Fachbereich Informatik und Mathematik
ISMI - Institut für Stochastik
& Mathematische Informatik

Analysis II

WS 2011

H. Dinges

15. Februar 2011

Inhaltsverzeichnis

1 Komplexe Differenzierbarkeit

In der Algebra behandelt man die Polynome als formale Rechengrößen, die man linear kombinieren, miteinander multiplizieren und (formal!) differenzieren kann. Man arbeitet mit den Koeffizientenfolgen. Die Polynome mit komplexen Koeffizienten können alternativ als komplexwertige Funktionen einer komplexen Variablen z verstanden werden. Die Addition und die Multiplikation werden dabei zu punktweisen Operationen; die Differentiation wird zu einer infinitesimalen 'lokalen' Operation. Wenn man unter dem lokalen Aspekt verallgemeinern will, dann gelangt man nicht zu den formalen Potenzreihen, (von welchen im Abschnitt 1.5 die Rede war), sondern zu den (in einem offenen Kreis $K_R \subseteq \mathbb{C}$) konvergenten Potenzreihen, von welchen schon einmal kurz im Abschnitt 6.3 bei der lokal gleichmäßigen Konvergenz die Rede war.

Wir werden uns in diesem Abschnitt mit etwas allgemeineren Typen von Funktionen auseinandersetzen, bei welchen wie bei den Potenzreihen die 'komplexe Differentiation' eine zentrale Rolle spielt. Dies sind die Funktionen auf einem Gebiet $\subseteq \mathbb{C}$, die sich lokal durch Potenzreihen darstellen lassen. Sie heissen die analytischen Funktionen.

Wenn man bei den analytischen Funktionen auf einem Gebiet $\subseteq \mathbb{C}$ das Konzept der Stammfunktion verfolgt, dann kommen topologische Fragen in den Blick. Man muss sich mit Kurvenintegralen befassen. Wir werden im zweiten Unterabschnitt etwas weiter ausholen. Wir diskutieren rektifizierbare Kurven in einem vollständigen metrischen Raum, sowie die Integration von 1-Formen entlang solcher Kurven. Dabei werden auch die Begriffe 'zusammenhängend' und 'bogenzusammenhängend' zu diskutieren sein.

Im dritten Unterabschnitt behandeln wir dann die analytischen Funktionen ausgehend vom geometrisch konzipierten Begriff der Holomorphie.

Wir werfen dann nochmals einen Blick auf die sog. elementaren Funktionen, die jetzt aber als Funktionen einer komplexen Veränderlichen verstanden werden. Und wir beschliessen unseren Einblick in die 'Funktionentheorie' mit Rechenaufgaben zur Verlagerung von Integrationswegen.

1.1 Die analytischen Funktionen auf einer Kreisscheibe

Definition 1.1 (Der Funktionenraum $\mathcal{A}_R$).
Eine komplexwertige Funktion $f(z)$, die im offenen Kreis $K_R = \{z : |z| < R\} \subset \mathbb{C}$ durch eine konvergente Potenzreihe dargestellt wird, nennt man eine in K_R analytische Funktion: $f(z) = \sum_0^\infty a_n \cdot z^n$. Die Menge dieser Funktionen wird mit $\mathcal{A}_R$ bezeichnet.

Zur Erinnerung: Koeffizientenfolgen, Identitätssatz, kompakte Konvergenz.
Wir haben oben in 6.3 bewiesen, dass die in K_R analytischen Funktionen genau durch die Folgen $(a_0, a_1, a_2, \ldots)$ mit $\limsup |a_n|^{1/n} \leq 1/R$ dargestellt werden.

Die Koeffizientenfolge $(a_n)_n$ (und damit die Funktion) ist eindeutig bestimmt durch die Werte der Funktion in einer beliebigen abzählbar unendlichen Menge, die sich im Nullpunkt häuft.

Man sagt von einer Folge $(f^{(N)})_N$ in $\mathcal{A}_R$, dass sie gegen $f \in \mathcal{A}_R$ (im Sinne der kompakten Konvergenz) konvergiert, wenn gilt

$$\sup\{\, |f^{(N)}(z) - f(z)| : \quad |z| \leq r \,\} \xrightarrow{N\to\infty} 0 \quad \text{für alle } r < R$$

Man spricht auch von lokal gleichmäßiger Konvergenz. Wir haben gesehen, dass dieser Konvergenzbegriff durch eine translationsinvariante Metrik $d_R(\cdot,\cdot)$ beschrieben werden kann. Wir haben behauptet, aber nicht bewiesen, dass der metrische Raum $\big(\mathcal{A}_R, d_R(\cdot,\cdot)\big)$ vollständig ist.

Wir werden allgemeiner die Funktionen auf einem Gebiet $G \subseteq \mathbb{C}$ betrachten, die sich lokal durch Potenzreihen darstellen lassen; man nennt sie die auf G analytischen Funktionen. Der Raum $\mathcal{A}_G$ dieser Funktionen ist vollständig in dem folgenden Sinn: Wenn eine Folge auf G analytischer Funktionen lokal gleichmäßig eine Cauchy-Folge ist, dann ist die Limesfunktion eine auf G analytische Funktion. Die Folge der abgeleiteten Funktionen konvergiert lokal gleichmäßg gegen die Ableitung der Grenzfunktion. Um das zu beweisen holen wir weiter aus.

Satz 1.1.1 (Multiplikation in $\mathcal{A}_R$ und formale Differentiation)**.**
Die formale Differentiation liefert eine surjektive lineare Abbildung des Funktionenraums $\mathcal{A}_R$ auf sich: $\quad D: \quad \mathcal{A}_R \ni f(z) = \sum_0^\infty a_n \cdot z^n \longmapsto \sum_1^\infty n \cdot a_n \cdot z^{n-1} = f'(z) \in \mathcal{A}_R$
Es sind genau die konstanten Funktionen, die auf die Nullfunktion abgebildet werden.
$\mathcal{A}_R$ ist eine Funktionenalgebra, auf welcher die Differentiation als Derivation wirkt, d. h.

$$D(f \cdot g) = f \cdot D(g) + g \cdot D(f).$$

Beweis. *Es gilt* $\limsup |n \cdot a_n|^{1/n} = \limsup |a_n|^{1/n}$*. Die Potenzreihe der derivierten Funktion hat also denselben Konvergenzradius wie die Funktion selber.*
Zu $g(z) = \sum_0^\infty b_n \cdot z^n \in \mathcal{A}_R$ *ist* $S(g) = \sum_0^\infty \frac{1}{n+1} b_n \cdot z^{n+1}$ *eine in* K_R *analytische Funktion mit der Ableitung* g*. Die Derivation ist also auf* $\mathcal{A}_R$ *surjektiv.*
Es seien $(a_n)_n$ *und* $(b_n)_n$ *Folgen mit* $\limsup |a_n| \leq 1/R$, $\limsup |b_n| \leq 1/R$*; und es sei* $c_n = \sum_{k=0}^n a_k b_{n-k}$ *für* $n = 0, 1, 2, \ldots$*. Zu jedem* $r < R$ *existiert eine Konstante* C*, sodass* $|a_n| \leq C \cdot (1/r)^n$, $|b_n| \leq C \cdot (1/r)^n$ *für alle* n*. Es gilt dann*

$$|c_n| \leq (n+1) \cdot C^2 \cdot (1/r)^k \cdot (1/r)^{n-k} \quad \textit{und daher} \quad \limsup |c_n|^{1/n} \leq 1/r.$$

Da diese Abschätzung für alle $r < R$ *gilt, hat die Potenzreihe* $\sum c_n z^n$ *einen Konvergenzradius* $\geq R$*.*
Die Produktregel für die formale Differentiation als Operation auf den Koeffizientenfolgen haben wir im Abschnitt über formale Potenzreihen bewiesen.

Bemerke: Die formale Derivation k-mal auf $f(z) = \sum_0^\infty a_n \cdot z^n$ angewandt liefert

$$D^k(f) = \sum_{n=k}^\infty [n]_k a_n z^{n-k} = \sum_{m=0}^\infty [m+k]_k a_{m+k} z^m,$$

mit der Notation $[n]_k = n(n-1)(n-2)\cdots(n-k+1)$ $\quad$ (‘n untere Faktorielle k’)

Satz 1.1.2 (Differenzenquotienten).
Es sei $f(z) = \sum a_n z^n \in \mathcal{A}_R$ *und* $f'(z) = \sum n a_n z^{n-1}$.
Für jedes $r < R$ *gibt es dann eine Konstante* $C(r)$*, so daß*

$$\left|\frac{f(w) - f(z)}{w - z} - f'(z)\right| \leq C(r) \cdot |w - z| \qquad \text{für alle } z, w \text{ mit } |z| \leq r, |w| \leq r.$$

Die Konstante kann abgeschätzt werden: $C(r) \leq \frac{1}{2} \sum n(n-1)|a_n| r^{n-2}$.

Beweis.
$$\frac{w^n - z^n}{w - z} = w^{n-1} + w^{n-2}z + \ldots + wz^{n-2} + z^{n-1}$$
$$\frac{w^n - z^n}{w - z} - nz^{n-1} = \left(w^{n-1} - z^{n-1}\right) + \left(w^{n-2} - z^{n-2}\right)z + + \cdots + (w - z)z^{n-2}$$

Man kann aus jedem Summanden den Faktor $w - z$ *herausziehen; der Multiplikator besteht dann insgesamt aus* $\frac{1}{2}n(n-1)$ *Termen mit Betrag* $\leq r^{n-2}$.

Bemerke: Das Resultat besagt insbesondere, das sich die (zunächst formal auf der Grundlage der Koeffizienten definierte) Ableitung f' als lokal gleichmäßiger Limes von 'Differenzenquotienten' ergibt.

Satz 1.1.3 (Umzentrieren).
Sei R *der Konvergenzradius von* $\sum a_n z^n$*; und sei* $|z_0| < R$*. Es existieren dann* $b_0, b_1, b_2, \ldots$*, so daß*

$$\sum b_k (z - z_0)^k = f(z) = \sum a_n z^n \quad \text{für } |z - z_0| < R - |z_0|.$$

Bevor wir uns dem Beweis zuwenden, bemerken wir: Die Koeffizienten b_n sind eindeutig bestimmt. Eine Konsequenz des Satzes ist eine Verschärfung des sog. Identitätssatzes für konvergente Potenzreihen: Wenn eine Funktion in K_R durch eine Potenzreihe dargestellt wird und auf einer Folge verschwindet, die sich im Inneren von K_R häuft, dann ist sie die Nullfunktion, und das bedeutet, dass alle Koeffizienten verschwinden.

Die Koeffizienten b_k kann man nicht wie bei den Polynomen (Horner–Schema) durch einen finiten Algorithmus bestimmen. Wir zeigen, dass man sie als Werte von Reihen oder als Werte von Integralen gewinnen kann. Zunächst die Reihendarstellung:

Satz 1.1.4 (Reihendarstellung der Koeffizienten).
Es sei $\sum a_n z^n$ *eine Potenzreihe mit dem Konvergenzradius* R*. Ist* z_0 *ein Punkt im Konvergenzkreis und* $b_k = \frac{1}{k!} f^{(k)}(z_0) = \sum_m \frac{[m+k]_k}{k!} a_{m+k} z_0^m$ *dann gilt*

$$\sum_k b_k (z - z_0)^k = \sum_n a_n z^n \qquad \text{für} \qquad |z - z_0| < R - |z_0|.$$

Beweis.
$$\frac{[m+k]_k}{k!} = \frac{(m+k)!}{m!k!} = \binom{n}{k} \quad \text{mit } n = m + k.$$
$$\sum_k \frac{1}{k!} f^{(k)}(z_0)(z - z_0)^k = \sum_{k,m} \frac{(m+k)!}{m!\,k!} a_{m+k} z_0^m (z - z_0)^k =$$
$$= \sum_n a_n \sum_k \binom{n}{k} z_0^{n-k} (z - z_0)^k = \sum a_n \big(z_0 + (z - z_0)\big)^n = f(z)$$

nach dem Binomialsatz $(a + b)^n = \sum \binom{n}{k} a^k b^{n-k}$.

In konkreten Beispielen wird man die Koeffizienten der umzentrierten Potenzreihe nicht immer mit einem allgemeinen Verfahren zu bestimmen suchen; manchmal führen einfachere adhoc-Rechnungen zum Ziel.

Beispiel 1.1.1. (Die geometrische Reihe)
Die Funktion $f(z) = \frac{1}{1-z}$, definiert für $z \in \mathbb{C} \setminus \{1\}$, ist in $\{|z| < 1\}$ durch eine Potenzreihe mit Zentrum 0 darstellbar. In $\{z : |z - z_0| < |1 - z_0|\}$ ist sie in eine Potenzreihe mit dem Zentrum z_0 darstellbar.

$$f(z) \;=\; \frac{1}{1-z} = \sum z^n = 1 + z + z^2 + z^3 + \ldots \quad \text{für } |z| < 1$$

$$\frac{1}{1-z} = \frac{1}{(1-z_0)-(z-z_0)} = \frac{1}{1-z_0} \cdot \frac{1}{1-\frac{z-z_0}{1-z_0}} = \sum b_k (z-z_0)^k$$

$$\text{für } |z - z_0| < |1 - z_0|, \quad \text{wobei} \qquad b_k = (1-z_0)^{-(k+1)} \;=\; \frac{1}{k!} f^{(k)}(z_0).$$

Beispiel 1.1.2. (Der Hauptwert des Logarithmus $\ln(1+z)$ im Kreis $|z| < 1$)
Wir wissen schon: Wenn $|a_n|$ beschränkt ist, dann ist der Konvergenzradius von $\sum a_n z^n$ mindestens 1. Betrachten wir z.B.

$$f(z) \;=\; z - \tfrac{1}{2}z^2 + \tfrac{1}{3}z^3 - \tfrac{1}{4}z^4 + \ldots \qquad \text{für } \; |z| < 1$$

Die Ableitung kennen wir bereits $\qquad f'(z) = 1 - z + z^2 - z^3 + \ldots = \frac{1}{1+z}$.
Es sollte daher nicht verwundern, daß unsere Potenzreihe die Funktion $\ln(1+z)$ in $|z| < 1$ darstellt. Genauer: Sie stellt dort den sog. Hauptwert des Logarithmus dar. Für $|z_0| < 1$ und $|z - z_0| < 1 - |z_0|$ gilt nach unserem Satz $f(z) = \sum_0^\infty c_k \cdot (z - z_0)^k$ mit gewissen Koeffizienten c_k. Es gilt für diese z

$$\begin{aligned} f(z) = \ln\big(1 - z_0 + (z - z_0)\big) &= \ln(1 - z_0) + \ln\big(1 + \tfrac{z-z_0}{1-z_0}\big) \\ &= \ln(1 - z_0) - \sum_{k=1}^{\infty} (z_0 - 1)^{-k} \cdot (z - z_0)^k. \end{aligned}$$

Hinweis: Wir bemerken, dass die Potenzreihe mit dem Zentrum z_0 nicht nur im Kreis $\{|z - z_0| < 1 - |z_0|\}$ konvergiert, sondern sogar im größeren Kreis $\{|z - z_0| < |1 - z_0|\}$. Auch für die z im größeren Kreis stellt die Reihe den Hauptwert des Logarithmus dar; denn keiner dieser Kreise trifft die negative reelle Achse.

Beispiel 1.1.3. (Newtons Binomialreihe)
Bei der Diskussion der Gamma-Funktion haben wir das Wachstumsverhalten der Binomialkoeffizienten bestimmt: $\qquad \binom{\alpha}{k} = O\big(k^{-\alpha-1}\big) \quad$ für $k \to \infty$, α fest.
Das zeigt, dass Newtons Binomialreihe im Einheitskreis konvergiert. Wir zeigen, dass man die dort dargestellte Funktion auch anders beschreiben kann, nämlich

$$\sum_0^\infty \binom{\alpha}{n} = \big(1 + z\big)^\alpha \qquad \text{für } |z| < 1.$$

Der Sinn der allgemeinen Potenzen im Komplexen muss jetzt noch kommentiert werden. Beginnen wir mit der n-ten Wurzel für $n \in \{2, 3, \dots\}$: Für jede komplexe Zahl $w \neq 0$ gibt es genau n verschiedene n-te Wurzeln: sie gehen auseinander hervor durch die Multiplikation mit den n-ten Einheitswurzeln $e^{2\pi i/n \cdot k}$. Es existiert genau ein $z = |z|e^{i\phi}$ mit $z^n = w$, $-\pi/n < \phi \leq \pi/n$; diese Zahl nennt man den Hauptwert der n-ten Wurzel. Es gilt offenbar $z = e^{1/n \cdot \ln w}$, wo $\ln(\cdot)$ der Hauptwert des Logarithmus ist. Diese Definition benützt man für alle α. Wenn α keine ganze Zahl ist, dann ist also die Funktion $w^\alpha = \varepsilon^{\alpha \cdot \ln w}$ unstetig in den Punkten der negativen reellen Achse. Wenn man den Definitionsbereich auf die geschlitzte Zahlenebene G einschränkt, erhält man eine analytische Funktion. Die oben aufgestellte Behauptung lautet in präzisen Worten: Die Funktion $(1+z)^\alpha$ wird im Einheitskreis durch Newtons Binomialreihe dargestellt. Die Ableitung ist $\alpha \cdot (1+z)^{\alpha-1}$.

Potenzreihen und einseitige Fourier–Reihen

Sei $\sum_0^\infty a_n z^n = f(z)$ für $|z| < R$; und sei $r < R$. Die Einschränkung von $f(\cdot)$ auf die Kreislinie $\left\{z = re^{it} : t \in \mathbb{R}/2\pi\right\}$ liefert dann eine 2π–periodische Funktion

$$h(t) = f(re^{it}) = \sum_n c_n e^{int} \text{ für } t \in \mathbb{R} \qquad \text{mit} \quad c_n = \begin{cases} 0 & \text{für } n < 0 \\ a_n r^n & \text{für } n \geq 0, \end{cases}$$

Hier können wir jetzt die Formel für die Koeffizienten einer (hier absolut summablen) trigonometrischen Reihe anwenden. $\big(|c_n| \leq \text{const } q^n$ mit $q < 1.\big)$

$$\frac{1}{n!} f^{(n)}(0) \;=\; a_n = \frac{1}{r^n} \cdot \frac{1}{2\pi} \int_{-\pi}^{+\pi} e^{-int} h(t)\, dt \;=\; \frac{1}{2\pi} \int_{-\pi}^{+\pi} \frac{1}{(re^{it})^{n+1}} f(re^{it}) re^{it} dt$$

Die aufwendigere Formel am Ende dient der Verallgemeinerungen der kommenden Verallgemeinerungen. Wir werden dann auch allgemeinere ‘Umlaufsintegrale’ in den Blick nehmen.

Mit ähnlichen Integralen gewinnen wir jetzt auch noch die Ableitungen von $f(\cdot)$ in beliebigen z_0 mit Inneren des Konvergenzkreises.

Satz 1.1.5 (Cauchy’s Integralformel für Kreise)**.**

Sei $\sum a_n z^n = f(z)$ *für* $|z| < R$ *und* $r < R$. *Für* $|z_0| < r$ *und* $n = 0, 1, 2, \dots$ *gilt dann*

$$\frac{1}{n!} f^{(n)}(z_0) = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \frac{1}{(re^{it} - z_0)^{n+1}} f(re^{it}) re^{it} dt = \frac{1}{2\pi i} \oint \frac{1}{(\zeta - z_0)^{n+1}} f(\zeta)\; d\zeta.$$

Beweis. *Wir beweisen die erste Gleichung, auf die zweite kommen wir zurück, wenn wir die hier benützte Notation eingeführt haben*

1. *Wir beweisen die Integralformel zunächst nur für Polynome. Es genügt, den Fall* $r = 1$ *zu studieren: Es ist* $|z_0| < 1$ *und integriert wird über den Einheitskreis.*

   Für das Monom $p(z) = (z - z_0)^m \quad (m = 0, 1, 2, \ldots)$ *liefert die linke Seite als Werte der Ableitungen im Punkt* z_0

$$\frac{1}{n!} p^{(n)}(z_0) = \begin{cases} 1 & \textit{für } n = m \\ 0 & \textit{für } n \in \{0, 1, \ldots, m-1\} \cup \{m+1, m+2, \ldots\} \end{cases}$$

   Der Integrand auf der anderen Seite ist

   $\frac{1}{(e^{it} - z_0)^{n+1}} (e^{it} - z_0)^m e^{it}\, dt$

2. *Für* $n = m$ *liefert die Integration wegen* $\frac{1}{(1 - z_0 e^{-it})} = 1 + z_0 e^{-it} + (z_0 e^{-it})^2 + \ldots$

$$\frac{1}{2\pi} \int_{-\pi}^{+\pi} \frac{1}{(e^{it} - z_0)} e^{it} dt = \frac{1}{2\pi} \int_{-\pi}^{+\pi} \frac{1}{(1 - z_0 e^{-it})} dt = 1,$$

   Beachte: Hier wurde die Annahme $|z_0| < 1$ *wesentlich benützt.*

3. *für* $n \in \{0, 1, \ldots, m-1\}$ *ist der Binomialsatz auf den Integranden anwendbar*

$$(e^{it} - z_0)^{\ell} e^{it} \quad \textit{mit} \quad \ell = 0, 1, \ldots, m$$

   Integration jedes Terms über die volle Periode liefert 0.

4. *für* $n \in \{m+1, m+2, \ldots\}$ *liefert die binomische Reihe das gewünschte Resultat* $= 0$.

5. *Daraus ergibt sich also in der Tat für das Monom* $p(z) = (z - z_0)^m$

$$\frac{1}{(1 - z_0 e^{-it})^{\ell}} e^{-it\ell} e^{it} = \sum_{k=\ell-1}^{\infty} b_k e^{-itk}, \quad \ell \geq 2.$$

   Wegen der Linearität haben wir die Integralformel für beliebige Polynome .

6. *Auch für allgemeine* $\sum_0^{\infty} a_n z^n = f(z)$ *genügt es, den Fall* $r = 1$ *zu betrachten.*

   Fixiere n. *Gleichmäßig für* $z_0 \in \{z : |z| \leq 1 - \delta\}$ *gilt*

$$\frac{1}{n!} f^{(n)}(z_0) = \lim_{N \to \infty} \sum_0^N d_k z_0^k \qquad \textit{mit gewissen } d_k.$$

   Wenn wir N *genügend groß wählen, ist der Rest im Betrag* $< \varepsilon$ *gleichmäßig in* z_0.

7. $f(e^{it}) = \lim\limits_{N\to\infty} \sum_0^N a_k e^{ikt} = \lim\limits_N s_N(e^{it})$.

Wenn wir N genügend groß wählen, dann ist der Fehler $< \varepsilon\delta^{n+1}$; daher gilt für den Integranden

$$\left|\frac{1}{(e^{it}-z_0)^{n+1}}\right| \left|f(e^{it}) - s_N(e^{it})\right| < \varepsilon.$$

Für die Größen, deren Gleichheit die Integralformel behauptet, kommt es also (bis auf ε) nur auf eine passende Partialsumme an. Für Polynome ist die Integralformel aber bereits bewiesen.

Satz 1.1.6 (Die Vollständigkeit von $\mathcal{A}_R$)**.**
Es sei $(f_N(\cdot))_N$ eine Folge von Funktionen, die in $\{z : |z| < R\}$ in eine Potenzreihe entwickelbar sind. Wenn für jedes $r < R$ die Einschränkungen auf die Kreislinie $\{\zeta : |\zeta| = r\}$ ($r < R$) gleichmäßig konvergieren, dann konvergiert die Funktionenfolge f_N gleichmäßig auf $\{z : |z| < r\}$ für jedes $r < R$. Die Grenzfunktion $\widetilde{f}$ wird auf $\{z : |z| < R\}$ durch eine Potenzreihe dargestellt. Ihre Koeffizienten sind die Limiten der Koeffizienten der f_N.

Beweis. *Für $r < R$ betrachte die 2π–periodischen Funktionen* $\quad h_N(t) = f_N(re^{it}) \quad t \in \mathbb{R}$.
Da die $(h_N(t))_N$ nach Voraussetzung gleichmäßig konvergieren, konvergieren ebenfalls gleichmäßig auf $\{z : |z| \leq (1-\delta)r\}$ die Integrale

$$f_N(z) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \frac{1}{re^{it}-z} h_N(t) re^{it} dt \qquad f_N'(z) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \frac{1}{(re^{it}-z)^2} h_N(t) re^{it} dt$$

Wenn $\widetilde{h}(t) = \lim h_N(t)$, *dann gilt* $\qquad \widetilde{f}(z) = \frac{1}{2\pi}\int_{-\pi}^{+\pi} \frac{1}{re^{it}-z}\widetilde{h}(t) re^{it} dt \quad$ *auf* $\{z : |z| < r\}$.

Definition 1.2 (Die auf G analytischen Funktionen)**.**
Eine auf einer offenen Menge $G \subseteq \mathbb{C}$ definierte komplexwertige Funktion f heisst analytisch auf G, wenn es zu jedem $\tilde{z} \in G$ einen Kreis $\{z : |z - \tilde{z}| < \tilde{r}\}$ gibt, in welchem sie durch eine Potenzreihe dargestellt wird

$$f(z) - f(\tilde{z}) \;=\; \sum_{k=1}^{\infty} b_k \cdot (z - \tilde{z})^k \qquad \text{für} \quad |z - \tilde{z}| < \tilde{r}.$$

Anmerkung: Wenn man die Sprech- und Denkweisen des 18. Jahrhunderts im Blick hat, dann wird man lieber von einer ‘lokal analytischen’ Funktion sprechen. Euler dachte bekanntlich bei den Funktionen an ‘analytische Ausdrücke’, die aus der ‘veränderlichen Größe’ und ’eigentlichen Zahlgrößen’ zusammengesetzt sind. In diese Kategorie fallen für ihn sehr wohl die Potenzreihen, aber wohl kaum irgendwelche mit disparaten Formeln zu präsentierende mathematischen Objekte.

Beispiel 1.1.4.
Die Funktion $f(z) = \frac{1}{z}$ ist eine analytische Funktion auf den dem Gebiete $G = \mathbb{C} \setminus \{0\}$, auf der sog. (im Nullpunkt) punktierten Zahlenebene.

Sind $p(z)$ und $q(z)$ Polynome, und ist $N = \{z_1, \ldots, z_m\}$ die Nullstellenmenge von $q(z)$, so ist die gebrochenrationale Funktion $f(z) = \frac{p(z)}{q(z)}$ eine analytische Funktion auf $\mathbb{C} \setminus N$.

Der Hauptwert der n-ten Wurzel $\sqrt[n]{z}$, eingeschränkt auf die entlang der negativen reellen Achse geschlitzten Zahlenebene $\mathbb{C} \setminus (-\mathbb{R}_+)$ ist dort eine analytische Funktion. Ebenso der Hauptwert des Logarithmus $L(z)$

$$\mathbb{C} \setminus (-\mathbb{R}_+) \ni z = |z| \cdot e^{i\phi} \longmapsto L(z) = \ln |z| + i\phi \quad \text{mit } |\phi| < \pi \ ,$$

sowie die allgemeinen Potenzen $f_\alpha(z) = e^{\alpha \cdot L(z)}$.

Die analytischen Funktionen als Abbildungen
Die Menge $\mathcal{A}_G$ der auf G analytischen Funktionen (mit den auf G punktweisen Operationen) ist eine Funktionenalgebra mit einer ausgezeichneten Derivation D. Wir wollen jetzt noch weitere Konstruktionen in den Blick nehmen.

Zur Erinnerung: Bei den formalen Potenzreihen hat man die Operation des Einsetzens: Wenn $B(\cdot)$ eine formale Potenzreihe ist, dann ist $B(A(\cdot))$ für jedes $A(\cdot)$ mit $A(0) = 0$ eine wohldefinierte formale Potenzreihe mit

$$D\big(B(A(\cdot))\big) = (DB)(A(\cdot)) \cdot (DA)(\cdot) \qquad \text{(Kettenregel).}$$

Der n-te Koeffizient c_n in der Reihe $C(\cdot) = B(A(\cdot))$ ergibt sich mit endlichen Operationen aus den Koeffizienten $a_1, \ldots, a_n$; $b_1, \ldots, b_n$.
Eine interessante Frage ist auch die nach einer 'Einsetzungsinversen' oder 'Inversen' einer formalen Potenzreihe $A(\cdot)$ mit $A(0) = 0$, $A'(0) \neq 0$.

Hier bei den analytischen Funktionen müssen wir uns um die Definitionsbereiche kümmern und mit Umzentrierungen argumentieren.

Satz 1.1.7. *Ist $f(\cdot)$ analytisch auf G und $g(\cdot)$ analytisch auf dem Bildbereich $f(G)$, so ist die zusammengesetzte Funktion $h(\cdot) = g(f(\cdot))$ analytisch auf G. Eine analytische Funktion ist lokal invertierbar in den Punkten $\tilde{z}$ mit $f'(\tilde{z})$. Genauer: Es existiert eine Umgebung $\tilde{U} \ni \tilde{z}$ und eine im Bildbereich $f(\tilde{U})$ analytische Funktion h, sodass $h(f(z)) = z$ für $z \in \tilde{U}$.*

Beweis. *Sei $\tilde{z} \in G$ und $\tilde{w} = f(\tilde{z})$,* $\quad g(\tilde{w} + w) = g(\tilde{w}) + \sum b_n w^n$ *für $|w| < r$.*
Wenn $\tilde{r}$ so klein ist, dass für $|\tilde{z}| < \tilde{r}$ der Betrag von $f(\tilde{z} + z) - f(\tilde{z}) = \sum_1^\infty a_k z^k$ kleiner ist als r, dann haben wir

$$g(f(\tilde{z}+z)) = g(\tilde{w}+(f(\tilde{z}+z)-f(\tilde{z})) = g(\tilde{w})+\sum b_n\big(f(\tilde{z}+z)-f(\tilde{z})\big)^n = g(\tilde{w})+\sum_1 c_m z^m.$$

Bei der Bestimmung einer lokalen Umkehrung kann man sich auf das vereinfachte Newton-Verfahren stützen. Die lokale Umkehrung ergibt sich für die w in einer Umgebung von

$\tilde{w} = f(\tilde{z})$ *als Limes der Funktionenfolge*

$$h^{(0)}(\cdot) = \tilde{z}, \qquad h^{(n+1)}(\cdot) = h^{(n)}(\cdot) - \frac{1}{f'(\tilde{z})} \cdot \big(f(h^{(n)}(\cdot)) - (\cdot)\big).$$

Die Funktionenfolge konvergiert gleichmäßig in einer Umgebung; die Koeffizienten in der Potenzreihendarstellung sind die, die man auch durch Koeffizientenvergleich erhält.

Beispiel 1.1.5.
Es sei f analytisch auf G. Für jedes Polynom p ist $p(f)$ auf G analytisch; ausserdem ist e^f auf G analytisch. Wenn f auf G den Wert 0 nicht annimmt, dann ist auch die reziproke Funktion $\frac{1}{f}$ auf G analytisch.

Interessant ist die Frage, wie man für eine Funktion f, die den Wert 0 nicht annimmt, den Logarithmus definieren kann. Gibt es eine Funktion h mit $e^h = f$? Wenn es eine solche Funktion h gibt, dann erfüllt auch $h + 2\pi i \cdot k$ mit $k \in \mathbb{Z}$ das Verlangte.

Eine entsprechende Frage kann man z. B. auch für die Arcussinusfunktion stellen. Wir werden uns damit befassen.

Ausblick auf 'Stammfunktionen' im Komplexen

Zu jeder polynomialen Funktion $f(z)$ auf $\mathbb{C}$ gibt es eine Stammfunktion; diese (bis auf eine additive Konstante eindeutig bestimmte) polynomiale Funktion $F(z)$ kann man wie im Reellen als unbestimmtes Integral gewinnen. Genauer: Man kann $F(\cdot)$ durch sogenannte komplexe Kurvenintegrale gewinnen:

$$F(z) - F(z_0) = \int_{\mathfrak{C}} f(\zeta)\, d\zeta, \qquad \text{wo } \mathfrak{C} \text{ eine Kurve von } z_0 \text{ nach } z \text{ ist}$$

Solche Kurvenintegrale studiert man nun auch für solche 'holomorphe Formen' $f(z)\,dz$, wo f lokal durch Potenzreihen dargestellt wird. Das Kurvenintegral betrachtet als Funktion des Endpunkts z hat da ebenfalls die Eigenschaften einer Stammfunktion: $F'(z) = f(z)$. Die Sache ist aber insofern komplizierter, als der Integrationsweg manchmal nicht unwichtig ist. Beispielsweise liefert die Integration der Form $\frac{1}{z}\,dz$ (vom Punkt 1 aus) nicht notwendigerweise den Hauptwert des Logarithmus. (Wir erinnern uns. dass die Exponentialfunktion im Komplexen nicht injektiv ist und daher nicht wirklich eine Umkehrfunktion besitzt.)

Wir werden uns mit der Kurvenintegration holomorpher Formen befassen, nachdem wir Kurven und Kurvenintegrale in einem allgemeineren Rahmen diskutiert haben.

1.2 Kurven und Integration entlang rektifizierbarer Kurven

Sprechweise (Äquivalente parametrisierte Kurven)**.**
Eine parametrisierte Kurve in der Menge S ist eine Abbildung $\gamma(\cdot)$ eines kompakten Intervalls $[a,\ b]$ in S. $\gamma(a)$ heisst der Startpunkt oder Ausgangspunkt oder Anfangspunkt der Kurve, $\gamma(b)$ heisst der Endpunkt.

Wir werden uns hauptsächlich für rektifizierbare Kurven in einem metrischen Raum $\big(S, d(\cdot,\cdot)\big)$ interesssieren. Zunächst befassen wir uns allgemeiner mit stetigen Kurven in einem topologischen Raum $\big(S, \mathfrak{U})\big)$.

Zwei parametrisierte Kurven $\{\gamma'(t) : t \in [a',\ b']\}$ und $\{\gamma''(s) : s \in [a'',\ b'']\}$ heissen äquivalent, wenn es eine stetige, striktwachsende und surjektive Abbildung $T(\cdot)$ gibt,

$$T : \quad [a'',\ b''] \longrightarrow [a',\ b'] \qquad \text{sodass} \quad \forall\, s \in [a'',\ b''] : \quad \gamma''(s) = \gamma'(T(s)).$$

Wir haben es tatsächlich mit einer Äquivalenzrelation zu tun. Die Umkehrabbildung zu $T(\cdot)$ ist eine stetige striktwachsende surjektive Abbildung $S(\cdot)$ mit $\gamma'(t) = \gamma''(S(t))$ für alle t. Wenn $S(\cdot)$ und $T(\cdot)$ stetig striktwachsend surjektiv sind mit passenden Definitionsbereichen

$$[a',\ b'] \xrightarrow{S} [a'',\ b''] \xrightarrow{T} [a''',\ b'''],$$

dann ist auch die zusammengesetzte Abbildung $T(S(\cdot))$ stetig striktwachsend surjektiv.

Sprechweise.
Eine Kurve in der Menge S ist eine Äquivalenzklasse $\mathfrak{C}$ von parametrisierten Kurven. Die Repräsentanten $\gamma(\cdot)$ heissen die Parametrisierungen der Kurve $\mathfrak{C}$.

Eine Kurve $\mathfrak{C}$ in einem Hausdorff-Raum $\big(S, \mathfrak{U})\big)$ heisst eine stetige Kurve, wenn die Repräsentanten stetige Abbildungen sind. Mit $\alpha(\mathfrak{C})$ bezeichnen wir den Startpunkt, mit $\beta(\mathfrak{C})$ den Endpunkt der Kurve. Von einer stetigen Kurve mit $\alpha(\mathfrak{C}) = P, \beta(\mathfrak{C}) = Q$ sagen wir, dass sie von P nach Q führt: man nennt sie auch ein stetige Verbindungskurve für die Punkte P und Q.

Das Bild des kompakten Intervalls unter $\gamma(\cdot)$ als Punktmenge in S heisst die Spur der Kurve; wir notieren $\sigma(\mathfrak{C})$.

Wir werden einige rein topologische Eigenschaften der Spuren von Kurven benötigen:

Satz 1.2.1. *Die Spur einer stetigen Kurve ist eine kompakte zusammenhängende Menge.*

Beweis. *Die Kompaktheit ergibt sich aus dem allgemeinen Satz: Kompakte Mengen werden durch stetige Abbildungen φ auf kompakte Mengen abgebildet. Das sieht man folgendermaßen: Sei K kompakt und $\{V_\alpha : \alpha \in I\}$ eine offene Überdeckung der Bildmenge $\varphi(K)$. Die vollen Urbilder $U_\alpha = \varphi^{-1}(V_\alpha)$ bilden dann eine offene Überdeckung von K. Es existiert eine endliche Teilüberdeckung $\{U_{\alpha_1}, \ldots, U_{\alpha_N}\}$ und die entsprechenden V_{α_k} sind eine endliche Teilüberdeckung von $\varphi(K)$.*

Eine Teilmenge B *eines topologischen Raums heisst, wenn sie nur auf triviale Weise mit zwei in der Relativtopologie offenen disjunkten Mengen überdeckt werden kann, d. h. wenn*

$$U_0, U_1 \text{ offen} \quad U_0 \cup U_1 \supseteq B; \quad U_0 \cap U_1 \cap B = \emptyset \quad \Longrightarrow \quad U_0 \cap B = \emptyset \text{ oder } U_1 \cap B = \emptyset.$$

Die Intervalle sind die zusammenhängenden Teilmengen der reellen Achse (ohne Beweis!).

Durch stetige Abbildungen werden zusammenhängende Mengen in zusammenhängende Mengen abgebildet. Das sieht man folgendermaßen: Das Bild $\varphi(B)$ *der zusammenhängenden Menge* B *sei durch die disjunkte offenen Mengen* V_0, V_1 *überdeckt. Die vollen Urbilder* $U_0 \cap B$, $U_1 \cap B$ *sind offen in der Relativtopologie auf* B*; ausserdem sind sie disjunkt und überdecken* B*. Eines der beiden trifft* B *nicht, die entsprechende Menge* V *trifft* $\varphi(B)$ *nicht.*

Wir bemerken, dass sich der Zwischenwertsatz von Bolzano daraus ergibt, dass die Intervalle die zusammenhängenden Teilmengen von $\mathbb{R}$ sind.

Satz 1.2.2. *Eine Teilmenge* B *eines Hausdorff-Raums ist jedenfalls dann zusammenhängend, wenn zu jedem Punktepaar* P, Q *eine stetige Kurve existiert, welche von* P *nach* Q *führt. (Man nennt sie in diesem Fall bogenzusammenhängend)*

Beweis. *Sei* B *von disjunkten offenen Mengen* U_0, U_1 *überdeckt. Wir führen die Annahme, dass beide nichtleer sind, zum Widerspruch. Seien* $P_0 \in U_0, P_1 \in U_1$ *und sei* $\mathfrak{C}$ *eine stetige Verbindungskurve. Für diese zusammenhängende Menge ist aber der Durchschnitt mit einer der Mengen* U_k *leer.*

Der folgende Satz zeigt, dass die Begriffe 'zusammenhängend' und ˚ bogenzusammenhängend' für die offenen Mengen in einem $\mathbb{R}^n$zusammenfallen. Die offenen zusammenhängenden Menge nennt man **Gebiete**.

Satz 1.2.3. *Es sei* $\big(S, \mathfrak{U})\big)$ *ein HRaB, welcher eine Basis besitzt, dessen Elemente* B *die Eigenschaft besitzen, dass man in* B *jeden Punkt mit jedem Punkt verbinden kann. Jede zusammenhängende offene Menge ist dann bogenzusammenhängend.*

Beweis. *Sei* G *eine offene Menge, die im Sinne der topologischen Definition zusammenhängend ist; und es sei* U *die Menge der Punkte, die man von einem gegebenen Punkt* P_0 *aus mit einer stetigen Kurve erreichen kann.* $V = G \setminus U$ *ist die Menge derjenigen Punkte, die man nicht von* P_0 *aus erreichen kann. Beide Mengen sind offen, wie wir sofort sehen werden.* V *ist also die leere Menge. Wenn* $P \in U$*, dann liegt* P *in einem* B*; jeden Punkt von* B *kann man von* P *aus, also auch von* P_0 *aus erreichen.* U *ist offen. Sei* $P \in V$*. Jedes* B*, welches* P *enthält, ist disjunkt zu* U*, weil man andernfalls* P *von* P_0 *aus erreichen könnte.* V *ist also offen.*

Wenn man zwei Punkte in einem Gebiet überhaupt mit einer stetigen Kurve verbinden kann, dann kann das i. Allg. auf sehr verschiedene Weisen geschehen. Unter Umständen will man aber von gewissen Kurven von P nach Q sagen, dass sie nicht sehr verschieden sind; man führt eine Äquivalenzrelation ein. Die wichtigste Äquivalenzrelation ist die sog. Homotopie.

Sprechweise. Es seien $\mathfrak{C}'$ und $\mathfrak{C}''$ Kurven von P nach Q im Gebiet G.
$\{\gamma'(t) : t \in [0, 1]\}$ und $\{\gamma''(s) : s \in [0, 1]\}$ seien Parametrisierungen. Man sagt, die Kurven seien homotop, wenn eine stetige Abbildung existiert

$$\gamma(\cdot,\cdot): \quad [0,1]\times[0,1] \longrightarrow G \qquad \begin{cases} \text{mit} & \forall u \quad \gamma(u,0) = P,\ \gamma(u,1) = Q \\ \text{und} & \gamma(0,\cdot) = \gamma'(\cdot),\ \gamma(1,\cdot) = \gamma''(\cdot) \end{cases}$$

Es existiert also eine Familie von Kurven von P nach Q, $\{\gamma^u(\cdot) : u \in [0, 1]\}$, welche $\gamma'(\cdot)$ in $\gamma''(\cdot)$ 'deformiert'. Die Frage, ob es eine solche Deformation gibt, hängt offenbar nicht von den gewählten Parametrisierungen der Kurven ab. Wir haben eine Äquivalenzrelation in der Menge der Kurven von P nach Q. Die Äquivalenzklassen heissen die Homotopieklassen der Kurven von P nach Q.

Besonders interessant ist die Menge der Äquivalenzklassen von 'geschlossenen' Kurven, die von P nach P führen. Geschlossene Kurven von P nach P kann man nämlich hintereinanderschalten; und die Homotopieklasse der zusammengesetzten Kurve hängt nur von den Homotopieklassen der 'Faktoren' ab. Man erhält eine (im Allg. nichtkommutative) Gruppe, wo das Einselement durch die Kurven repräsentiert werden, die man auf den Punkt P zusammenziehen kann. Wenn für einen bogenzusammenhängenden Raum diese Gruppe trivial ist, wenn man also jede geschlossene Kurve auf einen Punkt zusammenziehen kann, dann heisst der Raum **einfach zusammenhängend**.

Beispiel. Das Gebiet G sei die im Nullpunkt punktierte komplexe Ebene $G = \mathbb{C}\setminus\{0\}$. Das Gebiet ist bogenzusammenhängend; man kann je zwei Punkte durch eine stetige Kurve miteinander verbinden. Zwei solche Kurven kann man genau dann in G stetig ineinander deformieren, wenn sich die beiden Kurven gleich oft um den Nullpunkt schlingen.

Man sollte die Bedingung genauer formulieren: Sei $\{z(t) : t \in [a, b]\}$ eine Parametrisierung der Kurve $\mathfrak{C}$ von P nach Q, $P = z(a)$, $Q = z(b)$. Wir schreiben die Punkte der Kurve in Polarkoordinaten $z(t) = r(t)\cdot e^{i\phi(t)}$. Der 'Winkel' $\phi(t)$ ist nur bis auf ein ganzzahliges Vielfaches von 2π eindeutig bestimmt. Wenn wir den Winkel so wählen, dass er sich mit t stetig ändert, dann ist die gesamte Winkelveränderung $\phi(b) - \phi(a)$ durch die Kurve $\mathfrak{C}$ eindeutig bestimmt. Bei einer geschlossenen Kurve mit $\phi(b)-\phi(a) = k\cdot 2\pi$ nennt man die ganze Zahl $k(\mathfrak{C})$ die Umlaufszahl der Kurve (in Bezug auf den Nullpunkt). Die nullhomotopen Kurven sind die mit Umlaufszahl $= 0$; man kann sie auf den Punkt $P = Q$ zusammenziehen. (Ohne Beweis!)

Interessant sind die doppelpunktfreien geschlossenen Kurven in der komplexen Ebene; sie heissen die **Jordan-Kurven**. Es gilt der

Satz 1.2.4 (Jordan'scher Kurvensatz)**.** *Jede doppelpunktfreie geschlossene Kurve im zweidimensionalen reellaffinen Raum zerlegt diesen in genau zwei Zusammenhangskomponenten. Die Umlaufszahl bzgl. eines Punkts P kann nur zwei Werte annehmen, den Wert 0 für die Punkte im 'Äusseren' und den Wert 1 (bzw.-1) für die Punkte im 'Inneren'. (Ohne Beweis!)*

Den Wert $+1$ für die inneren Punkte erhält man, wenn die Jordan-Kurve **'im mathematisch positiven Sinn'** durchlaufen wird.

Konstruktion: Unterteilte Kurven
Es seien $\mathfrak{C}_1,\ \mathfrak{C}_2$ stetige Kurven mit $\beta(\mathfrak{C}_1) = \alpha(\mathfrak{C}_2)$. Wir können die Kurven dann zusammenfügen zu einer stetigen Kurve von $\alpha(\mathfrak{C}_1)$ nach $\beta(\mathfrak{C}_2)$. Wir bezeichnen diese zusammengefügte Kurve mit $\mathfrak{C} = \mathfrak{C}_1 \vec{\cup}\, \mathfrak{C}_2$.

Es sei $\{\gamma(t) : t \in [a,\ b]\}$ eine Parametrisierung der Kurve $\mathfrak{C}$. Eine Unterteilung der Intervalls $\big(\mathfrak{Z} : \quad a = t_0 < t_1 < \cdots < t_N = b\big)$ liefert dann ein N-Tupel von parametrisierten Kurven $\{\gamma(t) : t \in [t_{k-1},\ t_k]\}$, wo der Endpunkt der k-ten Kurve der Startpunkt der $(k+1)$-ten Kurve ist (für $k = 1, 2, \ldots, N-1$). Wir haben $\mathfrak{C} = \mathfrak{C}_1 \vec{\cup}\, \mathfrak{C}_2 \vec{\cup} \cdots \vec{\cup}\, \mathfrak{C}_N$.

Die Idee der Unterteilung einer stetigen Kurve $\mathfrak{C}$ ist offenbar nicht an die Wahl einer Parametrisierung gebunden.

Es ist klar, was eine Verfeinerung einer Unterteilung der Kurve $\mathfrak{C}$ ist. Wir notieren $\mathfrak{Z}_2 \sqsupseteq \mathfrak{Z}_1$, wenn $\mathfrak{Z}_2$ feiner ist als $\mathfrak{Z}_1$. Wir sagen, dass der Feinheitsgrad einer Folge von Unterteilungen $(\mathfrak{Z}_n)_n$ nach 0 strebt für $n \to \infty$, wenn in einer (und damit in jeder) Parametrisierung $\max\{|t_k^{(n)} - t_{k-1}^{(n)}| : k\}$ nach Null strebt für $n \to \infty$.

Definition 1.3 (Kurvenlänge)**.**
Für eine Kurve $\mathfrak{C}$ in einem metrischen Raum $\big(S, d(\cdot,\cdot)\big)$ definiert man die Kurvenlänge

$$L(\mathfrak{C}) \ = \ \sup\{\, L_{\mathfrak{Z}}(\mathfrak{C}) : \ \mathfrak{Z}\,\},$$

wobei $L_{\mathfrak{Z}}(\mathfrak{C}) = \sum_{k=1}^{N} d\big(\alpha(\mathfrak{C}_k), \beta(\mathfrak{C}_k)\big)$ für die Unterteilung $\mathfrak{C} = \mathfrak{C}_1 \vec{\cup}\, \mathfrak{C}_2 \vec{\cup} \cdots \vec{\cup}\, \mathfrak{C}_N$, und das Supremum über alle Unterteilungen zu erstrecken ist.

Die Kurvenlänge kann $+\infty$ sein. Wenn sie endlich ist, dann nennt man die Kurve rektifizierbar. Die Kurvenlänge ist additiv in dem Sinn

$$\mathfrak{C} = \mathfrak{C}_1 \vec{\cup}\, \mathfrak{C}_2 \quad \Longrightarrow \quad L(\mathfrak{C}) = L(\mathfrak{C}_1) + L(\mathfrak{C}_2).$$

Hinweis: Ein Thema der klassischen Integrationstheorie mit vielen beliebten Übungsaufgaben ist die Längenmessung für doppelpunktfreie glatte Kurven im euklidischen $\mathbb{R}^n$. Für uns ist das Thema hier nicht interessant. Es soll aber bemerkt werden: Wenn die Kurve $\mathfrak{C}$ doppelpunktfrei ist, dann hängt die Länge nur von der Spur der Kurve ab; doppelpunktfreie Kurven mit derselben Spur haben also dieselbe Länge. C. Caratheodory hat im Jahr 1914 in einer bahnbrechenden Arbeit „Lineares Maß “eine Theorie der Mengenfunktionen entwickelt, die die Längenmessung auf den Wegen der modernen Maßtheorie behandelt.

Definition 1.4 (Das Kurvenintegral einer 1-Form)**.**
Es sei $\mathfrak{C}$ eine rektifizierbare Kurve in einem metrischen Raum $\big(S, d(\cdot,\cdot)\big)$. Auf ihrer Spur sei $f(\cdot)$ eine stetige und $g(\cdot)$ eine bzgl. der induzierten Metrik Lipschitz-stetige Funktion Man definiert

$$\int_{\mathfrak{C}} f \cdot dg = \lim_{n\to\infty} \sum_k f(t_{k-1}^{(n)}) \cdot \left(g\left(t_k^{(n)}\right) - g\left(t_{k-1}^{(n)}\right)\right).$$

wobei der Limes über eine Folge von Unterteilungen, deren Feinheitsgrad nach 0 strebt, zu erstrecken ist. Der Limes heisst das Integral der 1-Form $f \cdot dg$ entlang der Kurve $\mathfrak{C}$.

Verweis auf Integrationstheorie: Wir wollen uns hier nicht mit dem Beweis aufhalten, dass der Limes existiert und unabhängig ist von der Wahl der Folge von Unterteilungen; denn wir werden in der abstrakten Maßtheorie allgemeinere Fragen dieser Art behandeln. Der Zusammenhang der maßtheoretischen Konstruktionen mit der aktuellen Situation der Kurvenintegrale ergibt sich aus der folgenden Beobachtung, die wir später noch genauer erläutern werden: Für jede Parametrisierung unserer rektifizierbaren Kurve $\{\gamma(t) : t \in [a,\ b]\}$ ist die zurückgenommene Funktion $G(t) = g(\gamma(t))$ eine Funktion beschränkter Schwankung. Das folgt aus der Annahme, dass die Kurve rektifizierbar und die Funktion g Lipschitz-stetig ist. Das Kurvenintegral ergibt sich als das Stieltjes-Integral der stetigen zurückgenommenen Funktion $F(t) = f(\gamma(t))$.

$$\int_{\mathfrak{C}} f \cdot dg \ = \int_a^b F(t)\ dG(t).$$

Wenn f im Betrag $< \varepsilon$ ist, und g entlang der Kurve die Totalvariation $\leq \bar{g}$ besitzt, dann gilt die Abschätzung $\left|\int_{\mathfrak{C}} f \cdot dg\right| < \varepsilon \cdot \bar{g}$.

Durch das Zurücknehmen (‘pullback’) der Funktionen f und g werden die maßtheoretischen Aspekte sehr einfach. Die eigentliche Bedeutung der Integration entlang von Kurven liegt aber nicht in der Maßtheorie; sie liegt im Bereich der Geometrie. Wir kommen darauf zurück, wenn wir uns mit glatten Mannigfaltigkeiten befassen. Die Integranden der Kurvenintegrale sind dort die sog. Pfaff'schen Formen $\omega = \sum_k f_k \cdot dg_k$, wo die f_k stetige und die g_k stetig differenzierbare Funktionen sind. Den Beweis des folgenden Satz verschieben wir in die Integrationstheorie.

Satz 1.2.5.
Es sei $\mathfrak{C}$ eine rektifizierbare Kurve in einem metrischen Raum; und sei g Lipschitz-stetig auf ihrer Spur. Es gilt dann

1. *Das Funktional* $I(\cdot) :\ f \longmapsto \int_{\mathfrak{C}} f \cdot dg$ *ist linear auf dem Vektorraum der stetigen Funktionen,* $\quad I(\alpha_1 f_1\ +\ \alpha_2 f_2) = \alpha_1 I(f_1)\ +\ \alpha_2 I(f_2)$,
2. *Das Kurvenintegral ist additiv bei Unterteilungen*
$$\mathfrak{C} = \mathfrak{C}_1 \vec{\cup}\, \mathfrak{C}_2 \quad \Longrightarrow \quad \int_{\mathfrak{C}} f\ dg = \int_{\mathfrak{C}_1} f\ dg + \int_{\mathfrak{C}_2} f\ dg,$$
3. *Ist auch f Lipschitzstetig, so gilt die Regel der partiellen Integration*
$$\int_{\mathfrak{C}} f \cdot dg + \int_{\mathfrak{C}} g \cdot df = \big(f \cdot g\big)(\beta(\mathfrak{C})) - \big(f \cdot g\big)(\alpha(\mathfrak{C})) \qquad \textit{kurz notiert} \ = \ \big[f \cdot g\big]_{\alpha(\mathfrak{C})}^{\beta(\mathfrak{C})},$$

Zum Beweis der Formel für die partielle Integration bemerken wir
$F(t_{k-1}) \cdot \big(G(t_k) - G(t_{k-1})\ +\ G(t_k) \cdot \big(F(t_k) - F(t_{k-1})\big)\ =\ G(t_k) \cdot F(t_k) - F(t_{k-1}) \cdot G(t_{k-1})$.

Beispiel.
Die Regel von der partiellen Integration liefert bemerkenswerte Formeln, wie z. B.

$$2 \cdot \int_{\mathfrak{C}} g \cdot dg = \left[g^2\right]_{\alpha(\mathfrak{C})}^{\beta(\mathfrak{C})} = \int_{\mathfrak{C}} d(g^2),$$

und allgemeiner $\int_{\mathfrak{C}} g^n \cdot dg = \frac{1}{n+1} \cdot \int_{\mathfrak{C}} d(g^{n+1})$. Die Funktionen f und g können übrigens auch komplexwertig sein.

Mit speziellen komplexen Kurvenintegralen über Gebieten $G \subset \mathbb{C}$ werden wir uns im nächsten Abschnitt beschäftigen. Wir bemerken schon vorab:
Wenn $p(z)$ ein Polynom ist, und $P'(z) = p(z)$, dann gilt $\int_{\mathfrak{C}} p(z)dz = P(\beta(\mathfrak{C})) - P(\alpha(\mathfrak{C}))$.

Ein einfaches Approximationsargument liefert für Funktionen $F(z)$, welche im Kreis K_R durch Potenzreihen dargestellt werden, die folgende Version des sog. Fundamentalsatzes der Differential- und Integralrechnung: Wenn $f(z) = F'(z)$, dann gilt

$$\int_{\mathfrak{C}} f(z)dz = F(\beta(\mathfrak{C})) - F(\alpha(\mathfrak{C})) \quad \text{für jede ganz im Kreis verlaufende Kurve.}$$

Interessanter wird die Sache, wenn wir analytische Funktionen auf allgemeineren Gebieten entlang von Kurven integrieren. Damit befassen wir uns im Folgenden.

1.3 Die holomorphen Funktionen sind analytische Funktionen

In diesem Abschnitt nähern wir uns dem Begriff der analytischen Funktion auf einem Weg, den man einen geometrischen Weg nennen kann.

Sprechweise (Holomorphie)**.**
Eine komplexwertige Funktion $f(z)$, die in einer Umgebung des Punkts $\tilde{z} \in \mathbb{C}$ definiert ist, heisst in $\tilde{z}$ komplex differenzierbar, wenn eine komplexe Zahl a existiert, sodass gilt

$$f(z) - f(\tilde{z}) \;=\; \tilde{a} \cdot (z - \tilde{z}) \;+\; o(|z - \tilde{z}|) \qquad \text{für} \quad z \to \tilde{z}.$$

Man notiert $\tilde{a} = f'(\tilde{z})$. Die Funktion heisst holomorph in $\tilde{z}$, wenn sie in allen Punkten in einer vollen Umgebung von $\tilde{z}$ komplex differenzierbar ist; sie heisst holomorph in der offenen Menge $G \subseteq \mathbb{C}$, wenn sie in allen Punkten in G holomorph ist.

