\varepsilon}{4}\Big] \\
&\quad + \mathbf{P}\Big[\exists\tilde{u} \in \{u_1, \ldots, u_{I_n \cdot n}\} : r_n|\mathbf{E}\{\sin(\tilde{u}X_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n}\sin(\tilde{u}X_{j,M_n})| > \frac{\varepsilon}{4}\Big]
\end{aligned}$$

$$\leq I_n n\Big(\max_{k\in\{1,\dots,I_n n\}} \mathbf{P}\big[|\mathbf{E}\{\cos(u_k X_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n}\cos(u_k X_{j,M_n})| > \frac{\varepsilon}{4r_n}\big]$$

$$+ \max_{k\in\{1,\dots,I_n n\}} \mathbf{P}\big[|\mathbf{E}\{\sin(u_k X_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n}\sin(u_k X_{j,M_n})| > \frac{\varepsilon}{4r_n}\big]\Big)$$

$$\leq 2 \cdot I_n \cdot n \cdot 2\exp\left(\frac{-2n(\frac{\varepsilon}{4r_n})^2}{\frac{1}{n}\sum_{j=1}^{n}|1-(-1)|^2}\right)$$

$$= 4 \cdot I_n \cdot n \exp\left(-C\frac{n}{(r_n)^2}\right)$$

für eine geeignete Konstante $C > 0$, welche nicht von n abhängt. Nach Voraussetzung existieren ein $k \in \mathbb{N}$ und ein $C > 0$, so dass $I_n \cdot n \leq Cn^k$ für alle bis auf endlich viele $n \in \mathbb{N}$ gilt. Außerdem existiert ein $C' > 0$, so dass $\frac{n}{(r_n)^2} \geq C'(\log(n))^2$ für alle $n \in \mathbb{N}$ gilt. Da die Reihe

$$\sum_{n=1}^{\infty} n^k \exp(-C(\log(n))^2)$$

für alle $k \in \mathbb{N}$ und $C > 0$ konvergiert, folgt

$$\sum_{n=1}^{\infty} \mathbf{P}[r_n T_{2,n} > \varepsilon] \leq K_\varepsilon + \sum_{n=K_\varepsilon+1}^{\infty} n^k \exp(-C(\log(n))^2) < \infty,$$

da sich die ersten K_ε Summanden alle durch 1 abschätzen lassen. Das Lemma von Borel-Cantelli (siehe [46] oder [47] impliziert nun, dass $r_n T_{2,n}$ fast sicher gegen 0 konvergiert. Also konvergieren $r_n T_{1,n}, r_n T_{2,n}$ und $r_n T_{3,n}$ fast sicher gegen 0 und der Beweis ist beendet. □

Lemma 3.6. *Sei X eine $\mathbb{R}^d$-wertige Zufallsvariable mit $\mathbf{E}\{|X^2|\} < \infty$. Sei $r_n \frac{\log(n)}{\sqrt[4]{n}}$ beschränkt und es gelte $I_n \leq Cn^k$ für ein $C > 0$ und ein $k \in \mathbb{N}$. Dann gilt*

$$r_n \sup_{u_1\in[0,I_n]} \left|\frac{\partial}{\partial u_2}\varphi_X(u_1,0,\dots,0) - \frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(u_1,0,\dots,0)\right| \to 0$$

fast sicher.

Beweis. Da X integrabel ist, ist die charakteristische Funktion von X partiell differenzierbar und es gilt

$$\frac{\partial}{\partial u_2}\varphi_X(u_1,0,\dots,0) = \mathbf{E}\{iX^{(2)}\exp(iu_1X^{(1)})\},$$

weiterhin gilt

$$\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(u_1,0,\dots,0) = \frac{1}{n}\sum_{j=1}^{n} iX_j^{(2)}\exp(iu_1X_j^{(1)}).$$

Nun gehen wir vor wie im Beweis von Lemma 3.5. Sei $M_n = \sqrt[4]{n}$. Dann folgt $\frac{r_n}{M_n} \to 0$. Weiter gilt

$$\sup_{u_1\in[0,I_n]} \left|\frac{\partial}{\partial u_2}\varphi_X(u_1,0,\dots,0) - \frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(u_1,0,\dots,0)\right|$$

$$
\begin{aligned}
&\leq \sup_{u_1\in[0,I_n]} \left|\mathbf{E}\{iX^{(2)}\exp(iu_1X^{(1)})\} - \mathbf{E}\{iX^{(2)}_{M_n}\exp(iu_1X^{(1)}_{M_n})\}\right| \\
&+ \sup_{u_1\in[0,I_n]} \left|\mathbf{E}\{iX^{(2)}_{M_n}\exp(iu_1X^{(1)}_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n} iX^{(2)}_{j,M_n}\exp(iu_1X^{(1)}_{j,M_n})\right| \\
&+ \sup_{u_1\in[0,I_n]} \left|\frac{1}{n}\sum_{j=1}^{n} iX^{(2)}_{j,M_n}\exp(iu_1X^{(1)}_{j,M_n}) - \frac{1}{n}\sum_{j=1}^{n} iX^{(2)}_{j}\exp(iu_1X^{(1)}_{j})\right| \\
&= T_{1,n} + T_{2,n} + T_{3,n}.
\end{aligned}
$$

Es genügt also zu zeigen, dass $r_nT_{1,n}$, $r_nT_{2,n}$ und $r_nT_{3,n}$ für $n\to\infty$ fast sicher gegen 0 konvergieren. Wir beginnen mit $T_{1,n}$. Aus der elementaren Ungleichung

$$
|ab-cd| \leq |a-c||b| + |c||b-d|
$$

folgt

$$
\begin{aligned}
r_nT_{1,n} &\leq r_n \sup_{u\in[0,I_n]} \mathbf{E}\{|i(X^{(2)} - X^{(2)}_{M_n})||\exp(iuX^{(1)})|\} \\
&\quad + r_n \sup_{u\in[0,I_n]} \mathbf{E}\{|iX^{(2)}_{M_n}||\exp(iuX^{(1)}) - \exp(iuX^{(1)}_{M_n})|\} \\
&\leq r_n\ \mathbf{E}\{|X^{(2)}|\mathbb{1}_{\{|X^{(2)}|\geq M_n\}}\} + r_n\mathbf{E}\{2M_n\mathbb{1}_{\{|X^{(1)}|\geq M_n\}}\}.
\end{aligned}
$$

Außerdem gilt

$$
r_n\mathbf{E}\{|X^{(2)}|\mathbb{1}_{\{|X^{(2)}|\geq M_n\}}\} = r_n\int |x|\mathbb{1}_{\{|x|\geq M_n\}}\mathbf{P}_{X^{(2)}} \leq \frac{r_n}{M_n}\mathbf{E}\{|X^{(2)}|^2\} \to 0
$$

und

$$
\begin{aligned}
r_n\mathbf{E}\{2M_n\mathbb{1}_{\{|X^{(1)}|\geq M_n\}}\} &= 2r_nM_n\int \mathbb{1}_{\{|x|\geq M_n\}}\mathbf{P}_{X^{(1)}} \\
&\leq 2r_nM_n\frac{1}{(M_n)^2}\mathbf{E}\{|X^{(1)}|^2\} \to 0.
\end{aligned}
$$

Somit folgt, dass $T_{1,n}$ gegen 0 konvergiert. Nun zu $T_{3,n}$. Analog zur Abschätzung von $T_{1,n}$ erhalten wir

$$
r_nT_{3,n} \leq r_n\frac{1}{n}\sum_{j=1}^{n}|X^{(2)}_j|\mathbb{1}_{\{|X^{(2)}_j|\geq M_n\}} + r_n\frac{1}{n}\sum_{j=1}^{n}2M_n\mathbb{1}_{\{|X^{(1)}|\geq M_n\}}.
$$

Nun verwenden die Abschätzung $\mathbb{1}_{\{|x|\geq M_n\}} \leq \frac{|x|}{M_n}$ und $\mathbb{1}_{\{|x|\geq M_n\}} \leq \frac{|x|^2}{(M_n)^2}$. Zusammen mit dem starken Gesetz der großen Zahlen erhalten wir

$$
r_n\frac{1}{n}\sum_{j=1}^{n}|X^{(2)}_j|\mathbb{1}_{\{|X^{(2)}_j|\geq M_n\}} \leq \frac{r_n}{M_n}\frac{1}{n}\sum_{j=1}^{n}\left|X^{(2)}\right|^2 \to 0 \quad \text{f.s.}
$$

und

$$
r_n\frac{1}{n}\sum_{j=1}^{n}2M_n\mathbb{1}_{\{|X^{(1)}|\geq M_n\}} \leq \frac{2r_n}{M_n}\frac{1}{n}\sum_{j=1}^{n}\left|X^{(1)}\right|^2 \to 0 \quad \text{f.s..}
$$

Insgesamt folgt also $r_nT_{3,n} \to 0$ fast sicher. Es bleibt noch zu zeigen, dass $r_nT_{2,n}$ fast sicher gegen 0 konvergiert. Wie im Beweis von Lemma 3.5 diskretisieren wir und verwenden das Lemma von Borell-Cantelli. Für $u_1, u_0 \in [0, I_n]$ gilt

$$\left|iX^{(2)}_{M_n} \cos(iu_1X^{(1)}_{M_n}) - iX^{(2)}_{M_n} \cos(iu_0X^{(1)}_{M_n})\right| \leq (M_n)^2|u_1 - u_0|$$

und

$$\left|iX^{(2)}_{M_n} \sin(iu_1X^{(1)}_{M_n}) - iX^{(2)}_{M_n} \sin(iu_0X^{(1)}_{M_n})\right| \leq (M_n)^2|u_1 - u_0|.$$

Daraus folgen, analog zu (3.6), die Ungleichungen

$$\begin{aligned}&\Big||\mathbf{E}\{iX^{(2)}_{M_n} \cos(uX^{(1)}_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n} iX^{(2)}_{j,M_n} \cos(uX^{(1)}_{j,M_n})| \\ &\quad - |\mathbf{E}\{iX^{(2)}_{M_n} \cos(\tilde{u}X^{(1)}_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n} iX^{(2)}_{j,M_n} \cos(\tilde{u}X^{(1)}_{j,M_n})|\Big| \\ &\leq \frac{2(M_n)^2}{n}\end{aligned}$$

und

$$\begin{aligned}&\Big||\mathbf{E}\{iX^{(2)}_{M_n} \sin(uX^{(1)}_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n} iX^{(2)}_{j,M_n} \sin(uX^{(1)}_{j,M_n})| \\ &\quad - |\mathbf{E}\{iX^{(2)}_{M_n} \sin(u_0X^{(1)}_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n} iX^{(2)}_{j,M_n} \sin(u_0X^{(1)}_{j,M_n})|\Big| \\ &\leq \frac{2(M_n)^2}{n}.\end{aligned}$$

Ist nun $u_1 = 0$ und $u_j = u_{j-1} + \frac{1}{n}$ für $j \in \{1, \dots I_n \cdot n\}$, so existiert zu jedem $u \in [0, I_n]$ ein $\tilde{u} \in \{u_0, u_1, \dots, u_{I_n n}\}$ mit $|u - \tilde{u}| \leq \frac{1}{n}$. Wir verwenden das Lemma von Borell-Cantelli, um die Konvergenz von $r_nT_{2,n}$ zu zeigen. Sei $\varepsilon > 0$. Nach Voraussetzung existiert ein $n_0 \in \mathbb{N}$, so dass

$$\frac{2(M_n)^2}{n} \leq \frac{\varepsilon}{4r_n}$$

für alle $n \geq n_0$ gilt. Für solche n folgt mit der Ungleichung von Hoeffding

$$\begin{aligned}&\mathbf{P}\Big[r_nT_{2,n} > \varepsilon\Big] \\ &\leq \mathbf{P}\Big[r_n\Big(\sup_{u\in[0,I_n]} |\mathbf{E}\{iX^{(2)}_{M_n} \cos(uX^{(1)}_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n} iX^{(2)}_{j,M_n} \cos(uX^{(1)}_{j,M_n})| \\ &\quad + \sup_{u\in[0,I_n]} |\mathbf{E}\{iX^{(2)}_{M_n} \sin(uX^{(1)}_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n} iX^{(2)}_{j,M_n} \sin(uX^{(1)}_{j,M_n})|\Big) > \varepsilon\Big] \\ &\leq \mathbf{P}\Big[r_n \sup_{u\in[0,I_n]} |\mathbf{E}\{iX^{(2)}_{M_n} \cos(uX^{(1)}_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n} iX^{(2)}_{j,M_n} \cos(uX^{(1)}_{j,M_n})| > \frac{\varepsilon}{2}\Big] \\ &\quad + \mathbf{P}\Big[r_n \sup_{u\in[0,I_n]} |\mathbf{E}\{iX^{(2)}_{M_n} \sin(uX^{(1)}_{M_n})\} - \frac{1}{n}\sum_{j=1}^{n} iX^{(2)}_{j,M_n} \sin(uX^{(1)}_{j,M_n})| > \frac{\varepsilon}{2}\Big]\end{aligned}$$

$$
\begin{aligned}
&\leq \mathbf{P}\Big[\exists \tilde{u} \in \{u_1, \dots, u_{I_n \cdot n}\} : r_n |\mathbf{E}\{iX^{(2)}_{M_n} \cos(\tilde{u} X^{(1)}_{M_n})\} \\
&\qquad - \frac{1}{n}\sum_{j=1}^{n} iX^{(2)}_{j,M_n} \cos(\tilde{u} X^{(1)}_{j,M_n})| > \frac{\varepsilon}{4}\Big] \\
&\quad + \mathbf{P}\Big[\exists \tilde{u} \in \{u_1, \dots, u_{I_n \cdot n}\} : r_n |\mathbf{E}\{iX^{(2)}_{M_n} \sin(\tilde{u} X^{(1)}_{M_n})\} \\
&\qquad - \frac{1}{n}\sum_{j=1}^{n} iX^{(2)}_{j,M_n} \sin(\tilde{u} X^{(1)}_{j,M_n})| > \frac{\varepsilon}{4}\Big] \\
&\leq I_n n \Big(\max_{k \in \{1,\dots,I_n n\}} \mathbf{P}\Big[|\mathbf{E}\{X^{(2)}_{M_n} \cos(u_k X^{(1)}_{M_n})\} \\
&\qquad - \frac{1}{n}\sum_{j=1}^{n} X^{(2)}_{j,M_n} \cos(u_k X^{(1)}_{j,M_n})| > \frac{\varepsilon}{4r_n}\Big] \\
&\quad + \max_{k \in \{1,\dots,I_n n\}} \mathbf{P}\Big[|\mathbf{E}\{X^{(2)}_{M_n} \sin(u_k X^{(1)}_{M_n})\} \\
&\qquad - \frac{1}{n}\sum_{j=1}^{n} X^{(2)}_{j,M_n} \sin(u_k X^{(1)}_{j,M_n})| > \frac{\varepsilon}{4r_n}\Big]\Big) \\
&\leq 2I_n \cdot 2n \cdot \exp\left(\frac{-2n(\frac{\varepsilon}{4r_n})^2}{\frac{1}{n}\sum_{j=1}^{n} |M_n - (-M_n)|^2} \right) \\
&= 4I_n \cdot n \cdot \exp\left(-C \frac{n}{(r_n)^2 M_n^2} \right) \\
&\leq Cn^k \cdot \exp(-C' \log(n)^2).
\end{aligned}
$$

Hierbei haben wir im letzten Schritt genutzt, dass $r_n \frac{\log(n) M_n}{\sqrt{n}}$ nach Voraussetzung beschränkt ist. Da die Reihe

$$\sum_{n=1}^{\infty} n^k \exp(-C(\log(n))^2)$$

für alle $k \in \mathbb{N}$ und $C > 0$ konvergiert, folgt

$$\sum_{n=1}^{\infty} \mathbf{P}[r_n T_{2,n} > \varepsilon] \leq K_\varepsilon + \sum_{n=1}^{\infty} n^k \exp((\log(n))^2) < \infty,$$

da sich die ersten K_ε Summanden alle durch 1 abschätzen lassen. Da $\varepsilon > 0$ beliebig war, impliziert das Lemma von Borel-Cantelli nun, dass $r_n T_{2,n}$ fast sicher gegen 0 konvergiert. Also konvergieren $r_n T_{1,n}, r_n T_{2,n}$ und $r_n T_{3,n}$ fast sicher gegen 0 und der Beweis ist beendet. □

Lemma 3.7. *Seien die 4-ten Momente von $X^{(1)}$, $X^{(2)}$ und $X^{(3)}$ aus dem in Definition 2.3 eingeführten Modell jeweils endlich und seien $X_1, X_2, \dots$ unabhängig und identisch verteilt. Ist*

$$\hat{a}_2 := \frac{\sum_{j=1}^{n} X_j^{(2)} X_j^{(3)}}{\sum_{j=1}^{n} X_j^{(1)} X_j^{(3)}},$$

und gilt $\mathbf{E}\{Z^2\} \neq 0$, dann gilt

$$\frac{\sqrt{n}}{\log(n)} |\hat{a}_2 - a_2| \to 0$$

fast sicher.

Beweis. Da die 4-ten Momente von $X^{(1)}, X^{(2)}$ und $X^{(3)}$ endlich sind, sind ihre Produkte $X^{(1)} \cdot X^{(2)}$ und $X^{(1)} \cdot X^{(3)}$ quadratisch integrierbar. Da $X_1, X_2, \dots$ unabhängig und identisch verteilt sind, sind auch $X^{(1)}, X_1^{(1)}, X_2^{(1)}, \dots$ sowie $X^{(2)}, X_1^{(2)}, X_2^{(2)}, \dots$ und $X^{(3)}, X_1^{(3)}, X_2^{(3)}, \dots$ jeweils unabhängig und identisch verteilt. Aus dem Gesetz vom iterierten Logarithmus (siehe [40, Theorem 31.1]) folgt, dass

$$\frac{\sqrt{n}}{\sqrt{2\log(\log(n))}}\left|\frac{1}{n}\sum_{i=1}^{n} X_i^{(1)} \cdot X_i^{(3)} - \mathbf{E}\{X^{(1)} \cdot X^{(3)}\}\right|$$

und

$$\frac{\sqrt{n}}{\sqrt{2\log(\log(n))}}\left|\frac{1}{n}\sum_{i=1}^{n} X_i^{(2)} \cdot X_i^{(3)} - \mathbf{E}\{X^{(2)} \cdot X^{(3)}\}\right|$$

mit Wahrscheinlichkeit 1 beschränkt sind. Daraus folgt, dass fast sicher

$$\frac{\sqrt{n}}{\log(n)}\left|\frac{1}{n}\sum_{i=1}^{n} X_i^{(1)} \cdot X_i^{(3)} - \mathbf{E}\{X^{(1)} \cdot X^{(3)}\}\right| \to 0$$

und

$$\frac{\sqrt{n}}{\log(n)}\left|\frac{1}{n}\sum_{i=1}^{n} X_i^{(2)} \cdot X_i^{(3)} - \mathbf{E}\{X^{(2)} \cdot X^{(3)}\}\right| \to 0$$

gilt. Wie wir bereits im Beweis von Lemma 2.2 gesehen haben, gilt außerdem

$$a_2 = \frac{\mathbf{E}\{X^{(2)} \cdot X^{(3)}\}}{\mathbf{E}\{X^{(1)} \cdot X^{(3)}\}}.$$

Aus der elementaren Ungleichung

$$\begin{aligned}\left|\frac{x_n}{y_n} - \frac{x}{y}\right| &= \frac{|x_n y - xy + xy - x y_n|}{|y_n y|} \\ &\leq \frac{|y||x_n - x|}{|y y_n|} + \frac{|x||y_n - y|}{|y y_n|}\end{aligned}$$

folgt mit $x = \mathbf{E}\{X^{(2)} \cdot X^{(3)}\}$, $y = \mathbf{E}\{X^{(1)} \cdot X^{(3)}\}|$, $x_n = \frac{1}{n}\sum_{i=1}^n X_i^{(2)} \cdot X_i^{(3)}$ und $y_n = \frac{1}{n}\sum_{i=1}^n X_i^{(1)} \cdot X_i^{(3)}$, dass

$$\begin{aligned}&\frac{\sqrt{n}}{\log(n)}|\hat{a}_2 - a_2| \\ &\leq \frac{|\mathbf{E}\{X^{(2)} \cdot X^{(3)}\}| \cdot \frac{\sqrt{n}}{\log(n)}|\frac{1}{n}\sum_{i=1}^n X_i^{(1)} \cdot X_i^{(3)} - \mathbf{E}\{X_i^{(1)} \cdot X_i^{(3)}\}|}{|\frac{1}{n}\sum_{i=1}^n X_i^{(1)} \cdot X_i^{(3)} \cdot \mathbf{E}\{X_i^{(1)} \cdot X_i^{(3)}\}|} \\ &\quad + \frac{|\mathbf{E}\{X^{(1)} \cdot X^{(3)}\}| \cdot \frac{\sqrt{n}}{\log(n)}|\frac{1}{n}\sum_{i=1}^n X_i^{(2)} \cdot X_i^{(3)} - \mathbf{E}\{X_i^{(2)} \cdot X_i^{(3)}\}|}{|\frac{1}{n}\sum_{i=1}^n X_i^{(1)} \cdot X_i^{(3)} \cdot \mathbf{E}\{X_i^{(1)} \cdot X_i^{(3)}\}|}\end{aligned}$$

gilt. In beiden Brüchen konvergiert der Nenner fast sicher gegen

$$\mathbf{E}\{X^{(1)} \cdot X^{(3)}\}^2 = \mathbf{E}\{Z^2\}a_3^2 \neq 0.$$

Außerdem konvergiert der zweite Faktor im Zähler fast sicher gegen 0. Also folgt

$$\frac{\sqrt{n}}{\log(n)}|\hat{a}_2 - a_2| \to 0 \quad \text{f.s.}$$

Damit ist der Beweis beendet. $\square$

Lemma 3.8. *Sei $\hat{K}_n$ definiert wie in (2.3) und sei $\varphi_{X^{(1)}}$ nullstellenfrei. Falls $u_n \to \infty$ gilt und falls Δ_n eine Nullfolge ist, so folgt*

$$\hat{K}_n \to \infty \quad \textit{f.s.}$$

Beweis. Nach Vorraussetzung gilt $u_n \to \infty$, es genügt also, wenn wir zeigen, dass $\hat{x}_n := \max\{x \in \mathbb{R} : \inf_{u\in[0,x]}|\hat{\varphi}_{X_1^n}(u,0,\dots,0)| \geq \Delta_n\}$ fast sicher gegen ∞ konvergiert. Für $x_0 \in [0,\infty)$ setzen wir $I = [0,x_0]$. Nach Vorraussetzung ist $|\varphi_{X^{(1)}}|$ auf I nullstellenfrei und nimmt ein Minimum an. Wir bezeichnen dieses als $\varepsilon > 0$. Für n groß genug gilt außerdem $I \subseteq [0,n]$. NUn verwenden wir Lemma 3.5 und sehen, dass mit Wahrscheinlichkeit 1 ein zufälliges $\hat{C}$ existiert, so dass

$$\sup_{u\in I}|\varphi_{X^{(1)}}(u) - \hat{\varphi}_{X_1^{n(1)}}(u)| \leq C\frac{\log(n)}{\sqrt{n}}$$

gilt. Abhängig von $\hat{C}$ wählen wir ein zufälliges $\hat{n}_0$, so dass $\Delta_n \leq \frac{\varepsilon}{2}$ und $\frac{\hat{C}\log(n)}{\sqrt{n}} \leq \frac{\varepsilon}{2}$ für alle $n \geq \hat{n}_0$ gilt. Für $n \geq \hat{n}_0$ folgt, dass es mit Wahrscheinlichkeit 1

$$x_0 \in \{x \in \mathbb{R} : \inf_{u\in[0,x]}|\hat{\varphi}_{X_1^n}(u,0,\dots,0)| \geq \Delta_n\}$$

gilt. Also gilt auch $\hat{x}_n \geq x_0$ mit Wahrscheinlichkeit 1 für alle $n \geq n_0$. Da x_0 beliebig war, ist die Behauptung damit bewiesen. $\square$

Lemma 3.9. *Seien X,Z und ε Zufallsvariablen wie in Definition 2.3. Sei T_n von$\mathbb{R}^n$ nach$\mathbb{R}$ wie in Definition 2.4 gegeben durch*

$$\begin{aligned} T_n(z) = \sup_{u\in[0,\hat{K}_n]} &\left|\hat{\varphi}_{X_1^n}(u,0,\dots) - \hat{\varphi}_z(u)\cdot\hat{\varphi}_{\hat{\varepsilon}_1^n}(u)\right| \\ &+ \left|\frac{\delta}{\delta u_2}\hat{\varphi}_{X_1^n}(u,0,\dots) - \hat{a}_2\frac{\delta}{\delta u}\hat{\varphi}_z(u)\cdot\hat{\varphi}_{\hat{\varepsilon}_1^n}(u)\right| \end{aligned}$$

und sei $\hat{z}_1^n = (\hat{z}_1,\dots,\hat{z}_n)$ das Minimum von T_n unter der Nebenbedingung $\frac{1}{n}\sum_{j=1}^n|\hat{z}_j|^2 \leq \frac{1}{n}\sum_{j=1}^n|X_j^{(1)}|^2 + 1$. Sei $\hat{K}_n$ wie in (2.3) mit $u_n \leq cn^k$ für ein $c > 0$, ein $k > 0$ und alle $n \in \mathbb{N}$ und $\lim_{n\to\infty}\Delta_n = 0$. Außerdem seien die 4-ten Momente von Z und $\varepsilon_1,\dots,\varepsilon_d$ endlich. Dann gilt

$$\frac{\sqrt[4]{n}}{\log(n)}T_n(\hat{z}_1^n) \to 0$$

fast sicher.

Beweis. Um $T_n(\hat{z}_1^n)$ abzuschätzen ersetzen wir die geschätzten Daten durch die wahren Daten. Da $\hat{z}_1^n$ das Minimum von T_n unter der Nebenbedingung (2.5) ist, folgt

$$T_n(\hat{z}_1^n) \leq T_n(Z_1^n)$$

falls

$$\frac{1}{n}\sum_{j=1}^{n}|Z_j|^2 \leq \frac{1}{n}\sum_{j=1}^{n}|X_j^{(1)}|^2 + 1 \tag{3.7}$$

gilt. Nach Vorraussetzung gilt $\mathbf{E}Z^2 < \infty$ und $\mathbf{E}\varepsilon_1^2 < \infty$. Nach dem starken Gesetz der großen Zahlen konvergiert $\frac{1}{n}\sum_{j=1}^{n}|Z_j|^2$ fast sicher gegen $\mathbf{E}Z^2$, da die Z_j unabhängig und identisch verteilt sind. Ebenso konvergiert $\frac{1}{n}\sum_{j=1}^{n}|X_j^{(1)}|^2 + 1$ fast sicher gegen $\mathbf{E}\{(X^{(1)})^2\} + 1$. Da Z und ε_1 unabhängig sind, gilt außerdem

$$\mathbf{E}\{Z^2\} = \mathbf{E}\{(X^{(1)})^2\} - \underbrace{\mathbf{E}\{\varepsilon_1^2\}}_{\geq 0} \leq \mathbf{E}\{(X^{(1)})^2\}. \tag{3.8}$$

Daher ist auch (3.7) mit Wahrscheinlichkeit 1 für n groß genug erfüllt. Nun verwenden wir die Identitäten (3.1) und (3.2) sowie die Dreiecksungleichung und erhalten

$$\begin{aligned}
T_n(Z_1^n) = & \sup_{u\in[0,\hat{K}_n]} \Big[\big|\hat{\varphi}_{X_1^n}(u,0,\dots) - \hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u)\big| \\
& + \left|\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(u,0,\dots) - \hat{a}_2\frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u)\right|\Big] \\
\leq & \sup_{u\in[0,u_n]} \big|\hat{\varphi}_{X_1^n}(u,0,\dots) - \varphi_X(u,0,\dots)\big| \\
& + \sup_{u\in[0,u_n]} |\varphi_X(u,0,\dots) - \varphi_Z(u)\varphi_{\varepsilon_1}(u)| \\
& + \sup_{u\in[0,u_n]} \big|\varphi_Z(u)\varphi_{\varepsilon_1}(u) - \hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u)\big| \\
& + \sup_{u\in[0,u_n]} \left|\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(u,0,\dots) - \frac{\partial}{\partial u_2}\varphi_X(u,0,\dots)\right| \\
& + \sup_{u\in[0,u_n]} \left|\frac{\partial}{\partial u_2}\varphi_X(u,0,\dots) - a_2\frac{\partial}{\partial u}\varphi_Z(u)\varphi_{\varepsilon_1}(u)\right| \\
& + \sup_{u\in[0,u_n]} \left|a_2\frac{\partial}{\partial u}\varphi_Z(u)\varphi_{\varepsilon_1}(u) - \hat{a}_2\frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u)\right|.
\end{aligned}$$

Sei nun

$$T_{1,n} = \sup_{u\in[0,u_n]} \big|\hat{\varphi}_{X_1^n}(u,0,\dots) - \varphi_X(u,0,\dots)\big|,$$

$$T_{2,n} = \sup_{u\in[0,u_n]} \big|\varphi_Z(u)\varphi_{\varepsilon_1}(u) - \hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u)\big|,$$

$$T_{3,n} = \sup_{u\in[0,u_n]} \left|\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(u,0,\dots) - \frac{\partial}{\partial u_2}\varphi_X(u,0,\dots)\right|$$

sowie

$$T_{4,n} = \sup_{u\in[0,u_n]} \left|a_2\frac{\partial}{\partial u}\varphi_Z(u)\varphi_{\varepsilon_1}(u) - \hat{a}_2\frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u)\right|.$$

Wegen

$$\sup_{u\in[0,u_n]} |\varphi_X(u,0,\dots) - \varphi_Z(u)\varphi_{\varepsilon_1}(u)| = 0$$

und

$$\sup_{u\in[0,u_n]}\left|\frac{\partial}{\partial u_2}\varphi_X(u,0,\dots) - a_2\frac{\partial}{\partial u}\varphi_Z(u)\varphi_{\varepsilon_1}(u)\right| = 0$$

gilt dann $T_n(Z_1^n) \le T_{1,n} + T_{2,n} + T_{3,n} + T_{4,n}$ und es genügt, die Konvergenz dieser vier Ausdrücke nachzuweisen. Wir beginnen mit $T_{2,n}$. Da die 4-ten Momente von Z und ε endlich sind, sind auch die 4-ten Momente von $X^{(1)}$ endlich. Also können wir die Lemmata 3.5 bis 3.7 anwenden. Aus Lemma 3.5 folgt, dass $\frac{\sqrt{n}}{\log(n)}T_{1,n}$ fast sicher gegen 0 konvergiert. Wegen

$$\begin{aligned}
T_{2,n} &\le \sup_{u\in[0,u_n]}\left[|\varphi_Z(u)|\left|\varphi_{\varepsilon_1}(u) - \hat{\varphi}_{\varepsilon_{1,1}^n}(u)\right| + \left|\hat{\varphi}_{\varepsilon_{1,1}^n}(u)\right|\left|\varphi_Z(u) - \hat{\varphi}_{Z_1^n}(u)\right|\right] \\
&\le \sup_{u\in[0,u_n]}\left|\varphi_{\varepsilon_1}(u) - \hat{\varphi}_{\varepsilon_{1,1}^n}(u)\right| + \sup_{u\in[0,u_n]}\left|\varphi_Z(u) - \hat{\varphi}_{Z_1^n}(u)\right|
\end{aligned}$$

folgt auch die Konvergenz von $\frac{\sqrt{n}}{\log(n)}T_{2,n}$ aus Lemma 3.5. Aus Lemma 3.6 folgt, dass $\frac{\sqrt[4]{n}}{\log(n)}T_{3,n}$ fast sicher gegen 0 konvergiert. Nun zu $T_{4,n}$. Es gilt

$$\begin{aligned}
T_{4,n} &= \sup_{u\in[0,u_n]}\left|a_2\frac{\partial}{\partial u}\varphi_Z(u)\varphi_{\varepsilon_1}(u) - \hat{a}_2\frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u)\right| \\
&= \sup_{u\in[0,u_n]}\left|a_2\frac{\partial}{\partial u}\varphi_Z(u)\varphi_{\varepsilon_1}(u) - a_2\frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u)\right. \\
&\qquad \left. + a_2\frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u) - \hat{a}_2\frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u)\right| \\
&\le \sup_{u\in[0,u_n]}|a_2|\left|\frac{\partial}{\partial u}\varphi_Z(u)\varphi_{\varepsilon_1}(u) - \frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u)\right| \\
&\qquad + \sup_{u\in[0,u_n]}\left|\frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u)\right||\hat{a}_2 - a_2| \\
&= T'_{4,n} + T''_{4,n}
\end{aligned}$$

Aus Lemma 3.7 folgt, dass $\frac{\sqrt{n}}{\log(n)}|\hat{a}_2 - a_2|$ mit Wahrscheinlichkeit 1 gegen 0 konvergiert. Außerdem gilt

$$\left|\hat{\varphi}_{\varepsilon_1^n}(u)\right| \le \frac{1}{n}\sum_{j=1}^n|\exp(iu\varepsilon_j)| \le 1$$

und

$$\begin{aligned}
\left|\frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\right| &= \left|\frac{\partial}{\partial u}\frac{1}{n}\sum_{j=1}^n\exp(iuZ_j)\right| \\
&= \frac{1}{n}\left|\sum_{j=1}^n iZ_j\exp(iuZ_j)\right| \\
&\le \frac{1}{n}|Z_j||\exp(iuZ_j)| \\
&\le \frac{1}{n}\sum_{j=1}^n|Z_j|.
\end{aligned}$$

Nach dem starken Gesetz der großen Zahlen konvergiert $\frac{1}{n}\sum_{j=1}^{n}|Z_j|$ mit Wahrscheinlichkeit 1 gegen $\mathbf{E}\{|Z|\} < \infty$. Also ist auch $\left|\frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\right|$ mit Wahrscheinlichkeit 1 beschänkt. Damit folgt

$$\frac{\sqrt{n}}{\log(n)}T''_{4,n} = \frac{\sqrt{n}}{\log(n)} \sup_{u\in[0,u_n]} \left|\frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u)\right| |\hat{a}_2 - a_2| \to 0 \quad \text{f.s.}$$

Es bleibt also nur noch zu zeigen, dass auch $\frac{\sqrt[4]{n}}{\log(n)}|T'_{4,n}$ mit Wahrscheinlichkeit 1 gegen 0 konvergiert. Da $|a_2| \in \mathbb{R}$ weder von u noch n abhängt, reicht es $\left|\frac{\partial}{\partial u}\varphi_Z(u)\varphi_{\varepsilon_1}(u) - \frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u)\right|$ abzuschätzen. Es gilt

$$\begin{aligned}
&\left|\frac{\partial}{\partial u}\varphi_Z(u)\varphi_{\varepsilon_1}(u) - \frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u)\right| \\
&\leq \left|\frac{\partial}{\partial u}\varphi_Z(u)\varphi_{\varepsilon_1}(u) - \frac{\partial}{\partial u}\varphi_Z(u)\hat{\varphi}_{\varepsilon_1^n}(u)\right| \\
&\quad + \left|\frac{\partial}{\partial u}\varphi_Z(u)\hat{\varphi}_{\varepsilon_1^n}(u) - \frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u)\right| \\
&\leq \left|\frac{\partial}{\partial u}\varphi_Z(u)\right| \left|\varphi_{\varepsilon_1}(u) - \hat{\varphi}_{\varepsilon_1^n}(u)\right| + \left|\hat{\varphi}_{\varepsilon_1^n}(u)\right| \left|\frac{\partial}{\partial u}\varphi_Z(u) - \frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\right|.
\end{aligned}$$

Aus Lemma 3.6 folgt $\frac{\sqrt[4]{n}}{\log(n)}\left|\frac{\partial}{\partial u}\varphi_Z(u) - \frac{\partial}{\partial u}\hat{\varphi}_{Z_1^n}(u)\right| \to 0$ fast sicher und aus Lemma 3.5 folgt $\frac{\sqrt{n}}{\log(n)}\left|\varphi_{\varepsilon_1}(u) - \hat{\varphi}_{\varepsilon_1^n}(u)\right| \to$ fast sicher. Außerdem gilt $|\hat{\varphi}_{\varepsilon_1^n}(u)| \leq 1$. Zudem gilt nach Voraussetzung $\mathbf{E}\{|Z|\} < \infty$, also folgt (siehe beispielsweise [38, Satz 15.31])

$$\frac{\partial}{\partial u}\varphi_Z(u) = i\int Z\exp(iuZ)\,\mathbf{P}$$

was wiederum

$$\left|\frac{\partial}{\partial u}\varphi_Z(u)\right| \leq \int |Z|\,\mathbf{P} = \mathbf{E}\{|Z|\} < \infty$$

impliziert. Insgesamt folgt, dass auch $\frac{\sqrt[4]{n}}{\log(n)}T'_{4,n}$ fast sicher gegen 0 konvergiert und dass somit auch mit Wahrscheinlichkeit 1

$$\frac{\sqrt[4]{n}}{\log(n)}T_n(\hat{z}_1^n) \to 0$$

gilt. Damit ist der Beweis beendet. □

Bemerkung. Betrachten wir den Beweis von Lemma 3.9 noch einmal genauer. Aus den Abschätzungen für $T_{1,n}$ und $T_{2,n}$ folgt, dass auch

$$\frac{\sqrt{n}}{\log(n)}\left|\hat{\varphi}_{X_1^n}(u,0,\dots) - \hat{\varphi}_{Z_1^n}(u)\hat{\varphi}_{\varepsilon_1^n}(u)\right|$$

mit Wahrscheinlichkeit 1 gegen 0 konvergiert.

Beweis von Lemma 3.3. Unter den Voraussetzungen von Lemma 3.3 sind auch die Voraussetzungen der Lemmata 3.5 bis 3.9 erfüllt. Wir werden diese im Beweis verwenden.

