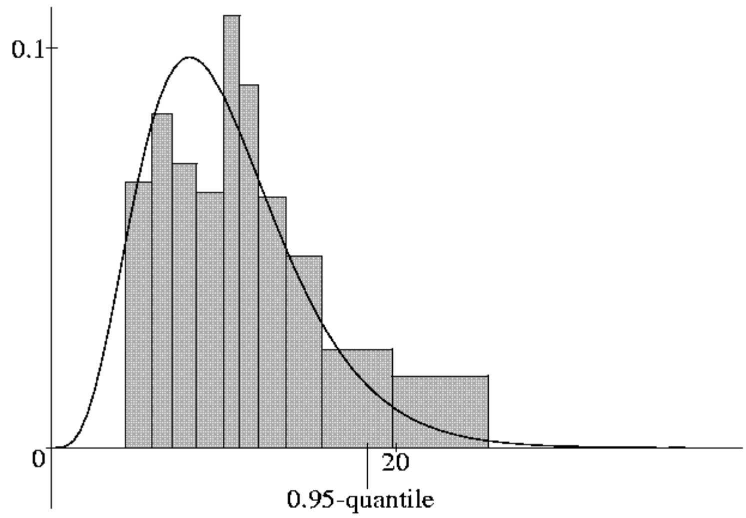
UNIVERSITÄT POTSDAM
Institut für Mathematik

Dualitätsformeln für Brownsche Bewegung und für eine Irrfahrt mit Anwendung am Konvergenzergebnis von Donsker

Diplomarbeit von Rüdiger Murr



Mathematische Statistik und
Wahrscheinlichkeitstheorie

Universität Potsdam – Institut für Mathematik

Mathematische Statistik und Wahrscheinlichkeitstheorie

Dualitätsformeln für Brownsche Bewegung und für eine Irrfahrt mit Anwendung am Konvergenzergebnis von Donsker

Rüdiger Murr
Institut für Mathematik der Universität Potsdam

Preprint 2008/05

Oktober 2008

Impressum

Herausgeber: Mathematische Statistik und Wahrscheinlichkeitstheorie
am Institut für Mathematik

Adresse: Universität Potsdam
PF 60 15 53
14415 Potsdam
Telefon:
Fax: +49-331-977 1500
E-mail: +49-331-977 1578
neisse@math.uni-potsdam.de

ISSN 1613-3307

Diplomarbeit im Fachbereich Mathematik der Universität Potsdam

Dualitätsformeln der Brownschen Bewegung und einer Irrfahrt, Anwendung an Donskers Konvergenzsatz

Betreuerin: Prof. Dr. Sylvie Roelly

Diplomarbeit von Rüdiger Murr

September 2008

Inhaltsverzeichnis

0.1 Danksagung

Ich möchte mich ganz herzlich bei allen bedanken, die mich bei der Anfertigung der Diplomarbeit unterstützt haben.
Mein besonderer Dank gilt Frau Prof. Sylvie Roelly. Auf sie geht nicht nur dir ursprüngliche Idee der Arbeit zurück. Ihre Anregungen und Kommentare sowie die Möglichkeit der Kontrolle und Diskussion während der Entstehung der Diplomarbeit waren eine große Hilfe. Auch während des Studiums gingen entscheidende Impulse von Frau Prof. Roelly aus, die mir eine Orientierung in der Stochastik sowie einen Auslandsaufenthalt ermöglichten.
Mein Dank gilt auch Herrn Prof. Christian Leonard, der ebenfalls für die Grundidee der Arbeit verantwortlich ist. Trotz der räumlichen Entfernung sind einige seiner wertvollen Anregungen in die Ausarbeitung mit eingegangen.
Ebenfalls danke ich der gesamten Arbeitsgruppe für Wahrscheinlichkeitstheorie aus Potsdam, die mehr oder weniger direkt helfenden Einfluss auf meine Arbeit hatte.
Auf dem Weg des Studiums haben mich viele Menschen begleitet. Vor allem meiner Familie danke ich für das ständig große Interesse und die Unterstützung.

0.2 Einleitung

Diese Diplomarbeit basiert auf der Theorie Wahrscheinlichkeitsmaße auf dem Raum der stetigen Funktionen im Intervall $[0, 1]$. Speziell für das Wienermaß auf diesem Raum wurde die in der Arbeit vorgestellte Dualitätsformel entwickelt. Von Hausmann und Bismut wurde 1978 bzw. 1981 gezeigt, dass das Wienermaß diese Dualitätsformel erfüllt (siehe Kapitel 2.1). Von Roelly und Zessin wurde dann 1991 gezeigt, dass daraus, dass die Dualitätsformel von einem Wahrscheinlichkeitsmaß auf dem Raum stetiger Funktionen erfüllt wird, automatisch folgt, dass dieses Wahrscheinlichkeitsmaß ein Wienermaß sein muss (siehe Kapitel 2.2). Man hat also eine Charakterisierung des Wienermaßes auf dem Raum stetiger Funktionen gefunden. Zum Vergleich sind zwei andere Charakterisierungen in Kapitel 1.1 zu finden.

Die Idee dieser Diplomarbeit ist nun, eine ähnliche Dualitätsformel für die symmetrische Irrfahrt zu finden. Auch diese Dualitätsformel soll charakteristisch, dass heißt notwendig und hinreichend, sein. Damit, dass solch eine Dualitätsformel existiert, beschäftigt sich Kapitel 3.1. Das eine Folge von Zufallsvariablen, die diese Dualitätsformel erfüllen, eine Irrfahrt sein muss, wird in Kapitel 3.2 behandelt. Der Begriff einer symmetrischen Irrfahrt wie er aus der Literatur bekannt ist, musste allerdings abgeschwächt werden, um eine Charakterisierung zu erhalten. Warum das so ist, wird in Kapitel 3 diskutiert.

Eine Rechtfertigung des Namens Dualitätsformel im Falle der Irrfahrt, der ja an den Namen Dualitätsformel vom Wienermaß her angelehnt ist, wird dann in Kapitel 4 geliefert. Dort wird gezeigt, dass die Dualitätsformel des Wienermaß asymptotisch von der Irrfahrt erfüllt wird. Natürlich bezieht sich die Konvergenzaussage auf eine entsprechend renormierte Version der Irrfahrt. Die Folgerung ist dann, zusammen mit einem Beweis der Straffheit der Wahrscheinlichkeitsmaße der Irrfahrt, ein Beweis des Konvergenztheorems von Donsker. Damit schließt diese Arbeit.

Ein Ausblick und auch ein Grund dafür, das dieses Thema gewählt wurde, ist, dass man die Konvergenz anderer diskreter Prozesse gegen ihr Pendant auf Pfadräumen ebenfalls mit Dualitätsformeln beweisen könnte beziehungsweise möchte.

1 Allgemeines und Grundlagen

Gegeben sei ein Wahrscheinlichkeitsraum $(\Omega, \mathcal{F}, P)$. In dieser Arbeit geht es zu großen Teilen um die Charakterisierung der Wahrscheinlichkeitsmaße stetiger Zufallsfunktionen. Eine stetige Zufallsfunktion X ist eine meßbare Abbildung

$$X : (\Omega, \mathcal{F}, P) \longmapsto (C(0,1), \mathcal{C}, P_X),$$

wobei $C(0,1)$ der Raum stetiger Funktionen über dem Intervall $[0,1]$ ist

$$C(0,1) := \{x : [0,1] \to \mathbb{R} \mid x \text{ ist stetig}\},$$

dessen Elemente wir später allgemein mit x bezeichnen werden. Außerdem verwenden wir die Schreibweise $x_t = x(t)$ für $t \in [0,1]$. $\mathcal{C}$ ist die $\sigma-$Algebra auf $C(0,1)$ erzeugt durch die offenen Mengen bezüglich der Supremumsnorm

$$\|x\|_\infty := \sup_{t \in [0,1]} |x(t)|.$$

Schließlich ist P_X das Bildmaß von P unter X. Wir werden die Bildmaße im weiteren Verlauf bei Bedarf konkretisieren. Wichtig ist der Fakt, dass schon die endlichdimensionalen Zylindermaße ein Maß auf $(C(0,1), \mathcal{C})$ charakterisieren. Dies ist eine Folgerung aus dem Theorem monotoner Klassen. Man kann diesen Sachverhalt unter anderem in [RW1] nachschlagen. Die Charakterisierung eines Maßes auf $(C(0,1), \mathcal{C})$ durch endlichdimensionale Zylindermaße wird im Folgenden immer wieder stillschweigend verwendet, und ist im Prinzip der Schlüssel, der uns die Tür zur Brownschen Bewegung öffnen wird.
Für ein festes $\omega \in \Omega$ bezeichnen wir mit

$$t \mapsto X_t(\omega), \ t \in [0,1]$$

einen Pfad der Zufallsfunktion X. Dieser ist nach Voraussetzung stetig für alle $\omega \in \Omega$. Eine Filtration auf $(\Omega, \mathcal{F}, P)$ ist eine Menge von $\sigma-$Algebren $(\mathcal{G}_t)_{0 \leq t \leq 1}$ für die

$$\forall s, t \in [0,1], \ s \leq t \quad \Rightarrow \quad \mathcal{G}_s \subset \mathcal{G}_t \subset \mathcal{F}$$

gilt. X ist genau dann $(\mathcal{G}_t)_{0 \leq t \leq 1}-$adaptiert, wenn X_t eine $\mathcal{G}_t$ messbare Zufallsvariable ist für alle $t \in [0,1]$. Bezüglich der sogenannten kanonischen Filtration

$$(\mathcal{F}_t)_{0 \leq t \leq 1} \text{ mit } \mathcal{F}_t := \sigma(X_s, 0 \leq s \leq t), \ t \in [0,1],$$

ist X immer adaptiert.
Um Wahrscheinlichkeitsmaße auf $(C(0,1), \mathcal{C})$ zu charakterisieren, werden wir verschiedene Funk-

tionenräume gebrauchen:

$$C_b := \{\phi : \mathbb{R}^n \to \mathbb{R} | \phi \text{ stetig und beschränkt}\},$$

$$C_b^\infty := \{\phi : \mathbb{R}^n \to \mathbb{R} | \phi \text{ beliebig oft differenzierbar, alle Ableitungen beschränkt}\},$$

um welchen $\mathbb{R}^n$ es sich handelt, wird aus den jeweiligen Zusammenhängen erkennbar sein.

1.1 Die Brownsche Bewegung

Wir beschäftigen uns viel mit dem Wienermaß bzw. der Brownschen Bewegung auf $C(0,1)$. Neben der Definition sollen deswegen einige alternative Charakterisierungen gegeben werden.

Definition 1.1 (Brownsche Bewegung)**.** Ein stetiger stochastischer Prozess $(X_t)_{0 \leq t \leq 1}$, adaptiert auf die Filtration $(\mathcal{G}_t)_{0 \leq t \leq 1}$, ist eine $(\mathcal{G}_t)-$Brownsche Bewegung mit Anfangsverteilung μ (Wahrscheinlichkeitsmaß auf $\mathbb{R}$), wenn gilt

i) $\mathcal{L}(X_0) = \mu$;

ii) für alle $s, t \in [0,1]$ mit $s \leq t$ ist $X_t - X_s$ unabhängig von $\mathcal{G}_s$, und ist normalverteilt mit Erwartungswert 0 und Varianz $t - s$.

Anmerkung 1.1 (Zur Anfangsverteilung). Gegebenenfalls werden stillschweigend noch einige Integrabilitätsbedingungen an die Verteilung μ der Zufallsvariable X_0 gestellt. Da die Definitionen der Brownschen Bewegung in der Literatur recht unterschiedlich in diesem Punkt sind, werden wir darauf nicht weiter eingehen und an den jeweiligen Stellen die Existenz einiger Momente der Startverteilung einfach implizit voraussetzen. Es geht bei der Definition der Brownschen Bewegung hauptsächlich darum, dass die Zuwächse normalverteilt und unabhängig sind.

Anmerkung 1.2 (Zur Brownschen Bewegung). Die Brownsche Bewegung ist benannt nach dem englischen Botaniker Robert Brown (lebte und wirkte im frühen 19. Jh.). Er erwähnte in einer 1828 veröffentlichten Arbeit die Zitterbewegung von Blütenpollen in einem Wasserbad. Nach eigenen Aussagen war er aber nicht der erste, dem dieses Phänomen aufgefallen ist.
Seine Veröffentlichung zog eine längere Debatte über die Ursachen der Zitterbewegung nach sich. So schloss man zum Beispiel organische Bewegungsursachen aus, nachdem auch anorganische Steinkrümel die gleiche Zitterbewegung wie ihre verwandten organischen Krümel zeigten (laut [Nel] diente auch ein Steinkrümel der Sphinx als Experimentationsobjekt). E. Nelson sagt über den Beitrag Robert Browns zur Erforschung der Zitterbewegung: “His contribution was to establish Brownian motion as an important phenomenon, to demonstrate clearly its presence in inorganic as well as organic matter, and to refute by experiment facile mechanical explanations of the phenomenon.“
Ein Durchbruch zu weiteren theoretischen Entwicklungen kam dann im Jahre 1905 durch einen Impuls von Albert Einstein. Er berechnete die Übergangswahrscheinlichkeitsdichte der Brownschen Bewegung, wie sie auch in unserer Definition implizit angegeben ist. Im Jahre 1923 definierte dann Wiener die Brownsche Bewegung auf streng mathematische Weise, die Definition hat sich seitdem im Prinzip nicht geändert. Von ihm stammte auch der erste Existenzbeweis. Erläuterungen zu diesem Beweis findet man unter anderen in [RW1] (Chapter I.6).

Seitdem wurde das mathematische Objekt Brownsche Bewegung immer mehr weiterentwickelt. Eine gute Hinführung zum Objekt Brownsche Bewegung findet man zum Beispiel in [KS] (Chapter 2), dort werden auch viele spezielle Eigenschaften entwickelt. Ein sehr schöner geschichtlicher Abriss der Brownschen Bewegung findet sich in [Nel] (Chapter 2-4).
Wie schon erwähnt, befindet sich ein Existenzbeweis der Brownschen Bewegung, der sich nach eigener Aussage eng an den ersten Beweis von Wiener hält, in [RW1] (Chapter I.6). Ein alternativer Beweis, der mehr in Richtung der Anwendung von Donsker's Theorem orientiert ist (mehr zu diesem Ergebnis kommt später), findet sich in [Bil3] (Section 8).

Anmerkung 1.3 (Zur Notation). Die von einer Brownschen Bewegung mit beliebiger Anfangsverteilung erzeugte Wahrscheinlichkeitsverteilung auf $(C(0,1),\mathcal{C})$ nennt man Wienermaß $\mathbb{W}$. Für den Erwartungswert bezüglich eines Wienermaßes schreiben wir einfach $\mathbb{E}$. Gilt $X_0 = y$ für ein $y \in \mathbb{R}$ fast sicher (im Folgenden abgekürzt mit $f.s.$), so schreiben wir $\mathbb{W}_y$ beziehungsweise $\mathbb{E}_y$. Für Erwartungswerte bezüglich anderer Wahrscheinlichkeitsmaße schreiben wir einfach E, oder speziell für $\rho \in \mathcal{P}(C(0,1))$ schreiben wir E_ρ.

Theorem 1.1 (Charakterisierung über das charakteristische Funktional)**.** *Ein stetiger stochastischer Prozess $(X_t)_{0\leq t\leq 1}$, adaptiert zur Filtration $(\mathcal{G}_t)_{0\leq t\leq 1}$, ist genau dann eine $(\mathcal{G}_t)-$Brownsche Bewegung, wenn für alle Treppenfunktionen g auf $[0,1]$ gilt*

$$E(e^{i\int_0^1 g(s)dX_s}) = e^{-\frac{1}{2}\int_0^1 g^2(s)ds}. \tag{1.1}$$

Beweis. Die Gleichung ist sinnvoll, da das stochastische Integral bezüglich einer Treppenfunktion immer definiert ist (als endliche Summe von Zufallsvariablen), und die linke Seite der Gleichung wegen der Beschränktheit offensichtlich integrierbar ist.
Sei Z eine Zufallsvariable mit $\mathcal{L}(Z) = \mathcal{N}(m,\sigma^2)$ ($\mathcal{L}(Z)$ bezeichnet die Wahrscheinlichkeitsverteilung von Z), dass heißt Z ist normalverteilt mit Erwartungswert m und Varianz σ^2. Wir wissen, dass die Verteilung von Z eineindeutig charakterisiert ist durch die charakteristische Funktion $\phi_Z(a) := E(e^{iaZ})$. Die charakteristische Funktion von Z ergibt sich über quadratische Ergänzung wie folgt:

$$\begin{aligned}
E(e^{iaZ}) &= c\int\limits_{\mathbb{R}} e^{iaz}e^{-\frac{1}{2}\frac{(z-m)^2}{\sigma^2}}dz \\
&= c\int\limits_{\mathbb{R}} e^{-\frac{1}{2\sigma^2}(z^2-2mz+m^2-2\sigma^2 iaz)}dz \\
&= c\int\limits_{\mathbb{R}} e^{-\frac{1}{2\sigma^2}(z-(m+\sigma^2 ia))^2)}e^{-\frac{1}{2\sigma^2}(m^2-(m+\sigma^2 ia)^2)}dz \\
&= e^{-\frac{1}{2\sigma^2}(m^2-m^2-2m\sigma^2 ia+\sigma^4 a)} \\
&= e^{ima-\frac{1}{2}\sigma^2 a^2}.
\end{aligned}$$

Einen alternativen Beweis dafür findet man zum Beispiel in [Bau] (Kapitel V, §22, Beispiel nach Satz 22.7). Ist nun $Z = (Z_1,\ldots,Z_N)$ ein Vektor unabhängiger normalverteilter Zufallsvariablen mit $\mathcal{L}(Z_j) = \mathcal{N}(m_j,\sigma_j^2)$, so gilt für die charakteristische Funktion des Zufallsvektors (wir verwenden hier die Schreibweisen $a = (a_1, \dots, a_N) \in \mathbb{R}^N$, $\langle a, Z\rangle := \sum_{j=1}^N a_j Z_j$)

$$\begin{aligned} E(e^{i\langle a,Z\rangle}) &= \prod_{j=1}^N E(e^{ia_jZ_j}) \\ &= \prod_{j=1}^N e^{im_ja_j - \frac{1}{2}\sigma_j^2 a_j^2} \\ &= e^{i\sum_{j=1}^N m_ja_j - \frac{1}{2}\sum_{j=1}^N \sigma_j^2 a_j^2}. \end{aligned} \tag{1.2}$$

Die charakteristische Funktion ist wiederum eineindeutig für die $N-$dimensionale Normalverteilung.

Die Aussage des Theorems spiegelt genau diese Charakterisierung wieder. Um das zu erkennen, schreiben wir die Gleichung (1.1) aus. Die Treppenfunktion g habe o.B.d.A. (ohne Beschränkung der Allgemeinheit) die Form

$$g = \sum_{j=1}^N a_j 1_{[t_{j-1}, t_j[} \quad \text{mit} \quad 0 \le t_0 \le \dots \le t_N \le 1.$$

Diese Treppenfunktion setzen wir in Gleichung (1.1) auf der linken Seite ein, und erhalten

$$E(e^{i\int_0^1 g(s)dX_s}) = E(e^{i\sum_{j=1}^N a_j(X_{t_j} - X_{t_{j-1}})}).$$

Auf der rechten Seite ergibt sich

$$e^{-\frac{1}{2}\int_0^1 g^2(s)ds} = e^{-\frac{1}{2}\sum_{j=1}^N a_j^2(t_j - t_{j-1})},$$

und insgesamt

$$E(e^{i\sum_{j=1}^N a_j(X_{t_j} - X_{t_{j-1}})}) = e^{-\frac{1}{2}\sum_{j=1}^N a_j^2(t_j - t_{j-1})}.$$

Mit der Gleichung (1.2) für die mehrdimensionale Normalverteilung erkennen wir, dass die $(X_{t_j} - X_{t_{j-1}})_{j=1,\dots,N}$ unabhängig voneinander sind, und normalverteilt mit $\mathcal{L}(X_{t_j} - X_{t_{j-1}}) = \mathcal{N}(0, (t_j - t_{j-1}))$. Das entspricht genau der Definition der Brownschen Bewegung. Aufgrund der erwähnten Eineindeutigkeit der Charakterisierung durch die charakteristische Funktion ist das Theorem bewiesen (denn damit sind die endlichdimensionalen Zylindermaße auf $(C(0,1), \mathcal{C})$ charakterisiert). □

Theorem 1.2 (Charakterisierung von Lévy)**.** *Ein stetiger stochastischer Prozess* $(X_t)_{0\le t\le 1}$*, adaptiert zur Filtration* $(\mathcal{G}_t)_{0\le t\le 1}$*, ist genau dann eine* $(\mathcal{G}_t)-$*Brownsche Bewegung, wenn er ein stetiges lokales Martingal mit quadratischer Variation* t *ist.*

Beweis. Sei X eine stetige Brownsche Bewegung auf $[0,1]$. Unter Bezug auf Anmerkung 1.1 setzen wir zusätzlich voraus, dass $X_0 \in L^2$ ist. Dann ist $X_t \in L^1$ für alle $t \in [0,1]$. Die Integrierbarkeitsbedingung aus der Definition lokaler Martingale ist daher erfüllt. Die Brownsche Bewegung ist sogar ein Martingal, denn für $s, t \in [0,1]$, $s \le t$ gilt

$$\begin{aligned}\mathbb{E}(X_t|\mathcal{G}_s) &= \mathbb{E}(X_t - X_s|\mathcal{G}_s) + \mathbb{E}(X_s|\mathcal{G}_s)\\ &= X_s,\end{aligned}$$

aufgrund der Unabhängigkeit und Zentriertheit der Zuwächse und der $\mathcal{G}_s-$Meßbarkeit von X_s. Um zu zeigen, dass die quadratische Variation von $(X_t)_{0\le t\le 1}$ gleich t ist, müssen wir beweisen, dass $(X_t^2 - t)_{0\le t\le 1}$ ein Martingal ist. Die Integrierbarkeitsbedingung ist wieder wegen der Definition erfüllt. Außerdem gilt

$$\begin{aligned}\mathbb{E}(X_t^2 - t|\mathcal{G}_s) &= \mathbb{E}((X_t - X_s)^2 + 2X_tX_s - X_s^2|\mathcal{G}_s) - t\\ &= \mathbb{E}((X_t - X_s)^2) + \mathbb{E}(2(X_t - X_s)X_s|\mathcal{G}_s) + \mathbb{E}(2X_s^2 - X_s^2|\mathcal{G}_s) - t\\ &= t - s + \mathbb{E}(2(X_t - X_s)X_s) + X_s^2 - t\\ &= X_s^2 - s.\end{aligned}$$

Somit ist $(X_t^2 - t)_{0\le t\le 1}$ ein Martingal und t die quadratische Variation der Brownschen Bewegung. Bleibt übrig zu zeigen, dass das Kriterium des Theorems hinreichend ist. Wähle $u \in \mathbb{R}$ fest und setze $f(y,t) = \exp(iuy + \frac{u^2}{2}t)$. Sei $(Z_t)_{0\le t\le 1}$ der stetige stochastische Prozess definiert durch $Z_t = f(X_t, t)$. Da $f \in C^2$ (bezeichnet die zweimal stetig differenzierbaren Funktionen), können wir die Itô-Formel anwenden. Dazu benötigen wir die in Frage kommenden quadratischen Variationen (bezeichnet mit $([.,.]_t)_{0\le t\le 1}$). Wir verwenden zur Berechnung der quadratischen Variation Rechenregeln die man zum Beispiel in [RY] findet (speziell Chapter IV, Proposition 1.18). Da t ein Prozess mit endlicher Variation ist (wir missbrauchen die Notation t etwas, und verwenden sie abwechselnd für den deterministischen Prozess auf [0,1] der die identische Abbildung bezeichnet, und für einen Wert aus [0,1]), haben wir automatisch $[X,t] \equiv 0$ und $[t,t] \equiv 0$. Nach Voraussetzung gilt außerdem $[X,X]_t = t$. Die verwendeten partiellen Ableitungen der Funktion f sind

$$\partial_1 f(y,t) = iuf(y,t), \qquad \partial_2 f(y,t) = \frac{u^2}{2} f(y,t), \qquad \partial_1^2 f(y,t) = -u^2 f(y,t).$$

Wir bezeichnen hier, wie auch allgemein, die partielle Ableitung einer beliebigen Funktion f nach ihrer $j-$ten Variablen mit $\partial_j f$. Setzen wir alles in die Itô-Formel ein, ergibt sich

$$\begin{aligned}Z_t &= 1 + iu\int_0^t Z_s dX_s + \tfrac{u^2}{2}\int_0^t Z_s ds - \tfrac{u^2}{2}\int_0^t Z_s ds\\ &= 1 + iu\int_0^t Z_s dX_s.\end{aligned}$$

Dies ist die exponentiale Integralgleichung (siehe [RW2], Chapter IV.19). Da X ein lokales Martingal ist, muss auch Z (als Itô-Integral eines lokalen Martingals) eines sein. Außerdem ist Z durch $\exp(\frac{u^2}{2})$ beschränkt, deshalb sogar ein Martingal. Da $f(y,t) \neq 0$ für alle $y \in \mathbb{R}$, $0 \le t \le 1$ (und $u \in \mathbb{R}$), ist der stetige stochastische Prozess $((f(X_t,t))^{-1})_{0\le t\le 1}$ definiert und wir erhalten wegen der Martingaleigenschaft des Prozesses Z

$$\begin{aligned} E(e^{(iu(X_t - X_s))}|\mathcal{G}_s) &= E(e^{iuX_t + \frac{u^2}{2}t} e^{-iuX_s - \frac{u^2}{2}s}|\mathcal{G}_s)e^{-\frac{u^2}{2}(t-s)} \\ &= E(f(X_t, t)(f(X_s, s))^{-1}|\mathcal{G}_s)e^{-\frac{u^2}{2}(t-s)} \\ &= (f(X_s, s))^{-1}E(f(X_t, t)|\mathcal{G}_s)e^{-\frac{u^2}{2}(t-s)} \\ &= (f(X_s, s))^{-1}f(X_s, s)e^{-\frac{u^2}{2}(t-s)} \\ &= e^{-\frac{u^2}{2}(t-s)}, \qquad f.s.\ \forall u \in \mathbb{R}. \end{aligned}$$

Mit Theorem 1.1 folgt, dass $X_t - X_s$ unabhängig von $\mathcal{G}_s$ und normalverteilt mit Erwartungswert 0 und Varianz $(t - s)$ ist. Es handelt sich also um eine Brownsche Bewegung. □

2 Die Dualitätsformel des Wienermaßes

Die Dualitätsformel ist eine Art partielle Integrationsformel unter dem Wienermaß. Sie heißt Dualitätsformel, da sie einen Differentiationsoperator in Relation zu einem Integrationsoperator setzt (auf einem Raum von Funktionalen auf $C(0,1)$). Diese werden wir im Folgenden erstmal definieren.