Bekanntlich heisst eine komplexwertige Funktion auf einem Gebiet $G \subseteq \mathbb{C}$, eine analytische Funktion, wenn es zu jedem $\tilde{z}$ einen Kreis $\{z : |z - \tilde{z}| < \tilde{r}\}$ gibt, in welchem sie durch eine Potenzreihe dargestellt wird

$$f(z) - f(\tilde{z}) \;=\; \sum_{k=1}^{\infty} b_k \cdot (z - \tilde{z})^k \qquad \text{für} \quad |z - \tilde{z}| < \tilde{r}.$$

Wir machen uns jetzt daran, den für die Funktionentheorie zentralen Satz zu beweisen, welcher besagt, dass jede in einem Gebiet G holomorphe Funktion analytisch ist; man nennt ihn mit Recht den Hauptsatz über holomorphe Funktionen. Der Schlüssel zum Beweis ist das Lemma von E. Goursat (1858- 1936). Mit diesem Satz ist es Goursat um 1900 zur großen Erleichterung der Mathematikergemeinde gelungen, einen herrschenden Richtungsstreit um den Gegenstandsbereich der Funktionentheorie beizulegen. Der Satz zeigt, dass die geometrisch gefärbte Funktionentheorie im Sinne von Riemann dieselben Funktionen zum Gegenstand hat wie die analytisch gefärbte Funktionentheorie im Sinne von Weierstraß. Technisch gesehen, zeigt das Lemma zunächst einmal mit einem Kompaktheitsargument, dass holomorphe Funktionen lokal Stammfunktionen besitzen. Daraus ergibt sich dann für beliebige holomorphe Funktionen die Darstellbarkeit durch die (für analytische Funktionen schon früher bekannte) Cauchy'sche Integralformel. Diese Integralformel hat den Vorzug, dass sie nicht nur über die Funktion, sondern auch über alle ihre Ableitungen Auskunft gibt.

Zuerst diskutieren wir aber noch die komplexe Differenzierbarkeit in ihrer Beziehung zur sog. totalen Differenzierbarkeit im Sinne der reellen Funktionentheorie.

Satz 1.3.1. *Es sei* $f(z)$ *eine komplexwertige Funktion auf einem Gebiet* $G \subseteq \mathbb{C}$. *Wir schreiben* $z = x + iy$ *und* $f(z) = u(x,y) + i\,v(x,y)$. *Ist* f *im Punkt* z *komplex differenzierbar mit* $f'(z) = \alpha + i\ \beta$, *so gibt es eine Matrix* $J = \begin{pmatrix} u_x & u_y \\ v_x & v_y \end{pmatrix} = \begin{pmatrix} \alpha & -\beta \\ \beta & \alpha \end{pmatrix}$, *sodass gilt*

$$\begin{aligned} u(x+h, y+k) - u(x,y) \quad &= u_x \cdot h + u_y \cdot k + \; o(\sqrt{h^2 + k^2}) \\ v(x+h, y+k) - v(x,y) \quad &= v_x \cdot h + v_y \cdot k + \; o(\sqrt{h^2 + k^2}) \end{aligned}$$

Beweis. *Mit* $w = h + ik$ *haben wir*

$$\big(u(x+h, y+k) - u(x,y)\big) + i\ \big(v(x+h, y+k) - v(x,y)\big) = f(z+w) - f(z) =$$
$$= \big(\alpha + i\ \beta\big) \cdot \big(h + i\ k\big) + \quad o(|w|) = (\alpha \cdot h - \beta \cdot k) + i(\beta \cdot h + \alpha k) + \quad o(|w|).$$

Sprechweise (Die Cauchy-Riemannschen Differentialgleichungen)**.**
Von einem Paar reellwertiger Funktionen $u(x,y), v(x,y)$ auf einem Gebiet $G \subseteq \mathbb{R}^2$ sagt man, dass es die Cauchy-Riemannschen Differentialgleichungen erfüllt, wenn die Funktionen total differenzierbar sind und für ihre partiellen Ableitungen gilt

$$u_x = v_y, \qquad v_x = -u_y.$$

Unser Satz besagt: Das Paar u, v erfüllt die Cauchy-Riemann'schen Differentialgleichungen genau dann, wenn $u + i\ v$, betrachtet als Funktion der komplexen Variablen $z = x + i\ y$ holomorph ist. Aus dem Hauptsatz über holomorphe Funktionen, den wir beweisen werden, folgt, dass $u(\cdot,\cdot)$ und $v(\cdot,\cdot)$ unendlich oft differenzierbar sind und sogar lokal durch Potenzreihen in den Variablen x, y dargestellt werden können. Dies impliziert insbesondere, dass die zweiten Ableitungen existieren mit $u_{xx} + u_{yy} = 0$; $v_{xx} + v_{yy} = 0$. Diesen Tatbestand drückt man so aus, dass man sagt, die Funktionen u und v seien **harmonische Funktionen** im Gebiet G. Wir kommen darauf später zurück.

Beispiel. Die Funktion $f(z) = \frac{1}{z}$ ist holomorph in der punktierten komplexen Ebene $G = \mathbb{C} \setminus \{0\}$ mit der Ableitung $f'(z) = -\frac{1}{z^2}$. Für $z \neq 0$ und $w \to 0$ gilt nämlich

$$f(z+w) - f(z) = \frac{1}{z+w} - \frac{1}{z} = \frac{-w}{z(z+w)} = \Big(-\frac{1}{z^2}\Big) \cdot w \quad + o(|w|).$$

Das Funktionenpaar $u(x,y) = \frac{x}{x^2+y^2}$, $v(x,y) = \frac{-y}{x^2+y^2}$ erfüllt die Cauchy-Riemann'schen Differentialgleichungen.

Satz 1.3.2 (Lemma von Goursat)**.**
Es sei $f(z)$ *holomoph im Gebiet* G*. Für jedes kompakte Dreieck* $\Delta \subset G$ *verschwindet das Kurvenintegral über den Rand des Dreiecks:*

$$\int_{\partial\Delta} f(z)\ dz = 0.$$

Beweis. *Der Rand des Dreiecks besteht aus drei Strecken, die aneinander anschliessend durchlaufen werden. Der Umfang sei* L*. Wenn wir die Mittelpunkte der Seiten verbinden, erhalten wir ein Dreieck mit dem halben Umfang* $\frac{1}{2}L$*; und das Dreieck wird in vier kongruente Dreiecke zerlegt. Die Summe der Randintegrale ist gleich dem Randintegral des ursprünglichen großen Dreiecks. Unter den kleinen Dreiecken gibt es mindestens eines, sagen wir* Δ_1*, mit* $\left|\int_{\partial\Delta_1} f(z)\ dz\right| \geq \frac{1}{4}\left|\int_{\partial\Delta} f(z)\ dz\right|$*. Wir zerteilen dieses* Δ_1 *in derselben Weise und wählen* Δ_2*; so fahren wir fort und erhalten eine absteigende Folge kompakter Dreiecke* $\Delta \supset \Delta_1 \supset \Delta_2 \supset \cdots$*, die einen einpunktigen Durchschnitt* $\{\tilde{z}\}$ *hat. Für die* $z \in \Delta_n$ *gilt* $|z - \tilde{z}| \leq 2^{-n}L$*. Da* f *im Punkt* $\tilde{z}$ *differenzierbar ist, existiert ein Zahl* $\tilde{a}$*, sodass gilt*

$$\forall\varepsilon > 0\,\exists N\,\forall z \in \Delta_N \qquad \left|f(z) - f(\tilde z) - \tilde a \cdot (z - \tilde z)\right| \le \varepsilon \cdot |z - \tilde z| \le \varepsilon \cdot 2^{-N}.$$

Der Umfang von Δ_N *ist* $2^{-N}L$. *Wegen* $\int_{\partial\Delta_N}\big(f(\tilde z) + \tilde a(z - \tilde z)\big)\,dz = 0$ *haben wir*

$$\Big|\int_{\partial\Delta} f(z)\,dz\Big| \le 4^N \cdot \Big|\int_{\partial\Delta_N} f(z)\,dz\Big| \le 4^N \cdot \varepsilon \cdot 2^{-N} \cdot 2^{-N}L \le \varepsilon.$$

Man kann das Resultat folgendermaßen aussprechen: Das Kurvenintegral von $f(z)dz$ über den Streckenzug $[z_0, z_1] \vec\cup [z_1, z_2]$ liefert dasselbe Resultat wie das über die Strecke $[z_0, z_2]$, wenn f in einer Umgebung des eingeschlossenen Dreiecks holomorph ist.

Betrachten wir die holomorphe Funktion in einem Teilgebiet S der Zahlenebene, welches sternförmig ist in Bezug auf einen Punkt $\tilde z$; sternförmig soll heissen, dass mit $z \in S$ auch die Verbindungsstrecke $[\tilde z, z]$ zu S gehört. Sei $F(z)$ der Wert des Kurvenintegrals über diese Strecke. Mit z liegt auch eine kleine kreisförmige Umgebung K im Stern. Ist $z+w$ ein Punkt in diese kleinen Umgebung, dann ist nach dem Lemma von Goursat die Differenz $F(z + w) - F(z)$ der Wert des Kurvenintegrals über die Strecke $[z, z + w]$

$$F(z + w) - F(z) = \int_{[z,z+w]} f(\zeta)d\zeta = \int_{[z,z+w]} f(z)d\zeta + \int_{[z,z+w]} \big(f(\zeta) - f(z)\big)d\zeta.$$

Das zweite Integral ist im Betrag kleiner als $\varepsilon \cdot |w|$, wenn der Kreis K genügend klein gewählt ist. Das erste Integral ist $f(z) \cdot w$. Somit ist $F(\cdot)$ in allen Punkten von S komplex differenzierbar mit der Ableitung $F' = f$. Wir haben somit in S eine Stammfunktion zu f :

$$F(z + w) = F(z) + f(z) \cdot w + o(|w|) \quad \text{für} \quad w \to 0.$$

Wenn $\mathfrak{C}$ ein von $\tilde z$ ausgehender Polygonzug mit dem Endpunkt z ist: $\mathfrak{C} = \mathfrak{C}_1 \vec\cup \mathfrak{C}_2 \vec\cup \cdots \vec\cup \mathfrak{C}_N$, dann beweist man leicht (z. B. durch vollständige Induktion nach N)

$$\int_{\mathfrak{C}} f(\zeta)\,d\zeta = \sum_k \int_{\mathbb{C}_k} f(\zeta)\,d\zeta = \sum_k \big(F(z_k) - F(z_{k-1})\big) = F(z).$$

Es folgt sofort $\int_{\mathfrak{C}} f(\zeta)\,d\zeta = F(\beta(\mathfrak{C})) - F(\alpha(\mathfrak{C})) = \int_{\mathfrak{C}} dF$ für beliebige Polygonzüge innerhalb des Sterns. Die Gleichung gilt schliesslich nicht nur für Polygonzüge, sondern für beliebige stetige Kurven, die zwei Punkte innerhalb des Sterns verbinden. Wir müssen mit den Kurvenintegralen nicht in einem Stern bleiben. Wenn zwei Kurven in einem Gebiet homotop sind, dann liefert die Integration der holomorphen Form $f(z) \cdot dz$ entlang der beiden Kurven denselben Wert. Diese Beobachtung als bekannt als der

Satz 1.3.3 (Homotopie-Version von Cauchy's Integralsatz)**.**
Ist $f(z) \cdot dz$ *holomorph im Gebiet* $G \subseteq \mathbb{C}$, *und sind* $\mathfrak{C}'$, $\mathfrak{C}''$ *homotope Kurven in* G, *dann gilt* $\int_{\mathfrak{C}'} f(\zeta)\,d\zeta = \int_{\mathfrak{C}''} f(\zeta)\,d\zeta$.

In einem einfach zusammenhängenden Gebiet sind je zwei Kurven von z_0 nach z homotop, und das liefert

Satz 1.3.4 (Stammfunktionen in einem einfach zusammenhängenden Gebiet).
Ist $f(z) \cdot dz$ *holomorph im einfach zusammenhängenden Gebiet* $G \subseteq \mathbb{C}$*, dann existiert in* G *eine Funktion* $F(z)$*, sodass für alle in* G *verlaufenden Kurven gilt*

$$\int_{\mathfrak{C}} f(\zeta)\ d\zeta = F\big(\beta(\mathfrak{C})\big) - F\big(\alpha(\mathfrak{C})\big) = \int_{\mathfrak{C}} dF$$

Sei jetzt wieder $f(z) \cdot dz$ eine holomorphe 1-Form auf irgendeinem Gebiet G. Seien $\mathfrak{C}'$, $\mathfrak{C}''$ zwei Kurven von von z_0 nach z, um sei $\mathfrak{C}''_-$ die umgekehrt durchlaufene Kurve $\mathfrak{C}''$. Die geschlossene Kurve $\mathfrak{C}' \vec{\cup} \mathfrak{C}''_-$ ist genau dann nullhomotop, wenn $\mathfrak{C}'$ und $\mathfrak{C}''$ homotop sind. Die Homotopie-Version des Cauchy'schen Integralsatzes besagt also.

Satz. *Wenn* $f(z) \cdot dz$ *in* G *holomorph ist, dann verschwindet entlang jeder nullhomotopen Kurve das Umlaufsintegral.*

Hinweis: Es sind nicht nur die nullhomotopen Kurven, entlang derer jede holomorphe Form das Umlaufintgral $= 0$ liefert. Es gibt eine tieferliegende Version des Cauchy'schen Integralsatzes; diese besagt, dass die Umlaufsintegrale auch entlang von 'null-homologen' Kurven verschwinden. ('Homologie-Version des Cauchy'schen Integralsatzes'.) Wir müssen es Spezialveranstaltungen (zur algebraischen Topologie oder zur Funktionentheorie) überlassen, die Begriffe der Homologietheorie zu entwickeln.

Wir bleiben bei der Homotopie-Version und betrachten die holomorphen Formen lokal:

Satz 1.3.5 (Lokale Stammfunktionen).
Zu jeder in G *holomorphen Funktion* $f(z)$ *und jedem* $\tilde{z} \in G$ *existiert eine Umgebung* $\tilde{U} \ni \tilde{z}$ *und dort eine Funktion* $\tilde{F}$*, sodass* $\tilde{F}' = f$ *auf* $\tilde{U}$.

Stückweise glatte Integrationswege: Wenn man spezielle Kurvenintegrale in der komplexen Ebene auswerten will, dann benützt man i. d. Regel geschickt gewählte Parametrisierungen der Kurve. Für einen Polygonzug bietet sich eine Parametrisierung an, die auf jedem Teilstück eine konstante Geschwindigkeit hat.

$$z(t) = z(t_{k-1}) + \nu_k \cdot (t - t_{k-1}) \quad \text{für } t \in [t_{k-1}, t_k] \quad \text{mit } \nu_k = \tfrac{z(t_k) - z(t_{k-1})}{t_k - t_{k-1}}.$$

Mit $\dot{z}(t) = \nu_k$ für $t \in [t_{k-1}, t_k]$ erhalten wir wir $\int_{\mathfrak{C}} f(\zeta)\ d\zeta = \int_{t_0}^{t_N} f(z(t)) \cdot \dot{z}(t)\ dt$. Für etwas kompliziertere Kurven sind die stückweise stetig differenzierbaren Parametrisierungen beliebt. Auch hier ist die Parametermenge unterteilt, und auf jedem Teilstück ist $z(t)$ stetig differenzierbar. Auch in diesem allgemeineren Fall gilt die Formel.

Beispiel. Es sei $f(z)\ dz$ eine holomorphe Form auf einem Gebiet, welche die Peripherie des Einheitskreises $\mathfrak{C}_1$ umfasst. Das Kurvenintegral über die positiv durchlaufene Kreislinie ist dann

$$\int_{\mathfrak{C}_1} f(\zeta)\ d\zeta = \int_0^{2\pi} f(z(t)) \cdot ie^{it}\ dt.$$

Im Spezialfall $f(z)\ dz$ ergibt sich

$$\int_{\mathfrak{C}_1} \frac{1}{\zeta}\ d\zeta = \int_0^{2\pi} \frac{1}{e^{it}} \cdot ie^{it}\ dt = 2\pi i.$$

Der Beweis ist klar, wenn wir für die positiv durchlaufene Kreislinie die Parametrisierung durch den Winkel wählen $\mathfrak{C}_1 = \{z(t) = e^{it} : t \in [0, 2\pi]\}$.

Umlaufszahlen

Das Umlaufsintegral der holomorphen Form $\frac{1}{z}\ dz$ spielt eine wichtige Rolle in der Funktionentheorie. Es sei $\mathfrak{C}$ eine geschlossene Kurve, die den Ursprung nicht trifft, mit einer stückweise glatten Parametrisierung : $\mathfrak{C} = \{z(t) = r(t) \cdot e^{i\phi(t)} : t \in [a, b]\}$. Zwar ist das Argument $\phi(t)$ für die einzelnen Punkte $z(t)$ nur bis auf ein Vielfaches von 2π bestimmt; das wird uns aber nicht stören, wenn wir bei der Wahl nur darauf achten, dass $\phi(\cdot)$ stetig ist. Wir erhalten

$$\int_{\mathfrak{C}} \frac{1}{\zeta}\ d\zeta = 2\pi i \cdot n(0, \mathfrak{C}) \qquad \text{wo } n(0, \mathfrak{C}) \text{ die Umlaufszahl ist}$$

genauer gesagt: die Umlaufszahl der Kurve im Bezug auf den Nullpunkt. In der Tat

$$\int_a^b \frac{1}{r(t) \cdot e^{i\phi(t)}} \cdot \left(\dot{r}(t) \cdot e^{i\phi(t)} + r(t) \cdot i\dot{\phi}(t)\ e^{i\phi(t)}\right)\ dt = \Big[\ln r\Big]_a^b + \Big[\phi\Big]_a^b = 2\pi i \cdot n(0, \mathfrak{C}),$$

Für die Formen $z^m\ dz$ mit $z \in \mathbb{Z} \setminus (-1)$ erhalten wir den Integralwert $= 0$ für jede den Nullpunkt meidende geschlossene Kurve. Für eine Kreislinie, die den Nullpunkt n-mal umschlingt ergibt sich das aus der Rechnung

$$\int_{\mathfrak{C}_1} \zeta^m\ d\zeta = \int_0^{n \cdot 2\pi} e^{imt} \cdot ie^{it}\ dt = 0$$

Für den allgemeinen Fall deformieren wir den Integrationsweg in der punktierten Ebene zu einer solchen Kreislinie.

Die Berechnung funktioniert natürlich ebenso für Formen $(z - z_0)^m \cdot dz$. Wenn $\mathfrak{C}$ den Punkt z_0 $n(z_0, \mathfrak{C})$-mal umschlingt, dann liefert das Umlaufsintegral

$$\frac{1}{2\pi i} \cdot \int_{\mathfrak{C}} \frac{1}{\zeta - z_0}\ d\zeta = n(z_0, \mathfrak{C}). \qquad \int_{\mathfrak{C}} (\zeta - z_0)^m\ d\zeta = 0 \quad \text{für } m \neq -1.$$

Satz 1.3.6 (Cauchy's Integralformel).
Sei $f(z)$ holomorph im Gebiet G, und sei $\mathfrak{C}$ eine in G verlaufende rektifizierbare doppelpunktfreie Kurve mit positiver Orientierung. Dann gilt

$$\frac{1}{2\pi i} \int_{\mathfrak{C}} \frac{1}{\zeta - z} \cdot f(\zeta)\ d\zeta = f(z) \qquad \textit{für jedes } z \textit{ im Innern von } \mathfrak{C}.$$

Beweis. *Wir zerlegen das Umlaufsintegral:*

$$\tfrac{1}{2\pi i}\int_{\mathfrak{C}}(\zeta-z)^{-1}\cdot f(z)\ d\zeta\ +\ \tfrac{1}{2\pi i}\int_{\mathfrak{C}}(\zeta-z)^{-1}\cdot\big(f(\zeta)-f(z)\big)\ d\zeta\ = I_1+I_2.$$

Den Integrationsweg kann man im punktierten Gebiet $G\backslash\{z\}$ zu einem kleinen Kreis deformieren, welcher den Punkt z einfach positiv umschlingt. Dabei ändern sich die Integrale nicht. Das erste Integral liefert nach den obigen Rechnungen den gewünschten Wert $f(z)$. Der Integrand des zweiten ist wegen der komplexen Differenzierbarkeit von f im Punkt z im Betrag kleiner als ε, wenn der Kreis genügend klein gewählt ist.

Satz 1.3.7 (Potenzreihendarstellung der holomorphen Funktionen).
Die Annahmen seien die zu Cauchy's Integralformel. Die Länge der Kurve sei $L<\infty$. $G_{\mathfrak{C}}$ bezeichne das Gebiet im Inneren der Jordankurve $\mathfrak{C}$. Zu jedem $z_0\in G_{\mathfrak{C}}$ existiert dann ein Kreis $K_0=\{z:|z-z_0|<r_0\}$, in welchem f durch eine Potenzreihe dargestellt wird: $f(z)=\sum_0^\infty b_n(z-z_0)^n$ für $z\in K_0$: und für die Koeffizienten gilt

$$b_n\ =\frac{1}{n!}\cdot f^{(n)}(z_0)\ =\ \frac{1}{2\pi i}\int_{\mathfrak{C}}\frac{1}{(\zeta-z)^{n+1}}\cdot f(\zeta)\ d\zeta.$$

Beweis. *Für ein $\delta>0$ sei $r_0=(1-\delta)\cdot\inf\{\,|\zeta-z_0|:\ \zeta\in\mathfrak{C}\}$.*
Zu jedem $\varepsilon>0$ existiert ein N, sodass

$$\begin{aligned}(\zeta-z)^{-1}\ &=\ (\zeta-z_0)^{-1}\cdot\Big(1-\tfrac{z-z_0}{\zeta-z_0}\Big)^{-1}\\ &=(\zeta-z_0)^{-1}\cdot\sum_0^N\left(\frac{z-z_0}{\zeta-z_0}\right)^n+R_N\quad \textit{mit } |R_N|<\varepsilon\quad \textit{für } z\in K_0,\ \zeta\in\mathfrak{C}.\end{aligned}$$

Integration liefert $\quad\left|\int_{\mathfrak{C}}R_N\cdot f(\zeta)\ d\zeta\right|<\varepsilon\cdot L\cdot\sup\{|f(\zeta)|:\zeta\in\mathfrak{C}\}$ $\quad$ *für alle* $z\in K_0$

$$\frac{1}{2\pi i}\int_{\mathfrak{C}}\frac{1}{\zeta-z}\cdot f(\zeta)\ d\zeta\ =\frac{1}{2\pi i}\int_{\mathfrak{C}}\sum_0^\infty(z-z_0)^n\cdot(\zeta-z_0)^{-(n+1)}\cdot f(\zeta)\ d\zeta.$$

Wir haben also $f(z)=\sum_0^\infty b_n\cdot(z-z_0)^n$ gleichmäßig in K_0; und die Koeffizienten einer konvergenten Potenzreihe sind bekanntermaßen bis auf den Faktor $n!$ die Ableitungen von f im Zentralpunkt z_0.

Kommentar: Die hier gefundene Integraldarstellung der holomorphen Funktion f hat eine Konsequenz, die aus der Perspektive der Differenzierbarkeitstheorie im Reellen nicht zu erwarten ist: Aus der komplexen Differenzierbarkeit in jedem Punkt eines Gebiets $G\subseteq\mathbb{C}$ folgt die Stetigkeit der Ableitung und die lokale Darstellbarkeit durch Potenzreihen, d. h. die Analytizität. Wir schreiben dies nochmals in Formeln

$$\begin{aligned}&\forall\tilde z\in G\ \exists\tilde a\in\mathbb{C}.\quad f(z)-f(\tilde z)=\tilde a\cdot(z-\tilde z)+o(|z-\tilde z|)\quad \text{für } |z-\tilde z|\to 0\\ &\Longrightarrow\qquad \tilde a \text{ als Funktion von } \tilde z \text{ ist analytisch in } G.\end{aligned}$$

In untechnischer Sprache könnte man die Begründung folgendermaßen rekapitulieren:

Die Holomorphie impliziert, dass für kompakte Dreiecke in G das Umlaufsintegral verschwindet. Würde es für ein Dreieck nicht verschwinden, dann gäbe es (aufgrund des Kompaktheitsarguments im Lemma von Goursat) einen Punkt $\tilde{z}$, in welchem die komplexe Differenzierbarkeit verletzt ist. Das Verschwinden der Umlaufsintegrale impliziert die Existenz lokaler Stammfunktionen und damit die Gleichheit der Kurvenintegrale über homotope Kurven. Für festes $\tilde{z}$ ist die Funktion $(\zeta - \tilde{z})^{-1} \cdot (f(\zeta) - f(\tilde{z}))$ im punktierten Gebiet $G \setminus \{\tilde{z}\}$ holomorph. Das Umlaufsintegral I_2 über einen kleinen Kreis verschwindet, während das Umlaufsintegral I_1 den Funktionswert $2\pi i f(\tilde{z})$ liefert. Dies beweist die Cauchy'sche Integralformel, und aus dieser ergibt sich die Analytizität.

Technische Ergänzung zur Integraldarstellung der Ableitungen:
Es ist lehrreich, auch die Abweichung des Differenzenquotienten $\frac{1}{w-z}(f(w)-f(z))$ von der Ableitung $f'(z)$ durch ein Kurvenintegral darzustellen, nämlich

$$\frac{1}{2\pi i}\int_{\mathfrak{C}}\Big(\frac{1}{\zeta - w} - \frac{1}{\zeta - z} - (w-z)\frac{1}{(\zeta - z)^2}\Big) \cdot f(\zeta)\ d\zeta$$

Der Integrand ist $(w-z)^2 \cdot \frac{1}{(\zeta-z)^2\cdot(\zeta-w)}$. Wenn die Punkte eine Abstand $\geq \delta > 0$ von der Kurve halten, dann ist der Unterschied des Differenzenquotienten zur Ableitung in einem der Punkte ein gleichmäßiges $O(|(w-z)|$.
Für die höheren Ableitungen können wir ebenso argumentieren. Wir benötigen dazu nur eine Verallgemeinerung der eben benützten algebraischen Umformung

$$\tfrac{1}{B} - \tfrac{1}{A} - \big(\tfrac{1}{B} - \tfrac{1}{A}\big) \cdot AB \cdot \tfrac{1}{A^2} = (A-B)^2 \cdot \tfrac{1}{A^2 B}$$

Allgemeiner gilt: Für jedes $n \in \mathbb{N}$ existiert ein homogenes Polynom vom Grad $n+2$, sodass

$$\big(\tfrac{1}{B^n} - \tfrac{1}{A^n}\big) - \big(\tfrac{1}{B} - \tfrac{1}{A}\big) \cdot AB \cdot \tfrac{n}{A^{n+1}} = (A-B)^2 \cdot p_{n+2}(\tfrac{1}{A}, \tfrac{1}{B}).$$

Wir erhalten eine gleichmäßige Abschätzung für Paare z, w, die deutlichen Abstand zur Kurve halten in dem Sinn $\inf\{|z-\zeta| : \zeta \in \mathfrak{C}\} > \delta, \quad \inf\{|w-\zeta| : \zeta \in \mathfrak{C}\} > \delta.$
Es existiert eine Zahl C_δ, sodass für diese Paare gilt $\Big|\frac{1}{w-z}\big(f(w) - f(z)\big) - f'(z)\Big| < |w-z| \cdot C_\delta.$

Man könnte nun im Sinne der Euler'schen Terminologie sagen: Der Cauchy'sche Integralsatz zu einer Jordankurve präsentiert eine in einer Umgebung des Inneren holomorphe Funktion durch einen 'geschlossenen' d. h 'analytischen' Ausdruck. Holomorphe Funktionen in einem einfach zusammenhängenden Gebiet und ihre Ableitungen können also auch in der alten Euler'schen Terminologie als analytische Funktionen gelten.

1.4 Die klassischen elementaren Umkehrfunktionen

Eine komplexwertige Funktion $f(z)$ auf einem Gebiet G kann auch als Abbildung $f(\cdot)$: $G \longrightarrow \mathbb{C}$ verstanden werden. Wenn wir von Abbildungen sprechen, dann denken wir vor allem an das Hintereinanderschalten und an das Invertieren, (während wir beim Sprechen über Funktionen in erster Linie an das Linearkombinieren und das (punktweise) Multiplizieren denken.)

Die lokalen Aspekte des Einsetzens und des Invertierens haben wir schon früher angesprochen; hier funktioniert der Kalkül der Potenzreihen ('Koeffizientenvergleich') in Verbindung mit dem Umzentrieren. Wir wissen: Wenn f holomorph ist, dann existiert zu jedem Punkt $\tilde{z}$ mit $f'(\tilde{z}) \neq 0$ eine Umgebung, in welcher die Abbildung umkehrbar ist, und die Umkehrabbildung ist holomorph. Kurz gesagt: Die im Gebiet G holomorphe Abbildung $f(\cdot)$ ist im Gebiet $G \cap \{z : f'(z) \neq 0\}$ lokal umkehrbar.

Eine ganz andere Thematik ist nun aber die der globalen stetigen Umkehrbarkeit. Wir diskutieren unter diesem Gesichtspunkt die sog. elementaren Funktionen.

1)Die Logarithmusfunktion

Die Exponentialabbildung bildet die komplexe Ebene surjektiv (aber nicht injektiv!) auf die punktierte Zahlenebene $\mathbb{C} \setminus \{0\}$ ab. $\exp(a + ib) = e^a \cdot e^{ib}$. Sie bildet z_1 und z_2 genau dann in denselben Punkt ab, wenn $z_2 - z_1$ ein ganzzahliges Vielfaches von $2\pi i$ ist. Die Parallelen zur reellen Achse werden auf Strahlen (vom Nullpunkt aus) abgebildet. Die Parallelen zur imaginären Achse werden auf Kreise mit dem Mittelpunkt 0 abgegebildet.

Eine interessante Frage mit einer überzeugenden Antwort ist die folgende:

Gegeben ist eine holomorphe Funktion $f(z)$ auf einem Gebiet G, die nirgends in G den Wert 0 annimmt. Gibt es eine stetige Funktion $g(z)$ auf G mit $f(z) = \exp(g(z))$ auf G? Wenn $g(z)$ eine solche Funktion ist, dann leistet auch $g(z) + 2\pi i \cdot k$ mit $k \in \mathbb{Z}$ das Verlangte. Wenn wir für ein vorgegebenes $\tilde{z}$ den Wert $g(\tilde{z})$ vorgeben (selbstverständlich unter Beachtung $\exp(g(\tilde{z})) = f(\tilde{z})$) dann kann das damit eindeutig bestimmte g als eine (stetige!) Version von $\ln f$ bezeichnet werden.

Satz 1.4.1 (Das Logarithmieren einer holomorphen Funktion)**.**
Ist $f(z)$ eine holomorphe Funktion ohne Nullstellen in einem Gebiet G, und besitzt $\dfrac{f'(z)}{f(z)}$ in G eine Stammfunktion, so existiert eine holomorphe Funktion $g(z)$ mit $\exp(g(z)) = f(z)$.

Beweis. *Wenn g das verlangte leistet, dann gilt nach der Kettenregel*

$$\frac{d}{dz}\exp(g(z)) = \exp(g(z)) \cdot g'(z); \qquad \textit{also} \quad g'(z) = \frac{f'(z)}{f(z)}.$$

Wenn $\dfrac{f'(z)}{f(z)}$ eine Stammfunktion $g(z) = \tilde{g}(z) + c$ besitzt und die Konstante c so gewählt ist, dass $\exp(g(\tilde{z})) = f(\tilde{z})$ für ein beliebig vorgegebenes $\tilde{z}$, dann leistet g das Verlangte.

Beispiel 1.4.1 (Der Hauptwert des Logarithmus).
Zur Einschränkung f der Funktion z auf die im Negativen geschlitzte Zahlenebene $G_{\ln} =$

$\mathbb{C}\setminus(-\mathbb{R}_+)$ gibt eine Funktion g mit $\exp(g(\cdot)) = f(\cdot)$ und $g(1) = 0$. Diese Funktion auf $G_{\ln}$ heisst der Hauptwert des natürlichen Logarithmus: $g(re^{i\phi}) = \ln r + i\phi$ mit $|\phi| < \frac{\pi}{2}$.

Beispiel 1.4.2 (Der Logarithmus des Cosinus).
Bezeichne $N_{\cos}$ die Nullstellenmenge der Cosinusfunktion, $N_{\cos} = \{\frac{\pi}{2} + k\cdot\pi :\ k \in \mathbb{N}\}$, und sei G die 'doppeltgeschlitzte' Zahlenebene $G = \mathbb{C}\setminus\left\{x :\ x \text{ reell mit } |x| \geq \frac{\pi}{2}\right\}$. Auf G existiert dann eine Funktion g, sodass $\exp(g(\cdot)) = \cos(\cdot)$ und $g(0) = 0$. Sie ist reellwertig im Intervall $(-\frac{\pi}{2}, \frac{\pi}{2})$.Man bezeichnet sie manchmal als den Hauptwert von $\ln\cos(\cdot)$.

Beispiel 1.4.3 (Die logarithmische Ableitung der Gamma-Funktion).
Die Gammafunktion $\Gamma(\cdot)$ ist holomorph in $G_\Gamma = \mathbb{C}\setminus\{0, -1, -2, \ldots\}$, und sie hat nirgends Nullstellen. Ihre Einschränkung auf das Gebiet $G_{\ln} = \mathbb{C}\setminus(-\mathbb{R}_+)$, die im Negativen geschlitzte Zahlenebene, (einem einfach zusammenhängenden Gebiet,) besitzt eine Darstellung $\Gamma(\cdot) = \exp(g(\cdot))$. Die Ableitung der hier gewonnenen Funktion $g(z) = \ln\Gamma(z)$ nennt man die logarithmische Ableitung der Gamma-Funktion; sie ist sogar auf der vielfach gepunkteten Menge G_Γ holomorph. Wir zeigen, wie man sie durch eine eindrucksvolle Reihe darstellen kann, eine Reihe, die eng verwandt ist mit der berühmten Cosinusreihe, die wir später präsentieren. Wir wissen

$$\frac{1}{\Gamma(z)} = G(z) = \lim_{n\to\infty} P_n(z) \qquad \text{mit} \quad P_n(z) = \frac{z(z+1)\cdots(z+n)}{n!\cdot n^z}$$

(im Sinne der lokal gleichmäßigen Konvergenz auf dem Gebiet G_Γ).

Für $\Re z > 0$ ist der (Hauptwert des) Logarithmus wohldefiniert, und es gilt dort

$$\frac{d}{dz}\big(\ln P_n(z)\big) = \frac{1}{z} + \frac{1}{z+1} + \cdots + \frac{1}{z+n} - \ln n.$$

Der Limes für $n \to \infty$ ist die in $\{\Re z > 0\}$ holomorphe Funktion $\frac{G'}{G} = -\frac{\Gamma'}{\Gamma}$.
Man könnte das Resultat auch suggestiv so schreiben

$$-\frac{d}{dz}\ln\Gamma(z) = \frac{1}{z} + \sum_1^\infty\Big(\frac{1}{z+k} - \frac{1}{k}\Big) + \lim_{n\to\infty}\Big(\sum_1^n \frac{1}{k} - \ln n\Big).$$

Die Summe $\sum_1^\infty\left(\frac{1}{z+k} - \frac{1}{k}\right)$ existiert im Sinne der gleichmäßigen Konvergenz auf Kompakten. Für die z mit negativen Realteil bekommen wir die gesuchte Funktion mittels der Funktionalgleichung

$$z\cdot\Gamma(z) = \Gamma(z+1), \qquad \frac{\Gamma'}{\Gamma}(z) = -\frac{1}{z} + \frac{\Gamma'}{\Gamma}(z+1).$$

Bemerke: Die Reihendarstellung zeigt, dass die Gammafunktion, eingeschränkt auf $\mathbb{R}_+$, logarithmisch konvex ist. $(\ln\Gamma)''(x) > 0$ für $x > 0$.

2)Die Arcustangensfunktion

Die Tangensfunktion ist π-periodisch und holomorph auf $\mathbb{C}\setminus N_{\cos}$.

$$\tan z = \frac{\sin z}{\cos z} = \frac{1}{i}\cdot\frac{e^{iz} - e^{-iz}}{e^{iz} + e^{-iz}} = \frac{1}{i}\cdot\frac{e^{2iz} - 1}{e^{2iz} + 1}$$

Die Gleichung $w = \tan z$ hat für $w \neq \pm i$ hat genau eine Lösung im senkrechten Streifen $\left\{z:\ -\frac{\pi}{2} < \Re z \leq \frac{\pi}{2}\right\}$; denn

$$\left(e^{iz} + e^{-iz}\right) \cdot iw = \left(e^{iz} - e^{-iz}\right); \quad \Longleftrightarrow \quad e^{2iz} = -\frac{w-i}{w+i};$$

und die lineargebrochene Abbildung $w \mapsto \frac{w-i}{w+i}$ bildet die erweiterte Zahlenebene $\bar{\mathbb{C}} = \mathbb{C} \cup \infty$ bijektiv auf sich ab. Die Tangensfunktion ist im gesamten Definitionsgebiet lokal umkehrbar; denn die Ableitung $(\tan z)' = 1 + (\tan z)^2$. verschwindet nirgends. Eine globale Umkehrung des Tangens auf einem Teilgebiet $D \subseteq \mathbb{C} \setminus N_{\cos}$ nennt man eine Version des Arcustangens. Zwei Versionen des Arcustangens unterscheiden sich um ein ganzzahliges Vielfaches von π. Wenn $\tan(g(z))) = f(z)$ auf G, dann gilt nach der Kettenregel $\frac{d}{dz} \tan g(z) = \left(1 + \tan^2 g(z)\right) \cdot g'(z)$, also $g'(z) = \frac{f'}{1+f^2}(z)$. Somit gilt

Satz 1.4.2. *Für eine holomorphe Funktion $f(z)$ auf einem Gebiet G (welche die Werte $\pm i$ nicht annimmt), sind die folgenden Bedingungen äquivalent.*

(i) Es existiert eine stetige Funktion $g(z)$, sodass $f(z) = \tan g(z)$ auf G.

(ii) Die Funktion $\dfrac{f'}{1+f^2}$ besitzt eine Stammfunktion auf G.

Beispiel (Der Hauptwert des Arcustangens).
Ist $N = \{w:\ w = ib \text{ mit } b \text{ reell}, |b| \geq 1\}$, dann ist $\mathbb{C} \setminus N$ vom Nullpunkt aus gesehen ein Sterngebiet, in welchem die Funktion $\frac{1}{1+w^2}$ holomorph ist. Wenn man in diesem Sterngebiet vom Nullpunkt aus das unbestimmte Integral nach w bestimmt, erhält man eine Version des Arcustangens $a(w)$. Diese bildet $\mathbb{C}\setminus N$ bijektiv auf den offenen Vertikalstreifen $G = \{z:\ |\Re z| < \frac{\pi}{2}\}$ ab, und die Tangensfunktion ist die Umkehrabbildung. Interessant ist, wie die Tangensfunktion die Randgeraden des Vertikalstreifens abbildet. Für $z = \frac{\pi}{2} + ia$ mit a reell haben wir

$$\tan(\tfrac{\pi}{2} + ia) = \frac{1}{i} \cdot \frac{\exp(2(\frac{\pi}{2} + ia)) - 1}{\exp(2(\frac{\pi}{2} + ia)) + 1} = \frac{1}{i} \cdot \frac{-\exp(-2a)) - 1}{-\exp(-2a)) + 1} = i \cdot (\tanh a)^{-1},$$

und der hyperbolische Tangens bildet die reelle Achse auf das Intervall $(-1, +1)$ ab.

Für das Folgende ist es bequemer, statt des Tangens den Cotangens zu studieren

$$\cot z = \frac{1}{\tan z} = \tan(z + \tfrac{\pi}{2}) \quad \text{für } z \neq k\pi.$$

Die singulären Stellen des Cotangens sind die ganzzahligen Vielfachen von π. Nur in der Nähe dieser Stellen ist der Betrag des Cotangens groß. Genauer: Für jedes $a > 0$ ist der Betrag $|\cot z|$ beschränkt im Gebiet $G_a = \left\{z:\ |z - k \cdot \pi| > a \text{ für alle } k \in \mathbb{Z}\right\}$. Dies sieht man folgendermaßen: $\cot z = w \Leftrightarrow e^{2iz} = \frac{w-i}{w+i}$. Die hier auftretende lineargebrochene

Abbildung bildet den unendlichfernen Punkt nach 1 an, und das Äussere eines Kreises mit großem Radius $\{w: \; |w| > M\}$ auf das Innere eines Kreises mit kleinem Radius. Nur für die e^{2iz} in diesem kleinen Kreis nimmt $\cot z$ Werte mit Betrag $> M$ an.— Wir werden das später benützen, wenn wir die berühmte lokal gleichmäßig konvergente Cotangensreihe herleiten.

3)Die Arcussinusfunktion

Die Sinusfunktion $\sin z$ ist 2π-periodisch und in der gesamten komplexen Ebene holomorph. Sie ist aber nicht überall lokal invertierbar, denn die Ableitung $\cos z$ verschwindet in der Menge $N_{\cos}$. Dies sind die Punkte, in welchen der Sinus einen der Werte ± 1 annimmt. Wir bemerken $\sin' z = \pm\sqrt{1-\sin^2 z}$. Für eine Funktion $f(z)$ auf einem Gebiet G, welche die Werte ± 1 nicht annimmmt, suchen wir eine Funktion $g(z)$ mit

$$f(z) = \sin g(z); \qquad \text{insbesondere} \quad f'(z) = \pm\sqrt{1-f^2(z)} \cdot g'(z).$$

Welches Vorzeichen hier in Betracht kommt, ist zunächst nicht klar. Die Gleichung $w = \frac{1}{2i} \cdot \left(e^{iz} - e^{-iz}\right)$ ist eine quadratische Gleichung für e^{iz}, welche genau zwei Lösungen besitzt:

$$\begin{aligned} \sin z = w &\iff e^{iz} = iw \pm \sqrt{1-w^2} \\ &\iff z = \tfrac{1}{i}\ln\left(iw \pm \sqrt{1-w^2}\right) + 2\pi k = \left(\arcsin w + 2\pi k\right). \end{aligned}$$

Warnung: Die Schreibweise $\arcsin(\cdot)$ ist so nicht zulässig, weil wir nicht festgelegt haben, welche der Lösungen der quadratischen Gleichung zu wählen ist. Wir müssen bedachtsamer vorgehen, wenn wir die Aufgabe lösen wollen:

Beispiel 1.4.4. Sei $f(z)$ die Einschränkung der Funktion z auf die doppeltgeschlitzte Ebene $G = \mathbb{C} \setminus \{x: \; x \text{ reell und } |x| \geq 1\}$. Vom Nullpunkt aus gesehen ist G ein Sternbereich. Für $\zeta \in G$ sind $1-\zeta$ und $1+\zeta$ nicht negativ reell. Wir präzisieren für $\zeta \in G$

$$\frac{1}{\sqrt{1-\zeta^2}} = \exp\left(-\tfrac{1}{2}\ln(1-\zeta)\right) \cdot \exp\left(-\tfrac{1}{2}\ln(1+\zeta)\right)$$

wo $\ln(\cdot)$ den Hauptwert bedeutet. Das Kurvenintegral im Sternbereich $g(z) = \int_0^z \frac{1}{\sqrt{1-\zeta^2}}\, d\zeta$. löst das Problem. Man nennt diese im Nullpunkt verschwindende Stammfunktion manchmal den den Hauptwert des Arcussinus.

Sei nun allgemeiner $f(z)$ eine holomorphe Funktion in einem einfach zusammenhängenden Gebiet G, welche die Werte ± 1 nicht annimmt. Es existieren dann genau zwei holomorphe Funktionen $h(z)$ und $-h(z)$, sodass $h^2(z) = 1 - f^2(z)$. Sei $\tilde{z}$ irgendein Punkt im Gebiet. Mit $\tilde{c}$, sodass $f(\tilde{z}) = \sin\tilde{c}$, erfüllt dann eine der beiden Stammfunktionen $g_+(z) = \tilde{c} + \int_{\tilde{z}}^z \frac{1}{h(\zeta)}\, d\zeta$, $\qquad g_-(z) = \tilde{c} - \int_{\tilde{z}}^z \frac{1}{h(\zeta)}\, d\zeta$ die Bedingung $f(z) = \sin g(z)$. Diese Funktion kann als eine Version von $\arcsin f(z)$ bezeichnet werden.

4) Die Joukowski-Abbildung und ihre Umkehrung

Wir diskutieren eine Abbildung der gepunkteten komplexen Ebene, die nicht nur als Rechenbeispiel weithin beliebt ist, sondern auch (in früheren Zeiten) dazu benützt wurde, gewisse Phänomene der Strömungstheorie zu erläutern.

$$f(\cdot): \quad \mathbb{C} \setminus \{0\} \ni z \longmapsto \tfrac{1}{2}\big(z + \tfrac{1}{z}\big).$$

Die Ableitung $f'(z) = \frac{1}{2}(1 - z^{-2})$ verschwindet in den Punkten ± 1, in allen anderen Punkten ist f lokal umkehrbar. Die Punkte ± 1 sind Fixpunkte der Abbildung. Es gilt $f(\frac{1}{z}) = f(z)$. Zu jedem $w \neq \pm 1$ existieren genau zwei Urbilder $z_{\pm} = w \pm \sqrt{1 - w^2}$; und die sind zueinander reziprok. Die Punkte auf der Peripherie des Einheitskreises werden auf das Intervall $I = [-1, +1]$ abgebildet: $f(e^{i\phi}) = \cos\phi$. Das Innere und das Äussere des Einheitskreises werden umkehrbar eindeutig auf die Menge $\mathbb{C} \setminus I$ abgebildet.

Interessant sind die Bilder der Strahlen durch den Nullpunkt und die Bilder der Kreise mit dem Mittelpunkt 0. Diese Kreise werden zu konfokalen Ellipsen. Und was die Strahlen durch den Nullpunkt betrifft, so gilt: Der Teilstrahl ausserhalb des Einheitskreises wird zu einem halben Hyperbelbogen, und ebenso der Teilstrahl innerhalb des Einheitskreises. In der Tat gilt für den Kreis mit dem Radius r

$$f(r \cdot e^{i\phi}) = \tfrac{1}{2} r \cdot e^{i\phi} + \tfrac{1}{2r} e^{-i\phi} = \tfrac{1}{2}(r + \tfrac{1}{r}) \cos\phi + \tfrac{i}{2}(r - \tfrac{1}{r}) \sin\phi = a \cdot \cos\phi + ib \cdot \sin\phi.$$

Wenn der Kreis mit Radius $r > 1$ im positiven Sinn durchlaufen wird, dann durchläuft der Bildpunkt die Ellipse ebenfalls im positiven Sinn; bei den Kreisen mit Radius $r < 1$ kehrt sich der Durchlaufungssinn um. Die Brennpunkte aller Ellipsen sind ± 1; denn

$$a^2 - b^2 = \tfrac{1}{4}(r + \tfrac{1}{r})^2 - \tfrac{1}{4}(r - \tfrac{1}{r})^2 \; = \; 1.$$

Betrachten wir jetzt die Bilder der Strahlen

$$\begin{gathered} f(r \cdot e^{i\phi}) = \cos\phi \cdot \tfrac{1}{2}(r + \tfrac{1}{r}) + \sin\phi \cdot \tfrac{i}{2}(r - \tfrac{1}{r}) = \alpha \cdot \cosh\psi + i\beta \cdot \sinh\psi \\ \text{mit} \qquad \alpha = \cos\phi, \; \beta = \sin\phi, \qquad \psi = \ln r. \end{gathered}$$

Wenn r das Intervall $[1, \infty]$ durchläuft, dann durchläuft ψ das Intervall $[0, \infty]$ und der Bildpunkt $z(\psi) = \alpha \cdot \cosh\psi + i\beta \cdot \sinh\psi$ den Hyperbelbogen, der auf der reellen Achse im Punkt $\alpha = \cos\phi$ startet. Wenn wir mit r unter 1 gehen, dann läuft der Bildpunkt denselben Hyperbelbogen zurück. Alle unsere Hyperbeln haben die Brennpunkte ± 1; denn $\alpha^2 + \beta^2 = \cos^2\phi + \sin^2\phi = 1$.

Hinweis: In der einfachsten Anwendung auf die Strömungstheorie interessiert man sich dafür, wie die Umkehrabbildung $f^{-1}(w)$ die Parallelen zur reellen Achse in den Aussenraum des Einheitskreises abbildet. Die Vorstellung ist die, dass diese die Strömungslinien einer Strömung beschreiben, welche einen Zylinder mit kreisförmigen Querschnitt umströmt. — R. Feynman in seinen berühmten Lecture notes (Kapitel 41) nennt die dazugehörigen Ideen die Theorie des trockenen Wassers, der er eine elaboriertere Theorie des nassen Wassers gegenüberstellt.

1.5 Rechenbeispiele zur Verlagerung von Integrationswegen

Wenn man bedenkt, dass $\frac{1}{1+x^2}$ auf $\mathbb{R}$ die Ableitung der Arcustangensfunktion ist, der Umkehrfunktion der Tangensfunktion, die das Intervall $(-\pi/2, +\pi/2)$ bijektiv auf $\mathbb{R}$ abbildet, dann ist klar

$$\frac{1}{2\pi} \cdot \int_{-\infty}^{\infty} \frac{1}{1+x^2}\, dx = \frac{1}{2}.$$

Hier wollen uns hier mit weiteren speziellen bestimmten Integralen dieser Art befassen, um damit die Technik des Verlagerns von Integrationswegen erläutern. Integrieren wir beispielsweise die Form $\frac{1}{1+z^2}\, dz$ über die geschlossene Kurve $\mathfrak{C}_T$, die zunächst entlang der reellen Achse von $-T$ nach $+T$ läuft und dann auf dem Halbkreisbogen in der oberen Halbebene zurück zum Ausgangspunkt. Der Beitrag des Halbkreisbogens ist im Betrag $< \varepsilon$, wenn T genügend groß ist. Andererseits kann man die die Kurve $\mathfrak{C}_T$ im Holomorphiegebiet homotop verformen in eine Kreisline mit kleinem Radius, welche den Punkt $+i$ einmal im positiven Sinn umschlingt. Wegen $\frac{1}{1+z^2} = \frac{1}{2i}\left(\frac{1}{z-i} - \frac{1}{z+i}\right)$ haben wir

$$\frac{1}{2\pi}\int_{\mathfrak{C}_T} \frac{1}{1+z^2}\, dz = \frac{1}{2\pi}\int_{\mathfrak{C}_T} \frac{1}{2i}\frac{1}{z-i}\, dz = \frac{1}{4\pi i}\oint \frac{1}{z-i}\, dz = \frac{1}{2}.$$

Als Verallgemeinerung gewinnen wir die sog. charakteristischen Funktionen der symmetrischen Cauchy-Dichten:

Satz 1.5.1 (Die symmetrischen Cauchy-Dichten)**.**
Für $a > 0$ *heisst* $p_a(x)\, dx = \frac{1}{\pi}\frac{a}{a^2+x^2}\, dx$ *die symmetrische Cauchy-Dichte zum Parameter* a. *Es gilt für* $t \in \mathbb{R}$

$$\varphi(t) = \frac{1}{\pi}\int_{-\infty}^{\infty} e^{itx} \cdot \frac{a}{a^2+x^2}\, dx = \exp(-a \cdot |t|).$$

Beweis. *Eine Streckung der Integrationsvariablen zeigt, dass es genügt, den Fall* $a = 1$ *zu behandeln. Wegen* $\varphi(-t) = \varphi(t)$ *genügt es, positive* t *zu studieren. Mit den Kurven* $\mathfrak{C}_T$ *von oben erhalten wir für die Umlaufsintegrale wie oben*

$$\frac{1}{2\pi}\int_{\mathfrak{C}_T} e^{itz} \cdot \frac{1}{1+z^2}\, dz = \frac{1}{2\pi}\int_{\mathfrak{C}_T} e^{itz} \cdot \frac{1}{2i}\frac{1}{z-i}\, dz = e^{-t}.$$

Wegen $t \geq 0$ *ist der Beitrag des Halbkreisbogens im Betrag* $< \varepsilon$, *wenn* T *groß ist.*

Satz 1.5.2 (Die sog. Laplace-Integrale)**.**

$$\lim_{T\to\infty} \int_{-T}^{T} \frac{\sin x}{x}\, dx = \pi;$$

und allgemeiner $$\lim_{T\to\infty} \int_{-T}^{T} \frac{x \cdot \sin ax}{x^2+b^2}\, dx = \pi \cdot e^{-ab} \qquad \text{für } a, b > 0.$$

Beweis. *Wir integrieren über die Kurve $\mathfrak{C}_T$ von oben. Sie umschlingt den Punkt ib, aber nicht den Punkt $-ib$. Die Cauchy'sche Integralformel zeigt*

$$\frac{1}{2\pi i}\int_{\mathfrak{C}_T} \frac{e^{iaz}}{z-ib}\,dz = e^{-ab}, \qquad \frac{1}{2\pi i}\int_{\mathfrak{C}_T} \frac{e^{iaz}}{z+ib}\,dz = 0.$$

Für großes T liefert nur der Integrationsweg entlang der reellen Achse einen wesentlichen Beitrag zum Umlaufsintegral. Die Abschätzung benützt wesentlich, dass für die z mit großem positiven Imaginärteil der Betrag des Integranden klein ist; $|e^{iaz}| = e^{-a\cdot\Im z}$; die Länge des Integrationswegs fällt demgegenüber nicht ins Gewicht.

$$\begin{aligned} 2\pi\cdot e^{-ab} &= \lim_{T\to\infty}\frac{1}{2ib}\cdot\int_{-T}^{T}\frac{1}{z^2+b^2}\cdot\left(e^{iaz}(z+ib)+e^{iaz}(z-ib)\right)\,dz \\ &= \lim_{T\to\infty}\frac{1}{2ib}\cdot\int_{-T}^{T}\frac{1}{z^2+b^2}\cdot\left(2z\cdot(\cos az+ib\sin az)\right)dz \end{aligned}$$

Auf ähnlichen Wegen bestätigen wir jetzt das Integral

$$\frac{1}{2\pi}\cdot\int_{-\infty}^{\infty}\frac{1}{1+x^4}\,dx = \frac{1}{2\cdot\sqrt{2}}.$$

Der Nenner der gebrochenrationalen Funktion $\frac{1}{1+z^4}$ hat Nullstellen in den primitiven achten Einheitswurzeln $\frac{1}{\sqrt{2}}\cdot(\pm 1\pm i)$. Wir notieren $z_1 = e^{\frac{1}{8}2\pi i} = \frac{1}{\sqrt{2}}\cdot(+1+i)$. Die weiteren primitiven achten Einheitswurzeln sind dann $z_3 = z_1^3,\ z_5 = z_1^5,\ z_7 = z_1^7$. Es gilt

$$\frac{1}{z^4+1} = \frac{1}{(z-z_1)(z-z_3)(z-z_5)(z-z_7)} = \frac{-1}{4}\left(\frac{z_1}{z-z_1}+\frac{z_3}{z-z_3}+\frac{z_5}{z-z_5}+\frac{z_7}{z-z_7}\right).$$

Die Kurve $\mathfrak{C}_T$ umschlingt die Punkte z_1 und z_3 einfach im positiven Sinn. Also gilt

$$\frac{1}{2\pi}\int_{\mathfrak{C}_T}\frac{1}{1+z^4}\,dz = \frac{1}{2\pi}\cdot\frac{-1}{4}\cdot 2\pi i\cdot(z_1+z_3) = \frac{1}{2\cdot\sqrt{2}}.$$

Satz 1.5.3 (Einfache Umlaufsintegrale für gebrochenrationale Funktionen)**.**
Es sei $q(z)$ ein Polynom vom Grad m, welches einfache Nullstellen besitzt in den Punkten $z_1,\ldots,z_m$. $p(z)$ sei ein Polynom vom Grad $< m$. $\mathfrak{C}$ sei eine positiv orientierte Jordankurve, in deren Innenbereich die Punkte $z_1,\ldots,z_l$ liegen, und die übrigen z_k ausserhalb. Es gilt dann

$$\frac{1}{2\pi i}\int_{\mathfrak{C}}\frac{p(z)}{q(z)}\,dz = \sum_{k=1}^{l}\frac{p(z_k)}{q'(z_k)}.$$

Beweis.
Gemäß der Partialbruchzerlegung gibt es Zahlen $a_1, \ldots, a_m$*, sodass* $\frac{p(z)}{q(z)} = \sum_{k=1}^m \frac{a_k}{z-z_k}$ *Die Summanden zu* $k = 1, \ldots, l$ *liefern den Beitrag* a_k *zum gesuchten Umlaufsintegral. Es müssen also nur noch die* a_k *berechnet werden.* a_k *ergibt sich als Limes*

$$a_k = \lim_{z \to z_k} (z - z_k) \cdot \frac{p(z)}{q(z)} = \frac{p(z_k)}{q'(z_k)}.$$

Sprechweise (Polstellen, Residuen).
Sei $f(z)$ eine Funktion, die im punktierten Kreis $\{z : 0 < |z - \tilde{z}| < r\}$ holomorph ist. Wenn für ein $m \in \mathbb{N}$ die Funktion $(z - \tilde{z})^m \cdot f(z)$ im ganzen Kreis holomorph ist

$$f(z) = \frac{a_{-m}}{(z - \tilde{z})^m} + \cdots + \frac{a_{-1}}{z - \tilde{z}} + a_0 + a_1(z - \tilde{z}) + \cdots$$

mit $a_{-m} \neq 0$, dann sagt man: die Funktion besitzt in $\tilde{z}$ einen Pol der Ordnung m mit dem Residuum a_{-1}. Man bemerkt: integriert man f über einen kleinen Kreis $\mathfrak{C}$, der die Polstelle $\tilde{z}$ umschliesst, dann erhält man das Residuum

$$\frac{1}{2\pi i} \cdot \int_{\mathfrak{C}} f(\zeta)\, d\zeta = a_{-1}.$$

Damit gewinnen wir eine Verallgemeinerung des oben studierten Laplace-Integrale

Satz. *Es seien* p *und* q *Polynome,* q *vom Grad* $= m$*,* p *vom Grad* $< m$*. Die in der oberen Halbebene liegenden Nullstellen von* q *seien* $z^{(1)}, \ldots, z^{(l)}$*, die Residuen der gebrochenrationalen Funktion* $f(z) = \frac{p(z)}{q(z)}$ *in diesen Punkten seien* $c^{(1)}, \ldots, c^{(l)}$*. Es gilt dann für jedes* $s > 0$

$$\lim_{T \to \infty} \frac{1}{2\pi i} \cdot \int_{-T}^{T} e^{isx} \cdot f(x)\, dx = \sum_{1}^{l} c^{(j)} \cdot \exp\bigl(isz^{(j)}\bigr).$$

Die Idee lässt sich weiter verallgemeinern; sie wird die Residuenmethode genannt. Wir behandeln hier noch ein berühmtes Beispiel, und überlassen die weitere Diskussion der 'Residuenmethode' einer Spezialveranstaltung über Funktionentheorie.