Außerdem verwenden wir erneut die Logarithmen aus dem Beweis von Lemma 2.2, deren Definition wir an dieser Stelle kurz wiederholen. Wir definieren

$$L_{\varphi_Z}(v) := \int_0^v \frac{\frac{\partial}{\partial u}\varphi_Z(u)}{\varphi_Z(u)}\,du \quad \text{und} \quad L_{\hat{\varphi}_{\hat{z}_1^n}}(v) := \int_0^v \frac{\frac{\partial}{\partial u}\hat{\varphi}_{\hat{z}_1^n}(u)}{\hat{\varphi}_{\hat{z}_1^n}(u)}\,du,$$

dann gilt

$$\exp(L_{\varphi_Z}(v)) = \exp\left(\int_0^v \frac{\frac{\partial}{\partial u}\varphi_Z(u)}{\varphi_Z(u)}\,du\right) = \varphi_Z(v)$$

und

$$\exp(L_{\hat{\varphi}_{\hat{z}_1^n}}(v)) = \exp(L_{\hat{\varphi}_{\hat{z}_1^n}}(v)) = \exp\left(\int_0^v \frac{\frac{\partial}{\partial u}\hat{\varphi}_{\hat{z}_1^n}(u)}{\hat{\varphi}_{\hat{z}_1^n}(u)}\,du\right) = \hat{\varphi}_{\hat{z}_1^n}(v).$$

$L_{\varphi_Z}(v)$ ist also ein Logarithmus von $\varphi_Z(v)$ und $L_{\hat{\varphi}_{\hat{z}_1^n}}(v)$ ist ein Logarithmus von $\hat{\varphi}_{\hat{z}_1^n}(v)$. Damit folgt

$$\begin{aligned}
|\hat{\varphi}_{\hat{z}_1^n}(u) - \varphi_Z(u)| &\leq |\exp(L_{\varphi_Z}(u)) - \exp(L_{\hat{\varphi}_{\hat{z}_1^n}}(u))| \\
&\leq |\exp(L_{\hat{\varphi}_{\hat{z}_1^n}}(u))||\exp(L_{\varphi_Z}(u) - L_{\hat{\varphi}_{\hat{z}_1^n}}(u)) - 1| \\
&\leq |\exp\left(\mathrm{Re}(L_{\varphi_Z}(u) - L_{\hat{\varphi}_{\hat{z}_1^n}}(u))\right) - 1| \\
&\quad + |\exp\left(i\,\mathrm{Im}\,(L_{\varphi_Z}(u) - L_{\hat{\varphi}_{\hat{z}_1^n}}(u))\right) - 1|,
\end{aligned}$$

wobei wir ausgenutzt haben, dass

$$\begin{aligned}
|\exp(L_{\hat{\varphi}_{\hat{z}_1^n}}(u))| = |\hat{\varphi}_{\hat{z}_1^n}(u)| &= |\frac{1}{n}\sum_{j=1}^n \exp(i \cdot u \cdot \hat{z}_j)| \\
&\leq \frac{1}{n}\sum_{j=1}^n |\exp(i \cdot u \cdot \hat{z}_j)| = 1
\end{aligned}$$

gilt. Wegen $\exp(0) = 1$ können wir den ersten Summanden der rechten Seite, also $|\exp(\mathrm{Re}(L_{\varphi_Z}(u) - L_{\hat{\varphi}_{\hat{z}_1^n}}(u))) - 1|$, unter Verwendung des Mittelwertsatzes abschätzen. Den zweiten Summanden der rechten Seite, also

$$|\exp\left(i\,\mathrm{Im}\,(L_{\varphi_Z}(u) - L_{\hat{\varphi}_{\hat{z}_1^n}}(u))\right) - 1|,$$

formen wir mit der Eulerformel

$$\exp(it) = \cos(t) + i\sin(t) \quad (t \in \mathbb{R})$$

weiter um. Es gilt, für ein geeignetes $\xi \in [0, \mathrm{Re}((L_{\varphi_Z}(u) - L_{\hat{\varphi}_{\hat{z}_1^n}}(u))]$,

$$\begin{aligned}
|\hat{\varphi}_{\hat{z}_1^n}(u) - \varphi_Z(u)| &\leq |\exp\left(\mathrm{Re}(L_{\varphi_Z}(u) - L_{\hat{\varphi}_{\hat{z}_1^n}}(u))\right) - \exp(0)| \\
&\quad + |\exp\left(i\,\mathrm{Im}\,(L_{\varphi_Z}(u) - L_{\hat{\varphi}_{\hat{z}_1^n}}(u))\right) - 1| \\
&\leq \exp(\xi)|\mathrm{Re}((L_{\varphi_Z}(u) - L_{\hat{\varphi}_{\hat{z}_1^n}}(u))| \\
&\quad + |\cos(\mathrm{Im}\,(L_{\varphi_Z}(u) - L_{\hat{\varphi}_{\hat{z}_1^n}}(u))) - 1|
\end{aligned}$$

$$
\begin{aligned}
&\quad + |i \sin(\mathrm{Im}\,(L_{\varphi_Z}(u) - L_{\hat\varphi_{\hat z_1^n}}(u)))| \\
&\le \exp(\xi)|\mathrm{Re}((L_{\varphi_Z}(u) - L_{\hat\varphi_{\hat z_1^n}}(u))| \\
&\quad + |\cos(\mathrm{Im}\,(L_{\varphi_Z}(u) - L_{\hat\varphi_{\hat z_1^n}}(u))) - 1| \\
&\quad + |\sin(\mathrm{Im}\,(L_{\varphi_Z}(u) - L_{\hat\varphi_{\hat z_1^n}}(u))) - 0| \\
&\le \exp(\xi)|\mathrm{Re}((L_{\varphi_Z}(u) - L_{\hat\varphi_{\hat z_1^n}}(u))| + 2|L_{\varphi_Z}(u) - L_{\hat\varphi_{\hat z_1^n}}(u)|.
\end{aligned}
$$

Dabei haben wir für die letzte Abschätzung ausgenutzt, dass Kosinus und Sinus beide Lipschitz-stetig sind, und dass

$$|\mathrm{Im}\,(L_{\varphi_Z}(u) - L_{\hat\varphi_{\hat z_1^n}}(u)))| \le |L_{\varphi_Z}(u) - L_{\hat\varphi_{\hat z_1^n}}(u)|$$

gilt. Gilt nun $|L_{\varphi_Z}(v) - L_{\hat\varphi_{\hat z_1^n}}(v)| \to 0$ fast sicher, so exisitert ein n_0, so dass

$$\exp(\xi) \le \max_{x \in [0, \mathrm{Re}((L_{\varphi_Z}(u) - L_{\hat\varphi_{\hat z_1^n}}(u))]} \le 2$$

für alle $n \ge n_0$ gilt. Außerdem gilt

$$|\mathrm{Re}(L_{\varphi_Z}(u) - L_{\hat\varphi_{\hat z_1^n}}(u)))| \le |L_{\varphi_Z}(u) - L_{\hat\varphi_{\hat z_1^n}}(u)|.$$

Es genügt also zu zeigen, dass

$$r_n \sup_{v \in [0, \hat K_n]} |L_{\varphi_Z}(v) - L_{\hat\varphi_{\hat z_1^n}}(v)| \tag{3.9}$$

fast sicher gegen 0 konvergiert. Um $|L_{\varphi_Z}(v) - L_{\hat\varphi_{\hat z_1^n}}(v)|$ abzuschätzen, verwenden wir die aus dem Beweis von Lemma 2.2 bekannten Identitäten

$$\varphi_X(u_1, 0, \dots, 0) = \varphi_Z(u_1)\varphi_{\varepsilon_1}(u_1)$$

und

$$\frac{\partial}{\partial u_2}\varphi_X(u_1, 0, \dots, 0) = a_2 \frac{\partial}{\partial u}\varphi_Z(u_1)\varphi_{\varepsilon_1}(u_1).$$

Sein nun $u_0 \in \mathbb{R}^n$ gegeben durch $u_0 = (u, 0, \dots)$ und $v_0 \in \mathbb{R}^n$ gegeben durch $v_0 = (v, 0, \dots)$. Es gilt

$$
\begin{aligned}
|L_{\varphi_Z}(v) - L_{\hat\varphi_{\hat z_1^n}}(v)| &\le \left| \int_0^v \frac{\frac{\partial}{\partial u}\varphi_Z(u)}{\varphi_Z(u)}\, du - \int_0^v \frac{\frac{\partial}{\partial u}\hat\varphi_{\hat z_1^n}(u)}{\hat\varphi_{\hat z_1^n}(u)}\, du \right| \\
&\le \left| \int_0^v \frac{1}{a_2} \frac{\frac{\partial}{\partial u_2}\varphi_X(u_0)}{\varphi_X(u_0)}\, du - \int_0^v \frac{\frac{\partial}{\partial u}\hat\varphi_{\hat z_1^n}(u)}{\hat\varphi_{\hat z_1^n}(u)}\, du \right| \\
&\le \left| \int_0^v \frac{1}{a_2} \frac{\frac{\partial}{\partial u_2}\varphi_X(u_0)}{\varphi_X(u_0)}\, du - \int_0^v \frac{1}{\hat a_2} \frac{\frac{\partial}{\partial u_2}\hat\varphi_{X_1^n}(u_0)}{\hat\varphi_{X_1^n}(u_0)}\, du \right| \\
&\quad + \left| \int_0^v \frac{1}{\hat a_2} \frac{\frac{\partial}{\partial u_2}\hat\varphi_{X_1^n}(u_0)}{\hat\varphi_{X_1^n}(u_0)}\, du - \int_0^v \frac{\hat a_2 \frac{\partial}{\partial u}\hat\varphi_{\hat z_1^n}(u)\hat\varphi_{\hat\varepsilon_1^n}(u)}{\hat a_2 \hat\varphi_{\hat z_1^n}(u)\hat\varphi_{\hat\varepsilon_1^n}(u)}\, du \right| \\
&= T_{1,n} + T_{2,n}.
\end{aligned}
$$

Außerdem gilt

$$\sup_{v\in[0,\hat{K}_n]} T_{1,n} = \sup_{v\in[0,\hat{K}_n]} \left| \int_0^v \frac{1}{a_2} \frac{\frac{\partial}{\partial u_2}\varphi_X(u_0)}{\varphi_X(u_0)}\, du - \int_0^v \frac{1}{\hat{a}_2} \frac{\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(u_0)}{\hat{\varphi}_{X_1^n}(u_0)}\, du \right|$$
$$\leq \hat{K}_n \sup_{v\in[0,\hat{K}_n]} \left| \frac{1}{a_2} \frac{\frac{\partial}{\partial u_2}\varphi_X(v_0)}{\varphi_X(v_0)} - \frac{1}{\hat{a}_2} \frac{\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(v_0)}{\hat{\varphi}_{X_1^n}(v_0)} \right|$$

und

$$\sup_{v\in[0,\hat{K}_n]} T_{2,n} = \sup_{v\in[0,\hat{K}_n]} \left| \int_0^v \frac{1}{\hat{a}_2} \frac{\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(u_0)}{\hat{\varphi}_{X_1^n}(u_0)}\, du - \int_0^v \frac{\hat{a}_2 \frac{\partial}{\partial u}\hat{\varphi}_{\hat{z}_1^n}(u)\hat{\varphi}_{\hat{\varepsilon}_1^n}(u)}{\hat{a}_2 \hat{\varphi}_{\hat{z}_1^n}(u)\hat{\varphi}_{\hat{\varepsilon}_1^n}(u)}\, du \right|$$
$$\leq \hat{K}_n \sup_{v\in[0,\hat{K}_n]} \left| \frac{1}{\hat{a}_2} \frac{\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(v_0)}{\hat{\varphi}_{X_1^n}(v_0)} - \frac{\hat{a}_2 \frac{\partial}{\partial u}\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)}{\hat{a}_2 \hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)} \right|.$$

Als Nächstes bringen wir die Brüche auf einen gemeinsamen Nenner. Es folgt

$$\sup_{v\in[0,\hat{K}_n]} T_{1,n}$$
$$\leq \hat{K}_n \sup_{v\in[0,\hat{K}_n]} \left| \frac{\hat{a}_2 \frac{\partial}{\partial u_2}\varphi_X(v_0)\hat{\varphi}_{X_1^n}(v_0) - a_2 \frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(v_0)\varphi_X(v_0)}{a_2\hat{a}_2\varphi_X(v_0)\hat{\varphi}_{X_1^n}(v_0)} \right|$$
$$= \hat{K}_n \sup_{v\in[0,\hat{K}_n]} \left| \frac{\hat{a}_2 \frac{\partial}{\partial u_2}\varphi_X(v_0)\hat{\varphi}_{X_1^n}(v_0)}{a_2\hat{a}_2\varphi_X(v_0)\hat{\varphi}_{X_1^n}(v_0)} - \frac{a_2 \frac{\partial}{\partial u_2}\varphi_X(v_0)\varphi_X(v_0)}{a_2\hat{a}_2\varphi_X(v_0)\hat{\varphi}_{X_1^n}(v_0)} \right.$$
$$\left. + \frac{a_2 \frac{\partial}{\partial u_2}\varphi_X(v_0)\varphi_X(v_0)}{a_2\hat{a}_2\varphi_X(v_0)\hat{\varphi}_{X_1^n}(v_0)} - \frac{a_2 \frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(v_0)\varphi_X(v_0)}{a_2\hat{a}_2\varphi_X(v_0)\hat{\varphi}_{X_1^n}(v_0)} \right|$$
$$\leq \hat{K}_n \sup_{v\in[0,\hat{K}_n]} \frac{1}{|a_2\hat{a}_2\varphi_X(v_0)\hat{\varphi}_{X_1^n}(v_0)|} \left[\left| \frac{\partial}{\partial u_2}\varphi_X(v_0) \right| \left| \hat{a}_2\hat{\varphi}_{X_n^1}(v_0) - a_2\varphi_X(v_0) \right| \right.$$
$$\left. + |a_2\varphi_X(v_0)| \left| \frac{\partial}{\partial u_2}\varphi_X(v_0) - \frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(v_0) \right| \right]$$

und

$$\sup_{v\in[0,\hat{K}_n]} T_{2,n}$$
$$\leq \hat{K}_n \sup_{v\in[0,\hat{K}_n]} \left| \frac{\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(v_0)\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v) - \hat{a}_2 \frac{\partial}{\partial u}\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)\hat{\varphi}_{X_1^n}(v_0)}{\hat{a}_2\hat{\varphi}_{X_1^n}(v_0)\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)} \right|$$
$$= \hat{K}_n \sup_{v\in[0,\hat{K}_n]} \left| \frac{\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(v_0)\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)}{\hat{a}_2\hat{\varphi}_{X_1^n}(v_0)\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)} - \frac{\hat{a}_2 \frac{\partial}{\partial u}\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)}{\hat{a}_2\hat{\varphi}_{X_1^n}(v_0)\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)} \right.$$
$$\left. + \frac{\hat{a}_2 \frac{\partial}{\partial u}\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)}{\hat{a}_2\hat{\varphi}_{X_1^n}(v_0)\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)} - \frac{\hat{a}_2 \frac{\partial}{\partial u}\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)\hat{\varphi}_{X_1^n}(v_0)}{\hat{a}_2\hat{\varphi}_{X_1^n}(v_0)\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)} \right|$$
$$\leq \hat{K}_n \sup_{v\in[0,\hat{K}_n]} \frac{1}{\left|\hat{a}_2\hat{\varphi}_{X_1^n}(v_0)\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)\right|} \left[\left|\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)\right| \right.$$
$$\left| \frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(v_0) - \hat{a}_2 \frac{\partial}{\partial u}\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v) \right|$$

$$+ \left|\hat{a}_2 \frac{\partial}{\partial u} \hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)\right| \left|\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v) - \hat{\varphi}_{X_n^1}(v_0)\right|\Bigg].$$

Im nächsten Schritt schätzen wir die Nenner

$$|a_2\hat{a}_2\varphi_X(v_0)\hat{\varphi}_{X_1^n}(v_0)| \quad \text{und} \quad |\hat{a}_2\hat{\varphi}_{X_1^n}(v_0)\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)|$$

nach unten ab. Nach Wahl von $\hat{K}_n$ in Definition 2.4 gilt $|\hat{\varphi}_{X_n^1}(v, 0, \dots)| \geq \Delta_n$ für alle $v \in [0, \hat{K}_n]$.

Wegen $u_n \leq n^k$ können wir Lemma 3.5 verwenden und es gilt

$$\lim_{n\to\infty} \frac{\sqrt{n}}{\log(n)} \sup_{u\in[0,\hat{K}_n]} |\varphi_X(u_0) - \hat{\varphi}_{X_1^n}(u_0)| = 0$$

fast sicher. Da für n groß genug nach Voraussetzung $\frac{\log(n)}{\sqrt{n}} \leq \frac{\Delta_n}{2}$ gilt, folgt aus der umgekehrten Dreiecksungleichung, dass mit Wahrscheinlichkeit 1 für n groß genug auch

$$\begin{aligned} |\varphi_X(u_0)| &\geq |\hat{\varphi}_{X_1^n}(u_0) + \varphi_X(u_0) - \hat{\varphi}_{X_1^n}(u_0)| \\ &\geq |\hat{\varphi}_{X_1^n}(u_0)| - |\varphi_X(u_0) - \hat{\varphi}_{X_1^n}(u_0)| \\ &\geq \Delta_n - \frac{\log(n)}{\sqrt{n}} \\ &\geq \frac{\Delta_n}{2} \end{aligned} \tag{3.10}$$

für alle $u \in [0, \hat{K}_n]$ gilt. Für $u \in [0, \hat{K}_n]$ gilt also $\left|\varphi_X(u_0)\hat{\varphi}_{X_1^n}(u_0)\right| \geq \frac{\Delta_n^2}{2}$. Den zweiten Nenner schätzen wir unter Verwendung von Lemma 3.9 beziehungsweise der auf das Lemma folgenden Bemerkung ab. Aus dieser folgt, dass mit Wahrscheinlichkeit 1

$$\frac{\sqrt{n}}{\log(n)} \sup_{u\in[0,K_n]} \left|\hat{\varphi}_{X_1^n}(u, 0, \dots) - \hat{\varphi}_z(u) \cdot \hat{\varphi}_{\hat{\varepsilon}_1^n}(u)\right| \to 0$$

gilt. Außerdem konvergiert nach Lemma 3.7 auch der Faktor $\hat{a}_2$ fast sicher gegen $a_2 \neq 0$. Mit Wahrscheinlichkeit 1 gilt also $\hat{a}_2 \geq \frac{a_2}{2}$ für n groß genug. Wir nutzen erneut die umgekehrte Dreiecksungleichung und die Tatsache, dass für n groß genug $\frac{\log(n)}{\sqrt{n}} \leq \frac{\Delta_n}{2}$ gilt, und erhalten, dass mit Wahrscheinlichkeit 1 für n groß genug

$$\left|\hat{a}_2\hat{\varphi}_{X_1^n}(v_0)\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)\right| \geq a_2\frac{\Delta_n^2}{4} \tag{3.11}$$

gilt. Wir verwenden (3.10) und (3.11) um $\sup_{v\in[0,\hat{K}_n]} T_{1,n}$ und $\sup_{v\in[0,\hat{K}_n]} T_{1,n}$ weiter abzuschätzen. Für n groß genug gilt

$$\begin{aligned} \sup_{v\in[0,\hat{K}_n]} T_{1,n} \leq \hat{K}_n \sup_{v\in[0,\hat{K}_n]} C_1 \frac{1}{(\Delta_n)^2} \Bigg[&\left|\frac{\partial}{\partial u_2}\varphi_X(v_0)\right| \left|\hat{a}_2\hat{\varphi}_{X_n^1}(v_0) - a_2\varphi_X(v_0)\right| \\ &+ |a_2\varphi_X(v_0)| \left|\frac{\partial}{\partial u_2}\varphi_X(v_0) - \frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(v_0)\right|\Bigg] \end{aligned}$$

und

$$\sup_{v\in[0,\hat{K}_n]} T_{2,n}$$

$$\leq \hat{K}_n \sup_{v\in[0,\hat{K}_n]} C_2 \frac{1}{(\Delta_n)^2}\Big[\left|\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)\right| \left|\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(v_0) - \hat{a}_2\frac{\partial}{\partial u_1}\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)\right|$$
$$+ \left|\hat{a}_2\frac{\partial}{\partial u_1}\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)\right| \left|\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v) - \hat{\varphi}_{X_n^1}(v_0)\right|\Big],$$

wobei die Konstanten C_1 und C_2 nur von a_2 abhängen. Als Nächstes etablieren wir obere Schranken für die Faktoren $\left|\frac{\partial}{\partial u_2}\varphi_X(u)\right|$ und $\left|\hat{a}_2\frac{\delta}{\delta u}\hat{\varphi}_z(u)\cdot\hat{\varphi}_{\hat{\varepsilon}_1^n}(u)\right|$. Es gilt

$$\begin{aligned}
\left|\frac{\partial}{\partial u_2}\varphi_X(u)\right| &\leq \left|\frac{\partial}{\partial u_2}\int \exp(i\sum_{j=1}^d x_j u_j)\mathbf{P}_X\right| \\
&\leq \int \left|\frac{\partial}{\partial u_2}\exp(i\sum_{j=1}^d x_j u_j)\right|\mathbf{P}_X \\
&\leq \int |x_2| \left|\exp(i\sum_{j=1}^d x_j u_j)\right|\mathbf{P}_X \\
&\leq \int |X|\mathbf{P}_X \\
&\leq \mathbf{E}\{|X|\} < \infty,
\end{aligned}$$

da nach Voraussetzung $\mathbf{E}\{Z\} < \infty$ und $\mathbf{E}\{\varepsilon_1\} = \cdots = \mathbf{E}\{\varepsilon_d\} = 0$ gilt, was $\mathbf{E}\{|X|\} < \infty$ impliziert. Wir benötigen noch eine Variante dieser Abschätzung für die geschätzten Daten $\hat{z}_1, \ldots, \hat{z}_n$. Aus Lemma 3.9 folgt, dass mit Wahrscheinlichkeit 1

$$\sup_{u\in[0,\hat{K}]} |\frac{\delta}{\delta u_2}\hat{\varphi}_{X_1^n}(u,0,\ldots) - \hat{a}_2\frac{\delta}{\delta u}\hat{\varphi}_z(u)\cdot\hat{\varphi}_{\hat{\varepsilon}_1^n}(u)| \to 0 \tag{3.12}$$

gilt. Da die Daten $X_1, \ldots, X_n$ unabhängig und identisch verteilt sind, gilt für alle $u' \in \mathbb{R}^d$ nach dem starken Gesetz der großen Zahlen

$$\begin{aligned}
\frac{\delta}{\delta u_2}\hat{\varphi}_{X_1^n}(u) &= \frac{\delta}{\delta u_2}\sum_{j=1}^n \exp\left(i\sum_{l=1}^d X_j^{(l)}u_l\right) \\
&= \sum_{j=1}^n X_j^{(2)}\exp(i\left(i\sum_{l=1}^d X_j^{(l)}u_l\right) \to \mathbf{E}\Big[X^{(2)}\exp\left(i\sum_{l=1}^d X^{(l)}u_l\right)\Big]
\end{aligned}$$

fast sicher. Für $u = u_0 = (u, 0, \ldots)$ gilt also

$$\left|\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(u_0)\right| \to \left|\mathbf{E}\{X^{(2)}\exp(iX^{(1)}u)\}\right| \leq \mathbf{E}\{|X^{(2)}|\} < \infty$$

fast sicher. Zusammen mit (3.12) folgt, dass mit Wahrscheinlichkeit 1 für n groß genug

$$\left|\hat{a}_2\frac{\delta}{\delta u}\hat{\varphi}_z(u)\cdot\hat{\varphi}_{\hat{\varepsilon}_1^n}(u)\right| \leq \mathbf{E}\{|X^{(2)}|\} + 1$$

gilt. Zusammen mit den bekannte Abschätzungen

$$\sup_{v\in\mathbb{R}}|\varphi_X(v_0)| \leq 1, \quad \sup_{v\in\mathbb{R}}|\hat{\varphi}_{\hat{z}_1^n}(v)| \leq 1 \text{ und } \sup_{v\in\mathbb{R}}|\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)| \leq 1$$

erhalten wir, dass eine Konstante C existiert, so dass mit Wahrscheinlichkeit 1 für n groß genug

$$\sup_{v\in[0,\hat{K}_n]} T_{1,n} \leq C\frac{\hat{K}_n}{(\Delta_n)^2}\Bigg[\sup_{v\in[0,\hat{K}_n]}\left|\hat{a}_2\hat{\varphi}_{X_n^1}(v_0)-a_2\varphi_X(v_0)\right| + \sup_{v\in[0,\hat{K}_n]}\left|\frac{\partial}{\partial u_2}\varphi_X(v_0)-\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(v_0)\right|\Bigg]$$

sowie

$$\sup_{v\in[0,\hat{K}_n]} T_{2,n} \leq C\frac{\hat{K}_n}{(\Delta_n)^2}\sup_{v\in[0,\hat{K}_n]}\Bigg[\left|\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(v_0)-\hat{a}_2\frac{\partial}{\partial u_1}\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)\right| + \left|\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)-\hat{\varphi}_{X_n^1}(v_0)\right|\Bigg]$$

gilt. Es müssen also nur noch die Differenzen abgeschätzt werden. Nach Lemma 3.9 gilt

$$\frac{\sqrt[4]{n}}{\log(n)}\sup_{v\in[0,\hat{K}_n]}\left|\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(v_0)-\hat{a}_2\frac{\partial}{\partial u_1}\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)\right| + \left|\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)-\hat{\varphi}_{X_n^1}(v_0)\right| \to 0$$

fast sicher, aus Lemma 3.6 folgt

$$\frac{\sqrt[4]{n}}{\log(n)}\sup_{v\in[0,\hat{K}_n]}\left|\frac{\partial}{\partial u_2}\varphi_X(v_0)-\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(v_0)\right| \to 0$$

fast sicher und aus Lemma 3.5 und Lemma 3.7 folgt

$$\begin{aligned}
&\frac{\sqrt[4]{n}}{\log(n)}\sup_{v\in[0,\hat{K}_n]}\left|\hat{a}_2\hat{\varphi}_{X_n^1}(v_0)-a_2\varphi_X(v_0)\right|\\
&\leq \frac{\sqrt{n}}{\log(n)}\sup_{v\in[0,\hat{K}_n]}\left|\hat{a}_2\hat{\varphi}_{X_n^1}(v_0)-a_2\varphi_X(v_0)\right|\\
&\leq \frac{\sqrt{n}}{\log(n)}\sup_{v\in[0,\hat{K}_n]}\left|\hat{a}_2\hat{\varphi}_{X_n^1}(v_0)-a_2\hat{\varphi}_{X_n^1}(v_0)+a_2\hat{\varphi}_{X_n^1}(v_0)-a_2\varphi_X(v_0)\right|\\
&\leq \frac{\sqrt{n}}{\log(n)}\sup_{v\in[0,\hat{K}_n]}\left|\hat{\varphi}_{X_n^1}(v_0)\right|\left|\hat{a}_2-a_2\right| + \frac{\sqrt{n}}{\log(n)}\sup_{v\in[0,\hat{K}_n]}|a_2|\left|\hat{\varphi}_{X_n^1}(v_0)-\varphi_X(v_0)\right|\\
&\leq \frac{\sqrt{n}}{\log(n)}|\hat{a}_2-a_2| + |a_2|\frac{\sqrt{n}}{\log(n)}\sup_{v\in[0,\hat{K}_n]}\left|\hat{\varphi}_{X_n^1}(v_0)-\varphi_X(v_0)\right| \to 0
\end{aligned}$$

fast sicher. Zusammen mit $r_n = \frac{\Delta_n^2}{\hat{K}_n}\frac{\sqrt[4]{n}}{\log(n)}$ folgt also

$$\begin{aligned}
&r_n\sup_{v\in[0,\hat{K}_n]}|L_{\varphi_Z}(v)-L_{\hat{\varphi}_{\hat{z}_1^n}}(v)|\\
&\leq r_n\sup_{v\in[0,\hat{K}_n]} T_{1,n} + r_n\sup_{v\in[0,\hat{K}_n]} T_{2,n}
\end{aligned}$$

$$
\begin{aligned}
&\leq \frac{\sqrt[4]{n}}{\log(n)} \sup_{v\in[0,\hat{K}_n]} \left|\hat{a}_2\hat{\varphi}_{X_n^1}(v_0) - a_2\varphi_X(v_0)\right| \\
&\quad + \frac{\sqrt[4]{n}}{\log(n)} \sup_{v\in[0,\hat{K}_n]} \left|\frac{\partial}{\partial u_2}\varphi_X(v_0) - \frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(v_0)\right| \\
&\quad + \frac{\sqrt[4]{n}}{\log(n)} \sup_{v\in[0,\hat{K}_n]} \left[\left|\frac{\partial}{\partial u_2}\hat{\varphi}_{X_1^n}(v_0) - \hat{a}_2\frac{\partial}{\partial u_1}\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v)\right|\right. \\
&\quad \left. + \left|\hat{\varphi}_{\hat{z}_1^n}(v)\hat{\varphi}_{\hat{\varepsilon}_1^n}(v) - \hat{\varphi}_{X_n^1}(v_0)\right|\right] \to 0
\end{aligned}
$$

fast sicher. Damit ist die Hilfsaussage (3.9) gezeigt und der Beweis beendet. □

3.2 Beweis der Konvergenz des Verteilungsschätzers

Beweis von Theorem 2.6. Es gilt

$$
|\hat{\mu}_n([a,b]) - \mathbf{P}_Z([a,b])| = \left|\sum_{j=1}^{n} \mathbb{1}_{[a,b]}(\hat{z}_j) - \int \mathbb{1}_{[a,b]}\mathbf{P}_Z\right|.
$$

Wir wollen Lemma 3.2 verwenden und müssen daher $\mathbb{1}_{[a,b]}$ durch trigonometrische Polynome approximieren. Da ein stetiges trigonometrisches Polynom nicht gleichmäßig gegen die unstetige Funktion $\mathbb{1}_{[a,b]}$ konvergieren kann, konstruieren wir zunächste eine Lipschitz-stetige Approximation der Indikatorfunktion. Sei $0 < \varepsilon < \frac{b-a}{2}$. Wir approximieren die Indikatorfunktion von $[a,b]$ von oben und von unten durch Lipschitz-stetige Funktionen. Die Funktionen

$$
f^{a,b}(x) := \begin{cases} 1 \text{ für } x \in [a,b] \\ 0 \text{ für } x \in [a-\varepsilon, b+\varepsilon]^C \\ \frac{1}{\varepsilon}x - \frac{a}{\varepsilon} + 1 \text{ für } x \in [a-\varepsilon, a) \\ \frac{-1}{\varepsilon}x + \frac{b}{\varepsilon} + 1 \text{ für } x \in (b, b+\varepsilon] \end{cases} \tag{3.13}
$$

und

$$
f_{a,b}(x) := \begin{cases} 1 \text{ für } x \in [a+\varepsilon, b-\varepsilon] \\ 0 \text{ für } x \in [a,b]^C \\ \frac{1}{\varepsilon}x - \frac{a}{\varepsilon} \text{ für } x \in [a, a+\varepsilon) \\ \frac{-1}{\varepsilon}x + \frac{b}{\varepsilon} \text{ für } x \in (b-\varepsilon, b] \end{cases} \tag{3.14}
$$

sind Lipschitz-stetig mit Lipschitz-Konstante $\frac{1}{\varepsilon}$ und es gilt

$$
0 \leq f_{a,b} \leq \mathbb{1}_{[a.b]} \leq f^{a,b} \leq 1
$$

sowie

$$
f^{a,b} - f_{a,b} = |f_{a,b} - f^{a,b}| \leq \mathbb{1}_{[a-\varepsilon,a+\varepsilon]} + \mathbb{1}_{[b-\varepsilon,b+\varepsilon]}. \tag{3.15}
$$

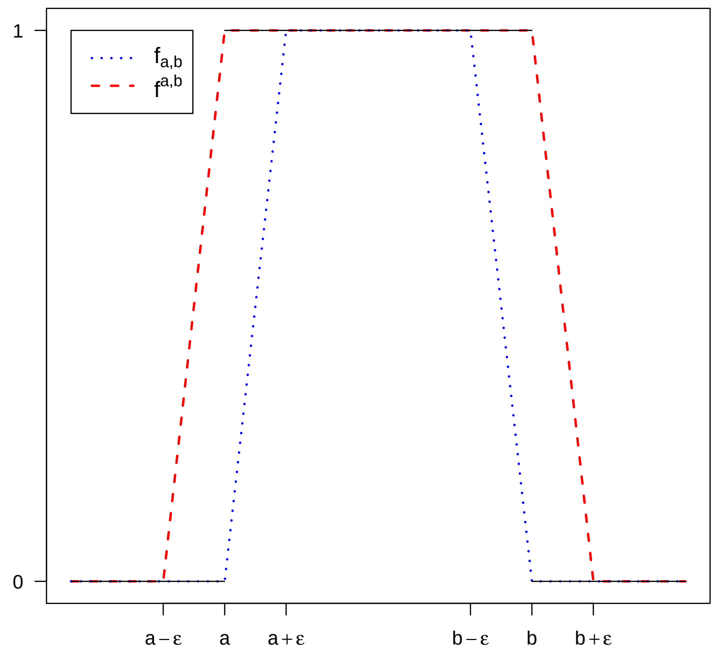

Abbildung 1: Graphen von $f_{a,b}$ und $f^{a,b}$

Hieraus folgt

$$\begin{aligned}
&\frac{1}{n}\sum_{j=1}^{n} f_{a,b}(\hat{z}_j) - \int f^{a,b}\mathbf{P}_Z \\
&\leq \frac{1}{n}\sum_{j=1}^{n} \mathbb{1}_{[a,b]}(\hat{z}_j) - \int \mathbb{1}_{[a,b]}\mathbf{P}_Z \qquad (3.16)\\
&\leq \frac{1}{n}\sum_{j=1}^{n} f^{a,b}(\hat{z}_j) - \int f_{a,b}\mathbf{P}_Z.
\end{aligned}$$

Gilt nun $\frac{1}{n}\sum_{j=1}^{n} \mathbb{1}_{[a,b]}(\hat{z}_j) - \int \mathbb{1}_{[a,b]}\mathbf{P}_Z \leq 0$, so folgt

$$\begin{aligned}
\Big|\frac{1}{n}\sum_{j=1}^{n}\mathbb{1}_{[a,b]}(\hat{z}_j) - \int \mathbb{1}_{[a,b]}\mathbf{P}_Z\Big| &\leq \left|\frac{1}{n}\sum_{j=1}^{n} f_{a,b}(\hat{z}_j) - \int f^{a,b}\mathbf{P}_Z\right| \\
&\leq \max\Big\{\Big|\frac{1}{n}\sum_{j=1}^{n} f_{a,b}(\hat{z}_j) - \int f^{a,b}\,\mathbf{P}_Z\Big|, \Big|\frac{1}{n}\sum_{j=1}^{n} f^{a,b}(\hat{z}_j) - \int f_{a,b}\,\mathbf{P}_Z\Big|\Big\}.
\end{aligned}$$

Gilt andererseits $\frac{1}{n}\sum_{j=1}^{n} \mathbb{1}_{[a,b]}(\hat{z}_j) - \int \mathbb{1}_{[a,b]}\mathbf{P}_Z \geq 0$, so folgt

$$\begin{aligned}
\Big|\frac{1}{n}\sum_{j=1}^{n}\mathbb{1}_{[a,b]}(\hat{z}_j) - \int \mathbb{1}_{[a,b]}\mathbf{P}_Z\Big| &\leq \left|\frac{1}{n}\sum_{j=1}^{n} f^{a,b}(\hat{z}_j) - \int f_{a,b}\mathbf{P}_Z\right| \\
&\leq \max\Big\{\Big|\frac{1}{n}\sum_{j=1}^{n} f_{a,b}(\hat{z}_j) - \int f^{a,b}\,\mathbf{P}_Z\Big|, \Big|\frac{1}{n}\sum_{j=1}^{n} f^{a,b}(\hat{z}_j) - \int f_{a,b}\,\mathbf{P}_Z\Big|\Big\}.
\end{aligned}$$

Weiterhin gilt

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{j=1}^{n} f_{a,b}(\hat{z}_j) - \int f^{a,b}\mathbf{P}_Z\right| \\
&\leq \left|\frac{1}{n}\sum_{j=1}^{n} f_{a,b}(\hat{z}_j) - \int f_{a,b}\mathbf{P}_Z\right| + \left|\int f^{a,b}\mathbf{P}_Z - \int f_{a,b}\mathbf{P}_Z\right|
\end{aligned}$$

und

$$\begin{aligned}
&\left|\frac{1}{n}\sum_{j=1}^{n} f^{a,b}(\hat{z}_j) - \int f_{a,b}\mathbf{P}_Z\right| x \\
&\leq \left|\frac{1}{n}\sum_{j=1}^{n} f^{a,b}(\hat{z}_j) - \int f^{a,b}\mathbf{P}_Z\right| + \left|\int f^{a,b}\mathbf{P}_Z - \int f_{a,b}\mathbf{P}_Z\right|.
\end{aligned}$$

Für $a, b, c \in \mathbb{R}^+$ gelten außerdem die elementaren Ungleichungen

$$\max\{a, b\} \leq a + b \text{ und } \max\{a + c, b + c|\} \leq c + \max\{a, b\}.$$

Mit $a = \left|\frac{1}{n}\sum_{j=1}^n f_{a,b}(\hat{z}_j) - \int f_{a,b}\mathbf{P}_Z\right|$, $b = \left|\frac{1}{n}\sum_{j=1}^n f^{a,b}(\hat{z}_j) - \int f^{a,b}\mathbf{P}_Z\right|$ und $c = \left|\int f^{a,b}\mathbf{P}_Z - \int f_{a,b}\mathbf{P}_Z\right|$ folgt also insgesamt

$$\left|\frac{1}{n}\sum_{j=1}^n \mathbb{1}_{[a,b]}(\hat{z}_j) - \int \mathbb{1}_{[a,b]}\mathbf{P}_Z\right| \leq \left|\frac{1}{n}\sum_{j=1}^n f_{a,b}(\hat{z}_j) - \int f_{a,b}\mathbf{P}_Z\right| \tag{3.17}$$

$$+\left|\frac{1}{n}\sum_{j=1}^n f^{a,b}(\hat{z}_j) - \int f^{a,b}\mathbf{P}_Z\right| + \left|\int f^{a,b}\mathbf{P}_Z - \int f_{a,b}\mathbf{P}_Z\right| \tag{3.18}$$

Es genügt also zu zeigen, dass die drei Summanden auf der linken Seite der Ungleichung fast sicher gegen 0 konvergieren.

Im zweiten Schritt approximieren wir $f^{a,b}$ und $f_{a,b}$ durch geeignete periodische Funktionen. Sei $\alpha_n = \frac{1}{2}n^{\frac{1}{40}}$, $d_n = \lfloor \hat{K}_n^{\frac{3}{2}} \rfloor$ und $s_n = \lfloor \hat{K}_n^{\frac{-1}{2}} \rfloor$, wobei $\hat{K}_n$ der zufällige Parameter aus Definition 2.4 der geschätzten Daten ist. Sei nun $\mathcal{G}_n = \mathcal{G}_n(\alpha_n, d_n, s_n)$ die Menge von trigonometischen Polynomen aus Definition 3.1. Aus den Voraussetzungen des Theorems folgt, dass $\varphi_{X^{(1)}}$ nullstellenfrei ist. Auch alle anderen Vorraussetzung von Lemma 3.8 sind erfüllt. Es gilt also $\lim_{n\to\infty} \hat{K}_n = \infty$ fast sicher. Dies impliziert, dass $s_n \to 0$ fast sicher gilt. Es gibt also mit Wahrscheinlichkeit 1 ein $n_0 \in \mathbb{N}$, so dass $[a-\varepsilon, b+\varepsilon] \subseteq [-\frac{\pi}{s_n}, \frac{\pi}{s_n}]$ für alle $n \geq n_0$ gilt. Für solche n seien $\overline{f_{a,b}}$ und $\overline{f^{a,b}}$ die $\frac{2\pi}{s_n}$-periodischen Fortsetzungen von $f_{a,b}$ und $f^{a,b}$. Da $f_{a,b}$ und $\overline{f_{a,b}}$ auf $[-\pi(s_n)^{-1}, \pi(s_n)^{-1}]$ übereinstimmen und $f_{a,b}$ außerhalb dieser Menge verschwindet, gilt

$$|f_{a,b}(x) - \overline{f_{a,b}}(x)| \leq \mathbb{1}_{\{|x|\geq \pi(s_n)^{-1}\}},$$

da $\overline{f_{a,b}}$ durch 1 beschränkt ist. Da auch $f^{a,b}$ und $\overline{f^{a,b}}$ auf $[-\pi(s_n)^{-1}, \pi(s_n)^{-1}]$ übereinstimmen und $f^{a,b}$ außerhalb dieser Menge verschwindet, gilt ebenso

$$|f^{a,b}(x) - \overline{f^{a,b}}(x)| \leq \mathbb{1}_{\{|x|\geq \pi(s_n)^{-1}\}},$$

da auch $\overline{f^{a,b}}$ durch 1 beschränkt ist. Das impliziert

$$\left|\int f_{a,b}\mathbf{P}_Z - \int \overline{f_{a,b}}\mathbf{P}_Z\right| \leq \mathbf{P}_Z\left[-\infty, \frac{-\pi}{s_n}\right] + \mathbf{P}_Z\left[\frac{\pi}{s_n}, \infty\right] \tag{3.19}$$

und

$$\left|\int f^{a,b}\mathbf{P}_Z - \int \overline{f^{a,b}}\mathbf{P}_Z\right| \leq \mathbf{P}_Z\left[-\infty, \frac{-\pi}{s_n}\right] + \mathbf{P}_Z\left[\frac{\pi}{s_n}, \infty\right].$$

Nun erarbeiten wir eine Abschätzung für die Summen. Wegen $\mathbb{1}_{\{|\hat{z}_j|>\pi(s_n)^{-1}\}} \leq \frac{s_n^2|\hat{z}_j|^2}{\pi^2}$ gilt

$$\begin{aligned}\left|\frac{1}{n}\sum_{j=1}^n f_{a,b}(\hat{z}_j) - \frac{1}{n}\sum_{j=1}^n \overline{f_{a,b}}(\hat{z}_j)\right| &\leq \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{|\hat{z}_j|>\pi(s_n)^{-1}\}} \\ &\leq \frac{s_n^2}{\pi^2}\frac{1}{n}\sum_{j=1}^n |\hat{z}_j|^2.\end{aligned} \tag{3.20}$$

Nach Definition erfüllen die geschätzten Daten $\hat{z}_j$ die Nebenbedingung (2.5), und es folgt

$$\Big|\frac{1}{n}\sum_{j=1}^{n} f_{a,b}(\hat{z}_j) - \frac{1}{n}\sum_{j=1}^{n}\overline{f_{a,b}}(\hat{z}_j)\Big| \leq \frac{s_n^2}{\pi^2}\frac{1}{n}\sum_{j=1}^{n}|X_j^{(1)}|^2 + 1. \tag{3.21}$$

Auf die gleiche Art und Weise lässt sich die analoge Abschätzung

$$\Big|\frac{1}{n}\sum_{j=1}^{n} f^{a,b}(\hat{z}_j) - \frac{1}{n}\sum_{j=1}^{n}\overline{f^{a,b}}(\hat{z}_j)\Big| \leq \frac{s_n^2}{\pi^2}\frac{1}{n}\sum_{j=1}^{n}|X_j^{(1)}|^2 + 1$$

zeigen. Nun approximieren wir die periodischen Fortsetzungen durch trigonometische Polynome aus $\mathcal{G}_n$. $\mathcal{G}_n$ enthält alle trigonometrischen Polynome mit Periode $\frac{2\pi}{s_n}$ vom Grad d_n, deren Koeffizienten durch α_n beschränkt sind. Für alle $f : \mathbb{R} \to \mathbb{R}$ und alle $j \in \mathbb{N}$ gelten die elementaren Abschätzungen

$$|a_j| = \left|\frac{s_n}{\pi}\int_0^{\frac{2\pi}{s_n}} f(x)\sin(js_nx)\,dx\right| \leq 2|f|_\infty$$

und

$$|b_j| = \left|\frac{s_n}{\pi}\int_0^{\frac{2\pi}{s_n}} f(x)\cos(js_nx)\,dx\right| \leq 2|f|_\infty.$$

Da $\overline{f_{a,b}}$ und $\overline{f^{a,b}}$ durch 1 beschränkt sind und für n groß genug $\alpha_n = n^{\frac{1}{40}} > 1$ gilt, sind also die Fouriersummen $S_{n,\overline{f_{a,b}}}$ von $\overline{f_{a,b}}$ und $S_{n,\overline{f^{a,b}}}$ von $\overline{f^{a,b}}$ für n groß genug in $\mathcal{G}_n$ enthalten. Ist f eine Lipschitz-stetige Funktion mit Lipschitz-Konstante λ und Periode $2s_n$, so ist $x \mapsto f\big(\frac{s_n}{\pi}\cdot x\big)$ eine 2π periodische und Lipschitz-stetige Funktion mit Lipschitz-Konstante $\frac{s_n}{\pi}\cdot\lambda$. Unter verwendung dieser Reskalierung können wir [48, Korrolar 1] anwenden. Also existiert eine konstante $C \in \mathbb{R}$, so dass

$$|S_{n,\overline{f_{a,b}}} - \overline{f_{a,b}}|_\infty \leq \frac{C}{\varepsilon}\frac{\log(d_n)}{s_nd_n} \tag{3.22}$$

und

$$|S_{n,\overline{f^{a,b}}} - \overline{f^{a,b}}|_\infty \leq \frac{C}{\varepsilon}\frac{\log(d_n)}{s_nd_n} \tag{3.23}$$

für alle $n \in \mathbb{N}$ gilt. Nun fügen wir die periodische Approximation und die trigonometische Approximation in (3.17) ein. Es gilt

$$\begin{aligned}
\Big|\frac{1}{n}\sum_{j=1}^{n}\mathbb{1}_{[a,b]}(\hat{z}_j) - \int\mathbb{1}_{[a,b]}\mathbf{P}_Z\Big| &\leq \left|\frac{1}{n}\sum_{j=1}^{n} f_{a,b}(\hat{z}_j) - \int f_{a,b}\mathbf{P}_Z\right| \\
&+ \left|\frac{1}{n}\sum_{j=1}^{n} f^{a,b}(\hat{z}_j) - \int f^{a,b}\mathbf{P}_Z\right| + \left|\int f^{a,b}\mathbf{P}_Z - \int f_{a,b}\mathbf{P}_Z\right| \\
&\leq \Big|\frac{1}{n}\sum_{j=1}^{n} f_{a,b}(\hat{z}_j) - \frac{1}{n}\sum_{j=1}^{n}\overline{f_{a,b}}(\hat{z}_j)\Big| + \Big|\frac{1}{n}\sum_{j=1}^{n}\overline{f_{a,b}}(\hat{z}_j) - \frac{1}{n}\sum_{j=1}^{n} S_{n,\overline{f_{a,b}}}(\hat{z}_j)\Big| \\
&+ \Big|\frac{1}{n}\sum_{j=1}^{n} S_{n,\overline{f_{a,b}}}(\hat{z}_j) - \int S_{n,\overline{f_{a,b}}}\mathbf{P}_Z\Big| + \Big|\int S_{n,\overline{f_{a,b}}}\,\mathbf{P}_Z - \int\overline{f_{a,b}}\,\mathbf{P}_Z\Big|
\end{aligned}$$

$$
\begin{aligned}
&+ \Big| \int \overline{f_{a,b}}\, \mathbf{P}_Z - \int f_{a,b}\, \mathbf{P}_Z \Big| \\
&+ \Big| \frac{1}{n} \sum_{j=1}^{n} f^{a,b}(\hat{z}_j) - \frac{1}{n} \sum_{j=1}^{n} \overline{f^{a,b}}(\hat{z}_j) \Big| + \Big| \frac{1}{n} \sum_{j=1}^{n} \overline{f^{a,b}}(\hat{z}_j) - \frac{1}{n} \sum_{j=1}^{n} S_{n,\overline{f^{a,b}}}(\hat{z}_j) \Big| \\
&+ \Big| \frac{1}{n} \sum_{j=1}^{n} S_{n,\overline{f^{a,b}}}(\hat{z}_j) - \int S_{n,\overline{f^{a,b}}} \mathbf{P}_Z \Big| + \Big| \int S_{n,\overline{f^{a,b}}}\, \mathbf{P}_Z - \int \overline{f^{a,b}}\, \mathbf{P}_Z \Big| \\
&+ \Big| \int \overline{f^{a,b}}\, \mathbf{P}_Z - \int f^{a,b}\, \mathbf{P}_Z \Big| \\
&+ \Big| \int f^{a,b}\, \mathbf{P}_Z - \int f_{a,b}\, \mathbf{P}_Z \Big| \\
= &\sum_{j=1}^{11} T_j. \qquad (3.24)
\end{aligned}
$$