Zunächst benötigen wir einen sinnvollen Differentiationsbegriff auf dem Pfadraum $C(0,1)$. Wir verwenden dazu das allgemeine in der Einleitung gegebene Set-up $X : (\Omega, \mathcal{F}, P) \to (C(0,1), \mathcal{C}, P_X)$. Sei $W^{1,2}$ die Teilmenge der Funktionale aus $L^2(P_X)$, die L^2-differenzierbar in folgendem Sinne sind: Es existiert ein $(D_t F(x))_{0 \le t \le 1,\, x \in C(0,1)}$ in $L^2(dt|_{[0,1]} \cdot P_X)$ so dass für alle $g \in L^2(dt|_{[0,1]})$ die folgende Gleichung in $L^2(P_X)$ gilt:

$$D_g F(x) = \lim_{\varepsilon \to 0} \varepsilon^{-1}(F(x + \varepsilon \int_0^{\cdot} g(t)dt) - F(x)) = \int_0^1 g(t) D_t F(x) dt.$$

Es handelt sich dabei um eine Gâteaux-Ableitung (eine allgemeine Definition der Gâteaux-Differenzierbarkeit ist in [Wer], Kapitel III.5, Definition III.5.1 zu finden). Außerdem versehen wir $W^{1,2}$ mit der Sobolev-Norm

$$\|F\|_{1,2} = \|F\|_2 + \|D_{\cdot} F\|_2 \,, \tag{2.1}$$

wobei

$$\|F\|_2 = E(F^2(X))^{\frac{1}{2}},$$

$$\|D_{\cdot} F\|_2 = E(\int_0^1 (D_s F)^2(X) ds)^{\frac{1}{2}}.$$

Wir bezeichnen mit $\mathcal{W}$ die zylindrischen Funktionale der Form $F(x) = \phi(x_{s_1}, \ldots, x_{s_N})$ wobei $0 \le s_1 \le \cdots \le s_N \le 1$ und ϕ eine C_b^∞ Funktion ist.

Anmerkung 2.1. Es gilt $\mathcal{W} \subset W^{1,2}$ wobei die Definition von $\mathcal{W}$ nicht vom Wahrscheinlichkeitsmaß P_X auf $(C(0,1), \mathcal{C})$ abhängt (die Ableitung existiert immer und ist quadratintegrabel). Speziell ergibt sich für $F \in \mathcal{W}$ die Gleichung

$$D_t F(x) = \sum_{i=1}^N \partial_i \phi(x_{s_1}, \ldots, x_{s_N}) 1_{[0,s_i[}(t). \tag{2.2}$$

Beweis. Mit Hilfe der Taylorentwicklung (siehe Theorem 5.1 im Anhang) kommt man darauf, dass für $F(x) = \phi(x_{s_1}, \ldots, x_{s_N})$ und $g \in L^2(dt|_{[0,1]})$ folgende Gleichungen für alle $x \in C(0,1)$ gelten (wie bereits im Beweis zu Theorem 1.1 bezeichnet $\langle .,. \rangle$ das Skalarprodukt im $\mathbb{R}^N$):

$$\begin{aligned}
D_g F(x) &= \lim_{\varepsilon\to 0} \varepsilon^{-1} \left(F(x + \varepsilon \int_0^{\cdot} g(t)dt) - F(x) \right) \\
&= \lim_{\varepsilon\to 0} \varepsilon^{-1} \left(\phi(x_{s_1} + \varepsilon \int_0^{s_1} g(t)dt, \ldots, x_{s_N} + \varepsilon \int_0^{s_N} g(t)dt) - \phi(x_{s_1}, \ldots, x_{s_N}) \right) \\
&= \lim_{\varepsilon\to 0} \varepsilon^{-1} \left(\varepsilon \left\langle \nabla\phi, (\int_0^{s_1} g(t)dt, \ldots, \int_0^{s_N} g(t)dt) \right\rangle + O(\varepsilon^2) \right) \\
&= \sum_{j=1}^{N} \partial_j \phi(x_{s_1}, \ldots, x_{s_N}) \int_0^{s_j} g(t)dt \\
&= \int_0^1 g(t) \sum_{j=1}^{N} \partial_j \phi(x_{s_1}, \ldots, x_{s_N}) 1_{[0,s_j[}(t)\ dt \\
&= \int_0^1 g(t)\ D_t F(x)\ dt.
\end{aligned}$$

Da die Gleichungen für alle $x \in C(0,1)$ gelten, gelten sie auch fast sicher. Im dritten Schritt wenden wir dann einfach die beschränkte Konvergenz an um die Gleichheit in L^2 zu haben. Wir verwenden dabei die Abschätzung der Formel (5.1):

$$\begin{aligned}
& E\left(\left(\varepsilon^{-1} \left(\phi(x_{s_1} + \varepsilon \int_0^{s_1} g(t)dt, \ldots, x_{s_N} + \varepsilon \int_0^{s_N} g(t)dt) - \phi(x_{s_1}, \ldots, x_{s_N}) \right.\right.\right. \\
& \qquad \left.\left.\left. - \sum_{j=1}^{N} \partial_j \phi(x_{s_1}, \ldots, x_{s_N}) \varepsilon \int_0^{s_j} g(t)dt \right) \right)^2 \right) \\
& \le E(\varepsilon^{-2}(K \sum_{j,k=1}^{N} \varepsilon^2 \int_0^{s_j} g(t)dt \int_0^{s_k} g(t)dt)^2) \\
& \le K'\varepsilon^2 \longrightarrow 0 \qquad \text{für } \varepsilon \to 0.
\end{aligned}$$

Eine solche Konstante K' existiert, da g nach Voraussetzung quadratintegrabel ist. Die Behauptung, dass $\mathcal{W} \subset W^{1,2}$ folgt dann aufgrund der Definition von $W^{1,2}$. $\square$

Anmerkung 2.2. Es ist egal, ob wir die Funktionale $F \in \mathcal{W}$ als Funktion $F(x) = \phi(x_{s_1}, \ldots, x_{s_N})$ oder als Funktion der Zuwächse $F(x) = \tilde{\phi}(x_{s_1}, x_{s_2} - x_{s_1}, \ldots, x_{s_N} - x_{s_{N-1}})$ schreiben (sie haben immer beide Darstellungen). Für die Gateaux-Ableitung gilt bei der von den Zuwächsen abhängigen Schreibweise dann

$$D_t F(x) = \sum_{j=1}^{N} \partial_j \tilde{\phi}(x_{s_1}, x_{s_2} - x_{s_1}, \ldots, x_{s_N} - x_{s_{N-1}}) 1_{[s_{j-1}, s_j[}(t)$$

Beweis. Sei $F \in \mathcal{W}$ gegeben. Definiere $h : \mathbb{R}^N \mapsto \mathbb{R}^N$ als lineare Transformation $(y_1, \ldots, y_N) \mapsto (y_1, y_2 - y_1, \ldots, y_N - y_{N-1})$. Dann ist $h^{-1} : \mathbb{R}^N \mapsto \mathbb{R}^N$ die lineare Transformation $(y_1, \ldots, y_N) \mapsto (y_1, y_2 + y_1, \ldots, y_N + \cdots + y_1)$. Wir erkennen hier leicht, dass h ein unendlich oft differenzierbarer Diffeomorphismus ist. Ist $F(x) = \phi(x_{s_1}, \ldots, x_{s_N})$, so ist äquivalent dazu $F(x) = \tilde{\phi}(x_{s_1}, x_{s_2} - x_{s_1}, \ldots, x_{s_N} - x_{s_{N-1}})$ mit $\tilde{\phi} := \phi \circ h^{-1}$ (und $\phi \circ h^{-1}$ ist in C_b^∞). Ein Funktional $F \in \mathcal{W}$ besitzt also immer beide Darstellungen.

Mit Hilfe der Kettenregel (angewandt auf $\phi = \tilde{\phi} \circ h$) erhalten wir

$$\partial_1 \phi = \partial_1 \tilde{\phi} - \partial_2 \tilde{\phi}, \quad \partial_2 \phi = \partial_2 \tilde{\phi} - \partial_3 \tilde{\phi}\ , \quad \dots\ , \quad \partial_N \phi = \partial_N \tilde{\phi}.$$

Für die Ableitung $D_t F$ lässt sich deshalb folgern (formal setzen wir in der folgenden Rechnung $\partial_{N+1} \tilde{\phi} \equiv 0$), dass

$$\begin{aligned} D_t F(x) &= \sum_{j=1}^{N} \partial_j \phi(x_{s_1}, \dots, x_{s_N}) 1_{[0,s_j[}(t) \\ &= \sum_{j=1}^{N} (\partial_j \tilde{\phi}(x_{s_1}, \dots, x_{s_N} - x_{s_{N-1}}) - \partial_{j+1} \tilde{\phi}(x_{s_1}, \dots, x_{s_N} - x_{s_{N-1}})) 1_{[0,s_j[}(t) \\ &= \sum_{j=1}^{N} \partial_j \tilde{\phi}(x_{s_1}, x_{s_2} - x_{s_1}, \dots, x_{s_N} - x_{s_{N-1}}) 1_{[s_{j-1},s_j[}(t). \end{aligned}$$

□

Anmerkung 2.3. Bezüglich der Norm $\|.\|_{1,2}$ ist $\mathcal{W}$ dicht in $W^{1,2}$. Dies ist ein Ergebnis aus der Maßtheorie (Anwendung des Satzes monotoner Klassen und der L^2-Konvergenz) und wird hier nicht bewiesen.

2.1 Wienermaß erfüllt Dualitätsformel

Die folgende Gleichung (2.3) wurde bereits von Hausmann (siehe [Hau]) erkannt und später von Bismut (siehe [Bis], Theorem 2.1) bewiesen. Dabei werden allerdings allgemeinere Fälle als der hier Angegebene behandelt. Es geht hauptsächlich darum, dass die Funktion g auch ein stochastischer Prozess sein kann (so ist g in [Bis] ein adaptierter Prozess im $\mathbb{R}^n$).
Für unsere auf spezielle Anwendungen orientierten Zwecke genügt aber das Folgende (in [RZ1] erwähnte) mit relativ grundlegenden Mitteln zu beweisende Theorem. Wir verwenden dabei die abkürzende Schreibweise $\delta(g) = \int_0^1 g(s) dX_s$, falls das stochastische Integral denn definiert ist. Hinreichend für die Existenz wäre zum Beispiel, dass X ein Semimartingal ist (und die Brownsche Bewegung ist ein Semimartingal, siehe Theorem 1.2) und für g fast sicher $\int_0^1 g^2(s) d\left\langle X\right\rangle_s \ \leq +\infty$ gilt (für die Brownsche Bewegung bedeutet das $g \in L^2(dt|_{[0,1]})$).

Theorem 2.1. *Unter einem Wienermaß gilt für jedes Funktional $F \in W^{1,2}$ und alle $g \in L^2(dt|_{[0,1]})$ die Dualitätsformel*

$$\mathbb{E}(F\delta(g)) = \mathbb{E}(D_g F). \tag{2.3}$$

Anmerkung 2.4 (Zur Anfangsbedingung der Wienermaße). In den folgenden Berechnungen gehen wir aus Notationsgründen anfänglich über zu dem Wienermaß mit deterministischer Anfangsbedingung $X_0 = y\ f.s.$ mit $y \in \mathbb{R}$. Um die Anfangsbedingungen zu verallgemeinern (d.h. $\mathcal{L}(X_0) = \mu$), verwenden wir eine Zerlegung von $\mathbb{E}$. Für ein allgemeines Pfadfunktional F gilt

$$\mathbb{E}(F(X)) = \int_{\mathbb{R}} \mathbb{E}_y(F(X)) \mu(dy).$$

Das diese Gleichung einen Sinn macht, kann man in [KS], Chapter 2.5.1 nachlesen (zumindest vom Aspekt der Meßbarkeit her). Da diese Mischung über die Anfangsbedingung als Integral linear ist, ergibt sich automatisch der allgemeine Fall der Dualitätsformel vermöge

$$\int_A \mathbb{E}_y(D_g F)\mu(dy) = \int_A \mathbb{E}_y(F\delta(g))\mu(dy).$$

wobei A eine beliebige Borelmenge in $\mathbb{R}$ ist.

Beweis von Theorem 2.1. Wir beweisen die Behauptung zuerst für Funktionale F aus $\mathcal{W}$ und Treppenfunktionen g. Die Behauptung folgt dann allgemein für $W^{1,2}$ beziehungsweise $L^2(dt|_{[0,1]})$ durch L^2-Grenzwertbildungen.
Der Beweis für $F \in \mathcal{W}$, wobei wir uns hier an die Schreibweise $F(x) = \phi(x_{s_1}, \dots, x_{s_N} - x_{s_{N-1}})$ mit $0 \leq s_1 \leq \dots \leq s_N \leq 1$ halten, folgt im Prinzip daraus, dass die lineare Hülle von Funktionen der Art (formal ist hier $y_0 = 0$)

$$\phi_\Lambda \in C_b^\infty : \ \phi_\Lambda := e^{i\sum_{j=1}^N \lambda_j(y_j - y_{j-1})} \text{ mit } \Lambda = (\lambda_1, \dots \lambda_N) \in \mathbb{R}^N \tag{2.4}$$

dicht in C_b^∞ liegt. Aufgrund der Linearität der Dualitätsformel auf beiden Seiten bezüglich g können wir den Beweis speziell auf die Treppenfunktion $g(s) = 1_{[0,t[}(s)$, mit $t \in [0,1]$ fest, beschränken:

$$\begin{aligned}
\mathbb{E}(D_{(\alpha_1 g_1 + \alpha_2 g_2)}F) &= \mathbb{E}(\int_0^1 (\alpha_1 g_1(s) + \alpha_2 g_2(s))D_s F(X)ds) \\
&= \mathbb{E}(\int_0^1 \alpha_1 g_1(s) D_s F(X)ds + \int_0^1 \alpha_2 g_2(s) D_s F(X)ds) \\
&= \alpha_1 \mathbb{E}(\int_0^1 g_1(s) D_s F(X)ds) + \alpha_2 \mathbb{E}(\int_0^1 g_2(s) D_s F(X)ds) \\
&= \alpha_1 \mathbb{E}(D_{g_1}F) + \alpha_2 \mathbb{E}(D_{g_2}F)
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}(F(X)\delta(\alpha_1 g_1 + \alpha_2 g_2)) &= \mathbb{E}(F(X)\int_0^1 (\alpha_1 g_1(s) + \alpha_2 g_2(s))dX_s) \\
&= \alpha_1 \mathbb{E}(F(X)\int_0^1 g_1(s)dX_s) + \alpha_2 \mathbb{E}(F(X)\int_0^1 g_2(s)dX_s) \\
&= \alpha_1 \mathbb{E}(F(X)\delta(g_1)) + \alpha_2 \mathbb{E}(F(X)\delta(g_2)).
\end{aligned}$$

Wir gehen bis auf Widerruf davon aus, dass $t \leq s_N$, denn wenn $t > s_N$ haben wir

$$\begin{aligned}
\mathbb{E}_y(F(X)\delta(g)) &= \mathbb{E}_y(\phi(X_{s_1}, \dots, X_{s_N} - X_{s_{N-1}}) \ (X_t - y)) \\
&= \mathbb{E}_y(\phi(X_{s_1}, \dots, X_{s_N} - X_{s_{N-1}}) \ (X_{s_N} - y)) \\
&\quad + \mathbb{E}_y(\phi(X_{s_1}, \dots, X_{s_N} - X_{s_{N-1}}) \ (X_t - X_{s_N})) \\
&= \mathbb{E}_y(\phi(X_{s_1}, \dots, X_{s_N} - X_{s_{N-1}}) \ (X_{s_N} - y)) + 0,
\end{aligned} \tag{2.5}$$

aufgrund der Unabhängigkeit und Zentriertheit der Zuwächse.
Wir beginnen mit dem Fall $N = 1$. Sei ein $F \in \mathcal{W}$ gegeben mit $F(x) = \phi(x_s)$. Auf der linken Seite

der Dualitätsformel steht dann

$$\mathbb{E}_y(F(X)\delta(g)) = \mathbb{E}_y(\phi(X_s)(X_t - y)),$$

und auf der rechten Seite

$$\mathbb{E}_y(D_g F(X)) = t\,\mathbb{E}_y(\phi'(X_s)).$$

Aus der Definition 1.1 kennen wir die endlichdimensionalen Dichten (auf Zylindermengen) der Brownschen Bewegung, und können den Erwartungswert als das Integral

$$\mathbb{E}_y(\phi(X_s)(X_t - y)) = C\int_{\mathbb{R}^2} \phi(y_2)(y_1 - y)e^{-\frac{1}{2}(\frac{(y_1-y)^2}{t}+\frac{(y_2-y_1)^2}{s-t})}dy_1dy_2$$

ausdrücken. C ist der Normierungsfaktor der Dichte und wird keine weitere Rolle spielen. Wir substituieren y_1 durch $y_2 - y_1$, und es ergibt sich

$$-C\int_{\mathbb{R}^2} \phi(y_2)(y_2 - y_1 - y)e^{-\frac{1}{2}(\frac{(y_2-y_1-y)^2}{t}+\frac{(y_1)^2}{s-t})}dy_1dy_2.$$

Jetzt führen wir eine partielle Integration nach y_2 durch:

$$\left[(-t)\phi(y_2)e^{-\frac{1}{2}(\frac{(y_2-y_1-y)^2}{t}+\frac{(y_1)^2}{s-t})}\right]_{y_2=-\infty}^{\infty} + (-t)C\int_{\mathbb{R}^2} \phi'(y_2)e^{-\frac{1}{2}(\frac{(y_2-y_1-y)^2}{t}+\frac{(y_1)^2}{s-t})}dy_1dy_2.$$

Der erste Term fällt aufgrund der Beschränktheit von ϕ weg. Im zweiten Term substituieren wir wieder y_1 durch $y_2 - y_1$ und erhalten

$$C\int_{\mathbb{R}^2} \phi'(y_2)te^{-\frac{1}{2}(\frac{(y_1-y)^2}{t}+\frac{(y_2-y_1)^2}{s-t})}dy_1dy_2.$$

Das ist aber genau unsere gesuchte Größe, denn

$$\mathbb{E}_y(\phi'(X_s)t) = C\int_{\mathbb{R}^2} \phi'(y_2)te^{-\frac{1}{2}(\frac{(y_1-y)^2}{t}+\frac{(y_2-y_1)^2}{s-t})}dy_1dy_2.$$

Zusammen mit der Gleichung (2.5) haben wir für den eindimensionalen Fall

$$\mathbb{E}_y(\phi(X_s)X_t) = (s \wedge t)\mathbb{E}_y(\phi'(X_s)) \tag{2.6}$$

bewiesen. Das diese Formel insbesondere für $\phi_\lambda(y) = e^{i\lambda y}$ gilt, werden wir gleich benötigen.
Sei jetzt $N \in \mathbb{N}$ beliebig. Wir wollen zeigen, dass die Dualitätsformel für $F(x) = \phi_\Lambda(x_{s_1}, \ldots, x_{s_N} - x_{s_{N-1}})$ gilt (ϕ_Λ wie in (2.4) definiert). Wir berechnen (diesmal ohne die Konvention $t \leq s_N$, formal

setzen wir $s_0 = 0$)

$$
\begin{aligned}
& \mathbb{E}_y(\phi_\Lambda(X_{s_1}, \dots, X_{s_N} - X_{s_{N-1}})(X_t - y)) \\
\overset{(*)}{=} \;& \mathbb{E}_y(\phi_{\lambda_1}(X_{s_1}) \cdots \phi_{\lambda_N}(X_{s_N} - X_{s_{N-1}}) \sum_{j=1}^{N} (X_{s_j \wedge t} - X_{s_{j-1} \wedge t})) \\
= \;& \sum_{j=1}^{N} \mathbb{E}_y(\phi_{\lambda_1}(X_{s_1}) \cdots \phi_{\lambda_N}(X_{s_N} - X_{s_{N-1}})(X_{s_j \wedge t} - X_{s_{j-1} \wedge t})) \\
\overset{(*)}{=} \;& \mathbb{E}_y(\phi_{\lambda_1}(X_{s_1})(X_{s_1 \wedge t} - y))\mathbb{E}_y(\phi_{\lambda_2}(X_{s_2} - X_{s_1}) \cdots \phi_{\lambda_N}(X_{s_N} - X_{s_{N-1}})) + \dots \\
& + \mathbb{E}_y(\phi_{\lambda_N}(X_{s_N} - X_{s_{N-1}})(X_{s_N \wedge t} - X_{s_{N-1} \wedge t}))\mathbb{E}_y(\phi_{\lambda_1}(X_{s_1}) \cdots \phi_{\lambda_{N-1}}(X_{s_{N-1}} - X_{s_{N-2}})) \\
\overset{(2.6)}{=} \;& \mathbb{E}_y(\phi'_{\lambda_1}(X_{s_1})(s_1 \wedge t))\mathbb{E}_y(\phi_{\lambda_2}(X_{s_2} - X_{s_1}) \cdots \phi_{\lambda_N}(X_{s_N} - X_{s_{N-1}})) + \dots \\
& + \mathbb{E}_y(\phi'_{\lambda_N}(X_{s_N} - X_{s_{N-1}})(s_N \wedge t - s_{N-1} \wedge t))\mathbb{E}_y(\phi_{\lambda_1}(X_{s_1}) \cdots \phi_{\lambda_{N-1}}(X_{s_{N-1}} - X_{s_{N-2}})) \\
\overset{(*)}{=} \;& \sum_{j=1}^{N} \mathbb{E}_y(\partial_j \phi_\Lambda(X_{s_1}, \dots, X_{s_N} - X_{s_{N-1}})(s_j \wedge t - s_{j-1} \wedge t)) \\
= \;& \mathbb{E}_y(\int_0^1 1_{[0,t[}(s) \sum_{j=1}^{N} \partial_j \phi_\Lambda(X_{s_1}, \dots, X_{s_N} - X_{s_{N-1}}) 1_{[s_{j-1}, s_j[}(s) ds).
\end{aligned}
$$

An den mit $(*)$ gekennzeichneten Stelle gebrauchen wir die Unabhängigkeit der Zuwächse, an der mit (2.6) gekennzeichneten verwenden wir zusätzlich zu Gleichung (2.6) noch die Markoveigenschaft der Brownschen Bewegung. Die Markoveigenschaft wird angewendet um

$$\mathbb{E}_y(\phi_{\lambda_j}(X_{s_j} - X_{s_{j-1}})(X_{s_j \wedge t} - X_{s_{j-1} \wedge t})) = \mathbb{E}_y(\phi'_{\lambda_j}(X_{s_j} - X_{s_{j-1}})(s_j \wedge t - s_{j-1} \wedge t))$$

zu erhalten, denn

$$
\begin{aligned}
& \mathbb{E}_y(\phi_{\lambda_j}(X_{s_j} - X_{s_{j-1}})(X_{s_j \wedge t} - X_{s_{j-1} \wedge t})) \\
& = \mathbb{E}_y(\mathbb{E}_y(\phi_{\lambda_j}(X_{s_j} - X_{s_{j-1}})(X_{s_j \wedge t} - X_{s_{j-1} \wedge t}) | \mathcal{F}_{s_{j-1}})) \\
& = \mathbb{E}_y(\mathbb{E}_{X_{s_{j-1}}}(\phi_{\lambda_j}(\hat{X}_{s_j - s_{j-1}})(\hat{X}_{s_j \wedge t - s_{j-1} \wedge t} - \hat{X}_0)),
\end{aligned}
$$

wobei $\hat{X}$ eine von X unabhängige Brownsche Bewegung ist. Auf den Term innerhalb des Erwartungswertes lässt sich die eindimensionale Formel anwenden, und dann müssen wir uns nur noch klarmachen, dass

$$(s_j - s_{j-1}) \wedge (s_j \wedge t - s_{j-1} \wedge t) = (s_j \wedge t - s_{j-1} \wedge t).$$

Nach Vergleich mit Anmerkung 2.2 sehen wir, dass die Dualitätsformel für Funktionen des Typs ϕ_Λ und Treppenfunktionen g gilt. Das werden wir jetzt erweitern auf Funktionale aus $W^{1,2}$ und Funktionen aus $L^2(dt|_{[0,1]})$, aber erstmal für Funktionale aus $\mathcal{W}$. Wir behandeln das Ganze in diesem Beweis noch mit etwas mehr Bedacht und Detailfreude als in späteren Beweisen in denen die gleichen Techniken aber wieder auftauchen werden.

Wir gehen jetzt mit Anmerkung 2.4 über zu dem Erwartungswert bezüglich eines Wienermaßes mit allgemeinen Anfangsbedingungen. Das, was wir bisher bewiesen haben, ist also die Dualitätsformel (2.3) für Funktionale $F \in \mathcal{W}$ mit $F(x) = \phi_\Lambda(x_{s_1}, \dots, x_{s_N} - x_{s_{N-1}})$ und Treppenfunktionen g.

Die Linearkombinationen von Funktionen des Typs ϕ_Λ liegen dicht in jedem C_b^∞. Deshalb gibt es für jedes $F \in \mathcal{W}$ mit $F(x) = \phi(x_{s_1}, \dots, x_{s_N} - x_{s_{N-1}})$, $\phi \in C_b^\infty$, eine Folge $(F_n)_{n\geq 1} \subset \mathcal{W}$ mit $F_n(x) = \phi_n(x_{s_1}, \dots, x_{s_N} - x_{s_{N-1}})$, $\phi_n \in C_b^\infty$ die gleichmäßig gegen ϕ konvergiert (d.h. bezüglich der Supremumsnorm auf $\mathbb{R}^N$), und für die bereits die Dualitätsformel gilt. Die gleichmäßige Konvergenz impliziert die fast sichere Konvergenz

$$F_n(X) \to F(X) \; f.s.,$$

denn es existiert ein $n_0 \in \mathbb{N}$, so dass

$$\|\phi_n - \phi\|_\infty < \varepsilon, \quad \forall n \geq n_0,$$

und dies impliziert

$$\mathbb{W}(\sup_{n\geq n_0} |\phi_n(X_{s_1}, \dots, X_{s_N} - X_{s_{N-1}}) - \phi(X_{s_1}, \dots, X_{s_N} - X_{s_{N-1}})| > \varepsilon) = 0,$$

und damit natürlich

$$\lim_{m\to\infty} \mathbb{W}(\sup_{n\geq m} |\phi_n(X_{s_1}, \dots, X_{s_N} - X_{s_{N-1}}) - \phi(X_{s_1}, \dots, X_{s_N} - X_{s_{N-1}})| > \varepsilon) = 0.$$

Da es sich bei $(\phi_n)_{n\geq 1}$ um eine konvergente Folge in C_b^∞ handelt, ist $\sup_n ||\phi_n||_\infty < K$ für eine Konstante $K \in \mathbb{N}$. Ebenso ist $\sup_n ||\partial_j \phi_n||_\infty < K_2$ für alle $j \in \{1, \dots, N\}$. Nach Voraussetzung ist $KX_t \in L^1$ und trivialerweise auch $K_2 \in L^1$. Deshalb können wir auf beiden Seiten der Dualitätsformel den Satz von der majorisierten Konvergenz anwenden (siehe [RW1], II.21) und erhalten

$$\mathbb{E}(F\delta(g)) = \lim_{n\to\infty} \mathbb{E}(F_n\delta(g)) = \lim_{n\to\infty} \mathbb{E}(D_g F_n) = \mathbb{E}(D_g F).$$

Die Dualitätsformel ist damit für $F \in \mathcal{W}$ und g Treppenfunktion bewiesen.