Satz 1.5.4 (Die Cotangensreihe).
Für ein kleines $a > 0$ *sei* $G_a = \bigl\{z : |z - k \cdot \pi| > a$ *für alle* $k \in \mathbb{Z}\bigr\}$*. Es gilt*

$$\cot z = \frac{1}{z} + \sum_{k=1}^{\infty} \frac{2z}{z^2 - k^2\pi^2}$$

gleichmäßig in jedem kompakten Teilbereich von G_a*.*

Beweis. *Die Cotangensfunktion* $\cot z = \frac{\cos z}{\sin z}$ *ist holomorph in* $\mathbb{C} \setminus \{\pi \cdot \mathbb{Z}\}$ *und sie hat einen Pol der Ordnung eins mit dem Residuum* $= 1$ *in jedem Punkt* $k\pi$*. Der Cotangens ist beschränkt in jedem* G_a*. Die Kreislinie* $\mathfrak{C}_n = \{\zeta(t) = (n+\frac{1}{2})\pi \cdot e^{it} : \ 0 \le t \le 2\pi\}$ *ist in* G_a *enthalten, wenn* $a < \frac{\pi}{2}$*. Wir wählen* a *klein.*

Für $z \in G_a$ *betrachten wir das Integral der Funktion* $f_z(\zeta) = \frac{1}{\zeta^2 - z^2} \cdot \cot z$ *über die Kurven* $\mathfrak{C}_n$ *mit großem* n*. Der Integrand ist auf den Kurven* $O(n^{-2})$ *und zwar gleichmäßig für* z *in jeder kompakten Teilmenge von* G_a*. Da der Integrationsweg nur die Länge* $c \cdot n$ *hat, ist der Betrag des Integrals* $O(n^{-1})$*.*

Die Funktion $f_z(\zeta)$ *hat Residuen* $(k^2\pi^2 - z^2)^{-1}$ *in den Punkten* $\zeta = k\pi$*; und ausserdem die Residuen* $\frac{1}{2z}\cot z$ *in den Positionen* $\zeta = \pm z$*; denn* $\frac{1}{(\zeta-z)(\zeta+z)} = \frac{1}{2z}\left(\frac{1}{\zeta-z} - \frac{1}{\zeta+z}\right)$.

Im Innern der Kurve $\mathfrak{C}_n$ *liegen die Punkte* $k\pi$ *mit* $|k| \le n$*, und die Punkte* z *mit* $|z| < (n+\frac{1}{2})\pi$*. Für diese* z *ergibt das Umlaufsintegral*

$$\frac{1}{2\pi i}\int_{\mathfrak{C}_n} \frac{\cot\zeta}{\zeta^2 - z^2}\, dz \ = \ \frac{1}{z}\cot z + \sum_{|k|\le n} \frac{1}{k^2\pi^2 - z^2}.$$

Wir multiplizieren mit z *und beachten, dass* k *und* $-k$ *denselben Beitrag liefern.Im Limes erhalten wir die Behauptung.*

Anmerkung: Suggestiv (wiewohl formal bedenklich) ist die folgende 'Umformung' des Resultats

$$\frac{d}{dz}\big(\ln\sin z\big) \ = \ \cot z \ = \ \lim_{n\to\infty} \ln\left(z \cdot \prod_{k=1}^{n}(1 - \tfrac{z^2}{k^2\pi^2})\right).$$

Sie verweist auf verwandte Sätze, die wir hier jedoch nicht genauer betrachten wollen:

$$\begin{aligned}
\frac{\sin \pi z}{\pi z} &= \lim_{n\to\infty}\prod_{1}^{n}\left(1 - \frac{z^2}{k^2}\right), && \text{(Produktdarstellung des Sinus)}\\
-\cot z &= \frac{d}{dz}\big(\ln\Gamma(z) + \ln\Gamma(1-z)\big), &&\\
\frac{\pi}{\sin z} &= \Gamma(z)\cdot\Gamma(1-z) && \text{(Ergänzungssatz)}
\end{aligned}$$

Weitere Beispiele für das Verlegen des Integrationswegs.

Satz 1.5.5 (Die doppeltexponentielle Dichte)**.**

$$\varphi(s) = \int_{-\infty}^{\infty} e^{isx}\cdot \tfrac{1}{2}e^{-|x|}\, dx \ = \ \frac{1}{1+s^2} \quad \textit{für } s \in \mathbb{R}.$$

Beweis. *Das Integral ist reellwertig und hängt nur von* $|s|$ *ab. Wir betrachten für* $s > 0$

$$\varphi_T(s) = 2 \cdot \int_0^T e^{isx} \cdot \tfrac{1}{2} e^{-x}\, dx \quad \textit{und wir zeigen} \quad \lim_{T\to\infty} \big(\Re\varphi_T(s)\big) = \frac{1}{1+s^2}.$$

Dazu bemerken wir: $(1-is)\int_0^T e^{-x(1-is)}\, dx$ *ist das Kurvenintegral der Form* $e^{-z}\, dz$ *entlang der Kurve* $\mathfrak{C}_T^{(s)} = \big\{z(x) = (1-is)\cdot x : \; x \in [0,T]\big\}$, *einem 'Strahl' in der rechten Halbebene. Wir betrachten das Umlaufsintegral über das Dreieck mit den Ecken* 0, T, $(1-is)\cdot T$. *Der Beitrag des 'vertikalen' Stücks zum Integral ist exponentiell klein für* $T \to \infty$. *Man kann daher den Integrationsweg* $\mathfrak{C}_T^{(s)}$ *deformieren zum Weg vom Nullpunkt nach* T *(entlang der reellen Achse) mit einem beliebig kleinen Fehler, wenn* T *groß ist.*

$$\lim_{T\to\infty} \varphi_T(s) \;=\; \int_0^\infty e^{-x(1-is)}\, dx \;=\; \frac{1}{1-is}\int_0^\infty e^{-x}\, dx \;=\; \frac{1}{1-is} \;=\; \frac{1+is}{1+s^2}$$

Satz 1.5.6 (Die Gamma-Dichten)**.**

$$\varphi_\alpha(s) = \int_0^\infty e^{isx} \cdot \frac{1}{\Gamma(\alpha)} \cdot x^{\alpha-1} \cdot e^{-x}\, dx \;=\; (1-is)^{-\alpha} \quad \textit{für } \alpha > 0,\ s \in \mathbb{R}.$$

Beweis. *Bei der* α*-ten Potenz ist natürlich der Hauptwert gemeint; der Wert des Integrals ist* $=1$ *für* $s = 0$*; die Gamma-Dichten sind Wahrscheinlichkeitsdichten auf* $\mathbb{R}_+$.

$$(1-is)^\alpha \cdot \varphi_\alpha(s) = \frac{1}{\Gamma(\alpha)} \cdot \int_0^\infty e^{-x(1-is)} \cdot \big((1-is)x\big)^\alpha \, \frac{dx}{x}.$$

Wir können wie oben den Integrationsweg entlang des Strahls zum Integrationsweg entlang $\mathbb{R}_+$ *deformieren, und erhalten den Wert* $=1$ *für alle* s.

Satz 1.5.7 (Die charakteristische Funktion der Normalverteilung)**.**

$$\int_{-\infty}^\infty e^{isx} \cdot \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\tfrac{1}{2}x^2\right)\, dx \;=\; \exp\!\big(-\tfrac{1}{2}\cdot s^2)\big)$$

Beweis. *Wir wissen bereits* $\quad 1 = \int_{-\infty}^\infty \frac{1}{\sqrt{2\pi}} \cdot \exp\left(-\frac{1}{2}\cdot x^2\right)$. *Nun gewinnen wir*

$$\exp\!\big(\tfrac{1}{2}\cdot s^2)\big) \cdot \int_{-\infty}^\infty e^{isx} \cdot \exp\left(-\tfrac{1}{2}x^2\right)\, dx = \int_{-\infty}^\infty \exp\!\big(\tfrac{1}{2}\cdot (x-is)^2\big)\, dx = \int_{\mathfrak{C}_s} \exp(-\frac{1}{2}z^2)\, dz$$

wo $\mathfrak{C}_s$ *die verschobene reelle Achse ist. Die Verschiebung ändert das Integral nicht, weil für jedes* s *die Integrale über die Kurven von* T *nach* $T - is$ *für großes* $|T|$ *vernachlässigt werden können. Das Integral ist also* $\sqrt{2\pi}$ *für alle* s.

Wir wollen jetzt die Form $\exp(-\frac{1}{2}z^2)\, dz$ auch noch über andere Kurven integrieren. Die unbestimmten Integrale

$$C(T) = \int_0^T \cos t^2\, dt \qquad \text{und} \qquad S(T) = \int_0^T \sin t^2\, dt$$

heissen die Fresnel-Integrale. (Sie treten in der Theorie der Lichtbeugung an einem Spalt auf.) Wir wollen zeigen, dass beide für $T \to \infty$ gegen den Wert $\frac{1}{2}\cdot\sqrt{\frac{\pi}{2}}$ konvergieren.

Satz 1.5.8 (Die Fresnel-Integrale).

$$\lim_{T\to\infty}\int_0^T \cos t^2\ dt \ = \ \frac{1}{2}\cdot\sqrt{\tfrac{\pi}{2}} \ = \ \lim_{T\to\infty}\int_0^T \sin t^2\ dt.$$

Beweis. *Wenn wir die Form* $\exp(-\frac{1}{2}z^2)\ dz$ *vom Nullpunkt ausgehend auf geradem Weg* $\mathfrak{C}_T$ *nach* $T\cdot(1+i)$ *integrieren, erhalten wir das Integral*

$$\int_{\mathfrak{C}_T}\exp(-\frac{1}{2}z^2)\ dz \ = \ \int_0^T \exp(-it^2)\cdot(1+i)\ dt \ = \ (1+i)\cdot\int_0^T\big(\cos t^2 - i\cdot\sin t^2\big)\ dt.$$

Wir wissen bereits $\frac{1}{\sqrt{2\pi}}\int_0^\infty \exp(-\frac{1}{2}x^2)\ dx = \frac{1}{2}$. *Wir beweisen die zur Aussage des Satzes äquivalente Gleichung*

$$\lim_{T\to\infty}\quad (1+i)\cdot\big(C(T)-i\cdot S(T)\big) = \frac{1}{2}\cdot\sqrt{2\pi}$$

indem wir den Integrationsweg verlagern. Wir integrieren über den Rand des Dreiecks mit den Ecken $0,\ T,\ (1+i)T,$ *und wir weisen nach, dass das Integral über die Strecke von* T *nach* $(1+i)T$ *vernachlässigt werden kann. Wenn wir die Kurve parametrisieren* $\big\{z(v) = T + iTv:\quad v\in[0,1]\big\}$, *dann hat dieses Integral die Gestalt*

$$\int_0^1 \exp\big(-\tfrac{1}{2}T^2(1+iv)^2\big)\ iT\ dv \ = \ \int_0^1 \exp\big(-\tfrac{1}{2}T^2(1-v^2)\big)\cdot\big(\cos(T^2v)+i\ \sin(T^2v)\big)\ iT\ dv$$

Sinus und Cosinus sind im Betrag ≤ 1 *und für die Funktion* $h(v) = \exp\big(-\frac{1}{2}T^2(1-v^2)\big)$ *gilt* $\quad \int_0^1 h(v)\ dv \le \exp\big(-\frac{1}{2}T^2)\big)\cdot\int_0^1\exp\big(+\frac{1}{2}T^2v)\big)\ dv = \dfrac{2}{T^2}\big[1-e^{-\frac{1}{2}T^2}\big].$

Wir bemerken $\quad S(T) = \int_0^T \sin t^2\ dt = \int_0^{\sqrt{T}} \sin u\ \frac{1}{2\sqrt{u}}\ du.\quad$ Die Integrale über die u-Intervalle $[(k-1)\pi,\ k\pi]$ sind abwechselnd positiv und negativ und im Betrag fallend. Die Funktion $S(\cdot)$ ist für alle $T>0$ positiv mit dem Limes $S(\infty) = \frac{1}{2}\cdot\sqrt{\frac{\pi}{2}}$. Dasselbe gilt für $C(\cdot)$ (ohne Beweis!). Die Kurve mit den cartesischen Koordinaten $\big(C(T), S(T)\big)$ ist bekannt als die Cornu'sche Spirale. Der Kurvenparameter T kann als die Bogenlänge vom Nullpunkt aus interpretiert werden.

Wir wollen unsere Form $\quad \frac{1}{\sqrt{2\pi}}\cdot\exp(-\frac{1}{2}z^2)\ dz \quad$ von $(1-i)(-Ta)$ nach $(1-i)(+Ta)$ integrieren. Mit der Parametrisierung $\quad \big\{z(v) = (1-i)Tv : v\in[-a,+a]\big\}$ und $T^2 = \frac{1}{\varepsilon}$ erhalten wir

$$\tfrac{1+i}{2} + o(1) = \tfrac{1}{\sqrt{2\pi\varepsilon}}\cdot\int_{-a}^{a}\exp\big(\frac{i}{\varepsilon}v^2\big)\ dv.$$

Durch partielle Integration finden wir weitere Formeln dieser Art:

$$\tfrac{1+i}{2}\cdot\tfrac{\varepsilon}{2i} + o(\varepsilon) = \tfrac{1}{\sqrt{2\pi\varepsilon}}\cdot\int_{-a}^{a}\exp\big(\frac{i}{\varepsilon}v^2\big)\cdot v^2\ dv.$$

Der wesentliche Beitrag zum Integral kommt im Limes von einer beliebig klein gewählte Umgebung des Nullpunkts. Das ist auch so, wenn wir etwas allgemeinere Integranden einführen: **Ausblick:** In der mathematischen Physik kennt man das Prinzip der stationären Phase. Im einfachsten Spezialfall besagt es: Wenn $g(v)$ eine glatte Funktion ist, die ausserhalb eines Intervalls $[-a, +a]$ verschwindet, dann gilt für $\varepsilon \to 0$

$$\tfrac{1}{\sqrt{2\pi\varepsilon}} \cdot \int \exp\Big(\frac{i}{\varepsilon}v^2\Big) \cdot g(v)\ dv = \tfrac{1+i}{2} \cdot \Big[g(0) + \tfrac{1}{4i}\varepsilon \cdot g''(0) + O(\varepsilon^2)\Big].$$

Es gibt diesen Typ von Aussage auch in höheren Dimensionen. Sei $K(v)$ eine zweimal stetig differenzierbare konvexe Funktion in einer Umgebung des Nullpunkts im $\mathbb{R}^d$ mit $K(0) = 0, \quad K'(0) = 0, \quad K''(0)$ nichtsingulär. Es gibt dann eine komplexe Zahl c_σ vom Betrag 1, die nur von der Signatur σ der Hesse-Matrix $K''(0)$ abhängt, sodass gilt

$$\lim_{\varepsilon\to 0} \quad \Big(\frac{1}{\sqrt{2\pi\varepsilon}}\Big)^d \cdot \sqrt{|\det K''(0)|} \cdot \int \exp\Big(\frac{i}{\varepsilon}K(v)\Big) \cdot g(v)\ dv \ = \ c_\sigma \cdot g(0) + O(\varepsilon).$$

(für positivdefinites $K''(0)$ ist $c_\sigma = (\frac{1+i}{\sqrt{2}})^d$.)
Im Problemkreis des Prinzips der stationären Phase interessiert man sich übrigens auch für den Fall, wo $K(\cdot)$ im Gebiet mehrere 'kritische' Punkte aufweist; das sind Punkte P_j, in welchen der Gradient K' verschwindet, (während die Hesse-Matrix nichtsingulär ist). Man findet das ausgeführt in dem (besonders auch wegen den gleichermaßen eleganten und effektiven Einführung in die Geometrie der glatten Mannigfaltigkeiten) sehr empfehlenswerten Lehrbuch

P. Bamberg & S. Sternberg: „*A course in mathematics for students of physics 2*“. Seite 750 ff. Cambridge University Press, 1990.

Anmerkung: Wir haben unsere speziellen bestimmten Integrale hier nur als Rechenaufgaben behandelt. Wir werden später auch Inhaltliches dazu sagen; hier können wir nur einige Andeutungen machen.

1) Wenn $p(x) \geq 0$ mit $\int_{-\infty}^{\infty} p(x)\ dx = 1$, dann nennt man $p(x)\ dx$ eine Wahrscheinlichkeitsdichte (welche absolutstetig ist relativ zum Lebesgue-Maß). Solche Wahrscheinlichkeitsdichten kann man auch durch ihre 'charakterischen Funktionen' beschreiben:

$$\varphi(t) = \int e^{itx} \cdot p(x)\ dx \qquad \text{für } t \in \mathbb{R}.$$

Man gewinnt auf diesem Wege Funktionen, die man positivdefinite Funktionen nennt.

Der Leser sollte sich nicht wundern, wenn da dieselben Rechenbeispiele auftauchen; es gibt nur wenige durch elementare Funktionen dargestellte Dichten, für welche die charakteristische Funktion ebenfalls eine elementare Funktion ist. Inhaltliche Zusammenhänge werden durch diese 'Beispiele' nicht etabliert.

2) Eine wichtige Eigenschaft des Übergangs von den Wahrscheinlichkeitsdichten zu den charakteristischen Funktionen ist die, dass die chakteristische Funktion des Faltungsprodukts zweier Dichten das punktweise Produkt der charakteristischen Funktionen ist.

$$r(x) = \int_{-\infty}^{\infty} p(y) \cdot q(x-y)\ dy;$$
$$\varphi(t) = \int e^{itx} \cdot p(x)\ dx; \qquad \psi(t) = \int e^{itx} \cdot q(x)\ dx,$$
$$\chi(t) = \varphi(t) \cdot \psi(t) = \int e^{itx} \cdot r(x)\ dx$$

3) So, wie es eine Theorie der Fourierkoeffizienten für 2π-periodische quadratintegrable Funktionen gibt, so gibt es auch eine Theorie der Fourier-Integrale für die quadratintegrablen Funktionen auf der reellen Achse (und übrigens auch im $\mathbb{R}^d$). Für glatte schnellabfallende Funktionen $f(x)$ gewinnt man die Fouriertransformierte $\hat{f}(t)$ durch eine Integration und zwar

$$\hat{f}(t) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} e^{-itx} \cdot f(x)\ dx,$$
$$f(x) = \frac{1}{\sqrt{2\pi}} \cdot \int_{-\infty}^{\infty} e^{+itx} \cdot \hat{f}(t)\ dt.$$

(Was hier schnellabfallend und glatt heisst, erläutern wir später im Anschluss an die Integrationstheorie. Dort werden auch besprechen, wie die Fourier-Tranformierte einer beliebigen quadratintegrablen Funktion zu definieren ist.) Auch für die Theorie der Fourier-Integrale bieten sich die oben vorgestellten Beispiele an, wenn man partout elementare Funktionen vorführen will. Das berühmteste Paar ist natürlich das Paar

$$f(x) = \frac{1}{\sqrt{2\pi} \cdot \sigma} \exp\bigl(-\frac{1}{\sigma^2}x^2\bigr); \qquad \hat{f}(t) = \frac{\sigma}{\sqrt{2\pi}} \exp\bigl(-\sigma^2 t^2\bigr).$$

Ein weiteres berühmtes Paar entsteht aus der 'Dreiecksfunktion' $f_a(x) = \frac{1}{2a} \cdot \bigl(1 - \frac{|x|}{2a}\bigr)^+$. Die Fourier-Transformierte ist (bis auf einen Faktor) das kontinuierliche Analogon zum Fejér-Kern:

$$\hat{f}_a(t) = \frac{1}{\sqrt{2\pi}} \cdot \bigl(\tfrac{1}{at}\bigr)^2 \sin^2(at).$$

In der Sprache der Stochastik heisst $f_a(x)\ dx$ die Dreiecksdichte mir der Varianz $\frac{2}{3} \cdot a^2$. Die charakteristische Funktion ist $\varphi_a(t) = \bigl(\frac{1}{at}\bigr)^2 \sin^2(at)$.

Die Wahrscheinlichkeitsdichte $\frac{1}{2\pi}\bigl(\frac{1}{at}\bigr)^2 \sin^2(at)\ dt$ ist zwar für große a auf eine kleine Umgebung des Nullpunkts konzentriert; sie hat aber unendliche Varianz.

2 Maße und Integrale

Die Techniken der Integration haben ihre Wurzeln in den Flächen- und Rauminhaltsberechnungen. In der Renaissancezeit (um 1550 herum) wurden die antiken Integrationsmethoden (aus Büchern von Heron und Archimedes) wieder aufgenommen; Volumen- und Schwerpunktsberechnungen wurden beliebte Themen. Unmittelbar nach den ersten Wegbereitern entstanden die großen Arbeiten von Kepler, Cavalieri und Torricelli, in denen Methoden entwickelt wurden, die schließlich zur Erfindung der Differential- und Integralrechnung führten. Typisch für diese Autoren war die Bereitschaft, die archimedische Strenge zugunsten von solchen Gedankengängen aufzugeben, die häufig auf unstrengen, manchmal 'atomaren' Voraussetzungen beruhten. Genaueres über die alte Geschichte findet man z. B. in den bereits erwähnten Büchern

Dirk J. Struik: *Abriss der Geschichte der Mathematik. VEB Deutscher Verlag der Wissenschaften Berlin 1963.*

C. H. Edwards: *The Historical Development of the Calculus. Springer-Verlag 1979.*

Wenn man heute in der Schulmathematik Integration betreibt, dann konzentriert man sich in der Regel auf die Integration von stetigen Funktionen auf einem Intervall. Eine tragende Rolle spielen dort die unbestimmten Integrale als Stammfunktionen; man pflegt im sog. Fundamentalsatz der Differential- und Integralrechnung die Beziehungen zur Idee der Differenzierbarkeit. — Die Integrationstheorie, die wir hier entwickeln, knüpft an die ältere Auffassung von Integration an, die auf die Zeit vor Newton und Leibniz zurückgeht; das Suchen nach Stammfunktionen ist kein Thema.

Über die bestimmten Integrale lernt man in der Schule: Das Integral ist die Fläche unter der Kurve, wobei aber die Fläche unterhalb der Achse negativ zu zählen ist. Dies ist die Idee des Cauchy-Integrals. Sie wird in diesem Abschnitt sehr weitgehend verallgemeinert. Dabei zielen wir nicht nur auf das klassische Lebesque-Integral, welchem das Lebesgue-Maß zugrundeliegt. Wir studieren Integrale bzgl. allgemeinerer (vorzüglich normierter) Maße; im einfachsten Fall sind das die sog. Stieltjes-Integrale.

Ein auffälliger Unterschied zur Integration im Sinne der Schulmathematik besteht darin, dass über abstrakte Grundräume (auf welchen ein Maß gegeben ist) integriert wird. Wir bleiben nicht bei Maßen über dem $\mathbb{R}^n$. Die Grundräume sind messbare Räume.

Den Begriff des messbaren Raums wollen wir im Folgenden abstrakt und (bis zu einem gewissen Grad) parallel zum allgemeinen Begriff des topologischen Raums entwickeln. Die Besonderheiten, die sich daraus ergeben, dass der Grundraum möglicherweise auch einmal ein kompakter Raum ist oder ein vollständiger metrischer Raum, werden wir in dieser Einführung nicht in die Tiefe verfolgen.

Wir haben vor allem diejenige 'abstrakte' Maß- und Integrationstheorie im Auge, die in der modernen Stochastik benötigt, gepflegt und weiterentwickelt wird. In dieser Theorie werden die Maßberechnungen als die Berechnungen von Wahrscheinlichkeiten interpretiert; und die Werte von Integralen werden als die Erwartungswerte von Zufallsgrößen gedeutet. Wir überlassen es aber der Grundvorlesung zur Stochastik, den Begriffe der Zufallsgrößen und den Begriff der Wahrscheinlichkeitsbewertung zu entwickeln.

2.1 Die Daniell-Fortsetzung eines Elementarintegrals

Eine stetige Funktion auf dem abgeschlossenen Einheitsintervall $[0,1]$ ist gleichmäßig stetig und beschränkt. Jeder solchen Funktion f ordnet man (seit Cauchy) ihr ('bestimmtes') Integral zu:

$$\tilde{I}(f) = \int_0^1 f(x)\ dx = \int f(x)\ dx.$$

Wir wollen an dieser Stelle nicht über Untersummen und Obersummen reden. Das wohldefinierte Funktional $\tilde{I}(\cdot)$ auf dem Vektorverband $\tilde{\mathfrak{E}} = \mathcal{C}\big([0,1]\big)$ ist unser Ausgangspunkt. Wir nennen dieses Funktional $\tilde{I}(\cdot)$ das **Cauchy-Integral.** – Das klassische Lebesgue-Integral über dem Einheitsintervall wird sich durch eine Fortsetzung ergeben; auf eine andere beliebte Fortsetzung, das sog. Riemann-Integral werden wir nicht speziell eingehen. Diese und andere Fortsetzungen des Cauchy-Integrals unterscheiden sich nur durch ihren Definitionsbereich; wenn eine Funktion im Durchschnitt der Definitionsbereiche liegt, dann stimmen die Integralwerte für alle die gebräuchlichen Fortsetzungen überein.

Konstruktion des Cauchy-Stieltjes-Integrals:
Es sei $F(\cdot)$ eine monotone Funktion auf $[0,1]$ mit $F(0) = 0,\ F(1) = 1$, die rechtsseitig stetig ist. (Man nennt eine solche Funktion die Verteilungsfunktion eines Wahrscheinlichkeitsmaßes auf $[0,1]$).

Es sei $\big(\mathfrak{Z}: \quad 0 = t_0 < t_1 < \cdots < t_N = 1\big)$ eine Unterteilung des Einheitsintervalls und $f(\cdot)$ eine stetige Funktion, also $f \in \tilde{\mathfrak{E}}$. Wir definieren dann

$$\int_0^1 f\ dF \ = \lim_{n\to\infty} \sum_k f(t_{k-1}^{(n)}) \cdot \left(F\left(t_k^{(n)}\right) - F\left(t_{k-1}^{(n)}\right)\right),$$

wo der Limes über eine Folge von Unterteilungen, deren Feinheitsgrad nach 0 strebt, zu erstrecken ist. Der Limes heisst das Stieltjes-Integral der Funktion f bzgl. F. (oder genauer bzgl. des durch F gegebenen Wahrscheinlichkeitsmaßes.) Das Funktional

$$\tilde{I}(\cdot): \quad \tilde{\mathfrak{E}} \ni\ f \longmapsto \tilde{I}(f) = \int_0^1 f\ dF$$

heisst das Cauchy-Stieltjes-Integral bzgl. F.

Die Existenz des Limes ergibt sich aus der gleichmäßigen Stetigkeit von f. Zu jedem $\varepsilon > 0$ existiert ein $\delta > 0$, sodass gilt: Wenn die Zerlegung die Feinheit $< \delta$ hat, dann gilt

$$f(x) = \sum_1^N \tilde{f}(t_{i-1}) \cdot 1_{(t_{i-1},t_i]}(x)\ +\ R_\varepsilon(x) \quad \text{mit } |R_\varepsilon(x)| < \varepsilon \text{ für } x \in (0,1].$$

und die Summe $\sum_k f(t_{k-1}^{(n)}) \cdot \left(F\left(t_k^{(n)}\right) - F\left(t_{k-1}^{(n)}\right)\right)$ liegt bis auf ε bei $\int_0^1 f\ dF$.

Hinweis: Das Kurvenintegral einer 1-Form $f\,dg$ über eine rektifizierbare Kurve, welches wir früher definiert haben, benötigt eine kleine Verallgemeinerung der Konstruktion. An die Stelle der monotonen Funktion tritt dort die Differenz zweier monotoner Funktionen; man spricht von einer Funktion mit beschränkter Schwankung. Diese Verallgemeinerung heisst die Integration bzgl. einer **Ladungsverteilung**; wir wollen sie hier nicht weiter verfolgen. Die Monotonie von F wird für das Folgende wesentlich sein.

Sprechweise. Ein Vektorverband reellwertiger Funktionen $\tilde{\mathfrak{E}}$ (auf irgendeiner Grundmenge Ω) wird ein Stone'scher Verband genannt, wenn mit f auch $f \wedge 1$ zu $\tilde{\mathfrak{E}}$ gehört.

Definition 2.1 (Elementar-Integral)**.**
Ein reellwertiges Funktional $\tilde{I}(\cdot)$ auf einem Stone'schen Verband $\tilde{\mathfrak{E}}$ heisst ein Elementarintegral auf $\tilde{\mathfrak{E}}$, wenn gilt

1. $f \leq g \quad \Rightarrow \quad \tilde{I}(f) \leq \tilde{I}(g)$; (Monotonie)
2. $\tilde{I}(f+g) \ = \ \tilde{I}(f) + \tilde{I}(g), \qquad \tilde{I}(\alpha \cdot f) = \alpha \cdot \tilde{I}(f)$ für alle $\alpha \in \mathbb{R}$; (Linearität)
3. $f_n \searrow 0 \quad \Rightarrow \quad \tilde{I}(f_n) \searrow 0.$ (Monotone Stetigkeit)

(Der Pfeil $\searrow$ bedeutet die punktweise absteigende Konvergenz. Wir notieren manchmal auch $\downarrow$. Entsprechend sind die Symbole $\nearrow$ und $\uparrow$ zu verstehen.)

Die fundamentale Bedeutung der dritten Eigenschaft, der monotonen Stetigkeit, war Cauchy nicht bewusst. Sie erwies sich erst in der Integrationstheorie nach Lebesgue (1902). Für die Cauchy-Stieltjes-Integrale ergibt sie sich aus dem Satz von Dini. Nach diesem Satz impliziert nämlich die monotone Konvergenz stetiger Funktionen gegen die Nullfunktion auf einem kompakten Grundraum die gleichmäßige Konvergenz; und die Stetigkeit bei gleichmäßiger Konvergenz liegt auf der Hand.

Die monotone Stetigkeit liefert die Grundlage für die Fortsetzung des Elementarintegrals auf größere Funktionenklassen. Entscheidend ist das

Satz 2.1.1. *Sei* $\tilde{I}(\cdot)$ *ein Elementarintegral auf* $\tilde{\mathfrak{E}}$*. Es gilt dann*

$$f_n \uparrow f', \quad g_n \uparrow g', \quad f' \geq g' \quad \Longrightarrow \quad \lim \uparrow \tilde{I}(f_n) \geq \lim \uparrow \tilde{I}(g_n);$$
$$f_n \uparrow f', \quad g_n \downarrow g'', \quad f' \geq g'' \quad \Longrightarrow \quad \lim \uparrow \tilde{I}(f_n) \geq \lim \downarrow \tilde{I}(g_n).$$

Man beachte, dass die punktweisen Limiten f', g' *und* g'' *im Allgemeinen nicht zu* $\tilde{\mathfrak{E}}$ *gehören. Im Falle des Cauchy-Integrals handelt es sich um unterhalbstetige bzw. oberhalbstetige Funktionen.*

Beweis *Wir benützen wie üblich die Bezeichnung* $h\backslash k = (h-k)^+ = h - h \wedge k$.
In der ersten Situation steigt die Folge $(f_n)_n$ *über jedes* g_k. *Die Folge* $(g_k \backslash f_n)_n$ *steigt ab nach* 0. *Also gilt* $\tilde{I}(g_k \backslash f_n) < \varepsilon$ *für genügend großes* n.

$$\begin{aligned} \tilde{I}(f_n) - \tilde{I}(g_k) &= \tilde{I}(f_n \backslash g_k) - \tilde{I}(g_k \backslash f_n) > 0 - \varepsilon; \\ \lim \uparrow \tilde{I}(f_n) &> \tilde{I}(g_k) - \varepsilon \quad \text{für alle } \varepsilon > 0; \\ \lim \uparrow \tilde{I}(f_n) &\geq \tilde{I}(g_k) \quad \text{für alle } k. \end{aligned}$$

In der zweiten Situation ist die Funktionenfolge $g_n \backslash f_n$ *absteigend gegen die Nullfunktion. Für große* n *gilt also* $\tilde{I}(g_n \backslash f_n) < \varepsilon$ *und*

$$\begin{aligned} \tilde{I}(g_n) - \tilde{I}(f_n) &= \tilde{I}(g_n \backslash f_n) - \tilde{I}(f_n \backslash g_n) < \varepsilon; \\ \lim \downarrow \tilde{I}(g_n) - \lim \uparrow \tilde{I}(f_n) &< \varepsilon \quad \text{für alle } \varepsilon > 0. \end{aligned}$$

Satz 2.1.2 (Konstruktion des Verbandskegels $\mathfrak{E}^\uparrow$).
Es sei $\tilde{\mathfrak{E}}$ *ein Stone'scher Vektorverband (über irgendeiner Grundmenge* Ω*); (Die Elemente* $f \in \tilde{\mathfrak{E}}$ *nennen wir die elementaren Funktionen.)*
Und es sei $\mathfrak{E}^\uparrow$ *die Menge derjenigen Funktionen, die man als Limes einer aufsteigenden Folge elementarer Funktionen gewinnen kann. Es gilt dann*

1. $f, g \in \mathfrak{E}^\uparrow \implies f+g, \quad f \wedge g, \quad f \vee g \quad \in \mathfrak{E}^\uparrow,$
2. $f \in \mathfrak{E}^\uparrow,\ \alpha \in \mathbb{R}_+ \implies \alpha \cdot f \in \mathfrak{E}^\uparrow$
3. $f_1 \leq f_2 \leq \cdots$ *mit* $f_n \in \mathfrak{E}^\uparrow \implies f = \lim \uparrow f_n \in \mathfrak{E}^\uparrow$

Man sagt: $\mathfrak{E}^\uparrow$ ist ein Verbandskegel, der aufsteigend-σ-vollständig ist. Man beachte, dass die Funktionen in $\mathfrak{E}^\uparrow$ den Wert $+\infty$ aber nicht den Wert $-\infty$ annehmen können.

Beispiel. Die elementaren Funktionen seien die stetigen Funktionen auf einem metrisierbaren Raum. Die Funktionen f in $\mathfrak{E}^\uparrow$ sind dann die unterhalbstetigen Funktionen, die eine stetige Minorante besitzen. Die punktweisen Suprema stetiger Funktionen sind nämlich unterhalbstetig; und in einem metrisierbaren Raum kann man jede unterhalbstetige Funktion, die eine stetige Minorante besitzt, als aufsteigenden Limes einer Folge stetiger Funktionen darstellen. Das sieht man folgendermaßen:

1. Die Indikatorfunktion einer offenen Menge U kann man als aufsteigenden Limes stetiger Funktionen darstellen. Oder, äquivalent dazu: die Indikatorfunktion einer abgeschlossenen Menge F kann als absteigender Limes stetiger Funktionen gewonnen werden:
$$1_F(\cdot) = \lim \downarrow \left(1 - n \cdot d(\cdot, F)\right)^+,$$
wo $d(\cdot, F)$ den Abstand von der Menge F bezeichnet.

2. Wenn $g_1, g_2, \dots$ nichtnegative Funktionen sind, die als aufsteigender Limes nichtnegativer stetiger Funktionen dargestellt sind, dann kann man auch die Summe $\sum_1^\infty g_k$ und das Supremum $\bigvee_1^\infty g_k$ so darstellen. Zu $g_k = \lim_n \uparrow f_k^{(n)}$ betrachte nämlich
$$h^{(n)} = f_0^{(n)} + f_1^{(n-1)} + \cdots + f_n^{(0)}, \qquad \text{bzw.} \qquad k^{(n)} = f_0^{(n)} \vee f_1^{(n-1)} \vee \cdots \vee f_n^{(0)}.$$
Die $h^{(n)}$ streben aufsteigend gegen die Summe, die $k^{(n)}$ gegen das Supremum.

3. Wir wollen hier zunächst einmal nur die nichtnegativen unterhalbstetigen f aufsteigend approximieren. Für $n = 1, 2, \dots$ und $k = 0, 1, 2, \dots$ sei $f_k^{(n)}$ die Indikatorfunktion der offenen Menge $\{f > \frac{k}{2^n}\}$, und $g^{(n)} = \frac{1}{2^n} \sum_k f_k^{(n)}$.
Offenbar ergibt sich $g^{(n)}(\omega)$ aus $f(\omega)$ durch Abrunden auf das nächste ganzzahlige Vielfache von $\frac{1}{2^n}$. Die $g^{(n)}$ liefern eine aufsteigende Approximation von f.

4. Wenn h unterhalbstetig ist mit der stetigen Minorante $\tilde{h}$, dann approximieren wir $h - \tilde{h}$ und addieren zu den approximierenden Funktionen die stetige Funktion $\tilde{h}$.

Satz 2.1.3 (Aufsteigende Fortsetzung).
Jedes Elementarintegral $\tilde{I}(\cdot)$ *auf* $\tilde{\mathfrak{E}}$ *besitzt genau eine aufsteigend-*σ*-stetige Fortsetzung. Diese Fortsetzung* $I^\uparrow(\cdot)$ *ist ausserdem isoton, additiv und positivhomogen, d. h.*

1. $f \leq g \quad \Rightarrow \quad I^\uparrow(f) \leq I^\uparrow(g)$;

2. $I^\uparrow(f + g) \;=\; I^\uparrow(f) + I^\uparrow(g), \qquad I^\uparrow(\alpha \cdot f) = \alpha \cdot I^\uparrow(f)$ *für alle* $\alpha \in \mathbb{R}_+$;

3. $f_1 \leq f_2 \leq \cdots \quad \textit{mit} \quad f_n \in \mathfrak{E}^\uparrow \quad \Longrightarrow \quad I^\uparrow(\lim \uparrow f_n) = \lim \uparrow I^\uparrow(f_n)$

Beweis Wir haben bereits gesehen, dass für zwei aufsteigende Folgen elementarer Funktionen die Integralwerte gegen denselben Grenzwert streben, und dass das so gewonnene Funktional $I^\uparrow(\cdot)$ isoton ist. Man beachte, dass es auch den Wert $+\infty$ annehmen kann. Die Additivität und die positive Homogenität liegen auf der Hand. Sei (f_n) eine aufsteigende Folge mit $\lim \uparrow f_n = f$. Die $f_n \in \mathfrak{E}^\uparrow$ seien als aufsteigende Limiten gegeben: $f_n = \lim_m \uparrow \tilde{f}_{nm}$. Die Folge der elementaren Funktionen $\tilde{k}_n = \tilde{f}_{1n} \vee \cdots \vee \tilde{f}_{nn}$ ist aufsteigend mit $\tilde{k}_n \leq f_n$ und $\lim \uparrow \tilde{k}_n = f$. Wir haben also $f \in \mathfrak{E}^\uparrow, I^\uparrow(f) = \lim \uparrow \tilde{I}(\tilde{k}_n) = \lim \uparrow I^\uparrow(f_n)$.
Wir betrachten jetzt auch den zu $\mathfrak{E}^\uparrow$ diametralen Kegel $\mathfrak{E}^\downarrow = -\mathfrak{E}^\uparrow = \left\{g = -f : \ f \in \mathfrak{E}^\uparrow\right\}$. Es handelt sich um einen Verbandskegel, der gegenüber absteigenden Limiten abgeschlossen ist. Interessant sind nun diejenigen Funktionen, die sich 'knapp' einschliessen lassen zwischen eine Majorante aus $\mathfrak{E}^\uparrow$ und eine Minorante aus $\mathfrak{E}^\downarrow$.

Definition 2.2 (Daniell-Integrabilität).
Es sei $\tilde{I}(\cdot)$ ein Elementarintegral. Eine Funktion h heisst $\tilde{I}(\cdot)$-Daniell-integrabel, wenn gilt
$$\forall \varepsilon < 0 \; \exists \, f \in \mathfrak{E}^\uparrow, g \in \mathfrak{E}^\downarrow : \quad \big(g \leq h \leq f\big) \;\wedge\; \big(I^\uparrow(f) - I^\downarrow(g) < \varepsilon\big).$$

Die Menge aller Daniell-Integrablen Funktionen bezeichnen wir mit $\mathfrak{E}^*$; und für $h \in \mathfrak{E}^*$ setzen wir

$$I^*(h) = \inf\{I^\uparrow(f) : f \geq h\} = \sup\{I^\downarrow(g) : g \leq h\}.$$

Satz 2.1.4.
Die Daniell-Fortsetzung übernimmt Eigenschaften des Elementarintegrals. Es gilt

1. *Für jede elementare Funktion* $\tilde{f}$ *gilt* $I^*(\tilde{f}) = \tilde{I}(\tilde{f})$.
2. $h \in \mathfrak{E}^*, \alpha \in \mathbb{R} \implies \alpha h \in \mathfrak{E}^*, I^*(\alpha h) = \alpha I^*(h)$
3. $h^{(1)}, h^{(2)} \in \mathfrak{E}^* \implies h^{(1)} \vee h^{(2)}, h^{(1)} \wedge h^{(2)} \in \mathfrak{E}^*$;
4. $h = h^{(1)} + h^{(2)}$ *mit* $h^{(1)}, h^{(2)} \in \mathfrak{E}^* \implies h \in \mathfrak{E}^*, I^*(h) = I^*(h^{(1)} + h^{(2)})$.
5. $h^{(1)} \leq h^{(2)} \leq \cdots$ *mit* $h^{(n)} \in \mathfrak{E}^*,\ \lim \uparrow I^*(h^{(n)}) < \infty$
$$\implies h = \lim \uparrow h^{(n)} \in \mathfrak{E}^*,\ I^*(h) = \lim \uparrow I^*(h^{(n)}).$$
(' Satz von der monotonen Konvergenz ')

Beweis. *Die beiden ersten Aussagen des Satzes sind offensichtlich.*
Wir bemerken: Die Integrabilitätsbedingung kann man auch formulieren, ohne von den Konstrukten $I^\uparrow(\cdot), I^\downarrow(\cdot)$ *expliziten Gebrauch zu machen:*

$$h \in \mathfrak{E}^* \iff \forall \varepsilon < 0\ \exists (\tilde{f}_n), (\tilde{g}_n) :\ \lim \uparrow \tilde{f}_n \geq h \geq \lim \downarrow \tilde{g}_n, \quad \lim \uparrow \tilde{I}(\tilde{f}_n \backslash \tilde{g}_n) < \varepsilon.$$

Wir beweisen die Aussagen 3) und 4):
Zu $h^{(1)}, h^{(2)} \in \mathfrak{E}^*$ *und* $\varepsilon < 0$ *wählen wir monotone Folgen elementarer Funktionen mit*

$$\lim \uparrow \tilde{f}_n^{(i)} \geq h^{(i)} \geq \lim \downarrow \tilde{g}_n^{(i)}, \quad \lim \uparrow \tilde{I}(\tilde{f}_n^{(i)} \backslash \tilde{g}_n^{(i)}) < \varepsilon/2.$$

Eine Einschließung bis auf ε *für* $h^{(1)} \vee h^{(2)}$ *wird geleistet von den monotonen Folgen* $\tilde{f}_n = \tilde{f}_n^{(1)} \vee \tilde{f}_n^{(2)}, \quad \tilde{g}_n = \tilde{g}_n^{(1)} \vee \tilde{g}_n^{(2)}$. *Analog verfahren wir für die Summe. Das Minimum erledigt sich wie das Maximum, wenn wir am Nullpunkt spiegeln.*
Für den Beweis der fünften Aussage, dem sog. Satz von der monotonen Konvergenz, wählen wir zu jedem $h^{(n)}$ *Funktionen* $f^{(n)} \in \mathfrak{E}^\uparrow$ *und* $g^{(n)} \in \mathfrak{E}^\downarrow$ *mit*

$$f^{(n)} \geq h^{(n)} \geq g^{(n)}, \qquad I^\uparrow(f^{(n)}) - I^\downarrow(g^{(n)}) < \tfrac{1}{2^n} \cdot \varepsilon.$$

Die Funktion $f' = \lim \uparrow (f^{(1)} \vee \cdots \vee f^{(n)})$ *dominiert den aufsteigenden Limes* $\lim \uparrow h^{(n)}$.
Andererseits: Für genügend großes N *erfüllt* $g'' = g^{(1)} \vee \cdots \vee g^{(N)} \in \mathfrak{E}^\downarrow$ *die Bedingungen*

$$f' \geq h \geq g'', \qquad I^\uparrow(f') - I^\downarrow(g'') < \varepsilon.$$

Bei den Rechnungen könnte es übrigens nützlich sein, wenn man den folgenden elementaren **Hilfssatz** im Auge hat: Sind a, b, c nichtnegative Zahlen oder nichtnegative Funktionen, dann gelten für die Operationen $\wedge, \backslash, \vee$ die Rechenregeln

$$\begin{aligned} &a = a \wedge b + a \backslash b, \qquad a \vee b = a + b \backslash a, \quad (a \backslash b) \backslash c = a \backslash (b + c) \\ &(a + b) \backslash c = a \backslash c + b \backslash (c \backslash a), \quad (a + b) \wedge c = a \wedge c + b \wedge (c \backslash a) \\ &(a \vee b) \backslash c \quad = \big(a + b \backslash a\big) \backslash c = a \backslash c + b \backslash \big(a \vee c\big) \quad \leq a \backslash c + b \backslash c \end{aligned}$$

Der Fortsetzungssatz von Daniell hat eine Schwäche, die wir nur durch sehr behutsames einigermaßen kaschieren konnten: Die Daniell-integrierbaren Funktionen sind **numerische Funktionen** in dem Sinn, dass sie Werte in $\mathbb{R} \cup \{\pm\infty\}$ annehmen. Die Menge $\mathfrak{E}^*$ der Daniell-integrierbaren Funktionen ist kein Vektorraum; die punktweise Addition ist in denjenigen Punkten ω nicht definiert, wo der eine Summand den Wert $+\infty$ und der anderen den Wert $-\infty$ annimmt.

Gleichheit fast überall

Die Schwierigkeiten können dadurch überwunden werden, dass man zu Äquivalenzklassen von Daniell-integrablen Funktionen übergeht: zwei Funktionen werden als äquivalent erklärt, wenn sie sich nur auf einer 'Nullmenge' unterscheiden (man sagt, 'wenn sie fastüberall gleich sind'). Für die 'Nullmengen' und die 'Gleichheit fastüberall' (notiert $f \sim g$ f.ü.) muß offenbar das Folgende geklärt werden:

1. Zu jeder Daniell-integrablen Funktion gibt es äquivalente endlichwertige Daniell-integrablen Funktionen mit demselben Integral.
2. Die Maximums- und Minimumsbildung sowie die Vektorraumoperationen sind verträglich mit der Äquivalenzrelation
3. $f_i \sim g_i$ für $i = 1, 2, \ldots \quad \Longrightarrow \quad \sup f_i = \sup g_i, \quad \inf f_i = \inf g_i,$

Mit dem Begriff der **Nullmenge** werden wir uns in der allgemeinen Theorie der Messbarkeit noch ausführlich zu beschäftigen haben. Hier geht es zunächst einmal etwas spezieller um die Nullfunktionen und Nullmengen für ein gegebenes Elementarintegral.

Definition 2.3.

Eine numerische Funktion $h^\star$ heißt **Nullfunktion** für das Elementarintegral $\tilde{I}(\cdot)$ auf $\tilde{\mathfrak{E}}$, wenn gilt

$$\forall \varepsilon > 0 \, \exists \, f \in \mathfrak{E}^\uparrow \ : |h^\star| \leq f \ \text{ und } \ I^\uparrow(f) < \varepsilon \, .$$

Eine Teilmenge N des Grundraums heißt **Nullmenge** bzgl. $\tilde{I}(\cdot)$, wenn die Indikatorfunktion 1_N eine Nullfunktion ist.

Hinweis: Wir interessieren uns eigentlich nur für bairesche Nullfunktionen und bairesche Nullmengen. Wir wollen die Baire-Eigenschaft aber nicht immer erwähnen, wenn sie für die Schlüsse keine Bedeutung hat.

Satz 2.1.5.
Die Vereinigung von abzählbar vielen Nullmengen ist eine Nullmenge. Das Supremum von abzählbar vielen Nullfunktionen ist eine Nullfunktion.

Beweis.
Es genügt offenbar, nichtnegative Nullfunktionen zu betrachten.
Wir bemerken : Wenn $h_1^\star$ *und* $h_2^\star$ *Nullfunktionen sind, dann auch* $h_1^\star \vee h_2^\star$.
Wir zeigen : Wenn $h_1^\star \le h_2^\star \le \ldots$ *eine aufsteigende Folge von nichtnegativen Nullfunktionen ist, dann ist auch* $h^\star := \lim \uparrow h_j$ *eine Nullfunktion.*
Wir wählen $f_j \in \mathfrak{E}^\uparrow$ *so dass*

$$h_j^\star \le f_j \quad \textit{und} \quad I(f_j) \le \frac{1}{2^j} \cdot \varepsilon \ .$$

Für die aufsteigende Folge $g_j = f_1 \vee f_2 \vee \ldots \vee f_j$ *haben wir*

$$\lim \uparrow g_j \ge h^\star \quad \textit{und} \quad I^\uparrow(\lim \uparrow g_j) \le \varepsilon \ .$$

Satz 2.1.6.
Eine Funktion $h^\star$ *ist genau dann eine Nullfunktion, wenn* $\{\omega \ : h^\star(\omega) \neq 0\}$ *eine Nullmenge ist.*

Beweis.
Sei $N := \{\omega \ : |h^\star(\omega)| > 0\}$ *eine Nullmenge. Für jedes* $n \in \mathbb{N}$ *ist dann* $n \cdot 1_N$ *eine Nullfunktion. Der aufsteigende Limes dieser Nullfunktionen ist eine Nullfunktion* $\ge h^\star$. *Sei* $|h^\star|$ *eine Nullfunktion.* $(|n \cdot h^\star| \wedge 1)_{n\in\mathbb{N}}$ *ist eine Folge von Nullfunktionen, welche gegen* 1_N *aufsteigt.*

Daniell-Maß
Einen zweiten Ausweg aus dem Problem mit den unendlichen Funktionswerten der zu integrierenden Funktionen liefert der Übergang zu dem Maß zu einem Daniell-Integral. Der Übersichtlichkeit halber wollen wir im Folgenden annehmen, dass unser Elementarintegral normiert ist in dem Sinne, dass die Konstanten Elementarfunktionen sind mit $\tilde{I}(1) = 1$. Das Daniell-Integral führt in diesem Fall zu einem Wahrscheinlichkeitsmaß im Sinne der folgenden Definitionen

Definition 2.4 (σ-Algebra)**.**
Ein System $\mathfrak{A}$ von Teilmengen einer Grundmenge Ω heißt eine **Mengenalgebra** wenn

i) $\emptyset \in \mathfrak{A}\ , \qquad \Omega \in \mathfrak{A},$

ii) $A \in \mathfrak{A} \quad \Longrightarrow \quad \Omega \backslash A \in \mathfrak{A}$

iii) $A, B \in \mathfrak{A} \quad \Longrightarrow \quad A \cap B \in \mathfrak{A} \quad \big(\text{und} \quad A \cup B \in \mathfrak{A}\big).$

vi) Eine Mengenalgebra $\mathfrak{A}$ über Ω heißt eine σ-Algebra, wenn zusätzlich gilt

$$A_1, A_2, \ldots \in \mathfrak{A} \quad \Longrightarrow \quad \bigcup^{\infty} A_n \in \mathfrak{A} \quad \Big(\text{und} \quad \bigcap^{\infty} A_n \in \mathfrak{A}\Big).$$

Definition 2.5 (Maß, Wahrscheinlichkeitsmaß)**.**
Ein Maß ist eine nichtnegative Funktion auf einer σ-Algebra $\mathfrak{A}$ mit den Eigenschaften

i) $\mu(\emptyset) = 0\,; \qquad \mu(A) \geq 0$ für alle $A \in \mathfrak{A}$,

ii) $A \subseteq B \quad \Longrightarrow \quad \mu(A) \leq \mu(B)$

iii) $A_1, A_2, \ldots$ paarweise disjunkt $\quad \Longrightarrow \quad \mu\left(\sum_1^\infty A_i\right) = \sum_1^\infty \mu(A_i)$.

Wenn $\mu(\Omega) = 1$, dann spricht man von einem **Wahrscheinlichkeitsmaß**.
Man spricht auch von einer normierten nichtnegativen σ-additiven Mengenfunktion. Wir schreiben auch kurz W-Maß.

Definition 2.6. Es sei $\mathrm{I}^*(\cdot)$ die Daniell-Fortsetzung eines Elementarintegrals über der Grundmenge Ω. Eine Teilmenge $A \subset \Omega$ heisst Daniell-meßbar (bzgl. $\mathrm{I}^*(\cdot)$), wenn die Indikatorfunktion Daniell-integrierbar ist. In diesem Fall heisst $\mu(A) = \mathrm{I}^*(1_A)$ das Daniell-Maß der Menge A.

Satz 2.1.7. *Wenn* $\mathrm{I}^*(\cdot)$ *das Daniell-Integral zu einem normierten Elementarintegral ist, dann ist die Gesamtheit* $\mathfrak{A}^*$ *aller Daniell-meßbaren Mengen eine* σ*-Algebra und* $\mu(\cdot)$ *ist ein Wahrscheinlichkeitsmaß auf* $\mathfrak{A}^*$*.*

Der Beweis ist trivial. Wir wollen uns noch ein wenig mit Definitionsbereichen befassen. Die σ-Algebra $\mathfrak{A}^*$ der Daniell-messbaren Mengen hängt von dem Elementarintegral ab. Dies wird in der Maßtheorie als störend empfunden; man arbeitet lieber mit einer σ-Algebra, die allein vom Vektorverband $\tilde{\mathfrak{E}}$ abhängt.