Nun schätzen wir die Terme T_j mithilfe der Ungleichungen aus den vorherigen Schritten ab. Wir verwenden (3.21) für T_1 und T_6, (3.22) und (3.23) für T_2, T_4, T_7, und T_9, (3.19) für T_5 und T_10, sowie (3.15) für T_{11}. T_3 und T_8 schätzen wir wegen $S_{n,\overline{f_{a,b}}} \in \mathcal{G}_n$ und $S_{n,\overline{f^{a,b}}} \in \mathcal{G}_n$ durch

$$
T_3 = \Big| \frac{1}{n} \sum_{j=1}^{n} S_{n,\overline{f_{a,b}}}(\hat{z}_j) - \int S_{n,\overline{f_{a,b}}} \mathbf{P}_Z \Big| \le \sup_{g \in \mathcal{G}_m} \Big| \sum_{j=1}^{n} g(\hat{z}_j) - \int g\, \mathbf{P}_Z \Big|
$$

und

$$
T_8 = \Big| \int S_{n,\overline{f^{a,b}}}\, \mathbf{P}_Z - \int \overline{f^{a,b}}\, \mathbf{P}_Z \Big| \le \sup_{g \in \mathcal{G}_m} \Big| \sum_{j=1}^{n} g(\hat{z}_j) - \int g\, \mathbf{P}_Z \Big|
$$

ab. Insgesamt erhalten wir also die Abschätzung

$$
\begin{aligned}
\Big| \frac{1}{n} \sum_{j=1}^{n} \mathbb{1}_{[a,b]}(\hat{z}_j) - \int \mathbb{1}_{[a,b]}\, \mathbf{P}_Z \Big| &\le 2 \left(\frac{s_n}{\pi} \right)^2 \left(\frac{1}{n} \sum_{j=1}^{n} |X_j^{(1)}|^2 + 1 \right) \\
&+ 4 \frac{C}{\varepsilon} \frac{\log(d_n)}{s_n d_n} \\
&+ 2 \sup_{g \in \mathcal{G}_m} \Big| \sum_{j=1}^{n} g(\hat{z}_j) - \int g\, \mathbf{P}_Z \Big| \\
&+ 2\mathbf{P}_Z \left[-\infty, \frac{-\pi}{s_n} \right] + 2\mathbf{P}_Z \left[\frac{\pi}{s_n}, \infty \right] \\
&+ \mathbf{P}_Z[a-\varepsilon, a+\varepsilon] + \mathbf{P}_Z[b-\varepsilon, b+\varepsilon].
\end{aligned}
$$

Sei nun $\varepsilon_0 > 0$ beliebig. Nach Voraussetzung gilt $\mathbf{P}_Z[a] = \mathbf{P}_Z[b] = 0$. Da $\mathbf{P}_Z$ stetig von oben ist, gilt

$$
\mathbf{P}_Z[a-\varepsilon, a+\varepsilon] + \mathbf{P}_Z[b-\varepsilon, b+\varepsilon] \le \varepsilon_0
$$

für ε klein genug. Nach dem starken Gesetz der großen Zahlen konvergiert $\frac{1}{n}\sum_{j=1}^{n} |(X_j^{(1)})|^2 + 1$ fast sicher gegen $\mathbf{E}\{(X^{(1)})^2\} + 1$. Zusammen mit $s_n \to 0$ folgt

$$
\left(\frac{s_n}{\pi} \right)^2 \left(\frac{1}{n} \sum_{j=1}^{n} |X_i^{(1)}|^2 + 1 \right) \to 0 \quad \text{f.s.}
$$

Wegen $s_n \to 0$ fast sicher und

$$\lim_{n\to\infty} \frac{\log(d_n)}{s_n d_n} = \lim_{n\to\infty} \frac{\frac{3}{2}\log(\hat{K}_n)}{\hat{K}_n} = 0$$

fast sicher, konvergieren auch $4\frac{C}{\varepsilon}\frac{\log(d_n)}{s_n d_n}$ und $2\mathbf{P}_Z\left[-\infty, \frac{-\pi}{s_n}\right] + 2\mathbf{P}_Z\left[\frac{\pi}{s_n}, \infty\right]$ fast sicher gegen 0. Es bleibt also nur noch zu zeigen, dass auch

$$\sup_{g\in\mathcal{G}_m}\left|\sum_{j=1}^{n} g(\hat{z}_j) - \int g\,\mathbf{P}_Z\right|$$

mit Wahrscheinlichkeit 1 gegen 0 konvergiert. Hierzu wollen wir Lemma 3.2 verwenden. Um die benötigte Konvergenz der empirischen charakteristischen Funktion der geschätzten Daten $\hat{\varphi}_{\hat{z}_1^n}$ zu gewährleisten, verwenden wir Korollar 3.4. Nach Voraussetzung erfüllen die Zufallsvariablen X, Z und ε die Bedingungen von Definition 2.3 und die vierten Momente von Z und $\varepsilon_1, \ldots, \varepsilon_d$ sind endlich. Außerdem sind die geschätzten Daten $\hat{z}_1^n$ definiert wie in Definition 2.4 und es gilt $\frac{2\log(n)}{\sqrt{n}} \le \Delta_n$ sowie $u_n \le \Delta_n^2 n^{\frac{1}{10}}$ und $u_n \le n^{\alpha} \le n^{\frac{1}{20}}$ für alle $n \in \mathbb{N}$. Damit sind die Bedingungen von Korollar 3.4 mit $r_n = n^{\frac{1}{10}}$ erfüllt. Es folgt also, dass mit Wahrscheinlichkeit 1

$$n^{\frac{1}{10}} \sup_{u\in[0,\hat{K}_n]} |\hat{\varphi}_{\hat{z}_1^n}(u) - \varphi_Z(u)| \to 0$$

gilt. Nach Wahl von α_n, s_n und d_n gilt $s_n d_n \le \hat{K}_n$ und

$$2\alpha_n d_n \le 2\alpha_n \hat{K}_n^{\frac{3}{2}} \le n^{\frac{1}{40}} u_n^{\frac{3}{2}} \le n^{\frac{1}{40}} n^{\frac{3}{40}} \le n^{\frac{1}{10}} = r_n.$$

Also sind die Bedingungen von Lemma 3.2 mit $r'_n = 1$ erfüllt und es folgt

$$2\sup_{g\in\mathcal{G}_m}\left|\sum_{j=1}^{n} g(\hat{z}_j) - \int g\,\mathbf{P}_Z\right| \to 0 \quad \text{f.s.}$$

Damit ist der Beweis beendet. □

Beweis von Theorem 2.7. Da unter den Voraussetzungen des Theorems die Voraussetzungen von Lemma 3.8 ebenfalls erfüllt sind, gilt mit Wahrscheinlichkeit 1 $\hat{K}_n \to \infty$ und somit auch $s_n \to 0$ fast sicher. Wir betrachten für jedes $n \in \mathbb{N}$ nur Intervalle $[a, b]$, die in $\left[-\frac{\pi}{s_n} + 1, \frac{\pi}{s_n} - 1\right] \subset \left[-\frac{\pi}{s_n}, \frac{\pi}{s_n}\right]$ enthalten sind. Sei $\varepsilon' > 0$ so, dass $2\beta + \varepsilon' < 1$ gilt. Für n groß genug gilt dann

$$\varepsilon_n := \frac{1}{(\hat{K}_n)^{\beta+\varepsilon'}} < \frac{1}{2}.$$

Somit liegen die Träger der Funktionen

$$f_n^{a,b}(x) := \begin{cases} 1 \text{ für } x \in [a,b] \\ 0 \text{ für } x \in [a-\varepsilon_n, b+\varepsilon_n]^C \\ \frac{1}{\varepsilon_n}x - \frac{a}{\varepsilon_n} + 1 \text{ für } x \in [a-\varepsilon_n, a) \\ \frac{-1}{\varepsilon_n}x + \frac{b}{\varepsilon_n} + 1 \text{ für } x \in (b, b+\varepsilon_n] \end{cases} \tag{3.25}$$

und

$$f_{n,a,b}(x) := \begin{cases} 1 \text{ für } x \in [a+\varepsilon_n, b-\varepsilon_n] \\ 0 \text{ für } x \in [a,b]^C \\ \frac{1}{\varepsilon_n}x - \frac{a}{\varepsilon_n} \text{ für } x \in [a, a+\varepsilon_n) \\ \frac{-1}{\varepsilon_n}x + \frac{b}{\varepsilon_n} \text{ für } x \in (b-\varepsilon_n, b] \end{cases} \tag{3.26}$$

für alle $a, b \in \left[-\frac{\pi}{s_n}+1, \frac{\pi}{s_n}-1\right]$ in $\left[-\frac{\pi}{s_n}, \frac{\pi}{s_n}\right]$. Wir verfahren wie im Beweis von Theorem 2.6 und approximieren die Indikatorfunktion $\mathbb{1}_{[a,b]}$ zunächst durch die Lipschitz-stetigen Funktionen $f_{n,a,b}$ und $f_n^{a,b}$, approximieren diese durch ihre $\frac{2\pi}{s_n}$-periodischen Fortsetzungen $\overline{f_{n,a,b}}$ und $\overline{f_n^{a,b}}$, welche wiederum durch ihre Fouriersummen $S_{n,\overline{f_{n,a,b}}}$ und $S_{n,\overline{f_n^{a,b}}}$ aus $\mathcal{G}_n(\alpha_n, s_n, d_n)$ approximiert werden, wobei wir $\alpha_n = 1$, $s_n = \frac{1}{\sqrt{\lfloor \hat{K}_n \rfloor}}$ und $d_n = \lfloor \hat{K}_n \rfloor^{\frac{3}{2}}$ wählen.

Aus Gleichung (3.24) sowie den Abschätzungen (3.15) bis (3.23) folgt also für n groß genug

$$\begin{aligned} &\hat{K}_n^\beta \sup_{a,b \in [\frac{-\pi}{s_n}+1, \frac{\pi}{s_n}-1]} \left| \frac{1}{n} \sum_{j=1}^n \mathbb{1}_{[a,b]}(\hat{z}_j) - \int \mathbb{1}_{[a,b]} \, \mathbf{P}_Z \right| \\ &\leq \hat{K}_n^\beta \sup_{a,b \in [\frac{-\pi}{s_n}+1, \frac{\pi}{s_n}-1]} \Bigg[2 \sup_{g \in G_n} \Big| \sum_{j=1}^n g(\hat{z}_j) - \int g \, \mathbf{P}_Z \\ &\quad + C \frac{1}{\varepsilon_n} \frac{\log(d_n)}{s_n d_n} \\ &\quad + 2\mathbf{P}_Z\big[-\infty, \frac{-\pi}{s_n}\big] + 2\mathbf{P}_Z\big[\frac{\pi}{s_n}, \infty\big] \\ &\quad + \mathbf{P}_Z[a-\varepsilon_n, a+\varepsilon_n] + \mathbf{P}_Z[b-\varepsilon_n, b+\varepsilon_n] \\ &\quad + 2\left(\frac{s_n}{\pi}\right)^2 \left(\frac{1}{n} \sum_{j=1}^n (X_i^{(1)})^2 + 1\right) \Bigg] \\ &= \hat{K}_n^\beta \sup_{a,b \in [\frac{-\pi}{s_n}+1, \frac{\pi}{s_n}-1]} \big[T_{1,n} + T_{2,n} + T_{3,n} + T_{4,n} + T_{5,n} \big]. \end{aligned} \tag{3.27}$$

Wir zeigen nun, dass alle 5 Terme fast sicher gegen 0 konvergieren. Wir beginnen mit $T_{4,n}$, welcher als einziger tatsächlich von a und b abhängt. Da Z nach Voraussetzung eine beschränkte Dichte f besitzt, gilt

$$\begin{aligned} &\hat{K}_n^\beta \sup_{a,b \in [\frac{-\pi}{s_n}+1, \frac{\pi}{s_n}-1]} T_{4,n} \\ &\leq \hat{K}_n^\beta \sup_{a,b \in [\frac{-\pi}{s_n}+1, \frac{\pi}{s_n}-1]} [\mathbf{P}_Z[a-\varepsilon, a+\varepsilon] + \mathbf{P}_Z[b-\varepsilon, b+\varepsilon]] \\ &\leq \hat{K}_n^\beta 2 \sup_{a \in \mathbb{R}} \mathbf{P}_Z[a, a+2\varepsilon_n] \\ &\leq \hat{K}_n^\beta 4|f|_\infty \varepsilon_n = 4|f|_\infty \frac{\hat{K}_n^\beta}{\hat{K}_n^{\beta+\varepsilon'}} \\ &= 4|f|_\infty \hat{K}_n^{-\varepsilon'}. \end{aligned}$$

Da mit Wahrscheinlichkeit Eins $\hat{K}_n \to \infty$ gilt, folgt, dass auch mit Wahrscheinlichkeit Eins $\hat{K}_n^\beta \sup_{a,b\in[\frac{-\pi}{s_n}+1,\frac{\pi}{s_n}-1]} T_{4,n} \to 0$ gilt. Als Zweites wenden wir uns $T_{2,n}$ zu. Nach Wahl von β und ε' gilt $2\beta+\varepsilon' < 1$. Zusammen mit $s_n = \frac{1}{\sqrt{\lfloor \hat{K}_n \rfloor}}$ und $d_n = \lfloor \hat{K}_n \rfloor^{\frac{3}{2}}$ folgt

$$
\begin{aligned}
\hat{K}_n^\beta T_{2,n} &\le C\hat{K}_n^{2\beta+\varepsilon'} \log(\lfloor \hat{K}_n \rfloor^{\frac{3}{2}}) \frac{1}{\lfloor \hat{K}_n \rfloor} \\
&\le C\hat{K}_n^{2\beta+\varepsilon'} \log(\lfloor \hat{K}_n \rfloor^{\frac{3}{2}}) \frac{1}{\hat{K}_n} \frac{1}{1-\frac{1}{\hat{K}_n}}.
\end{aligned}
$$

Aus der fast sicheren Konvergenz von $\hat{K}_n$ gegen ∞ folgt somit, dass mit Wahrscheinlichkeit 1 auch $\hat{K}_n^\beta T_{2,n} \to 0$ gilt. Nun schätzen wir $T_{3,n}$ ab. Es gilt

$$
\mathbb{1}_{\{z>\frac{\pi}{s_n}\}} \le \left|\frac{zs_n}{\pi}\right|
$$

und

$$
\mathbb{1}_{\{z<\frac{-\pi}{s_n}\}} \le \left|\frac{zs_n}{-\pi}\right|.
$$

Dies impliziert

$$
\begin{aligned}
\hat{K}_n^\beta T_{3,n} &= \hat{K}_n^\beta \left(\mathbf{P}_Z[-\infty,\frac{-\pi}{s_n}] + 2\mathbf{P}_Z[\frac{\pi}{s_n},\infty]\right) \\
&\le \hat{K}_n^\beta 2\int \mathbb{1}_{\{Z<\frac{-\pi}{s_n}\}}\,\mathbf{P} + 2\int \mathbb{1}_{\{Z>\frac{\pi}{s_n}\}}\,\mathbf{P} \\
&\le 4\pi \hat{K}_n^\beta \int |Z| s_n \mathbf{P} = 4\pi \mathbf{E}\{|Z|\}\hat{K}_n^\beta s_n \\
&= 4\pi \mathbf{E}\{|Z|\}\hat{K}_n^\beta \frac{1}{\sqrt{\lfloor \hat{K}_n \rfloor}}.
\end{aligned}
$$

Nach Voraussetzung gilt $\beta < \frac{1}{2}$ und $\hat{K}_n$ konvergiert fast sicher gegen ∞. Somit konvergiert auch $\hat{K}_n^\beta T_{3,n}$ für $n \to \infty$ fast sicher gegen 0, da nach Voraussetzung $\mathbf{E}\{|Z|\} < \infty$ gilt.

Um die Konvergenz von $\hat{K}_n^\beta T_{5,n}$ zu zeigen, verwenden wir das starke Gesetz der großen Zahlen. Da die X_j unabhängig und identisch verteilt sind, konvergiert die Folge $\frac{1}{n}\sum_{j=1}^n (X_j^{(1)})^2 + 1$ nach dem starken Gesetz der großen Zahlen fast sicher gegen $\mathbf{E}|X^{(1))}|^2+1$. Nach Gleichung (3.8) gilt $\mathbf{E}|X^{(1))}|^2+1 < \infty$. Daher ist die Folge $\frac{1}{n}\sum_{j=1}^n (X_j^{(1)})^2 + 1$ mit Wahrscheinlichkeit 1 beschränkt. Mit Wahrscheinlichkeit 1 gilt außerdem $\hat{K}_n^\beta s_n = \hat{K}_n^\beta \frac{1}{\sqrt{\lfloor \hat{K}_n \rfloor}} \to 0$. Also folgt auch $\hat{K}_n^\beta T_{5,n} \to 0$ fast sicher. Nun müssen wir nur noch $T_{1,n}$ verarzten. Da die Vorraussetzungen von Lemma 3.3 erfüllt sind, gilt

$$
\frac{\Delta_n^2}{\hat{K}_n}\frac{\sqrt[4]{n}}{\log(n)} \sup_{u\in[0,\hat{K}_n]} |\hat{\varphi}_{\hat{z}_1^n}(u) - \varphi_Z(u)| \to 0 \quad \text{f.s.}
$$

Nun wollen wir Lemma 3.2 mit $r_n = \frac{\Delta_n^2}{\hat{K}_n}\frac{\sqrt[4]{n}}{\log(n)}$ und $r_n' = \hat{K}_n^\beta$ verwenden. Nach Wahl von s_n und d_n gilt

$$
s_n d_n \le \hat{K}_n.
$$

Da u_n und Δ_n die Bedingung (2.6) erfüllen, gilt auch

$$4\hat{K}_n^{\frac{5}{2}+\beta}\Delta_n^{-2} \leq 4u_n^{\frac{5}{2}+\beta}\Delta_n^{-2} \leq \frac{\sqrt[4]{n}}{\log(n)},$$

was

$$4\alpha_n d_n \hat{K}_n^\beta \leq \frac{\Delta^2}{\hat{K}_n}\frac{\sqrt[4]{n}}{\log(n)}$$

impliziert. Also folgt aus Lemma 3.2, dass mit Wahrscheinlichkeit 1

$$\hat{K}_n^\beta T_{1,n} \to 0$$

gilt. Insgesamt gilt also

$$\hat{K}_n^\beta \sup_{a,b\in[\frac{-\pi}{s_n}+1,\frac{\pi}{s_n}-1]} \left|\frac{1}{n}\sum_{j=1}^n \mathbb{1}_{[a,b]}(\hat{z}_j) - \int \mathbb{1}_{[a,b]}\,\mathbf{P}_Z\right| \to 0 \quad \text{f.s.}$$

Damit ist der Beweis beendet. □

3.3 Beweis der Konvergenz des Schätzers der Verteilungsfunktion

Beweis von Theorem 2.9. Unter den Annahmen des Theorems sind auch die Voraussetzungen von Theorem 2.6 erfüllt, welches wir im Beweis verwenden werden. Sei $t \in \mathbb{R}$ beliebig und sei $t' \in \mathbb{R}$ mit $t' < 0$ und $t' < t$. Wegen $\lim_{n\to\infty} s_n = 0$ gibt es ein $n_0 \in \mathbb{N}$, so dass $\frac{-\pi}{s_n} + 1 < t'$ und $\frac{\pi}{s_n} > t$ für alle $n \geq n_0$ gilt. Für solche n folgt

$$\begin{aligned}
|F(t) - \hat{F}_n(t)| &= |\mu\left((-\infty,t]\right) - \hat{\mu}_n\left([\frac{-\pi}{s_n}+1,t]\right)| \\
&\leq |\mu\left((-\infty,t]\right) - \mu\left([t',t]\right)| + |\mu\left([t',t]\right) - \hat{\mu}_n\left([t',t)\right]| \\
&\quad + |\hat{\mu}_n\left([t',t]\right) - \hat{\mu}_n\left([\frac{-\pi}{s_n}+1,t]\right)| \\
&= \mu\left((-\infty,t')\right) + \left|\mu\left([t',t]\right) - \hat{\mu}_n\left([t',t]\right)\right| + \hat{\mu}\left([\frac{-\pi}{s_n}+1,t')\right) \\
&= T_{1,n} + T_{2.n} + T_{3,n}.
\end{aligned}$$

Es genügt also zu zeigen, das $T_{1,n}$, $T_{2,n}$ und $T_{3,n}$ für $n \to \infty$ jeweils fast sicher gegen 0 konvergieren. Sei dazu $\varepsilon > 0$ beliebig. Die geschätzten Daten $\hat{z}_j$ erfüllen die Nebenbedingung (2.5) aus Definition 2.4, daher gilt

$$\begin{aligned}
T_{3,n} &= \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{[\frac{-\pi}{s_n}+1,t')}(\hat{z}_j) \leq \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{\hat{z}_j \leq t'\}} \\
&\leq \frac{1}{n}\sum_{j=1}^n \mathbb{1}_{\{(t')^2 \leq \hat{z}_j^2\}} \leq \frac{1}{(t')^2}\frac{1}{n}\sum_{j=1}^n \hat{z}_j^2 \\
&\leq \frac{1}{(t')^2}\left[\frac{1}{n}\sum_{j=1}^n |X_j^{(1)}|^2 + 1\right].
\end{aligned} \tag{3.28}$$

Nach dem starken Gesetz der großen Zahlen konvergiert $\frac{1}{n}\sum_{j=1}^{n}|X_j^{(1)}|^2+1$ mit Wahrscheinlichkeit 1 und ist daher mit Wahrscheinlichkeit 1 durch eine zufällige Konstante $\hat{C}$ beschränkt. Die Abschätzung (3.28) gilt für n groß genug für alle t' mit $t'<0$ und $t'<t$, und somit auch für dasjenige, zufällige $\hat{t}'$, für welches

$$\frac{1}{(t')^2}\left[\frac{1}{n}\sum_{j=1}^{n}|X_j^{(1)}|^2+1\right]\leq\varepsilon$$

für alle $n\in\mathbb{N}$ und alle $t'\leq\hat{t}'$ gilt. Für ein ausreichend kleines dieser t' gilt auch

$$T_{1,n}=\mu\left((-\infty,t')\right)\leq\varepsilon$$

für alle $n\in\mathbb{N}$. Nach Theorem 2.6 gilt dann

$$T_{2,n}=\Big|\mu\left([t',t]\right)-\hat{\mu}_n\left([t',t]\right)\Big|\to 0$$

fast sicher. Damit ist der Beweis beendet. □

Beweis von Theorem 2.10. Zunächst eine Vorüberlegung. Ist F die Verteilungsfunktion von Z, und gilt $t\in(-\infty,\frac{-\pi}{s_n}+1]$, so folgt aufgrund der Monotonie von F, und da für solche t nach Definition $\hat{F}_n(t)=0$ gilt, die Abschätzung

$$\begin{aligned}|F(t)-\hat{F}_n(t)|&\leq F\left(\frac{-\pi}{s_n}+1\right)=\int_{-\infty}^{\frac{-\pi}{s_n}+1}1\mathbf{P}_Z\\&\leq\int_{-\infty}^{\frac{-\pi}{s_n}+1}\frac{|z|}{|\frac{-\pi}{s_n}+1|}\mathbf{P}_Z(dz)\\&\leq\frac{s_n}{|s_n-\pi|}\mathbf{E}\{|Z|\}.\end{aligned}\tag{3.29}$$

Für $t\geq\frac{\pi}{s_n}-1$ erhalten wir analog

$$|F(t)-\hat{F}_n(t)|\leq 1-F\left(\frac{\pi}{s_n}-1\right)=\int_{\frac{\pi}{s_n}-1}^{\infty}1\mathbf{P}_Z\tag{3.30}$$

$$\leq\frac{s_n}{|\pi-s_n|}\mathbf{E}\{|Z|\}.\tag{3.31}$$

Also gilt

$$\begin{aligned}&\sup_{t\ \in\mathbb{R}}\Big|F(t)-\hat{F}_n(t)\Big|\\&=\max\Big\{\sup_{t\leq\frac{-\pi}{s_n}+1}\Big|F(t)-\hat{F}_n(t)\Big|,\sup_{t\in(\frac{-\pi}{s_n}+1,\frac{\pi}{s_n}-1)}\Big|F(t)-\hat{F}_n(t)\Big|,\\&\quad\sup_{t\geq\frac{\pi}{s_n}-1}\Big|F(t)-\hat{F}_n(t)\Big|\Big\}\\&=\max\left\{\sup_{t\leq\frac{-\pi}{s_n}+1}|F(t)|,\sup_{t\in(\frac{-\pi}{s_n}+1,\frac{\pi}{s_n}-1)}\Big|F(t)-\hat{F}_n(t)\Big|,\sup_{t\geq\frac{\pi}{s_n}-1}|1-F(t)|\right\}\end{aligned}$$

$$
\begin{aligned}
&\leq \sup_{t \leq \frac{-\pi}{s_n}+1} |F(t)| + \sup_{t \in (\frac{-\pi}{s_n}+1, \frac{\pi}{s_n}-1)} \left| F(t) - \hat{F}_n(t) \right| + \sup_{t \geq \frac{\pi}{s_n}-1} |1 - F(t)| \\
&\leq \frac{2s_n}{|s_n - \pi|} \mathbf{E}\{|Z|\} + \sup_{t \in [\frac{-\pi}{s_n}+1, \frac{\pi}{s_n}-1]} \left| F(t) - \hat{F}_n(t) \right|,
\end{aligned}
$$

wobei wir im letzten Schritt die Abschätzungen (3.29) und (3.30) verwendet haben. Da unter den Annahmen des Theorems die Voraussetzungen von Lemma 3.8 erfüllt sind, konvergiert $s_n = \frac{1}{\sqrt{\lfloor \hat{K}_n \rfloor}}$ fast sicher gegen 0. Wir müssen also nur noch zeigen, dass auch $\sup_{t \in (\frac{-\pi}{s_n}+1, \frac{\pi}{s_n}-1)} |F(t) - \hat{F}_n(t)|$ mit Wahrscheinlichkeit 1 gegen 0 konvergiert. Es gilt

$$
\begin{aligned}
&\sup_{t \in [\frac{-\pi}{s_n}+1, \frac{\pi}{s_n}-1]} \left| \hat{F}_n(t) - F(t) \right| \\
&\leq \sup_{t \in [\frac{-\pi}{s_n}+1, \frac{\pi}{s_n}-1]} \left| \hat{F}_n(t) - (F(t) - F(-\frac{\pi}{s_n}+1)) \right| + \left| F(-\frac{\pi}{s_n}+1) \right| \\
&= \sup_{t \in [\frac{-\pi}{s_n}+1, \frac{\pi}{s_n}-1]} \left| \hat{\mu}_n([-\frac{\pi}{s_n}+1, t]) - \mu([-\frac{\pi}{s_n}+1, t]) \right| + \left| F(-\frac{\pi}{s_n}+1) \right| \\
&\leq \sup_{a,b \in [-\frac{\pi}{s_n}+1, \frac{\pi}{s_n}-1]} |\hat{\mu}_n([a,b]) - \mu([a,b])| + \left| F(-\frac{\pi}{s_n}+1) \right|.
\end{aligned}
$$

Unter den Vorraussetzungen des Theorems sind auch die Voraussetzungen von Theorem 2.7 mit $\beta = 0$ erfüllt. Also lässt sich Theorem 2.7 anwenden und es folgt

$$
\sup_{a,b \in [-\frac{\pi}{s_n}+1, \frac{\pi}{s_n}-1]} |\hat{\mu}_n([a,b]) - \mu([a,b])| \to 0 \quad \text{f.s.}
$$

Für den zweiten Summanden verwenden wir erneut die Abschätzung (3.29). Wegen

$$
\left| F(-\frac{\pi}{s_n}+1) \right| \leq \frac{s_n}{|s_n - \pi|} \mathbf{E}\{|Z|\}
$$

konvergiert auch $|F(-\frac{\pi}{s_n}+1)|$ fast sicher gegen 0. Damit ist der Beweis beendet. □

3.4 Beweis der Konvergenz des Dichteschätzers

Bevor wir uns der Konvergenz des Schätzers in L_1 annehmen, benötigen wir zunächst ein Hilfsresultat zur punktweisen Konvergenz des Dichteschätzers.

Proposition 3.10. *Sei $\beta \in \mathbb{R}$ mit $0 \leq \beta < \frac{1}{2}$ und sei $n_0 \in \mathbb{N}$. Sei Δ_n eine reelle Folge mit $\frac{2\log(n)}{\sqrt{n}} \leq \Delta_n$ für alle n in $\mathbb{N}$ mit $n \geq n_0$ und $\lim_{n\to\infty} \Delta_n = 0$. Sei u_n eine reelle Folge mit $\lim_{n\to\infty} u_n = \infty$ und $u_n \leq n^\alpha$ für ein $\alpha \in \mathbb{R}$ und alle $n \in \mathbb{N}$ mit $n \geq n_0$. Seien u_n und Δ_n außerdem so, dass*

$$
4u_n^{\frac{5}{2}+\beta} \frac{1}{\Delta_n^2} \leq \frac{\sqrt[4]{n}}{\log(n)} \tag{3.32}
$$

für alle $n \in \mathbb{N}$ mit $n \geq n_0$ gilt. Seien X, Z und ε wie in Definition 2.3 und seien außerdem die 4-ten Momente von Z und $\varepsilon_1, \ldots, \varepsilon_d$ endlich und Z besitze eine beschränkte Dichte

f. Seien die geschätzten Daten $\hat{z}_1^n$ und der Parameter $\hat{K}_n$ ausgehend von $X_1, \dots, X_n$ sowie Δ_n und u_n definiert wie in Definition 2.4 und sei $\hat{\mu}_n$ definiert wie in Definition 2.5. Sei h_n eine Folge von Bandbreiten mit $h_n \geq \hat{K}_n^{-\beta}$ und sei $\hat{f}_n$ definiert wie in (2.8). Ist dann $(\Omega, \mathcal{B}, \mathbf{P})$ der zugrundeliegende Wahrscheinlichkeitsraum, so existiert eine Menge $\Omega' \subseteq \Omega$ mit $\mathbf{P}[\Omega'] = 1$ und eine Menge $A \subseteq \mathbb{R}$ mit $\lambda(A^c) = 0$, so dass

$$\hat{f}_n(t, \omega) \to f(t)$$

für alle $(t, \omega) \in (\Omega') \times A$ gilt.

Bemerkung. Proposition 3.10 impliziert, dass $\hat{f}_n(t)$ für lebesgue-fast-alle $t \in \mathbb{R}$ fast sicher gegen f konvergiert.

Beweis. Wir wählen $s_n = \frac{1}{\sqrt{\lfloor \hat{K}_n \rfloor}}$. Dann sind unter den Voraussetzungen des Theorems auch die Voraussetzungen von Theorem 2.7 erfüllt. Sei nun $B_{h_n}(t)$ die Kugel mit Radius h_n um t. Da f eine Dichte ist, und daher auch integrabel ist, folgt wegen $h_n \to 0$ aus dem Differentationstheorem von Lebesgue (siehe beispielsweise [49, Theorem 5.6.2]), dass

$$f(t) = \lim_{n\to\infty} \frac{1}{\lambda(B_{h_n}(t))} \int_{B_{h_n}(t)} f(y)\, dy = \lim_{n\to\infty} \frac{\mu(B_{h_n}(t))}{\lambda(B_{h_n}(t))}$$

für lebesgue-fast-alle $t \in \mathbb{R}$ gilt. Da die Voraussetzungen von Lemmma 3.8 erfüllt sind, gilt für ein solches t für n groß genug $B_{h_n}(t) \subseteq [\frac{-\pi}{s_n} + 1, \frac{\pi}{s_n} - 1]$. Wegen $\lambda(B_{h_n}(t)) \geq 2\hat{K}_n^{-\beta}$ folgt daraus, unter Verwendung von Theorem 2.7, dass

$$\begin{aligned}
&\lim_{n\to\infty} \left| \frac{\mu(B_{h_n}(t)) - \hat{\mu}_n(B_{h_n}(t))}{\lambda(B_{h_n}(t)}) \right| \\
&\leq 2 \lim_{n\to\infty} \hat{K}_n^{\beta} \sup_{a,b \in [\frac{-\pi}{s_n}+1, \frac{\pi}{s_n}-1]} |\mu([a,b]) - \hat{\mu}_n([a,b])| = 0
\end{aligned}$$

mit Wahrscheinlichkeit 1 gilt. Da die rechte Seite nicht mehr von t abhängt, ist auch die Nullmenge, auf der $\left| \frac{\mu(B_{h_n}(t)) - \hat{\mu}_n(B_{h_n}(t))}{\lambda B_{h_n}(t)} \right|$ nicht konvergiert, für alle $t \in \mathbb{R}$ identisch. Insgesamt folgt also

$$\begin{aligned}
f(t) &= \lim_{n\to\infty} \frac{\mu(B_{h_n}(t))}{\lambda(B_{h_n}(t))} = \lim_{n\to\infty} \frac{\mu(B_{h_n}(t)) - \hat{\mu}_n(B_{h_n}(t))}{\lambda(B_{h_n}(t))} + \frac{\hat{\mu}_n(B_{h_n}(t))}{\lambda(B_{h_n}(t))} \\
&= \lim_{n\to\infty} \frac{\hat{\mu}_n(B_{h_n}(t))}{\lambda(B_{h_n}(t))} = \lim_{n\to\infty} \hat{f}_n(t) \quad \text{f.s.}
\end{aligned}$$

für alle Lebesguepunkte von f und damit auch lebesgue-fast-überall. □

Bemerkung. Aus dem Differentationstheorem von Lebesgue folgt ebenfalls, dass die Menge der Lebesguepunkte L_f, auf welcher $\hat{f}_n(t) \to f(t)$ gilt, nicht von der Wahl des Repräsentanten abhängt.

Beweis von Theorem 2.12. Die Voraussetzungen von Theorem 2.12 stimmen mit den Voraussetzungen von Proposition 3.10 überein. Wir können also die punktweise Konvergenz von $\hat{f}_n$ ausnutzen um die L_1-Konvergenz des Dichteschätzers zu beweisen. Da f und $\hat{f}_n$ beides Dichten sind, gilt nach dem Lemma von Scheffé

$$\int |f(t) - \hat{f}_n(t)|\, dt = 2 \int \left(f(t) - \hat{f}_n(t)\right)_+ dt.$$

Wobei $\big(f(t)-\hat{f}_n(t)\big)_+$ den Positivteil von $f(t)-\hat{f}_n(t)$ bezeichnet. Nun gilt $\big(f(t)-\hat{f}_n(t)\big)_+ \leq f(t)$ für alle $t \in \mathbb{R}$ und alle $n \in \mathbb{N}$. Da f eine Dichte ist, folgt, dass f eine integrierbare Majorante für $\big(f(t)-\hat{f}_n(t)\big)_+$ ist. Mit dominierter Konvergenz und unter Verwendung von Proposition 3.10 folgt

$$\begin{aligned}
&\lim_{n\to\infty}\int |f(t)-\hat{f}_n(t)|\,dt \leq 2\lim_{n\to\infty}\int \big(f(t)-\hat{f}_n(t)\big)_+\,dt \\
&= 2\int \lim_{n\to\infty}\big(f(t)-\hat{f}_n(t)\big)_+\,dt \\
&\leq 2\int \lim_{n\to\infty}|f(t)-\hat{f}_n(t)|\,dt \\
&= 0 \quad \text{f.s.}
\end{aligned} \tag{3.33}$$

Damit ist die Behauptung gezeigt. □

Lemma 3.11. *Sei $\beta \in \mathbb{R}$ mit $0 \leq \beta < \frac{1}{2}$ und sei $n_0 \in \mathbb{N}$. Sei Δ_n eine reelle Folge mit $\frac{2\log(n)}{\sqrt{n}} \leq \Delta_n$ für alle n in $\mathbb{N}$ mit $n \geq n_0$ und $\lim_{n\to\infty}\Delta_n = 0$. Sei u_n eine reelle Folge mit $\lim_{n\to\infty}u_n = \infty$ und $u_n \leq n^\alpha$ für ein $\alpha \in \mathbb{R}$ und alle $n \in \mathbb{N}$ mit $n \geq n_0$. Seien u_n und Δ_n außerdem so, dass*

$$4u_n^{\frac{5}{2}+\beta}\frac{1}{\Delta_n^2} \leq \frac{\sqrt[4]{n}}{\log(n)} \tag{3.34}$$

für alle $n \in \mathbb{N}$ mit $n \geq n_0$ gilt. Seien X, Z und ε wie in Definition 2.3 und seien außerdem die 4-ten Momente von Z und $\varepsilon_1,\ldots,\varepsilon_d$ endlich und Z besitze eine beschränkte Dichte f. Sei f außerdem (p,C)-glatt für ein $p \leq 2$. Seien die geschätzten Daten $\hat{z}_1^n$ und der Parameter $\hat{K}_n$ ausgehend von $X_1,\ldots,X_n$ sowie Δ_n und u_n definiert wie in Definition 2.4 und sei $\hat{\mu}_n$ definiert wie in Definition 2.5. Sei $\gamma < \frac{p}{1+p}\beta$ und sei $h_n = \hat{K}_n^{\gamma-\beta}$. Sei $\hat{f}_n$ definiert wie in (2.8). Ist dann $(\Omega,\mathcal{B},\mathbf{P})$ der zugrundeliegende Wahrscheinlichkeitsraum, so existiert eine Menge $\Omega' \subseteq \Omega$ mit $\mathbf{P}[\Omega'] = 1$, so dass

$$\hat{K}_n^\gamma|\hat{f}_n(t)-f(t)| \to 0$$

für alle $\omega \in \Omega'$ und alle $t \in \mathbb{R}$ gilt.

Beweis. Der Beweis verläuft analog zum Beweis von Proposition 3.10. Es gilt

$$\begin{aligned}
\hat{K}_n^\gamma\left|f(t)-\hat{f}_n(t)\right| &\leq \hat{K}_n^\gamma\left|f(t)-\frac{\mu(B_{h_n}(t))}{\lambda(B_{h_n}(t))}\right| + \hat{K}_n^\gamma\left|\frac{\mu(B_{h_n}(t))}{\lambda(B_{h_n}(t))}-\frac{\hat{\mu}_n(B_{h_n}(t))}{\lambda(B_{h_n}(t))}\right| \\
&\leq \hat{K}_n^\gamma T_{1,n} + \hat{K}_n^\gamma T_{2,n}.
\end{aligned}$$

Wir zeigen zuerst, dass $\hat{K}_n^\gamma T_{1,n}$ gegen 0 konvergiert und unterscheiden dazu die Fälle $p \leq 1$ und $1 < p \leq 2$. Wir beginnen mit $p \leq 1$. Es gilt

$$\begin{aligned}
\hat{K}_n^\gamma T_{1,n} &= \left|f(t)-\frac{\int_{B_{h_n}(t)}f(x)\,dx}{2h_n}\right| \\
&= \hat{K}_n^\gamma\frac{1}{2h_n}\left|\int_{B_{h_n}(t)} f(t)-f(x)\,dx\right| \\
&\leq \hat{K}_n^\gamma \max_{x\in B_{h_n}(t)}|f(x)-f(t)|
\end{aligned}$$

$$
\begin{aligned}
&\leq C\hat{K}_n^{\gamma} h_n^p = C\hat{K}_n^{\gamma}\hat{K}_n^{p\gamma-p\beta} \\
&= C\hat{K}_n^{(1+p)\gamma-p\beta}.
\end{aligned}
$$

Nach Voraussetzung gilt $\hat{K}_n \to \infty$ fast sicher und $(1+p)\gamma < p\beta$. Also folgt, dass auch $\hat{K}_n^{\gamma} T_{1,n}$ fast sicher gegen 0 konvergiert. Nun zum Fall $1 < p \leq 2$. Hierzu zunächst eine Vorüberlegung. Für $t \in \mathbb{R}$ sei

$$
g(h) := f(t+h) - f(t) - f'(t)h.
$$

Da die Ableitung von f nach Vorraussetzung $(p-1, C)$-glatt ist, ist die Ableitung von g ebenfalls $(p-1, C)$-glatt. Weiterhin gilt $g(0) = g'(0) = 0$, woraus

$$
\begin{aligned}
|f(t+h) - f(t) - f'(t)h| = |g(h)| &= \left| \int_0^h g'(x)\, dx \right| \\
&\leq |h| \int_0^1 |g'(hx)|\, dx \\
&\leq |h| \int_0^1 C|h|^{p-1}x^{p-1}\, dx \\
&= \frac{C|h|^p}{p} \leq C|h|^p
\end{aligned}
\tag{3.35}
$$

folgt. Mit $x = t + h$ folgt nun

$$
\begin{aligned}
\hat{K}_n^{\gamma} T_{1,n} &= \hat{K}_n^{\gamma} \frac{1}{2h_n} \left| \int_{B_{h_n}(t)} f(t) - f(x)\, dx \right| \\
&= \hat{K}_n^{\gamma} \frac{1}{2h_n} \left| \int_{B_{h_n}(0)} f(t) - f(t+h)\, dh \right|.
\end{aligned}
$$

Zusammen mit

$$
\int_{B_{h_n}(0)} f'(t)h\, dh = 0
$$

und der Abschätzung (3.35) folgt also

$$
\begin{aligned}
\hat{K}_n^{\gamma} T_{1,n} &= \hat{K}_n^{\gamma} \frac{1}{2h_n} \left| \int_{B_{h_n}(0)} f'(t)h + f(t) - f(t+h)\, dh \right| \\
&\leq \hat{K}_n^{\gamma} \max_{h \in B_{h_n}(0)} \left| f'(t)h + f(t) - f(t+h) \right| \\
&\leq C\hat{K}_n^{\gamma} |h_n|^p = \hat{K}_n^{(1+p)\gamma-p\beta}
\end{aligned}
$$

für h_n klein genug. Wegen $\hat{K}_n \to \infty$ fast sicher und $(1+p)\gamma < p\beta$ folgt also, dass auch für $1 < p \leq 2$ mit Wahrscheinlichkeit 1

$$
\hat{K}_n^{\gamma} T_{1,n} \to 0
$$

gilt.