Die Erweiterung auf $F \in W^{1,2}$ und $g \in L^2(dt|_{[0,1]})$ ergibt sich wie folgt:

Sei $g \in L^2(dt|_{[0,1]})$ beliebig und $F \in \mathcal{W}$. Da die Treppenfunktionen dicht in L^2 liegen, können wir eine Folge $(g_n)_{n\geq 1}$ von Treppenfunktionen finden, die in L^2 gegen g konvergiert.Wir verwenden jetzt, dass die L^2-Norm auf dem Skalarprodukt $\int_0^1 f(s)g(s)ds$ für $f, g \in L^2(dt|_{[0,1]})$ basiert. Jede stark konvergente Folge ist auch schwach konvergent in L^2 (Schwarzsche Ungleichung, siehe zum Beispiel [Els], Kap. VI, §5, Folgerung 5.8). Wir erinnern uns daran, dass für jedes $x \in C(0,1)$ nach Voraussetzung $(D_t F(x))_{0\leq t\leq 1} \in L^2(dt|_{[0,1]})$ ist, und schließen wegen schwacher Folgenkonvergenz auf

$$\int_0^1 g(t) D_t F(x) dt = \lim_{n\to\infty} \int_0^1 g_n(t) D_t F(x) dt.$$

Für jedes $x \in C(0,1)$ ist die Folge $(\int_0^1 g_n(t) D_t F(x) dt)_{n\geq 1}$ konvergent und damit beschränkt. Da $F \in \mathcal{W}$ gibt es eine Konstante $K \in \mathbb{R}^+$ so dass $\sup_{x\in C(0,1)} |D_t F(x)| \leq \sup_{x\in C(0,1)} |D_1 F(x)| \leq K$ gleichmäßig für $t \in [0,1]$ (man erinnere sich für die erste Ungleichung an die Gleichung (2.2)).

Damit kommen wir für alle $x \in C(0,1)$ auf

$$|\int_0^1 g(t)D_tF(x)dt| \leq K \sup_{n\geq 1}\int_0^1 |g_n(t)|dt.$$

Das Supremum auf der rechten Seite der Gleichung existiert, da die Folge $(\int_0^1 |g_n(t)|dt)_{n\geq 1}$ gegen den Grenzwert $\int_0^1 |g(t)|dt$ konvergiert, denn $(\mathrm{sign}(g(t)))_{0\leq t\leq 1} \in L^2(dt|_{[0,1]})$.
Etwas das Punktweise auf $C(0,1)$ gilt, gilt auch fast sicher. Damit können die majorisierte Konvergenz anwenden (Vgl. [Bil3], Chap. II, Theorem 21.2) und erhalten

$$\mathbb{E}(\int_0^1 g(t)D_tF(X)dt) = \lim_{n\to\infty}\mathbb{E}(\int_0^1 g_n(t)D_tF(X)dt).$$

Auf der anderen Seite der Dualitätsformel ergibt sich aus der Definition des Wienerintegrals (siehe dazu [Nel], Chapter 7, Theorem 7.1)

$$\int_0^1 g(s)dX_s = \lim_{n\to\infty}\int_0^1 g_n(s)dX_s, \quad \text{in } L^2(\mathbb{W}).$$

Auch die Norm auf $L^2(\mathbb{W})$ basiert auf einem Skalarprodukt, dass definiert ist durch $\mathbb{E}_y(ZY)$ für $Z, Y \in L^2(\mathbb{W})$. Aus der schwachen Folgenkonvergenz ergibt sich dann analog zu obiger Beweisführung

$$\mathbb{E}(F(X)\int_0^1 g(s)dX_s) = \lim_{n\to\infty}\mathbb{E}(F(X)\int_0^1 g_n(s)dX_s),$$

da $F \in \mathcal{W} \subset L^2(\mathbb{W})$. Das heißt, es gilt die Dualitätsformel für $F \in \mathcal{W}$ und $g \in L^2(dt|_{[0,1]})$.
Nach Anmerkung 2.3 können wir zu beliebigem $F \in W^{1,2}$ eine Folge $(F_n)_{n\geq 1} \subset \mathcal{W}$ finden die bezüglich der Norm $\|.\|_{1,2}$ (wurde in (2.1) definiert) gegen F konvergiert. Nun ist g ebenso wie die D_tF_n in $L^2(dt|_{[0,1]}\cdot\mathbb{W})$ $(g(X,s) = g(s))$, und wegen schwacher Folgenkonvergenz (diesmal bezüglich des Skalarprodukts $\mathbb{E}_y(\int_0^1 u(s,X)v(s,X)ds)$ für $u, v \in L^2(dt|_{[0,1]}\cdot\mathbb{W})$) ergibt sich deshalb

$$\mathbb{E}(D_gF) = \lim_{n\to\infty}\mathbb{E}(D_gF_n).$$

Für die Konvergenz von $(F_n)_{n\geq 1}$ können wir genauso argumentieren (diesmal wieder für $L^2(\mathbb{W})$), da nach Definition des Wienerintegrals $\delta(g) \in L^2(\mathbb{W})$, und erhalten

$$\mathbb{E}(F\delta(g)) = \lim_{n\to\infty}\mathbb{E}(F_n\delta(g)).$$

Die Dualitätsformel (2.3) unter einem Wienermaß ist damit bewiesen. □

Anmerkung 2.5 (Dualitätstheorem für Pfadfunktionale aus $\mathcal{W}$). Entsprechend der Anmerkung 2.2 gibt es jetzt zwei äquivalente Versionen des Dualitätstheorems 2.3 für Funktionale aus $F \in \mathcal{W}$. Da beide später nochmals auftauchen, wollen wir sie hier notieren. Für $F(x) = \phi(x_{s_1}, \dots, x_{s_N})$ haben wir

$$\mathbb{E}(\phi(X_{s_1}, \dots, X_{s_N})\delta(g)) = \mathbb{E}(\sum_{j=1}^N \partial_j\phi(X_{s_1}, \dots, X_{s_N})\int_0^{s_j} g(t)dt). \tag{2.7}$$

Für $F(x) = \tilde{\phi}(x_{s_1}, x_{s_2} - x_{s_1}, \dots, x_{s_N} - x_{s_{N-1}})$ haben wir

$$\mathbb{E}(\tilde{\phi}(X_{s_1}, \dots, X_{s_N} - X_{s_{N-1}})\delta(g)) = \mathbb{E}(\sum_{j=1}^{N} \partial_j \tilde{\phi}(X_{s_1}, \dots, X_{s_N} - X_{s_{N-1}}) \int_{s_{j-1}}^{s_j} g(t)dt). \tag{2.8}$$

2.2 Dualitätsformel charakterisiert Wienermaß

Nachdem wir im ersten Kapitel verschiedene Charakterisierungen der Brownschen Bewegung gegeben haben, und im ersten Paragraphen dieses Kapitels gezeigt haben, dass die Brownsche Bewegung die Dualitätsformel 2.3 erfüllt, wollen wir nun zeigen, dass sich die Brownsche Bewegung auch durch die Dualitätsformel charakterisieren läßt (d.h. die Dualitätsformel zu erfüllen, ist für ein Wienermaß hinreichend und notwendig). Das folgende Theorem geht auf Roelly und Zessin zurück (siehe z.B. [RZ1]). Der Beweis ist eine etwas detailliertere Version des dort ausgeführten Beweises.

Theorem 2.2. *Sei $\rho \in \mathcal{P}(C(0,1))$ ein Wahrscheinlichkeitsmaß für das der kanonische Koordinatenprozess integrierbar ist, d.h. $E_\rho(|X_t|) < +\infty$ für $t \in [0,1]$. Wenn die Gleichung*

$$E_\rho(F\delta(g)) = E_\rho(D_g F).$$

für jedes Funktional F aus $\mathcal{W}$ und jede Treppenfunktion g auf $[0,1]$ gilt, dann ist ρ ein Wienermaß.

Beweis. Zuerst überprüfen wir, ob die obige Gleichung überhaupt Sinn macht unter den gegebenen Voraussetzungen. g ist eine Treppenfunktion und hat daher o.B.d.A. die Form

$$g = a_1 1_{[t_0,t_1[} + a_2 1_{[t_1,t_2[} + \cdots + a_N 1_{[t_{N-1},t_N[}, \qquad a_i \in \mathbb{R},\ 0 \le t_0 \le \cdots \le t_N \le 1.$$

Das stochastische Integral $\delta(g)$ wird dann zur einfachen Summe

$$\delta(g) = a_1(X_{t_1} - X_{t_0}) + \cdots + a_N(X_{t_N} - X_{t_{N-1}}) = \sum_{j=1}^{N} a_j(X_{t_j} - X_{t_{j-1}}).$$

Auf der linken Seite ist das Produkt $F\delta(g)$ ρ-integrabel, da F beschränkt, und $\delta(g)$ nach Voraussetzung ρ-integrabel ist. Auf der rechten Seite ist $D_g F$ beschränkt durch das Produkt aus $\sup_{t\in[0,1]} |g(t)|$ und $\sup_{x\in C(0,1)} \|D_. F(x)\|_2$ und somit ρ-integrabel.
Sei nun $\hat{\rho}$ das folgende, zu ρ assoziierte Funktional:

$$\hat{\rho}(g) = E_\rho(\exp(i\delta(g))), \quad g \text{ Treppenfunktion auf } [0,1].$$

Da $\exp(i\delta(g))$ fast sicher durch 1 beschränkt ist, ist das Funktional für alle Treppenfunktionen wohldefiniert. Für alle $\lambda \in \mathbb{R}$ gilt dann

$$\frac{d}{d\lambda}\hat{\rho}(\lambda g) = E_\rho(i\delta(g)\exp(i\lambda\delta(g))).$$

Man darf hier unter dem Integralzeichen differenzieren, denn

i) $\exp(i\lambda \sum_{j=1}^{N} a_j(x_{t_j} - x_{t_{j-1}})) \in L^1(\rho)$ für $\forall\lambda \in \mathbb{R}$;

ii) für alle $x \in C(0,1)$ existiert die partielle Ableitung

$$\partial_\lambda \exp(i\lambda \sum_{j=1}^{N} a_j(x_{t_j} - x_{t_{j-1}})) = i \sum_{j=1}^{N} a_j(x_{t_j} - x_{t_{j-1}}) \exp(i\lambda \sum_{j=1}^{N} a_j(x_{t_j} - x_{t_{j-1}}));$$

iii) allgemein gilt für $x \in C(0,1)$ und $\lambda \in \mathbb{R}$ die Ungleichung

$$\left| i \sum_{j=1}^{N} a_j(x_{t_j} - x_{t_{j-1}}) \exp(i\lambda \sum_{j=1}^{N} a_j(x_{t_j} - x_{t_{j-1}})) \right| \leq \left| \sum_{j=1}^{N} a_j(x_{t_j} - x_{t_{j-1}}) \right|$$

und nach Voraussetzung ist $\delta(g) \in L^1(\rho)$;

(vergleiche hierzu Theorem 5.3 im Anhang).
Wie man leicht sieht, ist $F_\lambda(x) := i \exp(i\lambda\delta(g))$ ein Funktional aus $\mathcal{W}$. Wir können also die Gleichung aus der Voraussetzung des Theorems auf F_λ anwenden und erhalten

$$\begin{aligned} \tfrac{d}{d\lambda}\hat{\rho}(\lambda g) &= E_\rho(F_\lambda \delta(g)) \\ &= E_\rho(D_g F_\lambda) \\ &= E_\rho(\textstyle\int_0^1 g(s) D_s F_\lambda ds) \\ &\overset{(*)}{=} E_\rho(\textstyle\int_0^1 i\lambda g(s)^2 F_\lambda ds) \\ &= i\lambda \textstyle\int_0^1 g^2(s) ds E_\rho(F_\lambda) \\ &= -\lambda \textstyle\int_0^1 g^2(s) ds \hat{\rho}(\lambda g), \qquad \hat{\rho}(0) = 1. \end{aligned} \tag{2.9}$$

Der vorletzte Schritt funktioniert mit dem Satz von Fubini (vertauschen der Integrale) und der Linearität des Erwartungswerts. Außerdem steht $(*)$ für

$$\begin{aligned} D_t F_\lambda &= \sum_{j=1}^{N} \partial_j F_\lambda(x_{t_1}, \ldots, x_{t_N}) 1_{[0,t_j[}(t) \\ &= i \left(\sum_{j=1}^{N-1} i\lambda(a_j - a_{j+1}) \exp(i\lambda\delta(g)) 1_{[0,t_j[}(t) \right) + i\lambda a_N \exp(i\lambda\delta(g)) 1_{[0,t_N[}(t) \\ &= i\lambda F_\lambda \left(\sum_{j=1}^{N-1} (a_j - a_{j+1}) 1_{[0,t_j[}(t) + a_N 1_{[0,t_N[}(t) \right) \\ &= i\lambda F_\lambda \sum_{j=1}^{N} a_j (1_{[0,t_j[}(t) - 1_{[0,t_{j-1}[}(t)) \\ &= i\lambda g(t)\ F_\lambda. \end{aligned}$$

In (2.9) erkennen wir eine gewöhnliche Differentialgleichung mit Lösung $\hat{\rho}(\lambda g) = \exp -\lambda \int_0^1 \frac{1}{2} g^2(s) ds$. Wir setzen $\lambda = 1$, womit $\hat{\rho}(g) = E_\rho(\exp(i\delta(g))) = \exp - \int_0^1 \frac{1}{2} g^2(s) ds$ für alle Treppenfunktionen g gilt. Vergleicht man mit Theorem 1.1 erkennt man, dass ρ ein Wienermaß ist ($X_0 \in L^1$ gilt nach Voraussetzung). $\square$

Anmerkung 2.6 (Zur Intervallgröße). Alles bisher bewiesene kann auf das Intervall $[0, T]$ (für $T \in \mathbb{R}^+$ fest), beziehungsweise auf $C(0, T)$ kanonisch erweitert werden. Da die entsprechenden Theoreme dann für alle $C(0, T)$ mit $T \in \mathbb{R}^+$ gelten, gelten sie auch für $C(\mathbb{R}^+)$ (der Pfadraum stetiger Funktionen auf $\mathbb{R}^+$). Die Beweise und Theoreme sind dann in lokalisierten Versionen zu geben.
Als Beispiel dafür findet man in der Arbeit [RZ1] (Théorème 2) eine Version des Theorems 2.2 für das Wienermaß auf dem Raum $C(\mathbb{R}^+)$.

3 Die diskrete Dualitätsformel der Irrfahrt

Dieses Kapitel soll etwas Ähnliches wie die Dualitätsformel der Brownschen Bewegung für die Irrfahrt aufdecken. Dazu geben wir zuerst die Definition der Standardversion einer symmetrischen Irrfahrt.

Definition 3.1 (symmetrische Irrfahrt)**.** Eine symmetrische Irrfahrt ist eine Folge von Zufallsvariablen $(S_n)_{n\geq 0}$, die definiert ist vermöge

$$S_n = \sum_{j=1}^{n} Y_j \ , \qquad (S_0 = 0),$$

wobei $(Y_j)_{j\geq 1}$ voneinander unabhängige, identisch verteilte (abgekürzt mit iid.) Zufallsvariablen sind mit $E(Y_1) = 0,\ Var(Y_1) = 1$.

Im Verlauf der weiteren mathematischen Ausformulierung der diskreten Dualitätsformel ergab es sich allerdings, dass die Standardversion der symmetrischen Irrfahrt schon ein Zuviel an Voraussetzungen mitbringt. Das Ziel war nämlich eigentlich, eine Dualitätsformel zu finden, die eine symmetrische Irrfahrt charakterisiert. Damit die noch folgende diskrete Dualitätsformel zu einer Charakterisierung wird, muß man die Bedingungen im Vergleich zu einer symmetrischen Irrfahrt noch etwas abschwächen.
Bevor wir einen geeigneteren Begriff einer symmetrischen Irrfahrt angeben, führen wir eine Notation ein. Sei dazu $(Y_j)_{j\geq 1}$ eine Folge beliebiger quadratintegrabler Zufallsvariablen. Dann ist $\sigma(Y_{k_1}, \ldots, Y_{k_N})$ für $k_1, \ldots, k_N \in \mathbb{N}$ die von $Y_{k_1}, \ldots, Y_{k_N}$ erzeugte $\sigma-$Algebra. Für den bedingten Erwartungswert schreiben wir dann

$$E(Y_l | Y_{k_1}, \ldots, Y_{k_N}) := E(Y_l | \sigma(Y_{k_1}, \ldots, Y_{k_N}))$$

(siehe zum Beispiel [Kle] (Kapitel 8.2) für eine Definition des bedingten Erwartungswerts). Bei der Definition der folgenden, verallgemeinerte symmetrische Irrfahrt genannten Folge von Zufallsvariablen, spielt der bedingte Erwartungswert eine Schlüsselrolle.

Definition 3.2 (verallgemeinerte symmetrische Irrfahrt)**.** Sei $(S_n)_{n\geq 1}$ eine Folge quadratintegrabler zentrierter Zufallsvariablen, $S_0 := 0$. Wir definieren die Folge der Differenzen der S_n vermöge $Y_j := S_j - S_{j-1}$. Wir nennen $(S_n)_{n\geq 0}$ genau dann eine verallgemeinerte symmetrische Irrfahrt, wenn für alle $N \in \mathbb{N}$ und alle $l \in \mathbb{N}$ gilt

$$E(Y_l | Y_{j_1}, \ldots, Y_{j_N},\ j_\alpha \neq l\ \forall \alpha) = E(Y_l) = 0, \tag{3.1}$$

und für alle $l, k, m \in \mathbb{N}$ mit $k \leq l \leq k+m$ gilt

$$\begin{aligned} &E(Y_l | Y_k + \cdots + Y_{k+m},\ Y_{j_1}, \ldots, Y_{j_N},\ j_\alpha \notin \{k, \ldots, k+m\}\ \forall \alpha) \\ &= E(Y_k | Y_k + \cdots + Y_{k+m},\ Y_{j_1}, \ldots, Y_{j_N},\ j_\alpha \notin \{k, \ldots, k+m\}\ \forall \alpha). \end{aligned} \tag{3.2}$$

Anmerkung 3.1 (Varianz der verallgemeinerten symmetrischen Irrfahrt). Die zweiten Momente der oben definierten Folge $(Y_j)_{j\geq 1}$ einer verallgemeinerten symmetrischen Irrfahrt sind gleich (wir setzen deshalb definitorisch $Var(Y_1) = 1$, wie bei der symmetrischen Irrfahrt)

Beweis. Da die S_n quadratintegrierbar sind, sind es auch die Y_j, denn

$$\begin{aligned} E(Y_j^2) &= E((S_j - S_{j-1})^2) \\ &= E(S_j^2) + 2E(S_j S_{j-1}) + E(S_{j-1}^2) \\ &\leq E(S_j^2) + 2E(S_j^2)^{\frac{1}{2}} E(S_{j-1}^2)^{\frac{1}{2}} + E(S_{j-1}^2) < \infty \end{aligned}$$

(jeder L^2 ist ein Vektorraum). Nach Gleichung (3.2) aus der Definition der verallgemeinerten symmetrischen Irrfahrt gilt für $l_1, l_2 \in \mathbb{N}, l_1 < l_2$ und eine entsprechend quadratintegrable Funktion ϕ

$$E(Y_{l_1} \phi(Y_{l_1} + Y_{l_1+1} + \cdots + Y_{l_2})) = E(Y_{l_2} \phi(Y_{l_1} + Y_{l_1+1} + \cdots + Y_{l_2})).$$

Nach Voraussetzung ist $\phi(Y_{l_1} + \cdots + Y_{l_2}) = Y_{l_1} + \cdots + Y_{l_2}$ quadratintegrabel. Einsetzen dieser Funktion liefert

$$E(Y_{l_1}^2) + E(Y_{l_1} Y_{l_1+1}) + \cdots + E(Y_{l_1} Y_{l_2}) = E(Y_{l_2} Y_{l_1}) + E(Y_{l_2} Y_{l_1+1}) + \cdots + E(Y_{l_2}^2).$$

Ein Spezialfall von Gleichung (3.1) ist $E(Y_l | Y_k) = 0$ für $k \neq l$. Damit bleibt aus obiger Gleichung nur

$$E(Y_{l_1}^2) = E(Y_{l_2}^2)$$

übrig, unsere Behauptung (da $l_1, l_2 \in \mathbb{N}$ beliebig). □

Wir werden in dem folgenden Theorem versuchen, den noch ungewohnten Begriff einer verallgemeinerten symmetrischen Irrfahrt zwischen die aus der Literatur bekannten Begriffe der symmetrischen Irrfahrt und des zeitdiskreten Martingals einzuordnen. Wir benötigen dazu die geeignete Definition eines Martingals, auf die wir uns dann im folgenden Theorem beziehen werden.

Definition 3.3 (quadratintegrierbares Martingal)**.** Sei $(S_n)_{n \geq 1}$ eine Folge quadratintegrierbarer zentrierter Zufallsvariablen, $S_0 := 0$. Die Folge $(S_n)_{n \geq 0}$ ist genau dann ein Martingal bezüglich der Filtration $(\mathcal{F}_n)_{n \geq 0}$, wenn

$$E(S_n | \mathcal{F}_{n-1}) = S_{n-1} \quad \text{fast sicher } \forall n \in \mathbb{N}. \tag{3.3}$$

Äquivalent dazu wäre es, von der Folge der Differenzen $(Y_j)_{j \geq 1}$ mit $Y_j := S_j - S_{j-1}$ zu verlangen, dass

$$E(Y_n | \mathcal{F}_{n-1}) = 0 \quad \text{fast sicher } \forall n \in \mathbb{N}. \tag{3.4}$$

Man nennt die Folge $(Y_j)_{j \geq 1}$ dann eine Folge von Martingaldifferenzen. Nimmt man als Filtration die kanonische Filtration, kann man hier auch zum Vergleich mit (3.1) fordern, dass

$$E(Y_l | Y_j,\ j < l) = E(Y_l) = 0. \tag{3.5}$$

Theorem 3.1. *Es gelten die Implikationen*

$$
\begin{gathered}
\textit{symmetrische Irrfahrt} \\
\Downarrow \\
\textit{verallgemeinerte symmetrische Irrfahrt} \\
\Downarrow \\
\textit{Martingal},
\end{gathered}
\tag{3.6}
$$

und die Implikationen sind streng.