Definition 2.7 (Baire'sche Mengen)**.**
$\tilde{\mathfrak{E}}$ sei ein Stone'scher Vektorverband über Ω. $\mathfrak{E}^\beta$ sei die kleinste Menge von numerischen Funktionen, welche $\tilde{\mathfrak{E}}$ umfaßt und gegenüber der Bildung abzählbarer Suprema und Infima abgeschlossen ist. Die Elemente $f \in \mathfrak{E}^\beta$ heißen die **Baire-Funktionen** zu $\tilde{\mathfrak{E}}$.
Die Mengen $B \subseteq \Omega$, für welche 1_B eine Baire-Funktion ist, heißen die **Baire'schen Mengen** zu $\tilde{\mathfrak{E}}$.

Wir bemerken: Für jede elementare $\tilde{f}$ und jedes reelle a sind die Mengen $\{\tilde{f} > a\}$ und $\{\tilde{f} \geq a\}$ Baire'sche Mengen. In der Tat gilt

$$\big(n(\tilde{f} - \tilde{f} \wedge a)\big) \wedge 1 \nearrow 1_{\{\tilde{f} > a\}}.$$

Man zeigt leicht den Satz

Satz 2.1.8. *Die Gesamtheit* $\mathfrak{B}$ *aller Baire'schen Mengen zum Stone'schen Verband* $\tilde{\mathfrak{E}}$ *ist der kleinste* σ*-Ring, welcher alle Mengen der Gestalt* $\{\omega : \tilde{f}(\omega) > 0\}$ *mit* $\tilde{f} \in \tilde{\mathfrak{E}}$ *enthält. Eine numerische Funktion* $h(\cdot)$ *auf* Ω *ist genau dann eine Baire-Funktion, wenn*

$$\{\omega : h(\omega) \leq a\} \in \mathfrak{B} \quad \textit{für alle } a \in \mathbb{R} .$$

Beispiel (Baire'sche und Borel'sche Teilmengen eines HRaB).

Sei $\tilde{\mathfrak{E}}$ die Menge aller beschränkten stetigen Funktionen auf einem metrisierbaren HRaB. $\mathfrak{E}^{\uparrow}$ ist dann die Menge aller unterhalbstetigen Funktionen mit einer stetigen Minorante. Eine Indikatorfunktion 1_A gehört genau dann zu $\mathfrak{E}^{\uparrow}$, wenn A offen ist. Die Baire'schen Mengen sind in diesem Fall die borelschen Teilmengen des Raums S. In allgemeinen Hausdorff- Räumen definiert man die Borel-Algebra als die von den offenen Mengen erzeugte σ-Algebra; die Elemente der Borel-Algebra heissen die Borel-Mengen. In topologischen Räumen, die nicht metrisierbar sind, kann es Borel-Mengen geben, die nicht baire'sch sind.

Satz 2.1.9. *Für jede nichtnegative (bzgl.* $\tilde{\mathfrak{E}}$*) Baire'sche Funktion* f *ergibt sich das Daniell-Integral als das Integral einer fallenden Funktion auf* $\mathbb{R}_+$*:*

$$I^*(f) = \int_0^\infty \mu\{\omega : f(\omega) > a\}\, da,$$

Beweis. *Für die Indikatorfunktion* $f_k^{(n)}$ *der Menge* $\{\omega : f(\omega) > \frac{k}{2^n}\}$ *haben wir* $I^*(f_k^{(n)}) = \mu\{\omega : f(\omega) > \frac{k}{2^n}\}$*. Wie oben beim Studium der unterhalbstetigen Funktionen auf einem metrisierbaren Raum betrachten wir die auf das kleinste ganze Vielfache von* $\frac{1}{2^n}$ *abgerundeten Funktionen* $g^{(n)} = \frac{1}{2^n}\sum_0^\infty f_k^{(n)}$*. Die Funktionen* $\bar{F}(a) = \mu\{\omega : f(\omega) > a\}$ *und* $\bar{G}^{(n)}(a) = \mu\{\omega : g^{(n)}(\omega) \geq a\}$ *sind monoton fallend auf* $\mathbb{R}_+$ *und für* $a = \frac{k}{2^n}$ *gilt* $G^{(n)}(a) = \bar{F}(a)$*. Die Funktionenfolge* $G^{(n)}(\cdot)$ *konvergiert aufsteigend gegen* $\bar{F}(\cdot)$ *in allen Stetigkeitspunkten von* $\bar{F}(\cdot)$*. Die Integrale* $\int_0^\infty \bar{G}^{(n)}(a)\, da$ *konvergieren ebenfalls*

$$\frac{1}{2^n}\sum_0^\infty I^*(f_k^{(n)}) = I^*(g^{(n)}) = \int_0^\infty \bar{G}^{(n)}(a)\, da \quad \nearrow \quad \int_0^\infty \bar{F}(a)\, da = I^*(f).$$

Das Lebesgue-Integral als Spezialfall des Daniell-Integrals: Denken wir nochmals an das klassische Cauchy-Integral über einem kompakten Intervall. Riemann hat wie Cauchy die Fläche unter dem Graphen einer nichtnegativen Funktion $\{(y,x) : 0 \leq y \leq f(x)\}$ so approximiert, dass er die x-Achse fein unterteilt hat. Lebesgue hat betont, dass es genauso natürlich ist, die y-Achse zu unterteilen und das Maß der Mengen $\{x : a < f(x) \leq b\}$ für kleine Intervalle $(a, b]$ zur Flächenmessung heranzuziehen. Für großes n sollte

$$\sum \tfrac{k}{2^n} \cdot \lambda\{x : \tfrac{k}{2^n} < f(x) \leq \tfrac{k+1}{2^n}\} \;=\; \tfrac{1}{2^n} \cdot \sum \lambda\{x : \tfrac{k}{2^n} < f(x)\}.$$

eine genaue Approximation der 'Fläche unter der Kurve' ergeben. Diese Idee führt vom Lebegue-Maß $\lambda(\cdot)$ zum Lebesgue-Integral.

2.2 Die Konstruktion von Maßen nach Caratheodory

Nehmen wir einmal an, es wäre möglich, 'jeder' Teilmenge A des d-dimensionalen Anschauungsraums eine „Länge“ $\lambda^*(A)$ (die auch $+\infty$ sein kann) zuzuordnen. Welche Eigenschaften könnten oder sollten wir von der Mengenfunktion $\lambda^*(\cdot)$ erwarten?

Wenn wir einen *axiomatischen Zugang* zum Begriff der Länge einer Punktmenge versuchen möchten, dann würden wir jedenfalls postulieren:

1. Für eine Strecke $\overline{PQ}$ sollte der euklidische Abstand der Endpunkte herauskommen, d.h.
$$\lambda^*\left(\overline{PQ}\right) = \quad \mathrm{dist}(P, Q) \quad .$$
Für einen Polygonzug ist ebenfalls klar, was die Länge ist.

2. Auch für eine doppelpunktfreie stetig differenzierbare Kurven haben wir eine Vorstellung, was die Länge sein sollte, nämlich
$$\lambda^*(\mathfrak{C}) = \int_0^1 \|\dot{\gamma}(t)\| dt = \int_0^1 \sqrt{\sum (\dot{x}^j(t))^2} dt$$

3. Die Längenmessung sollte verschiebungsinvariant und drehungsinvariant sein

4. Für disjunkte Mengen, die einen positiven Abstand voneinander haben, sollte die „Länge“ der Vereinigung gleich der Summe der „Längen“ sein. (Eingeschränkte Additivität). Wenn man allerdings an zwei Mengen denkt, die zwar disjunkt, aber doch recht 'verzahnt' nahe beieinander liegen, dann scheint fraglich, ob eine Längenmessung $\lambda^*(\cdot)$ auch dafür additiv sein kann.

5. In jedem Fall würden wir aber erwarten

   i) $A \subseteq B \Rightarrow \lambda^*(A) \leq \lambda^* B)$ (Monotonie)

   ii) $\lambda^*(A \cup B) \leq \lambda^*(A) + \lambda^*(B)$ (Subadditivität)

   iii) $\lambda^*\left(\bigcup^{\infty} A_i\right) \leq \sum^{\infty} \lambda^*(A_i)$ (Vereinigungsbeschränktheit)

Auf der Grundlage der Axiome gewinnen wir bereits für einige Mengen konkrete Zahlenwerte. Denken wir z.B. an die Menge Q der rationalen Punkte im Einheitsintervall und die Menge $[0, 1]\backslash Q$ der irrationalen Punkte. Wenn man Q eine echt positive Länge zuerkennen möchte, dann käme man in Schwierigkeiten mi der Additivität. Die Additivität zusammen mit der Verschiebungsinvarianz impliziert, daß die Menge Q die Länge 0 und die Menge $[0, 1]\backslash Q$ die Länge 1 bekommen sollte.
Die Forderung der Subadditivität erscheint ziemlich schwach. Die Frage ist, für wie allgemeine Paare von disjunkten Mengen die Additivität der Längenmessung garantiert werden

kann. Eine höchst befriedigende Antwort gibt ein Satz von Caratheodory (1873-1950) aus dem Jahre 1914: Solange man bei Borel'schen Mengen bleibt, gilt nicht nur Additivität, sondern sogar σ-Additivität.

Der Beweis ist von einer Art, die vor 100 Jahren neuartig erschien, aber doch wohlwollend aufgenommen wurde. Nach dem Vorschlag von Caratheodory definiert man zunächst (für ein vorgegebenes $\rho > 0$) eine Mengenfunktion $L_\rho(\cdot)$, die zwar so noch nicht als Längenmessung in Frage kommt, aber doch im Limes $\rho \longrightarrow 0$ etwas sehr Brauchbares liefert, ein 'äusseres' Maß mit bemerkenswerten Eigenschaften.

Caratheodorys Konstruktion: Gegeben $A \subseteq \mathbb{R}^d$ und $\rho > 0$.
Wir überdecken A mit ρ-feinen offenen Mengen U_i so, daß die Summe der Durchmesser $\sum d(U_i)$ möglichst klein ist. Das Infimum nennen wir $L_\rho(A)$. $L_\rho(A)$ wächst, wenn $\rho \searrow 0$. Der Limes sei mit $L^*(A)$ bezeichnet.

Beispiele für Mengen $A \subseteq [0,1]$:

1) $A = Q \cap [0,1]$. Wir legen ein Intervall der Länge $\frac{\varepsilon}{2^n}$ um den rationalen Punkt r_n, wobei $r_1, r_2 \ldots$ eine Abzählung der Menge A ist. Die Summe der Durchmesser ist $= 2\varepsilon$. $L_\rho(A) = 0$ für alle ρ.

2) $A =$ Cantor'sches Diskontinuum.

Wenn wir k-mal den Prozeß des Herausschneidens vollzogen haben, dann bleiben 2^k Teilintervalle der Länge $\left(\frac{1}{3}\right)^k$ übrig. Die Gesamtlänge ist $\left(\frac{2}{3}\right)^k$. Die Überdeckung ist zulässig für $\rho > \frac{1}{3^k}$. Also

$$L_\rho(A) \le \left(\frac{2}{3}\right)^k \quad \text{für} \quad \rho > \left(\frac{1}{3}\right)^k, \; L_\rho(A) = 0 \quad \text{für} \quad \rho > \left(\frac{1}{3}\right)^k \quad .$$

Also gilt $L^x(A) = 0$ für die überabzählbare Menge A.

3) $A = [0,1]$. Wir versuchen möglichst sparsam mit ρ-feinen offenen Mengen zu überdecken. Es gilt $\sum_i d(U_i) \ge 1$ für alle ρ. Der Beweis baut auf die Kompaktheit von $\mathbb{R}$, oder elementar gesprochen, auf den Satz von Heine-Borel. Das Argument von Borel: Die Summe der Durchmesser der überdeckenden offenen Mengen kann nicht klein sein. Es reichen schon endlich viele; und wenn man mit endlich vielen Mengen überdeckt, ist die Summe der Durchmesser mindestens 1.Mit dieser Entdeckung von E. Borel beginnt (nach Meinung vieler Mathematikhistoriker) die moderne Maß- und Integrationstheorie.

4) Für die Teilmengen $A \subseteq \mathbb{R}^1$ wird der Limes $\rho \to 0$ nicht benötigt. $L^*(A)$ kann direkt als ein Infimum gewonnen werden. Für Menge $A \subseteq \mathbb{R}^d$ mit $d \ge 2$ braucht man das ρ sehr wohl. A sei z.B. der Graph der Kurve $t \longmapsto t \cdot \sin \frac{1}{t}\big(t \in [0,1]\big)$. Für $\rho > 0$ haben wir $L_\rho(A)$ endlich. Aber $\lim L_\rho(A) = +\infty$. Die Kurve hat unendliche Länge.

5) Sei $[t_0, t_1] \ni t \longmapsto \gamma(t)$ eine stetig differenzierbare doppelpunktfreie Kurve. Man kann leicht beweisen: $L^*\big(\gamma(\cdot)\big)$ ist das Supremum der Längen von einbeschriebenen Polygonzügen. Und das ist in der Tat die Länge im Sinne der klassischen Differential- und Integralrechnung.

Definition 2.8 (Äusseres Maß).
Eine nichtnegative Funktion $\mu^*(\cdot)$ (Wert $+\infty$ erlaubt) auf einer σ-Algebra $\mathfrak{A}^*$ heißt ein **äußeres Maß**, wenn sie monoton und vereinigungsbeschränkt ist, d.h. wenn gilt

i) $0 \leq \mu^*(A) \leq +\infty \quad$ für alle $\quad A,$ (‘Erweiterte Positivität’)

ii) $B \subseteq A \quad \Longrightarrow \quad \mu^*(B) \leq \mu^*(A),$ (’Monotonie’)

iii) $\mu^*\left(\bigcup^{\infty} A_k\right) \leq \sum^{\infty} \mu^*(A_k)\,.$ (‘Vereinigungsbeschränktheit’)

Unsere Mengenfunktionen $L_\rho(\cdot)$ und $L^*(\cdot)$ sind offenbar äußere Maße. Sie sind auch translationsinvariant. Für $L^* = \lim L_\rho(\cdot)$ haben wir darüber hinaus:
Wenn A_1 und A_2 positiven Abstand haben, dann gilt $L^*(A_1 + A_2) = L^*(A_1) + L^*(A_2)$. Auf diese Eigenschaften von $L^*(\cdot)$ kommen wir später zurück. Hier beschäftigen wir uns zunächst mit einem ganz allgemeinen äußeren Maß. Eine geniale Erfindung von Caratheodory ist der Begriff des Zerlegers für $\mu^*(\cdot)$. (Caratheodory gebrauchte die Bezeichnung „meßbare Menge“; diese Bezeichnung ist aber heutzutage anderweitig vergeben).

Definition 2.9 (Zerleger für μ^*).
A heißt Zerleger für das äußere Maß $\mu^*(\cdot)$, wenn

$$\mu^*(W) = \mu^*(W\backslash A) + \mu^*(W \cap A) \quad \text{für alle } W \quad .$$

Satz 2.2.1.
Die Gesamtheit $\mathfrak{A}$ aller Zerleger für $\mu^*(\cdot)$ ist eine σ-Algebra und die Einschränkung $\mu(\cdot)$ von $\mu^*(\cdot)$ auf $\mathfrak{A}$ ist ein Maß.

Beweis.
1) Eine Menge A ist schon dann ein Zerleger für das äußere Maß $\mu^*(\cdot)$, wenn für alle W mit $\mu^*(W) < \infty$ gilt

$$\mu^*(W) \geq \mu^*(W \cap A) + \mu^*(W\backslash A) \quad .$$

2) Wir zeigen, daß die Gesamtheit $\mathfrak{A}$ der Zerleger eine Mengenalgebra ist. Offensichtlich gilt

$$\emptyset \in \mathfrak{A}, \Omega \in \mathfrak{A} \quad \textit{und} \quad A \in \mathfrak{A} \Longrightarrow \Omega\backslash A \in \mathfrak{A} \quad .$$

Der Nachweis, daß mit $A, B \in \mathfrak{A}$ der Durchschnitt $D = A \cap B$ ein Zerleger ist, erfordert etwas Mühe. Wir zeigen, daß für alle W mit $\mu^*(W) < \infty$ gilt

$$\mu^*(W\backslash D) = \mu^*(W\backslash A) + \mu^*(W \cap A) - \mu^*(W \cap D) \quad .$$

Zu diesem Zweck zeigen wir

$$\begin{array}{rrcl} \textit{(i)} & \mu^*(W\backslash D) & = & \mu^*(W\backslash A) + \mu^*(W \cap A)\backslash D) \\ \textit{(ii)} & \mu^*(W \cap A) & = & \mu^*\big((W \cap A)\backslash D\big) + \mu^*(W \cap D) \quad . \end{array}$$

Zum Beweis von (i) wenden wir die Zerlegereigenschaft von A *auf* $W_1 := W\backslash D$ *an, wobei wir bemerken*

$$W_1\backslash A = (W\backslash D)\backslash A = W\backslash A\,; \quad W_1 \cap A = (W\backslash D) \cap A = (W \cap A)\backslash D \quad .$$

Zum Beweis von (ii) wenden wir die Zerlegereigenschaft von B *auf* $W_2 := W \cap A$ *an, wobei wir bemerken*

$$W_2\backslash B = (W \cap A)\backslash(A \cap B) = (W \cap A)\backslash D\,; \quad W_2 \cap B = W \cap D \quad .$$

3) Seien $A_1, A_2 \ldots$ Zerleger mit $A_1 \subseteq A_2 \subseteq \ldots$ und $A = \bigcup^{\infty} A_n$.

Wir zeigen, daß A ein Zerleger ist, indem wir für alle W mit $\mu^*(W) < \infty$ zeigen

$$\begin{array}{rrcll} \text{(i)} & \mu^*(W) & \geq & \mu^*(W\backslash A) + \mu^*(W \cap A_n) & \text{für alle } n \\ \text{(ii)} & \mu^*(W \cap A) & = & \lim \nearrow \mu^*(W \cap A_n) & . \end{array}$$

Die Behauptung (i) ist evident wegen $W\backslash A \subseteq W \cap A_n$. Zum Nachweis von (ii) zeigen wir zunächst

$$\mu^*(W \cap A_{n+1}) = \mu^*(W \cap A_n) + \mu^*\big(W \cap (A_{n+1} - A_n)\big)$$

indem wir bemerken

$$\big(W \cap A_{n+1}\big) \cap A_n = W \cap A_n; \qquad (W \cap A_{n+1})\backslash A_n = W \cap (A_{n+1} - A_n)$$

und A_n ist ein Zerleger. Es folgt

$$\mu^*(W \cap A_{n+1}) = \mu^*(W \cap A_1) + \mu^*\big(W \cap (A_2 - A_1)\big) + \ldots + \mu^*\big(W \cap (A_{n+1} - A_n)\big)$$

Nun gilt $W \cap A \subseteq \bigcup_{n=1}^{\infty} \big(W \cap (A_n - A_{n-1})\big)$ und wegen der Vereinigungsbeschränktheit

$$\mu^*(W \cap A) \leq \sum^{\infty} \mu^*\big(W \cap (A_n - A_{n-1})\big) = \lim_n \nearrow \mu^*(W \cap A_n) \quad \text{q.e.d.}$$

4) Sei $\tilde{W}$ eine beliebige Menge und

$$\tilde{\mu}(A) = \mu^*\left(\tilde{W} \cap A\right) \quad \text{für alle} \quad A \in \mathfrak{A} \quad \text{(System der Zerleger)}\,.$$

Wir zeigen, daß $\tilde{\mu}(\cdot)$ ein Maß ist: Seien A, B disjunkte Zerleger. Es gilt dann

$$\begin{array}{rcl} \Big(\tilde{W} \cap (A + B)\Big)\backslash A & = & \tilde{W} \cap B, \ \Big(\tilde{W} \cap (A + B)\Big) \cap A = \tilde{W} \cap A \\ \tilde{\mu}(A + B) & = & \tilde{\mu}(B) + \tilde{\mu}(A) \quad . \end{array}$$

Seien $A_1 \subseteq A_2 \subseteq \ldots$ Zerleger und $A = \bigcup\limits^{\infty} A_n$. Es gilt

$$\tilde{\mu}(A) = \mu^*\left(\tilde{W} \cap A\right) = \lim \nearrow \mu^*\left(W \cap A_n\right) = \lim \nearrow \tilde{\mu}\left(A_n\right) \quad .$$

Also ist $\tilde{\mu}(\cdot)$ aufsteigend stetig.
Damit ist der Satz vollständig bewiesen.

Äussere Maße und die daraus abgeleiteten Maße sind nur dann interessant, wenn die σ-Algebra der Zerleger hinreichend reichhaltig ist. Wir werden sehen, dass in unserem Fall die offenen Mengen Zerleger sind. Die Herleitung stützt sich auf die 'eingeschränkte Additivität' des äusseren Maßes: Für Mengen mit positivem Abstand voneinander gilt

$$\mu^*(A_1 + A_2) = \mu^* A_1 + \mu^* A_2.$$

Satz 2.2.2. *Ist $\mu^*(\cdot)$ ein äußeres Maß auf der Gesamtheit aller Teilmengen eines metrischen Raums $\big(E, d(\cdot,\cdot)\big)$ mit*

$$d(A_1, A_2) > 0 \qquad \Longrightarrow \qquad \mu^*(A_1 + A_2) = \mu^* A_1 + \mu^* A_2.$$

so ist jede offene Menge G ein Zerleger für μ^*.
Die Einschränkung von μ^* auf die Borelalgebra ist also ein Maß.

Beweis.
Sei F das Komplement F von G. Wir müssen für alle W mit $\mu^*(W) < \infty$ zeigen

$$\mu^*(W) \geq \mu^*(W \cap G) + \mu^*(W \cap F) \quad .$$

Für $n = 1, 2, \ldots$ sei $\qquad G_n = \left\{x \,:\, dist(x, F) > \frac{1}{n}\right\} \quad .$
Da jedes G_n positiven Abstand von F hat, haben wir

$$\mu^*(W) \geq \mu^*\left(W \cap G_n\right) + \mu^*(W \cap F) \quad ,$$

Der Satz ist bewiesen, wenn wir zeigen.

$$\mu^*(W \cap G) = \lim \nearrow \mu^*\left(W \cap G_n\right) \quad .$$

Die Subadditivität sagt $\qquad \mu^*\big(W \cap G\big) \leq \mu^*\big(W \cap G_n\big) + \mu^*\big(W \cap (G - G_n)\big).$
Wir zeigen, dass zu jedem $\varepsilon > 0$ ein n existiert, sodass

$$\mu^*\big(W \cap (G - G_n)\big) < \varepsilon.$$

Das ergibt sich folgendermaßen: Für jedes n haben G_n und $E - G_{n+1}$ positiven Abstand. Wenn nämlich x und y Punkte sind mit $dist(x, F) > \frac{1}{n}$, $\mathrm{dist}(y, F) \leq \frac{1}{n+1}$,
dann gilt $dist(x, y) \geq \frac{1}{n} - \frac{1}{n+1} = \frac{1}{n(n+1)}$.

Die Mengen $G_2 - G_1$, $G_4 - G_3$ ... *haben paarweise positiven Abstand. Für ihre Vereinigung* R_1 *gilt*

$$\mu^*(W \cap R_1) = \mu^*\big(W \cap (G_2 - G_1)\big) + \mu^*\big(W \cap (G_4 - G_3)\big) + \ldots$$

Dasselbe gilt für die Mengen $G_3 - G_2$, $G_5 - G_4$, ... *und ihre Vereinigung* R_2

$$\mu^*(W \cap R_2) = \mu^*\big(W \cap (G_3 - G_2)\big) + \mu^*\big(W \cap (G_5 - G_4)\big) \ldots .$$

Die Summen konvergieren wegen $\mu^*(W) < \infty$. *Wenn* n *genügend groß ist, dann gilt*

$$\mu^*\big(W \cap (G - G_n)\big) \le \sum_{k=n}^{\infty} \mu^*\big(W \cap (G_{k+1} - G_k)\big) < \varepsilon.$$

Caratheodory hat nicht nur das lineare Maß eingeführt. Auch sein Vorschlag für die Definition des „Oberflächenmaßes“ im $\mathbb{R}^d$ hat allgemeine Anerkennung gefunden. Caratheodory konstruiert folgendermaßen.

Weitere Konstruktionen:

Für eine offene konvexe Teilmenge K des $\mathbb{R}^d$ sei $d^{(2)}(K)$ der maximale zweidimensionale Schatten, d.h. das Maximum der Flächen von orthogonalen Projektionen von K auf einen zweidimensionalen Teilraum des $\mathbb{R}^d$.

Wir arbeiten nun mit konvexen K mit Durchmesser $< \rho$.

Eine beliebige Menge $A \subseteq \mathbb{R}^d$ soll nun möglichst sparsam mit ρ-feinen konvexen K_i überdeckt werden; man konstruiert

$$L_\rho^{(2)}(A) = \inf\Big\{ \sum d^{(2)}(K_i) \;:\; \bigcup^{\infty} K_i \supseteq A \,,\; d(K_i) < \rho \quad \text{für alle } i \Big\}$$

Der aufsteigende Limes $L^{*(2)}(A) = \lim_{\rho \searrow 0} L_\rho^{(2)}(A)$ erfüllt die Voraussetzungen der beiden Sätze. Die Einschränkung auf die Borelalgebra liefert eine Flächenmessung für Mengen im $\mathbb{R}^d$. Man sieht leicht, dass für sich für ein Rechteck im $\mathbb{R}^d$ die klassische Fläche ergibt: Fläche = Länge × Breite. Auch für allgemeinere ‘Flächenstücke’ A, denen man in der klassischen Infinitesimalrechnung einen Flächeninhalt zumisst, liefert Caratheodorys Konstruktion das gewünschte Ergebnis.

F. Hausdorff hat 1919 weitere metrische äußere Maße im $\mathbb{R}^n$ konstruiert, die in den letzten Jahren im Anschluß an einige Aufsätze von B. Mandelbrot sehr populär geworden sind. Eine sehr schöne Darstellung gibt

Falconer, K.: *Fractal Geometry*, Wiley, 1990.

Hausdorff konstruierte insbesondere für $s > 0$

$$\mu_\rho^{(s)}(A) = \inf\Big\{ \sum \big(d(U_i)\big)^s \;:\; \bigcup^{\infty} U_i \supseteq A \,,\; d(U_i) < \rho \Big\} \quad .$$

Der Limes $\mu^{(s)}(\cdot) = \lim \nearrow \mu_\rho^{(s)}(\cdot)$ ist offenbar ein Maß auf der Borelalgebra.
Für die meisten B hat man $\mu^{(s)}(B) = \infty$ oder $\mu^{(s)}(B) = 0$.
Wenn $\mu^{(s)}(B) < \infty$, dann gilt offenbar $\mu^{(t)}(B) = 0$ für alle $t > s$.
Das Infimum der s, mit $\mu^{(s)}(B) < \infty$ heißt die **Hausdorff-Dimension** von B.
Man kann in der Tat zeigen, dass die in den $\mathbb{R}^d$ eingebeteten k-dimensionalen glatten Mannigfaltigkeiten die Hausdorff-Dimension k haben. Für nichtglatte Mengen B findet man auch nichtganzzahlige Werte der Hausdorff-Dimension.

Hinweis für Stochastiker:
Der zweidimensionale Graph der eindimensionalen Brown'schen Bewegung hat (fast sicher) die Dimension $\frac{3}{2}$. Diese Brown'schen Pfade im $\mathbb{R}^3$ haben (fast sicher) die Hausdorff-Dimension 2. Die Brown'schen Graphen und die Brown'schen Pfade sind mit Erfolg sehr genau untersucht worden, nicht nur auf ihre Dimension hin.
Zur genaueren Untersuchung dient die folgende Verallgemeinerung des Ansatzes von Hausdorff.

Konstruktion
Sei $h(\cdot)\ :\ \mathbb{R}^+ \to \mathbb{R}^+$ eine isotone Funktion mit $h(0) = 0$. Für Teilmengen B des $\mathbb{R}^d$ definiert man das Hausdorff-Maß von B bzgl. $h(\cdot)$ als Limes

$$\mu^{(h)}(B) = \lim_{\rho\to 0} \uparrow \mu_\rho^{(h)}(B) = \lim_{\rho\to 0} \uparrow \inf\Big\{ \sum^{\infty} h\big(d\,(U_i)\,\big)\ :\ \bigcup^{\infty} U_i \supseteq B\ ,\ d\,(U_i) < \rho \Big\} \quad .$$

Bemerke: Das Hausdorff-Maß zur Dimensionsfunktion $h(\cdot)$ hängt monoton von $h(\cdot)$ ab. Je steiler $h(\cdot)$ im Nullpunkt ansteigt, desto größer ist das entsprechende Hausdorff-Maß (für alle B). Genauer gesagt: Sind $h(\cdot)$ und $k(\cdot)$ Funktionen, so daß für ein δ gilt $h(d) \leq k(d)$ für alle $d \in (0, \delta)$, dann gilt $\mu^{(h)}(B) \leq \mu^{(k)}(B)$ für alle Borel'schen B.

Beispiele
1) Es stellt sich heraus, daß die Brown'schen Pfade im $\mathbb{R}^n$ nicht nur für die Funktion d^2 endliches Hausdorff-Maß haben, sondern sogar für die steiler ansteigende Funktion

$$h(d) = d^2 \cdot \log\log\left(\frac{1}{d}\right) \quad .$$

Dieses $h(\cdot)$ liefert eine wirklich genaue Ausmessung der Brown'schen Pfade. Man kann beweisen : Für jede Dimension $n \geq 3$ gibt es eine Konstante c_n, so daß für fast alle Brown'schen Pfade im Zeitabschnitt $[0, T]$ gilt

$$L^{(h)}\Big(B_\omega[0, T]\Big) = T \cdot c_n \quad .$$

2) In den Dimensionen $n = 1$ und $n = 2$ sind die Verhältnisse etwas anders. Für die Graphen der eindimensionalen Brown'schen Bewegung beschreibt

$$h(d) = d^{3/2} \cdot \sqrt{\log\log \frac{1}{d}}$$

die genaue Hausdorff-Dimension. Es gilt

$$L^h\Big(\Gamma_\omega([0,T])\Big) = c \cdot T \quad \text{fast sicher} \quad .$$

Diese feinen Resultate sind noch nicht sehr alt. Siehe

S.J. Taylor, *The measure theory of random fractals*, Math. Proc. Cambridge Phil. Soc. 100 (1986), S. 383-406.

2.3 Inhalte, Prämaße, Maße und Treppenfunktionen

Definition 2.10 (Mengenring und Mengenalgebra).
Ein nichtleeres System $\mathfrak{R}$ von Teilmengen einer Grundmenge Ω heisst ein Mengenring über Ω, wenn gilt

$$A,\, B \in \mathfrak{R} \implies A \cup B,\; A \backslash B \in \mathfrak{R}.$$

Wenn ausserdem $\Omega \in \mathfrak{R}$, dann spricht man von einer Mengenalgebra über Ω.
Eine positive Linearkombination von Indikatorfunktionen

$$h = \sum \alpha_j \cdot 1_{R_j} \quad \text{mit} \quad \alpha_j \geq 0,\, R_j \in \mathfrak{R}$$

heisst eine nichtnegative $\mathfrak{R}$-Treppenfunktion.

Bemerke: Man hat in $\mathfrak{R}$ auch die Durchschnittsbildung $A \cap B = A \backslash (A \backslash B)$ und die 'symmetrische Differenz' $A \triangle B = A \backslash B + B \backslash A$. Und das System $\big(\mathfrak{R}, \emptyset, \triangle, \cap\big)$ ist in der Tat ein kommutativer Ring im Sinne der Algebra mit $\triangle$ als Addition und $\cap$ als Multiplikation. Man sieht das sofort, wenn man $R \in \mathfrak{R}$ mit der Funktion identifiziert, die Werte aus dem Körper $\mathbb{Z}/2\mathbb{Z}$ annimmt, den Wert 1 auf R und den Wert 0 auf dem Komplement.

Definition 2.11 (Inhalte und Prämaße).
Eine nichtnegative Funktion $\rho(\cdot)$ auf einem Mengenring heisst ein Inhalt, wenn gilt

$$\rho(\emptyset) = 0, \qquad \rho(A \cap B) + \rho(A \cap B) = \rho(A) + \rho(B).$$

Ein Inhalt wird ein Prämaß genannt, wenn er absteigend stetig ist in dem Sinn

$$A_1 \supseteq A_2 \supseteq \cdots \qquad \bigcap A_n = \emptyset \quad \implies \quad \lim \downarrow \rho(A_n) = 0.$$

Satz 2.3.1. *Es bezeichne $\mathfrak{R}^{\mathrm{Tr}+}$ die Menge der nichtnegativen $\mathfrak{R}$-Treppenfunktion und $\mathfrak{R}^{\mathrm{In}+}$ die Menge der Inhalte auf $\mathfrak{R}$. Es gilt*

1. *$\mathfrak{R}^{\mathrm{Tr}+}$ ist ein Verbandskegel mit punktweiser Addidition und punktweiser Maximums- und Minimumbildung*

2. *$\mathfrak{R}^{\mathrm{In}+}$ ist ein Verbandskegel mit punktweiser Addidition und punktweiser Ordnung (aber nicht mit 'punktweiser' Maximums- und Minimumbildung!). Es gilt*

$$\rho_1 \vee \rho_2 + \rho_1 \wedge \rho_2 = \rho_1 + \rho_2.$$

3. *Für alle $A \in \mathfrak{R}$ gilt*

$$\big(\rho_1 \vee \rho_2\big)(A) = \sup\Big\{\sum \max\{\rho_1(A_j),\, \rho_2(A_j)\} : \quad \sum A_j = A\Big\},$$
$$\big(\rho_1 \backslash \rho_2\big)(A) = \sup\Big\{\sum \big(\rho_1(A_j) - \rho_2(A_j)\big)^+ : \quad \sum A_j = A\Big\},$$

wo das Supremum über alle endlichen disjunkten Partitionen zu erstrecken ist.

4. Für alle Tripel $\rho_1, \rho_2, \rho_3 \in \mathfrak{R}^{\mathrm{In}+}$ *gilt*

$$\begin{aligned} \rho_1 = \rho_1 \wedge \rho_2 + \rho_1\backslash\rho_2, &\qquad (\rho_1\backslash\rho_2)\backslash\rho_3 = \rho_1\backslash(\rho_2+\rho_3), \\ (\rho_1+\rho_2)\backslash\rho_3 = \rho_1\backslash\rho_3 + \rho_2\backslash(\rho_3\backslash\rho_1), &\qquad (\rho_1+\rho_2)\wedge\rho_3 = \rho_1\wedge\rho_3 + \rho_2\wedge(\rho_3\backslash\rho_1), \\ (\rho_1\vee\rho_2)\backslash\rho_3 \quad = \big(\rho_1+\rho_2\backslash\rho_1\big)\backslash\rho_3 = \rho_1\backslash\rho_3 + \rho_2\backslash\big(\rho_1\vee\rho_3\big) &\quad \le \rho_1\backslash\rho_3+\rho_2\backslash\rho_3. \end{aligned}$$

Für den Beweis von 3) bemerken wir, dass die Verfeinerung einer Partition eine Vergrößerung der betreffenden Summe bewirkt.

Beispiel 2.3.1.
Es sei $\Omega = \sum A_j$ eine abzählbare Partition der Grundmenge. $\mathfrak{R}$ sei der Mengenring aller endlichen Vereinigungen von Atomen der Partitionen. Jedem Atom A_j sei eine nichtnegative Zahl $p_j = \rho(A_j)$ zugeordnet, das 'Gewicht des Atoms'. Es gibt genau eine Fortsetzung zu einem Inhalt, nämlich $\rho(A) = \sum_{\{j:A_j\subseteq A\}} p_j$.

Das System aller Mengen A, für welche entweder A selbst oder das Komplement $\Omega\backslash A$ zu $\mathfrak{R}$ gehört, ist die erzeugte Mengenalgebra.

Warnung: Wir bemerken möglicherweise eine gewisse strukturelle Ähnlichkeit (mit einer Art Dualität) zwischen den Verbandskegeln $\mathfrak{R}^{\mathrm{Tr}+}$ und $\mathfrak{R}^{\mathrm{In}+}$. Man darf sich aber nicht irreleiten lassen. Der große Unterschied besteht darin, dass man in $\mathfrak{R}^{\mathrm{Tr}+}$ den Begriff einer punktweise monoton fallenden Folge hat, zu welcher es keine interessante Entsprechung im Kegel $\mathfrak{R}^{\mathrm{In}+}$ gibt.

Die interessanten Inhalte sind die Prämaße. Die sollen hier genauer studiert werden. Wir bemerken: Wenn ρ ein Prämaß ist, dann ist auch jeder kleinere Inhalt ein Prämaß.

Satz 2.3.2 (Der Prämaß-Anteil eines Inhalts)**.**
Zu jedem Inhalt ρ *auf dem Mengenring* $\mathfrak{R}$ *gibt es eine größtes Prämaß unter* ρ*. Für dieses Prämaß* ρ_σ *gilt*

$$\rho_\sigma(A) = \inf\Big\{\sum \rho(R_j) : \quad \sum R_j = A\Big\} \quad \textit{für } A \in \mathfrak{R},$$

wo das Infimum über alle abzählbaren disjunkten Partitionen zu erstrecken ist.

Beweis. *Es sei* $A = \sum_{n=1}^{\infty} A_n$ *eine abzählbare Partition von* $A \in \mathfrak{R}$*. Es gilt*

$$\sum \rho_\sigma(A_n) = \inf\big\{\sum_{n,j} \rho(R_{n,j}) : \quad \sum_j R_{n,j} = A_n \quad \textit{für alle } n\big\}.$$

Das Infimum geht hier über alle Partitionen von A*, die feiner sind als* $A = \sum_{n=1}^{\infty} A_n$*. Daher gilt* $\rho_\sigma(A) \le \sum_n \rho_\sigma(A_n)$*. Die Mengenfunktion* ρ_σ *ist also monoton und vereinigungsbeschränkt. Es seien* A, B *disjunkt und* $A+B = \sum R_j$ *so fein, dass*

$$\rho_\sigma(A+B) \ge \varepsilon + \sum \rho(R_j) = \sum\big(\rho(A\cap R_j) + \rho(B\cap R_j)\big) \ \ge \varepsilon + \rho_\sigma(A) + \rho_\sigma(B).$$

Also ist ρ_σ ein Prämaß unter $\leq \rho$. Ist nun τ ein weiteres Prämaß $\leq \rho$ auf $\mathfrak{R}$, so gilt

$$\tau(A) = \inf\left\{\sum \tau(R_j): \quad \sum R_j = A\right\} \quad \leq \quad \rho_\sigma(A) \quad \textit{für alle } A \in \mathfrak{R}.$$

Somit erweist sich ρ_σ als das größte Prämaß $\leq \rho$ auf $\mathfrak{R}$. In der Tat leistet die Konstruktion $\rho^*(B) = \inf\left\{\sum \rho(R_j): \quad \sum R_j \supseteq B\right\}$ auf der von $\mathfrak{R}$ erzeugten σ-Algebra die Fortsetzung des Prämaßes ρ_σ zu einem Maß.

Satz 2.3.3 (Das σ-Ideal der Nullmengen zu einem Prämaß). *Es sei ρ ein Prämaß auf dem Mengenring $\mathfrak{R}$ über Ω, und $\mathfrak{A}$ die von $\mathfrak{R}$ erzeugte σ-Algebra. Eine Menge $N \in \mathfrak{A}$ heisst eine ρ-Nullmenge, wenn zu jedem $\varepsilon > 0$ eine abzählbare Überdeckung existiert*

$$N \subseteq \sum A_j, \quad \textit{sodass gilt} \quad \sum \rho(A_j) < \varepsilon.$$

Die Menge $\mathfrak{N}_\rho$ der ρ-Nullmengen ist ein σ-Ideal in $\mathfrak{A}$ in dem Sinn

i) $N \in \mathfrak{N}_\rho, \quad A \in \mathfrak{A} \quad \Longrightarrow \quad A \cap N \in \mathfrak{N}_\rho,$

ii) $N_1, N_2, \ldots \quad \in \mathfrak{N}_\rho \quad \Longrightarrow \quad \bigcup N_n \in \mathfrak{N}_\rho$

Beweis. *Die erste Behauptung ist trivial. Für den Beweis der zweiten wählen wir zur Nullmenge N_n eine Überdeckung mit $\sum \rho(A_j^{(n)}) < \varepsilon/2^n$. Alle überdeckenden Mengen zusammen liefern eine ε-knappe Überdeckung von $\bigcup N_n$.*

Beschränkte und σ-endliche Inhalte

Wir nehmen hier immer an, dass Inhalte endlichwertig sind. Eine wichtige Klasse sind die beschränkten Inhalte. Ein beschränkter Inhalt ρ mit $\sup\{\rho(R): R \in \mathfrak{R}\} = \|\rho\| < \infty$ kann dadurch zu einem Inhalt auf der erzeugten Mengenalgebra $\mathfrak{A}$ fortgesetzt werden, dass der Menge $\Omega - R$ der Wert $\|\rho\| - \rho(R)$ zugeordnet wird. Die Zahl $\rho(R)$ heisst die Totalvariation des beschränkten Inhalts ρ. Es gibt auch noch andere Fortsetzungen zu einem Inhalt auf $\mathfrak{A}$.

Es sei $\mathfrak{R}$ ein Mengenring über Ω mit der Eigenschaft, dass es eine abzählbare Familie gibt mit $\bigcup R_n = \Omega$. Wir werden sehen, dass es in diesem Fall zu jedem Prämaß genau eine Fortsetzung zu einem Maß auf der erzeugten σ-Algebra gibt. Maße, die so entstehen (und auch den Wert $+\infty$ annehmen können), heissen σ-endliche Maße. Die Hausdorff-Maße im vorigen Unterabschnitt sind typischerweise nicht σ-endlich. Im Folgenden denken wir prinzipiell an σ-endliche Maße, auch wenn manche der Sätze die σ-Endlichkeit nicht fordern müssten.

Didaktische Anmerkung: Die Beschränkung auf σ-endliche Maße hat einen didaktischen Grund. Die Fragen, was anders ist bei den Maßen, die nicht σ-endlich sind, überlassen wir gerne einer Spezialvorlesung über Maßtheorie. Die didaktische Strategie ist ähnlich wie früher schon in der Topologie: Als wir uns mit Hausdorff-Räumen beschäftigten, brauchten wir nicht immer die Existenz einer abzählbaren Basis. Wir wollen aber in dieser Anfängervorlesung diese Front nicht bearbeiten, und denken (wenn nichts Gegenteiliges gesagt wird) prinzipiell nur an die Räume mit abzählbarer Basis.

Die Prämaße ρ mit $\|\rho\| = 1$ führen zu normierten Maßen, die man bekanntlich auch Wahrscheinlichkeitsmaße nennt. Bei den Wahrscheinlichkeitsmaßen, die in der Stochastik besonderes Interesse finden, ist einiges anders (und einfacher zu formulieren), als bei den σ-endlichen Maßen, die auch den Wert $+\infty$ annehmen können.

Beispiel 2.3.2. 1. Sei Ω eine abzählbar unendliche Menge, und $\mathfrak{R}$ das Mengenring der endlichen Teilmengen. Den Elementen von Ω seien Gewichte $p(\omega) \geq 0$ zugewiesen, und es sei $\tilde{\rho}(A) = \sum_{\omega \in A} p(\omega)$. Es handelt sich um ein σ-endliches Prämaß.

2. Sei Ω eine abzählbar unendliche Menge, und $\mathfrak{R}$ derjenigen Mengen, die entweder selbst endlich sind oder ein endliches Komplement besitzen. Es seien Gewichte gegeben mit $\bar{p} = 1 - \sum p(\omega) \geq 0$. Für das Komplement einer endlichen Menge A sei $\tilde{\rho}(\complement A) = 1 - \sum_{\omega \in A} p(\omega)$. Für $\bar{p} > 0$ ergibt das einen Inhalt, der kein Prämaß ist. Für $\bar{p} = 0$ gibt es genau eine Fortsetzung zu einem Maß auf der Potenzmenge.

3. Sei $\mathfrak{S}$ das System aller beschränkten halboffenen Intervalle $(a, b]$ auf der reellen Achse (linksseitig offen, rechtsseitig abgeschlossen!). $\mathfrak{S}$ ist durchschnittsabgeschlossen und die Gesamtheit aller (disjunkten) Vereinigungen bilden einen Mengenring über $\mathbb{R}$. $F(\cdot)$ sei wachsend und rechtsseitig stetig auf $\Omega = \mathbb{R}$. Es gibt genau ein Maß ρ auf der Borelalgebra mit $\rho\big((a, b]\big)$. Es handelt sich um die Daniell-Fortsetzung des Stieltjes-integrals $I^*(\cdot)$ zu $F(\cdot)$, eingeschränkt auf die Borelalgebra. Es gilt $I(h) = \int h(\omega)\, dF(\omega)$ für jede positive Borel-messbare Funktion h.

4. Die halboffenen Intervalle im $\mathbb{R}^2$ sind die 'Rechtecke'

$$\big\{\omega : \ \omega = (x_1, x_2) \quad \text{mit} \quad a_1 < x_1 \leq b_1,\ a_2 < x_2 \leq b_2\big\} = (a_1, b_1] \times (a_1, b_1].$$

Das System $\mathfrak{S}$ aller Rechtecke ist durchschnittsstabil; das System aller (disjunkten) Vereinigungen von Rechtecken ist ein Mengenring $\mathfrak{R}$ oder sogar eine Mengenalgebra, wenn man nichtbeschränkte Rechtecke zulässt. $\mathfrak{R}$ erzeugt die Borelalgebra $\mathfrak{B}$ über $\mathbb{R}^2$. Ein σ-endliches Borelmaß ρ ist eindeutig bestimmt durch seine Werte auf den beschänkten Rechtecken. Es existiert eine Funktion $F(\cdot, \cdot)$, sodass gilt

$$\rho\big((a_1, b_1] \times (a_2, b_2]\big) = F(b_1, b_2) - F(a_1, b_2) - F(b_1, a_2) + F(a_1, a_2) \geq 0.$$

Es ist etwas umständlich, die Eigenschaft der 'Verteilungsfunktion' $F(\cdot, \cdot)$ zu formulieren, die der Rechtstetigkeit im eindimensionalen Fall entspricht — und die Fachleute sind heute der Meinung, dass es die Mühe nicht lohnt; es gibt andere Wege, die Borelmaße über dem $\mathbb{R}^2$ zu beschreiben.

5. Manche borel'sche Wahrscheinlichkeitsmaße ρ auf dem Raum $\mathbb{R}^2$ lassen sich durch Dichten bzgl. des Lebesgue'schen Maßes beschreiben.

$$p(x, y)\, dx\, dy \qquad \text{mit} \qquad p(\cdot, \cdot) \geq 0, \quad \int\!\!\int p(x, y)\, dx\, dy = 1.$$

Man notiert in diesem Falle für die borel'schen Menge $B \subseteq \mathbb{R}^2$

$$\rho(B) = \int_B p(x,y)\ dx\ dy = \int 1_B(x,y) \cdot p(x,y)\ dx\ dy \quad \text{oder kurz} \quad = \int 1_B p\ d\lambda.$$

Dabei bezeichnet λ das Lebesgue-Maß, ein Maß, welches aus vielen Gründen eine besondere Rolle spielt. Das Maß λ ist bis auf eine Normierungskonstante eindeutig festgelegt durch die Eigenschaft, dass es invariant ist gegenüber allen Translationen. Das Lebesgue-Maß erscheint aus der Sicht einer allgemeineren Theorie als das Haar'sche Maß auf der lokalkompakten Gruppe $\mathbb{R}^2$. — Wir werden immer wieder darauf zurückkommen.

Satz 2.3.4 (Der große Fortsetzungssatz)**.**
Jedes σ-endliche Prämaß $\tilde{\rho}$ auf einem Mengenring $\mathfrak{R}$ besitzt genau eine Fortsetzung zu einem Maß ρ auf der erzeugten σ-Algebra $\mathfrak{A}$.

Beweis. *Für die Konstruktion gehen wir vor wie Caratheodory: Wir definieren zuerst ein äusseres Maß ρ^* und zeigen dann, dass jedes $A \in \mathfrak{R}$ ein Zerleger sind. Wir setzen*

$$\rho^*(W) = \inf\left\{\sum \rho(A_j): \quad \bigcup A_j \supseteq W\right\} \quad \textit{für } W \textit{ beliebig},$$

wo das Infimum über alle abzählbaren $\mathfrak{R}$-Überdeckungen zu erstrecken ist. Wir zeigen

$$\rho^*(W) + \varepsilon \geq \rho^*(W \cap A) + \rho^*(W \setminus A) \quad \textit{für } W \textit{ mit } \rho^*(W) < \infty \textit{ und } A \in \mathfrak{R}.$$

Für W wählen wir eine ε-knappe Überdeckung $W \subseteq \bigcup A_j$ mit $\sum \tilde{\rho}(A_j) < \rho^*(W) + \varepsilon$; und haben dann

$$W \cap A \subseteq \bigcup A_j \cap A; \qquad W \setminus A \subseteq \bigcup A_j \setminus A,$$
$$\rho^*(W \cap A) + \rho^*(W \setminus A) \leq \sum \left[\tilde{\rho}(A_j \cap A) + \tilde{\rho}(A_j \setminus A)\right] = \sum \tilde{\rho}(A_j) \leq \rho^*(W) + \varepsilon.$$

Die Einschränkung von ρ^* auf die σ-Algebra ist ein Maß; und die Prämaß-Eigenschaft von $\tilde{\rho}$ garantiert, dass es eine Fortsetzung. ist.

Für den Beweis der Eindeutigkeit holen wir weiter aus.

Sprechweise.
Ein Mengensystem $\mathfrak{D}$ über der Grundmenge Ω heisst **Dynkin-System**, wenn gilt

i) $\Omega \in \mathfrak{D}$,

ii) $A \in \mathfrak{D} \implies \Omega - A \in \mathfrak{D}$,

iii) $A_1, A_2, \ldots \in \mathfrak{D}$ und A_i paarweise disjunkt $\implies \sum_1^\infty A_i \in \mathfrak{D}$.

Wir bemerken: In einem Dynkin-System gilt $\big(A, B \in \mathfrak{D}$ und $A \subseteq B\big) \;\Rightarrow\; B-A \in \mathfrak{D}$. Es gilt nämlich $(B-A)^c = B^c + A$. Dynkinsysteme sind somit stabil gegenüber (abzählbarer!) disjunkter Vereinigung und gegenüber 'echter' Differenzbildung.

Ein Dynkinsystem ist genau dann eine σ-Algebra, wenn es durchschnittsstabil ist.

Satz 2.3.5 (Durchschnittsstabile Erzeugendensysteme)**.**
Sei $\mathfrak{S}$ ein durchschnittsstabiles Mengensystem und $\mathfrak{D}$ das davon erzeugte Dynkin-System. Dann ist $\mathfrak{D}$ durchschnittsstabil, d. h. $\mathfrak{D}$ ist die erzeugte σ-Algebra $\mathfrak{A}_\sigma$.

Beweis.
Für jedes feste $S \in \mathfrak{S}$ ist das Mengensystem $\mathfrak{D}_S = \{A : A \cap S \in \mathfrak{D}\}$ ein Dynkin-System, welches $\mathfrak{S}$ umfasst. Es gilt also

$$\forall_{S \in \mathfrak{S}} \quad \mathfrak{D}_S \supseteq \mathfrak{D}; \qquad \text{d. h.} \qquad S \in \mathfrak{S},\ D \in \mathfrak{D} \Rightarrow S \cap D \in \mathfrak{D}.$$

Für ein festes $D \in \mathfrak{D}$ ist $\mathfrak{D}_D = \{A : A \cap D \in \mathfrak{D}\}$ ein Dynkin-System, welches $\mathfrak{D}$ umfasst. Es gilt daher

$$\forall_{D \in \mathfrak{D}} \quad \mathfrak{D}_D \supseteq \mathfrak{D}; \qquad \text{d. h.} \qquad D \in \mathfrak{D},\ A \in \mathfrak{D} \Rightarrow A \cap D \in \mathfrak{D},$$

was zu beweisen war.

Mit Hilfe dieses Satzes komplettieren wir den Beweis des großen Fortsetzungssatzes.

Satz 2.3.6 (Eindeutigkeitssatz).
Seien μ, ν σ-endliche Maße auf einer σ-Algebra $\mathfrak{A}$ über einem Grundraum Ω mit

i) Es existiert ein durchschnittsstabiles Erzeugendensystem $\mathfrak{S}$ mit

$$\mu(S) = \nu(S) \quad \textit{für alle } S \in \mathfrak{S}$$

ii) In $\mathfrak{S}$ existiert eine gegen Ω aufsteigende Folge $(A_n)_n$ mit $\mu(A_n) = \nu(A_n) < \infty$.

Dann gilt $\mu = \nu$.

Beweis. *Es genügt, den Satz für Wahrscheinlichkeitsmaße zu beweisen; denn die Maße $\mu(\cdot \cap A_n)$ und $\nu(\cdot \cap A_n)$ sind bis auf eine Normierung Wahrscheinlichkeitsmaße; und wenn $\mu(A \cap A_n) = \nu(A \cap A_n)$ für alle $A \in \mathfrak{A}$ und alle n, dann sind μ und ν gleich.*

Der Beweis ergibt sich aus der Beobachtung, dass das Mengensystem $\mathfrak{D} = \big\{D : \mu(D) = \nu(D)\big\}$ ein Dynkinsystem ist, welches von einem durchschnittsstabilen Mengensystem erzeugt wird.