Nun zu $T_{2,n}$. Nach Voraussetzung gilt

$$
\hat{K}_n^{\gamma} T_{2,n} \leq \frac{\hat{K}_n^{\gamma}}{2h_n} |\mu(B_{h_n}(t)) - \hat{\mu}_n(B_{h_n}(t))|
$$

$$
\begin{aligned}
&\leq \frac{1}{2}\hat{K}_n^{\gamma}\hat{K}_n^{\beta-\gamma} \sup_{a,b\in[\frac{-\pi}{s_n}+1,\frac{\pi}{s_n}-1]} |\mu([a,b]) - \hat{\mu}_n([a,b])| \\
&= \frac{1}{2}\hat{K}_n^{\beta} \sup_{a,b\in[\frac{-\pi}{s_n}+1,\frac{\pi}{s_n}-1]} |\mu([a,b]) - \hat{\mu}_n([a,b])|
\end{aligned}
$$

für n groß genug. Unter den Voraussetzungen der Proposition sind auch die Voraussetzungen von Theorem 2.7 erfüllt. Also konvergiert die linke Seite, die nicht mehr von t abhängt, fast sicher gegen 0. Damit konvergiert auch $\hat{K}_n^{\gamma}T_{2,n}$ fast sicher gegen 0 und der Beweis ist beendet. □

Beweis von Theorem 2.14. Unter den Annahmen von Theorem 2.14 sind die Voraussetzungen von Lemma 3.11 erfüllt. Die Behauptung folgt völlig analog zum Beweis von Theorem 2.12 unter Verwendung von Lemma 3.11 anstelle von Proposition 3.10. □

3.5 Beweis der deterministischen Konvergenzgeschwindigkeit

Beweis von Proposition 2.15. Für $x \in [0, A_n]$ gilt

$$
\begin{aligned}
\inf_{u\in[0,x]} \left|\hat{\varphi}_{X_1^n}(u,0,\dots,0)\right| &= \inf_{u\in[0,x]} \left|\hat{\varphi}_{X_1^n}(u,0,\dots,0) - \varphi_{X^{(1)}}(u) + \varphi_{X^{(1)}}(u)\right| \\
&\geq \inf_{u\in[0,A_n]} \left|\hat{\varphi}_{X_1^n}(u,0,\dots,0) - \varphi_{X^{(1)}}(u) + \varphi_{X^{(1)}}(u)\right| \\
&\geq \inf_{u\in[0,A_n]} \left\{|\varphi_{X^{(1)}}(u)| - \left|\hat{\varphi}_{X_1^n}(u,0,\dots,0) - \varphi_{X^{(1)}}(u)\right|\right\} \\
&\geq \inf_{u\in[0,A_n]} |\varphi_{X^{(1)}}(u)| - \sup_{u\in[0,A_n]} \left|\varphi_{X^{(1)}}(u) - \hat{\varphi}_{X_1^n}(u,0,\dots,0)\right|.
\end{aligned}
$$

Nach Voraussetzung gilt $\inf_{u\in[0,A_n]} |\varphi_{X^{(1)}}(u)| \geq (1+\varepsilon)\Delta_n$. Nach Lemma 3.5 existiert mit Wahrscheinlichkeit 1 ein zufälliges $C \in \mathbb{R}$, so dass

$$
\sup_{u\in[0,A_n]} \left|\varphi_{X^{(1)}}(u) - \hat{\varphi}_{X_1^n}(u,0,\dots,0)\right| \leq C\frac{\log(n)}{\sqrt{n}}
$$

für n groß genug gilt. Damit folgt, dass mit Wahrscheinlichkeit 1

$$
\inf_{u\in[0,x]} \left|\hat{\varphi}_{X_1^n}(u,0,\dots,0)\right| \geq (1+\varepsilon)\Delta_n - C\frac{\log(n)}{\sqrt{n}}
$$

für n groß genug gilt. Nun zeigen wir, dass sich die rechte Seite für n genügend groß nach unten durch Δ_n abschätzen lässt. nach Voraussetzung gilt $\Delta_n \geq 0$ für n groß genug. Für solche n ist die Ungleichung

$$
(1+\varepsilon)\Delta_n - C\frac{\log(n)}{\sqrt{n}} \geq \Delta_n
$$

offensichtlich zu

$$
(1+\varepsilon) - C\frac{\log(n)}{\sqrt{n}}\frac{1}{\Delta_n} \geq 1 \tag{3.36}
$$

äquivalent. Nach Voraussetzung gilt aber $\frac{\sqrt{n}}{\log(n)}\Delta_n \to \infty$, was

$$
\lim_{n\to\infty} C\frac{\log(n)}{\sqrt{n}}\frac{1}{\Delta_n} = 0
$$

impliziert. Also ist für n groß genug auch die Ungleichung (3.36) efüllt und es gilt

$$(1+\varepsilon)\Delta_n - C\frac{\log(n)}{\sqrt{n}} \geq \Delta_n$$

für n groß genug. Also gilt mit Wahrscheinlichkeit 1 für alle $x \in [0, A_n]$

$$\inf_{u\in[0,x]} |\hat{\varphi}_{X_1^n}(u, 0, \dots, 0)| \geq \Delta_n.$$

Zusammen mit der Definition von $\hat{x}_n$ folgt daraus, dass mit Wahrscheinlichkeit 1

$$\hat{x}_n \geq A_n$$

für n groß genug gilt. Damit ist der Beweis beendet. $\square$

Beweis von Theorem 2.16. Im Beweis von Theorem 2.16 wollen wir Theorem 2.14 und Proposition 2.15 verwenden. Wir beginnen also, indem wir überprüfen ob die Voraussetzungen dieser beiden Resultate unter den Voraussetzungen von Theorem 2.16 erfüllt sind. Offensichtlich gilt $\lim_{n\to\infty} \Delta_n = \lim_{n\to\infty} n^{-\frac{1}{20}} = 0$ und $\lim_{n\to\infty} u_n = \lim_{n\to\infty} n^{\frac{1}{29}} = \infty$. Außerdem gilt

$$\frac{2\log(n)}{\sqrt{n}} \leq n^{\frac{-1}{20}}$$

für n groß genug, sowie

$$\frac{\sqrt{n}}{\log(n)} n^{\frac{-1}{20}} = \frac{1}{\log(n)} n^{\frac{1}{2}-\frac{1}{20}} \to \infty$$

und $u_n = n^{\frac{1}{29}} \leq n$ für alle $n \in \mathbb{N}$. Sei $\beta = \frac{4}{10}$, dann gilt

$$4u_n^{\frac{5}{2}+\beta} \frac{1}{\Delta_n^2} = 4n^{\frac{1}{29}\cdot(\frac{5}{2}+\frac{4}{10})} n^{\frac{1}{10}} = 4n^{\frac{1}{10}+\frac{1}{10}} \leq \frac{\sqrt[4]{n}}{\log(n)}$$

für n groß genug. Also sind die Bedingungen an Δ_n und u_n aus Theorem 2.14 für $\beta = \frac{4}{10}$ erfüllt. Außerdem sind auch die Bedingungen an Δ_n aus Proposition 2.15 erfüllt. Nun überprüfen wir, ob die zusätzlichen Annahmen an Z,X und ε_1 zu den Annahmen aus Definition 2.3 konsistent sind. Nach Voraussetzung sind Z und ε_1 unabhahängig und jeweils mit Erwartungswert 0 und Varianz σ_1^2 beziehungsweise σ_2^2 normalverteilt, also ist auch $X^{(1)}$ normalverteilt es gilt $\mathbf{E}\{X\} = 0$ sowie $\mathbf{V}\{X\} = \sigma_1^2 + \sigma_2^2$. Da Z normalverteilt ist, sind außerdem alle Momente von Z endlich und es gilt $\mathbf{E}\{Z^2\} = \sigma_1^2 > 0$. Da $X^{(1)}$ normalverteilt ist, gilt für die charakteristische Funktion $\varphi_{X^{(1)}}$

$$\varphi_{X^{(1)}}(u) = e^{-\frac{1}{2}(\sigma_1^2+\sigma_2^2)u^2}.$$

Also ist die charakteristische Funktion von $X^{(1)}$ nullstellenfrei. Nach Voraussetzung sind die 4-ten Momente von $\varepsilon_2, \dots, \varepsilon_d$ endlich und die charakteristischen Funktionen von $\varepsilon_2, \dots, \varepsilon_d$ sind nullstellenfrei. Zusammen mit der nullstellenfreiheit von $\varphi_Z(u) = e^{-\frac{1}{2}\sigma_1^2 u^2}$ und der Unabhängigkeit von $Z, \varepsilon_2, \dots, \varepsilon_d$ erhällt man die nullstellenfreiheit der charakteristischen Funktionen von $X^{(2)}, \dots, X^{(d)}$. Damit erfüllen die Zufallsvariablen X, Z und ε die Bedingungen von Definition 2.3 sowie die zusätzlichen Momentenbedingungen aus Theorem 2.14.

Nun überprüfen wir ob die Dichte f von Z die entsprechenden Bedingungen aus Theorem 2.14 erfüllt. Da Z normalverteilt ist und $\mathbf{E}\{Z\} = 0$ und $\mathbf{V}\{Z\} = \sigma_1^2$ gilt besitzt Z die Dichte

$$f_Z(t) = \frac{1}{\sqrt{2\pi}\sigma_1} e^{-\frac{t^2}{2\sigma_1^2}}$$

offensichtlich ist f beliebig oft differenzierbar und die Ableitungen sind jeweils beschränkt. Die Dichte f_Z ist also $(2, C)$-glatt für ein geeignetes $C \in \mathbb{R}^+$. Also lässt sich Theorem 2.14 mit

$$\gamma = \frac{8}{31} < \frac{8}{30} = \frac{2}{3} \cdot \frac{4}{10} = \frac{p}{1+p} \cdot \beta$$

anwenden. Es gilt also

$$\hat{K}_n^{\frac{8}{31}} \int |f(t) - \hat{f}_n(t)| \, dt \to 0$$

fast sicher. Die Aussage des Theorems folgt also, wenn wir zeigen können, dass mit Wahrscheinlichkeit 1

$$\hat{K}_n \geq \Big(\frac{1}{\sigma_1^2 + \sigma_2^2} \log(\frac{n^{\frac{1}{10}}}{4})\Big)^{\frac{1}{2}}$$

für n groß genug gilt. Hierzu verwenden wir Proposition 2.15. Sei $\sigma^2 = \sigma_1^2 + \sigma_2^2$. Dann gilt

$$\varphi_{X^{(1)}}(u) = e^{-\frac{1}{2}\sigma^2 u^2}.$$

Wegen $\Delta_n \to 0$ gilt $-\frac{2}{\sigma} \log(2\Delta_n) > 0$ für n groß genug. Für solche n betrachten wir nun

$$A_n = \sqrt{-\frac{2}{\sigma^2} \log(2\Delta_n)} = \Big(\frac{1}{\sigma_1^2 + \sigma_2^2} \log(\frac{n^{\frac{1}{10}}}{4})\Big)^{\frac{1}{2}}.$$

Dann ist $\frac{A_n}{n}$ beschränkt. Da außerdem die charakteristische Funktion von $X^{(1)}$ auf $[0, A_n]$ streng monoton fallend ist, gilt

$$\begin{aligned}
\inf_{u \in [0, A_n]} |\varphi_{X^{(1)}}(u)| &= |\varphi_{X^{(1)}}(A_n)| \\
&= \exp\left(-\frac{1}{2}\sigma^2 A_n^2\right) \\
&= \exp\left(-\frac{1}{2}\sigma^2 \left(\sqrt{-\frac{2}{\sigma^2}\log(2\Delta_n)}\right)^2\right) \\
&= 2\Delta_n \\
&= (1+1)\Delta_n
\end{aligned}$$

für n groß genug. Also sind die Vorausetzung von Proposition 2.15 für $\varepsilon = 1$ erfüllt. Es folgt, dass mit Wahrscheinlichkeit 1

$$\sup\left\{x \in \mathbb{R}^+ : \inf_{u \in [0,x]} |\hat{\varphi}_{X_1^n}(u, 0, \ldots, 0)| \geq \Delta_n\right\} \geq A_n = \sqrt{-\frac{2}{\sigma^2}\log(2\Delta_n)}$$

für n groß genug gilt. Andererseits gilt auch

$$A_n = \sqrt{-\frac{2}{\sigma^2}\log(2\Delta_n)} \leq n^{\frac{1}{29}} = u_n$$

für n groß genug. Zusammen mit der Definition von $\hat{K}_n$,

$$\hat{K}_n := \min\{u_n, \max\{x \in \mathbb{R} : \inf_{u \in [0,x]} |\hat{\varphi}_{X_1^n}(u, 0, \ldots, 0)| \geq \Delta_n\}\}$$

folgt, dass $\hat{K}_n \geq A_n$ für n groß genug gilt. Damit ist der Beweis beendet. $\square$

4 Regressionsschätzung für latente Variablen

4.1 Modell

Wie wir in Kapitel 2 gesehen haben, lässt sich die Verteilung, die Verteilungsfunktion und die Dichte einer latenten Variable Z aus den Realisierungen der linear von Z abhängigen manifesten Variablen schätzen. Hierzu wurden Realisierungen von Z geschätzt, welche dann in bekannten Schätzern für Verteilung, Verteilungsfunktion und Dichte verwendet werden konnten. In diesem Kapitel werden wir diesen Ansatz auf ein lineares-latentes-Variablen-Modell mit 2 latenten Variablen verallgemeinern und zur Schätzung der Regressionsfunktion verwenden.

Seien X und Y jeweils $\mathbb{R}^{d_X}$- und $\mathbb{R}^{d_Y}$-wertige manifeste Zufallsvariablen. Wir gehen im Folgenden davon aus, dass jede Komponente der manifesten Variablen von höchstens einer der reellwertigen latenten Variablen Z_1, Z_2 abhängt, und dass der zugrundeliegende Zusammenhang linear ist. Unser vorläufiges Modell ist also durch

$$\begin{pmatrix} X^{(1)} \\ X^{(2)} \\ \vdots \\ X^{(d)} \\ Y^{(1)} \\ Y^{(2)} \\ \vdots \\ Y^{(l)} \end{pmatrix} = \begin{pmatrix} 1 \cdot Z_1 \\ a_2 \cdot Z_1 \\ \vdots \\ a_d \cdot Z_1 \\ 1 \cdot Z_2 \\ b_2 \cdot Z_2 \\ \vdots \\ b_l \cdot Z_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_d \\ \delta_1 \\ \delta_2 \\ \vdots \\ \delta_l \end{pmatrix} \tag{4.1}$$

gegeben. Hierbei sind $a_2, a_3, \ldots, a_d, b_1, b_2, \ldots b_l$ die reellen Koeffizienten des Modells, und Z_1, Z_2, ε_1, $\ldots \varepsilon_d$, δ_1, $\ldots$, δ_l sind reellwertige Zufallsvariablen, wobei die Störterme ε_1, $\ldots \varepsilon_d$,δ_1, $\ldots$, δ_l zentriert sind, d.h. es gilt

$$\mathbf{E}\{\varepsilon_j\} = \mathbf{E}\{\delta_k\} = 0$$

für alle $j \in \{1, \ldots, d\}$ und alle $k \in \{1, \ldots, l\}$. Da die Regressionsfunktion

$$m(x) := \mathbf{E}\{Z_2 | Z_1 = x\}$$

eindeutig durch die gemeinsame Verteilung von (Z_1, Z_2) bestimmt ist, kann eine Schätzung von m ausgehend von Realisierungen von

$$(X, Y) = (X^1(1), \ldots, X^{(d)}, Y^{(1)}, \ldots Y^{(l)})$$

nur gelingen, falls die Verteilung von (Z_1, Z_2) durch die Verteilung von (X, Y) eindeutig bestimmt ist. Genau wie im univariaten Fall von Lemma 2.2 benötigt man hierfür stärkere Voraussetzungen. Das folgende Lemma aus [34] liefert hinreichende Bedingungen.

Lemma 4.1. *Seien $X^{(1)}, \ldots X^{(d)}$, $Y^{(1)}, \ldots, Y^{(l)}$, Z_1, Z_2, $\varepsilon_1, \ldots, \varepsilon_d, \delta_1, \ldots, \delta_l$ integrierbare Zufallsvariablen, welche die in Gleichung (4.1) erfüllen. Gilt außerdem*

1) $d \geq 3$ und $l \geq 3$.

2) Die Koeffizienten $a_2, a_3 \in \mathbb{R}$ und $b_2, b_3 \in \mathbb{R}$ sind alle von 0 verschieden.

3) Z_1 und Z_2 sind quadratisch integrierbar mit $\mathbf{E}\{Z_1^2\} > 0$ und $\mathbf{E}\{Z_2^2\} > 0$.

4) Die Zufallsvariablen (Z_1, Z_2), $\varepsilon_1, \ldots, \varepsilon_d$, $\delta_1, \ldots, \delta_l$ sind unabhängig.

5) Die charakteristische Funktion von (X, Y) besitzt keine Nullstellen.

So folgt

$$\varphi_{(X,Y)}(u_1, 0, \ldots, 0, v_1, 0, \ldots, 0) = \varphi_{(Z_1,Z_2)}(u_1, v_1) \cdot \varphi_{\varepsilon_1}(u_1) \cdot \varphi_{\delta_1}(v_1),$$

$$\frac{\partial}{\partial u_2}\varphi_{(X,Y)}(u_1, 0, \ldots, 0, v_1, 0, \ldots, 0) = a_2 \cdot \frac{\partial}{\partial u_1}\varphi_{(Z_1,Z_2)}(u_1, v_1) \cdot \varphi_{\varepsilon_1}(u_1) \cdot \varphi_{\delta_1}(v_1)$$

und

$$\frac{\partial}{\partial v_2}\varphi_{(X,Y)}(u_1, 0, \ldots, 0, v_1, 0, \ldots, 0) = b_2 \cdot \frac{\partial}{\partial v_1}\varphi_{(Z_1,Z_2)}(u_1, v_1) \cdot \varphi_{\varepsilon_1}(u_1) \cdot \varphi_{\delta_1}(v_1).$$

Außerdem sind $a_2, \ldots, a_d$,$b_2, \ldots, b_l$ sowie die Verteilung von (Z_1, Z_2) und die Verteilung der Fehler $\varepsilon_1, \ldots, \varepsilon_d, \delta_1, \ldots, \delta_l$ eindeutig durch die Verteilung von $(X^{(1)}, \ldots, X^{(d)}$,$Y^{(1)}, \ldots, Y^{(l)})$ bestimmt.

Wir nehmen die Voraussetzungen $1)$ bis $5)$ mit in die Definition unseres Modells auf und erhalten:

Definition 4.2. Seien $a_2, \ldots a_d, b_1, \ldots, b_l \in \mathbb{R}$ und seien Z_1, Z_2, ε_1, $\ldots\varepsilon_d$, δ_1, $\ldots$, δ_l reellwertige Zufallsvariablen, mit

$$\mathbf{E}\{\varepsilon_j\} = \mathbf{E}\{\delta_k\} = 0$$

für alle $j \in \{1, \ldots, d\}$ und alle $k \in \{1, \ldots, l\}$. Weiterhin seien die Bedingungen $1) - 5)$ aus Lemmma 4.1 erfüllt. Dann ist das lineare-latente-Variablen-Modell in zwei Variablen gegeben durch

$$\begin{pmatrix} X^{(1)} \\ X^{(2)} \\ \vdots \\ X^{(d)} \\ Y^{(1)} \\ Y^{(2)} \\ \vdots \\ Y^{(l)} \end{pmatrix} = \begin{pmatrix} 1 \cdot Z_1 \\ a_2 \cdot Z_1 \\ \vdots \\ a_d \cdot Z_1 \\ 1 \cdot Z_2 \\ b_2 \cdot Z_2 \\ \vdots \\ b_l \cdot Z_2 \end{pmatrix} + \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_d \\ \delta_1 \\ \delta_2 \\ \vdots \\ \delta_l \end{pmatrix}. \tag{4.2}$$

Nun können wir mit der Konstruktion unseres Schätzers beginnen.

4.2 Definition des Regressionsschätzers

In diesem Unterkapitel werden wir den Regressionsschätzer $\hat{m}_n$ von m konstruieren. Ausgangspunkt sind dabei unabhängige und identisch verteilte Kopien der manifesten Variablen X, Y. Seien

$$(X, Y), (X_1, Y_1), (X_2, Y_2), \ldots$$

unabhängig und identisch verteilt. Im folgenden seien $Z_{1,i}$, $Z_{2,i}$, $\varepsilon_{1,i}$, $\dots$, $\delta_{l,i}$ die zu $X_i = (X_i^{(1)}, \dots, X_i^{(d)})^T$ and $Y_i = (Y_i^{(1)}, \dots, Y_i^{(l)})^T$ gehörigen latenten Variablen und Fehlerterme.

Unser Ziel ist, aus den Daten

$$\mathcal{D}_n = \{(X_1, Y_1), \dots, (X_n, Y_n)\}$$

einen Schätzer $m_n(\cdot) = m_n(\cdot, \mathcal{D}_n) : \mathbb{R} \to \mathbb{R}$ der durch

$$m(x) := \mathbf{E}\{Z_2 | Z_1 = x\}$$

definierten Regressionsfunktion $m : \mathbb{R} \to \mathbb{R}$ zu konstruieren, so dass der L^2-Fehler

$$\int |m_n(z_1) - m(z_1)|^2 \mathbf{P}_{Z_1}(dz_1)$$

klein ist.

Als erstes schätzen wir hierfür die Koeffizienten $a_j, j \in \{1, \dots, d\}$ und $b_k, k \in \{1, \dots, l\}$ des Modells (4.2). Dazu gehen wir wie im univariaten Fall vor. Nach Definition 4.2 gilt $a_1 = b_1 = 1$. Also müssen diese Koeffizienten nicht geschätzt werden, wir definieren $\hat{a}_1 = \hat{b}_1 = 1$. Da (Z_1, Z_2) von den Fehlertermen unabhängig ist, gilt für $i \neq j$

$$\begin{aligned}
\mathbf{E}\{X^{(j)} \cdot X^{(i)}\} &= \mathbf{E}\{(a_j Z_1 + \varepsilon_j) \cdot (a_i Z_1 + \varepsilon_i)\} \\
&= \mathbf{E}\{a_j a_i Z_1^2\} + \mathbf{E}\{\varepsilon_j \cdot a_i Z_1\} + \mathbf{E}\{\varepsilon_i \cdot a_j Z_1\} + \mathbf{E}\{\varepsilon_i \cdot \varepsilon_j\} \\
&= a_j a_i \mathbf{E}\{Z_1^2\} + a_i \mathbf{E}\{\varepsilon_j\}\mathbf{E}\{Z_1\} + a_j \mathbf{E}\{\varepsilon_i\}\mathbf{E}\{Z_1\} + \mathbf{E}\{\varepsilon_i\} \cdot \mathbf{E}\{\varepsilon_j\} \\
&= a_j a_i \mathbf{E}\{Z_1^2\}.
\end{aligned}$$

Wobei wir verwendet haben, dass $\mathbf{E}\{\varepsilon_j\} = 0$ für alle $j \in \{1, \dots, d\}$ gilt. Nun ist Z_1, Z_2 quadratisch integrierbar, damit folgt

$$\frac{\mathbf{E}\{X^{(2)} \cdot X^{(3)}\}}{\mathbf{E}\{X^{(1)} \cdot X^{(3)}\}} = \frac{a_2 a_3 \mathbf{E}\{Z_1^2\}}{a_1 a_3 \mathbf{E}\{Z_1^2\}} = a_2$$

sowie

$$\frac{\mathbf{E}\{X^{(2)} \cdot X^{(j)}\}}{\mathbf{E}\{X^{(1)} \cdot X^{(2)}\}} = \frac{a_2 a_j \mathbf{E}\{Z_1^2\}}{a_1 a_2 \mathbf{E}\{Z_1^2\}} = a_j$$

für $j \in \{3, \dots, d\}$. Hierbei haben wir verwendet, dass nach Voraussetzung $a_1 = 1$ und $0 < \mathbf{E}\{Z_1^2\} < \infty$ gilt. Völlig analog folgt

$$b_2 = \frac{\mathbf{E}\{Y^{(2)} \cdot Y^{(3)}\}}{\mathbf{E}\{Y^{(1)} \cdot Y^{(3)}\}} \quad \text{und} \quad b_k = \frac{\mathbf{E}\{Y^{(2)} \cdot Y^{(k)}\}}{\mathbf{E}\{Y^{(1)} \cdot Y^{(2)}\}} \quad \text{für } k \in \{3, \dots, l\}.$$

Wir erhalten unsere Schätzer, indem wir die Erwartungswerte durch empirische Mittel ersetzen

$$\hat{a}_2 = \frac{\frac{1}{n}\sum_{i=1}^n X_i^{(2)} \cdot X_i^{(3)}}{\frac{1}{n}\sum_{i=1}^n X_i^{(1)} \cdot X_i^{(3)}} \quad \text{und} \quad \hat{a}_j = \frac{\frac{1}{n}\sum_{i=1}^n X_i^{(2)} \cdot X_i^{(j)}}{\frac{1}{n}\sum_{i=1}^n X_i^{(1)} \cdot X_i^{(2)}}$$

und

$$\hat{b}_2 = \frac{\frac{1}{n}\sum_{j=1}^n Y_j^{(2)} \cdot Y_j^{(3)}}{\frac{1}{n}\sum_{j=1}^n Y_j^{(1)} \cdot Y_j^{(3)}} \quad \text{und} \quad \hat{b}_k = \frac{\frac{1}{n}\sum_{j=1}^n Y_j^{(2)} \cdot Y_j^{(k)}}{\frac{1}{n}\sum_{j=1}^n Y_j^{(1)} \cdot Y_j^{(2)}}$$

für $j,k>2$.

Ist $\hat{z}_{1,i}$ ein Schätzer für $Z_{1,i}$ so erhalten wir wegen

$$X^{(j)} - a_j Z_1 = \varepsilon_j \quad (j \in \{1,\dots,d\})$$

auch einen Schätzer $\hat{\varepsilon}_{j,i}$ von $\varepsilon_{j,i} = X_i^{(j)} - a_j Z_{1,i}$ durch

$$\hat{\varepsilon}_{j,i} = X_i^{(j)} - \hat{a}_j \hat{z}_{1,i}.$$

Die Schätzer für $\delta_{k,i}$ werden analog durch

$$\hat{\delta}_{k,i} = Y_i^{(k)} - \hat{b}_k \hat{z}_{2,i}$$

definiert. Wir benötigen also nur noch einen Schätzer $(\hat{z}_{1,i}, \hat{z}_{2,i})$ von $(Z_{1,i}, Z_{2,i})$. Fallst (X,Y), (Z_1,Z_2), ε_1 und δ_1 die Bedingungen aus Definition 4.2 erfüllen, so folgt aus Lemma 4.1, dass die charakteristischen Funktionen von (X,Y), (Z_1,Z_2), ε_1 und δ_1 und ihre partiellen Ableitungen die folgenden 3 Gleichungen erfüllen:

$$\varphi_{(X,Y)}(u_1,0,\dots,0,v_1,0,\dots,0) = \varphi_{(Z_1,Z_2)}(u_1,v_1) \cdot \varphi_{\varepsilon_1}(u_1) \cdot \varphi_{\delta_1}(v_1),$$

$$\frac{\partial}{\partial u_2}\varphi_{(X,Y)}(u_1,0,\dots,0,v_1,0,\dots,0) = a_2 \cdot \frac{\partial}{\partial u_1}\varphi_{(Z_1,Z_2)}(u_1,v_1) \cdot \varphi_{\varepsilon_1}(u_1) \cdot \varphi_{\delta_1}(v_1)$$

und

$$\frac{\partial}{\partial v_2}\varphi_{(X,Y)}(u_1,0,\dots,0,v_1,0,\dots,0) = b_2 \cdot \frac{\partial}{\partial v_1}\varphi_{(Z_1,Z_2)}(u_1,v_1) \cdot \varphi_{\varepsilon_1}(u_1) \cdot \varphi_{\delta_1}(v_1).$$

Wir versuchen nun $(\hat{z}_{1,i}, \hat{z}_{2,i})$ so zu wählen, dass diese Gleichungen zwischen den entsprechenden empririschen charakteristischen Funktionen näherungsweise gelten. Wie in Kapitel 2.2 benötigen wir hierzu eine Folge

$$\hat{K}_n = \hat{K}_n(X_1,\dots,X_n,Y_1,\dots,Y_n) \in \mathbb{N}_0$$

von Zufallsvariablen, die Sicherstellen, dass der Betrag der empirischen charakteristischen Funktion von $(X,Y)_1^n = (X_1,Y_1),(X_2,Y_2),\dots,(X_n,Y_n)$, die durch

$$\hat{\varphi}_{(X^{(1)},Y^{(1)})_1^n}(u,v) := \sum_{J=1}^{n} \exp\left(iuX_j^{(1)} + ivY_j^{(1)}\right)$$

gegeben ist auf einem entsprechenden Intervall einen hinreichenden Abstand zu 0 einhällt. Wir definier $\hat{K}_N$ durch

$$\hat{K}_n := \min\left\{\lfloor n^{\frac{1}{11}} \rfloor, \max\{K \in \mathbb{N} : \inf_{u,v\in[-\sqrt{K},\sqrt{K}]} \left|\hat{\varphi}_{(X^{(1)},Y^{(1)})_1^n}(u\pi, v\pi)\right| \geq \frac{1}{\log(n)}\}\right\}. \quad (4.3)$$

Im Unterschied zu den $\hat{K}_n$ aus der Verteilungsschätzung welche in Definition 2.4 eingeführt werden, gibt es für die hier definierten $\hat{K}_n$ keine steuerenden Folgen u_n und Δ_n. Da wir für den Regressionsschätzer keine Aussagen zur Konvergenzgeschwindigkeit beweisen werden, verzichten wir auf die Einführung entsprechender Folgen um die Notation in den Folgenden Theoremen zu vereinfachen.

Da $\varphi_{(X,Y)}$ nach Voraussetzung nullstellenfrei ist, liegt die Vermutung nahe, dass $\hat{K}_n$ gegen ∞ konvergiert. Dass dies tatsächlich der Fall ist, werden wir in Lemma 5.3 beweisen. Wir definieren $\hat{L}_n := \sqrt{\hat{K}_n}$ und $\hat{L}_n^* := 2 \cdot \hat{L}_n = 2 \cdot \sqrt{\hat{K}_n}$. Zur kontruktion des Samples verwenden wir empirische charakteristische Funktionen. Gilt $x \in \mathbb{R}^d \times R^n$ und $y \in \mathbb{R}^l \times \mathbb{R}^d$ so definieren wir $\hat{\varphi}_{(x,y)} : \mathbb{R}^d \times \mathbb{R}^l \mapsto \mathbb{C}$ durch

$$\hat{\varphi}_{(x,y)}(u,v) = \sum_{j=1}^{n} \exp\left(i\langle x_j, u\rangle + i\langle y_j, v\rangle\right).$$

Dann konstruieren wir ein geschätztes Sample

$$(\hat{z}_1, \hat{z}_2)_1^n = (\hat{z}_{1,1}, \hat{z}_{2,1}), \ldots, (\hat{z}_{1,n}, \hat{z}_{2,n})$$

wwww von (Z_1, Z_2), indem wir

$$\begin{aligned}
&T_n((\hat{z}_1, \hat{z}_2)_1^n) := \\
&\sup_{u_1,v_1 \in [-2\hat{K}_n, 2\hat{K}_n]} \Bigg\{ \Bigg| \hat{\varphi}_{(X^{(1)},Y^{(1)})_1^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}, \frac{\pi v_1}{\hat{L}_n^*}\Big) \\
&\qquad - \hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}, \frac{\pi v_1}{\hat{L}_n^*}\Big) \cdot \hat{\varphi}_{\hat{\varepsilon}_{1,1}^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}\Big) \cdot \hat{\varphi}_{\hat{\delta}_{1,1}^n}\Big(\frac{\pi v_1}{\hat{L}_n^*}\Big) \Bigg| \\
&+ \Bigg| \frac{\partial}{\partial u_2} \hat{\varphi}_{(X,Y)_1^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}, 0, \ldots, 0, \frac{\pi v_1}{\hat{L}_n^*}, 0, \ldots, 0\Big) \\
&\qquad - \hat{a}_2 \cdot \frac{\partial}{\partial u_1} \hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}, \frac{\pi v_1}{\hat{L}_n^*}\Big) \cdot \hat{\varphi}_{\hat{\varepsilon}_{1,1}^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}\Big) \cdot \hat{\varphi}_{\hat{\delta}_{1,1}^n}\Big(\frac{\pi v_1}{\hat{L}_n^*}\Big) \Bigg| \\
&+ \Bigg| \frac{\partial}{\partial v_2} \hat{\varphi}_{(X,Y)_1^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}, 0, \ldots, 0, \frac{\pi v_1}{\hat{L}_n^*}, 0, \ldots, 0\Big) \\
&\qquad - \hat{b}_2 \cdot \frac{\partial}{\partial v_1} \hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}, \frac{\pi v_1}{\hat{L}_n^*}\Big) \cdot \hat{\varphi}_{\hat{\varepsilon}_{1,1}^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}\Big) \cdot \hat{\varphi}_{\hat{\delta}_{1,1}^n}\Big(\frac{\pi v_1}{\hat{L}_n^*}\Big) \Bigg| \Bigg\}
\end{aligned} \tag{4.4}$$

unter der Nebenbedingung $\frac{1}{n} \cdot \sum_{j=1}^n |\hat{z}_{2,j}|^6 \leq \sqrt{\hat{K}_n}$ minimieren. Ausgehend von den geschätzten Daten schätzen wir die Regressionsfunktion m durch einen Kleinste-Quadrate-Schätzer. Da wir, ähnlich wie beim Verteilungsschätzer, die Konvergenz von trigonometrischen Polynomen nutzen wollen, minimieren wir die Fehlerquadrate über einem entsprechenden Raum von trigonometrischen Polynomen.

Definition 4.3. Sei $\mathcal{F}_n = F_n(\hat{K}_n, \beta_n)$ gegeben durch

$$\begin{aligned}
\mathcal{F}_n = \Big\{ f : \mathbb{R} \to \mathbb{R} \,\Big|\, & \\
& f(z) = \sum_{k=0}^{\hat{K}_n} a_k \cos\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z\right) + \sum_{k=0}^{\hat{K}_n} b_k \sin\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z\right), \\
& a_k, b_k \in [-\beta_n, \beta_n] \Big\}.
\end{aligned} \tag{4.5}$$

Dann ist der zu F_n gehörige Kleinste-Quadrate-Schätzer $\hat{m}_n$ von m gegeben durch

$$\hat{m}_n(\cdot) = \mathrm{argmin}_{f \in \mathcal{F}_n} \frac{1}{n} \sum_{i=1}^{n} |\hat{z}_{i,2} - f(\hat{z}_{i,1})|^2. \tag{4.6}$$

4.3 Konsistenz des Regressionsschätzers

Theorem 4.4. *Seien die Bedingungen* $1)$ *bis* $5)$ *aus Definition 4.2 erfüllt. Seien die* 6*-ten Momente von* Z_1 *und* Z_2 *sowie* $\varepsilon_1, \ldots, \varepsilon_d, \delta_1, \ldots, \delta_l$ *endlich sei* $\hat{K}_n$ *definiert wie in (4.3) und* $\mathcal{F}_n$ *definiert wie in (4.5) mit* $\hat{L}_n^* = 2 \cdot \hat{L}_n = 2 \cdot \sqrt{\hat{K}_n}$*. Weiterhin seien die geschätzten Daten* $(\hat{z}_1, \hat{z}_2)_1^n = \{(\hat{z}_{1,1}, \hat{z}_{2,1}), \ldots, (\hat{z}_{1,n}, \hat{z}_{2,n}$ *definiert als Minimum der Funktion* T_n *aus Gleichung (4.4) unter der Nebenbedingung* $\frac{1}{n} \cdot \sum_{j=1}^{n} |\hat{z}_{2,j}|^6 \leq \sqrt{\hat{K}_n}$*. Ist dann der Regressionsschätzer* $\hat{m}_n$ *definiert wie in (4.6), so konvergiert der* L_2*-Fehler von* $\hat{m}_n$ *fast sicher gegen* 0*. Es gilt also*

$$\int |m_n(x) - m(x)|^2 \mathbf{P}_{Z_1}(dx) \to 0 \quad (n \to \infty)$$

mit Wahrscheinlichkeit 1*.*

Theorem 4.4 zeigt, dass es möglich ist für latente Variablen einen stark konsistenten, d.h. einen in L^2 fast sicher konvergentn, Regressionsschätzer m_n zu konstruieren, für den keinerlei Annahmen an die Beschaffenheit der Regressionsfunktion m benötigt werden. Tatsächlich gehen die benötigten Voraussetzungen an die Verteilung der zugrundeliegenden latenten Variablen (Z_1, Z_2) nur geringfügig über die in Lemma 4.1 vorgestellten Mindestbedingungen hinaus. Lediglich die Momentenbedingung $3)$ aus Lemma 4.1 wurde verschärft, indem zusätzlich angenommen wurde, dass auch die 6-ten Momente von Z_1, Z_2 und den Fehlertermen $\varepsilon_1, \ldots, \varepsilon_d, \delta_1, \ldots, \delta_l$ endlich seien.

5 Beweise zu Kapitel 4

5.1 Beweis der Konvergenz der geschätzten Daten

Wir beginnen mit dem Beweis von Lemma 4.1.

Beweis von Lemma 4.1. Ein Beweis findet sich beispielsweise in [34]. Der Vollständigkeit halber folgt eine leicht modifizierte Variante des dortigen Beweises.