Beweis. Wir zeigen zuerst, dass jede symmetrische Irrfahrt eine verallgemeinerte symmetrische Irrfahrt ist. Sei $(S_n)_{n\geq 0}$ eine symmetrische Irrfahrt wie in Definition 3.1. Quadratintegrabilität wurde in beiden Definitionen vorausgesetzt. Außerdem haben wir wegen Unabhängigkeit und Zentriertheit bereits, dass für $l \in \mathbb{N}$ gilt

$$E(Y_l|Y_j,\ j \neq l) = E(Y_l) = 0,$$

die Bedingung (3.1) ist damit bereits erfüllt. Seien $l, k, m \in \mathbb{N}$ mit $l \in \{k, \dots, k+m\}$, außerdem seien $j_1, \dots j_N$ natürliche Zahlen mit $j_1, \dots, j_N \notin \{k, \dots, k+m\}$. Wir zeigen jetzt, dass

$$E(Y_l|Y_k + \cdots + Y_{k+m}, Y_{j_1}, \dots, Y_{j_N}) = E(Y_k|Y_k + \cdots + Y_{k+m}, Y_{j_1}, \dots, Y_{j_N})$$

gilt, womit dann auch die Forderung (3.2) erfüllt wäre. Dies ist äquivalent dazu, dass für alle $\phi \in C_b$

$$E(Y_l\phi(Y_k + \cdots + Y_{k+m}, Y_{j_1}, \dots, Y_{j_N})) = E(Y_k\phi(Y_k + \cdots + Y_{k+m}, Y_{j_1}, \dots, Y_{j_N}))$$

erfüllt ist. Dann ist die Behauptung aber klar, wie man sich am einfachsten verdeutlichen kann, wenn man die Erwartungswerte explizit als Integrale notiert. Sei dazu $\mathcal{L}(Y_1) = \mu$, dann

$$
\begin{aligned}
&\textstyle\int y_l\phi(y_k + \cdots + y_{k+m}, y_{j_1}, \dots, y_{j_N})\mu(dy_k)\cdots\mu(dy_{k+m})\mu(dy_{j_1})\cdots\mu(dy_{j_N}) \\
&= \textstyle\int y_k\phi(y_k + \cdots + y_{k+m}, y_{j_1}, \dots, y_{j_N})\mu(dy_k)\cdots\mu(dy_{k+m})\mu(dy_{j_1})\cdots\mu(dy_{j_N}),
\end{aligned}
\tag{3.7}
$$

da das Problem sich hier auf den Namen des Index reduziert.
Jetzt beweisen wir, dass jede verallgemeinerte symmetrische Irrfahrt ein Martingal ist. Um die Behauptung zu zeigen, müssen wir nur die Bedingung (3.1) an die verallgemeinerte symmetrische Irrfahrt anwenden. Wir verwenden die spezielle Variante der Gleichung (3.1)

$$E(Y_k|Y_1, \dots, Y_{k-1}) = 0, \tag{3.8}$$

und definieren dazu die kanonische Filtration vermöge

$$\mathcal{F}_k := \sigma(Y_1, \dots, Y_k).$$

Dann wird (3.8) zu

$$E(Y_k|\mathcal{F}_{k-1}) = 0,$$

anders ausgedrückt haben wir

$$E(S_k - S_{k-1}|\mathcal{F}_{k-1}) = 0.$$

Wir wissen aber, aus der Definition der Filtration $(\mathcal{F}_n)_{n\geq 1}$, dass S_{k-1} eine $\mathcal{F}_{k-1}-$meßbare Zufallsvariable ist, also

$$E(S_k - S_{k-1}|\mathcal{F}_{k-1}) = E(S_k|\mathcal{F}_{k-1}) - S_{k-1} = 0,$$

und daraus folgt die fast sicher geltende Gleichung

$$E(S_k|\mathcal{F}_{k-1}) = S_{k-1}.$$

Wir haben damit gezeigt (die Integrabilitätsbedingung ist nach Definition vorausgesetzt), dass $(S_n)_{n\geq 0}$ ein $(\mathcal{F}_n)_{n\geq 0}-$Martingal ist ($\mathcal{F}_0$ ist die triviale $\sigma-$Algebra).
Jetzt wollen wir zeigen, dass die gezeigten Implikationen auch streng sind. Beginnen wir dazu mit einer heuristischen Betrachtung. Das jede verallgemeinerte symmetrische Irrfahrt eine symmetrische Irrfahrt ist, gilt natürlich im Allgemeinen nicht. Um uns das zu verdeutlichen vergleichen wir $(\phi, \tilde{\phi} \in L^2)$

$$\begin{aligned} \text{symmetrische Irrfahrt} \quad &\leftrightarrow \text{verallgemeinerte symmetrische Irrfahrt} \\ E(\phi(Y_1)\tilde{\phi}(Y_2)) = E(\phi(Y_1))E(\tilde{\phi}(Y_2)) \quad &\leftrightarrow E(Y_1\phi(Y_2)) = E(Y_1)E(\phi(Y_2)), \\ E(\phi(Y_1)) = E(\phi(Y_2)) \quad &\leftrightarrow E(\phi(E(Y_1|Y_1+Y_2))) = E(\phi(E(Y_2|Y_1+Y_2))). \end{aligned}$$

Wir setzen aber bei der verallgemeinerten symmetrischen Irrfahrt viel mehr voraus als Unkorreliertheit, denn Unkorreliertheit würde hier nur bedeuten dass $E(Y_1Y_2) = E(Y_1)E(Y_2)$, während wir ja voraussetzen, dass $E(Y_1\phi(Y_2)) = E(Y_1)E(\phi(Y_2))$. Das nicht jedes Martingal eine verallgemeinerte symmetrische Irrfahrt ist liegt unter anderem an der Zusatzbedingung (3.2) der Definition. Um die heuristischen Argumente zu Beweisen bedarf es natürlich einiger Gegenbeispiele, die im Folgenden erbracht werden.
Wir beginnen mit dem Gegenbeispiel dafür, dass nicht jede verallgemeinerte symmetrische Irrfahrt eine symmetrische Irrfahrt ist. Sei $(X_n)_{n\geq 1}$ eine Folge von iid. integrierbaren und zentrierten Zufallsvariablen, sei $(Z_n)_{n\geq 1}$ eine davon unabhängige Sequenz von iid. Zufallsvariablen mit $\mathcal{L}(Z_1) = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$ (Bernoulliverteilung). Weiterhin sei W eine integrierbare Zufallsvariable, unabhängig von $(X_n)_{n\geq 1}$ und $(Z_n)_{n\geq 1}$. Wir definieren $Y_k := X_k + (-1)^{Z_k}W$, und $S_n := \sum_{k=1}^n Y_k$ $(S_0 = 0)$. Dann ist $(S_n)_{n\geq 0}$ ein verallgemeinerte symmetrische Irrfahrt.
Wir überprüfen zuerst die Eigenschaft (3.1). Es genügt zu zeigen, dass

$$E(Y_l|Y_{j_1}, \dots, Y_{j_N}) = 0$$

für $l \notin \{j_1, \ldots, j_N\}$. Sei $\phi \in C_b(\mathbb{R}^N)$ eine stetige und beschränkte Funktion. Dann gilt

$$
\begin{aligned}
&E(Y_l \phi(Y_{j_1}, \ldots, Y_{j_N})) \\
&= E((X_l + (-1)^{Z_l} W)\phi(X_{j_1} + (-1)^{Z_{j_1}} W, \ldots, X_{j_N} + (-1)^{Z_{j_N}} W)) \\
&= E((-1)^{Z_l} W \phi(X_{j_1} + (-1)^{Z_{j_1}} W, \ldots, X_{j_N} + (-1)^{Z_{j_N}} W)) \\
&= -\tfrac{1}{2} E(W \phi(X_{j_1} + (-1)^{Z_{j_1}} W, \ldots, X_{j_N} + (-1)^{Z_{j_N}} W)) \\
&\quad + \tfrac{1}{2} E(W \phi(X_{j_1} + (-1)^{Z_{j_1}} W, \ldots, X_{j_N} + (-1)^{Z_{j_N}} W)) \\
&= 0.
\end{aligned}
$$

Im zweiten Schritt wurde die Unabhängigkeit von X_l, im dritten die von Z_l verwendet. Wir müssen jetzt nur noch die Eigenschaft (3.2) überprüfen. Das Argument ähnelt dem aus (3.7). Seien $l, k, m, j_1, \ldots j_N \in \mathbb{N}$ so gewählt, dass $l \in \{k, \ldots, k+m\}$, $j_1, j_2 \ldots, j_N \notin \{k, \ldots, k+m\}$. Um (3.2) zu zeigen, genügt es wieder

$$
E(Y_l | Y_k + \cdots + Y_{k+m}, Y_{j_1}, \ldots, Y_{j_N}) = E(Y_k | Y_k + \cdots + Y_{k+m}, Y_{j_1}, \ldots, Y_{j_N}) \tag{3.9}
$$

zu beweisen. Sei dazu $\phi \in C_b$. Wir müssen zeigen, dass die Wahl des Index $l \in \{k, \ldots, k+m\}$ keinen Einfluß auf das Ergebnis von

$$
E(Y_l \phi(Y_k + \cdots + Y_{k+m}, Y_{j_1}, \ldots, Y_{j_N}))
$$

hat. Zur Vereinfachung der Notation nehmen wir o.B.d.A. an, dass $k = 1$, $j_1 = m+2$, $j_N = m+p$ gilt. Dann haben wir

$$
\begin{aligned}
&E(Y_l \phi(Y_1 + \cdots + Y_{m+1}, Y_{m+2}, \ldots, Y_{m+p})) \\
&= \int (x_l + (-1)^{z_l}) w \cdot \\
&\qquad \cdot \phi(x_1 + (-1)^{z_1} w + \cdots + x_{m+1} + (-1)^{z_{m+1}} w, x_{m+2} + (-1)^{z_{m+2}} w, \ldots, x_{m+p} + (-1)^{z_{m+p}} w) \cdot \\
&\qquad \cdot P_X(dx_1) \cdots P_X(dx_{m+p}) P_Z(dz_1) \cdots P_Z(dz_{m+p}) P_W(dw).
\end{aligned}
$$

Dieses Integral ist dank der Unabhängigkeit der Zuvallsvariablen nicht vom speziellen $l \in \{1, \ldots, m+1\}$ abhängig. Es handelt sich hier um einen recht allgemeinen Zugang, verallgemeinerte symmetrische Irrfahrten zu konstruieren, die keine symmetrischen Irrfahrten im strengeren Sinne sind. Allgemeiner können wir den Zugang ungefähr so ausdrücken: Gegeben seien von einander unabhängige Folgen $(X_j^{(1)})_{j\geq 1}, \ldots, (X_j^{(p)})_{j\geq 1}$ von iid. Zufallsvariablen, und eine davon unabhängige Zufallsvariable W. Dann sei $Y_j := f(X_j^{(1)}, \ldots, X_j^{(p)}, W)$, wobei f so gewählt wird, dass die Y_j quadratintegrabel und bedingt unter W zentriert sind. Dann ist die Folge $(Y_j)_{j\geq 1}$ eine verallgemeinerte symmetrische Irrfahrt, denn bedingt unter W sind die Y_j eine iid. Sequenz. Genauer gesagt haben wir

$$E(Y_l\phi(Y_{j_1},\dots,Y_{j_N})) = E(E(Y_l\phi(Y_{j_1},\dots,Y_{j_N})|W)) = 0$$

und damit die Eigenschaft (3.1). Um (3.2) zu erhalten bedingen wir ebenfalls einfach auf W und nutzen dann die Unabhängigkeit und identische Verteilung unter dieser Bedingung aus:

$$E(Y_l\phi(Y_k+\cdots+Y_{k+m},Y_{j_1},\dots,Y_{j_N}))$$

$$= E(E(Y_l\phi(Y_k+\cdots+Y_{k+m},Y_{j_1},\dots,Y_{j_N})|W)$$

$$= E(E(Y_k\phi(Y_k+\cdots+Y_{k+m},Y_{j_1},\dots,Y_{j_N})|W).$$

Das nicht jede verallgemeinerte symmetrische Irrfahrt dieser Art eine Irrfahrt sein muss, beweisen wir jetzt an einem Zahlenbeispiel. Dazu kehren wir zurück zu der anfänglich konstruierten verallgemeinerten symmetrischen Irrfahrt $(S_n)_{n\geq 0}$. Wir haben bereits festgelegt, dass P_Z eine Bernoulliverteilung ist. Sei weiterhin $P_X = \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1$ und $P_W = \frac{1}{2}\delta_1 + \frac{1}{2}\delta_2$. Währen Y_1 und Y_2 unabhängig, müsste die Gleichung $P(Y_1 = 3, Y_2 = 3) = P(Y_1 = 3)P(Y_2 = 3)$ gelten. Dem ist aber nicht so, denn

$$P(Y_1 = 3, Y_2 = 3)$$

$$= P(X_1 = 1, Z_1 = 0, X_2 = 1, Z_2 = 0, W = 2)$$

$$= P(X_1 = 1)P(Z_1 = 0)P(X_2 = 1)P(Z_2 = 0)P(W = 2) = \tfrac{1}{32},$$

wohingegen die rechte Seite

$$P(Y_1 = 3)P(Y_2 = 3)$$

$$= P(X_1 = 1, Z_1 = 0, W = 2)P(X_2 = 1, Z_2 = 0, W = 2)$$

$$= P(X_1 = 1)P(Z_1 = 0)P(W = 2)P(X_2 = 1)P(Z_2 = 0)P(W = 2) = \tfrac{1}{64}$$

ergibt. Die Folge $(Y_j)_{j\geq 1}$ ist also nicht stochastisch unabhängig, und es kann sich bei $(S_n)_{n\geq 0}$ nicht um eine symmetrische Irrfahrt handeln.

Jetzt liefern wir ein Beispiel dafür, dass nicht jedes Martingal eine verallgemeinerte symmetrische Irrfahrt ist. Sei $(X_n)_{n\geq 1}$ eine Folge von iid. integrierbaren und zentrierten Zufallsvariablen, definiere $S_n := \sum_{k=1}^{n} \frac{1}{k} X_k$, $S_0 = 0$. Dann ist die Folge $(S_n)_{n\geq 0}$ ein Martingal bezüglich der kanonischen Filtration:

$$E(S_n|\mathcal{F}_{n-1}) = E(\frac{1}{n}X_n + S_{n-1}|\mathcal{F}_{n-1}) = S_{n-1}.$$

Die Folge erfüllt aber im Allgemeinen nicht die Bedingung (3.2) an eine verallgemeinerte symmetrische Irrfahrt. Das machen wir uns an einem Zahlenbeispiel deutlich. Wir legen dazu die Verteilung von X_1 fest:

$$\mathcal{L}(X_1) := \frac{1}{2}\delta_{-1} + \frac{1}{2}\delta_1.$$

Wir zeigen jetzt, dass unter diesen Bedingungen

$$E(Y_1|Y_1 + Y_2) \neq E(Y_2|Y_1 + Y_2) \tag{3.10}$$

gilt, wobei $Y_k := S_k - S_{k-1} = \frac{1}{k}X_k$. Die Zufallsvariable $Y_1 + Y_2$ nimmt fast sicher einen der Werte $\{-1\frac{1}{2},\ -\frac{1}{2},\ \frac{1}{2},\ 1\frac{1}{2}\}$ an. Wir betrachten das Ereignis $Y_1 + Y_2 = 1\frac{1}{2}$. Dann gilt für die bedingte Erwartung von Y_1

$$\begin{aligned}
E(Y_1|Y_1 + Y_2 = 1\tfrac{1}{2}) &= E(Y_1 1_{Y_1=-1} + Y_1 1_{Y_1=1}|Y_1 + Y_2 = 1\tfrac{1}{2}) \\
&= -1P(Y_1 = -1|Y_1 + Y_2 = 1\tfrac{1}{2}) + 1P(Y_1 = 1|Y_1 + Y_2 = 1\tfrac{1}{2}) \\
&= 1,
\end{aligned}$$

denn wir haben nach Konstruktion die bedingten Wahrscheinlichkeitenmaße $P_{Y_1|Y_1+Y_2=1.5} = \delta_1$ und $P_{Y_2|Y_1+Y_2=1.5} = \delta_{\frac{1}{2}}$, was offensichtlich der Forderung (3.2) widerspricht. Für die bedingte Erwartung von Y_2 gilt dann genauso

$$\begin{aligned}
E(Y_2|Y_1 + Y_2 = 1\tfrac{1}{2}) &= E(Y_1 1_{Y_1=-\frac{1}{2}} + Y_1 1_{Y_1=\frac{1}{2}}|Y_1 + Y_2 = 1\tfrac{1}{2}) \\
&= -\tfrac{1}{2}P(Y_2 = -\tfrac{1}{2}|Y_1 + Y_2 = 1\tfrac{1}{2}) + \tfrac{1}{2}P(Y_2 = \tfrac{1}{2}|Y_1 + Y_2 = 1\tfrac{1}{2}) \\
&= \tfrac{1}{2}.
\end{aligned}$$

Damit ist die Ungleichung (3.10) bewiesen. □

Anmerkung 3.2 (Martingaldifferenzen). Wir wollen die Eigenschaft (3.5) von Martingaldifferenzen mit der Eigenschaft (3.1) der verallgemeinerten symmetrischen Irrfahrt vergleichen. Das die Forderung (3.1) an eine Folge $(Y_j)_{j\geq 1}$ von Zufallsvariablen die Eigenschaft (3.5) impliziert ist klar, bzw. kann aus dem obigen Beweis entnommen werden. Die umgekehrte Implikation ist aber im Allgemeinen falsch. Um das zu zeigen, konstruieren wir ein Gegenbeispiel. Sei dazu $(X_k)_{k\geq 1}$ eine Folge von iid. integrierbaren Zufallsvariablen mit $\mathcal{L}(X_1) = \frac{2}{3}\delta_1 + \frac{1}{3}\delta_{-2}$ und $(Z_k)_{k\geq 1}$ eine Folge von iid. Zufallsvariablen mit $\mathcal{L}(Z_1) = \frac{1}{2}\delta_0 + \frac{1}{2}\delta_1$. Die durch $Y_1 := X_1$, $Y_k := X_k + (-1)^{Z_k}Y_{k-1}$ definierte Folge von Zufallsvariablen erfüllt dann die Eigenschaft (3.5), aber nicht die Eigenschaft (3.1), denn

$$E(Y_n|Y_1, \ldots, Y_{n-1}) = -\frac{1}{2}Y_{n-1} + \frac{1}{2}Y_{n-1} = 0,$$

wohingegen

$$E(Y_1|Y_2 = -3) = -2P(X_1 = -2|X_1 = -2, X_2 = -2, Z_2 = 1) = -2,$$

denn $P_{Y_1|Y_2=-3} = \delta_{-2}$.

Anmerkung 3.3 (Beispiel einer verallgemeinerten symmetrischen Irrfahrt). Die Frage, die sich bei dem neu eingeführten Begriff der verallgemeinerten symmetrischen Irrfahrt stellt, ist natürlich die, ob er überhaupt Relevanz in der Anwendung hat, dass heißt, ob es gewichtige Beispiele gibt, die eine verallgemeinerte symmetrische Irrfahrt sind. Ein solches Beispiel wollen wir jetzt geben.

Es handelt sich um eine spezielle Art von Gibbsmaßen auf $\mathbb{R}^{\mathbb{N}}$, die auf Potentialfunktionen beruhen die supereven Potential genannt werden (frei übersetzt: supergerade Potentialfunktion). Die entsprechende Begriffe aus der Theorie der Gibbsmaße setzen wir hier voraus (siehe [Geo]) Eine Potential Φ heißt supergerade, wenn

$$\Phi_A(y_j, j \in A) = \Phi_A(\theta_j y_j, j \in A), \ \theta_j \in \{-1, 1\}, \ A \subset \mathbb{N} \text{ mit } |A| < +\infty$$

gilt. Solche Gibbs-Felder erfüllen die Forderung (3.1) an die verallgemeinerte symmetrische Irrfahrt (für einen Beweis, siehe Lemma 1 aus [NP]). Damit solche Gibbs-Felder die Eigenschaft (3.2) ebenfalls erfüllen, muss noch eine zusätzliche Symmetriebedingung an die Potentiale gestellt werden. Das sieht man, wenn man die entsprechende Stelle (3.7) im Beweis des Theorems 3.1 betrachtet. Die dort geforderte Vertauschbarkeit der Indizes (die letztendlich eine Indifferenz gegenüber speziellen Substitutionen ist) läßt sich entsprechend auf das Potential übersetzen. Sei dazu $A \subset \mathbb{N}$ eine endliche Teilmenge der natürlichen Zahlen, τ bezeichne eine beliebige Permutation der Elemente von A. Dann fordern wir die Symmetrie

$$\Phi_A(y_j, j \in A) = \Phi_A(y_{\tau(j)}, j \in A).$$

Unter dieser Symmetriebedingung ist, wie man leicht mit der Betrachtung des Arguments für (3.7) nachvollziehen kann, das Gibbs-Feld eine verallgemeinerte symmetrische Irrfahrt. Ein Beispiel für solch ein Potential wäre

$$\Phi_A(y_j, j \in A) = \exp\left(-\sup_{j \in A} |y_j| \cdot |A|^{\gamma}\right)$$

mit $\gamma \in \mathbb{R}$ (vergleiche dazu [NP] und [Nah]).

Soviel zur ersten Einordnung des Begriffs einer verallgemeinerten symmetrischen Irrfahrt. Warum wir zwei verschiedene Typen von Irrfahrten definieren und warum wir die Eigenschaft (3.1) statt der Eigenschaft (3.5) von Martingaldifferenzen verwenden wird, wie bereits angekündigt, in den nächsten Abschnitten noch etwas klarer werden.

3.1 Verallgemeinerte symmetrische Irrfahrt erfüllt diskrete Dualitätsformel

Für die Folge $(S_n)_{n \geq 0}$ einer verallgemeinerten symmetrischen Irrfahrt lässt sich eine diskrete Dualitätsformel finden, die der Dualitätsformel der Brownschen Bewegung ähnelt.

Theorem 3.2. *$(S_n)_{n \geq 0}$ sei eine wie in Definition 3.2 definierte verallgemeinerte symmetrische Irrfahrt. Sei $N \in \mathbb{N}$ beliebig, $0 \leq k_1 \leq \cdots \leq k_N$, $k_j \in \mathbb{N}$ und $(g_j)_{j \geq 1} \subset \mathbb{R}$ beliebig, wobei nur endlich viele Folgenglieder verschieden von Null sind. Sei weiterhin $\phi : \mathbb{R}^N \to \mathbb{R}$ eine meßbare Funktion, quadratintegrabel bezüglich $(S_{k_1}, \ldots, S_{k_N} - S_{k_{N-1}})$ und $(S_{k_1}, \ldots, S_{k_j - 1} - S_{k_{j-1}}, \ldots, S_{k_N} - S_{k_{N-1}})$*

für $j \in \{1, \ldots, N\}$. Dann gilt die diskrete Dualitätsformel

$$
\begin{aligned}
& E(\phi(S_{k_1}, S_{k_2} - S_{k_1}, \ldots, S_{k_N} - S_{k_{N-1}}) \sum_{j=1}^{\infty} g_j Y_j) \\
&= (\sum_{j=1}^{k_1} g_j) E((\phi(S_{k_1}, \ldots, S_{k_N} - S_{k_{N-1}}) - \phi(S_{k_1 - 1}, \ldots, S_{k_N} - S_{k_{N-1}})) Y_{k_1}) + \ldots \\
&\quad + (\sum_{j=k_{N-1}+1}^{k_N} g_j) E((\phi(S_{k_1}, \ldots, S_{k_N} - S_{k_{N-1}}) - \phi(S_{k_1}, \ldots, S_{k_N - 1} - S_{k_{N-1}})) Y_{k_N})
\end{aligned}
\tag{3.11}
$$

Anmerkung 3.4 (Über die Analogien zwischen den Dualitätsformeln). Vergleicht man die eindimensionalen Versionen (das heißt $N = 1$ in der Definition) der Dualitätsformeln der Brownschen Bewegung und der symmetrischen Irrfahrt

$$
(\sum_{j=1}^{k} g_j) E((\phi(S_k) - \phi(S_{k-1})) Y_k) = E(\phi(S_k) \sum_{j=1}^{\infty} g_j Y_j)
$$

und

$$
\int_0^s g(t) dt \; \mathbb{E}_0(\phi'(X_s)) = \mathbb{E}_0(\phi(X_s) \int_0^1 g(t) dX_t),
$$

kann man, nach multiplizieren mit einer Eins, folgende Analogien zwischen der diskreten und stetigen Version erkennen:

- $\frac{\phi(S_k) - \phi(S_{k-1})}{Y_k}$ entspricht $\phi'(X_s)$,
- $\sum_{j=1}^{k} g_j Y_k^2$ entspricht $\int_0^s g(t) dt$,
- $\phi(S_k)$ entspricht $\phi(X_s)$,
- $\sum_{j=1}^{\infty} g_j Y_j$ entspricht $\int_0^1 g(t) dX_t$.

Im zweiten Punkt erwartet man eigentlich Y_j^2, und sieht dann eine quadratische Variation gegen t konvergieren. Allerdings spielt es in späteren Ausführungen keine besondere Rolle, dass wir hier die Y_k^2 vorfinden. Das liegt vor allem an der speziellen Wahl der g_j. Wir werden diese Analogien später nochmal wiedersehen, dann im mehrdimensionalen Bereich, im Rahmen des Beweises des Theorems von Donsker mit der Dualitätsformel der Brownschen Bewegung in Kapitel 4. Dort werden sie allerdings auch in etwas modifizierter Form verwendet.

Beweis von Theorem 3.2. Da die Gleichheit für alle Folgen $(g_j)_{j \geq 1}$ mit nur endlich vielen Folgengliedern verschieden von Null gelten soll, genügt es, wegen der Linearität der diskreten Dualitätsformel auf beiden Seiten, zu zeigen, dass

$$
\begin{aligned}
& E(\phi(S_{k_1}, \ldots, S_{k_N} - S_{k_{N-1}}) Y_l) \\
&= E((\phi(S_{k_1}, \ldots, S_{k_j} - S_{k_{j-1}}, \ldots, S_{k_N} - S_{k_{N-1}}) - \phi(S_{k_1}, \ldots, S_{k_j - 1} - S_{k_{j-1}}, \ldots, S_{k_N} - S_{k_{N-1}})) Y_{k_j})
\end{aligned}
$$

falls $l \in \{k_{j-1}+1, \dots, k_j\}$ für ein $j \in \{1, \dots, N\}$ und sonst

$$E(\phi(S_{k_1}, \dots, S_{k_N} - S_{k_{N-1}})Y_l) = 0.$$

Das ist aber nach Voraussetzung direkt erfüllt. Genauer gesagt, ist für den Fall $l \in \{k_{j-1} + 1, \dots, k_j\}$ für ein $j \in \{1, \dots, N\}$ nach voraussetzen von $E(Y_{k_j}|S_{k_1}, \dots, S_{k_j-1} - S_{k_{j-1}}, \dots, S_{k_N} - S_{k_{N-1}}) = 0$ (dies ist nach Definition, Forderung (3.1), der verallgemeinerten symmetrischen Irrfahrt erfüllt)

$$E(\phi(S_{k_1}, \dots, S_{k_j-1} - S_{k_{j-1}}, \dots, S_{k_N} - S_{k_{N-1}})Y_{k_j}) = 0,$$

womit zu beweisen bleibt, dass

$$E(\phi(S_{k_1}, \dots, S_{k_N} - S_{k_{N-1}})Y_l) = E(\phi(S_{k_1}, \dots, S_{k_N} - S_{k_{N-1}})Y_{k_j}).$$

Da wir aber gesagt haben, dass $l \in \{k_{j-1} + 1, \dots, k_j\}$ und natürlich automatisch $k_j \in \{k_{j-1} + 1, \dots, k_j\}$, ist auch diese Gleichung automatisch nach Definition erfüllt. Für den Fall $l \notin \{k_{j-1} + 1, \dots, k_j\}$ für alle $j \in \{1, \dots, N\}$ gilt

$$E(Y_l|S_{k_1}, \dots, S_{k_N} - S_{k_{N-1}}) = 0,$$

nach Definition der bedingten Erwartung also automatisch

$$E(\phi(S_{k_1}, \dots, S_{k_N} - S_{k_{N-1}})Y_l) = 0.$$

Auch hier hatten wir eigentlich nichts zu beweisen, womit dann das Theorem bewiesen ist. □

Anmerkung 3.5 (Dualitätsformel für die symmetrische Irrfahrt). Für den Fall einer symmetrischen Irrfahrt können wir noch einen alternativen Beweis angeben. Dieser verwendet explizit die Unabhängigkeit, und die identische Verteilung der Y_j. Der „Vorteil“ dieses Beweises ist, dass er das Argument (3.7) aus dem Beweis zum Theorem 3.1, in dem der Erwartungswert als Integral direkt notiert wird, nicht explizit verwendet. Stattdessen kürzen sich hier kanonische Summen die mit Hilfe der Eigenschaften der besonderen Testfunktion, einer Exponentialfunktion, konstruiert werden. Der Schlüssel dazu ist die Gleichheit der charakteristischen Funktionen der Y_j, das ist also der Beweisteil der das Argument (3.7) umgeht. Wir beweisen nur die eindimensionale diskrete Dualitätsformel.

Beweis. Die Summenobergrenze k auf der rechten Seite der Formel ergibt sich, wenn man sich überlegt, dass

$$\begin{aligned} E(\phi(S_k) \sum_{j=1}^{\infty} g_j Y_j) &= E(\phi(S_k) \sum_{j=1}^{k} g_j Y_j) + E(\phi(S_k) \sum_{j=k+1}^{\infty} g_j Y_j) \\ &= E(\phi(S_k) \sum_{j=1}^{k} g_j Y_j) + 0 \end{aligned}$$

aufgrund der Unabhängigkeit und Zentriertheit der $(Y_j)_{j\geq 1}$. Der restliche Beweis setzt deshalb immer voraus, dass es ein $l \leq k$ gibt mit $g_j = 0$, $\forall j > l$.