Satz 2.3.7 (Das Elementarintegral zu einem Prämaß).
Ein Inhalt $\tilde{\rho}$ auf dem Mengenring $\mathfrak{R}$ liefert ein monotones positivlinares Funktional $\tilde{I}(\cdot)$ auf dem Verbandskegel $\mathfrak{R}^{\mathsf{Tr}+}$ der positiven $\mathfrak{R}$-Treppenfunktionen , wenn man definiert

$$\tilde{I}(h) = \sum \alpha_j \cdot \tilde{\rho}(A_j) \qquad \textit{für} \quad h = \sum \alpha_j \cdot 1_{A_j}.$$

Wenn $\tilde{\rho}$ ein Prämaß ist, dann gilt für absteigende Folgen von Treppenfunktionen

$$h_1 \geq h_2 \geq \quad \textit{mit} \quad \lim \downarrow h_n = 0 \ \textit{(punktweise)} \Longrightarrow \quad \lim \downarrow \tilde{I}(h_n) = 0.$$

Beweis. *Eine $\mathfrak{R}$-Treppenfunktion kann man auf sehr viele Weisen als Linearkombination von Indikatorfunktionen darstellen. Es gilt zu beweisen*

$$\sum a_j \cdot 1_{A_j} = h = \sum b_k \cdot 1_{B_k} \quad \Longrightarrow \quad \sum a_j \cdot \tilde{\rho}(A_j) = \sum b_k \cdot \tilde{\rho}(B_k).$$

1) Wir betrachten zuerst den Fall, wo sowohl die A_j als auch die B_k paarweise disjunkt sind. Wir können diese Tupel durch Mengen $A_0, B_0 \in \mathfrak{R}$ ergänzen, sodass wir zwei Partitionen einer Menge $\tilde{\Omega}$ erhalten, auf deren Komplement alle beteiligen Funktionen verschwinden. (Die Koeffizienten zu 1_{A_0} bzw. 1_{B_0} sind natürlich $a_0 = 0 = b_0$ zu setzen.) Es gilt

$$h = \sum_{j,k} c_{jk} \cdot 1_{A_j \cap B_k} \quad \textit{mit} \quad A_j \cap B_k \neq \emptyset \Rightarrow c_{jk} = a_j = b_k.$$

Es gilt daher wegen $B_k = \sum_j A_j \cap B_k, \quad A_j = \sum_k A_j \cap B_k$

$$\sum b_k \cdot \tilde{\rho}(B_k) = \sum_{j,k} c_{jk} \cdot \tilde{\rho}(A_j \cap B_k) = \sum a_j \cdot \tilde{\rho}(A_j).$$

2) Sei nun $h = \sum b_k \cdot 1_{B_k}$ *und* $\sum_j A_j = \tilde{\Omega}$ *eine Partition, die einen Mengenring erzeugt, welcher alle* B_k *enthält.*

$$B_k = \sum_{\{j:\, A_j \subseteq B_k\}} A_j, \qquad \tilde{\rho}(B_k) = \sum_{\{j:\, A_j \subseteq B_k\}} \tilde{\rho}(A_j) \quad \textit{für alle } k$$

$$h = \sum_{\{(j,k):\, A_j \subseteq B_k\}} b_k \cdot 1_{A_j} = \sum_j a_j \cdot 1_{A_j} \quad \textit{mit} \quad a_j = \sum_{\{k:\, A_j \subseteq B_k\}} b_k,$$

$$\sum b_k \cdot \tilde{\rho}(B_k) = \sum_{\{(j,k):\, A_j \subseteq B_k\}} b_k \cdot \tilde{\rho}(A_j \cap B_k) = \sum_j a_j \cdot \tilde{\rho}(A_j).$$

Da alle Darstellungen von h *mit paarweise disjunkte* A_j *denselben Wert liefern, liefern alle Darstellungen denselben Wert* $\tilde{I}(h)$. *Das Funktional* $\tilde{I}(\cdot)$ *ist auf* $\mathfrak{R}^{Tr+}$ *wohldefiniert. Es ist monoton und positivlinear.*

3) Wenn $0 = c_0 < c_1 < \cdots < c_M$ *die möglichen Werte von* h *sind, dann gilt*

$$h = \sum_m c_m \cdot 1_{C_m} \quad \textit{mit} \quad C_m = \{h = c_m\}$$

Bemerkenswert ist hier nun, dass wir $\tilde{I}(h)$ *als ein gewöhnliches Integral gewinnen können, als das Integral einer abnehmenden Sprungfunktion auf* $\mathbb{R}_+$. *In der Tat haben wir zunächst einmal für Funktionen, die (ausser der* 0*) nur einen einzigen Wert* c *annehmen können, d. h. für die Vielfachen von Indikatorfunktionen*

$$h = c \cdot 1_C \quad \Longrightarrow \quad \tilde{I}(c \cdot 1_C) = c \cdot \tilde{\rho}(C) = \int_0^\infty \tilde{\rho}(\{h > \lambda\})\, d\lambda;$$

für $h = \sum h_j = \sum a_j \cdot 1_{A_j}$ *mit paarweise disjunkten* A_j *wegen* $\quad \{h > \lambda\} = \sum \{h_j > \lambda\}$

$$\int_0^\infty \tilde{\rho}(\{h > \lambda\})\, d\lambda = \sum \int_0^\infty \tilde{\rho}(\{h_j > \lambda\})\, d\lambda = \tilde{I}(\sum h_j) = \tilde{I}(h).$$

4) Für eine punktweise nach 0 *absteigende Funktionenfolge* $(h_n)_n$ *gilt*

$$\forall \varepsilon > 0 \quad \big\{h_n > \varepsilon\big\} \searrow \emptyset \quad \textit{und daher für jedes Prämaß} \qquad \tilde{\rho}\{h_n > \varepsilon\} \downarrow 0, \quad \textit{sowie}$$

$$\tilde{I}(h_n) = \int_0^\infty \tilde{\rho}\{h_n > \lambda\}\, d\lambda \searrow 0.$$

Die Flächen unter den auf $\mathbb{R}_+$ *fallend nach Null strebenden 'Kurven'* $\quad \bar{F}_n(\lambda) = \tilde{\rho}\{h > \lambda\}$ *fallen nach Null. — Um das einzusehen, braucht man keine elaborierte Integrationstheorie.*

Wir könnten nun die Fortsetzungsidee von Daniell anwenden, um das Integral für Äquivalenzklassen integrabler messbarer Funktionen aus dem eben konstruierten Elementarintegral zu gewinnen. (Die Baire-σ-Algebra dazu ist die vom Mengenring $\mathfrak{R}$ erzeugte σ-Ring, welcher hier wegen der vorausgesetzten σ-Endlichtkeit des Prämaßes die erzeugte σ-Algebra $\mathfrak{A}$ ist.) Ein solches Vorgehen hat den Schönheitsfehler, dass man in den Zwischenschritten dieser Konstruktion die Definitionsbereiche der Mengenfunktionen und Funktionale etwas aus dem Auge verliert. Eine Unbequemlichkeit bereiten auch die Funktionswerte $\pm\infty$. Wir lassen daher die für die Fortsetzungsidee von Daniell essentielle Vektorraumstruktur vorerst beiseite. Stattdessen nützen wir die Kegel-Struktur der Menge aller nichtnegativen numerischen Funktionen.

Satz 2.3.8.
Es sei $\mathfrak{R}$ ein Mengenring über Ω_1 mit einer ausschöpfenden Folge $\Omega = \bigcup^\infty \mathsf{R}_n$.
$\mathfrak{A}$ sei die erzeugte σ-Algebra und $\mathfrak{A}^+$ der Verbandskegel der nichtnegativen $\mathfrak{A}$-messbaren numerischen Funktionen.
Zu jedem Prämaß ρ auf $\mathfrak{R}$ existiert dann genau ein Funktional $\mathrm{I}^+(\cdot)$ auf $\mathfrak{A}^+$ mit

i) $\mathrm{I}^+(1_\mathsf{R}) = \rho(\mathsf{R})$ *für alle* $\mathsf{r} \in \mathfrak{R}$,

ii) $\mathrm{I}^+(\alpha f + \beta g) = \alpha \mathrm{I}^+(f) + \beta \mathrm{I}^+(g)$ *für alle* $\alpha, \beta \in \mathbb{R}_+,\ f, g \in \mathfrak{A}^+$,

iii) $f_1 \leq f_2 \leq \cdots \quad \lim\uparrow f_n = f \quad \Longrightarrow \quad \lim\uparrow \mathrm{I}^+(f_n) = \mathrm{I}^+(f)$.

Beweis. *Die Anforderungen an I^+ implizieren, dass die Einschränkung auf die 1_A mit $A \in \mathfrak{A}$ eine Fortsetzung von ρ zu einem Maß μ ist; und dieses Maß ist σ-endlich. Wir haben bereits gesehen, dass es nur eine solche Fortsetzung gibt. Es ist klar, wie die Fortsetzung auf die $\mathfrak{A}$-Treppenfunktionen auszusehen hat. Und jedes $f \in \mathfrak{A}$ kann als aufsteigender Limes von solchen Treppenfunktionen gewonnen werden; ein Beispiel liefern die auf das nächste ganzzahlige Vielfache von $\frac{1}{2^n}$ abgerundeten Funktionen. Es gilt*

$$\mathrm{I}^+(f) = \int_0^\infty \mu(\{f > y\})\ dy.$$

Produktmaße und der Satz von Fubini .

Satz (Produkt-Ring)**.**
Es sei $\mathfrak{R}_1$ ein Mengenring über Ω_1 und $\mathfrak{R}_2$ ein Mengenring über Ω_2. Wir betrachten über dem cartesischen Produkt $\Omega = \Omega_1 \times \Omega_2$ das System $\mathfrak{S}$ aller Rechtecke $S = R_1 \times R_2$ und das System $\mathfrak{R}$ aller disjunkten Vereinigungen von Rechtecken. Es gilt

$$S', S'' \in \mathfrak{S} \quad \Longrightarrow \quad S' \cap S'' \in \mathfrak{S},\ S' \setminus S'' \in \mathfrak{R}.$$

Das System $\mathfrak{R}$ ist ein Mengenring.(Er wird mit $\mathfrak{R}_1 \otimes \mathfrak{R}_2$ bezeichnet.)

Der Beweis ist trivial.

Satz 2.3.9 (Produkt-Inhalt)**.**
Die Bezeichnungen seien wie eben. $\tilde\rho_1$ sei ein Inhalt auf $\mathfrak{R}_1$, $\tilde\rho_2$ ein Inhalt auf $\mathfrak{R}_2$. Es gibt dann genau einen Inhalt $\tilde\rho$ auf $\mathfrak{R}_1 \otimes \mathfrak{R}_2$ mit $\tilde\rho\big(R_1 \times R_2\big) = \tilde\rho(R_1) \cdot \rho_2(R_2)$ für alle Rechtecke. (Er wird mit $\tilde\rho_1 \otimes \tilde\rho_2$ bezeichnet.)

Wenn die 'Faktoren' $\tilde\rho_1, \tilde\rho_2$ Prämaße sind, dann ist auch das Produkt $\tilde\rho = \tilde\rho_1 \otimes \tilde\rho_2$ ein Prämaß.

Beweis. *Die ersten Aussagen sind trivial. Wir müssen zeigen*

$$A^{(1)} \supseteq A^{(2)} \supseteq \cdots \quad \in \mathfrak{R}_1 \otimes \mathfrak{R}_2, \quad \bigcap A^{(n)} = \emptyset \quad \Longrightarrow \quad \tilde{\rho}(A^{(n)}) \searrow 0.$$

Zuerst führen wir eine Notation ein, die in vielen Zusammenhängen nützlich ist.
Wie in der Konstruktion oben verstehen wir unsere Inhalte als Funktionale auf den betreffenden Systemen von Treppenfunktionen. Für eine $\mathfrak{R}_2$*-Treppenfunktion* $h_2 \in \mathfrak{R}_2^{\mathsf{Tr}+}$ *notieren wir* $\tilde{I}_2(h_2) = \int h_2(\omega_2)\ d\tilde{\rho}(\omega_2)$. *Entsprechend notieren wir die Elementarintegrale* $\tilde{I}_1(h_1)$ *für* $h_1 \in \mathfrak{R}_1^{\mathsf{Tr}+}$. *Für die Indikatorfunktion eines Rechtecks* $h = 1_{R_1 \times R_2}$ *notieren wir*

$$\int h(\omega_1, \omega_2)\ d\tilde{\rho}(\omega_1, \omega_2) = \int 1_{R_1}(\omega_1)\Big(\int 1_{R_2}(\omega_2)\Big)\ d\tilde{\rho}(\omega_1).$$

Für $h \in \big(\mathfrak{R}_1 \otimes \mathfrak{R}_2\big)^{\mathsf{Tr}+}$ *und* $\omega_1 \in \Omega_1$ *nennen wir die Funktion* $h_{\omega_1}(\cdot) = h(\omega_1, \cdot)$ *den Schnitt über* ω_1. *Das* ρ_2*- Integral des Schnitts über* ω_1 *hängt offenbar in der Form einer* $\mathfrak{R}_1$*−Treppenfunktion von* ω_1 *ab. Aus der Linearität des Elementarintegrals* $\tilde{I}$ *ergibt sich für beliebige* $h \in \big(\mathfrak{R}_1 \otimes \mathfrak{R}_2\big)^{\mathsf{Tr}+}$

$$\int h(\omega_1, \omega_2)\ d\tilde{\rho}(\omega_1, \omega_2) = \int\Big(\int h_{\omega_1}(\omega_2)\ d\tilde{\rho}(\omega_2)\Big)\ d\tilde{\rho}(\omega_1).$$

Für $h^{(n)} = 1_{A^{(n)}}$ *bilden die Schnitte über* ω_1 *eine Folge von* $\mathfrak{R}_2$*-Treppenfunktionen (die übrigens nur die Werte* 0 *und* 1 *annehmen), die gegen die Nullfunktion abfallen.*
$h^{(n)}_{\omega_1}(\cdot) = h^{(n)}(\omega_1, \cdot) \searrow 0$. *Die Prämaß-Eigenschaft von* ρ_2 *garantiert, dass die* ρ_2*- Integrale nach* 0 *fallen; und die Prämaß-Eigenschaft von* ρ_1 *liefert*

$$\tilde{\rho}(A^{(n)}) = \int h^{(n)}(\omega_1, \omega_2)\ d\tilde{\rho}(\omega_1, \omega_2) = \int\Big(\int h^{(n)}_{\omega_1}(\omega_2)\ d\tilde{\rho}(\omega_2)\Big)\ d\tilde{\rho}(\omega_1) \searrow 0.$$

Satz 2.3.10 (Satz von Fubini).
Gegeben seien σ*-endliche Prämaße* $\tilde{\rho}_i$ *auf* $\mathfrak{R}_i$ $(i = 1, 2)$. $\mathfrak{A}$ *bezeichne die von* $\mathfrak{R}_1 \otimes \mathfrak{R}_2$ *erzeugte* σ*-Algebra über* $\Omega_1 \times \Omega_2$ *. Das 'iterierte Integral'*

$$\mathfrak{A} \ni A \ \longmapsto \qquad \int\Big(\int 1_A(\omega_1, \omega_2)\ d\tilde{\rho}(\omega_2)\Big)\ d\tilde{\rho}(\omega_1).$$

ist die eindeutig bestimmte monotone Fortsetzung ρ *des Prämaßes* $\tilde{\rho} = \tilde{\rho}_1 \otimes \tilde{\rho}_2$.
Für beliebige $h \in \mathfrak{A}^+$ *gilt*

$$\int h\ d\rho = \int\Big(h(\omega_1, \omega_2)\ d\tilde{\rho}(\omega_2)\Big)\ d\tilde{\rho}(\omega_1).$$

Beweis. *Die Gesamtheit derjenigen nichtnegativen Funktionen* $h(\omega_1, \omega_2)$, *für welche jeder Schnitt* $h(\omega_1, \cdot)$ $\mathfrak{A}_2$*-messbar ist, enthält die nichtnegativen Treppenfunktionen; sie ist gegenüber monotoner Limesbildung abgeschlossen und umfasst daher* $\mathfrak{A}^+$. *Mit anderen*

Worten: Für jedes $f \in \mathfrak{A}^+$ *sind alle Schnitte messbar. Mit demselben Argument sieht man, dass die* ρ_2*-Integrale dieser Schnitte in* $\mathfrak{A}_1$*-messbarer Weise von* ω_1 *abhängen. Das iterierte Integral ist daher ein wohldefiniertes Funktional. Es ist positivlinear und aufsteigendstetig, und es liefert für die Indikatoren der Rechtecke die gewünschten Werte. Es ist daher gleich dem Funktional* $I^+(\cdot)$ *zum Produktprämaß* $\tilde{\rho}_1 \otimes \tilde{\rho}_2$.

Hinweis: Man kann natürlich die Reihenfolge der Faktoren des cartesischen Produkts Ω_i vertauschen. Der Satz liefert somit insbesondere die Vertauschbarkeit der Reihenfolge der Integrationen. In älteren Büchern wird diese Vertauschbarkeit der Integrationsreihenfolge der Satz von Fubini genannt. Die Mathematiker des 19. Jahrhunderts kämpften deswegen mit der Frage der Vertauschbarkeit, weil sie sich auch um Integrale von Funktionen h bemühen wollten, für die der Positivteil und der Negativteil unendliches Integral haben. $\int h^+ \, d\rho = +\infty = \int h^+ \, d\rho$. Die damals ins Auge gefassten Ansätze zu einer Verallgemeinerung des Integralbegriffs haben sich als nicht tragfähig herausgestellt. 'Uneigentliche Integrale' und 'bedingt integrable Funktionen' gibt es nicht in der modernen Integrationstheorie. Die Ausdrücke, die in der 'Integrationstheorie' des 19. Jahrhunderts gelegentlich behandelt werden sollten, werden in der modernen Integrationstheorie als Limiten von Integralen verstanden und mit der gebotenen Vorsicht behandelt.

Wenn man bei den Integranden f bleibt, für welche f^+ und f^- endliches Integral haben, dann gibt es keine Probleme mit dem schrittweisen Integrieren.

2.4 Integrabilität und die Banachräume $L^p(\Omega, \mathfrak{A}, \rho)$

Sprechweise.

Eine Menge Ω wird zu einem **messbaren Raum**, indem man eine σ-Algebra $\mathfrak{A}$ als das System der messbaren Mengen auszeichnet.

Ein messbarer Raum $(\Omega, \mathfrak{A})$ wird zu einem **Maßraum**, indem man ein σ-endliches Maß ρ auszeichnet.

Eine Maßraum $(\Omega, \mathfrak{A}, \rho)$ mit $\rho(\Omega) = 1$ heisst ein **Wahrscheinlichkeitsraum.**

In den vorhergehenden Unterabschnitten haben wir uns um die Konstruktion von Maßen bemüht.

Ausgehend von einem Elementarintegral $\tilde{I}$ auf einem Stone'schen Vektorverband $\tilde{\mathfrak{E}}$ haben wir ein Maß auf dem σ-Ring der $\tilde{\mathfrak{E}}$-Baire'schen Mengen konstruiert. Es handelt sich um eine σ-Algebra, wenn wir fordern (und das wollen wir für das Folgende), dass eine abzählbare Familie von Funktionen $\tilde{f}_i$ existiert mit $\bigcup\{\tilde{f}_i > 0\} = \Omega$.

Die Konstruktionsmethode für die Hausdorff-Maße im zweiten Unterabschnitt führt typischerweise auf Maße, die nicht σ-endlich sind. Wir werden sie nicht weiterverfolgen.

Die im dritten Unterabschnitt durchgeführte Fortsetzung eines (bei uns immer endlichwertigen) Prämaßes $\tilde{\rho}$ auf einem Mengenring $\mathfrak{R}$ zu einem Maß führt zu einem σ-endlichen Maß ρ auf der von $\mathfrak{R}$ erzeugten σ-Algebra $\mathfrak{A}$, wenn eine abzählbare Familie von Mengen R_i existiert mit $\bigcup R_i = \Omega$.

Alle diese Konstruktionen haben zu einem Funktional auf einem Verbandskegel von Funktionen mit Werten in $\mathbb{R}_+ \cup \{+\infty\}$ geführt, welches auch den Wert $+\infty$ annehmen kann. Über dieses Funktional I^+ und seinen Definitionsbereich $\mathfrak{A}^+$ soll jetzt genaueres gesagt werden:

Sprechweise (Messbare numerische Funktionen)**.**
Es sei $\mathfrak{A}$ eine σ-Algebra über Ω. Eine Funktion mit Werten in $\mathbb{R}_+ \cup \{\pm\infty\}$ heisst eine messbare numerische Funktion, wenn $\{f > \lambda\} \in \mathfrak{A}$ für alle λ.

Wenn die Funktion Werte in $\mathbb{R}_+ \cup \{+\infty\}$ hat, dann spricht man von einer nichtnegativen $\mathfrak{A}$-messbaren numerischen Funktion. Die Menge dieser Funktionen bezeichnen wir mit $\mathfrak{A}^+$.

Satz 2.4.1. *Die Funktionen in $\mathfrak{A}^+$ kann man positivlinear kombinieren; ausserdem gilt*

$$h_1, h_2, \ \ldots \quad \in \mathfrak{A}^+ \quad \Longrightarrow \quad \sup h_n \in \mathfrak{A}^+, \quad \inf h_n \in \mathfrak{A}^+.$$

$\mathfrak{A}^+$ ist also ein σ-vollständiger Verbandskegel numerischer Funktionen.

Zum Beweis, dass mit h_1, h_2 auch die Summe $h_1 + h_2$ zu $\mathfrak{A}^+$ gehört, bemerken wir $\{h_1 + h_2 > \lambda\} = \bigcup_r \{h_1 + h_2 > r\} \cap \{h_1 + h_2 > \lambda - r\}$, wenn r die abzählbare Menge der rationalen Zahlen durchläuft. Ebenso behandelt man Maximum und Minimum. Die übrigen Behauptungen sind trivial.

Wenn $(\Omega, \mathfrak{A}, \rho)$ ein Maßraum ist, dann definieren wir auf dem σ-vollständigen Verbandskegel $\mathfrak{A}^+$ das Funktional (welches auch den Wert $+\infty$ annehmen kann)

$$I^+(h) = \int h \; d\rho = \int_0^\infty \rho\{h > \lambda\} \; d\lambda.$$

Mit den oben entwickelten Argumenten ergibt sich der

Satz 2.4.2 (Die Integration positiver messbarer Funktionen).
Das Funktional $\mathrm{I}^+(\cdot)$ *zu* $\big(\Omega, \mathfrak{A}, \rho\big)$ *ist monoton mit*

1. $\mathrm{I}^+(a \cdot h + b \cdot k) = a \cdot \mathrm{I}^+(h) + b \cdot \mathrm{I}^+(k)$ *für alle* $a, b \geq 0$, $h, k \in \mathfrak{A}^+$,
2. $h_1 \leq h_2 \leq \cdots \in \mathfrak{A}^+ \implies h = \lim \uparrow h_n \in \mathfrak{A}^+$ *und* $\lim \uparrow \mathrm{I}^+(h_n) = \mathrm{I}^+(h)$.
3. $h_1 \geq h_2 \geq \cdots \in \mathfrak{A}^+$ *mit* $\mathrm{I}^+(h_1) < \infty \implies h = \lim \downarrow h_n \in \mathfrak{A}^+$ *und* $\lim \downarrow \mathrm{I}^+(h_n) = \mathrm{I}^+(h)$.

Warnung: Die Zusatzbedingung bei der absteigenden Stetigkeit 3) ist notwendig. Es gibt gegen Null absteigende Folgen von Funktionen, die allesamt das 'Integral' $+\infty$ haben.

Es ergeben sich unmittelbar zwei ausserordentlich nützliche Sätze

Satz 2.4.3 (Lemma von Fatou).
Für jeder Folge $f_n)_n \in \mathfrak{A}^+$ *gilt*

$$\int \liminf f_n \, d\rho \leq \liminf \int f_n \, d\rho.$$

Beweis. *Betrachte* $g_n = \sup\{f_m; m \geq n$ *Es handelt sich um eine absteigende Folge mit* $\lim \downarrow g_n = \liminf f_n$. *Das Integral des absteigenden Limes ist nicht größer als der absteigende Limes der Integrale* $\lim \downarrow \int g_n \, d\rho \geq \liminf \int f_n \, d\rho$.

Satz 2.4.4 (Satz von der majorisierten Konvergenz).
Es sei $(f_n)_n \in \mathfrak{A}^+$ *eine fastüberall konvergierende Folge*
$\Big(\liminf f_n = f = \limsup f_n$ *ausserhalb einer Nullmenge* $\Big)$.
Wenn ein h *existiert mit* $\forall n \; h \geq |f_n|$ *und* $\int h \, d\rho < \infty$, *dann gilt*

$$\int f_n \, d\rho \to \int f \, d\rho.$$

Beweis. *Das Lemma von Fatou, angewandt auf die Folge* $(h - f_n)_n$ *liefert*

$$\int h \, d\rho - \int \limsup f_n \geq \int h \, d\rho - \limsup \int f_n \, d\rho,$$

zusammen mit dem Fatou'schen Lemma also $\limsup \int f_n \, d\rho \leq \int f \, d\rho \leq \liminf \int f_n \, d\rho$.

Beispiel. Für $x \geq 0$ sei $f_1(x) = \frac{x}{1+x^4}$ und $f_n(x) = n \cdot f_1(nx)$. Es gilt $\int_0^\infty f_n(x) \, dx = \int_0^\infty f(x) \, dx = \frac{1}{2} \int_0^\infty \frac{1}{1+u^2} \, du$ für alle n, während $f_n(x) = \frac{n^2}{1+n^2 x^2} \to 0$ für alle x. Hier existiert keine integrable Funktion über allen f_n.

Hinweis: Wir haben den Satz von der majorisierten Konvergenz hier nur für nichtnegative f formuliert. Er gilt allgemeiner für fastsicher konvergierende Folgen von Funktionen, deren Absolutbetrag fastüberall von einer Funktion h mit $I^+(h) < \infty$ dominiert sind. Eine Verallgemeinerung dieses Satzes werden wir beweisen, wenn wir die sog. 'gleichgradig integrablen' fastüberall konvergenten Folgen behandeln.

Zuerst müssen wir aber Konstruktionen und Sprechweisen entwickeln, die sich auf das Rechnen mit numerischen Funktionen, die nicht notwendigerweise nichtnegativ sind, beziehen.

Sprechweise (Der Begriff 'ρ-fastüberall').
Von einer Funktion $h \in \mathfrak{A}^+$ sagen wir

1. h ist ρ-fastüberall endlich, wenn $\rho\{h = +\infty\} = 0$, $\quad \big(h < \infty \quad \rho - \text{f.ü.}\big)$
2. h ist ρ-Nullfunktion, wenn $\rho\{h > 0\} = 0$, $\quad \big(h = 0 \quad \rho - \text{f.ü.}\big)$
3. k ist ρ-fastüberall kleiner oder gleich h, wenn $\rho\{h > k\} = 0$, $\quad \big(k \leq h \quad \rho - \text{f.ü.}\big)$

Satz 2.4.5 ('Fastsichere Gleichheit' oder 'Gleichheit ρ-fastüberall').
Das Maß ρ liefert eine Äquivalenzrelation auf dem Verbandskegel $\big(\mathfrak{A}^+, +, \leq\big)$. in

$$k = h \quad \rho - \textit{f.ü.} \quad \Longleftrightarrow \quad k \leq h \quad \textit{und} \quad h \leq k \quad \rho - \textit{f.ü.}$$

Die Äquivalenzrelation ist verträglich mit der Addition und mit der Ordnungsrelation. Es gilt $k \leq h \quad \rho - \textit{f.ü.}$ *genau dann, wenn*

$$\forall \lambda \quad \{k > \lambda\} \subseteq \{h > \lambda\} \quad \rho - \textit{f.ü.} \qquad \textit{oder} \qquad \forall \lambda \quad \rho\Big(\{k > \lambda\} \setminus \{h > \lambda\}\Big) = 0$$

Für jede Folge von Äquivalenzklassen ist das Supremum und das Infimum eine wohldefinierte Äquivalenzklasse, und es gilt

$$\big(h = \sup h_n \quad \rho - \textit{f.ü.}\big) \quad \Longleftrightarrow \quad \forall \lambda > 0 \quad \big\{h > \lambda\big\} = \bigcup \big\{h_n > \lambda\big\} \quad \rho - \ \textit{f.ü.}$$

Der Beweis liegt auf der Hand.
Bemerke: Für jede Folge von Äquivalenzklassen $(h_n)_n$ ist der Limessuperior und der Limesinferior wohldefiniert und es gilt $\liminf h_n \leq \limsup h_n \quad \rho$- f.ü.
Man sagt, dass die Funktionenfolge $(h_n)_n$ ρ-fastüberall gegen h konvergiert, und notiert $\big(h_n \to h \quad \rho - \text{f.ü.}\big)$, wenn gilt $\liminf h_n = h = \limsup h_n \quad \rho$- f.ü.

Hinweis: Die ρ-fastsichere Konvergenz ist ein sonderbarer Begriff in dem Sinne, dass sie sich nicht durch eine Metrik beschreiben lässt, wie wir sehen werden. Trotzdem erweist sie sich (in Verbindung mit anderen Annahmen über die Folgen) in vielen Anwendungen als sehr nützlich. Wir werden sie später (in einem allgemeineren Rahmen) in Verbindung mit der sog. stochastische Konvergenz weiter behandeln.

Vektorraumstruktur
Die Menge der Äquivalenzklassen von $\mathfrak{A}$-messbaren $\mathbb{R}_+ \cup \{+\infty\}$-wertigen Funktionen bezeichnen wir mit $\big(\mathfrak{A}^+, \sim\big)$. Für den Augenblick bezeichne $\big(\mathfrak{A}^+_e, \sim\big)$ den Teilkegel der Äquivalenzklassen von ρ-fastüberall endlichwertigen Funktionen. Auf $\big(\mathfrak{A}^+_e, \sim\big)$ ist die Verknüpfung $(f, g) \mapsto f \backslash g = f - f \wedge g \; = \; \big(f - g\big)^+$ wohldefiniert. Es gelten die bekannten Regeln wie z. B. $\quad f \vee g = f + g \backslash f = g + f \backslash g$.

Auf der Menge der Paare (f, g) definieren wir die (bei der Differenzenbildung übliche) Äquivalenzrelation : $\qquad (f, g) \equiv (h, k) \quad \Leftrightarrow \quad f + k = h + g$
Die äquivalenten Paare werden durch die reellwertigen $\mathfrak{A}$-messbaren Funktionen repräsentiert mit der ausgezeichneten Darstellung

$$(f, g) \;\equiv\; (f \backslash g,\; g \backslash f) \quad \text{kurz geschrieben} \quad \equiv f - g = f \backslash g \;-\; g \backslash f.$$

Die Menge dieser 'Differenzen' bezeichnen wir (für den Augenblick) mit $\big(\mathfrak{A}_e, \sim\big)$, und wir nennen die Elemente die fastüberall endlichwertigen $\mathfrak{A}$-messbaren Funktionen. Es handelt sich um einen Vektorverband mit der Eigenschaft: Wenn eine Folge $f_1, f_2, \quad \in \big(\mathfrak{A}_e, \sim\big)$ überhaupt eine obere Schanke in $\big(\mathfrak{A}_e, \sim\big)$ besitzt, dann besitzt sie eine kleinste obere Schranke $f = \sup f_n$ $(\rho -$ f.ü.); und die Folge $(g_n)_n = (f_1 \vee f_2 \vee \cdots \vee f_n)_n$ konvergiert ρ-fastüberall gegen f. – Entsprechendes gilt für die Infimumbildung für Folgen in $\big(\mathfrak{A}_e, \sim\big)$: $\inf f_n$ $(\rho -$ f.ü.).

Es sind gewisse Teilräume des Raums $\big(\mathfrak{A}_e, \sim\big)$, die für die Integrationstheorie und die Funktionalanalysis interessant sind. Diese sollen jetzt diskutiert werden.

Summabilität; der Raum $L^1\big(\Omega, \mathfrak{A}, \rho\big)$
Die Funktion $f \in \mathfrak{A}_e$ heisst ρ-summabel oder ρ-integrabel, wenn $I^+(f^+) < \infty$, $I^+(f^-) < \infty$. Die Menge der ρ-summablen f wird mit $\mathcal{L}^1\big(\Omega, \mathfrak{A}, \rho\big)$ bezeichnet. Es handelt sich um einen Vektorverband von Funktionen mit der punktweisen Addition und der punktweisen Maximums- und Minimumsbildung.

Die Menge der Äquivalenzklassen dieser Funktionen mit $L^1\big(\Omega, \mathfrak{A}, \rho\big)$. Wir haben bereits gesehen, dass es sich um einen Vektorverband handelt. Wir müssen uns mit einem zu diesem Raum passenden Konvergenzbegriff befassen.

Satz 2.4.6. *Der Vektorraum* $L^1\big(\Omega, \mathfrak{A}, \rho\big)$ *wird durch die* 1*-Norm zu einem Banachraum:*

$$\|f\|_1 = \int |f| \; d\rho \; = \; I^+(f^+) \; + \; I^+(f^-).$$

Jede beschränkte messbare Funktion m *liefert ein beschränktes lineares Funktional*

$$L^1\big(\Omega, \mathfrak{A}, \rho\big) \ni f \longmapsto \quad \ell(f) = \int m \cdot f \; d\rho.$$

Beweis. *Offensichtlich ist* $\|\cdot\|_1$ *eine Norm auf dem Raum* L^1. *Die einzige Behauptung des Satzes, die wir noch nicht bewiesen haben, ist die Vollständigkeit. Es sei* $(g_n)_n$ *eine Cauchy-Folge. Es genügt zu zeigen, dass eine geeignet gewählte Teilfolge konvergiert. Wir* *wählen eine Teilfolge* $(f_n)_n$ *so, dass* $\|f_{n+1}-f_n\|_1 < \frac{1}{n^4}$. *Wir benützen die sog. Markov'sche Ungleichung*

$$\rho\big(\{|g|\geq a\}\big) \leq \frac{1}{a}\cdot\|g\|_1 \quad \textit{für alle } a>0,$$

die sich trivialerweise aus der Abschätzung $\quad |g| \geq a\cdot 1_{\{|g|\geq a\}} \quad$ *ergibt.*

Für die Mengen $A_n = \{|f_{n+1}-f_n| \geq \frac{1}{n^2}\}$ *gilt* $\rho(A_n)\leq\frac{1}{n^2}$, *und daher* $\rho\big(\bigcup_{m>n}A_m\big) \leq \sum_{m>n}\rho(A_m) < \frac{1}{n}$. *Die Menge der* ω, *die in unendlich vielen der* A_m *liegen, ist die Nullmenge* $\bigcap_n\bigcup_{m>n}A_m$. *Für die übrigen Punkte gilt* $\sum|f_{n+1}(\omega)-f_n(\omega)|<\infty$ *und es existiert der Limes* $\tilde{f}$: $f_n(\omega)\to\tilde{f}(\omega)$. *Diese Funktion ist summabel nach dem Lemma von Fatou:*

$$\int|\tilde{f}|\;d\rho = \int\liminf_n|f_n|\;d\rho \leq \liminf_n\int|f_n|\;d\rho < \infty.$$

Wir zeigen nun, dass die Folge f_n *im* L^1*-Sinn gegen die so identifizierte Funktion* $\tilde{f}$ *konvergiert. Wir wählen* N *so groß, dass für alle* $n>N$ *gilt* $\int|f_n-f_N|\;d\rho < \varepsilon$.
Wieder nach den Lemma von Fatou gilt

$$\int|\tilde{f}-f_N|\;d\rho = \int\liminf_n|f_n-f_N|\;d\rho \leq \liminf_n\int|f_n-f_N|\;d\rho < \varepsilon.$$

Rückblicke: Erinnern wir uns hier an das Fortsetzungsverfahren nach Daniell für ein Elementarintegral $\tilde{I}$ auf einem Stone'schen Vektorverband $\tilde{\mathfrak{E}}$.

(Wir nehmen an, dass es eine abzählbare Menge von Elementarfunktionen gibt, sodass $\Omega=\bigcup\{\tilde{f}_i>0\}$, dass also der von den $\{\tilde{f}>0\}$ erzeugte σ-Ring der $\tilde{\mathfrak{E}}$-Baire'schen Mengen eine σ-Algebra $\mathfrak{B}$ ist.)

Wir erhalten eine Semi-Norm auf $\tilde{\mathfrak{E}}$, wenn wir definieren $\|\tilde{f}\|_1=\tilde{I}(|\tilde{f}|)$. Auf dem Vektorraum der Äquivalenzklassen $\big(\tilde{\mathfrak{E}},\sim\big)$ ist das eine Norm. Aus unseren Überlegungen folgt, dass die Vervollständigung bzgl. dieser Norm den Raum $L^1\big(\Omega,\mathfrak{B},\rho\big)$ liefert, wo ρ die Einschränkung des Daniell-Maßes auf $\mathfrak{B}$ ist.

In dem Spezialfall, wo $\tilde{\mathfrak{E}}$ der Vektorverband der stetigen Funktionen (auf irgendeinem HRaB) ist, gingen wir in einem ersten Schritt zu den halbstetigen Funktionen, zu den Funktionenkegeln $\mathfrak{E}^{\uparrow}$ und $\mathfrak{E}^{\downarrow}$. Die Äquivalenzklassen von Funktionen in $\mathfrak{E}^{\uparrow}\cup\mathfrak{E}^{\downarrow}$ sind noch nicht alle Elemente des Banachraums $L^1\big(\Omega,\mathfrak{B},\rho\big)$. Man kommt aber zum Ziel mit den Äquivalenzklassen der Funktionen aus $\mathfrak{E}^{\uparrow\downarrow}\cup\mathfrak{E}^{\downarrow\uparrow}$ Wir haben in der Tat gesehen, dass für jedes Daniell-integrable h einschachtelnde Funktionen aus diese Funktionenmenge existieren.

Man könnte andererseits versuchen, ohne Äquivalenzbildung weiterzumachen mit monotoner Limesbildung, und weitere Funktionenräume konstruieren $\mathfrak{E}^{\uparrow\downarrow},\mathfrak{E}^{\downarrow\uparrow},\mathfrak{E}^{\uparrow\downarrow\uparrow},\ldots$. Man kann zeigen, dass man damit i. A nicht bis zur σ-Algebra $\mathfrak{B}$ gelangen kann. (Ohne Beweis!)

Als unsere Integralkonstruktion von den Prämaßen ausging, machten wir uns nur an einer Stelle, nämlich bei der Konstruktion von Produktmaßen Gedanken darüber, wie man zu einem Prämaß kommt.

Bei der Daniell-Konstruktion zu den Stietles-Integralen haben wir gründlicher angesetzt. Auf der Grundlage des Satzes von Dini erhalten wirklich ein Maß auf dem Stone'schen Verbands der stetigen Funktionen auf einem kompakten Raum. Die Konstruktion beweist offenbar auch den

Satz (Riesz'scher Darstellungssatz). *Es sei* $\tilde{\mathfrak{E}}$ *der Raum der stetigen Funktionen auf einem kompakten HRaB mit der Supremumsnorm. Jedes positive Linearform* $\ell(\cdot)$ *ist dann gegeben durch ein endliches Borel-Maß* μ $\quad \ell(f) = \int f \, d\mu$.

Die p- summablen Funktionen. In der Funktionalanalysis (nicht so sehr in der Stochastik) interessieren neben $L^1(\Omega, \mathfrak{A}, \rho)$ auch noch die Banachräume $L^p(\Omega, \mathfrak{A}, \rho)$ (für $p > 1$) und ganz besonders der Hilbertraum $L^2(\Omega, \mathfrak{A}, \rho)$.

Sprechweise. Es sei $(\Omega, \mathfrak{A}, \rho)$ ein σ-endlicher Maßraum. Eine komplexwertige $\mathfrak{A}$-messbare Funktion h heisst p-summabel, wenn $|h|^p$ summabel ist. Für solche Funktionen definiert man die p-Norm

$$N_p(h) = \left(\int |h|^p \, d\rho \right)^{1/p} = \|f\|_p.$$

Satz 2.4.7 (Hölder'sche Ungleichung).
Ist $p \geq 1$, $\frac{1}{p} + \frac{1}{q} = 1$, *und* h p*-summabel,* k q*-summabel, so ist das Produkt* $h \cdot k$ *summabel und es gilt*

$$\int |h \cdot k| \, d\rho \leq N_p(h) \cdot N_q(k).$$

Für jedes h *existiert ein* k*, sodass* $\langle h, k \rangle = \int h \cdot k \, d\rho = N_p(h) \cdot N_q(k)$*. Es gilt also*

$$N_p(h) = \sup\left\{ |\langle h, k \rangle| : \quad N_q(k) \leq 1 \right\}.$$

Beweis. *Es genügt, den Fall* $N_p(h) = 1 = N_q(k)$ *zu behandeln.*
Bekanntlich gilt $a^{1/p} \cdot b^{1/q} \leq \frac{1}{p} \cdot a + \frac{1}{q} \cdot b$*. für alle* $a, b \geq 0$*. Angewandt auf die Funktionen* $a = |h|^p$, $\quad b = |k|^q$ *erhalten wir* $|h \cdot k| \leq \frac{1}{p} \cdot |h|^p + \frac{1}{q} \cdot |k|^q$*. Und die Integration liefert die erste Behauptung.*

Zu h *mit* $N_p(h) = 1$ *assoziieren wir zunächst* $\tilde{k} = |h|^{p/q}$*. Es gilt* $N_q(\tilde{k}) = 1$ *und wegen* $a = |h|^p = |\tilde{k}|^q = b$ *haben wir* $|h| \cdot |\tilde{k}| = a^{1/p} \cdot b^{1/q} = \frac{1}{p} \cdot a + \frac{1}{q} \cdot b = \frac{1}{p} \cdot |h|^p + \frac{1}{q} \cdot |\tilde{k}|^q$*. Wenn* $h = |h| e^{i\phi}$*, dann leistet* $k = |\tilde{k}| e^{-i\phi}$*, das Verlangte.*

Eine offensichtliche Konsequenz ist die Subadditivität des Funktionals $N_p(\cdot)$; und das ist die Minkowski'sche Ungleichung $\quad \|f + g\|_p \leq \|f\|_p + \|g\|_p$.

Notation. Die Menge der p-summablen Funktionen wird mit $\mathcal{L}^p(\Omega, \mathfrak{A}, \rho)$ bezeichnet; die Menge der Äquivalenzklassen mit $L^p(\Omega, \mathfrak{A}, \rho)$ Man notiert $\| \cdot \|_p = N_p(\cdot)$.

Satz 2.4.8. *Der Raum* $L^p\big(\Omega, \mathfrak{A}, \rho\big)$ *wird durch die* p*-Norm* $\|\cdot\|_p = N_p(\cdot)$ *zu einem Banachraum. Der Dualraum zu* $L^p\big(\Omega, \mathfrak{A}, \rho\big)$ *ist der Raum* $L^q\big(\Omega, \mathfrak{A}, \rho\big)$ *für alle* $p \in [1, +\infty)$.

Beweis. *Wir beweisen die Vollständigkeit von* $L^p\big(\Omega, \mathfrak{A}, \rho\big)$ *für* $p > 1$ *ebenso wie für* $p = 1$. *Jetzt verlangen wir von der Teilfolge, die den Limes identifizieren soll:*

$$\int |f_{n+1} - f_n|^p \, d\rho = \big(N_p(f_{n+1} - f_n)\big)^p \le n^{-2-2p}.$$

Aufgrund der Markovschen Ungleichung $\rho\{|g| > \alpha\} \le \alpha^{-p} \cdot (N_q(g))^p$ *gewinnen wir Mengen* $A_n = \{|f_{n+1} - f_n| \ge \frac{1}{n^2}\}$ *mit* $\rho(A_n) < \frac{1}{n^2}$. *Sonst ist alles wie oben. Den Beweis, dass es auf* $L^p\big(\Omega, \mathfrak{A}, \rho\big)$ *neben den von den* $k \in L^q$ *erzeugten keine weiteren stetigen Linearformen gibt, verschieben wir auf den Abschnitt über unbestimmte Integrale.*

Notation. Für eine $\mathfrak{A}$-messbare Funktion h definiert das wesentliche Supremum:

$$\mathrm{ess\,sup}(h) = \inf\big\{\lambda : \ \rho\{h > \lambda\} = 0\big\} \qquad \text{und man notiert} \quad \|h\|_\infty = \mathrm{ess\,sup}(|h|).$$

Die Funktionen mit $\|h\|_\infty < \infty$ nennt man die wesentlich beschränkten Funktionen. Ihre Gesamtheit wird mit $\mathcal{L}^\infty\big(\Omega, \mathfrak{A}, \rho\big)$ bezeichnet, die Menge der Äquivalenzklassen mit $L^\infty\big(\Omega, \mathfrak{A}, \rho\big)$. Offenbar wird der Raum $L^\infty\big(\Omega, \mathfrak{A}, \rho\big)$ durch die Norm $\|\cdot\|_\infty$ zu einem Banachraum. Konvergenz bedeutet die gleichmäßige Konvergenz der Äquivalenzklassen.

Satz 2.4.9. *Wenn* f *summabel ist und* h *wesentlich beschränkt, dann ist das Produkt summabel mit* $\big|\int f \cdot h \, d\rho\big| \le \|f\|_1 \cdot \|h\|_\infty$. *Es gilt*

$$\begin{aligned} \|f\|_1 = &\quad \sup\big\{ |\langle f, k\rangle| : \quad \|h\|_\infty \le 1 \big\} \\ \|h\|_\infty = &\quad \sup\big\{ |\langle g, h\rangle| : \quad \|g\|_1 \le 1 \ \big\} \end{aligned}$$

$L^\infty\big(\Omega, \mathfrak{A}, \rho\big)$ *ist der Dualraum von* $L^1\big(\Omega, \mathfrak{A}, \rho\big)$.
Man beachte aber: Der Dualraum von $L^\infty\big(\Omega, \mathfrak{A}, \rho\big)$ *ist i.Allg. echt größer als der Raum* L^1 *(ohne Beweis!).*

Beweis. *Wenn* $\|h\|_\infty = 1$, *dann ist* $\ell(f) = \int f \cdot h$ *ein beschränktes lineares Funktional auf dem Raum* $\big(L^1, \ \|\cdot\big)$ *mit Norm* ≤ 1; *denn für jedes* $f \in L^1$ *gilt* $\big|\int f \cdot h \, d\rho\big| = \|f\|_1$.

Für $f \in L^1$ *und* $h = 1_{\{f>0\}} - 1_{\{f<0\}}$ *gilt* $\|f\|_1 = \int f \cdot h \, d\rho$.
Ist $\|h\|_\infty = 1$, *so gilt für jedes* $\varepsilon > 0$ $\rho\big(\{h > 1 - \varepsilon\}\big) > 0$ *oder* $\rho\big(\{-h > 1 - \varepsilon\}\big) > 0$. *Wir betrachten den ersten Fall. Wegen der* σ*-Endlichkeit des Maßes* ρ *existiert eine Teilmenge* $A_\varepsilon \subseteq \{h > 1 - \varepsilon\}$ *mit endlichem positivem Maß* $c_\varepsilon = \rho(A_\varepsilon)$. *Die Funktion* $g = (c_\varepsilon)^{-1} \cdot 1_{A_\varepsilon}$ *erfüllt* $\|g\|_1 = 1$ *und* $\int g \cdot h \, d\rho \ge 1 - \varepsilon$.

Im Unterschied zu den Fällen $< \infty$ *wird hier das Supremum* $\|h\|_\infty = \sup\big\{ |\langle g, h\rangle| : \|g\|_1 \le 1 \ \big\}$ *nicht notwendigerweise angenommen.*

Den Beweis, dass es auf dem L^1 *keine anderen stetigen Linearformen gibt als die von den* $h \in L^\infty$ *erzeugten Linearformen* $\langle \cdot, h\rangle$, *verschieben wir auf später.*

Hinweis: Aus der Sicht der Theorie der topologischen Vektorräume hat das Paar der Räume L^1, L^∞ die ausserordentlich unangenehme Eigenschaft der fehlenden Reflexivität.

Der Struktur des Raums $L^\infty(\Omega, \mathfrak{A}, \rho)$ hat viele Besonderheiten, die für die Maßtheorie eine Herausforderung sind. Man bemerke, dass alle Maße mit demselben Nullmengensystem zum gleichen Raum L^∞ führen. Familien von Maßen mit denselben Nullmengen sind ein aktuelles Thema der höheren Maßtheorie.

Beispiel 2.4.1. Es sei Ω der Raum $\mathbb{R}/2\pi$ mit dem verschiebungsinvarianten Wahrscheinlichkeitsmaß $\frac{1}{2\pi}\mathsf{d}t$. Für eine stetige 2π-periodische Funktion f sei also $\int f\, d\rho = \frac{1}{2\pi}\int f(t)\, dt$, wo das Integral über eine volle Periode zu erstrecken ist. Die Elemente von $L^p\big(\mathbb{R}/2\pi, \frac{1}{2\pi}dt\big)$ sind gegeben die Äquivalenzklassen borelmessbaren 2π-periodischer Funktionen f mit $\|f\|_p^p = \frac{1}{2\pi}\int |f(t)|^p\, dt$. Für $1 \le p < \infty$ liegen die trigonometrischen Polynome dicht in $L^p\big(\mathbb{R}/2\pi, \frac{1}{2\pi}dt\big)$. Man kann nämlich die Indikatorfunktion eines Intervalls durch trigonometrische Polynome in der p-Norm approximieren (ohne Beweis!); und man kann die $f \in L^p$ durch Treppenfunktionen in der p-Norm approximieren.

Beispiel 2.4.2. Besonders übersichtlich ist der normierte Vektorraum $L^2\big(\mathbb{R}/2\pi, \frac{1}{2\pi}dt\big)$; denn seine Elemente ν können auch durch die trigonometrischen Reihen $\sum c_n e^{int}$ mit quadratsummablen Koeffizientenfolgen $(c_n)_n$ repräsentiert werden. Die Darstellung ist umkehrbar eindeutig mit $\|\nu\|^2 = \sum |c_n|^2 = \frac{1}{2\pi}\int |f(t)|^2\, dt < \infty$. Wir haben eine Isometrie

$$\varphi : L^2\Big(\mathbb{R}/2\pi, \frac{1}{2\pi}\Big) \longrightarrow \ell^2(\mathbb{Z}).$$

Die Einträge c_n von $\varphi(f)$ heissen die Fourier-Koeffizienten der Äquivalenzklasse $f \in L^2$. Sie ergeben sich durch eine Integration $c_n = \frac{1}{2\pi}\int f\, dt$.

Zur Geschichte: Die Entdeckung, dass man die Elemente der Vervollständigung des Raums der trigonometrischen Polynome in der 2-Norm durch Äquivalenzklassen von Funktionen beschreiben kann, geht auf F.Riesz und E. Fischer (1906) zurück. Die Voraussetzung für diese Einsicht war natürlich das Verständnis von Nullmengen in der Integrationstheorie von Lebesgue (1902).

Zum Abschluss des Teilabschnitts wollen wir noch auf direktem Wege beweisen, dass der Banachraum l^p der p-summablen Folgen tatsächlich durch die Vervollständigung des Raums der finiten Folgen entsteht, wenn man diesen mit der p-Norm ausstattet. ($p \neq \infty$) Die Elemente des Raums nennt man wohl besser komplexe Gewichtungen der abzählbaren Grundmenge Ω mit dem Zählmaß. Wir können $\Omega = \mathbb{Z}$ annehmen. Da es in diesem Fall keine Probleme mit Nullmengen gibt, treten die übrigen für den Beweis benötigten maßtheoretischen Ideen besonders klar hervor. Es sind das im Wesentlichen das Lemma von Fatou und die punktweise Konvergenz als Vorstufe zur Norm-Konvergenz.

Satz 2.4.10. *Die Vervollständigung des Raums der finiten Gewichtungen von $\mathbb{Z}$ in der p-Norm (bzgl. des Zählmaßes) ist der Raum l^p der p-summablen Folgen (d. h. Gewichtungen).*

Beweis. *Es sei* $\left(\mathbf{c}^{(k)}\right)_{k=1,2,\ldots}$ *eine Folge von finiten (komplexen) Gewichtungen der Grundmenge* $\mathbb{Z}$*:* $\mathbf{c}^{(k)} = \{c_n^{(k)} : n \in \mathbb{Z}\}$*. Die Folge sei eine Cauchy-Folge bezüglich der* p*-Norm. Offensichtlich konvergieren die Gewichte in jeder Position:* $\forall n : \quad \lim_k c_n^{(k)} = \tilde{c}_n$*. Wir müssen zeigen, dass die Gewichtung* $\tilde{\mathbf{c}}$ p*-summabel ist, und dass unsere Cauchy-Folge im Sinne der* p*-Norm dagegen konvergiert. Das Lemma von Fatou garantiert*

$$\sum_n |\tilde{c}_n|^p \leq \liminf_k \sum_n |c_n^{(k)}|^p = \liminf_k \big(N_p(\mathbf{c}^{(k)})\big)^p < \infty.$$

Die Grenzgewichtung ist also p*-summabel.*

Sei K *so groß, dass für alle* $k > K$ *gilt* $\quad N_p(\mathbf{c}^{(k)} - \mathbf{c}^{(K)}) < \varepsilon$*, d. h.* $\sum_n |c_n^{(k)} - c_n^{(K)}|^p < \varepsilon^p$*. Nach dem Lemma von Fatou gilt dann*

$$\sum_n |\tilde{c}_n - c_n^{(K)}|^p = \sum_n \liminf_k |c_n^{(k)} - c_n^{(K)}|^p \leq \liminf_k \sum_n |c_n^{(k)} - c_n^{(K)}|^p < \varepsilon^p.$$

Die $\mathbf{c}^{(k)}$ *liegen also schliesslich beliebig nah an der Grenzgewichtung* $\tilde{\mathbf{c}}$*. Das Argument funktioniert auch für Cauchy-Folgen nichtfiniter* p*-summabler Gewichtungen.*

2.5 Unbestimmte Integrale und der Satz von Radon-Nikodym

In der elementaren Analysis ist 'unbestimmtes Integral' ein Synonym für 'Stammfunktion'. In der Maßtheorie ist das ganz anders. Das unbestimmte Integral einer ρ-summablen Funktion h ist eine Mengenfunktion $\nu(\cdot) = \int_{\cdot} h \; d\rho$. Im Falle $h \geq 0$ handelt es sich um ein endliches Maß, im allgemeinen Fall um ein 'signiertes Maß', das heisst um die Differenz zweier endlicher Maße.

Definition 2.12. Es sei $h \in L^1\big(\Omega, \mathfrak{A}, \rho\big)$ und $\nu(A) = \int 1_A \cdot h \; d\rho$ für $A \in \mathfrak{A}$. Die Mengenfunktion $\nu(\cdot)$ heisst dann das signierte Maß mit der Dichte h bzgl. ρ oder auch das unbestimmte Integral von h bzgl. ρ.

Wenn zwei Funktionen h_1 und h_2 sich nur auf einer ρ-Nullmenge unterscheiden, dann liefern sie dasselbe unbestimmte Integral. Wenn h_1 und h_2 nicht ρ-fastsicher gleich sind, dann sind die unbestimmten Integrale verschieden. Man notiert

$$d\nu = h \cdot \; d\rho \qquad \text{oder auch} \quad \frac{d\nu}{d\rho} = h \quad \rho\text{-f.ü.}$$

Erste Bemerkungen: Das unbestimmte Integral einer Funktion $h \geq 0$ mit endlichem Integral ist ein endliches Maß mit der Gesamtmasse $\nu(\Omega) = \int h \; d\rho$. Die Additivität von $\nu(\cdot)$ ist offensichtlich. Die σ-Additivität ergibt sich aus dem Satz von der monotonen Konvergenz: Wenn nämlich $A_1 \supset A_2 \supset \cdots$ mit $\bigcap A_n = \emptyset$, dann gilt $h \geq h \cdot 1_{A_n} \searrow 0$. Die Folge der Integrale fällt nach 0 nach dem Satz von der majorisierten Konvergenz.

Jede ρ-Nullmenge ist eine ν-Nullmenge. Man sagt deswegen, dass das endliche Maß ν totalstetig ist bzgl. ρ und man notiert $\nu \ll \rho$. Mit dem Begriff der Absolutsteitigkeit müssen wir uns gründlich beschäftigen.

Bevor wir uns der allgemeinen Theorie zuwenden, erinnern wir uns kurz an die Stieltjes-Integrale und -Maße auf der Borelalgebra $\mathfrak{B}$ über $\mathbb{R}$. In diesem Rahmen ist nämlich das Phänomen absolutstetiger Maße zuerst beachtet worden.

Satz.
Ist $G(t)$ eine nichtfallende rechtsstetige Funktion auf $\mathbb{R}$ mit $\lim_{T\to\infty}\big(G(T)-G(-T)\big) < \infty$, so gibt es genau ein endliches Maß ν auf $\mathfrak{B}$ mit $\nu\big((a,b]\big) = G(b) - G(a)$.

Jedes endliche Maß auf $\mathfrak{B}$ entsteht auf diese Weise, und die 'Verteilungsfunktion' G ist bis auf eine additive Konstante eindeutig bestimmt.

Spezialfälle: Wenn eine Funktion $p(t)$ existiert mit $G(b) - G(a) = \int_a^b p(t) \; dt$, dann sagt man, dass das Stieltjes-Maß ν totalstetig (oder 'absolutstetig') ist bzgl. des Lebesgue-Maßes. Man notiert $d\nu(t) = p(t) \; dt$ oder auch $\frac{d\nu}{dt} = p$ (Lebesque-f.ü.)

Bei einem totalstetigen W-Maß hat man $p(t) \geq 0$, $\int_{-\infty}^{+\infty} p(t) \; dt = 1$.

Neben den absolutstetigen Maßen gibt es weitere besondere Typen von Maßen. Wenn G eine Sprungfunktion ist, dann heisst das dazugehörige Maß ein diskretes Maß. In diesem Falle existiert eine abzählbare Familie von Paaren $(t_i, p_i)_{i\in I}$, sodass gilt

$G(t) = \sum_{\{i:\, t_i \leq t\}} p_i$. Man sagt, dass das diskrete Maß durch die Gewichtung des Punkte mit den Gewichten p_i entsteht. Man notiert manchmal $\mu = \sum_{i \in I} p_i \cdot \delta_{t_i}$.

Ein beliebtes konkretes Beispiel für ein atomloses Maß, welches nicht totalstetig ist ergibt sich aus der Cantorfunktion, die wir an anderer Stelle vorgestellt haben. Von Maßen dieser Art sagt man, dass sie singulär sind, genauer: singulär zum Lebesgue-Maß.