Die erste Aussage des Lemmas lässt sich einfach nachrechnen. Gilt (4.2) so folgt aus der Unabhängigkeit von $(Z_1, Z_2), \varepsilon_1, \dots, \varepsilon_d, \delta_1, \dots, \delta_l$, dass die charakteristische Funktion $\varphi_{(X,Y)}$ von (X, Y) durch

$$\begin{aligned}
&\varphi_{(X,Y)}(u_1, \dots, u_d, v_1, \dots, v_l) \\
&= \mathbf{E}\left\{\exp\left(i \cdot \sum_{j=1}^{d} u_j \cdot X^{(j)} + i \cdot \sum_{k=1}^{l} v_k \cdot Y^{(k)}\right)\right\} \\
&= \mathbf{E}\left\{\exp\left(i \cdot \sum_{j=1}^{d} u_j \cdot (a_j \cdot Z_1 + \varepsilon_j) + i \cdot \sum_{k=1}^{l} v_k \cdot (b_k \cdot Z_2 + \delta_k)\right)\right\} \\
&= \mathbf{E}\Bigg\{\exp\left(i \cdot (\sum_{j=1}^{d} u_j \cdot a_j \cdot Z_1 + \sum_{k=1}^{l} v_k \cdot b_k \cdot Z_2)\right) \cdot \prod_{j=1}^{d} \exp\left(i \cdot u_j \cdot \varepsilon_j\right) \\
&\qquad \cdot \prod_{k=1}^{l} \exp\left(i \cdot v_k \cdot \delta_k\right)\Bigg\} \\
&= \varphi_{(Z_1,Z_2)}\left(\sum_{j=1}^{d} u_j \cdot a_j \ , \ \sum_{k=1}^{l} v_k \cdot b_k\right) \cdot \prod_{j=1}^{d} \varphi_{\varepsilon_j}(u_j) \cdot \prod_{k=1}^{l} \varphi_{\delta_k}(v_k)
\end{aligned}$$

gegeben ist. Mit

$$\varphi_{\varepsilon_j}(0) = \varphi_{\delta_k}(0) = 1 \quad (j = 2, \dots, d, k = 2, \dots, l)$$

und

$$\varphi'_{\varepsilon_2}(0) = i \cdot \mathbf{E}\varepsilon_2 = 0 = \varphi'_{\delta_2}(0)$$

folgt daher

$$\varphi_{(X,Y)}(u_1, 0, \dots, 0, v_1, 0, \dots, 0) = \varphi_{(Z_1,Z_2)}(u_1, v_1) \cdot \varphi_{\varepsilon_1}(u_1) \cdot \varphi_{\delta_1}(v_1),$$

sowie

$$\begin{aligned}
&\frac{\partial}{\partial u_2}\varphi_{(X,Y)}(u_1, 0, \dots, 0, v_1, 0, \dots, 0) \\
&= a_2 \cdot \frac{\partial}{\partial u_1}\varphi_{(Z_1,Z_2)}(u_1, v_1) \cdot \varphi_{\varepsilon_1}(u_1) \cdot \varphi_{\delta_1}(v_1) \\
&\quad + \varphi_{(Z_1,Z_2)}(u_1, v_1) \cdot \varphi_{\varepsilon_1}(u_1) \cdot \varphi_{\delta_1}(v_1) \cdot \varphi'_{\varepsilon_2}(0) \\
&= a_2 \cdot \frac{\partial}{\partial u_1}\varphi_{(Z_1,Z_2)}(u_1, v_1) \cdot \varphi_{\varepsilon_1}(u_1) \cdot \varphi_{\delta_1}(v_1)
\end{aligned}$$

und

$$\frac{\partial}{\partial v_2}\varphi_{(X,Y)}(u_1,0,\ldots,0,v_1,0,\ldots,0) = b_2 \cdot \frac{\partial}{\partial v_1}\varphi_{(Z_1,Z_2)}(u_1,v_1)\cdot\varphi_{\varepsilon_1}(u_1)\cdot\varphi_{\delta_1}(v_1).$$

Es bleibt noch zu zeigen, dass die Verteilung von (Z_1, Z_2) eindeutig durch die Verteilung von (X, Y) bestimmt ist. Angenommen, es gelten die Beziehungen aus (4.2) und es gilt

$$\begin{pmatrix} X^{(1)} \\ X^{(2)} \\ \vdots \\ X^{(d)} \\ Y^{(1)} \\ Y^{(2)} \\ \vdots \\ Y^{(l)} \end{pmatrix} = \begin{pmatrix} 1\cdot\tilde{Z}_1 \\ \tilde{a}_2\cdot\tilde{Z}_1 \\ \vdots \\ \tilde{a}_d\cdot\tilde{Z}_1 \\ 1\cdot\tilde{Z}_2 \\ \tilde{b}_2\cdot\tilde{Z}_2 \\ \vdots \\ \tilde{b}_l\cdot\tilde{Z}_2 \end{pmatrix} + \begin{pmatrix} \tilde{\varepsilon}_1 \\ \tilde{\varepsilon}_2 \\ \vdots \\ \tilde{\varepsilon}_d \\ \tilde{\delta}_1 \\ \tilde{\delta}_2 \\ \vdots \\ \tilde{\delta}_l \end{pmatrix}$$

in Verteilung, wobei die Zufallsvariablen $\tilde{Z}_1, \tilde{Z}_2, \tilde{\varepsilon}_1, \ldots, \tilde{\delta}_l$ die Bedingungen von Definition 4.2 erfüllen. Nach Definition gilt $a_1 = \tilde{a}_1 = b_1 = \tilde{b}_1 = 1$. Aus der Unabhängigkeit folgt

$$a_j a_i \mathbf{E}\{Z_1^2\} = \mathbf{E}\{X^{(j)}\cdot X^{(i)}\} = \tilde{a}_j\tilde{a}_i\mathbf{E}\{\tilde{Z}_1^2\}$$

für $i, j \in \{1, \ldots, d\}$. Für $j, i \in \{1, 2, 3\}$ ist die linke Seite nicht Null, also muss für $j, i \in \{1, 2, 3\}$ auch die rechte Seite von Null verschieden sein. Es folgt

$$a_2 = \frac{\mathbf{E}\{X^{(2)}\cdot X^{(3)}\}}{\mathbf{E}\{X^{(1)}\cdot X^{(3)}\}} = \tilde{a}_2.$$

Analog folgt auch $a_j = \tilde{a}_j$ und $b_i = \tilde{b}_i$ für alle $j \in \{1, \ldots, d\}$ und alle $i \in \{1, \ldots, l\}$. Wie im univariaten Fall verwenden wir nun Logarithmen der charakteristischen Funktionen. Offensichtlich sind die Funktionen

$$\varphi_v(u):\ u \mapsto \varphi_{(Z_1,Z_2)}(u,v) \quad \text{and} \quad \varphi_u(v):\ v \mapsto \varphi_{(Z_1,Z_2)}(u,v)$$

für jedes $v \in \mathbb{R}$ beziehungsweise $u \in \mathbb{R}$ jeweils stetig. Außerdem ist aus der Wahrscheinlichkeitstheorie bekannt, dass die charakteristische Funktion einer Zufallsvariablen stetig differenzierbar ist, falls die entsprechenden Momente existieren (siehe beispielsweise [40, Theorem 25.2]). Da Z_1 und Z_2 nach quadratisch integrierbar sind, sind beide integrierbar. Also sind auch die Funktionen

$$\varphi_v(u):\ u \mapsto \varphi_{(Z_1,Z_2)}(u,v) \quad \text{and} \quad \varphi_u(v): v \mapsto \varphi_{(Z_1,Z_2)}(u,v)$$

für feste v beziehungsweise u jeweils stetig differenzierbar. Ist aber $f : \mathbb{R} \to \mathbb{C}$ stetig differenzierbar und auf einem Intervall $I \subset \mathbb{R}$ nullstellenfrei, so können wir $L_f(t)$ für $t \in I$ durch

$$L_f(t) := w_0 + \int_{t_0}^{t} \frac{f'(\tau)}{f(\tau)}\, d\tau$$

definieren. Hierbei ist w_0 eine komplexe Zahl, welche $\exp(w_0) = f(t_0)$ erfüllt. Das so definierte $L_f(t)$ ist ein Logarithmus von $f(t)$, es gilt also $\exp(L_f(t)) = f(t)$ (siehe [39, Seite 123ff.]). Nach Voraussetzung ist die charakteristische Funktion $\varphi_{(Z_1,Z_2)}$ nullstellenfrei. Damit sind auch die obigen Funktionen φ_v und φ_u für jedes $u \in \mathbb{R}$ und jedes $v \in \mathbb{R}$ nullstellenfrei. Damit ist

$$t \mapsto \int_0^t \frac{\frac{\partial}{\partial u}\varphi_{(Z_1,Z_2)}(z,0)}{\varphi_{(Z_1,Z_2)}(z,0)}\, dz$$

für jedes $t \in \mathbb{R}$ wegen $\log(\varphi_{(Z_1,Z_2)}(0,0)) = \log(1) = 0$ ein Logarithmus von $\varphi_{(Z_1,Z_2)}(t,0)$. Wir wenden die selbe Technik erneut an und erhalten

$$L_{\varphi_u}(v) = \int_0^u \frac{\frac{\partial}{\partial u}\varphi_{(Z_1,Z_2)}(z,0)}{\varphi_{(Z_1,Z_2)}(z,0)}\, dz + \int_0^v \frac{\frac{\partial}{\partial v}\varphi_{(Z_1,Z_2)}(u,z)}{\varphi_{(Z_1,Z_2)}(u,z)}\, dz. \tag{5.1}$$

Wegen

$$\log(\varphi_{(Z_1,Z_2)}(u,0)) = \int_0^u \frac{\frac{\partial}{\partial u}\varphi_{(Z_1,Z_2)}(z,0)}{\varphi_{(Z_1,Z_2)}(z,0)}\, dz$$

folgt $\exp(L_{\varphi_u}(v)) = \varphi_{(Z_1,Z_2)}(u,v)$ für $(u,v) \in \mathbb{R}^2$. Analog erhalten wir auch einen Logarithmus von $\varphi_{(\tilde{Z}_1,\tilde{Z}_2)}$. Zusammen mit den Beziehungen aus dem ersten Teil des Beweises folgt

$$\begin{aligned}\varphi_{(Z_1,Z_2)} = &\exp\Big(\int_0^v \frac{1}{b_2}\cdot\frac{\frac{\partial}{\partial v_2}\varphi_{(X,Y)}(u,0,\ldots,0,s,0,\ldots,0)}{\varphi_{(X,Y)}(u,0,\ldots,0,s,0,\ldots,0)}ds\Big)\\ &+\exp\Big(\int_0^u \frac{1}{a_2}\cdot\frac{\frac{\partial}{\partial u_2}\varphi_{(X,Y)}(t,0,\ldots,0)}{\varphi_{(X,Y)}(t,0,\ldots,0)}dt\Big)\end{aligned}$$

und

$$\begin{aligned}\varphi_{(\tilde{Z}_1,\tilde{Z}_2)} = &\exp\Big(\int_0^v \frac{1}{\tilde{b}_2}\cdot\frac{\frac{\partial}{\partial v_2}\varphi_{(X,Y)}(u,0,\ldots,0,s,0,\ldots,0)}{\varphi_{(X,Y)}(u,0,\ldots,0,s,0,\ldots,0)}ds\Big)\\ &+\exp\Big(\int_0^u \frac{1}{\tilde{a}_2}\cdot\frac{\frac{\partial}{\partial u_2}\varphi_{(X,Y)}(t,0,\ldots,0)}{\varphi_{(X,Y)}(t,0,\ldots,0)}dt\Big).\end{aligned}$$

Aus $\tilde{a}_2 = a_2$ und $\tilde{b_2} = b_2$ folgt also $\varphi_{(\tilde{Z}_1,\tilde{Z}_2)} = \varphi_{(Z_1,Z_2)}$. Zusammen mit

$$\begin{aligned}\varphi_{(\tilde{Z}_1,\tilde{Z}_2)}(\tilde{a}_j\cdot u_j,0)\cdot\varphi_{\tilde{\varepsilon}_j}(u_j) &= \varphi_{(X,Y)}(0,\ldots,0,u_j,0,\ldots,0)\\ &= \varphi_{(Z_1,Z_2)}(a_j\cdot u_j,0)\cdot\varphi_{\varepsilon_j}(u_j)\end{aligned}$$

und den anlogen Gleichungen für die charakteristischen Funktionen der δ_i folgt daraus

$$\varphi_{\varepsilon_j} = \varphi_{\tilde{\varepsilon}_j} \quad \text{und} \quad \varphi_{\delta_i} = \varphi_{\tilde{\delta}_i}$$

für $j \in \{1,\ldots,d\}$ und $i \in \{1,\ldots,l\}$. Aus dem Eindeutigkeitssatz für charakteristische Funktionen (siehe [40], Theorem 23.4) folgt, dass die Verteilung von (Z_1,Z_2) mit der Verteilung von $\tilde{Z}_1,\tilde{Z}_2$ übereinstimmen, und dass für alle $j \in \{1,\ldots,d\}$ und $i \in \{1,\ldots,l\}$ die Verteilungen von ε_j und $\tilde{\varepsilon}_j$ und die Verteilungen von δ_i und $\tilde{\delta}_i$ übereinstimmen. Damit ist der Beweis beendet. □

Wir wollen zeigen, dass die geschätzten Daten die latenden Variablen „gut“ approximieren. Da wir bei der Konstruktion die Beziehungen zwischen den charakteristischen Funktionen der latenten Variablen und der manifesten Variablen nachgeahmt haben, wäre es wünschenswert, wenn die empirische charakteristische Funktion der geschätzten Daten

$$\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(u,v) = \frac{1}{n}\sum_{j=1}^{n} \exp\left(i(u\hat{z}_{1,j} + v\hat{z}_{2,j})\right)$$

gegen die charakteristische Funktion der latenten Variablen, $\varphi_(Z_1, Z_2)$, konvergiert. Wie im univariaten Fall der Verteilungs und Dichteschätzung benötigen wir eine ganze Reihe von Hilfsresultaten.

Lemma 5.1. *Seien $W^{(1)}$ und $V^{(1)}$ reellwertige Zufallsvariablen mit $\mathbf{E}\{|W^{(1)}|\} < \infty$ und $\mathbf{E}\{|V^{(1)}|\} < \infty$ und seien $(W^{(1)}, V^{(1)})$,$(W_1^{(1)}, V_1^{(1)}), \ldots$ unabhängig und identisch verteilt. Dann gilt*

$$\frac{\sqrt{n}}{\log(n)} \sup_{u,v\in[-n,n]} \left|\varphi_{(W^{(1)},V^{(1)})}(u\pi, v\pi) - \hat{\varphi}_{(W^{(1)},V^{(1)})_1^n}(u\pi, v\pi)\right| \to 0$$

fast sicher.

Bemerkung. In Theorem 4.4 wird angenommen, dass Z_1, Z_2, ε_1 und δ_1 integrabel sind. Somit sind auch $X^{(1)}$ und $Y^{(1)}$ integrabel. Liegt $u \in [-2\hat{K}_n, 2\hat{K}_n]$, so folgt wegen $\hat{L}_n^* = 2\hat{L}_n = 2\sqrt{\hat{K}_n}$, dass auch $\frac{u}{\hat{L}_n^*} \in [-\hat{L}_n, \hat{L}_n]$ gilt. Nach Voraussetzung gilt außerdem $\hat{K}_n \leq n^{\frac{1}{11}}$, was $[-\hat{L}_n, \hat{L}_n] \subseteq [-n, n]$ impliziert. Insgesamt gilt also

$$\begin{aligned}
&\frac{\sqrt{n}}{\log(n)} \sup_{u,v\in[-2\hat{K}_n,2\hat{K}_n]} \left|\varphi_{(Z_1,Z_2)}\left(\frac{u\pi}{\hat{L}_n^*}, \frac{v\pi}{\hat{L}_n^*}\right) - \hat{\varphi}_{(Z_1,Z_2)_1^n}\left(\frac{u\pi}{\hat{L}_n^*}, \frac{v\pi}{\hat{L}_n^*}\right)\right| \\
&\leq \frac{\sqrt{n}}{\log(n)} \sup_{u,v\in[-\hat{L}_n,\hat{L}_n]} \left|\varphi_{(Z_1,Z_2)}(u\pi, v\pi) - \hat{\varphi}_{(Z_1,Z_2)_1^n}(u\pi, v\pi)\right| \\
&\leq \frac{\sqrt{n}}{\log(n)} \sup_{u,v\in[-n,n]} \left|\varphi_{(Z_1,Z_2)}(u\pi, v\pi) - \hat{\varphi}_{(Z_1,Z_2)_1^n}(u\pi, v\pi)\right|.
\end{aligned}$$

Mit Lemma 5.1 folgt somit

$$\frac{\sqrt{n}}{\log(n)} \sup_{u,v\in[-2\hat{K}_n,2\hat{K}_n]} \left|\varphi_{(Z_1,Z_2)}\left(\frac{u\pi}{\hat{L}_n^*}, \frac{v\pi}{\hat{L}_n^*}\right) - \hat{\varphi}_{(Z_1,Z_2)_1^n}\left(\frac{u\pi}{\hat{L}_n^*}, \frac{v\pi}{\hat{L}_n^*}\right)\right| \to 0 \quad a.s.$$

Sowie die analogen Aussagen für $(X^{(1)}, Y^{(1)})$, ε_1 und δ_1.

Beweis. Wie im Beweis von Lemma 3.5 stutzen wir die $W^{(1)}$ und $V^{(1)}$. Sei $M_n = \sqrt{n}$. Es gilt

$$\begin{aligned}
&\sup_{u,v\in[-n,n]} \left|\varphi_{(W^{(1)},V^{(1)})}(u\pi, v\pi) - \hat{\varphi}_{(W^{(1)},V^{(1)})_1^n}(u\pi, v\pi)\right| \\
&\leq \sup_{u,v\in[-n,n]} \left|\varphi_{(W^{(1)},V^{(1)})}(u\pi, v\pi) - \varphi_{(W_{M_n}^{(1)},V_{M_n}^{(1)})}(u\pi, v\pi)\right|
\end{aligned}$$

$$
\begin{aligned}
&+ \sup_{u,v\in[-n,n]} \left|\varphi_{(W^{(1)}_{M_n},V^{(1)}_{M_n})}(u\pi, v\pi) - \hat{\varphi}_{(W^{(1)}_{M_n},V^{(1)}_{M_n})_1^n}(u\pi, v\pi)\right| \\
&+ \sup_{u,v\in[-n,n]} \left|\hat{\varphi}_{(W^{(1)}_{M_n},V^{(1)}_{M_n})_1^n}(u\pi, v\pi) - \hat{\varphi}_{(W^{(1)},V^{(1)})_1^n}(u\pi, v\pi)\right| \\
&= T_{1,n} + T_{2,n} + T_{3,n}.
\end{aligned}
$$

Es genügt also zu zeigen, dass $\frac{\sqrt{n}}{\log(n)} \cdot T_{1,n}$, $\frac{\sqrt{n}}{\log(n)} \cdot T_{2,n}$ und $\frac{\sqrt{n}}{\log(n)} \cdot T_{3,n}$ fast sicher gegen 0 konvergieren. Wir beginnen mit $T_{1,n}$. Wir verwenden die Eulerschen Identität $\exp(ix) = \cos(x) + i\sin(x)$, die Tatsache, dass für alle $z \in \mathbb{C}$ die Abschätzung $|z| \leq |\Re(z)| + |\mathrm{Im}\,(z)|$ gitl, und die trivialen Abschätzungen $|\cos(x)| \leq 1$ und $|\sin(x)| \leq 1$. Damit folgt

$$
\begin{aligned}
T_{1,n} \leq &\sup_{u,v\in[-n,n]} \mathbf{E}\Big\{\left|\cos(W^{(1)}u\pi + V^{(1)}v\pi) - \cos(W^{(1)}_{M_n}u\pi + V^{(1)}_{M_n}v\pi)\right| \\
&+ \left|\sin(W^{(1)}u\pi + V^{(1)}v\pi) - \sin(W^{(1)}_{M_n}u\pi + V^{(1)}_{M_n}v\pi)\right|\Big\} \\
\leq &\sup_{u,v\in[-n,n]} \mathbf{E}\Big\{\left|\cos(W^{(1)}u\pi + V^{(1)}v\pi) - \cos(W^{(1)}_{M_n}u\pi + V^{(1)}_{M_n}v\pi)\right| \\
&\cdot [\mathbb{1}_{\{|W^{(1)}|\geq M_n\}} + \mathbb{1}_{\{|V^{(1)}|\geq M_n\}}] \\
&+ \left|\sin(W^{(1)}u\pi + V^{(1)}v\pi) - \sin(W^{(1)}_{M_n}u\pi + V^{(1)}_{M_n}v\pi)\right| \\
&\cdot [\mathbb{1}_{\{|W^{(1)}|\geq M_n\}} + \mathbb{1}_{\{|V^{(1)}|\geq M_n\}}]\Big\} \\
\leq &\, 4 \sup_{u,v\in[-n,n]} \mathbf{E}\left\{\mathbb{1}_{\{|W^{(1)}|\geq M_n\}} + \mathbb{1}_{\{|V^{(1)}|\geq M_n\}}\right\} \\
\leq &\, 4\mathbf{P}[|W^{(1)}| \geq M_n] + 4\mathbf{P}[|V^{(1)}| \geq M_n].
\end{aligned}
$$

Da $W^{(1)}$ und $V^{(1)}$ integrabel sind folgt, genau wie in (3.4),

$$
4\mathbf{P}[|W^{(1)}| \geq M_n] \leq 4\mathbf{E}\left\{|W^{(1)}|\right\}\frac{1}{M_n}
$$

und

$$
4\mathbf{P}[|V^{(1)}| \geq M_n] \leq 4\mathbf{E}\left\{|V^{(1)}|\right\}\frac{1}{M_n}.
$$

Damit folgt

$$
\frac{\sqrt{n}}{\log(n)}T_{1,n} \leq \frac{\sqrt{n}}{\log(n)}\frac{4}{M_n}\left(\mathbf{E}\left\{|V^{(1)}|\right\} + \mathbf{E}\left\{|V^{(1)}|\right\}\right).
$$

Somit folgt $\lim_{n\to\infty} \frac{\sqrt{n}}{\log(n)}T_{1,n} = 0$ aus $M_n = \sqrt{n}$.

Um $T_{3,n}$ abzuschätzen, benutzen wir die gleichen Argumente wie für $T_{1,n}$. Es gilt

$$
\begin{aligned}
T_{3,n} &\leq 4\left(\frac{1}{n}\sum_{j=1}^{n}\mathbb{1}_{\{|W_j^{(1)}|\geq M_n\}} + \frac{1}{n}\sum_{j=1}^{n}\mathbb{1}_{\{|V_j^{(1)}|\geq M_n\}}\right) \\
&\leq \frac{4}{M_n}\left(\frac{1}{n}\sum_{j=1}^{n}|W_j^{(1)}| + |V_j^{(1)}|\right).
\end{aligned} \tag{5.2}
$$

Nach dem starken Gesetz der großen Zahlen konvergiert der Faktor

$$\frac{1}{n}\sum_{j=1}^{n}|W_j^{(1)}| + |V_j^{(1)}|$$

mit Wahrscheinlichkeit 1 und ist daher auch mit Wahrscheinlichkeit 1 beschränkt. Also folgt

$$\frac{\sqrt{n}}{\log(n)}T_3 \to 0$$

mit Wahrscheinlichkeit 1.

Um die Konvergenz von $T_{2,n}$ zu beweisen, diskretisieren wir. Um den Diskretisierungsfehler abzuschätzen, verwenden wir die Lipschitz-Stetigkeit von $\sin(x)$ und $\cos(x)$ sowie, dass $W_{M_n}^{(1)}$ und $V_{M_n}^{(1)}$ durch M_n beschränkt sind. Damit folgt

$$\begin{aligned}
&\big|\varphi_{(W_{M_n}^{(1)},V_{M_n}^{(1)})}(u,v) - \varphi_{(W_{M_n}^{(1)},V_{M_n}^{(1)})}(u_0,v_0)\big| \\
&\quad\le |\varphi_{(W_{M_n}^{(1)},V_{M_n}^{(1)})}(u,v) - \varphi_{(W_{M_n}^{(1)},V_{M_n}^{(1)})}(u,v_0)\big| \\
&\qquad + |\varphi_{(W_{M_n}^{(1)},V_{M_n}^{(1)})}(u,v_0) - \varphi_{(W_{M_n}^{(1)},V_{M_n}^{(1)})}(u_0,v_0)\big| \\
&\quad\le |\cos(W_{M_n}^{(1)}u + V_{M_n}^{(1)}v) - \cos(W_{M_n}^{(1)}u + V_{M_n}^{(1)}v_0)| \\
&\qquad + |\sin(W_{M_n}^{(1)}u + V_{M_n}^{(1)}v) - \sin(W_{M_n}^{(1)}u + V_{M_n}^{(1)}v_0)| \\
&\qquad + |\cos(W_{M_n}^{(1)}u + V_{M_n}^{(1)}v_0) - \cos(W_{M_n}^{(1)}u_0 + V_{M_n}^{(1)}v_0)| \\
&\qquad + |\sin(W_{M_n}^{(1)}u + V_{M_n}^{(1)}v_0) - \sin(W_{M_n}^{(1)}u_0 + V_{M_n}^{(1)}v_0)| \\
&\quad\le 2|W_{M_n}^{(1)}u + V_{M_n}^{(1)}v - W_{M_n}^{(1)}u - V_{M_n}^{(1)}v_0| \\
&\qquad + 2|W_{M_n}^{(1)}u + V_{M_n}^{(1)}v_0 - W_{M_n}^{(1)}u_0 - V_{M_n}^{(1)}v_0| \\
&\quad\le 2M_n\big(|u-u_0| + |v-v_0|\big).
\end{aligned}$$

Völlig analog folgt

$$\big|\hat{\varphi}_{(X_{M_n}^{(1)},Y_{M_n}^{(1)})_1^n}(u,v) - \hat{\varphi}_{(X_{M_n}^{(1)},Y_{M_n}^{(1)})_1^n}(u_0,v_0)\big| \le 2M_n\big(|u-u_0| + |v-v_0|\big).$$

Sei

$$I_j := [-n,n] \cap [j/n - 1/2n, j/n + 1/2n] \quad (j \in \mathbb{Z},\ |j| \le n^2)$$

eine Zerlegung von $[-n,n]$ in Intervalle von höchstens Länge $\frac{1}{n}$. Zusammen mit unseren Überlegungen zur Lipschitz-Stetigkeit folgt

$$\sup_{u\in I_j, v\in I_k}\big|\varphi_{(W_{M_n}^{(1)},V_{M_n}^{(1)})}(u\pi,v\pi) - \varphi_{(W_{M_n}^{(1)},V_{M_n}^{(1)})}(\frac{j}{n}\pi,\frac{k}{n}\pi)\big| \le \frac{2\pi M_n}{n},$$

sowie eine analoge Abschätzunge für den Diskretisierungsfehler von $\hat{\varphi}_{(W_{M_n}^{(1)},V_{M_n}^{(1)})_1^n}$. Da $\frac{M_n}{n} = \frac{1}{\sqrt{n}}$ gegen 0 konvergiert, existiert also für jedes $C > 0$ ein $n_0 \in \mathbb{N}$, so dass für alle $n \ge n_0$,

$$\frac{\sqrt{n}}{\log(n)}\sup_{u,v\in[-n,n]}\left|\varphi_{(W_{M_n}^{(1)},V_{M_n}^{(1)})}(u\pi,v\pi) - \hat{\varphi}_{(W_{M_n}^{(1)},V_{M_n}^{(1)})_1^n}(u\pi,v\pi)\right| > C$$

die Existenz eines $(\frac{j}{n}, \frac{k}{n})$ mit $|j|, |k| \leq n^2$ und

$$\frac{\sqrt{n}}{\log(n)} \left| \varphi_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})}(\frac{j}{n} \cdot \pi, \frac{k}{n} \cdot \pi) - \hat{\varphi}_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})^n_1}(\frac{j}{n} \cdot \pi, \frac{k}{n} \cdot \pi) \right| > \frac{C}{2} \tag{5.3}$$

impliziert. Also gil

$$\begin{aligned}
&\mathbf{P}\left\{ \frac{\sqrt{n}}{\log(n)} \sup_{u,v \in [-n,n]} \left| \varphi_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})}(u\pi, v\pi) - \hat{\varphi}_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})^n_1}(u\pi, v\pi) \right| > C \right\} \\
&\leq \mathbf{P}\Big\{ \exists j \in \mathbb{Z} \cap [-n^2, n^2],\ \exists k \in \mathbb{Z} \cap [-n^2, n^2] : \\
&\quad \frac{\sqrt{n}}{\log(n)} \left| \varphi_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})}(\frac{j}{n} \cdot \pi, \frac{k}{n} \cdot \pi) - \hat{\varphi}_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})^n_1}(\frac{j}{n} \cdot \pi, \frac{k}{n} \cdot \pi) \right| > \frac{C}{2} \Big\} \\
&\leq \mathbf{P}\Big\{ \bigcup_{\substack{j \in \mathbb{Z} \cap [-n^2, n^2] \\ k \in \mathbb{Z} \cap [-n^2, n^2]}} \frac{\sqrt{n}}{\log(n)} \big| \varphi_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})}(\frac{j}{n} \cdot \pi, \frac{k}{n} \cdot \pi) \\
&\quad - \hat{\varphi}_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})^n_1}(\frac{j}{n} \cdot \pi, \frac{k}{n} \cdot \pi) \big| > \frac{C}{2} \Big\} \\
&\leq \sum_{\substack{j \in \mathbb{Z} \cap [-n^2, n^2] \\ k \in \mathbb{Z} \cap [-n^2, n^2]}} \mathbf{P}\Big\{ \big| \varphi_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})}(\frac{j}{n} \cdot \pi, \frac{k}{n} \cdot \pi) \\
&\quad - \hat{\varphi}_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})^n_1}(\frac{j}{n} \cdot \pi, \frac{k}{n} \cdot \pi) \big| > \frac{C}{2} \cdot \frac{\log(n)}{\sqrt{n}} \Big\} \\
&\leq (2n^2 + 1)^2 \cdot \max_{\substack{j \in \mathbb{Z} \cap [-n^2, n^2] \\ k \in \mathbb{Z} \cap [-n^2, n^2]}} \mathbf{P}\Big\{ \Big| \varphi_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})}(\frac{j}{n} \cdot \pi, \frac{k}{n} \cdot \pi) \\
&\quad - \hat{\varphi}_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})^n_1}(\frac{j}{n} \cdot \pi, \frac{k}{n} \cdot \pi) \Big| > \frac{C}{2} \cdot \frac{\log(n)}{\sqrt{n}} \Big\}.
\end{aligned} \tag{5.4}$$

Nun zerlegen wir φ und $\hat{\varphi}$ in Real- und Imaginärteil und wenden die Ungleichung von Hoeffding (siehe beispielsweise [17], Lemma A.3) an. Für $\varepsilon > 0$ gilt

$$\begin{aligned}
&\mathbf{P}\left\{ \left| \varphi_{(W^{(1)}, V^{(1)})}(u, v) - \hat{\varphi}_{(W^{(1)}, V^{(1)})^n_1}(u, v) \right| > \varepsilon \right\} \\
&\leq \mathbf{P}\left\{ \big| Re\big(\varphi_{(W^{(1)}, V^{(1)})}(u, v) - \hat{\varphi}_{(W^{(1)}, V^{(1)})^n_1}(u, v) \big) \big| > \frac{\varepsilon}{\sqrt{2}} \right\} \\
&\quad + \mathbf{P}\left\{ \big| Im\big(\varphi_{(W^{(1)}, V^{(1)})}(u, v) - \hat{\varphi}_{(W^{(1)}, V^{(1)})^n_1}(u, v) \big) \big| > \frac{\varepsilon}{\sqrt{2}} \right\} \\
&\leq 2 \cdot \left(2e^{\frac{-n \cdot \varepsilon^2}{4}} \right) = 4e^{\frac{-n \cdot \varepsilon^2}{4}}.
\end{aligned}$$

Zusammen mit der vorherigen Abschätzung erhalten wird

$$\begin{aligned}
&\mathbf{P}\left\{ \frac{\sqrt{n}}{\log(n)} \sup_{u,v \in [-n,n]} \left| \varphi_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})}(u\pi, v\pi) - \hat{\varphi}_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})^n_1}(u\pi, v\pi) \right| > C \right\} \\
&\leq (2n^2 + 1)^2 \cdot \max_{\substack{j \in \mathbb{Z} \cap [-n^2, n^2] \\ k \in \mathbb{Z} \cap [-n^2, n^2]}} \mathbf{P}\Big\{ \sup_{u,v \in [-n,n]} \Big| \varphi_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})}(\frac{j}{n} \cdot \pi, \frac{k}{n} \cdot \pi)
\end{aligned}$$

$$- \hat{\varphi}_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})^n_1}(\frac{j}{n} \cdot \pi, \frac{k}{n} \cdot \pi)\Big| > \frac{C}{2} \cdot \frac{\log(n)}{\sqrt{n}}\}$$

$$\leq 4 \cdot (2n^2+1)^2 \cdot e^{-n \cdot \left(\frac{C\log(n)}{2\sqrt{n}}\right)^2 \cdot \frac{1}{4}}$$

$$= 4 \cdot (2n^2+1)^2 \cdot e^{-C^2 \log(n)^2/16}.$$

Da die Folge $4 \cdot (2n^2+1)^2 \cdot e^{-C^2 \log(n)^2/16}$ sumierbar ist, folgt daraus, dass für jedes $C > 0$

$$\sum_{n=1}^{\infty} \mathbf{P}\left\{ \frac{\sqrt{n}}{\log(n)} \sup_{u,v \in [-n,n]} \left| \varphi_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})}(u\pi, v\pi) - \hat{\varphi}_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})^n_1}(u\pi, v\pi) \right| > C \right\}$$
$$< \infty$$

gilt. Aus dem ersten Lemma von Borel-Cantelli folgt damit, dass

$$\frac{\sqrt{n}}{\log(n)} \sup_{u,v \in [-n,n]} \left| \varphi_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})}(u\pi, v\pi) - \hat{\varphi}_{(W^{(1)}_{M_n}, V^{(1)}_{M_n})^n_1}(u\pi, v\pi) \right|$$

fast sicher gegen 0 konvergiert. Zusammen mit der Konvergenz von $T_{1,n}$ und $T_{3,n}$ impliziert dies die Behauptung des Lemmas. □

Lemma 5.2. *Seien $V = (V^{(1)}, \dots, V^{(d)})$ und $W = (W^{(1)}, \dots, W^{(l)})$ $\mathbb{R}^d$- beziehungsweise $\mathbb{R}^l$-wertige Zufallsvariablen und seien $V^{(1)}$, $W^{(1)}$, $V^{(2)}$ und $W^{(2)}$ quadratisch integrierbar. Außerdem seien (V, W),$(V_1, W_1), \dots$ unabhängig und indentisch verteilt. Dann gilt*

$$\frac{\sqrt[4]{n}}{\log(n)} \sup_{u_1,v_1 \in [-n,n]} \Big| \frac{\partial}{\partial u_2} \varphi_{(V,W)}(u_1\pi, 0, \dots, 0, v_1\pi, 0, \dots, 0)$$
$$- \frac{\partial}{\partial u_2} \hat{\varphi}_{(V,W)^n_1}(u_1\pi, 0, \dots, 0, v_1\pi, 0, \dots, 0) \Big| \to 0$$

fast sicher. Die analoge Aussage gilt, wenn man $\frac{\partial}{\partial u_2}$ durch $\frac{\partial}{\partial v_2}$ ersetzt.

Bemerkung. Wie schon Lemma 5.1 lässt sich auch Lemma 5.2 auf die Zufallsvariablen Z_1, Z_2, ε_1, δ_1 sowie $X^{(1)}$ und $Y^{(1)}$ aus Definition 4.2 anwenden. Es gilt

$$\frac{\sqrt[4]{n}}{\log(n)} \sup_{u_1,v_1 \in [-2\hat{K}_n, 2\hat{K}_n]} \Big| \frac{\partial}{\partial u_2} \varphi_{(X,Y)}(\frac{u_1\pi}{\hat{L}^*_n}, 0, \dots, 0, \frac{v_1\pi}{\hat{L}^*_n}, 0, \dots, 0)$$
$$- \frac{\partial}{\partial u_2} \hat{\varphi}_{(X,Y)^n_1}(\frac{u_1\pi}{\hat{L}^*_n}, 0, \dots, 0, \frac{v_1\pi}{\hat{L}^*_n}, 0, \dots, 0) \Big| \to 0 \quad \text{f.s.}$$

Ähliche Aussagen gelten für $\frac{\partial}{\partial v_2}\varphi_{(X,Y)}$, $\frac{\partial}{\partial z_1}\varphi_{(Z_1,Z_2)}$ und $\frac{\partial}{\partial z_2}\varphi_{(Z_1,Z_2)}$.

Beweis von Lemma 5.2. Wie aus der Wahrscheinlichkeitstheorie bekannt ist, ist die charakteristische Funktion einer Zufallsvariablen k-mal partiell differenzierbar, falls die absoluten Momente k-ter Ordnung endlich sind (siehe [40, Theorem 25.2]). Nach Voraussetzung sind $V^{(2)}$ und $W^{(2)}$ quadratisch integrierbar. Die partielle Ableitung von $\varphi_{(V,W)}$ existieren also und besitzen die Form

$$\frac{\partial}{\partial u_2} \varphi_{(V,W)}(u_1, 0, \dots, 0, v_1, 0, \dots, 0) = \mathbf{E}\Big\{ iV^{(2)} e^{i(u_1 V^{(1)} + 0 \cdot V^{(2)} + \dots + v_1 W^{(1)} + \dots + 0)} \Big\}$$

$$= \mathbf{E}\Big\{ iV^{(2)} e^{i(u_1 V^{(1)} + v_1 W^{(1)})} \Big\}.$$

Auch die partielle Ableitung der empirischen charakteristischen Funktion lässt sich einfach berechnen. Es gilt

$$\frac{\partial}{\partial u_2} \hat{\varphi}_{(V,W)_1^n}(u_1, 0, \ldots, 0, v_1, 0, \ldots, 0) = \frac{1}{n} \sum_{j=1}^{n} iV_j^{(2)} e^{i(u_1 V_j^{(1)} + v_1 W_j^{(1)})}.$$

Im folgenden gehen wir vor wie im Beweis von Lemma 5.1. Mit $M_n := \sqrt[4]{n}$ gilt

$$\begin{aligned}
&\sup_{u,v \in [-n,n]} \Big| \mathbf{E}\Big\{ iV^{(2)} e^{i(u\pi V^{(1)} + v\pi W^{(1)})} \Big\} - \frac{1}{n} \sum_{j=1}^{n} iV_j^{(2)} e^{i(u\pi V_j^{(1)} + v\pi W_j^{(1)})} \Big| \\
&\leq \sup_{u,v \in [-n,n]} \Big| \mathbf{E}\Big\{ iV^{(2)} e^{i\pi(uV^{(1)} + vW^{(1)})} \Big\} - \mathbf{E}\Big\{ iV_{M_n}^{(2)} e^{i\pi(uV_{M_n}^{(1)} + vW_{M_n}^{(1)})} \Big\} \Big| \\
&\quad + \sup_{u,v \in [-n,n]} \Big| \mathbf{E}\Big\{ iV_{M_n}^{(2)} e^{i\pi(uV_{M_n}^{(1)} + vW_{M_n}^{(1)})} \Big\} - \frac{1}{n} \sum_{j=1}^{n} iV_{M_n,j}^{(2)} e^{i\pi(uV_{M_n,j}^{(1)} + vW_{M_n,j}^{(1)})} \Big| \\
&\quad + \sup_{u,v \in [-n,n]} \Big| \frac{1}{n} \sum_{j=1}^{n} iV_{M_n,j}^{(2)} e^{i\pi(uV_{M_n,j}^{(1)} + vW_{M_n,j}^{(1)})} - \frac{1}{n} \sum_{j=1}^{n} iV_j^{(2)} e^{i\pi(uV_j^{(1)} + vW_j^{(1)})} \Big|. \\
&= T_{1,n} + T_{2,n} + T_{3,n}.
\end{aligned}$$

Wir beginnen mit $T_{1,n}$. Aus der elementaren Ungleichung

$$|ab - cd| \leq |a - c| \cdot |b| + |c| \cdot |b - d| \quad (a, b, c, d \in \mathbb{R})$$

folgt

$$\begin{aligned}
&\Big| \mathbf{E}\Big\{ iV^{(2)} e^{i\pi(uV^{(1)} + vW^{(1)})} \Big\} - \mathbf{E}\Big\{ iV_{M_n}^{(2)} e^{i\pi(uV_{M_n}^{(1)} + vW_{M_n}^{(1)})} \Big\} \Big| \\
&\leq \mathbf{E}\Big\{ \Big| V^{(2)} \big(\cos(\pi(uV^{(1)} + vW^{(1)})) + i \sin(\pi(uV^{(1)} + vW^{(1)})) \big) \\
&\quad - V_{M_n}^{(2)} \big(\cos(\pi(uV_{M_n}^{(1)} + vW_{M_n}^{(1)})) + i \sin(\pi(uV_{M_n}^{(1)} + vW_{M_n}^{(1)})) \big) \Big| \\
&\leq \mathbf{E}\Big\{ \Big| V^{(2)} - V_{M_n}^{(2)} \Big| \, \Big| \cos(\pi(uV^{(1)} + vW^{(1)})) + i \sin(\pi(uV^{(1)} + vW^{(1)})) \Big| \\
&\quad + \Big| V_{M_n}^{(2)} \Big| \, \Big| \cos(\pi(uV^{(1)} + vW^{(1)})) + i \sin(\pi(uV^{(1)} + vW^{(1)})) \\
&\quad - \cos(\pi(uV_{M_n}^{(1)} + vW_{M_n}^{(1)})) + i \sin(\pi(uV_{M_n}^{(1)} + vW_{M_n}^{(1)})) \Big| \Big\} \\
&\leq \mathbf{E}\Big\{ \Big| V^{(2)} - V_{M_n}^{(2)} \Big| \Big[\Big| \cos(\pi(uV^{(1)} + vW^{(1)})) \Big| + \Big| \sin(\pi(uV^{(1)} + vW^{(1)})) \Big| \Big] \Big\} \\
&\quad + \Big| V_{M_n}^{(2)} \Big| \Big[\Big| \cos(\pi(uV^{(1)} + vW^{(1)})) - \cos(\pi(uV_{M_n}^{(1)} + vW_{M_n}^{(1)})) \Big| \\
&\quad + \Big| \sin(\pi(uV^{(1)} + vW^{(1)})) - \sin(\pi(uV_{M_n}^{(1)} + vW_{M_n}^{(1)})) \Big| \Big] \\
&\leq 2\mathbf{E}\Big\{ \Big| V^{(2)} - V_{M_n}^{(2)} \Big| \mathbb{1}_{\{|V^{(2)}| \geq M_n\}} \Big\} \\
&\quad M_n \Big[\Big| \pi(uV^{(1)} + vW^{(1)}) - \pi(uV_{M_n}^{(1)} + vW_{M_n}^{(1)}) \Big| \\
&\quad \Big| \pi(uV^{(1)} + vW^{(1)}) - \pi(uV_{M_n}^{(1)} + vW_{M_n}^{(1)}) \Big| \Big]
\end{aligned}$$

$$\leq 2\mathbf{E}\Big\{\Big|V^{(2)} - V^{(2)}_{M_n}\Big| \mathbb{1}_{\{|V^{(2)}| \geq M_n\}}\Big\}$$
$$+ 4M_n\mathbf{E}\{\mathbb{1}_{\{|V^{(1)}| \geq M_n\}} + \mathbb{1}_{\{|W^{(1)}| \geq M_n\}}\}$$
$$\leq 2\mathbf{E}\left\{|V^{(2)}|\frac{|V^{(2)}|}{M_n}\right\}$$
$$+ 4M_n\mathbf{E}\left\{\frac{|V^{(1)}|^2}{M_n^2}\right\} + 4M_n\mathbf{E}\left\{\frac{|W^{(1)}|^2}{M_n^2}\right\}$$
$$\leq \frac{C}{M_n}$$

für ein geeignetes $C \in \mathbb{R}$, da $V^{(1)}$, $V^{(2)}$ und $W^{(1)}$ nach Voraussetzung quadratisch integrierbar sind. Also gilt auch

$$\frac{\sqrt[4]{n}}{\log(n)}T_{1,n} \leq C\frac{1}{\log(n)} \to 0.$$

Wie schon im Beweis von Lemma 3.5 verwenden wir zur Abschätzung von $T_{3,n}$ die selben Argumente wie zur Abschätzung von $T_{1,n}$. Es gilt

$$\Big|\frac{1}{n}\sum_{j=1}^{n} iV^{(2)}_{M_n,j}e^{i\pi(uV^{(1)}_{M_n,j}+vW^{(1)}_{M_n,j})} - \frac{1}{n}\sum_{j=1}^{n} iV^{(2)}_j e^{i\pi(uV^{(1)}_j+vW^{(1)}_j)}\Big|$$
$$\leq 2\sum_{j=1}^{n}|V^{(2)}_j|\frac{|V^{(2)}_j|}{M_n}$$
$$+ 4M_n\sum_{j=1}^{n}\frac{|V^{(1)}_j|^2}{M_n^2} + 4M_n\sum_{j=1}^{n}\frac{|W^{(1)}_j|^2}{M_n^2}.$$

Da die auftretenden Summen nach dem starken Gesetz der großen Zahlen jeweils fast sicher konvergieren, konvergiert auch $\frac{\sqrt[4]{n}}{\log(n)}T_{3,n}$ fast sicher gegen 0.