Wir werden die Formel erstmal für $\phi_\lambda(y) := e^{i\lambda y}$ mit $\lambda \in \mathbb{R}$ beweisen. Für ϕ_λ gilt:

$$\phi_\lambda(y_1 + y_2) = \phi_\lambda(y_1)\phi_\lambda(y_2) \quad \Rightarrow \quad \phi_\lambda(y_1 + y_2) - \phi_\lambda(y_1) = \phi_\lambda(y_1)(\phi_\lambda(y_2) - 1).$$

Wir definieren die folgende Funktion:

$$\Psi(\lambda) := E(e^{i\lambda Y_1}) = E(\phi_\lambda(Y_1)).$$

Ψ ist wohldefiniert, da ϕ_λ beschränkt ist. Außerdem gilt

$$E(\phi_\lambda(S_j)) = \Psi(\lambda)^j \quad \text{und} \quad E(iY_j\phi_\lambda(Y_j)) = \Psi'(\lambda).$$

Die Vertauschung von Integration und Differenziation wird genauso wie im Beweis zu Theorem 2.2 begründet.
Die erwähnten Eigenschaften der Funktion ϕ_λ angewandt auf die Irrfahrt liefern

$$\begin{aligned}
\phi_\lambda(S_k) \quad &= 1 + \sum_{j=1}^{k} (\phi_\lambda(S_j) - \phi_\lambda(S_{j-1})) \\
&= 1 + \sum_{j=1}^{k} \phi_\lambda(S_{j-1})(\phi_\lambda(Y_j) - 1) \\
\Rightarrow E(\phi_\lambda(S_k) \sum_{p=1}^{l} g_p Y_p) \quad &= E(\sum_{p=1}^{l} g_p Y_p) + E(\sum_{p=1}^{l} g_p Y_p \sum_{j=1}^{k} \phi_\lambda(S_{j-1})(\phi_\lambda(Y_j) - 1)) \\
&= \sum_{p=1}^{l} \sum_{j=1}^{k} \; g_p \; E(Y_p \phi_\lambda(S_{j-1})(\phi_\lambda(Y_j) - 1)).
\end{aligned}$$

Aus der letzten Summe berechnen wir die einzelnen Summanden:

- Fall $j < p : E(Y_p \phi_\lambda(S_{j-1})(\phi_\lambda(Y_j) - 1)) = 0$
- Fall $j = p : E(Y_j \phi_\lambda(S_{j-1})(\phi_\lambda(Y_j) - 1)) = \Psi^{j-1}(\lambda)\Psi'(\lambda)(-i)$
- Fall $j > p :$

$$\begin{aligned}
E(Y_p \phi_\lambda(S_{j-1})(\phi_\lambda(Y_j) - 1)) \quad &= E(Y_p \phi_\lambda(S_{p-1})\phi_\lambda(Y_p)\phi_\lambda(S_{j-1} - S_p)(\phi_\lambda(Y_j) - 1)) \\
&= (-i)\Psi^{p-1}(\lambda)\Psi^{j-1-p}(\lambda)(\Psi(\lambda) - 1)\Psi'(\lambda) \\
&= (-i)(\Psi(\lambda) - 1)\Psi^{j-2}(\lambda)\Psi'(\lambda).
\end{aligned}$$

Damit können wir die Summe umschreiben (wir kürzen formal $\Psi(\lambda)$ ab mit Ψ, und setzen aus dem gleichen Grund $G_j = \sum_{p=1}^{j} g_p$):

$$\begin{aligned}
&\sum_{p=1}^{l}\sum_{j=1}^{k} g_p \; E(Y_p\phi_\lambda(S_{j-1})(\phi_\lambda(Y_j)-1)) \\
&= (-i)\sum_{j=2}^{k} G_{l\wedge(j-1)}(\Psi-1)\Psi^{j-2}\Psi' + (-i)\sum_{j=1}^{l} g_j \;\Psi^{j-1}\Psi' \\
&= (-i)\Psi'\left[\sum_{j=2}^{l} G_{j-1}\Psi^{j-1} - \sum_{j=2}^{l+1} G_{j-1}\Psi^{j-2} + \sum_{j=l+1}^{k} G_l\Psi^{j-1} - \sum_{j=l+2}^{k} G_l\Psi^{j-2} + \sum_{j=1}^{l} g_j\Psi^{j-1}\right] \\
&= (-i)\Psi'\left[\sum_{j=2}^{l} G_j\Psi^{j-1} - \sum_{j=1}^{l} G_j\Psi^{j-1} + g_1\Psi^0 + \sum_{j=l+1}^{k} G_l\Psi^{j-1} - \sum_{j=l+1}^{k-1} G_l\Psi^{j-1}\right] \\
&= (-i)\Psi'\left[-g_1\Psi^0 + g_1\Psi^0 + G_l\Psi^{k-1}\right] \\
&= (-i)G_l\Psi^{k-1}\Psi' \\
&= E(G_l\phi_\lambda(S_{k-1})\phi_\lambda(Y_k)Y_k) \\
&= E(G_l(\phi_\lambda(S_k) - \phi_\lambda(S_{k-1}))Y_k)
\end{aligned}$$

Schreiben wir die berechnete Gleichung aus, sehen wir, dass wir für die Funktionen vom Typ ϕ_λ die diskrete Dualitätsformel

$$\Rightarrow E(\phi_\lambda(S_k)\sum_{j=1}^{l} g_jY_j) = (\sum_{j=1}^{l} g_j)E((\phi_\lambda(S_k) - \phi_\lambda(S_{k-1}))Y_k)$$

erhalten haben. Die Erweiterung auf Funktionen, die entsprechend quadratintegrabel sind, erfolgt mit den gleichen Techniken wie im Beweis zu Theorem 2.1. □

3.2 Diskrete Dualitätsformel charakterisiert verallgemeinerte symmetrische Irrfahrt

Die „Rückrichtung“ wird vorerst kaum aufregender werden. Die Beweise sind eigentlich keine Beweise, sondern sind schon so in den Definitionen enthalten, dass man sie eigentlich direkt ablesen kann. Wir werden die Situation für die symmetrische Irrfahrt, wie schon in der Hinrichtung von verallgemeinerter symmetrischer Irrfahrt zu diskreter Dualitätsformel, gesondert betrachten.

Theorem 3.3. *Sei $(S_n)_{n\geq 1}$ eine Folge quadratintegrabler Zufallsvariablen, $S_0 = 0$ (definiere wie immer $Y_j := S_j - S_{j-1}$, $\forall j \geq 1$). Gilt für diese Folge die diskrete Dualitätsformel (3.11), mit entsprechend quadratintegrablen Funktionen ϕ, so ist $(S_n)_{n\geq 0}$ eine verallgemeinerte symmetrische Irrfahrt im Sinne von Definition 3.2.*

Beweis. Setzen wir $\phi \equiv 1$ ergibt sich

$$E(\sum_{j=1}^{\infty} g_jY_j) = 0,$$

woraus folgt, dass die S_n zentriert sind.
Jetzt wählen wir für ein beliebiges entsprechend quadratintegrables ϕ beliebige natürliche Zahlen

$0 \leq k_1 \leq \cdots \leq k_N$ und l, so dass $l \notin \{k_{j-1}+1, \ldots, k_j\}$ für ein $j \in \{1, \ldots, N\}$. Außerdem setzen wir $g_l = 1$ und alle anderen Null. Die diskrete Dualitätsformel ergibt dann automatisch, dass der bedingte Erwartungswert gleich dem Erwartungswert ist, denn auf der rechten Seite der diskreten Dualitätsformel (3.11) kommt g_l in keiner der Summen vor, also

$$E(\phi(S_{k_1}, S_{k_2} - S_{k_1}, \ldots, S_{k_N} - S_{k_{N-1}})Y_l) = 0,$$

für alle entsprechend quadratintegrablen Funktionen ϕ. Und damit gilt

$$E(Y_l | S_{k_1}, S_{k_2} - S_{k_1}, \ldots, S_{k_N} - S_{k_{N-1}}) = 0,$$

Was natürlich äquivalent ist zur Forderung (3.1) der Definition der verallgemeinerten symmetrischen Irrfahrt. Sei jetzt $l \in \{k_{j-1}+1, \ldots, k_j\}$. Da wir gerade bewiesen haben, dass der bedingte Erwartungswert gleich dem Erwartungswert ist, falls $l \notin \{k_{j-1}+1, \ldots, k_j\}$, wird die rechte Seite der diskreten Dualitätsformel zu

$$E((\phi(S_{k_1}, \ldots, S_{k_j} - S_{k_{j-1}}, \ldots, S_{k_N} - S_{k_{N-1}}) - \phi(S_{k_1}, \ldots, S_{k_j - 1} - S_{k_{j-1}}, \ldots, S_{k_N} - S_{k_{N-1}}))Y_{k_j})$$
$$= E(\phi(S_{k_1}, \ldots, S_{k_j} - S_{k_{j-1}}, \ldots, S_{k_N} - S_{k_{N-1}})Y_{k_j}).$$

Insgesamt also die Gleichung

$$E(\phi(S_{k_1}, \ldots, S_{k_j} - S_{k_{j-1}}, \ldots, S_{k_N} - S_{k_{N-1}})Y_l)$$
$$= E(\phi(S_{k_1}, \ldots, S_{k_j} - S_{k_{j-1}}, \ldots, S_{k_N} - S_{k_{N-1}})Y_{k_j}),$$

für alle entsprechend quadratintegrablen Funktionen ϕ. Damit haben wir dann die $f.s.$ geltende Gleichung

$$E(Y_l | S_{k_1}, \ldots, S_{k_j} - S_{k_{j-1}}, \ldots, S_{k_N} - S_{k_{N-1}})$$
$$= E(Y_{k_j} | S_{k_1}, \ldots, S_{k_j} - S_{k_{j-1}}, \ldots, S_{k_N} - S_{k_{N-1}}),$$

und wir haben alle benötigten Eigenschaften einer symmetrischen Irrfahrt bewiesen, denn diese Gleichung ist natürlich äquivalent zur Forderung (3.2) der Definition der verallgemeinerten symmetrischen Irrfahrt. □

Anmerkung 3.6 (Zu den Momenten). Sei $(S_n)_{n \geq 0}$ eine verallgemeinerte symmetrische Irrfahrt. Nach Anmerkung 3.1 sind die ersten beiden Momente der Y_j identisch. Wir wollen jetzt mit Hilfe der Dualitätsformel (also im Prinzip mit Hilfe der dazu äquivalenten Definition) zeigen, dass auch höhere Momente, unter zusätzlichen Annahmen an die Integrabilität der Y_j, gleich sind. Dazu verwenden wir die diskrete Dualitätsformel in der eindimensionalen Variante, mit $g_l = 1$ für ein $l \in \mathbb{N}$ und alle anderen g_j sind Null. Die Dualitätsformel lautet dann

$$E(\phi(S_k)Y_l) = E((\phi(S_k) - \phi(S_{k-1}))Y_k).$$

Wir werden nur die Gleichheit gewisser Momente von Y_1 und Y_2 beweisen, die Gleichheit der Momente der anderen Y_j folgt dann induktiv. Setzen wir also $k = 2$ und $l = 1$, dann wird die eindimensionale diskrete Dualitätsformel zu

$$E(\phi(S_2)Y_1) = E(\phi(S_2)Y_2).$$

Um die Gleichheit der dritten Momenten der verallgemeinerten symmetrischen Irrfahrt zu beweisen, benötigen wir die Zusatzannahme, dass $\phi(S_n) = S_n^2$ quadratintegrabel ist für alle $n \in \mathbb{N}$ (man benötigt also die Existenz der vierten Momente der Y_j). Die Herleitung funktioniert im Prinzip genauso wie bei der Herleitung der Gleichheit der zweiten Momente in Anmerkung 3.1, deshalb sind die Erklärungen jetzt weniger ausführlich. Setzen wir also $\phi(y) = y^2$, dann

$$\begin{aligned}
E(\phi(S_2)Y_1) &= E(\phi(S_2)Y_2) \\
\Rightarrow E((Y_1 + Y_2)^2 Y_1) &= E((Y_1 + Y_2)^2 Y_2) \\
\Rightarrow E(Y_1^3) + 2E(Y_1^2 Y_2) + E(Y_1 Y_2^2) &= E(Y_1^2 Y_2) + 2E(Y_1 Y_2^2) + E(Y_2^3) \\
\Rightarrow E(Y_1^3) &= E(Y_2^3).
\end{aligned}$$

Setzen wir weiterhin voraus, dass $\phi(S_n) = S_n^3$ quadratintegrabel ist für alle $n \in \mathbb{N}$, erhalten wir die Gleichheit der vierten Momente:

$$\begin{aligned}
E((Y_1 + Y_2)^3 Y_1) &= E((Y_1 + Y_2)^3 Y_2) \\
\Rightarrow E(Y_1^4) + 3E(Y_1^3 Y_2) + 3E(Y_1^2 Y_2^2) + E(Y_1 Y_2^3) &= E(Y_1^3 Y_2) + 3E(Y_1^2 Y_2^2) + 3E(Y_1 Y_2^3) + E(Y_2^4) \\
\Rightarrow E(Y_1^4) &= E(Y_2^4).
\end{aligned}$$

Dieses System funktioniert allerdings nicht mehr um die Gleichheit der fünften Momente herzuleiten. Wir nehmen an, dass S_n^4 quadratintegrabel ist für alle $n \in \mathbb{N}$, und erhalten

$$\begin{aligned}
E((Y_1 + Y_2)^4 Y_1) &= E((Y_1 + Y_2)^4 Y_2) \\
\Rightarrow E(Y_1^5) + 6E(Y_1^3 Y_2^2) + 4E(Y_1^2 Y_2^3) &= E(Y_2^5) + 6E(Y_1^2 Y_2^3) + 4E(Y_1^3 Y_2^2).
\end{aligned}$$

Hier wird zusätzlich die Gleichung $E(Y_1^2 Y_2^3) = E(Y_1^3 Y_2^2)$ benötigt. Diese Gleichung ist für die symmetrische Irrfahrt erfüllt, was wir uns im nächsten Theorem zunutze machen.

Theorem 3.4. *Sei $(S_n)_{n \geq 0}$ eine Folge von Zufallsvariablen mit $S_0 = 0$ und S_n^p quadratintegrabel für alle $p, n \in \mathbb{N}$ (das heißt $S_n \in \bigcap_{p=1}^{\infty} L^p$), außerdem sei die durch $Y_j := S_j - S_{j-1}$ definierte Folge von Zufallsvariablen unabhängig. Die diskrete Dualitätsformel (3.11) sei für $(S_n)_{n \geq 0}$ erfüllt. Dann ist $(S_n)_{n \geq 0}$ eine symmetrische Irrfahrt.*

Beweis. Wir wissen nach Theorem 3.3 bereits, dass $(S_n)_{n \geq 1}$ eine verallgemeinerte symmetrische Irrfahrt ist. Da die Y_j unabhängig sind, muss nur noch gezeigt werden, dass sie auch identisch verteilt sind. Dazu müssen wir nur zeigen, dass alle Momente der Y_j gleich sind. Den Anfang haben wir gerade schon in obiger Anmerkung geleistet. Da die Festlegung $Var(Y_1) = 1$ willkürlich war, setzen wir sie hier ebenfalls als gegeben voraus. Wir verwenden die Funktion

$$\phi(y) = y^p \quad \Rightarrow \phi \in L^2.$$

Wir zeigen jetzt die Gleichheit aller Momente von Y_1 und Y_2. Induktiv ergibt sich daraus die Gleichheit der Momente von Y_k und Y_l für beliebige $k, l \in \mathbb{N}$, denn wir können dann die Gleichheit der Momente von Y_2 und Y_3 auf die gleiche Weise zeigen, induktiv dann für Y_{k-1} und Y_k.
Der Beweis der Gleichheit der $p+1-$ten Momente, wenn bereits gezeigt ist, das die ersten p Momente gleich sind, funktioniert genauso wie die Beispielrechnungen in der obigen Anmerkung. Wir erinnern uns dazu an die binomische Formel

$$(a+b)^p = \sum_{k=0}^{p} \binom{p}{k} a^{p-k} b^k,$$

und setzen in die Dualitätsformel ein:

$$\begin{aligned} E((Y_1+Y_2)^p Y_1) &= E(((Y_1+Y_2)^p - Y_1^p)Y_2) = E((Y_1+Y_2)^p Y_2) \\ \Rightarrow \sum_{k=0}^{p} E(\binom{p}{k} Y_1^{p-k+1} Y_2^k) &= \sum_{k=0}^{p} E(\binom{p}{k} Y_1^{p-k} Y_2^{k+1}) \\ \Rightarrow E(Y_1^{p+1}) + \sum_{k=1}^{p} \binom{p}{k} E(Y_1^{p-k+1} Y_2^k) &= E(Y_2^{p+1}) + \sum_{k=0}^{p-1} \binom{p}{k} E(Y_1^{p-k} Y_2^{k+1}) \\ \Rightarrow E(Y_1^{p+1}) + \sum_{k=1}^{p} \binom{p}{k} E(Y_1^{p-k+1}) E(Y_2^k) &= E(Y_2^{p+1}) + \sum_{k=1}^{p} \binom{p}{k} E(Y_2^{p-k+1}) E(Y_1^k) \\ \Rightarrow E(Y_1^{p+1}) &= E(Y_2^{p+1}) \end{aligned}$$

Im vorletzten Schritt haben wir umsummiert und verwendet, dass $\binom{p}{k} = \binom{p}{p-k}$ gilt. Der letzte Schritt hat dann die Induktionsvoraussetzung verwendet, dass die Momente bis p gleich sind. Aus der Momentengleichheit von Y_1 und Y_2 folgt dann die Gleichheit der Verteilungen.
Jetzt nutzen wir die Dualitätsformel

$$E(\phi(S_{k+1} - S_{k-1})Y_k) = E((\phi(S_{k+1} - S_{k-1}) - \phi(S_k - S_{k-1}))Y_{k+1}),$$

um auf die gleiche Art von der Momentgleichheit der $(Y_j)_{1 \leq j \leq k}$ und der Momentgleichheit von Y_k und Y_{k+1} auf die Momentgleichheit der $(Y_j)_{1 \leq j \leq k+1}$ zu schließen. Wir erhalten vermöge dieser Induktion, dass die $(Y_j)_{j \geq 1}$ identisch verteilt sind. Damit ist $(S_n)_{n \geq 0}$ eine symmetrische Irrfahrt im Sinne unserer Definition 3.1. $\square$

4 Donskers Theorem und die Dualitätsformeln

Der erste Beweis des Konvergenzergebnisses von Donsker wurde im Jahr 1951 vom Namenspatron M. Donsker selbst geführt. In seinem Artikel [Don] beweist er das Theorem ohne explizite Verwendung der Techniken der Straffheit von Wahrscheinlichkeitsmaßen auf $C(0,1)$. Vorarbeit wurde insofern geleistet, als dass die schwache Konvergenz gewisser Pfadfunktionale bereits bewiesen wurde (u.a. von Erdös und Kac). So kannte man unter anderem die Grenzverteilungen von (entsprechend normierten Versionen von) $\sum_{j=1}^n S_j$, $\max\{S_1,\ldots,S_n\}$, $\max\{|S_1|,\ldots,|S_n|\}$, $\sum_{j=1}^n S_j^2$ und $\sum_{j=1}^n |S_j|$.

Das Theorem von Donsker wollen wir hier auf dem Raum stetiger Funktionen auf $[0,1]$ beweisen. Den Beweis führen wir für die symmetrische Irrfahrt. In den folgenden Beweisen wird dann an geeigneter Stelle auf die explizite Verwendung von Eigenschaften, die wir bei der verallgemeinerten symmetrischen Irrfahrt aus Definition 3.2 nicht hätten, hingewiesen. Die symmetrische Irrfahrt aus Definition 3.1 bisher ist alles andere als stetig, und der Zeitindex sind die natürlichen Zahlen. Wir verwenden deshalb eine geeignete stetige Version die wir jetzt entwickeln werden. Zuerst machen wir die Irrfahrt stetig. Wir wollen dazu im Zeitindex die Zahlen $t \in [0,1]$ zulassen. Dazu stauchen wir alle Partialsummen S_k bis zum einem festen Index n in das Intervall $[0,1]$ hinein. Es ergibt sich eine Zufallsfunktion auf $[0,1]$ die auch von $n \in \mathbb{N}$ abhängt:

$$S_t^{(n)} := \sum_{j=1}^{\lfloor tn \rfloor} Y_j + (tn - \lfloor tn \rfloor) Y_{\lfloor tn \rfloor + 1},$$

wobei die leere Summe formal 0 gesetzt wird. Wir wollen, dass diese Zufallsfunktionen in n auf $C(0,1)$ konvergieren. Dazu fehlt noch ein Normierungsfaktor vor der Zufallsfunktion. Wenn in dieser Zufallsfunktion $t = 1$ gesetzt wird, hat man wieder die Werte der symmetrischen Irrfahrt. Damit ist eigentlich klar, dass als Normierungsfaktor nur $1/\sqrt{n}$ in Frage kommt, denn der zentrale Grenzwertsatz ist auf die Konvergenz von $S_1^{(n)}$ anwendbar. Hier ist direkt nachvollziehbar, dass es sich beim Invarianzprinzip von Donsker um das unendlich dimensionale Analogon zum zentralen Grenzwertsatz handelt (dementsprechend es sich beim Wienermaß um das Analogon zur Normalverteilung handelt). Wir vermuten also, dass die Zufallsfunktion

$$\hat{S}_t^{(n)} := \frac{1}{\sqrt{n}} \left(\sum_{j=1}^{\lfloor tn \rfloor} Y_j + (tn - \lfloor tn \rfloor) Y_{\lfloor tn \rfloor + 1} \right).$$

schwach im Raum $C(0,1)$ konvergieren wird. Die modifizierte Notation in folgender Definition verwenden wir um klarzumachen, dass es sich jetzt nicht mehr um eine symmetrische Irrfahrt, sonder um Zufallsfunktionen auf $C(0,1)$ handelt, die bestimmte Eigenschaften der symmetrischen Irrfahrt erben.

Anmerkung 4.1 (Warum nicht die verallgemeinerte symmetrische Irrfahrt?). Im Jahr 1961 bewies Billingsley als erster ein Konvergenzergebnis im Sinne des zentralen Grenzwertsatzes für stationäre, ergodische Martingale (siehe [Bil1]), zwei Jahre später bewies Ibragimov selbiges Theorem mit etwas anderen Mitteln (siehe [Ibr] Theorem 1, eine Abschätzung der Konvergenzrate findet sich in Theorem 2). Diese Ergebnisse lassen sich natürlich leicht auf endlichdimensionale Martingale erweitern. Die Erweiterung auf ein funktionales Invarianzprinzip, ähnlich dem Konvergenzergeb-

nis von Donsker, folgte dann. So kann man zum Beispiel in [Bil2], Theorem 23.1 das Ergebnis für stationäre, ergodische Martingale finden.
Die Frage ist nun, warum sollte man nicht für eine renormierte stetige Irrfahrt basierend auf der verallgemeinerten symmetrischen Irrfahrt definieren, und für dieses den Beweis der schwachen Konvergenz gegen das Wienermaß durchführen. Ein Indikator, dass es funktionieren könnte, ist zum Beispiel die Existenz eines zentralen Grenzwertsatzes für die Gibbs-Felder aus Anmerkung 3.3. Ob für diese speziellen Felder auch ein funktionales Invarianzprinzip existiert, ist mir allerdings nicht bekannt, da für diese Felder noch Stationarität und Ergodizität gezeigt werden müsste. Zwischen der verallgemeinerten symmetrischen Irrfahrt und dem stationären, ergodischen Martingal kann man zudem nicht wirkliche Vergleiche ziehen. Es gibt zwar verallgemeinerte symmetrische Irrfahrten, die auch ein stationäres, ergodisches Martingal sind (das Gegenbeispiel dafür, das nicht jede verallgemeinerte symmetrische Irrfahrt eine symmetrische Irrfahrt ist aus dem Beweis zu Theorem 3.1 gehört dazu), die Vermutung ist aber, dass es sowohl stationäre, ergodische Martingale gibt, die keine verallgemeinerten Irrfahrten sind (vergleiche dazu das Gegenbeispiel zu (3.5) impliziert nicht (3.1) aus Anmerkung 3.2), als auch verallgemeinerte symmetrische Irrfahrten die keine stationären, ergodischen Martingale sind (die Stationarität impliziert identische Verteilung der Y_j, diese ist bei verallgemeinerten symmetrischen Irrfahrten nicht evident, und konnte in Theorem 3.4 nur unter Hinzunahme der Unabhängigkeit bewiesen werden). Den Begriff einer verallgemeinerten symmetrischen Irrfahrt mussten wir aber trotz aller Probleme bilden, da sonst die exakte charakteristische diskrete Dualitätsformel Restglieder enthalten hätte, womit sie nicht mehr exakt und höchstens asymptotisch charakteristisch gewesen wäre.
Bei den folgenden Ausführungen sind deshalb immer wieder Bemerkungen gemacht worden, die Beweisteile, in denen die speziellen Eigenschaften einer symmetrischen Irrfahrt benutzt werden, herausgehoben sind.

Definition 4.1 (renormierte stetige Irrfahrt)**.** Sei $(\xi_j)_{j\geq 1}$ eine Folge von identisch verteilten, unabhängigen Zufallsvariablen mit Erwartungswert 0 und Varianz 1. Für jedes $n \in \mathbb{N}$ definieren wir die Zufallsfunktion $X^{(n)} \in C(0,1)$ durch

$$X_t^{(n)} := \frac{1}{\sqrt{n}} \left(\sum_{j=1}^{\lfloor tn \rfloor} \xi_j + (tn - \lfloor tn \rfloor)\xi_{\lfloor tn \rfloor + 1} \right), \quad t \in [0,1].$$

Dann heißt die Folge $(X^{(n)})_{n\geq 1}$ renormierte stetige Irrfahrt.

$\lfloor tn \rfloor$ bezeichnet die größte natürliche Zahl kleiner als tn (entsprechend bezeichnet $\lceil tn \rceil$ die kleinste natürliche Zahl größer als tn).

Anmerkung 4.2 (über die Beschränkung auf stetige Funktionen). Wir haben recht viel Arbeit darauf verwendet, dass wir die Konvergenz auf dem Raum stetiger Pfade $C(0,1)$ beweisen können. Eine andere Möglichkeit wäre eine renormierte Irrfahrt auf $D(0,1)$ zu verwenden. $D(0,1)$ ist der Raum der càdlàg Funktionen auf $[0,1]$. Càdlàg steht für „continue à droite, limitée à gauche", dass heißt rechtsstetig mit Grenzwert von links. Eine renormierte Irrfahrt könnte zum Beispiel eine Treppenfunktion auf diesem Raum sein. Wir verwenden aber die stetige Version, da wir dann nicht extra den Schritt auf $C(0,1)$ zurückgehen müssen und das Grenzmaß auf die stetigen Funktionen einschränken. Das heißt, um die Dualitätsformel (2.3) für die Identifizierung des Wahr-

scheinlichkeitsmaßes anwenden zu können, müssten wir das Grenzmaß der renormierten stetigen Irrfahrt auf $C(0,1)$ beschränken, was nicht trivial ist.
Einen Beweis des Konvergenztheorems von Donsker auf dem Raum $D(0,1)$ gibt es zum Beispiel in [Bil3] (Section 14, Theorem 14.2).