Maße, die zueinander singulär sind (oder 'disjunkt', wie man auch sagt), werden uns im allgemeinen Rahmen weiter beschäftigen.

Man bemerke, dass eine konvexe Kombination von Verteilungsfunktionen eine Verteilungsfunktion ergibt. Bei den dazugehörigen Wahrscheinlichkeitsmaßen spricht man von einer (konvexen) Mischung. Es ist leicht zu sehen, dass man jedes Wahrscheinlichkeitsmaß auf $\mathfrak{B}$ in eindeutiger Weise in drei Bestandteile zerlegen kann, einen diskreten, einen bzgl. des Lebesgue-Maßes totalstetigen und einen zum Lebesgue-Maß singulären atomlosen Anteil.

Bei den Stieltjes-Maßen ist $\int_{-\infty}^{+\infty} f(t)\ dG(t)$ eine gebräuchliche Bezeichnung für das Integral der (nichtnegativen) Funktion f. Das passt zu der Notation $\int f\ dG$, die uns bei den Kurvenintegralen begegnet ist. G war da allerdings nicht notwendigerweise steigend. Die Voraussetzungen waren dort so, dass die Funktion G als Differenz zweier steigender Funktionen geschrieben werden kann.

Konvention: Bei unseren Überlegungen zum Integral werden wir uns, wenn immer das möglich ist auf die Integration nichtnegativer Funktionen beschränken. Wenn der Integralwert $+\infty$ nicht zuzulassen, wird das gesagt. Sollten irgendwo signierte Maße oder Funktionen auftauchen, die auch negative Werte annehmen, werden wir diese Objekte immer schnell in ihre Positiv- und Negativteil aufspalten. Maße, die nicht σ-endlich sind, werden nicht in Betracht gezogen.

Definition 2.13 (Absolutstetiger Anteil).
Es seien ν und ρ Inhalte auf dem Mengenring $\mathfrak{R}$ über Ω mit $\nu(\cdot)$ beschränkt. Der Inhalt $\nu_a = \lim \uparrow_{M\to\infty} \nu \wedge (M \cdot \rho)$ heisst der bzgl. ρ absolutstetige Anteil von ν. Der Inhalt $\nu_s = \lim \downarrow_{M\to\infty} \nu \backslash (M \cdot \rho) = \nu - \nu_a$ heisst der bzgl. ρ singuläre Anteil von ν.
Man sagt, dass ν bzgl. ρ absolutstetig (oder 'totalstetig') ist, und notiert $\nu \ll \rho$, wenn der singuläre Anteil das Nullmaß ist.

Es ist bequem den Begriff Totalstetigkeit auch dann zu gebrauchen, wenn ρ ein σ-endliches Maß auf einer Mengenalgebra $\mathfrak{R}$ ist. Es gilt $\nu \ll \rho \iff 1_R \cdot \nu \ll 1_R \cdot \rho$ für alle r mit $\rho(R) < \infty$. Vom Inhalt (oder Maß) ν verlangen wir, dass $\nu(\cdot)$ auf $\mathfrak{R}$ beschänkt ist.

Wir werden sehen: Wenn der Definitionsbereich der Inhalte ρ und ν eine σ-Algebra ist, dann bedeutet Totalstetigkeit, dass jede ρ-Nullmenge auch ν-Nullmenge ist.

$$\nu \ll \rho \iff \quad \rho(N) = 0 \Rightarrow \nu(N) = 0.$$

Wir wollen hier aber zunächst bei den Prämaßen ν bleiben. Das System der Nullmengen für ein Prämaß ρ ist nämlich i. Allg. nicht so reichhaltig, dass man die Totalstetigkeitseigenschaft darauf basieren könnte. Hier müssen wir mit Mengen von kleinem ρ-Maß operieren.

Satz 2.5.1.
Es seien ρ und ν endliche Prämaße auf einer Mengenalgebra $\mathfrak{R}$ über Ω.

Wenn ν nicht bzgl. ρ totalstetig ist, wenn also $\big(\nu\backslash M\rho\big)(\Omega) \searrow \beta > 0$, dann existiert zu jedem M und $\varepsilon > 0$ eine Menge $\Omega_{M,\varepsilon} \in \mathfrak{R}$ mit
$$\rho(\Omega_{M,\varepsilon}) < \tfrac{1}{M} \cdot \nu(\Omega), \qquad \nu(\Omega_{M,\varepsilon}) > (1-\varepsilon)\beta.$$
Wenn $\nu \ll \rho$, dann gilt
$$\forall \varepsilon > 0 \,\exists\, \delta > 0 \;\forall A \qquad \big(\rho(A) < \delta\big) \Rightarrow \big(\nu(A) < \varepsilon\big).$$

Beweis. *Bekanntlich gilt $\big(\nu\backslash\rho\big)(\Omega) = \sup\big\{\sum_i\big(\nu(R_i) - \rho(R_i)\big)^+ : \; \sum R_i = \Omega\big\}$, wo das Supremum über alle disjunkten Zerlegungen von R zu erstrecken ist. Die Summe vergrößert sich, wenn man die Zerlegung verfeinert. Für eine gegebene Partition von Ω sei $\Omega_{M,\varepsilon} = \sum_{i \in I_+} R_i$, wo I_+ die Menge der Indizes ist, für welche $\nu(R_i) - \rho(R_i) > 0$.*
Zu jedem M und $\varepsilon > 0$ können wir die Zerlegung $\Omega = \sum R_i$ so fein wählen, dass für die Menge $\Omega_{M,\varepsilon}$ gilt $\nu(\Omega_{M,\varepsilon}) - M \cdot \rho(\Omega_{M,\varepsilon}) > (1-\varepsilon)\beta$.

Im Falle $\nu \ll \rho$ wählen wir M so groß, dass $\big(\nu\backslash M\rho\big)(\Omega) < \varepsilon/2$ und $\delta = \frac{\varepsilon}{2M}$. Für jede Menge A mit $\rho(A) < \delta$ gilt $\quad \nu(A) = \big(\nu\backslash M\rho\big)(A) + \big(\nu\wedge M\rho\big)(A) < \varepsilon/2 + M\cdot\rho(A) < \varepsilon.$

Hinweise: 1) Die Bedeutung des Begriffs der Totalstetigkeit hat als erster G. Vitali (1905) erkannt. Bei Vitali ist es eine Bedingung an eine Funktion F auf einem kompakten Intervall, die zum Längenmaß in Beziehung gesetzt wird.

Definition (G. Vitali)**.** Eine reelle oder komplexwerige Funktion F auf dem Intervall $[a, b]$ heisse absolut stetig, wenn zu jedem $\varepsilon > 0$ ein $\delta > 0$ existiert, sodass
$$\sum \big|F(b_k) - F(a_k)\big| \;<\; \varepsilon$$
gilt für alle Zerlegungen $a \le a_1 < b_1 \le a_2 < \cdots \le a_n < b_n \le b$ mit $\sum\big(b_k - a_k\big) < \delta$.

2) Es wird bei Vitali nicht angenommen, dass F reellwertig und monoton ist. Realteil und Imaginärteil einer absolut stetigen Funktion sind jedoch offenbar Funktionen mit beschränkter Schwankung und lassen sich daher als Differenz von monoton steigenden Funktionen schreiben.

3) Lebesgue, Vitali und andere haben aufgeklärt, dass die absolute Stetigkeit von F garantiert, dass die Funktion in Lebesgue-fast allen Punkten differenzierbar ist, und dass die Ableitung $f = F'$ Lebesgue-integrierbar ist und dass gilt $\int_c^d f(t)\; dt = F(d) - F(c)$ für alle $[c, d] \subseteq [a, b]$.

4) Wir wollen in unserer Maßtheorie darauf nicht eingehen. Die punktweise Differentiation ist für die Maßtheorie uninteressant. Wenn wir von einer absolutstetigen Funktion F auf einem Intervall $[a, b]$ sprechen, dann meinen wir eine Funktion, die als unbestimmtes Integral einer Lebesgue-integrablen Funktion f gewonnen werden kann. Die Berechtigung dazu liefert ein Satz von Vitali, der besagt, dass eine Funktion genau dann absolut stetig ist im obigen Sinn, wenn sie ein unbestimmtes Integral ist.

5)Wir werden sehen, dass diese Kennzeichnung der Totalstetigkeit konform geht mit dem allgemeinen Begriff der Totalstetigkeit von endlichen Maßen, den wir oben vorgestellt haben: Totalstetigkeit bedeutet (nach dem fundamentalen Satz von Radon-Nikodym) die Existenz einer 'Dichte'.

Satz 2.5.2.
Es sei $\big(\Omega, \mathfrak{A}, \rho\big)$ ein σ-endlicher Maßraum. Ein endliches Maß ν ist genau dann absolutsstetig, wenn jede ρ-Nullmenge eine ν-Nullmenge ist.

Beweis. *Die hier zu beweisende Kennzeichnung der Absolutstetigkeit $\nu \ll \mu$ mittels der Nullmengen benötigt ganz wesentlich, dass der Definitionsbereich $\mathfrak{A}$ eine σ-Algebra ist. Für den Beweis des Satzes nehmen wir an, dass ν nicht absolutstetig ist, und wir konstruieren eine Menge A^* mit $\rho(A^*) = 0$ und $\nu(A^*) > 0$. Sei also ε^* eine positive Zahl, sodass zu jedem $n \in \mathbb{N}$ eine Menge A_n existiert mit $\mu(A_n) < \frac{1}{2^n}$ und $\rho(A_n) \geq \varepsilon^*$. Für die Menge $B_m = \bigcup_{n>m} A_n$ gilt $\mu(B_m) < \frac{1}{2^m}$ und $\nu(B_m) \geq \varepsilon^*$. Die B_m bilden eine absteigende Folge. Der Durchschnitt $A^* = \bigcap B_m = \bigcap_m \bigcup_{n>m} A_n$ leistet das Gewünschte. (Man nennt die Menge A^* manchmal den Limessuperior der Mengenfolge $(A_n)_n$.)*

Bemerkung: Die hier bewiesene Kennzeichnung der Absolutstetigkeit ist kurz und bündig; sie läßt sich so auch wörtlich auf Paare σ-endlicher Maße übertragen. Sie sagt dann nichts anderes, dass die Einschränkungen dieser Maße auf Mengen mit endlichem Maß totalstetig sind.

Die oben abgeleitete $\varepsilon\delta$-Definition passt nur im Falle endlicher Maße. Sie hat aber große Vorzüge. Erstens passt sie auch auf endliche(!) Prämaße. Und zweitens lässt sich an sie in offensichtlicher Weise der (später zu diskutierende) Begriff der gleichgradigen Absolutstetigkeit von Familien von Maßen oder Prämaßen anknüpfen.

Satz 2.5.3 (Existenz disjunkter Träger für disjunkte Maße)**.**
Sind μ und ν irgendwelche σ-endliche Maße, so existieren messbare Mengen $\tilde{M}$ und $\tilde{N} = \complement\tilde{M}$, sodass gilt

$$\big(\mu\backslash\nu\big)(B) = 0 \quad \textit{für alle } B \subseteq \tilde{N}$$
$$\big(\nu\backslash\mu\big)(B) = 0 \quad \textit{für alle } A \subseteq \tilde{M}$$

Beweis. *1) Es genügt, den Fall endlicher Maße zu untersuchen. Wenn nämlich $\Omega = \sum \Omega_k$ mit $\big(\mu + \nu\big)(\Omega_k) < \infty$, dann kann man die Aufspaltung in Trägermengen auf jedem Ω_k separat vornehmen.*

2) Man kann die Aussage des Satzes auch so formulieren: Wenn μ und ν disjunkte endliche Maße sind in dem Sinn

$$0 = \mu \wedge \nu\ (\Omega) = \inf\Big\{\sum \mu(R_i) \wedge \nu(R_i) :\ \sum R_i = \Omega\Big\},$$

dann existieren disjunkte Mengen $\tilde{M}$ *und* $\tilde{N} = \complement\tilde{M}$*, sodass*

$$\mu(R) = \big(\mu\backslash\nu\big)(R) = \mu(R \cap \tilde{M}), \qquad \nu(R) = \big(\nu\backslash\mu\big)(R) = \nu(R \cap \tilde{N}) \quad \textit{für alle } R.$$

Die Mengen $\tilde{M}$, $\tilde{N}$ *sind bis auf eine* $(\mu\backslash\nu + \nu\backslash\mu)$*-Nullmenge eindeutig bestimmt.*

3) Zu jedem $\varepsilon = \varepsilon_n$ *existiert eine Partition der Grundmenge* $\Omega = \sum_i R_i^{(n)}$ *sodass*

$$\sum \big(\mu(R_i) - \nu(R_i)\big)^+ \geq \ \big(\mu\backslash\nu\big)\ (\Omega) - \varepsilon_n$$
$$\sum \big(\mu(R_i) - \nu(R_i)\big)^- \geq \ \big(\nu\backslash\mu\big)\ (\Omega) - \varepsilon_n$$

Es sei $I^+ = \{i : \mu(R_i > \nu(R_i)\}$, $\quad I^- = \complement I^+$, *und*

$$M^{(n)} = \sum_{i \in I^+} R_i^{(n)}, \qquad N^{(n)} = \sum_{i \in I^-} R_i^{(n)} = \complement M^{(n)}.$$

Für jede Partition einer Menge $B \subseteq N^{(n)}$, $B = \sum_j S_j$ *gilt* $\sum \big(\mu(S_j - \nu(S_j)\big)^+ < \varepsilon_n$, *also* $\mu\backslash\nu\,(B) < \varepsilon_n$.

Wir wählen eine summable Folge ε_n *und dazu passende Partitionen* $\Omega = \sum_i R_i^{(n)}$, *und gewinnen dazu*

$$\tilde{N} = \bigcap_m \bigcup_{n \geq m} N^{(n)}, \qquad \tilde{M} = \bigcap_m \bigcup_{n \geq m} M^{(n)}$$

Für $B \subseteq \bigcup_{n \geq m} N^{(n)}$ *gilt* $\mu\backslash\nu\,(B) < \sum_{n \geq m} \varepsilon_n$.
Für B *im (absteigenden) Durchschnitt* $\tilde{N}$ *gilt* $\mu\backslash\nu(B) = 0$. *Der Limes superior der Mengenfolge* $N^{(n)}$ *wird vom Maß* $\mu\backslash\nu$ *nicht getroffen.*
Entsprechendes gilt für das Maß $\nu\backslash\mu$*; Träger* $\tilde{N} = \limsup_n N^{(n)}$ *ist ein Träger dieses Maßes.*

Der eben bewiesene Satz ist im Wesentlichen der Zerlegungssatz von H. Hahn aus dem Jahr 1921. Er passt auch auf die Zerlegung eines endlichen Maßes in den singulären und den absolutstetigen Anteil

$$\nu = \nu_a + \nu_s \quad \text{mit} \qquad \nu_a = \lim_{M \to \infty} \uparrow \nu \wedge (M \cdot \rho) \qquad \nu_s = \lim_{M \to \infty} \downarrow \nu\backslash(M \cdot \rho).$$

Den folgenden Satz formulieren wir als Vorbereitung für den Beweis des Satzes von Radon-Nikodym aus dem Jahr 1930.

Satz. *Sei* ν *ein endliches Maß auf dem* σ*-endlichen Maßraum* $\big(\Omega, \mathfrak{A}, \rho\big)$. *Zu jedem* $r > 0$ *existiert dann eine Menge* Ω_r *sodass*

$$A \subseteq \Omega_r \Rightarrow \nu(A) \geq r \cdot \rho(A), \qquad B \subseteq \complement\Omega_r \Rightarrow \nu(A) \leq r \cdot \rho(A).$$

Wenn $r > s$, *dann ist* $\Omega_r\backslash\Omega_s$ *eine* ν*-Nullmenge. Es gilt* $\nu \ll \rho$ *genau dann, wenn* $\nu(\Omega_r) \searrow 0$ *für* $\to \infty$.
Wenn Ω_r' *und* Ω_r'' *das Verlangte leisten, dann gilt* $\nu(C) = r \cdot \mu(C)$ *für alle* $C \subseteq \Omega_r' \triangle \Omega_r''$.

Wenn wir für alle rationalen r *eine Wahl* $\tilde{\Omega}_r$ *treffen und dazu konstruieren* $\Omega_y = \bigcup_{r>y} \tilde{\Omega}_r$*, dann leistet* Ω_y *das Verlangte für jedes* $y > 0$*, und es existiert eine Funktion* h *mit* $\{\omega : h(\omega) > y\} = \Omega_y$ *für alle* $y \in \mathbb{R}_+$*.*

Der erste Teil ist eine Spezialisierung des obigen Satzes auf die zueinander disjunkten Maße $\big(\nu\backslash(r\cdot\rho)\big) \wedge \big((r\cdot\rho)\backslash\nu\big) = 0$. Die weiteren Teile liegen auf der Hand. Für jede Funktion h ist $\big\{\{h(\cdot) > y\} : \quad y \in \mathbb{R}_+\big\}$ eine rechtsstetige Schar von Mengen. Die Funktion $\bar{H}_B(y) = \rho(B \cap \{h > y\})$ ist rechtsstetig und abnehmend als Funktion von $y \in \mathbb{R}_+$ für jede messbare Menge B. Wir werden sehen, dass ihr Integral (in der üblichen Weise über die Halbachse $\mathbb{R}_+$ erstreckt den Wert $\nu(B)$ ergibt.

Satz 2.5.4 (Satz von Radon-Nikodym)**.**
Sei $(\Omega, \mathfrak{A}, \rho)$ *ein* σ*-endlicher Maßraum und* ν *ein endliches Maß auf* $\mathfrak{A}$*, welches bzgl.* ρ *totalstetig ist. Dann existiert eine messbare Funktion* h *sodass gilt*

$$\nu(B) = \int 1_B h \, d\rho \quad \textit{für alle } B \in \mathfrak{A}.$$

Die Funktion ist eindeutig bestimmt bis auf eine ρ*-Nullfunktion.*

Beweis. *Wir konstruieren wie oben Mengen* Ω_r *für die Zahlen* $r = k \cdot 2^{-n}$*. Wir können annehmen* $r > s \Rightarrow \Omega_r \subseteq \Omega_s$*.*

Die Absolutstetigkeit garantiert $\nu(\Omega_r) \searrow 0$ *für* $r \to 0$*. Wir können annehmen* $\Omega_\infty = \bigcap_r \Omega_r = \emptyset$*, denn das gesuchte* h *verschwindet ohnehin auf dieser Menge.*

Wir beschränken unsere Untersuchung zunächst auf Mengen B *mit* $B \subseteq \Omega_\delta$ *für ein* $\delta > 0$*. Wir haben daher* $\nu(B) \geq \delta \cdot \rho(B)$*. Dadurch bekommen wir die gesuchte Funktion* h *auf der verkleinerten Grundmenge* Ω_δ *mit endlichem* ρ*-Maß* $\rho(\Omega_\delta) \leq \frac{1}{\delta}\nu(\Omega_\delta)$*. Auf der Restmenge hat das gesuchte* h *Werte* $\leq \delta$ ρ*–fastüberall, während sie auf* Ω_δ *fastüberall Werte* $\geq \delta$ *hat.*

Für festes n *setzen wir zur Abkürzung* $\Omega_{k2^{-n}} = \Omega_k$ *und* $\Omega_k^{k+1} = \Omega_k - \Omega_{k+1}$ *für* $k = 1, 2, \ldots$*. Es gilt also*

$$\begin{aligned} \Omega_1 &= \Omega_1^2 + \Omega_2^3 + \Omega_3^4 + \cdots \\ \Omega_2 &= \qquad\;\; \Omega_2^3 + \Omega_3^4 + \cdots \\ &\ldots\ldots\ldots \end{aligned}$$

Aus der Summe der Indikatorfunktionen gewinnen wir die Funktion

$$h^{(n)} = \tfrac{1}{2^n} \cdot \sum_1^\infty 1_{\Omega_k} = \tfrac{1}{2^n} \cdot 1_{\Omega_1^2} + \tfrac{2}{2^n} \cdot 1_{\Omega_2^3} + \tfrac{3}{2^n} \cdot 1_{\Omega_3^4} + \cdots$$

Für $A \subseteq \Omega_k^{k+1}$ *gilt* $\quad \frac{k}{2^n}\rho(A) \leq \nu(A) \leq \frac{k+1}{2^n}\rho(A)$*.* $\qquad$ *Daher gilt für jedes* B

$$\nu(B) = \sum_1^\infty \nu(B \cap \Omega_k^{k+1}) \geq \sum \tfrac{k}{2^n}\rho(B \cap \Omega_k^{k+1}) = \int 1_B h^{(n)} \, d\rho.$$

Die Abschätzung ist bis auf den Fehler $\frac{1}{2^n}\rho(B)$ *genau, wenn* B *in einem* Ω_δ *mit* $\delta > \frac{1}{2^n}$ *enthalten ist; denn*

$$\nu(B) \le \int 1_B h^{(n)}\ d\rho + \frac{1}{2^n}\sum_1^\infty \rho(B\cap\Omega_k^{k+1}) = \int 1_B h^{(n)}\ d\rho + \frac{1}{2^n}\rho(B).$$

Wenn wir von n *nach* $n+1$ *gehen, dann vergrößert sie die Funktion* $h^{(n)}$ *an manchen Stellen um den Wert* $\frac{1}{2^{n+1}}$. *Im Limes* $n\to\infty$ *erhalten wir eine Funktion* h *mit*

$$s \le h(\omega) \le r \Leftrightarrow \omega \in \Omega_s - \Omega_r.$$

$h^{(n)}(\omega)$ *entsteht aus* $h(\omega)$ *durch Abrunden auf das nächste ganzzahlige Vielfache von* $\frac{1}{2^n}$. *Mit* $\Omega_s = \{h > s\}$ *(für die irrationalen s) haben wir* $h(\omega) = \lim \uparrow h^{(n)}(\omega) = \int_0^\infty 1_{\Omega_s}(\omega)\ ds$. *und somit nach dem Satz von der monotonen Konvergenz und dem Satz von Fubini*

$$\nu(B) = \int 1_B h\ d\rho = \int\int 1_B 1_{\Omega_s}\ d\rho\ ds = \int_0^\infty \rho(B\cap\Omega_s)\ ds.$$

Die Formel gilt auch, wenn B *in keinem* Ω_δ *enthalten ist. Wir brauchen nur* $\nu(\Omega) < \infty$. *Dass* h, *die Radon-Nikodym-Dichte* ρ*-fast überall eindetig bestimmt ist, haben wir bereits am Anfang des Abschnitts diskutiert.*

Eine Anwendung: In der elementaren mathematischen Statistik lernt man das Lemma von Neyman und Pearson. Es geht da um den Test einer einfachen Hypothese gegen eine einfache Alternative (auf dem 'Niveau' α.) Die Nullhypothese ist durch ein W-Maß μ beschrieben, die Alternative durch ein W-Maß ν. Gesucht ist ein 'Ablehnungsbereich' A mit

$$\mu(A) \le \alpha \qquad \nu(A) = \max!.$$

Der Fehler erster Art ist durch α limitiert; die 'Macht des Tests' soll möglichst groß sein. Der obige Satz sagt, wie ein optimaler Ablehnungsbereich A_α auszusehen hat. Wir formulieren und beweisen das Resultat hier für die übersichtlichste Situation.

Satz. *Es seien* μ *und* ν *Wahrscheinlichkeitsmaße mit* $\nu \ll \mu$, $d\nu = h\ d\mu$ *und es sei* $A = \Omega_r = \{h > r\}$. *Wenn* $\mu(\Omega_r) = \alpha$, *dann gilt*

$$\sup\big\{\nu(B):\ \mu(B)\le\alpha\big\} = \nu(\Omega_r).$$

Beweis. *Für eine beliebige Menge* B *gilt*

$$\big(\nu(B)-\nu(A)\big) - r\cdot\big(\mu(B)-\mu(A)\big) = \int\big(1_B - 1_A\big)(h-r)\ d\mu =$$
$$= \int\big(1_{B\setminus A} - 1_{A\setminus B}\big)\big((h-r)^+ - (h-r)^-\big)\ d\mu \le 0,$$

denn auf der Menge $B\backslash A$ *verschwindet die Funktion* $(h-r)^+$*, während auf* $A\backslash B$ *die Funktion* $(h-r)^-$ *verschwindet. Wenn* B *eine Menge ist mit* $\mu(B) \le \mu(A) = \alpha$*, dann haben wir* $\nu(B) - \nu(A) \le r\big(\mu(B) - \mu(A)\big) \le 0$. *Die Funktionen*

$$\alpha(r) = \mu\big(\{h > r\}\big) \qquad \textit{und} \qquad \beta(r) = \nu\big(\{h > r\}\big)$$

beschreiben also den Zusammenhang zwischen dem 'Niveau' und der 'Macht' des optimalen Tests der einfachen 'Nullhypothese' μ *gegen die einfache 'Alternative'* ν.

Als eine weitere 'Anwendung' des Satzes von Radon-Nikodym vollenden wir die Aufgabe, die Dualräume der Banachräume $L^p\big(\Omega, \mathfrak{A}, \rho\big)$ $(1 \le p < \infty)$ zu bestimmen. Wir haben (mit Hilfe der Hölder'schen Ungleichung) gesehen, dass jedes $h \in L^q\big(\Omega, \mathfrak{A}, \rho\big)$ eine stetige Linearform definiert. Wir müssen noch beweisen, dass jede stetige Linearform so zustande kommt. Dazu brauchen wir den Begriff des signierten Inhalts.

Definition 2.14. Ein reeller signierter Inhalt τ auf dem Mengenring $\mathfrak{R}$ ist eine reellwertige additive Mengenfunktion mit

$$\tau^+(R) = \sup\big\{\tau(R') :\ R' \subseteq R\big\} < \infty \qquad \text{für alle } R \in \mathfrak{R}$$

Eine komplexwertige Funktion auf $\mathfrak{R}$ heisst ein komplexer Inhalt, wenn Real- und Imaginärteil signierte Inhalte sind.

Satz. *Wenn* τ *auf* $\mathfrak{R}$ *ein signierter Inhalt ist, dann gilt für alle* $R \in \mathfrak{R}$

$$\tau^+(R) = \sup\big\{\sum (\tau(R_k))^+ :\ \sum R_k = R\big\} < \infty \qquad \textit{für jedes } R \in \mathfrak{R}$$

wobei das Supremum über alle finiten Partitionen von R *zu erstrecken ist.* τ^+ *und ebenso*

$$\tau^-(R) = \sup\big\{\sum \big(\tau(R_k)\big)^- :\ \sum R_k = R\big\}.$$

sind Inhalte, und es gilt $\tau = \tau^+ - \tau^-$, $\quad \tau^+ \wedge \tau^- = 0$. *Eine additive Mengenfunktion auf* $\mathfrak{R}$ *ist genau dann ein signierter Inhalt, wenn sie die Differenz zweier Inhalte ist.*

Ein signierter Inhalt τ *hat genau die Stetigkeitseigenschaft*

$$R_1 \supseteq R_2 \supseteq \cdots \qquad \bigcap R_n = \emptyset \quad \Longrightarrow \quad \lim \tau(R_n) = 0.$$

wenn $\tau^+ + \tau^-$ *ein Prämaß ist.*

Der Beweis ist ein Übungsaufgabe.

Satz (L^q ist der Dualraum von L^p).
Ist $\ell(\cdot)$ *eine beschränkte Linearform auf dem Banachraum* $L^p(\Omega, \mathfrak{A}, \rho\big)$ $(1 \le p < \infty)$*, so existiert ein* $g \in L^q$, $(1/p + 1/q = 1)$*, sodass*

$$\ell(f) = \int g \cdot f \, d\rho \qquad \textit{für alle } f \in L^p.$$

Beweis. *Wenn wir die Linearform* $\ell(\cdot)$ *auf die Menge der Indikatorfunktionen* 1_A *mit* $\rho(A) < \infty$ *einschränken, dann erhalten wir eine komplexwertige additive Mengenfunktion* τ *auf einem Mengenring. Sie ist ein komplexwertiger Inhalt, der bzgl.* ρ *totalstetig ist. Es existiert nämlich eine Konstante* M*, sodass*

$$|\tau(A)| = |\ell(1_A)| \leq M \cdot \|1_A\|_p = M \cdot (\rho(A))^{1/p}.$$

Für jede nach $\emptyset$ *absteigende Mengenfolge* A_n *streben die Normen* $\|1_{A_n}\|_p$ *(für alle* $p < \infty$*) nach* 0 *und daher auch die Werte* $\ell(1_{A_n})$*. Auf jeder Menge* $\tilde{A}$ *mit* $\rho(\tilde{A}) < \infty$ *existiert (nach dem Satz von Radon-Nikodym) eine* ρ*-fastüberall eindeutig bestimmte Funktion* $\tilde{g} = g \cdot 1_{\tilde{A}}$ *sodass für alle messbaren* $A \subseteq \tilde{A}$ *gilt* $\ell(A) = \int g \cdot 1_A \; d\rho$*. Die Gleichung setzt sich wegen der Linearität fort zu* $\ell(f) = \int g \cdot f \; d\rho$ *für alle* $f = \sum \alpha_k 1_{A_k}$*.*

Da $\ell(\cdot)$ *nach Voraussetzung in der* p*-Norm beschränkt ist, setzt sich die Gleichung auf den gesamten Banachraum fort, und* g *ist* q*-integrabel; denn*

$$\|g\|_q = \sup\Big\{\Big|\int g \cdot f \; d\rho\Big| : \|f\|_p \leq 1\Big\} = \sup\Big\{|\ell(f)| : \|f\|_p \leq 1\Big\}.$$

Für jedes $p \in [1, \infty)$ *ist die* q*-Norm wirklich die zur* $p - \mathrm{Norm}$ *duale Norm*

2.5.1 Stochastische Konvergenz und gleichgradige Integrierbarkeit

Wenn eine Folge $(f_n)_n$ in $L^p(\Omega, \mathfrak{A}, \rho)$ konvergiert, $\|f_n - f\|_p \to 0$, dann existiert eine Teilfolge, entlang welcher die Folge ρ-fastüberall konvergiert, $\left|f_{n_k} - f\right|(\cdot) \to 0$ fastüberall. Wir haben das oben bewiesen, mit dem Ziel, die Vollständigkeit der Räume L^p zeigen.

Die fastsichere Konvergenz ist weder notwendig noch hinreichend für die Normkonvergenz; und die im Satz von der majorisierten Konvergenz geforderte Bedingung ist nicht notwendig. — Wir wollen die Verhältnisse genauer studieren.

Satz 2.5.5. *Ist* $(f_n)_n$ *eine Cauchy-Folge im* $L^p(\Omega, \mathfrak{A}, \rho)$*, so gilt*

$$\forall\ \alpha,\ \beta > 0\ \exists\ N\ \forall\ m, n \geq N \quad \rho(\{|f_n - f_m| \geq \alpha\}) < \beta.$$

Beweis. *Es gilt*

$$\alpha^p \cdot \rho\{|f - g| \geq \alpha\} \leq \int 1_{\{|f-g|\geq\alpha\}} \cdot \left|f - g\right|^p \, d\rho \leq \int \left|f - g\right|^p \, d\rho = \|f - g\|_p^p$$

Wir wählen N *so groß, dass für alle* $m, n \geq N$ *gilt* $\frac{1}{\alpha}\|f_n - f_m\|_p < \beta^{1/p}$.

Eine in der p-Norm konvergente Folge ist also stochastisch konvergent im Sinn der

Definition 2.15 (Stochasische Konvergenz)**.**
Man sagt von einer Folge messbarer Abbildungen eines σ-endlichen Maßraums in einen metrischen Raum

$$\big(\Omega, \mathfrak{A}, \rho\big) \xrightarrow{f_n} \big(S, d(\cdot,\ \cdot)\big),$$

sie sei eine stochastische Cauchy-Folge oder eine Cauchy-Folge dem Maße nach, wenn für jedes $A \in \mathfrak{A}$ mit $\rho(A) < \infty$ gilt

$$\forall \alpha, \beta > 0\ \exists N\ \forall m, n \geq N \quad \rho\big(A \cap \{d(f_n,\ f_m) \geq \alpha\}\big) < \beta.$$

Man sagt, dass sie stochastisch (oder dem Maße nach) gegen f konvergiert, wenn für alle A mit $\rho(A) < \infty$ gilt

$$\forall \alpha, \beta > 0\ \exists N\ \forall n \geq N \quad \rho\big(A \cap \{d(f_n,\ f) \geq \alpha\}\big) < \beta.$$

Satz. *Wenn eine Folge messbarer Abbildungen* (f_n) *fastüberall eine Cauchy-Folge ist, dann ist sie eine stochastische Cauchy-Folge; wenn sie fast überall gegen* f *konvergiert, dann konvergiert sie auch stochastisch gegen* f.

Beweis. *Der Beweis ist sofort an der folgenden Formel abzulesen:* $(f_n)_n$ *ist genau dann f.ü.- Cauchy-Folge, wenn für alle* A *mit* $\rho(A) < \infty$ *gilt*

$$\forall \alpha, \beta > 0\ \exists N \quad \rho\big(A \cap \bigcup_{m,n \geq N} \{d(f_n,\ f_m) \geq \alpha\}\big) < \beta.$$

Bei den gegen ein f *konvergenten Folgen gilt das Entsprechende.*

Satz 2.5.6 (Fastüberall konvergente Teilfolgen).
Jede stochastisch konvergente Cauchy-Folge mit Werten in einem vollständigen metrischen Raum besitzt eine Teilfolge, die fastüberall konvergiert.

Beweis. *Wir betrachten die Situation zuerst auf einem endlichen Maßraum. Es seien $(\alpha_k), (\beta_k)$ summable Folgen positiver Zahlen, und N_k so, dass*

$$\forall m, n \geq N_k \quad \rho\big(\{d(f_n,\ f_m) \geq \alpha_k\}\big) < \beta_k.$$

Wir wählen die Folge N_k strikt wachsend und beweisen, dass die Teilfolge $(g_k) = (f_{N_k})$ fastüberall konvergiert.

Vorab eine einfache Bemerkung über Cauchy-Folgen in einem vollständigen metrischen Raum: Ist $(P_k)_k$ eine Folge in einem metrischen Raum mit $d(P_k, P_{k+1}) < \alpha_k$, $\sum \alpha_k < \infty$, so ist sie eine Cauchy-Folge. (Man könnte sie eine schell konvergente Folge nennen.)

Ist $(g_k)_k$ eine Folge messbarer Abbildungen in einen vollständigen metrischen Raum mit

$$\rho\big(\{d(g_k, g_{k+1}) \geq \alpha_k\}\big) < \beta_k; \quad \sum \alpha_k < \infty, \quad \sum \beta_k < \infty,$$

so ist ist sie fastüberall konvergent. Zum Beweis bemerken wir, dass $(g_k(\omega))_k$ konvergiert für alle ω in

$$\Omega_K = \big\{\omega :\ \forall k \geq K \quad d(g_k(\omega),\ g_{k+1}(\omega)) < \alpha_k\big\}.$$

Die Komplemente der Ω_K bilden eine absteigende Folge; ihr Durchschnitt umfasst die Menge der ω, in welchen die Folge $g_k(\omega)$ nicht konvergiert. Diese ist eine Nullmenge, weil die Folge $\rho\Big(\complement\Omega_K\Big) < \sum_{k\geq K} \beta_k$ nach 0 fällt.

Sei nun $(\varphi_n)_n$ eine stochastisch konvergente Folge auf einem σ-endlichen Maßraum $\big(\Omega, \mathfrak{A}, \rho\big)$ und $\Omega = \sum A_k$ mit $\rho(A_k) < \infty$ für alle k. Wir wählen eine Teilfolge, die fastüberall auf A_1 konvergiert, davon eine Teilfolge , die auf A_2 konvergiert, ... Die Diagonalfolge konvergiert fastüberall auf Ω.

Beispiel 2.5.1. Wir präsentieren eine Folge messbarer Funktionen $(f_n)_n$ auf dem Einheitsintervall $[0,\ 1]$ mit dem Lebesgue-Maß, welche stochastisch aber nicht fastüberall gegen die Nullfunktion konvergiert: Für $n = 2^N + k \quad (N \in \mathbb{N}, k = 1, 2, \ldots, 2^N)$ sei f_n die Indikatorfunktion des Intervalls $I_n = \big(\frac{k-1}{2^N},\ \frac{k}{2^N}\big]$. Für jede Zahlenfolge, die langsamer als n anwächst, $\alpha_n = o(n)$, konvergiert $\alpha_n \cdot f_n$ stochastisch und in der 1-Norm gegen die Nullfunktion. Für Folgen α_n mit $\liminf \alpha_n > 0$ konvergiert sie nicht fastüberall nach 0.

Satz 2.5.7 (Teilfolgen-Charakterisierung).
Eine Folge messbarer Abbildungen (f_n) auf einem W-Raum ist genau dann stochastisch konvergent, wenn jede Teilfolge eine fastüberall konvergente Teilfolge besitzt mit ein und demselben Limes. Dieser Limes ist f auch der stochastische Limes.

Der Beweis liegt auf der Hand. Diese Charakterisierung der stochastischen Konvergenz liefert manchmal bequeme Beweise für Sätze, die man schon für die fastsichere Konvergenz kennt.

Satz. *Ist* $(f_n)_n$ *eine stochastisch konvergente Folge mit Werten im metrischen Raum* $\big(S, d(\cdot,\cdot)\big)$, $f_n \to f$*; und ist* ψ *eine stetige Abbildung* $\big(S, d(\cdot,\cdot)\big) \longrightarrow \big(T, e(\cdot,\cdot)\big)$*, so konvergiert die Folge* $g_n = \psi(f_n)$ *stochastisch gegen* $g = \psi(f)$.

Konstruktion: Die stochastische Konvergenz kann (im Gegensatz zur fastsicheren Konvergenz) durch eine Metrik beschrieben werden:

Wenn $\big(\Omega, \mathfrak{A}, \rho\big)$ ein W-Raum ist, dann kann man z. B. folgendermaßen vorgehen:

Es seien f und g Abbildungen in einem metrischen Raum $\big(S, d(\cdot\,\cdot)\big)$. Man sagt, g liegt in der (α, β)-Umgebung von f und notiert

$$g \in U_{\alpha,\beta}(f) \quad \Longleftrightarrow \quad \rho\big(\{d(f,g) \geq \alpha\}\big) < \beta$$

Offenbar gilt $\quad g \in U_{\alpha,\beta}(f) \quad \Longleftrightarrow \quad f \in U_{\alpha,\beta}(g)\quad$ und ausserdem

$$g \in U_{\alpha',\beta'}(f), \quad h \in U_{\alpha'',\beta''}(g) \quad \Longrightarrow \quad h \in U_{\alpha'+\alpha'',\beta'+\beta''}(f).$$

Man definiert dann $\quad d_{\text{stoch}}(f,g) = \inf\{\delta :\ g \in U_{\delta,\delta}(f)\}$. Dies ist eine Metrik, welche die Topologie der stochastischen Konvergenz auf einem endlichen Maßraum erzeugt. Im Falle nichtendlicher Maßraume muss man etwas behutsamer umgehen.

Bevor wir uns wieder den L^1-Cauchy-Folgen zuwenden bemerken wir noch einen Satz über Bildmaße

Satz 2.5.8. *Auf dem W-Raum* $(\Omega, \mathfrak{A}, \rho)$ *sei* $(\varphi_n)_n$ *eine Folge von Abbildungen in den vollständigen metrischen Raum, die stochastisch konvergiert:* $\varphi_n \to \varphi$*. Es bezeichne* ν_n *das Bildmaß unter* φ_n *und* ν *das Bildmaß unter* φ*. Für jede beschränkte gleichmäßig stetige Funktion* h *gilt dann*

$$\int h(P)\ d\nu_n \longrightarrow \int h(P)\ d\nu. \qquad \text{(`Schwache Konvergenz')}$$

Beweis. *Wir wählen* δ *so klein, dass* $\quad d(P,Q) < \delta \quad \Rightarrow \quad |h(P) - h(Q)| < \varepsilon/2.$
Und wir wählen N *so groß, dass* $\quad n \geq N \quad \Rightarrow \rho\big(\{d(\varphi_n(\omega), \varphi(\omega)) \geq \delta\}\big) < \frac{\varepsilon}{4M}$*, wo* $M = \sup|h(P)|$*. Es gilt dann für* $n \geq N$

$$\int \big|h(\varphi_n(\omega)) - h(\varphi(\omega))\big|\ d\rho \leq 2M \cdot \rho\big(\{d(\varphi_n(\omega), \varphi(\omega)) \geq \delta\}\big) + \varepsilon/2 \cdot \rho(\Omega) < \varepsilon.$$

Die folgende 'Anwendung' des Teilfolgensatzes bringt uns zurück zum Raum $L^1(\Omega, \mathfrak{A}, \rho)$. Es handelt sich um eine Erweiterung des Satzes von der majorisierten Konvergenz, den wir oben (auf der Grundlage des Lemma von Fatou) für dominierte fastüberall konvergente Folgen bewiesen haben.

Satz (Majorisierte Konvergenz).
Es sei $(f_n)_n$ *eine Folge numerischer Funktionen auf einem* σ*-endlichen Maßraum* $\big(\Omega, \mathfrak{A}, \rho\big)$*, welche eine stochastische Cauchy-Folge ist. Wenn eine integrierbare Funktion* g *existiert,*

sodass $|f_n| \leq g$ *für alle* n*, dann ist der stochastische Grenzwert eine integrable Funktion* f*, und es gilt*

$$\int \big|f_n - f\big| \; d\rho \to 0, \qquad \textit{und für alle } A \in \mathfrak{A} \quad \int_A f_n \; d\rho \to \int_A f \; d\rho.$$

Diesen Satz werden wir nun mit Hilfe des Begriffs der gleichgradigen Integrierbarkeit verallgemeinern. Die konvergenten Folgen im Banachraum $L^1\big(\Omega, \mathfrak{A}, \rho\big)$ haben über die stochastische Konvergenz hinaus noch weitere bemerkenswerte Eigenschaften. Sie sind, wie wir sehen werden, gleichgradig integrabel im Sinne der

Definition 2.16 (Gleichgradige Integrabilität)**.**
Eine Familie numerischer Funktionen auf einem Maßraum $\big\{f_i : i \in I\big\}$ heisst gleichgradig integrabel, wenn zu jedem $\varepsilon > 0$ eine integrable Funktion $h \geq 0$ existiert, sodass

$$\int_{\{|f_i| \geq h\}} |f_i| \; d\rho < \varepsilon \qquad \text{für alle } i \in I.$$

Bemerkungen:

1. Wenn eine Familie von einer Funktion h majorisiert wird, dann ist sie offensichtlich gleichgradig integrabel. Wir werden sehen, dass gleichgradige Integrabilität ähnlich gute Dienste leistet, wenn es darum geht, Limesbildung und Integration zu vertauschen.

2. Die gleichgradige Integrierbarkeit ist ein Eigenschaft, die man auch als eine Eigenschaft der Familie der endlichen Maße $\big\{d\nu_i = |f_i| \; d\rho : i \in I\big\}$ verstehen kann. In dieser Betrachtung ist sie äquivalent mit der Bedingung der gleichgradigen Totalstetigkeit im Sinne der folgenden Definition.

Definition 2.17 (Gleichgradige Totalstetigkeit)**.**
Es sei $\big\{\nu_i : i \in I\big\}$ eine Familie von endlichen Maßen auf einem σ-endlichen Maßraum $\big(\Omega, \mathfrak{A}_\sigma, \rho\big)$. Man sagt, dass die Familie gleichgradig totalstetig ist, wenn gilt

i) $\sup_i\{\nu_i(\Omega)\} < \infty$,

ii) $\forall \varepsilon > 0 \; \exists \mu$ W- Maß und $\delta > 0 \quad \forall A \in \mathfrak{A}_\sigma \qquad \mu(A) < \delta \Rightarrow \big(\forall i \in I \quad \nu_i(A) < \varepsilon\big)$

Eine Familie von beschränkten Prämaßen $\big\{\rho_i : i \in I\big\}$ auf einer Mengenalgebra $\mathfrak{A}$, welche die σ-Algebra $\mathfrak{A}_\sigma$ erzeugt, heisst gleichgradig totalstetig, wenn die Eigenschaften auf der erzeugenden Mengenalgebra erfüllt sind. Die Abschätzungen übertragen sich nämlich auf die Fortsetzungen.

Eine Familie von signierten Prämaßen $\big\{\rho_i : i \in I\big\}$ heisst gleichgradig totalstetig, wenn die Familien $\big\{\rho_i^+ : i \in I\big\}$ und $\big\{\rho_i^- : i \in I\big\}$ gleichgradig totalstetig sind.

Satz 2.5.9. *Die Familie* $\{d\nu_i = f_i \, d\rho\}$ *mit beschränkten* $\int |f_i| \, d\rho$ *ist genau dann gleichgradig totalstetig, wenn die Familie* $\{f_i : i \in I\}$ *gleichgradig integrierbar ist.*

Beweis. *Die auftretenden signierten Maße bzw. Prämaße denken wir uns stets in Positiv- und Negativteil zerlegt. Es genügt, den Satz für Familie nichtnegativer Funktionen zu beweisen.*

Es sei $\{f_i : i \in I\}$ *eine Familie nichtnegativer Funktionen auf dem* σ*-endlichen Maßraum* $(\Omega, \mathfrak{A}_\sigma, \rho)$*; und es sei* $\varepsilon > 0$*. Wir nennen (für diesen Beweis) eine integrable Funktion* h *eine* ε*-stützende Funktion, wenn gilt*

$$\int_{\{f_i \geq h\}} f_i \, d\rho < \varepsilon \qquad \textit{für alle } i \in I.$$

(In dieser Sprache ist also eine Familie genau dann gleichgradig integrabel, wenn zu jedem $\varepsilon > 0$ *eine* ε*-stützende Funktion existiert.)*

Es sei also h *eine* ε*-stützende Funktion für die gegebene Familie. Die Familie der Gesamtgewichte ist dann beschränkt,* $\sup\{\int f_i \, d\rho : i \in I\} < \infty$ *; denn*

$$\int f_i \, d\rho \leq \int_{\{f_i \geq h\}} f_i \, d\rho + \int h \, d\rho \leq \int h \, d\rho + \varepsilon.$$

Sei $y = \int h \, d\rho$ *und* μ *das W-Maß* $d\mu = \frac{1}{y} h \, d\rho$*. Mit* $\Omega_i = \{f_i \geq h\}$ *haben für alle* $A \in \mathfrak{A}_\sigma$

$$\nu_i(A) = \nu_i(A \cap \Omega_i) + \nu_i(A \setminus \Omega_i) \leq \varepsilon + \int_{A \setminus \Omega_i} f_i \, d\rho \leq \varepsilon + \int_A h \, d\rho = \varepsilon + y \cdot \mu(A).$$

Für alle A *mit* $\mu(A) \leq \delta = \frac{1}{y} \cdot \varepsilon$ *gilt* $\nu_i(A) < 2\varepsilon$ *für alle* $i \in I$*. Das W-Maß* μ *ist also ein* 2ε*-stützendes W-Maß im Sinne der*

Sprechweise: Es sei $\{\nu_i : i \in I\}$ *eine Familie von Maßen mit* $\sup_i\{\nu_i(\Omega)\} < \infty$*; und es sei* $\varepsilon > 0$*. Wir nennen ein W-Maß* μ *ein* ε*-stützendes W-Maß, wenn* $\delta > 0$ *existiert, sodass* $\mu(A) < \delta \Rightarrow \forall i \, \nu_i(A) < \varepsilon$*.*
(In dieser Sprache ist also eine Familie $\{\nu_i : i \in I\}$ *genau dann gleichgradig totalstetig, wenn* $\sup_i\{\nu_i(\Omega)\} < \infty$ *und zu jedem* $\varepsilon > 0$ *ein* ε*-stützendes W-Maß existiert.)*

Es sei also μ *ein* ε*-stützendes W-Maß zu* δ *für die Familie der* $\nu_i = f_i \, d\rho$*, wo* ρ *ein* σ*-endliches Maß ist. O.B. d. A können wir* $\mu \ll \rho$ *annehmen,* $d\mu = \tilde{h} \, d\rho$*. Wir zeigen nun, dass für genügend großes* $y \in \mathbb{R}_+$ *die Funktion* $h = y \cdot \tilde{h}$ *eine* ε*-stützende Funktion für die Familie* $\{f_i : i \in I\}$ *ist. Für die Mengen* $\Omega_i = \{f_i \geq y \cdot \tilde{h}\}$ *haben wir in der Tat*

$$y \cdot \mu(\Omega_i) = \int_{\Omega_i} y \cdot \tilde{h} \, d\rho \leq \int_{\Omega_i} f_i \, d\rho \leq \nu(\Omega_i) \leq M < \infty.$$

Für $\frac{1}{y} M = \delta$ *ergibt das* $\mu(\Omega_i) < \delta, \quad \forall i \, \nu_i(\Omega_i) < \varepsilon$*, und somit* $\int_{\{f_i \geq y \cdot \tilde{h}\}} f_i \, d\rho < \varepsilon$.

Beispiel 2.5.2. Es sei Ω eine abzählbare Menge und $I = \Omega$. Jedem i sei ein Maß ν_i zugeordnet, welches auf einen Punkt konzentriert ist, und zwar auf den Punkt i selber: p_i sei das Gewicht. ν_i ist also totalstetig bzgl. des σ-endlichen Zählmaßes mit der 'Dichte' f_i, die im Punkt i den Wert p_i hat und sonst überall den Wert 0. Die Familie $\{\nu_i : i \in I\}$ ist genau dann gleichgradig absolutstetig, wenn die Zahlen p_i beschränkt sind und für jedes $\varepsilon > 0$ die Menge $\Omega_\varepsilon = \{\omega : p(\omega) > \varepsilon\}$ endlich ist. Die Notwendigkeit der Bedingung $|\Omega_\varepsilon| < \infty$ zeigen wir später mit dem Argument, dass es andernfalls eine Folge $\{f_{i_n}\}$ gibt, die stochastisch gegen die Nullfunktion strebt, während die Folge der Integrale $\int |f_{i_n}|$ nicht nach 0 strebt.

Hier beweisen wir, dass die angegebenen Bedingungen hinreichend sind. Ein ε-stützendes Wahrscheinlichkeitsmaß μ_ε ist z. B. die uniforme Verteilung auf Ω_ε. Jeder Punkt in Ω_ε bekommt hier das Gewicht $\frac{1}{|\Omega_\varepsilon|}$, die übrigen das Gewicht 0. Eine Menge A mit $\mu_\varepsilon(A) < \delta = \frac{1}{|\Omega_\varepsilon|}$ enthält keinen Punkt von Ω_ε; für Mengen A, die disjunkt zu Ω_ε sind, gilt $\nu_i(A) < \varepsilon$ für alle i.

Bei Familien von Maßen auf einem endlichen Maßraum ist die Beschreibung der gleichgradigen Totalstetigkeit etwas übersichtlicher.

Satz. *Eine Familie $\{\nu_i : i \in I\}$ auf einem W-Raum $(\Omega, \mathfrak{A}_\sigma, \mu)$ ist genau dann gleichgradig totalstetig, wenn gilt*

$$\sup_i\{\nu_i(\Omega)\} < \infty \qquad \textit{und} \qquad \forall\, \varepsilon > 0\, \exists \delta > 0\, \forall A \in \mathfrak{A}_\sigma \quad \mu(A) < \delta \Rightarrow \forall i\, \nu_i(A) < \varepsilon.$$

Eine Familie integrabler Funktionen $\{f_i : i \in I\}$ auf einem endlichen Maßraum ist genau dann gleichgradig integrierbar, wenn gilt

$$\forall\, \varepsilon > 0\, \exists M\, \forall i \quad \int_{\{|f_i| \geq M\}} |f_i|\, d\mu < \varepsilon.$$

Beweis. *Wir müssen zeigen, dass man für eine gleichgradige Familie für alle ε das endliche Grundmaß als ein ε-stützendes Maß verwenden kann. Wir wiederholen das Argument im Beweis des Lemma von Neyman und Pearson. Dort haben wir gezeigt:*
Wenn $\nu \ll \mu$ und $\mu(\frac{d\nu}{d\mu} \geq y) = \alpha, \quad \nu(\frac{d\nu}{d\mu} \geq y) = \beta$, dann gilt $\mu(B) \leq \alpha \Rightarrow \nu(B) \leq \beta$.
Es sei nun μ_ε ein ε-stützendes W-Maß mit einem gewissen $\delta < 0$, o. B. d. A totalstetig bzgl. des endlichen Grundmaßes $d\mu_\varepsilon = h_\varepsilon\, d\mu$. Wir wählen M so groß, dass

$$\Big(\mu_\varepsilon \backslash M\ \mu\Big)(\Omega) < \delta/2 \qquad \text{und} \qquad M \cdot \mu_\varepsilon(\{|h_\varepsilon| \geq M\}) < \delta/2 \qquad \textit{sodass also}$$

$$\mu_\varepsilon(\{|h_\varepsilon| \geq M\}) = M \cdot \mu(\{|h_\varepsilon| \geq M\}) + \Big(\mu_\varepsilon \backslash M\ \mu\Big)(\Omega) < \delta \quad \textit{und somit}$$

$$\mu(A) < \frac{1}{M}\delta/2 \quad \Rightarrow \quad \mu_\varepsilon(A) < \delta \quad \Rightarrow \quad \forall i\, \nu_i(A) < \varepsilon.$$

Stochastische Konvergenz, gleichgradige Integrierbarkeit und L^1-Konvergenz

Satz 2.5.10. *Wenn eine Folge in $L^1(\Omega, \mathfrak{A}, \rho)$ konvergiert, dann konvergiert sie stochastisch und sie ist gleichgradig integrierbar. Die beiden Eigenschaften sind auch hinreichend für die Norm-Konvergenz.*

Beweis. *Die stochastische Konvergenz haben wir in verschärfter Form bewiesen; wir mussten uns nicht auf Mengen von endlichem Maß einschränken. Es geht darum, die gleichgradige Integrierbarkeit zu beweisen. Die Folge der Integrale ist beschränkt. Es genügt also nachzuweisen, dass es zu jedem $\varepsilon > 0$ eine integrierbare Funktion h und ein $\delta > 0$ gibt mit*

$$\forall A \in \mathfrak{A} \quad \int_A h \; d\rho < \delta \quad \Longrightarrow \quad \forall n \int_A |f_n| \; d\rho < \varepsilon.$$

Wir wählen N so groß, dass $\|f_n - f\|_1 \leq \varepsilon/2$ für alle $n \geq N$.
Die Funktion $h = |f_1| \vee |f_2| \vee \cdots \vee |f_N| \vee |f|$ leistet das Verlangte mit $\delta = \varepsilon/2$; denn

$$\int_A |f_n| \; d\rho \leq \|f_n - f\|_1 + \int_A |f| \; d\rho \leq \quad \varepsilon/2 + \int_A h \; d\rho.$$

Die zweite Behauptung formulieren und beweisen wir in einem separaten Satz.

Satz 2.5.11. *Die Folge integrabler Funktionen $(f_n)_n$ konvergiere stochastisch gegen f. Genau dann, wenn sie gleichgradig integrierbar ist, konvergiert sie auch in der L^1-Norm.*

Beweis. *Die eine Richtung haben wir eben bewiesen. Wir betrachten also eine stochastisch konvergente Folge, die gleichgradig integrierbar ist.Als Vorbereitung bemerken wir:*
Wenn eine Familie $\{\nu_i : i \in I\}$ gleichgradig totalstetig ist, dann auch ihre konvexe Hülle.
Wenn nämlich $\nu = \sum \lambda_i \nu_i$ mit $\lambda_i \geq 0$, $\sum \lambda_i = 1$, dann gilt $\forall A \; \nu(A) \leq \sup \nu_i(A)$.
Wenn $\{f_n : n \in \mathbb{N}\}$ gleichgradig integrabel ist, dann auch die Familie aller $f_{mn} = f_n - f_m$.
Zu $\varepsilon > 0$ wählen wir ein integrables h so, dass

$$\int_{\{|f_{mn}| \geq h\}} |f_{mn}| \; d\rho < \varepsilon \quad \textit{für alle } m, n.$$

Wir werden ein N finden, sodass für $m, n \geq N$ gilt $\int_{\{|f_{mn}|<h\}} |f_{mn}| \; d\rho < 2\varepsilon$.
Zuerst behandeln wir den Teil des Integrationsbereichs, in welchem h sehr klein ist. Wir wählen $\eta > 0$ so klein, dass $\int_{\{h<\eta\}} h \; d\rho < \varepsilon$. Wir haben dann auch

$$\int_{\{|f_{mn}|<h\} \cap \{h<\eta\}} |f_{mn}| \; d\rho < \; \varepsilon \quad \textit{für alle } m, n$$

Auf dem restlichen Integrationsbereich $\{|f_{mn}| < h\} \cap \{h \geq \eta\}$ ist die bzgl. des endlichen Maßes $h \; d\rho$ integrable Konstante 1 eine Majorante der Funktionen $\frac{1}{h} f_{mn}$. Da die Doppelfolge stochastisch gegen 0 konvergiert. liefert der Satz von der majorisierten Konvergenz bei genügend großem N für alle $m, n \geq N$

$$\int_{\{|f_{mn}|<h\} \cap \{h \geq \eta\}} \left(\frac{1}{h} \cdot |f_{mn}|\right) \cdot h \; d\rho < \; \varepsilon.$$

Für $m, n \geq N$ *gilt also* $\int \left| f_n - f_m \right| \, d\rho < \varepsilon$. *Somit ist gezeigt, dass die Folge* $(f_n)_n$ *eine Cauchy-Folge im Raum* L^1 *ist. Sie konvergiert also auch im Sinne der Norm gegen die Grenzfunktion.*

Satz. *Die Folge nichtnegativer integrabler Funktionen* f_n *konvergiere stochastisch gegen die integrable Funktion* f. *Die Folge konvergiert genau dann auch in der Norm, wenn* $\int (f_n - f) \, d\rho \to 0$.