Um die Konvergenz von $T_{2,n}$ zu zeigen, diskretisieren wir wieder. Es gilt die Lipschitz-Abschätzung

$$|iX^{(2)}_{M_n}e^{i\pi(uX^{(1)}_{M_n}+vY^{(1)}_{M_n})} - iX^{(2)}_{M_n}e^{i\pi(u_0X^{(1)}_{M_n}+v_0Y^{(1)}_{M_n})}|$$
$$\leq 2\pi(M_n)^2(|u-u_0| + |v-v_0|).$$

Sei nun wieder

$$I_j := [-n,n] \cap [j/n - 1/2n, j/n + 1/2n] \quad (j \in \mathbb{Z},\ |j| \leq n^2)$$

eine Zerlegung von $[-n,n]$ in Intervalle von höchstens Länge $\frac{1}{n}$. Dann ist der maximale Diskretisierungsfehler

$$\sup_{u\in I_j, v\in I_k}\left|iX^{(2)}_{M_n}e^{i\pi(uX^{(1)}_{M_n}+vY^{(1)}_{M_n})} - iX^{(2)}_{M_n}e^{i\pi(u_0X^{(1)}_{M_n}+v_0Y^{(1)}_{M_n})}\right|$$

durch $K\frac{(M_n)^2}{n} = \frac{K}{\sqrt{n}}$ für eine geeignete Konstante K. Sein $C > 0$. Für n groß genug gilt dann

$$\frac{1}{\sqrt{n}} \leq \frac{C}{2}\cdot\frac{\log(n)}{\sqrt[4]{n}}.$$

Für solche n impliziert

$$\frac{\sqrt[4]{n}}{\log(n)} \sup_{u,v\in[-n,n]} \Big|\mathbf{E}\Big\{iV^{(2)}_{M_n} e^{i\pi(uV^{(1)}_{M_n}+vW^{(1)}_{M_n})}\Big\} - \frac{1}{n}\sum_{j=1}^{n} iV^{(2)}_{M_n,j} e^{i\pi(uV^{(1)}_{M_n,j}+vW^{(1)}_{M_n,j})}\Big| > C$$

die Existenz eines $(\frac{j}{n},\frac{k}{n})$ mit $|j|,|k|\le n^2$ für welches

$$\frac{\sqrt[4]{n}}{\log(n)}\Big|\mathbf{E}\Big\{iV^{(2)}_{M_n} e^{i\pi(\frac{j}{n}V^{(1)}_{M_n}+\frac{k}{n}W^{(1)}_{M_n})}\Big\} - \frac{1}{n}\sum_{j=1}^{n} iV^{(2)}_{M_n,j} e^{i\pi(\frac{j}{n}V^{(1)}_{M_n,j}+\frac{k}{n}W^{(1)}_{M_n,j})}\Big| > \frac{C}{2}$$

gilt. Nun wenden wir wieder die gleichen Methoden wie in der Abschätzung (5.4) sowie die Hoeffding-Ungleichung an und erhalten

$$\begin{aligned}
&\mathbf{P}\Big\{\frac{\sqrt[4]{n}}{\log(n)} \sup_{u,v\in[-n,n]} \Big|\mathbf{E}\Big\{iX^{(2)}_{M_n} e^{i(uX^{(1)}_{M_n}+vY^{(1)}_{M_n})}\Big\} \\
&\qquad - \frac{1}{n}\sum_{j=1}^{n} iX^{(2)}_{M_n,j} e^{i(uX^{(1)}_{M_n,j}+vY^{(1)}_{M_n,j})}\Big| \ge C\Big\} \\
&\le (2n^2+1)^2 \cdot \max_{\substack{j\in\mathbb{Z}\cap[-n^2,n^2]\\ k\in\mathbb{Z}\cap[-n^2,n^2]}} \mathbf{P}\Big\{\Big|\mathbf{E}\{iX^{(2)}_{M_n} e^{i(\frac{j}{n}X^{(1)}_{M_n}+\frac{k}{n}Y^{(1)}_{M_n})}\} \\
&\qquad - \frac{1}{n}\sum_{j=1}^{n} iX^{(2)}_{M_n,j} e^{i(\frac{j}{n}X^{(1)}_{M_n,j}+\frac{k}{n}Y^{(1)}_{M_n,j})}\Big| > \frac{C}{2}\cdot\frac{\log(n)}{\sqrt[4]{n}}\Big\} \\
&\le 4\cdot(2n^2+1)^2\cdot \exp\Big(-n\big(\frac{C\cdot\log(n)}{2\sqrt[4]{n}}\big)^2\cdot\frac{1}{4(M_n)^2}\Big) \\
&\le C'n^3\cdot\exp\big(-C''\log(n)^2\big)
\end{aligned}$$

für zwei geeignete Konstanten $C', C'' \in \mathbb{R}^+$. Nun gilt allerdings

$$\sum_{n=1}^{\infty} C'n^3\cdot\exp\big(-C''\log(n)^2\big) < \infty$$

für alle $C', C'' \in \mathbb{R}^+$. Also folgt aus dem ersten Lemma von Borel-Cantelli, dass mit Wahrscheinlichkeit 1 auch $\frac{\sqrt[4]{n}}{\log(n)}T_{2,n}$ gegen 0 konvergiert. Damit folgt die erste Aussage des Lemmas. Die zweite Aussage des Lemmas folgt völlig analog. □

Lemma 5.3. *Sei $\hat{K}_n$ definiert wie in (4.3). Sind $X^{(1)}$ und $Y^{(1)}$ absolut integrabel und verschwindet $\varphi_{(X^{(1)},Y^{(1)})}$ nirgends, so gilt*

$$\hat{K}_n \to \infty$$

fast sicher.

Beweis. Es genügt zu zeigen, dass

$$\max\{K\in\mathbb{N}: \inf_{u,v\in[-2\cdot\sqrt{K},2\cdot\sqrt{K}]} \big|\hat{\varphi}_{(X^{(1)},Y^{(1)})_1^n}(u\pi, v\pi)\big| \ge \frac{1}{\log(n)}\}$$

fast sicher bestimmt gegen ∞ divergiert. Dazu genügt es zu zeigen, dass für jedes Intervall $I \subset \mathbb{R}$ mit Wahrscheinlichkeit 1 ein n_0 existiert, so dass

$$\inf_{u,v \in I} \left|\hat{\varphi}_{(X^{(1)},Y^{(1)})_1^n}(u\pi, v\pi)\right| \geq \frac{1}{\log(n)}$$

für alle $n \geq n_0$ gilt. Hierzu verwenden wir Lemma 5.1. Da $X^{(1)}$ und $Y^{(1)}$ absolut integrabel sind, gilt

$$\frac{\sqrt{n}}{\log(n)} \sup_{u,v \in [-n,n]} \left|\varphi_{(X^{(1)},Y^{(1)})}(u\pi, v\pi) - \hat{\varphi}_{(X^{(1)},Y^{(1)})_1^n}(u\pi, v\pi)\right| \to 0$$

mit Wahrscheinlichkeit 1. Außerdem ist $\varphi_{(X_1,Y_1)}$ stetig und besitzt keine Nullstellen. Also existiert für jedes Interval I eine Konstante $C_I > 0$, so dass

$$\varphi_{(X_1,Y_1)}(u, v) > C_I$$

für alle $(u, v) \in I^2$ gilt. Also gilt mit Wahrscheinlichkeit 1, für n groß genug,

$$\frac{1}{\log n} < \frac{C_I}{2} \quad \text{und} \quad \sup_{u,v \in I} |\hat{\varphi}_{(X_1,Y_1)_1^n}(u, v) - \varphi_{(X_1,Y_1)}(u, v)| \leq \frac{C_I}{2}.$$

Dies wiederum impliziert, dass mit Wahrscheinlichkeit 1, für n groß genug, die Ungleichung

$$\begin{aligned}
&\min_{u,v \in I} |\hat{\varphi}_{(X^{(1)},Y^{(1)})_1^n}(u, v)| \\
&\geq \min_{u,v \in I} |\varphi_{(X_1,Y_1)}(u, v)| - \sup_{u,v \in I} |\hat{\varphi}_{(X_1,Y_1)_1^n}(u, v) - \varphi_{(X_1,Y_1)}(u, v)| \\
&\geq C_I - \frac{1}{2} C_I > \frac{1}{\log(n)}
\end{aligned}$$

gilt. Da I beliebig war, beendet dies den Beweis. □

Lemma 5.4. *Es seien $a_3 \neq 0$ und $b_3 \neq 0$ und seien $\mathbf{E}\{(X^{(1)})^4\}, \ldots, \mathbf{E}\{(Y^{(l)})^4\}$ jeweils endlich. Außerdem seien $\mathbf{E}Z_1^2 > 0$, $\mathbf{E}Z_2^2 > 0$. Dann gilt*

$$\frac{\sqrt{n}}{\log(n)} |\hat{a}_2 - a_2| \to 0 \quad \textit{und} \quad \frac{\sqrt{n}}{\log(n)} |\hat{b}_2 - b_2| \to 0$$

fast sicher.

Beweis. Für $\hat{a}_2$ wurde die Aussage bereits in Lemma 3.7 gezeigt. Der Beweis lässt sich ohne Änderungen auf das in Definition 4.2 vorgestellte Modell übertragen. Die Aussage für $\hat{b}_2$ lässt sich völlig analog beweisen. □

Lemma 5.5. *Seien die Annahmen 1) bis 5) aus Lemma 4.1 erfüllt. Außerdem seien die 6-ten Momente von Z_1 und Z_2 sowie $\varepsilon_1, \ldots, \delta_l$ endlich. Ist T_n definiert wie in (4.4) und das geschätzte Sample $(\hat{z}_1, \hat{z}_2)_1^n$ definiert als Minimum von T_n unter der Nebenbedingung $\frac{1}{n} \cdot \sum_{j=1}^n |\hat{z}_{2,j}|^6 \leq \sqrt{\hat{K}_n}$, so gilt*

$$\frac{\sqrt[4]{n}}{\log(n)} T_n((\hat{z}_1, \hat{z}_2)_1^n) \to 0$$

fast sicher.

Beweis. Wir wiederhohlen die Definiton von T_n. Es gilt

$$
\begin{aligned}
&T_n((\hat{z}_1,\hat{z}_2)_1^n) := \\
&\sup_{u_1,v_1\in[-2\hat{K}_n,2\hat{K}_n]} \Bigg\{ \Bigg| \hat{\varphi}_{(X^{(1)},Y^{(1)})_1^n}(\frac{\pi u_1}{\hat{L}_n^*},\frac{\pi v_1}{\hat{L}_n^*}) \\
&\qquad - \hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(\frac{\pi u_1}{\hat{L}_n^*},\frac{\pi v_1}{\hat{L}_n^*}) \cdot \hat{\varphi}_{\hat{\varepsilon}_{1,1}^n}(\frac{\pi u_1}{\hat{L}_n^*}) \cdot \hat{\varphi}_{\hat{\delta}_{1,1}^n}(\frac{\pi v_1}{\hat{L}_n^*}) \Bigg| \\
&+ \Bigg| \frac{\partial}{\partial u_2}\hat{\varphi}_{(X,Y)_1^n}(\frac{\pi u_1}{\hat{L}_n^*},0,\dots,0,\frac{\pi v_1}{\hat{L}_n^*},0,\dots,0) \\
&\qquad - \hat{a}_2 \cdot \frac{\partial}{\partial u_1}\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(\frac{\pi u_1}{\hat{L}_n^*},\frac{\pi v_1}{\hat{L}_n^*}) \cdot \hat{\varphi}_{\hat{\varepsilon}_{1,1}^n}(\frac{\pi u_1}{\hat{L}_n^*}) \cdot \hat{\varphi}_{\hat{\delta}_{1,1}^n}(\frac{\pi v_1}{\hat{L}_n^*}) \Bigg| \\
&+ \Bigg| \frac{\partial}{\partial v_2}\hat{\varphi}_{(X,Y)_1^n}(\frac{\pi u_1}{\hat{L}_n^*},0,\dots,0,\frac{\pi v_1}{\hat{L}_n^*},0,\dots,0) \\
&\qquad - \hat{b}_2 \cdot \frac{\partial}{\partial v_1}\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(\frac{\pi u_1}{\hat{L}_n^*},\frac{\pi v_1}{\hat{L}_n^*}) \cdot \hat{\varphi}_{\hat{\varepsilon}_{1,1}^n}(\frac{\pi u_1}{\hat{L}_n^*}) \cdot \hat{\varphi}_{\hat{\delta}_{1,1}^n}(\frac{\pi v_1}{\hat{L}_n^*}) \Bigg| \Bigg\}.
\end{aligned}
$$

Wobei $\hat{z}_1$ und $\hat{z}_2$ als Minimum von T_n unter der Nebenbedingung

$$
\frac{1}{n}\sum_{j=1}^{n}|\hat{z}_{2,j}|^6 \le \sqrt{\hat{K}_n}
$$

gewählt sind. Sind nun $(Z_1,Z_2),(Z_{1,1},Z_{2,1}),\dots,(Z_{1,n},Z_{2,n})$ unabhängig und identisch verteilt, so ist $\tilde{T}_n$ gegeben, indem wir die $\hat{z}_{1,j}$ durch $Z_{1,j}$, die $\hat{z}_{2,j}$ durch $Z_{2,j}$, und die $\hat{\varepsilon}_{1,j} = X_j^{(1)} - \hat{z}_{1,j}$ durch $\tilde{\varepsilon}_{1,j} := X_j^{(1)} - Z_{1,j}$ und die $\hat{\delta}_{1,j} = Y_j^{(1)} - \hat{z}_{2,j}$ durch $\tilde{\delta}_{1,j} := Y_j^{(1)} - Z_{2,j}$ ersetzen. Nach Voraussetzung gilt $\mathbf{E}Z_2^6 < \infty$ und $\frac{1}{n}\sum_{j=1}^{n} Z_{2,j}^6$ konvergiert gegen $\mathbf{E}Z_2^6$ fast sicher. Wir verwenden Lemma 5.3, und erhalten $\hat{K}_n \to \infty$ fast sicher. Mit Wahrscheinlichkeit 1 gilt also

$$
T_n \le \tilde{T}_n
$$

für n groß genug. Die aus Lemma 4.1 bekannten Gleichungen für die charakteristischen Funktionen von (X,Y) und (Z_1,Z_2) liefern nun

$$
\begin{aligned}
\tilde{T}_n \le &\sup_{u_1,v_1\in[-2\hat{K}_n,2\hat{K}_n]} \Bigg\{ \Bigg| \hat{\varphi}_{(X^{(1)},Y^{(1)})_1^n}(\frac{\pi u_1}{\hat{L}_n^*},\frac{\pi v_1}{\hat{L}_n^*}) - \varphi_{(X^{(1)},Y^{(1)})}(\frac{\pi u_1}{\hat{L}_n^*},\frac{\pi v_1}{\hat{L}_n^*}) \Bigg| \\
&+ \Bigg| \varphi_{(Z_1,Z_2)}(\frac{\pi u_1}{\hat{L}_n^*},\frac{\pi v_1}{\hat{L}_n^*}) \cdot \varphi_{\varepsilon_1}(\frac{\pi u_1}{\hat{L}_n^*}) \cdot \varphi_{\delta_1}(\frac{\pi v_1}{\hat{L}_n^*}) \\
&\qquad - \hat{\varphi}_{(Z_1,Z_2)_1^n}(\frac{\pi u_1}{\hat{L}_n^*},\frac{\pi v_1}{\hat{L}_n^*}) \cdot \hat{\varphi}_{\varepsilon_{1,1}^n}(\frac{\pi u_1}{\hat{L}_n^*}) \cdot \hat{\varphi}_{\delta_{1,1}^n}(\frac{\pi v_1}{\hat{L}_n^*}) \Bigg| \\
&+ \Bigg| \frac{\partial}{\partial u_2}\hat{\varphi}_{(X,Y)_1^n}(\frac{\pi u_1}{\hat{L}_n^*},0,\dots,\frac{\pi v_1}{\hat{L}_n^*},0,\dots) \\
&\qquad - \frac{\partial}{\partial u_2}\varphi_{(X,Y)}(\frac{\pi u_1}{\hat{L}_n^*},0,\dots,\frac{\pi v_1}{\hat{L}_n^*},0,\dots) \Bigg| \\
&+ \Bigg| a_2 \frac{\partial}{\partial u}\varphi_{(Z_1,Z_2)}(\frac{\pi u_1}{\hat{L}_n^*},\frac{\pi v_1}{\hat{L}_n^*}) \cdot \varphi_{\varepsilon_1}(\frac{\pi u_1}{\hat{L}_n^*}) \cdot \varphi_{\delta_1}(\frac{\pi v_1}{\hat{L}_n^*})
\end{aligned}
$$

$$
\begin{aligned}
&\qquad - \hat{a}_2 \frac{\partial}{\partial u} \hat{\varphi}_{(Z_1,Z_2)_1^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}, \frac{\pi v_1}{\hat{L}_n^*}\Big) \cdot \hat{\varphi}_{\varepsilon_{1,1}^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}\Big) \cdot \hat{\varphi}_{\delta_{1,1}^n}\Big(\frac{\pi v_1}{\hat{L}_n^*}\Big)\Big| \\
&+ \Big| \frac{\partial}{\partial v_2} \hat{\varphi}_{(X,Y)_1^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}, 0, \dots, \frac{\pi v_1}{\hat{L}_n^*}, 0, \dots\Big) \\
&\qquad - \frac{\partial}{\partial v_2} \varphi_{(X,Y)}\Big(\frac{\pi u_1}{\hat{L}_n^*}, 0, \dots, \frac{\pi v_1}{\hat{L}_n^*}, 0, \dots\Big)\Big| \\
&+ \Big| b_2 \frac{\partial}{\partial v} \varphi_{(Z_1,Z_2)}\Big(\frac{\pi u_1}{\hat{L}_n^*}, \frac{\pi v_1}{\hat{L}_n^*}\Big) \cdot \varphi_{\varepsilon_1}\Big(\frac{\pi u_1}{\hat{L}_n^*}\Big) \cdot \varphi_{\delta_1}\Big(\frac{\pi v_1}{\hat{L}_n^*}\Big) \\
&\qquad - \hat{b}_2 \frac{\partial}{\partial v} \hat{\varphi}_{(Z_1,Z_2)_1^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}, \frac{\pi v_1}{\hat{L}_n^*}\Big) \cdot \hat{\varphi}_{\varepsilon_{1,1}^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}\Big) \cdot \hat{\varphi}_{\delta_{1,1}^n}\Big(\frac{\pi v_1}{\hat{L}_n^*}\Big)\Big| \Big\} \\
&= \sup_{u_1,v_1 \in [-2\hat{K}_n, 2\hat{K}_n]} \sum_{j=1}^{6} T_{j,n} \\
&\leq \sum_{j=1}^{6} \sup_{u_1,v_1 \in [-2\hat{K}_n, 2\hat{K}_n]} T_{j,n}.
\end{aligned}
$$

Um die Summanden weiter abzuschätzen benötigen wir das folgende Hilfsresultat.Ist $k \in \mathbb{N}$ und sind $a_{1,n}, \dots, a_{k,n}$ jeweils reelle Folgen mit den jeweiligen Grenzwerten $a_1, \dots .a_k$, so folgt durch wiederhohltes Anwenden der Dreiecksungleichung die Abschätzung

$$
\left| \prod_{j=1}^{k} a_{j,n} - \prod_{j=1}^{k} a_j \right| \leq \sum_{l=1}^{k} \left[\Big(\prod_{1 \leq j < l} |a_j| \Big) |a_{l,n} - a_l| \Big(\prod_{l < j \leq k} |a_{j,n}| \Big) \right]. \tag{5.5}
$$

Wobei wir das Produkt über die leere Indexmenge als 1 definieren. Nun verwenden wir die Lemmata 5.1, 5.2 und 5.4 um die Konvergenz der obigen Summanden zu beweisen. Da nach Voraussetzung die 6-ten Momente von Z_1,Z_2 und $\varepsilon_1, \dots, \delta_l$ endlich sind und $|\hat{K}| \leq \lfloor n^{\frac{1}{11}} \rfloor$ gilt, sind die Voraussetzungen von Lemma 5.1 für $(X^{(1)}, Y^{(1)})_1^n$ erfüllt. Es folgt

$$
\begin{aligned}
\frac{\sqrt[4]{n}}{\log(n)} T_{1,n} &= \frac{\sqrt[4]{n}}{\log(n)} \sup_{u_1,v_1 \in [-2\hat{K}_n, 2\hat{K}_n]} \Big|\hat{\varphi}_{(X^{(1)},Y^{(1)})_1^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}, \frac{\pi v_1}{\hat{L}_n^*}\Big) \\
&\quad - \varphi_{(X^{(1)},Y^{(1)})}\Big(\frac{\pi u_1}{\hat{L}_n^*}, \frac{\pi v_1}{\hat{L}_n^*}\Big)\Big| \\
&\leq \frac{\sqrt{n}}{\log(n)} \sup_{u_1,v_1 \in [-n,n]} |\hat{\varphi}_{(X^{(1)},Y^{(1)})_1^n}(\pi u_1, \pi v_1) - \varphi_{(X^{(1)},Y^{(1)})}(\pi u_1, \pi v_1)| \to 0
\end{aligned}
$$

fast sicher. Aus Lemma 5.1 folgt ebenfalls, dass

$$
\frac{\sqrt[4]{n}}{\log(n)} \sup_{u_1,v_1 \in [-2\hat{K}_n, 2\hat{K}_n]} \left| \varphi_{(Z_1,Z_2)}\Big(\frac{\pi u_1}{\hat{L}_n^*}, \frac{\pi v_1}{\hat{L}_n^*}\Big) - \hat{\varphi}_{(Z_1,Z_2)_1^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}, \frac{\pi v_1}{\hat{L}_n^*}\Big) \right|
$$

sowie

$$
\frac{\sqrt[4]{n}}{\log(n)} \sup_{u_1,v_1 \in [-2\hat{K}_n, 2\hat{K}_n]} \left| \varphi_{\varepsilon_1}\Big(\frac{\pi u_1}{\hat{L}_n^*}\Big) - \hat{\varphi}_{\varepsilon_{1,1}^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}\Big) \right|
$$

und

$$
\frac{\sqrt[4]{n}}{\log(n)} \sup_{u_1,v_1 \in [-2\hat{K}_n, 2\hat{K}_n]} \left| \varphi_{\delta_1}\Big(\frac{\pi u_1}{\hat{L}_n^*}\Big) - \hat{\varphi}_{\delta_{1,1}^n}\Big(\frac{\pi u_1}{\hat{L}_n^*}\Big) \right|
$$

fast sicher gegen 0 konvergieren. Daraus folgt, zusammen mit dem Hilfsresultat (5.5) , dass auch

$$\frac{\sqrt[4]{n}}{\log(n)}T_{2,n} = \frac{\sqrt[4]{n}}{\log(n)}|\varphi_{(Z_1,Z_2)}(\frac{\pi u_1}{\hat{L}_n^*},\frac{\pi v_1}{\hat{L}_n^*})\cdot\varphi_{\varepsilon_1}(\frac{\pi u_1}{\hat{L}_n^*})\cdot\varphi_{\delta_1}(\frac{\pi v_1}{\hat{L}_n^*})$$
$$-\hat{\varphi}_{(Z_1,Z_2)_1^n}(\frac{\pi u_1}{\hat{L}_n^*},\frac{\pi v_1}{\hat{L}_n^*})\cdot\hat{\varphi}_{\varepsilon_{1,1}^n}(\frac{\pi u_1}{\hat{L}_n^*})\cdot\hat{\varphi}_{\delta_{1,1}^n}(\frac{\pi v_1}{\hat{L}_n^*})|$$

fast sicher gegen 0 konvergiert. Aus der Momentenbedingung für Z_1,Z_2, und $\varepsilon_1,\ldots,\delta_l$ folgt, das (X,Y) auch die Voraussetzungen von Lemma 5.2 erfüllt. Damit folgt die Konvergenz von

$$\frac{\sqrt[4]{n}}{\log(n)}\sup_{u_1,v_1\in[-2\hat{K}_n,2\hat{K}_n]}T_{3,n} \text{ und } \frac{\sqrt[4]{n}}{\log(n)}\sup_{u_1,v_1\in[-2\hat{K}_n,2\hat{K}_n]}T_{5,n}$$

analog zur Konvergenz von $T_{1,n}$. Aus der Wahrscheinlichkeitstheorie ist bekannt, dass die Ableitungen der charakteristischen Funktion einer Zufallsvariablen beschränkt sind, falls die entsprechenden Momente endlich sind (siehe beispielsweise [40, Satz 25.2]). Da die 6-ten Momente von Z_1 und Z_2 endlich sind, sind die Ableitungen $\frac{\partial}{\partial v}\varphi_{(Z_1,Z_2)}$ und $\frac{\partial}{\partial u}\varphi_{(Z_1,Z_2)}$ beschränkt. Zusammen mit dem Hilfsresultat (5.5) sowie Lemma 5.2 und Lemma 5.4 folgt damit die gewünschte Konvergenz von $\frac{\sqrt[4]{n}}{\log(n)}\sup_{u_1,v_1\in[-2\hat{K}_n,2\hat{K}_n]}T_{4,n}$ und $\frac{\sqrt[4]{n}}{\log(n)}\sup_{u_1,v_1\in[-2\hat{K}_n,2\hat{K}_n]}T_{6,n}$. Insgesamt folgt also

$$\frac{\sqrt[4]{n}}{\log(n)}T_n((\hat{z}_1,\hat{z}_2)_1^n)\to 0$$

fast sicher. Damit ist der Beweis beendet. □

Lemma 5.6. *Sei T_n definiert wie in (4.4) und sei $(\hat{z}_1,\hat{z}_2)_1^n$ das Minimum von T_n unter der Nebenbedingung $\frac{1}{n}\cdot\sum_{j=1}^n\hat{z}_{2,j}^6\le\sqrt{\hat{K}_n}$. Sei $\hat{K}_n$ als Funktion der Daten $(X_1,Y_1),\ldots,(X_n,Y_n)$ definiert wie in (4.3). Sei $\beta_n = n^{\frac{1}{11}}$, und $\hat{L}_n^* = 2\hat{L}_n = 2\sqrt{\hat{K}_n}$. Unter den Annahmen von Theorem 4.4 gilt dann*

$$\hat{K}_n^3\beta_n^2\sup_{|u|\le 2\cdot\hat{K}_n,|v|\le 2\cdot\hat{K}_n}\left|\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(\frac{u\pi}{\hat{L}_n^*},\frac{v\pi}{\hat{L}_n^*})-\varphi_{(Z_1,Z_2)}(\frac{u\pi}{\hat{L}_n^*},\frac{v\pi}{\hat{L}_n^*})\right|\to 0$$

fast sicher.

Beweis. Nach den Voraussetzungen von Theorem 4.4 sind die 6-ten Momente von Z_1 und Z_2 sowie von $\varepsilon_1,\ldots,\varepsilon_d$,$\delta_1,\ldots,\delta_l$ endlich. Daher sind auch die 6-ten Momente der Komponenten $X^{(j)}$ von X und $Y^{(k)}$ von Y endlich. Also lassen sich die Lemmata 5.1 und 5.2 auf alle in Theorem 4.4 auftretenden Zufallsvariablen anwenden. Weiterhin sind auch die für Lemma 5.5 benötigten Annahmen an unser Modell Teil der Voraussetzungen. Somit sind Voraussetzungen von Lemma 5.4 und Lemma 5.5 erfüllt. Außerdem wird in Theorem 4.4 vorausgesetzt, dass die charakteristische Funktion von $(X^{(1)},Y^{(1)})$ keine Nullstellen besitzt. Somit ist auch Lemma 5.3 anwendbar. Also können wir die Lemmata 5.1 − 5.5 für den Beweis von Lemma 5.6 verwenden. Nun zum eigentlichen Beweis. Nach Lemma 5.5 gilt

$$\sup_{u,v\in[-2\hat{K}_n,2\hat{K}_n]}|\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(\frac{u\pi}{\hat{L}_n^*},\frac{v\pi}{\hat{L}_n^*})\cdot\hat{\varphi}_{\hat{\varepsilon}_{1,1}^n}(\frac{u\pi}{\hat{L}_n^*})\cdot\hat{\varphi}_{\hat{\delta}_{1,1}^n}(\frac{v\pi}{\hat{L}_n^*})-\hat{\varphi}_{(X^1,Y^1)_1^n}(\frac{u\pi}{\hat{L}_n^*},\frac{v\pi}{\hat{L}_n^*})|$$

$$\leq \frac{\log(n)}{\sqrt[4]{n}} \leq \frac{1}{2\log(n)}$$

mit Wahrscheinlichkeit 1 für n groß genug. Andererseits folgt

$$\begin{aligned}\frac{1}{\log(n)} &\leq \inf_{u,v \in [-\sqrt{\hat{K}_n}, \sqrt{\hat{K}_n}]} \left|\hat{\varphi}_{(X^{(1)},Y^{(1)})_1^n}(u\pi, v\pi)\right| \\ &= \inf_{|u| \leq 2\hat{K}_n, |v| \leq 2\hat{K}_n} |\hat{\varphi}_{(X^{(1)},Y^{(1)})_1^n}(\frac{u\pi}{\hat{L}_n^*}, \frac{v\pi}{\hat{L}_n^*})|\end{aligned}$$

aus $\hat{L}_n^\star = 2\sqrt{\hat{K}_n}$ und unserer Definition von $\hat{K}_n$ in (4.3). Damit folgt fast sicher

$$\inf_{u,v \in [-2\hat{K}_n, 2\hat{K}_n]} |\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(\frac{u\pi}{\hat{L}_n^*}, \frac{v\pi}{\hat{L}_n^*}) \cdot \hat{\varphi}_{\hat{\varepsilon}_{1,1}^n}(\frac{u\pi}{\hat{L}_n^*}) \cdot \hat{\varphi}_{\hat{\delta}_{1,1}^n}(\frac{v\pi}{\hat{L}_n^*})| \geq \frac{1}{2\log(n)}$$

für n groß genug. Es genügt also zu zeigen, dass

$$(\hat{K}_n)^3 \beta_n^2 \sup_{|u| \leq 2\hat{K}_n, |v| \leq 2\hat{K}_n} \left|\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(\frac{u\pi}{\hat{L}_n^*}, \frac{v\pi}{\hat{L}_n^*}) - \varphi_{(Z_1,Z_2)}(\frac{u\pi}{\hat{L}_n^*}, \frac{v\pi}{\hat{L}_n^*})\right| \to 0 \quad \text{f.s.}$$

unter der Annahme gilt, dass auf $[-2\hat{K}_n, 2\hat{K}_n]^2 := [-2\hat{K}_n, 2\hat{K}_n] \times [-2\hat{K}_n, 2\hat{K}_n]$

$$|\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(\frac{u\pi}{\hat{L}_n^*}, \frac{v\pi}{\hat{L}_n^*}) \cdot \hat{\varphi}_{\hat{\varepsilon}_{1,1}^n}(\frac{u\pi}{\hat{L}_n^*}) \cdot \hat{\varphi}_{\hat{\delta}_{1,1}^n}(\frac{v\pi}{\hat{L}_n^*})| \geq \frac{1}{2\log(n)}$$

gilt. Wir nutzen die Äquivalenz

$$u \in [-2\hat{K}_n, 2\hat{K}_n] \Leftrightarrow \frac{u\pi}{\hat{L}_n^*} \in \left[-\pi\sqrt{\hat{K}_n}, \pi\sqrt{\hat{K}_n}\right],$$

um die Notation der auftretenden Terme zu vereinfachen. Außerdem benötigen wir die Logarithmen der charakteristischen Funktionen aus dem Beweis von Lemma 4.1. Wir definieren $L_{\varphi_u}(v)$ wie in (5.1). Dann gilt $\exp(L_{\varphi_u}(v)) = \varphi_{(Z_1,Z_2)}(u,v)$ für $(u,v) \in [-2\hat{K}_n, 2\hat{K}_n]^2$. Analog zur Konstruktion in (5.1) erhalten wir auch einen Logarithmus $L_{\hat{\varphi}_{u,n}}(v)$ von $\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(u,v)$. Um die Lesbarkeit des nächsten Schrittes zu erleichtern, führen wir die Kurzschreibweisen $L_{\hat{\varphi}_{u,n}}(v) = b_n$ und $L_{\varphi_u}(v) = b$ ein. Damit folgt

$$\begin{aligned}&\left|\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(u,v) - \varphi_{(Z_1,Z_2)}(u,v)\right| = \left|\exp\big(L_{\varphi_u}(v)\big) - \exp\big(L_{\hat{\varphi}_{u,n}}(v)\big)\right| \\ &= \left|\exp(b_n)\right| \cdot \left|\exp(b - b_n) - 1\right| \\ &= \left|\exp(b_n)\right| \cdot \left|\exp(Re(b - b_n)) \cdot \exp(i \cdot Im(b - b_n)) - 1\right| \\ &= \left|\exp(b_n)\right| \cdot \Big|\big(\exp(Re(b - b_n)) - 1\big) \cdot \exp(i \cdot Im(b - b_n)) \\ &\quad + \exp(i \cdot Im(b - b_n)) - 1\Big|.\end{aligned}$$

Es gilt

$$|\exp(b_n)| = |\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(u,v)| \leq 1.$$

Als Nächstes verwenden wir $\exp(0) = 1$ und den Mittelwertsatz. Daraus folgt, dass ein $\xi \in \mathbb{R}$ mit $|\xi| \in [0, Re(b - b_n)]$ existiert, so dass die rechte Seite der Ungleichung durch

$$\left|\exp(\xi) \cdot |Re(b - b_n)|\right| \cdot 1 + \left|\cos\big(Im(b - b_n)\big) + i \sin\big(Im(b - b_n)\big) - 1\right|$$

$$\leq \exp(|Re(b-b_n)|)\cdot|b-b_n|+2|b-b_n|$$

beschränkt ist. Wobei wir hier die Monotonie der Exponentialfunktion und die Lipschitz-Stetigkeit von Sinus und Cosinus genutzt haben. Falls nun b_n fast sicher gegen b konvergiert, so folgt auch

$$\exp(|Re(b-b_n)|)\to 1 \quad \text{f.s.}$$

Zusammen mit der obigen Abschätzung für $\left|\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(u,v)-\varphi_{(Z_1,Z_2)}(u,v)\right|$, zeigt dies, dass es zum Beweis des Lemmas genügt

$$(\hat{K}_n)^3\beta_n^2 \sup_{|u|\leq\pi\sqrt{\hat{K}_n},|v|\leq\pi\sqrt{\hat{K}_n}} |L_{\hat{\varphi}_{u,n}}(v)-L_{\varphi_u}(v)|\to 0 \quad \text{f.s.} \tag{5.6}$$

zu zeigen.

Seien nun die Vektoren $(u,0,\dots,0,s,0,\dots,0)$ und $(t,0,\dots,0,0,0,\dots,0)$ mit den Kurzschreibweisen (u_0,s_0) und $(t_0,0)$ denotiert. Wir nutzen erneut die Beziehungen zwischen den charakteristischen Funktionen von (X,Y), Z_1,Z_2 und den jeweiligen Ableitungen aus dem Beweis von Lemma 4.1. Damit gilt

$$\begin{aligned}
|b-b_n| &= \Big|\int_0^u \frac{\frac{\partial}{\partial u}\varphi_{(Z_1,Z_2)}(t,0)}{\varphi_{(Z_1,Z_2)}(t,0)}\,dt+\int_0^v \frac{\frac{\partial}{\partial v}\varphi_{(Z_1,Z_2)}(u,s)}{\varphi_{(Z_1,Z_2)}(u,s)}\,ds \\
&\quad -\int_0^u \frac{\frac{\partial}{\partial u}\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(t,0)}{\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(t,0)}\,dt-\int_0^v \frac{\frac{\partial}{\partial v}\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(u,s)}{\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(u,s)}\,ds\Big| \\
&= \Big|\int_0^u \frac{1}{a_2}\cdot\frac{\frac{\partial}{\partial u_2}\varphi_{(X,Y)}(t_0,0)}{\varphi_{(X,Y)}(t_0,0)}dt+\int_0^v \frac{1}{b_2}\cdot\frac{\frac{\partial}{\partial v_2}\varphi_{(X,Y)}(u_0,s_0)}{\varphi_{(X,Y)}(u_0,s_0)}ds \\
&\quad -\int_0^u \frac{\frac{\partial}{\partial u}\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(t,0)}{\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(t,0)}\,dt-\int_0^v \frac{\frac{\partial}{\partial v}\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(u,s)}{\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(u,s)}\,ds\Big| \\
&\leq \Big|\int_0^u \frac{1}{a_2}\cdot\frac{\frac{\partial}{\partial u_2}\varphi_{(X,Y)}(t_0,0)}{\varphi_{(X,Y)}(t_0,0)}dt-\int_0^u \frac{1}{\hat{a}_2}\cdot\frac{\frac{\partial}{\partial u_2}\hat{\varphi}_{(X,Y)_1^n}(t_0,0)}{\hat{\varphi}_{(X,Y)_1^n}(t_0,0)}dt \\
&\quad +\int_0^u \frac{1}{\hat{a}_2}\cdot\frac{\frac{\partial}{\partial u_2}\hat{\varphi}_{(X,Y)_1^n}(t_0,0)}{\hat{\varphi}_{(X,Y)_1^n}(t_0,0)}dt-\int_0^u \frac{\frac{\partial}{\partial u}\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(t,0)}{\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(t,0)}\,dt\Big| \\
&\quad +\Big|\int_0^v \frac{1}{b_2}\cdot\frac{\frac{\partial}{\partial v_2}\varphi_{(X,Y)}(u_0,s_0)}{\varphi_{(X,Y)}(u_0,s_0)}ds-\int_0^v \frac{1}{\hat{b}_2}\cdot\frac{\frac{\partial}{\partial v_2}\hat{\varphi}_{(X,Y)_1^n}(u_0,s_0)}{\hat{\varphi}_{(X,Y)_1^n}(u_0,s_0)}ds \\
&\quad +\int_0^v \frac{1}{\hat{b}_2}\cdot\frac{\frac{\partial}{\partial v_2}\hat{\varphi}_{(X,Y)_1^n}(u_0,s_0)}{\hat{\varphi}_{(X,Y)_1^n}(u_0,s_0)}ds-\int_0^v \frac{\frac{\partial}{\partial v}\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(u,s)}{\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(u,s)}\,ds\Big| \\
&\leq \Big|\int_0^u \frac{1}{a_2}\cdot\frac{\frac{\partial}{\partial u_2}\varphi_{(X,Y)}(t_0,0)}{\varphi_{(X,Y)}(t_0,0)}dt-\int_0^u \frac{1}{\hat{a}_2}\cdot\frac{\frac{\partial}{\partial u_2}\hat{\varphi}_{(X,Y)_1^n}(t_0,0)}{\hat{\varphi}_{(X,Y)_1^n}(t_0,0)}dt\Big| \\
&\quad +\Big|\int_0^u \frac{1}{\hat{a}_2}\cdot\frac{\frac{\partial}{\partial u_2}\hat{\varphi}_{(X,Y)_1^n}(t_0,0)}{\hat{\varphi}_{(X,Y)_1^n}(t_0,0)}dt-\int_0^u \frac{\frac{\partial}{\partial u_2}\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(t,0)}{\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(t,0)}dt\Big| \\
&\quad +\Big|\int_0^v \frac{1}{b_2}\cdot\frac{\frac{\partial}{\partial v_2}\varphi_{(X,Y)}(u_0,s_0)}{\varphi_{(X,Y)}(u_0,s_0)}ds-\int_0^v \frac{1}{\hat{b}_2}\cdot\frac{\frac{\partial}{\partial v_2}\hat{\varphi}_{(X,Y)_1^n}(u_0,s_0)}{\hat{\varphi}_{(X,Y)_1^n}(u_0,s_0)}ds\Big| \\
&\quad +\Big|\int_0^v \frac{1}{\hat{b}_2}\cdot\frac{\frac{\partial}{\partial v_2}\hat{\varphi}_{(X,Y)_1^n}(u_0,s_0)}{\hat{\varphi}_{(X,Y)_1^n}(u_0,s_0)}ds-\int_0^v \frac{\frac{\partial}{\partial v_2}\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(u,s)}{\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(u,s)}ds\Big|
\end{aligned}$$

$$= S_{1,n} + S_{2,n} + S_{3,n} + S_{4,n}.$$

Bevor wir zeigen können, dass (5.6) gilt, benötigen wir einige Abschätzungen für $S_{1,n}$, $S_{2,n}$, $S_{3,n}$ und $S_{4,n}$. Wir beginnen mit der Abschätzung von $S_{1,n}$. Es gilt

$$\begin{aligned}
&\sup_{|u|,|v|\leq \pi\sqrt{\hat{K}_n}} S_{1,n} \\
&\leq \pi\sqrt{\hat{K}_n} \sup_{|u|,|v|\leq \pi\sqrt{\hat{K}_n}} \left| \frac{1}{a_2} \cdot \frac{\frac{\partial}{\partial u_2}\varphi_{(X,Y)}(u_0,v_0)}{\varphi_{(X,Y)}(u_0,v_0)} - \frac{1}{\hat{a}_2} \cdot \frac{\frac{\partial}{\partial u_2}\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0)}{\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0)} \right|.
\end{aligned}$$

Sind $a_n \in \mathbb{R}$, $b_n \in \mathbb{R}\setminus\{0\}$ Folgen, die jeweils gegen $a \in \mathbb{R}$ und $b \in \mathbb{R}\setminus\{0\}$ konvergieren, so gilt die elementare Ungleichung

$$\begin{aligned}
\left|\frac{a}{b} - \frac{a_n}{b_n}\right| &= \left|\frac{ab_n - a_n b}{bb_n}\right| \\
&= \left|\frac{ab_n - ab + ab - a_n b}{bb_n}\right| \leq \left|\frac{a(b-b_n)}{bb_n}\right| + \left|\frac{b(a-a_n)}{bb_n}\right|.
\end{aligned}$$

Mit $a = \frac{\partial}{\partial u_2}\varphi_{(X,Y)}(u_0,v_0)$, $b = a_2\varphi_{(X,Y)}(u_0,v_0)$, $a_n = \frac{\partial}{\partial u_2}\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0)$ und $b_n = \hat{a}_2\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0)$ folgt damit

$$\begin{aligned}
&\left| \frac{\frac{\partial}{\partial u_2}\varphi_{(X,Y)}(u_0,v_0)}{a_2\varphi_{(X,Y)}(u_0,v_0)} - \frac{\frac{\partial}{\partial u_2}\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0)}{\hat{a}_2\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0)} \right| \\
&\leq \left| \frac{\left(\frac{\partial}{\partial u_2}\varphi_{(X,Y)}(u_0,v_0) - \frac{\partial}{\partial u_2}\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0)\right) \cdot a_2\varphi_{(X,Y)}(u_0,v_0)}{a_2\varphi_{(X,Y)}(u_0,v_0) \cdot \hat{a}_2\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0)} \right| \\
&\quad + \left| \frac{\frac{\partial}{\partial u_2}\varphi_{(X,Y)}(u_0,v_0) \cdot \left(a_2\varphi_{(X,Y)}(u_0,v_0) - \hat{a}_2\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0)\right)}{a_2\varphi_{(X,Y)}(u_0,v_0) \cdot \hat{a}_2\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0)} \right| \\
&= S_{1,n}^1 + S_{1,n}^2.
\end{aligned}$$