Wie schon erwähnt konvergiert nach dem berühmten Invarianzprinzip von Donsker die Folge $(X^{(n)})_{n\in\mathbb{N}}$ in Verteilung gegen die Standard Brownsche Bewegung ($X_0 = 0$ $f.s.$) auf $C(0,1)$. Der hier geführte Beweis wird sich an dem Beweis in [Bil3] (Section 8) orientieren (dazu Theorem 5.5 im Anhang). Zuerst wird gezeigt, dass die Folge $(\mathcal{L}(X^{(n)}))_{n\in\mathbb{N}}$ der Verteilungen von $(X^{(n)})_{n\in\mathbb{N}}$ straff ist. In [Bil3] wird dann die Grenzverteilung auf $C(0,1)$ mit Hilfe der endlichdimensionalen Grenzverteilungen identifiziert. Wir verwenden zur Identifizierung stattdessen die Dualitätsformeln. Zur Erinnerung und für die Vollständigkeit werden wir nun erst einmal die Straffheit der Folge $(\mathcal{L}(X^{(n)}))_{n\in\mathbb{N}}$ beweisen (siehe [Bil3], Section 8).

4.1 Straffheit der renormierten stetigen Irrfahrt

Da $P(X_0^{(n)} = 0) = 1$, muß nur noch gezeigt werden, dass

$$\begin{aligned} &\forall\epsilon > 0,\ \lim_{\delta\to 0}\lim_{n\to\infty} P(\omega_\delta(X^{(n)}) > \epsilon) = 0, \\ &\text{mit} \quad \omega_\delta(x) := \sup_{0\le s\le t\le 1,\ t-s<\delta} |x(t) - x(s)|\,,\ \delta \in]0,1[. \end{aligned} \tag{4.1}$$

Genaueres zur Straffheit von Wahrscheinlichkeitsmaßen findet sich im Anhang (Theorem 5.6). Zum Beweis benötigen wir einige Vorbereitung.

Lemma 4.1 (siehe [Bil3], Section 8)**.** *Definiere* $S_n := \xi_1 + \cdots + \xi_n$ ($S_0 := 0$). *Angenommen, dass*

$$\lim_{\lambda\to\infty}\limsup_{n\to\infty} \lambda^2 P\left(\max_{k\le n}|S_k| \ge \lambda\sqrt{n}\right) = 0, \tag{4.2}$$

dann ist $(X^{(n)})_{n\in\mathbb{N}}$ *straff.*

Beweis. Gleich zu Beginn benötigen wir einen kleinen Hilfssatz:

Hilfssatz 4.1 (siehe [Bil3], Section 8)**.** *Sei* $0 = t_0 < \cdots < t_N = 1$, *und* $\min_{1<j<N}(t_j - t_{j-1}) \ge \delta$. *Dann gilt für beliebiges* $x \in C(0,1)$

$$\omega_\delta(x) \le 3 \max_{1\le j\le N} \sup_{t_{j-1}\le s\le t_j} |x(s) - x(t_{j-1})|\,,$$

und für ein beliebiges Wahrscheinlichkeitsmaß P *auf* $C(0,1)$*,*

$$P(x : \omega_\delta(x) \ge 3\epsilon) \le \sum_{j=1}^{N} P(x : \sup_{t_{j-1}\le s\le t_j} |x(s) - x(t_{j-1})| \ge \epsilon).$$

Beweis des Hilfssatzes. Für den Beweis benötigen wir

$$M := \max_{1\le j\le N} \sup_{t_{j-1}\le s\le t_j} |x(s) - x(t_{j-1})|\,.$$

Wenn s und t im selben Intervall $I_j := [t_{j-1}, t_j]$ liegen, dann gilt nach Definition von M:

$$|x(s) - x(t)| \leq |x(s) - x(t_{j-1})| + |x(t) - x(t_{j-1})| \leq 2M.$$

Wenn s und t in benachbarten Intervallen I_j und I_{j+1} liegen, dann

$$|x(s) - x(t)| \leq |x(s) - x(t_{j-1})| + |x(t_{j-1}) - x(t_j)| + |x(t_j) - x(t)| \leq 3M.$$

Wenn $|s - t| \leq \delta$, dann müssen s und t nach Definition der $(t_j)_{j \in \{0,\ldots,N\}}$ im gleichen, oder in benachbarten Intervallen liegen (hier sieht man, warum die Bedingung, dass $t_N - t_{N-1} > \delta$ nicht gestellt werden muss und auch nicht gestellt wird; gleiches gilt für $t_1 - t_0 > \delta$). Damit folgt

$$\begin{aligned}
\omega_\delta(x) &= \sup_{0 \leq s \leq t \leq 1,\, t-s<\delta} |x(t) - x(s)| \\
&= \sup_{0 \leq s \leq t \leq 1} \{ \sup_{|t-s|<\delta} |x(t) - x(s)| \} \\
&\leq \sup_{0 \leq s \leq t \leq 1} \{3M\} \\
&= 3M \\
&= 3 \max_{1 \leq j \leq N} \sup_{t_{j-1} \leq s \leq t_j} |x(s) - x(t_{j-1})| .
\end{aligned}$$

Die zweite Aussage können wir nun aufgrund der Subadditivität von Wahrscheinlichkeitsmaßen erhalten:

$$\begin{aligned}
P(x : \omega_\delta(x) \geq 3\epsilon) &\leq P(x : \max_{1 \leq j \leq N} \sup_{t_{j-1} \leq s \leq t_j} |x(s) - x(t_{j-1})| \geq \epsilon) \\
&= P(\bigcup_{j=1}^{N} \{x : \sup_{t_{j-1} \leq s \leq t_j} |x(s) - x(t_{j-1})| \geq \epsilon\}) \\
&\leq \sum_{j=1}^{N} P(x : \sup_{t_{j-1} \leq s \leq t_j} |x(s) - x(t_{j-1})| \geq \epsilon)
\end{aligned}$$

□

Durch einfache Anwendung des Hilfssatzes (gleiche Voraussetzungen für die t_j) erhalten wir

$$P(\omega_\delta(X^{(n)}) \geq 3\epsilon) \leq \sum_{j=1}^{N} P(\sup_{t_{j-1} \leq s \leq t_j} |X_s^{(n)} - X_{t_{j-1}}^{(n)}| \geq \epsilon).$$

Dieser Ausdruck läßt sich leichter analysieren, wenn wir $t_j = \frac{m_j}{n}$ setzen für natürliche Zahlen $0 = m_0 < m_1 < \cdots < m_N = n$. Aufgrund des polygonalen Charakters der Zufallsfunktionen $X^{(n)}$ wird das Supremum aus obiger Gleichung dann zu einem Maximum von Differenzen der Art $|S_k - S_{m_{j-1}}|/\sqrt{n}$:

$$P(\omega_\delta(X^{(n)}) \geq 3\epsilon) \leq \sum_{j=1}^{N} P(\max_{m_{j-1} \leq k \leq m_j} \frac{|S_k - S_{m_{j-1}}|}{\sqrt{n}} \geq \epsilon).$$

Da die $(\xi_n)_{n\geq 1}$ eine iid.-Folge sind, können wir dies weiter vereinfachen zu

$$P(\omega_\delta(X^{(n)}) \geq 3\epsilon) \leq \sum_{j=1}^{N} P(\max_{k\leq m_j - m_{j-1}} |S_k| \geq \sqrt{n}). \tag{4.3}$$

Aufgrund der Forderung an die t_j hält die Ungleichung nur dann, wenn

$$\frac{m_j}{n} - \frac{m_{j-1}}{n} \geq \delta, \quad 1 < j < N.$$

Um weiter zu vereinfachen, wählen wir ein geeignetes $m \in \mathbb{N}$ und setzen $m_j = jm$ für $0 \leq j < N$ und $m_N = n$. Wir müssen die Ungleichung $m_j - m_{j-1} = m \geq n\delta$ für $j \leq N$ erfüllen, setze dazu $m = \lceil n\delta \rceil$. Außerdem muß $(N-1)m < nm < Nm$ gelten, setze $N = \lceil n/m \rceil$. Um hier sinnvolle Definitionen zu erhalten, ist es klar, dass δ klein genug, und dass n groß genug gewählt sein muss. Da wir in der Straffheitsbedingung (4.1) aber sowieso den Grenzwert bilden, ist diese Wahl gerechtfertigt. Dann gilt

i) $m_N - m_{N-1} \leq m$ (man erinnere sich hier daran, dass die Bedingung $m_N - m_{N-1} > n\delta$ am Rand nicht erfüllt sein muss);

ii) für δ klein genug ist $N \leq \frac{2}{\delta}$, denn

$$N = \left\lceil \frac{n}{m} \right\rceil = \left\lceil \frac{n}{\lceil n\delta \rceil} \right\rceil \leq \left\lceil \frac{n}{n\delta} \right\rceil = \left\lceil \frac{1}{\delta} \right\rceil \leq \frac{1}{\delta} + 1 \leq \frac{2}{\delta} \qquad \text{für } \delta \leq 1;$$

iii) für n groß genug bei festem δ gilt $\frac{n}{m} \geq \frac{1}{2\delta}$, denn

$$\frac{n}{m} = \frac{n}{\lceil n\delta \rceil} \geq \frac{n}{n\delta + 1} \to \frac{1}{\delta} > \frac{1}{2\delta}.$$

Mit diesen drei Punkten, mit oben bewiesener Summenabschätzung und dem Fakt dass $m_1 = m$ und $m_N - m_{N-1} \leq m$ gilt, haben wir

$$\begin{aligned} P(\omega_\delta(X^{(n)}) \geq 3\epsilon) \quad & \leq \sum_{j=1}^{N} P(\max_{k\leq m_j - m_{j-1}} |S_k| \geq \epsilon\sqrt{n}) \\ & \overset{i)}{\leq} \sum_{j=1}^{N} P(\max_{k\leq m} |S_k| \geq \epsilon\sqrt{n}) \\ & = NP(\max_{k\leq m} |S_k| \geq \epsilon\sqrt{n}) \\ & \overset{ii)}{\leq} \tfrac{2}{\delta} P(\max_{k\leq m} |S_k| \geq \epsilon\sqrt{n}) \\ & \overset{iii)}{\leq} \tfrac{2}{\delta} P(\max_{k\leq m} |S_k| \geq \epsilon\sqrt{\tfrac{m}{2\delta}}). \end{aligned}$$

Wir definieren λ als Funktion von δ und ϵ: $\lambda = \epsilon/\sqrt{2\delta}$ ($\Rightarrow \delta = \frac{\epsilon^2}{2\lambda^2}$). Als Ausdruck von λ ergibt sich dann die Ungleichung

$$P(\omega_\delta(X^{(n)}) \geq 3\epsilon) \leq \frac{4\lambda^2}{\epsilon^2} P(\max_{k\leq m} |S_k| \geq \lambda\sqrt{m}).$$

$\epsilon > 0$ ist fest, wir wählen ein beliebiges $\eta > 0$. Nach Voraussetzung 4.2 gibt es dann ein λ (λ groß genug), so dass

$$\frac{4\lambda^2}{\epsilon^2} \limsup_{m\to\infty} P(\max_{k\leq m} |S_k| \geq \lambda\sqrt{m}) < \eta.$$

Sind λ und damit auch δ fest und λ groß genug, bzw. δ klein genug, erhalten wir also

$$\limsup_{n\to\infty} P(\omega_\delta(X^{(n)}) \geq 3\epsilon) \leq \frac{4\lambda^2}{\epsilon^2} \limsup_{n\to\infty} P(\max_{k\leq\lceil n\delta\rceil} |S_k| \geq \lambda\sqrt{\lceil n\delta\rceil}) < \eta.$$

Da η beliebig gewählt war, haben wir

$$\lim_{\lambda\to\infty} \frac{4\lambda^2}{\epsilon^2} \limsup_{n\to\infty} P(\max_{k\leq\lceil n\delta\rceil} |S_k| \geq \lambda\sqrt{\lceil n\delta\rceil}) = 0,$$

und damit auch

$$\lim_{\delta\to 0} \lim_{n\to\infty} P(\omega_\delta(X^{(n)}) \geq 3\epsilon) = 0.$$

Dies ist nun bis auf den unwichtigen Faktor 3 (ϵ ist ja beliebig) genau unser Straffheitskriterium (4.1). $\square$

Anmerkung 4.3 (zur Version mit verallgemeinerter symmetrischer Irrfahrt). Im Beweis des Lemmas wurde nur einmal, und zwar für die Gleichung (4.3), verwendet, dass wir statt einer verallgemeinerten symmetrischen Irrfahrt eine symmetrische Irrfahrt zu Grunde gelegt haben. Der Beweis zur Straffheit der Folge $(\mathcal{L}(X^{(n)}))_{n\geq 1}$ wird aber vor allem im nächsten Theorem die Eigenschaften einer symmetrischen Irrfahrt nutzen, und zwar explizit in der Anwendung des zentralen Grenzwertsatzes. Das zeigt, dass ein Beweis der Straffheit einer Folge $(\mathcal{L}(X^{(n)}))_{n\geq 1}$ die nur auf der verallgemeinerten symmetrischen Irrfahrt basiert, andere Techniken verwende müsste (wenn es denn überhaupt möglich ist, das zu beweisen).

Theorem 4.1. *Die Folge* $(\mathcal{L}(X^{(n)}))_{n\in\mathbb{N}}$ *ist straff.*

Beweis. Der Beweis zielt jetzt natürlich genau darauf ab, unser neues Kriterium (4.2) zu verwenden. Wir betrachten die großen und kleinen Werte von k im Maximum auf der rechten Seite von (4.2) separat. Nach dem zentralen Grenzwertsatz gilt für k_λ groß genug, und $k_\lambda \leq k \leq n$

$$P(|S_k| \geq \lambda\sqrt{n}) \leq P(|S_k| \geq \lambda\sqrt{k}) = P(\frac{|S_k|}{\sqrt{k}} \geq \lambda) < 3/\lambda^4,$$

denn für eine standartnormalverteilte Zufallsvariable N ($\mathcal{L}(N) = N(0,1)$) gilt nach der Chebychev Ungleichung

$$P(|N| \geq \lambda) < E(N^4)/\lambda^4 = 3/\lambda^4.$$

Für $k < k_\lambda$ verwenden wir ebenfalls Chebychev's Ungleichung:

$$P(|S_k| \geq \lambda\sqrt{n}) \leq \frac{\mathrm{Var}(S_k)}{\lambda^2 n} = \frac{k}{\lambda^2 n} \leq \frac{k_\lambda}{\lambda^2 n}.$$

Das in der Ungleichung 4.2 gesuchte Maximum ist daher beschränkt durch $(3/\lambda^4) \vee (k_\lambda/\lambda^2 n)$. Wir erhalten

$$
\begin{aligned}
0 &\leq \lim_{\lambda\to\infty} \limsup_{n\to\infty} \lambda^2 P\left(\max_{k\leq n} |S_k| \geq \lambda\sqrt{n}\right) \\
&\leq \lim_{\lambda\to\infty} \limsup_{n\to\infty} \lambda^2 \left(\tfrac{3}{\lambda^4} \vee \tfrac{k_\lambda}{\lambda^2 n}\right) \\
&= \lim_{\lambda\to\infty} \lim_{n\to\infty} \left(\tfrac{3}{\lambda^2} \vee \tfrac{k_\lambda}{n}\right) \\
&= \lim_{\lambda\to\infty} \tfrac{3}{\lambda^2} \\
&= 0
\end{aligned}
$$

Nach Anwendung des Kriteriums aus Lemma 4.1 erhalten wir also die Straffheit der Irrfahrt. □

4.2 Konvergenz der Irrfahrt

Um das Ergebnis von Donsker zu beweisen, bleibt es jetzt noch übrig, das Wahrscheinlichkeitsmaß zu identifizieren gegen das die Folge $(\mathcal{L}(X^{(n)}))_{n\geq 1}$ konvergiert. Wir wollen dies vermöge der Dualitätsformeln erreichen.
Wir wissen zunächst einmal, dass das Wienermaß durch die Dualitätsformel der Brownschen Bewegung charakterisiert wird. Wenn man die Aussage des Theorems 2.2 beachtet, erkennt man, dass es genügt zu zeigen, dass die stetige Irrfahrt für Funktionen des Typs $\mathcal{W}$ die Dualitätsformel der Brownschen Bewegung asymptotisch erfüllen. Im welchem Sinne asymptotisch und mit welcher Konvergenzgeschwindigkeit wird das nächste Theorem aufdecken.

Theorem 4.2 (approximative Dualitätsformel)**.** *Sei $F \in \mathcal{W}$, g Treppenfunktion auf $[0,1]$. Wir verlangen als Zusatzannahme an die Definition der renormierten stetigen Irrfahrt, dass $\xi_1 \in L^3$ gilt. Diese renormierte stetige Irrfahrt erfüllt die approximative Dualitätsformel*

$$
E(F(X^{(n)}) \int_0^1 g(t)dX_t^{(n)}) = E(D_g F(X^{(n)})) + O(n^{-\frac{1}{2}}). \tag{4.4}
$$

Anmerkung 4.4 (das große Landau-Symbol)*.* Der Ausdruck $O(n^{-\frac{1}{2}})$ steht dafür, dass in der Gleichung (4.4) eine Abweichung vorkommt die asymptotisch von der Größenordnung $n^{-\frac{1}{2}}$ sein wird. Das Landau-Symbol ist im weiteren Verlauf immer mit diesem asymptotischen Hintergrund zu verstehen. Ebenso ist das kleine Wörtchen approximativ bezüglich der Dualitätsformel genau in diesem asymptotischen Sinne zu verstehen.
Aus dem Beweis wird hervorgehen, dass es eine positive Konstante $K \in \mathbb{R}^+$ gibt, so dass

$$
\left| E(F(X^{(n)}) \int_0^1 g(t)dX_t^{(n)}) - E(D_g F(X^{(n)})) \right| \leq K\, n^{-\frac{1}{2}}, \tag{4.5}
$$

für alle n.

Beweis des Theorems 4.2. Da $F \in \mathcal{W}$ ist, werden wir der Anmerkung 2.2 folgend, im Beweis zwischen den zwei verschiedenen Schreibweisen

$$
F(x) = \phi(x_{s_1}, \ldots, x_{s_N})
$$

und

$$F(x) = \tilde{\phi}(x_{s_1}, x_{s_2} - x_{s_1}, \dots, x_{s_N} - x_{s_{N-1}})$$

kommentarlos hin und her wechseln. Außerdem werden wir das Theorem vorerst nur für die spezielle Treppenfunktion $g = 1_{[0,t[}$, für $t \in [0,1]$ beliebig, beweisen. Wir werden später sehen, dass sich das Ergebnis sehr einfach auf allgemeine Treppenfunktionen erweitern lässt.
Der Beweis funktioniert durch sukzessive Approximationen mit Hilfe der Taylorformel (5.1). Alle Approximationen gelten in L^1, implizieren also eine Konvergenz der Erwartungswerte gegeneinander. Das Theorem kann natürlich keine L^1-Konvergenz zum Ergebnis haben, da die im Beweis anzuwendende diskrete Dualitätsformel ja eine Gleichung zwischen Erwartungswerten ist.
Sei $n \in \mathbb{N}$ fest gewählt. Damit wir überhaupt der diskreten Dualitätsformel im Kontext mit der renormierten stetigen Irrfahrt einen Sinn geben können, müssen die Zeitpunkte $s_1, \dots, s_N, t$ Vielfache von $1/n$ sein. Wir nehmen an, dass

$$s_1 = \frac{k_1}{n}, \dots, s_N = \frac{k_N}{n}, t = \frac{l}{n}, \quad \text{mit } k_1, \dots, k_N, l \in \mathbb{N},\ 0 \leq k_1 \leq, \dots, \leq k_N \leq n.$$

Betrachten wir nun die Definitionen der symmetrischen Irrfahrt 3.1 und die Definition 4.1 der renormierten stetigen Irrfahrt. Wir erkennen, dass sich folgende Elemente, im Kontext der Dualitätsformeln (siehe dazu Anmerkung 3.4), miteinander identifizieren lassen:

- $g_j = 1$ für alle $0 \leq j \leq l$, $g_j = 0$ sonst;
- $Y_j \mathrel{\hat{=}} \frac{1}{\sqrt{n}}\xi_j$, $j \in \mathbb{N}$;
- $S_k \mathrel{\hat{=}} X^{(n)}_{\frac{k}{n}} = \frac{1}{\sqrt{n}} \sum_{j=1}^{k} \xi_j$, für alle $0 \leq k \leq n$.

Der erste Schritt wird sein, die diskreten Differenzen auf der rechten Seite der mehrdimensionalen Dualitätsformel (3.11) durch Ableitungen der Funktion $\tilde{\phi}$ zu ersetzen. Nach Vergleich mit der diskreten Dualitätsformel (3.11) erhalten wir für unsere speziell gewählten Zeitpunkte die Gleichung

$$\begin{aligned}
&E(F(X^{(n)}) \int_0^1 g(s) dX^{(n)}_s) \\
&= E(\tilde{\phi}(X^{(n)}_{s_1}, \dots, X^{(n)}_{s_N} - X^{(n)}_{s_{N-1}}) X^{(n)}_t) \\
&= \tfrac{k_1 \wedge l}{\sqrt{n}}\ E((\tilde{\phi}(X^{(n)}_{s_1}, .., X^{(n)}_{s_N} - X^{(n)}_{s_{N-1}}) - \tilde{\phi}(X^{(n)}_{s_1 - \frac{1}{n}}, .., X^{(n)}_{s_N} - X^{(n)}_{s_{N-1}}))\xi_{k_1}) + \dots \\
&\quad + \tfrac{k_N \wedge l - (k_{N-1} - 1)\wedge l}{\sqrt{n}}\ E((\tilde{\phi}(X^{(n)}_{s_1}, \dots, X^{(n)}_{s_N} - X^{(n)}_{s_{N-1}}) - \tilde{\phi}(X^{(n)}_{s_1}, \dots, X^{(n)}_{s_N - \frac{1}{n}} - X^{(n)}_{s_{N-1}}))\xi_{k_N}).
\end{aligned}$$

Mit Hilfe der Taylorformel (5.1) werden wir jetzt die Differenzen der Funktion in den Erwartungswerten ersetzen. Die Taylorentwicklung gilt natürlich punktweise für die Funktion $\tilde{\phi}$, und damit auch fast sicher bezüglich des Wahrscheinlichkeitsmaßes von $X^{(n)}$. Für beliebiges $j \in \{1, \dots, N\}$ gilt dann

$$\begin{aligned}
&\tilde{\phi}(X^{(n)}_{s_1}, \dots, X^{(n)}_{s_j} - X^{(n)}_{s_{j-1}}, \dots, X^{(n)}_{s_N} - X^{(n)}_{s_{N-1}}) - \tilde{\phi}(X^{(n)}_{s_1}, \dots, X^{(n)}_{s_j - \frac{1}{n}} - X^{(n)}_{s_{j-1}}, \dots, X^{(n)}_{s_N} - X^{(n)}_{s_{N-1}}) \\
&= \partial_j \tilde{\phi}(X^{(n)}_{s_1}, \dots, X^{(n)}_{s_j - \frac{1}{n}} - X^{(n)}_{s_{j-1}}, \dots, X^{(n)}_{s_N} - X^{(n)}_{s_{N-1}}) \tfrac{1}{\sqrt{n}}\xi_{k_j} + O_1((\tfrac{1}{\sqrt{n}}\xi_{k_j})^2).
\end{aligned}$$

Das Restglied $O_1((\frac{1}{\sqrt{n}}\xi_{k_j})^2)$ der Taylorformel ist betragsmäßig beschränkt durch $K_1|\frac{1}{\sqrt{n}}\xi_{k_j}|^2$, wobei K_1 eine positive Konstante ist die von den zweiten partiellen Ableitungen der Funktion $\tilde{\phi}$ abhängt. Wir wählen K_1 so, dass die Beschränkung uniform für $j \in \{1, \ldots, N\}$ gilt. Wir wenden auf beiden Seiten den Erwartungswert an. Was wir dadurch bekommen ist zwar keine Konvergenz in L^1, aber nach Ausnutzung der Existenz des dritten Moments von ξ_1, eine Konvergenz der Erwartungswerte gegeneinander, denn

$$\sqrt{n}|E(O_1((\frac{1}{\sqrt{n}}\xi_{k_1})^2)\xi_{k_1})| \leq K_1 E(\frac{1}{\sqrt{n}}|\xi_{k_1}|^3) = K_1' n^{-\frac{1}{2}},$$

für ein K_1 dass groß genug gewählt wurde, und $K_1' := K_1 \cdot E(|\xi_1|^3)$. Das erste Mal, dass wir die Existenz des dritten Moments von ξ_1 benutzen. In der Schreibweise mit dem Landau-Symbol kommen wir also zu dem Ergebnis (unter Berücksichtigung, dass $E(\xi_1^2) = 1$), dass

$$\begin{aligned}
&\tfrac{k_1 \wedge l}{\sqrt{n}}\ E((\tilde{\phi}(X_{s_1}^{(n)}, .., X_{s_N}^{(n)} - X_{s_{N-1}}^{(n)}) - \tilde{\phi}(X_{s_1 - \frac{1}{n}}^{(n)}, .., X_{s_N}^{(n)} - X_{s_{N-1}}^{(n)}))\xi_{k_1}) + \ldots \\
&\quad + \tfrac{k_N \wedge l - (k_{N-1}-1)\wedge l}{\sqrt{n}}\ E((\tilde{\phi}(X_{s_1}^{(n)}, \ldots, X_{s_N}^{(n)} - X_{s_{N-1}}^{(n)}) - \tilde{\phi}(X_{s_1}^{(n)}, \ldots, X_{s_N - \frac{1}{n}}^{(n)} - X_{s_{N-1}}^{(n)}))\xi_{k_N}) \\
&= \tfrac{k_1 \wedge l}{n}\ E(\partial_1 \tilde{\phi}(X_{s_1 - \frac{1}{n}}^{(n)}, \ldots, X_{s_N}^{(n)} - X_{s_{N-1}}^{(n)})\xi_{k_1}^2) + \ldots \\
&\quad + \tfrac{k_N \wedge l - (k_{N-1}-1)\wedge l}{n}\ E(\partial_N \tilde{\phi}(X_{s_1}^{(n)}, \ldots, X_{s_N - \frac{1}{n}}^{(n)} - X_{s_{N-1}}^{(n)})\xi_{k_N}^2) + \sqrt{n}|E(O_1((\tfrac{1}{\sqrt{n}}\xi_{k_1})^2)\xi_{k_1})| \\
&= \tfrac{k_1 \wedge l}{n}\ E(\partial_1 \tilde{\phi}(X_{s_1 - \frac{1}{n}}^{(n)}, \ldots, X_{s_N}^{(n)} - X_{s_{N-1}}^{(n)})) + \ldots \\
&\quad + \tfrac{k_N \wedge l - (k_{N-1}-1)\wedge l}{n}\ E(\partial_N \tilde{\phi}(X_{s_1}^{(n)}, \ldots, X_{s_N - \frac{1}{n}}^{(n)} - X_{s_{N-1}}^{(n)})) + O_1(n^{-\frac{1}{2}}).
\end{aligned}$$

Im zweiten Schritt verwenden wir stillschweigend, dass

$$\left|\frac{k_j \wedge l - (k_{j-1} - 1) \wedge l}{n}\right| \leq 1, \quad \forall j \in \{1, \ldots, N\}.$$

Außerdem haben wir hier das einzige Mal im ganzen Beweis die Unabhängigkeit der $(\xi_j)_{j \geq 1}$ verwendet, um zu erhalten (hier für k_1), dass

$$E(\partial_1 \tilde{\phi}(X_{s_1 - \frac{1}{n}}^{(n)}, \ldots, X_{s_N}^{(n)} - X_{s_{N-1}}^{(n)})\xi_{k_1}^2) = E(\partial_1 \tilde{\phi}(X_{s_1 - \frac{1}{n}}^{(n)}, \ldots, X_{s_N}^{(n)} - X_{s_{N-1}}^{(n)})E(\xi_{k_1}^2). \tag{4.6}$$

An dieser Relation stören zwei Dinge. Erstens, dass im Faktor vor der Funktion nicht genau $s_j \wedge t - s_{j-1} \wedge t$ steht, wie es in der Dualitätsformel der Brownschen Bewegung steht. Wir wissen aber, dass die ersten Ableitungen von ϕ beschränkt sind, und damit gilt

$$\begin{aligned}
&\tfrac{k_j \wedge l - (k_{j-1}-1)\wedge l}{n} \partial_j \tilde{\phi}(X_{s_1}^{(n)}, \ldots, X_{s_j - \frac{1}{n}}^{(n)} - X_{s_{j-1}}^{(n)}, \ldots, X_{s_N}^{(n)} - X_{s_{N-1}}^{(n)}) \\
&= (s_j \wedge t - s_{j-1} \wedge t)\partial_j \tilde{\phi}(X_{s_1}^{(n)}, \ldots, X_{s_j - \frac{1}{n}}^{(n)} - X_{s_{j-1}}^{(n)}, \ldots, X_{s_N}^{(n)} - X_{s_{N-1}}^{(n)}) + O_2(\tfrac{1}{n}),
\end{aligned}$$

wobei das Restglied $O_2(\frac{1}{n})$ beschränkt ist durch $K_2|\frac{1}{n}|$. K_2 ist analog zu K_1 eine positive Konstante, die so gewählt werden kann, dass die Beschränkung uniform für $j \in \{1, \ldots, N\}$ gilt (da das Restglied keine Zufallsvariable ist, gilt diese Beschränkung auch unter dem Erwartungswert).