Beweis. *Nach dem Lemma von Fatou wissen wir* $\liminf \int (f_n - f) \, d\rho \geq 0$. *Wenn wir nämlich eine Menge endlichen Maßes* A *wählen und zu einer beliebig vorgegebenen Teilfolge eine Teilfolge fastsicherer Konvergenz, dann liefert uns das entlang dieser Teilfolge die Abschätzung* $\int_A f \leq \liminf \int_A f_{n_k}$.

Die Folge $(f_n \wedge f)_n$ *ist durch die integrable Funktion* f *majorisiert und daher gleichgradig integrierbar. Nach dem eben bewiesenen Satz konvergiert sie in der Norm gegen* f. $\int \left| f - f \wedge f_n \right| = \int f \backslash f_n \to 0$.

Auf der anderen Seite haben wir $\int \left(f_n - f \right) = \int \left(f_n \backslash f - f \backslash f_n \right)$. *Die Folge dieser Integrale konvergiert genau dann nach* 0, *wenn* $\int f_n \backslash f \to 0$, *also genau dann wenn* $\|f - f_n\|_1 = \int \left| f - f_n \right| = \int f \backslash f_n + \int f_n \backslash f \to 0$.

Wir fassen die Resultate zusammen:

Satz 2.5.12. *Die Folge integrabler Funktionen* $(f_n)_n$ *konvergiere stochastisch gegen* f. *Die folgenden zusätzlichen Eigenschaften der Folge sind dann äquivalent*

1. *Die Folge konvergiert in der Norm,*

2. *Die Folge* $(|f_n|)_n$ *ist gleichgradig integrierbar,*

3. *Es gilt* $\lim \int |f_n| \, d\rho = \int |f| \, d\rho$.

Die Totalvariation als Norm im Vektorraum der signierten Maße.
Es sei $\left(\Omega, \mathfrak{A}_\sigma\right)$ ein messbarer Raum und $\mathfrak{R}$ ein abzählbarer Mengenring, der $\mathfrak{A}_\sigma$ erzeugt. Ausserdem fordern wir, dass eine abzählbare Überdeckung des Gesamtraums durch $\mathfrak{R}$-Mengen existiert; $\Omega = \sum R_j$. Die Menge aller beschränkten signierten Prämaße τ ist dann ein normierter Vektorverband $\left(V, \|\cdot\|\right)$ bzgl. der sog. Totalvariationsnorm:

$$\|\tau\|_{TV} = \|\tau^+\| + \|\tau^-\| = \quad \sup\{\tau(R) : \; R \in \mathfrak{R}\} + \sup\{-\tau(R) : \; R \in \mathfrak{R}\}.$$

Wenn ρ irgendein Maß ist, bezüglich dessen τ^+ und τ^- totalstetig sind, dann existiert eine integrable Funktion h, sodass $\tau(R) = \int_R h \, d\rho$ für alle $R \in \mathfrak{R}$. Und es gilt $\|\tau\|_{TV} = \int |h| \, d\rho$. Der normierte Vektorverband $\left(V, \|\cdot\|\right)$ ist vollständig. Es handelt sich um einen Banachverband.

Wenn eine Folge τ_n in $\left(V_+, \|\cdot\|_{TV}\right)$ gegen τ konvergiert; dann gilt trivialerweise $\forall R \quad \tau_n(R) \to \tau(R)$. Andererseits: Wenn eine Folge $(\tau_n)_n$ in V_+ 'punktweise' konvergiert:

$\tau_n(R) \to \alpha(R)$ für alle $R \in \mathfrak{R}$, dann ist der Limes $\alpha(\cdot)$ ein Inhalt, aber nicht notwendigerweise ein Prämaß. Den obigen Satz können wir dahingehend interpretieren, dass $\alpha(\cdot)$ genau dann ein Prämaß ist, wenn die τ_n gleichgradig totalstetig sind.

Da wir die Abzählbarkeit von $\mathfrak{R}$ vorausgesetzt haben, existiert zu jeder in der Norm beschränkten Folge τ_n eine Teilfolge, welch 'punktweise' konvergiert.

Für eine Familie $\{\tau_i : i \in I\}$ ist die gleichgradige Totalstetigkeit eine notwendige und hinreichende Bedingung für die bedingte Kompaktheit.

2.5.2 W-Maße auf $\mathbb{R}/2\pi$ und auf $\mathbb{R}^p$; Faltung, Schwache Konvergenz.

Notation. Bisher haben wir die Wahrscheinlichkeitsmaße meistens mit Buchstaben wie $\mu, \nu, \ldots$ bezeichnet. Die Bezeichnungen $d\mu, d\nu, \ldots$ sind ebenfalls guter Brauch; die Bezeichnung der unbestimmten Integrale mit $h\ d\mu$, $k\ d\nu, \ldots.$ bedarf keiner Erläuterung. Wenn ein Wahrscheinlichkeits ν totalstetig ist bezüglich des Lebesgue-Maßes (auf einem n-dimensionalen reell-affinen Raum), dann notiert man gern $d\nu(x) = p(x)\ dx$ mit $p(\cdot) \geq 0$, $\int p(x)\ dx = 1$. In Analogie zu dieser Notation schreiben wir auch sonst $d\nu(\omega) = h(\omega)\ d\mu(\omega)$ oder $= h(\omega)\ \mu(d\omega)$, mit $h(\cdot) \geq 0$, $\int h(\omega)\ d\mu(\omega) = 1$.
Das Integral der Funktion f bzgl. $d\nu = h\ d\mu$ erscheint demgemäß in den Formen

$$\int f\ d\nu = \int fh\ d\mu = \int (fh)(\omega)\ d\mu(\omega) = \int (fh)(\omega)\ \mu(d\omega).$$

Definition 2.18 (Charakteristische Funktion)**.** In der Stochastik wird jedem Wahrscheinlichkeitsmaß auf einem reellen Vektorraum seine charakterische Funktion zugeordnet.

$$d\mu \longmapsto \varphi(\cdot) = \int e^{i\langle \cdot\,, x\rangle} d\mu(x) \quad \text{auf dem Dualraum.}$$

(Wir denken vor allem an den Raum $\mathbb{R}^p_{Sp}$ der reellen p-Spalten x, wo dann die Linearformen $\langle t\,,\cdot\,\rangle$ durch die die p-Zeilen t gegeben sind. Der Dualraum wird mit $\mathbb{R}^p_Z$ bezeichnet.) Die Dimension wird hier p genannt und nicht n, weil der Buchstabe n immer wieder benötigt wird, wenn irgendwelche Folgen ins Spiel kommen.

Beispiel.

1. Es sei μ die uniforme Verteilung auf dem Intervall $[-\frac{1}{2}, \frac{1}{2}]$, also $d\mu(x) = 1_{[-\frac{1}{2},\frac{1}{2}]}(x)\ dx$. Die charakteristische Funktion ist
$$\varphi(t) = \int_{-1/2}^{1/2} e^{itx}\ dx = \frac{1}{it}\big[e^{itx}\big]_{-1/2}^{1/2} = \frac{2}{t}\sin(t/2).$$

2. Wenn man μ mit sich selbst faltet, dann erhält man die Dreiecksdichte über $[-1, 1]$: $\big(\mu * \mu\big)(dx) = (1-|x|)^+\ dx$. Die charakteristische Funktion ist
$$\int e^{itx}(1-|x|)^+\ dx = 2\cdot\int_0^1 (1-x)\cos(tx)\ dx = \frac{2}{t^2}(1-\cos t) = \frac{4}{t^2}\left(\sin\frac{t}{2}\right)^2.$$

3. Die charakteristische Funktion der Standardnormalverteilung haben wir oben mit der Methode der Verlagerung des Integrationswegs berechnet
$$\int e^{itx}\tfrac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}\ dx = e^{-\frac{1}{2}t^2}\cdot\int \tfrac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}(x-it)^2}\ dx = e^{-\frac{1}{2}t^2}.$$

4. Es gibt noch einige weitere mit elementaren Funktionen darstellbare Dichten, deren charakteristische Funktion ebenfalls durch elementare Funktionen dargestellt werden. Wir haben oben die Rechnungen durchgeführt für die Cauchy-Verteilung und die ‘doppeltexponentielle’ Verteilung
$$\int e^{itx}\,\tfrac{1}{\pi}\,\frac{1}{1+x^2}\,dx = e^{-|t|};\qquad\qquad \int e^{itx}\,\tfrac{1}{2}\,e^{-|x|}\,dx = \frac{1}{1+t^2}.$$

5. Bemerkenswert einfach ist auch die Rechnung für die Gamma- Verteilungen:
$$\int_0^\infty e^{itx}\,\frac{1}{\Gamma(\alpha)}\,x^{\alpha-1}\cdot e^{-x}\,dx = \int_0^\infty \frac{1}{\Gamma(\alpha)}\,x^{\alpha}\cdot e^{-x(1-it)}\,\frac{1}{x}dx = (1-it)^{-\alpha}.$$

6. Wenn ein Maß μ auf $\mathbb{Z}$ konzentriert ist, dann ist die charakteristische Funktion eine gleichmäßig konvergente trigonometrische Reihe mit nichtnegativen Koeffizienten.
$$\varphi(t) = \sum p_n e^{int} \qquad \text{mit} \quad p_n = \mu(\{n\}).$$

Uns soll es hier nicht so sehr um einzelne Maße und ihre charakteristischen Funktionen gehen. Wir wollen den Raum $\mathcal{M}_1(\mathbb{R}^p)$ der borel'schen W-Maße auf $\mathbb{R}^p$ diskutieren.

Satz 2.5.13. *Für jedes borelsche W-Maß auf dem $\mathbb{R}^p_{Sp}$ ist die charakteristische Funktion eine stetige komplexwertige Funktion auf dem Raum $\mathbb{R}^p_Z$ mit den Eigenschaften*

i) $\varphi(0) = 1, \qquad |\varphi(t)| \le 1$

ii) $\varphi(-t) = \overline{\varphi(t)}$ *für alle* t

iii) Für jedes N *und jedes Tupel* $t_1, t_2, \ldots, t_N$ *ist die* $N \times N$*- Matrix* H *mit den Einträgen* $h_{lk} = \varphi(t_k - t_l)$ *eine positiv semidefinite Matrix.*

Beweis. *Zu* $\varepsilon > 0$ *wählen wir* M *so groß, dass* $\mu\{\|x\| \le M\} \ge 1 - \varepsilon/4$.
$\delta > 0$ *sei so klein, dass* $|y| < \delta \Rightarrow |e^{iy} - 1| < \varepsilon/2$. *Für* $\|h\| < \frac{1}{M}\delta$ *gilt dann*
$$\big|\varphi(t+h) - \varphi(t)\big| = \left|\int e^{it\cdot x}\big[e^{ih\cdot x} - 1\big]\,d\mu(x)\right| \le 2\cdot\mu\{\|x\| > M\} + \varepsilon/2 < \varepsilon.$$

Wir beweisen nun die Eigenschaft iii), die man die positive Definitheit der Funktion $\varphi(\cdot)$ *nennt. (Eigentlich müsste es Semidefinitheit heissen.)*
Sind $\xi_1, \xi_2, \ldots, \xi_N$ *irgendwelche komplexen Zahlen, so gilt mit* $h(x) = \sum \xi_k \cdot e^{i\langle t_k, x\rangle}$
$$\sum_{k,l=1}^N \bar{\xi}_l\cdot\varphi(t_k - t_l)\cdot\xi_k = \sum_{kl}\bar{\xi}_l\cdot\int e^{i\langle (t_k - t_l), x\rangle}\,d\mu(x)\cdot\xi_k$$
$$= \int\Big(\overline{\sum_l \xi_l e^{i\langle t_l, x\rangle}}\Big)\cdot\Big(\sum_k \xi_l e^{i\langle t_k\cdot x\rangle}\Big)\,d\mu(x) = \int\big|h(x)\big|^2\,d\mu(x).$$

Man kann zeigen, dass für jede stetige Funktion $\varphi(\cdot)$ mit den Eigenschaften i), ii), iii) genau ein Wahrscheinlichkeitsmaß mit der charakteristischen Funktion φ existiert.

Hinweis: ('Schwache Konvergenz der Maße') Wenn eine Folge charakteristischer Funktionen $\big(\varphi_n(\cdot)\big)_n$ gleichmäßig auf Kompakten konvergiert, dann konvergieren die dazugehörigen Wahrscheinlichkeitsmaße μ_n im Sinne der sog. schwachen Konvergenz gegen ein Limesmaß μ :

$$\int f(x)\ d\mu_n(x) \to \int f(x)\ d\mu(x) \qquad \text{für alle stetigen beschränkten} \quad f(\cdot).$$

Wir werden uns diesem Faktum in mehreren Schritten nähern.

Faltung und Glättung

Ein traditioneller Grund für die Wertschätzung der charakteristischen Funktionen ist die Tatsache, dass sich die Faltung von Wahrscheinlichkeitsverteilungen bei den charakteristischen Funktionen sehr einfach widerspiegelt: Die charakteristische Funktion des 'Faltungsprodukts' $\rho = \mu * \nu$ ist das punktweise Produkt der charakteristische Funktionen.

Satz 2.5.14 (Die Produktformel für charakteristische Funktionen)**.**
Ist ρ *das Faltungsprodukt der W-Maße* μ *und* ν*:* $\rho = \mu * \nu$*, so gilt*

$$\int e^{i\langle t,x\rangle} d\mu(x) = \varphi(t), \quad \int e^{i\langle t,x\rangle} d\nu(x) = \psi(t) \qquad \int e^{i\langle t,x\rangle} d\rho(x) = \varphi(t)\cdot\psi(t).$$

Beispiel. Man kann die Rechnung in den Beispielen 1) und 2) als Bestätigung verstehen, dass Faltung einer uniformen Verteilung mit sich selbst tatsächlich eine Dreiecksverteilung ergibt. Das Beispiel 5) zeigt, dass die Faltung der Gamma-Verteilung zum Parameter α mit der Gamma-Verteilung zu β die Gamma-Verteilung zu $\alpha+\beta$ ergibt.

Für den Beweis der Produktformel benötigen wir eine passende Notation. Nicht besonders übersichtlich ist die Notation, die sich auf die Idee der verschobenen Maße stützt:

$$\mu * \nu(B) = \int \nu(B-y)\ d\mu(y) \qquad \text{insbesondere} \qquad \delta_y * \nu(\cdot) = \nu(\cdot\ -y)$$

Die Stochastiker machen sich das Leben leicht, indem sie unabhängige Zufallsvektoren mit den betreffenden Verteilungen ins Spiel bringen. Wenn X ein Zufallsvektor mit der Verteilung $d\mu$ ist, und, davon stochastisch unabhängig, Y ein Zufallsvektor mit der Verteilung $d\nu$, dann schreiben die Stochastiker

$$\varphi(t) = \mathfrak{E}e^{i\langle t,X\rangle}, \quad \psi(t) = \mathfrak{E}e^{i\langle t,Y\rangle} \quad \Longrightarrow \quad \varphi(t)\cdot\psi(t) = \mathfrak{E}e^{i\langle t,X\rangle}\cdot\mathfrak{E}e^{i\langle t,Y\rangle} = \mathfrak{E}e^{i\langle t,X+Y\rangle}$$

Für Verteilungen mit Dichte haben wir die bekannte Formel

$$\mu(dx) = f(x)\ dx, \quad \nu(dx) = g(x)\ dx$$
$$\Longrightarrow \quad \big(\mu*\nu\big)(dx) = h(x)\ dx \quad \text{mit} \quad r(x) = \int f(x-y)g(y)\ dy,$$

Auch diese Charakterisierung des Faltunsprodukts ist nicht immer bequem. Der übersichtlichste Beschreibung ist zweifellos die folgende: Wenn μ und ν W-Maße sind dann ist $\mu*\nu$ das Bildmaß des Produktmaßes $\mu\otimes\nu$ unter des Addition. Man kann es auch so sagen: Das Faltungsprodukt ist dadurch gekennzeichnet, dass für alle beschränkten messbaren $a(\cdot)$ gilt:

$$\int a(z)\ \big(\mu*\nu\big)(dz) = \int\int a(x+y)\ \big(\mu\otimes\nu\big)(dx,dy).$$

Speziell für die Funktionen $a(z) = e^{i\langle t,z\rangle}$ ergibt das die Produktformel

$$\int e^{i\langle t,z\rangle}\ \big(\mu*\nu\big)(dz) = \int\int e^{i\langle t,x\rangle}\cdot e^{i\langle t,y\rangle}\ \big(\mu\otimes\nu\big)(dx,dy) = \int e^{i\langle t,x\rangle}\ \mu(dx)\cdot\int e^{i\langle t,y\rangle}\ \nu(dy).$$

Assoziativität und Distributivität des Faltungsprodukts

Der Zugang über die Produktmaße zeigt auch sehr einfach die Assoziativität des Faltungsprodukts. Das Faltungsprodukt $\mu*\nu*\rho$ ist das Bildmaß des Produktmaßes $\mu\otimes\nu\otimes\rho$ unter der Addition.

Die (reellen) Linearkombinationen von W-Maßen sind bekanntlich die signierten Maße. Wenn man das Faltungsprodukt in der offensichtlichen Weise auf diesen reellen Vektorraum fortsetzt, dann hat man auch die Distributivgesetze; man hat eine (reelle) Algebra. Wenn man diese Algebra mit der Totalvariationsnorm ausstattet, erhält man eine (reelle) Banachalgebra. Die Eigenschaft $\|\tau_1*\tau_2\|\leq\|\tau_1\|\cdot\|\tau_2\|$ ergibt sich aus der Tatsache, dass das Faltungsprodukt zweier W-Maße ein W-Maß ist. Die Vollständigkeit ergibt sich aus der Vollständigkeit der Räume $L^1(\Omega,\mu)$: Ist nämlich $\tau_1,\tau_2,\ldots$ eine Cauchyfolge im Raum der signierten Borel-Maße, dann existiert ein W-Maß μ mit $\tau_n\ll\mu$ für alle n. $d\tau_n = f_n\ d\mu$. Und im Raum aller bzgl. μ totalstetigen signierten Maße ist die Totalvariation durch die 1-Norm der Dichten beschrieben. Die Cauchy-Folgen in irgendeiner L^1-Norm konvergieren in dieser L^1-Norm.

Bemerke: Man kann so auch eine komplexe Banachalgebra konstruieren. Man verliert da aber die Verbandsstruktur; die wird uns hier aber wichtig sein. Am Ende interessiert vor allem die konvexe Menge der W-Maße $\mathcal{M}_1^+(\mathbb{R}^p)$, und manchmal auch der konvexe Kegel $\mathcal{M}^+(\mathbb{R}^p)$der endlichen Maße, eher selten der geordnete Vektorraum $\mathcal{M}^\pm(\mathbb{R}^p)$ der signierten Maße.

Eine wichtige Teilalgebra ist die Menge derjenigen signierten Maße, die bzgl. des Lebesgue-Maß $\lambda(dx) = dx$ totalstetig sind. Diese Teilalgebra ist nichts anderes als der Raum $L^1(\mathbb{R}^n,dx)$, ausgestattet mit dem Faltungsprodukt.

Die reellen (oder komplexen) Banachräume $L^1(\Omega,\mathfrak{A},\rho)$ zu einem allgemeinen σ-endlichen Maßraum sind im vorigen Abschnitt ausführlich untersucht worden. Hier soll nun die Banachalgebra $L^1(\mathbb{R}^p,\mathcal{O},+,*)$ als Teilalgebra von $\mathcal{M}^\pm(\mathbb{R}^p)$.

Im Raum $\mathcal{M}_1(\mathbb{R}^p)$ interessieren wir uns für vor allem für den Begriff der schwachen Konvergenz. Die Totalvariationsnorm wird dabei nur eine dienende Rolle spielen.

Daß die signierten Maße mit Dichte in der Tat eine Teilalgebra von $\mathcal{M}^\pm(\mathbb{R}^p)$ bilden, zeigt der

Satz 2.5.15. *Wenn eines der W-Maße* μ *oder* ν *eine Dichte bzgl. des Lebesgue-Maßes besitzt, dann auch das Faltungsprodukt.*

Beweis. *Für jede Nullfunktion im Sinne des Lebegue-Maßes* $a(\cdot)$ *verschwindet das iterierte Integral* $\int\left(\int a(x+y)p(x)\ dx\right)\ d\nu(y)$, *weil alle verschobenen Funktionen* $a(\,\cdot+y)$ *Nullfunktionen sind. Das Faltungsprodukt ist also totalstetig bzgl. des Lebegue-Maßes.*

Satz 2.5.16. *Es sei* $(\nu_n)_n$ *eine Folge von W-Maßen mit* $\nu(U) \to 1$ *für jede Umgebung* U *des Nullpunkts. Für jedes W-Maß mit Dichte* $\mu(dx) = f(x)\ dx$ *gilt dann* $\|\nu_n * \mu - \mu\| \to 0$.

Beweis. *Es sei* $\mathfrak{R}$ *ein Ring aller elementaren Rechtecke* $R = (a_1, b_1] \times \ldots \times (a_p, b_p]$. *Es genügt, die Behauptung für diejenigen* μ *nachzuweisen, deren Dichte eine Indikatorfunktion* 1_R *ist. Man kann nämlich bekanntlich jede Dichte* f *durch elementare Treppenfunktionen in der* L^1*-Norm approximieren; und ist* $d\mu_\varepsilon(x) = f_\varepsilon(x)\ dx$ *eine elementare Treppenfunktion mit* $\|f - f_\varepsilon\|_1 < \varepsilon$ *und , so gilt* $\left\|\nu * \left(\mu - \mu_\varepsilon\right)\right\| < \varepsilon$ *für jedes W-Maß* ν.

Betrachten wir eine Nullfolge (x_n) *und dazu die* δ*-Maße* $\nu_n = \delta_{x_n}$. *Wenn* $d\mu = f(x)\ dx$, *dann ist die Dichte von* $\nu_n * \mu$ *die 'verschobene' Funktion* $f(x - x_n)$. *Für eine Indikatorfunktion* $f = 1_R$ *gilt offensichtlich* $\int\left|f(x - x_n) - f(x)\right|\ dx \to 0$. *Genauer:*

$$\forall \varepsilon > 0\ \exists U \quad \text{Umgebg. von } 0: \forall x \in U \quad \|\delta_x * \mu - \mu\| < \varepsilon.$$

Wenn das Gewicht des W-Maßes ν *bis auf* ε' *auf eine solche Umgebung konzentriert ist, dann ist auch* $\|\nu * \mu - \mu\|$ *klein. Das sieht man folgendermaßen: Wenn* $f(x)$ *die Dichte von* μ *ist, dann hat das signierte Maß* $\nu * \mu - \mu$ *die Dichte* $g(x) = \int\left[f(x - y) - f(x)\right]\ d\nu(y)$. *Die Norm wird abgeschätzt*

$$\int |g(x)|\ dx \le \int\left|f(x - y) - f(x)\right|\ d\nu(y)\ dx \le \int_U \|\delta_y * \mu - \mu\|\ d\nu(y) + \nu\left(\complement U\right) \cdot 2\|\mu\|.$$

Glättung: Die Approximation von Maßen μ (oder ihren Dichten) durch Faltungsprodukte $\nu_n * \mu$ (bzw. deren Dichten) mit W-Maßen ν_n, die sich wie im Satz auf kleine Umgebungen des Nullpunkts konzentrieren, nennt man Glättung. Für verschiedene Zwecke bieten sich verschiedene glättende Folgen (ν_n) an. Im Abschnitt über Differenzierbarkeit werden wir die sog. Dirac-Folgen schätzen lernen. Diese haben unendlich oft differenzierbare Dichten, die auf kompakte Umgebungen konzentriert sind, die sich auf den Nullpunkt zusammenziehen. Man konstruiert z. B. ausgehend von einer nichtnegativen unendlich oft differenzierbaren Funktion ϕ mit $\int \phi(x)\ dx = 1$, welche ausserhalb einer kompakten Umgebung des Nullpunkts verschwindet, die W-Maße $d\nu_\sigma(x) = \phi(\frac{1}{\sigma}x)\ \left(\frac{1}{\sigma}\right)^p dx$. Für eine Nullfolge σ_n ergibt sich eine glättende Folge. Die Glättung mit Dirac-Folgen im strengen Sinn liefert unendlich oft differenzierbare Dichten (bzw. Funktionen).

Für Spezialfälle des Satzes Stone-Weierstraß haben wir früher schon konkrete glättende Folgen ins Feld geführt, die im strengen Sinn keine Dirac-Folgen sind, weil sie nämlich keinen beschränkten Träger haben. Im Abschnitt über Fourier-Integrale wird uns die Glättung mit den sog. Gauss-Kernen gute Dienste leisten. Im folgenden Teilabschnitt über Maße auf dem Raum $\mathbb{R}/2\pi$ werden wir die Glättung mit den sog. Fejér-Kernen heranziehen.

Schwache Konvergenz im Raum $\mathcal{M}(\mathbb{R}/2\pi)$
Bevor wir uns ernstlich mit den W-Maßen auf dem Raum $\mathbb{R}^p_{Sp}$ und ihren charakteristischen Funktionen befassen, betrachten wir zuerst einen technisch etwas einfacheren Fall. Wir studieren die W-Maße auf dem kompakten Raum $\mathbb{R}/2\pi$ und ihre charakteristischen Folgen.

Definition 2.19 (Charakteristische Folgen)**.** Jedem W-Maß $\mu(dt)$ auf dem Raum $\mathbb{R}/2\pi$ wird seine charakterische Folge zugeordnet: $(r_n)_{n\in\mathbb{Z}}$ zu. $r_n = \int e^{int}d\mu$ heisst der n-te charakteristische Koeffizient des W-Maßes μ auf $\mathbb{R}/2\pi$.

Bemerke: Es besteht hier eine Ähnlichkeit zum Begriff des Fourier-Koeffizienten. Man definiert bekanntlich den n-ten Fourier-Koeffizienten der 2π-periodischen integrablen Funktion $f(t)$: $c_n = \frac{1}{2\pi}\int e^{-int}f(t)\ dt$. Man sollte aber den Unterschied nicht verwischen. Die Folge der Fourier-Koeffizienten strebt immer nach 0 für $n \to \pm\infty$ ('Lemma von Riemann- Lebesgue'). Wenn f quadratintegrabel ist, dann ist die Folge der Fourier-Koeffizienten quadratsummabel. Die Folgen charakteristischer Koeffizienten konvergieren nicht notwendigerweise nach 0 für $n \to \pm\infty$.

Beispiel 2.5.3.

1. Für das auf den Punkt $\tilde{t}$ konzentrierte Wahrscheinlichkeitsmaß $\delta_{\tilde{t}}$ ergibt sich die charakteristische Folge $r_n = \tilde{\zeta}^n$ mit $\tilde{\zeta} = e^{i\tilde{t}}$. Es gilt $r_{n+m} = r_n \cdot r_m$.

   Ist $(r_n)_n$ die Folge der charakteristischen Koeffizienten von μ, so hat das verschobene Maß $\delta_{\tilde{t}} * \mu$ die charakteristischen Koeffizienten $(r_n \cdot e^{in\tilde{t}})_n$.

2. Es sei M eine natürliche Zahl und $d\mu^{(M)}(t)$ die Gleichverteilung auf der Punktmenge $\left\{2\pi\frac{k}{M} :\ k = 1, \ldots, M\right\}$, also $d\mu^{(M)}(t) = \frac{1}{M}\sum_1^M \delta_{2\pi\frac{k}{M}}$. Die charakteristische Folge hat den Wert 1 in den Vielfachen von M und sonst den Wert 0.

   Für die kontinuierliche Gleichverteilung $d\mu(t) = \frac{1}{2\pi}dt$ ist der nullte charakteristische Koeffizient $= 1$ und alle weiteren $= 0$.

3. Es sei $d\rho(t)$ eine Wahrscheinlichkeitsverteilung mit einer Dichte der Form

$$d\rho(t) = h(t)\tfrac{1}{2\pi}\ dt = \sum r_n e^{-int}\tfrac{1}{2\pi}\ dt \qquad \text{mit} \quad h(\cdot) \geq 0, \quad \int h(t)\tfrac{1}{2\pi}\ dt = 1.$$

   Die charakteristische Folge ist $(r_n)_{n\in\mathbb{Z}}$; denn $\int e^{int}\ d\rho(t) = r_n$.
   Wir werden unten diskutieren, wie man einer summablen Folge (r_n) ansehen kann, ob die trigonometrische Reihe $\sum r_n e^{-int}$ eine nichtnegative Funktion liefert.

Eine spezielle Glättung:
Wir diskutieren die charakteristische Folge einer Verteilung, welche mit einem Fejér-Kern geglättet ist. Sei also $F_N(t)\frac{1}{2\pi}\ dt = \sum\left(1 - \frac{|n|}{N}\right)^+ e^{int}\frac{1}{2\pi}\ dt$. Wir haben früher durch eine explizite Rechnung bewiesen, dass es sich tatsächlich um W-Dichten handelt, welche sich für großes N auf die Nähe des Nullpunkts konzentrieren. Für jede Verteilung mit Dichte $d\mu(t) = f(t)\frac{1}{2\pi}dt$ konvergieren die geglätteten Dichten im L^1-Sinn gegen $f(\cdot)$, wie wir oben für den analogen Fall des $\mathbb{R}^p$ bewiesen haben.

Andererseits kennen wir die charakteristische Folge der W-Dichte $d\nu_N(t) = F_N(t)\frac{1}{2\pi}\,dt$. Sie hat die 'Dreiecksform' $r^{(N)}(n) = \left(1 - \frac{|n|}{N}\right)^+$.

Die Produktformel für charakteristische Folgen sagt uns also insbesondere: Wenn man ein W-Maß μ mit dieser Verteilung faltet, dann erhält man also ein W-Maß $\nu_N * \mu$, dessen chakteristische Koeffizienten für alle n mit $|n| \geq N$ verschwinden; die Glättung führt auf W-Maße, deren Dichten trigonometrische Polynome sind, eine Tatsache, die wir schon an anderer Stelle festgestellt haben. (Dort interessierte uns übrigens auch die Frage, inwieweit für gewisse f auch punktweise oder gleichmäßige Konvergenz herrscht.)

Satz 2.5.17 (Charakteristische Folgen sind positivdefinite Folgen).
Für jedes W-Maß auf $\mathbb{R}/2\pi$ ist die charakteristische Folge eine Zahlenfolge mit

i) $r(0) = 1, \qquad |r(n)| \leq 1$

ii) $r(-n) = \overline{r(n)}$ *für alle* n

iii) Für jedes N *ist die* $N \times N$*- Matrix* H *mit den Einträgen* $h_{mn} = r(n-m)$ *eine positiv semidefinite Matrix.*

Beweis. *Für ein beliebiges* N*-Tupel komplexer Zahlen* $\xi_1, \xi_2, \ldots \xi_N$ *gilt*

$$\sum_{m,n=1}^{N} \bar{\xi}_m \cdot \int e^{i(n-m)} d\mu \cdot \xi_n = \int \left|\xi(t)\right|^2 \, d\mu(t) \qquad \textit{mit} \qquad \xi(t) = \sum_n \xi_n e^{int}.$$

Beispiel 2.5.4.
Ein interessantes Beispiel der nichtnegativen Doppelsumme liefert das Tupel $\xi_n^{(s,N)} = e^{-isn}$ für $n = 1, 2, \ldots, N$, und $= 0$ für $n > N$. Wir erhalten damit die Nichtnegativität von trigonometrischen Polynomen, die uns oben bei der Faltung mit den Fejér-Kernen begegnet sind.

$$\begin{aligned}
\int \Big|\sum_1^N e^{i(t-s)n}\Big|^2 \, d\mu(t) &= \sum_{m,n=1}^{N} \bar{\xi}_m \cdot r_{n-m} \cdot \xi_n = \sum_{m,n=1}^{N} r_{n-m} \cdot e^{-is(n-m)} = \\
&= N \cdot r_0 + (N-1) \cdot r_1 e^{-is} + (N-1) \cdot r_{-1} e^{is} + (N-2) \cdot r_2 e^{-2is} + \cdots \\
&= N \cdot \sum \left(1 - \tfrac{|n|}{N}\right)^+ r_n \cdot e^{-ins} \geq 0 \quad \text{für positiv semidefinite Folgen } (r_n)_n.
\end{aligned}$$

Sprechweise. Eine Zahlenfolge $(r_n)_{n\in\mathbb{Z}}$ mit den Eigenschaften i), ii), ii) wird eine normierte positiv definite Folge genannt. (Eigentlich müsste man von einer positiv semidefiniten Folge sprechen.) Wir bemerken, dass die Eigenschaft ii) nicht eigens aufgeführt werden müsste; denn sie ist in iii) als Spezialfall enthalten.

Satz 2.5.18 (Eindeutige Bestimmtheit und schwache Konvergenz).
Ein W-Maß μ auf (der Borel-Algebra über) $\mathbb{R}/2\pi$ ist durch die Folge seiner charakteristischen Koeffizinten eindeutig bestimmt. Wenn für eine Folge von W-Maßen $\mu^{(N)}$ die

charakteristischen Koeffizienten gegen die charakteristischen Koeffizienten des W-Maßes μ *konvergieren. dann gilt*

$$\int f(t)\ d\mu^{(N)}(t) \to \int f(t)\ d\mu(t) \qquad \text{für alle stetigen } 2\pi\text{-periodischen Funktionen } f(\cdot).$$

Beweis. *Wir müssen zeigen, dass das* μ*-Integral* $\langle\mu, f(\cdot)\rangle$ *für alle stetigen* 2π*-periodischen Funktionen* f *durch die charakteristischen Koeffizienten* $\langle\mu, e^{in(\cdot)}\rangle$ *eindeutig bestimmt ist. Für die trigonometrischen Polynome ergibt sich das aus der Linearität. Wenn* f *durch ein trigonometrisches Polynom gleichmäßig bis auf* ε *approximiert wird, dann gilt auch* $\left|\langle\mu, f(\cdot) - p_\varepsilon\rangle\right| < \varepsilon$ *für alle W-Maße* μ*. Der Satz von Stone-Weierstraß vollendet den Beweis der eindeutigen Bestimmtheit.*

Wenn für $N \to \infty$ *alle charakteristischen Koeffizienten konvergieren ('punktweise Konvergenz'), dann ergibt das* $\lim_{N\to\infty}\langle\mu^{(N)} - \mu, p(\cdot)\rangle = 0$ *für jedes trigonometrische Polynom. Zur stetigen Funktion* f *sei* p_ε *so gewählt, dass* $\sup\{|f(t) - p_\varepsilon(t)| :\ t \in \mathbb{R}/2\pi\} < \varepsilon$ *(Gleichmäßige Approximation gemäß dem Satz von Stone-Weierstraß!)*
Es gilt dann $\left|\langle\mu^{(N)} - \mu, f - p_\varepsilon\rangle\right| < 2\varepsilon$ *für alle* N*. Wenn wir* N *so groß wählen, dass* $\left|\langle\mu^{(N)} - \mu, p_\varepsilon(\cdot)\rangle\right| < \varepsilon$*, dann haben wir* $\left|\langle\mu^{(N)} - \mu, f\rangle\right| < 4\varepsilon$.

Satz.
Wenn die charakteristische Folge $(m_n)_n$ *des W-Maßes* μ *summabel ist, dann besitzt* μ *besitzt eine Dichte, und zwar* $d\mu(t) = \sum m_n e^{-int}\ \frac{1}{2\pi}\ dt$.
Die Faltung mit den Fejér-Kernen approximiert das W-Maß μ *mit den charakteristischen Koeffizienten* m_n *durch die W-Dichten*

$$d\mu^{(N)}(t) = \sum\left(1 - \tfrac{|n|}{N}\right)^+ \cdot m_n e^{-int}\ \tfrac{1}{2\pi}\ dt$$

Diese W-Maße mit Dichten konvergieren für $N \to \infty$ *im Sinne der schwachen Konvergenz gegen* μ.

Beweis. *Die positive Definitheit der Folge garantiert die Nichtnegativität der trigonometrischen Polynome* $\sum\left(1 - \frac{|n|}{N}\right)^+ \cdot m_n e^{-int}$*, wie wir oben gesehen haben. Es handelt sich um trigonometrische Polynome.* $\mu^{(N)}$ *entsteht aus* μ *durch Faltung mit dem* N*-ten Fejér-Kern.* $\mu^{(N)} = \nu_N * \mu$*. Die charakteristischen Koeffizienten* $\left(1 - \frac{|n|}{N}\right)^+ \cdot m_n$ *konvergieren offenbar gegen die charakteristischen Koeffizienten von* μ*. Man beachte, dass es hier nicht um* L^1*-Konvergenz geht, und schon gar nicht um gleichmäßige Konvergenz; es geht um schwache Konvergenz. Insbesondere wird das* δ*-Maß im Punkt* $\tilde{t}$ *approximiert durch*

$$\sum\left(1 - \tfrac{|n|}{N}\right)^+ e^{-in(t-\tilde{t})}\ \tfrac{1}{2\pi}\ dt \quad = \quad F^{(N)}(t - \tilde{t})\ \tfrac{1}{2\pi}\ dt \quad = \quad \frac{1}{N}\left(\frac{\sin\frac{N}{2}(t - \tilde{t})}{\sin\frac{1}{2}(t - \tilde{t})}\right)^2\ \frac{1}{2\pi}\ dt$$

Die schwache Kompaktheit von $\mathcal{M}_1(\mathbb{R}/2\pi)$.
Die konvexe Menge der normierten positiv semidefiniten Folge auf $\mathbb{Z}$ ist offenbar folgenkompakt im Sinne der punktweisen Konvergenz. Aus jeder Folge normierter positivdefiniter Folgen kann man eine in allen Positionen konvergente Teilfolge auswählen, und die Limesfolge ist eine normierte positivdfinite Folge. Wegen der eineindeutigen Entsprechung ergibt sich die Kompaktheit der konvexen Menge $\mathcal{M}_1\big(\mathbb{R}/2\pi\big)$ bzgl. der schwachen Konvergenz: Aus jeder Folge von W-Maßen auf dem kompakten Raum $\mathbb{R}/2\pi$ kann man eine Teilfolge auswählen, die schwach konvergiert.

Die Wahrscheinlichkeitsmaße auf $\mathbb{R}$
Eine gute Stütze bei den Untersuchungen der konvexen Menge $\mathcal{M}_1\big(\mathbb{R}/2\pi\big)$ ist die Kompaktheit des Grundraums. Bei der Menge $\mathcal{M}_1\big(\mathbb{R}\big)$ der W-Maße auf dem lokalkompakten Raum $\mathbb{R}^d$ ist die Sachlage nur wenig komplizierter. Der entscheidende Satz lautet.

Satz 2.5.19. *Wenn eine Folge charakteristischer Funktionen gleichmäßig auf Kompakten konvergiert, dann ist die Grenzfunktion die charakteristische Funktion eines W-Maßes und wir haben schwache Konvergenz gegen dieses Maß.*

Eine Familie von W-Maßen auf dem $\mathbb{R}^p$ ist genau dann bedingt kompakt in der Topologie der schwachen Konvergenz, wenn die charakteristischen Funktionen im Nullpunkt gleichgradig stetig sind.

Wir werden diesen Satz hier nicht beweisen. Wir machen ihn plausibel, indem wir (für den Fall p=1) Überlegungen zu den charakteristischen Funktionen präsentieren, die auf den obigen Sätzen über charakteristische Folgen aufbauen. Zunächst übertragen wir den Eindeutigkeitssatz auf den Fall der W-Maße auf $\mathbb{R}$.

Satz 2.5.20 (Eindeutige Bestimmtheit).
Ein W-Maß auf $\mathbb{R}$ ist durch seine charakteristische Funktion eindeutig bestimmt.

Beweis. *Wenn wir die charakteristische Funktion in den Punkten $t_k = \frac{k}{M}$, $k \in \mathbb{Z}$ betrachten, haben wir eine positiv semidefinite Folge. Die Werte $\varphi(t_k)$ legen das μ-Integral fest auf allen $(M \cdot 2\pi)$-periodischen Funktionen der Form* $p^{(M)}(x) = \sum_k c_k e^{i\frac{k}{M}} \cdot x$. *Damit liegt auch das μ-Integral $\int f(x)\ d\mu(x)$ fest für alle diejenigen Funktionen f, die sich gut durch durch solche $p^{(M)}$ approximieren lassen.*

Sei f stetig mit $\|f\|_\infty \leq 1$, und sei $\varepsilon > 0$ vorgegeben.
Es existiert M so, dass $\mu\big[-M\pi, +M\pi\big] > 1 - \varepsilon$, und es existiert (nach dem Satz von Stone-Weierstraß) $p^{(M)}$ so, dass

$$\sup\Big\{\big|f(x) - p^{(M)}(x)\big|\colon\ x \in \big[-M\pi, +M\pi\big]\Big\} < \varepsilon.$$

Wir haben $|p^{(M)}(x)| < 1 + \varepsilon$ *und* $\big|f(x) - p^{(M)}(x)\big| < 2 + \varepsilon$ *für alle x, sowie*

$$\Big|\int f\ d\mu - \int p^{(M)}\ d\mu\Big| \leq$$

$$\leq \int_{-M\pi}^{M\pi} \big|f(x) - p^{(M)}(x)\big|\ d\mu(x)\ +\ (2+\varepsilon)\cdot \mu\Big(\complement\big[-M\pi, +M\pi\big]\Big) < 3\varepsilon.$$

Bemerke: Es sei φ eine stetige positivdefinite Funktion mit $\varphi(0) = 1$. Es existiert dann genau ein W-Maß $\mu^{(M)}$ auf dem Intervall $\big(-M\pi, +M\pi\big]$, sodass mit

$$\int e^{it_k x}\ d\mu^{(M)}(x) = \varphi(t_k) = \int e^{it_k x}\ d\mu(x) \quad \text{für alle } k \in \mathbb{Z}.$$

Für die Folge $M = 2^m$ strebt die Folge der Maße $\mu^{(M)}$ gegen ein Maß $\mu^{(\infty)}$ mit der charakteristischen Funktion $\varphi(\cdot)$. Die Stetigkeit von φ im Nullpunkt garantiert, dass die $\mu^{(M)}$ bis auf ein vorgegebenes ε auf ein geeignetes beschränktes Intervall konzentriert sind. (Ohne Beweis!) Das Maß $\mu^{(M)}$, aufgefasst als Maß auf $\mathbb{R}/M\cdot 2\pi$ ergibt sich aus $\mu^{(\infty)}$ durch den Übergang zu den Äquivalenzklassen modulo $M \cdot 2\pi$.

Interessanter als diese Approximation von μ ist eine andere Approximation des Maßes zur positiv definiten Funktion φ. Dazu treffen wir einige Vorbereitungen. Wir diskutieren die kontinuierlichen Analoga der Dirichlet- und Fejér-Kerne. Wir nennen sie die kontinuierlichen Kerne zur Spannweite $[-T, +T]$

$$d^{(T)}(x) = \int_{-T}^{T} \cos ax\ da = \frac{2}{x}\sin(Tx); \qquad f^{(T)}(x) = \frac{1}{T}\int_0^T d^{(u)}(x)\ du = T \cdot \left(\frac{\sin(\frac{1}{2}Tx)}{\frac{1}{2}Tx}\right)^2$$

$$\int_{-\infty}^{\infty} \tfrac{1}{2\pi} f^{(T)}(x)\ dx\ = 1 = \lim_{M\to\infty}\int_{-M}^{M} \tfrac{1}{2\pi} d^{(T)}(x)\ dx \qquad \text{für alle } T.$$

Die Aussage, dass $\varphi(\cdot)$ positivdefinit ist lässt sich offenbar verallgemeinern

$$\sum_{k,l=1}^{N} \bar{\xi}_l \cdot \varphi(t_k - t_l) \cdot \xi_k \geq 0 \qquad \text{für alle Belegungen } \xi(\cdot) \text{ der Punktemenge } \{t_k\}$$

$$\int\int \overline{\xi(b)} \cdot \varphi(a-b) \cdot \xi(a) \geq 0 \qquad \text{für 'gute' Belegungen } \xi(\cdot) \text{ eines Intervalls } [-T, +T]$$

Eine gute Belegung in diesem Sinn ist gewiss die Belegung $\xi^{(s,T)}(\cdot)$, die ausserhalb des Intervalls $[-T, +T]$ verschwindet und für $|a| < T$ den Wert e^{-iay} hat.
In diesem Fall hängt der Integrand im Integrationsbereich nur von der Differenz der Argumente ab: $\quad \overline{\xi(b)} \cdot \varphi(a-b) \cdot \xi(a) = e^{-i(a-b)y} \cdot \varphi(a-b)$.
Die Läge des Integrationswegs für festes $\quad u = a - b \quad$ ist $\quad 2T\big(1 - \frac{|u|}{2T}\big)^+$. Schrittweises Integrieren ergibt

$$\frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T} e^{-i(a-b)y} \cdot \varphi(a-b)\ da\ db = \frac{1}{2T}\int e^{-iuy} \cdot \varphi(u) \cdot \big(1 - \tfrac{|u|}{2T}\big)^+\ du$$

Auf der anderen Seite gilt für die charakteristischen Funktion $\varphi(\cdot)$ des Wahrscheinlichkeitsmaßes μ wegen $\int_{-T}^{T} e^{ia(x-y)}\ da = \frac{2}{(y-x)}\sin(T(y-x))$

$$\frac{1}{4T^2}\int_{-T}^{T}\int_{-T}^{T} e^{iby} \cdot \left(\int e^{-ibx} \cdot e^{iax}\ d\mu(x)\right) \cdot e^{-iay}\ da\ db = \int\left(\frac{\sin(T(y-x))}{T(y-x)}\right)^2 d\mu(x).$$

Dies zeigt

$$\int e^{-iuy}\cdot\varphi(u)\cdot\bigl(1-\tfrac{|u|}{2T}\bigr)^{+}\,du \;=\; \int f^{(2T)}(y-x)\,d\mu(x) = 2\pi\cdot m^{(2T)}(y).$$

Das Faltungprodukt von μ mit dem kontinuierlichen Fejér-Kern $\nu_{2T}(dy)=\frac{1}{2\pi}f^{(2T)}(y)\,dy$ ist ein W-Maß $\nu_{2T}*\mu$, welches bzgl. des Lebegue-Maßes totalstetig ist Die rechte Seite ist seine Dichte (mit der Normierung $\int m^{(2T)}(y)\,dy=1$ für alle T).

Grenzübergang $T\to\infty$: Wenn $\varphi(\cdot)$ integrabel ist, dann konvergiert die linke Seite gleichmäßig auf Kompakten gegen die stetige Funktion $2\pi\cdot m^{(\infty)}(\cdot)=\int e^{-iu(\cdot)}\cdot\varphi(u)\,du$.

Wir können auch von dem Ausdruck auf der rechten Seite ausgehend argumentieren: Nehmen wir an, dass μ eine stetige Dichte hat, $d\mu(x)=m(x)\,dx$. Die Familie der kontinuierlichen Fejér-Kerne leistet eine Regularisierung im Sinne der Dirac-Folgen. Die rechte Seite strebt daher (gleichmäßig auf Kompakten) gegen $m^{(\infty)}(x)$. Wir nennen das Ergebnis den Satz von der expliziten Inversion für summable charakteristische Funktionen

Satz 2.5.21. *Wenn für ein W-Maß auf $\mathbb{R}/2\pi$ die charakteristische Folge summabel ist, dann besitzt es eine Dichte bzgl. der uniformen Verteilung*

$$m_n=\int e^{int}\,d\mu(t)\quad\textit{summabel}\qquad\Longrightarrow$$
$$d\mu(t)=m(t)\cdot\tfrac{1}{2\pi}\,dt\quad\textit{mit}\qquad m(t)=\sum m_n\cdot e^{-int}.$$

Wenn für ein W-Maß auf $\mathbb{R}$ die charakteristische Funktion Lebesgue-integrabel ist, dann besitzt das Maß eine Dichte bzgl. des Lebesgue-Maßes

$$\varphi(t)=\int e^{itx}\,d\mu(x)\quad\textit{mit}\quad\int\bigl|\varphi(t)\bigr|\,dt<\infty\qquad\Longrightarrow$$
$$d\mu(x)=m(x)\,dx\quad\textit{mit}\qquad m(x)=\tfrac{1}{2\pi}\int e^{-isx}\cdot\varphi(s)\,ds.$$

Eine Faustregel: Wenn wir die beiden Integralausdrücke zusammenbringen, und das iterierte Integral (unkorrekterweise!) in ein Doppelintgral umschreibt, dann ergeben sich merkwürdige Formeln

$$\begin{aligned}\varphi(t)=\int e^{ity}m(y)\,dy&=\frac{1}{2\pi}\int e^{ity}\Bigl(\int e^{-isy}\varphi(s)\,ds\Bigr)\,dy\\&=\frac{1}{2\pi}\int\int e^{i(t-s)y}\varphi(s)\,dy\,ds;\\m(x)=\frac{1}{2\pi}\int e^{-iux}\varphi(u)\,du&=\frac{1}{2\pi}\int e^{-iux}\Bigl(\int e^{iuy}m(y)\,dy\Bigr)\,du\\&=\frac{1}{2\pi}\int\int e^{iu(y-x)}m(y)\,du\,dy.\end{aligned}$$

Andererseits schreibt man gerne im Sinne der Faltung mit dem 'Dirac-Kern' $\delta(\cdot)$

$$\varphi(t) = \int \varphi(s)\ \delta(t-s)\ ds, \qquad m(x) = \int m(y)\ \delta(x-y)\ dy.$$

Und so lassen sich manche Leute verleiten zu den Formeln

$$\delta(t-s) = \frac{1}{2\pi}\int e^{i(t-s)y}dy = \frac{1}{2\pi}\int \overline{e^{isy}}\cdot e^{ity}dy.$$

Sie nennen das die Orthogonalitätsrelationen in der Familie $\big\{f_t(\cdot) = e^{it(\cdot)}:\ t \in \mathbb{R}\big\}$, in Analogie zu den bekannten Orthogonalitätsrelationen in der diskreten Schar 2π-periodischer Funktionen $\big\{f_n(\cdot) = e^{in(\cdot)};\ n \in \mathbb{Z}\big\}$. Es scheint, dass die Routiniers sich durch solche (mathematisch nicht korrekten) Formeln nicht in die Irre führen lassen.

Als Mathematiker muß man beanstanden, dass die Funktion

$$\psi(s,y) = \frac{1}{2\pi}e^{i(t-s)y}\cdot\varphi(s) \quad \text{auf} \quad \big(\mathbb{R}\times\mathbb{R}, ds\otimes dy\big)$$

nicht integrabel ist. Der Satz von Fubini ist nicht anwendbar.

Hinweis auf Fourier-Integrale:
Die Formeln in der 'expliziten Inversion' erscheinen auch in der Theorie der Fourier-Integrale. Dort zeigt man, dass für manche (aber nicht alle) Elemente f im Hilbertraum $L^2\big(\mathbb{R}^p, dx\big)$ der bzgl. des Lebesgue-Maßes quadratintegrablen Funktionen eine explizite Inversion der Fourier-Transformation durch Integration bewerkstelligt werden kann:

$$\hat{f}(x) = \big(\tfrac{1}{2\pi}\big)^p\cdot\int e^{-itx}\ f(t)\ dt \quad\Longleftrightarrow\quad f(t) = \int e^{itx}\ \hat{f}(x)\ dx.$$

Wenn man sich um die Positivität bzw. positive Definitheit keine Gedanken zu machen hat, dann bieten sich Approximationen an, die übersichtlicher sind als die mit den Fejér-Kernen. Wir kommen später darauf zurück.

Hinweis auf Gruppencharaktere.
Eine beschränkte stetige komplexwertige Funktion $\chi(\cdot)$ auf einer topologischen Gruppe $\big(G, e, \circ\big)$ nennt man einen Gruppencharakter, wenn sie multiplikativ ist: $\chi(g\circ h) = \chi(g)\cdot\chi(h)$ für alle $g, h \in G$. Man sieht sofort, dass als Funktionswerte nur Zahlen vom Betrag 1 in Betracht kommen.

Das (punktweise!) Produkt zweier Gruppencharaktere ist offenbar wieder ein Gruppencharakter. Die Menge der Gruppencharaktere ist eine kommutative Gruppe von komplexwertigen Funktionen auf G. Man nennt sie manchmal die duale Gruppe G^*.

Für die Gruppe $\mathbb{R}/2\pi$ liefert jedes $n \in \mathbb{Z}$ einen Gruppencharakter $e^{in(\cdot)}$. Es gibt keine weiteren stetigen Gruppencharaktere. Die Gruppe der Gruppencharaktere auf $\mathbb{R}/2\pi$ ist isomorph zu $\mathbb{Z}$.

Die Menge $\mathcal{M}^{\pm}\big(\mathbb{R}/2\pi\big)$ der signierten Maße auf der topologischen Gruppe $\mathbb{R}/2\pi$ ist eine Algebra mit der Faltung. Die Abbildung, die jedem W- Maß seine charakteristische Folge zuordnet, ist multiplikativ (im Sinne der punktweisen Multiplikation der 'Funktionen' auf $\mathbb{Z}$.)

Andererseits: Für die Gruppe $\mathbb{Z}$ liefert jedes $\tilde{t}$ einen Gruppencharakter $e^{i(\cdot)\tilde{t}}$ Es gibt keine weiteren Gruppencharaktere auf $\mathbb{Z}$. Die charakteristischen Folgen der δ-Maße auf $\mathbb{R}/2\pi$ sind also die Gruppencharaktere auf $\mathbb{Z}$. Die Menge der Gruppencharaktere auf $\mathbb{Z}$ ist isomorph zur Gruppe $\mathbb{R}/2\pi$.

Die Menge der summablen Folgen (die man auch als signierte Gewichtungen auf $\mathbb{Z}$ verstehen kann) ist eine Algebra bzgl. der Faltung. Die Abbildung, die der summablen Folge $(c_n)_n$ die stetige 2π-periodische Funktion $\sum c_n e^{int}$ zuordnet, ist multiplikativ (im Sinne der punktweisen Multiplikation der Funktionen auf $\mathbb{R}/2\pi$).

Diese Konstruktionen mögen erklären, warum man manchmal sagt, dass die Gruppen $\mathbb{R}/2\pi$ und $\mathbb{Z}$ zueinander dual sind.

Auch für die additive Gruppe der p-Zeilen $\big(\mathbb{R}^p_Z, 0, +\big)$ ist die Menge der stetigen Charaktere leicht zu bestimmen. Für jede p-Spalte x ist $\langle\cdot, x\rangle$ eine Linearform und $\chi(t) = e^{i\langle t,x\rangle}$ ein stetiger Gruppencharakter auf dem Raum der p-Zeilen; und das sind alle stetigen Gruppencharaktere. Man könnte sagen, dass die duale Gruppe hier also in eineindeutiger Beziehung steht zum Dualraum im Sinne der Vektorraumtheorie.— Die Idee der Dualität hat viele (mehr oder weniger intuitive) Konkretisierungen; sie ist aber nicht leicht allgemein zu beschreiben.