Nach Definition von $\hat{K}_n$ gilt

$$\begin{aligned}
&\hat{K}_n := \\
&\min\left\{ \lfloor n^{\frac{1}{11}} \rfloor, \max\{K \in \mathbb{N} : \inf_{u,v\in[-\sqrt{K},\sqrt{K}]} |\hat{\varphi}_{(X^{(1)},Y^{(1)})_1^n}(u\pi, v\pi)| \geq \frac{1}{\log(n)}\} \right\}.
\end{aligned}$$

Daraus folgt, dass $|\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0)| \geq \frac{1}{\log(n)}$ für $u,v \in [-\pi\sqrt{\hat{K}_n}, \pi\sqrt{\hat{K}_n}]$ gilt. Aus Lemma 5.1 folgt, dass

$$\sup_{u,v\in[-\pi\sqrt{\hat{K}_n},\pi\sqrt{\hat{K}_n}]} \left|\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0) - \varphi_{(X,Y)}(u_0,v_0)\right| \leq \frac{\log(n)}{\sqrt{n}}$$

mit Wahrscheinlichkeit 1 für n groß genug gilt. Für n groß genug gilt aber ebenfalls $\frac{\log(n)}{\sqrt{n}} < \frac{1}{2\log(n)}$. Zusammen mit

$$\inf_{u,v\in[-\pi\sqrt{\hat{K}_n},\pi\sqrt{\hat{K}_n}]} |\varphi_{(X,Y)}(u_0,v_0)| \geq \inf_{u,v\in[-\pi\sqrt{\hat{K}_n},\pi\sqrt{\hat{K}_n}]} |\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0)|$$

$$- \sup_{u,v \in [-\pi\sqrt{\hat{K}_n}, \pi\sqrt{\hat{K}_n}]} \left| \hat{\varphi}_{(X,Y)_1^n}(u_0,v_0) - \varphi_{(X,Y)}(u_0,v_0) \right|$$

folgt, dass mit Wahrscheinlichkeit 1 für n groß genug auch

$$\inf_{u,v \in [-\pi\sqrt{\hat{K}_n}, \pi\sqrt{\hat{K}_n}]} |\varphi_{(X,Y)}(u_0,v_0)| \geq \frac{1}{2\log(n)}$$

gilt. Außerdem konvergiert $\hat{a}_2$ nach Lemma 5.4 fast sicher gegen $a_2 \neq 0$. Es exisitert also eine Konstante $C > 0$, welche nur von a abhängt, so dass mit Wahrscheinlichkeit 1 für n groß genug $\hat{a}_2 \geq C$ gilt. Damit sind die Nenner von $S_{1,n}^1$ und $S_{1,n}^2$ mit Wahrscheinlichkeit 1 für n groß genung nach untern durch $C\frac{1}{2\log(n)^2}$ beschränkt. Da der Betrag der charakteristischen Funktion durch 1 beschränkt ist, folgt

$$\begin{aligned}
&\frac{\sqrt{n}}{\log(n)^4} \sup_{|u|,|v| \leq \pi\sqrt{\hat{K}_n}} S_{1,n}^1 \\
&= \frac{\sqrt{n}}{\log(n)^4} \sup_{|u|,|v| \leq \pi\sqrt{\hat{K}_n}} \Bigg| \frac{\left(\frac{\partial}{\partial u_2}\varphi_{(X,Y)}(u_0,v_0)\right)}{a_2\varphi_{(X,Y)}(u_0,v_0) \cdot \hat{a}_2\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0)} \\
&\quad - \frac{\frac{\partial}{\partial u_2}\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0)\big) \cdot a_2\varphi_{(X,Y)}(u_0,v_0)}{a_2\varphi_{(X,Y)}(u_0,v_0) \cdot \hat{a}_2\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0)} \Bigg| \\
&\leq \frac{2|a_2|}{C} \frac{\sqrt{n}}{\log(n)^2} \left| \frac{\partial}{\partial u_2}\varphi_{(X,Y)}(u_0,v_0) - \frac{\partial}{\partial u_2}\hat{\varphi}_{(X,Y)_1^n}(u_0,v_0) \right|.
\end{aligned}$$

Zusammen mit Lemma 5.2 impliziert dies, dass mit Wahrscheinlichkeit 1 eine zufällige Konstante C' existiert, so dass

$$\frac{\sqrt{n}}{\log(n)^4} \sup_{|u|,|v| \leq \pi\sqrt{\hat{K}_n}} S_{1,n}^1 \leq C'$$

für alle $n \in \mathbb{N}$ gilt. Um eine analoge Abschätzung für $S_{1,n}^2$ zu erhalten, stellen wir zunächst fest, dass $\frac{\partial}{\partial u_2}\varphi_{(X_{M_n},Y_{M_n})}(u_0,v_0)$ durch eine Konstante beschränkt ist, da alle Momente von $X^{(2)}$ endlich sind. Also können wir erneut Lemma 5.1 und Lemma 5.4 anwenden. Damit erhalten wir, dass

$$\frac{\sqrt{n}}{\log(n)^4} \sup_{|u|,|v| \leq \pi\sqrt{\hat{K}_n}} S_{1,n}^2,$$

und somit auch,

$$\frac{\sqrt{n}}{\log(n)^4} \sup_{|u|,|v| \leq \pi\sqrt{\hat{K}_n}} S_{1,n}$$

mit Wahrscheinlichkeit 1 durch eine jeweils zufällig Konstante beschränkt sind. Nun werden wir analoge Abschätzungen für die Terme $S_{2,n}$ bis $S_{4,n}$ herleiten. Alle Argumente, die wir zur Abschätzung von $S_{1,n}$ verwendet haben, lassen sich auf $S_{3,n}$ übertragen und liefern die selben oberen und unteren Schranken. Um Schranken für $S_{2,n}$ und $S_{4,n}$ herzuleiten modifizieren wir die obigen Argumente. Wie wir zu Beginn des Beweises festgestellt haben, können wir o.B.d.A. davon ausgehen, dass $|\hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}(u,v) \cdot \hat{\varphi}_{\hat{\varepsilon}_{1,1}^n}(u) \cdot \hat{\varphi}_{\hat{\delta}_{1,1}^n}(v)|$ größer ist als $\frac{1}{2\log(n)}$. Außerdem ist

$$|\hat{\varphi}_{(X,Y)_1^n}(t_0,0)| = \left|\frac{1}{n}\sum_{l=1}^{n} \exp\left(i \cdot (t \cdot X^{(1)} + 0 \cdot Y^{(1)}\right)\right|$$

offensichtlich durch 1 beschränkt und

$$\begin{aligned}
|\frac{\partial}{\partial u_2}\hat{\varphi}_{(X,Y)_1^n}(t_0,0)| &= |\frac{1}{n}\sum_{j=1}^{n} iX_j^{(2)} e^{i(tX_j^{(1)}+0Y_j^{(1)})}| \\
&\leq \frac{1}{n}\sum_{j=1}^{n}|X_j^{(2)}| \to \mathbf{E}\{|X^{(2)}|\} \quad \text{f.s.}
\end{aligned}$$

impliziert, dass $\frac{\partial}{\partial u_2}\hat{\varphi}_{(X,Y)_1^n}(t_0,0)$ mit Wahrscheinlichkeit 1 beschränkt ist. Also können wir S_2 und S_4 genau wie S_1 zerlegen, und die gewünschte Abschätzung folgt mit Lemma 5.5 anstelle von Lemma 5.1 und Lemma 5.2.

Nun benutzen wir die Abschätzungen für $|b - b_n|$ um zu zeigen, dass (5.6) gilt. Nach Definition gilt $\beta_n = n^{\frac{1}{11}}$ und $\hat{K}_n \leq n^{\frac{1}{11}}$. Mit Wahrscheinlichkeit 1 gibt es ein (zufälliges) $C > 0$, so dass

$$\begin{aligned}
&(\hat{K}_n)^3\beta_n^2 \sup_{|u|\leq\pi\sqrt{\hat{K}_n},|v|\leq\pi\sqrt{\hat{K}_n}} |L_{\hat{\varphi}_{u,n}}(v) - L_{\varphi_u}(v)| \\
&\leq n^{\frac{5}{11}} \sup_{|u|\leq\pi\sqrt{\hat{K}_n},|v|\leq\pi\sqrt{\hat{K}_n}} (S_{1,n} + S_{2,n} + S_{3,n} + S_{4,n}) \\
&\leq C \cdot n^{\frac{5}{11}} \cdot \frac{\log(n)^4}{\sqrt{n}}
\end{aligned}$$

gilt. Da die rechte Seite der Ungleichung gegen 0 konvergiert, ist der Beweis beendet. □

5.2 Beweis der Konvergenz des Regressionsschätzers

Für die folgenden Beweise benötigen wir einen auf die dem Regressionsschätzer zugrundliegende Menge von Funktionen $\mathcal{F}_n$ abgestimmte Menge von trigonometrischen Polynomen in zwei Variablen.

Definition 5.7. Sei $\mathcal{G}_n = \mathcal{G}_n(\hat{K}_n, \beta_n)$ die Mengen von Funktionen g mit

$$\begin{aligned}
g(z_1,z_2) \quad = \quad & \sum_{j,k=0}^{2\cdot\hat{K}_n} a_{j,k} \cdot \cos\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_1\right) \cdot \cos\left(j \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_2\right) \\
&+ \sum_{j,k=0}^{2\cdot\hat{K}_n} b_{j,k} \cdot \cos\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_1\right) \cdot \sin\left(j \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_2\right) \\
&+ \sum_{j,k=0}^{2\cdot\hat{K}_n} c_{j,k} \cdot \sin\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_1\right) \cdot \cos\left(j \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_2\right) \\
&+ \sum_{j,k=0}^{2\cdot\hat{K}_n} d_{j,k} \cdot \sin\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_1\right) \cdot \sin\left(j \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_2\right),
\end{aligned}$$

wobei $a_{j,k}, b_{j,k}, c_{j,k}, d_{j,k} \in [-10 \cdot \hat{K}_n \cdot \beta_n^2, 10 \cdot \hat{K}_n \cdot \beta_n^2]$ gilt.

Nun beweisen wir eine multivariate Variante von Lemma 3.2 für die trigonometrischen Polynome aus G_n.

Lemma 5.8. *Es sei $(\hat{K}_n)_n$ eine Folge von reellwertigen Zufallsvariablen und es gelte $\hat{K}_n \to \infty$ fast sicher. Sei $(\beta_n)_n$ ebenfalls eine Folge von reellwertigen Zufallsvariablen. Sei $\mathcal{G}_n$ in abhängigkeit von $\hat{K}_n$ und β_n definiert wie in Definition 5.7. Gilt dann*

$$\hat{K}_n^3 \cdot \beta_n^2 \cdot \sup_{k,j\in\mathbb{Z},|k|\leq 2\cdot\hat{K}_n,|j|\leq 2\cdot\hat{K}_n} \left| \hat{\varphi}_{(\hat{z}_1,\hat{z}_2)_1^n}\left(j \cdot \frac{\pi}{\hat{L}_n^*}, k \cdot \frac{\pi}{\hat{L}_n^*}\right) - \varphi_{(Z_1,Z_2)}\left(j \cdot \frac{\pi}{\hat{L}_n^*}, k \cdot \frac{\pi}{\hat{L}_n^*}\right)\right| \to 0$$

mit Wahrscheinlichkeit 1, so folgt

$$\sup_{g\in\mathcal{G}_n} \left| \frac{1}{n}\sum_{l=1}^n g(\hat{z}_{1,l}, \hat{z}_{2,l}) - \int g(Z_1, Z_2)\mathbf{P}_{(Z_1,Z_2)}(d(z_1,z_2))\right| \to 0 \quad \textit{f.s.}$$

Beweis. Sei

$$A_{n,1} := \sup_{k,j\in\mathbb{Z},|k|\leq 2\cdot\hat{K}_n,|j|\leq 2\cdot\hat{K}_n} \left| \frac{1}{n}\sum_{l=1}^n \cos\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot \hat{z}_{1,l}\right) \cdot \cos\left(j \cdot \frac{\pi}{\hat{L}_n^*} \cdot \hat{z}_{2,l}\right) - \int \cos\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_1\right) \cdot \cos\left(j \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_2\right) \mathbf{P}_{(Z_1,Z_2)}(d(z_1,z_2))\right|,$$

$$A_{n,2} := \sup_{k,j\in\mathbb{Z},|k|\leq 2\cdot\hat{K}_n,|j|\leq 2\cdot\hat{K}_n} \left| \frac{1}{n}\sum_{l=1}^n \cos\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot \hat{z}_{1,l}\right) \cdot \sin\left(j \cdot \frac{\pi}{\hat{L}_n^*} \cdot \hat{z}_{2,l}\right) - \int \cos\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_1\right) \cdot \sin\left(j \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_2\right) \mathbf{P}_{(Z_1,Z_2)}(d(z_1,z_2))\right|,$$

$$A_{n,3} := \sup_{k,j\in\mathbb{Z},|k|\leq 2\cdot\hat{K}_n,|j|\leq 2\cdot\hat{K}_n} \left| \frac{1}{n}\sum_{l=1}^n \sin\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot \hat{z}_{1,l}\right) \cdot \cos\left(j \cdot \frac{\pi}{\hat{L}_n^*} \cdot \hat{z}_{2,l}\right) - \int \sin\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_1\right) \cdot \cos\left(j \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_2\right) \mathbf{P}_{(Z_1,Z_2)}(d(z_1,z_2))\right|$$

und

$$A_{n,4} := \sup_{k,j\in\mathbb{Z},|k|\leq 2\cdot\hat{K}_n,|j|\leq 2\cdot\hat{K}_n} \left| \frac{1}{n}\sum_{l=1}^n \sin\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot \hat{z}_{1,l}\right) \cdot \sin\left(j \cdot \frac{\pi}{\hat{L}_n^*} \cdot \hat{z}_{2,l}\right) - \int \sin\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_1\right) \cdot \sin\left(j \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_2\right) \mathbf{P}_{(Z_1,Z_2)}(d(z_1,z_2))\right|.$$

Nach Voraussetzung konvergiert das Produkt von $\hat{K}_n^3\beta_n^2$ und dem Betrag von

$$\begin{aligned}\Theta_{j,k} :=& \frac{1}{n}\sum_{l=1}^{n}\exp\big(i\big(\frac{\pi\cdot j}{\hat{L}_n^*}\hat{z}_{1,l}+\frac{\pi\cdot k}{\hat{L}_n^*}\hat{z}_{2,l}\big)\big)\\ &-\mathbf{E}\left\{\exp\big(i\cdot\big(\frac{\pi\cdot j}{\hat{L}_n^*}\cdot Z_1+\frac{\pi\cdot k}{\hat{L}_n^*}\cdot Z_2\big)\big)\right\}\end{aligned}$$

gleichmäßig für jedes $(j,k)\in\{-2\cdot\hat{K}_n,-2\cdot\hat{K}_n+1,\dots,2\cdot\hat{K}_n\}^2$ fast sicher gegen 0. Also konvergiert auch $\hat{K}_n^3\beta_n^2\cdot|(\Theta_{j,k}+\Theta_{j,(-k)})|$ gleichmäßig fast sicher gegen 0. Für $x,y\in\mathbb{R}$ gilt

$$Re\exp(ix)=\cos(x)\quad\text{und}\quad\cos(x+y)+\cos(x-y)=2\cos(x)\cos(y).$$

Damit erhalten wir

$$\begin{aligned}&Re\left(\exp\big(i\big(\frac{\pi\cdot j}{\hat{L}_n^*}\hat{z}_{1,l}+\frac{\pi\cdot k}{\hat{L}_n^*}\hat{z}_{2,l}\big)\big)\right)+Re\left(\exp\big(i\big(\frac{\pi\cdot j}{\hat{L}_n^*}\hat{z}_{1,l}+\frac{\pi\cdot(-k)}{\hat{L}_n^*}\hat{z}_{2,l}\big)\big)\right)\\ &=\cos\big(\frac{\pi\cdot j}{\hat{L}_n^*}\hat{z}_{1,l}+\frac{\pi\cdot k}{\hat{L}_n^*}\hat{z}_{2,l}\big)+\cos\big(\frac{\pi\cdot j}{\hat{L}_n^*}\hat{z}_{1,l}+\frac{\pi\cdot(-k)}{\hat{L}_n^*}\hat{z}_{2,l}\big)\\ &=2\cos\big(\frac{\pi\cdot j}{\hat{L}_n^*}\hat{z}_{1,l}\big)\cdot\cos\big(\frac{\pi\cdot k}{\hat{L}_n^*}\hat{z}_{2,l}\big).\end{aligned}$$

Zusammen mit

$$\begin{aligned}Re\left(\Theta_{j,k}+\Theta_{j,(-k)}\right) =& \frac{1}{n}\sum_{l=1}^{n}\left(Re\left(\exp\big(i\big(\frac{\pi\cdot j}{\hat{L}_n^*}\hat{z}_{1,l}+\frac{\pi\cdot k}{\hat{L}_n^*}\hat{z}_{2,l}\big)\big)\right)\right.\\ &\left.+Re\left(\exp\big(i\big(\frac{\pi\cdot j}{\hat{L}_n^*}\hat{z}_{1,l}+\frac{\pi\cdot(-k)}{\hat{L}_n^*}\hat{z}_{2,l}\big)\big)\right)\right)\\ &-\mathbf{E}\left\{Re\left(\exp\big(i\big(\frac{\pi\cdot j}{\hat{L}_n^*}\cdot Z_1+\frac{\pi\cdot(-k)}{\hat{L}_n^*}\cdot Z_2\big)\big)\right)\right.\\ &\left.+Re\left(\exp\big(i\big(\frac{\pi\cdot j}{\hat{L}_n^*}\cdot Z_1+\frac{\pi\cdot k}{\hat{L}_n^*}\cdot Z_2\big)\big)\right)\right\},\end{aligned}$$

liefert dies

$$\hat{K}_n^3\beta_n^2\sup_{k,j\in\mathbb{Z},|k|\le 2\cdot\hat{K}_n,|j|\le 2\cdot\hat{K}_n}|Re(\Theta_{j,k}+\Theta_{j,(-k)})|=2\cdot\hat{K}_n^3\beta_n^2A_{n,1}.$$

Da

$$\begin{aligned}&\hat{K}_n^3\beta_n^2\sup_{k,j\in\mathbb{Z},|k|\le 2\cdot\hat{K}_n,|j|\le 2\cdot\hat{K}_n}|Re(\Theta_{j,k}+\Theta_{j,(-k)})|\\ \le\ &\hat{K}_n^3\beta_n^2\sup_{k,j\in\mathbb{Z},|k|\le 2\cdot\hat{K}_n,|j|\le 2\cdot\hat{K}_n}|(\Theta_{j,k}+\Theta_{j,(-k)})|\to 0\qquad\text{f.s.}\end{aligned}$$

gilt, folgt, dass auch $\hat{K}_n^3\beta_n^2A_{n,1}\to 0$ fast sicher gilt. Mit den Identitäten

$$Im(\exp(ix))=\sin(x),$$

$$\sin(x-y)+\sin(x+y)=2\sin(x)\cos(y)$$

und

$$\cos(x-y)-\cos(x+y)=2\sin(x)\sin(y)$$

erhalten wir durch analoge Argumente, das auch

$$\hat{K}_n^3\beta_n^2 A_{n,j}\to 0 \quad \text{f.s.}$$

für $j\in\{2,3,4\}$ gilt. Insgesamt folgt

$$\begin{aligned}
&\sup_{g\in\mathcal{G}_n}\left|\frac{1}{n}\sum_{l=1}^{n} g(\hat{z}_{1,l},\hat{z}_{2,l})-\int g(z_1,z_2)\mathbf{P}_{(Z_1,Z_2)}(d(z_1,z_2))\right|\\
&=\sup_{\substack{a_{j,k},\,b_{j,k},\,c_{j,k},\,d_{j,k}\\ \in[-10\cdot\hat{K}_n\beta_n^2,10\cdot\hat{K}_n\beta_n^2]}}\Bigg|\frac{1}{n}\sum_{l=1}^{n}\sum_{j,k=0}^{2\cdot\hat{K}_n}\Big[a_{j,k}\cos\big(k\frac{\pi}{\hat{L}_n^*}\hat{z}_{1,l}\big)\cos\big(j\frac{\pi}{\hat{L}_n^*}\hat{z}_{2,l}\big)\\
&\quad+b_{j,k}\cos\big(k\frac{\pi}{\hat{L}_n^*}\hat{z}_{1,l}\big)\sin\big(j\frac{\pi}{\hat{L}_n^*}\hat{z}_{2,l}\big)+c_{j,k}\sin\big(k\frac{\pi}{\hat{L}_n^*}\hat{z}_{1,l}\big)\cos\big(j\frac{\pi}{\hat{L}_n^*}\hat{z}_{2,l}\big)\\
&\quad+d_{j,k}\sin\big(k\frac{\pi}{\hat{L}_n^*}\hat{z}_{1,l}\big)\sin\big(j\frac{\pi}{\hat{L}_n^*}\hat{z}_{2,l}\big)\Big]-\int\sum_{j,k=0}^{2\hat{K}_n}\Big[a_{j,k}\cos\big(k\frac{\pi}{\hat{L}_n^*}z_1\big)\cos\big(j\frac{\pi}{\hat{L}_n^*}z_2\big)\\
&\quad+b_{j,k}\cos\big(k\frac{\pi}{\hat{L}_n^*}z_1\big)\sin\big(j\frac{\pi}{\hat{L}_n^*}z_2\big)+c_{j,k}\sin\big(k\frac{\pi}{\hat{L}_n^*}z_1\big)\cos\big(j\frac{\pi}{\hat{L}_n^*}z_2\big)\\
&\quad+d_{j,k}\sin\big(k\frac{\pi}{\hat{L}_n^*}z_1\big)\sin\big(j\frac{\pi}{\hat{L}_n^*}z_2\big)\Big]d\mathbf{P}_{(Z_1,Z_2)}(d(z_1,z_2))\Bigg|\\
&\le 10\cdot\hat{K}_n\cdot\beta_n^2\sum_{j,k=0}^{2\cdot\hat{K}_n}\Big|\frac{1}{n}\sum_{l=1}^{n}\cos\big(k\frac{\pi}{\hat{L}_n^*}\hat{z}_{1,l}\big)\cos\big(j\frac{\pi}{\hat{L}_n^*}\hat{z}_{2,l}\big)\\
&\qquad-\int\cos\big(k\frac{\pi}{\hat{L}_n^*}z_1\big)\cos\big(j\frac{\pi}{\hat{L}_n^*}z_2\big)\,d\mathbf{P}_{(Z_1,Z_2)}(d(z_1,z_2))\Big|\\
&\quad+\Big|\frac{1}{n}\sum_{l=1}^{n}\cos\big(k\frac{\pi}{\hat{L}_n^*}\hat{z}_{1,l}\big)\sin\big(j\frac{\pi}{\hat{L}_n^*}\hat{z}_{2,l}\big)\\
&\qquad-\int\cos\big(k\frac{\pi}{\hat{L}_n^*}z_1\big)\sin\big(j\frac{\pi}{\hat{L}_n^*}z_2\big)\,d\mathbf{P}_{(Z_1,Z_2)}(d(z_1,z_2))\Big|\\
&\quad+\Big|\frac{1}{n}\sum_{l=1}^{n}\sin\big(k\frac{\pi}{\hat{L}_n^*}\hat{z}_{1,l}\big)\cos\big(j\frac{\pi}{\hat{L}_n^*}\hat{z}_{2,l}\big)\\
&\qquad-\int\sin\big(k\frac{\pi}{\hat{L}_n^*}z_1\big)\cos\big(j\frac{\pi}{\hat{L}_n^*}z_2\big)\,d\mathbf{P}_{(Z_1,Z_2)}(d(z_1,z_2))\Big|\\
&\quad+\Big|\frac{1}{n}\sum_{l=1}^{n}\sin\big(k\frac{\pi}{\hat{L}_n^*}\hat{z}_{1,l}\big)\sin\big(j\frac{\pi}{\hat{L}_n^*}\hat{z}_{2,l}\big)\\
&\qquad-\int\sin\big(k\frac{\pi}{\hat{L}_n^*}z_1\big)\sin\big(j\frac{\pi}{\hat{L}_n^*}z_2\big)\,d\mathbf{P}_{(Z_1,Z_2)}(d(z_1,z_2))\Big|\\
&\le 10\cdot\hat{K}_n\beta_n^2\cdot(2\cdot\hat{K}_n+1)^2\cdot\sum_{r=1}^{4}A_{n,r}\le 90\cdot\hat{K}_n^3\beta_n^2\cdot\sum_{r=1}^{4}A_{n,r}
\end{aligned}$$

was den Beweis beendet. $\square$

Lemma 5.9. *Seien $Z_1, Z_2, X, Y, \varepsilon, \delta$ Zufallsvariablen wie in Definition 4.2. Sei $\mathcal{F}_n$ die Menge von Funktionen gegeben durch*

$$f(z_1) = \sum_{k=0}^{\hat{K}_n} a_k \cdot \cos\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_1\right) + \sum_{k=0}^{\hat{K}_n} b_k \cdot \sin\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_1\right)$$

wobei $a_k, b_k \in [-\beta_n, \beta_n]$ gilt. Wir wählen $\hat{K}_n$ abhängig von den Daten $X_1^{(1)}, \ldots, X_n^{(1)}$, $Y_1^{(1)}, \ldots, Y_n^{(1)}$ wie in (4.3) angegeben. Sei T_n wie in Gleichung (4.4) und seien die geschätzten Daten $(\hat{z}_1, \hat{z}_2)_1^n$ definiert als minimum von T_n under der Nebenbedingung $\frac{1}{n} \cdot \sum_{j=1}^n |\hat{z}_{2,j}|^6 \leq \sqrt{\hat{K}_n}$. Sei m_n gegeben durch

$$m_n(\cdot) = \arg\min_{f \in \mathcal{F}_n} \frac{1}{n} \sum_{i=1}^n |\hat{z}_{i,2} - f(\hat{z}_{i,1})|^2.$$

Gilt nun $\hat{L}_n^* = 2 \cdot \hat{L}_n = 2 \cdot \sqrt{\hat{K}_n}$ sowie $\beta_n = n^{\frac{1}{11}}$ und sind die 6-ten Momente $\mathbf{E}Z_1^6$ und $\mathbf{E}Z_2^6$ endlich und gilt

$$\begin{aligned} \hat{K}_n^3 \cdot \beta_n^2 \cdot \sup_{k,j \in \mathbb{Z}, |k| \leq 2\hat{K}_n, |j| \leq 2\hat{K}_n} \Bigg| &\hat{\varphi}_{(\hat{z}_1, \hat{z}_2)_1^n}\left(j \cdot \frac{\pi}{\hat{L}_n^*}, k \cdot \frac{\pi}{\hat{L}_n^*}\right) \\ &- \varphi_{(Z_1, Z_2)}\left(j \cdot \frac{\pi}{\hat{L}_n^*}, k \cdot \frac{\pi}{\hat{L}_n^*}\right)\Bigg| \to 0 \end{aligned}$$

fast sicher, so folgt

$$\int |m_n(z_1) - m(z_1)|^2 \mathbf{P}_{Z_1}(dz_1) \to 0$$

fast sicher.

Beweis. Zunächst zerlegen wir den L_2-Fehler des Kleinste-Quadrate-Schätzers. Wir verwenden eine Variation der klassischen Dekomposition(siehe [17, Lemma 10.1]). Für $\delta > 0$ gilt

$$\begin{aligned} &\int |m_n(z_1) - m(z_1)|^2 \mathbf{P}_{Z_1}(dz_1) \\ &= \mathbf{E}\Big\{|m_n(Z_1) - Z_2|^2 \Big| \mathcal{D}_n\Big\} - \mathbf{E}\Big\{|m(Z_1) - Z_2|^2\Big\} \\ &= \mathbf{E}\Big\{|m_n(Z_1) - Z_2|^2 \Big| \mathcal{D}_n\Big\} - (1+\delta)^8 \inf_{f \in \mathcal{F}_n} \mathbf{E}\{|f(Z_1) - Z_2|^2\} \\ &\quad + (1+\delta)^8 \left[\inf_{f \in \mathcal{F}_n} \mathbf{E}\{|f(Z_1) - Z_2|^2\} - \mathbf{E}\Big\{|m(Z_1) - Z_2|^2\Big\}\right] \\ &\quad + (1+\delta)^8 \mathbf{E}\Big\{|m(Z_1) - Z_2|^2\Big\} - \mathbf{E}\Big\{|m(Z_1) - Z_2|^2\Big\} \\ &= T_{1,n} + T_{2,n} + T_{3,n}. \end{aligned}$$

Im folgenden werden wir Schrittweise obere Schranken für $T_{1,n}, T_{2,n}$ und $T_{3,n}$ herleiten und mit Hilfe dieser die Konvergenz von $\int |m_n(z_1) - m(z_1)|^2 \mathbf{P}_{Z_1}(dz_1)$ beweisen. Der erste Schritt ist der einfachste. Sei $\varepsilon > 0$ beliebig. Da $\mathbf{E}\{|m(Z_1) - Z_2|^2\}$ nicht von n abhängt, können wir δ so wählen, dass

$$T_{3,n} = ((1+\delta)^8 - 1)\mathbf{E}\{|m(Z_1) - Z_2|^2\} \leq \varepsilon$$

gitl. Im zweiten Schritt nehmen wir uns $T_{2,n}$ an. Wir verenden die Standard-Zerlegung des L_2-Risikos die wir bereits in der Einleitung als Gleichung (1.3) vorgestellt haben. Diese liefert

$$T_{2,n} = (1+\delta)^8 \inf_{f \in \mathcal{F}_n} \int |f(z_1) - m(z_1)|^2 \mathbf{P}_{Z_1}(dz_1).$$

Da die Regressionsfunktion m in $L_2(\mathbf{P}_{Z_1})$ beliebig genau durch glatte Funktionen mit kompaktem Träger approximiert werden kann (siehe beispielsweise [17, Theorem A.1]), genügt es zu zeigen, dass solche Funktionen in $L_2(\mathbf{P}_{Z_1})$ beliebig genau durch Funktionen aus $\mathcal{F}_n$ approximiert werden können. Sei m^* eine glatte Funktion mit kompaktem Träger und B ein Intervall, welches den Träger enthält. Unter den Voraussetzungen des Lemmas sind auch die Voraussetzungen von Lemma 5.3 erfüllt und es gilt gilt $\hat{K}_n \to \infty$ fast sicher. Aus unserer Wahl der $\hat{L}_n^*$ folgt damit $\hat{L}_n^* \to \infty$ fast sicher. Also ist der Support von m^* für große n fast sicher in

$$\hat{I}_n := [-\hat{L}_n^*, \hat{L}_n^*]$$

enthalten. Für solche n setzen wir nun m^* Lipschitz-stetig und $2\hat{L}_n^*$-periodisch von $[-\hat{L}_n^*, \hat{L}_n^*]$ nach $\mathbb{R}$ fort. Für x in $\mathbb{R}$ exisitert ein eindeutiges $x_0 \in [-\hat{L}_n^*, \hat{L}_n^*]$, so dass $x - x_0 = a \cdot 2\hat{L}_n^*$ für ein $a \in \mathbb{Z}$ gilt. Diese, von n abhängige Fortsetzung von m^* ist dann für alle $x \in \mathbb{R}$ durch

$$m_n^*(x) := m^*(x_0)$$

gegeben. Offensichtlich ist m_n^* durch $\|m^*\|_\infty$ betragsmäßig beschränkt. Nach Definition sind die Koeffizienten der Fourier-Reihe von m_n^* durch $\|m_n^*\|_\infty$ und daher auch durch $\|m^*\|_\infty$ beschränkt. Ist nun $S_{\hat{K}_n, m_n^*}$ die $\hat{K}_n$-te Fourier-Summe von m_n^*, so ist diese wegen $\beta_n = n^{\frac{1}{11}} \to \infty$ für genügenden große $n \in \mathbb{N}$ in $\mathcal{F}_n$ enthalten. Für solche n gilt

$$\begin{aligned}
&(1+\delta)^8 \inf_{f \in \mathcal{F}_n} \int |f(z_1) - m^*(z_1)|^2 \mathbf{P}_{Z_1}(dz_1) \\
&\leq (1+\delta)^8 \int |S_{\hat{K}_n, m_n^*}(z_1) - m^*(z_1)|^2 \mathbf{P}_{Z_1}(dz_1) \\
&\leq \left(\|S_{\hat{K}_n, m_n^*} - m^*\|_{\infty, \hat{I}_n}\right)^2 + \int_{z_1 \notin \hat{I}_n} S_{\hat{K}_n, m_n^*}(z_1)^2 \mathbf{P}_{Z_1}(dz_1) \\
&\leq \left(\|S_{\hat{K}_n, m_n^*} - m_n^*\|_\infty\right)^2 + \int_{z_1 \notin \hat{I}_n} S_{\hat{K}_n, m_n^*}(z_1)^2 \mathbf{P}_{Z_1}(dz_1),
\end{aligned} \tag{5.7}$$

da m^* außerhalb von $\hat{I}_n = [-\hat{L}_n^*, \hat{L}_n^*]$ verschwindet und m^* und m_n^* auf I_n übereinstimmen. Wir zeigen zunächst, dass der erste Summand der rechten Seite von (5.7) fast sicher gegen 0 konvergiert. Da m^* als glatte Funktion mit kompaktem Träger Lipschitz-stetig ist, ist auch jedes der m_n^* Lipschitz-stetig und alle diese Funktionen lassen die gleiche Konstante K als Lipschitz-Konstante zu.

Analog zum vorgehen im Beweis von Theorem 2.6 reskalieren wir sowohl m_n^* als auch die zugehörige Fourier-Summe auf 2π-periodische Funktionen, und dann den Approximationsfehler in der Supremumsnorm unter Verwendung von Korrolar 1 aus [48] abschätzen.

Es gilt, da die trigonometischen Polynome in $\mathcal{F}_n$ von Grad $\hat{K}_n$ sind,

$$\|S_{\hat{K}_n,m_n^*} - m_n^*\|_\infty \leq C\frac{\hat{L}_n \ln(\hat{K}_n)}{\hat{K}_n} \tag{5.8}$$

für eine geeignete Konstante $C > 0$. Nun zum zweiten Summanden der rechten Seite von (5.7). Wir nutzen erneut die obige Abschätzung für den Approximationsfehler von $S_{\hat{K}_n,m_n^*}$ bezüglich der Supremumsnorm sowie die Ungleichung

$$(a+b)^2 \leq 2a^2 + 2b^2$$

und erhalten

$$\begin{aligned}\int_{z_1 \notin \hat{I}_n} S_{\hat{K}_n,m_n^*}(z_1)^2 \mathbf{P}_{Z_1}(dz_1) &\leq \int_{\{|z_1| \geq \hat{L}_n^*\}} \|S_{\hat{K}_n,m_n^*}\|_\infty^2 \mathbf{P}_{Z_1}(dz_1) \\ &\leq 2\|S_{\hat{K}_n,m_n^*} - m_n^*\|_\infty^2 + 2\int_{\{|z_1| \geq \hat{L}_n^*\}} \|m_n^*\|_\infty^2 \mathbf{P}_{Z_1}(dz_1) \\ &\leq 2\left(C\frac{\hat{L}_n \ln(\hat{K}_n)}{\hat{K}_n}\right)^2 + 2\|m^*\|_\infty^2 \cdot \int_{\{|z_1| \geq \hat{L}_n^*\}} 1\mathbf{P}_{Z_1}(dz_1)\end{aligned} \tag{5.9}$$

für eine geeignete Konstante $C > 0$. Unter den Voraussetzungen des Lemmas sind auch die Voraussetzungen von Lemma 5.3 erfüllt und es gilt gilt $\hat{K}_n \to \infty$ fast sicher. Zusammen mit $\hat{L}_n^* = 2 \cdot \hat{L}_n = 2 \cdot \sqrt{\hat{K}_n}$ folgt, dass mit Wahrscheinlichkeit 1 sowohl

$$\|S_{\hat{K}_n,m_n^*} - m_n^*\|_\infty \to$$

als auch

$$\int_{z_1 \notin \hat{I}_n} S_{\hat{K}_n,m_n^*}(z_1)^2 \mathbf{P}_{Z_1}(dz_1) \to 0$$

gilt. Insgesamt folgt also $T_{2,n} \to 0$ fast sicher.

In Schritt 3 des Beweises werden wir nun die Konvergenz von $T_{1,n}$ zeigen. Hierzu verwenden wir, ähnlich unserem Vorgehen im zweiten Schritt, erneut ein Approximation durch trigonometrische Polynome. Da wir allerdings Funktionen $f(z_1) - z_2$ approxierem wollen, benötigten wir zur Approximation nun trigonometischen Polynome in zwei Variablen. Wir beginnen mit der periodischen Approximation der Identität. Sei $h_{1,n} : [-\hat{L}_n, \hat{L}_n] \to \mathbb{R}$ gegeben durch $h_{1,n}(z_2) := z_2$. Wir definieren $\tilde{h}_{1,n} : \mathbb{R} \to \mathbb{R}$ als $4\hat{L}_n = 2\hat{L}_n^*$-periodische Fortsetzung von $h_{1,n}$ durch

$$x \mapsto \begin{cases} 2\hat{L}_n - x & \text{für } x \in [\hat{L}_n, 3\hat{L}_n] \\ x & \text{für } x \in (-\hat{L}_n, \hat{L}_n). \end{cases}$$

Nun ist $\tilde{h}_{1,n}$ für alle $n \in \mathbb{N}$ eine periodische Fortsetzung von $h_{1,n}$ welche mit Lipschitz-Konstante 1 Lipschitz-stetig ist. Die Koeffizienten der Fourier-Reihe von $\tilde{h}_{1,n}$ sind durch $\|\tilde{h}_{1,n}\|_\infty = \hat{L}_n$ beschränkt.

Wegen

$$\hat{L}_n \leq \hat{K}_n \leq n^{\frac{1}{11}} = \beta_n,$$

liegt die $\hat{K}_n$-nte Fourier-Summe $g_{1,n}$ von $\tilde{h}_{1,n}$ in $\mathcal{F}_n$. Damit können wir

$$g_{f_n,n}(z_1, z_2) := f_n(z_1) - g_{1,n}(z_2)$$

für $f_n \in \mathcal{F}_n$ definieren. Für alle $f_n \in \mathcal{F}_n$ ist $g_{f_n,n}$ ein trigonometrisches Polynom in zwei Variablen vom Grad $\hat{K}_n$ und Periode $2\hat{L}_n^*$ in beiden Variablen. Reskalieren der Periode auf 2π und Anwenden des Korollar 1 aus [48] liefert

$$\begin{aligned}
&\sup_{z_1,z_2\in[-\hat{L}_n^*,\hat{L}_n^*]} \left|\left(f_n(z_1) - \tilde{h}_{1,n}(z_2)\right) - g_{f_n,n}(z_1,z_2)\right| \\
&= \sup_{z_2\in[-\hat{L}_n^*,\hat{L}_n^*]} \left|\tilde{h}_{1,n}(z_2) - g_{1,n}(z_2)\right| \to 0
\end{aligned}$$

für $n \to \infty$. Da $\tilde{h}_{1,n}$ eine Fortsetzung von $h_{1,n}$ ist, approximiert $g_{f_n,n}$ für $z_1, z_2 \in [-\hat{L}_n, \hat{L}_n]$ gleichmäßig $f_n(z_1) - z_2$. Außerdem hängt der Approximationsfehler nicht von f_n ab, da f_n nach Konstruktion in $g_{f_n,n}$ enthalten ist.