Zweitens stört, dass im Index des Arguments der partiellen Ableitung, jeweils an $j-$ter Stelle nicht der Zeitpunkt s_j auftaucht, sondern $s_j - 1/n$. Auch das können wir mit der Taylorformel korrigieren. Wir wenden diese dafür auf die ersten Ableitungen an:

$$\begin{aligned}
&\partial_j \tilde{\phi}(X^{(n)}_{s_1}, \ldots, X^{(n)}_{s_j - \frac{1}{n}} - X^{(n)}_{s_{j-1}}, \ldots, X^{(n)}_{s_N} - X^{(n)}_{s_{N-1}}) \\
&= \partial_j \tilde{\phi}(X^{(n)}_{s_1}, \ldots, X^{(n)}_{s_j} - X^{(n)}_{s_{j-1}}, \ldots, X^{(n)}_{s_N} - X^{(n)}_{s_{N-1}}) + O_3(\tfrac{1}{\sqrt{n}}\xi_{k_j}),
\end{aligned}$$

wobei das Restglied $O_3(\frac{1}{\sqrt{n}}\xi_{k_j})$ analog zu den anderen Rechnungen beschränkt werden kann durch $K_3|\frac{1}{\sqrt{n}}\xi_{k_j}|$ mit einer positiven Konstante K_3, die von den zweiten partiellen Ableitungen von $\tilde{\phi}$ abhängt und so gewählt werden kann, dass die Beschränkung uniform in $j \in \{1, \ldots, N\}$ gilt. Wir schreiben dann $K_3' := K_3 \cdot E(|\xi_1|)$ und erkennen sofort, dass $K_3' \leq K_3$.
Fassen wir die drei Approximationsschritte zusammen. Unter Zuhilfenahme der Dreiecksungleichung und addieren einiger Nullen erhalten wir

$$\begin{aligned}
&\Big| E(\tilde{\phi}(X^{(n)}_{s_1}, \ldots, X^{(n)}_{s_N} - X^{(n)}_{s_{N-1}}) X^{(n)}_t) \\
&\quad - \sum_{j=1}^{N} (s_j \wedge t - s_{j-1} \wedge t) E(\partial_j \tilde{\phi}(X^{(n)}_{s_1}, \ldots, X^{(n)}_{s_j} - X^{(n)}_{s_{j-1}}, \ldots, X^{(n)}_{s_N} - X^{(n)}_{s_{N-1}}))\Big| \\
&\leq K_1' n^{-\frac{1}{2}} + K_2 n^{-1} + K_3 n^{-\frac{1}{2}}.
\end{aligned} \tag{4.7}$$

Es gilt also die Gleichung (4.4), für festes $n \in \mathbb{N}$, und ein zylindrisches Pfadfunktional F, das nur von Zeitpunkten abhängt, die Vielfache von $1/n$ sind.
Bisher verwendeten wir $s_1, \ldots, s_N, t$ Vielfache von $1/n$. Jetzt verallgemeinern wir das Ganze auf beliebige (aber wie üblich geordnete) Zeitpunkte $s_1, \ldots, s_N, t$ aus dem Intervall $[0, 1]$. Die Verallgemeinerung läuft über eine Zurückführung auf Vielfache von $1/n$. Dazu definieren wir allgemein

$$t^{(n)} := \frac{\lfloor nt \rfloor}{n}, \qquad \Rightarrow |t^{(n)} - t| \leq \frac{1}{n}.$$

Die Frage ist jetzt, welchen Fehler ein Abrunden auf Vielfache von $1/n$ innerhalb der verwendeten Erwartungswerte verursacht. In jedem Falle muss dieses Abrunden geschehen, um die diskrete Dualitätsformel anwenden zu können. Wir verwenden jetzt die Variante $F(x) = \phi(x_{s_1}, \ldots, x_{s_N})$. Die folgenden Gleichungen gelten auch allgemein für Funktionen $\phi \in C_b^\infty$, also auch für beliebige Ableitungen von ϕ:

$$\begin{aligned}
&\phi(X^{(n)}_{s_1^{(n)}}, \ldots, X^{(n)}_{s_N^{(n)}}) - \phi(X^{(n)}_{s_1}, \ldots, X^{(n)}_{s_N}) \\
&= \phi(X^{(n)}_{s_1^{(n)}}, \ldots, X^{(n)}_{s_N^{(n)}}) - \phi(X^{(n)}_{s_1^{(n)}} + (s_1 - s_1^{(n)})\tfrac{1}{\sqrt{n}}\xi_{s_1^{(n)}+1}, \ldots, X^{(n)}_{s_N^{(n)}} + (s_N - s_N^{(n)})\tfrac{1}{\sqrt{n}}\xi_{s_N^{(n)}+1}) \\
&= O_4(\sum_{j=1}^{N} n^{-\frac{3}{2}}\xi_{k_j}).
\end{aligned}$$

Es handelt sich natürlich wieder um eine Anwendung der Taylorformel (5.1). Es kann also wieder eine positive Konstante K_4 gewählt werden, so dass das Restglied $O_4(\sum_{j=1}^{N} n^{-\frac{3}{2}}\xi_{k_j})$ beschränkt ist durch $K_4 \sum_{j=1}^{N} n^{-\frac{3}{2}}|\xi_{k_j}|$. Wir wählen K_4 gleich so groß, dass auch beim Abrunden der Zeiten in der ersten Ableitungen von ϕ die Restglieder mit K_4 kontrolliert werden können (das heißt, dass die gleiche Rechnung mit zum Beispiel $\partial_1\phi$ auch ein Restglied hat, das mit Hilfe der Konstanten K_4 kontrolliert werden kann). Wir setzen gleich $K_4' := NK_4E(|\xi_1|) \leq NK_4$. Die Größenordnung in n kommt zum Einen durch die Zeitdifferenzen zustande die durch $1/n$ majorisiert werden, zum Anderen durch den Faktor $1/\sqrt{n}$ vor den ξ_{k_j}. Zur Vereinfachung der Notation schreiben wir formal

$$\phi^{(n)}_{s_1,\ldots,s_N} := \phi(X^{(n)}_{s_1},\ldots,X^{(n)}_{s_N}),$$

Für die ersten Ableitungen, und damit für die rechte Seite der approximativen Dualitätsformel (4.4) erhalten wir

$$\begin{aligned}
&\sum_{j=1}^N \partial_j\phi^{(n)}_{s_1^{(n)},\ldots,s_N^{(n)}}t^{(n)} - \sum_{j=1}^N \partial_j\phi^{(n)}_{s_1,\ldots,s_N}t \\
&= \sum_{j=1}^N \partial_j\phi^{(n)}_{s_1^{(n)},\ldots,s_N^{(n)}}t^{(n)} - \sum_{j=1}^N \partial_j\phi^{(n)}_{s_1,\ldots,s_N}(t^{(n)} + (t - t^{(n)})) \\
&= \sum_{j=1}^N (\partial_j\phi^{(n)}_{s_1^{(n)},\ldots,s_N^{(n)}} - \partial_j\phi^{(n)}_{s_1,\ldots,s_N})t^{(n)} + \sum_{j=1}^N \phi^{(n)}_{s_1,\ldots,s_N}(t - t^{(n)}) \\
&= t^{(n)}\sum_{i=1}^N O_4(\sum_{j=1}^N n^{-\frac{3}{2}}\xi_{k_j}) + O_5(\tfrac{1}{n}).
\end{aligned}$$

Das Restglied $O_5(\frac{1}{n})$ kann wegen der Beschränktheit der ϕ mit einer Konstante K_5 kontrolliert werden durch $K_5\frac{1}{n}$. Das Restglied $t^{(n)}\sum_{j=1}^N O_4(\sum_{j=1}^N n^{-\frac{3}{2}}\xi_{k_j})$ kann, wie schon erwähnt, durch $t^{(n)}\sum_{i=1}^N K_4\sum_{j=1}^N |n^{-\frac{3}{2}}\xi_{k_j}|$ beschränkt werden. Man erkennt, dass das Restglied beschränkt ist durch NK_4'. Es ergibt sich die fast sichere Abschätzung

$$|\sum_{j=1}^N \partial_j\phi^{(n)}_{s_1^{(n)},\ldots,s_N^{(n)}}t^{(n)} - \sum_{j=1}^N \partial_j\phi^{(n)}_{s_1,\ldots,s_N}t| \leq NK_4\sum_{j=1}^N |n^{-\frac{3}{2}}\xi_{k_j}| + K_5\frac{1}{n},$$

und deshalb die Konvergenz in L^1 zu

$$\begin{aligned}
&E(|\sum_{j=1}^N \partial_j\phi^{(n)}_{s_1^{(n)},\ldots,s_N^{(n)}}t^{(n)} - \sum_{j=1}^N \partial_j\phi^{(n)}_{s_1,\ldots,s_N}t|) \\
&\leq NK_4n^{-\frac{3}{2}}\textstyle\sum_{j=1}^N E(|\xi_{k_j}|) + K_5\frac{1}{n} \\
&\leq NK_4'n^{-\frac{3}{2}} + K_5\tfrac{1}{n}.
\end{aligned}$$

Wichtig ist für uns wiederum nur die Konvergenz der Erwartungswerte gegeneinander, hier von der Ordnung $1/n$, dass heißt

$$E(\sum_{j=1}^N \partial_j\phi(X^{(n)}_{s_1^{(n)}},\ldots,X^{(n)}_{s_N^{(n)}})t^{(n)}) = E(\sum_{j=1}^N \partial_j\phi(X^{(n)}_{s_1},\ldots,X^{(n)}_{s_N})t) + O(n^{-1}), \tag{4.8}$$

Und das Restglied ist beschränkt durch $(NK_4'+K_5)n^{-1}$. Auf der anderen Seite der Dualitätsformel gehen wir ähnlich vor:

$$
\begin{aligned}
&\phi^{(n)}_{s_1^{(n)},\ldots,s_N^{(n)}} X^{(n)}_{t^{(n)}} - \phi^{(n)}_{s_1,\ldots,s_N} X^{(n)}_t \\
&= \phi^{(n)}_{s_1^{(n)},\ldots,s_N^{(n)}} X^{(n)}_{t^{(n)}} - \phi^{(n)}_{s_1,\ldots,s_N} (X^{(n)}_{t^{(n)}} + (t - t^{(n)}) \tfrac{1}{\sqrt{n}} \xi_{t^{(n)}+1}) \\
&= (\phi^{(n)}_{s_1^{(n)},\ldots,s_N^{(n)}} - \phi^{(n)}_{s_1,\ldots,s_N}) \cdot X^{(n)}_{t^{(n)}} + \phi^{(n)}_{s_1,\ldots,s_N} (t - t^{(n)}) \tfrac{1}{\sqrt{n}} \xi_{t^{(n)}+1} \\
&= O_4(\sum_{j=1}^{N} n^{-\frac{3}{2}} \xi_{k_j} X^{(n)}_{t^{(n)}}) + O_6(n^{-\frac{3}{2}} \xi_{t^{(n)}+1}).
\end{aligned}
$$

Das Restglied $O_4(\sum_{j=1}^N n^{-\frac{3}{2}} \xi_{k_j} X^{(n)}_{t^{(n)}})$ kann vom Betrag her kontrolliert werden durch $K_4 \sum_{j=1}^N n^{-\frac{3}{2}} |\xi_{k_j} X^{(n)}_{t^{(n)}}|$ mit der schon erwähnten Konstante K_4. Das Restglied $O_6(n^{-\frac{3}{2}} \xi_{t^{(n)}+1})$ kann wegen der Beschränktheit der ϕ kontrolliert werden, sagen wir mit einer positive Konstante K_6 durch $K_6 n^{-\frac{3}{2}} |\xi_{t^{(n)}+1}|$, also unter dem Erwartungswert durch $K_6' := K_6 n^{-\frac{3}{2}} E(|\xi_1|)$. Wieder erhalten wir hier eine Konvergenz in L^1:

$$
\begin{aligned}
&E(|\phi(X^{(n)}_{s_1}, \ldots, X^{(n)}_{s_N}) X^{(n)}_t - \phi(X^{(n)}_{s_1^{(n)}}, \ldots, X^{(n)}_{s_N^{(n)}}) X^{(n)}_{t^{(n)}}|) \\
&\leq K_4 \textstyle\sum_{j=1}^N n^{-\frac{3}{2}} E(|\xi_{k_j} X^{(n)}_{t^{(n)}}|) + K_6 n^{-\frac{3}{2}} E(|\xi_{t^{(n)}+1}|) \\
&\leq N K_4 n^{-\frac{3}{2}} \textstyle\sum_{i=1}^n E(|\xi_{k_j}|^2) + K_6' n^{-\frac{3}{2}} \\
&= N K_4 n^{-\frac{3}{2}} n + K_6' n^{-\frac{3}{2}} \\
&= N K_4 n^{-\frac{3}{2}} n + K_6' n^{-\frac{3}{2}} \\
&\leq K_7 n^{-\frac{1}{2}},
\end{aligned}
$$

für eine genügend groß gewählte Konstante K_7 die nicht von n abhängt. Wichtig ist wieder nur, dass die Erwartungswerte gegeneinander konvergieren mit Geschwindigkeit $1/\sqrt{n}$, dass heißt

$$
E(\phi(X^{(n)}_{s_1}, \ldots, X^{(n)}_{s_N}) X^{(n)}_t) = E(\phi(X^{(n)}_{s_1^{(n)}}, \ldots, X^{(n)}_{s_N^{(n)}}) X^{(n)}_{t^{(n)}}) + O(n^{-\frac{1}{2}}). \tag{4.9}
$$

Nun haben wir recht sorgfältig alle benötigten kleinen Veränderungen auf beiden Seiten der Dualitätsformel durchgeführt und immer die Schranke für den Approximationsfehler im Auge behalten. Der große Moment, in dem wir die erhaltenen Approximationen zusammenfügen, ist jetzt gekommen. Wir fassen nochmal unsere Voraussetzungen zusammen. Sei $F \in \mathcal{W}$, wobei $F(x) = \phi(x_{s_1}, \ldots, x_{s_N}) = \tilde{\phi}(x_{s_1}, x_{s_2} - x_{s_1}, \ldots, x_{s_N} - x_{s_{N-1}})$. Hier sollen $0 \leq s_1 \leq \cdots \leq s_N \leq 1$ beliebig sein, ebenso wie $t \in [0, 1]$. Weiterhin ist $g = 1_{[0,t[}$ Dann haben wir die approximative

Dualitätsformel

$$
\begin{aligned}
& E(F(X^{(n)})\int_0^1 g(s)dX_s^{(n)}) \\
= \quad & E(\phi(X_{s_1}^{(n)},\dots,X_{s_N}^{(n)})X_t^{(n)}) \\
\overset{4.9}{=} \quad & E(\phi(X_{s_1^{(n)}}^{(n)},\dots,X_{s_N^{(n)}}^{(n)})X_{t^{(n)}}^{(n)}) + O(n^{-\frac{1}{2}}) \\
= \quad & E(\tilde{\phi}(X_{s_1^{(n)}}^{(n)},\dots,X_{s_N^{(n)}}^{(n)} - X_{s_{N-1}^{(n)}}^{(n)})X_{t^{(n)}}^{(n)}) + O(n^{-\frac{1}{2}}) \\
\overset{4.7}{=} \quad & E((s_1^{(n)}\wedge t^{(n)})\partial_1\tilde{\phi}(X_{s_1^{(n)}}^{(n)},..,X_{s_N^{(n)}}^{(n)} - X_{s_{N-1}^{(n)}}^{(n)})) + \dots \\
& +E((s_N^{(n)}\wedge t^{(n)} - s_{N-1}^{(n)}\wedge t^{(n)})\partial_N\tilde{\phi}(X_{s_1^{(n)}}^{(n)},..,X_{s_N^{(n)}}^{(n)} - X_{s_{N-1}^{(n)}}^{(n)})) + O(n^{-\frac{1}{2}}) \\
= \quad & E((s_1^{(n)}\wedge t^{(n)})\partial_1\phi(X_{s_1^{(n)}}^{(n)},\dots,X_{s_N^{(n)}}^{(n)})) + \dots \\
& +E((s_N^{(n)}\wedge t^{(n)})\partial_N\phi(X_{s_1^{(n)}}^{(n)},\dots,X_{s_N^{(n)}}^{(n)})) + O(n^{-\frac{1}{2}}) \\
\overset{4.8}{=} \quad & E((s_1\wedge t)\partial_1\phi(X_{s_1}^{(n)},\dots,X_{s_N}^{(n)})) + \dots \\
& +E((s_N\wedge t)\partial_N\phi(X_{s_1}^{(n)},\dots,X_{s_N}^{(n)})) + O(n^{-\frac{1}{2}}) \\
= \quad & E(\int_0^1 1_{[0,t[}(s)\sum_{j=1}^{N}\partial_j\phi(X_{s_1}^{(n)},\dots,X_{s_N}^{(n)})1_{[0,s_j[}(s)ds) + O(n^{-\frac{1}{2}}) \\
= \quad & E(D_gF(X^{(n)})) + O(n^{-\frac{1}{2}}).
\end{aligned}
$$

Damit ist die Gleichung (4.4) für $F \in \mathcal{W}$ und $g = 1_{[0,t[}$ bewiesen. Die Verallgemeinerung auf allgemeine Treppenfunktionen g ist jetzt einfach eine Folgerung aus der approximativen Linearität der Gleichung (4.4). Sei zum Beispiel $g = \sum_{j=0}^{p} a_j 1_{[t_{j-1},t_j[}$ gegeben, dann kann man g auch schreiben als $\sum_{j=0}^{p}\lambda_j 1_{[0,t_j[}$ wobei die Koeffizienten λ_j linear von den a_j abhängen. Sollten wir die Gleichung (4.4) für Treppenfunktionen der Form $1_{[0,t[}$ bewiesen haben, ergibt sich für unser g

$$
\begin{aligned}
E(F(X^{(n)})\int_0^1 g(t)dX_t^{(n)}) &= \textstyle\sum_{j=0}^{p}\lambda_j E(F(X^{(n)})\int_0^1 1_{[0,t_j[}(t)dX_t^{(n)}) \\
&= \textstyle\sum_{j=0}^{p}\lambda_j(E(D_{1_{[0,t_j[}}F(X^{(n)})) + O(n^{-\frac{1}{2}})) \\
&= E(D_gF(X^{(n)})) + \textstyle\sum_{j=0}^{p}\lambda_j O(n^{-\frac{1}{2}}) \\
&= E(D_gF(X^{(n)})) + O(n^{-\frac{1}{2}}),
\end{aligned}
$$

da es sich bei dem Landau-Symbol um eine asymptotische Aussage einer Konvergenz gegen Null handelt, und die Multiplikation mit einer endlichen Konstante ($p\max\{\lambda_0,\dots\lambda_p\}$) nichts an dieser ändert.

Aus dem Beweis geht natürlich auch hervor, dass die Gleichung (4.5) für ein $K \in \mathbb{R}_+$ erfüllt ist. Außerdem hängt dieses K von der Treppenfunktion g, von der Funktion ϕ und deren Ableitungen, und vom dritten Moment von ξ_1 ab. □

Anmerkung 4.5 (zur Version mit verallgemeinerter symmetrischer Irrfahrt). Wir haben nur einmal verwendet, dass die renormierte stetige Irrfahrt auf einer symmetrischen Irrfahrt basiert, und nicht auf einer verallgemeinerten symmetrischen Irrfahrt. Um die Gleichung (4.6) zu erhalten verwenden wir im Prinzip, dass (Schreibweisen wie in Definition 3.2)

$$E(\xi_l^2|S_{k_1}-S_{k_0},\dots,S_{k_N}-S_{k_{N-1}})=E(\xi_l^2)=1,$$

für $l\notin\{k_{j-1}+1,\dots,k_j\}$ für alle $j\in\{1,\dots,N\}$. Für die verallgemeinerte symmetrische Irrfahrt hatten wir eine entsprechende Gleichung für das erste Moment vorausgesetzt:

$$E(\xi_l|S_{k_1}-S_{k_0},\dots,S_{k_N}-S_{k_{N-1}})=E(\xi_l)=0.$$

Würden wir diese Gleichung für das zweite Moment mit voraussetzen, könnten wir die approximative Dualitätsformel auch für eine Folge von Zufallsfunktionen $(X^{(n)})_{n\geq 1}$, die auf einer verallgemeinerten symmetrischen Irrfahrt basieren, beweisen. Das Problem wäre dann aber immer noch, wie in Anmerkung 4.3 schon erwähnt, der Beweis der Straffheit dieser Folge von Zufallsfunktionen.

Eine renormierte stetige Irrfahrt erfüllt also asymptotisch vom Grade $n^{-\frac{1}{2}}$ die Dualitätsformel des Wienermaß mit Anfangsbedingung $\mathbb{W}(X_0=0)=1$. Die Vermutung liegt natürlich nahe, dass sich nun zusammen mit der Charakterisierung des Wienermaßes durch die Dualitätsformel (Theorem 2.2) und mit der Straffheit der renormierten stetigen Irrfahrt (Theorem 4.1) das Theorem von Donsker beweisen lässt. Das dem wirklich so ist, wird im nächsten Theorem gezeigt.