Eine stetige komplexwertige Funktion ϕ auf einer topologischen Gruppe G wird eine positivdefinite Funktion genannt, wenn $\sum_{gh} \bar{\xi}(h) \cdot \phi(h^{-1} \circ g) \cdot \xi(g) \geq 0$ für jede Funktion auf G $\xi(\cdot)$, die nur an endlich vielen Stellen $\neq 0$ ist. Die normierten $(\phi(e) = 1)$ positivdefiniten Funktionen bilden eine konvexe Menge der speziellen Art, welcher man Simplex nennt. Die Charaktere sind spezielle positivdefinite Funktionen; sie sind (in Fällen wie den unseren) die Extremalpunkte der konvexen Menge. Die positivdefiniten Funktionen sind (untechnisch gesprochen) die 'konvexen Mischungen' der Charaktere.

2.5.3 Kurzschwänzige W-Verteilungen: Momente und Kumulanten

Für viele W-Maße kann man aus dem Verhalten der charakterischen Funktion in der Nähe des Nullpunkts bedeutsame Einsichten über das Maß gewinnen. Das ist besonders deutlich bei den W-Maßen, die auf eine beschränkte Menge des $\mathbb{R}^p$ konzentriert sind. Die Forderung des kompakten Trägers kann aber auch abgeschwächt werden.

Sprechweise (Momenten- und kumulantenerzeugende Funktion)**.**
Wir sagen von einem W-Maß $d\mu(x)$, dass es kurze Schwänze hat, wenn eine Umgebung V des Nullpunkts existiert, sodass

$$M(\theta) = \int e^{\langle\theta,\ x\rangle}\ d\mu(x)\ < \infty \quad \text{für} \quad \theta \in V.$$

Die Funktion $M(\cdot)$ heisst die momentenerzeugende Funktion, ihr Logarithmus

$$\psi(\theta) = \ln M(\theta) \quad \text{und} \quad = \infty \text{ für} \quad \theta \notin V$$

heisst die kumulantenerzeugende Funktion. Die konvexe Menge V heisst der Endlichkeitsbereich für $d\mu(x)$.

Beispiel. Wir berechnen die kumulantenerzeugende Funktion $\psi_\alpha(\theta)$ zur Gammaverteilung zum Parameter $\alpha > 0$.

$$\int_0^\infty e^{\theta x}\frac{1}{\gamma(\alpha)}x^{\alpha-1}\,e^{-x}\ dx = (\theta-1)^{-\alpha} \quad \text{für } \theta < 1$$

Die Konvexität des Endlichheitsbereichs V einer kumulantenerzeugenden Funktion ist klar, wenn man bedenkt, dass $M(\cdot)$ eine konvexe Überlagerung der konvexen Funktionen $k_x(\cdot) = \exp(\langle\cdot\ ,\ x\rangle)$ ist. Wir werden später sehen, dass die momentenereugenden Funktionen nicht nur konvex, sondern sogar logarithmisch konvex sind.

Man kann die momentenerzeugende Funktion auf einen Bereich im $\mathbb{C}^n$ ausdehnen, und damit einen Zusammenhang herstellen zu den charakterischen Funktionen einer Schar von Verteilungen $\big\{\mu_\theta :\ \theta \in \mathrm{V}\big\}$.

Sprechweise. Wenn $d\mu(x)$ ein kurzschwänziges W-Maß ist und V der Endlichkeitsbereich der momentenerrzeugenden Funktion $M(\cdot) = \exp\big(\psi(\cdot)\big)$, dann heissen die W-Maße

$$d\mu_\theta(x) = e^{\langle\theta,\ x\rangle}\ d\mu(x)\cdot e^{-\psi(\theta)}, \quad \text{für } \theta \in V$$

die getilteten Verteilungen.

Wir stellen fest, dass

$$\varphi_\theta(t) = \frac{1}{M(\theta)}\cdot M(\theta+it) = \int e^{\langle\theta+it\ ,\ x\rangle}\ d\mu(x)\cdot e^{-\psi(\theta)}$$

die charakteristische Funktion des getilteten W-Maßes μ_θ ist.

Konzentrieren wir uns jetzt zunächst auf den eindimensionalen Fall. Die momentenerzeugende Funktion $M(z) = \int e^{zx}\, d\mu(x)$ ist holomorph im Streifen $\{z : \Re z \in V^o\}$, wo V^o das Innere des Endlichkeitsbereichs bezeichnet. Sie und ihr Logarithmus besitzen Potenzreihendarstellungen in einem Kreis um den Nullpunkt

$$M(z) = 1+\alpha_1 z + \tfrac{1}{2!}\alpha_2 z^2 + \cdots = \sum_0^\infty \tfrac{1}{k!}\alpha_k z^k \qquad \text{für } |z| < R$$

$$\psi(z) = \ln M(z) = \kappa_1 z + \tfrac{1}{2!}\kappa_2 z^2 + \cdots = \sum_1^\infty \tfrac{1}{k!}\kappa_k z^k \qquad \text{für } |z| < r.$$

Der Koeffizient α_k heisst das $k$te Moment des W-Maßes μ, κ_k heisst der k-te Kumulant. Die Zahl $\alpha_1 = \kappa_1$ heisst der Erwartungswert, die Zahl $\kappa_2 = \alpha_2 - \alpha_1^2$ heisst die Varianz. Eine Interpretation der höheren Momente ergibt sich aus der Darstellung

$$M(z) = \int \Big(\sum \tfrac{1}{k!}(zx)^k\Big)\, d\mu(x) = \sum \tfrac{1}{k!}\Big(\int x^k\, d\mu(x)\Big)\cdot z^k.$$

Hinweis: Man hat sich früher tiefsinnige Gedanken gemacht, wie man einer Folge $(\alpha_k)_{k\in\mathbb{N}}$ ansehen kann, ob sie die Folge der Momente einer W-Maßes ist. Das Stichwort lautet Hausdorffs Momentenproblem. Es hat sich als bequemer erwiesen, die charakteristischen Funktionen (als Funktionen auf $\mathbb{R}$) zu betrachten.

$$\varphi(t) = M(it) = \sum \tfrac{1}{k!}\alpha_k (it)^k \qquad \text{für kleine } |t|.$$

Die Potenzreihendarstellung um den Nullpunkt hat nicht immer den gewünschten Aussagewert. Dagegen hat die konvexe Funktion $\psi(\theta) = \ln M(\theta)$ in neuerer Zeit im Zusammenhang mit der Theorie der sog. großen Abweichungen verstärkte Aufmerksamkeit gefunden.

Satz 2.5.22 (Logarithmische Konvexität der momentenerzeugenden Funktion).
Die kumulantenerzeugende Funktion ψ einer kurzschwänzigen Verteilung im $\mathbb{R}^n$ ist konvex und im Inneren des Endlichkeitsbereichs unendlich oft differenzierbar mit positiv definiter Hesse-Matrix $\psi''(\theta)$.

Beweis. *Wir zeigen für* $\theta_1, \theta_2 \in V$, $\lambda \in (0,1)$

$$M\big((1-\lambda)\theta_1 + \lambda\theta_2\big) \le \big(M(\theta_1)\big)^{1-\lambda} \cdot \big(M(\theta_2)\big)^{\lambda}.$$

Die Formel erscheint vertrauter, wenn wir die Notation verändern: $1-\lambda = \frac{1}{p}, \lambda = \frac{1}{q}$,

$$a(x) = \frac{1}{M(\theta_1)} \cdot \exp(\langle\theta_1, x\rangle), \qquad b(x) = \frac{1}{M(\theta_2)} \cdot \exp(\langle\theta_2, x\rangle).$$

Wir benützen die bekannte Ungleichung $a^{1/p} \cdot b^{1/q} \leq \frac{1}{p}a + \frac{1}{q}b$ *in allen Positionen* x.

$$\int \frac{1}{p} a(x)\ d\mu(x) + \int \frac{1}{q} b(x)\ d\mu(x) = 1,$$

$$1 \geq \int a^{1/p}(x) \cdot b^{1/q}(x)\ d\mu(x) = \Big[M(\theta_1)^{1/p} \cdot M(\theta_2)^{1/q}\Big]^{-1} \cdot \int \exp(\langle \frac{1}{p}\theta_1 + \frac{1}{q}\theta_2, x\rangle)\ d\mu(x).$$

Beispiel. 1) Die Funktion $M(\alpha) = \Gamma(\alpha + 1)$ ist eine momentenerzeugende Funktion mit dem Endlichkeitsbereich $V = (-1, \infty)$. Es gilt nämlich $M(0) = 1$ und

$$M(\alpha) = \int_0^\infty e^{\alpha \ln x} \cdot e^{-x}\ dx = \int_{-\infty}^\infty e^{\alpha y} \cdot \exp(-e^y) \cdot e^y\ dy.$$

2) Die n-dimensionale Normalverteilung mit der Covarianzmatrix $C = Q^{-1}$ ist kurzschwänzig. Ihre kumulantenerzeugende Funktion ist die quadratische Form $\psi(\theta) = \frac{1}{2}\theta \cdot C \cdot \theta^T$. In der Tat gilt

$$\big(\tfrac{1}{\sqrt{2\pi}}\big)^n \int \exp(\theta \cdot x) \cdot \exp\big(-\tfrac{1}{2}x^T \cdot Q \cdot x\big)\sqrt{\det Q}\ dx = \exp\big(\tfrac{1}{2}\theta \cdot C \cdot \theta^T\big).$$

Der Schwerpunkt eines W-Maßes: Die Momente eines W-Maßes im $\mathbb{R}^m$ sind auch dann eng verbunden mit den Ableitungen der charakteristischen Funktion im Nullpunkt, wenn die Schwänze nicht exponentiell abfallen.

Sprechweise. Man sagt von einem W-Maß μ auf einem m-dimensionalen reellen Vektorraum, dass es einen Schwerpunkt besitzt, wenn für jede Linearform $\ell(\cdot)$ das Integral $\int \ell(x)\ d\mu(x)$ existiert. Der Schwerpunkt m ist dadurch gekennzeichnet, dass für jede Linearform gilt

$$\langle \ell, m\rangle = \int \ell(x)\ d\mu(x)$$

Man sagt, dass es zweite Momente besitzt, wenn für alle $\ell(\cdot)$ das Integral $\int (\ell(x))^2\ d\mu(x)$ existiert.

Eine wichtige Aussage über W-Maße mit Schwerpunkt ist der

Satz 2.5.23 (Jensen's Ungleichung)**.** *Wenn* m *der Schwerpunkt des W-Maßes* μ *ist, dann gilt für jede konvexe Funktion* $k(\cdot)$

$$k(m) = k\Big(\int x\ d\mu(x)\Big) \leq \int k(x)\ d\mu(x).$$

Beweis. *Die konvexe Funktion* k *besitzt eine affine Minorante. Das Integral von* k *kann* $+\infty$ *sein; in diesem erweiterten Sinn existiert es aber jedenfalls. Für jede affine Minorante* $\ell \leq k$ *gilt*

$$\ell(m) = \int \ell(x)\ d\mu(x) \leq \int k(x)\ d\mu(x).$$

Wenn k *unterhalbstetig und konvex ist (und das ist der interessante Fall), dann gilt*

$$k(m) = \sup\{\ell(m) : \ell \textit{ ist affine Minorante}\} \qquad \textit{also} \quad k(m) \le \int k(x)\ d\mu(x).$$

Für den allgemeinen Fall können wir annehmen, dass μ *nicht auf einen echten affinen Unterraum konzentriert ist. Der Schwerpunkt liegt dann im Inneren des Endlichkeitsbereich. Die Jensen'sche Ungleichung ergibt sich.*

Den Schwerpunkt eines W-Maßes, wenn er denn existiert, kann von der charakteristischen Funktion abgelesen werden. Wenn alles genügend regulär zugeht, dann kann man bekanntlich die Differentiation mit der Integration vertauschen, und man erhält dann

$$\varphi(t) = \int e^{i\langle t,\ x\rangle}\ d\mu(x); \qquad \varphi'(t) = i \cdot \int e^{i\langle t,\ x\rangle} \cdot x\ d\mu(x).$$

Der folgende Satz präzisiert die Idee. Es genügt, den eindimensionalen Fall ($m = 1$) zu betrachten.

Satz 2.5.24. *Es sei* $d\mu(x)$ *ein W-Maß auf* $\mathbb{R}$ *mit* $M_n = \int |x|^n\ d\mu(x) < \infty$. *Die charakteristische Funktion ist dann* n*-mal stetig differenzierbar mit der* n*-ten Ableitung*

$$\varphi^{(n)}(t) = (i)^n \cdot \int e^{itx} \cdot x^n\ d\mu(x).$$

Mit den Koeffizienten $\quad \alpha_k = \int x^k\ d\mu(x)$ *für* $k = 1, \dots n$ $\quad$ *ergibt sich*

$$\left|\varphi(t) - 1 - it\alpha_1 - \tfrac{1}{2!}(it)^2\alpha_2 - \cdots - \tfrac{1}{(n-1)!}(it)^{n-1}\alpha_{n-1}\right| \le \tfrac{1}{n!}t^n M_n$$

Beweis. *Wir beweisen das durch vollständige Induktion nach* n. *Für* $n = 1$ *haben wir*

$$\begin{aligned}\frac{1}{ih}\big[\varphi(t+h) - \varphi(t)\big] &= \int e^{itx}\frac{e^{ihx}-1}{ih}\ d\mu(x)\\ &= \int e^{itx} \cdot x\ d\mu(x) \quad + \quad \int e^{itx}\Big[\frac{e^{ihx}-1}{ihx} - 1\Big] \cdot x\ d\mu(x).\end{aligned}$$

Das zweite Integral strebt nach 0 *(nach dem Satz von der majorisierten Konvergenz); die Konvergenz ist sogar gleichmäßig in* t. *Ebenso funktioniert der Schritt von* $n-1$ *nach* n.

$$\begin{aligned}\frac{1}{ih}\big[\varphi^{(n-1)}(t+h) - \varphi^{(n-1)}(t)\big] &= \int e^{itx}\frac{e^{ihx}-1}{ih} \cdot x^{n-1}\ d\mu(x)\\ &= \int e^{itx} \cdot x^n\ d\mu(x) \quad + \quad \int e^{itx}\Big[\frac{e^{ihx}-1}{ihx} - 1\Big] \cdot x^n\ d\mu(x)\end{aligned}$$

Für die zweite Aussage benötigen wir eine Abschätzung der elementaren Funktionen

$$R_{n-1}(u) \ = \ e^{iu} - 1 - iu - \tfrac{1}{2!}(iu)^2 - \cdots - \tfrac{1}{(n-1)!}(iu)^{n-1}$$

Es gilt $\quad R_0(u) \ = \ e^{iu} - 1 \ = \ \frac{1}{i}\int_0^u e^{iv}\ dv \quad$ *und allgemein* $\quad R_n(u) \ = \ \frac{1}{i}\int_0^u R_{n-1}(v)\ dv$.
Durch vollständige Induktion gewinnt man $\quad \left|R_{n-1}(u)\right| \le \frac{1}{n!}|u|^n$.
Der Ausdruck, den es abzuschätzen gilt, ist $\quad \int R_{n-1}(tx)\ d\mu(x)$.

Wenn der Schwerpunkt existiert, dann ist der nächstwichtige Fall ist der Fall der zweiten Ableitung.Es zeigt sich, dass die Existenz des zweiten Moment nicht nur hinreichend, sondern auch notwendig ist für die Existenz der zweiten Ableitung der charakteristischen Funktion. Die Hesse-Matrix der charakteristischen Funktion nennt man die Covarianzmatrix des W-Maßes. Wir behandeln den eindimensionalen Fall: Wenn das zweite Moment unendlich ist, dann ist die charakteristische Funktion im Nullpunkt nicht zweimal differenzierbar. Für den Realteil $u(h) = \Re\varphi(h) = \int \cos(hx)\ d\mu(x)$ gilt dann nämlich für $h \to 0$

$$\frac{1-u(h)}{h^2} = \int \frac{1-\cos\ hx}{(hx)^2} \cdot x^2\ d\mu(x) \longrightarrow \infty$$

Ein Beispiel, wo nicht einmal das erste absolute Moment existiert, ist die Cauchy-Verteilung: Die charakteristische Funktion $\exp(-|t|)$ ist im Nullpunkt nicht differenzierbar.

Im m-dimensionalen Fall mit $\int \|x\|^2\ d\mu(x) < \infty$ und $m = \frac{1}{i}\varphi'(0)$gilt

$$\varphi(t) = 1 + itm - \frac{1}{2!} t \cdot M_2 \cdot t^T + o(\|t\|^2) \quad \text{für } t \to 0$$

$$\text{wobei} \qquad M_2 = \varphi''(0) = \int x \cdot x^T\ d\mu(x).$$

Wenn wir das Maß so verschieben, dass sein Schwerpunkt in den Nullpunkt zu liegen kommt, dann bedeutet das für die charakteristische Funktion die Multiplikation mit e^{-imt} und wir erhalten

$$\varphi(t) \cdot e^{-itm} = 1 - \frac{1}{2!}\, t \cdot C \cdot t^T + o(\|t\|^2) \quad \text{wo}$$

$$C = \int (x-m) \cdot (x-m)^T\ d\mu(x) \quad \text{die sog. Covarianzmatrix ist}$$

$$\ln\varphi(t) - itm \ = \ -\frac{1}{2!}\, t \cdot C \cdot t^T + o(\|t\|^2)$$

Für die kumulantenerzeugende Funktion einer wirklich kurzschwänzigen Verteilung lautet die entsprechende Formel

$$\ln M(\theta + it) = \psi(\theta + it) = it\psi'(\theta) - \frac{1}{2}\ t \cdot \psi''(\theta) \cdot t^T + o(\|t\|^2).$$

wo $\psi'(\theta)$ der Erwartungswert und $\psi''(\theta)$ die Kovarianzmatrix des Maßes $d\mu_\theta$ ist.

$$\psi'(\theta) = \int x\ d\mu_\theta(x); \qquad \psi''(\theta) = \int \big(x - \psi'(\theta)\big) \cdot \big(x - \psi'(\theta)\big)^T\ d\mu_\theta(x).$$

Die Formeln für $\psi'(\theta)$ und $\psi''(\theta)$ kann man auch durch Differentiation unter dem Integralzeichen gewinnen; es gibt keine Schwierigkeiten mit der Regularität, weil es ja schliesslich um (im Streifen) holomorphe Funktionen geht. Allenfalls die Matrix-Notation ist nachdenkenswert.

$$\psi' = \frac{M'}{M} = \int e^{\theta x}\, x\, d\mu(x) \cdot e^{-\psi(\theta)} = \int x\, d\mu_\theta(x),$$

$$\psi'' = \frac{M''}{M} - \left(\frac{M'}{M}\right) \cdot \left(\frac{M'}{M}\right)^T =$$

$$= \int e^{\theta x}\, x \cdot x^T\, d\mu(x) \cdot e^{-\psi(\theta)} - \left(\int e^{\theta x}\, x\, d\mu(x) \cdot e^{-\psi(\theta)}\right) \cdot \left(\int e^{\theta x}\, x\, d\mu(x) \cdot e^{-\psi(\theta)}\right)^T =$$

$$= \int x \cdot x^T\, d\mu_\theta(x) - \left(\int x\, d\mu_\theta(x)\right) \cdot \left(\int x\, d\mu_\theta(x)\right)^T =$$

$$= \int \big(x - \psi'(\theta)\big) \cdot \big(x - \psi'(\theta)\big)^T\, d\mu_\theta(x).$$

So ist also $\psi'(\theta)$ der Schwerpunkt und $\psi''(\theta)$ die Covarianzmatrix des getilteten Maßes.

2.5.4 Fourier-Integrale und Fourier-Transformation

Vorbemerkung und Ausblick:
Die Gruppe $G = \mathbb{R}/2\pi$ trägt ein translationsinvariantes W-Maß,; man bezeichnet es üblicherweise mit $d\lambda(t) = \frac{1}{2\pi}\,dt$. Für eine integrable Funktion $f(t)$ definiert man die Folge der Fourier-Koeffizienten $c_n = \frac{1}{2\pi}\int e^{-int}\,dt$. Man versteht diese Folge als eine Funktion $c(\cdot) = \hat{f}(\cdot)$ auf der 'dualen Gruppe' $G^* = \mathbb{Z}$, die man mit dem Zählmaß als translationsinvariantem Maß ausstattet.

Wenn $f(t)$ sogar quadratintegrabel ist, dann ist die 'Fouriertransformierte' $\hat{f}(\cdot)$ quadratsummabel mit $\sum |\hat{f}(n)|^2 = \frac{1}{2\pi}\int |f(t)|^2\,dt$. Wenn $c(\cdot) = \hat{f}(\cdot)$ ausserdem summabel ist, dann gilt $f(t) = \sum c_n \cdot e^{int}$. In diesem Fall ist f nicht nur quadratintegrabel, sondern darüberhinaus stetig auf $\mathbb{R}/2\pi$. Auf der anderen Seite gibt es keinen Grund anzunehmen, dass eine stetige 2π-periodische Funktion eine summable Folge von Fourier-Koeffizienten besitzt. Die Sache ist etwas verwickelter.

Befriedigend ist die isometrische Entsprechung zwischen $l^2(\mathbb{Z})$ und $L^2\big(\mathbb{R}/2\pi, d\lambda(t)\big)$, die auf dichtliegenden Teilvektorräumen direkt durch Integrationen bzw. Summationen zu berechnen ist, und dann auf die vervollständigten Hilbert-Räume fortzusetzen ist.

Es ist lehrreich, die Konstruktion beim normierten Vektorraum $\big(V, \|\cdot\|_2\big)$ der trigonometrischen Polynomen zu beginnen. Die Vektoren v werden einerseits durch 2π-periodische Funktionen dargestellt und andererseits durch finite Folgen repräsentiert. Die Elemente der Vervollständigung $\big(\bar{V}, \|\cdot\|\big)$ sind bekanntlich zunächst einmal Äquivalenzklassen von Cauchy-Folgen; sie können durch quadratsummable Folgen dargestellt werden, wie wir bewiesen haben. Es gibt aber auch noch andere Darstellungen. Auf der Grundlage der Integrationstheorie von Lebesgue haben E. Fischer und F. Riesz 1906 (unabhängig voneinander) gezeigt, dass man die Elemente der Vervollständigung auch durch Äquivalenzklassen quadratintegrierbarer repräsentieren kann.

$$c(\cdot) \in l^2 \longleftrightarrow v \in \bar{V} \longleftrightarrow f(\cdot) \in L^2 \qquad \text{mit} \qquad \sum |c_n|^2 = \|v\|^2 = \frac{1}{2\pi}\int |f(t)|^2\,dt.$$

In dieser Betrachtungsweise ähnelt die Fourier-Transformation $f(\cdot) \leftrightarrow \hat{f}(\cdot)$ einem Koordinatenwechsel. Allerdings hat man nur im Darstellungsraum l^2 eine Orthonormalbasis $\mathcal{B}$ ausgezeichnet. Mit diesen Basisvektoren hat es eine besondere Bewandtnis: sie sind die normierten Eigenvektoren für die Verschiebungen der Gruppe $\mathbb{R}/2\pi$: $U_s : f(\cdot) \mapsto f(\cdot + s)$. In der Tat gilt

$$U_s v_n = U_s(e^{in(\cdot)}) = e^{in(\cdot+s)} = e^{ins} \cdot e^{in(\cdot)} = \lambda_n \cdot v_n.$$

Auf der anderen Seite haben wir die Gruppe der Verschiebungen des Raums $l^2(\mathbb{Z})$. Sie wirken in der Darstellung der v als Multiplikationsoperatoren: $V_n : f(\cdot) \mapsto e^{in(\cdot)} \cdot f(\cdot)$. Zu diesen Operatoren gibt es keine Eigenvektoren; allenfalls könnte man Funktionen, die ausserhalb einer kleinen Umgebung von t_0 verschwinden, approximative Eigenvektoren (zum Eigenwert e^{int_0}) nennen. Für Situationen dieser Art hat man in der Hilbertraum-Theorie den Begriff der kontinuierlichen Spektralzerlegung entwickelt. Wenn man die Sachlage so auffasst, dann erscheinen die Folgen $c(\cdot)$ und die Äquivalenzklassen von Funktionen $f(\cdot)$ als die 'Amplitudenfunktionen' desselben Vektors in verschiedenen Spektralzerlegungen des abstrakten Hilbert-Raums $\big(\bar{V}, \|\cdot\|\big)$

Das spezielle zur Fourier-Transformation gehörende Paar von Spektralzerlegungen spielt eine besondere Rolle in Anbetracht der Gruppenstruktur in $\mathbb{R}/2\pi$ sowie der Gruppenstruktur in der 'dualen' Gruppe $\mathbb{Z}$. Analoge Betrachtungen kann man auch anstellen für das Paar der Gruppen $\mathbb{R}^p_Z$ und $\mathbb{R}^p_{Sp}$. Die Betrachtungsweise spielt eine wichtige Rolle für die Mathematik der Quantenmechanik; mit der Fourier-Transformation beschreibt man dort den Übergang von den Ortskoordinaten zu den Impulskoordinaten eines quantenmechanischen Zustands.

Wir wollen die Ideen der Basistransformation für einen abstrakten Hilbertraums hier nicht weiter verfolgen. Stattdessen werden wir im Folgenden die Konstruktion der Fourier-Integrale und Fourier-Transformationen in einem elementaren Ansatz in Angriff nehmen. Die Überlegungen erscheinen daher eher als Übungen in elementarer Integrationstheorie.

Fourier-Integrale

Definition 2.20. Für eine komplexwertige Lebesgue-integrable Funktion $f(t)$ auf dem Raum der reellen p-Zeilen $\mathbb{R}^p_Z$ definiert man das Fourier-Integral

$$\hat{f}(x) = \left(\tfrac{1}{2\pi}\right)^p \int e^{-it\cdot x} f(t)\ dt \quad \text{für} \quad x \in \mathbb{R}^p_{Sp}.$$

Die Fourier-Integrale $\hat{f}(\cdot)$ sind offenbar beschränkte gleichmäßig stetige Funktionen. Weiterhin gilt

Satz ('Lemma von Riemann-Lebesgue')**.**
Jedes Fourier-Integral verschwindet im Unendlichen: $\lim_{\|x\|\to\infty} \hat{f}(x) = 0.$

Beweis. *Es genügt, den Satz für die Indikatorfunktion eines Rechtecks* $f(t) = 1_R(t) = 1_{[a_1,b_1)}(t_1)\cdot\dots\cdot 1_{[a_p,b_p)}(t_p)$ *zu beweisen; denn jede integrable Funktion kann in der* L^1*-Norm durch Linearkombinationen solcher Funktionen approximiert werden; und für ein* $r(t)$ *mit* $\int |r(t)|\ dt \le \varepsilon$ *gilt* $\sup_x |\hat{r}(x)| \le \left(\frac{1}{2\pi}\right)^p \varepsilon$. *Für* $a < b$ *und* $c = \frac{1}{2}(a+b)$, $h = \frac{1}{2}(b-a)$ *gilt*

$$\frac{1}{2\pi}\int e^{-itx} 1_{[a,b)}(t)\ dt = \frac{1}{2\pi} e^{-icx} \int_{-h}^{h} e^{-iux}\ du = \frac{1}{2\pi} e^{-icx}\cdot 2h \cdot \frac{\sin(hx)}{hx}.$$

Diese Funktionen konvergieren nach 0 für $x \to \pm\infty$. *Man bemerke aber, dass nicht schnell genug abfallen, um Lebesgue-integrabel zu sein.*

Sprechweise 2.5.1. Wir sagen eine Funktion $f(t)$ sei direkt Fourier-invertierbar, wenn sie und ihr Fourier-Integral $\hat{f}(x)$ integrabel sind und ausserdem

$$f(t) = \int e^{it\cdot x} \hat{f}(x)\ dx \qquad \text{für fast alle } t \in \mathbb{R}^p_Z.$$

Wir werden sehen, dass die Integrabilitätsbedingungen bereits hinreichend sind. Der Beweis braucht aber einige Vorbereitungen. Zuerst verweisen wir auf Beispiele, für die wir schon früher die Rechnungen durchgeführt haben.

Beispiel. 1. $f(t) = e^{-|t|}$, $\hat{f}(x) = \frac{1}{\pi} \cdot \frac{1}{1+x^2}$; $\int e^{itx} \frac{1}{\pi} \cdot \frac{1}{1+x^2}\, dx = e^{-|t|}$;

2. $f(t) = \left(\frac{\sin t}{t}\right)^2$, $\hat{f}(x) = \frac{1}{2\pi} \int_{-\infty}^{\infty} e^{-itx} \left(\frac{\sin t}{t}\right)^2\, dt = \frac{1}{2}\left(1 - \frac{|x|}{2}\right)^+$
$\left(\frac{\sin t}{t}\right)^2 = \int e^{itx} \frac{1}{2}\left(1 - \frac{|x|}{2}\right)^+\, dx.$

3. $g_\varepsilon(t) = \exp\left(-\frac{\varepsilon^2}{2}(t_1^2 + \cdots + t_p^2)\right)$; $\hat{g}_\varepsilon(x) = \left(\frac{1}{\sqrt{2\pi}\varepsilon}\right)^p \cdot \exp\left(-\frac{1}{2\varepsilon^2}(x_1^2 + \cdots + x_p^2)\right)$.

Satz 2.5.25 (Der Umkehrsatz für Fourier-Integrale)**.**
Wenn $f(t)$ und $\hat{f}(x)$ integrabel sind, dann ist $f(\cdot)$ direkt Fourier- invertierbar, d. h.

$$\hat{f}(x) = \left(\tfrac{1}{2\pi}\right)^p \int e^{-it\cdot x} f(t)\, dt; \qquad f(t) = \int e^{it\cdot x} \hat{f}(x)\, dx.$$

Beweis. *Wir beweisen das zunächst für geeignet geglättete $h(\cdot)$. Für ein direkt Fourier-invertierbares $g(\cdot)$ sei $h(t) = \int g(t-s)\cdot f(s)$. Nach der Produktformel für die Faltung*

$$\begin{aligned}
\hat{h}(x) &= \left(\tfrac{1}{2\pi}\right)^p \int e^{-it\cdot x} h(t)\, dt = \left(\tfrac{1}{2\pi}\right)^p \int e^{-is\cdot x} g(s)\, dt \cdot \int e^{-is\cdot x} f(s)\, ds \\
&= \hat{g}(x) \cdot \int e^{-is\cdot x} f(s)\, ds \qquad \textit{und daher} \\
\int e^{it\cdot x} \hat{h}(x)\, dx &= \int e^{it\cdot x} \hat{g}(x) \cdot \int e^{-is\cdot x} f(s)\, ds\, dx \\
&= \int \left(\int e^{i(t-s)\cdot x} \hat{g}(x)\, dx \right) \cdot f(s)\, ds \\
&= \int \quad g(t-s) \cdot f(s)\, ds = \quad h(t).
\end{aligned}$$

Wichtig ist hier die Produktintegrabilität von $\hat{g}(x) \cdot f(s)$.

Für $g(t)\, dt = \hat{g}_\varepsilon(t)\, dt$, dem Gausskern zur Varianz ε^2 haben wir eine passende Glättung. Für $f_\varepsilon = f * \hat{g}_\varepsilon$ erhalten wir

$$\begin{aligned}
\hat{f}_\varepsilon(x) &= \quad \hat{f}(x) \cdot \exp\left(-\frac{\varepsilon^2}{2}\|x\|^2\right), \\
f_\varepsilon(t) &= \quad \int e^{it\cdot x} \hat{f}(x) \cdot \exp\left(-\frac{\varepsilon^2}{2}\|x\|^2\right)\, dx.
\end{aligned}$$

Die linke Seite konvergiert nach dem Satz von der L^1-Stetigkeit der Verschiebung in der L^1-Norm nach f; die rechte konvergiert (nach dem Satz von der majorisierten Konvergenz wegen der Integrabilität von $\hat{f}$) gegen die stetige Funktion $\int e^{it\cdot x} \hat{f}(x)\, dx$.

Satz 2.5.26 (L^2-Isometrie zum Fourier-Integral).
Es sei $f(t)$ *eine direkt Fourier-invertierbare Funktion ist und* $\hat{f}(x)$ *ihr Fourier-Integral.*

$$f(t) = \int e^{it\cdot x}\hat{f}(x)\ dx; \qquad \hat{f}(x) = \left(\tfrac{1}{2\pi}\right)^p \int e^{-it\cdot x} f(t)\ dt.$$

Die Funktionen f *und* $\hat{f}$ *sind dann quadratintegrabel, und es gilt*

$$\left(\tfrac{1}{2\pi}\right)^p \int \left|f(t)\right|^2\ dt \quad = \quad \int \left|\hat{f}(x)\right|^2\ dx.$$

Beweis. *Aus typographischen Gründen bezeichnen wir die zu* f *komplex konjugierte Funktion mit* f^* *statt mit* $\bar{f}$. *Wir haben also* $|f|^2 = f^* \cdot f$.

Funktionen, die direkt Fourier-invertierbar sind, sind beschränkt, gleichmäßig stetig und integrabel, und deshalb auch quadratintegrabel.

Wir zeigen die (auf den ersten Blick) etwas allgemeinere Gleichheit

$$\left(\tfrac{1}{2\pi}\right)^p \int f^*(t)\cdot h(t)\ dt = \int \hat{f}(x)^* \cdot \hat{h}(x)\ dx.$$

für direkt Fourier-invertierbare f, h. *Der Satz von Fubini kann angewendet werden auf das Doppelintegral*

$$\left(\tfrac{1}{2\pi}\right)^p \int f^*(t)\cdot e^{it\cdot x}\cdot \hat{h}(x)\ dx\ dt.$$

Wenn wir zuerst nach x *integrieren erhalten wir den Ausdruck auf der linken Seite der behaupteten Gleichheit. Wenn wir zuerst nach* t *integrieren, dann müssen wir beachten, dass* $\left(\frac{1}{2\pi}\right)^p \int f^*(t)\cdot e^{it\cdot x}\ dt$ *die komples konjuierte Funktion zur Fourier-Tranformierten von* f *ist.*

Es erscheint lehrreich, wie oben im Beweis auch ein geglättetes $h(t) = \int g(t-s)\cdot k(s) ds$ *studieren. Zuerst beachten wir* $\hat{h}(x) = \hat{g}(x)\cdot\left(\int e^{-isx}k(s)\ ds\right)$. *Sodann ergibt sich*

$$\begin{aligned}\int \hat{f}^*(x)\cdot\hat{h}(x)\ dx &= \left(\tfrac{1}{2\pi}\right)^p \int \left(\int e^{itx} f^*(t)\ dt\right)\cdot \hat{g}(x)\cdot\left(\int e^{-isx}k(s)\ ds\right)\ dx = \\ &= \left(\tfrac{1}{2\pi}\right)^p \int f^*(t)\cdot\left(\int e^{i(t-s)x}\hat{g}(x)\ dx\right)\cdot k(s)\ dt\ ds = \\ &= \left(\tfrac{1}{2\pi}\right)^p \int f^*(t)\cdot g(t-s)\cdot k(s)\ ds\ dt = \left(\tfrac{1}{2\pi}\right)^p \int f^*(t)\cdot h(t)\ dt.\end{aligned}$$

Auch hier ist der Satz von Fubini anwendbar, weil die Funktion $f^*(t)\cdot\hat{g}(x)\cdot k(s)$ *integrabel ist bzgl. des Produktmaßes* $dt \otimes\ dx \otimes\ ds$.

Betrachten wir nun die spezielle Glättung mit den Gauss-Dichten $\hat{g}_\varepsilon(t)$ *Für die geglättete Dichte* h_ε *haben wir dann* $\quad \hat{h}_\varepsilon(x) = \hat{k}(x)\cdot\exp\left(-\frac{\varepsilon^2}{2}\|x\|^2\right)$.

Da $\widehat{f}, \widehat{h}$ *quadratintegrabel sind, ist das Produkt (nach der Hölder'schen Ungleichung) integrabel; und nach dem Satz von der majorisierten Konvergenz konvergiert die linke Seite (für* $\varepsilon \to 0$*) gegen* $\int \widehat{f}^*(x) \cdot \widehat{k}(x)\ dx$.

Auf der anderen Seite konvergieren die geglätteten h_ε *in der* L^1*-Norm gegen* k*; und die rechte Seite strebt gegen* $\left(\frac{1}{2\pi}\right)^p \int f^*(t) \cdot k(t)\ dt$. *Wenn wir die Formel auf den Spezialfall* $f = k$ *anwenden, ergibt sich, dass aus Quadratintegrabilität von* $\widehat{f}$ *die Quadratintegrabilität von* f *folgt und umgekehrt.*

Definition 2.21 (Fourier-Transformation)**.** Die durch das Fourier-Integral gegebene Abbildung wird eingeschränkt auf den Teilraum der quadratintegrablen direkt Fourier-invertierbaren Funktionen, und dann stetig fortgesetzt auf den Raum aller quadratintegrablen f. Diese Abbildung $\mathcal{F}$ heisst die Fourier-Transformation.

Die Fourier-Transformation ist, wie wir eben gesehen haben eine surjektive Abbildung auf den Raum $L^2\big(\mathbb{R}^p_{Sp}\big)$. Sie ist eine Isometrie in dem Sinn: Wenn $\mathcal{F}(f) = \widehat{f}$, dann gilt

$$\|f\|_2 \quad = \quad \sqrt{\left(\tfrac{1}{2\pi}\right)^p \int \left|f(t)\right|^2\ dt} \quad = \quad \sqrt{\int \left|\widehat{f}(x)\right|^2\ dx}..$$

Didaktischer Hinweis: Manche Lehrbücher finden es störend, dass auf dem Raum der Spalten eine andere Normierung des verschiebungsinvarianten Maßes vorgenommen werden muß als auf dem Raum der Zeilen. Sie führen daher einen Faktor ein in die Definition des Fourier-Integrals. Für die direkt Fourier-invertierbaren Funktionen in $L^1 \cap L^2$ erhalten sie dann das modifizierte Fourier-Integral und seine Umkehrung in der Form

$$\mathcal{F}_{\text{mod}}(f) = \left(\tfrac{1}{2\pi}\right)^{p/2} \int e^{-it\cdot x}\ f(t)\ dt : \qquad \mathcal{F}^{-1}_{\text{mod}}(\widehat{f}) = \left(\tfrac{1}{2\pi}\right)^{p/2} \int e^{it\cdot x}\ \widehat{f}(x)\ dx.$$

Andere Lehrbücher schätzen die vollständige Analogie zur Situation bei den Fourier-Reihen: Für die direkt Fourier-invertierbaren 2π-periodischen $f(t)$ und ihre Folgen der Fourier-Koeffizienten $\widehat{f}$ gilt dort bekanntlich

$$\widehat{f}(n) = \mathcal{F}(f) = \tfrac{1}{2\pi} \int e^{-int} f(t)\ dt; \qquad f(t) = \mathcal{F}^{-1}(\widehat{f}) = \sum \widehat{f}(n) e^{int}$$

$$\|\nu\|_2 \quad = \quad \sqrt{\left(\tfrac{1}{2\pi}\right) \int \left|f(t)\right|^2\ dt} \quad = \quad \sqrt{\sum |\widehat{f}(n)|^2}.$$

Das translationsinvariante W-Maß auf $\mathbb{R}/2\pi$ trifft sich hier mit dem Zählmaß auf $\mathbb{Z}$.

Eine physikalische Interpretation

Eine Funktion der Form $f(t; x, y, z) = A \cdot \exp\big(i(\omega t - k \cdot x)\big)$ heisst in der Theorie der Wellen die Amplitudenfunktion einer ebenen Welle mit der Kreisfrequenz ω und der Wellenzahl $k = (k_x, k_y, k_z)$. Die komplexe Zahl A heisst die komplexe Amplitude im

Nullpunkt. Setzen wir zunächst einmal $t = 0$. Eine Überlagerung von 'rein sinusförmigen' Amplituden heisst in diesem Zusammenhang ein Wellenpaket zur Zeit 0.

$$A_0(x) = \int \exp\bigl(-ik \cdot x\bigr) q(k)\ dk.$$

Dabei ist $q(\cdot)$ eine quadratintegrable Funktion, sodass auch A_0 quadratintegrabel ist - wobei man allerdings das Integral sum grano salis verstehen muss, wenn $q(\cdot)$ nicht auch integrabel ist. Man interessiert sich vor allem für solche Wellenpakete, wo die beiden W-Dichten $\text{const} \cdot \left|A_0(x)\right|^2\ dx$ und $\text{const} \cdot \left|q(k)\right|^2\ dk$ ('im Wesentlichen') auf einen nicht allzu großen Bereich konzentriert sind.

Ein Wellenpaket verändert sich in der Zeit nach einem Gesetz, welches sich aus der sog. Dispersionsrelation (im betreffenden Medium) ergibt. Die Dispersionrrelation gibt an, wie die Kreisfrequenz von der Wellenzahl abhängt, $\omega = \Omega(k)$.

In der Quantenmechanik des freien Teilchens erfährt das Tripel der Wellenzahlen die Interpretation als Impuls, die Kreisfrequenz versteht man als Energie. Dabei hat man eine Dimensionierung durch das Planck'sche Wirkungsquantum

$$E = \hbar\omega \qquad p = \hbar \cdot k.$$

Für die Amplitudenfunktion ergibt sich

$$A_t(x) = \int \exp\Bigl(\frac{i}{\hbar}(E(p)t - p \cdot x)\Bigr) Q(p)\ dp$$

mit der für das freie Teilchen gültigen Dispersionsrelation

$$E^2 - c^2 \cdot \|(p\|^2 = \text{const} = m^2 \cdot c^4$$

(wo c die Lichtgeschwindigkeit und m die Ruhemasse des Teilchens ist.)

Die Bewegung des Wellenpakets ist bestimmt durch die sog. Gruppengeschwindigkeit $v = \frac{d}{dp}E(p)$. Diese ist im konkreten Fall leicht zu berechnen: $v = \frac{1}{M}p$, wo $M = \sqrt{m^2 + \frac{1}{c^2}\|p\|^2}$, die bewegte Masse des Teilchens genannt wird.

'Gruppengeschwindigkeit'

Wir wollen noch kurz andeuten, was es mit der sog. Gruppengeschwindigkeit auf sich hat. Wir gehen aus von einer glatten quadratintegrablen Funktion $q(y) = |q(y)| \cdot e^{iT(y)}$ und nehmen an, dass $|q(y)|^2\ dy$ eine Wahrscheinlichkeitsdichte ist. Wir interessieren uns für die Funktion

$$\begin{aligned} A_t^{(\varepsilon)}(x)\ &:= \left(\tfrac{1}{\sqrt{2\pi\varepsilon}}\right)^p \int \exp\left(\tfrac{i}{\varepsilon}\bigl(\Omega(y)\cdot t - y \cdot x\bigr)\right) q(y)\ dy = \\ &= \left(\tfrac{1}{\sqrt{2\pi\varepsilon}}\right)^p \int \exp\left(\frac{i}{\varepsilon}K(y)\right) \cdot q(y)dy \\ &\text{mit}\ \ K(y) = K_{t,x}(y) = \Omega(y)\cdot t - y \cdot x\ . \end{aligned}$$

Mit der zweiten Schreibweise erinnern wir an die Überlegungen im Zusammenhang mit den Fresnel-Integralen. Wir haben dort argumentiert, dass (bei einem 'glatten' $q(\cdot)$) der Beitrag zum Integralwert $A_t^{(\varepsilon)}(x)$ im Limes $\varepsilon \to 0$ durch $q(Y(x))$ bestimmt ist, wenn es im Träger von $q(\cdot)$ genau einen Punkt $Y(x)$ gibt, in welchem $K'(\cdot)$ verschwindet. Im vorliegenden Fall ist $Y(x)$ die Lösung der Gleichung $\Omega'(Y(x)) = \frac{x}{t}$,. Das ist in der physikalischen Interpretation der Impuls zur Gruppengeschwindigkeit.

Wir wollen das hier noch etwas konkreter untersuchen. Wir schreiben $A_t^{(\varepsilon)}(x) = |a(x)| \cdot e^{iS(x)}$. Für jedes (t, ε) ist $|a(x)|^2\,dx$ eine Wahrscheinlichkeitsdichte im Raum der Spalten; denn wir können

$$A_t^{(\varepsilon)}(x) := \left(\frac{1}{\sqrt{2\pi\varepsilon}}\right)^p \int \exp(-\tfrac{i}{\varepsilon}\, y \cdot x) \cdot \exp\left(\tfrac{i}{\varepsilon}\, \Omega(y) \cdot t\right) q(y)\, dy$$

als ein Fourier-Integral (mit der passenden Normierung!) verstanden werden.

Beispiel 2.5.5. Bei sehr speziellen Annahmen über $q(y) \cdot e^{\frac{i}{\varepsilon}T(y)}$ können wir die transformierte Funktion $a(x) = |a(x)| \cdot e^{\frac{i}{\varepsilon}S(x)}$ tatsächlich explizit berechnen.

$$\begin{aligned} q(y) = |q(y)| \cdot e^{\frac{i}{\varepsilon}T(y)} = \text{const} \cdot \exp\left(-\tfrac{\sigma^2}{2}(y-\tilde{y})^2\right) \cdot \exp\left(\tfrac{i}{\varepsilon}(y-\tilde{y}) \cdot \tilde{x}\right) \quad &\Longrightarrow \\ \left(\frac{1}{\sqrt{2\pi\varepsilon}}\right)^p \int \exp(-\tfrac{i}{\varepsilon}\, y \cdot x) q(y)\, dy \quad = \quad a(x) = |a(x)| \cdot e^{\frac{i}{\varepsilon}S(x)} &= \\ = \text{const} \cdot \exp\left(-\tfrac{1}{2(\varepsilon\sigma)^2}(x-\tilde{x})^2\right) \cdot \exp\left(-\tfrac{i}{\varepsilon}\tilde{y} \cdot x\right). \end{aligned}$$

Das Resultat wird einprägsamer, wenn wir die übliche Bezeichnung für die Dichten der p-dimensionalen Normalverteilungen $\mathcal{N}(\tilde{y}, \frac{1}{\sigma^2} \cdot \mathrm{I})$ und $\mathcal{N}(\tilde{x}, (\varepsilon\sigma)^2 \cdot \mathrm{I})$ benützen.

$$\begin{aligned} q(y) \quad &= \quad n\left(\tilde{y}, \sigma^{-2}\right) \cdot \exp\left(\tfrac{i}{\varepsilon} y \cdot \tilde{x}\right) \cdot \text{const} \\ a(x) \quad &= \quad n\left(\tilde{x}, (\varepsilon\sigma)^2\right) \cdot \exp\left(-\tfrac{i}{\varepsilon}\tilde{y} \cdot x\right) \cdot \text{const} \end{aligned}$$

Wir halten fest, dass ein linearer Phasenfaktor $\frac{i}{\varepsilon} \cdot T(y)$ eine Verschiebung der W-Dichte hervorruft. Die Varianz ändert sich so, dass das Produkt $= \varepsilon^2$ ist. Letzteres ist ein Phänomen, welches (in allgemeineren Situationen) als Heisenbergs Unschärferelation bekannt ist. Eine schärfere Konzentration im Impulsraum ruft eine verringerte Schärfe im Ortsraum hervor; und das Produkt der Unschärfen ist nach unten durch die Planck'sche Konstante bestimmt: ($\Delta p \times \Delta x \geq \hbar$ in sehr vereinfachter Notation).

Differentiation im Schwartz-Raum

Satz 2.5.27 (Gliedweise differenzierbare Fourier-Reihen)**.**
Es sei $f = \sum c_n e^{int}$ *mit* $\sum n|c_n| < \infty$. *Dann ist* f *stetig differenzierbar und es gilt* $\frac{1}{i} \cdot f'(t) = \sum n c_n e^{int}$.

Es sei f *eine stetig differenzierbare* 2π*-periodische Funktion mit* $\int f\ dt = 0$*, deren Ableitung summable Fourier-Koeffizienten besitzt,* $\frac{1}{i}\cdot f'(t) = \sum d_n e^{int}$ *mit* $\sum |d_n| < \infty$*. Dann gilt* $d_0 = 0$ *und* $f(t) = \sum \frac{1}{n} d_n e^{int}$.

Beweis. *Wir beweisen nur die erste Aussage. Die Differenzenquotienten haben die Form*

$$\frac{1}{h}\big[f(t+h) - f(t)\big] = \sum c_n e^{int} n \cdot [\tfrac{e^{inh}-1}{nh}]$$

Da die Funktion $\frac{1}{x}\sin x$ *beschränkt ist, folgt die Behauptung aus dem Satz von der majorisierten Konvergenz.*

Satz 2.5.28 (Differenzierbare Fourier-Integrale).
Wenn für eine integrable Funktion $f(t)$ *auf* $\mathbb{R}$ *die Funktion* $x \cdot \hat{f}(x)$ *integrabel ist. dann ist* f *stetig differenzierbar, und es gilt* $x \cdot \hat{f}(x) = \frac{1}{i}\hat{f}'(x)$.

Es sei $f(t)$ *eine stetig differenzierbare Funktion auf* $\mathbb{R}$ *mit* $f(t) \to 0$ *für* $|t| \to \infty$*. Wenn die Ableitung* $f'(t)$ *direkt Fourier-invertierbar ist, dann ist auch* f *direkt Fourier-invertierbar und es gilt* $x \cdot \hat{f}(x) = \frac{1}{i}\hat{f}'(x)$.

Beweis. *Wenn eine stetig differenzierbare Funktion* $f(t)$ *integrabel ist, dann verschwindet das Integral der Ableitung. Die Ableitung besitzt eine Stammfunktion, die im Unendlichen verschwindet; wir nennen sie die ausgezeichnete Stammfunktion. Die übrigen Stammfunktionen unterscheiden sich um eine Konstante und sind nicht integrabel. Sei nun* $f'(t)$ *direkt Fourier-invertierbar, und sei* f *die ausgezeichnete Stammfunktion. Partielle Integration ergibt dann*

$$\hat{f}'(x) = \int e^{-itx} f'(t)\ dt = \big[e^{-itx} f(t)\big]_{-\infty}^{\infty} + \int ixe^{-itx} f(t)\ dt = ix \cdot \hat{f}(x).$$

Der Beweis der ersten Behauptung funktioniert wie im diskreten Fall; wir überlassen ihn dem Leser.

Satz 2.5.29. *Wenn sowohl* $f(t)$ *als auch* $t \cdot f(t)$ *direkt Fourier-invertierbar sind, dann ist* $\hat{f}$ *stetig differenzierbar und die Ableitung* $\hat{f}'$ *ist bis auf den Faktor* i *die Fourier-Transformierte von* $t \cdot f(t)$.

$$i \cdot \frac{d}{dx}\hat{f}(x) \quad = \quad \widehat{tf}\ (x)$$

Beweis. *Die Differenzenquotienten konvergieren nach dem Satz von der majorisierten Konvergenz*

$$i \cdot \frac{1}{h}\big(\hat{f}(x+h) - f(t)\big) = \frac{1}{2\pi}\int e^{-itx}\frac{1}{-ith}\big(e^{-ith} - 1\big) \cdot t \cdot f(t)\ dt$$

Notation. Ein p-Tupel nichtnegativer ganzer Zahlen $\alpha = (\alpha_1, \ldots, \alpha_p)$ heisst ein Multiindex vom Gewicht k, wenn $k = |\alpha| = \sum \alpha_j$.

Für ein p-Tupel $t = (t_1, \ldots, t_p)$ definiert man $t^\alpha = t_1^{\alpha_1} \cdot \cdots \cdot t_p^{\alpha_p}$. Mit M^α bezeichnen wir den Operator der Multiplikation mit t^α. Für eine komplexwertigen Funktion $f = f(t_1, \ldots, t_p)$ ist $g = M^\alpha f$ die Funktion $g(t_1, \ldots, t_p) = t^\alpha \cdot f(t_1, \ldots, t_p)$. Wir benützen die Notation nicht nur für Funktionen $f(t)$ auf dem Zeilenraum $\mathbb{R}^p_Z$, sondern auch für Funktionen $F(x)$ auf dem Spaltenraum $\mathbb{R}^p_{Sp}$. $\quad G = M^\alpha F$, wenn $G(x) = x^\alpha \cdot F(x)$.

Sprechweise 2.5.2. Eine komplexwertige Funktion f von n reellen Variablen heisst k-mal stetig differenzierbar oder auch k-glatt, wenn die partiellen Ableitungen $\partial^\alpha f$ für alle Multiindizes mit Gewicht $\leq k$ existieren und stetig sind. Wir benützen die Notation ∂^α nicht nur für die für die glatten Funktionen $f(t)$ auf dem Zeilenraum Raum $\mathbb{R}^p_Z$, sondern auch für Funktionen $F(x)$ auf dem Spaltenraum $\mathbb{R}^p_{Sp}$.

Die Differentialoperatoren vom Gewicht 1 nennen wir die elementaren partiellen Ableitungen; wir bezeichnen sie auch mit ∂_j für $j = 1, \ldots, p$.

Man sollte wissen, und wir werden darüber im Kapitel über stetige Differenzierbarkeit ausführlich sprechen: Wenn f zweimal stetig differenzierbar ist, dann gilt $\partial_j(\partial_k f) = \partial_k(\partial_j f)$ für alle j, k. Für eine k-mal stetig differenzierbare Funktion f erhält man (für jeden Multiindex $\alpha = (\alpha_1, \ldots, \alpha_p)$ vom Gewicht $\leq k$) die Funktion $\partial^\alpha f$, indem man f gleichgültig in welcher Reihenfolge, α_1-mal den Operator ∂_1, α_2-mal den Operator ∂_2, usw. anwendet.

Definition. Eine Funktion von n reellen Variablen f heisst eine Schwartz-Funktion, wenn sie beliebig oft stetig differenzierbar ist und für alle Paare von Multiindizes α, β die Funktion $M^\alpha \partial^\beta f$ beschränkt ist. Die Menge der Schwartz-Funktionen auf dem Zeilenraum wird mit $\mathcal{S}(\mathbb{R}^p_Z)$ bezeichnet; $\mathcal{S}(\mathbb{R}^p_{Sp})$ bezeichnet den Schwartz-Raum über dem Spaltenraum.

Die beiden Schwartz-Räume sind offenbar Vektorräume, die auch gegenüber punktweiser Multiplikation abgeschlossen sind. Jeder der Multiplikationsoperatoren M^α und jeder der Differentialoperatoren ∂^α bildet den Schwartz-Raum in sich ab. Diese Operatoren sind lineare Operatoren. Für die elementaren Differentialoperatoren gilt die bekannte Produktregel: $\partial_j(f \cdot g) = (\partial_j f) \cdot g + (\partial_j g) \cdot f$. Eine interessante Konsequenz ist die Kommutatorrelation $\quad \partial_j \circ M_j - M_j \circ \partial_j = \mathrm{id} \quad$ für $j = 1, \ldots, p$.

Satz 2.5.30. *Das Fourier-Integral liefert eine bijektive Abbildung der Schwartz-Räume*

$$\mathcal{F}: \quad S(\mathbb{R}^p_Z) \longrightarrow \mathcal{S}(\mathbb{R}^p_{Sp}); \qquad\qquad f \mapsto \left(\tfrac{1}{2\pi}\right)^p \int e^{-it(\cdot)}\, f(t)\, dt = \hat{f}.$$

Die Umkehrabbildung hat die Form

$$\mathcal{F}^{-1}: \quad S(\mathbb{R}^p_{Sp}) \longrightarrow \mathcal{S}(\mathbb{R}^p_Z); \qquad\qquad F \mapsto \int e^{i(\cdot)t}\, F(x)\, dx = \check{F}.$$

Für jeden Multiindex α gilt

$$\mathcal{F}(M^\alpha f) = i^{|\alpha|} \cdot \partial^\alpha(\mathcal{F} f); \qquad \mathcal{F}(\partial^\alpha f) = i^{|\alpha|} \cdot M^\alpha(\mathcal{F} f)$$
$$\mathcal{F}^{-1}(M^\alpha f) = (-i)^{|\alpha|} \cdot \partial^\alpha(\mathcal{F}^{-1} f); \qquad \mathcal{F}^{-1}(\partial^\alpha f) = (-i)^{|\alpha|} \cdot M^\alpha(\mathcal{F}^{-1} f)$$

Die Beweise haben wir oben für die elementaren Operatoren geführt; und daraus ergibt sich alles.

Index