Nun zerlegen wir $T_{1,n}$. Es gilt

$$\begin{aligned}
T_{1,n} = \sup_{f\in\mathcal{F}_n} \Big[&\mathbf{E}\big\{|m_n(Z_1) - Z_2|^2 \big| \mathcal{D}_n\big\} - (1+\delta)^4 \frac{1}{n}\sum_{j=1}^n |m_n(\hat{z}_{1,j}) - \hat{z}_{2,j}|^2 \\
&+ (1+\delta)^4 \Big(\frac{1}{n}\sum_{j=1}^n |m_n(\hat{z}_{1,j}) - \hat{z}_{2,j}|^2 - \frac{1}{n}\sum_{j=1}^n |f(\hat{z}_{1,j}) - \hat{z}_{2,j}|^2\Big) \\
&+ (1+\delta)^4 \frac{1}{n}\sum_{j=1}^n |f(\hat{z}_{1,j}) - \hat{z}_{2,j}|^2 - (1+\delta)^8 \mathbf{E}\{|f(Z_1) - Z_2|^2\}\Big] \\
= \sup_{f\in\mathcal{F}_n} &\big[T_{1,n}^{(1)} + T_{1,n}^{(2)} + T_{1,n}^{(3)}\big].
\end{aligned}$$

Da m_n der Kleinste-Quadrate-Schätzer ist, gilt $T_{1,n}^{(2)} \le 0$. Also folgt

$$T_{1,n} \le \sup_{f\in\mathcal{F}_n} \big[T_{1,n}^{(1)} + T_{1,n}^{(3)}\big] = T_{1,n}^{(1)} + \sup_{f\in\mathcal{F}_n} T_{1,n}^{(3)}.$$

Die beiden Summanden auf der rechten Seite formen wir mithilfe der elementaren Ungleichung

$$(a+b)^2 - (1+\delta)a^2 \le (1+\frac{1}{\delta})b^2 \quad (\delta > 0\,; a, b \in \mathbb{R}) \tag{5.10}$$

weiter um. Es gilt

$$\begin{aligned}
T_{1,n}^{(1)} &= \mathbf{E}\big\{|m_n(Z_1) - Z_2|^2 \big| \mathcal{D}_n\big\} - (1+\delta)^4 \cdot \frac{1}{n}\sum_{j=1}^n |m_n(\hat{z}_{1,j}) - \hat{z}_{2,j}|^2 \\
&= \mathbf{E}\big\{|m_n(Z_1) - Z_2|^2 \big| \mathcal{D}_n\big\} - (1+\delta)\cdot \mathbf{E}\big\{|m_n(Z_1) - Z_{2,\hat{L}_n}|^2 \big| \mathcal{D}_n\big\} \\
&+ (1+\delta)\cdot \mathbf{E}\big\{|m_n(Z_1) - Z_{2,\hat{L}_n}|^2 \big| \mathcal{D}_n\big\} - (1+\delta)^2 \cdot \mathbf{E}\big\{g_{m_n,n}(Z_1, Z_{2,\hat{L}_n})^2 \big| \mathcal{D}_n\big\} \\
&+ (1+\delta)^2 \cdot \mathbf{E}\big\{g_{m_n,n}(Z_1, Z_{2,\hat{L}_n})^2 \big| \mathcal{D}_n\big\} - (1+\delta)^3 \cdot \mathbf{E}\big\{g_{m_n,n}(Z_1, Z_2)^2 \big| \mathcal{D}_n\big\} \\
&+ (1+\delta)^3 \cdot \Big(\mathbf{E}\big\{g_{m_n,n}(Z_1, Z_2)^2 \big| \mathcal{D}_n\big\} - \frac{1}{n}\sum_{j=1}^n g_{m_n,n}(\hat{z}_{1,j}, \hat{z}_{2,j})^2\Big)
\end{aligned}$$

$$+ (1+\delta)^3 \cdot \frac{1}{n} \sum_{j=1}^n g_{m_n,n}(\hat{z}_{1,j}, \hat{z}_{2,j})^2 - (1+\delta)^4 \cdot \frac{1}{n} \sum_{j=1}^n |m_n(\hat{z}_{1,j}) - \hat{z}_{2,j}|^2.$$

Mit (5.10) folgt

$$\begin{aligned}
T_{1,n}^{(1)} &\leq \left(1+\frac{1}{\delta}\right) \mathbf{E}\Big\{ \Big(m_n(Z_1) - Z_2 - \Big(m_n(Z_1) - Z_{2,\hat{L}_n} \Big) \Big)^2 \Big| \mathcal{D}_n \Big\} \\
&\quad + (1+\frac{1}{\delta})(1+\delta) \mathbf{E}\Big\{ \Big(m_n(Z_1) - Z_{2,\hat{L}_n} - g_{m_n,n}(Z_1, Z_{2,\hat{L}_n}) \Big)^2 \Big| \mathcal{D}_n \Big\} \\
&\quad + (1+\frac{1}{\delta})(1+\delta)^2 \mathbf{E}\Big\{ \Big(g_{m_n,n}(Z_1, Z_{2,\hat{L}_n}) - g_{m_n,n}(Z_1, Z_2) \Big)^2 \Big| \mathcal{D}_n \Big\} \\
&\quad + (1+\delta)^3 \Big[\mathbf{E}\Big\{ g_{m_n,n}(Z_1, Z_2)^2 \Big| \mathcal{D}_n \Big\} - \frac{1}{n} \sum_{j=1}^n g_{m_n,n}(\hat{z}_{1,j}, \hat{z}_{2,j})^2 \Big] \\
&\quad + (1+\delta)^3 (1+\frac{1}{\delta}) \frac{1}{n} \sum_{j=1}^n \Big(g_{m_n,n}(\hat{z}_{1,j}, \hat{z}_{2,j}) - (m_n(\hat{z}_{1,j}) - \hat{z}_{2,j}) \Big)^2 \\
&= \sum_{k=1}^5 \tilde{T}_{1,n}^{(k)}.
\end{aligned}$$

Mit der selben Technik können wir auch $T_{1,n}^{(3)}$ abschätzen. Dabei integrieren wir jedoch nicht über die zufälligen Stutzungshöhen $\hat{L}_n$, weswegen wir den in $T_{1,n}^{(3)}$ auftretenden Erwartungswert durch das Integral bezüglich $\mathbf{P}_{(Z_1,Z_2)}$ ersetzen. Es gilt

$$\begin{aligned}
T_{1,n}^{(3)} &= (1+\delta)^4 \frac{1}{n} \sum_{j=1}^n |f(\hat{z}_{1,j}) - \hat{z}_{2,j}|^2 - (1+\delta)^8 \mathbf{E}\{|f(Z_1) - Z_2|^2\} \\
&= (1+\delta)^4 \frac{1}{n} \sum_{j=1}^n |f(\hat{z}_{1,j}) - \hat{z}_{2,j}|^2 - (1+\delta)^5 \frac{1}{n} \sum_{j=1}^n |g_{f,n}(\hat{z}_{1,j}, \hat{z}_{2,j})|^2 \\
&\quad + (1+\delta)^5 \Big[\frac{1}{n} \sum_{j=1}^n \Big(g_{f,n}(\hat{z}_{1,j}, \hat{z}_{2,j}) \Big)^2 - \int g_{f,n}(z_1, z_2)^2 \mathbf{P}_{(Z_1,Z_2)} \Big] \\
&\quad + (1+\delta)^5 \int g_{f,n}(z_1, z_2)^2 \mathbf{P}_{(Z_1,Z_2)} - (1+\delta)^6 \int g_{f,n}(z_1, z_{2,\hat{L}_n})^2 \mathbf{P}_{(Z_1,Z_2)} \\
&\quad + (1+\delta)^6 \int g_{f,n}(z_1, z_{2,\hat{L}_n})^2 \mathbf{P}_{(Z_1,Z_2)} - (1+\delta)^7 \int (f(z_1) - z_{2,\hat{L}_n})^2 \mathbf{P}_{(Z_1,Z_2)} \\
&\quad (1+\delta)^7 \int (f(z_1) - z_{2,\hat{L}_n})^2 \mathbf{P}_{(Z_1,Z_2)} - (1+\delta)^8 \mathbf{E}\{|f(Z_1) - Z_2|^2\}
\end{aligned}$$

Mit (5.10) folgt also

$$\begin{aligned}
T_{1,n}^{(3)} &\leq \left(1+\frac{1}{\delta}\right) (1+\delta)^4 \frac{1}{n} \sum_{j=1}^n \Big(f(\hat{z}_{1,j}) - \hat{z}_{2,j} - g_{f,n}(\hat{z}_{1,j}, \hat{z}_{2,j}) \Big)^2 \\
&\quad + (1+\delta)^5 \Big[\frac{1}{n} \sum_{j=1}^n \Big(g_{f,n}(\hat{z}_{1,j}, \hat{z}_{2,j}) \Big)^2 - \int g_{f,n}(z_1, z_2)^2 \mathbf{P}_{(Z_1,Z_2)} \Big] \\
&\quad + (1+\delta)^5 (1+\frac{1}{\delta}) \int \Big(g_{f,n}(z_1, z_2) - g_{f,n}(z_1, z_{2,\hat{L}_n}) \Big)^2 \mathbf{P}_{(Z_1,Z_2)}
\end{aligned}$$

$$
\begin{aligned}
&+ (1+\delta)^6 (1+\frac{1}{\delta}) \int \Big(g_{f,n}(z_1, z_{2,\hat{L}_n}) - \big(f(z_1) - z_{2,\hat{L}_n}\big)\Big)^2 \mathbf{P}_{(Z_1,Z_2)} \\
&+ (1+\delta)^7 (1+\frac{1}{\delta}) \int \Big(f(z_1) - z_{2,\hat{L}_n} - \big(f(z_1) - z_2\big)\Big)^2 \mathbf{P}_{(Z_1,Z_2)} \\
= &\sum_{k=6}^{10} \tilde{T}_{1,n}^{(k)}.
\end{aligned}
$$

Es gilt $m_n \in \mathcal{F}_n$ nach Definition des Schätzers. Also genügt es zu zeigen, dass die obigen Terme gegen 0 konvergieren, wenn wir m_n durch ein $f \in \mathcal{F}_n$ ersetzen, und dass diese Konvergenz gleichmäßig in $f \in \mathcal{F}_n$ ist. Dazu unterteilen wir die obigen Ausdrücke in 4 grobe Klassen. Im folgenden bezeichnen wir $\tilde{T}_{1,n}^{(1)}$, $\tilde{T}_{1,n}^{(3)}$, $\tilde{T}_{1,n}^{(8)}$ und $\tilde{T}_{1,n}^{(10)}$ als Terme vom Typ i. Für diese lässt sich die gewünschte Konvergenz aus der Konvergenz von $Z_{2,\hat{L}_n}$ gegen Z_2 folgern. $\tilde{T}_{1,n}^{(2)}$ und $\tilde{T}_{1,n}^{(9)}$ sind vom Typ ii. Hier werden wir die Eigenschaften von $g_{f,n}$ ausnutzen. $\tilde{T}_{1,n}^{(5)}$ und $\tilde{T}_{1,n}^{(6)}$ bilden den Typ iii, für den wir Abschätzungstechniken für Typ i und Typ ii kombinieren werden. Die verbleibenden Terme $\tilde{T}_{1,n}^{(4)}$ und $\tilde{T}_{1,n}^{(7)}$ konstituieren Typ iv. Terme vom Typ iv konvergieren nach Lemma 5.8 gegen 0 sofern diese anwendbar ist.

Wir beginnen mit den Termen vom Typ i. Für $f \in \mathcal{F}_n$ gilt

$$
g_{f,n}(x,y) = f(x) - g_{1,n}(y),
$$

wobei wie zuvor $g_{1,n}$ unsere zu Beginn von Schritt 3 definierte trigonometrische Approximation der Identität ist. Es gilt

$$
\begin{aligned}
\tilde{T}_{1,n}^{(1)} &= \left(1+\frac{1}{\delta}\right) \mathbf{E}\left\{ |Z_2 - Z_{2,\hat{L}_n}|^2 | \mathcal{D}_n \right\} \\
\tilde{T}_{1,n}^{(3)} &= \left(1+\frac{1}{\delta}\right) (1+\delta)^2 \mathbf{E}\left\{ |g_{1,n}(Z_2) - g_{1,n}(Z_{2,\hat{L}_n})|^2 | \mathcal{D}_n \right\} \\
\tilde{T}_{1,n}^{(8)} &= \left(1+\frac{1}{\delta}\right) (1+\delta)^5 \int \Big(g_{1,n}(z_2) - g_{1,n}(z_{2,\hat{L}_n}) \Big)^2 \mathbf{P}_{(Z_1,Z_2)}
\end{aligned}
$$

und

$$
\tilde{T}_{1,n}^{(10)} = \left(1+\frac{1}{\delta}\right) (1+\delta)^7 \int \Big(z_2 - z_{2,\hat{L}_n} \Big)^2 \mathbf{P}_{(Z_1,Z_2)}.
$$

Also sind die in den Termen vom Typ i auftretenden Differenzen unabhängig von $f \in \mathcal{F}_n$. Nun leiten wir eine Lipschitz-Konstante für $g_{1,n}$ her. Wegen $g_{1,n} \in \mathcal{F}_n$ gilt

$$
g_{1,n}(z_1) \quad = \quad \sum_{k=0}^{\hat{K}_n} a_k \cdot \cos\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_1 \right) + \sum_{k=0}^{\hat{K}_n} b_k \cdot \sin\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_1 \right).
$$

Wobei die Koeffizienten $a_0, \ldots, a_{\hat{K}_n}$ und $b_0, \ldots, b_{\hat{K}_n}$ durch $\hat{L}_n$ beschränkt sind, da es sich um die Fourier-Koeffizienten von $\tilde{h}_{1,n}$ handelt. Da Cosinus und Sinus Lipschitz-stetig mit Lipschitz-Konstante 1 sind, folgt damit

$$\begin{aligned}
|g_{1,n}(x) - g_{1,n}(y)| &= \Bigg| \sum_{k=0}^{\hat{K}_n} a_k \cdot \left[\cos\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot x\right) - \left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot y\right)\right] \\
&\quad + \sum_{k=0}^{\hat{K}_n} b_k \left[\cdot \sin\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot x\right) - \cdot \sin\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot y\right)\right] \Bigg| \\
&\leq \sum_{k=0}^{\hat{K}_n} [|a_k| + |b_k|] \left|\frac{\pi}{\hat{L}_n^*}\right| |x-y| \\
&\leq 2\pi \hat{L}_n \hat{K}_n \frac{1}{\hat{L}_n^*} |x-y|.
\end{aligned} \tag{5.11}$$

Somit existiert ein $C \in \mathbb{R}$ für welches $C\hat{K}_n$ eine Lipschitz-Konstante von $g_{1,n}$ ist. Da δ fest ist, genügt es zu zeigen, dass

$$\hat{K}_n \int \left(z_{2,\hat{L}_n} - z_2\right)^2 \mathbf{P}_{(Z_1,Z_2)} \to 0 \tag{5.12}$$

mit Wahrscheinlichkeit 1 gilt. Nach Voraussetzung gilt $\mathbf{E}\{|Z_2|^5\} < \infty$. Außerdem gilt

$$\begin{aligned}
(x - x_{\hat{L}_n})^2 &\leq x^2 \mathbb{1}_{\{|x| \geq \hat{L}_n\}} \\
&\leq x^2 \frac{|x|^3}{\hat{L}_n^3}
\end{aligned} \tag{5.13}$$

für alle $x \in \mathbb{R}$. Damit folgt

$$\hat{K}_n \int \left(z_{2,\hat{L}_n} - z_2\right)^2 \mathbf{P}_{(Z_1,Z_2)} \leq \frac{1}{\hat{L}_n} \mathbf{E}\{|Z_2|^5\} \to 0$$

fast sicher, da nach Voraussetzung $\hat{L}_n \to \infty$ fast sicher gilt. Wenden wir (5.13) auf $\mathbf{E}\{(Z_2 - Z_{2,\hat{L}_n})^2 | \mathcal{D}_n\}$ an, so folgt

$$\mathbf{E}\{(Z_2 - Z_{2,\hat{L}_n})^2 | \mathcal{D}_n\} \leq \mathbf{E}\{|Z_2|^5\} \frac{1}{\hat{L}_n^3},$$

da Z_2 von $\mathcal{D}_n$ unabhängig ist und da $\hat{L}_n$ messbar bezüglich $\mathcal{D}_n$ ist. Also folgt auch für die Terme mit der bedingten Erwartung die gewünschte Konvergenz.

Als Nächstes behandeln wir die Ausdrücke vom Typ ii. Diese sind

$$\tilde{T}_{1,n}^{(2)} = (1 + \frac{1}{\delta})(1+\delta) \mathbf{E}\Big\{ \Big(m_n(Z_1) - Z_{2,\hat{L}_n} - g_{m_n,n}(Z_1, Z_{2,\hat{L}_n}) \Big)^2 \Big| \mathcal{D}_n \Big\}$$

und

$$\tilde{T}_{1,n}^{(9)} = (1+\delta)^6 (1 + \frac{1}{\delta}) \int \Big(g_{f,n}(z_1, z_{2,\hat{L}_n}) - \big(f(z_1) - z_{2,\hat{L}_n}\big) \Big)^2 \mathbf{P}_{(Z_1,Z_2)}$$

Wie zuvor hängen die Terme wegen

$$g_{f,n}(x,y) = f(x) - g_{1,n}(y)$$

nicht von m_n und f ab. Es genügt daher zu zeigen, dass jeweils mit Wahrscheinlichkeit 1

$$\mathbf{E}\{(Z_{2,\hat{L}_n} - g_{1,n}(Z_{2,\hat{L}_n}))^2|\mathcal{D}_n\} \to 0$$

und

$$\int (g_{1,n}(z_{2,\hat{L}_n}) - z_{2,\hat{L}_n})^2 \mathbf{P}_{(Z_1,Z_2)} \to 0$$

gilt. Wegen $Z_{2,\hat{L}_n} \in [-\hat{L}_n, \hat{L}_n]$ und da

$$\tilde{h}_{1,n}(x) = \begin{cases} 2\hat{L}_n - x & \text{für } x \in [\hat{L}_n, 3\hat{L}_n] \\ x & \text{für } x \in (-\hat{L}_n, \hat{L}_n). \end{cases}$$

auf $[-\hat{L}_n, \hat{L}_n]$ mit der Identität übereinstimmt, genügt es, die gleichmäßige Konvergenz von $g_{1,n}$ gegen $\tilde{h}_{1,n}$ zu zeigen. Nach Definiton ist $\tilde{h}_{1,n}$ Lipschitz-stetig, und die Lipschitz-Konstante hängt nicht von n ab. Da $g_{1,n}$ gerade das n-te Fourier-Polynom von $\tilde{h}_{1,n}$ ist, und da beide Funktionen $2\hat{L}_n^*$-periodisch sind, können wir Korollar 1 aus [48] auf die reskalierten Funktionen $x \mapsto g_{1,n}\big(\frac{\hat{L}_n^*}{\pi} \cdot x\big)$ und $x \mapsto \tilde{h}_{1,n}\big(\frac{\hat{L}_n^*}{\pi} \cdot x\big)$ anwenden. Damit folgt

$$\|g_{1,n} - \tilde{h}_{1,n}\|_\infty \leq C \frac{\hat{L}_n \ln(\hat{K}_n)}{\hat{K}_n}.$$

Die rechte Seite der Gleichung konvergiert fast sicher gegen 0, da $\hat{L}_n = \sqrt{\hat{K}_n}$ gilt und $\hat{K}_n$ nach Lemma 5.3 fast sicher gegen ∞ konvergiert. Da $\hat{L}_n$ messbar bezüglich $\mathcal{D}_n$ ist, und da $\mathbf{P}_{(Z_1,Z_2)}$ ein Wahrscheinlichkeitsmaß ist, konvergieren also auch

$$\mathbf{E}\Big\{\Big(m_n(Z_1) - Z_{2,\hat{L}_n} - g_{m_n,n}(Z_1, Z_{2,\hat{L}_n})\Big)^2 \Big|\mathcal{D}_n\Big\} \leq \mathbf{E}\left\{C \frac{\hat{L}_n \ln(\hat{K}_n)}{\hat{K}_n} |\mathcal{D}_n\right\}$$

und

$$\int \Big(g_{f,n}(z_1, z_{2,\hat{L}_n}) - \big(f(z_1) - z_{2,\hat{L}_n}\big)\Big)^2 \mathbf{P}_{(Z_1,Z_2)} \leq C \frac{\hat{L}_n \ln(\hat{K}_n)}{\hat{K}_n} \int 1 \mathbf{P}_{(Z_1,Z_2)}$$

fast sicher gegen 0 und die Konvergenzgeschwindigkeit hängt weder von m_n noch von $f \in \mathcal{F}_n$ ab.

Nun zu Typ iii. Da δ konstant ist, können wir die ersten beiden Faktoren von $\tilde{T}_{1,n}^{(5)}$ und $\tilde{T}_{1,n}^{(6)}$ jeweils vernachlässigen. Nach unserer Definition von $g_{m_n,n}$ und $g_{f,n}$ gilt

$$\begin{aligned} &\frac{1}{n} \sum_{j=1}^{n} \Big(f(\hat{z}_{1,j}) - \hat{z}_{2,j} - g_{f,n}(\hat{z}_{1,j}, \hat{z}_{2,j})\Big)^2 \\ &= \frac{1}{n} \sum_{j=1}^{n} \Big(m_n(\hat{z}_{1,j}) - \hat{z}_{2,j} - g_{m_n,n}(\hat{z}_{1,j}, \hat{z}_{2,j})\Big)^2 \\ &= \frac{1}{n} \sum_{j=1}^{n} (\hat{z}_{2,j} - g_{1.n}(\hat{z}_{2,j}))^2 = J_n. \end{aligned}$$

Also hängen sowohl $\tilde{T}_{1,n}^{(5)}$ als auch $\tilde{T}_{1,n}^{(6)}$ nicht von $f \in \mathcal{F}_n$ ab. Um J_n nach oben abzuschätzen stutzen wir nun auch die $\hat{z}_{2,j}$ mit der zufälligen Stutzungshöhe $\hat{L}_n$. Sei

$$\hat{z}_{2,j,\hat{L}_n} = \begin{cases} \hat{z}_2 \text{ falls } |\hat{z}_2| \leq \hat{L}_n \\ \operatorname{sign}(\hat{z}_2) \text{ sonst.} \end{cases}$$

Nun gilt

$$\begin{aligned} J_n &\leq 3\frac{1}{n}\sum_{j=1}^{n}\Big(|\hat{z}_{2,j} - \hat{z}_{2,j,\hat{L}_n}|^2 + |\hat{z}_{2,j,\hat{L}_n} - g_{1,n}(\hat{z}_{2,j,\hat{L}_n})|^2 \\ &\qquad + |g_{1,n}(\hat{z}_{2,j,\hat{L}_n}) - g_{1,n}(\hat{z}_{2,j})|^2\Big) \\ &= 3 \cdot (J_{1,n} + J_{2,n} + J_{3,n}). \end{aligned}$$

Da $\hat{z}_{2,j,\hat{L}_n}$ durch $\hat{L}_n$ beschränkt ist, können wir für $J_{2,n}$ die gleichen Abschätzungen wie für die Terme vom Typ ii verwenden. Also gilt $J_{2,n} \to 0$ fast sicher. Aus der Abschätzung (5.11) ist bekannt, dass ein $C > 0$ exisitert, so dass $C\hat{K}_n$ eine Lipschitz-Konstante für $g_{1,n}$ ist. Damit folgt

$$J_{3,n} \leq C^2\hat{K}_n \sum_{j=1}^{n}|\hat{z}_{2,j,\hat{L}_n} - \hat{z}_{2,j}|^2.$$

Nun benutzen wir eine ähnliche Abschätzung wie in (5.13). Es gilt

$$\begin{aligned} \sum_{j=1}^{n}|\hat{z}_{2,j} - \hat{z}_{2,j,\hat{L}_n}|^2 &\leq \sum_{j=1}^{n}(\hat{z}_{2,j})^2\frac{|\hat{z}_{2,j}|^4}{(\hat{L}_n)^4} \\ &= \frac{1}{(\hat{L}_n)^4}\frac{1}{n}\sum_{j=1}^{n}|\hat{z}_{2,j}|^6 \\ &\leq (\hat{K}_n)^{-1.5}, \end{aligned}$$

da $\frac{1}{n} \cdot \sum_{i=1}^{n}|\hat{z}_{2,j}|^6 \leq \sqrt{\hat{K}_n}$ nach Definition der geschätzten Daten sowie $\hat{L}_n = \sqrt{\hat{K}_n}$ gilt. Aus $\hat{K}_n \to \infty$ fast sicher folgt nun $J_{1,n} \to 0$ und $J_{3,n} \to 0$ fast sicher.

Nun müssen nur noch die Ausdrücke vom Typ iv, also

$$\tilde{T}_{1,n}^{(4)} = (1+\delta)^3\Big[\mathbf{E}\Big\{g_{m_n,n}(Z_1,Z_2)^2\Big|\mathcal{D}_n\Big\} - \frac{1}{n}\sum_{j=1}^{n}g_{m_n,n}(\hat{z}_{1,j},\hat{z}_{2,j})^2\Big]$$

und

$$\tilde{T}_{1,n}^{(7)} = (1+\delta)^5\Big[\frac{1}{n}\sum_{j=1}^{n}\Big(g_{f,n}(\hat{z}_{1,j},\hat{z}_{2,j})\Big)^2 - \int g_{f,n}(z_1,z_2)^2\mathbf{P}_{(Z_1,Z_2)}\Big],$$

behandelt werden. Liegen $g_{f,n}^2$ und $g_{m_n,n}$ in $\mathcal{G}_n$, so folgt

$$\tilde{T}_{1,n}^{(4)} \leq \sup_{g \in \mathcal{G}_n}\left|\frac{1}{n}\sum_{l=1}^{n}g(\hat{z}_{1,l},\hat{z}_{2,l}) - \int g(Z_1,Z_2)\mathbf{P}_{(Z_1,Z_2)}(d(z_1,z_2))\right|$$

und

$$\sup_{f\in\mathcal{F}_n} \tilde{T}_{1,n}^{(7)} \leq \sup_{g\in\mathcal{G}_n} \left| \frac{1}{n}\sum_{l=1}^{n} g(\hat{z}_{1,l},\hat{z}_{2,l}) - \int g(Z_1,Z_2)\mathbf{P}_{(Z_1,Z_2)}(d(z_1,z_2)) \right|.$$

Da unter den Voraussetzungen des Lemmas auch die Voraussetzungen von Lemma 5.6 und damit auch die Voraussetzungen von Lemma 5.8 erfüllt sind, konvergieren die rechten Seiten der Ungleichungen mit Wahrscheinlichkeit 1 gegen 0. Da m_n und f in $\mathcal{F}_n$ liegen, genügt es also zu zeigen, dass für jedes f aus $\mathcal{F}_n$ auch $g_{f,n}^2 \in \mathcal{G}_n$ gilt um den Beweis zu beenden.

Wir zeigen nur die wichtigsten Umformungen der umfangreichen aber unkomplizierten benötigten Rechnung. Wir verwenden die trigonometrischen Identitäten

$$\sin(x)\sin(y) = \frac{1}{2}\big(\cos(x-y) - \cos(x+y)\big), \tag{5.14}$$

$$\cos(x)\cos(y) = \frac{1}{2}\big(\cos(x-y) + \cos(x+y)\big) \tag{5.15}$$

und

$$\sin(x)\cos(y) = \frac{1}{2}\big(\sin(x-y) + \sin(x+y)\big). \tag{5.16}$$

Sei $f_n \in \mathcal{F}_n$, dann gilt

$$\begin{aligned} g_{f_n,n}^2(z_1,z_2) &= \big(f_n(z_1) - g_{f_n,n}(z_2)\big)^2 \\ &= f_n(z_1)^2 + g_{1,n}(z_2)^2 - 2f_n(z_1)g_{1,n}(z_2). \end{aligned}$$

Da f und $g_{1,n}$ in $\mathcal{F}_n$ liegen, gibt es Koeffizienten $a_0, \tilde{a}_0, \ldots, a_{\hat{K}_n}, \tilde{a}_{\hat{K}_n}$ und $b_0, \tilde{b}_0, \ldots, b_{\hat{K}_n}, \tilde{b}_{\hat{K}_n}$ mit $a_j, \tilde{a}_j, b_j, \tilde{b}_j \in [-\beta_n, \beta_n]$ und

$$f(z_1) = \sum_{k=0}^{\hat{K}_n} a_k \cdot \cos\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_1\right) + \sum_{k=0}^{\hat{K}_n} b_k \cdot \sin\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_1\right)$$

sowie

$$g(z_2) = \sum_{k=0}^{\hat{K}_n} \tilde{a}_k \cdot \cos\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_2\right) + \sum_{k=0}^{\hat{K}_n} \tilde{b}_k \cdot \sin\left(k \cdot \frac{\pi}{\hat{L}_n^*} \cdot z_2\right).$$

Also folgt

$$\begin{aligned} f_n(z_1)\cdot g_{1,n}(z_2) = \sum_{k,j=0}^{\hat{K}_n} \Bigg(& a_k\tilde{a}_j \cos\left(k\cdot\frac{\pi}{\hat{L}_n^*}\cdot z_1\right)\cos\left(j\cdot\frac{\pi}{\hat{L}_n^*}\cdot z_2\right) \\ & + b_k\tilde{a}_j \sin\left(k\cdot\frac{\pi}{\hat{L}_n^*}\cdot z_1\right)\cos\left(j\cdot\frac{\pi}{\hat{L}_n^*}\cdot z_2\right) \\ & + a_k\tilde{b}_j \cos\left(k\cdot\frac{\pi}{\hat{L}_n^*}\cdot z_1\right)\sin\left(j\cdot\frac{\pi}{\hat{L}_n^*}\cdot z_2\right) \end{aligned}$$

$$+b_k\tilde{b}_j \sin\left(k\cdot\frac{\pi}{\hat{L}_n^*}\cdot z_1\right)\sin\left(j\cdot\frac{\pi}{\hat{L}_n^*}\cdot z_2\right)\Bigg).$$

Damit ist $2f_n(z_1)\cdot g_{1,n}(z_2)$ ein trigonometrisches Polynom vom Grad $\hat{K}_n$ in zwei Variablen dessen Koeffizienten durch $2\beta_n^2$ beschränkt sind. Außerdem gilt

$$\begin{aligned}
f_n(z_1)^2 &= \sum_{k,j=0}^{\hat{K}_n}\Bigg(a_ka_j\cos\left(k\cdot\frac{\pi}{\hat{L}_n^*}\cdot z_1\right)\cos\left(j\cdot\frac{\pi}{\hat{L}_n^*}\cdot z_1\right)\\
&\quad + b_ka_j\sin\left(k\cdot\frac{\pi}{\hat{L}_n^*}\cdot z_1\right)\cos\left(j\cdot\frac{\pi}{\hat{L}_n^*}\cdot z_1\right)\\
&\quad + a_kb_j\cos\left(k\cdot\frac{\pi}{\hat{L}_n^*}\cdot z_1\right)\sin\left(j\cdot\frac{\pi}{\hat{L}_n^*}\cdot z_1\right)\\
&\quad + b_kb_j\sin\left(k\cdot\frac{\pi}{\hat{L}_n^*}\cdot z_1\right)\sin\left(j\cdot\frac{\pi}{\hat{L}_n^*}\cdot z_1\right)\Bigg)\\
&= \sum_{k,j=0}^{\hat{K}_n}\frac{a_ka_j}{2}\left[\cos\left(\frac{(k-j)\pi}{\hat{L}_n^*}\cdot z_1\right)+\cos\left(\frac{(k+j)\pi}{\hat{L}_n^*}\cdot z_1\right)\right]\\
&\quad + \frac{b_ka_j}{2}\left[\sin\left(\frac{(k-j)\pi}{\hat{L}_n^*}\cdot z_1\right)+\sin\left(\frac{(k+j)\pi}{\hat{L}_n^*}\cdot z_1\right)\right]\\
&\quad + \frac{a_kb_j}{2}\left[\sin\left(\frac{(j-k)\pi}{\hat{L}_n^*}\cdot z_1\right)+\sin\left(\frac{(j+k)\pi}{\hat{L}_n^*}\cdot z_1\right)\right]\\
&\quad + \frac{b_kb_j}{2}\left[\cos\left(\frac{(k-j)\pi}{\hat{L}_n^*}\cdot z_1\right)-\cos\left(\frac{(k+j)\pi}{\hat{L}_n^*}\cdot z_1\right)\right]\\
&= \sum_{k,j=0}^{\hat{K}_n}\frac{a_ka_j}{2}\left[\cos\left(\frac{(k-j)\pi}{\hat{L}_n^*}\cdot z_1\right)+\cos\left(\frac{(k+j)\pi}{\hat{L}_n^*}\cdot z_1\right)\right]\\
&\quad + \frac{b_ka_j}{2}\left[\sin\left(\frac{(k-j)\pi}{\hat{L}_n^*}\cdot z_1\right)+\sin\left(\frac{(k+j)\pi}{\hat{L}_n^*}\cdot z_1\right)\right]\\
&\quad + \frac{a_kb_j}{2}\left[-\sin\left(\frac{(k-j)\pi}{\hat{L}_n^*}\cdot z_1\right)+\sin\left(\frac{(j+k)\pi}{\hat{L}_n^*}\cdot z_1\right)\right]\\
&\quad + \frac{b_kb_j}{2}\left[\cos\left(\frac{(k-j)\pi}{\hat{L}_n^*}\cdot z_1\right)-\cos\left(\frac{(k+j)\pi}{\hat{L}_n^*}\cdot z_1\right)\right].
\end{aligned}$$

Wobei wir die trigonometrischen Identitäten (5.14) bis (5.16) sowie die Symetrie des Sinus verwendet haben. Umsortieren und Zusammenfassen der Summanden liefert

$$f_n(z_1)^2 = \sum_{p=0}^{2\hat{K}_n}\Bigg(\cos\left(\frac{\pi p}{L_n^*}z_1\right)\left[\sum_{\substack{0\le k\le\hat{K}_n\\ 0\le j\le\hat{K}_n\\ k-j=p}}\frac{a_ka_j+b_kb_j}{2}+\sum_{\substack{0\le k\le\hat{K}_n\\ 0\le j\le\hat{K}_n\\ k+j=p}}\frac{a_ka_j-b_kb_j}{2}\right]$$

$$+\sin\left(\frac{\pi p}{L_n^*}z_1\right)\left[\sum_{\substack{0\le k\le \hat{K}_n\\ 0\le j\le \hat{K}_n\\ k-j=p}}\frac{a_jb_k-a_kb_j}{2}+\sum_{\substack{0\le k\le \hat{K}_n\\ 0\le j\le \hat{K}_n\\ k+j=p}}\frac{a_jb_k+a_kb_j}{2}\right].$$

Da es für jedes $p\in\{0,\dots,2\hat{K}_n\}$ höchstens $\hat{K}_n+1$ Möglichkeiten gibt, k und j aus $\{0,\dots,\hat{K}_n\}$ zu wählen, so dass $k-j=p$ gilt, und da es für jedes $p\in\{0,\dots,2\hat{K}_n\}$ höchstens $\hat{K}_n+1$ Möglichkeiten gibt, k und j aus $\{0,\dots,\hat{K}_n\}$ zu wählen, so dass $k+j=p$ gilt, folgt, dass $f_n(z_1)^2$ ein trigonometrisches Polynom vom Grad $2\cdot\hat{K}_n$ ist, und alle Koeffizienten durch $(2\cdot\hat{K}_n+2)\cdot\beta_n^2$ beschränkt sind. Da $g_{1,n}$ ebenfalls in $\mathcal{F}_n$ liegt, ist auch $g_{1,n}^2$ ein trigonometrisches Polynom vom Grad $2\cdot\hat{K}_n$, dessen Koeffizienten durch $(2\cdot\hat{K}_n+2)\cdot\beta_n^2$ beschränkt sind. Mit Wahrscheinlichkeit 1 gilt $\hat{K}_n\ge 1$ für n groß genug. Für solche n gilt nach den Rechenregeln für Polynome, dass $g_{f_n,n}^2$ ein trigonometrisches Polynom vom Grad $2\cdot\hat{K}_n$ ist, dessen Koeffizienten durch

$$2(2\cdot\hat{K}_n+2)\beta_n^2+2\beta_n^2=4\hat{K}_n+6\beta_n^2\le 10\cdot\hat{K}_n\cdot\beta_n^2$$

beschränkt sind. Somit liegt $g_{f_n,n}^2$ mit Wahrscheinlichkeit 1 für n groß genug in $\mathcal{G}_n$. Wir können also Lemma 5.8 anwenden und erhalten

$$T_{1,n}^{(4)}\to 0 \quad\text{und}\quad T_{1,n}^{(7)}\to 0$$

fast sicher. Damit ist der Beweis beendet. □

Beweis von Theorem 4.4. Unter den Annahmen von Theorem 4.4 sind alle Voraussetzungen von Lemma 5.6 erfüllt. Damit folgt Theorem 4.4 direkt aus Lemma 5.9. □

Literatur

[1] V. Glivenko, "Sulla determinazione empirica delle leggi di probabilita," *Giorn. Ist. Ital. Attuari*, vol. 4, no. 1, pp. 92–99, 1933.

[2] L. Devroye and L. Györfi, "No empirical probability measure can converge in the total variation sense for all distributions," *The Annals of Statistics*, vol. 18, no. 3, pp. 1496–1499, 1990.

[3] H. Scheffé, "A useful convergence theorem for probability distributions," *The Annals of Mathematical Statistics*, vol. 18, no. 3, pp. 434–438, 1947.

[4] E. Parzen, "On estimation of a probability density function and mode," *The annals of mathematical statistics*, vol. 33, no. 3, pp. 1065–1076, 1962.

[5] M. Rosenblatt, "Remarks on some nonparametric estimates of a density function," *The Annals of Mathematical Statistics*, vol. 27, no. 3, pp. 832–837, 1956.

[6] R. Mnatsakanov and E. Khmaladze, "On L1-convergence of statistical kernel estimators of distribution densities," in *Soviet Mathematics Doklady*, vol. 23, pp. 633–636, 1981.

[7] L. Devroye, "The equivalence of weak, strong and complete convergence in L1 for kernel density estimates," *The Annals of Statistics*, vol. 11, no. 3, pp. 896–904, 1983.

[8] Y.-S. Chow, S. Geman, and L.-D. Wu, "Consistent cross-validated density estimation," *The Annals of Statistics*, vol. 11, no. 1, pp. 25–38, 1983.

[9] L. Devroye and G. Lugosi, *Combinatorial methods in density estimation*. Springer Science & Business Media, 2012.

[10] L. Devroye, "Asymptotic performance bounds for the kernel estimate," *The Annals of Statistics*, vol. 16, no. 3, pp. 1162–1179, 1988.

[11] L. Devroye, "A universal lower bound for the kernel estimate," *Statistics & Probability Letters*, vol. 8, no. 5, pp. 419–423, 1989.

[12] S. Abou-Jaoude, "Sur une condition nécessaire et suffisante de L1-convergence presque complete de lâestimateur de la partition fixe pour une densité," *CR Acad. Sci. Paris Sér. AB*, vol. 283, p. 16, 1976.

[13] M. Rosenblatt, "Curve estimates," *The Annals of Mathematical Statistics*, vol. 42, no. 6, pp. 1815–1842, 1971.

[14] J. Johannes, "Deconvolution with Unknown Error Distribution," *The Annals of Statistics*, vol. 37, no. 5A, pp. 2301–2323, 2009.

[15] A. Meister, *Deconvolution Problems in Nonparametric Statistics*. Berlin: Springer Verlag, 2009.

[16] A.-K. Bott, L. Devroye, and M. Kohler, "Estimation of a distribution from data with small measurement errors," *Electronic Journal of Statistics*, vol. 7, pp. 2457–2476, 2013.

[17] L. Györfi, M. Kohler, A. Krzyżak, and H. Walk, *A Distribution-Free Theory of Nonparametric Regression*. New York: Springer Verlag, 2002.

[18] C. J. Stone, "Consistent nonparametric regression," *Annals of Statistics*, vol. 5, no. 4,

pp. 595–645, 1977.

[19] L. Devroye, L. Györfi, A. Krzyzak, and G. Lugosi, "On the strong universal consistency of nearest neighbor regression function estimates," *The Annals of Statistics*, vol. 22, no. 3, pp. 1371–1385, 1994.

[20] L. Györfi, "Recent results on nonparametric regression estimate and multiple classification," *Problems of Control and Information Theory*, vol. 10, no. 1, pp. 43–52, 1981.

[21] E. A. Nadaraya, "On estimating regression," *Theory of Probability and its Applications*, vol. 9, no. 1, pp. 141–142, 1964.

[22] G. S. Watson, "Smooth regression analysis," *Sankhya Series A*, vol. 26, pp. 359–372, 1964.

[23] G. Lugosi and K. Zeger, "Nonparametric estimation via empirical risk minimization," *IEEE Transactions on Information Theory*, vol. 41, no. 3, pp. 677–687, 1995.

[24] M. Kohler and A. Krzyżak, "Nonparametric regression estimation using penalized least squares," *IEEE Transactions on Information Theory*, vol. 47, no. 7, pp. 3054–3058, 2001.

[25] G. Lugosi and K. Zeger, "Nonparametric estimation via empirical risk minimization," *IEEE Transactions on information theory*, vol. 41, no. 3, pp. 677–687, 1995.

[26] D. D. Cox, "Approximation of least squares regression on nested subspaces," *The Annals of Statistics*, vol. 16, no. 2, pp. 713–732, 1988.

[27] L. Birgé and P. Massart, "Rates of convergence for minimum contrast estimators," *Probability Theory and Related Fields*, vol. 97, no. 1-2, pp. 113–150, 1993.

[28] S. M. Schennach, "Nonparametric regression in the presence of measurement error," *Econometric Theory*, vol. 20, no. 06, pp. 1046–1093, 2004.

[29] K. A. Bollen, "Latent variables in psychology and the social sciences," *Annual review of psychology*, vol. 53, no. 1, pp. 605–634, 2002.

[30] A. Skrondal and S. Rabe-Hesketh, *Generalized latent variable modeling: Multilevel, longitudinal, and structural equation models*. Crc Press, 2004.

[31] R. E. Schumacker and G. A. Marcoulides, *Interaction and nonlinear effects in structural equation modeling.* Lawrence Erlbaum Associates Publishers, 1998.

[32] D. Connes, E. Ronchetti, and M.-P. Victoria-Feser, "Goodness of fit for generalized linear latent variables models," *Journal of the American Statistical Association*, vol. 105, no. 491, pp. 1126–1134, 2010.

[33] D. Paul, E. Bair, T. Hastie, and R. Tibshirani, ""Preconditioning" for feature selection and regression in high-dimensional problems," *Annals of Statistics*, vol. 36, no. 4, pp. 1595–1618, 2008.

[34] A. Kelava, M. Kohler, A. Krzyżak, and T. F. Schaffland, "Nonparametric estimation of a latent variable model," *Journal of Multivariate Analysis*, vol. 154, no. C, pp. 112–134, 2017.

[35] H. W. Marsh, Z. Wen, and K.-T. Hau, "Structural equation models of latent interaction and quadratic effects," in *Structural Equation Modeling: A Second Course*

(G. R. Hancock and R. Mueller, eds.), pp. 225–265, Charlotte, NC.: Information Age Publishing, 2006.

[36] L. Devroye and L. Györfi, *Nonparametric density estimation: the L1 view*, vol. 119. John Wiley & Sons Incorporated, 1985.

[37] G. Pólya, "Remarks on characteristic functions," in *Proc. First Berkeley Conf. on Math. Stat. and Prob*, pp. 115–123, 1949.

[38] A. Klenke, *Wahrscheinlichkeitstheorie*. Springer-Lehrbuch Masterclass, Springer Berlin Heidelberg, 2008.

[39] S. Lang, *Complex Analysis*. New York: Springer Verlag, 1999.

[40] H. Bauer, *Wahrscheinlichkeitstheorie*. Walter de Gruyter, 1991.

[41] J. Elstrodt, *Maß- und Integrationstheorie*. Springer-Lehrbuch, Springer Berlin Heidelberg, 2007.

[42] A. S. Besicovitch, *Almost periodic functions*, vol. 29. Dover publications New York, 1954.

[43] R. L. Cooke, "Almost-periodic functions," *The American Mathematical Monthly*, vol. 88, no. 7, pp. 515–526, 1981.

[44] L. Amerio and G. Prouse, *Almost-periodic functions and functional equations*. Springer Science & Business Media, 2013.

[45] E. Lukacs, *Characteristic Functions*. Griffin books of cognate interest, Hafner Publishing Company, 1970.

[46] M. Émile Borel, "Les probabilités dénombrables et leurs applications arithmétiques," *Rendiconti del Circolo Matematico di Palermo (1884-1940)*, vol. 27, no. 1, pp. 247–271, 1909.

[47] F. P. Cantelli, "Sulla probabilita come limite della frequenza," *Atti Accad. Naz. Lincei*, vol. 26, no. 1, pp. 39–45, 1917.

[48] D. Jackson, *The Theory of Approximation*. New York: American Mathematical Society, 1930.

[49] V. I. Bogachev, *Measure theory*, vol. 1. Springer Science & Business Media, 2007.

Wissenschaftlicher Werdegang

20.12.1986	Geboren in Köln
2006	Abitur, Gymnasium am Wirteltor, Düren
2010	Bachelor of Science in Mathematik, Technische Universität Darmstadt
2012	Master of Science in Mathematik, Technische Universität Darmstadt
2012–2017	Wissenschaftlicher Mitarbeiter und Promotionsstudium am Fachbereich Mathematik, Arbeitsgruppe Stochastik, Technische Universität Darmstadt