Theorem 4.3 (Konvergenz der renormierten stetigen Irrfahrt). *Die renormierte stetige Irrfahrt konvergiert in Verteilung gegen eine Brownsche Bewegung mit Anfangsbedingung $X_0=0$.*

Beweis. Da die Sequenz $(\mathcal{L}(X^{(n)}))_{n\geq 1}$ straff ist, enthält jede aufsteigende Folge (n') natürlicher Zahlen eine Unterfolge (n'') natürlicher Zahlen, so dass $\mathcal{L}(X^{(n'')})\to Q$ (im Sinne der schwachen Konvergenz von Maßen) für $(n'')\to\infty$ und ein noch näher zu bestimmendes Wahrscheinlichkeitsmaß Q auf $C(0,1)$ (Vergleiche dazu [Bil3], Section 5). Wir nennen X ein Element aus einem Wahrscheinlichkeitsraum, dass die Verteilung Q hat. Sollten wir beweisen können, das Q die Dualitätsformel der Brownschen Bewegung für $F\in\mathbb{W}$ und g Treppenfunktion auf $[0,1]$ erfüllt, wissen wir nach Theorem 2.2, dass Q ein Wienermaß ist. Wir werden jetzt also im Prinzip zeigen, dass alle endlichdimensionalen Verteilungen der renormierten stetigen Irrfahrt gegen die endlichdimensionalen Verteilungen des Wienermaß mit Startpunkt Null konvergieren. Damit ist dann das Theorem von Donsker bewiesen.
Sei im weiteren Verlauf $F\in\mathbb{W}$ und g Treppenfunktion auf $[0,1]$. Da beide jeweils nur von endlich vielen Zeitpunkten abhängen, können wir die Funktionen so umschreiben, dass sie von gleichen Zeitpunkten abhängen. Sei

$$F(x)=\phi(x_{s_1},\dots,x_{s_M}),\ g=\sum_{j=1}^M g_j 1_{[t_{j-1},t_j[},\quad M\in\mathbb{N},\ t_0=0.$$

Da die Projektion $\pi_{s_1,\dots,s_M}:C(0,1)\to\mathbb{R}^M$, definiert durch

$$\pi_{s_1,\dots,s_M}:x\longmapsto(x_{s_1},\dots,x_{s_M})$$

stetig ist, haben wir die schwache Konvergenz des Bildmaßes (vergleiche [Bil3], Theorem 2.7):

$$(X_{s_1}^{(n'')}, \ldots, X_{s_M}^{(n'')}) \xrightarrow{\mathcal{L}} (X_{s_1}, \ldots, X_{s_M}).$$

Da die Ableitungen von ϕ stetig und beschränkt sind (siehe Anmerkung 2.1), folgt für die linke Seite der approximativen Dualitätsformel

$$\lim_{n'' \to \infty} E(D_g F(X^{(n'')})) = E(D_g F(X)),$$

nach Definition der schwachen Konvergenz, beziehungsweise dem Portemanteau-Theorem (siehe Anhang, Theorem 5.2). Es bleibt also noch zu beweisen, dass

$$E(F(X) \int_0^1 g(s) dX_s) \stackrel{?}{=} \lim_{n'' \to \infty} E(F(X^{(n'')}) \int_0^1 g(s) dX_s^{(n'')}),$$

denn da bekannt ist, dass ein Grenzwert existiert, ist bereits

$$\lim_{n'' \to \infty} E(F(X^{(n'')}) \int_0^1 g(s) dX_s^{(n'')}) = \lim_{n'' \to \infty} E(D_g F(X^{(n'')})) = E(D_g F(X))$$

gezeigt (das Restglied konvergiert wie bewiesen gegen Null).
Das Problem ist, dass $\phi(y_1, \ldots, y_M) \sum_{j=1}^M g_j(y_j - y_{j-1})$ keine beschränkte Funktion ist. Aber immerhin ist sie noch stetig. Wir werden die Funktion deshalb auf stetige Weise beschränken und versuchen den Fehler der dabei entsteht klein zu halten. Dabei helfen uns bestimmte Eigenschaften des Wahrscheinlichkeitsvektors $(X_{s_1}^{n''}, \ldots, X_{s_M}^{n''} - X_{s_{M-1}}^{n''})_{n''}$.
Wir wollen die Summenbestandteile $(y_j - y_{j-1})$ uniform in j und auf stetige Art beschränken. Dazu definieren wir die stetige Abschneidefunktion

$$h_K(y) := y 1_{\{|y| \leq K\}} + \operatorname{sign}(y) K^2 (K + \frac{1}{K} - |y|) 1_{\{K < |y| \leq K + \frac{1}{K}\}}, \quad y \in \mathbb{R}.$$

$K \in \mathbb{R}^+$ ist eine Konstante die noch speziell gewählt werden wird. In jedem Fall sieht man leicht, dass diese Funktion beschränkt und stetig ist. Es gilt

$$\begin{aligned} & |(y_j - y_{j-1}) - h_K(y_j - y_{j-1})| \\ & \leq |(y_j - y_{j-1}) - (y_j - y_{j-1}) 1_{\{|y_j - y_{j-1}| \leq K\}}|. \end{aligned}$$

Der Fehler, den wir in L^1 durch die Ersetzung der nichtbeschränkten Identität mit der beschränkten Funktion erhalten, kann also abgeschätzt werden vermöge

$$
\begin{aligned}
&E(|\phi(X_{s_1}^{(n'')},\dots,X_{s_M}^{(n'')})\sum_{j=1}^M g_j(X_{s_j}^{(n'')}-X_{s_{j-1}}^{(n'')})\\
&\quad -\phi(X_{s_1}^{(n'')},\dots,X_{s_M}^{(n'')})\textstyle\sum_{j=1}^M g_j\ h_K(X_{s_j}^{(n'')}-X_{s_{j-1}}^{(n'')})|)\\
&\le \|\phi\|_\infty \sum_{j=1}^M g_j E(|(X_{s_j}^{(n'')}-X_{s_{j-1}}^{(n'')})-(X_{s_j}^{(n'')}-X_{s_{j-1}})1_{\{|X_{s_j}^{(n'')}-X_{s_{j-1}}^{(n'')}|\le K\}}|)\\
&= C\sum_{j=1}^M g_j E(|X_{s_j}^{(n'')}-X_{s_{j-1}}^{(n'')})|1_{\{|X_{s_j}^{(n'')}-X_{s_{j-1}}^{(n'')}|>K\}})
\end{aligned}
$$

Das ist das Kriterium der uniformen Integrabilität. Wenn wir also zeigen, dass die Folgen $(X_{s_j}^{(n'')}-X_{s_{j-1}}^{(n'')})_{(n'')}$ uniform integrabel sind, können wir diese Größe uniform für alle $n''\in\mathbb{N}$ beschränken. Glücklicherweise gilt

$$
E((X_{s_j}^{(n'')}-X_{s_{j-1}}^{(n'')})^2)\le E((X_1^{(n'')})^2)=\frac{1}{n''}\sum_{j=1}^{n''}Var(\xi_1)=1.
$$

Die Folgen sind also alle in L^2 beschränkt, damit auch uniform integrabel. Sei also $\varepsilon>0$ beliebig. Wir wählen $K(\varepsilon)\in\mathbb{R}^+$ groß genug, so dass

$$
C\sum_{j=1}^M g_j E(|X_{s_j}^{(n'')}-X_{s_{j-1}}^{(n'')})|1_{\{|X_{s_j}^{(n'')}-X_{s_{j-1}}^{(n'')}|>K(\epsilon)\}})<\varepsilon,
$$

und können damit sagen

$$
\begin{aligned}
&E(\phi(X_{s_1}^{(n'')},\dots,X_{s_M}^{(n'')})\sum_{j=1}^M g_j(X_{s_j}^{(n'')}-X_{s_{j-1}}^{(n'')}))\\
&=E(\phi(X_{s_1}^{(n'')},\dots,X_{s_M}^{(n'')})\sum_{j=1}^M g_j\ h_{K(\varepsilon)}(X_{s_j}^{(n'')}-X_{s_{j-1}}^{(n'')})|)+O(\varepsilon),
\end{aligned}
$$

wobei das Restglied vom Betrage beschränkt ist durch ε, und nicht von n'' abhängt.
Mit der stetigen, beschränkten Funktion können wir dann das Portemanteau-Theorem anwenden:

$$
\begin{aligned}
&\lim_{n''\to\infty}E(\phi(X_{s_1}^{(n'')},\dots,X_{s_M}^{(n'')})\sum_{j=1}^M g_j(X_{s_j}^{(n'')}-X_{s_{j-1}}^{(n'')}))\\
=\;&\lim_{n''\to\infty}(E(\phi(X_{s_1}^{(n'')},\dots,X_{s_M}^{(n'')})\sum_{j=1}^M g_j\ h_{K(\varepsilon)}(X_{s_j}^{(n'')}-X_{s_{j-1}}^{(n'')}))+O(\varepsilon))\\
=\;&E(\phi(X_{s_1},\dots,X_{s_M})\sum_{j=1}^M g_j\ h_{K(\varepsilon)}(X_{s_j}-X_{s_{j-1}}))+O(\varepsilon)
\end{aligned}
$$

Die $(X_{s_j}-X_{s_{j-1}})$ sind nun quadratintegrable Zufallsvariablen, damit auch als einzelne Zufallsvariablen uniform integrabel. Denn die Funktion $f(y)=y^2$ ist stetig, damit gilt $\mathcal{L}((X_{s_j}^{(n'')}-X_{s_{j-1}}^{(n'')})^2)\to\mathcal{L}((X_{s_j}-X_{s_{j-1}})^2)$, und damit (vergleiche [Bil3], Theorem 3.4, als Theorem 5.7 im Anhang) ergibt sich dann

$$E((X_{s_j} - X_{s_{j-1}})^2) \leq \lim \inf_{n'' \to \infty} E((X_{s_j}^{(n'')} - X_{s_{j-1}}^{(n'')})^2) \leq 1.$$

Wir können also $K(\varepsilon)$ wenn notwendig etwas vergrößern, so dass gilt:

$$\begin{aligned}
&E(|\phi(X_{s_1}, \ldots, X_{s_M}) \sum_{j=1}^{M} g_j(h_{K(\varepsilon)}(X_{s_j} - X_{s_{j-1}}) - (X_{s_j} - X_{s_{j-1}}))|) \\
&\leq C \sum_{j=1}^{M} g_j E(|h_{K(\varepsilon)}(X_{s_j} - X_{s_{j-1}}) - (X_{s_j} - X_{s_{j-1}})|) \\
&\leq C \sum_{j=1}^{M} g_j E(|X_{s_j} - X_{s_{j-1}}| 1_{\{|X_{s_j} - X_{s_{j-1}}| > K(\varepsilon)\}}) \\
&\leq \varepsilon,
\end{aligned}$$

also

$$\begin{aligned}
&E(\phi(X_{s_1}, \ldots, X_{s_M}) \sum_{j=1}^{M} g_j \; h_{K(\varepsilon)}(X_{s_j} - X_{s_{j-1}})) \\
&= E(\phi(X_{s_1}, \ldots, X_{s_M}) \sum_{j=1}^{M} g_j(X_{s_j} - X_{s_{j-1}})) + O(\varepsilon),
\end{aligned}$$

und das Restglied ist betragsmäßig beschränkt durch ε. Zusammengenommen haben wir damit

$$\begin{aligned}
&\lim_{n'' \to \infty} E(\phi(X_{s_1}^{(n'')}, \ldots, X_{s_M}^{(n'')}) \sum_{j=1}^{M} g_j(X_{s_j}^{(n'')} - X_{s_{j-1}}^{(n'')})) \\
&= E(\phi(X_{s_1}, \ldots, X_{s_M}) \sum_{j=1}^{M} g_j(X_{s_j} - X_{s_{j-1}})) + O(\varepsilon).
\end{aligned}$$

Da $\varepsilon > 0$ beliebig gewählt war:

$$\begin{aligned}
&\lim_{n'' \to \infty} E(\phi(X_{s_1}^{(n'')}, \ldots, X_{s_M}^{(n'')}) \sum_{j=1}^{M} g_j(X_{s_j}^{(n'')} - X_{s_{j-1}}^{(n'')})) \\
&= E(\phi(X_{s_1}, \ldots, X_{s_M}) \sum_{j=1}^{M} g_j(X_{s_j} - X_{s_{j-1}})).
\end{aligned}$$

Für die Grenzverteilung Q gilt also mit jeder Funktion $F \in \mathcal{W}$ und jeder Treppenfunktion g in $[0, 1]$

$$E(F(X) \int_0^1 g(s) dX_s) = E(D_g F(X)).$$

Das ist natürlich die Dualitätsformel des Wienermaßes (2.3), Q ist also das Wienermaß mit Anfangsbedingung $X_0 = 0 \; f.s.$ (vergleiche [Bil3], Theorem 7.5). Damit ist das Konvergenzergebnis von Donsker bewiesen. □

Anmerkung 4.6 (Zur Beweistechnik der schwachen Konvergenz auf $C(0, 1)$). Der angeführte Beweis hält sich nicht exakt an die Formulierung des Theorems 5.5 aus dem Anhang. Während dort Straffheit und Konvergenz der endlichdimensionalen Verteilungen zur Konvergenz der Maßfolge führen, nutzen wir hier die Straffheit um konvergente Teilfolgen von Teilfolgen ausfindig zu machen, und identifizieren das Grenzmaß dann immer als ein Wienermaß. Das wir trotzdem die gewünschte schwache Konvergenz der Folge $(\mathcal{L}(X^{(n)}))_{n \geq 1}$ gegen das Wienermaß $\mathbb{W}_0$ erhalten, kann man sich

wie folgt klarmachen:
Sei S ein beliebiger metrischer Raum und $(x_n)_{n\geq 1}$ eine Folge von Elementen aus S. Zu jeder beliebigen Teilfolge $(x_{n_k})_{k\geq 1}$ von $(x_n)_{n\geq 1}$ gebe es eine Teilfolge $(x_{n_{k_l}})_{l\geq 1}$ von $(x_{n_k})_{k\geq 1}$ die gegen $y \in S$ konvergiert (das haben wir für unsere Folge von Wahrscheinlichkeitsmaßen bewiesen). Dann konvergiert auch die Folge $(x_n)_{n\geq 1}$ gegen y. Zum Beweis nehme man an, dies sei nicht der Fall, dann existiert eine Teilfolge $(x_{n_k})_{k\geq 1}$ und ein $\varepsilon > 0$ so dass $|x_{n_k} - y| > \varepsilon$ für alle $k \geq 1$ gilt. Dies ist aber ein Widerspruch zur Annahme, dass eine Teilfolge $(x_{n_{k_l}})_{l\geq 1}$ von $(x_{n_k})_{k\geq 1}$ existiert die gegen y konvergiert. Also konvergiert $(x_n)_{n\geq 1}$ gegen y.
Jetzt müssen wir uns nur noch daran erinnern, dass der Raum der Maße auf $C(0,1)$ metrisierbar ist (zum Beispiel mit der Prohorov-Metrik, siehe dazu Kapitel 13.2 in [Kle]), und wir erkennen, dass die schwache Konvergenz der Folge $(\mathcal{L}(X^{(n)}))_{n\geq 1}$ bewiesen wurde.

Anmerkung 4.7 (Zum Beweis des Konvergenzergebnis). Der oben geführte Beweis des Theorems 4.3 ist aufgrund seiner Anschaulichkeit recht ausführlich gehalten. Das Ganze funktioniert aber auch viel einfacher und kürzer, indem man einfach nur Theorem 3.5 aus [Bil3] (siehe Theorem 5.8 im Anhang) anwendet. Damit erhält man die gleiche Möglichkeit die Limiten und Erwartungswerte zu vertauschen. Dem Theorem als Zauberkasten verwendet liegen allerdings die gleichen hier im Beweis vorgeführten Berechnungen zu Grunde. Der Vollständigkeit halber hier noch der alternative Beweis.

Alternativer Beweis von Theorem 4.3. Wir kürzen einfach den Beweisteil ab, der mit Hilfe einer stetigen Abschneidefunktion beweist, dass

$$\begin{aligned}&\lim_{n''\to\infty} E(\phi(X_{s_1}^{(n'')}, \dots, X_{s_M}^{(n'')}) \sum_{j=1}^{M} g_j(X_{s_j}^{(n'')} - X_{s_{j-1}}^{(n'')}))\\ &= E(\phi(X_{s_1}, \dots, X_{s_M}) \sum_{j=1}^{M} g_j(X_{s_j} - X_{s_{j-1}})).\end{aligned}$$

Wie schon im ersten Beweis erwähnt ist der Wahrscheinlichkeitsvektor $(X_{s_1}^{(n'')}, \dots, X_{s_M}^{(n'')})$ uniform integrierbar, denn es gilt

$$E((X_{s_j}^{(n'')} - X_{s_{j-1}}^{(n'')})^2) \leq E((X_1^{(n'')})^2) = \frac{1}{n''} \sum_{j=1}^{n''} Var(\xi_1) = 1.$$

Damit kann man leicht schlussfolgern, dass auch die Zufallsvariablen

$$\phi(X_{s_1}^{(n'')}, \dots, X_{s_M}^{(n'')}) g_j(X_{s_j}^{(n'')} - X_{s_{j-1}}^{(n'')})$$

für alle $j \in \{1, \dots, M\}$ uniform integrierbar sind, denn ϕ ist nach Voraussetzung beschränkt. Wir können also einfach Theorem 5.8 aus dem Anhang anwenden, und erhalten die Konvergenz

$$E(\phi(X_{s_1}^{(n'')}, \dots, X_{s_M}^{(n'')}) g_j(X_{s_j}^{(n'')} - X_{s_{j-1}}^{(n'')})) \to E(\phi(X_{s_1}, \dots, X_{s_M}) g_j(X_{s_j} - X_{s_{j-1}})).$$

Damit ist das Theorem 4.3 bewiesen. □

Anmerkung 4.8 (zur Version mit verallgemeinerter symmetrischer Irrfahrt). Der Beweis des Theorems 4.3 verwendet nicht mehr explizit die speziellen Eigenschaften der symmetrischen Irrfahrt. Natürlich benötigt man aber die vorher bewiesenen Ergebnisse der Straffheit und der approximativen Dualitätsformel.

5 Anhang

Theorem 5.1 (Taylor-Formel im $\mathbb{R}^N$)**.** *Für $\phi \in C_b^\infty$, $y, h \in \mathbb{R}^N$ gilt die Taylor-Formel*

$$\begin{aligned}\phi(y+h) &= \phi(y) + \tfrac{1}{1!}\sum_{j=1}^{N} \partial_j \phi(y) h_j + \tfrac{1}{2!}\sum_{j_1,j_2=1}^{N} \partial_{j_1 j_2}\phi(y) h_{j_1}\cdot h_{j_2} + \dots \\ &\quad \dots + \tfrac{1}{p!}\sum_{j_1,\dots,j_p=1}^{N} \partial_{j_1\dots j_p}\phi(y) h_{j_1}\cdots h_{j_p} + R(y,h),\end{aligned}$$

mit dem Restglied

$$R(y,h) = \int_0^1 \frac{(1-s)^p}{p!} \sum_{j_1,\dots j_{p+1}=1}^{N} \partial_{j_1\dots j_{p+1}}\phi(y+sh) h_{j_1}\cdots h_{j_{p+1}} ds.$$

Es gilt die Abschätzung

$$|R(y,h)| \le \frac{|h|^{p+1}}{(p+1)!} \sup_{0\le s\le 1} \sqrt{\sum_{j_1,\dots j_{p+1}=1}^{N} |\partial_{j_1\dots j_{p+1}}\phi(y+sh)|^2},$$

da $\phi \in C_b^\infty$ ist, kann eine Konstante $K \in \mathbb{R}^+$ so gewählt werden, dass

$$|R(y,h)| \le K \sum_{j_1,\dots j_{p+1}=1}^{N} |h_{j_1}\cdots h_{j_{p+1}}|. \tag{5.1}$$

Beweis. Vergleiche dazu die Version in [Bae] (Kapitel 5, Satz 5.6) die als Vorlage für die Formulierung dieses Theorems gedient hat. Eine einfachere Version mit Beweis befindet sich in [For] (Kapitel I, §7, Satz 2). Eine allgemeinere Version des Theorems in Bezug auf die Funktionenklasse findet sich in [Wer] (Satz III.5.5). □

Theorem 5.2 (Portemanteau Theorem, Alexandrow)**.** *Seien $X, X_1, X_2, \dots$ Zufallselementen in einen metrischen Raum S, dann sind folgende Bedingungen äquivalent:*

i) Die Folge $(X_n)_{n\ge1}$ konvergiert schwach gegen X, geschrieben $X_n \xrightarrow{\mathcal{L}} X$ (zur Erinnerung: für alle $\phi \in C_b(S,\mathbb{R})$ muss $E(\phi(X_n)) \to E(\phi(X))$ gelten);

ii) $\liminf P(X_n \in G) \ge P(X \in G)$ *für alle offenen Mengen $G \subset S$;*

iii) $\limsup P(X_n \in F) \le P(X \in F)$ *für alle geschlossenen Mengen $F \subset S$;*

iv) $P(X_n \in B) \to P(X \in B)$ für alle Borel-Mengen $B \subset S$ mit $P(X \in \partial B) = 0$.

Beweis. Das Theorem wurde aus [Kal] übernommen (Theorem 3.25), ein Beweis findet sich an eben dieser Stelle. □

Theorem 5.3 (Differentiation unter dem Integralzeichen)**.** *Es seien $I \subset \mathbb{R}$ ein Intervall, $\lambda_0 \in I$, und $F : I \times \Omega \to \mathbb{R}$ (Ω meßbarer Raum) habe folgende Eigenschaften:*

i) Für alle $\lambda \in I$ gilt $F(\lambda, .) \in L^1$.

ii) Die partielle Ableitung $\frac{\partial F}{\partial \lambda}(\lambda_0, x)$ *existiert für alle* $x \in \Omega$.

iii) Es existiert positive Funktion $G \in L^1$, *so dass für alle* $\lambda \in I$ *und* $x \in \Omega$ *gilt:*

$$\left| \frac{\partial F}{\partial \lambda}(\lambda, x) \right| \leq G(x).$$

Dann ist die Funktion $\hat{\rho} : I \to \mathbb{R}$,

$$\hat{\rho}(\lambda) := \int_\Omega F(\lambda, x) dP(x) = E(F(\lambda, X)), \quad \lambda \in I$$

im Punkt λ_0 *differenzierbar,* $\frac{\partial F}{\partial \lambda}(\lambda_0, .)$ *ist integrierbar, und es gilt*

$$\frac{d}{d\lambda}\hat{\rho}(\lambda_0) = \int_\Omega \frac{\partial F}{\partial \lambda}(\lambda_0, x) dP(x) = E(\frac{\partial F}{\partial \lambda}(\lambda_0, X)).$$

Beweis. Satz 5.7, Kapitel IV. aus [Els] wurde hier einfach übernommen. Beweissuchende seien dorthin verwiesen. □

Theorem 5.4 (Itô-Formel)**.** *Sei* $f : \mathbb{R}^n \to \mathbb{R}$ *zweimal stetig differenzierbar, sei* $X_t = (X_t^1, \ldots, X_t^n)$ *ein stetiges Semimartingal in* $\mathbb{R}^n$ *(das heißt, jedes* X^i *ist ein stetiges Semimartingal). Dann:*

$$f(X_t) = f(X_0) + \sum_{i=1}^n \int_0^t \partial_i f(X_s) dX_s^i + \frac{1}{2} \sum_{i,j=1}^n \int_0^t \partial_{ij} f(X_s) d[X^i, X^j]_s.$$

Beweis. Das Theorem wurde aus [RW2] übernommen (Chapter IV, Theorem 32.8). Ein Beweis befindet sich an ebendieser Stelle. □

Theorem 5.5 (Konvergenz von Wahrscheinlichkeitsmaßen auf $C(0,1)$)**.** *Seien* $P, P_1, P_2, \ldots$ *Wahrscheinlichkeitsmaße auf* $C(0,1)$ *(ausgestattet mit der Topologie der gleichmäßigen Konvergenz). Wenn die endlichdimensionalen Verteilungen von* P_n *schwach gegen die von* P *konvergieren, und wenn* $(P_n)_{n \geq 1}$ *straff ist, dann konvergiert* P_n *schwach gegen* P.

Beweis. Das Theorem ist aus [Bil3] entnommen (Chapter 2, Theorem 7.1), Erläuterungen und ein Beweis finden sich dort. □

Theorem 5.6 (Straffheit von Wahrscheinlichkeitsmaßen)**.** *Die Folge* $(P_n)_{n \geq 1}$ *von Wahrscheinlichkeitsmaßen auf* $C(0,1)$ *ist genau dann straff, wenn folgende zwei Bedingungen gelten:*

i) Für alle $\eta > 0$ *existiert ein* $a \in \mathbb{R}^+$ *und ein* $n_0 \in \mathbb{N}$, *so dass*

$$P_n(|x(0)| \geq a) \leq \eta, \quad n \geq n_0.$$

ii) Für alle $\varepsilon, \eta > 0$ *existiert ein* $\delta \in (0,1)$ *und ein* $n_0 \in \mathbb{N}$, *so dass*

$$P_n(\omega_\delta(x) \geq \varepsilon) \leq \eta, \quad n \geq n_0.$$

Beweis. Das Theorem ist aus [Bil3] entnommen (Chapter 2, Theorem 7.3), Erläuterungen und ein Beweis finden sich dort. □

Theorem 5.7 (obere Schranke des Erwartungswerts). *Seien* $X, X_1, X_2, \ldots$ *Zufallselementen in einen metrischen Raum* S. *Wenn* X_n *schwach gegen* X *konvergiert, so gilt* $E(|X|) \leq \liminf E(|X_n|)$.

Beweis. Das Theorem ist aus [Bil3] entnommen (Chapter 1, Theorem 3.4), ein Beweis findet sich dort. □

Theorem 5.8 (Konvergenz des Erwartungswerts). *Seien* $X, X_1, X_2, \ldots$ *Zufallselementen in einen metrischen Raum* S. *Ist* $(X_n)_{n \geq 1}$ *uniform integrierbar und konvergiert* X_n *schwach gegen* X, *dann ist* X *integrierbar, und* $E(X_n) \to E(X)$.

Beweis. Das Theorem ist aus [Bil3] entnommen (Chapter 1, Theorem 3.5), ein Beweis findet sich dort. □

Literatur

[Bae] G. Bärwolff (2005): Höhere Mathematik für Naturwissenschaftler und Ingenieure, erste Auflage, Spektrum, München.

[Bau] H. Bauer (1991): Wahrscheinlichkeitstheorie, vierte überarbeitete Auflage, Berlin, De Gruyter.

[Bil1] P. Billingsley (1961): The Lindeberg-Lévy theorem for martingales, Proc. Amer. Math. Soc. 12, 1961, 788-792.

[Bil2] P. Billingsley (1968): Convergence of Probability Measures, first edition, New York, Wiley.

[Bil3] P. Billingsley (1999): Convergence of Probability Measures, second edition, New York, Wiley.

[Bis] J.-M. Bismut (1981): Martingales, the Malliavin Calculus and Hypoellipticity under general Hörmander's Conditions, Z. Wahr. Ver. Geb., vol. 56, 1981, p. 469-505.

[Don] M.D. Donsker (1951): An Invariance Principle for Certain Probability Limit Theorems, Mem. Am. Math. Soc. 6, 1951, p.1-12, Rhode Island.

[Els] J. Elstrodt (2005): Maß- und Integrationstheorie, vierte Auflage, Berlin, Springer.

[For] O. Forster (2005): Differentialrechnung im $\mathbb{R}^n$, gewöhnliche Differentialgleichungen, sechste Auflage, Vieweg, Wiesbaden.

[Geo] H.-O. Georgii: Gibbs Measures and Phase Transitions, first edition, Berlin, de Gruyter

[Hau] U. Hausmann (1978): Funktionals of Ito Processes as Stochastic Integrals, SIAM J. Control and Optimisation, vol. 16, 1978, p. 252-269.

[Ibr] I.A. Ibragimov (1963): A central limit theorem for a class of dependent random variables, Theory Probab. Appl. 8, 1963, 83-94 (transl. by R.A. Silverman).

[Kal] O. Kallenberg (1997): Foundation of Modern Probability, first edition, Springer, New-York.

[KS] I. Karatzas, S.E. Shreve (1999): Brownian Motion and Stochastic Calculus, second edition, New York, Springer.

[Kle] A. Klenke (2006): Wahrscheinlichkeitstheorie, Springer, Berlin.

[Nah] B. Nahapetian (1995): Billingsley-Ibragimov theorem for martingale-difference random fields and its applications to some models of classical statistical physics, C. R. Acad. Sci. Paris, t. 320, Série I, p. 1539-1544, 1995.

[NP] B. Nahapetian, A.Petrosian (1992): Martingale-difference Gibbs random fields and central limit theorem, Ann. Acad. Ski. Fenn., Ser. A I Math., 17, 1992, pp. 105-110.

[Nel] E. Nelson (2001): Dynamical Theories of Brownian Motion, second edition, Princeton University Press.

[Pro] P.E. Protter (2005): Stochastic Integration and Differential Equations, second edition, Version 2.1, Berlin, Springer.

[RY] D. Revuz, M. Yor (1998): Continous Martingales and Brownian Motion, second edition, New York, Springer.

[RZ1] S. Roelly, H. Zessin (1991): Une caractérisation des diffusions pas le calcul des variations stochastiques, C.R. Acad. Sci. Paris, t.313, Série I, p.309-312.

[RZ2] S. Roelly, H. Zessin (1993): Une caractérisation des mesures de Gibbs sur $C(0,1)^{\mathbb{Z}^d}$ par le calcul des variations stochastiques, Ann. Inst. Henri Poincaré, Vol. 29, n°3, p.327-338.

[RW1] L.C.G. Rogers, D. Williams (1994): Diffusions, Markov Processes, and Martingales, Volume One: Foundations, second edition, New York, Wiley.

[RW2] L.C.G. Rogers, D. Williams (1994): Diffusions, Markov Processes, and Martingales, Volume Two: Itô Calculus, first edition, New York, Wiley.

[Wer] D. Werner (2007): Funktionalanalysis, sechste Auflage, Springer, Berlin.

Schlusserklärung: Hiermit erkläre ich, dass ich die vorliegende Diplomarbeit selbstständig verfasst und keine anderen als die angegebenen Quellen und Hilfsmittel benutzt habe.

Ort,Datum:

Unterschrift: